Implementing MPLS OAM

MPLS Operations, Administration, and Maintenance (OAM) helps service providers to monitor label-switched paths (LSPs) and quickly isolate MPLS forwarding problems to assist with fault detection and troubleshooting in an MPLS network. This module describes MPLS LSP Ping and Traceroute features which can be used for failure detection and troubleshooting of MPLS networks.

**MPLS LSP Ping**

The MPLS LSP Ping feature is used to check the connectivity between Ingress LSR and egress LSRs along an LSP. MPLS LSP ping uses MPLS echo request and reply messages, similar to Internet Control Message Protocol (ICMP) echo request and reply messages, to validate an LSP. While ICMP echo request and reply messages validate IP networks, MPLS echo and reply messages validate MPLS networks. The MPLS echo request packet is sent to a target router through the use of the appropriate label stack associated with the LSP to be validated. Use of the label stack causes the packet to be forwarded over the LSP itself. The destination IP address of the MPLS echo request packet is different from the address used to select the label stack. The destination IP address is defined as a 127.x.y.z/8 address and it prevents the IP packet from being IP switched to its destination, if the LSP is broken.

An MPLS echo reply is sent in response to an MPLS echo request. The reply is sent as an IP packet and it is forwarded using IP, MPLS, or a combination of both types of switching. The source address of the MPLS echo reply packet is an address obtained from the router generating the echo reply. The destination address is the source address of the router that originated the MPLS echo request packet. The MPLS echo reply destination port is set to the echo request source port.

The following figure shows MPLS LSP ping echo request and echo reply paths.
By default, the `ping mpls ipv4` command tries to determine the Forwarding Equivalence Class (FEC) being used automatically. However, this is only applicable at head-end and works only if the FEC at the destination is same as the source. If the source and destination FEC types are not the same, the `ping mpls ipv4` command may fail to identify the targeted FEC type. You can overcome this limitation by specifying the FEC type in MPLS LSP ping using the `fec-type` command option. If the user is not sure about the FEC type at the transit or the destination, or it may change through network, use of the `generic` FEC type command option is recommended. Generic FEC is not coupled to a particular control plane and allows path verification when the advertising protocol is unknown, or may change during the path of the echo request. If you are aware of the destination FEC type, specify the target FEC as BGP or LDP.

Configuration Examples

This example shows how to use MPLS LSP ping to test the connectivity of an IPv4 LDP LSP. The destination is specified as a Label Distribution Protocol (LDP) IPv4 address.

```
RP/0/RP0/CPU0:router# ping mpls ipv4 10.1.1.2/32 verbose

Sun Nov 15 11:27:43.070 UTC
Sending 5, 100-byte MPLS Echos to 10.1.1.2/32, timeout is 2 seconds, send interval is 0 msec:

Type escape sequence to abort.

! size 100, reply addr 10.1.0.2, return code 3
! size 100, reply addr 10.1.0.2, return code 3
! size 100, reply addr 10.1.0.2, return code 3
! size 100, reply addr 10.1.0.2, return code 3
! size 100, reply addr 10.1.0.2, return code 3
Success rate is 100 percent (5/5), round-trip min/avg/max = 2/2/4 ms
```

In this example, the destination is specified as a Label Distribution Protocol (LDP) IPv4 prefix and Forwarding Equivalence Class (FEC) type is specified as generic.

```
RP/0/RP0/CPU0:router# ping mpls ipv4 10.1.1.2/32 fec-type generic

Wed Nov 25 03:36:33.143 UTC
Sending 5, 100-byte MPLS Echos to 10.1.1.2/32, timeout is 2 seconds, send interval is 0 msec:

! size 100, reply addr 10.1.0.2, return code 3
! size 100, reply addr 10.1.0.2, return code 3
! size 100, reply addr 10.1.0.2, return code 3
! size 100, reply addr 10.1.0.2, return code 3
! size 100, reply addr 10.1.0.2, return code 3
Success rate is 100 percent (5/5), round-trip min/avg/max = 2/2/4 ms
```

Type escape sequence to abort.

