

**CloudEngine 8800 and 6800 Series Switches
V200R023C00**

Configuration Guide - MPLS

Issue	01
Date	2023-11-15



Copyright © Huawei Technologies Co., Ltd. 2023. All rights reserved.

No part of this document may be reproduced or transmitted in any form or by any means without prior written consent of Huawei Technologies Co., Ltd.

Trademarks and Permissions



HUAWEI and other Huawei trademarks are trademarks of Huawei Technologies Co., Ltd.

All other trademarks and trade names mentioned in this document are the property of their respective holders.

Notice

The purchased products, services and features are stipulated by the contract made between Huawei and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

Huawei Technologies Co., Ltd.

Address: Huawei Industrial Base
Bantian, Longgang
Shenzhen 518129
People's Republic of China

Website: <https://e.huawei.com>

Security Declaration

Product Life Cycle

Huawei's regulations on product life cycle are subject to the Product End of Life Policy. For details about the policy, see the following website: <https://support.huawei.com/ecolumnsweb/en/warranty-policy>

Vulnerability

Huawei's regulations on product vulnerability management are subject to "Vul. Response Process". For details about the policy, see the following website: <https://www.huawei.com/en/psirt/vul-response-process>

For enterprise customers who need to obtain vulnerability information, visit: <https://securitybulletin.huawei.com/enterprise/en/security-advisory>

Preconfigured Digital Certificate

Huawei has released the Huawei Preset Digital Certificate Disclaimer for the preconfigured digital certificates delivered with devices. For details about the disclaimer, visit the following website: <https://support.huawei.com/enterprise/en/bulletins-service/ENEWS2000015789>

Life Cycle of Product Documentation

Huawei released the Huawei Product Documentation Lifecycle Policy for after-sales customer documentation. For details about this policy, see the website of Huawei's official website.: <https://support.huawei.com/enterprise/en/bulletins-website/ENEWS2000017761>

About This Document

Intended Audience

This document is intended for network engineers responsible for CE series switches configuration and management. You should be familiar with basic Ethernet knowledge and have extensive experience in network deployment and management.

Symbol Conventions

The symbols that may be found in this document are defined as follows.

Symbol	Description
 NOTICE	Indicates a potentially hazardous situation which, if not avoided, could result in equipment damage, data loss, performance deterioration, or unanticipated results. NOTICE is used to address practices not related to personal injury.
 NOTE	Supplements the important information in the main text. NOTE is used to address information not related to personal injury, equipment damage, and environment deterioration.

Command Conventions

The command conventions that may be found in this document are defined as follows.

Convention	Description
Boldface	The keywords of a command line are in boldface .

Convention	Description
<i>Italic</i>	Command arguments are in <i>italics</i> .
[]	Items (keywords or arguments) in brackets [] are optional.
{ x y ... }	Optional items are grouped in braces and separated by vertical bars. One item is selected.
[x y ...]	Optional items are grouped in brackets and separated by vertical bars. One item is selected or no item is selected.
{ x y ... }*	Optional items are grouped in braces and separated by vertical bars. A minimum of one item or a maximum of all items can be selected.
[x y ...]*	Optional items are grouped in brackets and separated by vertical bars. Several items or no item can be selected.
&<1-n>	The parameter before the & sign can be repeated 1 to n times.
#	A line starting with the # sign is comments.

Interface Numbering Conventions

Interface numbers used in this manual are examples. In device configuration, use the existing interface numbers on devices.

Security Conventions

- Password setting
 - When configuring a password, the cipher text is recommended. To ensure device security, change the password periodically.
 - When you configure a password in plain text that starts and ends with %^%#.....%^%# (the password can be decrypted by the device), the password is displayed in the same manner as the configured one in the configuration file. Do not use this setting. After the system master key is set using the **set master-key** command, do not start and end the key with %@%# because the string starting and ending with %@%# is considered as a valid cipher-text key.
 - When you configure a password in cipher text, different features cannot use the same cipher-text password. For example, the cipher-text password set for the AAA feature cannot be used for other features.
 - After the system software is downgraded and the switch restarts with the configuration of the higher version, AAA, VTY, serial interface login, and

SNMP user passwords become invalid. As a result, users fail to log in to the switch using the passwords and the switch is disconnected from the network management system.

To address this problem, take the following measures:

- i. If no password is configured for the console port, log in to the device through the console port, and reconfigure AAA and password for users such as VTY and SNMP users. For security purposes, the console port password is recommended.
- ii. If a password is configured for login through the console port, the password becomes invalid after the downgrade and you cannot log in to the switch through the console port. In this case, contact Huawei technical support engineers for assistance.

- Encryption algorithm

Currently, the device uses the following encryption algorithms: DES, 3DES, AES, DSA, RSA, DH, ECDH, HMAC, SHA1, SHA2, PBKDF2, scrypt, and MD5. The encryption algorithm depends on the applicable scenario. Use the recommended encryption algorithm; otherwise, security defense requirements may be not met.

- For the symmetrical encryption algorithm, use AES with the key of 256 bits or more.
- When you need to use an asymmetric cryptography, RSA (3072-bit or longer key) is recommended. In addition, use different key pairs for encryption and signature.
- For the digital signature, RSA (3072-bit or longer key) or DSA (2048-bit or longer key) is recommended.
- For key negotiation, DH (2048-bit or longer key) or ECDH (256-bit or longer key) is recommended.
- For the hash algorithm, use SHA with the key of 256 bits or more.
- For the HMAC algorithm, use HMAC-SHA2.
- DES, 3DES, RSA and AES are reversible encryption algorithm. If protocols are used for interconnection, the locally stored password must be reversible.
- SHA1, SHA2, and MD5 are irreversible encryption algorithm. When configuring a password for local administrator, it is recommended that you use the SHA2 irreversible encryption algorithm.
- To prevent brute force cracking of the user password, the iteration algorithm is added to the password on the basis of salts. The iteration algorithm uses PBKDF2 or scrypt key export algorithm.
- The ECB mode has a poor capability of defending against plaintext playback attacks, so ECB is not recommended for password encryption.
- In SSH2.0, the symmetric cryptography using the CBC mode may undergo the plaintext-recovery attack to cause a data leak. Therefore, the CBC mode is not recommended for SSH2.0.

- Data

Some data (such as MAC or IP addresses of terminals) may be obtained or used during operation or fault location of your purchased products, services, features, so you have an obligation to make privacy policies and take measures according to the applicable law of the country to protect data.

- The terms mirrored port, port mirroring, traffic mirroring, and mirroring in this manual are mentioned only to describe the product's function of communication error or failure detection, and do not involve collection or processing of any personal information or communication data of users.

Reference Standards and Protocols

To obtain reference standards and protocols, log in to [Huawei official website](#), search for "protocol compliance list", and download the *Huawei CloudEngine Switches Protocol Compliance List*.

Declaration

- This manual is only a reference for you to configure your devices. The contents in the manual, such as command line syntax, and command outputs, are based on the device conditions in the lab. The manual provides instructions for general scenarios, but do not cover all usage scenarios of all product models. The contents in the manual may be different from your actual device situations due to the differences in software versions, models, and configuration files. The manual will not list every possible difference. You should configure your devices according to actual situations.
- The specifications provided in this manual are tested in lab environment (for example, the tested device has been configured with a certain type of cards or only one protocol is run on the device). Results may differ from the listed specifications when you attempt to obtain the maximum values with multiple functions enabled on the device.
- In this document, public IP addresses may be used in feature introduction and configuration examples and are for reference only unless otherwise specified.

Mapping Between Product Software and Related Products

For details about the mapping between product software and related products, see [CloudEngine Switches Software Versions](#). Select the correct matching version. Otherwise, functions may be unavailable.

Contents

About This Document.....	iii
1 MPLS Basics.....	1
1.1 Overview of MPLS.....	1
1.2 Understanding MPLS.....	2
1.2.1 Basic MPLS Architecture.....	2
1.2.2 MPLS Label.....	4
1.2.3 LSP Setup.....	6
1.2.4 MPLS Forwarding.....	7
1.2.5 LSP Connectivity Check.....	10
1.3 Application Scenarios for MPLS.....	12
1.3.1 MPLS VPN.....	12
1.4 MPLS Common Configuration.....	13
1.4.1 (Optional) Configuring the MPLS TTL Processing Mode.....	13
1.4.2 (Optional) Configuring MPLS MTU.....	15
1.4.3 (Optional) Configuring the PHP Feature.....	16
1.4.4 Enabling the MPLS Trap Function.....	17
1.4.5 (Optional) Configuring the Function of Removing MPLS Labels.....	21
2 MPLS LDP Configuration.....	23
2.1 Overview of MPLS LDP.....	23
2.2 Understanding MPLS LDP.....	24
2.2.1 Basic Concepts of LDP.....	24
2.2.2 LDP Working Mechanism.....	25
2.2.2.1 LDP Overview.....	25
2.2.2.2 LDP Session Setup.....	26
2.2.2.3 LDP LSP Setup.....	28
2.2.3 LDP Security Mechanisms.....	32
2.2.4 LDP Reliability.....	34
2.2.4.1 Overview of LDP Reliability.....	34
2.2.4.2 BFD for LDP LSP.....	35
2.2.4.3 LDP-IGP Synchronization.....	36
2.2.4.4 Auto LDP FRR.....	39
2.2.4.5 LDP GR.....	41

2.2.4.6 LDP NSR.....	43
2.2.4.7 Local-and-Remote LDP Session.....	44
2.2.5 LDP Session Protection.....	45
2.3 Summary of MPLS LDP Configuration Tasks.....	46
2.4 Licensing Requirements and Limitations for MPLS LDP.....	48
2.5 Default Settings for MPLS LDP.....	51
2.6 Configuring Basic Functions of MPLS LDP.....	51
2.6.1 Configuring the LSR ID.....	52
2.6.2 Enabling Global MPLS	52
2.6.3 Enabling Global MPLS LDP.....	53
2.6.4 Configuring LDP Sessions.....	53
2.6.5 (Optional) Configuring LDP Session Protection.....	56
2.6.6 (Optional) Configuring an LDP Transport Address.....	57
2.6.7 (Optional) Configuring Timers for LDP Session.....	58
2.6.8 (Optional) Configuring LDP Loop Detection.....	63
2.6.9 (Optional) Configuring the LDP Label Policies.....	64
2.6.10 (Optional) Disabling a Device from Distributing Labels to Remote Peers.....	65
2.6.11 (Optional) Configuring a Policy for Triggering LDP LSP Establishment.....	67
2.6.12 (Optional) Disabling LDP LSP Flapping Suppression.....	69
2.6.13 (Optional) Disabling a Device from Forwarding Unknown TLVs.....	70
2.6.14 Verifying the Configuration of Basic MPLS LDP Functions.....	70
2.7 Configuring Static BFD to Detect an LDP LSP.....	71
2.7.1 Configuring BFD with Specific Parameters on the Ingress Node.....	72
2.7.2 Configuring BFD with Specific Parameters on the Egress Node.....	74
2.7.3 Verifying the Configuration of Static BFD for LDP LSPs.....	75
2.8 Configuring Synchronization Between LDP and IGP.....	75
2.8.1 Enabling Synchronization Between LDP and IGP.....	76
2.8.2 (Optional) Blocking Synchronization Between LDP and IGP on an Interface.....	79
2.8.3 (Optional) Setting the Hold-down Timer Value.....	81
2.8.4 (Optional) Setting the Hold-max-cost Timer Value.....	83
2.8.5 (Optional) Setting the Delay Timer Value.....	86
2.8.6 (Optional) Configuring LDP Graceful Deletion.....	87
2.8.7 Verifying the Configuration of Synchronization Between LDP and IGP.....	88
2.9 Configuring Auto LDP FRR.....	88
2.10 Configuring LDP GR Helper.....	90
2.11 Configuring LDP Security Mechanisms.....	92
2.11.1 Configuring LDP MD5 Authentication.....	93
2.11.2 Configuring LDP Keychain Authentication.....	95
2.11.3 Configuring the LDP GTSM.....	96
2.11.4 Verifying the Configuration of LDP Security Mechanisms.....	97
2.12 Configuring Non-labeled Public Network Routes to Be Iterated to LSPs.....	97
2.13 Maintaining MPLS LDP.....	98

2.13.1 Disabling LDP Session Flapping Suppression.....	98
2.13.2 Disabling LDP Interface Flapping Suppression.....	99
2.13.3 Resetting LDP.....	99
2.13.4 Clearing LDP Statistics.....	100
2.13.5 Monitoring the LDP Running Status.....	100
2.13.6 Checking the LSP Connectivity.....	101
2.14 Configuration Examples for MPLS LDP.....	101
2.14.1 Example for Configuring Static BFD to Detect LDP LSPs.....	102
2.15 Troubleshooting MPLS LDP.....	108
2.15.1 LDP Session Alternates Between Up and Down States.....	108
2.15.2 LDP Session Is Down.....	109
2.15.3 LDP LSP Alternates Between Up and Down States.....	110
2.15.4 LDP LSP Is Down.....	110
3 MPLS QoS Configuration.....	112
3.1 Overview of MPLS QoS.....	112
3.2 Understanding MPLS QoS.....	113
3.2.1 MPLS DiffServ.....	113
3.2.2 MPLS DiffServ Tunnel Modes.....	115
3.3 Licensing Requirements and Limitations for MPLS QoS.....	117
3.4 Default Settings for MPLS QoS.....	118
3.5 Configuring MPLS QoS.....	121
3.5.1 Configuring MPLS Public QoS (CE6870EI).....	121
3.5.2 Configuring MPLS Public QoS (CE Switches Excluding CE6870EI).....	123
3.5.3 Configuring MPLS L3VPN QoS (CE6870EI).....	126
3.5.4 Configuring MPLS L3VPN QoS (CE Switches Excluding CE6870EI).....	126
3.5.5 Configuring MPLS L2VPN QoS (CE6870EI).....	128
3.5.6 Configuring MPLS L2VPN QoS (Except CE6870EI).....	130

1 MPLS Basics

This chapter describes MPLS basics.

[1.1 Overview of MPLS](#)

[1.2 Understanding MPLS](#)

[1.3 Application Scenarios for MPLS](#)

[1.4 MPLS Common Configuration](#)

1.1 Overview of MPLS

Definition

The Multiprotocol Label Switching (MPLS) protocol is used on Internet Protocol (IP) backbone networks. MPLS uses connection-oriented label switching on connectionless IP networks. By combining Layer 3 routing technologies and Layer 2 switching technologies, MPLS leverages the flexibility of IP routing and the simplicity of Layer 2 switching.

MPLS is based on Internet Protocol version 4 (IPv4). The core MPLS technology can be extended to multiple network protocols, such as Internet Protocol version 6 (IPv6), Internet Packet Exchange (IPX), and Connectionless Network Protocol (CLNP). "Multiprotocol" in MPLS means that multiple network protocols are supported.

MPLS is used for tunneling but not a service or an application. MPLS supports multiple protocols and services. Moreover, it ensures security of data transmission.

Purpose

IP-based routing serves well on the Internet in the mid 90s, but IP technology can be inefficient at forwarding packets because software must search for routes using the longest match algorithm. As a result, the forwarding capability of IP technology can act as a bottleneck.

In contrast, Asynchronous transfer mode (ATM) technology uses labels of fixed length and maintains a label table that is much smaller than a routing table.

Compared to IP, ATM is more efficient at forwarding packets. ATM is a complex protocol, however, with high deployment costs, that hinder its widespread use.

Because traditional IP technology is simple and costs little to deploy, a combination of IP and ATM capabilities would be ideal. This has sparked the emergence of MPLS technology.

MPLS was created to increase forwarding rates. Unlike IP routing and forwarding, MPLS analyzes a packet header only on the edge of the network and not at each hop. MPLS therefore reduces packet processing time.

The use of hardware-based functions based on application-specific integrated circuits (ASICs) has made IP routing far more efficient, so MPLS is no longer needed for its high-speed forwarding advantages. However, MPLS does not support multi-layer labels, and its forwarding plane is connection-oriented. For these reasons, MPLS is widely used for virtual private network (VPN), and quality of service (QoS).

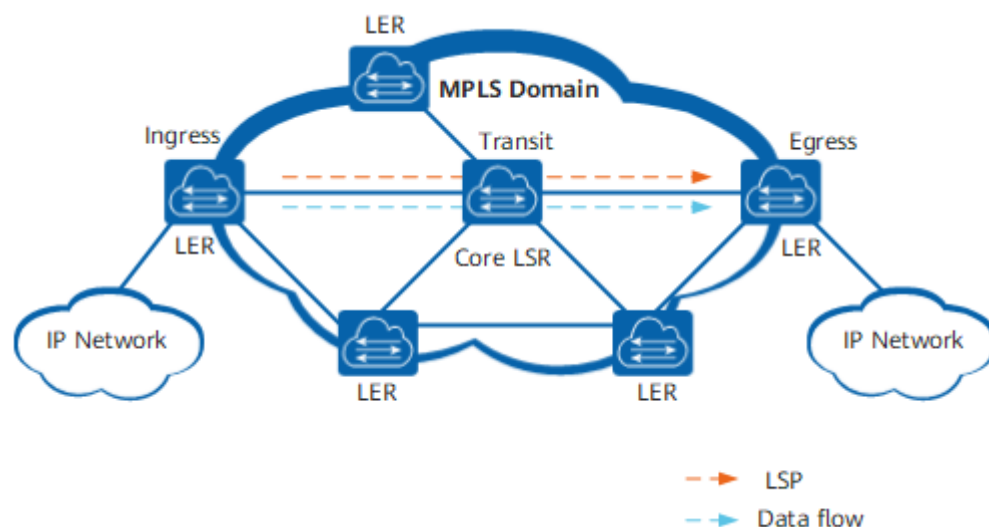
1.2 Understanding MPLS

1.2.1 Basic MPLS Architecture

MPLS Network Structure

Figure 1-1 shows a typical MPLS network structure. Packets are forwarded on an MPLS network based on labels. In **Figure 1-1**, network devices that swap MPLS labels and forward packets are label switching routers (LSRs), which form an MPLS domain. LSRs that reside at the edge of the MPLS domain and connect to other networks are called label edge routers (LERs), and LSRs within the MPLS domain are core LSRs.

Figure 1-1 MPLS network structure



When IP packets reach an MPLS network, the ingress LER analyzes the packets and then adds appropriate labels to them. All LSRs on the MPLS network forward packets based on labels. When IP packets leave the MPLS network, the egress LER pops the labels.

A path along which IP packets are transmitted on an MPLS network is called a label switched path (LSP). An LSP is a unidirectional path in the same direction data packets traverse.

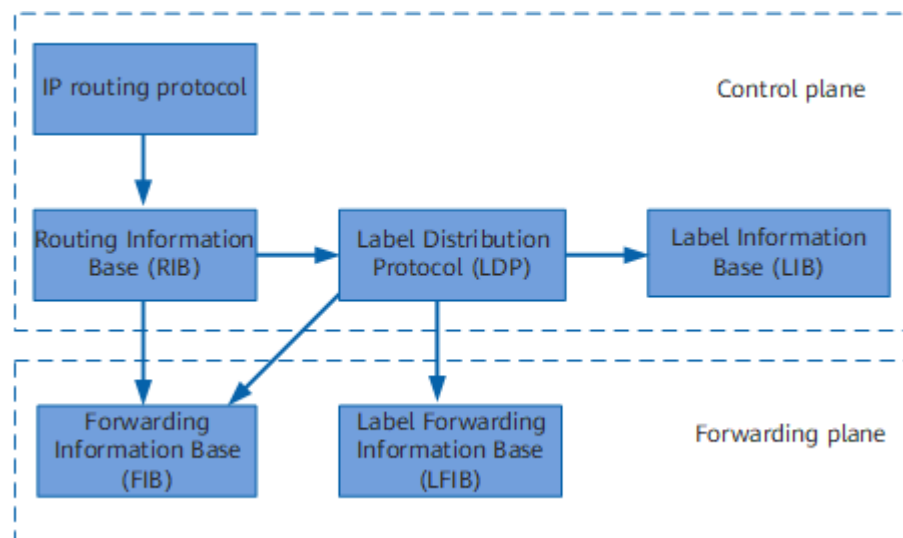
As shown in [Figure 1-1](#), the LER at the starting point of an LSP is the ingress node, and the LER at the end of the LSP is the egress node. The LSRs between the ingress node and egress node along the LSP are transit nodes. An LSP may have zero, one, or several transit nodes and only one ingress node and one egress node.

On an LSP, MPLS packets are sent from the ingress to the egress. In this transmission direction, the ingress node is the upstream node of the transit nodes, and the transit nodes are the downstream nodes of the ingress node. Similarly, transit nodes are the upstream nodes of the egress node, and the egress node is the downstream node of the transit nodes.

MPLS Architecture

[Figure 1-2](#) shows the MPLS architecture, which consists of a control plane and a forwarding plane.

Figure 1-2 MPLS architecture



The MPLS architecture has the following parts:

- Control plane: generates and maintains routing and label information
 - Routing information base (RIB): is generated by IP routing protocols and used to select routes.
 - Label distribution protocol (LDP): allocates labels, creates a label information base (LIB), and establishes and tears down LSPs.
 - Label information base (LIB): is generated by LDP and used to manage labels.

- Forwarding plane (data plane): forwards IP packets and MPLS packets
 - Forwarding information base (FIB): is generated based on routing information obtained from the RIB and used to forward common IP packets.
 - Label forwarding information base (LFIB): is created by LDP on an LSR and used to forward MPLS packets.

1.2.2 MPLS Label

Forwarding Equivalence Class

A forwarding equivalence class (FEC) is a collection of packets with the same characteristics. Packets of the same FEC are forwarded in the same way on an MPLS network.

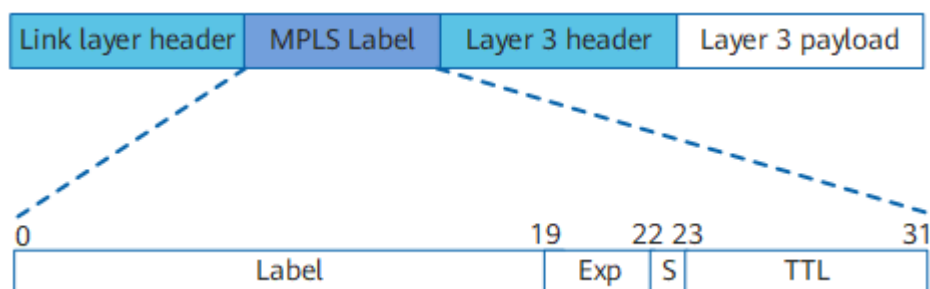
FECs can be identified by the source address, destination address, source port, destination port, and VPN. For example, in IP forwarding, packets matching the same route based on the longest match algorithm belong to an FEC.

Label

A label is a short, fixed-length (4 bytes) identifier that is only locally significant. A label identifies an FEC to which a packet belongs. In some cases, such as load balancing, an FEC can be mapped to multiple incoming labels. Each label, however, represents only one FEC on a device.

Compared with an IP packet, an MPLS packet has the additional 4-byte MPLS label. The MPLS label is between the link layer header and the network layer header, and allows use of any link layer protocol. [Figure 1-3](#) shows position of an MPLS label and fields in the MPLS label.

Figure 1-3 MPLS label encapsulation format



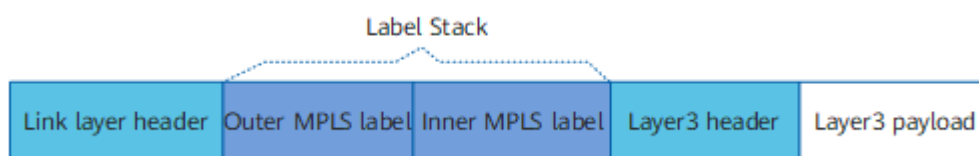
An MPLS label contains the following fields:

- Label: 20-bit label value.
- Exp: 3-bit, used as an extension value. Generally, this field is used as the class of service (CoS) field. When congestion occurs, devices prioritize packets that have a larger value in this field.
- S: 1-bit value indicating the bottom of a label stack. MPLS supports nesting of multiple labels. When the S field is 1, the label is at the bottom of the label stack.

- TTL: time to live. This 8-bit field is the same as the TTL field in IP packets.

A label stack is an arrangement of labels. In [Figure 1-4](#), the label next to the Layer 2 header is the top of the label stack (outer MPLS label), and the label next to the Layer 3 header is the bottom of the label stack (inner MPLS label). An MPLS label stack can contain an unlimited number of labels. Currently, MPLS label stacks can be applied to MPLS VPN.

Figure 1-4 Label stack



The label stack organizes labels according to the rule of Last-In, First-Out. The labels are processed from the top of the stack.

Label Space

The label space is the value range of the label, and the space is organized in the following ranges:

- 0 to 15: special labels. For details about special labels, see [Table 1-1](#).
- 16 or above: label space for dynamic signaling protocols, such as Label Distribution Protocol (LDP) and MultiProtocol Border Gateway Protocol (MP-BGP).

Table 1-1 Special labels

Label Value	Label	Description
0	IPv4 Explicit NULL Label	The label must be popped out (removed), and the packets must be forwarded based on IPv4. If the egress node allocates a label with the value of 0 to the penultimate hop LSR, the penultimate hop LSR pushes label 0 to the top of the label stack and forwards the packet to the egress node. When the egress node detects that the label of the packet is 0, the egress node pops the label out.
1	Router Alert Label	A label that is only valid when it is not at the bottom of a label stack. The label is similar to the Router Alert Option field in IP packets. After receiving such a label, the node sends it to a local software module for further processing. Packet forwarding is determined by the next-layer label. If the packet needs to be forwarded continuously, the node pushes the Router Alert Label to the top of the label stack again.

Label Value	Label	Description
2	IPv6 Explicit NULL Label	The label must be popped out, and the packets must be forwarded based on IPv6. If the egress node allocates a label with the value of 2 to the LSR at the penultimate hop, the LSR pushes label 2 to the top of the label stack and forwards the packet to the egress node. When the egress node recognizes that the value of the label carried in the packet is 2, the egress node immediately pops it out.
3	Implicit NULL Label	When the label with the value of 3 is swapped on an LSR at the penultimate hop, the LSR pops the label out and forwards the packet to the egress node. Upon receiving the packet, the egress node forwards the packet in IP forwarding mode or according to the next layer label.
4 to 13	Reserved	None.
14	OAM Router Alert Label	A label for operation, administration and maintenance (OAM) packets over an MPLS network. MPLS OAM sends OAM packets to monitor LSPs and report faults. OAM packets are transparent on transit nodes and the penultimate LSR.
15	Reserved	None.

1.2.3 LSP Setup

Before forwarding packets, MPLS must allocate labels to packets and establish an LSP.

Label Distribution Protocols for Dynamic LSPs

Dynamic LSPs are established using label distribution protocols. As the control protocol or signaling protocol for MPLS, a label distribution protocol defines FECs, distributes labels, and establishes and maintains LSPs.

MPLS can use the following protocols for label distribution:

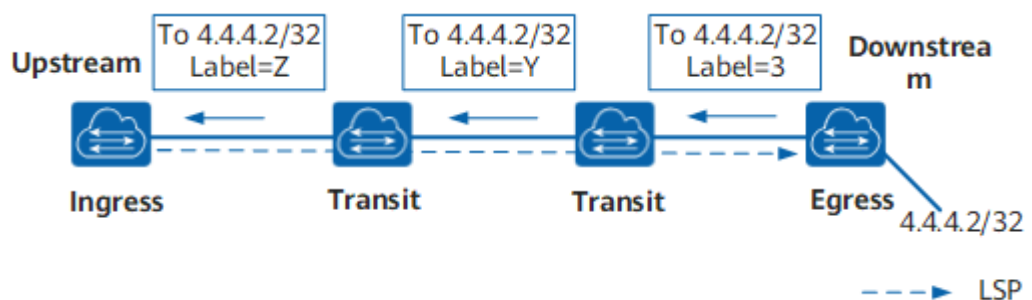
- LDP
The Label Distribution Protocol (LDP) is designed for distributing labels. It sets up an LSP hop by hop according to IGP and BGP routing information.
For details about LDP implementation, see Understanding MPLS LDP in the [2 MPLS LDP Configuration](#).
- MP-BGP
MP-BGP is an extension to BGP and allocates labels to MPLS VPN routes and inter-AS VPN routes.

For details about MP-BGP implementation, see BGP Configuration in *CloudEngine 9800, 8800, 6800, 5800 Series Switches Configuration Guide - IP Routing*.

Procedure for Establishing Dynamic LSPs

MPLS labels are distributed from downstream LSRs to upstream LSRs. As shown in [Figure 1-5](#), a downstream LSR identifies FECs based on the IP routing table, allocates a label to each FEC, and records the mapping between labels and FECs. The downstream LSR then encapsulates the mapping into a message and sends the message to the upstream LSR. As this process proceeds on all the LSRs, the LSRs create a label forwarding information base (LFIB) and establish an LSP.

Figure 1-5 Establishing a dynamic LSP



1.2.4 MPLS Forwarding

MPLS Forwarding Process

Basic Concepts

Label operations involved in MPLS packet forwarding include push, swap, and pop:

- **Push:** When an IP packet enters an MPLS domain, the ingress node adds a new label to the packet between the Layer 2 header and the IP header. Alternatively, an LSR adds a new label to the top of the label stack.
- **Swap:** When a packet is transferred within the MPLS domain, a local node swaps the label at the top of the label stack in the MPLS packet for the label allocated by the next hop according to the label forwarding table.
- **Pop:** When a packet leaves the MPLS domain, the label is popped out of (removed from) the MPLS packet.

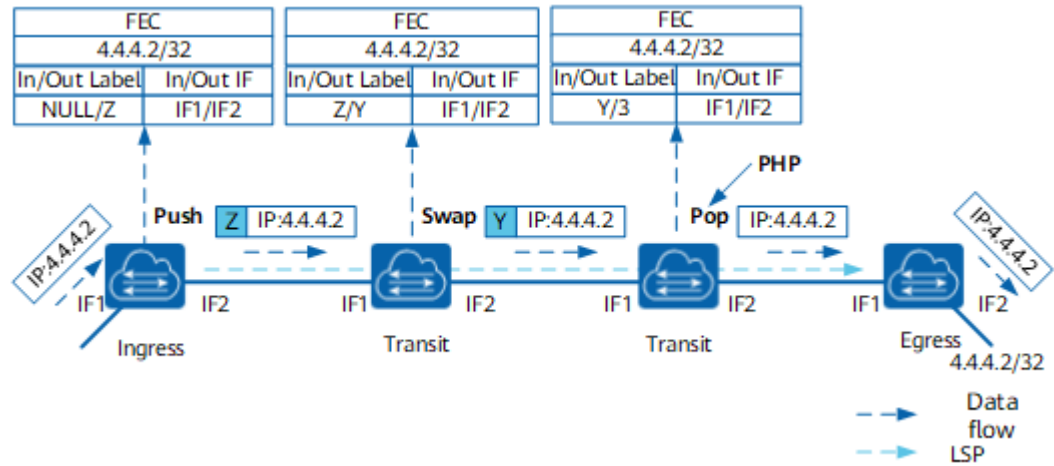
A label is invalid at the last hop of an MPLS domain. The penultimate hop popping (PHP) feature applies. On the penultimate node, the label is popped out of the packet to reduce the size of the packet that is forwarded to the last hop. Then, the last hop directly forwards the IP packet or forwards the packet by using the second label.

By default, PHP is configured on the egress node. The egress node supporting PHP allocates the label with the value of 3 to the penultimate hop.

Basic Forwarding Process

LSPs that support PHP are used in the following example to describe how MPLS packets are forwarded.

Figure 1-6 Basic MPLS forwarding process



As shown in [Figure 1-6](#), the LSRs have distributed MPLS labels and set up an LSP with the destination address of 4.4.4.2/32. MPLS packets are forwarded as follows:

1. The ingress node receives an IP packet destined for 4.4.4.2. Then, the ingress node adds Label Z to the packet and forwards it.
2. When the downstream transit node receives the labeled packet, the node replaces Label Z by Label Y.
3. When the transit node at the penultimate hop receives the packet with Label Y, the node pops out Label Y because the label value is 3. The transit node then forwards the packet to the egress node as an IP packet.
4. The egress node receives the IP packet and forwards it to 4.4.4.2/32.

MPLS TTL Processing

This section describes how MPLS processes the TTL and responds to TTL timeout.

