
cisco
You make possible

Troubleshooting VXLAN BGP EVPN

Shridhar V Dhodapkar CCIE#6367
Technical Leader
@shridhardh

BRKDCN-3040

cisco Live!
Barcelona | January 27-31, 2020

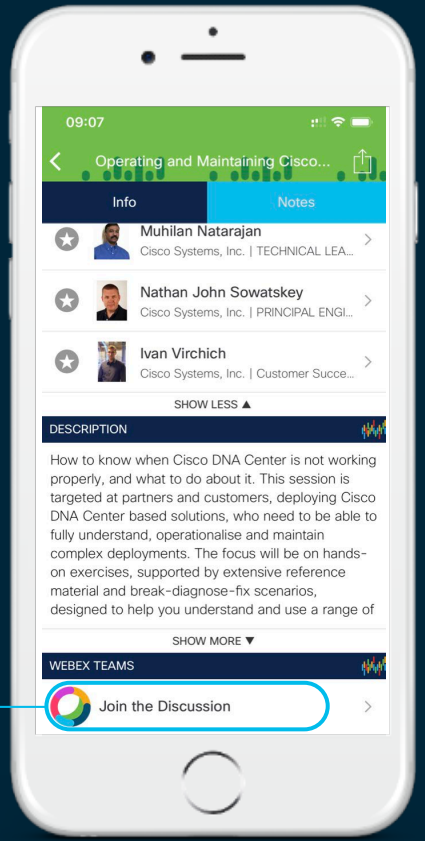
Cisco Webex Teams

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



Agenda

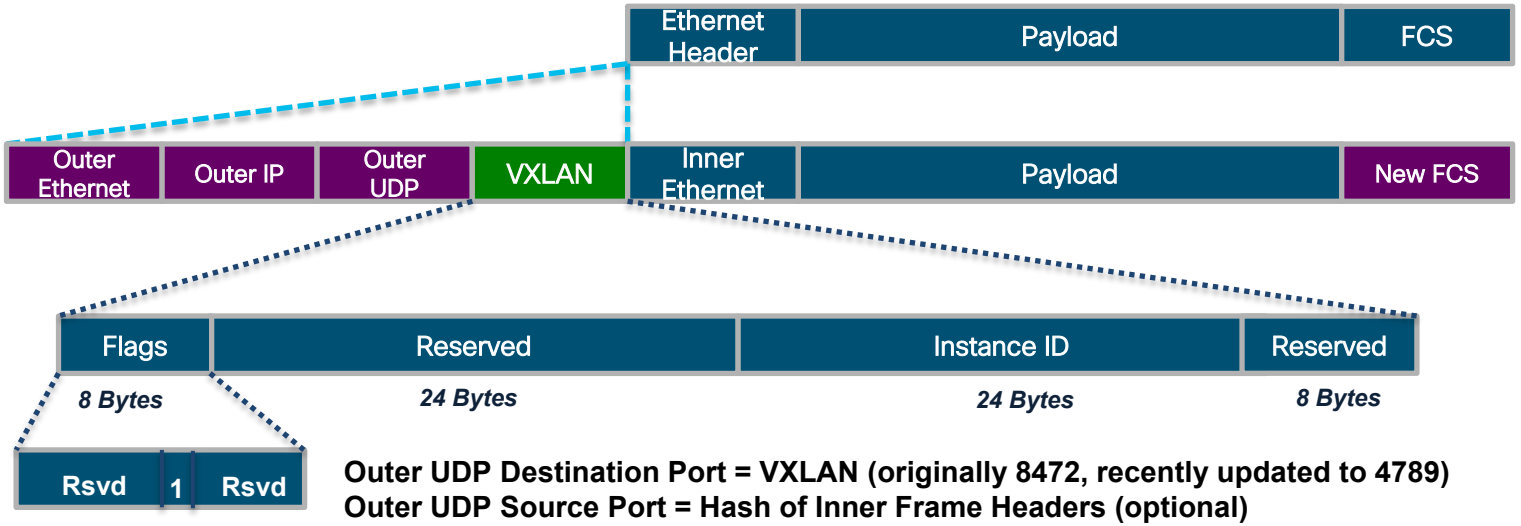
- Introduction
- VXLAN Packet Flow
- Nexus 9000 Components
  - Control Plane Troubleshooting
- Troubleshooting BUM Traffic
- Troubleshooting Tenant Routed Multicast
- Troubleshooting Tools
- Conclusion