!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 2/2/3 ms

In this example, the destination is specified as a Label Distribution Protocol (LDP) IPv4 prefix and the FEC type is specified as BGP.

RP/0/RP0/CPU0:router# ping mpls ipv4 10.1.1.2/32 fec-type bgp

Wed Nov 25 03:38:33.143 UTC
Sending 5, 100-byte MPLS Echos to 10.1.1.2/32,
timeout is 2 seconds, send interval is 0 msec:


Type escape sequence to abort.

!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 2/2/3 ms

**MPLS LSP Traceroute**

The MPLS LSP Traceroute feature is used to isolate the failure point of an LSP. It is used for hop-by-hop fault localization and path tracing. The MPLS LSP Traceroute feature relies on the expiration of the Time to Live (TTL) value of the packet that carries the echo request. When the MPLS echo request message hits a transit node, it checks the TTL value and if it is expired, the packet is passed to the control plane, else the message is forwarded. If the echo message is passed to the control plane, a reply message is generated based on the contents of the request message.

The following figure shows an MPLS LSP traceroute example with an LSP from LSR1 to LSR4.
By default, the `traceroute mpls ipv4` command tries to determine the Forwarding Equivalence Class (FEC) being used automatically. However, this is only applicable at head-end and works only if the FEC at the destination is same as the source. If the source and destination FEC types are not the same, the `traceroute mpls ipv4` command may fail to identify the targeted FEC type. You can overcome this limitation by specifying the FEC type in MPLS LSP traceroute using the `fec-type` command option. If the user is not sure about the FEC type at the transit or the destination, or it may change through network, use of the `generic` FEC type command option is recommended. Generic FEC is not coupled to a particular control plane and allows path verification when the advertising protocol is unknown, or may change during the path of the echo request. If you are aware of the destination FEC type, specify the target FEC as BGP or LDP.

**Configuration Examples**

This example shows how to use the `traceroute` command to trace to a destination.

```
RP/0/RP0/CPU0:router# traceroute mpls ipv4 10.1.1.2/32 destination 127.0.0.3 127.0.0.6 2
```

Sat Jan 27 03:50:23.746 UTC

Tracing MPLS Label Switched Path to 10.1.1.2/32, timeout is 2 seconds


Type escape sequence to abort.

Destination address 127.0.0.3
0 10.2.1.2 MRU 1500 [Labels: 24000 Exp: 0] 
L 1 10.2.1.1 MRU 1500 [Labels: implicit-null Exp: 0] 8 ms 
! 2 10.1.0.2 3 ms 

Destination address 127.0.0.5
0 10.2.1.2 MRU 1500 [Labels: 24000 Exp: 0] 
L 1 10.2.1.1 MRU 1500 [Labels: implicit-null Exp: 0] 5 ms 
! 2 10.1.0.2 2 ms 

This example shows how to use the `traceroute` command and how to specify the maximum number of hops for the traceroute to traverse by specifying the `ttl` value.
Overview of P2MP TE Network

A Point to Multipoint (P2MP) TE network contains the following elements:

- **Headend Router**

  The headend router, also called the source or ingress router, is responsible for initiating the signaling messages that set up the P2MP TE LSP. The headend router can also be a branch point, which means the router performs packet replication and the sub-LSPs split into different directions.
• **Midpoint Router**
  The midpoint router is where the sub-LSP signaling is processed. The midpoint router can be a branch point.

• **Tailend Router**
  The tailend router, also called the destination, egress, or leaf-node router, is where sub-LSP signaling ends. The router which is one of potentially many destinations of the P2MP TE LSP.

• **Bud Router**
  A bud router is a midpoint and tailend router at the same time. An LSR that is an egress LSR, but also has one or more directly connected downstream LSRs.

• **Branch Router**
  A branch router is either a midpoint or tailend router at any given time.

• **Transit Router**
  A transit router is an LSR that is not an egress router, but also has one or more directly connected downstream routers.

• A P2MP tunnel consists of one or more sub-LSPs. All sub-LSPs belonging to the same P2MP tunnel employ the same constraints, protection policies, and so on, which are configured at the headend router.