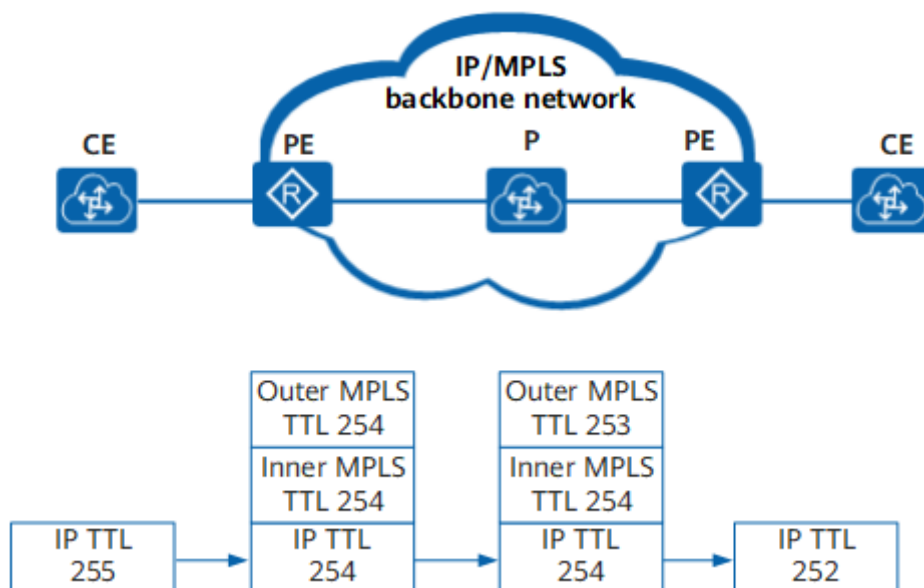
MPLS TTL Processing Modes

The TTL field in an MPLS label is 8 bits long. The TTL field is the same as that in an IP packet header. MPLS processes the TTL to prevent loops and implement traceroute.

MPLS TTL processing modes include: Uniform and Pipe modes. By default, MPLS processes the TTL in Uniform mode. The two modes work as follows:

- **Uniform mode**
When IP packets enter an MPLS network, the ingress node decreases the IP TTL by one and copies this new value to the MPLS TTL field. The TTL field in MPLS packets is processed in standard mode. The egress node decreases the MPLS TTL by one and maps this new value to the IP TTL field. [Figure 1-7](#) shows how the TTL field is processed on the transmission path.

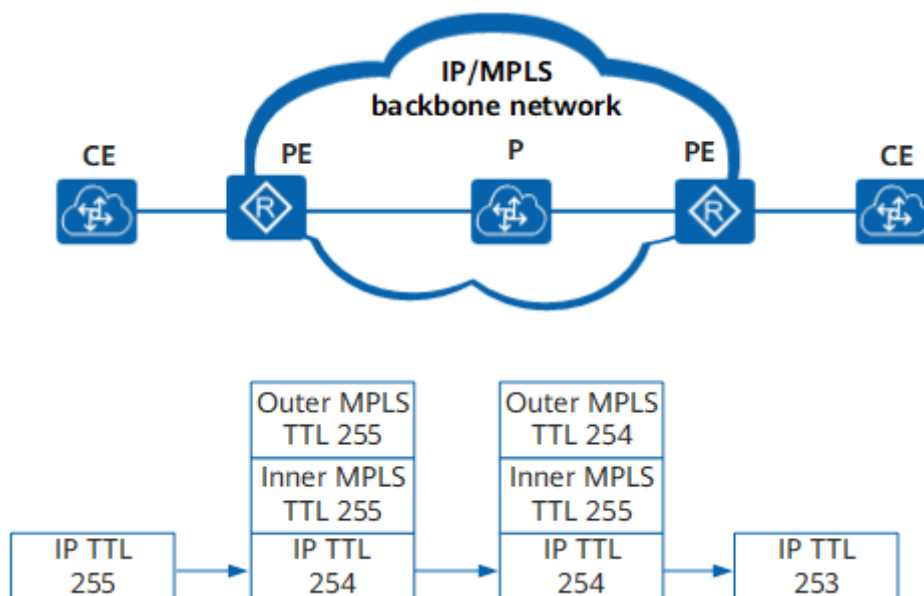
Figure 1-7 TTL processing in Uniform mode



- Pipe mode

As shown in [Figure 1-8](#), the ingress node decreases the IP TTL by one and the MPLS TTL remains constant. The TTL field in MPLS packets is processed in standard mode. The egress node decreases the IP TTL by one. In Pipe mode, the IP TTL only decreases by one on the ingress node and one on the egress node when packets travel across an MPLS network.

Figure 1-8 TTL processing in Pipe mode



In MPLS VPN applications, the MPLS backbone network needs to be shielded to ensure network security. The Pipe mode is recommended for private network packets.

ICMP Response Packet

On an MPLS network, when an LSR receives an MPLS packet with the TTL value of 1, the LSR generates an ICMP response packet.

The LSR returns the ICMP response packet to the sender in the following ways:

- If the LSR has a reachable route to the sender, the LSR directly sends the ICMP response packet to the sender through the IP route.
- If the LSR has no reachable route to the sender, the LSR forwards the ICMP response packet along the LSP. The egress node forwards the ICMP response packet to the sender.

In most cases, the received MPLS packet contains only one label and the LSR responds to the sender with the ICMP response packet using the first method. If the MPLS packet contains multiple labels, the LSR uses the second method.

The MPLS VPN packets may contain only one label when they arrive at an autonomous system boundary router (ASBR) on the MPLS VPN. These devices have no IP routes to the sender, so they use the second method to reply to the ICMP response packets.

1.2.5 LSP Connectivity Check

Overview of LSP Connectivity Check

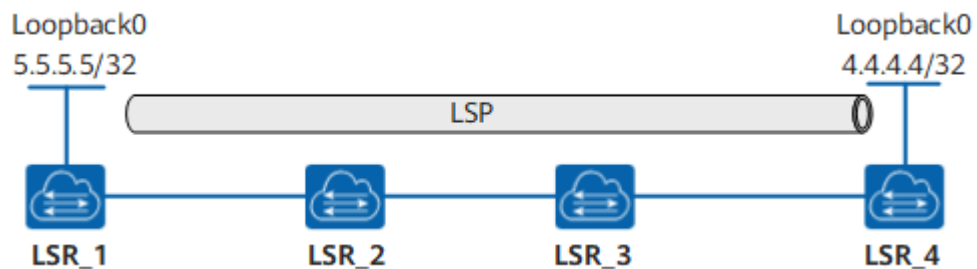
On an MPLS network, the control panel used for setting up an LSP cannot detect the failure in data forwarding of the LSP. This makes network maintenance difficult. The MPLS ping and tracer mechanisms detect LSP errors and locate faulty nodes.

MPLS ping is used to check network connectivity. MPLS tracer is used to check the network connectivity, and to locate network faults. Similar to IP ping and tracer, MPLS ping and tracer use MPLS echo request packets and MPLS echo reply packets to check LSP availability. MPLS echo request packets and echo reply packets are both encapsulated into User Datagram Protocol (UDP) packets. The UDP port number of the MPLS echo request packet is 3503, which can be identified only by MPLS-enabled devices.

An MPLS echo request packet carries FEC information to be detected, and is sent along the same LSP as other packets with the same FEC. In this manner, the connectivity of the LSP is checked. MPLS echo request packets are forwarded to the destination end using MPLS, while MPLS echo reply packets are forwarded to the source end using IP. Routers set the destination address in the IP header of the MPLS echo request packets to 127.0.0.1/8 (local loopback address) and the TTL value is 1. In this way, MPLS echo request packets are not forwarded using IP forwarding when the LSP fails so that the failure of the LSP can be detected.

MPLS Ping

Figure 1-9 MPLS network



As shown in [Figure 1-9](#), LSR_1 establishes an LSP to LSR_4. LSR_1 performs MPLS ping on the LSP by performing the following steps:

1. LSR_1 checks whether the LSP exists. If the LSP does not exist, an error message is displayed and the MPLS ping stops. If the LSP exists, LSR_1 performs the following operations.
2. LSR_1 creates an MPLS echo request packet and adds 4.4.4.4 to the destination FEC in the packet. In the IP header of the MPLS echo request packet, the destination address is 127.0.0.1/8 and the TTL value is 1. LSR_1 searches for the corresponding LSP, adds the LSP label to the MPLS echo request packet, and sends the packet to LSR_2.
3. Transit nodes LSR_2 and LSR_3 forward the MPLS echo request packet based on MPLS. If MPLS forwarding on a transit node fails, the transit node returns an MPLS echo reply packet carrying the error code to LSR_1.
4. If no fault exists along the MPLS forwarding path, the MPLS echo request packet reaches the LSP egress node LSR_4. LSR_4 returns a correct MPLS echo reply packet after verifying that the destination IP address 4.4.4.4 is the loopback interface address. MPLS ping is complete.

MPLS Tracert

As shown in [Figure 1-9](#), LSR_1 performs MPLS tracert on LSR_4 (4.4.4.4/32) by performing the following steps:

1. LSR_1 checks whether an LSP exists to LSR_4. If the LSP does not exist, an error message is displayed and the tracert stops. If the LSP exists, LSR_1 performs the following operations.
2. LSR_1 creates an MPLS echo request packet and adds 4.4.4.4 to the destination FEC in the packet. In the IP header of the MPLS echo request packet, the destination address is 127.0.0.1/8. Then LSR_1 adds the LSP label to the packet, sets the MPLS TTL value to 1, and sends the packet to LSR_2. The MPLS echo request packet contains a downstream mapping type-length-value (TLV) that carries downstream information about the LSP at the current node, such as next-hop address and outgoing label.
3. Upon receiving the MPLS echo request packet, LSR_2 decreases the MPLS TTL by one and finds that TTL times out. LSR_2 then checks whether the LSP exists and the next-hop address and whether the outgoing label of the

downstream mapping TLV in the packet is correct. If so, LSR_2 returns a correct MPLS echo reply packet that carries the downstream mapping TLV of LSR_2. If not, LSR_2 returns an incorrect MPLS echo reply packet.

4. After receiving the correct MPLS echo reply packet, LSR_1 resends the MPLS echo request packet that is encapsulated in the same way as step 2 and sets the MPLS TTL value to 2. The downstream mapping TLV of this MPLS echo request packet is replicated from the MPLS echo reply packet. LSR_2 performs common MPLS forwarding on this MPLS echo request packet. If TTL times out when LSR_3 receives the MPLS echo request packet, LSR_3 processes the MPLS echo request packet and returns an MPLS echo reply packet in the same way as step 3.
5. After receiving a correct MPLS echo reply packet, LSR_1 repeats step 4, sets the MPLS TTL value to 3, replicates the downstream mapping TLV in the MPLS echo reply packet, and sends the MPLS echo request packet. LSR_2 and LSR_3 perform common MPLS forwarding on this MPLS echo request packet. Upon receiving the MPLS echo request packet, LSR_4 repeats step 3 and verifies that the destination IP address 4.4.4.4 is the loopback interface address. LSR_4 returns an MPLS echo reply packet that does not carry the downstream mapping TLV. MPLS tracer is complete.

When routers return the MPLS echo reply packet that carries the downstream mapping TLV, LSR_1 obtains information about each node along the LSP.

1.3 Application Scenarios for MPLS

1.3.1 MPLS VPN

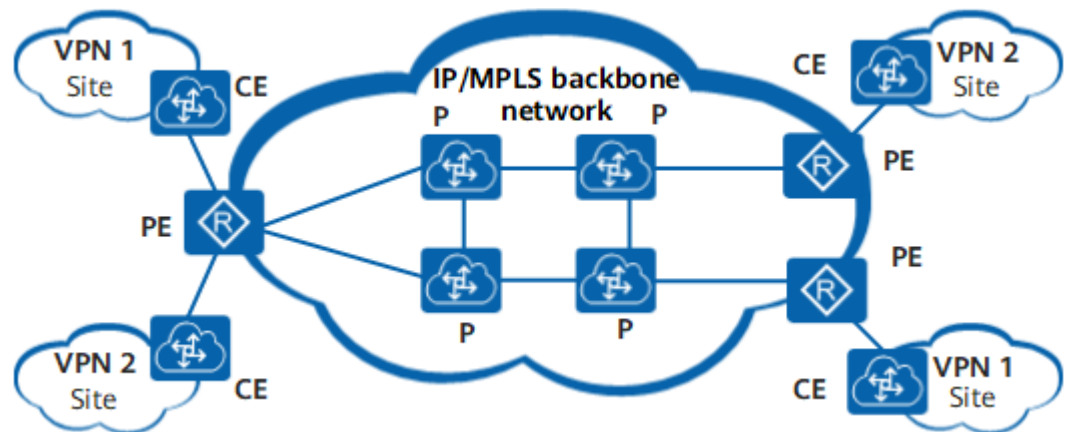
Traditional VPNs transmit private network data over the public network using tunneling protocols, such as the Generic Routing Encapsulation (GRE), Layer 2 Tunneling Protocol (L2TP), and Point to Point Tunneling Protocol (PPTP). MPLS LSPs are set up by swapping labels, and data packets are not encapsulated or encrypted. Therefore, MPLS is an appropriate technology for VPN implementation.

MPLS VPN can build a private network with security similar to a Frame Relay (FR) network. On MPLS VPN networks, customer devices do not need to set up tunnels such as GRE and L2TP tunnels, so the network delay is minimized.

As shown in [Figure 1-10](#), the MPLS VPN connects private network branches through LSPs to form a unified network. The MPLS VPN also controls the interconnection between VPNs. [Figure 1-10](#) shows the devices on an MPLS VPN network.

- A customer edge (CE) is deployed on the edge of a customer network. It can be a router, a switch, or a host.
- A provider edge (PE) is deployed on the edge of an IP/MPLS backbone network.
- A provider (P) device on an IP/MPLS backbone network is not directly connected to CEs. The provider device only needs to provide basic MPLS forwarding capabilities and does not maintain VPN information.

Figure 1-10 MPLS VPN



An MPLS VPN has the following characteristics:

- PEs manage VPN users, set up LSPs between PEs, and advertise routing information between users in a VPN.
- PEs use MP-BGP to advertise VPN routing information.
- The MPLS-based VPN supports IP address multiplexing between sites as well as the interconnection of different VPNs.

1.4 MPLS Common Configuration

1.4.1 (Optional) Configuring the MPLS TTL Processing Mode

Context

MPLS processes the TTL in the following modes:

- MPLS TTL processing modes
In MPLS applications, the MPLS backbone network needs to be shielded to ensure network security. The MPLS Pipe mode on the ingress node is recommended for private network packets. To reflect the path where packets pass, use the MPLS Uniform mode on the ingress node.
- Path where ICMP response packets are transmitted
By default, when the received MPLS packet contains only one label, the LSR directly sends an ICMP response packet to the sender using an IP route. When the received MPLS packet contains multiple labels, the LSR sends an ICMP response packet to the sender along an LSP.
However, MPLS packets received by the devices may contain only one label in some cases. The devices have no IP routes to the sender, so they forward the ICMP response packets along an LSP.

Procedure

- Configuring the MPLS TTL processing mode.
For the CE6870EI, run the following commands.

- a. Run **system-view**
The system view is displayed.
- b. Run **mpls mode { short-pipe [ttl *ttlvalue* exp *expvalue*] | uniform }**
The MPLS TTL processing mode is configured.
By default, the MPLS TTL processing mode is Uniform.
- c. Run **commit**
The configuration is committed.

For non-CE6870EI, run the following commands.

- a. Run **system-view**
The system view is displayed.
 - b. Run **mpls**
The MPLS view is displayed.
 - c. Run **mpls ldp ttl-mode { pipe | uniform }**
The mode for MPLS LDP to process TTLs in packets is configured.
By default, MPLS LDP processes TTLs in packets in uniform mode.
 - d. Run **quit**
The system view is displayed.
 - e. Run **commit**
The configuration is committed.
- Configuring the path where ICMP response packets are transmitted
 - a. Run **system-view**
The system view is displayed.
 - b. Run **mpls**
The MPLS view is displayed.
 - c. Run **ttl expiration pop disable**
The device is configured to transmit ICMP response packets along an LSP.

Or run **undo ttl expiration pop disable**:

The device is configured to transmit ICMP response packets using an IP route.

By default, the LSR returns an ICMP packet using the local IP route if the received MPLS TTL-expired packet contains one label.
 - d. Run **quit**
The system view is displayed.
 - e. Run **commit**
The configuration is committed.

----End

1.4.2 (Optional) Configuring MPLS MTU

Context

The size of the maximum transmission unit (MTU) determines the maximum number of bytes that can be transmitted by the sender at a time. If the MTU exceeds the maximum number of bytes supported by the receiver or a transit device, packets are fragmented or even discarded, which increases the network transmission load. In this manner, devices have to calculate the MTU before the communication to ensure that sent packets reach the receiver successfully.

$$\text{LDP MTU} = \text{Min} \{ \text{All MTUs advertised by all downstream devices, MTU of the local outbound interface} \}$$

A downstream LSR uses the preceding formula to calculate an MTU, adds it to the MTU TLV in a Label Mapping message, and sends the Label Mapping message to the upstream device. If an MTU value changes (such as when the local outbound interface or its configuration is changed), an LSR recalculates an MTU and sends a Label Mapping message carrying the new MTU to its upstream LSR. The relationships between the MPLS MTU and the interface MTU are as follows:

- If an interface MTU but not an MPLS MTU is configured on an interface, the interface MTU is used.
- If both an MPLS MTU and an interface MTU are configured on an interface, the smaller value between the MPLS MTU and the interface MTU is used.

MPLS determines the size of MPLS packets on the ingress node according to the LDP MTU to prevent the transit node from forwarding large-sized MPLS packets.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Run **mpls ldp**

The MPLS LDP view is displayed.

Step 3 Run the following commands as required.

- Run **mtu-signalling disable**

The LSR is disabled from sending Label Mapping messages carrying MTU TLVs.

By default, the switch enables Label Mapping messages from carrying an MTU TLV.

- Run **mtu-signalling apply-tlv**

The LSR is configured to send Label Mapping messages carrying MTU TLVs that comply with RFC 3988.

By default, the switch sends Label Mapping messages carrying Huawei proprietary MTU TLV.

Step 4 Run **quit**

The system view is displayed.

Step 5 Run **interface** *interface-type interface-number*

The view of an MPLS-enabled interface is displayed.

Step 6 On an Ethernet interface, run **undo portswitch**

The interface is switched to Layer 3 mode.

By default, an Ethernet interface works in Layer 2 mode.

The mode switching function takes effect when the interface only has attribute configurations (for example, **shutdown** and **description** configurations). Alternatively, if configuration information supported by both Layer 2 and Layer 3 interfaces exists (for example, **mode lacp** and **lacp system-id** configurations), no configuration that is not supported after the working mode of the interface is switched can exist. If unsupported configurations exist on the interface, delete the configurations first and then run the **undo portswitch** command.

 NOTE

If many Ethernet interfaces need to be switched to Layer 3 mode, run the **undo portswitch batch interface-type { interface-number1 [to interface-number2] } <1-10>** command in the system view to switch these interfaces to Layer 3 mode in batches.

Step 7 Run **mpls mtu mtu**

An MPLS MTU is configured on the interface.

By default, the MTU of MPLS packets is equal to the interface MTU.

Step 8 Run **commit**

The configuration is committed.

----End

1.4.3 (Optional) Configuring the PHP Feature

Context

No label needs to be swapped on the egress node of an LSP. PHP can be configured on the egress node to allow the LSR at the penultimate hop to pop out the label from an MPLS packet and send the packet to the egress node. After receiving the packet, the egress node directly forwards the packet through an IP link or according to the next layer label. PHP helps reduce the burden on the egress node.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Run **mpls**

The MPLS view is displayed.

Step 3 Run **label advertise { explicit-null | implicit-null | non-null }**

The label allocated to the LSR at the penultimate hop is configured.

The egress node can allocate different labels to the PHP based on the parameter setting.

- **implicit-null**: default value, which indicates that PHP is supported. If this parameter is configured, the egress node allocates an implicit null label with the value of 3 to the LSR at the penultimate hop.
- **explicit-null**: PHP is not supported. If this parameter is configured, the egress node allocates an explicit null label with the value of 0 to the LSR at the penultimate hop. The **explicit-null** parameter can be configured when MPLS QoS attributes are used.
- **non-null**: PHP is not supported. If this parameter is configured, the egress allocates a common label with a value greater than or equal to 16 to the LSR at the penultimate hop. The **non-null** parameter can be configured when MPLS QoS attributes are used.

 NOTE

After removing MPLS labels, switches except CE6870EI does not load balance MPLS packets through an Eth-Trunk based on the IP header by default. In the PHP scenario, you can configure the **mpls top-label src-ip dst-ip 2nd-label** command on the penultimate hop to load balance traffic to an Eth-Trunk. You can modify the load balancing profile based on the traffic characteristics to optimize the load balancing effect.

After the **label advertise** command is run to change the label distribution mode on the egress node, the modification takes effect on new LSPs but not on existing LSPs. To enable the modification to take effect on the existing LSPs, run the **reset mpls ldp** or **lsp-trigger** command.

Step 4 Run **commit**

The configuration is committed.

----End

1.4.4 Enabling the MPLS Trap Function

Context

To facilitate operation and maintenance and determine the running status of the MPLS network, configure the MPLS trap function so that the device can notify the NMS of the LDP session status.

Procedure

- Configure the trap function for LDP.
 - a. Run **system-view**

The system view is displayed.
 - b. Run **snmp-agent trap enable feature-name ldp [trap-name trapname]**

The trap function is enabled for the MPLS LDP module.

By default, the trap function is disabled for the MPLS LDP module.
- Configure the trap function for LSPM.

a. Run **system-view**

The system view is displayed.

b. Run **snmp-agent trap enable feature-name mpls_lspm [trap-name trapname]**

The trap function is enabled for the MPLS LSPM module.

By default, the trap function is disabled for the MPLS LSPM module.

When performing the following steps to configure alarm thresholds, pay attention to the following points:

- To configure the alarm function for dynamic label usage, specify **hwmplsdynamiclabelthresholdexceed** and **hwmplsdynamiclabelthresholdexceedclear** to enable the threshold exceeding alarm and clear alarm when configuring **trap-name**. When the usage of dynamic labels exceeds the upper threshold or falls below the lower threshold, the system generates a threshold exceeding alarm or clear alarm.
- To configure the LSP usage alarm function, specify **hwmplslspthresholdexceed** and **hwmplslspthresholdexceedclear** to enable the threshold exceeding alarm and clear alarm when configuring **trap-name**. When the LSP usage exceeds the upper threshold or falls below the lower threshold, the system generates a threshold exceeding alarm or clear alarm.
- To configure the alarm function for MPLS resource usage, specify **hwmplsresourcethresholdexceed** and **hwmplsresourcethresholdexceedclear** to enable the threshold exceeding alarm and clear alarm when configuring **trap-name**. When the usage of MPLS resources exceeds the upper threshold or falls below the lower threshold, the system generates a threshold exceeding alarm or clear alarm.

c. Run **mpls**

The MPLS view is displayed.

d. Run **mpls dynamic-label-number threshold-alarm upper-limit upper-limit-value lower-limit lower-limit-value**

The upper and lower thresholds for dynamic label usage are set.

You can set the following parameters:

- *upper-limit-value*: a percent indicating the upper limit of dynamic labels. If dynamic label usage reaches the upper limit, an alarm is generated. An upper limit less than or equal to 95% is recommended.
- *lower-limit-value*: a percent indicating the lower limit of dynamic labels. If dynamic label usage falls below the lower limit, a clear alarm is generated.
- The value of *upper-limit-value* must be greater than that of *lower-limit-value*.

By default, the upper limit is 80%, and the lower limit is 70%, which are recommended.

After the **snmp-agent trap enable feature-name mpls_lspm trap-name { hwmplsdynamiclabeltotalcountexceed | hwmplsdynamiclabeltotalcountexceedclear }** command is run to enable the device to generate limit-reaching alarms and their clear alarms, the following situations occur:

- If the number of dynamic labels reaches the maximum number of dynamic labels supported by a device, a limit-reaching alarm is generated.
 - If the number of dynamic labels falls below 95% of the maximum number of dynamic labels supported by the device, a clear alarm is generated.
- e. Run **mpls ldp-lsp-number threshold-alarm upper-limit *upper-limit-value* lower-limit *lower-limit-value***

The upper and lower thresholds of alarms for LDP LSP usage are configured.

The parameters in this command are described as follows:

- *upper-limit-value* specifies the upper threshold of alarms for LDP LSP usage. An alarm is generated when the proportion of established LDP LSPs to total supported LDP LSPs reaches the upper limit.
- *lower-limit-value* specifies the lower threshold of clear alarms for LDP LSP usage. A clear alarm is generated when the proportion of established LDP LSPs to total supported LDP LSPs falls below the lower limit.
- The value of *upper-limit-value* must be greater than that of *lower-limit-value*.

By default, the upper limit of the alarm threshold for LDP LSP usage is 80%, the lower limit of the clear alarm threshold for LDP LSP usage is 75%, which are recommended.

After the **snmp-agent trap enable feature-name mpls_lspm trap-name { hwmplslsptotalcountexceed | hwmplslsptotalcountexceedclear }** command is run to enable LSP limit-crossing alarm and LSP limit-crossing clear alarm, an alarm is generated in the following situations:

- If the total number of LDP LSPs reaches the upper limit, a limit-crossing alarm is generated.
 - If the total number of LDP LSPs falls below 95% of the upper limit, a limit-crossing clear alarm is generated.
- f. Run **mpls remote-adjacency-number threshold-alarm upper-limit *upper-limit-value* lower-limit *lower-limit-value***

The conditions that trigger the threshold-reaching alarm and its clear alarm for remote LDP adjacencies are configured.

Note the following issues when configuring trigger conditions:

- *upper-limit-value*: upper alarm threshold for the proportion of established remote LDP adjacencies to all supported ones.
- *lower-limit-value*: lower alarm threshold for the proportion of established remote LDP adjacencies to all supported ones.
- The value of *upper-limit-value* must be greater than that of *lower-limit-value*.

By default, the upper alarm threshold is 80%, and the lower alarm threshold is 75%, which are recommended.

When the number of remote LDP adjacencies exceeds the upper threshold or falls below the lower threshold, the system generates a threshold exceeding alarm or clear alarm.

- g. Run **mpls outsegment-number threshold-alarm upper-limit *upper-limit-value* lower-limit *lower-limit-value***

The alarm threshold for LDP OutSegment number is configured.

The parameters in this command are described as follows:

- *upper-limit-value* specifies the upper limit of the alarm threshold for LDP OutSegment number.
- *lower-limit-value* specifies the lower limit of the clear alarm threshold for LDP OutSegment number.
- The value of *upper-limit-value* must be greater than that of *lower-limit-value*.

By default, the upper limit of the alarm threshold for LDP OutSegment number is 80%, the lower limit of the clear alarm threshold for LDP OutSegment number is 75%, which are recommended.

When the LDP OutSegment number exceeds the upper threshold or falls below the lower threshold, the system generates a threshold exceeding alarm or clear alarm.

----End

Checking the Configuration

- Run the **display snmp-agent trap feature-name ldp all** command to check the status of all traps on the MPLS LDP module.
- Run the **display snmp-agent trap feature-name mpls lspm all** command to check all trap messages on the MPLS LSPM module.
- Run the **display default-parameter mpls management** command to check default configurations of the MPLS management module.

1.4.5 (Optional) Configuring the Function of Removing MPLS Labels

Context

In some scenarios, an MPLS-incapable device needs to be able to remove MPLS labels to obtain IP packet information in MPLS packets and redirect the IP packet information to the analyzer for traffic analysis.

To meet this requirement, configure the function of removing MPLS labels on the interfaces of the device. Then the device offsets MPLS packets. If the packets match the traffic policy on the device, the device removes the MPLS labels from the packets and forwards the packets through the outbound interface.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Enter the interface view.

- Run **interface** *interface-type interface-number*
The Ethernet interface view is displayed.
- Run **interface Eth-trunk** *trunk-id*
The Eth-Trunk interface view is displayed.

Step 3 Run **mpls pop-label all**

The function of removing MPLS labels is enabled on the interface. By default, the function of removing MPLS labels is disabled.

NOTE

Only the CE5881, CE6820, CE6881, CE6881E, CE6881K, CE6863, CE6863K and CE6863E support this command.

 NOTE

- MPLS label removal is supported only for L3VPN MPLS+IPv4 (without Header Options) and L2VPN MPLS+ETH+IPv4 (without Header Options). These packets may be incorrectly identified as IPv4 packets during strong verification. A maximum of five layers of labels can be removed.
- MPLS label removal is not supported for MPLS+IPv4 (with Header Options) and MPLS+ETH (without the IP header) packets.
- MPLS labels with the control word cannot be removed. The control word is used in the ping trace, BFD, and VPLS scenarios to detect the PW connectivity. In practice, the control word is seldom used for L3VPN data traffic.
- MPLS label removal is supported only for MPLS UC packets (with the outer Ethernet header being 0x8847) but not for MPLS MC packets (with the outer Ethernet header being 0x8848).
- MPLS labels can be removed from MPLS packets only when MPLS label removal is enabled on inbound interfaces and the action of redirecting inner IPv4 packets matching an ACL rule to a physical interface or an LAG interface is configured.
- Only Layer 2 interfaces (physical interfaces and LAG interfaces) support MPLS label removal. VLANIF interfaces, sub-interfaces, Layer 3 main interfaces, stack interfaces, peer-link interfaces, and LAG member interfaces do not support this function.
- In the MPLS label removal scenario, an ACL rule **if-match-any** needs to be configured. This rule has the lowest priority, and packets matching this rule will be redirected to the specified physical interface or LAG interface. If this rule is not configured, the forwarding behavior of the packets that do not match ACLs is unpredictable.
- For the MPLS packets received by the interface that has MPLS label removal enabled, other services such as VLAN check, incoming statistics collection, and MAC address learning are processed normally.

Step 4 Run commit

The configuration is committed.

----**End**

2 MPLS LDP Configuration

This chapter describes how to configure Multiprotocol Label Switching Label Distribution Protocol (MPLS LDP). MPLS LDP defines the messages and procedures for distributing labels. MPLS LDP is used by Label Switching Routers (LSRs) to negotiate session parameters, distribute labels, and then establish Label Switched Paths (LSPs).

[2.1 Overview of MPLS LDP](#)

[2.2 Understanding MPLS LDP](#)

[2.3 Summary of MPLS LDP Configuration Tasks](#)

[2.4 Licensing Requirements and Limitations for MPLS LDP](#)

[2.5 Default Settings for MPLS LDP](#)

[2.6 Configuring Basic Functions of MPLS LDP](#)

[2.7 Configuring Static BFD to Detect an LDP LSP](#)

[2.8 Configuring Synchronization Between LDP and IGP](#)

[2.9 Configuring Auto LDP FRR](#)

[2.10 Configuring LDP GR Helper](#)

[2.11 Configuring LDP Security Mechanisms](#)

[2.12 Configuring Non-labeled Public Network Routes to Be Iterated to LSPs](#)

[2.13 Maintaining MPLS LDP](#)

[2.14 Configuration Examples for MPLS LDP](#)

[2.15 Troubleshooting MPLS LDP](#)

2.1 Overview of MPLS LDP

Definition

The Label Distribution Protocol (LDP) is a control protocol of Multiprotocol Label Switching (MPLS) that functions like a signaling protocol on a traditional network.

LDP classifies forwarding equivalence classes (FECs), distributes labels, and establishes and maintains label switched paths (LSPs). LDP defines messages used in the label distribution process as well as procedures for processing these messages.

Purpose

MPLS is highly scalable because it allows multiple labels in a packet and has a connection-oriented forwarding plane. This scalability enables an MPLS/IP network to provide a variety of services. Label switching routers (LSRs) on an MPLS network use LDP to map Layer 3 routing information to Layer 2 switched paths, and establish LSPs at the network layer.

LDP is widely used to provide VPN services because of its simple deployment and configuration, abilities to set up LSPs dynamically based on routing information, and support for a large number of LSPs.

2.2 Understanding MPLS LDP

2.2.1 Basic Concepts of LDP

LDP Peers

Two LSRs that use LDP to set up an LDP session and exchange label messages are LDP peers.

LDP peers learn labels from each other over the LDP session between them.

LDP Adjacency

When an LSR receives a Hello message from a peer, an LDP adjacency is set up between the two LSRs. Two types of LDP adjacencies are used:

- Local adjacency: adjacency discovered by multicasting a Hello message (link Hello message)
- Remote adjacency: adjacency discovered by unicasting a Hello message (targeted Hello message)

LDP maintains peer information based on adjacencies. The type of a peer depends on the type of its adjacency. A peer can be maintained by multiple adjacencies. If a peer is maintained by both local and remote adjacencies, the peer is a local-and-remote peer.

LDP Session

LSRs exchange messages over an LDP session that include label mapping and release messages. LDP sessions can be set up only between LDP peers. The following types of LDP sessions are available:

- Local LDP session: set up between two LSRs that are directly connected
- Remote LDP session: set up between two LSRs that are directly or indirectly connected

An LSR can set up local and remote LDP sessions simultaneously.

Differences and Relations Among the LDP Adjacency, Peer, and Session

Differences

Differences among the LDP adjacency, peer, and session are as follows:

- An LDP adjacency is a TCP connection established after two devices exchange Hello messages with each other. The LDP adjacency is based on a link between two interconnected interfaces.
- LDP peers refer to two devices that run LDP to exchange label messages over an established TCP connection.
- An LDP session is a series of processes of exchanging label messages between two LDP peers.

Relations

The association between LDP adjacencies, peers, and sessions is summarized as follows: Before setting up an LDP session, you need to establish a link that establishes a TCP connection. The link is an adjacency. After the adjacency is established, the two devices exchange label information to form a peer relationship. Finally, LDP sessions are established between peers. It may be specifically described as follows:

- LDP maintains the existence of the peers through adjacencies. The type of peer is determined by the type of the adjacency that maintains the peer.
- A peer can be maintained using multiple adjacencies. If a peer is maintained by both local and remote adjacencies, the peer is a local and remote coexistent peer.
- Only LDP peers can establish LDP sessions.

