



You make **possible**



Multicast in ACI

John Weston, Technical Marketing Engineer
Intent Based Networking Group

BRKACI-2608

CISCO *Live!*

Barcelona | January 27-31, 2020



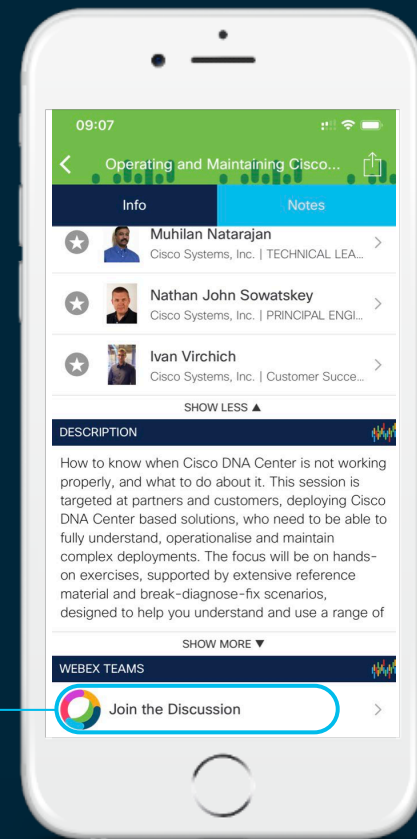
Cisco Webex Teams

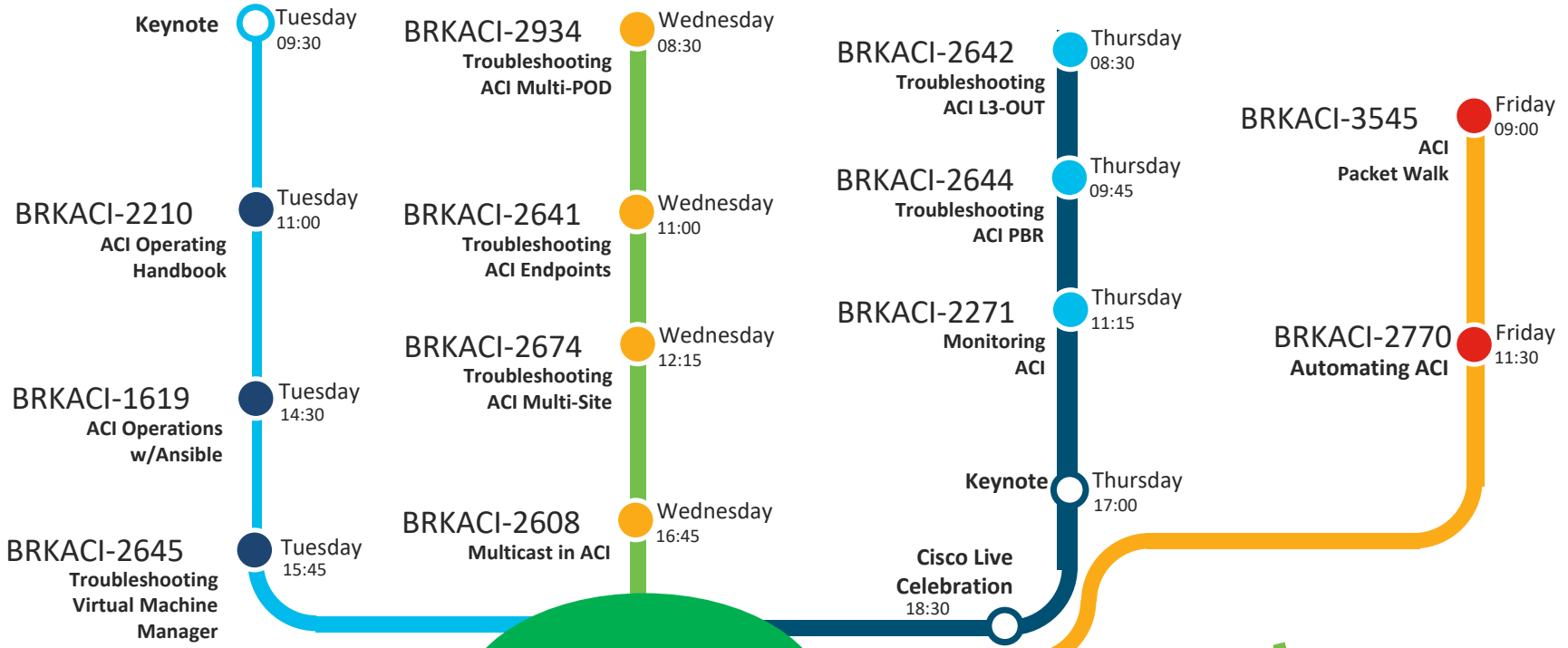
Questions?

Use Cisco Webex Teams to chat with the speaker after the session

How

- 1 Find this session in the Cisco Events Mobile App
- 2 Click “Join the Discussion”
- 3 Install Webex Teams or go directly to the team space
- 4 Enter messages/questions in the team space





CISCO *Live!*

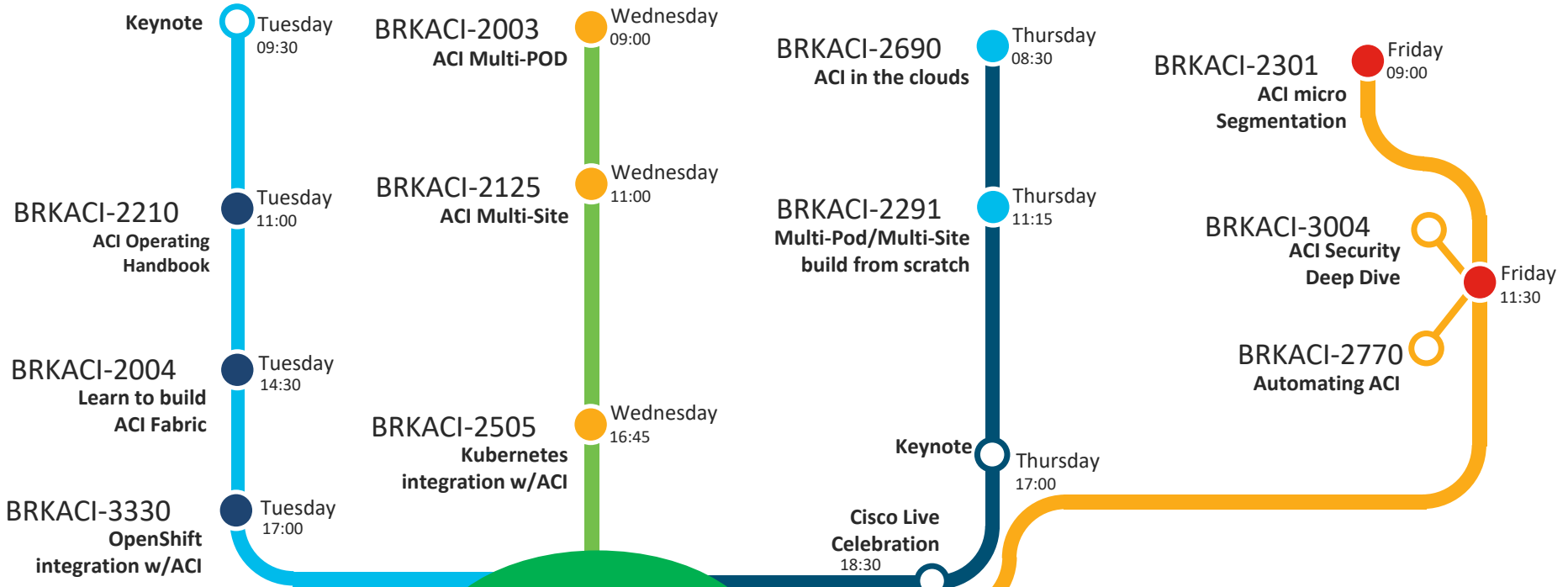
ACI Operations



CISCO *Live!*



Technology



Session Objectives



At the end of the session, the participants should be able to:

- Understand how tenant multicast traffic is forwarded in the ACI fabric

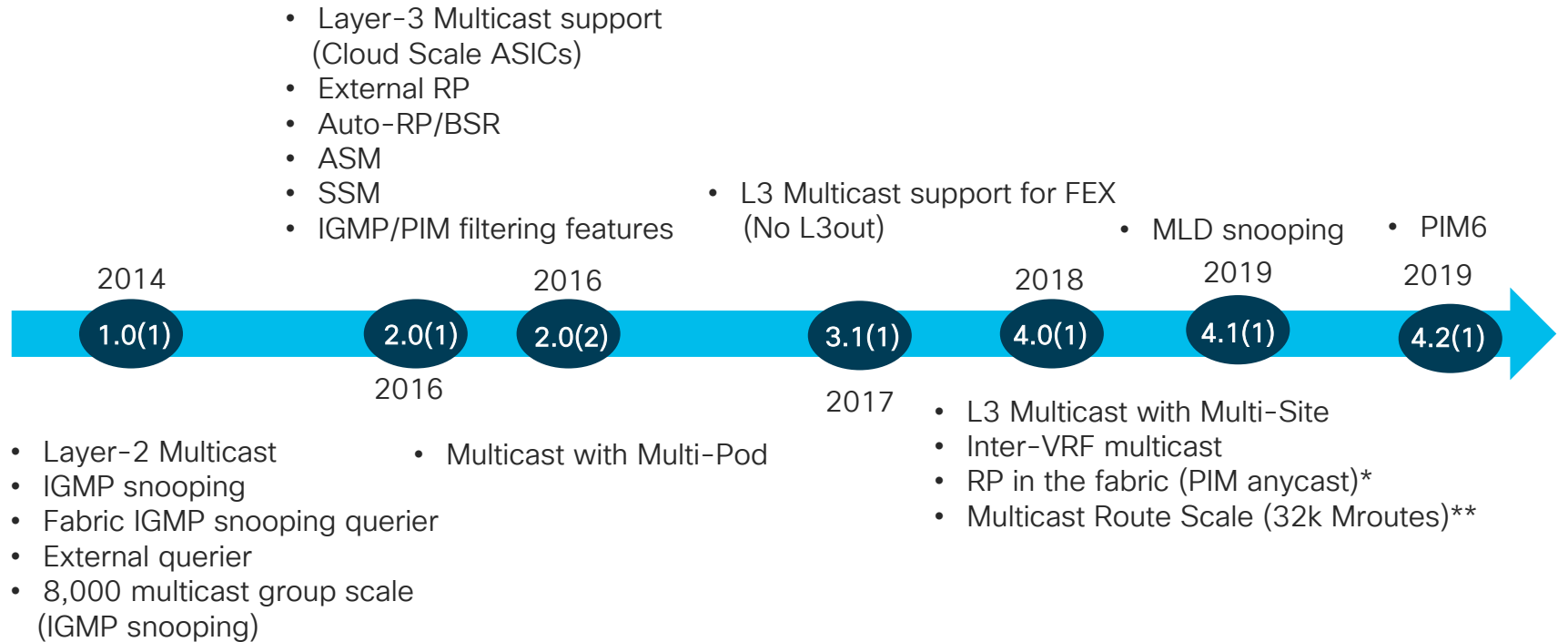
Initial assumption:

- The audience already has good knowledge of ACI main concepts (Tenant, BD, EPG, L3Out, Multi-Pod, Multi-Site, etc.)
- The audience already has a good understanding of multicast routing (IGMP, PIM)

Agenda

- Introduction
- Multicast Data Plane
- Multicast Control Plane
- Layer-2 IP Multicast
- Layer-3 Multicast
- Multicast Configuration
- Multicast in a Multi-Pod Fabric
- Multicast in a Multi-Site Architecture
- Inter-VRF Multicast
- Multicast Troubleshooting

ACI Multicast Features Timeline

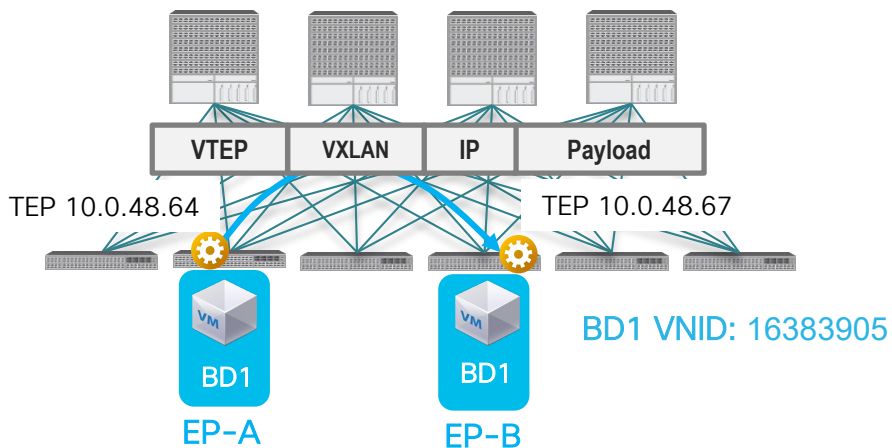


*RP in the fabric not supported with Multi-Site
 **FX only

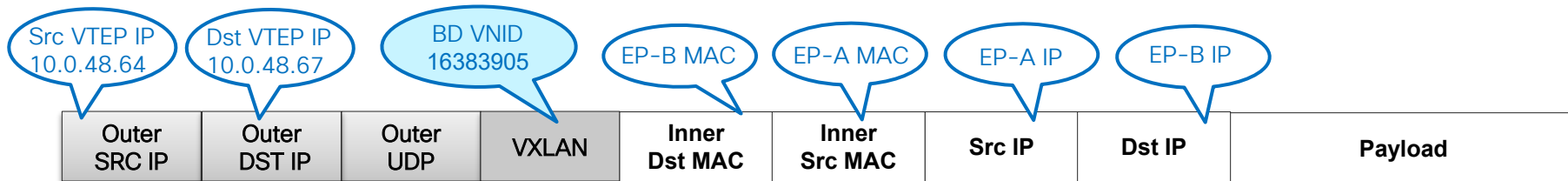
ACI Multicast Data Plane

Layer-2 Unicast VXLAN packet

⚙️ = VTEP (VXLAN Tunnel Endpoint)



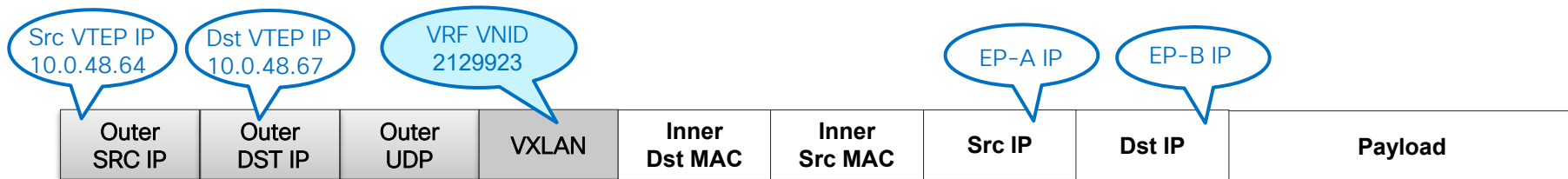
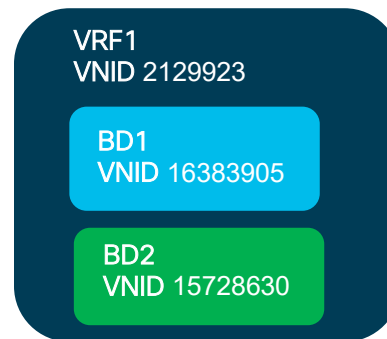
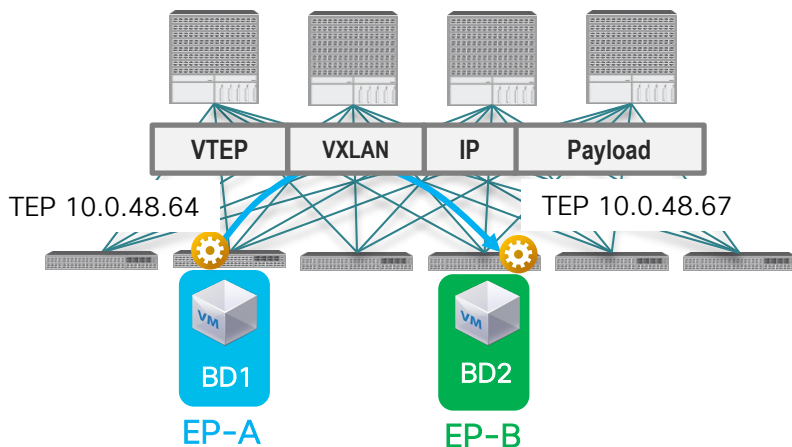
- ACI is a VXLAN fabric
- Unicast packets sent across the fabric will be encapsulated in a unicast VXLAN packet
- The outer source and destination IP addresses are VXLAN tunnel endpoints
- The inner packet carries tenant traffic



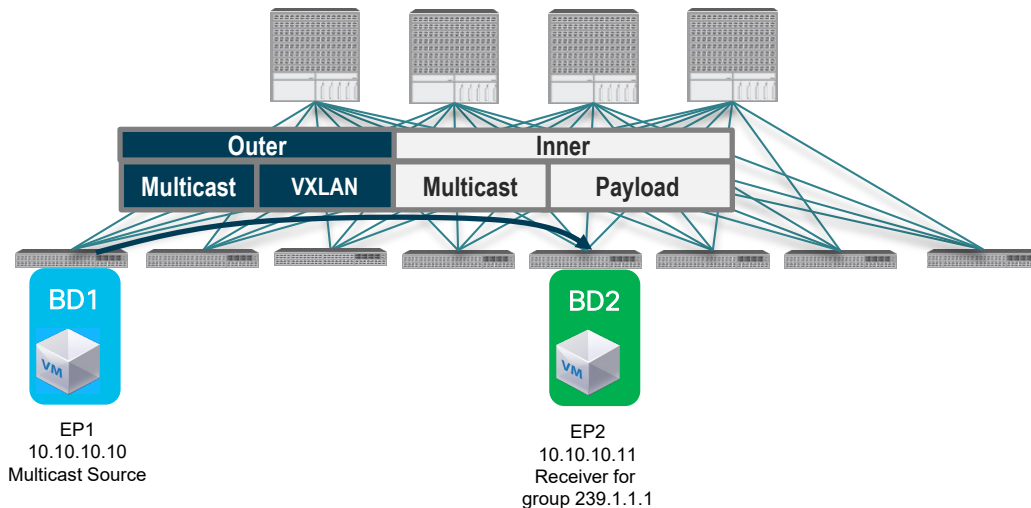
Review

Layer-3 Unicast VXLAN packet

⚙️ = VTEP (VXLAN Tunnel Endpoint)



