

Optimized Layer 2 Overlay Multicast

- Multicast Layer 2 EVPN Type 6 routes, on page 1
- Feature support and limitations for multicast Layer 2 EVPN Type-6 routes, on page 6
- Configuration options for Layer 2 EVPN route Type 6, on page 8
- Commands for multicast Layer 2 EVPN route types, on page 10
- Command output examples for multicast Layer 2 EVPN route Type 6, on page 11

Multicast Layer 2 EVPN Type 6 routes

A multicast Layer 2 EVPN Type-6 route is a BGP route type that:

- enables selective multicast traffic forwarding over a VXLAN fabric,
- advertises interest in specific multicast groups from VTEPs using SMET routes, and
- optimizes bandwidth by only replicating multicast traffic to VTEPs with interested receivers.

In data center topologies, many hosts can be attached to a VLAN or subnet, possibly spanning multiple PODs. In VXLAN-based data centers, VXLAN Tunnel Endpoints (VTEPs) are devices connected to the fabric that create VXLAN tunnels to send Layer 2 traffic between VLAN segments.

Sources located on a VTEP can send IP multicast (S,G) traffic toward receivers by encapsulating packets in VXLAN on the source VTEP and flooding to all remote VTEPs. Receiving VTEPs decapsulate and forward to interested receivers. However, indiscriminate flooding wastes bandwidth if no receivers are interested.

- Type 6 EVPN Selective Multicast (SMET) Routes provide a solution to efficiently send IP multicast traffic between sources and receivers over the VXLAN fabric.
- Hosts attached to VTEPs express interest in multicast groups by sending membership reports on a VLAN; the VTEP snoops the IGMP reports and advertises a Type 6 EVPN SMET route to other remote VTEPs through BGP, enabling all VTEPs attached to the fabric to learn about receivers interested in specified multicast groups.
- VTEPs (leaf nodes) will only send IP multicast (S,G) traffic originated from locally connected sources over the VXLAN fabric if there is at least one receiver on a remote VTEP that is interested in the traffic, thereby optimizing the use of the fabric by not sending multicast traffic if there are no interested receivers.
- If ingress-replication is used in the underlay, the traffic can be further optimized by only sending multicast traffic to those VTEPs that have attached receivers.

Note Cisco NX-OS does not currently support selective multicast delivery over a unicast underlay.

RFC 9251 describes the procedures and NLRI formats for Type 6 EVPN SMET routes.

Functionality of Multicast Layer 2 EVPN Type-6 routes

Layer 2 EVPN Route Type-6 (SMET) routes for IPv4 (IGMP) routes have been supported in NX-OS as L2TRM**.

Beginning with Cisco NX-OS Release 10.5(3)F, the Multicast Layer 2 EVPN Route Type-6 supports new functionality for IPv4 IGMP and IPv6 MLD snooping.

Beginning with Cisco NX-OS Release 10.5(3)F, the Multicast Layer 2 EVPN Route Type-6 will support the following functionality:

- IPv4 IGMP EVPN Type-6 SMET Routes using L2RIB instead of next-generation multicast VPN (NGMVPN).
- IPv6 MLD Snooping EVPN Type-6 SMET Routes.

Generating IGMP EVPN routes using L2RIB

Starting from Cisco NX-OS Release 10.5(3)F, added the following capabilities to Type-6 EVPN SMET route functionality.

- Multicast router or querier behind the EVPN domain, to support (*, *) route advertisement or receipt.
- Multiversion (IGMPv2/v3/v2+v3) advertisement capability

These capabilities are in compliance with RFC 9251.