2.2.2 LDP Working Mechanism

2.2.2.1 LDP Overview

LDP defines the label distribution process and messages transmitted during label distribution. LSRs use LDP to map Layer 3 routing information to Layer 2 switched paths, and set up an LSP.

LDP Messages

LDP defines the following messages:

- Discovery message: used to announce or maintain an LSR on a network. For example, Hello messages are discovery messages.
- Session message: used to establish, maintain, and terminate sessions between LDP peers. For example, Initialization and Keepalive messages are session messages.
- Advertisement message: used to create, modify, and delete label mappings for FECs.
- Notification message: used to provide advisory and error information.

To ensure the reliability of message transmission, LDP uses Transmission Control Protocol (TCP) transport for Session, Advertisement, and Notification messages. LDP uses User Datagram Protocol (UDP) transport only for transmitting the Discovery message.

LDP Working Process

The LDP working process has the following phases:

1. LDP Session Setup

After LSRs send Hello messages to discover LDP peers, the LSRs establish an LDP session. LDP peers periodically send Hello and Keepalive messages to maintain the LDP session.

- LDP peers periodically send Hello messages to maintain the adjacency. If an LSR does not receive any Hello message from a peer before the Hello timer expires, the local LSR deletes the adjacency. After the adjacency is deleted, the local LSR sends a Notification message to terminate the LDP session.
- LDP peers periodically send Keepalive messages over the LDP session to maintain the LDP session. If an LSR does not receive a Keepalive message from a peer before the Keepalive timer expires, the local LSR terminates the TCP connection and sends a Notification message to terminate the LDP session.

2. LDP LSP Setup

After a session is established, LDP peers advertise the mappings between FECs and labels and set up an LSP.

2.2.2.2 LDP Session Setup

LSRs use LDP discovery mechanisms to discover LDP peers and establish an LDP session. An LDP LSP can be set up to transmit services only after an LDP session is set up.

LDP Discovery Mechanisms

LSRs use LDP discovery mechanisms to discover LDP peers. LSRs can use the following types of LDP discovery mechanisms:

- Basic discovery mechanism: used to discover directly-connected LSR peers on a link.
An LSR periodically sends LDP Link Hello messages to implement the mechanism and establish a local LDP session.
LDP Link Hello messages are encapsulated in UDP packets with the multicast destination address 224.0.0.2. If an LSR receives an LDP Link Hello message on an interface, an LDP peer is connected to this interface.
- Extended discovery mechanism: used to discover LSR peers that are not directly connected on a link.
An LSR periodically sends LDP Target Hello messages to a specified destination IP address according to the mechanism to establish a remote LDP session.

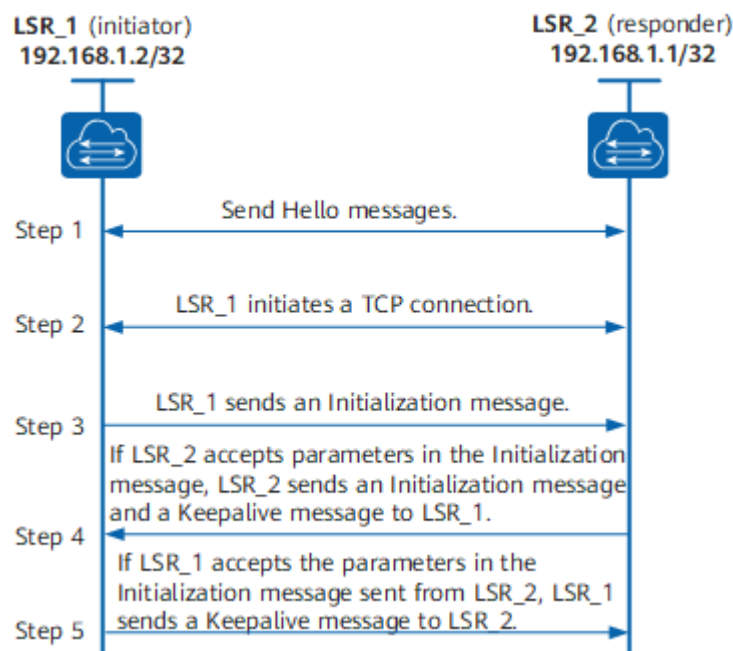
The LDP Target Hello messages are encapsulated in UDP packets with the unicast destination IP addresses. If an LSR receives an LDP Target Hello message, an LDP peer is connected to the LSR.

LDP Session Setup Process

Two LSRs exchange Hello messages to trigger the establishment of an LDP session.

Figure 2-1 shows how an LDP session is set up.

Figure 2-1 Process of establishing an LDP session



The LDP session setup process consists of the following steps:

- Two LSRs send Hello messages to each other.
Each Hello message contains the transport address (device IP address) that the two LSRs use to establish an LDP session.
- The LSR with a larger transport address initiates a TCP connection.
As shown in **Figure 2-1**, LSR_1 initiates a TCP connection and LSR_2 waits for the TCP connection request.
- After the TCP connection is successfully established, LSR_1 sends an Initialization message to negotiate with LSR_2 about parameters used for establishing the LDP session.
These parameters include the LDP version, label distribution mode, Keepalive timer value, maximum packet data unit (PDU) length, and label space.
- If LSR_2 accepts parameters in the Initialization message, LSR_2 sends an Initialization message and a Keepalive message to LSR_1.
If LSR_2 rejects the parameters in the Initialization message, LSR_2 sends a Notification message to LSR_1 to stop the establishment of the LDP session.

Parameters in the Initialization message include the LDP version, label distribution mode, Keepalive timer value, maximum PDU length, and label space.

5. If LSR_1 accepts the parameters in the Initialization message sent from LSR_2, LSR_1 sends a Keepalive message to LSR_2.

If LSR_1 rejects the parameters in the Initialization message, LSR_1 sends a Notification message to LSR_2 to stop the establishment of the LDP session.

After both LSR_1 and LSR_2 have accepted Keepalive messages from each other, an LDP session is established between them.

2.2.2.3 LDP LSP Setup

LDP peers send Label Request and Mapping messages to advertise FEC-to-label mappings and establish LSPs based on the mappings. Label distribution and management depend on the label advertisement mode, label distribution control mode, and label retention mode.

Label Advertisement and Management

Label Advertisement Modes

An LSR assigns a label to a specified FEC and notifies its upstream LSR of the label. This means that labels are allocated by a downstream LSR and distributed from downstream to upstream.

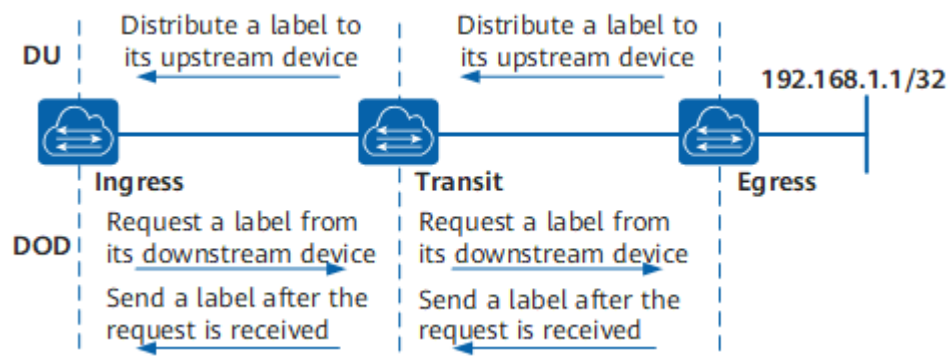
As described in [Table 2-1](#), two label advertisement modes are available. The label advertisement modes on upstream and downstream LSRs must be the same.

Table 2-1 Label advertisement modes

Label Advertisement Mode	Definition	Description
Downstream Unsolicited (DU) mode	An LSR distributes labels to a specified FEC without having to receive Label Request messages from its upstream LSR. When the DU mode is used, LDP supports label distribution for all non-downstream peers by default. Each node can send Label Mapping messages to all non-downstream peers.	As shown in Figure 2-2 , the downstream egress triggers the establishment of an LSP destined for the FEC 192.168.1.1/32 using a host route and sends a Label Mapping message to the upstream transit node to advertise the label of the host route to 192.168.1.1/32.

Label Advertisement Mode	Definition	Description
Downstream on Demand (DoD) mode	An LSR distributes labels to a specified FEC only after receiving Label Request messages from its upstream LSR.	As shown in Figure 2-2 , the downstream egress triggers the establishment of an LSP destined for the FEC 192.168.1.1/32 in host mode. The upstream ingress sends a Label Request message to the downstream egress. After receiving the message, the downstream egress sends a Label Mapping message to the upstream LSR.

Figure 2-2 DU and DoD



Label Distribution Control Modes

The label distribution control mode refers to a label distribution method used on an LSR during LSP establishment.

Two label distribution control modes are available.

- Independent mode: A local LSR can distribute a label bound to an FEC and then inform the upstream LSR, without waiting for the label distributed by the downstream LSR.
- Ordered mode: An LSR advertises the mapping between a label and an FEC to its upstream LSR only when this LSR is the outgoing node of the FEC or receives the Label Mapping message of the next hop for the FEC.

[Table 2-2](#) describes the combination between the label distribution control mode and label advertisement mode.

Table 2-2 Combination between the label distribution control mode and label advertisement mode

Label Distribution Control Mode	DU Mode	DoD Mode
Independent Mode	A transit LSR can assign a label to the ingress node without waiting for the label assigned by the egress node.	The directly-connected ingress transit node that sends a Label Request message replies with a label without waiting for the label assigned by the egress node.
Ordered Mode	The LSR (the transit LSR in Figure 2-2) must receive a Label Mapping message from the downstream LSR (the egress node in Figure 2-2). Then, the transit LSR can distribute a label to the ingress node in the diagram.	The directly connected transit node of the ingress node that sends the Label Request message must receive a Label Mapping message from the downstream (the egress node in the diagram). Then, the transit node can distribute a label to the ingress node in Figure 2-2 .

Label Retention Modes

The label retention mode refers to the way an LSR processes the label mapping received but not immediately used.

The label mapping that an LSR receives may or may not originate at the next hop.

As described in [Table 2-3](#), two label retention modes are available.

Table 2-3 Label retention modes

Label Retention Mode	Definition	Description
Liberal mode	Upon receiving a Label Mapping message from a neighbor LSR, an LSR retains the message regardless of whether the neighbor LSR is its next hop.	When the next hop of an LSR changes due to a change in network topology, note that: • In Liberal mode, the LSR can use the previous label sent by a non-next hop to quickly reestablish an LSP. This requires more memory and

Label Retention Mode	Definition	Description
Conservative mode	Upon receiving a Label Mapping message from a neighbor LSR, an LSR retains the message only when the neighbor LSR is its next hop.	label space than in conservative mode. • In Conservative mode, the LSR only retains labels sent by the next hop. This saves memory and label space but slows down the reestablishment of an LSP. Conservative mode and DoD mode are used together to set up LSRs with limited label space.

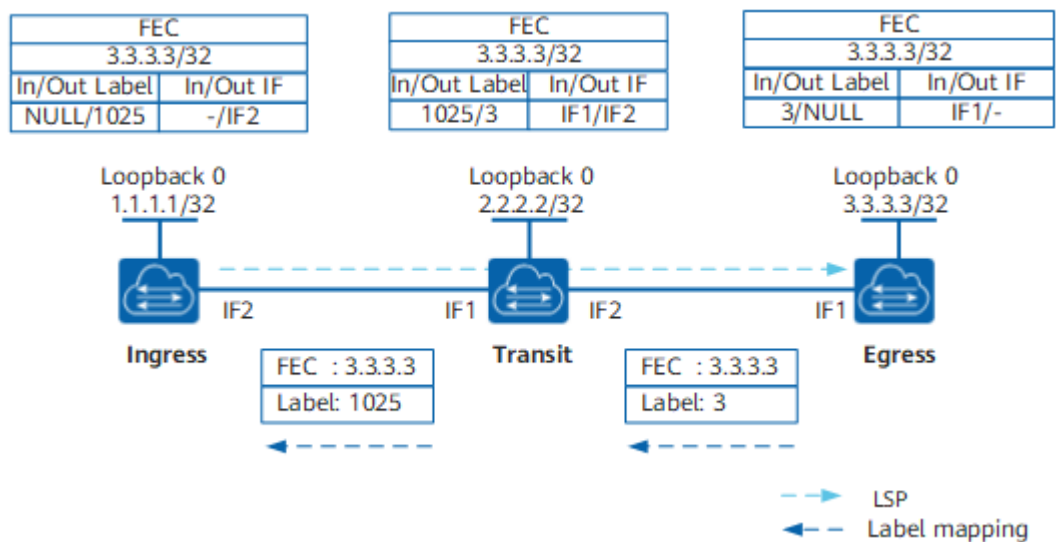
Currently, the following combinations are supported:

- DU label advertisement mode, ordered label control mode, and liberal label retention mode (default mode)

LDP LSP Setup Process

LSP setup is the process of mapping an FEC to a label and advertising the mapping to neighboring LSRs. **Figure 2-3** shows how an LSP is set up in DU and Ordered mode.

Figure 2-3 LDP LSP setup process



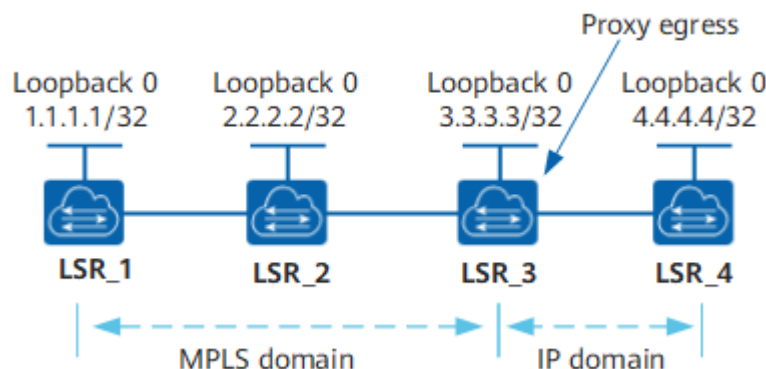
The LDP LSP setup process consists of the following steps:

1. By default, upon a route change, if an edge node (egress) finds a new host route that does not belong to any existing FEC, the egress node creates an FEC for the route.
2. If the egress node has available labels, it distributes a label for the new FEC and sends a Label Mapping message to the upstream node. The Label Mapping message contains the distributed label and FEC.
3. After receiving the Label Mapping message, the upstream transit node checks whether the sender (egress node) is the next hop of the FEC. If so, the transit node adds the label-to-FEC mapping in the Label Mapping message to its label forwarding table, and then sends the Label Mapping message of the specified FEC to the upstream LSR (ingress).
4. After receiving the Label Mapping message, the ingress node checks whether the sender (transit node) is the next hop of the FEC. If so, the ingress node adds the label-to-FEC mapping in the Label Mapping message to its label forwarding table. An LSP is established, and the packets of this FEC can be forwarded based on labels.

This process sets up a common LDP LSP. Another type of LSP is the proxy egress LSP. A proxy egress establishes LSPs using routes in which the next-hop addresses are not local addresses. If penultimate hop popping (PHP) is enabled, an LSR at the penultimate hop is a specific proxy egress along an LSP. A proxy egress is configured manually. The proxy egress applies to a network with MPLS-incapable switches or helps load-balance traffic based on Border Gateway Protocol (BGP) routes.

On the network shown in [Figure 2-4](#), LSR_1, LSR_2, and LSR_3 are in an MPLS domain. LSR_4 is not enabled with MPLS LDP or does not support MPLS LDP. If a policy is configured to use all IGP routes to establish LDP LSPs, LSR_3 functions as a proxy egress and becomes the penultimate hop of the route. This allows LSR_1, LSR_2, and LSR_3 to establish LDP LSPs to LSR_4.

Figure 2-4 Proxy egress



2.2.3 LDP Security Mechanisms

To ensure security of LDP packets, MPLS provides three security mechanisms: Message-digest algorithm 5 (MD5), Keychain, and Generalized TTL Security Mechanism (GTSM).

Keychain is more secure than MD5 authentication, and only one of these mechanisms can be used for an LDP peer. GTSM protects a device against attacks of invalid LDP packets and can be used with MD5 authentication or Keychain.

MD5 Authentication

MD5 authentication is a standard digest algorithm defined in RFC. A typical application of MD5 is to calculate a message digest to prevent message spoofing. The MD5 message digest is a unique result calculated by an irreversible character string conversion. If a message is modified during transmission, a different digest is generated. After the message arrives at the receiver, the receiver can determine whether the packet has been modified by comparing the received digest with the pre-calculated digest.

MD5 generates a unique digest for an information segment, so LDP MD5 authentication can prevent LDP packets from being modified. This authentication is stricter than common checksum verification of TCP. The MD5 authentication process is as follows:

1. Before an LDP session message is sent over a TCP connection, the sender pads the TCP header with a unique digest. The digest is calculated using the MD5 algorithm based on the TCP header, LDP message, and configured password.
2. Upon receiving the TCP packet, the receiver obtains the TCP header, digest, and LDP session message, and then uses MD5 to calculate a digest based on the received TCP header, LDP session message, and locally stored password. The receiver compares the calculated digest with the received one to check whether the packet has been modified.

A password can be set in either cipher text or simple text. The simple-text password is directly saved in the configuration file. The cipher-text password is saved in the configuration file after being encrypted using a special algorithm. However, the character string entered by the user is used to calculate the digest, regardless of whether the password is in simple text or cipher text. That is, the cipher-text password does not participate in MD5 calculation. As devices from different vendors use proprietary password encryption algorithms, this digest calculation method shields differences of password encryption algorithms used on different devices.

Keychain Authentication

Compared with MD5, Keychain is an enhanced encryption algorithm that calculates a message digest for the same LDP message to prevent the message from being modified.

During Keychain authentication, a group of passwords are defined to form a password string. Each password is specified with encryption and decryption algorithms such as MD5 algorithm and SHA-1, and is configured with the validity period. When sending or receiving a packet, the system selects a valid password based on the user's configuration. Within the valid period of the password, the system uses the encryption algorithm matching the password to encrypt the packet before sending it out, or uses the decryption algorithm matching the password to decrypt the packet before accepting it. In addition, the system automatically uses a new password after the previous password expires, preventing the password from being decrypted.

The Keychain authentication password, the encryption and decryption algorithms, and the password validity period that constitute a Keychain configuration node are configured using different commands. A Keychain configuration node requires at least one password and encryption and decryption algorithms.

LDP GTSM

GTSM protects services by checking whether the TTL value in the IP header is within the pre-defined range. The prerequisites for using GTSM are as follows:

- The TTL of normal packets between routers is determined.
- It is difficult to change the TTL value.

LDP GTSM refers to GTSM implementation over LDP.

To protect the device against attacks, GTSM verifies the TTL in a packet. LDP GTSM is applied to LDP packets between neighbor or adjacent (based on a fixed number of hops) devices. The TTL range is preset on each device for packets from other devices. With GTSM enabled, if the TTL of an LDP packet received by a device configured with LDP is out of the TTL range, the packet is considered invalid and is discarded. This protects the upper-layer protocols.

2.2.4 LDP Reliability

2.2.4.1 Overview of LDP Reliability

MPLS provides multiple reliability technologies to ensure high reliability of key services transmitted over LDP LSPs. The following table describes these reliability technologies.

Table 2-4 LDP reliability technologies

Reliability Technology	Description	Function
Fault detection	Rapidly detects faults on LDP LSPs of an MPLS network and triggers protection switching.	• 2.2.4.2 BFD for LDP LSP
Traffic protection	Ensures that traffic is switched to the backup LDP LSP and minimizes packet loss when a working LDP LSP fails.	• 2.2.4.3 LDP-IGP Synchronization • 2.2.4.4 Auto LDP FRR • 2.2.4.7 Local-and-Remote LDP Session
	Ensures nonstop forwarding on the forwarding plane when the control plane fails on a node.	• 2.2.4.5 LDP GR • 2.2.4.6 LDP NSR

2.2.4.2 BFD for LDP LSP

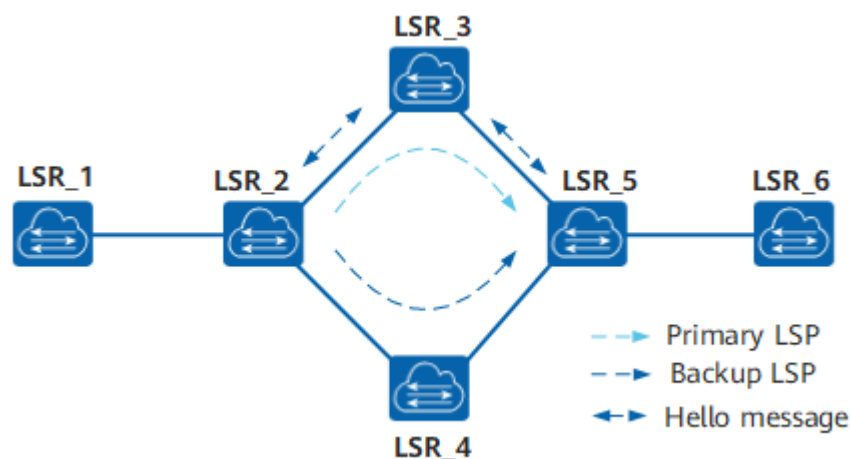
Bidirectional Forwarding Detection (BFD) can quickly detect faults on an LDP LSP and trigger a traffic switchover upon an LDP LSP failure to improve network reliability.

Background

If a node or link along a working LDP LSP fails, traffic is switched to the backup LSP. The path switchover time depends on the speed of fault detection and traffic switching. A slow path switchover causes long-time traffic loss. Fast traffic switching can be ensured by **Auto LDP FRR**. Because the fault detection mechanism of LDP is slow, however, LDP FRR alone cannot solve the preceding problem.

As shown in **Figure 2-5**, an LSR periodically sends Hello messages to its neighboring LSRs to advertise its existence on the network and maintain adjacencies. An LSR creates a Hello timer for each neighbor to maintain an adjacency. Each time the LSR receives a Hello message, the LSR resets the Hello timer. If the Hello timer expires before the LSR receives a new Hello message, the LSR considers that the adjacency is terminated. This mechanism cannot detect link faults quickly, especially when a Layer 2 device is deployed between LSRs.

Figure 2-5 BFD for LDP LSPs



BFD can quickly detect faults on an LDP LSP and trigger a traffic switchover upon an LDP LSP failure, minimizing packet loss and improving network reliability.

Implementation

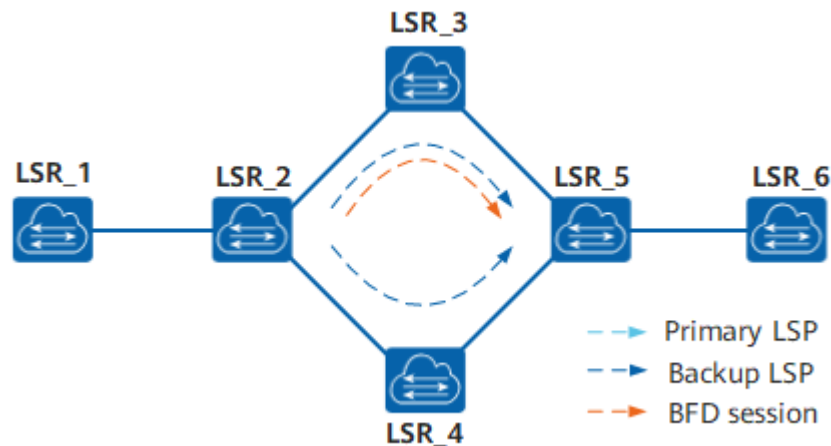
BFD for LDP LSP is implemented by establishing a BFD session between two nodes on both ends of an LSP and binding the session to the LSP. BFD rapidly detects LSP faults and triggers a traffic switchover. When BFD monitors a unidirectional LDP LSP, the reverse path of the LDP LSP can be an IP link, an LDP LSP, or a traffic engineering (TE) tunnel.

A BFD session that monitors LDP LSPs is negotiated in static mode: The negotiation of a BFD session is performed using the local and remote

discriminators that are manually configured for the BFD session to be established. On a local LSR, you can bind an LSP with a specified next-hop IP address to a BFD session with a specified peer IP address.

BFD uses the asynchronous mode to check LSP continuity. That is, the ingress and egress periodically send BFD packets to each other. If one end does not receive BFD packets from the other end within a detection period, BFD considers the LSP Down and sends an LSP Down message to the LSP management (LSPM) module.

Figure 2-6 BFD for LDP LSPs



2.2.4.3 LDP-IGP Synchronization

Context

LDP-IGP synchronization is used to synchronize the states of LDP and an IGP so as to minimize the traffic loss time if a fault occurs on the network.

On a network with primary and backup links, if the primary link fails, both IGP and LSP traffic is switched to the backup link. After the primary link recovers, IGP routes are switched back to the primary link before LDP convergence is complete. Before being established, the LSP along the primary link takes time to make preparations, such as adjacency restoration, which leads to traffic drop. If an LDP session or adjacency on the primary link fails, the LSP along the primary link is deleted. However, the IGP still uses the primary link. As a result, LSP traffic cannot be switched to the backup link, and is continuously dropped.

NOTE

LDP-IGP synchronization supports only OSPFv2 and IS-IS in IPv4.

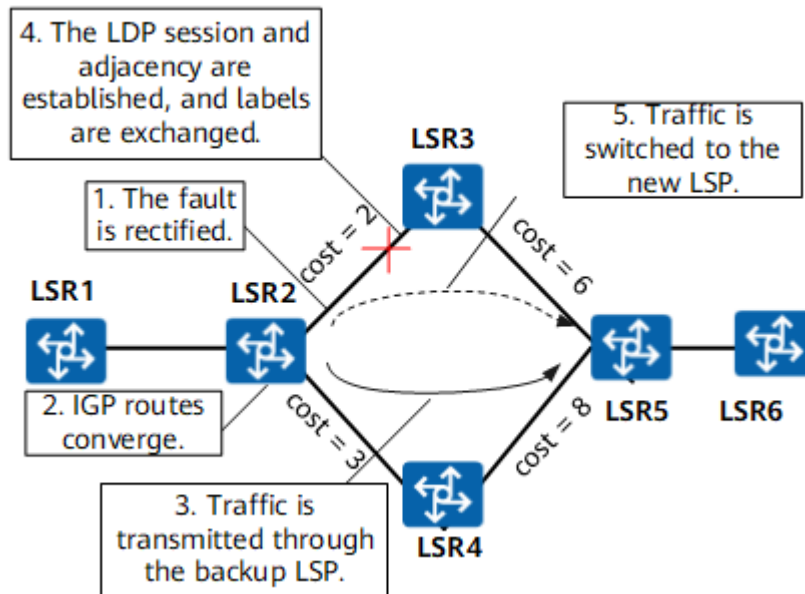
LDP-IGP synchronization allows you to set an IGP cost to delay a route switchback until LDP convergence is complete. Specifically, before the LSP on the primary link is established, the LSP on the backup link is retained to continue forwarding traffic. The LSP on the backup link is torn down only after the LSP on the primary link is established successfully.

LDP-IGP synchronization timers are as follows:

- Hold-max-cost timer
- Delay timer

Implementation

Figure 2-7 Switchback with LDP-IGP synchronization enabled



- The network shown in [Figure 2-7](#) has primary and backup links. When the primary link recovers from a fault, traffic is switched from the backup link back to the primary link. During the traffic switchback, the backup LSP cannot be used once IGP route convergence is complete. As a result, before a new LSP is set up over the primary link, traffic is dropped. To prevent this issue, LDP-IGP synchronization can be configured to delay IGP route switchback until LDP convergence is complete. Before convergence of the primary LSP completes, the backup LSP is retained to continue to forward traffic. The backup LSP is torn down after the primary LSP is successfully established. The process of LDP-IGP synchronization is as follows:
 - The link fault is rectified.
 - The IGP advertises the maximum cost of the primary link, delaying the IGP route switchback.
 - Traffic is still forwarded over the backup LSP.
 - After the LDP session and adjacency are established, Label Mapping messages are exchanged to instruct the IGP to start synchronization.
 - The IGP advertises the normal cost of the primary link, and its routes converge to the original forwarding path. The LSP is reestablished and entries are delivered to the forwarding table (within milliseconds).
- If an LDP session or adjacency on the primary link fails, the LSP along the primary link is deleted. However, the IGP still uses the primary link. As a result, LSP traffic cannot be switched to the backup link, and is continuously dropped. To prevent this problem, you can configure LDP-IGP synchronization.

If an LDP session or adjacency fails, LDP informs the IGP that the LDP session or adjacency is faulty. In this case, IGP advertises the maximum cost of the faulty link. In this way, both the route and LSP are switched to the backup link. The process is as follows:

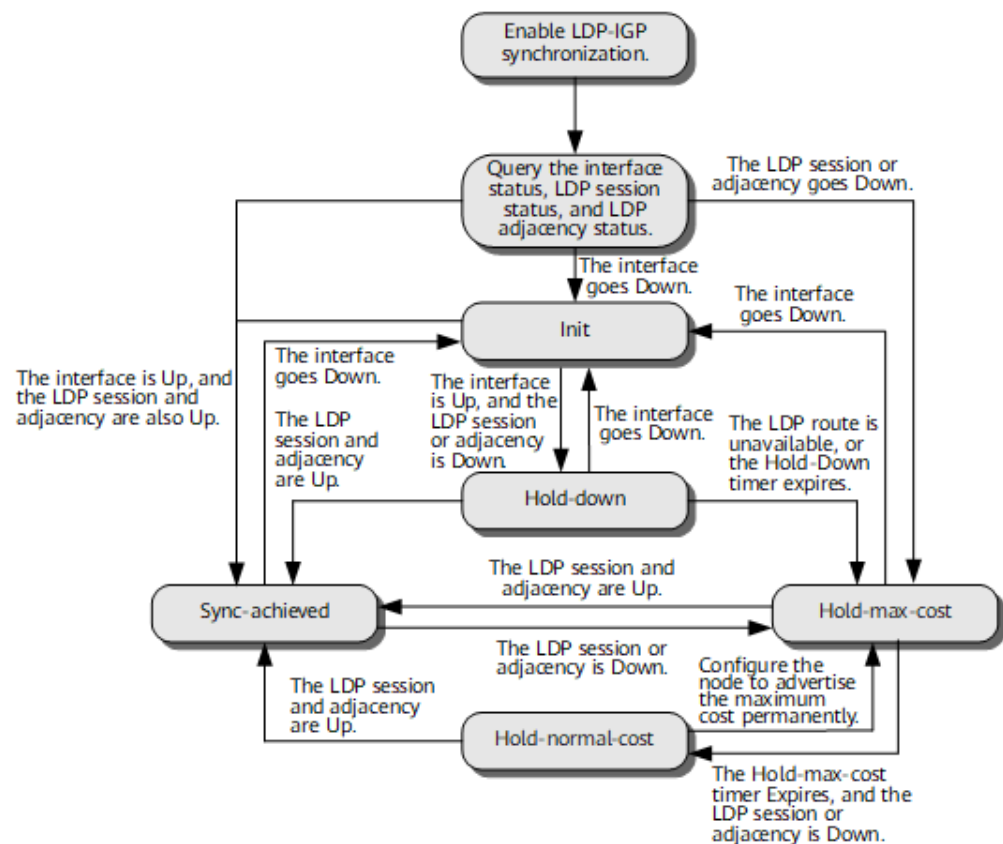
- The LDP session or adjacency between nodes on the primary link is faulty.
- LDP informs the IGP that the LDP session or adjacency along the primary link is faulty. IGP advertises the maximum cost of the primary link.
- The IGP route switches to the backup link.
- An LSP is set up over the backup link, and then forwarding entries are delivered.

To prevent a continuous failure to reestablish the LDP session or adjacency, you can configure the Hold-max-cost timer to permanently advertise the maximum cost so that traffic is always transmitted over the backup link until the LDP session and LDP adjacency are reestablished on the primary link.

- LDP-IGP synchronization state transition mechanism

After LDP-IGP synchronization is enabled on an interface, the IGP queries the state of the interface, LDP session, and LDP adjacency according to the process shown in [Figure 2-8](#), enters the corresponding state according to the query result, and then transitions the state according to [Figure 2-8](#).