Multicast VXLAN encapsulation

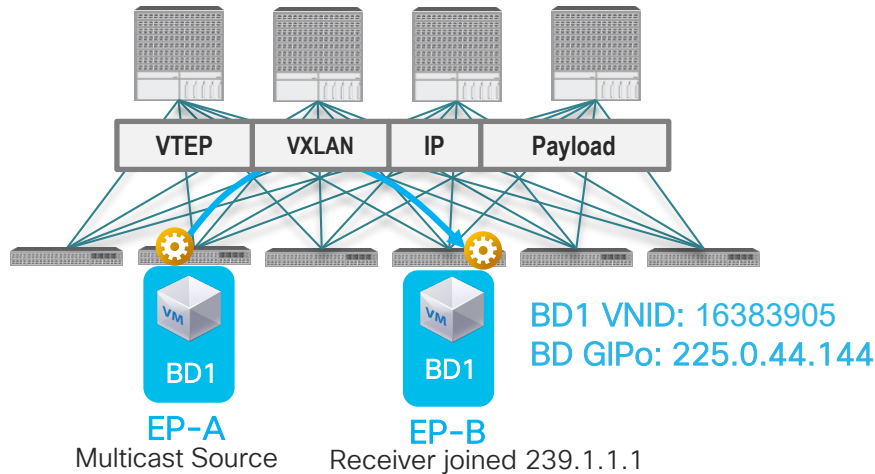


- Multicast traffic is also encapsulated in VXLAN across the fabric
- Tenant multicast traffic (inner packet) is encapsulated in an outer VXLAN multicast packet

Note: Also used for other multi-destination BUM traffic (Broadcast, unknown unicast, multicast)

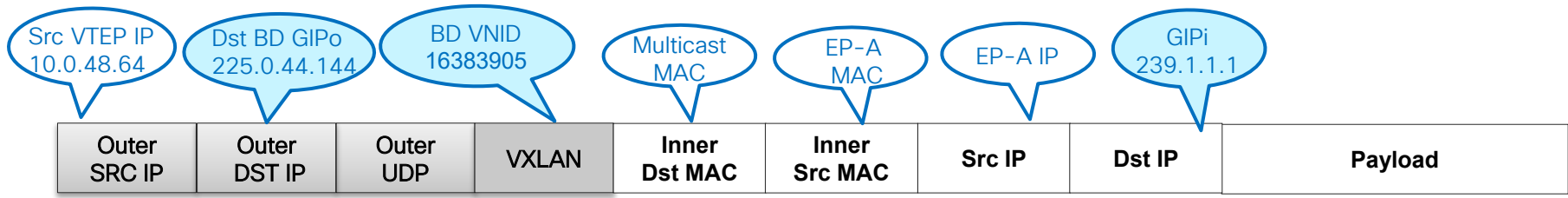
Layer-2 IP Multicast VXLAN packet

⚙️ = VTEP (VXLAN Tunnel Endpoint)

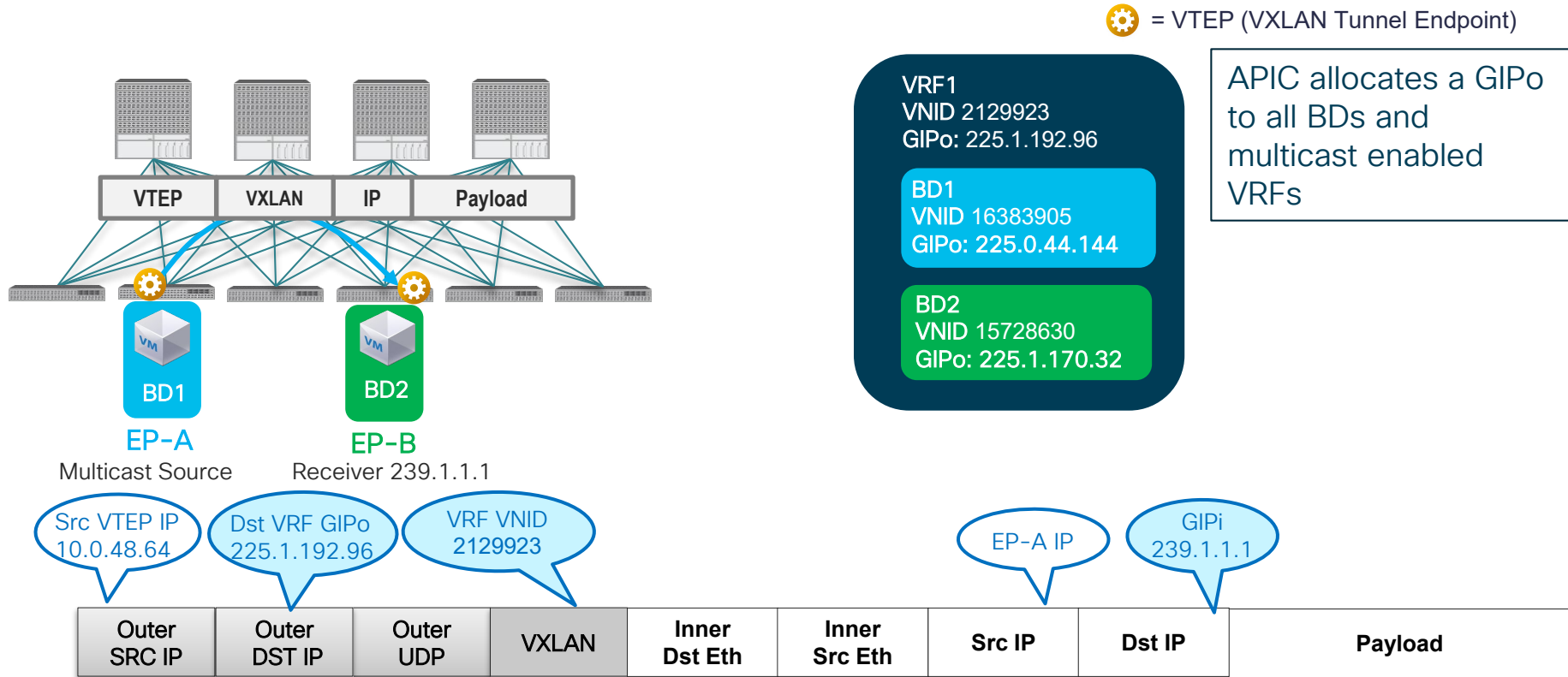


ACI Multicast terminology

- **GIPi:** Group IP Inner address. Multicast address in the inner VXLAN packet. Tenant multicast traffic running in the tenant VRF
- **GIPo:** Group IP Outer address. Multicast address in the outer VXLAN packet. This is the multicast address used for distributing multicast traffic across the fabric



Layer-3 Multicast VXLAN packet



ACI Unicast and Multicast VXLAN Summary

Single pod and Multi-Pod fabrics*



For Your Reference

Tenant Layer-2 Unicast Packet

SRC VTEP	DST VTEP	Outer UDP	BD VNID	Dst MAC	Src MAC	Src IP	Dst IP	Payload
-----------------	-----------------	------------------	----------------	----------------	----------------	---------------	---------------	----------------

Tenant Layer-3 Unicast Packet

SRC VTEP	DST VTEP	Outer UDP	VRF VNID	Fabric MAC	Fabric MAC	Src IP	Dst IP	Payload
-----------------	-----------------	------------------	-----------------	-------------------	-------------------	---------------	---------------	----------------

Tenant Layer-2 Multicast Packet

SRC VTEP	BD GIPo	Outer UDP	BD VNID	Mcast MAC	Src MAC	Src IP	GIPi	Payload
-----------------	----------------	------------------	----------------	------------------	----------------	---------------	-------------	----------------

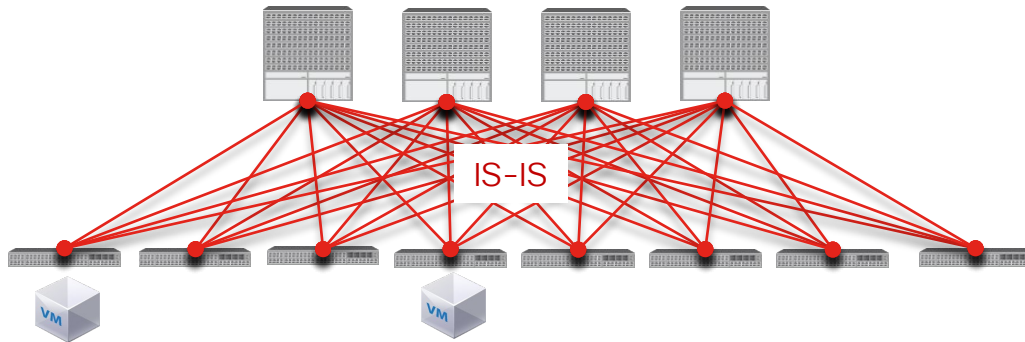
Tenant Layer-3 Multicast Packet

SRC VTEP	VRF GIPo	Outer UDP	VRF VNID	Fabric MAC	Fabric MAC	Src IP	GIPi	Payload
-----------------	-----------------	------------------	-----------------	-------------------	-------------------	---------------	-------------	----------------

*Multisite, remote leaf, and vPod use Head end replication to send multicast

Multicast Control Plane

Multicast Underlay Control Plane

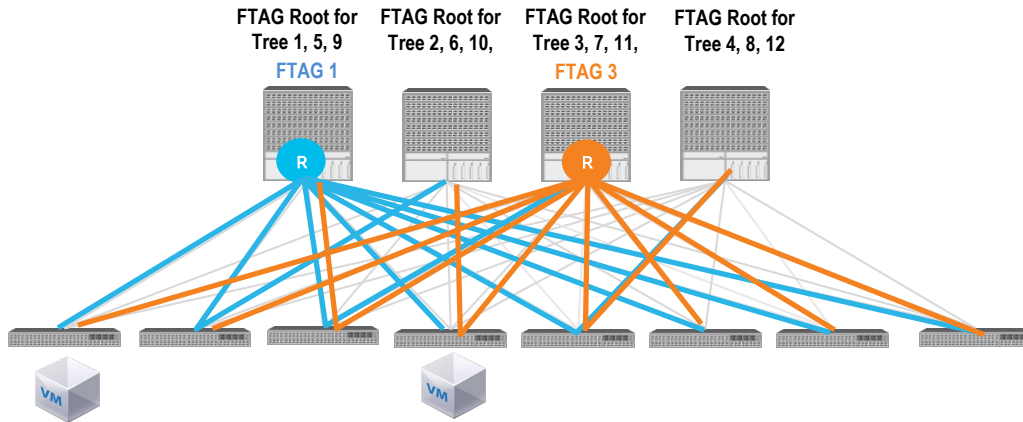


Underlay Control Plane

- ACI uses the GIPo range 225.0.0.0/15 configured during fabric setup by default (configurable)
- The underlay multicast groups are separate from tenant multicast groups
- APIC assigns GIPo addresses to BDs and VRFs
- PIM is not used in the underlay
- GIPo groups are advertised using IS-IS GM-LSPs

Multicast Underlay Forwarding Plane

FTAGs (Forwarding Tags)

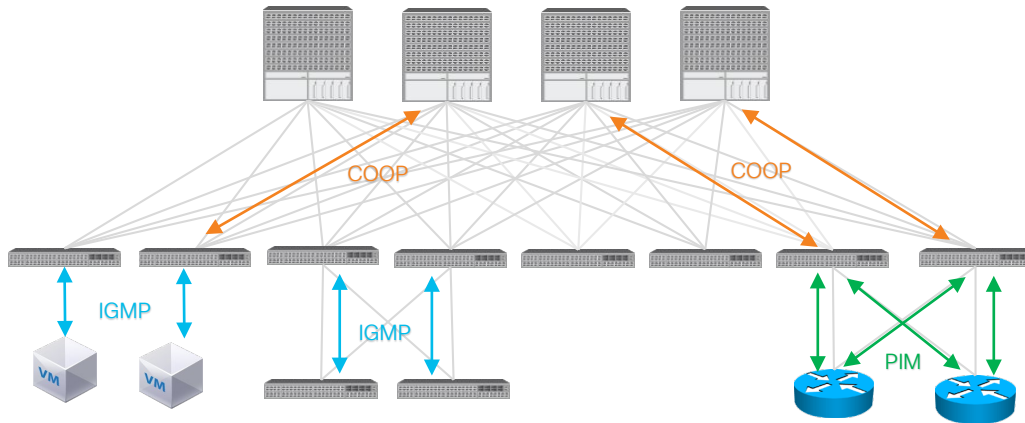


Underlay Control Plane

- IS-IS is used to build loop free distribution trees called FTAG trees.
- All FTAG trees extend to all nodes (loop free topology)
- Multiple trees achieve load balancing across the fabric for multicast traffic
- There are 16 FTAG trees but only 12 trees are used for user traffic

Multicast Overlay Control Plane

IGMP/PIM/COOP



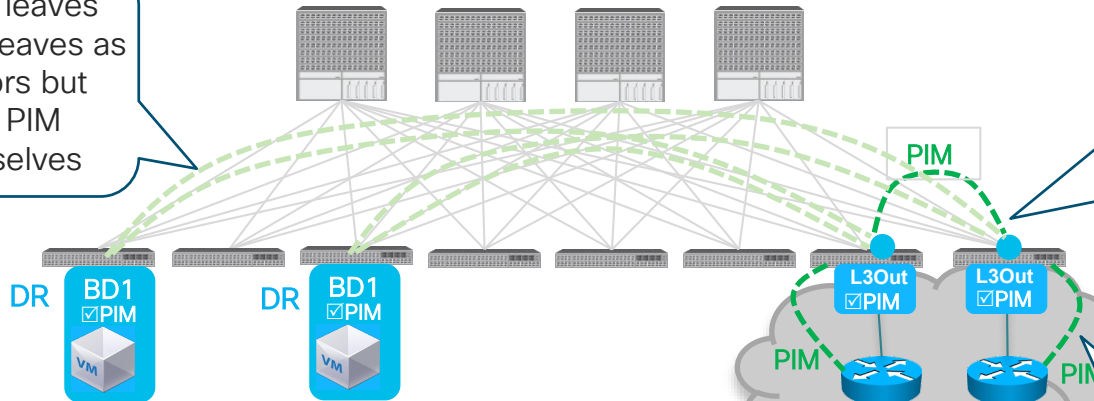
Overlay Control Plane

- IGMP is used by hosts to join multicast group (version 2, version 3 supported)
- COOP is used within the fabric for advertising multicast interest to spines and border leaves
- PIM is used on L3Out connections and Bridge Domain SVIs

Multicast Overlay Control Plane

PIM Modes

- Non-border leaves see border leaves as PIM neighbors but do not send PIM hellos themselves



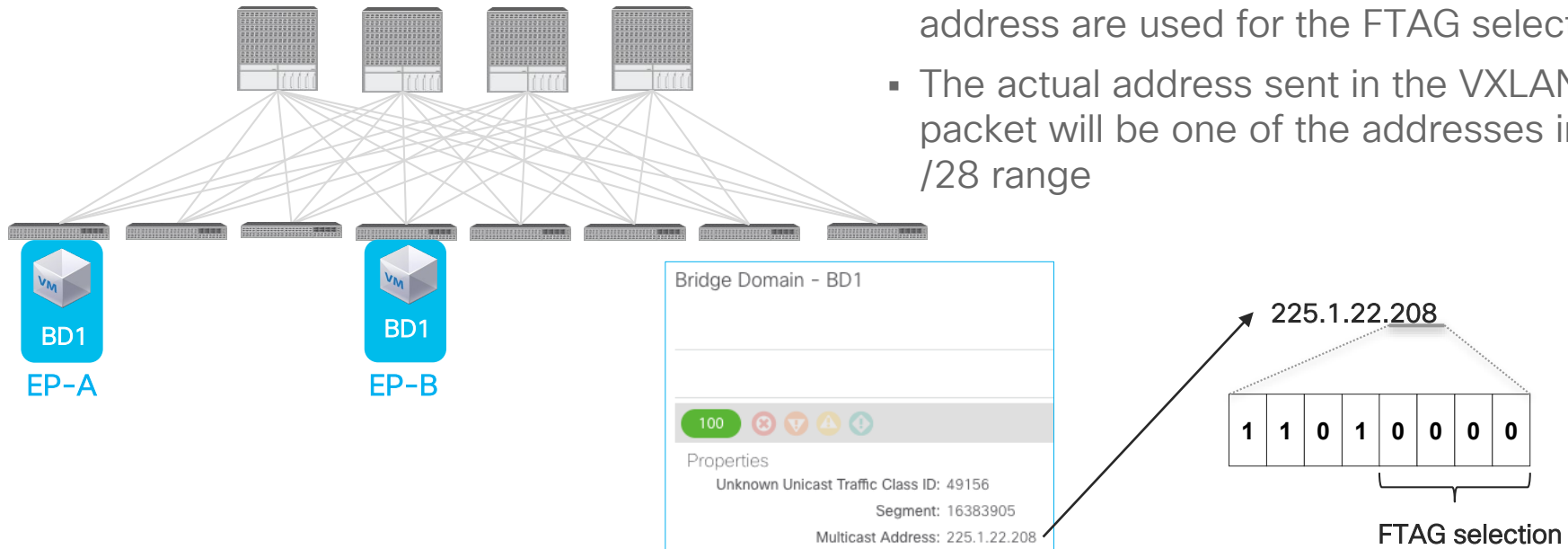
- PIM enabled L3Outs must also have loopbacks configured
- Border leaves form PIM neighbors with other border leaves across fabric

- Non-border leaves run PIM on bridge domain SVIs
- PIM on non-border leaves runs in passive mode and will not peer with any external devices

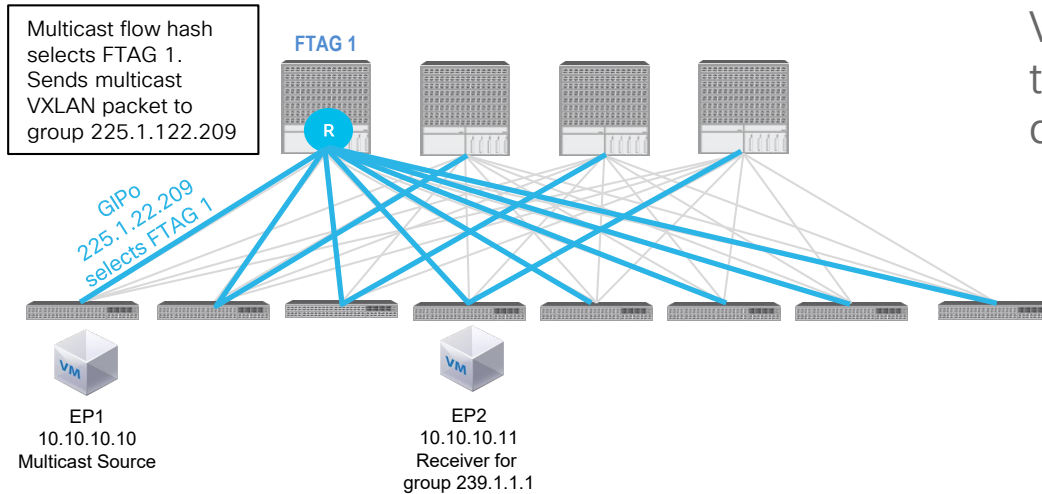
- Border leaves run full PIM protocol and peer with external PIM routers
- PIM enabled L3Outs are required for L3 multicast even if source and receivers are inside the fabric