Generating MLD Snooping SMET route

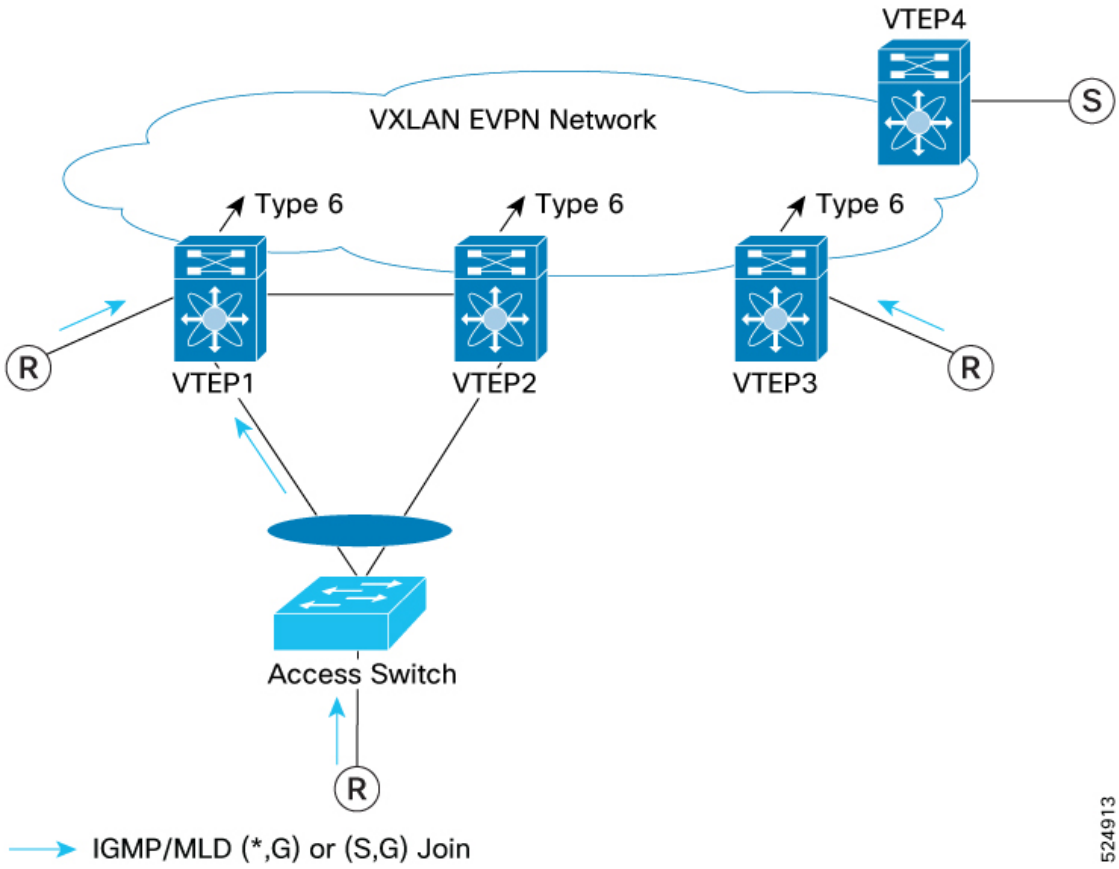
Before Cisco NX-OS Release 10.5(3)F, SMET route generation was limited to IPv4 IGMP snooped entries. Beginning with Cisco NX-OS Release 10.5(3)F, Type-6 SMET route advertisement for MLD snooped routes is supported where L2RIB and BGP will handle IPv4 or IPv6 SMET routes from IGMP/MLD on the transmit side and receive side, filtering and forwarding routes appropriately. Multicast router and multiversion capability (MLDv1/v2/v1+v2) will be supported.

Supported topologies for multicast Layer 2 EVPN Type-6 routes

EVPN Route Type-6 will support vPC, vPC Fabric peering, ESI and non-vPC topologies.

Sample topologies and associated control plane flow for these topologies are as mentioned below.