Figure 2-8 LDP-IGP synchronization state transition



NOTE

Note the differences when different IGP protocols are used:

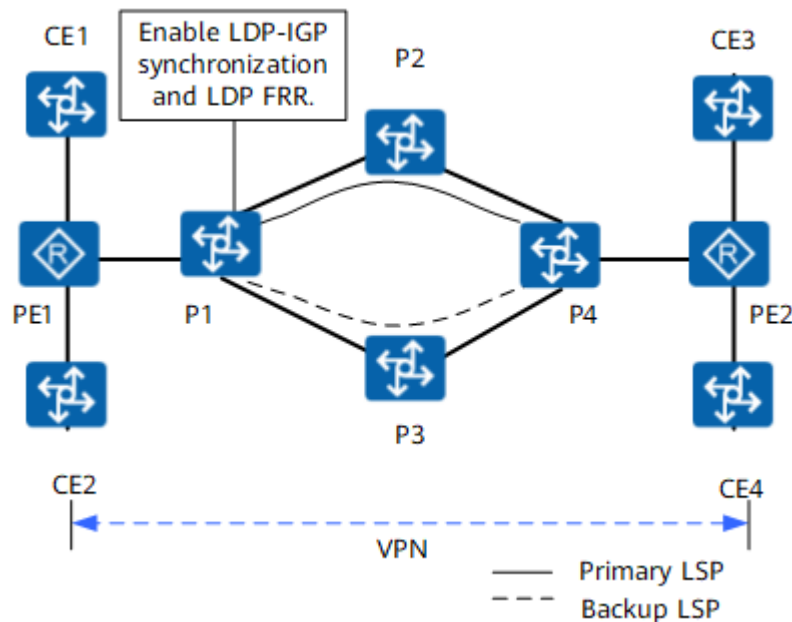
- When OSPF is used, the state transits based on the flowchart shown in [Figure 2-8](#).
- When IS-IS is used, the Hold-normal-cost state does not exist. After the Hold-max-cost timer expires, IS-IS advertises the normal link cost, but the Hold-max-cost state is displayed even though this state does not exist.

Usage Scenario

LDP-IGP synchronization applies to the following scenario:

On the network shown in [Figure 2-9](#), primary and backup links are connected. Both LDP-IGP synchronization and LDP FRR can be deployed.

Figure 2-9 LDP-IGP synchronization



Benefits

This feature reduces the packet loss rate during a primary/backup link switchover and improves the reliability of the entire network.

2.2.4.4 Auto LDP FRR

Auto LDP fast reroute (FRR) provides link backup on an MPLS network. When the primary LSP fails, traffic is quickly switched to the backup LSP, minimizing traffic loss.

Background

On an MPLS network, when the primary link fails, IP FRR ensures fast IGP route convergence and switches traffic to the backup link. However, a new LSP needs to

be established, which causes traffic loss. If the LSP fails (for some reason other than a primary link failure), traffic is restored until a new LSP is established, causing traffic interruption for a long time. Auto LDP FRR is used on an MPLS network to address these issues.

Auto LDP FRR, using the liberal label retention mode of LDP, obtains a liberal label, assigns a forwarding entry to the label, and then delivers the forwarding entry to the forwarding plane as the backup forwarding entry for the primary LSP. When the interface goes Down (as detected by the interface itself or by BFD) or the primary LSP fails (as detected by BFD), traffic is quickly switched to the backup LSP.

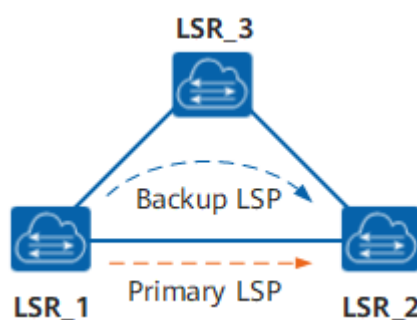
Concepts

Auto LDP FRR: This automatic approach depends on IP FRR. A backup LSP can be established and its forwarding entries can be delivered only when the source of the liberal label matches the backup route. That is, the liberal label is obtained from the outbound interface and next hop of the backup route, and the backup LSP triggering conditions are met. By default, LDP LSP setup is triggered by a 32-bit backup route.

Implementation

In liberal label retention mode, an LSR can receive a Label Mapping message of an FEC from any neighboring LSR. However, only the Label Mapping message sent by the next hop of the FEC can be used to generate a label forwarding table for LSP setup. In contrast, Auto LDP FRR can generate an LSP as the backup of the primary LSP based on Label Mapping messages that are not from the next hop of the FEC. Auto LDP FRR establishes forwarding entries for the backup LSP and adds the forwarding entries to the forwarding table. If the primary LSP fails, traffic is switched to the backup LSP quickly to minimize traffic loss.

Figure 2-10 Auto LDP FRR - triangle topology



In [Figure 2-10](#), the optimal route from LSR_1 to LSR_2 is LSR_1-LSR_2. A suboptimal route is LSR_1-LSR_3-LSR_2. After receiving a label from LSR_3, LSR_1 compares the label with the route from LSR_1 to LSR_2. Because LSR_3 is not the next hop of the route from LSR_1 to LSR_2, LSR_1 stores the label as a liberal label. If a route is available for the source of the liberal label, LSR_1 assigns a forwarding entry to the liberal label as the backup forwarding entry, and then

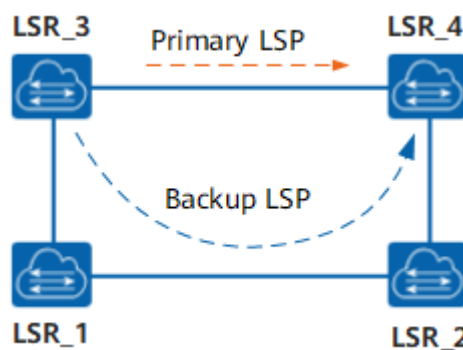
delivers this forwarding entry to the forwarding plane with the primary LSP. In this way, the primary LSP is associated with the backup LSP.

Auto LDP FRR is triggered when an interface failure is detected by the interface itself or BFD, or a primary LSP failure is detected by BFD. After Auto LDP FRR is complete, traffic is switched to the backup LSP using the backup forwarding entry. Then the route is converged from LSR_1-LSR_2 to LSR_1-LSR_3-LSR_2. An LSP is established on the new path (the original backup LSP) and the original primary LSP is deleted. Traffic is forwarded along the new LSP of LSR_1-LSR_3-LSR_2.

Usage Scenario

Figure 2-10 shows a typical application environment of Auto LDP FRR. Auto LDP FRR functions well in a triangle topology but may not take effect in some situations in a rectangle topology.

Figure 2-11 Auto LDP FRR - rectangle topology



As shown in **Figure 2-11**, if the optimal route from LSR_1 to LSR_4 is LSR_1-LSR_2-LSR_4 (with no other route for load balancing), LSR_3 receives a liberal label from LSR_1 and is bound to Auto LDP FRR. If the link between LSR_3 and LSR_4 fails, traffic is switched to the route of LSR_3-LSR_1-LSR_2-LSR_4. No loop occurs in this situation.

However, if optional routes from LSR_1 to LSR_4 are available for load balancing (LSR_1-LSR_2-LSR_4 and LSR_1-LSR_3-LSR_4), LSR_3 may not receive a liberal label from LSR_1 because LSR_3 is a downstream node of LSR_1. Although LSR_3 receives a liberal label and is configured with Auto LDP FRR, traffic may still be forwarded to LSR_3 after the traffic switching, leading to a loop. The loop exists until the route from LSR_1 to LSR_4 is converged to LSR_1-LSR_2-LSR_4.

2.2.4.5 LDP GR

LDP Graceful Restart (GR) ensures uninterrupted traffic transmission during a protocol restart or active/standby switchover because the forwarding plane is separated from the control plane.

Background

On an MPLS network, when the GR Restarter restarts a protocol or performs an active/standby switchover, label forwarding entries on the forwarding plane are deleted, interrupting data forwarding.

LDP GR can address this issue and therefore improve network reliability. During a protocol restart or active/standby switchover, LDP GR retains label forwarding entries because the forwarding plane is separated from the control plane. The device still forwards packets based on the label forwarding entries, ensuring data transmission. After the protocol restart or active/standby switchover is complete, the GR Restarter can restore to the original state with the help of the GR Helper.

Concepts

LDP GR is a high-reliability technology based on non-stop forwarding (NSF). A GR process involves GR Restarter and GR Helper devices:

- GR Restarter: has GR capability and restarts a protocol.
- GR Helper: assists in the GR process as a GR-capable neighbor of the GR Restarter.

NOTE

The device can function only as the GR Helper.

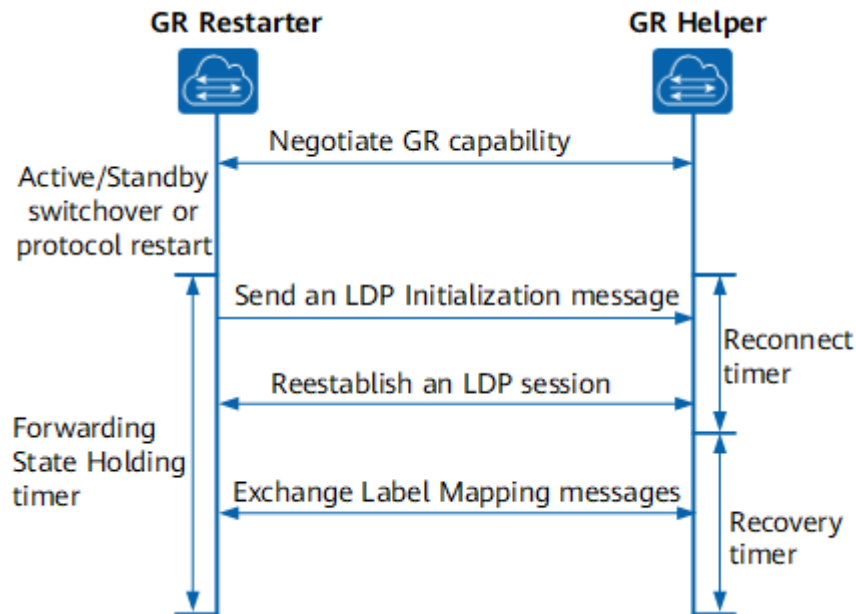
LDP GR uses the following timers:

- Forwarding State Holding timer: specifies the duration of the LDP GR process.
- Reconnect timer: controls the time during which the GR Helper waits for LDP session reestablishment. After a protocol restart or active/standby switchover occurs on the GR Restarter, the GR Helper detects that the LDP session with the GR Restarter is Down. The GR Helper then starts this timer and waits for the LDP session to be reestablished before the timer expires.
- Recovery timer: controls the time during which the GR Helper waits for LSP recovery. After the LDP session is reestablished, the GR Helper starts this timer and waits for the LSP to recover before the timer expires.

Implementation

[Figure 2-12](#) shows LDP GR implementation.

Figure 2-12 LDP GR implementation



LDP GR works as follows:

1. An LDP session is set up between the GR Restarter and GR Helper. The GR Restarter and GR Helper negotiate GR capabilities during LDP session setup.
2. When restarting a protocol or performing an active/standby switchover, the GR Restarter starts the Forwarding State Holding timer, retains label forwarding entries, and sends an LDP Initialization message to the GR Helper. When the GR Helper detects that the LDP session with the GR Restarter is Down, the GR Helper retains label forwarding entries of the GR Restarter and starts the Reconnect timer.
3. After the protocol restart or active/standby switchover, the GR Restarter reestablishes an LDP session with the GR Helper. If an LDP session is not reestablished before the Reconnect timer expires, the GR Helper deletes label forwarding entries of the GR Restarter.
4. After the GR Restarter reestablishes an LDP session with the GR Helper, the GR Helper starts the Recovery timer. Before the Recovery timer expires, the GR Restarter and GR Helper exchange Label Mapping messages over the LDP session. The GR Restarter restores forwarding entries with the help of the GR Helper, and the GR Helper restores forwarding entries with the help of the GR Restarter. After the Recovery timer expires, the GR Helper deletes all forwarding entries that have not been restored.
5. After the Forwarding State Holding timer expires, the GR Restarter deletes label forwarding entries and the GR is complete.

2.2.4.6 LDP NSR

LDP Non-Stop Routing (LDP NSR) ensures nonstop data transmission on the control plane and forwarding plane when an active/standby switchover occurs on a device, without the help of neighboring nodes. For details about NSR, see NSR in

CloudEngine 9800, 8800, 6800, 5800 Series Switches Configuration Guide - Reliability.

NOTE

LDP NSR is enabled on the device by default and does not need to be configured.

LDP NSR is implemented through synchronization between the active and standby control boards. LDP NSR synchronizes the following key data between the active and standby control boards:

- LDP control block
- LSP forwarding entries
- Cross connect (XC) information that describes the cross connection between a forwarding equivalence class (FEC) and an LSP
- Labels, including the following types:
 - LDP LSP labels on a public network
 - Virtual Circuit (VC) labels in Martini Virtual Private LAN Service (VPLS) networking

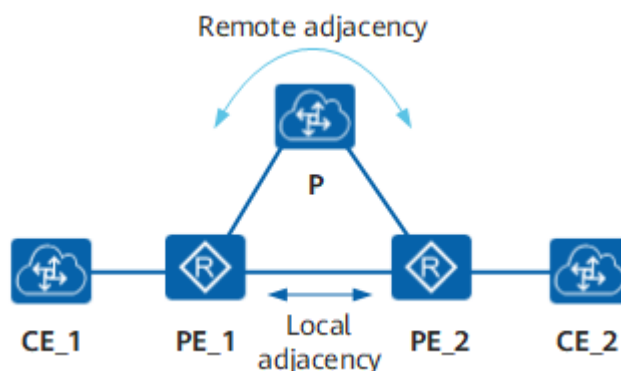
2.2.4.7 Local-and-Remote LDP Session

A local node can set up both local and remote LDP adjacencies with an LDP peer. That is, the peer is maintained by both local and remote LDP adjacencies.

As shown in [Figure 2-13](#), when the local LDP adjacency is deleted because the link associated with the adjacency fails, the type of the peer may change but the peer status remains unchanged. Depending on the adjacency type, the peer type can be local, remote, or local-and-remote.

When the link is faulty or recovering, the peer type may change as well as the corresponding session type. However, the session stays Up in this process and is not deleted or set to Down.

Figure 2-13 Local-and-remote LDP session between two nodes



A typical application of local-and-remote LDP is with a Layer 2 virtual private network (L2VPN). As shown in [Figure 2-13](#), L2VPN services are configured on PE_1 and PE_2. When the direct link between PE_1 and PE_2 is disconnected and then recovers, the changes in the peer and session types are as follows:

1. PE_1 and PE_2 have MPLS LDP enabled and establish a local LDP session. Then PE_1 and PE_2 are configured as remote peers and establish a remote LDP session. PE_1 and PE_2 maintain both local and remote adjacencies. In this case, a local-and-remote LDP session exists between PE_1 and PE_2. L2VPN messages are transmitted over this LDP session.
2. When the physical link between PE_1 and PE_2 goes Down, the local LDP adjacency goes Down. The route between PE_1 and PE_2 is reachable through P, so the remote LDP adjacency is still Up. The session type changes to a remote session. Since the session is still Up, L2VPN is uninformed of the session type change and does not delete the session. This avoids the neighbor disconnection and recovery process and therefore reduces the service interruption time.
3. When the physical link between PE_1 and PE_2 recovers, the local LDP adjacency goes Up. The session is restored to a local-and-remote and remains Up. Again L2VPN is not informed of the session type change and does not delete the session. This reduces the service interruption time.

2.2.5 LDP Session Protection

LDP session protection is an enhancement to the basic peer discovery mechanism. If the basic peer discovery mechanism fails, LDP session protection uses an extended peer discovery mechanism to maintain a session between LDP peers. After the basic peer discovery mechanism recovers, LDP can use it to rapidly converge routes and reestablish an LSP.

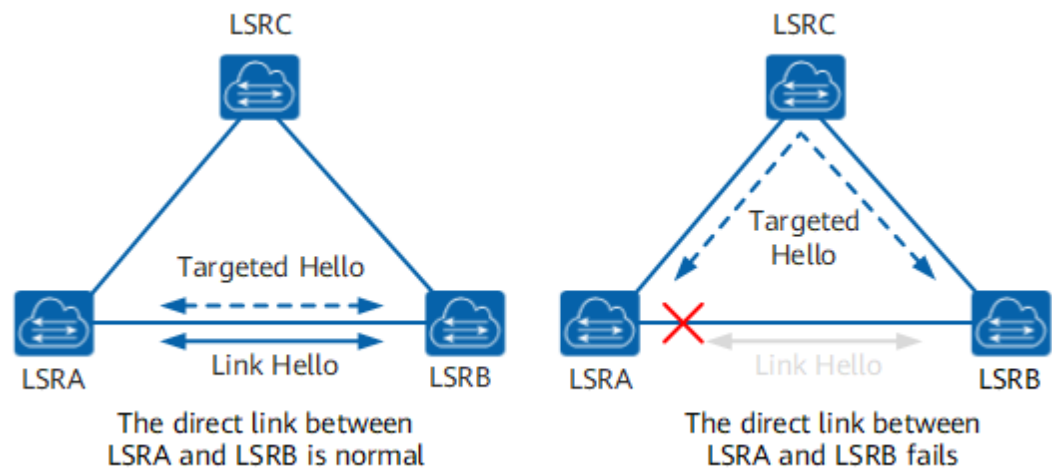
Background

If a direct link for a local LDP session fails, the LDP adjacency is torn down, and the session and labels are deleted. After the direct link recovers, the local LDP session is reestablished and labels are re-distributed so that an LSP can be reestablished over the session. Before the LSP is reestablished, however, LDP LSP traffic is dropped.

To speed up LDP LSP convergence and minimize packet loss, the device implements LDP session protection. LDP session protection helps maintain an LDP session, eliminating the need to reestablish an LDP session or re-distribute labels.

Implementation

As shown in [Figure 2-14](#), LDP session protection is configured on the nodes at both ends of a local LDP session. The remote LDP session established using Targeted Hello messages backs up the local LDP session established using Link Hello messages, implementing LDP session protection.

Figure 2-14 LDP session protection

As shown in [Figure 2-14](#), if the direct link between LSRA and LSRB fails, the adjacency established using Link Hello messages is torn down. Because the indirectly connected link is working properly, the remote adjacency established using Targeted Hello messages remains. Therefore, the LDP session is maintained by the remote adjacency, and the mapping between FECs and labels for the session also remains. After the direct link recovers, the local LDP session can rapidly restore LSP information. There is no need to reestablish the LDP session or re-distribute labels, which minimizes the time required for LDP session convergence.

Session Holdtime

In addition to LDP session protection, a session holdtime can be configured. After a local adjacency established using Link Hello messages is torn down, a remote adjacency established using Targeted Hello messages continues to maintain an LDP session within the configured session holdtime. If the local adjacency does not recover after the session holdtime elapses, the remote adjacency is torn down, and the LDP session maintained using the remote adjacency is also torn down. If the session holdtime is not specified, the remote adjacency permanently maintains the LDP session.

2.3 Summary of MPLS LDP Configuration Tasks

After basic functions of MPLS LDP are configured, you can build an MPLS network using LDP. To ensure network reliability and security, you need to perform other configuration in addition to MPLS LDP.

[Table 2-5](#) describes the MPLS LDP configuration tasks.

Table 2-5 MPLS LDP configuration tasks

Scenario	Description	Task
Configure basic functions of MPLS LDP	You can build an MPLS network and establish LDP LSPs only after basic functions of MPLS LDP are configured.	2.6 Configuring Basic Functions of MPLS LDP
Configure LDP reliability	The following reliability technologies can be used to improve MPLS network reliability: • BFD for LDP LSPs: quickly detects faults on an LDP LSP and triggers a traffic switchover upon an LDP LSP failure, minimizing packet loss. • Synchronization between LDP and IGP: solves the traffic loss problem when the primary LSP is faulty in networking where there are the primary and backup LSPs. • Auto LDP fast reroute (LDP FRR): provides link backup on an MPLS network. When the primary LSP fails, traffic is quickly switched to the backup LSP, minimizing traffic loss. • LDP GR: ensures uninterrupted traffic transmission when an active/standby switchover or a protocol restart occurs on the neighboring device (GR Restarter). LDP GR helps GR Restarter to restart.	2.7 Configuring Static BFD to Detect an LDP LSP 2.8 Configuring Synchronization Between LDP and IGP 2.9 Configuring Auto LDP FRR 2.10 Configuring LDP GR Helper
Configure LDP security mechanisms	LDP security mechanisms ensure security of LDP messages.	2.11 Configuring LDP Security Mechanisms

Scenario	Description	Task
Configure non-labeled public network routes to be iterated to LSPs	After this configuration is performed on access devices, service data is forwarded to the Internet through tunnels. By doing this, core devices of the carrier do not need to learn many Internet routes, saving the routing table storage space and CPU resources.	2.12 Configuring Non-labeled Public Network Routes to Be Iterated to LSPs

2.4 Licensing Requirements and Limitations for MPLS LDP

Involved Network Elements

Other network elements are not required.

License Requirements

MPLS is a basic function of the switch, and as such is controlled by the license for basic software functions. The license for basic software functions has been loaded and activated before delivery. You do not need to manually activate it.

Version Requirements

Table 2-6 Products and minimum version supporting this feature

Product Model	Minimum Version Required
CE6856HI	V200R002C50
CE6857EI	V200R005C10
CE6857E	V200R020C10
CE6857F	V200R021C00
CE6865EI	V200R005C00
CE6865E	V200R021C00
CE8850E-32CQ-EI	V200R021C10
CE6870-48S6CQ-EI	V200R001C00
CE6870-48S6CQ-EI-A	V200R022C00

Product Model	Minimum Version Required
CE6870-48T6CQ-EI	V200R002C50
CE8850-32CQ-EI	V200R002C50
CE8850-64CQ-EI	V200R005C00
CE8861EI	V200R005C10
CE8868EI	V200R005C10

Feature Limitations

When configuring MPLS LDP on the switch, pay attention to the following points:

Limitations on MPLS LDP:

- CE switches cannot function as PE devices on an MPLS network.
- RIP-1 cannot run on the MPLS network.
- LDP cannot be restarted during LDP GR.
- For the CE6870EI functioning as the penultimate hop device:
 - If the next hop is an LSP with implicit null label 3 and an LSP with a non-3 label and the LSPs form an ECMP, only one LSP is selected to forward traffic, the ECMP load balancing function becomes invalid, and packets are lost during a link switchover.
 - If the next hop is an LSP with implicit null label 3 and an LSP with a non-3 label and the LSPs form an FRR, only the primary LSP is selected to forward traffic and the LDP FRR function becomes invalid in versions earlier than V200R019C00. In V200R019C00 and later versions, packets are discarded during an LDP FRR switchover and the switchover has the following impacts:
 - When the primary LSP uses label 3 and traffic is forwarded through the backup LSP, the traffic is forwarded using the inner label after the outer label is popped out. In this case, the TTL of the inner label in the packet sent by the device is decreased by 1. As a result, the tracer path is incomplete.
 - When the primary LSP uses a non-3 label and traffic is forwarded through the backup LSP, label 0 is added to the packets before they are forwarded.
- If received MPLS packets contain four or more layers of labels, the CE6870EI does not load balance the packets based on the IP header using Eth-Trunk or ECMP after popping out the outermost label.
- For CE6870EI series switches, if MPLS packets with four or more labels are received, Eth-Trunk or ECMP load balancing cannot be implemented based on the inner IP header.
- After removing MPLS labels, switches except CE6870EI do not load balance MPLS packets through an Eth-Trunk based on the IP header by default. In the PHP scenario, you can configure the **mpls top-label src-ip dst-ip 2nd-label**

command on the penultimate hop to load balance traffic to an Eth-Trunk. You can modify the load balancing profile based on the traffic characteristics to optimize the load balancing effect.

- Layer 3 sub-interfaces on the CE6865EI, CE6865E, CE8850E-32CQ-EI, CE6857EI, CE6857E, CE6857F, CE8861EI, and CE8868EI do not support MPLS.
- The following operations lead to deletion or re-creation of LDP sessions, causing MPLS service interruption. Exercise caution when performing these operations:
 - Change or delete the remote peer IP address or use the IP address as the IP address of a local interface.
 - Change the LDP transport address.
 - Modify the timeout interval of the Keepalive hold timer.
 - Configure LDP MD5 or LDP Keychain authentication.
 - Enable or disable LDP GR or modify the timeout interval of LDP GR-related timers.
- When the CE6870EI is not configured with the MPLS PHP function and uses the inner label to forward packets after the outer label pops up (for example, the switch serves as an ASBR in the inter-AS VPN Option B or inter-AS VPN Option C scenario), the TTL value in the inner label minus 1 is set as the TTL value in the label of packets forwarded by the switch. As a result, an incomplete tracer path is displayed.
- On an MPLS network where packets contain multiple labels, the tracer function may fail to work properly.

Limitations on coexistence of MPLS LDP and other features:

- For CE6870EI series switches, if MPLS functions have been configured on the device, interfaces on the MPLS forwarding path cannot have TRILL functions enabled. Otherwise, MPLS functions do not take effect.
- When the system resource mode is set to **large-arp** or UFT hybrid mode where ARP entries are specified, the device does not support the MPLS function. After the MPLS function is configured, you cannot change the system resource mode to **large-arp** or specify ARP entries to UFT hybrid mode.
- MPLS cannot be used together with MUX VLAN.
- When a CE6865EI, CE6865E, CE8850E-32CQ-EI, CE6870EI functions as a PE node on a VPLS network and the MPLS TTL processing mode is set to uniform, the TTL field of the MPLS label has a fixed value of 255.
- In V200R005C10 and later versions, stacked switches do not support MPLS.
- In the PHP scenarios where non-CE6870EI switches are deployed, you are not advised to configure BFD on the penultimate hop device. Otherwise, extra local BFD session resources are occupied. In addition, when the discriminator of the local BFD session is the same as that of the BFD session established between PEs, the local BFD session cannot go Up.
- When the CE6856HI, CE6857EI, CE6857E, CE6857F, CE6865EI, CE6865E, CE8850E-32CQ-EI, CE8861EI, or CE8868EI switch functions as a penultimate hop device, the implicit null label is enabled, and LSPs whose next hop is label 3 or non-label 3 form multiple paths:

- The switch prefers the LSP with the smallest next-hop IP address among all outbound interfaces as the forwarding path and does not support ECMP fast update.
- In the FRR scenario, only the primary path is selected, and FRR is not supported.
- When LDP load balancing and FRR are used together, OSPF is not supported.

2.5 Default Settings for MPLS LDP

Table 2-7 Default settings for MPLS LDP

Parameter	Default Setting
Global MPLS capability	Disabled
Global MPLS LDP capability	Disabled
Link-Hello send timer	5 seconds
Link-Hello hold timer	15 seconds
Target-Hello send timer	15 seconds
Target-Hello hold timer	45 seconds
Keepalive send timer	15 seconds
Keepalive hold timer	45 seconds
Exponential backoff timer	Initial value: 15 seconds; maximum value: 120 seconds
LDP for Open Shortest Path First (OSPF)	Disabled
LDP for Intermediate System to Intermediate System (IS-IS)	Disabled
LDP GR	Disabled

2.6 Configuring Basic Functions of MPLS LDP

You can build an MPLS network only after basic functions of MPLS LDP are configured.

Pre-configuration Tasks

Before configuring basic functions of MPLS LDP, configure static routes or an IGP to ensure that IP routes between LSRs are reachable.

NOTE

Routing Information Protocol version 1 (RIP-1) cannot be used.

Configuration Procedure

Configure basic functions of MPLS LDP according to the following sequence.

2.6.1 Configuring the LSR ID

Context

An LSR ID identifies an LSR on a network. An LSR does not have the default LSR ID, and you must configure an LSR ID for it. To enhance network reliability, you are advised to use the IP address of a loopback interface on the LSR as the LSR ID.

Perform the following steps on each node in an MPLS domain.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Run **mpls lsr-id** *lsr-id*

The LSR ID of the local node is configured.

By default, no LSR ID is set.

Step 3 Run **commit**

The configuration is committed.

----End

2.6.2 Enabling Global MPLS

Context

You can perform other MPLS configurations only after enabling global MPLS.

Perform the following steps on each node in an MPLS domain.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Run **mpls**

MPLS is enabled globally and the MPLS view is displayed.

By default, no node is enabled with MPLS.

Step 3 Run **commit**

The configuration is committed.

----End

2.6.3 Enabling Global MPLS LDP

Context

You can perform other MPLS LDP configurations only after enabling global MPLS LDP.

Perform the following steps on each node in an MPLS domain.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Run **mpls ldp**

MPLS LDP is enabled globally and the MPLS LDP view is displayed.

By default, LDP is not enabled globally.

Step 3 (Optional) Run **lsr-id lsr-id**

The LSR ID is set for an LDP instance.

By default, the LSR ID of the LDP instance is the LSR ID of the local node. It is recommended that the default value be used.

In certain networking where VPN instances are used, such as BGP/MPLS IP VPN networking, if the VPN address and the LSR ID overlap, you need to configure LSR IDs for LDP instances to ensure that TCP connections can be correctly set up.

Step 4 Run **commit**

The configuration is committed.

----End

2.6.4 Configuring LDP Sessions

Context

The MPLS LDP session is classified into local LDP sessions and remote LDP sessions. You can choose one of the following configurations according to your requirements:

- **Configuring a local LDP session (based on interfaces)**

In most cases, you need to configure a local LDP session when deploying MPLS LDP services.

- **Configuring local LDP sessions (based on routes)**

When the MPLS LDP function is configured, the MPLS LDP function must be enabled globally first and then on all the required interfaces. If too many interfaces need to be enabled with the function, the configuration method wastes too much time and energy, and cannot ensure the configuration accuracy.

After the MPLS LDP function is enabled globally, the function is automatically enabled on the IGP-enabled interfaces and the IGP-based LDP sessions are configured. This simplifies configuration and guarantees the configuration accuracy.

- **Configuring a remote LDP session**

In most cases, remote LDP sessions are not established between adjacent LSRs. A remote LDP session is used for configuring a VPLS in Martini mode.

A local LDP session and a remote LDP session can coexist. That is, two LSRs can establish a local LDP session and a remote LDP session simultaneously. In this case, configurations of the local and remote LDP sessions at both ends must be the same.

Procedure

- Configuring a local LDP session (based on interfaces)

Perform the following steps on two directly connected LSRs.

- a. Run **system-view**

The system view is displayed.

- b. Run **interface** *interface-type interface-number*

The view of the interface on which the LDP session is to be set up is displayed.

- c. On an Ethernet interface, run **undo portswitch**

The interface is switched to Layer 3 mode.

By default, an Ethernet interface works in Layer 2 mode.

The mode switching function takes effect when the interface only has attribute configurations (for example, **shutdown** and **description** configurations). Alternatively, if configuration information supported by both Layer 2 and Layer 3 interfaces exists (for example, **mode lacp** and **lacp system-id** configurations), no configuration that is not supported after the working mode of the interface is switched can exist. If unsupported configurations exist on the interface, delete the configurations first and then run the **undo portswitch** command.

NOTE

If many Ethernet interfaces need to be switched to Layer 3 mode, run the **undo portswitch batch** *interface-type { interface-number1 [to interface-number2] }* &<1-10> command in the system view to switch these interfaces to Layer 3 mode in batches.

- d. Run **mpls**

MPLS is enabled on the interface.

By default, no interface is enabled with MPLS.

- e. Run **mpls ldp**

MPLS LDP is enabled on the interface.

By default, no interface is enabled with LDP.

f. Run **commit**

The configuration is committed.

- Configuring local LDP sessions (based on routes)

Configure automatic LDP deployment based on Intermediate System to Intermediate System (IS-IS).

a. Run **system-view**

The system view is displayed.

b. Run **isis** [*process-id*]

The IS-IS view is displayed.

c. Run **mpls ldp auto-config**

The MPLS LDP function is automatically enabled on the IS-IS interfaces.

After this command is run, the MPLS LDP function is automatically enabled on all the interfaces that can establish IS-IS neighboring relationships in the IS-IS process. To disable the MPLS LDP function on some interfaces, you can run the **isis mpls ldp auto-config disable** command.

d. Run **commit**

The configuration is committed.

Configure automatic LDP deployment based on Open Shortest Path First (OSPF).

a. Run **system-view**

The system view is displayed.

b. Run **ospf** [*process-id*]

The OSPF view is displayed.

c. Run **mpls ldp auto-config**

The MPLS LDP function is automatically enabled on the OSPF interfaces.

After this command is run, the MPLS LDP function is automatically enabled on all the interfaces that can establish OSPF neighboring relationships in the OSPF process. To disable the MPLS LDP function on some interfaces, you can run the **ospf mpls ldp auto-config disable** command.

d. Run **commit**

The configuration is committed.

- Configuring a remote MPLS LDP session

Perform the following steps on the LSRs on both ends of a remote LDP session.

a. Run **system-view**

The system view is displayed.

b. Run **mpls ldp remote-peer** *remote-peer-name*

The remote peer is created and the remote peer view is displayed.

- c. (Optional) Run **description** *description-value*

The description of a remote peer is configured.

- d. Run **remote-ip** *ip-address*

The IP address of the remote MPLS LDP peer is configured.

By default, the IP address of the remote LDP peer is not configured.

This IP address must be the LSR ID of the remote MPLS LDP peer. If the LSR IDs of the LDP instance and the local node are different, use the LSR ID of the LDP instance.

NOTICE

- Modifying or deleting the IP address of a remote peer leads to deletion of the remote LDP session and MPLS service interruption.
- After the IP address of the remote peer is configured using the **remote-ip** *ip-address* command, the value of *ip-address* cannot be used as the IP address of the local interface. Otherwise, the remote session will be interrupted, causing MPLS service interruption.