GIPO and FTAG selection

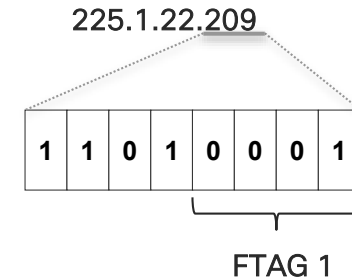
- A GIPO address is allocated to BDs and multicast enabled VRFs.
- The GIPO address displayed in the GUI is a /28 address. The last four bits of this address are used for the FTAG selection
- The actual address sent in the VXLAN packet will be one of the addresses in the /28 range



GIPO and FTAG selection

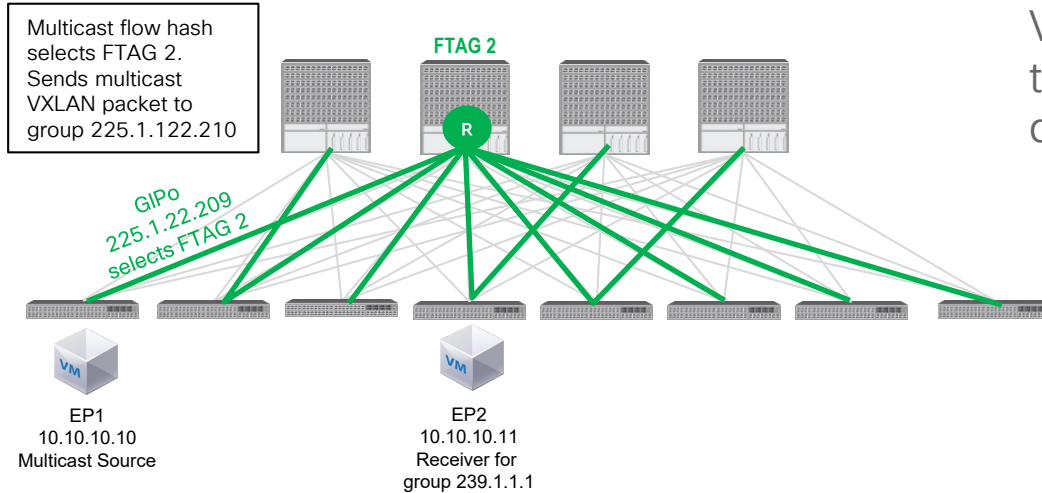


- When the source leaf sends a multicast VXLAN packet it selects an FTAG using the last four bits in the last octet of the destination multicast address (GIPO)

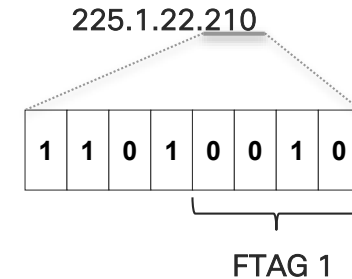


GIPO: 225.1.22.208
FTAG range: 225.1.22.209-220

GIPo and FTAG selection



- When the source leaf sends a multicast VXLAN packet it selects an FTAG using the last four bits in the last octet of the destination multicast address (GIPo)



GIPo: 225.1.22.208
FTAG range: 225.1.22.209-220

Multicast GIPo Usage



- Multicast is also used in the overlay for other multi-destination traffic. Non-IP multicast, Spanning tree, Broadcast, Unknown Unicast. The BD GIPo is used for other multi-destination traffic

	Non-Multicast Routing Enabled BD	Multicast Routing Enabled BD
Broadcast	BD GIPo	BD GIPo
Unknown Unicast Flood	BD GIPo	BD GIPo
Multicast	BD GIPo	VRF GIPo

Layer-2 IP Multicast

Layer-2 Multicast

- In this section layer-2 multicast refers to IP multicast packets forwarded on a layer-2 network segment (BD/subnet)
- It is not Layer-2 non-IP multicast packets. (multicast packets with a destination multicast MAC address without an IP header)
- Also excludes link local multicast (224.0.0.0/24). Link local multicast is always forwarded to all ports in the BD

Layer-2 IP multicast forwarding Overview

Supported from release 1.0

Layer-2 multicast is only forwarded within the bridge domain

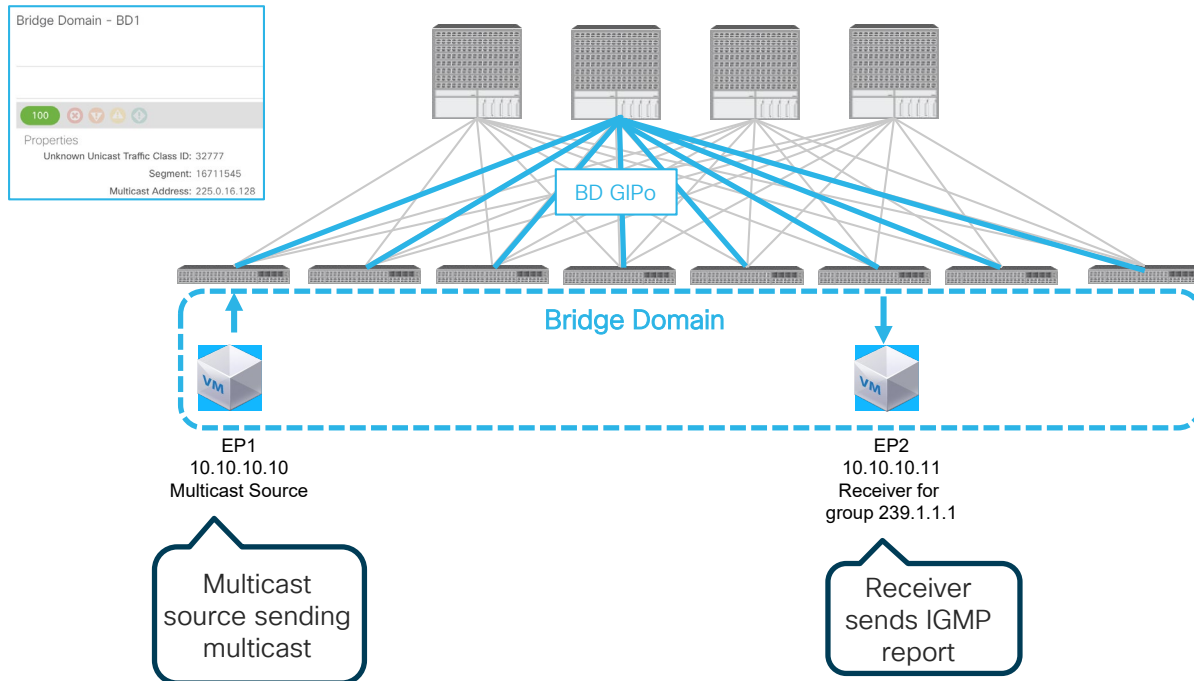
Forwarded using the BD GIPo address

Layer-2 multicast is supported on all generation leaf and spine switches

Use case: Layer-2 multicast in bridge domain

Multicast sources and receivers in the same subnet (bridge domain)

No L3 multicast requirement



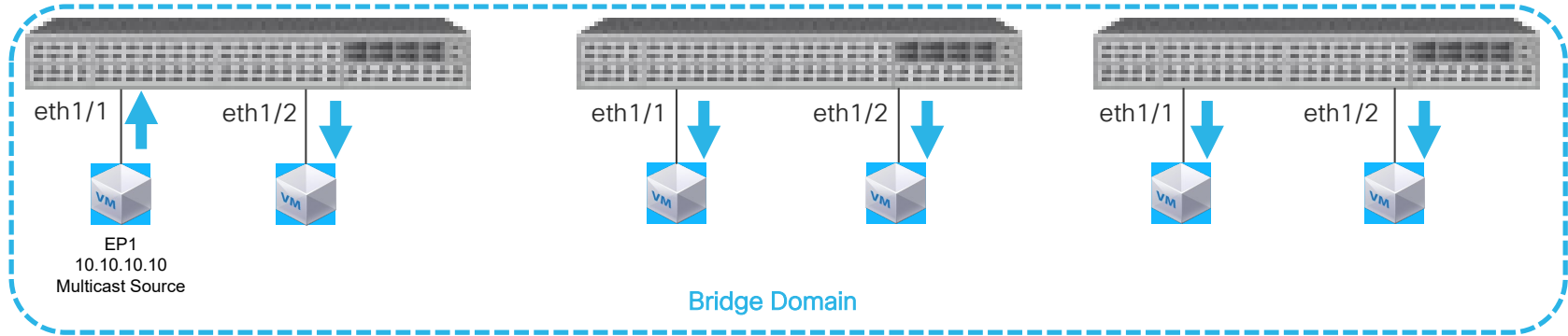
- Multicast will be forwarded to all leaf switches where the BD is deployed using the BD GIPo
- Each Leaf switch will receive a copy of the multicast stream and forward multicast out front panel ports based on the BD **Unknown Multicast** configuration
 - Flood
 - Optimized Multicast Flood

BD Flood Modes

Multicast sources and receivers in the same bridge domain

Unknown Multicast: Flood

Multicast is “unknown” when there are no IGMP snooping entries for the group (per leaf)



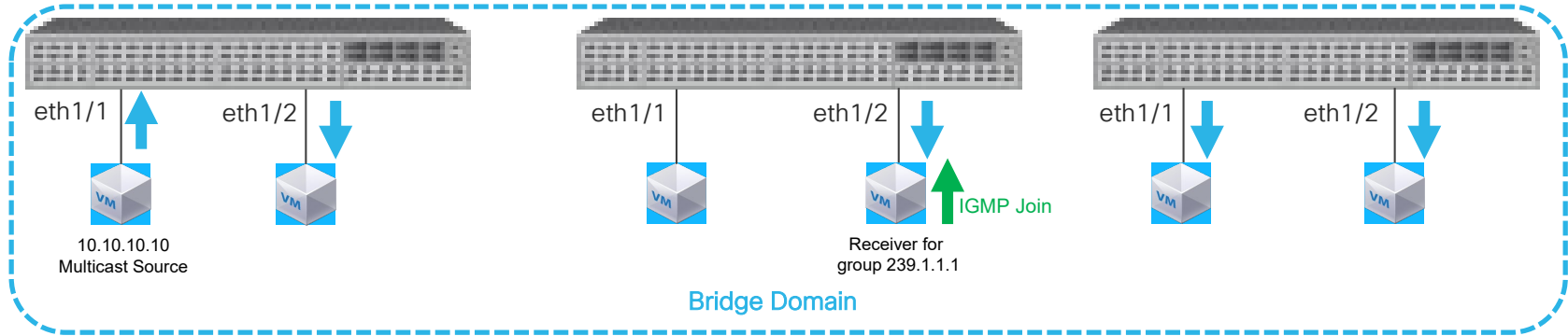
Unknown multicast is forwarded out all ports in the BD

BD Flood Modes

Multicast sources and receivers in the same bridge domain

Unknown Multicast: Flood

If a group is known (IGMP snooping table entry) streams for that group will be forwarded based on the IGMP snooping table (per leaf)



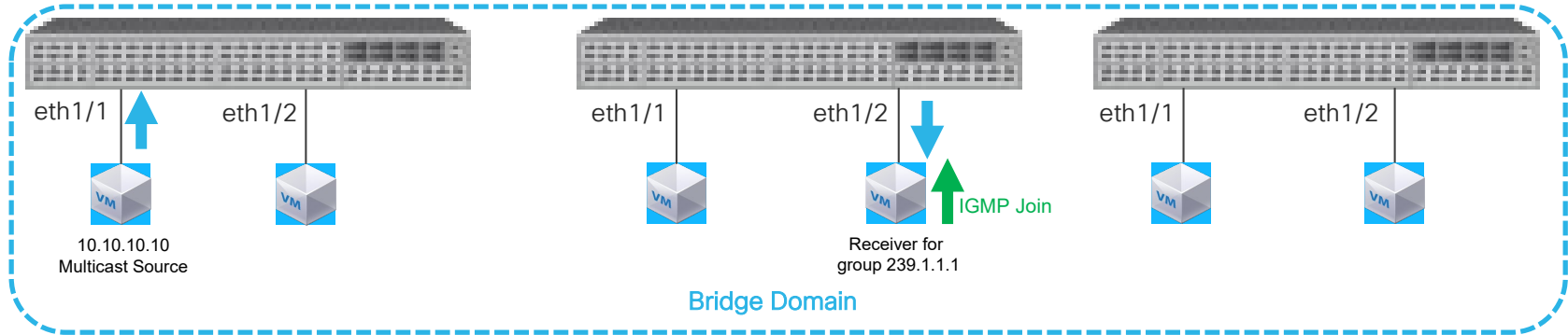
Known multicast forwarded based on IGMP snooping table (per leaf)

BD Flood Modes

Multicast sources and receivers in the same subnet (bridge domain)

Unknown Multicast: Optimized Flood

Optimized flood forwards only based on IGMP snooping table across all leaf switches where BD is deployed



Unknown multicast is forwarded based on IGMP snooping table

Which BD mode to use?

Bridge Domains without PIM enabled



- **Optimized Flood** is similar to IGMP snooping where multicast is only sent to interested receivers.
 - Requires an IGMP snooping querier to trigger report messages from hosts
- **Flood** will forward multicast out all ports in the BD. Receivers will receive multicast in the absence of a querier but can consume bandwidth and result in unnecessary flooding of multicast traffic

IGMP Snooping Querier



For Your Reference

- 1 Configure an IGMP snooping policy for the BD. Enable the querier option
- 2 Enable querier under the BD subnet

Create IGMP Snoop Policy

Name:

Description:

Admin State: Disabled Enabled

Control: Fast leave Enable querier

Query Interval (sec):

Query Response Interval (sec):

Last Member Query Interval (sec):

Start Query Count:

Start Query Interval (sec):

Subnet - 192.168.1.1/24

IP Address: 192.168.1.1/24

Description:

Treat as virtual IP address:

Make this IP address primary:

Scope: Private to VRF Advertised Externally Shared between VRFs

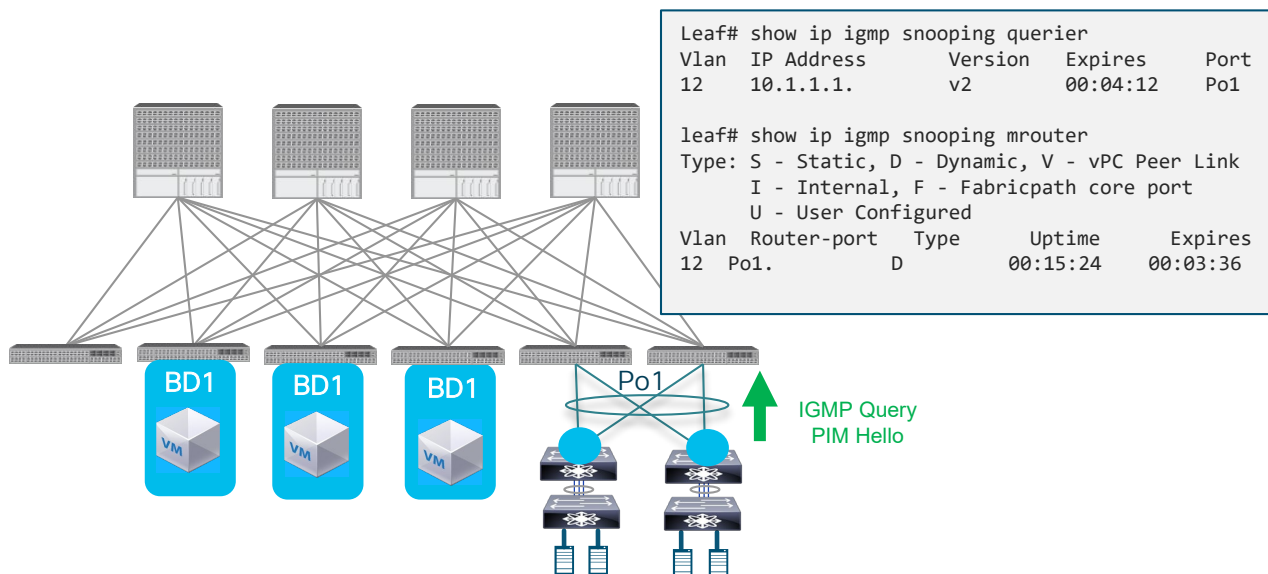
Subnet Control: No Default SVI Gateway Querier IP

L3 Out for Route Profile:

Route Profile:

Use case: Layer-2 multicast with external querier

Bridge domain is Layer-2 only with gateway outside of fabric



- Fabric supports external querier
- A port that receives an IGMP query or PIM hello will become an IGMP snooping mrouter port
- The external queries will be forwarded out all BD ports
- IGMP reports and all multicast traffic will be sent out mrouter ports

Layer-3 Multicast

Layer-3 Multicast

Scale, SW/HW Support

HW

Supported spines

- All supported spines

Supported Leaf

- All 2nd Gen (Cloud Scale)
- -EX
- -FX
- -GX

Scale

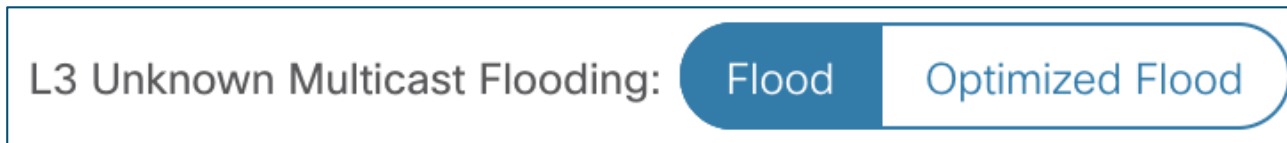
- Supports up to 8,000 Mroutes (fabric wide and per leaf) Default profile
- Supports up to 32,000 Mroutes (FX only) High Dual Stack profile