Layer 2 EVPN network with vPC



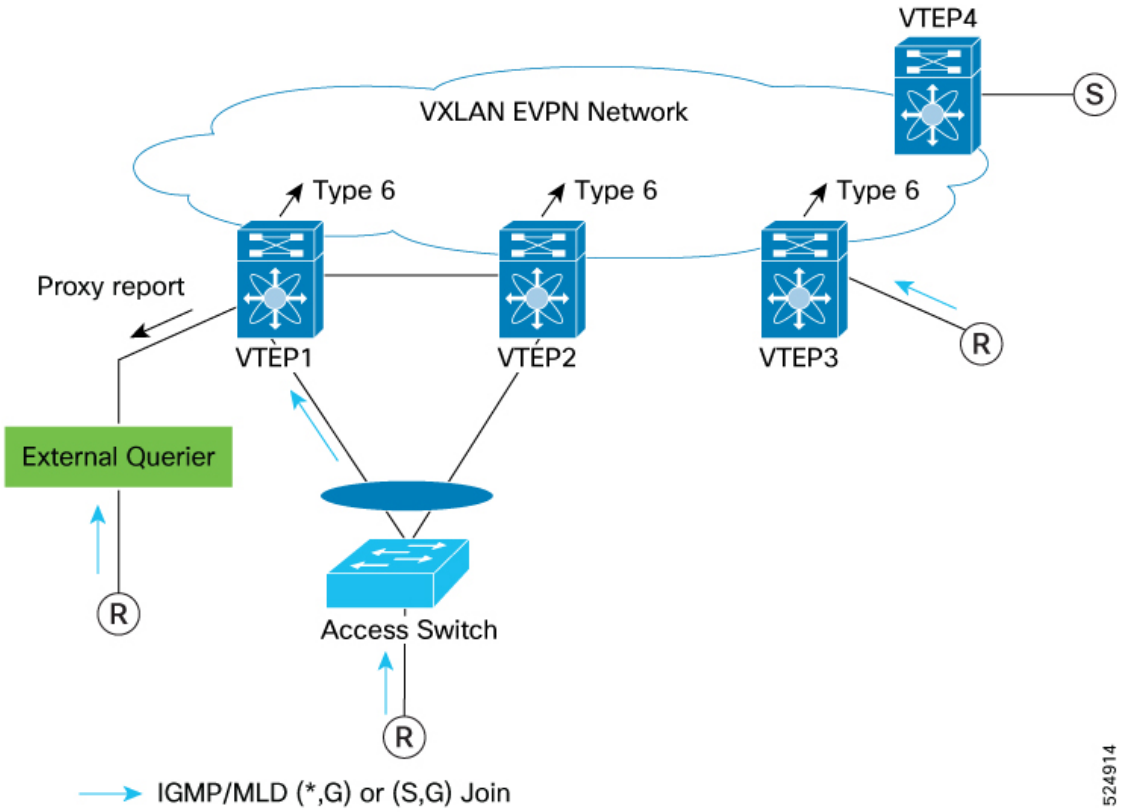
In this example topology, there are 4 VTEPs - a pair of vPC VTEPs (VTEP1 and VTEP2), and two standalone VTEPs (VTEP3 and VTEP4). Receivers (R) are attached to

- an orphan port behind VTEP1,
- access switch behind VPC complex VTEP1 and VTEP2, and
- standalone switch VTEP3.

When these receivers send (*, G) and/or (S,G) IGMP / MLD membership reports that are received by the VTEP1, VTEP2 and VTEP3, each VTEP with attached receiver originates a Type-6 EVPN SMET route type, which is received by all VTEPs, including VTEP4 to which the source is attached.

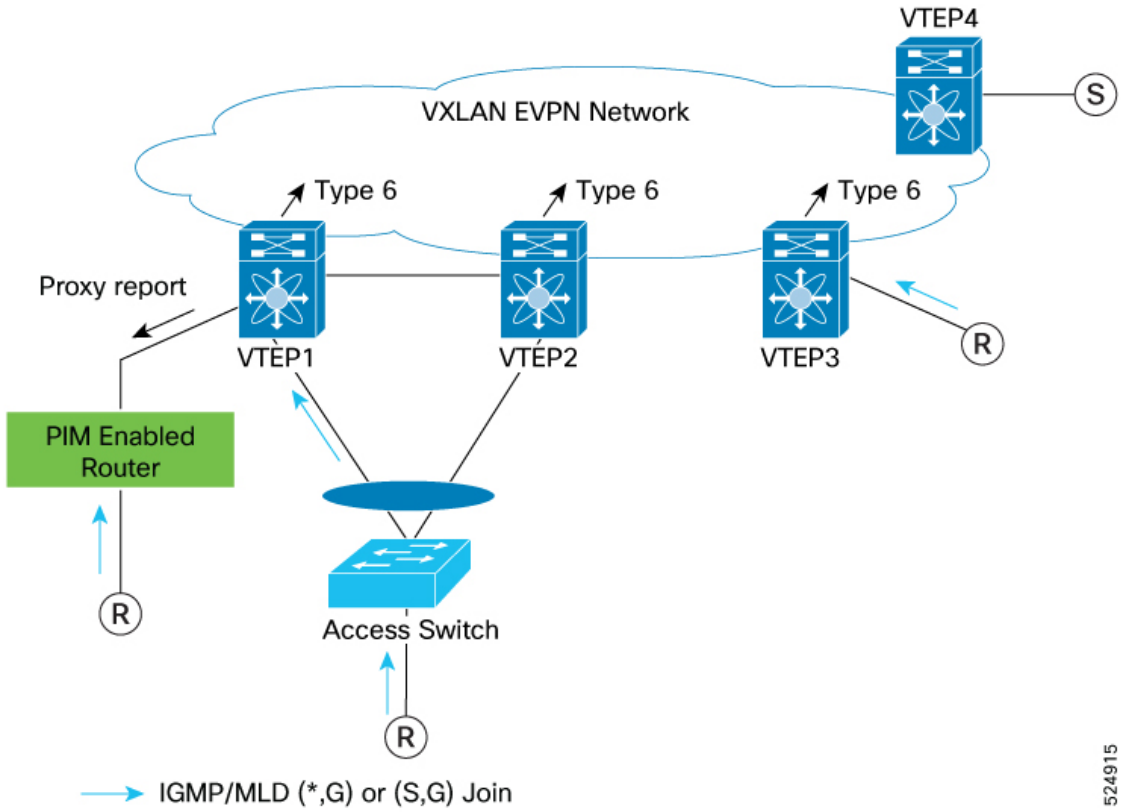
When the source ‘S’ behind VTEP4 sends multicast (S,G) traffic, the traffic is encapsulated in VXLAN by VTEP4 and sent over the fabric toward receivers; receiving VTEPs VTEP1, VTEP2 and VTEP3 decapsulate the VXLAN packet and forward toward local networks, as shown in the attached figure. Note that VTEP4 will only forward multicast traffic if it receives at least one EVPN Type-6 EVPN SMET route from other VTEPs in the fabric.

