Next Generation MULTICAST In-band Signaling (VRF MLDP: Profile 6)

Contents

Introduction

This document describes the In-band Signaling VRF MLDP which is Profile 6 for Next Generation Multicast over VPN (mVPN). It uses an example and the implementation in Cisco IOS in order to illustrate the behaviour.

Background Information

Multicast Label Distribution Protocol (MLDP) in-band signalling to enable the MLDP core to create (S,G) or (*,G) state without using out-of-band signalling such as Border Gateway Protocol (BGP) or Protocol Independent Multicast (PIM).

MLDP-supported multicast VPN (MVPN) allows VPN multicast streams to be aggregated over a VPN-specific tree.

No customer state is created in the MLDP core, there is the only state for default and data multicast distribution trees (MDTs).

In certain scenarios, the state created for VPN streams is limited and does not appear to be a risk or limiting factor. In these scenarios, MLDP can build in-band MDTs that are transit Label Switched Paths (LSPs).

Trees used in a VPN space are MDTs. Trees used in the global table are transit point-to-multipoint (P2MP) or multipoint-to-multipoint (MP2MP) LSPs.

In both cases, a single multicast stream (VPN or not) is associated with a single LSP in the MPLS core. The stream information is encoded in the Forwarding Equivalence Class (FEC) of the LSP. This is in-band signalling.

LSM provides benefits when compared to GRE core tunnels that are currently used to transport customer traffic in the core and It leverages the MPLS infrastructure for transporting IP multicast packets, providing a common data plane for unicast and multicast.

MLDP Signaling

MLDP Signalling provides two functions:

- To discover the FEC (and its associated Opaque Value) for a MP LSP
- To assign a multicast flow to a MP LSP

In-Band Signaling

- Opaque Value is used to map an MP LSP to an IP multicast flow.
- Contents of the opaque value is derived from the multicast flow.

Overlay or Out-Of-Band Signaling

- Uses an overlay protocol to map an MP LSP to an IP multicast flow.
- Opaque value can be assigned by the ingress root PE or statically configured.
- MP LSP is created on-demand or can be pre-configured.
- Allows aggregating multicast streams onto a single MP LSP.

Label Distribution Protocol (LDP) Typed Wildcard Forward Equivalence Class (FEC) # RFC5918

The Typed Wildcard FEC Element refers to all FECs of the specified type that meet the constraint. It specifies a 'FEC Element Type' and an optional constraint, which is intended to provide additional information.

The format of the Typed Wildcard FEC Element is:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Typed (0x05)  | FEC Element   | Len FEC Type  |                 |
| Wildcard      | Type          | Info          |                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+                 |
|                                                                 |
~          Additional FEC Type-specific Information               ~
|                  (Optional)                                     |
|                                               +-+-+-+-+-+-+-+-+-+
|                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Typed Wildcard: One-octet FEC Element Type (0x05).

LDP [RFC5036] distributes labels for Forwarding Equivalence Classes (FECs). LDP uses FEC TLVs in LDP messages to specify FECs.

An LDP FEC TLV includes one or more FEC elements. A FEC element includes a FEC type and an optional type-dependent value.

RFC 5036 specifies two FEC types (Prefix and Wildcard), and other documents specify additional FEC types; e.g., see [RFC4447] and [MLDP].

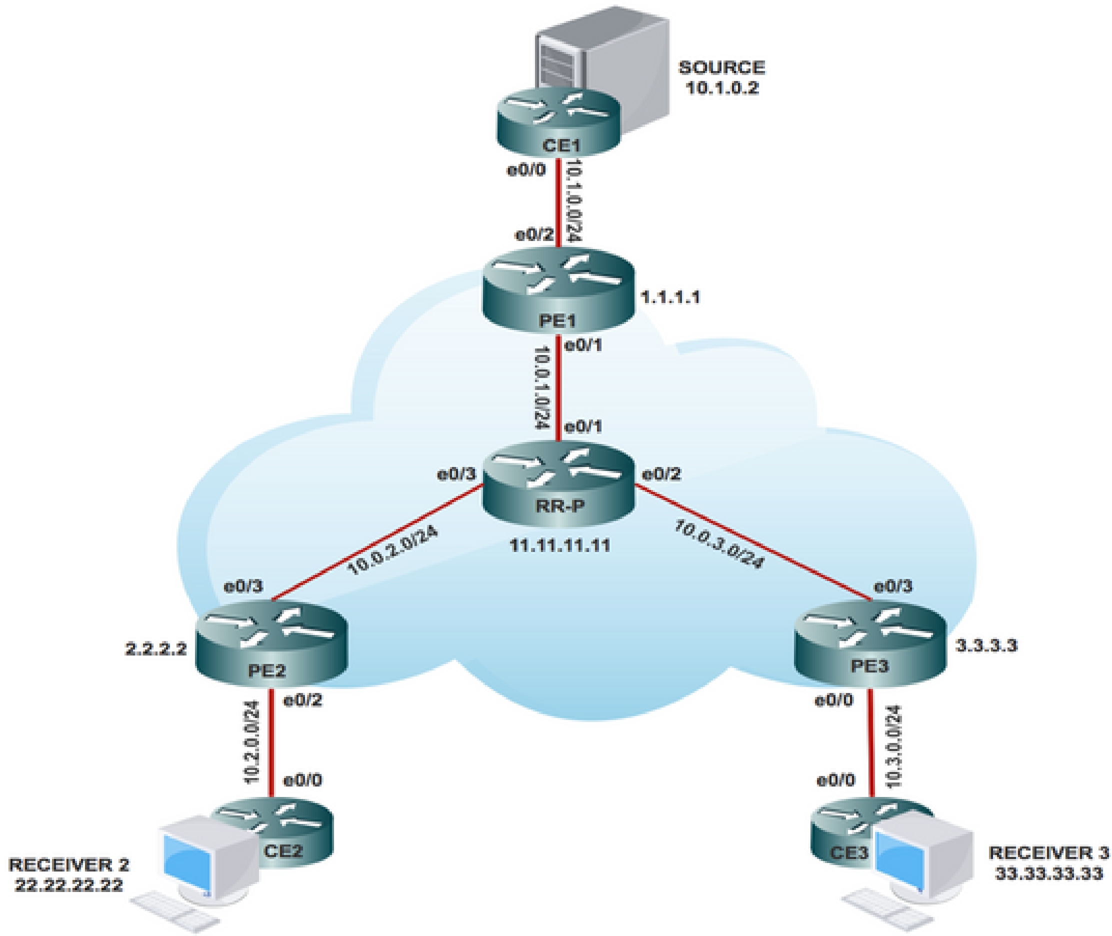
As specified by RFC 5036, the Wildcard FEC Element refers to all FECs relative to an optional constraint.