**Figure 3: Elements of P2MP TE Network** illustrates the elements of P2MP TE network.

Figure 3: Elements of P2MP TE Network

P2MP TE tunnels build on the features that exist in basic point-to-point TE tunnels. The P2MP TE tunnels have the following characteristics:

• There is one source (headend) but more than one destination (tailend).

• They are unidirectional.
• They are explicitly routed.
• Multiple sub-LSPs connect the headend router to various tailend routers.

**P2MP Ping**

The P2MP ping feature is used to check the connectivity between Ingress LSR and egress LSR, along a P2MP LSP. The Ingress LSR sends the P2MP echo request message along the specified P2MP LSP. All egress LSRs which receive the P2MP echo request message from the ingress LSR must send a P2MP echo reply message to the ingress LSR, according to the reply mode specified in the P2MP echo request message.

**P2MP Traceroute**

The P2MP traceroute feature is used to isolate the failure point of a P2MP LSP. Traceroute can be applied to all nodes in the P2MP tree. However, you can select a specific traceroute target through the P2MP Responder Identifier TLV. An entry in this TLV represents an responder-id or a transit node. This is only the case for P2MP TE LSPs.

---

**Note**

Only P2MP TE LSP IPv4 is supported. If the Responder Identifier TLV is missing, the echo request requests information from all responder-ids.

---

**MPLS OAM Support for BGP 3107**

The MPLS OAM Support for BGP 3107 feature provides support for ping, traceroute and treetrace (traceroute multipath) operations for LSPs signaled via BGP for the IPv4 unicast prefix FECs in the default VRF, according to the RFC 3107 - Carrying Label Information in BGP-4. This feature adds support for MPLS OAM operations in the seamless MPLS architecture deployments, i.e., combinations of BGP and LDP signaled LSPs.

**Configuration Examples: P2MP Ping and P2MP Traceroute**

This example shows an extract of the P2MP ping command.

```
RP/0/RP0/CPU0:router# ping mpls traffic-eng tunnel-mte 10
Sending 1, 100-byte MPLS Echos to tunnel-mte10,
    timeout is 2.2 seconds, send interval is 0 msec, jitter value is 200 msec:

Codes: '!' - success, 'Q' - request not sent, '.' - timeout,
    'L' - labeled output interface, 'B' - unlabeled output interface,
    'D' - DS Map mismatch, 'F' - no FEC mapping, 'f' - FEC mismatch,
    'M' - malformed request, 'm' - unsupported tlvs, 'N' - no rx label,
    'P' - no rx intf label prot, 'p' - premature termination of LSP,
    'R' - transit router, 'I' - unknown upstream index,
    'X' - unknown return code, 'x' - return code 0, 'd' - DDMAP

Type escape sequence to abort.

Request #1
  ! reply addr 192.168.222.2
  ! reply addr 192.168.140.2
  ! reply addr 192.168.170.1
```
This example shows an extract of the P2MP ping command with the jitter option.

RP/0/RP0/CPU0:router# ping mpls traffic-eng tunnel-mte 10 jitter 300

Sending 1, 100-byte MPLS Echos to tunnel-mte10,
  timeout is 2.3 seconds, send interval is 0 msec, jitter value is 300 msec:

Codes: '!' - success, 'Q' - request not sent, '.' - timeout,
  'L' - labeled output interface, 'B' - unlabeled output interface,
  'D' - DS Map mismatch, 'F' - no FEC mapping, 'f' - FEC mismatch,
  'M' - malformed request, 'm' - unsupported tlvs, 'N' - no rx label,
  'R' - transit router, 'I' - unknown upstream index,
  'X' - unknown return code, 'x' - return code 0, 'd' - DDMAP

Type escape sequence to abort.

Request #1
  ! reply addr 192.168.222.2
  ! reply addr 192.168.140.2
  ! reply addr 192.168.170.1

Success rate is 100 percent (3 received replies/3 expected replies),
  round-trip min/avg/max = 148/191/256 ms

This example shows an extract of the P2MP ping command with the ddmap option.