Layer 2 EVPN network with vPC and external querier



In this example topology, there is an external querier attached to the orphan port, and an election is done between the VTEP1 and External Querier to determine the Querier for the network. For the case where the external querier wins the election, VTEP1 views the external querier as attached to an ‘mrouter’ port; VTEP1 advertises a (*,*) route toward the EVPN network. Any multicast traffic received from source ‘S’ is forwarded toward the external querier. Furthermore, reports received from VTEP3 over SMET are sent as a proxy report toward the querier. There are variants of the ‘external querier’ topology, where the querier is attached to a standalone node, or attached to the vPC complex, and the case where the VTEP wins the Querier election. These variants are supported, though the specific IGMP / MLD report and traffic flows depend on the specific querier location or querier election winner.

Layer 2 EVPN network with vPC and PIM enabled router



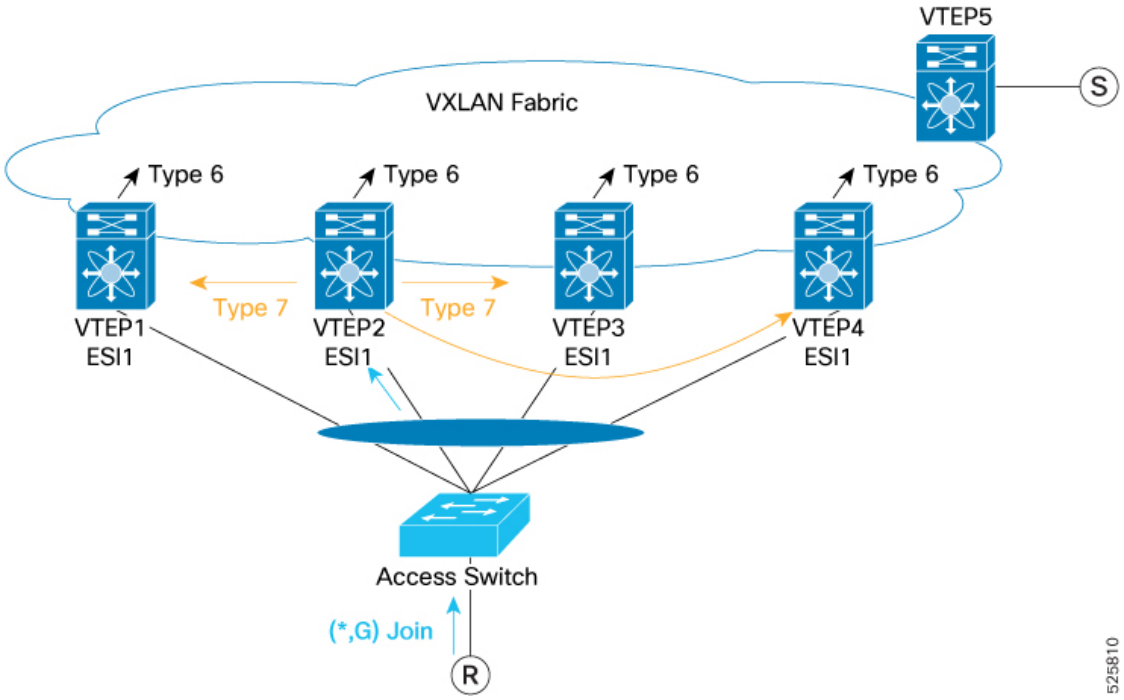
In this example topology, there is a PIM enabled router to the orphan port. In this case, vPC VTEP1 receives PIM hellos from the PIM enabled router, and views the orphan port as a mrouter port; VTEP1 will advertise a (*,*) route toward the EVPN network. Any multicast traffic received from source ‘S’ is forwarded toward the PIM enabled router. Furthermore, reports received from VTEP3 over SMET are sent as a proxy report toward the PIM enabled router. There are variants of the ‘PIM router’ topology, where the PIM enabled router is attached to a standalone node; however, the PIM enabled router attached to the vPC complex access switch is not supported.