The only constraint RFC 5036 specifies is one that limits the scope of the Wildcard FEC Element to "all FECs bound to a given label".

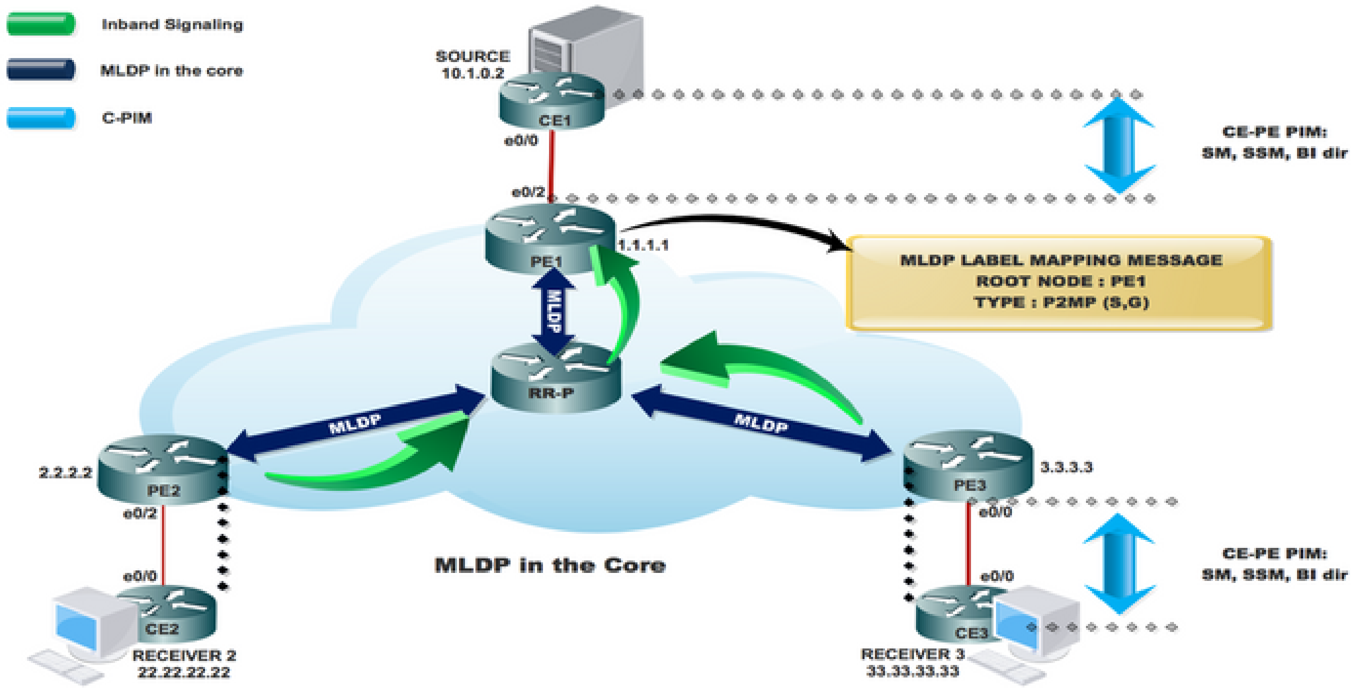
The RFC 5036 specification of the Wildcard FEC Element have these deficiencies that limit its utility:

- The Wildcard FEC Element is untyped. There are situations where it would be useful to be able to refer to all FECs of a given type (as another constraint).
- Use of the Wildcard FEC Element is limited to Label Withdraw and Label Release messages only. There are situations where it would be useful to have a Wildcard FEC Element, with type constraint, in Label Request messages.

Topology



Overlay Signaling



Configuration

Step 1. Enable MPLS MLDP in Core nodes.

 # mpls mldp logging

Step 2. Enable MLDP INBAND SIGNALING in CORE.

On PE1, PE2 and PE3

 # ip multicast vrf INBAND-MLDP mpls mldp

 # ip pim vrf INBAND-MLDP mpls source loopback 0

Step 3. Enable PIM SM in all CE interfaces and PE VRF interface.

On CE1, CE2, CE3 and all VRF interfaces PE1, PE2 and PE3

 # interface x/x

 # ip pim sparse-mode

 # interface loopback x/x

 # ip pim sparse-mode

> **Note**: Enable **PIM-Mode** only in CE facing interfaces on Provider edge routers; not required in the core.

Step 4. Enable multicast in the VRF.

On PE1, PE2 and PE3

 # ip multicast-routing vrf INBAND-MLDP

Step 5. Enable VRF on PE-CE interface x/x of the PE router.

 # interface x/x

 # ip vrf forwarding INBAND-MLDP

Step 6. Configure mode SSM in CE and PE nodes (VRF only).

On CE nodes

 # ip pim ssm default

On PE1, PE2, PE3 under VRF

 # ip pim vrf INBAND-MLDP ssm default

Step 7. Configure IGMP group SSM 232.1.1.1 (receiver).

