Implementing Layer-3 Multicast Routing

Want to deliver messages like corporate communications or newsletters to subscribed members using a minimum of network bandwidth?

With the traditional method like unicast, you can send messages from one source to one destination. Each host added to the network consumes bandwidth and it’s a challenge to reduce the load on the traffic.

On the other hand, broadcast sends messages to all the hosts in the network and not to the selected members.

Enable Multicast routing to deliver data traffic efficiently from a single source to multiple users or selected members or even a group. It’s scalable and yet reduces the load on the traffic.

Learn about Multicast

Many applications such as video conferencing, corporate communications, distance learning, and distribution of software, stock quotes, and news involve multiple participants. Multicast is naturally suitable for this communication paradigm.

Unlike unicast and broadcast, multicast allows a host to send a single data stream to a subset of hosts (group transmission) at about the same time. The IP hosts subscribed to a group are known as group members.

A multicast address is chosen from the multicast group. The sender uses that group address as the destination address of a datagram to reach all members of the group

Packets delivered to group members are identified by a single multicast group address. Multicast packets are delivered to a group using best-effort reliability, just like IP unicast packets.

Membership in a multicast group is dynamic; hosts can join and leave at any time. There's no restriction on the location or number of members in a multicast group. A host can be a member of more than one multicast group at a time.

To send messages, multicast routing uses the following components:

- The sender or the source address
- The receiver or the multicast address

  The receiver can be a group of members and are identified by a single multicast group address that falls under the IP Class D address range from 224.0.0.0 through 239.255.255.255. A multicast address is chosen for the receivers in a multicast group. Senders use that group address as the destination address of a datagram to reach all members of the group.

Note Any host, regardless of whether it's a member of a group or not, can send to a group. However, only the members of a group receive the message

- A protocol to identify the selected users to send a message.

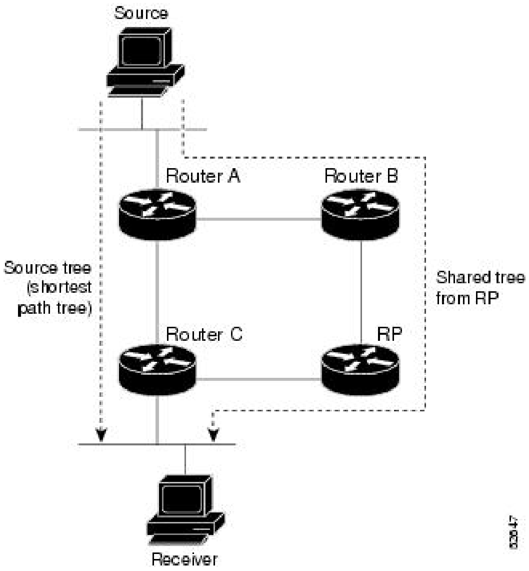
Cisco IOS XR Software supports the following protocols to implement multicast routing:

- IGMP (IPv4): Use IGMP to allow hosts (IPv4) to communicate with routers to express the interest to receive multicast traffic on specific groups. Use Multicast Listener Discovery (MLD v1/2) for IPv6.
- Protocol Independent Multicast in sparse mode (PIM-SM): Use PIM-SM between routers to track which multicast packets to forward to each other and to their directly connected LANs.
- Protocol Independent Multicast in Source-Specific Multicast (PIM-SSM): PIM-SSM is similar to PIM-SM. Hosts use PIM-SSM to report interest in receiving packets from specific source addresses.

  PIM-SSM is made possible by IGMPv3 and MLDv2. Hosts can now indicate interest in specific sources using IGMPv3 and MLDv2. SSM doesn't require a rendezvous point (RP) to operate.

This image shows IGMP and PIM-SM operating in a multicast environment.

Figure 1: Multicast Routing Protocols



- Key Protocols and Features Supported in the Cisco IOS XR Software Multicast Routing Implementation, on page 3
- Prerequisites for Implementing Multicast Routing, on page 3
- Restrictions for Implementing Multicast Routing, on page 4
- Configuring Multicast, on page 4
- Internet Group Management Protocol , on page 5
- Protocol Independent Multicast, on page 7
- Multicast-Only Fast Reroute, on page 22
- Multicast Source Discovery Protocol, on page 23

- Multicast Nonstop Forwarding, on page 24
- Multicast Configuration Submodes, on page 24
- Understanding Interface Configuration Inheritance, on page 25
- Understanding Interface Configuration Inheritance Disablement, on page 26
- Understanding Enabling and Disabling Interfaces, on page 27
- Controlling Source Information on MSDP Peer Routers, on page 27
- Multicast Routing Information Base, on page 28
- Multicast Forwarding Information Base, on page 28
- MSDP MD5 Password Authentication, on page 29
- Label Switch Multicast, on page 30
- mLDP Loop-Free Alternative Fast Reroute, on page 32
- Multicast Route Statistics, on page 44

Key Protocols and Features Supported in the Cisco IOS XR Software Multicast Routing Implementation

Table 1: Supported Features for IPv4 and IPv6

Feature	IPv4 Support	IPv6 Support
Auto-RP	Yes	No
BGP	Yes	Yes
BSR	Yes	Yes
Dynamic host registration	Yes (IGMP v2/3)	Yes (MLD v1/2)
Explicit tracking of hosts, groups, and channels	Yes (IGMP v3)	Yes
MSDP	Yes	No
Multicast NSF	Yes	Yes
OOR handling	Yes	Yes
PIM-SM	Yes	Yes
PIM-SSM	Yes	Yes
PIM-SSM Mapping	Yes	Yes

Prerequisites for Implementing Multicast Routing

- You must be in a user group associated with a task group that includes the proper task IDs. The command reference guides include the task IDs required for each command. If you suspect user group assignment is preventing you from using a command, contact your AAA administrator for assistance.

- You must be familiar with IPv4 and IPv6 multicast routing configuration tasks and concepts.
- Unicast routing must be operational.
- To enable multicast VPN, configure a VPN routing and forwarding (VRF) instance.

Restrictions for Implementing Multicast Routing

- The following features are not supported:
    - Multicast VPN
    - InterAS Option A
    - PIM Bidir
- IPv6 Multicast destination addresses are only allowed with a /96 mask. IPv6 Multicast destination address should vary only in the last 32 bits of the group address. If they vary outside this range, they might map to the same entry in the hardware.
-

Configuring Multicast

To configure multicast, perform the following configuration:

```
Router#configure
Router(config)# multicast-routing
Router(config-mcast)#address-family ipv4
Router(config-mcast-default-ipv4)#interface all enable
Router(config-mcast-default-ipv4)#exit
Router(config-mcast)#router igmp
Router(config-igmp)#version 3
Router(config-igmp)#commit
Tue Feb  4 04:43:37.679 UTC
Router(config-igmp)#exit
Router(config)#exit
```

Verification

```
Router#show pim ipv4 group-map
Tue Feb  4 04:48:29.003 UTC

IP PIM Group Mapping Table
(* indicates group mappings being used)
(+ indicates BSR group mappings active in MRIB)

Group Range         Proto Client   Groups RP address      Info

224.0.1.39/32*      DM    perm     0      0.0.0.0
224.0.1.40/32*      DM    perm     0      0.0.0.0
224.0.0.0/24*       NO    perm     0      0.0.0.0
232.0.0.0/8*        SSM   config   0      0.0.0.0
224.0.0.0/4*        SM    static   0      0.0.0.0         RPF: Null,0.0.0.0
```

To view the PIM topology table information for a specific group or all groups.

```
Router#show pim topology
IP PIM Multicast Topology Table
Entry state: (*/S,G)[RPT/SPT] Protocol Uptime Info
Entry flags: KAT - Keep Alive Timer, AA - Assume Alive, PA - Probe Alive
    RA - Really Alive, IA - Inherit Alive, LH - Last Hop
    DSS - Don't Signal Sources,  RR - Register Received
    SR - Sending Registers, SNR - Sending Null Registers
    E - MSDP External, EX - Extranet
    MFA - Mofrr Active, MFP - Mofrr Primary, MFB - Mofrr Backup
    DCC - Don't Check Connected, ME - MDT Encap, MD - MDT Decap
    MT - Crossed Data MDT threshold, MA - Data MDT Assigned
    SAJ - BGP Source Active Joined, SAR - BGP Source Active Received,
    SAS - BGP Source Active Sent, IM - Inband mLDP, X - VxLAN
Interface state: Name, Uptime, Fwd, Info
Interface flags: LI - Local Interest, LD - Local Dissinterest,
    II - Internal Interest, ID - Internal Dissinterest,
    LH - Last Hop, AS - Assert, AB - Admin Boundary, EX - Extranet,
    BGP - BGP C-Multicast Join, BP - BGP Source Active Prune,
    MVS - MVPN Safi Learned, MV6S - MVPN IPv6 Safi Learned

(*,224.0.1.40) DM Up: 00:56:47 RP: 0.0.0.0
JP: Null(never) RPF: Null,0.0.0.0 Flags: LH DSS
  Loopback0                   00:56:47  off LI II LH

(21.5.7.2,232.1.1.1)SPT SSM Up: 00:00:44
JP: Join(00:00:05) RPF: Null,0.0.0.0 Flags:
  FourHundredGigE0/0/0/11     00:00:44  fwd LI LH
```

Internet Group Management Protocol

Cisco IOS XR Software provides support for Internet Group Management Protocol (IGMP) over IPv4.

IGMP provides a means for hosts to indicate which multicast traffic they are interested in and for routers to control and limit the flow of multicast traffic throughout the network. Routers build state by means of IGMP and MLD messages; that is, router queries and host reports.

A set of queries and hosts that receive multicast data streams from the same source is called a *multicast group*. Hosts use IGMP and MLD messages to join and leave multicast groups.

Note IGMP messages use group addresses, which are Class D IP addresses. The high-order four bits of a Class D address are 1110. Host group addresses can be in the range 224.0.0.0 to 239.255.255.255. The address 224.0.0.0 is guaranteed not to be assigned to any group. The address 224.0.0.1 is assigned to all systems on a subnet. The address 224.0.0.2 is assigned to all routers on a subnet.

IGMP Versions

The following points describe IGMP versions 2, and 3:

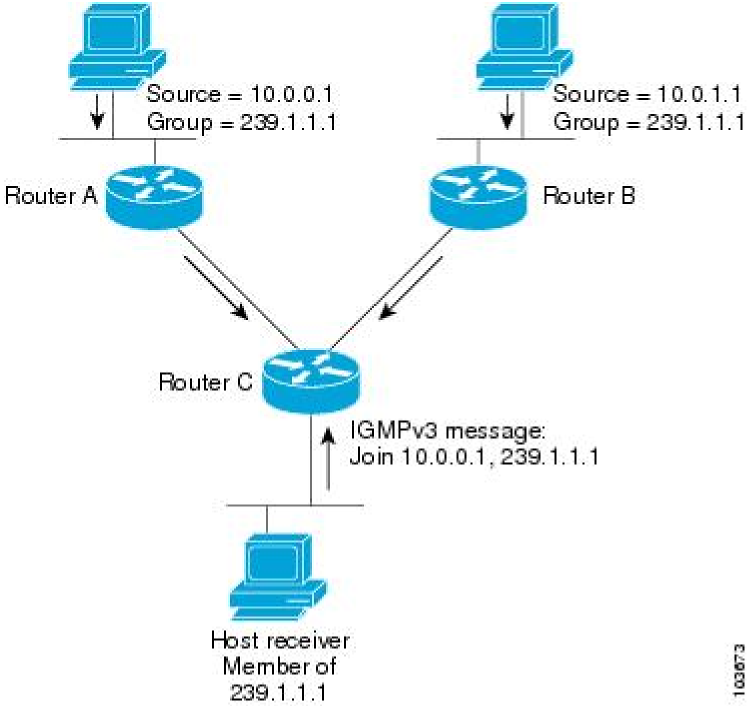
- IGMP Version 2 extends IGMP allowing such features as the IGMP query timeout and the maximum query-response time. See RFC 2236.

- IGMP Version 3 permits joins and leaves for certain source and group pairs instead of requesting traffic from all sources in the multicast group.

IGMP Routing Example

The below image illustrates two sources, 10.0.0.1 and 10.0.1.1, that are multicasting to group 239.1.1.1. The receiver wants to receive traffic addressed to group 239.1.1.1 from source 10.0.0.1 but not from source 10.0.1.1. The host must send an IGMPv3 message containing a list of sources and groups (S, G) that it wants to join and a list of sources and groups (S, G) that it wants to leave. Router C can now use this information to prune traffic from Source 10.0.1.1 so that only Source 10.0.0.1 traffic is being delivered to Router C.

Figure 2: IGMPv3 Signaling



Note When configuring IGMP, ensure that all systems on the subnet support the same IGMP version. The router does not automatically detect Version 1 systems. Configure the router for Version 2 if your hosts do not support Version 3.

Configuring IGMP Per Interface States Limit

The IGMP Per Interface States Limit sets a limit on creating OLEs for the IGMP interface. When the set limit is reached, the group is not accounted against this interface but the group can exist in IGMP context for some other interface.