-
- e. Run **commit**

The configuration is committed.

----End

2.6.5 (Optional) Configuring LDP Session Protection

Context

If the direct link of a local LDP session between two devices fails, an LDP adjacency for the LDP session is torn down. The LDP session and related labels are also deleted. After the direct link recovers, the LDP session is reestablished and labels are re-distributed so that an LDP LSP over the session can converge. During this process, LDP LSP traffic is dropped.

To speed up LDP LSP convergence and minimize traffic loss, LDP session protection can be used. With LDP session protection, the LDP session does not need to be reestablished, and labels do not need to be re-distributed during the preceding process.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Run **mpls ldp**

The MPLS-LDP view is displayed.

Step 3 Run **session protection** [**peer-group** *peer-group-name*] [**duration** { **infinite** | *time-value* }]

LDP session protection is enabled.

By default, LDP session protection is disabled.

Step 4 Run **commit**

The configuration is committed.

----End

2.6.6 (Optional) Configuring an LDP Transport Address

Context

LDP sessions are established based on TCP connections. Before two LSRs establish an LDP session, they need to check the LDP transport address of each other, and then establish a TCP connection.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Run **interface** *interface-type interface-number*

The view of the interface on which the LDP session is to be set up is displayed.

Step 3 On an Ethernet interface, run **undo portswitch**

The interface is switched to Layer 3 mode.

By default, an Ethernet interface works in Layer 2 mode.

The mode switching function takes effect when the interface only has attribute configurations (for example, **shutdown** and **description** configurations). Alternatively, if configuration information supported by both Layer 2 and Layer 3 interfaces exists (for example, **mode lacp** and **lacp system-id** configurations), no configuration that is not supported after the working mode of the interface is switched can exist. If unsupported configurations exist on the interface, delete the configurations first and then run the **undo portswitch** command.

NOTE

If many Ethernet interfaces need to be switched to Layer 3 mode, run the **undo portswitch batch interface-type { interface-number1 [to interface-number2] } <1-10>** command in the system view to switch these interfaces to Layer 3 mode in batches.

Step 4 Run **mpls ldp transport-address** { *interface-type interface-number* | **interface** }

An LDP transport address is specified.

The default transport address for a node on a public network is the LSR ID of the node, and the default transport address for a node on a private network is the primary IP address of an interface on the node.

If LDP sessions are to be established over multiple links connecting two LSRs, LDP-enabled interfaces on either LSR must use the default transport address or the same transport address. If multiple transport addresses are configured on an LSR, only one transport address can be used to establish only one LDP session.

NOTICE

Changing an LDP transport address interrupts an LDP session. Exercise caution when running this command.

Step 5 Run commit

The configuration is committed.

----End

2.6.7 (Optional) Configuring Timers for LDP Session

Context

[Table 2-8](#) describes the timers for an LDP session.

Table 2-8 Timers for an LDP session

LDP Timer	Description	Suggestion
Hello send timer: - Link-Hello send timer (for only local LDP sessions) - Target-Hello send timer (for only remote LDP sessions)	Used to send Hello messages periodically to notify a peer LSR of the local LSR's presence and establish a Hello adjacency.	On an unstable network, decrease the value of a Hello send timer, speeding up network fault detection.

LDP Timer	Description	Suggestion
Hello hold timer: • Link-Hello hold timer (for only local LDP sessions) • Target-Hello hold timer (for only remote LDP sessions)	Used to exchange Hello messages periodically between two LDP peers to maintain the Hello adjacency. If no Hello message is received after the Hello hold timer expires, the Hello adjacency is torn down.	On a network with unstable links or a large number of packets, increase the value of the Hello hold timer, preventing the LDP session from being torn down and set up frequently.
Keepalive send timer	Used to send Keepalive messages periodically, maintaining the LDP sessions.	On an unstable network, set a smaller value for a Keepalive send timer, speeding up network fault detection.
Keepalive hold timer	Used to send LDP PDUs over an LDP session, maintaining the LDP session. If no LDP PDU is received after the Keepalive hold timer expires, the TCP connection is closed and the LDP session is terminated.	On a network with unstable links, increase the value of the Keepalive hold timer, preventing the LDP session from flapping.
Exponential backoff timer	Started by an LSR that plays an active role after an LDP Initialization message sent by the LSR to another LSR that plays a passive role fails to be processed or parameters carried in the message are rejected. The LSP that plays the active role periodically resends an LDP Initialization message to initiate an LDP session before the Exponential backoff timer expires.	• When a device is upgraded, prolong the period for the active role to retry setting up a session. In this case, you can set larger initial and maximum values for the Exponential backoff timer. • When a device that bears services tends to alternate between Up and Down, shorten the period for the active role to retry setting up a session. In this case, you can set smaller initial and maximum values for the Exponential backoff timer.

 NOTE

When local and remote LDP sessions coexist, the timeout interval of the Keepalive hold timer of the local and remote LDP sessions must be the same.

Procedure

- Configuring timers for a local LDP session

- a. Run **system-view**

The system view is displayed.

- b. Run **interface** *interface-type interface-number*

The view of an interface on which an LDP session is to be established is displayed.

- c. On an Ethernet interface, run **undo portswitch**

The interface is switched to Layer 3 mode.

By default, an Ethernet interface works in Layer 2 mode.

The mode switching function takes effect when the interface only has attribute configurations (for example, **shutdown** and **description** configurations). Alternatively, if configuration information supported by both Layer 2 and Layer 3 interfaces exists (for example, **mode lacp** and **lacp system-id** configurations), no configuration that is not supported after the working mode of the interface is switched can exist. If unsupported configurations exist on the interface, delete the configurations first and then run the **undo portswitch** command.

 NOTE

If many Ethernet interfaces need to be switched to Layer 3 mode, run the **undo portswitch batch** *interface-type { interface-number1 [to interface-number2] }* &<1-10> command in the system view to switch these interfaces to Layer 3 mode in batches.

- d. Run **mpls ldp timer hello-send** *interval*

A link Hello send timer is configured.

The default value of a link Hello send timer is one third of the value of a link Hello hold timer.

Effective value of a link Hello send timer = Min {Configured value of the link Hello send timer, one third of the value of the link Hello hold timer}

- e. Run **mpls ldp timer hello-hold** *interval*

A link Hello hold timer is configured.

The default value of a link Hello hold timer is 15, in seconds.

The smaller value between two configured link Hello hold timers on both ends of the LDP session takes effect.

The timer must be longer than the time a device takes to perform a master/slave main control board switchover. If the timer is set to less than the switchover time, a protocol intermittent interruption occurs during a switchover. The default timer value is recommended.

f. Run **mpls ldp timer keepalive-send** *interval*

A Keepalive send timer is configured.

The default value of a Keepalive send timer is one third of the value of the Keepalive hold timer.

Effective value of a Keepalive send timer = Min { Configured value of the Keepalive send timer, one third of the value of the Keepalive hold timer }

If more than one LDP-enabled links connect two LSRs, the values of Keepalive send timers for all links must be the same. Otherwise, LDP sessions become unstable.

g. Run **mpls ldp timer keepalive-hold** *interval*

A Keepalive hold timer is configured.

The default value of a Keepalive hold timer is 45, in seconds.

The smaller value between two configured Keepalive hold timers on both ends of the LDP session takes effect.

If more than one LDP-enabled links connect two LSRs, the values of Keepalive hold timers for all links must be the same. Otherwise, LDP sessions may fail to be set up.

The timer must be longer than the time a device takes to perform a master/slave main control board switchover. If the timer is set to less than the switchover time, a protocol intermittent interruption occurs during a switchover. The default timer value is recommended.

NOTICE

Changing the Keepalive hold timer value in an instance will interrupt the MPLS service in the instance because the LDP session must be reestablished.

h. Configure an Exponential backoff timer.

i. Run **quit**

The system view is displayed.

ii. Run **mpls ldp**

The MPLS LDP view is displayed.

iii. Run **backoff timer** *init max*

An Exponential backoff timer is configured.

By default, the initial value is 15 and the maximum value is 120, in seconds. Setting the initial value of the Exponential backoff timer to be greater than or equal to 15s and the maximum value to be greater than or equal to 120s is recommended.

i. Run **commit**

The configuration is committed.

- Configuring timers for a remote LDP session

a. Run **system-view**

The system view is displayed.

b. Run **mpls ldp remote-peer** *remote-peer-name*

The remote MPLS LDP peer view is displayed.

c. Run **mpls ldp timer hello-send** *interval*

The target Hello send timer is configured.

The default value of the target Hello send timer is one third of the value of a target Hello hold timer that takes effect.

Effective value of a target Hello send timer = Min { Configured value of the target Hello send timer, One third of the value of the target Hello hold timer }

d. Run **mpls ldp timer hello-hold** *interval*

The target Hello hold timer is configured.

The default value of the target Hello hold timer is 45, in seconds.

The smaller value between two configured target Hello hold timers on both ends of the LDP session takes effect.

The timer must be longer than the time a device takes to perform a master/slave main control board switchover. If the timer is set to less than the switchover time, a protocol intermittent interruption occurs during a switchover. The default timer value is recommended.

e. Run **mpls ldp timer keepalive-send** *interval*

A Keepalive send timer is configured.

The default value of a Keepalive send timer is one third of the value of the Keepalive hold timer.

Effective value of a Keepalive send timer = Min { Configured value of the Keepalive send timer, one third of the value of the Keepalive hold timer }

If more than one LDP-enabled links connect two LSRs, the values of Keepalive send timers for all links must be the same. Otherwise, LDP sessions become unstable.

f. Run **mpls ldp timer keepalive-hold** *interval*

A Keepalive hold timer is configured.

The default value of a Keepalive hold timer is 45, in seconds.

The smaller value between two configured Keepalive hold timers on both ends of the LDP session takes effect.

If more than one LDP-enabled links connect two LSRs, the values of Keepalive hold timers for all links must be the same. Otherwise, LDP sessions may fail to be set up.

The timer must be longer than the time a device takes to perform a master/slave main control board switchover. If the timer is set to less than the switchover time, a protocol intermittent interruption occurs during a switchover. The default timer value is recommended.

NOTICE

Changing the Keepalive hold timer value in an instance may interrupt the MPLS service in the instance because the LDP session must be reestablished.

- g. Configure an Exponential backoff timer.
 - i. Run **quit**
The system view is displayed.
 - ii. Run **mpls ldp**
The MPLS LDP view is displayed.
 - iii. Run **backoff timer *init max***
An Exponential backoff timer is configured.
By default, the initial value is 15 and the maximum value is 120, in seconds. Setting the initial value of the Exponential backoff timer to be greater than or equal to 15s and the maximum value to be greater than or equal to 120s is recommended.
- h. Run **commit**
The configuration is committed.

----End

2.6.8 (Optional) Configuring LDP Loop Detection

Context

The device does not support LDP loop detection. If the neighbor of a node supports loop detection and requires the same loop detection function on both ends of an LDP session, configure LDP loop detection on the local node to ensure the establishment of an LDP session.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Run **mpls ldp**

The MPLS LDP view is displayed.

Step 3 Run **loop-detect**

The device is enabled to advertise the loop detection capability during initialization of LDP sessions.

By default, a device does not advertise loop detection capability during initialization of LDP sessions.

Step 4 Run **commit**

The configuration is committed.

----End

2.6.9 (Optional) Configuring the LDP Label Policies

Context

The LSR distributes labels to both upstream and downstream LDP peers, which increases the LDP LSP convergence speed. However, receiving and sending Label Mapping messages result in the establishments of a large number of LSPs, which wastes resources. To reduce the number of LSPs and save memory, use the following policies:

- Configure the LDP outbound policy.
Configure LDP outbound policy to restrict the sending of Label Mapping messages.
- Configure the LDP split horizon policy.
Access devices on the MPLS network have low performance. If LDP distributes labels to all peers, a large number of LSPs will be established, which cannot be processed by the LSR. The split horizon policy is recommended.

Procedure

- Configure an outbound LDP policy.
 - a. Run **system-view**
The system view is displayed.
 - b. Run **mpls ldp**
The MPLS LDP view is displayed.
 - c. Run **ipv4-family**
The MPLS-LDP-IPv4 view is displayed.
 - d. Perform either of the following steps to apply the outbound policy that allows Label Mapping messages for specified routes to be sent to a specified LDP peer:
 - To configure an outbound policy that allows Label Mapping messages for specified IGP routes to be sent to a specified LDP peer, run **outbound peer { peer-id | peer-group peer-group-name | all } fec { none | host | ip-prefix prefix-name }**
 - To configure an outbound policy that allows Label Mapping messages for specified labeled Border Gateway Protocol (BGP) routes to be sent to a specified LDP peer, run **outbound peer { peer-id | peer-group peer-group-name | all } bgp-label-route { none | ip-prefix prefix-name }**

To apply a policy associated with a single Forwarding Equivalence Class (FEC) range to an LDP peer group or all LDP peers to which the local LSR sends Label Mapping messages, configure either the **peer-group peer-group-name** or **all** parameter in the command.

NOTE

If multiple outbound policies are configured for a specified LDP peer, the first configured one takes effect. For example, the following two outbound policies are configured:

```
outbound peer 2.2.2.2 fec host
outbound peer peer-group group1 fec none
```

As group1 also contains an LDP peer with *peer-id* of 2.2.2.2, the following outbound policy takes effect:

```
outbound peer 2.2.2.2 fec host
```

If two outbound policies are configured in sequence and the **peer** parameters in the two commands are the same, the second command overwrites the first one. For example, the following two outbound policies are configured:

```
outbound peer 2.2.2.2 fec host
outbound peer 2.2.2.2 fec none
```

The second configuration overwrites the first one. This means that the following outbound policy takes effect on the LDP peer with *peer-id* of 2.2.2.2:

```
outbound peer 2.2.2.2 fec none
```

e. Run **commit**

The configuration is committed.

- Configure an LDP split horizon policy.

a. Run **system-view**

The system view is displayed.

b. Run **mpls ldp**

The MPLS LDP view is displayed.

c. Run **undo outbound peer { *peer-id* | all } split-horizon disable**

A split horizon policy is configured to distribute labels to only upstream LDP peers.

By default, LDP split horizon is enabled for all LDP peers so that LSRs allocate labels only to upstream LDP peers.

d. Run **commit**

The configuration is committed.

----End

Follow-up Procedure

To delete all outbound policies, run the **undo** command multiple times to delete them one by one, or run the **undo outbound peer all** command to delete them simultaneously. The first method takes a long time.

2.6.10 (Optional) Disabling a Device from Distributing Labels to Remote Peers

Context

In MPLS L2VPN scenarios using LDP (including Martini VPLS), PEs at both ends need to establish a remote LDP session. The remote LDP session is only used to

transmit Label Mapping messages, so LDP is not required. By default, LDP allocates common LDP labels to remote peers. Many useless idle labels are generated, wasting LDP labels.

To solve the preceding problem, disable a device from distributing labels to remote peers to save system resources. You can use either of the following modes:

- In the LDP view, disable the PE from distributing labels to all remote peers.
- In the view of a specified remote peer, disable the PE from distributing labels to the specified remote peer.

NOTE

When TE services are transmitted over a backbone network in the LDP over TE scenario, do not perform this configuration.

Procedure

- Disable a device from distributing labels to a specified remote peer.
 - a. Run **system-view**
The system view is displayed.
 - b. Run **mpls ldp remote-peer** *remote-peer-name*
The remote MPLS LDP peer view is displayed.
 - c. Run **remote-ip** *ip-address* **pwe3**
LDP is prevented from allocating public network labels to a specified remote peer device.
By default, the IP address of the remote LDP peer is not configured.
 - d. Run **commit**
The configuration is committed.
- Disable a device from distributing LDP labels to all remote peers.
If an explicit configuration has been performed to allow a device to distribute labels to a specified remote LDP peer, run **clear remote-ip pwe3** to delete the configuration, you can perform this step.
 - a. Run **system-view**
The system view is displayed.
 - b. Run **mpls ldp**
The MPLS LDP view is displayed.
 - c. Run **remote-peer pwe3**
LDP is prevented from allocating public network labels to all remote peer devices.
By default, an LSR is permitted to distribute public network labels to all remote peers.
 - d. Run **commit**
The configuration is committed.

----End

2.6.11 (Optional) Configuring a Policy for Triggering LDP LSP Establishment

Context

After MPLS LDP is enabled, LSPs are automatically established. If no policy is configured, an increasing number of LSPs are established, wasting resources.

- Configure the **lsp-trigger** command on the ingress and egress nodes to trigger LSP setup based on routes. This setting controls the number of LSPs and saves network resources.

The **lsp-trigger** command can be configured in the MPLS view and MPLS-LDP-IPv4 view. If the **lsp-trigger** command is configured in both the MPLS view and MPLS-LDP-IPv4 view, the **lsp-trigger** command in the MPLS-LDP-IPv4 view takes effect.

- Configure the **propagate mapping** command on the transit node to allow LDP to use routes matching specified conditions to establish transit LSPs. For the routes that do not match specified conditions, the local device does not send Label Mapping messages to the upstream device, which reduces the number of LSPs and saves network resources.

The **propagate mapping** command can be configured in the MPLS-LDP view and MPLS-LDP-IPv4 view. If the **propagate mapping** command is configured in both the MPLS-LDP view and MPLS-LDP-IPv4 view, the latest configuration takes effect.

By default, the **lsp-trigger** command is used. If policies cannot be configured on the ingress and egress nodes, configure the **propagate mapping** command.

Procedure

- Perform the following steps on the ingress and egress nodes:

Configuration in the MPLS view:

- a. Run **system-view**

The system view is displayed.

- b. Run **mpls**

The MPLS view is displayed.

- c. Run the following commands as required.

- Run **lsp-trigger { all | host | ip-prefix *ip-prefix-name* | none }**

A policy for triggering LSP establishment based on static and IGP routes is configured.

By default, the policy is **host**. This policy allows LDP to use 32-bit host-address routes (except 32-bit host-address of interfaces) to establish LSPs.

NOTE

LSPs can be established using exactly matching routes on LSRs. For example, an exactly matching host route to an IP address with a 32-bit mask of a loopback interface can be used to trigger LSP establishment.

- Run **lsp-trigger bgp-label-route** [**ip-prefix** *ip-prefix-name*] [**not-only-host**]

A policy for triggering LSP establishment based on labeled public network BGP routes is configured.

By default, LDP does not allocate labels to labeled BGP routes on the public network.

- d. Run **proxy-egress disable**

The device is disabled from establishing proxy egress LSPs.

By default, a device is enabled to establish proxy egress LSPs.

If a policy allows a device to use all static and IGP routes to establish LSPs or use an IP address prefix list to establish LSPs, the policy also triggers proxy egress LSP establishment. However, the proxy egress LSPs may be unavailable, which wastes system resources. To prevent this problem, run the **proxy-egress disable** command to disable a device from establishing such proxy egress LSPs.

- e. Run **commit**

The configuration is committed.

Configuration in the MPLS-LDP-IPv4 view:

- a. Run **system-view**

The system view is displayed.

- b. Run **mpls**

The MPLS view is displayed.

- c. Run **proxy-egress disable**

The device is disabled from establishing proxy egress LSPs.

By default, a device is enabled to establish proxy egress LSPs.

If a policy allows a device to use all static and IGP routes to establish LSPs or use an IP address prefix list to establish LSPs, the policy also triggers proxy egress LSP establishment. However, the proxy egress LSPs may be unavailable, which wastes system resources. To prevent this problem, run the **proxy-egress disable** command to disable a device from establishing such proxy egress LSPs.

- d. Run **quit**

The system view is displayed.

- e. Run **mpls ldp**

The MPLS LDP view is displayed.

- f. Run **ipv4-family**

The MPLS-LDP-IPv4 view is displayed.

- g. Run **lsp-trigger** { **all** | **host** | **ip-prefix** *ip-prefix-name* | **none** }

A policy for triggering LSP establishment based on static and IGP routes is configured.

By default, the policy is **host**. This policy allows LDP to use 32-bit host-address routes (except 32-bit host-address of interfaces) to establish LSPs.

 NOTE

LSPs can be established using exactly matching routes on LSRs. For example, an exactly matching host route to an IP address with a 32-bit mask of a loopback interface can be used to trigger LSP establishment.

h. Run **commit**

The configuration is committed.

• Perform the following steps on the transit nodes:

a. Run **system-view**

The system view is displayed.

b. Run **mpls ldp**

The MPLS LDP view is displayed.

c. (Optional) Run **ipv4-family**

The MPLS-LDP-IPv4 view is displayed.

d. Run **propagate mapping for ip-prefix** *ip-prefix-name*

A policy for triggering transit LSP establishment is configured.

By default, LDP uses all routes without filtering them to establish transit LSPs.

e. Run **commit**

The configuration is committed.

----End

2.6.12 (Optional) Disabling LDP LSP Flapping Suppression

Context

After an LDP LSP goes Up, it goes Down due to a protocol or interface failure. A device attempts to reestablish the LDP LSP, which maximizes the LDP LSP protocol hard convergence. If LDP LSP alternates between Up and Down when a downstream node frequently sends a label to an upstream node or a label is withdrawn, CPU usage increases. To prevent label suppression, configure LDP LSP flapping suppression.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Run **mpls**

The MPLS view is displayed.

Step 3 Run **mpls ldp**

MPLS LDP is enabled globally.

Step 4 Run **protocol-packets suppress disable**

LDP LSP flapping suppression is disabled.

By default, LDP LSP flapping suppression is enabled.

Step 5 Run **commit**

The configuration is committed.

----End

2.6.13 (Optional) Disabling a Device from Forwarding Unknown TLVs

A local device can be disabled from forwarding unknown TLVs to peers.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Run **mpls ldp**

MPLS LDP is globally enabled.

Step 3 Run **propagate mapping unknown-tlv disable**

The device is disabled from forwarding unknown TLVs.

By default, the device is enabled to forward unknown TLVs.

If an upstream device cannot process unknown TLVs, network problems may occur. In this case, you can run this command to disable the local device from forwarding unknown TLVs.

Step 4 Run **commit**

The configuration is committed.

----End

2.6.14 Verifying the Configuration of Basic MPLS LDP Functions

Prerequisites

The configurations of the MPLS LDP function are complete.

Procedure

- Run the **display default-parameter mpls management** command to check default configurations of the MPLS management module.

- Run the **display default-parameter mpls ldp** command to check the default configurations of MPLS LDP.
- Run the **display mpls interface** [*interface-type interface-number*] [**verbose**] command to check information about MPLS-enabled interfaces.
- Run the **display mpls label all summary** command to check allocation information about all MPLS labels.
- Run the **display mpls ldp** [**all** | **all verbose**] command to check LDP information.
- Run the **display mpls ldp interface** [*interface-type interface-number* | **all** | **verbose**] command to check information about LDP-enabled interfaces.
- Run the **display mpls ldp adjacency** [**interface** *interface-type interface-number* | **remote**] [**peer** *peer-id*] [**verbose**] command to check information about LDP adjacencies.
- Run the **display mpls ldp adjacency statistics** command to check statistics about LDP adjacencies.
- Run the **display mpls ldp session** [*peer-id* | **all** | **verbose**] command to check the LDP session status.
- Run the **display mpls ldp session statistics** command to check statistics about sessions between LDP peers.
- Run the **display mpls ldp peer** [*peer-id* | **all** | **verbose**] command to check information about LDP peers.
- Run the **display mpls ldp peer statistics** command to check statistics about LDP peers.
- Run the **display mpls ldp remote-peer** [*remote-peer-name* | **peer-id** *lsp-id*] command to check information about the LDP remote peer.
- Run the **display mpls ldp lsp** [**all**] command to check LDP LSP information.
- Run the **display mpls ldp lsp statistics** command to check statistics about LDP LSPs.
- Run the **display mpls route-state** [{ **exclude** | **include** } { **idle** | **ready** | **settingup** } | *destination-address mask-length*] [**verbose**] command to check the dynamic LSP route.
- Run the **display mpls lsp** [**verbose**] command to check LSP information.
- Run the **display mpls lsp statistics** command to check statistics about the LSPs that are in the Up state and the number of the LSPs that are activated on the ingress, transit, and egress nodes.

----End

2.7 Configuring Static BFD to Detect an LDP LSP

Context

When static BFD monitors an LDP LSP, pay attention to the following points:

- BFD is bound to only the ingress node of an LDP LSP.
- One LSP is bound to only one BFD session.

- The detection only supports the LDP LSP that is triggered to establish by the host route.
- The forwarding modes on the forwarding path and reverse path can be different (for example, an IP packet is sent from the source to the destination through an LSP, and is sent from the destination to the source in IP forwarding mode), but the forwarding path and reverse path must be established over the same link. If they use different links, BFD cannot identify the faulty path when a fault is detected.

Static BFD for LDP LSPs fast detects faults on an LDP LSP. Static BFD for LDP LSPs can be flexibly deployed, but needs to be manually controlled.

Pre-configuration Tasks

Before configuring static BFD to detect an LDP LSP, complete the following task:

- Configure a local LDP session. For details, see [2.6 Configuring Basic Functions of MPLS LDP](#).

Configuration Procedure

Configure static BFD for LDP LSPs according to the following sequence.

2.7.1 Configuring BFD with Specific Parameters on the Ingress Node

Context

BFD parameters on the ingress node include the local and remote discriminators, intervals for sending and receiving BFD packets, and local BFD detection multiplier. The BFD parameters affect BFD session setup.

You can adjust the local detection time according to the network situation. On an unstable link, if a small detection time is used, a BFD session may flap. You can increase the detection time of the BFD session.

NOTE

Actual interval for the local device to send BFD packets = MAX { locally configured interval for sending BFD packets, remotely configured interval for receiving BFD packets }

Actual interval for the local device to receive BFD packets = MAX { remotely configured interval for sending BFD packets, locally configured interval for receiving BFD packets }

Local detection time = Actual interval for receiving BFD packets x Remotely configured BFD detection multiplier

Perform the following steps on the ingress node of an LSP:

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Run **bfd**

This node is enabled with the global BFD function. The global BFD view is displayed.

By default, global BFD is disabled.

Step 3 Run **quit**

Return to the system view.

Step 4 Run **bfd session-name bind ldp-lsp peer-ip ip-address nexthop ip-address**
[**interface interface-type interface-number**]

The BFD session is bound to a dynamic LSP.

When the IP address of the egress node on the LSP to be detected is borrowed or lent, an interface must be specified.

Step 5 Set local and remote discriminators of a BFD session.

- Run **discriminator local discr-value**
The local discriminator is configured.
- Run **discriminator remote discr-value**
The remote discriminator is configured.

 NOTE

The local and remote identifiers on both ends of a BFD session must be consistent with each other; otherwise, the session cannot be established correctly. In addition, the local and remote identifiers cannot be modified after configuration.

Step 6 (Optional) Run **min-tx-interval interval**

The interval for sending BFD packets is set on the local device.

By default, the interval for sending BFD packets is 1000 ms. The default value is recommended.

Step 7 (Optional) Run **min-rx-interval interval**

The interval for receiving BFD packets is set on the local device.

By default, the interval for receiving BFD packets is 1000 ms. The default value is recommended.

Step 8 (Optional) Run **detect-multiplier multiplier**

The local BFD detection multiplier is set.

The default value is 3.

Step 9 Run **commit**

The configuration is committed.

----End

Follow-up Procedure

When the BFD session is established and its status is Up, the BFD starts to detect failure in an LDP LSP.

When the LDP LSP is deleted, the BFD status turns Down.

2.7.2 Configuring BFD with Specific Parameters on the Egress Node

Context

BFD parameters on the egress node include the local and remote discriminators, intervals for sending and receiving BFD packets, and local BFD detection multiplier. The BFD parameters affect BFD session setup.

You can adjust the local detection time according to the network situation. On an unstable link, if a small detection time is used, a BFD session may flap. You can increase the detection time of the BFD session.

NOTE

Actual interval for the local device to send BFD packets = MAX {locally configured interval for sending BFD packets, remotely configured interval for receiving BFD packets}

Actual interval for the local device to receive BFD packets = MAX {remotely configured interval for sending BFD packets, locally configured interval for receiving BFD packets}

Local detection time = Actual interval for receiving BFD packets x Remotely configured BFD detection multiplier

Perform the following steps on the egress node of the LSP.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Run **bfd**

This node is enabled with global BFD. The global BFD view is displayed.

By default, global BFD is disabled.

Step 3 Run **quit**

Return to the system view.

Step 4 Configure a reverse tunnel to inform the ingress node of a fault if the fault occurs. The reverse tunnel can be the IP link, or LDP LSP. To ensure that BFD packets are received and sent along the same path, an LSP tunnel is preferentially used to inform the egress node of an LSP fault. If the configured reverse tunnel requires BFD detection, configure a pair of BFD sessions for it. Run the following commands as required.

- For the IP link, run **bfd session-name bind peer-ip ip-address [vpn-instance vpn-name] [source-ip ip-address]**
- For the LDP LSP, run **bfd session-name bind ldp-lsp peer-ip ip-address nexthop ip-address [interface interface-type interface-number]**

Step 5 Set local and remote discriminators of a BFD session.

- Run **discriminator local discr-value**
The local discriminator is configured.

- Run **discriminator remote** *discr-value*
The remote discriminator is configured.

 NOTE

The local identifier and remote identifier on both ends of a BFD session must accord with each other. The session cannot be established correctly otherwise. In addition, the local identifier and remote identifier cannot be modified after configuration.

Step 6 (Optional) Run **min-tx-interval** *interval*

The interval for sending BFD packets is set on the local device.

By default, the interval for sending BFD packets is 1000 ms. The default value is recommended.

Step 7 (Optional) Run **min-rx-interval** *interval*

The interval for receiving BFD packets is set on the local device.

By default, the interval for receiving BFD packets is 1000 ms. The default value is recommended.

Step 8 (Optional) Run **detect-multiplier** *multiplier*

The local BFD detection multiplier is set.

The default value is 3.

Step 9 Run **commit**

The configuration is committed.

----End

2.7.3 Verifying the Configuration of Static BFD for LDP LSPs

Prerequisites

The configurations of the static BFD for LDP LSP are complete.

Procedure

- Run the **display bfd session { all | static }** command to check information about the BFD session.
- Run the **display bfd statistics session { all | static }** command to check statistics about BFD.

----End

2.8 Configuring Synchronization Between LDP and IGP

Synchronization between LDP and IGP applies to MPLS networks where primary and backup LSPs exist. LSPs are established between LSRs based on IGP. When the LDP session or LDP adjacency on the primary LSP fails (not due to a link failure) or the faulty primary LSP is restored, you can enable synchronization between LDP and IGP to prevent traffic interruption caused by the active/standby switchover.

Pre-configuration Tasks

Before configuring synchronization between LDP and IGP, configure a local LDP session. For details, see [2.6 Configuring Basic Functions of MPLS LDP](#).

Configuration Procedure

Enabling synchronization between LDP and IGP is mandatory and other tasks are optional.

2.8.1 Enabling Synchronization Between LDP and IGP

Context

Synchronization between LDP and IGP can be configured in either of the following modes:

- Enable this function in the interface view.
This mode allows synchronization between LDP and IGP to be enabled on interfaces. This mode applies to the scenario where a few interfaces need to support this function.
- Enable this function in an IGP process.
This mode allows synchronization between LDP and IGP to be enabled on all interfaces in the IGP process. This mode applies to the scenario where many interfaces on a node need to support this function.

Procedure

- If OSPF is used as an IGP, perform the following steps:

Enable synchronization between LDP and OSPF on an interface.

- a. Run **system-view**

The system view is displayed.

- b. Run **interface** *interface-type interface-number*

The interface view is displayed.

- c. On an Ethernet interface, run **undo portswitch**

The interface is switched to Layer 3 mode.

By default, an Ethernet interface works in Layer 2 mode.

The mode switching function takes effect when the interface only has attribute configurations (for example, **shutdown** and **description** configurations). Alternatively, if configuration information supported by both Layer 2 and Layer 3 interfaces exists (for example, **mode lacp** and **lacp system-id** configurations), no configuration that is not supported after the working mode of the interface is switched can exist. If unsupported configurations exist on the interface, delete the configurations first and then run the **undo portswitch** command.

 NOTE

If many Ethernet interfaces need to be switched to Layer 3 mode, run the **undo portswitch batch** *interface-type* { *interface-number1* [**to** *interface-number2*] } &<1-10> command in the system view to switch these interfaces to Layer 3 mode in batches.

d. Run **ospf ldp-sync**

Synchronization between LDP and OSPF is enabled on the protected interface.

By default, synchronization between LDP and OSPF is disabled on an interface.

 NOTE

To enable LDP-OSPF synchronization on a multi-area adjacency interface, run the **ospf ldp-sync multi-area** *area-id* command.

e. Run **commit**

The configuration is committed.

Enable synchronization between LDP and OSPF in an OSPF process.

a. Run **system-view**

The system view is displayed.

b. Run **ospf** [*process-id*]

A specified OSPF process is started and the OSPF view is displayed.

process-id specifies an OSPF process. If *process-id* is not specified, the default OSPF process ID of the system is 1.

c. Run **area** *area-id*

The OSPF area view is displayed.

d. Run **ldp-sync enable**

Synchronization between LDP and OSPF is enabled on all interfaces in the specified OSPF area.