On receiver 2 and 3

CE #interface x/x

 # ip pim join-group 232.1.1.1 source 10.1.0.2

Verify

IGP, MPLS LDP, BGP runs fine across our network end to end.

In this section, verification is done to check the VPN AF adjacency in the core/aggregation network. Adjacency is checked between CE-PE and the control-plane is also checked along with the Data plane for VPN traffic over MPLS network.

To verify that the local and remote customer edge (CE) devices can communicate across the Multiprotocol Label Switching (MPLS) core, perform these tasks:

Task 1: Verify Physical Connectivity.

- Verify all the connected interface are **UP**.

Task 2: Verify BGP Address Family VPNv4 unicast.

- Verify that BGP is enabled in all the routers for AF VPNv4 unicast and BGP neighbours are **UP**.
- Verify that BGP VPNv4 unicast table has all the Customer prefixes.

Task 3: Verify Multicast Traffic end to end.

- Check PIM neighborship with the connected PIM neighbour.
- Verify that multicast state is created in the VRF.

On PE VRF mRIB entry on PE1, PE2 and PE3

- Verify that VRF (S, G) mFIB entry, packet getting incremented in software forwarding.
- Verify ICMP packets getting reach from CE to CE.

Task 4: Verify MPLS CORE.

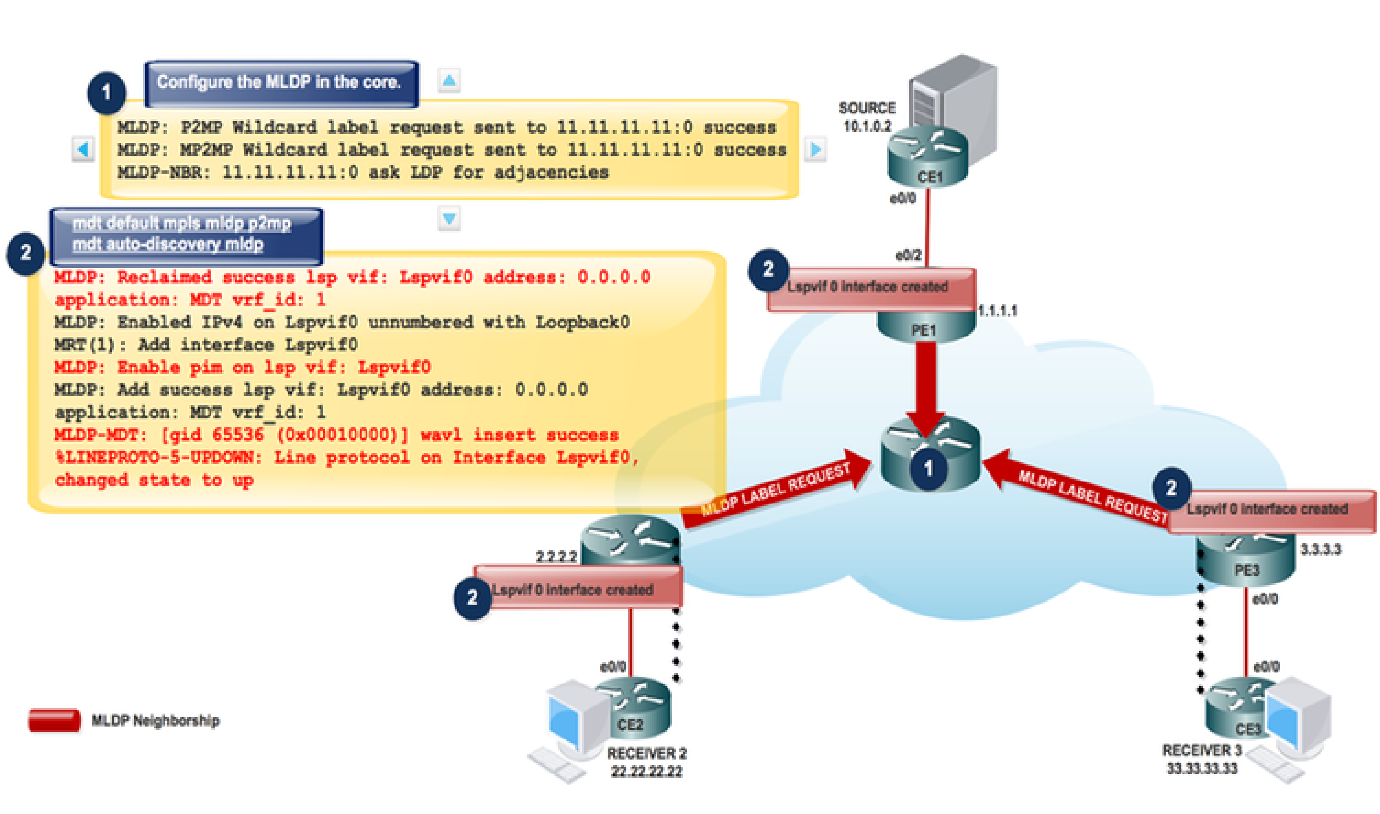
- Verify the MPLS LSP core.
- Verify MPLS forwarding inside the core as per design.
- Verify the MPLS P2MP LSP ping work.

How to Build a Control Plane?

Verify the control plane that Label imposition occurs when the PE router forwards based on the IP header and adds an MPLS label to the packet upon entering an MPLS network.

In the direction of label imposition, the router switches packets based on a CEF table lookup to find the next hop and adds the appropriate label information stored in the FIB for the destination. When a router performs label swapping in the core on an MPLS packet, the router does an MPLS table lookup. The router derives this MPLS table (LFIB) from information in the CEF table and the Label Information Base (LIB).

