Troubleshoot EVPN VxLAN TRM on Catalyst 9000 Switches

Contents

Introduction

Prerequisites

- Requirements
- Components Used

Background Information

Terminology

Verify

- Verification Common to all Scenarios
  - Verify NVE peering
  - Verify the RPF Interface in The Tenant VRF
  - Verify Multicast Control Plane Uses BGP
  - Verify MDT Group
- Scenario 1. AnyCast RP (SPT-only trees) IPv4 & IPv6
  - Network Diagram
  - Verify BGP EVPN and MVPN Routes
  - Verify the TRM group Leaf-01 (FHR)
  - Verify the TRM group Leaf-02 (LHR)
- Scenario 2: PIM SSM in the Fabric
  - Network Diagram
  - Verify the Sequence of Events Required for this Scenario
- Scenario 3: Single RP Inside the Fabric (Regular Sparse-Mode)
  - Network Diagram
  - Verify the Sequence of Events Required for this Scenario
- Scenario 4: RP Outside the Fabric (RP Imported from Border Leaf-02 from IP space)
  - Network Diagram
  - Verify Border switch imports from IP to Fabric
- Scenario 5: Data MDT
  - Verify MDT Data Group
  - Debug MDT Data Group

Troubleshoot

- Undetected Multicast Sources
- Other Helpful Debugs
- Sources and Receivers Outside the Fabric
- eBGP Multiple AS (Spine to Spine) Topology
- Register Tunnel with Symmetric L2VNI (FHR Stuck in PIM Register State)

Related Information

Introduction

This document describes how to troubleshoot issues with TRM (Tenant Routed Multicast) over EVPN

VxLAN.

Prerequisites

- It is recommended that you are familiar with Unicast EVPN VxLAN feature, BGP and MVPN (Multicast Virtual Private Network).
- Additionally, you must understand how multicast operates, and multicast concepts

Requirements

This guide assumes BGP, NVE peers are already correct. If there are issues with basic EVPN VxLAN bring up (Unicast ping failure, BGP, NVE peers down, and so on) please reference BGP, EVPN, route/switch troubleshoot guides as necessary.

Feature availability in each code release

Release	**Feature**
17.1.1	TRMv4 with Anycast RP
17.3.1	TRMv4 with External RP or Single RP
17.3.1	TRMv6 with Anycast RP
17.3.1	TRMv6 with External RP or Single RP
17.3.1	TRMv4 with MVPN interworking (profile11) with Single RP on the Fabric Side
17.6.2 & 17.7.1	TRMv4 Data mdt with Anycast RP, External RP, or Single RP

Components Used

The information in this document is based on these software and hardware versions:

- C9300
- C9400
- C9500
- C9600

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. If your network is live, ensure that you understand the potential impact of any command.

Note: Consult the appropriate configuration guide for the commands that are used in order to enable

---

these features on other Cisco platforms.

Background Information

To configure EVPN TRM consult: BGP EVPN VXLAN Configuration Guide, Cisco IOS XE Amsterdam 17.3.x

Tenant Routed Multicast (TRM) is a BGP-EVPN based solution that enables multicast routing between sources and receivers connected on VTEPS in VxLAN fabric [RFC7432]. TRM relies on routes present in the unicast EVPN to discover multicast source and multicast RP. Like with NG-MVPN, Multicast source and receiver information is propagated by BGP protocol among VTEPs configured with the BGP MVPN address family. No PIM/IGMP packets are sent to VxLAN fabric from a TRM VTEP.

The key problem that TRM solves is the ability of multicast senders and receivers that are located in different VLANs but in the same VRF, to be able to communicate with each other. Without TRM, multicast traffic is sent as part of the same BUM (Broadcast, Unicast, and Multicast) infrastructure in the underlay, which can be a multicast tree or an ingress-replication. This infrastructure is built per VLAN and as a result, while multicast sources and receivers on the same VLAN can communicate, those that are on different VLANs cannot. With TRM, Multicast is moved out of BUM and clubbed together under the parent VRF. Due to this, multicast communication is fully enabled, irrespective of the VLANs in which the source or the receiver reside.

TRM provides multi-tenancy aware multicast forwarding between senders and receivers within the same or different subnets local or across VTEPs. See guide BGP EVPN VXLAN Configuration Guide, Cisco IOS XE Amsterdam 17.3.x for more details

How to Orient yourself in this guide:

- Guide is broken into 4 Scenarios based on the RP location.
- A scenario can refer to CLI examples not directly in the section you are in. For example, SSM Scenario 2 refers you to to Scenario 1 to understand how to read certain CLIs.
- **Only Scenario 1 covers IPv4 & IPv6** as the concepts are fundamentally the same for both address families.
- **The requirements listed in these Scenarios assume the Source and Receiver are directly attached to the VTEPs** (See Related Information Section 'Sources and Receivers Outside the Fabric' for more on this).

Scenarios	**IPv4/v6**	**What is covered in each**
Common details for all other scenarios	IPv4	Always required: MVPN underlay and NVE are healthy, RPF check to any TRM Source is the L3VNI.
AnyCast RP (each VTEP is an RP with a common RP IP)	IPv4/v6	BGP, PIM, IGMP, MFIB, and FED commands in full detail for both IPv4/v6 plus capture examples
No RP (SSM Overlay)	IPv4	SSM Specific information. (Refer to scenario 1 for common information)

RP Inside the Fabric (one common RP for the fabric)	IPv4	BGP, PIM, IGMP, MFIB, and FED commands in full detail for IPv4
RP Outside the Fabric (the RP is not in the fabric)	IPv4	IP to Fabric Border Specific information. (Refer to scenario 3 for common information)
RP Inside the Fabric (one common RP for the fabric) with symmetric L2VNI	IPv4	Caveats for use of a single RP in fabric when the VNI is on both Sender and Receiver VTEPs. (Refer to Scenario 3 for common information)

In this troubleshooting document, comments have been added at the end of certain lines of the outputs of show commands. This has been done to highlight or explain a specific aspect of that line of output. If a comment begins in a new line, then it refers to the line of output that precedes the comment. This notation has been used throughout the document to highlight the comments inside the outputs of show commands:

<#root>

```
<-- Text highlighted in this format inside a command's output represents a comment.


    This is done for explanation purpose only and is not part of the command's output.
```

Terminology

EVPN	Ethernet Virtual Private Network	Extension that allows BGP to transport Layer 2 MAC and Layer 3 IP information is EVPN and uses Multi-Protocol Border Gateway Protocol (MP-BGP) as the protocol to distribute reachability information that pertains to the VXLAN overlay network.
VXLAN	Virtual Extensible LAN (Local Area Network)	VXLAN is designed to overcome the inherent limitations of VLANs and STP. It is a proposed IETF standard [RFC 7348] to provide the same Ethernet Layer 2 network services as VLANs do, but with greater flexibility. Functionally, it is a MAC-in- UDP encapsulation protocol that runs as a virtual overlay on a Layer 3 underlay network.
VTEP	Virtual Tunnel Endpoint	This is the device that does the encapsulation and de-encapsulation
NVE	Network Virtualization Edge (Interface)	Logical interface where the encapsulation and de-encapsulation occur
VNI	VXLAN network identifier	Uniquely identifies each Layer 2 subnet or segment. There are two types of VNI:

		Symmetric (L2VNI): VTEPs have same VNI Asymmetric (L3VNI): VTEPs do not have same VNI and are routed via a single common VNI.
MDT	Multicast Distribution Tree	The multicast tree built between VTEPs for encapsulation and tunnelling of Tenant Multicast Traffic.
BUM	Broadcast, Unknown Unicast, Multicast	BUM traffic is sent via the Mcast group tied to the VNI under the NVE configuration.
RP	Rendezvous Point	A role that a device performs when in PIM Sparse Mode. The common meeting point for Multicast sources and Receivers.
AnyCast (RP)	AnyCast Rendezvous Point	Two or more RPs are configured with the same IP address on loopback interfaces. FHR registers to nearest RP based on Unicast routing.
RPT (Tree)	Root Path Tree	Also called the Shared or *,G tree. This path is toward the RP
SPT (Tree)	Shortest Path Tree	The shortest path to the Source as determined by the Unicast Routing Table
FHR	First Hop Router	The device that is Directly Connected (ARP adjacent) to the Source. The FHR Registers source info with the RP.
LHR	Last Hop Router	The device where the Receiver is connected
RPF	Reverse Path Forwarding	The Unicast path back to the Source. Incoming multicast packets are not accepted/forwarded unless they are received the same path as the unicast routing table. ('ip multicast multipath' use cases excluded).
MRIB	Multicast Routing Information Base	The software multicast routing table, also called mroute table. The control plane maintains state of multicast entries including: Source/Group information, Incoming outgoing interfaces, uptime and expiration timers. Contains (*, G), (S, G), and (*, G/m) MRIB entries
MFIB	Multicast Forwarding Information Base	The multicast equivalent of CEF. Populated by updates from MRIB, and this is the forwarding plane representation of the multicast route. Contains (*, G), (S, G), and (*, G/m) MFIB entries.
FED	Forwarding Engine Driver	The hardware forwarding data plane. Populated by information from the forwarding plane, FED is the driver that programs the the Application

		Specific Integrated Circuit (ASIC) (hardware) layer and contains the hardware representation of a multicast entry. FED contains the information that is used to write, rewrite, and forward packets. When validating a multicast entry MRIB, MFIB and FED commands should be used to verify the Control, Forwarding, and Data Planes respectively.
IIF	Incoming Interface	PIM enabled interface which is also the unicast RPF upstream path back to the Source. (seen in show ip mroute)
OIF	Outgoing Interface	PIM enabled interface that is downstream toward the Receiver. (seen in show ip mroute)

Verify

Verification Common to all Scenarios

This first section covers the base requirements **necessary for any of the scenarios**.

- Ensure the required NVE peers are up
- Ensure that RPF interface toward the Source in the Tenant VRF is the L3VNI SVI. If the RPF interface is not the L3VNI SVI, BGP does not send a type-7 join route. In any scenario, RPF interface must point to this Interface.
- Ensure the Underlay path (MDT tunnel) between peers is complete.
- Ensure that BGP is used for multicast control-plane (use MVPN versus PIM)

Note: This section is applicable to both IPv4 & IPv6 Tenant multicast verification.

Verify NVE peering

Check to ensure the NVE peers are up between VTEPs for any of the scenarios in this guide

- NVE peers are formed by addresses learned from BGP.

```
<#root>
Leaf-01#
sh nve peers

Interface VNI    Type Peer-IP       RMAC/Num_RTs   eVNI    state flags UP time
nve1      50901  L3CP 172.16.254.4  7c21.0dbd.9548 50901   UP    A/-/4 01:54:11 <-- IPv4 peering


with Leaf 02
```

```
nve1      50901  L3CP 172.16.254.4  7c21.0dbd.9548 50901   UP    A/M/6 17:48:36 <-- IPv6 peering with Le


Leaf-02#
sh nve peers

Interface VNI    Type Peer-IP       RMAC/Num_RTs   eVNI    state flags UP time
nve1      50901  L3CP 172.16.254.3  10b3.d56a.8fc8 50901   UP    A/-/4 01:55:44 <-- IPv4 peering with Le


nve1      50901  L3CP 172.16.254.3  10b3.d56a.8fc8 50901   UP    A/M/6 17:56:19 <-- IPv6 peering with Le
```

Verify the RPF Interface in The Tenant VRF

If this interface is any interface other than the L3VNI SVI, BGP does not originate an MVPN Type-7 join.

- If you do not see this interface, please confirm that there is no issue with the configuration that would make the route back to the Source an interface that is not the L3VNI.

```
<#root>
Leaf-03#
sh ip rpf vrf green 10.1.101.11 <-- Multicast source IP

RPF information for ? (10.1.101.11)

 RPF interface: Vlan901                <-- RPF interface is the L3VNI SVI


RPF neighbor: ? (172.16.254.3)        <-- Underlay Next hop IP


RPF route/mask: 10.1.101.0/24         <-- Network prefix for the Source

  RPF type: unicast (bgp 65001)
  Doing distance-preferred lookups across tables
  RPF topology: ipv4 multicast base, originated from ipv4 unicast base
```

Verify Multicast Control Plane Uses BGP

- **mdt overlay use-bgp:** informs devices to use BGP MVPN type 5/6/7 as the signal protocol (versus PIM messages)
- **spt-only:** additional keyword informs device to use only SPT trees in AnyCast RP Scenario. Since each VTEP is an RP **no MVPN Type-6 route is used.**

<#root>

```
Leaf-01
!
vrf definition green
rd 1:1
!
address-family ipv4
mdt auto-discovery vxlan
```

mdt default vxlan 239.1.1.1 <-- Defines MDT default underlay group address

mdt overlay use-bgp [spt-only] <-- Required for VTEP to use MVPN Type 5/6/7 versus PIM for multicast

Verify MDT Group

The MDT group is common to all scenarios as this is the outer tunnel group the TRM group is encapsulated in.

Check that the MDT group is correctly programmed on Source side

- Incoming interface of MDT group is the Source side Loopback
- Outgoing interface of MDT group is the Underlay Interface

Verify Leaf-01: the MDT mroute is correct in MRIB/MFIB

<#root>

```
Leaf-01#
```

sh ip mroute 239.1.1.1 172.16.254.3

```
(
```

172.16.254.3

```
,
```

239.1.1.1

```
), 00:46:35/00:02:05, flags: FTx
  Incoming interface:
```

Loopback1

```
, RPF nbr
```

0.0.0.0

<-- IIF is local loopback with 0.0.0.0 RPF indicating local

```
  Outgoing interface list:
```

GigabitEthernet1/0/2

, Forward/Sparse, 00:46:35/00:03:12

<-- OIF is the underlay uplink

Leaf-01#

sh ip mfib 239.1.1.1 172.16.254.3

(172.16.254.3,239.1.1.1) Flags: HW

  SW Forwarding: 2/0/150/0, Other: 1/1/0

HW Forwarding: 1458/0/156/0

, Other: 0/0/0

<-- Hardware counters indicate the entry is operating in hardware and forwarding packets

Null0 Flags: A NS **<--- Null0 (originated locally)**

GigabitEthernet1/0/2

 Flags: F NS

<-- OIF is into the Underlay (Global route table)

    Pkts: 0/0/1 Rate: 0 pps

Verify Leaf-01: FED entries for the MDT group

<#root>

Leaf-01#

sh platform software fed switch active ip mfib 239.1.1.1/32 172.16.254.3 detail <-- the detail option gi

MROUTE ENTRY

vrf 0

 (

172.16.254.3, 239.1.1.1/32

)

<-- vrf 0 = global for this MDT S,G pair

  HW Handle: 139738317079128 Flags:

RPF interface: Null0

(1)):

<-- Leaf-01 the Source (Null0)

HW Handle:139738317079128 Flags:A
Number of OIF: 2
Flags: 0x4

Pkts : 71 <-- packets that used this adjacency (similar to mfib command, but shown at the FED

OIF Details:

Null0 A

<-- The incoming interface is Local Loopback1 and A-Accept flag set

GigabitEthernet1/0/2

F

NS

<-- The Underlay Outgoing Interface and F-Forward flag set

Htm: 0x7f175cc0beb8 Si: 0x7f175cc0a6b8

Di: 0x7f175cc09df8

Rep_ri: 0x7f175cc0a1d8

<-- The DI (dest index) handle

DI details

Handle:0x7f175cc09df8 Res-Type:ASIC_RSC_DI Res-Switch-Num:255 Asic-Num:255 Feature-ID:AL_FID_L3_MULTICAS
priv_ri/priv_si Handle:(nil) Hardware Indices/Handles:

index0:0x538d

mtu_index/l3u_ri_index0:0x0

index1:0x538d

mtu_index/l3u_ri_index1:0x0

Brief Resource Information (ASIC_INSTANCE# 1)

Destination index = 0x538d

```
pmap = 0x00000000 0x00000002
pmap_intf : [GigabitEthernet1/0/2]  <-- FED has the correct programming for the OIF

============================================================
```

Check that the MDT group is correctly programmed on Receiver side

- Incoming interface of MDT group is the RPF interface back to the Source side Loopback
- Outgoing interface of MDT group is Encap/Decap Tunnel interface

Verify Leaf-02: the MDT mroute is correct in MRIB/MFIB

```
<#root>
Leaf-02#
sh ip mroute 172.16.254.3 239.1.1.1         <-- This is the Global MDT group

(
172.16.254.3
,
239.1.1.1
), 00:23:35/00:01:09, flags: JTx
<-- Source is Leaf-01 Lo1 IP

  Incoming interface: GigabitEthernet1/0/2, RPF nbr 172.16.24.2
  Outgoing interface list:

Tunnel0
, Forward/Sparse, 00:23:35/00:00:24
<-- Decap Tunnel

Leaf-02#
sh ip mfib 239.1.1.1 172.16.254.3

Default                                          <-- Global routing table

  (172.16.254.3,239.1.1.1) Flags: HW
    SW Forwarding: 1/0/150/0, Other: 0/0/0

HW Forwarding: 5537/0/168/0, Other: 0/0/0  <-- Hardware counters indicate the entry is operating in hard

    GigabitEthernet1/0/2 Flags: A             <-- Accept via Underlay (Global) interface
```

```
    Tunnel0, VXLAN Decap Flags: F NS           <-- Forward to VxLAN decap Tunnel
      Pkts: 0/0/1 Rate: 0 pps
```