Layer-3 Multicast Overview

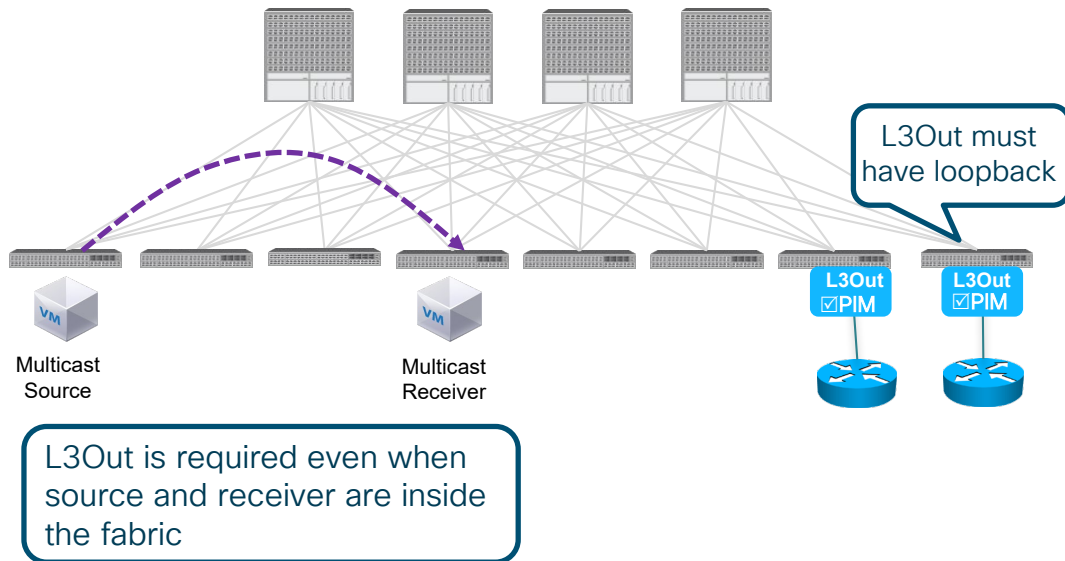
- Supported Modes
 - PIM-SM (ASM)
 - PIM-SSM
- External Rendezvous Point (RP)
 - Static
 - Auto-RP
 - BSR
- Fabric RP (Anycast RP) 4.0
 - No Auto-RP or BSR
 - Not supported with Multi-Site (coming soon)
- Inter-VRF multicast 4.0
- Multi-Site support 4.0
- PIM and IGMP filters
- Multicast traffic not filtered by contracts
- Not supported with remote leaf

Layer-3 Multicast BD Flood/OMF and IGMP snooping

- IGMP snooping must be enabled for PIM enabled BDs
- Multicast packets will only be forwarded out BD ports if there is an IGMP snooping entry
- It is not required to change the L3 Unknown Multicast Flooding setting (this differs from non-PIM enabled BDs)

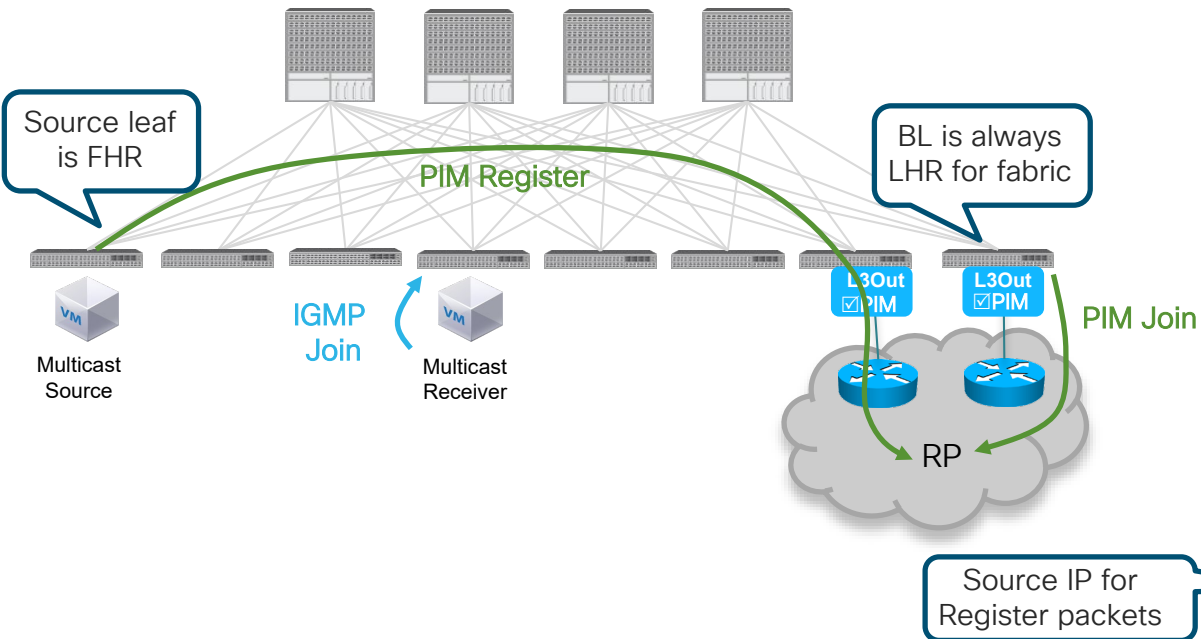


Layer-3 Multicast: Role of L3Outs



- L3Outs are a requirement for multicast routing
- The L3Out must be configured with a loopback address and have PIM enabled
 - PIM is supported on L3Outs with routed or routed sub-interfaces and includes Layer-3 port-channels
 - Not supported on L3Outs with SVI interfaces
- For ASM the L3Out provides a path to the external RP or can function as the RP (4.0 feature)

Layer-3 Multicast: External RP



External RP Options

- Static RP (supports multiple RPs)
- Auto-RP (MA Filters)
- BSR (BSR Filters)

Any Source Multicast (ASM)

Shared Range Policy

RouteMap:

Source, Group(S,G) Expiry Policy

RouteMap:

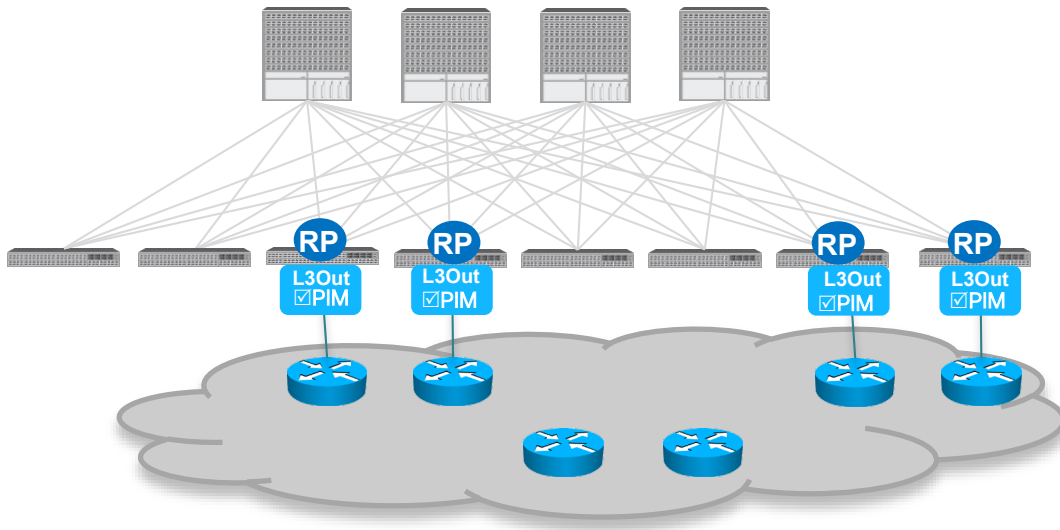
Expiry (seconds):

Register Traffic Policy

Max Rate (packets per second):

Source IP:

Layer-3 Multicast: Fabric RP

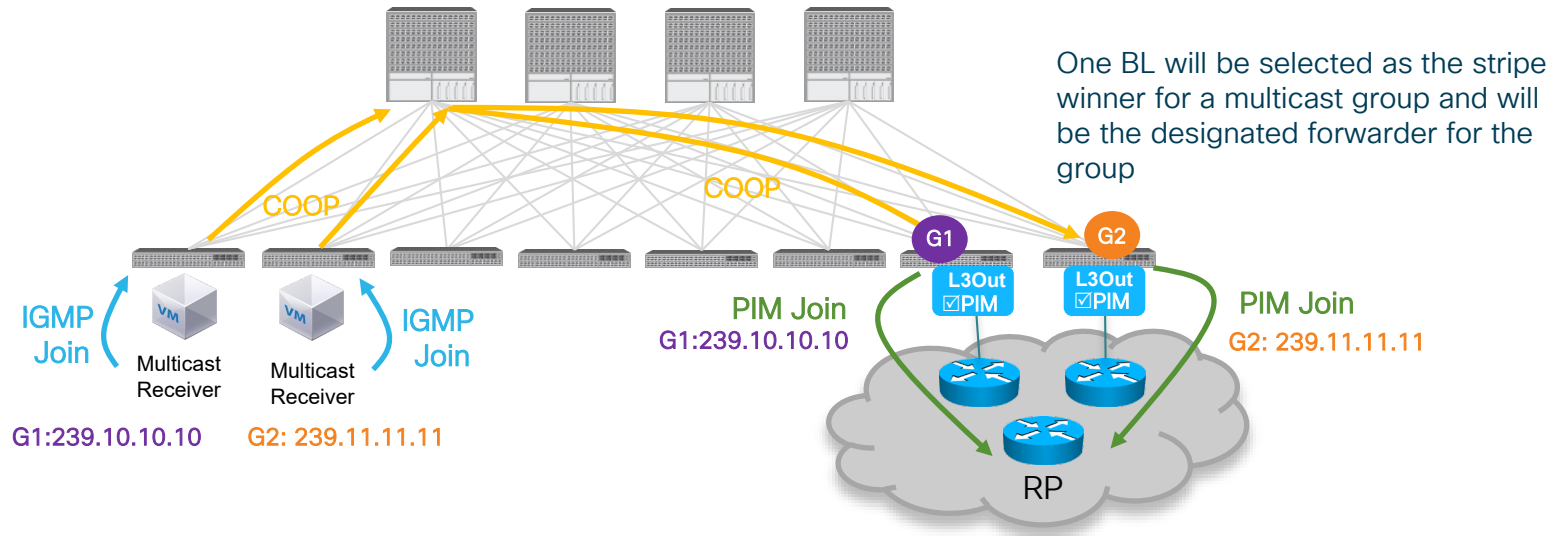


Fabric RP

- Anycast RP (equivalent to PIM anycast, does not run MSDP)
- Does not support peering anycast RP with external RP
- All PIM enabled border leaves will become anycast RP members (per VRF)
- Required for inter-VRF multicast

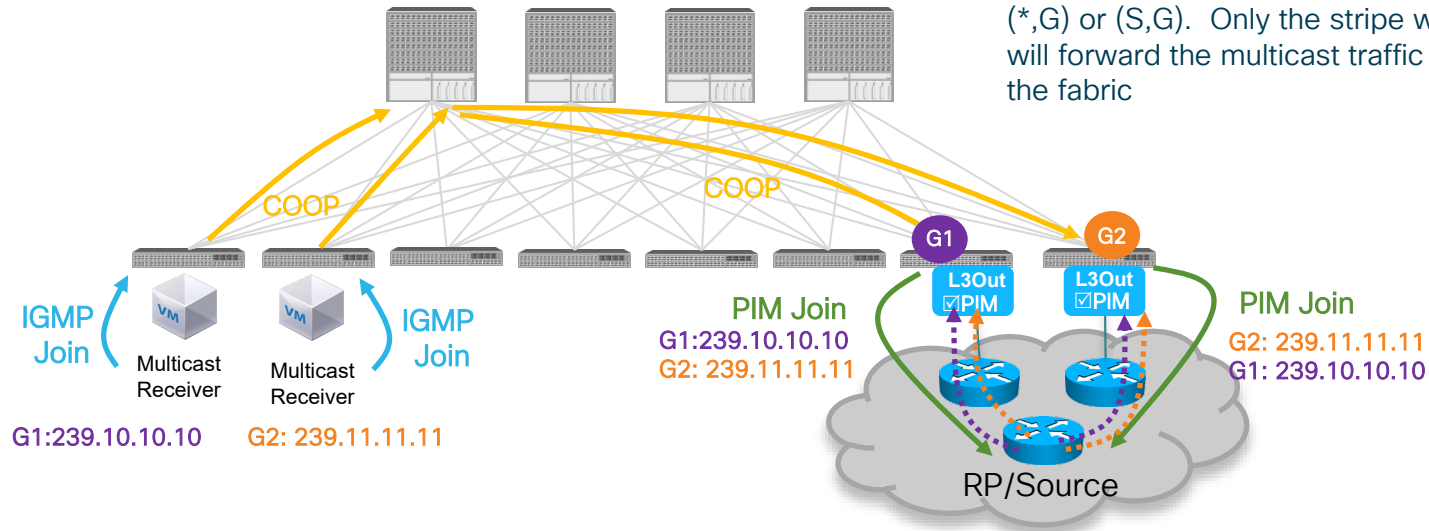
Layer-3 Multicast: Stripe Winners

The COOP database maintains multicast group interest in the fabric and is distributed to all active border leaves in the fabric



Layer-3 Multicast: Fast Convergence

Fast Convergence is a feature where all border leaves (stripe winners and non stripe winners) will send joins for (*,G) or (S,G). Only the stripe winner will forward the multicast traffic into the fabric



External network has multicast state.
Improves convergence time after a failure

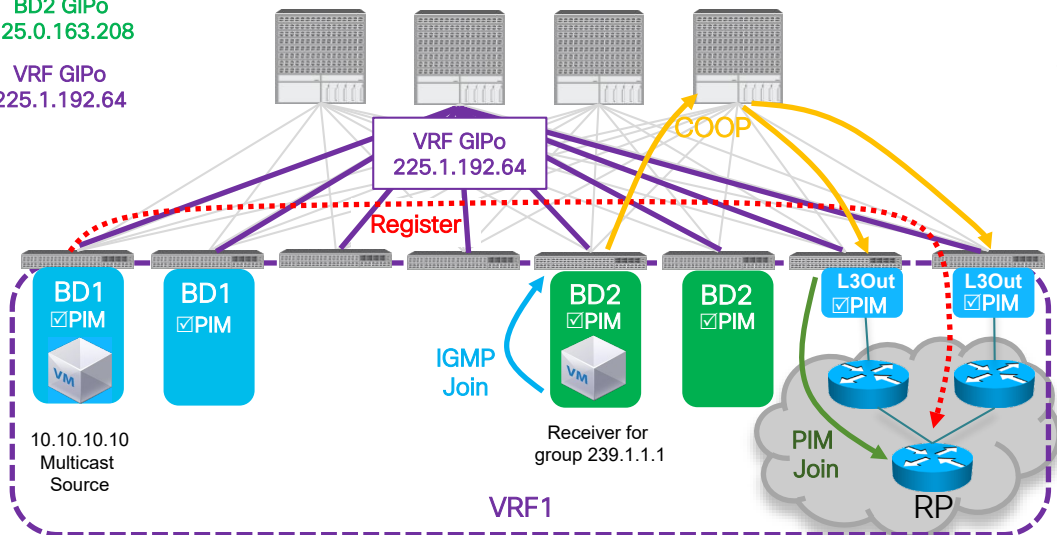
Use Case: Layer-3 Multicast

Source and Receiver in the Fabric (ASM)

BD1 GIPo
225.1.22.208

BD2 GIPo
225.0.163.208

VRF GIPo
225.1.192.64

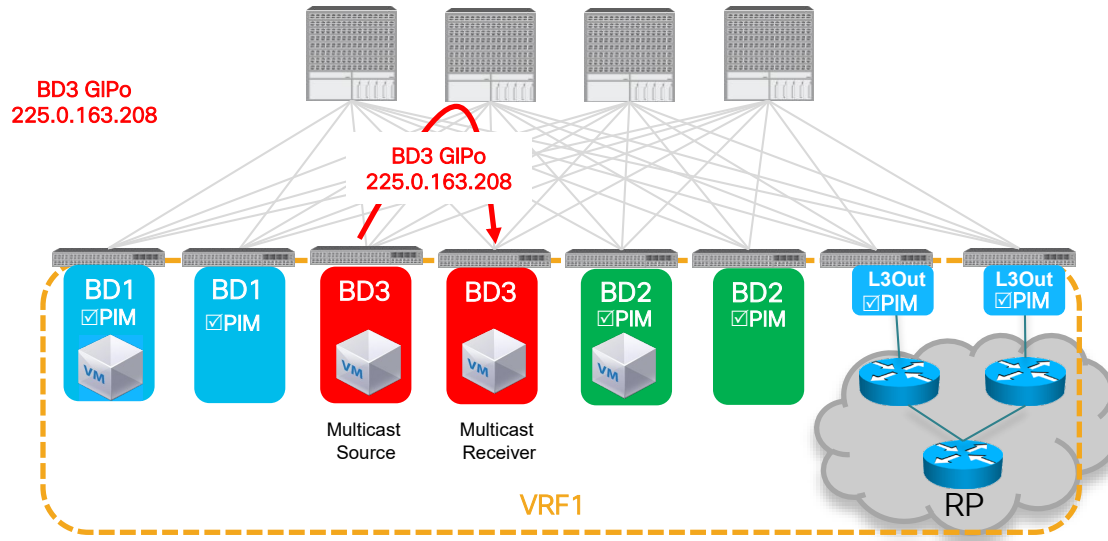


ACI switches do not continuously send register packets to the RP. Only a single packet is sent to the RP. The leaf will send periodic null registers to the RP.