Label disposition occurs when the PE router receives an MPLS packet, makes a forwarding decision based on the MPLS label, removes the label, and sends an IP packet. The PE router uses the LFIB for path determination for a packet in this direction.As stated previously, a special iBGP session facilitates the advertisement of VPNv4 prefixes and their labels between PE routers. At the advertising PE, BGP allocates labels for the VPN prefixes learned locally and installs them in the LFIB, which is the MPLS forwarding table.



Step 1. Once you configure the MLDP in the core. These messages exchange.

```
MLDP: P2MP Wildcard label request sent to 11.11.11.11:0 success
MLDP: MP2MP Wildcard label request sent to 11.11.11.11:0 success
MLDP-MFI: Enabled MLDP MFI client on Lspvif0; status =  ok
LDP Peer 11.11.11.11:0 re-announced
MLDP-NBR: 11.11.11.11:0 UP sess_hndl: 1, (old ID: 0.0.0.0:0)
mLDP-RW: Sending RW notification message to process: mLDP Process
mLDP-RW: RW Tracking started for: 11.11.11.11
MLDP-LDP: [id 0] Wildcard label request from: 11.11.11.11:0 label: 0 root: 6.2.0.0 Opaque_len: 0
sess_hndl: 0x1
MLDP-LDP: [id 0] Wildcard label request from: 11.11.11.11:0 label: 0 root: 8.2.0.0 Opaque_len: 0
sess_hndl: 0x1

Neighbor 11.11.11.11 request for the label request to PE1.
```

Use this Debug to check the preceding establishment:

 # debug mpls mldp all

> **Note**: Respond to Typed Wildcard Label Requests received from peer by replaying its label database for prefixes. Make use of Typed Wildcard Label Requests towards peers to request a replay of peer label database for prefixes.

Step 2. Enable INBAND SIGNALING in VRF.

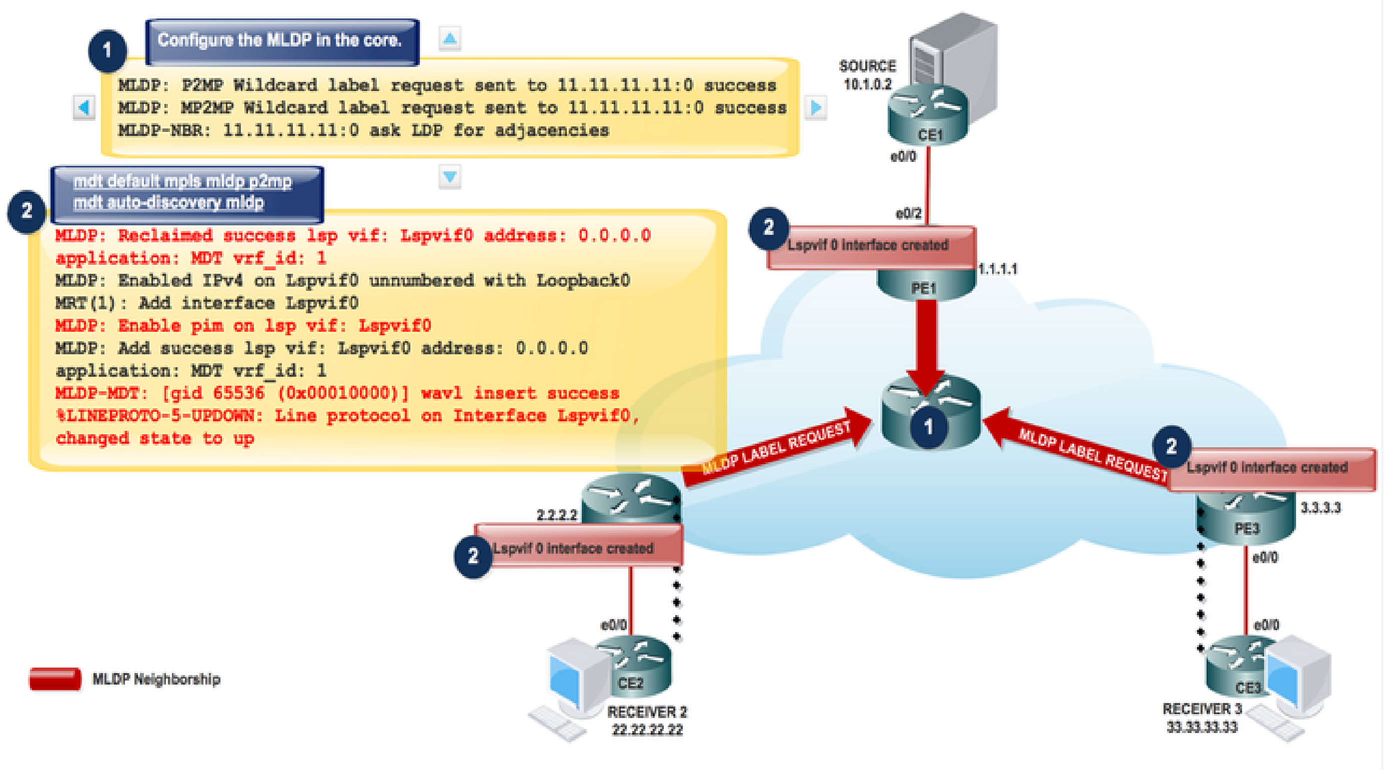
```
PE1 # Config t
# ip pim vrf MLDP-INBAND mpls source loopback 0
# ip multicast vrf MLDP-INBAND mpls mldp
```

```
MLDP: Enabled IPv4 on Lspvif0 unnumbered with Loopback0
MLDP-MFI: Enable lsd on int failed; not registered;
MLDP: Enable pim on lsp vif: Lspvif0
MLDP: Add success lsp vif: Lspvif0 address: 0.0.0.0 application: MLDP vrf_id: 1
MLDP-DB: Replaying database events for opaque type value: 250
%LINEPROTO-5-UPDOWN: Line protocol on Interface Lspvif0, changed state to up
PIM(1): Check DR after interface: Lspvif0 came up!
PIM(1): Changing DR for Lspvif0, from 0.0.0.0 to 1.1.1.1 (this system)
%PIM-5-DRCHG: VRF MLDP-INBAND: DR change from neighbor 0.0.0.0 to 1.1.1.1 on interface Lspvif0

Use this Debug to check the preceding establishment
# debug ip pim vrf LDP-INBAND6


PE1#sh interfaces lspvif 0
Lspvif0 is up, line protocol is up
  Hardware is
  Interface is unnumbered. Using address of Loopback0 (1.1.1.1)
  MTU 17940 bytes, BW 8000000 Kbit/sec, DLY 5000 usec,
     reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation LOOPBACK, loopback not set
```