Verify Leaf-02: FED entries for the MDT group

```
<#root>
Leaf-02#
sh platform software fed switch active ip mfib 239.1.1.1/32 172.16.254.3 detail
MROUTE ENTRY
 vrf 0
 (
172.16.254.3, 239.1.1.1/32
)
<-- vrf 0 = global for this MDT S,G pair

  HW Handle: 140397391831832 Flags:
RPF interface: GigabitEthernet1/0/2
(57)):
<-- RPF interface to 172.16.254.3

  HW Handle:140397391831832 Flags:A
Number of OIF: 2
Flags: 0x4
Pkts : 1585           <-- packets that used this adjacency (similar to mfib command, but shown at the FE

OIF Details:
     Tunnel0 F NS               <-- Send to decap tunnel to remove VxLAN header

          (Adj: 0x73 )          <-- Tunnel0 Adjacency

     GigabitEthernet1/0/2 A     <-- Accept MDT packets from this interface

Htm: 0x7fb0d0f1f388 Si: 0x7fb0d0f1dc08 Di: 0x7fb0d0ed0438 Rep_ri: 0x7fb0d0ed07a8
RI details                      <-- Rewrite Index is used for VxLAN decapsulation

----------
```

```
Handle:0x7fb0d0ed07a8 Res-Type:ASIC_RSC_RI_REP Res-Switch-Num:255 Asic-Num:255 Feature-ID:AL_FID_L3_MUL
priv_ri/priv_si Handle:(nil) Hardware Indices/Handles: index0:0x38 mtu_index/l3u_ri_index0:0x0 index1:0

Brief Resource Information (ASIC_INSTANCE# 0)
---------------------------------------
ASIC# 0
Replication list :
------------------

Total #ri : 6
Start_ri : 56
Common_ret : 0

Replication entry

rep_ri 0x38

 #elem = 1
0)

ri[0]=0xE803

 Dynamic port=88ri_ref_count:1 dirty=0

Leaf-02#

sh platform hardware fed sw active fwd-asic resource asic all rewrite-index range 0xE803 0xE803


ASIC#:0 RI:59395

Rewrite_type:

AL_RRM_REWRITE_L2_PAYLOAD_

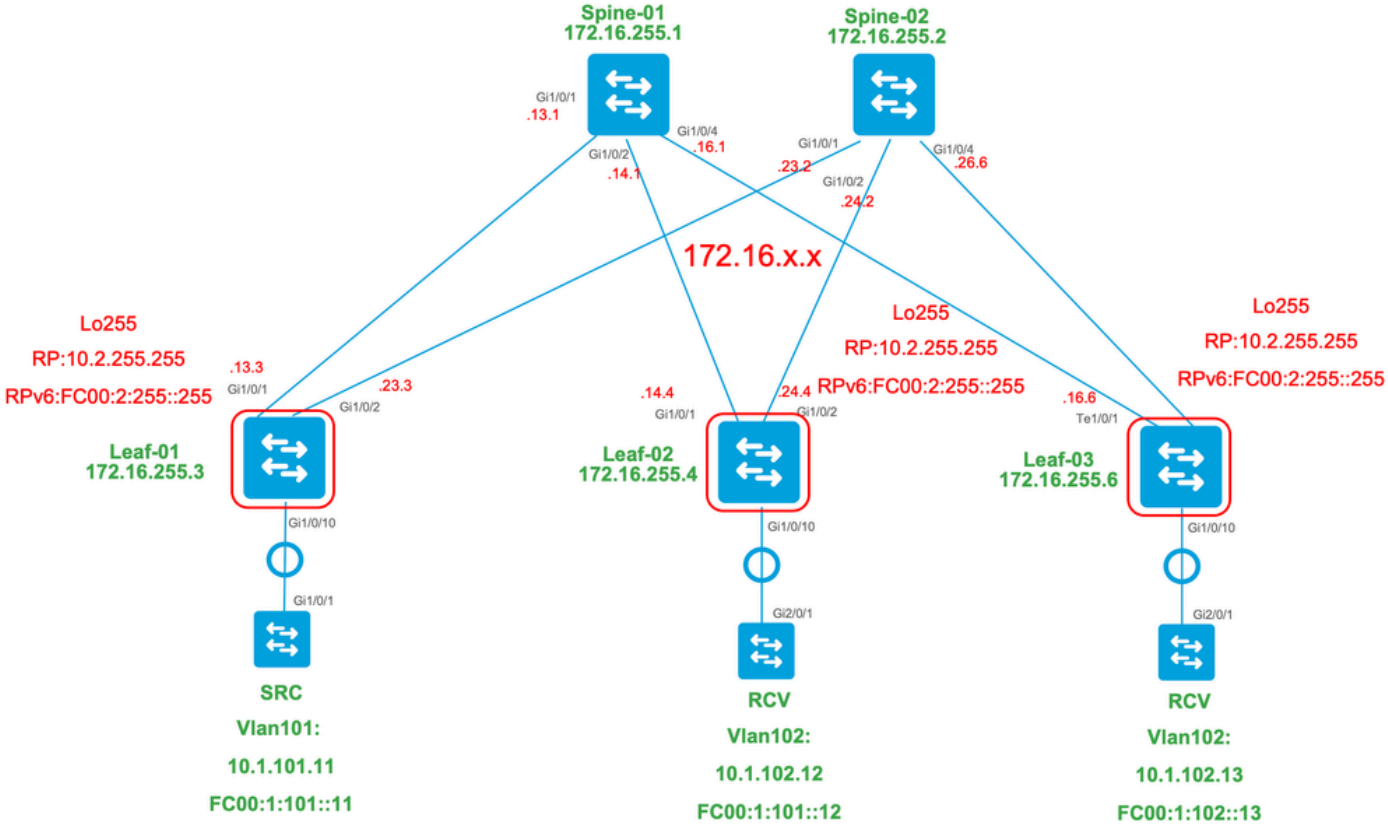
IPV4_EVPN_DECAP

(118) Mapped_rii:LVX_EVPN_DECAP(246)
<...snip...>
```

Scenario 1. AnyCast RP (SPT-only trees) IPv4 & IPv6

In this mode there is an RP located on every VTEP. These VTEPs do not synchronize learned Sources via MSDP and there is no Shared tree. Instead the MDT mode uses BGP information to create only SPT multicast trees. This mode is interchangeably called as SPT-only mode or distributed Anycast-RP mode. In this mode, each VTEP is the PIM RP. Thus, the (*,G) tree at each site is truncated at the local VTEP itself. There is no need to send (*,G) joins or MVPN RT-6 across the fabric.

Network Diagram



For this mode, consider 3 BGP route-types:

1. EVPN Route-type 2. This allows the other PEs that need to construct a C-Multicast route (MVPN type6/7) back to the origin PE, to attach the proper C-multicast Import RT so that the originator PE can import the C-Multicast route (RFC 6514 11.1.3) [RFC6514]. The usage of this VRI is dependent on the command "**mdt overlay use-bgp**" VRF command.
2. MVPN Route-type 5. This is the same as in MVPN, and is the advertisement of a mulitcast Source/Group available
3. MVPN Route-type 7. Information from the IGMP or MLD layer & from EVPN Type-2 are used to create this BGP type join. The Type-7 drives the creation of the MRIB OIF on the Source side.

EVPN Type-2 requirements:

1. Directly Connected multicast source comes online.
2. FHR (source VTEP) verifies ARP (or ND) and CEF adjacency (confirms the source is Directly Connected).
3. FHR originates the EVPN Type-2 BGP update

MVPN Type-5 requirements:

1. The requirement for the source Direct connect is resolved
2. RP is local so FHR registers to itself
3. FHR originates MVPN Type-5 BGP update

MVPN Type-7 requirements:

1. EVPN Type-2 entry is present (required to construct the C-Multicast route type-7 with correct VRI and sent from Source VTEP)
2. MVPN Type-5 entry is present (required to resolve what Source/Group pair is available for SPT join)
3. IGMP or MLD membership report has been received and processed by the LHR VTEP

4. LHR VTEP RPF interface is the Fabric L3VNI interface

Tip: At the egress LHR VTEP PIM checks the path toward the Source. PIM must find a route in the RIB that is the L3VNI as the RPF interface. If the L3VNI is not configured properly, is down, and so on. the VTEP does not attempt to create the type-7 BGP join.

Verify BGP EVPN and MVPN Routes

Verify Leaf-01: the EVPN Type-2 is created

<#root>

IPv4

Leaf-01#

sh bgp l2vpn evpn all route-type 2 0 F4CFE24334C5 10.1.101.11

...or you can also use:

Leaf-01#

sh bgp l2vpn evpn detail [2][172.16.254.3:101][0][48][F4CFE24334C5][32][10.1.101.11]/24

BGP routing table entry for [2][172.16.254.3:101][0][48][F4CFE24334C5][32][10.1.101.11]/24, version 6
Paths: (1 available, best #1,

table evi_101

)
  Advertised to update-groups:
     1
  Refresh Epoch 1
  Local

:: (via default) from 0.0.0.0 (172.16.255.3) <-- Leaf-01 locally created

      Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
      EVPN ESI: 00000000000000000000, Label1 10101, Label2 50901
      Extended Community: RT:1:1 RT:65001:101 MVPN AS:65001:0.0.0.0

        MVPN VRF:172.16.255.3:2

 ENCAP:8 Router MAC:10B3.D56A.8FC8

 <-- MVPN VRI RT is part of the EVPN Type-2

      Local irb vxlan vtep:

        vrf:green, l3-vni:50901 <-- Vrf and VxLAN tag

        local router mac:10B3.D56A.8FC8

        core-irb interface:Vlan901 <-- L3VNI SVI

```
        vtep-ip:172.16.254.3                <-- Leaf-01 VTEP


      rx pathid: 0, tx pathid: 0x0
      Updated on Dec 16 2020 17:40:29 UTC



### IPv6 ###

Leaf-01#

sh bgp l2vpn evpn all route-type 2 0 F4CFE24334C1 FC00:1:101::11

...or you can also use:
Leaf-01#

sh bgp l2vpn evpn detail [2][172.16.254.3:101][0][48][F4CFE24334C1][128][FC00:1:101::11]/36

BGP routing table entry for [2][172.16.254.3:101][0][48][F4CFE24334C1][128][FC00:1:101::11]/36, version
Paths: (1 available, best #1, table evi_101)
  Advertised to update-groups:
     1
  Refresh Epoch 1
  Local


:: (via default) from 0.0.0.0 (172.16.255.3)                    <-- Leaf-01 locally created


      Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
      EVPN ESI: 00000000000000000000, Label1 10101, Label2 50901
      Extended Community: RT:1:1 RT:65001:101 MVPN AS:65001:0.0.0.0


 MVPN VRF:172.16.255.3:2

 ENCAP:8 Router MAC:10B3.D56A.8FC8

<-- MVPN VRI RT is part of the EVPN Type-2


      Local irb vxlan vtep:


   vrf:green, l3-vni:50901


        local router mac:10B3.D56A.8FC8


core-irb interface:Vlan901       <-- L3VNI SVI



 vtep-ip:172.16.254.3             <-- Leaf-01 VTEP


      rx pathid: 0, tx pathid: 0x0
      Updated on Mar 22 2021 19:54:18 UTC
```

Verify Leaf-01: ARP/IPv6 ND and EVPN Debugs show ARP/ND is learned, then Route-type 2 created and

sent

<#root>

IPv4

Leaf-01#

sh debugging

ARP:

ARP packet debugging is on

BGP L2VPN EVPN:

BGP updates debugging is on for address family: L2VPN E-VPN
BGP update events debugging is on for address family: L2VPN E-VPN

*Dec 17 17:00:06.480:

IP ARP: rcvd rep src 10.1.101.11 f4cf.e243.34c5

, dst 10.1.101.11 Vlan101

tableid 2 <-- Multicast Source ARP

*Dec 17 17:00:06.481:

BGP: EVPN Rcvd pfx: [2]

[172.16.254.3:101][0][48][F4CFE24334C5][32][10.1.101.11]/24, net flags: 0

<-- BGP Triggered Type-2 creation

*Dec 17 17:00:06.481:

TRM communities added to sourced RT2 <-- TRM extended VRI communities being injected into EVPN Type-2

*Dec 17 17:00:06.481:

BGP(10): update modified for [2]

[172.16.254.3:101][0][48][F4CFE24334C5][32][10.1.101.11]/30

<-- Modifying the update

*Dec 17 17:00:06.481: BGP(10): 172.16.255.1 NEXT_HOP set to vxlan local vtep-ip 172.16.254.3 for net [2
*Dec 17 17:00:06.481: BGP(10): update modified for [2][172.16.254.3:101][0][48][F4CFE24334C5][32][10.1.
*Dec 17 17:00:06.481: BGP(10): (base) 172.16.255.1

send UPDATE

(format)

[2]

[172.16.254.3:101][0][48][F4CFE24334C5][32][10.1.101.11]/30, next 172.16.254.3, metric 0, path Local, e

MVPN VRF:172.16.255.3:2

ENCAP:8 Router MAC:10B3.D56A.8FC8

<--- Final update sent to RR with standard EVPN community info and required MVPN community attributes

IPv6

Leaf-01#

debug ipv6 nd

ICMP Neighbor Discovery events debugging is on
ICMP ND HA events debugging is ON

IPv6 ND:

Mar 23 14:29:51.935:

ICMPv6-ND: (Vlan101,FC00:1:101::11) Resolution request

Mar 23 14:29:51.935: ICMPv6-ND: (Vlan101,FC00:1:101::11) DELETE -> INCMP
Mar 23 14:29:51.935: ICMPv6-ND HA: in Update Neighbor Cache: old state 6 new state 0
Mar 23 14:29:51.935: ICMPv6-ND HA: add or delete entry not synced as no peer detected
Mar 23 14:29:51.936: ICMPv6-ND: (Vlan101,FC00:1:101::11) Sending NS
Mar 23 14:29:51.936: ICMPv6-ND: (Vlan101,FC00:1:101::11) Queued data for resolution
Mar 23 14:29:51.953:

ICMPv6-ND: (Vlan101,FC00:1:101::11) Received NA from FC00:1:101::11

Mar 23 14:29:51.953:

ICMPv6-ND: Validating ND packet options: valid

Mar 23 14:29:51.953:

ICMPv6-ND: (Vlan101,FC00:1:101::11) LLA f4cf.e243.34c1

Mar 23 14:29:51.953: ICMPv6-ND HA: modify entry not synced as no peer detected
Mar 23 14:29:51.953:

ICMPv6-ND: (Vlan101,FC00:1:101::11) INCMP -> REACH <-- peer is reachable

Leaf-01#

debug bgp l2vpn evpn updates

Leaf-01#

debug bgp l2vpn evpn updates events

BGP L2VPN EVPN:

Mar 23 14:11:56.462:

BGP: EVPN Rcvd pfx: [2][172.16.254.3:101][0][48][F4CFE24334C1][128][FC00:1:101::11]/36,

net flags: 0

<-- BGP Triggered Type-2 creation

Mar 23 14:11:57.462:

TRM communities added to sourced RT2

ar 23 14:11:57.474:

BGP(10): update modified for [2]

[172.16.254.3:101][0][48][F4CFE24334C1][128]

[FC00:1:101::11]/42

Mar 23 14:11:57.474: BGP(10): 172.16.255.1 NEXT_HOP set to vxlan local vtep-ip 172.16.254.3 for net [2]
Mar 23 14:11:57.474: BGP(10): update modified for [2][172.16.254.3:101][0][48][F4CFE24334C1][128][FC00:
Mar 23 14:11:57.474: BGP(10): (base) 172.16.255.1

send UPDATE

(format)

[2]

[172.16.254.3:101][0][48][F4CFE24334C1][128][FC00:1:101::11]/42, next 172.16.254.3, metric 0, path Loca

MVPN VRF:172.16.255.3:2

ENCAP:8 Router MAC:10B3.D56A.8FC8

<--- Final update sent to RR with standard EVPN community info and required MVPN community attributes

Verify Leaf-02: Source side **Route-type 2** is learnt in BGP on Receiver side

<#root>

IPv4

Leaf-02#

sh bgp l2vpn evpn all | b 10.1.101.11

* i

[2]

[172.16.254.3:101][0][48][F4CFE24334C5][32][10.1.101.11]/24

<-- Remote VTEP route-type 2

172.16.254.3 0 100 0 ?

***>i 172.16.254.3 0 100 0 ? <-- IP of Leaf01 Lo1**

Leaf-02#

sh bgp l2vpn evpn route-type 2 0 F4CFE24334C5 10.1.101.11

...or you can also use:
Leaf-02#

sh bgp l2vpn evpn detail [2][172.16.254.3:101][0][48][F4CFE24334C5][32][10.1.101.11]/24

BGP routing table entry for [2][172.16.254.3:101][0][48][F4CFE24334C5][32][10.1.101.11]/24, version 175
Paths: (2 available, best #2, table

EVPN-BGP-Table) <-- In BGP EVPN table
Flag: 0x100

Not advertised to any peer
Refresh Epoch 2
Local

172.16.254.3

(metric 3) (via default) from 172.16.255.2 (172.16.255.2)
Origin incomplete, metric 0, localpref 100, valid, internal
EVPN ESI: 00000000000000000000, Label1 10101,

Label2 50901

Extended Community: RT:1:1 RT:65001:101

MVPN AS:65001:0.0.0.0

MVPN VRF:172.16.255.3:2

ENCAP:8

Router MAC:10B3.D56A.8FC8

Originator: 172.16.255.3, Cluster list: 172.16.255.2
rx pathid: 0, tx pathid: 0
Updated on Dec 14 2020 19:58:57 UTC

MVPN AS:65001:0.0.0.0 <-- MVPN Autonomous System
MVPN VRF:172.16.255.3:2 <-- VRI Extended Community to be used in MVPN Type-7
Router MAC:10B3.D56A.8FC8 <-- Leaf-01 RMAC
Label2 50901 <-- L3VNI 50901

IPv6

Leaf-02#

sh bgp l2vpn evpn all | b FC00:1:101::11

* i [2][172.16.254.3:101][0][48][F4CFE24334C1][128][FC00:1:101::11]/36
172.16.254.3 0 100 0 ?

***>i 172.16.254.3 0 100 0 ? <-- IP of Leaf01 Lo1**

Leaf-02#

sh bgp l2vpn evpn route-type 2 0 F4CFE24334C1 FC00:1:101::11

...or you can also use:
Leaf-02#

sh bgp l2vpn evpn detail [2][172.16.254.3:101][0][48][F4CFE24334C1][128][FC00:1:101::11]/36

BGP routing table entry for

[2]

[172.16.254.3:101][0][48][

F4CFE24334C1

][128][

FC00:1:101::11

]/36, version 659
Paths: (2 available, best #2,

table EVPN-BGP-Table

)

<-- In BGP EVPN table

  Flag: 0x100
  Not advertised to any peer
  Refresh Epoch 2
  Local

172.16.254.3

 (metric 3) (via default) from 172.16.255.2 (172.16.255.2)
      Origin incomplete, metric 0, localpref 100, valid, internal
      EVPN ESI: 00000000000000000000, Label1 10101,

Label2 50901

      Extended Community: RT:1:1 RT:65001:101 MVPN

AS:65001:0.0.0.0

MVPN VRF:172.16.255.3:2

 ENCAP:8

Router MAC:10B3.D56A.8FC8

      Originator: 172.16.255.3, Cluster list: 172.16.255.2
      rx pathid: 0, tx pathid: 0
      Updated on Mar 23 2021 14:11:57 UTC

MVPN AS:65001:0.0.0.0 <-- MVPN Autonomous System

```
MVPN VRF:172.16.255.3:2   <-- VRI Extended Community to be used in MVPN Type-7
Router MAC:10B3.D56A.8FC8 <-- Leaf-01 RMAC
Label2 50901              <-- L3VNI 50901
```

Verify Leaf-02: Source Route-type 5 is learnt in BGP on Receiver VTEP Leaf-02

<#root>

IPv4

Leaf-02#

sh bgp ipv4 mvpn all route-type 5 10.1.101.11 226.1.1.1

...or you can also use:
Leaf-02#

sh bgp ipv4 mvpn detail [5][1:1][10.1.101.11][226.1.1.1]/18

BGP routing table entry for

[5]

[1:1]

[10.1.101.11][226.1.1.1]

/18, version 72

<-- Type-5 contains advertised S,G pair

Paths: (2 available, best #1,

table MVPNv4-BGP-Table

, not advertised to EBGP peer)

<-- In BGP IPv4 MVPN table

  Flag: 0x100
  Not advertised to any peer
  Refresh Epoch 1
  Local

    172.16.255.3

 (metric 3) from 172.16.255.2 (172.16.255.2)

<-- Loopback0 of Leaf-01

      Origin incomplete, metric 0, localpref 100, valid, internal
      Community: no-export
      Extended Community: RT:1:1

Originator: 172.16.255.3

, Cluster list: 172.16.255.2

```
      rx pathid: 0, tx pathid: 0
      Updated on Dec 15 2020 16:54:53 UTC


### IPv6 ###


Leaf-02#
sh bgp ipv6 mvpn all route-type 5 FC00:1:101::11 FF06:1::1
...or you can also use:
Leaf-02#
sh bgp ipv6 mvpn detail [5][1:1][FC00:1:101::11][FF06:1::1]/42
BGP routing table entry for
[5]
[1:1]
[FC00:1:101::11][FF06:1::1]
/42, version 11
<-- Type-5 contains advertised S,G pair


Paths: (2 available, best #1,
table MVPNV6-BGP-Table
, not advertised to EBGP peer)
<-- In BGP IPv6 MVPN table


  Flag: 0x100
  Not advertised to any peer
  Refresh Epoch 1
  Local


172.16.255.3
 (metric 3) from 172.16.255.2 (172.16.255.2)
<-- Loopback0 of Leaf-01


      Origin incomplete, metric 0, localpref 100, valid, internal
      Community: no-export
      Extended Community: RT:1:1


 Originator: 172.16.255.3
, Cluster list: 172.16.255.2


      rx pathid: 0, tx pathid: 0
      Updated on Mar 23 2021 15:13:06 UTC
```

Verify Leaf-02: has needed BGP info from Leaf-01 to create the Type-7. Final requirement is IGMP or

MLD has processed a membership report which informs the VTEP there is an interested Receiver.