Note PIM enabled router attached to vPC complex access switch is not supported.

Layer 2 EVPN with ESI

Beginning with Cisco NX-OS Release 10.6(1)F, the support for ESI port-channels has been added. For more information, see Supported L2 multicast topologies and IGMP control flows . Correspondingly, support for EVPN Type-6 SMET route generation for a receiver R attached to an ESI port channel is also supported, as shown.

Figure 1: Layer 2 EVPN with ESI



In the ESI case, receivers can send a (*,G) and (S,G) IGMP or MLD membership report over the ESI port-channel. When the report is received by one of the VTEPs, this VTEP will originate a Type-6 EVPN SMET route across the EVPN network. Separately, the VTEP also sends an EVPN Type-7 Report-Sync route to ESI peers. When ESI peers receive the Type-7 routes, these VTEPs originate a Type-6 route to the EVPN network. Note that there is currently no support for the DF only to originate the Type-6 route, as is recommended in RFC 9251.

Feature support and limitations for multicast Layer 2 EVPN Type-6 routes

This topic outlines the supported features, configuration constraints, and platform compatibility for multicast Layer 2 EVPN route Type 6.

Generic configuration guidelines and limitations

- If **advertise evpn multicast** is configured, it is recommended to disable NVE as a static router port using these command:
  - **ip igmp snooping disable-nve-static-router-port**
  - **ipv6 mld snooping disable-nve-static-router-port**
- The **advertise evpn multicast** command is required for originating Type 6 SMET routes; IPv4 IGMP Snooped Routes are used for IPv4 and IPv6 MLD Snooped Routes are used for IPv6.

Note If **advertise evpn multicast** is not configured, IPv4 IGMP will flood IGMP control packets in the fabric. However, IPv6 MLD will not flood MLD control packets in the fabric.

- The NGMVPN feature is no longer necessary to originate Type 6 routes, though it can remain without impact. The configuration of "feature ngmvpn" was previously required as part of "L2TRM" for the Type 6 EVPN SMET route advertisement.
- Beginning with Cisco NX-OS Release 10.5(3)F, **show fabric multicast ipv4 l2-mroute** command is deprecated, there are equivalent commands in L2RIB to view the route state in L2RIB component.
- PIM packets will be flooded in the fabric.
- Beginning with Cisco NX-OS Release 10.6(1)F, EVPN route Type 6, 7 and 8 supports interoperability through RFC6625 encoding.

Route Type advertising and receiving support

- Supports advertisement and receipt of IGMPv2, IGMPv3, and combinations of IGMPv2 and IGMPv3.
- Supports advertisement and receipt of MLDv1, MLDv2, and combinations of MLDv1, MLDv2.
- Supports advertisement and receipt of (*,G) IGMPv3 and (*,G) MLDv2 entries with the exclude flag
- Does not support IGMPv1 SMET route origination (as per RFC 9251).
- Does not support advertisement of S,G routes with the exclude flag; these are advertised as (*, G) exclude routes.
- Supports receipt of (S,G) routes with the exclude flag by treating as (*,G) routes.
- Supports advertisement of EVPN Type 6, 7 and 8 routes with multicast source length as 0 for (*, G) routes and both multicast source length and multicast group length set to 0 for (*, *) routes when the **advertise evpn multicast format rfc6625** command is enabled.