The border leaf seeing that the source is local to the fabric (RPF is fabric interface) will send a PIM prune towards the RP

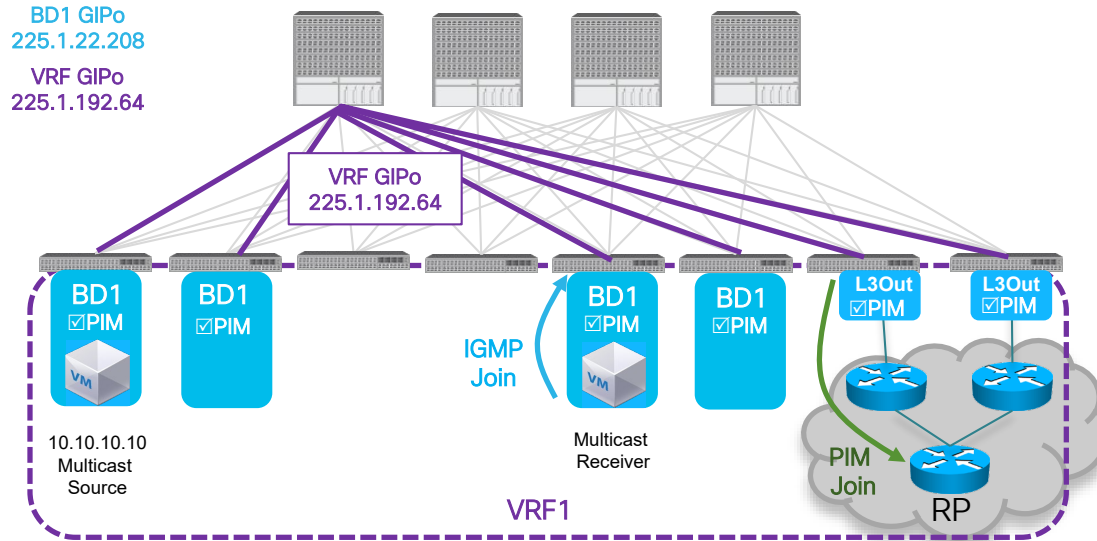
BDs without PIM enabled in multicast enabled VRF

- Multicast in BD without PIM enabled
- IGMP Querier (fabric or external)
 - OMF mode

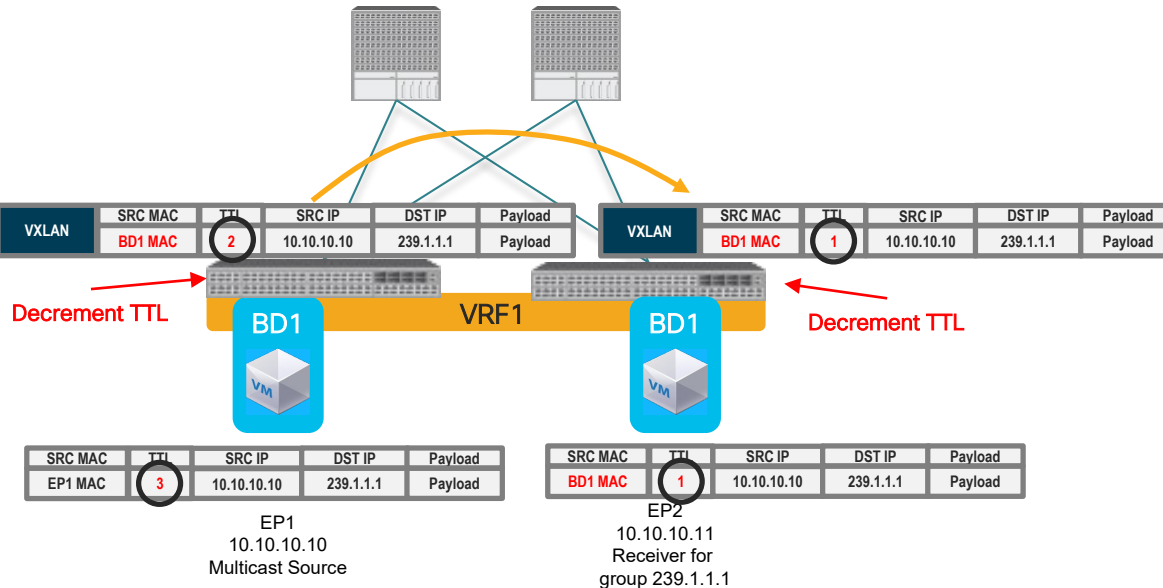


Use Case: Layer-2 Multicast with PIM enabled BDs

Source and Receiver in the Fabric Same BD



Layer-2 IP Multicast in PIM enabled Bridge Domains



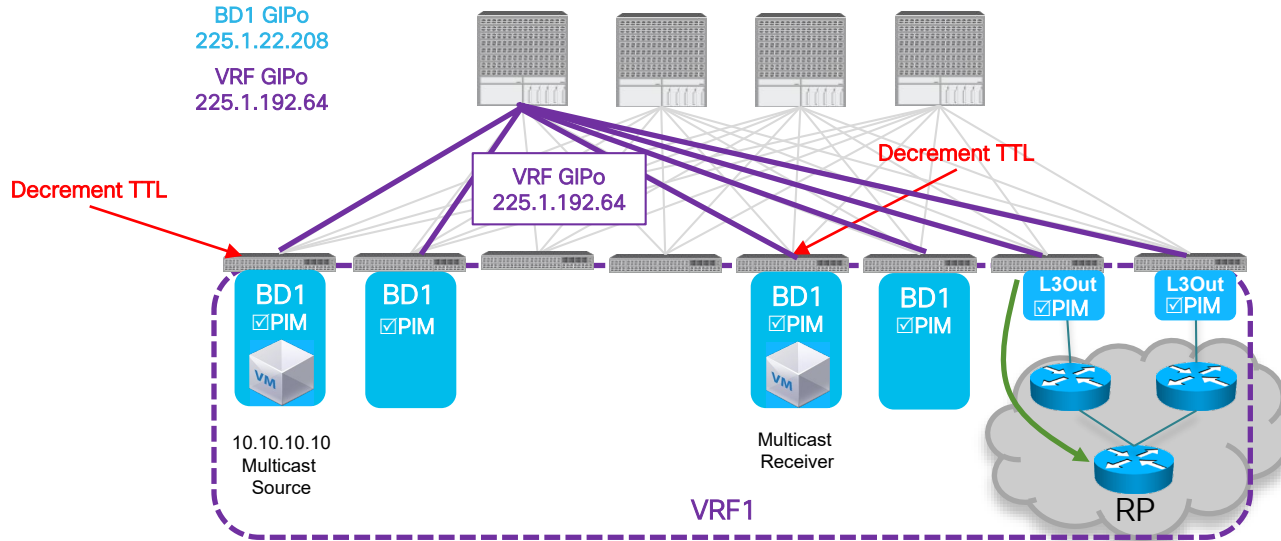
Layer-2 multicast forwarding behavior for PIM enabled BDs

- Routing first approach. All IP multicast is routed
- TTL will be decremented twice. Once on the ingress node and once on the egress node (regardless of number of transit nodes)
- RP must be defined for ASM
- Source MAC will be rewritten to BD MAC

*Excludes link local multicast (224.0.0/24)

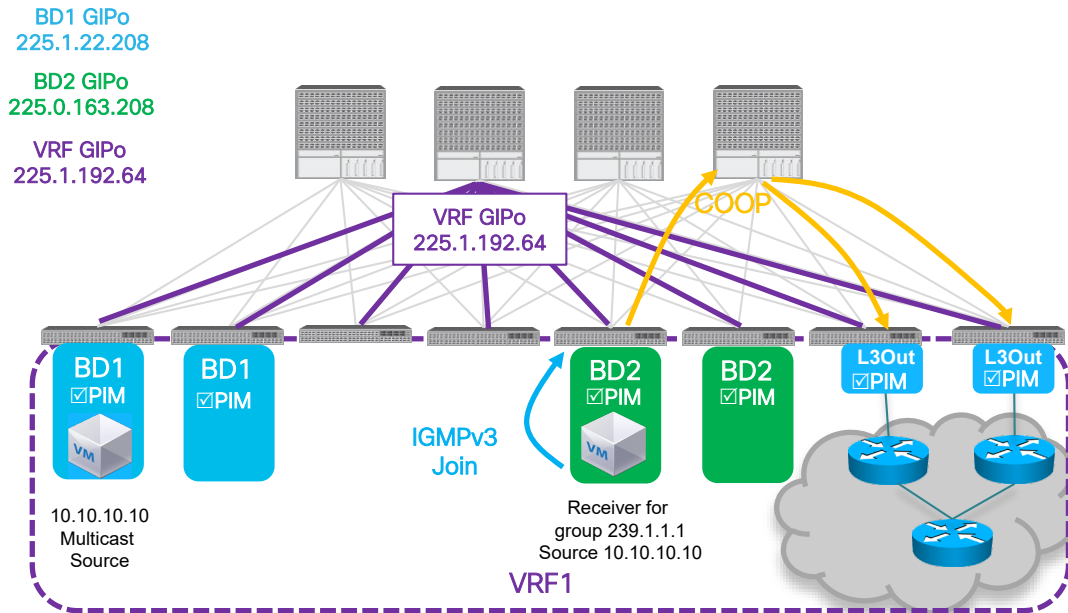
Use Case: Layer-3 Multicast

Source and Receiver in the Fabric (ASM)



Use Case: Layer-3 Multicast

PIM-SSM

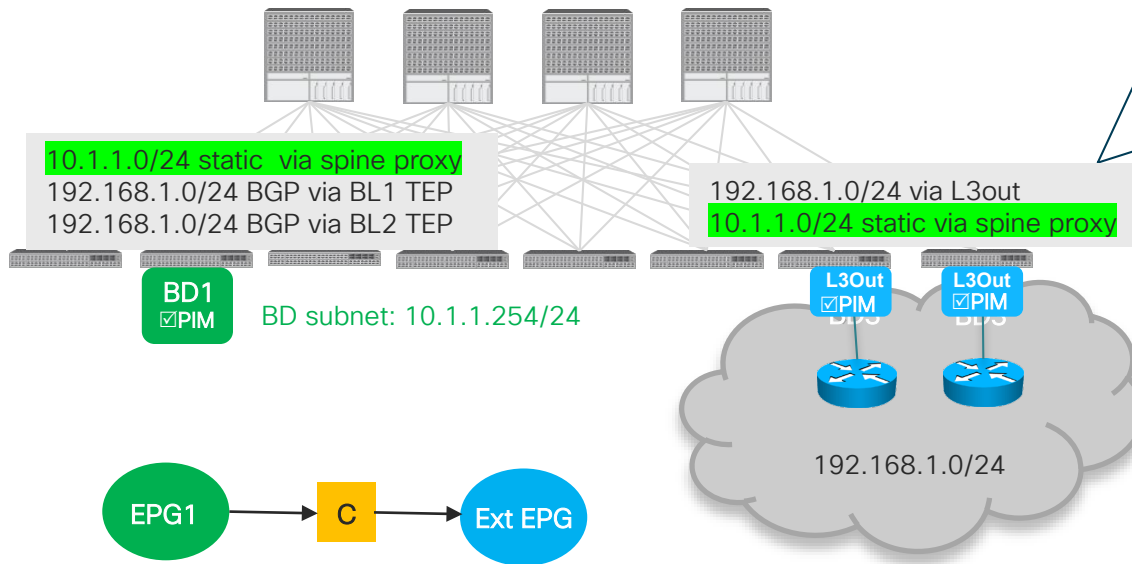


PIM-SSM Support

- Default SSM multicast range 232.0.0.0/8
- Configuration of different SSM range supported
- PIM enabled BDs run IGMPv2. Must enable IGMPv3 in BD IGMP policy
- Supports IGMP SSM Translate (allows IGMP version 2 hosts to join SSM groups)

Contract Requirement

Multicast traffic is not filtered by contracts
Contracts are required only for the routing requirement



1. BD Subnet is not programmed on BL. BL cannot advertise BD subnet out L3Out
2. BD subnet will be programmed on BL only if policy allows traffic to BL subnet
 - Contract
 - vzAny Contract
 - Preferred Group
 - Unenforced VRF

IPv6 Multicast

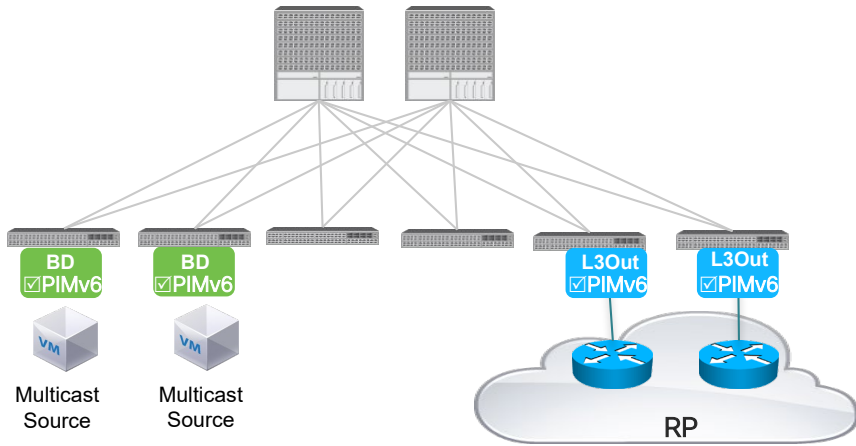


IPv6 Multicast

- MLD snooping support added in release 4.1(1)
- Supports MLDv1 and MLDv2*
- Supports up to 2,000 IPv6 multicast groups across the fabric
- Not supported on 1st generation leaf switches (EX/FX/FX2 only)
- Supports the same flooding modes as IPv4 (Flood/Optimized Flood)
- IPv6 Multicast routing (PIMv6) supported from 4.2(1) release

* MLDv2 forwarded in hardware based on (*,G) lookup

IPv6 Layer-3 Multicast



- L3Outs are required for multicast routing
- L3Outs must have a Global IPv6 unicast address configured (can be the same loopback used for IPv4/router-id)
- IPv6 tenant multicast is carried across the fabric tunnel (VRF GIPo IPv4 multicast address)
- Dual stack will require both an IPv4 and IPv6 unicast loopback address
- Only external RP supported
- Only external receivers supported in the 4.2(1) release

Multicast in a Multi-Pod Fabric



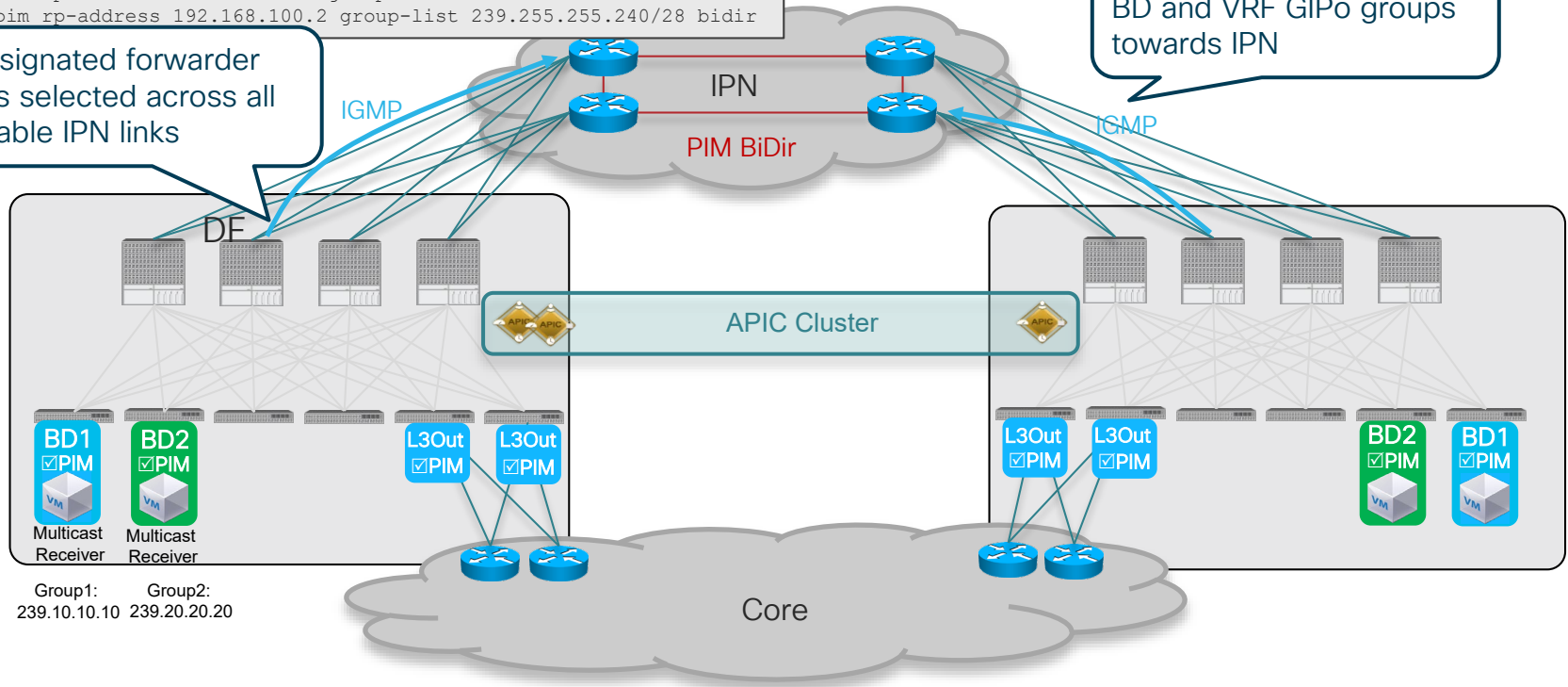
ACI Multi-Pod – Underlay Control Plane

IPN Configured PIM Bidir for fabric GIPO range

```
vrf context ipn  
ip pim rp-address 10.101.200.1 group-list 225.0.0.0/15 bidir  
ip pim rp-address 192.168.100.2 group-list 239.255.255.240/28 bidir
```

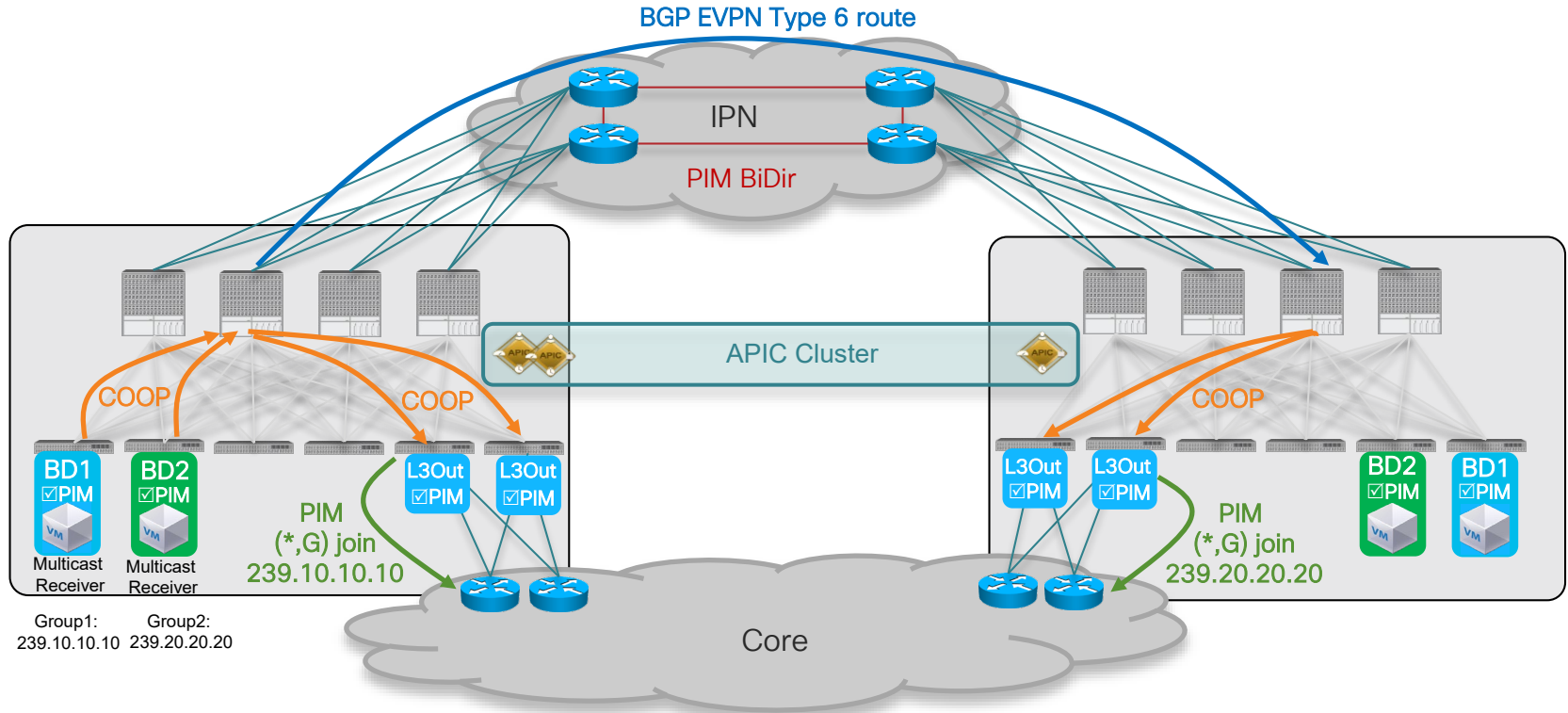
A designated forwarder link is selected across all available IPN links