<#root>

IPv4

Leaf-02#

sh ip igmp snooping groups vlan 102

```
Vlan     Group                Type            Version         Port List
------------------------------------------------------------------------
102      226.1.1.1
```

igmp

```
            v2
```

Gi1/0/10

<-- Receiver joined on Gi1/0/10

IPv6

Leaf-02#

sh ipv6 mld vrf green groups detail
Interface: Vlan102   <-- Join on Vlan 102

Group: FF06:1::1   <-- Group joined

```
Uptime:         06:38:25
Router mode:    EXCLUDE (Expires: 00:02:14)
Host mode:      INCLUDE
```

Last reporter: FE80::46D3:CAFF:FE28:6CC1 <-- MLD join from Receiver link-local address

Source list is empty   <-- ASM join, no sources listed

Leaf-02#

sh ipv6 neighbors vrf green
IPv6 Address

```
                                Age Link-layer Addr State Interface
```

FE80::46D3:CAFF:FE28:6CC1

```
                    0
```

44d3.ca28.6cc1

```
  REACH Vl102
```

<-- Receiver IP & MAC

Leaf-02#sh ipv6 mld snooping address vlan 102 <-- If MLD snooping is on, it can be checked as well

```
Vlan     Group               Type             Version         Port List
------------------------------------------------------------------------
102
FF06:1::1


  mld
             v2
Gi1/0/10        <-- Receiver joined on Gi1/0/10
```

Verify Leaf-02: MVPN Debugs show Route-type 7 is created when IGMP/MLD membership report arrives and required EVPN Type-2 and Type-5 are already installed.

```
<#root>
### IPv4 ###

Leaf-02#
debug bgp ipv4 mvpn updates

Leaf-02#
debug bgp ipv4 mvpn updates events


*Dec 14 19:41:57.645: BGP[15] MVPN:
add c-route, type 7
, bs len 0 asn=0,
rd=1:1
,
*Dec 14 19:41:57.645:
source=10.1.101.11/4,

*Dec 14 19:41:57.645:
group=226.1.1.1/4,

*Dec 14 19:41:57.645:
nexthop=172.16.254.3
,
<-- Source is via Leaf-01 IP

*Dec 14 19:41:57.645: len left = 0
*Dec 14 19:41:57.645: BGP[14] MVPN umh lookup: vrfid 2, source 10.1.101.11
*Dec 14 19:41:57.645: BGP[4] MVPN umh lookup: vrfid 2, source 10.1.101.11, net 1:1:10.1.101.11/32, 1:1:
```

0x10B:172.16.255.3:2

,
*Dec 14 19:41:57.646:

BGP: MVPN(15) create local route [7][172.16.254.3:101][65001][10.1.101.11/32][226.1.1.1/32]/22

*Dec 14 19:41:57.646:

BGP[15] MVPN: add c-route, type 7, bs len 0 asn=65001, rd=1:1,

IPv6

Leaf-02#

debug bgp ipv6 mvpn updates

Leaf-02#

debug bgp ipv6 mvpn updates events

Mar 23 15:46:11.171: BGP[16] MVPN:

add c-route, type 7

, bs len 0 asn=0, rd=1:1,
Mar 23 15:46:11.171:

source=FC00:1:101::11/16,

Mar 23 15:46:11.171:

group=FF06:1::1/16,

Mar 23 15:46:11.171:

nexthop=::FFFF:172.16.254.3

,

<-- IPv4 next hop of Leaf-01

Mar 23 15:46:11.171: len left = 0
Mar 23 15:46:11.171: BGP[19] MVPN umh lookup: vrfid 2, source FC00:1:101::11
Mar 23 15:46:11.171: BGP[5] MVPN umh lookup: vrfid 2, source FC00:1:101::11, net [1:1]FC00:1:101::11/12

0x10B:172.16.255.3:2

,

Mar 23 15:46:11.172: BGP: MVPN(16) create local route [7][172.16.254.3:101][65001][FC00:1:101::11][FF06:
Mar 23 15:46:11.172: BGP[16] MVPN: add c-route, type 7, bs len 0 asn=65001, rd=1:1,

Verify Leaf-01: The MVPN Type-7 received from Leaf-02

<#root>

IPv4

Leaf-01#

sh bgp ipv4 mvpn all route-type 7 172.16.254.3:101 65001 10.1.101.11 226.1.1.1

...or you can also use:
Leaf-01#

sh bgp ipv4 mvpn detail [7][172.16.254.3:101][65001][10.1.101.11/32][226.1.1.1/32]/22

BGP routing table entry for

[7][172.16.254.3:101]

[65001][10.1.101.11/32][226.1.1.1/32]/22, version 76
Paths: (2 available, best #1, table

MVPNv4-BGP-Table

)

<-- In BGP IPv4 MVPN table

  Not advertised to any peer
  Refresh Epoch 1
  Local

    172.16.255.4

 (metric 3) from 172.16.255.2 (172.16.255.2)

<-- loopback of Leaf-02 Receiver VTEP

      Origin incomplete, metric 0, localpref 100, valid, internal

      Extended Community: RT:172.16.255.3:2   <-- The VRI derived from EVPN Type-2 and ad

      Originator: 172.16.255.4, Cluster list: 172.16.255.2
      rx pathid: 0, tx pathid: 0
      Updated on Dec 15 2020 14:14:38 UTC

IPv6

Leaf-01#

sh bgp ipv6 mvpn all route-type 7 172.16.254.3:101 65001 FC00:1:101::11 FF06:1::1

...or you can also use:
Leaf-01#

sh bgp ipv6 mvpn detail [7][172.16.254.3:101][65001][FC00:1:101::11][FF06:1::1]/46

BGP routing table entry for

[7][172.16.254.3:101]

[65001][FC00:1:101::11][FF06:1::1]/46, version 45
Paths: (2 available, best #1, table

MVPNV6-BGP-Table

)

<-- In BGP IPv6 MVPN table

  Not advertised to any peer
  Refresh Epoch 1
  Local

    172.16.255.4

 (metric 3) from 172.16.255.1 (172.16.255.1)

<-- loopback of Leaf-02 Receiver VTEP

      Origin incomplete, metric 0, localpref 100, valid, internal, best

      Extended Community: RT:172.16.255.3:2   <-- The VRI derived from EVPN Type-2 and added to the MVP

      Originator: 172.16.255.4, Cluster list: 172.16.255.1
      rx pathid: 0, tx pathid: 0x0
      Updated on Mar 23 2021 15:46:11 UTC

Verify Leaf-01: MVPN Debugs show Route-type 7 received with the MVPN VRI Route-Target

<#root>

*Dec 17 16:16:31.923: BGP(15): 172.16.255.2

rcvd UPDATE w/ attr: nexthop 172.16.255.4

, origin ?, localpref 100, metric 0, originator 172.16.255.4, clusterlist 172.16.255.2,

extended community RT:172.16.255.3:2 <-- VRI RT

*Dec 17 16:16:31.923: BGP(15): 172.16.255.2

rcvd [7]

[172.16.254.3:101][65001][10.1.101.11/32][226.1.1.1/32]/22

<-- Received MVPN Type-7

<...only update from Spine-02 172.16.255.2 ...>

*Dec 17 16:16:31.923: BGP(15): skip vrf default table RIB route [7][172.16.254.3:101][65001][10.1.101.1
*Dec 17 16:16:31.924: BGP(15): add RIB route (0:0)[7][1:1][65001][10.1.101.11/32][226.1.1.1/32]/22

(Skipping IPv6, see the debugs demonstrated in previous steps)

Verify Leaf-02: Complete BGP table contains the Leaf-01 EVPN Type-2 and MVPN Type-5, and the Type-7 generated by Receiver Leaf-02

<#root>

IPv4

Leaf-02#

sh bgp l2vpn evpn all | b 10.1.101.11

* i

[2]

[172.16.254.3:101][0][48][F4CFE24334C5][32][10.1.101.11]/24

<-- Remote VTEP route-type 2

172.16.254.3 0 100 0 ?

***>i 172.16.254.3 0 100 0 ? <-- IP of Leaf01 Lo1**

Leaf-02#

sh bgp ipv4 mvpn all

Network Next Hop Metric LocPrf Weight Path

Route Distinguisher: 1:1

(default for vrf green)

<-- default RD for vrf green

*>i

[5][1:1][10.1.101.11][226.1.1.1]

/18

<-- Type-5, source & group

172.16.255.3

0 100 0 ?

<-- Next hop Leaf-01 IP

* i 172.16.255.3 0 100 0 ?

Route Distinguisher: 172.16.254.3:101 <-- MVPN RD sent from Source Leaf-01

*>

[7]

[172.16.254.3:101][65001][10.1.101.11/32][226.1.1.1/32]/22

<-- Type-7 BGP Join Entry

0.0.0.0

32768

?

<-- Locally created (0.0.0.0) by Leaf-02

IPv6

Leaf-02#

sh bgp l2vpn evpn all | b FC00:1:101::11

* i

[2]

[172.16.254.3:101][0][48][F4CFE24334C1][128][FC00:1:101::11]/36

<-- Remote VTEP route-type 2

172.16.254.3 0 100 0 ?

***>i 172.16.254.3 0 100 0 ? <-- IP of Leaf-01 Lo1**

Leaf-02#

sh bgp ipv6 mvpn all

Network Next Hop Metric LocPrf Weight Path

Route Distinguisher: 1:1

(default for vrf green)

<-- default RD for vrf green

*>i

[5][1:1][FC00:1:101::11][FF06:1::1]

/42

<-- Type-5, source & group

172.16.255.3

0 100 0 ?

<-- IPv4 Next hop Leaf-01 IP

 * i   172.16.255.3   0 100   0 ?

Route Distinguisher: 172.16.254.3:101   <-- MVPN RD sent from Source Leaf-0

 *>

[7]

[172.16.254.3:101][65001][FC00:1:101::11][FF06:1::1]/46

<-- Type-7 BGP Join Entry

::   **32768**

 ?

<-- Locally created (::) by Leaf-02

Verify the TRM group Leaf-01 (FHR)

Check that the MDT and TRM groups are correctly formed on Source side.

- Incoming interface of TRM group is the SVI associated to the Client VRF
- Outgoing interface of TRM group is the L3VNI SVI

Verify Leaf-01: the TRM group MRIB/MFIB

<#root>

IPv4

Leaf-01#

sh ip mroute vrf green 226.1.1.1 10.1.101.11

(10.1.101.11, 226.1.1.1), 02:57:56/00:03:14,

flags: FTGqx <-- Flags: BGP S-A Route

  Incoming interface:

Vlan101

, RPF

nbr 0.0.0.0   <-- Local to Vlan101 Direct connected source

  Outgoing interface list:

Vlan901

, Forward/Sparse, 02:57:56/stopped

<-- OIF is VxLAN L3VNI

Leaf-01#

sh ip mfib vrf green 226.1.1.1 10.1.101.11

VRF green <-- Tenant VRF

(10.1.101.11,226.1.1.1) Flags: HW
  SW Forwarding: 1/0/100/0, Other: 0/0/0

HW Forwarding: 5166/0/118/0, Other: 0/0/0 <-- Hardware counters indicate the entry is operating in hardw

Vlan101 Flags: A <-- Accept flag set on Connected Source SVI

  Vlan102 Flags: F NS
    Pkts: 0/0/1 Rate: 0 pps

Vlan901, VXLAN v4 Encap (50901, 239.1.1.1) Flags: F <-- Forward via Vlan 901. Use MDT group 239.1.1.1, v

    Pkts: 0/0/0 Rate: 0 pps

IPv6

Leaf-01#

sh ipv6 mroute vrf green

(FC00:1:101::11, FF06:1::1), 01:01:00/00:01:08,

flags: SFTGq <-- Flags: q - BGP S-A Route, G - BGP Signal Received

  Incoming interface:

Vlan101

RPF nbr: FE80::F6CF:E2FF:FE43:34C1 <-- link local address of Source

  Immediate Outgoing interface list:

Vlan901

, Forward, 01:01:00/never

<-- OIF is VxLAN L3VNI

```
Leaf-01#
sh ipv6 mfib vrf green FF06:1::1
VRF green   <-- Tenant VRF

(FC00:1:101::11,FF06:1::1) Flags: HW

   SW Forwarding: 0/0/0/0, Other: 1/0/1

HW Forwarding: 1968/0/118/0, Other: 0/0/0 <-- Hardware counters indicate the entry is operating in hardw

 Vlan101 Flags: A NS                        <-- Accept flag set on Connected Source SVI

 Vlan901, VXLAN v4 Encap (50901, 239.1.1.1) Flags: F <-- Forward via Vlan 901. Use MDT group 239.1.1.1,
     Pkts: 0/0/0 Rate: 0 pps
```

Verify Leaf-01: the TRM group in FED

```
<#root>
### IPv4 ###

Leaf-01#
sh platform software fed switch active ip mfib vrf green 226.1.1.1/32 10.1.101.11

Multicast (S,G) Information
VRF            : 2     <-- VRF ID 2 = vrf green (from "show vrf detail")

Source Address : 10.1.101.11
HTM Handler    : 0x7f175cc08578
SI Handler     : 0x7f175cc06ea8
DI Handler     : 0x7f175cc067c8
REP RI handler : 0x7f175cc06b38
Flags          : {Svl}
Packet count   : 39140       <-- packets that used this adjacency (similar to mfib command, but shown at

State          : 4
RPF
            :
     Vlan101   A             <-- Accept on Vlan 101 in Tenant vrf green
```

```
OIF             :
     Vlan102     F NS
       Vlan101   A

       Vlan901   F {Remote}


<-- Forward via L3VNI interface


             (Adj: 0x6a )     <-- Adjacency for this entry


### IPv6 ###

Leaf-01#

sh plat soft fed switch active ipv6 mfib vrf green FF06:1::1/128 FC00:1:101::11


Multicast (S,G) Information

VRF             : 2           <-- VRF ID 2 = vrf green (from "show vrf detail")


Source Address : fc00:1:101::11
HTM Handler    : 0x7fba88d911b8
SI Handler     : 0x7fba88fc4348
DI Handler     : 0x7fba88fc8dc8
REP RI handler : 0x7fba88fc8fd8
Flags          : {Svl}

Packet count   : 2113


<-- packets that used this adjacency (similar to mfib command, but shown at the FED layer)


State          : 4

RPF             :


     Vlan101   A {Remote}    <-- Accept on Vlan 101 in Tenant vrf green (says remote, but this is a loca


OIF             :
     Vlan101     A {Remote}

 Vlan901     F {Remote}


<-- Forward via L3VNI interface



  (Adj: 0x7c )      <-- Adjacency for this entry
```

Verify Leaf-01: the Adjacency is correct

<#root>

IPv4

Leaf-01#

sh platform software fed switch active ip adj

IPV4 Adj entries

dest if_name dst_mac si_hdl ri_hdl

adj_id

Last-modified
---- ------- ------- ------ ------

239.1.1.1

nve1.VNI50901

4500.0000.0000 0x7f175ccd8c38 0x7f175ccd8de8 0x60

0x6a

2020/12/16 17:39:55.747

***** Adjacency 0x6a details *****

Destination =

the MDT tunnel multicast group 239.1.1.1

Interface =

nve1.VNI50901 (the L3VNI 50901)

IPv6

Leaf-01#

sh platform software fed switch active ipv6 adj
IPV6 Adj entries

dest if_name dst_mac si_hdl ri_hdl

adj_id

Last-modified
---- ------- ------- ------ ------

239.1.1.1

```
nve1.VNI50901

  4500.0000.0000  0x7fba88cf9fc8  0x7fba88cfa248 0x60

  0x7c

    2021/03/22 19:54:09.831


*** Adjacency 0x7c details ***
Destination =

the MDT tunnel multicast group 239.1.1.1

Interface   =

nve1.VNI50901 (the L3VNI 50901)
```

Verify the TRM group Leaf-02 (LHR)

Check that the MDT and TRM groups are correctly formed on Receiver side.

- Incoming interface of TRM group is the SVI associated to the L3VNI
- Outgoing interface of TRM group is the Client SVI where the IGMP join was processed.

Verify Leaf-02: the TRM (Tenant multicast route) route in MRIB/MFIB

```
<#root>

Leaf-02#

sh ip mroute vrf green 226.1.1.1 10.1.101.11      <-- The TRM Client group


(10.1.101.11, 226.1.1.1), 00:26:03/00:02:37, flags: TgQ

  Incoming interface: Vlan901, RPF nbr 172.16.254.3       <-- Via L3VNI, RPF to Leaf-01


  Outgoing interface list:

    Vlan102,

 Forward/Sparse, 00:26:03/00:03:10

<-- Client Receiver Vlan


Leaf-02#

sh ip mfib vrf green 226.1.1.1 10.1.101.11


VRF green                                          <--- The Tenant VRF


  (10.1.101.11,226.1.1.1) Flags: HW
    SW Forwarding: 1/0/100/0, Other: 0/0/0
```

HW Forwarding: 39013/0/126/0, Other: 0/0/0 <-- Hardware counters indicate the entry is operating in

Vlan901, VXLAN Decap Flags: A <-- L3VNI Accept and decapsulate from VxLAN

Vlan102 Flags: F NS <-- Forward to the Tenant Vlan

Pkts: 0/0/1 Rate: 0 pps

Verify Leaf-02: the TRM group in FED

<#root>

IPv4

Leaf-02#

sh platform software fed switch active ip mfib vrf green 226.1.1.1/32 10.1.101.11 detail <-- Use detail

MROUTE ENTRY vrf 2 (10.1.101.11, 226.1.1.1/32)
HW Handle: 140397391947768 Flags: {Svl}

RPF interface: Vlan901

(60)):

SVI <-- RPF interface = L3VNI SVI Vlan901

HW Handle:140397391947768 Flags:A {Remote}
Number of OIF: 2
Flags: 0x4

Pkts : 39387 <-- packets that used this adjacency (similar to mfib command, but shown at the FED la

OIF Details:

Vlan102 F NS <-- Client Vlan

Vlan901 A {Remote} <-- Accept interface is RPF to source via Remote EVPN next hop

(Adj: 0xf80003c1) <-- Adj for vlan 901(show plat soft fed sw active ipv4 adj)

Htm: 0x7fb0d0edfb48 Si: 0x7fb0d0ee9158 Di: 0x7fb0d0eca8f8 Rep_ri: 0x7fb0d0ef2b98

DI details <-- Dest index (egress interface) details

Handle:0x7fb0d0eca8f8 Res-Type:ASIC_RSC_DI Res-Switch-Num:255 Asic-Num:255 Feature-ID:AL_FID_L3_MULTICA

priv_ri/priv_si Handle:(nil) Hardware Indices/Handles: index0:0x538b mtu_index/l3u_ri_index0:0x0 index1

Brief Resource Information

(ASIC_INSTANCE# 1)

<-- Gi1/0/10 is mapped to instance 1

Destination index = 0x538b

pmap = 0x00000000 0x00000200

pmap_intf : [GigabitEthernet1/0/10] <-- Maps to Gi1/0/10, the port toward the client

