

3Com Switch 8800 Configuration Guide

www.3com.com

Part No. DUA1750-2BAA01

Published: December 2005

Copyright © 2005, 3Com Corporation. All rights reserved. No part of this documentation may be reproduced in any form or by any means or used to make any derivative work (such as translation, transformation, or adaptation) without written permission from 3Com Corporation.

3Com Corporation reserves the right to revise this documentation and to make changes in content from time to time without obligation on the part of 3Com Corporation to provide notification of such revision or change.

3Com Corporation provides this documentation without warranty, term, or condition of any kind, either implied or expressed, including, but not limited to, the implied warranties, terms or conditions of merchantability, satisfactory quality, and fitness for a particular purpose. 3Com may make improvements or changes in the product(s) and/or the program(s) described in this documentation at any time.

If there is any software on removable media described in this documentation, it is furnished under a license agreement included with the product as a separate document, in the hard copy documentation, or on the removable media in a directory file named LICENSE.TXT or !LICENSE.TXT. If you are unable to locate a copy, please contact 3Com and a copy will be provided to you.

UNITED STATES GOVERNMENT LEGEND

If you are a United States government agency, then this documentation and the software described herein are provided to you subject to the following:

All technical data and computer software are commercial in nature and developed solely at private expense. Software is delivered as "Commercial Computer Software" as defined in DFARS 252.227-7014 (June 1995) or as a "commercial item" as defined in FAR 2.101(a) and as such is provided with only such rights as are provided in 3Com's standard commercial license for the Software. Technical data is provided with limited rights only as provided in DFAR 252.227-7015 (Nov 1995) or FAR 52.227-14 (June 1987), whichever is applicable. You agree not to remove or deface any portion of any legend provided on any licensed program or documentation contained in, or delivered to you in conjunction with, this User Guide.

Unless otherwise indicated, 3Com registered trademarks are registered in the United States and may or may not be registered in other countries.

3Com, the 3Com logo, are registered trademarks of 3Com Corporation.

Intel and Pentium are registered trademarks of Intel Corporation. Microsoft, MS-DOS, Windows, and Windows NT are registered trademarks of Microsoft Corporation. UNIX is a registered trademark in the United States and other countries, licensed exclusively through X/Open Company, Ltd.

All other company and product names may be trademarks of the respective companies with which they are associated.

3Com Corporation
350 Campus Drive
Marlborough, MA
01752-3064

About This Manual

Release Notes

This manual applies to 3Com Switch 8800.

Related Manuals

The related manuals are listed in the following table.

Manual	Content
3Com Switch 8800 Installation Guide	It provides information for the system installation, booting, hardware/software maintenance & monitoring.
3Com Switch 8800 Command Reference Guide	It introduces all commands available in the Switch 8800, as well as a command index.

Organization

3Com Switch 8800 Configuration Guide consists of the following parts:

- **MPLS**
This module introduces the configuration on MPLS and BGP/MPLS VPN features.
- **STP**
This module introduces the configuration on STP feature.
- **Security**
This module presents the configuration on 802.1x, AAA and RADIUS protocols, and TACACS+ protocol.
- **Reliability**
This module focuses on VRRP and HA configurations.
- **System Management**
This module details the configuration involved in system management and maintenance, including file management, system maintenance, SNMP, RMON, NTP protocol, SSH terminal services, and network management.

- **PoE**
This module focuses on power over Ethernet (PoE) configuration.
- **NAT & VPLS**
This module presents the configurations on NAT and VPLS.
- **Appendix**
This appendix offers the acronyms in this manual.

Intended Audience

The manual is intended for the following readers:

- Network engineers
- Network administrators
- Customers who are familiar with network fundamentals

Conventions

The manual uses the following conventions:

I. General conventions

Convention	Description
Arial	Normal paragraphs are in Arial.
Boldface	Headings are in Boldface .
Courier New	Terminal Display is in Courier New.

II. Command conventions

Convention	Description
Boldface	The keywords of a command line are in Boldface .
<i>italic</i>	Command arguments are in <i>italic</i> .
[]	Items (keywords or arguments) in square brackets [] are optional.
{ x y ... }	Alternative items are grouped in braces and separated by vertical bars. One is selected.
[x y ...]	Optional alternative items are grouped in square brackets and separated by vertical bars. One or none is selected.
{ x y ... } *	Alternative items are grouped in braces and separated by vertical bars. A minimum of one or a maximum of all can be

Convention	Description
	selected.
[x y ...]*	Optional alternative items are grouped in square brackets and separated by vertical bars. Many or none can be selected.
#	A line starting with the # sign is comments.

III. GUI conventions

Convention	Description
< >	Button names are inside angle brackets. For example, click the <OK> button.
[]	Window names, menu items, data table and field names are inside square brackets. For example, pop up the [New User] window.
/	Multi-level menus are separated by forward slashes. For example, [File/Create/Folder].

IV. Keyboard operation

Format	Description
<Key>	Press the key with the key name inside angle brackets. For example, <Enter>, <Tab>, <Backspace>, or <A>.
<Key1+Key2>	Press the keys concurrently. For example, <Ctrl+Alt+A> means the three keys should be pressed concurrently.
<Key1, Key2>	Press the keys in turn. For example, <Alt, A> means the two keys should be pressed in turn.

V. Mouse operation

Action	Description
Select	Press and hold the primary mouse button (left mouse button by default).
Click	Select and release the primary mouse button without moving the pointer.
Double-Click	Press the primary mouse button twice continuously and quickly without moving the pointer.
Drag	Press and hold the primary mouse button and move the pointer to a certain position.

VI. Symbols

Eye-catching symbols are also used in the manual to highlight the points worthy of special attention during the operation. They are defined as follows:



Caution: Means reader be extremely careful during the operation.



Note: Means a complementary description.

Table of Contents

Chapter 1 Product Overview	1-1
1.1 Product Overview.....	1-1
1.2 Function Features	1-1
Chapter 2 Logging into Switch	2-1
2.1 Setting Up Configuration Environment Through the Console Port.....	2-1
2.2 Setting Up Configuration Environment Through Telnet.....	2-3
2.2.1 Connecting a PC to the Switch Through Telnet.....	2-3
2.2.2 Telneting a Switch Through Another Switch	2-4
2.3 Setting Up Configuration Environment Through a Dial-up the Modem	2-5
Chapter 3 Command Line Interface.....	3-1
3.1 Command Line Interface	3-1
3.2 Command Line View.....	3-1
3.3 Features and Functions of Command Line	3-7
3.3.1 Online Help of Command Line	3-7
3.3.2 Displaying Characteristics of Command Line	3-9
3.3.3 History Command of Command Line	3-9
3.3.4 Common Command Line Error Messages.....	3-10
3.3.5 Editing Characteristics of Command Line.....	3-10
Chapter 4 User Interface Configuration	4-1
4.1 User Interface Overview	4-1
4.2 User Interface Configuration.....	4-2
4.2.1 Entering User Interface View	4-2
4.2.2 Define the Login Header	4-2
4.2.3 Configuring Asynchronous Port Attributes	4-3
4.2.4 Configuring Terminal Attributes.....	4-4
4.2.5 Managing Users	4-6
4.2.6 Configuring Modem Attributes.....	4-10
4.2.7 Configuring Redirection.....	4-10
4.3 Displaying and Debugging User Interface	4-11
Chapter 5 Management Interface Configuration	5-1
5.1 Management Interface Overview.....	5-1
5.2 Management Interface Configuration	5-1
Chapter 6 Ethernet Port Configuration	6-1
6.1 Ethernet Port Overview.....	6-1
6.2 Ethernet Port Configuration	6-1
6.2.1 Entering Ethernet Port View.....	6-2
6.2.2 Enabling/Disabling an Ethernet Port	6-2
6.2.3 Setting Ethernet Port Description.....	6-2

6.2.4 Setting the Duplex Attribute of the Ethernet Port	6-2
6.2.5 Setting Speed on the Ethernet Port	6-2
6.2.6 Setting the Cable Type for the Ethernet Port	6-3
6.2.7 Enabling/Disabling Flow Control for the Ethernet Port	6-3
6.2.8 Permitting/Forbidding Jumbo Frame to Pass the Ethernet Port	6-3
6.2.9 Setting the Ethernet Port Broadcast Suppression Ratio	6-4
6.2.10 Setting the Ethernet Port Mode	6-4
6.2.11 Setting the Link Type for the Ethernet Port	6-4
6.2.12 Adding the Ethernet Port to Specified VLANs	6-5
6.2.13 Setting the Default VLAN ID for the Ethernet Port	6-5
6.2.14 Setting the VLAN VPN Feature	6-6
6.2.15 Copying Port Configuration to Other Ports	6-6
6.2.16 Setting Port Hold Time	6-7
6.2.17 Setting the Ethernet Port in Loopback Mode	6-7
6.3 Displaying and Debugging Ethernet Port	6-8
6.4 Ethernet Port Configuration Example	6-8
6.5 Ethernet Port Troubleshooting	6-9
Chapter 7 Link Aggregation Configuration	7-1
7.1 Overview	7-1
7.1.1 Introduction to Link Aggregation	7-1
7.1.2 Introduction to LACP	7-1
7.1.3 Aggregation Types	7-1
7.1.4 Load Sharing	7-2
7.2 Link Aggregation Configuration	7-3
7.2.1 Enabling/Disabling LACP at Port	7-4
7.2.2 Creating/Deleting an Aggregation Group	7-4
7.2.3 Adding/Deleting an Ethernet Port into/from an Aggregation Group	7-4
7.2.4 Setting/Deleting Aggregation Group Description	7-5
7.2.5 Configuring System Priority	7-5
7.2.6 Configuring Port Priority	7-6
7.3 Displaying and Debugging Link Aggregation	7-6
7.4 Link Aggregation Configuration Example	7-7
Chapter 8 VLAN Configuration	8-1
8.1 VLAN Overview	8-1
8.2 Configuring VLAN	8-2
8.2.1 Creating/Deleting a VLAN	8-2
8.2.2 Specifying a Description Character String for a VLAN or VLAN interface	8-2
8.2.3 Creating/Removing a VLAN Interface	8-3
8.2.4 Shutting down/Bringing up a VLAN Interface	8-3
8.3 Configuring Port-Based VLAN	8-4
8.3.1 Adding Ethernet Ports to a VLAN	8-4
8.4 Configuring Protocol-Based VLAN	8-4

8.4.1 Creating/Deleting a VLAN Protocol Type.....	8-4
8.4.2 Associating/Dissociating a Port with/from a Protocol-Based VLAN.....	8-5
8.5 Displaying VLAN.....	8-5
8.6 VLAN Configuration Example.....	8-6
Chapter 9 GARP/GVRP Configuration.....	9-1
9.1 Configuring GARP.....	9-1
9.1.1 GARP Overview.....	9-1
9.1.2 Setting the GARP Timer.....	9-2
9.1.3 Displaying and Debugging GARP.....	9-3
9.2 Configuring GVRP.....	9-3
9.2.1 GVRP Overview.....	9-3
9.2.2 Enabling/Disabling Global GVRP.....	9-4
9.2.3 Enabling/Disabling Port GVRP.....	9-5
9.2.4 Setting the GVRP Registration Type.....	9-5
9.2.5 Displaying and Debugging GVRP.....	9-6
9.2.6 GVRP Configuration Example.....	9-6
Chapter 10 Super VLAN Configuration.....	10-1
10.1 Super VLAN Overview.....	10-1
10.2 Configuring a Super VLAN.....	10-1
10.2.1 Configuring a Super VLAN.....	10-1
10.2.2 Super VLAN Configuration Example.....	10-3
Chapter 11 IP Address Configuration.....	11-2
11.1 Introduction to IP Address.....	11-2
11.1.1 IP Address Classification and Representation.....	11-2
11.1.2 Subnet and Mask.....	11-5
11.2 Configuring IP Address.....	11-6
11.2.1 Configuring the Hostname and Host IP Address.....	11-6
11.2.2 Configuring the IP Address of the VLAN Interface.....	11-7
11.3 Displaying and debugging IP Address.....	11-7
11.4 IP Address Configuration Example.....	11-7
11.5 Troubleshooting IP Address Configuration.....	11-8
Chapter 12 ARP Configuration.....	12-1
12.1 Introduction to ARP.....	12-1
12.2 Configuring ARP.....	12-2
12.2.1 Manually Adding/Deleting Static ARP Mapping Entries.....	12-2
12.2.2 Configuring the Dynamic ARP Aging Timer.....	12-3
12.2.3 Enabling/Disabling the Checking Function of ARP Entry.....	12-3
12.3 Displaying and Debugging ARP.....	12-3
12.4 Enabling/Disabling the Scheme of Preventing Attack from Packets.....	12-4
12.4.1 Introduction to the Scheme of Preventing Attack from Packets.....	12-4

Chapter 13 DHCP Configuration	13-1
13.1 Introduction to DHCP.....	13-1
13.1.1 How DHCP Works.....	13-1
13.2 Configuring General DHCP.....	13-3
13.2.1 Enabling/Disabling DHCP Service.....	13-4
13.2.2 Configuring Processing Method of DHCP Packets.....	13-4
13.2.3 Enabling/Disabling Fake DHCP Server Detection.....	13-5
13.3 Configuring DHCP Server.....	13-6
13.3.1 Creating a Global DHCP IP Address Pool.....	13-6
13.3.2 Configuring IP Address Assignment Mode.....	13-7
13.3.3 Forbidding Specified IP Addresses to Be Automatically Assigned.....	13-9
13.3.4 Configuring Lease Time For DHCP Address Pool.....	13-10
13.3.5 Configuring DHCP Client Domain Names.....	13-11
13.3.6 Configuring DNS Server Address for DHCP Clients.....	13-12
13.3.7 Configuring NetBIOS Server Address for DHCP Clients.....	13-13
13.3.8 Configuring NetBIOS Node Type for DHCP Clients.....	13-15
13.3.9 Configuring Custom DHCP Options.....	13-16
13.3.10 Configuring Outbound Gateway Address for DHCP Clients.....	13-17
13.3.11 Configuring Parameters for DHCP Server to Send Ping Packets.....	13-17
13.3.12 Displaying and Debugging the DHCP Server.....	13-18
13.3.13 Clearing the Configuration Information of the DHCP Server.....	13-19
13.3.14 DHCP Server Configuration Example.....	13-19
13.4 Configuring DHCP Relay.....	13-21
13.4.1 Introduction to DHCP Relay.....	13-21
13.4.2 Configuring DHCP Relay.....	13-22
13.4.3 Displaying and Debugging DHCP Relay.....	13-23
13.4.4 DHCP Relay Configuration Example.....	13-24
Chapter 14 DNS Configuration	14-1
14.1 Introduction to DNS.....	14-1
14.1.1 Static Domain Name Resolution.....	14-1
14.1.2 Dynamic Domain Name Resolution.....	14-1
14.2 Configuring Static Domain Name Resolution.....	14-2
14.3 Configuring Dynamic Domain Name Resolution.....	14-2
14.3.1 Enable/Disable Static Domain Name Resolution.....	14-2
14.3.2 Configure the IP Address of Domain Name Server.....	14-3
14.3.3 Configure Domain Name Suffix.....	14-3
14.4 Displaying and Debugging Domain Name Resolution.....	14-4
14.5 DNS Configuration Example.....	14-4
14.6 Troubleshooting Domain Name Resolution Configuration.....	14-5
Chapter 15 IP Performance Configuration	15-1
15.1 Configuring IP Performance.....	15-1
15.1.1 Configuring TCP Attributes.....	15-1

15.2 Displaying and Debugging IP Performance.....	15-2
15.3 Troubleshooting IP Performance.....	15-3
Chapter 16 IP Routing Protocol Overview	16-5
16.1 Introduction to IP Route and Routing Table	16-5
16.1.1 IP Route and Route Segment.....	16-5
16.1.2 Route Selection through the Routing Table	16-6
16.2 Routing Management Policy.....	16-8
16.2.1 Routing Protocols and the Preferences of the Corresponding Routes	16-8
16.2.2 Supporting Load Sharing and Route Backup.....	16-9
16.2.3 Routes Shared Between Routing Protocols.....	16-10
Chapter 17 Static Route Configuration	17-1
17.1 Introduction to Static Route	17-1
17.1.1 Static Route.....	17-1
17.1.2 Default Route	17-1
17.2 Configuring Static Route.....	17-2
17.2.1 Configuring a Static Route	17-2
17.2.2 Configuring a Default Route.....	17-3
17.2.3 Deleting All the Static Routes.....	17-3
17.3 Displaying and Debugging Static Route	17-4
17.4 Typical Static Route Configuration Example	17-4
17.5 Troubleshooting Static Route Faults	17-5
Chapter 18 RIP Configuration	18-1
18.1 Introduction to RIP	18-1
18.1.1 RIP Operation Mechanism.....	18-1
18.1.2 RIP Enabling and Running.....	18-2
18.2 Configuring RIP.....	18-2
18.2.1 Enabling RIP and Entering RIP View.....	18-3
18.2.2 Enabling RIP on the Specified Network Segment.....	18-3
18.2.3 Configuring Unicast of the Packets.....	18-4
18.2.4 Configuring Split Horizon	18-4
18.2.5 Setting Additional Routing Metric.....	18-5
18.2.6 Configuring RIP to Import Routes of Other Protocols.....	18-5
18.2.7 Configuring Route Filtering	18-6
18.2.8 Disabling RIP to Receive Host Route	18-7
18.2.9 Enabling RIP-2 Route Aggregation Function.....	18-8
18.2.10 Setting the RIP Preference	18-8
18.2.11 Specifying RIP Version of the Interface	18-8
18.2.12 Configuring RIP Timers.....	18-9
18.2.13 Configuring RIP-1 Zero Field Check of the Interface Packet.....	18-10
18.2.14 Specifying the Operating State of the Interface	18-10
18.2.15 Setting RIP-2 Packet Authentication.....	18-11
18.3 Displaying and Debugging RIP.....	18-12

18.4 Typical RIP Configuration Example	18-12
18.5 Troubleshooting RIP Faults	18-14
Chapter 19 OSPF Configuration	19-1
19.1 OSPF Overview	19-1
19.1.1 Introduction to OSPF	19-1
19.1.2 Process of OSPF Route Calculation	19-1
19.1.3 OSPF Packets	19-2
19.1.4 LSA Type	19-3
19.1.5 Basic Concepts Related to OSPF	19-3
19.1.6 OSPF Features Supported by the Switch 8800	19-5
19.2 Configuring OSPF	19-6
19.2.1 Configuring Router ID	19-7
19.2.2 Enabling OSPF	19-7
19.2.3 Entering OSPF Area View	19-8
19.2.4 Specifying an Interface to Run OSPF	19-8
19.2.5 Configuring OSPF to Import Routes of Other Protocols	19-9
19.2.6 Configuring OSPF to Import Default Routes	19-11
19.2.7 Configuring OSPF Route Filtering	19-12
19.2.8 Configuring the Route Summary of OSPF	19-13
19.2.9 Setting OSPF Route Preference	19-15
19.2.10 Configuring OSPF Timers	19-15
19.2.11 Configuring the Network Type on the OSPF Interface	19-17
19.2.12 Configuring NBMA Neighbors for OSPF	19-18
19.2.13 Setting the Interface Priority for DR Election	19-18
19.2.14 Configuring an Interval Required for Sending LSU Packets	19-20
19.2.15 Configuring the Cost for Sending Packets on an Interface	19-20
19.2.16 Configuring to Fill the MTU Field When an Interface Transmits DD Packets	19-20
19.2.17 Setting a Shortest Path First (SPF) Calculation Interval for OSPF	19-21
19.2.18 Disabling the Interface to Send OSPF Packets	19-21
19.2.19 Configuring OSPF Authentication	19-22
19.2.20 Configuring OSPF Virtual Link	19-23
19.2.21 Configuring Stub Area of OSPF	19-24
19.2.22 Configuring NSSA Area of OSPF	19-24
19.2.23 Configuring OSPF and Network Management System (NMS)	19-26
19.2.24 Resetting the OSPF Process	19-27
19.3 Displaying and Debugging OSPF	19-27
19.4 Typical OSPF Configuration Example	19-29
19.4.1 Configuring DR Election Based on OSPF Priority	19-29
19.4.2 Configuring OSPF Virtual Link	19-31
19.5 Troubleshooting OSPF Faults	19-32
Chapter 20 Integrated IS-IS Configuration	20-1
20.1 Introduction to Integrated IS-IS	20-1

20.1.1 Terms of IS-IS Routing Protocol	20-1
20.1.2 Two-level Structure of IS-IS Routing Protocol	20-2
20.1.3 NSAP Structure of IS-IS Routing Protocol	20-4
20.1.4 IS-IS Routing Protocol Packets.....	20-5
20.2 Configuring Integrated IS-IS	20-6
20.2.1 Enabling IS-IS and Entering the IS-IS View.....	20-7
20.2.2 Setting Network Entity Title	20-7
20.2.3 Enabling IS-IS on the Specified Interface	20-7
20.2.4 Setting Priority for DIS Election.....	20-8
20.2.5 Setting Router Type	20-8
20.2.6 Setting Interface Circuit Level	20-9
20.2.7 Configuring IS-IS to Import Routes of Other Protocols.....	20-9
20.2.8 Configuring IS-IS Route Filtering	20-10
20.2.9 Configuring IS-IS Routing Leak.....	20-11
20.2.10 Setting IS-IS Route Summary.....	20-11
20.2.11 Setting to Generate Default Route.....	20-12
20.2.12 Setting the Preference of IS-IS Protocol.....	20-12
20.2.13 Configuring IS-IS Route Metric Type	20-13
20.2.14 Setting IS-IS Link State Routing Cost.....	20-13
20.2.15 Configuring IS-IS Timers.....	20-13
20.2.16 Setting IS-IS Authentication	20-16
20.2.17 Setting the Mesh Group of the Interface.....	20-17
20.2.18 Setting Overload Flag Bit	20-18
20.2.19 Setting to Discard the LSPs with Checksum Errors.....	20-18
20.2.20 Setting to Log the Peer Changes.....	20-19
20.2.21 Setting LSP Refreshment Interval.....	20-19
20.2.22 Setting Lifetime of LSP.....	20-19
20.2.23 Setting Parameters Related to SPF.....	20-20
20.2.24 Enabling/Disabling the Interface to Send Packets.....	20-21
20.2.25 Resetting All the IS-IS Data Structure.....	20-22
20.2.26 Resetting the Specified IS-IS Peer	20-22
20.3 Displaying and Debugging Integrated IS-IS	20-22
20.4 Typical Integrated IS-IS Configuration Example.....	20-23
Chapter 21 BGP Configuration	21-1
21.1 BGP/MBGP Overview.....	21-1
21.1.1 Introduction to BGP	21-1
21.1.2 BGP Message Types	21-2
21.1.3 BGP Routing Mechanism.....	21-2
21.1.4 MBGP.....	21-3
21.1.5 BGP Peer and Peer Group	21-4
21.2 Configuring BGP	21-4
21.2.1 Enabling BGP.....	21-5

21.2.2	Configuring Basic Features for BGP Peer	21-5
21.2.3	Configuring application features of a BGP peer (group)	21-8
21.2.4	Configuring Route Filtering of a Peer (group)	21-12
21.2.5	Configuring Network Routes for BGP Distribution	21-14
21.2.6	Configuring the Interaction Between BGP and IGP	21-14
21.2.7	Configuring BGP Route Summarization	21-15
21.2.8	Configuring BGP Route Filtering.....	21-16
21.2.9	Configuring BGP Route Dampening	21-17
21.2.10	Configuring BGP Preference.....	21-18
21.2.11	Configuring BGP Timer	21-19
21.2.12	Configuring the Local Preference	21-19
21.2.13	Configuring MED for AS.....	21-20
21.2.14	Comparing the MED Routing Metrics from the Peers in Different ASs.....	21-20
21.2.15	Configuring BGP Route Reflector	21-21
21.2.16	Configuring BGP AS Confederation Attribute	21-22
21.2.17	Clearing BGP Connection	21-24
21.2.18	Refreshing BGP Routes.....	21-24
21.3	Displaying and Debugging BGP	21-24
21.4	Typical BGP Configuration Example	21-26
21.4.1	Configuring BGP AS Confederation Attribute	21-26
21.4.2	Configuring BGP Route Reflector	21-28
21.4.3	Configuring BGP Routing.....	21-30
21.5	Troubleshooting BGP	21-33
Chapter 22	IP Routing Policy Configuration	22-1
22.1	Introduction to IP Routing Policy.....	22-1
22.1.1	Filter	22-1
22.1.2	Routing Policy Application.....	22-2
22.2	Configuring IP Routing Policy	22-2
22.2.1	Configuring a Route-policy	22-3
22.2.2	Configuring ip-prefix	22-6
22.2.3	Configuring the AS Path List.....	22-7
22.2.4	Configuring a Community Attribute List	22-8
22.2.5	Importing Routing Information Discovered by Other Routing Protocols	22-8
22.2.6	Configuring Route Filtering	22-9
22.3	Displaying and Debugging the Routing Policy.....	22-10
22.4	Typical IP Routing Policy Configuration Example	22-11
22.4.1	Configuring to Filter the Received Routing Information	22-11
22.5	Troubleshooting Routing Policy	22-12
Chapter 23	IP Multicast Overview	23-4
23.1	IP Multicast Overview	23-4
23.1.1	Problems with Unicast/Broadcast	23-4
23.1.2	Advantages of Multicast	23-6

23.1.3 Application of Multicast	23-7
23.2 Implementation of IP Multicast.....	23-7
23.2.1 IP Multicast Addresses.....	23-7
23.2.2 IP Multicast Protocols.....	23-9
23.3 RPF Mechanism for IP Multicast Packets	23-10
Chapter 24 IGMP Snooping Configuration	24-1
24.1 IGMP Snooping Overview	24-1
24.1.1 IGMP Snooping Principle	24-1
24.1.2 Implement IGMP Snooping	24-2
24.2 IGMP Snooping Configuration	24-4
24.2.1 Enabling/Disabling IGMP Snooping	24-4
24.2.2 Configuring Router Port Aging Time	24-5
24.2.3 Configuring Maximum Response Time.....	24-5
24.2.4 Configuring Aging Time of Multicast Group Member Ports	24-6
24.2.5 Configuring Unknown Multicast Packets not Broadcasted within a VLAN	24-6
24.3 Displaying and debugging IGMP Snooping.....	24-7
24.4 IGMP Snooping Configuration Example.....	24-7
24.4.1 Enable IGMP Snooping.....	24-7
24.5 Troubleshoot IGMP Snooping	24-8
Chapter 25 Multicast VLAN Configuration.....	25-1
25.1 Multicast VLAN Overview	25-1
25.2 Multicast VLAN Configuration.....	25-1
25.3 Multicast VLAN Configuration Example.....	25-2
Chapter 26 Common Multicast Configuration.....	26-1
26.1 Introduction to Common Multicast Configuration.....	26-1
26.2 Common Multicast Configuration.....	26-1
26.2.1 Enabling Multicast	26-1
26.2.2 Configuring multicast route number limit.....	26-2
26.2.3 Clearing MFC Forwarding Entries or Its Statistic Information.....	26-2
26.2.4 Clearing Route Entries from the Kernel Multicast Routing Table	26-2
26.3 Controlled Multicast Configuration.....	26-3
26.3.1 Controlled Multicast Overview.....	26-3
26.3.2 Configuring Controlled Multicast	26-3
26.3.3 Controlled Multicast Configuration Example	26-4
26.4 Displaying and Debugging Common Multicast Configuration	26-5
Chapter 27 IGMP Configuration	27-1
27.1 IGMP Overview.....	27-1
27.1.1 Introduction to IGMP	27-1
27.2 IGMP Configuration	27-2
27.2.1 Enabling Multicast	27-3
27.2.2 Enabling IGMP on an Interface.....	27-3

27.2.3 Configuring the IGMP Version	27-3
27.2.4 Configuring the Interval to Send IGMP Query Message.....	27-4
27.2.5 Configuring the Interval and the Number of Querying IGMP Packets	27-4
27.2.6 Configuring the Present Time of IGMP Querier	27-5
27.2.7 Configuring Maximum Response Time for IGMP Query Message.....	27-5
27.2.8 Configuring the limit of IGMP groups on an interface	27-6
27.2.9 Configuring a Router to Join Specified Multicast Group	27-6
27.2.10 Limiting Multicast Groups that an Interface Can Access	27-7
27.2.11 Deleting IGMP Groups Joined on an Interface	27-8
27.3 Displaying and Debugging IGMP	27-9
Chapter 28 PIM-DM Configuration	28-1
28.1 PIM-DM Overview	28-1
28.1.1 Introduction to PIM-DM	28-1
28.1.2 PIM-DM Working Principle	28-1
28.2 PIM-DM Configuration	28-3
28.2.1 Enabling Multicast	28-3
28.2.2 Enabling PIM-DM	28-3
28.2.3 Configuring the Time Intervals for Ports to Send Hello Packets.....	28-4
28.2.4 Entering the PIM View.....	28-4
28.2.5 Configuring the Filtering of Multicast Source/Group	28-5
28.2.6 Configuring the Filtering of PIM Neighbor.....	28-5
28.2.7 Configuring the Maximum Number of PIM Neighbor on an Interface.....	28-5
28.2.8 Clearing multicast route entries from PIM routing table.....	28-6
28.2.9 Clearing PIM Neighbors	28-6
28.3 Displaying and Debugging PIM-DM.....	28-6
28.4 PIM-DM Configuration Example	28-7
Chapter 29 PIM-SM Configuration	29-1
29.1 PIM-SM Overview	29-1
29.1.1 Introduction to PIM-SM	29-1
29.1.2 PIM-SM Working Principle	29-1
29.1.3 Preparations before Configuring PIM-SM	29-2
29.2 PIM-SM Configuration	29-3
29.2.1 Enabling Multicast	29-4
29.2.2 Enabling PIM-SM	29-4
29.2.3 Entering the PIM View.....	29-4
29.2.4 Configuring the Time Intervals for Ports to Send Hello Packets.....	29-4
29.2.5 Configuring Candidate-BSRs	29-4
29.2.6 Configuring Candidate-RPs	29-5
29.2.7 Configuring Static RP.....	29-6
29.2.8 Configuring the PIM-SM Domain Border	29-6
29.2.9 Configuring the filtering of multicast source/group.....	29-7
29.2.10 Configuring the filtering of PIM neighbor.....	29-7

29.2.11 Configuring RP to Filter the Register Messages Sent by DR	29-7
29.2.12 Limiting the range of legal BSR.....	29-7
29.2.13 Limiting the range of legal C-RP	29-8
29.2.14 Clearing multicast route entries from PIM routing table	29-8
29.2.15 Clearing PIM Neighbors	29-8
29.3 Displaying and Debugging PIM-SM.....	29-8
29.4 PIM-SM Configuration Example	29-9
Chapter 30 MSDP Configuration.....	30-1
30.1 MSDP Overview.....	30-1
30.1.1 Introduction.....	30-1
30.1.2 Working Principle	30-2
30.2 MSDP Configuration	30-4
30.2.1 Enabling MSDP	30-4
30.2.2 Configuring MSDP Peers	30-5
30.2.3 Configuring Static RPF Peers	30-5
30.2.4 Configuring Originating RP	30-6
30.2.5 Configuring SA Caching State	30-6
30.2.6 Configuring the Maximum Number of SA caching.....	30-7
30.2.7 Requesting Source Information of MSDP Peers.....	30-7
30.2.8 Controlling the Source Information Created.....	30-7
30.2.9 Controlling the Source Information Forwarded	30-8
30.2.10 Controlling the Received Source Information	30-9
30.2.11 Configuring MSDP Mesh Group.....	30-10
30.2.12 Configuring the MSDP Connection Retry Period.....	30-10
30.2.13 Shutting MSDP Peers Down	30-11
30.2.14 Clearing MSDP Connections, Statistics and SA Caching Configuration	30-11
30.3 Displaying and Debugging MSDP	30-12
30.4 MSDP Configuration Examples	30-13
30.4.1 Configuring Static RPF Peers	30-13
30.4.2 Configuring Anycast RP	30-14
30.4.3 MSDP Integrated Networking.....	30-18
Chapter 31 MBGP Multicast Extension Configuration	31-1
31.1 MBGP Multicast Extension Overview	31-1
31.1.1 Introduction.....	31-1
31.1.2 MBGP Extension Attributes for Multicast.....	31-1
31.1.3 MBGP Operating Mode and Message Type	31-2
31.2 MBGP Multicast Extension Configuration.....	31-3
31.2.1 Enabling MBGP Multicast Extension Protocol	31-3
31.2.2 Specifying Network Routes Notified by MBGP Multicast Extension.....	31-4
31.2.3 Configuring the MED Value for an AS	31-4
31.2.4 Comparing MED Values from Different AS Neighbor Paths.....	31-4
31.2.5 Configuring Local Preference.....	31-5

31.2.6 Configuring MBGP Timer	31-5
31.2.7 Configuring MBGP Peer (Group)	31-5
31.2.8 Configuring MBGP Route Aggregation	31-9
31.2.9 Configuring an MBGP Route Reflector	31-9
31.2.10 Configure MBGP Community Attributes	31-10
31.2.11 Importing IGP Routing Information into MBGP	31-10
31.2.12 Defining AS Path List and Routing Policy	31-10
31.2.13 Configuring MBGP Route Filtering	31-11
31.2.14 Resetting BGP Connections	31-11
31.3 Displaying and Debugging MBGP Configuration.....	31-11
31.4 MBGP Multicast Extension Configuration Example.....	31-12
Chapter 35 MPLS Architecture.....	35-2
35.1 MPLS Overview	35-2
35.2 MPLS Basic Concepts	35-3
35.2.1 FEC	35-3
35.2.2 Label.....	35-3
35.2.3 LDP	35-6
35.3 MPLS Architecture.....	35-6
35.3.1 MPLS Network Structure.....	35-6
35.3.2 Forwarding Labeled Packets.....	35-7
35.3.3 Establishing LSP	35-7
35.3.4 LSP Tunnel and Hierarchy	35-9
35.4 MPLS and other Protocols.....	35-10
35.4.1 MPLS and Routing Protocols	35-10
35.5 MPLS Application	35-10
35.5.1 MPLS VPN	35-10
Chapter 36 MPLS Basic Capability Configuration	36-1
36.1 MPLS Basic Capability Overview	36-1
36.2 MPLS Configuration.....	36-1
36.2.1 Defining MPLS LSR ID.....	36-1
36.2.2 Enabling MPLS and Entering MPLS View	36-2
36.2.3 Configuring the Topology-Driven LSP Setup Policy	36-2
36.2.4 Configuring Static LSP	36-2
36.3 LDP Configuration	36-3
36.3.1 Enabling LDP protocol	36-3
36.3.2 Enabling LDP on VLAN interface.....	36-4
36.3.3 Configuring Remote-Peer for Extended Discovery Mode.....	36-4
36.3.4 Configuring session parameters	36-5
36.3.5 Configuring LDP Loop Detection Control.....	36-7
36.3.6 Configuring LDP Authentication Mode Between Every Two Routers	36-8
36.4 Displaying and Debugging MPLS.....	36-8
36.4.1 Displaying and Debugging MPLS	36-8

36.4.2 Displaying and Debugging LDP	36-10
36.5 Typical MPLS Configuration Example	36-11
36.6 Troubleshooting MPLS Configuration.....	36-15
Chapter 37 BGP/MPLS VPN Configuration.....	37-1
37.1 BGP/MPLS VPN Overview	37-1
37.1.1 BGP/MPLS VPN Model.....	37-2
37.1.2 BGP/MPLS VPN Implementation.....	37-5
37.1.3 Nested BGP/MPLS VPN Implementation	37-7
37.1.4 Hierarchical BGP/MPLS VPN Implementation.....	37-7
37.1.5 Introduction to OSPF Multi-instance	37-8
37.1.6 Introduction to Multi-Role Host.....	37-9
37.2 BGP/MPLS VPN Configuration.....	37-10
37.2.1 Configuring CE Router	37-10
37.2.2 Configuring PE Router	37-12
37.2.3 Configuring P Router.....	37-25
37.3 Displaying and Debugging BGP/MPLS VPN.....	37-25
37.4 Typical BGP/MPLS VPN Configuration Example	37-27
37.4.1 Integrated BGP/MPLS VPN Configuration Example.....	37-27
37.4.2 Hybrid BGP/MPLS VPN Configuration Example	37-33
37.4.3 Extranet Configuration Example	37-40
37.4.4 Hub&Spoke Configuration Example.....	37-44
37.4.5 CE Dual-home Configuration Example	37-50
37.4.6 Cross-domain BGP/MPLS VPN Configuration Example	37-56
37.4.7 Cross-Domain BGP/MPLS VPN Configuration Example — Option C.....	37-61
37.4.8 Hierarchical BGP/MPLS VPN Configuration Example.....	37-68
37.4.9 OSPF Multi-instance sham link Configuration Example	37-72
37.4.10 Nested BGP/MPLS VPN Configuration Example	37-77
37.4.11 OSPF Multi-instance CE Configuration Example.....	37-83
37.4.12 Multi-Role Host Configuration Example	37-85
37.5 Troubleshooting	37-90
Chapter 38 MSTP Region-configuration	38-1
38.1 Introduction to MSTP	38-1
38.1.1 MSTP Concepts	38-2
38.1.2 MSTP Principles.....	38-7
38.1.3 MSTP Implementation on the Switch.....	38-12
38.2 Configuring MSTP	38-12
38.2.1 Configuring the MST Region for a Switch.....	38-13
38.2.2 Specifying the Switch as a Primary or a Secondary Root bridge	38-15
38.2.3 Configuring the MSTP Running Mode	38-16
38.2.4 Configuring the Bridge Priority for a Switch	38-17
38.2.5 Configuring the Max Hops in an MST Region.....	38-18
38.2.6 Configuring the Switching Network Diameter	38-18

38.2.7 Configuring the Time Parameters of a Switch	38-19
38.2.8 Setting the Timeout Factor of a Specific Bridge.....	38-21
38.2.9 Configuring the Max Transmission Speed on a Port	38-21
38.2.10 Configuring a Port as an Edge Port or Non-edge Port	38-22
38.2.11 Configuring the Path Cost of a Port	38-23
38.2.12 STP Path Cost Calculation Standards on STP port.....	38-24
38.2.13 Configuring the Priority of a Port.....	38-26
38.2.14 Configuring the Port (Not) to Connect with the Point-to-Point Link.....	38-27
38.2.15 Configuring the mCheck Variable of a Port.....	38-28
38.2.16 Configuring the Switch Protection Function	38-30
38.2.17 Enabling/Disabling MSTP on the Device	38-32
38.2.18 Enable/Disable Address Table Reset on Specified Port.....	38-33
38.2.19 Enabling/Disabling ARP Address Update	38-33
38.2.20 Enabling/Disabling MSTP on a Port.....	38-34
38.3 Displaying and Debugging MSTP	38-34
38.4 Typical MSTP Configuration Example.....	38-35
Chapter 39 802.1x Configuration	39-2
39.1 802.1x Overview	39-2
39.1.1 802.1x Standard Overview.....	39-2
39.1.2 802.1x System Architecture	39-3
39.1.3 802.1x Authentication Process.....	39-4
39.1.4 Implementing 802.1x on Ethernet Switches.....	39-4
39.2 802.1x Configuration.....	39-5
39.2.1 Enabling/Disabling 802.1x.....	39-5
39.2.2 Setting the Port Access Control Mode	39-6
39.2.3 Setting Port Access Control Method	39-7
39.2.4 Checking the Users that Log on the Switch via Proxy	39-7
39.2.5 Setting Supplicant Number on a Port.....	39-8
39.2.6 Setting the Authentication in DHCP Environment.....	39-8
39.2.7 Configuring Authentication Method for 802.1x User	39-8
39.2.8 Enabling/Disabling Guest VLAN	39-9
39.2.9 Setting the Maximum times of authentication request message retransmission.....	39-10
39.2.10 Configuring 802.1x Timers	39-10
39.2.11 Enabling/Disabling quiet-period Timer	39-11
39.3 Displaying and Debugging 802.1x.....	39-12
39.4 802.1x Configuration Example.....	39-12
Chapter 40 AAA and RADIUS/TACACS+ Protocol Configuration	40-1
40.1 AAA and RADIUS/TACACS+ Protocol Overview	40-1
40.1.1 AAA Overview	40-1
40.1.2 RADIUS Protocol Overview	40-1
40.1.3 TACACS+ Protocol Overview	40-2
40.1.4 Implementing AAA/RADIUS on a Switch	40-5

40.2 AAA Configuration	40-6
40.2.1 Creating/Deleting an ISP Domain	40-6
40.2.2 Configuring Relevant Attributes of an ISP Domain	40-7
40.2.3 Configuring Self-Service Server URL	40-8
40.2.4 Creating/Deleting a Local User	40-9
40.2.5 Setting the Attributes of a Local User	40-10
40.2.6 Disconnecting a User by Force	40-11
40.2.7 Configuring Dynamic VLAN Delivering	40-11
40.3 Configuring RADIUS Protocol.....	40-12
40.3.1 Creating/Deleting a RADIUS scheme	40-13
40.3.2 Setting IP Address and Port Number of a RADIUS Server	40-13
40.3.3 Setting the RADIUS Packet Encryption Key	40-15
40.3.4 Setting the Response Timeout Timer of a RADIUS Server	40-16
40.3.5 Setting the Retransmission Times of RADIUS Request Packets	40-16
40.3.6 Enabling the Selection Of Radius Accounting Option.....	40-17
40.3.7 Setting a Real-time Accounting Interval.....	40-17
40.3.8 Setting the Maximum Times of Real-time Accounting Request Failing to be Responded	40-18
40.3.9 Enabling/Disabling Stopping Accounting Request Buffer	40-19
40.3.10 Setting the Maximum Retransmitting Times of Stopping Accounting Request.....	40-19
40.3.11 Setting the Supported Type of RADIUS Server	40-20
40.3.12 Setting RADIUS Server State	40-20
40.3.13 Setting the Username Format Transmitted to RADIUS Server	40-21
40.3.14 Setting the Unit of Data Flow that Transmitted to RADIUS Server.....	40-21
40.3.15 Creating/Deleting a Local RADIUS authentication Server.....	40-22
40.4 Configuring TACACS+ Protocol	40-22
40.4.1 Creating a HWTACAS Scheme	40-23
40.4.2 Configuring TACACS+ Authentication Servers.....	40-23
40.4.3 Configuring TACACS+ Authorization Servers	40-24
40.4.4 Configuring TACACS+ Accounting Servers and the Related Attributes.....	40-25
40.4.5 Configuring the Source Address for TACACS+ Packets Sent by NAS	40-26
40.4.6 Setting a Key for Securing the Communication with TACACS Server	40-26
40.4.7 Setting the Username Format Acceptable to the TACACS Server.....	40-26
40.4.8 Setting the Unit of Data Flows Destined for the TACACS Server.....	40-27
40.4.9 Setting Timers Regarding TACACS Server	40-27
40.5 Displaying and Debugging AAA and RADIUS Protocol	40-29
40.6 AAA and RADIUS/TACACS+ Protocol Configuration Examples.....	40-30
40.6.1 Configuring Authentication at Remote RADIUS Server	40-30
40.6.2 Configuring Authentication at Local RADIUS Authentication Server.....	40-32
40.6.3 Configuring Authentication at Remote TACACS Server	40-32
40.7 Troubleshooting AAA and RADIUS/TACACS+	40-34

Chapter 41 VRRP Configuration	41-1
41.1 Introduction to VRRP	41-1
41.2 Configuring VRRP	41-3
41.2.1 Enabling/Disabling the Function to Ping the Virtual IP Address	41-3
41.2.2 Enabling/Disabling the Check of TTL Value of VRRP Packet	41-4
41.2.3 Setting Correspondence Between Virtual IP Address and MAC Address	41-4
41.2.4 Adding/Deleting a Virtual IP Address	41-5
41.2.5 Configuring the Priority of Switches in the Virtual Router	41-5
41.2.6 Configuring Preemption and Delay for a Switch Within a Virtual Router	41-6
41.2.7 Configuring Authentication Type and Authentication Key	41-7
41.2.8 Configuring Virtual Router Timer	41-8
41.2.9 Configuring Switch to Track a Specified Interface	41-8
41.3 Displaying and debugging VRRP	41-9
41.4 VRRP Configuration Example	41-10
41.4.1 VRRP Single Virtual Router Example	41-10
41.4.2 VRRP Tracking Interface Example	41-11
41.4.3 Multiple Virtual Routers Example	41-13
41.5 Troubleshooting VRRP	41-14
Chapter 42 HA Configuration	42-1
42.1 Introduction to HA	42-1
42.2 Configuring HA	42-1
42.2.1 Restarting the Slave System Manually	42-2
42.2.2 Starting the Master-Slave Switchover Manually	42-2
42.2.3 Enabling/Disabling Automatic Synchronization	42-2
42.2.4 Synchronizing the Configuration File Manually	42-3
42.2.5 Configuring the Load Mode of the Fabric and Slave Board	42-3
42.3 Displaying and Debugging HA Configuration	42-4
Chapter 43 File System Management	43-4
43.1 File System Configuration	43-4
43.1.1 File System Overview	43-4
43.1.2 Directory Operation	43-5
43.1.3 File Operation	43-5
43.1.4 Storage Device Operation	43-6
43.1.5 Setting the Prompt Mode of the File System	43-7
43.2 Configuration File Management	43-7
43.2.1 Configuration File Management Overview	43-7
43.2.2 Displaying the Current-Configuration and Saved-Configuration of Ethernet Switch	43-8
43.2.3 Modifying and Saving the Current-Configuration	43-9
43.2.4 Erasing Configuration Files from Flash Memory	43-9
43.2.5 Configuring the Name of the Configuration File Used for the Next Startup	43-9
43.3 FTP Configuration	43-10
43.3.1 FTP Overview	43-10

43.3.2 Enabling/Disabling FTP Server	43-12
43.3.3 Configuring the FTP Server Authentication and Authorization	43-12
43.3.4 Configuring the Running Parameters of FTP Server	43-13
43.3.5 Displaying and Debugging FTP Server	43-13
43.3.6 Disconnecting an FTP User	43-13
43.3.7 Introduction to FTP Client	43-14
43.3.8 FTP Client Configuration Example	43-14
43.3.9 FTP Server Configuration Example	43-15
43.4 TFTP Configuration	43-17
43.4.1 TFTP Overview	43-17
43.4.2 Downloading Files by Means of TFTP	43-18
43.4.3 Uploading Files by Means of TFTP	43-18
43.4.4 TFTP Client Configuration Example	43-19
Chapter 44 MAC Address Table Management	44-1
44.1 MAC Address Table Management Overview	44-1
44.2 MAC Address Table Management Configuration	44-2
44.2.1 Setting MAC Address Table Entries	44-2
44.2.2 Setting MAC Address Aging Time	44-2
44.3 Maximum MAC Address Number Learned by Ethernet Port and Forwarding Option Configuration	44-3
44.3.1 Maximum MAC Address Number Learned by a Port and Forwarding Option Configuration Tasks	44-4
44.3.2 Configuring Maximum MAC Address Number Learned by Ethernet Port and Forwarding Option Example	44-5
44.4 Displaying and Debugging MAC Address Tables	44-5
44.5 Resetting MAC Addresses	44-6
44.6 MAC Address Table Management Configuration Example	44-6
Chapter 45 Device management	45-1
45.1 Device Management Overview	45-1
45.2 Device Management Configuration	45-1
45.2.1 Rebooting the Ethernet Switch	45-1
45.2.2 Enabling the Timing Reboot Function	45-1
45.2.3 Designating the APP Adopted on Next Booting	45-2
45.2.4 Upgrading BootROM	45-3
45.2.5 Setting Slot Temperature Limit	45-3
45.2.6 Updating Service Processing Boards	45-3
45.3 Displaying and Debugging Device Management	45-4
45.4 Device Management Configuration Example	45-5
45.4.1 Using the Switch as an FTP Client to Implement the Remote Upgrade	45-5
45.4.2 Use the Switch as an FTP Server to Implement the Remote Upgrade	45-7
Chapter 46 System Maintenance and Debugging	46-1
46.1 Basic System Configuration	46-1

46.1.1	Setting a Name for a Switch.....	46-1
46.1.2	Setting the System Clock.....	46-1
46.1.3	Setting the Time Zone.....	46-1
46.1.4	Setting the Summer Time.....	46-2
46.2	Displaying the State and Information of the System.....	46-2
46.3	System Debugging.....	46-3
46.3.1	Enabling/Disabling the Terminal Debugging.....	46-3
46.3.2	Displaying Diagnostic Information.....	46-4
46.4	Testing Tools for Network Connection.....	46-5
46.4.1	ping.....	46-5
46.4.2	ping-distribute enable.....	46-5
46.4.3	tracert.....	46-6
46.5	Logging Function.....	46-6
46.5.1	Introduction to Info-center.....	46-6
46.5.2	Info-center Configuration.....	46-10
46.5.3	Sending the Configuration Information to the Loghost.....	46-14
46.5.4	Sending the Configuration Information to Console terminal.....	46-16
46.5.5	Sending the Configuration Information to Telnet Terminal or Dumb Terminal.....	46-19
46.5.6	Sending the Configuration Information to the Log Buffer.....	46-21
46.5.7	Sending the Configuration Information to the Trap Buffer.....	46-23
46.5.8	Sending the Configuration Information to SNMP Network Management.....	46-25
46.5.9	Displaying and Debugging Info-center.....	46-27
46.5.10	Configuration Examples of Sending Log to the Unix Loghost.....	46-28
46.5.11	Configuration examples of sending log to Linux loghost.....	46-30
46.5.12	Configuration Examples of Sending Log to the Console Terminal.....	46-32
Chapter 47	SNMP Configuration.....	47-1
47.1	SNMP Overview.....	47-1
47.2	SNMP Versions and Supported MIB.....	47-1
47.3	Configuring SNMP.....	47-3
47.3.1	Setting Community Names.....	47-3
47.3.2	Setting the System Information.....	47-4
47.3.3	Enabling/Disabling SNMP Agent to Send Trap.....	47-4
47.3.4	Setting the Destination Address of Trap.....	47-5
47.3.5	Setting Lifetime of Trap Message.....	47-5
47.3.6	Setting the Engine ID of a Local or Remote Device.....	47-6
47.3.7	Setting/Deleting an SNMP Group.....	47-6
47.3.8	Setting the Source Address of Trap.....	47-7
47.3.9	Adding/Deleting a User to/from an SNMP Group.....	47-7
47.3.10	Creating/Updating View Information or Deleting a View.....	47-8
47.3.11	Setting the Size of the SNMP Packet Sent/Received by an Agent.....	47-8
47.3.12	Disabling SNMP Agent.....	47-8
47.4	Displaying and Debugging SNMP.....	47-9

47.5 SNMP Configuration Example	47-9
Chapter 48 RMON Configuration	48-1
48.1 RMON Overview	48-1
48.2 Configuring RMON	48-1
48.2.1 Adding/Deleting an Entry to/from the Event Table.....	48-2
48.2.2 Adding/Deleting an Entry to/from the Alarm Table.....	48-2
48.2.3 Adding/Deleting an Entry to/from the Extended RMON Alarm Table	48-3
48.2.4 Adding/Deleting an Entry to/from the History Control Table	48-4
48.2.5 Adding/Deleting an Entry to/from the Statistics Table.....	48-5
48.3 Displaying and Debugging RMON.....	48-5
48.4 RMON Configuration Example	48-6
Chapter 49 NTP Configuration	49-1
49.1 Brief Introduction to NTP	49-1
49.1.1 NTP Functions.....	49-1
49.1.2 Basic Operating Principle of NTP.....	49-1
49.2 NTP Configuration	49-2
49.2.1 Configuring NTP Operating Mode.....	49-2
49.2.2 Configuring NTP ID Authentication	49-6
49.2.3 Setting NTP Authentication Key.....	49-6
49.2.4 Setting Specified Key as Reliable	49-7
49.2.5 Designating an Interface to Transmit NTP Messages	49-7
49.2.6 Setting NTP Master Clock.....	49-8
49.2.7 Setting Authority to Access a Local Ethernet Switch	49-8
49.2.8 Setting Maximum Local Sessions	49-9
49.3 Displaying and Debugging NTP.....	49-9
49.4 NTP Configuration Example	49-10
49.4.1 Configuring a NTP Server	49-10
49.4.2 NTP Peer Configuration Example	49-11
49.4.3 Configure NTP Broadcast Mode	49-13
49.4.4 Configure NTP Multicast Mode	49-14
49.4.5 Configure Authentication-Enabled NTP Server Mode	49-16
Chapter 50 SSH Terminal Service.....	50-1
50.1 SSH Terminal Service	50-1
50.1.1 SSH Overview	50-1
50.1.2 SSH Server Configuration.....	50-3
50.1.3 SSH Client Configuration	50-12
50.1.4 Displaying and Debugging SSH.....	50-13
50.1.5 SSH Server Configuration Example.....	50-13
50.1.6 SSH Client Configuration Example	50-16
50.2 SFTP Service.....	50-17
50.2.1 SFTP Overview	50-17
50.2.2 SFTP Server Configuration	50-17

50.2.3 SFTP Client Configuration	50-18
50.2.4 SFTP Configuration Example.....	50-22
Chapter 51 PoE Configuration	51-1
51.1 PoE Overview	51-1
51.1.1 PoE on the Switch.....	51-1
51.1.2 External PSE4500-A Power System.....	51-2
51.2 PoE Configuration.....	51-2
51.2.1 PoE Configuration Tasks	51-3
51.3 Comprehensive Configuration Example	51-5
Chapter 52 PoE PSU Supervision Configuration	52-1
52.1 Introduction to PoE PSU Supervision.....	52-1
52.2 AC Input Alarm Thresholds Configuration	52-1
52.2.1 AC Input Alarm Thresholds Configuration Tasks.....	52-1
52.2.2 AC Input Alarm Thresholds Configuration Example	52-2
52.3 DC Output Alarm Thresholds Configuration	52-2
52.3.1 DC Output Alarm Thresholds Configuration Tasks.....	52-3
52.3.2 DC Output Alarm Thresholds Configuration Example	52-3
52.4 Displaying PoE Supervision Information	52-4
52.5 PoE PSU Supervision Configuration Example	52-4

Chapter 1 Product Overview

1.1 Product Overview

The Switch 8800 is a large-capacity, modularized L2/L3 switch. It is mainly designed for broadband MAN, backbone, switching core and convergence center of large-sized enterprise network and campus network. It provides diverse services and can be used in constructing a stable and high-performance IP network.

The Switch 8800 supports the following services:

- Internet broadband access
- MAN, enterprise/campus networking
- Providing multicast service and multicast routing and supporting multicast audio and video services.

1.2 Function Features

Table 1-1 Function features

Features	Implementation
VLAN	Supports VLAN compliant with IEEE 802.1Q Standard Supports port-based and MAC-based VLAN Supports GARP VLAN Registration Protocol (GVRP)
STP protocol	Supports Spanning Tree Protocol (STP) / Multiple Spanning Tree Protocol (MSTP), compliant with IEEE 802.1D/IEEE 802.1s Standard
Flow control	Supports IEEE 802.3x flow control (full-duplex) Supports back-pressure based flow control (half-duplex)
Broadcast Suppression	Supports Broadcast Suppression
Multicast	Supports Internet Group Management Protocol Snooping (IGMP Snooping) Supports Internet Group Management Protocol (IGMP) Supports Protocol-Independent Multicast-Dense Mode (PIM-DM) Supports Protocol-Independent Multicast-Sparse Mode (PIM-SM) Supports Multicast Source Discovery Protocol (MSDP) Supports Multiprotocol BGP (MBGP)

Features	Implementation
IP routing	<ul style="list-style-type: none"> Supports static routing Supports Routing Information Protocol (RIP) v1/v2 Supports Open Shortest Path First (OSPF) Supports Border Gateway Protocol (BGP) Supports Intermediate System-to-Intermediate System intra-domain routing information exchange protocol (IS-IS) Supports IP routing policy
DHCP Relay	Supports Dynamic Host Configuration Protocol (DHCP) Relay
Link aggregation	Supports link aggregation, including two kinds of link aggregation LACPs (link aggregation control protocols): static aggregation and dynamic aggregation.
Mirror	<ul style="list-style-type: none"> Supports the port-based mirror Supports flow mirroring of copying messages to CPU
Quality of Service (QoS)	<ul style="list-style-type: none"> Supports traffic classification Supports bandwidth control Supports congestion control Supports traffic shaping and traffic supervision Supports queues of different priority on the port Queue scheduling: supports Strict Priority Queuing (SP), Weighted Round Robin (WRR), and SP+WRR
Security features	<ul style="list-style-type: none"> Supports Multi-level user management and password protect Supports 802.1X authentication Supports Packet filtering
MPLS	<ul style="list-style-type: none"> Supports Multiprotocol Label Switching (MPLS) basic function Supports MPLS L3 VPN
Management and Maintenance	<ul style="list-style-type: none"> Supports Command Line Interface configuration Supports local configuration via Console port and AUX port Supports Local and remote configuration through Telnet on Ethernet port Supports Remote configuration through dialing with modem via the AUX port. Supports SNMP management (Supports Quidview NMS and RMON MIB Group 1, 2, 3 and 9) Supports system log Supports level alarms Supports output of the debugging information Supports PING and Tracert Supports the remote maintenance via Telnet and Modem
Loading and updating	<ul style="list-style-type: none"> Supports to load and upgrade software via XModem protocol Supports to load and upgrade software via File Transfer Protocol (FTP) and Trivial File Transfer Protocol (TFTP)

Chapter 2 Logging into Switch

2.1 Setting Up Configuration Environment Through the Console Port

Step 1: As shown in the figure below, to set up the local configuration environment, connect the serial port of a PC (or a terminal) to the Console port of the switch with the Console cable.

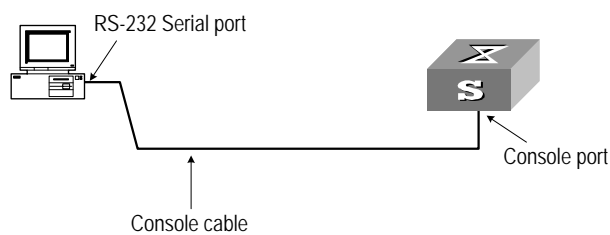


Figure 2-1 Set up the local configuration environment through the Console port

Step 2: Run terminal emulator (such as Terminal on Windows 3X or the Hyper Terminal on Windows 9X) on the Computer. Set the terminal communication parameters as follows: Set the baud rate to 9600, databit to 8, parity check to none, stopbit to 1, flow control to none and select the terminal type as VT100.



Figure 2-2 Set up new connection



Figure 2-3 Configure the port for connection

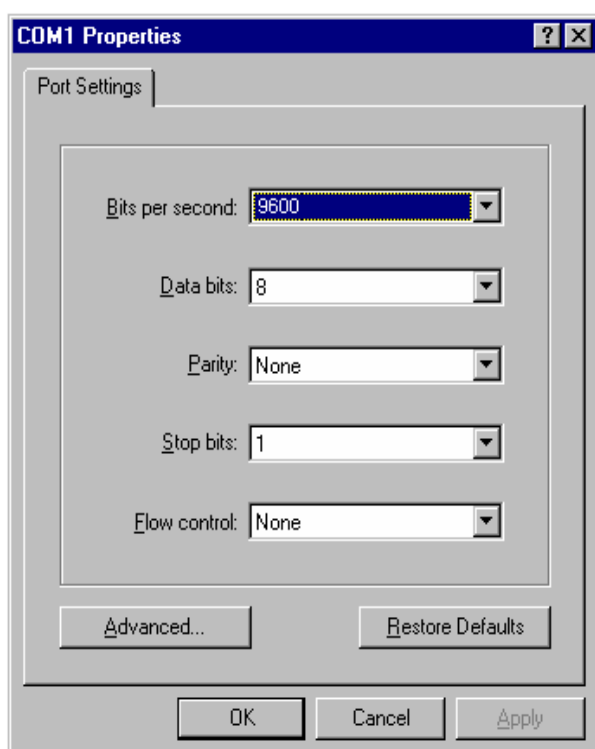


Figure 2-4 Set communication parameters

Step 3: The switch is powered on. Display self-test information of the switch and prompt you to press Enter to show the command line prompt such as <SW8800>.

Step 4: Input a command to configure the switch or view the operation state. Input a “?” for an immediate help. For details of specific commands, refer to the following chapters.

2.2 Setting Up Configuration Environment Through Telnet

2.2.1 Connecting a PC to the Switch Through Telnet

After you have correctly configured IP address of a VLAN interface for a switch via Console port (using **ip address** command in VLAN interface view), and added the port (that connects to a terminal) to this VLAN (using **port** command in VLAN view), you can telnet this switch and configure it.

Step 1: Before logging into the switch through telnet, you need to configure the Telnet user name and password on the switch through the console port.

Note:

By default, the password is required for authenticating the Telnet user to log in the switch. If a user logs in via the Telnet without password, he will see the prompt "Login password has not been set !".

```
<SW8800> system-view
Enter system view , return user view with Ctrl+Z.
[SW8800] user-interface vty 0
[SW8800-ui-vty0] set authentication password simple xxxx (xxxx is the preset login
password of Telnet user)
```

Step 2: To set up the configuration environment, connect the Ethernet port of the PC to that of the switch via the LAN, as shown in Figure 2-5.

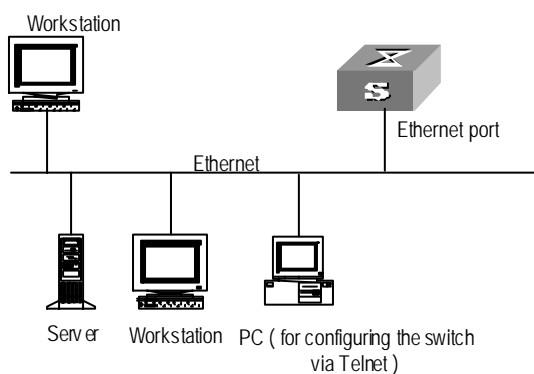


Figure 2-5 Set up configuration environment through telnet

Step 3: Run Telnet on the PC and input the IP address of the VLAN connected to the PC port, as shown in Figure 2-6.

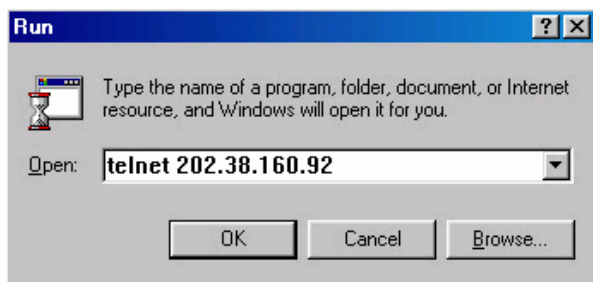


Figure 2-6 Run Telnet

Step 4: The terminal displays “Login authentication!” and prompts the user to input the logon password. After you input the correct password, it displays the command line prompt (such as <SW8800>). If the prompt “All user interfaces are used, please try later! The connection was closed by the remote host!” appears, it indicates that the maximum number of Telnet users that can be accessed to the switch is reached at this moment. In this case, please reconnect later. At most 5 Telnet users are allowed to log on to the Switch 8800 simultaneously.

Step 5: Use the corresponding commands to configure the switch or to monitor the running state. Enter “?” to get the immediate help. For details of specific commands, refer to the following chapters.

Note:

- When configuring the switch via Telnet, do not modify the IP address of it unless necessary, for the modification might cut the Telnet connection.
 - By default, when a Telnet user passes the password authentication to log on to the switch, he can access the commands at Level 0.
-

2.2.2 Telneting a Switch Through Another Switch

After a user has logged into a switch, he or she can configure another switch through the switch via Telnet. The local switch serves as Telnet client and the peer switch serves as Telnet server. If the ports connecting these two switches are in a same local network, their IP addresses must be configured in the same network segment. Otherwise, the two switches must establish a route that can reach each other.

As shown in the figure below, after you telnet to a switch, you can run **telnet** command to log in and configure another switch.



Figure 2-7 Provide Telnet Client service

Step 1: Configure the Telnet user name and password on the Telnet Server through the console port.

Note:

By default, the password is required for authenticating the Telnet user to log in the switch. If a user logs in via the Telnet without password, he will see the prompt “Login password has not been set !.”.

```
<SW8800> system-view
System View: return to User View with Ctrl+Z
[SW8800] user-interface vty 0
[SW8800-ui-vty0] set authentication password simple xxxx (xxxx is the preset login
password of Telnet user)
```

Step 2: The user logs in the Telnet Client (switch). For the login process, refer to the section describing “Connecting a PC to the Switch through Telnet”.

Step 3: Perform the following operations on the Telnet Client:

```
<SW8800> telnet xxxx (xxxx can be the hostname or IP address of the Telnet Server. If it is the
hostname, you need to use the ip host command to specify.)
```

Step 4: Enter the preset login password and you will see the prompt such <SW8800>. If the prompt “All user interfaces are used, please try later! The connection was closed by the remote host!” appears, it indicates that the maximum number of Telnet users that can be accessed to the switch is reached at this moment. In this case, please connect later.

Step 5: Use the corresponding commands to configure the switch or view its running state. Enter “?” to get the immediate help. For details of specific commands, refer to the following chapters.

2.3 Setting Up Configuration Environment Through a Dial-up the Modem

Step 1: Authenticate the Modem user via the Console port of the switch before he logs in the switch through a dial-up Modem.

Note:

By default, the password is required for authenticating the Modem user to log in the switch. If a user logs in via the Modem without password, he will see the prompt “Login password has not been set !.”

```
<SW8800> system-view
System View: return to User View with Ctrl+Z..
[SW8800] user-interface aux 0
[SW8800-ui-aux0] set authentication password simple xxxx (xxxx is the preset login
password of the Modem user.)
```

Step 2: As shown in the figure below, to set up the remote configuration environment, connect the Modems to a PC (or a terminal) serial port and the switch AUX port respectively.

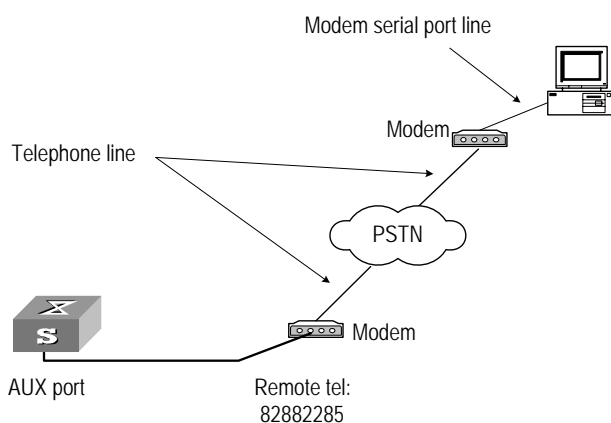


Figure 2-8 Set up remote configuration environment

Step 3: Dial for connection to the switch, using the terminal emulator and Modem on the remote end. The number dialed shall be the telephone number of the Modem connected to the switch. See the two figures below.



Figure 2-9 Set the dialed number

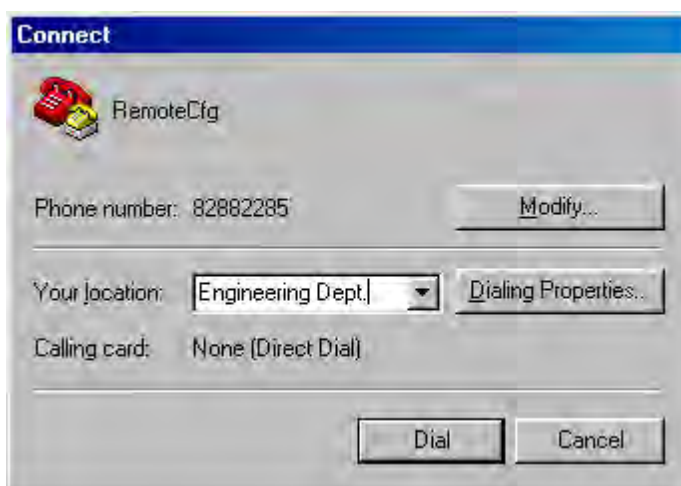


Figure 2-10 Dial on the remote PC

Step 4: Enter the preset login password on the remote terminal emulator and wait for the prompt such as <SW8800>. Then you can configure and manage the switch. Enter “?” to get the immediate help. For details of specific commands, refer to the following chapters.

Note:

By default, when a Modem user logs in, he can access the commands at Level 0.

Chapter 3 Command Line Interface

3.1 Command Line Interface

The Switch 8800 provides a series of configuration commands and command line interfaces for configuring and managing the switch. The command line interface has the following characteristics:

- Local configuration via the Console port and AUX port.
- Local or remote configuration via Telnet.
- Remote configuration through dialing with modem via the AUX port.
- Hierarchy command protection to avoid the unauthorized users accessing switch.
- Enter a “?” to get immediate online help.
- Provide network testing commands, such as Tracert and Ping, to fast troubleshoot the network.
- Provide various detailed debugging information to help with network troubleshooting.
- Log in and manage other switch directly, using the Telnet command.
- Provide FTP service for the users to upload and download files.
- Provide the function similar to Doskey to execute a history command.
- The command line interpreter searches for target not fully matching the keywords. It is ok for you to key in the whole keyword or part of it, as long as it is unique and not ambiguous.

3.2 Command Line View

The Switch 8800 provides hierarchy protection for the command lines to avoid unauthorized user accessing illegally.

Commands are classified into four levels, namely visit level, monitoring level, configuration level and management level. They are introduced as follows:

- Visit level: Commands of this level involve command of network diagnosis tool (such as **ping** and **tracert**), command of switch between different language environments of user interface (**language-mode**) and **telnet** command etc. The operation of saving configuration file is not allowed on this level of commands.
- Monitoring level: Commands of this level, including the **display** command and the **debugging** command, are used to system maintenance, service fault diagnosis, etc. The operation of saving configuration file is not allowed on this level of commands.
- Configuration level: Service configuration commands, including routing command and commands on each network layer, are used to provide direct network service to the user.

- Management level: They are commands that influence basis operation of the system and system support module, which plays a support role on service. Commands of this level involve file system commands, FTP commands, TFTP commands, XModem downloading commands, user management commands, and level setting commands.

At the same time, login users are classified into four levels that correspond to the four command levels respectively. After users of different levels log in, they can only use commands at the levels that are equal to or lower than its own level.

In order to prevent unauthorized users from illegal intrusion, user will be identified when switching from a lower level to a higher level with **super** [*level*] command. User ID authentication is performed when users at lower level switch to users at higher level. In other words, user password of the higher level is needed (Suppose the user has set the **super password** [*level level*] { **simple** | **cipher** } *password*.) For the sake of confidentiality, on the screen the user cannot see the password that he entered. Only when correct password is input for three times, can the user switch to the higher level. Otherwise, the original user level will remain unchanged.

Different command views are implemented according to different requirements. They are related to one another. For example, after logging in the switch, you will enter user view, in which you can only use some basic functions such as displaying the running state and statistics information. In user view, key in **system-view** to enter system view, in which you can key in different configuration commands and enter the corresponding views.

The command line provides the following views:

- User view
- System view
- Port view
- VLAN view
- VLAN interface view
- Local-user view
- User interface view
- FTP Client view
- SFTP Client view
- MST region view
- PIM view
- MSDP view
- IPv4 multicast sub-address family view
- RIP view
- OSPF view
- OSPF area view
- BGP view
- IS-IS view

- Route policy view
- Basic ACL view
- Advanced ACL view
- Layer-2 ACL view
- Conform-level view
- WRED index view
- RADIUS server group view
- ISP domain view
- MPLS view
- VPNv4 sub-address family view
- VPN-instance sub-address family view
- Remote-peer view
- VSI-LDP view
- VSI view
- TACACS+ view
- Port group view
- Lanswitch view

The following table describes the function features of different views and the ways to enter or quit.

Table 3-1 Function feature of command view

Command view	Function	Prompt	Command to enter	Command to exit
User view	Show the basic information about operation and statistics	<SW8800 >	Enter right after connecting the switch	quit disconnects to the switch
System view	Configure system parameters	[SW8800]	Key system-view in user view	quit or return returns to user view

Command view	Function	Prompt	Command to enter	Command to exit
Port view	Ethernet port view: Configure Ethernet port parameters	[SW8800-Ethernet2/1/1]	100M Ethernet port view Key in interface ethernet 2/1/1 in system view	quit returns to system view return returns to user view
		[SW8800-GigabitEthernet2/1/1]	GigabitEthernet port view Key in interface gigabitethernet 2/1/1 in system view	
		[SW8800-10-GigabitEthernet2/1/1]	10G Ethernet port view Key in interface 10-gigabitethernet 2/1/1 in system view	
VLAN view	Configure VLAN parameters	[SW8800-Vlan1]	Key in vlan 1 in system view	quit returns to system view return returns to user view
VLAN interface view	Configure IP interface parameters for a VLAN or a VLAN aggregation	[SW8800-Vlan-interface1]	Key in interface vlan-interface 1 in system view	quit returns to system view return returns to user view
Local-user view	Configure local user parameters	[SW8800-Iuser-user1]	Key in local-user user1 in system view	quit returns to system view return returns to user view
User interface view	Configure user interface parameters	[SW8800-ui0]	Key in user-interface 0 in system view	quit returns to system view return returns to user view
FTP Client view	Configure FTP Client parameters	[ftp]	Key in ftp in user view	quit returns to system view
SFTP Client view	Configure SFTP Client parameters	sftp-client >	Key in sftp ip-address in system view	quit returns to system view return returns to user view
MST region view	Configure MST region parameters	[SW8800-mst-region]	Key in stp region-configuration in system view	quit returns to system view return returns to user view

Command view	Function	Prompt	Command to enter	Command to exit
PIM view	Configure PIM parameters	[SW8800-PIM]	Key in pim in system view	quit returns to system view return returns to user view
MSDP view	Configure MSDP parameters	[SW8800-msdp]	Key in msdp in system view	quit returns to system view return returns to user view
IPv4 multicast sub-addresses family view	Enter the IPv4 multicast sub-address family view to configure MBGP multicast extension parameters	[SW8800-bgp-af-multicast]	Key in ipv4-family multicast in BGP view	quit returns to BGP view return returns to user view
RIP view	Configure RIP parameters	[SW8800-rip]	Key in rip in system view	quit returns to system view return returns to user view
OSPF view	Configure OSPF parameters	[SW8800-ospf]	Key in ospf in system view	quit returns to system view return returns to user view
OSPF area view	Configure OSPF area parameters	[SW8800-ospf-0.0.0.1]	Key in area 1 in OSPF view	quit returns to OSPF view return returns to user view
BGP view	Configure BGP parameters	[SW8800-bgp]	Key in bgp 100 in system view	quit returns to system view return returns to user view
IS-IS view	Configure IS-IS parameters	[SW8800-isis]	Key in isis in system view	quit returns to system view return returns to user view
Route policy view	Configure route policy parameters	[SW8800-route-policy]	Key in route-policy policy1 permit node 10 in system view	quit returns to system view return returns to user view
Basic ACL view	Define the rule of basic ACL	[SW8800-acl-basic-2000]	Key in acl number 2000 in system view	quit returns to system view return returns to user view

Command view	Function	Prompt	Command to enter	Command to exit
Advanced ACL view	Define the rule of advanced ACL	[SW8800-acl-adv-3000]	Key in acl number 3000 in system view	quit returns to system view return returns to user view
Layer-2 ACL view	Define the rule of layer-2 ACL	[SW8800-acl-link-4000]	Key in acl number 4000 in system view	quit returns to system view return returns to user view
Conform-level view	Configure the "DSCP + Conform-level Service group" mapping table and "EXP + Conform-level->service parameters" mapping table and "Local-precedence + Conform-level 802.1p priority" mapping table	[SW8800-conform-level-0]	Key in qos conform-level 0 in system view	quit returns to system view return returns to user view
WRED index view	Configure WRED parameters	[SW8800-wred-0]	Key in wred 0 in system view	quit returns to system view return returns to user view
RADIUS server group view	Configure radius parameters	[SW8800-radius-1]	Key in radius scheme 1 in system view	quit returns to system view return returns to user view
ISP domain view	Configure ISP domain parameters	[SW8800-isp-3Com163.net]	Key in domain 3Com163.net in system view	quit returns to system view return returns to user view
MPLS view	Configure MPLS parameters	[SW8800-mpls]	Key in mpls in system view	quit returns to system view return returns to user view
VPNv4 subaddresses family view	Configure VPNv4 address family parameters	[SW8800-bgp-af-vpn]	Key in ipv4-family vpnv4 in BGP view	quit returns to system view return returns to user view

Command view	Function	Prompt	Command to enter	Command to exit
VPN-instance subaddress family view	Configure VPN instance subaddress family parameters	[SW8800-bgp-af-vpn-instance]	Key in ipv4-family vpn-instance vpna in BGP/RIP view	quit returns to system view return returns to user view
Remote-peer view	Configure MPLS peer group parameters	[SW8800-mpls-remote1]	Key in mpls remote1	quit returns to system view return returns to user view
VSI-LDP view	Configure some VPLS features	[SW8800-vsi-3Com-ldp]	Key in vsi 3Com in system view Key in pwsignal ldp in vsi view	quit returns to vsi view return returns to user view
VSI view	Specify VPLS mode	[SW8800-vsi-3Com]	Key in vsi 3Com in system view	quit returns to system view return returns to user view
TACACS+ view	Configure TACACS+ protocol parameters	[SW8800-tacacs+-3Com]	Key in tacacs+ scheme 3Com in system view	quit returns to system view return returns to user view
Port group view	Combine the ports with the same configuration, omitting repeated configuration procedure	[SW8800-port-group X]	Key in port-group X in system view	quit returns to system view return returns to user view
Lanswitch view	Enter lanswitch view. After entering the specified lanswitch view, you can use the specified Ethernet switch.	[SW8800-lanswitchX /X/X-/]	Key in lanswitch X/X/X-/ in HGMP view	quit returns to HGMP view return returns to user view

3.3 Features and Functions of Command Line

3.3.1 Online Help of Command Line

The command line interface provides the following online help modes.

- Full help

- Partial help

You can get the help information through these online help commands, which are described as follows.

1) Input “?” in any view to get all the commands in it and corresponding descriptions.

```
<SW8800> ?
```

```
User view commands:
```

```
language-mode  Specify the language environment
ping           Ping function
quit          Exit from current command view
super         Privilege current user a specified priority level
telnet        Establish one TELNET connection
tracert       Trace route function
```

2) Input a command with a “?” separated by a space. If this position is for keywords, all the keywords and the corresponding brief descriptions will be listed.

```
<SW8800> language-mode ?
```

```
chinese  Chinese environment
english  English environment
```

3) Input a command with a “?” separated by a space. If this position is for parameters, all the parameters and their brief descriptions will be listed.

```
[SW8800] garp timer leaveall ?
```

```
INTEGER<65-32765>  Value of timer in centiseconds
                    (LeaveAllTime > (LeaveTime [On all ports]))
                    Time must be multiple of 5 centiseconds
```

```
[SW8800] garp timer leaveall 300 ?
```

```
<cr>
```

<cr> indicates no parameter in this position. The next command line repeats the command, you can press <Enter> to execute it directly.

4) Input a character string with a “?”, then all the commands with this character string as their initials will be listed.

```
<SW8800> p?
```

```
ping  pwd
```

5) Input a command with a character string and “?”, then all the key words with this character string as their initials in the command will be listed.

```
<SW8800> display ver?
```

```
version
```

6) Input the first letters of a keyword of a command and press <Tab> key. If no other keywords are headed by this letters, then this unique keyword will be displayed automatically.

7) To switch to the Chinese display for the above information, perform the language-mode command.

3.3.2 Displaying Characteristics of Command Line

Command line interface provides the following display characteristics:

- For users' convenience, the instruction and help information can be displayed in both English and Chinese.
- For the information to be displayed exceeding one screen, pausing function is provided. In this case, users can have three choices, as shown in the table below.

Table 3-2 Functions of displaying

Key or Command	Function
Press <Ctrl+C> when the display pauses	Stop displaying and executing command.
Enter a space when the display pauses	Continue to display the next screen of information.
Press <Enter> when the display pauses	Continue to display the next line of information.

3.3.3 History Command of Command Line

Command line interface provides the function similar to that of DosKey. The commands entered by users can be automatically saved by the command line interface and you can invoke and execute them at any time later. History command buffer is defaulted as 10. The operations are shown in the table below.

Table 3-3 Retrieve history command

Operation	Key	Result
Display history command	display history-command	Display history command by user inputting
Retrieve the previous history command	Up cursor key <↑> or <Ctrl+P>	Retrieve the previous history command, if there is any.
Retrieve the next history command	Down cursor key <↓> or <Ctrl+N>	Retrieve the next history command, if there is any.

Note:

Cursor keys can be used to retrieve the history commands in Windows 3.X Terminal and Telnet. However, in Windows 9X HyperTerminal, the cursor keys ↑ and ↓ do not work, because Windows 9X HyperTerminal defines the two keys differently. In this case, use the combination keys <Ctrl+P> and <Ctrl+N> instead for the same purpose.

3.3.4 Common Command Line Error Messages

All the input commands by users can be correctly executed, if they have passed the grammar check. Otherwise, error messages will be reported to users. The common error messages are listed in the following table.

Table 3-4 Common command line error messages

Error messages	Causes
Unrecognized command	Cannot find the command.
	Cannot find the keyword.
	Wrong parameter type.
	The value of the parameter exceeds the range.
Incomplete command	The input command is incomplete.
Too many parameters	Enter too many parameters.
Ambiguous command	The parameters entered are not specific.

3.3.5 Editing Characteristics of Command Line

Command line interface provides the basic command editing function and supports to edit multiple lines. A command cannot longer than 256 characters. See the table below.

Table 3-5 Editing functions

Key	Function
Common keys	Insert from the cursor position and the cursor moves to the right, if the edition buffer still has free space.
Backspace	Delete the character preceding the cursor and the cursor moves backward.
Leftwards cursor key <←> or <Ctrl+B>	Move the cursor a character backward
Rightwards cursor key <→> or <Ctrl+F>	Move the cursor a character forward
Up cursor key <↑> or <Ctrl+P> Down cursor key <↓> or <Ctrl+N>	Retrieve the history command.

<Tab>	Press <Tab> after typing the incomplete key word and the system will execute the partial help: If the key word matching the typed one is unique, the system will replace the typed one with the complete key word and display it in a new line; if there is not a matched key word or the matched key word is not unique, the system will do no modification but display the originally typed word in a new line.
-------	---

Chapter 4 User Interface Configuration

4.1 User Interface Overview

User interface configuration is another way provided by the switch to configure and manage the port data.

The Switch 8800 supports the following configuration methods:

- Local configuration via the Console port and AUX port
- Local and remote configuration through Telnet on Ethernet port
- Remote configuration through dialing with modem via the AUX port.

According to the above-mentioned configuration methods, there are three types of user interfaces:

- Console user interface

Console user interface is used to log in the switch via the Console port. A switch can only have one Console user interface.

- AUX user interface

AUX user interface is used to log in the switch locally or remotely with a modem via the AUX port. A switch can only have one AUX user interface. The local configuration for it is similar to that for the Console user interface.

- VTY user interface

VTY user interface is used to telnet the switch. A switch can have up to five VTY user interface.

User interface is numbered in the following two ways: absolute number and relative number.

I. Absolute number

The user interfaces for IP PBX fall into three types and they are sequenced as follows: console interface (CON), auxiliary interface (AUX) and virtual interface (VTY). A switch has one CON, one AUX and multiple VTYS. The first absolute number is designated as 0; the second one is designated as 1; and so on. This method can specify a unique user interface or a group of interfaces.

It follows the rules below.

- Console user interface is numbered as the first interface designated as user interface 0.
- AUX user interface is numbered as the second interface designated as user interface 1.

- VTY is numbered after AUX user interface. The absolute number of the first VTY is incremented by 1 than the AUX user interface number.

II. Relative number

The relative number is in the format of “user interface type” + “number”. The “number” refers to the internal number for each user interface type. This method can only specify one interface or one group of interfaces for a user interface type instead of different user interface types.

It follows the rules below:

- Number of Console user interface: console 0.
- Number of AUX user interface: AUX 0.
- Number of VTY: The first VTY interface is designated as VTY 0; the second one is designated as VTY 1, and so on.

4.2 User Interface Configuration

The following sections describe the user interface configuration tasks.

- Entering User Interface View
- Define the Login Header
- Configuring Asynchronous Port Attributes
- Configuring Terminal Attributes
- Managing Users
- Configuring Modem Attributes
- Configuring Redirection

4.2.1 Entering User Interface View

The following command is used for entering a user interface view. You can enter a single user interface view or multi user interface view to configure one or more user interfaces respectively.

Perform the following configuration in system view.

Table 4-1 Enter user interface view

Operation	Command
Enter a single user interface view or multi user interface views	user-interface [<i>type</i>] <i>first-number</i> [<i>last-number</i>]

4.2.2 Define the Login Header

The following command is used for configuring the displayed header when user login.

When the users log in the switch, if a connection is activated, the **login** header will be displayed. After the user successfully logs in the switch, the **shell** header will be displayed.

Perform the following configuration in system view.

Table 4-2 Configure the login header.

Operation	Command
Configure the login header	header [shell incoming login] <i>text</i>
Remove the login header configured	undo header [shell incoming login]

Note that if you press <Enter> after typing any of the three keywords **shell**, **login** and **incoming** in the command, then what you type after the word header is the contents of the login information, instead of identifying header type.

4.2.3 Configuring Asynchronous Port Attributes

The following commands can be used for configuring the attributes of the asynchronous port in asynchronous interactive mode, including speed, flow control, parity, stop bit and data bit.

Perform the following configurations in user interface (Console and AUX user interface only) view.

I. Configuring the transmission speed

Table 4-3 Configure the transmission speed

Operation	Command
Configure the transmission speed	speed <i>speed-value</i>
Restore the default transmission speed	undo speed

By default, the transmission speed on an asynchronous port is 9600bps.

II. Configuring flow control

Table 4-4 Configure flow control

Operation	Command
Configure the flow control	flow-control { hardware none software }
Restore the default flow control mode	undo flow-control

By default, the flow control on an asynchronous port is none, that is, no flow control will be performed.

III. Configuring parity

Table 4-5 Configure parity

Operation	Command
Configure parity mode	parity { even mark none odd space }
Restore the default parity mode	undo parity

By default, the parity on an asynchronous port is none, that is, no parity bit.

IV. Configuring the stop bit

Table 4-6 Configure the stop bit

Operation	Command
Configure the stop bit	stopbits { 1 1.5 2 }
Restore the default stop bit	undo stopbits

By default, an asynchronous port supports 1 stop bit.

Note that setting 1.5 stop bits is not available on the Switch 8800.

V. Configuring the data bit

Table 4-7 Configure the data bit

Operation	Command
Configure the data bit	databits { 7 8 }
Restore the default data bit	undo databits

By default, an asynchronous port supports 8 data bits.

4.2.4 Configuring Terminal Attributes

The following commands can be used for configuring the terminal attributes, including enabling/disabling terminal service, disconnection upon timeout, lockable user interface, configuring terminal screen length and history command buffer size.

Perform the following configuration in user interface view. Perform **lock** command in user view.

I. Enabling/disabling terminal service

After the terminal service is disabled on a user interface, you cannot log in to the switch through the user interface. However, the user logged in through the user interface before disabling the terminal service can continue his operation. After such user logs out, he cannot log in again. In this case, a user can log in to the switch through the user interface only when the terminal service is enabled again.

Table 4-8 Enable/disable terminal service

Operation	Command
Enable terminal service	shell
Disable terminal service	undo shell

By default, terminal service is enabled on all the user interfaces.

Note the following points:

- For the sake of security, the **undo shell** command can only be used on the user interfaces other than Console user interface.
- You cannot use this command on the user interface via which you log in.
- You will be asked to confirm before using **undo shell** on any legal user interface.

II. Configuring idle-timeout

Table 4-9 Configure idle-timeout

Operation	Command
Configure idle-timeout	idle-timeout <i>minutes</i> [<i>seconds</i>]
Restore the default idle-timeout	undo idle-timeout

By default, idle-timeout is enabled and set to 10 minutes on all the user interfaces. That is, the user interface will be disconnected automatically after 10 minutes without any operation.

idle-timeout 0 means disabling idle-timeout.

III. Locking user interface

This configuration is to lock the current user interface and prompt the user to enter the password. This makes it impossible for others to operate in the interface after the user leaves.

Table 4-10 Lock user interface

Operation	Command
Lock user interface	lock

IV. Setting the screen length

If a command displays more than one screen of information, you can use the following command to set how many lines to be displayed in a screen, so that the information can be separated in different screens and you can view it more conveniently.

Table 4-11 Set the screen length

Operation	Command
Set the screen length	screen-length <i>screen-length</i>
Restore the default screen length	undo screen-length

By default, the terminal screen length is 24 lines.

screen-length 0 indicates to disable screen display separation function.

V. Setting the history command buffer size

Table 4-12 Set the history command buffer size

Operation	Command
Set the history command buffer size	history-command max-size <i>value</i>
Restore the default history command buffer size	undo history-command max-size

By default, the size of the history command buffer is 10, that is, 10 history commands can be saved.

4.2.5 Managing Users

The management of users includes the setting of user logon authentication method, level of command which a user can use after logging on, level of command which a user can use after logging on from the specifically user interface, and command level.

I. Configuring the authentication method

The following command is used for configuring the user login authentication method to deny the access of an unauthorized user.

Perform the following configuration in user interface view.

Table 4-13 Configure the authentication method

Operation	Command
Configure the authentication method	authentication-mode { password scheme none }

By default, terminal authentication is not required for local users log in via the Console port. However, password authentication is required for local users and remote Modem users log in via the AUX port, and Telnet users log in through Ethernet port.

1) Perform local password authentication to the user interface

Using **authentication-mode password** command, you can perform local password authentication. That is, you need use the command below to configure a login password in order to login successfully.

Perform the following configuration in user interface view.

Table 4-14 Configure the local authentication password

Operation	Command
Configure the local authentication password	set authentication password { cipher simple }password
Remove the local authentication password	undo set authentication password

Configure for password authentication when a user logs in through a VTY 0 user interface and set the password to 3Com.

```
[SW8800] user-interface vty 0
[SW8800-ui-vty0] authentication-mode password
[SW8800-ui-vty0] set authentication password simple 3Com
```

2) Perform local or remote authentication of username and password to the user interface

Using **authentication-mode scheme** command, you can perform local or remote authentication of username and password. The type of the authentication depends on your configuration. For detailed information, see “Security” section.

In the following example, local username and password authentication are configured.

Perform username and password authentication when a user logs in through VTY 0 user interface and set the username and password to zbr and 3Com respectively.

```
[SW8800-ui-vty0] authentication-mode scheme
[SW8800-ui-vty0] quit
[SW8800] local-user zbr
[SW8800-luser-zbr] password simple 3Com
[SW8800-luser-zbr] service-type telnet
```

3) No authentication

```
[SW8800-ui-vty0] authentication-mode none
```

Note:

By default, password is required to be set for authenticating local users and remote Modem users log in via the AUX port, and Telnet users log in through Ethernet port. If no password has been set, the following prompt will be displayed “Login password has not been set !.”

If the **authentication-mode none** command is used, the local and Modem users via the AUX port and Telnet users will not be required to input password.

II. Setting the command level used after a user logging in

The following command is used for setting the command level used after a user logging in.

Perform the following configuration in local-user view.

Table 4-15 Set the command level used after a user logging in

Operation	Command
Set command level used after a user logging in	service-type telnet [level <i>level</i>]
Restore the default command level used after a user logging in	undo service-type telnet [level]

By default, the specified logon user can access the commands at Level 0.

III. Setting the command level used after a user logs in from a user interface

You can use the following command to set the command level after a user logs in from a specific user interface, so that a user is able to execute the commands at such command level.

Perform the following configuration in user interface view.

Table 4-16 Set the command level used after a user logging in from a user interface

Operation	Command
Set command level used after a user logging in from a user interface	user privilege level <i>level</i>
Restore the default command level used after a user logging in from a user interface	undo user privilege level

By default, you can access the commands at Level 3 after logging in through the Console user interface, and the commands at Level 0 after logging in through the AUX or VTY user interface.

Note:

When a user logs in the switch, the command level that it can access depends on two points. One is the command level that the user itself can access, the other is the set command level of this user interface. If the two levels are different, the former will be taken. For example, the command level of VTY 0 user interface is 1, however, you have the right to access commands of level 3; if you log in from VTY 0 user interface, you can access commands of level 3 and lower.

IV. Setting the command priority

The following command is used for setting the priority of a specified command in a certain view. The command levels include visit, monitoring, configuration, and management, which are identified with 0 through 3 respectively. An administrator assigns authorities as per user requirements.

Perform the following configuration in system view.

Table 4-17 Set the command priority

Operation	Command
Set the command priority in a specified view.	command-privilege level level view view command
Restore the default command level in a specified view.	Undo command-privilege view view command

V. Setting input protocol for a user terminal

You can use the following command to set input protocol for a user terminal. The input protocol type can be TELNET, SSH or all.

Perform the following configuration in user interface view.

Table 4-18 Set input protocol for a user terminal

Operation	Command
Set input protocol for a user terminal	protocol inbound { all telnet ssh }

By default, the input protocol type for a user terminal is all.

4.2.6 Configuring Modem Attributes

When logging in the switch via the Modem, you can use the following commands to configure these parameters.

Perform the following configuration in AUX user interface view.

Table 4-19 Configure Modem attributes

Operation	Command
Set the interval since the system receives the RING until CD_UP	modem timer answer <i>seconds</i>
Restore the default interval since the system receives the RING until CD_UP	undo modem timer answer
Configure auto answer	modem auto-answer
Configure manual answer	undo modem auto-answer
Configure to allow call-in	modem call-in
Configure to bar call-in	undo modem call-in
Configure to permit call-in and call-out.	modem both
Configure to disable call-in and call-out	undo modem both

4.2.7 Configuring Redirection

I. Send command

The following command can be used for sending messages between user interfaces.

Perform the following configuration in user view.

Table 4-20 Configure to send messages between different user interfaces.

Operation	Command
Configure to send messages between different user interfaces.	send { all <i>number</i> <i>type number</i> }

II. Auto-execute command

The following command is used to automatically run a command after you log in. After a command is configured to be run automatically, it will be automatically executed when you log in again.

This command is usually used to automatically execute **telnet** command on the terminal, which will connect the user to a designated device automatically.

Perform the following configuration in user interface view.

Table 4-21 Configure to automatically run the command

Operation	Command
Configure to automatically run the command	auto-execute command <i>text</i>
Configure not to automatically run the command	undo auto-execute command

Note the following points:

- After executing this command, the user interface can no longer be used to carry out the routine configurations for the local system. Use this command with caution.
- Make sure that you will be able to log in the system in some other way and cancel the configuration, before you use the **auto-execute command** command and save the configuration.

Telnet 10.110.100.1 after the user logs in through VTY0 automatically.

```
[SW8800-ui-vty0] auto-execute command telnet 10.110.100.1
```

When a user logs on via VTY 0, the system will run **telnet** 10.110.100.1 automatically.

4.3 Displaying and Debugging User Interface

After the above configuration, execute **display** command in any view to display the running of the user interface configuration, and to verify the effect of the configuration.

Execute **free** command in user view to clear a specified user interface.

Table 4-22 Display and debug user interface

Operation	Command
Clear a specified user interface	free user-interface [<i>type</i>] <i>number</i>
Display the user application information of the user interface	display users [all]
Display the physical attributes and some configurations of the user interface	display user-interface [<i>type number</i> <i>number</i>] [summary]

Chapter 5 Management Interface Configuration

5.1 Management Interface Overview

The Switch 8800 provides a 10/100Base-TX management interface on the Fabric. The management interface can connect a background PC for software loading and system debugging, or a remote network management station for remote system management.

5.2 Management Interface Configuration

The following sections describe management interface configuration tasks.

- Configuring interface IP address
- Enabling/disabling the interface
- Setting interface description
- Displaying current system information
- Test network connectivity (**ping**, **tracert**)

See the Port and System Management parts of this manual for details.



Caution:

Only the management interface configured with an IP address can run normally.

Chapter 6 Ethernet Port Configuration

6.1 Ethernet Port Overview

The Switch 8800 provides conventional Ethernet ports, fast Ethernet ports, 1000 Mbps Ethernet ports and 10 Gbps Ethernet ports. The configurations of these Ethernet ports are basically the same, which will be described in the following sections.

6.2 Ethernet Port Configuration

The following sections describe Ethernet port configuration tasks:

- Entering Ethernet Port View
- Enabling/Disabling an Ethernet Port
- Setting Ethernet Port Description
- Setting the Duplex Attribute of the Ethernet Port
- Setting Speed on the Ethernet Port
- Setting the Cable Type for the Ethernet Port
- Enabling/Disabling Flow Control for the Ethernet Port
- Permitting/Forbidding Jumbo Frame to Pass the Ethernet Port
- Setting the Ethernet Port Broadcast Suppression Ratio
- Setting the Ethernet Port Mode
- Setting the Link Type for the Ethernet Port
- Adding the Ethernet Port to Specified VLANs
- Setting the Default VLAN ID for the Ethernet Port
- Setting the VLAN VPN Feature
- Copying Port Configuration to Other Ports
- Setting Port Hold Time
- Setting the Ethernet Port in Loopback Mode

6.2.1 Entering Ethernet Port View

Before configuring the Ethernet port, enter Ethernet port view first.

Perform the following configuration in system view.

Table 6-1 Enter Ethernet port view

Operation	Command
Enter Ethernet port view	<code>interface { interface_type interface_num interface_name }</code>

6.2.2 Enabling/Disabling an Ethernet Port

After configuring the related parameters and protocol of the port, you can use **undo shutdown** command to enable the port. If you do not want a port to forward data any more, use **shutdown** command to disable it.

Perform the following configuration in Ethernet port view.

Table 6-2 Enable/disable an Ethernet port

Operation	Command
Disable an Ethernet port	shutdown
Enable an Ethernet port	undo shutdown

By default, the port is enabled.

6.2.3 Setting Ethernet Port Description

To distinguish the Ethernet ports, you can use the following command to make some necessary descriptions.

Perform the following configuration in Ethernet port group view.

Table 6-3 Set Ethernet port description

Operation	Command
Set an Ethernet port description	description <i>text</i>
Delete the Ethernet port description	undo description

By default, an Ethernet port has no description.

6.2.4 Setting the Duplex Attribute of the Ethernet Port

To configure a port to send and receive data packets at the same time, set it to full-duplex. To configure a port to either send or receive data packets at a time, set it to half-duplex. If the port has been set to auto-negotiation mode, the local and peer ports will automatically negotiate about the duplex mode.

Perform the following configuration in Ethernet port view.

Table 6-4 Set the duplex attribute for the Ethernet port

Operation	Command
Set duplex attribute for Ethernet port	duplex { auto full half }
Restore the default duplex attribute of Ethernet port	undo duplex

Note that, 10/100 Mbps electrical Ethernet port can operate in full-duplex, half-duplex or auto-negotiation mode. The 10/100/1000 Mbps electrical Ethernet port can operate in full duplex, half duplex or auto-negotiation mode. When the port operates at 1000 Mbps or in auto mode, the duplex mode can be set to **full** (full duplex) or **auto** (auto-negotiation). The optical 100/1000 Mbps and 10 Gbps Ethernet ports work in full duplex mode without user intervention.

The port defaults the **auto** (auto-negotiation) mode.

6.2.5 Setting Speed on the Ethernet Port

You can use the following command to set the speed on the Ethernet port. If the speed is set to auto-negotiation mode, the local and peer ports will automatically negotiate about the port speed.

Perform the following configuration in Ethernet port view.

Table 6-5 Set speed on the Ethernet port

Operation	Command
Set Ethernet port speed	speed { 10 100 1000 10000 auto }
Restore the default speed on Ethernet port	undo speed

Note that, the 10/100 Mbps electrical Ethernet port can operate at 10 Mbps, 100 Mbps and in auto mode. You can set it accordingly. The 10/100/1000Mbps electrical Ethernet port can operate at 10 Mbps, 100 Mbps, or 1000 Mbps as per different requirements. However in half duplex mode, the port cannot operate at 1000 Mbps or in auto mode. The 100 Mbps optical Ethernet port supports 100 Mbps; the 1000 Mbps optical Ethernet port supports 1000 Mbps; the 10 Gbps optical Ethernet port supports 10 Gbps without user intervention.

By default, the speed of the port is in **auto** mode.

6.2.6 Setting the Cable Type for the Ethernet Port

The Ethernet port supports the straight-through and cross-over network cables. The following command can be used for configuring the cable type.

Perform the following configuration in Ethernet port view.

Table 6-6 Set the type of the cable connected to the Ethernet port

Operation	Command
Set the type of the cable connected to the Ethernet port	mdi { across auto normal }

Operation	Command
Restore the default type of the cable connected to the Ethernet port	undo mdi

Note that, the settings only take effect on 10/100 Mbps and 10/100/1000 Mbps electrical ports.

By default, the cable type is **auto** (auto-recognized). That is, the system can automatically recognize the type of cable connecting to the port.

6.2.7 Enabling/Disabling Flow Control for the Ethernet Port

After enabling flow control in both the local and the peer switch, if congestion occurs in the local switch, the switch will inform its peer to pause packet sending. Once the peer switch receives this message, it will pause packet sending, and vice versa. In this way, packet loss is reduced effectively. The flow control function of the Ethernet port can be enabled or disabled through the following command.

Perform the following configuration in Ethernet port view.

Table 6-7 Enable/disable flow control for the Ethernet port

Operation	Command
Enable Ethernet port flow control	flow-control
Disable Ethernet port flow control	undo flow-control

By default, Ethernet port flow control is disabled.

6.2.8 Permitting/Forbidding Jumbo Frame to Pass the Ethernet Port

The Ethernet port may encounter the jumbo frame exceeding the standard frame length, when switching large throughput data like transmitting files. This command can forbid or permit the jumbo frame to pass the Ethernet port.

Perform the following configuration in Ethernet port view.

Table 6-8 Permit/forbid jumbo frame to pass the Ethernet port

Operation	Command
Permit jumbo frame to pass the Ethernet port	jumboframe enable [<i>jumboframe-value</i>]
Forbid jumbo frame to pass the Ethernet port	undo jumboframe enable

By default, the jumbo frame is permitted to pass the Ethernet port.

Note that, the values can be consecutive, but the effective values are discrete. The effective frame length for the FE port is 1552. The effective frame length for the GE port and 10 GE port is 1552, 9022, 9192 and 10240. You can execute the **display interface** command to view the configured effective value for the port.

6.2.9 Setting the Ethernet Port Broadcast Suppression Ratio

You can use the following commands to restrict the broadcast traffic. Once the broadcast traffic exceeds the value set by the user, the system will maintain an appropriate broadcast packet ratio by discarding the overflow traffic, so as to suppress broadcast storm, avoid suggestion and ensure the normal service. The parameter is taken the maximum wire speed ratio of the broadcast traffic allowed on the port. The smaller the ratio is, the smaller the broadcast traffic is allowed. If the ratio is 100%, it means not to perform broadcast storm suppression on the port.

Perform the following configuration in Ethernet port view.

Table 6-9 Set the Ethernet port broadcast suppression ratio

Operation	Command
Set Ethernet port broadcast suppression ratio	broadcast-suppression <i>pct</i>
Restore the default Ethernet port broadcast suppression ratio	undo broadcast-suppression

By default, 50% broadcast traffic is allowed to pass,

6.2.10 Setting the Ethernet Port Mode

Most ports adopt the LAN mode for general data exchange. The port must work in WAN mode, however, if it needs special frame format for data transfer (such as in fiber transmission). You can configure network mode available on the port using the **port-mode** command.

Perform the following configuration in Ethernet port view.

Table 6-10 Set the Ethernet port mode

Operation	Command
Set the Ethernet port mode	port-mode { wan lan }
Restore the default Ethernet port mode	undo port-mode

By default, the Ethernet port works in LAN mode.

Note that only GE or 10GE port supports this command.

6.2.11 Setting the Link Type for the Ethernet Port

Ethernet port can operate in three different link types, access, hybrid, and trunk types. The access port carries one VLAN only, used for connecting to the user's computer. The trunk port can belong to more than one VLAN and receive/send the packets on multiple VLANs, used for connection between the switches. The hybrid port can also carry more than one VLAN and receive/send the packets on multiple VLANs, used for connecting both the switches and user's computers. The difference between the hybrid port and the trunk port is that the hybrid port allows the packets from multiple VLANs to be sent without tags, but the trunk port only allows the packets from the default VLAN to be sent without tags.

Perform the following configuration in Ethernet port view.

Table 6-11 Set the link type for the Ethernet port

Operation	Command
Configure the port as access port	port link-type access
Configure the port as hybrid port	port link-type hybrid
Configure the port as trunk port	port link-type trunk
Restore the default link type, that is, the access port	undo port link-type

You can configure three types of ports concurrently on the same switch, but you cannot switch between trunk port and hybrid port. You must turn it first into access port and then set it as other type. For example, you cannot configure a trunk port directly as hybrid port, but first set it as access port and then as hybrid port.

By default, the port is access port.

6.2.12 Adding the Ethernet Port to Specified VLANs

The following commands are used for adding an Ethernet port to a specified VLAN. The access port can only be added to one VLAN, while the hybrid and trunk ports can be added to multiple VLANs.

Perform the following configuration in Ethernet port view.

Table 6-12 Add the Ethernet port to specified VLANs

Operation	Command
Add the current access port to a specified VLAN	port access vlan <i>vlan_id</i>
Add the current hybrid port to specified VLANs	port hybrid vlan <i>vlan_id_list</i> { tagged untagged }

Operation	Command
Add the current trunk port to specified VLANs	port trunk permit vlan { <i>vlan_id_list</i> all }
Remove the current access port from to a specified VLAN	undo port access vlan
Remove the current hybrid port from to specified VLANs	undo port hybrid vlan <i>vlan_id_list</i>
Remove the current trunk port from specified VLANs	undo port trunk permit vlan { <i>vlan_id_list</i> all }

Note that the access port shall be added to an existing VLAN other than VLAN 1. The VLAN to which Hybrid port is added must have been existed.

After adding the Ethernet port to specified VLANs, the local port can forward packets of these VLANs. The hybrid and trunk ports can be added to multiple VLANs, thereby implementing the VLAN intercommunication between peers. For the hybrid port, you can configure to tag some VLAN packets, based on which the packets can be processed differently.

6.2.13 Setting the Default VLAN ID for the Ethernet Port

Since the access port can only be included in one VLAN only, its default VLAN is the one to which it belongs. The hybrid port and the trunk port can be included in several VLANs, it is necessary to configure the default VLAN ID. If the default VLAN ID has been configured, the packets without VLAN Tag will be forwarded to the port that belongs to the default VLAN. When sending the packets with VLAN Tag, if the VLAN ID of the packet is identical to the default VLAN ID of the port, the system will remove VLAN Tag before sending this packet.

Perform the following configuration in Ethernet port view.

Table 6-13 Set the default VLAN ID for the Ethernet port

Operation	Command
Set the default VLAN ID for the hybrid port	port hybrid pvid vlan <i>vlan_id</i>
Set the default VLAN ID for the trunk port	port trunk pvid vlan <i>vlan_id</i>
Restore the default VLAN ID of the hybrid port to the default value	undo port hybrid pvid
Restore the default VLAN ID of the trunk port to the default value	undo port trunk pvid

Note that: to guarantee the proper packet transmission, the default VLAN ID of local hybrid port or Trunk port should be identical with that of the hybrid port or Trunk port on the peer switch.

By default, the VLAN of hybrid port and trunk port is VLAN 1 and that of the access port is the VLAN to which it belongs

6.2.14 Setting the VLAN VPN Feature

VLAN Tag consists of 12 bits (defined by IEEE802.1Q), so Ethernet Switches can support up to 4k VLANs. In networking, especially in MAN (metropolitan area network), a large numbers of VLANs are required to segment users. In this case, 4k VLANs are not enough.

VLAN VPN feature can provide duplex VLAN Tags to a packet, i.e. mark the packet with another VLAN Tag besides the original one, thus to provide 4k x 4k VLANs to meet user's demands. At the same time, VLAN VPN feature provides the following functions: using the original VLAN Tag to differentiate users and services, and using the new VLAN Tag to load service and VPN users. These make VLAN configuration simple and practicable. Thus, Ethernet Switches can meet the requirement of MAN.

If VLAN VPN is enabled on a port, all the packets (no matter whether it carries a VLAN Tag or not) will be given a new Tag that specifies the default VLAN of this port. Therefore, the packets that have had a VLAN Tag get two Tags, and the packets that have not had a VLAN Tag get one.

Perform the following configuration in Ethernet port view.

Table 6-14 Set the VLAN VPN feature

Operation	Command
Enable the VLAN VPN feature	vlan-vpn enable
Disable the VLAN VPN feature	undo vlan-vpn

Note that if anyone of GVRP, STP, and 802.1x has been enabled on a port, VLAN VPN cannot be enabled on it.

By default, the port VLAN VPN is disabled.

6.2.15 Copying Port Configuration to Other Ports

To keep the configuration of other ports consistent with a specified port, you can use **copy configuration** command to copy the configuration of that specified port to other ports. Such configurations may involve: STP setting, QoS setting, LACP setting, and port setting. The detailed table is as follows:

Table 6-15 Configurations that can be copied

Attribute	Detailed Setting
STP setting	Enable/disable STP
	Port priority

	Path cost
	Link attributes(point-to-point or not)
	Port mCheck
	Max transmission speed
	Enable/disable root protection
	Enable/disable loop protection
	Edge or non-edge port
	Reset ARP or not
QoS setting	Define/apply flow template
	Traffic reshaping
	Traffic redirection
	Packet filtering
	Priority re-assignment
	Traffic statistics
	Traffic mirroring
	Rate limiting
Port setting	Permitted VLAN ID
	Default VLAN ID
	Add ports to VLAN
	Default 802.1p priority
	Port speed, duplex mode
	Port link type
LACP	Enable/disable LACP on the port

Note:

- Using copy configuration command will clear protocol VLAN attributes of the destination port, but it can not copy protocol VLAN attributes of source port to the destination port.
 - Using the **copy configuration** command, you can only copy the configurations of Ethernet ports, GigabitEthernet ports and aggregation groups.
-

Perform the following configuration in system VLAN

Table 6-16 Copy port configuration to other ports

Operation	Command
Copy port configuration to other ports	copy configuration source { <i>interface-type interface-number</i> <i>interface-name</i> aggregation-group <i>agg-id</i> } destination { <i>interface_list</i> [aggregation-group <i>agg-id</i>] aggregation-group <i>agg-id</i> }

Note that if the copy source is an aggregation group, take the port with minimum Active as the source; if the copy destination is an aggregation group, make the configurations of all group member ports identical with that of the source. You cannot specify a dynamic aggregation group as the copy destination.

6.2.16 Setting Port Hold Time

When you use the **shutdown/undo shutdown** command on ports too frequently, the switch may fail. Therefore, you can configure port hold time to prohibit frequent change of the port status.

Perform the following configuration in system view.

Table 6-17 Set the port hold time

Operation	Command
Set the port hold time	link-status hold <i>hold-time</i>
Restore the default value	undo link-status hold

By default, the port hold time is set to 3 seconds.

6.2.17 Setting the Ethernet Port in Loopback Mode

Perform the following configuration in Ethernet port view.

Table 6-18 Set the Ethernet port in loopback mode

Operation	Command
Set the Ethernet port in loopback mode	loopback { external internal }
Remove loopback configuration on the port	undo loopback

By default, the Ethernet port is set in loopback mode. At present, the Switch 8800 does not support external loopback mode.

6.3 Displaying and Debugging Ethernet Port

After the above configuration, execute **display** command in any view to display the running of the Ethernet port configuration, and to verify the effect of the configuration.

Execute **reset** command in user view to clear the statistics information of the port.

Table 6-19 Display and debug Ethernet port

Operation	Command
Display all the information of the port	display interface { <i>interface_type</i> <i>interface_type interface_num</i> [<i>packets</i>] <i>interface_name</i> }
Display hybrid port or trunk port	display port { hybrid trunk }
Display the information of VLAN VPN	display port vlan-vpn
Display the statistics information of the port	display counters [<i>rate</i>] { inbound outbound } interface [<i>interface-type</i>]
Clear the statistics information of the port	reset counters interface [<i>interface_type</i> <i>interface_type interface_num</i> <i>interface_name</i>]

Note:

- The Switch 8800 does not support external loopback mode.
 - When 802.1x is enabled on the port, its statistics information can not be cleared.
 - By default, the **display counters** command displays the statistic information of all the ports.
-

6.4 Ethernet Port Configuration Example

I. Network requirements

Switch A is connected to Switch B through Trunk port GigabitEthernet2/1/1. Configure the Trunk port with default VLAN ID, so that: when receiving the packets without VLAN Tag, the port can forward them to the member ports belonging to the default VLAN; when it sending the packets with VLAN Tag and the packet VLAN ID is the default VLAN ID, the Trunk port remove the packet VLAN Tag and forward the packet.

II. Network diagram



Figure 6-1 Network diagram for Ethernet port configuration

III. Configuration procedure

The following configurations are used for Switch A. Please configure Switch B in the similar way.

Enter the Ethernet port view of GigabitEthernet2/1/1.

```
[SW8800] interface gigabitethernet2/1/1
```

Set the GigabitEthernet2/1/1 as a trunk port and allows VLANs 2, 6 through 50, and 100 to pass.

```
[SW8800-GigabitEthernet2/1/1] port link-type trunk
```

```
[SW8800-GigabitEthernet2/1/1] port trunk permit vlan 2 6 to 50 100
```

Create the VLAN 100.

```
[SW8800] vlan 100
```

Configure the default VLAN ID of GigabitEthernet2/1/1 as 100.

```
[SW8800-GigabitEthernet2/1/1] port trunk pvid vlan 100
```

6.5 Ethernet Port Troubleshooting

Symptom 1: Default VLAN ID configuration fails.

Solution: Take the following steps:

- Execute the **display interface** or **display port** command to check if the port is a trunk port or a hybrid port. If it is neither of them, configure it as a trunk or hybrid port.
- Then configure the default VLAN ID.

Symptom 2: The port is in down status.

Solution: Please check

- If the cable connection is correct and if the optical fiber cable is inversely connected.
- If the **shutdown** command is used on the port.
- If the right optical module is inserted.

Chapter 7 Link Aggregation Configuration

7.1 Overview

7.1.1 Introduction to Link Aggregation

Link aggregation means aggregating several ports together to implement the outgoing/incoming payload balance among the member ports and enhance the connection reliability. Link aggregation may be manual aggregation, dynamic LACP aggregation or static LACP aggregation. For the member ports in an aggregation group, their basic configurations must be the same. That is, if one is a trunk port, others must also be; when it turns into access port, then others must change to access port.

Basic configuration includes STP setting, QoS setting, VLAN setting, and port setting. The STP setting includes STP enabling/disabling, link attribute (point-to-point or not), STP priority, path cost, max transmission speed, loop protection, root protection, edge port or not. The QoS setting includes traffic limiting, priority marking, default 802.1p priority, bandwidth assurance, congestion avoidance, traffic redirection, traffic statistics. The VLAN setting includes permitted VLAN types, default VLAN ID. The port setting includes port link type.

One Switch 8800 can support up to 728 aggregation groups (seven load sharing aggregation groups at most), with each group containing a maximum of eight ports.

Note:

The Switch 8800 also supports trans-board aggregation. The trans-board aggregation is the same as the intra-board aggregation.

7.1.2 Introduction to LACP

Link aggregation control protocol (LACP) based on the IEEE802.3ad standard can be used in dynamic link aggregation. An LACP-enabled port sends link aggregation control protocol data units (LACPDUs) to tell the peer about its system priority, system MAC address, port priority, port number and operation key. After receiving the information from the sender, the receiver compares it with the locally saved information about other ports, chooses member ports for the aggregation group and reaches agreement about if a port can join or leave a dynamic aggregation group.

During port aggregation, LACP generates a configuration mix according to the port configuration (rate, duplex, basic configuration, management key), which is called an

operation key. The management key of an LACP-enabled dynamic aggregation port is 0 by default. The management key of an LACP-enabled static aggregation port is the same as the aggregation group ID. In a dynamic aggregation group, the member ports must have the same operation key. In manual and static aggregation groups, the active ports have the same operation key.

7.1.3 Aggregation Types

Port aggregation can be divided into manual aggregation, dynamic LACP aggregation and static LACP aggregation.

I. Manual aggregation and static LACP aggregation

Both manual aggregation and static LACP aggregation are configured manually, and cannot be added or removed automatically by the system. A manual or static LACP aggregation group must contain a member port at least. In the case of one port in an aggregation group, the unique method for you to remove the port from the aggregation group is to delete the aggregation group. By default, the system disables the LACP for the manual aggregation port. You are prohibited to enable the LACP for the manual aggregation port. By default, the system enables the LACP for the static aggregation port. When a static aggregation group is removed, the member ports will form one or more dynamic LACP aggregation groups with LACP enabled. You are prohibited to disable the LACP for the static aggregation port.

In the manual and static aggregation groups, a port maybe in active or inactive state. The port in active state can tranceive user service packets, but the port in inactive state cannot. The active port with the minimum port number serves as the master port, while others as slave ports.

In a manual aggregation group, the system sets the ports to active or inactive state based on these rules:

- Based on the descending order of priority levels from full duplex/high speed, to full duplex/low-speed, to half duplex/high speed and till half duplex/low speed, the system sets the port with the highest priority to active state, and others to inactive state.
- The system sets to inactive state the ports which cannot aggregate with the master port, due to hardware limit (such as trans-board aggregation is forbidden).
- The system sets to inactive state the ports with basic configurations different from the active port.

In a static aggregation group, the system sets the ports to active or inactive state based on these rules:

- Based on the descending order of priority levels from full duplex/high speed, to full duplex/low-speed, to half duplex/high speed and till half duplex/low speed, the system sets the port with the highest priority to active state, and others to inactive state.

- The system sets to inactive state the active port connecting to the different peer devices, or the port connecting to the same peer device but locating in the different aggregation group.
- The system sets to inactive state the ports which cannot be aggregated with the port, due to hardware limit (for example, trans-board aggregation is forbidden).
- The system sets to inactive state the ports with basic configurations different from the active port.

Since only a defined number of ports can be added in an aggregation group, then if the active ports in an aggregation group exceed the maximum threshold for that group, the system shall set some ports with smaller port numbers (in ascending order) as active ports and others as inactive ports. Both active and inactive ports can transceive LACP protocol, but the inactive ports cannot forward user service packets.

II. Dynamic LACP aggregation

The system can create/delete automatically dynamic LACP aggregations, and you cannot add/delete member ports into/from dynamic LACP aggregation. The system can also aggregate one port, which is called single port aggregation. The dynamic LACP aggregation LACP is in enabled state. The system can only aggregate the ports with the same speed, duplex attribute, device connection, basic configuration.

Since only a defined number of ports can be added in an aggregation group, then if the current member ports in an aggregation group exceed the maximum threshold for that group, the system shall set some ports with smaller device ID (system priority + system MAC address) and smaller port ID (port priority + port number) as active ports, and others as inactive ports. If the maximum threshold is not exceeded, all member ports are active ports. Both active and inactive ports can transceive LACP protocol, but the inactive ports cannot forward user service packets. In an aggregation group, the active port with the minimum port number serves as the master port, while others as slave ports. When comparing device ID, the system compare system priority first, and then system MAC address in the case of the same system priority. The smaller device ID is regarded as higher priority. When comparing port ID, the system compares port priority first, and then port number in the case of the same port priority. The smaller port ID is regarded as higher priority. If the device ID changes to higher priority, the active and inactive state of the member ports in an aggregation group depends on the device port ID. You can also set system and port priority to define active and inactive ports.

7.1.4 Load Sharing

I. Types of Load sharing

In terms of load balancing, link aggregation may be load balancing aggregation and non-load balancing aggregation. The Switch 8800 allocates IP packet load sharing according to destination and source IP addresses. The switches allocate non-IP packet load sharing according to source and destination MAC addresses. You can check

protocol types in determining if to use IP or MAC addresses. The packet with 0800 ETYPE Ethernet field is IP packet. In general, the system only provides limited resources. The system will always allocate hardware aggregation resources to the load balancing aggregation groups with higher priority levels. When the load sharing aggregation resources are used up for existing aggregation groups, newly-created aggregation groups will be non-load sharing ones. The priority levels (in descending order) for allocating load sharing aggregation resources are as follows:

- Aggregation groups of special ports with hardware aggregation resources included
- Aggregation groups including special ports which require hardware aggregation resources
- Aggregation groups that probably reach the maximum rate after the resources are allocated to them
- Aggregation groups with the minimum master port numbers if they reach the equal rate with other groups after the resources are allocated to them

When aggregation groups of higher priority levels appear, the aggregation groups of lower priority levels release their hardware resources. For single-port aggregation groups, if they can transceive packets normally without occupying hardware resources, they shall not occupy the resources.

II. Port state

In a aggregation group, its ports may be in active or inactive state and only the active ports can transceive user service packets, but not inactive ports. The active port with the minimum port number serves as the master port, while others as slave ports.

In a aggregation group, the system sets the ports to active or inactive state based on these rules:

- Based on the descending order of priority levels from full duplex/high speed, to full duplex/low-speed, to half duplex/high speed and till half duplex/low speed, the system sets the port with the highest priority to active state, and others to inactive state.
- The system sets to inactive state the ports which cannot aggregate with the master port, due to hardware limit.
- The system sets to inactive state the ports with basic configurations different from the master port.

Since only a defined number of ports can be supported in an aggregation group, then if the active ports in an aggregation group exceed the port quantity threshold for that group, the system shall set some ports with smaller port numbers (in ascending order) as active ports and others as inactive ports. The active ports can transceive user service packets, but not inactive ports.

A load sharing aggregation group may contain several active ports, but a non-load sharing aggregation group can only have one active port, while others as inactive ports.

7.2 Link Aggregation Configuration

The following sections describe link aggregation tasks:

- Enabling/Disabling LACP at Port
- Creating/Deleting an Aggregation Group
- Adding/Deleting an Ethernet Port into/from an Aggregation Group
- Setting/Deleting Aggregation Group Description
- Configuring System Priority
- Configuring Port Priority

Note:

- When configuring an aggregation group, the status of GVRP feature configured on the master port is reserved, but that on the slave port is disabled.
 - When adding a port to an existing aggregation group, the GVRP feature on the port is disabled.
 - When the master port leaves an aggregation group, the status of GVRP feature on both the group and port is reserved; when a slave port leaves an aggregation group, the GVRP feature on the port is disabled.
 - When configuring GVRP feature on any port in an aggregation group, the configuration is mapped to the master port of the group.
 - When querying the GVRP feature configured on any port in an aggregation group, the returned result is about the master port of the group.
-

For details, refer to the “VLAN” part of this manual

7.2.1 Enabling/Disabling LACP at Port

You should first enable LACP at the ports before performing dynamic aggregation, so that both parties can agree on adding/deleting the ports into/from a dynamic LACP aggregation group.

Perform the following configuration in Ethernet port view.

Table 7-1 Enable/disable LACP at port

Operation	Command
Enable LACP at the port	lACP enable
Disable LACP at the port	undo lACP enable

By default, LACP is not enabled at the port.

Note that:

- You cannot enable LACP at the mirroring port, the port with static MAC address configured, and the port with static ARP configured, port with 802.1x enabled.
- You are inhibited to enable LACP at the port in a manual aggregation group.
- You can add a port with LACP enabled into a manual aggregation group, but then the LACP will be disabled on it automatically. Or you can add a port with LACP disabled into a static LACP aggregation group, and then the LACP will be enabled automatically.

7.2.2 Creating/Deleting an Aggregation Group

You can use the following command to create/delete a aggregation group. When you delete a aggregation group, all its member ports are disaggregated.

Perform the following configuration in system view.

Table 7-2 Create/delete an aggregation group

Operation	Command
Create an aggregation group	link-aggregation group <i>agg-id</i> mode { manual static }
Delete an aggregation group	undo link-aggregation group <i>agg-id</i>

During creating an aggregation group, if it already exists in the system but contains no member port, it changes to the new type; if it already exists in the system and contains member ports, then you can only change a dynamic or static LACP aggregation group to a manual one, or a dynamic LACP aggregation group to a static one. In the former case, LACP shall be disabled at the member ports automatically, while in the latter case, LACP shall remain enabled.

Note:

There are three types of link aggregation: manual aggregation, static aggregation and dynamic aggregation.

7.2.3 Adding/Deleting an Ethernet Port into/from an Aggregation Group

You can add/delete ports into/from an aggregation group.

Perform the following configuration in corresponding view.

Table 7-3 Add/delete an Ethernet port into/from an aggregation group

Operation	Command
Add an Ethernet port into the aggregation group (Ethernet port view)	port link-aggregation group <i>agg-id</i>
Delete an Ethernet port from the aggregation port (Ethernet port view)	undo port link-aggregation group
Aggregate Ethernet ports (system view)	link-aggregation <i>interface_name1</i> to <i>interface_name2</i> [both]

Note that:

- You cannot add the mirrored port, port with static MAC address configured, port with static ARP configured, port with 802.1x enabled, and VPN port into an aggregation group.
- You must delete the aggregation group, instead of the port, if the aggregation group contains only one port.
- When master port enables VLAN VPN, aggregation is permitted in the system. Because the link type of slave port will always keep same as that of master port. When master port and slave port disable VLAN VPN, aggregation is permitted in the system, it is average aggregation. After the port enabling VLAN VPN, aggregation is not permitted in the system, at the same time, the system will tell users that the slave port in the aggregation group conflict with the master port on VLAN VPN.

7.2.4 Setting/Deleting Aggregation Group Description

Perform the following configuration in system view.

Table 7-4 Set/delete aggregation group description

Operation	Command
Set an aggregation group description	link-aggregation group <i>agg-id</i> description <i>aname</i>
Delete the aggregation group description	undo link-aggregation group <i>agg-id</i> description

By default, an aggregation group has no description.

Note:

If you save the current configuration using the **save** command, the static and dynamic LACP aggregation groups and their description strings remains on the system after rebooting, but not the dynamic LACP aggregation groups, or their description strings.

7.2.5 Configuring System Priority

The LACP refers to system IDs to determine if the member ports are active or inactive for a dynamic LACP aggregation group. The system ID consists of two-byte system priority and six-byte system MAC (system ID = system priority + system MAC). In comparing system IDs, the system first compares system priority values; if they are equal, then it compares system MAC addresses. The smaller system ID is considered prior. Changing system priority may affect the priority levels of member ports, and further their active or inactive state.

Perform the following configuration in system view.

Table 7-5 Configure system priority

Operation	Command
Configure system priority	lACP system-priority <i>system-priority-value</i>
Restore the default system priority	undo lACP system-priority

By default, system priority is 32,768.

7.2.6 Configuring Port Priority

The LACP compares system IDs first and then port IDs (if system IDs are the same) to determine if the member ports are active or inactive for a dynamic LACP aggregation group. If the ports in an aggregation group exceed the port quantity threshold for that group, the system sets some ports with smaller port IDs as active ports and others as inactive ports. The port ID consists of two-byte port priority and two-byte port number, that is, port ID = port priority + port number. The system first compares port priority values and then port numbers and the small port ID is considered prior.

Perform the following configuration in Ethernet port view.

Table 7-6 Configure port priority

Operation	Command
Configure port priority	lACP port-priority <i>port-priority-value</i>
Restore the default port priority	undo lACP port-priority

By default, port priority is 32,768.

7.3 Displaying and Debugging Link Aggregation

After the above configuration, execute the **display** command in any view to display the running of the link aggregation configuration, and to verify the effect of the configuration.

In user view, execute the reset command to clear statistics on the LACP-enabled port, and the **debugging** command to enable LACP debugging.

Table 7-7 Display and debug link aggregation

Operation	Command
Display summary information of all aggregation groups	display link-aggregation summary
Display detailed information of a specific aggregation group	display link-aggregation verbose <i>agg-id</i>
Display the local device ID	display lacp system-id
Display detailed link aggregation information at the port	display link-aggregation interface { <i>interface-type interface-number</i> <i>interface-name</i> } [to { <i>interface-type interface-num</i> <i>interface-name</i> }]
Clear LACP statistics on the port	reset lacp statistics [interface { <i>interface-type interface-number</i> <i>interface-name</i> } [to { <i>interface-type interface-num</i> <i>interface-name</i> }]]
Disable/enable LACP state debugging	[undo] debugging lacp state [interface { <i>interface-type interface-number</i> <i>interface-name</i> } [to { <i>interface-type interface-num</i> <i>interface-name</i> }]] [{ actor-churn mux partner-churn ptx rx }* all]
Disable/enable LACP packet debugging	[undo] debugging lacp packet [interface { <i>interface-type interface-number</i> <i>interface-name</i> } [to { <i>interface-type interface-num</i> <i>interface-name</i> }]]
Disable/enable link aggregation error debugging	[undo] debugging link-aggregation error
Disable/enable link aggregation event debugging	[undo] debugging link-aggregation event

7.4 Link Aggregation Configuration Example

I. Network requirements

Switch A connects switch B with three aggregation ports, numbered as Ethernet2/1/1 to Ethernet2/1/3, so that incoming/outgoing load can be balanced among the member ports.

II. Network diagram

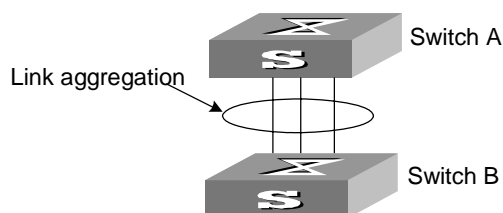


Figure 7-1 Network diagram for link aggregation configuration

III. Configuration procedure

The following only lists the configuration for switch A, and that on switch B is similar.

1) In manual aggregation mode

Create aggregation group 1.

```
[SW8800] link-aggregation group 1 mode manual
```

Add Ethernet ports Ethernet2/1/1 to Ethernet2/1/3 into aggregation group 1.

```
[SW8800] interface ethernet2/1/1
[SW8800-Ethernet2/1/1] port link-aggregation group 1
[SW8800-Ethernet2/1/1] interface ethernet2/1/2
[SW8800-Ethernet2/1/2] port link-aggregation group 1
[SW8800-Ethernet2/1/2] interface ethernet2/1/3
[SW8800-Ethernet2/1/3] port link-aggregation group 1
```

2) In static LACP aggregation mode

Create aggregation group 1.

```
[SW8800] link-aggregation group 1 mode static
```

Add Ethernet ports Ethernet2/1/1 to Ethernet2/1/3 into aggregation group 1.

```
[SW8800] interface ethernet2/1/1
[SW8800-Ethernet2/1/1] port link-aggregation group 1
[SW8800-Ethernet2/1/1] interface ethernet2/1/2
[SW8800-Ethernet2/1/2] port link-aggregation group 1
[SW8800-Ethernet2/1/2] interface ethernet2/1/3
[SW8800-Ethernet2/1/3] port link-aggregation group 1
```

3) In dynamic LACP aggregation mode

Enable LACP on Ethernet ports Ethernet2/1/1 to Ethernet2/1/3.

```
[SW8800] interface ethernet2/1/1
[SW8800-Ethernet2/1/1] lacp enable
[SW8800-Ethernet1/1/1] interface ethernet2/1/2
[SW8800-Ethernet2/1/2] lacp enable
[SW8800-Ethernet2/1/2] interface ethernet2/1/3
[SW8800-Ethernet2/1/3] lacp enable
```

You must set basic configuration, rate and duplex attribute consistent at both ends to aggregate successfully the LACP-enabled ports into a dynamic aggregation group and achieve load sharing.

Chapter 8 VLAN Configuration

8.1 VLAN Overview

Virtual local area network (VLAN) groups the devices in a LAN logically, not physically, into segments to form virtual workgroups. IEEE issued the IEEE 802.1Q in 1999 to standardize the VLAN implementations.

The VLAN technology allows network administrators to logically divide a physical LAN into different broadcast domains or the so-called virtual LANs. Every VLAN contains a group of workstations with the same demands. The workstations, physically separated, are not necessarily on the same physical LAN segment.

You can establish VLANs of the following types on switches:

- Port-based
- MAC address-based
- IP multicast-based (A multicast group can be a VLAN.)
- Network layer-based (A VLAN can be established by the network layer addresses or protocols of the hosts.)

With the VLAN technology, the broadcast and unicast traffic within a VLAN will not be forwarded to other VLANs. This is helpful to control network traffic, save device investment, simplify network management and enhance security.

8.2 Configuring VLAN

The following sections describe VLAN configuration tasks:

- Creating/Deleting a VLAN
- Specifying a Description Character String for a VLAN or VLAN interface
- Creating/Removing a VLAN Interface
- Shutting down/Bringing up a VLAN Interface

8.2.1 Creating/Deleting a VLAN

You can use the following commands to create/delete a VLAN. If the VLAN to be created exists, the system enters the VLAN view directly. Otherwise, the system creates the VLAN first, and then enters the VLAN view.

Perform the following configuration in system view.

Table 8-1 Create/Delete a VLAN

Operation	Command
Create a VLAN and enter the VLAN view	vlan <i>vlan_id</i>
Delete a specified VLAN	undo vlan { <i>vlan_id</i> [<i>to</i> <i>vlan_id</i>] all }

Note that the default VLAN, namely VLAN 1, cannot be deleted.

8.2.2 Specifying a Description Character String for a VLAN or VLAN interface

You can use the following commands to specify a description character string for a VLAN or VLAN interface.

Perform the following configuration in VLAN view or VLAN interface view.

Table 8-2 Specify a description character string for a VLAN or VLAN interface

Operation	Command
Specify a description character string for a VLAN or VLAN interface	description <i>string</i>
Restore the default description of the current VLAN or VLAN interface	undo description

By default, the description character string of a VLAN is the VLAN ID of the VLAN, such as VLAN 0001. The description character string of a VLAN interface is the VLAN interface name, such as Vlan-interface1 Interface.

8.2.3 Creating/Removing a VLAN Interface

You can use the following commands to create/remove a VLAN interface. To implement the network layer function on a VLAN interface, IP address and mask should be set to the VLAN interface. For corresponding configuration, refer to “Network protocol” part in this manual.

Perform the following configuration in system view.

Table 8-3 Create/remove a VLAN interface

Operation	Command
Create a new VLAN interface and enter the VLAN interface view	interface vlan-interface <i>vlan_id</i>
Remove a specified VLAN interface	undo interface vlan-interface <i>vlan_id</i>

Create a VLAN before creating a VLAN interface corresponding to the VLAN.

8.2.4 Shutting down/Bringing up a VLAN Interface

You can use the following commands to shut down/bring up a VLAN interface.

Perform the following configuration in VLAN interface view.

Table 8-4 Shut down/bring up a VLAN interface

Operation	Command
Shut down a VLAN interface	shutdown
Bring up a VLAN interface	undo shutdown

Shutting down or bringing up a VLAN interface has no effect on the UP/DOWN status of the Ethernet ports in this VLAN.

By default, when all the Ethernet ports in a VLAN are in the DOWN state, this VLAN interface is also DOWN. When there are one or more Ethernet ports in the UP state, this VLAN interface is also UP.

8.3 Configuring Port-Based VLAN

8.3.1 Adding Ethernet Ports to a VLAN

You can use the following commands to add the Ethernet ports to a VLAN.

Perform the following configuration in VLAN view.

Table 8-5 Add Ethernet ports to a VLAN

Operation	Command
Add Ethernet ports to a VLAN	port <i>interface_list</i>
Remove Ethernet ports from a VLAN	undo port <i>interface_list</i>

By default, the system adds all the ports to a default VLAN whose ID is 1.

Note that you can add/remove the trunk and Hybrid ports to/from a VLAN by the **port/undo port** commands in Ethernet port view, but not in VLAN view.

8.4 Configuring Protocol-Based VLAN

The following sections describe the protocol-based VLAN configuration tasks:

- Creating/Deleting a VLAN Protocol Type
- Associating/Dissociating a Port with/from a Protocol-Based VLAN

8.4.1 Creating/Deleting a VLAN Protocol Type

You can use the following commands to create/delete a VLAN protocol type.

Perform the following configuration in VLAN view.

Table 8-6 Create/Delete a VLAN protocol type

Operation	Command
Create a VLAN protocol type	protocol-vlan protocol { ip <i>ip_address</i> [<i>net_mask</i>] mode { ethernetii etype <i>etype_id</i> llc dsap <i>dsap_id</i> ssap <i>ssap_id</i> snap etype <i>etype_id</i> } }
Delete an existing VLAN protocol type	undo protocol-vlan protocol { <i>protocol_index</i> [to <i>protocol_end</i>] all }

8.4.2 Associating/Dissociating a Port with/from a Protocol-Based VLAN

Perform the following configuration in Ethernet port view.

Table 8-7 Associate/Dissociate a port with/from a protocol-based VLAN

Operation	Command
Associate a port with a protocol-based VLAN	port hybrid protocol-vlan vlan <i>vlan-id</i> { <i>vlan-protocol_list</i> all }
Remove a port from a protocol-based VLAN	undo port hybrid protocol-vlan vlan <i>vlan-id</i> { <i>vlan-protocol_list</i> all }

Note:

- The port to be associated with a protocol-based VLAN must be of Hybrid type and in this VLAN.
- The same protocol can be configured in the different VLANs, but cannot be configured repeatedly in the same VLAN.
- A port cannot be associated with different VLANs with the same protocols configured.
- You cannot delete a protocol-based VLAN that has ports associated with.
- You cannot delete a protocol-based VLAN on a port while the port is associated with the VLAN.

8.5 Displaying VLAN

After the above configuration, execute the **display** command in any view to display the running of the VLAN configuration, and to verify the configuration.

Table 8-8 Display VLAN

Operation	Command
Display the related information about the VLAN interface	display interface vlan-interface [<i>vlan_id</i>]
Display the related information about the VLAN	display vlan [<i>vlan_id to vlan_id</i> all static dynamic]
Display the protocol information and protocol index configured on the specified VLAN	display vlan-protocol vlan { <i>vlan_list</i> all }
Display the protocol information and protocol index configured on the specified port	display vlan-protocol interface { <i>interface_list</i> all }

8.6 VLAN Configuration Example

I. Network requirements

- Create VLAN2 and VLAN3.
- Add Ethernet3/1/1 and Ethernet4/1/1 to VLAN2.
- Add Ethernet3/1/2 and Ethernet4/1/2 to VLAN3.

II. Network diagram

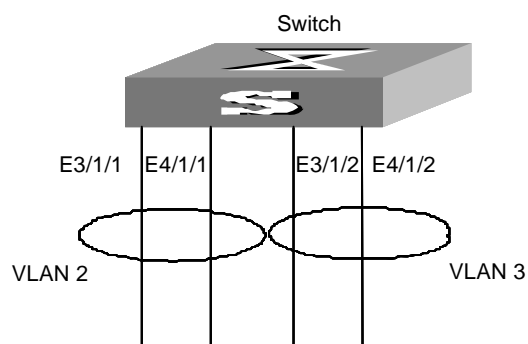


Figure 8-1 Network diagram for VLAN configuration

III. Configuration procedure

Create VLAN 2 and enter its view.

```
[SW8800] vlan 2
```

Add Ethernet3/1/1 and Ethernet4/1/1 to VLAN2.

```
[SW8800-vlan2] port ethernet3/1/1 ethernet4/1/1
```

Create VLAN 3 and enters its view.

```
[SW8800-vlan2] vlan 3
```

Add Ethernet3/1/2 and Ethernet4/1/2 to VLAN3.

```
[SW8800-vlan3] port ethernet3/1/2 ethernet4/1/2
```

Chapter 9 GARP/GVRP Configuration

9.1 Configuring GARP

9.1.1 GARP Overview

Generic attribute registration protocol (GARP) offers a mechanism that is used by the members in the same switching network to distribute, propagate and register such information as VLAN and multicast addresses.

GARP does not exist in a switch as an entity. A GARP participant is called GARP application. The main GARP applications at present are GVRP (GARP VLAN registration protocol) and GMRP. For details, refer to section 9.2 "Configuring GVRP" and section "Configuring Multicast". When a GARP participant is on a port of the switch, this port corresponds to a GARP participant.

The GARP mechanism enables the configuration information on one GARP member to be propagated rapidly across the whole switching network. A GARP member can be a terminal workstation or a bridge. The GARP member can notify other members to register or remove its attribute information by sending declarations or withdrawing declarations. It can also register or remove the attribute information of other GARP members according to the received declarations/withdrawn declarations.

GARP members exchange information by sending messages. There are mainly three types of GARP messages, Join, Leave, and LeaveAll. When a GARP participant wants to register its attribute information with other switches, it sends the Join message outward. When it wants to remove some attribute information from other switches, it sends the Leave message. The LeaveAll timer is started simultaneously when each GARP participant is enabled and the LeaveAll message is sent upon expiration. The Join and Leave messages cooperate to ensure the logout and the re-registration of a message. The message exchange enables all the to-be-registered attribute information to be propagated to all the switches across the same switching network.

The destination MAC addresses of the packets of the GARP participants are specific multicast MAC addresses. A GARP-supporting switch classifies the packets received from the GARP participants and processes them with corresponding GARP applications (GVRP or GMRP).

GARP and GMRP are described in details in the IEEE 802.1P standard (which has been added to the IEEE802.1D standard). Quidway series switches fully support the GARP compliant with the IEEE standards.

The following section describes the GARP configuration task:

- Setting the GARP Timer

Note:

- The value of GARP timer will be used in all the GARP applications, including GVRP and GMRP, running in one switched network.
- In one switched network, the GARP timers on all the switching devices should be set to the same value. Otherwise, GARP application cannot work normally.

9.1.2 Setting the GARP Timer

GARP timers include Hold timer, Join timer, Leave timer and LeaveAll timer.

The GARP participant sends the Join Message regularly when Join timer times out so that other GARP participants can register its attribute values.

When the GARP participant wants to remove some attribute values, it sends the Leave Message. The GARP participant that receives the message starts the Leave timer. If the Join Message is not received again before the Leave timer expires, the GARP attribute values are removed.

LeaveAll timer is started as soon as the GARP participant is enabled. The LeaveAll message is sent upon timeout so that other GARP participants remove all the attribute values of this participant. Then, LeaveAll timer is restarted and a new cycle begins.

When the switch receives some GARP registration information, it does not send the Join Message immediately. Instead, it enables a Hold timer and sends the Join Message upon timeout of the Hold timer. In this way, all the VLAN registration information received within the time specified by the Hold timer can be sent in one frame so as to save the bandwidth resources.

Configure Hold timer, Join timer and Leave timer in Ethernet port view. Configure LeaveAll timer in system view.

Table 9-1 Set the GARP timer

Operation	Command
Set GARP Hold timer, Join timer and Leave timer	garp timer { hold join leave } timer_value
Set GARP LeaveAll timer	garp timer leaveall timer_value
Restore the default settings of GARP Hold timer, Join timer and Leave timer	undo garp timer { hold join leave }
Restore the default settings of GARP LeaveAll timer	undo garp timer leaveall

By default, Hold timer is 10 centiseconds, Join timer is 20 centiseconds, Leave timer is 60 centiseconds, and LeaveAll timer is 1000 centiseconds.

Note that, the value of Join timer should be no less than the doubled value of Hold timer, and the value of Leave timer should be greater than the doubled value of Join timer and smaller than the Leaveall timer value. Besides, you must set the value of the Join timer in terms of 5 centiseconds. Otherwise, the system will prompt message of error.

The value range of a timer varies with the values of other timers. So if the value of a timer you want to set is not within the available value range, you can change the value range by changing the values of other related timers.

- The lower limit of Hold timer is 10 centiseconds. You can change its upper limit by changing the value of Join timer.
- You can change the lower limit and upper limit of Join timer by changing the value of Hold timer and Leave timer respectively.
- You can change the lower limit and upper limit of Leave timer by changing the value of Join timer and LeaveAll timer respectively.
- The upper limit of LeaveAll timer is 32765 centiseconds. You can change its lower limit by changing the value of Leave timer.

9.1.3 Displaying and Debugging GARP

After the above configuration, execute the **display** command in any view to display the running of GARP configuration, and to verify the configuration.

Execute the **reset** command in user view to reset the configuration of GARP. Execute the **debugging** command in user view to debug the configuration of GARP.

Table 9-2 Display and debug GARP

Operation	Command
Display GARP statistics information	display garp statistics [interface interface-list]
Display GARP timer	display garp timer [interface interface-list]
Clear GARP statistics information	reset garp statistics [interface interface-list]
Enable GARP event debugging	debugging garp event
Disable GARP event debugging	undo debugging garp event

9.2 Configuring GVRP

9.2.1 GVRP Overview

GARP VLAN Registration Protocol (GVRP) is a GARP application. Based on GARP operating mechanism, GVRP provides maintenance of the dynamic VLAN registration information in the switch and propagates the information to other switches. All the

GVRP-supporting switches can receive VLAN registration information from other switches and dynamically update the local VLAN registration information including the active members and through which port those members can be reached. All the GVRP-supporting switches can propagate their local VLAN registration information to other switches so that the VLAN information can be consistent on all GVRP-supporting devices in one switching network. The VLAN registration information propagated by GVRP includes both the local static registration information configured manually and the dynamic registration information from other switches.

GVRP is described in details in the IEEE 802.1Q standard. Quidway series switches fully support the GARP compliant with the IEEE standards.

Main GVRP configuration includes:

- Enabling/Disabling Global GVRP
- Enabling/Disabling Port GVRP
- Setting the GVRP Registration Type

In the above-mentioned configuration tasks, GVRP should be enabled globally before it is enabled on the port. Configuration of GVRP registration type can only take effect after the port GVRP is enabled. Besides, GVRP must be configured on the Trunk port.

Note:

- When you configure an aggregation group, the GVRP feature configured on the master port is unchanged, but that on the slave port is disabled.
 - When you add a port to an existing aggregation group, the GVRP feature on the port is disabled.
 - When the master port leaves an aggregation group, the GVRP feature on both the group and port is unchanged; when a slave port leaves an aggregation group, the GVRP feature on the port is disabled.
 - When you configure GVRP feature on any port in an aggregation group, the configuration is mapped to the master port of the group.
 - When you query the GVRP feature configured on any port in an aggregation group, the returned result is about the master port of the group.
-

9.2.2 Enabling/Disabling Global GVRP

You can use the following command to enable/disable global GVRP.

Perform the following configurations in system view.

Table 9-3 Enable/disable global GVRP

Operation	Command
Enable global GVRP	gvrp
Disable global GVRP	undo gvrp

By default, global GVRP is disabled.

9.2.3 Enabling/Disabling Port GVRP

You can use the following command to enable/disable the GVRP on a port.

Perform the following configurations in Ethernet port view.

Table 9-4 Enable/disable port GVRP

Operation	Command
Enable port GVRP	gvrp
Disable port GVRP	undo gvrp

GVRP should be enabled globally before it is enabled on the port. The GVRP can only be enabled/disabled on Trunk ports.

By default, port GVRP is disabled.

9.2.4 Setting the GVRP Registration Type

The GVRP registration types include **normal**, **fixed** and **forbidden** (refer to IEEE 802.1Q).

- When an Ethernet port is set to be in **normal** registration mode, the dynamic and manual creation, registration and deregistration of VLAN are allowed on this port.
- When a Trunk port is set as **fixed**, the port is not allowed to dynamically register/deregister a VLAN, it only propagates information about static VLANs that are manually configured instead of that of dynamic VLANs. That is, a Trunk port that is of fixed type only permits manually configured VLANs even you configure it to permit all VLANs.
- When an Ethernet port is set to be in **forbidden** registration mode, all the VLANs except VLAN1 will be deregistered and no other VLANs can be created and registered on this port.

Perform the following configuration in Ethernet port view.

Table 9-5 Set the GVRP registration type

Operation	Command
Set GVRP registration type	gvrp registration { normal fixed forbidden }
Restore the default GVRP registration type	undo gvrp registration

By default, GVRP registration type is **normal**.

9.2.5 Displaying and Debugging GVRP

After the above configuration, execute the **display** command in any view to display the running of GVRP configuration, and to verify the configuration.

Execute the **debugging** command in user view to debug the configuration of GVRP.

Table 9-6 Display and debug GVRP

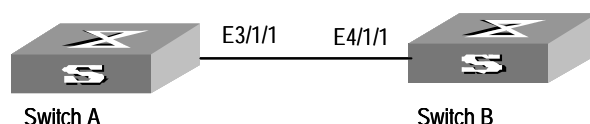
Operation	Command
Display GVRP statistics information	display gvrp statistics [interface interface-list]
Display GVRP global status information	display gvrp status
Enable GVRP packet or event debugging	debugging gvrp { packet event }
Disable GVRP packet or event debugging	undo debugging gvrp { packet event }

9.2.6 GVRP Configuration Example

I. Network requirements

To dynamically register and update VLAN information among switches, GVRP needs to be enabled on the switches.

II. Network diagram

**Figure 9-1** GVRP configuration example

III. Configuration procedure

Configure Switch A:

Enable GVRP globally.

```
[SW8800] gvrp
```

Set Ethernet3/1/1 as a Trunk port and allows all the VLANs to pass through.

```
[SW8800] interface ethernet3/1/1
```

```
[SW8800-Ethernet3/1/1] port link-type trunk
```

```
[SW8800-Ethernet3/1/1] port trunk permit vlan all
```

Enable GVRP on the Trunk port.

```
[SW8800-Ethernet3/1/1] gvrp
```

Configure Switch B:

Enable GVRP globally.

```
[SW8800] gvrp
```

Set Ethernet4/1/1 as a Trunk port and allows all the VLANs to pass through.

```
[SW8800] interface ethernet4/1/1
```

```
[SW8800-Ethernet4/1/1] port link-type trunk
```

```
[SW8800-Ethernet4/1/1] port trunk permit vlan all
```

Enable GVRP on the Trunk port.

```
[SW8800-Ethernet4/1/1] gvrp
```

Chapter 10 Super VLAN Configuration

10.1 Super VLAN Overview

Super VLAN is also called VLAN aggregation: A *super VLAN* contains multiple *sub VLANs*. A super VLAN can be configured with an IP address of the virtual port, while a sub VLAN cannot be configured with the IP address of the virtual port. Each sub VLAN is a broadcast domain. Different sub VLANs are isolated at Layer 2. When users in a sub VLAN need to communicate with each other, they use the IP address of the virtual interface of the super VLAN as the IP address of the gateway. The IP address is shared by multiple VLANs. Therefore IP addresses are saved. If different sub VLANs want to communicate with each other at Layer 3, or a sub VLAN communicates with other networks, you must enable ARP proxy (by default it is disabled). The address resolution protocol (ARP) proxy can forward and process ARP request and response packets so that the isolated sub VLANs can communicate with each other at Layer 3.

10.2 Configuring a Super VLAN

Super VLAN configuration includes:

- Configuring a Super VLAN

10.2.1 Configuring a Super VLAN

Note:

- You can configure multiple super VLANs for a switch. The configured VLAN port and IP address configurations are the same as common VLAN configurations.
 - A sub VLAN configuration is the same as a common VLAN configuration. The following table describes the specific commands to configure a sub VLAN. For detailed information, refer to Chapter 8 “VLAN Configuration”.
 - By default, ARP proxy is enabled for super VLANs and disabled on the sub VLANs.
-

You can configure a super VLAN as follows:

Table 10-1 Configure a super VLAN

Number	Item	Command	Description
1	Enter system view	<SW8800> system-view	—

Number	Item	Command	Description
2	Enter VLAN view	[SW8800] vlan <i>vlan_id</i>	Required
3	Set the VLAN type to super VLAN	[SW8800-vlan4093] supervlan	Required. The VLAN_ID is the configured VLAN ID in the range 1 to 4,094.
4	Create a sub VLAN and enter sub VLAN view	[SW8800] vlan <i>vlan_id</i>	Required
5	Add Ethernet ports to sub VLANs	[SW8800] port <i>interface_list</i>	Optional
6	Configure the mapping relation between super VLANs and sub VLANs	[SW8800-vlan4093] subvlan <i>sub-vlan-list</i>	Required. The view is the VLAN view of a super VLAN.
7	Display configuration information	<SW8800> display super vlan [<i>supervlan_id</i>]	Optional. You can execute the display super vlan command in any view.

To cancel the configurations, use the corresponding **undo** commands.



Caution:

- Super VLANs cannot contain ports.
 - After you set the VLAN type to super VLAN, the ARP proxy is automatically enabled on the VLAN port, and you do not need to configure the proxy.
 - When a super VLAN exists, the ARP proxy should be enabled on the corresponding VLAN port.
 - The default VLAN cannot be set to a super VLAN.
 - You can add multiple ports (non-uplink port) to each sub VLAN.
 - You cannot configure a virtual port for a sub VLAN.
 - If the **undo subvlan** command is not followed by *vlan_id*, the mapping relationship between all sub VLANs and specified super VLANs is removed; if the **undo subvlan** command is followed by *vlan_id*, the mapping relationship between the specified sub VLANs and specified super VLANs is removed.
-

10.2.2 Super VLAN Configuration Example

I. Network requirements

Super VLAN 10 and sub VLANs including VLAN 2, VLAN 3 and VLAN 5 need configuring. VLAN2 contains port 1 and 2; VLAN3 contains port 3 and 4; VLAN5 contains port 5 and 6. These sub VLANs are isolated at Layer 2. It is required that these sub VLANs communicate with each other at Layer 3.

II. Network diagram

Omitted

III. Configuration procedure

```
[SW8800] vlan 10
[SW8800-vlan10] supervlan
[SW8800-vlan10] vlan 2
[SW8800-vlan2] port ethernet3/1/1 ethernet3/1/2
[SW8800-vlan2] vlan 3
[SW8800-vlan3] port Ethernet3/1/3 ethernet3/1/4
[SW8800-vlan3] vlan 5
[SW8800-vlan5] port ethernet3/1/5 ethernet3/1/6
[SW8800-vlan5] vlan 10
[SW8800-vlan10] subvlan 2 3 5
[SW8800-vlan10] interface vlan 10
[SW8800-Vlan-interface10] ip address 10.110.1.1 255.255.255.0
```

Note:

By default ARP proxy on super VLANs is enabled, and disabled on sub VLANs.

Chapter 11 IP Address Configuration

11.1 Introduction to IP Address

11.1.1 IP Address Classification and Representation

An IP address is a 32-bit address allocated to a device that accesses the Internet. It consists of two fields: net-id field and host-id field. IP addresses are allocated by Network Information Center (NIC) of American Defense Data Network (DDN). To manage IP addresses conveniently, IP addresses are classified into five types. See the following figure.

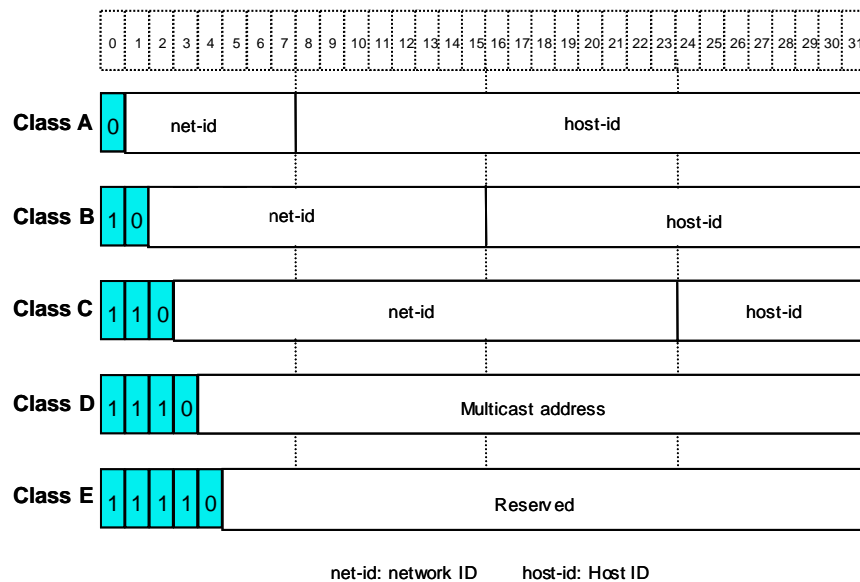


Figure 11-1 Five classes of IP address

Here, Class A, Class B and Class C addresses are unicast addresses, while Class D addresses are multicast ones and class E addresses are reserved for special applications in future. The first three types are commonly used.

The IP address is in dotted decimal format. Each IP address contains four integers in dotted decimal notation. Each integer corresponds to one byte, for example, 10.110.50.101.

When using IP addresses, note that some of them are reserved for special uses, and are seldom used. The IP addresses you can use are listed in the following table.

Table 11-1 IP address classes and ranges

Network class	Address range	IP network range available	Note
A	0.0.0.0 to 127.255.255.255	1.0.0.0 to 126.0.0.0	<p>Host ID with all the digits being 0 indicates that the IP address is the network address, and is used for network routing.</p> <p>Host ID with all the digits being 1 indicates the broadcast address, that is, broadcast to all hosts on the network.</p> <p>IP address 0.0.0.0 is used for the host that is not put into use after starting up.</p> <p>The IP address with network ID being 0 indicates the current network and its network can be cited by the router without knowing its network number.</p> <p>The IP addresses with the format of 127.X.Y.Z are reserved for self-loop test and the packets sent to these addresses are not output to the line. The packets are processed internally and regarded as input packets.</p>
B	128.0.0.0 to 191.255.255.255	128.0.0.0 to 191.254.0.0	<p>Host ID with all the digits being 0 indicates that the IP address is the network address, and is used for network routing.</p> <p>Host ID with all the digits being 1 indicates the broadcast address, that is, broadcast to all hosts on the network.</p>
C	192.0.0.0 to 223.255.255.255	192.0.0.0 to 223.255.254.0	<p>Host ID with all the digits being 0 indicates that the IP address is the network address, and is used for network routing.</p> <p>Host ID with all the digits being 1 indicates the broadcast address, that is, broadcast to all hosts on the network.</p>

Network class	Address range	IP network range available	Note
D	224.0.0.0 to 239.255.255.255	None	Addresses of class D are multicast addresses, among which: <ul style="list-style-type: none"> • IP address 224.0.0.0 is reserved and will not be allocated. Those from 224.0.0.1 to 224.0.0.255 are reserved for routing protocols and other protocols that are used to discover and maintain routes. • Those from 239.0.0.0 to 239.255.255.255 are used for local multicast management. • Those from 224.0.0.255 to 238.255.255.255 are for users.
E	240.0.0.0 to 255.255.255.254	None	The addresses are reserved for future use.
Other addresses	255.255.255.255	255.255.255.255	255.255.255.255 is used as a Local Area Network (LAN) broadcast address.

11.1.2 Subnet and Mask

Nowadays, with rapid development of the Internet, IP addresses are depleting very fast. The traditional IP address allocation method wastes IP addresses greatly. In order to make full use of the available IP addresses, the concept of mask and subnet is proposed.