By default, synchronization between LDP and OSPF is disabled on all interfaces in an OSPF area.

e. Run **commit**

The configuration is committed.

- If IS-IS is used as an IGP, perform the following steps:

Enable synchronization between LDP and IS-IS on an interface.

a. Run **system-view**

The system view is displayed.

b. Run **interface** *interface-type interface-number*

The interface view is displayed.

c. On an Ethernet interface, run **undo portswitch**

The interface is switched to Layer 3 mode.

By default, an Ethernet interface works in Layer 2 mode.

The mode switching function takes effect when the interface only has attribute configurations (for example, **shutdown** and **description** configurations). Alternatively, if configuration information supported by both Layer 2 and Layer 3 interfaces exists (for example, **mode lacp** and **lacp system-id** configurations), no configuration that is not supported after the working mode of the interface is switched can exist. If unsupported configurations exist on the interface, delete the configurations first and then run the **undo portswitch** command.

 NOTE

If many Ethernet interfaces need to be switched to Layer 3 mode, run the **undo portswitch batch** *interface-type* { *interface-number1* [**to** *interface-number2*] } &<1-10> command in the system view to switch these interfaces to Layer 3 mode in batches.

- d. Run **isis enable** *process-id*

IS-IS is enabled.

- e. Run **isis ldp-sync**

Synchronization between LDP and IS-IS is enabled on the protected interface.

By default, synchronization between LDP and IS-IS is disabled on an interface.

- f. Run **commit**

The configuration is committed.

Enable synchronization between LDP and IS-IS in an IS-IS process.

- a. Run **system-view**

The system view is displayed.

- b. Run **isis** [*process-id*]

A specified IS-IS process is started and the IS-IS view is displayed.

process-id specifies an IS-IS process. If *process-id* is not specified, the default IS-IS process ID of the system is 1.

- c. Run **ldp-sync enable**

Synchronization between LDP and IS-IS is enabled on all interfaces in the specified IS-IS process.

By default, synchronization between LDP and IS-IS is disabled on all interfaces in an IS-IS process.

- d. Run **commit**

The configuration is committed.

----End

2.8.2 (Optional) Blocking Synchronization Between LDP and IGP on an Interface

Context

After the **ldp-sync enable** command is run in an IGP process, the following situations occur:

- On a P2P network, the interfaces whose neighbor status is Up implement synchronization between LDP and IGP.
- On a broadcast network, if the IGP is OSPF, the interfaces (whose neighbor status is Up) connecting the DR and a non-DR (or BDR) implement synchronization between LDP and IGP; if the IGP is IS-IS, the interfaces (whose neighbor status is Up) connecting the DIS and a non-DIS implement synchronization between LDP and IGP.

For the interfaces on a device bearing key services, ensure that this device detours around the backup path. You can block synchronization between LDP and IGP on a specified interface of the switch.

Procedure

- If OSPF is used for IGP, perform the following steps on the interfaces on both ends of the link between the node (where the primary LSP and the backup LSP diverge from each other) and its LDP peer on the primary LSP.

- a. Run **system-view**

The system view is displayed.

- b. Run **interface** *interface-type interface-number*

The OSPF interface view is displayed.

- c. On an Ethernet interface, run **undo portswitch**

The interface is switched to Layer 3 mode.

By default, an Ethernet interface works in Layer 2 mode.

The mode switching function takes effect when the interface only has attribute configurations (for example, **shutdown** and **description** configurations). Alternatively, if configuration information supported by both Layer 2 and Layer 3 interfaces exists (for example, **mode lacp** and **lacp system-id** configurations), no configuration that is not supported after the working mode of the interface is switched can exist. If unsupported configurations exist on the interface, delete the configurations first and then run the **undo portswitch** command.

NOTE

If many Ethernet interfaces need to be switched to Layer 3 mode, run the **undo portswitch batch** *interface-type { interface-number1 [to interface-number2] }* &<1-10> command in the system view to switch these interfaces to Layer 3 mode in batches.

- d. Run **ospf ldp-sync block**

Synchronization between LDP and OSPF is blocked on the interface.

By default, synchronization between LDP and OSPF is not blocked on an interface.

 NOTE

To block LDP-OSPF synchronization on a multi-area adjacency interface, run the **ospf ldp-sync block multi-area** *area-id* command.

e. Run **commit**

The configuration is committed.

- If IS-IS is used for IGP, perform the following steps on the interfaces on both ends of the link between the node (where the primary LSP and the backup LSP diverge from each other) and its LDP peer on the primary LSP.

a. Run **system-view**

The system view is displayed.

b. Run **interface** *interface-type interface-number*

The IS-IS interface view is displayed.

c. On an Ethernet interface, run **undo portswitch**

The interface is switched to Layer 3 mode.

By default, an Ethernet interface works in Layer 2 mode.

The mode switching function takes effect when the interface only has attribute configurations (for example, **shutdown** and **description** configurations). Alternatively, if configuration information supported by both Layer 2 and Layer 3 interfaces exists (for example, **mode lacp** and **lacp system-id** configurations), no configuration that is not supported after the working mode of the interface is switched can exist. If unsupported configurations exist on the interface, delete the configurations first and then run the **undo portswitch** command.

 NOTE

If many Ethernet interfaces need to be switched to Layer 3 mode, run the **undo portswitch batch** *interface-type { interface-number1 [to interface-number2] }* &<1-10> command in the system view to switch these interfaces to Layer 3 mode in batches.

d. Run **isis ldp-sync block**

Synchronization between LDP and IS-IS is blocked on the interface.

By default, synchronization between LDP and IS-IS is not blocked on an interface.

e. Run **commit**

The configuration is committed.

----End

2.8.3 (Optional) Setting the Hold-down Timer Value

Context

On a device that has LDP-IGP synchronization enabled, if the active physical link recovers, an IGP enters the Hold-down state, and a Hold-down timer starts. Before the Hold-down timer expires, the IGP delays establishing an IGP neighbor relationship until an LDP session or LDP adjacency is established over the active link so that the LDP session over and IGP route for the active link can become available simultaneously.

NOTE

If IS-IS is used, you can set the value of the Hold-down timer on a specified interface or set the value of the Hold-down timer for all IS-IS interfaces in the IS-IS view.

If different Hold-down values on an interface and in an IS-IS process are set, the setting on the interface takes effect.

Procedure

- If OSPF is used as an IGP, perform the following steps:

- a. Run **system-view**

The system view is displayed.

- b. Run **interface** *interface-type interface-number*

The interface view is displayed.

- c. On an Ethernet interface, run **undo portswitch**

The interface is switched to Layer 3 mode.

By default, an Ethernet interface works in Layer 2 mode.

The mode switching function takes effect when the interface only has attribute configurations (for example, **shutdown** and **description** configurations). Alternatively, if configuration information supported by both Layer 2 and Layer 3 interfaces exists (for example, **mode lacp** and **lacp system-id** configurations), no configuration that is not supported after the working mode of the interface is switched can exist. If unsupported configurations exist on the interface, delete the configurations first and then run the **undo portswitch** command.

NOTE

If many Ethernet interfaces need to be switched to Layer 3 mode, run the **undo portswitch batch** *interface-type { interface-number1 [to interface-number2] }* &<1-10> command in the system view to switch these interfaces to Layer 3 mode in batches.

- d. Run **ospf timer ldp-sync hold-down** *value*

The interval during which OSPF waits for an LDP session or LDP adjacency to be established is set.

By default, the Hold-down timer value is 10 seconds.

 NOTE

To configure a value for the Hold-down timer on an OSPF multi-area adjacency interface, run the **ospf timer ldp-sync hold-down value multi-area area-id** command.

- e. Run **commit**

The configuration is committed.

- If IS-IS is used as an IGP, perform the following steps:

Set the Hold-down timer on a specified IS-IS interface.

- a. Run **system-view**

The system view is displayed.

- b. Run **interface interface-type interface-number**

The interface view is displayed.

- c. On an Ethernet interface, run **undo portswitch**

The interface is switched to Layer 3 mode.

By default, an Ethernet interface works in Layer 2 mode.

The mode switching function takes effect when the interface only has attribute configurations (for example, **shutdown** and **description** configurations). Alternatively, if configuration information supported by both Layer 2 and Layer 3 interfaces exists (for example, **mode lacp** and **lacp system-id** configurations), no configuration that is not supported after the working mode of the interface is switched can exist. If unsupported configurations exist on the interface, delete the configurations first and then run the **undo portswitch** command.

 NOTE

If many Ethernet interfaces need to be switched to Layer 3 mode, run the **undo portswitch batch interface-type { interface-number1 [to interface-number2] } &<1-10>** command in the system view to switch these interfaces to Layer 3 mode in batches.

- d. Run **isis timer ldp-sync hold-down value**

The interval during which IS-IS waits for an LDP session or LDP adjacency to be established is set.

By default, the Hold-down timer value is 10 seconds.

- e. Run **commit**

The configuration is committed.

Set the Hold-down timer on all IS-IS interfaces in a specified IS-IS process.

- a. Run **system-view**

The system view is displayed.

- b. Run **isis [process-id]**

The IS-IS process view is displayed.

c. Run **timer ldp-sync hold-down** *value*

The Hold-down timer is set, which enables all IS-IS interfaces within an IS-IS process to delay establishing IS-IS neighbor relationships until LDP sessions or LDP adjacency are established.

By default, the Hold-down timer value is 10 seconds.

d. Run **commit**

The configuration is committed.

----End

2.8.4 (Optional) Setting the Hold-max-cost Timer Value

Context

If an LDP session or LDP adjacency over the active link fails but an IGP route for the active link is reachable, a node that has LDP-IGP synchronization enabled uses a Hold-max-cost timer to enable an IGP to advertise LSAs or LSPs carrying the maximum route cost, which delays IGP route convergence until an LDP session or LDP adjacency is established. Therefore, an IGP route for a standby link and an LDP session over the standby link can become available simultaneously.

You can set the Hold-max-cost timer value in either of the following methods:

- Setting the Hold-max-cost timer value in the interface view

You can set the Hold-max-cost timer value on a specified interface. This mode applies to the scenario where a few interfaces need to use the Hold-max-cost timer.

- Setting the Hold-max-cost timer value in the IGP process

After you set the Hold-max-cost timer value in the IGP process, the Hold-max-cost timers on all interfaces in the IGP process are set to this value. This mode applies to the scenario where many interfaces on a node need to use the Hold-max-cost timer.

Select parameters based on networking requirements:

- If an IGP carries only LDP services, configure the parameter **infinite** to ensure that a selected IGP route is kept consistent with the LDP LSP.
- If an IGP carries multiple types of services including LDP services, set the value of the parameter *value* to ensure that a teardown of LDP sessions or LDP adjacency does not affect IGP route selection or other services.

Procedure

- If OSPF is used as an IGP, perform the following steps:

Set the Hold-max-cost timer on a specified OSPF interface.

a. Run **system-view**

The system view is displayed.

b. Run **interface** *interface-type interface-number*

The interface view is displayed.

- c. On an Ethernet interface, run **undo portswitch**

The interface is switched to Layer 3 mode.

By default, an Ethernet interface works in Layer 2 mode.

The mode switching function takes effect when the interface only has attribute configurations (for example, **shutdown** and **description** configurations). Alternatively, if configuration information supported by both Layer 2 and Layer 3 interfaces exists (for example, **mode lacp** and **lacp system-id** configurations), no configuration that is not supported after the working mode of the interface is switched can exist. If unsupported configurations exist on the interface, delete the configurations first and then run the **undo portswitch** command.

 NOTE

If many Ethernet interfaces need to be switched to Layer 3 mode, run the **undo portswitch batch** *interface-type* { *interface-number1* [**to** *interface-number2*] } &<1-10> command in the system view to switch these interfaces to Layer 3 mode in batches.

- d. Run **ospf timer ldp-sync hold-max-cost** { *value* | **infinite** }

The interval for advertising the maximum cost in the LSAs of local LSRs through OSPF is set.

By default, the local device running OSPF keeps advertising the maximum cost by sending LSAs.

 NOTE

To set a value for the hold-max-cost timer on an OSPF multi-area adjacency interface, run the **ospf timer ldp-sync hold-max-cost** { *value* | **infinite** } **multi-area** *area-id* command.

- e. Run **commit**

The configuration is committed.

Set the Hold-max-cost timer on all OSPF interfaces in a specified OSPF process.

- a. Run **system-view**

The system view is displayed.

- b. Run **ospf** [*process-id*]

A specified OSPF process is started and the OSPF view is displayed.

- c. Run **area** *area-id*

The OSPF area view is displayed.

- d. Run **timer ldp-sync hold-max-cost** { **infinite** | *interval* }

The Hold-max-cost timer is set, which enables OSPF to keep advertising LSPs carrying the maximum route cost on all interfaces within an OSPF process.

By default, the local device running OSPF keeps advertising the maximum cost by sending LSAs.

e. Run **commit**

The configuration is committed.

- If IS-IS is used as an IGP, perform the following steps:

Set the Hold-max-cost timer on a specified IS-IS interface.

a. Run **system-view**

The system view is displayed.

b. Run **interface** *interface-type interface-number*

The interface view is displayed.

c. On an Ethernet interface, run **undo portswitch**

The interface is switched to Layer 3 mode.

By default, an Ethernet interface works in Layer 2 mode.

The mode switching function takes effect when the interface only has attribute configurations (for example, **shutdown** and **description** configurations). Alternatively, if configuration information supported by both Layer 2 and Layer 3 interfaces exists (for example, **mode lacp** and **lacp system-id** configurations), no configuration that is not supported after the working mode of the interface is switched can exist. If unsupported configurations exist on the interface, delete the configurations first and then run the **undo portswitch** command.

 NOTE

If many Ethernet interfaces need to be switched to Layer 3 mode, run the **undo portswitch batch** *interface-type { interface-number1 [to interface-number2] }* &<1-10> command in the system view to switch these interfaces to Layer 3 mode in batches.

d. Run **isis timer ldp-sync hold-max-cost { value | infinite }**

The value of the Hold-max-cost timer is set.

By default, the local device running IS-IS keeps advertising the maximum cost in Link State PDUs.

e. Run **commit**

The configuration is committed.

Set the Hold-max-cost timer on all IS-IS interfaces in a specified IS-IS process.

a. Run **system-view**

The system view is displayed.

b. Run **isis** [*process-id*]

A specified IS-IS process is restarted and the IS-IS view is displayed.

c. Run **timer ldp-sync hold-max-cost { infinite | interval }**

The Hold-max-cost timer is set, which enables IS-IS to keep advertising LSPs carrying the maximum route cost on all interfaces within an IS-IS process.

By default, the local device running IS-IS keeps advertising the maximum cost in Link State PDUs.

- d. Run **commit**

The configuration is committed.

----End

2.8.5 (Optional) Setting the Delay Timer Value

Context

When an LDP session or LDP adjacency is reestablished on a faulty link, LDP starts the Delay timer to wait for the establishment of an LSP. After the Delay timer times out, LDP notifies the IGP that synchronization between LDP and IGP is complete.

You can set the Delay timer value in either of the following methods:

- Setting the Delay timer value in the interface view
You can set the Delay timer value on a specified interface. This mode applies to the scenario where a few interfaces need to use the Delay timer.
- Setting the Delay timer value in the MPLS LDP view
After you set the Delay timer value in the MPLS LDP view, the Delay timers on all interfaces in the MPLS LDP view are set to this value. This mode applies to the scenario where many interfaces on a node need to use the Delay timer.

Procedure

- Perform the following steps in the MPLS-LDP view.
 - a. Run **system-view**
The system view is displayed.
 - b. Run **mpls ldp**
The MPLS-LDP view is displayed.
 - c. Run **igp-sync-delay timer value**
The period of waiting for the LSP setup after the establishment of the LDP session or LDP adjacency is set.
By default, the value of the delay timer is 10s.
 - d. Run **commit**
The configuration is committed.
- Perform the following steps in the interface view.
 - a. Run **system-view**
The system view is displayed.
 - b. Run **interface interface-type interface-number**
The interface view is displayed.
 - c. On an Ethernet interface, run **undo portswitch**

The interface is switched to Layer 3 mode.

By default, an Ethernet interface works in Layer 2 mode.

The mode switching function takes effect when the interface only has attribute configurations (for example, **shutdown** and **description** configurations). Alternatively, if configuration information supported by both Layer 2 and Layer 3 interfaces exists (for example, **mode lacp** and **lacp system-id** configurations), no configuration that is not supported after the working mode of the interface is switched can exist. If unsupported configurations exist on the interface, delete the configurations first and then run the **undo portswitch** command.

NOTE

If many Ethernet interfaces need to be switched to Layer 3 mode, run the **undo portswitch batch** *interface-type { interface-number1 [to interface-number2] }* &<1-10> command in the system view to switch these interfaces to Layer 3 mode in batches.

- d. Run **mpls ldp timer igp-sync-delay** *value*

The period of waiting for the LSP setup after the establishment of the LDP session or LDP adjacency is set.

By default, the value of the delay timer is 10s.

- e. Run **commit**

The configuration is committed.

----End

2.8.6 (Optional) Configuring LDP Graceful Deletion

Context

LDP graceful deletion can be configured in the synchronization between LDP and IGP or LDP FRR scenario to speed up traffic switching and ensure uninterrupted traffic during traffic switching, improving reliability of the entire network.

If the physical and protocol status of the primary link is proper but the LDP session on the primary link goes Down, Synchronization between LDP and IGP enables LDP to inform the IGP of the primary link fault, and the IGP configures the maximum cost value for the primary link. LDP then immediately instructs the upstream device to withdraw labels and allocates labels to the upstream device because a new LSP is set up on the secondary link, prolonging the LSP convergence time. As a result, packet loss occurs.

After the LDP session on the faulty link goes Down, LDP does not immediately instruct the upstream device to withdraw labels but keeps the labels and LSP and allows traffic to be transmitted on the primary link until the LSP convergence is complete on the secondary link. This ensures uninterrupted traffic and speeds up the synchronization between LDP and IGP.

Perform the following steps on the LSR configured with synchronization between LDP and IGP.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Run **mpls ldp**

The MPLS-LDP view is displayed.

Step 3 Run **graceful-delete**

LDP graceful deletion is enabled.

Step 4 (Optional) Run **graceful-delete timer time**

The value of the graceful deletion timer is set.

After the LDP session goes Down, LDP does not instruct the upstream device to withdraw labels until the graceful deletion timer expires.

By default, the value of the graceful deletion timer is 5 seconds.

 NOTE

If the value of the graceful deletion timer is too large, the invalid LSPs will be kept for a long time, consuming system resources.

Step 5 Run **commit**

The configuration is committed.

----End

2.8.7 Verifying the Configuration of Synchronization Between LDP and IGP

Prerequisites

The configurations of the synchronization between LDP and IGP function are complete.

Procedure

- Run the **display mpls ldp** command to check the global LDP configuration.
- Run the **display ospf [process-id] ldp-sync interface { all | interface-type interface-number }** command to check information about synchronization between LDP and OSPF on an interface.
- Run the **display isis [process-id] ldp-sync interface** command to check information about synchronization between LDP and IS-IS on the interface.

----End

2.9 Configuring Auto LDP FRR

By configuring Auto LDP FRR, you can quickly switch traffic to the backup LSP when a link fails, which ensures uninterrupted traffic transmission.

Pre-configuration Tasks

- Before configuring Auto LDP FRR, complete the following tasks:
 - Configure a local LDP session. For details, see [2.6 Configuring Basic Functions of MPLS LDP](#).
 - Configure Auto IGP FRR. For details, see *CloudEngine 9800, 8800, 6800, 5800 Series Switches Configuration Guide - IP Routing*.

Context

After IGP auto FRR is enabled using the **frr** command, auto LDP FRR is automatically enabled. If you need to adjust the backup LDP LSP establishment policy, configure auto LDP FRR on the ingress or transit node.

NOTE

The CE9860EI does not support Auto LDP FRR.

Procedure

- **Configuring Auto LDP FRR**
 - a. Run **system-view**

The system view is displayed.
 - b. Run **mpls ldp**

The MPLS-LDP view is displayed.
 - c. (Optional) Run **ipv4-family**

The MPLS-LDP-IPv4 view is displayed.
 - d. Run **auto-frr lsp-trigger { all | host | ip-prefix ip-prefix-name | none }**

A policy for triggering backup LDP LSP establishment is configured.

By default, LDP uses backup routes to addresses with 32-bit masks to set up backup LSPs.

NOTE

 - The **auto-frr lsp-trigger** command takes effect in either the MPLS LDP view or MPLS-LDP-IPv4 view. If the command is configured in both views, only the latest configuration takes effect.
 - If both the **auto-frr lsp-trigger** command and the **lsp-trigger** command are run, the established backup LSPs satisfy both the policy for triggering LDP LSP establishment and the policy for triggering backup LDP LSP establishment.
 - e. (Optional) Configuring LDP Graceful Deletion

If the primary link goes Down and the LDP session on the primary link goes Down accordingly, LDP immediately instructs the upstream device to withdraw labels and triggers LDP FRR, and the LSP convergence on the secondary link requires LDP to allocate labels to the upstream device again, prolonging the convergence time and traffic switching using FRR. As a result, packet loss occurs.

If LDP graceful deletion is configured and the LDP session goes Down, LDP delays the LDP session deletion and reserves the relevant labels and

LSP. The LSP on the secondary link does not require LDP to allocate labels to the upstream device again, shortening traffic switching using LDP FRR and reducing packet loss.

i. Run **graceful-delete**

LDP graceful deletion is enabled.

By default, LDP graceful deletion is disabled.

ii. Run **graceful-delete timer time**

The value of the graceful deletion timer is set.

By default, the value of the graceful deletion timer is 5 seconds.

After the LDP session goes Down, forwarding entries on the LSR remain unchanged for the time specified by the graceful deletion timer.

 NOTE

If the value of the graceful deletion timer is too large, the invalid LSPs will be kept for a long time, consuming system resources.

f. Run **commit**

The configuration is committed.

----End

Verifying the Configuration

Run the **display mpls lsp** command to check information about LSPs enabled with LDP Auto FRR.

2.10 Configuring LDP GR Helper

Configure the local node as the GR Helper to help the neighbor complete LDP GR.

Pre-configuration Tasks

Before configuring LDP GR Helper, complete the following tasks:

- Configure a local LDP session. For details, see [2.6 Configuring Basic Functions of MPLS LDP](#).
- Configure IGP GR. For details, see *CloudEngine 9800, 8800, 6800, 5800 Series Switches IP Unicast Routing Configuration Guide*.

Context

[Table 2-9](#) describes timers used during LDP GR.

Table 2-9 Timers used during LDP GR

Timer	Description	Suggestion
Reconnect timer	After the GR Restarter performs an active/standby switchover, the GR Helper detects that the LDP session with the GR Restarter fails, and then starts the Reconnect timer and waits for reestablishment of the LDP session. The value of the Reconnect timer that takes effect on the GR Helper is the smaller one between the value of the Neighbor-liveness timer set on the GR Helper and the value of Reconnect timer set on the GR Restarter.	When a network with a large number of routes is faulty, you can increase the value of the Reconnect timer to avoid that all the LDP sessions cannot recover within the default timeout period 300s.
Recovery timer	After the LDP session is reestablished, the GR Helper starts the Recovery timer and waits for the recovery of the LSP. The value of the Recovery timer that takes effect on the GR Helper is the smaller one between the value of the Recovery timer set on the GR Helper and the value of Recovery timer set on the GR Restarter.	When a network with a large number of routes is faulty, you can increase the value of the Recovery timer to avoid that all the LSPs cannot recover within the default timeout period 300s.
Neighbor-liveness timer	The value of the neighbor-liveness timer defines the LDP GR period. The value of the Neighbor-liveness timer on the GR Restarter is the same as that of the Forwarding State Holding timer.	When the number of LSPs on a network is small, you can set a smaller value for the Neighbor-liveness timer to shorten the GR period.

 NOTE

Enabling or disabling LDP GR, or changing the LDP GR timer value cause LDP session reestablishment. To disable a device from re-establishing LDP sessions when LDP GR is enabled or disabled, or the LDP GR timer value is changed, run the **no-renegotiate session-parameter-change graceful-restart** command.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Run **mpls ldp**

The MPLS LDP view is displayed.

Step 3 Run **graceful-restart**

The LDP GR function is enabled.

By default, the LDP GR function is disabled.

Step 4 (Optional) Run **graceful-restart timer reconnect** *time*

The Reconnect timer for the LDP session is set.

By default, the Reconnect timer is set to 300 seconds.

Step 5 (Optional) Run **graceful-restart timer recovery** *time*

The LSP Recovery timer is set.

By default, the LSP Recovery timer is set to 300 seconds.

Step 6 (Optional) Run **graceful-restart timer neighbor-liveness** *time*

The Neighbor-liveness timer is set.

By default, the Neighbor-liveness timer is 600 seconds.

Step 7 (Optional) Run **no-renegotiate session-parameter-change graceful-restart**

The device is disabled from re-establishing LDP sessions when LDP GR is enabled or disabled, or the LDP GR timer value is changed.

By default, enabling or disabling LDP GR, or changing the LDP GR timer value cause LDP session reestablishment.

Step 8 Run **commit**

The configuration is committed.

----End

Verifying the Configuration

- Run the **display mpls ldp [all [verbose]]** command to check information about LDP.
- Run the **display mpls ldp session [all | verbose]** command to check information about the LDP session.

2.11 Configuring LDP Security Mechanisms

LDP security mechanisms such as LDP MD5 authentication, LDP Keychain authentication, and LDP GTSM can be configured to meet high network security requirements.

Pre-configuration Tasks

Before configuring LDP security features, configure basic functions of MPLS LDP.
For details, see [2.6 Configuring Basic Functions of MPLS LDP](#).

Configuration Procedure

You can perform the following configuration tasks in any sequence as required.
You can configure only either of LDP MD5 authentication and Keychain authentication for one neighbor at the same time.

2.11.1 Configuring LDP MD5 Authentication

Context

MD5 authentication can be configured for a TCP connection over which an LDP session is established, improving security. Note that the peers of an LDP session can be configured with different encryption modes, but must be configured with a single password.

The MD5 algorithm is easy to configure and generates a single password which can be changed only manually. MD5 authentication applies to the network requiring short-period encryption.

Keychain authentication and MD5 authentication cannot be both configured on a single LDP peer. Note that MD5 encryption algorithm cannot ensure security. Keychain authentication is recommended.

NOTICE

Configuring LDP MD5 authentication may cause LDP session reestablishment, deletion of the LSP associated with the deleted LDP session, and MPLS service interruption.

NOTE

LDP authentication configurations are prioritized in descending order: for a single peer, for a specified peer group, for all peers. Keychain and MD5 configurations of the same priority are mutually exclusive. Keychain or MD5 authentication can be configured simultaneously for a specified LDP peer, for this LDP peer in a specified peer group, and for all LDP peers. The configuration with a higher priority takes effect. For example, if MD5 authentication is configured for Peer1 and then keychain authentication is configured for all LDP peers, MD5 authentication takes effect on Peer1. Keychain authentication takes effect on other peers.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Run **mpls ldp**

The MPLS-LDP view is displayed.

Step 3 Configure LDP MD5 authentication

- Configure LDP MD5 authentication for a single LDP peer.

Run **md5-password { plain | cipher } *peer-lsr-id* password**

MD5 authentication is configured and a password is set.

By default, LDP MD5 authentication is not performed between LDP peers.

The password can be set in either simple text or ciphertext. A simple password is a pre-configured character string that is recorded in a configuration file as it is. A ciphertext password is a character string that is encrypted using a specified algorithm before being recorded in a configuration file.

If you configure a simple password, it will be saved in the configuration file in simple text that has a high security risk. Therefore, configuring a ciphertext password is recommended. To improve the device security, periodically change the password.

- Configure LDP MD5 authentication for LDP peers in a specified LDP peer group.

- a. Run **md5-password { plain | cipher } peer-group *ip-prefix-name* password**

MD5 authentication is enabled and a password is set for LDP peers in a specified LDP peer group.

An IP prefix list can be specified using *ip-prefix-name* to define the range of IP addresses in a group. Before using an IP prefix list, ensure that the IP prefix list must have been created.

- b. (Optional) Run **authentication exclude peer *peer-id***

The device is disabled from authenticating a specified LDP peer.

By default, after LDP MD5 authentication is enabled for a specified LDP peer group, MD5 authentication takes effect on all LDP peers in the group. To disable the device from authenticating a specified LDP peer in the group, perform this step.

- Configure LDP MD5 authentication for all LDP peers.

- a. Run **md5-password { plain | cipher } all password**

MD5 authentication is enabled and a password is set for all LDP peers.

- b. (Optional) Run **authentication exclude peer *peer-id***

The device is disabled from authenticating a specified LDP peer.

By default, after LDP MD5 authentication is enabled for all LDP peers, MD5 authentication takes effect on all LDP peers. To disable the device from authenticating a specified LDP peer, perform this step.

Step 4 Run commit

The configuration is committed.

----End

2.11.2 Configuring LDP Keychain Authentication

Context

To help improve LDP session security, keychain authentication can be configured for a TCP connection over which an LDP session has been established.

Keychain authentication involves a set of passwords and uses a new password when the previous one expires. Keychain authentication is complex to configure and applies to a network requiring high security.

You cannot configure keychain authentication and MD5 authentication for a neighbor at the same time.

Before configuring LDP keychain authentication, configure keychain globally. For details about the keychain configuration, see the *Keychain Configuration in CloudEngine 9800, 8800, 6800, 5800 Series Switches Configuration Guide - Security*.

NOTICE

Configuring LDP keychain authentication may cause LDP session reestablishment, deletion of the LSP associated with the deleted LDP session, and MPLS service interruption.

NOTE

LDP authentication configurations are prioritized in descending order: for a single peer, for a specified peer group, for all peers. Keychain and MD5 configurations of the same priority are mutually exclusive. Keychain or MD5 authentication can be configured simultaneously for a specified LDP peer, for this LDP peer in a specified peer group, and for all LDP peers. The configuration with a higher priority takes effect. For example, if MD5 authentication is configured for Peer1 and then keychain authentication is configured for all LDP peers, MD5 authentication takes effect on Peer1. Keychain authentication takes effect on other peers.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Run **mpls ldp**

The MPLS-LDP view is displayed.

Step 3 Configure LDP keychain authentication.

- Configure LDP keychain authentication for a specified LDP peer.
Run **authentication key-chain peer *peer-id* name *keychain-name***
LDP keychain is enabled and a keychain name is specified.
By default, LDP keychain authentication is not performed between LDP peers.
- Configure LDP keychain authentication for LDP peers in a specified LDP peer group.

- a. Run **authentication key-chain peer-group** *ip-prefix-name name keychain-name*
LDP keychain is enabled and a keychain name is specified for a specified LDP peer group.
An IP prefix list can be specified using *ip-prefix-name* to define the range of IP addresses in a group. Before using an IP prefix list, ensure that the IP prefix list must have been created.
- b. (Optional) Run **authentication exclude peer** *peer-id*
The device is disabled from authenticating a specified LDP peer.
By default, after LDP keychain authentication is enabled for a specified LDP peer group, keychain authentication takes effect on all LDP peers in the group. To disable the device from authenticating a specified LDP peer, perform this step.
- Configure LDP keychain authentication for all LDP peers.
 - a. Run **authentication key-chain all name** *keychain-name*
LDP keychain is enabled and a keychain name is specified for all LDP peers.
 - b. (Optional) Run **authentication exclude peer** *peer-id*
The device is disabled from authenticating a specified LDP peer.
By default, after LDP keychain authentication is enabled for all LDP peers, keychain authentication takes effect on all LDP peers. To disable the device from authenticating a specified LDP peer, perform this step.

Step 4 Run **commit**

The configuration is committed.

----End

2.11.3 Configuring the LDP GTSM

Context

To protect device from attacks, Generalized TTL Security Mechanism (GTSM) checks the TTL value of a packet to check whether the packet is valid. To check the TTL value of an LDP packet exchanged between LDP peers, enable GTSM on LDP peers and set the TTL range. If the TTL of an LDP packet is out of the TTL range, the LDP packet is considered as an invalid attack packet and discarded. This prevents the CPU from processing a large number of forged LDP packets. In this way, the upper layer protocols are protected.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Run **mpls ldp**

The MPLS-LDP view is displayed.

Step 3 Run **gtsm peer ip-address valid-ttl-hops hops**

The LDP GTSM is configured.

By default, no LDP peer is configured with the GTSM.

hops is the maximum number of valid hops permitted by the GTSM. If a TTL value carried in a received packet is in a specified range of $[255 - hops + 1, 255]$, the packet is accepted; if the TTL value is out of the range, the packet is discarded.

Step 4 Run **commit**

The configuration is committed.

----End

2.11.4 Verifying the Configuration of LDP Security Mechanisms

Prerequisites

The configurations of LDP security features are complete.

Procedure

- Run the **display mpls ldp session verbose** command to check the configurations of LDP MD5 authentication and LDP keychain authentication and TCP-AO authentication.
- Run the **display gtsm statistics { slot-id | all }** command to check GTSM statistics.