The following configuration sets a limit on the number of group memberships created on an interface as a result of receiving IGMP or MLD membership reports.

```
router igmp | mld [vrf <vrfname>]
            interface <ifname>
                        (no) maximum groups-per-interface <max> [threshold <threshold>]
[<acl>]
            !
!
```

where,

`<ifname>` is the interface name

`<max>` is the maximum limit on the groups

`<threshold>` is the threshold number of groups at which point a syslog warning message will be issued

`<acl>` provides an option for selective accounting. If provided, only groups or (S,G)s that are permitted by the ACL is accounted against the limit. Groups or (S, G)s that are denied by the ACL are not accounted against the limit. If not provided, all the groups are accounted against the limit.

The following messages are displayed when the threshold limit is reached for IGMP:

```
igmp[1160]: %ROUTING-IPV4_IGMP-4-OOR_THRESHOLD_REACHED : Threshold for Maximum number of
group per interface has been reached 3: Groups joining will soon be throttled.
Config a higher max or take steps to reduce states

igmp[1160]: %ROUTING-IPV4_IGMP-4-OOR_LIMIT_REACHED : Maximum number of group per interface
 has been reached 6: Groups joining is throttled.
Config a higher max or take steps to reduce states
```

Limitations

- If a user has configured a maximum of 20 groups and has reached the maximum number of groups, then no more groups can be created. If the user reduces the maximum number of groups to 10, the 20 joins will remain and a message of reaching the maximum is displayed. No more joins can be added until the number of groups has reached less than 10.
- If a user already has configured a maximum of 30 joins and add a max of 20, the configuration occurs displaying a message that the maximum has been reached. No state change occurs and also no more joins can occur until the threshold number of groups is brought down below the maximum number of groups.

Protocol Independent Multicast

Protocol Independent Multicast (PIM) is a routing protocol designed to send and receive multicast routing updates. Proper operation of multicast depends on knowing the unicast paths towards a source or an RP. PIM relies on unicast routing protocols to derive this reverse-path forwarding (RPF) information. As the name PIM implies, it functions independently of the unicast protocols being used. PIM relies on the Routing Information Base (RIB) for RPF information.

If the multicast subsequent address family identifier (SAFI) is configured for Border Gateway Protocol (BGP), or if multicast intact is configured, a separate multicast unicast RIB is created and populated with the BGP multicast SAFI routes, the intact information, and any IGP information in the unicast RIB. Otherwise, PIM gets information directly from the unicast SAFI RIB. Both multicast unicast and unicast databases are outside of the scope of PIM.

The Cisco IOS XR implementation of PIM is based on RFC 4601 Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification. For more information, see RFC 4601 and the Protocol Independent Multicast (PIM): Motivation and Architecture Internet Engineering Task Force (IETF) Internet draft.

Note Cisco IOS XR Software supports PIM-SM, PIM-SSM, and PIM Version 2 only. PIM Version 1 hello messages that arrive from neighbors are rejected.

PIM-Sparse Mode

Typically, PIM in sparse mode (PIM-SM) operation is used in a multicast network when relatively few routers are involved in each multicast. Routers do not forward multicast packets for a group, unless there is an explicit request for traffic. Requests are accomplished using PIM join messages, which are sent hop by hop toward the root node of the tree. The root node of a tree in PIM-SM is the rendezvous point (RP) in the case of a shared tree or the first-hop router that is directly connected to the multicast source in the case of a shortest path tree (SPT). The RP keeps track of multicast groups, and the sources that send multicast packets are registered with the RP by the first-hop router of the source.

As a PIM join travels up the tree, routers along the path set up the multicast forwarding state so that the requested multicast traffic is forwarded back down the tree. When multicast traffic is no longer needed, a router sends a PIM prune message up the tree toward the root node to prune (or remove) the unnecessary traffic. As this PIM prune travels hop by hop up the tree, each router updates its forwarding state appropriately. Ultimately, the forwarding state associated with a multicast group or source is removed. Additionally, if prunes are not explicitly sent, the PIM state will timeout and be removed in the absence of any further join messages.

PIM-SM is the best choice for multicast networks that have potential members at the end of WAN links.

PIM-Source Specific Multicast

When PIM-SM is used with SSM, multi-cast routing is easier to manage. This is because RPs (rendezvous points) are not required and therefore, no shared trees (*,G) are built.

There is no specific IETF document defining PIM-SSM. However, RFC4607 defines the overall SSM behavior.

In the rest of this document, we use the term PIM-SSM to describe PIM behavior and configuration when SSM is used.

PIM in Source-Specific Multicast operation uses information found on source addresses for a multicast group provided by receivers and performs source filtering on traffic.

- By default, PIM-SSM operates in the 232.0.0.0/8 multicast group range for IPv4. To configure these values, use the **ssm range** command.
- If SSM is deployed in a network already configured for PIM-SM, only the last-hop routers must be upgraded with Cisco IOS XR Software that supports the SSM feature.
- No MSDP SA messages within the SSM range are accepted, generated, or forwarded.
- SSM can be disabled using the **ssm disable** command.
- The ssm allow-override command allows SSM ranges to be overridden by more specific ranges.

In many multicast deployments where the source is known, protocol-independent multicast-source-specific multicast (PIM-SSM) mapping is the obvious multicast routing protocol choice to use because of its simplicity.

Typical multicast deployments that benefit from PIM-SSM consist of entertainment-type solutions like the ETTH space, or financial deployments that completely rely on static forwarding.

In SSM, delivery of data grams is based on (S,G) channels. Traffic for one (S,G) channel consists of datagrams with an IP unicast source address S and the multicast group address G as the IP destination address. Systems receive traffic by becoming members of the (S,G) channel. Signaling is not required, but receivers must subscribe or unsubscribe to (S,G) channels to receive or not receive traffic from specific sources. Channel subscription signaling uses IGMP to include mode membership reports, which are supported only in Version 3 of IGMP (IGMPv3).

To run SSM with IGMPv3, SSM must be supported on the multicast router, the host where the application is running, and the application itself. Cisco IOS XR Software allows SSM configuration for an arbitrary subset of the IP multicast address range 224.0.0.0 through 239.255.255.255.

When an SSM range is defined, existing IP multicast receiver applications do not receive any traffic when they try to use addresses in the SSM range, unless the application is modified to use explicit (S,G) channel subscription.

Benefits of PIM-SSM over PIM-SM

PIM-SSM is derived from PIM-SM. However, whereas PIM-SM allows for the data transmission of all sources sending to a particular group in response to PIM join messages, the SSM feature forwards traffic to receivers only from those sources that the receivers have explicitly joined. Because PIM joins and prunes are sent directly towards the source sending traffic, an RP and shared trees are unnecessary and are disallowed. SSM is used to optimize bandwidth utilization and deny unwanted Internet broad cast traffic. The source is provided by interested receivers through IGMPv3 membership reports.

PIM-SM and PIM-SSM

Protocol Independent Multicast (PIM) is a multicast routing protocol used to create multicast distribution trees, which are used to forward multicast data packets. PIM is an efficient IP routing protocol that is “independent” of a routing table, unlike other multicast protocols such as Multicast Open Shortest Path First (MOSPF) or Distance Vector Multicast Routing Protocol (DVMRP).

Cisco IOS XR Software supports Protocol Independent Multicast in sparse mode (PIM-SM) and Protocol Independent Multicast in Source-Specific Multicast (PIM-SSM), permitting these modes to operate on your router at the same time.

PIM-SM and PIM-SSM supports one-to-many applications by greatly simplifying the protocol mechanics for deployment ease.

- PIM in sparse mode operation is used in a multicast network when relatively few routers are involved in each multicast and these routers do not forward multicast packets for a group, unless there is an explicit request for the traffic.
- PIM in Source-Specific Multicast operation uses information found on source addresses for a multicast group provided by receivers and performs source filtering on traffic.
    - By default, PIM-SSM operates in the 232.0.0.0/8 multicast group range for IPv4 and ff3x::/32 (where x is any valid scope) in IPv6. To configure these values, use the **ssm range** command.
    - If SSM is deployed in a network already configured for PIM-SM, only the last-hop routers must be upgraded with Cisco IOS XR Software that supports the SSM feature.
    - No MSDP SA messages within the SSM range are accepted, generated, or forwarded.

Restrictions for PIM-SM and PIM-SSM

- Interoperability with SSM:

  PIM-SM operations within the SSM range of addresses change to PIM-SSM. In this mode, only PIM (S,G) join and prune messages are generated by the router, and no (S,G) RP shared tree or (*,G) shared tree messages are generated.

- **IGMP Version:**

  To report multicast memberships to neighboring multicast routers, hosts use IGMP, and all routers on the subnet must be configured with the same version of IGMP.

  A router running Cisco IOS XR Software does not automatically detect Version 1 systems. You must use the **version** command in router IGMP configuration submode to configure the IGMP version.

Configuring PIM-SSM for Use in a Legacy Multicast Deployment

Deploying PIM-SSM in legacy multicast-enabled networks can be problematic, because it requires changes to the multicast group management protocols used on the various devices attached to the network. Host, routers, and switches must all be upgraded in such cases.

To support legacy hosts and switches in a PIM-SSM deployment, this router offers a configurable mapping feature. Legacy group membership reports for groups in the SSM group range are mapped to a set of sources providing service for that set of (S,G) channels.

Restrictions for PIM-SSM Mapping

PIM-SSM mapping does not modify the SSM group range. Instead, the legacy devices must report group membership for desired groups in the SSM group range.

Configuration Example

To creconfigure PIM-SSM for use in a legacy multicast deployment, you must complete the following configurations:

1. Configuring a Set of Access Control Lists for Static SSM Mapping
2. Configuring a Set of Sources for SSM Mapping

Configuration

To configure a set of access control lists (ACLs) where each ACL describes a set of SSM groups to be mapped to one or more sources:

```
Router#configure
Tue Feb  4 05:15:56.544 UTC
Router(config)#ipv4 access-list mc3
Router(config-ipv4-acl)#permit 1 host 232.1.1.2 any
Router(config-ipv4-acl)#commit
Tue Feb  4 05:16:28.752 UTC
Router(config-ipv4-acl)#exit
Router(config)#exit
Router:ios#
```

To configure a set of sources mapped by SSM groups:

```
Router#configure
Router(config)#router igmp vrf vrf20
Router(config-igmp-vrf20)#ssm map static 232.1.1.1 mc2
Router(config-igmp-vrf20)#exit
Router(config-igmp)#commit
```

Configuring PIM Per Interface States Limit

The PIM Per Interface States Limit sets a limit on creating OLEs for the PIM interface. When the set limit is reached, the group is not accounted against this interface but the group can exist in PIM context for some other interface.

The following configuration sets a limit on the number of routes for which the given interface may be an outgoing interface as a result of receiving a PIM J/P message.

```
router pim | pim6 [vrf <vrfname>]
interface <ifname>
          maximum route-interfaces <max> [threshold <threshold>] [<acl>]
!
!
```

where,

`<ifname>` is the interface name

`<max>` is the maximum limit on the groups

`<threshold>` is the threshold number of groups at which point a syslog warning message will be issued

`<acl>` provides an option for selective accounting. If provided, only groups or (S,G)s that are permitted by the ACL is accounted against the limit. Groups or (S, G)s that are denied by the ACL are not accounted against the limit. If not provided, all the groups are accounted against the limit.

The following messages are displayed when the threshold limit is reached for PIM:

```
pim[1157]: %ROUTING-IPV4_PIM-4-CAC_STATE_THRESHOLD : The interface GigabitEthernet0_2_0_0
threshold number (4) allowed states has been reached.
State creation will soon be throttled. Configure a higher state limit value or take steps
to reduce the number of states.

pim[1157]: %ROUTING-IPV4_PIM-3-CAC_STATE_LIMIT : The interface GigabitEthernet0_2_0_0 maximum
 number (5) of allowed states has been reached.
State creation will not be allowed from here on. Configure a higher maximum value or take
steps to reduce the number of states
```

Limitations

- If a user has configured a maximum of 20 groups and has reached the maximum number of groups, then no more groups/OLEs can be created. If the user now decreases the maximum number to 10, the 20 joins/OLE will remain and a message of reaching the max is displayed. No more joins/OLE can be added at this point until it has reached less than 10.

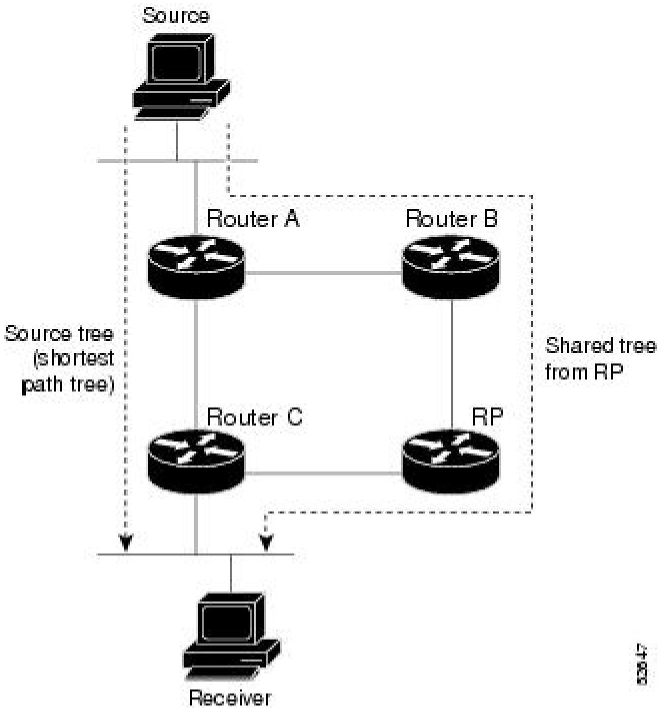
- If a user already has configured a maximum of 30 joins/OLEs and add a max of 20, the configuration occurs displaying a message that the max has been reached. No states will change but no more joins/OLEs can happen until the number is brought down below the maximum number of groups.