A mask is a 32-bit number corresponding to an IP address. The number consists of 1s and 0s. Principally, these 1s and 0s can be combined randomly. However, the first consecutive bits are set to 1s when you design a mask. The mask divides the IP address into two parts: subnet address and host address. The part of IP address that corresponds to the bits 1s in the mask indicates the subnet address and the other part of IP address indicate the host address. If there is no subnet division, then its subnet mask is the default value and the length of "1" indicates the net-id length. Therefore, for IP addresses of classes A, B and C, the default values of corresponding subnet mask are 255.0.0.0, 255.255.0.0 and 255.255.255.0 respectively.

The mask can be used to divide a Class A network containing more than 16,000,000 hosts or a Class B network containing more than 60,000 hosts into multiple small networks. Each small network is called a subnet. For example, for the Class B network address 138.38.0.0, the mask 255.255.224.0 can be used to divide the network into eight subnets: 138.38.0.0, 138.38.32.0, 138.38.64.0, 138.38.96.0, 138.38.128.0,

138.38.160.0, 138.38.192.0 and 138.38.224.0 (Refer to the following figure). Each subnet can contain more than 8000 hosts.

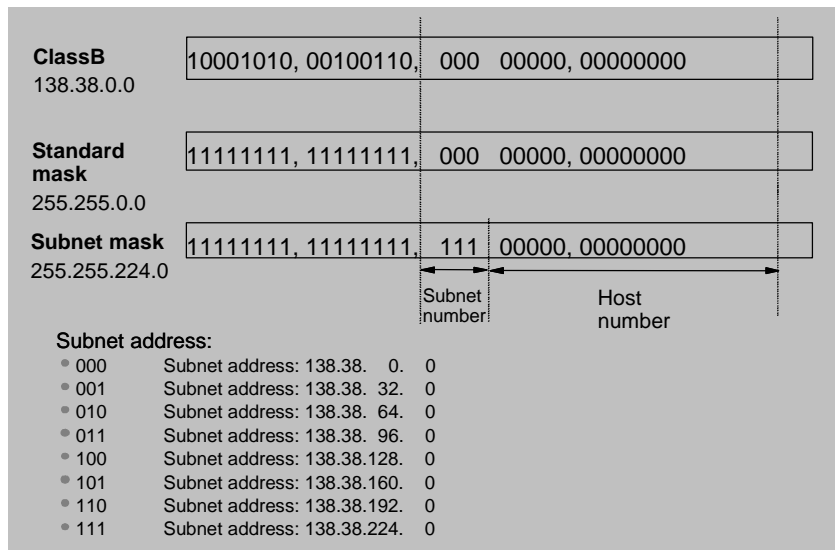


Figure 11-2 Subnet division of an IP address

11.2 Configuring IP Address

The following sections describe IP address configuration tasks:

- Configuring the Hostname and Host IP Address
- Configuring the IP Address of the VLAN Interface

11.2.1 Configuring the Hostname and Host IP Address

Using this command, you can associate a host name with an IP address. After that, when using an application like telnet, you can use the host name instead of the IP address that is hard to memorize, and the system automatically translates the host name to the IP address.

Perform the following configuration in system view.

Table 11-2 Configure the host name and the corresponding IP address

Operation	Command
Configure the host name and the corresponding IP address	ip host <i>hostname ip-address</i>
Cancel the host name and the corresponding IP address	undo ip host <i>hostname [ip-address]</i>

By default, there is no host name associated to any host IP address.

11.2.2 Configuring the IP Address of the VLAN Interface

You can configure an IP address for every VLAN interface of the switch. Generally, it is enough to configure one IP address for an interface. You can also configure ten IP addresses for an interface at most, so that it can be connected to several subnets. Among these IP addresses, one is the primary IP address and all others are secondary. Perform the following configuration in VLAN interface view.

Table 11-3 Configure an IP address for a VLAN interface

Operation	Command
Configure an IP address for a VLAN interface	ip address <i>ip-address</i> { <i>mask</i> <i>mask-length</i> } [sub]
Delete an IP address of a VLAN interface	undo ip address <i>ip-address</i> { <i>mask</i> <i>mask-length</i> } [sub]

By default, the IP address of a VLAN interface is null.

11.3 Displaying and debugging IP Address

After the above configuration, execute the **display** command in any view to display the IP addresses configured on interfaces of the network device, and to verify the effect of the configuration.

Table 11-4 Display and debug IP address

Operation	Command
Display all hosts on the network and the corresponding IP addresses	display ip host
Display the configurations of a VLAN interface	display ip interface <i>vlan-interface</i> <i>vlan-id</i>

11.4 IP Address Configuration Example

I. Network requirements

Configure the IP address as 129.2.2.1 and subnet mask as 255.255.255.0 for the VLAN interface 1 of the switch.

II. Network diagram

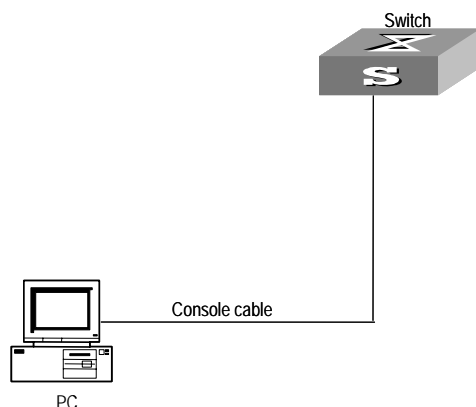


Figure 11-3 Network diagram for IP address configuration

III. Configuration procedure

Enter VLAN interface 1.

```
[SW8800] interface vlan-interface 1
```

Configure the IP address for VLAN interface 1.

```
[SW8800-Vlan-interface1] ip address 129.2.2.1 255.255.255.0
```

11.5 Troubleshooting IP Address Configuration

Fault 1: The switch cannot ping through a certain host in the LAN.

Troubleshooting can be performed as follows:

- 1) Check the configuration of the switch. Use the **display arp** command to view the ARP entry table that the switch maintains.
- 2) Check which VLAN includes the port of the switch used to connect to the host. Check whether the VLAN has been configured with a VLAN interface. Then check whether the IP address of the VLAN interface and that of the host are on the same network segment.
- 3) If the configuration is correct, enable the ARP debugging on the switch, and check whether the switch can correctly send and receive ARP packets. If it can only send ARP packets but cannot receive them, errors may occur on the Ethernet physical layer.

Chapter 12 ARP Configuration

12.1 Introduction to ARP

Address Resolution Protocol (ARP) is used to resolve an IP address into a MAC address.

I. Necessity of ARP

An IP address cannot be directly used for communication between network devices because network devices can only identify MAC addresses. An IP address is only an address of a host in the network layer. To send the data packets transmitted through the network layer to the destination host, MAC address of the host is required. So the IP address must be resolved into a MAC address.

II. ARP implementation procedure

When two hosts on the Ethernet need to communicate with each other, they must know the MAC addresses of each other. Every host maintains the IP-MAC address translation table, which is known as the ARP mapping table. A series of maps between IP addresses and MAC addresses of other hosts which recently communicate with the local host are stored in the ARP mapping table. When a dynamic ARP mapping entry is not in use for a specified period of time, the host removes it from the ARP mapping table so as to save the memory space and shorten the interval for the switch to search ARP mapping table.

Suppose there are two hosts on the same network segment: Host A and Host B. The IP address of Host A is IP_A and the IP address of Host B is IP_B. Host A will transmit messages to Host B. Host A checks its own ARP mapping table first to know whether there are corresponding ARP entries of IP_B in the table. If the corresponding MAC address is found, Host A uses the MAC address in the ARP mapping table to encapsulate the IP packet in frame and sends it to Host B. If the corresponding MAC address is not found, Host A stores the IP packet in the queue waiting for transmission, create an ARP request packet and broadcast it throughout the Ethernet. The ARP request packet contains the IP address of Host B and IP address and MAC address of Host A. Since the ARP request packet is broadcasted, all hosts on the network segment can receive the request. However, only the requested host (namely, Host B) needs to process the request. Host B first stores the IP address and the MAC address of the request sender (Host A) in the ARP request packet in its own ARP mapping table. Then, Host B generates an ARP reply packet by adding its own MAC address into the packet, and then send it to Host A. The reply packet is directly sent to Host A in stead of being broadcasted. Receiving the reply packet, Host A extracts the IP address and the

corresponding MAC address of Host B and adds them to its own ARP mapping table. Then Host A sends Host B all the packets standing in the queue.

Normally, dynamic ARP takes effect and automatically searches for the resolution from the IP address to the Ethernet MAC address without the help of an administrator.

12.2 Configuring ARP

The ARP mapping table can be maintained dynamically or manually. Usually, the manually configured mapping from the IP addresses to the MAC addresses is known as static ARP. The user can display, add or delete the entries in the ARP mapping table through relevant manual maintenance commands.

The following sections describe static ARP configuration tasks:

- Manually Adding/Deleting Static ARP Mapping Entries
- Configuring the Dynamic ARP Aging Timer
- Enabling/Disabling the Checking Function of ARP Entry

12.2.1 Manually Adding/Deleting Static ARP Mapping Entries

Perform the following configuration in system view.

Table 12-1 Manually add/delete static ARP mapping entries

Operation	Command
Manually add a static ARP mapping entry	arp static <i>ip-address mac-address</i> [<i>vlan-id</i> { <i>interface_type interface_num</i> <i>interface_name</i> } <i>vpn-instance-name</i>]
Manually delete a static ARP mapping entry	undo arp <i>ip-address</i>

By default, the ARP mapping table is empty and the address mapping is obtained through dynamic ARP.

Note that:

- As long as a switch operates, its static ARP mapping entries remain valid unless you perform operations that make ARP invalid, such as change or remove VLAN virtual interfaces, remove a VLAN, or remove an interface from a VLAN. These operations cause the corresponding ARP mapping entries to be automatically removed.
- The *vlan-id* argument must be the ID of a VLAN that has been created by the user, and the Ethernet port specified behind this parameter must belong to the VLAN.
- As for the *vpn-instance-name* argument, you must provide the VPN-instance name of an existing MPLS VPN for it.
- ARP map entries with port parameters cannot be configured for aggregated ports.

12.2.2 Configuring the Dynamic ARP Aging Timer

For purpose of flexible configuration, the system provides the following commands to assign dynamic ARP aging period. When the system learns a dynamic ARP entry, its aging period is based on the current value configured.

Perform the following configuration in system view.

Table 12-2 Configure the dynamic ARP aging timer

Operation	Command
Configure the dynamic ARP aging timer	arp timer aging <i>aging-time</i>
Restore the default dynamic ARP aging time	undo arp timer aging

By default, the aging time of dynamic ARP aging timer is 20 minutes.

12.2.3 Enabling/Disabling the Checking Function of ARP Entry

You can use the following command to control the device whether to learn the ARP entry where the MAC address is a multicast MAC address.

Perform the following configuration in system view.

Table 12-3 Enable/Disable the checking function of ARP entry

Operation	Command
Enable the checking of ARP entry, that is, the device does not learn the ARP entry where the MAC address is a multicast MAC address	arp check enable
Disable the checking of ARP entry, that is, the device learns the ARP entry where the MAC address is a multicast MAC address	undo arp check enable

By default, the checking of ARP entry is enabled, that is, the device does not learn the ARP entry where the MAC address is a multicast MAC address.

12.3 Displaying and Debugging ARP

After the above configuration, execute the **display** command in any view to display the running of the ARP configuration, and to verify the effect of the configuration.

Execute the **reset** command in user view to clear ARP mapping table. Execute the **debugging** command in user view to debug ARP configuration.

Table 12-4 Display and debug ARP

Operation	Command
Display ARP mapping table	display arp [<i>ip-address</i> [dynamic static] [{ begin include exclude } <i>text</i>]]
Display the current setting of the dynamic ARP aging timer	display arp timer aging
Reset ARP mapping table	reset arp [dynamic static interface { <i>interface_type</i> <i>interface_num</i> <i>interface_name</i> } all]
Enable ARP information debugging	debugging arp { error info packet }
Disable ARP information debugging	undo debugging arp { error info packet }

12.4 Enabling/Disabling the Scheme of Preventing Attack from Packets

12.4.1 Introduction to the Scheme of Preventing Attack from Packets

A scheme of preventing attack from packets is designed against some typical attack modes on the 8800 series switches. The scheme can prevent attacks from IP, ARP, 802.1x and unknown multicast packets.

- IP packet attack: Means that a Switch 8800 receives too many IP packets whose destination addresses and VLAN port address are in the same segment. The switch has no corresponding forwarding entries for the packets, therefore they are sent to the CPU, occupying lots of CPU resource and even affecting normal data forwarding.
- ARP packet attack: Means that a Switch 8800 receives lots of ARP request packets with the same or similar source media access control (MAC) addresses, affecting normal ARP learning.
- 802.1x packet attack: Means that a Switch 8800 receives lots of 802.1x authentication packets with the same or similar source MAC addresses, consequently occupying the CPU resources.

Perform the following configuration in system view.

Table 12-5 Enable/Disable the scheme of preventing attack from packets

Operation	Command
Enable/Disable the scheme of preventing attack from packets	anti-attack { arp dot1x ip }{ disable enable }

By default, the scheme of preventing attack from IP packets is enabled; the scheme of preventing attack from ARP packets and dot1x packets is disabled.

Chapter 13 DHCP Configuration

13.1 Introduction to DHCP

13.1.1 How DHCP Works

This is a world where networks are ever-growing in both size and complexity, and the network configuration is getting more and more complex. As is often the case, the number of hosts in a network exceeds that of the available IP addresses, and position changes of hosts (when users carry their laptops from here to there, or move to a wireless network) require reassigned new IP addresses. Dynamic host configuration protocol (DHCP) is designed to accommodate this context. DHCP adopts client/server model, where DHCP clients send requests to the DHCP server dynamically and the DHCP server in turn returns corresponding configuration information (such as IP addresses) according to the policies configured for it.

A typical DHCP implementation comprises a DHCP server and multiple DHCP clients (PCs or laptops). Figure 13-1 illustrates a network that employs DHCP.

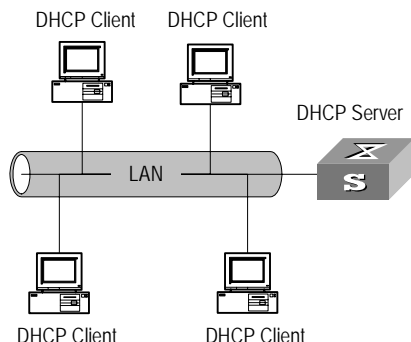


Figure 13-1 Network diagram for DHCP

I. IP address assignment

1) IP address assignment policy

Different types of clients have different requirements for IP addresses. Servers usually require long-term fixed IP addresses, some hosts may require automatically assigned long-term fixed IP addresses, and some hosts may only require dynamically assigned temporary IP addresses.

A DHCP server provides three policies to meet these requirements.

- Manual IP address assignment. The administrator assigns fixed IP addresses to DHCP clients that are of special uses, such as a WWW server.

- Automatic IP address assignment. The DHCP server automatically assigns fixed IP addresses to DHCP clients when they connect to the network for the first time. After that, the IP addresses are always occupied by the DHCP clients.
- Dynamic IP address assignment. The DHCP server leases IP addresses to DHCP clients for predetermined period of time and reclaims them at the expiration of the period. In this case, a DHCP client must reapply for an IP address regularly. This is the common case for normal users.

IP address assignment orderThe DHCP server assigns IP addresses except the forbidden ones to clients in the following orders.

- IP addresses in the address pool of the DHCP server that are statically bound to the MAC addresses of the DHCP clients.
 - IP addresses that are reclaimed by the DHCP server. That is, those in the Requested IP Addr Option fields of DHCP Discover packets sent by DHCP clients.
 - The first available IP address in the address pool the DHCP server finds.
 - The first expired or once conflicted IP address it finds. A DHCP server returns an error if it cannot find any available IP address from all these types of IP addresses when assigning an IP address.
- 2) Types of address pools of DHCP server
- Global address pool, valid for the entire switch. An address pool of this type is created using the **dhcp server ip-pool** command in system view.
 - VLAN interface address pool, valid for a specific VLAN interface. An address pool of this type is created by the system when the VLAN interface is configured with a legal unicast IP address and you specify to assign IP addresses in VLAN interface address pool using the **dhcp select interface** command in VLAN interface view. The address range of the available addresses is that of the network segment the VLAN interface resides.

II. Communications between DHCP clients and DHCP server

To obtain valid dynamic IP addresses, the DHCP clients exchange different information with the DHCP server in different phases. Usually, three modes are involved:

1) First round registration

A DHCP client goes through the following four steps when it accesses a network for the first time:

- Discovery. The DHCP client tries to find a DHCP server by broadcasting a DHCP_Discover packet in the network. (Only DHCP servers respond to this type of packet.)
- Provision. Each DHCP server that receives the DHCP_Discover packet selects an available IP address from an address pool and sends a DHCP_Offer packet that carries the selected IP address and other configuration information to the DHCP client.

- Selection. The DHCP client only receives the first arriving DHCP_Offer packet if there are DHCP_Offer packets from several DHCP servers. Then, it retrieves the IP address carried in the packet, and broadcasts a DHCP_Request packet to each DHCP server. The packet contains the IP address carried by the DHCP_Offer packet.
- Acknowledgement. Upon receiving the DHCP_Request packet, the DHCP server that owns the IP address the DHCP_Request packet carries sends a DHCP_ACK packet to the DHCP client. And then the DHCP client binds TCP/IP protocol components to its network adapter.
- IP addresses offered by other DHCP servers (if any) through DHCP_Offer packets but not selected by the DHCP client are still available for other clients.

2) Second round registration

A second round registration goes through the following steps:

- After going through the first round registration successfully and logging out, when the DHCP client logs on to the network again, it directly broadcasts a DHCP_Request packet that contains the IP address assigned to it in the first round registration instead of a DHCP_Discover packet. .
- Upon receiving the DHCP_Request packet, if the IP address carried in the packet is still available, the DHCP server owning the IP address answers with a DHCP_ACK packet to enable the DHCP client to use the IP address again.
- If the IP address is not available (for example, it is occupied by other DHCP client), the DHCP server answers with a DHCP_NAK packet, which enables the DHCP client to go through steps in the first round registration.

3) Prolonging the lease time of IP address

An IP address assigned dynamically is valid for a specified lease time and will be reclaimed by the DHCP server when the time expires. So the DHCP client must update the lease to prolong the lease time if it is to use the IP address for a longer time.

By default, a DHCP client updates its IP address lease automatically by sending a DHCP_Request packet to the DHCP server when half of the lease time elapses. The DHCP server, in turn, answers with a DHCP_ACK packet to notify the DHCP client of the new lease.

13.2 Configuring General DHCP

General DHCP configuration refers to those that are applicable to both DHCP server and DHCP relay.

The following sections describe the general DHCP configuration tasks:

- Enabling/Disabling DHCP Service
- Configuring Processing Method of DHCP Packets
- Enabling/Disabling Fake DHCP Server Detection

13.2.1 Enabling/Disabling DHCP Service

For both DHCP server and DHCP relay, you must enable the DHCP service first before performing other DHCP configurations. The other related DHCP configurations take effect only after the DHCP service is enabled.

Perform the following configuration in system view.

Table 13-1 Enable/Disable DHCP service

Operation	Command
Enable DHCP service	dhcp enable
Disable DHCP service	undo dhcp enable

DHCP service is disabled by default.

13.2.2 Configuring Processing Method of DHCP Packets

You can perform the configurations listed in the following tables on your switch. After that, the switch processes the DHCP packets it received from DHCP clients in the methods you have configured.

Perform the following configuration in VLAN interface view to configure the processing method of DHCP packets for current VLAN interface.

Table 13-2 Configure the processing method for current VLN interface

Operation	Command
Specify to forward DHCP packets to local DHCP server and let the local server assign IP addresses in global address pools to DHCP clients	dhcp select global
Specify to forward DHCP packets to local DHCP server and let the local server assign IP addresses in VLAN interface address pool to DHCP clients	dhcp select interface
Specify to forward DHCP packets to remote DHCP servers. In this case, the current switch operates as a DHCP relay, and IP addresses are assigned by DHCP servers located in other networks	dhcp select relay
Revert to the default processing mode	undo dhcp select

Perform the following configuration in system view to configure the processing method of DHCP packets for multiple VLAN interfaces.

Table 13-3 Configure the processing method for multiple VLAN interfaces

Operation	Command
Specify to forward DHCP packets to local DHCP server and let the local server assign IP addresses in global address pools to DHCP clients	dhcp select global { interface vlan-interface <i>vlan_id</i> [to vlan-interface <i>vlan_id</i>] all }
Specify to forward DHCP packets to local DHCP server and let the local server assign IP addresses in VLAN interface address pool to DHCP clients	dhcp select interface { interface vlan-interface <i>vlan_id</i> [to vlan-interface <i>vlan_id</i>] all }
Specify to forward DHCP packets to remote DHCP servers. In this case, the current switch operates as a DHCP relay, and IP addresses are assigned by DHCP servers located in other networks	dhcp select relay { interface vlan-interface <i>vlan_id</i> [to vlan-interface <i>vlan_id</i>] all }
Revert to the default processing mode	undo dhcp select { interface vlan-interface <i>vlan_id</i> [to vlan-interface <i>vlan_id</i>] all }

By default, DHCP packets are processed in **global** method. That is, DHCP packets are forwarded to local DHCP server and IP addresses in global address pools are assigned.

13.2.3 Enabling/Disabling Fake DHCP Server Detection

If an unauthorized DHCP server exists in a network, it also answers when users in the network request IP addresses, and then interacts with the DHCP clients. This causes that the users cannot obtain correct IP addresses to access network. This kind of DHCP servers are known as fake DHCP servers.

With fake DHCP server detection enabled, the switch can record information (such as the IP addresses) about the DHCP servers. This helps administrators to detect fake DHCP servers in time and take proper measures.

Perform the following configuration in system view.

Table 13-4 Enable/Disable fake DHCP server detection

Operation	Command
Enable fake DHCP server detection	dhcp server detect
Disable fake DHCP server detection	undo dhcp server detect

Fake DHCP server detection is disabled by default.

13.3 Configuring DHCP Server

The following sections describe the DHCP server configuration tasks:

- Creating a Global DHCP IP Address Pool
- Configuring IP Address Assignment Mode
- Forbidding Specified IP Addresses to Be Automatically Assigned
- Configuring Lease Time For DHCP Address Pool
- Configuring DHCP Client Domain Names
- Configuring DNS Server Address for DHCP Clients
- Configuring NetBIOS Server Address for DHCP Clients
- Configuring NetBIOS Node Type for DHCP Clients
- Configuring Custom DHCP Options
- Configuring Outbound Gateway Address for DHCP Clients
- Configuring Parameters for DHCP Server to Send Ping Packets

Note:

Some of the above DHCP configurations can be performed for global IP address pools, IP address pool of current VLAN interface, or IP address pools of multiple specified VLAN interface respectively. They are:

- Configuring lease time for DHCP address pool
 - Configuring DHCP client domain names
 - Configuring DNS server address for DHCP clients
 - Configuring NetBIOS server address for DHCP clients
 - Configuring NetBIOS node type for DHCP clients
 - Configuring DHCP custom options
-

13.3.1 Creating a Global DHCP IP Address Pool

An IP address pool contains IP addresses that can be assigned to DHCP clients. In response to DHCP request sent by a DHCP client, the DHCP server selects an appropriate IP address pool based on your configuration, choose an available IP address from the pool, and sends the IP address and other parameters (such as the lease time of the IP address) to the DHCP client. At present, you can configure up to 128 global DHCP address pools for a DHCP server.

The address pools of a DHCP server are hierarchically grouped like a tree. The root holds the IP address of the network segment, the branches hold the subnet IP addresses, and finally, the leaves hold the IP addresses of DHCP clients, which are manually bound to the corresponding network adapters. Such a structure enables configurations to be inherited. That is, configurations of the network segment can be inherited by its subnets, whose configurations in turn can be inherited by their clients.

So, you can configure the parameters (such as domain name) that are common to all levels in the address pool structure or some subnets only for the network segment or for corresponding subnets.

The **display dhcp server tree** command displays the tree-like structure of address pool, where address pools on the same level are sorted by the time they are created.

The **dhcp server ip-pool** command can be used to create a global DHCP address pool and enter the corresponding address pool view. If the address pool already exists, this command brings you to the address pool view directly.

Perform the following configuration in system view.

Table 13-5 Create a global DHCP address pool

Operation	Command
Create a DHCP address pool and enter the corresponding DHCP address pool view	dhcp server ip-pool <i>pool-name</i>
Remove a DHCP address pool	undo dhcp server ip-pool <i>pool-name</i>

By default, no global DHCP address pool is created.

Note that a VLAN interface address pool is created by the system after a legal unicast IP address is assigned to the VLAN interface and you specify to assign IP addresses in VLAN interface address pool by using the **dhcp select interface** command in VLAN interface view.

13.3.2 Configuring IP Address Assignment Mode

IP address can be assigned in two modes: static binding and dynamic assignment. You can statically bind an IP address in an address pool to the MAC address of a client or configure a address range to allow the DHCP server dynamic allocate the addresses in the range to DHCP clients. The two modes cannot coexist in a global DHCP address pool, but they can coexist in a VLAN interface address pool (but those that are dynamically assigned have the same network segment as that of the IP address of the VLAN interface).

For the dynamic assignment mode, you must specify the range of the addresses to be dynamically assigned. A global DHCP address pool whose IP addresses are statically bound to DHCP clients is actually a special kind of DHCP address pool.

I. Configuring static address binding for a global DHCP address pool

fixed IP address to the MAC address of a DHCP client who needs fixed IP address. After that, when the client requests for an IP address, the DHCP server finds (according to the MAC address) and assigns the fixed IP address to the client. At present, only one-to-one MAC-IP binding is supported for global DHCP address pool.

Perform the following configuration in DHCP address pool view.

Table 13-6 Configure static address binding for a global DHCP address pool

Operation	Command
Configure an IP address to be statically bound	static-bind ip-address <i>ip-address</i> [mask netmask]
Free a statically bound IP address	undo static-bind ip-address
Configure a MAC address to be statically bound	static-bind mac-address <i>mac-address</i>
Free a statically bound MAC address	undo static-bind mac-address

IP addresses in a global DHCP address pool are not statically bound by default.

Note:

The **static-bind ip-address** command and the **static-bind mac-address** command must be used together as a pair when you configure static binding entries. When you re-execute the command pair with the same IP address/MAC address, the newly configured IP address/MAC address overwrites the existing one.

II. Configuring static address binding for a VLAN interface address pool

At present, a VLAN interface DHCP address pool supports one-to-multiple MAC-IP address binding.

Perform the following configuration in VLAN interface view.

Table 13-7 Configure static address binding for a VLAN interface address pool

Operation	Command
Configure static address binding for the current VLAN interface address pool	dhcp server static-bind ip-address <i>ip-address mac-address mac-address</i>
Remove a statically bound IP address entry	undo dhcp server static-bind { ip-address <i>ip-address</i> mac-address <i>mac-address</i> }

IP addresses in the address pool of a VLAN interface are not statically bound by default.

**Caution:**

A binding in a VLAN interface address pool cannot be overwritten directly. If an IP-to-MAC address binding entry is configured and you want to modify it, you must remove it and redefine a new one.

III. Configuring dynamic IP address assignment

If you specify to assign IP addresses dynamically, that is, IP addresses are leased permanently or temporarily, you need to configure an available address range.

Perform the following configuration in DHCP address pool view.

Table 13-8 Configure an address range for dynamic IP address assignment

Operation	Command
Configure an address range for dynamic IP address assignment	network <i>ip-address</i> [mask <i>netmask</i>]
Remove an dynamic assignment address range	undo network

By default, no IP address range is configured for dynamic IP address assignment.

Each DHCP address pool can be configured with only one address range. If you execute the **network** command multiple times, then only the last configured address range works.

13.3.3 Forbidding Specified IP Addresses to Be Automatically Assigned

You can use the command here to prevent a DHCP server from assigning IP addresses that are already occupied by such network devices as gateways and file transfer protocol (FTP) servers to other DHCP clients to avoid IP address conflicts.

Perform the following configuration in system view.

Table 13-9 Forbid specified IP addresses to be automatically assigned

Operation	Command
Forbid specified IP addresses to be automatically assigned	dhcp server forbidden-ip <i>low-ip-address</i> [<i>high-ip-address</i>]
Cancel the forbiddance	undo dhcp server forbidden-ip <i>low-ip-address</i> [<i>high-ip-address</i>]

All IP addresses in a DHCP address pool can be automatically assigned by default.

You can set multiple IP address ranges that are not assigned automatically by executing the **dhcp server forbidden-ip** command multiple times.

13.3.4 Configuring Lease Time For DHCP Address Pool

You can configure different lease times for different DHCP address pools. But you can configure only one lease time for one DHCP address pool and all the address in the same pool will have the same lease time.

I. Configuring a lease time for a global DHCP address pool

Perform the following configuration in DHCP address pool view.

Table 13-10 Configure a lease time for a global DHCP address pool

Operation	Command
Configure a lease time for a global DHCP address pool	expired { day <i>day</i> [hour <i>hour</i> [minute <i>minute</i>]] unlimited }
Restore the lease time of a global DHCP address pool to the default value	undo expired

II. Configuring a lease time for current VLAN interface

Perform the following configuration in VLAN interface view.

Table 13-11 Configure a lease time for current VLAN interface

Operation	Command
Configure a lease time for DHCP address pool of current VLAN interface	dhcp server expired { day <i>day</i> [hour <i>hour</i> [minute <i>minute</i>]] unlimited }
Restore the lease time of DHCP address pool of current VLAN interface to the default value	undo dhcp server expired

III. Configuring a lease time for multiple VLAN interfaces

Perform the following configuration in system view.

Table 13-12 Configure a lease time for multiple VLAN interfaces

Operation	Command
Configure a lease time for DHCP address pools of multiple VLAN interfaces	dhcp server expired { day <i>day</i> [hour <i>hour</i> [minute <i>minute</i>]] unlimited } { interface <i>vlan-interface</i> <i>vlan_id</i> [to <i>vlan-interface</i> <i>vlan_id</i>] all }

Operation	Command
Restore the lease time of DHCP address pools of multiple VLAN interfaces to the default value	undo dhcp server expired { interface vlan-interface <i>vlan_id</i> [to vlan-interface <i>vlan_id</i>] all }

The default lease times for global address pools and VLAN interface address pools are all one day.

13.3.5 Configuring DHCP Client Domain Names

You can configure a domain name used by DHCP clients for each address pool on a DHCP server.

I. Configuring a DHCP client domain name for a global DHCP address pool

Perform the following configuration in DHCP address pool view.

Table 13-13 Configure a DHCP client domain name for a global DHCP address pool

Operation	Command
Configure a DHCP client domain name for a global DHCP address pool	domain-name <i>domain-name</i>
Remove the DHCP client domain name configured for a global DHCP address pool	undo domain-name

II. Configuring a DHCP client domain name for current VLAN interface

Perform the following configuration in VLAN interface view.

Table 13-14 Configure a DHCP client domain name for current VLAN interface

Operation	Command
Configure a DHCP client domain name for the DHCP address pool of the current VLAN interface	dhcp server domain-name <i>domain-name</i>
Remove the DHCP client domain name configured for the DHCP address pool of the current VLAN interface	undo dhcp server domain-name

III. Configuring a DHCP client domain name for multiple VLAN interfaces

Perform the following configuration in system view.

Table 13-15 Configure a DHCP client domain name for multiple VLAN interfaces

Operation	Command
Configure a DHCP client domain name for DHCP address pools of multiple VLAN interfaces	dhcp server domain-name <i>domain-name</i> { interface vlan-interface <i>vlan_id</i> [to vlan-interface <i>vlan_id</i>] all }
Remove the DHCP client domain name configured for DHCP address pools of multiple VLAN interfaces	undo dhcp server domain-name <i>domain-name</i> { interface vlan-interface <i>vlan_id</i> [to vlan-interface <i>vlan_id</i>] all }

By default, global address pools and VLAN interface address pools are not configured with any DHCP client domain name.

If you execute the **dhcp server domain-name** command multiple times, the newly configured DHCP client domain name overwrites the existing one.

13.3.6 Configuring DNS Server Address for DHCP Clients

When a host uses a domain name to access the Internet, the domain name must be translated into an IP address. Domain name system (DNS) is responsible for the translation. Therefore, when a DHCP server assigns an IP address to a DHCP client, it must also send a DNS server address to the client. At present, you can configure up to eight DNS server addresses for one DHCP address pool.

I. Configuring DNS server address for a global DHCP address pool

Perform the following configuration in DHCP address pool view.

Table 13-16 Configure DNS server address for a global DHCP address pool

Operation	Command
Configure one or more DNS server addresses for a global DHCP address pool	dns-list <i>ip-address</i> [<i>ip-address</i>]
Remove one or all DNS server addresses configured for a global DHCP address pool	undo dns-list { <i>ip-address</i> all }

II. Configuring DNS server address for current VLAN interface

Perform the following configuration in VLAN interface view.

Table 13-17 Configure DNS server address for current VLAN interface

Operation	Command
Configure one or more DNS server addresses for the DHCP address pool of the current VLAN interface	dhcp server dns-list <i>ip-address</i> [<i>ip-address</i>]
Remove one or all DNS server addresses configured for the DHCP address pool of the current VLAN interface	undo dhcp server dns-list { <i>ip-address</i> all }

III. Configuring DNS server address for multiple VLAN interfaces

Perform the following configuration in system view.

Table 13-18 Configure DNS server address for multiple VLAN interfaces

Operation	Command
Configure one or more DNS server addresses for the DHCP address pools of multiple VLAN interfaces	dhcp server dns-list <i>ip-address</i> [<i>ip-address</i>] { interface vlan-interface <i>vlan_id</i> [to vlan-interface <i>vlan_id</i>] all }
Remove one or all DNS server addresses configured for the DHCP address pools of multiple VLAN interfaces	undo dhcp server dns-list { <i>ip-address</i> all } { interface vlan-interface <i>vlan_id</i> [to vlan-interface <i>vlan_id</i>] all }

By default, no DNS server address is configured for global and VLAN interface address pools.

If you execute the **dhcp server dns-list** command multiple times, the newly configured IP addresses overwrite the existing ones.

13.3.7 Configuring NetBIOS Server Address for DHCP Clients

For clients running a Windows operating system and communicating through the NetBIOS protocol, translations between host name and IP address are carried out by Windows Internet Naming Service (WINS) servers. So you need to perform configurations concerning WINS for these clients. At present, you can configure up to eight NetBIOS server addresses for a DHCP address pool.

I. Configuring NetBIOS server address for a global DHCP address pool

Perform the following configuration in DHCP address pool view.

Table 13-19 Configure NetBIOS server address for a global DHCP address pool

Operation	Command
Configure one or more NetBIOS server addresses for a global DHCP address pool	nbns-list <i>ip-address</i> [<i>ip-address</i>]
Remove one or all NetBIOS server addresses configured for a global DHCP address pool	undo nbns-list { <i>ip-address</i> all }

II. Configuring NetBIOS server address for current VLAN interface

Perform the following configuration in VLAN interface view.

Table 13-20 Configure NetBIOS server address for current VLAN interface

Operation	Command
Configure one or more NetBIOS server addresses for the DHCP address pool of current VLAN interface	dhcp server nbns-list <i>ip-address</i> [<i>ip-address</i>]
Remove one or all NetBIOS server addresses configured for the DHCP address pool of the current VLAN interface	undo dhcp server nbns-list { <i>ip-address</i> all }

III. Configuring NetBIOS server address for multiple VLAN interfaces

Perform the following configuration in system view.

Table 13-21 Configure NetBIOS server address for multiple VLAN interfaces

Operation	Command
Configure one or more NetBIOS server addresses for the DHCP address pools of multiple VLAN interfaces	dhcp server nbns-list <i>ip-address</i> [<i>ip-address</i>] { interface <i>vlan-interface</i> <i>vlan_id</i> [to <i>vlan-interface</i> <i>vlan_id</i>] all }
Remove one or all NetBIOS server addresses configured for the DHCP address pools of multiple VLAN interfaces	undo dhcp server nbns-list { <i>ip-address</i> all } { interface <i>vlan-interface</i> <i>vlan_id</i> [to <i>vlan-interface</i> <i>vlan_id</i>] all }

By default, no NetBIOS server address is configured for global and VLAN interface address pools.

If you execute the **dhcp server nbns-list** command multiple times, the newly configured IP addresses overwrite the existing ones.

13.3.8 Configuring NetBIOS Node Type for DHCP Clients

For DHCP clients communicating in wide area network (WAN) by NetBIOS protocol, the mapping between their host names and IP addresses must be established. According to the ways they establish their mappings, NetBIOS nodes fall into the following four types:

- b-node: Nodes of this type establish their mappings by broadcasting. (b stands for broadcast.)
- p-node: Nodes of this type establish their mappings by communicating with NetBIOS server. (p stands for peer-to-peer.)
- m-node: Nodes of this type are p nodes which take some broadcast features. (m stands for mixed.)
- h-node: Nodes of this type are b nodes which take peer-to-peer mechanism. (h stands for hybrid.)

I. Configuring NetBIOS node type for a global DHCP address pool

Perform the following configuration in DHCP address pool view.

Table 13-22 Configure a NetBIOS node type for a global DHCP address pool

Operation	Command
Configure the NetBIOS node type for a global DHCP address pool	netbios-type { b-node h-node m-node p-node }
Cancel the NetBIOS node type configuration for a global DHCP address pool	undo netbios-type

II. Configuring NetBIOS node type for current VLAN interface

Perform the following configuration in VLAN interface view.

Table 13-23 Configure a NetBIOS node type for current VLAN interface

Operation	Command
Configure the NetBIOS node type for DHCP clients of the current VLAN interface DHCP address pool	dhcp server netbios-type { b-node h-node m-node p-node }
Remove NetBIOS node type configured for DHCP clients of the current VLAN interface DHCP address pool	undo dhcp server netbios-type

III. Configuring NetBIOS node type for multiple VLAN interfaces

Perform the following configuration in system view.

Table 13-24 Configure a NetBIOS node type for multiple VLAN interfaces

Operation	Command
Configure NetBIOS node types for DHCP clients of multiple VLAN interface DHCP address pools	dhcp server netbios-type { b-node h-node m-node p-node } { interface vlan-interface <i>vlan_id</i> [to vlan-interface <i>vlan_id</i>] all }
Remove NetBIOS node type configurations of multiple VLAN interface DHCP address pools	undo dhcp server netbios-type { interface vlan-interface <i>vlan_id</i> [to vlan-interface <i>vlan_id</i>] all }

By default, the DHCP clients of global and VLAN interface address pools are all of h-node type.

13.3.9 Configuring Custom DHCP Options

With the evolvement of DHCP, new options come forth continuously. To utilize these options, you can manually add them to the property list of a DHCP server.

I. Configuring custom DHCP options for a global DHCP address pool

Perform the following configuration in DHCP address pool view.

Table 13-25 Configure a custom DHCP options for a global DHCP address pool

Operation	Command
Configure a custom DHCP option for a global DHCP address pool	option code { ascii <i>ascii-string</i> hex <i>hex-string</i> ip-address <i>ip-address</i> [<i>ip-address</i>] }
Remove a custom DHCP option configured for a global DHCP address pool	undo option code

II. Configuring custom DHCP options for current VLAN interface

Perform the following configuration in VLAN interface view.

Table 13-26 Configure custom DHCP options for current VLAN interface

Operation	Command
Configure a custom DHCP option for DHCP address pool of the current VLAN interface	dhcp server option code { ascii <i>ascii-string</i> hex <i>hex-string</i> ip-address <i>ip-address</i> [<i>ip-address</i>] }
Remove a custom DHCP option configured for the DHCP address pool of the current VLAN interface	undo dhcp server option code

III. Configuring custom DHCP options for multiple VLAN interfaces

Perform the following configuration in system view.

Table 13-27 Configure custom DHCP options for multiple VLAN interfaces

Operation	Command
Configure a custom DHCP option for DHCP address pools of multiple VLAN interfaces	dhcp server option code { ascii <i>ascii-string</i> hex <i>hex-string</i> ip-address <i>ip-address</i> [<i>ip-address</i>] } { interface vlan-interface <i>vlan_id</i> [to vlan-interface <i>vlan_id</i>] all }
Remove a custom DHCP option configured for DHCP address pools of multiple VLAN interfaces	undo dhcp server option code { interface vlan-interface <i>vlan_id</i> [to vlan-interface <i>vlan_id</i>] all }

If you execute the **dhcp server option** command multiple times, the newly configured option overwrites the existing one.

13.3.10 Configuring Outbound Gateway Address for DHCP Clients

An outbound gateway enables DHCP clients to access external network devices. Packets destined for external networks are forwarded by outbound gateways. At present, you can configure up to eight IP addresses for outbound gateways.

Perform the following configuration in DHCP address pool view.

Table 13-28 Configure outbound gateway address for DHCP clients

Operation	Command
Configure one or more outbound gateway addresses for DHCP clients	gateway-list <i>ip-address</i> [<i>ip-address</i>]
Remove one or all outbound gateway addresses configured for DHCP clients	undo gateway-list { <i>ip-address</i> all }

By default, no outbound gateway address is configured for DHCP clients.

If you execute the **gateway-list** command multiple times, the newly configured IP addresses overwrite the existing ones.

13.3.11 Configuring Parameters for DHCP Server to Send Ping Packets

To avoid address conflict caused by reassigning an in-use IP address, before assigning an IP address to a DHCP client, the DHCP server detects the network using the **ping** instructions to ensure the IP address is not occupied. The DHCP server determines whether an IP address is reachable by sending specified number of ping packets. It waits for response packet for a specified period after sending each of these packets. If

the DHCP server receives no response after sending all these packets, it considers the IP address is not used by other devices in this network and assigns the IP address to this DHCP client. Otherwise, it does not assign the IP address.

Perform the following configuration in system view.

Table 13-29 Configure parameters for DHCP server to send ping packets

Operation	Command
Set the maximum number of ping packets the DHCP is allowed to send	dhcp server ping packets <i>number</i>
Revert to the default maximum number	undo dhcp server ping packets
Set the maximum duration for the DHCP server to wait for response to a ping packet	dhcp server ping timeout <i>milliseconds</i>
Revert to the default maximum duration	undo dhcp server ping timeout

By default, the DHCP server sends up to 2 ping packets to test an IP address and waits for a response for up to 500 milliseconds before it sends another ping packet.

Note that the DHCP server detects address conflict by ping packets, whereas a DHCP client does this by ARP packets.

13.3.12 Displaying and Debugging the DHCP Server

After the above configuration, you can execute the **display** command in any view to display operating information about the DHCP server to verify your configuration, and execute the **debugging** command to enable debugging for the DHCP server

Execute the following command in any view.

Table 13-30 Display the configuration information about the DHCP server

Operation	Command
Display the statistics about DHCP address conflicts	display dhcp server conflict { all ip <i>ip-address</i> }
Display information about lease-expired addresses in DHCP address pool(s). The lease-expired IP addresses in an address pool are assigned to other DHCP clients as needed if the address pool runs out of its available IP addresses	display dhcp server expired { ip <i>ip-address</i> pool [<i>pool-name</i>] interface [vlan-interface <i>vlan_id</i>] all }
Display the ranges of available (unassigned) IP addresses in DHCP address pools	display dhcp server free-ip

Operation	Command
Display the information about IP address binding in DHCP address pool(s)	display dhcp server ip-in-use { ip <i>ip-address</i> pool [<i>pool-name</i>] interface [vlan-interface <i>vlan_id</i>] all }
Display the statistics about the DHCP server	display dhcp server statistics
Display the information about the tree-like structure of DHCP address pool(s)	display dhcp server tree { pool [<i>pool-name</i>] interface [vlan-interface <i>vlan_id</i>] all }

Perform the following configuration in user view.

Table 13-31 Enable/Disable debugging for the DHCP server

Operation	Command
Disable debugging for the DHCP server	undo debugging dhcp server { all error event packet }
Enable debugging for the DHCP server	debugging dhcp server { all error event packet }

13.3.13 Clearing the Configuration Information of the DHCP Server

You can clear the configuration information of the DHCP server by executing the **reset** command in user view.

Perform the following configuration in user view.

Table 13-32 Clear the configuration information of the DHCP server

Operation	Command
Clear the statistics about DHCP address conflicts	reset dhcp server conflict { ip <i>ip-address</i> all }
Clear the information about dynamically bound DHCP addresses	reset dhcp server ip-in-use { all interface [vlan-interface <i>vlan_id</i>] ip <i>ip-address</i> pool [<i>pool-name</i>] }
Clear the statistics about the DHCP server	reset dhcp server statistics

13.3.14 DHCP Server Configuration Example

I. Network requirements

As shown in Figure 13-2, two DHCP clients at the same network segment (10.110.0.0) are connected to the following switch through a port in VLAN2. The switch, acting as a

DHCP server, is supposed to assign IP addresses to the two DHCP clients without the help of any DHCP Relay.

II. Network diagram

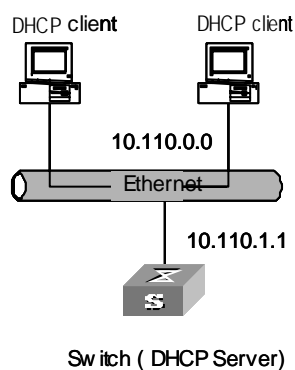


Figure 13-2 Network diagram for DHCP server

III. Configuration procedure

Enter system view.

```
<SW8800>system-view
```

Create VLAN2.

```
[SW8800]vlan 2
```

Enter VLAN interface view and create Vlan-interface 2.

```
[SW8800]interface Vlan-interface 2
```

Assign an IP address to Vlan-interface 2.

```
[SW8800-Vlan-interface2]ip address 10.110.1.1 255.255.0.0
```

Specify to assign IP addresses in the interface address pool to DHCP clients.

```
[SW8800-Vlan-interface2]dhcp select interface
```

Specify to assign IP addresses in global address pool to DHCP clients (it is also the default configuration).

```
[SW8800-Vlan-interface2]dhcp select global
```

Or execute the following command to revert to the default.

```
[SW8800-Vlan-interface2]undo dhcp select
```

Configure a global address pool.

```
[8505Tlhy]dhcp server ip-pool 1
```

```
[8505Tlhy-dhcp-1]network 10.110.0.0 mask 255.255.0.0
```

```
[8505Tlhy-dhcp-1]gateway-list 10.110.1.1
```


13.4 Configuring DHCP Relay

13.4.1 Introduction to DHCP Relay

This is a world where networks are ever-growing in both size and complexity, and the network configuration is getting more and more complex. As is often the case, the number of hosts in a network exceeds that of the available IP addresses, and position changes of hosts (when users carry their laptops from here to there, or move to a wireless network) require reassigned new IP addresses. Dynamic host configuration protocol (DHCP) is designed to accommodate this context. DHCP adopts client/server model, where DHCP clients send requests to the DHCP server dynamically and the DHCP server in turn returns corresponding configuration information according to the policies configured for it.

Early implementations of DHCP only work when DHCP clients and DHCP servers are in the same subnet. That is, they cannot work across networks. So, to implement dynamic host configuration, you must deploy at least one DHCP server in each subnet, and this is obviously uneconomical. DHCP Relay is designed to resolve this problem. Through a DHCP relay, DHCP clients in a LAN can communicate with DHCP servers in other subnets to acquire IP addresses. This enables DHCP clients of multiple networks to share a common DHCP server and thus enables you to save your cost and perform centralized administration. Figure 13-3 illustrates a typical DHCP Relay application.

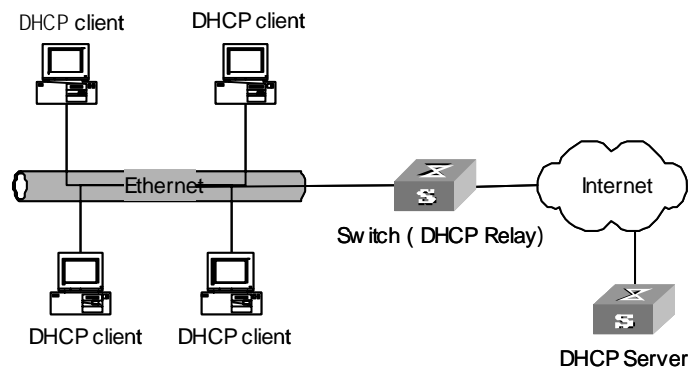


Figure 13-3 Network diagram for DHCP Relay

The dynamic host configuration procedure with DHCP relay is as follows:

- A DHCP client broadcasts configuration request packet in the local network when it starts up and initializes the configuration.
- If a DHCP server exists in the network, it processes the configuration request packet directly without the help of a DHCP Relay.
- If no DHCP server exists in the network, the network device serving as a DHCP Relay in the network appropriately processes the configuration request packet and forwards it to a specified DHCP server located in another network.

- After receiving the packet, the DHCP server generates configuration information accordingly and sends it to the DHCP client through the DHCP Relay to complete the dynamic configuration of the DHCP client.

Note that the entire configuration procedure may go through multiple times of such interactions.

13.4.2 Configuring DHCP Relay

DHCP Relay configuration includes the following: The following text describes the DHCP Relay configuration tasks:

- Configuring a DHCP server for a VLAN interface
- Configure user address entries for a DHCP Server
- Enable/Disable DHCP security on a VLAN interface
- Enabling/Disabling fake DHCP server detecting

I. Configuring a DHCP server for a VLAN interface

You can execute the **ip relay address** command to configure the DHCP packet processing mode on VLAN interface as relay and a corresponding DHCP server for a VLAN interface.

Perform the following configuration in VLAN interface view.

Table 13-33 Configure a corresponding DHCP server for a VLAN interface

Operation	Command
Configure a corresponding DHCP server for current VLAN interface	ip relay address <i>ip_address</i>
Remove the DHCP server configured for current VLAN interface	undo ip relay address { <i>ip_address</i> all }

No DHCP server is configured for a VLAN interface by default.

Note that when configuring a new DHCP server for a VLAN that already has a DHCP server configured for it, the newly configured one does not overwrite the existing ones. Both the new and the old ones are valid. You can configure up to 20 DHCP server addresses for a VLAN interface.

II. Configure user address entries for a DHCP Server

In a VLAN that has DHCP Relay configured, to enable a DHCP client using a legal fixed IP address to pass the address checking of the DHCP security feature, you must add a static address entry for the DHCP client. A static address entry indicates the relation between a fixed IP address and a MAC address.

Perform the following configuration in system view.

Table 13-34 Configure user address entries for DHCP server

Operation	Command
Add a user address entry for DHCP server	dhcp relay security <i>ip_address</i> <i>mac_address</i> static
Remove a user address entry for DHCP server	undo dhcp relay security <i>ip_address</i>

III. Enable/Disable DHCP security on a VLAN interface

If you enable the DHCP security feature on a VLAN interface, the switch performs user address checking on the VLAN interface to prevent unauthorized binding request. If you disable the DHCP security feature on a VLAN interface, the switch does not perform user address checking on the VLAN interface.

Perform the following configuration in VLAN interface view.

Table 13-35 Enable/disable DHCP security on a VLAN interface

Operation	Command
Enable DHCP security on a VLAN interface	dhcp relay security address-check enable
Disable DHCP security on a VLAN interface	dhcp relay security address-check disable

DHCP security is disabled on a VLAN interface by default.

13.4.3 Displaying and Debugging DHCP Relay

After the above configuration, you can execute the **display** command in any view to display running information about DHCP Relay to verify your configuration.

Execute the **debugging** command in user view to debug DHCP Relay.

Table 13-36 Display and debug DHCP Relay

Operation	Command
Display information about the DHCP servers configured for VLAN interface	display dhcp relay address { interface vlan-interface <i>vlan-id</i> all }
Display information about legal user address entries for DHCP server	display dhcprelay-security [<i>ip_address</i>]
Enable debugging for DHCP Relay	debugging dhcp relay { all packet error event }

Operation	Command
Disable debugging for DHCP Relay	undo debugging dhcp relay { all packet error event }

13.4.4 DHCP Relay Configuration Example

I. Network requirements

As shown in Figure 13-4, two DHCP clients located at the same network segment (10.110.0.0) are connected to a switch through a port in VLAN 2. The switch, acting as a DHCP relay, is supposed to forward DHCP packets between the two DHCP clients and the DHCP server with the IP address of 202.38.1.2.

II. Network diagram

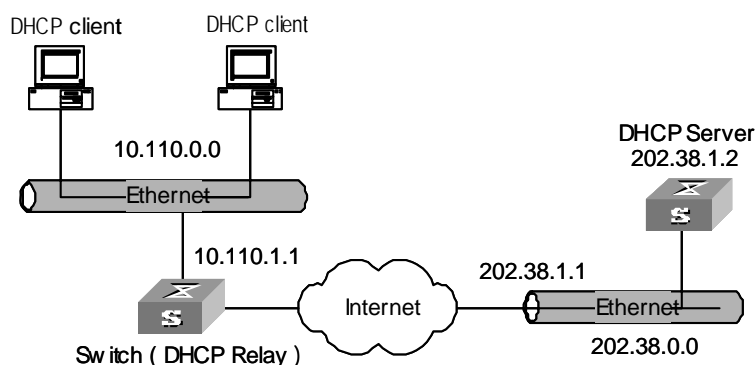


Figure 13-4 Network diagram for DHCP Relay

III. Configuration procedure

Enter system view.

```
<SW8800>system-view
```

Create VLAN 2.

```
[SW8800]vlan 2
```

Create Vlan-interface 2 and enter VLAN interface view.

```
[SW8800]interface Vlan-interface 2
```

Assign an IP address to Vlan-interface 2.

```
[SW8800-Vlan-interface2]ip address 10.110.1.1 255.255.0.0
```

Specify to forward DHCP packets to a remote DHCP server.

```
[SW8800-Vlan-interface2]dhcp select relay
```

Configure the IP address of the DHCP server to which VLAN 2 sends DHCP packets.

```
[SW8800-Vlan-interface2]ip relay address 202.38.1.2
```

Note:

Besides the above configurations for DHCP Relay, you need to configure address pool on the DHCP server and make sure the DHCP server and the switch interface connecting the two DHCP clients is routing reachable with each other.

Chapter 14 DNS Configuration

14.1 Introduction to DNS

Used in the TCP/IP application, Domain Name System (DNS) is a distributed database which provides the translation between domain name and the IP address. In this way, the user can use domain names that are easy to memorize and meaningful, and never needs to keep obscure IP addresses in mind.

There are two kinds of domain name resolutions: static domain name resolution and dynamic domain name resolution, which supplement each other in real application. On resolving a domain name, use the static resolution first. If it fails, use the dynamic resolution method. You can put some common domain names into the static domain name resolution table to raise the domain name resolution efficiency greatly.

14.1.1 Static Domain Name Resolution

Static domain resolution is to establish maps between domain name and the IP address manually. When you perform some applications using domain names, the system can obtain the IP address of the specified domain name by searching the static domain name resolution table.

14.1.2 Dynamic Domain Name Resolution

Dynamic domain name resolution is implemented by inquiring the domain name server. As a DNS client, the switch sends an inquiry request to the domain name server, and the domain name server searches the related IP address of the domain name in its own database and sends it back to the switch. If the domain name server judges that the domain name does not belong to the local domain, it forwards the request to the upper level domain name resolution server till the resolution is finished.

Dynamic domain name resolution supports the buffer function. It stores each successful domain name/IP address mapping that is resolved dynamically in the dynamic domain name buffer. When the same domain name is searched next time, it can be read directly from the buffer, without requesting the domain name server. The aged mapping in the buffer is deleted after a certain period of time to ensure the updated contents can be got from the domain name server timely. The aging time is set by the domain name server and obtained by the switch from the protocol packet.

Dynamic domain name resolution supports the domain name suffix list function. You can set some domain name suffixes beforehand and input part of the domain name field during the domain name resolution, then the system adds different suffixes to the input domain name automatically for resolution. For example, if a user wants to search the domain name "3Com.com", he can configure the "com" in the suffix list and input

“3Com”. Then the system connects the input domain name with the suffix into “3Com.com” automatically to search. When the domain name suffix is used, if the input domain name does not include “.”, like “3Com”, the system regards it as a host name and add a domain name suffix to search. After all the domain names are failed to be searched out in this way, the system finally searches with the primarily input domain name. If the input domain name does include “.”, like “www.3Com”, the system searches with it directly. The system adds each suffix to search one by one only after the search fails. If the input domain name contains a “.” in the final position, like “3Com.com.”, it indicates that the domain name suffix needs not to be added. The system removes the last “.” from the input domain name and search with the remaining part. Succeeded or not, the system returns to the originally input domain name. Put it more specifically, if the last character of the input domain name is “.”, the system only searches according to characters before the “.” rather than matches the domain name. In this sense, the last “.” is also called “search terminator”.

14.2 Configuring Static Domain Name Resolution

You can use this command to map the host name to the host IP address. When you use applications like Telnet, you can use the host name directly, and the system translates it into the IP address, rather than the obscure IP address.

Perform the following configuration in system view.

Table 14-1 Configure host name and the corresponding IP address

Operation	Command
Configure host name and the corresponding IP address	ip host <i>hostname ip-address</i>
Cancel host name and the corresponding IP address	undo ip host <i>hostname [ip-address]</i>

Each host can have only one IP address. If you configure a host name more than once, then the IP address configured at last is effective.

14.3 Configuring Dynamic Domain Name Resolution

Dynamic domain name resolution configuration includes:

- Enable/Disable Static Domain Name Resolution
- Configure the IP Address of Domain Name Server
- Configure Domain Name Suffix

14.3.1 Enable/Disable Static Domain Name Resolution

You can use the following command to enable dynamic domain name resolution. However, since dynamic domain name resolution may take some time, you can disable

this function when you do not want to perform dynamic domain name resolution sometimes.

Perform the following configuration in system view.

Table 14-2 Enable/disable dynamic domain name resolution

Operation	Command
Enable dynamic domain name resolution	dns resolve
Disable dynamic domain name resolution	undo dns resolve

By default, dynamic domain name resolution is disabled.

14.3.2 Configure the IP Address of Domain Name Server

You are required to configure the domain name sever if you need to use the function of the dynamic domain name resolution. In this way, you can send the inquiry request packets to the appropriate sever. The system supports up to six domain name servers.

Perform the following configuration in system view.

Table 14-3 Configure the IP address of the domain name sever

Operation	Command
Configure the IP address of the domain name sever	dns server <i>ip-address</i>
Delete the IP address of the domain name sever	undo dns server [<i>ip-address</i>]

14.3.3 Configure Domain Name Suffix

You can use the following command to configure domain name suffix list. By configuring this, you can just input part of the domain name and the system automatically adds the preconfigured suffix to perform the resolution. The system supports up to 10 domain name suffixes.

Perform the following configuration in system view.

Table 14-4 Configure domain name suffix

Operation	Command
Configure domain name suffix	dns domain <i>domain-name</i>
Delete domain name suffix	undo dns domain [<i>domain-name</i>]

14.4 Displaying and Debugging Domain Name Resolution

After the above configuration, you can execute the **display** command in any view to view the running states of the domain name resolution, and verify the configuration results through the displayed information.

Execute the **reset** command in user view to clear the dynamic domain name buffer. Execute the **debugging** command to debug the domain name resolution.

Table 14-5 Display and debug the domain name resolution

Operation	Command
Display the static domain name resolution table	display ip host
Display the information on domain name sever	display dns server [dynamic]
Display the information on domain name suffix list	display dns domain [dynamic]
Display the information on the dynamic domain name buffer	display dns dynamic-host
Clear dynamic domain name buffer	reset dns dynamic-host
Enable the debugging for the domain name resolution	debugging dns
Disable the debugging for the domain name resolution	undo debugging dns

14.5 DNS Configuration Example

I. Network requirements

As the client, the switch uses dynamic domain name resolution. The IP address of the domain name server is 172.16.1.1. The configured suffix of the domain name is "com". There is a route between the switch and the server.

II. Network diagram

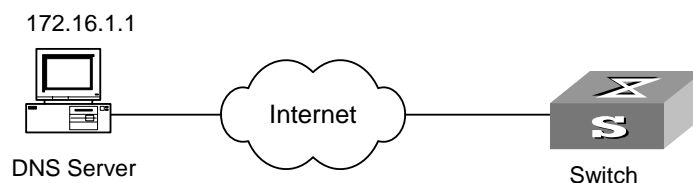


Figure 14-1 Network diagram for DNS client

III. Configuraiton procedure

Enable dynamic domain name resolution

```
[SW8800] dns resolve
```

Configure the IP address of the domain name server to 172.16.1.1.

```
[SW8800] dns server 172.16.1.1
```

Configure the domain name suffix as com.

```
[SW8800] dns domain com
```

Ping a host with the specified domain name.

```
[SW8800] ping ftp
```

```
Trying DNS server (172.16.1.1)
```

```
PING ftp.com (200.200.200.200): 56 data bytes, press CTRL_C to break
```

```
Reply from 200.200.200.200: bytes=56 Sequence=1 ttl=128 time=2 ms
```

```
Reply from 200.200.200.200: bytes=56 Sequence=2 ttl=128 time=2 ms
```

```
Reply from 200.200.200.200: bytes=56 Sequence=3 ttl=128 time=2 ms
```

```
Reply from 200.200.200.200: bytes=56 Sequence=4 ttl=128 time=2 ms
```

```
Reply from 200.200.200.200: bytes=56 Sequence=5 ttl=128 time=2 ms
```

```
--- ftp.com ping statistics ---
```

```
5 packet(s) transmitted
```

```
5 packet(s) received
```

```
0.00% packet loss
```

```
round-trip min/avg/max = 2/2/2 ms
```

The routing configuration between the switch and the domain name sever is omitted here, and refer to the related chapter for the configuration.

14.6 Troubleshooting Domain Name Resolution Configuration

Fault: Domain name resolution fails.

Troubleshoot: Perform the following procedures:

- Check whether the domain name resolution function is enabled.
- Check whether the IP address of the domain name sever is correctly configured.
- Check whether there is a correct route between the domain name sever and the switch.
- Check whether there is network connection failure, such as network cable break, loose connection, and so on.

Chapter 15 IP Performance Configuration

15.1 Configuring IP Performance

IP performance configuration includes:

- Configuring TCP Attributes

15.1.1 Configuring TCP Attributes

TCP attributes that can be configured include:

- **synwait timer:** When sending the syn packets, TCP starts the synwait timer. If response packets are not received before synwait timeout, the TCP connection is terminated. The timeout of synwait timer ranges from 2 to 600 seconds and it is 75 seconds by default.
- **finwait timer:** When the TCP connection state turns from FIN_WAIT_1 to FIN_WAIT_2, finwait timer is started. If FIN packets are not received before finwait timer timeout, the TCP connection is terminated. The timeout of finwait timer ranges from 76 to 3600 seconds and it is 675 seconds by default.
- The receiving/sending buffer size of the connection-oriented socket is in the range from 1 to 32 KB and is 8 KB by default.

Perform the following configuration in System view.

Table 15-1 Configure TCP attributes

Operation	Command
Configure timeout time for the synwait timer in TCP	tcp timer syn-timeout <i>time-value</i>
Restore the default timeout time of the synwait timer	undo tcp timer syn-timeout
Configure timeout time for the FIN_WAIT_2 timer in TCP	tcp timer fin-timeout <i>time-value</i>
Restore the default timeout time of the FIN_WAIT_2 timer	undo tcp timer fin-timeout
Configure the socket receiving/sending buffer size of TCP	tcp window <i>window-size</i>
Restore the socket receiving/sending buffer size of TCP to default value	undo tcp window

15.2 Displaying and Debugging IP Performance

After the above configuration, execute the **display** command in any view to display the running of the IP performance configuration, and to verify the effect of the configuration. Execute the **reset** command in user view to clear IP, TCP and UDP statistics information.

Execute the **debugging** command to debug IP performance.

Table 15-2 Display and debug IP performance

Operation	Command
Display TCP connection state	display tcp status
Display TCP connection statistics data	display tcp statistics
Display UDP statistics information	display udp statistics
Display IP statistics information	display ip statistics
Display ICMP statistics information	display icmp statistics
Display the current socket information of the system	display ip socket [socktype <i>sock-type</i>] [<i>task-id</i> <i>socket-id</i>]
Display the summary of the Forwarding Information Base (FIB)	display fib
Display the FIB entries matching the destination IP address (range)	display fib <i>ip_address1</i> [{ <i>mask1</i> <i>mask-length1</i> } [<i>ip_address2</i> { <i>mask2</i> <i>mask-length2</i> } longer] longer]
Display the FIB entries matching a specific ACL	display fib acl { <i>number</i> <i>name</i> }
Display the FIB entries which are output from the buffer according to regular expression and related to the specific character string	display fib { { begin include exclude } <i>text</i> }
Display the FIB entries matching the specific prefix list	display fib ip-prefix <i>listname</i>
Display the total number of FIB entries	display fib statistics
Reset IP statistics information	reset ip statistics
Reset TCP statistics information	reset tcp statistics
Reset UDP statistics information	reset udp statistics
Enable the debugging of IP packets	debugging ip packet [acl <i>acl-number</i>]
Disable the debugging of IP packets	undo debugging ip packet
Enable the debugging of ICMP packets	debugging ip icmp
Disable the debugging of ICMP packets	undo debugging ip icmp

Operation	Command
Enable the debugging of UDP connections	debugging udp packet [<i>task-id</i> <i>socket-id</i>]
Disable the debugging of UDP connections	undo debugging udp packet [<i>task-id</i> <i>socket-id</i>]
Enable the debugging of TCP connections	debugging tcp packet [<i>task-id</i> <i>socket-id</i>]
Disable the debugging of TCP connections	undo debugging tcp packet [<i>task-id</i> <i>socket-id</i>]
Enable the debugging of TCP events	debugging tcp event [<i>task-id</i> <i>socket-id</i>]
Disable the debugging of TCP events	undo debugging tcp event [<i>task-id</i> <i>socket-id</i>]
Enable the debugging of the MD5 authentication	debugging tcp md5
Disable the debugging of the MD5 authentication	undo debugging md5

15.3 Troubleshooting IP Performance

Fault: IP layer protocol works normally but TCP and UDP cannot work normally.

Troubleshoot: In the event of such a fault, you can enable the corresponding debugging information output to view the debugging information.

- Use the **display** command to view the running information of IP performance and make sure that the PCs used by the user is running normally.
- Use the **terminal debugging** command to output the debugging information to the console.
- Use the **debugging udp packet** command to enable the UDP debugging to trace the UDP packet.

The following are the UDP packet formats:

```
UDP output packet:
Source IP address:202.38.160.1
Source port:1024
Destination IP Address 202.38.160.1
Destination port: 4296
task = ROUT(15)
socketid = 6,
src = 192.168.1.1:520,
dst = 255.255.255.255:520,
datalen = 24
```

- Use the **debugging tcp packet** command to enable the TCP debugging to trace the TCP packets.

Operations include:

```
<SW8800> terminal debugging
<SW8800> debugging tcp packet
```

Then the TCP packets received or sent can be checked in real time. Specific packet formats include:

```
TCP output packet:
Source IP address:202.38.160.1
Source port:1024
Destination IP Address 202.38.160.1
Destination port: 4296
Sequence number :4185089
Ack number: 0
Flag :SYN
Packet length :60
Data offset: 10
task = ROUT(15)
socketid = 5
state = Established
src = 172.16.1.2
Source port:1025
dst = 172.16.1.1
Destination port: 4296
seq = 1921836502
ack = 4192768493
flag = ACK
window = 16079
```

Chapter 16 IP Routing Protocol Overview

Note:

A router that is referred to in the following or its icon represents a generalized router or a Switch 8800 running routing protocols. To improve readability, this will not be described in the other parts of the manual.

For the configuration of VPN instance, refer to the MPLS chapter in this book.

16.1 Introduction to IP Route and Routing Table

16.1.1 IP Route and Route Segment

Routers are implemented for route selection in the Internet. A router works in the following way: The router selects an appropriate path (through a network) according to the destination address of the packet it receives and forwards the packet to the next router. The last router in the path is responsible for submitting the packet to the destination host.

In Figure 16-1, R stands for a router. A packet sent from Host A to Host C should go through two routers and the packet is transmitted through two hops. Therefore, when a node (router) is connected to another node through a network, they are in the same route segment and are deemed as adjacent in the Internet. That is, the adjacent routers refer to two routers connected to the same network. The number of route segments between a router and hosts in the same network counted as zero. In Figure 16-1, the bold arrows represent these route segments. Which physical links comprise which route segment is not a concern of a router however.

with the mask 255.255.0.0 is located will be 129.102.0.0. It is made up of several consecutive "1"s, which can also be expressed in the dotted decimal format.

- Output interface: It indicates an interface through which an IP packet should be forwarded.
- Next hop address: It indicates the next router that an IP packet will pass through.
- Priority added to the IP routing table for a route: There may be different next hops to the same destination. These routes may be discovered by different routing protocols, or they can just be the static routes configured manually. The one with the highest priority (the smallest numerical value) will be selected as the current optimal route.
- Path cost: Cost to forward data by the route.

According to different destinations, the routes can be divided into:

- Subnet route: The destination is a subnet.
- Host route: The destination is a host

In addition, according to whether the network of the destination host is directly connected to the router, there are the following types of routes:

- Direct route: The router is directly connected to the network where the destination resides.
- Indirect route: The router is not directly connected to the network where the destination resides.

In order to limit the size of the routing table, an option is available to set a default route. All the packets that fail to find the suitable entry will be forwarded through this default route.

In a complicated Internet as shown in Figure 16-2, the number in each network is the network address, and R stands for a router. The router R8 is connected with three networks, so it has three IP addresses and three physical ports, and its routing table is shown in the diagram below:

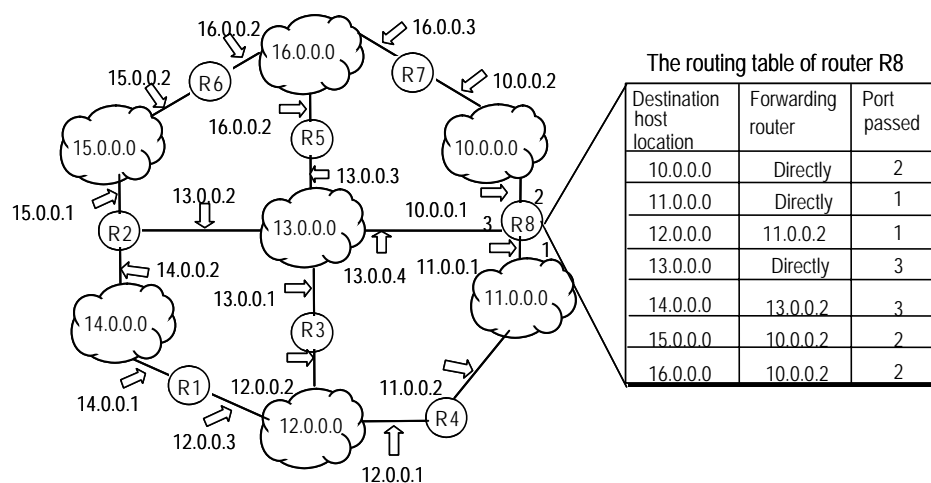


Figure 16-2 The routing table

The Switch 8800 supports the configuration of a series of dynamic routing protocols such as RIP, OSPF, IS-IS and BGP, as well as the static routes. In addition, the running switch will automatically obtain some direct routes according to the port state and user configuration.

16.2 Routing Management Policy

For the Switch 8800, you can configure manually the static route to a specific destination, and configure dynamic routing protocol to interact with other routers on the network. The routing algorithm can also be used to discover routes. For the configured static routes and dynamic routes discovered by the routing protocol, the Switch 8800 implements unified management. That is, the static routes configured by the user are managed together with the dynamic routes discovered by the routing protocol. The static routes and the routes learned or configured by different routing protocols can also be shared with each other.

16.2.1 Routing Protocols and the Preferences of the Corresponding Routes

Different routing protocols (as well as the static configuration) may generate different routes to the same destination, but not all these routes are optimal. In fact, at a certain moment, only one routing protocol can determine a current route to a specific destination. Thus, each of these routing protocols (including the static configuration) is set with a preference, and when there are multiple routing information sources, the route discovered by the routing protocol with the highest preference will become the current route. Routing protocols and the default preferences (the smaller the value is, the higher the preference is) of the routes learned by them are shown in Table 16-1.

In the table, 0 indicates a direct route. 255 indicates any route from unreliable sources.

Table 16-1 Routing protocols and the default preferences for the routes learned by them

Routing protocol or route type	The preference of the corresponding route
DIRECT	0
OSPF	10
IS-IS	15
STATIC	60
RIP	100
OSPF ASE	150
OSPF NSSA	150
IBGP	256
EBGP	256

Routing protocol or route type	The preference of the corresponding route
UNKNOWN	255

Apart from direct routing, IBGP and EBGP, the preferences of various dynamic routing protocols can be manually configured to meet the user requirements. In addition, the preferences for individual static routes can be different.

16.2.2 Supporting Load Sharing and Route Backup

I. Load sharing

The Switch 8800 supports static equivalent route, permitting to configure multiple routes that reach the same destination and use the same precedence. After you configured static equivalent routes, a packet can reach the same destination through multiple different paths, whose precedence levels are equal. When there is no route that can reach the same destination with a higher precedence, the multiple routes will be adopted. Thus, the router will forward the packets to the destination through these paths according to a certain algorithm so as to implement load sharing.

For the same destination, a specified routing protocol may find multiple different routes with the same precedence and different next hops. If the routing protocol has the highest precedence among all active routing protocols, these multiple routes will be regarded as currently valid routes. Thus, load sharing of IP traffic is ensured in terms of routing protocols.

The Switch 8800 supports eight routes to implement load sharing.

II. Route backup

The Switch 8800 supports route backup. When the main route fails, the system will automatically switch to a backup route to improve the network reliability.

In order to achieve static route backup, the user can configure multiple routes to the same destination according to actual situations. One of the routes has the highest precedence and is called as main route. The other routes have descending precedence levels and are called as backup routes. Normally, the router sends data via main route. When the line fails, the main route will hide itself and the router will choose one from the left routes as a backup route whose precedence is higher than others' to send data. In this way, the switchover from the main route to the backup route is implemented. When the main route recovers, the router will restore it and re-select route. As the main route has the highest precedence, the router still chooses the main route to send data. This process is the automatic switchover from the backup route to the main route.

16.2.3 Routes Shared Between Routing Protocols

As the algorithms of various routing protocols are different, different protocols may generate different routes, thus bringing about the problem of how to resolve the differences when different routes are generated by different routing protocols. The Switch 8800 can import the information of another routing protocol. Each protocol has its own route importing mechanism. For details, refer to the description about "Importing an External Route" in the operation manual of the corresponding routing protocol.

Chapter 17 Static Route Configuration

17.1 Introduction to Static Route

17.1.1 Static Route

A static route is a special route configured manually by an administrator. You can set up an interconnecting network with the static route configuration. The problem for such configuration is when a fault occurs to the network, the static route cannot change automatically to steer away from the node causing the fault, if without the help of an administrator.

In a relatively simple network, you only need to configure the static routes to make the router work normally. The proper configuration and usage of the static route can improve the network performance and ensure the bandwidth of the important applications.

All the following routes are static routes:

- Reachable route: A normal route is of this type. That is, the IP packet is sent to the next hop via the route marked by the destination. It is a common type of static routes.
- Unreachable route: When a static route to a destination has the "**reject**" attribute, all the IP packets to this destination will be discarded, and the source host will be informed that the destination is unreachable.
- Blackhole route: If a static route to a destination has the "**blackhole**" attribute, the outgoing interface of this route is the Null 0 interface regardless of the next hop address, and any IP packets addressed to this destination are dropped without notifying the source host.

The attributes "**reject**" and "**blackhole**" are usually used to control the range of reachable destinations of this router, and help troubleshooting the network.

17.1.2 Default Route

A default route is a special route. You can configure a default route using a static route. Some dynamic routing protocols can also generate default routes, such as OSPF and IS-IS.

In brief, a default route is used only when no suitable routing table entry is matched. That is, when no proper route is found, the default route is used. In a routing table, the default route is in the form of the route to the network 0.0.0.0 (with the mask 0.0.0.0). You can see whether the default route has been set by executing the **display ip routing-table** command. If the destination address of a packet fails in matching any entry of the routing table, the router will select the default route to forward this packet. If

there is no default route and the destination address of the packet fails in matching any entry in the routing table, this packet will be discarded, and an internet control message protocol (ICMP) packet will be sent to the originating host to inform that the destination host or network is unreachable.

17.2 Configuring Static Route

Static Route Configuration includes:

- Configuring a Static Route
- Configuring a Default Route
- Deleting All the Static Routes

17.2.1 Configuring a Static Route

Perform the following configurations in system view.

Table 17-1 Configure a static route

Operation	Command
Add a static route	ip route-static [vpn-instance <i>vpn-instance-name</i>]* <i>ip-address</i> { <i>mask</i> <i>mask-length</i> } { <i>interface-type</i> <i>interface-number</i> [vpn-instance <i>vpn-instance-name</i>] <i>gateway-address</i> } [preference <i>preference-value</i>] [reject blackhole]
Delete a static route	undo ip route-static [vpn-instance <i>vpn-instance-name</i>]* <i>ip-address</i> { <i>mask</i> <i>mask-length</i> } { <i>interface-type</i> <i>interface-number</i> [vpn-instance <i>vpn-instance-name</i>] <i>gateway-address</i> } [preference <i>preference-value</i>] [reject blackhole]

The parameters are explained as follows:

- IP address and mask

The IP address and mask are in a dotted decimal format. As "1"s in the 32-bit mask is required to be consecutive, the dotted decimal mask can also be replaced by the *mask-length* (which refers to the digits of the consecutive "1"s in the mask).

- Next hop address and NULL interface

When configuring a static route, you can specify the *gateway-address* to decide the next hop address, depending on the actual conditions.

In fact, for all the routing entries, the next hop address must be specified. When IP layer transmits an IP packet, it will first search the matching route in the routing table according to the destination address of the packet. Only when the next hop address of the route is specified can the link layer find the corresponding link layer address, and then forward the packet according to this address.

The packets sent to NULL interface, a kind of virtual interface, will be discarded at once. This can decrease the system load.

- Preference

Depending on the configuration of preference, you can achieve different route management policies. For example, to implement load sharing, you can specify the same preference for multiple routes to the same destination network. To implement route backup, you can specify different preferences for them.

- Other parameters

The attributes **reject** and **blackhole** respectively indicate the unreachable route and the blackhole route.

17.2.2 Configuring a Default Route

Perform the following configurations in system view.

Table 17-2 Configure a default route

Operation	Command
Configure a default route	ip route-static 0.0.0.0 { 0.0.0.0 0 } { <i>interface-type interface-number</i> <i>gateway-address</i> } [preference value] [reject blackhole]
Delete a default route	undo ip route-static 0.0.0.0 { 0.0.0.0 0 } [<i>interface-type interface-number</i> <i>gateway-address</i>] [preference value]

The meanings of parameters in the command are the same as those of the static route.

17.2.3 Deleting All the Static Routes

You can use the **undo ip route-static** command to delete one static route. The Switch 8800 also provides a special command for you to delete all static routes at one time, including the default routes.

Perform the following configuration in system view.

Table 17-3 Delete all static routes

Operation	Command
Delete all static routes	delete static-routes all
Delete all static routes of the VPN	delete vpn-instance <i>vpn-instance-name</i> static-routes all

17.3 Displaying and Debugging Static Route

After the above configuration, execute the **display** command in any view to display the running of the static route configuration, and to verify the effect of the configuration.

Table 17-4 Display and debug the routing table

Operation	Command
Display routing table summary	display ip routing-table
Display routing table details	display ip routing-table verbose
Display the detailed information of a specific route	display ip routing-table <i>ip_address</i> [<i>mask</i>] [longer-match] [verbose]
Display the route information in the specified address range	display ip routing-table <i>ip_address1 mask1 ip_address2 mask2</i> [verbose]
Display the route filtered through the specified basic access control list (ACL)	display ip routing-table acl { <i>acl-number</i> <i>acl-name</i> } [verbose]
Display the route information that is filtered through the specified ip prefix list	display ip routing-table ip-prefix <i>ip-prefix-number</i> [verbose]
Display the routing information discovered by the specified protocol	display ip routing-table protocol <i>protocol</i> [inactive verbose]
Display the tree routing table	display ip routing-table radix
Display the statistics of the routing table	display ip routing-table statistics
Display the routing information about the VPN instance	display ip routing-table vpn-instance <i>vpn-instance-name</i>

17.4 Typical Static Route Configuration Example

I. Network requirements

As shown in Figure 17-1, the masks of all the IP addresses are 255.255.255.0. It is required that all the hosts or the Switch 8800 can be interconnected in pairs by static route configuration.

II. Network diagram

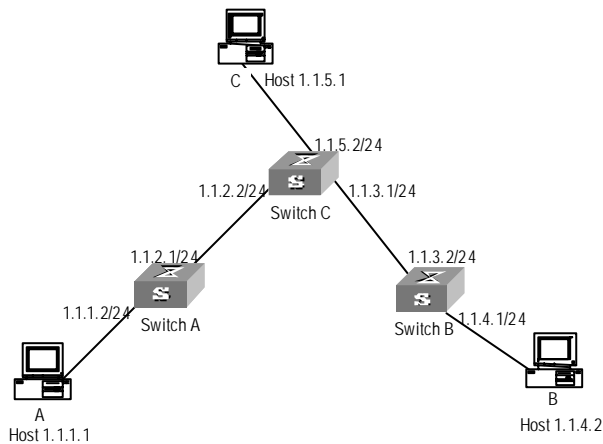


Figure 17-1 Network diagram for the static route configuration example

III. Configuration procedure

Configure the static route for Switch A

```
[Switch A] ip route-static 1.1.3.0 255.255.255.0 1.1.2.2
[Switch A] ip route-static 1.1.4.0 255.255.255.0 1.1.2.2
[Switch A] ip route-static 1.1.5.0 255.255.255.0 1.1.2.2
```

Configure the static route for Switch B

```
[Switch B] ip route-static 1.1.2.0 255.255.255.0 1.1.3.1
[Switch B] ip route-static 1.1.5.0 255.255.255.0 1.1.3.1
[Switch B] ip route-static 1.1.1.0 255.255.255.0 1.1.3.1
```

Configure the static route for Switch C

```
[Switch C] ip route-static 1.1.1.0 255.255.255.0 1.1.2.1
[Switch C] ip route-static 1.1.4.0 255.255.255.0 1.1.3.2
```

Configure the default gateway of the Host A to be 1.1.1.2

Configure the default gateway of the Host B to be 1.1.4.1

Configure the default gateway of the Host C to be 1.1.5.2

Then, all the hosts or switches in the figure can be interconnected in pairs.

17.5 Troubleshooting Static Route Faults

Symptom:

The switch is not configured with the dynamic routing protocol and both the physical status and the link layer protocol status of the interface is UP, but the IP packets cannot be forwarded normally.

Solution:

- Use the **display ip routing-table protocol static** command to view whether the configured static route is correct and in effect.

Chapter 18 RIP Configuration

18.1 Introduction to RIP

Routing Information Protocol (RIP) is a relatively simple interior gateway protocol (IGP), which is mainly applied to small scale networks.

It is easy to implement RIP. You can configure and maintain RIP more easily than OSPF and IS-IS, so RIP still has a wide application in actual networking.

18.1.1 RIP Operation Mechanism

I. RIP basic concepts

RIP is a kind of Distance-Vector (D-V) algorithm-based protocol and exchanges routing information via UDP packets.

It employs Hop Count to measure the distance to the destination host, which is called Routing Cost. In RIP, the hop count from a router to its directly connected network is 0, and that to a network which can be reached through another router is 1, and so on. To restrict the time to converge, RIP prescribes that the cost value is an integer ranging from 0 to 15. The hop count equal to or exceeding 16 is defined as infinite, that is, the destination network or the host is unreachable.

To improve the performance and avoid route loop, RIP supports Split Horizon and allows importing the routes discovered by other routing protocols.

II. RIP route database

Each router running RIP manages a route database, which contains routing entries to all the reachable destinations in the network. These routing entries contain the following information:

- Destination address: IP address of a host or a network.
- Next hop address: The interface address of the next router that an IP packet will pass through for reaching the destination.
- Output interface: The interface through which the IP packet should be forwarded.
- Cost: The cost for the router to reach the destination, which should be an integer in the range of 0 to 16.
- Timer: Duration from the last time that the routing entry is modified till now. The timer is reset to 0 whenever a routing entry is modified.

III. RIP timer

In RFC1058, RIP is controlled by the following timers: Period update, Timeout and Garbage-Collection.

- Period update is triggered periodically to send all RIP routes to all neighbors.
- If the RIP route is not updated (a router receives the update packets from the neighbor) when the Timeout timer expires, this route is regarded as unreachable. The cost is set to 16.
- If the Garbage-Collection timer expires, and the unreachable route receives no update packet from the same neighbor, the route will be completely deleted from the routing table.
- By default, the values of Period Update and Timeout timers are 30 seconds and 180 seconds respectively. The value of Garbage-collection timer is four times that of Period Update timer: 120 seconds.

18.1.2 RIP Enabling and Running

The following section describes the procedure:

- If RIP is enabled on a router for the first time, the router will broadcast or multicast the request packet to the adjacent routers. Upon receiving the request packet, the RIP on each adjacent router responds with a packet conveying its local routing table.
- After receiving the response packets, the router, which has sent the request, will modify its own routing table. At the same time, the router sends trigger modification packets to its adjacent routers running RIP and broadcasts modification information, following split horizon mechanism. After receiving trigger modification packets, the adjacent routers send trigger modification packets to their respective adjacent routers. As a result, each router can obtain and maintain the latest routing information.
- RIP broadcasts its routing table to the adjacent routers every 30 seconds. The adjacent routers will maintain their own routing table after receiving the packets and will select an optimal route, and then advertise the modification information to their respective adjacent network so as to make the updated route globally known. Furthermore, RIP uses the timeout mechanism to handle the out-timed routes so as to ensure the real-timeliness and validity of the routes.

RIP has become one of the actual standards of transmitting router and host routes by far. It can be used in most of the campus networks and the regional networks that are simple yet extensive. For larger and more complicated networks, RIP is not recommended.

18.2 Configuring RIP

1) RIP basic configuration

RIP basic configuration includes:

- Enabling RIP
- Enabling RIP on specified network

If the link, which does not support broadcast or multicast packets, runs RIP, you need to configure RIP to send any packet to the specified destination, establishing RIP neighbors correctly.

In NBMA link networking through a Frame Relay sub-interface and others, to ensure the routing information can be correctly transmitted, you possibly need to disable split horizon.

2) RIP route management

You can make the following configurations for RIP to advertise and receive routing information:

- Setting additional routing metric
 - Configuring RIP to import routers of other protocols
 - Configuring RIP route filtering
 - Disabling host route
 - Disabling RIP route aggregation
- ## 3) RIP configuration
- Configuring the RIP precedence
 - Configuring RIP timers
 - Configuring RIP-1 zero field check of the interface packet
 - Specifying RIP version of the interface
- ## 4) Configuration related to security

You can select the following configurations to improve RIP security during exchanging routing information, or control the area to transmit RIP packets.

- Setting RIP-2 packet authentication
- Specifying the operating state of the interface

18.2.1 Enabling RIP and Entering RIP View

Perform the following configurations in system view.

Table 18-1 Enable RIP and enter RIP view

Operation	Command
Enable RIP and enter the RIP view	rip
Disable RIP	undo rip

By default, RIP is not enabled.

18.2.2 Enabling RIP on the Specified Network Segment

To flexibly control RIP operation, you can enable RIP on the specified network segment so that the corresponding ports can receive and send RIP packets.

Perform the following configurations in RIP view.

Table 18-2 Enable RIP Interface

Operation	Command
Enable RIP on the specified network	network <i>network-address</i>
Disable RIP on the specified network	undo network <i>network-address</i>

Note that after the RIP task is enabled, you should also specify its operating network segment, for RIP only operates on the interface on the specified network segment. For an interface that is not on the specified network segment, RIP does not receive or send routes on it, nor forwards its interface route, as if this interface does not exist at all. *network-address* is the address of the enabled or disabled network, and it can also be configured as the IP network address of respective interfaces.

When a command **network** is used for an address, you can enable the network address of the port, which also includes the subnet addresses. For example, for **network** 129.102.1.1, you can see **network** 129.102.0.0 either using **display current-configuration** or using **display rip** command.

By default, RIP is disabled on all the interfaces after it is started up.

18.2.3 Configuring Unicast of the Packets

Usually, RIP sends packets using broadcast or multicast addresses. It exchanges routing information with non-broadcasting networks in unicast mode.

Perform the following configuration in RIP view.

Table 18-3 Configure unicast of the packets

Operation	Command
Configure unicast of the packets	peer <i>ip-address</i>
Cancel unicast of the packets	undo peer <i>ip-address</i>

By default, RIP does not send any packets to any unicast addresses.

It should be noted that **peer** should also be restricted by **rip work**, **rip output**, **rip input** and **network**.

18.2.4 Configuring Split Horizon

Split horizon means that the route received via an interface will not be sent via this interface again. To some extent, the split horizon is necessary for reducing routing loop. But in some special cases, split horizon must be disabled so as to ensure the correct advertisement of the routes at the cost of efficiency. For example, split horizon is disabled on a NBMA network if it runs RIP.

Perform the following configuration in interface view.

Table 18-4 Configure Split Horizon

Operation	Command
Enable split horizon	rip split-horizon
Disable split horizon	undo rip split-horizon

By default, split horizon of the interface is enabled.

18.2.5 Setting Additional Routing Metric

Additional routing metric is the input or output routing metric added to an RIP route. It does not change the metric value of the route in the routing table, but adds a specified metric value when the interface receives or sends a route.

Perform the following configuration in interface view.

Table 18-5 Set additional routing metric

Operation	Command
Set the additional routing metric of the route when the interface receives an RIP packet	rip metricin <i>value</i>
Disable the additional routing metric of the route when the interface receives an RIP packet	undo rip metricin
Set the additional routing metric of the route when the interface sends an RIP packet	rip metricout <i>value</i>
Disable the additional routing metric of the route when the interface sends an RIP packet	undo rip metricout

By default, the additional routing metric added to the route when RIP sends a packet is 1. The additional routing metric when RIP receives the packet is 0 by default.

Note:

The metricout configuration takes effect only on the RIP routes learnt by the router and RIP routes generated by the router itself. That is, it has no effect on the routes imported to RIP by other routing protocols.

18.2.6 Configuring RIP to Import Routes of Other Protocols

RIP allows users to import the route information of other protocols into the RIP routing table.

RIP can import the routes of Direct, Static, OSPF, IS-IS and BGP, etc.

Perform the following configuration in RIP view.

Table 18-6 Configure RIP to import routes of other protocols

Operation	Command
Configure RIP to import routes of other protocols	import-route <i>protocol</i> [cost <i>value</i> route-policy <i>route-policy-name</i>]*
Cancel the imported routing information of other protocols	undo import-route <i>protocol</i>
Set the default routing metric	default cost <i>value</i>
Restore the default routing metric	undo default cost

By default, RIP does not import the route information of other protocols.

If you do not specify the routing metric when importing a route, the default routing metric 1 is used.

18.2.7 Configuring Route Filtering

The router provides the route filtering function. You can configure the filter policy rules through specifying the ACL and ip-prefix for route import and advertisement. Besides, to import a route, the RIP packet of a specific router can also be received by designating a neighbor router.

Perform the following configuration in RIP view.

I. Configuring RIP to filter the received routes

Table 18-7 Configure RIP to filter the received routes

Operation	Command
Configure RIP to filter the received routing information advertised by the specified address	filter-policy gateway <i>ip-prefix-name</i> import
Cancel filtering the received routing information advertised by the specified address	undo filter-policy gateway <i>ip-prefix-name</i> import
Configure RIP to filter the received global routing information	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } import
Cancel filtering the received global routing information	undo filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } import

II. Configuring RIP to filter the routes advertised by RIP

Table 18-8 Configure RIP to filter the advertised routes

Operation	Command
Configure RIP to filter the advertised routing information	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } export [<i>routing-protocol</i>]
Cancel filtering the advertised routing information	undo filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } export [<i>routing-protocol</i>]

By default, RIP does not filter the received and advertised routing information.

Note:

- The **filter-policy import** command filters the RIP routes received from its neighbors, and the routes that cannot pass the filter will not be added to the routing table, and will not be advertised to the neighbors.
 - The **filter-policy export** command filters all the advertised routes, including routes imported by the **import-route** command, and RIP routes learned from the neighbors.
 - If the **filter-policy export** command does not specify which route to be filtered, then all the routes imported by the **import-route** command and the advertised RIP routes will be filtered.
-

18.2.8 Disabling RIP to Receive Host Route

In some special cases, the router can receive a lot of host routes, and these routes are of little help in route addressing but consume a lot of network resources. Routers can be configured to reject host routes by using the **undo host-route** command.

Perform the following configuration in RIP view.

Table 18-9 Disable RIP to receive host route

Operation	Command
Enable receiving host route	host-route
Disable receiving host route	undo host-route

By default, the router receives the host route.

18.2.9 Enabling RIP-2 Route Aggregation Function

The so-called route aggregation means that different subnet routes in the same natural network can be aggregated into one natural mask route for transmission when they are sent to the outside (i.e. other network). Route aggregation can be performed to reduce the routing traffic on the network as well as to reduce the size of the routing table.

RIP-1 only sends the route with natural mask, that is, it always sends routes in the route aggregation form. RIP-2 supports subnet mask and classless interdomain routing. To advertise all the subnet routes, the route aggregation function of RIP-2 can be disabled. Perform the following configuration in RIP view.

Table 18-10 Enable RIP-2 route aggregation function

Operation	Command
Enable the route aggregation function of RIP-2	summary
Disable the route aggregation function of RIP-2	undo summary

By default, RIP-2 route summarization is enabled.

18.2.10 Setting the RIP Preference

Each kind of routing protocol has its own preference, by which the routing policy will select the optimal one from the routes of different protocols. The greater the preference value is, the lower the preference becomes. The preference of RIP can be set manually.

Perform the following configuration in RIP view.

Table 18-11 Set the RIP Preference

Operation	Command
Set the RIP Preference	preference <i>value</i>
Restore the default value of RIP preference	undo preference

By default, the preference of RIP is 100.

18.2.11 Specifying RIP Version of the Interface

RIP has two versions, RIP-1 and RIP-2. You can specify the version of the RIP packets processed by the interface.

RIP-1 broadcasts the packets. RIP-2 can transmit packets by both broadcast and multicast. By default, multicast is adopted for transmitting packets. In RIP-2, the multicast address is 224.0.0.9. The advantage of transmitting packets in the multicast mode is that the hosts not operating RIP in the same network can avoid receiving RIP

broadcast packets. In addition, this mode can also make the hosts running RIP-1 avoid incorrectly receiving and processing the routes with subnet mask in RIP-2. When an interface is running in RIP-2 broadcast mode, the RIP-1 packets can also be received.

Perform the following configuration in interface view:

Table 18-12 Specify RIP version of the interface

Operation	Command
Specify the RIP version as RIP-1 for the interface	rip version 1
Specify the RIP version as RIP-2 for the interface	rip version 2 [broadcast multicast]
Restore the default RIP version running on the interface	undo rip version

By default, the interface receives and sends the RIP-1 packets. It will transmit packets in multicast mode when the interface RIP version is set to RIP-2.

18.2.12 Configuring RIP Timers

As mentioned previously, RIP has three timers: Period update, Timeout and Garbage-collection. Modification of these timers affects RIP convergence speed.

Perform the following configuration in RIP view.

Table 18-13 Configure RIP timers

Operation	Command
Configure RIP timers	timers { update <i>update-timer-length</i> timeout <i>timeout-timer-length</i> } *
Restore the default settings of RIP timers	undo timers { update timeout } *

The modification of RIP timers is validated immediately.

By default, the values of Period Update and Timeout timers are 30 seconds and 180 seconds respectively. The value of Garbage-collection timer is four times that of Period Update timer: 120 seconds.

In fact, you may find that the timeout time of Garbage-collection timer is not fixed. If Period Update timer is set to 30 seconds, Garbage-collection timer might range from 90 to 120 seconds.

Before RIP completely deletes an unreachable route from the routing table, it advertises the route by sending four Period Update packets with route metric of 16, so as to acknowledge all the neighbors that the route is unreachable. As routes cannot

always become unreachable at the point when a new period starts, the actual value of Garbage-collection timer is three to four times that of Period Update timer.

Note:

You must consider network performance when adjusting RIP timers, and configure all the routers that are running RIP, so as to avoid unnecessary traffic or network jitter.

18.2.13 Configuring RIP-1 Zero Field Check of the Interface Packet

According to the RFC1058, some fields in the RIP-1 packet must be 0, and they are called zero fields. Therefore, when an interface version is set as RIP-1, the zero field check should be performed on the packet. But if the value in the zero field is not zero, processing will be refused. As there is no zero field in the RIP-2 packet, this configuration is invalid for RIP-2.

Perform the following configuration in RIP view.

Table 18-14 Configure zero field check of the interface packet

Operation	Command
Configure zero field check on the RIP-1 packet	checkzero
Disable zero field check on the RIP-1 packet	undo checkzero

By default, RIP-1 performs zero field check on the packet.

18.2.14 Specifying the Operating State of the Interface

In interface view, you can specify the operating state of RIP on the interface. For example, whether RIP operates on the interface, namely, whether RIP update packets are sent and received on the interface. In addition, whether an interface sends or receives RIP update packets can be specified separately.

Perform the following configuration in interface view.

Table 18-15 Specify the operating state of the interface

Operation	Command
Enable the interface to run RIP	rip work
Disable the interface to run RIP	undo rip work
Enable the interface to receive RIP update packet	rip input
Disable the interface to receive RIP update packet	undo rip input

Operation	Command
Enable the interface to send RIP update packet	rip output
Disable the interface to send RIP update packet	undo rip output

The **undo rip work** command and the **undo network** command have similar but not all the same functions. Neither of the two commands configures an interface to receive or send RIP route. The difference also exists. RIP still advertises the routes of the interface applying the **undo rip work** command. However, other interfaces will not forward the routes of the interface applying the **undo network** command. It seems that the interface is removed.

In addition, **rip work** is functionally equivalent to both **rip input** and **rip output** commands.

By default, all interfaces except loopback interfaces both receive and transmit RIP update packets.

18.2.15 Setting RIP-2 Packet Authentication

RIP-1 does not support packet authentication. But when the interface operates RIP-2, the packet authentication can be configured.

RIP-2 supports two authentication modes: Simple authentication and MD5 authentication. MD5 authentication uses two packet formats: One follows RFC1723 and the other follows the RFC2082.

The simple authentication does not ensure security. The authentication key not encrypted is sent together with the packet, so the simple authentication cannot be applied to the case with high security requirements.

Perform the following configuration in Interface view:

Table 18-16 Set RIP-2 packet authentication

Operation	Command
Configure RIP-2 simple authentication key	rip authentication-mode simple <i>password-string</i>
Perform usual MD5 authentication on RIP-2 packets	rip authentication-mode md5 usual <i>key-string</i>
Perform nonstandard-compatible MD5 authentication on RIP-2 packets	rip authentication-mode md5 nonstandard <i>key-string key-id</i>
Disable RIP-2 packet authentication	undo rip authentication-mode

Before configuring MD5 authentication, you must configure MD5 type. The **usual** packet format follows RFC1723 and the **nonstandard** follows RFC2082.

18.3 Displaying and Debugging RIP

After the above configuration, execute the **display** command in any view to display the running of the RIP configuration, and to verify the effect of the configuration. Execute the **debugging** command in user view to debug the RIP module. Execute the **reset** command in RIP view to reset the system configuration parameters of RIP.

Table 18-17 Display and debug RIP

Operation	Command
Display the current RIP running state and configuration information.	display rip
Enable the RIP packet debugging information	debugging rip packet
Disable the RIP packet debugging information	undo debugging rip packet
Enable the debugging of RIP receiving packets	debugging rip receive
Disable the debugging of RIP receiving packets	undo debugging rip receive
Enable the debugging of RIP sending packet	debugging rip send
Disable the debugging of RIP sending packet	undo debugging rip send
Reset the system configuration parameters of RIP	reset

18.4 Typical RIP Configuration Example

I. Network requirements

As shown in Figure 18-1, switch C connects to the subnet 117.102.0.0 through the Ethernet port. The Ethernet ports of switches A and Switch B are respectively connected to the network 155.10.1.0 and 196.38.165.0. Switch C, Switch A and Switch B are connected via Ethernet 110.11.2.0. Correctly configure RIP to ensure that Switch C, Switch A and Switch B can interconnect with each other.

II. Network diagram

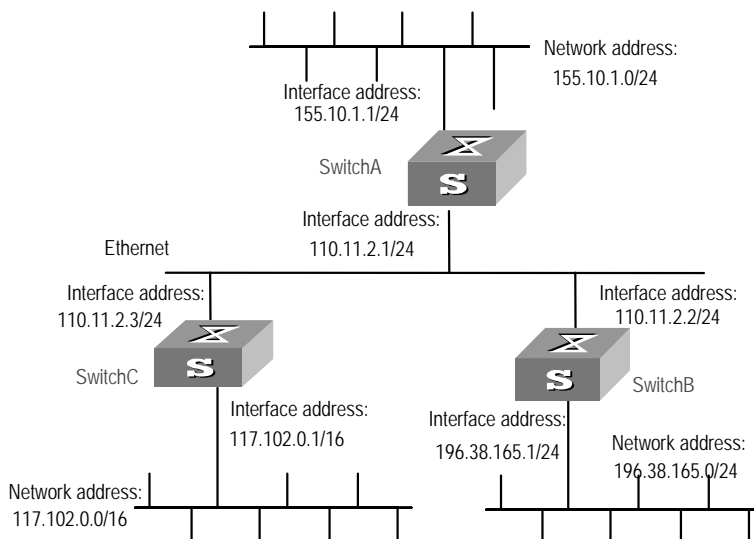


Figure 18-1 Network diagram for RIP configuration

III. Configuration procedure

Note:

The following configuration only shows the operations related to RIP. Before performing the following configuration, make sure the Ethernet link layer can work normally.

1) Configure Switch A

Configure RIP

```
[Switch A] rip
[Switch A-rip] network 110.11.2.0
[Switch A-rip] network 155.10.1.0
```

2) Configure Switch B

Configure RIP

```
[Switch B] rip
[Switch B-rip] network 196.38.165.0
[Switch B-rip] network 110.11.2.0
```

3) Configure Switch C

Configure RIP

```
[Switch C] rip
[Switch C-rip] network 117.102.0.0
```

```
[Switch C-rip] network 110.11.2.0
```

18.5 Troubleshooting RIP Faults

Symptom: The Switch 8800 cannot receive the update packets when the physical connection to the peer routing device is normal.

Solution: RIP does not operate on the corresponding interface (for example, the **undo rip work** command is executed) or this interface is not enabled through the **network** command. The peer routing device is configured to be in the multicast mode (for example, the **rip version 2 multicast** command is executed) but the multicast mode has not been configured on the corresponding interface of the local switch.

Chapter 19 OSPF Configuration

19.1 OSPF Overview

19.1.1 Introduction to OSPF

Open Shortest Path First (OSPF) is an Interior Gateway Protocol based on the link state developed by IETF. At present, OSPF version 2 (RFC2328) is used, which is available with the following features:

- **Applicable scope:** It can support networks in various sizes and can support several hundreds of routers at maximum.
- **Fast convergence:** It can transmit the update packets instantly after the network topology changes so that the change is synchronized in the AS.
- **Loop-free:** Since the OSPF calculates routes with the shortest path tree algorithm according to the collected link states, it is guaranteed that no loop routes will be generated from the algorithm itself.
- **Area partition:** It allows the network of AS to be divided into different areas for the convenience of management so that the routing information transmitted between the areas is abstracted further, hence to reduce the network bandwidth consumption.
- **Equal-cost multi-route:** Support multiple equal-cost routes to a destination.
- **Routing hierarchy:** OSPF has a four-level routing hierarchy. It prioritizes the routes to be intra-area, inter-area, external type-1, and external type-2 routes.
- **Authentication:** It supports the interface-based packet authentication so as to guarantee the security of the route calculation.
- **Multicast transmission:** Support multicast address to receive and send packets.

19.1.2 Process of OSPF Route Calculation

The routing calculation process of the OSPF protocol is as follows:

- Each OSPF-capable router maintains a Link State Database (LSDB), which describes the topology of the whole AS. According to the network topology around itself, each router generates a Link State Advertisement (LSA). The routers on the network transmit the LSAs among them by transmitting the protocol packets to each others. Thus, each router receives the LSAs of other routers and all these LSAs compose its LSDB.
- LSA describes the network topology around a router, so the LSDB describes the network topology of the whole network. Routers can easily transform the LSDB to a weighted directed graph, which actually reflects the topology architecture of the whole network. Obviously, all the routers get a graph exactly the same.

- A router uses the SPF algorithm to calculate the shortest path tree with itself as the root, which shows the routes to the nodes in the autonomous system. The external routing information is the leaf node. A router, which advertises the routes, also tags them and records the additional information of the autonomous system. Obviously, the routing tables obtained by different routers are different.

Furthermore, to enable individual routers to broadcast their local state information to the entire AS, any two routers in the environment should establish adjacency between them. In this case, however, the changes that any router takes will result in multiple transmissions, which are not only unnecessary but also waste the precious bandwidth resources. To solve this problem, “Designated Router” (DR) is defined in the OSPF. Thus, all the routers only send information to the DR for broadcasting the network link states in the network. Thereby, the number of router adjacent relations on the multi-access network is reduced.

OSPF supports interface-based packet authentication to guarantee the security of route calculation. Also, it transmits and receives packets by IP multicast (224.0.0.5 and 224.0.0.6).

19.1.3 OSPF Packets

OSPF uses five types of packets:

- Hello Packet:

It is the commonest packet, which is periodically sent by a router to its neighbor. It contains the values of some timers, DR, BDR and the known neighbor.

- Database Description (DD) Packet:

When two routers synchronize their databases, they use the DD packets to describe their own LSDBs, including the digest of each LSA. The digest refers to the HEAD of LSA, which uniquely identifies the LSA. This reduces the traffic size transmitted between the routers, since the HEAD of a LSA only occupies a small portion of the overall LSA traffic. With the HEAD, the peer router can judge whether it already has had the LSA.

- Link State Request (LSR) Packet:

After exchanging the DD packets, the two routers know which LSAs of the peer routers are lacked in the local LSDBs. In this case, they will send LSR packets requesting for the needed LSAs to the peers. The packets contain the digests of the needed LSAs.

- Link State Update (LSU) Packet:

The packet is used to transmit the needed LSAs to the peer router. It contains a collection of multiple LSAs (complete contents).

- Link State Acknowledgment (LSAck) Packet

The packet is used for acknowledging the received LSU packets. It contains the HEAD(s) of LSA(s) requiring acknowledgement.

19.1.4 LSA Type

I. Five basic LSA types

As mentioned previously, OSPF calculates and maintains routing information from LSAs. RFC2328 defines five LSA types as follows:

- Router-LSAs: Type-1. Each router generates Router-LSAs, which describe the link state and cost of the local router. Router-LSAs are broadcast within the area where the router is located.
- Network-LSAs: Type-2. DRs on the broadcast network and NBMA network generate Network-LSAs, which describe the link state of the local network. Network-LSAs are broadcast within the area where a DR is located.
- Summary-LSAs: Include Type-3 and Type-4. Area border routers (ABRs) generate Summary-LSAs. Summary-LSAs are broadcast within the area related to the LSA. Each Summary-LSA describes a route (inter-area route) to a certain destination in other areas of this AS. Type-3 Summary-LSAs describe the routes to networks (the destination is network). Type-4 Summary-LSAs describe the routes to autonomous system border routers (ASBRs).
- AS-external-LSAs: or ASE LSA, the Type-5. ASBRs generate AS-external-LSAs, which describe the routes to other ASs. AS-external-LSA packets are transmitted to the whole AS (except Stub areas). AS-external-LSAs can also describe the default route of an AS.

II. Type-7 LSA

RFC1587 (OSPF NSSA Option) adds a new LSA type: Type-7 LSAs.

According to RFC1587, Type-7 LSAs differ from Type-5 LSAs as follows:

- Type-7 LSAs are generated and released within a Not-So-Stubby Area (NSSA). Type-5 LSAs cannot be generated or released within a NSSA.
- Type-7 LSAs can only be released within an NSSA. When Type-7 LSAs reach an ABR, the ABR can convert part routing information of Type-7 LSAs into Type-5 LSAs and releases the information. Type-7 LSAs cannot be directly released to other areas or backbone areas.

19.1.5 Basic Concepts Related to OSPF

I. Router ID

To run OSPF, a router must have a router ID. If no ID is configured, the system will automatically pick an IP address from the IP addresses of the current interfaces as the Router ID. The following introduces how to choose a router ID. If loopback interface addresses exist, the system chooses the Loopback address with the greatest IP address value as the router ID. If no Loopback interface configured, then the address of the physical interface with the greatest IP address value will be the router ID.

II. DR and BDR

- Designated Router (DR)

In multi-access networks, if any two routers establish adjacencies, the same LSA will be transmitted repeatedly, wasting bandwidth resources. To solve this problem, the OSPF protocol regulates that a DR must be elected in a multi-access network and only the DR (and the BDR) can establish adjacencies with other routers in this network. Two non-DR routers or non-BDR routers cannot establish adjacencies and exchange routing information.

You cannot specify the DR in the segment. Instead, DR is elected by all the routers in the segment.

- Backup Designated Router (BDR)

If the DR fails for some faults, a new DR must be elected and synchronized with other routers on the segment. This process will take a relatively long time, during which, the route calculation is incorrect. To shorten the process, BDR is brought forth in OSPF. In fact, BDR is a backup for DR. DR and BDR are elected in the meantime. The adjacencies are also established between the BDR and all the routers on the segment, and routing information is also exchanged between them. After the existing DR fails, the BDR will become a DR immediately.

III. Area

The network size grows increasingly larger. If all the routers on a huge network are running OSPF, the large number of routers will result in an enormous LSDB, which will consume an enormous storage space, complicate the SPF algorithm, and add the CPU load as well. Furthermore, as a network grows larger, the topology becomes more likely to take changes. Hence, the network will always be in “turbulence”, and a great deal of OSPF packets will be generated and transmitted in the network. This will lower the network bandwidth utility. In addition, each change will cause all the routes on the network to recompute the route.

OSPF solves the above problem by partition an AS into different areas. Areas are logical groups of routers. The borders of areas are formed by routers. Thus, some routers may belong to different areas. A router connects the backbone area and a non-backbone area is called Area Border Router (ABR). An ABR can connect to the backbone area physically or logically.

IV. Backbone area and virtual link

- Backbone Area

After the area partition of OSPF, not all the areas are equal. In which, an area is different from all the other areas. Its area-id is 0 and it is usually called the backbone area.

- Virtual link

Since all the areas should be connected to the backbone area, virtual link is adopted so that the physically separated areas can still maintain the logic connectivity to the backbone area.

V. Route summary

An AS is divided into different areas that are interconnected via OSPF ABRs. The routing information between areas can be reduced through route summary. Thus, the size of routing table can be reduced and the calculation speed of the router can be improved. After calculating an intra-area route of an area, the ABR summarizes multiple OSPF routes into an LSA and sends it outside the area according to the configuration of summary.

For example, as shown in Figure 19-1, the Area 19 has three area intra-area routes: 19.1.1.0/24, 19.1.2.0/24 and 19.1.3.0/24. The three routes are summarized into one route 19.1.0.0/16 after you configured route summary. The RTA only generates an LSA, describing the summarized route.

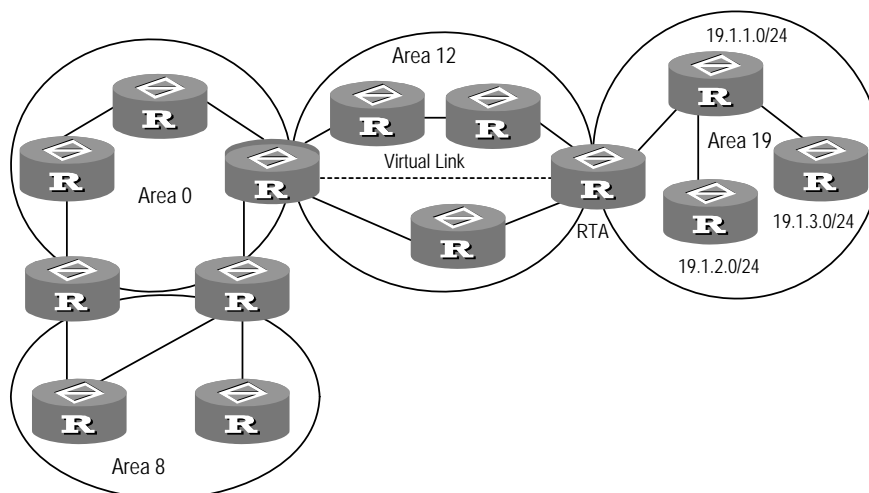


Figure 19-1 Area and route aggregation

19.1.6 OSPF Features Supported by the Switch 8800

The Switch 8800 supports the following OSPF features:

- Support stub areas: OSPF defines stub areas to decrease the overhead when the routers within the area receive ASE routes.
- Support NSSA: OSPF defines NSSA areas, surmounting the restriction of stub areas on topology. NSSA is the abbreviation of Not-So-Stubby Area.
- Support OSPF Multi-Process: A router runs multiple OSPF processes.
- Share the discovered routing information with other dynamic routing protocols: OSPF currently can import static routes and routes of other dynamic routing protocols such as RIP into the autonomous system of the router, or advertise the routing information discovered by OSPF to other routing protocols.

- Authenticator: OSPF provides clear text authenticator and MD5 encryption authenticator to authenticate packets transmitted between neighboring routers in the same area.
- Flexible configuration for the router port parameter: On the router port, you can configure the following OSPF parameters: output cost, Hello packet interval, retransmission interval, port transmission delay, route precedence, invalid time for adjacent routers, packet authentication mode, packet authenticator, and others.
- Virtual connection: Creates and configures virtual connections.
- Abundant debugging information: OSPF provides abundant debugging information, consequently helping users to diagnose failure.

19.2 Configuring OSPF

OSPF configuration needs cooperation among routers: intra-area, area boundary, and AS boundary. If none of OSPF parameters is configured, their default settings apply. In this case, sent and received packets are not authenticated, and an individual interface does not belong to the area of any AS. When reconfiguring a default parameter on one router, make sure that the same change is made on all other involved routers.

In various configurations, you must first enable OSPF, specify the interface and area ID before configuring other functions. But the configuration of the functions related to the interface is not restricted by whether the OSPF is enabled or not. It should be noted that after OSPF is disabled, the OSPF-related interface parameters also become invalid.

OSPF configuration includes:

- 1) OSPF basic configuration
 - Configuring Router ID
 - Enabling OSPF
 - Entering the OSPF area view
 - Enabling OSPF on the specified network
- 2) Configuration related to OSPF route
 - Configuring OSPF to import routes of other protocols
 - Configuring OSPF to import default routes
 - Configuring OSPF route filtering
 - Configuring OSPF route convergence
- 3) Some OSPF configurations
 - Configuring OSPF precedence
 - Setting the interface priority for DR election
 - Configuring OSPF timers
 - Configuring the time for the interface to send LSUs
 - Configuring the cost for sending packets on an interface
 - Configuring cost value for next hop
 - Configuring the network type on the OSPF interface
 - Configuring NBMA neighbors for OSPF

- Configuring to fill the MTU field when an interface transmits DD packets
- Setting an SPF calculation interval for OSPF
- 4) Configurations related to OSPF networking
 - Configuring OSPF authentication
 - Prohibit OSPF packet receiving/sending
 - Configuring OSPF virtual link
 - Configuring Stub area of OSPF
 - Configuring NSSA of OSPF
- 5) Configuration related to specific applications
 - Configuring OSPF and network management system
- 6) Others
 - Resetting the OSPF process

19.2.1 Configuring Router ID

Router ID is a 32-bit unsigned integer in IP address format that uniquely identifies a router within an AS. Router ID can be configured manually. If router ID is not configured, the system will select the IP address of an interface automatically. When you do that manually, you must guarantee that the IDs of any two routers in the AS are unique. A common undertaking is to set the router ID to be the IP address of an interface on the router.

Perform the following configuration in system view.

Table 19-1 Configure router ID

Operation	Command
Configure router ID	router id <i>router-id</i>
Remove the router ID	undo router id

To ensure stability of OSPF, the user should determine the division of router IDs and manually configure them when planning the network.

19.2.2 Enabling OSPF

Perform the following configuration in system view.

Table 19-2 Enable/Disable OSPF

Operation	Command
Enable OSPF and enter OSPF view	ospf [<i>process-id</i> [router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>]]
Disable one or all OSPF processes	undo ospf [<i>process-id</i>]

By default, OSPF is disabled.

When enabling OSPF, pay attention to the following points:

- The default OSPF process ID is 1. If no process ID is specified in the command, the default one is adopted.
- If a router is running multiple OSPF processes, you are recommended to use **router-id** in the command to specify different router IDs for different processes.

19.2.3 Entering OSPF Area View

OSPF divides an AS into different areas or logical groups of routers.

Perform the following configuration in OSPF view.

Table 19-3 Enter OSPF area view

Operation	Command
Enter OSPF area view	area <i>area-id</i>
Delete an OSPF area	undo area <i>area-id</i>

The *area-id* parameter identifies an area. It can be a decimal integer in the range of 0 to 4,294,967,295, or in the format of IP address. Regardless of how it is specified, it is displayed in the format of IP address.

Note that when you configure OSPF routers in the same area, you should apply most configuration data to the whole area. Otherwise, the neighboring routers cannot exchange information. This may even block routing information or create routing loops.

19.2.4 Specifying an Interface to Run OSPF

After using the **ospf** command to enable OSPF in system view, you must specify the network to run OSPF. An ABR router can be in different areas, while a network segment can only belong to an area. That is, you must specify a specific area for each port running OSPF.

Perform the following configuration in OSPF area view.

Table 19-4 Specifying an interface to run OSPF

Operation	Command
Specify an interface to run OSPF	network <i>ip-address ip-mask</i>
Disable OSPF on the interface	undo network <i>ip-address ip-mask</i>

The *ip-mask* argument is IP address wildcard shielded text (similar to the complement of the IP address mask).

19.2.5 Configuring OSPF to Import Routes of Other Protocols

The dynamic routing protocols on the router can share the routing information. As far as OSPF is concerned, the routes discovered by other routing protocols are always processed as the external routes of AS. In the **import-route** commands, you can specify the route cost type, cost value and tag to overwrite the default route receipt parameters (refer to “Configuring parameters for OSPF to import external routes”).

The OSPF uses the following four types of routes (ordered by priority):

- Intra-area route
- Inter-area route
- External route type 1
- External route type 2

Intra-area and inter-area routes describe the internal AS topology whereas the external routes describe how to select the route to the destinations beyond the AS.

The external routes type-1 refers to the imported IGP routes (such as static route and RIP). Since these routes are more reliable, the calculated cost of the external routes is the same as the cost of routes within the AS. Also, such route cost and the route cost of the OSPF itself are comparable. That is, cost to reach the external route type 1 = cost to reach the corresponding ASBR from the local router + cost to reach the destination address of the route from the ASBR.

The external routes type-2 refers to the imported EGP routes. Since these routes have lower credibility, OSPF assumes that the cost spent from the ASBR to reach the destinations beyond the AS is greatly higher than that spent from within the AS to the ASBR. So in route cost calculation, the former is mainly considered, that is, the cost spent to reach the external route type 2 = cost spent to the destination address of the route from the ASBR. If the two values are equal, then the cost of the router to the corresponding ASBR will be considered.

I. Configuring OSPF to import external routes

Perform the following configuration in OSPF view.

Table 19-5 Configure OSPF to import external routes

Operation	Command
Configure OSPF to import routes of other protocols	import-route <i>protocol</i> [cost <i>value</i> type <i>value</i> tag <i>value</i> route-policy <i>route-policy-name</i>]*
Cancel importing routing information of other protocols	undo import-route <i>protocol</i>

By default, OSPF will not import the routing information of other protocols. For a imported route, type is 2, cost is 1, and tag is 1 by default.

The routes that can be imported include Direct, Static, RIP, IS-IS, or BGP and in addition, the routes of other OSPF processes.

Note:

- It is recommended to configure the imported route type, cost and tag for the **import-route** command simultaneously. Otherwise, the later configuration will overwrite the former configuration.
 - After you configured the **import-route** command on the OSPF router to import external routing information, this OSPF router becomes an ASBR.
-

II. Configuring parameters for OSPF to import external routes

When the OSPF imports the routing information discovered by other routing protocols in the autonomous system, some additional parameters need configuring, such as default route cost and default tag of route distribution. Route tag can be used to identify the protocol-related information. For example, OSPF can use it to identify the AS number when receiving BGP.

Perform the following configuration in OSPF view.

Table 19-6 Configure parameters for OSPF to import external routes

Operation	Command
Configure the default cost for the OSPF to import external routes	default cost <i>value</i>
Restore the default cost for the OSPF to import external routes	undo default cost
Configure the default tag for the OSPF to import external routes	default tag <i>tag</i>
Restore the default tag for the OSPF to import external routes	undo default tag
Configure the default type of external routes that OSPF will import	default type { 1 2 }
Restore the default type of the external routes imported by OSPF	undo default type

By default, the type of imported route is type-2, the cost is 1 and the tag is 1 for a imported route.

III. Configuring the default interval and number for OSPF to import external routes

OSPF can import the external routing information and broadcast it to the entire autonomous system. Importing routes too often and importing too many external routes at one time will greatly affect the performance of the device. Therefore it is necessary to specify the default interval and number for the protocol to import external routes.

Perform the following configuration in OSPF view.

Table 19-7 Configure the default interval and number for OSPF to import external routes

Operation	Command
Configure the default interval for OSPF to import external routes	default interval <i>seconds</i>
Restore the default interval for OSPF to import external routes	undo default interval
Configure the upper limit to the routes that OSPF imports at a time	default limit <i>routes</i>
Restore the default upper limit to the external routes that can be imported at a time	undo default limit

By default, the interval for importing external routes is 1 second. The upper limit to the external routes imported is 1000 at a time.

19.2.6 Configuring OSPF to Import Default Routes

By default, there are no default routes in a common OSPF area (either a backbone area or a non-backbone area). Besides, the **import-route** command cannot be used to import the default route.

Use the **default-route-advertise** command to generate and advertise a default route in an OSPF route area. Note the following when you use this command:

- If you use the **default-route-advertise** command on an ASBR or ABR of a common OSPF area, the system generates a Type-5 LSA, advertising the default route in the OSPF route area.
- If you use the **default-route-advertise** command on an ASBR or ABR of an NSSA, the system generates a Type-7 LSA, advertising the default route in the NSSA.
- This command is invalid for a stub area or a totally stub area.
- For an ASBR, the system generates the corresponding Type-5 LSA or Type-7 LSA by default when a default route existed in the routing table.
- For an ABR, the system will generate a Type-5 LSA or Type-7 LSA no matter whether there is a default route in the routing table.

- The broadcasting scope of Type-5 LSA or Type-7 LSA advertising the default route is the same as that of the common Type-5 LSA or Type-7 LSA.

Perform the following configuration in OSPF view.

Table 19-8 Configure OSPF to import the default route

Operation	Command
Import the default route to OSPF	default-route-advertise [always cost <i>value</i> type <i>type-value</i> route-policy <i>route-policy-name</i>]*
Remove the imported default route	undo default-route-advertise [always cost type route-policy]*

By default, OSPF does not import the default route.

If you use the **always** keyword of this command, the system will generate a Type-5 or Type-7 LSA no matter whether there is default route in the routing table. Be cautious that the **always** keyword is only valid for an ASBR.

Because OSPF does not calculate the LSAs it generated during SPF calculation, there is no default route in the OSPF route on this router. To ensure the correct routing information, you should configure to import the default route on the router only connected to the external network.

Note:

- After the **default-route-advertise** command is configured on the OSPF router, this router becomes an ASBR. For the OSPF router, the **default-route-advertise** and **import-route** commands have the similar effect.
 - For the ABR or ASBR in the NSSA area, the **default-route-advertise** and **nssa default-route-advertise** commands have the same effect.
-

19.2.7 Configuring OSPF Route Filtering

Perform the following configuration in OSPF view.

I. Configuring OSPF to filter the received routes

Table 19-9 Enable OSPF to filter the received routes

Operation	Command
Disable filtering the received global routing information	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> gateway <i>ip-prefix-name</i> } import
Cancel filtering the received global routing information	undo filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> gateway <i>ip-prefix-name</i> } import

By default, OSPF will not filter the received routing information.

II. Configuring filtering the routes imported to OSPF

Use the **filter-policy export** command to configure the ASBR router to filter the external routes imported to OSPF. This command is only valid for the ASBR router.

Table 19-10 Enable OSPF to filter the imported routes of other routing protocols

Operation	Command
Enable OSPF to filter the routes advertised by other routing protocols	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } export [<i>routing-process</i>]
Disable OSPF to filter the advertised routes by other routing protocols	undo filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } export [<i>routing-process</i>]

By default, OSPF does not receive the routes advertised by other routing protocols.

Note:

- The **filter-policy import** command only filters the OSPF routes of this process received from the neighbors, and routes that cannot pass the filter will not be added to the routing table. This command only takes effect on ABR.
 - The **filter-policy export** command only takes effect on the routes imported by the **import-route** command. If you configure the switch with only the **filter-policy export** command, but without configuring the **import-route** command to import other external routes (including OSPF routes of different process), then the **filter-policy export** command does not take effect.
 - If the **filter-policy export** command does not specify which type of route is to be filtered, it takes effect on all routes imported by the local device using the **import-route** command.
-

19.2.8 Configuring the Route Summary of OSPF

I. Configuring the route summary of OSPF area

Route summary means that ABR can aggregate information of the routes of the same prefix and advertise only one route to other areas. An area can be configured with multiple aggregate segments, thereby OSPF can summarize them. When the ABR transmits routing information to other areas, it will generate Sum_net_Lsa (type-3 LSA) per network. If some continuous networks exist in this area, you can use the **abr-summary** command to summarize these segments into one segment. Thus, the ABR only needs to send an aggregated LSA, and all the LSAs in the range of the

aggregate segment specified by the command will not be transmitted separately. This can reduce the LSDB size in other areas.

Once the aggregated segment of a certain network is added to the area, all the internal routes of the IP addresses in the range of the aggregated segment will no longer be separately advertised to other areas. Only the route summary of the whole aggregated network will be advertised. But if the range of the segment is restricted by the keyword **not-advertise**, the route summary of this segment will not be advertised. This segment is represented by IP address and mask.

Route summary can take effect only when it is configured on ABRs.

Perform the following configuration in OSPF area view.

Table 19-11 Configure the route summary of OSPF area

Operation	Command
Configure route summary of OSPF area	abr-summary <i>ip-address</i> <i>mask</i> [advertise not-advertise]
Cancel route summary of OSPF area	undo abr-summary <i>ip-address</i> <i>mask</i>

By default, route summary is disabled on ABRs.

II. Configuring summarization of imported routes by OSPF

OSPF of the Switch 8800 supports route summarization of imported routes.

Perform the following configurations in OSPF view.

Table 19-12 Configure summarization of imported routes by OSPF

Operation	Command
Configure summarization of imported routes by OSPF	asbr-summary <i>ip-address</i> <i>mask</i> [not-advertise tag <i>value</i>]
Remove summarization of routes imported into OSPF	undo asbr-summary <i>ip-address</i> <i>mask</i>

By default, summarization of imported routes is disabled.

After the summarization of imported routes is configured, if the local router is an autonomous system border router (ASBR), this command summarizes the imported Type-5 LSAs in the summary address range. When NSSA is configured, this command will also summarize the imported Type-7 LSA in the summary address range.

If the local router works as an area border router (ABR) and a router in the NSSA, this command summarizes Type-5 LSAs transformed from Type-7 LSAs. If the router is not the router in the NSSA, the summarization is disabled.

19.2.9 Setting OSPF Route Preference

Since maybe multiple dynamic routing protocols are running on one router concurrently, the problem of route sharing and selection between various routing protocols occurs. The system sets a preference for each routing protocol, which will be used in tie-breaking in case different protocols discover the same route.

Perform the following configuration in OSPF view.

Table 19-13 Set OSPF route preference

Operation	Command
Configure a preference for OSPF for comparing with the other routing protocols	preference [ase] preference
Restore the default protocol preference	undo preference [ase]

By default, the OSPF preference is 10, and that of the imported external routing protocol is 150.

19.2.10 Configuring OSPF Timers

I. Setting the interval for Hello packet transmission

Hello packets are a kind of most frequently used packets, which are periodically sent to the adjacent router for discovering and maintaining the adjacency, and for electing DR and BDR. The user can set the hello timer.

According to RFC2328, the consistency of hello intervals between network neighbors should be kept. The hello interval value is in inverse proportion to the route convergence rate and network load.

Perform the following configuration in interface view.

Table 19-14 Set the interval for Hello packet transmission

Operation	Command
Set the hello interval of the interface	ospf timer hello seconds
Restore the default hello interval of the interface	undo ospf timer hello
Set the poll interval on the NBMA interface	ospf timer poll seconds
Restore the default poll interval	undo ospf timer poll

By default, p2p and broadcast interfaces send Hello packets every 10 seconds, and p2mp and nbma interfaces send packets every 30 seconds.

II. Setting a dead timer for the neighboring routers

The dead timer of neighboring routers refers to the interval in which a router will regard the neighboring router as dead if no Hello packet is received from it. The user can set a dead timer for the neighboring routers.

Perform the following configuration in interface view.

Table 19-15 Set a dead timer for the neighboring routers

Operation	Command
Configure a dead timer for the neighboring routers	ospf timer dead <i>seconds</i>
Restore the default dead interval of the neighboring routers	undo ospf timer dead

By default, the dead interval for the neighboring routers of p2p or broadcast interfaces is 40 seconds and that for the neighboring routers of p2mp or nbma interfaces is 120 seconds.

Note that both hello and dead timer will restore to the default values after the user modify the network type.

III. Setting an interval for LSA retransmission between neighboring routers

If a router transmits a Link State Advertisements (LSA) to the peer, it requires the acknowledgement packet from the peer. If it does not receive the acknowledgement packet within the retransmit time, it will retransmit this LSA to the neighbor. The value of retransmit is user-configurable.

Perform the following configuration in interface view.

Table 19-16 Set an interval for LSA retransmission between neighboring routers

Operation	Command
Configure the interval of LSA retransmission for the neighboring routers	ospf timer retransmit <i>interval</i>
Restore the default LSA retransmission interval for the neighboring routers	undo ospf timer retransmit

By default, the interval for neighboring routers to retransmit LSAs is 5 seconds.

The value of *interval* should be bigger than the roundtrip value of a packet.

Note that you should not set the LSA retransmission interval too small. Otherwise, unnecessary retransmission will be caused.

19.2.11 Configuring the Network Type on the OSPF Interface

The route calculation of OSPF is based upon the topology of the adjacent network of the local router. Each router describes the topology of its adjacent network and transmits it to all the other routers.

OSPF divides networks into four types by link layer protocol:

- Broadcast: If Ethernet or FDDI is adopted, OSPF defaults the network type to broadcast.
- Non-Broadcast Multi-access (**nbma**): If Frame Relay, ATM, HDLC or X.25 is adopted, OSPF defaults the network type to NBMA.
- Point-to-Multipoint (**p2mp**): OSPF will not default the network type of any link layer protocol to **p2mp**. A **p2mp** network is always changed from another type of network. The general undertaking is to change a partially connected NBMA network to **p2mp** network if the NBMA network is not fully connected.
- Point-to-point (**p2p**): If PPP or LAPB is adopted, OSPF defaults the network type to **p2p**.

NBMA means that a network is non-broadcast and multi-accessible. ATM is a typical example for it. The user can configure the polling interval to specify the interval for sending polling hello packets before the adjacency of the neighboring routers is formed.

Set the network type to NBMA if routers not supporting multicast addresses exist in a broadcast network.

Set the interface type to **p2mp** if not all the routers are directly accessible on an NBMA network.

Change the interface type to **p2p** if the router has only one peer on the NBMA network.

The differences between NBMA and **p2mp** are listed below:

- With OSPF, NBMA refers to the networks that are fully connected, non-broadcast and multi-accessible. However, a **p2mp** network is not necessarily fully connected.
- DR and BDR are required on a NBMA network but not on **p2mp** network.
- NBMA is the default network type. For example, if ATM is adopted as the link layer protocol, OSPF defaults the network type on the interface to NBMA, regardless of whether the network is fully connected. **p2mp** is not the default network type. No link layer protocols are regarded as **p2mp**. You must change the network type to **p2mp** by force. The commonest undertaking is to change a partially connected NBMA network to a **p2mp** network.
- NBMA forwards packets by unicast and requires configuring neighbors manually. **p2mp** forwards packets by multicast.

Perform the following configuration in interface view.

Table 19-17 Configure a network type for an OSPF interface

Operation	Command
Configure the network type on the interface	ospf network-type { broadcast nbma p2mp p2p }
Restore the default network type of the OSPF interface	undo ospf network-type

By default, OSPF determines the network type based on the link layer type. After the interface has been configured with a new network type, the original network type of the interface is removed automatically.

19.2.12 Configuring NBMA Neighbors for OSPF

For an NBMA network, some special configurations are required. Since an NBMA interface on the network cannot discover the adjacent router through broadcasting Hello packets, you must manually specify an IP address for the adjacent router for the interface, and specify whether the adjacent router is eligible for election.

Perform the following configuration in OSPF view.

Table 19-18 Configure the NBMA neighbors for OSPF

Operation	Command
Configure the NBMA neighbors for OSPF	peer ip-address [dr-priority dr-priority-number]
Remove the configured NBMA neighbors	undo peer ip-address

By default, the preference for NBMA neighbor is 1.

19.2.13 Setting the Interface Priority for DR Election

On a broadcast or NBMA network, a designated router (DR) and a backup designated router (BDR) must be elected.

The priority of a router interface determines the qualification of the interface in DR election. The router with the priority of 0 cannot be elected as the DR or BDR.

DR is not designated manually. Instead, it is elected by all the routers on the segment. Routers with the priorities larger than 0 in the network are eligible “candidates”. Votes are hello packets. Each router writes the expected DR in the packet and sends it to all the other routers on the segment. If two routers attached to the same segment concurrently declare themselves to be the DR, choose the one with higher priority. If the priorities are the same, choose the one with greater router ID. If the priority of a router is 0, it will not be elected as DR or BDR.

If DR fails due to some faults, the routers on the network must elect a new DR and synchronize with the new DR. The process will take a relatively long time, during which, the route calculation is incorrect. In order to speed up this process, OSPF puts forward the concept of BDR. In fact, BDR is a backup for DR. DR and BDR are elected in the meantime. The adjacencies are also established between the BDR and all the routers on the segment, and routing information is also exchanged between them. When the DR fails, the BDR will become the DR instantly. Since no re-election is needed and the adjacencies have already been established, the process is very short. But in this case, a new BDR should be elected. Although it will also take a quite long period of time, it will not exert any influence upon the route calculation.

Note the following:

- The DR on the network is not necessarily the router with the highest priority. Likewise, the BDR is not necessarily the router with the second highest priority. If a new router is added after DR and BDR election, it is impossible for the router to become the DR even if it has the highest priority.
- DR is based on the router interface in a certain segment. Maybe a router is a DR on one interface, but can be a BDR or DROther on another interface.
- DR election is only required for the broadcast or NBMA interfaces. For the **p2p** or **p2mp** interfaces, DR election is not required.

Perform the following configuration in interface view.

Table 19-19 Set the interface priority for DR election

Operation	Command
Configure the interface with a priority for DR election	ospf dr-priority <i>priority_num</i>
Restore the default interface priority	undo ospf dr-priority

By default, the priority of the interface is 1 in the DR election.

Use the **ospf dr-priority** and **peer** commands to set priorities with different usages:

- Use the **ospf dr-priority** command to set priority for DR selection.
- The priority you use the **peer** command to set indicates whether the adjacent router is eligible for election. If you specify the priority as 0 during neighbor configuration, the local router considers that this neighbor is not eligible for election, thus sending no Hello packets to this neighbor. This configuration can reduce the Hello packets on the network during DR and BDR selection. However, if the local router is DR or BDR, this router can also send Hello packets to the neighbor with priority 0 to establish adjacency relations.

19.2.14 Configuring an Interval Required for Sending LSU Packets

Trans-delay seconds should be added to the aging time of the LSA in an LSU packet. Setting the parameter like this mainly considers the time duration that the interface requires for transmitting a packet.

The user can configure the interval of sending LSU message. Obviously, more attention should be paid to this item over low speed networks.

Perform the following configuration in interface view.

Table 19-20 Configure an interval required for sending LSU packets

Operation	Command
Configure an interval for sending LSU packets	ospf trans-delay <i>seconds</i>
Restore the default interval for sending LSU packets	undo ospf trans-delay

By default, the LSU packets are transmitted per second.

19.2.15 Configuring the Cost for Sending Packets on an Interface

The user can control the network traffic by configuring different packet sending costs for different interfaces.

Perform the following configuration in interface view.

Table 19-21 Configure the cost for sending packets on an interface

Operation	Command
Configure the cost for sending packets on an interface	ospf cost <i>value</i>
Restore the default cost for packet transmission on the interface	undo ospf cost

For the Switch 8800, the default cost for running OSPF on the VLAN interface is 10.

19.2.16 Configuring to Fill the MTU Field When an Interface Transmits DD Packets

OSPF-running routers use Database Description (DD) packets to describe their own LSDBs during LSDB synchronization.

You can manually specify an interface to fill in the MTU field in a DD packet when it transmits the packet. The MTU should be set to the real MTU on the interface.

Perform the following configuration in interface view.

Table 19-22 Configure whether the MTU field will be filled in when an interface transmits DD packets

Operation	Command
Enable an interface to fill in the MTU field when transmitting DD packets	ospf mtu-enable
Disable the interface to fill the MTU field when transmitting DD packets	undo ospf mtu-enable

By default, the interface does not fill in the MTU field when transmitting DD packets. In other words, MTU in the DD packets is 0.

19.2.17 Setting a Shortest Path First (SPF) Calculation Interval for OSPF

Whenever the LSDB of OSPF takes changes, the shortest path requires recalculation. Calculating the shortest path upon change will consume enormous resources as well as affect the operation efficiency of the router. Adjusting the SPF calculation interval, however, can restrain the resource consumption due to frequent network changes.

Perform the following configuration in OSPF view.

Table 19-23 Set the SPF calculation interval

Operation	Command
Set the SPF calculation interval	spf-schedule-interval <i>seconds</i>
Restore the SPF calculation interval	undo spf-schedule-interval <i>seconds</i>

By default, the interval of SPF recalculation is five seconds.

19.2.18 Disabling the Interface to Send OSPF Packets

To prevent OSPF routing information from being acquired by the routers on a certain network, use the **silent-interface** command to disable the interface to transmit OSPF packets.

Perform the following configuration in OSPF view.

Table 19-24 Enable/Disable the interface to send OSPF packets

Operation	Command
Disable the interface to send OSPF packets	silent-interface <i>silent-interface-type</i> <i>silent-interface-number</i>
Enable the interface to send OSPF packets	undo silent-interface <i>silent-interface-type</i> <i>silent-interface-number</i>

By default, all interfaces are allowed to transmit and receive OSPF packets.

After an OSPF interface is set to be in silent status, the interface can still advertise its direct route. However, the OSPF hello packets of the interface will be blocked, and no neighboring relationship can be established on the interface. Thereby, the capability for OSPF to adapt to the networking can be enhanced, which will hence reduce the consumption of system resources. On a switch, this command can disable/enable the specified VLAN interface to send OSPF packets.

19.2.19 Configuring OSPF Authentication

I. Configuring the OSPF Area to Support Packet Authentication

All the routers in one area must use the same authentication mode (no authentication, simple text authentication or MD5 cipher text authentication). If the mode of supporting authentication is configured, all routers on the same segment must use the same authentication key. To configure a simple text authentication key, use the **authentication-mode simple** command. Use the **authentication-mode md5** command to configure the MD5 cipher text authentication key if the area is configured to support MD5 cipher text authentication mode.

Perform the following configuration in OSPF area view.

Table 19-25 Configure the OSPF area to support packet authentication

Operation	Command
Configure the area to support authentication type	authentication-mode { simple md5 }
Cancel the configured authentication mode	undo authentication-mode

By default, the area does not support packet authentication.

II. Configuring OSPF packet authentication

OSPF supports simple authentication or MD5 authentication between neighboring routers.

Perform the following configuration in interface view.

Table 19-26 Configure OSPF packet authentication

Operation	Command
Specify a password for OSPF simple text authentication on the interface	ospf authentication-mode simple <i>password</i>
Cancel simple authentication on the interface	undo ospf authentication-mode simple
Specify the interface to use MD5 authentication	ospf authentication-mode md5 <i>key_id key</i>

Operation	Command
Disable the interface to use MD5 authentication	undo ospf authentication-mode md5

By default, the interface is not configured with either simple authentication or MD5 authentication.

19.2.20 Configuring OSPF Virtual Link

According to RFC2328, after the area partition of OSPF, not all the areas are equal. In which, an area is different from all the other areas. Its area-id is 0.0.0.0 and it is usually called the backbone Area. The OSPF routes between non-backbone areas are updated with the help of the backbone area. OSPF stipulates that all the non-backbone areas should maintain the connectivity with the backbone area. That is, at least one interface on the ABR should fall into the area 0.0.0.0. If an area does not have a direct physical link with the backbone area 0.0.0.0, a virtual link must be created.

If the physical connectivity cannot be ensured due to the network topology restriction, a virtual link can satisfy this requirement. The virtual link refers to a logic channel set up through the area of a non-backbone internal route between two ABRs. Both ends of the logic channel should be ABRs and the connection can take effect only when both ends are configured. The virtual link is identified by the ID of the remote router. The area, which provides the ends of the virtual link with a non-backbone area internal route, is called the transit area. The ID of the transit area should be specified during configuration.

The virtual link is activated after the route passing through the transit area is calculated, which is equivalent to a **p2p** connection between two ends. Therefore, similar to the physical interfaces, you can also configure various interface parameters on this link, such as hello timer.

The "logic channel" means that the routers running OSPF between two ABRs only take the role of packet forwarding (the destination addresses of the protocol packets are not these routers, so these packets are transparent to them and the routers forward them as common IP packets). The routing information is directly transmitted between the two ABRs. The routing information herein refers to the type-3 LSAs generated by the ABRs, for which the synchronization mode of the routers in the area will not be changed.

Perform the following configuration in OSPF area view.

Table 19-27 Configure an OSPF virtual link

Operation	Command
Create and 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> simple <i>password</i> md5 <i>keyid key</i>]*
Remove the created virtual link	undo vlink-peer <i>router-id</i>

area-id and *router-id* have no default value. By default, hello timer is 10 seconds, retransmit 5 seconds, trans-delay 1 second, and the dead 40 seconds.

19.2.21 Configuring Stub Area of OSPF

Stub areas are some special areas, in which the ABRs do not propagate the learned external routes of the AS.

The stub area is an optional configuration attribute, but not every area conforms to the configuration condition. Generally, stub areas, located at the AS boundaries, are those non-backbone areas with only one ABR. Even if this area has multiple ABRs, no virtual links are established between these ABRs.

To ensure that the routes to the destinations outside the AS are still reachable, the ABR in this area will generate a default route (0.0.0.0) and advertise it to the non-ABR routers in the area.

Pay attention to the following items when configuring a stub area:

- The backbone area cannot be configured to be the stub area and the virtual link cannot pass through the stub area.
- If you want to configure an area to be the stub area, then all the routers in this area should be configured with this attribute.
- No ASBR can exist in a stub area. In other words, the external routes of the AS cannot be propagated in the stub area.

Perform the following configuration in OSPF area view.

Table 19-28 Configure stub area of OSPF

Operation	Command
Configure an area to be the stub area	stub [no-summary]
Remove the configured stub area	undo stub
Configure the cost of the default route transmitted by OSPF to the stub area	default-cost <i>value</i>
Remove the cost of the default route to the stub area	undo default-cost

By default, the stub area is not configured, and the cost of the default route to the stub area is 1.

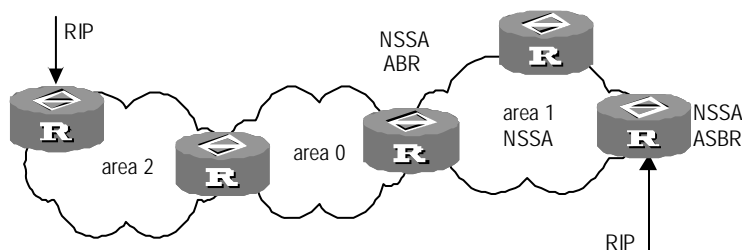
19.2.22 Configuring NSSA Area of OSPF

RFC1587 introduced a new type of area called NSSA area, and a new type of LSA called NSSA LSA (or Type-7 LSA).

NSSA areas are virtually variations of Stub areas. They are similar in many ways. Neither of them generates or imports AS-External-LSA (namely Type-5 LSA), and both of them can generate and import Type-7 LSA. Type-7 LSA is generated by ASBR of NSSA area, which can only be advertised in NSSA area. When Type-7 LSA reaches ABR of NSSA, ABR will select whether to transform Type-7 LSA into AS-External-LSA so as to advertise to other areas.

For example, in the network below, the AS running OSPF comprises three areas: Area 1, Area 2 and Area 0. Among them, Area 0 is the backbone area. Also, there are other two ASs respectively running RIP. Area 1 is defined as an NSSA area. After RIP routes of the Area 1 are propagated to the NSSA ASBR, the NSSA ASBR will generate type-7 LSAs which will be propagated in Area 1. When the type-7 LSAs reach the NSSA ABR, the NSSA ABR will transform it into type-5 LSA, which will be propagated to Area 0 and Area 2. On the other hand, RIP routes of the AS running RIP will be transformed into type-5 LSAs that will be propagated in the OSPF AS. However, the type-5 LSAs will not reach Area 1 because Area 1 is an NSSA. NSSAs and stub areas have the same approach in this aspect.

Similar to a stub area, the NSSA cannot be configured with virtual links.

**Figure 19-2** NSSA area

Perform the following configuration in OSPF area view.

Table 19-29 Configure NSSA of OSPF

Operation	Command
Configure an area to be the NSSA area	nssa [default-route-advertise no-import-route no-summary]*
Cancel the configured NSSA	undo nssa
Configure the default cost value of the route to the NSSA	default-cost <i>cost</i>
Restore the default cost value of the route to the NSSA area	undo default-cost

All the routers connected to the NSSA should use the **nssa** command to configure the area with the NSSA attribute.

The keyword **default-route-advertise** is used to generate default type-7 LSAs. When **default-route-advertise** is configured, a default type-7 LSA route will be generated on the ABR, even though no default route 0.0.0.0 is in the routing table. On an ASBR, however, a default type-7 LSA route can be generated only if the default route 0.0.0.0 is in the routing table.

Executing the keyword **no-import-route** on the ASBR will prevent the external routes that OSPF imported through the **import-route** command from being advertised to the NSSA. Generally, if an NSSA router is both ASBR and ABR, this keyword will be used.

The keyword **default-cost** is used on the ABR attached to the NSSA. Using this command, you can configure the default route cost on the ABR to NSSA.

By default, the NSSA is not configured, and the cost of the default route to the NSSA is 1.

19.2.23 Configuring OSPF and Network Management System (NMS)

I. Configuring OSPF MIB binding

After multiple OSPF processes are enabled, you can configure to which OSPF process MIB is bound.

Perform the following configuration in system view.

Table 19-30 Configure OSPF MIB binding

Operation	Command
Configure OSPF MIB binding	ospf mib-binding <i>process-id</i>
Restore the default OSPF MIB binding	undo ospf mib-binding

By default, MIB is bound to the first enabled OSPF process.

II. Configuring OSPF TRAP

You can configure the switch to send multiple types of SNMP TRAP packets in case of OSPF anomalies. In addition, you can configure the switch to send SNMP TRAP packets when a specific process is abnormal by specifying the process ID.

Perform the following configuration in system view.

Table 19-31 Enable/Disable OSPF TRAP function

Operation	Command
Enable OSPF TRAP function	snmp-agent trap enable ospf [<i>process-id</i>] [ifstatechange virifstatechange nbrstatechange virnbrstatechange ifcggerr virifcggerr ifauthfail virifauthfail ifrxbadpkt virifrxbadpkt txretransmit viriftxretransmit originatelsa maxagelsa lsdboverflow lsdbapproachoverflow]
Disable OSPF TRAP function	undo snmp-agent trap enable ospf [<i>process-id</i>] [ifstatechange virifstatechange nbrstatechange virnbrstatechange ifcggerr virifcggerr ifauthfail virifauthfail ifrxbadpkt virifrxbadpkt txretransmit viriftxretransmit originatelsa maxagelsa lsdboverflow lsdbapproachoverflow]

By default, OSPF TRAP function is disabled. That is, the switch does not send TRAP packets when any OSPF process is abnormal. The configuration is valid to all OSPF processes if you do not specify a process ID.

For detailed configuration of SNMP TRAP, refer to the module "System Management" in this manual.

19.2.24 Resetting the OSPF Process

If the **undo ospf** command is executed on a router and then the **ospf** command is used to restart the OSPF process, the previous OSPF configuration will lose. With the **reset ospf** command, you can restart the OSPF process without losing the previous OSPF configuration.

Perform the following configuration in user view.

Table 19-32 Reset OSPF processes

Operation	Command
Reset one or all OSPF processes	reset ospf [statistics] { all <i>process-id</i> }

Resetting the OSPF process can immediately clear invalid LSAs, and make the modified router ID effective or the DR and BDR are re-elected.

19.3 Displaying and Debugging OSPF

After the above configuration, execute the **display** command in any view to display the running of the OSPF configuration, and to verify the effect of the configuration. Execute the **debugging** command in user view to debug the OSPF module.

Table 19-33 Display and debug OSPF

Operation	Command
Display the brief information of the OSPF routing process	display ospf [<i>process-id</i>] brief
Display OSPF statistics	display ospf [<i>process-id</i>] cumulative
Display LSDB information of OSPF	display ospf [<i>process-id</i>] [<i>area-id</i>] lsdb [brief [asbr ase network nssa router summary] [<i>ip-address</i>] [originate-router <i>ip-address</i> self-originate]]
Display OSPF peer information	display ospf [<i>process-id</i>] peer [brief]
Display OSPF next hop information	display ospf [<i>process-id</i>] nexthop
Display OSPF routing table	display ospf [<i>process-id</i>] routing
Display OSPF virtual links	display ospf [<i>process-id</i>] vlink
Display OSPF request list	display ospf [<i>process-id</i>] request-queue
Display OSPF retransmission list	display ospf [<i>process-id</i>] retrans-queue
Display the information of OSPF ABR and ASBR	display ospf [<i>process-id</i>] abr-asbr
Display the summary information of OSPF imported routes	display ospf [<i>process-id</i>] asbr-summary [<i>ip-address mask</i>]
Display OSPF interface information	display ospf [<i>process-id</i>] interface
Display OSPF errors	display ospf [<i>process-id</i>] error
Display the state of the global OSPF debugging switches and the state of the debugging switches for each process	display debugging ospf
Enable OSPF packet debugging	debugging ospf packet [ack dd hello interface <i>interface-type interface-number</i> request update]
Disable OSPF packet debugging	undo debugging ospf packet [ack dd hello interface <i>interface-type interface-number</i> request update]
Enable OSPF event debugging	debugging ospf event

Operation	Command
Disable OSPF event debugging	undo debugging ospf event
Enable OSPF LSA packet debugging	debugging ospf lsa-originate
Disable OSPF LSA packet debugging	undo debugging ospf lsa-originate
Enable SPF debugging of OSPF	debugging ospf spf
Disable SPF debugging of OSPF	undo debugging ospf spf

19.4 Typical OSPF Configuration Example

19.4.1 Configuring DR Election Based on OSPF Priority

I. Network requirements

Four Switch 8800s, Switch A, Switch B, Switch C and Switch D, which can perform the router functions and run OSPF, are located on the same segment, as shown in the following figure.

Configure Switch A and Switch C as DR and BDR respectively. The priority of Switch A is 100, which is the highest on the network, so it is elected as the DR. Switch C has the second highest priority, that is, 2, so it is elected as the BDR. The priority of Switch B is 0, which means that it cannot be elected as the DR. Switch D does not have a priority, which takes 1 by default.

II. Network diagram

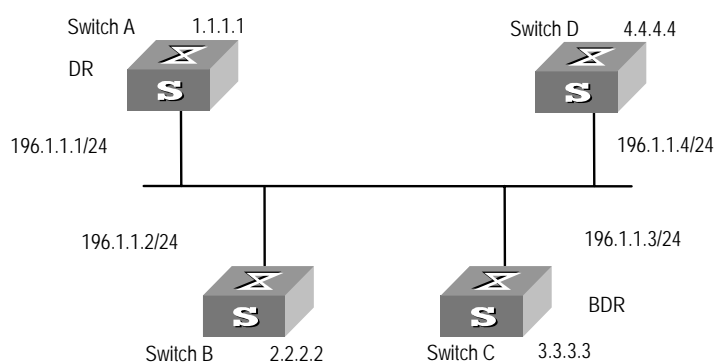


Figure 19-3 Network diagram for configuring DR election based on OSPF priority

III. Configuration procedure

Configure Switch A

```
[Switch A] interface Vlan-interface 1
[Switch A-Vlan-interface1] ip address 196.1.1.1 255.255.255.0
[Switch A-Vlan-interface1] ospf dr-priority 100
[Switch A] router id 1.1.1.1
[Switch A] ospf
[Switch A-ospf-1] area 0
[Switch A-ospf-1-area-0.0.0.0] network 196.1.1.0 0.0.0.255
```

Configure Switch B.

```
[Switch B] interface Vlan-interface 1
[Switch B-Vlan-interface1] ip address 196.1.1.2 255.255.255.0
[Switch B-Vlan-interface1] ospf dr-priority 0
[Switch B] router id 2.2.2.2
[Switch B] ospf
[Switch B-ospf-1] area 0
[Switch B-ospf-1-area-0.0.0.0] network 196.1.1.0 0.0.0.255
```

Configure Switch C.

```
[Switch C] interface Vlan-interface 1
[Switch C-Vlan-interface1] ip address 196.1.1.3 255.255.255.0
[Switch C-Vlan-interface1] ospf dr-priority 2
[Switch C] router id 3.3.3.3
[Switch C] ospf
[Switch C-ospf-1] area 0
[Switch C-ospf-1-area-0.0.0.0] network 196.1.1.0 0.0.0.255
```

Configure Switch D.

```
[Switch D] interface Vlan-interface 1
[Switch D-Vlan-interface1] ip address 196.1.1.4 255.255.255.0
[Switch D] router id 4.4.4.4
[Switch D] ospf
[Switch D-ospf-1] area 0
[Switch D-ospf-1-area-0.0.0.0] network 196.1.1.0 0.0.0.255
```

On Switch A, execute the **display ospf peer** command to display the OSPF peers. Note that Switch A has three peers.

The state of each peer is full, which means that adjacency is set up between Switch A and each peer. (Switch A and Switch C should set up adjacencies with all the routers on the network for them to be DR and BDR on the network respectively.) Switch A is DR, while Switch C is BDR on the network. And all the other neighbors are DR others (which means that they are neither DRs nor BDRs).

Change the priority of Switch B to 200

```
[Switch B-Vlan-interface2000] ospf dr-priority 200
```

On Switch A, execute the **display ospf peer** command to show its OSPF neighbors. Note the priority of Switch B has changed to 200, but it is still not the DR.

Only when the current DR is offline, will the DR be changed. Shut down Switch A, and execute the **display ospf peer** command on Switch D to display its neighbors. Note that the original BDR (Switch C) becomes the DR, and Switch B is BDR now.

If all Switches on the network are removed and added back again, Switch B will be elected as the DR (with the priority of 200), and Switch A becomes the BDR (with a priority of 100). To switch off and restart all of the switches will bring about a new round of DR/BDR selection.

19.4.2 Configuring OSPF Virtual Link

I. Network requirements

In Figure 19-4, Area 2 and Area 0 are not directly connected. Area 1 is required to be taken as a transit area for connecting Area 2 and Area 0. Configure a virtual link between Switch B and Switch C in Area 1.

II. Network diagram

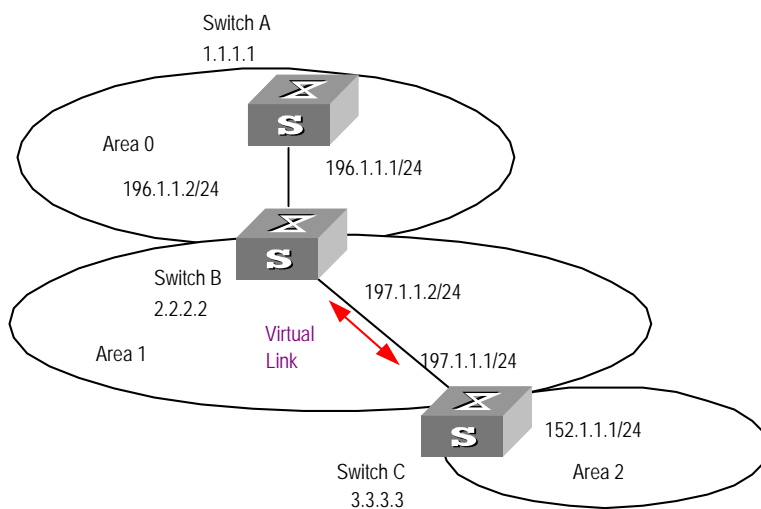


Figure 19-4 Network diagram for OSPF virtual link configuration

III. Configuration procedure

Configure Switch A

```
[Switch A] interface Vlan-interface 1
[Switch A-Vlan-interface1] ip address 196.1.1.1 255.255.255.0
[Switch A] router id 1.1.1.1
[Switch A] ospf
[Switch A-ospf-1] area 0
[Switch A-ospf-1-area-0.0.0.0] network 196.1.1.0 0.0.0.255
```

Configure Switch B

```
[Switch B] interface vlan-interface 7
[Switch B-Vlan-interface7] ip address 196.1.1.2 255.255.255.0
[Switch B] interface vlan-interface 8
[Switch B-Vlan-interface8] ip address 197.1.1.2 255.255.255.0
[Switch B] router id 2.2.2.2
[Switch B] ospf
[Switch B-ospf-1] area 0
[Switch B-ospf-1-area-0.0.0.0] network 196.1.1.0 0.0.0.255
[Switch B-ospf-1-area-0.0.0.0] quit
[Switch B-ospf-1] area 1
[Switch B-ospf-1-area-0.0.0.1] network 197.1.1.0 0.0.0.255
[Switch B-ospf-1-area-0.0.0.1] vlink-peer 3.3.3.3
```

Configure Switch C

```
[Switch C] interface Vlan-interface 1
[Switch C-Vlan-interface1] ip address 152.1.1.1 255.255.255.0
[Switch C] interface Vlan-interface 2
[Switch C-Vlan-interface2] ip address 197.1.1.1 255.255.255.0
[Switch C] router id 3.3.3.3
[Switch C] ospf
[Switch C-ospf-1] area 1
[Switch C-ospf-1-area-0.0.0.1] network 197.1.1.0 0.0.0.255
[Switch C-ospf-1-area-0.0.0.1] vlink-peer 2.2.2.2
[Switch C-ospf-1-area-0.0.0.1] quit
[Switch C-ospf-1] area 2
[Switch C-ospf-1-area-0.0.0.2] network 152.1.1.0 0.0.0.255
```


19.5 Troubleshooting OSPF Faults

Symptom 1: OSPF has been configured in accordance with the earlier-mentioned steps, but OSPF on the router cannot run normally.