==========

IPv6

Leaf-02#

sh platform software fed switch active ipv6 mfib vrf green FF06:1::1/128 FC00:1:101::11 detail

MROUTE ENTRY

vrf 2

(fc00:1:101::11, ff06:1::1/128)
HW Handle: 139852137577736 Flags: {Svl}

RPF interface: Vlan901

(62)): SVI

<-- RPF to Source L3VNI SVI 901

HW Handle:139852137577736

Flags:A {Remote}

Number of OIF: 2

Flags: 0x4 Pkts : 7445 <-- Packets use this Entry

OIF Details:

Vlan102 F NS <-- F - Forward. The OIF Vlan SVI 901

Vlan901 A {Remote}

(Adj: 0xf80003e2) <-- Adj for vlan 901 (show plat soft fed sw active ipv6 adj)

Htm: 0x7f31dcfee238 Si: 0x7f31dcfba5d8 Di: 0x7f31dcfc2358 Rep_ri: 0x7f31dcfcb1a8

DI details

```
----------
Handle:0x7f31dcfc2358 Res-Type:ASIC_RSC_DI Res-Switch-Num:255 Asic-Num:255 Feature-ID:AL_FID_L3_MULTICA
priv_ri/priv_si Handle:(nil) Hardware Indices/Handles: index0:0x5381 mtu_index/l3u_ri_index0:0x0 index1

Brief Resource Information

(ASIC_INSTANCE# 1)   <-- Gig1/0/10 is mapped to Instance 1


----------------------------------------

Destination index = 0x5381


pmap = 0x00000000 0x00000200

pmap_intf : [GigabitEthernet1/0/10]            <-- Maps to Gig1/0/10, the port toward the client


============================================================


Leaf-02#

sh platform software fed switch active ifm mappings


Interface                IF_ID

Inst


Asic

 Core Port SubPort Mac Cntx LPN GPN Type Active

GigabitEthernet1/0/10

     0x12

 1


0

   1    9      0      5   15   10  10  NIF  Y

 <-- Instance 1 of ASIC 0
```

Verify Leaf-02: Packet capture taken shows the outer MDT tunnel group with inner client traffic

```
<#root>

Leaf-02#

sh mon ca 1 parameter


monitor capture 1 interface GigabitEthernet1/0/2 IN
monitor capture 1 match any
```

```
monitor capture 1 buffer size 10
monitor capture 1 limit pps 1000


### IPv4 ###


Leaf-02#
sh mon capture 1 buffer detailed


Ethernet II, Src: 7c:21:0d:bd:2c:d6 (7c:21:0d:bd:2c:d6),
Dst: 01:00:5e:01:01:01
 (01:00:5e:01:01:01)
<-- MAC is matching 239.1.1.1



Type: IPv4 (0x0800)  <-- IPv4 outer packet


Internet Protocol Version 4,
Src: 172.16.254.3, Dst: 239.1.1.1 <- Leaf-01 Source IP and MDT outer tunnel Group


    0100 .... = Version: 4
    .... 0101 = Header Length: 20 bytes (5)
    Time to live: 253
User Datagram Protocol
, Src Port: 65287,
Dst Port: 4789 <-- VxLAN UDP port 4789



Virtual eXtensible Local Area Network


Flags: 0x0800,
VXLAN Network ID (VNI)


Group Policy ID: 0
VXLAN Network Identifier (VNI): 50901 <-- L3VNI value
   Type: IPv4
 (0x0800)
<-- IPv4


inner packet


Internet Protocol Version 4
```

,

Src: 10.1.101.11, Dst: 226.1.1.1 <-- Encapsulated IPv4 TRM group

    0100 .... = Version: 4
    Time to live: 254
    Protocol: ICMP (1)

(multiple lines removed from this example capture)

IPv6

Leaf-02#

sh mon capture 1 buffer detailed

Ethernet II,

Src: 7c:21:0d:bd:2c:d6

 (7c:21:0d:bd:2c:d6),

Dst: 01:00:5e:01:01:01

 (01:00:5e:01:01:01)

<-- DMAC is matching 239.1.1.1

    Type: IPv4 (0x0800)   <-- IPv4 outer packet

Internet Protocol Version 4, Src: 172.16.254.3, Dst: 239.1.1.1

    0100 .... = Version: 4
    .... 0101 = Header Length: 20 bytes (5)
    Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
        0000 00.. = Differentiated Services Codepoint: Default (0)
        .... ..00 = Explicit Congestion Notification: Not ECN-Capable Transport (0)
    Total Length: 150
    Identification: 0x4e4b (20043)
    Flags: 0x4000, Don't fragment
        0... .... = Reserved bit: Not set
        .1.. .... = Don't fragment: Set <-- DF flag=1. MTU can be an issue if too low in path
        ..0. .... = More fragments: Not set
    ...0 0000 0000 0000 = Fragment offset: 0
    Time to live: 253
    Protocol: UDP (17)

    Header checksum: 0x94f4 [validation disabled]
    [Header checksum status: Unverified]
    Source: 172.16.254.3
    Destination: 239.1.1.1
User Datagram Protocol,

Src Port: 65418, Dst Port: 4789 <-- VxLAN UDP port 4789

```
    Source Port: 65418

 Destination Port: 4789


<...snip...>


Virtual eXtensible Local Area Network


    Flags: 0x0800,
VXLAN Network ID (VNI)


        0... .... .... .... = GBP Extension: Not defined
        .... .... .0.. .... = Don't Learn: False


  .... 1... .... .... = VXLAN Network ID (VNI): True


        .... .... .... 0... = Policy Applied: False
        .000 .000 0.00 .000 = Reserved(R): 0x0000
    Group Policy ID: 0
    VXLAN Network Identifier (VNI): 50901 <-- L3VNID 50901


    Reserved: 0
Ethernet II, Src: 10:b3:d5:6a:00:00 (10:b3:d5:6a:00:00), Dst:
33:33:00:00:00:01
 (33:33:00:00:00:01)
<-- DMAC matches ff06:1::1


    Type: IPv6 (0x86dd)                  <-- IPv6 inner packet


Internet Protocol Version 6
,
Src: fc00:1:101::11, Dst: ff06:1::1 <-- Encapsulated IPv6 TRM group


    0110 .... = Version: 6
<...snip...>


    Source: fc00:1:101::11


    Destination: ff06:1::1


Internet Control Message Protocol v6
```

```
    Type: Echo (ping) request (128)
<...snip...>
```

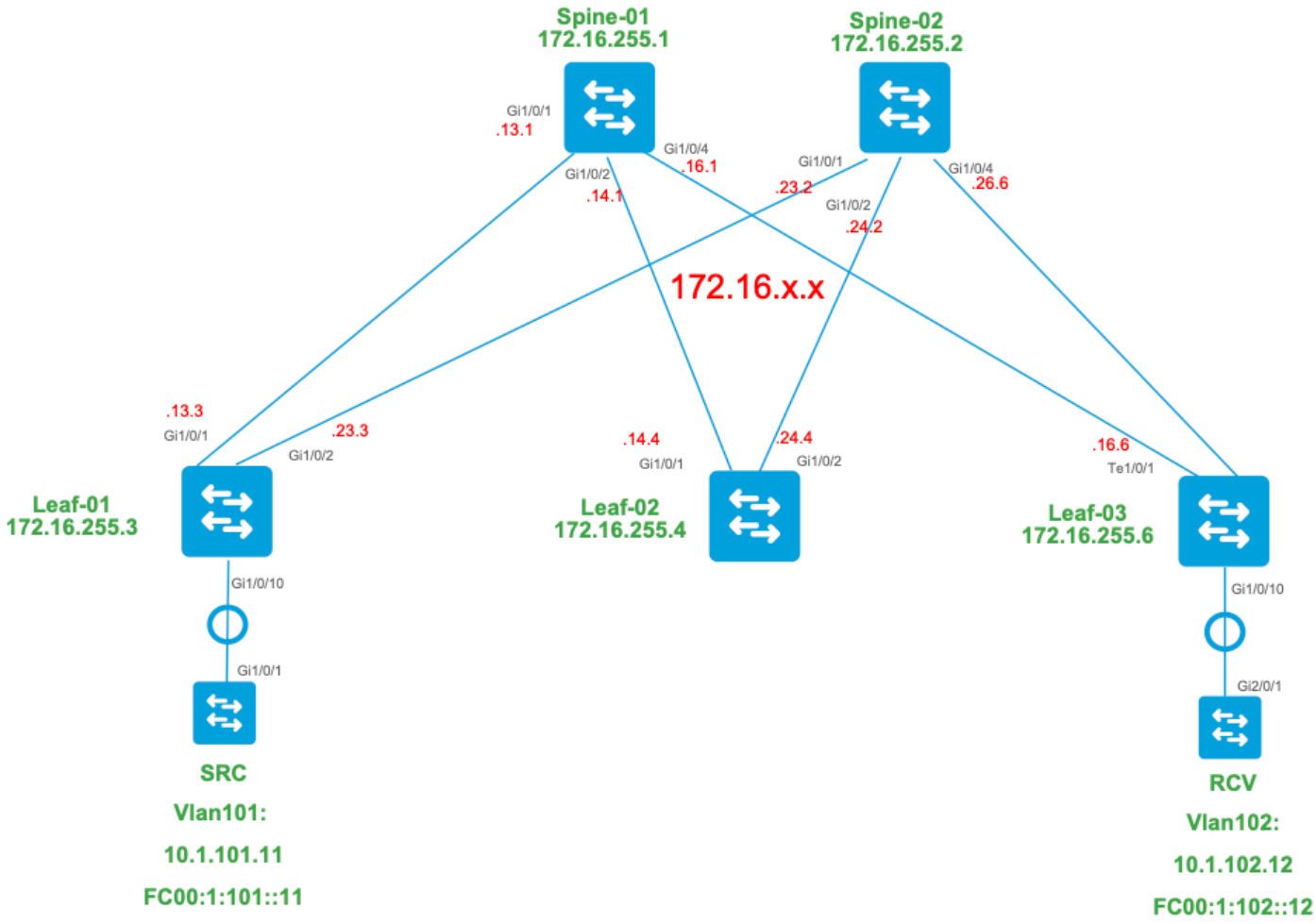
Scenario 2: PIM SSM in the Fabric

In this mode there is no RP in the Overlay, and no MVPN Type-5 or Type-7 are used (the Underlay continues to operate as PIM ASM). **In SSM, receiver sends and IGMPv3 S,G join** towards the LHR VTEP. This VTEP performs RPF lookup for the Source in the RIB. If L3VNI SVI is found as the RPF interface, the LHR VTEP sends MVPN RT-7 to the FHR VTEP who receives and installs this route. FHR VTEP then informs PIM to add L3VNI SVI as the Outgoing interface for the S,G mroute.

This section shows the differences from Scenario 1. The steps and methods that are the same are noted only in Scenario 1.

- See verification and debug steps for BGP and PIMfrom Scenario 1, as the BGP and PIM operations are the same

Network Diagram



For this mode, consider these BGP route-types and their origins

Created by: Source VTEP

- EVPN Route-type 2. Used to obtain Unicast and VRI info for the Source, and added to the C-

Multicast route (MVPN type-7) when the VTEP joins STP tree.

Created by: Receiver VTEP

- MVPN Route-type 7. Information from the IGMP or MLD layer & from EVPN Type-2 are used to create this BGP type join. The Type-7 drives the creation of the MRIB OIF on the Source side.

EVPN Type-2 requirements:

1. FHR (source VTEP) verifies ARP (or ND) and CEF adjacency (confirms the source is Directly Connected)..
2. FHR originates the EVPN Type-2 BGP update

MVPN Type-7 requirements:

1. EVPN Type-2 entry is present (required to construct the C-Multicast route type-7 with correct VRI and sent from Source VTEP)
2. Receiver VTEP: IGMPv3 Source specific membership report has been received and processed by the LHR VTEP
3. LHR VTEP RPF interface is the Fabric L3VNI interface

For this mode, there is added configuration required on the LHR VTEP to enbale SSM range, and process IGMPv3 membership reports

Configure Leaf-03: set the IGMP querier to Version 3 under the Tenant SVI

<#root>

```
interface Vlan102


vrf forwarding green
ip address 10.1.102.1 255.255.255.0
ip pim sparse-mode

ip igmp version 3  <-- Sets the version to V3


end
```

Verify Leaf-03: the IGMP querier is set to **version 3**

<#root>

```
Leaf-03#

sh ip igmp snooping querier vlan 102


IP address : 10.1.102.1   <-- IP is that of the Vlan102 SVI


IGMP version : v3         <-- Querier is now version 3
```

```
Port : Router            <-- Mrouter port is "Router" meaning querier is local to this VTEP

Max response time : 10s
Query interval : 60s
Robustness variable : 2
```

Enable Leaf-03: the SSM range required for the Tenant VRF

```
<#root>

Leaf-03(config)#

ip pim vrf green ssm


?


default

Use 232/8 group range for SSM          <-- Set to the normally defined SSM range


range

ACL for group range

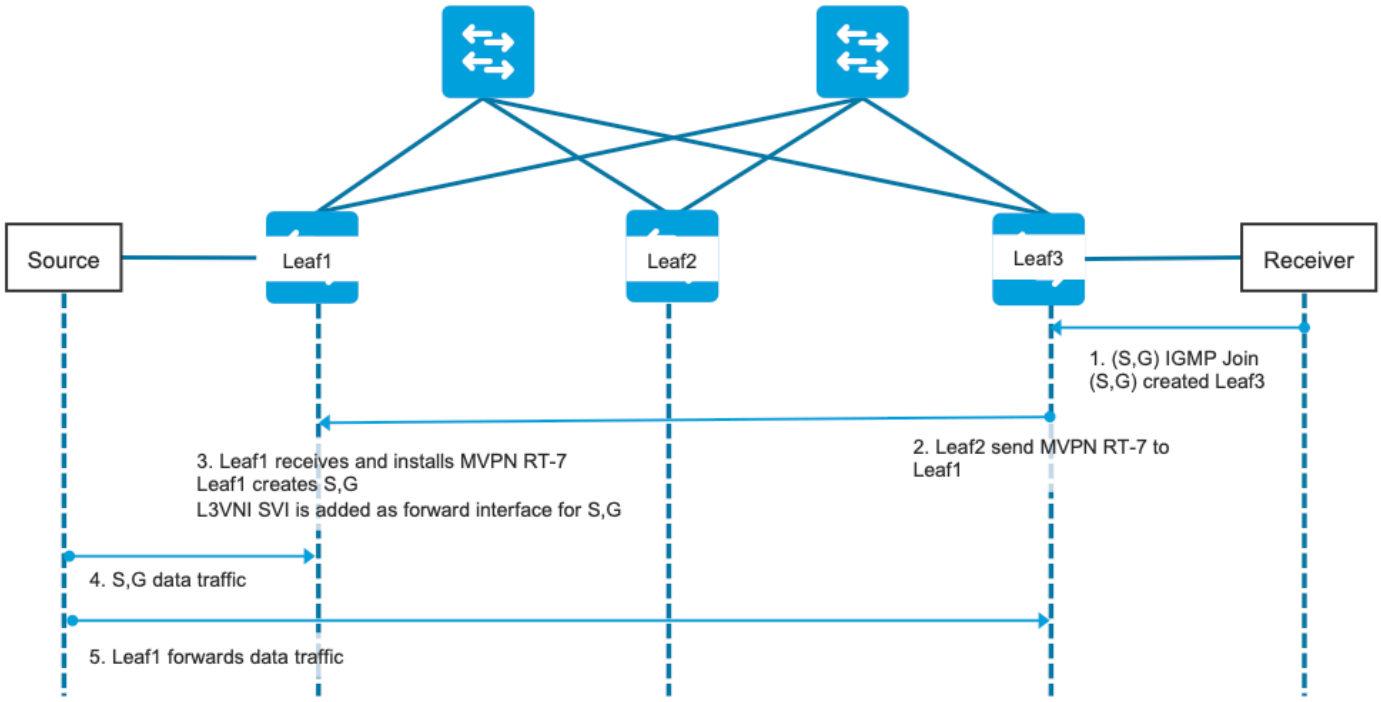
 to be used for SSM

<-- use an ACL to define a non-default SSM range
```



Tip: SSM groups do not create a *,G mroute. If you do see *,G for the group, verify your configuration is correct for SSM.

Verify the Sequence of Events Required for this Scenario



Step 0 EVPN (Leaf-03): Verify there is an EVPN prefix which BGP can find the VRI to use in the MVPN type-7.

<#root>

Leaf-03#

sh bgp l2vpn evpn all

```
BGP table version is 16, local router ID is 172.16.255.6
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
              t secondary path, L long-lived-stale,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

     Network          Next Hop            Metric LocPrf Weight Path
Route Distinguisher: 1:1 (default for vrf green)
* i
```

[2]

[172.16.254.3:101][0][48][F4CFE24334C1][32]

[10.1.101.11]

/24

172.16.254.3 0 100 0 ?

***>i** **172.16.254.3 0 100 0 ? <-- From Leaf-01**

Leaf-03#

sh bgp l2vpn evpn all route-type 2 0 F4CFE24334C1 10.1.101.11 <-- Detailed view of the EVPN type-2 e

```
BGP routing table entry for

[2]

[172.16.254.3:101][0][48][F4CFE24334C1][32][10.1.101.11]/24, version 283
Paths: (2 available, best #2,

table EVPN-BGP-Table

)
  Not advertised to any peer
  Refresh Epoch 1
  Local
    172.16.254.3 (metric 3) (via default) from 172.16.255.1 (172.16.255.1)
      Origin incomplete, metric 0, localpref 100, valid, internal, best
      EVPN ESI: 00000000000000000000, Gateway Address: 0.0.0.0, VNI Label 50901, MPLS VPN Label 0
      Extended Community: RT:1:1 MVPN AS:65001:0.0.0.0


 MVPN VRF:172.16.255.3:4

 ENCAP:8 Router MAC:10B3.D56A.8FC8

<-- BGP finds the VRI in this entry


      Originator: 172.16.255.3, Cluster list: 172.16.255.1
      rx pathid: 0, tx pathid: 0x0
      Updated on May 6 2021 16:17:06 UTC
```

Step 1 (Leaf-03): IGMPv3 Membership report received and contains a Source

```
<#root>

Leaf-03#

show ip igmp snooping groups vlan 102 226.1.1.1


Vlan

Group

                    Type

Version

      Port List
------------------------------------------------------------------------
102

226.1.1.1

                igmp

 v3

           Gi1/0/10


Leaf-03#

show ip igmp snooping groups vlan 102 226.1.1.1 sources <-- Specify "sources" to see Source information
```

```
Vlan      Group                    Type        Version      Port List
------------------------------------------------------------------------
```

Source information for group 226.1.1.1

```
:
Timers: Expired sources are deleted on next IGMP General Query
```

SourceIP

```
       Expires   Uptime
```

Inc Hosts

```
 Exc Hosts
-------------------------------------------------------
```

10.1.101.11

```
    00:01:20  00:02:58
```

1

```
         0
```

<-- Source specified in IGMP includes one source

Step 2 (Leaf-03): BGP is informed of this join, creates, and sends the Type-7 MVPN join.