- Local interest joins are added, even if the limit has reached and is accounted for it.

PIM Shared Tree and Source Tree (Shortest Path Tree)

In PIM-SM, the rendezvous point (RP) is used to bridge sources sending data to a particular group with receivers sending joins for that group. In the initial setup of state, interested receivers receive data from senders to the group across a single data distribution tree rooted at the RP. This type of distribution tree is called a shared tree or rendezvous point tree (RPT) as illustrated in the below image. Data from senders is delivered to the RP for distribution to group members joined to the shared tree.

Figure 3: Shared Tree and Source Tree (Shortest Path Tree)



Unless the **spt-threshold infinity** command is configured, this initial state gives way as soon as traffic is received on the leaf routers (designated router closest to the host receivers). When the leaf router receives traffic from the RP on the RPT, the router initiates a switch to a data distribution tree rooted at the source sending traffic. This type of distribution tree is called a **shortest path tree** or **source tree**. By default, the Cisco IOS XR Software switches to a source tree when it receives the first data packet from a source.

The following process describes the move from shared tree to source tree in more detail:

1. Receiver joins a group; leaf Router C sends a join message toward RP.
2. RP puts link to Router C in its outgoing interface list.
3. Source sends data; Router A encapsulates data in Register and sends it to RP.
4. RP forwards data down the shared tree to Router C and sends a join message toward Source. At this point, data may arrive twice at the RP, once encapsulated and once natively.
5. When data arrives natively (unencapsulated) at RP, RP sends a register-stop message to Router A.

6. By default, receipt of the first data packet prompts Router C to send a join message toward Source.
7. When Router C receives data on (S,G), it sends a prune message for Source up the shared tree.
8. RP deletes the link to Router C from outgoing interface of (S,G). RP triggers a prune message toward Source.

Join and prune messages are sent for sources and RPs. They are sent hop by hop and are processed by each PIM router along the path to the source or RP. Register and register-stop messages are not sent hop by hop. They are exchanged using direct unicast communication between the designated router that is directly connected to a source and the RP for the group.

Tip The **spt-threshold infinity** command lets you configure the router so that it never switches to the shortest path tree (SPT).

Multicast-Intact

The multicast-intact feature provides the ability to run multicast routing (PIM) when Interior Gateway Protocol (IGP) shortcuts are configured and active on the router. Both Open Shortest Path First, version 2 (OSPFv2), and Intermediate System-to-Intermediate System (IS-IS) support the multicast-intact feature. Multiprotocol Label Switching Traffic Engineering (MPLS-TE) and IP multicast coexistence is supported in Cisco IOS XR Software by using the **mpls traffic-eng multicast-intact** IS-IS or OSPF router command. See the Routing Configuration Guide for Cisco 8000 Series Routers for information on configuring multicast intact using IS-IS and OSPF commands.

You can enable multicast-intact in the IGP when multicast routing protocols (PIM) are configured and IGP shortcuts are configured on the router. IGP shortcuts are MPLS tunnels that are exposed to IGP. The IGPs route the IP traffic over these tunnels to destinations that are downstream from the egress router of the tunnel (from an SPF perspective). PIM cannot use IGP shortcuts for propagating PIM joins because reverse path forwarding (RPF) cannot work across a unidirectional tunnel.

When you enable multicast-intact on an IGP, the IGP publishes a parallel or alternate set of equal-cost next-hops for use by PIM. These next-hops are called **mcast-intact next-hops**. The mcast-intact next-hops have the following attributes:

- They are guaranteed not to contain any IGP shortcuts.
- They are not used for unicast routing but are used only by PIM to look up an IPv4 next hop to a PIM source.
- They are not published to the Forwarding Information Base (FIB).
- When multicast-intact is enabled on an IGP, all IPv4 destinations that were learned through link-state advertisements are published with a set equal-cost mcast-intact next-hops to the RIB. This attribute applies even when the native next-hops have no IGP shortcuts.
- In IS-IS, the max-paths limit is applied by counting both the native and mcast-intact next-hops together. (In OSPFv2, the behavior is slightly different.)

Designated Routers

Cisco routers use PIM-SM to forward multicast traffic and follow an election process to select a designated router (DR) when there is more than one router on a LAN segment.

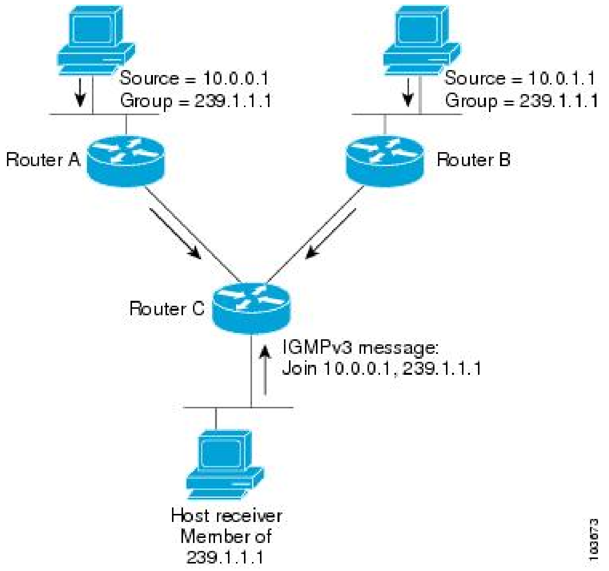
The designated router is responsible for sending PIM register and PIM join and prune messages toward the RP to inform it about host group membership.

If there are multiple PIM-SM routers on a LAN, a designated router must be elected to avoid duplicating multicast traffic for connected hosts. The PIM router with the highest IP address becomes the DR for the LAN unless you choose to force the DR election by use of the **dr-priority** command. The DR priority option allows you to specify the DR priority of each router on the LAN segment (default priority = 1) so that the router with the highest priority is elected as the DR. If all routers on the LAN segment have the same priority, the highest IP address is again used as the tiebreaker.

Note DR election process is required only on multi access LANs. The last-hop router directly connected to the host is the DR.

The figure "Designated Router Election on a Multiaccess Segment", below illustrates what happens on a multi access segment. Router A (10.0.0.253) and Router B (10.0.0.251) are connected to a common multi access Ethernet segment with Host A (10.0.0.1) as an active receiver for Group A. As the Explicit Join model is used, only Router A, operating as the DR, sends joins to the RP to construct the shared tree for Group A. If Router B were also permitted to send (*,G) joins to the RP, parallel paths would be created and Host A would receive duplicate multicast traffic. When Host A begins to source multicast traffic to the group, the DR's responsibility is to send register messages to the RP. Again, if both routers were assigned the responsibility, the RP would receive duplicate multicast packets.

Figure 4: Designated Router Election on a Multiaccess Segment



If the DR fails, the PIM-SM provides a way to detect the failure of Router A and to elect a failover DR. If the DR (Router A) were to become inoperable, Router B would detect this situation when its neighbor adjacency with Router A timed out. Because Router B has been hearing IGMP membership reports from Host A, it already has IGMP state for Group A on this interface and immediately sends a join to the RP when it becomes the new DR. This step reestablishes traffic flow down a new branch of the shared tree using Router B. Additionally, if Host A were sourcing traffic, Router B would initiate a new register process immediately after receiving the next multicast packet from Host A. This action would trigger the RP to join the SPT to Host A, using a new branch through Router B.

Note Two PIM routers are neighbors if there is a direct connection between them. To display your PIM neighbors, use the show pim neighbor command in EXEC mode.

- They are not used for unicast routing but are used only by PIM to look up an IPv4 next hop to a PIM source.
- They are not published to the Forwarding Information Base (FIB).
- When multicast-intact is enabled on an IGP, all IPv4 destinations that were learned through link-state advertisements are published with a set equal-cost mcast-intact next-hops to the RIB. This attribute applies even when the native next-hops have no IGP shortcuts.
- In IS-IS, the max-paths limit is applied by counting both the native and mcast-intact next-hops together. (In OSPFv2, the behavior is slightly different.)

Configuration Example

Configures the router to use DR priority 4 for TenGigE interface 0/0/0/1, but other interfaces will inherit DR priority 2:

```
Router#configure
Router(config)#router pim
Router(config-pim-default)#address-family ipv4
Router(config-pim-default-ipv4)#dr-priority 2
Router(config-pim-default-ipv4)#interface TenGigE0/0/0/1
Router(config-pim-ipv4-if)#dr-priority 4
Router(config-ipv4-acl)#commit
```

Running Configuration

```
Router#show run router pim
router pim
 address-family ipv4
  dr-priority 2
  spt-threshold infinity
  interface TenGigE 0/0/0/1
   dr-priority 4
   hello-interval 45
```

Verification

Verify if the parameters are set according to the configured values:

```
Router#show pim interface
PIM interfaces in VRF default
Address           Interface           PIM  Nbr   Hello  DR    DR Count Intvl  Prior
100.1.1.1        TenGigE0/0/0/1        on   1     45     4     this system
26.1.1.1         TenGigE0/0/0/26       on   1     30     2     this system
```

Rendezvous Points

When PIM is configured in sparse mode, you must choose one or more routers to operate as a rendezvous point (RP). A rendezvous point is a single common root placed at a chosen point of a shared distribution tree, as illustrated in PIM Shared Tree and Source Tree (Shortest Path Tree), on page 12. A rendezvous point can be either configured statically in each box or learned through a dynamic mechanism.

PIM DRs forward data from directly connected multicast sources to the rendezvous point for distribution down the shared tree. Data is forwarded to the rendezvous point in one of two ways:

- Encapsulated in register packets and unicast directly to the rendezvous point by the first-hop router operating as the DR.
- Multicast forwarded by the RPF forwarding algorithm, described in the Reverse-Path Forwarding, on page 19, if the rendezvous point has itself joined the source tree.

The rendezvous point address is used by first-hop routers to send PIM register messages on behalf of a host sending a packet to the group. The rendezvous point address is also used by last-hop routers to send PIM join and prune messages to the rendezvous point to inform it about group membership. You must configure the rendezvous point address on all routers (including the rendezvous point router).

A PIM router can be a rendezvous point for more than one group. Only one rendezvous point address can be used at a time within a PIM domain. The conditions specified by the access list determine for which groups the router is a rendezvous point.

You must manually configure a PIM router to function as a rendezvous point.

Configuration Example

The following example shows how to configure a static RP and allow backward compatibility:

```
RP/0/RP0/CPU0:ios#configure
Thu Jan 30 08:30:02.187 UTC
RP/0/RP0/CPU0:ios(config)#router pim
RP/0/RP0/CPU0:ios(config-pim)#old-register-checksum
RP/0/RP0/CPU0:ios(config-pim)#exit
RP/0/RP0/CPU0:ios(config)#ipv4 access-list rp-access
RP/0/RP0/CPU0:ios(config-ipv4-acl)#permit 239.1.1.0 0.0.255.255
RP/0/RP0/CPU0:ios(config-ipv4-acl)#commit
Thu Jan 30 08:31:22.679 UTC
RP/0/RP0/CPU0:ios(config-ipv4-acl)#
```

Auto-RP

Automatic route processing (Auto-RP) is a feature that automates the distribution of group-to-RP mappings in a PIM network. This feature has these benefits:

- It is easy to use multiple RPs within a network to serve different group ranges.
- It allows load splitting among different RPs.
- It facilitates the arrangement of RPs according to the location of group participants.
- It avoids inconsistent, manual RP configurations that might cause connectivity problems.

Multiple RPs can be used to serve different group ranges or to serve as hot backups for each other. To ensure that Auto-RP functions, configure routers as candidate RPs so that they can announce their interest in operating as an RP for certain group ranges. Additionally, a router must be designated as an RP-mapping agent that receives the RP-announcement messages from the candidate RPs, and arbitrates conflicts. The RP-mapping agent sends the consistent group-to-RP mappings to all remaining routers. Thus, all routers automatically determine which RP to use for the groups they support.

Tip By default, if a given group address is covered by group-to-RP mappings from both static RP configuration, and is discovered using Auto-RP or PIM BSR, the Auto-RP or PIM BSR range is preferred. To override the default, and use only the RP mapping, use the **rp-address override** keyword.

Note Auto-RP is not supported on VRF interfaces. Auto-RP Lite allows you to configure auto-RP on the CE router. It allows the PE router that has the VRF interface to relay auto-RP discovery, and announce messages across the core and eventually to the remote CE. Auto-RP is supported in only the IPv4 address family.