RP/0/RP0/CPU0:router# ping mpls traffic-eng tunnel-mte 10 ddmap

Sending 1, 100-byte MPLS Echos to tunnel-mte10,
  timeout is 2.2 seconds, send interval is 0 msec, jitter value is 200 msec:

Codes: '!' - success, 'Q' - request not sent, '.' - timeout,
  'L' - labeled output interface, 'B' - unlabeled output interface,
  'D' - DS Map mismatch, 'F' - no FEC mapping, 'f' - FEC mismatch,
  'M' - malformed request, 'm' - unsupported tlvs, 'N' - no rx label,
  'R' - transit router, 'I' - unknown upstream index,
  'X' - unknown return code, 'x' - return code 0, 'd' - DDMAP

Type escape sequence to abort.

Request #1
  ! reply addr 192.168.222.2
  ! reply addr 192.168.140.2
  ! reply addr 192.168.170.1

Success rate is 100 percent (3 received replies/3 expected replies),
  round-trip min/avg/max = 105/178/237 ms

RP/0/RP0/CPU0:router# show mpls traffic-eng tunnels p2mp 10
Mon Apr 12 12:14:35:55.075 EST
Signalling Summary:
  LSP Tunnels Process: running
RSVP Process: running
Forwarding: enabled
Periodic reoptimization: every 3600 seconds, next in 654 seconds
Periodic FRR Promotion: every 300 seconds, next in 70 seconds
Auto-bw enabled tunnels: 0 (disabled)

Name: tunnel-mte10
Status:
  Admin: up Oper: up (Up for 12w4d)

Config Parameters:
  Bandwidth: 0 kbps (CT0) Priority: 7 7 Affinity: 0x0/0xffffffff
  Metric Type: TE (default)
  Fast Reroute: Not Enabled, Protection Desired: None
  Record Route: Not Enabled

Destination summary: (3 up, 0 down, 0 disabled) Affinity: 0x0/0xffffffff
Auto-bw: disabled
Destination: 10.1.0.1
  State: Up for 12w4d
  Path options:
    path-option 1 dynamic [active]

Destination: 10.2.0.1
  State: Up for 12w4d
  Path options:
    path-option 1 dynamic [active]

Destination: 10.3.0.1
  State: Up for 12w4d
  Path options:
    path-option 1 dynamic [active]

History:
  Reopt. LSP:
    Last Failure:
      LSP not signalled, identical to the [CURRENT] LSP
      Date/Time: Thu Jan 14 02:49:22 EST 2010 [12w4d ago]

Current LSP:
  lsp-id: 10002 p2mp-id: 10 tun-id: 10 src: 10.0.0.1 extid: 10.0.0.1
  LSP up for: 12w4d
  Reroute Pending: No
  Inuse Bandwidth: 0 kbps (CT0)
  Number of S2Ls: 3 connected, 0 signaling proceeding, 0 down

S2L Sub LSP: Destination 10.1.0.1 Signaling Status: connected
S2L up for: 12w4d
  Sub Group ID: 1 Sub Group Originator ID: 10.1.0.1
  Path option path-option 1 dynamic (path weight 1)
  Path info (OSPF 1 area 0)
    192.168.222.2
    10.1.0.1

S2L Sub LSP: Destination 10.2.0.1 Signaling Status: connected
S2L up for: 12w4d
  Sub Group ID: 2 Sub Group Originator ID: 10.0.0.1
  Path option path-option 1 dynamic (path weight 2)
  Path info (OSPF 1 area 0)
    192.168.222.2
    192.168.140.3
    192.168.140.2
    10.2.0.1

S2L Sub LSP: Destination 10.3.0.1 Signaling Status: connected
S2L up for: 12w4d
Sub Group ID: 3  Sub Group Originator ID: 10.0.0.1
Path option path-option 1 dynamic  (path weight 2)
Path info (OSPF 1 area 0)
  192.168.222.2
  192.168.170.3
  192.168.170.1
  10.3.0.1

Reoptimized LSP (Install Timer Remaining 0 Seconds):
  None
Cleaned LSP (Cleanup Timer Remaining 0 Seconds):
  None
Displayed 1 (of 16) heads, 0 (of 0) midpoints, 0 (of 0) tails
Displayed 1 up, 0 down, 0 recovering, 0 recovered heads