Note: MPLS MLDP is not yet created as Receiver is not online yet.

When Receiver comes online:

Receiver 3 comes online and sends the PIM JOIN (S, G) messages to PE3.




```
PIM(1): Received v2 Join/Prune on Ethernet0/2 from 10.2.0.2, to us
PIM(1): Join-list: (10.1.0.2/32, 232.1.1.1), S-bit set
MRT(1): Create (*,232.1.1.1), RPF (unknown, 0.0.0.0, 2147483647/0)
MLDP: Interface Lspvif1 moved from VRF (default) to VRF MLDP-INBAND
MLDP: Enabled IPv4 on Lspvif1 unnumbered with Loopback0
MLDP-MFI: Enabled MLDP MFI client on Lspvif1; status =  ok
```

```
MRT(1): Add interface Lspvif1
MLDP: Enable pim on lsp vif: Lspvif1
MLDP: Add success lsp vif: Lspvif1 address: 1.1.1.1 application: MLDP vrf_id: 1


MLDP: LDP root 1.1.1.1 added
mLDP-RW: Sending RW notification message to process: mLDP Process
mLDP-RW: RW Tracking started for: 1.1.1.1
MLDP:  Route watch started for 1.1.1.1 topology: base ipv4
MLDP-DB: Added [vpnv4 10.1.0.2 232.1.1.1 1:1] DB Entry
MLDP-DB: [vpnv4 10.1.0.2 232.1.1.1 1:1] Added P2MP branch for MRIBv4(1) label
%MLDP-5-ADD_BRANCH: [vpnv4 10.1.0.2 232.1.1.1 1:1] Root: 1.1.1.1, Add P2MP branch MRIBv4(1)
remote label

MLDP: nhop 10.0.2.2 added
MLDP-NBR: 11.11.11.11:0 mapped to next_hop: 10.0.2.2
MLDP: Root 1.1.1.1 old paths: 0 new paths: 1
MLDP-DB: [vpnv4 10.1.0.2 232.1.1.1 1:1] Changing peer from none to 11.11.11.11:0
MLDP-DB: [vpnv4 10.1.0.2 232.1.1.1 1:1] Add accepting element nbr: 11.11.11.11:0
MLDP: [vpnv4 10.1.0.2 232.1.1.1 1:1] label mappping msg sent to 11.11.11.11:0 success
MLDP-DB: [vpnv4 10.1.0.2 232.1.1.1 1:1] path to peer: 11.11.11.11:0 changed None:0.0.0.0 to
Ethernet0/3:10.0.2.2
```

Any communication from Receiver (S,G) Join, will be converted to MLDP and all messages are traverse towards the Lspvif 1

With PIM JOIN (S,G) as MLDP is receiver-driven protocol, it starts building the MLDP database from Receiver to Source. This is Downstream Label allocation for P2MP MLDP.

P2MP MLDP

The P2MP packet transport is implemented using Resource Reservation Protocol (RSVP) P2MP – Traffic Engineering (P2MP-TE) and M2M packet transport are implemented through IPv4 Multicast VPN (MVPN) using multicast Label Distribution Protocol (MLDP).