<#root>

debug mvpn

debug ip igmp vrf green 226.1.1.1

```
May 6 17:11:08.500:
```

IGMP(6): Received v3 Report for 1 group on Vlan102 from 10.1.102.12

```
May 6 17:11:08.500:
```

IGMP(6): Received Group record for group 226.1.1.1, mode 5 from 10.1.102.12 for 1 sources <-- IGMPv3 typ

```
May 6 17:11:08.500: IGMP(6): WAVL Insert group: 226.1.1.1 interface: Vlan102 Successful
May 6 17:11:08.500: IGMP(6): Create source 10.1.101.11
May 6 17:11:08.500: IGMP(6): Updating expiration time on (10.1.101.11,226.1.1.1) to 180 secs
May 6 17:11:08.500: IGMP(6): Setting source flags 4 on (10.1.101.11,226.1.1.1)
May 6 17:11:08.500: IGMP(6): MRT Add/Update Vlan102 for (10.1.101.11,226.1.1.1) by 0

May 6 17:11:08.501:
```

MVPN: Received local route update for (10.1.101.11, 226.1.1.1) with RD: 1:1, Route Type: 7, flags: 0x00

May 6 17:11:08.501: MVPN: Route Type 7 added [(10.1.101.11, 226.1.1.1)] rd:1:1 send:1
May 6 17:11:08.501:

MVPN: Sending BGP prefix=[7:0 1:1 : (10.1.101.11,226.1.1.1)] len=23, nh 172.16.254.3, Originate route

May 6 17:11:08.501:

MVPN: Originate C-route, BGP remote RD 1:1

Leaf-03#

sh bgp ipv4 mvpn all

BGP table version is 10, local router ID is 172.16.255.6
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
              t secondary path, L long-lived-stale,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

     Network   Next Hop   Metric LocPrf Weight Path
Route Distinguisher: 1:1 (default for vrf green)
*>

[7][1:1][65001][10.1.101.11/32][226.1.1.1/32]/22   <-- Locally created Type-7

                      0.0.0.0   32768 ?

Leaf-03#

sh ip mroute vrf green 226.1.1.1   <-- for SSM you only see S,G and no *,G

IP Multicast Routing Table

<...snip...>

(10.1.101.11, 226.1.1.1), 00:29:12/00:02:46, flags: sTIg <-- s = SSM, I = Source Specific Join received,

Incoming interface: Vlan901

, RPF nbr 172.16.254.3

<-- RPF interface is the L3VNI

  Outgoing interface list:
    Vlan102, Forward/Sparse, 00:29:12/00:02:46

Step 3 (Leaf-01): Source Leaf receives and installs MVPN Type-7 join route, and informs PIM to install L3VNI OIF

<#root>

debug mvpn

debug ip pim vrf green 226.1.1.1

May 6 18:16:07.260: MVPN: Received BGP prefix=[7:65001 1:1 : (10.1.101.11,226.1.1.1)] len=23, nexthop: 1

May 6 18:16:07.260: MVPN: Received BGP route update for (10.1.101.11, 226.1.1.1) with RD: 1:1, Route Ty
May 6 18:16:07.260: MVPN:

Route Type 7 added [(10.1.101.11, 226.1.1.1), nh 172.16.255.6] rd:1:1 send:0, to us <-- add type-7 rou

May 6 18:16:07.260: PIM(4)[green]: Join-list: (10.1.101.11/32, 226.1.1.1), S-bit set, BGP C-Route
May 6 18:16:07.263:

PIM(4)[green]: Add Vlan901/0.0.0.0 to (10.1.101.11, 226.1.1.1), Forward state, by BGP SG Join <-- PIM a

May 6 18:16:07.264: PIM(4)[green]: Insert (10.1.101.11,226.1.1.1) join in nbr 10.1.101.11's queue
May 6 18:16:07.264:

MVPN(green[AF_IPv4]): Add (10.1.101.11, 226.1.1.1) intf Vlan901 olist Join state for BGP C-Rt type 7 Acc

Leaf-01#

sh bgp ipv4 mvpn all

<...snip...>
    Network   Next Hop   Metric LocPrf Weight Path
Route Distinguisher: 1:1 (default for vrf green)

***>i [7][1:1][65001][10.1.101.11/32][226.1.1.1/32]/22**

172.16.255.6

          0   100   0 ?

<-- Recieved from Reciever Leaf-03

* i   172.16.255.6   0   100   0 ?

Leaf-01#

sh ip mroute vrf green 226.1.1.1

<...snip...>

(10.1.101.11, 226.1.1.1), 00:42:41/stopped, flags: sTGx   <-- s = SSM Group, G = Received BGP

  Incoming interface: Vlan101, RPF nbr 10.1.101.11

```
Outgoing interface list:

    Vlan901, Forward/Sparse, 00:42:41/stopped                    <-- L3VNI installed as OIF interface
```

Step 4 & 5 (Leaf-01 & Leaf-03): Multicast arrives to the FHR leaf and is sent over fabric to LHR leaf. Summary of validation commands given here. You can check the detailed validation of these commands in Scenario 1.

<#root>

```
show ip mroute vrf green 226.1.1.1 count                                    <-- software mr

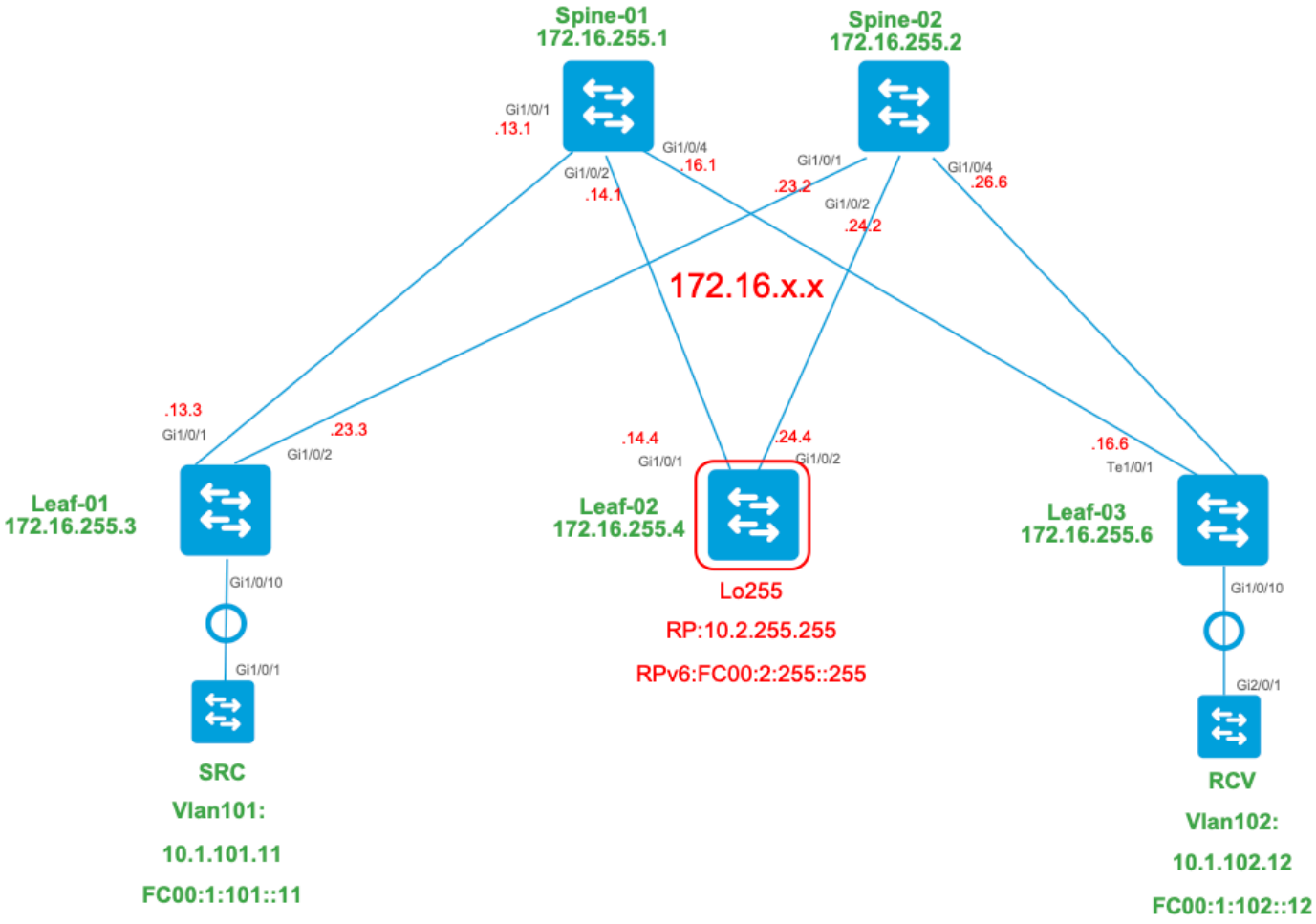
show ip mfib vrf green 226.1.1.1 <verbose>                                  <-- hardware mr

sh platform software fed switch active ip mfib vrf green 226.1.1.1/32 10.1.101.11 detail <-- ASIC entry
```

Scenario 3: Single RP Inside the Fabric (Regular Sparse-Mode)

This mode is interchangeably called as non-Anycast RP or external RP mode. In this mode, there is only one RP in the overlay. Thus, (*,G) tree in the overlay could span across multiple sites. BGP uses an MVPN RT-6 to advertise (*,G) membership across the fabric. If RP and FHR are at different sites, PIM registers are sent across the fabric. This is the default operational mode for PIM SM in the overlay.

Network Diagram



For this mode, consider these BGP route-types and their origins

Created by: Source VTEP

- EVPN Route-type 2. Used to obtain Unicast and VRI info for the Source, and added to the C-Multicast route (MVPN type-7) when the VTEP joins STP tree.
- MVPN Route-type 5. Source A-D route sent to VTEPs for S,G

Created by: RP VTEP

- EVPN Route-type 5. Used to obtain Unicast and VRI info for RP loopback. Loopback does not create Route-type 2, so type 5 is used.
- MVPN Route-type 7. This is the IGMP join + RT VRI details taken from the EVPN Type-2 and sent to the Source VTEP, and it drives the creation of the MRIB OIF.

Created by: Receiver VTEP

- MVPN Route-type 6**.** Route-type created by receiver VTEP to join the Shared Tree *,G (RPT tree) toward the RP.
- MVPN Route-type 7. Information from the IGMP or MLD layer & from EVPN Type-2 are used to create this BGP type join. The Type-7 drives the creation of the MRIB OIF on the Source side.

EVPN Type-2 requirements:

1. FHR (source VTEP) verifies ARP (or ND) and CEF adjacency (confirms the source is Directly Connected)..

2. FHR originates the EVPN Type-2 BGP update

EVPN Type-5 requirements:

1. RP loopback is configured and advertised into BGP

MVPN Type-5 requirements:

In this mode, Leaf at the source site advertises source active A-D messages for an (S,G) only if these two conditions are satisfied.

1. It receives traffic on the RPF interface towards the Source. (source is sends mcast to the FHR)
2. L3VNI SVI interface is added as a forwarding interface for (S,G) entry, as a result of an S,G join from the RP as part of the PIM registration process. (The L3VNI SVI is installed in the OIF list)

MVPN Type-6 requirements:

1. RP advertised its EVPN Type-5 route that contains its VRI and Unicast reachability details.
2. IGMP join received on LHR which triggers a BGP update toward RP

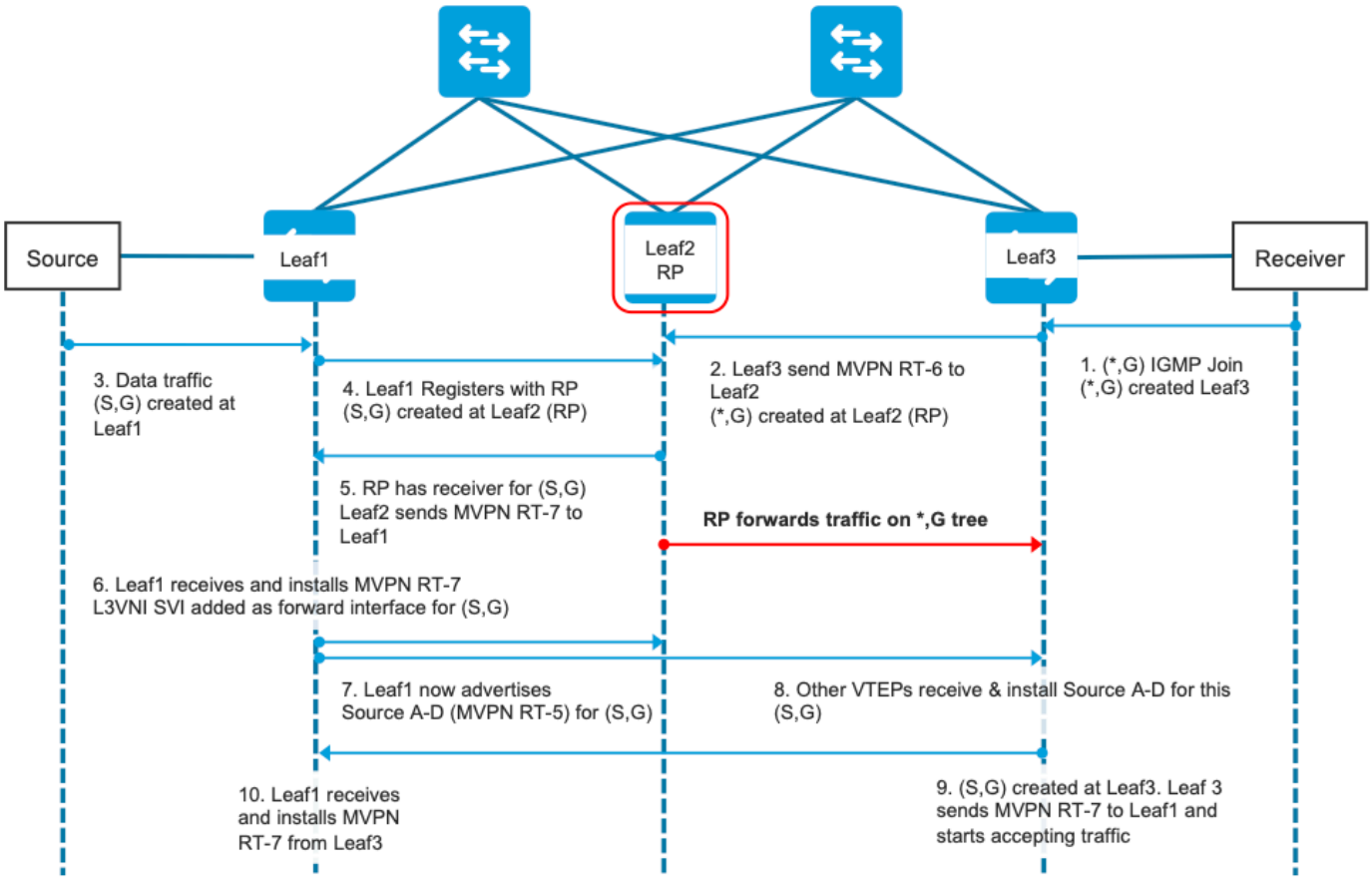
MVPN Type-7 requirements:

1. EVPN Type-2 entry is present (required to construct the C-Multicast route type-7 with correct VRI and sent from Source VTEP)
2. MVPN Type-5 entry is present (required to resolve what Source/Group pair is available for STP join)
3. Receiver VTEP: IGMP membership report has been received and processed by the LHR VTEP
4. RP VTEP: RP has received multicast Register packets, has EVPN routes, and has a receiver for S,G (learnt via the type-6)
5. LHR VTEP RPF interface is the Fabric L3VNI interface

 **Tip**: At the egress LHR VTEP PIM checks the path toward the Source. PIM must find a route in the RIB that is the L3VNI as the RPF interface. If the L3VNI is not configured properly, is down, and so on. the VTEP does not create the type-7 BGP join.

Verify the Sequence of Events Required for this Scenario

Validate the steps needed for the Receiver VTEP to initially join the Shared Tree, then cut over to the Shortest Path Tree. This involves checks of the BGP tables, IGMP, and MRIB creation states.



Step EVPN (Leaf-03): EVPN Type-5 from RP is learned on LHR. This is required for the receiver VTEP to create an MVPN Type-6 route

<#root>

```
Leaf-03#
```

```
sh bgp l2vpn evpn all route-type 5 0 10.2.255.255 32
```

```
...or you can also use:
Leaf-03#
```

```
sh bgp l2vpn evpn detail [5][1:1][0][32][10.2.255.255]/17
```

```
BGP routing table entry for [5][1:1][0][32][10.2.255.255]/17, version 25
Paths: (2 available, best #1, table EVPN-BGP-Table)
  Not advertised to any peer
  Refresh Epoch 2
  Local
```

```
    172.16.254.4
```

```
 (metric 3) (via default) from 172.16.255.1 (172.16.255.1)
```

```
<-- RP's global next hop IP
```

```
      Origin incomplete, metric 0, localpref 100, valid, internal, best
      EVPN ESI: 00000000000000000000, Gateway Address: 0.0.0.0, VNI Label 50901, MPLS VPN Label 0
      Extended Community: RT:1:1 MVPN AS:65001:0.0.0.0
```

```
        MVPN VRF:172.16.255.4:2
```

```
 ENCAP:8
```

Router MAC:7C21.0DBD.9548

```
      Originator: 172.16.255.4, Cluster list: 172.16.255.1
      rx pathid: 0, tx pathid: 0x0
      Updated on Jan 13 2021 19:09:31 UTC
  Refresh Epoch 2
  Local
```

MVPN VRF:172.16.255.4:2

<-- MVPN VRI

Router MAC:7C21.0DBD.9548 <-- Leaf-02 RMAC

Step 1 (Leaf-03): IGMP Membership report received

<#root>

```
Leaf-03#
```

sh ip igmp snooping groups

```
Vlan     Group            Type                Version        Port List
------------------------------------------------------------------------
102      224.0.1.40       igmp                v2             Gi1/0/10
```

102 226.1.1.1 igmp v2 Gi1/0/10 <-- Client has joined

Step 2 (Leaf-03): MVPN Type-6 created, sent to RP, and received by RP **(Leaf-02)**

<#root>

Type-6 from the Receiver VTEP perspective

```
Leaf-03#
```

sh bgp ipv4 mvpn all route-type 6 1:1 65001 10.2.255.255 226.1.1.1 <-- Source is RP Loopback

```
...or you can also use:
Leaf-03#
```

sh bgp ipv4 mvpn

detail [6][1:1][65001][10.2.255.255/32][226.1.1.1/32]/22

```
BGP routing table entry for [6][1:1][65001][10.2.255.255/32][226.1.1.1/32]/22, version 13
```

```
Paths: (1 available, best #1, table MVPNv4-BGP-Table)
  Advertised to update-groups:
      1
  Refresh Epoch 1
  Local
    0.0.0.0 from 0.0.0.0 (172.16.255.6) <-- Generated locally

      Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
      Extended Community: RT:172.16.255.4:2 <-- VRI Ext Comm added from EVPN Type-5

      rx pathid: 2, tx pathid: 0x0
      Updated on Jan 14 2021 14:51:29 UTC

#### Type-6 from the RP perspective  ###

Leaf-02#
sh bgp ipv4 mvpn all route-type 6 1:1 65001 10.2.255.255 226.1.1.1 <-- type-6, RD 1:1, AS 65001, Source/

...or you can also use:
Leaf-02#
sh bgp ipv4 mvpn detail [6][1:1][65001][10.2.255.255/32][226.1.1.1/32]/22

BGP routing table entry for
[6]
[1:1][65001][10.2.255.255/32][226.1.1.1/32]/22, version 25
Paths: (2 available, best #1, table MVPNv4-BGP-Table)
   Flag: 0x100
   Not advertised to any peer
   Refresh Epoch 2
   Local
     172.16.255.6 (metric 3) from 172.16.255.1 (172.16.255.1)
       Origin incomplete, metric 0, localpref 100, valid, internal, best
       Extended Community: RT:172.16.255.4:2 <-- Contains VRI learned from EVPN Type-5

      Originator: 172.16.255.6
, Cluster list: 172.16.255.1
<-- Sent from Leaf03 IP to RP

      rx pathid: 0, tx pathid: 0x0
      Updated on Jan 14 2021 14:54:29 UTC
```