Configuring Example

```
Router#configure
Router(config)# router pim
Router(config-pim-ipv4)# auto-rp candidate-rp GigabitEthernet0/1/0/1 scope 31 group-list 2
 bidir
Router(config-pim-ipv4)# auto-rp mapping-agent GigabitEthernet0/1/0/1 scope 20
Router(config-pim-ipv4)# exit
Router(config)# ipv4 access-list 2
Router(config-ipv4-acl)# permit 239.1.1.1 0.0.0.0
Router(config-ipv4-acl)#commit
```

This example shows that Auto-RP messages are prevented from being sent out of the GigabitEthernet interface 0/3/0/0. It also shows that access list 111 is used by the Auto-RP candidate and access list 222 is used by the boundary command to contain traffic on GigabitEthernet interface 0/3/0/0.

```
ipv4 access-list 111
 10 permit 224.1.0.0 0.0.255.255
 20 permit 224.2.0.0 0.0.255.255
!
!Access list 111 is used by the Auto-RP candidate.
!
ipv4 access-list 222
 10 deny any host 224.0.1.39
 20 deny any host 224.0.1.40
!
!Access list 222 is used by the boundary command to contain traffic (on
GigabitEthernet0/3/0/0) that is sent to groups 224.0.1.39 and 224.0.1.40.
!
router pim
 auto-rp mapping-agent loopback 2 scope 32 interval 30
 auto-rp candidate-rp loopback 2 scope 15 group-list 111 interval 30
multicast-routing
 interface hundredGigE 0/0/0/25
 boundary 222
!
```

PIM Bootstrap Router

The PIM bootstrap router (BSR) provides a fault-tolerant, automated RP discovery and distribution mechanism.

To avoid a single point of failure, you can configure several candidate BSRs in a PIM domain. A BSR is elected among the candidate BSRs automatically.

Candidates use bootstrap messages to discover which BSR has the highest priority. The candidate with the highest priority sends an announcement to all PIM routers in the PIM domain that it is the BSR.

Routers that are configured as candidate RPs unicast to the BSR the group range for which they are responsible. The BSR includes this information in its bootstrap messages and disseminates it to all PIM routers in the domain. Based on this information, all routers are able to map multicast groups to specific RPs. As long as a router is receiving the bootstrap message, it has a current RP map.

Configuration Example

Configures the router as a candidate BSR with a hash mask length of 30:

```
Router# configure
Router:(config)# router pim
Router:(config-pim)# bsr candidate-bsr 10.0.0.1 hash-mask-len 30
Router:(config-ipv4-acl)#commit
```

Configures the router to advertise itself as a candidate rendezvous point to the BSR in its PIM domain. Access list number 4 specifies the prefix associated with the candidate rendezvous point address 10.2.1.1 . This rendezvous point is responsible for the groups with the prefix 239.

```
RP/0/RP0/CPU0:ios#configure
Thu Jan 30 08:03:47.952 UTC
RP/0/RP0/CPU0:ios(config)#router pim
RP/0/RP0/CPU0:ios(config-pim)#bsr candidate-bsr 10.0.0.1 hash-mask-len 30
RP/0/RP0/CPU0:ios(config-pim)#bsr candidate-rp 172.3.2.1 group-list 4 bidir
RP/0/RP0/CPU0:ios(config-pim)#interface fourHundredGigE 0/0/0/1
RP/0/RP0/CPU0:ios(config-pim-ipv4-if)# bsr-border
RP/0/RP0/CPU0:ios(config-pim-ipv4-if)#exit
RP/0/RP0/CPU0:ios(config-pim-default-ipv4)#exit
RP/0/RP0/CPU0:ios(config-pim)#exit
RP/0/RP0/CPU0:ios(config)#ipv4 access-list 4
RP/0/RP0/CPU0:ios(config-ipv4-acl)#permit 239.1.1.1 0.255.255.255
RP/0/RP0/CPU0:ios(config-ipv4-acl)#commit
Thu Jan 30 08:05:36.780 UTC
RP/0/RP0/CPU0:ios(config-ipv4-acl)#exit
RP/0/RP0/CPU0:ios(config)#exit
```

Running Configuration

```
RP/0/RP0/CPU0:ios#show running-config router pim
Thu Jan 30 08:08:06.568 UTC
router pim
 address-family ipv4
  interface FourHundredGigE0/0/0/1
   bsr-border
  !
  bsr candidate-bsr 10.0.0.1 hash-mask-len 30 priority 1
  bsr candidate-rp 172.3.2.1 group-list 4 priority 192 interval 60 bidir
 !
!
```

Verification

Displays PIM candidate RP information for the BSR.

```
RP/0/RP0/CPU0:ios#show pim bsr candidate-rp
Thu Jan 30 08:08:32.851 UTC
PIM BSR Candidate RP Info

Cand-RP          mode  scope priority uptime    group-list
172.3.2.1        BD    16    192      00:00:00  4
```

Displays PIM candidate election information for the BSR.

```
RP/0/RP0/CPU0:ios#show pim bsr election
Thu Jan 30 08:08:58.846 UTC
PIM BSR Election State

Cand/Elect-State        Uptime   BS-Timer BSR                C-BSR

Inactive/Accept-Any     00:00:00 00:00:00 0.0.0.0 [0, 0]     10.0.0.1 [1, 30]
```

Displays PIM RP cache information for the BSR.

```
RP/0/RP0/CPU0:ios#show pim bsr rp-cache
Thu Jan 30 08:09:44.901 UTC
PIM BSR Candidate RP Cache
```

Displays group-to-PIM mode mapping.

```
RP/0/RP0/CPU0:ios#show pim ipv4 group-map
Thu Jan 30 08:10:14.793 UTC
No ranges found.
```

Reverse-Path Forwarding

Reverse-path forwarding (RPF) is an algorithm used for forwarding multicast datagrams. It functions as follows:

- If a router receives a datagram on an interface it uses to send unicast packets to the source, the packet has arrived on the RPF interface.
- If the packet arrives on the RPF interface, a router forwards the packet out the interfaces present in the outgoing interface list of a multicast routing table entry.
- If the packet does not arrive on the RPF interface, the packet is silently discarded to prevent loops.

PIM uses both source trees and RP-rooted shared trees to forward datagrams; the RPF check is performed differently for each, as follows:

- If a PIM router has an (S,G) entry present in the multicast routing table (a source-tree state), the router performs the RPF check against the IP address of the source for the multicast packet.
- If a PIM router has no explicit source-tree state, this is considered a shared-tree state. The router performs the RPF check on the address of the RP, which is known when members join the group.

Sparse-mode PIM uses the RPF lookup function to determine where it needs to send joins and prunes. (S,G) joins (which are source-tree states) are sent toward the source. (*,G) joins (which are shared-tree states) are sent toward the RP.

Multicast Non-Stop Routing

Multicast Non-Stop Routing (NSR) enables the router to synchronize the multicast routing tables on both the active and standby RSPs so that during an HA scenario like an RSP failover there is no loss of multicast data. Multicast NSR is enabled through the multicast processes being hot standby. Multicast NSR supports both Zero Packet Loss (ZPL) and Zero Topology Loss (ZTL). With Multicast NSR, there is less CPU churn and no multicast session flaps during a failover event.

Multicast NSR is enabled by default, however, if any unsupported features like BNG or Snooping are configured, Multicast performs Non-Stop Forwarding (NSF) functionality during failover events. When Multicast NSR is enabled, multicast routing state is synchronized between the active and standby RSPs. Once the synchronization occurs, each of the multicast processes signal the NSR readiness to the system. For the multicast processes to support NSR, the processes must be hot standby compliant. That is, the processes on active and standby RSPs both have to be in synchronization at all times. The active RSP receives packets from the network and makes local decisions while the standby receives packet from the network and synchronizes it with the active RSPs for all the local decisions. Once the state is determined, a check is performed to verify if the states are synchronized. If the states are synchronized, a signal in the form NSR_READY is conveyed to the NSR system.

With NSR, in the case of a failover event, routing changes are updated to the forwarding plane immediately. With NSF, there is an NSF hold time delay before routing changes can be updated.

Non-Supported Features

The following features are unsupported on NG NSR:

- IGMP and MLD Snooping
- BNG

Configuration Example

```
RP/0/RP0/CPU0:ios#configure
Fri Feb  7 08:53:51.603 UTC
RP/0/RP0/CPU0:ios(config)#router pim address-family ipv4
RP/0/RP0/CPU0:ios(config-pim-default-ipv4)#nsf lifetime 30
RP/0/RP0/CPU0:ios(config-pim-default-ipv4)#exit
RP/0/RP0/CPU0:ios(config-pim)#router igmp
RP/0/RP0/CPU0:ios(config-igmp)#nsf lifetime 30
RP/0/RP0/CPU0:ios(config-igmp)#commit
Fri Feb  7 08:54:45.747 UTC
RP/0/RP0/CPU0:ios(config-igmp)#exit
RP/0/RP0/CPU0:ios(config)#exit
RP/0/RP0/CPU0:ios#show igmp nsf
Fri Feb  7 08:55:02.046 UTC
IGMP Non-Stop Forwarding Status:
Multicast routing state: Normal
   NSF Lifetime:       00:00:30

RP/0/RP0/CPU0:ios#show mfib nsf
Fri Feb  7 08:55:12.462 UTC
IP MFWD Non-Stop Forwarding Status:
 NSF Lifetime:          00:15:00

On node 0/RP0/CPU0 :
Multicast routing state: Normal

RP/0/RP0/CPU0:ios#show mrib nsf
```

```
Fri Feb  7 08:55:24.228 UTC
IP MRIB Non-Stop Forwarding Status:
Multicast routing state: Normal
   NSF Lifetime:       00:01:30
RP/0/RP0/CPU0:ios#show pim nsf
Fri Feb  7 08:55:33.499 UTC
IP PIM Non-Stop Forwarding Status:
Multicast routing state: Normal
   NSF Lifetime:       00:00:30
RP/0/RP0/CPU0:ios#
```

Verification

Verify the state of NSF operation in IGMP.

```
RP/0/RP0/CPU0:ios#show igmp nsf
Fri Feb  7 08:55:02.046 UTC
IGMP Non-Stop Forwarding Status:
Multicast routing state: Normal
   NSF Lifetime:       00:00:30
```

Verify the state of NSF operation for the MFIB line cards.

```
RP/0/RP0/CPU0:ios#show mfib nsf
Fri Feb  7 08:55:12.462 UTC
IP MFWD Non-Stop Forwarding Status:
 NSF Lifetime:          00:15:00

On node 0/RP0/CPU0 :
Multicast routing state: Normal
```

Verify the state of NSF operation in the MRIB.

```
RP/0/RP0/CPU0:ios#show mrib nsf
Fri Feb  7 08:55:24.228 UTC
IP MRIB Non-Stop Forwarding Status:
Multicast routing state: Normal
   NSF Lifetime:       00:01:30
```

Verify the state of NSF operation for PIM.

```
RP/0/RP0/CPU0:ios#show pim nsf
Fri Feb  7 08:55:33.499 UTC
IP PIM Non-Stop Forwarding Status:
Multicast routing state: Normal
   NSF Lifetime:       00:00:30
RP/0/RP0/CPU0:ios#
```

Failure Scenarios in NSR

If a switchover occurs before all multicast processes issue an NSR_READY signal, the proceedings revert back to the existing NSF behavior. Also, on receiving the GO_ACTIVE signal from the multicast processes, the following events occur in processes that have not signaled NSR_READY:

1. IGMP starts the NSF timer for one minute.
2. PIM starts the NSF timer for two minutes.
3. MSDP resets all peer sessions that are not synchronized.

Multicast-Only Fast Reroute

Multicast-Only Fast Reroute (MoFRR) allows fast reroute for multicast traffic on a multicast router. MoFRR minimizes packet loss in a network when node or link failures occur (at the topology merge point). It works by making simple enhancements to multicast routing protocols.

MoFRR involves transmitting a multicast join message from a receiver towards a source on a primary path and transmitting a secondary multicast join message from the receiver towards the source on a backup path. Data packets are received from the primary and secondary paths. The redundant packets are discarded at topology merge points with the help of Reverse Path Forwarding (RPF) checks. When a failure is detected on the primary path, the repair occurs locally by changing the interface on which packets are accepted to the secondary interface, thus improving the convergence times in the event of a node or link failure on the primary path.

RIB-Based MoFRR

RIB-based MoFRR enables PIM to perform a fast convergence of specified routes or flows upon detecting a failure on any of the multiple equal-cost paths between the router and the source.

Restriction

Only RIB-Based MoFRR is supported.

Configuring RIB-Based MoFRR

When a failure is detected on one of multiple equal-cost paths between the router and the source, perform a fast convergence (MoFRR) of specified routes or flows using the **mofrr rib** command.

Configuration example

```
Router(config)# router pim
Router(pim)# mofrr rib route-list
```

Running Configuration

```
Router#show running-config router pim
router pim
 address-family ipv4
  mofrr
   rib route-list
  !
 !
!
```

Note The *route-list* keyword is a previously defined IPv4 access-list which has the source and group information to match the specific multicast route that needs MoFRR enabled. Please refer to *Modular QoS Configuration Guide* on how to configure an access-list.