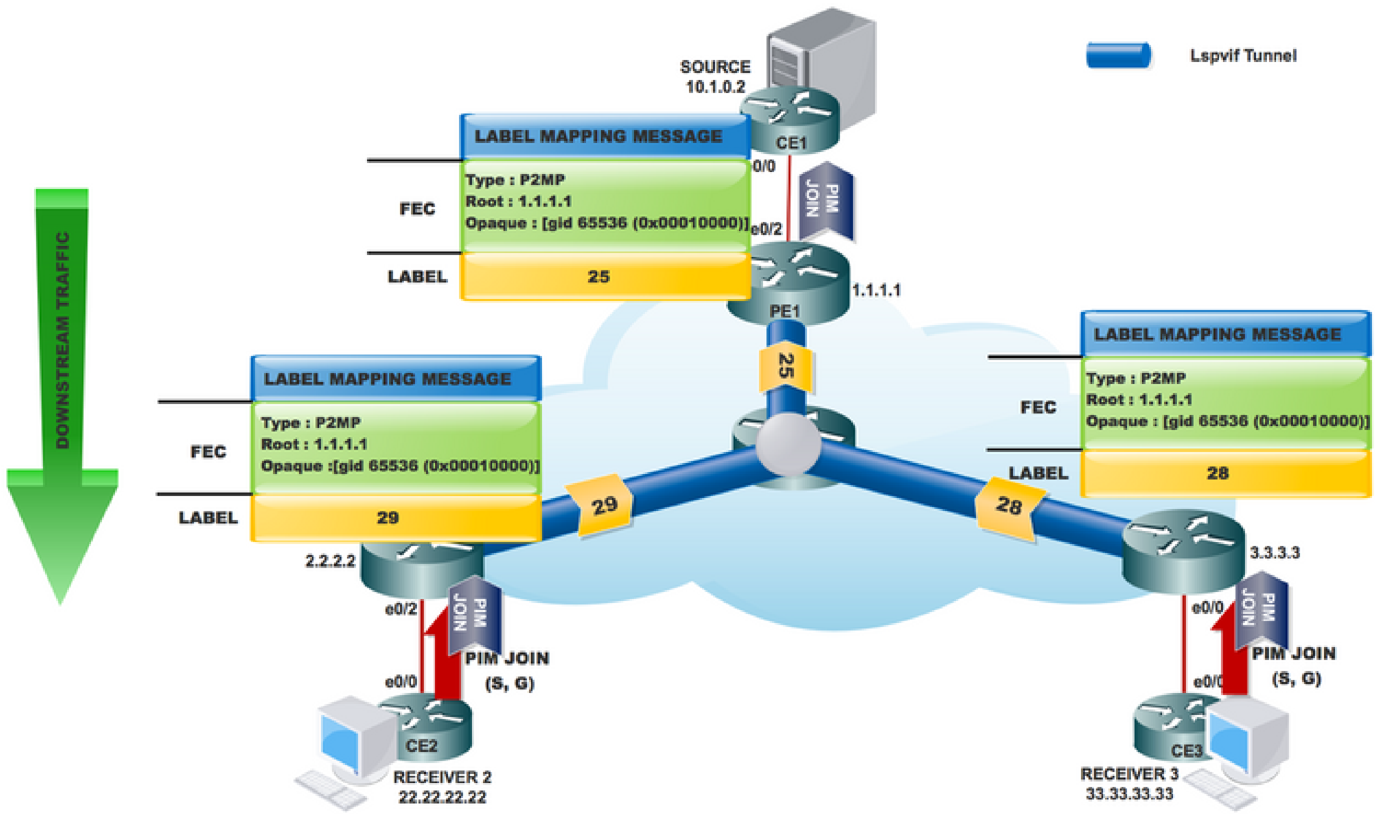
The packet is transported over three types of routers:

• Headend router: Encapsulates the IP packet with one or more labels.

• Midpoint router: Replaces the in-label with an out-label.

• Tailend router: Removes the label from the packet.

Packet Flow in MLDP-based MVPN Network For each packet coming in, MPLS creates multiple out-labels. Packets from the source network are replicated along the path to the receiver network. The CE1 router sends out the native IP multicast traffic. The PE1 router imposes a label on the incoming multicast packet and replicates the labelled packet towards the MPLS core network. When the packet reaches the core router (P), the packet is replicated with the appropriate labels for the MP2MP default MDT or the P2MP data MDT and transported to all the egress PEs. Once the packet reaches the egress PE, the label is removed and the IP multicast packet is replicated onto the VRF interface




```
PE1#sh mpls mldp database
  * For interface indicates MLDP recursive forwarding is enabled
  * For RPF-ID indicates wildcard value
  > Indicates it is a Primary MLDP MDT Branch

LSM ID : 1   Type: P2MP   Uptime : 00:23:11
  FEC Root           : 1.1.1.1 (we are the root)
  Opaque decoded     : [vpnv4 10.1.0.2 232.1.1.1 1:1]
  Opaque length      : 16 bytes
  Opaque value       : FA 0010 0A010002E801010100000001000000001
  Upstream client(s) :
    None
      Expires        : N/A            Path Set ID  : 1
  Replication client(s):
    11.11.11.11:0
      Uptime         : 00:23:11       Path Set ID  : None
      Out label (D)  : 21             Interface    : Ethernet0/1*
      Local label (U): None           Next Hop     : 10.0.1.2


RR-P#sh mpls mldp database
  * For interface indicates MLDP recursive forwarding is enabled
  * For RPF-ID indicates wildcard value
  > Indicates it is a Primary MLDP MDT Branch

LSM ID : 2   Type: P2MP   Uptime : 00:28:12
  FEC Root           : 1.1.1.1
  Opaque decoded     : [vpnv4 10.1.0.2 232.1.1.1 1:1]
  Opaque length      : 16 bytes
  Opaque value       : FA 0010 0A010002E801010100000001000000001
  Upstream client(s) :
    1.1.1.1:0    [Active]
      Expires        : Never          Path Set ID  : 2
      Out Label (U)  : None           Interface    : Ethernet0/1*
      Local Label (D): 21             Next Hop     : 10.0.1.1
```