Step 1 & 2 Debugs (Leaf-01): IGMP Report, EVPN source lookup, and MVPN Type-6 create

<#root>

debug ip igmp vrf green 226.1.1.1

debug bgp ipv4 mvpn updates

debug bgp ipv4 mvpn updates events

Client sends IGMP membership report

IGMP processes this IGMP report

*Feb 1 21:13:19.029: IGMP(2): Received v2 Report on Vlan102 from 10.1.102.12 for 226.1.1.1

<--- IGMP processes received report

*Feb 1 21:13:19.029: IGMP(2): Received Group record for group 226.1.1.1, mode 2 from 10.1.102.12 for 0
*Feb 1 21:13:19.029: IGMP(2): WAVL Insert group: 226.1.1.1 interface: Vlan102 Successful
*Feb 1 21:13:19.029: IGMP(2): Switching to EXCLUDE mode for 226.1.1.1 on Vlan102
*Feb 1 21:13:19.029: IGMP(2): Updating EXCLUDE group timer for 226.1.1.1
Feb 1 21:13:19.029: IGMP(2): MRT Add/Update Vlan102 for (,226.1.1.1) by 0

<--- Notify MRT to add Vlan 102 into Outgoing interface list

BGP is informed by IGMP, does an EVPN source lookup, creates the MVPN Type-6 route, sends to RR

(

Without the EVPN Type-5 prefix already in BGP you see IGMP debugs trigger, but no subsequent BGP debugs

*Feb 1 21:13:19.033: BGP[15] MVPN:

add c-route, type 6

, bs len 0 asn=0, rd=1:1,

<-- Start creation of Type-6 C-multicast Shared Tree Join

*Feb 1 21:13:19.033:

source=10.2.255.255

/4,

<-- RP loopback255

*Feb 1 21:13:19.033: group=226.1.1.1/4,

<-- Group IP

```
*Feb 1 21:13:19.033:
nexthop=172.16.254.4
,
<-- Global Next-Hop learned from EVPN VRI

*Feb 1 21:13:19.033: len left = 0
*Feb 1 21:13:19.033: BGP[14]
MVPN umh lookup:
 vrfid 2, source 10.2.255.255
<-- UMH (upstream multicast hop) as found in the RT of the EVPN type-5

*Feb 1 21:13:19.033: BGP[4] MVPN umh lookup: vrfid 2, source 10.2.255.255, net 1:1:10.2.255.255/32, 1:1
<-- EVPN info adding to MVPN

*Feb 1 21:13:19.033: BGP: MVPN(15) create local route [6][1:1][65001][10.2.255.255/32][226.1.1.1/32]/22
<--- MVPN creating type-6

*Feb 1 21:13:19.033: BGP[15] MVPN: add c-route, type 6, bs len 0 asn=65001, rd=1:1,
*Feb 1 21:13:19.033: source=10.2.255.255/4,
*Feb 1 21:13:19.033: group=226.1.1.1/4,
*Feb 1 21:13:19.033: nexthop=172.16.254.4,
*Feb 1 21:13:19.033: len left = 0
*Feb 1 21:13:19.033: BGP[14] MVPN umh lookup: vrfid 2, source 10.2.255.255
*Feb 1 21:13:19.033: BGP[4] MVPN umh lookup: vrfid 2, source 10.2.255.255, net 1:1:10.2.255.255/32, 1:1
*Feb 1 21:13:19.034: BGP(15): skip vrf default table RIB route [6][1:1][65001][10.2.255.255/32][226.1.1
*Feb 1 21:13:19.034: BGP(15): 172.16.255.1 NEXT_HOP self is set for sourced RT Filter for net [6][1:1][
*Feb 1 21:13:19.034: BGP(15): (base)
172.16.255.1 send UPDATE
 (format) [6][1:1][65001][10.2.255.255/32][226.1.1.1/32]/22, next 172.16.255.6, metric 0, path Local, e
<-- Advertise to RR
 (
172.16.255.1)
```

Step 3 & 4 (Leaf-01):From FHR perspective, validate the S,G create & Register events (S,G create & Register happen nearly at the same time)

3. Data traffic starts and S,G is created at FHR VTEP. The requirements noted in "Undetected Multicast Sources" section apply here.

4. Leaf-01 performs Source registration to RP via its PIM tunnel

```
<#root>
Leaf-01#
debug ip pim vrf green 226.1.1.1
```

PIM debugging is on

Leaf-01#

debug ip mrouting vrf green 226.1.1.1

IP multicast routing debugging is on

Debugs for PIM and Mroute show creation of S,G and PIM register encap event

Jan 29 18:18:37.602: PIM(2): Building Periodic (,G) Join / (S,G,RP-bit) Prune message for 226.1.1.1
*Jan 29 18:18:58.426:

MRT(2): (10.1.101.11,226.1.1.1), RPF install from /0.0.0.0 to Vlan101/10.1.101.11<-- S,G is creation me

*Jan 29 18:18:58.427:

PIM(2): Adding register encap tunnel (Tunnel4) as forwarding interface of (10.1.101.11, 226.1.1.1). <--

Jan 29 18:18:58.427: MRT(2): Set the F-flag for (, 226.1.1.1)
*Jan 29 18:18:58.427: MRT(2): Set the F-flag for (10.1.101.11, 226.1.1.1)
*Jan 29 18:18:58.428:

MRT(2): Create (10.1.101.11,226.1.1.1), RPF (Vlan101, 10.1.101.11, 0/0) <-- S,G is creation message (MR

*Jan 29 18:18:58.428: MRT(2): Set the T-flag for (10.1.101.11, 226.1.1.1)

Tunnel 4 is PIM Register tunnel (Encap: encapsulate in tunnel to RP)

Leaf-01#

sh int tunnel4

Tunnel4 is up, line protocol is up
  Hardware is Tunnel
  Description:

Pim Register Tunnel (Encap) for RP 10.2.255.255 on VRF green <-- VRF green for Leaf-02 RP

  Interface is unnumbered.

Using address of Loopback901 (10.1.255.1) <-- Local Loopback

S,G is created when Source sends data traffic

Leaf-01#

sh ip mroute vrf green 226.1.1.1

IP Multicast Routing Table
<...snip...>
Outgoing interface flags: H - Hardware switched, A - Assert winner, p - PIM Join
 Timers: Uptime/Expires
 Interface state: Interface, Next-Hop or VCD, State/Mode

(*, 226.1.1.1), 00:00:16/stopped, RP 10.2.255.255, flags: SPF
  Incoming interface: Vlan901, RPF nbr 172.16.254.4
  Outgoing interface list: Null

(10.1.101.11, 226.1.1.1)

, 00:00:16/00:02:47, flags: FTGqx

Incoming interface: Vlan101

,

RPF nbr 10.1.101.11

,

Registering <-- S,G created, in Register state, RPF IP is the /32 host prefix for this source

Outgoing interface list:

Vlan901

, Forward/Sparse, 00:00:16/00:02:43

<-- OIF is the L3VNI SVI

Checking S,G in Hardware

Leaf-01#

sh platform software fed switch active ip mfib vrf green 226.1.1.1/32 10.1.101.11 de

MROUTE ENTRY

vrf 2

 (10.1.101.11, 226.1.1.1/32)

<-- VRF 2 is the ID for vrf green

  HW Handle: 140213987784872 Flags: {Svl}

RPF interface: Vlan101

(59)): SVI

<-- RPF is Direct connected on a Local Subnet

  HW Handle:140213987784872 Flags:A
Number of OIF: 2
Flags: 0x4

Pkts : 336   <-- packets that used this adjacency (similar to mfib command, but shown at the FED l

OIF Details:

Vlan101 A <-- Accept interface is programmed correctly

Vlan901 F {Remote} <-- Forward interface is L3VNI SVI

(Adj: 0x5f) <-- Validate this Adj

Htm: 0x7f861cf071b8 Si: 0x7f861cf04838 Di: 0x7f861cf097a8 Rep_ri: 0x7f861ceecb38

Check ADJ 0x5f for next hop details

Leaf-01#

sh platform software fed switch active ip adj

IPV4 Adj entries
dest if_name dst_mac si_hdl ri_hdl pd_flags

adj_id

Last-modified
---- ------- ------- ------ ------ -------- ----- ---

239.1.1.1

nve1.VNI50901

4500.0000.0000 0x7f861ce659b8 0x7f861ce65b68 0x60

0x5f

2021/01/29 17:07:06.568

Dest = MDT default group 239.1.1.1

Outgoing Interface = Nve1 using L3 VNI 50901

Step 4 (Leaf-02): From RP perspective, confirm Source registration reaches the RP and S,G is created.

<#root>

PIM debugs showing PIM register event

Leaf-02#

debug ip pim vrf green 226.1.1.1

PIM debugging is on

Jan 29 18:21:35.500: PIM(2): Building Periodic (,G) Join / (S,G,RP-bit) Prune message for 226.1.1.1

***Jan 29 18:21:35.500: PIM: rp our address <-- Leaf-02 is the RP**

***Jan 29 18:21:41.005: PIM(2): Received v2 Register on Vlan901 from 10.1.255.1 <--- IP of Lo901 on Leaf-0**

***Jan 29 18:21:41.005: for 10.1.101.11, group 226.1.1.1**

***Jan 29 18:21:41.006: PIM(2): Adding register decap tunnel (Tunnel4) as accepting interface of (10.1.101**

*Jan 29 18:21:41.008: PIM(2): Upstream mode for (10.1.101.11, 226.1.1.1) changed from 1 to 2

Tunnel 4 is PIM Register tunnel (decap)

Leaf-02#

sh int tunnel 4

Tunnel4 is up, line protocol is up
  Hardware is Tunnel
  Description:

Pim Register Tunnel (Decap) for RP 10.2.255.255 on VRF green <-- decap side of register tunnel

  Interface is unnumbered.

Using address of Loopback255 (10.2.255.255) <-- RP IP

Mroute debugs show pim Register triggering S,G

Leaf-02#

debug ip mrouting vrf green 226.1.1.1

IP multicast routing debugging is on

*Jan 29 20:44:31.483: MRT(2):

(10.1.101.11,226.1.1.1)

,

RPF install from /0.0.0.0 to Vlan901/172.16.254.3 <-- RPF is to Leaf-01

*Jan 29 20:44:31.485: MRT(2):

Create (10.1.101.11,226.1.1.1), RPF (Vlan901, 172.16.254.3, 200/0) <-- Create the S,G

*Jan 29 20:44:33.458: MRT(2):

Set the T-flag for (10.1.101.11, 226.1.1.1) **<-- Set SPT bit for S,G**

S,G is created and traffic is now sent along the *,G shared tree
Leaf-02#sh ip mroute vrf green

```
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,
       L - Local, P - Pruned, R - RP-bit set, F - Register flag,
       T - SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet,
       X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement,
       U - URD, I - Received Source Specific Host Report,
       Z - Multicast Tunnel, z - MDT-data group sender,
       Y - Joined MDT-data group, y - Sending to MDT-data group,
       G - Received BGP C-Mroute, g - Sent BGP C-Mroute,
       N - Received BGP Shared-Tree Prune, n - BGP C-Mroute suppressed,
       Q - Received BGP S-A Route, q - Sent BGP S-A Route,
       V - RD & Vector, v - Vector, p - PIM Joins on route,
       x - VxLAN group, c - PFP-SA cache created entry,
       * - determined by Assert, # - iif-starg configured on rpf intf,
       e - encap-helper tunnel flag
Outgoing interface flags: H - Hardware switched, A - Assert winner, p - PIM Join
 Timers: Uptime/Expires
 Interface state: Interface, Next-Hop or VCD, State/Mode

(*, 226.1.1.1), 00:05:49/stopped, RP 10.2.255.255, flags:
```

SGx <-- Sparse, Received BGP C-Mroute

Incoming interface: Null, RPF nbr 0.0.0.0 **<-- RP is us (Incoming Interface Null with**

```
  Outgoing interface list:
    Vlan901, Forward/Sparse, 00:05:49/stopped

(
```

10.1.101.11, 226.1.1.1

```
), 00:01:22/00:01:41, flags:
```

PTXgx <-- Pruned, SPT bit, Sent BGP C-Mroute

```
  Incoming interface: Vlan901,
```

RPF nbr 172.16.254.3 **<-- Leaf-01 is RPF next hop**

```
  Outgoing interface list: Null
```

Step 5 (Leaf-02): RP has a receiver, so immediately created Type-7 MVPN Source Tree Join route

<#root>

Leaf-02#

sh ip mroute vrf green 226.1.1.1

<...snip...>

(*, 226.1.1.1)

, 00:02:22/00:00:37, RP 10.2.255.255, flags: SGx
  Incoming interface: Null, RPF nbr 0.0.0.0
  Outgoing interface list:

    Vlan901, Forward/Sparse, 00:02:22/00:00:37   <-- L3 VNI is populated from Receiver BGP Type-6 join

Debugs showing Type-7 creation from RP

Leaf-02#

debug bgp ipv4 mvpn updates

BGP updates debugging is on for address family: MVPNv4 Unicast

Leaf-02#

debug bgp ipv4 mvpn updates events

BGP update events debugging is on for address family: MVPNv4 Unicast

*Jan 29 18:21:41.008: BGP[15]

MVPN: add c-route, type 7

, bs len 0 asn=0, rd=1:1,
*Jan 29 18:21:41.008:

source=10.1.101.11/4,

*Jan 29 18:21:41.008:

group=226.1.1.1/4,

*Jan 29 18:21:41.008:

nexthop=172.16.254.3

,

<-- Leaf-01 Global next hop

*Jan 29 18:21:41.008: len left = 0
*Jan 29 18:21:41.008: BGP[14] MVPN umh lookup: vrfid 2, source 10.1.101.11
*Jan 29 18:21:41.008: BGP[4] MVPN umh lookup: vrfid 2, source 10.1.101.11, net 1:1:10.1.101.11/32, 1:1:

0x10B:172.16.255.3:2

,

<-- This is the VRI picked up from the EVPN Type-2

*Jan 29 18:21:41.009: BGP:

MVPN(15) create local route [7][172.16.254.3:101][65001][10.1.101.11/32][226.1.1.1/32]/22

*Jan 29 18:21:41.009:

BGP[15] MVPN: add c-route, type 7, bs len 0 asn=65001, rd=1:1,

*Jan 29 18:21:41.009: source=10.1.101.11/4,
*Jan 29 18:21:41.009: group=226.1.1.1/4,
*Jan 29 18:21:41.009: nexthop=172.16.254.3,
*Jan 29 18:21:41.009: len left = 0
*Jan 29 18:21:41.009: BGP[14] MVPN umh lookup: vrfid 2, source 10.1.101.11
*Jan 29 18:21:41.009: BGP[4] MVPN umh lookup: vrfid 2, source 10.1.101.11, net 1:1:10.1.101.11/32, 1:1:

Type-7 Locally created on RP and sent to Source Leaf-01

Leaf-02#

sh bgp ipv4 mvpn all

BGP table version is 81, local router ID is 172.16.255.4
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
x best-external, a additional-path, c RIB-compressed,
t secondary path, L long-lived-stale,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

     Network   Next Hop   Metric LocPrf Weight Path

Route Distinguisher: 172.16.254.3:101 <-- Note the VRI is learnt from Leaf-01

 *>

[7][172.16.254.3:101]

[65001]

[10.1.101.11/32][226.1.1.1/32]

/22

<-- [7] = type-7 for this S,G / VRI 172.16.254.3:101 learned from Leaf-01

  0.0.0.0

32768

 ?

<-- 0.0.0.0 locally originated

with local Weight

Step 6 (Leaf-01): Source Leaf-01 receives and installs MVPN Route-Type 7. (L3 VNI SVI is installed as a forwarding Interface for S,G)

<#root>

Received Type-7 from Leaf-02 RP

Leaf-01#

debug bgp ipv4 mvpn updates

BGP updates debugging is on for address family: MVPNv4 Unicast
Leaf-01#

debug bgp ipv4 mvpn updates events

BGP update events debugging is on for address family: MVPNv4 Unicast

*Jan 29 18:18:58.457:

BGP(15): 172.16.255.1 rcvd UPDATE w/ attr: nexthop 172.16.255.4, origin ?, localpref 100, metric 0, orig

*Jan 29 18:18:58.457: BGP(15): 172.16.255.1

rcvd [7][172.16.254.3:101][65001][10.1.101.11/32][226.1.1.1/32]/22 <-- Received |

*Jan 29 18:18:58.457: BGP(15): skip vrf default table RIB route [7][172.16.254.3:101][65001][10.1.101.1
*Jan 29 18:18:58.458: BGP(15): add RIB route (0:0)[7][1:1][65001][10.1.101.11/32][226.1.1.1/32]/22

PIM updated by MVPN to install L3 VNI in Outgoing Interface List

Leaf-01#

debug ip pim vrf green 226.1.1.1

PIM debugging is on
Leaf-01#

debug ip mrouting vrf green 226.1.1.1

IP multicast routing debugging is on

*Jan 29 18:18:58.458: PIM(2):

Join-list: (10.1.101.11/32, 226.1.1.1), S-bit set, BGP C-Route

*Jan 29 18:18:58.459: MRT(2):

```
WAVL Insert VxLAN interface: Vlan901 in (10.1.101.11,226.1.1.1) Next-hop: 239.1.1.1 VNI 50901 Successful


*Jan 29 18:18:58.459: MRT(2): set min mtu for (10.1.101.11, 226.1.1.1) 18010->9198
*Jan 29 18:18:58.460:

MRT(2): Add Vlan901/239.1.1.1/50901 to the olist of (10.1.101.11, 226.1.1.1), Forward state - MAC not bu


*Jan 29 18:18:58.460: PIM(2): Add Vlan901/0.0.0.0 to (10.1.101.11, 226.1.1.1), Forward state, by BGP SG
*Jan 29 18:18:58.460: MRT(2): Add Vlan901/239.1.1.1/50901to the olist of (10.1.101.11, 226.1.1.1), Forw
```

Step 7 (Leaf-01): Leaf-01 advertises MVPN Source A-D Type-5 for S,G

```
<#root>

Leaf-01#

debug bgp ipv4 mvpn updates


BGP updates debugging is on for address family: MVPNv4 Unicast
Leaf-01#

debug bgp ipv4 mvpn updates events


BGP update events debugging is on for address family: MVPNv4 Unicast

*Jan 29 18:18:58.461: BGP(15): nettable_walker

[5][1:1][10.1.101.11][226.1.1.1]/18 route sourced locally <-- BGP determines route is local to Leaf-01


*Jan 29 18:18:58.461: BGP(15): delete RIB route (0:0)[5][1:1][10.1.101.11][226.1.1.1]/18
*Jan 29 18:18:58.461: BGP(15): 172.16.255.1 NEXT_HOP self is set for sourced RT Filter for net [5][1:1]
*Jan 29 18:18:58.461: BGP(15): (base) 172.16.255.1

send UPDATE (format) [5][1:1][10.1.101.11][226.1.1.1]/18, next 172.16.255.3, metric 0, path Local, exter
```