Feature support

- Only supported in a Layer 2 fabric.
- Supports IPv4 multicast and ingress-replication (IR) underlay. However, the IR implementation does not support sending multicast traffic to only VTEP nodes with associated receivers.
- Does not support IPv6 multicast and IR underlay
- Border Gateways do not support re-origination of SMET routes learnt from another peer. However, for any natively learned IGMP interest via local hosts or on the nve1 interface via VxLAN Fabric, BGWoriginates the SMET Type-6 route which can be utilized by the remote VTEP/BGW for originating proxy reports to any PIM routers/IGMP querier connected to it.
- Does not support MLD Snooping with VXLAN on EoR.
- To ensure membership reports are received and processed by the VTEP, querier functionality must be enabled on VTEPs based on topology.

- For successful downgrade from Cisco NX-OS Release 10.6(1)F to a prior release, ensure that **format rfc6625** option is removed from **advertise evpn multicast format rfc6625** configuration, otherwise downgrade is impacted.

Scale limitations

- In case of vPC or vPC fabric peering, triggers like uplink flap, reload where the peer-link goes down, there can be momentary increase in platform programming resources causing exhaustion. If exhaustion occurs due to such triggers, follow these steps:
  - Use the **clear ip igmp snooping groups * vlan all** command to clear the respective IGMP and MLD snooping tables.
  - Use the **no advertise evpn multicast** command to disable the SMET route generation on the affected node to recover from any mis-programming.
  - Use the **advertise evpn multicast** command to enable the SMET route generation.

Table 1: Supported release and platform of multicast Layer 2 EVPN route Type 6

Release	Platform
10.5(3)F and later	Cisco Nexus 9300-FX/FX2/FX3/GX/GX2/H2R/H1 Series switches and Nexus 9500 Series switches with 9700-FX/GX/FX3 line cards

Configuration options for Layer 2 EVPN route Type 6

Multicast Layer 2 EVPN route Type 6 enables the advertisement of Source Multicast Ethernet Tag (SMET) routes in a BGP EVPN environment, supporting efficient multicast forwarding within a VXLAN fabric.

Key features and configuration options include:

- **SMET route generation**: Use the **advertise evpn multicast** command to generate and advertise Type 6 (SMET) routes in BGP EVPN. The procedure is similar to Configure Layer 2 Tenant Routed Multicast section.
- **IGMP Snooping Querier**: Configure the IGMP Snooping Querier for each multicast-enabled VXLAN VLAN on all Layer-2 leaf switches to support correct multicast group management.

Configure IPv4 IGMP Snooping

Use this task to configure IGMP snooping when multicast applications require efficient group management and traffic distribution in VXLAN EVPN environments.

Follow these steps to configure IPv4 IGMP Snooping.

Before you begin

Ensure that VXLAN EVPN is configured.

Procedure

Step 1 Run the **advertise evpn multicast** command in global configuration mode, to advertise Layer 2 multicast capability.

Example:

```
switch# configure terminal
switch(config)# advertise evpn multicast
```

Step 2 (Optional) Run the **advertise evpn multicast format rfc6625** command to enable RFC6625 encoding for EVPN route Type 6, 7 and 8, setting multicast source length to 0 for (*, G) routes and both multicast source length and multicast group length to 0 for (*, *) routes.

Example:

```
switch(config)# advertise evpn multicast format rfc6625
```

Note Use the **format rfc6625** option only if interoperability is required.

Use **advertise evpn multicast** command to revert to the original encoding method with multicast source length as 32/128 and multicast source address as 0.0.0.0/0::0 for (*, G) routes, and both lengths as 32/128 with addresses as 0.0.0.0/0::0 for (*, *) routes.

Step 3 Run the **ip igmp snooping vxlan** command to configure IGMP snooping for VXLANs.

Example:

```
switch(config)# ip igmp snooping vxlan
```

Step 4 Run the **vlan configuration** *vlan-id* command to enter configuration mode for VLAN 101.

Example:

```
switch(config)# vlan configuration 101
switch(config-vlan-config)#
```

Step 5 Run the **ip igmp snooping querier** *querier-ip-address* command to configure IGMP snooping querier for each multicast-enabled VXLAN VLAN.

Example:

```
switch(config-vlan-config)# ip igmp snooping querier 198.51.100.1
```

Querier must be configured on the switch.

IGMP snooping is configured for IPv4 multicast over VXLAN EVPN, enabling efficient multicast group management and reducing unnecessary traffic.

Configure IPv6 MLD Snooping

Configure IPv6 MLD Snooping to control IPv6 multicast traffic efficiently on a VXLAN EVPN-enabled switch.

Follow these steps to configure IPv6 MLD Snooping.