Spines send IGMP joins for BD and VRF GIPO groups towards IPN



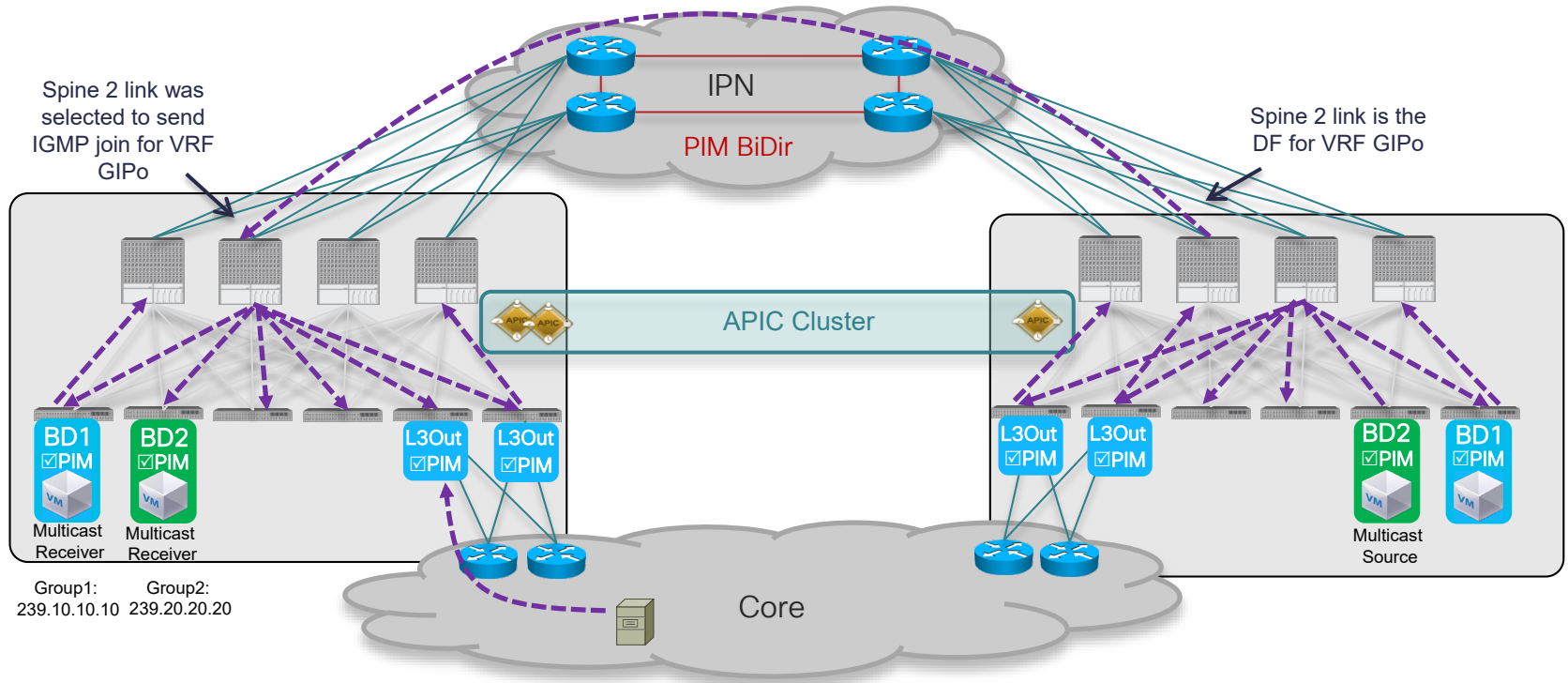
Group1: 239.10.10.10
Group2: 239.20.20.20

ACI Multi-Pod - Overlay Control Plane



ACI Multi-Pod – Data Plane

Multicast Forwarding between pods – Data Plane



Multicast Configuration



How do we configure multicast routing?



Three (or four) steps to enable L3 multicast

1. Enable Multicast at the VRF
2. Enable multicast for BDs where Multicast is required
3. Enable multicast for the L3Out
4. Configure and RP (for ASM)

1 Enable Multicast at the VRF level

Enable multicast for BDs and L3outs



The screenshot shows the Cisco GUI Multicast configuration page. On the left is a navigation tree with "Prod" selected, containing "Application Profiles", "Networking", "Bridge Domains", "VRFs", "Multicast", and "External Bridged Networks". The main area is titled "Multicast" and has a "Configuration" tab selected. Below the tabs, there is a "Bridge Domains" table and an "Interfaces" table. The "Bridge Domains" table has a red box around its header and contains rows for "Prod/BD1", "Prod/BD2", and "Prod/BD3". The "Interfaces" table has a red box around its header and contains a row for "L3Out".

Bridge Domains	IGMP Policy
Prod/BD1	
Prod/BD2	
Prod/BD3	

Interfaces	Interface Group	Interface	IGMP Policy	PIM Policy
L3Out				

2 Enable PIM for Bridge Domains

3 Enable PIM for L3out

Configure PIM Rendezvous Point (RP)



For Your Reference

4 Configure an RP address (ASM)

The screenshot shows the Cisco ICM configuration page for Multicast. The left sidebar contains a navigation tree with categories like 'Prod', 'Application Profiles', 'Networking', 'Bridge Domains', 'VRFs', and 'Multicast'. The main content area is titled 'Fabric RP Configuration' and has tabs for 'Interfaces', 'Rendezvous Points', 'Pattern Policy', 'PIM Setting', 'IGMP Setting', and 'Inter-VRF Multicast'. The 'Rendezvous Points' tab is active, showing two sections: 'Static RP' and 'Fabric RP'. In the 'Static RP' section, a table lists an entry with 'IP' as the 'RouteMap' and '10.100.100.100' as the address, which is highlighted with a red box. The 'Fabric RP' section is currently empty, displaying the message 'No items have been found. Select Actions to create a new item.' Below these sections are configuration options for 'Auto-RP' and 'Bootstrap Router (BSR)'. The 'Auto-RP' section includes checkboxes for 'Forward Auto-RP Updates' and 'Listen to Auto-RP Updates', and a dropdown for 'MA Filter'. The 'BSR' section includes checkboxes for 'Forward BSR Updates' and 'Listen to BSR Updates', and a dropdown for 'BSR Filter'. A blue callout box with the text 'Auto-RP and BSR options are also supported' has arrows pointing to these two sections.

Multicast Features



Multicast routing is enabled by enabling the VRF, BD, L3out and RP configuration for ASM but additional multicast features are supported

IGMP Features

IGMP Report Policy

IGMP static join

IGMP fast leaves

IGMP state-limit

IGMP SSM translate

PIM Features

PIM Authentication

PIM timers

PIM Join/Prune filters

PIM neighbor filters

PIM multicast domain boundary

Auto-RP

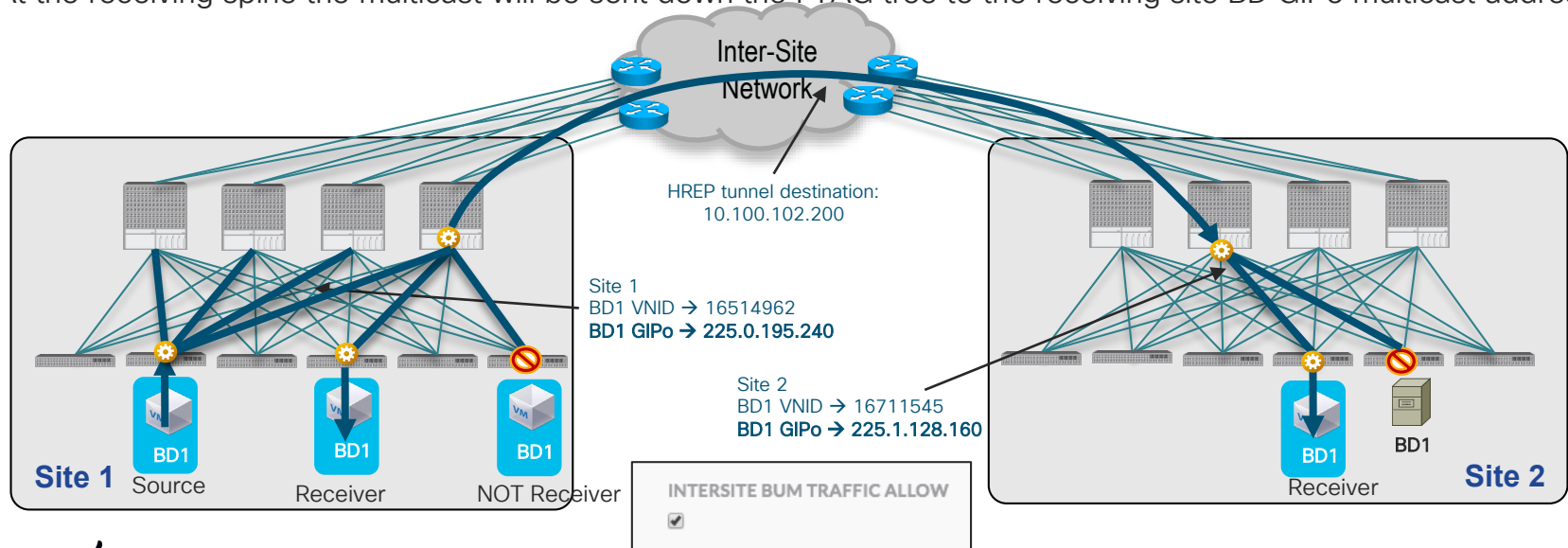
BSR

Multicast with Multi-Site Fabrics

Multicast Routing with Multi-Site

L2 Multicast over Multi-Site

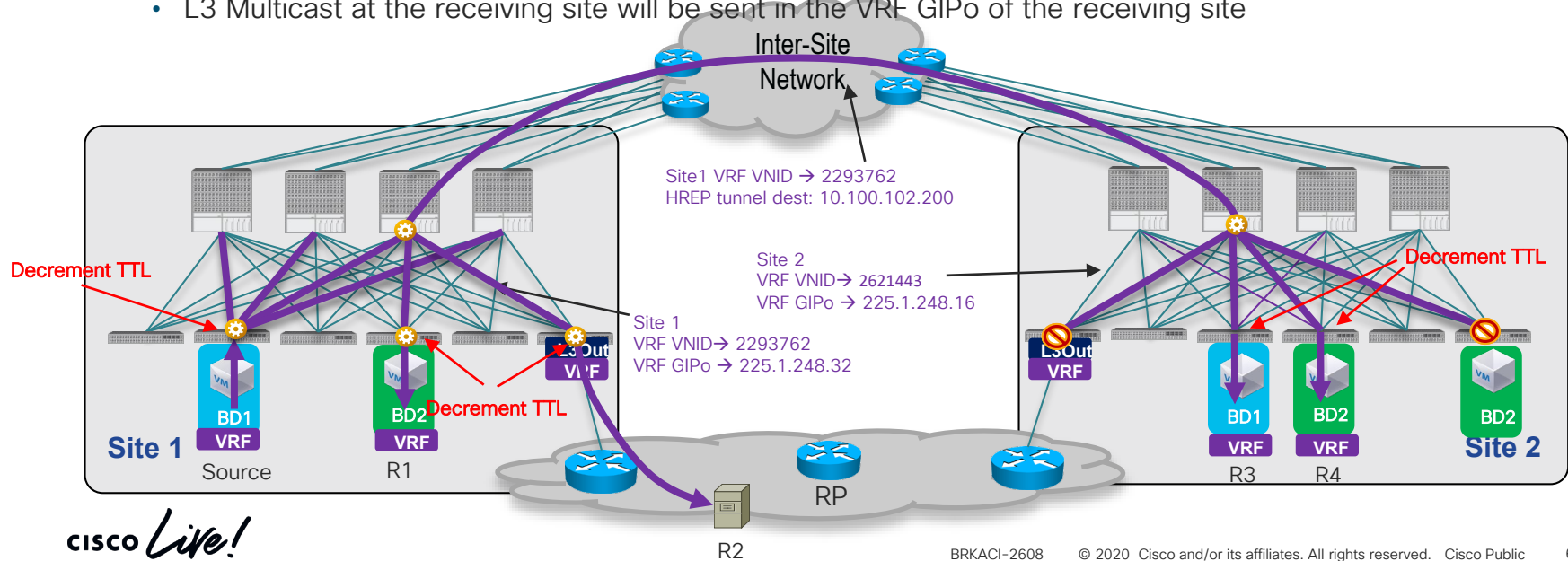
- Stretched BDs with BUM Traffic Enabled (no PIM configuration required)
- Within a site the L2 multicast is sent to the BD GIPo multicast address (unique per site) → reaches all the spines and the leaf nodes where the BD is defined (configuration driven)
- Spine elected as Designated Forwarder (DF) replicate the stream to each remote sites where the BD is stretched
- At the receiving spine the multicast will be sent down the FTAG tree to the receiving site BD GIPo multicast address



Multicast Routing with Multi-Site

L3 Multicast over Multi-Site (Source Inside the Fabric)

- Built as Routing-First Approach (decrement TTL at source and destination ACI leaf node)
- L3 Multicast is always sent to the VRF GIPO within a site (existing behavior)
- Between sites it is sent over the HREP tunnel to the Multicast TEP of the remote sites where the VRF is stretched (the VXLAN header will include the source site VRF VNID)
- L3 Multicast at the receiving site will be sent in the VRF GIPO of the receiving site



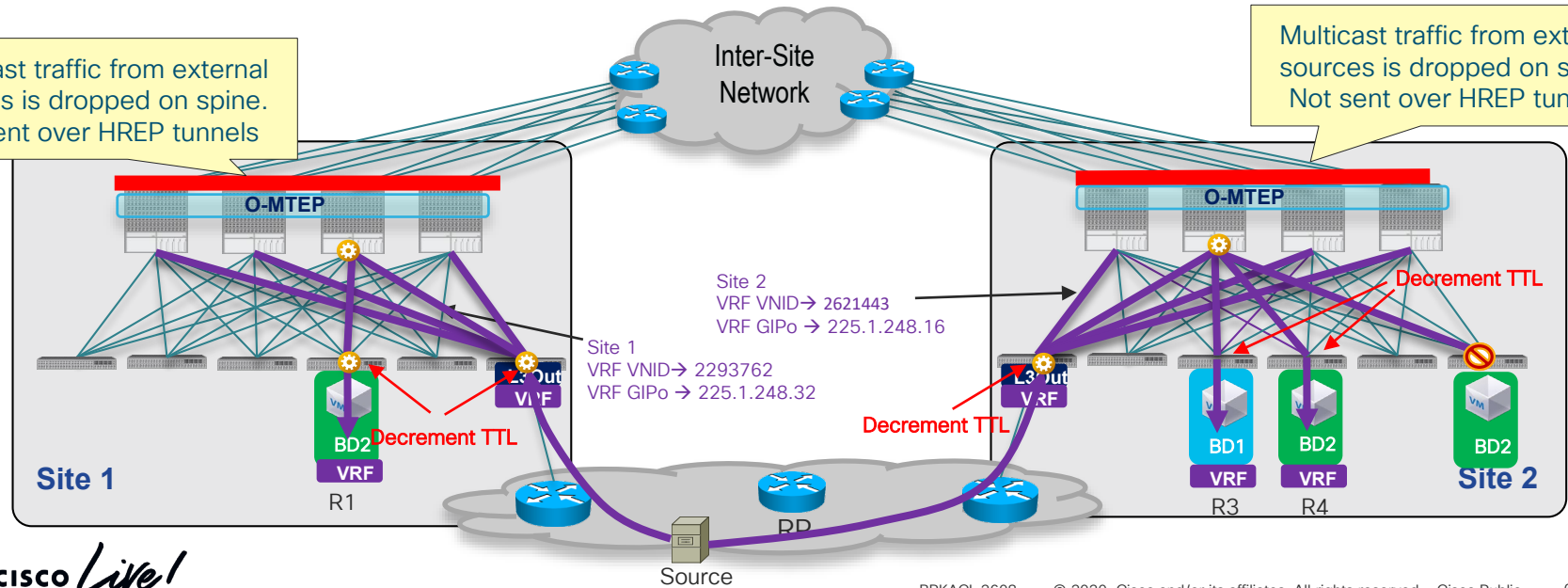
Multicast Routing with Multi-Site

L3 Multicast over Multi-Site (Source Outside the Fabric)

- Local L3Out must be used to receive traffic from an external source
- Multicast traffic from external sources dropped on the spines (to avoid traffic duplication)

Multicast traffic from external sources is dropped on spine. Not sent over HREP tunnels

Multicast traffic from external sources is dropped on spine. Not sent over HREP tunnels



Inter-VRF multicast

Inter-VRF Multicast

Allows multicast receivers in one VRF to receive multicast traffic from sources in another VRF

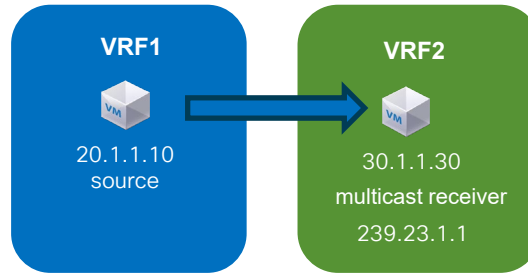
Always forwarded in the source VRF across the fabric