VxLAN Overview

VXLAN Concepts

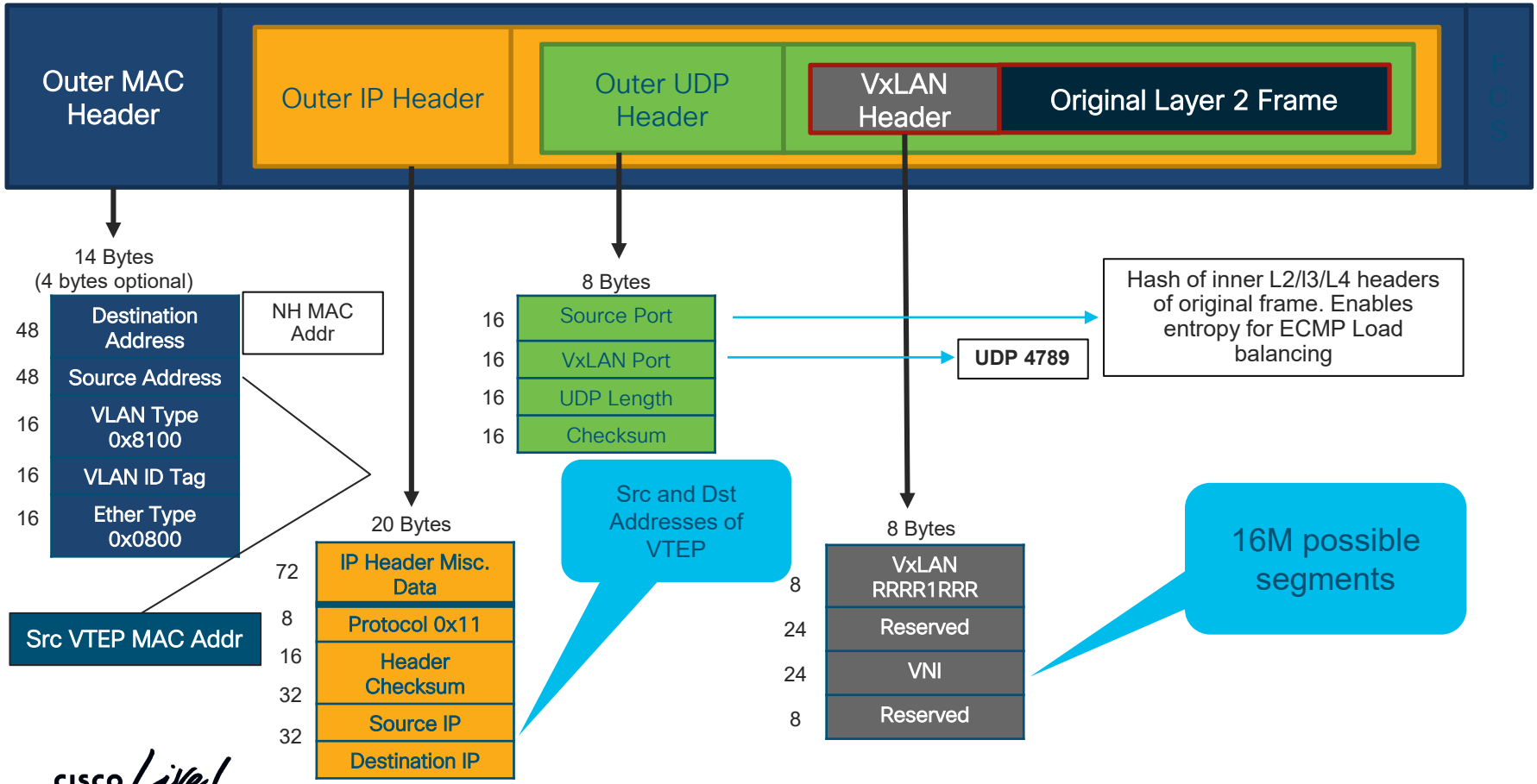
- VXLAN Overlay
  - A VXLAN Overlay or VXLAN segment is a Layer-2 broadcast domain identified by the VNID that extends or tunnels traffic from one VTEP to another.
- VXLAN Tunnel End Point (VTEP)
  - A VTEP is a device that provides both encapsulation and de-capsulation of classical Ethernet and VXLAN packets to and from a VXLAN segment
  - Each VTEP may have the following types of interfaces:
    - Switchport interfaces on the local LAN segment to support local endpoints
    - Layer-3 interfaces to the transport IP network
    - SVI interfaces
- VXLAN Gateway
  - A VTEP that bridges traffic between VXLAN segments