Solution: Check according to the following procedure.

Local troubleshooting: Check whether the protocol between two directly connected routers is in normal operation. The normal sign is the peer state machine between the two routers reaches the FULL state. (Note: On a broadcast or NBMA network, if the interfaces for two routers are in DROther state, the peer state machines for the two routers are in 2-way state, instead of FULL state. The peer state machine between DR/BDR and all the other routers is in FULL state.

- Execute the **display ospf peer** command to view peers.
- Execute the **display ospf interface** command to view OSPF information on the interface.
- Check whether the physical connections and the lower layer protocol operate normally. You can execute the **ping** command to test. If the local router cannot ping the peer router, it indicates that faults have occurred to the physical link and the lower layer protocol.
- If the physical link and the lower layer protocol are normal, check the OSPF parameters configured on the interface. The parameters should be the same parameters configured on the router adjacent to the interface. The same area ID should be used, and the networks and the masks should also be consistent. (The **p2p** or virtually linked segment can have different segments and masks.)
- Ensure that the dead timer on the same interface is at least four times the value of the hello timer.
- If the network type is NBMA, the peer must be manually specified, using the **peer ip-address** command.
- If the network type is broadcast or NBMA, there must be at least one interface with a priority greater than zero.
- If an area is set as the stub area, to which the routers are connected. The area on these routers must be also set as the stub area.
- The same interface type should be adopted for the neighboring routers.
- If more than two areas are configured, at least one area should be configured as the backbone area (that is to say, the area ID is 0).
- Ensure that the backbone area is connected to all other areas.
- The virtual links do not pass through the stub area.

Global troubleshooting: If OSPF cannot discover the remote routes yet in the case that the above steps are correctly performed, proceed to check the following configurations.

- If more than two areas are configured on a router, at least one area should be configured as the backbone area.

As shown in Figure 19-5: RTA and RTD are configured to belong to only one area, whereas RTB (area0 and area1) and RTC (area1 and area 2) are configured to belong to two areas. In which, RTB also belongs to area0, which is compliant with the requirement. However, none of the areas to which RTC belongs is area0. Therefore, a virtual link should be set up between RTC and RTB. Ensure that area2 and area0 (backbone area) is connected.

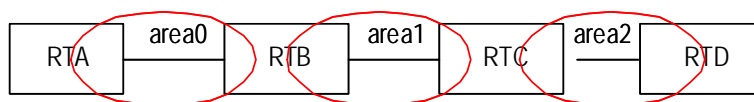


Figure 19-5 OSPF areas

- The backbone area (area 0) cannot be configured as the stub area and the virtual link cannot pass through the stub area. That is, if a virtual link has been set up between RTB and RTC, neither area1 nor area0 can be configured as a stub area. In the above figure, only area 2 can be configured as the stub area.
- Routers in the stub area cannot receive external routes.
- The backbone area must guarantee the connectivity of all nodes.

Chapter 20 Integrated IS-IS Configuration

20.1 Introduction to Integrated IS-IS

Intermediate System-to-Intermediate System (IS-IS) intra-domain routing information exchange protocol is designed by the international organization for standardization (ISO) for connection-less network protocol (CLNP). This protocol is a dynamic routing protocol. To let this protocol support IP routing, IETF expands and modifies IS-IS in RFC1195, applying the protocol to TCP/IP and OSI. The modified IS-IS is called Integrated IS-IS or Dual IS-IS.

IS-IS is a link state protocol, which uses shortest path first (SPF) algorithm. IS-IS and the OSPF protocol are similar in many aspects. As an interior gateway protocol (IGP), IS-IS is applied inside an AS.

20.1.1 Terms of IS-IS Routing Protocol

I. Terms of IS-IS routing protocol

- Intermediate System (IS). IS equals a router of TCP/IP. It is the basic unit in IS-IS protocol used for propagating routing information and generating routes. In the following text, the IS shares the same meaning with the router.
- End System (ES). It equals the host system of TCP/IP. ES does not process the IS-IS routing protocol, and therefore it can be ignored in the IS-IS protocol.
- Routing Domain (RD). A group of ISs exchange routing information with the same routing protocol in a routing domain.
- Area. Area is the division unit in the routing domain.
- Link State DataBase (LSDB). All the link states in the network form the LSDB. In an IS, at least one LSDB is available. The IS uses the SPF algorithm and the LSDB to generate its own routes.
- Link State Protocol Data Unit (LSPDU). In the IS-IS, each IS will generate an LSP which contains all the link state information of the IS. Each IS collects all the LSPs in the local area to generate its own LSDB.
- Network Protocol Data Unit (NPDU). It is the network layer packets of OSI and equals the IP packet of TCP/IP.
- Designated IS (DIS). It is the elected router on the broadcast network.
- Network Service Access Point (NSAP) is the network layer address of OSI. It identifies an abstract network service access point and describes the very network address structure for the OSI model.

II. Link types IS-IS routing protocol is applied to

IS-IS routing protocol can run on point to point Links, such as PPP, HDLC and others. IS-IS routing protocol can also run on broadcast links, such as Ethernet, Token-Ring and others. For a Non-Broadcast Multi-Access (NBMA) network such as ATM, you need to configure sub-interfaces and configure sub-interface type as P2P or broadcast network. IS-IS routing protocol cannot run on point to MultiPoint links.

20.1.2 Two-level Structure of IS-IS Routing Protocol

I. Two-level structure of IS-IS routing protocol

Two-level structure of IS-IS routing protocol is adopted in a route area to support large scale route network. A large route area can be divided into one or multiple areas. A Level-1 router manages the intra-area routes. A Level-2 router manages the inter-area routes.

II. Level-1 and Level-2

- Level-1 router

The Level-1 router is responsible for intra-area route. The Level-1 router and the Level-1 router or Level-1-2 router in the same area are neighbors. The Level-1 router maintains a Level-1 LSDB. This LSDB contains intra-area routing information. The packets sent to other areas are forwarded to the closest Level-2 router.

- Level-2 router

The Level-2 router is responsible for inter-area route. The Level-2 router and Level-2 routers or Level-1-2 routers in other areas are neighbors. The Level-2 router maintains a Level-2 LSDB. This LSDB contains inter-area routing information. The backbone (which is made up of all Level-2 routers) of a route area is responsible for inter-area communications. The Level-2 routers in the route area must be continuous to ensure the backbone continuity.

- Level-1-2 router

A Level-1-2 router is both a Level-1 router and a Level-2 router. At least one Level-1-2 router in each area connects the area to the backbone network. A Level-1-2 router maintains two LSDBs: the Level-1 LSDB for intra- area route and Level-2 LSDB for inter-area route.

Figure 20-1 illustrates a network running IS-IS routing protocol and composed of Routing Domain 1 and Routing Domain 2. Routing Domain 1 includes two areas, Area 1 and Area 2, and Routing Domain 2 only has Area 3. In Routing Domain 1, the three ISs connected by bold lines compose the area backbone. They are all Level-2 routers. The other 4 ISs not connected by bold line are Level-1 routers.

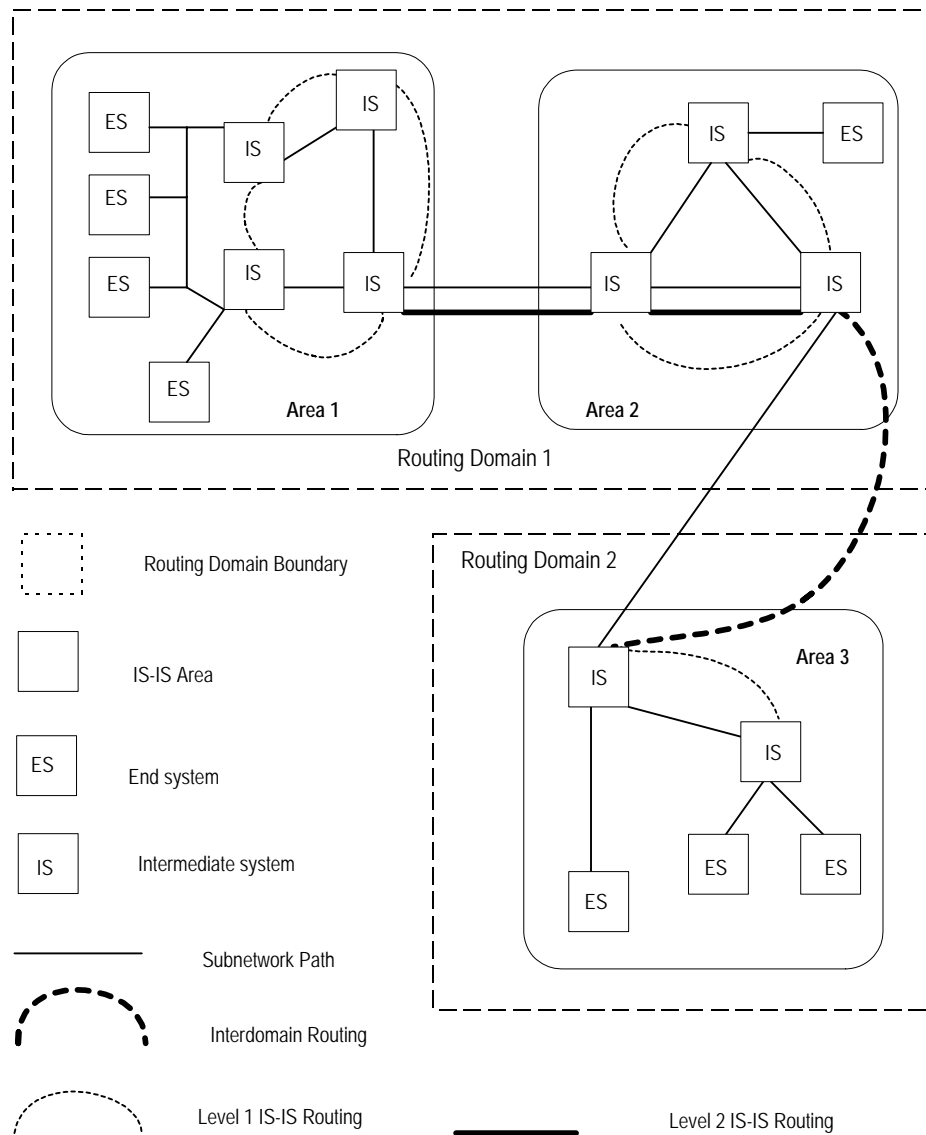


Figure 20-1 IS-IS topology

20.1.3 NSAP Structure of IS-IS Routing Protocol

I. Address structure

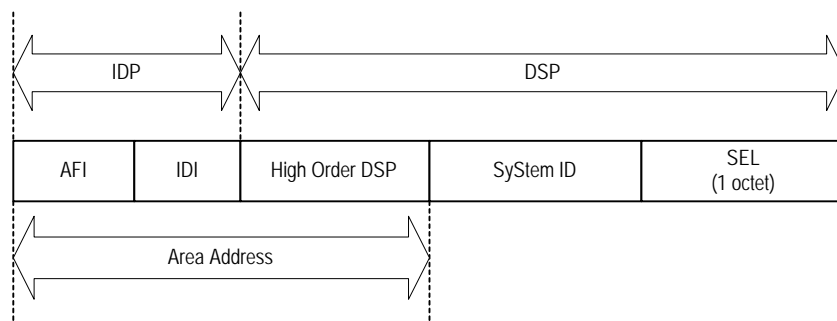


Figure 20-2 NSAP structure

OSI adopts the address structure as shown in Figure 20-2. NSAP includes initial domain part (IDP) and domain specific part (DSP). The IDP is defined by ISO; it consists of authority responsible for assigning the rest of the address and address format. The DSP is allocated by the authority specified in IDP. IDP and DSP are length-variable with a total length of 20 bytes.

- Area Address

IDP includes authority and format identifier (AFI) and initial domain identifier (IDI). AFI defines the format of IDI. DSP has several bytes. The combination of IDP and HO-DSP can identify a route area and an area of the route area, so the combination is called an area address.

In general, you only need to configure an area address for a router. The area addresses of all nodes are the same in an area. To support the seamless combination, segmentation and conversion, the Switch 8800 supports up to three area addresses.

- System ID

System ID uniquely identifies terminal system or router in a route area. You can select length for it. The System ID length is 48 bits (6 bytes). In general, you can obtain System ID according to Router_ID.

If the IP address 168.10.1.1 of the interface LoopBack0 serves as a router_ID for the router, you can use the following method to obtain the System ID:

Turn each part of the IP address 168.10.1.1 into three digits. Add 0 to the front of the part less than three digits.

Divide the expanded address 168.010.001.001 into three parts. Each part contains four digits.

You get the System ID 1680.1000.1001.

You can specify a System ID using different methods. However, you should ensure a System ID can uniquely identify a terminal system or a router.

- SEL

NSAP selector (SEL or N-SEL) functions as the protocol identifier of an IP address. Different transmission protocols correspond to different identifiers. All the SELs of IP are 00.

Because the address structure defines clearly an area, a Level-1 router can easily identify the packets not sent to the area where it is located. The Level-1 router forwards the packets to a Level-2 router.

The Level-1 router performs routing within areas by System IDs. If it detects the destination address of a packet does not belong to the area where it is located, it forwards the packet to its closest Level-2 router.

The Level-2 router performs intra-area routing according to the area address (IDP + HO-DSP).

II. NET

Network Entity Title (NET) indicates the network layer information, which contains no transfer layer information (SEL=0). You can regard it as a special NSAP.

In general, you can configure a NET for a router. If you will redivide an area (combine multiple areas or divide an area into multiple areas), you can configure multiple NETs to ensure correct routes in the case of reconfiguration. Because you can configure up to three area addresses, you can only configure up to three NETs.

For example, there is a NET 47.0001.aaaa.bbbb.cccc.00, in which,

Area=47.0001, System ID=aaaa.bbbb.cccc, SEL=00.

For example, there is a NET 01.1111.2222.4444.00, in which,

Area=01, System ID=1111.2222.4444, and SEL=00.

20.1.4 IS-IS Routing Protocol Packets

IS-IS packets are directly encapsulated in the data link frames and mainly divided into 3 kinds, Hello, LSP and SNP.

I. Hello packets

Hello packets, which is also called IIH (IS-to-IS Hello PDUs), can establish and maintain neighbor relations. The Level-1 router in a broadcast LAN forwards Level-1 LAN IIH; the Level-2 router in a broadcast LAN forwards Level-2 LAN IIH; non-broadcast network forwards Point-to-Point IIH.

II. LSP

Link state packet (LSP) can switch link state information. LSP can be divided into Level-1 LSP and Level-2 LSP. Level-2 routers transmit Level-2 LSPs; Level-1 routers transmit Level-1 LSPs; Level-1-2 routers transmit both Level-2 LSPs and Level-1 LSPs.

III. SNP

Sequence Number Packet (SNP) can confirm the LSPs last received from neighbors. SNPs function as acknowledge packets, but SNPs function more validly. SNP includes complete SNP (CSNP) and partial SNP (PSNP). SNP can be further divided into Level-1 CSNP, Level-2 CSNP, Level-1 PSNP and Level-2 PSNP.

PSNP only lists one or more last received LSP sequence numbers, and confirms multiple LSPs. When detecting asynchronous LSDBs, the system asks neighbors to send new LSPs by PSNPs.

CSNP contains all LSP digest information in a LSDB, synchronizing LSDBs for neighbor routers. On a broadcast network, a DIS sends CSNPs periodically (the default sending period is 10 seconds). On the point-to-point line, a DIS sends CSNPs only when the neighbors are established for the first time.

20.2 Configuring Integrated IS-IS

Among the following configurations, the configuration of enabling integrated IS-IS is required, while other configurations are optional.

IS-IS configuration includes:

- 1) IS-IS basic configuration
 - Enabling IS-IS and Entering the IS-IS View
 - Setting Network Entity Title
 -

Enabling IS-IS on the Specified Interface

- Setting Priority for DIS Election
 - Setting Router Type
 - Setting Interface Circuit Level
- 2) Configuration related to IS-IS route
- Configuring IS-IS to Import Routes of Other Protocols
 - Configuring IS-IS Route Filtering
 - Configuring IS-IS Routing Leak
 - Setting IS-IS Route Summary
 - Setting to Generate Default Route
- 3) Default route generation
- Setting the Preference of IS-IS Protocol
 -

- Configuring IS-IS Route Metric Type
- Setting IS-IS Link State Routing Cost
- Configuring IS-IS Timers
- Setting to Discard the LSPs with Checksum Errors
- Setting LSP Refreshment Interval
- Setting Lifetime of LSP
- Setting Parameters Related to SPF
- 4) Configuration related to IS-IS networking
 - Setting IS-IS Authentication
 - Setting Overload Flag Bit
 - Setting to Log the Peer Changes
 - Setting the Mesh Group of the Interface
 - Enabling/Disabling the Interface to Send Packets
- 5) Some operation commands
 - Resetting All the IS-IS Data Structure
 - Resetting the Specified IS-IS Peer

20.2.1 Enabling IS-IS and Entering the IS-IS View

After creating an IS-IS routing process, you should also activate this routing process at an interface that may correlate with another router. After that, the IS-IS protocol can be started and run.

Perform the following configuration in system view.

Table 20-1 Enable IS-IS and enter the IS-IS view

Operation	Command
Enable the IS-IS and enter the IS-IS view	isis [<i>tag</i>]

The *tag* argument identifies the IS-IS process. In the present version, just one IS-IS process is allowed.

By default, the IS-IS routing process is disabled.

20.2.2 Setting Network Entity Title

Network Entity Titles (hereafter referred to as NETs) defines the current IS-IS area address and the system ID of the router.

Perform the following configurations in IS-IS view.

Table 20-2 Set NET

Operation	Command
Set a NET	network-entity <i>network-entity-title</i>

Delete a NET	undo network-entity <i>network-entity-title</i>
--------------	--

The format of the *network-entity-title* argument is X...X.XXXXXXXXXXXXXX.XX, among which the first "X...X" is the area address, the twelve Xs in the middle is the System ID of the router. The last XX should be 00.

20.2.3 Enabling IS-IS on the Specified Interface

After enabling IS-IS, you need to specify on which Interfaces the IS-IS will be run.

Perform the following configuration in interface view.

Table 20-3 Enable IS-IS on the specified interface

Operation	Command
Enable IS-IS on the specified Interface	isis enable [tag]
Cancel this designation	undo isis enable [tag]

20.2.4 Setting Priority for DIS Election

In the broadcast network, the IS-IS needs to elect a DIS from all the routers.

When you need to select a DIS from the IS-IS neighbors on the broadcast network, you should select level-1 DIS and level-2 DIS respectively. The higher the priority is, the more possible it is selected. If there are two or more routers with the highest priority in the broadcast network, the one with the greatest MAC address will be selected. If all the adjacent routers' priorities are 0, the one with the greatest MAC address will be selected.

The DISs of Level-1 and Level-2 are elected separately. You can set different priorities for DIS election at different levels.

Perform the following configuration in interface view.

Table 20-4 Set priority for DIS election

Operation	Command
Set the priorities for DIS election on the interface	isis dis-priority value [level-1 level-2]
Restore the default priorities for DIS election on the interface	undo isis dis-priority [level-1 level-2]

By default, the interface priority is 64. If the level is not specified, it defaults to setting the priority of Level-1.

20.2.5 Setting Router Type

Based upon the position of the router, the levels can be divided into Level-1 (intra-domain router), Level-2 (inter-domain router) and Level-1-2 (that is, intra-domain router as well as inter-domain router).

Perform the following configuration in IS-IS view.

Table 20-5 Set the router type

Operation	Command
Set the router type	is-level { level-1 level-1-2 level-2 }
Restore the default router type	undo is-level

By default, the router type is **level-1-2**.

20.2.6 Setting Interface Circuit Level

Perform the following configuration in Interface view.

Table 20-6 Set the interface circuit level

Operation	Command
Set the interface circuit level	isis circuit-level [level-1 level-1-2 level-2]
Restore the default interface circuit level	undo isis circuit-level

Note:

Only when the router to which the interface belongs is of Level-1-2 type, is the modification to the interface circuit level meaningful. Otherwise, the type of the router determines the level of adjacency relation.

You can set the circuit level to limit what adjacency can be established for the interface. For example, Level-1 interface can only have Level-1 adjacency. Level-2 interface can only have Level-2 adjacency. For the Level-1-2 router, you can configure some interfaces to Level-2 to prevent transmitting Level-1 Hello packets to Level-2 backbone so as to save the bandwidth. However, Level-1 and Level-2 use the same kind of Hello packet over the **p2p** link, and therefore such setting is unnecessary in this case.

By default, the circuit-level on the interface is **level-1-2**.

20.2.7 Configuring IS-IS to Import Routes of Other Protocols

For IS-IS, the routes discovered by other routing protocols are processed as the routes outside the routing domain. When importing the routes of other protocols, you can specify the default cost for them.

When IS-IS imports routes, you can also specify to import the routes to Level-1, Level-2 or Level-1-2.

Perform the following configuration in IS-IS view.

Table 20-7 Import routes of other protocols

Operation	Command
Import routes of other protocols	import-route <i>protocol</i> [cost <i>value</i> type { external internal } [level-1 level-1-2 level-2] route-policy <i>route-policy-name</i>]*
Cancel importing routes from other protocols	undo import-route <i>protocol</i> [cost <i>value</i> type { external internal } [level-1 level-1-2 level-2] route-policy <i>route-policy-name</i>]*

If the level is not specified in the command for importing the route, it defaults to importing the routes into **level-2**.

protocol specifies the routing protocol sources that can be imported, which can be direct, static, rip, bgp, and ospf, etc.

By default, IS-IS does not import routing information from any other protocols.

For more about importing routing information, refer to the "Configuring IP Routing Policy" part.

20.2.8 Configuring IS-IS Route Filtering

IS-IS protocol can filter the received and advertised routes according to the access control list specified by *acl-number*.

Perform the following configuration in IS-IS view.

I. Configuring to filter the routes received by IS-IS

Table 20-8 Configure to filter the received routes

Operation	Command
Configure to filter the received routes	filter-policy <i>acl-number</i> import
Cancel filtering the received routes	undo filter-policy <i>acl-number</i> import

II. Configuring to filter the advertised routes

Table 20-9 Configure to filter the advertised routes

Operation	Command
Configure to filter the routes advertised by IS-IS	filter-policy <i>acl-number</i> export [<i>protocol</i>]
Configure not to filter the routes advertised by IS-IS	undo filter-policy <i>acl-number</i> export [<i>protocol</i>]

By default, IS-IS does not filter the route advertised by other routing protocols.

protocol specifies the routing protocol sources for advertising routes, which can be direct, static, rip, bgp, ospf, ospf-ase, and so on.

Note:

- The **filter-policy import** command only filters the ISIS routes received from the neighbors, and routes that cannot pass the filter will not be added to the routing table. This command takes effect on Level-1-2 routers.
- The **filter-policy export** command only takes effect to the routes imported by the **import-route** command. If you configure the switch with only the **filter-policy export** command, but without configuring the **import-route** command to import other external routes, then the **filter-policy export** command does not take effect.
- If the **filter-policy export** command does not specify which route to be filtered, then the all the routes imported by the **import-route** command will be filtered.

20.2.9 Configuring IS-IS Routing Leak

By virtual of IS-IS routing leak function, a Level-2 router can advertise the routing information of Level-1 areas and the Level-2 area it knows to a Level-1 router.

Perform the following configuration in IS-IS view.

Table 20-10 Configure IS-IS routing leak

Operation	Command
Enable IS-IS routing leak	import-route isis level-2 into level-1 [acl acl-number]
Disable IS-IS routing leak	undo import-route isis level-2 into level-1 [acl acl-number]

By default, a Level-2 router does not advertise its routing information to a Level-1 area.

20.2.10 Setting IS-IS Route Summary

Users can set the routes with the same next hops as one route in the routing table. Perform the following configurations in IS-IS view.

Table 20-11 Set a summary route

Operation	Command
Set a summary route	summary ip-address ip-mask [level-1 level-1-2 level-2]
Delete the summary route	undo summary ip-address ip-mask [level-1 level-1-2 level-2]

By default, the system disables route summarization.

20.2.11 Setting to Generate Default Route

In the IS-IS route domain, the Level-1 router only has the LSDB of the local area, so it can only generate the routes in the local areas. But the Level-2 router has the backbone LSDB in the IS-IS route domains and generates the backbone network routes only. If a Level-1 router in one area wants to forward the packets to other areas, it needs to first forward the packets to the closest Level-1-2 router in the local area along its default route. You do not need to configure the default Level-1 route, but need to manually configure the default Level-2 route.

Perform the following configurations in IS-IS view.

Table 20-12 Set to generate default route

Operation	Command
Set to generate default route	default-route-advertise [route-policy <i>route-policy-name</i>]
Set not to generate default route	undo default-route-advertise [route-policy <i>route-policy-name</i>]

The default route generated by this command will only be imported to the router at the same level.

20.2.12 Setting the Preference of IS-IS Protocol

In a router on which several routing protocols are concurrently operating, there is an issue of sharing and selecting the routing information among all the routing protocols. The system sets a preference for each routing protocol. When various routing protocols find the route to the same destination, the protocol with the higher preference will take effect.

Perform the following configuration in IS-IS view.

Table 20-13 Configure the preference of IS-IS protocol

Operation	Command
Configure the preference of IS-IS protocol	preference <i>value</i>
Restore the default preference of IS-IS protocol	undo preference

By default, the preference of IS-IS route is 15.

20.2.13 Configuring IS-IS Route Metric Type

IS-IS routing protocol has two styles of route metric:

- Narrow: The value of route metric ranges from 1 to 63.
- Wide: The value of route metric ranges from 1 to 16,777,215.

A router can choose either or both of the styles.

Perform the following configuration in IS-IS view.

Table 20-14 Configure the style for route metric values of IS-IS packets

Operation	Command
Configure the style for route metric values of IS-IS packets	cost-style { narrow wide wide-compatible { compatible narrow-compatible } [relax-spf-limit] }
Restore the default settings	undo cost-style

By default, IS-IS only receives and sends the packets whose route metric is in narrow style.

20.2.14 Setting IS-IS Link State Routing Cost

Users can configure the interface cost, namely, the default routing cost.

Perform the following configuration in interface view.

Table 20-15 Set IS-IS link state routing cost

Operation	Command
Set the routing cost of the interface	isis cost <i>value</i> [level-1 level-2]
Restore the default routing cost of the interface	undo isis cost [level-1 level-2]

If the level is not specified, the default setting is Level-1 routing cost.

The *value* argument is configured according to the link state of the interface.

By default, the routing cost of IS-IS on an interface is 10.

20.2.15 Configuring IS-IS Timers

I. Setting the Hello packet broadcast interval

The IS-IS periodically sends the Hello packets from the interface and the routers maintain the adjacency through the transmitting/receiving of the Hello packets. The Hello packet interval can be modified.

Perform the following configuration in interface view.

Table 20-16 Set the Hello packet broadcast interval

Operation	Command
Set Hello packet interval, measured in seconds.	isis timer hello <i>seconds</i> [level-1 level-2]
Restore the default Hello packet interval on the interface	undo isis timer hello [level-1 level-2]

Usually, on the broadcast links, there exist level-1 and level-2 hello packets. For different packets, different broadcast intervals should be set. However, there are two exceptions. One is when there is no level separation in the link, parameters of level-1 and level-2 need not be specified in the command (adopt the default values). So the system will set the broadcast intervals of all packets as that of the level-1 hello packet. The other is if hello packets are not separated according to level-1 and level-2 on the **p2p** links, the attribute of the packets need not be set either.

By default, Hello packets are transmitted on an interface every 10 seconds.

II. Setting the CSNP packet broadcast interval

The CSNP packet is transmitted by the DIS over the broadcast network to synchronize the link state database (LSDB). The CSNP packet is regularly broadcast over the broadcast network at an interval, which can be set by users.

Perform the following configuration in interface view.

Table 20-17 Set the CSNP packet broadcast interval

Operation	Command
Set the CSNP packet broadcast interval, measured in seconds	isis timer csnp <i>seconds</i> [level-1 level-2]
Restore the default CSNP packet broadcast interval on the interface	undo isis timer csnp [level-1 level-2]

If the level is not specified, it defaults to setting CSNP packet broadcast interval for Level-1.

By default, the CSNP packet is transmitted via interface every 10 seconds.

III. Setting the LSP packet transmission interval

LSP carries the link state records for propagation throughout the area.

Perform the following configuration in interface view.

Table 20-18 Set the LSP packet transmission interval

Operation	Command
Set LSP packet interval on the interface, measured in milliseconds.	isis timer lsp time
Restore the default LSP packet interval on the interface	undo isis timer lsp

By default, the LSP packet is transmitted via the interface every 33 milliseconds.

IV. Setting LSP packet retransmission interval

Over a **p2p** link, if the local end does not receive the response within a period of time after it sends an LSP packet, it considers that the originally transmitted LSP packet has been lost or dropped. In order to guarantee the transmission reliability, the local router will retransmit the original LSP packet.

Perform the following configuration in interface view.

Table 20-19 Set LSP packet retransmission interval

Operation	Command
Set the retransmission interval of the LSP packet over p2p links	isis timer retransmit seconds
Restore the default retransmission interval of the LSP packet over p2p links	undo isis timer retransmit

By default, the LSP packet is transmitted every five seconds over the **p2p** link.

V. Configuring number of invalid Hello packets for the interface

The router maintains the adjacency by sending/receiving Hello packets. When receiving no Hello packets from the peer within a time interval, the local router regards the neighbors are invalid. The time interval is called Holddown time for the IS-IS.

Setting invalid number of Hello packets can adjust the Holddown time in the IS-IS. That is to say, after continuously receiving no specified number of Hello packets, the router regards the neighbors are invalid.

Table 20-20 Set number of invalid Hello packets for the interface

Operation	Command
Set the number of invalid Hello packets	isis timer holding-multiplier value [level-1 level-2]
Restore the default setting	undo isis timer holding-multiplier [level-1 level-2]

By default, the number of the invalid Hello packets is set to 3.

If this command does not specify Level-1 or Level-2, the system regard the invalid Hello packets are set for both Level-1 and Level-2 routers.

20.2.16 Setting IS-IS Authentication

I. Setting interface authentication

The authentication password set on the interface is mainly used in the Hello packet so as to confirm the validity and correctness of its peers. The authentication passwords at the same level of all the interfaces of a network should be identical.

Perform the following configuration in interface view.

Table 20-21 Set interface authentication password

Operation	Command
Set authentication password	isis authentication-mode { simple md5 } <i>password</i> [{ level-1 level-2 } [ip osi]]
Delete authentication-mode password	undo isis authentication-mode { simple md5 } <i>password</i> [{ level-1 level-2 } [ip osi]]

By default, the interface is not configured with any authentication password nor performs authentication. If the level is not specified, it defaults to setting the authentication password of Level-1.

II. Setting IS-IS area or IS-IS routing domain authentication password

Users can configure the IS-IS area or the IS-IS routing domain with authentication password.

If area authentication is needed, the area authentication password will be encapsulated into the level-1 LSP, CSNP and PSNP packets, in the specified mode. If other routers in the same area also have started the area authentication, their authentication modes and passwords must be identical to those of their neighbors, so that they can work normally. Similarly, for domain authentication, the password will also be encapsulated into the level-2 LSP, CSNP and PSNP packets in the specified mode. If the routers in the backbone layer (level-2) also need domain authentication, their authentication mode and password must be identical to those of their neighbors.

Note that the passwords for authentication of the routers on the same network segment must be identical.

Perform the following configurations in IS-IS view.

Table 20-22 Set IS-IS authentication password

Operation	Command
Set authentication-mode password	area-authentication-mode { simple md5 } <i>password</i> [ip osi]
Delete authentication-mode password	undo area-authentication-mode { simple md5 } [ip osi]
Set routing domain authentication password	domain-authentication-mode { simple md5 } <i>password</i> [ip osi]
Delete routing domain authentication password	undo domain-authentication-mode { simple md5 } [ip osi]

By default, the system does not require password or perform authentication.

III. Setting the IS-IS to use the MD5 algorithm compatible with that of the other vendors

You must configure this command when the switch needs to authenticate the devices of other vendors using MD5 algorithm in IS-IS.

Perform the following configuration in IS-IS view.

Table 20-23 Set the IS-IS to use the MD5 algorithm compatible with that of the other vendors

Operation	Command
Set the IS-IS to use the MD5 algorithm compatible with that of the other vendors	md5-compatible
Set the IS-IS to use the default MD5 algorithm	undo md5-compatible

By default, the system uses the MD5 algorithm in IS-IS which is compatible with that of 3Com.

20.2.17 Setting the Mesh Group of the Interface

On a NBMA network, the interface of a router will flood the received LSP to other interfaces. However, this processing method applied to a network with higher connectivity and several **p2p** links will cause repeated LSP flooding and waste bandwidth.

To avoid such problem, you can configure several interfaces into a mesh group. The interface will flood it outside the group only.

Perform the following configuration in interface view.

Table 20-24 Set the mesh group of the interface

Operation	Command
Add an interface to a mesh group	isis mesh-group { <i>mesh-group-number</i> mesh-blocked }
Remove the interface from the mesh group	undo isis mesh-group

By default, the LSP is flooded normally from the interface. When configured with the **mesh-blocked** keyword, it will not flood the LSP to other interfaces.

Thus the IS-IS configuration tasks on the interface are finished. The following sections discuss how to configure other parameters of IS-IS.

20.2.18 Setting Overload Flag Bit

Sometimes, the router in the IS-IS domain may encounter some problems in operation thus errors may occur in the whole routing area. In order to avoid this problem, you can set the overload flag bit for this router.

When the overload threshold is set, other routers should not send this router the packets which should be forwarded by it.

Perform the following configurations in IS-IS view.

Table 20-25 Set overload flag bit

Operation	Command
Set overload flag bit	set-overload
Remove the overload flag bit	undo set-overload

By default, no over load bit is set.

20.2.19 Setting to Discard the LSPs with Checksum Errors

After receiving an LSP packet, the local IS-IS will calculate its checksum and compares the result with the checksum in the LSP packet. This process is the checksum authentication over the received LSP. By default, even when the checksum in the packet is not consistent with the calculated result, the LSP packet is not discarded. However, when not ignoring LSP checksum error is set with the **ignore-lsp-checksum-error** command, the LSP packet will be discarded if the checksum error is found.

Perform the following configuration in IS-IS view.

Table 20-26 Set to discard the LSPs with checksum errors

Operation	Command
Set to discard the LSP with checksum error	ignore-lsp-checksum-error
Set to ignore the LSP checksum error	undo ignore-lsp-checksum-error

By default, the LSP checksum error is ignored.

20.2.20 Setting to Log the Peer Changes

After peer changes log is enabled, the IS-IS peer changes will be output on the configuration terminal until the log is disabled.

Perform the following configuration in IS-IS view.

Table 20-27 Set to log the peer changes

Operation	Command
Enable peer changes log	log-peer-change
Disable peer changes log	undo log-peer-change

By default, the peer changes log is disabled.

20.2.21 Setting LSP Refreshment Interval

In order to ensure that the LSPs in the whole area can maintain the synchronization, all the current LSPs will be transmitted periodically.

Perform the following configuration in IS-IS view.

Table 20-28 Set LSP refreshment interval

Operation	Command
Set LSP refreshment interval	timer lsp-refresh <i>seconds</i>
Restore the default LSP refreshment interval	undo timer lsp-refresh

By default, LSP is refreshed every 900 seconds (15 minutes).

20.2.22 Setting Lifetime of LSP

When a router generates the LSP of the system, it will fill in the maximum lifetime of this LSP. When other routers receive this LSP, its life time will be reduced continuously as the time goes. If updated LSP has not been received before the old one times out, this LSP will be deleted from the LSDB.

Perform the following configuration in IS-IS view.

Table 20-29 Set Lifetime of LSP

Operation	Command
Set lifetime of LSP	timer lsp-max-age <i>seconds</i>
Restore the default LSP lifetime	undo timer lsp-max-age

By default, LSP can live for 1200 seconds (20 minutes).

20.2.23 Setting Parameters Related to SPF

I. Setting SPF calculation interval

When IS-IS LSDB changes, the router will compute the shortest path again. However, the immediate calculation upon every change will occupy too many resources and affect the efficiency of the router. In the case that SPF computing interval is set, when LSDB changes, SPF algorithm will be run after the SPF interval times out.

Perform the following configuration in IS-IS view.

Table 20-30 Set SPF calculation interval

Operation	Command
Set SPF calculation interval	timer spf <i>second</i> [level-1 level-2]
Restore default SPF calculation interval	undo timer spf [level-1 level-2]

If the level is not specified, it defaults to setting the SPF calculation interval of Level-1.

By default, SPF calculation runs every 10 seconds.

II. Setting SPF calculation in slice

When there is a large number of routes in the routing table (over 150,000), SPF calculation of IS-IS may occupy the system resources for a long time. To prevent such a case, SPF calculation can be set to perform in slice.

Perform the following configuration in IS-IS view.

Table 20-31 Set SPF calculation in slice

Operation	Command
Set the duration of one cycle in second of SPF calculation	spf-slice-size <i>seconds</i>
Restore the default configuration	undo spf-slice-size

By default, SPF calculation is not divided into slices but runs to the end once, which can also be implemented by setting the *seconds* argument to 0.

After slice calculation is set, the routes that are not processed once will be calculated in one second.

Normally, the user is not recommended to modify the default configuration. When the number of routes is between 150,000 and 200,000, it is recommended to set the *seconds* argument to 1, that is, the duration time for SPF calculation each time is 1 second.

III. Setting SPF to release CPU actively

To prevent SPF calculation from occupying the system resources for a long time, which affects the response speed of the console, SPF can be set to automatically release the system CPU resources after processing a certain number of routes and the unprocessed routes will be calculated in one second.

Perform the following configuration in IS-IS view.

Table 20-32 Set SPF to release CPU actively

Operation	Command
Specify the number of routes to process before releasing CPU	spf-delay-interval <i>number</i>
Restore the default configuration	undo spf-delay-interval

By default, CPU is released once when every 5000 routes are processed by the SPF of IS-IS.

20.2.24 Enabling/Disabling the Interface to Send Packets

To prevent the IS-IS routing information from being obtained by some router in a certain network, the **silent-interface** command can be used to prohibit sending IS-IS packets via the interface connecting with the router.

Perform the following configuration in IS-IS view.

Table 20-33 Enable/Disable the interface to send IS-IS packets

Operation	Command
Disable the interface to send IS-IS packets	silent-interface <i>silent-interface-type</i> <i>silent-interface-number</i>
Enable the interface to send IS-IS packets	undo silent-interface <i>silent-interface-type</i> <i>silent-interface-number</i>

By default, the interface is allowed to receive and send IS-IS packets.

The **silent-interface** command is only used to restrain the IS-IS packets not to be sent on the interface, but the interface routes can still be sent from other interfaces. On a switch, this command can disable/enable the specified VLAN interface to send IS-IS packets.

20.2.25 Resetting All the IS-IS Data Structure

When it is necessary to refresh some LSPs immediately, perform the following configuration in user view.

Table 20-34 Reset all the IS-IS data structures

Operation	Command
Reset the IS-IS data structure	reset isis all

By default, the IS-IS data structure is not cleared.

20.2.26 Resetting the Specified IS-IS Peer

When it is necessary to connect a specified peer again, perform the following configuration in user view.

Table 20-35 Reset the specified IS-IS peer

Operation	Command
Reset the specified IS-IS peer	reset isis peer <i>system-id</i>

By default, the IS-IS peer is not cleared.

20.3 Displaying and Debugging Integrated IS-IS

After completing the above configuration, execute the **display** command in any view to display the running state of the IS-IS configuration, and to verify the effect of the configuration. Execute the **debugging** command in user view to debug the IS-IS module.

Through the following configuration operations, you can view the LSDB of the IS-IS, the transmitting/receiving of various packets of the IS-IS and the SPF calculation so as to determine the IS-IS route maintenance conditions.

Table 20-36 Display and debug IS-IS

Operation	Command
Display IS-IS LSDB	display isis lsdb [[I1 I2 level-1 level-2] [[LSPID local] verbose]*]*

Display IS-IS SPF calculation log	display isis spf-log
Display IS-IS routing information	display isis route
Display IS-IS neighbor information	display isis peer [verbose]
Display mesh group information	display isis mesh-group
Enable IS-IS debugging	debugging isis { adjacency all authentication-error checksum-error circuit-information configuration-error datalink-receiving-packet datalink-sending-packet general-error interface-information memory-allocating receiving-packet-content self-originate-update sending-packet-content snp-packet spf-event spf-summary spf-timer task-error timer update-packet }
Disable IS-IS debugging	undo debugging isis { adjacency all authentication-error checksum-error circuit-information configuration-error datalink-receiving-packet datalink-sending-packet general-error interface-information memory-allocating receiving-packet-content self-originate-update sending-packet-content snp-packet spf-event spf-summary spf-timer task-error timer update-packet }

20.4 Typical Integrated IS-IS Configuration Example

I. Network requirements

As is shown in Figure 20-3, Switches A, B, C and D belong to the same autonomous system. The IS-IS routing protocol is running in these four switches so as to implement route interconnection. In the network design, switches A, B, C and D belong to the same area.

II. Network diagram

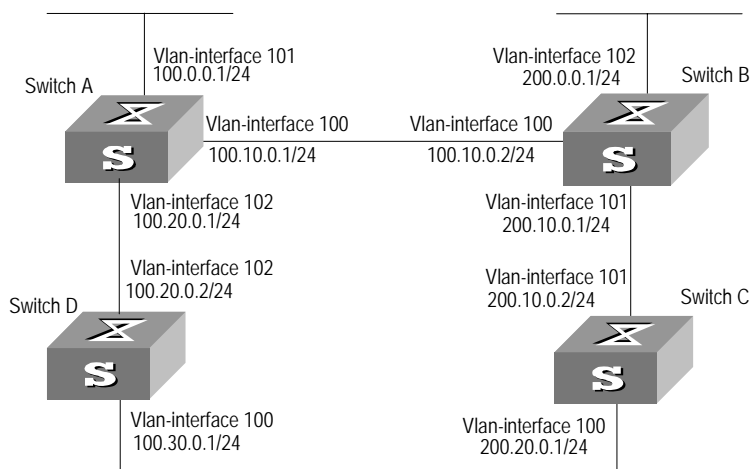


Figure 20-3 IS-IS configuration example

III. Configuration procedure

Configure Switch A

```
[Switch A] isis
[Switch A-isis] network-entity 86.0001.0000.0000.0005.00
[Switch A] interface vlan-interface 100
[Switch A-Vlan-interface100] ip address 100.10.0.1 255.255.255.0
[Switch A-Vlan-interface100] isis enable
[Switch A] interface vlan-interface 101
[Switch A-Vlan-interface101] ip address 100.0.0.1 255.255.255.0
[Switch A-Vlan-interface101] isis enable
[Switch A] interface vlan-interface 102
[Switch A-Vlan-interface102] ip address 100.20.0.1 255.255.255.0
[Switch A-Vlan-interface102] isis enable
```

Configure Switch B

```
[Switch B] isis
[Switch B-isis] network-entity 86.0001.0000.0000.0006.00
[Switch B] interface vlan-interface 101
[Switch B-Vlan-interface101] ip address 200.10.0.1 255.255.255.0
[Switch B-Vlan-interface101] isis enable
[Switch B] interface vlan-interface 102
[Switch B-Vlan-interface102] ip address 200.0.0.1 255.255.255.0
[Switch B-Vlan-interface102] isis enable
[Switch B] interface vlan-interface 100
[Switch B-Vlan-interface100] ip address 100.10.0.2 255.255.255.0
[Switch B-Vlan-interface100] isis enable
```

Configure Switch C

```
[Switch C] isis
[Switch C-isis] network-entity 86.0001.0000.0000.0007.00
[Switch C] interface vlan-interface 101
[Switch C-Vlan-interface101] ip address 200.10.0.2 255.255.255.0
[Switch C-Vlan-interface101] isis enable
[Switch C] interface vlan-interface 100
[Switch C-Vlan-interface100] ip address 200.20.0.1 255.255.255.0
[Switch C-Vlan-interface100] isis enable
```

Configure Switch D

```
[Switch D] isis
[Switch D-isis] network-entity 86.0001.0000.0000.0008.00
[Switch D] interface vlan-interface 102
[Switch D-Vlan-interface102] ip address 100.20.0.2 255.255.255.0
[Switch D-Vlan-interface102] isis enable
[Switch D] interface vlan-interface 100
[Switch D-Vlan-interface100] ip address 100.30.0.1 255.255.255.0
[Switch D-Vlan-interface100] isis enable
```

Chapter 21 BGP Configuration

21.1 BGP/MBGP Overview

21.1.1 Introduction to BGP

Border gateway protocol (BGP) is an inter-autonomous system (inter-AS) dynamic route discovery protocol. Three early versions of BGP are BGP-1 (RFC1105), BGP-2 (RFC1163) and BGP-3 (RFC1267). The current version is BGP-4 (RFC1771) that is applied to advertised structures and supports classless inter-domain routing (CIDR). Actually, BGP-4 is becoming the external routing protocol standard of the Internet, which is frequently used between ISPs.

The characteristics of BGP are as follows:

- BGP is an external gateway protocol (EGP). Different from such internal routing protocols as OSPF and RIP, it focuses on route propagation control and selection of best routes other than discovery and calculation of routes.
- It eliminates routing loop by adding AS path information to BGP routes.
- It enhances its own reliability by using TCP as the transport layer protocol.
- When routes are updated, BGP only transmits updated routes, which greatly reduces bandwidth occupation by route propagation and can be applied to propagation of a great amount of routing information on the Internet.
- BGP-4 supports CIDR, which is an important improvement to BGP-3.
- In consideration of management and security, users desire to perform control over outgoing and incoming routing information of each AS. BGP-4 provides abundant route policies to implement flexible filtering and selecting of routes.
- BGP-4 can be extended easily to support new developments of the network.

Note:

- CIDR handles IP addresses in an entirely new way, that is, it does not distinguish networks of Class A, Class B and Class C. For example, an invalid Class C network address 192.213.0.0 (255.255.0.0) can be expressed as 192.213.0.0/16 in CIDR mode, which is a valid super network. Here /16 means that the subnet mask is composed of the first 16 bits from the left.
 - The introduction of CIDR simplifies route aggregation. Actually, route aggregation is the process of aggregating several different routes, which turns advertisement processes of several routes to the advertisement of single route so as to simplify the routing table.
-

BGP runs on a router in any of the following modes:

- Internal BGP (IBGP)
- External BGP (EBGP)

The BGP is called IBGP when it runs in an AS and EBGP when it runs among different ASs.

21.1.2 BGP Message Types

BGP is driven by messages, which include the following types:

- Type 1, OPEN: The first message sent after the creation of a connection to create association between BGP peers.
- Type 2, UPDATE: The most important information in BGP system used to exchange routing information between peers. It is composed of up to three parts, that is, unreachable route, path attributes and network layer reachable information (NLRI).
- Type 3, NOTIFICATION: Used to notify errors.
- Type 4, KEEPALIVE: Used to check connectivity.
- Type 5, ROUTE-REFRESH: Used to advertise its own route refreshing capability.

The first four types are defined in RFC1771, while the last one is in RFC2918 (Route Refresh Capability for BGP-4).

21.1.3 BGP Routing Mechanism

On the first startup of the BGP system, the BGP router exchanges routing information with its peers by transmitting the complete BGP routing table, after that only update messages are exchanged. In the operating of the system, keepalive messages are received and transmitted to check the connections between various neighbors.

The router transmitting BGP messages is called a BGP speaker, which receives and generates new routing information continuously and advertises the information to the other BGP speakers. When a BGP speaker receives a new route advertisement from another AS, it will advertise the route, if the route is better than the current route that has been learned or is a new route, to all the other BGP speakers in the AS.

A BGP speaker calls peers other BGP speakers which exchange information with it and multiple related peers compose a peer group.

I. Route advertisement policy

In the Switch 8800, these policies are used by BGP when advertising routes:

- If there are multiple routes available, a BGP speaker only selects the optimum one.
- A BGP only advertises its own route to its peers.
- A BPG advertises the routes obtained from EBGP to all its BGP peers (including EBGP and IBGP peers).

- A BGP speaker does not advertise the routes obtained from IBGP to its IBGP peers.
- A BGP speaker advertises the routes obtained from IBGP to its IBGP peers (In the Switch 8800, BGP and IGP are asynchronous.)
- Once the connection is set up, a BGP speaker will advertise all its BGP routes to its peers.

II. Route selection policy

In the Switch 8800, these policies are adopted for BGP to select routes:

- First discard the routes unreachable to the next hop.
- First select the routes with the highest local preference.
- First select the routes rooted from the router itself.
- First select the routes with the least AS-paths.
- First select the routes with the lowest origin.
- First select the routes with the lowest MED value.
- First select the routes learned from EBGp.
- First select the routes advertised by the router with the lowest ID.

21.1.4 MBGP

I. MBGP overview

As described at the beginning of this chapter, BGP, as the practical exterior gateway protocol, is widely used in interconnection between autonomous systems. The traditional BGP-4 can only manage the routing information of IPv4 and has limitation in inter-AS routing when used in the application of other network layer protocols (such as IPv6 etc).

In order to support multiple network layer protocols, IETF extended BGP-4 and formed MBGP (Multiprotocol Extensions for BGP-4, multiple protocols extension of BGP-4). The present MBGP standard is RFC2858.

MBGP is backward compatible, that is, a router supporting BGP extension can be interconnected with a router that does not support it.

II. MBGP extension attributes

In the packets BGP-4 uses, three pieces of information related to IPv4 are carried in the update packet. They are Network Layer Reachability Information (NLRI), Next_Hop (The next hop address) in path attribute and Aggregator in path attribute (This attribute includes the BGP speaker address which forms the summary route).

When multiple network layer protocols are supported, it is necessary for BGP-4 to reflect the information of the specified network layer protocol to NLRI and the Next_Hop. Two new routing attributes are introduced in MBGP:

- **MP_REACH_NLRI:** Multiprotocol Reachable NLRI, used to advertise reachable routes and the next hop information.
- **MP_UNREACH_NLRI:** Multiprotocol Unreachable NLRI, used to delete unreachable routes.

These two attributes are optional non-transitive. Therefore, the BGP speaker that does not provide multiple protocols ability will ignore the information of them nor transfer them to other peers.

III. Address family

The network layer protocols are differentiated by address families in BGP. See RFC1700 (assigned numbers) for the possible values of these address families. The Switch 8800 provides various MBGP extended applications, including extension of multicast, VPN, and so on. Different extended applications should be configured in their own address family views.

For more information about the commands executed in MBGP address family view, see “Multicast Protocol” and “MPLS Configuration” of this manual.

21.1.5 BGP Peer and Peer Group

I. Definition of peer and peer group

A BGP speaker calls peers other BGP speakers which exchange information with it and multiple related peers compose a peer group.

II. Relationship between peer configuration and peer group configuration

In the Switch 8800, a BGP peer must belong to a peer group. If you want to configure a BGP peer, you need first to create a peer group and then add a peer into the group.

BGP peer group feature can simplify user configuration and improve route advertisement efficiency. When added into a peer group, a peer inherits all the configuration of the group.

If the configuration of a peer group changes, the configuration of its member peers also alters. Some attributes can be configured to a particular member peer by specifying its IP address. The attributes configured in this way is with higher priority than those by configuring for peer group. It should be noted that all member peers must use the same update policy as its group, but may use different ingress policy.

21.2 Configuring BGP

These categories are involved in BGP configuration:

- 1) Basic BGP configuration
 - Enabling BGP
 - Configuring Basic Features for BGP Peer
- 2) BGP peer configuration

- Configuring application features of a BGP peer (group)
- Configuring Route Filtering of a Peer (group)
- 3) BGP route configuration
 - Configuring Network Routes for BGP Distribution
 - Configuring the Interaction Between BGP and IGP
 - Configuring BGP Route Summarization
 - Configuring BGP Route Filtering
 - Configuring BGP Route Dampening
- 4) BGP protocol configuration
 - Configuring BGP Preference
 - Configuring BGP Timer
 - Configuring the Local Preference
 - Configuring MED for AS
- 5) BGP application configuration
 - Comparing the MED Routing Metrics from the Peers in Different ASs
- 6) BGP networking configuration
 - Configuring BGP Route Reflector
 - Configuring BGP AS Confederation Attribute
- 7) Others
 - Clearing BGP Connection
 - Refreshing BGP Routes

21.2.1 Enabling BGP

To enable BGP, local AS number should be specified. After the enabling of BGP, local router listens to BGP connection requests sent by adjacent routers. To make the local router send BGP connection requests to adjacent routers, refer to the configuration of the **peer** command. When BGP is disabled, all established BGP connections will be disconnected.

Perform the following configuration in system view.

Table 21-1 Enable/Disable BGP

Operation	Command
Enable BGP and enter the BGP view	bgp <i>as-number</i>
Disable BGP	undo bgp [<i>as-number</i>]

By default, BGP is not enabled.

21.2.2 Configuring Basic Features for BGP Peer

When configuring a MBGP peer (group), you should first configure AS ID for it and then enter the corresponding address family view to activate the association.

Perform the following configurations in BGP view.

I. Creating a peer group

A BGP peer must belong to a peer group. Before configuring a BGP peer, a peer group to which the peer belongs must be created first.

Table 21-2 Create a peer group

Operation	Command
Create a peer group	group <i>group-name</i> [internal external]
Delete the specified peer group	undo group <i>group-name</i>

There are two types of BGP peer group, IBGP and EBG. Using the **internal** keyword to create a IBGP peer group. You can use the **external** keyword to create an EBG peer group and sub-AS peer groups inside a confederation. *group-name* is locally significant.

The default type of BGP peer group is IBGP.

II. Configuring AS number of an EBG peer group

You can specify AS number for an EBG peer group, but IBGP needs no AS number. When a peer group is specified with an AS number, all its member peers inherit the AS number.

Table 21-3 Configure AS number of a EBG peer group

Operation	Command
Configure the AS number of the EBG peer group	peer <i>group-name</i> as-number <i>as-number</i>
Delete the AS number of the EBG peer group	undo peer <i>group-name</i> as-number <i>as-number</i>

If a peer group has peers, you cannot specify an AS number for the peer group. In addition, deleting the AS number of a peer group will delete all peers in it.

III. Adding a member to a peer group

A BGP peer must belong to a peer group. If you want to configure a BGP peer, you need first to create a peer group and then add a peer into the group.

Table 21-4 Create a peer group and add a member

Operation	Command
Add a peer to the peer group	peer <i>peer-address</i> group <i>group-name</i> [as-number <i>as-number</i>]

Delete a peer	undo peer <i>peer-address</i>
---------------	--------------------------------------

If you want to add a peer to an IBGP peer group, this command cannot specify AS numbers.

When a peer is added to an EBGP peer group and the peer group is defined with an AS number, all its member peers inherits the configuration of the group. If the AS number of the peer group is not specified, each peer added to it should be specified with its own AS number. AS numbers of peers in a same peer group can be different.

IV. Configuring the state of a peer/peer group

BGP peer/peer group has two types of state: enabled and disabled. The BGP speakers do not exchange routing information with the disabled peer or peer group.

Table 21-5 Configure the state of a peer/peer group

Operation	Command
Enable a peer/peer group	peer { <i>group-name</i> <i>peer-address</i> } enable
Disable a peer/peer group	undo peer { <i>group-name</i> <i>peer-address</i> } enable

By default, only BGP peer groups of IPv4 unicast address family are enabled. Other peer types or peer group types are disabled, consequently exchanging no routing information.

When exchanging routing information between BGP speakers, the peer group must be enabled first and then the peer should be added to the enabled peer group.

V. Configuring description of a peer (group)

Description of a peer (group) can be configured to facilitate network maintenance.

Table 21-6 Configure description of a peer (group)

Operation	Command
Configure description of a peer (group)	peer { <i>peer-address</i> <i>group-name</i> } description <i>description-line</i>
Delete description of a peer (group)	undo peer { <i>peer-address</i> <i>group-name</i> } description

By default, no BGP peer (group) description is set.

VI. Configuring timer of a peer (group)

The **peer timer** command is used to configure timers of a BGP peer (group), including the keep-alive message interval and the hold timer. The preference of this command is

higher than the **timer** command that is used to configure timers for the whole BGP peers.

Perform the following configuration in BGP view.

Table 21-7 Configure timer of a peer (group)

Operation	Command
Configure keep-alive message interval and hold timer of a peer (group)	peer { <i>group-name</i> <i>peer-address</i> } timer keep-alive <i>keepalive-interval</i> hold <i>holdtime-interval</i> }
Restore the default value of keep-alive message interval and hold timer of a peer (group)	undo peer { <i>group-name</i> <i>peer-address</i> } timer

By default, the keep-alive message is sent every 60 seconds and the value of the hold timer is 180 seconds.

VII. Configuring the interval at which route update messages are sent by a peer group

Table 21-8 Configure the interval at which route update messages are sent by a peer group

Operation	Command
Configure the route update message interval of a peer group	peer <i>group-name</i> route-update-interval <i>seconds</i>
Restore the default route update message interval of a peer group	undo peer <i>group-name</i> route-update-interval

By default, the intervals at which route update messages are sent by an IBGP and EBGP peer group are 5 seconds and 30 seconds respectively

21.2.3 Configuring application features of a BGP peer (group)

I. Configuring to permit connections with EBGP peer groups on indirectly connected networks

Generally, EBGP peers must be connected physically. Otherwise the command below can be used to perform the configuration to make them communicate with each other normally.

Perform the following configuration in BGP view.

Table 21-9 Configure to permit connections with EBGP peer groups on indirectly connected networks

Operation	Command
Configure to permit connections with EBGP peer groups on indirectly connected networks	peer <i>group-name</i> ebgp-max-hop [<i>tvl</i>]
Configure to permit connections with EBGP peer groups on directly connected network only	undo peer <i>group-name</i> ebgp-max-hop

By default, only the connections with EBGP peer groups on directly connected networks are permitted. *tvl* refers to time-to-live in the range of 1 to 255 with the default value as 64.

II. Configuring an IBGP peer group to be a client of a route reflector

Perform the following configuration in BGP view.

Table 21-10 Configure an IGMP peer group to be a client of a route reflector

Operation	Command
Configure a peer group to be a client of a route reflector	peer <i>group-name</i> reflect-client
Cancel the configuration of making the peer group as the client of the BGP route reflector	undo peer <i>group-name</i> reflect-client

This configuration can be applied to IBGP peer groups only.

By default, all IBGP peers in the autonomous system must be fully connected. Moreover, neighbors do not notify the learned IBGP routes.

III. Configuring to send default route to a peer group

If you only need to notify a default route between a pair of BGP peer instead of transmitting the default route within the whole network, you can use the **peer default-route-advertise** command.

Perform the following configuration in BGP view.

Table 21-11 Configure to send default route to a peer group

Operation	Command
Configure to send default route to a peer group	peer <i>group-name</i> default-route-advertise
Configure not to send default route to a peer group	undo peer <i>group-name</i> default-route-advertise

By default, a BGP speaker does not send default route to any peer group.

After you use the **peer default-route-advertise** command, the local router will send a default route with the next hop as itself to the peer unconditionally, even if there is no default route in BGP routing table.

IV. Configuring itself as the next hop when advertising routes

In general, when sending routes to the EBGP peer, the BGP speaker will set the next hop address of the routing information as the local address. When sending routes to the IBGP peer, the BGP speaker will not modify the next hop address.

In some networking conditions, when the routes are sent to the IBGP peer, you can configure the local address of the sender as the next hop, consequently ensuring the IBGP neighbors can find the correct next hop.

Perform the following configuration in BGP view.

Table 21-12 Configure itself as the next hop when advertising routes

Operation	Command
Configure itself as the next hop when advertising routes	peer <i>group-name</i> next-hop-local
Disable the specification of itself as the next hop when advertising routes	undo peer <i>group-name</i> next-hop-local

V. Removing private AS numbers while transmitting BGP update messages

Generally, the AS numbers (public AS numbers or private AS numbers) are included in the AS paths while transmitting BGP update messages. This command is used to configure certain outbound routers to ignore the private AS numbers while transmitting update messages.

Perform the following configuration in BGP view.

Table 21-13 Remove private AS numbers while transmitting BGP update messages

Operation	Command
Remove private AS numbers while transmitting BGP update messages	peer <i>group-name</i> public-as-only
Include private AS numbers while transmitting BGP update messages	undo peer <i>group-name</i> public-as-only

By default, the private AS numbers are included during BGP update messages transmission.

The configuration can only be applied to the peer group.

VI. Configuring to send the community attributes to a peer group

Perform the following configuration in BGP view.

Table 21-14 Configure to send the community attributes to a peer group

Operation	Command
Configure to send the community attributes to a peer group	peer <i>group-name</i> advertise-community
Configure not to send the community attributes to a peer group	undo peer <i>group-name</i> advertise-community

By default, the BGP speaker does not send the community attributes to a peer group.

VII. Configuring the repeating time of local AS

BGP records the passed AS numbers in the routing information, and checks route loop depending on whether the AS number are repeated. In some special applications, it is allowed to receive the routing information with the repeated AS number.

Perform the following configuration in BGP view.

Table 21-15 Configure the repeating time of local AS

Operation	Command
Configure the repeating time of local AS	peer { <i>group-name</i> <i>peer-address</i> } allow-as-loop [<i>number</i>]
Remove the repeating time of local AS	undo peer { <i>group-name</i> <i>peer-address</i> } allow-as-loop

By default, the allowed repeating time of local AS is set to 1.

VIII. Specifying the source interface of a route update packet

Generally, the system specified the source interface of a route update packet. When the interface fails to work, in order to keep the TCP connection valid, the interior BGP session can be configured to specify the source interface. This command is usually used on the Loopback interface.

Table 21-16 Specify the source interface of a route update packet

Operation	Command
Specify the source interface of a route update packet	peer { <i>peer-address</i> <i>group-name</i> } connect-interface <i>interface-type interface-name</i>
Use the best source interface	undo peer { <i>peer-address</i> <i>group-name</i> } connect-interface <i>interface-type interface-name</i>

By default, BGP uses the interface to establish BGP links for the source interface of a route update packet.

IX. Configuring BGP MD5 authentication password

BGP uses TCP as its transport layer. For the sake of high security, you can configure MD5 authentication password when setting up a TCP connection. In other words, BGP MD5 authentication just sets password for TCP connection, but not for authenticating BGP packets. The authentication is implemented by TCP.

Perform the following configuration in BGP view.

Table 21-17 Configure BGP MD5 authentication

Operation	Command
Configure MD5 authentication password	peer { <i>group-name</i> <i>peer-address</i> } password { cipher simple } <i>password</i>
Cancel MD5 authentication	undo peer { <i>group-name</i> <i>peer-address</i> } password

In BGP, no MD5 authentication is performed in setting up TCP connections by default.

Note:

The multicast extension configured in BGP view is also available in MBGP, since they use the same TCP link.

21.2.4 Configuring Route Filtering of a Peer (group)

The Switch 8800 supports filtering imported and advertised routes for peers (groups) through Route-policy, AS path list, ACL and ip prefix list.

The route filtering policy of advertised routes configured for each member of a peer group must be same with that of the peer group but their route filtering policies of ingress routes may be different.

Perform the following configuration in BGP view.

I. Configuring route policy for a peer (group)

Table 21-18 Configure route policy for a peer (group)

Operation	Command
Configure the ingress route policy for a peer (group)	peer { <i>peer-address</i> <i>group-name</i> } route-policy <i>route-policy-name</i> import
Remove the ingress route policy of a peer (group)	undo peer { <i>peer-address</i> <i>group-name</i> } route-policy <i>policy-name</i> import

Operation	Command
Configure the egress route policy for a peer group	peer <i>group-name</i> route-policy <i>route-policy-name</i> export
Remove the egress route policy of a peer group	undo peer <i>group-name</i> route-policy <i>route-policy-name</i> export

II. Configuring route filtering policy based on IP ACL for a peer (group)

Table 21-19 Configure route filtering policy based on IP ACL for a peer (group)

Operation	Command
Configure the ingress route filtering policy based on IP ACL for a peer (group)	peer { <i>peer-address</i> <i>group-name</i> } filter-policy <i>acl-number</i> import
Remove the ingress route filtering policy based on IP ACL of a peer (group)	undo peer { <i>peer-address</i> <i>group-name</i> } filter-policy <i>acl-number</i> import
Configure the egress route filtering policy based on IP ACL for a peer (group)	peer <i>group-name</i> filter-policy <i>acl-number</i> export
Remove the egress route filtering policy based on IP ACL for a peer (group)	undo peer <i>group-name</i> filter-policy <i>acl-number</i> export

III. Configuring route filtering policy based on AS path list for a peer (group)

Table 21-20 Configure route filtering policy based on AS path list for a peer (group)

Operation	Command
Configure the ingress route filtering policy based on AS path list for a peer (group)	peer { <i>peer-address</i> <i>group-name</i> } as-path-acl <i>acl-number</i> import
Remove the ingress route filtering policy based on AS path list of a peer (group)	undo peer { <i>peer-address</i> <i>group-name</i> } as-path-acl <i>acl-number</i> import
Configure the egress route filtering policy based on IP ACL for a peer group	peer <i>group-name</i> as-path-acl <i>acl-number</i> export
Remove the egress route filtering policy based on IP ACL for a peer group	undo peer <i>group-name</i> as-path-acl <i>acl-number</i> export

The *acl-number* argument indicates AS path list number, which you can use the **acl** command instead of the **ip as-path-acl** command to configure. For the detailed configuration, refer to Chapter 22 “IP Routing Policy Configuration”.

IV. Configuring route filtering policy based on address prefix list for a peer (group)

Table 21-21 Configure route filtering policy based on address prefix list for a peer (group)

Operation	Command
Configure the ingress route filtering policy based on address prefix list for a peer (group)	peer { <i>peer-address</i> <i>group-name</i> } ip-prefix <i>prefixname</i> import
Remove the ingress route filtering policy based on address prefix list of a peer (group)	undo peer { <i>peer-address</i> <i>group-name</i> } ip-prefix <i>prefixname</i> import
Configure the egress route filtering policy based on address prefix list for a peer group	peer <i>group-name</i> ip-prefix <i>prefixname</i> export
Remove the egress route filtering policy based on address prefix list for a peer group	undo peer <i>group-name</i> ip-prefix <i>prefixname</i> export

By default, route filtering based on address prefix list for a peer (group) is disabled.

21.2.5 Configuring Network Routes for BGP Distribution

Perform the following configuration in BGP view.

Table 21-22 Configure network routes for BGP distribution

Operation	Command
Configure the local network route for BGP distribution	network <i>ip-address</i> <i>address-mask</i> [route-policy <i>route-policy-name</i>]
Remove the local network route for BGP distribution	undo network <i>ip-address</i> <i>address-mask</i> [route-policy <i>route-policy-name</i>]

By default, no network route is configured for BGP distribution.

21.2.6 Configuring the Interaction Between BGP and IGP

I. Importing IGP routes

BGP can transmit the internal network information of local AS to other AS. To reach such objective, the network information about the internal system learned by the local router via IGP routing protocol can be transmitted.

Perform the following configuration in BGP view.

Table 21-23 Import IGP routing information

Operation	Command
Configure BGP to import routes of IGP protocol	import-route <i>protocol</i> [<i>process-id</i>] [med <i>med</i>] [route-policy <i>route-policy-name</i>]
Configure BGP not to import routes of IGP protocol	undo import-route <i>protocol</i>

The *protocol* argument specifies the imported source route protocols. The specified and imported source route protocols can be direct, static, rip, isis, ospf, ospf-ase, and ospf-nssa.

By default, BGP does not import the route information of other protocols.

After you configure the **import-route** command in a BGP view, you cannot import the default route of the imported source route protocols to BGP by default.

II. Configuring not to synchronize with IGP

If the local BGP is not set synchronous with the IGP and the next hop of the learned BGP route is reachable, the local BGP will add this BGP route into its routing table immediately after it learns the route, rather than waiting till the IGP also learns the route.

Perform the following configuration in BGP view.

Table 21-24 Configure not to synchronize with IGP

Operation	Command
Cancel the synchronization of BGP and IGP	undo synchronization

By default, BGP does not synchronize with IGP. The Switch 8800 does not support synchronization of BGP and IGP.

21.2.7 Configuring BGP Route Summarization

There are two modes of BGP route summarization:

- **summary**: The summary of the BGP subnet routes. After the configuration of the **summary**, the BGP will not be able to receive subnets imported by the IGP;
- **aggregate**: The aggregation of the BGP local routes. In general, the preference of the aggregation is higher than that of the summarization.

Perform the following configuration in BGP view.

Table 21-25 Configure BGP route summarization

Operation	Command
Configure the summary automatic function of the subnet routes	summary
Cancel the summary automatic function of the subnet routes	undo summary
Configure local route aggregation function	aggregate <i>address mask</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>]*
Cancel local route aggregation function	undo aggregate <i>address mask</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>]*

By default, the BGP will not perform local route aggregation.

21.2.8 Configuring BGP Route Filtering

I. Configuring BGP to filter the received route information

The routes received by the BGP can be filtered, and only those routes that meet the certain conditions will be received by the BGP.

Perform the following configuration in BGP view.

Table 21-26 Configure imported route filtering

Operation	Command
Configure received route filtering	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> [gateway <i>ip-prefix-name</i>] } import
Cancel the received route filtering	undo filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> [gateway <i>ip-prefix-name</i>] } import
Filter the received global routing information	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } import
Cancel the received global route filtering	undo filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } import

By default, the BGP will not filter the received routes.

II. Configuring to filter the routes advertised by the BGP

Perform the following configuration in the BGP view.

Table 21-27 Configure to filter the routes advertised by the BGP

Operation	Command
Configure to filter the routes advertised by the BGP	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } export [<i>routing-process</i>]
Cancel the filtering of the routes advertised by the BGP	undo filter-policy <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } export [<i>routing-process</i>]

By default, BGP does not receive the routing information advertised by other routing protocols.

Note:

- The **filter-policy import** command filters BGP route received from the neighbors. The routes that cannot pass the filter will not be added to the routing table, and will not be advertised to the neighbors.
 - The **filter-policy export** command filters all the advertised routes, including routes imported by using the **import-route** command, and BGP routes learned from the neighbors.
 - If the **filter-policy export** command does not specify which route to be filtered, then the all the routes imported by the **import-route** command and the advertised BGP routes will be filtered.
-

21.2.9 Configuring BGP Route Dampening

I. Configure BGP route dampening

The main possible reason for unstable route is the intermittent disappearance and re-emergence of the route that formerly existed in the routing table, and this situation is called the flapping. When the flapping occurs, update packet will be propagated on the network repeatedly, which will occupy much bandwidth and much processing time of the router. You have to find measures to avoid it. The technology controlling unstable route is called route dampening.

The dampening divides the route into the stable route and unstable route, the latter of which shall be suppressed (not to be advertised). The history performance of the route is the basis to evaluate the future stability. When the route flapping occurs, penalty will be given, and when the penalty reaches a specific threshold, the route will be

suppressed. With time going, the penalty value will decrease according to power function, and when it decreases to certain specific threshold, the route suppression will be eliminated and the route will be re-advertised.

Perform the following configuration in BGP view.

Table 21-28 Configure BGP route dampening

Operation	Command
Configure BGP route dampening	dampening [<i>half-life-reachable</i> <i>half-life-unreachable</i> <i>reuse</i> <i>suppress</i> <i>ceiling</i>] [route-policy <i>route-policy-name</i>]
Clear route attenuation information and eliminating the suppression of the route	reset dampening [<i>network-address</i> [<i>mask</i>]]
Cancel BGP route dampening	undo dampening

By default, route dampening is disabled.

II. Clear route attenuation information

Perform the following configuration in user view to clear route attenuation information.

Table 21-29 Clear route attenuation information

Operation	Command
Clear route attenuation information	reset bgp dampening [<i>network-address</i> [<i>mask</i>]]

After you use the **reset bgp dampening** command, the command will release the suppression of a suppressed route.

21.2.10 Configuring BGP Preference

Three types of routes may be involved in BGP: routes learned from external peers, routes learned from internal peers and local-originated routes. You can set preference values for the three types of route.

Perform the following configuration in BGP view.

Table 21-30 Configure BGP preference

Operation	Command
Configure BGP preference	preference <i>ebgp-value</i> <i>ibgp-value</i> <i>local-value</i>
Restore the default preference	undo preference

The *ebgp-value*, *ibgp-value* and *local-value* arguments are in the range of 1 to 256. By default, the first two is 256 and the last one is 130.

21.2.11 Configuring BGP Timer

After you established BGP connections between routers, a router sends Keepalive packets to the peer periodically. Otherwise, the routers regard the BGP connections are interrupted. If the router receives no Keepalive packets or any other types of packets within the set connection holdtime, the router regards the BGP connection has been interrupted and quits the BGP connection.

When a router establishes a BGP connection with the peer, the router will compare their holdtime and regard the smaller time as the negotiated holdtime. If the negotiation result is 0, the router does not send Keepalive packets and detect whether the holdtime exceeds.

Perform the following configuration in BGP view.

Table 21-31 Configure BGP timers

Operation	Command
Configure BGP timers	timer keep-alive <i>keepalive-interval</i> hold <i>holdtime-interval</i>
Restore the default timer value	undo timer

By default, the interval of sending keepalive packet is 60 seconds. The interval of sending holdtime packet is 180 seconds.

The reasonable maximum interval of sending Keepalive packets is one third of the interval of sending holdtime packet. The interval of sending Keepalive packets cannot be less than 1 second. As a result, if the holdtime is not 0 second, the minimum holdtime is 3 seconds.

21.2.12 Configuring the Local Preference

When BGP select routes, it will select the route of the highest local preference.

Perform the following configuration in BGP view.

Table 21-32 Configure the local preference

Operation	Command
Configure the local preference	default local-preference <i>value</i>
Restore the default local preference	undo default local-preference

The local preference is transmitted only when the IBGP peers exchange the update packets and it will not be transmitted beyond the local AS.

By default, the local preference is 100.

21.2.13 Configuring MED for AS

Multi-Exit Discriminators (MED) attribute is the external metric for a route. AS uses the local preference to select the route to the outside, and uses the MED to determine the optimum route for entering the AS. When a router running BGP gets routes with the same destination address but different next hops through different external peers, it will select the route of the smallest MED as the optimum route, provided that all the other conditions are the same.

Perform the following configuration in BGP view.

Table 21-33 Configure an MED metric for the system

Operation	Command
Configure an MED metric for the system	default med <i>med-value</i>
Restore the default MED metric of the system	undo default med

By default, MED metric is 0.

The router configured above only compares the route MED metrics of different EBGP peers in the same AS. Using the **compare-different-as-med** command, you can compare the route MED metrics of the peers in different ASs.

21.2.14 Comparing the MED Routing Metrics from the Peers in Different ASs

It is used to select the best route. The route with smaller MED value will be selected.

Perform the following configuration in BGP view.

Table 21-34 Compare the MED routing metrics from the peers in different ASs

Operation	Command
Compare the MED routing metrics from the peers in different ASs	compare-different-as-med
Configure not to compare the MED routing metrics from the peers in different ASs	undo compare-different-as-med

By default, MED comparison is not allowed among the routes from the neighbors in different ASs.

It is not recommended to use this configuration unless you can make sure that the ASs adopt the same IGP and routing method.

21.2.15 Configuring BGP Route Reflector

To ensure the interconnection between IBGP peers, it is necessary to establish a fully connected network. If there are many IBGP peers, large overhead is needed to establish a fully connected network.

Route reflecting can solve the problem. Route reflector is the centralized point of other routers, and other routers are called the clients. The client is the peer of the route reflector and switching the routing information with it. The route reflector will reflect the information in order among the clients.

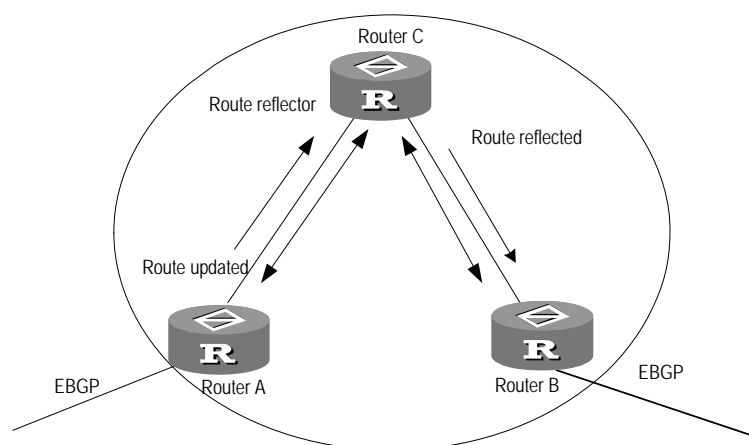


Figure 21-1 The route reflector diagram

In Figure 21-1, Router C is a route reflector with two peer clients: Router A and Router B. Router A sends to Router C the update packet from an external peer. Router C sends the update packet to Router B. After using reflecting technology, you do not need to establish a connection between Router A and Router B. You only need to connect Router C to Router A and Router B respectively.

If a BGP router is not either a reflector or client, we call the BGP router non-client. You still need connect non-clients to reflectors and non-clients.

You only need to configure route reflecting for the route reflector. When configuring the route reflector, you must specify the routers to serve as clients.

I. Configuring the route reflection between clients

Perform the following configuration in BGP view.

Table 21-35 Configure the route reflection between clients

Operation	Command
Enable route reflection between clients	reflect between-clients
Disable route reflection between clients	undo reflect between-clients

By default, the route reflection between clients is allowed. If the clients are fully connected, for the purpose of overhead reduction, it is recommended to use the **undo reflect between-clients** command to disable the route reflection between clients.

II. Configuring the cluster ID

Generally, there is only one route reflector in a cluster which is identified by the router ID of the route reflector.

Perform the following configuration in BGP view.

Table 21-36 Configure the Cluster_ID of the route reflector

Operation	Command
Configure the Cluster_ID of the route reflector	reflect cluster-id { <i>cluster-id</i> <i>address</i> }
Cancel the Cluster_ID of the route reflector	undo reflect cluster-id

The autonomous system possibly generates routing loop due to the route reflector in a cluster. After leaving a cluster, a routing update packet possibly tries to go back to the cluster. Because the routing update packet has not left an AS, the traditional AS path method cannot detect the loop inside the AS. When configuring route reflectors, you can use the following two methods to avoid loop inside the AS. One is to use the cluster ID; the other is to use Originator_ID of a route reflector.

If you configure Originator_ID improperly, the originator will discard the update packet when the update packet goes back to the originator. You do not need to configure Originator_ID. Originator_ID automatically takes effect when BGP is enabled.

21.2.16 Configuring BGP AS Confederation Attribute

Confederation provides the method to handle the booming IBGP network connections inside AS. It divides the AS into multiple sub-AS, in each of which all IBGP peers are fully connected, and are connected with other sub-AS of the confederation.

The shortcomings of confederation are that it is required that the route be re-configured upon switching from non-confederation to confederation solution, and that the logic topology be basically changed. Furthermore, the path selected via confederation may not be the best path if there is no manually-set BGP policy.

I. Configuring confederation_ID

In the eye of the BGP speakers that are not included in the confederation, multiple sub-ASs that belong to the same confederation are a whole. The external network does not need to know the status of internal sub-ASs, and the confederation ID is the AS number identifying the confederation as a whole.

Perform the following configuration in BGP view.

Table 21-37 Configure confederation_ID

Operation	Command
Configure confederation_ID	confederation id <i>as-number</i>
Cancel confederation_ID	undo confederation id

By default, the confederation_ID is not configured.

The configured confederation_ID and the existing AS number of a peer or peer group cannot be the same.

II. Configuring sub-AS belonging to the confederation

Configure confederation_ID first, and then configure the sub-AS belonging to the confederation. One confederation includes up to 32 sub-AS.

Perform the following configuration in BGP view.

Table 21-38 Configure sub-AS belonging to the confederation

Operation	Command
Configure a confederation consisting of which sub-ASs	confederation peer-as <i>as-number-1</i> [... <i>as-number-n</i>]
Cancel the specified sub-AS in the confederation	undo confederation peer-as [<i>as-number-1</i>] [... <i>as-number-n</i>]

By default, no autonomous system is configured as a member of the confederation.

The configured sub-AS number is valid only inside the confederation. In addition, the number cannot be the same as the AS number of a peer in the peer group for which you have not configured an AS number.

III. Configuring AS confederation attribute compatible with nonstandard

If it is necessary to perform the interconnection with the devices whose implementation mechanism is different from that of RFC1965, you must configure all the routers in the confederation.

Perform the following configuration in BGP view.

Table 21-39 Configure AS confederation attribute compatible with nonstandard

Operation	Command
Configure AS confederation attribute compatible with nonstandard router	confederation nonstandard
Cancel AS confederation attribute compatible with nonstandard router	undo confederation nonstandard

By default, the configured confederation is consistent with RFC1965.

21.2.17 Clearing BGP Connection

After the user changes BGP policy or protocol configuration, they must cut off the current connection so as to enable the new configuration.

Perform the following configuration in user view.

Table 21-40 Clear BGP connection

Operation	Command
Clear the connection between BGP and the specified peers	reset bgp <i>peer-address</i> [flap-info]
Clear all connections of BGP	reset bgp all
Clear the connections between the BGP and all the members of a group	reset bgp group <i>group-name</i>

21.2.18 Refreshing BGP Routes

It is required to re-compute associated route information when BGP routing policy changes.

Perform the following configuration in user view.

Table 21-41 Refresh BGP routes

Operation	Command
Refresh general BGP routes	refresh bgp { all <i>peer-address</i> group <i>group-name</i> } { import export }

The **import** keyword means to refresh the routes learned from the peers and the **export** keyword means to refresh routes advertised to the peers.

21.3 Displaying and Debugging BGP

After the above configuration, execute the **display** command in any view to display the running of the BGP configuration, and to verify the effect of the configuration. Execute the **reset** command in user view to clear the statistics of the configuration. Execute the **debugging** command in user view to debug the configuration. Execute the **reset** command in user view to reset the statistic information of BGP.

Table 21-42 Display and debug BGP

Operation	Command
Display the routing information in BGP routing table	display bgp routing-table [<i>ip-address</i> [<i>mask</i>]]
Display filtered AS path information in the BGP	display ip as-path-acl <i>acl-number</i>
Display CIDR routes	display bgp routing-table cidr
Display the routing information of the specified BGP community	display bgp routing-table community [<i>aa:nn</i> no-export-subconfed no-advertise no-export]* [whole-match]
Display the routing information allowed by the specified BGP community list	display bgp routing-table community-list <i>community-list-number</i> [whole-match]
Display BGP dampened paths	display bgp routing-table dampened
Display the routing information the specified BGP peer advertised or received	display bgp routing-table peer <i>peer-address</i> { advertised received } [<i>network-address</i> [<i>mask</i>] statistic]
Display the routes matching with the specified access-list	display bgp routing-table as-path-acl <i>acl-number</i>
Display route flapping statistics information	display bgp routing-table flap-info [{ regular-expression <i>as-regular-expression</i> } { as-path-acl <i>acl-number</i> } { <i>network-address</i> [<i>mask</i>] } { longer-match }]]]]
Display routes with different source ASs	display bgp routing-table different-origin-as
Display peers information	display bgp peer <i>peer-address</i> verbose display bgp peer [verbose]
Display the configured routing information	display bgp network
Display AS path information	display bgp paths <i>as-regular-expression</i>
Display peer group information	display bgp group [<i>group-name</i>]
Display the information on BGP routes which is mapped to a certain regular expression	display bgp routing-table regular-expression <i>as-regular-expression</i>
Display configured route-policy information	display route-policy [<i>policy-name</i>]
Enable/Disable information debugging of all BGP packets	[undo] debugging bgp all
Enable/Disable BGP event debugging	[undo] debugging bgp event

Operation	Command
Enable/Disable BGP Keepalive debugging	[undo] debugging bgp keepalive [receive send] [verbose]
Enable/Disable BGP Open debugging	[undo] debugging bgp open [receive send] [verbose]
Enable /Disable BGP packet debugging	[undo] debugging bgp packet [receive send] [verbose]
Enable/Disable BGP Update packet debugging	[undo] debugging bgp route-refresh [receive send] [verbose]
Enable/Disable information debugging of BGP normal functions.	[undo] debugging bgp normal
Enable/Disable BGP Update packet debugging	[undo] debugging bgp update [receive send] [verbose]
Reset BGP flap information	reset bgp flap-info [regular-expression <i>as-regular-expression</i> as-path-acl <i>acl-number</i> <i>network-address</i> [<i>mask</i>]]

21.4 Typical BGP Configuration Example

21.4.1 Configuring BGP AS Confederation Attribute

I. Network requirements

Divide the following AS 100 into three sub-AS: 1001, 1002, and 1003, and configure EBGP, confederation EBGP, and IBGP.

II. Network diagram

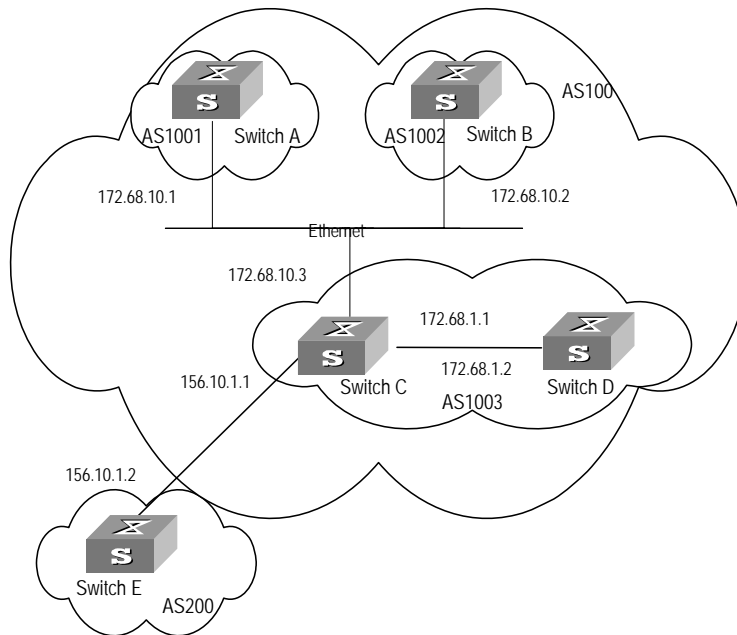


Figure 21-2 Network diagram for AS confederation configuration

III. Configuration procedure

Configure Switch A:

```
[Switch A] bgp 1001
[Switch A-bgp] confederation id 100
[Switch A-bgp] confederation peer-as 1002 1003
[Switch A-bgp] group confed1002 external
[Switch A-bgp] peer confed1002 as-number 1002
[Switch A-bgp] group confed1003 external
[Switch A-bgp] peer confed1003 as-number 1003
[Switch A-bgp] peer 172.68.10.2 group confed1002
[Switch A-bgp] peer 172.68.10.3 group confed1003
```

Configure Switch B:

```
[Switch B] bgp 1002
[Switch B-bgp] confederation id 100
[Switch B-bgp] confederation peer-as 1001 1003
[Switch B-bgp] group confed1001 external
[Switch B-bgp] peer confed1001 as-number 1001
[Switch B-bgp] group confed1003 external
[Switch B-bgp] peer confed1003 as-number 1003
[Switch B-bgp] peer 172.68.10.1 group confed1001
[Switch B-bgp] peer 172.68.10.3 group confed1003
```


Configure Switch C:

```
[Switch C] bgp 1003
[Switch C-bgp] confederation id 100
[Switch C-bgp] confederation peer-as 1001 1002
[Switch C-bgp] group confed1001 external
[Switch C-bgp] peer confed1001 as-number 1001
[Switch C-bgp] group confed1002 external
[Switch C-bgp] peer confed1002 as-number 1002
[Switch C-bgp] peer 172.68.10.1 group confed1001
[Switch C-bgp] peer 172.68.10.2 group confed1002
[Switch C-bgp] group ebgp200 external
[Switch C-bgp] peer 156.10.1.2 group ebgp200 as-number 200
[Switch C-bgp] group ibgp1003 internal
[Switch C-bgp] peer 172.68.1.2 group ibgp1003
```

21.4.2 Configuring BGP Route Reflector**I. Network requirements**

Switch B receives an update packet passing EBGP and transmits it to Switch C. Switch C is a reflector with two clients: Switch B and Switch D. When Switch C receives a route update from Switch B, it will transmit such information to Switch D. It is required to establish an IBGP connection between Switch B and Switch D, because Switch C reflects information to Switch D.

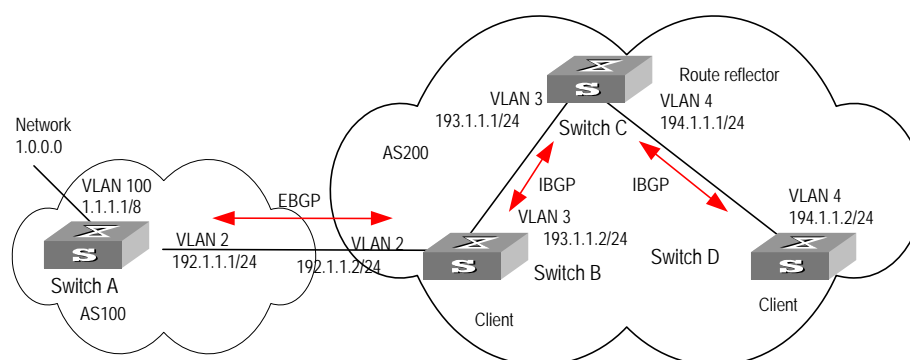
II. Network diagram

Figure 21-3 Network diagram for BGP route reflector configuration

III. Configuration procedure**1) Configure Switch A:**

```
[Switch A] interface vlan-interface 2
[Switch A-Vlan-interface2] ip address 192.1.1.1 255.255.255.0
[Switch A-Vlan-interface2] interface Vlan-interface 100
[Switch A-Vlan-interface100] ip address 1.1.1.1 255.0.0.0
```

```
[Switch A-Vlan-interface100] quit
[Switch A] bgp 100
[Switch A-bgp] network 1.0.0.0 255.0.0.0
[Switch A-bgp] group ex external
[Switch A-bgp] peer 192.1.1.2 group ex as-number 200
```

2) Configure Switch B:

Configure VLAN 2:

```
[Switch B] interface Vlan-interface 2
[Switch B-Vlan-interface2] ip address 192.1.1.2 255.255.255.0
```

Configure VLAN 3:

```
[Switch B] interface Vlan-interface 3
[Switch B-Vlan-interface3] ip address 193.1.1.2 255.255.255.0
```

Configure BGP peers.

```
[Switch B] bgp 200
[Switch B-bgp] group ex external
[Switch B-bgp] peer 192.1.1.1 group ex as-number 100
[Switch B-bgp] group in internal
[Switch B-bgp] peer 193.1.1.1 group in
```

3) Configure Switch C:

Configure VLAN 3:

```
[Switch C] interface Vlan-interface 3
[Switch C-Vlan-interface3] ip address 193.1.1.1 255.255.255.0
```

Configure VLAN 4:

```
[Switch C] interface vlan-Interface 4
[Switch C-Vlan-interface4] ip address 194.1.1.1 255.255.255.0
```

Configure BGP peers and route reflector.

```
[Switch C] bgp 200
[Switch C-bgp] group rr internal
[Switch C-bgp] peer rr reflect-client
[Switch C-bgp] peer 193.1.1.2 group rr
[Switch C-bgp] peer 194.1.1.2 group rr
```

4) Configure Switch D:

Configure VLAN 4:

```
[Switch D] interface vlan-interface 4
[Switch D-Vlan-interface4] ip address 194.1.1.2 255.255.255.0
```

Configure BGP peers

```
[Switch D] bgp 200
group in internal
[Switch D-bgp] peer 194.1.1.1 group in
```

Using the **display bgp routing-table** command, you can view BGP routing table on Switch B. Note: Switch B has known the existence of network 1.0.0.0.

Using the **display bgp routing-table** command ,you can view the BGP routing table on Switch D. Note: Switch D also knows the existence of network 1.0.0.0.

21.4.3 Configuring BGP Routing

I. Network requirements

This example illustrates how the administrators manage the routing via BGP attributes. All switches are configured with BGP, and IGP in AS 200 utilizes OSPF. Switch A is in AS 100, and Switch B, Switch C and Switch D are in AS 200. Switch A, Switch B, and Switch C operate EBGP. Switch B, Switch C and Switch D operate IBGP.

II. Network diagram

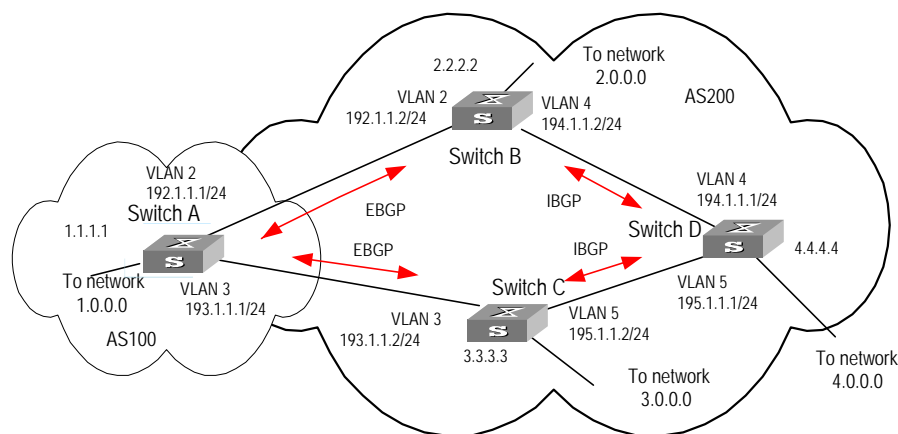


Figure 21-4 Networking diagram for BGP routing configuration

III. Configuration procedure

1) Configure Switch A:

```
[Switch A] interface Vlan-interface 2
[Switch A-Vlan-interface2] ip address 192.1.1.1 255.255.255.0
[Switch A] interface Vlan-interface 3
[Switch A-Vlan-interface3] ip address 193.1.1.1 255.255.255.0
```

Enable BGP

```
[Switch A] bgp 100
```

Specify the network that BGP sends to

```
[Switch A-bgp] network 1.0.0.0
```

Configure the peers

```
[Switch A-bgp] group ex192 external
[Switch A-bgp] peer 192.1.1.2 group ex192 as-number 200
```

```
[Switch A-bgp] group ex193 external
[Switch A-bgp] peer 193.1.1.2 group ex193 as-number 200
[Switch A-bgp] quit
```

Configure the MED attribute of Switch A

- Add ACL on Switch A, enable network 1.0.0.0.

```
[Switch A] acl number 2000
[Switch A-acl-basic-2000] rule permit source 1.0.0.0 0.255.255.255
[Switch A-acl-basic-2000] rule deny source any
```

- Define two route policies, one is called `apply_med_50` and the other is called `apply_med_100`. The first MED attribute with the route policy as network 1.0.0.0 is set as 50, while the MED attribute of the second is 100.

```
[Switch A] route-policy apply_med_50 permit node 10
[Switch A-route-policy] if-match acl 2000
[Switch A-route-policy] apply cost 50
[Switch A-route-policy] quit
[Switch A] route-policy apply_med_100 permit node 10
[Switch A-route-policy] if-match acl 2000
[Switch A-route-policy] apply cost 100
[Switch A-route-policy] quit
```

- Apply route policy `set_med_50` to egress route update of Switch C (193.1.1.2), and apply route policy `set_med_100` on the egress route of Switch B (192.1.1.2)

```
[Switch A] bgp 100
[Switch A-bgp] peer ex193 route-policy apply_med_50 export
[Switch A-bgp] peer ex192 route-policy apply_med_100 export
```

2) Configure Switch B:

```
[Switch B] interface vlan-interface 2
[Switch B-Vlan-interface2] ip address 192.1.1.2 255.255.255.0
[Switch B] interface vlan-interface 4
[Switch B-Vlan-interface4] ip address 194.1.1.2 255.255.255.0
[Switch B] ospf
[Switch B-ospf-1] area 0
[Switch B-ospf-1-area-0.0.0.0] network 194.1.1.0 0.0.0.255
[Switch B-ospf-1-area-0.0.0.0] network 192.1.1.0 0.0.0.255
[Switch B] bgp 200
[Switch B-bgp] undo synchronization
[Switch B-bgp] group ex external
[Switch B-bgp] peer 192.1.1.1 group ex as-number 100
[Switch B-bgp] group in internal
[Switch B-bgp] peer 194.1.1.1 group in
[Switch B-bgp] peer 195.1.1.2 group in
```

3) Configure Switch C:

```
[Switch C] interface Vlan-interface 3
```

```
[Switch C-Vlan-interface3] ip address 193.1.1.2 255.255.255.0
[Switch C] interface vlan-interface 5
[Switch C-Vlan-interface5] ip address 195.1.1.2 255.255.255.0
[Switch C] ospf
[Switch C-ospf-1] area 0
[Switch C-ospf-1-area-0.0.0.0] network 193.1.1.0 0.0.0.255
[Switch C-ospf-1-area-0.0.0.0] network 195.1.1.0 0.0.0.255
[Switch C] bgp 200
[Switch C-bgp] group ex external
[Switch C-bgp] peer 193.1.1.1 group ex as-number 100
[Switch C-bgp] group in internal
[Switch C-bgp] peer 195.1.1.1 group in
[Switch C-bgp] peer 194.1.1.2 group in
```

4) Configure Switch D:

```
[Switch D] interface vlan-interface 4
[Switch D-Vlan-interface4] ip address 194.1.1.1 255.255.255.0
[Switch D] interface vlan-interface 5
[Switch D-Vlan-interface5] ip address 195.1.1.1 255.255.255.0
[Switch D] ospf
[Switch D-ospf-1] area 0
[Switch D-ospf-1-area-0.0.0.0] network 194.1.1.0 0.0.0.255
[Switch D-ospf-1-area-0.0.0.0] network 195.1.1.0 0.0.0.255
[Switch D-ospf-1-area-0.0.0.0] network 4.0.0.0 0.255.255.255
[Switch D] bgp 200
[Switch D-bgp] group ex external
[Switch D-bgp] peer ex as-number 200
[Switch D-bgp] peer 195.1.1.2 group ex
[Switch D-bgp] peer 194.1.1.2 group ex
```

To enable the configuration, all BGP neighbors will be reset using the **reset bgp all** command.

After above configuration, due to the fact that the MED attribute of route 1.0.0.0 discovered by Switch C is less than that of Switch B, Switch D will first select the route 1.0.0.0 from Switch C.

If the MED attribute of Switch A is not configured, the local preference on Switch C is configured as follows:

Configure the local preference attribute of Switch C

- Add ACL 2000 on Switch C and permit network 1.0.0.0

```
[Switch C] acl number 2000
[Switch C-acl-basic-2000] rule permit source 1.0.0.0 0.255.255.255
[Switch C-acl-basic-2000] rule deny source any
```

- Define the route policy with the name of localpref, of those, the local preference matching ACL 2000 is set as 200, and that of not matching is set as 100.

```
[Switch C] route-policy localpref permit node 10
[Switch C-route-policy] if-match acl 2000
[Switch C-route-policy] apply local-preference 200
[Switch C-route-policy] route-policy localpref permit node 20
[Switch C-route-policy] apply local-preference 100
[Switch C-route-policy] quit
```

- Apply such route policy to the BGP neighbor 193.1.1.1 (Switch A)

```
[Switch C] bgp 200
[Switch C-bgp] peer 193.1.1.1 route-policy localpref import
```

By then, due to the fact that the Local preference attribute value (200) of the route 1.0.0.0 learned by Switch C is more than that of Switch B (Switch B is not configured with local Preference attribute, 100 by default), Switch D will also first select the route 1.0.0.0 from Switch C.

21.5 Troubleshooting BGP

Symptom 1: The neighborhood cannot be established (The Established state cannot be entered).

Solution: The establishment of BGP neighborhood needs the router able to establish TCP connection through port 179 and exchange Open packets correctly. Perform the check according to the following steps:

- Check whether the configuration of the neighbor's AS number is correct.
- Check whether the neighbor's IP address is correct.
- If using the Loopback interface, check whether the **connect-source loopback** command has been configured. By default, the router uses the optimal local interface to establish the TCP connection, not using the loopback interface.
- If it is the EBGP neighbor not directly connected, check whether the **peer ebgp-max-hop** command has been configured.
- Use the **ping** command to check whether the TCP connection is normal. Since one router may have several interfaces able to reach the peer, the extended **ping -a ip-address** command should be used to specify the source IP address sending ping packet.
- If the Ping operation fails, use the **display ip routing-table** command to check if there is available route in the routing table to the neighbor.
- If the Ping operation succeeds, check if there is an ACL denying TCP port 179. If the ACL is configured, cancel the denying of port 179.

Symptom 2: BGP route cannot be advertised correctly after route of IGP is imported with the **network** command.

Solution: Route imported by the **network** command should be same as a route in the current routing table, which should include destination segment and mask. Route

covering large network segment cannot be imported. For example, route 10.1.1.0/24 can be imported, while 10.0.0.0/8 may cause error.

Chapter 22 IP Routing Policy Configuration

22.1 Introduction to IP Routing Policy

When a router advertises or receives routing information, it possibly needs to implement some policies to filter the routing information, so as to receive or advertise the routing information which can meet the specified condition only. A routing protocol, e.g. RIP, may need import the routing information discovered by other protocols to enrich its routing knowledge. While importing the routing information, it possibly only needs import the information meeting the conditions and set some special attributes to make them meet its requirement.

For implementing the routing policy, you need define a set of matching rules by specifying the characteristics of the routing information to be filtered. You can set the rules based on such attributes like destination address and source address of the information. The matching rules can be set in advance and then used in the routing policy to advertise, receive and import the route information.

22.1.1 Filter

In the Switch 8800, five kinds of filters, Route-policy, acl, as-path, community-list, and ip-prefix, are provided to be called by the routing protocols. The following sections introduce these filters respectively.

I. acl

The access control list (ACL) used by routing policy can be divided into the following types:

- Number-based basic ACLs
- Name-based basic ACLs
- Number-based advanced ACLs
- Name-based advanced ACLs
- Number-based L2 ACLs
- Name-based L2 ACLs
- Number-based user ACLs
- Name-based user ACLs

For routing information filtering, the basic ACL is generally used. When users define the ACL, they will define the range of an IP address or subnet to the destination network segment address or the next-hop address of the routing information. If an advanced ACL is used, perform the matching operation by the specified source address range.

II. ip-prefix

The function of the ip-prefix is similar to that of the acl, but it is more flexible and easy for the users to understand. When the ip-prefix is applied to the routing information filtering, its matching objects are the destination address information domain of the routing information.

An ip-prefix is identified by the ip-prefix name. Each ip-prefix can include multiple list items, and each list item can independently specify the match range of the network prefix forms and is identified with an index-number. The index-number designates the matching check sequence in the ip-prefix.

During the matching, the router checks list items identified by the sequence-number in the ascending order. Once a single list item meets the condition, it means that it has passed the ip-prefix filtering and will not enter the testing of the next list item.

III. as-path

The as-path list is only used in the BGP. The routing information packet of the BGP includes an autonomous system path domain (During the process of routing information exchanging of the BGP, the autonomous system paths the routing information has passed through will be recorded in this domain). Targeting at the AS path domain, the as-path specifies the match condition.

IV. community-list

The community-list is only used in the BGP. The routing information packet of the BGP includes a community attribute domain to identify a community. Targeting at the community attribute, the community-list specifies the match condition.

22.1.2 Routing Policy Application

Two routing policy applications are as follows:

- When advertising/receiving routing information, the router filters the information according to the route policy, and receives or advertises the routing information which can meet the specified condition only.
- When importing other routes detected by other routing protocol, the router only imports the routing information, which can meet the specified condition only, according to the route policy.

22.2 Configuring IP Routing Policy

The routing policy configuration includes:

- 1) Filter configuration includes:
 - Configuring a Route-policy
 - Configuring Access Control List (ACL)
 -

- Configuring ip-prefix
- Configuring the AS Path List
- Configuring a Community Attribute List

Note:

For the configuration of ACL, refer to the QoS/ACL operation part of this manual.

- 2) Applications of routing policies include:
- Importing Routing Information Discovered by Other Routing Protocols
 - Configuring Route Filtering

22.2.1 Configuring a Route-policy

A route-policy can comprise multiple nodes. Each node is a unit for matching operation. The nodes will be tested against by *node-number*.

Each node consists of a group of **if-match** clauses and **apply** clauses.

- The **if-match** clauses define the matching rules. The different **if-match** clauses for a node have the relationship of “AND”. That is, the route must satisfy all the **if-match** clauses for the node to match the node before passing this node.
- The **apply** clauses define the executed action after the routing information passes the matching test. That is, the clause sets the routing information attribute.

I. Defining a route-policy

Perform the following configuration in system view.

Table 22-1 Define a route-policy

Operation	Command
Enter Route policy view	route-policy <i>route-policy-name</i> { permit deny } node <i>node-number</i>
Remove the specified route-policy	undo route-policy <i>route-policy-name</i> [permit deny node <i>node-number</i>]

The **permit** keyword specifies the matching mode for a defined node in the route-policy to be in permit mode. If a route satisfies all the **if-match** clauses of the node, it will pass the filtering of the node, and the **apply** clauses for the node will be executed without taking the test of the next node. If not, however, the route should take the test of the next node.

The **deny** keyword specifies the matching mode for a defined node in the route-policy to be in deny mode. In this mode, the **apply** clauses will not be executed. If a route

satisfies all the **if-match** clauses of the node, it will be denied by the node and will not take the test of the next node. If not, however, the route will take the test of the next node.

The nodes have the “OR” relationship. In other words, the router will test the route against the nodes in the route-policy in sequence. Once a node is matched, the route-policy filtering will be passed.

By default, the route-policy is not defined.

Note: If multiple nodes are defined in a route-policy, at least one of them should be in permit mode. Apply the route-policy to filter routing information. If the routing information does not match any node, the routing information will be denied by the route-policy. If all the nodes in the route-policy are in deny mode, all routing information will be denied by the route-policy.

II. Defining if-match clauses for a route-policy

The **if-match** clauses define the matching rules. That is, the filtering conditions that the routing information should satisfy for passing the route-policy. The matching objects are some attributes of routing information.

Perform the following configuration in route policy view.

Table 22-2 Define if-match conditions

Operation	Command
Match the AS path domain of the BGP routing information	if-match as-path <i>acl-number</i>
Cancel the matched AS path domain of the BGP routing information	undo if-match as-path
Match the community attribute of the BGP routing information	if-match community { <i>basic-community-number</i> [whole-match] <i>adv-community-number</i> }
Cancel the matched community attribute of the BGP routing information	undo if-match community
Match the destination address of the routing information	if-match { acl <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> }
Cancel the matched destination address of the routing information	undo if-match { acl ip-prefix }
Match the next-hop interface of the routing information	if-match interface { <i>interface-type</i> <i>interface-number</i> }
Cancel the matched next-hop interface of the routing information	undo if-match interface
Match the next-hop of the routing information	if-match ip next-hop { acl <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> }

Operation	Command
Cancel the matched next-hop of the routing information set by ACL	undo if-match ip next-hop
Cancel the matched next-hop of the routing information set by address prefix list	undo if-match ip next-hop ip-prefix
Match the routing cost of the routing information	if-match cost <i>value</i>
Cancel the matched routing cost of the routing information	undo if-match cost
Match the tag domain of the OSPF routing information	if-match tag <i>value</i>
Cancel the tag domain of the matched OSPF routing information	undo if-match tag

By default, no matching will be performed.

Note the following:

- The **if-match** clauses for a node in the route-policy have the relationship of “AND” for matching. That is, the route must satisfy all the clauses to match the node before the actions specified by the **apply** clauses can be executed.
- If no **if-match** clauses are specified, all the routes will pass the filtering on the node.

III. Defining apply clauses for a route-policy

The **apply** clauses specify actions, which are the configuration commands executed after a route satisfies the filtering conditions specified by the **if-match** clauses. Thereby, some attributes of the route can be modified.

Perform the following configuration in route policy view.

Table 22-3 Define apply clauses

Operation	Command
Add the specified AS number before the as-path series of the BGP routing information	apply as-path <i>as-number-1</i> [<i>as-number-2</i> [<i>as-number-3</i> ...]]
Cancel the specified AS number added before the as-path series of the BGP routing information	undo apply as-path
Set the community attribute in the BGP routing information	apply community [[<i>aa:nn</i> no-export-subconfed no-export no-advertise] * [additive] additive none]
Cancel the set community attribute in the BGP routing information	undo apply community

Operation	Command
Set the next-hop address of the routing information	apply ip next-hop <i>ip-address</i>
Cancel the next-hop address of the routing information	undo apply ip next-hop
Import the route to IS-IS level-1, level-2 or level-1-2	apply isis [level-1 / level-2 / level-1-2]
Remove the function of importing the route to IS-IS	undo apply isis
Set the local preference of the BGP routing information	apply local-preference <i>local-preference</i>
Cancel the local preference of the BGP routing information	undo apply local-preference
Set the routing cost of the routing information	apply cost <i>value</i>
Cancel the routing cost of the routing information	undo apply cost
Set the cost type of the routing information	apply cost-type [internal external]
Remove the setting of the cost type	undo apply cost-type
Set the route origin of the BGP routing information	apply origin { igp egp <i>as-number</i> incomplete }
Cancel the route origin of the BGP routing information	undo apply origin
Set the tag domain of the OSPF routing information	apply tag <i>value</i>
Cancel the tag domain of the OSPF routing information	undo apply tag

By default, perform no settings.

Note that if the routing information meets the match conditions specified in the route-policy and also notifies the MED value configured with the **apply cost-type internal** when notifying the IGP route to the EBGp peers, then this value will be regarded as the MED value of the IGP route. The preference configured with the **apply cost-type internal** command is lower than that configured with the **apply cost** command, but higher than that configured with the **default med** command.

22.2.2 Configuring ip-prefix

- A prefix-list is identified by an *ip-prefix-name*. Each IP prefix-list may include multiple entries each specifying an IP prefix matching range. IP prefix entries are identified by *index-numbers*. The order in which IP prefix entries are matched against depends on the order of their index numbers.

Perform the following configuration in system view.

Table 22-4 Define prefix-list

Operation	Command
Define prefix-list	ip ip-prefix <i>ip-prefix-name</i> [index <i>index-number</i>] { permit deny } <i>network len</i> [greater-equal <i>greater-equal</i>] [less-equal <i>less-equal</i>]
Remove prefix-list	undo ip ip-prefix <i>ip-prefix-name</i> [index <i>index-number</i>] permit deny]

During the matching, the router checks list items identified by the *index-number* in the ascending order. If only one list item meets the condition, it means that it has passed the ip-prefix filtering (will not enter the testing of the next list item).

Note that if more than one ip-prefix item are defined, then the match mode of at least one list item should be the **permit** mode. The list items of the **deny** mode can be firstly defined to rapidly filter the routing information not satisfying the requirement, but if all the items are in the **deny** mode, no route will pass the ip-prefix filtering. You can define an item of **permit** 0.0.0.0/0 **greater-equal** 0 **less-equal** 32 after the multiple list items in the **deny** mode so as to let all the other routes pass.

22.2.3 Configuring the AS Path List

The routing information packet of the BGP includes an autonomous system path domain. The as path-list can be used to match with the autonomous system path domain of the BGP routing information so as to filter the routing information, which does not conform to the requirements.

Perform the following configuration in the system view:

Table 22-5 Define the AS path list

Operation	Command
Define the AS path list	ip as-path-acl <i>acl-number</i> { permit deny } <i>as-regular-expression</i>
Delete the specified AS path list	undo ip as-path-acl <i>acl-number</i>

By default, no AS path list is defined.

22.2.4 Configuring a Community Attribute List

In BGP, community attribute is optional and transitive. Some community attributes known globally are called standard community attributes. Some community attributes are for special purpose. You can also define expanded community attribute.

A route can have one more community attributes. The speakers of multiple community attributes of a route can act according to one, several or all attributes. A router can select community attribute modification before transmitting routes to other peers.

Community lists, which identify community information, can be divided into basic-community-lists and advanced-community-lists. Basic-community-lists range from 1 to 99, while advanced-community-lists range from 100 to 199.

Perform the following configuration in system view.

Table 22-6 Configure a community attribute list

Operation	Command
Configure a basic community-list	ip community-list <i>basic-comm-list-number</i> { permit deny } [<i>aa:nn</i> internet no-export-subconfed no-advertise no-export]*
Configure an advanced community-list	ip community-list <i>adv-comm-list-number</i> { permit deny } <i>comm-regular-expression</i>
Cancel a community-list	undo ip community-list { <i>basic-comm-list-number</i> <i>adv-comm-list-number</i> }

By default, a BGP community attribute list is not configured.

22.2.5 Importing Routing Information Discovered by Other Routing Protocols

A routing protocol can import the routes discovered by other routing protocols to enrich its route information. The route-policy can be used for route information filtering to implement the purposeful redistribution. If the destination routing protocol importing the routes cannot directly reference the route costs of the source routing protocol, you should satisfy the requirement of the protocol by specifying a route cost for the imported route.

Perform the following configuration in routing protocol view.

Table 22-7 Configure to import the routes of other protocols

Operation	Command
Set to import routes of other protocols	import-route <i>protocol</i> [med <i>med</i> cost <i>cost</i>] [tag <i>value</i>] [type 1 2] [route-policy <i>route-policy-name</i>]

Operation	Command
Cancel the setting for importing routes of other protocols	undo import-route <i>protocol</i>

By default, the routes discovered by other protocols will not be advertised.

Note:

In different routing protocol views, the parameter options are different. For details, respectively refer to the **import-route** command in different protocols.

22.2.6 Configuring Route Filtering

I. Configuring to filter the received routes

Perform the following configuration in routing protocol view.

Define a policy to filter the routing information not satisfying the conditions while receiving routes with the help of an ACL or address prefix-list. **gateway** specifies that only the update packets from a particular neighboring router will be received.

Table 22-8 Configure to filter the received routes

Operation	Command
Configure to filter the received routing information advertised by the specified address	filter-policy gateway <i>ip-prefix-name</i> import
Cancel the filtering of the received routing information advertised by the specified address	undo filter-policy gateway <i>ip-prefix-name</i> import
Configure to filter the received global routing information	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } [gateway] import
Cancel the filtering of the received global routing information	undo filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } [gateway] import

II. Configuring to filter the advertised routes

You may define a route advertisement policy to filter advertised routing information. This can be done by referencing an ACL or IP prefix-list to filter routing information that does not meet the conditions, or by specifying a protocol to filter routing information of the protocol only.

Perform the following configuration in routing protocol view.

Table 22-9 Configure to filter the advertised routes

Operation	Command
Configure to filter the routes advertised by the protocol	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } export [<i>protocol</i>]
Cancel the filtering of the routes advertised by the protocol	undo filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } export [<i>protocol</i>]

By far, the route policy supports importing the routes discovered by the following protocols into the routing table:

direct: The hop (or host) to which the local interface is directly connected.

static: Route configured statically

rip: Route discovered by RIP

ospf: Route discovered by OSPF

ospf-ase: External route discovered by OSPF

ospf-nssa: NSSA route discovered by OSPF

isis: Route discovered by IS-IS

bgp: Route acquired by BGP

By default, the filtering of the received and advertised routes will not be performed.

22.3 Displaying and Debugging the Routing Policy

After the above configuration, execute the **display** command in any view to display the running of the routing policy configuration, and to verify the effect of the configuration.

Table 22-10 Display and debug the route policy

Operation	Command
Display the routing policy	display route-policy [<i>route-policy-name</i>]
Display the path information of the AS filter in BGP	display ip as-path-acl [<i>acl-number</i>]
Display the address prefix list information	display ip ip-prefix [<i>ip-prefix-name</i>]

22.4 Typical IP Routing Policy Configuration Example

22.4.1 Configuring to Filter the Received Routing Information

I. Network requirements

- Switch A communicates with Switch B, running OSPF protocol. The router ID of Switch A is 1.1.1.1, and that of Switch B is 2.2.2.2.
- Import three static routes through enabling the OSPF protocol on the Switch A.
- The route filtering rules can be configured on Switch B to make the received three static routes partially visible and partially shielded. It means that routes in the network segments 20.0.0.0 and 40.0.0.0 are visible while those in the network segment 30.0.0.0 are shielded.

II. Network diagram

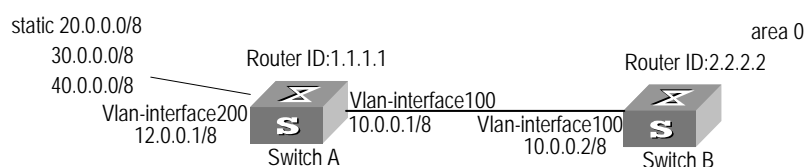


Figure 22-1 Network diagram for filtering the received routing information

III. Configuration procedure

1) Configure Switch A:

Configure the IP address of VLAN interface.

```
[Switch A] interface vlan-interface 100
[Switch A-Vlan-interface100] ip address 10.0.0.1 255.0.0.0
[Switch A] interface vlan-interface 200
[Switch A-Vlan-interface200] ip address 12.0.0.1 255.0.0.0
```

Configure three static routes.

```
[Switch A] ip route-static 20.0.0.1 255.0.0.0 12.0.0.2
[Switch A] ip route-static 30.0.0.1 255.0.0.0 12.0.0.2
[Switch A] ip route-static 40.0.0.1 255.0.0.0 12.0.0.2
```

Enable the OSPF protocol and specifies the number of the area to which the interface belongs.

```
[Switch A] router id 1.1.1.1
[Switch A] ospf
[Switch A-ospf-1] area 0
[Switch A-ospf-1-area-0.0.0.0] network 10.0.0.0 0.255.255.255
```

Import the static routes

```
[Switch A-ospf-1] import-route static
```

2) Configure Switch B:

Configure the IP address of VLAN interface.

```
[Switch B] interface vlan-interface 100
[Switch B-Vlan-interface100] ip address 10.0.0.2 255.0.0.0
```

Configure the access control list.

```
[Switch B] acl number 2000
[Switch B-acl-basic-2000] rule deny source 30.0.0.0 0.255.255.255
[Switch B-acl-basic-2000] rule permit source any
```

Enable OSPF protocol and specifies the number of the area to which the interface belongs.

```
[Switch B] router id 2.2.2.2
[Switch B] ospf
[Switch B-ospf-1] area 0
[Switch B-ospf-1-area-0.0.0.0] network 10.0.0.0 0.255.255.255
```

Configure OSPF to filter the external routes received.

```
[Switch B-ospf-1] filter-policy 2000 import
```

22.5 Troubleshooting Routing Policy

Symptom 1: Routing information filtering cannot be implemented in normal operation of the routing protocol

Solution: Check for the following faults:

- The if-match mode of at least one node of the Route-policy should be the **permit** mode. When a Route-policy is used for the routing information filtering, if a piece of routing information does not pass the filtering of any node, then it means that the route information does not pass the filtering of the Route-policy. When all the nodes of the Route-policy are in the **deny** mode, then all the routing information cannot pass the filtering of the Route-policy.
- The if-match mode of at least one list item of the ip-prefix should be the permit mode. The list items of the deny mode can be firstly defined to rapidly filter the routing information not satisfying the requirement, but if all the items are in the deny mode, any routes will not pass the ip-prefix filtering. You can define an item of permit 0.0.0.0/0 less-equal 32 after the multiple list items in the deny mode so as to let all the other routes pass the filtering (If less-equal 32 is not specified, only the default route will be matched).

Chapter 23 IP Multicast Overview

Note:

An Ethernet switch functions as a router when it runs IP multicast protocol. A router that is referred to in the following represents a generalized router or a layer 3 Ethernet switch running IP multicast protocol.

23.1 IP Multicast Overview

23.1.1 Problems with Unicast/Broadcast

The constant development of the Internet and increasing interaction of versatile data, voice and video information over the network, has promoted the emergence of new services like e-commerce, network conference, online auction, video on demand (VoD), and tele-education. These services require higher information security and greater rewards.

I. Data transmission in unicast mode

In unicast mode, every user that needs the information receives a copy through the channels the system separately establishes for them. See Figure 23-1.

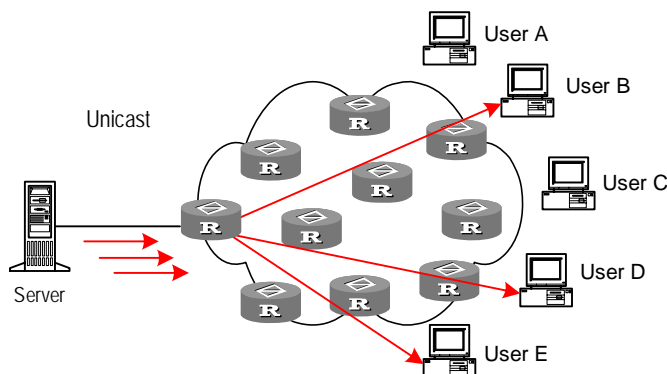


Figure 23-1 Data transmission in unicast mode

Suppose that Users B, D, and E need the information, the information source Server establishes transmission channels with every of them. Since the traffic in transmission increases with the number of users, excessive copies of the information would spread

over the network if there is a large number of users in need of this information. As the bandwidth would turn short, the unicast mode is incapable of massive transmission.

II. Data transmission in broadcast mode

In broadcast mode, every user on the network receives the information regardless of their needs. See Figure 23-2 Data transmission in broadcast mode.

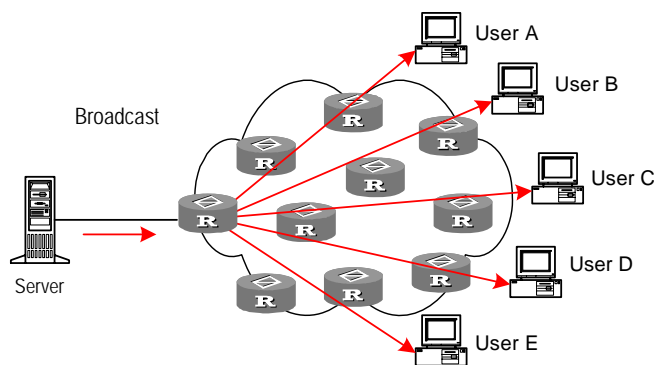


Figure 23-2 Data transmission in broadcast mode

Suppose the Users B, D, and E need the information, the information source Server broadcasts the information through the router; User A and User C can also receive the information. In that case, information security and rewards to services are not guaranteed. Moreover, bandwidth is terribly wasted when only a few part of users are in need of the information.

In short, the unicast mode is useful in networks with scattered users, and the multicast mode is suitable for networks with dense users. When the number of users is uncertain, the adoption of unicast or multicast mode results in low efficiency.

23.1.2 Advantages of Multicast

I. Multicast

IP multicast technology solves those problems. When some users in the network need specific information, it allows the multicast source to send the information only once. With the tree route established by the multicast routing protocol, the information will not be duplicated or distributed until it reaches the bifurcation point as far as possible. See Figure 23-3 Data transmission in multicast mode.

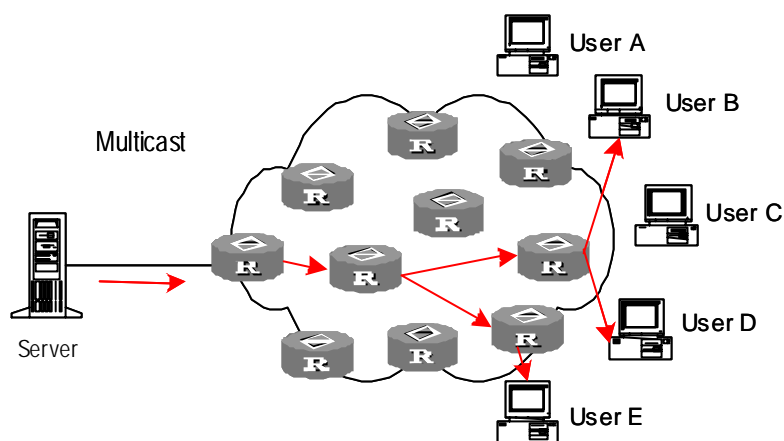


Figure 23-3 Data transmission in multicast mode

Suppose the Users B, D, and E need the information, they need to be organized into a receiver group to ensure that the information can reach them smoothly. The routers on the network duplicate and forward the information according to the distribution of these users in the group. Finally, the information is transmitted to the intended receivers B, D and E properly and correctly.

In multicast mode, the information sender is called the "multicast source", the receiver is called the "multicast group", and the routers for multicast information transmission are called "multicast routers". Members of a multicast group can scatter around the network; the multicast group therefore has no geographical limitation. It should be noted that a multicast source does not necessarily belong to a multicast group. It sends data to multicast groups but is not necessarily a receiver. Multiple sources can send packets to a multicast group simultaneously.

II. Advantages

The main advantages of multicast are:

- Enhanced efficiency: It reduces network traffic and relieves server and CPU of loads.
- Optimized performance: It eliminates traffic redundancy.
- Distributed application: It enables multipoint application.

23.1.3 Application of Multicast

IP multicast technology effectively implements point to multi-point forwarding with high speed, as saves network bandwidth a lot and can relieve network loads. It facilitates also the development of new value-added services in the Internet information service area that include online live show, Web TV, tele-education, telemedicine, network radio station and real-time audio/video conferencing. It takes a positive role in:

- Multimedia and streaming media application

- Occasional communication for training and cooperation
- Data storage and finance (stock) operation
- Point-to-multipoint data distribution

With the increasing popularity of multimedia services over IP network, multicast is gaining its marketplace. In addition, the multicast service becomes popular and prevalent gradually.

23.2 Implementation of IP Multicast

23.2.1 IP Multicast Addresses

In multicast mode, there are questions about where to send the information, how to locate the destination or know the receiver. All these questions can be narrowed down to multicast addressing. To guarantee the communication between a multicast source and a multicast group (that is, a group of receivers), the network layer multicast address (namely the IP multicast address) is required, along with the technique to correlate it with the link layer MAC multicast address. Following is the introduction to these two kinds of addresses.

I. IP Multicast Addresses

According to the definition in Internet Assigned Number Authority (IANA), IP addresses fall into four types: Class A, Class B, Class C and Class D. Unicast packets use IP addresses of Class A, Class B or Class C, depending on specific packet scales. Multicast packets use IP addresses of Class D as their destination addresses, but Class D IP addresses cannot be contained in the source IP field of IP packets.

During unicast data transmission, a packet is transmitted "hop-by-hop" from the source address to the destination address. However, in IP multicast environment, a packet has more than one destination address, or a group of addresses. All the information receivers are added to a group. Once a receiver joins the group, the data for this group address starts flowing to this receiver. All members in the group can receive the packets. This group is a multicast group.

Membership here is dynamic, and a host can join or leave the group at any time. A multicast group can be permanent or temporary. Some multicast group addresses are allocated by IANA, and the multicast group is called permanent multicast group. The IP addresses of a permanent multicast group are unchangeable, but its membership is changeable, and the number of members is arbitrary. It is quite possible for a permanent group to not a single member. Those not reserved for permanent multicast groups can be used by temporary multicast groups. Class D multicast addresses range from 224.0.0.0 to 239.255.255.255. More information is listed in Table 23-1 Ranges and meanings of Class D addresses.

Table 23-1 Ranges and meanings of Class D addresses

Class D address range	Description
224.0.0.0~224.0.0.255	Reserved multicast addresses (addresses of permanent groups). All but 224.0.0.0 can be allocated by routing protocols.
224.0.1.0~238.255.255.255	Multicast addresses available for users (addresses of temporary groups). They are valid in the entire network.
239.0.0.0~239.255.255.255	Multicast addresses for local management. They are valid only in the specified local range.

Reserved multicast addresses that are commonly used are described in the following table.

Table 23-2 Reserved multicast address list

Class D address range	Description
224.0.0.0	Base Address (Reserved)
224.0.0.1	Addresses of all hosts
224.0.0.2	Addresses of all multicast routers
224.0.0.3	Not for allocation
224.0.0.4	DVMRP routers
224.0.0.5	OSPF routers
224.0.0.6	OSPF DR
224.0.0.7	ST routers
224.0.0.8	ST hosts
224.0.0.9	RIP-2 routers
224.0.0.10	IGRP routers
224.0.0.11	Active agents
224.0.0.12	DHCP server/Relay agent
224.0.0.13	All PIM routers
224.0.0.14	RSVP encapsulation
224.0.0.15	All CBT routers
224.0.0.16	Specified SBM
224.0.0.17	All SBMS
224.0.0.18	VRRP

Class D address range	Description
.....

II. Ethernet Multicast MAC Addresses

When a unicast IP packet is transmitted on the Ethernet, the destination MAC address is the MAC address of the receiver. However, for a multicast packet, the destination is no longer a specific receiver but a group with unspecific members. Therefore, the multicast MAC address should be used.

As Internet Assigned Number Authority (IANA) provisions, the high 24 bits of a multicast MAC address are 0x01005e and the low 23 bits of a MAC address are the low 23 bits of a multicast IP address. The high twenty-fifth bit is 0, a fixed value.

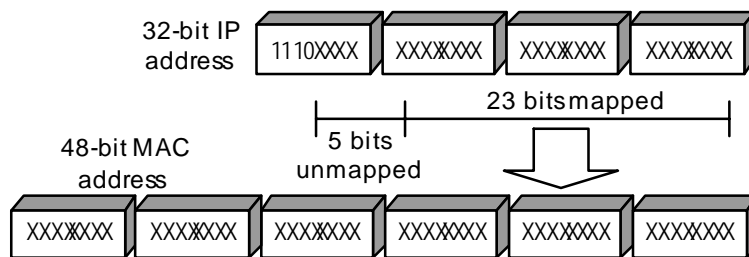


Figure 23-4 Mapping between a multicast IP address and an Ethernet MAC address

The first four bits of the multicast address are 1110, representing the multicast identifier. Among the rest 28 bits, only 23 bits are mapped to the MAC address, and the other five bits are lost. This may result in that 32 IP addresses are mapped to the same MAC address.

23.2.2 IP Multicast Protocols

IP multicast protocols mainly involve multicast group management protocols and multicast routing protocols. Their application positions are shown in Figure 23-5 Application positions of multicast-related protocols.

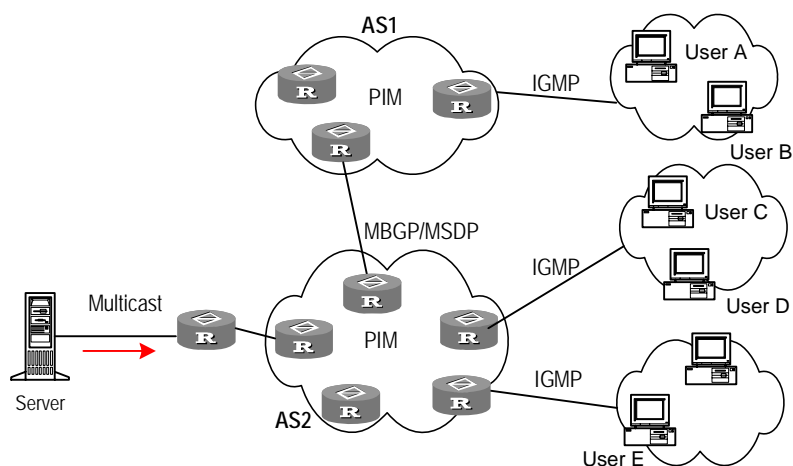


Figure 23-5 Application positions of multicast-related protocols

I. Multicast group management protocol

Multicast groups use Internet group management protocol (IGMP) as the management protocols. IGMP runs between the host and multicast router and defines the membership establishment and maintenance mechanism between them.

II. Multicast routing protocols

A multicast routing protocol runs between multicast routers to create and maintain multicast routes for correct and efficient forwarding of multicast packet. The multicast routing creates a loop-free data transmission path from one source to multiple receivers. The task of multicast routing protocols is to build up the distribution tree architecture. A multicast router can use multiple methods to build up a path for data transmission, that is, a distribution tree.

As in unicast routing, the multicast routing can also be intra-domain or inter-domain. Intra-domain multicast routing is rather mature and protocol independent multicast (PIM) is the most widely used intra-domain protocol, which can work in collaboration with unicast routing protocols. The inter-domain routing first needs to solve how to transfer routing information between ASs. Since the ASs may belong to different telecom carriers, the inter-domain routing information must contain carriers' policies, in addition to distance information. Currently, inter-domain routing protocols include multicast source discovery protocol (MSDP) and MBGP multicast extension.

23.3 RPF Mechanism for IP Multicast Packets

To ensure that multicast packets reach a router along the shortest path, the multicast router must check the receiving interface of multicast packets depending on the unicast routing table or a unicast routing table independently provided for multicast. This check mechanism is the basis for most multicast routing protocols to perform multicast forwarding, and is known as Reverse Path Forwarding (RPF) check. A multicast router

uses the source address of a received multicast packet to query the unicast routing table or the independent multicast routing table to determine that the receiving interface is on the shortest path from the receiving station to the source. If a source tree is used, the source address is the address of the source host sending the multicast packet. If a shared tree is used, the source address is the RP address of the shared tree. A multicast packet arriving at the router will be forwarded according to the multicast forwarding entry if it passes the RPF check, or else, it will be discarded.

Chapter 24 IGMP Snooping Configuration

24.1 IGMP Snooping Overview

24.1.1 IGMP Snooping Principle

IGMP Snooping (Internet Group Management Protocol Snooping) is a multicast control mechanism running on the Layer 2 Ethernet switch and it is used for multicast group management and control.

IGMP Snooping runs on the link layer. When receiving the IGMP messages transmitted between the host and router, the Layer 2 Ethernet switch uses IGMP Snooping to analyze the information carried in the IGMP messages. If the switch hears IGMP host report message from an IGMP host, it will add the host to the corresponding multicast table. If the switch hears IGMP leave message from an IGMP host, it will remove the host from the corresponding multicast table. The switch continuously listens to the IGMP messages to create and maintain MAC multicast address table on Layer 2. And then it can forward the multicast packets transmitted from the upstream router according to the MAC multicast address table.

When IGMP Snooping is disabled, the packets are broadcasted on Layer 2. See the following figure:

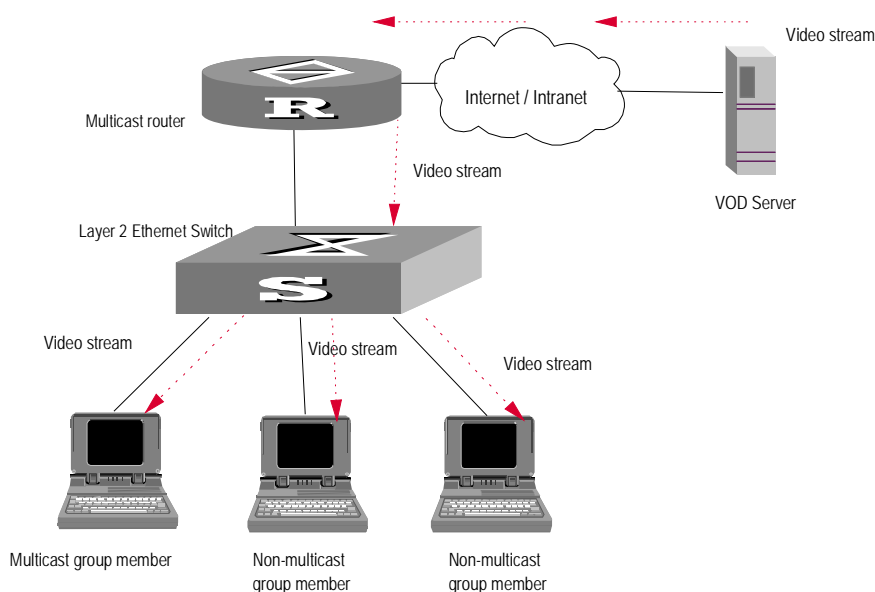


Figure 24-1 Multicast packet transmission without IGMP Snooping

When IGMP Snooping runs, the packets are multicast rather than broadcasted on Layer 2. See the following figure:

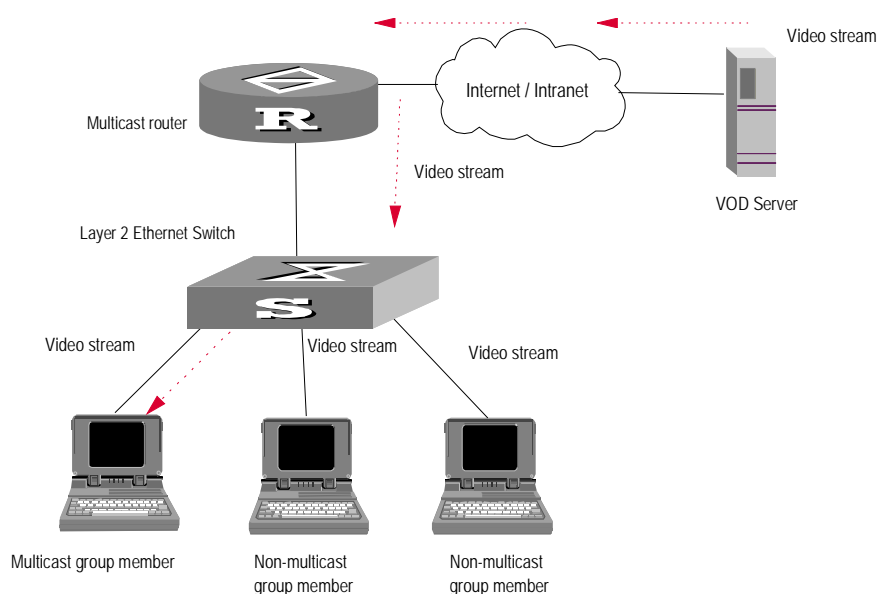


Figure 24-2 Multicast packet transmission when IGMP Snooping runs

24.1.2 Implement IGMP Snooping

I. Related concepts of IGMP Snooping

To facilitate the description, this section first introduces some related switch concepts of IGMP Snooping:

- Router Port: The port of the switch, directly connected to the multicast router.
- Multicast member port: The port connected to the multicast member. The multicast member refers to a host joined a multicast group.
- MAC multicast group: The multicast group is identified with MAC multicast address and maintained by the Ethernet switch.
- Router port aging time: Time set on the router port aging timer. If the switch has not received any IGMP general query message before the timer times out, it considers the port no longer as a router port.
- Multicast group member port aging time: When a port joins an IP multicast group, the aging timer of the port will begin timing. The multicast group member port aging time is set on this aging timer. If the switch has not received any IGMP report message before the timer times out, it transmits IGMP specific query message to the port.
- Maximum response time: When the switch transmits IGMP specific query message to the multicast member port, the Ethernet switch starts a response timer, which times before the response to the query. If the switch has not received any IGMP report message before the timer times out, it will remove the port from the multicast member ports

II. Implement Layer 2 multicast with IGMP Snooping

The Ethernet switch runs IGMP Snooping to listen to the IGMP messages and map the host and its ports to the corresponding multicast group address. To implement IGMP Snooping, the Layer 2 Ethernet switch processes different IGMP messages in the way illustrated in the figure below:

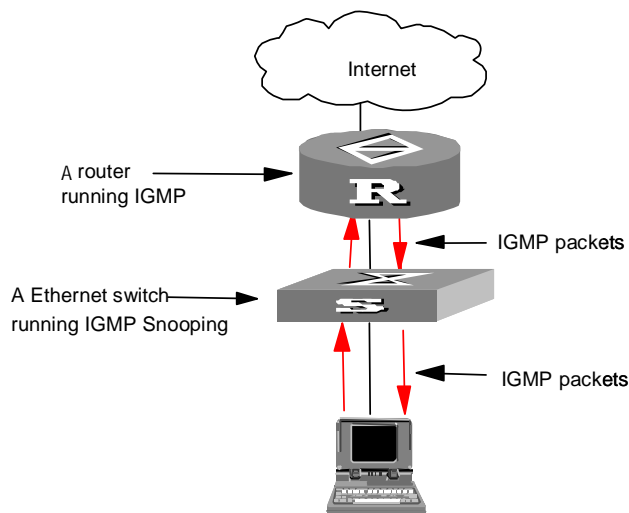


Figure 24-3 Implement IGMP Snooping

- IGMP general query message: Transmitted by the multicast router to the multicast group members to query which multicast group contains member. When an IGMP general query message arrives at a router port, the Ethernet switch will reset the aging timer of the port. When a port other than a router port receives the IGMP general query message, the Ethernet switch will notify the multicast router that a port is ready to join a multicast group and starts the aging timer for the port.
- IGMP specific query message: Transmitted from the multicast router to the multicast members and used for querying if a specific group contains any member. When received IGMP specific query message, the switch only transmits the specific query message to the IP multicast group which is queried.
- IGMP report message: Transmitted from the host to the multicast router and used for applying for joining a multicast group or responding to the IGMP query message. When received the IGMP report message, the switch checks if the MAC multicast group, corresponding to the IP multicast group the packet is ready to join exists.

If the corresponding MAC multicast group does not exist, the switch only notifies the router that a member is ready to join a multicast group, creates a new MAC multicast group, adds the port received the message to the group, starts the port aging timer, and then adds all the router ports in the native VLAN of the port into the MAC multicast

forwarding table, and meanwhile creates an IP multicast group and adds the port received the report message to it.

If the corresponding MAC multicast group exists but does not contains the port received the report message, the switch adds the port into the multicast group and starts the port aging timer. And then the switch checks if the corresponding IP multicast group exists.

If it does not exist, the switch creates a new IP multicast group and adds the port received the report message to it. If it exists, the switch adds the port to it.

If the MAC multicast group corresponding to the message exists and contains the port received the message, the switch will only reset the aging timer of the port.

- IGMP leave message: Transmitted from the multicast group member to the multicast router to notify that a host left the multicast group. When received a leave message of an IP multicast group, the Ethernet switch transmits the specific query message concerning that group to the port received the message, in order to check if the host still has some other member of this group and meanwhile starts a maximum response timer. If the switch has not receive any report message from the multicast group after the timer expires, the port will be removed from the corresponding MAC multicast group. If the MAC multicast group does not have any member, the switch will notify the multicast router to remove the branch from the multicast tree.

24.2 IGMP Snooping Configuration

The main IGMP Snooping configuration includes:

- Enabling/disabling IGMP Snooping
- Configuring the aging time of router port
- Configuring maximum response time
- Configuring the aging time of multicast group member port
- Configuring Unknown Multicast Packets not Broadcasted within a VLAN

Among the above configuration tasks, enabling IGMP Snooping is required, while others are optional for your requirements.

24.2.1 Enabling/Disabling IGMP Snooping

You can use the following commands to enable/disable IGMP Snooping to control whether MAC multicast forwarding table is created and maintained on Layer 2. First enable IGMP Snooping globally in system view, and then enable IGMP Snooping of the corresponding VLAN in VLAN view. The second step must be based on the first one.

Perform the following configuration in system view and VLAN view.

Table 24-1 Enabling/Disabling IGMP Snooping

Operation	Command
Enable/disable IGMP Snooping	igmp-snooping { enable disable }

By default, IGMP Snooping is disabled.

**Caution:**

- Although layer 2 and layer 3 multicast protocols can be configured in pair, they cannot run on the same VLAN or its corresponding VLAN interface at the same time. For example, if PIM or IGMP is enabled on a VLAN, then IGMP Snooping cannot operate on this VLAN.
 - IGMP Snooping functions only when it is enabled both in system view and in VLAN view. Otherwise, IGMP Snooping does not take effect.
-

24.2.2 Configuring Router Port Aging Time

This task is to manually configure the router port aging time. If the switch has not received any general query message from the router before the router port is aged, it will remove the port from all MAC multicast groups.

Perform the following configuration in system view.

Table 24-2 Configuring router port aging time

Operation	Command
Configure router port aging time	igmp-snooping router-aging-time <i>seconds</i>
Restore the default aging time	undo igmp-snooping router-aging-time

By default, the port aging time is 105s.

24.2.3 Configuring Maximum Response Time

This task is to manually configure the maximum response time. If the Ethernet switch receives no report message from a port in the maximum response time, it will remove the port from the multicast group.

Perform the following configuration in system view.

Table 24-3 Configuring the maximum response time

Operation	Command
Configure the maximum response time	igmp-snooping max-response-time <i>seconds</i>
Restore the default setting	undo IGMP-snooping max-response-time

By default, the maximum response time is 1 seconds.

24.2.4 Configuring Aging Time of Multicast Group Member Ports

This task is to manually set the aging time of the multicast group member port. If the switch receives no multicast group report message during the member port aging time, it will transmit the specific query message to that port and starts a maximum response timer.

Perform the following configuration in system view.

Table 24-4 Configuring aging time of the multicast member ports

Operation	Command
Configure aging time of the multicast member	igmp-snooping host-aging-time <i>seconds</i>
Restore the default setting	undo igmp-snooping host-aging-time

By default, the aging time of the multicast member is 260 seconds.

24.2.5 Configuring Unknown Multicast Packets not Broadcasted within a VLAN

This configuration task is to enable/disable the function of not broadcasting unknown multicast packets within a VLAN. If this function is not enabled but IGMP snooping enabled on VLAN, multicast packets are broadcasted on within the VLAN when the destination broadcast group has no member ports. When this function is enabled, however, multicast packets are only forwarded to the router port, but not broadcasted within the VLAN if no member port exists. Since the router port is that connected to the router with IGMP/PIM enabled and the router sends regularly IGMP Query and PIM Hello packets, the switch can identify the router port. If there is no router port, multicast packets shall be dropped, instead of being forwarded.

**Caution:**

If IGMP snooping is not enabled on the VLAN (nor Layer 3 multicast), unknown multicast packets are broadcasted within the VLAN no matter whether this function is enabled or not. That is, to make unknown multicast packets not be broadcasted with a VLAN, you must enable `igmp-snooping` in this VLAN and enable **`igmp-snooping nonflooding-enable`** globally.

Perform the following configuration in system view.

Table 24-5 Globally enable/disable multicast packets not broadcasted within a VLAN

Operation	Command
Enable multicast packets not to be broadcasted within a VLAN	<code>igmp-snooping nonflooding-enable</code>
Disable multicast packets not to be broadcasted within a VLAN	<code>undo igmp-snooping nonflooding-enable</code>

By default, unknown multicast packets are broadcasted within the VLAN.

24.3 Displaying and debugging IGMP Snooping

After the above configuration, execute **`display`** command in any view to display the running of the IGMP Snooping configuration, and to verify the effect of the configuration.

Table 24-6 Displaying and debugging IGMP Snooping

Operation	Command
Display the information about current IGMP Snooping configuration	<code>display igmp-snooping configuration</code>
Display IGMP Snooping statistics of received and sent messages	<code>display igmp-snooping statistics</code>
Display IP/MAC multicast group information in the VLAN	<code>display igmp-snooping group [vlan <i>vlanid</i>]</code>

24.4 IGMP Snooping Configuration Example

24.4.1 Enable IGMP Snooping

I. Networking requirements

To implement IGMP Snooping on the switch, you need to enable IGMP Snooping on the switch first. The switch is connected with the router via the router port, and connected with user PC through the non-router ports.

II. Networking diagram

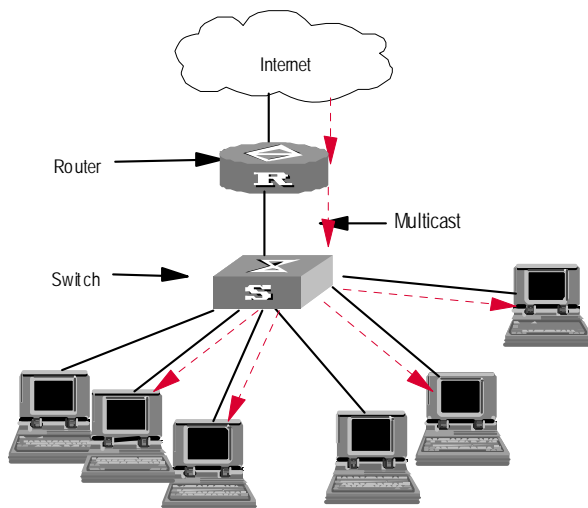


Figure 24-4 IGMP Snooping configuration networking

III. Configuration procedure

Suppose you need to enable IGMP Snooping on VLAN1. The procedures are as follows:

Display the current state of IGMP Snooping.

```
<SW8800> display igmp-snooping configuration
```

If IGMP Snooping is not enabled, enable it in system view.

```
[SW8800] igmp-snooping enable
```

Display the status of the VLAN1 interface, to check if PIM or IGMP is enabled on it.

```
[SW8800] display current-configuration interface Vlan-interface 1
```

If PIM or IGMP is not running on VLAN1, you can enable IGMP Snooping in VLAN view.

```
[SW8800-vlan1] igmp-snooping enable
```

24.5 Troubleshoot IGMP Snooping

Fault: Multicast function cannot be implemented on the switch.

Troubleshooting:

- 1) IGMP Snooping is disabled.
 - Input the **display current-configuration** command to display the status of IGMP Snooping.

- If IGMP Snooping is not enabled, input the **igmp-snooping enable** command in system view to enable IGMP Snooping. Then, use the same command in VLAN view to enable IGMP Snooping in the corresponding VLAN.
- 2) Multicast forwarding table set up by IGMP Snooping is wrong.
- Input the **display igmp-snooping group** command to display if the multicast group is the expected one.
 - If the multicast group created by IGMP Snooping is not correct, turn to professional maintenance personnel for help.
 - Continue with diagnosis 3 if the second step is completed.
- 3) Multicast forwarding table set up on the bottom layer is wrong.
- In any view, execute the **display mac-address vlan** command to check whether the MAC multicast forwarding table established in the VLAN identified by *vlanid* is consistent with that established by IGMP Snooping.
 - If they are not consistent, please contact the maintenance personnel for help.

Chapter 25 Multicast VLAN Configuration

25.1 Multicast VLAN Overview

Based on the current multicast on demand, when users in different VLANs request the service, multicast flow is duplicated in each VLAN and a great deal of bandwidth is wasted. To solve this problem, we provide the multicast VLAN feature. With this feature, you can add switch ports to a multicast VLAN and enable IGMP Snooping to allow users in different VLANs to share the same multicast VLAN. In this way, multicast flow is transmitted in one multicast VLAN instead of multiple user VLANs and bandwidth is greatly saved. Multicast VLAN is isolated from user VLANs, this guarantees both security and enough bandwidth. After you configure the multicast VLAN, multicast information flow can be transmitted to users continuously.

25.2 Multicast VLAN Configuration

Multicast VLAN is based on layer 2 multicast. The following table describes the multicast VLAN configuration tasks:

Table 25-1 Configure multicast VLAN for a layer 2 switch

Item	Command	Description
Enter system view	system-view	—
Enable IGMP Snooping in system view	igmp-snooping enable	Required
Enter VLAN view	vlan x	x: Vlan ID.
IGMP Snooping is enabled on the VLAN Enable IGMP Snooping in VLAN view	igmp-snooping enable	Required
Enable multicast VLAN	service-type multicast	Required
Quit VLAN view	quit	
Enter the view of the Ethernet port connected to the user	interface <i>interface_type</i> <i>interface_num</i>	<i>interface_type</i> <i>interface_num</i> : port type and number.
Define the port type to hybrid	port link-type hybrid	Required

Item	Command	Description
Add ports to corresponding VLANs	port hybrid vlan <i>vlan_id_list</i> untagged	Required

To cancel the configurations, use the corresponding **undo** commands.

Note:

- A port can only belong to one multicast VLAN.
 - The type of the ports connected to user terminals must be hybrid untagged.
-

25.3 Multicast VLAN Configuration Example

I. Network requirements

The following table describes the devices required in this example:

Table 25-2 Device number and description

Device	Description	Requirement
Switch A	Layer 3 switch	The IP address of VLAN 2 interface is 168.10.1.1. The port E1/1/1 belongs to VLAN 2 and is connected to the Workstation. The IP address of VLAN 10 interface is 168.20.1.1. The port E1/1/10 belongs to VLAN 10 and is connected to Switch B. Configure layer 3 multicast PIM DM and IGMP on VLAN 10
Switch B	Layer 2 switch	VLAN 2 contains the port E1/1/1 and VLAN 3 contains the port E1/1/2. The ports E1/1/1 and E1/1/2 are connected to PC1 and PC2 respectively. The port E1/1/10 is connected to Switch A.
PC 1	User 1	PC1 is connected to the port E1/1/1 of Switch B.
PC 2	User 2	PC2 is connected to the port E1/1/2 of Switch B.

Configure multicast VLAN to make users in VLAN 2 and VLAN 3 receive multicast flow through VLAN10.

II. Network diagram

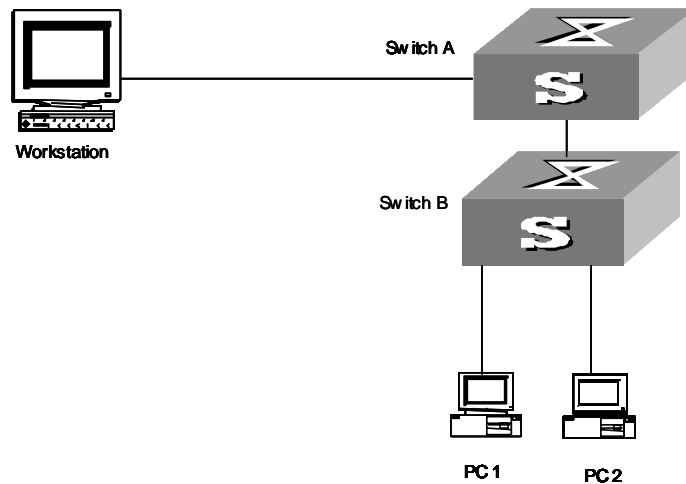


Figure 25-1 Network diagram for multicast VLAN configuration

III. Configuration procedure

Before performing the following configurations, you should configure the IP addresses and connect the devices correctly.

1) Configure Switch A

Configure the IP address of the VLAN 2 interface to 168.10.1.1. Enable the PIM DM protocol.

```
<Switch A> system-view
[Switch A] multicast routing-enable
[Switch A] vlan 2
[Switch A-vlan2] interface vlan-interface 2
[Switch A-Vlan-interface2] ip address 168.10.1.1 255.255.255.0
[Switch A-Vlan-interface2] pim dm
[Switch A-Vlan-interface2] quit
```

Configure the IP address of the VLAN 10 interface to 168.20.1.1. Enable the PIM DM and IGMP protocols.

```
[Switch A] vlan 10
[Switch A-vlan10] interface vlan-interface 10
[Switch A-Vlan-interface10] ip address 168.20.1.1 255.255.255.0
[Switch A-Vlan-interface10] pim dm
[Switch A-Vlan-interface10] igmp enable
[Switch A-Vlan-interface10] quit
[Switch A] interface e1/1/10
[Switch A-Ethernet1/1/10] port link-type trunk
[Switch A-Ethernet1/1/10] port trunk permit vlan 10
```

2) Configure Switch B

Enable IGMP Snooping.

```
<Switch B> system-view
[Switch B] igmp-snooping enable
```

Enable IGMP-Snooping on VLAN 2 and VLAN 3.

```
[Switch B] vlan 2
[Switch B-vlan 2] igmp-snooping enable
[Switch B-vlan 2]quit
[Switch B] vlan 3
[Switch B-vlan 3] igmp-snooping enable
```

Configure VLAN 10 as multicast VLAN. Enable IGMP Snooping.

```
[Switch B] vlan 10
[Switch B-vlan10] igmp-snooping enable
[Switch B-vlan10] service-type multicast
[Switch B-vlan10] quit
```

Define Ethernet 1/1/10 as trunk port. Add the port to VLAN 10.

```
[Switch B] interface Ethernet 1/1/10
[Switch B-Ethernet 1/1/10] port link-type trunk
[Switch B-Ethernet 1/1/10] port trunk vlan 10
[Switch B-Ethernet 1/1/10] quit
```

Define Ethernet 1/1/1 as hybrid port. Add the port to VLAN 2 and VLAN 10. Make the port carry no VLAN label when it transmits packets of VLAN 2 and VLAN 10. Set the default VLAN ID of the port to VLAN 2.

```
[Switch B] interface Ethernet 1/1/1
[Switch B-Ethernet 1/1/1] port link-type hybrid
[Switch B-Ethernet 1/1/1] port hybrid vlan 2 10 untagged
[Switch B-Ethernet 1/1/1] port hybrid pvid vlan 2
[Switch B-Ethernet 1/1/1] quit
```

Define Ethernet 1/1/2 as hybrid port. Add the port to VLAN 3 and VLAN 10. Make the port carry no VLAN label when it transmits packets of VLAN 3 and VLAN 10. Set the default VLAN ID of the port to VLAN 3.

```
[Switch B] interface Ethernet 1/1/1
[Switch B-Ethernet 1/1/2] port link-type hybrid
[Switch B-Ethernet 1/1/2] port hybrid vlan 3 10 untagged
[Switch B-Ethernet 1/1/2] port hybrid pvid vlan 3
[Switch B-Ethernet 1/1/2] quit
```


Chapter 26 Common Multicast Configuration

26.1 Introduction to Common Multicast Configuration

The multicast common configuration is for both the multicast group management protocol and the multicast routing protocol. The configuration includes enabling multicast, displaying multicast routing table and multicast forwarding table, etc.

26.2 Common Multicast Configuration

Common multicast configuration includes:

- Enabling multicast
- Configuring multicast route limit
- Clearing MFC (Multicast Forwarding Cache) forwarding entries or its statistic information
- Configuring controlled multicast
- Clearing route entries from the kernel multicast routing table

26.2.1 Enabling Multicast

Enable multicast first before enabling multicast routing protocol.