Receiving switch responsible for crossing VRFs

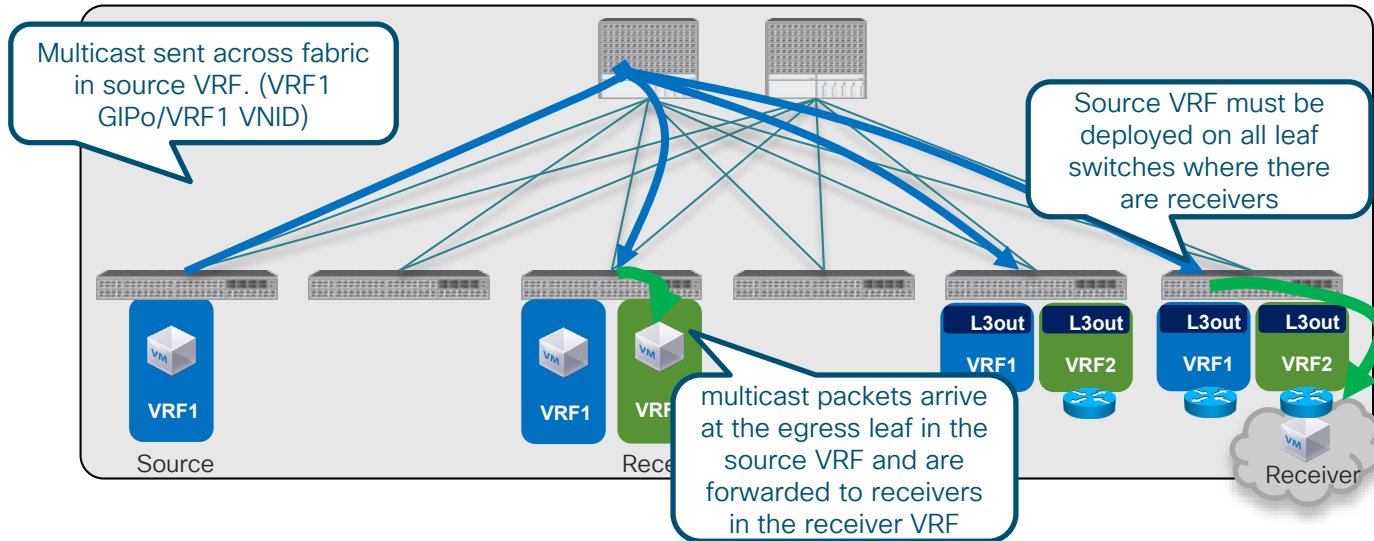
The RPF lookup for the source/RP is done in the source VRF

Requires source VRF to be present on all switches where receiver VRF is located

Inter-VRF Multicast Flow

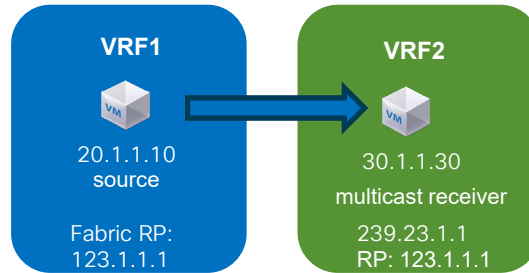


Note: The source VRF is not automatically programmed on the switch where the receiver VRF is located. Requires user to configure an EPG on the switch in the source VRF

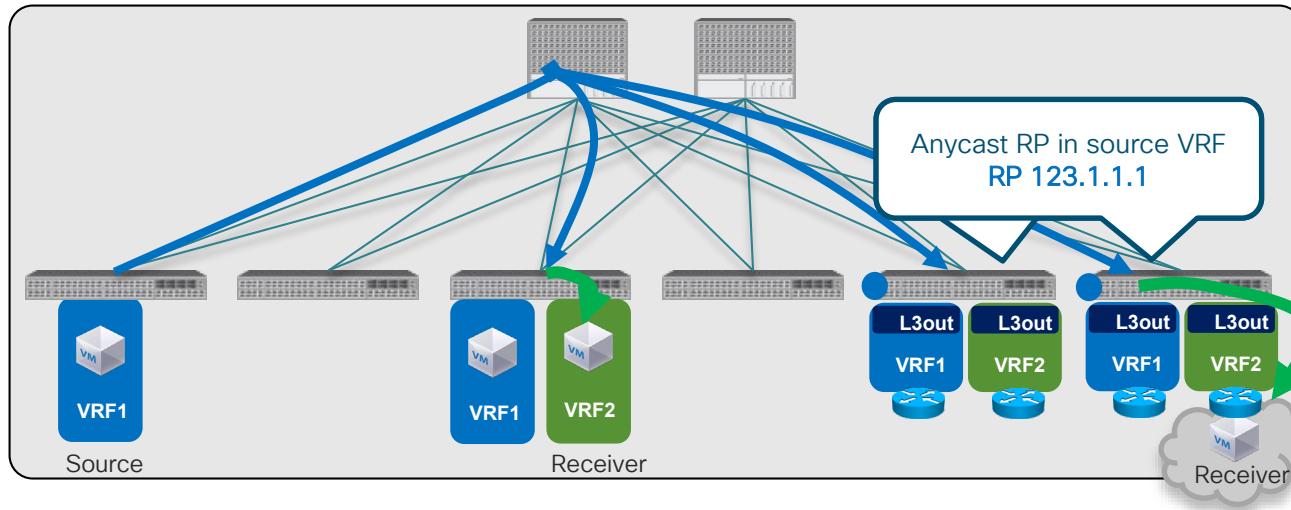


Inter-VRF Multicast Flow RP Requirements

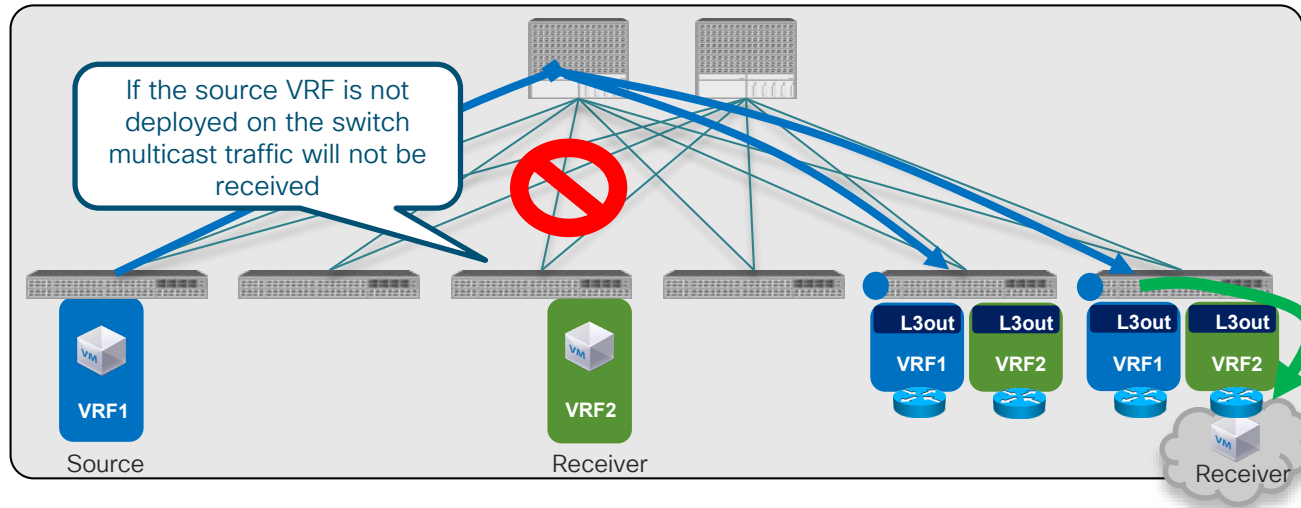
Requires RP in the fabric configured in the source VRF



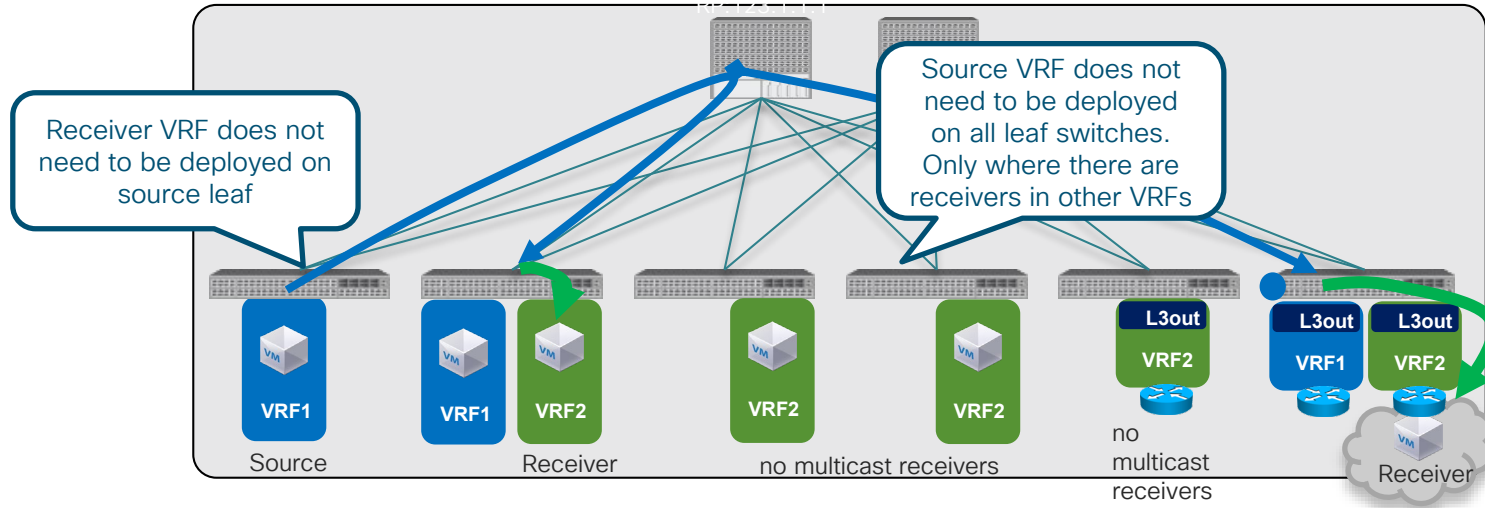
Static RP configured in receiver VRF



Inter-VRF Multicast source VRF not deployed on leaf with receivers: **Not Supported**



Inter-VRF Source VRF only deployed where required: Supported



Inter-VRF Configuration



For Your Reference

Multicast

Configuration Faults History

Interfaces **Rendezvous Points** Pattern Policy PIM Setting IGMP Setting Inter-VRF Multicast

Static RP

Fabric RP

Fabric RP in source VRF

IP RouteMap

123.1.1.1

Multicast

Configuration Faults History

Rendezvous Points Pattern Policy PIM Setting IGMP Setting **Inter-VRF Multicast**

Configure route-map with multicast group range and the VRF where the multicast sources are located

Tenant	Source VRF	RouteMap
ABC	VRF	abc-route-map

Note: sources for a multicast group cannot be in different VRFs

Troubleshooting

Layer-3 ASM Multicast States

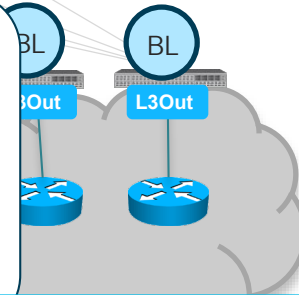
First Hop Router (FHR)

```
Leaf-1# show ip mroute 239.10.10.10 vrf CL:vrf
IP Multicast Routing Table for VRF "CL:vrf"

(10.10.1.10/32, 239.10.10.10/32), uptime: 00:00:14, ip pim
Incoming interface: Tunnel25, RPF nbr: 10.1.184.66 (pervasive)
Outgoing interface list: (count: 0)
```



1. FHR leaf will show (S,G) mroute with empty OIF for “show ip mroute” command
2. The “show forwarding distribution multicast route” command will display the outgoing tunnel interface



```
Leaf-1# vsh -c 'show forwarding distribution multicast route vrf CL:vrf'

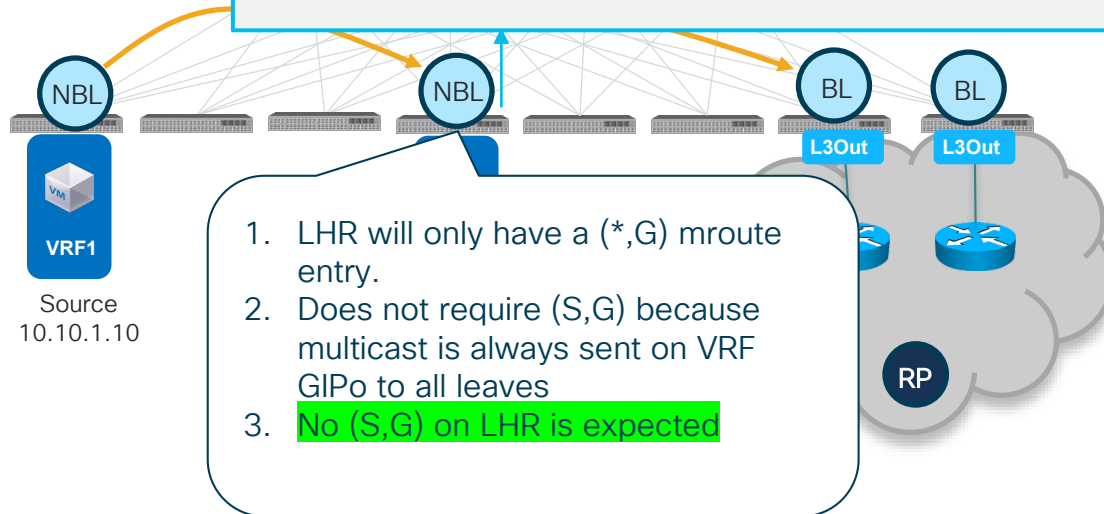
(10.10.1.10/32, 239.10.10.10/32), RPF Interface: Tunnel25, flags: 0
Received Packets: 2 Bytes: 130
Number of Outgoing Interfaces: 1
Outgoing Interface List Index: 8202
Tunnel25
```

Layer-3 ASM Multicast States

Last Hop Router (LHR)

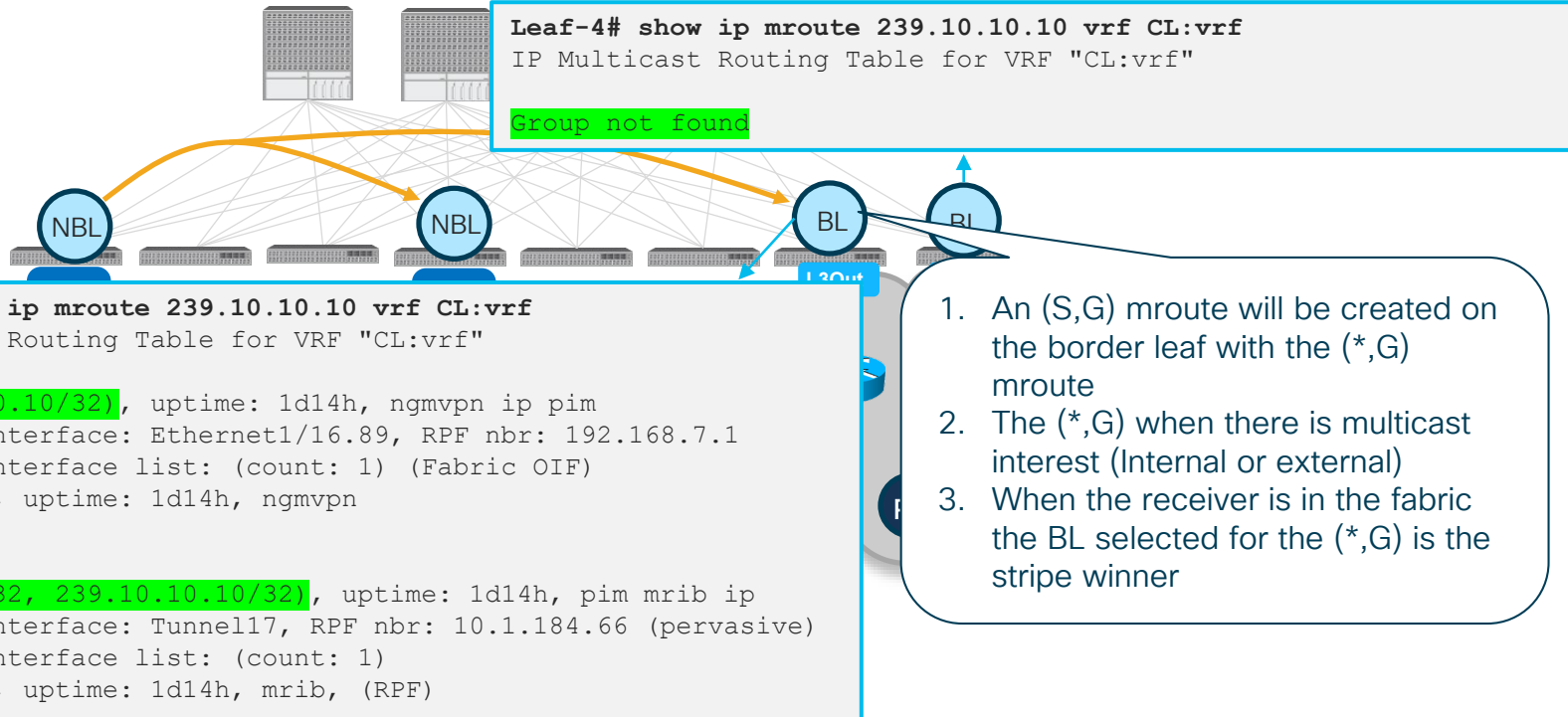
```
Leaf-2# show ip mroute 239.10.10.10 vrf CL:vrf
IP Multicast Routing Table for VRF "CL:vrf"

(*, 239.10.10.10/32), uptime: 17:08:54, igmp ip pim
Incoming interface: Tunnel24, RPF nbr: 10.1.176.64
Outgoing interface list: (count: 1)
  Vlan104, uptime: 17:08:54, igmp
```