----End

2.12 Configuring Non-labeled Public Network Routes to Be Iterated to LSPs

By default, non-labeled public network routes can be iterated to outgoing interfaces and next hops, but cannot be iterated to tunnels. After this feature is configured, non-labeled public network routes can be iterated to LSPs.

Pre-configuration Tasks

Before configuring non-labeled public network routes to be iterated to LSPs, complete the following tasks:

- Configure a local LDP session. For details, see [2.6 Configuring Basic Functions of MPLS LDP](#).
- Configure an IP prefix list if non-labeled public network routes to be iterated to LSPs need to be limited.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Run **route recursive-lookup tunnel** [**ip-prefix** *ip-prefix-name*] [**tunnel-policy** *policy-name*]

The non-label public network route is allowed to be iterated to the LSP to forward through MPLS.

By default, the non-label public network route can be iterated only to the outbound interface and the next hop but not the LSP tunnel.

If **ip-prefix** *ip-prefix-name* is not set, all static routes and non-labeled public BGP routes will be preferentially iterated to LSP tunnels.

Step 3 Run **commit**

The configuration is committed.

----End

Verifying the Configuration

After non-labeled public routes are iterated to LSPs, you can run the **display bgp routing-table network** command to view route iteration information.

2.13 Maintaining MPLS LDP

2.13.1 Disabling LDP Session Flapping Suppression

Context

If an LDP session goes down due to a protocol or interface fault, LDP immediately attempts to reestablish the LDP session to ensure the fastest LDP hard convergence. For an LDP session that alternates between up and down multiple times within a period of time, the involved upstream and downstream LSPs are frequently created and deleted, wasting resources. To prevent this problem, LDP session flapping suppression is enabled by default. For a stable LDP network, you can disable LDP session flapping suppression.

Procedure

Step 1 Run the **system-view** command to enter the system view.

Step 2 Run the **mpls** command to enable MPLS globally and enter the MPLS view.

Step 3 Run the **mpls ldp** command to enable LDP globally and enter the MPLS-LDP view.

Step 4 Run the **session suppress disable** command to disable LDP session flapping suppression.

Step 5 Run the **commit** command to commit the configuration.

----End

2.13.2 Disabling LDP Interface Flapping Suppression

Context

If LDP interface flapping suppression is disabled and an interface frequently flaps, LDP frequently sends Address and Address Withdraw messages to all LDP sessions. If there are a large number of sessions, the CPU usage of the device increases, causing protocol flapping. To prevent this problem, LDP interface flapping suppression is enabled by default. For a stable LDP network, you can disable LDP interface flapping suppression.

Procedure

- Step 1** Run the **system-view** command to enter the system view.
- Step 2** Run the **mpls** command to enable MPLS globally and enter the MPLS view.
- Step 3** Run the **mpls ldp** command to enable LDP globally and enter the MPLS-LDP view.
- Step 4** Run the **suppress-flapping interface disable** command to disable LDP interface flapping suppression.
- Step 5** Run the **commit** command to commit the configuration.

----End

2.13.3 Resetting LDP

Context

NOTICE

- Resetting LDP may temporarily affect the reestablishment of the LSP. Take care to reset LDP.
 - Resetting LDP is prohibited during the LDP GR.
-

Procedure

- Run the **reset mpls ldp** command to reset configurations of the global LDP instance in the user view.
- Run the **reset mpls ldp all** command to reset configurations on all LDP instances in the user view.
- Run the **reset mpls ldp peer *peer-id*** command to reset a specified peer in the user view.
- To reset all GR-capable LDP peers in global LDP instances, run the **reset mpls ldp graceful** command in the user view to make new configurations take effect, achieving uninterrupted service transmission during a restart.

- To reset a specified GR-capable LDP peer, run the **reset mpls ldp peer *peer-id* graceful** command in the user view to make new configurations take effect, achieving uninterrupted service transmission during a restart.

----End

2.13.4 Clearing LDP Statistics

Context

NOTICE

The cleared LDP statistics cannot be restored. Exercise caution when you use the following commands.

Procedure

- Run the **reset mpls ldp error packet { tcp | udp | all }** command in the user view to clear statistics on LDP error messages.
- Run the **reset mpls ldp event adjacency-down** command in the user view to clear statistics on LDP adjacencies in Down state.
- Run the **reset mpls ldp event session-down** command in the user view to clear statistics on LDP sessions in Down state.

----End

2.13.5 Monitoring the LDP Running Status

Context

In routine maintenance, you can run the following commands in any view to view the LDP running status.

Procedure

- Run the **display mpls ldp error packet { tcp [peer *peer-id*] | udp [interface *interface-type* *interface-num*] }** command to check statistics on LDP error messages.
- Run the **display mpls ldp event adjacency-down [interface *interface-type* *interface-number* | remote] [peer *peer-id*] [verbose]** command to check information about LDP adjacencies in Down state.
- Run the **display mpls ldp event session-down [*peer-id*] [verbose]** command to check information about LDP sessions in Down state.

----End

2.13.6 Checking the LSP Connectivity

Context

In MPLS, the control panel used for setting up an LSP cannot detect data forwarding failures on the LSP. This makes network maintenance difficult.

MPLS ping checks LSP connectivity, and MPLS traceroute locates network faults in addition to checking LSP connectivity.

MPLS ping and MPLS traceroute can be performed in any view.

Procedure

- (Optional) Run the **lspv mpls-lsp-ping cpu-defend *cpu-defend*** command to set an upper limit for the rate of sending MPLS echo request packets to the main control board.
- (Optional) Run the **lspv echo-reply compitable fec enable** command to enable the device to respond to MPLS echo request packets with MPLS echo reply packets that do not carry FEC information.

NOTE

If the device interworks with a non-Huawei device, perform the preceding step.

- Run the **ping lsp** [**-a** *source-ip* | **-c** *count* | **-exp** *exp-value* | **-h** *ttl-value* | **-m** *interval* | **-r** *reply-mode* | **-s** *packet-size* | **-t** *time-out* | **-v** | **-g**] * **ip destination-iphost mask-length** [*ip-address*] [**nexthop** *nexthop-address*] command to perform an MPLS ping test.
- Run the **tracert lsp** [**-a** *source-ip* | **-exp** *exp-value* | **-h** *ttl-value* | **-r** *reply-mode* | **-t** *time-out* | **-g**] * **ip destination-iphost mask-length** [*ip-address*] [**nexthop** *nexthop-address*] [**detail**] command to perform an MPLS traceroute test.

Follow-up Procedure

- Run the **display lspv statistics** command to check the LSPV test statistics. A large amount of statistical information is saved in the system after MPLS ping or traceroute tests are performed multiple times, which is unhelpful for problem analysis. To obtain more accurate statistics, run the **reset lspv statistics** command to clear LSPV test statistics before running the **display lspv statistics** command.
- Run the **lspv mpls-lsp-ping echo disable** command to disable response to MPLS Echo Request packets. It is recommended that you run this command after completing an MPLS ping or traceroute test to save system resources.

2.14 Configuration Examples for MPLS LDP

This section only provides configuration examples for individual features. For details about multi-feature configuration examples, feature-specific configuration examples, interoperation examples, protocol or hardware replacement examples, and industry application examples, see the *Typical Configuration Examples*.

2.14.1 Example for Configuring Static BFD to Detect LDP LSPs

Networking Requirements

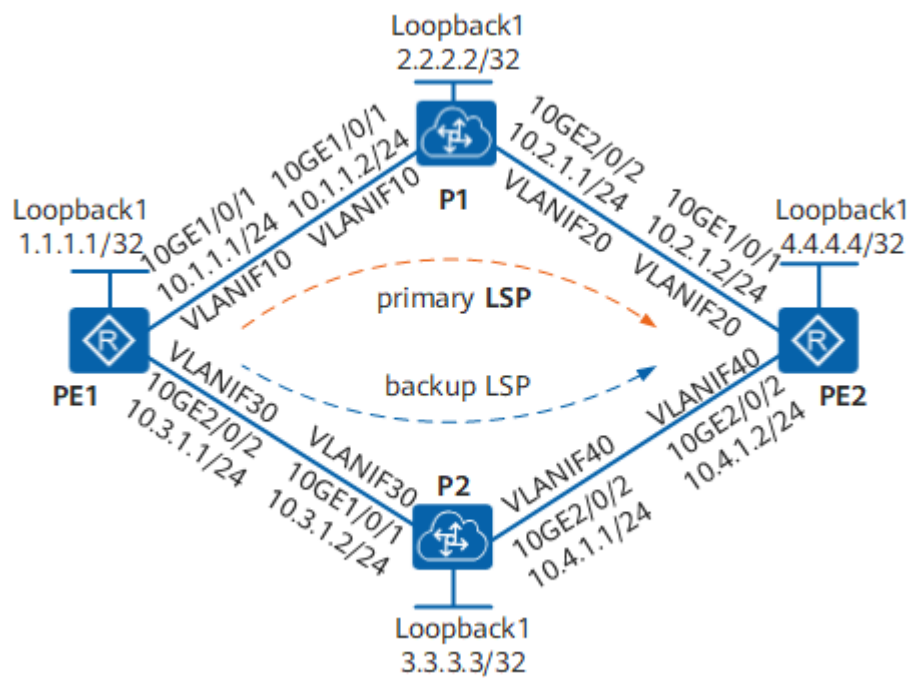
As shown in [Figure 2-15](#), the network topology is simple and stable, PEs and P are MPLS backbone network devices, and LDP LSPs are set up on the backbone network to transmit network services.

Network services, such as VoIP, online game, and online video service, have high requirements for real-timeness. Data loss caused by faulty links will seriously affect services. It is required that services be fast switched to the backup LSP when the primary LSP becomes faulty, minimizing packet loss. Static BFD for LDP LSPs is configured to fast detect LDP LSPs.

NOTICE

Ensure that STP is disabled in this scenario; otherwise, the primary LSP may be unavailable.

Figure 2-15 Networking diagram of configuring static BFD for LDP LSPs



Configuration Roadmap

The configuration roadmap is as follows:

1. Configure OSPF between the PEs and P to implement IP connectivity on the backbone network.
2. Configure local LDP sessions on PEs and P so that LDP LSPs can be set up to transmit network services.

3. Configure static BFD on PEs to fast detect LDP LSPs.

NOTE

CE switches cannot function as PE devices. The configurations of PE devices in this example are only for your reference. In real-world networking, you need to select the devices that can function as PE devices and configure the devices according to the product manual. It is recommended that NE routers be used as PE devices.

Procedure

- Step 1** Create VLANs and VLANIF interfaces on the switch, configure IP addresses for the VLANIF interfaces, and add physical interfaces to the VLANs.

Configure PE1. The configurations of P1, P2, and PE2 are similar to the configuration of PE1, and are not mentioned here.

```
<HUAWEI> system-view
[~HUAWEI] sysname PE1
[*HUAWEI] commit
[~PE1] interface loopback 1
[*PE1-LoopBack1] ip address 1.1.1.1 32
[*PE1-LoopBack1] quit
[*PE1] vlan batch 10 30
[*PE1] interface vlanif 10
[*PE1-Vlanif10] ip address 10.1.1.1 24
[*PE1-Vlanif10] quit
[*PE1] interface vlanif 30
[*PE1-Vlanif30] ip address 10.3.1.1 24
[*PE1-Vlanif30] quit
[*PE1] interface 10ge 1/0/1
[*PE1-10GE1/0/1] port link-type trunk
[*PE1-10GE1/0/1] port trunk allow-pass vlan 10
[*PE1-10GE1/0/1] quit
[*PE1] interface 10ge 2/0/2
[*PE1-10GE2/0/2] port link-type trunk
[*PE1-10GE2/0/2] port trunk allow-pass vlan 30
[*PE1-10GE2/0/2] quit
```

- Step 2** Configure OSPF to advertise the network segments connecting to interfaces on each node and to advertise the routes of hosts with LSR IDs.

Configure PE1. The configurations of P1, P2, and PE2 are similar to the configuration of PE1, and are not mentioned here.

```
[*PE1] ospf 1
[*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 10.1.1.0 0.0.0.255
[*PE1-ospf-1-area-0.0.0.0] network 10.3.1.0 0.0.0.255
[*PE1-ospf-1-area-0.0.0.0] quit
[*PE1-ospf-1] quit
```

- Step 3** Set the cost of VLANIF 30 on PE1 to 1000.

```
[*PE1] interface vlanif 30
[*PE1-Vlanif30] ospf cost 1000
[*PE1-Vlanif30] quit
[*PE1] commit
```

After the configuration is complete, run the **display ip routing-table** command on each node. You can see that the nodes learn routes from each other. The outbound interface of the route from PE1 to PE2 is VLANIF 10.

- Step 4** Configure local LDP sessions.

Configure PE1. The configurations of P1, P2, and PE2 are similar to the configuration of PE1, and are not mentioned here.

```
[~PE1] mpls lsr-id 1.1.1.1
[*PE1] mpls
[*PE1-mpls] quit
[*PE1] mpls ldp
[*PE1-mpls-ldp] quit
[*PE1] interface vlanif 10
[*PE1-Vlanif10] mpls
[*PE1-Vlanif10] mpls ldp
[*PE1-Vlanif10] quit
[*PE1] interface vlanif 30
[*PE1-Vlanif30] mpls
[*PE1-Vlanif30] mpls ldp
[*PE1-Vlanif30] quit
[*PE1] commit
```

Run the **display mpls ldp lsp** command. The command output shows that an LDP LSP destined for 4.4.4.4/32 is set up on PE1.

```
[~PE1] display mpls ldp lsp
LDP LSP Information
An asterisk (*) before an LSP means the LSP is not established
An asterisk (*) before a Label means the USCB or DSCB is stale
An asterisk (*) before an UpstreamPeer means the session is in GR state
An asterisk (*) before a DS means the session is in GR state
An asterisk (*) before a NextHop means the LSP is FRR LSP
```

DestAddress/Mask	In/OutLabel	UpstreamPeer	NextHop	OutInterface
1.1.1.1/32	3/NULL	2.2.2.2	127.0.0.1	InLoop0
1.1.1.1/32	3/NULL	3.3.3.3	127.0.0.1	InLoop0
*1.1.1.1/32	Liberal/1024		DS/2.2.2.2	
*1.1.1.1/32	Liberal/1024		DS/3.3.3.3	
2.2.2.2/32	NULL/3	-	10.1.1.2	Vlanif10
2.2.2.2/32	1024/3	2.2.2.2	10.1.1.2	Vlanif10
2.2.2.2/32	1024/3	3.3.3.3	10.1.1.2	Vlanif10
*2.2.2.2/32	Liberal/1025		DS/3.3.3.3	
3.3.3.3/32	NULL/1026	-	10.1.1.2	Vlanif10
3.3.3.3/32	1026/1026	2.2.2.2	10.1.1.2	Vlanif10
3.3.3.3/32	1026/1026	3.3.3.3	10.1.1.2	Vlanif10
*3.3.3.3/32	Liberal/3		DS/3.3.3.3	
4.4.4.4/32	NULL/1025	-	10.1.1.2	Vlanif10
4.4.4.4/32	1025/1025	2.2.2.2	10.1.1.2	Vlanif10
4.4.4.4/32	1025/1025	3.3.3.3	10.1.1.2	Vlanif10
*4.4.4.4/32	Liberal/1026		DS/3.3.3.3	

```
-----
TOTAL: 11 Normal LSP(s) Found, 5 Liberal LSP(s) Found
       0 FRR LSP(s) Found.
```

Step 5 Enable global BFD on the two nodes of the detected link.

Configure PE1.

```
[~PE1] bfd
[*PE1-bfd] quit
```

Configure PE2.

```
[~PE2] bfd
[*PE2-bfd] quit
```

Step 6 Bind the BFD session destined for the LDP LSP on PE1.

Configure PE1.

```
[*PE1] bfd pe1tope2 bind ldp-lsp peer-ip 4.4.4.4 nexthop 10.1.1.2 interface vlanif 10
[*PE1-bfd-lsp-session-pe1tope2] discriminator local 1
```

```
[*PE1-bfd-lsp-session-pe1tope2] discriminator remote 2
[*PE1-bfd-lsp-session-pe1tope2] commit
[~PE1-bfd-lsp-session-pe1tope2] quit
```

Step 7 On PE2, configure a BFD session that is bound to the IP link to notify PE1 of the detected faults on the LDP LSP.

Configure PE2.

```
[*PE2] bfd pe2tope1 bind peer-ip 1.1.1.1
[*PE2-bfd-session-pe2tope1] discriminator local 2
[*PE2-bfd-session-pe2tope1] discriminator remote 1
[*PE2-bfd-session-pe2tope1] commit
[~PE2-bfd-session-pe2tope1] quit
```

Step 8 Verify the configuration.

Run the **display bfd session all** command on PE1. The command output shows that the **State** field is displayed as **Up**.

```
[~PE1] display bfd session all
S: Static session
D: Dynamic session
IP: IP session
IF: Single-hop session
PEER: Multi-hop session
LDP: LDP session
LSP: Label switched path
TE: Traffic Engineering
AUTO: Automatically negotiated session
VXLAN: VXLAN session
(w): State in WTR
(*): State is invalid
Total UP/DOWN Session Number : 1/0
-----
Local    Remote    PeerIpAddr  State  Type      InterfaceName
-----
1        2         4.4.4.4     Up     S/LDP-LSP Vlanif10
-----
```

Run the **display bfd session all** command on PE2, and the command output that the **State** field is displayed as **Up**.

```
[~PE2] display bfd session all
S: Static session
D: Dynamic session
IP: IP session
IF: Single-hop session
PEER: Multi-hop session
LDP: LDP session
LSP: Label switched path
TE: Traffic Engineering
AUTO: Automatically negotiated session
VXLAN: VXLAN session
(w): State in WTR
(*): State is invalid
Total UP/DOWN Session Number : 1/0
-----
Local    Remote    PeerIpAddr  State  Type      InterfaceName
-----
2        1         1.1.1.1     Up     S/IP-PEER -
-----
```

----End

Configuration Files

- PE1 configuration file

```
#
sysname PE1
#
vlan batch 10 30
#
bfd
#
mpls lsr-id 1.1.1.1
#
mpls
#
mpls ldp
#
ipv4-family
#
interface Vlanif10
ip address 10.1.1.1 255.255.255.0
mpls
mpls ldp
#
interface Vlanif30
ip address 10.3.1.1 255.255.255.0
ospf cost 1000
mpls
mpls ldp
#
interface 10GE1/0/1
port link-type trunk
port trunk allow-pass vlan 10
#
interface 10GE2/0/2
port link-type trunk
port trunk allow-pass vlan 30
#
interface LoopBack1
ip address 1.1.1.1 255.255.255.255
#
ospf 1
area 0.0.0.0
network 1.1.1.1 0.0.0.0
network 10.1.1.0 0.0.0.255
network 10.3.1.0 0.0.0.255
#
bfd pe1tope2 bind ldp-lsp peer-ip 4.4.4.4 nexthop 10.1.1.2 interface Vlanif10
discriminator local 1
discriminator remote 2
#
return
```

- P1 configuration file

```
#
sysname P1
#
vlan batch 10 20
#
mpls lsr-id 2.2.2.2
#
mpls
#
mpls ldp
#
ipv4-family
#
interface Vlanif10
ip address 10.1.1.2 255.255.255.0
mpls
```

```
mpls ldp
#
interface Vlanif20
ip address 10.2.1.1 255.255.255.0
mpls
mpls ldp
#
interface 10GE1/0/1
port link-type trunk
port trunk allow-pass vlan 10
#
interface 10GE2/0/2
port link-type trunk
port trunk allow-pass vlan 20
#
interface LoopBack1
ip address 2.2.2.2 255.255.255.255
#
ospf 1
area 0.0.0.0
network 2.2.2.2 0.0.0.0
network 10.1.1.0 0.0.0.255
network 10.2.1.0 0.0.0.255
#
return
```

- P2 configuration file

```
#
sysname P2
#
vlan batch 30 40
#
mpls lsr-id 3.3.3.3
#
mpls
#
mpls ldp
#
ipv4-family
#
interface Vlanif30
ip address 10.3.1.2 255.255.255.0
mpls
mpls ldp
#
interface Vlanif40
ip address 10.4.1.1 255.255.255.0
mpls
mpls ldp
#
interface 10GE1/0/1
port link-type trunk
port trunk allow-pass vlan 30
#
interface 10GE2/0/2
port link-type trunk
port trunk allow-pass vlan 40
#
interface LoopBack1
ip address 3.3.3.3 255.255.255.255
#
ospf 1
area 0.0.0.0
network 3.3.3.3 0.0.0.0
network 10.3.1.0 0.0.0.255
network 10.4.1.0 0.0.0.255
#
return
```

- PE2 configuration file

```
#
sysname PE2
#
vlan batch 20 40
#
bfd
#
mpls lsr-id 4.4.4.4
#
mpls
#
mpls ldp
#
ipv4-family
#
interface Vlanif20
ip address 10.2.1.2 255.255.255.0
mpls
mpls ldp
#
interface Vlanif40
ip address 10.4.1.2 255.255.255.0
mpls
mpls ldp
#
interface 10GE1/0/1
port link-type trunk
port trunk allow-pass vlan 20
#
interface 10GE2/0/2
port link-type trunk
port trunk allow-pass vlan 40
#
interface LoopBack1
ip address 4.4.4.4 255.255.255.255
#
bfd pe2tope1 bind peer-ip 1.1.1.1
discriminator local 2
discriminator remote 1
#
ospf 1
area 0.0.0.0
network 4.4.4.4 0.0.0.0
network 10.2.1.0 0.0.0.255
network 10.4.1.0 0.0.0.255
#
return
```

2.15 Troubleshooting MPLS LDP

2.15.1 LDP Session Alternates Between Up and Down States

Fault Description

An LDP session alternates between Up and Down states when you add, change, or delete the LDP GR timer, LDP authentication, LDP Keepalive timer, or LDP transport address.

Procedure

- Step 1** Run the **display this** command in the LDP view to check whether LDP GR or LDP MTU is configured.

- If the following information is displayed:
mpls ldp
graceful-restart
LDP GR is configured.
- If information similar to the following is displayed:
mpls ldp
md5-password cipher 2.2.2.2 %%%#i@]DOR.^#F:v8OauMulclT]0\<CYVbliE:Mj04%##
or
mpls ldp
authentication key-chain peer 2.2.2.2 name kc1
LDP authentication is configured.

Step 2 Run the **display this** command in the interface view to check whether the LDP Keepalive timer or LDP transport address is configured.

- If information similar to the following is displayed:
mpls ldp
mpls ldp timer keepalive-hold 30
The LDP Keepalive timer is configured.
- If information similar to the following is displayed:
mpls ldp
mpls ldp transport-address interface
The LDP transport address is configured.

Step 3 After the preceding configurations are complete, wait for 10s and the LDP session becomes stable.

----End

2.15.2 LDP Session Is Down

Fault Description

An LDP session is Down after being established.

Procedure

Step 1 Check whether the interface where the LDP session is established is shut down.
Run the **display this** command in the interface view. If the following information is displayed:

```
shutdown
```

The interface is shut down.

If the interface is shut down, run the **undo shutdown** command to start the interface.

Step 2 Check whether the MPLS-related configurations are deleted.
Run the **display current-configuration** command to check whether MPLS-related configurations exist.

- If the output does not include the following information:
mpls
The MPLS configuration is deleted.
- If the output does not include the following information:
mpls ldp
The MPLS LDP configuration is deleted.

- If the output does not include the following information:
mpls ldp remote-peer
The remote LDP session is deleted.
- If MPLS-related configurations are deleted, run the corresponding commands to restore the configurations.

----End

2.15.3 LDP LSP Alternates Between Up and Down States

Fault Description

An LDP LSP alternates between Up and Down states after being established.

Procedure

- Check whether the LDP session flaps.
Run the **display mpls ldp session** command to check the displayed **Status** field. You are advised to run this command once every 1s. If the LDP session status switches between **Operational** and non-operational, the LDP session flap occurs.

If the LDP session flap occurs, rectify the fault by referring to [LDP Session Alternates Between Up and Down States](#).

----End

2.15.4 LDP LSP Is Down

Fault Description

An LDP LSP is Down after being established.

Procedure

- Step 1** Check whether the LDP session is correctly established.
- Run the **display mpls ldp session** command to check the displayed **Status** field. If LDP session status is **Operational**, the LDP session is established and in Up state. If LDP session status is not **Operational**, the LDP session is not established.
- If the LDP session is not established, rectify the fault by referring to [LDP Session Is Down](#).
- Step 2** Check whether the LSP establishment policy is configured.
- Run the **display this** command in the MPLS view. If information similar to the following is displayed:
lsp-trigger ip-prefix abc
Check whether the Down LSP is filtered out based on the IP prefix list **abc**.
 - Run the **display this** command in the MPLS-LDP view. If information similar to the following is displayed:
propagate mapping for ip-prefix abc
Check whether the Down LSP is filtered out based on the IP prefix list **abc**.
 - Run the **display ip ip-prefix** command in the system view. If information similar to the following is displayed:

```
index: 10      permit 10.1.1.1/32
index: 20      permit 10.2.2.2/32
```

The LSP can be established based only on routes 10.1.1.1/32 and 10.2.2.2/32.

- If the preceding IP prefix list is configured, add routing information of the Down LSP to the IP prefix list.

----End

3 MPLS QoS Configuration

On an MPLS network, MPLS QoS controls enterprise network traffic, and implements congestion avoidance and congestion management to reduce packet loss. In addition, MPLS QoS provides dedicated bandwidth for enterprise users or differentiated services (such as voice, video, and data services).

[3.1 Overview of MPLS QoS](#)

[3.2 Understanding MPLS QoS](#)

[3.3 Licensing Requirements and Limitations for MPLS QoS](#)

[3.4 Default Settings for MPLS QoS](#)

[3.5 Configuring MPLS QoS](#)

3.1 Overview of MPLS QoS

Definition

Multiprotocol Label Switching quality of service (MPLS QoS) is implemented using the Differentiated Services (DiffServ) model on an MPLS network. MPLS QoS provides differentiated services to meet diversified requirements.

Purpose

MPLS uses label-based forwarding to replace route-based forwarding and provides powerful and flexible functions to meet requirements of new applications. In addition, MPLS supports multiple network protocols including IPv4 and IPv6. MPLS has been widely used for building large-scale networks. On an MPLS network, IP QoS cannot be used to guarantee quality of services, so MPLS QoS is used.

Similar to the way IP QoS differentiates services based on priorities of IP packets, MPLS QoS differentiates data flows based on the EXP field and provides differentiated services for data flows. The use of MPLS QoS helps minimize delays and ensure low packet loss ratios for voice and video data streams.

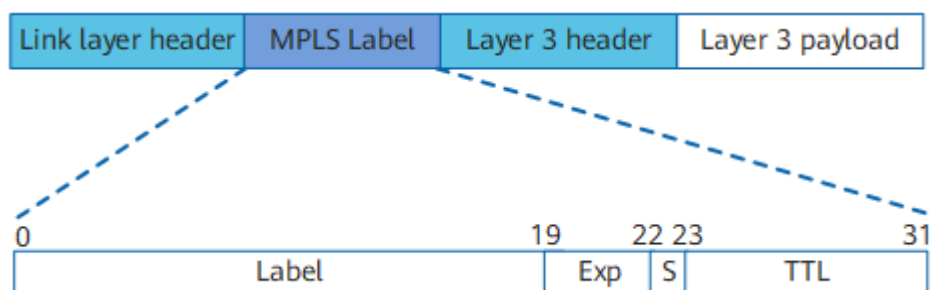
3.2 Understanding MPLS QoS

3.2.1 MPLS DiffServ

Implementation

In the DiffServ model, network edge nodes map a service to a service class based on QoS requirements of the service. A service class is identified by the differentiated service (DS) field (Type of Service (ToS) field) in IP packets or the PRI field (802.1p priority) in VLAN packets. Nodes on a backbone network apply preset policies to the service based on the DS or PRI field to ensure service quality. For details, see Priority Mapping Configuration (CE Switches Excluding CE6870EI) or Priority Mapping Configuration (CE6870EI) in *CloudEngine 9800, 8800, 6800, 5800 Series Switches Configuration Guide - QoS*. The service classification and label distribution mechanisms of DiffServ are similar to MPLS label distribution. MPLS DiffServ combines DS or PRI distribution with MPLS label distribution.

Figure 3-1 Fields in an MPLS packet



MPLS DiffServ maps the EXP field (as shown in [Figure 3-1](#)) to a per-hop behavior (PHB). LSRs forward MPLS packets based on EXP fields in the MPLS packets. MPLS DiffServ provides the following solutions for label switched path (LSP) setup:

- E-LSP: an LSP whose PHB is determined by the EXP field. E-LSP applies to a network with fewer than eight PHBs. A differentiated services code point (DSCP) or 802.1p priority is mapped to a specified EXP value that identifies a PHB. [Table 3-1](#) describes the mapping between PHBs and EXP values. Packets are forwarded based on labels, and the EXP field determines the packet scheduling algorithm and drop priority at each hop. An LSP transmits a maximum of eight PHB flows that are identified by the EXP field in the MPLS packet header. The EXP value can be configured by the Internet service provider (ISP) or mapped from the DSCP or 802.1p priority in a packet. In E-LSP, PHB information does not need to be transmitted by signaling protocols. Additionally, the label efficiency is high, and the label status is easy to maintain.

Table 3-1 Mapping between DiffServ PHBs and EXP values

PHB	EXP Value
BE	0
AF1	1
AF2	2
AF3	3
AF4	4
EF	5
CS6	6
CS7	7

- L-LSP: an LSP whose PHB is determined by both the label and EXP value. L-LSP applies to a network with any number of PHBs. During packet forwarding, the label of a packet determines the forwarding path and scheduling algorithm, whereas the EXP field determines the drop priority of the packet. Labels differentiate service flows, so service flows of a specified type are transmitted over the same LSP. This solution requires more labels and occupies a large number of system resources.

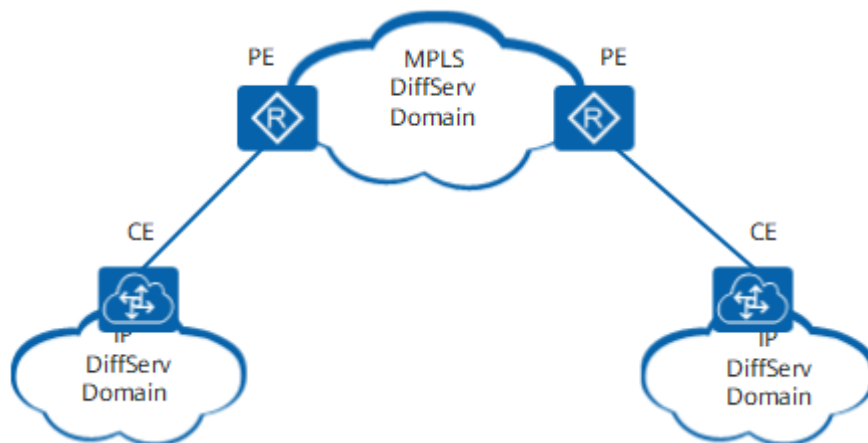
NOTE

The device supports only E-LSP.

DiffServ Domain

As shown in [Figure 3-2](#), DiffServ domains include MPLS DiffServ and IP DiffServ domains.

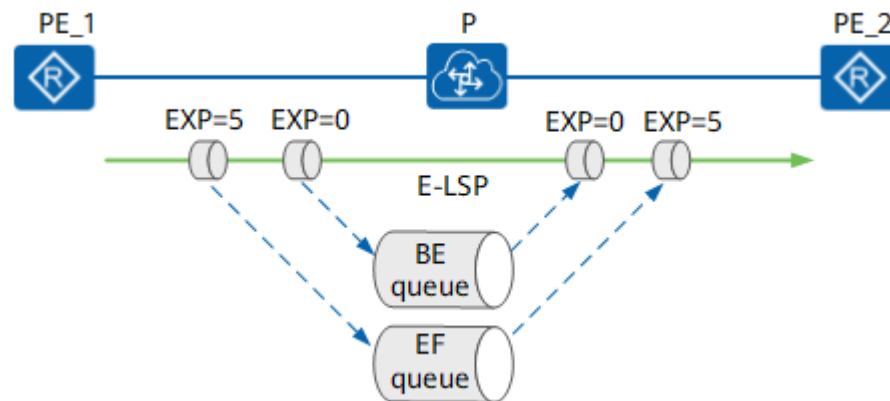
In the E-LSP solution, MPLS DiffServ manages and schedules the two DiffServ domains and implements bidirectional mapping between DSCP or 802.1p priorities and EXP priorities at the MPLS network edge.

Figure 3-2 DiffServ domain

As shown in [Figure 3-3](#), the MPLS DiffServ domain forwards MPLS packets based on EXP values and provides differentiated services.

When MPLS packets enter the P device, the P device classifies packets and maps EXP values in packets to class of service (CoS) values and drop priorities. After traffic classification, QoS implementations including traffic shaping, traffic policing, and congestion avoidance are the same as those on an IP network. When MPLS packets leave the P device, the packets carry EXP values.