Perform the following configuration in system view.

Table 26-1 Enabling multicast

Operation	Command
Enable multicast	multicast routing-enable
Disable multicast	undo multicast routing-enable

By default, multicast is disabled.



Caution:

Only when multicast is enabled can other multicast configuration become effective.

26.2.2 Configuring multicast route number limit

Because too many multicast routes may exhaust the router memory, you need to limit the number of multicast routes.

Perform the following configuration in system view.

Table 26-2 Configuring multicast route limit

Operation	Command
Configure multicast route limit	multicast route-limit <i>limit</i>
Restore multicast route limit to the default value	undo multicast route-limit

By default, the maximum multicast routes allowed by the system is 512.

26.2.3 Clearing MFC Forwarding Entries or Its Statistic Information

You can clear MFC forward entries or statistic information of FMC forward entries via the following command.

Perform the following configuration in user view.

Table 26-3 Clearing MFC forwarding entries or its statistic information

Operation	Command
Clear MFC forwarding entries or its statistic information	reset multicast forwarding-table [statistics] { all { <i>group-address</i> [mask { <i>group-mask</i> <i>group-mask-length</i> }] <i>source-address</i> [mask { <i>source-mask</i> <i>source-mask-length</i> }] incoming-interface { null <i>NULL-interface-number</i> <i>interface-type interface-number</i> } } * }

26.2.4 Clearing Route Entries from the Kernel Multicast Routing Table

You can clear route entries from the kernel multicast routing table, as well as MFC forwarding entries via the following command.

Perform the following configuration in user view.

Table 26-4 Clearing routing entries of multicast routing table

Operation	Command
Clear routing entries of multicast routing table	reset multicast routing-table { all { <i>group-address</i> [mask { <i>group-mask</i> <i>group-mask-length</i> }] <i>source-address</i> [mask { <i>source-mask</i> <i>source-mask-length</i> }] { incoming-interface <i>vlan-interface interface-number</i> } } * }

26.3 Controlled Multicast Configuration

26.3.1 Controlled Multicast Overview

The controlled multicast feature controls user's authority to join multicast groups. This feature is based on ports: users must first pass the 802.1x authentication set for their ports. Then they are allowed to join the multicast groups specifically configured for them but are prohibited from joining the unauthorized multicast groups. This gives you a way to control users' access to specific multicast groups.

After the distributed multicast is improved, some of the multicast modules operating previously in centralized mode are migrated to interface boards now. That is, the interface boards perform part of the multicast operations to reduce the burden of the main control board and the operation results are synchronized to the main control board. To control users' access to specific multicast groups in distributed mode, some of the controlled multicast modules operating previously in centralized mode are migrated to run in the distributed environment to make the controlled multicast operate in distributed mode.

If no user interfaces is added, the CLI commands under controlled multicast in distributed mode are consistent with that in centralized mode.

Prerequisites of multicast authentication:

- 1) DOT1X is enabled both globally and on ports. Otherwise, when you enable controlled multicast, all IGMP report messages are legal. Then the system allows users to join any group and cannot control the access to multicast groups.
- 2) The controlled multicast is based on port. The DOT1X mode on port must be port authentication. Otherwise, the system discards all IGMP report messages without any processing.

26.3.2 Configuring Controlled Multicast

Controlled multicast configuration tasks include:

- Enabling controlled multicast globally
- Configuring multicast address for specific user access
- Displaying the status of online controlled multicast members
- Displaying the debug information about controlled multicast

Perform the following configuration in local user view.

Table 26-5 Configure the controlled multicast

Operation	Command
Configure the controlled multicast	multicast <i>ip-address</i>
Remove the configuration	undo multicast { <i>ip-address</i> all }

**Caution:**

In local user view, before executing this command, you must configure user service type to LAN-ACCESS, which is the only one supported by controlled multicast at present.

26.3.3 Controlled Multicast Configuration Example

I. Network requirements

As shown in Figure 26-1, HostA and HostB join the multicast group. Layer 3 multicast is enabled on LSA, LSB, LSC and LSD. Controlled multicast is enabled on LSA and LSC. Because controlled multicast combines multicast with 802.1x, 802.1x should be enabled on LSA and LSC.

II. Network diagram

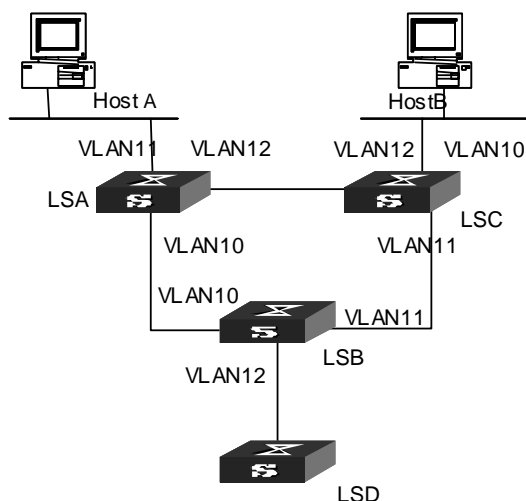


Figure 26-1 Network diagram for controlled multicast

III. Configuration procedure

Controlled multicast is combined with 802.1x, so you need to perform the following configuration beside multicast configuration:

Enable controlled multicast globally.

```
[SW8800] ip managed-multicast
```

Enable 802.1x globally.

```
[SW8800] dot1x
```

Enable 802.1x on the controlled ports (the access ports for LSA and LSC).

```
[SW8800-GigabitEthernet2/1/1] dot1x
[SW8800-GigabitEthernet2/1/2] dot1x
```

Configure the authentication mode on the controlled ports to port-based mode.

```
[SW8800-GigabitEthernet2/1/1] dot1x -method portbased
[SW8800-GigabitEthernet2/1/2] dot1x -method portbased
```

Create a local-user in system view. Then set the password and service type for the user.

```
[SW8800] local-user liu
[SW8800-luser-liu] password simple aaa
[SW8800-luser-liu] service-type lan-access
```

In user view, configure the allowed multicast group for the user to join.

```
[SW8800-luser-liu] multicast 227.1.1.1
```

26.4 Displaying and Debugging Common Multicast Configuration

After the above configuration, execute **display** command in any view to display the running of the multicast configuration, and to verify the effect of the configuration.

Execute **debugging** command in user view for the debugging of multicast.

Table 26-6 Displaying and Debugging Common Multicast Configuration

Operation	Command
Display the multicast routing table	display multicast routing-table [<i>group-address</i> [mask { <i>mask</i> <i>mask-length</i> }] <i>source-address</i> [mask { <i>mask</i> <i>mask-length</i> }] incoming-interface { vlan-interface <i>vlan-interface-number</i> register }]*
Display the multicast forwarding table	display multicast forwarding-table [<i>group-address</i> [mask { <i>mask</i> <i>mask-length</i> }] <i>source-address</i> [mask { <i>mask</i> <i>mask-length</i> }] incoming-interface { <i>interface-type</i> <i>interface-number</i> null <i>NULL-interface-number</i> register }]*
Enable multicast packet forwarding debugging	debugging multicast forwarding
Disable multicast packet forwarding debugging	undo debugging multicast forwarding
Enable multicast forwarding status debugging	debugging multicast status-forwarding
Disable multicast forwarding status debugging	undo debugging multicast status-forwarding

Operation	Command
Enable multicast kernel routing debugging	debugging multicast kernel-routing
Disable multicast kernel routing debugging	undo debugging multicast kernel-routing

The multicast routing tables can be layered as follows:

- Each multicast routing protocol has a multicast routing table of itself.
- All the multicast routing tables can be summarized into the multicast kernel routing tables.
- The multicast kernel routing tables should keep consistent with the multicast forwarding tables which actually control the forwarding of the multicast data packets.

The multicast forwarding tables are mainly used for debugging. Usually, users can view the multicast kernel routing tables to get the required information.

Chapter 27 IGMP Configuration

27.1 IGMP Overview

27.1.1 Introduction to IGMP

Internet Group Management Protocol (IGMP) is a protocol in the TCP/IP suite responsible for management of IP multicast members. It is used to establish and maintain multicast membership among IP hosts and their directly connected neighboring routers. IGMP excludes transmitting and maintenance of membership information among multicast routers, which are completed by multicast routing protocols. All hosts participating in multicast must implement IGMP.

Hosts participating in IP multicast can join and leave a multicast group at any time. The number of members of a multicast group can be any integer and the location of them can be anywhere. A multicast router does not need and cannot keep the membership of all hosts. It only uses IGMP to learn whether receivers (i.e., group members) of a multicast group are present on the subnet connected to each interface. A host only needs to keep which multicast groups it has joined.

IGMP is not symmetric on hosts and routers. Hosts need to respond to IGMP query messages from the multicast router, i.e., report the group membership to the router. The router needs to send membership query messages periodically to discover whether hosts join the specified group on its subnets according to the received response messages. When the router receives the report that hosts leave the group, the router will send a group-specific query packet (IGMP Version 2) to discover whether no member exists in the group.

Up to now, IGMP has three versions, namely, IGMP Version 1 (defined by RFC1112), IGMP Version 2 (defined by RFC2236) and IGMP Version 3. At present, IGMP Version 2 is the most widely used version.

IGMP Version 2 boasts the following improvements over IGMP Version 1:

I. Election mechanism of multicast routers on the shared network segment

A shared network segment means that there are multiple multicast routers on a network segment. In this case, all routers running IGMP on the network segment can receive the membership report from hosts. Therefore, only one router is necessary to send membership query messages. In this case, the router election mechanism is required to specify a router as the querier.

In IGMP Version 1, selection of the querier is determined by the multicast routing protocol. While IGMP Version 2 specifies that the multicast router with the lowest IP

address is elected as the querier when there are multiple multicast routers on the same network segment.

II. Leaving group mechanism

In IGMP Version 1, hosts leave the multicast group quietly without informing the multicast router. In this case, the multicast router can only depend on the timeout of the response time of the multicast group to confirm that hosts leave the group. In Version 2, when a host is intended to leave, it will send a leave group message if it is the host who responds to the latest membership query message.

III. Specific group query

In IGMP Version 1, a query of a multicast router is targeted at all the multicast groups on the network segment, which is known as General Query.

In IGMP Version 2, Group-Specific Query is added besides general query. The destination IP address of the query packet is the IP address of the multicast group. The group address domain in the packet is also the IP address of the multicast group. This prevents the hosts of members of other multicast groups from sending response messages.

IV. Max response time

The Max Response Time is added in IGMP Version 2. It is used to dynamically adjust the allowed maximum time for a host to respond to the group query message.

27.2 IGMP Configuration

After the multicast function is enabled, you must enable IGMP on the interface first and then perform other IGMP configurations.

- 1) IGMP basic configuration includes:
 - Enabling multicast
 - Enabling IGMP on an interface
- 2) IGMP advanced configuration includes:
 - Configuring a router to join specified multicast group
 - Controlling the access to IP multicast groups
 - Configuring the IGMP query message interval
 - Configuring the IGMP version
 - Configuring the IGMP querier present timer
 - Configuring the maximum query response time
 - Configuring the times of sending IGMP Group-Specific Query packet
 - Configuring the interval of sending IGMP Group-Specific Query packet
 - Configuring the limit of IGMP groups on an interface
 - Deleting IGMP Groups Joined on an Interface

27.2.1 Enabling Multicast

Only if the multicast function is enabled can the multicast-related configurations take effect.

Refer to Chapter 26 Common Multicast Configuration.

27.2.2 Enabling IGMP on an Interface

This configuration task is to enable IGMP on the interface which needs to maintain the multicast membership. After this, you can initiate IGMP feature configuration.

Perform the following configuration in VLAN interface view.

Table 27-1 Enabling/Disabling IGMP on an interface

Operation	Command
Enable IGMP on an interface	igmp enable
Disable IGMP on an interface	undo igmp enable

By default, IGMP is not enabled.

27.2.3 Configuring the IGMP Version

Perform the following configuration in VLAN interface view.

Table 27-2 Configuring the IGMP version

Operation	Command
Select the IGMP version that the router uses	igmp version { 2 1 }
Restore the default setting	undo igmp version

By default, IGMP Version 2 is used.



Caution:

All routers on a subnet must support the same version of IGMP. After detecting the presence of IGMP Version 1 system, a router cannot automatically switch to Version 1.

27.2.4 Configuring the Interval to Send IGMP Query Message

Multicast routers send IGMP query messages to discover which multicast groups are present on attached networks. Multicast routers send query messages periodically to refresh their knowledge of members present on their networks.

Perform the following configuration in VLAN interface view.

Table 27-3 Configuring the interval to send IGMP query message

Operation	Command
Configure the interval to send IGMP query message	igmp timer query <i>seconds</i>
Restore the default value	undo igmp timer query

When there are multiple multicast routers on a network segment, the querier is responsible for sending IGMP query messages to all hosts on the LAN.

By default, the interval is 60 seconds.

27.2.5 Configuring the Interval and the Number of Querying IGMP Packets

On the shared network, it is the query router (querier) that maintains IGMP membership on the interface. The following commands are used to configure the interval and times of sending IGMP group-specific query packets for the querier when it receives an IGMP leave message from a host.

- The host sends the IGMP Leave message.
- Upon receiving the message, IGMP querier sends the group-specific IGMP query message for specified times (defined by the *robust-value* in **igmp robust-count**, with the default value as 2) and at a time interval (defined by the *seconds* in **igmp lastmember-queryinterval**, with the default value as 1 second).
- When other hosts receive the message from the IGMP querier and are interested in this group, they return the IGMP Membership Report message within the defined maximum response time.
- If IGMP querier receives the report messages from other hosts within the period equal to $robust-value \times seconds$, it continues membership maintenance for this group.
- If it receives no report message from any other host within this period, it reckens this as timeout and ends membership maintenance for this group.

This command can be used only when the querier runs IGMP version 2, since a host running IGMP Version 1 does not send IGMP Leave Group message when it leaves a group.

Please perform the following configurations in VLAN interface view.

I. Configuring interval for querying IGMP packets

Table 27-4 Configuring interval for querying IGMP packets

Operation	Command
Configure interval for querying IGMP packets	igmp lastmember-queryinterval <i>seconds</i>
Restore te default query interval	undo igmp lastmember-queryinterval

By default, the interval is 1 second.

II. Configuring the number of last member querying

Table 27-5 Configuring the number of last member querying

Operation	Command
Configure number of last member querying	igmp robust-count <i>robust-value</i>
Restore the default number of querying	undo igmp robust-count

By default, an IGMP group-specific query message is sent for twice.

27.2.6 Configuring the Present Time of IGMP Querier

The IGMP querier present timer defines the period of time before the router takes over as the querier sending query messages, after the previous querier has stopped doing so.

Perform the following configuration in VLAN interface view.

Table 27-6 Configuring the present time of IGMP querier

Operation	Command
Change the present time of IGMP querier	igmp timer other-querier-present <i>seconds</i>
Restore the default value	undo igmp timer other-querier-present

By default, the value is 120 seconds. If the router has received no query message within twice the interval specified by the **igmp timer query** command, it will regard the previous querier invalid.

27.2.7 Configuring Maximum Response Time for IGMP Query Message

When a router receives a query message, the host will set a timer for each multicast group it belongs to. The value of the timer is randomly selected between 0 and the

maximum response time. When any timer becomes 0, the host will send the membership report message of the multicast group.

Setting the maximum response time reasonably can enable the host to respond to query messages quickly. In this case, the router can fast master the existing status of the members of the multicast group.

Perform the following configuration in VLAN interface view.

Table 27-7 Configuring the maximum response time for IGMP query message

Operation	Command
Configure the maximum response time for IGMP query message	igmp max-response-time <i>seconds</i>
Restore the maximum query response time to the default value	undo igmp max-response-time

The smaller the maximum query response time value, the faster the router prunes groups. The actual response time is a random value in the range from 1 to 25 seconds. By default, the maximum query response time is 10 seconds.

27.2.8 Configuring the limit of IGMP groups on an interface

If there is no limit to the number of IGMP groups added on a router interface or a router, the router memory may be exhausted, which may cause router failure.

You can set number limit for the IGMP groups added on the interface, but not the number limit for the IGMP groups added in the router, which is defined by the system.

Perform the following configuration in VLAN interface view.

Table 27-8 Configuring the limit of IGMP groups on an interface

Operation	Command
Configure the limit of IGMP groups on an interface	igmp group-limit <i>limit</i>
Restore the limit of IGMP groups on an interface to the default value	undo igmp group-limit

By default, the maximum number of IGMP groups on an interface is 512.

If the number of IGMP groups on an interface has exceeded the specified value during configuration, the existing IGMP groups will not be deleted.

27.2.9 Configuring a Router to Join Specified Multicast Group

Usually, the host operating IGMP will respond to IGMP query packet of the multicast router. In case of response failure, the multicast router will consider that there is no multicast member on this network segment and will cancel the corresponding path.

Configuring one interface of the router as multicast member can avoid such problem. When the interface receives IGMP query packet, the router will respond, thus ensuring that the network segment where the interface located can normally receive multicast packets.

For an Ethernet switch, you can configure a port in a VLAN interface to join a multicast group.

Perform the following configuration in the corresponding view.

Table 27-9 Configuring a router to join specified multicast group

Operation	Command
Configure the router to join a specified multicast group (in VLAN interface view)	igmp host-join <i>group-address</i> port { <i>interface_type interface_ num interface_name</i> } [to { <i>interface_type interface_ num interface_name</i> }]
Cancel the configuration (in VLAN interface view)	undo igmp host-join <i>group-address</i> port { <i>interface_type interface_ num interface_name</i> } [to { <i>interface_type interface_ num interface_name</i> }]
Configure the router to join a specified multicast group (in Ethernet port view)	igmp host-join <i>group-address</i> vlan <i>vlanid</i>
Cancel the configuration (in Ethernet port view)	undo igmp host-join <i>group-address</i> vlan <i>vlanid</i>

Note:

The above two configuration methods have the same result (both takes effect on port). You can select any one of them.

By default, a router joins no multicast group. Note that the specified port must belong to this VLAN interface on which IGMP is enabled. Otherwise, the configuration does not take effect.

27.2.10 Limiting Multicast Groups that an Interface Can Access

A multicast router learns whether there are members of a multicast group on the network via the received IGMP membership message. A filter can be set on an interface so as to limit the range of allowed multicast groups.

Perform the following configuration in the corresponding view.

Table 27-10 Limiting multicast groups an interface can access

Operation	Command
Limit the range of allowed multicast groups on current interface (in VLAN interface view)	igmp group-policy <i>acl-number</i> [1 2 port { <i>interface_type interface_num</i> <i>interface_name</i> } [to { <i>interface_type interface_num</i> <i>interface_name</i> }]]
Remove the filter set on the interface (in VLAN interface view)	undo igmp group-policy [port { <i>interface_type interface_num</i> <i>interface_name</i> } [to { <i>interface_type interface_num</i> <i>interface_name</i> }]]
Limit the multicast groups that the interface serves (in Ethernet port view)	igmp group-policy <i>acl-number</i> vlan <i>vlanid</i>
Cancel the filter configured on the interface (in Ethernet port view)	undo igmp group-policy vlan <i>vlanid</i>

Note:

Using the above two configuration methods to configure ports, you can obtain the same result. You can select any one.

By default, no filter is configured, that is, all multicast groups are allowed on the interface.

The **port** keyword only takes effect on VLAN interfaces. The port specified by the **port** keyword must belong to this VLAN interface. For the configuration in Ethernet port view, the **port** must belong to the VLAN interface specified by the command. Besides, IGMP is enabled on this VLAN interface. Otherwise, this command does not take effect.

27.2.11 Deleting IGMP Groups Joined on an Interface

This configuration task is to delete all IGMP groups joined on all interfaces or specific interfaces of the router, or to delete the IGMP groups at the specific address or in the specific network segment on the specific interfaces of the router.

Perform the following configuration in user view.

Table 27-11 Deleting IGMP groups joined on an interface

Operation	Command
Delete IGMP groups joined on an interface	reset igmp group { all interface <i>vlan-interface interface-number</i> { all <i>group-address</i> [<i>group-mask</i>] } }

After a group is deleted, it can be joined on an interface again.

27.3 Displaying and Debugging IGMP

After the above configuration, execute **display** command in any view to display the running of IGMP configuration, and to verify the effect of the configuration.

Execute **debugging** command in corresponding views for the debugging of IGMP.

Table 27-12 Displaying and debugging IGMP

Operation	Command
Display the information about members of IGMP multicast groups (any views)	display igmp group [<i>group-address</i> interface <i>vlan-interface interface-number</i>]
Display the IGMP configuration and running information about the interface (any views)	display igmp interface [<i>vlan-interface interface-number</i>]
Enable the IGMP information debugging (user view)	debugging igmp { all event host packet timer }
Disable the IGMP information debugging (user view)	undo debugging igmp { all event host packet timer }

Chapter 28 PIM-DM Configuration

28.1 PIM-DM Overview

28.1.1 Introduction to PIM-DM

PIM-DM (Protocol Independent Multicast, Dense Mode) belongs to dense mode multicast routing protocols. PIM-DM is suitable for small networks. Members of multicast groups are relatively dense in such network environments.

28.1.2 PIM-DM Working Principle

The working procedures of PIM-DM include neighbor discovery, flood & prune and graft.

I. Neighbor discovery

The PIM-DM router needs to use Hello messages to perform neighbor discovery when it is started. All network nodes running PIM-DM keep in touch with one another with Hello messages, which are sent periodically.

II. Flood&Prune

PIM-DM assumes that all hosts on the network are ready to receive multicast data. When a multicast source "S" begins to send data to a multicast group "G", after the router receives the multicast packets, the router will perform RPF check according to the unicast routing table first. If the RPF check is passed, the router will create an (S, G) entry and then flood the data to all downstream PIM-DM nodes. If the RPF check is not passed, that is, multicast packets enter from an error interface, the packets will be discarded. After this process, an (S, G) entry will be created in the PIM-DM multicast domain.

If the downstream node has no multicast group members, it will send a Prune message to the upstream nodes to inform the upstream node not to forward data to the downstream node. Receiving the prune message, the upstream node will remove the corresponding interface from the outgoing interface list corresponding to the multicast forwarding entry (S, G). In this way, a SPT (Shortest Path Tree) rooted at Source S is built. The pruning process is initiated by leaf routers first.

This process is called "flood & prune" process. In addition, nodes that are pruned provide timeout mechanism. Each router re-starts the "flood & prune" process upon pruning timeout. The consistent "flood & prune" process of PIM-DM is performed periodically.

During this process, PIM-DM uses the RPF check and the existing unicast routing table to build a multicast forwarding tree rooted at the data source. When a packet arrives, the router will first judge the correctness of the path. If the interface that the packet arrives is the one indicated by the unicast routing to the multicast source, the packet is regarded to be from the correct path. Otherwise, the packet will be discarded as a redundancy packet without the multicast forwarding. The unicast routing information as path judgment can come from any unicast routing protocol independent of any specified unicast routing protocol such as the routing information learned by RIP and OSPF

III. Assert mechanism

As shown in the following figure, both routers A and B on the LAN have their own receiving paths to multicast source S. In this case, when they receive a multicast packet sent from multicast source S, they will both forward the packet to the LAN. Multicast Router C at the downstream node will receive two copies of the same multicast packet.

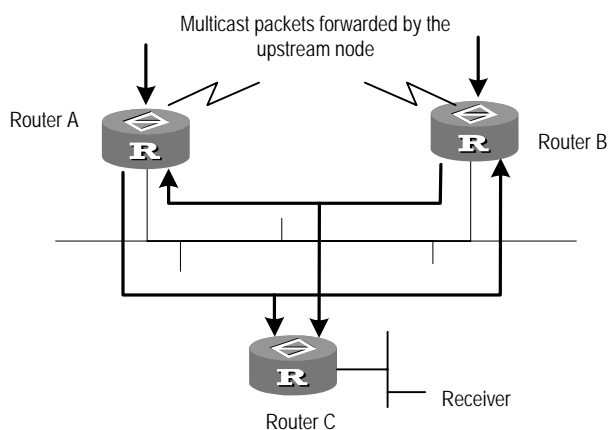


Figure 28-1 Assert mechanism diagram

When they detect such a case, routers need to select a unique sender by using the assert mechanism. Routers will send Assert packets to select the best path. If two or more than two paths have the same priority and metric, the path with a higher IP address will be the upstream neighbor of the (S, G) entry, which is responsible for forwarding the (S, G) multicast packet.

Note:

Currently assert mechanism is not available on the Switch 8800.

IV. Graft

When the pruned downstream node needs to be restored to the forwarding state, the node will send a graft packet to inform the upstream node.

28.2 PIM-DM Configuration

- 1) PIM-DM basic configuration includes:
 - Enabling multicast
 - Enabling PIM-DM
- 2) PIM-DM advanced configuration includes:
 - Configuring the time interval for ports to send Hello packets
 - Entering the PIM view
 - Configuring filtering of multicast source/group
 - Configuring filtering of PIM neighbor
 - Configuring the maximum number of PIM neighbor on an interface
 - Clearing PIM neighbors

28.2.1 Enabling Multicast

Refer to Chapter 26 Common Multicast Configuration.

28.2.2 Enabling PIM-DM

PIM-DM needs to be enabled in configuration of all interfaces.

After PIM-DM is enabled on an interface, it will send PIM Hello messages periodically and process protocol packets sent by PIM neighbors.

Perform the following configuration in VLAN interface view.

Table 28-1 Enabling PIM-DM

Operation	Command
Enable PIM-DM on an interface	pim dm
Disable PIM-DM on an interface	undo pim dm

It's recommended to configure PIM-DM on all interfaces in non-special cases. This configuration is effective only after the multicast routing is enabled in system view.

Once enabled PIM-DM on an interface, PIM-SM cannot be enabled on the same interface and vice versa.

28.2.3 Configuring the Time Intervals for Ports to Send Hello Packets

When protocol independent multicast (PIM) protocol is enabled for a port, the port sends Hello packets periodically. The time intervals to send Hello packets vary with the bandwidth and type of the connected networks.

Perform the following configuration in VLAN interface view.

Table 28-2 Configure the time intervals for ports to send Hello packets

Operation	Command
Configure the time intervals for ports to send Hello packets	pim timer hello <i>seconds</i>
Restore the default values of the time intervals	undo pim timer hello

You can configure different time intervals according to the actual networks. By default, the time interval for sending Hello packets is 30 seconds. In general, you need not modify the parameter *seconds*.

Note:

When you configure the time interval for a port to send Hello packets, the pim neighbor hold-time value automatically turns into 3.5 times the time interval value. Therefore you need not configure a value for pim neighbor hold-time.

The time interval can be configured only after the PIM protocol such as protocol independent multicast-dense mode (PIM-DM) protocol or protocol independent multicast-sparse mode (PIM-SM) protocol is enabled in VLAN interface view.

28.2.4 Entering the PIM View

Global parameters of PIM should be configured in PIM view.

Perform the following configuration in system view.

Table 28-3 Entering PIM view

Operation	Command
Enter PIM view	pim
Back to system view	undo pim

Using **undo pim** command, you can clear the configuration in PIM view, and back to system view.

28.2.5 Configuring the Filtering of Multicast Source/Group

You can set to filter the source (and group) address of multicast data packets via this command. When this feature is configured, the router filters not only multicast data, but the multicast data encapsulated in the registration packets.

Perform the following configuration in the PIM view.

Table 28-4 Configuring the filtering of multicast source/group

Operation	Command
Configure the filtering of multicast source/group	source-policy <i>acl-number</i>
Remove the configuration of filtering	undo source-policy

If resource address filtering is configured, as well as basic ACLs, then the router filters the resource addresses of all multicast data packets received. Those not matched will be discarded.

If resource address filtering is configured, as well as advanced ACLs, then the router filters the resource and group addresses of all multicast data packets received. Those not matched will be discarded.

28.2.6 Configuring the Filtering of PIM Neighbor

You can configure basic ACLs to filter the routers which can be PIM neighbors of the current interface.

Perform the following configuration in the VLAN interface view.

Table 28-5 Configuring the filtering of PIM neighbor

Operation	Command
Configure filtering of PIM neighbor	pim neighbor-policy <i>acl-number</i>
Remove the configuration of filtering	undo pim neighbor-policy

28.2.7 Configuring the Maximum Number of PIM Neighbor on an Interface

The maximum number of PIM neighbors of a router interface can be configured to avoid exhausting the memory of the router or router faults. The maximum number of PIM neighbors of a router is defined by the system, and is not open for modification.

Perform the following configuration in the VLAN interface view.

Table 28-6 Configuring the maximum number of PIM neighbor on an interface

Operation	Command
Configure the maximum number of PIM neighbor on an interface	pim neighbor-limit <i>limit</i>
Restore the limit of PIN neighbor to the default value	pim neighbor-limit

By default, the PIM neighbors on the interface are limited to 128.

If the number of PIM neighbors of an interface has exceeded the configured value by the time of configuration, the existing PIM neighbors will not be deleted.

28.2.8 Clearing multicast route entries from PIM routing table

Perform the following configuration in user view.

Table 28-7 Clearing multicast route entries from PIM routing table

Operation	Command
Clear multicast route entries from PIM routing table	reset pim routing-table { all { <i>group-address</i> [mask { <i>group-mask</i> <i>group-mask-length</i> }] <i>source-address</i> [mask { <i>source-mask</i> <i>source-mask-length</i> }] } { incoming-interface { <i>interface-type interface-number</i> null } } * }

28.2.9 Clearing PIM Neighbors

Perform the following configuration in user view.

Table 28-8 Resetting PIM neighbor

Operation	Command
Clear PIM neighbors	reset pim neighbor { all { <i>neighbor-address</i> interface <i>interface-type interface-number</i> } * }

28.3 Displaying and Debugging PIM-DM

After the above configuration, execute **display** command in any view to display the running of PIM-DM configuration, and to verify the effect of the configuration.

Execute **debugging** command in user view for the debugging of PIM-DM.

Table 28-9 Displaying and debugging PIM-DM

Operation	Command
Display the PIM multicast routing table	display pim routing-table [{ { *g [group-address [mask { mask-length mask }]] **rp [rp-address [mask { mask-length mask }]] } } { group-address [mask { mask-length mask }] source-address [mask { mask-length mask }] } * } incoming-interface { Vlan-interface Vlan-interface-number null } { dense-mode sparse-mode }] *
Display the PIM interface information	display pim interface [Vlan-interface Vlan-interface-number]
Display the information about PIM neighboring routers	display pim neighbor [interface Vlan-interface Vlan-interface-number]
Display BSR information	display pim bsr-info
Display RP information	display pim rp-info [group-address]
Enable the PIM debugging	debugging pim common { all event packet timer }
Disable the PIM debugging	undo debugging pim common { all event packet timer }
Enable the PIM-DM debugging	debugging pim dm { alert all mbr mrt timer warning { recv send } { all assert graft graft-ack join prune } }
Disable the PIM-DM debugging	undo debugging pim dm { alert all mbr mrt timer warning { recv send } { all assert graft graft-ack join prune } }

28.4 PIM-DM Configuration Example

I. Networking requirements

Lanswitch1 is connected to the multicast source through VLAN-interface 10, connected to Lanswitch2 through VLAN-interface 11 and connected to Lanswitch3 through VLAN-interface 12. Through running PIM-DM, you can implement multicast among RECEIVER 1, RECEIVER 12 and Multicast Source.

II. Networking diagram

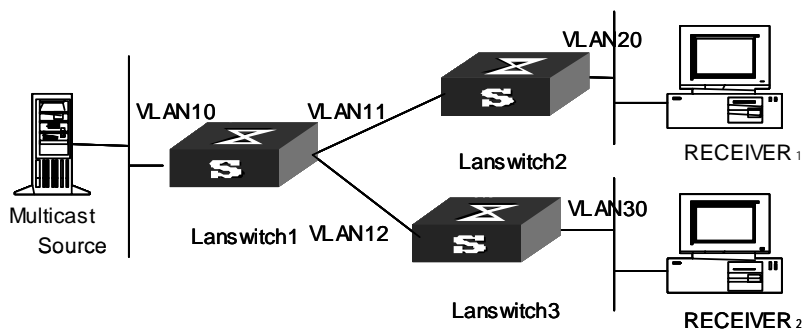


Figure 28-2 PIM-DM configuration networking

III. Configuration procedure

This section only introduces Lanswitch1 configuration procedure, while Lanswitch2 and Lanswitch3 configuration procedures are similar.

Enable the multicast routing protocol.

```
[SW8800] multicast routing-enable
```

Enable IGMP and PIM-DM on the interface.

```
[SW8800] vlan 10
[SW8800-vlan10] port ethernet 2/1/2
[SW8800-vlan10] quit
[SW8800] vlan 11
[SW8800-vlan11] port ethernet 2/1/4
[SW8800-vlan11] quit
[SW8800] vlan 12
[SW8800-vlan12] port ethernet 2/1/6
[SW8800-vlan12] quit
[SW8800] interface vlan-interface 10
[SW8800-vlan-interface10] ip address 1.1.1.1 255.255.0.0
[SW8800-vlan-interface10] igmp enable
[SW8800-vlan-interface10] pim dm
[SW8800-vlan-interface10] quit
[SW8800] interface vlan-interface 11
[SW8800-vlan-interface11] ip address 2.2.2.2 255.255.0.0
[SW8800-vlan-interface11] igmp enable
[SW8800-vlan-interface11] pim dm
[SW8800-vlan-interface11] quit
[SW8800] interface vlan-interface 12
[SW8800-vlan-interface12] ip address 3.3.3.3 255.255.0.0
[SW8800-vlan-interface12] igmp enable
[SW8800-vlan-interface12] pim dm
```

Chapter 29 PIM-SM Configuration

29.1 PIM-SM Overview

29.1.1 Introduction to PIM-SM

PIM-SM (Protocol Independent Multicast, Sparse Mode) belongs to sparse mode multicast routing protocols. PIM-SM is mainly applicable to large-scale networks with broad scope in which group members are relatively sparse.

Different from the flood & prune principle of the dense mode, PIM-SM assumes that all hosts do not need to receive multicast packets, unless there is an explicit request for the packets.

PIM-SM uses the RP (Rendezvous Point) and the BSR (Bootstrap Router) to advertise multicast information to all PIM-SM routers and uses the join/prune information of the router to build the RP-rooted shared tree (RPT), thereby reducing the bandwidth occupied by data packets and control packets and reducing the process overhead of the router. Multicast data flows along the shared tree to the network segments the multicast group members are on. When the data traffic is sufficient, the multicast data flow can switch over to the SPT (Shortest Path Tree) rooted on the source to reduce network delay. PIM-SM does not depend on the specified unicast routing protocol but uses the present unicast routing table to perform the RPF check.

Note that, the creation and interaction of the RPs and BSRs are implemented through periodical RP advertisements and BSR Bootstrap packets respectively. You can view the packets in the following debugging information:

```
<SW8800> debugging pim sm send ?
  assert      PIM-SM assertion packet sending debugging functions
  bootstrap   PIM-SM bootstrap packet sending debugging functions
  crpadv      PIM-SM RP candidate advertisement sending debugging functions
  jp          PIM-SM join/prune packet sending debugging functions
  reg         PIM-SM registration packet sending debugging functions
  regstop     PIM-SM registration-stop packet sending debugging functions
```

To make PIM-SM operate, you must configure candidate RPs and BSRs. BSRs collect and broadcast the information from candidate RPs.

29.1.2 PIM-SM Working Principle

The PIM-SM working process is as follows: neighbor discovery, building the RP-rooted shared tree (RPT), multicast source registration and SPT switchover etc. The neighbor discovery mechanism is the same as that of PIM-DM, which will not be described any more.

I. Build the RP shared tree (RPT)

When hosts join a multicast group G, the leaf routers that directly connect with the hosts send IGMP messages to learn the receivers of multicast group G. In this way, the leaf routers calculate the corresponding rendezvous point (RP) for multicast group G and then send join messages to the node of the next level toward the rendezvous point (RP). Each router along the path between the leaf routers and the RP will generate (*, G) entries in the forwarding table, indicating that all packets sent to multicast group G are applicable to the entries no matter from which source they are sent. When the RP receives the packets sent to multicast group G, the packets will be sent to leaf routers along the path built and then reach the hosts. In this way, an RP-rooted tree (RPT) is built as shown in the following figure.

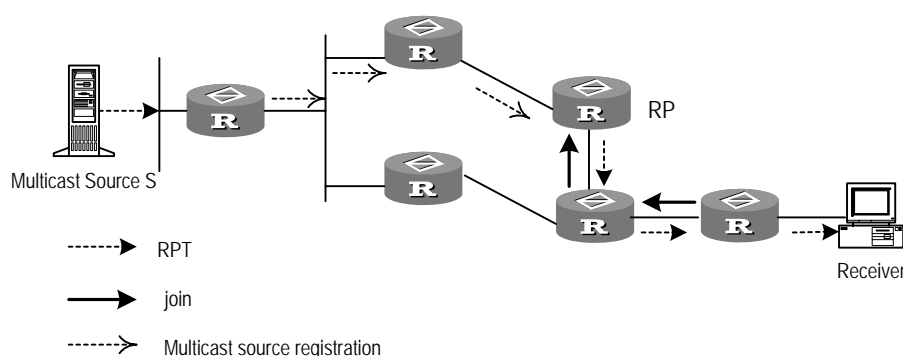


Figure 29-1 RPT schematic diagram

II. Multicast source registration

When multicast source S sends a multicast packet to the multicast group G, the PIM-SM multicast router directly connected to S will encapsulate the received packet into a registration packet and send it to the corresponding RP in unicast form. If there are multiple PIM-SM multicast routers on a network segment, the Designated Router (DR) will be responsible for sending the multicast packet.

29.1.3 Preparations before Configuring PIM-SM

I. Configuring candidate RPs

In a PIM-SM network, multiple RPs (candidate-RPs) can be configured. Each Candidate-RP (C-RP) is responsible for forwarding multicast packets with the destination addresses in a certain range. Configuring multiple C-RPs is to implement load balancing of the RP. These C-RPs are equal. All multicast routers calculate the RPs corresponding to multicast groups according to the same algorithm after receiving the C-RP messages that the BSR advertises.

It should be noted that one RP can serve multiple multicast groups or all multicast groups. Each multicast group can only be uniquely correspondent to one RP at a time rather than multiple RPs.

II. Configuring BSRs

The BSR is the management core in a PIM-SM network. Candidate-RPs send announcement to the BSR, which is responsible for collecting and advertising the information about all candidate-RPs.

It should be noted that there can be only one BSR in a network but you can configure multiple candidate-BSRs. In this case, once a BSR fails, you can switch over to another BSR. A BSR is elected among the C-BSRs automatically. The C-BSR with the highest priority is elected as the BSR. If the priority is the same, the C-BSR with the largest IP address is elected as the BSR.

III. Configuring static RP

The router that serves as the RP is the core router of multicast routes. If the dynamic RP elected by BSR mechanism is invalid for some reason, the static RP can be configured to specify RP. As the backup of dynamic RP, static RP improves network robusticity and enhances the operation and management capability of multicast network.

29.2 PIM-SM Configuration

- 1) PIM-SM basic configuration includes:
 - Enabling Multicast
 - Enabling PIM-SM
 - Entering the PIM view
 - Configuring candidate-BSRs
 - Configuring candidate-RPs
 - Configuring static RP
- 2) PIM-SM advanced configuration includes:
 - Configuring the PIM-SM domain boundary
 - Configuring the sending interval for the Hello packets of the interface
 - Configuring the filtering of multicast source/group
 - Configuring the filtering of PIM neighbor
 - Configuring the maximum number of PIM neighbor on an interface
 - Configuring RP to filter the register messages
 - Limiting the range of legal BSR
 - Limiting the range of legal C-RP
 - Clearing multicast route entries from PIM routing table
 - Clearing PIM neighbor

It should be noted that at least one router in an entire PIM-SM domain should be configured with Candidate-RPs and Candidate-BSRs.

29.2.1 Enabling Multicast

Refer to Chapter 26 Common Multicast Configuration.

29.2.2 Enabling PIM-SM

This configuration can be effective only after multicast is enabled.

Perform the following configuration in VLAN interface view.

Table 29-1 Enabling PIM-SM

Operation	Command
Enable PIM-SM on an interface	pim sm
Disable PIM-SM on an interface	undo pim sm

Repeat this configuration to enable PIM-SM on other interfaces. Only one multicast routing protocol can be enabled on an interface at a time.

Once enabled PIM-SM on an interface, PIM-DM cannot be enabled on the same interface and vice versa.

29.2.3 Entering the PIM View

Refer to 28.2.4 Entering the PIM View.

29.2.4 Configuring the Time Intervals for Ports to Send Hello Packets

In general, PIM-SM broadcasts Hello packets on the PIM-SM-enabled port periodically to detect PIM neighbors and determine the designated router (DR).

For details, refer to 28.2.3 Configuring the Time Intervals for Ports to Send Hello Packets.

29.2.5 Configuring Candidate-BSRs

In a PIM domain, one or more candidate BSRs should be configured. A BSR (Bootstrap Router) is elected among candidate BSRs. The BSR takes charge of collecting and advertising RP information.

The automatic election among candidate BSRs is described as follows:

One interface which has started PIM-SM must be specified when configuring the router as the candidate BSR.

At first, each candidate BSR considers itself as the BSR of the PIM-SM domain, and sends Bootstrap message by taking the IP address of the interface as the BSR address.

When receiving Bootstrap messages from other routers, the candidate BSR will compare the BSR address of the newly received Bootstrap message with that of itself. Comparison standards include priority and IP address. The bigger IP address is considered better when the priority is the same. If the priority of the former is higher, the candidate BSR will replace its BSR address and stop regarding itself as the BSR. Otherwise, the candidate BSR will keep its BSR address and continue to regard itself as the BSR.

Perform the following configuration in PIM view.

Table 29-2 Configuring candidate-BSRs

Operation	Command
Configure a candidate-BSR	c-bsr Vlan-interface <i>Vlan-interface-number</i> <i>hash-mask-len</i> [<i>priority</i>]
Remove the candidate-BSR configured	undo c-bsr

Candidate-BSRs should be configured on the routers in the network backbone. By default, no BSR is set. The default priority is 0.



Caution:

One router can only be configured with one candidate-BSR. When a candidate-BSR is configured on another interface, it will replace the previous configuration.

29.2.6 Configuring Candidate-RPs

In PIM-SM, the shared tree built by multicast routing data is rooted at the RP. There is a mapping from a multicast group to an RP. A multicast group can be mapped to only one RP. Different multicast groups can be mapped to the same RP or different RPs.

Perform the following configuration in PIM view.

Table 29-3 Configuring candidate-RPs

Operation	Command
Configure a candidate-RP	c-rp <i>interface-type interface-number</i> [group-policy <i>acl-number</i> priority <i>priority-value</i>]*

Operation	Command
Remove the candidate-RP configured	undo c-rp { <i>interface-type interface-number</i> all }

When configuring RP, if the range of the served multicast group is not specified, the RP will serve all multicast groups. Otherwise, the range of the served multicast group is the multicast group in the specified range. It is suggested to configure Candidate RP on the backbone router.

29.2.7 Configuring Static RP

Static RP serves as the backup of dynamic RP, so as to improve network robusticity. Perform the following configuration in PIM view.

Table 29-4 Configuring static RP

Operation	Command
Configure static RP	static-rp <i>rp-address</i> [<i>acl-number</i>]
Remove the configured static RP	undo static-rp

Basic ACL can control the range of multicast group served by static RP.

If static RP is in use, all routers in the PIM domain must adopt the same configuration. If the configured static RP address is the interface address of the local router whose state is UP, the router will function as the static RP. It is unnecessary to enable PIM on the interface that functions as static RP.

When the RP elected from BSR mechanism is valid, static RP does not work.

29.2.8 Configuring the PIM-SM Domain Border

After the PIM-SM domain border is configured, bootstrap messages can not cross the border in any direction. In this way, the PIM-SM domain can be split.

Perform the following configuration in VLAN interface view.

Table 29-5 Configuring the PIM-SM domain border

Operation	Command
Set the PIM-SM domain border	pim bsr-boundary
Remove the PIM-SM domain border configured	undo pim bsr-boundary

By default, no domain border is set. After this configuration is performed, a bootstrap message can not cross the border but other PIM packets can. This configuration can effectively divide a network into domains using different BSRs.

29.2.9 Configuring the filtering of multicast source/group

Refer to 28.2.5 Configuring the Filtering of Multicast Source/Group.

29.2.10 Configuring the filtering of PIM neighbor

Refer to 28.2.6 Configuring the Filtering of PIM Neighbor.

Refer to 28.2.7 Configuring the Maximum Number of PIM Neighbor on an Interface.

29.2.11 Configuring RP to Filter the Register Messages Sent by DR

In the PIM-SM network, the register message filtering mechanism can control which sources to send messages to which groups on the RP, i.e., RP can filter the register messages sent by DR to accept specified messages only.

Perform the following configuration in PIM view.

Table 29-6 Configuring RP to filter the register messages sent by DR

Operation	Command
Configure RP to filter the register messages sent by DR	register-policy <i>acl-number</i>
Cancel the configured filter of messages	undo register-policy

If an entry of a source group is denied by the ACL, or the ACL does not define operation to it, or there is no ACL defined, the RP will send RegisterStop messages to the DR to prevent the register process of the multicast data stream.



Caution:

Only the register messages matching the ACL permit clause can be accepted by the RP. Specifying an undefined ACL will make the RP deny all register messages.

29.2.12 Limiting the range of legal BSR

To prevent the legal BSR from being replaced maliciously in the network, you can limit the range of legal BSR. Other BSR messages beyond the range are not received by the router and thus ensure the BSR security.

Perform the following configuration in PIM view.

Table 29-7 Limiting the range of legal BSR

Operation	Command
Set the limit legal BSR range	bsr-policy <i>acl-number</i>
Restore to the default setting	undo bsr-policy

For detailed information of **bsr-policy**, please refer to the command manual.

29.2.13 Limiting the range of legal C-RP

To avoid C-RP spoofing, you can limit the range of legal C-RP and limit the groups that each C-RP servers.

Perform the following configuration in PIM view.

Table 29-8 Limiting the range of legal C-RP

Operation	Command
Set the limit legal C-RP range	crp-policy <i>acl-number</i>
Restore to the default setting	undo crp-policy

For detailed information of **crp-policy**, please refer to the command manual.

29.2.14 Clearing multicast route entries from PIM routing table

Refer to Chapter 28 PIM-DM Configuration.

29.2.15 Clearing PIM Neighbors

Refer to Chapter 28 PIM-DM Configuration.

29.3 Displaying and Debugging PIM-SM

After the above configuration, execute **display** command in any view to display the running of PIM-SM configuration, and to verify the effect of the configuration.

Execute **debugging** command in user view for the debugging of PIM-SM.

Table 29-9 Displaying and debugging PIM-SM

Operation	Command
Display the BSR information	display pim bsr-info
Display the RP information	display pim rp-info [<i>group-address</i>]
Enable the PIM-SM debugging	debugging pim sm { all mrt msdp verbose warning mbr { alert fresh } { recv send } { assert bootstrap crpadv jp reg regstop } timer { assert bsr crpadv jp jpgdelay mrt probe spt }
Disable the PIM-SM debugging	undo debugging pim sm { all mrt msdp verbose warning mbr { alert fresh } { recv send } { assert bootstrap crpadv jp reg regstop } timer { assert bsr crpadv jp jpgdelay mrt probe spt }

29.4 PIM-SM Configuration Example

I. Networking requirements

In actual network, we assume that the switches can intercommunicate and the IP address of each VLAN interface has been configured.

- LS_A is connected to LS_B through VLAN-interface10, connected to HostA through VLAN-interface11 and connected to LS_C through VLAN-interface12.
- LS_B is connected to LS_A through VLAN-interface10, connected to LS_C through VLAN-interface11 and connected to LS_D through VLAN-interface12.
- LS_C is connected to HostB through VLAN-interface10, connected to LS_B through VLAN-interface11 and connected to LS_A through VLAN-interface12.

Suppose that Host A is the receiver of the multicast group at 225.1.1.1. Host B begins transmitting data destined to 225.1.1.1. LS_A receives the multicast data from Host B via LS_B.

II. Networking diagram

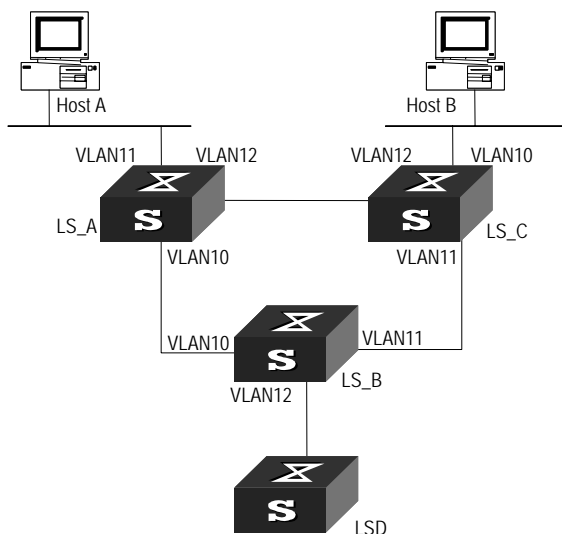


Figure 29-2 PIM-SM configuration networking

III. Configuration procedure

1) Configure LS_A

Enable PIM-SM.

```
[SW8800] multicast routing-enable
[SW8800] vlan 10
[SW8800-vlan10] port ethernet 2/1/2 to ethernet 2/1/3
[SW8800-vlan10] quit
[SW8800] interface vlan-interface 10
[SW8800-vlan-interface10] igmp enable
[SW8800-vlan-interface10] pim sm
[SW8800-vlan-interface10] quit
[SW8800] vlan 11
[SW8800-vlan11] port ethernet 2/1/4 to ethernet 2/1/5
[SW8800-vlan11] quit
[SW8800] interface vlan-interface 11
[SW8800-vlan-interfacel1] igmp enable
[SW8800-vlan-interfacel1] pim sm
[SW8800-vlan-interfacel1] quit
[SW8800] vlan 12
[SW8800-vlan12] port ethernet 2/1/6 to ethernet 2/1/7
[SW8800-vlan12] quit
[SW8800] interface vlan-interface 12
[SW8800-vlan-interface12] igmp enable
[SW8800-vlan-interface12] pim sm
```

```
[SW8800-vlan-interface12] quit
```

2) Configure LS_B

Enable PIM-SM.

```
[SW8800] multicast routing-enable
[SW8800] vlan 10
[SW8800-vlan10] port ethernet 2/1/2 to ethernet 2/1/3
[SW8800-vlan10] quit
[SW8800] interface vlan-interface 10
[SW8800-vlan-interface10] igmp enable
[SW8800-vlan-interface10] pim sm
[SW8800-vlan-interface10] quit
[SW8800] vlan 11
[SW8800-vlan11] port ethernet 2/1/4 to ethernet 2/1/5
[SW8800-vlan11] quit
[SW8800] interface vlan-interface 11
[SW8800-vlan-interface11] igmp enable
[SW8800-vlan-interface11] pim sm
[SW8800-vlan-interface11] quit
[SW8800] vlan 12
[SW8800-vlan12] port ethernet 2/1/6 to ethernet 2/1/7
[SW8800-vlan12] quit
[SW8800] interface vlan-interface 12
[SW8800-vlan-interface12] igmp enable
[SW8800-vlan-interface12] pim sm
[SW8800-vlan-interface12] quit
```

Configure the C-BSR.

```
[SW8800] pim
[SW8800-pim] c-bsr vlan-interface 10 30 2
```

Configure the C-RP.

```
[SW8800] acl number 2000
[SW8800-acl-basic-2000] rule permit source 225.0.0.0 0.255.255.255
[SW8800] pim
[SW8800-pim] c-rp vlan-interface 10 group-policy 2000
```

Configure PIM domain border.

```
[SW8800] interface vlan-interface 12
[SW8800-vlan-interface12] pim bsr-boundary
```

After VLAN-interface 12 is configured as domain border, the LS_D will be excluded from the local PIM domain and cannot receive the BSR information transmitted from LS_B any more.

3) Configure LS_C.

Enable PIM-SM.

```
[SW8800] multicast routing-enable
[SW8800] vlan 10
[SW8800-vlan10] port ethernet 2/1/2 to ethernet 2/1/3
[SW8800-vlan10] quit
[SW8800] interface vlan-interface 10
[SW8800-vlan-interface10] igmp enable
[SW8800-vlan-interface10] pim sm
[SW8800-vlan-interface10] quit
[SW8800] vlan 11
[SW8800-vlan11] port ethernet 2/1/4 to ethernet 2/1/5
[SW8800-vlan11] quit
[SW8800] interface vlan-interface 11
[SW8800-vlan-interface11] igmp enable
[SW8800-vlan-interface11] pim sm
[SW8800-vlan-interface11] quit
[SW8800] vlan 12
[SW8800-vlan12] port ethernet 2/1/6 to ethernet 2/1/7
[SW8800-vlan12] quit
[SW8800] interface vlan-interface 12
[SW8800-vlan-interface12] igmp enable
[SW8800-vlan-interface12] pim sm
[SW8800-vlan-interface12] quit
```

Chapter 30 MSDP Configuration

30.1 MSDP Overview

30.1.1 Introduction

Multicast source discovery protocol (MSDP) is used to discover multicast source information in other PIM-SM domains. No ISP would like to forward multicast traffic depending on the RP of competitors, though it has to obtain information from the source and distribute it among its members, regardless of the location of the source RP. MSDP is proposed to solve this problem. MSDP describes interconnection mechanism of multiple PIM-SM domains. MSDP allows the RPs of different domains to share the multicast source information, but all these domains must use PIM-SM as their intro-domain multicast routing protocol.

A RP configured with MSDP peer notifies all of its MSDP peers of the active multicast source message in its domain via SA (Source Active) message. In this way, multicast source information in a PIM-SM domain is transmitted to another PIM-SM domain.

MSDP peer relationship can be established between RPs in different domains or in a same domain, between a RP and a common router, or between common routers. The connection between MSDP peers is TCP connection.

MSDP makes a PIM-SM domain independent of the RP in another PIM-SM domain. After getting multicast source information in that domain, the receiver here can join directly to the SPT of the multicast source in that domain.

Another application of MSDP is Anycast RP. In a domain, configure a certain interface (usually Loopback interface) on different routers with a same IP address; designate these interfaces as C-RPs; and create MSDP peer relationship among them. After the unicast route convergence, the multicast source can select the nearest RP for registration, and the receiver can also select the nearest RP to add into its RPT. The RPs exchange individual registration source information via MSDP peers. Therefore, every RP knows all multicast sources of the entire domain; and every receiver on each RP can receive multicast data from all the multicast sources in the entire domain.

By initiating registration and RPT joining to the nearest RP, MSDP implements RP load sharing. Once an RP turns invalid, its original registered source and receivers will select another nearest RP, implementing redundant RP backup.

In addition, MSDP only accepts the SA messages from the correct paths and excludes redundant SA messages through RPF check mechanism, and prevents the flooding of SA messages among MSDP peers by configuring Mesh Group.

30.1.2 Working Principle

I. Identifying multicast source and receiving multicast data

As shown in Figure 30-1, the RPs of PIM-SM domains 1, 2 and 3 establish peer relationship between them. Domain 3 contains a group member.

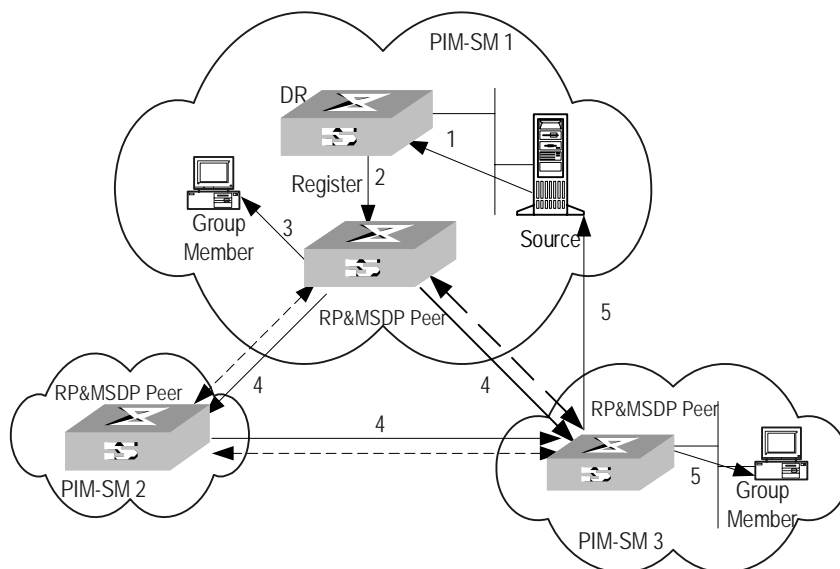


Figure 30-1 MSDP working principles (I)

When the multicast source in domain 1 sends data to the multicast group, the working process of the member in domain 3, from discovering the multicast source to receiving data from the source, includes the following:

- 1) The multicast source in PIM-SM domain 1 begins to send datagram.
- 2) The DR connected to the multicast source encapsulates the datagram into a Register packet and forward to the RP in domain 1.
- 3) The RP in domain 1 decapsulates the packet and forwards it along the RPT to all the members within the domain. The domain members can choose to take the path along SPT.
- 4) The RP in domain 1 generates an SA (Source Active) message for the MSDP peers (the RPs in PIM-SM domain 2 and domain 3). The SA message contains multicast source IP address, multicast group address and the address of the RP that generates the message. Besides, the RP in domain 1 encapsulates the first received multicast data into this SA message.
- 5) If there is any group member in the domain of an MSDP peer (in the figure, it is PIM-SM domain 3), the RP in this domain sends the multicast data encapsulated in the SA message to group members along the RPT and the join message to multicast source.
- 6) After the reverse forwarding path is created, the multicast source data is sent directly to the RP in domain 3, which then RP forwards the data along the RPT. In

this case, the last hop router connected with the group member in domain 3 can choose whether to switch to SPT.

II. Message forwarding and RPF check between MSDP peers

As shown in Figure 30-2 MSDP working principles (II), Switch A, Switch B, Switch C, Switch D, Switch E and Switch F belong to domain 1, domain 2 and domain 3 respectively. MSDP peer relationship is established between them, indicated with bi-directional arrows in the figure. Among them, Mesh Group is created among Switch B, Switch C and Switch D.

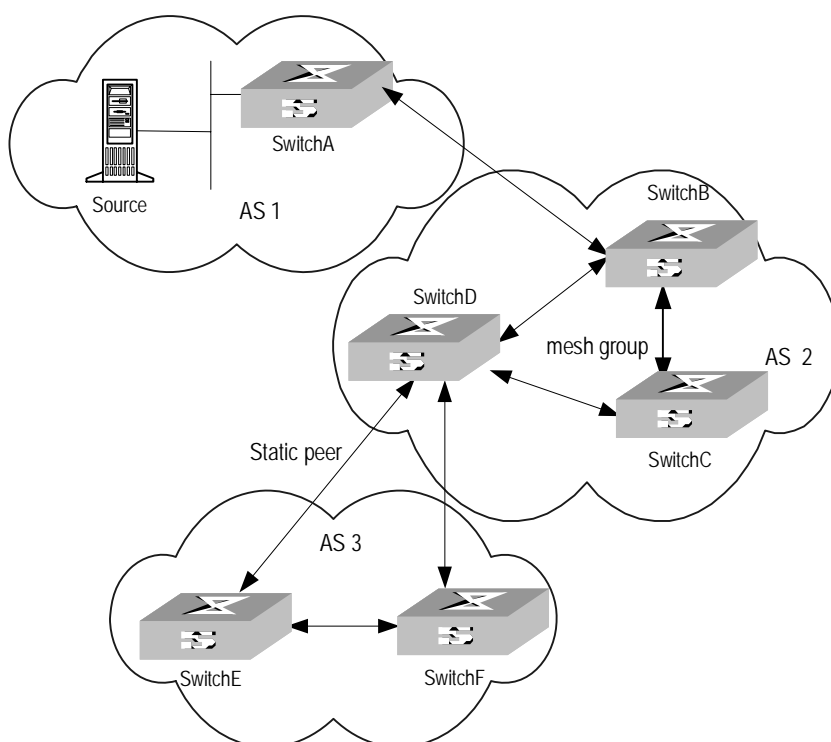


Figure 30-2 MSDP working principles (II)

The SA message forwarding and RPF check among these MSDP peers are illustrated as follows:

- 1) If the SA message is from a MSDP peer that is the RP of the multicast source as from Switch A to Switch B, it is received and forwarded to other peers.
- 2) If the SA message is from a MSDP peer that has only one peer as from Switch B to Switch A, it is received.
- 3) If the SA message is from a static RPF peer as from Switch D to Switch E, it is received and forwarded to other peers.
- 4) If the SA message is from a MSDP peer in Mesh Group as from Switch B to Switch D, it is received and forwarded to the peers outside the Mesh Group.
- 5) If the SA message is sent from a MSDP peer in a same domain, and the peer is the next hop along the optimal path to the RP in the domain of source, as in the

case when the message is from Switch E to Switch F, it is received and forwarded to other peers.

- 6) If the SA message is sent from a MSDP peer in a different domain which is the next autonomous domain along the optimal path to the RP in the domain of source, as from Switch D to Switch F, it is received and forwarded to other peers.
- 7) For other SA messages, they are neither received nor forwarded.

III. Precautions for configuration

The router operating MSDP must also run BGP or MBGP. It is recommended to use the same IP address of the MSDP peer with that of the BGP peer or MBGP peer. If neither BGP nor MBGP is in operation, a static RPF peer should be configured.

30.2 MSDP Configuration

- 1) Basic configuration tasks of MSDP include:
 - Enable MSDP
 - Configure MSDP peers
- 2) Advanced configuration tasks of MSDP include:
 - Configure static RPF peers
 - Configure Originating RP
 - Configure SA caching state
 - Configure the maximum number of SA caching
 - Request the source information of MSDP peers
 - Control the source information created
 - Control the source information forwarded
 - Control the received source information
 - Configure MSDP full connection group
 - Configure the MSDP connection retry period
 - Disable MSDP peers
 - Clear MSDP connection, statistics and SA cache

30.2.1 Enabling MSDP

To configure MSDP, you must enable MSDP first.

Please perform the following configurations in system view.

Table 30-1 Enabling MSDP

Operation	Command
Enable MSDP and enter MSDP view	msdp
Clear all MSDP configurations	undo msdp

30.2.2 Configuring MSDP Peers

To run MSDP, you need to configure MSDP peers locally.

Please perform the following configurations in MSDP view.

Table 30-2 Configuring MSDP peers

Operation	Command
Configure MSDP peers	peer peer-address connect-interface interface-type interface-number
Remove MSDP peer configuration	undo peer peer-address
Add description to a MSDP peer	peer peer-address description text
Remove the description	undo peer peer-address description text

The command to add description is optional.

If the local router is also in BGP Peer relation with a MSDP peer, the MSDP peer and the BGP peer should use the same IP address.

Not any two routers between which MSDP peer relationship has been established must run BGP or MBGP, so long as they have a BGP or MBGP route between them. If no BGP or MBGP route exists between them, then you must configure static RPF peers.

30.2.3 Configuring Static RPF Peers

Please perform the following configurations in MSDP view.

Table 30-3 Configuring static RPF peers

Operation	Command
Configure static RPF peers	static-rpf-peer peer-address [rp-policy list]
Remove static RPF peer configuration	undo static-rpf-peer peer-address

By default, no static RPF peer is configured.

The **peer** command must be configured before the configuration of **static-rpf-peer** command.

If only one MSDP peer is configured via the peer command, the MSDP peer will be regarded as the static RPF peer.

To configure multiple static RPF peers at the same time, take any of the two methods:

- Using **rp-policy** parameters universally: Multiple static RPF peers take effect at the same time and SA messages are filtered by the RP addresses contained according to the configured prefix list. If multiple static RPF peers using the same

rp-policy parameter are configured, any peer that receives an SA message will forward it to the other peers.

- Not using the **rp-policy** parameter universally: According to the configuration sequence, only the first static RPF peer whose connection state is UP is activated. All SA messages from the peer will be received and those from other static RPF peers will be discarded. Once the activated static RPF peer turns invalid (possibly out of configuration removed or connection interrupted), the following first static RPF peer with UP connection state according to the configuration sequence will assume its role.

30.2.4 Configuring Originating RP

During the creation of SA message, an MSDP peer can be configured to use the IP address of a specified interface as the RP address in its SA message.

Please perform the following configurations in MSDP view.

Table 30-4 Configuring Originating RP

Operation	Command
Configure an MSDP peer to use the IP address of a specified interface as the RP address of its SA message	originating-rp interface-type interface-number
Remove the above operation	undo originating-rp

By default, the RP address in SA message is the one configured by PIM.

30.2.5 Configuring SA Caching State

When SA messages are cached on a router, the new join-in groups can directly access all the active sources and join to the corresponding source tree, instead of waiting for the arrival of the next SA message.

Please perform the following configurations in MSDP view.

Table 30-5 Configuring SA caching state

Operation	Command
Configure SA caching state	cache-sa-enable
Disable SA caching state	undo cache-sa-enable

By default, the router caches the SA state, or rather the (S, G) entry when receiving an SA message.

Some memory is consumed as the join delay of groups is shortened by this configuration.

30.2.6 Configuring the Maximum Number of SA caching

To prevent DoS (Deny of Service) attacks, you can set the maximum number of SAs cached on the router.

Perform the following configuration in MSDP view.

Table 30-6 Configuring the maximum number of SA caching

Operation	Command
Configuring the maximum number of SA caching	peer <i>peer-address</i> sa-cache-maximum <i>sa-limit</i>
Restore the default configuration	undo peer <i>peer-address</i> sa-cache-maximum

By default, the maximum number of SA caching is 2048.

30.2.7 Requesting Source Information of MSDP Peers

When a new group joins, the router will send a SA request message to the specified MSDP peer, and the MSDP peer will respond with the SA messages it caches. If the MSDP peer does not enable the SA caching, the configuration is invalid.

Please perform the following configurations in MSDP view.

Table 30-7 Requesting source information of MSDP peers

Operation	Command
Configure the router to send SA request message to the specified MSDP peer when receiving the join message of a group	peer <i>peer-address</i> request-sa-enable
Restore the default configuration	undo peer <i>peer-address</i> request-sa-enable

The SA request message sent by a local RP will get the immediate response about all active sources.

By default, the router does not send SA request message to its MSDP peer when receiving the join message of a group. Instead, it waits for the arrival of SA message of the next period.

30.2.8 Controlling the Source Information Created

I. Filtering the multicast routing entries imported

RP filters the registered sources to control the information of the active sources advertised in SA message. MSDP peers can be configured to only advertise the

qualified (S, G) entries in the multicast routing table when creating SA messages, that is, to control the (S,G) entries imported from the multicast routing table to the domain.

Please perform the following configurations in MSDP view.

Table 30-8 Filtering the multicast routing entries imported

Operation	Command
Advertise only the (S, G) entries permitted by the ACL	import-source [<i>acl acl-number</i>]
Remove the above configuration	undo import-source

By default, only intra-domain sources are advertised in SA messages.

If the import-source command without **acl** parameter is executed, no source is advertised in SA messages.

II. Filtering SA request messages

Please perform the following configurations in MSDP view.

Table 30-9 Filtering SA request messages

Operation	Command
Filter all the SA request messages from a specified MSDP peer	peer <i>peer-address</i> sa-request-policy
Filter the SA request messages of the groups of a specified MSDP peer permitted by the basic ACL from	peer <i>peer-address</i> sa-request-policy <i>acl acl-number</i>
Remove the configuration of filtering SA request messages	undo peer <i>peer-address</i> sa-request-policy

By default, only the routers which caches SA messages can repond to SA request messages. Routers receive all SA request messages from its MSDP peers.

Multicast group addresses are described in ACL. If no ACL is specified, all SA request messages sent by the corresponding MSDP peer will be ignored. If an ACL is specified, only SA request messages of the groups permitted by the ACL will be processed.

30.2.9 Controlling the Source Information Forwarded

Controlling of source information also includes that of forwarding and receiving source information besides that of creating source information. The outbound filter or time to live (TTL) threshold of SA messages can be used to control the SA message forwarding. By default, all SA messages are forwarded to other MSDP peers.

I. Using MSDP outbound filter

MSDP outbound filter of are functional in:

- Filtering off all the (S, G) entries
- Forwarding only the SA messages permitted by the advanced ACL

Please perform the following configurations in MSDP view.

Table 30-10 Using MSDP outbound filter to control the source information forwarded

Operation	Command
Filter off all the SA messages to a specified MSDP peer	peer <i>peer-address</i> sa-policy export
Forward the SA messages permitted by the advanced ACL to a specified MSDP peer	peer <i>peer-address</i> sa-policy export [acl <i>acl-number</i>]
Remove the filtering over the source information forwarded	undo peer <i>peer-address</i> sa-policy export

II. Using TTL to filter SA messages with encapsulated data

An SA message with encapsulated data can reach the specified MSDP peer only when the TTL in its IP header is no less than the threshold. Therefore, the forwarding of SA messages with encapsulated data can be controlled by configuring the TTL threshold.

For example, you can set the TTL threshold for intra-domain multicast traffic as 10 if you wish to restrict SA messages with TTL less than or equal to 10 carrying encapsulated data from being propagated. If you set the TTL threshold greater than 10, then they can be propagated to outside.

Please perform the following configurations in MSDP view.

Table 30-11 Using TTL to filter SA messages with encapsulated data

Operation	Command
Filter off the multicast data encapsulated in the first SA message aiming at a specified MSDP peer	peer <i>peer-address</i> minimum-ttl <i>tth</i>
Remove the TTL threshold configuration	undo peer <i>peer-address</i> minimum-ttl

The default value of TTL threshold is 0.

30.2.10 Controlling the Received Source Information

Please perform the following configurations in MSDP view.

Table 30-12 Controlling the received source information

Operation	Command
Filter off the SA messages from a specified MSDP peer	peer <i>peer-address</i> sa-policy import
Receive the SA messages permitted by the advanced ACL from a specified MSDP peer	peer <i>peer-address</i> sa-policy import [acl <i>acl-number</i>]
Remove the filtering rule over received source information	undo peer <i>peer-address</i> sa-policy import

Similar to MSDP outbound filter in function, MSDP inbound filter controls the received SA messages. By default, the SA messages from all peers are accepted.

30.2.11 Configuring MSDP Mesh Group

Mesh Group is useful when full connection among MSDP peers is required but SA message flooding shall be prevented.

In a Mesh group, the SA messages from outside the group are forwarded to other members in the group, but the SA messages from peers inside the group will not be performed with Peer-RPF check or forwarded in the group. In this case, the overflow of SA messages is avoided and Peer-RPF is simplified, as BGP or MBGP is not required between MSDP peers.

Please perform the following configurations in MSDP view.

Table 30-13 Configuring MSDP full connection group

Operation	Command
Configure an MSDP peer to be a member of an MSDP Mesh Group	peer <i>peer-address</i> mesh-group <i>name</i>
Delete that member from the Group	undo peer <i>peer-address</i> mesh-group <i>name</i>

If an MSDP peer is configured into different mesh groups, only the last configuration is valid.

30.2.12 Configuring the MSDP Connection Retry Period

Perform the following configurations in MSDP view.

Table 30-14 Configuring the MSDP connection retry period

Operation	Command
Configuring the MSDP connection retry period	timer retry <i>seconds</i>
Restore the default value of MSDP connection retry interval	undo timer retry

By default, MSDP connection is retried at the interval of 30 seconds.

30.2.13 Shutting MSDP Peers Down

The session between MSDP peers can be cut off and re-activated as needed.

If a session between MSDP peers is cut off, the TCP connection will terminate with no retry effort, but the configuration information will be reserved.

Please perform the following configurations in MSDP view.

Table 30-15 Shutting MSDP peers down

Operation	Command
Shut a specified MSDP peer down	shutdown <i>peer-address</i>
Turn the MSDP peer up	undo shutdown <i>peer-address</i>

By default, MSDP peer is enabled.

30.2.14 Clearing MSDP Connections, Statistics and SA Caching Configuration

Perform the following configurations in user view.

Table 30-16 Clearing MSDP connections, statistics and SA caching configuration

Operation	Command
Clear a specified TCP connection and reset the counters of all MSDP information	reset msdp peer <i>peer-address</i>
Clear MSDP peer statistics	reset msdp statistics [<i>peer-address</i>]
Clear cached SA entries of MSDP	reset msdp sa-cache [<i>group-address</i>]

30.3 Displaying and Debugging MSDP

I. Displaying and Debugging MSDP

After the above configuration, execute **display** commands in any view to display the running information of MSDP and to verify the effect of the configuration.

Execute the **debugging** command in user view for the debugging of MSDP.

Table 30-17 Displaying and debugging MSDP configuration

Operation	Command
Display the numbers of sources and groups of SA messages from a specified autonomous domain	display msdp sa-count [<i>as-number</i>]
Display the details of a MSDP peer	display msdp peer-status [<i>peer-address</i>]
Display the (S,G) state learnt from MSDP peer	display msdp sa-cache [<i>group-address</i> [<i>source-address</i>]] [<i>autonomous-system-number</i>]
Display MSDP peer state	display msdp brief
Enable MSDP debugging	debugging msdp { all connect event packet source-active }

Note that only after the **cache-sa-enable** command is executed, will the **display msdp sa-count** command have output.

II. Tracing the Transmission Path of SA Messages on the Network

The **mtracert** command can be used in any view to trace the network path of multicast data from multicast source to destination receiver and locate faults.

Table 30-18 Tracing the transmission path of SA messages on the network

Operation	Command
Trace the transmission path of SA messages on the network	msdp-tracert { <i>source-address</i> } { <i>group-address</i> } { <i>rp-address</i> } [max-hops <i>max-hops</i>] [next-hop-info] [sa-info] [peer-info] [skip-hops <i>skip-hops</i>]

Locating information loss and reducing configuration faults can be realized by tracing the network path of the specified (S, G, RP) entries. After the transmission path of SA messages is determined, the overflow of SA messages can be avoided by the correct configuration.

30.4 MSDP Configuration Examples

30.4.1 Configuring Static RPF Peers

I. Networking requirements

In the following networking environment, four Switch 8800s all are in the PIM-SM domains with no BGP or MBGP running among them (Note that MBGP is not supported in the basic code; the extended option is required).

To enable Switch D to receive the specified source information from PIM-SM domains 1, 2 and 3, you can configure static RPF peers with the parameter **rp-policy**.

After the configuration is complete, Switch D will only receive SA messages permitted by the corresponding filtering policy from its static RPF peers.

II. Networking diagram

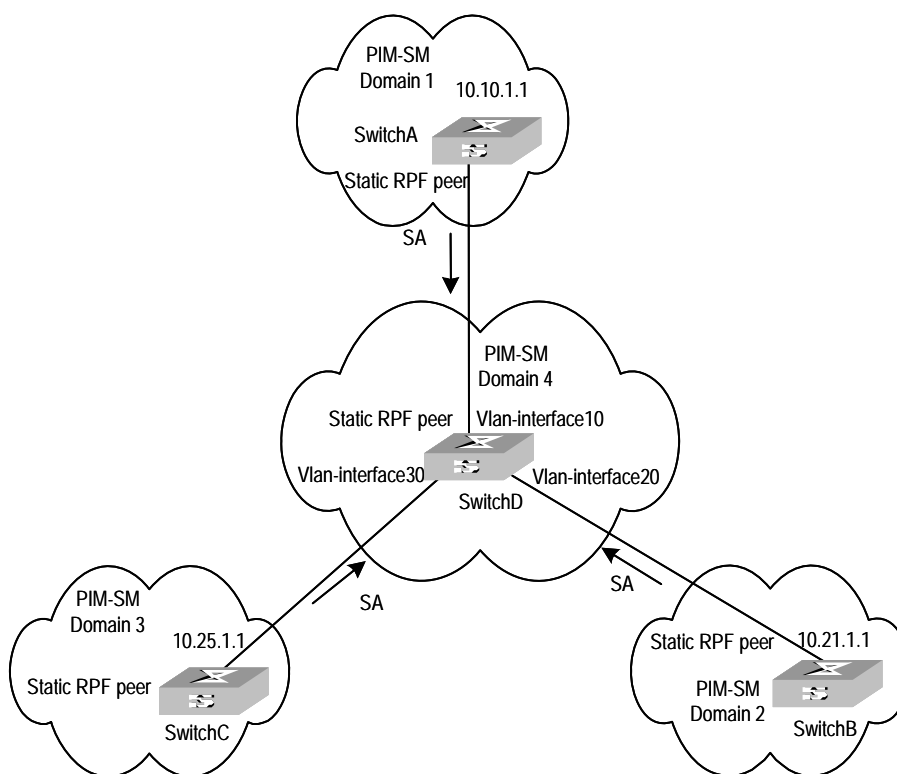


Figure 30-3 Configuring static RPF peers

III. Configuration procedure

Configure Switch A to be a static RPF peer of Switch D.

```
[SwitchD] ip ip-prefix list-a permit 10.10.0.0 16
[SwitchD] msdp
[SwitchD-msdp] peer 10.10.1.1 connect-interface Vlan-interface 10
```



```
[SwitchD-msdp] static-rpf-peer 10.10.1.1 rp-policy list-a
```

Configure Switch B to be a static RPF peer of Switch D.

```
[SwitchD] ip ip-prefix list-b permit 10.21.0.0 16
```

```
[SwitchD] msdp
```

```
[SwitchD-msdp] peer 10.21.1.1 connect-interface Vlan-interface 20
```

```
[SwitchD-msdp] static-rpf-peer 10.21.1.1 rp-policy list-b
```

Configure Switch C to be a static RPF peer of Switch D.

```
[SwitchD] ip ip-prefix list-c permit 10.25.0.0 16
```

```
[SwitchD] msdp
```

```
[SwitchD-msdp] peer 10.25.1.1 connect-interface Vlan-interface30
```

```
[SwitchD-msdp] static-rpf-peer 10.25.1.1 rp-policy list-c
```

30.4.2 Configuring Anycast RP

I. Networking requirements

To configure Anycast RP in the PIM-SM domain, establish MSDP peer relationship between Switch A and Switch B; use the address of loopback0 on Switch A and Switch B to send SA messages outside; set Loopback10 interface on Switch A and Switch B as BSR/RP and configure the Anycast RP address. In this way, when a unicast group member joins, the switch directly connected to the host can originate a join message to the nearest RP in the topology.

This example focuses on the configuration of Switch A and Switch B. Configuration performed on Switch E, Switch D and Switch C is omitted as it mainly concerns enabling multicast and enabling PIM-SM on the interfaces.

II. Networking diagram

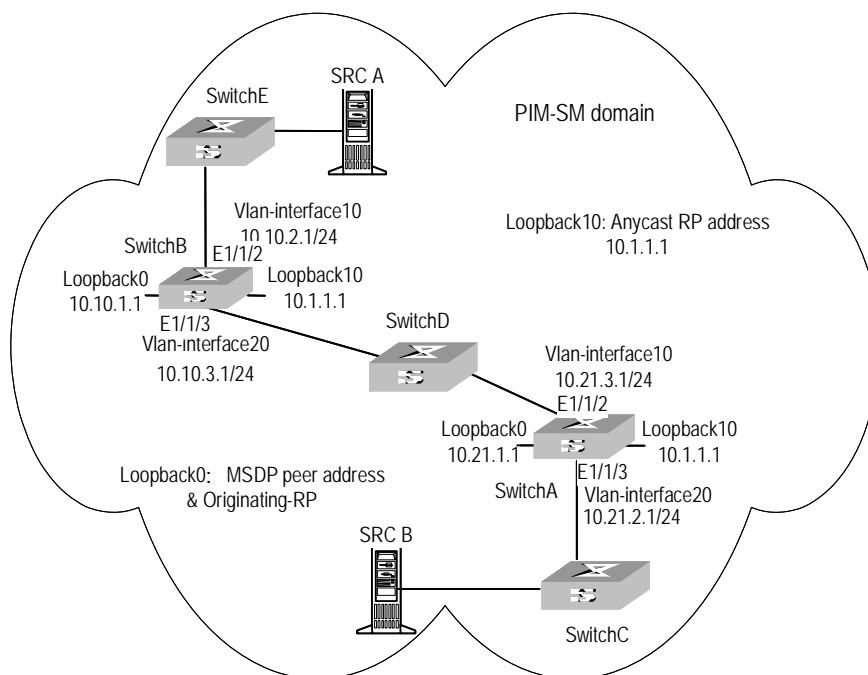


Figure 30-4 Networking diagram for Anycast RP configuration

III. Configuration procedure

1) Configure SwitchB:

Configure VLAN

```
<SwitchB> system-view
[SwitchB] vlan 10
[SwitchB-vlan10] port ethernet1/1/2
[SwitchB-vlan10] quit
[SwitchB] vlan 20
[SwitchB-vlan20] port ethernet1/1/3
[SwitchB-vlan20] quit
```

Enable multicast.

```
[SwitchB] multicast routing-enable
```

Configure the IP address of interface loopback0.

```
[SwitchB] interface loopback0
[SwitchB-LoopBack0] ip address 10.10.1.1 255.255.255.255
[SwitchB-LoopBack0] quit
```

Configure the IP address of interface loopback10 and enable IGMP and PIM-SM.

```
[SwitchB] interface loopback10
[SwitchB-LoopBack10] ip address 10.1.1.1 255.255.255.255
```

```
[SwitchB-LoopBack10] igmp enable
[SwitchB-LoopBack10] pim sm
[SwitchB-LoopBack10] quit
```

Configure the IP address of Vlan-interface10 and enable IGMP and PIM-SM.

```
[SwitchB] interface Vlan-interface10
[SwitchB-Vlan-interface10] ip address 10.10.2.1 255.255.255.0
[SwitchB-Vlan-interface10] igmp enable
[SwitchB-Vlan-interface10] pim sm
[SwitchB-Vlan-interface10] undo shutdown
[SwitchB-Vlan-interface10] quit
```

Configure the IP address of Vlan-interface20 and enable IGMP and PIM-SM.

```
[SwitchB] interface Vlan-interface20
[SwitchB-Vlan-interface20] ip address 10.10.3.1 255.255.255.0
[SwitchB-Vlan-interface20] igmp enable
[SwitchB-Vlan-interface20] pim sm
[SwitchB-Vlan-interface20] undo shutdown
[SwitchB-Vlan-interface20] quit
```

Configure OSPF

```
[SwitchB] ospf
[SwitchB-ospf-1] area 0
[SwitchB-ospf-1-area-0.0.0.0] network 10.10.2.0 0.255.255.255
[SwitchB-ospf-1-area-0.0.0.0] network 10.10.3.0 0.255.255.255
[SwitchB-ospf-1-area-0.0.0.0] network 10.1.1.1 0.0.0.0
[SwitchB-ospf-1-area-0.0.0.0] network 10.10.1.1 0.0.0.0
[SwitchB-ospf-1-area-0.0.0.0] quit
[SwitchB-ospf-1] quit
```

Configure Switch A as its MSDP peer.

```
[SwitchB] msdp
[SwitchB-msdp] peer 10.21.1.1 connect-interface loopback 0
```

Configure Originating RP.

```
[SwitchB-msdp] originating-rp loopback0
[SwitchB-msdp] quit
```

Configure C-RP and BSR.

```
[SwitchB] pim
[SwitchB-pim] c-rp loopback 10
[SwitchB-pim] c-bsr loopback 10 30
```

2) Configure Switch A:

Configure VLAN

```
<SwitchA> system-view
```

```
[SwitchA] vlan 10
[SwitchA-vlan10] port ethernet1/1/2
[SwitchA-vlan10] quit
[SwitchA] vlan 20
[SwitchA-vlan20] port ethernet1/1/3
[SwitchA-vlan20] quit
```

Enable multicast.

```
[SwitchA] multicast routing-enable
```

Configure the IP address of interface loopback0.

```
[SwitchA] interface loopback0
[SwitchA-LoopBack0] ip address 10.21.1.1 255.255.255.255
[SwitchA-LoopBack0] quit
```

Configure the IP address of interface loopback10 and enable IGMP and PIM-SM.

```
[SwitchA] interface loopback10
[SwitchA-LoopBack10] ip address 10.1.1.1 255.255.255.255
[SwitchA-LoopBack10] igmp enable
[SwitchA-LoopBack10] pim sm
[SwitchA-LoopBack10] quit
```

Configure the IP address of interface Vlan-interface20 and enable IGMP and PIM-SM.

```
[SwitchA] interface Vlan-interface20
[SwitchA-Vlan-interface20] ip address 10.21.2.1 255.255.255.0
[SwitchA-Vlan-interface20] igmp enable
[SwitchA-Vlan-interface20] pim sm
[SwitchA-Vlan-interface20] undo shutdown
[SwitchA-Vlan-interface20] quit
```

Configure the IP address of Vlan-interface10 and enable IGMP and PIM-SM.

```
[SwitchA] interface Vlan-interface10
[SwitchA-Vlan-interface10] ip address 10.21.3.1 255.255.255.0
[SwitchA-Vlan-interface10] igmp enable
[SwitchA-Vlan-interface10] pim sm
[SwitchA-Vlan-interface10] undo shutdown
[SwitchA-Vlan-interface10] quit
```

Configure OSPF route.

```
[SwitchA] ospf
[SwitchA-ospf-1] area 0
[SwitchA-ospf-1-area-0.0.0.0] network 10.21.2.0 0.255.255.255
[SwitchA-ospf-1-area-0.0.0.0] network 10.21.3.0 0.255.255.255
[SwitchA-ospf-1-area-0.0.0.0] network 10.1.1.1 0.0.0.0
[SwitchA-ospf-1-area-0.0.0.0] network 10.21.1.1 0.0.0.0
[SwitchA-ospf-1-area-0.0.0.0] quit
```

```
[SwitchA-ospf-1] quit
```

Configure Switch B as its MSDP peer.

```
[SwitchA] msdp
```

```
[SwitchA-msdp] peer 10.10.1.1 connect-interface loopback 0
```

Configure Originating RP.

```
[SwitchA-msdp] originating-rp loopback0
```

```
[SwitchA-msdp] quit
```

Configure C-RP and BSR.

```
[SwitchA] pim
```

```
[SwitchA-pim] c-rp loopback 10
```

```
[SwitchA-pim] c-bsr loopback 10 30
```

30.4.3 MSDP Integrated Networking

I. Networking requirement

In the following network, enable MSDP and configure an Anycast RP in PIM-SM domain 1; establish MSDP peer relationship among RPs across PIM-SM domains; and use MBGP between domains. For the related commands, refer to “Multicast Protocol Configuration”.

II. Networking diagram

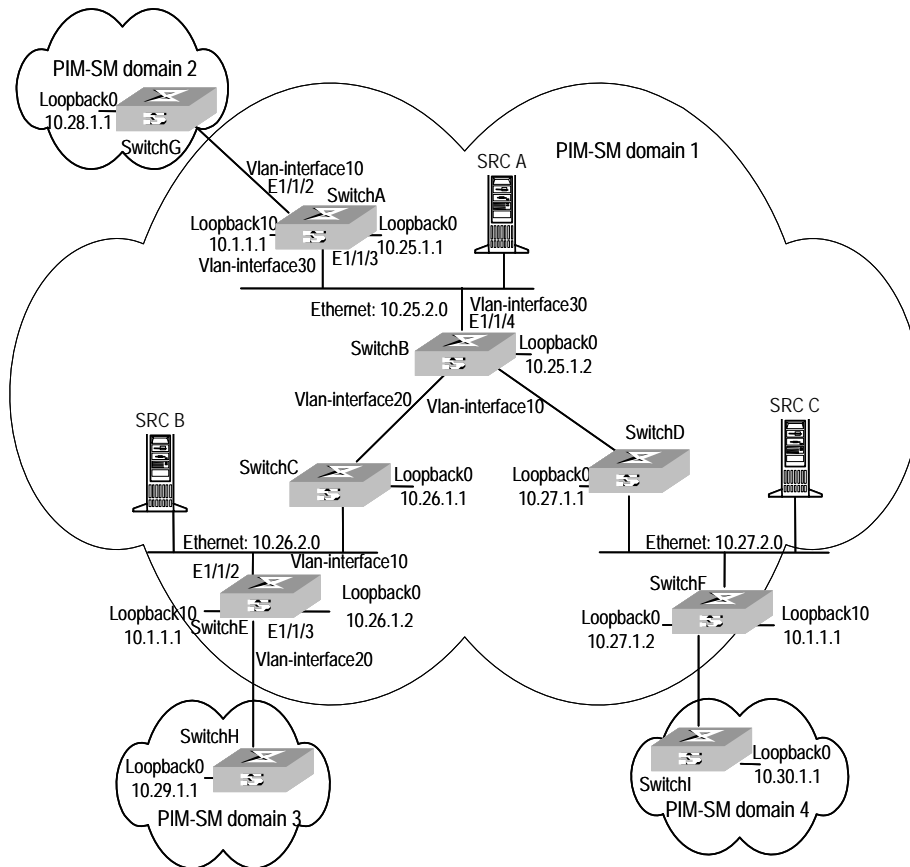


Figure 30-5 MSDP integrated networking

III. Configuration procedure

Note:

The follow procedure details multicast configuration, but briefs router configuration.

1) Configure Switch A:

Configuring VLAN

```
<SwitchA> system-view
[SwitchA] vlan 10
[SwitchA-vlan10] port ethernet1/1/2
[SwitchA-vlan10] quit
[SwitchA] vlan 30
[SwitchA-vlan30] port ethernet1/1/3
[SwitchA-vlan30] quit
```

Enable multicast.

```
[SwitchA] multicast routing-enable
```

Configure the IP address of interface loopback0 and enable PIM-SM.

```
[SwitchA] interface loopback0
[SwitchA-LoopBack0] ip address 10.25.1.1 255.255.255.255
[SwitchA-LoopBack0] pim sm
[SwitchA-LoopBack0] quit
```

Configure the IP address of interface loopback10 and enable PIM-SM.

```
[SwitchA] interface loopback10
[SwitchA-LoopBack10] ip address 10.1.1.1 255.255.255.255
[SwitchA-LoopBack10] pim sm
[SwitchA-LoopBack10] quit
```

Configure the IP address of Vlan-interface30 and enable IGMP and PIM-SM.

```
[SwitchA] interface Vlan-interface30
[SwitchA-Vlan-interface30] ip address 10.25.2.3 255.255.255.0
[SwitchA-Vlan-interface30] igmp enable
[SwitchA-Vlan-interface30] pim sm
[SwitchA-Vlan-interface30] undo shutdown
[SwitchA-Vlan-interface30] quit
```

Configure the IP address of Vlan-interface10 and enable IGMP and PIM-SM.

```
[SwitchA] interface Vlan-interface10
[SwitchA-Vlan-interface10] ip address 10.25.3.1 255.255.255.0
[SwitchA-Vlan-interface10] igmp enable
[SwitchA-Vlan-interface10] pim sm
[SwitchA-Vlan-interface10] undo shutdown
[SwitchA-Vlan-interface10] quit
```

Configure OSPF

```
[SwitchA] ospf
[SwitchA-ospf-1] area 0
[SwitchA-ospf-1-area-0.0.0.0] network 10.25.2.0 0.255.255.255
[SwitchA-ospf-1-area-0.0.0.0] network 10.1.1.1 0.0.0.0
[SwitchA-ospf-1-area-0.0.0.0] network 10.25.1.1 0.0.0.0
[SwitchA-ospf-1-area-0.0.0.0] quit
[SwitchA-ospf-1] quit
```

Configure BGP.

```
[SwitchA] bgp 100
[SwitchA-bgp] undo synchronization
[SwitchA-bgp] group in internal
[SwitchA-bgp] peer 10.26.1.2 group in
```

```
[SwitchA-bgp] peer 10.27.1.2 group in
[SwitchA-bgp] peer in connect-interface loopback0
[SwitchA-bgp] ipv4-family multicast
[SwitchA-bgp-af-mul] peer in enable
[SwitchA-bgp-af-mul] peer 10.26.1.2 group in
[SwitchA-bgp-af-mul] peer 10.27.1.2 group in
[SwitchA-bgp-af-mul] peer in next-hop-local
[SwitchA-bgp-af-mul] quit
[SwitchA-bgp] group ex external
[SwitchA-bgp] peer 10.28.1.1 group ex as-number 200
[SwitchA-bgp] peer ex next-hop-local
[SwitchA-bgp] peer ex default-route-advertise
[SwitchA-bgp] ipv4-family multicast
[SwitchA-bgp-af-mul] peer ex enable
[SwitchA-bgp-af-mul] peer 10.28.1.1 group ex
[SwitchA-bgp-af-mul] peer ex next-hop-local
[SwitchA-bgp-af-mul] quit
[SwitchA-bgp] quit
```

Configure MSDP peer, Mess Group and Originating RP.

```
[SwitchA] msdp
[SwitchA-msdp] peer 10.28.1.1 connect-interface loopback 0
[SwitchA-msdp] peer 10.26.1.2 connect-interface loopback 0
[SwitchA-msdp] peer 10.27.1.2 connect-interface loopback 0
[SwitchA-msdp] peer 10.26.1.2 mesh-group net
[SwitchA-msdp] peer 10.27.1.2 mesh-group net
[SwitchA-msdp] originating-rp loopback0
[SwitchA-msdp] quit
```

Configuring C-RP and BSR.

```
[SwitchA] pim
[SwitchA-pim] c-rp loopback 10
[SwitchA-pim] c-bsr loopback 0 30
```

2) Configure Switch E:

Configuring VLAN

```
<SwitchE> system-view
[SwitchE] vlan 10
[SwitchE-vlan10] port ethernet1/1/2
[SwitchE-vlan10] quit
[SwitchE] vlan 20
[SwitchE-vlan20] port ethernet1/1/3
[SwitchE-vlan20] quit
```

Enable multicast.


```
[SwitchE] multicast routing-enable
```

Configure the IP address of interface loopback0 and enable PIM-SM.

```
[SwitchE] interface loopback0
[SwitchE-LoopBack0] ip address 10.26.1.2 255.255.255.255
[SwitchE-LoopBack0] pim sm
[SwitchE-LoopBack0] quit
```

Configure the IP address of interface lookback10 and enable PIM-SM.

```
[SwitchE] interface loopback10
[SwitchE-LoopBack10] ip address 10.1.1.1 255.255.255.255
[SwitchE-LoopBack10] pim sm
[SwitchE-LoopBack10] quit
```

Configure the IP address of Vlan-interface10 and enable IGMP and PIM-SM.

```
[SwitchE] interface Vlan-interface10
[SwitchE-Vlan-interface10] ip address 10.26.2.3 255.255.255.0
[SwitchE-Vlan-interface10] igmp enable
[SwitchE-Vlan-interface10] pim sm
[SwitchE-Vlan-interface10] undo shutdown
[SwitchE-Vlan-interface10] quit
```

Configure the IP address of Vlan-interface20 and enable IGMP and PIM-SM.

```
[SwitchE] interface Vlan-interface20
[SwitchE-Vlan-interface20] ip address 10.26.3.1 255.255.255.0
[SwitchE-Vlan-interface20] igmp enable
[SwitchE-Vlan-interface20] pim sm
[SwitchE-Vlan-interface20] undo shutdown
[SwitchE-Vlan-interface20] quit
```

Configuring OSPF

```
[SwitchE] ospf
[SwitchE-ospf-1] area 0
[SwitchE-ospf-1-area-0.0.0.0] network 10.26.2.0 0.255.255.255
[SwitchE-ospf-1-area-0.0.0.0] network 10.1.1.1 0.0.0.0
[SwitchE-ospf-1-area-0.0.0.0] network 10.26.1.2 0.0.0.0
[SwitchE-ospf-1-area-0.0.0.0] quit
[SwitchE-ospf-1] quit
```

Configure BGP.

```
[SwitchE] bgp 100
[SwitchE-bgp] undo synchronization
[SwitchE-bgp] group in internal
[SwitchE-bgp] peer 10.25.1.1 group in
[SwitchE-bgp] peer 10.27.1.2 group in
[SwitchE-bgp] peer in connect-interface loopback0
```

```
[SwitchE-bgp] ipv4-family multicast
[SwitchE-bgp-af-mul] peer in enable
[SwitchE-bgp-af-mul] peer 10.25.1.1 group in
[SwitchE-bgp-af-mul] peer 10.27.1.2 group in
[SwitchE-bgp-af-mul] peer in next-hop-local
[SwitchE-bgp-af-mul] quit
[SwitchE-bgp] group ex external
[SwitchE-bgp] peer 10.29.1.1 group ex as-number 300
[SwitchE-bgp] peer ex default-route-advertise
[SwitchE-bgp] peer ex ebgp-max-hop 255
[SwitchE-bgp] ipv4-family multicast
[SwitchE-bgp-af-mul] peer ex enable
[SwitchE-bgp-af-mul] peer 10.29.1.1 group ex
[SwitchE-bgp-af-mul] peer ex next-hop-local
[SwitchE-bgp-af-mul] quit
[SwitchE-bgp] quit
```

Configure MSDP peer, Mess Group and Originating RP.

```
[SwitchE] msdp
[SwitchE-msdp] peer 10.29.1.1 connect-interface loopback 0
[SwitchE-msdp] static-rpf-peer 10.29.1.1
[SwitchE-msdp] peer 10.25.1.1 connect-interface loopback 0
[SwitchE-msdp] peer 10.27.1.2 connect-interface loopback 0
[SwitchE-msdp] peer 10.25.1.1 mesh-group net
[SwitchE-msdp] peer 10.27.1.2 mesh-group net
[SwitchE-msdp] originating-rp loopback0
[SwitchE-msdp] quit
[SwitchE] ip route-static 10.29.1.1 255.255.255.0 Vlan-interface20
```

Configure C-RP and BSR.

```
[SwitchE] pim
[SwitchE-pim] c-rp loopback 10
[SwitchE-pim] c-bsr loopback 0 30
```

Note:

The configuration on the switches other than SwitchA and SwitchE is omitted here.

Note:

MBGP is not supported in the basic code. The extended option is required for MBGP.

Chapter 31 MBGP Multicast Extension Configuration

31.1 MBGP Multicast Extension Overview

31.1.1 Introduction

At present, the most widely used inter-domain unicast routing protocol is BGP-4. Because the multicast topology may be different from the unicast topology, BGP-4 must be modified in order to implement the transmission of inter-domain multicast routing information. Some routers in the network may only support unicast rather than multicast and may not forward multicast packets since the particular policy requires that. To construct inter-domain multicast routing trees, you need to know the unicast routing information as well as the information of multicast-supporting parts of the network, namely, the multicast network topology.

BGP-4 has been proved to be an effective and stable inter-domain unicast routing protocol. Therefore, it is more rational to enhance and extend the BGP-4 protocol than to construct a new protocol. RFC2858 provisions the multi-protocol extension method for BGP. The extended BGP (MBGP, also written as BGP-4+) can not only carry IPv4 unicast routing information but also the routing information of other network layer protocols (such as multicast, IPv6). Carrying multicast routing information is only one of the extended functions. This chapter describes mainly MBGP extension for multicast.

MBGP enables unicast and multicast routing information to be exchanged through the same process but stored in different routing tables. As MBGP is an enhanced version of BGP-4, all the common policies and configuration methods that BGP-4 supports can be applied to multicast.

31.1.2 MBGP Extension Attributes for Multicast

To make MBGP support multicast, RFC2858 defines two new route attributes in the UPDATE message: MP_REACH_NLRI (multiprotocol reachable NLRI) and MP_UNREACH_NLRI (multiprotocol unreachable NLRI). They are all optional non-transitive attributes, that is, routers that do not support MBGP can ignore the information in the attributes and not forward the attributes.

Among the information carried by MP_REACH_NLRI and MP_UNREACH_NLRI, AFI (Address Family Identifier) and SAFI (Subsequent Address Family Identifier) can identify for which address family the information is. SAFI is a complement to NLRI (Network Layer Reachability Information), with value 1 for the unicast mode of NLRI, and value 2 for the multicast mode of NLRI.

I. MP_REACH_NLRI attribute

MP_REACH_NLRI is an optional non-transitive attribute, and can be used to:

- Send the routing information of a new reachable protocol.
- Send the next hop information about the new protocol with the same coding mode as that of NLRI.
- Enable the router to report part or all of the SNPAs (Sub-network Points of Attachment) saved in the local system.

II. MP_UNREACH_NLRI attribute

The MP_UNREACH_NLRI is an optional non-transitive attribute that can be used for the purpose of withdrawing one or multiple unfeasible routes from service. It includes the following fields:

- AFI and SAFI.
- Withdrawn Routes: Contains one or multiple NLRIs, in which are the unreachable destination addresses.

An UPDATE packet that contains the MP_UNREACH_NLRI is not required to carry any other path attributes.

These two attributes enables MBGP to carry multi-protocol information. MSBP therefore supports both unicast and multicast by constructing different topology maps to implement appropriate policies. Besides, MBGP may construct different inter-domain routes for unicast and multicast under a same policy.

31.1.3 MBGP Operating Mode and Message Type

MBGP runs on a router in the following two modes:

- IBGP (Internal BGP)
- EBGP (External BGP)

MBGP running in an autonomous system is called IBGP; MBGP running across autonomous systems is called EBGP.

MBGP offers four types of messages:

- Open Message
- Update Message
- Notification Message
- Keepalive Message

Open Message is the first message sent after the TCP connection is established. It is used to establish MBGP peer relationship. Notification Message is used to notify errors. Keepalive message is used to check the validity of a connection. Update Message is the most important information in the MBGP system, used to exchange routing information among peers. It consists of three parts at the most: MP_UNREACH_NLRI, Path Attributes and MP_REACH_NLRI.

31.2 MBGP Multicast Extension Configuration

Basic configuration tasks of MBGP multicast extension include:

- Enable MBGP multicast extension protocol
- Specify the network routes notified by the MBGP multicast extension

Advanced configuration tasks of MBGP multicast extension include:

- Configure the MED value for an AS
- Compare MED values from different AS neighbor paths
- Configure local preference
- Configure MBGP timer
- Configure MBGP Peer (group)
- Configure MBGP route aggregation
- Configure an MBGP route reflector
- Configure the MBGP community attributes
- Configure the interaction between MBGP and IGP
- Define AS path list and routing policy
- Configure MBGP route filtering
- Reset BGP connections

Note:

Only configuration tasks in IPv4 multicast sub-address family view are detailed below. Other tasks configured in BGP or system view are only briefed. For the detailed configuration, refer to the BGP Configuration and IP Routing policy sections in Routing Protocol of this manual.

31.2.1 Enabling MBGP Multicast Extension Protocol

To enable the MBGP multicast extension protocol, enter the IPv4 multicast sub-address family view.

A router does not start receiving MBGP connection requests instantly after the MBGP multicast extension protocol is enabled. To activate a router to originate MBGP connection requests to neighboring routers, refer to the **neighbor** configuration. Perform the following configuration in BGP view.

Table 31-1 Enabling MBGP multicast extension protocol

Operation	Command
Enter the MBGP multicast address family view	ipv4-family multicast
Remove the MBGP multicast address family view	undo ipv4-family multicast

By default, the system does not run the MBGP multicast extension protocol.

31.2.2 Specifying Network Routes Notified by MBGP Multicast Extension

The **network** command is used to specify the network routes to be advertised to MBGP peers, as well as the mask and route policy of this network route.

Perform the following configurations in IPV4 multicast sub-address family view.

Table 31-2 Specifying network routes notified by MBGP multicast extension

Operation	Command
Configure the network routes to be advertised by the local MBGP	network <i>ip-address</i> [<i>address-mask</i>] [route-policy <i>route-policy-name</i>]
Remove the network routes to be advertised by the local MBGP	undo network <i>ip-address</i> [<i>address-mask</i>] [route-policy <i>route-policy-name</i>]

By default, no route is advertised by the local MBGP.

The **network** command advertises only the precisely matched route, the one with prefix and mask completely conforming to the configuration. If no mask is specified, match goes by the natural network segment.

31.2.3 Configuring the MED Value for an AS

The MED configured in BGP view is valid for both unicast and multicast.

For the details of this configuration, refer to “BGP Configuration” of the Routing Protocol part of this manual.

31.2.4 Comparing MED Values from Different AS Neighbor Paths

Do not use this configuration unless you are sure that different ASs adopt the same IGP and route selection method. The configuration in BGP view works both in unicast and multicast.

For the details of this configuration, refer to “BGP Configuration” of the Routing Protocol part of this manual.

31.2.5 Configuring Local Preference

Different local preference can be configured as a reference of the MBGP route selection. When an MBGP router gets routes with the same destination but different next hops through different neighbors, it will choose the route with the highest local preference.

The configuration works both in unicast and multicast.

For the details of this configuration, refer to “BGP Configuration” of the Routing Protocol part of this manual.

31.2.6 Configuring MBGP Timer

After a router establishes MBGP connection with a peer, it sends Keepalive messages to the peer periodically to check for the smooth connection. If the router does not receive a single Keepalive message or any other kind of message from the peer within the defined connection Holdtime, it will think the MBGP connection broken and exit, and will process the routing information received through this connection as appropriate. Therefore, the Keepalive message sending interval and MBGP connection Holdtime are two parameters of great importance in MBGP mechanism.

The configuration works both in unicast and multicast.

For the details of this configuration, refer to “BGP Configuration” of the Routing Protocol part of this manual.

31.2.7 Configuring MBGP Peer (Group)

The use of MBGP peer groups is to simplify configuration. When configuring MBGP peers, you can create and configure a peer group in BGP view, and then add the peers into the group, since all peers in a group have the same configuration with the group. Then, enable this peer group in IPv4 multicast sub-address family view and add peers to this peer group to create MBGP peers and an MBGP peer group. In conclusion, to create MBGP peers/peer groups, you must configure them successfully in BGP view first.



Caution:

Configure the peer group under the guide of technical support engineers.

I. Creating a peer group with members

To configure a MBGP peer (group), configure a peer group in BGP view and add peers to this peer group. For details, refer to “BGP Configuration” in the Routing Protocol part.

II. Enabling a peer (group)

Please perform the following configurations in IPV4 multicast sub-address family view.

Table 31-3 Enabling a peer (group)

Operation	Command
Enable the specified peer (group)	peer <i>group-name</i> enable
Disable the specified peer (group)	undo peer <i>group-name</i> enable

III. Adding an MBGP peer to the group

Please perform the following configurations in IPV4 multicast sub-address family view.

Table 31-4 Adding an MBGP peer to the group

Operation	Command
Add an MBGP peer to the group	peer <i>peer-address</i> group <i>group-name</i>
Delete the MBGP peer	undo peer <i>peer-address</i>

IV. Advertising MBGP community attributes to a peer (group)

Please perform the following configurations in IPV4 multicast sub-address family view.

Table 31-5 Configuring to advertise the community attributes to a peer (group)

Operation	Command
Advertise the community attributes to a peer (group)	peer <i>group-name</i> advertise-community
Configure not to advertise the community attributes to a peer (group)	undo peer <i>group-name</i> advertise-community

By default, no community attribute is advertised to any peer (group).

V. Configuring a peer (group) as an MBGP route reflector client

Please perform the following configurations in IPV4 multicast sub-address family view.

Table 31-6 Configuring a peer (group) as an MBGP route reflector client

Operation	Command
Configure a peer (group) as an MBGP route reflector client	peer <i>group-name</i> reflect-client
Remove the above configuration	undo peer <i>group-name</i> reflect-client

By default, there is no route reflector in an AS.

It is generally unnecessary to configure this command for a peer group. This command is reserved for the occasional compatibility with the network equipments of other vendors.

VI. Configuring the local address as the next hop when advertising routes

This involves removing the next hop configuration in the routing information advertised to a peer (group) and configuring the local address as the next hop address. It is valid only for IBGP peers/peer groups.

Please perform the following configurations in IPV4 multicast sub-address family view.

Table 31-7 Configuring the local address as the next hop when advertising routes

Operation	Command
Configure the local address as the next hop when advertising routing information	peer <i>group-name</i> next-hop-local
Remove the above configuration	undo peer <i>group-name</i> next-hop-local

VII. Specifying the routing policy for a peer (group)

Please perform the following configurations in IPV4 multicast sub-address family view.

Table 31-8 Specifying the routing policy for a peer (group)

Operation	Command
Configure routing policy for incoming packets	peer { <i>group-name</i> <i>peer-address</i> } route-policy <i>policy-name</i> import
Remove incoming policy configuration	undo peer { <i>group-name</i> <i>peer-address</i> } route-policy <i>policy-name</i> import
Configure routing policy for outgoing packets	peer <i>group-name</i> route-policy <i>policy-name</i> export
Remove outgoing policy configuration	undo peer <i>group-name</i> route-policy <i>policy-name</i> export

By default, no routing policy is specified for any peer (group).

VIII. Configuring IP-ACL-based route filtering policy for a peer (group)

Please perform the following configurations in IPV4 multicast sub-address family view.

Table 31-9 Configuring IP-ACL-based route filtering policy for a peer (group)

Operation	Command
Configure filtering policy for incoming packets	peer { <i>group-name</i> <i>peer-address</i> } filter-policy <i>acl-number</i> import
Remove incoming policy configuration	undo peer { <i>group-name</i> <i>peer-address</i> } filter-policy <i>acl-number</i> import
Configure routing policy for outgoing packets	peer <i>group-name</i> filter-policy <i>acl-number</i> export
Remove outgoing policy configuration	undo peer <i>group-name</i> filter-policy <i>acl-number</i> export

By default, a peer (group) does not perform route filtering based on the IP ACL.

IX. Configuring AS-path-list-based route filtering policy for a peer (group)

Please perform the following configurations in IPV4 multicast sub-address family view.

Table 31-10 Configuring the AS-path-list-based route filtering policy for a peer (group)

Operation	Command
Configure filtering policy for incoming packets	peer { <i>group-name</i> <i>peer-address</i> } as-path-acl <i>acl-number</i> import
Remove incoming policy configuration	undo peer { <i>group-name</i> <i>peer-address</i> } as-path-acl <i>acl-number</i> import
Configure routing policy for outgoing packets	peer <i>group-name</i> as-path-acl <i>acl-number</i> export
Remove outgoing policy configuration	undo peer <i>group-name</i> as-path-acl <i>acl-number</i> export

By default, a peer (group) does not perform route filtering based on the AS path list.

X. Configuring prefix-list-based route filtering policy for a peer (group)

Please perform the following configurations in IPV4 multicast sub-address family view.

Table 31-11 Configuring prefix-list-based route filtering policy for a peer (group)

Operation	Command
Configure filtering policy for incoming packets	peer { <i>group-name</i> <i>peer-address</i> } ip-prefix <i>prefixname</i> import
Remove incoming policy configuration	undo peer { <i>group-name</i> <i>peer-address</i> } ip-prefix <i>prefixname</i> import
Configure routing policy for outgoing packets	peer <i>group-name</i> ip-prefix <i>prefixname</i> export

Operation	Command
Remove outgoing policy configuration	undo peer <i>group-name</i> ip-prefix <i>prefixname</i> export

By default, a peer (group) does not perform route filtering based on the prefix list.

31.2.8 Configuring MBGP Route Aggregation

MBGP supports the manual aggregation of routes. Manual aggregation aggregates the local MBGP routes. A series of parameters can be configured during manual route aggregation.

Please perform the following configurations in IPV4 multicast sub-address family view.

Table 31-12 Configuring MBGP route aggregation

Operation	Command
Configure the aggregation of local routes	aggregate <i>address mask</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>]*
Remove the aggregation of local routes	undo aggregate <i>address mask</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>]*

By default, MBGP does not aggregate local routes.

31.2.9 Configuring an MBGP Route Reflector

To ensure the interconnectivity among MBGP peers, it is necessary to establish fully-closed network among IBGP multicast peers. However, some internal MBGP multicast networks are very large, and it costs a good sum to establish a fully-closed network. Route reflector solves this problem. The core is to specify a router as the focus of the internal sessions. Multiple MBGP multicast routers can be peers of one central point, namely a multiple route reflector, which in turn creates peer relationship with other reflectors. The route reflector is the focus of other routers. The routers other than those reflectors are called clients. The clients are in peer with route reflects and exchange routing information with them. The route reflectors transfer (reflect) information between the clients in turn.

For the details of the principles and configurations, refer to “BGP Configuration” of the Routing Protocol part.

31.2.10 Configure MBGP Community Attributes

Within the MBGP, a community is a set of destinations with some characteristics in common. A community is not limited to a network or an AS has no physical boundary.

For details, refer to “BGP Configuration” in the Routing Protocol part.

31.2.11 Importing IGP Routing Information into MBGP

MBGP can advertise intra-area network information to other ASs. To this end, you can use MBGP to advertise the intra-area network information that local router gets through IGP routing protocol.

Please perform the following configurations in IPV4 multicast sub-address family view.

Table 31-13 Importing IGP routing information

Operation	Command
Import IGP Routing Information into MBGP	import-route <i>protocol</i> [route-policy <i>policy-name</i>] [med <i>med-value</i>]
Delete the imported IGP routing information	undo import-route <i>protocol</i>

By default, MBGP does not import any route of other protocols.

Parameter *Protocol* specifies the source routing protocols of import, which can be direct, static, rip, isis, ospf, ospf-ase or ospf-nssa at present.

31.2.12 Defining AS Path List and Routing Policy

To configure AS path list and routing policy you need to:

- Configure the regular expression of autonomous systems (in system view);

The UPDATE information of MBGP contains an AS_PATH domain. The autonomous system paths for MBGP routing information exchange is recorded in this domain.

- Define the routing policy (in system view);
- Define matching rules (in routing policy view);
- Define value assigning rules (in routing policy view)

For the detailed configuration of regular expression of AS, refer to “BGP Configuration” (section 6.2.4) of the Routing Protocol part of this manual. For the other configuration, refer to the “IP Routing Policy Configuration” of the Routing Protocol part of this manual.

31.2.13 Configuring MBGP Route Filtering

The route filtering configuration of MBGP is the same as that of unicast BGP.

For details, refer to “BGP Configuration” of the Routing Protocol part of this manual.

31.2.14 Resetting BGP Connections

After changing the MBGP policy or protocol configuration, users must disconnect the present BGP connection to make the new configuration effective.

For details, refer to “BGP Configuration” of the Routing Protocol part of this manual.

31.3 Displaying and Debugging MBGP Configuration

After the above configuration, execute **display** commands in any view to display the running information of MBGP, and to verify the effect of the configuration.

Execute the **debugging** command in user view for the debugging of MBGP.

Table 31-14 Displaying and debugging MBGP configuration

Operation	Command
Display an MBGP routing table	display bgp multicast routing-table [<i>ip-address</i> [<i>mask</i>]]
Display CIDR (classless inter-domain routing)	display bgp multicast routing-table cidr
Display the routing information about the specified MBGP community	display bgp multicast routing-table community [<i>aa.nn</i> / no-export-subconfed no-advertise no-export]* [whole-match]
Display the routes permitted by the specified MBGP community list	display bgp multicast routing-table community-list <i>community-list-number</i> [whole-match]
Display the routes with inconsistent source autonomous systems	display bgp multicast routing-table different-origin-as
Display the routing information to or from a specified multicast neighbor	display bgp multicast peer [<i>peer-address</i>] [verbose]
Display the routing information advertised by MBGP	display bgp multicast network
Display the peer group information	display bgp multicast group [<i>group-name</i>]
Display the AS path information matching the AS regular expression	display bgp multicast routing-table regular-expression <i>as-regular-expression</i>
Disable/enable debugging MBGP UPDATE packets	[undo] debugging bgp mp-update [receive send] [verbose]

31.4 MBGP Multicast Extension Configuration Example

I. Networking requirement

This example describes how the administrator uses the MBGP attributes to manage route selection. All switches are configured with MBGP. The IGP in AS200 uses OSPF. Switch A is AS100 and serves as the MBGP neighbor of Switch B and Switch C in AS200. Switch B and Switch C run IBGP for Switch D in AS200. Switch D is also in AS200.

II. Networking diagram

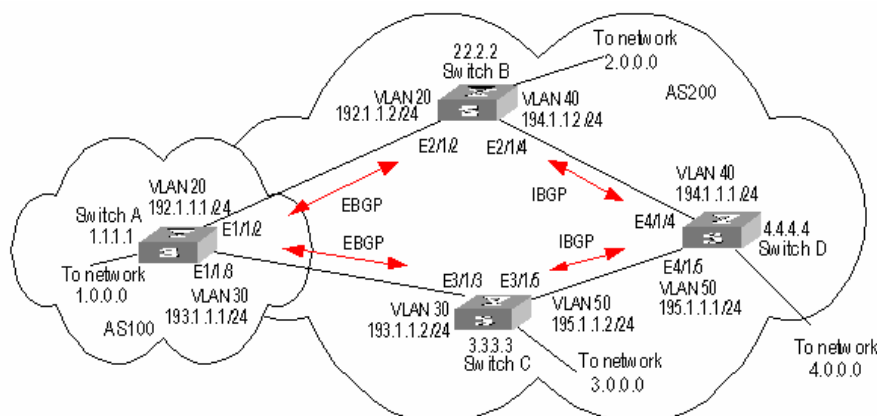


Figure 31-1 Networking diagram for MBGP path selection configuration

III. Configuration procedure

1) Configure Switch A:

```
[SwitchA] vlan 20
[SwitchA-vlan20] port ethernet1/1/2
[SwitchA-vlan20] quit
[SwitchA] interface vlan-interface 20
[SwitchA-Vlan-interface20] ip address 192.1.1.1 255.255.255.0
[SwitchA-Vlan-interface20] quit
[SwitchA] vlan 30
[SwitchA-vlan30] port ethernet1/1/3
[SwitchA-vlan30] quit
[SwitchA] interface vlan-interface 30
[SwitchA-Vlan-interface30] ip address 193.1.1.1 255.255.255.0
[SwitchA-Vlan-interface30] quit
```

Enable MBGP.

```
[SwitchA] bgp 100
[SwitchA-bgp] ipv4-family multicast
```

Specify target network for MBGP.

```
[SwitchA-bgp-af-mul] network 1.0.0.0
[SwitchA-bgp-af-mul] network 2.0.0.0
[SwitchA-bgp-af-mul] quit
```

Configure peers relationship.

```
[SwitchA-bgp] bgp 100
[SwitchA-bgp] group a1 external
[SwitchA-bgp] peer 192.1.1.2 group a1 as-number 200
[SwitchA-bgp] group a2 external
[SwitchA-bgp] peer 193.1.1.2 group a2 as-number 200
[SwitchA-bgp] ipv4-family multicast
[SwitchA-bgp-af-mul] peer a1 enable
[SwitchA-bgp-af-mul] peer a2 enable
```

Configure the MED attribute of Switch A.

- Add an ACL on Switch A to permit network 1.0.0.0.

```
[SwitchA] acl number 2000
[SwitchA-acl-basic-2000] rule permit source 1.0.0.0 0.255.255.255
[SwitchA-acl-basic-2000] rule deny source any
```

- Define two routing policies: `set_med_50` and `set_med_100`, providing two MED values for network 1.0.0.0 (50 and 100 respectively).

```
[SwitchA] route-policy set_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 set_med_100 permit node 10
[SwitchA-route-policy] if-match acl 2000
[SwitchA-route-policy] apply cost 100
```

- Apply the routing policy `set_med_50` to the exported route updates of Switch C (193.1.1.2). Apply the routing policy `set_med_100` to the exported route updates of Switch B (192.1.1.2).

```
[SwitchA] bgp 100
[SwitchA-bgp] ipv4-family multicast
[SwitchA-bgp-af-mul] peer a2 route-policy set_med_50 export
[SwitchA-bgp-af-mul] peer a1 route-policy set_med_100 export
```

2) Configure Switch B:

```
[SwitchB] vlan 20
[SwitchB-vlan20] port ethernet2/1/2
[SwitchB-vlan20] quit
[SwitchB] interface vlan-interface 20
[SwitchB-Vlan-interface20] ip address 192.1.1.2 255.255.255.0
[SwitchB-Vlan-interface20] quit
[SwitchB] vlan 40
[SwitchB-vlan40] port ethernet2/1/4
```

```
[SwitchB-vlan40] quit
[SwitchB] interface vlan-interface 40
[SwitchB-Vlan-interface40] ip address 194.1.1.2 255.255.255.0
[SwitchB-Vlan-interface40] quit
[SwitchB] ospf
[SwitchB-ospf-1] area 0
[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] network 192.1.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
[SwitchB-ospf-1] quit
[SwitchB] bgp 200
[SwitchB-bgp] undo synchronization
[SwitchB-bgp] group b1 external
[SwitchB-bgp] peer 192.1.1.1.1 group b1 as-number 100
[SwitchB-bgp] group b2 internal
[SwitchB-bgp] peer 194.1.1.1.1 group b2
[SwitchB-bgp] peer 195.1.1.1.2 group b2
[SwitchB-bgp] ipv4-family multicast
[SwitchB-bgp-af-mul] peer b1 enable
[SwitchB-bgp-af-mul] peer b2 enable
```

3) Configure Switch C:

```
[SwitchC] vlan 30
[SwitchC-vlan30] port ethernet3/1/3
[SwitchC-vlan30] quit
[SwitchC] interface vlan-interface 30
[SwitchC-Vlan-interface30] ip address 193.1.1.2 255.255.255.0
[SwitchC-Vlan-interface30] quit
[SwitchC] vlan 50
[SwitchC-vlan50] port ethernet3/1/5
[SwitchC-vlan50] quit
[SwitchC] interface vlan-interface 50
[SwitchC-Vlan-interface50] ip address 195.1.1.2 255.255.255.0
[SwitchC-Vlan-interface50] quit
[SwitchC] ospf
[SwitchC-ospf-1] 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
[SwitchC] bgp 200
[SwitchC-bgp] undo synchronization
[SwitchC-bgp] group c1 external
[SwitchC-bgp] peer 193.1.1.1 group c1 as-number 100
```



```
[SwitchC-bgp] group c2 internal
[SwitchC-bgp] peer 194.1.1.2 group c2
[SwitchC-bgp] peer 195.1.1.1 group c2
[SwitchC-bgp] ipv4-family multicast
[SwitchC-bgp-af-mul] peer c1 enable
[SwitchC-bgp-af-mul] peer c2 enable
```

Configure the local preference attribute of Switch C.

- Add ACL 2000 on Switch C to permit network 1.0.0.0.

```
[SwitchC] acl number 2000
[SwitchC-acl-basic-2000] rule permit source 1.0.0.0 0.255.255.255
[SwitchC-acl-basic-2000] quit
```

- Define the routing policy named "localpref". Set the local preference for the routes matching ACL 2000 to 200, and otherwise, to 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
[SwitchC] route-policy localpref permit node 20
[SwitchC-route-policy] apply local-preference 100
```

- Apply this routing policy to the inbound traffic from BGP neighbor 193.1.1.1 (Switch A).

```
[SwitchC] bgp 200
[SwitchC-bgp] ipv4-family multicast
[SwitchC-bgp-af-mul] peer 193.1.1.1 route-policy localpref import
```

4) Configure Switch D:

```
[SwitchD] vlan 40
[SwitchD-vlan40] port ethernet4/1/4
[SwitchD-vlan40] quit
[SwitchD] interface vlan-interface 40
[SwitchD-Vlan-interface40] ip address 194.1.1.1 255.255.255.0
[SwitchD-Vlan-interface40] quit
[SwitchD] vlan 50
[SwitchD-vlan50] port ethernet4/1/5
[SwitchD-vlan50] quit
[SwitchD] interface vlan-interface 50
[SwitchD-Vlan-interface50] ip address 195.1.1.1 255.255.255.0
[SwitchD-Vlan-interface50] quit
[SwitchD] ospf
[SwitchD-ospf-1] 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] network 4.0.0.0 0.0.0.255
```

```
[SwitchD-ospf-1-area-0.0.0.0] quit
[SwitchD-ospf-1] quit
[SwitchD] bgp 200
[SwitchD-bgp] undo synchronization
[SwitchD-bgp] group d1 internal
[SwitchD-bgp] peer 194.1.1.2 group d1
[SwitchD-bgp] peer 195.1.1.2 group d1
[SwitchD-bgp] ipv4-family multicast
[SwitchD-bgp-af-mul] peer d1 enable
```

To make the configuration effective, you need to use the **reset bgp all** command on all MBGP neighbors.

Chapter 32 ACL Configuration

32.1 ACL Overview

32.1.1 Introduction to ACL

A series match rules must be configured to recognize the packets before they are filtered. Only when packets are identified, can the network take corresponding actions, allowing or prohibiting them to pass, according to the preset policies. Access control list (ACL) is targeted to achieve these functions.

ACLs classify packets using a series of matching rules, which can be source addresses, destination addresses and port IDs. ACLs can be used globally on the switch or just at a port, through which the switch determines whether to forward or drop the packets.

The matching rules defined in ACLs can also be imported to differentiate traffic in other situations, for example, defining traffic classification rules in QoS.

An ACL rule can include many sub-rules, which may be defined for packets within different address ranges. Matching order is involved in matching an ACL.

I. ACLs being activated directly on hardware

ACLs can be delivered to hardware for traffic filtering and classification.

The cases when ACLs are sent directly to hardware include: referencing ACLs to provide for QoS functions, filtering and forwarding packets with ACLs.

II. ACLs being referenced by upper-level modules

ACLs may also be used to filter and classify packets processed by software. Then you can define matching order for the sub-rules in an ACL. Two matching modes are available in this case: **config** (user-defined order) and **auto** (depth first by the system). You cannot modify the matching order once you define it for an ACL rule, unless you delete the rule and redefine the matching order.

The cases when ACLs are referenced by upper-level modules include referencing ACLs to achieve routing policies, and using ACLs to control register users and so on.

Note:

Depth first principle means putting the statement with smaller packet range in the front. You can know the packet range by comparing IP address wildcards: The smaller the wildcard is, the smaller host range is. For example, the address 129.102.1.1 0.0.0.0 specifies the host 129.102.1.1 and address 129.102.1.1 0.0.255.255 specifies the segment 129.102.1.1 to 129.102.255.255. Then 129.102.1.1 is surely put in the front. Specifically, for the statements of basic ACL rules, directly compare the wildcards of source addresses and follow **config** order if the wildcards are equal; for the ACL rules used in port packet filtering, the rules configured with **any** are put to the end and other rules follow **config** order; for advanced ACL rules, first compare the wildcards of source addresses, then the wildcards of destination addresses if those of source addresses are equal, then the port IDs if the wildcards of destination addresses are still equal. Follow **config** order if port IDs are also equal.

Note:

The user-defined ACL matching order takes effect only when multiple rules of one ACL are applied at the same time. For example, an ACL has two rules. If the two rules are not applied simultaneously, even if you configure the matching order to be depth first, the switch still matches them according to their application order.

If one rule is a subset of another rule in an ACL, it is recommended to apply the rules according to the range of the specified packets. The rule with the smallest range of the specified data packets is applied first, and then other rules are applied based on this principle.

32.1.2 ACLs Supported

The switch supports these types of ACLs:

- Number-based basic ACLs
- Name-based basic ACLs
- Number-based advanced ACLs
- Name-based advanced ACLs
- Number-based L2 ACLs
- Name-based L2 ACLs
- Number-based user ACLs
- Name-based user ACLs

The requirements for the various ACLs available on the switch are listed in the following table.

Table 32-1 Requirements for defining ACLs

Item	Number range	Maximum number
Number-based basic ACL	2000 to 2999	1000
Number-based advanced ACL	3000 to 3999	1000
Number-based L2 ACL	4000 to 4999	1000
Number-based user ACL	5000 to 5999	1000
Name-based basic ACL	--	--
Name-based advanced ACL	--	--
Name-based L2 ACL	--	--
Name-based user ACL	--	--
Maximum sub-rules for an ACL	0 to 127	128
Maximum sub-rules for the system	--	12288

Table 32-2 Max ACL rules that can be activated on cards

Card	Max ACL rules per card/port
1-port 10GBASE-X (XENPAK) Advanced 2-port 10GBASE-X (XFP) Advanced 24-port 1000BASE-X (SFP) Advanced 24-port 10/100/1000BASE-T (RJ45) Advanced 48-port 10/100/1000BASE-T (RJ45) Access	1012
1-port 10GBASE-X (XENPAK) 2-port 10GBASE-X (XFP) 4-port 10GBASE-X (XFP) 12-port 1000BASE-X (SFP) 24-port 1000BASE-X (SFP) 24-port 10/100/1000BASE-T (RJ45) 48-port 10/100/1000BASE-T (RJ45)	1024 or 2048 ACL rules are based on the number of Packet Processors (PP) per blade. The 24-port blades and 4-port 10G blades both use 2 PPs, therefore the max ACL rules is 2k (2048). The rest of the blades have a single PP and 1024 max ACL rules. The system can define up to 128 rules per ACL for a maximum of $\text{Sum}((\text{number of ACLx}) \times (\text{number of rules per ACLx}))$.

32.2 ACL Configuration

The following table describes the ACL configuration tasks for interface cards.

Table 32-3 ACL configuration tasks

No.	Item	Command	Description
1	Enter the system view	<SW8800> system-view	—
2	Configure the time range	[SW8800] time-range	Optional
3	Define a flow template	[SW8800] flow-template user-defined slot slotid template-info	Optional
4	Enter the ACL view	[SW8800] acl	Required
5	Define sub-rules	[SW8800-acl-adv-3000] rule	Required
6	Exit ACL view Enter Ethernet port view	[SW8800-acl-adv-3000] quit [SW8800] interface Ethernet 5/1/1	—
7	Apply a defined flow template in the Ethernet port view	[SW8800-Ethernet5/1/1] flow-template user-defined	Optional. You can perform this operation only when a flow template has been previously defined.
8	Activate the ACL	[SW8800-Ethernet5/1/1] packet-filter inbound	Required

32.2.1 Configuring Time Range

You may set such items in time range configuration: The defined time range includes absolute time range and period time range. The absolute time range is in the form of hh:mm YYYY/MM/DD; the period time range is in the format of hh:mm, day.

Perform the following configurations in system view.

Table 32-4 Configure time range

Operation	Command
Create time range	time-range <i>time-name</i> { <i>start-time to end-time days-of-the-week</i> [from <i>start-time start-date</i>] [to <i>end-time end-date</i>] from <i>start-time start-date</i> [to <i>end-time end-date</i>] to <i>end-time end-date</i> }
Delete time range	undo time-range <i>time-name</i> [<i>start-time to end-time days-of-the-week</i> [from <i>start-time start-date</i>] [to <i>end-time end-date</i>] from <i>start-time start-date</i> [to <i>end-time end-date</i>] to <i>end-time end-date</i>]

start-time and *end-time days-of-the-week* define period time range together. *start-time start-date* and *end-time end-date* define absolute time range together.

If a time range only defines the period time range, the time range is only active within the period time range.

If a time range only defines the absolute time range, the time range is only active within the absolute time range.

If a time range defines the period time range and the absolute time range, the time range is only active when the period time range and the absolute time range are both matched. For example, a time range defines a period time range which is from 12:00 to 14:00 every Wednesday, and defines an absolute time range which is from 00:00 2004/1/1 to 23:59 2004/12/31. This time range is only active from 12:00 to 14:00 every Wednesday in 2004.

If neither starting time nor end time is specified, the time range is 24 hours (0:00 to 24:00).

If no end date is specified, the time range is from the date of configuration till the largest date available in the system.

Currently the largest time range is 1970/01/01 to 2100/12/31 in the system.

32.2.2 Defining and Applying Flow Template

I. Defining Flow Template

Flow template defines useful information used in flow classification. For example, a template defines a quadruple: source and destination IP, source and destination TCP ports, and then only those traffic rules including all these elements can be sent to target hardware and referenced for such QoS functions as packet filtering, traffic policing, priority re-labeling. Otherwise, the rules cannot be activated on the hardware and referenced.

Perform the following configurations in system view.

Table 32-5 Define flow template

Operation	Command
Define flow template	flow-template user-defined slot <i>slotid</i> <i>template-info</i>
Delete flow template	undo flow-template user-defined slot <i>slotid</i>

Note that the sum of all elements should not be more than 16 bytes in length. The following table lists the length of the elements involved.

Table 32-6 Length of template elements

Name	Description	Length in template
cos	802.1p priority	1 byte
dip	Destination IP field in IP packet header	4 bytes
dmac	Destination MAC field in Ethernet packet header	6 bytes
dport	Destination port field	2 bytes
dscp	DSCP field in IP packet header	1 byte
ip-precedence	IP precedence field in IP packet header	
tos	ToS field in IP packet header	
exp	EXP field in MPLS packet	
ethernet-protocol	Protocol field in Ethernet packet header	4 bytes
fragment-flags	Flag field of fragment in IP packed header	No bytes
icmp-code	ICMP code field	1 byte
icmp-type	ICMP type field	1 byte
ip-protocol	Protocol field in IP packet header	1 byte
sip	Source IP field in IP packet header	4 bytes
smac	MAC field in Ethernet packet header	6 bytes
sport	Source port field	2 bytes
tcp-flag	Flag field in TCP packet header	1 byte
vlanid	Vlan ID of the packet	2 bytes
c-tag-cos	802.1p priority in the Internal 802.1Q tag (internal tag of QinQ tag-in-tag application)	2 bytes
c-tag-vlanid	Vlan ID in the internal 802.1Q tag (internal tag of QinQ tag-in-tag application)	2 bytes
bt-flag	Flag for Bit Torrent peer-to-peer service	2 bytes

Note:

The numbers listed in the table are not the actual length of these elements in IP packets, but their length in flow template. DSCP field is one byte in flow template, but six bits in IP packets. You can judge the total length of template elements using these numbers. The dscp, exp, ip-precedence and tos fields jointly occupy one byte. One byte is occupied no matter you define one, two or three of these fields.

The fragment-flags field is 0 in length in flow template, so it can be ignored in calculating the total length of template elements.

You can either use the default template or define a flow template based on your needs.

Note:

Default flow template:

ip-protocol tcp-flag sport dport icmp-type icmp-code sip 0.0.0.0 dip 0.0.0.0

You cannot modify or delete the default flow template.

II. Applying Flow Template

Perform the following configurations in Ethernet port view or port group view to apply the user-defined flow template to current port or current port group.

Table 32-7 Apply flow template

Operation	Command
Apply the user-defined flow template	flow-template user-defined
Cancel the applied flow template	undo flow-template user-defined

32.2.3 Defining ACL

The switch supports several types of ACLs, which are described in this section.

Follow these steps to define an ACL

- 1) Enter the corresponding ACL view
- 2) Define ACL rules

Note:

- If the **time-range** keyword is not selected, the ACL will be effective at any time after being activated.
- You can define multiple rules for the ACL by using the **rule** command several times.
- If the ACL is sent directly to hardware for packet filtering and traffic classification, the **auto** matching order is available and the user-defined (**config**) matching order becomes ineffective. If the ACL is used in filtering or classifying the packets processed by software, the **config** matching order is available. You cannot modify the matching order once you define that for an ACL rule.
- By default, ACL rules are matched in **config** order.

I. Defining basic ACL

Basic ACLs only make rules and process packets according to the source IP addresses.

Perform the following configurations in the specified views.

Table 32-8 Define basic ACL

Operation	Command
Enter basic ACL view (system view)	acl { number <i>acl-number</i> name <i>acl-name</i> basic } [match-order { config auto }]
Define an ACL rule (basic ACL view)	rule [<i>rule-id</i>] { permit deny } [source { <i>source-addr</i> <i>wildcard</i> any } fragment time-range <i>name</i> vpn-instance <i>instance-name</i>]*
Delete an ACL rule (basic ACL view)	undo rule <i>rule-id</i> [source fragment time-range vpn-instance <i>instance-name</i>]*
Delete an ACL or all ACLs (system view)	undo acl { number <i>acl-number</i> name <i>acl-name</i> all }

II. Defining advanced ACL

Advanced ACLs define classification rules and process packets according to the attributes of the packets such as source and destination IP addresses, TCP/UDP ports used, and packet priority. ACLs support three types of priority schemes: ToS (type of service) priority, IP priority and DSCP priority.

Perform the following configurations in the specified view.

Table 32-9 Define advanced ACL

Operation	Command
Enter advanced ACL view (system view)	acl { number <i>acl-number</i> name <i>acl-name</i> advanced } [match-order { config auto }]
Define an ACL rule (advanced ACL view)	rule [<i>rule-id</i>] { permit deny } <i>protocol</i> [source { <i>source-addr wildcard</i> any }] [destination { <i>dest-addr wildcard</i> any }] [source-port <i>operator port1</i> [<i>port2</i>]] [destination-port <i>operator port1</i> [<i>port2</i>]] [icmp-type <i>type code</i>] [established] [[precedence <i>precedence</i> tos <i>tos</i>]*] [dscp <i>dscp</i>] [fragment] [time-range <i>name</i>] [vpn-instance <i>instance-name</i>]
Delete an ACL rule (advanced ACL view)	undo rule <i>rule-id</i> [source destination source-port destination-port icmp-type precedence tos dscp fragment time-range vpn-instance]*
Delete an ACL or all ACLs (system view)	undo acl { number <i>acl-number</i> name <i>acl-name</i> all }

Note that the *port1* and *port2* parameters in the command should be TCP/UDP ports for advanced applications. For some common ports, you can use mnemonic symbols to replace numbers. For example, you can use “bgp” to represent TCP port 179, which is for BGP protocol.

III. Defining L2 ACLs

L2 ACLs define the Layer 2 information such as source and destination MAC addresses, source VLAN ID, and L2 protocol type in their rules and process packets according to these attributes.

Perform the following configurations in the specified view.

Table 32-10 Define L2 ACLs

Operation	Command
Enter L2 ACL view (system view)	acl { number <i>acl-number</i> name <i>acl-name</i> link } [match-order { config auto }]
Define an ACL rule (L2 ACL view)	rule [<i>rule-id</i>] { permit deny } [cos <i>cos-value</i> { arp ip mpls [<i>l2lable-range</i>] [exp <i>exp-value</i>] nbx pppoe-control pppoe-data rarp } ingress { { <i>source-vlan-id</i> <i>source-mac-addr source-mac-wildcard</i> }* any } egress { <i>dest-mac-addr dest-mac-wildcard</i> any } time-range <i>name</i>]*
Delete an ACL rule (L2 ACL view)	undo rule <i>rule-id</i>
Delete an ACL or all ACLs (system view)	undo acl { number <i>acl-number</i> name <i>acl-name</i> all }

32.2.4 Activating ACL

After defining an ACL, you must activate it. This configuration activates those ACLs to filter or classify the packets forwarded by hardware.

For interface cards, perform the following configurations in Ethernet port view or port group view.

Table 32-11 Activate ACL

Operation	Command
Activate IP group ACL	packet-filter inbound ip-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i> [system-index <i>index</i>]]
Deactivate IP group ACL	undo packet-filter inbound ip-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i>]
Activate IP group ACL and link group ACL at same time	packet-filter inbound ip-group { <i>acl-number</i> <i>acl-name</i> } { rule <i>rule</i> link-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i> [system-index <i>index</i>]] link-group { <i>acl-number</i> <i>acl-name</i> } rule <i>rule</i> }
Deactivate IP group ACL and link group ACL at same time	undo packet-filter inbound ip-group { <i>acl-number</i> <i>acl-name</i> } { rule <i>rule</i> link-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i>] link-group { <i>acl-number</i> <i>acl-name</i> } rule <i>rule</i> }
Activate link group ACL	packet-filter inbound link-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i> [system-index <i>index</i>]]
Deactivate link group ACL	undo packet-filter inbound link-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i>]

system-index *index* here is the system index for an ACL rule. When delivering a rule, the system assigns an index to it, for convenience of later retrieval. You can also assign a system index for it when activating an ACL rule with this command. However, you are not recommended to assign a system index if not urgently necessary.

32.3 Displaying and Debugging ACL Configurations

After these configurations are completed, you can use the **display** command in any view to view ACL running to check configuration result. You can clear ACL statistics using the **display** command in user view.

Table 32-12 Display and debug ACL configurations

Operation	Command
Display time range configuration	display time-range { all <i>name</i> }

Display ACL configuration	display acl config { all <i>acl-number</i> <i>acl-name</i> }
Display ACL application information	display acl running-packet-filter { all interface { <i>interface-name</i> <i>interface-type interface-num</i> } vlan <i>vlan-id</i> }
Display configuration information of flow template	display flow-template [default interface <i>interface-type interface-num</i> slot <i>slotid</i> user-defined]
Clear ACL statistics	reset acl counter { all <i>acl-number</i> <i>acl-name</i> }

The **display acl config** command only displays the ACL matching information processed by the CPU. You can use the **display qos-interface traffic-statistic** commands to view the ACL matching information during data forwarding.

See the corresponding *Command Manual* for description of parameters.

32.4 ACL Configuration Example

32.4.1 Advanced ACL Configuration Example

I. Network requirements

The departments in the intranet are connected through 100 Mbps ports of the switches. The research and development (R&D) department is connected through the port Ethernet2/1/1. The wage server of the financial department is at 129.110.1.2. The requirement is to configure ACLs correctly to limit that the R&D department can only access the wage server at working time from 8:00 to 18:00.

II. Network diagram

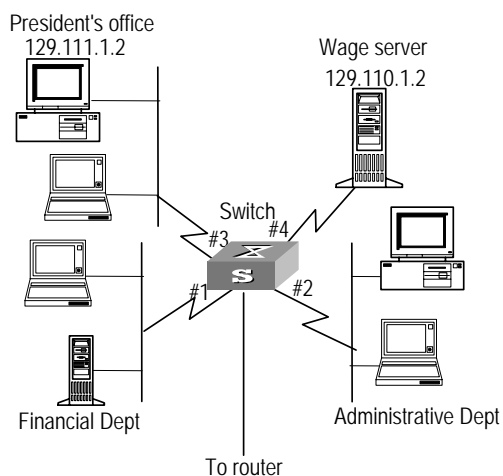


Figure 32-1 Network diagram for advanced ACL configuration

III. Configuration procedure

Note:

Only the commands concerning ACL configuration are listed here.

1) Define the time range.

Define the time range from 8:00 to 18:00.

```
[SW8800] time-range 3Com 8:00 to 18:00 working-day
```

2) Define inbound traffic to the wage server.

Create a name-based advanced ACL “traffic-of-payserver” and enter it.

```
[SW8800] acl name traffic-of-payserver advanced
```

Define ACL rule for the wage server.

```
[SW8800-acl-adv-traffic-of-payserver] rule 1 deny ip source any destination
129.110.1.2 0.0.0.0 time-range 3Com
```

3) Activate the ACL.

Activate the ACL “traffic-of-payserver”.

```
[SW8800-Ethernet2/1/1] packet-filter inbound ip-group traffic-of-payserver
```

32.4.2 Basic ACL Configuration Example

I. Network requirements

With proper basic ACL configuration, during the time range from 8:00 to 18:00 everyday the switch filters the packets from the host with source IP 10.1.1.1 (the host is connected through the port Ethernet2/1/1 to the switch.)

II. Network diagram

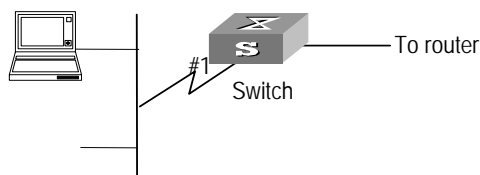


Figure 32-2 Network diagram for basic ACL configuration

III. Configuration procedure

Note:

Only the commands concerning ACL configuration are listed here.

- 1) Define the time range.

Define the time range from 8:00 to 18:00.

```
[SW8800] time-range 3Com 8:00 to 18:00 daily
```

- 2) Define the traffic with source IP 10.1.1.1.

Create a name-based basic ACL "traffic-of-host" and enter it.

```
[SW8800] acl name traffic-of-host basic
```

Define ACL rule for source IP 10.1.1.1.

```
[SW8800-acl-basic-traffic-of-host] rule 1 deny source 10.1.1.1 0 time-range 3Com
```

- 3) Activate the ACL.

Activate the ACL "traffic-of-host".

```
[SW8800-Ethernet2/1/1] packet-filter inbound ip-group traffic-of-host
```

32.4.3 L2 ACL Configuration Example

I. Network requirements

With proper L2 ACL configuration, during the time range from 8:00 to 18:00 everyday the switch filters the packets with source MAC 00e0-fc01-0101 and destination MAC 00e0-fc01-0303 (configuring at the port Ethernet2/1/1 to the switch.)

II. Network diagram

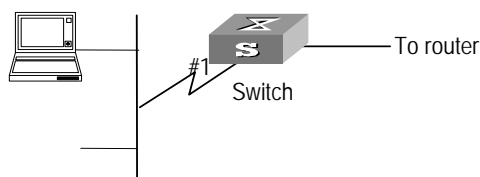


Figure 32-3 Network diagram for L2 ACL configuration

III. Configuration procedure

Note:

Only the commands concerning ACL configuration are listed here.

- 1) Define the time range.

Define the time range from 8:00 to 18:00.

```
[SW8800] time-range 3Com 8:00 to 18:00 daily
```

- 2) Define a user-defined flow template

```
[SW8800] flow-template user-defined slot 3 ethernet-protocol smac 0-0-0 dmac  
0-0-0
```

- 3) Define the traffic with source MAC 00e0-fc01-0101 and destination MAC 00e0-fc01-0303.

Create a name-based L2 ACL “traffic-of-link” and enter it.

```
[SW8800] acl name traffic-of-link link
```

Define ACL rule for the traffic with source MAC 00e0-fc01-0101 and destination MAC 00e0-fc01-0303.

```
[SW8800-acl-link-traffic-of-link] rule 1 deny ingress 00e0-fc01-0101 0-0-0  
egress 00e0-fc01-0303 0-0-0 time-range 3Com
```

```
[SW8800-acl-link-traffic-of-link] quit
```

- 4) Apply the user-defined flow template to the port and activate the ACL.

Apply the user-defined flow template to Ethernet2/1/1.

```
[SW8800] interface Ethernet2/1/1
```

```
[SW8800-Ethernet2/1/1] flow-template user-defined
```

Activate the ACL “traffic-of-link”.

```
[SW8800-Ethernet2/1/1] packet-filter inbound link-group traffic-of-link
```


Chapter 33 QoS Configuration

33.1 QoS Overview

Conventional packet network treats all packets equally. Each switch/router processes all packets in First-in-First-out (FIFO) mode and then transfers them to the destination in the best effort, but it provides no commitment and guarantee to such transmission performance as delay and jitter.

With fast growth of computer networks, more and more data like voice and video that are sensitive to bandwidth, delay and jitter are transmitted over the network. This makes growing demands on quality of service (QoS) of networks.

Ethernet technology is a widely-used network technology dominant for independent LANs and many LANs based on Ethernet are organic parts of the Internet. In addition, Ethernet access is becoming one of the major access modes for Internet users. Therefore it is inevitable to consider Ethernet QoS if we want to achieve point-to-point global QoS solution. Ethernet switching devices then naturally need to provide different QoS guarantee for different types of services, especially for those which are sensitive to delay and jitter.

The following terms are involved in QoS.

I. Flow

It refers to all packets passing through the switch.

II. Traffic classification

Traffic classification is the technology that identifies the packets with a specified attribute according to a specific rule. Classification rule refers to a packet filtering rule configured by an administrator. A classification rule can be very simple. For example, the switch can identify the packets of different priority levels according to the ToS (type of service) field in the packet headers. It can also be very complex. For example, it may contain information of the link layer (layer 2), network layer (layer 3) and transport layer (layer 4) and the switch classifies packets according to such information as MAC address, IP protocol, source address, destination address and port ID. Classification rule often is limited to the information encapsulated at the packet header, rarely using packet contents.

III. Packet filtering

Packet filtering refers to filtering operation applied to traffic flow. For example, the deny operation drops the traffic flow which matches the classification rule and allows other traffic to pass. Ethernet switches use complex classification rules, so that traffic flow can be filtered purposefully to enhance network security.

There are two key steps in packet filtering:

Step 1: Classify the traffic at the port according to a specific rule.

Step 2: Run filtering operation (deny or permit) to the identified traffic. By default, deny operation is selected.

IV. Traffic policing

QoS can police traffic at the ingress port, to provide better services with the limited network resources.

V. Redirection

You can re-specify forwarding port for packets, based on QoS policy.

VI. Traffic priority

Ethernet switches can provide priority tags, including ToS, DSCP, 802.1p, and so on, for specific packets. These priority tags are applicable to different QoS models.

The following describes IP priority, ToS priority, DCSP priority, Exp priority and 802.1p priority.

- 1) IP priority, ToS priority, DSCP priority and Exp priority

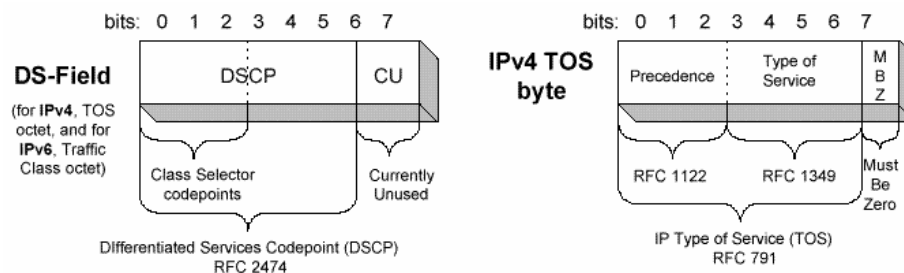


Figure 33-1 DS field and ToS byte

As shown in Figure 33-1, the ToS field in the IP header contains 8 bits. The first three bits represent IP priority, in the range of 0 to 7; bits 3-6 stand for ToS priority, in the range of 0 to 15. RFC2474 redefines the ToS field in IP packets as DS (differentiated services) field. The first six bits denote DSCP (differentiated services codepoint) priority, in the range of 0 to 63, the latter two bits are reserved. The first three bits (bit 0-2) of DSCP priority represent Exp priority, in the range of 0 to 7.

- 2) 802.1p priority

802.1p priority is stored in the header of L2 packets and is suitable for the case where only L2 QoS guarantee, not L3 header analysis, is required.

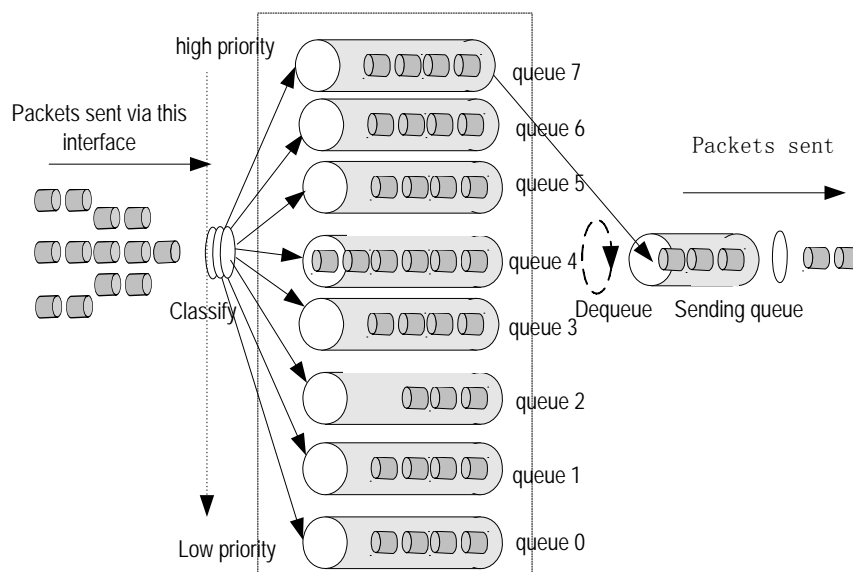


Figure 33-4 Priority queues

SP algorithm is designed for key services. One of the characteristics of key services is these services should be processed first to minimize response delay during switch congestion. For example, there are eight outbound queues at the port, numbered respectively as 7 to 0, with priority levels in descending order.

In SP mode, the system first sends those packets of higher priority in strict accordance with priority order. Only when packets in high priority queue are all sent can those in lower priority queue be sent. This manner of putting key-service packets into high priority queue and non-key service packets into low priority queue does ensure that key-service packets are sent first, while non-key service packets are sent during the interval when no key-service packets needs to be processed.

SP algorithm also has its disadvantages: If high priority queues always have packets for a long period, then the packets in low queues may die of hunger for being processed.

2) WRR algorithm

Each port supports eight outbound queues except that port of XP4 board only supports four queues. In WRR mode, the system processes the queues by turn, so every queue can have a service period.

See the case where the port supports eight outbound queues. Every queue is assigned with a weight value (respectively numbered as w_7 , w_6 , w_5 , w_4 , w_3 , w_2 , w_1 and w_0), which indicates the weight in obtaining resources. For a 100 Mbps port, the weight values are set as 50, 30, 10, 10, 50, 30, 10 and 10 (corresponding respectively to w_7 , w_6 , w_5 , w_4 , w_3 , w_2 , w_1 and w_0). The even the queue with the lowest priority can be allocated with a 5 Mbps bandwidth.

Another merit for WRR algorithm: Though the queues are scheduled by turn, they are not configured with fixed time quantum. If a queue has no packets, the system immediately schedules the next queue. Then bandwidth resources can be fully utilized.

VIII. Traffic mirroring

Traffic mirroring duplicates specified packets to CPU for network test and troubleshooting.

IX. Port mirroring

Port mirroring duplicates all packets at a specified port to the monitoring port for network test and troubleshooting.

X. Flow-based traffic statistics

The system can make traffic statistics based on flow for further analysis.

33.2 Introduction to Port Group-Based QoS Configuration

To configure QACL for a port group on the Switch 8800, you only need to create a port group and configure QACL for the group. Then the configuration becomes valid for all members in the group. This group-based QACL configuration saves you from configuring QACL for individual ports. After this configuration, the QACL configuration of each member port remains consistent forever.

33.2.1 Group-Based QoS Configuration Task

The following table describes the group-based QoS configuration tasks. (Suppose the flow template and ACL are defined already.)

Table 33-1 Group-based QoS Configuration tasks

Item	Command	Description
Enter system view	system-view	—
Enter port group view	port-group <i>index</i>	Required. <i>index</i> : group number. For a common interface board, it ranges from 1 to 128.
Add ports to the port group	port <i>interface_list</i>	Required. <i>interface_list</i> = { <i>interface_type</i> <i>interface_num</i> <i>interface_name</i> } [to { <i>interface_type</i> <i>interface_num</i> <i>interface_name</i> }] &<1-n>.

Item	Command	Description
Apply flow template	flow-template user-defined	Optional. Refer to section 32.2.2 "Defining and Applying Flow Template".
Activate ACL	packet-filter inbound	Optional. Refer to section 32.2.4 "Activating ACL".
Configure local precedence for port	priority priority-level	Optional. Refer to section 33.3.1 "Configuring Service Parameter Allocation Rule".
Configure traffic policing	traffic-limit inbound	Optional. Refer to section 33.3.2 "Configuring Traffic Policing".
Configure traffic shaping	traffic-shape [queue queue-id] max-rate burst-size	Optional. Refer to section 33.3.3 "Configuring Traffic Shaping".
Configure traffic priority	traffic-priority inbound	Optional. Refer to section 33.3.4 "Configuring Traffic Priority".
Configure traffic redirection	traffic-redirect inbound	Optional. Refer to section 33.3.5 "Configuring Traffic Redirection".
Configure queue scheduling algorithm	queue-scheduler wrr { group1 { queue-id queue-weight } &<1-8> group2 { queue-id queue-weight } &<1-8> }*	Optional. Refer to section 33.3.6 "Configuring Queue Scheduling".
Configure drop algorithm	drop-mode { tail-drop wred } [wred-index]	Optional. Refer to section 33.3.7 "Configuring WRED Parameters".
Configure traffic mirroring	mirrored-to inbound	Optional. Refer to section 33.3.8 "Configuring Traffic Mirroring".
Configure traffic statistics	traffic-statistic inbound	Optional. Refer to section 33.3.10 "Configuring Traffic Statistics".

Item	Command	Description
Display QoS configuration	display	You can execute the display command in any view to check the QoS configuration. Refer to section 33.3.11 "Displaying and Debugging QoS Configuration".

For the common interface boards except XP4, note that:

- The port group members must be on the same board and each port can only be added to one port group.
- The aggregated port cannot be added to the port group. If a port group member is to be aggregated, it exits from the port group automatically and the configuration of the primary port in an aggregated group overrides that of this port.
- When a single port is added to a port group, its configuration is overridden by that of the port group and the ACL rule cannot be applied to the port any longer.
- If there is no port in the port group, you cannot configure the QACL. If all the ports exist from the port group, the QACL configuration of the group still exists. And this configuration is applied again when new ports are added.

For the XP4 board, the system creates two port groups by default. One group contains ports 0 and 1, and the other contains ports 2 and 3. Their group numbers are $300 + 2 \times slot-no$ and $300 + 2 \times slot-no + 1$ (*slot-no* is the slot where the XP4 board locates) respectively, which are automatically assigned by the system. For example, when the XP4 board locates in slot 1, the group numbers are 302 and 303.

When configuring port groups for the XP4 board, pay attention to the following limitation:

- configure new port groups, and not allowed to join/remove any port into/from port group.
- The QACL commands are only allowed to be executed on port groups instead of individual ports.
- Traffic shaping is not supported.
- Port mirroring across groups are not supported. You can configure one incoming and one outgoing monitoring ports for each group (other kinds of interface boards have each of the monitoring ports per board).
- Four output queues are supported for queue scheduling.

33.2.2 Configuration Example for port group

I. Network requirements

Forward the packets sent from PC1 (IP 1.0.0.1), PC2(IP 2.0.0.1) during the time range from 8:00 to 18:00 every day to the address 3.0.0.1.

II. Network diagram

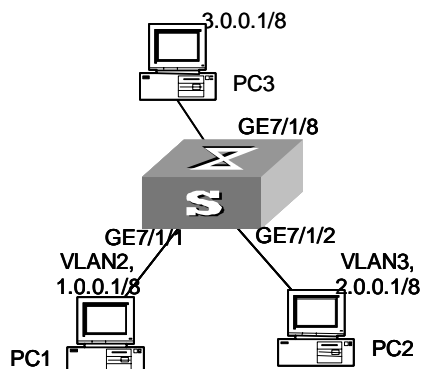


Figure 33-5 Network diagram for traffic redirection configuration

III. Configuration procedure

- 1) Define the time range.

Define the time range from 8:00 to 18:00.

```
[SW8800] time-range 3Com 8:00 to 18:00 daily
```

- 2) Define the traffic from PC1.

Create a number-based basic ACL 2000 and enter it.

```
[SW8800] acl number 2000
```

Define ACL rule for the traffic from PC1.

```
[SW8800-acl-basic-2000] rule 0 permit source 1.0.0.1 0 time-range 3Com
[SW8800-acl-basic-2000] quit
```

- 3) Create a port group.

Create port group 1 and enter the port group view.

```
[SW8800] port-group 1
```

Add the ports GE7/1/1 and GE7/1/2 to port group 1.