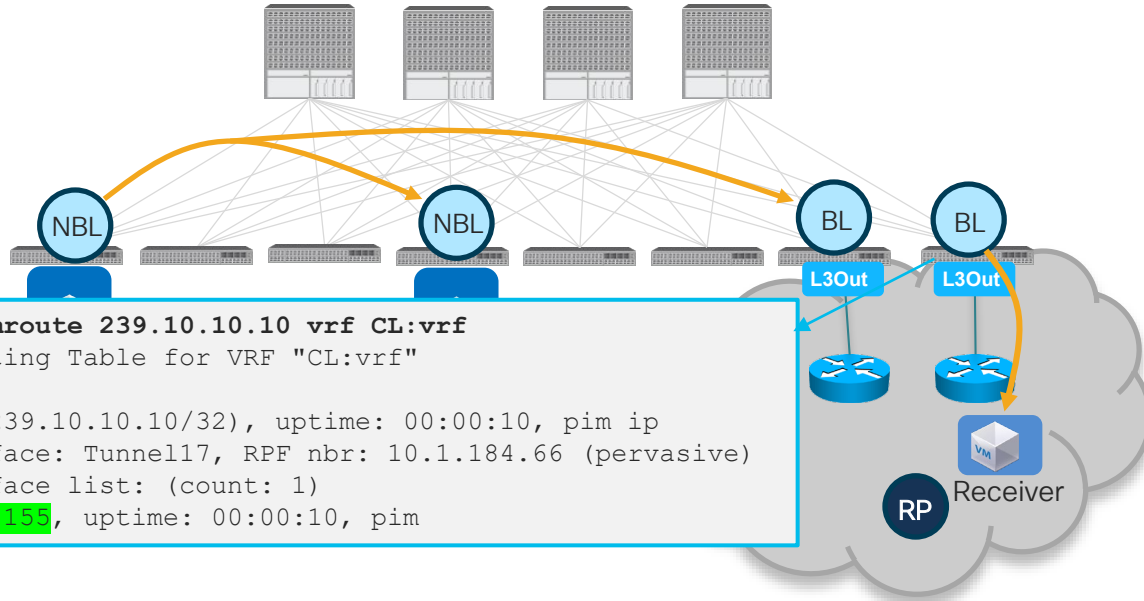
Layer-3 ASM Multicast States

Border Leaf



Layer-3 ASM Multicast States

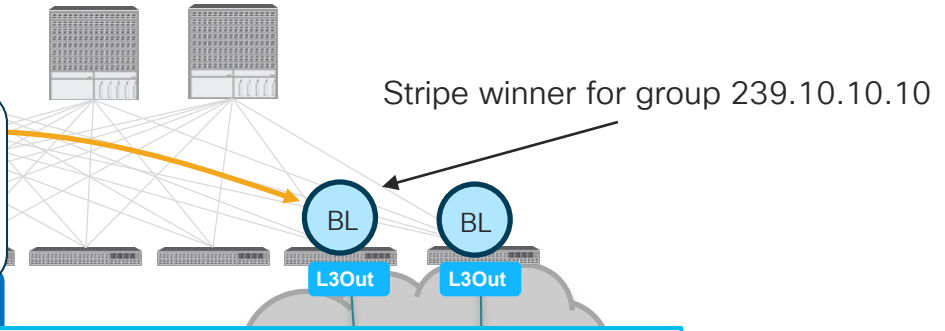
Border Leaf, External Receiver



Layer-3 ASM Multicast States

Stripe Winner

1. Command to find stripe winner can be executed on any leaf
2. Hash based on group and VRF VNID

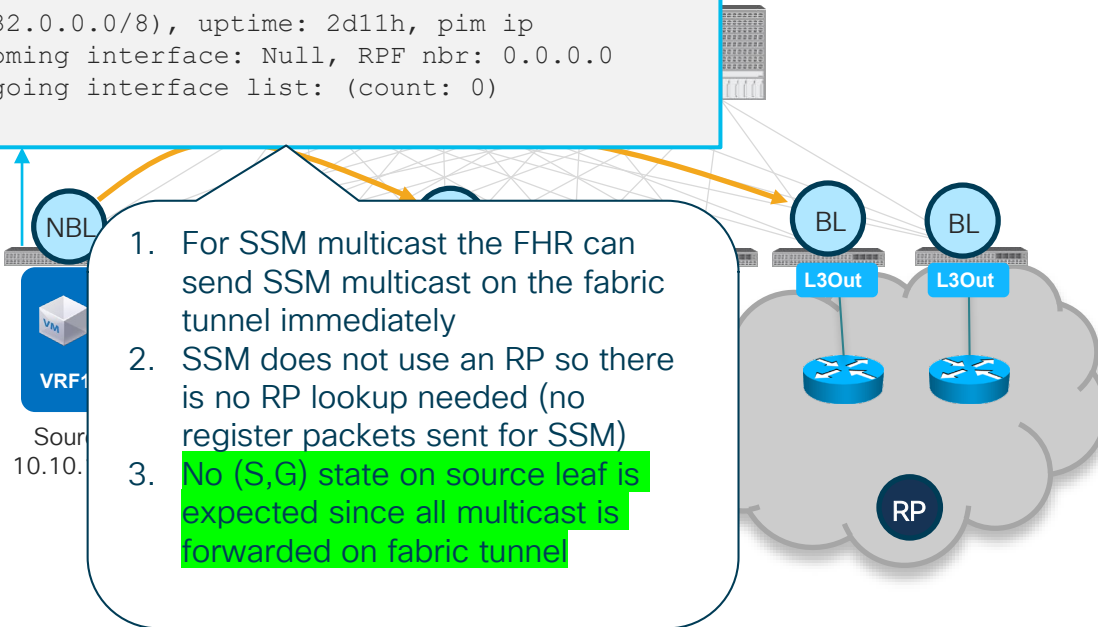


```
Leaf-3# vsh -c 'show ip pim internal stripe-winner 239.10.10.10 vrf CL:vrf'  
PIM Stripe Winner info for VRF "CL:vrf" (BL count: 2)  
(*, 239.10.10.10)  
BLs:  
Group hash 1262394260 VNID 3047428  
1.1.1.103 hash: 1948335499 (local)  
1.1.1.104 hash: 725726570  
Winner: 1.1.1.103 best_hash: 1948335499
```

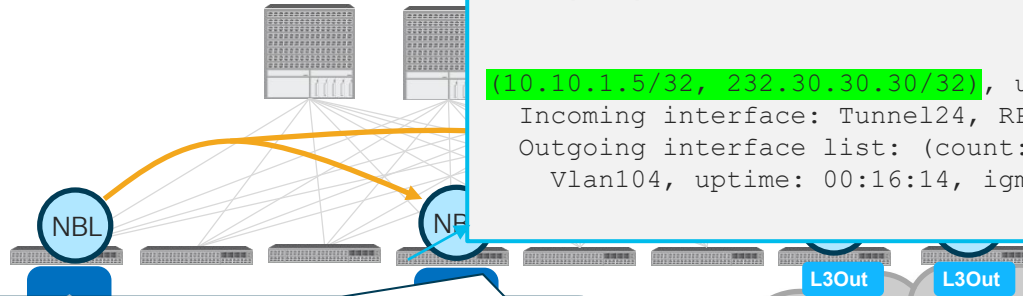
Layer-3 SSM Multicast First Hop Router (FHR)

```
Leaf-1# show ip mroute vrf CL:vrf
IP Multicast Routing Table for VRF "CL:vrf"

(*, 232.0.0.0/8), uptime: 2d11h, pim ip
  Incoming interface: Null, RPF nbr: 0.0.0.0
  Outgoing interface list: (count: 0)
```



Layer-3 SSM Multicast Last Hop Router (LHR)



```
Leaf-2# show ip mroute vrf CL:vrf
IP Multicast Routing Table for VRF "CL:vrf"

(*, 232.0.0.0/8), uptime: 2d11h, pim ip
  Incoming interface: Null, RPF nbr: 0.0.0.0
  Outgoing interface list: (count: 0)

(10.10.1.5/32, 232.30.30.30/32), uptime: 00:16:14, igmp ip pim
  Incoming interface: Tunnel24, RPF nbr: 10.1.184.66 (pervasive)
  Outgoing interface list: (count: 1)
    Vlan104, uptime: 00:16:14, igmp
```

1. Last Hop Router leaf will have (S,G) mroute entry
2. The (S,G) state is created by the (S,G) IGMP join

Group 232.30.30.30
Source: 10.10.1.5

Layer-3 ASM Multicast Troubleshooting

Is source leaf forwarding multicast?

1. If register stop is not received this indicates a possible configuration or routing issue
2. The FHR will not forward multicast if the RP is not configured "show ip pim rp" command to check
3. The FHR will not forward multicast if there is no route to the RP "show ip route" command to check
4. The FHR will forward multicast if the RP is configured and there is a route to the RP

```
Leaf-2# show ip mroute 239.10.10.10 detail vrf CL:vrf
IP Multicast Routing Table for VRF "CL:vrf"
```

```
(10.10.1.10/32, 239.10.10.10/32), uptime: 00:00:37, ip(0) pim(0)
Data Created: Yes
Received Register stop
Pervasive
VPC Flags
  RPF-Source Forwarder
Stats: 0/0 [Packets/Bytes], 0.000 bps
Incoming interface: Tunnel25, RPF nbr: 10.1.184.66 (pervasive)
Outgoing interface list: (count: 0)
```

```
Leaf-2# show ip mroute 239.10.10.10 detail vrf CL:vrf
IP Multicast Routing Table for VRF "CL:vrf"
```

```
(10.10.1.10/32, 239.10.10.10/32), uptime: 00:00:37, ip(0) pim(0)
Data Created: Yes
Pervasive
VPC Flags
  RPF-Source Forwarder
Stats: 0/0 [Packets/Bytes], 0.000 bps
Incoming interface: Tunnel25, RPF nbr: 10.1.184.66 (pervasive)
Outgoing interface list: (count: 0)
```

Layer-3 ASM Multicast Troubleshooting

Is receiver leaf receiving multicast?

Multicast not received

1. Receiver has joined multicast group (sent IGMP join report)
2. Incoming interface is Null indicating multicast is not being received
3. RPF is 0.0.0.0

```
Leaf-2# show ip mroute 239.10.10.10 vrf CL:vrf
```

```
IP Multicast Routing Table for VRF "CL:vrf"
```

```
(*, 239.10.10.10/32), uptime: 00:00:12, igmp ip pim  
Incoming interface: Null, RPF nbr: 0.0.0.0  
Outgoing interface list: (count: 1)  
Vlan104, uptime: 00:00:12, igmp
```

Multicast received

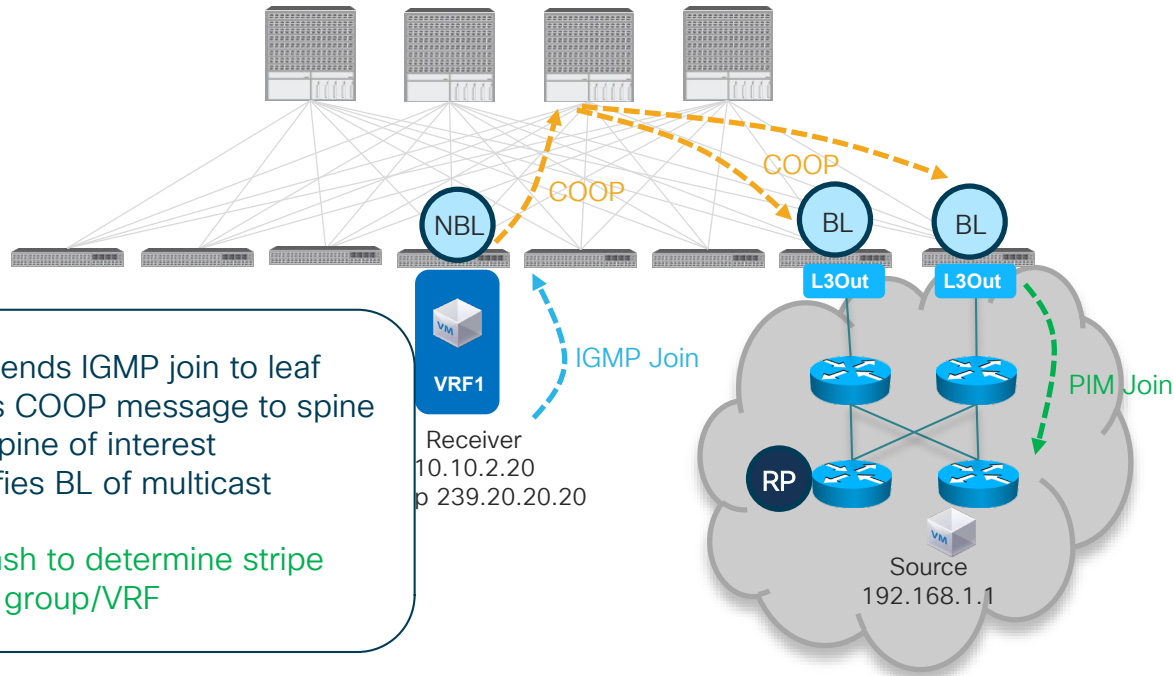
1. Incoming interface shows fabric tunnel interface
2. RPF neighbor is spine anycast proxy address.

```
Leaf-2# show ip mroute 239.10.10.10 vrf CL:vrf
```

```
IP Multicast Routing Table for VRF "CL:vrf"
```

```
(*, 239.10.10.10/32), uptime: 00:03:13, igmp ip pim  
Incoming interface: Tunnel124, RPF nbr: 10.1.176.64  
Outgoing interface list: (count: 1)  
Vlan104, uptime: 00:03:13, igmp
```

Border leaf loopback requirement



1. Receiver sends IGMP join to leaf
2. LHR sends COOP message to spine notifying spine of interest
3. Spine notifies BL of multicast interest
4. BLs run hash to determine stripe winner for group/VRF

Border Leaf loopback requirement

Failed state

1. No PIM neighbors will be formed over the fabric

```
Leaf-1# show ip pim neigh vrf CL:vrf
```

```
PIM Neighbor information for Dom:CL:vrf
Neighbor   Interface   Uptime   Expires   DRPriority Bidir   BFDState
```

1. Receiver sends IGMP join to leaf
2. LHR sends COOP message to spine notifying spine of interest
3. Spine notifies BL of multicast interest
4. **No stripe winner. Join not sent towards external RP/Source**

```
Leaf-103# show ip pim internal stripe-winner 239.10.10.11 vrf
CL:vrf
PIM Stripe Winner info for VRF "CL:vrf" (BL count: 0)
(*, 239.10.10.11)
BLs:
Group hash 158879015 VNID 3047428
Winner: 255.251.249.253 best_hash: 0
```

Border Leaf loopback requirement

Working state

1. PIM neighbors formed over the fabric

```
Leaf-1# show ip pim neigh vrf CL:vrf
```

```
PIM Neighbor information for Dom:CL:vrf
```

Neighbor	Interface	Uptime	Expires	DRPriority	Bidir	BFDState
1.1.1.104/32	tunnel25	00:00:08	00:01:36	1	no	n/a
1.1.1.103/32	tunnel25	00:00:08	00:01:42	1	no	n/a

1. Receiver sends IGMP join to leaf
2. LHR sends COOP message to spine notifying spine of interest
3. Spine notifies BL of multicast interest
4. BLs run hash to determine stripe winner for group/VRF

```
Leaf-103# show ip pim internal stripe-winner 239.20.20.20 vrf CL:vrf
```

```
PIM Stripe Winner info for VRF "CL:vrf" (BL count: 2)  
(*, 239.20.20.20)
```

```
BLs:
```

```
Group hash 738275646 VNID 3047428  
1.1.1.103 hash: 426404817 (local)  
1.1.1.104 hash: 1470373220
```

```
Winner: 1.1.1.104 best_hash: 1470373220
```

Troubleshoot multicast data plane with ELAM Assistant

System Tenants Fabric Virtual Networking L4-L7 Services Admin Operations **Apps** Integrations

Apps | Faults

Apps ElamAssistant



ACI Policy Viewer
by Cisco

It visualizes ACI network policies such as VRF, BD, EPG along with its domains such as physical domain

This application is installed in tabs:

- Tenant
- EPG



APIC Postman
by Cisco

A graph based application to create json/xml configuration for post operation to APIC

Open



ELAM Assistant
by Cisco

Help you perform ELAM(Embedded Logic Analyzer Module) on ACI nodes to capture a single packet at a time and analyze where the packet goes.

Open



Troubleshoot multicast data plane with ELAM Assistant

ELAM PARAMETERS

Name your capture:

Status	Node	Direction	Source I/F	Parameters
Report Ready	node-102	from fabriclink	any	dst ip 239

PIM Setting
VRF GiPo address: 225.1.193.32

L3 Header	
L3 Type	IPv4
Destination IP	239.12.12.12
Source IP	32.1.1.13
IP Protocol	0x11 (UDP)
DSCP	0
TTL	61
Do Not Fragment Bit	0x0 (0x0)

Outer L3 Header	
IPv4	225.1.193.33 (None(None) + ftag None)
Source IP	10.1.176.64 (F1-P1-Leaf-103)
IP Protocol	0x11 (UDP)
DSCP	57
TTL	31
Do not Fragment Bit	0x0 (0x0)

Outer destination address is 225.1.193.33 (FTAG 1)

225.1.193.33 (None(None) + ftag None)

Identifying FTAG root spines

System Tenants **Fabric** Virtual Networking L4-L7 Services Admin Operations Apps Integrations

Inventory Fabric Policies | Access Policies

Inventory

- Quick Start
- Topology
- Pod 1
 - F1-P1-Leaf105 (Node-105)
 - F1-P1-Leaf-101 (Node-101)
 - F1-P1-Leaf-102 (Node-102)
 - F1-P1-Leaf-103 (Node-103)
 - F1-P1-Leaf-104 (Node-104)
 - F1-P1-Spine-1001 (Node-1001)
 - Chassis
 - Interfaces
 - Protocols
 - BGP
 - COOP
 - IPV4
 - IPV6
 - ISIS
 - ISIS for VRF-overlay-1
 - Discovered Tunnel Endpoints
 - LSP Records
 - IS-IS Interfaces
 - Neighbors
 - IS-IS Routes
 - IS-IS Fabric Multicast Trees**
 - Traffic

IS-IS Fabric Multicast Trees

ID	Root Address	Root Port	Diameter	Origin	Diameter Alert	Operational
0	10.1.176.67	eth1/1.36	7	Learned through isis	Normal	Active
1	10.1.176.69	eth1/2.40	8	Learned through isis	Normal	Active
2	10.1.176.68	eth1/4.39	9	Learned through isis	Normal	Active
3	10.1.184.96	eth1/5.35	9	Learned through isis	Normal	Active
4	10.1.176.69	eth1/2.40	11	Learned through isis	Normal	Active
5	10.1.176.67	eth1/1.36	7	Learned through isis	Normal	Active
6	10.1.176.65	unspecified	6	Static	Normal	Active
7	10.1.176.65	unspecified	6	Static	Normal	Active
8	10.1.176.67	eth1/1.36	7	Learned through isis	Normal	Active
9	10.1.176.65	unspecified	6	Static	Normal	Active
10	10.1.176.65	unspecified	6	Static	Normal	Active
11	10.1.176.68	eth1/4.39	9	Learned through isis	Normal	Active
12	10.1.176.68	eth1/4.39	9	Learned through isis	Normal	Active
13	0.0.0.0	unspecified	0	Static	Normal	Inactive
14	0.0.0.0	unspecified	0	Static	Normal	Inactive
15	0.0.0.0	unspecified	0	Static	Normal	Inactive

Page 1 Of 1 Objects Per Page: 100 Displaying Object

Conclusions and Q&A

ACI Multicast Features

- Layer-3 Multicast support (Cloud Scale ASICs)
- External RP
- Auto-RP/BSR
- ASM
- SSM
- IGMP/PIM filtering features

- L3 Multicast support for FEX (No L3out)

- MLD snooping
- PIM6



- Layer-2 Multicast
- IGMP snooping
- Fabric IGMP snooping querier
- External querier
- 8,000 multicast group scale (IGMP snooping)

- Multicast with Multi-Pod

2017

- L3 Multicast with Multi-Site
- Inter-VRF multicast
- RP in the fabric (PIM anycast)*
- Multicast Route Scale (32k Mroutes)**

*RP in the fabric not supported with Multi-Site
 **FX only

Complete your online session survey



- Please complete your session survey after each session. Your feedback is very important.
- Complete a minimum of 4 session surveys and the Overall Conference survey (starting on Thursday) to receive your Cisco Live t-shirt.
- All surveys can be taken in the Cisco Events Mobile App or by logging in to the Content Catalog on ciscolive.com/emea.

Cisco Live sessions will be available for viewing on demand after the event at ciscolive.com.

Continue your education



Demos in the
Cisco campus



Walk-in labs



Meet the engineer
1:1 meetings



Related sessions



Thank you





You make **possible**