Verify RIB-Based MoFRR is Enabled

Verify that you have successfully configured RIB-Based MoFRR using the following CLI command. The command output shows the MoFRR-RIB flag, the primary and secondary RPF interfaces are enabled for MoFRR.

Sample Configuration:

```
Router# show pim topology src-ip-address/ grp-address detail
```

Verification Example:

```
Router# show pim topology 232.1.1.1 detail

IP PIM Multicast Topology Table

(100.1.1.1,232.1.1.1)SPT SSM Up: 00:00:35
JP: Join(00:00:14) RPF: FourHundredGigE0/0/0/2,100.1.1.1* MoFRR-RIB, Flags:
Up: MT clr (00:00:00) MDT: JoinSend N, Cache N/N/N, Misc (0x0,0/0)
Cache: Add 00:00:00, Rem 00:00:00. MT Cnt: Set 0, Unset 0. Joins sent 0
MDT-ifh 0x0/0x0, MT Slot none/ none
RPF-redirect BW usage: 0, Flags: 0x0, ObjID: 0x0
c-multicast-routing: PIM* BGPJP: 1w0d
RPF Table: IPv4-Unicast-default
RPF Secondary: FourHundredGigE0/0/0/1,100.1.1.1
  FourHundredGigE0/0/0/0      00:00:35  fwd LI LH
```

Note To ensure that MoFRR yields better convergence, prioritize the multicast source routes using IGP protocol for RPF check. Thus ensuring the routes are always taken first for SPF calculation in case of path changes.

```
Router(config)# router isis isp
Router(config-isis)#)#address-family ipv4 unicast
Router(config-isis-af)#spf prefix-priority critical ISIS-CRIT

Router#show running-config ipv4 prefix-list ISIS-CRIT
Wed May 27 01:26:58.653 PDT
ipv4 prefix-list ISIS-CRIT
10 permit 192.168.224.60/32 ge 32
11 permit 192.92.1.2/32 ge 32
```

Multicast Source Discovery Protocol

Multicast Source Discovery Protocol (MSDP) is a mechanism to connect multiple PIM sparse-mode domains. MSDP allows multicast sources for a group to be known to all rendezvous points (RPs) in different domains. Each PIM-SM domain uses its own RPs and need not depend on RPs in other domains.

An RP in a PIM-SM domain has MSDP peering relationships with MSDP-enabled routers in other domains. Each peering relationship occurs over a TCP connection, which is maintained by the underlying routing system.

MSDP speakers exchange messages called Source Active (SA) messages. When an RP learns about a local active source, typically through a PIM register message, the MSDP process encapsulates the register in an SA message and forwards the information to its peers. The message contains the source and group information for the multicast flow, as well as any encapsulated data. If a neighboring RP has local joiners for the multicast group, the RP installs the S, G route, forwards the encapsulated data contained in the SA message, and sends PIM joins back towards the source. This process describes how a multicast path can be built between domains.

Note Although you should configure BGP or Multiprotocol BGP for optimal MSDP interdomain operation, this is not considered necessary in the Cisco IOS XR Software implementation. For information about how BGP or Multiprotocol BGP may be used with MSDP, see the MSDP RPF rules listed in the Multicast Source Discovery Protocol (MSDP), Internet Engineering Task Force (IETF) Internet draft.

Restriction

Loop-Free Alternative Fast Reroute is not supported.

MSDP Configuration Submode

When you issue the **router msdp** command, the CLI prompt changes to “config-msdp,” indicating that you have entered router MSDP configuration submode.

Multicast Nonstop Forwarding

The Cisco IOS XR Software nonstop forwarding (NSF) feature for multicast enhances high availability (HA) of multicast packet forwarding. NSF prevents hardware or software failures on the control plane from disrupting the forwarding of existing packet flows through the router.

The contents of the Multicast Forwarding Information Base (MFIB) are frozen during a control plane failure. Subsequently, PIM attempts to recover normal protocol processing and state before the neighboring routers time out the PIM hello neighbor adjacency for the problematic router. This behavior prevents the NSF-capable router from being transferred to neighbors that will otherwise detect the failure through the timed-out adjacency. Routes in MFIB are marked as stale after entering NSF, and traffic continues to be forwarded (based on those routes) until NSF completion. On completion, MRIB notifies MFIB and MFIB performs a mark-and-sweep to synchronize MFIB with the current MRIB route information.

Multicast Configuration Submodes

Cisco IOS XR Software moves control plane CLI configurations to protocol-specific submodes to provide mechanisms for enabling, disabling, and configuring multicast features on a large number of interfaces.

Cisco IOS XR Software allows you to issue most commands available under submodes as one single command string from the global or XR config mode.

For example, the **ssm** command could be executed from the PIM configuration submode like this:

```
RP/0/RSP0/CPU0:router(config)# router pim
RP/0/RSP0/CPU0:router(config-pim)# address-family ipv4
RP/0/RSP0/CPU0:router(config-pim-default-ipv4)# ssm range
```

Alternatively, you could issue the same command from the global or XR config mode like this:

```
RP/0/RSP0/CPU0:router(config)# router pim ssm range
```

The following multicast protocol-specific submodes are available through these configuration submodes:

Multicast-Routing Configuration Submode

Basic multicast services start automatically without any explicit configuration required. The following multicast services are started automatically:

- MFWD
- MRIB
- PIM
- IGMP

Other multicast services require explicit configuration before they start. For example, to start the MSDP process, you must enter the **router msdp** command and explicitly configure it.

When you issue the **multicast-routing ipv4 or multicast-routing ipv6** command, all default multicast components (PIM, IGMP, MLD, MFWD, and MRIB) are automatically started, and the CLI prompt changes to “config-mcast-ipv4” or “config-mcast-ipv6”, indicating that you have entered multicast-routing configuration submode.

PIM Configuration Submode

When you issue the **router pim** command, the CLI prompt changes to “config-pim-ipv4,” indicating that you have entered the default pim address-family configuration submode.

To enter pim address-family configuration submode for IPv6, type the **address-family ipv6** keyword together with the **router pim** command before pressing Enter.

IGMP Configuration Submode

When you issue the **router igmp** command, the CLI prompt changes to “config-igmp,” indicating that you have entered IGMP configuration submode.

MLD Configuration Submode

When you issue the **router mld** command, the CLI prompt changes to “config-mld,” indicating that you have entered MLD configuration submode.

MSDP Configuration Submode

When you issue the **router msdp** command, the CLI prompt changes to “config-msdp,” indicating that you have entered router MSDP configuration submode.

Understanding Interface Configuration Inheritance

Cisco IOS XR Software allows you to configure commands for a large number of interfaces by applying command configuration within a multicast routing submode that could be inherited by all interfaces. To override the inheritance mechanism, you can enter interface configuration submode and explicitly enter a different command parameter.

For example, in the following configuration you could quickly specify (under router PIM configuration mode) that all existing and new PIM interfaces on your router will use the hello interval parameter of 420 seconds. However, Packet-over-SONET/SDH (POS) interface 0/1/0/1 overrides the global interface configuration and uses the hello interval time of 210 seconds.

```
RP/0/RP0/CPU0:router(config)# router pim
RP/0/RP0/CPU0:router(config-pim-default-ipv4)# hello-interval 420
RP/0/RP0/CPU0:router(config-pim-default-ipv4)# interface pos 0/1/0/1
RP/0/RP0/CPU0:router(config-pim-ipv4-if)# hello-interval 210
```

The following is a listing of commands (specified under the appropriate router submode) that use the inheritance mechanism:

```
router pim
  dr-priority
  hello-interval
  join-prune-interval

multicast-routing
  version
  query-interval
  query-max-response-time
  explicit-tracking
router mld
  interface all disable
  version
  query-interval
  query-max-response-time
  explicit-tracking

router msdp
  connect-source
  sa-filter
  filter-sa-request list
  remote-as
  ttl-threshold
```

Understanding Interface Configuration Inheritance Disablement

As stated elsewhere, Cisco IOS XR Software allows you to configure multiple interfaces by applying configurations within a multicast routing submode that can be inherited by all interfaces.

To override the inheritance feature on specific interfaces or on all interfaces, you can enter the address-family IPv4 or IPv6 submode of multicast routing configuration mode, and enter the **interface-inheritance disable** command together with the **interface** *type interface-path-id* or **interface all** command. This causes PIM or IGMP protocols to disallow multicast routing and to allow only multicast forwarding on those interfaces specified. However, routing can still be explicitly enabled on specified individual interfaces.

The following configuration disables multicast routing interface inheritance under PIM and IGMP generally, although forwarding enablement continues. The example shows interface enablement under IGMP of GigabitEthernet 0/6/0/3:

```
RP/0/RP0/CPU0:router# multicast-routing address-family ipv4
```

```
RP/0/RP0/CPU0:router(config-mcast-default-ipv4)# interface all enable
RP/0/RP0/CPU0:router(config-mcast-default-ipv4)# interface-inheritance disable

!

!
RP/0/RP0/CPU0:router(config)# router igmp
RP/0/RP0/CPU0:router(config-igmp)# vrf default
RP/0/RP0/CPU0:router(config-igmp)# interface GigabitEthernet0/6/0/0
RP/0/RP0/CPU0:router(config-igmp-name-if)# router enable
```

For related information, see Understanding Enabling and Disabling Interfaces, on page 27.

Understanding Enabling and Disabling Interfaces

When the Cisco IOS XR Software multicast routing feature is configured on your router, by default, no interfaces are enabled.

To enable multicast routing and protocols on a single interface or multiple interfaces, you must explicitly enable interfaces using the **interface** command in multicast routing configuration mode.

To set up multicast routing on all interfaces, enter the **interface all** command in multicast routing configuration mode. For any interface to be fully enabled for multicast routing, it must be enabled specifically (or be default) in multicast routing configuration mode, and it must not be disabled in the PIM and IGMP/MLD configuration modes.

For example, in the following configuration, all interfaces are explicitly configured from multicast routing configuration submode:

```
RP/0/RP0/CPU0:router(config)# multicast-routing
RP/0/RP0/CPU0:router(config-mcast)# interface all enable
```

To disable an interface that was globally configured from the multicast routing configuration submode, enter interface configuration submode, as illustrated in the following example:

```
RP/0/RP0/CPU0:router(config-mcast)# interface GigabitEthernet0pos 0/1/0/0
RP/0/RP0/CPU0:router(config-mcast-default-ipv4-if)# disable
```

Controlling Source Information on MSDP Peer Routers

Your MSDP peer router can be customized to control source information that is originated, forwarded, received, cached, and encapsulated.

When originating Source-Active (SA) messages, you can control to whom you will originate source information, based on the source that is requesting information.

When forwarding SA messages you can do the following:

- Filter all source/group pairs
- Specify an extended access list to pass only certain source/group pairs

- Filter based on match criteria in a route map

When receiving SA messages you can do the following:

- Filter all incoming SA messages from an MSDP peer
- Specify an extended access list to pass certain source/group pairs
- Filter based on match criteria in a route map

In addition, you can use time to live (TTL) to control what data is encapsulated in the first SA message for every source. For example, you could limit internal traffic to a TTL of eight hops. If you want other groups to go to external locations, you send those packets with a TTL greater than eight hops.

By default, MSDP automatically sends SA messages to peers when a new member joins a group and wants to receive multicast traffic. You are no longer required to configure an SA request to a specified MSDP peer.

Configuration Example

```
Router#configure
Router(config)# router msdp
Router(config-msdp)# sa-filter out router.cisco.com list 100
Router(config-msdp)# cache-sa-state 100
Router(config-msdp)# ttl-threshold 8
Router(config-msdp)# exit
Router(config)# ipv4 access-list 100 20 permit 239.1.1.1 0.0.0.0
Router(config)# commit
```

Multicast Routing Information Base

The Multicast Routing Information Base (MRIB) is a protocol-independent multicast routing table that describes a logical network in which one or more multicast routing protocols are running. The tables contain generic multicast routes installed by individual multicast routing protocols. There is an MRIB for every logical network (VPN) in which the router is configured. MRIBs do not redistribute routes among multicast routing protocols; they select the preferred multicast route from comparable ones, and they notify their clients of changes in selected attributes of any multicast route.

Multicast Forwarding Information Base

Table 2: Feature History Table

Feature Name	Release Information	Feature Description
MFIB scale enhancement	Release 7.3.15	This feature allows you to increase the route-scale for IPv4 SSM from 64K to 120K using the hw-module multicast route-scale command.

Multicast Forwarding Information Base (MFIB) is a protocol-independent multicast forwarding system that contains unique multicast forwarding entries for each source or group pair known in a given network. There is a separate MFIB for every logical network (VPN) in which the router is configured. Each MFIB entry resolves a given source or group pair to an incoming interface (IIF) for reverse-path forwarding (RPF) checking and an outgoing interface list (olist) for multicast forwarding.

Enable 120K Route-Scale for IPv4 SSM

Use the **hw-module multicast route-scale** command to enable the 120K route-scale for IPv4 SSM. Note that IPv6 supports only 64K route-scale.

```
Router# configure
Router(config)# hw-module multicast route-scale
```

See **hw-module multicast route-scale** command under the *Multicast Routing Forwarding Commands* chapter in *Multicast Command Reference for Cisco 8000 Series Routers*.