```
[SW8800-port-group1] port GigabitEthernet 7/1/1 GigabitEthernet 7/1/2
```

- 4) Modify the next hop for the packets from PC1.

Define the next hop for the packets from port group 1 as 3.0.0.1.

```
[SW8800-port-group1] traffic-redirect inbound ip-group 2000 rule 0 next-hop
3.0.0.1
```

33.3 QoS Configuration

The following sections describe QoS configuration tasks.

- Configuring Service Parameter Allocation Rule
- Configuring Traffic Policing

- Configuring Traffic Shaping
- Configuring Traffic Priority
- Configuring Traffic Redirection
- Configuring Queue Scheduling
- Configuring Traffic Mirroring
- Configuring Port Mirroring
- Configuring Traffic Statistics

Before initiating any of these QoS configuration tasks, you should first define the corresponding ACL. Then you can achieve packet filtering just by activating the right ACL.

To configure packet filtering, you need only to activate corresponding ACL. For more details, refer to the section 32.2.4 .

Some of QoS terms are listed in the following table.

Table 33-2 QoS terms

Term	Description
CoS	It has the same meaning as 802.1p priority. Both refer to the priority at packet header, with the value ranging from 0 to 7.
Service parameters	Switch allocates a set of parameters, which are used in achieving QoS functions, upon receiving a packet. Four items are included: 802.1p priority, DSCP priority, local precedence and drop precedence.
Drop-precedence	One of service parameters, ranging from 0 to 2. Drop precedence is allocated when the switch receives the packet and may be when the packet is processed. Allocating drop precedence to the packet is also called coloring the packet: the packet with drop precedence 2 as red, that with drop precedence 1 as yellow and that with drop precedence 0 as green. Drop precedence is referred to when switch needs to drop packets in its congestion.
Conform-Level	The result calculated from the user-defined CIR, CBS, EBS, PIR and actual traffic when the switch runs traffic policing, in the range of 0 to 2. It is used as a parameter in the traffic-limit command (here the value depends on the calculated result). It is also involved in the DSCP + Conform level → Service parameter mapping table which is used in re-allocating service parameters to a packet with the traffic-priority command. Then Conform-Level must be 0.

33.3.1 Configuring Service Parameter Allocation Rule

QoS is based on service parameters, a set of parameters for a packet, including 802.1p priority (CoS priority), DSCP priority, EXP priority, local precedence and drop precedence.

After receiving a packet, the switch allocates a set of service parameters to it according to a specific rule. The switch first gets its local precedence and drop precedence according to the packet 802.1p priority value, by searching in the CoS → Local-precedence mapping table and the CoS → Drop-precedence mapping table. Default values are available for the two mapping tables, but you can also configure the mapping tables according to your needs. If the switch fails in allocating local precedence for the packet, it configures the local precedence of the packet to be the precedence of the port that receives this packet. After obtaining the packet CoS value by inverse-searching the CoS → Local-precedence mapping table, the switch then gets its drop precedence from the CoS → Drop-precedence mapping table.

I. Configuring mapping table

Perform the following configurations in system view.

Table 33-3 Configure mapping tables

Operation	Command
Configure the CoS → Drop-precedence mapping table	qos cos-drop-precedence-map <i>cos0-map-drop-prec</i> <i>cos1-map-drop-prec</i> <i>cos2-map-drop-prec</i> <i>cos3-map-drop-prec</i> <i>cos4-map-drop-prec</i> <i>cos5-map-drop-prec</i> <i>cos6-map-drop-prec</i> <i>cos7-map-drop-prec</i>
Restore the default values of CoS → Drop-precedence mapping table	undo qos cos-drop-precedence-map
Configure the CoS → Local-precedence mapping table	qos cos-local-precedence-map <i>cos0-map-local-prec</i> <i>cos1-map-local-prec</i> <i>cos2-map-local-prec</i> <i>cos3-map-local-prec</i> <i>cos4-map-local-prec</i> <i>cos5-map-local-prec</i> <i>cos6-map-local-prec</i> <i>cos7-map-local-prec</i>
Restore the default values of CoS → Local-precedence mapping table	undo qos cos-local-precedence-map

By default, the switch obtains local precedence and drop precedence according to the default mapping values.

II. Configuring default local precedence for port

Perform the following configurations in Ethernet port view or port group view.

Table 33-4 Configure default local precedence for port

Operation	Command
Configure default local precedence for a port	priority <i>priority-level</i>
Restore the default local precedence for a port	undo priority

33.3.2 Configuring Traffic Policing

Traffic policing refers to rate limit based on traffic. If the traffic threshold is exceeded, corresponding measures will be taken, for example, dropping the excessive packets or re-defining their priority levels.

In the traffic supervision action, the switch uses the service parameters allocated according to the DSCP + Conform-Level → Service parameter mapping table and the EXP + Conform-Level → Service parameter mapping table and the 802.1p priority values allocated according to the Local-precedence + Conform-Level → 802.1p priority mapping table. So you should configure these three mapping tables or use their default values.

I. Configuring mapping tables

Perform the following configurations in the specified views.

Table 33-5 Configure mapping table

Operation	Command
Enter conform level view (System view)	qos conform-level <i>conform-level-value</i>
Configure the DSCP + Conform-Level → Service parameters mapping table (conform level view)	dscp <i>dscp-list</i> : <i>dscp-value</i> <i>exp-value</i> <i>cos-value</i> <i>local-precedence-value</i> <i>drop-precedence</i>
Restore the default values of the DSCP + Conform-Level → Service parameters mapping table (conform level view)	undo dscp <i>dscp-list</i>
Configure the EXP + Conform-Level → Service parameters mapping table (conform level view)	exp <i>exp-list</i> : <i>dscp-value</i> <i>exp-value</i> <i>cos-value</i> <i>local-precedence-value</i> <i>drop-precedence</i>
Restore the default values of the EXP + Conform-Level → Service parameters mapping table (conform level view)	undo exp <i>exp-list</i>
Configure the Local-precedence + Conform-Level → mapping table (conform level view)	local-precedence <i>cos-value0</i> <i>cos-value1</i> <i>cos-value2</i> <i>cos-value3</i> <i>cos-value4</i> <i>cos-value5</i> <i>cos-value6</i> <i>cos-value7</i>

Restore the default values of the Local-precedence + Conform-Level mapping table (conform level view) →	undo local-precedence
---	------------------------------

The system provides default mapping tables.

II. Configuring traffic policing

The purpose of this configuration task is to implement traffic policing on ACL-matched data streams, and then take normal actions on data streams within the traffic limit and take other actions (discarding packets, for example) on those exceeding the limit.

For interface cards, perform the following configurations in Ethernet port view or port group view.

Table 33-6 Configure traffic policing

Operation	Command
Configure traffic policing which only applies IP group ACL	traffic-limit inbound ip-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i> [system-index <i>index</i>]] [tc-index <i>index</i>] <i>cir</i> <i>cbs</i> <i>ebs</i> [<i>pir</i>] [conform { { remark-cos remark-drop-priority } * } remark-policed-service }] [exceed { forward drop }]
Remove traffic policing setting which only applies IP group ACL	undo traffic-limit inbound ip-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i>]
Configure traffic policing which applies IP group ACL and link group ACL at same time	traffic-limit inbound ip-group { <i>acl-number</i> <i>acl-name</i> } { rule <i>rule</i> link-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i> [system-index <i>index</i>]] link-group { <i>acl-number</i> <i>acl-name</i> } rule <i>rule</i> } [tc-index <i>index</i>] <i>cir</i> <i>cbs</i> <i>ebs</i> [<i>pir</i>] [conform { { remark-cos remark-drop-priority } * } remark-policed-service }] [exceed { forward drop }]
Remove traffic policing setting which applies IP group ACL and link group ACL at same time	undo traffic-limit inbound ip-group { <i>acl-number</i> <i>acl-name</i> } { rule <i>rule</i> link-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i>] link-group { <i>acl-number</i> <i>acl-name</i> } rule <i>rule</i> }
Configure traffic policing which only applies link group ACL	traffic-limit inbound link-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i> [system-index <i>index</i>]] [tc-index <i>index</i>] <i>cir</i> <i>cbs</i> <i>ebs</i> [<i>pir</i>] [conform { { remark-cos remark-drop-priority } * } remark-policed-service }] [exceed { forward drop }]
Remove traffic policing setting which only applies link group ACL	undo traffic-limit inbound link-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i>]

Note:

- The parameters of traffic policy must be the same if you configure the same tc-index for different traffic; otherwise the system prompts you for the wrong configuration.
 - For traffic policing configuration over the port group, all ports in the group shares the same bandwidth, that is, the traffic parameters you define take effect on all ports in the group.
-

Caution:

- Before configuring traffic policing, you must first define corresponding ACLs and configure the DSCP+ Conform-Level → Service parameters mapping table and the Local-precedence + Conform-Level → 802.1p priority mapping table.
-

You must first define the corresponding ACL and configure the DSCP + Conform-Level → Service parameters mapping table and Local-precedence + Conform-Level → mapping table before starting this configuration.

This configuration achieves traffic policing for the packets that match the ACL. If the traffic rate threshold is exceeded, corresponding measures will be taken, for example, dropping excessive packets.

system-index *index* here is the system index for an ACL rule. When delivering a rule, the system assigns an index to it, for convenience of later retrieval. You can also assign a system index for it when delivering an ACL rule with this command. However, you are not recommended to assign a system index if not urgently necessary.

tc-index *index* here is traffic policing index, in the range of 0 to 12288. If you configure the same index for different ACL rules during setting traffic policing, then the sum of traffic shall be limited by the traffic policing-related parameters predefined. For example, if CIR (committed information rate) of the traffic that matches ACL1 is set to 10 kbps and that for ACL2 to 10 kbps, and their traffic policing indexes are the same, then the average rate of the traffic that matches ACL1 and ACL2 shall be limited to 10kbps.

See the corresponding Command Manual for details of the commands.

33.3.3 Configuring Traffic Shaping

Traffic shaping controls the rate of outbound packets, to ensure they are sent at relatively average rates. Traffic shaping measure tries to match packet transmission rate with the capacity of downstream devices. Its major difference from traffic policing is: Traffic shaping buffers packets at over-threshold rates to make them sent at average

rates, while traffic policing drops excessive packets. Therefore, traffic shaping may increase transmission delay, but not for traffic policing.

Perform the following configurations in Ethernet port view or port group view.

Table 33-7 Configure traffic shaping

Operation	Command
Configure traffic shaping	traffic-shape [queue <i>queue-id</i>] <i>max-rate</i> <i>burst-size</i>
Remove traffic shaping setting	undo traffic-shape [queue <i>queue-id</i>]

The switch supports traffic shaping based on port, that is, all traffic on the port is shaped. It also supports traffic shaping for a specific queue. You can choose to achieve one of them by selecting different parameters in the command.

See the corresponding Command Manual for details of the commands.

33.3.4 Configuring Traffic Priority

This configuration re-labels priority value for the packets that match the ACL in these ways: using the service parameters allocated by the switch, re-allocating service parameters by searching the mapping table based on the packet DSCP value, re-allocating service parameters by searching the mapping table based on the specified DSCP value and EXP value, customizing service parameters for the packets.

For interface cards, perform the following configurations in Ethernet port view or port group view.

Table 33-8 Configure traffic priority

Operation	Command
Configure traffic priority which only applies IP group ACL	traffic-priority inbound ip-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i> [system-index <i>index</i>]] { auto remark-policed-service { trust-dscp dscp <i>dscp-value</i> untrusted dscp <i>dscp-value</i> cos <i>cos-value</i> local-precedence <i>local-precedence</i> drop-priority <i>drop-level</i> } }
Remove traffic priority setting which only applies IP group ACL	undo traffic-priority inbound ip-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i>]
Configure traffic priority which applies IP group ACL and link group ACL at same time	traffic-priority inbound ip-group { <i>acl-number</i> <i>acl-name</i> } { rule <i>rule</i> link-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i> [system-index <i>index</i>]] link-group { <i>acl-number</i> <i>acl-name</i> } rule <i>rule</i> } { auto remark-policed-service { trust-dscp dscp <i>dscp-value</i> untrusted dscp <i>dscp-value</i> cos <i>cos-value</i> local-precedence <i>local-precedence</i> drop-priority <i>drop-level</i> } }

Remove traffic priority setting which applies IP group ACL and link group ACL at same time	undo traffic-priority inbound ip-group { <i>acl-number</i> <i>acl-name</i> } { <i>rule rule</i> link-group { <i>acl-number</i> <i>acl-name</i> } [<i>rule rule</i>] link-group { <i>acl-number</i> <i>acl-name</i> } <i>rule rule</i> }
Configure traffic priority which only applies link group ACL	traffic-priority inbound link-group { <i>acl-number</i> <i>acl-name</i> } [<i>rule rule</i> [system-index <i>index</i>]] { auto remark-policed-service { trust-dscp dscp <i>dscp-value</i> untrusted dscp <i>dscp-value</i> cos <i>cos-value</i> local-precedence <i>local-precedence</i> drop-priority <i>drop-level</i> } }
Remove traffic priority setting which only applies link group ACL	undo traffic-priority inbound link-group { <i>acl-number</i> <i>acl-name</i> } [<i>rule rule</i>]

system-index *index* here is the system index for an ACL rule. When delivering a rule, the system assigns an index to it, for convenience of later retrieval. You can also assign a system index for it when delivering an ACL rule with this command. However, you are not recommended to assign a system index if not urgently necessary.

Note:

- For MPLS packets, other than that the *dscp-value* stands for their DSCP priority value, the three low-order bits of the value represent the EXP flag field. You set the EXP value when defining the *dscp-value*.
 - The DSCP + Conform-Level 0 → Service parameters mapping table and the EXP + Conform-Level → Service parameters mapping table (the mapping table for conform level 0) is used here.
-

See the corresponding Command Manual for details of the commands.

33.3.5 Configuring Traffic Redirection

Traffic redirection changes packet forwarding direction, to CPU, other ports, other IP addresses or other cards.

For interface cards, perform the following configurations in Ethernet port view or port group view.

Table 33-9 Configure traffic redirection

Operation	Command
Configure traffic redirection which only applies IP group ACL	traffic-redirect inbound ip-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i> [system-index <i>index</i>]] { cpu interface { <i>interface-name</i> <i>interface-type</i> <i>interface-num</i> } <i>destination-vlan</i> { I2-vpn I3-vpn } next-hop <i>ip-addr1</i> [<i>ip-addr2</i>] slot <i>slotid</i> <i>vlanid</i> }
Remove traffic redirection setting which only applies IP group ACL	undo traffic-redirect inbound ip-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i>]
Configure traffic redirection which applies IP group ACL and link group ACL at same time	traffic-redirect inbound ip-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i>] link-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i>]] { cpu interface { <i>interface-name</i> <i>interface-type</i> <i>interface-num</i> } <i>destination-vlan</i> { I2-vpn I3-vpn } next-hop <i>ip-addr1</i> [<i>ip-addr2</i>] slot <i>slotid</i> <i>vlanid</i> }
Remove traffic redirection setting which applies IP group ACL and link group ACL at same time	undo traffic-redirect inbound ip-group { <i>acl-number</i> <i>acl-name</i> } { rule <i>rule</i> link-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i>] link-group { <i>acl-number</i> <i>acl-name</i> } rule <i>rule</i> } or undo traffic-redirect inbound link-group { <i>acl-number</i> <i>acl-name</i> } { rule <i>rule</i> ip-group { <i>acl-number</i> <i>acl-name</i> } ip-group { <i>acl-number</i> <i>acl-name</i> } rule <i>rule</i> }
Configure traffic redirection which only applies link group ACL	traffic-redirect inbound link-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i> [system-index <i>index</i>]] { cpu interface { <i>interface-name</i> <i>interface-type</i> <i>interface-num</i> } <i>destination-vlan</i> { I2-vpn I3-vpn } next-hop <i>ip-addr1</i> [<i>ip-addr2</i>] slot <i>slotid</i> <i>vlanid</i> }
Remove traffic redirection setting which only applies link group ACL	undo traffic-redirect inbound link-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i>]

system-index *index* here is the system index for an ACL rule. When delivering a rule, the system assigns an index to it, for convenience of later retrieval. You can also assign a system index for it when delivering an ACL rule with this command. However, you are not recommended to assign a system index if not urgently necessary.

Note:

- Traffic redirection setting is only available for the permitted rules in the ACL.
 - The packet redirected to the CPU cannot be forwarded normally.
 - You can achieve policy route by selecting the **next-hop** keyword.
-

See the corresponding Command Manual for details of the commands.

33.3.6 Configuring Queue Scheduling

Each port supports eight outbound queues except that port of XP4 board only supports four queues. The switch puts the packets into the queues according to the local precedence of packets. Queue scheduling is used to resolve problems of resource contention by many packets. The switch supports SP algorithm and WRR algorithm.

Different outbound queues at the port may use different algorithms. The switch supports three scheduling modes:

- 1) All-SP scheduling mode
- 2) All-WRR mode: The outbound queues are divided into WRR queue 1 and WRR queue 2. The switch first schedules the queues in the WRR queue1. If no packets are waiting for being forwarded in WRR queue 1, then it begins to schedule the queues in WRR queue 2. By default, all queues at a port are in WRR queue 1.
- 3) SP plus WRR mode: The outbound queues are put into different scheduling groups. SP group uses SP algorithm, WRR groups use WRR algorithm. The select one queue respectively from SP group, WRR group 1 and WRR group 2 and schedule them using SP algorithm.

Perform the following configurations in Ethernet port view or port group view.

Table 33-10 Configure queue scheduling

Operation	Command
Configuring queue scheduling	<code>queue-scheduler wrr { group1 { queue-id queue-weight } &<1-8> group2 { queue-id queue-weight } &<1-8> }*</code>
Restore the default setting	<code>undo queue-scheduler [queue-id] &<1-8></code>

By default, the switch uses all-SP mode, so those queues not configured with WRR algorithm are SP mode.

See the corresponding Command Manual for details of the commands.

33.3.7 Configuring WRED Parameters

In the case of network congestion, the switch drops packets to release system resources. And then no packets are put into long-delay queues.

The switch allocates drop precedence for it when receiving a packet (also called coloring the packet). The drop precedence values range from 0 to 2, with 2 for red, 1 for yellow and 0 for green. In congestion, red packets will be first dropped, and green packets last.

You can configure drop parameters and thresholds by queue or drop precedence.

The following two drop modes are available:

- 1) Tail drop mode: Different queues (red, yellow and red) are allocated with different drop thresholds. When these thresholds are exceeded respectively, excessive packets will be dropped.
- 2) WRED drop mode: Drop precedence is taken into account in drop action. When only min-thresholds of red, yellow and green packets are exceeded, excessive packets are dropped randomly at given probability. But when max-thresholds of red, yellow and green packets are exceeded, all excessive packets will be dropped.

You must first configure WRED parameters for every outbound queue in defining drop precedence.

I. Configuring WRED parameters

The switch provides four sets of default WRED parameters, respectively numbered as 0 to 3. Each set includes 80 parameters, 10 parameters for each of the eight queues. The ten parameters are *green-min-threshold*, *yellow-min-threshold*, *red-min-threshold*, *green-max-threshold*, *yellow-max-threshold*, *red-max-threshold*, *green-max-prob*, *yellow-max-prob*, *red-max-prob* and *exponent*. Red, yellow and green packets respectively refer to those with drop precedence levels 2, 1 and 0.

Perform the following configurations in the specified views.

Table 33-11 Configure WRED parameters

Operation	Command
Enter WRED index view (system view)	wred <i>wred-index</i>
Restore the default WRED parameters (system view)	undo wred <i>wred-index</i>
Configure WRED parameters (WRED index view)	queue <i>queue-id green-min-threshold green-max-threshold green-max-prob yellow-min-threshold yellow-max-threshold yellow-max-prob red-min-threshold red-max-threshold red-max-prob exponent</i>
Restore the default WRED parameters (WRED index)	undo queue <i>queue-id</i>
Exit WRED index view (WRED index view)	quit

The command restores the parameters of the specified WRED index as the default setting. The command restores the WRED parameters related to the queue as the default setting.

The switch provides four sets of WRED parameters by default.

See the corresponding Command Manual for details of the commands.

II. Configuring drop algorithm

Please perform the following configurations in Ethernet port view.

Table 33-12 Configure drop algorithm

Operation	Command
Configure drop algorithm	drop-mode { tail-drop wred } [<i>wred-index</i>]
Restore the default algorithm	undo drop-mode

By default, tail drop mode is selected.

See the corresponding Command Manual for details of the commands.

33.3.8 Configuring Traffic Mirroring

Traffic mirroring duplicates the traffic that matches ACL rules to the CPU, for traffic analysis and monitoring.

Perform the following configurations in Ethernet port view or port group view.

Table 33-13 Configure traffic mirroring

Operation	Command
Configure traffic mirroring which only applies IP group ACL	mirrored-to inbound ip-group { <i>acl-number</i> <i>acl-name</i> } [rule rule [system-index index]] cpu
Remove traffic mirroring setting which only applies IP group ACL	undo mirrored-to inbound ip-group { <i>acl-number</i> <i>acl-name</i> } [rule rule]
Configure traffic mirroring which applies IP group ACL and link group ACL at same time	mirrored-to inbound ip-group { <i>acl-number</i> <i>acl-name</i> } { rule rule [system-index index] } link-group { <i>acl-number</i> <i>acl-name</i> } link-group { <i>acl-number</i> <i>acl-name</i> } rule rule } cpu
Remove traffic mirroring setting which applies IP group ACL and link group ACL at same time	undo mirrored-to inbound ip-group { <i>acl-number</i> <i>acl-name</i> } { rule rule [system-index index] } link-group { <i>acl-number</i> <i>acl-name</i> } rule rule }
Configure traffic mirroring which only applies link group ACL	mirrored-to inbound link-group { <i>acl-number</i> <i>acl-name</i> } [rule rule [system-index index]] cpu
Remove traffic mirroring setting which only applies link group ACL	undo mirrored-to inbound link-group { <i>acl-number</i> <i>acl-name</i> } [rule rule]

system-index index here is the system index for an ACL rule. When delivering a rule, the system assigns an index to it, for convenience of later retrieval. You can also assign

a system index for it when delivering an ACL rule with this command. However, you are not recommended to assign a system index if not urgently necessary.

See the corresponding Command Manual for details of the commands.

33.3.9 Configuring Port Mirroring

Port mirroring duplicates data on the monitored port to the designated monitoring port, for purpose of data analysis and supervision. The switch supports multiple-to-one mirroring, that is, you can duplicate packets from multiple ports to a monitoring port.

You can also specify the monitoring direction:

- Only inbound packets
- Only outbound packets

Perform the following configurations in system view.

Table 33-14 Configure port mirroring

Operation	Command
Configure port mirroring	mirroring-group <i>groupId</i> { inbound outbound } <i>mirroring-port-list</i> &<1-8> mirrored-to <i>monitor-port</i>
Remove port mirroring setting	undo mirroring-group <i>groupId</i>

You can implement port mirroring configuration by setting mirroring groups at the port. Up to 20 mirroring groups can be configured at a port, with each group including one monitoring port and multiple monitored ports.

Note:

The Switch 8800 supports cross-board mirroring, that is, the monitoring and monitored ports can be at different boards.

Consider these issues when configuring port mirroring:

- For intra-board mirroring, only one monitoring port can be configured for the mirroring groups in the same direction. For example, if one mirroring group, with port A as its monitoring port, has been configured on a board to monitor those received packets, you need to choose port A as its monitoring port when configuring a second mirroring group to monitor those received packets. The same restriction applies to the mirroring group to monitor those packets sent.
- For cross-board mirroring, only one monitoring port (which is on another board) can be configured for the mirroring groups in the same direction. For example, a mirroring group is configured on board 1, with port B on board 2 as its monitoring

port. You can only choose port B on board 2 as its monitoring port when configuring a second mirroring group in the same direction on board 1.

- One mirroring group can contain as many as 24 monitored ports at most.
- You can configure as many as 24 monitored ports for all the mirroring groups in transmit group.
- You can configure 24 mirroring groups in both directions in total.
- One port can act as a mirroring port and a mirrored port at the same time for a different mirroring group in a different direction.

More issues for the GV48 board (LSBM1GV48DA):

- For the mirroring (including incoming port mirroring and outgoing port mirroring) on the same GV48 board, only one monitoring port is allowed. For example, if you have configured a port mirroring group in a GV48 board, with port A as the monitoring port, then you can only choose port A as its monitoring port when configuring another port mirroring group.
- For all port groups configured in the system, only one monitoring port is allowed on the same GV48 board.

For the XP4 board, the system creates two port groups by default. One group contains ports 0 and 1, and the other contains ports 2 and 3. Pay attention to the following limitation on port mirroring:

- Port mirroring across groups are not supported. That is, in a port monitoring group, the monitoring port and monitored port can only be ports 0 and 1, or ports 2 and 3.
- A port group can contain one incoming and one outgoing monitoring ports (other interface boards have each of them per board).

See the corresponding Command Manual for details of the commands.

33.3.10 Configuring Traffic Statistics

Traffic statistics count packets of designated service traffic, that is, the packets match the defined ACL among those forwarded. You can view the information with the **display qos-interface traffic-statistic** command.

Perform the following configurations in Ethernet port or port group view.

Table 33-15 Configure traffic statistics

Operation	Command
Configure traffic statistics which only applies IP group ACL	traffic-statistic inbound ip-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i>] [system-index <i>index</i>] [tc-index <i>index</i>]
Remove traffic statistics setting which only applies IP group ACL	undo traffic-statistic inbound ip-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i>]

Configure traffic statistics which only applies link group ACL	traffic-statistic inbound link-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i>] [system-index <i>index</i>] [tc-index <i>index</i>]
Remove traffic statistics setting which only applies link group ACL	undo traffic-statistic inbound link-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i>]
Display traffic statistics for the port	display qos-interface [<i>interface-name</i> <i>interface-type</i> <i>interface-num</i>] traffic-statistic

Note:

The system counts the traffic on all ports in the group after you use the **traffic-statistic** command in port group view.

system-index *index* here is the system index for an ACL rule. When delivering a rule, the system assigns an index to it, for convenience of later retrieval. You can also assign a system index for it when delivering an ACL rule with this command. However, you are not recommended to assign a system index if not urgently necessary.

See the corresponding Command Manual for details of the commands.

33.3.11 Displaying and Debugging QoS Configuration

After these configurations are completed, you can use the **display** command in any view to view QoS running and check configuration result. You can clear QoS statistics using the **reset traffic-statistic** command in Ethernet port view or port group view.

Table 33-16 Display and debug QoS configurations

Operation	Command
Display traffic mirroring configuration of a port	display qos-interface [<i>interface-name</i> <i>interface-type</i> <i>interface-num</i>] mirrored-to
Display traffic priority configuration of a port	display qos-interface [<i>interface-name</i> <i>interface-type</i> <i>interface-num</i>] traffic-priority
Display traffic redirection configuration of a port	display qos-interface [<i>interface-name</i> <i>interface-type</i> <i>interface-num</i>] traffic-redirect
Display traffic statistics of a port	display qos-interface [<i>interface-name</i> <i>interface-type</i> <i>interface-num</i>] traffic-statistic
Display port mirroring configuration	display mirroring-group [<i>groupid</i>]
Display QoS configurations of all ports	display qos-interface [<i>interface-name</i> <i>interface-type</i> <i>interface-num</i>] all

Operation	Command
Display traffic limit configuration of a port	display qos-interface [<i>interface-name</i> <i>interface-type interface-num</i>] traffic-limit
Display queue scheduling configuration of a port	display qos-interface [<i>interface-name</i> <i>interface-type interface-num</i>] queue-scheduler
Display traffic shaping configuration of a port	display qos-interface [<i>interface-name</i> <i>interface-type interface-num</i>] traffic-shape
Display the parameter settings for traffic policing	display traffic-params [<i>traffic-index</i>]
Display QoS configuration of a VLAN	display qos-vlan [<i>vlan-id</i>] all
Display traffic priority configuration of a VLAN	display qos-vlan [<i>vlan-id</i>] traffic-priority
Display traffic limit configuration of a VLAN	display qos-vlan [<i>vlan-id</i>] traffic-limit
Display traffic direction configuration of a VLAN	display qos-vlan [<i>vlan-id</i>] traffic-redirect
Display traffic statistics of a VLAN	display qos-vlan [<i>vlan-id</i>] traffic-statistic
Display the DSCP + Conform-level — > Service parameter, EXP + Conform-level — > Service parameter and Local-precedence + Conform-level — > 802.1p priority mapping tables	display qos conform-level [<i>conform-level-value</i>] { dscp-policed-service-map [<i>dscp-list</i>] exp-policed-service-map local-precedence-cos-map }
Display the CoS — > Drop-precedence mapping table	display qos cos-drop-precedence-map
Display the CoS — > Local-precedence mapping table	display qos cos-local-precedence-map
Clear traffic statistics	reset traffic-statistic inbound { { ip-group { <i>acl-number</i> <i>acl-name</i> } rule <i>rule</i> link-group { <i>acl-number</i> <i>acl-name</i> } }* { ip-group { <i>acl-number</i> <i>acl-name</i> } link-group { <i>acl-number</i> <i>acl-name</i> } rule <i>rule</i> }* ip-group { <i>acl-number</i> <i>acl-name</i> } rule <i>rule</i> link-group { <i>acl-number</i> <i>acl-name</i> } rule <i>rule</i> }

See the corresponding Command Manual for description of display information and parameters.

33.4 Configuration Example

33.4.1 Traffic Shaping Configuration Example

I. Network requirements

Set traffic shaping for the outbound queue 2 at the port GE7/1/8: maximum rate 500kbps, burst size 12k bytes.

II. Network diagram

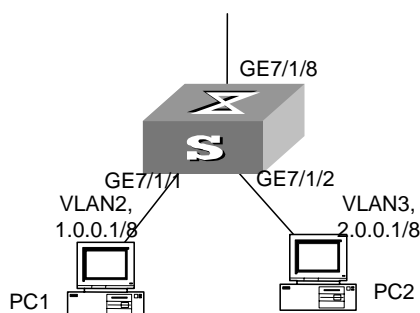


Figure 33-6 Network diagram for QoS configuration

III. Configuration procedure

- 1) Enter Ethernet port view.

```
[SW8800] interface GigabitEthernet 7/1/8
```

```
[SW8800-GigabitEthernet7/1/8]
```

- 2) Set traffic shaping for the outbound queue 2 at the port: maximum rate 500 Kbps, burst size 12 KB.

```
[SW8800-GigabitEthernet7/1/8] traffic-shape queue 2 500 12
```

33.4.2 Port Mirroring Configuration Example

I. Network requirements

Use one server to monitor the packets of two ports. R&D department is accessed from the port GE3/1/1 and sales department from the port GE3/1/2. The server is connected to the port GE3/1/8.

II. Network diagram

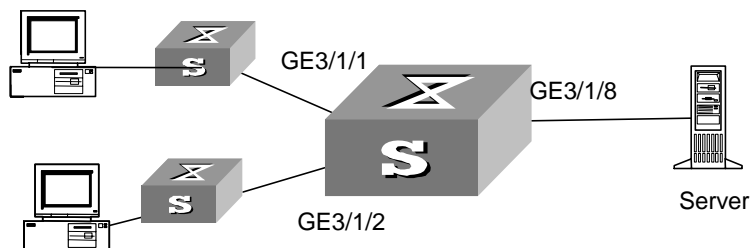


Figure 33-7 Networking for port mirroring configuration

III. Configuration procedure

Define a mirroring group, with monitoring port as GigabitEthernet3/1/8.

```
[SW8800] mirroring-group 1 inbound gigabitethernet3/1/1 gigabitethernet3/1/2
mirrored-to gigabitethernet3/1/8
[SW8800] mirroring-group 2 outbound gigabitethernet3/1/1
gigabitethernet3/1/2 mirrored-to gigabitethernet3/1/8
```

33.4.3 Traffic Priority Configuration Example

I. Network requirements

Re-allocate service parameters according to the mapping table for DSCP 63 for the packets from PC1 (IP 1.0.0.1) during the time range 8:00 to 18:00 everyday.

II. Network diagram

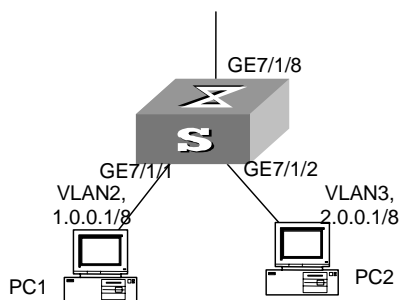


Figure 33-8 Network diagram for priority configuration

III. Configuration procedure

- 1) Define the time range.

Define the time range from 8:00 to 18:00.

```
[SW8800] time-range 3Com 8:00 to 18:00 daily
```

- 2) Define the traffic from PC1.

Create a number-based basic ACL 2000 and enter it.

```
[SW8800] acl number 2000
```

Define ACL rule for the traffic from PC1.

```
[SW8800-acl-basic-2000] rule 0 permit source 1.0.0.1 0 time-range 3Com
```

3) Define the CoS → Conform-Level mapping table.

Define the CoS → Conform-Level mapping table. The switch allocates drop precedence (all as 0 for the sake of simplification) for them when receiving packets.

```
[SW8800] qos cos-drop-precedence-map 0 0 0 0 0 0 0 0
```

The modified CoS → Conform-Level mapping table:

Table 33-17 Modified CoS → Conform-Level mapping table

CoS Value	Drop-precedence
0	0
1	0
2	0
3	0
4	0
5	0
6	0
7	0

4) Define the DSCP + Conform-Level → Service parameter mapping table.

Define the DSCP + Conform-Level → Service parameter mapping table. Allocate a set of service parameters for the packets from PC1 according the mapping table for DSCP 63.

```
[SW8800] qos conform-level 0
```

```
[SW8800-conform-level-0] dscp 63 : 32 4 4 4 0
```

The modified DSCP + Conform-Level → Service parameter mapping table:

Table 33-18 Modified DSCP + Conform-Level → Service parameter mapping table

DSCP	CL	Policed-DSCP	Policed-exp	Policed-802.1p	Policed-Localprec	Policed-Drop Precedence
63	0	32	4	4	4	0

5) Re-allocate service parameters for the packets from PC1.

Re-allocate service parameters for the packets from PC1.

```
[SW8800-GigabitEthernet7/1/1] traffic-priority inbound ip-group 2000
remark-policed-service dscp 63
```

33.4.4 Traffic Redirection Configuration Example

I. Network requirements

Forward the packets sent from PC1 (IP 1.0.0.1) during the time range from 8:00 to 18:00 every day to the address 2.0.0.1.

II. Network diagram

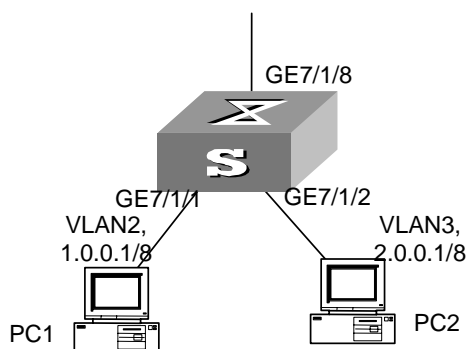


Figure 33-9 Network diagram for traffic redirection configuration

III. Configuration procedure

- 1) Define the time range.

Define the time range from 8:00 to 18:00.

```
[SW8800] time-range 3Com 8:00 to 18:00 daily
```

- 2) Define the traffic from PC1.

Create a number-based basic ACL 2000 and enter it.

```
[SW8800] acl number 2000
```

Define ACL rule for the traffic from PC1.

```
[SW8800-acl-basic-2000] rule 0 permit source 1.0.0.1 0 time-range 3Com
```

- 3) Modify the next hop for the packets from PC1.

Define the next hop for the packets from PC1 as 2.0.0.1.

```
[SW8800-GigabitEthernet7/1/1] traffic-redirect inbound ip-group 2000 rule 0
next-hop 2.0.0.1
```

33.4.5 Queue Scheduling Configuration Example

I. Network requirements

Modify the correspondence between 802.1p priority levels and local priority levels to change the mapping between 802.1p priority levels and queues. That is, put packets into outbound queues according to the new mapping. Use WRR algorithm for the queues 0 to 5 at the port GE7/1/1. Set the queues 0, 1 and 2 into WRR queue 1, with weight respectively as 20, 20 and 30; set the queues 3, 4 and 5 into WRR queue 2, with weight respectively as 20, 20 and 40. The queues 6 and 7 use SP algorithm. See Queue Scheduling for the default mapping.

Table 33-19 802.1p priority → Local precedence mapping table

802.1p priority	Local precedence
0	7
1	6
2	5
3	4
4	3
5	2
6	1
7	0

II. Network diagram

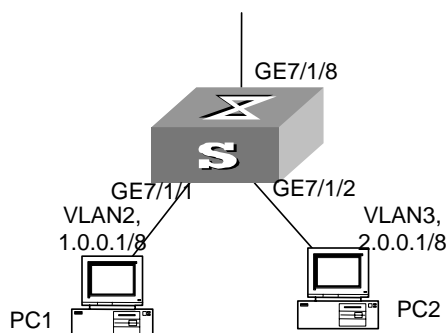


Figure 33-10 Network diagram for queue-schedule configuration

III. Configuration procedure

- 1) Re-specify the mapping between 802.1p priority and local precedence.

```
[SW8800] qos cos-local-precedence-map 7 6 5 4 3 2 1 0
```

- 2) Use WRR algorithm for the queues 0 to 5. Set the queues 0, 1 and 2 into WRR queue 1, with weight respectively as 20, 20 and 30; set the queues 3, 4 and 5 into WRR queue 2, with weight respectively as 20, 20 and 40. Use SP algorithm for the queues 6 and 7.

```
[SW8800-GigabitEthernet7/1/1] queue-scheduler wrr group1 0 20 1 20 2 30 group2
3 20 4 20 5 40
```

```
[SW8800] display qos-interface GigabitEthernet7/1/1 queue-scheduler
```

```
GigabitEthernet7/1/1 Port scheduling:
```

QID:	scheduling-group	weight
0 :	wrr , group1	20
1 :	wrr , group1	20
2 :	wrr , group1	30
3 :	wrr , group2	20
4 :	wrr , group2	20
5 :	wrr , group2	40
6 :	sp	0
7 :	sp	0

33.4.6 WRED Parameters Configuration Example

I. Network requirements

Set WRED parameters and drop algorithm for packets at the port GE7/1/1: Configure parameters for WRED 0; outbound queue ID is 7; *green-min-threshold* is 150; *green-max-threshold* is 500; *green-max-prob* is 5; *yellow-min-threshold* is 100; *yellow-max-threshold* is 150; *yellow-max-prob* is 10; *red-min-threshold* is 50; *red-max-threshold* is 100; *red-max-prob* is 15; *exponent* is 10; the port is in WRED drop mode; import the parameters of WRED 0.

II. Network diagram

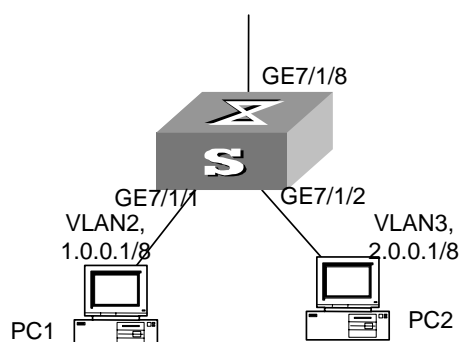


Figure 33-11 Network diagram for WRED parameters configuration

III. Configuration procedure

1) Configure WRED parameters

Configure parameters for WRED 0.

```
[SW8800] wred 0
[SW8800-wred-0] queue 7 150 500 5 100 150 10 50 100 15 10
```

2) Set drop algorithm and thresholds.

Define the port GE7/1/1 in WRED drop mode, set the parameters of WRED 0.

```
[SW8800-GigabitEthernet7/1/1] drop-mode wred 0
```

33.4.7 Traffic Statistics Configuration Example

I. Network requirements

Suppose the IP address of PC1 is 1.0.0.1 and that of PC2 is 2.0.0.1. The switch is up-linked through the port GE7/1/8. Count the packets sent from the switch to PC1 during the time range from 8:00 to 18:00 every day.

II. Network diagram

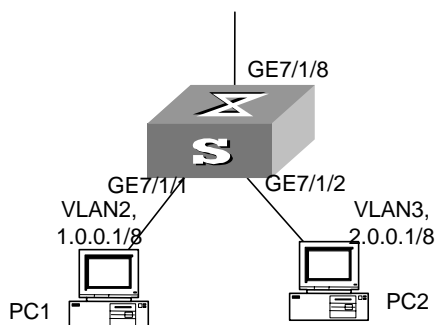


Figure 33-12 Network diagram for traffic statistics configuration

III. Configuration procedure

1) Define the time range.

Define the time range from 8:00 to 18:00.

```
[SW8800] time-range 3Com 8:00 to 18:00 daily
```

2) Define the traffic from PC1.

Define ACL rule for the traffic from PC1.

```
[SW8800] acl number 2000
[SW8800-acl-basic-2000] rule 0 permit source 1.0.0.1 0.0.0.0 time-range 3Com
```

3) Count the packets to PC1 and display the result using the **display** command.

```
[SW8800-GigabitEthernet7/1/1] traffic-statistic inbound ip-group 2000 rule 0
[SW8800] display qos-interface GigabitEthernet7/1/1 traffic-statistic
GigabitEthernet7/1/1: traffic-statistic
Inbound:
  Matches: Acl 2000 rule 0 running
    12002688 bytes (green 1270244416 byte(s), yellow 1895874880 byte(s), red
704683968 byte(s) )
    3333270 packets (green 0 byte(s), yellow 0 byte(s), red 0 byte(s) )
```

Chapter 34 Logon User ACL Control Configuration

34.1 Overview

As the Ethernet switches are used more and more widely over the networks, the security issue becomes even more important. The switches provide several logon and device accessing measures, mainly including TELNET access, SNMP access, and HTTP access (currently the Switch 8800 does not support it). The security control over the access measures is provided with the switches to prevent illegal users from logging on to and accessing the devices. There are two levels of security controls. At the first level, the user connection is controlled with ACL filter and only the legal users can be connected to the switch. At the second level, a connected user can log on to the device only if he can pass the password authentication.

This chapter mainly introduces how to configure the first level security control over these access measures, that is, how to configure to filter the logon users with ACL. For detailed description about how to configure the first level security, refer to “getting started” module of Operation Manual.

34.2 Configuring ACL for Telnet Users

This configuration can filter out malicious or illegal connection request before password authentication.

The following sections describe ACL configuration tasks.

- Defining ACL
- Importing ACL

34.2.1 Defining ACL

Currently number-based ACLs or advanced ACL can be imported, with the number ranging from 2000 to 3999.

Perform the following configurations in system view.

Table 34-1 Define basic ACL and advanced ACL

Operation	Command
Enter basic ACL (system view)	acl { number <i>acl-number</i> name <i>acl-name</i> basic } match-order { config auto }
Define a sub-rule (basic ACL view)	rule [<i>rule-id</i>] { permit deny } [source <i>source-addr wildcard</i> any] [fragment] [time-range <i>name</i>]

Operation	Command
Delete a sub-rule (basic ACL view)	undo rule <i>rule-id</i> [source] [fragment] [time-range]
Delete an ACL or all ACLs (system view)	undo acl { number <i>acl-number</i> name <i>acl-name</i> / all }
Enter advanced ACL view from system view	acl { number <i>acl-number</i> name <i>acl-name</i> advanced } [match-order { config auto }]
Define sub-rule(advanced ACL view)	rule [<i>rule-id</i>] { permit deny } <i>protocol</i> [source { <i>source-addr wildcard</i> any }] [destination { <i>dest-addr wildcard</i> any }] [source-port <i>operator port1</i> [<i>port2</i>]] [destination-port <i>operator port1</i> [<i>port2</i>]] [icmp-type <i>type code</i>] [established] [[precedence <i>precedence</i> tos <i>tos</i>]*] [dscp <i>dscp</i>] [fragment] [time-range <i>name</i>] [vpn-instance <i>instance-name</i>]
Delete a sub-rule(advanced ACL view)	undo rule <i>rule-id</i> [source destination source-port destination-port icmp-type precedence tos dscp fragment time-range vpn-instance]*
Delete an ACL or all ACLs (system view)	undo acl { number <i>acl-number</i> name <i>acl-name</i> / all }

You can define multiple rules for an ACL by using the **rule** command several times.

34.2.2 Importing ACL

You can import a defined ACL in user interface view to achieve ACL control.

Perform the following configurations respectively in system view and user interface view.

Table 34-2 Import ACL

Operation	Command
Enter user interface view (system view)	user-interface [<i>type</i>] <i>first-number</i>
Import the ACL (user interface view)	acl <i>acl-number</i> { inbound outbound }

See the Command Manual for details about these commands.

Note:

Currently the ACL control function of TELNET user can reference to the number-based ACLs and advanced ACLs.

34.2.3 Configuration Example

I. Network requirements

Only the Telnet users from 10.110.100.52 and 10.110.100.46 can access the switch.

II. Network diagram

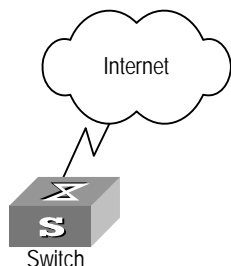


Figure 34-1 ACL configuration for Telnet users

III. Configuration procedure

Define a basic ACL.

```
[SW8800] acl number 2000 match-order config
[SW8800-acl-basic-2000] rule 1 permit source 10.110.100.52 0
[SW8800-acl-basic-2000] rule 2 permit source 10.110.100.46 0
[SW8800-acl-basic-2000] rule 3 deny source any
[SW8800-acl-basic-2000] quit
```

Import the ACL.

```
[SW8800] user-interface vty 0 4
[SW8800-user-interface-vty0-4] acl 2000 inbound
```

34.3 Configuring ACL for SNMP Users

The Switch 8800 supports remote network management (NM) and the user can use SNMP to access them. Proper ACL configuration can prevent illegal users from logging onto the switches.

The following sections describe ACL configuration tasks.

- Defining ACL
- Importing ACL

34.3.1 Defining ACL

Currently only number-based ACLs can be imported, with the number ranging from 2000 to 2999. See 34.3.1 “Defining ACL” for detailed configuration.

34.3.2 Importing ACL

Import the defined ACL into the commands with SNMP community, username and group name configured, to achieve ACL control over SNMP users.

Perform the following configurations in system view.

Table 34-3 Import ACL

Operation	Command
Import the defined ACL into the commands with SNMP community configured	snmp-agent community { read write } <i>community-name</i> [[mib-view <i>view-name</i>]] [acl <i>acl-number</i>]]*
Import the defined ACL into the commands with SNMP group name configured	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>] 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>]
Import the defined ACL into the commands with SNMP username configured	snmp-agent usm-user { v1 v2c } <i>user-name</i> <i>group-name</i> [acl <i>acl-number</i>] snmp-agent usm-user v3 <i>user-name</i> <i>group-name</i> [authentication-mode { md5 sha } <i>auth-password</i>] [privacy-mode des56 <i>priv-password</i>] [acl <i>acl-number</i>]

SNMP community is one of the features of SNMP v1 and SNMP v2, so you import the ACL into the commands with SNMP community configured, for the SNMP V1 and SNMP V2.

SNMP username or group name is one of the features of SNMP V2 and above, therefore you import the ACL into the commands with SNMP username or group name configured, for the SNMP V2 and above. If you import the ACL into both features, the switch will filter both features for the users.

Note:

You can import different ACLs in the three commands listed above.

See the Command Manual for details about these commands.

Note:

Currently you can import only the basic ACLs with digit IDs.

34.3.3 Configuration Example

I. Network requirements

Only SNMP users from 10.110.100.52 and 10.110.100.46 can access the switch.

II. Network diagram

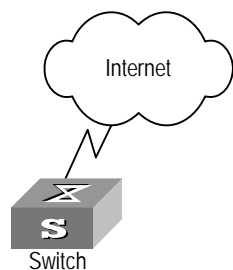


Figure 34-2 ACL configuration for SNMP users

III. Configuration procedure

Define a basic ACL.

```
[SW8800] acl number 2000 match-order config
[SW8800-acl-baisc-2000] rule 1 permit source 10.110.100.52 0
[SW8800-acl-baisc-2000] rule 2 permit source 10.110.100.46 0
[SW8800-acl-basic-2000] rule 3 deny source any
[SW8800-acl-baisc-2000] quit
```

Import the ACL.

```
[SW8800] snmp-agent community read 3Com acl 2000
[SW8800] snmp-agent group v3 3Comgroup acl 2000
[SW8800] snmp-agent usm-user v3 3Comuser 3Comgroup acl 2000
```

Chapter 35 MPLS Architecture

Note:

A Switch 8800 running MPLS can serve as a router. Routers mentioned in this manual can be either a router in common sense, or a layer 3 Ethernet switch running MPLS. To enable MPLS function on the Switch 8800, you must select the interface cards that support MPLS. Note that all the B cards do not support MPLS. For example, FT48C card supports MPLS, but FT48B does not.

35.1 MPLS Overview

MPLS (Multiprotocol Label Switching) encapsulates network layer packets with short and fixed-length labels. As the name implies, it supports multiple protocols, such as IP, IPv6, and IPX. And it allows a device to make forwarding decision based on the labels attached to the received packets without going through the complex routing table lookup procedures with IP. MPLS brings together the advantages of the connectionless control with IP and the connection-oriented forwarding with ATM. In addition to the support from IP routing and control protocols, its powerful and flexible routing functions allows it to accommodate to various emerging applications.

MPLS was initially proposed to accelerate the packet forwarding on routers, but it has been widely used in Traffic Engineering (TE), Virtual Private Network (VPN), and other aspects, and is becoming one of the most important standards on large scale IP networks.

35.2 MPLS Basic Concepts

35.2.1 FEC

Forwarding Equivalence Class (FEC) is an important concept in MPLS. MPLS is actually a kind of classify-and-forward technology. It categorizes packets with the same forwarding strategy (same destination addresses, same forwarding routes and same QoS levels) into one class, which is called a FEC. Generally, the FEC classification is based on network layer address. Packets of the same FEC are processed in the same way in MPLS network.

35.2.2 Label

I. Label definition

A label is a locally significant short identifier with fixed length, which is used to identify a FEC. When reaching at MPLS network ingress, packets are divided into different FECs, based on their FECs, different labels are encapsulated into the packets. Later forwarding is based on these labels.

II. Label structure

The structure of the label is shown in Figure 35-1.

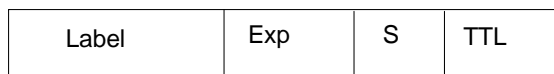


Figure 35-1 Label structure

Label is located between the link layer header and the network layer packet, with the length of four bytes. A label contains four fields:

Label: label value, 20 bits.

Exp: three bits, reserved, used for COS.

S: one bit, MPLS supports hierarchical label structure, namely multi-layer label. Value 1 refers to the label of bottom layer.

TTL: eight bits, with the same meaning as TTL in IP packet.

III. Label operations

1) Label mapping

There are two types of label mapping: label mapping at ingress routers, and label mapping in MPLS domain.

The first type of mapping is implemented at ingress label switching routers (LSR). The ingress LSRs group the incoming packets into multiple FECs based on certain principles, and then map corresponding labels to these FECs and record the mapping results into the label information base (LIB). In simple words, label mapping is to assign a label to a FEC.

The second type is also called incoming label mapping (ILM), that is, to map each input label to a series of next hop label forwarding entries (NHLFE). The packets are forwarded along the paths based on the mapping results.

2) Label encapsulation

Figure 35-2 illustrates label encapsulation in different media:

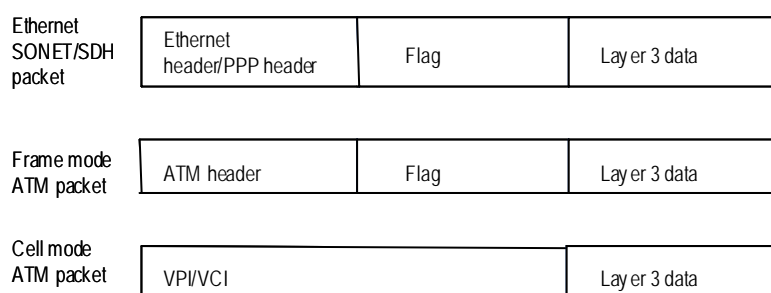


Figure 35-2 Label position in packet

In Ethernet packets and PPP packets, label stack lies between layer 2 header and layer 3 data, acting like a shim. In ATM cell mode packets, VPI/VCI is used as the label.

3) Label assignment and distribution

Label distribution refers to the process of creating a corresponding label switching path (LSP) for a FEC.

In the MPLS architecture, the decision to bind a particular label to a particular FEC is made by downstream LSR; after making the decision, the downstream LSR notifies the upstream LSR. That is to say, the label is assigned by the downstream LSR, and the assigned label is distributed from downstream to upstream.

Two label distribution modes are available in MPLS: downstream unsolicited (DU) mode and downstream on demand (DoD) mode.

- For a specific FEC, if LSR originates label assignment and distribution even without receiving label request message from upstream, it is in DU mode.
- For a specific FEC, if LSR begins label assignment and distribution only after receiving label request message from upstream, it is in DoD mode.

The upstream and downstream which have adjacency relation in label distribution should reach agreement on label distribution mode.

To distribute labels to its peer, the LSR can use Label Distribution Protocol (LDP) messages or make the labels carried on other routing protocol messages.

Note:

Upstream and downstream are just on a relative basis: For a packet forwarding process, the transmit router serves as upstream LSR and receive router serves as downstream LSR. Currently, the Switch 8800 adopts the DU label distribution mode.

4) Label assignment control mode

There are two modes to control the assignment and distribution of labels: independent mode and ordered mode.

In independent control mode, each LSR can send label mapping messages to the LSRs it connects to at anytime.

In ordered control mode, a LSR can send label mapping messages to upstream only when it receives a specific label mapping messages of the next hop of a FEC or the LSR serves as LSP (Label Switching Path) egress node.

Note:

Currently, the Switch 8800 adopts the ordered label control mode.

5) Label retention mode

There are two label-retention modes: liberal label retention mode and conservative label retention mode.

Suppose there are two LSRs: Ru and Rd. For a specific FEC, if LSR Ru has received the label binding from LSR Rd, in case Rd is not the next hop of Ru and Ru saves this binding, then it is the liberal label retention. And if Ru discards this binding, then it is the conservative label retention mode.

In case it is required that LSR is capable of adapting route variation rapidly, you can use the liberal label retention mode. In case it is required that a few labels are saved in LSR, you can use the conservative label retention mode.

Note:

Currently, the Switch 8800 adopts the liberal label retention mode.

35.2.3 LDP

Label distribution protocol (LDP) is the signaling control protocol in MPLS, which controls binding labels and FECs between LSRs and coordinates a series of procedures between LSRs.

35.3 MPLS Architecture

35.3.1 MPLS Network Structure

The basic composing unit of MPLS network is LSR (Label Switching Router). It runs MPLS control protocol and L3 routing protocol, exchanges routing messages with other LSRs and create the routing table, maps FECs with IP packet headers, binds FECs

with labels, distributes label binding messages, establishes and maintains label forwarding table.

The network consisting of LSRs is called MPLS domain. The LSR that is located at the edge of the domain is called edge LSR (LER, Labeled Edge Router). It connects an MPLS domain with a non-MPLS domain or with another MPLS domain, classifies packets, distributes labels (as ingress LER) and distributes labels (as egress LER). The ingress LER is termed as ingress and egress LER as egress.

The LSR that is located inside the domain is called core LSR, which provides functions such as label swapping and label distribution. The labeled packets are transmitted along the LSP (Label Switched Path) composed of a series of LSRs.

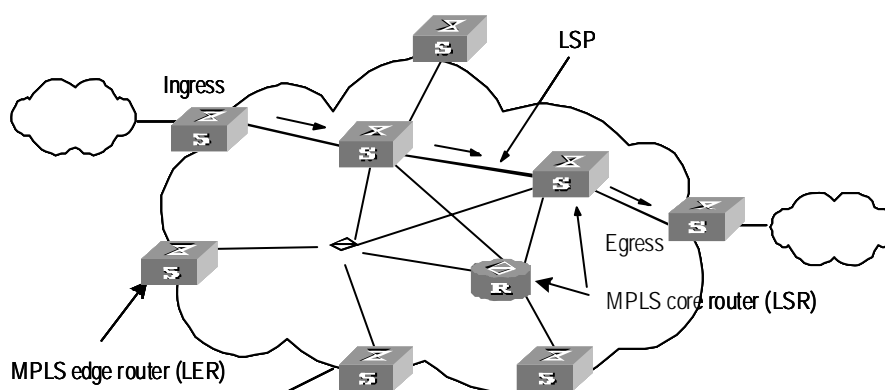


Figure 35-3 MPLS basic principle

35.3.2 Forwarding Labeled Packets

At the ingress, the packets entering the network are classified into FECs according to their characteristics. Usually, packets are classified into FECs according to the IP address prefix or host address. Packets in the same FEC pass through the same path (that is, LSP) in MPLS area. LSR assigns a short label of fixed length for the incoming FEC packet, and then forwards it through the corresponding interface.

On the LSR along the LSP, the mapping table of the import/export labels has been established (the element of this table is referred to as Next Hop Label Forwarding Entry (NHLFE)). When the labeled packet arrives, LSR only needs to find the corresponding NHLFE from the table according to the label and replace the original label with a new one, and then forwards the labeled packet. This process is called Incoming Label Map (ILM).

At the ingress, MPLS specifies a FEC for a specific packet, and the following routers only need to forward the packet by label switching, therefore this method is much simpler than general network layer forwarding and increases the forwarding speed.

35.3.3 Establishing LSP

Actually, the establishment of LSP refers to the process of binding FEC with the label, and then advertising this binding to the adjacent LSR on LSP. This process is implemented through LDP, which regulates the message in interactive processing and message structure between LSRs as well as routing mode.

I. LDP working process

Through sending Hello message periodically, an LSR finds its neighbor and then establish LDP session with the newly discovered adjacent LSR. By LDP session, the adjacent LSRs advertise such information as label switching mode, label space, session keepalive timer value to each other. LDP session is a TCP connection, which needs to be maintained through LDP message. In case there is not any other LDP message during the time period specified by the session keepalive timer value, and then it is necessary to send session keepalive message to maintain the existence of LDP session. Figure 35-4 illustrates the diagram of LDP label distribution.

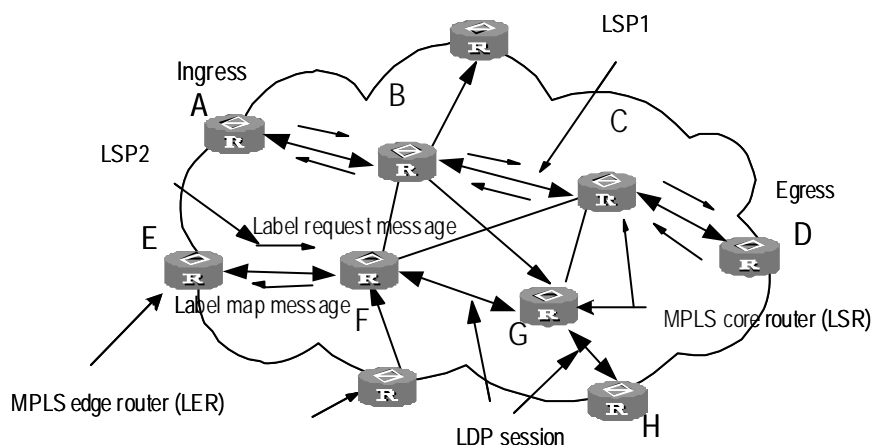


Figure 35-4 Label distribution process

For the label distribution mentioned previously, there are two modes: DoD and DU. The main difference between these two modes is that the label mapping messages are distributed actively or passively.

In DoD mode, the label is distributed in this way: the upstream LSR sends label request message (containing FEC descriptive information) to the downstream LSR, and the downstream LSR distributes label for this FEC, and then it sends the bound label back to the upstream LSR through label map message. The time when the downstream LSR feeds back the label map message depends on whether this LSR uses independent label control mode or sequential label control mode. When the sequential label control mode is used by the downstream LSR, the label map message is sent back to its upstream LSR if only it has received the label map message from its downstream LSR. And when the independent label control mode is used by the downstream LSR, then it will send label map message to its upstream LSR immediately, no matter whether it has

received the returned label map message from its downstream LSR. Usually, the upstream LSR selects the downstream LSR according to the information in its routing table. In Figure 35-4, LSRs on the way along LSP1 use the sequential label control mode, and the LSR F on LSP2 uses independent label control mode.

In DU mode, the label is distributed in the following way: when LDP session is established successfully, the downstream LSR will actively distribute label map message to its upstream LSR. And the upstream LSR saves the label map information and processes the received label map information according to the routing table.

II. LSP loop control

While establishing LSP in MPLS domain, it is also necessary to prevent the presence of path loop. Then, such two methods as maximum hop count and path vector can be used.

The maximum hop count method refers to that the hop-count information is contained in the message bound with the forwarding label, and the value pluses one for each hop. When the value exceeds the threshold value, it is considered that a loop presents, and the process for establishing LSP is terminated.

The path vector method refers to that the path information is recorded in the message bound with the forwarding label, and, for every hop, the corresponding router checks if its ID is contained in this record. If not, the router adds its ID into the record; and if yes, it indicates that a loop presents and the process for establishing LSP is terminated.

35.3.4 LSP Tunnel and Hierarchy

I. LSP tunnel

MPLS supports LSP tunnel technology. On an LSP path, LSR Ru and LSR Rd are both the upstream and the downstream for each other. However, the path between LSR Ru and LSR Rd may not be part of the path provided by routing protocol. MPLS allows establishing a new LSP path <Ru R1...Rn Rd> between LSR Ru and LSR Rd, and LSR Ru and LSR Rd are respectively the starting point and ending point of this LSP. The LSP between LSR Ru and LSR Rd is referred to as the LSP tunnel, which avoids the traditional encapsulated tunnel on the network layer. If the route along which the tunnel passes and the route obtained hop by hop from routing protocol is consistent, this tunnel is called hop-by-hop routing tunnel. And if the two routes are not consistent, then the tunnel of this type is called explicit routing tunnel.

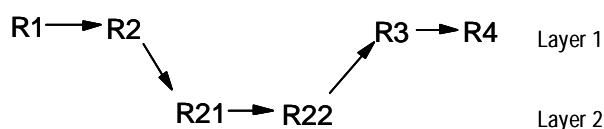


Figure 35-5 LSP tunnel

As shown in Figure 35-5, LSP <R2 R21 R22 R3> is a tunnel between R2 and R3.

II. Multi-layer label stack

In MPLS, a packet may carry multiple labels which are in the form of stack. Operations to the stack follow the “last in first out” principle and it is always the labels at the top of the stack that decide how to forward packets. Pushing label indicates to add a label into a outgoing packet, then the depth of the label stack is the former one plus 1, and the current label of the packet changes to the newly added one; popping a label indicates to remove a label from a packet, then the depth of the packet is the former one minus 1, and the current label of the packet changes to the label of its underlayer.

Multiple-layer label stack is used in LSP tunnel. When a packet travels in LSP tunnel, there will be multiple layers for the label of the packet. Then, at the ingress and egress of each tunnel, it is necessary to implement pushing and popping operation for the label stack. For each pushing operation, the label will be added with one layer. And there is no depth limitation for the label stack from MPLS.

The labels are organized according to the principle of “last in first out” in the label stack, and MPLS processes the labels beginning from the top of the stack.

If the depth of the label stack for a packet is m , it indicates that the label at the bottom of that stack is level 1 label, and the label at the top of the stack is level m label. A packet with no label can be regarded as a packet with empty label stack, that is, the depth of its label stack is 0.

35.4 MPLS and other Protocols

35.4.1 MPLS and Routing Protocols

When LDP establishes LSP in hop-by-hop mode, the next hop is determined by using the information, which is usually collected by such routing protocols as IGP, BGP in each LSR route forwarding table, on the way. However, LDP just uses the routing information indirectly, rather than associates with various routing protocols directly.

On the other hand, although LDP is the special protocol for implementing label distribution, it is not the sole protocol for label distribution. The existing protocols such as BGP, RSVP, after being extended, can also support MPLS label distribution. For some MPLS applications, it is also necessary to extend some routing protocols. For example, the application of MPLS VPN requires extending the BGP protocol, thus the BGP protocol can propagate VPN routing information.

35.5 MPLS Application

35.5.1 MPLS VPN

To transmit data stream of private network on public network, traditional VPN uses tunnel protocols like GRE, L2TP, and PPTP. LSP itself is a tunnel on public network, so

there are obvious advantages to implement VPN by MPLS. MPLS VPN connects the geographically different branches of private network by using LSP, forming a united network. MPLS VPN also supports the interconnection between different VPNs.

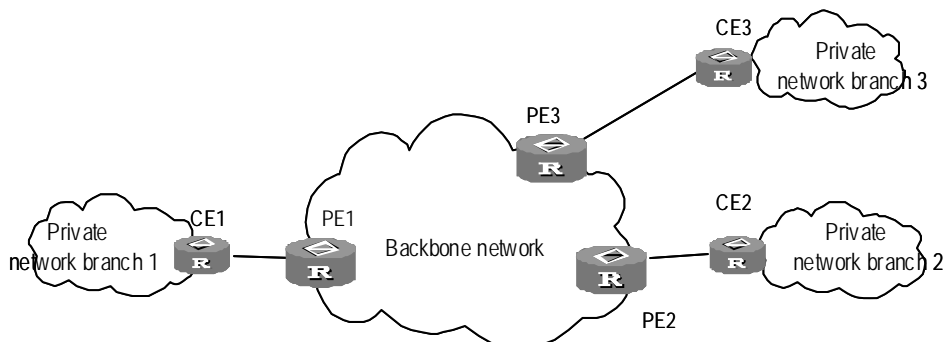


Figure 35-6 MPLS-Based VPN

The basic structure of MPLS-based VPN is shown in Figure 35-6. CE is the customer edge device, and it may either be a router or a switch, or perhaps a host. PE is a service provider edge router, which is located on the backbone network. PE is responsible for the management of VPN customers, establishing LSP connection between various PEs, route allocation among different branches of the same VPN customer.

Usually the route allocation between PEs is realized by using extended BGP. MPLS VPN supports the IP address multiplexing between different branches and the interconnection between different VPNs. Compared with traditional route, it is necessary to add branch and VPN identifier information in VPN route. So, it is necessary to extend BGP so as to carry VPN routing information.

Chapter 36 MPLS Basic Capability Configuration

36.1 MPLS Basic Capability Overview

Basic MPLS forwarding functions includes LDP session establishment and LSP path maintenance.

The typical configuration procedure for enabling basic MPLS functions on a routing switch is as follows:

- 1) Configure LSR ID
- 2) Enable MPLS
- 3) Enable LDP
- 4) Enter VLAN interface view and enable MPLS and LDP on the interface

Then the routing switch can provide MPLS forwarding and LDP signaling functions.

If you want to modify the default parameters or enable some special functions, for example, manually creating LSP or explicit route, you can configure according to the methods in configuration list. For some complicated functions, configuration combination may be required.

36.2 MPLS Configuration

The following sections describe the required configuration tasks for MPLS basic capability:

- Defining MPLS LSR ID
- Enabling MPLS and Entering MPLS View

The following sections describe the optional configuration tasks for MPLS basic capability:

- Configuring the Topology-Driven LSP Setup Policy
- Configuring Static LSP

36.2.1 Defining MPLS LSR ID

Before configuring any other MPLS command, it is necessary to configure LSR ID firstly. This ID is usually in IP address format and must be unique in the domain.

Perform the following configuration in the system view.

Table 36-1 Define MPLS LSR ID

Operation	Command
Define LSR ID	mpls lsr-id <i>ip-address</i>
Delete LSR ID	undo mpls lsr-id

By default, LSR ID is not defined.

36.2.2 Enabling MPLS and Entering MPLS View

In system view, you can first enable MPLS globally and enter MPLS view using the **mpls** command. Then you can directly enter MPLS view after using the **mpls** command in system view.

Use the **mpls** command in VLAN interface view to enable MPLS on the VLAN interface.

Table 36-2 Enter MPLS view

Operation	Command
Enable MPLS globally and enter MPLS view (system view) Enable MPLS on a VLAN interface (VLAN interface view)	mpls
Disable MPLS globally or on a VLAN interface (system or VLAN interface view)	undo mpls

By default, MPLS is not enabled.

36.2.3 Configuring the Topology-Driven LSP Setup Policy

It refers to specifying filtering policy as all or ip-prefix.

Perform the following configuration in MPLS view.

Table 36-3 Configure the topology-driven LSP setup policy

Operation	Command
Configure the topology-Driven LSP setup policy	lsp-trigger { all ip-prefix <i>ip-prefix</i> }
Use the default value, which only allows 32-bit IP to trigger ISP.	undo lsp-trigger { all ip-prefix <i>ip-prefix</i> }

36.2.4 Configuring Static LSP

You can manually set an LSR to be a node along an LSP, and place a limit on the traffic over the LSP. Depending on the position in an MPLS domain, an LSR along an LSP can

be the ingress node, an intermediate node (also called transit node), or the egress node. Note that an LSP operates normally only after all the LSRs along the LSP have been properly configured.

The **undo static-lsp** command is used to delete a specified LSP established manually.

Perform the following configuration in MPLS view.

Table 36-4 Set the local LSR to a node on a specified LSP

Operation	Command
Set the current LSR to the ingress node of the specified LSP	static-lsp ingress <i>lsp-name</i> { destination <i>dest-addr</i> { <i>addr-mask</i> <i>mask-length</i> } l2vpn } nexthop <i>next-hop-addr</i> } } out-label <i>out-label-value</i>
Cancel the ingress node setting of the specified LSP	undo static-lsp ingress <i>lsp-name</i>
Set the current LSR to an intermediate node along the specified LSP	static-lsp transit <i>lsp-name</i> [l2vpn] incoming-interface <i>interface-type</i> <i>interface-num</i> in-label <i>in-label-value</i> nexthop <i>next-hop-addr</i> out-label <i>out-label-value</i>
Cancel the intermediate node setting of the specified LSP	undo static-lsp transit <i>lsp-name</i>
Set the current LSR to the egress node of the specified LSP	static-lsp egress <i>lsp-name</i> [l2vpn] incoming-interface <i>interface-type</i> <i>interface-num</i> in-label <i>in-label-value</i>
Cancel the egress node setting of the specified LSP	undo static-lsp egress <i>lsp-name</i>

36.3 LDP Configuration

The following sections describe the required LDP configuration tasks for MPLS basic capability:

- Enabling LDP protocol
- Enabling LDP on VLAN interface

The following sections describe the optional LDP configuration tasks for MPLS basic capability:

- Configuring Remote-Peer for Extended Discovery Mode
- Configuring session parameters

36.3.1 Enabling LDP protocol

To configure LDP, first enable LDP.

Perform the following configuration in the system view.

Table 36-5 Enable/disable LDP view

Operation	Command
Enable LDP protocol	mpls ldp
Disable LDP	undo mpls ldp

By default, LDP is disabled.

36.3.2 Enabling LDP on VLAN interface

To make the VLAN interface support LDP, you must enable LDP function on virtual interface in VLAN interface mode. After enabling the LDP function, the virtual interface then sets up session. It begins to set up LSP if in topology-driven mode,.

Disabling LDP function on interface causes the break of all LDP session in VLAN interface, and all the LSP based on those sessions are deleted. So you must use this command with cautiously.

Perform the following configuration in the interface view.

Table 36-6 Enable/disable LDP on interface

Operation	Command
Enable LDP function on interface	mpls ldp enable
Disable LDP function on interface	mpls ldp disable

By default, the interface LDP function is disabled.

36.3.3 Configuring Remote-Peer for Extended Discovery Mode

The remote-peer configuration is mainly used for extended discovery mode so that this LSR can establish sessions with LSRs that are not directly connected with it at the link layer.

I. Create a remote-peer

Perform the following configuration in the system view.

Table 36-7 Create a remote-peer

Operation	Command
Create a remote-peer and enter remote-peer view	mpls ldp remote-peer <i>index</i>
Delete the corresponding remote-peer	undo mpls ldp remote-peer <i>index</i>

There is no default remote-peer.

II. Configuring an address for the remote-peer

You can specify the address of any LDP-enabled interface on the remote-peer or the address of the loopback interface on the LSR that has advertised the route as the address of the remote-peer.

Perform the following configuration in the remote-peer view.

Table 36-8 Configure a remote-peer address

Operation	Command
Configure a remote-peer address	remote-ip <i>remoteip</i>

remoteip: the IP address of remote peer. It should be the LSR ID of the peer LSR.

There is no default remote-peer.

36.3.4 Configuring session parameters

I. Configuring session hold-time

The LDP entity on the interface sends Hello packets periodically to find out LDP peer, and the established sessions must also maintain their existence by periodic message (if there is no LDP message, then Keepalive message must be sent).

Note:

There are two types of LDP sessions: basic and remote. Basic session can be established only on two direct-connect switches, while remote session can be on two switches which are not directly connected. You can only configure basic sessions in VLAN interface view and remote sessions in remote-peer view.

Caution:

Modifying the *holdtime* parameter results in re-establish the original session, as well as the LSP over this session. Here the session refers to basic session, but not remote session.

Perform the following configuration in VLAN interface view.

Table 36-9 Configure basic session hold-time

Operation	Command
Configure session hold-time	mpls ldp timer { session-hold session-holdtime hello hello-holdtime }
Return to the default value	undo mpls ldp timer { session-hold hello }

By default, the *session-holdtime* is 60 seconds and *hello-holdtime* is 15 seconds.

Perform the following configuration in remote-peer view.

Table 36-10 Configure remote session hold-time

Operation	Command
Configure session hold-time	mpls ldp timer { targeted-session-hold targeted-hello } { holdtime interval }
Return to the default value	undo mpls ldp timer { targeted-session-hold targeted-hello }

By default, **targeted-session-hold** *holdtime* is 60 seconds, and the interval is 24 seconds; **targeted-hello** *holdtime* is 45 seconds and the interval is 13 seconds.

II. Configuring hello transport-address

The transport-address discussed here refers to the address carried in the transport address TLV in hello messages. Generally, you can configure the transport-address to the MPLS LSR ID of the current LSR, but you can also configure the transport-address to other address flexibly as required by some applications.

Perform the following configuration in VLAN interface view.

Table 36-11 Configure hello transport-address

Operation	Command
Configure hello transport-address	mpls ldp transport-ip { interface ip-address }
Return to the default hello transport-address	undo mpls ldp transport-ip

Transport-address defaults to the MPLS LSR ID of the current LSR.

If there are multiple links connecting two neighboring LSRs, all the LDP-enabled interfaces on the links connecting LSR and its neighbor must have the same transport address. You are recommended to use the same interface address for all of them, that is, LSR-ID.

36.3.5 Configuring LDP Loop Detection Control

I. Enabling loop detection

It is used to enable or disable the loop detection function during LDP signaling process. The loop detection includes maximum hop count mode and path vector mode.

The maximum hop count method refers to that the hop-count information is contained in the message bound with the forwarding label, and the value pluses one for each hop. When the value exceeds the threshold value, it is considered that a loop presents, and the process for establishing LSP is terminated.

The path vector method refers to that the path information is recorded in the message bound with the forwarding label, and, for every hop, the corresponding router checks if its ID is contained in this record. If not, the router adds its ID into the record; and if yes, it indicates that a loop presents and the process for establishing LSP is terminated. When this method is used, if the defined maximum value is exceeded, it is considered that a loop happens and the LSP establishment fails.

Perform the following configuration in the system view.

Table 36-12 Enable loop detection

Operation	Command
Enable loop detection	mpls ldp loop-detect
Disable loop detection	undo mpls ldp loop-detect

By default, the loop detection is disabled.

II. Setting the maximum hop count for loop detection

When maximum hop count mode is adopted for loop detection, the maximum hop-count value can be defined. And if the maximum value is exceeded, it is considered that a loop happens and the LSP establishment fails.

Perform the following configuration in the system view.

Table 36-13 Set the maximum hop count for loop detection

Operation	Command
Set maximum hop count for loop detection	mpls ldp hops-count <i>hop-number</i>
Return to the default maximum hop count	undo mpls ldp hops-count

The maximum hop count defaults to 32.

III. Setting the maximum hop count in path vector mode

When path vector mode is adopted for loop detection, it is also necessary to specify the maximum value of LSP path. In this way, when one of the following conditions is met, it is considered that a loop happens and the LSP establishment fails.

- The record of this LSR already exists in the path vector recording table.
- The path hop count exceeds this maximum value.

Perform the following configuration in the system view.

Table 36-14 Set the maximum hop count in path vector mode

Operation	Command
Set the maximum hop count in path vector mode	mpls ldp path-vectors <i>pv-number</i>
Return to the default maximum hop count in path vector mode	undo mpls ldp path-vectors

The maximum hop count defaults to 32.

36.3.6 Configuring LDP Authentication Mode Between Every Two Routers

Perform the following configuration in VLAN interface view or remote-peer view.

Table 36-15 Configure LDP authentication mode

Operation	Command
Configure LDP authentication Mode	mpls ldp password [cipher simple] <i>password</i>
Remove LDP authentication	undo mpls ldp password

36.4 Displaying and Debugging MPLS

36.4.1 Displaying and Debugging MPLS

MPLS provides abundant display and debugging commands for monitoring LDP session state, tunnel, all the LSPs and their states, and so on. These commands are the powerful debugging and diagnosing tools.

I. Displaying static LSPs

After accomplishing the configuration tasks mentioned previously, you can execute the **display** command in any view to view the running state of a single or all the static LSPs and thus to evaluate the effect of the configurations.

Table 36-16 Display the static LSP information

Operation	Command
Display the static LSP information	display mpls static-lsp [include <i>text</i> verbose]

II. Displaying MPLS-enabled interfaces

After accomplishing the configuration tasks mentioned previously, you can execute the **display** command in any view to view the information related to the MPLS-enabled interfaces and thus to evaluate the effect of the configurations.

Table 36-17 Display information of the MPLS-enabled interfaces

Operation	Command
Display information of the MPLS-enabled interfaces	display mpls interface

III. Displaying LSP

Execute the following commands in any view to display the information related to MPLS LSP.

Table 36-18 Display the information about MPLS LSP

Operation	Command
Display the information about MPLS LSP	display mpls lsp [include <i>text</i> verbose]

IV. Debugging MPLS

You may execute the **debugging** command in user view to debug the information concerning all interfaces with MPLS function enabled.

As enabling debugging may affect the router performance, you are recommended to use this command when necessary. Execute the **undo** form of this command to disable the corresponding debugging.

Table 36-19 Enable/disable debugging for MPLS

Operation	Command
Enable debugging for MPLS LSP	debugging mpls lspm { agent all event ftn interface packet policy process vpn }
Disable debugging for MPLS LSP	undo debugging mpls lspm { agent all event ftn interface packet policy process vpn }

V. Trapping MPLS

This command is used to enable the trap function of MPLS during an LSP/LDP setup process.

Perform the following configuration in system view.

Table 36-20 Enable the trap function of MPLS

Operation	Command
Enable the LDP trap function of MPLS	snmp-agent trap enable ldp
Disable the LDP trap function of MPLS	undo snmp-agent trap enable ldp
Enable the LSP trap function of MPLS	snmp-agent trap enable lsp
Disable the LSP trap function of MPLS	undo snmp-agent trap enable lsp

36.4.2 Displaying and Debugging LDP

I. LDP display commands

VRP provides abundant MPLS monitoring commands for monitoring states of LSRs, LDP sessions, interfaces and peers. These commands are the powerful debugging and diagnosing tools.

After accomplishing the configuration tasks described earlier, you can execute the **display** command in any view to view the running state of LDP and thus to evaluate the effect of the configurations.

Table 36-21 Display LDP

Operation	Command
Display LDP information	display mpls ldp
Display buffer information for LDP	display mpls ldp buffer-info
Display LDP-enabled interface information	display mpls ldp interface
Display LDP saved label information	display mpls ldp lsp
Display information on all peers of LDP session	display mpls ldp peer
Display information of the remote-peers in the LDP sessions	display mpls ldp remote
Display states and parameters of LDP sessions	display mpls ldp session

II. LDP debugging commands

Execute **debugging** command in user view for the debugging of various messages related to LDP

Table 36-22 Enable/disable debugging for MPLS LDP

Operation	Command
Enable debugging for MPLS LDP	debugging mpls ldp { all main advertisement session pdu notification remote } [interface <i>interface-type interface-num</i>]
Disable debugging for MPLS LDP	undo mpls debugging ldp { all main advertisement session pdu notification remote } [interface <i>interface-type interface-num</i>]

all: Displays all LDP-related debugging information

main: Displays debugging information about LDP main tasks

advertisement: Displays debugging information in processing LDP advertisements

session: Displays debugging information in processing LDP session

pdu: Displays debugging information in processing PDU packets

notification: Displays debugging information in processing notifications

remote: Displays debugging information about all remote peers

interface-type Interface-num: Port type and port ID.

Use the **mpls ldp reset-session** command in VLAN interface to reset a specific LDP session on the VLAN interface.

Table 36-23 Reset LDP

Operation	Command
Reset a specific LDP session on the VLAN interface (VLAN interface view)	mpls ldp reset-session <i>peer-address</i>

36.5 Typical MPLS Configuration Example

I. Network requirements

Figure 36-1 illustrates a network with four switches, which connects to each other through Ethernet.

The four switches all support MPLS, and LSP can be established between any two switches with the routing protocol OSPF. LDP establishes LSP by using routing information of OSPF.

II. Network diagram

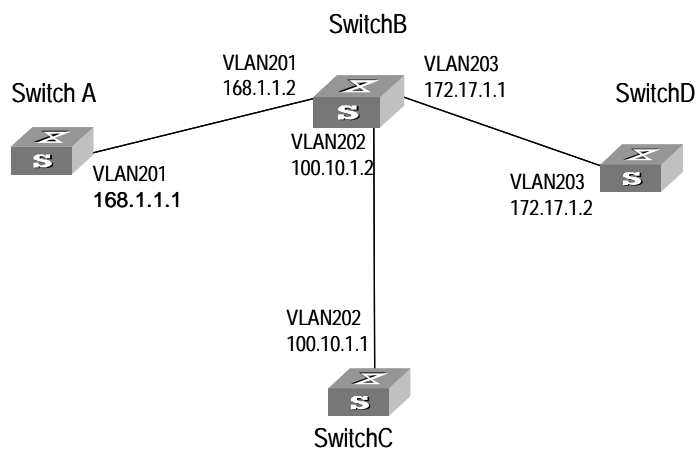


Figure 36-1 Network diagram

III. Configuration procedure

1) Configure Switch A

Configure LSR ID and enable MPLS and LDP.

```
[SW8800] mpls lsr-id 168.1.1.1
[SW8800] mpls
[SW8800-mpls] quit
[SW8800] mpls ldp
```

Configure IP address and enable MPLS and LDP for VLAN interface 201.

```
[SW8800] vlan 201
[SW8800-vlan201] port gigabitethernet 2/1/1
[SW8800-vlan201] quit
[SW8800] interface Vlan-interface 201
[SW8800-Vlan-interface201] ip address 168.1.1.1 255.255.0.0
[SW8800-Vlan-interface201] mpls
[SW8800-Vlan-interface201] mpls ldp enable
[SW8800-Vlan-interface201] mpls ldp transport-ip interface
```

Enable OSPF on the interface connecting Switch A with Switch B.

```
[SW8800] Router id 168.1.1.1
[SW8800] ospf
[SW8800-ospf-1] area 0
[SW8800-ospf-1-area-0.0.0.0] network 168.1.0.0 0.0.255.255
```

2) Configure Switch B

Configure LSR ID and enable MPLS and LDP.

```
[SW8800] mpls lsr-id 172.17.1.1
```

```
[SW8800] mpls
[SW8800-mpls] quit
[SW8800] mpls ldp
```

Configure IP address and enable MPLS and LDP for VLAN interface 201.

```
[SW8800] vlan 201
[SW8800-vlan201] port gigabitethernet 2/1/1
[SW8800-vlan201] quit
[SW8800] interface Vlan-Interface 201
[SW8800-Vlan-interface201] ip address 168.1.1.2 255.255.0.0
[SW8800-Vlan-interface201] mpls
[SW8800-Vlan-interface201] mpls ldp enable
[SW8800-Vlan-interface201] mpls ldp transport-ip interface
```

Configure IP address and enable MPLS and LDP for VLAN interface 203.

```
[SW8800] vlan 203
[SW8800-vlan203] port gigabitethernet 2/1/3
[SW8800-vlan203] quit
[SW8800] interface Vlan-Interface 203
[SW8800-Vlan-interface203] ip address 172.17.1.1 255.255.0.0
[SW8800-Vlan-interface203] mpls
[SW8800-Vlan-interface203] mpls ldp enable
[SW8800-Vlan-interface203] mpls ldp transport-ip interface
```

Configure IP address and enable MPLS and LDP for VLAN interface 202.

```
[SW8800] vlan 202
[SW8800-vlan202] port gigabitethernet 2/1/2
[SW8800-vlan202] quit
[SW8800] interface Vlan-interface 202
[SW8800-Vlan-interface202] ip address 100.10.1.2 255.255.255.0
[SW8800-Vlan-interface202] mpls
[SW8800-Vlan-interface202] mpls ldp enable
[SW8800-Vlan-interface202] mpls ldp transport-ip interface
[SW8800-Vlan-interface202] quit
```

Enable OSPF on the interfaces respectively connecting Switch B with Switch A, Switch D and Switch C.

```
[SW8800] Router id 172.17.1.1
[SW8800] ospf
[SW8800-ospf-1] area 0
[SW8800-ospf-1-area-0.0.0.0] network 168.1.0.0 0.0.255.255
[SW8800-ospf-1-area-0.0.0.0] network 172.17.0.0 0.0.255.255
[SW8800-ospf-1-area-0.0.0.0] network 100.10.1.0 0.0.0.255
[SW8800-ospf-1-area-0.0.0.0] quit
```

3) Configure Switch C

Configure LSR ID and enable MPLS and LDP.

```
[SW8800] mpls lsr-id 100.10.1.1
[SW8800] mpls
[SW8800-mpls] quit
[SW8800] mpls ldp
```

Configure IP address and enable LDP and MPLS for VLAN interface 202.

```
[SW8800] vlan 202
[SW8800-vlan202] port gigabitethernet 2/1/1
[SW8800-vlan202] quit
[SW8800] interface Vlan-interface 202
[SW8800-Vlan-interface202] ip address 100.10.1.1 255.255.255.0
[SW8800-Vlan-interface202] mpls
[SW8800-Vlan-interface202] mpls ldp enable
[SW8800-Vlan-interface202] quit
```

Enable OSPF on the interface connecting Switch C with Switch B.

```
[SW8800] Router id 100.10.1.1
[SW8800] ospf
[SW8800-ospf-1] area 0
[SW8800-ospf-1-area-0.0.0.0] network 100.10.1.0 0.0.0.255
```

4) Configure Switch D**Configure LSR ID and enable MPLS and LDP.**

```
[SW8800] mpls lsr-id 172.17.1.2
[SW8800] mpls
[SW8800-mpls] quit
[SW8800] mpls ldp
```

Configure IP address and enable MPLS and LDP for VLAN interface 203.

```
[SW8800] vlan 203
[SW8800-vlan203] port gigabitethernet 2/1/3
[SW8800-vlan203] quit
[SW8800] interface vlan-interface 203
[SW8800-Vlan-interface203] ip address 172.17.1.2 255.255.0.0
[SW8800-Vlan-interface203] mpls
[SW8800-Vlan-interface203] mpls ldp enable
```

Enable OSPF on the interface connecting Switch D with Switch B.

```
[SW8800] Router id 172.17.1.2
[SW8800] ospf
[SW8800-ospf-1] area 0
[SW8800-ospf-1-area-0.0.0.0] network 172.17.0.0 0.0.255.255
```

36.6 Troubleshooting MPLS Configuration

Symptom: Session cannot be setup with the peer after LDP is enabled on the interface.

Troubleshooting:

Cause 1: Loop detection configuration is different at the two ends.

Solution: Check loop detection configuration at both ends to see if one end is configured while the other end is not (this will result in session negotiation failure).

Cause 2: Local machine cannot get the route to peer LSR ID, so TCP connection cannot be set up and session cannot be established.

Solution: The default address for session transfer is MPLS LSR ID. The local machine should issue the LSR ID route (often the loopback address) and learn the peer LSR ID route.

Chapter 37 BGP/MPLS VPN Configuration

37.1 BGP/MPLS VPN Overview

Traditional VPN, for which layer 2 tunneling protocols (L2TP, L2F and PPTP, and so on.) or layer 3 tunnel technology (IPSec, GRE and so on.) is adopted, is a great success and is therefore widely used. However, along with the increase of the size of VPNs, the deficiency of traditional VPN in such aspects as expansibility and manageability becomes more and more obvious. In addition, QoS (Quality of Service) and security are also the difficult problem for traditional VPN.

Using the MPLS technology, service providers can implement the IP-based VPN services easily and enable their networks to meet the expansibility and manageability requirement for VPN. The VPN constructed by using MPLS also provides the possibility for the implementation of value-added service. Multiple VPNs can be formed from a single access point, and each VPN represents a different service, making the network able to transmit services of different types in a flexible way.

Product currently provides comparatively complete BGP/MPLS VPN networking capabilities:

- Address isolation, allowing the overlap of address of different VPNs and public networks.
- Supporting MBGP advertising VPN routing information through public network, establishing BGP/MPLS VPN.
- Forwarding VPN data stream over MPLS LSP.
- Providing MPLS VPN performance monitoring and fault detecting tools.

37.1.1 BGP/MPLS VPN Model

I. BGP/MPLS VPN model

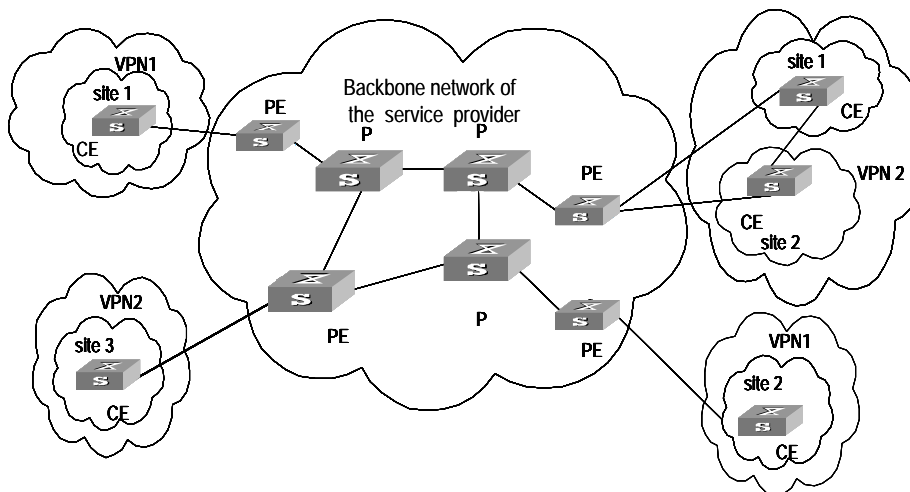


Figure 37-1 MPLS VPN model

As shown in Figure 37-1, MPLS VPN model contains three parts: CE, PE and P.

- CE (Customer Edge) device: It is a composing part of the customer network, which is usually connected with the service provider directly through an interface. It may be a router or a switch which cannot sense the existence of VPN.
- PE (Provider Edge) router: It is the Provider Edge router, namely the edge device of the provider network, which connects with your CE directly. In MPLS network, PE router processes all the operations for VPN. PE needs to possess MPLS basic forwarding capability.
- P (Provider) router: It is the backbone router in the provider network, which is not connected with CE directly. P router needs to possess MPLS basic forwarding capability.

The classification of CE and PE mainly depends on the range for the management of the provider and the customer, and CE and PE are the edges of the management ranges.

II. Nested BGP/MPLS VPN model

In a basic BGP/MPLS VPN model, the PEs are in the network of the service provider and are managed by the service provider.

When a VPN user wants to subdivide the VPN into multiple VPNs, the traditional solution is to configure these VPNs directly on the PEs of the service provider. This solution is easy to implement, but has the following disadvantages: the number of the VPNs carried on PEs may increase rapidly; the operator may have to perform more operations when required by a user to adjust the relation between the user's internal

VPNs. These disadvantages not only increase the network operating cost, but also bring relevant management and security issues.

The nested VPN is a better solution. Its main idea is to transfer VPNv4 route between PE and CE of common BGP MPLS/VPN such that user themselves can manage their internal VPN division, and the service provider can be saved from participating into users' internal VPN management.

The following figure shows the network model for nested VPN:

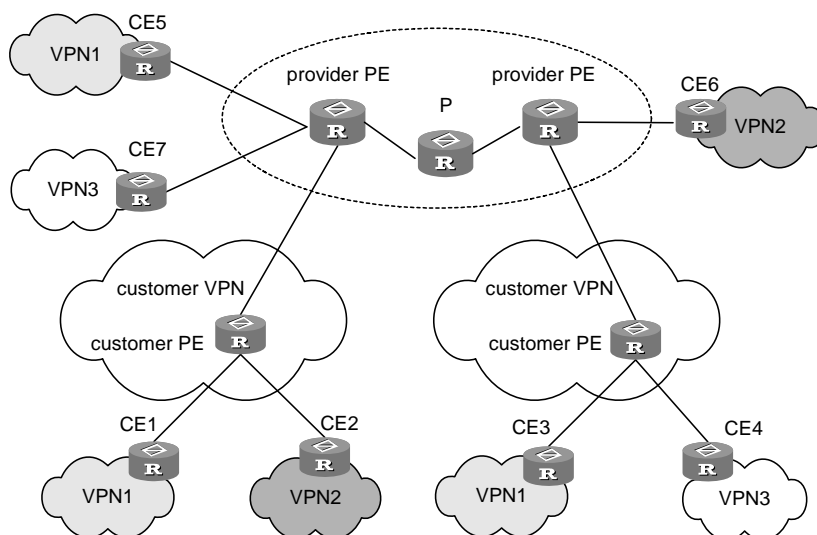


Figure 37-2 Network model for nested BGP/MPLS VPN

III. Basic concepts in BGP/MPLS VPN

1) vpn-instance

vpn-instance is an important concept in VPN routing in MPLS. In an MPLS VPN implementation, each site corresponds to a specific vpn-instance on PE (their association is implemented by binding vpn-instance to the VALN interface). If subscribers on one site belong to multiple VPNs, then the corresponding vpn-instance includes information about all these VPNs.

Specifically, such information should be included in vpn-instance: label forwarding table, IP routing table, the interfaces bound with vpn-instance, and the management information (RD, route filtering policy, member interface list, and so on). It includes the VPN membership and routing rules of this site.

PE is responsible for updating and maintaining the relationship between vpn-instance and VPN. To avoid data leakage from the VPN and illegal data entering into the VPN, each vpn-instance on the PE has an independent set of routing table and label forwarding table, in which the forwarding information of the message is saved

2) MBGP

MBGP (multiprotocol extensions for BGP-4, see RFC2283) propagates VPN membership information and routes between PE routers. It features backward

compatibility: It not only supports traditional IPv4 address family, but also supports other address families, for example, VPN-IPv4 address family. MP-BGP ensures that VPN private routes are only advertised within VPNs, as well as implementing communication between MPLS VPN members.

3) VPN-IPv4 address

VPN is just a private network, so it can use the same IP address to indicate different sites. But the IP address is supposed as unique when MP-BGP advertises CE routes between PE routers, so routing errors may occur for the different meaning in two systems. The solution is to switch IPv4 addresses to VPN-IPv4 address to generate globally unique addresses before advertising them, so PE routers is required to support MP-BGP.

A VPN-IPv4 address consists of 12 bytes, and the first eight bytes represent the RD (Route Distinguisher), which are followed by a 4-byte IPv4 address. The service providers can distribute RD independently. However, their special AS (Autonomous System) number must be taken as a part of the RD. After being processed in this way, even if the 4-byte IPv4 address contained in VPN-IPv4 address has been overlapped, the VPN-IPv4 address can still maintain globally unique. RD is only used within the carrier network to differentiate routes. When the RD is 0, a VPN-IPv4 address is just a IPv4 address in general sense.

The route received by PE from CE is the IPv4 route that needs to be redistributed into vpn-instance routing table, and in this case a RD needs to be added. It is recommended that the same RD be configured for all routes from the same user site.

IV. VPN Target attribute

VPN Target attribute is one of the MBGP extension community attributes and is used to limit VPN routing information advertisement. It identifies the set of sites that can use some route, namely by which Sites this route can be received, and the PE router can receive the route transmitted by which Sites. The PE routers connected with the site specified in VPN Target can all receive the routes with this attribute.

For PE routers, there are two sets of VPN Target attributes: one of them, referred to as Export Targets, is added to the route received from a direct-connect site in advertising local routes to remote PE routers. And the other one, known as Import Targets, is used to decide which routes can be imported into the routing table of this site in receiving routes from remote PE routers.

When matching the VPN Target attribute carried by the route to filter the routing information received by the PE router, if the export VPN target set of the received route contains identical items with the import VPN target set of the local end, the route is imported into the VPN routing table and then advertised to the connected CE . Otherwise, the route will be rejected.

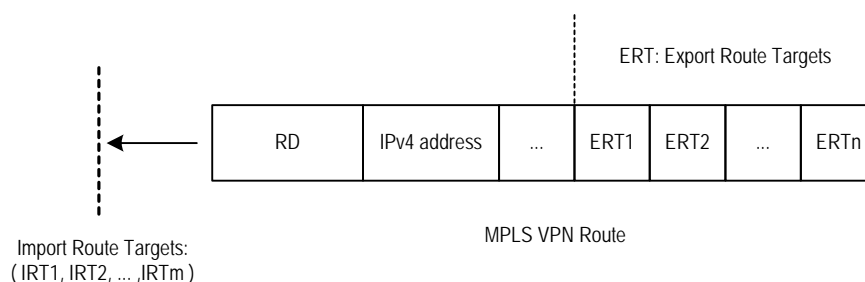


Figure 37-3 Route filtering through matching VPN Target attribute

Note:

The routes for other VPNs will not appear in the VPN's routing table by using VPN Target attribute to filter routing information received at PE router, so the CE-transmitted data will only be forwarded within the VPN.

37.1.2 BGP/MPLS VPN Implementation

BGP/MPLS VPN works on this principle: It uses BGP to propagate VPN private routing information on carrier backbone network, and uses MPLS to forward VPN service traffic.

The following are introductions to BGP/MPLS implementation from two aspects: advertising VPN routing information and forwarding VPN packets.

I. Advertising VPN routing information

Routing information exchange has the following four types:

1) Between CE and PE

A PE router can learn routing information about the CE connected to it through static route, RIP (supporting multi-instance), OSPF (supporting multi-instance) or EBGp, and imports it in a vpn-instance.

2) Between ingress PE and egress PE

The ingress PE router uses MP-BGP to send information across public network: It advertises routing information learned from CE to the egress PE router (with MPLS label) and learns the CE routing information learned at the egress PE router.

The internal connectivity among the VPN internal nodes is ensured through enabling IGP (for example, RIP and OSPF) or configuring static routes on the PEs.

LSP setup between PEs

LSPs must be set up between PEs for VPN data traffic forwarding with MPLS LSP. The PE router which receives packets from CE and create label protocol stack is called ingress LSR, while the BGP next hop (egress PE router) is egress LSR. Using LDP to create fully connected LSPs among PEs.

3) Between PE and CE

A CE can learn remote VPN routes from the PE connected through static routes, RIP, OSPF or EBGP.

With above-mentioned steps, reachable routes can be established between CEs, for transmission of VPN private routing information over public network.

II. Forwarding VPN packets

On the ingress PE, two-layer label stack is formed for each VPN packet:

Interior-layer label, also called MPLS label, is at the bottom of the label stack and distributed by M-BGP when the egress PE advertises routing information (in VPN forwarding table) to ingress GE. When VPN packets from public network reach the CE, they can be forwarded from the designated interface to the designated CE or site by searching for the target MPLS forwarding table according to the labels contained.

Exterior-layer label, known as LSP initialization label, distributed by MPLS LDP, is at the top of the label stack and indicates an LSP from the ingress PE to egress PE. By the switching of exterior-layer label, VPN packets can be forwarded along the LSP to the peer PE.

Figure 37-4 illustrates the details:

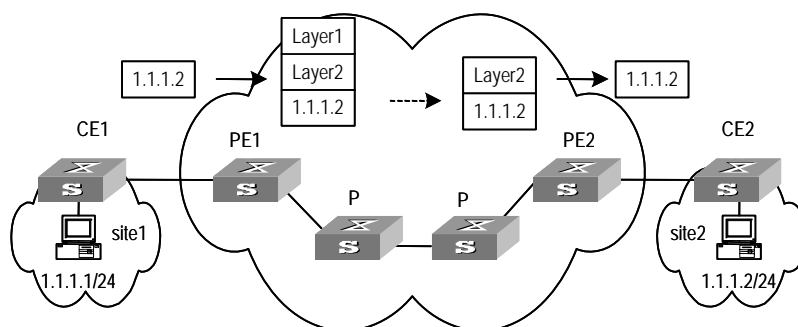


Figure 37-4 Forwarding VPN packets

- 1) Site 1 sends an IPv4 packet with the destination address 1.1.1.2 of to CE1. CE1 looks up the IP routing table for a matched entry and sends the packet to PE1 according to the matched entry.
- 2) Depending on the interface the packet reaches and the destination of it, PE1 looks up the VPN-instance entry to obtain interior-layer label, exterior-layer label, BGP next hop (PE2), and output interfaces. After the establishment of labels, PE1 forwards MPLS packets to the first P of LSP through output interface.

- 3) Each P router on LSP forwards MPLS packets using exterior-layer label to the penultimate-hop router, namely the P router before PE2. The penultimate-hop router extracts the exterior-layer and sends MPLS packet to PE2.
- 4) PE2 looks up in the MPLS forwarding table according to the interior-layer label and destination address to determine the egress interface for labeling operation and the packet. It then extracts the interior-layer label and forwards through the egress interface the IPv4 packet to CE2.
- 5) CE2 looks up in the routing table and sends the packet in normal IPv4 packet forwarding mode to the site2.

37.1.3 Nested BGP/MPLS VPN Implementation

When implementing a nested BGP/MPLS VPN, pay attention to the following items:

- No address overlap is allowed between user's internal sub-VPNs.
- To ensure the VPN routing information is correctly advertised over the backbone network, the VPN-Targets of the user VPN and the internal sub-VPNs cannot be overlapped and must be specified by the service provider.
- The provider PE and the customer PE must be directly connected and cannot exchange VPNv4 route in Multihop-EBGP mode.

Before configuring a nested BGP/MPLS VPN, you must complete the following tasks:

- Configuring IGP on the MPLS backbone network (including provider PE and P routers) to implement the IP connectivity on the backbone network.
- Configuring basic MPLS capability on the MPLS backbone network.
- Configuring MPLS LDP and setting up LDP LSP on the MPLS backbone network.
- Configuring BGP on the MPLS backbone network (create EBGP peers between provider PEs).
- Configuring basic MPLS capability on user-end network (including customer PEs).

37.1.4 Hierarchical BGP/MPLS VPN Implementation

As PE is required to aggregate multiple VPN routes on a BGP/MPLS VPN, it is prone to forming a bottleneck in a large-scale deployment or in the case that PE capacity is small.

Hierarchical BGP/MPLS VPN divides an MPLS VPN into several MPLS VPNs in a hierarchical network structure. Each VPN takes on a role depending on its level. There are high performance requirements in routing and forwarding on the PEs at the higher level of MPLS VPN, because they are primarily used for connecting the backbone networks and providing access service for huge VPN clients. However, such requirements are relatively low for PEs at the lower level of the network as they primarily function to access the VPN clients at the edges. Congruous with the IP network model, HoVPN model improves the scalability of BGP/MPLS VPN, and hence allows lower-layer MPLS VPNs comprising low-end equipment to provide MPLS VPN accessing and interconnect through the high-end MPLS VPN backbone.

As shown in Figure 37-5, the PEs directly connected with user devices are called UPE (underlayer PE or user-end PE); the devices in the core network connected with the UPEs are called SPE (superstratum PE or service-provider-end PE).

Hierarchical PEs have the same appearance as that of the traditional PEs and can coexist with other PEs in the same MPLS network.

UPEs are responsible for user access; they only maintain the routes of directly connected VPN sites, but not that of the remote sites. SPEs, however, are responsible for the maintenance and advertisement of VPN routes; they maintain all the routes of the VPNs connected by their UPEs, including the routes in both local and remote sites.

UPE and SPE are relative concepts. In a multi-layer PE architecture, an upper layer PE is an SPE for its lower layer PE, and a lower layer PE is an UPE for its upper layer PE.

The MBGP runs between SPE and UPE can be either MP-IBGP or MP-EBGP, depending on whether the SPE and the UPE are in the same AS.

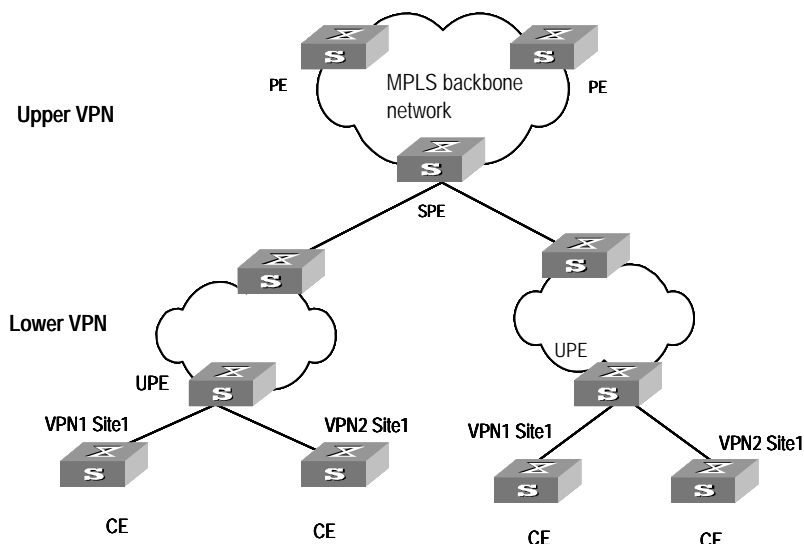


Figure 37-5 Hierarchical BGP/MPLS VPN

37.1.5 Introduction to OSPF Multi-instance

As one of the most popular IGP routing protocols, OSPF is used as an internal routing protocol in many VPNs. Using OSPF on PE-CE links brings convenience to you because in this case CE routers only need to support OSPF protocol, without the need of supporting other protocols, and network administrator only have to know the OSPF protocol. If you want to transform conventional OSPF backbone into BGP/MPLS VPN, using OSPF between PE and CE can simplify this transform process.

Therefore IETF raised two new OSPF VPN extension drafts, to provide a complete solution to SPPF problems in BGP/MPLS VPN application when OSPF is used as PE-CE routing protocol. In this case, PE router must be able to run multiple OSPF instances, each of which corresponds to one VPN instance, owns an individual

interface, routing table, and sends VPN routing information over MPLS network using BGP/OSPF interaction.

If supporting OSPF multi-instance, one router can run multiple OSPF procedures, which can be bound to different VPN instances. In practice, you can create one OSPF instance for each service type. OSPF multi-instance can fully isolate different services in transmission, which can solve security problems with low cost to meet the needs of customers. Generally, OSPF multi-instance is run on PEs; The CE running OSPF multi-instance in the LAN is called multi-VPN-instance CE. At present, isolation of LAN services implements by VLAN function of the switch. OSPF Multi-VPN-Instance CE provides schemes of services isolation implemented on routers.

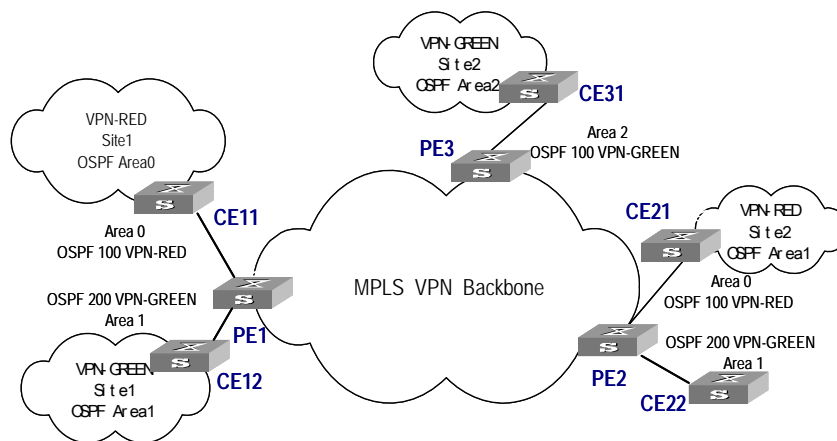


Figure 37-6 OSPF multi-instance application in MPLS/BGP VPN PE

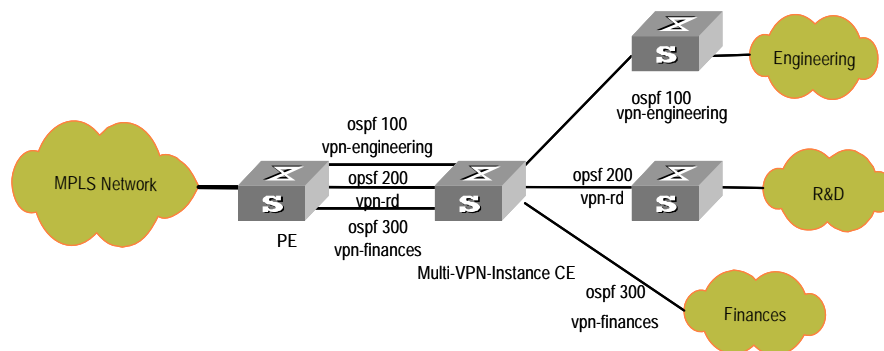


Figure 37-7 Multi-VPN-instance CE application in conventional LAN

37.1.6 Introduction to Multi-Role Host

The VPN attribute of the packets from a CE to its PE lies on the VPN bound with the ingress interface. This, in fact determines that all the CEs forwarded by the PE through the same ingress interface belong to the same VPN; but in actual network environments, a CE may need to access multiple VPNs through one physical interface. Though you can configure different logical interfaces to meet this need, this

compromised method brings additional configuration burden and has limitation in actual use.

To resolve this problem, the idea of multi-role host is generated. Specifically to say, this idea is to differentiate the accesses to different VPNs through configuring policy routing based on IP addresses, and transmit downstream data flow from PE to CE by configuring static routing. The static routing under multi-role host circumstance is different from common hosts; it is implemented by specifying an interface of another VPN as the egress interface through a static route in a VPN; and thus allowing one logical interface to access multiple VPNs.

37.2 BGP/MPLS VPN Configuration

Implementing BGP/MPLS VPN functions requires the following procedures in general: Configure basic information on PE, CE and P; establish the logical or physical link with IP capabilities from PE to PE; advertise and update VPN network information.

I. CE router

The configuration on CE is relative simple. Only static route, RIP, OSPF or EBGp configuration is needed for VPN routing information exchange with the PE connected, MPLS configuration is not needed.

II. PE router

The configuration on PE is relative complex. After the configuration, the PE implements MPLS/BGP VPN core functions.

The following sections describe the configuration tasks on a PE device:

- Configuring basic MPLS capability
- Defining BGP/MPLS VPN site
- Configuring PE-CE route exchanging
- Configuring PE-PE route exchanging

III. P router

The configuration on P device is relative simple. The main task is to configure MPLS basic capacity on the P device to support LDP and MPLS forwarding.

The following are detailed configurations.

37.2.1 Configuring CE Router

As a customer-side device, only basic configuration is required on a CE router, for routing information exchange with PE router. Currently route switching modes available include static route, RIP, OSPF, EBGp, and so on.

I. Creating static route

If you select static route mode for CE-PE route switching, you should then configure a private static route pointing to PE on CE.

Perform the following configuration in the system view.

Table 37-1 Create/delete a static route in VPN instance routing table

Operation	Command
Create a specified vpn-instance static route	ip route-static <i>ip-address</i> { <i>mask</i> <i>mask-length</i> } { <i>interface-name</i> <i>gateway-address</i> } [preference <i>preference-value</i>] [reject blackhole]
Delete a specified vpn-instance static route	undo ip route-static <i>ip-address</i> { <i>mask</i> <i>mask-length</i> } [<i>interface-name</i> <i>gateway-address</i>] [preference <i>preference-value</i>]

By default, the preference value for a static route is 60. You can also specify preference for a static route.

II. Configuring RIP

If you select RIP mode for CE-PE route switching, you should then configure RIP on CE. For detailed RIP configuration steps, see the RIP section in this guide

III. Configuring OSPF

If you select OSPF mode for CE-PE route switching, you should then configure OSPF on CE. For configuring OSPF, see the routing protocol section of this guide.

You must configure OSPF multi-instance to isolate services of different VPNs on CE router, which is now called Multi-VPN-Instance CE.

You can bind OSPF procedures with VPN with the following command in OSPF view.

Table 37-2 Configure the router as multi-VPN-instance CE

Operation	Command
Configure the router as multi-VPN-instance CE	vpn-instance-capability simple
Remove the configuration	undo vpn-instance-capability

IV. Configuring EBGp

If you select BGP mode for CE-PE route switching, you should then configure EBGp peer, import direct-connect route, static route and other IGP routes, for BGP to advertise VPN routes to PE.

37.2.2 Configuring PE Router

I. Configuring basic MPLS capability

It includes configuring MPLS LSR ID, enable MPLS globally and enable MPLS in the corresponding VLAN interface view.

See MPLS Basic Capacity Configure for details.

II. Defining BGP/MPLS VPN site

- 1) Create vpn-instance and enter vpn-instance view

The VPN instance is associated with a site. The VPN membership and routing rules of a site is configured in the corresponding VPN instance.

This command is used to create a new vpn-instance and enter the vpn-instance view, or directly enter the vpn-instance view if the vpn-instance already exists.

Perform the following configuration in the system view.

Table 37-3 Create a vpn-instance and enter vpn-instance view

Operation	Command
Create a vpn-instance and enter vpn-instance view	ip vpn-instance <i>vpn-instance-name</i>
Delete a vpn-instance	undo ip vpn-instance <i>vpn-instance-name</i>

By default, no vpn-instance is defined.

- 2) Configure RD for the vpn-instance

After PE router is configured with RD, when a VPN route learned from CE is imported into BGP, BGP attaches the RD in front of the IPv4 address. Then the general IPv4 address which may overlaps between several VPN IPv4 addresses in the VPN is turned into a globally unique VPN IPv4 address and thus ensure the correct routing in the VPN.

Perform the following configuration in vpn-instance view.

Table 37-4 Configure RD for the vpn-instance

Operation	Command
Configure RD for the vpn-instance	route-distinguisher <i>route-distinguisher</i>

The parameter in the above command has no default value. A vpn-instance works only when a RD is configured for it. Other parameters for a vpn-instance cannot be configured before configuring a RD for it.

To modify the RD, you must first delete the vpn-instance and reconfigure it.

3) Configure vpn-instance description

Perform the following configuration in vpn-instance view

Table 37-5 Configure vpn-instance description

Operation	Command
Configure vpn-instance description	description <i>vpn-instance-description</i>
Delete vpn-instance description	undo description

4) Configure vpn-target attribute for the vpn-instance

VPN-target attribute, a BGP extension community attribute, controls advertisement of VPN routing information.

The following is the advertisement controlling process of VPN routing information:

- When BGP is imported into a VPN route learned at CE, it associates a VPN-target extension community attribute list for the route. Usually the list is the VPN-instance output routing attribute list which is associated with CE.
- VPN instance defines input routing attribute list according to the **import-extcommunity** in VPN-target, defines the acceptable route range and import it.
- VPN instance modifies VPN-target attributes for the routes to be advertised, according to the **export-extcommunity** in VPN-target.

Like an RD, an extension community includes an ASN plus an arbitrary number or an IP address plus an arbitrary number. There are two types of formats:

The first one is related to autonomous system number (ASN), in the form of 16-bit ASN (can be 0 here): 32-bit user-defined number, for example, 100:1.

The second one is related to IP address, in the form of 32-bit IP address (can be 0.0.0.0 here):16-bit user-defined number, for example, 172.1.1.1:1.

Perform the following configuration in the vpn-instance view.

Table 37-6 Configure vpn-target attribute for the vpn-instance

Operation	Command
Configure vpn-target attribute for the vpn-instance	vpn-target <i>vpn-target-extcommunity</i> [import-extcommunity export-extcommunity both]
Delete the specified route-target attribute from the vpn-target attribute list associated with the vpn-instance	undo vpn-target <i>vpn-target-extcommunity</i> [import-extcommunity export-extcommunity both]

By default, the value is **both**. In general all sites in a VPN can be interconnected, and the `import-extcommunity` and `export-extcommunity` attributes are the same, so you can execute the command only with the **both** option.

Up to 16 `vpn-targets` can be configured with a command, and up to 20 `vpn-targets` can be configured for a VPN-instance.

5) Limit the maximum number of routes in a vpn-instance

This command is used to limit the maximum number of routes for a vpn-instance so as to avoid too many routes imported from a site.

Perform the following configuration in the `vpn-instance` view.

Table 37-7 Limit the maximum number of routes in the vpn-instance

Operation	Command
Limit the maximum number of routes in the <code>vpn-instance</code>	routing-table limit { <i>warn threshold</i> <i>simpleinteger</i> { <i>alarm-integer</i> syslog-alert }
Remove the maximum number limitation	undo routing-table limit

Integer is in the range of 1 to 65536 and *alarm-integer* is in the range of 1 to 100.

Note:

Changing the maximum route limit for VPN-instance will not affect the existing routing table. To make the new configuration take effect immediately, you should rebuild the corresponding routing protocol or perform **shutdown/undo shutdown** operation on the corresponding interface.

6) Configure packet redirection in hybrid MPLS VPN networking mode (in this mode, both the service card that does not support MPLS (e.g., B card) and the service card that supports MPLS are used) (optional).

When you configure MPLS/VPN service, you need not replace all the MPLS-incapable B cards with MPLS-capable cards, just add a MPLS-capable card in the B card environment and that is all right. To use the Ethernet port on B card for the connection on CE side, you must configure packet redirection on B card in Ethernet port view to redirect packets to the Ethernet port on the MPLS-capable card for MPLS processing.

- Configure basic ACL.

Perform the following configuration in the system view.

Table 37-8 Configure basic ACL

Operation	Command
Configure basic ACL	acl { number <i>acl-number</i> name <i>acl-name</i> basic } [match-order { config auto }]
Delete basic ACL	undo acl { number <i>acl-number</i> name <i>acl-name</i> all }

- Defines subrules for the basic ACL

Perform the following configuration in corresponding ACL view.

Table 37-9 Define subrules of the ACL

Operation	Command
Define the subrule of basic ACL	rule [<i>rule-id</i>] { permit deny } [source <i>source-addr wildcard</i> any] [fragment] [time-range <i>name</i>]
Delete the subrule of basic ACL	undo rule <i>rule-id</i> [source] [fragment] [time-range]

- Add Ethernet ports on the B card into a VLAN

Perform the following configuration in VLAN view.

Table 37-10 Add Ethernet ports into a VLAN

Operation	Command
Add one or a group of ports into a VLAN	port <i>interface_list</i>
Remove one or a group of ports from a VLAN	undo port <i>interface_list</i>

- Configure virtual interfaces for the above-mentioned VLAN.

Perform the following configuration in the system view.

Table 37-11 Configure VLAN interfaces for a VLAN

Operation	Command
Configure VLAN interface	interface <i>vlan-interface</i>
Delete VLAN interface	undo interface <i>vlan-interface</i>

- Configure packet redirection at the Ethernet port on B card.

Packet redirection falls into three types:

- IP address-based packet redirection (all IP packets are allowed to pass),
- VLAN ID-based packet redirection,

- Both VLAN ID and IP address-based packet redirection.

Perform the following configuration in the Ethernet port view.

Table 37-12 Configure packet redirection at the Ethernet port on the B card

Operation	Command
Configure packet redirection to the specific port on the supporting MPLS card.	traffic-redirect inbound { link-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i> [system-index <i>index</i>]] ip-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i>] } { [system-index <i>index</i>] } } interface { <i>interface-name</i> <i>interface-type</i> <i>interface-num</i> } { I2-vpn <i>destination-vlan</i> I3-vpn
Delete packet redirection configuration	undo traffic-redirect inbound { link-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i>] ip-group { <i>acl-number</i> <i>acl-name</i> } [rule <i>rule</i>] }

7) Associate interface with vpn-instance

VPN instance is associated with the direct-connect site through interface binding. When the packets from the site reach the PE router through the interface bound, then the PE can look routing information (including next hop, label, egress interface, and so on.) up in the corresponding vpn-instance.

This command can associate a vpn-instance with an interface.

Perform the following configuration in VLAN interface view.

Table 37-13 Associate interface with vpn-instance

Operation	Command
Associate interface with vpn-instance	ip binding vpn-instance <i>vpn-instance-name</i>
Remove the association of the interface with vpn-instance	undo ip binding vpn-instance <i>vpn-instance-name</i>

Caution:

As executing the **ip binding vpn-instance** command on an interface will delete the IP address of the interface, you must configure the IP address of the interface after executing that command when you bind the interface with a vpn-instance.

III. Configuring PE-CE route exchanging

These route exchanging modes are available between PE and CE: static route, RIP, OSPF, EBGp.

- 1) Configure static route on PE

You can configure a static route pointing to CE on PE for it to learn VPN routing information from CE.

Perform the following configuration in the system view.

Table 37-14 Configure static route in vpn-instance routing table

Operation	Command
Create a specific vpn-instance static route	ip route-static vpn-instance <i>vpn-instance-name1</i> <i>vpn-instance-name2</i> ... <i>ip-address</i> { <i>mask</i> <i>mask-length</i> } { <i>interface-name</i> [vpn-instance <i>vpn-nexthop-name</i> <i>vpn-nexthop-address</i>] } [preference <i>preference-value</i>] [reject blackhole]
Delete a specific vpn-instance static route	undo ip route-static vpn-instance <i>vpn-instance-name1</i> <i>vpn-instance-name2</i> ... <i>ip-address</i> { <i>mask</i> <i>mask-length</i> } { <i>interface-name</i> [vpn-instance <i>vpn-nexthop-name</i> <i>vpn-nexthop-address</i>] } [preference <i>preference-value</i>] [reject blackhole]

By default, the preference value for a static route is 60. You can also specify another preference for the static route you are configuring.

2) Configure RIP multi-instance

If you select RIP mode for CE-PE route switching, you should then specify running environment for RIP instance on PE. With this command, you can enter RIP view and import and advertise RIP instance in the view.

Perform the following configuration in the RIP view.

Table 37-15 Configure PE-CE RIP instance

Operation	Command
Create PE-CE RIP instance	ipv4-family [unicast] vpn-instance <i>vpn-instance-name</i>
Delete PE-CE RIP instance	undo ipv4-family [unicast] vpn-instance <i>vpn-instance-name</i>

Then configuring RIP multi-instance to import IBGP route.

For details about RIP configuration, see RIP configuration section in Routing Protocol of this manual.

3) Configure OSPF multi-instance on PE

If you select OSPF mode for CE-PE route switching, you should then configure OSPF multi-instance on PE. Other configurations, such as MPLS basic configuration, vpn-instance configuration, do not change. Noted that when OSPF routes and direct-connect routes are imported in the VPN instance address family view, BGP

routes should also be imported into OSPF. Here only introduces OSPF multi-instance configuration in detail.

First step: Configure OSPF procedure.

Perform the following configuration in the system view.

Table 37-16 Configure OSPF procedure

Operation	Command
Configure an OSPF procedure	ospf <i>process-id</i> [router-id <i>router-id-number</i>] [vpn-instance <i>vpn-instance-name</i>]
Delete an OSPF procedure	undo ospf <i>process-id</i>

By default, the procedure index is 1.



Caution:

An OSPF procedure can only belong to one VPN instance, while one VPN instance may contain multiple OSPF procedures. By default, an OSPF procedure belongs to public network.

Step 2: Configure domain ID

The domain ID is used to identify an OSPF autonomous system (AS), and the same OSPF domain must have the same domain ID. One process can be configured with only one domain ID; different processes can be configured with the same domain ID or different domain IDs.

Perform the following configuration in the OSPF view.

Table 37-17 Configure domain ID

Operation	Command
Configure domain ID	domain-id { <i>id-number</i> <i>id-addr</i> }
Return to the default value	undo domain-id

By default, *id-number* is 0 and *id-addr* is 0.0.0.0.

It is recommended that all OSPF instances in a VPN are configured with either the same domain ID or the default value.

**Caution:**

The configured value will not take effect until the command **reset ospf** is executed.

Step 3: Configure tag for imported VPN route (optional)

If a VPN site links to multiple PEs, routing ring may present when the routes learned by MPLS/BGP are received by another PE router in being advertised by category-5/-7 LSA of a PE to the VPN site. To solve this problem, you should configure route-tag. It is recommended to configure identical route-tag for the PEs in the same VPN.

Perform the following configuration in the OSPF view.

**Caution:**

The configured value will not take effect until the command **reset ospf** is executed.

Table 37-18 Configure tag for imported VPN route

Operation	Command
Configure tag for imported VPN route	route-tag <i>tag-number</i>
Return to the default value	undo route-tag

tag-number is used to identify tag value; by default, the first two bytes are fixed, that is, 0xD000, and the last two bytes is AS number of local BGP. For example, the AS number of local BGP is 100, and then its default tag value is 3489661028 in decimal notation. This value is an integer ranging from 0 to 4294967295.

Step 4: Configure sham link (optional)

Sham links are required between two PEs when backdoor links (that is, the OSPF links that do not pass through the MPLS backbone network) exist between the two PEs and data is expected to be transmitted over the MPLS backbone. A sham link between two PEs is considered as a link in OSPF domain. Its source and destination addresses are both the loopback interface address with 32-bit mask, but this loopback interface should be bound to a VPN instance and direct-connect routes must be imported into BGP by BGP.

Perform the following configuration in the OSPF area view.

Table 37-19 Configure sham link

Operation	Command
Configure sham link	sham-link <i>source-addr destination-addr</i> [cost <i>cost-value</i>] [simple <i>password</i> md5 <i>keyid key</i>] [dead <i>seconds</i>] [hello <i>seconds</i>] [retransmit <i>seconds</i>] [trans-delay <i>seconds</i>]
Delete a sham link	undo sham-link <i>source-addr destination-addr</i>

By default, the cost value is 1, dead value is 40 seconds, hello value is 10 seconds, retransmit value is 5 seconds and trans-delay value is 1 second.

4) Configure EBGP on PE

If you select EBGP between PE and CE, you should configure a neighbor for each VPN in VPN instance address family sub-view, and import IGP route of CE.

Step 1: Configure peer group

Configuring peer group in VPN instance address family view.

Table 37-20 Configure peer group

Operation	Command
Configure a peer group	group <i>group-name</i> [internal external]
Delete the specified peer group	undo group <i>group-name</i>

By default, the peer group is configured as internal. When BGP mode is used for PE-CE route switching, they often belong to different ASs, so you should configure EBGP peer as external.

Step 2: Configure AS number for a specific neighbor and add group member to a peer group

When EBGP mode is used for PE-CE route switching, you should configure AS number for a specific neighbor for every CE VPN-instance.

Perform the following configuration in VPN instance address family view.

Table 37-21 Configure AS number for a specific neighbor

Operation	Command
Configure AS number for a specific neighbor	peer { <i>group-name</i> [<i>peer-address</i> group <i>group-name</i>] } as-number <i>as-number</i>
Delete the AS number of a specific neighbor	undo peer { <i>group-name</i> [<i>peer-address</i> group <i>group-name</i>] } as-number <i>as-number</i>

Step 3: Activate peer (group)

By default, BGP neighbor is active while MBGP neighbor is inactive. You should activate MBGP neighbor in VPNv4 sub-address family view.

Perform the following configuration in VPNv4 sub-address family view.

Table 37-22 Activate/deactivate peer (group)

Operation	Command
Activate the peer (group)	peer <i>group-name</i> enable
Deactivate the peer (group)	undo peer <i>group-name</i> enable

Step 4: Configure MBGP to import VPN route of direct-connect CE

To advertise correct VPN route over public network to other PEs with which BGP adjacency has been created, a PE must import the VPN routing information of the direct-connect CE into its MBGP routing table.

For example, if a static route is used between PE and CE, PE must import a static route in VPN-instance address family sub-view of MBGP (import-route static). If RIP is run between PE and CE, PE must import an RIP route in VPN-instance view of MBGP (import-route rip). If BGP is run between PE and CE, MBGP imports a direct-connect route.

Perform the following configuration in VPN instance address family sub-view.

Table 37-23 Import IGP route

Operation	Command
Import IGP route	import-route <i>protocol</i> [<i>process-id</i>] [med <i>med</i>]
Remove IGP route import	undo import-route <i>protocol</i>

Step 5: Configure BGP as asynchronous.

Perform the following configuration in VPN instance address family sub-view.

Table 37-24 Configure BGP asynchronous with IGP

Operation	Command
Configure BGP asynchronous with IGP	undo synchronization

By default, BGP is in asynchronous mode.

Step 6: Permit route loop configuration in Hub&Spoke networking (optional)

Generally speaking, PE-CE configuration is completed after you specify the AS number of neighbor; for the rest configuration, you can keep the system default values.

In the case of standard BGP, BGP tests routing loop via AS number to avoid generating routing loop. In the case of Hub&Spoke networking, however, PE carries the AS number of the local autonomous system when advertising the routing information to CE, if EBGP is run between PE and CE. Accordingly, the updated routing information will carry the AS number of the local autonomous system when route update is received from CE. In this case, PE will not accept the route update information.

This phenomenon can be avoided by executing the **peer allow-as-loop** command, which makes the PE still receives the route update information containing the local AS number from CE.

Perform the following configuration in IPv4 instance sub-address family view.

Table 37-25 Configure to allow/disable routing loop

Operation	Command
Configure to allow routing loop	peer { <i>group-name</i> <i>peer-address</i> } allow-as-loop <i>asn_limit</i>
Configure to disable routing loop	undo peer { <i>group-name</i> <i>peer-address</i> } allow-as-loop <i>asn_limit</i>

By default, the received route update information is not allowed to generate loop information.

Step 7: Configure BGP features.

IV. Configuring PE-PE route exchanging

To exchange VPN-IPv4 routing information between PEs, you should configure MP-IBGP on PEs.

Perform the following configuration in BGP view or PVN instance address family sub-view.

1) Configure IBGP

These steps are often required.

Step 1: Configure BGP as asynchronous.

Step 2: Configure BGP neighbor.

Note that BGP adjacency is established through loopback interface and the sub-net mask must be 32 bits.

Step 3: Permit BGP session over any operable TCP interface.

In general, BGP uses the best local address in TCP connection. To keep TCP connection available even when the interface involved fails, you can perform the following configuration to permit BGP session over any interface through which TCP connection with the peer can be set up. The command here is usually executed together with the Loopback interface.

Table 37-26 Permit BGP session over any operable TCP interface

Operation	Command
Permit BGP session over any operable TCP interface	peer { <i>peer-address</i> <i>group-name</i> } connect-interface { <i>interface-type</i> <i>interface_num</i> }
Use the best local address for TCP connection	undo peer { <i>peer-address</i> <i>group-name</i> } connect-interface

BGP creates BGP adjacency to the peer end using specific interfaces, which is usually the loopback interface. Because this interface is always in the up state, and thus reduces the strike brought by network shock.

2) Configure MP-IBGP

Step 1: Enter protocol address family view.

Perform the following configuration in BGP view.

Table 37-27 Configure VPNv4 address family

Operation	Command
Enter VPNv4 sub-address family view	ipv4-family vpnv4 [unicast]
Delete VPNv4 sub-address family view configuration	undo ipv4-family vpnv4 [unicast]

Step 2: Configure MBGP neighbor

Configure internal neighbor of MBGP in VPNv4 sub-address family view.

Table 37-28 Configure peer group

Operation	Command
Create a peer group	group <i>group-name</i> [internal external]
Delete a specific peer group	undo group <i>group-name</i>

Choose the internal keyword when using the command to create an IBGP peer group.

Step 3: Activate peer (group)

By default, BGP neighbor is active while MBGP neighbor is inactive. You must enable MBGP neighbor in VPNv4 sub-address family view.

Table 37-29 Enable/disable IBGP peer group

Operation	Command
Enable a peer group	peer <i>group-name</i> enable
Disable a specific peer group	undo peer <i>group-name</i> enable

Step 4: Configure the local address as the next hop in route advertisement (optional)

Since the default value is no configuration, you must show clearly to add in this configuration command when configuring MBGP of PE-PE.

Perform the following configuration in VPNv4 sub-address family view.

Table 37-30 Configure the local address as the next hop in route advertisement

Operation	Command
Configure the local address as the next hop in route advertisement	peer { <i>peer-address</i> <i>group-name</i> } next-hop-local peer { <i>peer-address</i> <i>group-name</i> } next-hop-local
Remove the configuration	undo peer { <i>peer-address</i> <i>group-name</i> } next-hop-local

Step 5: Transfer BGP update packet without AS number (optional)

Perform the following configuration in VPNv4 sub-address family view.

Table 37-31 Transfer BGP update packet without AS number

Operation	Command
Transfer BGP update packet without AS number	peer { <i>peer-address</i> <i>group-name</i> } public-as-only
Transfer BGP update packet with AS number	undo peer { <i>peer-address</i> <i>group-name</i> } public-as-only

Step 6: Advertise default route to the peer (group)

This command adds a default route which uses local address as the next hop on the PE SPE (system processing engine)

Table 37-32 Advertise default route to the peer (group)

Operation	Command
Advertise default route to the peer (group)	peer { <i>peer-address</i> <i>group-name</i> } default-route-advertise [vpn-instance <i>vpn-instancename</i>]

Operation	Command
Remove to advertise default route to the peer (group)	undo peer { <i>peer-address</i> <i>group-name</i> } default-route-advertise [vpn-instance <i>vpn-instancename</i>]

Step 7: Configure BGP neighbor as the UPE of BGP/MPLS VPN

This command is only used for UPE (user port function) of BGP/MPLS VPN.

Configuring the following commands in the VPNv4 sub-address family view.

Table 37-33 Configure BGP neighbor as the UPE of BGP/MPLS VPN

Operation	Command
Configure BGP neighbor as the UPE of BGP/MPLS VPN	peer <i>peer-address</i> upe
Disable the configuration	undo peer <i>peer-address</i> upe

37.2.3 Configuring P Router

P router does not maintain VPN routes, but do keep connection with public network and coordinate with PE in creating LSPs. These configurations are required on P router:

Step 1: Configure MPLS basic capacity and enable LDP on the interfaces connecting P router to PE router, for forwarding MPLS packets. See Chapter 36 MPLS Basic Capability Configuration.

Step 2: Enable OSPF protocol at the interfaces connecting P router to PE router and import direct-connect routes. See “OSPF” part in “Routing Protocol” for details.

37.3 Displaying and Debugging BGP/MPLS VPN

I. Displaying VPN address information from BGP table

After the above configuration, execute **display** command in any view to display the running of the VPNv4 information in BGP database configuration, and to verify the effect of the configuration.

Table 37-34 Display VPN address information from BGP table

Operation	Command
Display VPN address information from BGP table	display bgp vpnv4 { all route-distinguisher <i>rd-value</i> vpn-instance <i>vpn-instance-name</i> } { group network peer routing-table }

II. Displaying IP routing table associated with vpn-instance

After the above configuration, you can execute **display** command in any view to display the corresponding information in the IP routing tables related to vpn-instance, and to verify the effect of the configuration.

Table 37-35 Display IP routing table associated with vpn-instance

Operation	Command
Display IP routing table associated with vpn-instance	display ip routing-table vpn-instance <i>vpn-instance-name</i> [<i>[ip-address]</i> [verbose] statistics]

III. Displaying vpn-instance related information

After the above configuration, executing the **display** command in any view can display the vpn-instance related information, including its RD, description, the interfaces associated with it, and so on. You can view the information to verify the configuration effect.

Table 37-36 Display vpn-instance related information

Operation	Command
Display the vpn-instance related information, including its RD, description, the interfaces associated with it, and so on.	display ip vpn-instance [<i>vpn-instance-name</i> verbose]

IV. Debugging information concerning processing BGP

Execute **debugging** command in user view for the debugging of the related vpn-instance information.

Table 37-37 Enable the debugging for processing BGP

Operation	Command
Enable the debugging for processing BGP	debugging bgp { all event normal { keepalive mp-update open packet update route-refresh update } [receive send] [verbose] }
Disable the debugging	undo debugging bgp { { all event normal keepalive mp-update open packet update route-refresh } [receive send verbose] } { all event normal update }

V. Displaying MPLS I3vpn-Isp information

Table 37-38 Display MPLS I3vpn-Isp information

Operation	Command
Display MPLS I3vpn LSP information	display mpls I3vpn-Isp [verbose] include text
Display MPLS I3vpn LSP vpn-instance information	display mpls I3vpn-Isp [vpn-instance vpn-instance-name] [transit egress ingress] [include text verbose]

VI. Displaying sham link

Table 37-39 Display sham link

Operation	Command
Display sham link	display ospf [<i>process-id</i>] sham-link

37.4 Typical BGP/MPLS VPN Configuration Example

37.4.1 Integrated BGP/MPLS VPN Configuration Example

I. Network requirements

- VPN-A includes CE1 and CE3; VPN-B includes CE2 and CE4.
- Subscribers in different VPNs cannot access each other. The VPN-target attribute for VPN-A is 111:1 and that for VPN-B is 222:2.
- The PEs and P are switches supporting MPLS, and CEs are common layer 3 switches.

Note:

The configuration in this case is focused on:

- Configure EBGp to exchange VPN routing information between CEs and PEs.
 - Configure OSPF for inter-PE communication between PEs.
 - Configure MP-IBGP to exchange VPN routing information between PEs.
-

II. Network diagram

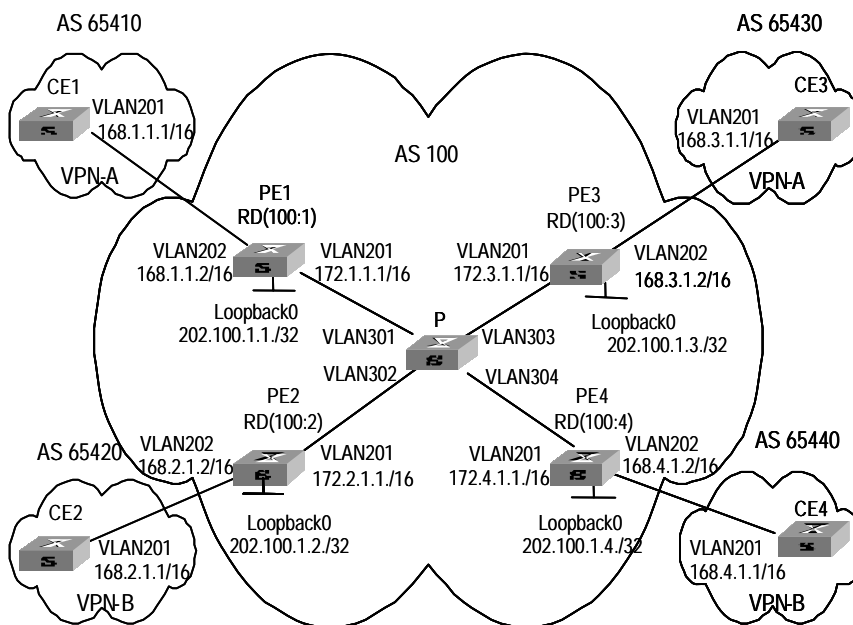


Figure 37-8 Network diagram for integrated BGP/MPLS VPN

III. Configuration procedure

The following are the configuration introduction to PE, CE and P switches.

1) Configure CE1.

Configure CE1 and PE1 as EBGP neighbors, import direct-connect routes and static routes to import intra-CE1 VPN routes into BGP and advertise to PE1. CE1 connects to PE1 through interface GigabitEthernet 2/1/1.

```
[CE1] vlan 201
[CE1-vlan201] port gigabitethernet 2/1/1
[CE1-vlan201] quit
[CE1] interface Vlan-interface 201
[CE1-Vlan-interface201] ip address 168.1.1.1 255.255.0.0
[CE1-Vlan-interface201] quit
[CE1] bgp 65410
[CE1-bgp] group 168 external
[CE1-bgp] peer 168.1.1.2 group 168 as-number 100
[CE1-bgp] import-route direct
[CE1-bgp] import-route static
```

Note:

The configuration on the other three CE switches (CE2 to CE4) is similar to that on CE1, the details are omitted here.

2) Configure PE1

Configure vpn-instance for VPN-A on PE1, as well as other associated attributes to control advertisement of VPN routing information.

```
[PE1] ip vpn-instance vpna
[PE1-vpn-vpna] route-distinguisher 100:1
[PE1-vpn-vpna] vpn-target 111:1 both
[PE1-vpn-vpna] quit
```

Configure PE1 and CE1 as MP-EBGP neighbors, import CE1 VPN routes learned into MBGP VPN-instance address family.

```
[PE1] bgp 100
[PE1-bgp] ipv4-family vpn-instance vpna
[PE1-bgp-af-vpn-instance] import-route direct
[PE1-bgp-af-vpn-instance] group 168 external
[PE1-bgp-af-vpn-instance] peer 168.1.1.1 group 168 as-number 65410
[PE1-bgp-af-vpn-instance] quit
[PE1-bgp] quit
```

Bind the VLAN interface connecting PE1 and CE1 to the VPN-A. Note that you should first configure association between the VLAN interface and VPN-instance, and then configure the IP address of the VLAN interface.

```
[PE1] vlan 202
[PE1-vlan202] port gigabitethernet 2/1/2
[PE1-vlan202] quit
[PE1] interface Vlan-interface 202
[PE1-Vlan-interface202] ip binding vpn-instance vpna
[PE1-Vlan-interface202] ip address 168.1.1.2 255.255.0.0
[PE1-Vlan-interface202] quit
```

Configure loopback interface. (For PE, the IP address for loopback interface must be a host address with 32-bit mask, to prevent the route is aggregated and then LSP cannot process correctly interior-layer labels.)

```
[PE1] interface loopback0
[PE1-LoopBack 0] ip address 202.100.1.1 255.255.255.255
[PE1-LoopBack 0] quit
```

Configure MPLS basic capacity and enable MPLS and LDP on VLAN interface connecting PE1 and P. Create LSP and achieve MPLS packet forwarding.

```
[PE1] mpls lsr-id 202.100.1.1
[PE1] mpls
[PE1-mpls] quit
[PE1] mpls ldp
[PE1] vlan 201
[PE1-vlan201] port gigabitethernet 2/1/1
[PE1-vlan201] quit
[PE1] interface Vlan-interface 201
[PE1-Vlan-interface201] ip address 172.1.1.1 255.255.0.0
[PE1-Vlan-interface201] mpls
[PE1-Vlan-interface201] mpls ldp enable
[PE1-Vlan-interface201] quit
```

Enable OSPF on the interface connecting PE1 and P and on the loopback interface, import direct-connect routes. Achieve inter-PE communication.

```
[PE1] ospf
[PE1-ospf-1] area 0
[PE1-ospf-1-area-0.0.0.0] network 172.1.0.0 0.0.255.255
[PE1-ospf-1-area-0.0.0.0] network 202.100.1.1 0.0.0.0
[PE1-ospf-1-area-0.0.0.0] quit
[PE1-ospf-1] import-route direct
[PE1-ospf-1] quit
```

Set up MP-IBGP adjacency between PEs to exchange inter-PE VPN routing information and activate MP-IBGP peer in VPNv4 sub-address family view.

```
[PE1] bgp 100
[PE1-bgp] group 202 internal
[PE1-bgp] peer 202.100.1.3 group 202
[PE1-bgp] peer 202.100.1.3 connect-interface loopback0
[PE1-bgp] ipv4-family vpnv4
[PE1-bgp-af-vpn] peer 202 enable
[PE1-bgp-af-vpn] peer 202.100.1.3 group 202
[PE1-bgp-af-vpn] quit
[PE1-bgp] quit
```

3) Configure P:

Configure MPLS basic capacity, enable LDP on the interfaces connecting P and PE for MPLS packet forwarding.

```
[P] mpls lsr-id 172.1.1.2
[P] mpls
[P-mpls] quit
[P] mpls ldp
[P] interface loopback0
[P-LoopBack 0] ip address 172.1.1.2 255.255.255.255
```

```
[P-LoopBack 0] quit
[P] vlan 301
[P-vlan301] port gigabitethernet 3/1/1
[P-vlan301] quit
[P] interface Vlan-interface 301
[P-Vlan-interface301] ip address 172.1.1.2 255.255.0.0
[P-Vlan-interface301] mpls
[P-Vlan-interface301] mpls ldp enable
[P-Vlan-interface301] quit
[P] vlan 302
[P-vlan302] port gigabitethernet 3/1/2
[P-vlan302] quit
[P] interface Vlan-interface 302
[P-Vlan-interface302] ip address 172.2.1.2 255.255.0.0
[P-Vlan-interface302] mpls
[P-Vlan-interface302] mpls ldp enable
[P-Vlan-interface302] quit
[P] vlan 303
[P-vlan303] port gigabitethernet 3/1/3
[P-vlan303] quit
[P] interface Vlan-interface 303
[P-Vlan-interface303] ip address 172.3.1.2 255.255.0.0
[P-Vlan-interface303] mpls
[P-Vlan-interface303] mpls ldp enable
[P-Vlan-interface303] quit
[P] vlan 304
[P-vlan304] port gigabitethernet 3/1/4
[P-vlan304] quit
[P] interface Vlan-interface 304
[P-Vlan-interface304] ip address 172.4.1.2 255.255.0.0
[P-Vlan-interface304] mpls
[P-Vlan-interface304] mpls ldp enable
[P-Vlan-interface304] quit
```

Enable OSPF protocol on the interfaces connecting P and PE, import direct-connect route to achieve inter-PE communication.

```
[P] ospf
[P-ospf-1] area 0
[P-ospf-1-area-0.0.0.0] network 172.1.1.0 0.0.255.255
[P-ospf-1-area-0.0.0.0] network 172.2.1.0 0.0.255.255
[P-ospf-1-area-0.0.0.0] network 172.3.1.0 0.0.255.255
[P-ospf-1-area-0.0.0.0] network 172.4.1.0 0.0.255.255
[P-ospf-1-area-0.0.0.0] quit
```

```
[P-ospf-1] import-route direct
```

4) Configure PE3

Note:

The configuration on PE3 is similar to that on PE1, you should pay more attention to VPN routing attribute setting on PE3 to get information about how to control advertisement of a same VPN routing information (with same VPN-target) over MPLS network.

Create VPN-instance for VPN-A on PE3, configure correlative attributes to control advertisement of VPN routing information.

```
[PE3] ip vpn-instance vpna
[PE3-vpn-vpna] route-distinguisher 100:3
[PE3-vpn-vpna] vpn-target 111:1 both
[PE3-vpn-vpna] quit
```

Set up MP-EBGP adjacency between PE3 and CE3, import intra-CE3 VPN routes learned into MBGP VPN-instance address family.

```
[PE3] bgp 100
[PE3-bgp] ipv4-family vpn-instance vpna
[PE3-bgp-af-vpn-instance] import-route direct
[PE3-bgp-af-vpn-instance] group 168 external
[PE3-bgp-af-vpn-instance] peer 168.3.1.1 group 168 as-number 65430
[PE3-bgp-af-vpn-instance] quit
[PE3-bgp] quit
```

Bind the interface connecting PE3 and CE3 to VPN-A.

```
[PE3] vlan 202
[PE3-vlan202] port gigabitethernet 2/1/2
[PE3-vlan202] quit
[PE3] interface Vlan-interface 202
[PE3-Vlan-interface202] ip binding vpn-instance vpna
[PE3-Vlan-interface202] ip address 168.3.1.2 255.255.0.0
[PE3-Vlan-interface202] quit
```

Configure loopback interface

```
[PE3] interface loopback0
[PE3-LoopBack 0] ip address 202.100.1.3 255.255.255.255
[PE3-LoopBack 0] quit
```

Configure MPLS basic capacity and enable MPLS and LDP on VLAN interface connecting PE3 and P. Creates LSP and achieve MPLS packet forwarding.

```
[PE3] mpls lsr-id 202.100.1.3
[PE3] mpls
[PE3-mpls] quit
[PE3] mpls ldp
[PE3] vlan 201
[PE3-vlan201] interface gigabitethernet 2/1/1
[PE3-vlan201] quit
[PE3] interface Vlan-interface 201
[PE3-Vlan-interface201] ip address 172.3.1.1 255.255.0.0
[PE3-Vlan-interface201] mpls
[PE3-Vlan-interface201] mpls ldp enable
[PE3-Vlan-interface201] quit
```

Enable OSPF on the interface connecting PE3 and P and the loopback interface, import direct-connect routes.

```
[PE3] ospf
[PE3-ospf-1] area 0
[PE3-ospf-1-area-0.0.0.0] network 172.3.0.0 0.0.255.255
[PE3-ospf-1-area-0.0.0.0] network 202.100.1.3 0.0.0.0
[PE3-ospf-1-area-0.0.0.0] quit
[PE3-ospf-1] import-route direct
[PE3-ospf-1-area-0.0.0.0] import-route direct
```

Set up MP-IBGP adjacency between PEs to exchange inter-PE VPN routing information.

```
[PE3] bgp 100
[PE3-bgp] group 202 internal
[PE3-bgp] peer 202.100.1.1 group 202 as-number 100
[PE3-bgp] peer 202.100.1.1 connect-interface loopback0
[PE3-bgp] ipv4-family vpnv4
[PE3-bgp-af-vpn] peer 202 enable
[PE3-bgp-af-vpn] peer 202.100.1.1 group 202
[PE3-bgp-af-vpn] quit
```

5) Configure PE2 and PE4

The configuration of PE2 and PE4 is similar to that of PE1 and PE3. The details are omitted here.

37.4.2 Hybrid BGP/MPLS VPN Configuration Example

I. Network requirements

- VPN-A includes CE-1 and CE-3; VPN-B includes CE-2 and CE-4

- Two Switch 8800s serve as PE devices, which support MPLS feature. CE-1 and CE-2 are two mid-range switches; a Layer 2 switch serves as both CE-3 and CE-4, which is accessed directly with users.
- Two PEs are configured with the same interface cards: Slot 2 holds the common interface card with FE ports (B card) and slot 3 holds the enhanced interface card with GE ports (C card).

II. Network diagram

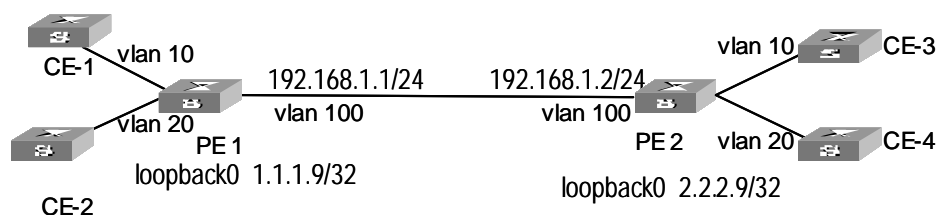


Figure 37-9 Network diagram for hybrid BGP/MPLS VPN

III. Configure procedure

1) Configure CE-1

Create EBGP neighborhood between CE-1 and PE 1, import direct-connect routes and static routes to import the VPN routes inside CE-1 to BGP and to advertise to PE 1, link CE-1 to PE 1 through the Ethernet0/1/0 port.

```
[CE1] interface ethernet 0/1/0
[CE1-Ethernet0/1/0] ip address 20.1.1.1 255.255.0.0
[CE1-Ethernet0/1/0] quit
[CE1] bgp 65410
[CE1-bgp] group 20 external
[CE1-bgp] peer 20.1.1.2 group 20 as-number 100
[CE1-bgp] import-route direct
[CE1-bgp] import-route static
```

Note:

The configuration on CE-2 is similar to that on CE-1, so its detailed configuration is omitted here.

2) Configure PE 1

Configure VPN-instance.

Note:

The configuration on VPN-B is similar to that on VPN-A and only VPN-A configuration is detailed here.

```
[PE1] ip vpn-instance vpna
[PE1-vpn-vpna] route-distinguisher 100:1
[PE1-vpn-vpna] vpn-target 100:1 both
[PE1-vpn-vpna] quit
```

Bind the VLAN interface connecting PE 1 and CE-1 to VPN-A.

```
[PE1] vlan 10
[PE1-vlan10] interface vlan-interface 10
[PE1-vlan-interface10] ip binding vpn-instance vpna
[PE1-vlan-interface10] ip address 20.1.1.2 255.255.255.0
[PE1-vlan-interface10] quit
```

Configure a basic IP ACL to allow redirection of all IP packets on CE.

```
[PE1] acl number 2000
[PE1-acl-basic-2000] rule 0 permit source any
[PE1-acl-basic-2000] quit
[PE1] interface Ethernet 2/1/1
[PE1-Ethernet2/1/1] traffic-redirect inbound ip-group 2000 rule 0 interface
GigabitEthernet 3/3/3 10 13-vpn
```

Create EBGP neighborhood between PE 1 and CE-1 and import the direct routes of the VPN-instance.

```
[PE1] bgp 100
[PE1-bgp] ipv4-family vpn-instance vpna
[PE1-bgp-af-vpn-instance] group 20 external
[PE1-bgp-af-vpn-instance] peer 20.1.1.1 group 20 as-number 65410
[PE1-bgp-af-vpn-instance] import-route direct
[PE1-bgp-af-vpn-instance] quit
[PE1-bgp] quit
```

Create MP-IBGP neighborhood between PE 1 and PE 2 to exchange VPN routing information between them, enable the IBGP peer in VPNv4 address family view.

```
[PE1] bgp 100
[PE1-bgp] group 2
[PE1-bgp] peer 2.2.2.9 group 2
[PE1-bgp] peer 2.2.2.9 connect-interface loopback0
[PE1-bgp] ipv4-family vpnv4
[PE1-bgp-af-vpn] peer 2 enable
[PE1-bgp-af-vpn] peer 2.2.2.9 group 2
```

Globally enable MPLS.

```
[PE1] mpls lsr-id 1.1.1.9
[PE1] mpls
[PE1] mpls ldp
```

Configure the public network interface and enable MPLS on it.

```
[PE1] interface loopback0
[PE1-LoopBack0] ip address 1.1.1.9 255.255.255.255
[PE1-LoopBack0] quit
[PE1] vlan 100
[PE1-vlan100] port GigabitEthernet 3/1/1
[PE1-vlan100] interface vlan-interface 100
[PE1-vlan-interface100] ip address 192.168.1.1 255.255.255.0
[PE1-vlan-interface100] mpls
[PE1-vlan-interface100] mpls ldp enable
[PE1-vlan-interface100] quit
```

Enable OSPF on the interface connecting PE 1 and PE 2 and on the loopback interface, import direct routes to achieve the intercommunication between PE 1 and PE 2.

```
[PE1] ospf 1 route-id 1.1.1.9
[PE1-ospf-1] area 0
[PE1-ospf-1-area-0.0.0.0] network 192.168.1.0 0.255.255.255
[PE1-ospf-1-area-0.0.0.0] network 1.1.1.9 0.0.0.0
[PE1-ospf-1-area-0.0.0.0] import-route direct
```

3) Configure PE 2

Note:

Successful redirection configuration clears the VLAN configuration on the destination port:

- If the destination port is not a trunk port, the redirection configuration changes the port as a trunk port and clears all existing settings on the port, including protocol VLAN and static ARP.
- If the destination port is a trunk port and redirection has never been configured on it, the redirection configuration clears all existing settings.
- When the redirection configuration is removed, the source port is deleted from the VLAN to which the packets are redirected, no matter whether the port belongs to the VLAN before the redirection configuration.

Configure the VPN-instance.

Note:

The configuration on VPN-B is similar to that on VPN-A and only VPN-A configuration is detailed here.

```
[PE2] ip vpn-instance vpna
[PE2-vpn-vpna] route-distinguisher 100:1
[PE2-vpn-vpna] vpn-target 100:1 both
[PE2-vpn-vpna] quit
```

Configure the VLAN interface connecting PE 2 with CE-3 and bind the VLAN interface to VPN-A.

```
[PE2] vlan 10
[PE2-vlan10] interface vlan-interface 10
[PE2-vlan-interface10] ip binding vpn-instance vpna
[PE2-vlan-interface10] ip address 20.2.1.2 255.255.255.0
[PE2-vlan-interface10] quit
```

Configure a user-defined flow template and a link ACL, and then perform the redirection configuration.

```
[PE2] flow-template user-defined slot 3 vlanid
[PE2] acl number 4000
[PE2-acl-link-4000] rule 0 permit ingress 10 egress any
[PE2-acl-link-4000] quit
[PE2] interface Ethernet 2/1/1
[PE2-Ethernet2/1/1] port link-type trunk
[PE2-Ethernet2/1/1] flow-template user-defined
[PE2-Ethernet2/1/1] traffic-redirect inbound link-group 4000 rule 0 interface
GigabitEthernet 3/3/3 10 13-vpn
```

Import the routes of the private network interface between PE 2 and CE-3.

```
[PE2] bgp 100
[PE2-bgp] ipv4-family vpn-instance vpna
[PE2-bgp-af-vpn-instance] import-route direct
[PE2-bgp-af-vpn-instance] quit
[PE2-bgp] quit
```

Create MP-IBGP neighborhood between PE 1 and PE 2 to exchange VPN routing information between them, enable the IBGP peer in VPNv4 address family view.

```
[PE2] bgp 100
[PE2-bgp] group 2
[PE2-bgp] peer 1.1.1.9 group 2
[PE2-bgp] peer 1.1.1.9 connect-interface loopback0
[PE2-bgp] ipv4-family vpnv4
```

```
[PE2-bgp-af-vpn] peer 2 enable
[PE2-bgp-af-vpn] peer 1.1.1.9 group 2
```

Globally enable MPLS.

```
[PE2] mpls lsr-id 2.2.2.9
[PE2] mpls
[PE2] mpls ldp
```

Configure the public network interface and enable MPLS on it.

```
[PE2] interface loopback0
[PE2-LoopBack0] ip address 2.2.2.9 255.255.255.255
[PE2-LoopBack0] quit
[PE2] vlan 100
[PE2-vlan100] port GigabitEthernet 3/1/1
[PE2-vlan100] interface vlan-interface 100
[PE2-vlan-interface100] ip address 192.168.1.2 255.255.255.0
[PE2-vlan-interface100] mpls
[PE2-vlan-interface100] mpls ldp enable
[PE2-vlan-interface100] quit
```

Enable OSPF on the interface connecting PE 1 and PE 2 and on the loopback interface, import direct routes to allow information exchange between PE 1 and PE 2.