Step 8 (Leaf-03): Receiver VTEP gets the Type-5 and installs Source A-D route for S,G

```
<#root>

Leaf-03#

debug bgp ipv4 mvpn updates


BGP updates debugging is on for address family: MVPNv4 Unicast
Leaf-03#

debug bgp ipv4 mvpn updates events


BGP update events debugging is on for address family: MVPNv4 Unicast

*Jan 29 19:18:53.318: BGP(15): 172.16.255.1 rcvd UPDATE w/ attr: nexthop 172.16.255.3, origin ?, localp

*Jan 29 19:18:53.319: BGP(15): 172.16.255.1 rcvd [5][1:1][10.1.101.11][226.1.1.1]/18   <-- Type-5 Receiv
```

```
*Jan 29 19:18:53.319: BGP(15): skip vrf default table RIB route [5][1:1][10.1.101.11][226.1.1.1]/18


Leaf-03#
```

sh bgp ipv4 mvpn all route-type 5 10.1.101.11 226.1.1.1
...or you can also use:

```
Leaf-03#
```

sh bgp ipv4 mvpn detail [5][1:1][10.1.101.11][226.1.1.1]/18

```
BGP routing table entry for
```

[5][1:1][10.1.101.11][226.1.1.1]/18

```
, version 41
```

<-- Type-5 A-D route from Leaf-01

```
Paths: (2 available, best #2, table MVPNv4-BGP-Table, not advertised to EBGP peer)
Flag: 0x100
Not advertised to any peer
Refresh Epoch 1
Local
```

172.16.255.3

```
 (metric 3) from 172.16.255.1 (172.16.255.1)
```

<-- Leaf-01 IP

```
    Origin incomplete, metric 0, localpref 100, valid, internal, best
    Community: no-export
     Extended Community: RT:1:1
```

Originator: 172.16.255.3

```
, Cluster list: 172.16.255.1
     rx pathid: 0, tx pathid: 0x0
     Updated on Jan 29 2021 19:18:53 UTC
```

Step 9 (Leaf-03): S,G is created, Leaf-03 sends MVPN Type-7 to join SPT tree, and starts to accept traffic

<#root>

```
debug ip mrouting vrf green 226.1.1.1
debug bgp ipv4 mvpn updates
debug bgp ipv4 mvpn updates events
```

Debug of Mrouting shows S,G create and call to BGP to create Type-7 BGP S,G join

```
*Feb 12 19:34:26.045:
```

MRT(2):

(10.1.101.11,226.1.1.1), RPF install from /0.0.0.0 to Vlan901/172.16.254.3 <-- RPF check done as first c

*Feb 12 19:34:26.046:

MRT(2):

Create (10.1.101.11,226.1.1.1), RPF (Vlan901, 172.16.254.3, 200/0) <-- RPF successful Creating S,G

*Feb 12 19:34:26.047: MRT(2): WAVL Insert interface: Vlan102 in (10.1.101.11,226.1.1.1) Successful
*Feb 12 19:34:26.047: MRT(2): set min mtu for (10.1.101.11, 226.1.1.1) 18010->9198
*Feb 12 19:34:26.047: MRT(2): Set the T-flag for (10.1.101.11, 226.1.1.1)
*Feb 12 19:34:26.048:

MRT(2):

Add Vlan102/226.1.1.1 to the olist of (10.1.101.11, 226.1.1.1)

, Forward state - MAC not built

<-- Adding Vlan102 Receiver SVI into OIF list

*Feb 12 19:34:26.048:

MRT(2): Set BGP Src-Active for (10.1.101.11, 226.1.1.1) <-- Signaling to BGP that this Source is seen a

BGP Type-7 created

Leaf-03#

sh bgp ipv4 mvpn all

Route Distinguisher:

172.16.254.3:101   <-- VRI Route Distinguisher

 *>

[7]

[

172.16.254.3:101]

[65001]

[10.1.101.11/32][226.1.1.1/32]

/22

<-- Type [7], VRI, S,G info

0.0.0.0

                            32768 ?

<-- created locally

```
Leaf-03#
```

```
sh ip mroute vrf green 226.1.1.1 10.1.101.11
```

```
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,
       L - Local, P - Pruned, R - RP-bit set, F - Register flag,
```

```
 T - SPT-bit set
```

```
, J - Join SPT, M - MSDP created entry, E - Extranet,
       X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement,
       U - URD, I - Received Source Specific Host Report,
       Z - Multicast Tunnel, z - MDT-data group sender,
       Y - Joined MDT-data group, y - Sending to MDT-data group,
       G - Received BGP C-Mroute,
```

```
g - Sent BGP C-Mroute
```

```
,
       N - Received BGP Shared-Tree Prune, n - BGP C-Mroute suppressed,
```

```
Q - Received BGP S-A Route
```

```
, q - Sent BGP S-A Route,
       V - RD & Vector, v - Vector, p - PIM Joins on route,
       x - VxLAN group, c - PFP-SA cache created entry,
       * - determined by Assert, # - iif-starg configured on rpf intf,
       e - encap-helper tunnel flag
Outgoing interface flags: H - Hardware switched, A - Assert winner, p - PIM Join
 Timers: Uptime/Expires
 Interface state: Interface, Next-Hop or VCD, State/Mode
```

```
(10.1.101.11, 226.1.1.1), 00:08:41/00:02:13,
```

```
flags: TgQ <-- SPT bit, Sent MVPN type-7, Received MVPN type-5
```

```
 Incoming interface: Vlan901, RPF nbr 172.16.254.3     <-- Receive from L3VNI via Leaf-01 IP next hop
```

```
  Outgoing interface list:
```

```
   Vlan102, Forward/Sparse, 00:08:41/00:02:22          <-- Send to host in Vlan 102
```

Step 10 (Leaf-01): Leaf-01 receives and installs MVPN Type-7 from Leaf-03

<#root>

```
debug bgp ipv4 mvpn updates
```

```
debug bgp ipv4 mvpn updates events
```

```
### Type-7 Received from Leaf-03 VTEP and installed into RIB ###
```

```
*Feb 12 19:55:29.000: BGP(15): 172.16.255.1
rcvd [7][172.16.254.3:101][65001][10.1.101.11/32][226.1.1.1/32]/22   <-- Type-7 from Leaf-03


*Feb 12 19:55:29.000: BGP(15): skip vrf default table RIB route [7][172.16.254.3:101][65001][10.1.101.1
*Feb 12 19:55:29.000: BGP(15): add RIB route (0:0)[7][1:1][65001][10.1.101.11/32][226.1.1.1/32]/22
```

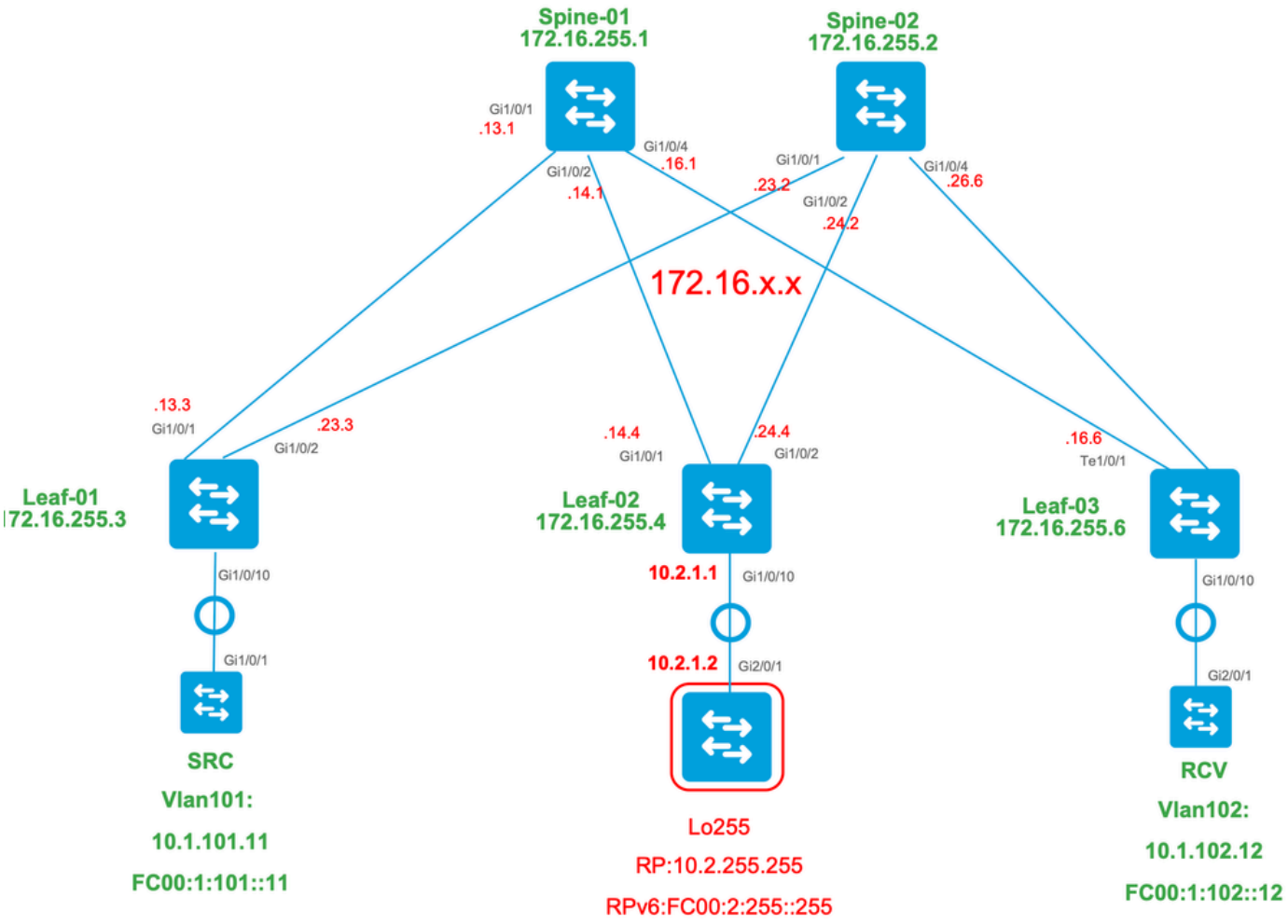
Scenario 4: RP Outside the Fabric (RP Imported from Border Leaf-02 from IP space)

This scenario is basically the same as Scenario 3. There is a single RP used by the Fabric overall. The difference is the RP IP must be imported from a non-fabric IP space into Fabric and advertised into BGP.

This section shows the differences from Scenario 3. The steps and methods that are the same are noted only in Scenario 3

- See **Verify the Sequence of Events Required for this Scenario** from Scenario 3 as the BGP and PIM operations are the same

Network Diagram



Verify Border switch imports from IP to Fabric

The primary difference with this design versus Scenario 3 is the need to first import the RP IP from the IP space to EVPN.

The Border needs to contain certain commands to import/export to and from fabric & IP spaces:

- **route-target <value> stitching** commands under the VRF configuration section
- **advertise l2vpn evpn** under the BGP vrf address-family

Verify (Leaf-02): Configuration

<#root>

```
Leaf-02#

sh run vrf green
Building configuration...

Current configuration : 1533 bytes

vrf definition green
rd 1:1
!
address-family ipv4
mdt auto-discovery vxlan
mdt default vxlan 239.1.1.1
mdt overlay use-bgp

route-target export 1:1



route-target import 1:1



route-target export 1:1 stitching    <-- BGP-EVPN fabric redistributes the stitching routes between the



route-target import 1:1 stitching


exit-address-family

Leaf-02#

sh run | sec router bgp



address-family ipv4 vrf green       <--- BGP VRF green address-family



advertise l2vpn evpn                <--- Use the 'advertise l2vpn evpn' command and 'export stitching' F


redistribute connected
redistribute static

redistribute ospf 2 match internal external 1 external 2   <-- Learning via external OSPF neighbor in VR


exit-address-family
```

Verify (Leaf-02): Prefix Import and Advertisement

<#root>

debug bgp vpnv4 unicast updates

debug bgp vpnv4 unicast updates events

debug bgp l2vpn evpn updates

debug bgp l2vpn evpn updates events

***Feb 15 15:30:54.407: BGP(4): redist event (1) request for 1:1:10.2.255.255/32**

*Feb 15 15:30:54.407: BGP(4) route 1:1:10.2.255.255/32 gw-1 10.2.1.2 src_proto (ospf) path-limit 1
*Feb 15 15:30:54.407: BGP(4): route 1:1:10.2.255.255/32 up
*Feb 15 15:30:54.407: bgp_ipv4set_origin: redist 1, opaque 0x0, net 10.2.255.255
*Feb 15 15:30:54.407: BGP(4): sourced route for 1:1:10.2.255.255/32 path 0x7FF8065EB9C0 id 0 gw 10.2.1.
*Feb 15 15:30:54.408: BGP(4): redistributed route 1:1:10.2.255.255/32 added gw 10.2.1.2
*Feb 15 15:30:54.408: BGP: topo green:VPNv4 Unicast:base Remove_fwdroute for 1:1:10.2.255.255/32

***Feb 15 15:30:54.408: BGP(4): 1:1:10.2.255.255/32 import vpn re-orig or locally sourced or learnt from C**

***Feb 15 15:30:54.409: BGP(10): update modified for [5][1:1][0][32][10.2.255.255]/17**

*Feb 15 15:30:54.409: BGP(10): 172.16.255.1

NEXT_HOP set to vxlan local vtep-ip 172.16.254.4

for net [5][1:1][0][32][10.2.255.255]/17 <-- Set NH to Leaf-02 loopback

*Feb 15 15:30:54.409: BGP(10): update modified for [5][1:1][0][32][10.2.255.255]/17

***Feb 15 15:30:54.409: BGP(10): (base) 172.16.255.1 send UPDATE (format) [5][1:1][0][32][10.2.255.255]/17**
<-- BGP EVPN Type update created from Non-fabric Imported prefix and sent to RR

Verify the NLRI is learned and Imported on Border Leaf-02

Leaf-02#

sh bgp vpnv4 unicast all

BGP table version is 39, local router ID is 172.16.255.4
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,

```
              r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
              t secondary path, L long-lived-stale,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

 Network Next Hop Metric LocPrf Weight Path
Route Distinguisher: 1:1 (default for vrf green)
```

AF-Private Import to Address-Family: L2VPN E-VPN, Pfx Count/Limit: 3/1000 <-- Prefix Import details. (N

```
 *>
```

10.2.255.255/32 10.2.1.2   2   32768 ?   <-- Locally redistributed, Next hop

```
Leaf-02#
```

sh bgp l2vpn evpn all route-type 5 0 10.2.255.255 32

```
...or you can also use:
Leaf-02#
```

sh bgp l2vpn evpn detail [5][1:1][0][32][10.2.255.255]/17

```
BGP routing table entry for
```

[5][1:1][0][32][10.2.255.255]

```
/17, version 69
Paths: (1 available, best #1, table EVPN-BGP-Table)
  Advertised to update-groups:
     2
  Refresh Epoch 1
```

Local, imported path from base

10.2.1.2 (via vrf green) from 0.0.0.0 (172.16.255.4)   <-- Imported to EVPN Fabric table fr

```
      Origin incomplete, metric 2, localpref 100, weight 32768, valid, external, best
      EVPN ESI: 00000000000000000000, Gateway Address: 0.0.0.0,
```

local vtep: 172.16.254.4, VNI Label 50901,

```
MPLS VPN Label 17
```

<-- VTEP IP of Leaf-02, L3VNI label

```
      Extended Community: RT:1:1 OSPF DOMAIN ID:0x0005:0x000000020200
        MVPN AS:65001:0.0.0.0
```

MVPN VRF:172.16.255.4:2

```
 ENCAP:8
```

<-- MVPN VRI created

```
        Router MAC:7C21.0DBD.9548 OSPF RT:0.0.0.0:2:0
        OSPF ROUTER ID:10.2.255.255:0
      rx pathid: 0, tx pathid: 0x0
```

```
      Updated on Feb 15 2021 15:30:54 UTC
```

Verify (Leaf-02): Border Path to RP

<#root>

Leaf-02#sh ip mroute vrf green

```
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,
       L - Local, P - Pruned, R - RP-bit set, F - Register flag,
       T - SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet,
       X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement,
       U - URD, I - Received Source Specific Host Report,
       Z - Multicast Tunnel, z - MDT-data group sender,
       Y - Joined MDT-data group, y - Sending to MDT-data group,
       G - Received BGP C-Mroute, g - Sent BGP C-Mroute,
       N - Received BGP Shared-Tree Prune, n - BGP C-Mroute suppressed,
       Q - Received BGP S-A Route, q - Sent BGP S-A Route,
       V - RD & Vector, v - Vector, p - PIM Joins on route,
       x - VxLAN group, c - PFP-SA cache created entry,
       * - determined by Assert, # - iif-starg configured on rpf intf,
       e - encap-helper tunnel flag
Outgoing interface flags: H - Hardware switched, A - Assert winner, p - PIM Join
 Timers: Uptime/Expires
 Interface state: Interface, Next-Hop or VCD, State/Mode
```

(*, 226.1.1.1)

, 2d21h/stopped,

RP 10.2.255.255

, flags: SJGx

<-- *,G for group and Non-fabric RP IP

Incoming interface: Vlan2001

,

RPF nbr 10.2.1.2 <-- RPF neighbor is populated for IP next hop outside VxLAN

Outgoing interface list:

Vlan901, Forward/Sparse, 01:28:47/stopped <-- Outgoing is L3VNI SVI

Scenario 5: Data MDT

Verify MDT Data Group

The MDT Data group is similar to ther MDT Default group where outer tunnel group for TRM to be encapsulated in. However, unlike the MDT Default, this group will only have VTEP's join this tree if they have interested receivers for the TRM group.