Note For the new route-scale to take effect, you must reload all the nodes on your router using the **reload** command.

```
Router# reload location all
```

MSDP MD5 Password Authentication

MSDP MD5 password authentication is an enhancement to support Message Digest 5 (MD5) signature protection on a TCP connection between two Multicast Source Discovery Protocol (MSDP) peers. This feature provides added security by protecting MSDP against the threat of spoofed TCP segments being introduced into the TCP connection stream.

MSDP MD5 password authentication verifies each segment sent on the TCP connection between MSDP peers. The **password clear** command is used to enable MD5 authentication for TCP connections between two MSDP peers. When MD5 authentication is enabled between two MSDP peers, each segment sent on the TCP connection between the peers is verified.

Note MSDP MD5 authentication must be configured with the same password on both MSDP peers to enable the connection between them. The 'password encrypted' command is used only for applying the stored running configuration. Once you configure the MSDP MD5 authentication, you can restore the configuration using this command.

MSDP MD5 password authentication uses an industry-standard MD5 algorithm for improved reliability and security.

Configuration Example

```
Router#configure
Router(config)#router msdp
Router(config-msdp)#peer 10.0.5.4
Router(config-msdp-peer)#password encrypted a34bi5m
Router(config-msdp-peer)#commit
```

Label Switch Multicast

Label Switch Multicast (LSM) is MPLS technology extensions to support multicast using label encapsulation. Next-generation MVPN is based on Multicast Label Distribution Protocol (mLDP), which can be used to build P2MP and MP2MP LSPs through a MPLS network. These LSPs can be used for transporting both IPv4 and IPv6 multicast packets, either in the global table or VPN context. mLDP is supported on core.

When router is positioned as the core router running mLDP, it only supports the Profiles 5, 6, 7, 12, 14, and 17.

When router is positioned as the core router running P2MP-TE core, it only supports the Profiles 8 and 10.

Benefits of LSM MLDP based MVPN

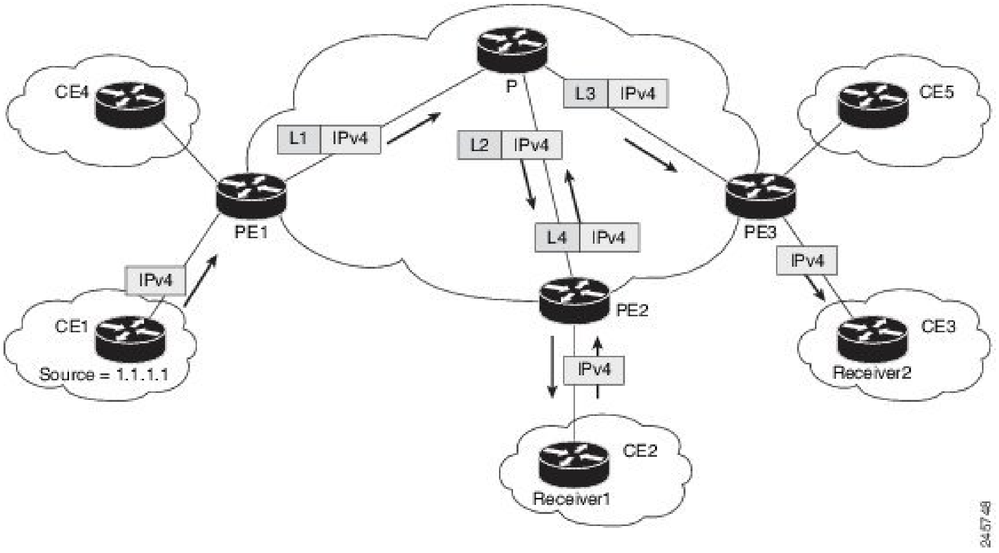
LSM provides these benefits when compared to GRE core tunnels that are currently used to transport customer traffic in the core:

- It leverages the MPLS infrastructure for transporting IP multicast packets, providing a common data plane for unicast and multicast.
- It eliminates the complexity associated PIM.

Configuring MLDP MVPN

The MLDP MVPN configuration enables IPv4 and IPV6 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 and edge network. For MLDP MVPN configuration to work, ensure that the global MPLS MLDP configuration is enabled. To configure MVPN extranet support, configure the source multicast VPN Routing and Forwarding (mVRF) on the receiver Provider Edge (PE) router or configure the receiver mVRF on the source PE. MLDP MVPN is supported for both intranet and extranet.

Figure 5: MLDP based MPLS Network on Core Routers



Packet Flow in mLDP-based Multicast VPN

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 Provider Edge1 (PE1) router imposes a label on the incoming multicast packet and replicates the labeled 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 (edge routers), the label is removed and the IP multicast packet is replicated onto the VRF interface. Basically, the packets are encapsulated at headend and decapsulated at tailend on the PE routers.

Multicast Label Distribution Protocol (MLDP) as Core Router

Multicast Label Distribution Protocol (MLDP) provides extensions to the Label Distribution Protocol (LDP) for the setup of point-to-multipoint (P2MP) and multipoint-to-multipoint (MP2MP) Label Switched Paths (LSPs) in Multiprotocol Label Switching (MPLS) networks.

MLDP eliminates the use of native multicast PIM to transport multicast packets across the core. In MLDP multicast traffic is label switched across the core. This saves a lot of control plane processing effort.

Restrictions

MLDP is supported only on core routers.

Configuration

For more information about MLDP configuration, see the *Enabling MLDP* section in the *Implementing MPLS Label Distribution Protocol* chapter of the *MPLS Configuration Guide for Cisco 8000 Routers.*

Point-to-Multipoint Traffic Engineering Label-Switched Multicast

IP multicast was traditionally used for IPTV broadcasting and content delivery services. Point-to-Multipoint (P2MP) Traffic-Engineering is fast replacing the IP multicast technique because of the various advantages of MPLS-TE, such as:

- Fast re-routing (FRR) and restoration in case of link/ node failure
- Bandwidth guarantee

Configuration

For more information about Point-to-Multipoint Traffic Engineering Label-Switched Multicast configuration, see the *Point-to-Multipoint Traffic-Engineering* section in the *Implementing MPLS Traffic Engineering* chapter of the *MPLS Configuration Guide for Cisco 8000 Routers.*

mLDP Loop-Free Alternative Fast Reroute

Table 3: Feature History Table

Feature Name	Release Information	Feature Description
mLDP Loop-Free Alternative Fast Reroute	Release 7.3.15	When this feature is enabled, mLDP relies on the Loop-Free Alternative algorithm to calculate the primary and backup, which is also referred as fast re-route path. During the event of a link failure, the router uses this precomputed backup path to send the multicast traffic. The fast switchover helps to reduce multicast traffic loss and the switchover time is less than 50 milliseconds.

Background

Generally, in a network, a network topology change, caused by a failure in a network, results in a loss of connectivity until the control plane convergence is complete. There can be various levels of loss of connectivity depending on the performance of the control plane, fast convergence tuning, and leveraged technologies of the control plane on each node in the network.

The amount of loss of connectivity impacts some loss-sensitive applications, which have severe fault tolerance (typically of the order of hundreds of milliseconds and up to a few seconds). In order to ensure that the loss of connectivity conforms to such applications, a technology implementation for data plane convergence is essential. **Fast Reroute (FRR)** is one of such technologies that is primarily applicable to the network core.

With the FRR solution, at each node, the backup path is pre-computed, and the traffic is routed through this backup path. As a result, the reaction to failure is local; immediate propagation of the failure and subsequent processing on to other nodes is not required. With FRR, if the failure is detected quickly, a loss of connectivity as low as 10s of milliseconds is achieved.

Loop-Free Alternative Fast Reroute

IP Loop Free Alternative FRR is a mechanism that enables a router to rapidly switch traffic to a pre-computed or a pre-programmed **loop-free alternative (LFA)** path, which is Data Plane Convergence, following either an adjacent link and node failure, or an adjacent link or node failure in both IP and LDP networks. The LFA path is used to switch traffic till the router installs the new primary next-hops based upon the changed network topology, which is Control Plane Convergence.

The goal of LFA FRR is to reduce the loss of connectivity to tens of milliseconds by using a pre-computed alternative next-hop, in the case where the selected primary next-hop fails.

There are two approaches to computing LFA paths:

- **Link-based (per-link):** In link-based LFA paths, all prefixes reachable through the primary (protected) link share the same backup information. This means that the whole set of prefixes sharing the same primary also shares the repair and FRR ability.

- **Prefix-based (per-prefix):** Prefix-based LFAs allow computing backup information for each prefix. This means that the repair and backup information computed for a given prefix using prefix-based LFA may be different from the one computed by link-based LFA.

  Protection against a node failure by rerouting traffic around the failed node (Node-protection support) is available with per-prefix LFA FRR on ISIS currently. It uses a tie-breaker mechanism in the code to select node-protecting backup paths.

The per-prefix LFA approach is preferred to the per-link LFA approach for the following reasons:

- Better node failure resistance.
- Better coverage: Each prefix is analyzed independently.
- Better capacity planning: Each flow is backed up on its own optimized shortest path.

mLDP LFA FRR

The point-to-point physical or bundle interface FRR mechanism is supported on mLDP. FRR with LFA backup is supported on mLDP. When there is a link failure, mLDP automatically sets up and chooses the backup path.

With this implementation, you must configure the physical or bundle interface for unicast traffic, so that the mLDP can act as an mLDP FRR.

LFA FRR support on mLDP is a per-prefix backup mechanism. As part of computing the LFA backup for a remote IP, the LFA backup paths for the loopback address of the downstream intermediate nodes are also computed. MLDP uses this small subset of information, by using the loopback address of the peer to compute the LFA backup path.

Note Both IPv4 and IPv6 traffic is supported on the mLDP LFA FRR solution.

MLDP LFA FRR - Features

- Supports both IPv4 and IPv6 multicast traffic carried by MLDP label.
- Supports all MLDP profiles and behaves as a MLDP core router only.
- Supports both ISIS and OSPF routing protocols

Advantages of LFA FRR

The following are the advantages of the LFA FRR solution:

- The backup path for the traffic flow is pre-computed, so that it help in faster convergence.
- Reaction to failure is local, an immediate propagation and processing of failure on to other nodes is not required.
- If the failure is detected in time, the loss of connectivity of up to 50 milliseconds can be achieved.
- The mechanism is locally significant and does not impact the Interior Gateway Protocol (IGP) communication channel.

- LFA next-hop can protect against:
  - a single link failure
  - failure of one of more links within a shared risk link group (SRLG)
  - any combination of the above
- Supports switchover time of less than 50 milliseconds.
- Supports switchover time to be independent of the number of multicast routes that has to be switched over.

Limitations of LFA FRR

The following are some of the known limitations of the LFA FRR solution:

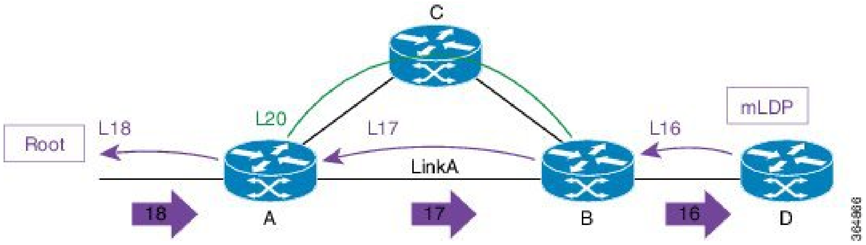
- When a failure that is more extensive than that which the alternate path was intended to protect occurs, there is the possibility of temporarily looping traffic (micro looping) until Control Plane Convergence.
- MLDP LFA FRR is supported only on MLDP core routers and not on the MLDP edge router.

MLDP LFA FRR - Workflow

To enable FRR for mLDP over physical or bundle interfaces, LDP session-protection feature has to be configured. The sequence of events that occur in an mLDP LFA FRR scenario is explained with the following example:

1. Step 1: MLDP LFA FRR - Initial Setup

   Figure 6: MLDP LFA FRR - Setup

   a. In this set up, Router A is the source provider edge router, and the next Hop is Router B. The primary path is Router A -> Router B - > Router D, and the backup path is from Router A -> Router C -> Router B -> Router D. The backup path is pre-computed by IGP through LFA prefix-based selection.

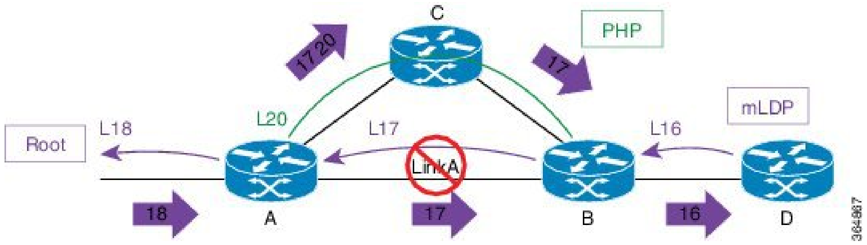
   b. Backup paths are configured for Link A or auto-tunnels are enabled.

   c. MLDP LSP is built from D, B, and A towards the root.

   d. Router A installs a downstream forwarding replication over link A to Router B. This entry has both the primary interface (Link A) and the backup paths programmed.