```
[PE2] ospf 1 route-id 2.2.2.9
[PE2-ospf-1] area 0
[PE2-ospf-1-area-0.0.0.0] network 192.168.1.0 0.255.255.255
[PE2-ospf-1-area-0.0.0.0] network 2.2.2.9 0.0.0.0
[PE2-ospf-1-area-0.0.0.0] import-route direct
```

Note:

Cautions and configuration limitations in hybrid networking:

- For a trunk port on the common interface card, you can configure to redirect MPLS VPN of multiple VLANs to the same destination port.
 - In a VLAN you can configure only one source port to redirect MPLS VPN to the destination port. It is unnecessary to make multiple redirection configurations in a VLAN which is redirected to the same destination port, since the VLAN can contain other ports (for Layer 2 forwarding).
 - In a VLAN you can configure only one destination port to which MPLS VPN is redirected and the destination port must be in loopback state. Make sure before redirection configuration that no other loopback port exists in the VLAN.
 - For a destination port on the enhanced interface card, you can make multiple redirection configurations and just need to trunk one VLAN for each redirection configuration. Loopback is set automatically on it after you make one redirection configuration on such a destination port, and no more redirection configurations can be made, so you must check that the destination port is not in manual shutdown state (the **shutdown** command is used) before making another redirection configuration.
 - You are recommended to bind the VLAN interface to the VPN after making MPLS VPN redirection configuration, to enable your configuration.
 - You cannot configure MPLS VPN redirection and protocol VLAN on the same port. That is, you cannot configure MPLS VPN redirection if you have enabled protocol VLAN, and vice versa. MPLS VPN redirection configuration clears all protocol VLANs on the destination port if there are any.
 - You cannot configure MPLS VPN redirection on the POS port or use the POS port as the destination port for MPLS VPN redirection configuration.
 - You cannot configure MPLS VPN redirection on the aggregation port or use the aggregation port as the destination port for MPLS VPN redirection configuration. If redirection configuration is made on the destination port on the common interface card or enhanced interface card, you cannot add it to any aggregation group.
 - If VRRP is enabled on the VLAN interface where the source port for MPLS VPN redirection configuration belongs to, removing and inserting the enhanced interface card triggers state transition of the VRRP group on the VLAN interface.
 - The trunk FE port can use only the VLAN of 1k for VPN access and MPLS forwarding, but you can define the starting VLAN ID. Then the VLANs which can pass the trunk FE port range from *vlan-id* to *vlan-id+1023*.
-

37.4.3 Extranet Configuration Example

I. Network requirements

Company A and Company B are located at City A and City B respectively. Their headquarters is located at City C. They respectively own VPN1 and VPN2.

In this case, VPN function is provided by MPLS. There are some shared resources at the City C for the two VPNs. All subscribers in both VPNs can access the shared resources, but VPN subscribers in City A and City B cannot access each other.

The two companies cannot use identical IP addresses, for they share the same VPN-instance at PE-C.

Note:

In the case the configuration is focused on controlling access authority of VPN subscribers at different cities by configuring different VPN-target attributes at different PEs.

II. Network diagram

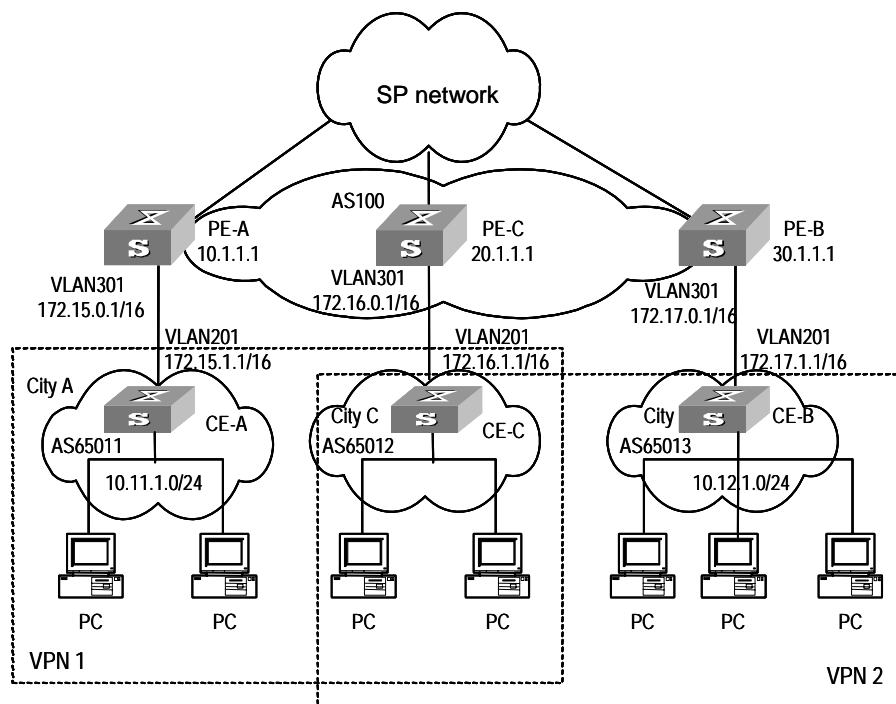


Figure 37-10 Network diagram for Extranet

III. Configuration procedure

Note:

This configuration procedure has omitted configurations between PE and P, and configurations on CEs. For these details refer to the former example.

1) Configure PE-A:

Configure VPN-instance 1 for VPN1 on PE-A, so that it can transceive VPN routing information of VPN-target 111:1.

```
[PE-A] ip vpn-instance vpn-instance 1
[PE-A-vpn-1] route-distinguisher 100:1
[PE-A-vpn-1] vpn-target 111:1 both
[PE-A-vpn-1] quit
```

Set up MP-EBGP adjacency between PE-A and CE-A, import intra-CE-A VPN routes learned into MBGP VPN-instance address family.

```
[PE-A] bgp 100
[PE-A-bgp] ipv4-family vpn-instance vpn-instance1
[PE-A-bgp-af-vpn-instance] import-route direct
[PE-A-bgp-af-vpn-instance] import-route static
[PE-A-bgp-af-vpn-instance] group 172 external
[PE-A-bgp-af-vpn-instance] peer 172.15.1.1 group 172 as-number 65011
[PE-A-bgp-af-vpn-instance] quit
[PE-A-bgp] quit
```

Bind VPN-instance1 with virtual interface of VLAN301 which connects CE-A.

```
[PE-A] vlan 301
[PE-A-vlan301] port gigabitethernet 3/1/1
[PE-A-vlan301] quit
[PE-A] interface Vlan-interface 301
[PE-A-Vlan-interface301] ip binding vpn-instance vpn-instance1
[PE-A-Vlan-interface301] ip address 172.15.0.1 255.255.0.0
[PE-A-Vlan-interface301] quit
```

Configure loopback interface

```
[PE-A] interface loopback 0
[PE-A-LoopBack0] ip address 10.1.1.1 255.255.255.255
[PE-A-LoopBack0] quit
```

Configure MPLS basic capacity.

```
[PE-A] mpls lsr-id 10.1.1.1
[PE-A] mpls
```

```
[PE-A-mpls] quit
[PE-A] mpls ldp
```

Set up MP-IBGP adjacency between PEs to exchange inter-PE VPN routing information and activate MP-IBGP peer in VPNv4 sub-address family view.

```
[PE-A] bgp 100
[PE-A-bgp] group 20 internal
[PE-A-bgp] peer 20.1.1.1 group 20
[PE-A-bgp] peer 20.1.1.1 connect-interface loopback 0
[PE-A-bgp] ipv4-family vpnv4
[PE-A-bgp-af-vpn] peer 20 enable
[PE-A-bgp-af-vpn] peer 20.1.1.1 group 20
[PE-A-bgp-af-vpn] quit
```

2) Configure PE-C.

Create a VPN-instance 2 on PE-C, so that it can transceive VPN routing information of VPN-target 111:1 and 222:2.

```
[PE-C] ip vpn-instance vpn-instance 2
[PE-C-vpn-2] route-distinguisher 100:2
[PE-C-vpn-2] vpn-target 111:1 both
[PE-C-vpn-2] vpn-target 222:2 both
[PE-C-vpn-2] quit
```

Set up MP-EBGP adjacency between PE-C and CE-C, import intra-CE-C VPN routes learned into MBGP VPN-instance address family.

```
[PE-C] bgp 100
[PE-C-bgp] ipv4-family vpn-instance vpn-instance2
[PE-C-bgp-af-vpn-instance] import-route direct
[PE-C-bgp-af-vpn-instance] import-route static
[PE-C-bgp-af-vpn-instance] group 172 external
[PE-C-bgp-af-vpn-instance] peer 172.16.1.1 group 172 as-number 65012
[PE-C-bgp-af-vpn-instance] quit
[PE-C-bgp] quit
```

Bind VPN-instance2 with the interface of VLAN301 which connects CE-C.

```
[PE-C] vlan 301
[PE-C-vlan301] port gigabitethernet 3/1/1
[PE-C-vlan301] quit
[PE-C] interface Vlan-interface 301
[PE-C-Vlan-interface301] ip binding vpn-instance vpn-instance2
[PE-C-Vlan-interface301] ip address 172.16.0.1 255.255.0.0
[PE-C-Vlan-interface301] quit
```

Configure loopback interface

```
[PE-C] interface loopback 0
```

```
[PE-C-LoopBack0] ip address 20.1.1.1 255.255.255.255
[PE-C-LoopBack0] quit
```

Configure MPLS basic capacity.

```
[PE-C] mpls lsr-id 20.1.1.1
[PE-C] mpls
[PE-C-mpls] quit
[PE-C] mpls ldp
```

Set up MP-IBGP adjacency between PEs to exchange inter-PE VPN routing information and activate MP-IBGP peer in VPNv4 sub-address family view.

```
[PE-C] bgp 100
[PE-C-bgp] group 10
[PE-C-bgp] peer 10.1.1.1 group 10
[PE-C-bgp] peer 10.1.1.1 connect-interface loopback 0
[PE-C-bgp] group 30
[PE-C-bgp] peer 30.1.1.1 group 30
[PE-C-bgp] peer 30.1.1.1 connect-interface loopback 0
[PE-C-bgp] ipv4-family vpnv4
[PE-C-bgp-af-vpn] peer 10 enable
[PE-C-bgp-af-vpn] peer 10.1.1.1 group 10
[PE-C-bgp-af-vpn] peer 30 enable
[PE-C-bgp-af-vpn] peer 30.1.1.1 group 30
[PE-C-bgp-af-vpn] quit
```

3) Configure PE-B:

Create VPN-instance 3 for VPN2 on PE-B, so that it can transceive VPN routing information of VPN-target 222:2.

```
[PE-B] ip vpn-instance vpn-instance 3
[PE-B-vpn-3] route-distinguisher 100:3
[PE-B-vpn-3] vpn-target 222:2 both
[PE-B-vpn-3] quit
```

Set up MP-EBGP adjacency between PE-B and CE-B, import intra-CE-B VPN routes learned into MBGP VPN-instance address family.

```
[PE-B] bgp 100
[PE-B-bgp] ipv4-family vpn-instance vpn-instance3
[PE-B-bgp-af-vpn-instance] import-route direct
[PE-B-bgp-af-vpn-instance] import-route static
[PE-B-bgp-af-vpn-instance] group 172 external
[PE-B-bgp-af-vpn-instance] peer 172.17.1.1 group 172 as-number 65013
[PE-B-bgp-af-vpn-instance] quit
[PE-B-bgp] quit
```

Bind VPN-instance3 with interface of VLAN301 which connects to CE-B.

```
[PE-B] vlan 301
[PE-B-vlan301] port gigabitethernet 3/1/1
[PE-B-vlan301] quit
[PE-B] interface Vlan-interface 301
[PE-B-Vlan-interface301] ip binding vpn-instance vpn-instance3
[PE-B-Vlan-interface301] ip address 172.17.0.1 255.255.0.0
[PE-B-Vlan-interface301] quit
```

Configure loopback interface

```
[PE-B] interface loopback 0
[PE-B-LoopBack0] ip address 30.1.1.1 255.255.255.255
[PE-B-LoopBack0] quit
```

Configure MPLS basic capacity.

```
[PE-B] mpls lsr-id 30.1.1.1
[PE-B] mpls
[PE-B-mpls] quit
[PE-B] mpls ldp
```

Set up MP-IBGP adjacency between PEs to exchange inter-PE VPN routing information and activate MP-IBGP peer in VPNv4 sub-address family view.

```
[PE-B] bgp 100
[PE-B-bgp] group 20
[PE-B-bgp] peer 20.1.1.1 group 20
[PE-B-bgp] peer 20.1.1.1 connect-interface loopback 0
[PE-B-bgp] ipv4-family vpnv4
[PE-B-bgp-af-vpn] peer 20 enable
[PE-B-bgp-af-vpn] peer 20.1.1.1 group 20
[PE-B-bgp-af-vpn] quit
```

37.4.4 Hub&Spoke Configuration Example

I. Network requirements

Hub&Spoke networking is also called central server networking. The site in the center is called hub-site, while the one not in the center is called spoke-site. The hub-site knows the routes to all other sites in the same VPN, and the spoke-site must send its traffic first to hub-site and then to the destination. Hub-site is the central node of spoke-sites.

A bank has a headquarters network and subsidiary networks, and it requires that the subsidiaries cannot directly exchange data with each other, but they can exchange data through the headquarters network which provides uniform control. In this case, Hub&Spoke networking topology is used: CE2 and CE3 are spoke-sites, while CE1 is a hub-site in the bank data center. CE1 controls communication between CE2 and CE3.

- Set up IBGP adjacency between PE1 and PE2 or PE1 and PE3, but not between PE2 and PE3, that is, VPN routing information cannot be exchanged between PE2 and PE3.
- Create two VPN-instances on PE1, import VPN routes of VPN-target 100:11 and 100:12, set VPN-target for VPN routes advertised as 100:2.
- Create a VPN-instance on PE2, import VPN routes of VPN-target 100:2, set VPN-target for VPN routes advertised as 100:11.
- Create a VPN-instance on PE3, import VPN routes of VPN-target 100:2, set VPN-target for VPN routes advertised as 100:12.

Then PE2 and PE3 can only learn their neighbor's routes through PE1.

Note:

In this case the configuration is focused on four points:

- Route advertisement can be controlled by VPN-target settings on different PEs.
 - Routing loop is permitted only once, so that PE can receive route update messages with AS number included from CE.
 - In Hub&Spoke networking, vpn-target of VPN-instance (VPN-instance3) which is used to release route on the PE1 cannot be the same with any vpn-target of VPN-instance (VPN-instance2) which is used to import route on PE1.
 - In Hub&Spoke networking, route-distinguisher rd2 (100:3) of VPN-instance which is used to release route on the PE1 cannot be the same with the route-distinguisher rd1 (100:1) or rd4 (100:4) of corresponding VPN-instances on each PE2 and PE3; rd 1 and rd4 can be the same or not.
-

II. Network diagram

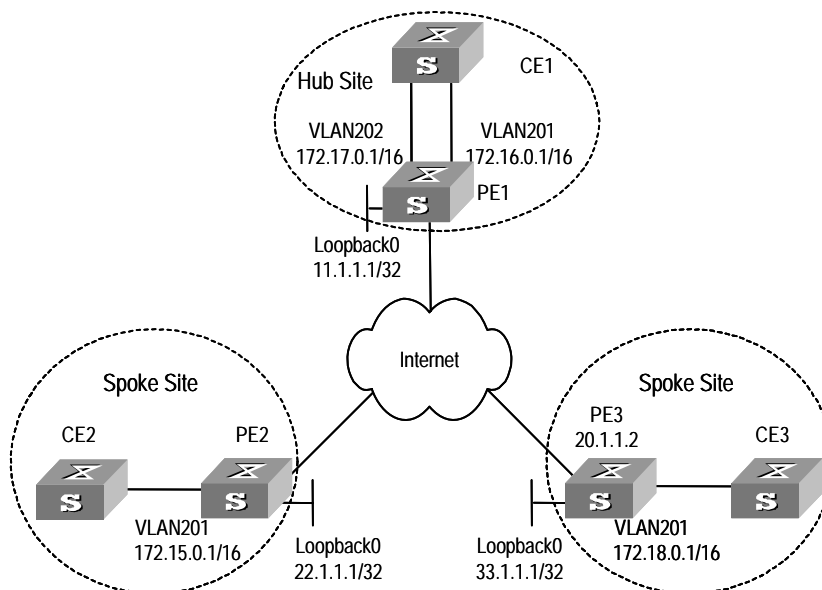


Figure 37-11 Network diagram for Hub&Spoke

III. Configuration procedure

Note:

The following contents are omitted in this case: MPLS basic capacity configuration between PEs, configuration between PE and P, configuration between CEs. For the details refer to 37.4.1 .

1) Configure PE1

Configure two VPN-instances on PE1, set specified VPN-target for the routes received from PE2 and PE3.

```
[PE1] ip vpn-instance vpn-instance2
[PE1-vpn-vpn-instance2] route-distinguisher 100:2
[PE1-vpn-vpn-instance2] vpn-target 100:11 import-extcommunity
[PE1-vpn-vpn-instance2] vpn-target 100:12 import-extcommunity
[PE1-vpn-instance2] quit
[PE1] ip vpn-instance vpn-instance3
[PE1-vpn-vpn-instance3] route-distinguisher 100:3
[PE1-vpn-vpn-instance3] vpn-target 100:2 export-extcommunity
[PE1-vpn-vpn-instance3] quit
```

Set up MP-EBGP adjacency between PE1 and CE1, import intra-CE1 VPN routes learned into MBGP VPN-instance address family, with one routing loop permitted.

```
[PE1] bgp 100
[PE1-bgp] ipv4-family vpn-instance vpn-instance2
[PE1-bgp-af-vpn-instance] import-route static
[PE1-bgp-af-vpn-instance] import-route direct
[PE1-bgp-af-vpn-instance] group 17216 external
[PE1-bgp-af-vpn-instance] peer 172.16.1.1 group 17216 as-number 65002
[PE1-bgp-af-vpn-instance] quit
[PE1-bgp] ipv4-family vpn-instance vpn-instance3
[PE1-bgp-af-vpn-instance] import-route static
[PE1-bgp-af-vpn-instance] import-route direct
[PE1-bgp-af-vpn-instance] group 17217 external
[PE1-bgp-af-vpn-instance] peer 172.17.1.1 group 17217 as-number 65002
[PE1-bgp-af-vpn-instance] peer 172.17.1.1 allow-as-loop 1
[PE1-bgp-af-vpn] quit
[PE1-bgp] quit
```

Bind VLAN interface connecting PE1 and CE1 to different VPN-instances. Bind the interface of the VLAN to which the Ethernet port GigabitEthernet 2/1/1 belongs to vpn-instance2, bind the interface of the VLAN to which the Ethernet port GigabitEthernet 2/1/2 belongs to vpn-instance3.

```
[PE1] vlan 201
[PE1-vlan201] port gigabitEthernet 2/1/1
[PE1-vlan201] quit
[PE1] interface Vlan-interface 201
[PE1-Vlan-interface201] ip binding vpn-instance vpn-instance2
[PE1-Vlan-interface201] ip address 172.16.0.1 255.255.0.0
[PE1-Vlan-interface201] quit
[PE1] vlan 202
[PE1-vlan202] port gigabitEthernet 2/1/2
[PE1-vlan202] quit
[PE1] interface Vlan-interface 202
[PE1-Vlan-interface202] ip binding vpn-instance vpn-instance3
[PE1-Vlan-interface202] ip address 172.17.0.1 255.255.0.0
[PE1-Vlan-interface202] quit
```

Configure loopback interface

```
[PE1] interface loopback 0
[PE1-LoopBack0] ip address 11.1.1.1 255.255.255.255
[PE1-LoopBack0] quit
```

Set up MP-IBGP adjacency between PEs to exchange inter-PE VPN routing information and activate MP-IBGP peer in VPNv4 sub-address family view.

```
[PE1] bgp 100
[PE1-bgp] group 22
```

```
[PE1-bgp] peer 22.1.1.1 group 22 as-number 100
[PE1-bgp] peer 22.1.1.1 connect-interface loopback 0
[PE1-bgp] group 33
[PE1-bgp] peer 33.1.1.1 group 33 as-number 100
[PE1-bgp] peer 33.1.1.1 connect-interface loopback 0
[PE1-bgp] ipv4-family vpnv4
[PE1-bgp-af-vpn] peer 22 enable
[PE1-bgp-af-vpn] peer 22.1.1.1 group 22
[PE1-bgp-af-vpn] peer 33 enable
[PE1-bgp-af-vpn] peer 33.1.1.1 group 33
[PE1-bgp-af-vpn] quit
```

2) Configure PE2

Create a VPN-instance on PE2, import VPN routing information of VPN-target 100:2 and advertise VPN routing information of VPN-target 100:1.

```
[PE2] ip vpn-instance vpn-instance1
[PE2-vpn-vpn-instance1] route-distinguisher 100:1
[PE2-vpn-vpn-instance1] vpn-target 100:11 export-extcommunity
[PE2-vpn-vpn-instance1] vpn-target 100:2 import-extcommunity
[PE2-vpn-vpn-instance1] quit
```

Set up MP-EBGP adjacency between PE2 and CE2, import intra-CE2 VPN routes learned into MBGP VPN-instance address family.

```
[PE2] bgp 100
[PE2-bgp] ipv4-family vpn-instance vpn-instance1
[PE2-bgp-af-vpn-instance] import-route static
[PE2-bgp-af-vpn-instance] import-route direct
[PE2-bgp-af-vpn-instance] group 172 external
[PE2-bgp-af-vpn-instance] peer 172.15.1.1 group 172 as-number 65001
[PE2-bgp-af-vpn-instance] quit
[PE2-bgp] quit
```

Bind the interface of the VLAN to which the port connecting PE2 and CE2 belongs to VPN-instance.

```
[PE2] vlan 201
[PE2-vlan201] port gigabitethernet 2/1/1
[PE2-vlan201] quit
[PE2] interface Vlan-interface 201
[PE2-Vlan-interface201] ip binding vpn-instance vpn-instance1
[PE2-Vlan-interface201] ip address 172.15.0.1 255.255.0.0
[PE2-Vlan-interface201] quit
```

Configure loopback interface

```
[PE2] interface loopback 0
[PE2-LoopBack0] ip address 22.1.1.1 255.255.255.255
```

```
[PE2-LoopBack0] quit
```

Set up MP-IBGP adjacency between PE2 and PE1 to exchange inter-PE VPN routing information and activate MP-IBGP peer in VPNv4 sub-address family view.

```
[PE2] bgp 100
[PE2] group 11
[PE2-bgp] peer 11.1.1.1 group 11 as-number 100
[PE2-bgp] peer 11.1.1.1 connect-interface loopback 0
[PE2-bgp] ipv4-family vpnv4
[PE2-bgp-af-vpn] peer 11 enable
[PE2-bgp-af-vpn] peer 11.1.1.1 allow-as-loop 1
[PE2-bgp-af-vpn] quit
[PE2-bgp] quit
```

3) Configure PE3

Create a VPN-instance on PE3, import VPN routing information of VPN-target 100:2 and advertise VPN routing information of VPN-target 100:12.

```
[PE3] ip vpn-instance vpn-instance2
[PE3-vpn-vpn-instance2] route-distinguisher 100:4
[PE3-vpn-vpn-instance2] vpn-target 100:12 export-extcommunity
[PE3-vpn-vpn-instance2] vpn-target 100:2 import-extcommunity
[PE3-vpn-vpn-instance2] quit
```

Set up MP-EBGP adjacency between PE3 and CE3 import intra-CE3 VPN routes learned into MBGP VPN-instance address family.

```
[PE3] bgp 100
[PE3-bgp] ipv4-family vpn-instance vpn-instance2
[PE3-bgp-af-vpn-instance] import-route static
[PE3-bgp-af-vpn-instance] import-route direct
[PE3-bgp-af-vpn-instance] group 172 external
[PE3-bgp-af-vpn-instance] peer 172.18.1.1 group 172 as-number 65001
[PE3-bgp-af-vpn-instance] quit
[PE3-bgp] quit
```

Bind the interface of the VLAN to which the port connecting PE3 and CE3 belongs to VPN-instance.

```
[PE3] vlan 201
[PE3-vlan201] port gigabitethernet 2/1/1
[PE3-vlan201] quit
[PE3] interface Vlan-interface 201
[PE3-Vlan-interface201] ip binding vpn-instance vpn-instance2
[PE3-Vlan-interface201] ip address 172.18.0.1 255.255.0.0
[PE3-Vlan-interface201] quit
```

Configure loopback interface

```
[PE3] interface loopback 0
[PE3-LoopBack0] ip address 33.1.1.1 255.255.255.255
[PE3-LoopBack0] quit
```

Set up MP-IBGP adjacency between PE3 and PE1 to exchange inter-PE VPN routing information and activate MP-IBGP peer in VPNv4 sub-address family view.

```
[PE3] bgp 100
[PE3-bgp] group 11
[PE3-bgp] peer 11.1.1.1 group 11
[PE3-bgp] peer 11.1.1.1 connect-interface loopback 0
[PE3-bgp] ipv4-family vpnv4
[PE3-bgp-af-vpn] peer 11 enable
[PE2-bgp-af-vpn] peer 11.1.1.1 group 11
[PE2-bgp-af-vpn] peer 11.1.1.1 allow-as-loop 1
[PE3-bgp-af-vpn] quit
[PE3-bgp] quit
```

37.4.5 CE Dual-home Configuration Example

I. Network requirements

For the applications which require high robustness of network, you may use CE dual-home networking mode.

CE1 and CE2 are dual-homed; they are connected to both PE1 and PE2. Three PEs are connected to each other so the links between them are backed up. CE3 and CE4 are single-homed; each of them is only connected to one PE.

CE1 and CE3 are in one VPN, and CE2 and CE4 are in another VPN. The two VPNs cannot intercommunicate with each other.

II. Network diagram

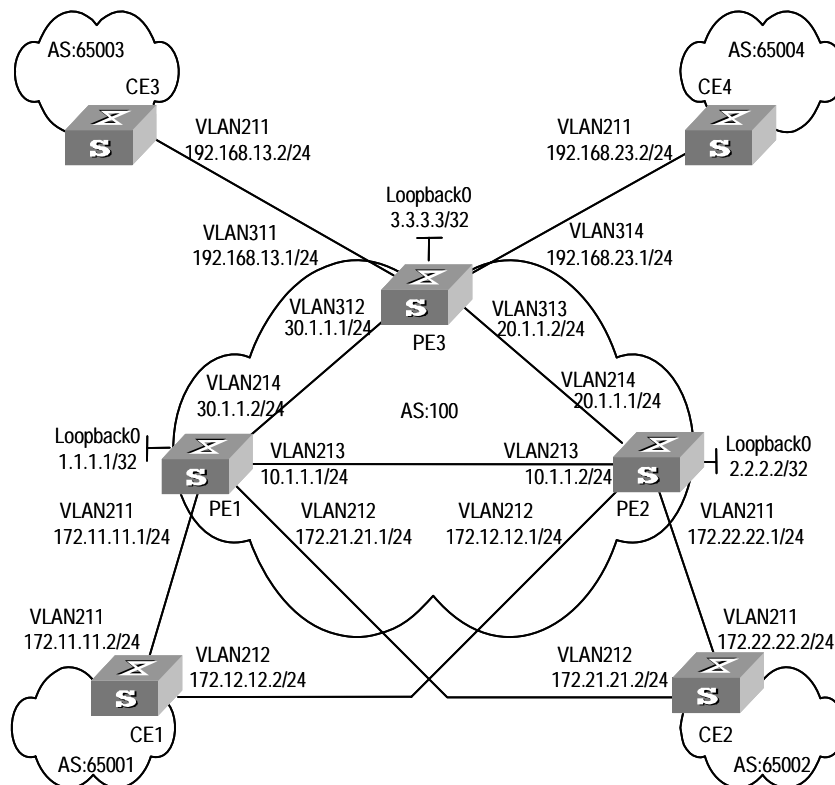


Figure 37-12 Network diagram for CE dual-home

III. Configuration procedure

Note:

The configuration of CE router is omitted in this case and you can refer to Section 37.4.1 Integrated BGP/MPLS VPN Configuration .

1) Configure PE1

Configure two VPN-instances 1.1 and 1.2 respectively for CE1 and CE2 on PE1, set different VPN-targets for them.

```
[PE1] ip vpn-instance vpn-instance1.1
[PE1-vpn-vpn-instance1.1] route-distinguisher 1.1.1.1:1
[PE1-vpn-vpn-instance1.1] vpn-target 1.1.1.1:1
[PE1-vpn-vpn-instance1.1] quit
[PE1] ip vpn-instance vpn-instance1.2
[PE1-vpn-vpn-instance1.2] route-distinguisher 2.2.2.2:2
[PE1-vpn-vpn-instance1.2] vpn-target 2.2.2.2:2
```

```
[PE1-vpn-vpn-instance1.2] quit
```

Set up MP-EBGP adjacency between PE1 and CE1, import intra-CE1 VPN routes learned into VPN-instance 1.1.

```
[PE1] bgp 100
[PE1-bgp] ipv4-family vpn-instance vpn-instance1.1
[PE1-bgp-af-vpn-instance] import-route direct
[PE1-bgp-af-vpn-instance] import-route static
[PE1-bgp-af-vpn-instance] group 17211 external
[PE1-bgp-af-vpn-instance] peer 172.11.11.2 group 17211 as-number 65001
[PE1-bgp-af-vpn-instance] quit
[PE1-bgp] quit
```

Set up MP-EBGP adjacency between PE1 and CE2, import intra-CE2 VPN routes learned into VPN-instance 1.2.

```
[PE1-bgp] ipv4-family vpn-instance vpn-instance1.2
[PE1-bgp-af-vpn-instance] import-route direct
[PE1-bgp-af-vpn-instance] import-route static
[PE1-bgp-af-vpn-instance] group 17221 external
[PE1-bgp-af-vpn-instance] peer 172.21.21.2 group 17221 as-number 65002
[PE1-bgp-af-vpn] quit
[PE1-bgp] quit
```

Bind the interface connecting PE1 and CE1 to VPN-instance 1.1 and interface connecting PE1 and CE2 to VPN-instance 1.2.

```
[PE1] vlan 211
[PE1-vlan211] port gigabitethernet 2/1/1
[PE1-vlan211] quit
[PE1] interface Vlan-interface 211
[PE1-Vlan-interface211] ip binding vpn-instance vpn-instance1.1
[PE1-Vlan-interface211] ip address 172.11.11.1 255.255.255.0
[PE1-Vlan-interface211] quit
[PE1] vlan 212
[PE1-vlan212] port gigabitethernet 2/1/2
[PE1-vlan212] quit
[PE1] interface Vlan-interface 212
[PE1-Vlan-interface212] ip binding vpn-instance vpn-instance1.2
[PE1-Vlan-interface212] ip address 172.21.21.1 255.255.255.0
[PE1-Vlan-interface212] quit
```

Configure loopback interface

```
[PE1] interface loopback 0
[PE1-LoopBack0] ip address 1.1.1.1 255.255.255.255
[PE1-LoopBack0] quit
```


Configure MPLS basic capacity, enable LDP on the interface connecting PE1 and PE2 and the interface connecting PE1 and PE3.

```
[PE1] mpls lsr-id 1.1.1.1
[PE1] mpls
[PE1-mpls] quit
[PE1] mpls ldp
[PE1] vlan 213
[PE1-vlan213] port gigabitethernet 2/1/3
[PE1-vlan213] quit
[PE1] interface Vlan-interface213
[PE1-Vlan-interface213] mpls
[PE1-Vlan-interface213] mpls ldp enable
[PE1-Vlan-interface213] mpls ldp transport-ip interface
[PE1-Vlan-interface213] ip address 10.1.1.1 255.255.255.0
[PE1-Vlan-interface213] quit
[PE1] vlan 214
[PE1-vlan214] port gigabitethernet 2/1/4
[PE1-vlan214] quit
[PE1] interface Vlan-interface 214
[PE1-Vlan-interface214] mpls
[PE1-Vlan-interface214] mpls ldp enable
[PE1-Vlan-interface214] mpls ldp transport-ip interface
[PE1-Vlan-interface214] ip address 30.1.1.2 255.255.255.0
[PE1-Vlan-interface214] quit
```

Enable OSPF on the interface connecting PE1 and PE2 and the interface connecting PE1 and PE3 and the loopback interface, to achieve inter-PE communication.

```
[PE1] Router-id 1.1.1.1
[PE1] ospf
[PE1-ospf-1] area 0
[PE1-ospf-1-area-0.0.0.0] network 1.1.1.1 0.0.0.0
[PE1-ospf-1-area-0.0.0.0] network 30.1.1.2 0.0.0.255
[PE1-ospf-1-area-0.0.0.0] network 10.1.1.1 0.0.0.255
[PE1-ospf-1-area-0.0.0.0] quit
[PE1-ospf-1] quit
```

Set up MP-IBGP adjacency between PEs to exchange inter-PE VPN routing information and activate MP-IBGP peer in VPNv4 sub-address family view.

```
[PE1] bgp 100
[PE1-bgp] group 2
[PE1-bgp] peer 2.2.2.2 group 2
[PE1-bgp] peer 2.2.2.2 connect-interface loopback 0
[PE1-bgp] group 3
```

```
[PE1-bgp] peer 3.3.3.3 group 3
[PE1-bgp] peer 3.3.3.3 connect-interface loopback 0
[PE1-bgp] ipv4-family vpnv4
[PE1-bgp-af-vpn] peer 2 enable
[PE1-bgp-af-vpn] peer 2.2.2.2 group 2
[PE1-bgp-af-vpn] peer 3 enable
[PE1-bgp-af-vpn] peer 3.3.3.3 group 3
[PE1-bgp-af-vpn] quit
```

2) Configure PE2

Note:

The configuration of PE2 is similar to that of PE1, so only VPN-instance configuration is detailed here.

Create two VPN-instances 2.1 and 2.2 respectively for CE1 and CE2 on PE2, configure different VPN-targets for them.

```
[PE2] ip vpn-instance vpn-instance2.1
[PE2-vpn-vpn-instance2.1] route-distinguisher 1.1.1.1:1
[PE2-vpn-vpn-instance2.1] vpn-target 1.1.1.1:1
[PE2-vpn-vpn-instance2.1] quit
[PE2] ip vpn-instance vpn-instance2.2
[PE2-vpn-vpn-instance2.2] route-distinguisher 2.2.2.2:2
[PE2-vpn-vpn-instance2.2] vpn-target 2.2.2.2:2
[PE2-vpn-vpn-instance2.2] quit
```

Set up MP-EBGP adjacency between PE2 and CE1, import intra-CE1 VPN routes learned into VPN-instance2.1.

```
[PE2] bgp 100
[PE2-bgp] ipv4-family vpn-instance vpn-instance2.1
[PE2-bgp-af-vpn-instance] import-route direct
[PE2-bgp-af-vpn-instance] import-route static
[PE2-bgp-af-vpn-instance] group 17212 external
[PE2-bgp-af-vpn-instance] peer 172.12.12.2 group 17212 as-number 65001
[PE2-bgp-af-vpn] quit
```

Set up MP-EBGP adjacency between PE2 and CE2, import intra-CE2 VPN routes learned into VPN-instance2.2.

```
[PE2-bgp] ipv4-family vpn-instance vpn-instance2.2
[PE2-bgp-af-vpn-instance] import-route direct
[PE2-bgp-af-vpn-instance] import-route static
[PE2-bgp-af-vpn-instance] group 17222 external
```

```
[PE2-bgp-af-vpn-instance] peer 172.22.22.2 group 17222 as-number 65002
[PE2-bgp-af-vpn] quit
[PE2-bgp] quit
```

Bind the interface connecting PE2 and CE1 to VPN-instance 2.1 and the interface connecting PE2 and CE2 to VPN-instance 2.2.

```
[PE2] vlan 212
[PE2-vlan212] port gigabitethernet 2/1/2
[PE2-vlan212] quit
[PE2] interface Vlan-interface 212
[PE2-Vlan-interface212] ip binding vpn-instance vpn-instance2.1
[PE2-Vlan-interface212] ip address 172.12.12.1 255.255.255.0
[PE2-Vlan-interface212] quit
[PE2] vlan 211
[PE2-vlan211] port gigabitethernet 2/1/1
[PE2-vlan211] quit
[PE2] interface Vlan-interface 211
[PE2-Vlan-interface211] ip binding vpn-instance vpn-instance2.2
[PE2-Vlan-interface211] ip address 172.22.22.1 255.255.255.0
[PE2-Vlan-interface211] quit
```

3) Configure PE3

Note:

Only the VPN-instance configuration of PE3 is detailed here, other configurations are similar to that of the PE1 and PE2, and are omitted here.

Create two VPN-instances 3.1 and 3.2 respectively for CE3 and CE4 on PE3, configure different VPN-targets for them.

```
[PE3] ip vpn-instance vpn-instance3.1
[PE3-vpn-vpn-instance3.1] route-distinguisher 1.1.1.1:1
[PE3-vpn-vpn-instance3.1] vpn-target 1.1.1.1:1
[PE3-vpn-vpn-instance3.1] quit
[PE3] ip vpn-instance vpn-instance3.2
[PE3-vpn-instance] route-distinguisher 2.2.2.2:2
[PE3-vpn-instance] vpn-target 2.2.2.2:2
[PE3-vpn-instance] quit
```

Set up MP-EBGP adjacency between PE3 and CE3, import intra-CE3 VPN routes learned into VPN-instance3.1.

```
[PE3] bgp 100
[PE3-bgp] ipv4-family vpn-instance vpn-instance3.1
```

```
[PE3-bgp-af-vpn-instance] import-route direct
[PE3-bgp-af-vpn-instance] import-route static
[PE3-bgp-af-vpn-instance] group 192 external
[PE3-bgp-af-vpn-instance] peer 192.168.13.2 group 192 as-number 65003
[PE3-bgp-af-vpn-instance] quit
[PE3-bgp] quit
```

Set up MP-EBGP adjacency between PE3 and CE4, import intra-CE4 VPN routes learned into VPN-instance3.2.

```
[PE3-bgp] ipv4-family vpn-instance vpn-instance3.2
[PE3-bgp-af-vpn-instance] import-route direct
[PE3-bgp-af-vpn-instance] import-route static
[PE3-bgp-af-vpn-instance] group 232 external
[PE3-bgp-af-vpn-instance] peer 192.168.23.2 group 232 as-number 65004
[PE3-bgp-af-vpn-instance] quit
[PE3-bgp] quit
```

Bind the interface connecting PE3 and CE3 to VPN-instance3.1 and the interface connecting PE3 and CE4 to VPN-instance 3.2.

```
[PE3] vlan 311
[PE3-vlan311] port gigabitethernet 3/1/1
[PE3-vlan311] quit
[PE3] interface Vlan-interface 311
[PE3-Vlan-interface311] ip binding vpn-instance vpn-instance3.1
[PE3-Vlan-interface311] ip address 192.168.13.1 255.255.255.0
[PE3-Vlan-interface311] quit
[PE3] vlan 314
[PE3-vlan314] port gigabitethernet 3/1/4
[PE3-vlan314] quit
[PE3] interface Vlan-interface 314
[PE3-Vlan-interface314] ip binding vpn-instance vpn-instance3.2
[PE3-Vlan-interface314] ip address 192.168.23.1 255.255.255.0
[PE3-Vlan-interface314] quit
```

37.4.6 Cross-domain BGP/MPLS VPN Configuration Example

I. Network requirements

A VPN subscriber has sites in both city A and B. Because of the geographical reason, site in City A accesses to the MPLS/VPN network of service provider in City A, and gets AS100 as the AS number; site in City B accesses to the MPLS/VPN network of service provider in City B, and gets AS200 as the AS number. The VPN goes through two ASs. CE1 and CE2 belong to VPN-A, while CE3 and CE4 belong to VPN-B.

II. Network diagram

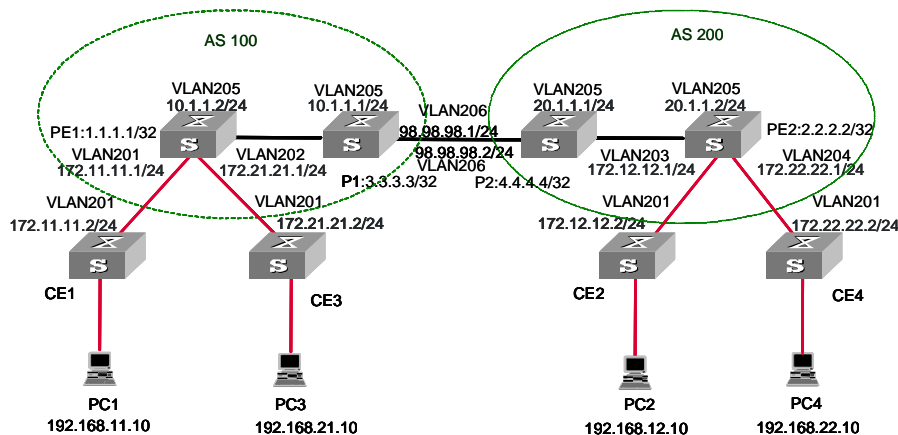


Figure 37-13 Network diagram for ASBR

III. Configuration procedure

1) Configure PE1

Enable MPLS and LDP.

```
[PE1] mpls lsr-id 1.1.1.1
[PE1] mpls
[PE1-mpls] quit
[PE1] mpls ldp
```

Configure the VLAN interface connecting CE.

```
[PE1] vlan 201
[PE1-vlan201] port gigabitethernet 2/1/1
[PE1-vlan201] quit
[PE1] vlan 202
[PE1-vlan202] port gigabitethernet 2/1/2
[PE1-vlan202] quit
```

Configure loopback interface.

```
[PE1] interface loopback 0
[PE1-LoopBack0] ip address 1.1.1.1 255.255.255.255
```

Configure VPN-instance.

```
[PE1] ip vpn-instance vpna
[PE1-vpn-vpna] route-distinguisher 100:1
[PE1-vpn-vpna] vpn-target 100:1 both
[PE1] ip vpn-instance vpnb
[PE1-vpn-vpnb] route-distinguisher 100:2
[PE1-vpn-vpnb] vpn-target 100:2 both
```

Configure VLAN interface connecting PE1 and P1.

```
[PE1] vlan 205
[PE1-vlan205] port gigabitethernet 2/2/1
[PE1-vlan205] quit
[PE1] interface Vlan-interface 205
[PE1-Vlan-interface205] mpls
[PE1-Vlan-interface205] mpls ldp enable
[PE1-Vlan-interface205] ip address 10.1.1.2 255.255.255.0
```

Bind the VLAN interface with the VPN-instance.

```
[PE1] interface Vlan-interface 201
[PE1-Vlan-interface201] ip binding vpn-instance vpna
[PE1-Vlan-interface201] ip address 172.11.11.1 255.255.255.0
[PE1-Vlan-interface201] quit
[PE1] interface Vlan-interface 202
[PE1-Vlan-interface202] ip binding vpn-instance vpnb
[PE1-Vlan-interface202] ip address 172.21.21.1 255.255.255.0
[PE1-Vlan-interface202] quit
```

Enable EBGP between PE and CE.

```
[PE1] bgp 100
[PE1-bgp] ipv4-family vpn-instance vpna
[PE1-bgp-af-vpn-instance] import-route direct
[PE1-bgp-af-vpn-instance] group 172-11 external
[PE1-bgp-af-vpn-instance] peer 172.11.11.2 group 172-11 as-number 65011
[PE1-bgp-af-vpn] quit
[PE1-bgp] ipv4-family vpn-instance vpnb
[PE1-bgp-af-vpn-instance] import-route direct
[PE1-bgp-af-vpn-instance] group 172-21 external
[PE1-bgp-af-vpn-instance] peer 172.21.21.2 group 172-21 as-number 65021
[PE1-bgp-af-vpn-instance] peer 172.21.21.2 next-hop-local
[PE1-bgp-af-vpn-instance] quit
```

Enable IBGP between PE-ASBRs.

```
[PE1-bgp] group 3 internal
[PE1-bgp] peer 3.3.3.3 group 3
[PE1-bgp] peer 3.3.3.3 connect-interface loopback0
[PE1-bgp] ipv4-family vpnv4
[PE1-bgp-af-vpn] peer 3 enable
[PE1-bgp-af-vpn] peer 3.3.3.3 group 3
[PE1-bgp-af-vpn] quit
[PE1-bgp] quit
```

2) Configure PE2

Configure MPLS.

```
[PE2] mpls lsr-id 2.2.2.2
```

```
[PE2] mpls
[PE2-mpls] quit
[PE2] mpls ldp
```

Configure the VLAN interface connecting CE.

```
[PE2] vlan 203
[PE2-vlan203] port gigabitethernet 2/1/1
[PE2-vlan203] quit
[PE2] vlan 204
[PE2-vlan204] port gigabitethernet 2/1/2
[PE2-vlan204] quit
```

Configure loopback interface.

```
[PE2] interface loopback 0
[PE2-LoopBack0] ip address 2.2.2.2 255.255.255.255
```

Configure VPN-instance.

```
[PE2] ip vpn-instance vpna
[PE2-vpn-vpna] route-distinguisher 200:1
[PE2-vpn-vpna] vpn-target 100:1 both
[PE2] ip vpn-instance vpb
[PE2-vpn-vpb] route-distinguisher 200:2
[PE2-vpn-vpb] vpn-target 100:2 both
```

Configure the VLAN interface connecting PE2 and P2.

```
[PE1] vlan 205
[PE1-vlan205] port gigabitethernet 2/2/1
[PE1-vlan205] quit
[PE1] interface Vlan-interface 205
[PE1-Vlan-interface205] mpls
[PE1-Vlan-interface205] mpls ldp enable
[PE1-Vlan-interface205] ip address 20.1.1.2 255.255.255.0
```

Bind the VLAN interface with the VPN-instance.

```
[PE2] interface Vlan-interface 203
[PE2-Vlan-interface203] ip binding vpn-instance vpna
[PE2-Vlan-interface203] ip address 172.12.12.1 255.255.255.0
[PE2-Vlan-interface203] quit
[PE2] interface Vlan-interface 204
[PE2-Vlan-interface204] ip binding vpn-instance vpb
[PE2-Vlan-interface204] ip address 172.22.22.1 255.255.255.0
[PE2-Vlan-interface204] quit
```

Enable EBGp between PE and CE.

```
[PE2] bgp 200
[PE2-bgp] ipv4-family vpn-instance vpna
```

```
[PE2-bgp-af-vpn-instance] import-route direct
[PE2-bgp-af-vpn-instance] group 172-12 external
[PE2-bgp-af-vpn-instance] peer 172.12.12.2 group 172-12 as-number 65012
[PE2-bgp] ipv4-family vpn-instance vpnb
[PE2-bgp-af-vpn-instance] import-route direct
[PE2-bgp-af-vpn-instance] group 172-22 external
[PE2-bgp-af-vpn-instance] peer 172.22.22.2 group 172-22 as-number 65022
[PE2-bgp-af-vpn-instance] quit
[PE2] quit
```

Enable IBGP between PE-ASBRs

```
[PE2-bgp] group 4
[PE2-bgp] peer 4.4.4.4 group 4
[PE2-bgp] peer 4.4.4.4 connect-interface loopback0
[PE2-bgp] ipv4-family vpnv4
[PE2-bgp-af-vpn] peer 4 enable
[PE2-bgp-af-vpn] peer 4.4.4.4 group 4
```

3) Configure P1 (P2 in similar way)

Configure MPLS basic capability.

```
[P1] mpls lsr-id 3.3.3.3
[P1] mpls
[P1-mpls] quit
[P1] mpls ldp
```

Configure the interface loopback 0.

```
[P1] interface loopback 0
[P1-LoopBack0] ip address 3.3.3.3 255.255.255.255
```

Configure VLAN interface connecting PE1.

```
[P1] vlan 205
[P1-vlan205] port gigabitethernet 2/1/1
[P1-vlan205] quit
[P1] interface Vlan-interface 205
[P1-Vlan-interface205] mpls
[P1-Vlan-interface205] mpls ldp enable
[P1-Vlan-interface205] ip address 10.1.1.1 255.255.255.0
[P1-Vlan-interface205] quit
```

Configure VLAN interface connecting PE2.

```
[P1] vlan 206
[P1-vlan206] port gigabitethernet 2/1/2
[P1-vlan206] quit
[P1] interface Vlan-interface 206
[P1-Vlan-interface206] mpls
```



```
[P1-Vlan-interface206] mpls ldp enable
[P1-Vlan-interface206] ip address 98.98.98.1 255.255.255.0
[P1-Vlan-interface206] quit
```

Configure IBGP neighbors and EBGP neighbors.

```
[P1] bgp 100
[P1-bgp] group 1 internal
[P1-bgp] peer 1.1.1.1 group 1
[P1-bgp] peer 1.1.1.1 connect-interface loopback0
[P1-bgp] group 4 external
[P1-bgp] peer 98.98.98.2 group 4 as-number 200
[P1-bgp] ipv4-family vpnv4
[P1-bgp-af-vpn] peer 1 enable
[P1-bgp-af-vpn] peer 1.1.1.1 group 1
[P1-bgp-af-vpn] peer 1 next-hop-local
[P1-bgp-af-vpn] peer 98 enable
[P1-bgp-af-vpn] peer 98.98.98.2 group 98
[P1-bgp-af-vpn] undo policy vpn-target
```

37.4.7 Cross-Domain BGP/MPLS VPN Configuration Example — Option C

I. Network requirements

CE1 and CE2 belong to the same VPN. CE1 accesses the MPLS network through PE1 in AS100; and CE2 accesses the MPLS network through PE2 in AS200.

The example adopts Option C to implement a cross-domain BGP/MPLS VPN, that is, the VPN routing is managed by the Multi-hop MP-EBGP which advertise label VPN-IPv4 routes between PEs.

II. Network diagram

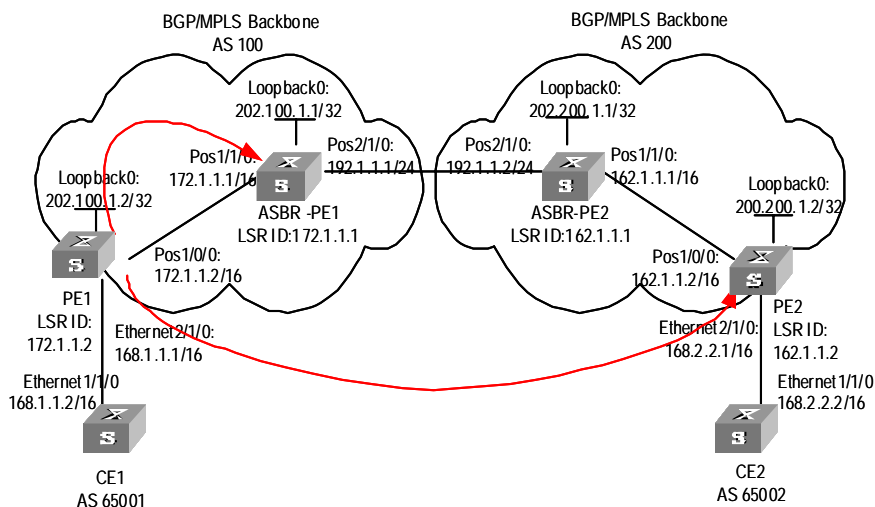


Figure 37-14 Network diagram for Multihop EBGP cross-domain VPN

III. Configuration procedure

- Configuring OSPF on the MPLS backbone network
 - Configuring basic MPLS capability on the MPLS backbone network
 - Configuring a VPN instance on PEs.
 - Configuring MP-BGP
- 1) Configure OSPF as the IGP protocol on the MPLS backbone network; making OSPFs on PEs can learn routes from each other. Create OSPF neighbor between ASBR-PE and PE in the same AS.

Configure PE1.

```
[PE1] interface loopback0
[PE1-LoopBack0] ip address 202.100.1.2 255.255.255.255
[PE1-LoopBack0] quit
[PE1] interface pos1/1/0
[PE1-Pos1/1/0] ip address 172.1.1.2 255.255.0.0
[PE1-Pos1/1/0] quit
[PE1] ospf
[PE1-ospf-1] area 0
[PE1-ospf-1-area-0.0.0.0] network 172.1.0.0 0.0.255.255
[PE1-ospf-1-area-0.0.0.0] network 202.100.1.2 0.0.0.0
[PE1-ospf-1-area-0.0.0.0] quit
[PE1-ospf-1] quit
```

Configure ASBR-PE1

```
[ASBR-PE1] interface loopback0
[ASBR-PE1-LoopBack 0] ip address 202.100.1.1 255.255.255.255
[ASBR-PE1-LoopBack 0] quit
[ASBR-PE1] interface pos1/1/0
[ASBR-PE1-Pos1/1/0] ip address 172.1.1.1 255.255.0.0
[ASBR-PE1-Pos1/1/0] quit
[ASBR-PE1] interface pos 2/1/0
[ASBR-PE1-Pos2/1/0] ip address 192.1.1.1 255.255.255.0
[ASBR-PE1-Pos2/1/0] quit
[ASBR-PE1] ospf
[ASBR-PE1-ospf-1] area 0
[ASBR-PE1-ospf-1-area-0.0.0.0] network 172.1.0.0 0.0.255.255
[ASBR-PE1-ospf-1-area-0.0.0.0] network 202.100.1.1 0.0.0.0
[ASBR-PE1-ospf-1-area-0.0.0.0] quit
[ASBR-PE1-ospf-1] quit
```

Configure PE2

```
[PE2] interface loopback0
[PE2-LoopBack0] ip address 202.200.1.2 255.255.255.255
[PE2-LoopBack0] quit
```

```
[PE2] interface pos1/1/0
[PE2-Pos1/1/0] ip address 162.1.1.2 255.255.0.0
[PE2-Pos1/1/0] quit
[PE2] ospf
[PE2-ospf-1] area 0
[PE2-ospf-1-area-0.0.0.0] network 162.1.0.0 0.0.255.255
[PE2-ospf-1-area-0.0.0.0] network 202.200.1.2 0.0.0.0
[PE2-ospf-1-area-0.0.0.0] quit
[PE2-ospf-1] quit
```

Configure ASBR-PE2

```
[ASBR-PE2] interface loopback0
[ASBR-PE2-LoopBack0] ip address 202.200.1.1 255.255.255.255
[ASBR-PE2-LoopBack0] quit
[ASBR-PE2] interface pos1/1/0
[ASBR-PE2-Pos1/1/0] ip address 162.1.1.1 255.255.0.0
[ASBR-PE2-Pos1/1/0] quit
[ASBR-PE2] interface Pos 2/1/0
[ASBR-PE2-Pos2/1/0] ip address 192.1.1.2 255.255.255.0
[ASBR-PE2-Pos2/1/0] quit
[ASBR-PE2] ospf
[ASBR-PE2-ospf-1] area 0
[ASBR-PE2-ospf-1-area-0.0.0.0] network 162.1.0.0 0.0.255.255
[ASBR-PE2-ospf-1-area-0.0.0.0] network 202.200.1.1 0.0.0.0
[ASBR-PE2-ospf-1-area-0.0.0.0] quit
[ASBR-PE2-ospf-1] quit
```

- 2) Configure basic MPLS capability on the MPLS backbone network to enable the network to forward VPN traffic.

Note:

MPLS must be enabled on the interfaces between the ASBR-PEs.

Configure basic MPLS capability on PE1 and enable LDP on the interface connected to ASBR-PE1.

```
[PE1] mpls lsr-id 172.1.1.2
[PE1-mpls] lsp-trigger all
[PE1-mpls] quit
[PE1] mpls ldp
[PE1-mpls-ldp] quit
[PE1] interface pos1/1/0
[PE1-Pos1/1/0] mpls
```

```
[PE1-Pos1/1/0] mpls ldp
[PE1-Pos1/1/0] quit
```

Configure basic MPLS capability on ASBR-PE1, enable LDP on the interface connected to PE1, and enable MPLS on the interface connected to ASBR-PE2.

```
[ASBR-PE1] mpls lsr-id 172.1.1.1
[ASBR-PE1-mpls] lsp-trigger all
[ASBR-PE1-mpls] quit
[ASBR-PE1] mpls ldp
[ASBR-PE1-mpls-ldp] quit
[ASBR-PE1] interface pos1/1/0
[ASBR-PE1-Pos1/1/0] mpls
[ASBR-PE1-Pos1/1/0] mpls ldp
[ASBR-PE1-Pos1/1/0] quit
[ASBR-PE1] interface pos2/1/0
[ASBR-PE1-Pos2/1/0] mpls
[ASBR-PE1-Pos2/1/0] quit
```

Configure basic MPLS capability on ASBR-PE2, enable LDP on the interface connected to PE2, and enable MPLS on the interface connected to ASBR-PE1.

```
[ASBR-PE2] mpls lsr-id 162.1.1.1
[ASBR-PE2-mpls] lsp-trigger all
[ASBR-PE2-mpls] quit
[ASBR-PE2] mpls ldp
[ASBR-PE2-mpls-ldp] quit
[ASBR-PE2] interface pos1/1/0
[ASBR-PE2-Pos1/1/0] mpls
[ASBR-PE2-Pos1/1/0] mpls ldp
[ASBR-PE2-Pos1/1/0] quit
[ASBR-PE2] interface pos2/1/0
[ASBR-PE2-Pos2/1/0] mpls
[ASBR-PE2-Pos2/1/0] quit
```

Configure basic MPLS capability on PE2 and enable LDP on the interface connected to ASBR-PE2.

```
[PE2] mpls lsr-id 162.1.1.2
[PE2-mpls] lsp-trigger all
[PE2-mpls] quit
[PE2] mpls ldp
[PE2-mpls-ldp] quit
[PE2] interface pos1/1/0
[PE2-Pos1/1/0] mpls
[PE2-Pos1/1/0] mpls ldp
[PE2-Pos1/1/0] quit
```

- 3) Create a VPN instance on each PE, and bind the instance to the interface connected to the corresponding CE.

Configure CE1

```
[CE1] interface ethernet 1
[CE1-Ethernet1] ip address 168.1.1.2 255.255.0.0
[CE1-Ethernet1] quit
```

Create a VPN instance on PE1 and bind it to the interface connected to CE1

```
[PE1] ip vpn-instance vpna
[PE1-vpn-vpn-vpna] route-distinguisher 100:2
[PE1-vpn-vpn-vpna] vpn-target 100:1 both
[PE1-vpn-vpn-vpna] quit
[PE1] interface ethernet 2/1/0
[PE1-Ethernet2/1/0] ip binding vpn-instance vpna
[PE1-Ethernet2/1/0] ip address 168.1.1.1 255.255.0.0
[PE1-Ethernet2/1/0] quit
```

Configure CE2

```
[CE2] interface ethernet 1
[CE2-Ethernet1] ip address 168.2.2.2 255.255.0.0
[CE2-Ethernet1] quit
```

Create a VPN instance on PE2 and bind it to the interface connected to CE2

```
[PE2] ip vpn-instance vpna
[PE2-vpn-instance] route-distinguisher 200:2
[PE2-vpn-instance] vpn-target 100:1 both
[PE2-vpn-instance] quit
[PE2] interface ethernet 2/1/0
[PE2-Ethernet2/1/0] ip binding vpn-instance vpna
[PE2-Ethernet2/1/0] ip address 168.2.2.1 255.255.0.0
[PE2-Ethernet2/1/0] quit
```

- 4) Configure MP-BGP, set up IBGP peer relation between PEs, and set up EBGP peer relation between PEs and their CEs.

Note:

- Enable the exchanging of label-carried IPv4 route between the following routers: PE1 and ASBR-PE1, PE2 and ASBR-PE2, ASBR-PE1 and ASBR-PE2.
 - Make each ASBR-PE change the next hop to its own when it advertises routes to the PE in the same AS.
 - Configure routing policy on each ASBR-PE as follows: make the ASBR-PE assign MPLS label when it advertises a route received from the PE in this AS to the ASBR-PE in the peer AS, and let the ASBR-PE assign a new MPLS label when it advertises a label-carried IPv4 route to the PE in this AS.
-

Configure CE1

```
[CE1] bgp 65001
[CE1-bgp] group 20 external
[CE1-bgp] peer 168.1.1.1 group 20 as-number 100
[CE1-bgp] quit
```

Configure PE1: set up EBGP peer relation with CE1, IBGP peer relation with ASBR-PE1, and Multihop MP-EBGP peer relation with PE2.

```
[PE1] bgp 100
[PE1-bgp] ipv4-family vpn-instance vpna
[PE1-bgp-af-vpn-instance] group 10 external
[PE1-bgp-af-vpn-instance] peer 168.1.1.2 group 10 as-number 65001
[PE1-bgp-af-vpn-instance] import-route direct
[PE1-bgp-af-vpn-instance] quit
[PE1-bgp] group 20
[PE1-bgp] peer 20 label-route-capability
[PE1-bgp] peer 202.100.1.1 group 20
[PE1-bgp] peer 202.100.1.1 connect-interface loopback0
[PE1-bgp] group 30 external
[PE1-bgp] peer 30 ebgp-max-hop
[PE1-bgp] peer 200.200.1.2 group 30 as-number 200
[PE1-bgp] peer 200.200.1.2 connect-interface loopback0
[PE1-bgp] ipv4-family vpnv4
[PE1-bgp-af-vpn] peer 30 enable
[PE1-bgp-af-vpn] peer 200.200.1.2 group 30
[PE1-bgp-af-vpn] quit
[PE1-bgp] quit
```

Configure ASBR-PE1: configure the route policy.

```
[ASBR-PE1] acl number 2001
[ASBR-PE1-acl-basic-2001] rule permit source 202.100.1.2 0
[ASBR-PE1-acl-basic-2001] rule deny source any
```

```
[ASBR-PE1-acl-basic-2001] quit
[ASBR-PE1] route-policy rtp-ebgp permit node 1
[ASBR-PE1-route-policy] if-match acl 2001
[ASBR-PE1-route-policy] apply mpls-label
[ASBR-PE1-route-policy] quit
[ASBR-PE1] route-policy rtp-ibgp permit node 10
[ASBR-PE1-route-policy] if-match mpls-label
[ASBR-PE1-route-policy] apply mpls-label
[ASBR-PE1-route-policy] quit
```

Configure ASBR-PE1: set up EBGP peer relation with ASBR-PE2, and IBGP peer relation with PE1.

```
[ASBR-PE1] bgp 100
[ASBR-PE1-bgp] import-route ospf
[ASBR-PE1-bgp] group 10 external
[ASBR-PE1-bgp] peer 10 label-route-capability
[ASBR-PE1-bgp] peer 10 route-policy rtp-ebgp export
[ASBR-PE1-bgp] peer 192.1.1.2 group 10 as-number 200
[ASBR-PE1-bgp] group 20
[ASBR-PE1-bgp] peer 20 label-route-capability
[ASBR-PE1-bgp] peer 20 next-hop-local
[ASBR-PE1-bgp] peer 20 route-policy rtp-ibgp export
[ASBR-PE1-bgp] peer 202.100.1.2 group 20
[ASBR-PE1-bgp] peer 202.100.1.2 connect-interface loopback0
[ASBR-PE1-bgp] quit
```

Configure CE2.

```
[CE2] bgp 65002
[CE2-bgp] group 10 external
[CE2-bgp] peer 168.2.2.1 group 10 as-number 200
[CE2-bgp] quit
```

Configure PE2: set up EBGP peer relation with CE2, IBGP peer relation with ASBR-PE2, and Multihop MP-EBGP peer relation with PE1.

```
[PE2] bgp 200
[PE2-bgp] ipv4-family vpn-instance vpna
[PE2-bgp-af-vpn-instance] group 10 external
[PE2-bgp-af-vpn-instance] peer 168.2.2.2 group 10 as-number 65002
[PE2-bgp-af-vpn-instance] import-route direct
[PE2-bgp-af-vpn-instance] quit
[PE2-bgp] group 20
[PE2-bgp] peer 20 label-route-capability
[PE2-bgp] peer 202.200.1.1 group 20
[PE2-bgp] peer 202.200.1.1 connect-interface loopback0
```

```
[PE2-bgp] group 30 external
[PE2-bgp] peer 30 ebgp-max-hop
[PE2-bgp] peer 202.100.1.2 group 30 as-number 100
[PE2-bgp] peer 202.100.1.2 connect-interface loopback0
[PE2-bgp] ipv4-family vpnv4
[PE2-bgp-af-vpn] peer 30 enable
[PE2-bgp-af-vpn] peer 202.100.1.2 group 30
[PE2-bgp-af-vpn] quit
[PE2-bgp] quit
```

Configure ASBR-PE2: configure the route policy.

```
[ASBR-PE2] acl number 2001
[ASBR-PE2-acl-basic-2001] rule permit source 200.200.1.2 0
[ASBR-PE2-acl-basic-2001] rule deny source any
[ASBR-PE2-acl-basic-2001] quit
[ASBR-PE2] route-policy rtp-ebgp permit node 1
[ASBR-PE2-route-policy] if-match acl 2001
[ASBR-PE2-route-policy] apply mpls-label
[ASBR-PE2-route-policy] quit
[ASBR-PE2] route-policy rtp-ibgp permit node 10
[ASBR-PE2-route-policy] if-match mpls-label
[ASBR-PE2-route-policy] apply mpls-label
[ASBR-PE2-route-policy] quit
```

Configure ASBR-PE2: set up EBGP peer relation with ASBR-PE1, and IBGP peer relation with PE2.

```
[ASBR-PE2] bgp 200
[ASBR-PE2-bgp] import-route ospf
[ASBR-PE2-bgp] group 10 external
[ASBR-PE2-bgp] peer 10 label-route-capability
[ASBR-PE2-bgp] peer 10 route-policy rtp-ebgp export
[ASBR-PE2-bgp] peer 192.1.1.1 group 10 as-number 100
[ASBR-PE2-bgp] group 20
[ASBR-PE2-bgp] peer 20 label-route-capability
[ASBR-PE2-bgp] peer 20 next-hop-local
[ASBR-PE2-bgp] peer 20 route-policy rtp-ibgp export
[ASBR-PE2-bgp] peer 202.200.1.2 group 20
[ASBR-PE2-bgp] peer 202.200.1.2 connect-interface loopback0
```

37.4.8 Hierarchical BGP/MPLS VPN Configuration Example

I. Network requirements

For those VPNs that have distinct hierarchy, an MPLS VPN covering a province and its cities, for example, incorporating the backbone network at the province level and the

networks at the city level into a single MPLS VPN will impose a high requirement in performance on the equipment on the entire network, in the event that the network topology size is large. However, the requirement in equipment performance can become lower if this MPLS VPN is separated into two VPNs, the network at the province level and the network at the city level, for example.

SPE acts as a PE on the network at the province level, and is connected with a downstream MPLS VPN at the city level. UPE acts as a PE on the network at the city level and provide access service for the VPN clients which are normally low-end routers.

II. Network diagram

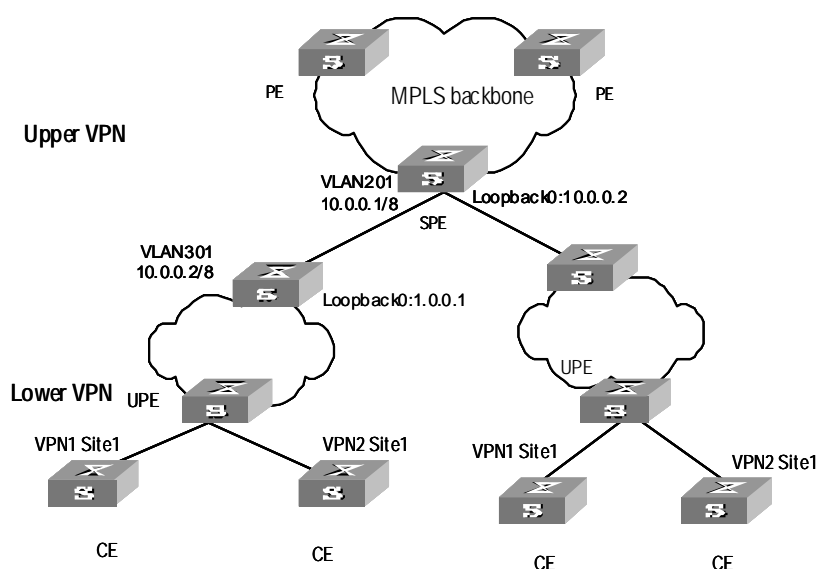


Figure 37-15 Network diagram for hierarchical BGP/MPLS VPN

III. Configuration procedure

Note:

This case only illustrates the configurations concerned with PEs in a hierarchical BGP/MPLS VPN.

1) Configure SPE

Configure the basic MPLS capability.

```
[SPE] mpls lsr-id 1.0.0.2
[SPE] mpls
[SPE-mpls] quit
```

```
[SPE] mpls ldp
```

Configure VPN-instance

```
[SPE] ip vpn-instance vpn1
[SPE-vpn-vpn1] route-distinguisher 100:1
[SPE-vpn-vpn1] vpn-target 100:1 both
```

Configure interfaces (So far as a PE router concerned, its LOOPBACK 0 interface must be assigned with a host address of 32-bit mask.

```
[SPE] vlan 201
[SPE-vlan201] port gigabitethernet 2/1/1
[SPE-vlan201] quit
[SPE] interface Vlan-interface 201
[SPE-Vlan-interface201] ip address 10.0.0.1 255.0.0.0
[SPE-Vlan-interface201] mpls
[SPE-Vlan-interface201] mpls ldp enable
[SPE-Vlan-interface201] quit
[SPE] interface loopback0
[SPE-LoopBack 0] ip address 1.0.0.2 255.255.255.255
[SPE-LoopBack 0] quit
```

Configure BGP

```
[SPE] bgp 100
[SPE] import direct
[SPE-bgp] group 1 internal
[SPE-bgp] peer 1.0.0.1 group 1
[SPE-bgp] peer 1 connect-interface LoopBack0
[SPE-bgp] ipv4-family vpn-instance vpn1
[SPE--bgp-af-vpn-instance] import direct
[SPE--bgp-af-vpn-instance] quit
[SPE-bgp] ipv4-family vpnv4
[SPE-bgp-af-vpn] peer 1 enable
[SPE-bgp-af-vpn] peer 1.0.0.1 group 1
[SPE-bgp-af-vpn] peer 1.0.0.1 upe
[SPE-bgp-af-vpn] peer 1.0.0.1 default-route-advertise vpn-instance vpn1
[SPE-bgp-af-vpn] quit
[SPE-bgp] quit
```

Configure OSPF

```
[SPE] ospf
[SPE] import-route direct
[SPE-ospf-1] area 0
[SPE-ospf-1-area-0.0.0.0] network 1.0.0.2 0.0.0.0
[SPE-ospf-1-area-0.0.0.0] network 10.0.0.0 0.255.255.255
```

2) Configure UPE

Configure the basic MPLS capability.

```
[UPE] mpls lsr-id 1.0.0.1
[UPE] mpls
[UPE-mpls] quit
[UPE] mpls ldp
```

Configure VPN-instance

```
[UPE] ip vpn-instance vpn1
[UPE-vpn-vpn1] route-distinguisher 100:1
[UPE-vpn-vpn1] vpn-target 100:1 both
```

Configure interfaces

```
[UPE] vlan 301
[UPE-vlan301] port gigabitethernet 2/2/1
[UPE-vlan301] quit
[UPE] interface Vlan-interface 301
[UPE-Vlan-interface301] mpls
[UPE-Vlan-interface301] mpls ldp enable
[UPE-Vlan-interface301] mpls ldp transport-ip interface
[UPE-Vlan-interface301] ip address 10.0.0.2 255.0.0.0
[UPE-Vlan-interface301] quit
[UPE] interface loopback0
[UPE-LoopBack 0] ip address 1.0.0.1 255.255.255.255
```

Configure BGP

```
[UPE] bgp 100
[UPE-bgp] group 1 internal
[UPE-bgp] peer 1.0.0.2 group 1
[UPE-bgp] ipv4-family vpn-instance vpn1
[UPE--bgp-af-vpn-instance] import direct
[UPE-bgp] ipv4-family vpnv4
[UPE-bgp-af-vpn] peer 1 enable
[UPE-bgp-af-vpn] peer 1.0.0.2 group 1
```

Configure OSPF

```
[UPE] ospf
[UPE-ospf-1] import-route direct
[UPE-ospf-1] area 0
[UPE-ospf-1-area-0.0.0.0] network 1.0.0.1 0.0.0.0
[UPE-ospf-1-area-0.0.0.0] network 10.0.0.2 0.255.255.255
[UPE-ospf-1-area-0.0.0.0] quit
```

37.4.9 OSPF Multi-instance sham link Configuration Example

I. Network requirements

As shown in the following picture, a company connects to a WAN through OSPF multi-instance function of a router. OSPF is bind to VPN1.MPLS VPN backbone runs between PEs and OSPF runs between PE and CE. Configure a sham link between PE1 and PE2 to ensure the traffic between CE1 and CE2 does not pass the backdoor link that directly connects CE1 and CE2.

II. Network diagram

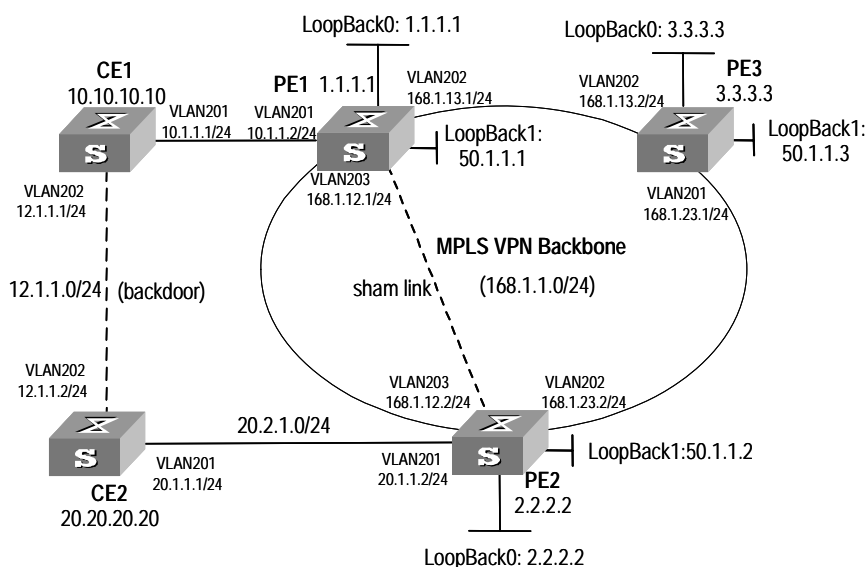


Figure 37-16 Network diagram for OSPF multi-instance

III. Configuration procedure

1) Configure PE1

Enable MPLS and LDP.

```
[PE1] mpls lsr-id 50.1.1.1
[PE1] mpls
[PE1-mpls] quit
[PE1] mpls ldp
```

Configure VPN-instance.

```
[PE1] ip vpn-instance vpn1
[PE1-vpn-vpn1] route-distinguisher 2:1
[PE1-vpn-vpn1] vpn-target 100:1 export-extcommunity
[PE1-vpn-vpn1] vpn-target 100:1 import-extcommunity
```

Configure VLAN interface.

```
[PE1] vlan 203
[PE1-vlan203] port gigabitethernet 2/1/3
[PE1-vlan203] quit
[PE1] interface Vlan-interface 203
[PE1-Vlan-interface203] ip address 168.1.12.1 255.255.255.0
[PE1-Vlan-interface203] mpls
[PE1-Vlan-interface203] mpls ldp enable
[PE1-Vlan-interface203] quit
[PE1] vlan 201
[PE1-vlan201] port gigabitethernet 2/1/1
[PE1-vlan201] quit
[PE1] interface Vlan-interface 201
[PE1-Vlan-interface201] ip binding vpn-instance vpn1
[PE1-Vlan-interface201] ip address 10.1.1.2 255.255.255.0
[PE1-Vlan-interface201] ospf cost 1
[PE1-Vlan-interface201] quit
[PE1] vlan 202
[PE1-vlan202] port gigabitethernet 2/1/2
[PE1-vlan202] quit
[PE1] interface Vlan-interface 202
[PE1-Vlan-interface202] ip address 168.1.13.1 255.255.255.0
[PE1-Vlan-interface202] ospf cost 1
[PE1-Vlan-interface202] mpls
[PE1-Vlan-interface202] mpls ldp enable
[PE1-Vlan-interface202] mpls ldp transport-ip interface
[PE1-Vlan-interface202] quit
[PE1] interface loopback0
[PE1-LoopBack0] ip binding vpn-instance vpn1
[PE1-LoopBack0] ip address 1.1.1.1 255.255.255.255
[PE1-LoopBack0] quit
[PE1] interface loopback1
[PE1-LoopBack1] ip address 50.1.1.1 255.255.255.255
```

Configure BGP peer.

```
[PE1] bgp 100
[PE1-bgp] undo synchronization
[PE1-bgp] group fc internal
[PE1-bgp] peer 50.1.1.2 group fc
[PE1-bgp] peer 50.1.1.2 connect-interface LoopBack1
[PE1-bgp] peer 50.1.1.3 group fc
```

Configure BGP and import OSPF routing and direct-connect route.

```
[PE1-bgp] ipv4-family vpn-instance vpn1
[PE1-bgp-af-vpn-instance] import-route ospf 100
```

```
[PE1-bgp-af-vpn-instance] import-route ospf-ase 100
[PE1-bgp-af-vpn-instance] import-route ospf-nssa 100
[PE1-bgp-af-vpn-instance] import-route direct
[PE1-bgp-af-vpn-instance] undo synchronization
```

Create and activate peer in MBGP.

```
[PE1-bgp-af-vpn] ipv4-family vpnv4
[PE1-bgp-af-vpn] peer fc enable
[PE1-bgp-af-vpn] peer fc advertise-community
[PE1-bgp-af-vpn] peer 50.1.1.2 group fc
```

Bind OSPF process to VPN-instance.

```
[PE1] ospf 100 router-id 1.1.1.1 vpn-instance vpn1
[PE1-ospf-100] import-route bgp
[PE1-ospf-100] area 0.0.0.0
[PE1-ospf-100-area-0.0.0.0] network 10.1.1.0 0.0.0.255
```

Configuring sham link

```
[PE1-ospf-100-area-0.0.0.1] sham-link 1.1.1.1 2.2.2.2
```

Configure the routes distributed to PE2 and PE3.

```
[PE1] ospf 1000
[PE1-ospf-1000] area 0
[SW8800-ospf-1000-area-0.0.0.0] network 168.12.1.0 0.0.0.255
[SW8800-ospf-1000-area-0.0.0.0] network 50.1.1.1 0.0.0.0
```

2) Configure PE2

Enable MPLS and LDP.

```
[PE2] mpls lsr-id 50.1.1.2
[PE2] mpls
[PE2-mpls] quit
[PE2] mpls ldp
```

Configure vpn-instance vpn1.

```
[PE2] ip vpn-instance vpn1
[PE2-vpn-vpn1] route-distinguisher 2:1
[PE2-vpn-vpn1] vpn-target 100:1 export-extcommunity
[PE2-vpn-vpn1] vpn-target 100:1 import-extcommunity
```

Configure VLAN interface.

```
[PE2] vlan 203
[PE2-vlan203] port gigabitethernet 2/1/3
[PE2-vlan203] quit
[PE2] interface Vlan-interface 203
[PE2-Vlan-interface203] ip address 168.1.12.2 255.255.255.0
[PE2-Vlan-interface203] mpls
```

```
[PE2-Vlan-interface203] mpls ldp enable
[PE2-Vlan-interface203] quit
[PE2] vlan 201
[PE2-vlan201] port gigabitethernet 2/1/1
[PE2-vlan201] quit
[PE2] interface Vlan-interface 201
[PE2-Vlan-interface201] ip binding vpn-instance vpn1
[PE2-Vlan-interface201] ip address 20.1.1.2 255.255.255.0
[PE2-Vlan-interface201] ospf cost 1
[PE2-Vlan-interface201] quit
[PE2] vlan 202
[PE2-vlan202] port gigabitethernet 2/1/2
[PE2-vlan202] quit
[PE2] interface Vlan-interface 202
[PE2-Vlan-interface202] ip address 168.1.23.2 255.255.255.0
[PE2-Vlan-interface202] ospf cost 1
[PE2-Vlan-interface202] mpls
[PE2-Vlan-interface202] mpls ldp enable
[PE2-Vlan-interface202] quit
[PE2] interface LoopBack0
[PE2-LoopBack0] ip binding vpn-instance vpn1
[PE2-LoopBack0] ip address 2.2.2.2 255.255.255.255
[PE2-LoopBack0] quit
[PE2] interface LoopBack1
[PE2-LoopBack1] ip address 50.1.1.2 255.255.255.255
```

Configure BGP.

```
[PE2] bgp 100
[PE2-bgp] undo synchronization
[PE2-bgp] group fc internal
[PE2-bgp] peer 50.1.1.1 group fc
[PE2-bgp] peer 50.1.1.1 connect-interface LoopBack1
[PE2-bgp] peer 50.1.1.3 group fc
```

Configure VPN-instance and import OSPF and direct-connect route.

```
[PE2-bgp] ipv4-family vpn-instance vpn1
[PE2-bgp-af-vpn-instance] import-route direct
[PE2-bgp-af-vpn-instance] import-route ospf-nssa 100
[PE2-bgp-af-vpn-instance] import-route ospf-ase 100
[PE2-bgp-af-vpn-instance] import-route ospf 100
[PE2-bgp-af-vpn-instance] undo synchronization
```

Configure MBGP and enable peer.

```
[PE2-bgp-af-vpn] ipv4-family vpnv4
```

```
[PE2-bgp-af-vpn] peer fc enable
[PE2-bgp-af-vpn] peer fc advertise-community
[PE2-bgp-af-vpn] peer 50.1.1.1 group fc
```

Configure OSPF and import BGP and direct-connect route.

```
[PE2] ospf 100 router-id 2.2.2.2 vpn-instance vpn1
[PE2-ospf-100] import-route bgp
[PE2-ospf-100] import-route static
[PE2-ospf-100] area 0.0.0.0
[PE2-ospf-100-area-0.0.0.0] network 20.1.1.0 0.0.0.255
```

Configuring sham link

```
[PE2-ospf-100-area-0.0.0.0] sham-link 2.2.2.2 1.1.1.1
```

Configure static route to PE1 and PE3.

```
[PE2] ip route-static 50.1.1.1 255.255.255.255 168.1.12.1
[PE2] ip route-static 50.1.1.3 255.255.255.255 168.1.23.3
```

Configure the routes distributed to PE1 and PE3.

[PE1] ospf 1000

```
[PE1-ospf-1000] area 0
[SW8800-ospf-1000-area-0.0.0.0] network 168.12.1.0 0.0.0.255
[SW8800-ospf-1000-area-0.0.0.0] network 50.1.1.1 0.0.0.0
```

3) Configure CE1.

Configure interfaces

```
[CE1] vlan 202
[CE1-vlan202] port gigabitethernet 2/1/2
[CE1-vlan202] quit
[CE1] interface Vlan-interface 202
[CE1-Vlan-interface202] ip address 12.1.1.1 255.255.255.0
[CE1-Vlan-interface202] ospf cost 100
[CE1-Vlan-interface202] quit
[CE1] vlan 201
[CE1-vlan201] port gigabitethernet 2/1/1
[CE1-vlan201] quit
[CE1] interface Vlan-interface 201
[CE1-Vlan-interface201] ip address 10.1.1.1 255.255.255.0
[CE1-Vlan-interface201] ospf cost 1
```

Configure OSPF.

```
[CE1] ospf 100 router-id 10.10.10.129
[CE1-ospf-100] import-route direct
[CE1-ospf-100] area 0.0.0.0
[CE1-ospf-100-area-0.0.0.0] network 10.1.1.0 0.0.0.255
```



```
[CE1-ospf-100-area-0.0.0.0] network 12.1.1.0 0.0.0.255
```

4) Configure CE2

Configure interface.

```
[CE2] vlan 202
[CE2-vlan202] port gigabitethernet 2/1/2
[CE2-vlan202] quit
[CE2] interface Vlan-interface 202
[CE2-Vlan-interface202] ip address 12.1.1.2 255.255.255.0
[CE2-Vlan-interface202] ospf cost 100
[CE2-Vlan-interface202] quit
[CE2] vlan 201
[CE2-vlan201] port gigabitethernet 2/1/1
[CE2-vlan201] quit
[CE2] interface Vlan-interface 201
[CE2-Vlan-interface201] ip address 20.1.1.1 255.255.255.0
[CE2-Vlan-interface201] ospf cost 1
```

Configure OSPF.

```
[CE2] ospf 100 router-id 20.20.20.20
[CE2-ospf-100] area 0.0.0.0
[CE2-ospf-100-area-0.0.0.0] network 12.1.1.0 0.0.0.255
[CE2-ospf-100-area-0.0.0.0] network 20.1.1.0 0.0.0.255
```

37.4.10 Nested BGP/MPLS VPN Configuration Example

I. Network requirements

A VPN user has multiple nodes to access the service provider's BGP/MPLS backbone network. And this VPN is divided into three sub-VPNs: VPN1, VPN2 and VPN3.

Some of the nodes of these sub-VPNs directly access a PE in the network, and some access a PE through the father VPN. That is, the adopted network structure is unsymmetrical.

This example mainly describes the configuration of VPN1; the configuration of other sub-VPNs is similar.

II. Network diagram

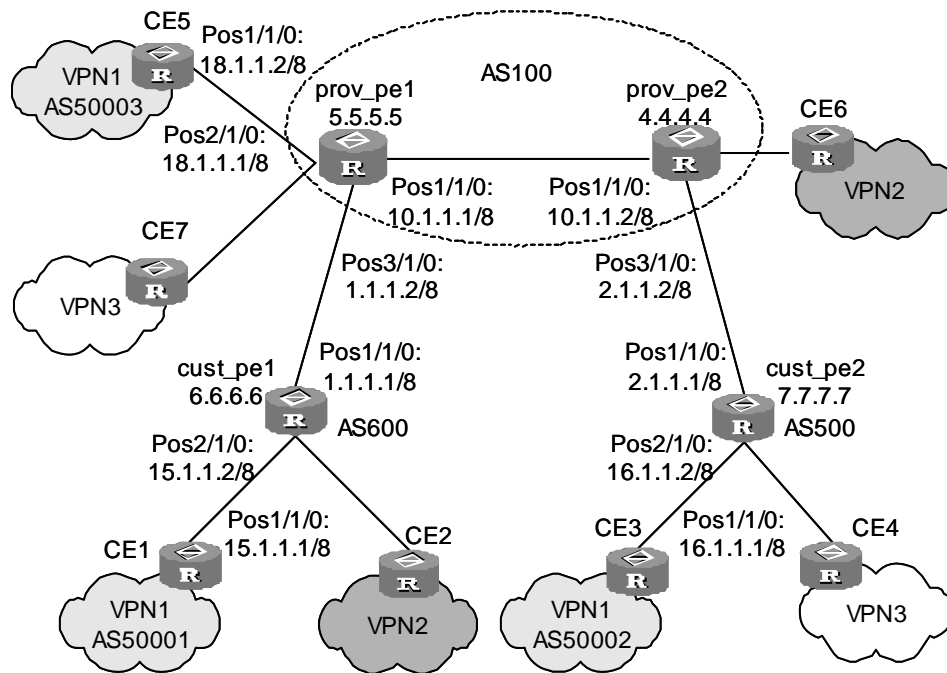


Figure 37-17 Network diagram for nested VPN

III. Configuration procedure

Note:

This procedure omits part of the configuration for CE router.

- 1) Configure IGP on the service provider's backbone network.

Configure prov_pe1

```
<SW8800> system-view
[SW8800] sysname prov_pe1
[prov_pe1] interface LoopBack0
[prov_pe1-LoopBack0] ip address 5.5.5.5 255.255.255.255
[prov_pe1-LoopBack0] quit
[prov_pe1] interface pos 1/1/0
[prov_pe1-Pos1/1/0] link-protocol ppp
[prov_pe1-Pos1/1/0] ip address 10.1.1.1 255.0.0.0
[prov_pe1-Pos1/1/0] quit
[prov_pe1] ospf
[prov_pe1-ospf] area 0
[prov_pe1-ospf-area-0.0.0.0] network 5.5.5.5 0.0.0.0
```

```
[prov_pe1-ospf-area-0.0.0.0] network 10.0.0.0 0.255.255.255
```

Configure prov_pe2

```
<SW8800> system-view
[SW8800] sysname prov_pe2
[prov_pe2] interface LoopBack0
[prov_pe2-LoopBack0] ip address 4.4.4.4 255.255.255.255
[prov_pe2-LoopBack0] quit
[prov_pe2] interface pos 1/1/0
[prov_pe2-Pos1/1/0] link-protocol ppp
[prov_pe2-Pos1/1/0] ip address 10.1.1.2 255.0.0.0
[prov_pe2] ospf
[prov_pe2-ospf] area 0
[prov_pe2-ospf-area-0.0.0.0] network 4.4.4.4 0.0.0.0
[prov_pe2-ospf-area-0.0.0.0] network 10.0.0.0 0.255.255.255
```

Configure basic MPLS capability and MPLS LDP on the backbone network.

Configure prov_pe1

```
[prov_pe1] mpls lsr-id 5.5.5.5
[prov_pe1] mpls ldp
[prov_pe1] interface pos 1/1/0
[prov_pe1-Pos1/1/0] mpls
[prov_pe1-Pos1/1/0] mpls ldp
[prov_pe1-Pos1/1/0] quit
```

Configure prov_pe2

```
[prov_pe2] mpls lsr-id 4.4.4.4
[prov_pe2] mpls ldp
[prov_pe2] interface pos 1/1/0
[prov_pe2-Pos1/1/0] mpls
[prov_pe2-Pos1/1/0] mpls ldp
[prov_pe2-Pos1/1/0] quit
```

Configure IBGP between provider PEs.

Configure prov_pe1

```
[prov_pe1] bgp 100
[prov_pe1-bgp] group ibgp internal
[prov_pe1-bgp] peer 4.4.4.4 group ibgp
[prov_pe1-bgp] peer 4.4.4.4 connect-interface LoopBack0
[prov_pe1-bgp] ipv4-family vpnv4
[prov_pe1-bgp-af-vpn] peer ibgp enable
[prov_pe1-bgp-af-vpn] peer ibgp next-hop-local
[prov_pe1-bgp-af-vpn] peer 4.4.4.4 group ibgp
[prov_pe1-bgp-af-vpn] quit
```

```
[prov_pe1-bgp] quit
```

Configure prov_pe2

```
[prov_pe2] bgp 100
[prov_pe2-bgp] group ibgp internal
[prov_pe2-bgp] peer 5.5.5.5 group ibgp
[prov_pe2-bgp] peer 5.5.5.5 connect-interface LoopBack0
[prov_pe2-bgp] ipv4-family vpnv4
[prov_pe2-bgp-af-vpn] peer ibgp enable
[prov_pe2-bgp-af-vpn] peer ibgp next-hop-local
[prov_pe2-bgp-af-vpn] peer 5.5.5.5 group ibgp
[prov_pe2-bgp-af-vpn] quit
[prov_pe2-bgp] quit
```

- 2) Create a VPN instance on provider PEs to access customer PEs and directly connected user CEs.

Configure prov_pe1

```
[prov_pe1] ip vpn-instance customer_vpn
[prov_pe1-vpn-instance] route-distinguisher 3:3
[prov_pe1-vpn-instance] vpn-target 3:3
[prov_pe1-vpn-instance] quit
[prov_pe1] ip vpn-instance vpn1
[prov_pe1-vpn-instance] route-distinguisher 1:1
[prov_pe1-vpn-instance] vpn-target 1:1
[prov_pe1-vpn-instance] vpn-target 3:3
[prov_pe1-vpn-instance] quit
[prov_pe1] interface pos 3/1/0
[prov_pe1-Pos3/1/0] ip binding vpn-instance customer_vpn
[prov_pe1-Pos3/1/0] link-protocol ppp
[prov_pe1-Pos3/1/0] ip address 1.1.1.2 255.0.0.0
[prov_pe1-Pos3/1/0] mpls
[prov_pe1-Pos3/1/0] quit
[prov_pe1] interface pos 2/1/0
[prov_pe1-Pos2/1/0] ip binding vpn-instance vpn1
[prov_pe1-Pos2/1/0] link-protocol ppp
[prov_pe1-Pos2/1/0] ip address 18.1.1.1 255.0.0.0
[prov_pe1-Pos2/1/0] quit
```

Configure prov_pe2

```
[prov_pe2] ip vpn-instance customer_vpn
[prov_pe2-vpn-instance] route-distinguisher 3:3
[prov_pe2-vpn-instance] vpn-target 3:3
[prov_pe2-vpn-instance] quit
[prov_pe2] interface pos 3/1/0
```

```
[prov_pe2-Pos3/1/0] ip binding vpn-instance customer_vpn
[prov_pe2-Pos3/1/0] link-protocol ppp
[prov_pe2-Pos3/1/0] ip address 2.1.1.2 255.0.0.0
[prov_pe2-Pos3/1/0] mpls
[prov_pe2-Pos3/1/0] quit
```

Configure cust_pe1

```
<SW8800> system-view
[SW8800] sysname cust_pe1
[cust_pe1] interface LoopBack0
[cust_pe1-LoopBack0] ip address 6.6.6.6 255.255.255.255
[cust_pe1-LoopBack0] quit
[cust_pe1] mpls lsr-id 6.6.6.6
[cust_pe1] interface pos 1/1/0
[cust_pe1-Pos1/1/0] link-protocol ppp
[cust_pe1-Pos1/1/0] ip address 1.1.1.1 255.0.0.0
[cust_pe1-Pos1/1/0] mpls
[cust_pe1-Pos1/1/0] quit
```

Configure cust_pe2

```
<SW8800> system-view
[SW8800] sysname cust_pe2
[cust_pe2] interface LoopBack0
[cust_pe2-LoopBack0] ip address 7.7.7.7 255.255.255.255
[cust_pe2-LoopBack0] quit
[cust_pe2] mpls lsr-id 7.7.7.7
[cust_pe2] interface pos 1/1/0
[cust_pe2-Pos1/1/0] link-protocol ppp
[cust_pe2-Pos1/1/0] ip address 2.1.1.1 255.0.0.0
[cust_pe2-Pos1/1/0] mpls
[cust_pe2-Pos1/1/0] quit
```

3) Configure EBGP between provider PE and customer PE.

Configure prov_pe1 to access the corresponding Customer PE.

```
[prov_pe1] route-policy comm permit node 10
[prov_pe1-route-policy-comm-10] if-match vpn-target 1:1
[prov_pe1-route-policy-comm-10] quit
[prov_pe1] bgp 100
[prov_pe1-bgp] ipv4-family vpn-instance customer_vpn
[prov_pe1-bgp-af-vpn-instance] group ebgp external
[prov_pe1-bgp-af-vpn-instance] undo peer ebgp enable
[prov_pe1-bgp-af-vpn-instance] peer 1.1.1.1 group ebgp as-number 600
[prov_pe1-bgp] ipv4-family vpnv4
[prov_pe1-bgp-af-vpn] nesting-vpn
```

```
[prov_pe1-bgp-af-vpn] peer ebgp vpn-instance customer_vpn enable
[prov_pe1-bgp-af-vpn] peer 1.1.1.1 vpn-instance customer_vpn group ebgp
[prov_pe1-bgp-af-vpn] peer 1.1.1.1 vpn-instance customer_vpn route-policy
comm import
[prov_pe1-bgp-af-vpn] quit
```

Configure prov_pe1 to access CE5

```
[prov_pe1-bgp] ipv4-family vpn-instance vpn1
[prov_pe1-bgp-af-vpn-instance] group ebgp external
[prov_pe1-bgp-af-vpn-instance] peer 18.1.1.2 group ebgp as-number 50003
```

Configure prov_pe2 to access the corresponding Customer PE.

```
[prov_pe2] route-policy com2 permit node 10
[prov_pe2-route-policy-com2-10] if-match vpn-target 1:1
[prov_pe2-route-policy-com2-10] quit
[prov_pe2] bgp 100
[prov_pe2-bgp] ipv4-family vpn-instance customer_vpn
[prov_pe2-bgp-af-vpn-instance] group ebgp external
[prov_pe2-bgp-af-vpn-instance] undo peer ebgp enable
[prov_pe2-bgp-af-vpn-instance] peer 2.1.1.1 group ebgp as-number 500
[prov_pe2-bgp] ipv4-family vpnv4
[prov_pe2-bgp-af-vpn] nesting-vpn
[prov_pe2-bgp-af-vpn] peer ebgp vpn-instance customer_vpn enable
[prov_pe2-bgp-af-vpn] peer 2.1.1.1 vpn-instance customer_vpn group ebgp
[prov_pe2-bgp-af-vpn] peer 2.1.1.1 vpn-instance customer_vpn route-policy
com2 import
```

Configure cust_pe1

```
[cust_pe1] bgp 600
[cust_pe1-bgp] group ebgp external
[cust_pe1-bgp] undo peer ebgp enable
[cust_pe1-bgp] peer 1.1.1.2 group ebgp as-number 100
[cust_pe1-bgp] ipv4-family vpnv4
[cust_pe1-bgp-af-vpn] peer ebgp enable
[cust_pe1-bgp-af-vpn] peer 1.1.1.2 group ebgp
```

Configure cust_pe2

```
[cust_pe2] bgp 500
[cust_pe2-bgp] group ebgp external
[cust_pe2-bgp] undo peer ebgp enable
[cust_pe2-bgp] peer 2.1.1.2 group ebgp as-number 100
[cust_pe2-bgp] ipv4-family vpnv4
[cust_pe2-bgp-af-vpn] peer ebgp enable
[cust_pe2-bgp-af-vpn] peer 2.1.1.2 group ebgp
```

- 4) On each Customer PE, configure the sub-VPN that accesses the network through the Customer PE.

Configure cust_pe1

```
[cust_pe1] ip vpn-instance vpn1
[cust_pe1-vpn-instance] route-distinguisher 1:1
[cust_pe1-vpn-instance] vpn-target 1:1
[cust_pe1-vpn-instance] quit
[cust_pe1] interface pos 2/1/0
[cust_pe1-Pos2/1/0] ip binding vpn-instance vpn1
[cust_pe1-Pos2/1/0] link-protocol ppp
[cust_pe1-Pos2/1/0] ip address 15.1.1.2 255.0.0.0
[cust_pe1-Pos2/1/0] quit
[cust_pe1] bgp 600
[cust_pe1-bgp] undo peer ebgp enable
[cust_pe1-bgp] ipv4-family vpn-instance vpn1
[cust_pe1-bgp-af-vpn-instance] group cegroup external
[cust_pe1-bgp-af-vpn-instance] peer 15.1.1.1 group cegroup as-number 50001
[cust_pe1-bgp-af-vpn-instance] quit
[cust_pe1-bgp] quit
```

Configure cust_pe2

```
[cust_pe2] ip vpn-instance vpn1
[cust_pe2-vpn-instance] route-distinguisher 1:1
[cust_pe2-vpn-instance] vpn-target 1:1
[cust_pe2] interface pos 2/1/0
[cust_pe2-Pos2/1/0] ip binding vpn-instance vpn1
[cust_pe2-Pos2/1/0] link-protocol ppp
[cust_pe2-Pos2/1/0] ip address 16.1.1.2 255.0.0.0
[cust_pe2-Pos2/1/0] quit
[cust_pe2] bgp 500
[cust_pe2-bgp] undo peer ebgp enable
[cust_pe2-bgp] ipv4-family vpn-instance vpn1
[cust_pe2-bgp-af-vpn-instance] group cegroup external
[cust_pe2-bgp-af-vpn-instance] peer 16.1.1.1 group cegroup as-number 50002
[cust_pe2-bgp-af-vpn-instance] quit
[cust_pe2-bgp] quit
```

37.4.11 OSPF Multi-instance CE Configuration Example

I. Network requirements

CE router in a VPN achieves service isolation by configuring multiple VPN instances.

II. Network diagram

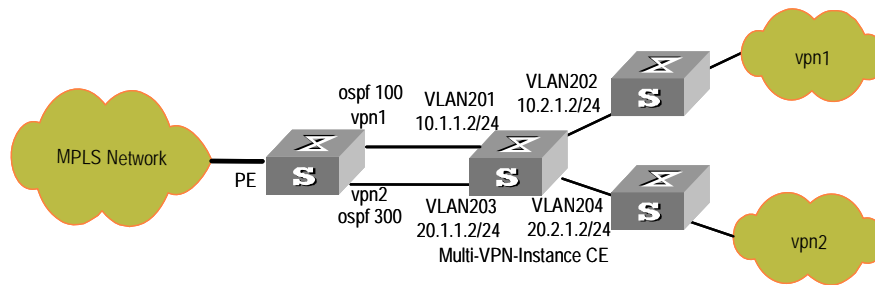


Figure 37-18 Network diagram for OSPF multi-instance CE configuration

III. Configuration procedure

1) Configuring CE router

Configure instance vpn1

```
[CE] ip vpn-instance vpn1
[CE-vpn-vpn1] route-distinguisher 100:1
[CE-vpn-vpn1] vpn-target 100:1 export-extcommunity
[CE-vpn-vpn1] vpn-target 100:1 import-extcommunity
```

Configure instance vpn2

```
[CE] ip vpn-instance vpn2
[CE-vpn-vpn2] route-distinguisher 200:1
[CE-vpn-vpn2] vpn-target 200:1 export-extcommunity
[CE-vpn-vpn2] vpn-target 200:1 import-extcommunity
```

Configure VLAN201

```
[CE] vlan 201
[CE-vlan201] port gigabitethernet 2/1/1
[CE-vlan201] quit
[CE] interface Vlan-interface 201
[CE-Vlan-interface201] ip binding vpn-instance vpn1
[CE-Vlan-interface201] ip address 10.1.1.2 255.255.255.0
```

Configure VLAN202

```
[CE] vlan 202
[CE-vlan202] port gigabitethernet 2/1/2
[CE-vlan202] quit
[CE] interface Vlan-interface 202
[CE-Vlan-interface202] ip binding vpn-instance vpn1
[CE-Vlan-interface202] ip address 10.2.1.2 255.255.255.0
[CE-Vlan-interface202] ospf cost 100
```

Configure VLAN203


```
[CE] vlan 203
[CE-vlan203] port gigabitethernet 2/1/3
[CE-vlan203] quit
[CE] interface Vlan-interface 203
[CE-Vlan-interface203] ip binding vpn-instance vpn2
[CE-Vlan-interface203] ip address 20.1.1.2 255.255.255.0
```

Configure VLAN204

```
[CE] vlan 204
[CE-vlan204] port gigabitethernet 2/1/4
[CE-vlan204] quit
[CE] interface Vlan-interface 204
[CE-Vlan-interface204] ip binding vpn-instance vpn2
[CE-Vlan-interface204] ip address 20.2.1.2 255.255.255.0
```

Configure ospf 100

```
[CE] ospf 100 vpn-instance vpn1
[CE-ospf-100] vpn-instance-capability simple
[CE-ospf-100] area 0.0.0.0
[CE-ospf-100-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[CE-ospf-100-area-0.0.0.0] network 10.2.1.0 0.0.0.255
```

Configure ospf 300

```
[CE] ospf 300 vpn-instance vpn2
[CE-ospf-300] vpn-instance-capability simple
[CE-ospf-300] area 0.0.0.1
[CE-ospf-300-area-0.0.0.1] network 20.1.1.0 0.0.0.255
[CE-ospf-300-area-0.0.0.1] network 20.2.1.0 0.0.0.255
```

37.4.12 Multi-Role Host Configuration Example

I. Network requirements

CE1 and CE3 belong to VPN1, and CE2 belong to VPN2.

The host PC2 with the IP address of 172.16.0.1 accesses the network through CE2. As a multi-role host, it can access both VPN1 and VPN2.

II. Network diagram

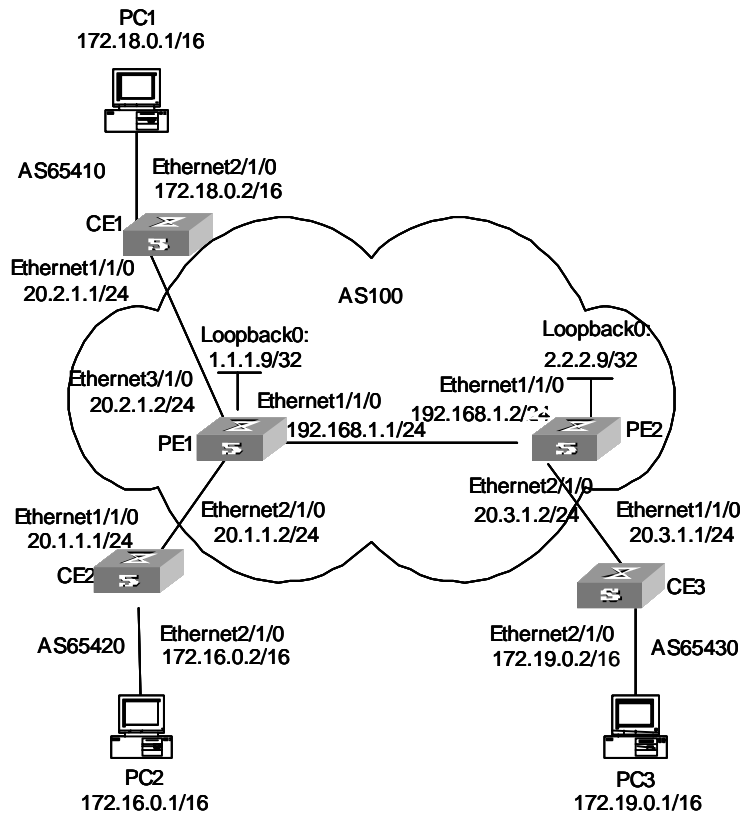


Figure 37-19 Network diagram for multi-role host application

III. Configuration procedure

- 1) Configure OSPF as the IGP protocol on the MPLS backbone network.

Configure OSPF on PE1:

```
[PE1] interface loopback 0
[PE1-LoopBack0] ip address 1.1.1.9 32
[PE1-LoopBack0] quit
[PE1] interface Ethernet1/1/0
[PE1-Ethernet1/1/0] ip address 192.168.1.1 24
[PE1-Ethernet1/1/0] quit
[PE1] ospf
[PE1-ospf-1] area 0
[PE1-ospf-1-area-0.0.0.0] network 192.168.1.0 0.0.0.255
[PE1-ospf-1-area-0.0.0.0] network 1.1.1.9 0.0.0.0
[PE1-ospf-1-area-0.0.0.0] quit
[PE1-ospf-1] quit
```

Configure OSPF on PE2:

```
[PE2] interface loopback 0
```

```
[PE2-LoopBack0] ip address 2.2.2.9 32
[PE2-LoopBack0] quit
[PE2] interface Ethernet1/1/0
[PE2-Ethernet1/1/0] ip address 192.168.1.2 24
[PE2-Ethernet1/1/0] quit
[PE2] ospf
[PE2-ospf-1] area 0
[PE2-ospf-1-area-0.0.0.0] network 192.168.1.0 0.0.0.255
[PE2-ospf-1-area-0.0.0.0] network 2.2.2.9 0.0.0.0
[PE2-ospf-1-area-0.0.0.0] quit
[PE2-ospf-1] quit
```

2) Configure basic MPLS capability and create VPN instances.

Configure basic MPLS capability on PE1:

```
[PE1] mpls lsr-id 1.1.1.9
[PE1-mpls] lsp-trigger all
[PE1-mpls] quit
[PE1] mpls ldp
[PE1-mpls-ldp] quit
[PE1] interface Ethernet1/1/0
[PE1-Ethernet1/1/0] mpls
[PE1-Ethernet1/1/0] mpls ldp
[PE1-Ethernet1/1/0] quit
```

Create VPN instances for VPN1 and VPN2 on PE1, bind Ethernet3/1/0 to VPN1, and bind Ethernet2/1/0 to VPN2.

```
[PE1] ip vpn-instance vpn1
[PE1-vpn-vpn1] route-distinguisher 100:1
[PE1-vpn-vpn1] vpn-target 100:1 both
[PE1-vpn-vpn1] quit
[PE1] ip vpn-instance vpn2
[PE1-vpn-vpn2] route-distinguisher 100:2
[PE1-vpn-vpn2] vpn-target 100:2 both
[PE1-vpn-vpn2] quit
[PE1] interface Ethernet3/1/0
[PE1-Ethernet3/1/0] ip binding vpn-instance vpn1
[PE1-Ethernet3/1/0] ip address 20.2.1.2 24
[PE1-Ethernet3/1/0] quit
[PE1] interface Ethernet2/1/0
[PE1-Ethernet2/1/0] ip binding vpn-instance vpn2
[PE1-Ethernet2/1/0] ip address 20.1.1.2 24
[PE1-Ethernet2/1/0] quit
```

Configure basic MPLS capability on PE2:

```
[PE2] mpls lsr-id 2.2.2.9
[PE2-mpls] lsp-trigger all
[PE2-mpls] quit
[PE2] mpls ldp
[PE2-mpls-ldp] quit
[PE2] interface Ethernet1/1/0
[PE2-Ethernet1/1/0] mpls
[PE2-Ethernet1/1/0] mpls ldp
[PE2-Ethernet1/1/0] quit
```

Create a VPN instance for VPN1 on PE2, and bind Ethernet2/1/0 to VPN1.

```
[PE2] ip vpn-instance vpn1
[PE2-vpn-vpn1] route-distinguisher 300:1
[PE2-vpn-vpn1] vpn-target 100:1 both
[PE2-vpn-vpn1] quit
[PE2] interface Ethernet2/1/0
[PE2-Ethernet2/1/0] ip binding vpn-instance vpn1
[PE2-Ethernet2/1/0] ip address 20.3.1.2 24
[PE2-Ethernet2/1/0] quit
```

Configure BGP.

Configure CE1:

```
[CE1] interface Ethernet1/1/0
[CE1-Ethernet1/1/0] ip address 20.2.1.1 24
[CE1-Ethernet1/1/0] quit
[CE1] bgp 65410
[CE1-bgp] import-route direct
[CE1-bgp] group 10 external
[CE1-bgp] peer 20.2.1.2 group 10 as-number 100
[CE1-bgp] quit
```

Configure CE2:

```
[CE2] interface Ethernet1/1/0
[CE2-Ethernet1/1/0] ip address 20.1.1.1 24
[CE2-Ethernet1/1/0] quit
[CE2] bgp 65420
[CE2-bgp] import-route direct
[CE2-bgp] group 10 external
[CE2-bgp] peer 20.1.1.2 group 10 as-number 100
[CE2-bgp] quit
```

Configure CE3:

```
[CE3] interface Ethernet1/1/0
[CE3-Ethernet1/1/0] ip address 20.3.1.1 24
```

```
[CE3-Ethernet1/1/0] quit
[CE3] bgp 65430
[CE3-bgp] import-route direct
[CE3-bgp] group 10 external
[CE3-bgp] peer 20.3.1.2 group 10 as-number 100
[CE3-bgp] quit
```

Configure PE1: set up IBGP peer relation with PE2 in BGP-VPNv4 sub-address family view; set up EBGP peer relation with CE2 in BGP-VPN instance view.

```
[PE1] bgp 100
[PE1-bgp] group 10
[PE1-bgp] peer 2.2.2.9 group 10
[PE1-bgp] peer 2.2.2.9 connect-interface loopback 0
[PE1-bgp] ipv4-family vpnv4
[PE1-bgp-af-vpn] peer 10 enable
[PE1-bgp-af-vpn] peer 2.2.2.9 group 10
[PE1-bgp-af-vpn] quit
[PE1-bgp] ipv4-family vpn-instance vpn1
[PE1-bgp-af-vpn-instance] import-route direct
[PE1-bgp-af-vpn-instance] group 20 external
[PE1-bgp-af-vpn-instance] peer 20.2.1.1 group 20 as-number 65410
[PE1-bgp-af-vpn-instance] quit
[PE1-bgp] ipv4-family vpn-instance vpn2
[PE1-bgp-af-vpn-instance] import-route direct
[PE1-bgp-af-vpn-instance] group 30 external
[PE1-bgp-af-vpn-instance] peer 20.1.1.1 group 30 as-number 65420
[PE1-bgp-af-vpn-instance] quit
[PE1-bgp] quit
```

Configure PE2: set up IBGP peer relation with PE1 in BGP-VPNv4 sub-address family view; set up EBGP peer relation with CE3 in BGP-VPN instance view.

```
[PE2] bgp 100
[PE2-bgp] group 10
[PE2-bgp] peer 1.1.1.9 group 10
[PE2-bgp] peer 1.1.1.9 connect-interface loopback 0
[PE2-bgp] ipv4-family vpnv4
[PE2-bgp-af-vpn] peer 10 enable
[PE2-bgp-af-vpn] peer 1.1.1.9 group 10
[PE2-bgp-af-vpn] quit
[PE2-bgp] ipv4-family vpn-instance vpn1
[PE2-bgp-af-vpn-instance] import-route direct
[PE2-bgp-af-vpn-instance] group 20 external
[PE2-bgp-af-vpn-instance] peer 20.3.1.1 group 20 as-number 65430
[PE2-bgp-af-vpn-instance] quit
```

```
[PE2-bgp] quit
```

Configure multi-role host feature.

Configure a default route pointing to PE1 on CE2.

```
[CE2] ip route-static 0.0.0.0 0.0.0.0 20.1.1.2
```

If routing protocol is used between CE2 and PE1, to avoid routing loop, disable PE1 from advertising any route to CE2. In this example, no routing protocol is used between PE1 and CE2; so, a static route for PC2 is directly configured on PE1 (you can also adopt other similar configuration here).

```
[PE1] ip route-static vpn-instance vpn2 172.16.0.0 16 20.1.1.1
```

Import the route of VPN1 to VPN2 using the VPN Target attribute.

```
[PE1] ip vpn-instance vpn2
```

```
[PE1-vpn-vpn2] vpn-target 100:1 import-extcommunity
```

37.5 Troubleshooting

I. Symptom 1

In central server topology networking mode, the local end switch (spoke PE) cannot learn the routing information of the peer end switch (spoke PE).

Solution:

- Check whether the BGP adjacent of spoke PE and hub PE is created correctly.
- Check whether the routing attributes import/export relation of each VPN-instance is correct.
- Check from the hub PE that whether the routing information between two VPN instances can be learnt by each other. if not, perform the following operation: check if the EBGP protocol runs between hub PE and hub CE, check whether the **peer peer-address allow-as-loop** command is configured between PE and CE.

II. Symptom 2

PE at the local end can learn private network route of the PE at peer end, but two PEs cannot intercommunicate with each other.

Solution:

- Check whether the loopback interface configured on the PE has the address with 32-bit mask.
- Check whether the tag of private network route is correct.
- Check whether the LDP session is established using the **display mpls ldp session** command.
- Check whether the LSP tunnel is established using the **display mpls lsp** command.

III. Symptom 3

In Hub&Spoke networking mode, spoke PE cannot learn the private networking route of Hub PE.

Solution:

- Check whether the LSP tunnel is established using the **display mpls lsp** command.
- Check whether the BGP adjacent is established correctly.
- Check whether the routing import/export relation of the VPN-instance is correct.
- Check whether allow-as-loop is configured between spoke PE and hub PE.

IV. Symptom 4

Fail to specify the loopback interface at the peer end as the BGP neighbor.

Solution:

- Check whether the local routing table has learnt the loopback interface routing information of the peer end using the **display ip routing-table** command.
- Check whether the address of the loopback interface at the peer end can be pinged using the **ping** command.
- Check whether the configuration information is correct using the **display current-configuration bgp** command; confirm that you have specified the local loopback interface as the interface to create adjacent interface with the peer end by using the **peer peer-address connect-interface** command; confirm that you have activate the neighbor in VPNv4 sub-address family view.
- Check whether the BGP information is correct on the PE at the peer end; check whether specified the local loopback interface as the interface to create adjacent with the peer end; and check whether you have configured VPN capacity.

V. Symptom 5

During ASBR configuration, VPN route interior label does not switch on the ASBR.

Solution:

- Check whether the VPN neighbor is created correctly using the **display bgp vpnv4 all peer** command.
- Check whether ASBR is configured with the **undo policy vpn-target** command. If not, configure this command.

Chapter 38 MSTP Region-configuration

38.1 Introduction to MSTP

MSTP stands for Multiple Spanning Tree Protocol, which is compatible with Spanning Tree Protocol (STP) and Rapid Spanning Tree Protocol (RSTP).

STP is not fast in state transition. Even on a point-to-point link or a edge port, it has to take an interval twice as long as forward delay before the port transits to the forwarding state.

RSTP converges fast, but has the following drawback like STP: all the network bridges in a LAN share one spanning tree and the redundant links cannot be blocked based on VLANs. Packets of all VLANs are forwarded along one spanning tree.

MSTP makes up for the drawback of STP and RSTP. It not only converges fast, but also allows the traffic of different VLANs to be distributed along their respective paths, which provides a better load-balance mechanism for the redundant links.

MSTP keeps a VLAN mapping table to associate VLANs with their spanning trees. Using MSTP, you can divide one switching network into multiple regions, each of which can have multiple spanning trees with each one independent of others. MSTP prunes the ring network into a loopfree tree to avoid the generation of loops and infinite circulations. It also provides multiple redundant paths for data forwarding to implement the load-balance mechanism of the VLAN data.

38.1.1 MSTP Concepts

There are 4 MST regions in Figure 38-1. Each region consists of four switches, all of which run MSTP. The following introduces the concept of MSTP with the help of this figure.

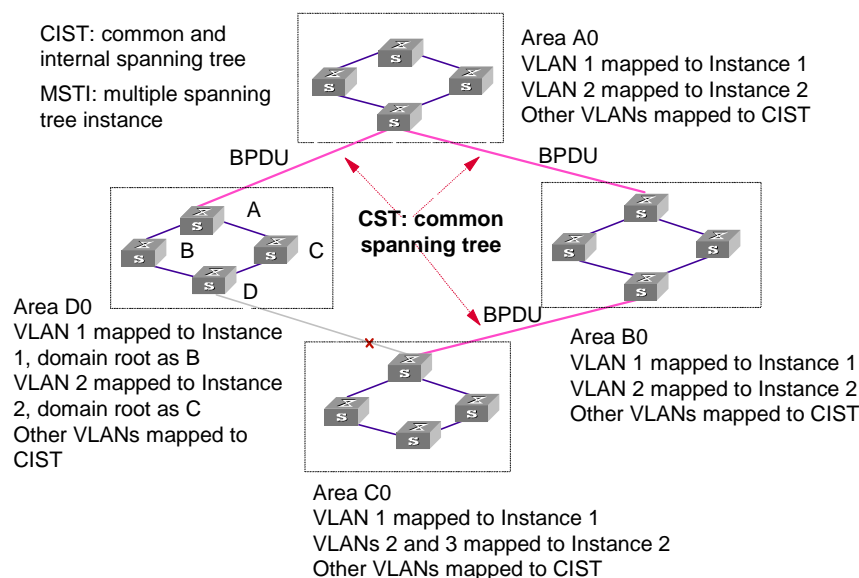


Figure 38-1 Basic MSTP concepts

I. MST region

Multiple Spanning Tree Regions: A multiple spanning tree region contains several switches and the network segments between them. These MSTP switches share the same region name, VLAN-spanning tree mapping configuration, and MSTP revision level configuration, and are connected directly. There can be several MST regions on a switching network. You can group several switches into a MST region, using MSTP configuration commands. For example, in Figure 38-1, the four switches in MST region A0 are configured with the same region name, the same VLAN mapping table (VLAN1 is mapped to instance 1, VLAN 2 is mapped to instance 2, other VLANs is mapped to instance CIST), and the same revision level (not indicated in Figure 38-1).

II. VLAN mapping table

The VLAN mapping table is an attribute of MST region. It is used for describing the mapping relationship of VLANs and spanning tree instances (STIs). For example, in the VLAN mapping table of MST region A0 in Figure 38-1, VLAN1 is mapped to instance 1, VLAN 2 is mapped to instance 2, other VLANs is mapped to CIST.

In the same region, the mapping relationship of VLANs and STIs must be consistent on all the switches in this region. Otherwise, VLAN and STI are not in the same region.

III. IST

Internal Spanning Tree (IST): a spanning tree in a MSTP region. The IST and the Common Spanning Tree (CST), together make up a Common and Internal Spanning Tree (CIST) for the entire switching network. The IST in a MST region is a fragment of the CIST. For example, every MST region in Figure 38-1 has an IST, which is a fragment of CIST.

IV. CST

Common Spanning Tree (CST): a LAN has only one CST. CST connects the spanning trees of all MST regions. Regard every MST region as a “switch”, and the CST is generated by the computing of “switches” through STP/RSTP. For example, the red line in Figure 38-1 indicates the CST.

V. CIST

Common and Internal Spanning Tree (CIST): A single spanning tree made up of ISTs and CST. It connects all switches in a switching network. CIST of Figure 38-1 is composed of ISTs in all MST regions and the CST.

VI. MSTI

Multiple Spanning Tree Instance (MSTI): multiple spanning trees can be generated with MSTP in an MST region and independent of one another. Such a spanning tree is called an MSTI. As shown in Figure 38-1, every MST region have many STIs. Each STI corresponds to a VLAN and is called a MSTI.

VII. Region root

The region root refers to the root of the IST and MSTI of the MST region. The spanning trees in an MST region have different topology and their region roots may also be different. For example, the region root of the STI 1 is the switch B and that of the STI 2 is the switch C, as shown in Figure 38-1.

VIII. Common Root Bridge

The Common Root Bridge refers to the root bridge of CIST. For example, the common root bridge is a certain switch in A0, as shown in Figure 38-1.

IX. Edge port

The edge port refers to the port located at the MST region edge, connecting different MST regions, MST region and STP region, or MST region and RSTP region. For MSTP calculation, the edge port shall take the same role on MSTI and CIST instance. For example, as shown in Figure 38-1, if a switch in region A0 connects to the first port on a switch in region D0, and the common root bridge of the whole switching network is in A0, then this first port is an edge port of region D0.

X. Port role

In the process of MSTP calculation, a port can serve as a designated port, root port, master port, alternate port, or backup port.

- The root port is the one through which the data are forwarded to the root.
- The designated port is the one through which the data are forwarded to the downstream network segment or switch.

- Master port is the port connecting the entire region to the Common Root Bridge and located on the shortest path between them.
- An alternate port is a backup of the master port, and also a backup port of a root port in the region. As a backup of the master port, an alternate port will become a new master port after a master port is blocked.
- If two ports of a switch are connected, there must be a loop. In this case, the switch blocks one of them. The blocked one is called a backup port.

A port can play different roles in different spanning tree instances.

The following figure illustrates the earlier-mentioned concepts for your better understanding. In this figure, the switch A, B, C, and D make up a MST region. Port 1 and 2 on switch A connects to the common root bridge; port 5 and 6 on switch C forms a loop; port 3 and 4 on switch D connects to other MST regions in the downstream direction.

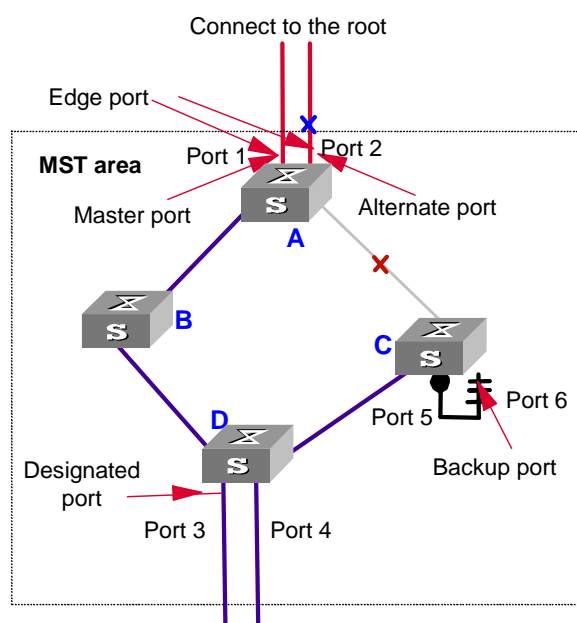


Figure 38-2 Port roles

XI. TC packet

Topology change (TC) means the structure of the MSTP spanning tree changes due to some bridge change or some port change on the network. In versatile routing platform (VRP) implementation, when a port state changes from discarding to forwarding, it means TC occurs.

The following section describes two kinds of STP packets:

1) MSTP BPDU packet

MSTP modules communicate with each other among bridges by MSTP BPDU packets. The following figure shows the MSTP BPDU packet format:

	Octet
Protocol Identifier	1–2
Protocol Version Identifier	3
BPDU Type	4
CIST Flags	5
CIST Root Identifier	6–13
CIST External Path Cost	14–17
CIST Regional Root Identifier	18–25
CIST Port Identifier	26–27
Message Age	28–29
Max Age	30–31
Hello Time	32–33
Forward Delay	34–35
Version 1 Length = 0	36
Version 3 Length	37–38
MST Configuration Identifier	39–89
CIST Internal Root Path Cost	90–93
CIST Bridge Identifier	94–101
CIST Remaining Hops	102
MSTI Configuration Messages (may be absent)	103–39 + <i>Version 3 Length</i>

Figure 38-3 BPDU packet format

	Octet
MSTI Flags	1
MSTI Regional Root Identifier	2–9
MSTI Internal Root Path Cost	10–13
MSTI Bridge Priority	14
MSTI Port Priority	15
MSTI Remaining Hops	16

Figure 38-4 MSTI information format of the last part in BPDU packets

Besides field root bridge priority, root path cost, local bridge priority and port priority, the field flags which takes one byte in an instance is also used for role selection. The following figure describes the meaning of its eight bits:

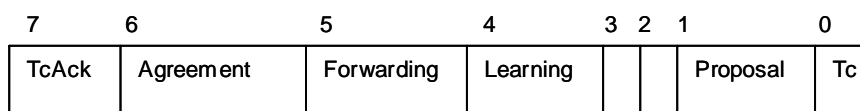


Figure 38-5 Meaning of 1-byte Flags in BPDU packets

The second and third bits together indicate MSTP port role.

2) TC packet

A TC packet is also an MSTP BPDU packet, but the lowest bit of its flags field is set to 1, which endows the TC packet with special meaning. So the TC packet has its special meaning. After receiving or detecting TC packets, a port will broadcast TC packets to tell the whole network the changed topology information at the fastest speed.

38.1.2 MSTP Principles

MSTP divides the entire Layer 2 network into several MST regions and calculates and generates CST for them. Multiple spanning trees are generated in a region and each of them is called an MSTI. The instance 0 is called IST, and others are called MSTI. Similar to RSTP, MSTP also use configuration messages to calculate and generate spanning trees, the difference is that it is the MSTP configuration information on the switches that is carried in the configuration messages.

I. CIST calculation

The CIST root is the highest-priority switch elected from the switches on the entire network through comparing their configuration BPDUs. MSTP calculates and generates IST in each MST region; at the same time it regards each MST region as a single "switch" and then calculates and generates the CST between the regions. The CST and ISTs together make up the CIST which connects all the switches in the whole switching network.

II. MSTI calculation

Inside an MST region, MSTP generates different MSTIs for different VLANs according to the association between VLAN and the spanning tree. The calculation process of MSTI is like that of RSTP.

The following introduces the calculation process of one MSTI.

The fundamental of STP is that the switches exchange a special kind of protocol packet (which is called configuration Bridge Protocol Data Units, or BPDU, in IEEE 802.1D) to decide the topology of the network. The configuration BPDU contains the information enough to ensure the switches to compute the spanning tree.

Figure 38-6 shows the Designated bridge and designated port.

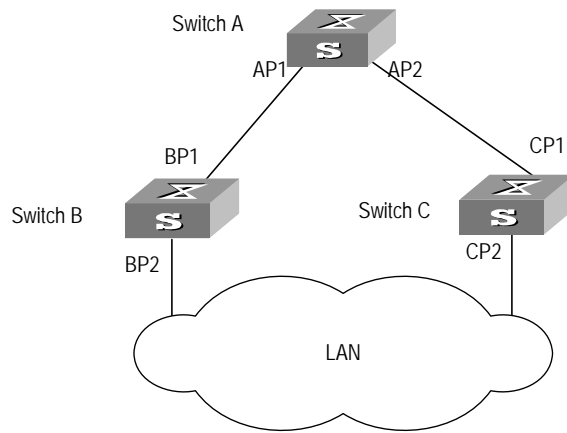


Figure 38-6 Designated bridge and designated port

For a switch, the designated bridge is a switch in charge of forwarding BPDU to the local switch via a port called the designated port accordingly. For a LAN, the designated bridge is a switch that is in charge of forwarding BPDU to the network segment via a port called the designated port accordingly. As illustrated in the Figure 38-6, Switch A forwards data to Switch B via the port AP1. To Switch B, the designated bridge is Switch A and the designated port is AP1. In the figure, Switch B and Switch C are connected to the LAN and Switch B forwards BPDU to LAN. So the designated bridge of LAN is Switch B and the designated port is BP2.

- The specific calculation process of STP algorithm.

The following example illustrates the calculation process of STP.

Figure 38-7 illustrates the practical network.

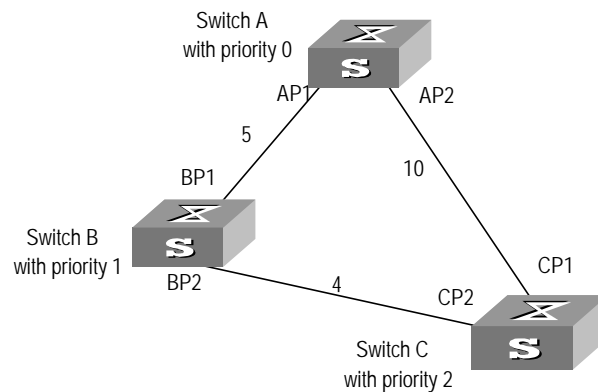


Figure 38-7 Ethernet switch networking

To facilitate the descriptions, only the first four parts of the configuration BPDU are described in the example. They are root ID (expressed as Ethernet switch priority), path cost to the root, designated bridge ID (expressed as Ethernet switch priority) and the designated port ID (expressed as the port number). As illustrated Figure 38-7, the

priorities of Switch A, B and C are 0, 1 and 2 and the path costs of their links are 5, 10 and 4 respectively.

1) Initial state

When initialized, each port of the switches generates the configuration BPDU taking itself as the root with a root path cost as 0, designated bridge IDs as their own switch IDs and the designated ports as their ports.

Switch A:

Configuration BPDU of AP1: {0, 0, 0, AP1}

Configuration BPDU of AP2: {0, 0, 0, AP2}

Switch B:

Configuration BPDU of BP1: {1, 0, 1, BP1}

Configuration BPDU of BP2: {1, 0, 1, BP2}

Switch C:

Configuration BPDU of CP2: {2, 0, 2, CP2}

Configuration BPDU of CP1: {2, 0, 2, CP1}

2) Select the optimum configuration BPDU

Every switch transmits its configuration BPDU to others. When a port receives a configuration BPDU with a lower priority than that of its own, the switch discards the message and keep the local BPDU unchanged. When the port receives a higher-priority configuration BPDU, the switch uses the content in the received configuration BPDU to change the content of the local BPDU of this port. Then the switch compare the configuration BPDU of this port to those of other ports on it to elect the optimum configuration BPDU.

The comparison rules are:

- The configuration BPDU with a smaller root ID has a higher priority.
- If the root IDs are the same, perform the comparison based on root path costs. The cost comparison is as follows: the path cost to the root recorded in the configuration BPDU plus the corresponding path cost of the local port is set as S, the configuration BPDU with a smaller S has a higher priority.
- If the costs of path to the root are also the same, compare in sequence the designated bridge ID, designated port ID and the ID of the port via which the configuration BPDU was received.

For the convenience of expression, this example supposes that the optimum configuration BPDU can be elected just by the comparison of root IDs.

3) Determine the root and designated ports, and update the configuration BPDU of designated ports.

The port receiving the optimum configuration BPDU is designated to be the root port, whose configuration BPDU remains unchanged. Switch calculates a designated port

BPDU for every port: substituting the root ID with the root ID in the configuration BPDU of the root port, the cost of path to root with the value made by the root path cost plus the path cost corresponding to the root port, the designated bridge ID with the local switch ID and the designated port ID with the local port ID.

Switch compares the calculated BPDU with the BPDU of corresponding port. If the BPDU of corresponding port is better, the port is blocked, and the BPDU of the port remains unchanged. The port will not forward data and only receive but not send BPDU. If the calculated BPDU is better, the port will be the designated port, and the port BPDU will be modified by the calculated BPDU and sent out regularly.

The comparison process of each switch is as follows.

Switch A:

AP1 receives the configuration BPDU from Switch B and finds out that the local configuration BPDU priority is higher than that of the received one, so it discards the received configuration BPDU. The configuration BPDU is processed on the AP2 in a similar way. Thus Switch A finds itself the root and designated bridge in the configuration BPDU of every port. It regards itself as the root, retains the configuration BPDU of each port and transmits configuration BPDU to others regularly thereafter. By now, the configuration BPDUs of the two ports are as follows:

Configuration BPDU of AP1: {0, 0, 0, AP1}.

Configuration BPDU of AP2: {0, 0, 0, AP2}.

Switch B:

BP1 receives the configuration BPDU from Switch A and finds that the received BPDU has a higher priority than the local one, so it updates its configuration BPDU.

BP2 receives the configuration BPDU from Switch C and finds that the local BPDU priority is higher than that of the received one, so it discards the received BPDU.

By now, the configuration BPDUs of each port are as follows: Configuration BPDU of BP1: {0, 0, 0, AP1}, Configuration BPDU of BP2: {1, 0, 1, BP2}.

Switch B compares the configuration BPDUs of the ports and selects the BP1 BPDU as the optimum one because the current configuration BPDU {0, 5, 0, AP1} of BP1 has a higher priority than the configuration BPDU {1, 0, 1, BP2} of BP2. Thus BP1 is elected as the root port and the configuration BPDUs of Switch B ports are updated as follows.

The configuration BPDU of the root port BP1 retains as {0, 5, 0, AP1}. BP2 updates root ID with that in the optimum configuration BPDU, the path cost to root with 5, sets the designated bridge as the local switch ID and the designated port ID as the local port ID. Thus, the configuration BPDU becomes {0, 5, 1, BP2}.

Then, all the designated ports of Switch B transmit the configuration BPDUs regularly.

Switch C:

CP2 receives from the BP2 of Switch B the configuration BPDU {1, 0, 1, BP2} that has not been updated and then the updating process is launched. The configuration BPDU is updated as {1, 0, 1, BP2}.

CP1 receives the configuration BPDU {0, 0, 0, AP2} from Switch A and Switch C launches the updating. The configuration BPDU is updated as {0, 0, 0, AP2}.

Now, the configuration BPDU of CP1 is {0, 10, 0, AP2}, which has a higher priority than that of CP2. By comparison, CP1 configuration BPDU is elected as the optimum one. The CP1 is thus specified as the root port without modifying its configuration BPDU. However, CP2 will be blocked and its BPDU also remains unchanged, but it will not receive the data (excluding the STP packets) forwarded from Switch B until spanning tree calculation is launched again by some new events. For example, the link from Switch B to Switch C is down or the port receives any better configuration BPDU.

CP2 will receive the updated configuration BPDU, {0, 5, 1, BP2}, from Switch B. Since this configuration BPDU is better than the old one, the old BPDU will be updated to {0, 5, 1, BP2}.

Meanwhile, CP1 receives the configuration BPDU from Switch A but its configuration BPDU is not updated and retain {0, 10, 0, AP2}.

By comparison, {0, 9, 1, BP2}, the configuration BPDU of CP2, is elected as the optimum one. Thus, CP2 is elected as the root port, whose BPDU will not change, while CP1 is blocked, its BPDU is retained, and will not receive the data forwarded from Switch A until spanning tree calculation is triggered again by some changes. For example, the link from Switch B to Switch C is down or the port receives any better configuration BPDU

Thus, the spanning tree is stabilized. The tree with the root bridge A is illustrated in the Figure 38-8.

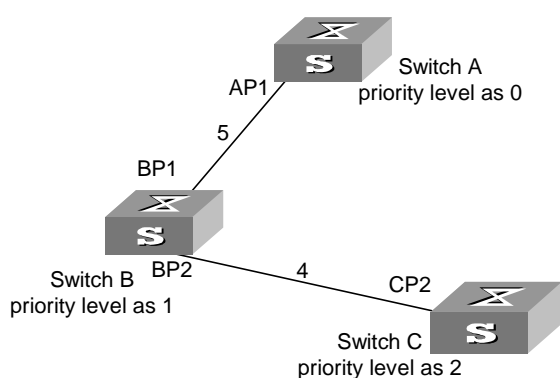


Figure 38-8 The final stabilized spanning tree

To facilitate the descriptions, the description of the example is simplified. For example, the root ID and the Designated bridge ID in actual calculation should comprise both switch priority and switch MAC address. Designated port ID should comprise port priority and port ID. In the updating process of a configuration BPDU, other

configuration BPDUs besides the first four items will make modifications according to certain rules. The basic calculation process is described below:

In addition, with identical priority, the path cost of an aggregation port is smaller than that of a non-aggregation port. Therefore, under identical root ID, path cost value and designated switch ID, the switch will generally select the aggregation port as the root port.

- Configuration BPDU forwarding mechanism in STP:

Upon the initiation of the network, all the switches regard themselves as the roots. The designated ports send the configuration BPDUs of local ports at a regular interval of HelloTime. If it is the root port that receives the configuration BPDU, the switch will enable a timer to time the configuration BPDU as well as increase MessageAge carried in the configuration BPDU by certain rules. If a path goes wrong, the root port on this path will not receive configuration BPDUs any more and the old configuration BPDUs will be discarded due to timeout. Hence, recalculation of the spanning tree will be initiated to generate a new path to replace the failed one and thus restore the network connectivity.

However, the new configuration BPDU as now recalculated will not be propagated throughout the network right away, so the old root ports and designated ports that have not detected the topology change will still forward the data through the old path. If the new root port and designated port begin to forward data immediately after they are elected, an occasional loop may still occur. In STP, a transitional state mechanism is thus adopted to ensure the new configuration BPDU has been propagated throughout the network before the root port and designated port begin to send data again. That is, the root port and designated port should undergo a transitional state for a period of Forward Delay before they enter the forwarding state.

And thus, the packets of a VLAN will be forwarded along the following path: in the MST region, the packets will be forwarded along the corresponding MSTI; among the regions, the packets will be forwarded along the CST.

38.1.3 MSTP Implementation on the Switch

MSTP is compatible with STP and RSTP. The MSTP switch can recognize both the STP and RSTP packets and calculate the spanning tree with them. Besides the basic MSTP functions, the Switch 8800 also provides some features that are easy to manage from users' point of view. These features include root bridge hold, secondary root bridge, ROOT protection, BPDU protection, loop protection, hot swapping of the interface boards, master/slave switchover, and so on. Note that the spanning tree needs to be calculated again when a master/slave switchover occurs.

38.2 Configuring MSTP

MSTP configuration includes:

- Configuring the MST Region for a Switch
- Specifying the Switch as a Primary or a Secondary Root bridge
- Configuring the MSTP Running Mode
- Configuring the Bridge Priority for a Switch
- Configuring the Max Hops in an MST Region
- Configuring the Switching Network Diameter
- Configuring the Time Parameters of a Switch
-

Setting the Timeout Factor of a Specific Bridge

- Configuring the Max Transmission Speed on a Port
- Configuring a Port as an Edge Port or Non-edge Port
- Configuring the Path Cost of a Port
- STP Path Cost Calculation Standards on STP port
- Configuring the Priority of a Port
- Configuring the Port (Not) to Connect with the Point-to-Point Link
- Configuring the mCheck Variable of a Port
- Configuring the Switch Protection Function
- Enabling/Disabling MSTP on the Device
- Enable/Disable Address Table Reset on Specified Port
- Enabling/Disabling MSTP on a Port

Only after MSTP is enabled on the device will other configurations take effect. Before enabling MSTP, you can configure the related parameters of the device and Ethernet ports, which will take effect upon enabling MSTP and stay effective even after resetting MSTP. The **check region-configuration** command can display the region parameters that have not yet taken effect. The **display current-configuration** command shows the parameters configured before MSTP is enabled. For those configured after MSTP is enabled, you can use the related **display** commands. For detailed information, refer to the “Display and Debug MSTP” section.

Note:

When GVRP and MSTP start on the switch simultaneously, GVRP packets will propagate along CIST which is a spanning tree instance. In this case, if you want to issue a certain VLAN through GVRP on the network, you should make sure that the VLAN is mapped to CIST when configuring the VLAN mapping table of MSTP. CIST is spanning tree instance 0.

38.2.1 Configuring the MST Region for a Switch

Which MST region a switch belongs to is determined with the configurations of the region name, VLAN mapping table, and MSTP revision level. You can perform the following configurations to put a switch into an MST region.

I. Entering MST region view

Perform the following configuration in system view.

Table 38-1 Enter MST region view

Operation	Command
Enter MST region view (from system view)	stp region-configuration
Restore the default settings of MST region	undo stp region-configuration

II. Configuring parameters for the MST region

Perform the following configuration in MST region view.

Table 38-2 Configure the MST region for a switch

Operation	Command
Configure the MST region name	region-name <i>name</i>
Restore the default MST region name	undo region-name
Configure VLAN mapping table	instance <i>instance-id</i> vlan <i>vlan-list</i>
Restore the default VLAN mapping table	undo instance <i>instance-id</i> [vlan <i>vlan-list</i>]
Configure the MSTP revision level of MST region	revision-level <i>level</i>
Restore the MSTP revision level of MST region	undo revision-level

An MST region can contain up to 49 spanning tree instances, among which the Instance 0 is IST and the Instances 1 through 48 are MSTIs. Upon the completion of the above configurations, the current switch is put into a specified MST region. Note that two switches belong to the same MST region only if they have been configured with the same MST region name, STI-VLAN mapping tables of an MST region, and the same MST region revision level.

Configuring the related parameters, especially the VLAN mapping table, of the MST region, will lead to the recalculation of spanning tree and network topology flapping. To bate such flapping, MSTP triggers to recalculate the spanning tree according to the configurations only if one of the following conditions is met:

- A user manually activates the configured parameters related to the MST region, using the **active region-configuration** command.
- A user enables MSTP using the **stp enable** command.

By default, the MST region name is the switch MAC address, all the VLANs in the MST region are mapped to the STI 0, and the MSTP region revision level is 0. You can restore the default settings of MST region, using the **undo stp region-configuration** command in system view.

III. Activating the MST region configuration, and exit the MST region view

Perform the following configuration in MST region view.

Table 38-3 Activate the MST region configuration and exit the MST region view

Operation	Command
Show the configuration information of the MST region under revision	check region-configuration
Manually activate the MST region configuration	active region-configuration
Exit MST region view	quit

38.2.2 Specifying the Switch as a Primary or a Secondary Root bridge

MSTP can determine the spanning tree root through calculation. You can also specify the current switch as the root, using the command provided by the switch.

You can use the following commands to specify the current switch as the primary or secondary root of the spanning tree.

Perform the following configuration in system view.

Table 38-4 Specify the switch as a primary or a secondary root bridge

Operation	Command
Specify the current switch as the primary root bridge of the specified spanning tree	stp [instance <i>instance-id</i>] root primary [bridge-diameter <i>bridgenum</i>] [hello-time <i>centi-seconds</i>]
Specify the current switch as the secondary root bridge of the specified spanning tree	stp [instance <i>instance-id</i>] root secondary [bridge-diameter <i>bridgenum</i>] [hello-time <i>centi-seconds</i>]
Specify current switch not to be the primary or secondary root	undo stp [instance <i>instance-id</i>] root

After a switch is configured as the primary root bridge or the secondary root bridge, users cannot modify the bridge priority of the switch.

You can configure the current switch as the primary or secondary root bridge of the STI (specified by the **instance *instance-id*** parameter). If the *instance-id* takes 0, the current switch is specified as the primary or secondary root bridge of the CIST.

The root types of a switch in different STIs are independent of one another. The switch can be a primary or secondary root of any STI. However, it cannot serve as both the primary and secondary roots of one STI.

If the primary root is down or powered off, the secondary root will take its place, unless you configure a new primary root. Of two or more configured secondary root bridges, MSTP selects the one with the smallest MAC address to take the place of the failed primary root.

When configuring the primary and secondary switches, you can also configure the network diameter and hello time of the specified switching network. For detailed information, refer to the configuration tasks “Configure switching network diameter” and “Configure the Hello Time of the switch”.

Note:

You can configure the current switch as the root of several STIs. However, it is not necessary to specify two or more roots for an STI. In other words, do not specify the root for an STI on two or more switches.

You can configure more than one secondary root for a spanning tree through specifying the secondary STI root on two or more switches.

Generally, you are recommended to designate one primary root and more than one secondary root for a spanning tree.

By default, a switch is neither the primary root nor the secondary root of the spanning tree.

38.2.3 Configuring the MSTP Running Mode

MSTP and RSTP are compatible and they can recognize the packets of each other. However, STP cannot recognize MSTP packets. To implement the compatibility, MSTP provides two operation modes, STP-compatible mode and MSTP mode. In STP-compatible mode, the switch sends STP packets via every port. In MSTP mode, the switch ports send MSTP or STP packets (when connected to the STP switch) and the switch provides multiple spanning tree function.

You can use the following command to configure MSTP running mode. MSTP can intercommunicate with STP. If there is a STP switch in the switching network, you may use the command to configure the current MSTP to run in STP-compatible mode. Otherwise, configure it to run in MSTP mode.

Perform the following configuration in system view.

Table 38-5 Configure the MSTP running mode

Operation	Command
Configure MSTP to run in STP-compatible mode	stp mode stp
Configure MSTP to run in MSTP mode	stp mode mstp

Restore the default MSTP running mode	undo stp mode
---------------------------------------	----------------------

Generally, if there is a STP switch on the switching network, the port connected to it will automatically transit from MSTP mode to STP-compatible mode. But the port cannot automatically transit back to MSTP mode after the STP switch is removed. In this case, you can execute the **stp mcheck** command to restore the MSTP mode.

By default, MSTP runs in MSTP mode.

38.2.4 Configuring the Bridge Priority for a Switch

Whether a switch can be elected as the spanning tree root depends on its Bridge priority. The switch configured with a smaller Bridge priority is more likely to become the root. An MSTP switch may have different priorities in different STIs.

You can use the following command to configure the Bridge priorities of the Designated bridge in different STIs.

Perform the following configuration in system view.

Table 38-6 Configure the Bridge priority for a switch

Operation	Command
Configure the Bridge priority of the Designated bridge	stp [instance <i>instance-id</i>] priority <i>priority</i>
Restore the default Bridge priority of the Designated bridge	undo stp [instance <i>instance-id</i>] priority

When configuring the switch priority with the **instance *instance-id*** parameter as 0, you are configuring the CIST priority of the switch.



Caution:

In the process of spanning tree root election, of two or more switches with the same Bridge priorities, the one has a smaller MAC address is elected as the root.

By default, the switch Bridge priority is 32768.

38.2.5 Configuring the Max Hops in an MST Region

The scale of MST region is limited by the max hops in an MST region, which is configured on the region root. As the BPDU travels from the spanning tree root, each time when it is forwarded by a switch, the max hops is reduced by 1. The switch

discards the configuration BPDU with 0 hops left. This makes it impossible for the switch beyond the max hops to take part in the spanning tree calculation, thereby limiting the scale of the MST region.

You can use the following command to configure the max hops in an MST region.

Perform the following configuration in system view.

Table 38-7 Configure the max hops in an MST region

Operation	Command
Configure the max hops in an MST region	stp max-hops <i>hop</i>
Restore the default max hops in an MST region	undo stp max-hops

The more the hops in an MST region, the larger the scale of the region. Only the max hops configured on the region root can limit the scale of MST region. Other switches in the MST region also apply the configurations on the region root, even if they have been configured with max hops.

By default, the max hop of an MST is 20.

38.2.6 Configuring the Switching Network Diameter

Any two hosts on the switching network are connected with a specific path carried by a series of switches. Among these paths, the one passing more switches than all others is the network diameter, expressed as the number of passed switches.

You can use the following command to configure the diameter of the switching network.

Perform the following configuration in system view.

Table 38-8 Configure the switching network diameter

Operation	Command
Configure the switching network diameter	stp bridge-diameter <i>bridgenum</i>
Restore the default switching network diameter	undo stp bridge-diameter

The network diameter is the parameter specifying the network scale. The larger the diameter is, the larger the scale of the network is.

When a user configures the network diameter on a switch, MSTP automatically calculates and sets the hello time, forward-delay time and maximum-age time of the switch to the desirable values.

Setting the network diameter takes effect on CIST only, but has no effect on MSTI.

By default, the network diameter is 7 and the three corresponding timers take the default values.

Note:

The **stp bridge-diameter** command configures the switching network diameter and determines the three MSTP time parameters (Hello Time, Forward Delay, and Max Age) accordingly.

38.2.7 Configuring the Time Parameters of a Switch

The switch has three time parameters, Forward Delay, Hello Time, and Max Age.

Forward Delay is the switch state transition mechanism. The spanning tree will be recalculated upon link faults and its structure will change accordingly. However, the configuration BPDU recalculated cannot be immediately propagated throughout the network. The temporary loops may occur if the new root port and designated port forward data right after being elected. Therefore the protocol adopts a state transition mechanism. It takes a Forward Delay interval for the root port and designated port to transit from the learning state to forwarding state. The Forward Delay guarantees a period of time during which the new configuration BPDU can be propagated throughout the network.

The switch sends Hello packet periodically at an interval specified by Hello Time to check if there is any link fault.

Max Age specifies when the configuration BPDU will expire. The switch will discard the expired configuration BPDU.

You can use the following command to configure the time parameters for the switch.

Perform the following configuration in system view.

Table 38-9 Configure the time parameters of a switch

Operation	Command
Configure Forward Delay on the switch	stp timer forward-delay <i>centiseconds</i>
Restore the default Forward Delay of the switch	undo stp timer forward-delay
Configure Hello Time on the switch	stp timer hello <i>centiseconds</i>
Restore the default Hello Time on the switch	undo stp timer hello
Configure Max Age on the switch	stp timer max-age <i>centiseconds</i>
Restore the default Max Age on the switch	undo stp timer max-age

Every switch on the switching network adopts the values of the time parameters configured on the root bridge of the CIST.

**Caution:**

The Forward Delay configured on a switch depends on the switching network diameter. Generally, the Forward Delay is supposed to be longer when the network diameter is longer. Note that too short a Forward Delay may redistribute some redundant routes temporarily, while too long a Forward Delay may prolong the network connection resuming. The default value is recommended.

A suitable Hello Time ensures the switch to detect the link fault on the network but occupy moderate network resources. The default value is recommended. If you set too long a Hello Time, when there is packet dropped over a link, the switch may consider it as a link fault and the network device will recalculate the spanning tree accordingly. However, for too short a Hello Time, the switch frequently sends configuration BPDU, which adds its burden and wastes the network resources.

Too short a Max Age may cause the network device frequently calculate the spanning tree and mistake the congestion as a link fault. However, if the Max Age is too long, the network device may not be able to discover the link fault and recalculate the spanning tree in time, which will weaken the auto-adaptation capacity of the network. The default value is recommended.

To avoid frequent network flapping, the values of Hello Time, Forward Delay and Maximum Age should guarantee the following formulas.

$$2 \times (\text{forward-delay} - 1 \text{ second}) \geq \text{maximum-age}$$
$$\text{maximum-age} \geq 2 \times (\text{hello time} + 1 \text{ second})$$

You are recommended to use the **stp root primary** command to specify the network diameter and Hello Time of the switching network, and then MSTP will automatically calculate and give the rather desirable values.

By default, Forward Delay is 15 seconds, Hello Time is 2 seconds, and Max Age is 20 seconds.

38.2.8 Setting the Timeout Factor of a Specific Bridge

A switch transmits hello packet regularly to the adjacent bridges to check if there is link failure. Generally, if the switch does not receive the STP packets from the upstream switch for 3 times of hello time, the switch will decide the upstream switch is dead and will recalculate the topology of the network. Then, in a steady network, the recalculation may be caused when the upstream is busy. In this case, user can redefine the timeout interval to a longer time to avoid this kind of meaningless recalculation.

You can use the following command to set the multiple value of hello time of a specified bridge.

Perform the following configurations in system view.

Table 38-10 Setting the timeout factor of a specific switch

Operation	Command
Set the timeout factor of a specified switch	stp timer-factor <i>number</i>
Restore the default timeout factor	undo stp timer-factor

It is recommended to set 5, 6 or 7 as the timeout factor in the steady network.

By default, the timeout factor of the switch is 3.

38.2.9 Configuring the Max Transmission Speed on a Port

The max transmission speed on a port specifies how many MSTP packets will be transmitted via the port every Hello Time.

The max transmission speed on a port is limited by the physical state of the port and the network structure. You can configure it according to the network conditions.

You can configure the max transmission speed on a port in the following ways.

I. Configuration in system view

Perform the following configuration in system view.

Table 38-11 Configure the max transmission speed on a port

Operation	Command
Configure the max transmission speed on a port	stp interface <i>interface-list</i> transmit-limit <i>packetnum</i>
Restore the default max transmission speed on a port	undo stp interface <i>interface-list</i> transmit-limit

II. Configuration in Ethernet port view

Perform the following configuration in Ethernet port view.

Table 38-12 Configure the max transmission speed on a port

Operation	Command
Configure the max transmission speed on a port	stp transmit-limit <i>packetnum</i>
Restore the default max transmission speed on a port	undo stp transmit-limit

You can configure the max transmission speed on a port with either of the earlier-mentioned measures. For more about the commands, refer to the *Command Manual*.

This parameter only takes a relative value without units. If it is set too large, too many packets will be transmitted during every Hello Time and too many network resources will be occupied. The default value is recommended.

By default, the max transmission speed on every Ethernet port of the switch is 3.

38.2.10 Configuring a Port as an Edge Port or Non-edge Port

An edge port refers to the port not directly connected to any switch or indirectly connected to a switch over the connected network.

You can configure a port as an edge port or non-edge port in the following ways.

I. Configuration in system view

Perform the following configuration in system view.

Table 38-13 Configure a port as an edge port or a non-edge port

Operation	Command
Configure a port as an edge port.	stp interface <i>interface-list</i> edged-port enable
Configure a port as a non-edge port.	stp interface <i>interface-list</i> edged-port disable
Restore the default setting of the port as a non-edge port.	undo stp interface <i>interface-list</i> edged-port

II. Configuration in Ethernet port view

Perform the following configuration in Ethernet port view.

Table 38-14 Configure a port as an edge port or a non-edge port

Operation	Command
Configure a port as an edge port.	stp edged-port enable
Configure a port as a non-edge port.	stp edged-port disable
Restore the default setting of the port as a non-edge port.	undo stp edged-port

You can configure a port as an edge port or a non-edge port with either of the earlier-mentioned measures.

After configured as an edge port, the port can fast transit from blocking state to forwarding state without any delay. You can only set the port connecting with the terminal as an edge port. In the case that BPDU protection has not been enabled on the switch, the configured edge port will turn into a non-edge port again when it receives BPDU from other ports. In the case that BPDU protection is enabled, the port will be disabled. The configuration of this parameter takes effect on all the STIs. In other words, if a port is configured as an edge port or non-edge port, it is configured the same on all the STIs.

It is better to configure the BPDU protection on the edged port, so as to prevent the switch from being attacked.

Before BPDU protection is enabled on the switch, the port runs as a non-edge port when it receives BPDU, even if the user has set it as an edge port.

If BPDU protection is enabled on the switch, the port is disabled. Only the network administrators can enable the port.

By default, all the Ethernet ports of the switch have been configured as non-edge ports.

Note:

It is better to configure the port directly connected with the terminal as an edge port, and enable the BPDU function on the port. That is, to realize fast state-transition and prevent the switch from being attacked.

38.2.11 Configuring the Path Cost of a Port

Path Cost is related to the speed of the link connected to the port. On the MSTP switch, a port can be configured with different path costs for different STIs. Thus the traffic from different VLANs can run over different physical links, thereby implementing the VLAN-based load-balancing.

You can configure the path cost of a port in the following ways.

I. Configuration in system view

Perform the following configuration in system view.

Table 38-15 Configure the path cost of a port

Operation	Command
Configure the path cost of a port.	stp interface <i>interface-list</i> [instance <i>instance-id</i>] cost <i>cost</i>
Restore the default path cost of a port.	undo stp interface <i>interface-list</i> [instance <i>instance-id</i>] cost

II. Configuration in Ethernet port view

Perform the following configuration in Ethernet port view.

Table 38-16 Configure the path cost of a port

Operation	Command
Configure the path cost of a port	stp [instance <i>instance-id</i>] cost <i>cost</i>
Restore the default path cost of a port.	undo stp [instance <i>instance-id</i>] cost

You can configure the path cost of a port with either of the earlier-mentioned measures. Upon the change of path cost of a port, MSTP will recalculate the port role and transit the state. When *instance-id* takes 0, it indicates to set the path cost on the CIST.

By default, MSTP is responsible for calculating the path cost of a port.

38.2.12 STP Path Cost Calculation Standards on STP port

The Switch 8800 uses its own legacy path calculation but both DOT1T and DOT1D-1998 path cost calculation can be used, as well. By default, the DOT1T is applied.

The port rate must be obtained first before calculating the path cost of a port as the path cost is associated with the port rate. The three standards use their own way to work out the port rate, based on which each standard calculates the path cost of the by certain algorithm.

I. DOT1T calculation standard

- 1) Calculating the rate
 - Aggregation port

The rate of either a primary or a secondary port in an aggregation port group is the sum of the port rates in the group. If a port is down, the rate is 0.

- Non-aggregation port

The actual rate counts.

2) Calculating the path cost

- Full-duplex and non-aggregation port at a rate less than 1 GE

$$\text{Path cost} = [200,000,000 / (\text{rate} \times 10)] - 1$$

- Other ports

$$\text{Path cost} = 200,000,000 / (\text{rate} \times 10)$$

II. DOT1D-1998 calculation standard

1) Calculating the rate

- Aggregation port

If the port is up, the actual rate counts. If the port is down, the rate is determined by that of the port which goes up first in the aggregation group. If all the ports in the aggregation group are down, the rate of the aggregation port is 0.

- Non-aggregation port

The actual rate counts.

2) Calculating the path cost

Table 38-17 details the correspondence between the rate range and the path cost values of the ports.

Table 38-17 Correspondence between the rate range and the path cost values

Rate range	Path cost value
[0, 10]	99 (for full-duplex port) 95 (for aggregation port) 100 (default)
(10, 100]	18 (for full-duplex port) 15 (for aggregation port) 19 (default)
(100,1000]	3 (for aggregation port) 4 (default)
(1000,10000]	2 (for aggregation port) 1 (default)
> 10000	1

III. The Switch 8800 legacy calculation standard

1) Calculating the rate

- Aggregation port

The rate of the primary port in an aggregation group is determined by the sum of the port rates in this group. No calculation is performed for secondary port.

- Non-aggregation port

The actual rate counts, but the rate is 0 if the port is down.

2) Calculating the path cost

Table 38-18 details the correspondence between the rate range and the value range of the path cost of the ports.

Table 38-18 Correspondence between the rate range and path cost range

Rate range	Path cost range
[0, 100]	2200 to (20 × rate)
(100,1000]	220 to the integer of [(0.2 × rate)]
(1000,10000]	22 to the integer of [(0.002 × rate)]
> 10000	1

You can specify the intended standard by using the following commands.

Perform the following configuration in system view.

Table 38-19 Specifying the standard to be followed in path cost calculation

Operation	Command
Specify the standard to be adopted when the switch calculates the default path cost for the connected link	stp pathcost-standard { dot1d-1998 dot1t legacy }
Restore the default standard to be used	undo stp pathcost-standard

By default, the switch calculates the default path cost of a port by the DOT1T standard.

38.2.13 Configuring the Priority of a Port

For spanning tree calculation, the port priority is an importance factor to determine if a port can be elected as the root port. With other things being equal, the port with the highest priority will be elected as the root port. On the MSTP switch, a port can have different priorities in different STIs and plays different roles respectively. Thus the traffic from different VLANs can run over different physical links, thereby implementing the VLAN-based load-balancing.

You can configure the port priority in the following ways.

I. Configuration in system view

Perform the following configuration in system view.

Table 38-20 Configure the port priority

Operation	Command
Configure the port priority.	stp interface <i>interface-list</i> instance <i>instance-id</i> port priority <i>priority</i>
Restore the default port priority.	undo stp interface <i>interface-list</i> instance <i>instance-id</i> port priority

II. Configuration in Ethernet port view

Perform the following configuration in Ethernet port view.

Table 38-21 Configure the port priority

Operation	Command
Configure the port priority.	stp [instance <i>instance-id</i>] port priority <i>priority</i>
Restore the default port priority.	undo stp [instance <i>instance-id</i>] port priority

You can configure the port priority with either of the earlier-mentioned measures. Upon the change of port priority, MSTP will recalculate the port role and transit the state. Generally, a smaller value represents a higher priority. If all the Ethernet ports of a switch are configured with the same priority value, the priorities of the ports will be differentiated by the index number. The change of Ethernet port priority will lead to spanning tree recalculation. You can configure the port priority according to actual networking requirements.

By default, the priority of all the Ethernet ports is 128.

38.2.14 Configuring the Port (Not) to Connect with the Point-to-Point Link

The point-to-point link directly connects two switches.

You can configure the port (not) to connect with the point-to-point link in the following ways.

I. Configuration in system view

Perform the following configuration in system view.

Table 38-22 Configure the port (not) to connect with the point-to-point link

Operation	Command
Configure the port to connect with the point-to-point link.	stp interface <i>interface-list</i> point-to-point force-true
Configure the port not to connect with the point-to-point link.	stp interface <i>interface-list</i> point-to-point force-false

Operation	Command
Configure MSTP to automatically detect if the port is directly connected with the point-to-point link.	stp interface <i>interface-list</i> point-to-point auto
Configure MSTP to automatically detect if the port is directly connected with the point-to-point link, as defaulted.	undo stp interface <i>interface-list</i> point-to-point

II. Configuration in Ethernet port view

Perform the following configuration in Ethernet port view.

Table 38-23 Configure the port (not) to connect with the point-to-point link

Operation	Command
Configure the port to connect with the point-to-point link.	stp point-to-point force-true
Configure the port not to connect with the point-to-point link.	stp point-to-point force-false
Configure MSTP to automatically detect if the port is directly connected with the point-to-point link.	stp point-to-point auto
Configure MSTP to automatically detect if the port is directly connected with the point-to-point link, as defaulted.	undo stp point-to-point

You can configure the port (not) to connect with the point-to-point link with either of the earlier-mentioned measures. For the ports connected with the point-to-point link, upon some port role conditions met, they can transit to forwarding state fast through transmitting synchronization packet, thereby reducing the unnecessary forwarding delay. If the parameter is configured as auto mode, MSTP will automatically detect if the current Ethernet port is connected with the point-to-point link.

Note:

For a link aggregation, only the master port can be configured to connect with the point-to-point link. If a port in auto-negotiation mode operates in full-duplex mode upon negotiation, it can be configured to connect with the point-to-point link.

This configuration takes effect on the CIST and all the MSTIs. The settings of a port whether to connect the point-to-point link will be applied to all the STIs to which the port belongs. Note that a temporary loop may be redistributed if you configure a port that is not physically connected with the point-to-point link as connected to such a link by force.

By default, the parameter is configured as **auto**.

38.2.15 Configuring the mCheck Variable of a Port

The port of an MSTP switch operates in either STP-compatible or MSTP mode.

Suppose a port of an MSTP switch on a switching network is connected to an STP switch, the port will automatically transit to operate in STP-compatible mode. However, the port stays in STP-compatible mode and cannot automatically transit back to MSTP mode when the STP switch is removed. In this case, you can perform mCheck operation to transit the port to MSTP mode by force.

You can use the following measure to perform mCheck operation on a port.

I. Configuration in system view

Perform the following configuration in system view.

Table 38-24 Configure the mCheck variable of a port

Operation	Command
Perform mCheck operation on a port.	stp interface <i>interface-list</i> mcheck

Note:

By default, MSTP runs in MSTP mode, which is compatible with RSTP and STP (This mode can recognize MSTP BPDU, STP config BPDU and RSTP config BPDU). However, the STP switch can only recognize config BPDU (STP BPDU) sent by the STP and RSTP bridges. After the switch running STP-compatible mode switches back to MSTP mode, it will not send MSTP BPDU if you do not execute the **stp mcheck** command. Therefore, the connected device still sends config BPDU (STP BPDU) to it, causing the same configuration exists in different regions and other problems. Remember to perform `stp interface mcheck` after modifying stp mode.

II. Configuration in Ethernet port view

Perform the following configuration in Ethernet port view.

Table 38-25 Configure the mCheck variable of a port

Operation	Command
Perform mCheck operation on a port.	stp mcheck

You can configure mCheck variable on a port with either of the earlier-mentioned measures. Note that the command can be used only if the switch runs MSTP. The command does not make any sense when the switch runs in STP-compatible mode.

38.2.16 Configuring the Switch Protection Function

An MSTP switch provides BPDU protection, Root protection functions, loop protection and TC-protection.

I. BPDU protection

For an access device, the access port is generally directly connected to the user terminal (for example, PC) or a file server, and the access port is set to an edge port to implement fast transition. When such a port receives BPDU packet, the system will automatically set it as a non-edge port and recalculate the spanning tree, which causes the network topology flapping. In normal cases, these ports will not receive STP BPDU. If someone forges BPDU to attack the switch, the network will flap. BPDU protection function is used against such network attacks.

II. Root protection

The primary and secondary root bridges of the spanning tree, especially those of ICST, shall be located in the same region. It is because the primary and secondary roots of CIST are generally placed in the core region with a high bandwidth in network design. In case of configuration error or malicious attack, the legal primary root may receive the BPDU with a higher priority and then lose its place, which causes network topology change errors. Due to the illegal change, the traffic supposed to travel over the high-speed link may be pulled to the low-speed link and congestion will occur on the network. Root protection function is used against such problems.

III. Loop protection

The root port and other blocked ports maintain their states according to the BPDUs sent by uplink switch. Once the link is blocked or has trouble, then the ports cannot receive BPDUs and the switch will select root port again. In this case, the former root port will turn into specified port and the former blocked ports will enter forwarding state, as a result, a link loop will be generated.

After the loop protection is enabled, for the root port, its role will not change, but its state will change. For the blocked port, its role will change, but its state will maintain in discarding. The blocked port does not forward packets, thus avoiding link loop.

Note:

For the loop protection-enabled port, if the port participates in STP calculation, all the instances of the port will be always set to be in discarding state regardless of the port role.

IV. TC-protection

As a general rule, the switch deletes the corresponding entries in the MAC address table and ARP table upon receiving TC-BPDU packets. Under malicious attacks of TC-BPDU packets, the switch shall receive a great number of TC-BPDU packets in a very short period. Too frequent delete operations shall consume huge switch resources and bring great risk to network stability.

When the protection from TC-BPDU packet attack is enabled, the switch just perform one delete operation in a specified period (generally, 15 seconds) after receiving TC-BPDU packets, as well as monitoring whether it receives TC-BPDU packets during this period. Even if it detects a TC-BPDU packet is received in a period shorter than the specified interval, the switch shall not run the delete operation till the specified interval is reached. This can avoid frequent delete operations on the MAC address table and ARP table.

You can use the following command to configure the protection functions of the switch. Perform the following configuration in corresponding configuration modes.

Table 38-26 Configure the switch protection function

Operation	Command
Configure BPDU protection of the switch (from system view)	stp bpdu-protection
Restore the disabled BPDU protection state as defaulted (from system view)	undo stp bpdu-protection
Configure Root protection of the switch (from system view)	stp interface <i>interface-list</i> root-protection
Restore the disabled Root protection state as defaulted (from system view)	undo stp interface <i>interface-list</i> root-protection
Configure Root protection of the switch (from Ethernet port view)	stp root-protection
Restore the disabled Root protection state as defaulted (from Ethernet port view)	undo stp root-protection
Configure loop protection function of the switch (from Ethernet port view)	stp loop-protection
Restore the disabled loop protection state, as defaulted (from Ethernet port view)	stp loop-protection

Operation	Command
Configure TC protection of the switch (from system view)	stp tc-protection enable
Disable TC protection (from system view)	stp tc-protection disable

By default, only the protection from TC-BPDU packet attack is enabled on the switch. BPDU protection, Root protection and loop protection are disabled.

After configured with BPDU protection, the switch will disable the edge port through MSTP which receives a BPDU, and notify the network manager at same time. These ports can be resumed by the network manager only.

The port configured with Root protection only plays a role of designated port on every instance. Whenever such port receives a higher-priority BPDU, that is, it is about to turn into non-designated port, it will be set to listening state and not forward packets any more (as if the link to the port is disconnected). If the port has not received any higher-priority BPDU for a certain period of time thereafter, it will resume the normal state.

For one port, only one configuration can be effective among loop protection, Root protection and Edge port configuration at the same moment.

Note:

The port configured with loop protection can only turn into discarding state on every instance. That such a port receives no configuration message for a long time indicates that it is about to change its state and role. Only the port role changes but the port discarding state remains unchanged, and no packets are forwarded. In this way, if the peer end cannot send BPDU packets due to error operation, and the port enters forwarding state directly for not receiving configuration message for a long time, no loop will be generated by enabling the loop protection.

By default, the switch does not enable BPDU protection or Root protection.

38.2.17 Enabling/Disabling MSTP on the Device

You can use the following command to enable MSTP on the device.

Perform the following configuration in system view.

Table 38-27 Enable/Disable MSTP on a device

Operation	Command
Enable MSTP on a device.	stp enable
Disable MSTP on a device.	stp disable
Restore the disable state of MSTP, as defaulted.	undo stp

Only if MSTP has been enabled on the device will other MSTP configurations take effect. If MSTP is disabled on the device, MSTP cannot be enabled on a port.

By default, MSTP is disabled.

38.2.18 Enable/Disable Address Table Reset on Specified Port

When a TC/TCN packet is received on a port, the system performs whole bridge traverse decision on the reset-arp enable/disable status. If reset-arp is enabled on the port, and the STP port is in active state, the system reset the MAC and dynamic ARP address tables on the port. In the case of TC/TCN entries of instance 0, the system removes the ARP entries of all instances. In the case of TC/TCN entries of other instances, the system removes the instance's ARP entry.

By default, this function is disabled.

Perform the following configuration in Ethernet port view.

Table 38-28 Enable/disable the reset of MAC and dynamic ARP address tables on a port

Operation	Command
Enable/Disable the reset of MAC and dynamic ARP address tables on a port of the device	stp reset-arp { enable disable }

By default, this function is disabled.

38.2.19 Enabling/Disabling ARP Address Update

ARP update is based on the following hypothesis: There are lots of bidirectional multicast and broadcast packets in the actual network. After the network topology changes, the MAC and ARP addresses may become invalid if the system does not delete ARP and MAC entries. However, if the peer sends a multicast or broadcast packet, the local port will learn ARP and MAC addresses. Then the system can find the corresponding ARP entries and update these entries according to the new port correspondence. This planning is also called: ARP address updates with MAC address.

The **stp update-arp** command can be executed in the system view for the system to determine whether to adopt ARP address update flexibly. If ARP address update is

disabled, upon receiving TC/TCN packets, the port broadcasts TC packets to delete the MAC address entries of the port in the STP active state on the bridge.

Perform the following configuration in system view.

Table 38-29 Enable/disable ARP address update

Operation	Command
Enable/disable ARP address update	stp update-arp { enable disable }

By default, ARP address update is enabled.

Note:

In general, the **STP update-arp disable** command works together with the **stp reset-arp enable** command in the port view. That is, the system removes MAC and ARP entries of the port after receiving TC/TCN packets.

38.2.20 Enabling/Disabling MSTP on a Port

You can use the following command to enable/disable MSTP on a port. You may disable MSTP on some Ethernet ports of a switch to spare them from spanning tree calculation. This is a measure to flexibly control MSTP operation and save the CPU resources of the switch.

MSTP can be enabled/disabled on a port through the following ways.

I. Configuration in system view

Perform the following configuration in system view.

Table 38-30 Enable/Disable MSTP on a port

Operation	Command
Enable MSTP on a port.	stp interface <i>interface-list</i> enable
Disable MSTP on a port.	stp interface <i>interface-list</i> disable

II. Configuration in Ethernet port view

Perform the following configuration in Ethernet port view.

Table 38-31 Enable/Disable MSTP on a port

Operation	Command
Enable MSTP on a port.	stp enable
Disable MSTP on a port.	stp disable

You can enable/disable MSTP on a port with either of the earlier-mentioned measures. Note that redundant route may be generated after MSTP is disabled.

By default, MSTP is enabled on all the ports after it is enabled on the device.

38.3 Displaying and Debugging MSTP

After the above configuration, execute the **display** command in any view to display the running of the MSTP configuration, and to verify the effect of the configuration. Execute the **reset stp [interface *interface-list*]** command in user view to clear the statistics of MSTP module. Execute the **debugging** command in user view to debug the MSTP module

Table 38-32 Display and debug MSTP

Operation	Command
Display the MSTP information about the current switch.	display stp
Display the configuration information about the current port and the switch.	display stp instance <i>instance-id</i> [interface <i>interface-list</i>] [brief]
Display the current configurations of the specified service board.	display stp slot <i>number</i> [brief]
Display the configuration information about the region.	display stp region-configuration
Display TC statistics	display stp tc [instance <i>instanceid</i>] { all detected received sent }
Clear the MSTP statistics information.	reset stp [interface <i>interface-list</i>]
Enable event debugging of MSTP.	debugging stp event
Enable packet debugging of MSTP.	debugging stp packet
Enable/Disable MSTP (packet receiving/transmitting, event, error) debugging on the port.	[undo] debugging stp [interface <i>interface-list</i>] { packet event }
Enable/Disable the global MSTP debugging.	[undo] debugging stp { global-event global-error all }
Enable/Disable specified STI debugging	[undo] debugging stp instance <i>instance-id</i>

Operation	Command
Enable STP global error or event debugging	<code>debugging stp { global-error global-event }</code>
Disable STP global error or event debugging	<code>undo debugging stp { global-error global-event }</code>

38.4 Typical MSTP Configuration Example

I. Network requirements

MSTP provides different forwarding paths for packets of different VLANs. The configurations are as follows: all the switches in the network belong to the same MST domain, packets of VLAN 10 travels along instance 1, packets of VLAN 30 travels along instance 3, packets of VLAN 40 travels along instance 4, and that of VLAN 20 travels along instance 0.

In the following network diagram, Switch A and Switch B are devices of the convergence layer, Switch C and Switch D are devices of the access layer. VLAN 10 and 30 function at the distribution and access layers, and VLAN 40 functions at the access layer only. So the root of instance 1 can be configured as Switch A, root of instance 3 can be Switch B, and root of instance 4 can be Switch C.

II. Network diagram

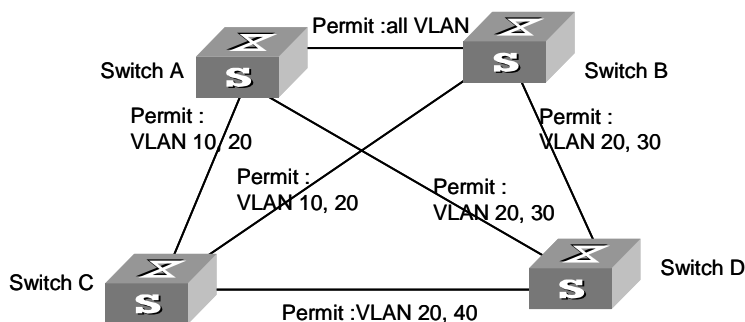


Figure 38-9 Network diagram for MSTP configuration

Note:

The explanations on the above figure which goes like “permit: XXXX” means that packets of these VLANs are permitted to pass.

III. Configuration procedure

1) Configurations on Switch A

MST region

```
[SW8800] stp region-configuration
[SW8800-mst-region] region-name example
[SW8800-mst-region] instance 1 vlan 10
[SW8800-mst-region] instance 3 vlan 30
[SW8800-mst-region] instance 4 vlan 40
[SW8800-mst-region] revision-level 0
```

Manually activate MST region configuration.