Required Configuration

<#root>

```
vrf definition green
rd 1:1
!
address-family ipv4
mdt auto-discovery vxlan
mdt default vxlan 239.1.1.1
```

mdt data vxlan 239.1.2.0 0.0.0.255 <-- Defines MDT Data underlay group address range

mdt data threshold 1

<-- Defines the threshold before cutting over to the Data group (In Kilobits per second)

```
mdt overlay use-bgp spt-only
route-target export 1:1
route-target import 1:1
route-target export 1:1 stitching
route-target import 1:1 stitching
exit-address-family
!
```

Check that the MDT group is correctly programmed on Source side

- Incoming interface of MDT group is the Source side Loopback
- Outgoing interface of MDT group is the Underlay Interface

Verify Leaf-01: the MDT mroute is correct in MRIB/MFIB

<#root>

```
Leaf-01#
```

sh ip mroute 239.1.2.0 172.16.254.3

```
<snip>
```

(172.16.254.3, 239.1.2.0)

```
, 00:01:19/00:02:10, flags: FT
  Incoming interface:
```

Loopback1

```
, RPF nbr
```

```
0.0.0.0


<-- IIF is local loopback with 0.0.0.0 RPF indicating local

  Outgoing interface list:
    TenGigabitEthernet1/0/1
, Forward/Sparse, 00:01:19/00:03:10
<-- OIF is the underlay uplink


Leaf-01#
sh ip mfib 239.1.2.0 172.16.254.3


<snip>
(172.16.254.3,239.1.2.0) Flags: HW
  SW Forwarding: 2/0/828/0, Other: 0/0/0
  HW Forwarding: 450/2/834/13
, Other: 0/0/0
<-- Hardware counters indicate the entry is operating in hardware and forwarding packets


  Null0 Flags: A                              <-- Null0 (Originated locally)


  TenGigabitEthernet1/0/1
 Flags: F NS
<-- OIF is into the Underlay (Global routing table)

    Pkts: 0/0/0 Rate: 0 pps
```

Verify Leaf-01: FED entries for the MDT group

```
<#root>
Leaf-01#
show platform software fed switch active ip mfib 239.1.2.0/32 172.16.254.3 detail <-- The detail option

MROUTE ENTRY
vrf 0 (172.16.254.3, 239.1.2.0/32)  <-- vrf 0 = global for this MDT Data S,G pair

  HW Handle: 140028029798744 Flags:
RPF interface: Null0
```

```
(1)):
<-- Leaf-01 is the Source(Null0)


  HW Handle:140028029798744 Flags:A
Number of OIF: 2
Flags: 0x4 Pkts : 570
<-- Packets that used this adjacency (similar to the mfib command, but shown at the FED layer)


OIF Details:
     TenGigabitEthernet1/0/1 F NS               <-- The Underlay Outgoing Interface and F-Forward flag


     Null0 A                                    <-- The Incoming Interface is local loopback1 and A-Acc


Htm: 0x7f5ad0fa48b8 Si: 0x7f5ad0fa4258
Di: 0x7f5ad0fa8948
 Rep_ri: 0x7f5ad0fa8e28
<--The DI (dest index) handle




DI details


----------
Handle:0x7f5ad0fa8948 Res-Type:ASIC_RSC_DI Res-Switch-Num:255 Asic-Num:255 Feature-ID:AL_FID_L3_MULTICAS
priv_ri/priv_si Handle:(nil) Hardware Indices/Handles:
index0:0x536e
 mtu_index/l3u_ri_index0:0x0
index1:0x536e
 mtu_index/l3u_ri_index1:0x0 index2:0x536e mtu_index/l3u_ri_index2:0x0 index3:0x536e mtu_index/l3u_ri_i
<snip>
Brief Resource Information (ASIC_INSTANCE# 3)
---------------------------------------
Destination index = 0x536e


pmap = 0x00000000 0x00000001
pmap_intf : [TenGigabitEthernet1/0/1]  <--FED has the correct programing of the OIF


============================================================
```

Check that the MDT group is correctly programmed on Receiver side

- Incoming interface of MDT group is the RPF interface back to the Source side Loopback
- Outgoing interface of MDT group is Encap/Decap Tunnel interface

Verify Leaf-02: the MDT mroute is correct in MRIB/MFIB

```
<#root>

Leaf-03#

sh ip mroute 239.1.2.0 172.16.254.3              <-- This is the Global MDT Data Group


<snip>
(

172.16.254.3, 239.1.2.0

), 00:06:12/00:02:50, flags: JTx

<-- Source is Leaf-01 Loopback1 IP


  Incoming interface: TenGigabitEthernet1/0/1, RPF nbr 172.16.26.2
  Outgoing interface list:

    Tunnel0

, Forward/Sparse, 00:06:12/00:02:47

<-- Decap Tunnel



Leaf-03#

sh ip mfib 239.1.2.0 172.16.254.3


<snip>

Default                                                 <-- Global Routing Table


  (

172.16.254.3,239.1.2.0

) Flags: HW
    SW Forwarding: 2/0/828/0, Other: 0/0/0

    HW Forwarding: 760/2/846/13

, Other: 0/0/0

<-- Hardware counters indicate the entry is operating in hardware and forwarding packets


    TenGigabitEthernet1/0/1 Flags: A                    <-- Accept via Underlay (Global) interface


    Tunnel0, VXLAN Decap Flags: F NS                    <-- Forward to VxLAN Decap Tunnel

      Pkts: 0/0/2 Rate: 0 pps
```

Verify Leaf-02: FED entries for the MDT group

```
<#root>

Leaf-03#

show platform software fed switch active ip mfib 239.1.2.0/32 172.16.254.3 detail

MROUTE ENTRY

vrf 0 (172.16.254.3, 239.1.2.0/32)  <-- vrf 0 = global for this MDT Data S,G pair

  HW Handle: 140592885196696 Flags:

RPF interface: TenGigabitEthernet1/0/1

(55)):

<-- RPF Interface to 172.16.254.3

  HW Handle:140592885196696 Flags:A
Number of OIF: 2
Flags: 0x4

Pkts : 800                          <-- packets that used this adjacency (similar to mfib command, but

OIF Details:

     TenGigabitEthernet1/0/1 A                  <-- Accept MDT packets from this interface

     Tunnel0 F NS                               <-- Forward to Decap Tunnel to remove VxLAN header

          (Adj: 0x3c )                          <-- Tunnel0 Adjacency

Htm: 0x7fde54fb7d68 Si: 0x7fde54fb50d8 Di: 0x7fde54fb4948 Rep_ri: 0x7fde54fb4c58

<snip>

RI details                                      <-- Rewrite Index is used for VxLAN decapsulation

----------
Handle:0x7fde54fb4c58 Res-Type:ASIC_RSC_RI_REP Res-Switch-Num:255 Asic-Num:255 Feature-ID:AL_FID_L3_MUL
priv_ri/priv_si Handle:(nil) Hardware Indices/Handles: index0:0x1a mtu_index/l3u_ri_index0:0x0 index1:0

Brief Resource Information (ASIC_INSTANCE# 0)
---------------------------------------
ASIC# 0
Replication list :
------------------

Total #ri : 6
Start_ri : 26
```

```
Common_ret : 0

Replication entry

rep_ri 0x1A

 #elem = 1
0)

ri[0]=0xE803

 Dynamic port=88ri_ref_count:1 dirty=0
<snip>

Leaf-03#

show platfomr software fed switch active fwd-asic resource asic all rewrite-index range 0xE803 0XE803


ASIC#:0 RI:59395

Rewrite_type

:AL_RRM_REWRITE_L2_PAYLOAD_

IPV4_EVPN_DECAP

(118) Mapped_rii:LVX_EVPN_DECAP(143)
<snip>
```

Debug MDT Data Group

Use the MVPN debug to check the Data MDT cutover event

Source Side VTEP

```
<#root>

Leaf#

debug mvpn


<snip>
*Mar 27 12:12:11.115: MVPN: Received local withdraw for (10.1.101.11, 239.1.1.1) with RD: 1:1, Route Ty
*Mar 27 12:12:11.115: MVPN: Sending BGP prefix=[5: 1:1 : (10.1.101.11,239.1.1.1)] len=19, nh 0.0.0.0, W
*Mar 27 12:12:11.115: MVPN: Route Type 5 deleted [(10.1.101.11, 239.1.1.1), nh 0.0.0.0] rd:1:1 send:1
*Mar 27 12:12:11.115: MVPN: Received BGP prefix=[5: 1:1 : (10.1.101.11,239.1.1.1)] len=19, nexthop: UNK
*Mar 27 12:12:11.115: MVPN: Received BGP withdraw for (10.1.101.11, 239.1.1.1) with RD: 1:1, Route Type
*Mar 27 12:13:00.430: MVPN: Received local route update for (10.1.101.11, 239.1.1.1) with RD: 1:1, Rout
*Mar 27 12:13:00.431: MVPN: Route Type 5 added [(10.1.101.11, 239.1.1.1), nh 0.0.0.0] rd:1:1 send:1
*Mar 27 12:13:00.431: MVPN: RP 10.2.255.255 updated in newly created route
*Mar 27 12:13:00.431: MVPN: Sending BGP prefix=[5: 1:1 : (10.1.101.11,239.1.1.1)] len=19, nh 0.0.0.0, O
*Mar 27 12:13:00.431: MVPN: Received BGP prefix=[5: 1:1 : (10.1.101.11,239.1.1.1)] len=19, nexthop: UNK
*Mar 27 12:13:00.431: MVPN: Received BGP withdraw for (10.1.101.11, 239.1.1.1) with RD: 1:1, Route Type
*Mar 27 12:13:17.151:

MVPN(green[AF_IPv4]): Successfully notified nve fordatamdt adjacency create 239.1.2.0


<-- Notify NVE about creating DATA MDT
```

*Mar 27 12:13:17.151:

MVPN: Received local update <104:0x00:0>(172.16.254.3, 239.1.2.0) next_hop:0.0.0.0 router_id:172.16.255.

*Mar 27 12:13:17.151:

MVPN: LSM AD route added [(10.1.101.11,239.1.1.1) : <104:0x00:0>(172.16.254.3, 239.1.2.0)] orig:172.16.2

*Mar 27 12:13:17.151:

MVPN(green[AF_IPv4]): Sending VxLAN BGP AD prefix=[3:172.16.255.3 1:1 : (10.1.101.11,239.1.1.1)] len=23,

*Mar 27 12:13:17.151:

MVPN(green[AF_IPv4]): Originate VxLAN BGP AD rt:3

*Mar 27 12:13:17.151:

MVPN(green[AF_IPv4]): VXLAN MDT-Data, node added for (10.1.101.11,239.1.1.1) MDT: 239.1.2.0

Leaf-01#

Receiver Side VTEP

<#root>

Leaf#

debug mvpn

<snip>
*Mar 27 12:27:54.920: MVPN: Received BGP prefix=[5: 1:1 : (10.1.101.11,239.1.1.1)] len=19, nexthop: 172
*Mar 27 12:27:54.920: MVPN: Received BGP route update for (10.1.101.11, 239.1.1.1) with RD: 1:1, Route
*Mar 27 12:27:54.920: MVPN: Route Type 5 found [(10.1.101.11, 239.1.1.1), nh 172.16.255.3]rd:1:1 send:0
*Mar 27 12:27:54.920: MVPN: Received BGP prefix=[5: 1:1 : (10.1.101.11,239.1.1.1)] len=19, nexthop: UNK
*Mar 27 12:27:54.920: MVPN: Received BGP withdraw for (10.1.101.11, 239.1.1.1) with RD: 1:1, Route Type
*Mar 27 12:27:54.920: MVPN: Route Type 5 deleted [(10.1.101.11, 239.1.1.1), nh 172.16.255.3] rd:1:1 sen
*Mar 27 12:28:27.648: MVPN: Received BGP prefix=[5: 1:1 : (10.1.101.11,239.1.1.1)] len=19, nexthop: UNK
*Mar 27 12:28:27.657: MVPN: Received BGP withdraw for (10.1.101.11, 239.1.1.1) with RD: 1:1, Route Type
*Mar 27 12:28:44.235: MVPN: Received BGP prefix=[5: 1:1 : (10.1.101.11,239.1.1.1)] len=19, nexthop: 172
*Mar 27 12:28:44.235: MVPN: Received BGP route update for (10.1.101.11, 239.1.1.1) with RD: 1:1, Route
*Mar 27 12:28:44.235: MVPN: Route Type 5 added [(10.1.101.11, 239.1.1.1), nh 172.16.255.3] rd:1:1 send:
*Mar 27 12:29:00.956: MVPN: Received BGP prefix=[3:172.16.255.3 1:1 : (10.1.101.11,239.1.1.1)] len=23,
*Mar 27 12:29:00.956: MVPN: Received BGP prefix=[3:172.16.255.3 1:1 : (10.1.101.11,239.1.1.1)] len=23,
*Mar 27 12:29:00.956:

MVPN: Received BGP update <104:0x00:50901>(172.16.254.3, 239.1.2.0) next_hop:172.16.255.3 router_id:172.

*Mar 27 12:29:00.956:

MVPN: LSM AD route added [(10.1.101.11,239.1.1.1) : <104:0x00:50901>(172.16.254.3, 239.1.2.0)] orig:172.

*Mar 27 12:29:00.957:

MVPN(green[AF_IPv4]): Activating PE (172.16.255.3, 1:1) ad route refcnt:1 control plane refcnt: 0

```
*Mar 27 12:29:00.958:
MVPN(green[AF_IPv4]): Successfully notified datamdt group for NVE (239.1.2.0, TRUE, FALSE)

*Mar 27 12:29:00.958: MVPN: Received BGP update <104:0x00:50901>(172.16.254.3, 239.1.2.0) next_hop:172.
Leaf-03#
```

Troubleshoot

Undetected Multicast Sources

Before you look at why a multicast flow does not work, it is important to understand the relationship of ARP and multicast forwarding

Usually when a host becomes active and sends traffic, ARP entries are completed by the regular source detection pro
multicast sources, it is possible that source starts to send traffic and L2 plane at the FHR processes this multicast traf
the source.

ARP completion plays an important role in the TRM functionality for two reasons.

1. **"Directly connected" check** at the first hop router invokes a FIB API which in turn depends upon ARP comp
   towards the multicast source is not completed, CEF adjacency towards the source remains incomplete and dire
   FALSE.
2. Source detection triggers **advertisement of EVPN RT-2 in the EVPN fabric**. This EVPN route installed in L
   the RPF route towards the source. Thus, if the source is undetected, RPF for the (S,G) entry cannot be found.
   NULL or a less specific route (if present) is installed in the RIB.

Please ensure that ARP is resolved and the Source is reachable within the EVPN fabric.

Other Helpful Debugs

In this section are other debugs that can be helpful in isolation of TRM issues

- **debug mvpn** (all MVPN events, see Scenario 2 for example)
- **debug ip|ipv6 pim <vrf>** (PIM protocol activity)
- **debug ip mrib <vrf> trans** (MRIB, classic PIM translation)
- **debug ip mfib <vrf> pak|ps|fs** (Packet forwarding| Process switching| Fast switching)

Sources and Receivers Outside the Fabric

In some cases, the Source and/or receiver can live one or more L3 hops away from the fabric VTEP(s).

This is a valid design, but changes what EVPN route type carriers the VRI, and what process is responsible for the c
VTEP.

- If Source is outside the fabric the ingress VTEP sees the source via a PIM neighbor, not a Directly Connected Receiver VTEP. The VRI is contained in this Type-5.
- If Receiver is outside the fabric, the join comes via a PIM join IGMP. Information in the PIM join is used to c

eBGP Multiple AS (Spine to Spine) Topology

In some cases the topology can require BGP to send update information to another AS/Fabric.

It is possible for up to 30 seconds elapse for BGP control plane information to converge, and multicast to start to wo

- **This is due to the default eBGP advertisement interval of 30 seconds**.
- If there is an issue with long convergence times due to delay in BGP updates, the eBGP advertisement i updates more frequently.
- Consult the BGP Configuration guide in the Reference section of this article for more information on th

eBGP inter-as **requires** an additional command

Use the inter-as keyword for the MVPN address family routes to cross the BGP autonomous system (AS) boundarie

<#root>

```
Border-Leaf(config-vrf-af)#
```

```
mdt auto-discovery vxlan inter-as
```

Register Tunnel with Symmetric L2VNI (FHR Stuck in PIM Register State)

In cases where the VNI exists on the FHR and on other VTEPs, it is possible to have the FHR become stuck in

This is due to the fact that the PIM Register Tunnel source IP is the AnyCast gateway. When the RP receives a PIM the right VTEP to send the register stop since the IP is common for multiple devices.

PIM Register Tunnel Route Issue

(Leaf-01) This is the actual FHR: Sends Register messages to RP

<#root>

```
Leaf-01#sh ip pim vrf green tunnel
```

```
Tunnel5*
Type : PIM Encap
RP : 10.2.255.255
```

```
Source : 10.1.101.1 <-- Source of Register Tunnel
```

```
State : UP
Last event : Created (00:33:28)
```

(Leaf-03): This VTEP (and possibly others) contains the same SVI and IP address as the FHR

<#root>

Leaf-03#sh ip pim vrf green tunnel

Tunnel4
Type : PIM Encap
RP : 10.2.255.255

Source : 10.1.101.1 <-- Source of Register Tunnel

State : UP
Last event : Created (00:11:53)

(Leaf-01): The FHR remains stuck in Register (It does not receive a register-stop from the RP)

<#root>

Leaf-01#

show ip mroute vrf green 226.1.1.1 10.1.101.11

(10.1.101.11, 226.1.1.1), 02:02:19/00:02:22, flags: PFT
Incoming interface: Vlan101, RPF nbr 10.1.101.11,

Registering <-- Leaf-01 is stuck in register state

Outgoing interface list: Null

(Leaf-02) This is the RP: In this case it also owns the same AnyCast IP as the FHR, and thereby sends the register-

If RP does not have the l2vni but 2 or 3 other vteps do, register-stop could be sent to the wrong VTEP as the RP has

<#root>

Leaf-02#

sh ip route vrf green 10.1.101.1

Routing Table: green
Routing entry for 10.1.101.1/32

Known via "connected"

, distance 0, metric 0 (connected)
Routing Descriptor Blocks:
*

directly connected, via Vlan101 <-- Leaf-02 sees IP as Connected, and sends the Register-stop to itself

Route metric is 0, traffic share count is 1

(Leaf-02): Debug on RP shows the issue where RP has this route as Connected Local

```
<#root>

Leaf-02#

debug ip pim vrf green 226.1.1.1


PIM debugging is on
*May 26 17:33:15.797: PIM(2)[green]:

Received v2 Register on Vlan901 from 10.1.101.1                    <-- Received from Leaf-01 with Sou


*May 26 17:33:15.797: PIM(2)[green]:

Send v2 Register-Stop to 10.1.101.1 for 10.1.101.11, group 226.1.1.1 <-- Sending Register-stop to FHR


*May 26 17:33:15.797: PIM(2)[green]:

Received v2 Register-Stop on Vlan101 from 10.2.255.255             <-- Leaf-02 receives its own Regis


*May 26 17:33:15.797: PIM(2)[green]:

for source 10.1.101.11, group 226.1.1.1                              <-- S,G the Stop is for


*May 26 17:33:15.797: PIM(2)[green]:

Clear Registering flag to 10.2.255.255 for (10.1.101.11/32, 226.1.1.1) <-- Done with Register event


*May 26 17:33:17.801: PIM(2)[green]:

Received v2 Register on Vlan901 from 10.1.101.1                      <-- Another Register messages fr


*May 26 17:33:17.801: PIM(2)[green]: Send v2 Register-Stop to 10.1.101.1 for 10.1.101.11, group 226.1.1
*May 26 17:33:17.802: PIM(2)[green]: Received v2 Register-Stop on Vlan101 from 10.2.255.255
*May 26 17:33:17.802: PIM(2)[green]: for source 10.1.101.11, group 226.1.1.1
*May 26 17:33:17.802: PIM(2)[green]: Clear Registering flag to 10.2.255.255 for (10.1.101.11/32, 226.1.
```

PIM Register Tunnel Route Issue Solution

The **solution** is to use a unique Loopback IP on all VTEPs and use the configuration noted in this section.

```
<#root>

Leaf-01#

sh run int lo 901


interface Loopback901

vrf forwarding green <-- Loopback is in the Tenant VRF
```

```
ip address 10.1.255.1
 255.255.255.255
<-- IP is unique to the VTEP

ip pim sparse-mode
Leaf-02(config)#
ip pim vrf green register-source loopback 901 <-- force the Register Source to use the Loopback


Leaf-01#
sh ip pim vrf green tunnel

Tunnel5
Type : PIM Encap     <-- Register Encapsulation tunnel


RP : 10.2.255.255    <-- RP IP is the Tunnel destination


Source : 10.1.255.1  <-- Loopback 901 is the Tunnel source

State : UP
Last event : Created (02:45:58)
Leaf-02#
show bgp l2vpn evpn all | beg 10.1.255.1

*>i
[5]
[1:1][0][32]
[10.1.255.1]
/17
          172.16.254.3
       0          100    0 ?
<-- Only one entry and next hop


to Leaf-01
```

Related Information

EVPN VxLAN TRM Configuration Guide

EVPN VxLAN Unicast Troubleshooting

MVPN Configuration Guide 17.3.x (Catalyst 9300 Switches)

MVPN Configuration Guide 17.3.x (Catalyst 9500 Switches)

BGP configuration guide