RP/0/RP0/CPU0:router# ping mpls traffic-eng tunnel-mte 10 lsp id 10002
Mon Apr 12 12:14:04.532 EST
Sending 1, 100-byte MPLS Echos to tunnel-mte10,
  timeout is 2.2 seconds, send interval is 0 msec, jitter value is 200 msec:

Codes: '!' - success, 'Q' - request not sent, '.' - timeout,
  'L' - labeled output interface, 'B' - unlabeled output interface,
  'D' - DS Map mismatch, 'F' - no FEC mapping, 'E' - FEC mismatch,
  'M' - malformed request, 'm' - unsupported tlvs, 'N' - no rx label,
  'P' - no rx intf label prot, 'p' - premature termination of LSP,
  'R' - transit router, 'I' - unknown upstream index,
  'X' - unknown return code, 'x' - return code 0, 'd' - DDMAP

Type escape sequence to abort.

Request #1
  ! reply addr 192.168.222.2
  ! reply addr 192.168.170.1
  ! reply addr 192.168.140.2

Success rate is 100 percent (3 received replies/3 expected replies),
  round-trip min/avg/max = 128/153/167 ms

This example shows an extract of the P2MP ping command with the responder-id.

RP/0/RP0/CPU0:router# ping mpls traffic-eng tunnel-mte 10 responder-id 10.3.0.1
Mon Apr 12 12:15:34.205 EST
Sending 1, 100-byte MPLS Echos to tunnel-mte10,
  timeout is 2.2 seconds, send interval is 0 msec, jitter value is 200 msec:

Codes: '!' - success, 'Q' - request not sent, '.' - timeout,
  'L' - labeled output interface, 'B' - unlabeled output interface,
  'D' - DS Map mismatch, 'F' - no FEC mapping, 'E' - FEC mismatch,
  'M' - malformed request, 'm' - unsupported tlvs, 'N' - no rx label,
  'P' - no rx intf label prot, 'p' - premature termination of LSP,
  'R' - transit router, 'I' - unknown upstream index,
  'X' - unknown return code, 'x' - return code 0, 'd' - DDMAP

Type escape sequence to abort.

Request #1
  ! reply addr 192.168.170.1

Success rate is 100 percent (1 received reply/1 expected reply),
  round-trip min/avg/max = 179/179/179 ms
This example shows an extract of the P2MP traceroute command with the ttl option.

```
RP/0/RP0/CPU0:router# traceroute mpls traffic-eng tunnel-mte 10  ttl 4
Mon Apr 12 12:16:50.095 EST
Tracing MPLS MTE Label Switched Path on tunnel-mte10, timeout is 2.2 seconds
Codes: '!' - success, 'Q' - request not sent, '.' - timeout,
       'L' - labeled output interface, 'B' - unlabeled output interface,
       'D' - DS Map mismatch, 'F' - no FEC mapping, 'f' - FEC mismatch,
       'M' - malformed request, 'm' - unsupported tlvs, 'N' - no rx label,
       'P' - no rx intf label prot, 'p' - premature termination of LSP,
       'R' - transit router, 'I' - unknown upstream index,
       'X' - unknown return code, 'x' - return code 0, 'd' - DDMAP
Type escape sequence to abort.

! 1 192.168.222.2 186 ms [Estimated Role: Bud]
   [L] DDMAP 0: 192.168.140.2 192.168.140.2 MRU 1500 [Labels: 16001 Exp: 0]

! 2 192.168.222.2 115 ms [Estimated Role: Bud]
   [L] DDMAP 0: 192.168.140.2 192.168.140.2 MRU 1500 [Labels: 16001 Exp: 0]

! 2 192.168.140.2 213 ms [Estimated Role: Egress]
! 2 192.168.170.1 254 ms [Estimated Role: Egress]

! 3 192.168.222.2 108 ms [Estimated Role: Bud]
   [L] DDMAP 0: 192.168.140.2 192.168.140.2 MRU 1500 [Labels: 16001 Exp: 0]

! 3 192.168.170.1 164 ms [Estimated Role: Egress]
! 3 192.168.140.2 199 ms [Estimated Role: Egress]

! 4 192.168.170.1 198 ms [Estimated Role: Egress]
! 4 192.168.222.2 206 ms [Estimated Role: Bud]
   [L] DDMAP 0: 192.168.140.2 192.168.140.2 MRU 1500 [Labels: 16001 Exp: 0]
```