```
[SW8800-mst-region] active region-configuration
```

Specify Switch A as the root of instance 1

```
[SW8800] stp instance 1 root primary
```

2) Configurations on Switch B

MST region.

```
[SW8800] stp region-configuration
[SW8800-mst-region] region-name example
[SW8800-mst-region] instance 1 vlan 10
[SW8800-mst-region] instance 3 vlan 30
[SW8800-mst-region] instance 4 vlan 40
[SW8800-mst-region] revision-level 0
```

Manually activate MST region configuration.

```
[SW8800-mst-region] active region-configuration
```

Specify Switch B as the root of instance 3

```
[SW8800] stp instance 3 root primary
```

3) Configurations on Switch C

MST region.

```
[SW8800] stp region-configuration
[SW8800-mst-region] region-name example
[SW8800-mst-region] instance 1 vlan 10
[SW8800-mst-region] instance 3 vlan 30
[SW8800-mst-region] instance 4 vlan 40
[SW8800-mst-region] revision-level 0
```

Manually activate MST region configuration.

```
[SW8800-mst-region] active region-configuration
```

Specify Switch C as the root of instance 4.

```
[SW8800] stp instance 4 root primary
```

4) Configurations on Switch D

MST region

```
[SW8800] stp region-configuration
[SW8800-mst-region] region-name example
[SW8800-mst-region] instance 1 vlan 10
[SW8800-mst-region] instance 3 vlan 30
[SW8800-mst-region] instance 4 vlan 40
[SW8800-mst-region] revision-level 0
```

Manually activate MST region configuration.

```
[SW8800-mst-region] active region-configuration
```

Chapter 39 802.1x Configuration

39.1 802.1x Overview

39.1.1 802.1x Standard Overview

IEEE 802.1x (hereinafter simplified as 802.1x) is a port-based network access control protocol that is used as the standard for LAN user access authentication.

In the LANs complying with the IEEE 802 standards, the user can access the devices and share the resources in the LAN through connecting the LAN access control device like the LAN Switch. However, in telecom access, commercial LAN (a typical example is the LAN in the office building) and mobile office etc., the LAN providers generally hope to control the user's access. In these cases, the requirement on the above-mentioned "Port Based Network Access Control" originates.

As the name implies, "Port Based Network Access Control" means to authenticate and control all the accessed devices on the port of LAN access control device. If the user's device connected to the port can pass the authentication, the user can access the resources in the LAN. Otherwise, the user cannot access the resources in the LAN. It equals that the user is physically disconnected.

802.1x defines port based network access control protocol and only defines the point-to-point connection between the access device and the access port. The port can be either physical or logical. The typical application environment is as follows: Each physical port of the LAN Switch only connects to one user workstation (based on the physical port) and the wireless LAN access environment defined by the IEEE 802.11 standard (based on the logical port), etc.

39.1.2 802.1x System Architecture

The system using the 802.1x is the typical C/S (Client/Server) system architecture. It contains three entities, which are illustrated in the following figure: Supplicant System, Authenticator System and Authentication Sever System.

The LAN access control device needs to provide the Authenticator System of 802.1x. The devices at the user side such as the computers need to be installed with the 802.1x client Supplicant software, for example, the 802.1x client provided by Microsoft Windows XP. The 802.1x Authentication Sever system normally stays in the carrier's AAA center.

Authenticator and Authentication Sever exchange information through EAP (Extensible Authentication Protocol) frames. The Supplicant and the Authenticator exchange information through the EAPoL (Extensible Authentication Protocol over LANs) frame defined by IEEE 802.1x. Authentication data are encapsulated in the EAP frame, which

is to be encapsulated in the packets of other AAA upper layer protocols (e.g. RADIUS) so as to go through the complicated network to reach the Authentication Server. Such procedure is called EAP Relay.

There are two types of ports for the Authenticator. One is the Uncontrolled Port, and the other is the Controlled Port. The Uncontrolled Port is always in bi-directional connection state. The user can access and share the network resources any time through the ports. The Controlled Port will be in connecting state only after the user passes the authentication. Then the user is allowed to access the network resources.

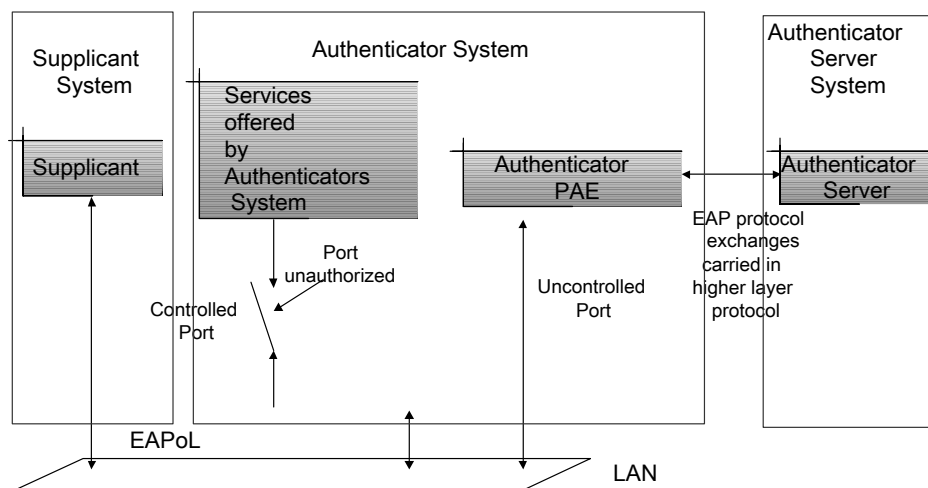


Figure 39-1 802.1x system architecture

39.1.3 802.1x Authentication Process

802.1x configures EAP frame to carry the authentication information. The Standard defines the following types of EAP frames:

- EAP-Packet: Authentication information frame, used to carry the authentication information.
- EAPoL-Start: Authentication originating frame, actively originated by the Supplicant.
- EAPoL-Logoff: Logoff request frame, actively terminating the authenticated state.
- EAPoL-Key: Key information frame, supporting to encrypt the EAP packets.
- EAPoL-Encapsulated-ASF-Alert: Supports the Alerting message of Alert Standard Forum (ASF).

The EAPoL-Start, EAPoL-Logoff and EAPoL-Key only exist between the Supplicant and the Authenticator. The EAP-Packet information is re-encapsulated by the Authenticator System and then transmitted to the Authentication Server System. The EAPoL-Encapsulated-ASF-Alert is related to the network management information and terminated by the Authenticator.

802.1x provides an implementation solution of user ID authentication. However, 802.1x itself is not enough to implement the scheme. The administrator of the access device should configure the AAA scheme by selecting RADIUS or local authentication so as to assist 802.1x to implement the user ID authentication. For detailed description of AAA, refer to the corresponding AAA configuration.

39.1.4 Implementing 802.1x on Ethernet Switches

The Switch 8800 not only supports the port access authentication method regulated by 802.1x, but also extends and optimizes it in the following way:

- Support to connect several End Stations in the downstream via a physical port.
- The access control (or the user authentication method) can be based on port or MAC address.

In this way, the system becomes much securer and easier to manage.

39.2 802.1x Configuration

The configuration tasks of 802.1x itself can be fulfilled in system view of the Ethernet switch. After the global 802.1x is enabled, the user can configure the 802.1x state of the port. The configured items will take effect after the global 802.1x is enabled.

Note:

When 802.1x is enabled on a port, the max number of MAC address learning which is configured by the command **mac-address max-mac-count** cannot be configured on the port, and vice versa.

The following sections describe 802.1x configuration tasks.

- Enabling/Disabling 802.1x
- Setting the Port Access Control Mode
- Setting Port Access Control Method
-

Checking the Users that Log on the Switch via Proxy

- Setting Supplicant Number on a Port
- Setting the Authentication in DHCP Environment
- Configuring Authentication Method for 802.1x User
- Enabling/Disabling Guest VLAN
- Setting the Maximum times of authentication request message retransmission
-

Configuring 802.1x Timers

- Enabling/Disabling quiet-period Timer

Among the above tasks, the first one is compulsory, otherwise 802.1x will not take any effect. The other tasks are optional. You can perform the configurations at requirements.

39.2.1 Enabling/Disabling 802.1x

The following command can be used to enable/disable the 802.1x on the specified port or globally. When it is used in system view, if the parameter *interface-list* is not specified, 802.1x will be globally enabled. If the parameter *interface-list* is specified, 802.1x will be enabled on the specified port. When this command is used in Ethernet port view, the parameter *interface-list* cannot be input and 802.1x can only be enabled on the current port.

Perform the following configuration in system view or Ethernet port view.

Table 39-1 Enable/Disable 802.1x

Operation	Command
Enable the 802.1x	dot1x [interface <i>interface-list</i>]
Disable the 802.1x	undo dot1x [interface <i>interface-list</i>]

By default, 802.1x authentication has not been enabled globally and on any port.

You cannot configure 802.1x on a port before you enable it globally. And you must disable 802.1x on each port before you disable 802,1x globally.

39.2.2 Setting the Port Access Control Mode

The following commands can be used for setting 802.1x access control mode on the specified port. When no port is specified, the access control mode of all ports is configured.

Perform the following configuration in system view or Ethernet port view.

Table 39-2 Set the port access control mode

Operation	Command
Set the port access control mode	dot1x port-control { authorized- force unauthorized-force auto } [interface <i>interface-list</i>]
Restore the default access control mode of the port	undo dot1x port-control [interface <i>interface-list</i>]

auto (automatic identification mode, which is also called protocol control mode). That is, the initial state of the port is unauthorized. It only permits EAPoL packets receiving/transmitting and does not permit the user to access the network resources. If the authentication flow is passed, the port will be switched to the authorized state and permit the user to access the network resources.

The **authorized-force** keyword specifies the port to operate in authorized-force mode. Ports in this mode are always authorized. Users can access a network through this kind of port without being authorized.

The **unauthorized-force** keyword specifies the port to operate in unauthorized-force mode. Ports in this mode are always unauthorized. They do not respond to authorization requests. Users cannot access a network through this kind of port.

By default, the mode of 802.1x performing access control on the port is **auto** (automatic identification mode).

39.2.3 Setting Port Access Control Method

The following commands are used for setting 802.1x access control method on the specified port. When no port is specified in system view, the access control method of all ports is configured.

Perform the following configuration in system view or Ethernet port view.

Table 39-3 Set port access control method

Operation	Command
Set port access control method	dot1x port-method { macbased portbased } [interface <i>interface-list</i>]
Restore the default port access control method	undo dot1x port-method [interface <i>interface-list</i>]

The **macbased** keyword specifies to authenticate each user accessing through the port. And disconnection of a user does not affect other users. Whereas if you specify the **portbased** keyword, users can access a network without being authenticated if a user passes the authentication previously. But these users are denied when the one who passes the authentication first goes offline.

By default, 802.1x authentication method on the port is **macbased**. That is, authentication is performed based on MAC addresses.

39.2.4 Checking the Users that Log on the Switch via Proxy

The following commands are used for checking the users that log on the switch via proxy.

Perform the following configuration in system view or Ethernet port view.

Table 39-4 Check the users that log on the switch via proxy

Operation	Command
Enable the check for access users via proxy	dot1x supp-proxy-check { logoff trap } [interface <i>interface-list</i>]
Cancel the check for access users via proxy	undo dot1x supp-proxy-check { logoff trap } [interface <i>interface-list</i>]

These commands take effect on the ports specified by the *interface-list* parameter when executed in system view. The parameter *interface-list* cannot be input when the command is executed in Ethernet Port view and it has effect only on the current interface. After globally enabling proxy user detection and control in system view, only if you enable this feature on a specific port can this configuration take effects on the port.

39.2.5 Setting Supplicant Number on a Port

The following commands are used for setting number of users allowed by 802.1x on specified port. When no port is specified, all the ports accept the same number of supplicants.

Perform the following configuration in system view or Ethernet port view.

Table 39-5 Setting maximum number of users via specified port

Operation	Command
Set maximum number of users via specified port	dot1x max-user <i>user-number</i> [interface <i>interface-list</i>]
Restore the maximum number of users on the port to the default value	undo dot1x max-user [interface <i>interface-list</i>]

By default, 802.1x allows up to 1024 supplicants on each port for the Switch 8800, and a Switch 8800 can accommodate a total of 2048 supplicants.

39.2.6 Setting the Authentication in DHCP Environment

If in DHCP environment the users configure static IP addresses, you can set 802.1x to disable the switch to trigger the user ID authentication over them with the following command.

Perform the following configuration in system view.

Table 39-6 Set the Authentication in DHCP Environment

Operation	Command
Disable the switch to trigger the user ID authentication over the users who configure static IP addresses in DHCP environment	dot1x dhcp-launch
Enable the switch to trigger the authentication over them	undo dot1x dhcp-launch

By default, the switch can trigger the user ID authentication over the users who configure static IP addresses in DHCP environment.

39.2.7 Configuring Authentication Method for 802.1x User

The following commands can be used to configure the authentication method for 802.1x user. Three kinds of methods are available: PAP authentication (RADIUS server must support PAP authentication), CHAP authentication (RADIUS server must support CHAP authentication), EAP relay authentication (switch send authentication information to RADIUS server in the form of EAP packets directly and RADIUS server must support EAP authentication).

Perform the following configuration in system view.

Table 39-7 Configure authentication method for 802.1x user

Operation	Command
Configure authentication method for 802.1x user	dot1x authentication-method { chap pap eap md5-challenge }
Restore the default authentication method for 802.1x user	undo dot1x authentication-method

By default, CHAP authentication is used for 802.1x user authentication.

39.2.8 Enabling/Disabling Guest VLAN

If Guest VLAN is enabled, a switch broadcasts active authentication packets to all 802.1x-enabled ports. The ports not sending response packets are added to Guest VLAN when the maximum number of re-authentications is reached. Users in a Guest VLAN can utilize resources in the Guest VLAN without undergoing the 802.1x authentication, but they can utilize the resources outside the Guest VLAN only when they have passed the 802.1x authentication. In this way, unauthenticated users can still perform operations such as accessing some resources with the 802.1x client not installed, and upgrading 802.1x client.

Perform the following configuration in system view or Ethernet interface view.

Table 39-8 Enable/Disable Guest VLAN

Operation	Command
Enable Guest VLAN	dot1x guest-vlan <i>vlan-id</i> [interface <i>interface-list</i>]
Disable Guest VLAN	undo dot1x guest-vlan <i>vlan-id</i> [interface <i>interface-list</i>]

Note that:

- Guest VLAN is only supported when the switch performs port-based authentication.
- A switch can have only one Guest VLAN.
- Users who are not authenticated, fail to be authenticated, or are offline are all members of the Guest VLAN.
- Guest VLANs can only be configured on access ports.
- You must use an existing VLAN ID, and the corresponding VLAN cannot be a super VLAN.
- You must perform corresponding configuration manually to isolate the Guest VLAN from other VLAN interfaces,.

39.2.9 Setting the Maximum times of authentication request message retransmission

The following commands are used for setting the maximum retransmission times of the authentication request message that the switch sends to the supplicant.

Perform the following configuration in system view.

Table 39-9 Set the maximum times of the authentication request message retransmission

Operation	Command
Set the maximum times of the authentication request message retransmission	dot1x retry <i>max-retry-value</i>
Restore the default maximum retransmission times	undo dot1x retry

By default, the *max-retry-value* is 2. That is, the switch can retransmit the authentication request message to a supplicant for 2 times at most.

39.2.10 Configuring 802.1x Timers

The following commands are used for configuring the 802.1x timers.

Perform the following configuration in system view.

Table 39-10 Configure 802.1x timers

Operation	Command
Configure timers	dot1x timer { handshake-period handshake-period-value quiet-period quiet-period-value tx-period tx-period-value supp-timeout supp-timeout-value server-timeout server-timeout-value }
Restore default settings of the timers	undo dot1x timer { handshake-period quiet-period tx-period supp-timeout server-timeout }

handshake-period: This timer begins after the user has passed the authentication. After setting handshake-period, system will send the handshake packet by the period. Suppose the dot1x retry time is configured as N, the system will consider the user having logged off and set the user as logoff state if system doesn't receive the response from user for consecutive N times.

handshake-period-value: Handshake period. The value ranges from 1 to 1024 in units of second and defaults to 30.

quiet-period: Specifies the quiet timer. If an 802.1x user has not passed the authentication, the Authenticator will keep quiet for a while (which is specified by quiet-period timer) before launching the authentication again. During the quiet period, the Authenticator does not do anything related to 802.1x authentication.

quiet-period-value: Specifies how long the quiet period is. The value ranges from 10 to 120 in units of second and defaults to 60.

server-timeout: Specifies the timeout timer of an Authentication Server. If an Authentication Server has not responded before the specified period expires, the Authenticator will resend the authentication request.

server-timeout-value: Specifies how long the duration of a timeout timer of an Authentication Server is. The value ranges from 100 to 300 in units of second and defaults to 100 seconds.

supp-timeout: Specifies the authentication timeout timer of a Supplicant. After the Authenticator sends Request/Challenge request packet which requests the MD5 encrypted text, the supp-timeout timer of the Authenticator begins to run. If the Supplicant does not respond back successfully within the time range set by this timer, the Authenticator will resend the above packet.

supp-timeout-value: Specifies how long the duration of an authentication timeout timer of a Supplicant is. The value ranges from 10 to 120 in units of second and defaults to 30.

tx-period: Specifies the transmission timeout timer. After the Authenticator sends the Request/Identity request packet which requests the user name or user name and password together, the tx-period timer of the Authenticator begins to run. If the Supplicant does not respond back with authentication reply packet successfully, then the Authenticator will resend the authentication request packet.

tx-period-value: Specifies how long the duration of the transmission timeout timer is. The value ranges from 10 to 120 in units of second and defaults to 30.

39.2.11 Enabling/Disabling quiet-period Timer

You can use the following commands to enable/disable a quiet-period timer of an Authenticator (such as a Switch 8800). If an 802.1x user has not passed the authentication, the Authenticator will keep quiet for a while (which is specified by **dot1x timer quiet-period** command) before launching the authentication again. During the quiet period, the Authenticator does not do anything related to 802.1x authentication.

Perform the following configuration in system view.

Table 39-11 Enable/Disable a quiet-period timer

Operation	Command
Enable a quiet-period timer	dot1x quiet-period
Disable a quiet-period timer	undo dot1x quiet-period

By default, **quiet-period** timer is disabled.

39.3 Displaying and Debugging 802.1x

After the above configuration, execute **display dot1x** command in any view to display the running of the 802.1x configuration, and to verify the effect of the configuration. Execute **reset dot1x statistics** command in user view to reset 802.1x statistics. Execute **debugging** command in user view to debug 802.1x.

Table 39-12 Display and debug 802.1x

Operation	Command
Display the configuration, running and statistics information of 802.1x	display dot1x [sessions statistics enabled-interface] [interface interface-list]
Reset the 802.1x statistics information	reset dot1x statistics [interface interface-list]

Enable the error/event/packet/all debugging of 802.1x	debugging dot1x { error event packet all }
Disable the error/event/packet/all debugging of 802.1x.	undo debugging dot1x { error event packet all }

39.4 802.1x Configuration Example

I. Network requirements

As shown in Figure 39-2, the workstation of a user is connected to the port Ethernet 3/1/1 of the Switch.

The switch administrator will enable 802.1x on all the ports to authenticate the supplicants so as to control their access to the Internet. The access control mode is configured as based on the MAC address

All the supplicants belong to the default domain 3Com163.net, which can contain up to 30 users. RADIUS authentication is performed first. If there is no response from the RADIUS server, local authentication will be performed. For accounting, if the RADIUS server fails to account, the user will be disconnected. In addition, when the user is accessed, the domain name does not follow the user name. Normally, if the user's traffic is less than 2000 Byte/s consistently over 20 minutes, he will be disconnected.

A server group, consisting of two RADIUS servers at 10.11.1.1 and 10.11.1.2 respectively, is connected to the switch. The former one acts as the primary-authentication/secondary-accounting server. The latter one acts as the secondary-authentication/primary-accounting server. Set the encryption key as "name" when the system exchanges packets with the authentication RADIUS server and "money" when the system exchanges packets with the accounting RADIUS server. Configure the system to retransmit packets to the RADIUS server if no response received in 5 seconds. Retransmit the packet no more than 5 times in all. Configure the system to transmit a real-time accounting packet to the RADIUS server every 15 minutes. The system is instructed to transmit the user name to the RADIUS server after removing the user domain name from the user name.

The user name of the local 802.1x access user is localuser and the password is localpass (input in plain text). The idle cut function is enabled.

II. Network diagram

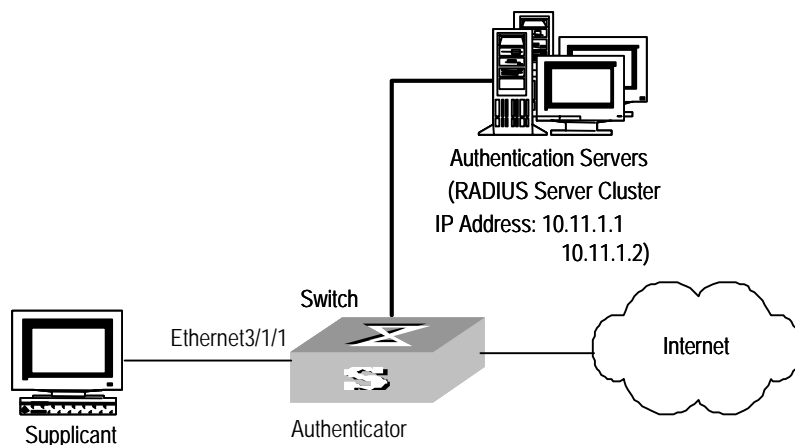


Figure 39-2 Enable 802.1x and RADIUS to perform AAA on the supplicant

III. Configuration procedure

Note:

The following examples concern most of the AAA/RADIUS configuration commands. For details, refer to the chapter AAA and RADIUS/TACACS+ Protocol Configuration. The configurations of access user workstation is omitted.

RADIUS server configuration is carried out in terms of RADIUS schemes. A RADIUS scheme actually can either be a stand-alone RADIUS server or two mutually backed up RADIUS servers with the same configuration and different IP addresses. So, for each RADIUS scheme, you need to configure the IP addresses for the primary and secondary RADIUS servers, and the shared key.

Enable 802.1x globally.

```
[SW8800] dot1x
```

Enable the 802.1x performance on the specified port Ethernet 3/1/1.

```
[SW8800] dot1x interface Ethernet 3/1/1
```

Set the access control mode. (This command could not be configured, when it is configured as MAC-based by default.)

```
[SW8800] dot1x port-method macbased interface Ethernet 3/1/1
```

Create the RADIUS scheme radius1 and enters its configuration mode.

```
[SW8800] radius scheme radius1
```

Set IP address of the primary authentication/accounting RADIUS servers.

```
[SW8800-radius-radius1] primary authentication 10.11.1.1
```

```
[SW8800-radius-radius1] primary accounting 10.11.1.2
```

Set the IP address of the secondary authentication/accounting RADIUS servers.

```
[SW8800-radius-radius1] secondary authentication 10.11.1.2
```

```
[SW8800-radius-radius1] secondary accounting 10.11.1.1
```

Set the encryption key when the system exchanges packets with the authentication RADIUS server.

```
[SW8800-radius-radius1] key authentication name
```

Set the encryption key when the system exchanges packets with the accounting RADIUS server.

```
[SW8800-radius-radius1] key accounting money
```

Set the timeouts and times for the system to retransmit packets to the RADIUS server.

```
[SW8800-radius-radius1] timer 5
```

```
[SW8800-radius-radius1] retry 5
```

Set the interval for the system to transmit real-time accounting packets to the RADIUS server.

```
[SW8800-radius-radius1] timer realtime-accounting 15
```

Configure the system to transmit the user name to the RADIUS server after removing the domain name.

```
[SW8800-radius-radius1] user-name-format without-domain
```

```
[SW8800-radius-radius1] quit
```

Create the user domain 3Com163.net and enters its configuration mode.

```
[SW8800] domain 3Com163.net
```

Specify radius1 as the RADIUS scheme for the users in the domain 3Com163.net.

```
[SW8800-isp-3Com163.net] radius-scheme radius1
```

Set a limit of 30 users to the domain 3Com163.net.

```
[SW8800-isp-3Com163.net] access-limit enable 30
```

Enable idle cut function for the user and set the idle cut parameter in the domain 3Com163.net.

```
[SW8800-isp-3Com163.net] idle-cut enable 20 2000
```

Add a local supplicant and sets its parameter.

```
[SW8800] local-user localuser
```

```
[SW8800-luser-localuser] service-type lan-access
```

```
[SW8800-luser-localuser] password simple localpass
```

Chapter 40 AAA and RADIUS/TACACS+ Protocol Configuration

40.1 AAA and RADIUS/TACACS+ Protocol Overview

40.1.1 AAA Overview

Authentication, Authorization and Accounting (AAA) provide a uniform framework used for configuring these three security functions to implement the network security management.

The network security mentioned here refers to access control and it includes:

- Which user can access the network server?
- Which service can the authorized user enjoy?
- How to keep accounts for the user who is using network resource?

Accordingly, AAA shall provide the following services:

- Authentication: authenticates if the user can access the network sever.
- Authorization: authorizes the user with specified services.
- Accounting: traces network resources consumed by the user.

Generally, AAA adopts Client/Server architecture, with its client running at the managed side and its server centralizes and stores the user information. Therefore AAA framework takes good scalability, and is easy to realize the control and centralized management of user information.

40.1.2 RADIUS Protocol Overview

As mentioned above, AAA is a management framework, so it can be implemented by some protocols. RADIUS is such a protocol frequently used.

I. What is RADIUS

Remote Authentication Dial-In User Service, RADIUS for short, is a kind of distributed information switching protocol in Client/Server architecture. RADIUS can prevent the network from interruption of unauthorized access and it is often used in the network environments requiring both high security and remote user access. For example, it is often used for managing a large number of scattering dial-in users who use serial ports and modems. RADIUS system is the important auxiliary part of Network Access Server (NAS).

After RADIUS system is started, if the user wants to have right to access other network or consume some network resources through connection to NAS (dial-in access server

in PSTN environment or Ethernet switch with access function in Ethernet environment), NAS, namely RADIUS client end, will transmit user AAA request to the RADIUS server. RADIUS server has a user database recording all the information of user authentication and network service access. When receiving user's request from NAS, RADIUS server performs AAA through user database query and update and returns the configuration information and accounting data to NAS. Here, NAS controls supplicant and corresponding connections, while RADIUS protocol regulates how to transmit configuration and accounting information between NAS and RADIUS.

NAS and RADIUS exchange the information with UDP packets. During the interaction, both sides encrypt the packets with keys before uploading user configuration information (like password etc.) to avoid being intercepted or stolen.

Note:

The authentication and authorization of a RADIUS scheme cannot be performed separately.

II. RADIUS operation

RADIUS server generally uses proxy function of the devices like access server to perform user authentication. The operation process is as follows: First, the user send request message (the client username and encrypted password is included in the message) to RADIUS server. Second, the user will receive from RADIUS server various kinds of response messages in which the ACCEPT message indicates that the user has passed the authentication, and the REJECT message indicates that the user has not passed the authentication and needs to input username and password again, otherwise he will be rejected to access.

40.1.3 TACACS+ Protocol Overview

I. TACACS+ SPECIALITY

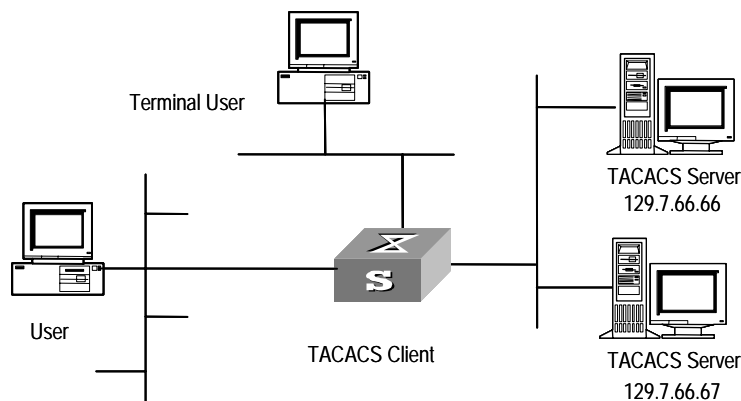
TACACS+ is an enhanced security protocol based on TACACS (RFC1492). Similar to the RADIUS protocol, it implements AAA for different types of users through communications with TACACS servers in the Server/Client model. TACACS+ can be used for the authentication, authorization and accounting of PPP and VPDN access users and login users.

Compared with RADIUS, TACACS+ provides more reliable transmission and encryption, and therefore is more suitable for security control. The following table lists the primary differences between TACACS+ and RADIUS protocols:

Table 40-1 TACACS+ vs. RADIUS

TACACS+	RADIUS
Adopts TCP, providing more reliable network transmission.	Adopts UDP.
Encrypts the entire packet except for the standard TACACS+ header.	Encrypts only the password field in authentication packets.
Separates authentication from authorization. For example, you can use RADIUS to authenticate but TACACS+ to authorize.	Binds authentication with authorization.
Suitable for security control.	Suitable for accounting.
Supports the authorization of different users to use the configuration commands of the routing module of the switch.	Not support.

Working as a client of TACACS+, the switch sends the username and password to the TACACS server for authentication, as shown in the following figure:

**Figure 40-1** Network diagram for TACACS+

II. Basic message exchange procedures in TACACS+

For example, use TACACS+ to implement authentication, authorization, and accounting for a telnet user. The basic message exchange procedures are as follows:

- A user requests access to the switch; the TACACS client sends a start-authentication packet to TACACS server upon receiving the request.
- The TACACS server sends back an authentication response requesting for the username; the TACACS client asks the user for the username upon receiving the response.
- The TACACS client sends an authentication continuance packet carrying the username after receiving the username from the user.

- The TACACS server sends back an authentication response, requesting for the login password. Upon receiving the response, the TACACS client requests the user for the login password.
- After receiving the login password, the TACACS client sends an authentication continuance packet carrying the login password to the TACACS server.
- The TACACS server sends back an authentication response indicating that the user has passed the authentication.
- The TACACS client sends the user authorization packet to the TACACS server.
- The TACACS server sends back the authorization response, indicating that the user has passed the authorization.
- Upon receipt of the response indicating an authorization success, the TACACS client pushes the configuration interface of the switch to the user.
- The TACACS client sends a start-accounting request to the TACACS server.
- The TACACS server sends back an accounting response, indicating that it has received the start-accounting request.
- The user logs off; the TACACS client sends a stop-accounting request to the TACACS server.
- The TACACS server sends a stop-accounting response to the client, which indicates it has received the stop-accounting request packet.

The following figure illustrates the basic message exchange procedures:

Figure 40-2 illustrates the basic message exchange procedures.

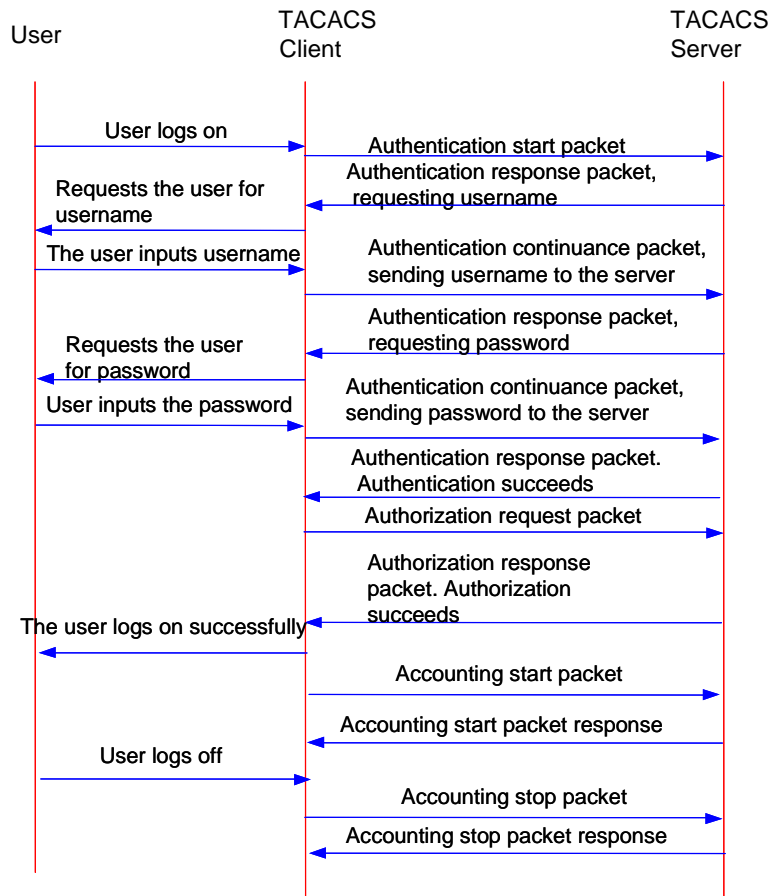


Figure 40-2 Basic message exchange procedures

40.1.4 Implementing AAA/RADIUS on a Switch

By now, we understand that in the above-mentioned AAA/RADIUS framework, a Switch 8800, serving as the user access device (NAS), is the client end of RADIUS. In other words, the AAA/RADIUS concerning client-end is implemented on the Switch 8800. Figure 40-3 illustrates the RADIUS authentication network including the Switch 8800.

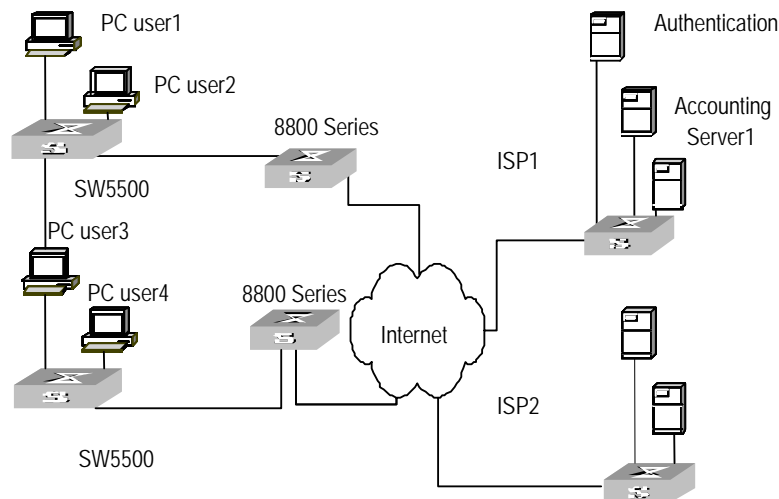


Figure 40-3 Network diagram for using RADIUS to authenticate

40.2 AAA Configuration

The following sections describe AAA configuration tasks.

- Creating/Deleting an ISP Domain
- Configuring Relevant Attributes of an ISP Domain
- Configuring Self-Service Server URL
- Creating/Deleting a Local User
- Setting the Attributes of a Local User
- Disconnecting a User by Force
- Configuring Dynamic VLAN Delivering

Among the above configuration tasks, creating ISP domain is compulsory, otherwise the supplicant attributes cannot be distinguished. The other tasks are optional. You can configure them at requirements.

40.2.1 Creating/Deleting an ISP Domain

What is Internet Service Provider (ISP) domain? To make it simple, ISP domain is a group of users belonging to the same ISP. Generally, for a username in the `userid@isp-name` format, taking `gw20010608@3Com163.net` as an example, the `isp-name` (i.e. `3Com163.net`) following the `@` is the ISP domain name. When a Switch 8800 controls user access, as for an ISP user whose username is in `userid@isp-name` format, the system will take `userid` part as username for identification and take `isp-name` part as domain name.

The purpose of introducing ISP domain settings is to support the multi-ISP application environment. In such environment, one access device might access users of different ISP. Because the attributes of ISP users, such as username and password formats, etc, may be different, it is necessary to differentiate them through setting ISP domain. In the

Switch 8800 ISP domain view, you can configure a complete set of exclusive ISP domain attributes on a per-ISP domain basis, which includes AAA policy (RADIUS scheme applied etc.)

For the Switch 8800, each supplicant belongs to an ISP domain. Up to 16 domains can be configured in the system. If a user has not reported its ISP domain name, the system will put it into the default domain.

Perform the following configuration in system view.

Table 40-2 Create/Delete an ISP domain

Operation	Command
Create ISP domain or enter the view of a specified domain	domain <i>isp-name</i>
Remove a specified ISP domain	undo domain <i>isp-name</i>
Enable the default ISP domain specified by <i>isp-name</i>	domain default enable <i>isp-name</i>
Restore the default ISP domain to "system"	domain default disable

By default, a domain named "system" has been created in the system. The attributes of "system" are all default values.

40.2.2 Configuring Relevant Attributes of an ISP Domain

The relevant attributes of ISP domain include the adopted RADIUS scheme, ISP domain state, maximum number of supplicants, accounting optional enable/disable state, address pool definition, IP address assignment for PPP domain users, and user idle-cut enable/disable state where:

- The adopted RADIUS scheme is the one used by all the users in the ISP domain. The RADIUS scheme can be used for RADIUS authentication or accounting. By default, the default RADIUS scheme is used. The command shall be used together with the commands of setting RADIUS server and server cluster. For details, refer to the following Configuring RADIUS section of this chapter. If local is configured as the first scheme, only the local scheme will be adopted, neither RADIUS nor TACACS+ scheme will be adopted. When local scheme is adopted, only authentication and authorization will be performed, accounting will not be performed. None has the same effect as local. The usernames used for local authentication carry no domain name, so if the local scheme is configured, pay attention not to add domain name to the username when you configure a local user.
- Every ISP domain has two states: active and block. If an ISP domain is in active state, the users in it are allowed to request network services, while in block state, its users are inhibit from requesting any network service, which will not affect the

users already online. An ISP is in active state once it is created, that is, at that time, all the users in the domain are allowed to request network services.

- Maximum number of supplicants specifies how many supplicants can be contained in the ISP. For any ISP domain, there is no limit to the number of supplicants by default.
- The idle cut function means: If the traffic from a certain connection is lower than the defined traffic, cut off this connection.
- The PPP access users can obtain IP addresses through the PPP address negotiation function.

Perform the following configuration in ISP domain view.

Table 40-3 Configure relevant attributes of an ISP domain

Operation	Command
Configure the AAA scheme used by an ISP domain	scheme { radius-scheme <i>radius-scheme-name</i> [local] TACACS+-scheme <i>TACACS+-scheme-name</i> [local] local none }
Restore the default AAA scheme used by an ISP domain	undo scheme { radius-scheme TACACS+-scheme none }
Specify the ISP domain state to be used	state { active block }
Set a limit to the amount of supplicants	access-limit { disable enable <i>max-user-number</i> }
Restore the limit to the default setting	undo access-limit
Enable accounting to be optional	accounting optional
Disable accounting to be optional	undo accounting optional
Set the idle	idle-cut { disable enable <i>minute</i> <i>flow</i> }
Define an address pool to assign IP addresses to users	ip pool <i>pool-number</i> <i>low-ip-address</i> [<i>high-ip-address</i>]
Delete the specified address pool	undo ip pool <i>pool-number</i>

By default, the local scheme is adopted, an ISP domain is in active state once it is created, no limit is set to the amount of supplicants, accounting optional is disabled, idle-cut is disabled, and no IP address pool is defined.

40.2.3 Configuring Self-Service Server URL

The **self-service-url enable** command must be incorporated with a RADIUS server (such as a CAMS server) that supports self-service. Self-service means that users can manage their accounts and card numbers by themselves. And a server with the self-service software is called a self-service server.

Once this function is enabled on the switch, users can locate the self-service server through the following operations:

- Select "Change user password" on the 802.1x client.
- After the client opens the default explorer (IE or NetScape), locate the specified URL page used to change the user password on the self-service server.
- Change user password on this page.

Perform the following configuration in ISP domain view.

Table 40-4 Configure the self-service server URL

Operation	Command
Configure self-service server URL and configure the URL address used to change the user password on the self-service server	self-service-url enable <i>url-string</i>
Remove the configuration of self-service server URL	self-service-url disable

By default, self-service server URL is not configured on the switch.

Note that, if "?" is contained in the URL, you must replace it with "|" when inputting the URL in the command line.

The "Change user password" option is available only when the user passes the authentication; otherwise, this option is in grey and unavailable.

40.2.4 Creating/Deleting a Local User

A local user is a group of users set on NAS. The username is the unique identifier of a user. A supplicant requesting network service may use local authentication only if its corresponding local user has been added onto NAS.

Perform the following configuration in system view.

Table 40-5 Create/Delete a local user

Operation	Command
Add a local user	local-user <i>user-name</i>
Delete all the local users	undo local-user all
Delete a local user by specifying its type	undo local-user { <i>user-name</i> all [service-type { lan-access ftp telnet ppp ssh terminal }] }

By default, there is no local user in the system.

40.2.5 Setting the Attributes of a Local User

The attributes of a local user include its password display mode, state, service type and some other settings.

I. Setting the password display mode

Perform the following configuration in system view.

Table 40-6 Set the method that a local user uses to display password

Operation	Command
Set the mode that a local user uses to display password	local-user password-display-mode { cipher-force auto }
Cancel the mode that the local user uses to display password	undo local-user password-display-mode

Where, **auto** means that the password display mode will be the one specified by the user at the time of configuring password (see the **password** command in the following table for reference), and **cipher-force** means that the password display mode of all the accessing users must be in cipher text.

II. Setting/Removing the attributes of a local user

Perform the following configuration in local user view.

Table 40-7 Set/Remove the attributes concerned with a specified user

Operation	Command
Set a password for a specified user	password { simple cipher } <i>password</i>
Remove the password set for the specified user	undo password
Set the state of the specified user	state { active block }
Set a service type for the specified user	service-type { ftp [ftp-directory <i>directory</i>] lan-access ppp [call-number <i>call-number</i> callback-nocheck callback-number <i>callback-number</i>] ssh [level <i>level</i> telnet terminal] telnet [level <i>level</i> ssh terminal] terminal [level <i>level</i> ssh telnet] }
Cancel the service type of the specified user	undo service-type { ftp [ftp-directory <i>directory</i>] lan-access ppp [call-number <i>call-number</i> callback-nocheck callback-number <i>callback-number</i>] ssh [level <i>level</i> telnet terminal] telnet [level <i>level</i> ssh terminal] terminal [level <i>level</i> ssh telnet] }

Operation	Command
Set the priority of the specified user	level <i>level</i>
Restore the default priority of the specified user	undo level
Configure the attributes of lan-access users	attribute { ip <i>ip-address</i> mac <i>mac-address</i> idle-cut <i>second</i> access-limit <i>max-user-number</i> vlan <i>vlanid</i> location { nas-ip <i>ip-address</i> port <i>portnum</i> port <i>portnum</i> }*
Remove the attributes defined for the lan-access users	undo attribute { ip mac idle-cut access-limit vlan location }*

By default, users are not authorized to any service, all their priorities are 0.

40.2.6 Disconnecting a User by Force

Sometimes it is necessary to disconnect a user or a category of users by force. The system provides the following command to serve for this purpose.

Perform the following configuration in system view.

Table 40-8 Disconnect a user by force

Operation	Command
Disconnect a user by force	cut connection { all access-type { dot1x gcm mac-authentication } domain <i>domain-name</i> interface <i>interface-type interface-number</i> ip <i>ip-address</i> mac <i>mac-address</i> radius-scheme <i>radius-scheme-name</i> vlan <i>vlanid</i> ucibindex <i>ucib-index</i> user-name <i>user-name</i> }

40.2.7 Configuring Dynamic VLAN Delivering

Dynamic VLAN delivering enables an Ethernet switch to monitor network resources available to users by adding the ports to which the authenticated users connect to different VLANs according to the properties delivered by RADIUS servers. To work with Guest VLAN, ports are usually configured to perform port-based authentications. (If you configure a port to perform MAC address-based authentication, it can have only one user connected.)

At present, a switch supports VLAN IDs delivered by RADIUS servers to be of string type. The port is added to the VLANs on a switch with their IDs matching the one delivered by the RADIUS servers. If this kind of VLANs does not exist, the VLAN delivering fails and the user fails to pass the authentication.

Perform the following configuration in system view.

Table 40-9 Configure VLAN delivering mode

Operation	Command
Configure the VLAN delivering mode to be of string type	private-group-id mode standard
Revert to the default VLAN delivering mode.	undo private-group-id mode standard

By default, a VLAN ID cannot be a string.

40.3 Configuring RADIUS Protocol

For the Switch 8800, the RADIUS protocol is configured on the per RADIUS scheme basis. In real networking environment, a RADIUS scheme can be an independent RADIUS server or a set of primary/secondary RADIUS servers with the same configuration but two different IP addresses. Accordingly, attributes of every RADIUS scheme include IP addresses of primary and secondary servers, shared key and RADIUS server type etc.

Actually, RADIUS protocol configuration only defines some necessary parameters using for information interaction between NAS and RADIUS Server. To make these parameters take effect on an ISP domain, you must configure the ISP domain to use the RADIUS scheme configured with these parameters in ISP domain view. For more about the configuration commands, refer to the AAA Configuration section above.

The following sections describe RADIUS protocol configuration tasks.

- Creating/Deleting a RADIUS scheme
- Setting IP Address and Port Number of a RADIUS Server
- Setting the RADIUS Packet Encryption Key
- Setting the Response Timeout Timer of a RADIUS Server
- Setting the Retransmission Times of RADIUS Request Packets
- Enabling the Selection Of Radius Accounting Option
- Setting a Real-time Accounting Interval
- Setting the Maximum Times of Real-time Accounting Request Failing to be Responded
- Enabling/Disabling Stopping Accounting Request Buffer
- Setting the Maximum Retransmitting Times of Stopping Accounting Request
- Setting the Supported Type of RADIUS Server
- Setting RADIUS Server State
- Setting the Username Format Transmitted to RADIUS Server
- Setting the Unit of Data Flow that Transmitted to RADIUS Server
- Creating/Deleting a Local RADIUS authentication Server

Among the above tasks, creating RADIUS scheme and setting IP address of RADIUS server are required, while other tasks are optional and can be performed as your requirements.

40.3.1 Creating/Deleting a RADIUS scheme

As mentioned above, RADIUS protocol configurations are performed on the per RADIUS scheme basis. Therefore, before performing other RADIUS protocol configurations, it is compulsory to create the RADIUS scheme and enter its view.

You can use the following commands to create/delete a RADIUS scheme.

Perform the following configuration in system view.

Table 40-10 Create/Delete a RADIUS server group

Operation	Command
Create a RADIUS server group and enter its view	radius scheme <i>radius-server-name</i>
Delete a RADIUS server group	undo radius scheme <i>radius-server-name</i>

Several ISP domains can use a RADIUS server group at the same time. You can configure up to 16 RADIUS schemes, including the default server group named as System.

By default, the system has a RADIUS scheme named "system" whose attributes are all default values.

40.3.2 Setting IP Address and Port Number of a RADIUS Server

After creating a RADIUS scheme, you are supposed to set IP addresses and UDP port numbers for the RADIUS servers, including primary/secondary authentication/authorization servers and accounting servers. So you can configure up to 4 groups of IP addresses and UDP port numbers. However, at least you have to set one group of IP address and UDP port number for each pair of primary/secondary servers to ensure the normal AAA operation.

You can use the following commands to configure the IP address and port number for RADIUS schemes.

Perform the following configuration in RADIUS scheme view.

Table 40-11 Set IP Address and Port Number of RADIUS Server

Operation	Command
Set IP address and port number of primary RADIUS authentication/authorization server.	primary authentication <i>ip-address [port-number]</i>
Restore IP address and port number of primary RADIUS authentication/authorization or server to the default values.	undo primary authentication
Set IP address and port number of primary RADIUS accounting server.	primary accounting <i>ip-address [port-number]</i>
Restore IP address and port number of primary RADIUS accounting server or server to the default values.	undo primary accounting
Set IP address and port number of secondary RADIUS authentication/authorization server.	secondary authentication <i>ip-address [port-number]</i>
Restore IP address and port number of secondary RADIUS authentication/authorization or server to the default values.	undo secondary authentication
Set IP address and port number of secondary RADIUS accounting server.	secondary accounting <i>ip-address [port-number]</i>
Restore IP address and port number of secondary RADIUS accounting server or server to the default values.	undo secondary accounting

By default, as for the "system" RADIUS scheme created by the system:

The IP address of the primary authentication server is 127.0.0.1, and the UDP port number is 1645.

The IP address of the secondary authentication server is 0.0.0.0, and the UDP port number is 1812.

The IP address of the primary accounting server is 127.0.0.1, and the UDP port number is 1646

The IP address of the secondary accounting server is 0.0.0.0, and the UDP port number is 1813;

As for the newly created RADIUS scheme:

The IP address of the primary/secondary authentication server is 0.0.0.0, and the UDP port number of this server is 1812;

The IP address of the primary/secondary accounting server is 0.0.0.0, and the UDP port number of this server is 1813;

In real networking environments, the above parameters shall be set according to the specific requirements. For example, you may specify 4 groups of different data to map

4 RADIUS servers, or specify one of the two servers as primary authentication/authorization server and secondary accounting server and the other one as secondary authentication/authorization server and primary accounting server, or you may also set 4 groups of exactly same data so that every server serves as a primary and secondary AAA server.

To guarantee the normal interaction between NAS and RADIUS server, you are supposed to guarantee the normal routes between RADIUS/TACACS+ server and NAS before setting IP address and UDP port of the RADIUS/TACACS+ server. In addition, because RADIUS/TACACS+ protocol uses different UDP ports to receive/transmit authentication/authorization and accounting packets, you shall set two different ports accordingly. Suggested by RFC2138/2139, authentication/authorization port number is 1812 and accounting port number is 1813. However, you may use values other than the suggested ones. (Especially for some earlier RADIUS/TACACS+ Servers, authentication/authorization port number is often set to 1645 and accounting port number is 1646.)

The RADIUS/TACACS+ service port settings on the Switch 8800 are supposed to be consistent with the port settings on the RADIUS server. Normally, RADIUS accounting service port is 1813 and the authentication/authorization service port is 1812.

Note:

For a Switch 8800, the default RADIUS scheme authentication/authorization port is 1645, the accounting port is 1646. And port 1812 and 1813 are for other schemes.

40.3.3 Setting the RADIUS Packet Encryption Key

RADIUS client (switch system) and RADIUS server use MD5 algorithm to encrypt the exchanged packets. The two ends verify the packet through setting the encryption key. Only when the keys are identical can both ends to accept the packets from each other end and give response.

You can use the following commands to set the encryption key for RADIUS packets.

Perform the following configuration in RADIUS scheme view.

Table 40-12 Set RADIUS packet encryption key

Operation	Command
Set RADIUS authentication/authorization packet encryption key	key authentication <i>string</i>
Restore the default RADIUS authentication/authorization packet encryption key	undo key authentication

Set RADIUS accounting packet encryption key	key accounting <i>string</i>
Restore the default RADIUS accounting packet encryption key	undo key accounting

By default, the encryption keys of RADIUS authentication/authorization and accounting packets are all “3Com”.

40.3.4 Setting the Response Timeout Timer of a RADIUS Server

After RADIUS (authentication/authorization or accounting) request packet has been transmitted for a period of time, if NAS has not received the response from RADIUS server, it has to retransmit the request to guarantee RADIUS service for the user.

You can use the following command to set response timeout timer of RADIUS server.

Perform the following configuration in RADIUS scheme view.

Table 40-13 Set the response timeout timer of a RADIUS server

Operation	Command
Set response timeout timer of RADIUS server	timer <i>second</i>
Restore the response timeout timer of RADIUS server to default value	undo timer

By default, timeout timer of a RADIUS server is 3 seconds.

40.3.5 Setting the Retransmission Times of RADIUS Request Packets

Since RADIUS protocol uses UDP packet to carry the data, the communication process is not reliable. If the RADIUS server has not responded NAS before timeout, NAS has to retransmit RADIUS request packet. If it transmits more than the specified *retry-times*, NAS considers the communication with the current RADIUS server has been disconnected, and turn to send request packet to other RADIUS server.

You can use the following command to set retransmission times of RADIUS request packet.

Perform the following configuration in RADIUS scheme view.

Table 40-14 Set the retransmission times of RADIUS request packets

Operation	Command
Set retransmission times of RADIUS request packet	retry <i>retry-times</i>
Restore the default value of retransmission times	undo retry

By default, RADIUS request packet will be retransmitted up to three times.

40.3.6 Enabling the Selection Of Radius Accounting Option

If no RADIUS server is available or if RADIUS accounting server fails when the **accounting optional** is configured, the user can still use the network resource, otherwise, the user will be disconnected.

Perform the following configuration in RADIUS scheme view.

Table 40-15 Enable the selection of RADIUS accounting option

Operation	Command
Enable the selection of RADIUS accounting option	accounting optional
Disable the selection of RADIUS accounting option	undo accounting optional

By default, selection of RADIUS accounting option is disabled.

40.3.7 Setting a Real-time Accounting Interval

To implement real-time accounting, it is necessary to set a real-time accounting interval. After the attribute is set, NAS will transmit the accounting information of online users to the RADIUS server regularly.

You can use the following command to set a real-time accounting interval.

Perform the following configuration in RADIUS scheme view.

Table 40-16 Set a real-time accounting interval

Operation	Command
Set a real-time accounting interval	timer realtime-accounting <i>minute</i>
Restore the default value of the interval	undo timer realtime-accounting

minute specifies the real-time accounting interval in minutes. The value shall be a multiple of 3.

The value of *minute* is related to the performance of NAS and RADIUS server. The smaller the value is, the higher the performances of NAS and RADIUS are required. When there are a large amount of users (more than 1000, inclusive), we suggest a larger value. The following table recommends the ratio of *minute* value to the number of users.

Table 40-17 Recommended real-time accounting intervals for different number of users

Number of users	Real-time accounting interval in minutes
1 to 99	3
100 to 499	6
500 to 999	12
≥1000	≥15

By default, *minute* is set to 12 minutes.

40.3.8 Setting the Maximum Times of Real-time Accounting Request Failing to be Responded

RADIUS server usually checks if a user is online with timeout timer. If the RADIUS server has not received the real-time accounting packet from NAS for long, it will consider that there is device failure and stop accounting. Accordingly, it is necessary to disconnect the user at NAS end and on RADIUS server synchronously when some unpredictable failure exists. The Switch 8800 supports setting maximum times of real-time accounting request failing to be responded. NAS will disconnect the user if it has not received real-time accounting response from RADIUS server for some specified times.

You can use the following command to set the maximum times of real-time accounting request failing to be responded.

Perform the following configuration in RADIUS scheme view.

Table 40-18 Set the maximum times of real-time accounting request failing to be responded

Operation	Command
Set maximum times of real-time accounting request failing to be responded	retry realtime-accounting <i>retry-times</i>
Restore the maximum times to the default value	undo realtime-accounting retry

How to calculate the value of *retry-times*? Suppose that RADIUS server connection will timeout in T and the real-time accounting interval of NAS is t, then the integer part of the result from dividing T by t is the value of *count*. Therefore, when applied, T is suggested the numbers which can be divided exactly by t.

By default, the real-time accounting request can fail to be responded no more than 5 times.

40.3.9 Enabling/Disabling Stopping Accounting Request Buffer

Because the stopping accounting request concerns account balance and will affect the amount of charge, which is very important for both the subscribers and the ISP, NAS shall make its best effort to send the request to RADIUS accounting server. Accordingly, if the request from the Switch 8800 to RADIUS accounting server has not been responded, switch shall save it in the local buffer and retransmit it until the server responds or discards the messages after transmitting for specified times. You can use the following command to set whether or not to save the stopping accounting requests.

Perform the following configuration in RADIUS scheme view.

Table 40-19 Enable/Disable stopping accounting request buffer

Operation	Command
Enable stopping accounting request buffer	stop-accounting-buffer enable
Disable stopping accounting request buffer	undo stop-accounting-buffer enable

By default, the stopping accounting request will be saved in the buffer.

40.3.10 Setting the Maximum Retransmitting Times of Stopping Accounting Request

Because the stopping accounting request concerns account balance and will affect the amount of charge, which is very important for both the subscribers and the ISP, NAS shall make its best effort to send the message to RADIUS accounting server. Accordingly, if the request from the Switch 8800 to RADIUS accounting server has not been responded, switch shall save it in the local buffer and retransmit it until the server responds or discards the messages after transmitting for specified times. Use the following command to set the maximum retransmission times.

Perform the following configuration in RADIUS scheme view.

Table 40-20 Set the maximum retransmitting times of stopping accounting request

Operation	Command
Set the maximum retransmitting times of stopping accounting request	retry stop-accounting retry-times
Restore the maximum retransmitting times of stopping accounting request to the default value	undo retry stop-accounting

By default, the stopping accounting request can be retransmitted for up to 500 times.

40.3.11 Setting the Supported Type of RADIUS Server

The Switch 8800 supports the standard RADIUS protocol and the extended RADIUS service platforms, such as IP Hotel, 201+ and Portal.

You can use the following command to set the supported types of RADIUS servers.

Perform the following configuration in RADIUS scheme view.

Table 40-21 Set the supported type of RADIUS scheme

Operation	Command
Set the Supported Type of RADIUS Server	server-type { 3Com standard }
Restore the Supported Type of RADIUS Server to the default setting	undo server-type

By default, the newly created RADIUS scheme supports the server of **standard** type, while the "system" RADIUS scheme created by the system supports the server of **3Com** type

40.3.12 Setting RADIUS Server State

For the primary and secondary servers (no matter it is an authentication/authorization server or accounting server), if the primary is disconnected to NAS for some fault, NAS will automatically turn to exchange packets with the secondary server. However, after the primary one recovers, NAS will not resume the communication with it at once, instead, it continues communicating with the secondary one. When the secondary one fails to communicate, NAS will turn to the primary one again. The following commands can be used to set the primary server to be **active** manually, in order that NAS can communicate with it right after the troubleshooting.

When the primary and secondary servers are both **active** or **block**, NAS will send the packets to the primary server only.

Perform the following configuration in RADIUS scheme view.

Table 40-22 Set RADIUS server state

Operation	Command
Set the state of primary RADIUS server	state primary { accounting authentication } { block active }
Set the state of secondary RADIUS sever	state secondary { accounting authentication } { block active }

By default, the state of each server in RADIUS scheme server group is **active**.

40.3.13 Setting the Username Format Transmitted to RADIUS Server

As mentioned above, the supplicants are generally named in `userid@isp-name` format. The part following "@" is the ISP domain name. The Switch 8800 will put the users into different ISP domains according to the domain names. However, some earlier RADIUS servers reject the username including ISP domain name. In this case, you have to remove the domain name before sending the username to the RADIUS server. The following command of switch decides whether the username to be sent to RADIUS server carries ISP domain name or not.

Perform the following configuration in RADIUS scheme view.

Table 40-23 Set the username format transmitted to RADIUS server

Operation	Command
Set Username Format Transmitted to RADIUS Server	<code>user-name-format { with-domain without-domain }</code>

Note:

If a RADIUS scheme is configured not to allow usernames including ISP domain names, the RADIUS scheme shall not be simultaneously used in more than one ISP domain. Otherwise, the RADIUS server will regard two users in different ISP domains as the same user by mistake, if they have the same username (excluding their respective domain names.)

By default, as for the newly created RADIUS scheme, the username sent to RADIUS servers includes an ISP domain name; as for the "system" RADIUS scheme created by the system, the username sent to RADIUS servers excludes the ISP domain name.

40.3.14 Setting the Unit of Data Flow that Transmitted to RADIUS Server

The following command defines the unit of the data flow sent to RADIUS server.

Perform the following configuration in RADIUS scheme view.

Table 40-24 Set the unit of data flow transmitted to RADIUS server

Operation	Command
Set the unit of data flow transmitted to RADIUS server	<code>data-flow-format data { byte giga-byte kilo-byte mega-byte } packet { giga-byte kilo-byte mega-byte one-packet }</code>
Restore the unit to the default setting	<code>undo data-flow-format</code>

By default, the default data unit is byte and the default data packet unit is one packet.

40.3.15 Creating/Deleting a Local RADIUS authentication Server

RADIUS service, which adopts authentication/authorization/accounting servers to manage users, is widely used in the Switch 8800. Besides, local authentication/authorization service is also used in these products and it is called local RADIUS function, i.e. realize basic RADIUS function on the switch.

Perform the following configuration in system view.

Table 40-25 Creating/Deleting a local RADIUS authentication server

Operation	Command
Create a local RADIUS authentication server	local-server nas-ip <i>ip-address</i> key <i>password</i>
Delete a local RADIUS authentication server	undo local-server nas-ip <i>ip-address</i>

By default, the IP address of local RADIUS authentication server group is 127.0.0.1 and the password is 3Com.

When using local RADIUS server function, note that,

- 1) The number of UDP port used for authentication/authorization is 1645 and that for accounting is 1646.
- 2) The *password* configured by **local-server** command must be the same as that of the RADIUS authentication/authorization packet configured by the command **key authentication** in radius scheme view.
- 3) The Switch 8800 serving as a local RADIUS authentication server currently only supports the CHAP and PAP authentication modes; they do not support the MD5-challenge mode.

40.4 Configuring TACACS+ Protocol

The following sections describe TACACS+ configuration tasks.

- Creating a HWTACAS Scheme
-

Configuring TACACS+ Authentication Servers

- Configuring TACACS+ Authorization Servers
- Configuring TACACS+ Accounting Servers and the Related Attributes
- Configuring the Source Address for TACACS+ Packets Sent by NAS
- Setting a Key for Securing the Communication with TACACS Server
- Setting the Username Format Acceptable to the TACACS Server
- Setting the Unit of Data Flows Destined for the TACACS Server
- Setting Timers Regarding TACACS Server

Note:

Pay attention to the following when configuring a TACACS server:

- TACACS+ server does not check whether a scheme is being used by users when changing most of HWTACS attributes, unless you delete the scheme.
- By default, the TACACS server has no key.

In the above configuration tasks, creating TACACS+ scheme and configuring TACACS authentication/authorization server are required; all other tasks are optional and you can determine whether to perform these configurations as needed.

40.4.1 Creating a HWTACAS Scheme

As aforementioned, TACACS+ protocol is configured scheme by scheme. Therefore, you must create a TACACS+ scheme and enter TACACS+ view before you perform other configuration tasks.

Perform the following configuration in system view.

Table 40-26 Create a TACACS+ scheme

Operation	Command
Create a TACACS+ scheme and enter TACACS+ view	TACACS+ <i>TACACS+-scheme-name</i> scheme
Delete a TACACS+ scheme	undo TACACS+ <i>TACACS+-scheme-name</i> scheme

By default, no TACACS+ scheme exists.

If the TACACS+ scheme you specify does not exist, the system creates it and enters TACACS+ view. In TACACS+ view, you can configure the TACACS+ scheme specifically.

The system supports up to 16 TACACS+ schemes. You can only delete the schemes that are not being used.

40.4.2 Configuring TACACS+ Authentication Servers

Perform the following configuration in TACACS+ view.

Table 40-27 Configure TACACS+ authentication servers

Operation	Command
Configure the TACACS+ primary authentication server	primary authentication <i>ip-address</i> [<i>port</i>]
Delete the TACACS+ primary authentication server	undo primary authentication
Configure the TACACS+ secondary authentication server	secondary authentication <i>ip-address</i> [<i>port</i>]
Delete the TACACS+ secondary authentication server	undo secondary authentication

The primary and secondary authentication servers cannot use the same IP address. The default port number is 49.

If you execute this command repeatedly, the new settings will replace the old settings.

A TACACS scheme authentication server can be deleted only when no active TCP connection used to send authentication packets is using the server.

40.4.3 Configuring TACACS+ Authorization Servers

Perform the following configuration in TACACS+ view.

Table 40-28 Configure TACACS+ authorization servers

Operation	Command
Configure the primary TACACS+ authorization server	primary authorization <i>ip-address</i> [<i>port</i>]
Delete the primary TACACS+ authorization server	undo primary authorization
Configure the secondary TACACS+ authorization server	secondary authorization <i>ip-address</i> [<i>port</i>]
Delete the secondary TACACS+ authorization server	undo secondary authorization

Note:

If only authentication and accounting servers are configured and no authorization server is configured, both authentication and accounting can be performed normally for the ftp, telnet, and ssh users, but the priority of these users is 0 (that is, the lowest privilege level) by default,

The primary and secondary authorization servers cannot use the same IP address. The default port number is 49.

If you execute this command repeatedly, the new settings will replace the old settings.

40.4.4 Configuring TACACS+ Accounting Servers and the Related Attributes

I. Configuring TACACS+ accounting servers

Perform the following configuration in TACACS+ view.

Table 40-29 Configure TACACS+ accounting servers

Operation	Command
Configure the primary TACACS accounting server	primary accounting <i>ip-address</i> [<i>port</i>]
Delete the primary TACACS accounting server	undo primary accounting
Configure the secondary TACACS accounting server	secondary accounting <i>ip-address</i> [<i>port</i>]
Delete the secondary TACACS accounting server	undo secondary accounting

Do not configure the same IP address for the primary accounting server and the secondary accounting server. Otherwise, an error occurs.

By default, a TACACS accounting server uses an all-zero IP address and port 49.

If you execute the **primary accounting** or **secondary accounting** command repeatedly, the newly configured settings overwrite the corresponding existing settings.

You can delete a TACACS scheme only when no active TCP connection used to send authentication packets uses the server.

II. Enabling stop-accounting packet retransmission

Perform the following configuration in TACACS+ view.

Table 40-30 Configure stop-accounting packet retransmission

Operation	Command
Enable stop-accounting packet retransmission and set the allowed maximum number of transmission attempts	retry stop-accounting <i>retry-times</i>
Disable stop-accounting packet retransmission	undo retry stop-accounting
Clear the stop-accounting request packets that have no response	reset stop-accounting-buffer TACACS+-scheme <i>TACACS+-scheme-name</i>

By default, stop-accounting packet retransmission is enabled, and the maximum number of transmission attempts is 300.

40.4.5 Configuring the Source Address for TACACS+ Packets Sent by NAS

Perform the following configuration in the corresponding view.

Table 40-31 Configure the source address for TACACS+ packets sent by the NAS

Operation	Command
Configure the source address for TACACS+ packets sent from the NAS (TACACS+ view)	nas-ip <i>ip-address</i>
Delete the configured source address for TACACS+ packets sent from the NAS (TACACS+ view)	undo nas-ip
Configure the source address for TACACS+ packets sent from the NAS (System view)	TACACS+ nas-ip <i>ip-address</i>
Cancel the configured source address for TACACS+ packets sent from the NAS (System view)	undo TACACS+ nas-ip

The TACACS+ view takes precedence over the system view when configuring the source address for TACACS+ packets sent from the NAS.

By default, the source address is not specified, and the virtual interface of the VLAN that contains the port to which the server connects for packet sending is used as the source address.

40.4.6 Setting a Key for Securing the Communication with TACACS Server

When using a TACACS server as an AAA server, you can set a key to improve the communication security between the switch and the TACACS server.

Perform the following configuration in TACACS+ view.

Table 40-32 Set a key for securing the communication with the TACACS+ server

Operation	Command
Configure a key for securing the communication with the accounting, authorization or authentication server	key { accounting authorization authentication } string
Delete the configuration	undo key { accounting authorization authentication }

No key is configured by default.

40.4.7 Setting the Username Format Acceptable to the TACACS Server

Username is usually in the “userid@isp-name” format, with the domain name following “@”.

If a TACACS server does not accept the username with domain name, you can remove the domain name and resend it to the TACACS server.

Perform the following configuration in TACACS+ view.

Table 40-33 Set the username format acceptable to the TACACS server

Operation	Command
Send username with domain name	user-name-format with-domain
Send username without domain name	user-name-format without-domain

By default, each username sent to a TACACS server contains a domain name.

40.4.8 Setting the Unit of Data Flows Destined for the TACACS Server

Perform the following configuration in TACACS+ view.

Table 40-34 Set the unit of data flows destined for the TACACS server

Operation	Command
Set the unit of data flows destined for the TACACS server	data-flow-format data { byte giga-byte kilo-byte mega-byte } data-flow-format packet { giga-packet kilo-packet mega-packet one-packet }
Restore the default unit of data flows destined for the TACACS server	undo data-flow-format { data packet }

The default data flow unit is byte.

40.4.9 Setting Timers Regarding TACACS Server

I. Setting the response timeout timer

Since TACACS+ is implemented on the basis of TCP, server response timeout or TCP timeout may terminate the connection to the TACACS server.

Perform the following configuration in TACACS+ view.

Table 40-35 Set the response timeout timer

Operation	Command
Set the response timeout time	timer response-timeout <i>seconds</i>
Restore the default setting	undo timer response-timeout

The default response timeout timer is set to 5 seconds.

II. Setting the quiet timer for the primary TACACS server

Perform the following configuration in TACACS+ view.

Table 40-36 Set the quiet timer for the primary TACACS server

Operation	Command
Set the quiet timer for the primary TACACS server	timer quiet <i>minutes</i>
Restore the default setting	undo timer quiet

The **timer quiet** command is used to make the switch ignore users' requests for server within the time configured in this command in case the communication between the switch and the server is terminated. In that case, the switch can send users' requests to the server only after it has waited a time no less than the time configured with this command for the communication to be resumed.

By default, the primary TACACS server must wait five minutes before it can resume the active state. The time ranges from 1 to 255.

III. Setting a realtime accounting interval

The setting of real-time accounting interval is necessary to real-time accounting. After an interval value is set, the NAS transmits the accounting information of online users to the TACACS accounting server periodically.

Perform the following configuration in TACACS+ view.

Table 40-37 Set a real-time accounting interval

Operation	Command
Set a real-time accounting interval	timer realtime-accounting <i>minutes</i>
Restore the default real-time accounting interval	undo timer realtime-accounting

The interval is in minutes and must be a multiple of 3.

The setting of real-time accounting interval somewhat depends on the performance of the NAS and the TACACS server: a shorter interval requires higher device performance. You are therefore recommended to adopt a longer interval when there are a large number of users (more than 1000, inclusive). The following table lists the numbers of users and the recommended intervals.

Table 40-38 Numbers of users and the recommended intervals

Number of users	Real-time accounting interval (in minutes)
1 to 99	3
100 to 499	6
500 to 999	12
<i>f</i> 1000	<i>f</i> 15

The real-time accounting interval defaults to 12 minutes.

40.5 Displaying and Debugging AAA and RADIUS Protocol

After the above configuration, execute **display** command in any view to display the running of the AAA and RADIUS/TACACS+ configuration, and to verify the effect of the configuration. Execute **reset** command in user view to reset AAA and RADIUS/TACACS+ statistics, etc. Execute **debugging** command in user view to debug AAA and RADIUS/TACACS+.

Table 40-39 Display and debug AAA and RADIUS/TACACS+ protocol

Operation	Command
Display the configuration information of the specified or all the ISP domains	display domain [<i>isp-name</i>]

Operation	Command
Display related information of user's connection	display connection { access-type { dot1x gcm } domain <i>isp-name</i> interface <i>interface-type</i> <i>interface-number</i> ip <i>ip-address</i> mac <i>mac-address</i> radius-scheme <i>radius-scheme-name</i> vlan <i>vlanid</i> ucibindex <i>ucib-index</i> user-name <i>user-name</i> }
Display related information of the local user	display local-user [domain <i>isp-name</i> idle-cut { disable enable } service-type { ftp lan-access ppp ssh telnet terminal } state { active block } user-name <i>user-name</i> vlan <i>vlanid</i>]
Display the statistics of local RADIUS server group	display local-server { statistics nas-ip }
Display the configuration information of all the RADIUS server groups or a specified one	display radius [<i>radius-server-name</i>]
Display the statistics of RADIUS packets	display radius statistics
Display the stop-accounting requests saved in buffer without response	display stop-accounting-buffer { radius-scheme <i>radius-scheme-name</i> session-id <i>session-id</i> time-range <i>start-time stop-time</i> user-name <i>user-name</i> }
Reset the statistics of RADIUS server	reset radius statistics
Display the specified or all the TACACS+ schemes	display TACACS+ [<i>TACACS+-server-name</i>]
Display the TACACS+ stop-accounting requests saved in buffer without response	display stop-accounting-buffer TACACS+-scheme <i>TACACS+-scheme-name</i>
Delete the stop-accounting requests saved in buffer without response	reset stop-accounting-buffer { radius-scheme <i>radius-scheme-name</i> session-id <i>session-id</i> time-range <i>start-time stop-time</i> user-name <i>user-name</i> }
Delete the TACACS+ stop-accounting requests saved in buffer without response	reset stop-accounting-buffer TACACS+-scheme <i>TACACS+-scheme-name</i>
Reset the statistics of TACACS+ server	reset TACACS+ statistics { accounting authentication authorization all }
Enable RADIUS packet debugging	debugging radius packet
Disable RADIUS packet debugging	undo debugging radius packet

Operation	Command
Enable debugging of local RADIUS authentication server	debugging local-server { all error event packet }
Disable debugging of local RADIUS authentication server	undo debugging local-server { all error event packet }
Enable TACACS+ debugging	debugging TACACS+ { all error event message receive-packet send-packet }
Disable TACACS+ debugging	undo debugging TACACS+ { all error event message receive-packet send-packet }

40.6 AAA and RADIUS/TACACS+ Protocol Configuration Examples

For the hybrid configuration example of AAA/RADIUS/TACACS+ protocol and 802.1x protocol, refer to section 39.4 “802.1x Configuration Example”. It will not be detailed here.

40.6.1 Configuring Authentication at Remote RADIUS Server

Note:

Configuring Telnet user authentication at the remote server is similar to configuring FTP users. The following description is based on Telnet users.

I. Network Requirements

In the environment as illustrated in the following figure, it is required to achieve through proper configuration that the RADIUS server authenticates the Telnet users to be registered.

One RADIUS server (as authentication server) is connected to the switch and the server IP address is 10.110.91.146. The password for exchanging messages between the switch and the authentication server is "expert". The switch cuts off domain name from username and sends the left part to the RADIUS server.

II. Network Topology

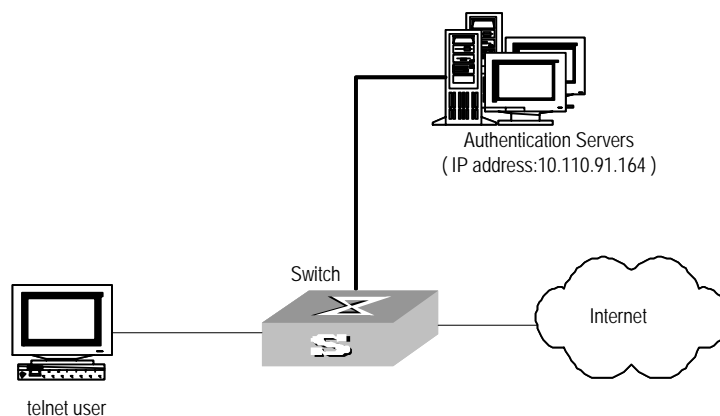


Figure 40-4 Network diagram for the remote RADIUS authentication of Telnet users

III. Configuration procedure

Add a Telnet user.

Omitted

Note:

For details about configuring FTP and Telnet users, refer to User Interface Configuration of *Getting Started Operation* in this manual.

Configure remote authentication mode for the Telnet user, i.e. scheme mode.

```
[SW8800-ui-vty0-4] authentication-mode scheme
```

Configure domain.

```
[SW8800] domain cams
[SW8800-isp-cams] quit
```

Configure RADIUS scheme.

```
[SW8800] radius scheme cams
[SW8800-radius-cams] primary authentication 10.110.91.146 1812
[SW8800-radius-cams] key authentication expert
[SW8800-radius-cams] server-type 3Com
[SW8800-radius-cams] user-name-format without-domain
```

Associate the domain with RADIUS.

```
[SW8800-radius-cams] quit
[SW8800] domain cams
[SW8800-isp-cams] radius-scheme cams
```

40.6.2 Configuring Authentication at Local RADIUS Authentication Server

Local RADIUS authentication of Telnet/FTP users is similar to the remote RADIUS authentication described in section 40.6.1 . But you should modify the server IP address in Figure 40-4 of section 40.6.1 to 127.0.0.1, authentication password to 3Com, the UDP port number of the authentication server to 1645.

Note:

For details about local RADIUS authentication of Telnet/FTP users, refer to “40.3.15 Creating/Deleting a Local RADIUS authentication Server”.

40.6.3 Configuring Authentication at Remote TACACS Server

I. Network requirements

Configure the switch to use a TACACS server to provide authentication and authorization services to login users (see the following figure).

Connect the switch to one TACACS server (which acting as a AAA server) with the IP address 10.110.91.164. On the switch, set the shared key for AAA packet encryption to “expert”. Configure the switch to send usernames to the TACACS server with *isp-name* removed.

On the TACACS server, set the shared key for encrypting the packets exchanged with the switch to “expert” .

II. Network diagram

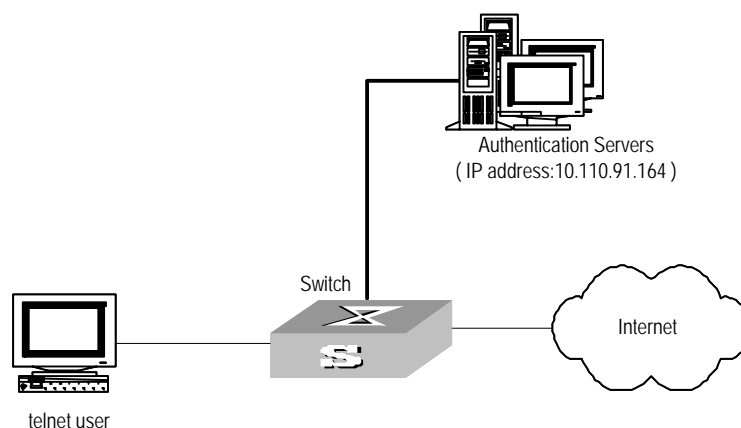


Figure 40-5 Network diagram for TACACS authentication

III. Configuration procedure

Configure the Telnet user.

Here it is omitted.

Note:

The configuration of the FTP and Telnet users can refer to User Interface Configuration of *Getting Started Operation* section of this manual.

Configure a TACACS+ scheme.

```
[SW8800] TACACS+ scheme hwtac
[SW8800-TACACS+-hwtac] primary authentication 10.110.91.164
[SW8800-TACACS+-hwtac] primary authorization 10.110.91.164
[SW8800-TACACS+-hwtac] key authentication expert
[SW8800-TACACS+-hwtac] key authorization expert
[SW8800-TACACS+-hwtac] user-name-format without-domain
[SW8800-TACACS+-hwtac] quit
```

Associate the domain with the TACACS+ scheme.

```
[SW8800] domain TACACS+
[SW8800-isp-TACACS+] scheme TACACS+-scheme hwtac
```

40.7 Troubleshooting AAA and RADIUS/TACACS+

RADIUS/TACACS+ protocol is located on the application layer of TCP/IP protocol suite. It mainly specifies how to exchange user information between NAS and RADIUS/TACACS+ server of ISP. So it is very likely to be invalid.

I. Symptom: User authentication/authorization always fails

Solution:

- The username may not be in the `userid@isp-name` format or NAS has not been configured with a default ISP domain. Please use the username in proper format and configure the default ISP domain on NAS.
- The user may have not been configured in the RADIUS/TACACS+ server database. Check the database and make sure that the configuration information of the user does exist in the database.
- The user may have input a wrong password. So please make sure that the applicant inputs the correct password.
- The encryption keys of RADIUS/TACACS+ server and NAS may be different. Please check carefully and make sure that they are identical.

- There might be some communication fault between NAS and RADIUS/TACACS+ server, which can be discovered through pinging RADIUS/TACACS+ server from NAS. So please ensure the normal communication between NAS and RADIUS/TACACS+ server.

II. Symptom: RADIUS/TACACS+ packet cannot be transmitted to RADIUS/TACACS+ server.

Solution:

- The communication lines (on physical layer or link layer) connecting NAS and RADIUS/TACACS+ server may not work well. So please ensure the lines work well.
- The IP address of the corresponding RADIUS/TACACS+ server may not have been set on NAS. Please set a proper IP address for RADIUS/TACACS+ server.
- UDP ports of authentication/authorization and accounting services may not be set properly. So make sure they are consistent with the ports provided by RADIUS/TACACS+ server.

III. Symptom: After being authenticated and authorized, the user cannot send charging bill to the RADIUS/TACACS+ server.

Solution:

- The accounting port number may be set improperly. Please set a proper number.
- The accounting service and authentication/authorization service are provided on different servers, but NAS requires the services to be provided on one server (by specifying the same IP address). So please make sure the settings of servers are consistent with the actual conditions.

Chapter 41 VRRP Configuration

41.1 Introduction to VRRP

Virtual Router Redundancy Protocol (VRRP) is a fault-tolerant protocol. In general, a default route (for example, 10.100.10.1 as shown in the following internetworking diagram) will be configured for every host on a network, so that the packets destined to some other network segment from the host will go through the default route to the Layer 3 Switch, implementing communication between the host and the external network. If Switch is down, all the hosts on this segment taking Switch as the next-hop on the default route will be disconnected from the external network.

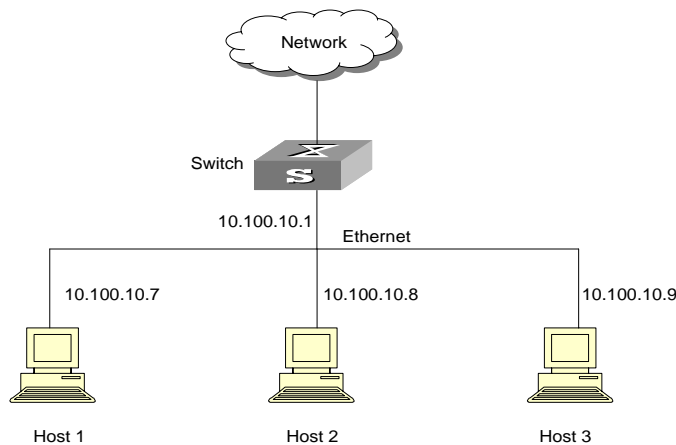


Figure 41-1 Network diagram for LAN

VRRP, designed for LANs with multicast and broadcast capabilities (such as Ethernet) settles the above problem. The diagram below is taken as an example to explain the implementation principal of VRRP. VRRP combines a group of LAN switches (including a Master and several Backups) into a virtual router.

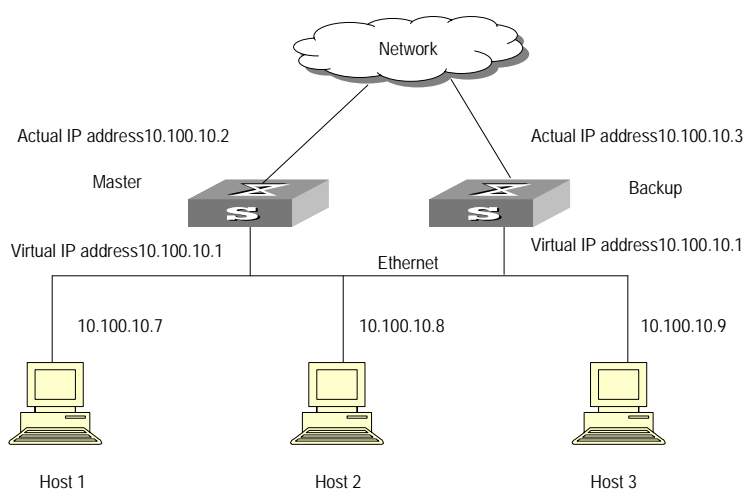


Figure 41-2 Network diagram for virtual router

This virtual router has its own IP address: 10.100.10.1 (which can be the interface address of a switch within the virtual router). The switches within the virtual router have their own IP addresses (such as 10.100.10.2 for the Master switch and 10.100.10.3 for the Backup switch). The host on the LAN only knows the IP address of this virtual router 10.100.10.1 (usually called as virtual IP address of virtual router), but not the specific IP addresses 10.100.10.2 of the Master switch and 10.100.10.3 of the Backup switch. They configure their own default routes as the IP address of this virtual router: 10.100.10.1. Therefore, hosts within the network will communicate with the external network through this virtual router. If a Master switch in the virtual group breaks down, another Backup switch will function as the new Master switch to continue serving the host with routing to avoid interrupting the communication between the host and the external networks.

41.2 Configuring VRRP

The following sections describe the VRRP configuration tasks:

- Enabling/Disabling the Function to Ping the Virtual IP Address
- Enabling/Disabling the Check of TTL Value of VRRP Packet
- Setting Correspondence Between Virtual IP Address and MAC Address
- Adding/Deleting a Virtual IP Address
- Configuring the Priority of Switches in the Virtual Router
- Configuring Preemption and Delay for a Switch Within a Virtual Router
- Configuring Authentication Type and Authentication Key
- Configuring Virtual Router Timer
- Configuring Switch to Track a Specified Interface

41.2.1 Enabling/Disabling the Function to Ping the Virtual IP Address

This operation enables or disables the function to ping the virtual IP address of the virtual router. The standard protocol of VRRP does not support the ping function, then the user cannot judge with **ping** command whether an IP address is used by the virtual router. If the user configure the IP address for the host same as the virtual IP address of the virtual router, then all messages in this segment will be forwarded to the host.

So the Switch 8800 provides the ping function to ping the virtual IP address of the virtual router.

Perform the following configuration in system view.

Table 41-1 Enable/disable the ping function

Operation	Command
Enable to ping the virtual IP address	vrrp ping-enable
Disable to ping the virtual IP address	undo vrrp ping-enable

By default, the function to ping the virtual IP address is disabled.

You should set the ping function before configuring the virtual router. If a virtual router is already established on the switch, you cannot perform this configuration any more.

41.2.2 Enabling/Disabling the Check of TTL Value of VRRP Packet

This operation configures whether to check TTL value of VRRP packet on the switch. The TTL value must be 225. If the switch find TTL is not 225 when receiving VRRP packet, the packet will be discarded.

Perform the following configuration in VLAN interface view.

Table 41-2 Enable/disable the check of TTL value of VRRP packet

Operation	Command
Disable the check of TTL value of VRRP packet	vrrp un-check ttl
Enable the check of TTL value of VRRP packet	undo vrrp un-check ttl

By default, the switch checks TTL value of VRRP packets.

41.2.3 Setting Correspondence Between Virtual IP Address and MAC Address

This operation sets correspondence between the virtual IP address and the MAC address. In the standard protocol of VRRP, the virtual IP address of the virtual router corresponds to the virtual MAC address, to ensure correct data forwarding in the sub-net.

Due to the chips installed, some switches support matching one IP address to multiple MAC addresses.

The Switch 8800 not only guarantees correct data forwarding in the sub-net, but also support such function: the user can choose to match the virtual IP address with the real MAC address or virtual MAC address of the routing interface.

The following commands can be used to set correspondence between the IP address and the MAC address.

Perform the following configuration in system view.

Table 41-3 Set correspondence between virtual IP address and MAC address

Operation	Command
Set correspondence between the virtual IP address and the MAC address	vrrp method { real-mac virtual-mac }
Set the correspondence to the default value	undo vrrp method

By default, the virtual IP address of the virtual router corresponds to the virtual MAC address.

You should set correspondence between the virtual IP address of the virtual router and the MAC address before configuring the virtual router. Otherwise, you cannot configure the correspondence.

If you set correspondence between the IP address of the virtual router and the real MAC address, you can configure only one virtual router on VLAN interface.

41.2.4 Adding/Deleting a Virtual IP Address

The following command is used for assigning a virtual IP address of the local segment to a virtual router or removing an assigned virtual IP address of a virtual router from the virtual address list.

Perform the following configuration in VLAN interface view.

Table 41-4 Add/delete a virtual IP address

Operation	Command
Add a virtual IP address	vrrp vrid <i>virtual-router-ID</i> virtual-ip <i>virtual-address</i>
Delete a virtual IP address	undo vrrp vrid <i>virtual-router-ID</i> [virtual-ip <i>virtual-address</i>]

The *virtual-router-ID* covers the range from 1 to 255.

The *virtual-address* can be an unused address in the network segment where the virtual router resides, or the IP address of an interface in the virtual router. If the IP address is of the switch in the virtual router, it can also be configured as *virtual-address*. In this case, the switch will be called an IP Address Owner. When adding the first IP address to a virtual router, the system will create a new virtual router accordingly. When adding a new address to this virtual router thereafter, the system will directly add it into the virtual IP address list.

After the last virtual IP address is removed from the virtual router, the whole virtual router will also be removed. That is, there is no more virtual router on the interface any more and any configuration of it is invalid accordingly.

41.2.5 Configuring the Priority of Switches in the Virtual Router

The status of each switch in the virtual router will be determined by its priority in VRRP. The switch with the highest priority will become the Master.

Perform the following configuration in VLAN interface view.

Table 41-5 Configure the priority of switches in the virtual router.

Operation	Command
Configure the priority of switches in the virtual router.	vrrp vrid <i>virtual-router-ID</i> priority
Clear the priority of switches in the virtual router.	undo vrrp vrid <i>virtual-router-ID</i> priority

The priority ranges from 0 to 255. The greater the number, the higher the priority. However the value can only be taken from 1 to 254. The priority 0 is reserved for special use and 255 is reserved for the IP address owner by the system.

By default, the priority is 100.

Note:

The priority for IP address owner is always 255, which cannot be configured otherwise.

41.2.6 Configuring Preemption and Delay for a Switch Within a Virtual Router

Once a switch in the virtual router becomes the Master switch, so long as it still functions properly, other switches, even configured with a higher priority later, cannot become the Master switch unless they are configured to work in preemption mode. The switch in preemption mode will become the Master switch, when it finds its own priority

is higher than that of the current Master switch. Accordingly, the former Master switch will become the Backup switch.

Together with preemption settings, a delay can also be set. In this way, a Backup will wait for a period of time before becoming a Master. In an unstable network if the Backup switch has not received the packets from the Master switch punctually, it will become the Master switch. However, the failure of Backup to receive the packets may be due to network congestion, instead of the malfunction of the Master switch. In this case, the Backup will receive the packet after a while. The delay settings can thereby avoid the frequent status changing.

Perform the following configuration in VLAN interface view.

Table 41-6 Configure preemption and delay for a switch within a virtual router

Operation	Command
Enable the preemption mode and configure a period of delay.	<code>vrrp vrid virtual-router-ID preempt-mode [timer delay delay-value]</code>
Disable the preemption mode.	<code>undo vrrp vrid virtual-router-ID preempt-mode</code>

The delay ranges from 0 to 255, measured in seconds. By default, the preemption mode is preemption with a delay of 0 second.

Note:

If preemption mode is cancelled, the delay time will automatically become 0 second.

41.2.7 Configuring Authentication Type and Authentication Key

VRRP provides following authentication types:

- **simple**: Simple character authentication
- **md5**: MD5 authentication

In a network under possible security threat, the authentication type can be set to **simple**. Then the switch will add the authentication key into the VRRP packets before transmitting it. The receiver will compare the authentication key of the packet with the locally configured one. If they are the same, the packet will be taken as a true and legal one. Otherwise it will be regarded as an illegal packet to be discarded. In this case, an authentication key not exceeding 8 characters should be configured.

In a totally unsafe network, the authentication type can be set to **md5**. The switch will use the authentication type and MD5 algorithm provided by the Authentication Header

to authenticate the VRRP packets. In this case an authentication key not exceeding 8 characters should be configured.

Those packets failing to pass the authentication will be discarded and a trap packet will be sent to the network management system.

Perform the following configuration in VLAN interface view.

Table 41-7 Configure authentication type and authentication key

Operation	Command
Configure authentication type and authentication key	vrrp authentication-mode authentication-type authentication-key
Remove authentication type and authentication key	undo vrrp authentication-mode

The authentication key is case sensitive.

Note:

The same authentication type and authentication key should be configured for all VLAN interfaces that belong to the virtual router.

41.2.8 Configuring Virtual Router Timer

The Master switch advertises its normal operation state to the switches within the VRRP virtual router by sending them VRRP packets regularly (at *adver-interval*). And the backup switch only receives VRRP packets. If the Backup has not received any VRRP packet from the Master after a period of time (specified by *master-down-interval*), it will consider the Master as down, and then take its place and become the Master.

You can use the following command to set a timer and adjust the interval, *adver-interval*, between Master transmits VRRP packets. The *master-down-interval* of the Backup switch is three times that of the *adver-interval*. The excessive network traffic or the differences between different switch timers will result in *master-down-interval* timing out and state changing abnormally. Such problems can be solved through prolonging the *adver-interval* and setting delay time. *adver-interval* is measured in seconds.

Perform the following configuration in VLAN interface view.

Table 41-8 Configure virtual router timer

Operation	Command
Configure virtual router timer	vrpp vrid <i>virtual-router-ID</i> timer advertise <i>adver-interval</i>
Clear virtual router timer	undo vrpp vrid <i>virtual-router-ID</i> timer advertise

By default, *adver-interval* is configured to be 1.

41.2.9 Configuring Switch to Track a Specified Interface

VRRP interface track function has expanded the backup function. Backup is provided not only to the interface where the virtual router resides, but also to some other malfunctioning switch interface. By implementing the following command you can track some interface.

If the interface which is tracked is DOWN, the priority of the switch including the interface will reduce automatically by the value specified by *value-reduced*, thus resulting in comparatively higher priorities of other switches within the virtual router, one of which will turn to Master switch so as to track this interface.

Perform the following configuration in VLAN interface view.

Table 41-9 Configure switch to track a specified interface

Operation	Command
Configure the switch to track a specified interface	vrpp vrid <i>virtual-router-ID</i> track vlan-interface <i>interface-num</i> [reduced <i>value-reduced</i>]
Stop tracking the specified interface	undo vrpp vrid <i>virtual-router-ID</i> track [vlan-interface <i>interface-num</i>]

By default, *value-reduced* is taken 10.

Note:

When the switch is an IP address owner, its interfaces cannot be tracked.

If the interface tracked is up again, the corresponding priority of the switch, including the interface, will be restored automatically

You can only track up to eight interfaces in one virtual router.

41.3 Displaying and debugging VRRP

After the above configuration, execute **display** command in any view to display the running of the VRRP configuration, and to verify the configuration. Execute **debugging** command in user view to debug VRRP configuration.

Table 41-10 Display and debug VRRP

Operation	Command
Display VRRP state information	display vrrp [interface vlan-interface <i>interface-num</i> [<i>virtual-router-ID</i>]]
Display VRRP statistics information	display vrrp statistics [vlan-interface <i>interface-num</i> [<i>virtual-router-ID</i>]]
Display VRRP summary information	display vrrp summary
Clear the statistics information about VRRP	reset vrrp statistics [vlan-interface <i>interface-num</i> [<i>virtual-router-ID</i>]]
Enable VRRP debugging.	debugging vrrp { state packet error }
Disable VRRP debugging.	undo debugging vrrp { state packet error }

You can enable VRRP debugging to check its running. You may choose to enable VRRP packet debugging (*option* as packet), VRRP state debugging (*option* as state), and/or VRRP error debugging (*option* as error). By default, VRRP debugging is disabled.

41.4 VRRP Configuration Example

41.4.1 VRRP Single Virtual Router Example

I. Networking requirements

Host A uses the VRRP virtual router which combines switch A and switch B as its default gateway to access host B on the Internet.

VRRP virtual router information includes: virtual router ID1, virtual IP address 202.38.160.111, switch A as the Master and switch B as the Backup allowed preemption.

II. Networking diagram

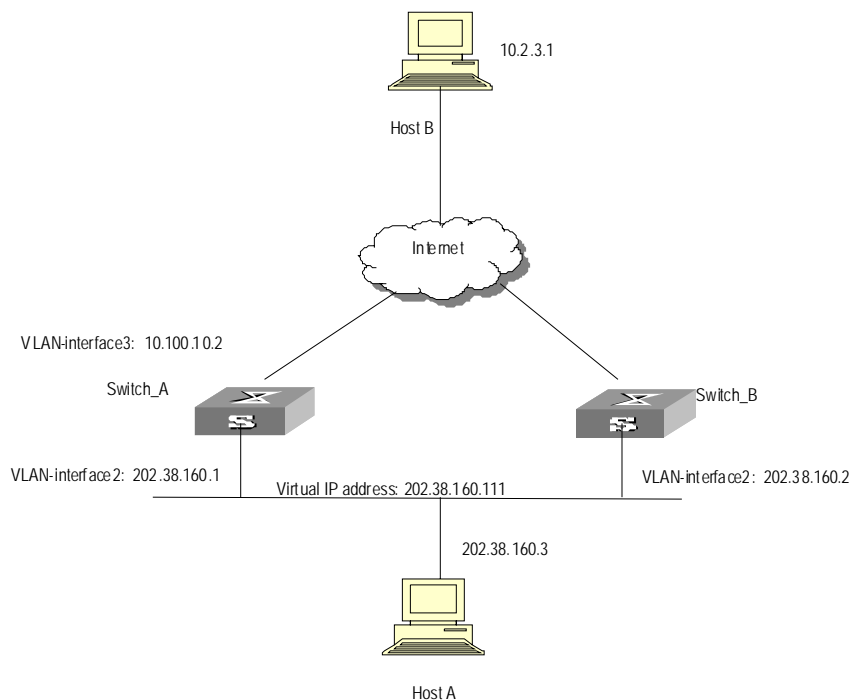


Figure 41-3 Network diagram for VRRP configuration

III. Configuration Procedure

Configure switch A

Configure VLAN 2.

```
[LSW-A] vlan 2
```

```
[LSW-A-vlan2] interface vlan 2
```

```
[LSW-A-vlan-interface2] ip address 202.38.160.1 255.255.255.0
```

```
[LSW-A-vlan-interface2] quit
```

Configure VRRP.

```
[LSW-A] vrrp ping-enable
```

```
[LSW-A] interface vlan 2
```

```
[LSW_A-vlan-interface2] vrrp vrid 1 virtual-ip 202.38.160.111
```

```
[LSW_A-vlan-interface2] vrrp vrid 1 priority 110
```

```
[LSW-A-vlan-interface2] vrrp vrid 1 preempt-mode
```

Configure switch B

Configure VLAN2.

```
[LSW-B] vlan 2
```

```
[LSW-B-vlan2] interface vlan 2
```



```
[LSW-B-vlan-interface2] ip address 202.38.160.2 255.255.255.0
```

```
[LSW-B-vlan-interface2] quit
```

Configure VRRP.

```
[LSW-B] vrrp ping-enable
```

```
[LSW-B] interface vlan 2
```

```
[LSW-B-vlan-interface2] vrrp vrid 1 virtual-ip 202.38.160.111
```

```
[LSW-B-vlan-interface2] vrrp vrid 1 preempt-mode
```

The virtual router can be used soon after configuration. Host A can configure the default gateway as 202.38.160.111.

Under normal conditions, switch A functions as the gateway, but when switch A is turned off or malfunctioning, switch B will function as the gateway instead.

Configure preemption mode for switch A, so that it can resume its gateway function as the Master after recovery.

41.4.2 VRRP Tracking Interface Example

I. Networking requirements

Even when switch A is still functioning, it may want switch B to function as gateway when the Internet interface connected with it does not function properly. This can be implemented by configuration of tracking interface.

In simple language, the virtual router ID is set as 1 with additional configurations of authorization key and timer.

II. Networking diagram

See Figure 41-3.

III. Configuration Procedure

Configure switch A

Configure VLAN2.

```
[LSW-A] vlan 2
```

```
[LSW-A-vlan2] interface vlan 2
```

```
[LSW-A-vlan-interface2] ip address 202.38.160.1 255.255.255.0
```

```
[LSW-A-vlan-interface2] quit
```

Enable the function to ping the virtual IP address of virtual router.

```
[SW8800LSW-A ] vrrp ping-enable
```

Create a virtual router.

```
[LSW-A] interface vlan 2
```

```
[LSW_A-vlan-interface2] vrrp vrid 1 virtual-ip 202.38.160.111
```

Set the priority for the virtual router.

```
[LSW_A-vlan-interface2] vrrp vrid 1 priority 110
```

Set the authentication key for the virtual router.

```
[LSW_A-vlan-interface2] vrrp authentication-mode md5 switch
```

Set Master to send VRRP packets every 5 seconds.

```
[LSW_A-vlan-interface2] vrrp vrid 1 timer advertise 5
```

Track an interface.

```
[LSW_A-vlan-interface2] vrrp vrid 1 track vlan-interface 3 reduced 30
```

Configure switch B

Configure VLAN2.

```
[LSW-B] vlan 2
```

```
[LSW-B-vlan2] interface vlan 2
```

```
[LSW-B-vlan-interface2] ip address 202.38.160.2 255.255.255.0
```

```
[LSW-B-vlan-interface2] quit
```

Enable the function to ping the virtual IP address of virtual router.

```
[SW8800LSW-B] vrrp ping-enable
```

Create a virtual router.

```
[LSW-B] interface vlan 2
```

```
[LSW_B-vlan-interface2] vrrp vrid 1 virtual-ip 202.38.160.111
```

Set the authentication key for the virtual router.

```
[LSW_B-vlan-interface2] vrrp authentication-mode md5 switch
```

Set Master to send VRRP packets every 5 seconds.

```
[LSW_B-vlan-interface2] vrrp vrid 1 timer advertise 5
```

Under normal conditions, switch A functions as the gateway, but when the interface vlan-interface 3 of switch A is down, its priority will be reduced by 30, lower than that of switch B so that switch B will preempt the Master for gateway services instead.

When vlan-interface3, the interface of switch A, recovers, this switch will resume its gateway function as the Master.

41.4.3 Multiple Virtual Routers Example

I. Networking requirements

A Switch can function as the backup switch for many virtual routers.

Such a multi-backup configuration can implement load balancing. For example, switch A as the Master switch of virtual router 1 can share the responsibility of the backup

switch for virtual router 2 and vice versa for switch B. Some hosts employ virtual router 1 as the gateway, while others employ virtual router 2 as the gateway. In this way, both load balancing and mutual backup are implemented.

II. Networking diagram

See Figure 41-3.

III. Configuration Procedure

Configure switch A

Configure VLAN2.

```
[LSW-A] vlan 2
[LSW-A-vlan2] interface vlan 2
[LSW-A-vlan-interface2] ip address 202.38.160.1 255.255.255.0
```

Create virtual router 1.

```
[LSW_A-vlan-interface2] vrrp vrid 1 virtual-ip 202.38.160.111
```

Set the priority for the virtual router.

```
[LSW_A-vlan-interface2] vrrp vrid 1 priority 150
```

Create virtual router 2.

```
[LSW_A-vlan-interface2] vrrp vrid 2 virtual-ip 202.38.160.112
```

Configure switch B

Configure VLAN2.

```
[LSW-B] vlan 2
[LSW-B-vlan2] interface vlan 2
[LSW-B-vlan-interface2] ip address 202.38.160.2 255.255.255.0
```

Create virtual router 1.

```
[LSW_B-vlan-interface2] vrrp vrid 1 virtual-ip 202.38.160.111
```

Create virtual router 2.

```
[LSW_B-vlan-interface2] vrrp vrid 2 virtual-ip 202.38.160.112
```

Set the priority for the virtual router.

```
[LSW_B-vlan-interface2] vrrp vrid 2 priority 110
```

Note:

Multiple virtual routers are often used in actual network applications.

41.5 Troubleshooting VRRP

As the configuration of VRRP is not very complicated, almost all the malfunctions can be found through viewing the configuration and debugging information. Here are some possible failures you might meet and the corresponding troubleshooting methods.

I. Fault 1: Frequent prompts of configuration errors on the console

This indicates that an incorrect VRRP packet has been received. It may be because of the inconsistent configuration of another switch within the virtual router, or the attempt of some devices to send out illegal VRRP packets. The first possible fault can be solved through modifying the configuration. And as the second possibility is caused by the malicious attempt of some devices, non-technical measures should be resorted to.

II. Fault 2: More than one Masters existing within the same virtual router

There are also 2 reasons. One is short time coexistence of many Master switches, which is normal and needs no manual intervention. Another is the long time coexistence of many Master switches, which may be because switches in the virtual router cannot receive VRRP packets from each other, or receive some illegal packets.

To solve such problems, an attempt should be made to ping among the many Master switches, and if such an attempt fails, check the device connectivity. If they can be pinged, check the VRRP configuration. For the configuration of the same VRRP virtual router, complete consistence for the number of virtual IP addresses, each virtual IP address, timer duration and authentication type must be guaranteed.

III. Fault 3: Frequent switchover of VRRP state

Such problem occurs when the virtual router timer duration is set too short. So the problem can be solved through prolonging this duration or configuring the preemption delay.

Chapter 42 HA Configuration

42.1 Introduction to HA

HA (high availability) is to achieve a high availability of the system and to recover the system as soon as possible in the event of Fabric failures so as to shorten the MTBF (Mean Time Between Failure) of the system.

The functions of HA are mainly implemented by the application running on the Fabric and slave board. The two boards are working in the master-slave mode: one board works in master mode, the other work in slave mode. If the master-slave system detects a fault in the Fabric, a hot master-slave switchover will be performed automatically. The slave board will try to connect and control the system bus while the original Fabric will try to disconnect from the bus. Thus, the master-slave switchover of the active system is completed, and at the same time the original Fabric is reset to recover as soon as possible and then function as the slave board. Even if the Fabric fails, the slave board can also take its role to ensure the normal operation, and the system can recover as soon as possible.

The Switch 8800 supports hot swap of Fabric and slave board. The hot swap of Fabricss will cause master-slave switchover.

The Switch 8800 supports manual master-slave switchover. You can change the current board state manually by executing command.

The configuration file of slave is copied from the Fabric at the same time. This can ensure that the slave system continues to operate in the same configuration as that of the original active system after the slave system has taken place of the active system. The Switch 8800 supports automatic synchronization. The active system stores its configuration file and backup the configuration file to the slave system simultaneously when the master's configuration file is modified, ensuring the consistency of the configurations of the active system and slave system.

Besides, the system can monitor the power supply and the working environment of the system and give timely alarms to avoid the escalation of failures and ensure safe operations of the system.

42.2 Configuring HA

The following sections describe the HA configuration tasks:

- Restarting the Slave System Manually
- Starting the Master-Slave Switchover Manually
- Enabling/Disabling Automatic Synchronization
- Synchronizing the Configuration File Manually

- Configuring the Load Mode of the Fabric and Slave Board

42.2.1 Restarting the Slave System Manually

In the environment in which the slave system is available, the user can restart the slave system manually.

Perform the following configuration in user view.

Table 42-1 Restart the slave system manually

Operation	Command
Restart the slave system manually	slave restart

42.2.2 Starting the Master-Slave Switchover Manually

In the environment in which the slave board is available and master in real-time backup state, the user can inform the slave board of a master-slave switchover by using a command if he expects the slave board to operate in place of the Fabric. After the switchover, the slave board will control the system and the original Fabric will be forced to reset.

Perform the following configuration in user view.

Table 42-2 Start the master-slave switchover manually

Operation	Command
Start the master-slave switchover manually	slave switchover

The switchover manually will be ineffective if user set the system forbid master-slave switchover manually.

42.2.3 Enabling/Disabling Automatic Synchronization

The Switch 8800 supports automatic synchronization. The active system stores its configuration file and backup the configuration file to the slave system simultaneously when the master's configuration file is modified, ensuring the consistency of the configurations of the active system and slave system.

You can enable/disable automatic synchronize of the Switch 8800.

Perform the following configuration in system view.

Table 42-3 Enable/Disable automatic synchronization

Operation	Command
Enable automatic synchronization	slave auto-update config
Disable automatic synchronization	undo slave auto-update config

By default, the automatic synchronization of system is enabled.

42.2.4 Synchronizing the Configuration File Manually

Although the system can perform the synchronization automatically, the synchronization can occur only when the Fabric saves its configuration file. If the user expects to determine the backup of the configuration file by himself, he can do it manually to backup the configuration file saved in the Fabric.

Perform the following configuration in user view.

Table 42-4 Synchronize the configuration file manually

Operation	Command
Synchronize the configuration file manually	slave update configuration

This operation can backup the configuration file to the slave board only if a slave system is available. The configuration file will be fully copied once at every time the operation is executed.

42.2.5 Configuring the Load Mode of the Fabric and Slave Board

The Switch 8800 supports two kinds of load modes (balance and single) between the Fabric and slave board. You can use the **xbar** command to configure XBar (cross bar) load mode.

Perform the following configuration in system view.

Table 42-5 Configure the XBar load mode

Operation	Command
Configure the load mode of the Fabric and slave board	xbar [load-balance load-single]

By default, the load mode of the Fabric and slave board is **load-single**.

**Caution:**

When a single Fabric is in position, the load-balance mode is not effective and the Fabric changes to the load-single mode automatically.

42.3 Displaying and Debugging HA Configuration

After the above configuration, execute **display** command in relevant view to display the running of the ACL configuration, and to verify the configuration. Execute **debugging** command in user view to enable HA module debugging function.

Perform the following configuration in relevant view.

Table 42-6 Display and debug HA configuration

Operation	Command
Display the status of the Fabric and slave board(any view)	display switchover state [<i>slot-id</i>]
Display the load mode of the Fabric and slave board(system view)	display xbar
Enable the debugging information output of the HA module(user view)	debugging ha { all event message state }
Disable the debugging information output of the HA module(user view)	undo debugging ha { all event message state }

Chapter 43 File System Management

43.1 File System Configuration

43.1.1 File System Overview

The Ethernet switch provides a file system module for user's efficient management over the storage devices such as flash memory. The file system offers file access and directory management, mainly including creating the file system, creating, deleting, modifying and renaming a file or a directory and opening a file.

By default, the file system needs user's confirmation before executing the commands, such as deleting or overwriting a file, which may make losses.

Based on the operated objects, the file system operation can be divided as follows. The following sections describe the file system configuration tasks.

- Directory Operation
- File Operation
- Storage Device Operation
- Note: The error message "% Device can't be found or file can't be found in the directory" can indicate that the CF card is not formatted.
- Setting the Prompt Mode of the File System

Note:

The Switch 8800 supports master board and slave board. The two boards both have file system. User can operate the file on the two boards. In the case user operate the file on slave board, the file directory or URL should be started with "slot[No.]#flash:/", the [No.] is the slave board number. For example, suppose slot 1 is slave board, "text.txt" file URL on slave board should be "slot1#flash:/text.txt".

Note:

The limitation on the names of directories and files on switch are as follows:

- The name of one direction or file can be up to 64 characters long.
 - The total number of characters including device, directory and file names can be up to 136 characters long.
-

43.1.2 Directory Operation

The file system can be used to create or delete a directory, display the current working directory, and display the information about the files or directories under a specified directory. You can use the following commands to perform directory operations.

Perform the following configuration in user view.

Table 43-1 Directory operation

Operation	Command
Create a directory	mkdir <i>directory</i>
Delete a directory	rmdir <i>directory</i>
Display the current working directory	pwd
Display the information about directories or files	dir [/ all] [<i>file-url</i>]
Change the current directory	cd <i>directory</i>

43.1.3 File Operation

The file system can be used to delete or undelete a file and permanently delete a file. Also, it can be used to display file contents, rename, copy and move a file and display the information about a specified file. You can use the following commands to perform file operations.

Perform the following configuration in user view.

Table 43-2 File operation

Operation	Command
Delete a file	delete [/ unreserved] <i>file-url</i>
Undelete a file	undelete <i>file-url</i>
Delete a file from the recycle bin permanently	reset recycle-bin [<i>file-url</i>]
View contents of a file	more <i>file-url</i>
Rename a file	rename <i>fileurl-source fileurl-dest</i>
Copy a file	copy <i>fileurl-source fileurl-dest</i>
Move a file	move <i>fileurl-source fileurl-dest</i>
Display the information about directories or files	dir [/ all] [<i>file-url</i>]
Execute the specified batch file (system view)	execute <i>filename</i>

**Caution:**

When you use the **delete** command without the **unreserved** option to delete a file, the file is in fact saved in the recycle bin and still occupies some of the storage space. So, the frequent uses of this command may results in insufficient storage space of the Ethernet switch; in this case, you should find out the unused files kept in the recycle bin and permanently delete them with the **reset recycle-bin** command to reclaim the storage space.

Note:

The directory and file names on the switch have the following limitation:

- The maximum length of a directory or file name is 64 characters.
- The maximum length of a full path name (containing the device name, directory name and file name) is 136 characters.
- The **move** command takes effect only when the source and destination files are in the same device.

43.1.4 Storage Device Operation

The file system can be used to format a specified memory device. You can use the following commands to format a specified memory device.

Switch supports compact flash (CF) card. After a CF card is inserted successfully, you can use such common commands as **dir**, **cd**, **copy**, **delete**, **move** to perform operations on the files in the card. You can also stop the CF card through a command before dismounting it.

Perform the following configuration in user view.

Table 43-3 Storage device operation

Operation	Command
Format the storage device	format <i>filesystem</i>
Restore the space of the storage device	fixdisk <i>device</i>
Delete the CF card	umount <i>device</i>

Note: The error message “% Device can’t be found or file can’t be found in the directory” can indicate that the CF card is not formatted.

43.1.5 Setting the Prompt Mode of the File System

The following command can be used for setting the prompt mode of the current file system.

Perform the following configuration in system view.

Table 43-4 File system operation

Operation	Command
Set the file system prompt mode.	file prompt { alert quiet }

43.2 Configuration File Management

43.2.1 Configuration File Management Overview

The management module of configuration file provides a user-friendly operation interface. It saves the configuration of the Ethernet switch in the text format of command line to record the whole configuration process. Thus you can view the configuration information conveniently.

The format of configuration file includes:

- It is saved in the command format.
- Only the non-default constants will be saved
- The organization of commands is based on command views. The commands in the same command mode are sorted in one section. The sections are separated with a blank line or a comment line (A comment line begins with exclamation mark "#").
- Generally, the sections in the file are arranged in the following order: system configuration, Ethernet port configuration, vlan interface configuration, routing protocol configuration and so on.
- It ends with "end".

The following sections describe configuration file management tasks.

- Displaying the Current-Configuration and Saved-Configuration of Ethernet Switch
- Modifying and Saving the Current-Configuration
- Erasing Configuration Files from Flash Memory
- Configuring the Name of the Configuration File Used for the Next Startup.

43.2.2 Displaying the Current-Configuration and Saved-Configuration of Ethernet Switch

After being powered on, the system will read the configuration files from Flash Memory for the initialization of the device. (Such configuration files are called saved-configuration files). If there is no configuration file in Flash Memory, the system

will begin the initialization with the default parameters. Relative to the saved-configuration, the configuration in effect during the operating process of the system is called current-configuration. You can use the following commands to display the current-configuration and saved-configuration information of the Ethernet switch.

Perform the following configuration in any view.

Table 43-5 Display the configurations of the Ethernet switch

Operation	Command
Display the saved-configuration information of the Ethernet switch	display saved-configuration
Display the current-configuration information of the Ethernet switch	display current-configuration [controller interface <i>interface-type</i> [<i>interface-number</i>] configuration [<i>configuration</i>]] [{ begin exclude include } <i>regular-expression</i>]
Display the running configuration of the current view	display this

Note:

The configuration files are displayed in their corresponding saving formats.

43.2.3 Modifying and Saving the Current-Configuration

You can modify the current configuration of Ethernet switch through the CLI. Use the **save** command to save the current-configuration in the Flash Memory, and the configurations will become the saved-configuration when the system is powered on for the next time.

Perform the following configuration in user view.

Table 43-6 Save the current-configuration

Operation	Command
Save the current-configuration	save [<i>file-name</i>]

Even if the problems like reboot and power-off occur during , the configuration file can be still saved to Flash.

43.2.4 Erasing Configuration Files from Flash Memory

The **reset saved-configuration** command can be used to erase configuration files from Flash Memory. The system will use the default configuration parameters for initialization when the Ethernet switch is powered on for the next time.

Perform the following configuration in user view.

Table 43-7 Erase configuration files from Flash Memory

Operation	Command
Erase configuration files from Flash Memory	reset saved-configuration

You may erase the configuration files from the Flash in the following cases:

- After being upgraded, the software does not match with the configuration files.
- The configuration files in flash are damaged. (A common case is that a wrong configuration file has been downloaded.)

43.2.5 Configuring the Name of the Configuration File Used for the Next Startup.

Perform the following configuration in user view.

Table 43-8 Configure the name of the configuration file used for the next startup

Operation	Command
Configure the name of the configuration file used for the next startup	startup saved-configuration <i>cfgfile</i>

cfgfile is the name of the configuration file and its extension name can be “.cfg”. The file is stored in the root directory of the storage devices.

After the above configuration, execute **display** command in any view to display the running of the configuration files, and to verify the effect of the configuration.

Table 43-9 Display the information of the file used at startup

Operation	Command
Display the information of the file used at startup	display startup

43.3 FTP Configuration

Note:

The system supports FTP services over VPN.

43.3.1 FTP Overview

FTP (File Transfer Protocol) is a universal method for transmitting files on the Internet and IP networks. In this method, files are copied from one system to another. FTP supports definite file types (such as ASCII and Binary) and file structures (byte stream and record). Even now, FTP is still used widely, while most users transmit files by Email and Web.

FTP, a TCP/IP protocol on the application layer, is used for transmitting files between a remote server and a local host.

The Ethernet switch provides the following FTP services:

- FTP server: You can run FTP client program to log in the server and access the files on it.
- FTP client: You can run the `ftp X.X.X.X` command (where, X.X.X.X represents the IP address of the remote FTP server) to set up a connection between the Ethernet switch and a remote FTP server to access the files on the remote server.

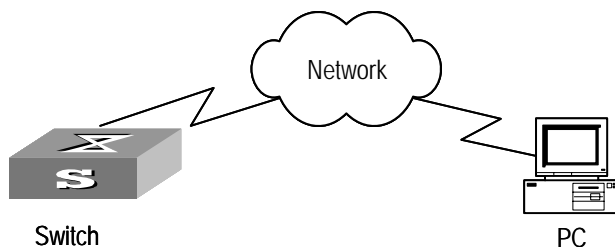


Figure 43-1 FTP configuration

Table 43-10 lists the configuration of the switch as FTP client.

Table 43-10 Configuration of the switch as FTP client

Device	Configuration	Default	Description
Switch	Log into the remote FTP server directly with the ftp command.	—	You need first get FTP user command and password, and then log into the remote FTP server. Then you can get the directory and file authority.
PC	Start FTP server and make such settings as username, password, and authority.	—	—

Table 43-11 lists the configuration of the switching as FTP server.

Table 43-11 Configuration of the switch as FTP server

Device	Configuration	Default	Description
Switch	Start FTP server.	FTP server is disabled.	You can view the configuration information of FTP server with the display ftp-server command.
	Configure authentication and authorization for FTP server.	—	Configure username, password and authorized directory for FTP users.
	Configure running parameters for FTP server.	—	Configure timeout time value for FTP server.
PC	Log into the switch from FTP client.	—	—



Caution:

The prerequisite for normal FTP function is that the switch and PC are reachable.

43.3.2 Enabling/Disabling FTP Server

You can use the following commands to enable/disable the FTP server on the switch. Perform the following configuration in system view.

Table 43-12 Enable/disable FTP Server

Operation	Command
Enable the FTP server	ftp server enable
Disable the FTP server	undo ftp server

FTP server supports multiple users to access at the same time. A remote FTP client sends request to the FTP server. Then, the FTP server will carry out the corresponding operation and return the result to the client.

By default, FTP server is disabled.

43.3.3 Configuring the FTP Server Authentication and Authorization

The authorization information of FTP server includes the path to the desired directory for FTP users. The FTP server service is available only for the authenticated and

authorized users. You can use the following commands to configure FTP server authentication and authorization. The authorization information of FTP server includes the top working directory provided for FTP clients.

Perform the following configuration in corresponding view.

Table 43-13 Configure the FTP Server Authentication and Authorization

Operation	Command
Create new local user and enter local user view(system view)	local-user { <i>username</i> multicast [domain <i>domain-name</i>] <i>ipaddress</i> password-display-mode { auto cipher-force } }
Delete local user(system view)	undo local-user { <i>username</i> all [service-type { ftp lan-access telnet ppp ssh terminal }] multicast [domain <i>domain-name</i>] <i>ipaddress</i> password-display-mode }
Configure password for local user(local user view)	password { cipher simple } <i>password</i>
Configure service type for local user(local user view)	service-type { ftp [ftp-directory <i>directory</i>] lan-access telnet [level <i>level</i>] }
Cancel password for local user(local user view)	undo password
Cancel service type for local user(local user view)	undo service-type { ftp [ftp-directory] lan-access telnet [level <i>level</i>] }

Only the clients who have passed the authentication and authorization successfully can access the FTP server.

43.3.4 Configuring the Running Parameters of FTP Server

You can use the following commands to configure the connection timeout of the FTP server. If the FTP server receives no service request from the FTP client for a period of time, it will cut the connection to it, thereby avoiding the illegal access from the unauthorized users. The period of time is FTP connection timeout.

Perform the following configuration in system view.

Table 43-14 Configuring FTP server connection timeout

Operation	Command
Configure FTP server connection timeouts	ftp timeout <i>minute</i>
Restoring the default FTP server connection timeouts	undo ftp timeout

By default, the FTP server connection timeout is 30 minutes.

43.3.5 Displaying and Debugging FTP Server

After the above configuration, execute **display** command in any view to display the running of the FTP Server configuration, and to verify the effect of the configuration.

Table 43-15 Display and debug FTP Server

Operation	Command
Display FTP server	display ftp-server
Display the connected FTP users.	display ftp-user

The **display ftp-server** command can be used for displaying the configuration information about the current FTP server, including the maximum amount of users supported by FTP server and the FTP connection timeout. The **display ftp-user** command can be used for displaying the detail information about the connected FTP users.

43.3.6 Disconnecting an FTP User

Perform the following configuration in system view.

Table 43-16 Disconnect an FTP user

Operation	Command
Disconnect an FTP user.	ftp disconnect <i>user-name</i>

43.3.7 Introduction to FTP Client

As an additional function provided by Ethernet switch, FTP client is an application module and has no configuration functions. The switch connects the FTP clients and the remote server and inputs the command from the clients for corresponding operations (such as creating or deleting a directory).

43.3.8 FTP Client Configuration Example

I. Network requirements

The switch serves as FTP client and the remote PC as FTP server. The configuration on FTP server: Configure an FTP user named as switch, with password hello and with read & write authority over the Switch root directory on the PC. The IP address of a VLAN interface on the switch is 1.1.1.1, and that of the PC is 2.2.2.2. The switch and PC are reachable.

The switch application `switch.app` is stored on the PC. Using FTP, the switch can download the `switch.app` from the remote FTP server and upload the `vrpcfg.cfg` to the FTP server under the switch directory for backup purpose.

II. Network diagram

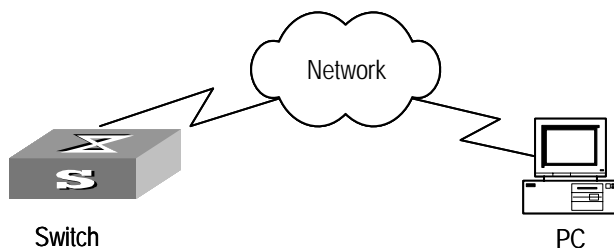


Figure 43-2 Network diagram for FTP configuration

III. Configuration procedure

- 1) Configure FTP server parameters on the PC: a user named as switch, password hello, read and write authority over the Switch directory on the PC.
- 2) Configure the switch

Log into the switch through the Console port locally or Telnet remotely.

Then type in the right command in user view to establish FTP connection, then correct username and password to log into the FTP server.

```
<SW8800> ftp 2.2.2.2
Trying ...
Press CTRL+K to abort
Connected.
220 WFTPD 2.0 service (by Texas Imperial Software) ready for new user
User(none):switch
331 Give me your password, please
Password:*****
230 Logged in successfully
[ftp]
```



Caution:

If the Flash Memory of the switch is not enough, you need to first delete the existing programs in the Flash Memory and then upload the new ones.

Enter the authorized directory of the FTP server.

```
[ftp] cd switch
```

Use the **put** command to upload the vrpcfg.cfg to the FTP server.

```
[ftp] put vrpcfg.cfg
```

Use the **get** command to download the switch.app from the FTP server to the Flash directory on the FTP server.

```
[ftp] get switch.app
```

Use the **quit** command to release FTP connection and return to user view.

```
[ftp] quit  
<SW8800>
```

Use the **boot boot-loader** command to specify the downloaded program as the application at the next login and reboot the switch.

```
<SW8800> boot boot-loader switch.app  
<SW8800> reboot
```

43.3.9 FTP Server Configuration Example

I. Network requirements

Switch serves as FTP server and the remote PC as FTP client. The configuration on FTP server: Configure an FTP user named as switch, with password hello and with read & write authority over the flash root directory on the PC. The IP address of a VLAN interface on the switch is 1.1.1.1, and that of the PC is 2.2.2.2. The switch and PC are reachable.

The switch application switch.app is stored on the PC. Using FTP, the PC can upload the switch.app from the remote FTP server and download the vrpcfg.cfg from the FTP server for backup purpose.

II. Network diagram

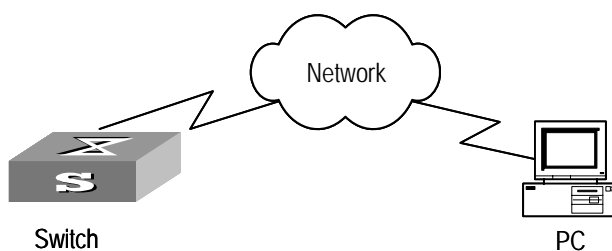


Figure 43-3 Network diagram for FTP configuration

III. Configuration procedure

- 1) Configure the switch

Log into the switch through the console port locally or Telnet remotely, and start FTP function and set username, password and file directory.

```
[SW8800] ftp server enable
[SW8800] local-user switch
[SW8800-luser-switch] service-type ftp ftp-directory flash:
[SW8800-luser-switch] password simple hello
```

- 2) Run FTP client on the PC and establish FTP connection. Upload the switch.app to the switch under the Flash directory and download the vrpcfg.cfg from the switch. FTP client is not shipped with the switch, so you need to buy it separately.

**Caution:**

If the Flash Memory of the switch is not enough, you need to first delete the existing programs in the Flash Memory and then upload the new ones.

- 3) When the uploading is completed, initiate file upgrade on the switch.

Use the **boot boot-loader** command to specify the downloaded program as the application at the next login and reboot the switch.

```
<SW8800> boot boot-loader switch.app
<SW8800> reboot
```

43.4 TFTP Configuration

43.4.1 TFTP Overview

Trivial File Transfer Protocol (TFTP) is a simple file transmission protocol. It is initially designed for the booting of free-disk systems (work stations or X terminals in general). Compared with FTP, another file transmission protocol, TFTP has no complicated interactive access interface or authentication control, and therefore it can be used when there is no complicated interaction between the clients and server. TFTP is implemented on the basis of UDP.

TFTP transmission is originated from the client end. To download a file, the client sends a request to the TFTP server and then receives data from it and sends acknowledgement to it. To upload a file, the client sends a request to the TFTP server and then transmits data to it and receives the acknowledgement from it. TFTP transmits files in two modes, binary mode for program files and ASCII mode for text files.

The administrator needs to configure the IP addresses of TFTP client and server before configuring TFTP, and makes sure that the route between the client and server is reachable.

The switch can only function as a TFTP client.

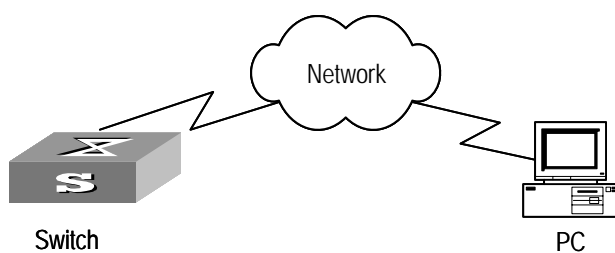


Figure 43-4 TFTP configuration

Table 43-17 lists the configuration of the switch as TFTP client.

Table 43-17 Configuration of the switch as TFTP client

Device	Configuration	Default	Description
Switch	Configure IP address for the VLAN interface of the switch, in the same network segment as that of TFTP server.	—	TFTP is right for the case where no complicated interactions are required between the client and server. Make sure that the route is reachable between the switch and the TFTP server.
	Use the ftfp command to log into the remote TFTP server for file uploading and downloading.	—	—
PC	Start TFTP server and set authorized TFTP directory.	—	—

43.4.2 Downloading Files by Means of TFTP

To download a file, the client sends a request to the TFTP server and then receives data from it and sends acknowledgement to it. You can use the following commands to download files by means of TFTP.

Perform the following configuration in user view.

Table 43-18 Download files by means of TFTP

Operation	Command
Download files by means of TFTP	ftfp <i>ftfp-server</i> get <i>source-file</i> [<i>dest-file</i>]

In the command, *ftfp-server* indicates the IP address or host name of TFTP server; *source-file* indicates the file information to be downloaded from TFTP server; *dest-file* indicates the name of the file downloaded on switch.

43.4.3 Uploading Files by Means of TFTP

To upload a file, the client sends a request to the TFTP server and then transmits data to it and receives the acknowledgement from it. You can use the following commands to upload files.

Perform the following configuration in user view.

Table 43-19 Upload files by means of TFTP

Operation	Command
Upload files by means of TFTP	<code>tftp tftp-server put source-file [dest-file]</code>

In the command, *source-file* indicates the file to be uploaded to server; *dest-file* indicates the saving directory on TFTP server; *tftp-server* indicates the IP address or host name of TFTP server.

43.4.4 TFTP Client Configuration Example

I. Network requirements

The switch serves as TFTP client and the remote PC as TFTP server. Authorized TFTP directory is set on the TFTP server. The IP address of a VLAN interface on the switch is 1.1.1.1, and that of the PC is 1.1.1.2.

The switch application switch.app is stored on the PC. Using TFTP, the switch can download the switch.app from the remote TFTP server and upload the vrpcfg.cfg to the TFTP server under the switch directory for backup purpose.

II. Network diagram

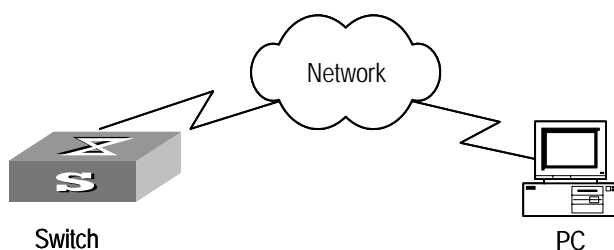


Figure 43-5 Network diagram for TFTP configuration

III. Configuration procedure

- 1) Start TFTP server on the PC and set authorized TFTP directory.
- 2) Configure the switch

Log into the switch (through local console or remote Telnet, refer to the Getting Started for login information), and then enter the system view.

```
<SW8800> system-view
[SW8800]
```

**Caution:**

If the Flash Memory of the switch is not enough, you need to first delete the existing programs in the Flash Memory and then upload the new ones.

Configure IP address 1.1.1.1 for the VLAN interface, ensure the port connecting the PC is also in this VLAN (VLAN 1 in this example).

```
[SW8800] interface vlan 1
[SW8800-vlan-interface1] ip address 1.1.1.1 255.255.255.0
[SW8800-vlan-interface1] quit
```

Enter system view and download the switch.app from the TFTP server to the Flash Memory of the switch.

```
<SW8800> tftp 1.1.1.2 get switch.app switch.app
```

Upload the vrpcfg.cfg to the TFTP server.

```
<SW8800> tftp 1.1.1.2 put vrpcfg.cfg vrpcfg.cfg
```

Use the **boot boot-loader** command to specify the downloaded program as the application at the next login and reboot the switch.

```
<SW8800> boot boot-loader switch.app
<SW8800> reboot
```


Chapter 44 MAC Address Table Management

44.1 MAC Address Table Management Overview

An Ethernet Switch maintains a MAC address table for fast forwarding packets. A table entry includes the MAC address of a device and the port ID of the Ethernet switch connected to it. The dynamic entries (not configured manually) are learned by the Ethernet switch. The Ethernet switch learns a MAC address in the following way: after receiving a data frame from a port (assumed as port A), the switch analyzes its source MAC address (assumed as MAC_SOURCE) and considers that the packets destined at MAC_SOURCE can be forwarded through the port A. If the MAC address table contains the MAC_SOURCE, the switch will update the corresponding entry; otherwise, it will add the new MAC address (and the corresponding forwarding port) as a new entry to the table.

The system forwards the packets whose destination addresses can be found in the MAC address table directly through the hardware and broadcasts those packets whose addresses are not contained in the table. The network device will respond after receiving a broadcast packet and the response contains the MAC address of the device, which will then be learned and added into the MAC address table by the Ethernet switch. The consequent packets destined the same MAC address can be forwarded directly thereafter.

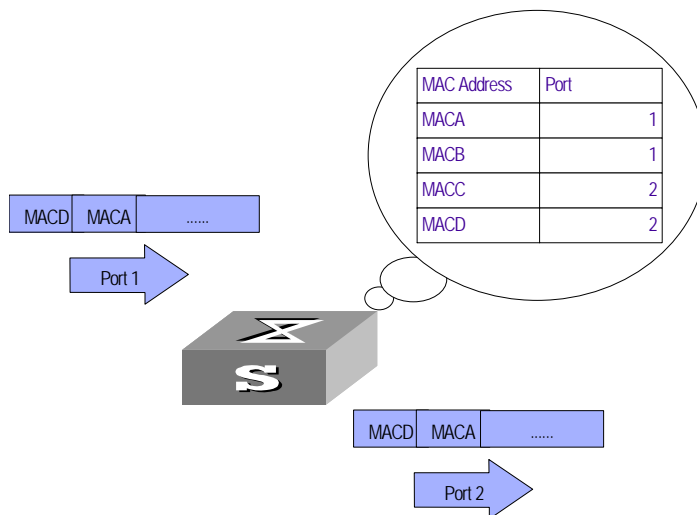


Figure 44-1 The Ethernet switch forwards packets with MAC address table

The Ethernet switch also provides the function of MAC address aging. If the switch receives no packet for a period of time, it will delete the related entry from the MAC address table. However, this function takes no effect on the static MAC addresses.

You can configure (add or modify) the MAC address entries manually according to the actual networking environment. The entries can be static ones or dynamic ones.

44.2 MAC Address Table Management Configuration

The following sections describe the MAC address table management configuration tasks.

- Setting MAC Address Table Entries
- Setting MAC Address Aging Time
- Maximum MAC Address Number Learned by Ethernet Port and Forwarding Option Configuration

44.2.1 Setting MAC Address Table Entries

Administrators can manually add, modify, or delete the entries in MAC address table according to the actual needs. They can also delete all the (unicast) MAC address table entries related to a specified port or delete a specified type of entries, such as dynamic entries or static entries.

You can use the following commands to add, modify, or delete the entries in MAC address table.

Perform the following configuration in system view.

Table 44-1 Set MAC address table entries

Operation	Command
Add/Modify an address entry	mac-address { static dynamic } <i>mac-addr</i> interface { <i>interface-name</i> <i>interface-type</i> <i>interface-num</i> } vlan <i>vlan-id</i>
Delete an address entry	undo mac-address [static dynamic] [<i>mac-addr</i> [interface { <i>interface-name</i> <i>interface-type</i> <i>interface-num</i> }] vlan <i>vlan-id</i> interface { <i>interface-name</i> <i>interface-type</i> <i>interface-num</i> } vlan <i>vlan-id</i>]

44.2.2 Setting MAC Address Aging Time

The setting of an appropriate aging time can effectively implement the function of MAC address aging. Too long or too short aging time set by subscribers will cause the problem that the Ethernet switch broadcasts a great amount of data packets without MAC addresses, which will affect the switch operation performance.

If aging time is set too long, the Ethernet switch will store a great number of out-of-date MAC address tables. This will consume MAC address table resources and the switch will not be able to update MAC address table according to the network change.

If aging time is set too short, the Ethernet switch may delete valid MAC address table. You can use the following commands to set the MAC address aging time for the system.

Perform the following configuration in system view.

Table 44-2 Set the MAC address aging time for the system

Operation	Command
Set the dynamic MAC address aging time	mac-address timer { aging age no-aging }
Restore the default MAC address aging time	undo mac-address timer aging

In addition, this command takes effect on all the ports. However the address aging only functions on the dynamic addresses (the learned or configured as age entries by the user).

By default, the *aging-time* is 300 seconds. With the **no-aging** parameter, the command performs no aging on the MAC address entries.



Caution:

The dynamic MAC address aging is completed during the second aging cycle.

44.3 Maximum MAC Address Number Learned by Ethernet Port and Forwarding Option Configuration

With MAC address learning, an Ethernet switch can obtain MAC addresses of every network devices on network segments connecting to a port. As for packets destined to those MAC addresses, the switch directly uses hardware to forward them. An overlarge MAC address table may cause the low forwarding performance of the switch.

You can control the number of entries of the MAC address table by setting the maximum number of MAC addresses learned by a port. if you set the value to *count*, and when the number of MAC addresses learned by the port reaches this value, this port will no longer learn any more MAC addresses.

You can also set the switch to drop corresponding packets when the number of MAC addresses learned by the port exceeds the configured threshold.

44.3.1 Maximum MAC Address Number Learned by a Port and Forwarding Option Configuration Tasks

Maximum MAC address number learned by a port and forwarding option configuration tasks are described in the following table:

Table 44-3 Configure the maximum number of MAC addresses learned by a port and forwarding option

Sequence number	Configuration item	Command	Description
1	Enter system view	<SW8800> system-view	—
2	Enter Ethernet port view	[SW8800] interface { <i>interface_type</i> <i>interface_num</i> <i>interface_name</i> }	The port can be an Ethernet port or a GigabitEthernet port. Ethernet port view prompt is related to the port you choose
3	Set the maximum number of MAC addresses learned by an Ethernet port	[SW8800-EthernetX/1/X] mac-address max-mac-count <i>count</i> or [SW8800-GigabitEthernetX/1/X] mac-address max-mac-count <i>count</i>	By default, the switch has no limit on the maximum number of MAC addresses learned by a port
4	Set the switch to drop the packets whose source MAC addresses are not learned by the port when the number of MAC addresses learned exceeds the threshold value	[SW8800-EthernetX/1/X] undo mac-address max-mac-count enable forward or [SW8800-GigabitEthernetX/1/X] undo mac-address max-mac-count enable forward	By default, the switch forwards packets whose source MAC addresses are not learned by the port when the number of MAC addresses learned exceeds the threshold value

Sequence number	Configuration item	Command	Description
5	Set the maximum number of MAC addresses learned by an Ethernet port, and when the current number of MAC addresses exceeds the threshold value, whether the switch forwards packets or gives the network administrator an alarm,	[SW8800-EthernetX/1/X] mac-address max-mac-count <i>count</i> or [SW8800-EthernetX/1/X] mac-address max-mac-count enable forward alarm	By default, the switch has no limit on the maximum number of MAC addresses learned by a port.

Use the corresponding undo command to cancel the configuration.

44.3.2 Configuring Maximum MAC Address Number Learned by Ethernet Port and Forwarding Option Example

I. Network requirements

- Set the maximum number of MAC addresses learned by Ethernet port Ethernet3/1/3 to 600
- Set the switch to drop the packets whose source MAC addresses are not learned by the port when the number of MAC addresses learned exceeds 600

II. Configuration procedure

1) Enter system view.

```
<SW8800> system-view
```

```
[SW8800]
```

2) Enter Ethernet port view.

```
[SW8800] interface ethernet 3/1/3
```

3) Set the maximum number of MAC addresses learned by Ethernet port Ethernet3/1/3 to 600.

```
[SW8800-Ethernet3/1/3] mac-address max-mac-count 600
```

4) Set the switch to drop the packets whose source MAC addresses are not learned by the port when the number of MAC addresses learned exceeds 600.

```
[SW8800-Ethernet3/1/3] undo mac-address max-mac-count enable forward
```

44.4 Displaying and Debugging MAC Address Tables

After the above configuration, execute the **display** command in any view to display the running of the MAC address table configuration, and to verify the effect of the configuration.

Execute the **debugging** command in user view to debug MAC address table configuration.

Table 44-4 Display and debug MAC address tables

Operation	Command
Display the information in the address table	display mac-address [<i>mac-addr</i> [vlan <i>vlan-id</i>] [static dynamic] [interface { <i>interface-name</i> / <i>interface-type interface-num</i> }] [vlan <i>vlan-id</i>] [count]]
Display the aging time of dynamic address table entries	display mac-address aging-time

44.5 Resetting MAC Addresses

After configuration, use the **reset mac-address** command in user view to reset the configured mac-address table information.

Table 44-5 Reset MAC addresses

Operation	Command
Reset mac-address table information	reset mac-address { all dynamic static interface { <i>interface_type interface_num</i> / <i>interface_name</i> } Vlan <i>vlan_number</i> }

44.6 MAC Address Table Management Configuration Example

I. Network requirements

The user logs into the switch through the Console port to configure the address table management. It is required to set the address aging time to 500s and add a static address 00e0-fc35-dc71 to Ethernet2/1/2 in vlan1.

II. Network diagram

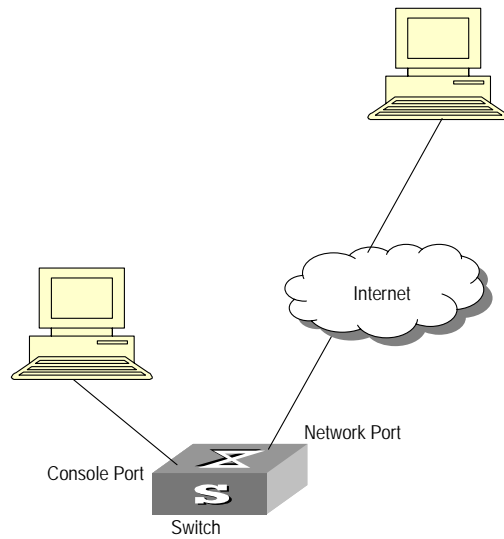


Figure 44-2 Network diagram for address table management configuration

III. Configuration procedure

Enter the system view of the switch.