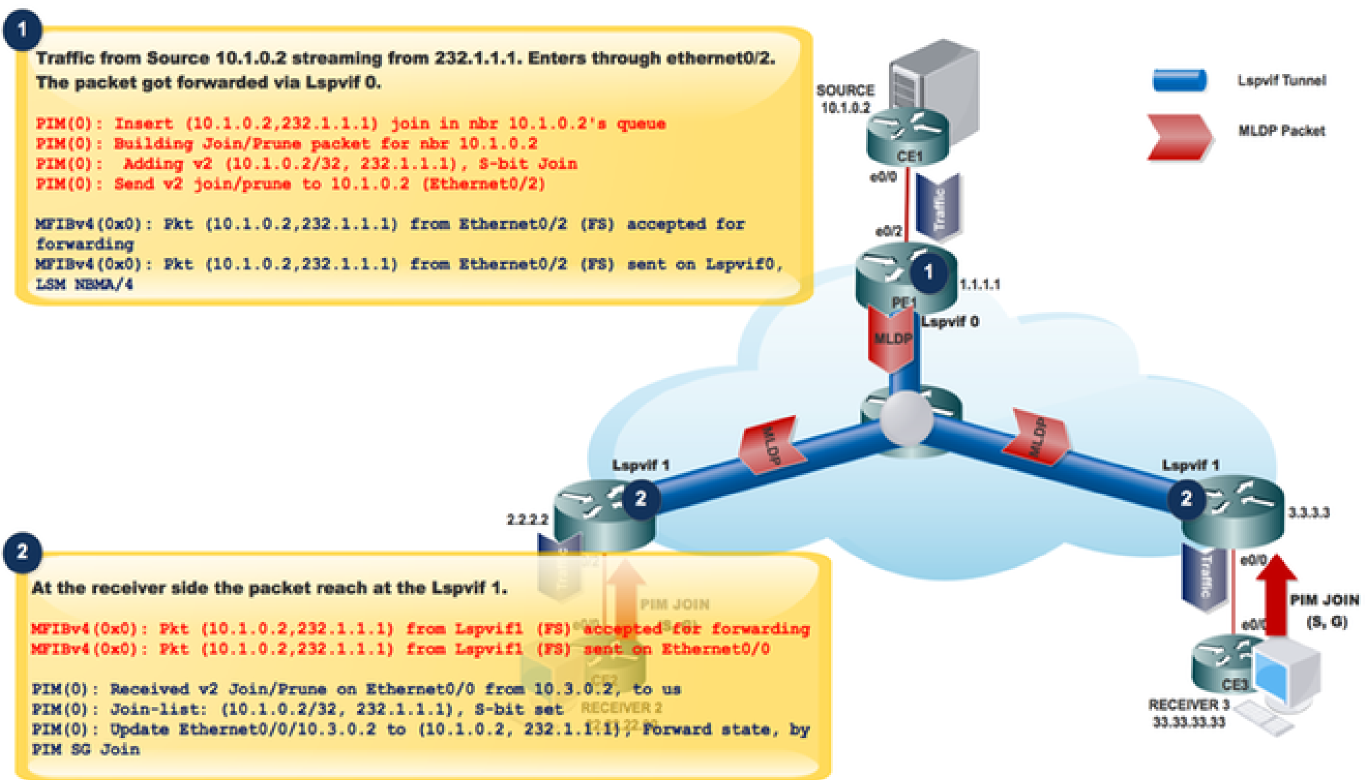
```
  Replication client(s):
    3.3.3.3:0
      Uptime         : 00:28:12      Path Set ID  : None
      Out label (D)  : 26            Interface    : Ethernet0/2*
      Local label (U): None          Next Hop     : 10.0.3.1
    2.2.2.2:0
      Uptime         : 00:24:41      Path Set ID  : None
      Out label (D)  : 25            Interface    : Ethernet0/3*
      Local label (U): None          Next Hop     : 10.0.2.1


RR-P#sh mpls forwarding-table labels 21
Local      Outgoing   Prefix           Bytes Label   Outgoing   Next Hop
Label      Label      or Tunnel Id     Switched      interface
21         26         [vpnv4 10.1.0.2 232.1.1.1 1:1]   \
                                       0             Et0/2      10.0.3.1
           25         [vpnv4 10.1.0.2 232.1.1.1 1:1]   \
                                       0             Et0/3      10.0.2.1
```

MRIB created on PE devices:

```
PE1#sh ip mroute vrf MLDP-INBAND 232.1.1.1 verbose
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group
       T - SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet,
       U - URD, I - Received Source Specific Host Report,

(10.1.0.2, 232.1.1.1), 00:00:17/00:02:42, flags: sTI
  Incoming interface: Ethernet0/2, RPF nbr 10.1.0.2
  Outgoing interface list:
    Lspvif0, LSM ID: 1, Forward/Sparse, 00:00:17/00:02:42
```

When Source starts streaming :

When the multicast source starts sending traffic, [10.1.0.2, 232.1.1.1] happen as shown in this image.



Traffic from Source 10.1.0.2 streaming from 232.1.1.1. Enters through ethernet0/2.

The packet got forwarded via Lspvif 0.

```
PIM(0): Insert (10.1.0.2,232.1.1.1) join in nbr 10.1.0.2's queue
PIM(0): Building Join/Prune packet for nbr 10.1.0.2
PIM(0):  Adding v2 (10.1.0.2/32, 232.1.1.1), S-bit Join
PIM(0): Send v2 join/prune to 10.1.0.2 (Ethernet0/2)
```

```
MFIBv4(0x0): Pkt (10.1.0.2,232.1.1.1) from Ethernet0/2 (FS) accepted for forwarding
MFIBv4(0x0): Pkt (10.1.0.2,232.1.1.1) from Ethernet0/2 (FS) sent on Lspvif0, LSM NBMA/4
```

```
     36 28.764034     10.1.0.2            232.1.1.1          ICMP     118 Echo (ping) request  id=0x0001,
Frame 36: 118 bytes on wire (944 bits), 118 bytes captured (944 bits) on interface 0
Ethernet II, Src: aa:bb:cc:00:10:10 (aa:bb:cc:00:10:10), Dst: aa:bb:cc:00:30:10 (aa:bb:cc:00:30:10)
MultiProtocol Label Switching Header, Label: 24, Exp: 0, S: 1, TTL: 254
Internet Protocol Version 4, Src: 10.1.0.2, Dst: 232.1.1.1
Internet Control Message Protocol
```

This packet get tunneled into the Lspvif 0.