This example shows an extract of the P2MP traceroute command with the responder-id option.

```
RP/0/RP0/CPU0:router# traceroute mpls traffic-eng tunnel-mte 10 responder-id 10.3.0.1
Mon Apr 12 12:18:01.994 EST
Tracing MPLS MTE Label Switched Path on tunnel-mte10, timeout is 2.2 seconds
Codes: '!' - success, 'Q' - request not sent, '.' - timeout,
       'L' - labeled output interface, 'B' - unlabeled output interface,
       'D' - DS Map mismatch, 'F' - no FEC mapping, 'f' - FEC mismatch,
       'M' - malformed request, 'm' - unsupported tlvs, 'N' - no rx label,
       'P' - no rx intf label prot, 'p' - premature termination of LSP,
       'R' - transit router, 'I' - unknown upstream index,
       'X' - unknown return code, 'x' - return code 0, 'd' - DDMAP
Type escape sequence to abort.

d 1 192.168.222.2 113 ms [Estimated Role: Branch]
   [L] DDMAP 0: 192.168.140.2 192.168.140.2 MRU 1500 [Labels: 16001 Exp: 0]

d 2 192.168.222.2 118 ms [Estimated Role: Branch]
   [L] DDMAP 0: 192.168.140.2 192.168.140.2 MRU 1500 [Labels: 16001 Exp: 0]
```

This example shows an extract of the P2MP traceroute command with the responder-id option.
This example shows an extract of the P2MP traceroute command with the jitter option.

```
RP/0/RP0/CPU0:router# traceroute mpls traffic-eng tunnel-mte 10 responder-id 10.3.0.1 ttl 4 jitter 500
Mon Apr 12 12:19:00.292 EST
Tracing MPLS MTE Label Switched Path on tunnel-mte10, timeout is 2.5 seconds
Codes: '!' - success, 'Q' - request not sent, '.' - timeout,
      'L' - labeled output interface, 'B' - unlabeled output interface,
      'D' - DS Map mismatch, 'F' - no FEC mapping, 'f' - FEC mismatch,
      'M' - malformed request, 'm' - unsupported tlvs, 'N' - no rx label,
      'P' - no rx intf label prot, 'p' - premature termination of LSP,
      'R' - transit router, 'I' - unknown upstream index,
      'X' - unknown return code, 'x' - return code 0, 'd' - DDMAP
Type escape sequence to abort.
```

d  1 192.168.222.2 238 ms [Estimated Role: Branch]
  [L] DDMAP 0: 192.168.140.2 192.168.140.2 MRU 1500 [Labels: 16001 Exp: 0]
  ! 2 192.168.170.1 244 ms [Estimated Role: Egress]

d  3 192.168.222.2 141 ms [Estimated Role: Branch]
  [L] DDMAP 0: 192.168.140.2 192.168.140.2 MRU 1500 [Labels: 16001 Exp: 0]
  ! 3 192.168.170.1 204 ms [Estimated Role: Egress]

d  2 192.168.170.1 188 ms [Estimated Role: Branch]
  [L] DDMAP 0: 192.168.140.2 192.168.140.2 MRU 1500 [Labels: 16001 Exp: 0]
  ! 3 192.168.170.1 290 ms [Estimated Role: Egress]

d  4 192.168.170.1 110 ms [Estimated Role: Branch]
  [L] DDMAP 0: 192.168.140.2 192.168.140.2 MRU 1500 [Labels: 16001 Exp: 0]
  ! 4 192.168.170.1 174 ms [Estimated Role: Egress]

d  3 192.168.170.1 204 ms [Estimated Role: Egress]
  [L] DDMAP 0: 192.168.140.2 192.168.140.2 MRU 1500 [Labels: 16001 Exp: 0]
  ! 4 192.168.170.1 327 ms [Estimated Role: Egress]