Before you begin

Before you begin, ensure that

- VXLAN EVPN is configured.
- The **system mld snooping** command must be configured, followed by a reload.
- Use the **hardware access-list tcam region ing-sup** *size* command to enable the ACL TCAM **ing-sup** region as specified.

```
switch(config)# hardware access-list tcam region ing-sup 768
              switch(config)# hardware access-list tcam region ing-racl 1024
```

- Use the **copy running-config startup-config** command and reload the switch or line cards to apply the configuration changes.

Procedure

Step 1 Run the **ipv6 mld snooping** command in global configuration mode, to enable MLD snooping on the switch.

Example:

```
switch(config)# ipv6 mld snooping
```

Step 2 Run the **ipv6 mld snooping vxlan** command to configure MLD snooping for VXLANs.

Example:

```
switch(config)# ipv6 mld snooping vxlan
```

Step 3 Run the **vlan configuration** *vlan-id* command to enter configuration mode for VLAN 101.

Example:

```
switch(config)# vlan configuration 501
switch(config-vlan-config)#
```

Step 4 Run the **ip mld snooping querier** *querier-ip-address* command to configure the MLD snooping querier for each multicast-enabled VXLAN VLAN.

Example:

```
switch(config-vlan-config)# ipv6 mld snooping querier fe80:0:0:1::1
```

Querier must be configured on the switch.

IPv6 MLD snooping is enabled, ensuring multicast traffic is sent only to interested receivers and reducing unnecessary network flooding.

Commands for multicast Layer 2 EVPN route types

Use these commands to display multicast Layer 2 EVPN route types configuration information:

Command	Purpose
show ip igmp snooping evpn	Displays IGMP snooping information for IPv4 EVPN route type.
show ipv6 mld snooping evpn	Displays MLD snooping information for IPv6 EVPN route type.
show ip igmp snooping evpn smet	Displays IGMP snooping information for IPv4 SMET route type.
show ipv6 mld snooping evpn smet	Displays MLD snooping information for IPv6 SMET route type.
show ip igmp snooping remote groups	Displays IGMP snooping information for remote installed SMET and Joinsync routes.
show ipv6 mld snooping remote groups	Displays MLD snooping information for IPv6 for remote installed SMET and Joinsync routes.
show ip igmp snooping mrouter	Displays IGMP mrouter ports added by SMET.
show ipv6 mld snooping mrouter	Displays MLD mrouter ports added by SMET.
show l2route smet {**topology** *topo-id* | **all**} [**detail**]	Displays L2RIB state of the SMET routes.
Show bgp l2vpn evpn route-type 6	Displays BGP route Type 6 state.

Command output examples for multicast Layer 2 EVPN route Type 6

This topic provide reference examples of command outputs and legend explanations for verifying Multicast Layer 2 EVPN Route Type 6 features.

show ip igmp snooping evpn

Displays whether the EVPN Route Type 6 SMET feature is configured.

```
switch# show ip igmp snooping evpn
IGMP Snoop EVPN Parameters:
Advertise EVPN : Configured
ESI-MH : Disabled
L2RIB : UP [Initialized 1] [handle 0x42ac020]
EVPN-RIB : UP
Last Trigger : ADD AdvertiseEvpnMcast [id:101]
          : Start:00:02:13 [133914.65048] End:00:02:13 [133914.388186]
```

show ip igmp snooping evpn smet

Displays locally originated and remotely received SMET routes and associated versions.

```
switch # show ip igmp snooping evpn smet

VLAN       VNI        Group                   Source                  Local/Remote Flags

11         1011       225.0.0.11              0.0.0.0                 Local        v2
```

```
11         1011       225.0.0.11              0.0.0.0                 Remote       E v2
v3
11         1011       225.0.0.11              11.1.1.108              Local        I v3

11         1011       225.0.0.11              192.0.2.10              Remote       I v3

11         1011       225.0.1.11              0.0.0.0                 Local        v2
```

show ip igmp snooping remote groups

Displays the entire set of ports associated with an IGMP entry. Also, there is a **detail** option to show details of each route and associated ports.

```
switch# show ip igmp snooping remote groups
Type: S - Static, D - Dynamic, R - Router port, F - Fabricpath core port

Vlan  Group Address      Ver  Type  Port list
11    225.0.0.11         v2   S     nve1
        11.1.1.108            S     nve1
11    225.0.1.11         v2   S     nve1
```

show ip igmp snooping mrouter

Displays the router ports that are locally advertised or remotely learned. Remotely learned router port through SMET is marked as 'M' with NVE as the router port.