2. Step 2: Link Failure

Figure 7: Link Failure



- When a failure occurs on Link A:

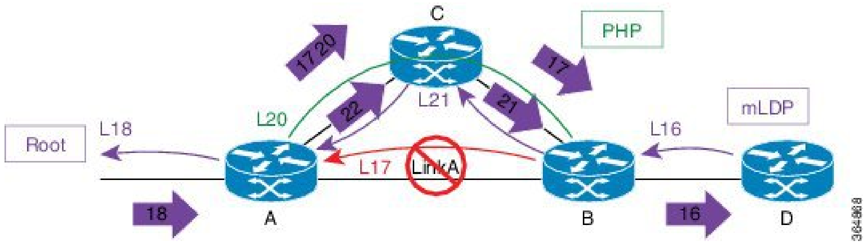
  a. Traffic over link A is rerouted over the backup path with same MLDP Label 17 (inner label) , plus a unicast label 20 (outer label) towards mid Router C.

  b. Router C performs penultimate hop popping (PHP) and removes the outer label 20.

  c. Router B receives the mLDP packets with label 17 and forwards to Router D.

3. Step 3: Re-optimization

Figure 8: Re-optimization



During re-optimization:

a. mLDP is notified that the root is reachable through Router C, and mLDP converges. With this, a new mLDP path is built to router A through Router C.

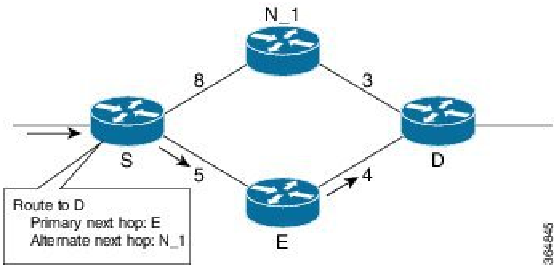
b. Router A forwards packets natively with old label 17 and also new label 22.

c. Router B drops traffic carried from new label 22 and forwards traffic with label 17.

MLDP LFA FRR - Behavior

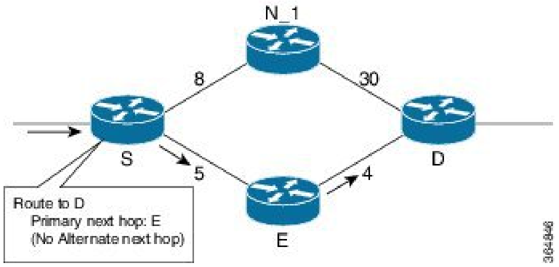
In the following scenarios, S is source router, D is the destination router, E is primary next hop, and N_1 is the alternative next hop.

Figure 9: LFA FRR Behavior - LFA Available



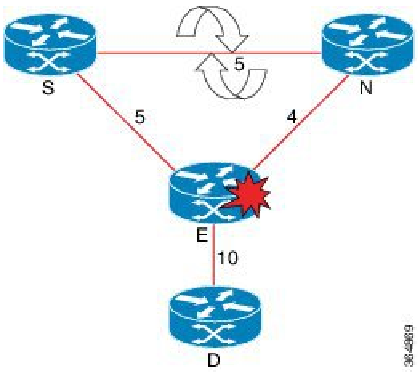
With LFA FRR, the source router S calculates an alternative next hop N_1 to forward traffic towards the destination router D through N_1, and installs N_1 as a the alternative next hop. On detecting the link failure between routers S and E, router S stops forwarding traffic destined for router D towards E through the failed link; instead it forwards the traffic to a pre-computed alternate next hop N_1, until a new SPF is run and the results are installed.

Figure 10: LFA FRR Behavior - LFA Not Available



In the above scenario, if the link cost between the next hop N_1 and the destination router D is increased to 30, then the next hop N_1 would no longer be a loop-free alternative. (The cost of the path, from the next hop N_1 to the destination D through the source S, would be 17, while the cost from the next hop N_1 directly to destination D would be 30). Thus, the existence of a LFA next hop is dependent on the topology and the nature of the failure, for which the alternative is calculated.

Figure 11: Link Protecting LFA



In the above illustration, if router E fails, then both router S and router N detects a failure and switch to their alternates, causing a forwarding loop between both routers S and N. Thus, the link protecting LFA causes loop on node failure; however, this can be avoided by using a down-stream path, which can limit the coverage of alternates. Router S will be able to use router N as a downstream alternate, however, router N can't use S. Therefore, N would have no alternate and would discard the traffic, thus avoiding the micro looping.

Configuring MLDP Loop-Free Alternative Fast Reroute

The following section describes the configurations to enable LFA FRR:

Note mLDP FRR relies on the IGP protocol, you can either configure OSPF or ISIS.

Configuring Router OSPF LFA FRR

mLDP FRR relies on the IGP protocol, you can configure either OSPF or ISIS.

The OSPF LFA FRR uses a precomputed alternate next hop to reduce failure reaction time when the primary next hop fails. It lets you configure a per-prefix LFA path that redirects traffic to a next hop other than the primary neighbor.

Configuration Example

```
Router# configure
Wed Apr  7 08:54:52.769 UTC
Router(config)# router ospf 0
Router(config-ospf)# area 0
Router(config-ospf-ar)# interface Bundle-Ether10

/*Enabling Per Prefix LFA*/
Router(config-ospf-ar-if)# fast-reroute per-prefix

/*To add interfaces to LFA Candidate List:*/
Router(config-ospf-ar-if)# fast-reroute per-prefix lfa-candidate interface Bundle-Ether10

/*To exclude interface from backup*/
Router(config-ospf-ar-if)# fast-reroute per-prefix exclude interface Bundle-Ether10

/*To restrict the backup interface to the LFA candidate list:*/
Router(config-ospf-ar-if)# fast-reroute per-prefix use-candidate-only enable
Router(config-ospf-ar-if)# commit
```

Running Configuration

```
router ospf 0
 area 0
  interface Bundle-Ether10
   fast-reroute per-prefix
   fast-reroute per-prefix exclude interface Bundle-Ether10
   fast-reroute per-prefix lfa-candidate interface Bundle-Ether10
   fast-reroute per-prefix use-candidate-only enable
  !
 !
!
```

Configuring Router ISIS LFA FRR

IS-IS computes LFA next-hop routes for the forwarding plane to use in case of primary path failures. LFA is computed per prefix.

Configuration Example

```
Router# configure
Router(config)# router isis MCAST
Router(config-isis)# net 49.0001.0000.0000.0001.00
Router(config-isis-af)# interface HundredGigE0/0/24
Router(config-isis-if-af)# address-family ipv4 unicast
/*configure per-prefix Link based LFA*/
Router(config-isis-if-af)# fast-reroute per-prefix
```

Running Configuration

```
!
router isis MCAST
 net 49.0001.0000.0000.0001.00
 interface HundredGigE0/0/0/24
  address-family ipv4 unicast
   fast-reroute per-prefix  <---- configure per-prefix Link based LFA
  !
 !
!
```

Configuring Bidirectional Forwarding Detection

When a local interface is down, it can take a long delay for the remote peer to detect the link disconnection. To quickly detect if the remote interface is down, the physical port and bundle interfaces must have Bidirectional Forwarding Detection (BFD) to ensure faster failure detection.

```
Router#configure
Router(config)#router ospf 0
Router(config-ospf)#nsr
Router(config-ospf)#router-id 21.21.21.21
Router(config-ospf)#nsf cisco
Router(config-ospf)#address-family ipv4 unicast
Router(config-ospf)#area 0
Router(config-ospf-ar)#bfd minimum-interval 3
Router(config-ospf-ar)#bfd fast-detect
Router(config-ospf-ar)#bfd multiplier 2
Router(config-ospf-ar)#fast-reroute per-prefix
Router(config-ospf-ar)#mpls traffic-eng
Router(config-ospf-ar)#interface Bundle-Ether100.1
Router(config-ospf-ar-if)#bfd fast-detect
Router(config-ospf-ar-if)#fast-reroute per-prefix

Router(config-ospf-ar)#interface Bundle-Ether100.2
Router(config-ospf-ar-if)#bfd fast-detect
Router(config-ospf-ar-if)#fast-reroute per-prefix
Router(config-ospf-ar-if)#commit
```

Running Configuration

```
router ospf 0
 nsr
 router-id 21.21.21.21
 nsf cisco
 address-family ipv4 unicast
 area 0
  bfd minimum-interval 3
  bfd fast-detect   <---- configure bfd fast-detect
  bfd multiplier 2
  fast-reroute per-prefix
  mpls traffic-eng
  interface Bundle-Ether100.1
   bfd fast-detect    <---- configure bfd fast-detect under the "protected" inerface
   fast-reroute per-prefix
  !
  interface Bundle-Ether100.2
   bfd fast-detect   <---- configure bfd fast-detect under the "protected" inerface
   fast-reroute per-prefix
  !
```

For bundle main interface, configure BFD under the bundle interface:

Note The **bundle minimum-active links** is required if LACP is not configured on the bundle members.

```
interface Bundle-Ether101
bfd address-family ipv4 timers start 60
bfd address-family ipv4 timers nbr-unconfig 3600
bfd address-family ipv4 multiplier 2
bfd address-family ipv4 destination 44.2.0.4
bfd address-family ipv4 fast-detect
bfd address-family ipv4 minimum-interval 3
ipv4 address 44.2.0.1 255.255.255.0
ipv6 address 44:2::1/64
bundle minimum-active links 3
!
```

For LACP, configure **mode active** under bundle member:

```
interface HundredGigE0/0/0/22
bundle id 101 mode active
!
interface HundredGigE0/0/0/28
bundle id 101 mode active
!
interface HundredGigE0/0/0/29
bundle id 101 mode active
!
```

For physical interface and subinterface and bundle subinterface, configure BFD under IGP, for example ISIS:

```
router isis ring
interface HundredGigE0/0/0/22
  bfd minimum-interval 10
  bfd multiplier 2
  bfd fast-detect ipv4
  point-to-point
  address-family ipv4 unicast
   fast-reroute per-prefix
   metric 1000
   weight 1000
  !
!
```

Configuring MPLS LFA FRR

Configuration Example

Configure session protection to support MLDP LFA FRR:

```
Router# configure
Router(config)# mpls ldp
Router(config-ldp)# nsr
Router(config-ldp)# graceful-restart
Router(config-ldp)# router-id 20.20.20.20
Router(config-ldp)# session protection
Router(config-ldp)# address-family ipv4
Router(config-ldp-af)#commit
```

Show Running Configuration

```
mpls ldp
  nsr
  graceful-restart
 !
 nsr
 graceful-restart
 router-id 20.20.20.20
 session protection
 address-family ipv4
 !
!
```

Make Before Break Configuration for LFA FRR

Make Before Break (MBB) is an inherent nature of MLDP. In MBB configuration, configure **forwarding recursive** to enable LFA FRR feature. If forwarding recursive is not configured, MLDP uses non-recursive method to select MLDP core facing interface towards next hop. The detailed configuration steps and an example follows.

```
Router(config)# mpls ldp
Router(config-ldp)# log
Router(config-ldp-log)# neighbor
Router(config-ldp-log)# nsr
Router(config-ldp-log)# graceful-restart
Router(config-ldp-log)# mldp
Router(config-ldp-mldp)# address-family ipv4
Router(config-ldp-mldp-af)# forwarding recursive
Router(config-ldp-mldp-af)# make-before-break delay 60
Router(config-ldp-mldp-af)# commit
```

Configuring Make Before Break Delay and Delete

By default, MBB is set to 10 seconds. You can configure different MBB timing to determine when the merge node starts to accept the new label.