At the receiver Side:

At the receiver side the packet reach at the Lspvif 1.

```
MFIBv4(0x0): Pkt (10.1.0.2,232.1.1.1) from Lspvif1 (FS) accepted for forwarding
MFIBv4(0x0): Pkt (10.1.0.2,232.1.1.1) from Lspvif1 (FS) sent on Ethernet0/0
```

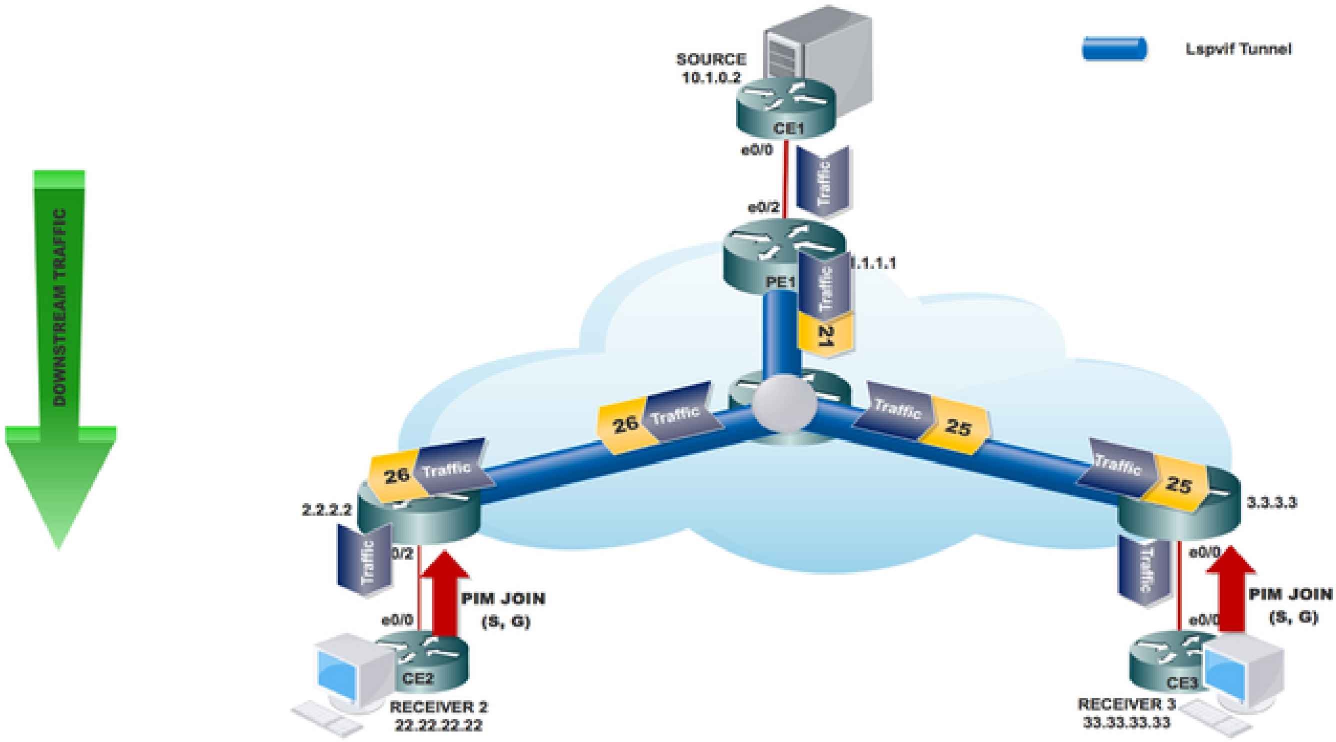
```
PIM(0): Received v2 Join/Prune on Ethernet0/0 from 10.3.0.2, to us
PIM(0): Join-list: (10.1.0.2/32, 232.1.1.1), S-bit set
PIM(0): Update Ethernet0/0/10.3.0.2 to (10.1.0.2, 232.1.1.1), Forward state, by PIM SG Join
```

When packet hit the PE1, it checks the LSM ID to forward the traffic, which label to impose in the Multicast packet.

This image shows the verification of LSPVIF interface.

```
PE1#sh ip mroute vrf MLDP-INBAND 232.1.1.1 verbose
IP Multicast Routing Table
Flags: s - SSM Group, C - Connected,
       T - SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet,
       I - Received Source Specific Host Report,

((10.1.0.2, 232.1.1.1), 00:07:07/stopped, flags: sTI
  Incoming interface: Ethernet0/2, RPF nbr 10.1.0.2
  Outgoing interface list:
    Lspvif0, LSM ID: 1, Forward/Sparse, 00:07:07/00:01:51
```

```
PE1#sh mpls mldp database

LSM ID : 1   Type: P2MP   Uptime : 00:40:23
  FEC Root            : 1.1.1.1 (we are the root)
  Opaque decoded      : [vpnv4 10.1.0.2 232.1.1.1 1:1]
  Opaque length       : 16 bytes
  Opaque value        : FA 0010 0A010002E80101010000000100000001
  Upstream client(s) :
    None
      Expires         : N/A            Path Set ID  : 1
  Replication client(s):
    11.11.11.11:0
      Uptime          : 00:40:23       Path Set ID  : None
      Out label (D)   : 21             Interface    : Ethernet0/1*
      Local label (U): None            Next Hop     : 10.0.1.2
```



For each packet coming in, MPLS creates multiple out-labels. Packets from the source network are replicated along the path to the receiver network. The CE1 router sends out the native IP multicast traffic. The PE1 router imposes a label on the incoming multicast packet and replicates the labelled packet towards the MPLS core network.

When the packet reaches the core router (P), the packet is replicated with the appropriate labels for the MP2MP default MDT or the P2MP data MDT and transported to all the egress PEs. Once the packet reaches the egress PE, the label is removed and the IP multicast packet is replicated onto the VRF interface.

The Conclusion

The MLDP MVPN configuration enables IPv4 multicast packet delivery using MPLS. This configuration uses MPLS labels to construct default and data Multicast Distribution Trees (MDTs).

The MPLS replication is used as a forwarding mechanism in the core network. For MLDP MVPN configuration to work, ensure that the global MPLS MLDP configuration is enabled.

Related Information

- **https://tools.ietf.org/html/rfc5918**
- **https://tools.ietf.org/html/rfc4447**
- **https://www.cisco.com/c/en/us/td**/docs/ios-xml/ios/ipmulti_lsm/configuration/15-sy/imc-lsm-15-sy-book.pdf
- **Technical Support & Documentation - Cisco Systems**