Note

- The "internally created" peer-link router port is not advertised in SMET.
- If "ip igmp snooping disable-nve-router-port" is not configured, then NVE is added as a static router port.

```
switch # show ip igmp snooping mrouter
Type: S - Static, D - Dynamic, V - vPC Peer Link
      I - Internal, F - Fabricpath core port
      C - Co-learned, U - User Configured
      P - learnt by Peer, M - learnt by SMET
Vlan  Router-port   Type      Uptime      Expires
16    nve1          M         05:33:42    never
17    nve1          M         05:49:10    never
18    nve1          M         05:49:10    never
19    nve1          M         05:49:10    never
20    nve1          M         05:49:10    never
```

show l2route smet all detail

Displays the visibility of the Type 6 SMET route data at the L2RIB level. Remotely learned routes show the associated route-originators for the route, and the flags from each router originator.

```
switch# show l2route smet all detail
V1: MLD version 1, V2: IGMP/MLD version 2, V3: IGMP version 3, I: Include Group Type, E:
Exclude Group Type
Topology    Source Group IP Addresses                 Producer
----------- ------------------------------------------------------------------------

11          (0.0.0.0, 225.0.0.11)                         None
            Sent To: IGMPSN
            Total Counters (V1/V2/V3/I/E): 0/1/1/0/1
```

```
11          (0.0.0.0, 225.0.0.11)                          BGP
            Total Counters (V1/V2/V3/I/E): 0/2/1/0/1
            Originator router IP: 192.0.2.20, Flags: (V2)
            Originator router IP: 198.51.100.11, Flags: (V2,V3,E)

11          (0.0.0.0, 225.0.0.11)                          IGMP
            Sent To: BGP
            Total Counters (V1/V2/V3/I/E): 0/1/0/0/0
            Originator router IP: Local, Flags: (V2)

11          (11.1.1.108, 225.0.0.11)                       None
            Sent To: IGMPSN
            Total Counters (V1/V2/V3/I/E): 0/0/1/1/0

11          (11.1.1.108, 225.0.0.11)                       BGP
            Total Counters (V1/V2/V3/I/E): 0/0/4/4/0
            Originator router IP: 192.0.2.20, Flags: (V3,I)
            Originator router IP: 198.51.100.11, Flags: (V3,I)
            Originator router IP: 198.51.100.13, Flags: (V3,I)
            Originator router IP: 203.0.113.14, Flags: (V3,I)
```

Table 2: Route and Producer types

Route type	Producer	Description
Locally Advertised Routes	IGMP	These routes are locally advertised
Remotely Learned Routes	BGP	These routes show the associated route-originators for the route.
Routes Sent to IGMP	None	-

show bgp l2vpn evpn route-type 6

Displays the details of the received and advertised Type 6 EVPN routes at the BGP level.

```
switch # show bgp l2vpn evpn route-type 6
BGP routing table information for VRF default, address family L2VPN EVPN
Route Distinguisher: 192.0.2.20:32778
BGP routing table entry for [6]:[0]:[0]:[0.0.0.0]:[32]:[225.0.0.11]:[32]:[192.0.2.20]:[2]/144,
 version 204525
Paths: (1 available, best #1)
Flags: (0x000002) (high32 00000000) on xmit-list, is not in l2rib/evpn, is not in HW

  Advertised path-id 1
  Path type: internal, path is valid, is best path, no labeled nexthop, is extd
             Imported to 1 destination(s)
             Imported paths list: L2-1011
  AS-Path: NONE, path sourced internal to AS
    203.0.113.7 (metric 45) from 192.0.2.20 (192.0.2.20)
      Origin IGP, MED not set, localpref 100, weight 0
      Extcommunity: RT:100:1011 ENCAP:8

  Path-id 1 not advertised to any peer
BGP routing table entry for [6]:[0]:[0]:[0.0.0.0]:[32]:[225.0.1.11]:[32]:[192.0.2.20]:[2]/144,
 version 203912
Paths: (1 available, best #1)
Flags: (0x000002) (high32 00000000) on xmit-list, is not in l2rib/evpn, is not in HW

  Advertised path-id 1
  Path type: internal, path is valid, is best path, no labeled nexthop, is extd
             Imported to 1 destination(s)
```

```
           Imported paths list: L2-1011
AS-Path: NONE, path sourced internal to AS
  203.0.113.7 (metric 45) from 192.0.2.20 (192.0.2.20)
    Origin IGP, MED not set, localpref 100, weight 0
    Extcommunity: RT:100:1011 ENCAP:8
```