Figure 3-3 E-LSP



3.2.2 MPLS DiffServ Tunnel Modes

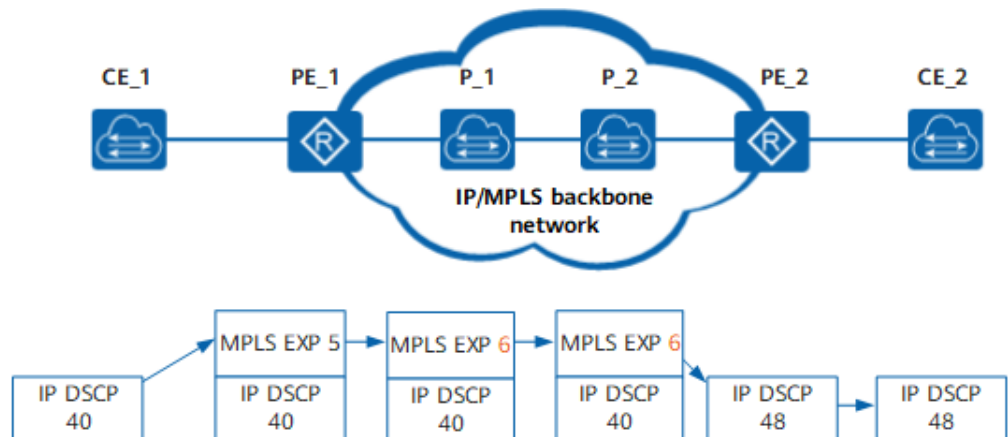
An MPLS DiffServ domain supports three tunnel modes:

- **Uniform:** Packet priorities are uniformly defined on the IP network and the MPLS network, so the priorities are globally valid. On the ingress node, each packet is assigned a label and its DSCP priority is mapped to an EXP value. A change in the EXP value on the MPLS network determines the PHB used when the packet leaves the MPLS network. The egress node maps the EXP value to the DSCP priority. [Figure 3-4](#) shows priority mapping in a uniform tunnel. P_1 changes the MPLS EXP value to 6. PE_2 changes the DSCP priority to 48. (PE_2 maps EXP value 6 to DSCP priority 48 by default.)

NOTE

The preceding description is based on a non-PHP scenario. In a PHP scenario, packets received by PE_2 do not carry MPLS EXP values; therefore, PE_2 does not implement MPLS QoS on the packets. To use MPLS QoS in a PHP scenario, it is recommended that you configure related configurations on the penultimate hop instead of the egress node, or configure the egress node to allocate an explicit null label to the penultimate hop.

Figure 3-4 Uniform mode

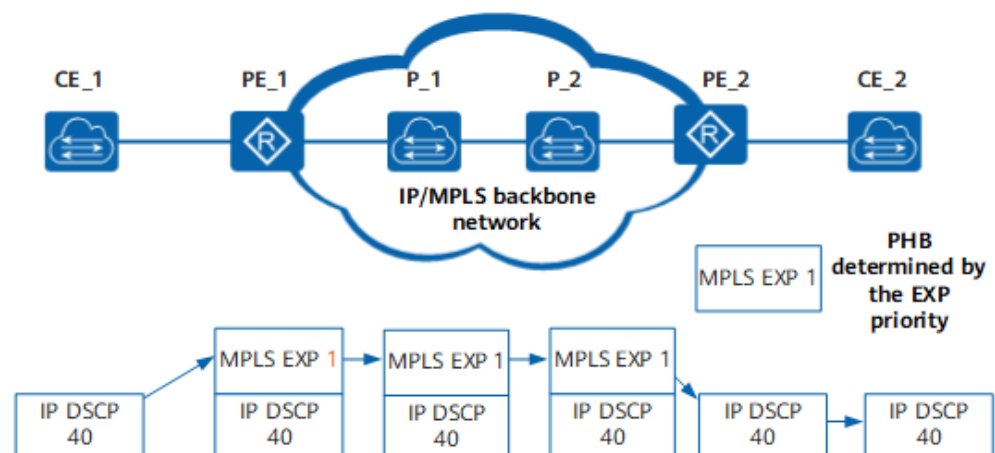


- Pipe: The EXP value can be manually configured, and the ingress node adds this EXP value to MPLS packets. Any change in the EXP value is valid only on the MPLS network. The egress node selects the PHB for MPLS packets according to the EXP value. When the packets leave the MPLS network, their DSCP priority is still valid. [Figure 3-5](#) shows priority mapping in a pipe tunnel. PE_1 changes the MPLS EXP values to 1. PE_2 retains the DSCP priority of packets and selects a PHB based on the MPLS EXP value.

NOTE

The preceding description is based on a non-PHP scenario. In a PHP scenario, packets received by PE_2 do not carry MPLS EXP values; therefore, PE_2 does not implement MPLS QoS on the packets. To use MPLS QoS in a PHP scenario, it is recommended that you configure related configurations on the penultimate hop instead of the egress node, or configure the egress node to allocate an explicit null label to the penultimate hop.

Figure 3-5 Pipe mode

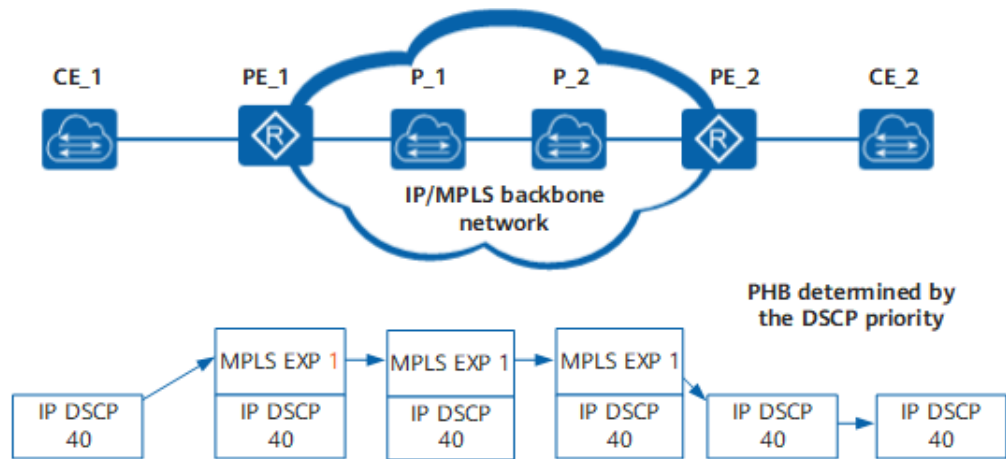


- Short pipe: The EXP value can be manually configured, and the ingress node adds this EXP value to MPLS packets. Any change in the EXP value is valid only on the MPLS network. The egress node selects the PHB for MPLS packets according to the DSCP priority. When the packets leave the MPLS network, their DSCP priority is still valid. [Figure 3-6](#) shows priority mapping in a short-pipe tunnel. PE_1 changes the MPLS EXP values to 1. PE_2 retains the DSCP priority of packets and selects a PHB on the DSCP priority.

NOTE

The preceding description is based on a non-PHP scenario.

Figure 3-6 Short pipe mode



3.3 Licensing Requirements and Limitations for MPLS QoS

Involved Network Elements

Other network elements are not required.

License Requirements

MPLS QoS is a basic function of the switch, and as such is controlled by the license for basic software functions. The license for basic software functions has been installed and activated on the device before delivery. You do not need to manually activate it.

Version Requirements

Table 3-2 Products and minimum version supporting this feature

Product Model	Minimum Version Required
CE6856HI	V200R002C50
CE6857EI	V200R005C10
CE6857E	V200R020C10
CE6857F	V200R021C00
CE6865EI	V200R005C00

Product Model	Minimum Version Required
CE6865E	V200R021C00
CE8850E-32CQ-EI	V200R021C10
CE6870-48S6CQ-EI	V200R001C00
CE6870-48S6CQ-EI-A	V200R022C00
CE6870-48T6CQ-EI	V200R002C50
CE8850-32CQ-EI	V200R002C50
CE8850-64CQ-EI	V200R005C00
CE8861EI	V200R005C10
CE8868EI	V200R005C10

Feature Limitations

When configuring MPLS QoS on the switch, pay attention to the following points:

- CE switches cannot function as PE devices on an MPLS network.
- When configuring MPLS QoS or MPLS L3VPN QoS on the CE6870EI switch, you can configure the MPLS DiffServ mode only on the ingress and egress nodes. The MPLS DiffServ mode does not take effect on the transit node.
- In V200R005C00, if the CE6856HI functions as the ingress node of the VPLS network, the **MPLS TTL** field has a fixed value of 255.
- When the CE6870EI switch is not configured with the MPLS PHP function and uses the inner label to forward packets after the outer label pops up (for example, the switch serves as an ASBR in the inter-AS VPN Option B or inter-AS VPN Option C scenario), the EXP in the label of the forwarded packets cannot be obtained from outbound priority mappings based on the PHB, and the EXP in the inner label is used.
- The user-configured EXP value is preferred as the EXP value in the outer label of L2VPN packets on the CE6870EI. If no EXP value is configured, the EXP value in the inner label is used as the EXP value in the outer label.
In addition, MPLS QoS is developed based on MPLS LDP. For other configuration notes about MPLS LDP, see [MPLS LDP Configuration Notes](#).
- When the CE6856HI functions as the penultimate hop in an MPLS PHP scenario and receives single-tagged packets, the **qos phb marking dscp enable** command does not take effect.
- For the CE6870EI, in the MPLS PHP scenario in uniform mode, the user-defined mapping between DSCP priorities and EXP priorities does not take effect for outgoing packets.

3.4 Default Settings for MPLS QoS

By default, the mappings in the DiffServ domain are as follows:

- **Table 3-3** lists the mappings from DSCP priorities to EXP priorities in MPLS packets on other models excluding the CE6870EI.
- **Table 3-4** lists the mappings from DSCP priorities to EXP priorities in MPLS packets on the CE6870EI.
- **Table 3-5** lists the mappings between EXP priorities in MPLS packets, DSCP priorities, and PHBs and colors.

Table 3-3 Mapping from PHBs and colors to EXP priorities in the outbound direction in the DiffServ domain (CE8861EI, CE8868EI, CE8850EI, CE6865EI, CE6865E, CE8850E-32CQ-EI, CE6857EI, CE6857E, CE6857F, and CE6856HI)

PHB	Color	EXP
BE	green	0
BE	yellow	0
BE	red	0
AF1	green	1
AF1	yellow	1
AF1	red	1
AF2	green	2
AF2	yellow	2
AF2	red	2
AF3	green	3
AF3	yellow	3
AF3	red	3
AF4	green	4
AF4	yellow	4
AF4	red	4
EF	green	5
EF	yellow	5
EF	red	5
CS6	green	6
CS6	yellow	6
CS6	red	6
CS7	green	7
CS7	yellow	7

PHB	Color	EXP
CS7	red	7

Table 3-4 Mapping from DSCP priorities to EXP priorities in the outbound direction in the DiffServ domain

DSCP Priority	EXP Priority	DSCP Priority	EXP Priority
0	0	32	4
1	0	33	4
2	0	34	4
3	0	35	4
4	0	36	4
5	0	37	4
6	0	38	4
7	0	39	4
8	1	40	5
9	1	41	5
10	1	42	5
11	1	43	5
12	1	44	5
13	1	45	5
14	1	46	5
15	1	47	5
16	2	48	6
17	2	49	6
18	2	50	6
19	2	51	6
20	2	52	6
21	2	53	6
22	2	54	6
23	2	55	6
24	3	56	7

DSCP Priority	EXP Priority	DSCP Priority	EXP Priority
25	3	57	7
26	3	58	7
27	3	59	7
28	3	60	7
29	3	61	7
30	3	62	7
31	3	63	7

Table 3-5 Mapping between EXP priorities in MPLS packets, DSCP priorities, and PHBs and colors in the inbound direction in the DiffServ domain

EXP Priority	CoS Value	Color	DSCP
0	BE	Green	0
1	AF1	Green	10
2	AF2	Green	18
3	AF3	Green	26
4	AF4	Green	34
5	EF	Green	46
6	CS6	Green	48
7	CS7	Green	56

3.5 Configuring MPLS QoS

3.5.1 Configuring MPLS Public QoS (CE6870EI)

Pre-configuration Tasks

After you configure MPLS Public QoS, differentiated services are provided for different services on the MPLS public network based on the service priorities.

Before configuring MPLS QoS, complete the following tasks:

- Configure IP addresses for interfaces to ensure that neighboring nodes are reachable at the network layer.
- Configure an IGP to advertise the network segments connecting to interfaces on each node and the routes of hosts with LSR IDs.

- Configure MPLS and MPLS LDP.

Context

To implement certain QoS functions on an MPLS network, devices on the network need to determine the packet priorities based on the MPLS labels. Therefore, you need to map the MPLS labels to EXP priorities.

MPLS QoS can differentiate different services on an MPLS network. If you want to retain priorities carried in original packets regardless of the change in the EXP field, set the DiffServ mode to short pipe. If you set the DiffServ mode to uniform, priorities carried in original packets may be changed. In uniform mode, the egress node determines the per-hop behavior (PHB) based on EXP priorities of packets. In short pipe mode, the egress node determines the PHB based on Differentiated Service Code Point (DSCP) priorities of packets.

NOTE

The MPLS DiffServ mode configured on a transit device does not take effect.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Create a DiffServ domain and configure priority mapping.

For details, see Priority Mapping Configuration in *CloudEngine 9800, 8800, 6800, 5800 Series Switches Configuration Guide - QoS*.

- Performing the following steps on the ingress:
 - a. Run **diffserv domain** *ds-domain-name*
A DiffServ domain is created and the DiffServ domain view is displayed.
The **default** domain defines default mappings from packet priorities to PHBs. You can modify the mapping defined in the **default** domain but cannot delete the **default** domain. In addition to the DiffServ domain **default**, a maximum of six user-defined DiffServ domains can be created.
 - b. Run **mpls-exp-outbound** *dscp-value* **map** *exp-value*
The DSCP priority is mapped to the EXP priority of MPLS packets in the outbound direction of an interface.
For details on the default mapping from the DSCP priority to the EXP priority, see **mpls-exp-outbound**.
The DSCP priority in this command is the priority carried in packets of a device's outbound interface.
 - c. Run **quit**
Return to the system view.
- Performing the following steps on the egress:
 - a. Run **diffserv domain** *ds-domain-name*
A DiffServ domain is created and the DiffServ domain view is displayed.
The **default** domain defines default mappings from packet priorities to PHBs. You can modify the mapping defined in the **default** domain but

cannot delete the **default** domain. In addition to the DiffServ domain **default**, a maximum of six user-defined DiffServ domains can be created.

- b. Run **mpls-exp-inbound** *exp-value* **phb** *service-class*

The EXP priority of MPLS packets is mapped to the PHB in the inbound direction of an interface.

For details on the default mapping from the EXP priority to the PHB, see **mpls-exp-inbound**.

The EXP priority of MPLS packets is mapped to green in the inbound direction of an interface.

- c. Run **quit**

Return to the system view.

 NOTE

To apply a user-defined DiffServ domain to a Layer 3 interface on the ingress node receiving the packets or a Layer 3 interface on the egress node receiving the packets, run the **trust upstream** *ds-domain-name* command in the Layer 3 interface view.

Step 3 Configure the MPLS DiffServ mode.

Run **mpls mode** { **short-pipe** [[**ttl** *ttl-value* **exp** *exp-value*] | **domain** *ds-name*] | **uniform** [**domain** *ds-name*] }

The MPLS DiffServ mode is set.

By default, the MPLS DiffServ mode is uniform.

 NOTE

When running the **mpls mode short-pipe ttl** *ttl-value* **exp** *exp-value* command, specify the same EXP value for the outer and inner MPLS labels. If the parameter **exp** is not specified, the inbound interface performs priority mapping for incoming packets based on the original priorities.

When the MPLS DiffServ mode is uniform, the egress node maps EXP priorities of packets to DSCP priorities. To map PHBs to DSCP priorities, configure the **qos phb marking dscp enable** command in the system view.

Step 4 Run **commit**

The configuration is committed.

----End

Checking the Configuration

Run the **display diffserv domain** [**brief** | *ds-domain-name*] command to view the DiffServ domain configuration.

3.5.2 Configuring MPLS Public QoS (CE Switches Excluding CE6870EI)

After you configure MPLS Public QoS, differentiated services are provided for different services on the MPLS public network based on the service priorities.

Pre-configuration Tasks

Before configuring MPLS QoS, complete the following tasks:

- Configure IP addresses for interfaces to ensure that neighboring nodes are reachable at the network layer.
- Configure an IGP to advertise the network segments connecting to interfaces on each node and the routes of hosts with LSR IDs.
- Configure MPLS and MPLS LDP.

Context

To implement certain QoS functions on an MPLS network, devices on the network need to determine the packet priorities based on the MPLS labels. Therefore, you need to map the MPLS labels to EXP priorities.

MPLS QoS can differentiate different services on an MPLS network. If you want to retain priorities carried in original packets regardless of the change in the EXP field, set the DiffServ mode to pipe or short pipe. If you set the DiffServ mode to uniform, priorities carried in original packets may be changed. In uniform and pipe modes, the egress node determines the per-hop behavior (PHB) based on EXP priorities of packets. In short pipe mode, the egress node determines the PHB based on Differentiated Service Code Point (DSCP) priorities of packets.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Create a DiffServ domain and configure priority mapping.

For details, see Priority Mapping Configuration in *CloudEngine 9800, 8800, 6800, 5800 Series Switches Configuration Guide - QoS*.

A DiffServ domain comprises the connected DiffServ nodes, which use the same service policy and implement the same PHBs.

When traffic enters a device, the device maps packet priorities to PHBs and colors, and performs congestion management based on PHBs and congestion avoidance based on colors. When traffic flows out of the device, the device maps PHBs and colors of packets to priorities. The downstream device provides QoS services based on packet priorities.

1. Run **diffserv domain** *ds-domain-name*

A DiffServ domain is created and the DiffServ domain view is displayed.

The **default** domain defines default mappings from packet priorities to PHBs. You can modify the mapping defined in the **default** domain but cannot delete the **default** domain. In addition to the DiffServ domain **default**, a maximum of seven user-defined DiffServ domains can be created.

2. Define a traffic policy on the device.

- Run **exp-inbound** *exp-value* **phb** *service-class* *color*

The EXP priority of MPLS packets is mapped to the PHB and color in the inbound direction of an interface.

For details on the default mapping from EXP priorities of MPLS packets to PHBs and colors in the inbound direction of an interface in a DiffServ domain, see **exp-inbound**.

- Run **exp-outbound service-class color map exp-value**

The PHB and color in original packets are mapped to the EXP priority of MPLS packets in the outbound direction of an interface.

For details on the default mapping from PHBs and colors in original packets to EXP priorities of MPLS packets in the outbound direction of an interface in a DiffServ domain, see **exp-outbound**.

3. Run **quit**

Return to the system view.

 NOTE

To apply a user-defined DiffServ domain to a Layer 2 interface on the ingress node receiving the packets or a Layer 2 interface on the egress node receiving the packets, run the **trust upstream ds-domain-name** command in the Layer 2 interface view.

Step 3 Configure the MPLS DiffServ mode.

To map priorities of incoming packets to PHBs and colors based on the mappings defined in a DiffServ domain, bind the DiffServ domain to the inbound interface of the packets. The system then maps priorities of packets to PHBs and colors based on the mappings in the DiffServ domain.

To map PHBs and colors of outgoing packets to priorities based on the mappings defined in a DiffServ domain, bind the DiffServ domain to the outbound interface of the packets. The system then maps PHBs and colors of outgoing packets to priorities based on the mappings in the DiffServ domain.

1. Run **mpls**

The MPLS view is displayed.

2. Run **mpls ldp qos-mode { pipe exp exp-value | short-pipe [exp exp-value] | uniform } [domain ds-name]**

The MPLS DiffServ mode is set.

By default, the MPLS DiffServ mode is uniform.

If no DiffServ domain name is specified, the default domain is used.

 NOTE

- When running the **mpls ldp qos-mode { pipe exp exp-value | short-pipe exp exp-value }** command, specify the EXP value for MPLS labels. When the DiffServ mode is short pipe and the **exp exp-value** parameter is not specified, the device maps the original priorities of packets on the inbound interface to the EXP values.
- If the DiffServ mode of the MPLS network is set to uniform, you need to run the **qos phb marking dscp enable** command on the penultimate hop in a PHP scenario and the egress node in a non-PHP scenario; otherwise, the EXP values of packets cannot be mapped to DSCP priorities.
- If the DiffServ mode of the MPLS network is set to pipe, DSCP priorities of original packets may be changed if the **qos phb marking dscp enable** command is configured on the outbound interface of the egress node. The changed value is determined by the DiffServ domain in use.
- In an implicit PHP scenario, the egress does not implement MPLS QoS on the packets. To use MPLS QoS in pipe or uniform mode in this scenario, it is recommended that you configure related configurations on the penultimate hop instead of the egress node, or configure the egress node to allocate an explicit null label to the penultimate hop.

Step 4 Run **commit**

The configuration is committed.

----End

Checking the Configuration

Run the **display diffserv domain [brief | ds-domain-name]** command to view the DiffServ domain configuration.

3.5.3 Configuring MPLS L3VPN QoS (CE6870EI)

Context

After you configure MPLS VPN QoS, differentiated services are provided for different services on the MPLS L3VPN network based on the service priorities.

The configuration of MPLS L3VPN QoS is similar to the configuration of QoS on the MPLS public network. For detailed configuration, see [3.5.1 Configuring MPLS Public QoS \(CE6870EI\)](#).

3.5.4 Configuring MPLS L3VPN QoS (CE Switches Excluding CE6870EI)

Context

After you configure MPLS VPN QoS, differentiated services are provided for different services on the MPLS L3VPN network based on the service priorities.

To implement certain QoS functions on an MPLS network, devices on the network need to determine the packet priorities based on the MPLS labels. Therefore, you need to map the MPLS labels to EXP priorities.

MPLS QoS can differentiate different L3VPN services on an MPLS network. If you want to retain priorities carried in original packets regardless of the change in the

EXP field, set the DiffServ mode of a VPN instance to pipe or short pipe. If you set the DiffServ mode to uniform, priorities carried in original packets may be changed. In uniform and pipe modes, the egress node determines the per-hop behavior (PHB) based on EXP priorities of packets. In short pipe mode, the egress node determines the PHB based on Differentiated Service Code Point (DSCP) priorities of packets.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Create a DiffServ domain and configure priority mapping.

For details, see Priority Mapping Configuration in *CloudEngine 9800, 8800, 6800, 5800 Series Switches Configuration Guide - QoS*.

A DiffServ domain comprises the connected DiffServ nodes, which use the same service policy and implement the same PHBs.

When traffic enters a device, the device maps packet priorities to PHBs and colors, and performs congestion management based on PHBs and congestion avoidance based on colors. When traffic flows out of the device, the device maps PHBs and colors of packets to priorities. The downstream device provides QoS services based on packet priorities.

1. Run **diffserv domain** *ds-domain-name*

A DiffServ domain is created and the DiffServ domain view is displayed.

The **default** domain defines default mappings from packet priorities to PHBs. You can modify the mapping defined in the **default** domain but cannot delete the **default** domain. In addition to the DiffServ domain **default**, a maximum of seven DiffServ domains can be created.

2. Define a traffic policy on the device.

– Run **exp-inbound** *exp-value phb service-class color*

The EXP priority of MPLS packets is mapped to the PHB and color in the inbound direction of an interface.

For details on the default mapping from EXP priorities of MPLS packets to PHBs and colors in the inbound direction of an interface in a DiffServ domain, see **exp-inbound**.

– Run **exp-outbound** *service-class color map exp-value*

The PHB and color in original packets are mapped to the EXP priority of MPLS packets in the outbound direction of an interface.

For details on the default mapping from PHBs and colors in original packets to EXP priorities of MPLS packets in the outbound direction of an interface in a DiffServ domain, see **exp-outbound**.

3. Run **quit**

Return to the system view.

Step 3 Configure the DiffServ mode for a VPN instance.

To map priorities of incoming packets to PHBs and colors based on the mappings defined in a DiffServ domain, bind the DiffServ domain to the inbound interface of

the packets. The system then maps priorities of packets to PHBs and colors based on the mappings in the DiffServ domain.

To map PHBs and colors of outgoing packets to priorities based on the mappings defined in a DiffServ domain, bind the DiffServ domain to the outbound interface of the packets. The system then maps PHBs and colors of outgoing packets to priorities based on the mappings in the DiffServ domain.

1. Run **ip vpn-instance** *vpn-instance-name*
The VPN instance view is displayed.
2. Run **qos mode { pipe exp exp-value | { short-pipe exp exp-value | uniform } domain ds-name }**

The DiffServ mode is set for the VPN instance.

By default, the DiffServ mode of a VPN instance is uniform, and the default domain is used.

NOTE

- When running the **qos mode { pipe exp exp-value | short-pipe exp exp-value domain ds-name }** command, specify the EXP value for MPLS private network labels.
- If the DiffServ mode of a VPN instance in the MPLS L3VPN scenario is set to uniform, you need to run the **qos phb marking dscp enable** command on the penultimate hop in a PHP scenario and the egress node in a non-PHP scenario; otherwise, the EXP values of packets cannot be mapped to DSCP priorities.
- If the DiffServ mode of a VPN instance in the MPLS L3VPN scenario is set to pipe, DSCP priorities of original packets may be changed if the **qos phb marking dscp enable** command is configured on the outbound interface of the egress node. The changed value is determined by the DiffServ domain in use.

Step 4 Run **commit**

The configuration is committed.

----End

Checking the Configuration

Run the **display diffserv domain [brief | ds-domain-name]** command to view the DiffServ domain configuration.

3.5.5 Configuring MPLS L2VPN QoS (CE6870EI)

Context

After you configure MPLS VPN QoS, differentiated services are provided for different services on the MPLS L2VPN network based on the service priorities.

On an MPLS L2VPN network, you can configure MPLS L2VPN QoS to differentiate services in a VPN. To retain the priorities of original packets, set the DiffServ mode to pipe. To change the priorities of original packets, set the DiffServ mode to uniform.

Procedure

Step 1 Run **system-view**

The system view is displayed.

Step 2 Create a DiffServ domain and configure priority mapping.

- Performing the following steps on the ingress:
 - a. Run **diffserv domain** *ds-domain-name*
A DiffServ domain is created and the DiffServ domain view is displayed.
The **default** domain defines default mappings from packet priorities to PHBs. You can modify the mapping defined in the **default** domain but cannot delete the **default** domain. In addition to the DiffServ domain **default**, a maximum of six user-defined DiffServ domains can be created.
 - b. Run **exp-outbound** *service-class color map exp-value*
The PHB and color in original packets are mapped to the EXP priority of VPLS packets in the outbound direction of an interface.
For details on the default mapping from PHBs and colors in original packets to EXP priorities of VPLS packets in the outbound direction of an interface in a DiffServ domain, see **exp-outbound**.
 - c. Run **quit**
Return to the system view.
- Performing the following steps on the egress:
 - a. Run **diffserv domain** *ds-domain-name*
A DiffServ domain is created and the DiffServ domain view is displayed.
The **default** domain defines default mappings from packet priorities to PHBs. You can modify the mapping defined in the **default** domain but cannot delete the **default** domain. In addition to the DiffServ domain **default**, a maximum of six user-defined DiffServ domains can be created.
 - b. Run **exp-inbound** *exp-value phb service-class color*
The EXP priority of VPLS packets is mapped to the PHB in the inbound direction of an interface.
For details on the default mapping from EXP priorities of VPLS packets to PHBs and colors in the inbound direction of an interface in a DiffServ domain, see **exp-inbound**.
 - c. Run **quit**
Return to the system view.

 NOTE

If a user-defined DiffServ domain is used:

1. When a VLANIF interface is connected to an L2VPN, you need to run the **trust upstream ds-domain-name** command in the VLAN view of the egress node to apply this domain.
2. When a BD is connected to an L2VPN, you need to run the **trust upstream ds-domain-name** command in the BD view or Layer 2 sub-interface view to apply this domain. The configuration in the Layer 2 sub-interface view takes precedence over that in the BD view.

Step 3 Run **mpls l2vpn mode { pipe | uniform } [domain ds-name]**

The DiffServ mode for MPLS L2VPN is specified and applied to a DiffServ domain.

When the DiffServ mode for MPLS L2VPN is Uniform, the 802.1p priorities of packets can only be mapped to EXP priorities.

 NOTE

On the CE6870EI:

1. When a VLANIF interface is connected to an L2VPN and the DiffServ mode is set to pipe, the egress node will map PHBs to 802.1p priorities for outgoing packets if the interface encapsulation type is changed to Ethernet using the **encapsulation ethernet** command. The mappings are pre-defined by the system and cannot be changed.
2. When a BD is connected to an L2VPN and the DiffServ mode is set to uniform, packets will retain their 802.1p priorities when leaving the network if a QinQ transparent transmission Layer 2 sub-interface, Dot1q transparent transmission Layer 2 sub-interface, or default Layer 2 sub-interface (allowing tagged packets) is connected to the L2VPN.
3. When a BD is connected to an L2VPN and the DiffServ mode is set to pipe, the egress node will map PHBs to 802.1p priorities for outgoing packets if a QinQ termination Layer 2 sub-interface or Dot1q termination Layer 2 sub-interface is connected to the L2VPN. The mappings are pre-defined by the system and cannot be changed.

Step 4 Run commit

The configuration is committed.

----End

Checking the Configuration

Run the **display diffserv domain [brief | *ds-domain-name*]** command to view the DiffServ domain configuration.

3.5.6 Configuring MPLS L2VPN QoS (Except CE6870EI)

Context

You can configure MPLS L2VPN QoS to distinguish services with different priorities on an MPLS L2VPN network and provide differentiated services accordingly.

On an MPLS L2VPN network, you can configure MPLS L2VPN QoS to distinguish different services from the same VPN.

If you want to retain priorities carried in original packets when the EXP priority changes, set the DiffServ mode to short pipe. If you set the DiffServ mode to uniform or pipe, priorities carried in original packets may be changed. In uniform and pipe modes, the egress node determines the per-hop behavior (PHB) based on EXP priorities of packets. In short pipe mode, the egress node determines the PHB based on 802.1p priorities of packets.

Procedure

Step 1 Run system-view

The system view is displayed.

Step 2 Run diffserv domain *ds-domain-name*

A DiffServ domain is created and the DiffServ domain view is displayed.

The **default** domain defines default mappings from packet priorities to PHBs. You can modify the mapping defined in the **default** domain but cannot delete the **default** domain. In addition to the **default** domain, a maximum of seven DiffServ domains can be created.

Step 3 Configure priority mapping.

Choose either of the following tasks:

- **DiffServ mode: uniform**

- a. Run **exp-outbound** *service-class color map exp-value*

The PHBs and colors are mapped to the EXP priorities for outgoing MPLS packets on the ingress node.

For details about the default mapping from PHBs and colors to the EXP priorities for outgoing packets in a DiffServ domain, see **exp-outbound**.

- b. Run **exp-inbound** *exp-value phb service-class color*

The EXP priorities are mapped to PHBs and colors for incoming MPLS packets on the egress node.

For details about the default mapping from the EXP priorities to PHBs and colors for incoming packets in a DiffServ domain, see **exp-inbound**.

- c. Run **quit**

Return to the system view.

- **DiffServ mode: pipe**

- a. Run **exp-inbound** *exp-value phb service-class color*

The EXP priorities are mapped to PHBs and colors for incoming MPLS packets on the egress node.

For details about the default mapping from the EXP priorities to PHBs and colors for incoming packets in a DiffServ domain, see **exp-inbound**.

If you want to retain the 802.1p priorities of original packets and make them effective after packets leave the MPLS network, run the **qos phb marking 8021p disable** command on the outbound interface of the egress node to disable mapping from PHBs to the 8021.p priorities.

- b. Run **quit**

Return to the system view.

- **DiffServ mode: short pipe**

- a. Run **exp-outbound** *service-class color map exp-value*

The PHBs and colors are mapped to the EXP priorities for outgoing MPLS packets on the ingress node.

For details about the default mapping from PHBs and colors to the EXP priorities for outgoing packets in a DiffServ domain, see **exp-outbound**.

If *exp-value* is specified, skip this step.

In the inbound direction of the interface on the egress node, the system sends packets to the queue whose internal priority is 0 and sets the color to green by default.

- b. Run **quit**

Return to the system view.

Step 4 Configure an MPLS DiffServ mode.

1. Run **vsi** *vsi-name* **bd-mode**

The VSI view is displayed.

2. Run **diffserv-mode** { **pipe exp** *exp-value* | **short-pipe** [**exp** *exp-value*] | **uniform** } [**domain** *ds-name*]

An MPLS DiffServ mode is configured.

- If *exp-value* is specified, the EXP priority in an MPLS private network label will be changed to this value.
- When you set the DiffServ mode to short pipe, the ingress node performs priority mapping based on the 802.1p priorities of original packets if **exp** *exp-value* is not specified for packets to be transmitted over the LSP. If the packets do not carry VLAN tags, the 802.1p priority is fixed as 0.

Step 5 Run **commit**

The configuration is committed.

----End

Checking the Configuration

Run the **display diffserv domain** [**brief** | *ds-domain-name*] command to view the DiffServ domain configuration.