In this configuration example, the MBB (delay) period is set of 90 seconds. The merge node starts accepting new label 90 seconds after detecting the link disconnection towards the head node. The delete delay is set to 60 seconds; that is, when MBB expires, the time period after which the merge node sends old label delete request to head node is 60 seconds. The default value is zero. The range of delete delay is from 30 to 60, for scale LSPs.

```
Router# configure
Router(config)# mpls ldp
Router(config-ldp)# mldp
Router(config-ldp-mldp)# address-family ipv4
Router(config-ldp-mldp-af)# make-before-break delay 90
Router(config-ldp-mldp-af)# make-before-break delay 90 60
Router(config-ldp-mldp-af)# commit
```

Verification of MLDP Configuration

Use the following show commands to verify the mLDP LFA FRR configuration:

The following example shows how to verify mLDP Neighbor:

```
Router# show mrib regdb
Tue Mar 23 17:45:27.762 UTC
  NH  addr          : 45.45.45.45    <--- Next Hop Peer's Loopback Address
  Destination vrf   : default
  Regdb Entry Type  : Label
  IP Ole count      : 0x0
  Label Ole count   : 0x2
  MLC Ole count     : 0x0
  ECD registered    : YES
  ECD stale         : NO
  ECD Information   : 55a76e23e000
  ECD Length        : 50
  Number of notifs  : 1
```

The following example shows how to verify the mLDP traffic. The zero in the following example indicates that there's no mLDP packet forwarding out of that outgoing interface.

```
Router# show mpls forwarding p2mp
Local Outgoing Prefix Outgoing Next Hop Bytes
Label Label or ID Interface Switched
------ ----------- ------------------ ------------ --------------- ------------
40127 40018 mLDP/IR: 0x003e9 Hu0/3/0/0 61.154.2.50 113972905604
40045 mLDP/IR: 0x003e9 Hu0/3/0/26 82.154.5.2 113339018314
40128 40019 mLDP/IR: 0x003ea Hu0/1/0/17/1 97.1.1.5 113972768600
40046 mLDP/IR: 0x003ea Hu0/3/0/32 82.154.4.2 0
40129 40020 mLDP/IR: 0x003eb Hu0/1/0/17/1 97.1.1.5 113972234482
40047 mLDP/IR: 0x003eb Hu0/3/0/26 82.154.5.2 0
40130 40021 mLDP/IR: 0x003ec Hu0/1/0/5 10.10.10.17 113972828144
40048 mLDP/IR: 0x003ec Hu0/3/0/26 82.154.5.2 0
40131 40022 mLDP/IR: 0x003ed Hu0/1/0/17/1 97.1.1.5 113973181832
40049 mLDP/IR: 0x003ed Hu0/3/0/32 82.154.4.2 0
40132 40023 mLDP/IR: 0x003ee Hu0/1/0/17/1 97.1.1.5 113972294384
40050 mLDP/IR: 0x003ee Hu0/3/0/26 82.154.5.2 0
40133 40024 mLDP/IR: 0x003ef Hu0/3/0/0 61.154.2.50 113972687482
```

The following example shows how to view the list of mLDP in the router:

```
Router# show mrib mpls forwarding detail
LSP information (mLDP) :
  LSM-ID: 0x00014, Role: Mid
    Incoming Label       : 24028
    Transported Protocol : <unknown>
    Explicit Null        : None
    IP lookup            : disabled
    Platform information : MCGID: 56633, Tunnel RIF: -1, RIF VRF: -1  <--- The local label
 24028 has MCGID: 56633,

          used for programming label's FAP ID bitmask
    Outsegment Info #1 [M/Swap, Recursive]:
      OutLabel: 24027, NH: 45.45.45.45, ID: 0x14, Sel IF: Bundle-Ether101(V)  <---Primary
path BE101 (HundredGigE0/2/0/17)
        UL IF: HundredGigE0/2/0/17, Node-ID: 0x9
      Backup Tunnel: Un:0x0  Backup State: Ready, NH: 0.0.0.0, MP Label: 0
        Backup Sel IF: Bundle-Ether102(V), UL IF: HundredGigE0/0/0/13, Node-ID: 0x1
<-----Backup path BE102 (HundredGigE0/0/0/13).

LSP information (mLDP) :
  LSM-ID: 0x00015, Role: Mid
    Incoming Label       : 24029
    Transported Protocol : <unknown>
    Explicit Null        : None
    IP lookup            : disabled
    Platform information : MCGID: 56634, Tunnel RIF: -1, RIF VRF: -1

    Outsegment Info #1 [M/Swap, Recursive]:
      OutLabel: 24028, NH: 45.45.45.45, ID: 0x15, Sel IF: Bundle-Ether101(V)
        UL IF: HundredGigE0/2/0/18, Node-ID: 0x8
      Backup Tunnel: Un:0x0  Backup State: Ready, NH: 0.0.0.0, MP Label: 0
        Backup Sel IF: Bundle-Ether102(V), UL IF: HundredGigE0/0/0/27, Node-ID: 0x2
```

The following example shows how to view the details of a specific mLDP.

```
Router# show mpls forwarding labels 24028 detail
Tue Mar 23 17:47:28.962 UTC
Local  Outgoing    Prefix             Outgoing     Next Hop        Bytes
Label  Label       or ID              Interface                    Switched
------ ----------- ----------------- ------------ --------------- ------------
24028             mLDP/IR: 0x00014 (0x00014)
       Updated Mar 23 17:28:29.946
       mLDP/IR LSM-ID: 0x00014, MDT: 0x0
       Flags:IP Lookup:not-set, Expnullv4:not-set, Expnullv6:not-set
             Payload Type v4:not-set, Payload Type v6:not-set, l2vpn:not-set
             Head:not-set, Tail:not-set, Bud:not-set, Peek:not-set, inclusive:not-set
             Ingress Drop:not-set, Egress Drop:not-set
             RPF-ID:0, Encap-ID:0
             Disp-Tun:[ifh:0x0, label:-]
             Platform Data [28]:
              { 0 0 221 57 0 0 0 4
                0 0 0 2 0 0 0 0
                144 207 44 47 255 255 255 255
                0 0 0 0 }
       mpls paths: 1, local mpls paths: 1, protected mpls paths: 1

       24027       mLDP/IR: 0x00014 (0x00014)    \
                                  BE101        44.2.0.4        1130065367760
         Updated: Mar 23 17:28:29.952
         My Nodeid:0x2000
         Interface Nodeids:
           [ 0x9 - - - - - - - - -  ]
         Interface Handles:
           [ 0x1000218 - - - - - - - - -  ]
         Backup Interface Nodeids:
           [ 0x1 - - - - - - - - - ]
         Backup Interface Handles:
           [ 0x240 - - - - - - - - - ]
         Packets Switched: 1121096595
```

The following example shows how to see the mLDP neighbors:

```
Router# show mpls mldp neighbors
Sat May  9 06:37:06.877 UTC
mLDP neighbor database
 MLDP peer ID      : 20.20.20.20:0, uptime 01:38:38 Up,
  Capabilities     : GR, Typed Wildcard FEC, P2MP, MP2MP, MBB
  Target Adj       : No
  Upstream count   : 1
  Branch count     : 2
  LDP GR           : Enabled
                   : Instance: 1
  Label map timer  : never
  Policy filter in : None
  Path count       : 2
  Path(s)          : 21.20.100.1       Bundle-Ether100 LDP
                   : 21.20.20.2        Bundle-Ether20.1 LDP
  Adj list         : 21.20.20.2        Bundle-Ether20.1
                   : 21.20.100.1       Bundle-Ether100
  Peer addr list   : 172.18.51.116
                   : 20.20.20.20
                   : 20.22.5.1
                   : 22.20.23.2
                   : 21.20.20.2
                   : 21.20.100.1

 MLDP peer ID      : 22.22.22.22:0, uptime 01:38:38 Up,
  Capabilities     : GR, Typed Wildcard FEC, P2MP, MP2MP
  Target Adj       : No
  Upstream count   : 0
  Branch count     : 2
  LDP GR           : Enabled
                   : Instance: 1
  Label map timer  : never
  Policy filter in : None
  Path count       : 2
  Path(s)          : 22.21.23.1        TenGigE0/1/0/3 LDP
                   : 22.21.20.1        TenGigE0/2/1/0 LDP
  Adj list         : 22.21.20.1        TenGigE0/2/1/0
                   : 22.21.23.1        TenGigE0/1/0/3
  Peer addr list   : 172.18.51.118
                   : 22.22.22.22
                   : 20.22.5.2
                   : 22.2.9.1
                   : 22.20.23.1
                   : 22.21.20.1
                   : 22.21.23.1#
```

Multicast Route Statistics

Table 4: Feature History Table

Feature Name	Release Information	Feature Description
Multicast Route Statistics	Release 7.3.1	When enabled, this feature provides statistics on the number of packets received for a multicast route. This information may be useful for monitoring and billing purposes.

Multicast route statistic provides information about the multicast routes. The multicast statistics information includes the rate at which packets are received and the number of packets received.

Cisco IOSXR Software counters are always present. To enable per-prefix counters only in hardware, use the **accounting per-prefix** command. When per-prefix counters are enabled, existing, and new (S, G) and (*, G) routes are assigned ingress counters, except for the following:

- Default multicast routes
- IPv4 (*, G) routes configured with prefix length less than 32.
- IPv6 (*, G) routes configured with prefix length less than 128.

If there is limited number of counters available and you want to enable counters on particular prefixes for troubleshooting purposes, you can configure **hw-module route-stats** to enable accounting for multicast routing for a limited number of routes.

For more information, see the **hw-module route-stats** command to configure a filter to choose which (S.G) routes have statistics enabled.

Restrictions

Supports multicast route statistics for ingress direction

Configuring multicast route stats

Perform the following to configure multicast route stats:

- Configure rate per route
- Enable per-prefix counters
- Create Access Control List
- Enable multicast route statistics on a particular prefix

Configuration Example

The following example shows how to enable multicast route statistics for IPv4:

```
/*Configure rate per route*/
Router# configure
Router(config)# multicast-routing
Router(config-mcast)# address-family ipv4
Router(config-mcast-default-ipv4)# rate-per-route

/*Enable per-prefix counters*/
Router# configure
Router(config)# multicast-routing
Router(config-mcast)# address-family ipv4
Router(config-mcast-default-ipv4)# accounting per-prefix

/*Create ACL*/

Router(config)# ipv4 access-list mcast-counter
Router(config-acl)# 10 permit ipv4 host 10.1.1.2 host 224.2.151.1
Router(config-acl)# 30 permit ipv4 10.1.1.0/24 232.0.4.0/22
Router(config-acl)# 50 permit ipv4 192.168.0.0/24 232.0.4.0/22
Router(config-acl)#  commit
Router(config-acl)# exit

/*Enable multicast route statistics on a particular prefixe*\

Router(config)# hw-module route-stats l3mcast vrf default ipv4 mcast-counter
```

Similarly, you can enable route statistics for IPv6 address:

```
/*Configure rate per route*/
Router# configure
Router(config)# multicast-routing
Router(config-mcast)# address-family ipv6
Router(config-mcast-default-ipv4)# rate-per-route

/*Enable per-prefix counters*/
Router# configure
Router(config)# multicast-routing
Router(config-mcast)# address-family ipv6
Router(config-mcast-default-ipv4)# accounting per-prefix

/*Create ACL*/
Router# configure
Router(config)# ipv6 access-list mcast-counter
Router(config-ipv6-acl)# 10 permit ipv6 any any
Router(config-ipv6-acl)# 20 deny udp any any
Router(config-ipv6-acl)# commit
Router(config-ipv6-acl)# exit

/*Enable multicast route statistics on a particular prefixe*\

Router(config)# hw-module route-stats l3mcast vrf default ipv6 mcast-counter
```

Verification

```
Router# show mfib route statistics location 0/RP0/CPU0
Thu Aug 13 19:16:58.321 UTC

IP Multicast Forwarding Rates
(Source Address, Group Address)
Incoming rate:
Node: (Incoming node) : pps/bps
Outgoing rate:
Node: (Outgoing node) : pps/bps

(192.168.0.0,232.0.4.0)
Incoming rate :
Node : 0/RP0/CPU0 : 749 / 1007969
Outgoing rate :
```

```
Node : 0/RP0/CPU0 : 0 / 0
RP/0/RP0/CPU0:ios#
```

To clear the Multicast Forwarding Information Base (MFIB) route packet hardware counters, use the **clear mfib platform route statistics** command.

Note To clear an ingress statistics of a route, you can get the `stats-ole` location for a specified route using the **show mrib route detail** command.

A `stats-ole` is programmed on one of the line cards for a particular route and helps report ingress statistics for a particular route.

If you know the `stats-ole` location, you can clear the ingress counters for a route on that location. If you do not know the `stats-ole` location, you can use the option `location all` instead, which helps to find the specific `stats-ole` location and clear the ingress counters.

The following example shows how to find the `stats-ole` location:

```
Router # show mrib vrf vrf15 route 18.18.15.2 225.0.0.1 detail

IP Multicast Routing Information Base
Entry flags: L - Domain-Local Source, E - External Source to the Domain,
    C - Directly-Connected Check, S - Signal, IA - Inherit Accept,
    IF - Inherit From, D - Drop, MA - MDT Address, ME - MDT Encap,
    MD - MDT Decap, MT - MDT Threshold Crossed, MH - MDT interface handle
    CD - Conditional Decap, MPLS - MPLS Decap, MF - MPLS Encap, EX - Extranet
Interface flags: F - Forward, A - Accept, IC - Internal Copy,
    NS - Negate Signal, DP - Don't Preserve, SP - Signal Present,
    II - Internal Interest, ID - Internal Disinterest, LI - Local Interest,
    LD - Local Disinterest, DI - Decapsulation Interface
    EI - Encapsulation Interface, MI - MDT Interface, LVIF - MPLS Encap,
    EX - Extranet

18.18.15.2 225.0.0.1) Ver: 0x4df RPF nbr: 18.18.15.2 Flags: RPF, MRID: 60638, MCGID: 61036,

Stats T [R/S/I]: 0/11/0  /* 0/11/CPU0 is the stats-ole location. */
    Up: 01:45:14
     Incoming Interface List
                           Bundle-Ether43.80 Flags: A, Up: 01:45:14
     Outgoing Interface List
                           HundredGigE0/3/0/22.180 Flags: F NS, Up: 01:45:14
```

From the earlier example, you know that `stats-ole` location is `0/11/CPU0`. You can now clear the ingress stats using `0/11/CPU0` location.

```
Route# clear mfib platform route statistics location 0/11/CPU0
```