```
<SW8800> system-view
```

Add a MAC address (specify the native VLAN, port and state).

```
[SW8800] mac-address static 00e0-fc35-dc71 interface ethernet2/1/2 vlan 1
```

Set the address aging time to 500s.

```
[SW8800] mac-address timer 500
```

Display the MAC address configurations in any view.

```
[SW8800] display mac-address interface ethernet2/1/2
```

MAC ADDR	VLAN ID	STATE	PORT INDEX	AGING TIME(s)
00-e0-fc-35-dc-71	1	Static	Ethernet2/1/2	NOAGED
00-e0-fc-17-a7-d6	1	Learned	Ethernet2/1/2	500
00-e0-fc-5e-b1-fb	1	Learned	Ethernet2/1/2	500
00-e0-fc-55-f1-16	1	Learned	Ethernet2/1/2	500

--- 4 mac address(es) found on port Ethernet2/1/2 ---

Chapter 45 Device management

45.1 Device Management Overview

With the device management function, the Ethernet Switch can display the current running state and event debugging information about the slots, thereby implementing the maintenance and management of the state and communication of the physical devices. In addition, there is a command available for rebooting the system, when some function failure occurs.

45.2 Device Management Configuration

The main device management tasks for you are to check the status of the boards, CPU, and the memory usage of the switch.

The following sections describe the configuration tasks for device management:

- Rebooting the Ethernet Switch
- Enabling the Timing Reboot Function
- Designating the APP Adopted on Next Booting
- Upgrading BootROM
- Setting Slot Temperature Limit
- Updating Service Processing Boards

45.2.1 Rebooting the Ethernet Switch

It would be necessary for users to reboot the Ethernet switch when failure occurs.

Perform the following configuration in user view.

Table 45-1 Reboot Ethernet switch

Operation	Command
Root Ethernet switch	reboot [slot <i>slot-no</i>]

45.2.2 Enabling the Timing Reboot Function

After you enable the timing reboot function on the switch, the switch will be rebooted on the specified time.

Perform the following configuration in user view, and **display schedule reboot** command can be performed in any view.

Table 45-2 Enable the Timing Reboot Function

Operation	Command
Enable the timing reboot function of the switch, and set specified time and date	schedule reboot at <i>hh:mm</i> [<i>yyyy/mm/dd</i>]
Enable the timing reboot function of the switch, and set waiting time	schedule reboot delay { <i>hhh:mm</i> <i>mmm</i> }
Cancel the parameter configuration of timing reboot function of the switch	undo schedule reboot
Check the parameter configuration of the reboot terminal service of the current switch	display schedule reboot

Note:

The precision of switch timer is 1 minute. The switch will reboot in one minute when time comes to the specified rebooting point.

45.2.3 Designating the APP Adopted on Next Booting

APP refers to the host application deployed on switch. In the case that there are several APPs in the Flash Memory, you can use this command to designate the APP adopted when booting the Ethernet switch next time.

Perform the following configuration in user view.

Table 45-3 Designate the APP adopted when booting the Ethernet switch next time

Operation	Command
Designate the APP adopted when booting the Ethernet switch next time	boot boot-loader <i>file-url</i> [slot <i>slot-no</i>]

Note:

The Switch 8800 supports master board and slave board. The two boards both have file system. User can operate the file on the two boards. In the case user designate the APP adopted on slave board next time, the file directory or URL should be started with "slot[No.]#flash:", the [No.] is the slave board number. For example, suppose slot 1 is slave board, "text.txt" file URL on slave board should be "slot1#flash:/text.txt".

45.2.4 Upgrading BootROM

You can use followed command to upgrade the BootROM with the BootROM program in the Flash Memory. This configuration task facilitates the remote upgrade. You can upload the BootROM program file from a remote end to the switch by FTP and then use this command to upgrade the BootROM.

Perform the following configuration in user view.

Table 45-4 Upgrade BootROM

Operation	Command
Upgrade BootROM	boot bootrom <i>file-url slot slot-num-list</i>

Note:

The Switch 8800 supports master board and slave board. The two boards both have file system. User can operate the file on the two boards. In the case user upgrade the BootROM adopted on slave board, the file directory or URL should be started with “slot[No.]#flash:”, the [No.] is the slave board number. For example, suppose slot 1 is slave board, “text.txt” file URL on slave board should be “slot1#flash:/text.txt”.

When you are upgrading the BootROM on a slave board, the boot code file must be present in the local flash.

45.2.5 Setting Slot Temperature Limit

The switch system alarms when the temperature on a slot exceeds the preset limit.

Perform the following configuration in user view.

Table 45-5 Set slot temperature limit

Operation	Command
Set slot temperature limit	temperature-limit <i>slot down-value up-value</i>
Restore temperature limit to default value	undo temperature-limit <i>slot</i>

45.2.6 Updating Service Processing Boards

The size of the flash for a main control board in a Switch 8800 is 16 MB, while the size of current host software including the host application of service processing board reaches over 15MB. If a compact flash (CF) card is not configured, the current flash cannot provide enough room to save loading files. Therefore for a Switch 8800 with the

main control board of a 16 MB flash, the service processing board cannot be updated according to the original procedure. To update it, you need to execute the following command to download host software containing the app file of service processing board host application to the system's synchronous dynamic random access memory (SDRAM).

Note:

If you configure a CF card or the flash room of a subsequent main control board expands to 32MB, you need not to change the method to update boards. Then when loading files you only need to choose the APP files containing the application file of service processing board to update common interface boards and service processing boards.

The error message "% Device can't be found or file can't be found in the directory" can indicate that the CF card is not formatted.

Perform the following configuration in system view.

Table 45-6 Update service processing boards

Operation	Command
Download the host software of service processing board to the system memory	update l3plus slot <i>slot-no</i> filename <i>file-name</i> ftpsrvr <i>server-name</i> username <i>user-name</i> password <i>password</i> [port <i>port-num</i>]

45.3 Displaying and Debugging Device Management

After the above configuration, execute **display** command in any view to display the running of the device management configuration, and to verify the effect of the configuration.

Table 45-7 Display and Debug device management

Operation	Command
Display the module types and running states of each card.	display device [detail [shelf <i>shelf-no</i>] [frame <i>frame-no</i>] [slot <i>slot-no</i>]]
Display the application deployed on next startup	display boot-loader
Display the running state of the built-in fans.	display fan [<i>fan-id</i>]
Display the Used status of switch memory	display memory [slot <i>slot-no</i>]
Display the state of the power.	display power [<i>powe-ID</i>]

Display CPU occupancy

display cpu [slot slot-no]

45.4 Device Management Configuration Example

45.4.1 Using the Switch as an FTP Client to Implement the Remote Upgrade

I. Network requirements

The user logs into the switch using Telnet, downloads the application from the FTP server to the flash memory of the switch, and implements remote upgrade using the right commands.

The switch serves as an FTP client and the remote PC as an FTP server. The configuration on the FTP server is as follows: an FTP user is configured with the name switch, the password hello and the read & write authority over the Switch root directory on the PC. The IP address of a VLAN interface on the switch is 1.1.1.1, and the IP address of the PC is 2.2.2.2. The switch and PC are reachable with each other.

The switch applications switch.app and boot.app are stored on the PC. Using FTP, these files can be downloaded from the remote FTP server to the switch.

II. Network diagram

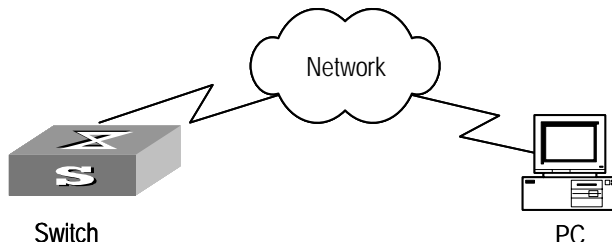


Figure 45-1 Network diagram for FTP configuration

III. Configuration procedure

- 1) Configure FTP server parameters on the PC: a user named as switch, password hello, read & write authority over the Switch directory on the PC. No further details are provided here.
- 2) Configure the switch

The switch has been configured with a Telnet user named as user, as 3-level user, with password hello, requiring username and password authentication.

Use the **telnet** command to log into the switch.

```
<SW8800>
```

**Caution:**

If the flash memory of the switch is not enough, you need to first delete the existing programs in the flash memory and then download the new ones to the memory.

Enter the corresponding command in user view to establish FTP connection. Then enter correct username and password to log into the FTP server.

```
<SW8800> ftp 2.2.2.2
Trying ...
Press CTRL+K to abort
Connected.
220 WFTPD 2.0 service (by Texas Imperial Software) ready for new user
User(none):switch
331 Give me your password, please
Password:*****
230 Logged in successfully
[ftp]
```

Enter the authorized directory of the FTP server.

```
[ftp] cd switch
```

Use the **get** command to download the switch.app and boot.app files from the FTP server to the flash directory on the FTP client.

```
[ftp] get switch.app
[ftp] get boot.app
```

Use the **quit** command to release FTP connection and return to user view.

```
[ftp] quit
<SW8800>
```

Upgrade the BootROM of main board 0.

```
<SW8800> boot bootrom boot.app slot 0
```

Use the **boot boot-loader** command to specify the downloaded program as the application at the next login and reboot the switch.

```
<SW8800> boot boot-loader switch.app
<SW8800>display boot-loader
The app to boot at the next time is: flash:/Switch.app
The app to boot of board 0 at this time is: flash:/PLAT.APP
<SW8800> reboot
```

45.4.2 Use the Switch as an FTP Server to Implement the Remote Upgrade

I. Network requirements

The switch serves as an FTP server and the PC as an FTP client. The configuration on the FTP server is as follows: an FTP user is configured with the name switch, the password hello and the read & write authority over the root directory of the switch. The IP address of a VLAN interface on the switch is 1.1.1.1, and the IP address of the PC is 2.2.2.2. The switch and PC are reachable with each other.

The switch application switch.app is stored on the PC. Using FTP, this file can be uploaded from the PC to the switch remotely, and the configuration file vrpcfg.txt on the switch can be downloaded to the PC as a backup.

II. Network diagram

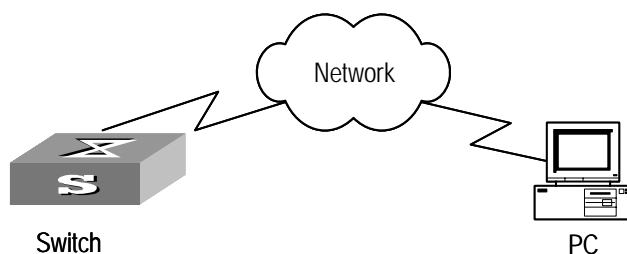


Figure 45-2 Network diagram for FTP configuration

III. Configuration procedure

1) Configure the switch

Log into the switch through the console port locally or through telnet remotely (refer to the getting start module for details about the login modes).

```
<SW8800>
```

Enable FTP on the switch; configure a username, password and path.

```
[SW8800] ftp server enable
```

```
[SW8800] local-user switch
```

```
[SW8800-luser-switch] service-type ftp ftp-directory flash:
```

```
[SW8800-luser-switch] password simple hello
```

2) Run the FTP client program on the PC to set up an FTP connection with the switch. Then upload the switch program switch.app to the flash root directory on the switch and download the configuration file vrpcfg.txt from the switch. The FTP client program is not provided along with the switch, so, it is for you to purchase and install it.

**Caution:**

If the flash memory on the switch is not sufficient, delete the original application program in the flash before uploading the new one into the flash of the switch.

3) After uploading, performs upgrading on the switch.

<SW8800>

You can use the **boot boot-loader** command to specify the new file as the application program on the next booting and reboot the switch to implement the upgrading of the application program.

```
<SW8800> boot boot-loader switch.app
```

```
<SW8800> reboot
```

Chapter 46 System Maintenance and Debugging

46.1 Basic System Configuration

The basic system configuration and management include:

- Switch name setting
- System clock setting
- Time zone setting
- Summer time setting

46.1.1 Setting a Name for a Switch

Perform the operation of **sysname** command in the system view.

Table 46-1 set a name for a Switch

Operation	Command
Set the switch name	sysname <i>sysname</i>
Restore the switch name to default value	undo sysname

46.1.2 Setting the System Clock

Perform the following configuration in user view.

Table 46-2 Set the system clock

Operation	Command
Set the system clock	clock datetime <i>HH:MM:SS YYYY/MM/DD</i>

46.1.3 Setting the Time Zone

You can configure the name of the local time zone and the time difference between the local time and the standard Universal Time Coordinated (UTC).

Perform the following configuration in user view.

Table 46-3 Set the time zone

Operation	Command
Set the local time	clock timezone <i>zone_name</i> { add minus } <i>HH:MM:SS</i>
Restore to the default UTC time zone	undo clock timezone

By default, the UTC time zone is adopted.

46.1.4 Setting the Summer Time

You can set the name, starting and ending time of the summer time.

Perform the following configuration in user view.

Table 46-4 Set the summer time

Operation	Command
Set the name and range of the summer time	clock summer-time <i>zone_name</i> { one-off repeating } <i>start-time start-date end-time end-date offset-time</i>
Remove the setting of the summer time	undo clock summer-time

By default, the summer time is not set.

46.2 Displaying the State and Information of the System

The switch provides the **display** command for displaying the the system state and statistics information.

For the **display** commands related to each protocols and different ports, refer to the relevant chapters. The following **display** commands are used for displaying the system state and the statistics information.

Perform the following operations in any view.

Table 46-5 The **display** commands of the system

Operation	Command
Display the system clock	display clock
Display the system version	display version
Display the state of the debugging	display debugging [interface { <i>interface-name</i> <i>interface-type interface-number</i> }] [<i>module-name</i>]

Operation	Command
Display the information about the optical module connected with a in-place optical port on current frame	display fiber-module or display fiber-module [<i>interface-type</i> <i>interface-number</i> <i>interface-name</i>]

46.3 System Debugging

46.3.1 Enabling/Disabling the Terminal Debugging

The Ethernet switch provides various ways for debugging most of the supported protocols and functions, which can help you diagnose and address the errors.

The following switches can control the outputs of the debugging information:

- Protocol debugging switch controls the debugging output of a protocol.
- Terminal debugging switch controls the debugging output on a specified user screen.

The figure below illustrates the relationship between two switches.

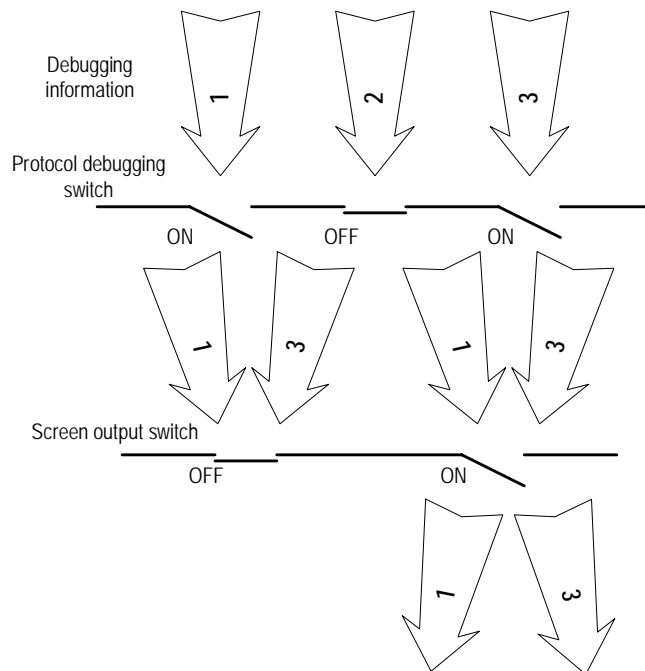


Figure 46-1 Debug output

You can use the following commands to control the above-mentioned debugging.

Perform the following operations in user view.

Table 46-6 Enabling/Disabling the debugging

Operation	Command
Enable the protocol debugging	debugging { all [timeout interval] <i>module-name</i> [<i>debugging-option</i>] }
Disable the protocol debugging	undo debugging { all <i>module-name</i> [<i>debugging-option</i>] }
Enable the terminal debugging	terminal debugging
Disable the terminal debugging	undo terminal debugging

For more about the usage and format of the debugging commands, refer to the relevant chapters.

Note:

Since the debugging output will affect the system operating efficiency, do not enable the debugging without necessity, especially use the **debugging all** command with caution. When the debugging is over, disable all the debugging.

46.3.2 Displaying Diagnostic Information

When the Ethernet switch does not run well, you can collect all sorts of information about the switch to locate the source of fault. However, each module has its corresponding display command, you can use **display diagnostic-information** command.

You can perform the following operations in any view.

Table 46-7 displaying diagnostic information

Operation	Command
display diagnostic information	display diagnostic-information

Note:

When using the display diagnostic-information command to keep track of Ethernet switch, you should execute the command at least twice so that you can compare the information for locating problem.

46.4 Testing Tools for Network Connection

46.4.1 ping

The **ping** command can be used to check the network connection and if the host is reachable.

Perform the following configuration in any view.

Table 46-8 The **ping** command

Operation	Command
Support IP ping	ping [ip] [-a <i>ip-address</i> -c <i>count</i> -d -h <i>tth</i> -i { <i>interface-type interface-num</i> <i>interface-name</i> } -n -p <i>pattern</i> -q -r -s <i>packet-size</i> -t <i>timeout</i> -tos <i>tos</i> -v -vpn-instance <i>vpn-instance-name</i>]* <i>host</i>

The output of the command includes:

- The response to each ping message. If no response packet is received when time is out, "Request time out" information appears. Otherwise, the data bytes, the packet sequence number, TTL, and the round-trip time of the response packet will be displayed.
- The final statistics, including the number of the packets the switch sent out and received, the packet loss ratio, the round-trip time in its minimum value, mean value and maximum value.

46.4.2 ping-distribute enable

Use the **ping-distribute enable** command to enable the ping distribution function.

Use the **undo ping-distribute enable** command to disable the ping distribution function.

Perform the following configuration in system view.

Table 46-9 Enable/disable the ping distribution function

Operation	Command
Enable the ping distribution function	ping-distribute enable
Disable the ping distribution function	undo ping-distribute

By default, the ping distribution function is enabled.

46.4.3 tracert

The **tracert** is used for testing the gateways passed by the packets from the source host to the destination one. It is mainly used for checking if the network is connected and analyzing where the fault occurs in the network.

The execution process of tracert is described as follows: Send a packet with TTL value as 1 and the first hop sends back an ICMP error message indicating that the packet cannot be sent, for the TTL is timeout. Re-send the packet with TTL value as 2 and the second hop returns the TTL timeout message. The process is carried over and over until the packet reaches the destination. The purpose to carry out the process is to record the source address of each ICMP TTL timeout message, so as to provide the route of an IP packet to the destination.

Perform the following configuration in any view.

Table 46-10 The **tracert** command

Operation	Command
Trace route	tracert [-a <i>source-IP</i> -f <i>first-TTL</i> -m <i>max-TTL</i> -p <i>port</i> -q <i>num-packet</i> -vpn-instance <i>vpn-instance-name</i> -w <i>timeout</i>] <i>string</i>

46.5 Logging Function

46.5.1 Introduction to Info-center

The Info-center is an indispensable part of the Ethernet switch. It serves as an information center of the system software modules. The logging system is responsible for most of the information outputs, and it also makes detailed classification to filter the information efficiently. Coupled with the debugging program, the info-center provides powerful support for the network administrators and the R&D personnel to monitor the operating state of networks and diagnose network failures.

When the log information is output to terminal or log buffer, the following parts will be included:

% <priority> Timestamp Sysname Module name/Severity/Digest: Content

For example:

```
%Jun 7 05:22:03 2003 SW8800 IFNET/6/UPDOWN:Line protocol on interface
Ethernet2/1/2, changed state to UP
```

When the log information is output to info-center, the first part will be "<Priority>".

For example:

```
% <189>Jun 7 05:22:03 2003 SW8800 IFNET/6/UPDOWN:Line protocol on interface
Ethernet0/0/0, changed state to UP
```

The description of the components of log information is as follows:

1) %

In practical output, some of the information is started with the % character, which means a logging is necessary.

2) Priority

The priority is computed according to following formula: $\text{facility} * 8 + \text{severity} - 1$. The default value for the facility is 23. The range of severity is 1~8, and the severity will be introduced in separate section.

Priority is only effective when information is send to log host. There is no character between priority and timestamp.

3) Timestamp

If the logging information is send to the log host, the default format of timestamp is date
The date format of timestamp is " Mmm dd hh:mm:ss yyyy".

" Mmm " is month field, such as: Jan, Feb, Mar, Apr, May, Jun, Jul, Aug, Sep, Oct, Nov, Dec.

"dd" is day field, if the day is little than 10th, one blank should be added, such as " 7".

"hh:mm:ss" is time field, "hh" is from 00 to 23, "mm" and "ss" are from 00 to 59.

"yyyy" is year field.

4) Sysname

The sysname is the host name, the default value is "SW8800".

User can change the host name through **sysname** command.

Notice: There is a blank between sysname and module name.

5) Module name

The module name is the name of module which create this logging information, the following sheet list some examples:

Table 46-11 The module name field

Module name	Description
8021X	802.1X module
ACL	Access control list module
ADBM	MAC address management module
ARP	Address resolution protocol module
BGP	Border gateway protocol module
CFM	Configuration file management module
CMD	Command module
default	Default settings for all the modules

Module name	Description
DEV	Device management module
DHCP	Dynamic host configuration protocol module
DIAGCLI	Diagnosis module
DNS	Domain name server module
DRVMPLS	Multiprotocol label switching drive module
DRVL2	Layer 2 drive module
DRVL3	Layer 3 drive module
DRVL3MC	Layer 3 multicast module
MPLS	MPLS drive module
DRVQACL	QACL drive module
DRVVPLS	Virtual private LAN service drive module
ETH	Ethernet module
FTPS	FTP server module
HA	High availability module
IFNET	Interface management module
IGSP	IGMP snooping module
IP	Internet protocol module
ISIS	Intermediate system-to-intermediate system intradomain routing protocol module
L2INF	L2 interface management module
L2V	L2 VPN module
LACL	LAN switch ACL module
LDP	label distribution protocol module
LINKAGG	LINKAGG module
LQOS	LAN switch QoS module
LS	Local server module
LSPAGENT	Label switched path agent module
LSPM	Label switch path management module
MIX	Dual system management module
MMC	MMC module
MODEM	Modem module
MPLSFW	MPLS forward module

Module name	Description
MPM	Multicast port management module
MSDP	Multicast source discovery protocol module
MSTP	Multiple spanning tree protocol module
NAT	Network address translation module
NTP	Network time protocol module
OSPF	Open shortest path first module
PHY	Physical sublayer & physical layer module
PPP	Point to point protocol module
PSSINIT	PSSINIT module
RDS	RADIUS module
RM	Routing management module
RMON	Remote monitor module
RSA	RSA (Revest, Shamir and Adleman) encryption module
RTPRO	Routing protocol module
SHELL	User interface module
SNMP	Simple network management protocol module
SOCKET	Socket module
SSH	Secure shell module
SYSM	System manage veneer module
SYSMIB	System MIB module
TAC	Terminal access controller module
TELNET	Telnet module
USERLOG	User calling logging module
VFS	Virtual file system module
VLAN	Virtual local area network module
VOS	Virtual operate system module
VRRP	VRRP (virtual router redundancy protocol) module
VTY	VTY (virtual type terminal) module

Notice: There is a slash (/) between module name and severity.

6) Severity

Switch information falls into three categories: log information, debugging information and trap information. The info-center classifies every kind of information into 8 severity or urgent levels. The log filtering rule is that the system prohibits outputting the information whose severity level is greater than the set threshold. The more urgent the logging packet is, the smaller its severity level is. The level represented by “emergencies” is 1, and that represented by “debugging” is 8. Therefore, when the threshold of the severity level is “debugging”, the system will output all the information. Definition of severity in logging information is as followed.

Table 46-12 Info-center-defined severity

Severity	Value	Description
emergencies	1	The extremely emergent errors
alerts	2	The errors that need to be corrected immediately.
critical	3	Critical errors
errors	4	The errors that need to be concerned but not critical
warnings	5	Warning, there might exist some kinds of errors.
notifications	6	The information should be concerned.
informational	7	Common prompting information
debugging	8	Debugging information

Notice: There is a slash between severity and digest.

7) Digest

The digest is abbreviation, it represent the abstract of contents.

Notice: There is a colon between digest and content. The digest can be up to 32 characters long.

46.5.2 Info-center Configuration

Switch supports 7 output directions of information.

The system assigns a channel in each output direction by default. See the table below.

Table 46-13 Numbers and names of the channels for log output

Output direction	Channel number	Default channel name
Console	0	console
Monitor	1	monitor
Info-center loghost	2	loghost
Trap buffer	3	trapbuf

Logging buffer	4	logbuf
snmp	5	snmpagent
Log file	6	logfile

Note:

The settings in the 7 directions are independent from each other. The settings will take effect only after enabling the information center.

The info-center of Ethernet Switch has the following features:

- Support to output log in 7 directions, i.e., Console, monitor to Telnet terminal, logbuffer, loghost, trapbuffer, and SNMP log file.
 - The log is divided into 8 levels according to the significance and it can be filtered based on the levels.
 - The information can be classified in terms of the source modules and the information can be filtered in accordance with the modules.
 - The output language can be selected between Chinese and English.
- 1) Sending the configuration information to the loghost.

Table 46-14 Send the configuration information to the loghost

Device	Configuration	Default value	Configuration description
Switch	Enable info-center	By default, info-center is enabled.	Other configurations are valid only if the info-center is enabled.
	Set the information output direction to the loghost	—	The configuration about the loghost on the switch and that on loghost must be the same; otherwise the information cannot be sent to the loghost correctly.
	Set information source	—	You can define which modules and information to be sent out and the time-stamp format of information, and so on. You must turn on the switch of the corresponding module before defining output debugging information.
Loghost	Refer to configuration cases for related log host configuration	—	—

- 2) Sending the configuration information to the console terminal.

Table 46-15 Send the configuration information to the console terminal.

Device	Configuration	Default value	Configuration description
Switch	Enable info-center	By default, info-center is enabled.	Other configurations are valid only if the info-center is enabled.
	Set the information output direction to the Console	—	—
	Set information source	—	You can define which modules and information to be sent out and the time-stamp format of information, and so on. You must turn on the switch of the corresponding module before defining output debugging information.
	Enable terminal display function	—	You can view debugging information after enabling terminal display function

3) Sending the configuration information to the monitor terminal

Table 46-16 Send the configuration information to the monitor terminal

Device	Configuration	Default value	Configuration description
Switch	Enable info-center	By default, info-center is enabled.	Other configurations are valid only if the info-center is enabled.
	Set the information output direction to the monitor	—	—
	Set information source	—	You can define which modules and information to be sent out and the time-stamp format of information, and so on. You must turn on the switch of the corresponding module before defining output debugging information.
	Enable the terminal display function and this function for the corresponding information	—	For Telnet terminal and dumb terminal, to view the information, you must enable the current terminal display function using the terminal monitor command.

4) Sending the configuration information to the log buffer.

Table 46-17 Send the configuration information to the log buffer

Device	Configuration	Default value	Configuration description
Switch	Enable info-center	By default, info-center is enabled.	Other configurations are valid only if the info-center is enabled.
	Set the information output direction to the logbuffer	—	You can configure the size of the log buffer at the same time.
	Set information source	—	You can define which modules and information to be sent out and the time-stamp format of information, and so on. You must turn on the switch of the corresponding module before defining output debugging information.

5) Sending the configuration information to the trap buffer.

Table 46-18 Send the configuration information to the trap buffer

Device	Configuration	Default value	Configuration description
Switch	Enable info-center	By default, info-center is enabled.	Other configurations are valid only if the info-center is enabled.
	Set the information output direction to the trapbuffer	—	You can configure the size of the trap buffer at the same time.
	Set information source	—	You can define which modules and information to be sent out and the time-stamp format of information, and so on. You must turn on the switch of the corresponding module before defining output debugging information.

6) Sending the configuration information to SNMP

Table 46-19 Send the configuration information to SNMP

Device	Configuration	Default value	Configuration description
Switch	Enable info-center	By default, info-center is enabled.	Other configurations are valid only if the info-center is enabled.

	Set the information output direction to SNMP	—	—
	Set information source	—	You can define which modules and information to be sent out and the time-stamp format of information, and so on. You must turn on the switch of the corresponding module before defining output debugging information.
	Configure SNMP features	—	Refer to Chapter 47 SNMP Configuration
Network management workstation	The same as the SNMP configuration of the switch	—	—

46.5.3 Sending the Configuration Information to the Loghost

To send configuration information to the loghost, follow the steps below:

- 1) Enabling info-center

Perform the following configuration in system view.

Table 46-20 Enable/disable info-center

Operation	Command
Enable info-center	info-center enable
Disable info-center	undo info-center enable

Note:

Info-center is enabled by default. After info-center is enabled, system performances are affected when the system processes much information because of information classification and outputting.

- 2) Configuring to output information to the loghost

Perform the following configuration in system view.

Table 46-21 Configure to output information to the loghost

Operation	Command
Output information to the loghost	info-center loghost <i>host-ip-addr</i> [channel { <i>channel-number</i> <i>channel-name</i> } facility <i>local-number</i> language { chinese english }]*
Cancel the configuration of outputting information to loghost	undo info-center loghost <i>host-ip-addr</i>

Note that the IP address of log host must be correct.

Note:

Ensure to enter the correct IP address using the **info-center loghost** command to configure loghost IP address. If you enter a loopback address, the system prompts of invalid address appears.

3) Configuring information source on the switch

By this configuration, you can define the information that sent to console terminal is generated by which modules, information type, information level, and so on.

Perform the following configuration in system view.

Table 46-22 Define information source

Operation	Command
Define information source	info-center source { <i>modu-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> }*]*
Cancel the configuration of information source	undo info-center source { <i>modu-name</i> default all } channel { <i>channel-number</i> <i>channel-name</i> }

modu-name specifies the module name; **default** represents all the modules; **all** represents all the information filter configuration in channelnum channel except default; **level** refers to the severity levels; *severity* specifies the severity level of information. The information with the level below it will not be output. *channel-number* specifies the channel number and *channel-name* specifies the channel name.

When defining the information sent to the loghost, *channel-number* or *channel-name* must be set to the channel that corresponds to loghost direction.

Every channel has been set with a default record, whose module name is **default** and the module number is 0xffff0000. However, for different channels, the default record may have different default settings of log, trap and debugging. When there is no specific configuration record for a module in the channel, use the default one.

Note:

If you want to view the debugging information of some modules on the switch, you must select **debugging** as the information type when configuring information source, meantime using the **debugging** command to turn on the debugging switch of those modules.

You can use the following commands to configure log information, debugging information and the time-stamp output format of trap information. This configuration will affect the timestamp of the displayed information.

Perform the following configuration in system view:

Table 46-23 Configure the output format of time-stamp

Operation	Command
Configure the output format of the time-stamp	info-center timestamp { log trap debugging } { boot date none }
Output time-stamp is disabled	undo info-center timestamp { log trap debugging }

4) Configuring the loghost

The configuration on the loghost must be the same with that on the switch. For related configuration, see the configuration examples in the later part.

46.5.4 Sending the Configuration Information to Console terminal

To send configuration information to console terminal, follow the steps below:

1) Enabling info-center

Perform the following configuration in system view.

Table 46-24 Enable/disable info-center

Operation	Command
Enable info-center	info-center enable
Disable info-center	undo info-center enable

Note:

Info-center is enabled by default. After info-center is enabled, system performances are affected when the system processes much information because of information classification and outputting.

2) Configuring to output information to console terminal

Perform the following configuration in system view.

Table 46-25 Configure to output information to console terminal

Operation	Command
Output information to Console	info-center console channel { <i>channel-number</i> <i>channel-name</i> }
Cancel the configuration of outputting information to Console	undo info-center console channel

3) Configuring information source on the switch

By this configuration, you can define the information that sent to console terminal is generated by which modules, information type, information level, and so on.

Perform the following configuration in system view:

Table 46-26 Define information source

Operation	Command
Define information source	info-center source { <i>modu-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> }*]*
Cancel the configuration of information source	undo info-center source { <i>modu-name</i> default all } channel { <i>channel-number</i> <i>channel-name</i> }

modu-name specifies the module name; **default** represents all the modules; **all** represents all the information filter configuration in channelnum channel except default; **level** refers to the severity levels; *severity* specifies the severity level of information. The information with the level below it will not be output. *channel-number* specifies the channel number and *channel-name* specifies the channel name.

When defining the information sent to the console terminal, *channel-number* or *channel-name* must be set to the channel that corresponds to Console direction.

Every channel has been set with a default record, whose module name is **default** and the module number is 0xffff0000. However, for different channels, the default record

may have different default settings of log, trap and debugging. When there is no specific configuration record for a module in the channel, use the default one.

Note:

If you want to view the debugging information of some modules on the switch, you must select **debugging** as the information type when configuring information source, meantime using the **debugging** command to turn on the debugging switch of those modules.

You can use the following commands to configure log information, debugging information and the time-stamp output format of trap information.

This configuration will affect the timestamp of the displayed information.

Perform the following configuration in system view:

Table 46-27 Configure the output format of time-stamp

Operation	Command
Configure the output format of the time-stamp	info-center timestamp { log trap debugging } { boot date none }
Output time-stamp is disabled	undo info-center timestamp { log trap debugging }

4) Enable terminal display function

To view the output information at the console terminal, you must first enable the corresponding log, debugging and trap information functions at the switch.

For example, if you have set the log information as the information sent to the console terminal, now you need to use the **terminal logging** command to enable the terminal display function of log information on the switch, then you can view the information at the console terminal.

Perform the following configuration in user view:

Table 46-28 Enable terminal display function

Operation	Command
Enable terminal display function of debugging information	terminal debugging
Disable terminal display function of debugging information	undo terminal debugging
Enable terminal display function of log information	terminal logging
Disable terminal display function of log information	undo terminal logging

Enable terminal display function of trap information	terminal trapping
Disable terminal display function of trap information	undo terminal trapping

46.5.5 Sending the Configuration Information to Telnet Terminal or Dumb Terminal

To send configuration information to Telnet terminal or dumb terminal, follow the steps below:

- 1) Enabling info-center

Perform the following configuration in system view.

Table 46-29 Enable/disable Info-center

Operation	Command
Enable info-center	info-center enable
Disable info-center	undo info-center enable

Note:

Info-center is enabled by default. After info-center is enabled, system performances are affected when the system processes much information because of information classification and outputting.

- 2) Configuring to output information to Telnet terminal or dumb terminal

Perform the following configuration in system view.

Table 46-30 Configure to output information to Telnet terminal or dumb terminal

Operation	Command
Output information to Telnet terminal or dumb terminal	info-center monitor channel { <i>channel-number</i> <i>channel-name</i> }
Cancel the configuration of outputting information to Telnet terminal or dumb terminal	undo info-center monitor channel

- 3) Configuring information source on the switch

By this configuration, you can define the information that sent to Telnet terminal or dumb terminal is generated by which modules, information type, information level, and so on.

Perform the following configuration in system view:

Table 46-31 Define information source

Operation	Command
Define information source	info-center source { <i>modu-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> }*]*
Cancel the configuration of information source	undo info-center source { <i>modu-name</i> default all } channel { <i>channel-number</i> <i>channel-name</i> }

modu-name specifies the module name; **default** represents all the modules; **all** represents all the information filter configuration in channelnum chunnel except default; **level** refers to the severity levels; *severity* specifies the severity level of information. The information with the level below it will not be output. *channel-number* specifies the channel number and *channel-name* specifies the channel name.

When defining the information sent to Telnet terminal or dumb terminal, *channel-number* or *channel-name* must be set to the channel that corresponds to monitor direction.

Every channel has been set with a default record, whose module name is **default** and the module number is 0xffff0000. However, for different channels, the default record may have different default settings of log, trap and debugging. When there is no specific configuration record for a module in the channel, use the default one.

Note:

When there are more than one Telnet users or monitor users at the same time, some configuration parameters should be shared among the users, such as module-based filtering settings and severity threshold. When a user modifies these settings, it will be reflected on other clients.

Note:

If you want to view the debugging information of some modules on the switch, you must select **debugging** as the information type when configuring information source, meantime using the **debugging** command to turn on the debugging switch of those modules.

You can use the following commands to configure log information, debugging information and the time-stamp output format of trap information.

This configuration will affect the timestamp of the displayed information.

Perform the following configuration in system view:

Table 46-32 Configure the output format of time-stamp

Operation	Command
Configure the output format of the time-stamp	info-center timestamp { log trap debugging } { boot date none }
Output time-stamp is disabled	undo info-center timestamp { log trap debugging }

4) Enabling terminal display function

To view the output information at the Telnet terminal or dumb terminal, you must first enable the terminal display function, and then the corresponding terminal display function of log information on the switch.

For example, if you have set the log information as the information sent to the Telnet terminal or dumb terminal, now you need to use the terminal monitor command to enable the terminal display function and the **terminal logging** command to enable the terminal display function of log information on the switch, then you can view the information at the Telnet terminal or dumb terminal.

Perform the following configuration in user view:

Table 46-33 Enable terminal display function

Operation	Command
Enable terminal display function of log, debugging and trap information	terminal monitor
Disable terminal display function of the above information	undo terminal monitor
Enable terminal display function of debugging information	terminal debugging
Disable terminal display function of debugging information	undo terminal debugging
Enable terminal display function of log information	terminal logging
Disable terminal display function of log information	undo terminal logging
Enable terminal display function of trap information	terminal trapping
Disable terminal display function of trap information	undo terminal trapping

46.5.6 Sending the Configuration Information to the Log Buffer

To send configuration information to the log buffer, follow the steps below:

1) Enabling info-center

Perform the following configuration in system view.

Table 46-34 Enable/disable info-center

Operation	Command
Enable info-center	info-center enable
Disable info-center	undo info-center enable

Note:

Info-center is enabled by default. After info-center is enabled, system performances are affected when the system processes much information because of information classification and outputting.

2) Configuring to output information to the log buffer

Perform the following configuration in system view.

Table 46-35 Configure to output information to log buffer

Operation	Command
Output information to log buffer	info-center logbuffer [channel { <i>channel-number</i> <i>channel-name</i> } size <i>buffersize</i>]*
Cancel the configuration of outputting information to log buffer	undo info-center logbuffer [channel size]

3) Configuring information source on the switch

By this configuration, you can define the information that sent to log buffer is generated by which modules, information type, information level, and so on.

Perform the following configuration in system view:

Table 46-36 Define information source

Operation	Command
Define information source	info-center source { <i>modu-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> }*]*
Cancel the configuration of information source	undo info-center source { <i>modu-name</i> default all } channel { <i>channel-number</i> <i>channel-name</i> }

modu-name specifies the module name; **default** represents all the modules; **all** represents all the information filter configuration in channelnum channel except default; **level** refers to the severity levels; *severity* specifies the severity level of information. The information with the level below it will not be output. *channel-number* specifies the channel number and *channel-name* specifies the channel name.

When defining the information sent to log buffer, *channel-number* or *channel-name* must be set to the channel that corresponds to logbuffer direction.

Every channel has been set with a default record, whose module name is **default** and the module number is 0xffff0000. However, for different channels, the default record may have different default settings of log, trap and debugging. When there is no specific configuration record for a module in the channel, use the default one.

Note:

If you want to view the debugging information of some modules on the switch, you must select **debugging** as the information type when configuring information source, meantime using the **debugging** command to turn on the debugging switch of those modules.

You can use the following commands to configure log information, debugging information and the time-stamp output format of trap information.

Perform the following configuration in system view:

Table 46-37 Configure the output format of time-stamp

Operation	Command
Configure the output format of the time-stamp	info-center timestamp { log trap debugging } { boot date none }
Output time-stamp is disabled	undo info-center timestamp { log trap debugging }

46.5.7 Sending the Configuration Information to the Trap Buffer

To send configuration information to the trap buffer, follow the steps below:

- 1) Enabling info-center

Perform the following configuration in system view.

Table 46-38 Enable/disable info-center

Operation	Command
Enable info-center	info-center enable

Disable info-center	undo info-center enable
---------------------	--------------------------------

Note:

Info-center is enabled by default. After info-center is enabled, system performances are affected when the system processes much information because of information classification and outputting.

2) Configuring to output information to the trap buffer

Perform the following configuration in system view.

Table 46-39 Configure to output information to trap buffer

Operation	Command
Output information to trap buffer	info-center trapbuffer [size <i>buffersize</i> channel { <i>channel-number</i> <i>channel-name</i> }]*
Cancel the configuration of outputting information to trap buffer	undo info-center trapbuffer [channel size]

3) Configuring information source on the switch

By this configuration, you can define the information that sent to trap buffer is generated by which modules, information type, information level, and so on.

Perform the following configuration in system view:

Table 46-40 Define information source

Operation	Command
Define information source	info-center source { <i>modu-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> }*]*
Cancel the configuration of information source	undo info-center source { <i>modu-name</i> default all } channel { <i>channel-number</i> <i>channel-name</i> }

modu-name specifies the module name; **default** represents all the modules; **all** represents all the information filter configuration in channelnum channel except default; **level** refers to the severity levels; *severity* specifies the severity level of information. The information with the level below it will not be output. *channel-number* specifies the channel number and *channel-name* specifies the channel name.

When defining the information sent to trap buffer, *channel-number* or *channel-name* must be set to the channel that corresponds to trapbuffer direction.

Every channel has been set with a default record, whose module name is **default** and the module number is 0xffff0000. However, for different channels, the default record may have different default settings of log, trap and debugging. When there is no specific configuration record for a module in the channel, use the default one.

Note:

If you want to view the debugging information of some modules on the switch, you must select **debugging** as the information type when configuring information source, meantime using the **debugging** command to turn on the debugging switch of those modules.

You can use the following commands to configure log information, debugging information and the time-stamp output format of trap information.

This configuration will affect the timestamp of the displayed information.

Perform the following configuration in system view:

Table 46-41 Configuring the output format of time-stamp

Operation	Command
Configure the output format of the time-stamp	info-center timestamp { log trap debugging } { boot date none }
Output time-stamp is disabled	undo info-center timestamp { log trap debugging }

46.5.8 Sending the Configuration Information to SNMP Network Management

To send configuration information to SNMP NM, follow the steps below:

- 1) Enabling info-center

Perform the following configuration in system view.

Table 46-42 Enable/disable info-center

Operation	Command
Enable info-center	info-center enable
Disable info-center	undo info-center enable

Note:

Info-center is enabled by default. After info-center is enabled, system performances are affected when the system processes much information because of information classification and outputting.

2) Configuring to output information to SNMP NM

Perform the following configuration in system view.

Table 46-43 Configure to output information to SNMP NM

Operation	Command
Output information to SNMP NM	info-center snmp channel { <i>channel-number</i> <i>channel-name</i> }
Cancel the configuration of outputting information to SNMP NM	undo info-center snmp channel

3) Configuring information source on the switch

By this configuration, you can define the information that sent to SNMP NM is generated by which modules, information type, information level, and so on.

Perform the following configuration in system view:

Table 46-44 Define information source

Operation	Command
Define information source	info-center source { <i>modu-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> }*]*
Cancel the configuration of information source	undo info-center source { <i>modu-name</i> default all } channel { <i>channel-number</i> <i>channel-name</i> }

modu-name specifies the module name; **default** represents all the modules; **all** represents all the information filter configuration in channelnum channel except default **level** refers to the severity levels; *severity* specifies the severity level of information. The information with the level below it will not be output. *channel-number* specifies the channel number and *channel-name* specifies the channel name.

When defining the information sent to SNMP NM, *channel-number* or *channel-name* must be set to the channel that corresponds to SNMP direction.

Every channel has been set with a default record, whose module name is **default** and the module number is 0xffff0000. However, for different channels, the default record may have different default settings of log, trap and debugging. When there is no specific configuration record for a module in the channel, use the default one.

Note:

If you want to view the debugging information of some modules on the switch, you must select **debugging** as the information type when configuring information source, meantime using the **debugging** command to turn on the debugging switch of those modules.

You can use the following commands to configure log information, debugging information and the time-stamp output format of trap information.

This configuration will affect the timestamp of the displayed information.

Perform the following configuration in system view:

Table 46-45 Configure the output format of time-stamp

Operation	Command
Configure the output format of the time-stamp	info-center timestamp { log trap debugging } { boot date none }
Output time-stamp is disabled	undo info-center timestamp { log trap debugging }

4) Configuring of SNMP and network management workstation on the switch

You have to configure SNMP on the switch and the remote workstation to ensure that the information is correctly sent to SNMP NM. Then you can get correct information from network management workstation. SNMP configuration on switch refers to Chapter 47 SNMP Configuration

46.5.9 Displaying and Debugging Info-center

After the above configuration, execute the **display** command in any view to view the running state of the info-center. You also can authenticate the effect of the configuration by viewing displayed information. Execute the **reset** command in user view to clear statistics of info-center.

Perform the following configuration in user view. The **display** command still can be performed in any view.

Table 46-46 Display and debug info-center

Operation	Command
Display the content of information channel	display channel [<i>channel-number</i> <i>channel-name</i>]
Display configuration of system log and memory buffer	display info-center
Display the attribute of logbuffer and the information recorded in logbuffer	display logbuffer [summary] [level [<i>levelnum</i> emergencies alerts critical debugging errors informational notifications warnings]] [[begin exclude include text]] [size <i>sizenum</i>]
Display the summary information recorded in logbuffer	display logbuffer summary [level severity]
Display the attribute of trapbuffer and the information recorded in trapbuffer	display trapbuffer [summary] [level [<i>levelnum</i> emergencies alerts critical debugging errors informational notifications warnings]] [size <i>sizenum</i>]
Clear information in memory buffer	reset logbuffer
Clear information in trap buffer	reset trapbuffer

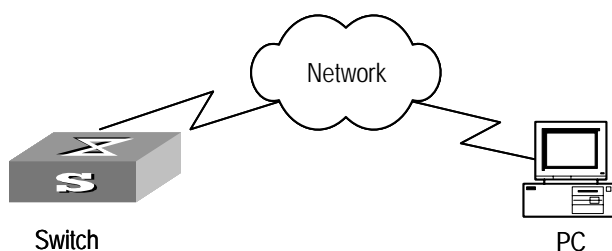
46.5.10 Configuration Examples of Sending Log to the Unix Loghost

I. Network requirements

The network requirements are as follows:

- Sending the log information of the switch to Unix loghost
- The IP address of the loghost is 202.38.1.10
- The information with the severity level above informational will be sent to the loghost
- The output language is English
- The modules that allowed to output information are ARP and IP

II. Network diagram

**Figure 46-2** Network diagram

III. Configuration steps

1) Configuration on the switch

Enable info-center

```
[SW8800] info-center enable
```

Set the host with the IP address of 202.38.1.10 as the loghost; set the severity level threshold value as informational, set the output language to English; set that the modules which are allowed to output information are ARP and IP.

```
[SW8800] info-center loghost 202.38.1.10 facility local4 language english
```

```
[SW8800] info-center source arp channel loghost log level informational
```

```
[SW8800] info-center source ip channel loghost log level informational
```

2) Configuration on the loghost

This configuration is performed on the loghost. The following example is performed on SunOS 4.0 and the operation on Unix operation system produced by other manufactures is generally the same to the operation on SunOS 4.0.

Step 1: Perform the following command as the super user (root).

```
mkdir /var/log/SW8800
```

```
touch /var/log/SW8800/information
```

Step 2: Edit file /etc/syslog.conf as the super user (root), add the following selector/actor pairs.

```
SW8800 configuration messages
```

```
local4.info /var/log/SW8800/information
```

Note:

Note the following points when editing /etc/syslog.conf:

- The note must occupy a line and start with the character #.
 - There must be a tab other than a space as the separator in selector/actor pairs.
 - No redundant space after file name.
 - The device name and the acceptant log information level specified in /etc/syslog.conf must be consistent with info-center loghost and info-center loghost a.b.c.d facility configured on the switch. Otherwise, the log information probably cannot be output to the loghost correctly.
-

Step 3: After the establishment of information (log file) and the revision of /etc/syslog.conf, you should send a HUP signal to syslogd (system daemon), through the following command, to make syslogd reread its configuration file /etc/syslog.conf.

```
ps -ae | grep syslogd
```

```
147
```

```
kill -HUP 147
```

After the above operation, the switch system can record information in related log files.

Note:

To configure facility, severity, filter and the file syslog.conf synthetically, you can get classification in great detail and filter the information.

46.5.11 Configuration examples of sending log to Linux loghost

I. Network requirements

The Network requirements are as follows:

- Sending the log information of the switch to Linux loghost
- The IP address of the loghost is 202.38.1.10
- The information with the severity level above informational will be sent to the loghost
- The output language is English
- All modules are allowed to output information

II. Network diagram

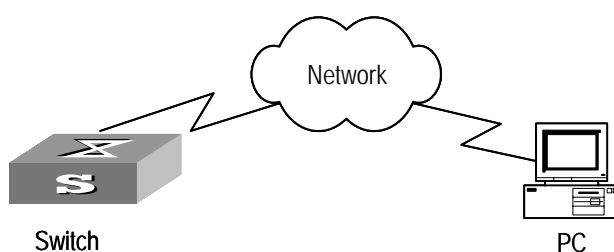


Figure 46-3 Network diagram

III. Configuration procedure

1) Configuration on the switch

Enable info-center

```
[SW8800] info-center enable
```

Set the host with the IP address of 202.38.1.10 as the loghost; set the severity level threshold value as informational, set the output language to English; set all the modules are allowed output information.

```
[SW8800] info-center loghost 202.38.1.10 facility local7 language english  
[SW8800] info-center source default channel loghost log level informational
```

2) Configuration on the loghost

This configuration is performed on the loghost.

Step 1: Perform the following command as the super user (root).

```
mkdir /var/log/SW8800
touch /var/log/SW8800/information
```

Step 2: Edit file `/etc/syslog.conf` as the super user (root), add the following selector/actor pairs.

```
SW8800 configuration messages
local7.info /var/log/SW8800/information
```

Note:

Note the following points when editing `/etc/syslog.conf`:

- The note must occupy a line and start with the character `#`.
- There must be a tab other than a space as the separator in selector/actor pairs.
- No redundant space after file name.
- The device name and the acceptant log information level specified in `/etc/syslog.conf` must be consistent with info-center loghost and info-center loghost a.b.c.d facility configured on the switch. Otherwise, the log information probably cannot be output to the loghost correctly.

Step 3: After the establishment of information (log file) and the revision of `/etc/syslog.conf`, you should view the number of `syslogd` (system daemon) through the following command, kill `syslogd` daemon and reuse `-r` option the start `syslogd` in daemon.

```
ps -ae | grep syslogd
147
kill -9 147
syslogd -r &
```

Note:

For Linux loghost, you must ensure that `syslogd` daemon is started by `-r` option.

After the above operation, the switch system can record information in related log files.

Note:

To configure facility, severity, filter and the file syslog.conf synthetically, you can get classification in great detail and filter the information.

46.5.12 Configuration Examples of Sending Log to the Console Terminal

I. Network requirements

The network requirements are as follows:

- Sending the log information of the switch to console terminal
- The information with the severity level above informational will be sent to the console terminal
- The output language is English

The modules that allowed to output information are ARP and IP

II. Network diagram

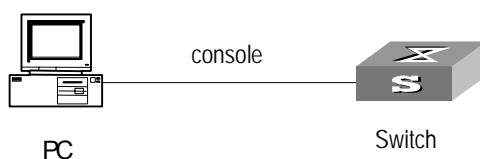


Figure 46-4 Network diagram

III. Configuration procedure

1) Configuration on the switch

Enable info-center.

```
[SW8800] info-center enable
```

Configure console terminal log output; allow modules ARP and IP to output information; the severity level is restricted within the range of emergencies to informational.

```
[SW8800] info-center console channel console
```

```
[SW8800] info-center source arp channel console log level informational
```

```
[SW8800] info-center source ip channel console log level informational
```

Enable terminal display function.

```
<SW8800> terminal logging
```

Chapter 47 SNMP Configuration

47.1 SNMP Overview

By far, the Simple Network Management Protocol (SNMP) has gained the most extensive application in the computer networks. SNMP has been put into use and widely accepted as an industry standard in practice. It is used for ensuring the transmission of the management information between any two nodes. In this way, network administrators can easily search and modify the information on any node on the network. In the meantime, they can locate faults promptly and implement the fault diagnosis, capacity planning and report generating. SNMP adopts the polling mechanism and provides the most basic function set. It is most applicable to the small-sized, fast-speed and low-cost environment. It only requires the unverified transport layer protocol UDP; and is thus widely supported by many other products.

In terms of structure, SNMP can be divided into two parts, namely, Network Management Station and Agent. Network Management Station is the workstation for running the client program. At present, the commonly used NM platforms include Sun NetManager and IBM NetView. Agent is the server software operated on network devices. Network Management Station can send GetRequest, GetNextRequest and SetRequest messages to the Agent. Upon receiving the requests from the Network Management Station, Agent will perform Read or Write operation according to the message types, generate and return the Response message to Network Management Station. On the other hand, Agent will send Trap message on its own initiative to the Network Management Station to report the events whenever the device encounters any abnormalities such as restart.

47.2 SNMP Versions and Supported MIB

To uniquely identify the management variables of a device in SNMP messages, SNMP adopts the hierarchical naming scheme to identify the managed objects. It is like a tree. A tree node represents a managed object, as shown in the figure below. Thus the object can be identified with the unique path starting from the root.

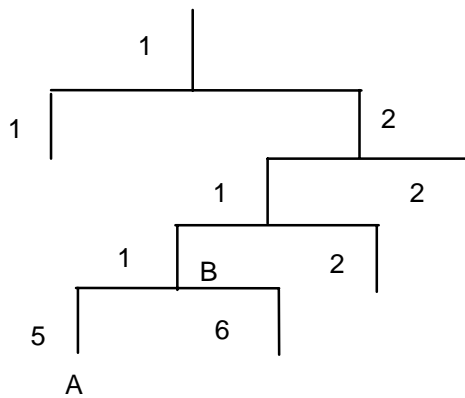


Figure 47-1 Architecture of the MIB tree

The MIB (Management Information Base) is used to describe the hierarchical architecture of the tree and it is the set defined by the standard variables of the monitored network device. In the above figure, the managed object B can be uniquely specified by a string of numbers {1.2.1.1}. The number string is the Object Identifier of the managed object.

The current SNMP Agent of Ethernet switch supports SNMP V1, V2C and V3. The MIBs supported are listed in the following table.

Table 47-1 MIBs supported by the Ethernet Switch

MIB attribute	MIB content	References
Public MIB	MIB II based on TCP/IP network device	RFC1213
	BRIDGE MIB	RFC1493
		RFC2675
	RIP MIB	RFC1724
	RMON MIB	RFC2819
	Ethernet MIB	RFC2665
	OSPF MIB	RFC1253
IF MIB	RFC1573	

MIB attribute	MIB content	References
Private MIB	DHCP MIB	—
	QACL MIB	—
	ADBM MIB	—
	RSTP MIB	—
	VLAN MIB	—
	Device management	—
	Interface management	—

47.3 Configuring SNMP

The following sections describe the SNMP configuration tasks.

- Setting Community Names
- Setting the System Information
- Enabling/Disabling SNMP Agent to Send Trap
- Setting the Destination Address of Trap
- Setting Lifetime of Trap Message
- Setting the Engine ID of a Local or Remote Device
- Setting/Deleting an SNMP Group
- Setting the Source Address of Trap
- Adding/Deleting a User to/from an SNMP Group
- Creating/Updating View Information or Deleting a View
- Setting the Size of the SNMP Packet Sent/Received by an Agent
- Disabling SNMP Agent

47.3.1 Setting Community Names

- SNMP V1 and SNMPV2C adopt the community name authentication scheme. SNMP Community is named with a character string, which is called community name. SNMP community name defines the relationship between SNMP manager and SNMP agent. The community name functions like a password, that is, it controls the access of the SNMP manager to the SNMP agent. You can choose to specify one or more community name-related features: Define MIB views of all the accessible MIB subsets.
- Define the read-only or read-write access mode of the community name to the MIB. The community with read-only authority can only query the device information, whereas the community with read-write authority can also configure the device.

Perform the following configuration in system view.

Table 47-2 Set community names

Operation	Command
Set the community name and the access authority	snmp-agent community { read write } <i>community-name</i> [[mib-view <i>view-name</i>] [acl <i>acl-list</i>]]
Remove the community name and the access authority	undo snmp-agent community <i>community-name</i>

47.3.2 Setting the System Information

System information includes the ID and the contact method of the administrator, the location of the Ethernet switch and the version of the SNMP.

The ID and the contact method of the administrator is a character string describing the contact information used for the system maintenance. Through this information, the device maintenance staffs can obtain the manufacturer information of the device so as to contact the manufacturer in case the device is in trouble. You can use the following command to set the contact information.

The location information of the Ethernet switch is a management variable of the system group in MIB, which represents the location of the managed device.

Perform the following configuration in system view.

Table 47-3 Set the system information

Operation	Command
Set the system information	snmp-agent sys-info { contact <i>sysContact</i> location <i>sysLocation</i> version { { v1 v2c v3 }* all } }
Restore the default information	undo snmp-agent sys-info { { contact location }* version { { v1 v2c v3 }* all } }

By default, the version is SNMPv3

47.3.3 Enabling/Disabling SNMP Agent to Send Trap

The managed device transmits trap without request to the Network Management Station to report some critical and urgent events (such as restart).

You can use the following commands to enable or disable the managed device to send trap message.

Perform the following configuration in corresponding views.

Table 47-4 Enable/disable SNMP Agent to send Trap

Operation	Command
Enable the sending of trap(system view)	snmp-agent trap enable [standard [authentication] [coldstart] [linkdown] [linkup] bgp [backwardtransition] [established] vrrp [authfailure newmaster]]
Disable the sending of trap(system view)	undo snmp-agent trap enable [standard [authentication] [coldstart] [linkdown] [linkup] bgp [backwardtransition] [established] vrrp [authfailure newmaster]]
Enable the switch ports to send SNMP trap messages (Ethernet port view or VLAN interface view)	enable snmp trap updown
Disable the switch port to send SNMP trap messages (Ethernet port view or VLAN interface view)	undo enable snmp trap updown

By default, the current port or VLAN interface sends trap messages.

47.3.4 Setting the Destination Address of Trap

You can use the following commands to set or delete the destination address of the trap.

Perform the following configuration in system view.

Table 47-5 Set the destination address of trap

Operation	Command
Set the destination address of trap	snmp-agent target-host trap address udp-domain <i>host-addr</i> [udp-port <i>udp-port-number</i>] params securityname <i>community-string</i> [v1 v2c v3 [authentication privacy]]
Delete the destination address of trap	undo snmp-agent target-host <i>host-addr</i> securityname <i>community-string</i>

47.3.5 Setting Lifetime of Trap Message

You can use the following command to set lifetime of Trap message. Trap message that exists longer than the set lifetime will be dropped.

Perform the following configuration in system view.

Table 47-6 Set the lifetime of Trap message

Operation	Command
Set lifetime of Trap message	snmp-agent trap life <i>seconds</i>
Restore lifetime of Trap message	undo snmp-agent trap life

By default, the lifetime of Trap message is 120 seconds.

47.3.6 Setting the Engine ID of a Local or Remote Device

You can use the following commands to set the engine ID of a local or remote device.

Perform the following configuration in system view.

Table 47-7 Set the engine ID of a local or remote device

Operation	Command
Set the engine ID of the device	snmp-agent local-engineid <i>engineid</i>
Restore the default engine ID of the device.	undo snmp-agent local-engineid

The engine ID of the device is in hexadecimal notation and has at least five characters, which can be IP address, MAC address or self-defined text. It defaults to the enterprise number + the device information.

47.3.7 Setting/Deleting an SNMP Group

You can use the following commands to set or delete an SNMP group.

Perform the following configuration in system view.

Table 47-8 Set/Delete an SNMP Group

Operation	Command
Set an SNMP 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-list</i>] 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-list</i>]
Delete an SNMP group	undo snmp-agent group { v1 v2c } <i>group-name</i> undo snmp-agent group v3 <i>group-name</i> [authentication privacy]

47.3.8 Setting the Source Address of Trap

You can use the following commands to set or remove the source address of the trap. Perform the following configuration in system view.

Table 47-9 Set the source address of trap

Operation	Command
Set the Source Address of Trap	snmp-agent trap source <i>interface-name</i> <i>interface-num</i>
Remove the source address of trap	undo snmp-agent trap source

Note:

Currently, this command takes effect only on the interfaces with vlan-interface type.

47.3.9 Adding/Deleting a User to/from an SNMP Group

You can use the following commands to add or delete a user to/from an SNMP group. Perform the following configuration in system view.

Table 47-10 Add/Delete a user to/from an SNMP group

Operation	Command
Add a user to an SNMP group.	snmp-agent usm-user { v1 v2c } <i>username</i> <i>groupname</i> [acl <i>acl-list</i>] snmp-agent usm-user v3 <i>username</i> <i>groupname</i> [authentication-mode { md5 sha } <i>authpassstring</i> [privacy-mode { des56 <i>privpassstring</i> }]] [acl <i>acl-list</i>]
Delete a user from an SNMP group.	undo snmp-agent usm-user { v1 v2c } <i>username</i> <i>groupname</i> undo snmp-agent usm-user v3 <i>username</i> <i>groupname</i> { local engineid <i>engine-id</i> }

You must first configure the SNMP engine ID before configuring the remote user for an agent, because the engine ID is required during the authentication. If you forget to configure the engine ID before adding a user, the operation of adding this user will fail.

For SNMP V1 and V2c, this operation is adding a new community name, while for SNMP V3, this operation is adding a user for an SNMP group.

47.3.10 Creating/Updating View Information or Deleting a View

You can specify the view to control the access to the MIB by SNMP manager. You can use either the predefined views or the self-defined views. You can use the following commands to create, update the information of views or delete a view.

Perform the following configuration in system view.

Table 47-11 Create/Update view information or delete a view

Operation	Command
Create/Update information	<code>snmp-agent mib-view { included excluded } view-name oid-tree</code>
Delete a view	<code>undo snmp-agent mib-view view-name</code>

47.3.11 Setting the Size of the SNMP Packet Sent/Received by an Agent

You can use the following commands to set the size of SNMP packet sent/received by an agent.

Perform the following configuration in system view.

Table 47-12 Set the size of the SNMP packet sent/received by an agent

Operation	Command
Set the size of the SNMP packet sent/received by an agent	<code>snmp-agent packet max-size byte-count</code>
Restore the default size of the SNMP packet sent/received by an agent	<code>undo snmp-agent packet max-size</code>

The agent can receive/send the SNMP packets of the sizes ranging from 484 to 17940, measured in bytes. By default, the size of an SNMP packet is 1500 bytes.

47.3.12 Disabling SNMP Agent

To disable SNMP Agent, perform the following configuration in system view.

Table 47-13 Disable snmp agent

Operation	Command
Disable snmp agent	<code>undo snmp-agent</code>

If users disable NMP Agent, it will be enabled whatever **snmp-agent** command is configured thereafter.

47.4 Displaying and Debugging SNMP

After the above configuration, execute the **display** command in any view to display the running of the SNMP configuration, and to verify the effect of the configuration.

Table 47-14 Display and debug SNMP

Operation	Command
Display the statistics information about SNMP packets	display snmp-agent statistics
Display the engine ID of the active device	display snmp-agent { local-engineid remote-engineid }
Display the group name, the security mode, the states for all types of views, and the storage mode of each group of the switch.	display snmp-agent group [<i>group-name</i>]
Display SNMP user information in the group user table	display snmp-agent usm-user [engineid <i>engineid</i> group <i>groupname</i> username <i>username</i>]*
Display the current community name	display snmp-agent community [read write]
Display the current MIB view	display snmp-agent mib-view [exclude include { viewname <i>mib-view</i> }]
Display the contact character strings, location character strings, and the SNMP version of the system	display snmp-agent sys-info [contact location version]*

47.5 SNMP Configuration Example

I. Network requirements

Network Management Station and the Ethernet switch are connected through the Ethernet. The IP address of Network Management Station is 129.102.149.23 and that of the VLAN interface on the switch is 129.102.0.1. Perform the following configurations on the switch: setting the community name and access authority, administrator ID, contact and switch location, and enabling the switch to send trap packets.

II. Network diagram

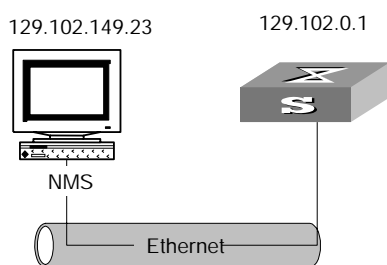


Figure 47-2 Network diagram for SNMP configuration

III. Configuration procedure

Enter the system view.

```
<SW8800> system-view
```

Set the community name, group and user.

```
[SW8800] snmp-agent sys-info version all
[SW8800] snmp-agent community write public
[SW8800] snmp-agent mib include internet 1.3.6.1
[SW8800] snmp-agent group v3 managev3group write internet
[SW8800] snmp-agent usm v3 managev3user managev3group
```

Set the VLAN interface 2 as the interface for network management. Add port GigabitEthernet 2/1/3 to the VLAN 2. This port will be used for network management. Set the IP address of VLAN interface 2 as 129.102.0.1.

```
[SW8800] vlan 2
[SW8800-vlan2] port gigabitethernet 2/1/3
[SW8800-vlan2] interface vlan 2
[SW8800-Vlan-interface2] ip address 129.102.0.1 255.255.0.0
```

Enable SNMP agent to send the trap to network management station whose IP address is 129.102.149.23. The SNMP community is public.

```
[SW8800] snmp-agent trap enable standard authentication
[SW8800] snmp-agent trap enable standard coldstart
[SW8800] snmp-agent trap enable standard linkup
[SW8800] snmp-agent trap enable standard linkdown
[SW8800] snmp-agent target-host trap address udp-domain 129.102.149.23
udp-port 5000 params securityname public
```

IV. Configure network management system

The PC on which the network management resides requires for login configuration. As for Mib-Browser, the login configuration is as follows: SNMPV1/V2 logs in using the default community name public, and the SNMPV3 logs in using managev3user.

Users can query and configure the Ethernet switch through the network management system. For details, see the manuals for the network management products.

Chapter 48 RMON Configuration

48.1 RMON Overview

Remote Network Monitoring (RMON) is a type of IETF-defined MIB. It is the most important enhancement to the MIB II standard. It mainly used for monitoring the data traffic on a segment and even on a whole network. It is one of the widely used Network Management standards by far.

RMON is implemented fully based on the SNMP architecture (which is one of its outstanding advantages) and compatible with the existing SNMP framework, and therefore it is unnecessary to adjust the protocol. RMON includes NMS and the Agent running on the network devices. On the network monitor or detector, RMON Agent tracks and accounts different traffic information on the segment connected to its port, such as the total number of packets on a segment in a certain period of time or that of the correct packets sent to a host. ROMN helps the SNMP monitor the remote network device more actively and effectively, which provides a highly efficient means for the monitoring of the subnet operations. RMON can reduce the communication traffic between the NMS and the agent, thus facilitates an effective management over the large interconnected networks.

RMON allows multiple monitors. It can collect data in two ways.

- One is to collect data with a special RMON probe. NMS directly obtains the management information from the RMON probe and controls the network resource. In this way, it can obtain all the information of RMON MIB
- Another way is to implant the RMON Agent directly into the network devices (for example router, switch and HUB), so that the devices become network facilities with RMON probe function. RMON NMS uses the basic SNMP commands to exchange data information with SNMP Agent and collect NM information. However, limited by the device resources, normally, not all the data of RMON MIB can be obtained with this method. In most cases, only four groups of information can be collected. The four groups include alarm information, event information, history information and statistics information.

The Ethernet Switch implements RMON in the second method by far. With the RMON-supported SNMP Agent running on the network monitor, NMS can obtain such information as the overall traffic of the segment connected to the managed network device port, the error statistics and performance statistics, thereby implementing the management (generally remote management) over the network.

48.2 Configuring RMON

Note:

Before configuring RMON, you must ensure that the SNMP agent is properly configured. See Chapter 50 SSH Terminal Service for the SNMP agent configuration.

The following sections describe the RMON configuration tasks.

- Adding/Deleting an Entry to/from the Event Table
- Adding/Deleting an Entry to/from the Alarm Table
- Adding/Deleting an Entry to/from the Extended RMON Alarm Table
- Adding/Deleting an Entry to/from the History Control Table
- Adding/Deleting an Entry to/from the Statistics Table

48.2.1 Adding/Deleting an Entry to/from the Event Table

RMON event management defines the event ID and the handling of the event.

You can handle the event in the following ways:

- Keeping logs
- Sending the trap messages to NMS
- Keeping logs and sending the trap messages to NMS

Perform the following configuration in system view.

Table 48-1 Add/delete an entry to/from the event table

Operation	Command
Add an entry to the event table	<code>rmon event event-entry [description string] { log trap trap-community log-trap log-trapcommunity none } [owner rmon-station]</code>
Delete an entry from the event table	<code>undo rmon event event-entry</code>

48.2.2 Adding/Deleting an Entry to/from the Alarm Table

RMON alarm management can monitor the specified alarm variables such as the statistics on a port. When a value of the monitored data exceeds the defined threshold, an alarm event will be generated. And then the events are handled according to the definition, which is decided in the event management.

Note:

Before adding an entry to the alarm table, you need to define the event referenced in the alarm table by using the **rmon event** command.

Perform the following configuration in system view.

Table 48-2 Add/delete an entry to/from the alarm table

Operation	Command
Add an entry to the alarm table	rmon alarm <i>entry-number</i> <i>alarm-variable</i> <i>sampling-time</i> { delta absolute } rising-threshold <i>threshold-value1</i> <i>event-entry1</i> falling-threshold <i>threshold-value2</i> <i>event-entry2</i> [owner <i>text</i>]
Delete an entry from the alarm table	undo rmon alarm <i>entry-number</i>

After you defined the alarm entry, the system then processes the entry in the following way:

- 1) Sampling the defined alarm-variable according to the time interval *sampling-time* that you have set
- 2) Comparing the sampled value with the configured threshold and handling them in the way described in the following table

Table 48-3 Handling the alarm entry

Case	Processing
The sampled value is greater than the configured upper limit <i>threshold-value1</i>	The defined event <i>event-entry1</i> is triggered
The sampled value is less than the configured lower limit <i>threshold-value2</i>	The defined event <i>event-entry2</i> is triggered

48.2.3 Adding/Deleting an Entry to/from the Extended RMON Alarm Table

You can use the command to add/delete an entry to/from the extended RMON alarm table. The extended alarm entry performs mathematical operation to the sampled value of the alarm variable, and then the result will be compared with the configured threshold to implementing the alarm function.

Note:

Before adding extended alarm entry, you need to define the referenced event in the extended alarm entry by using the **rmon event** command.

You can define up to 50 prialarm entries.

Perform the following configuration in system view.

Table 48-4 Add/delete an entry to/from the extended RMON alarm table

Operation	Command
Add an entry to the extended RMON alarm table	rmon prialarm <i>entry-number</i> <i>alarm-var</i> [<i>alarm-des</i>] <i>sampling-timer</i> { delta absolute changeratio } rising-threshold <i>threshold-value1</i> <i>event-entry1</i> falling-threshold <i>threshold-value2</i> <i>event-entry2</i> entrytype { forever cycle } [cycle-period] [owner <i>text</i>]
Delete an entry from the extended RMON alarm table	undo rmon prialarm <i>entry-number</i>

After you define the extended alarm entry, the system processes the entry in the following way:

- 1) Sampling the defined prialarm-formula according to the time interval *sampling-time* that you have set
- 2) Performing the operation to the sampled value according to the defined formula *prialarm-formula*
- 3) Comparing the result with the configured threshold and handling them in the way described in the following table

Table 48-5 Handling the extended alarm entry

Case	Processing
The result is greater than the configured upper limit <i>threshold-value1</i>	The defined event <i>event-entry1</i> is triggered
The result is less than the configured lower limit <i>threshold-value2</i>	The defined event <i>event-entry2</i> is triggered

48.2.4 Adding/Deleting an Entry to/from the History Control Table

The history data management helps you set the history data collection, periodical data collection and storage of the specified ports. The sampling information includes the utilization ratio, error counts and total number of packets.

You can use the following commands to add/delete an entry to/from the history control table.

Perform the following configuration in Ethernet port view.

Table 48-6 Add/delete an entry to/from the history control table

Operation	Command
Add an entry to the history control table.	rmon history <i>entry-number</i> buckets <i>number</i> interval <i>sampling-interval</i> [owner <i>text-string</i>]

Delete an entry from the history control table.	undo rmon history <i>entry-number</i>
---	--

History control entry calculates various data at the sampling time interval. You can use the **display rmon history** command to view the information of the history control entry.

48.2.5 Adding/Deleting an Entry to/from the Statistics Table

The RMON statistics management concerns the port usage monitoring and error statistics when using the ports. The statistics include collision, CRC and queuing, undersize packets or oversize packets, timeout transmission, fragments, broadcast, multicast and unicast messages and the usage ratio of bandwidth.

You can use the following commands to add/delete an entry to/from the statistics table.

Perform the following configuration in Ethernet port view.

Table 48-7 Add/delete an entry to/from the statistics table

Operation	Command
Add an entry to the statistics table	rmon statistics <i>entry-number</i> [owner <i>text-string</i>]
Delete an entry from the statistics table	undo rmon statistics <i>entry-number</i>

Statistics entry calculates the accumulated information starting from the time defined by an event. You can use the **display rmon history** command to view the information of the statistics entry.

48.3 Displaying and Debugging RMON

After the above configuration, execute the **display** command in any view to display the running of the RMON configuration, and to verify the effect of the configuration.

Table 48-8 Display and debug RMON

Operation	Command
Display the RMON statistics	display rmon statistics [<i>port-num</i>]
Display the history information of RMON	display rmon history [<i>port-num</i>]
Display the alarm information of RMON	display rmon alarm [<i>alarm-table-entry</i>]
Display the extended alarm information of RMON	display rmon prialarm [<i>prialarm-table-entry</i>]
Display the RMON event	display rmon event [<i>event-table-entry</i>]

Display the event log of RMON	display rmon eventlog [<i>event-number</i>]
-------------------------------	---

48.4 RMON Configuration Example

I. Network requirements

Set an entry in RMON Ethernet statistics table for the Ethernet port performance, which is convenient for network administrators' query.

II. Network diagram

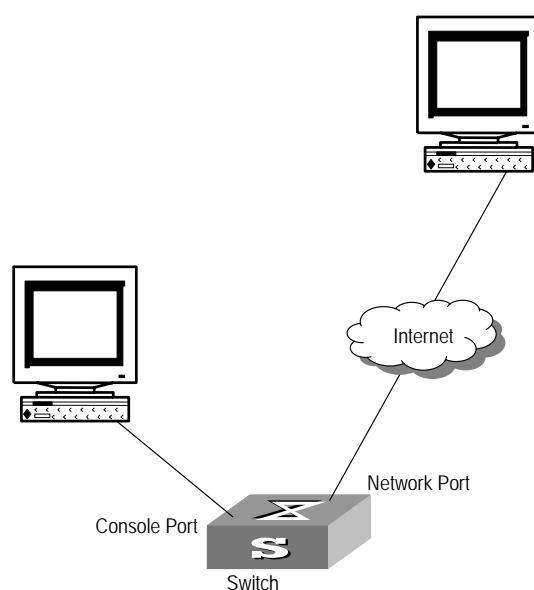


Figure 48-1 Network diagram for RMON configuration

III. Configuration procedure

Configure RMON.

```
[SW8800-Ethernet2/1/1] rmon statistics 1 owner 3Com-rmon
```

View the configurations in user view.

```
<SW8800> display rmon statistics Ethernet 2/1/1
```

```
Statistics entry 1 owned by 3Com-rmon is VALID.
```

```
Gathers statistics of interface Ethernet2/1/1. Received:
octets           : 270149,   packets           : 1954
broadcast packets :1570   ,   multicast packets:365
undersized packets :0     ,   oversized packets:0
fragments packets :0     ,   jabbers packets :0
CRC alignment errors:0   ,   collisions       :0
Dropped packet events (due to lack of resources):0
```


Packets received according to length (in octets):

64	:644	,	65-127	:518	,	128-255	:688
256-511	:101	,	512-1023	:3	,	1024-1518	:0

Chapter 49 NTP Configuration

49.1 Brief Introduction to NTP

49.1.1 NTP Functions

As the network topology gets more and more complex, it becomes important to synchronize the clocks of the equipment on the whole network. Network Time Protocol (NTP) is the TCP/IP that advertises the accurate time throughout the network.

NTP ensures the consistency of the following applications:

- For the increment backup between the backup server and client, NTP ensures the clock synchronization between the two systems.
- For multiple systems that coordinate to process a complex event, NTP ensures them to reference the same clock and guarantee the right order of the event.
- Guarantee the normal operation of the inter-system (Remote Procedure Call).
- Record for an application when a user logs in to a system, a file is modified, or some other operation is performed.

49.1.2 Basic Operating Principle of NTP

The following figure illustrates the basic operating principle of NTP:

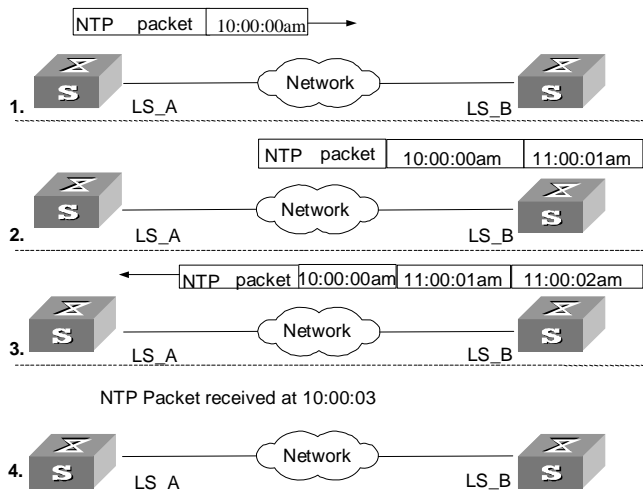


Figure 49-1 Basic operating principle of NTP

In the figure above, Ethernet Switch A and Ethernet Switch B are connected through the Ethernet port. They have independent system clocks. Before implement automatic clock synchronization on both switches, we assume that:

- Before synchronizing the system clocks on Ethernet Switch A and B, the clock on Ethernet Switch A is set to 10:00:00am, and that on B is set to 11:00:00am.
- Ethernet Switch B serves as an NTP time server. That is, Ethernet Switch A synchronizes the local clock with the clock of B.
- It takes 1 second to transmit a data packet from either A or B to the opposite end.

The system clocks are synchronized as follows:

- Ethernet Switch A sends an NTP packet to Ethernet Switch B. The packet carries the timestamp 10:00:00am (T_1) that tells when it left Ethernet Switch A.
- When the NTP packet arrives at Ethernet Switch B, Ethernet Switch B adds a local timestamp 11:00:01am (T_2) to it.
- When the NTP packet leaves Ethernet Switch B, Ethernet Switch B adds another local timestamp 11:00:02am (T_3) to it.
- When Ethernet Switch A receives the acknowledgement packet, it adds a new timestamp 10:00:03am (T_4) to it.

Now Ethernet Switch A collects enough information to calculate the following two important parameters:

- The delay for a round trip of an NTP packet traveling between the Switch A and B:
Delay = $(T_4 - T_1) - (T_3 - T_2)$.
- Offset of Ethernet Switch A clock relative to Ethernet Switch B clock: offset = $((T_2 - T_1) + (T_4 - T_3)) / 2$.

In this way, Ethernet Switch A uses the above information to set the local clock and synchronize it with the clock on Ethernet Switch B.

The operating principle of NTP is briefly introduced above. For details, refer to RFC1305.

49.2 NTP Configuration

NTP is used for time synchronization throughout a network. The following sections describe the NTP configuration tasks.

- Configuring NTP Operating Mode
- Configuring NTP ID Authentication
- Setting NTP Authentication Key
- Setting Specified Key as Reliable
- Designating an Interface to Transmit NTP Messages
- Setting NTP Master Clock
- Setting Authority to Access a Local Ethernet Switch
- Setting Maximum Local Sessions

49.2.1 Configuring NTP Operating Mode

You can set the NTP operating mode of an Ethernet Switch according to its location in the network and the network structure. For example, you can set a remote server as the

time server of the local equipment. In this case the local Ethernet Switch works as an NTP client. If you set a remote server as a peer of the local Ethernet Switch, the local equipment operates in symmetric active mode. If you configure an interface on the local Ethernet Switch to transmit NTP broadcast packets, the local Ethernet Switch will operate in broadcast mode. If you configure an interface on the local Ethernet Switch to receive NTP broadcast packets, the local Ethernet Switch will operate in broadcast client mode. If you configure an interface on the local Ethernet Switch to transmit NTP multicast packets, the local Ethernet Switch will operate in multicast mode. Or you may also configure an interface on the local Ethernet Switch to receive NTP multicast packets, the local Ethernet Switch will operate in multicast client mode.

- Configure NTP server mode
- Configure NTP peer mode
- Configure NTP broadcast server mode
- Configure NTP broadcast client mode
- Configure NTP multicast server mode
- Configure NTP multicast client mode

I. Configuring NTP Server Mode

Set a remote server whose ip address is *ip-address* as the local time server. *ip-address* specifies a host address other than a broadcast, multicast or reference clock IP address. In this case, the local Ethernet Switch operates in client mode. In this mode, only the local client synchronizes its clock with the clock of the remote server, while the reverse synchronization will not happen.

Perform the following configuration in system view.

Table 49-1 Configure NTP time server

Operation	Command
Configure NTP time server	ntp-service unicast-server <i>ip-address</i> [version <i>number</i> authentication-keyid <i>keyid</i> source-interface { <i>interface-name</i> <i>interface-type</i> <i>interface-number</i> } priority]*
Cancel NTP server mode	undo ntp-service unicast-server <i>ip-address</i>

NTP version number *number* ranges from 1 to 3 and defaults to 3; the authentication key ID *keyid* ranges from 1 to 4294967295; *interface-name* or *interface-type* *interface-number* specifies an interface, from which the source IP address of the NTP packets sent from the local Ethernet Switch to the time server will be taken, the interface can be VLAN interface and Loopback interface; **priority** indicates the time server will be the first choice.

II. Configuring NTP Peer Mode

Set a remote server whose ip address is *ip-address* as the peer of the local equipment. In this case, the local equipment operates in symmetric active mode. *ip-address* specifies a host address other than a broadcast, multicast or reference clock IP address. In this mode, both the local Ethernet Switch and the remote server can synchronize their clocks with the clock of opposite end.

Perform the following configuration in system view.

Table 49-2 Configure NTP peer mode

Operation	Command
Configure NTP peer mode	ntp-service unicast-peer <i>ip-address</i> [version <i>number</i> authentication-key <i>keyid</i> source-interface { <i>interface-name</i> <i>interface-type</i> <i>interface-number</i> } priority]*
Cancel NTP peer mode	undo ntp-service unicast-peer <i>ip-address</i>

NTP version number *number* ranges from 1 to 3 and defaults to 3; the authentication key ID *keyid* ranges from 1 to 4294967295; *interface-name* or *interface-type interface-number* specifies an interface, from which the source IP address of the NTP packets sent from the local Ethernet Switch to the peer will be taken, the interface can be VLAN interface and Loopback interface; **priority** indicates the peer will be the first choice for time server.

III. Configuring NTP Broadcast Server Mode

Designate an interface on the local Ethernet Switch to transmit NTP broadcast packets. In this case, the local equipment operates in broadcast mode and serves as a broadcast server to broadcast messages to its clients regularly.

Perform the following configuration in VLAN interface view.

Table 49-3 Configure NTP broadcast server mode

Operation	Command
Configure NTP broadcast server mode	ntp-service broadcast-server [authentication-keyid <i>keyid</i> version <i>number</i>]*
Cancel NTP broadcast server mode	undo ntp-service broadcast-server

NTP version number *number* ranges from 1 to 3 and defaults to 3; the authentication key ID *keyid* ranges from 1 to 4294967295; This command can only be configured on the interface where the NTP broadcast packets will be transmitted.

IV. Configuring NTP Broadcast Client Mode

Designate an interface on the local Ethernet Switch to receive NTP broadcast messages and operate in broadcast client mode. The local Ethernet Switch listens to the broadcast from the server. When it receives the first broadcast packets, it starts a brief client/server mode to switch messages with a remote server for estimating the network delay. Thereafter, the local Ethernet Switch enters broadcast client mode and continues listening to the broadcast and synchronizes the local clock according to the arrived broadcast message.

Perform the following configuration in VLAN interface view.

Table 49-4 Configure NTP broadcast client mode

Operation	Command
Configure NTP broadcast client mode	ntp-service broadcast-client
Disable NTP broadcast client mode	undo ntp-service broadcast-client

This command can only be configured on the interface where the NTP broadcast packets will be received.

V. Configuring NTP Multicast Server Mode

Designate an interface on the local Ethernet Switch to transmit NTP multicast packets. In this case, the local equipment operates in multicast mode and serves as a multicast server to multicast messages to its clients regularly.

Perform the following configuration in VLAN interface view.

Table 49-5 Configure NTP multicast server mode

Operation	Command
Configure NTP multicast server mode	ntp-service multicast-server [<i>ip-address</i>] [authentication-keyid <i>keyid</i> ttl <i>tll-number</i> version <i>number</i>]*
Cancel NTP multicast server mode	undo ntp-service multicast-server [<i>ip-address</i>]

NTP version number *number* ranges from 1 to 3 and defaults to 3; the authentication key ID *keyid* ranges from 1 to 4294967295; *tll-number* of the multicast packets ranges from 1 to 255; And the multicast IP address defaults to 224.0.1.1. Actually, for the Switch 8800, you can set 224.0.1.1 as the multicast IP address only.

This command can only be configured on the interface where the NTP multicast packet will be transmitted.

VI. Configuring NTP Multicast Client Mode

Designate an interface on the local Ethernet Switch to receive NTP multicast messages and operate in multicast client mode. The local Ethernet Switch listens to the multicast from the server. When it receives the first multicast packets, it starts a brief client/server mode to switch messages with a remote server for estimating the network delay. Thereafter, the local Ethernet Switch enters multicast client mode and continues listening to the multicast and synchronizes the local clock by the arrived multicast message.

Perform the following configuration in VLAN interface view.

Table 49-6 Configure NTP multicast client mode

Operation	Command
Configure NTP multicast client mode	ntp-service multicast-client [<i>ip-address</i>]
Cancel NTP multicast client mode	undo ntp-service multicast-client

Multicast IP address *ip-address* defaults to 224.0.1.1; this command can only be configured on the interface where the NTP multicast packets will be received. Actually, for the Switch 8800, you can set 224.0.1.1 as the multicast IP address only.

49.2.2 Configuring NTP ID Authentication

Enable NTP authentication, set MD5 authentication key, and specify the reliable key. A client will synchronize itself by a server only if the server can provide a reliable key.

Perform the following configuration in system view.

Table 49-7 Configure NTP authentication

Operation	Command
Enable NTP authentication	ntp-service authentication enable
Disable NTP authentication	undo ntp-service authentication enable

49.2.3 Setting NTP Authentication Key

This configuration task is to set NTP authentication key.

Perform the following configuration in system view.

Table 49-8 Configure NTP authentication key

Operation	Command
Configure NTP authentication key	ntp-service authentication-keyid <i>number</i> authentication-mode md5 <i>value</i>
Remove NTP authentication key	undo ntp-service authentication-keyid <i>number</i>

Key number *number* ranges from 1 to 4294967295; the key *value* contains 1 to 32 ASCII characters.

49.2.4 Setting Specified Key as Reliable

This configuration task is to set the specified key as reliable.

Perform the following configuration in system view.

Table 49-9 Set the specified key as reliable

Operation	Command
Set the specified key as reliable	ntp-service reliable authentication-keyid <i>key-number</i>
Cancel the specified reliable key.	undo ntp-service reliable authentication-keyid <i>key-number</i>

Key number *key-number* ranges from 1 to 4294967295

49.2.5 Designating an Interface to Transmit NTP Messages

If the local equipment is configured to transmit all the NTP messages, these packets will have the same source IP address, which is taken from the IP address of the designated interface.

Perform the following configuration in system view.

Table 49-10 Designate an interface to transmit NTP messages

Operation	Command
Designate an interface to transmit NTP messages	ntp-service source-interface { <i>interface-name</i> <i>interface-type interface-number</i> }
Cancel the interface to transmit NTP messages	undo ntp-service source-interface

An interface is specified by *interface-name* or *interface-type interface-number*, and the interface can be VLAN interface and Loopback interface at present. The source address of the packets will be taken from the IP address of the interface. If the

ntp-service unicast-server or **ntp-service unicast-peer** command also designates a transmitting interface, use the one designated by them.

49.2.6 Setting NTP Master Clock

This configuration task is to set the external reference clock or the local clock as the NTP master clock.

Perform the following configuration in system view.

Table 49-11 Set the external reference clock or the local clock as the NTP master clock

Operation	Command
Set the external reference clock or the local clock as the NTP master clock.	ntp-service refclock-master [<i>ip-address</i>] [<i>stratum</i>]
Cancel the NTP master clock settings	undo ntp-service refclock-master [<i>ip-address</i>]

ip-address specifies the IP address 127.127.1.u of a reference clock, in which u ranges from 0 to 3. *stratum* specifies how many stratum the local clock belongs to and ranges from 1 to 15.

The IP address defaults 127.127.1.0, and the stratum defaults to 8.

49.2.7 Setting Authority to Access a Local Ethernet Switch

Set authority to access the NTP services on a local Ethernet Switch. This is a basic and brief security measure, compared to authentication. An access request will be matched with **peer**, **server**, **server only**, and **query only** in an ascending order of the limitation. The first matched authority will be given.

Perform the following configuration in system view.

Table 49-12 Set authority to access a local Ethernet switch

Operation	Command
Set authority to access a local Ethernet switch	ntp-service access { query synchronization server peer } <i>acl-number</i>
Cancel settings of the authority to access a local Ethernet switch	undo ntp-service access { query synchronization server peer }

IP address ACL number is specified through the *acl-number* parameter and ranges from 2000 to 2999. The meanings of other authority levels are as follows:

query: Allow control query for the local NTP service only.

synchronization: Allow request for local NTP time service only.

server: Allow local NTP time service request and control query. However, the local clock will not be synchronized by a remote server.

peer: Allow local NTP time service request and control query. And the local clock will also be synchronized by a remote server.

49.2.8 Setting Maximum Local Sessions

This configuration task is to set the maximum local sessions.

Perform the following configurations in system view.

Table 49-13 Set the maximum local sessions

Operation	Command
Set the maximum local sessions	ntp-service max-dynamic-sessions <i>number</i>
Resume the maximum number of local sessions	undo ntp-service max-dynamic-sessions

number specifies the maximum number of local sessions, ranges from 0 to 100, and defaults to 100.

49.3 Displaying and Debugging NTP

After completing the above configurations, you can use the **display** command to show how NTP runs and verify the configurations according to the outputs.

In user view, you can use the **debugging** command to debug NTP.

Table 49-14 Display and debug NTP

Operation	Command
Display the status of NTP service	display ntp-service status
Display the status of sessions maintained by NTP service	display ntp-service sessions [verbose]
Display the brief information about every NTP time server on the way from the local equipment to the reference clock source.	display ntp-service trace
Enable NTP debugging	debugging ntp-service

49.4 NTP Configuration Example

49.4.1 Configuring a NTP Server

I. Network requirements

On SW88001, set local clock as the NTP master clock at stratum 2. On SW88002, configure SW88001 as the time server in server mode and set the local equipment as in client mode. (Note: SW88001 supports to configure the local clock as the master clock)

II. Network diagram

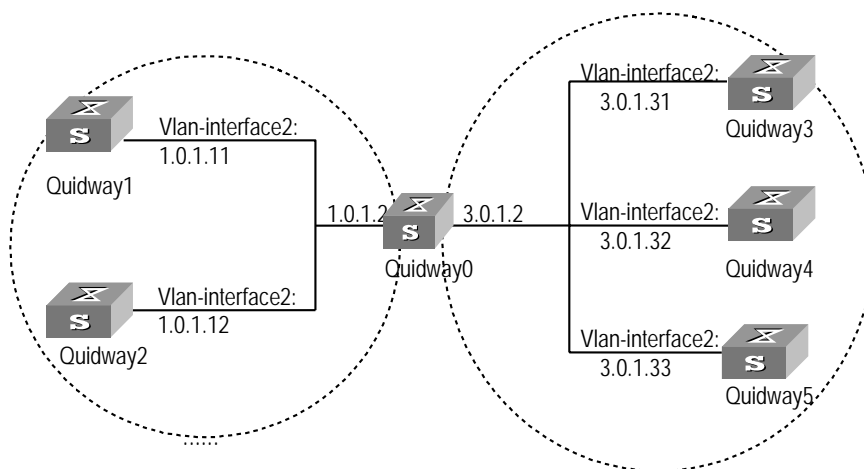


Figure 49-2 Typical NTP configuration network diagram

III. Configuration procedure

Configure Ethernet Switch SW88001:

Enter system view.

```
<SW88001> system-view
```

Set the local clock as the NTP master clock at stratum 2.

```
[SW88001] ntp-service refclock-master 2
```

Configure Ethernet Switch SW88002:

Enter system view.

```
<SW88002> system-view
```

Set SW88001 as the NTP server.

```
[SW88002] ntp-service unicast-server 1.0.1.11
```

The above examples synchronized SW88002 by SW88001. Before the synchronization, the SW88002 is shown in the following status:

```
[SW88002] 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: 2^17
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)

```

After the synchronization, SW88002 turns into the following status:

```

[SW88002] display ntp-service status
Clock status: synchronized
Clock stratum: 3
Reference clock ID: 1.0.1.11
Nominal frequency: 60.0002 Hz
Actual frequency: 60.0002 Hz
Clock precision: 2^17
Clock offset: -9.8258 ms
Root delay: 27.10 ms
Root dispersion: 49.29 ms
Peer dispersion: 10.94 ms
Reference time: 19:21:32.287 UTC Oct 24 2004(C5267F3C.49A61E0C)

```

By this time, SW88002 has been synchronized by SW88001 and is at stratum 3, higher than SW88001 by 1.

Display the sessions of SW88002 and you will see SW88002 has been connected with SW88001.

```

[SW88002] display ntp-service sessions
source          reference    stra reach poll  now offset  delay disper
*****
[12345]1.0.1.11  LOCAL(0)   3    377  64  16  -0.4  0.0  0.9
note: 1 source(master),2 source(peer),3 selected,4 candidate,5 configured

```

49.4.2 NTP Peer Configuration Example

I. Network requirements

On SW88003, set local clock as the NTP master clock at stratum 2. On SW88002, configure SW88001 as the time server in server mode and set the local equipment as in client mode. At the same time, SW88005 sets SW88004 as its peer. (Note: SW88003 supports to configure the local clock as the master clock)

II. Network diagram

See Figure 7-2.

III. Configuration procedure

Configure Ethernet Switch SW88003:

Enter system view.

```
<SW88003> system-view
```

Set the local clock as the NTP master clock at stratum 2.

```
[SW88003] ntp-service refclock-master 2
```

Configure Ethernet Switch SW88004:

Enter system view.

```
<SW88004> system-view
```

Set SW88001 as the NTP server at stratum 3 after synchronization.

```
[SW88004] ntp-service unicast-server 3.0.1.31
```

Configure Ethernet Switch SW88005: (SW88004 has been synchronized by SW88003)

Enter system view.

```
<SW88005> system-view
```

Set the local clock as the NTP master clock at stratum 1.

```
[SW88005] ntp-service refclock-master 1
```

After performing local synchronization, set SW88004 as a peer.

```
[SW88005] ntp-service unicast-peer 3.0.1.32
```

The above examples configure SW88004 and SW88005 as peers and configure SW88005 as in active peer mode and SW88004 in passive peer mode. Since SW88005 is at stratum 1 and SW88004 is at stratum 3, synchronize SW88004 by SW88005.

After synchronization, SW88004 status is shown as follows:

```
[SW88004] display ntp-service status
```

```
Clock status: synchronized
```

```
  Clock stratum: 2
```

```
  Reference clock ID: 3.0.1.31
```

```
  Nominal frequency: 60.0002 Hz
```

```
  Actual frequency: 60.0002 Hz
```

```
  Clock precision: 2^17
```

```
  Clock offset: -9.8258 ms
```

```
  Root delay: 27.10 ms
```

```
  Root dispersion: 49.29 ms
```

```
  Peer dispersion: 10.94 ms
```

```
Reference time: 19:21:32.287 UTC Oct 24 2004(C5267F3C.49A61E0C)
```

By this time, SW88004 has been synchronized by SW88005 and it is at stratum 2, or higher than SW88005 by 1.

Display the sessions of SW88004 and you will see SW88004 has been connected with SW88005.

```
[Quidwa4] display ntp-service sessions
source          reference  stra reach poll  now offset  delay disper
*****
[12345]3.0.1.33  LOCAL(0)   2    377  64  16    0.0    0.0    0.9
note: 1 source(master),2 source(peer),3 selected,4 candidate,5 configured
```

49.4.3 Configure NTP Broadcast Mode

I. Network requirements

On SW88003, set local clock as the NTP master clock at stratum 2 and configure to broadcast packets from Vlan-interface2. Configure SW88004 and SW88001 to listen to the broadcast from their Vlan-interface2 respectively. (Note: SW88003 supports to configure the local clock as the master clock)

II. Network diagram

See Figure 7-2.

III. Configuration procedure

Configure Ethernet Switch SW88003:

Enter system view.

```
<SW88003> system-view
```

Set the local clock as the NTP master clock at stratum 2.

```
[SW88003] ntp-service refclock-master 2
```

Enter Vlan-interface2 view.

```
[SW88003] interface vlan-interface 2
```

Set it as broadcast server.

```
[SW88003-Vlan-Interface2] ntp-service broadcast-server
```

Configure Ethernet Switch SW88004:

Enter system view.

```
<SW88004> system-view
```

Enter Vlan-interface2 view.

```
[SW88004] interface vlan-interface 2
```

```
[SW88004-Vlan-Interface2] ntp-service broadcast-client
```

Configure Ethernet Switch SW88001:

Enter system view.

```
<SW88001> system-view
```

Enter Vlan-interface2 view.

```
[SW88001] interface vlan-interface 2
[SW88001-Vlan-Interface2] ntp-service broadcast-client
```

The above examples configured SW88004 and SW88001 to listen to the broadcast through Vlan-interface2, SW88003 to broadcast packets from Vlan-interface2. Since SW88001 and SW88003 are not located on the same segment, they cannot receive any broadcast packets from SW88003, while SW88004 is synchronized by SW88003 after receiving its broadcast packet.

After the synchronization, you can find the state of SW88004 as follows:

```
[SW88004] display ntp-service status
clock status: synchronized
clock stratum: 3
reference clock ID: LOCAL(0)
nominal frequency: 100.0000 Hz
actual frequency: 100.0000 Hz
clock precision: 2^17
clock offset: 0.0000 ms
root delay: 0.00 ms
root dispersion: 10.94 ms
peer dispersion: 10.00 ms
reference time: 20:54:25.156 UTC Mar 7 2002(C0325201.2811A112)
```

By this time, SW88004 has been synchronized by SW88003 and it is at stratum 3, higher than SW88003 by 1.

Display the status of SW88004 sessions and you will see SW88004 has been connected to SW88003.

```
[SW88002] display ntp-service sessions
source           reference      stra reach poll  now offset  delay disper
[12345]127.127.1.0 LOCAL(0)      7  377  64  57  0.0  0.0  1.0
[5]1.0.1.11      LOCAL(0)      3   0   64  -   0.0  0.0  0.0
[5]128.108.22.44 0.0.0.0      16  0   64  -   0.0  0.0  0.0
note: 1 source(master),2 source(peer),3 selected,4 candidate,5 configured
```

49.4.4 Configure NTP Multicast Mode

I. Network requirements

SW88003 sets the local clock as the master clock at stratum 2 and multicast packets from Vlan-interface2. Set SW88004 and SW88001 to receive multicast messages from

their respective Vlan-interface2. (Note: SW88003 supports to configure the local clock as the master clock)

II. Network diagram

See Figure 7-2.

1) Configuration procedure

Configure Ethernet Switch SW88003:

Enter system view.

```
<SW88003> system-view
```

Set the local clock as a master NTP clock at stratum 2.

```
[SW88003] ntp-service refclock-master 2
```

Enter Vlan-interface2 view.

```
[SW88003] interface vlan-interface 2
```

Set it as a multicast server.

```
[SW88003-Vlan-Interface2] ntp-service multicast-server
```

Configure Ethernet Switch SW88004:

Enter system view.

```
<SW88004> system-view
```

Enter Vlan-interface2 view.

```
[SW88004] interface vlan-interface 2
```

Enable multicast client mode.

```
[SW88004-Vlan-Interface2] ntp-service multicast-client
```

Configure Ethernet Switch SW88001:

Enter system view.

```
<SW88001> system-view
```

Enter Vlan-interface2 view.

```
[SW88001] interface vlan-interface 2
```

Enable multicast client mode.

```
[SW88001-Vlan-Interface2] ntp-service multicast-client
```

The above examples configure SW88004 and SW88001 to receive multicast messages from Vlan-interface2, SW88003 multicast messages from Vlan-interface2. Since SW88001 and SW88003 are not located on the same segments, SW88001 cannot receive the multicast packets from SW88003, while SW88004 is synchronized by SW88003 after receiving the multicast packet.

49.4.5 Configure Authentication-Enabled NTP Server Mode

I. Network requirements

SW88001 sets the local clock as the NTP master clock at stratum 2. SW88002 sets SW88001 as its time server in server mode and itself in client mode and enables authentication. (Note: SW88001 supports to configure the local clock as the master clock)

II. Network diagram

See Figure 7-2.

III. Configuration procedure

Configure Ethernet Switch SW88001:

Enter system view.

```
<SW88001> system-view
```

Set the local clock as the master NTP clock at stratum 2.

```
[SW88001] ntp-service refclock-master 2
```

Configure Ethernet Switch SW88002:

Enter system view.

```
<SW88002> system-view
```

Set SW88001 as time server.

```
[SW88002] ntp-service unicast-server 1.0.1.11
```

Enable authentication.

```
[SW88002] ntp-service authentication enable
```

Set the key.

```
[SW88002] ntp-service authentication-keyid 42 authentication-mode md5  
aNiceKey
```

Set the key as reliable.

```
[SW88002] ntp-service reliable authentication-keyid 42  
[Qudiway2] ntp-service unicast-server 1.0.1.11 authentication-keyid 42
```

The above examples synchronized SW88002 by SW88001. Since SW88001 has not been enabled authentication, it cannot synchronize SW88002. And now let us do the following additional configurations on SW88001 :

Enable authentication.

```
[SW88001] ntp-service authentication enable
```

Set the key.

```
[SW88001] ntp-service authentication-keyid 42 authentication-mode md5  
aNiceKey
```

Configure the key as reliable.

```
[SW88001] ntp-service reliable authentication-keyid 42
```

Chapter 50 SSH Terminal Service

50.1 SSH Terminal Service

50.1.1 SSH Overview

This chapter introduces the secure shell (SSH) feature. When a user telnets to the switch from an insecure network, the SSH feature can provide secure information and powerful authentication functionality, thereby protecting the switch from attacks such as IP address spoofing and clear text password interception attacks.

The switch can act as either SSH server or SSH client. When used as an SSH server, the switch supports multiple connections with SSH clients; when used as an SSH client, the switch supports SSH connections with the SSH server-enabled switch, UNIX hosts, and so on.

Currently, the switch supports SSH 2.0.

Figure 50-1 and Figure 50-2 illustrate two methods for establishing an SSH channel between a client and the server:

- Connect through a LAN
- Connect through a WAN

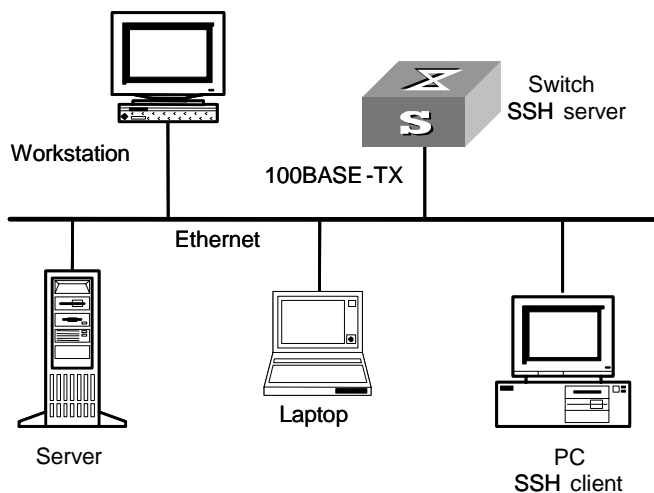


Figure 50-1 Establish an SSH channel through a LAN

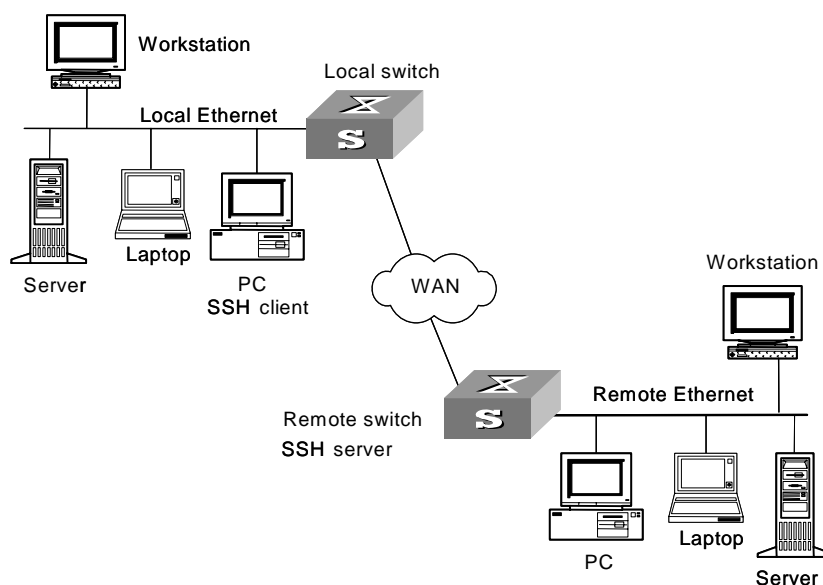


Figure 50-2 Establish an SSH channel through a WAN

To establish an SSH authentication secure connection, the server and the client must go through the following five phases:

- 1) Version number negotiation:
 - The client sends a TCP connection request.
 - After the TCP connection is established, the server and the client negotiate the version number.
 - If the negotiation succeeds, the key algorithm negotiation phase starts; otherwise, the server tears down the TCP connection.
- 2) Key algorithm negotiation:
 - The server generates a RSA key pair randomly, and sends the public key in the key pair to the client.
 - The client uses the public key from the server and a random number generated locally (in length of eight bytes) as parameters to calculate the session key.
 - Using the public key from the server, the client encrypts the random number for calculating the session key and sends the result to the server.
 - Using the local private key, the server decrypts the data sent by the client and obtains the random number used by the client.
 - The server uses the public key and the random number from the client as parameters to calculate the session key with the same algorithm as on the client. The resulting key is 16 bytes long.

On completion of the above steps, the server and the client obtains the same session key. During the session, both ends use the same session key to perform encryption and decryption, thereby guaranteeing the security of data transfer.

- 3) Authentication mode negotiation:
 - The client sends its username information to the server.

- The server initiates a procedure to authenticate the user. If the server is configured not to authenticate the user, the process proceeds to session request phase directly.
- The client employs an authentication mode to authenticate the server till the authentication succeeds or the server tears down the connection because of timeout.

Note:

SSH provides two authentication modes: password authentication and RSA authentication.

1) Password authentication procedure:

- The client sends the username and password to the server;
- The server compares the username and password sent from the client with the local configuration. If it finds an exact match, the authentication succeeds.

2) RSA authentication procedure:

- The server configures an RSA public key for the client;
- The client sends its RSA public key member module to the server;
- The server performs validity authentication on the member module. If the authentication succeeds, the server generates a random number, encrypts it using the RSA public key from the client, and sends the encrypted information back to the client;
- Both the server and the client uses the random number and the session ID with the length of 16 characters as parameters to calculate the authentication data;
- The client sends the authentication data it generates to the server;
- The server compares the authentication data from the client with that locally calculated. If they match, the authentication succeeds.

-
- 4) Session request: If the authentication succeeds, the client sends a session request to the server. When the server has successfully processed the request, SSH enters the interactive session phase.
- 5) Interactive session: The client and the server exchange data till the session is over.

50.1.2 SSH Server Configuration

The following table describes the SSH server configuration tasks.

Table 50-1 SSH2.0 Configuration tasks

Num	Item	Command	Description
1	Entering system view	<SW8800> system-view	-

Num	Item	Command	Description
2	Entering VTY type of user interface view	[SW8800] user-interface vty X X	-
3	Configure the protocol supported by current user interface	[SW8800-ui-vtyX-X] protocol inbound { all ssh telnet }	Optional
4	Returning to system view	[SW8800-ui-vtyX-X] quit	-
5	Generating a local RSA key pair	[SW8800] rsa local-key-pair create	Required
	Destroying a local RSA key pair	[SW8800] rsa local-key-pair destroy	
6	Configure the SSH user authentication mode	[SW8800] ssh user username authentication-type { password rsa password-publickey all }	Required By default, users are unable to log in.
7	Configure the updating cycle of the server key	[SW8800] ssh server rekey-interval hours	Optional By default, the system does not update the server key.
8	Configure the SSH authentication timeout	[SW8800] ssh server timeout seconds	Optional By default, it is 60 seconds.
9	Configure the number of SSH authentication retries	[SW8800] ssh server authentication-retries times	Optional By default, it is three times.
10	Enter public key view	[SW8800] rsa peer-public-key key-name	Required
	Generate RSA key using key generator tool	See Generating the Client Public Key.	
11	Entering public key edit view to edit the key	[SW8800-rsa-public-key] public-key-code begin	Required
12	Exiting public key edit view	[SW8800-rsa-public-key] public-key-code end	Required
13	Specifying the public key for an SSH user	[SW8800] ssh user username assign rsa-key keyname	Required
14	Configure first-authentication SSH server	[SW8800] ssh client first-time enable	Optional By default, the system does not perform the first authentication.

Num	Item	Command	Description
15	Configure the SSH compatibility mode	[SW8800] ssh server compatible_ssh1x enable	Optional By default, the server is compatible with the SSH1.x client.

I. Configuring the protocol the current user interface supports

Use this configuration task to specify the protocol the current user interface supports. Perform the following configuration in VTY user interface view.

Table 50-2 Configure the protocol the current user interface supports

Operation	Command
Configure the protocol the current user interface supports	protocol inbound { all pad ssh telnet }

By default, the system supports all protocols.



Caution:

- If the supported protocol configured in the user interface is SSH, make sure to configure the authentication mode for logging into the user interface to authentication-mode scheme (using AAA authentication mode).
- If the authentication mode is configured as **authentication-mode password** or **authentication-mode none**, the configuration of **protocol inbound ssh** will fail, and vice versa.

II. Generating or destroying an RSA key pair

Use this configuration task to generate or destroy an RSA key pair (including the host key and server key) of the server. The naming conventions for the keys are *switchname* + *host* and *switchname* + *server* respectively.

After this command is entered, the system prompts you to input the number of the key pair bits. Pay attention to the following:

- The host key and the server key must have a difference of at least 128 bits in length.

- The minimum and maximum lengths for the host key and the server key are 512 bits and 2048 bits respectively.

Perform the following configuration in system view.

Table 50-3 Generate an RSA key pair

Operation	Command
Generate an RSA key pair	rsa local-key-pair create
Destroy an RSA key pair	rsa local-key-pair destroy



Caution:

- Generating the RSA key pair of the server is the first step to perform after SSH login.
- This command needs to be performed only once; you need not re-perform it after rebooting the switch.
- If a key pair exists before the configuration, a prompt will appear asking if you want to replace it.

III. Configuring the user authentication mode

Use this configuration task to specify the authentication mode for an SSH user. You must specify an authentication mode for a new user; otherwise, the new user will not be able to log in.

Perform the following configuration in system view.

Table 50-4 Configure the authentication mode for an SSH user

Operation	Command
Configure the authentication mode for an SSH user	ssh user <i>username</i> authentication-type { password rsa password-publickey all }
Restore the default unable-to-login mode	undo ssh user <i>username</i> authentication-type

By default, no login authentication mode is specified, that is, SSH users are unable to log in.

IV. Configuring the updating cycle of the server key

Use this configuration task to set the updating cycle of the server key to secure the SSH connection in best effort.

Perform the following configuration in system view

Table 50-5 Configure the updating cycle of the server key

Operation	Command
Configure the updating cycle of the server key	ssh server rekey-interval <i>hours</i>
Cancel the updating cycle configuration	undo ssh server rekey-interval

By default, the system does not update the server key.

V. Configuring the authentication timeout

Use this configuration task to set the authentication timeout of SSH connections.

Perform the following configuration in system view.

Table 50-6 Set the SSH authentication timeout

Operation	Command
Set the SSH authentication timeout	ssh server timeout <i>seconds</i>
Restore the default SSH authentication timeout	undo ssh server timeout

By default, the authentication timeout is 60 seconds.

VI. Configuring the number of authentication retries

Use this configuration task to set the number of authentication retries an SSH user can request for a connection, thereby preventing illegal behaviors such as malicious guessing.

Perform the following configuration in system view.

Table 50-7 Configure the number of SSH authentication retries

Operation	Command
Configure the number of SSH authentication retries	ssh server authentication-retries <i>times</i>
Restore the default number of SSH authentication retries	undo ssh server authentication-retries

By default, the number of authentication retries is 3.

VII. Entering the public key view

Use this configuration command to enter the public key view and specify the name of the public key of the client.

Perform the first configuration in the following table in system view.

Table 50-8 Public key configuration

Operation	Command
Enter the public key view	rsa peer-public-key <i>key-name</i>
Exit the public view and return to the system view	peer-public-key end

Note:

The configuration commands are applicable to the environments where the server employs RSA authentication on SSH users. If the server adopts password authentication on SSH users, these configurations are not necessary.

VIII. Entering the public key edit view

After entering the public key view by the **rsa peer-public-key** command, you can use the **public-key-code begin** command to enter the public key edit view and input the public key of the client.

When inputting the public key, you may type spaces between the characters (the system will delete the spaces automatically), or press <Enter> and then continue to input the key. Note that the public key must be a hexadecimal string coded in the public key format.

Perform the following configuration in public key view.


Table 50-9 Enter the public key edit view

Operation	Command
Enter the public key edit view	public-key-code begin

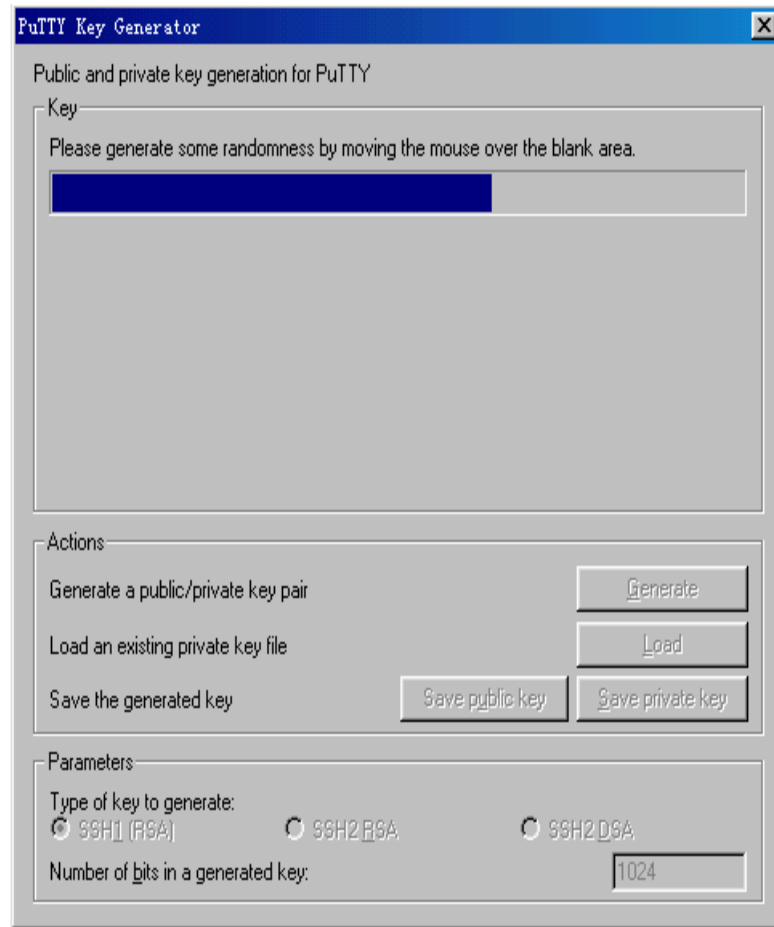
IX. Generating the Client Public Key

The client public key is generated using the PuTTY Key Generator application. Perform the following procedure to generate the key.

Table 50-10 Generate the Client Public Key

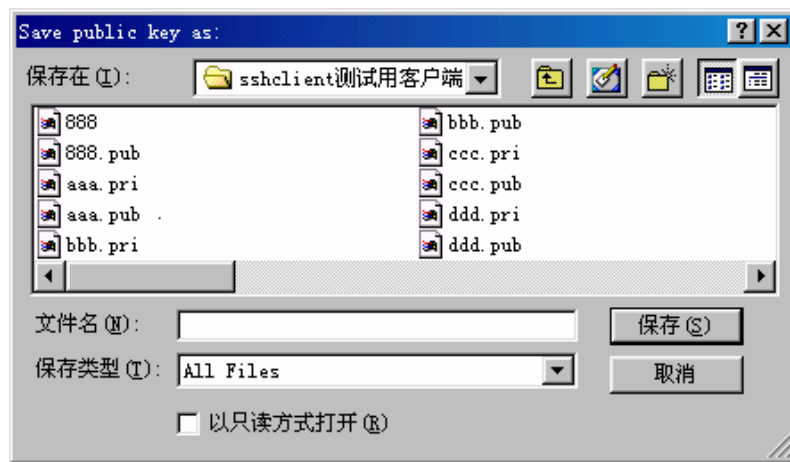
Operation	Command
Run the PuTTY Key Generator application	 puttygen.exe

While the Generator is running, move your mouse over the blank area of the window.

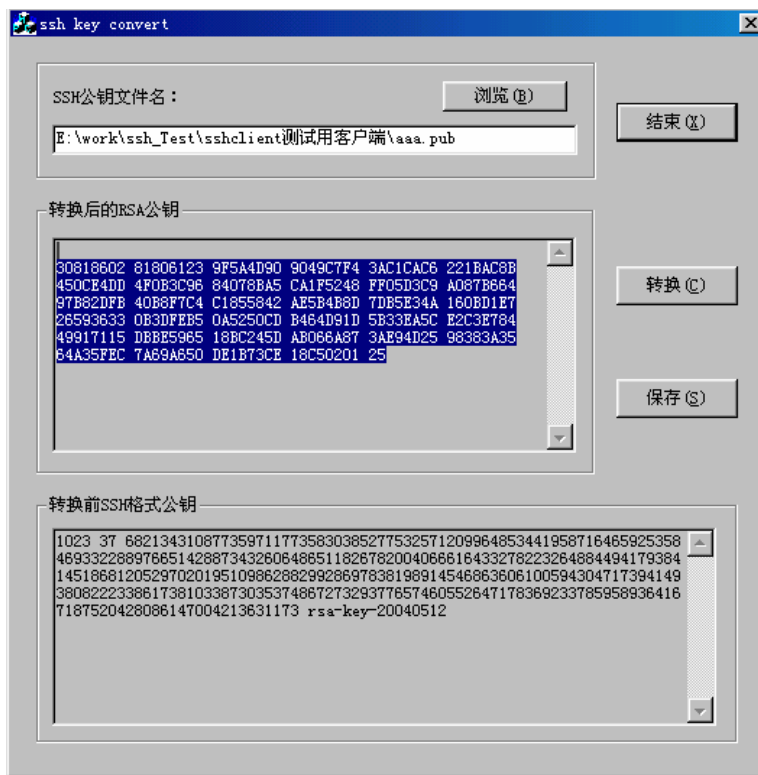


Save the pair of keys as publickey and privatekey.

File names are **aaa.pub** and **aaa.pri**



Convert the file **aaa.pub** into key configuration data in Hex.



Convert the converted result into the CLI of the switch unit

```
[SW8800-rsa-key-code]
[SW8800-rsa-key-code]30818602 81806123 9F5A4D90 9049C7F4 3AC1CAC6 221BAC8B
[SW8800-rsa-key-code]450CE4DD 4F0B3C96 84078BA5 CA1F5248 FF05D3C9 A087B664
[SW8800-rsa-key-code]97B82DFB 40B8F7C4 C1855842 AE5B4B8D 7DB5E34A 160BD1E7
[SW8800-rsa-key-code]26593633 0B3DFEB5 0A5250CD B464D91D 5B33EA5C E2C3E784
[SW8800-rsa-key-code]49917115 DBBE5965 18BC245D AB066A87 3AE94D25 98383A35
[SW8800-rsa-key-code]64A35FEC 7A69A650 DE1B73CE 18C50201 25
[SW8800-rsa-key-code]
[SW8800-rsa-key-code]public-key-code end
[SW8800-rsa-public-key]peer-public-key ?
end
```

Exit from editing the peer public key	[SW8800-rsa-public-key]peer-public-key end
---------------------------------------	--

```
[SW8800]
[SW8800]dis rsa peer-public-key

=====
Key name: aaa
Key address:
=====
Key Code:
308186
028180
61239F5A 4D909049 C7F43AC1 CAC6221B AC8B450C E4DD4F0B 3C968407 8BA5CA1F
5248FF05 D3C9A087 B66497B8 2DFB40B8 F7C4C185 5842AE5B 4B8D7DB5 E34A160B
D1E72659 36330B3D FEB50A52 50CDB464 D91D5B33 EA5CE2C3 E7844991 7115DBBE
596518BC 245DAB06 6A873AE9 4D259838 3A3564A3 5FEC7A69 A650DE1B 73CE18C5
0201
25

[SW8800]
```

X. Exiting the public key edit view

Use this configuration task to return from the public key edit view to the public key view and save the input public key. Before saving the input public key, the system will check the validity of the key:

- If the public key string contains any illegal character, the configured key is invalid;
- If the configured key is valid, it will be saved to the public key list.

Perform the following configuration in public key edit view.

Table 50-11 Exit the public key edit view

Operation	Command
Exit the public key edit view	public-key-code end

XI. Specifying the public key for an SSH user

Use this configuration task to specify an existing public key for an SSH user.

Perform the following configuration in system view.

Table 50-12 Specify the public key for an SSH user

Operation	Command
Specify the public key for an SSH user	ssh user <i>username</i> assign rsa-key <i>keyname</i>
Cancel the corresponding relationship between the user and the public key	undo ssh user <i>username</i> assign rsa-key

XII. Configuring the server compatibility mode

Use this configuration task to set whether the server should be compatible with the SSH 1.x client.

Perform the following configuration in system view.

Table 50-13 Configure the compatibility mode

Operation	Command
Set the server to be compatible with the SSH 1.x client	ssh server compatible_ssh1x enable
Set the server to be incompatible with the SSH 1.x client	undo ssh server compatible_ssh1x

By default, the server is compatible with the SSH 1.x client.

50.1.3 SSH Client Configuration

The following sections describe the SSH client configuration tasks.

- Set to perform the first-time authentication on the SSH server to be accessed
- Specifying the public key of the server
- Configuring the first-time authentication of the server

I. Starting the SSH client

Use this configuration task to enable the the SSH client, establish the connection with the server, and carry out interactive session.

Perform the following configuration in system view.

Table 50-14 Start the SSH client

Operation	Command
Start the SSH client	ssh2 { <i>host-ip</i> <i>host-name</i> } [<i>port-num</i>] [prefer_kex { dh_group1 dh_exchange_group }] [prefer_ctos_cipher { des 3des aes128 }] [prefer_stoc_cipher { des 3des aes128 }] [prefer_ctos_hmac { sha1 sha1_96 md5 md5_96 }] [prefer_stoc_hmac { sha1 sha1_96 md5 md5_96 }]

II. Specifying the public key of the server

Use this configuration task to allocate a existent public key to the client.

Perform the following configuration in system view.

Table 50-15 Specify the public key of the server

Operation	Command
Specify the public key of the server	ssh client <i>server-ip</i> assign rsa-key <i>keyname</i>
Cancel the corresponding relationship between the server and the public key	undo ssh client <i>server-ip</i> assign rsa-key

III. Configuring the first-time authentication of the server

Use this configuration task to configure or cancel the first-time authentication of the server performed by the SSH client.

The first-time authentication means that when the SSH client accesses the server for the first time in the case that there is no local copy of the server's public key, the user can choose to proceed to access the server and save a local copy of the server's public key; when the client accesses the server next time, it uses the saved public key to authenticate the server.

Perform the following configuration in system view.

Table 50-16 Configure the first-time authentication of the server

Operation	Command
Configure the first-time authentication of the server	ssh client first-time enable
Cancel the first-time authentication of the server	undo ssh client first-time

By default, the client does not perform the first-time authentication.

50.1.4 Displaying and Debugging SSH

On completion of the above configurations, you can use the **display** command in any view to view the operation of the configured SSH and further verify the result of the configurations. You can also debug SSH by performing the **debugging** command in user view.

Table 50-17 Display information relevant to SSH

Operation	Command
Display the public key of the host key pair and the server key pair of the server	display rsa local-key-pair public
Display the public key of the specified RSA key pair of the client	display rsa peer-public-key [brief name <i>keyname</i>]
Display the SSH status information and session information	display ssh server { status session }
Display information about the SSH user	display ssh user-information [<i>username</i>]
Enable SSH debugging	debugging ssh server { vty index all }
Disable SSH debugging	undo debugging ssh server { vty index all }

50.1.5 SSH Server Configuration Example

I. Network requirements

As shown in Figure 50-3, a PC (SSH client) running SSH 2.0-enabled client software establishes a local connection with the switch (SSH server) to better guarantee the security of exchanged information.

II. Network diagram

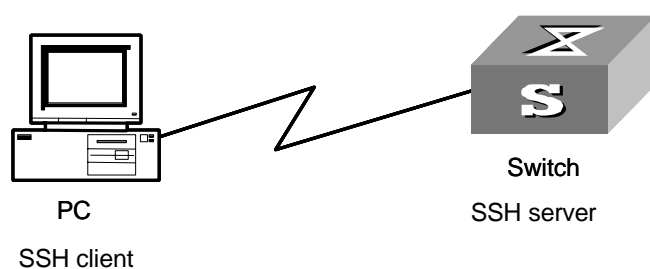


Figure 50-3 Network diagram for SSH server

III. Configuration procedure

- 1) Generate the RSA key.

```
[SW8800] rsa local-key-pair create
```

Note:

If the configuration for generating the local key has already been completed, skip this step.

- 2) Set the user login authentication mode.

The following shows the configuration methods for both password authentication and RSA public key authentication.

- Password authentication.

Create the local user client001, and set the authentication mode of the user interface to AAA.

```
[SW8800] user-interface vty 0 4
[SW8800-ui-vty0-4] authentication-mode scheme
```

Specify the login protocol for user client001 as SSH.

```
[SW8800-ui-vty0-4] protocol inbound ssh
[SW8800] local-user client001
[SW8800-luser-client001] password simple 3Com
[SW8800] ssh user client001 authentication-type password
```

Note:

You can use the default values for SSH authentication timeout and retries. After completing the above configurations, you can run the SSH 2.0-enabled client software on any other terminal connected with the switch and access the switch with the username client001 and password 3Com.

- RSA public key authentication.

Create the local user client001, and set the authentication mode of the user interface to AAA.

```
[SW8800] user-interface vty 0 4
[SW8800-ui-vty0-4] authentication-mode scheme
```

Specify the login protocol for user client002 as SSH.

```
[SW8800-ui-vty0-4] protocol inbound ssh
```

Set the authentication mode for the remote user on the switch to publickey.

```
[SW8800] ssh user client002 authentication-type publickey
```

Using the SSH 2.0-enabled client software, randomly generate an RSA key pair and send the public key to the server.

Configure the public key of the client. Refer to [Generating the Client Public Key](#) for details.

```
[SW8800] rsa peer-public-key SW8800002
[SW8800-rsa-public-key] public-key-code begin
[SW8800-rsa-key-code] 308186028180739A291ABDA704F5D93DC8FDF84C427463
[SW8800-rsa-key-code] 1991C164B0DF178C55FA833591C7D47D5381D09CE82913
[SW8800-rsa-key-code] D7EDF9C08511D83CA4ED2B30B809808EB0D1F52D045DE4
[SW8800-rsa-key-code] 0861B74A0E135523CCD74CAC61F8E58C452B2F3F2DA0DC
[SW8800-rsa-key-code] C48E3306367FE187BDD944018B3B69F3CBB0A573202C16
[SW8800-rsa-key-code] BB2FC1ACF3EC8F828D55A36F1CDDC4BB45504F020125
[SW8800-rsa-key-code] public-key-code end
[SW8800-rsa-public-key] peer-public-key end
[SW8800]
```

#Allocate an existent public key SW8800002 to user client002.

```
[SW8800] ssh user client002 assign rsa-key SW8800002
```

Start the SSH client software on the terminal preserving the RSA private key, and perform the corresponding configurations to establish the SSH connection.

50.1.6 SSH Client Configuration Example

I. Network requirements

As shown in Figure 50-4:

- Switch A is used as an SSH client.
- Switch B is used as the SSH server, and the IP address is 10.165.87.136.

II. Network diagram

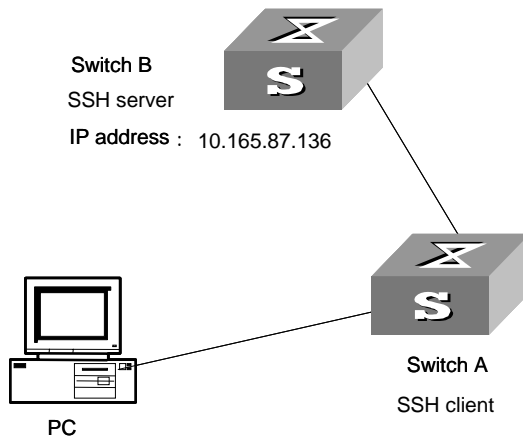


Figure 50-4 Network diagram for SSH client

III. Configuration procedure

Configure the client to perform the first-time authentication of the server.

- Employ password authentication mode, and start using the default encryption algorithm.

Log onto the SSH2 server with IP address 10.165.87.136.

```
[SW8800] ssh2 10.165.87.136
Please input the username:sshuser1
Trying 10.165.87.136
Press CTRL+K to abort
Connected to 10.165.87.136 ...
Enter password:
*****
*           All rights reserved (1997-2004)           *
*           Without the owner's prior written consent, *
*no decompiling or reverse-engineering shall be allowed.*
*****
<SW8800>
```

Configure the client to authenticate the server for the first time.

```
<SW8800> system-view
```

```
[SW8800] ssh client first-time enable
```

Access the remote server and perform operations.

- Employ RSA public key authentication mode, and start using the corresponding encryption algorithm configured.

```
[SW8800] ssh2 10.165.87.136 22 perfer_kex dh_group1 perfer_ctos_cipher des
perfer_stoc_cipher 3des perfer_ctos_hmac md5 perfer_stoc_hmac md5
```

```
Please input the username: client003
```

```
Trying 10.165.87.136...
```

```
Press CTRL+K to abort
```

```
Connected to 10.165.87.136...
```

```
The Server is not autherncated.Do you continue access it?(Y/N):y
```

```
Do you want to save the server's public key?(Y/N):y
```

```
*****
```

```
*           All rights reserved (1997-2004)           *
```

```
*           Without the owner's prior written consent,   *
```

```
*no decompiling or reverse-engineering shall be allowed.*
```

```
*****
```

```
<SW8800>
```

Configure the client to authenticate the server for the first time.

```
<SW8800> sys
```

```
[SW8800] ssh client first-time enable
```

Access the remote server and perform operations.

50.2 SFTP Service

50.2.1 SFTP Overview

Secure FTP is established on SSH connections, which makes remote users able to securely log in to the switch and perform file management and transfer operations such as system upgrade, and thereby providing higher security for data transfer. At the same time, since the switch can be used as a client, users can log in to remote devices to transfer files securely.

50.2.2 SFTP Server Configuration

SFTP server configuration tasks are described in this section:

- Configuring the service type to be used
- Starting the SFTP server

I. Configuring the service type to be used

Use this configuration task to set the SSH service type to be used.

Perform the following configuration in system view.

Table 50-18 Configure the service type to be used

Operation	Command
Configure the service type to be used	ssh user <i>username</i> service-type { telnet sftp all }
Restore the default service type	undo ssh user <i>username</i> service-type

By default, the service type is **telnet**.

II. Starting the SFTP server

Perform the following configuration in system view.

Table 50-19 Start the SFTP server

Operation	Command
Start the SFTP server	sftp server enable
Shut down the SFTP server	undo sftp server enable

By default, the SFTP server is shut down.

50.2.3 SFTP Client Configuration

The following table describes the SFTP client configuration tasks.

Table 50-20 SFTP client configuration tasks

Num	Item	Command	Description
1	Enter system view	<SW8800> system-view	-
2	Starting the SFTP client	[SW8800] sftp ipaddr [prefer_kex { dh_group1 dh_exchange_group }] [prefer_ctos_cipher { des 3des aes128 }] [prefer_stoc_cipher { des 3des aes128 }] [prefer_ctos_hmac { sha1 sha1_96 md5 md5_96 }] [prefer_stoc_hmac { sha1 sha1_96 md5 md5_96 }]	Required

Num	Item	Command	Description
3	Shut down the SFTP client	sftp-client> bye	Optional
		sftp-client> exit	
		sftp-client> quit	
4	SFTP directory operation	Chang the current directory sftp-client> cd [<i>remote-path</i>]	-
		Return to the upper directory sftp-client> cdup	
		Display the current directory sftp-client> pwd	
		Display the file list in the specified directory sftp-client> dir [<i>remote-path</i>]	
		sftp-client> ls [<i>remote-path</i>]	
Delete a directory on the server sftp-client> rmdir <i>remote-path</i>			
5	SFTP file operation	Change the name of the specified file on the server sftp-client> rename <i>oldname newname</i>	Optional
		Download a file from the remote server sftp-client> get <i>remote-file</i> [<i>local-file</i>]	
		Upload a local file to the remote server sftp-client> put <i>local-file</i> [<i>remote-file</i>]	
		Display the file list in the specified directory sftp-client> dir [<i>remote-path</i>]	
		sftp-client> ls [<i>remote-path</i>]	
		Delete a file from the server sftp-client> remove <i>remote-file</i> sftp-client> delete <i>remote-file</i>	
6	Command help on the client	sftp-client> help [<i>command</i>]	Optional

I. Starting the SFTP client

Use this configuration task to start the SFTP client program, establish a connection with the remote SFTP server, and enter the SFTP client view.

Perform the following configuration in system view.

Table 50-21 Start the SFTP client

Operation	Command
Start the SFTP client	<code>sftp ipaddr [prefer_kex { dh_group1 dh_exchange_group }] [prefer_ctos_cipher { des 3des aes128 }] [prefer_stoc_cipher { des 3des aes128 }] [prefer_ctos_hmac { sha1 sha1_96 md5 md5_96 }] [prefer_stoc_hmac { sha1 sha1_96 md5 md5_96 }]</code>

II. Shutting down the SFTP client

Use this configuration task to shut down the SFTP client program.

Perform the following configuration in SFTP client view.

Table 50-22 Shut down the SFTP client

Operation	Command
Shut down the SFTP client	<code>bye</code>
	<code>exit</code>
	<code>quit</code>

Note:

The three commands, **bye**, **exit**, and **quit**, have the same functionality. You can also use the **quit** command in port group view.

III. SFTP directory operations

As shown in Table 50-23, available SFTP directory operations include: change or display the current directory, create or delete a directory, display the specified file or directory.

Perform the following configuration in SFTP client view.

Table 50-23 SFTP directory operations

Operation	Command
Change the current directory	<code>cd remote-path</code>
Return to the upper directory	<code>cdup</code>
Display the current directory	<code>pwd</code>
Display the list of files in the specified	<code>dir [remote-path]</code>

Operation	Command
directory	ls [<i>remote-path</i>]
Create a new directory on the server	mkdir <i>remote-path</i>
Delete a directory from the server	rmdir <i>remote-path</i>

Note:

The **dir** command and the **ls** command have the same functionality.

IV. SFTP file operations

As shown in Table 50-24, available SFTP file operations include: change the name of a file, download a file, upload a file, display the list of files, and delete a file.

Perform the following configuration in SFTP user view.

Table 50-24 SFTP file operations

Operation	Command
Change the name of the specified file on the server	rename <i>old-name new-name</i>
Download a file from the remote server	get <i>remote-file</i> [<i>local-file</i>]
Upload a local file to the remote server	put <i>local-file</i> [<i>remote-file</i>]
Display the list of files in the specified directory	dir [<i>remote-path</i>]
	ls [<i>remote-path</i>]
Delete a file from the server	delete <i>remote-file</i>
	remove <i>remote-file</i>

Note:

- The **dir** command and the **ls** command have the same functionality.
 - The **delete** command and the **remove** command have the same functionality.
-

V. Displaying help information

Use this command to display command-relevant help information such as the format of the command, parameter configurations, and so on.

Perform the following configuration in SFTP client view.

Table 50-25 Display help information for client commands

Operation	Command
Display help information for client commands	help [<i>command-name</i>]

50.2.4 SFTP Configuration Example

I. Network requirements

As shown in Figure 50-5:

- Switch B is used as the SFTP server, and its IP address is 10.111.27.91;
- Switch B is used as the SFTP client;
- An SFTP user is configured with the username 8040 and password SW8800.

II. Network diagram

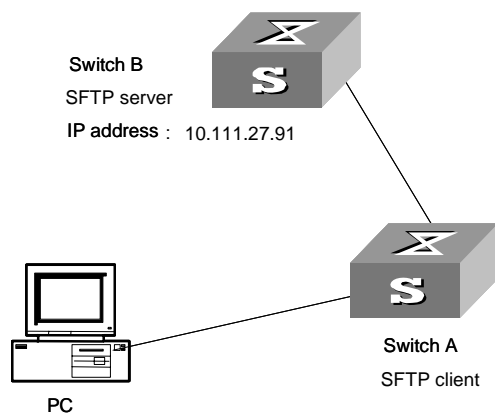


Figure 50-5 Network diagram for SFTP

III. Configuration procedure

- 1) Configure Switch B.

Start the SFTP server.

```
[SW8800] sftp-server enable
```

Specify the service type as SFTP.

```
[SW8800] ssh user 8040 service-type sftp
```

Set the authentication mode to password.

```
[SW8800] ssh user 8040 authentication-type password
```

- 2) Configure Switch A.

Configure the server with a public key whose name is the IP address of the server.

```
[SW8800] rsa peer-public-key 10.111.27.91
```



```
[SW8800-rsa-public-key] public-key-code begin
[SW8800-rsa-key-code] 308186028180739A291ABDA704F5D93DC8FDF84C427463
[SW8800-rsa-key-code] 1991C164B0DF178C55FA833591C7D47D5381D09CE82913
[SW8800-rsa-key-code] D7EDF9C08511D83CA4ED2B30B809808EB0D1F52D045DE4
[SW8800-rsa-key-code] 0861B74A0E135523CCD74CAC61F8E58C452B2F3F2DA0DC
[SW8800-rsa-key-code] C48E3306367FE187BDD944018B3B69F3CBB0A573202C16
[SW8800-rsa-key-code] BB2FC1ACF3EC8F828D55A36F1CDDC4BB45504F020125
[SW8800-rsa-key-code] public-key-code end
[SW8800-rsa-public-key] peer-public-key end
[SW8800] ssh client 10.111.27.91 assign rsa-key 10.111.27.91
```

Establish the SSH connection between the client and the server.

```
[SW8800] ssh2
Please input the username:8040
Trying
Press CTRL+K to abort
Connected to 10.111.27.91 ...
Enter password:SW8800
*****
*           All rights reserved (1997-2004)           *
*           Without the owner's prior written consent, *
*no decompiling or reverse-engineering shall be allowed.*
*****
<SW8800>
```

Establish a connection with the remote SFTP server and enter the SFTP client view.

```
<SW8800> sys
[SW8800] sftp 10.111.27.91
```

Display the current directory of the server, delete file z, and check if the directory has been deleted successfully.

```
sftp-client> dir
-rwxrwxrwx  1 noone  nogroup  1759 Aug 23 06:52 vrpcfg.cfg
-rwxrwxrwx  1 noone  nogroup    225 Aug 24 08:01 pubkey2
-rwxrwxrwx  1 noone  nogroup    283 Aug 24 07:39 pubkey1
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
Remove this File?(Y/N)
flash:/zy
File successfully Removed

sftp-client> dir
-rwxrwxrwx  1 noone  nogroup  1759 Aug 23 06:52 vrpcfg.cfg
```

```
-rwxrwxrwx 1 noone nogroup 225 Aug 24 08:01 pubkey2
-rwxrwxrwx 1 noone nogroup 283 Aug 24 07:39 pubkey1
drwxrwxrwx 1 noone nogroup 0 Sep 01 06:22 new
-rwxrwxrwx 1 noone nogroup 225 Sep 01 06:55 pub
```

Create a new directory new1, and check if the new directory has been created successfully.

```
sftp-client> mkdir new1
New path created
sftp-client> dir
-rwxrwxrwx 1 noone nogroup 1759 Aug 23 06:52 vrpcfg.cfg
-rwxrwxrwx 1 noone nogroup 225 Aug 24 08:01 pubkey2
-rwxrwxrwx 1 noone nogroup 283 Aug 24 07:39 pubkey1
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
```

Change the directory name new1 to new2, and check if the directory name has been changed successfully.

```
sftp-client> rename new1 new2
sftp-client> dir
-rwxrwxrwx 1 noone nogroup 1759 Aug 23 06:52 vrpcfg.cfg
-rwxrwxrwx 1 noone nogroup 225 Aug 24 08:01 pubkey2
-rwxrwxrwx 1 noone nogroup 283 Aug 24 07:39 pubkey1
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 file pubkey2 from the server to a local device, and change the file name to pu.

```
sftp-client> get pubkey2 pu
Downloading file successfully ended
```

Upload local file pu to the server, change the file name to puk, and check if the operations are successful.

```
sftp-client> put pu puk
Uploading file successfully ended
sftp-client> dir
-rwxrwxrwx 1 noone nogroup 1759 Aug 23 06:52 vrpcfg.cfg
-rwxrwxrwx 1 noone nogroup 225 Aug 24 08:01 pubkey2
-rwxrwxrwx 1 noone nogroup 283 Aug 24 07:39 pubkey1
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 pu
-rwxrwxrwx 1 noone nogroup 283 Sep 02 06:36 puk
```

```
sftp-client>
```

```
Exit SFTP.
```

```
sftp-client> quit
```

```
Bye
```

```
<SW8800>
```

Chapter 51 PoE Configuration

51.1 PoE Overview

51.1.1 PoE on the Switch

The Switch 8800 supports power-over-Ethernet (PoE). Equipped with external power supply and PoE-capable cards, Switch 8800s can provide 48 VDC power for remote powered devices (PDs, such as IP phones, WLAN APs, and Network cameras) through twisted pairs.

- The Switch 8800 supports IEEE802.3af standard. While they can also supply power to PDs noncompliant with the standard.

The power supply of the Switch 8800 is administered by the Fabric; each PoE card on the switch can be viewed as a power sourcing equipment (PSE), which administers the power supplying of all the ports on it independently.

The Switch 8800 can transmit data and supply power in the mean time through the signal lines (1, 3, 2, and 6) of the category-3/5 twisted pairs. Using converters, they can also supply power to the PDs that can be powered only through spare lines (4, 5, 7, and 8).

- The Switch 8800 supplies power through the Ethernet electrical ports on the service cards. Each service card can supply power to up to 48 remote devices at the maximum distance of 100 m (328 feet).
- The maximum power that can be supplied by each Ethernet port to its PD is 15.4 W.
- When supplying power to remote devices, the maximum total power that can be provided by the Switch 8800 is 4500 W (220 V)/2250 W (110V). The switch determines whether or not to supply power to the next remote PD it discovered depending on the total power it currently supply.

Note:

- When a remote PD is powered by an Switch 8800, the PD needs not have any external power supply.
 - If the remote PD has an external power supply, the Switch 8800 and the external power supply will be redundant with each other for the PD.
-

51.1.2 External PSE4500-A Power System

If PSE4500-A power system is taken as the external power supply of the switch, the power distribution is as follows:

- 1) Input voltage: 90 VAC to 160 VAC
 - One PSU (power supply unit) of the PSE4500-A power system can provide 1200 W of power.
- 2) Input voltage: 160 VAC to 264 VAC
 - One PSU of the PSE4500-A power system can provide 2500 W of power.
 - If the PSE4500-A power modules are in 2+1 redundancy, then each module provides a power of 1500 W.

51.2 PoE Configuration

The Switch 8800 can automatically detect any connected device that needs a remote power supply and feeds power to this device.

- Depending on your actual network requirement, you can set the maximum PoE power totally supplied by the switch through the command line.
- You can set the maximum PoE power supplied by a card through the command line.
- You can also control the PoE on each PoE port independently through the command line. The control includes: enabling/disabling the PoE feature, and setting the maximum PoE power, the PoE mode and the PoE priority on the port.

51.2.1 PoE Configuration Tasks

The following table describes the PoE configuration tasks on the Switch 8800.

Table 51-1 PoE configuration tasks on the Switch 8800

No	Item	Command	Description
1	Enter system view	system-view	—
2	Enter Ethernet port view	interface <i>interface-type</i> <i>interface-number</i>	As a result of this command, a port view prompt is displayed, which varies with the port type you selected.
3	Enable PoE on the port	poe enable	By default, PoE is disabled on a port.
4	Set the maximum PoE power supplied by the port	poe max-power <i>max-power</i>	You can set the maximum PoE power supplied by a port depending on the power of the actual PD. By default, the <i>max-power</i> is 16800 mW.
5	Set the PoE mode on the port	poe mode { signal spare auto }	The Switch 8800 supports only signal line PoE mode. By default, the PoE mode on a port is signal.
6	Set the PoE priority on the port	poe priority { critical high low }	You can set the PoE priority on a port depending on the practical situation. By default, the PoE priority on a port is low.
7	Display the PoE state of a specific or all ports of the switch	display poe interface [<i>interface-name</i> <i>interface-type</i> <i>interface-num</i>]	You can execute this command in any view. Executing the display poe interface command without any option displays the PoE status of all the ports.
8	Display the PoE power information of a specific or all ports of the switch	display poe interface power [<i>interface-name</i> <i>interface-type</i> <i>interface-num</i>]	You can execute this command in any view. Executing the display poe interface power command without any option displays the PoE power information about all the ports.
9	Display the PoE status and PoE power information of each card	display poe pse	You can execute this command in any view

To cancel the configurations, use the corresponding **undo** commands. For details about the parameters, refer to the *Command Manual*.

Note:

- When setting the maximum PoE power supplied by the switch, you must set it to a value greater than the total power that has been distributed to the cards. Otherwise, the command cannot be executed successfully. The maximum power ranges from 3000 mW to 16800 mW.
 - Before setting the maximum power supplied by a card, make sure the remaining power of the switch is no less than the full power of the card, and the power you can set for a card ranges from 37 W to 806 W.
 - The reserved power for a blank slot will be recycled automatically by the system if you insert a PoE-incapable card into the slot.
 - When a card is almost fully loaded and a new PD is added, the switch will respond to the PD according to the PoE priority set on the port.
 - The PoE priority of each port is based on its card. In other words, the switch cannot compare the priorities of ports on different cards.
-

51.3 Comprehensive Configuration Example

I. Network requirements

- Two PoE-capable cards are installed in slots 3 and 5 on a Switch 8800.
- GigabitEthernet3/1/1 through GigabitEthernet3/1/48 are connected with IP phones and GigabitEthernet5/1/1 through GigabitEthernet5/1/48 are connected with access point (AP) devices.
- The IP phones connected to GigabitEthernet3/1/23 and GigabitEthernet3/1/24 do not need PoE.
- GigabitEthernet3/1/48 is reserved for the use of network management, so it needs higher priority.
- Slot 3 is provided with 400 W power and slot 5 is provided with full power.
- The input power of the AP device connected to GigabitEthernet5/1/15 cannot be greater than 9000 mW.

II. Network diagram

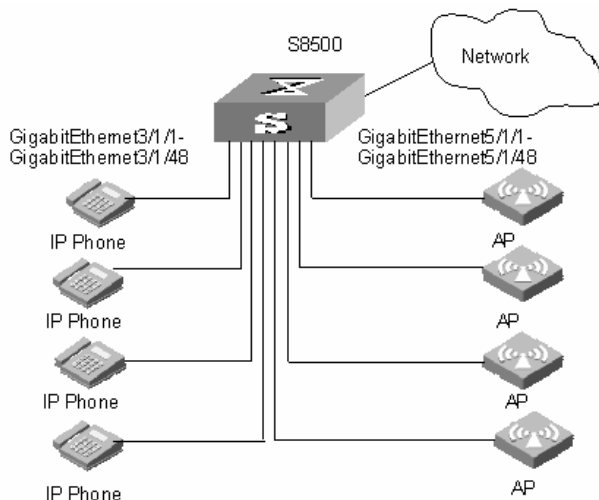


Figure 51-1 PoE remote power supplying

III. Configuration procedure

Set the maximum power to 400 W on the card in slot 3. By default, the power of each card is full, so the power on the card in slot 5 need not be configured.

```
[SW8800] poe max-power 400 slot 3
```

Enable PoE on the ports GigabitEthernet3/1/1 through GigabitEthernet3/1/48.

```
[SW8800-GigabitEthernet3/1/1] poe enable
```

```
[SW8800-GigabitEthernet3/1/2] poe enable
```

```
[SW8800-GigabitEthernet3/1/3] poe enable
```

Go on the configuration till the port GigabitEthernet3/1/48.

Enable PoE on the ports GigabitEthernet5/1/1 through GigabitEthernet5/1/48.

```
[SW8800-GigabitEthernet5/1/1] poe enable
```

```
[SW8800-GigabitEthernet5/1/2] poe enable
```

```
[SW8800-GigabitEthernet5/1/3] poe enable
```

Go on the configuration till the port GigabitEthernet5/1/48.

Set the PoE priority of the port GigabitEthernet3/1/48 to critical, the PD connected with GigabitEthernet3/1/48 will be powered in precedence on the premise that other ports' power supplying is not interrupted.

```
[SW8800-GigabitEthernet3/1/48] poe priority critical
```

Set the maximum PoE power on the GigabitEthernet5/1/15 port to 9000 mW.

```
[SW8800] interface GigabitEthernet5/1/15
```

```
[SW8800-GigabitEthernet5/1/15] poe max-power 9000
```


Chapter 52 PoE PSU Supervision Configuration

52.1 Introduction to PoE PSU Supervision

The PoE-capable Switch 8800 can monitor the external PoE PSUs through the power supervision module on the PoE external power system.

The PoE PSU supervision module enables you to:

- Set the alarm thresholds for the AC input voltages of the PoE PSUs.
- Set the alarm thresholds for the DC output voltages of the PoE PSUs.
- Query PSU information such as voltage and power.

52.2 AC Input Alarm Thresholds Configuration

You can set the AC input alarm thresholds for the PoE PSUs to enable the Switch 8800 to monitor the AC input voltages of the PSUs in real time through the PoE supervision module.

52.2.1 AC Input Alarm Thresholds Configuration Tasks

Table 52-1 AC input alarm thresholds configuration tasks

No	Item	Command	Description
1	Enter system view	system-view	—
2	Set the overvoltage alarm threshold of AC input (upper threshold) for the PoE PSUs	poe-power input-thresh upper string	Required, and the max voltage is 264.0 V.
3	Set the undervoltage alarm threshold of AC input (lower threshold) for the PoE PSUs	poe-power input-thresh lower string	Required, and the min voltage is 90.0 V.
4	Display the AC input state of each PoE PSU	display poe-power ac-input state	Optional, and you can execute this command in any view.

Note:

- You can set the thresholds to any appropriate values in the range, but make sure the lower threshold is less than the upper threshold.
 - For 220 VAC input, it is recommended to set the upper threshold to 264 V and the lower threshold to 181 V.
 - For 110 VAC input, it is recommended to set the upper threshold to 132 V and the lower threshold to 90 V.
-

52.2.2 AC Input Alarm Thresholds Configuration Example

I. Network requirements

- Set the overvoltage alarm threshold of AC input for PoE PSUs to 264.0 V.
- Set the undervoltage alarm threshold of AC input for PoE PSUs to 181.0 V.

II. Configuration procedure

Enter system view.

```
<SW8800> system-view
```

Set the overvoltage alarm threshold of AC input for PoE PSUs to 264.0 V.

```
[SW8800] poe-power input-thresh upper 264.0
```

Set the undervoltage alarm threshold of AC input for PoE PSUs to 181.0 V.

```
[SW8800] poe-power input-thresh lower 181.0
```

Display the information about the AC input for the PoE PSUs.

```
[SW8800] display poe-power ac-input state
```

52.3 DC Output Alarm Thresholds Configuration

You can set the DC output alarm thresholds for the PoE PSUs to enable the Switch 8800 to monitor the DC output voltages of the PSUs in real time through the PoE supervision module.

52.3.1 DC Output Alarm Thresholds Configuration Tasks

Table 52-2 DC output alarm thresholds configuration tasks

No	Operation	Command	Description
1	Enter system view	system-view	—
2	Set the overvoltage alarm threshold of DC output (upper threshold) for the PoE PSUs	poe-power output-thresh upper <i>string</i>	Required, and the range is 55.0 V to 57.0 V.
3	Set the undervoltage alarm threshold of DC output (lower threshold) for the PoE PSUs	poe-power output-thresh lower <i>string</i>	Required, and the range is 45.0 V to 47.0 V.
4	Display the DC output state of the PoE PSUs.	display poe-power dc-output state	Optional, and you can execute this command in any view.
5	Display the DC output voltage/current value of the PoE PSUs	display poe-power dc-output value	Optional, and you can execute this command in any view.

Note:

For both 220 VAC and 110 VAC input, it is recommended to set the upper threshold to 57.0 V and the lower threshold to 45.0 V.

52.3.2 DC Output Alarm Thresholds Configuration Example

I. Network requirements

- Set the overvoltage alarm threshold of DC output for the PoE PSUs to 57.0 V.
- Set the undervoltage alarm threshold of DC output for the PoE PSUs to 45.0 V.

II. Configuration procedure

Enter system view.

```
<SW8800> system-view
```

Set the overvoltage alarm threshold of DC output for the PoE PSUs to 57.0 V.

```
[SW8800] poe-power output-thresh upper 57.0
```

Set the undervoltage alarm threshold of DC output for the PoE PSUs to 45.0 V.

```
[SW8800] poe-power output-thresh lower 45.0
```

Display the DC output state of the PoE PSUs.

```
[SW8800] display poe-power dc-output state
```

Display the DC output voltage/current values of the PoE PSUs.

```
[SW8800] display poe-power dc-output value
```

52.4 Displaying PoE Supervision Information

After completing the above configurations, you can execute the **display** command in any view to query the PoE state of the switch. Then you can view the display output to check the effect of these configurations.

Table 52-3 Display PoE supervision information

No	Operation	Command	Description
1	Display the basic information about the PoE PSUs.	display supervision-module information	You can execute this command in any view.
2	Display detailed alarm information about the PoE PSUs.	display poe-power alarm	You can execute this command in any view.
3	Display the number and state of the switches of the PoE PSUs.	display poe-power switch state	You can execute this command in any view.

For details about display output, refer to the *Command Manual*.

52.5 PoE PSU Supervision Configuration Example

I. Network requirements

- Insert a PoE-capable card into slot 3 of the Switch 8800.
- Connect GigabitEthernet3/1/1 to GigabitEthernet3/1/48 to IP phones.
- Set the AC input and DC output alarm thresholds to appropriate values.

II. Network diagram

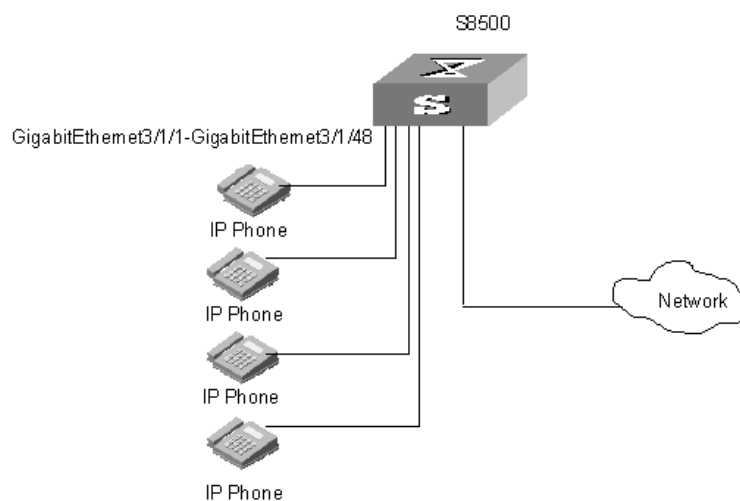


Figure 52-1 Network diagram for PoE supervision configuration

III. Configuration procedure

Enter system view.

```
<SW8800> system-view
```

Set the overvoltage alarm threshold of AC input for PoE PSUs to 264.0 V.

```
[SW8800] poe-power input-thresh upper 264.0
```

Set the undervoltage alarm threshold of AC input for PoE PSUs to 181.0 V.

```
[SW8800] poe-power input-thresh lower 181.0
```

Set the overvoltage alarm threshold of DC output for the PoE PSUs to 57.0 V.

```
[SW8800] poe-power output-thresh upper 57.0
```

Set the undervoltage alarm threshold of DC output for the PoE PSUs to 45.0 V.

```
[SW8800] poe-power output-thresh lower 45.0
```