VXLAN Encapsulation



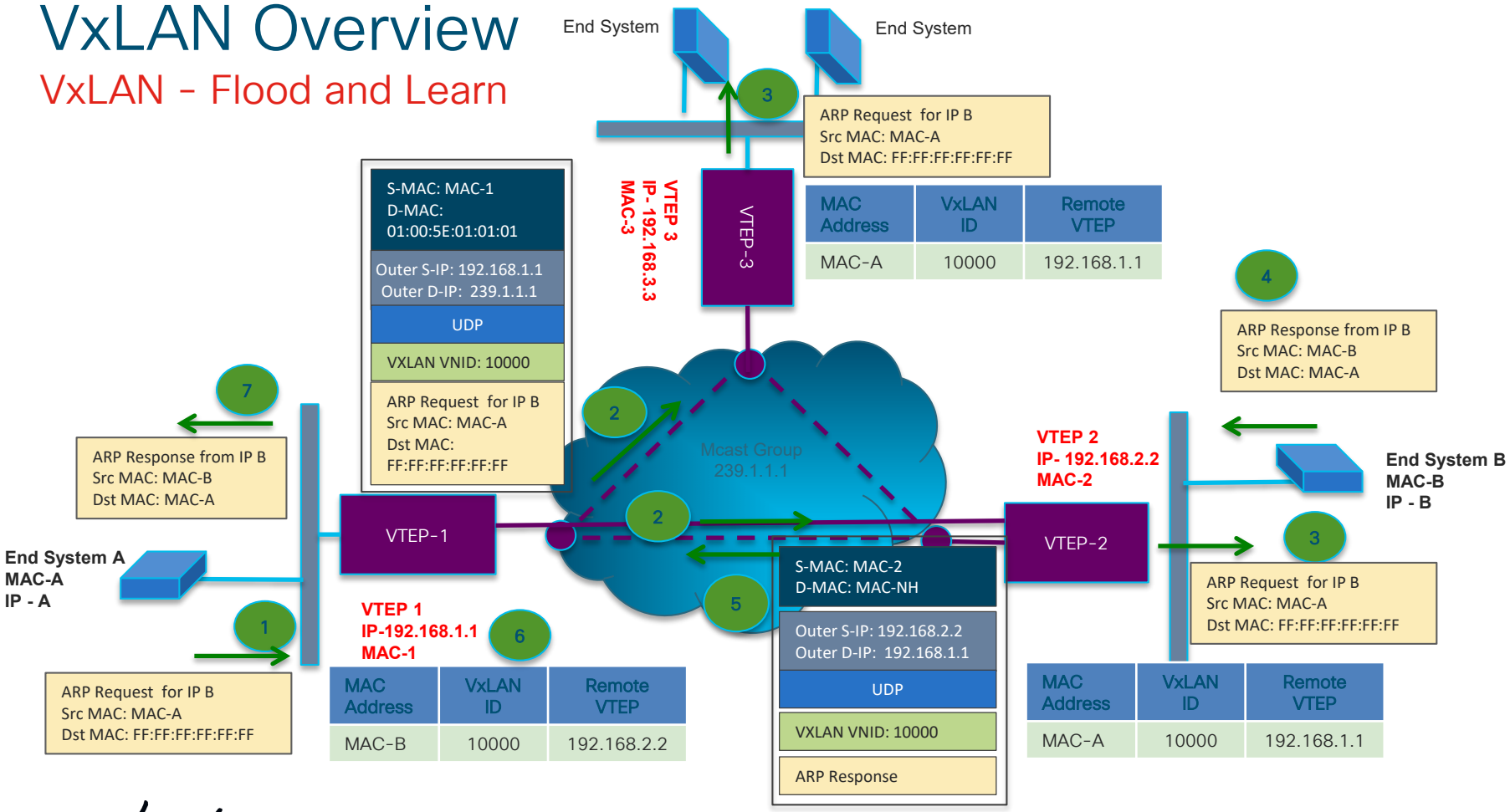
- The outer IP header has the source IP and destination IP of the VTEP endpoints
- The outer Ethernet header has the source MAC of the source VTEP and the destination MAC of the immediate Layer-3 next hop

VxLAN Packet Structure



VxLAN Overview

VxLAN – Flood and Learn



VxLAN Overview

VxLAN – Flood and Learn

- Data Plane learning technique for VxLAN
- VNI's are mapped to a multicast group on a VTEP
- Local MACs are learnt over a VLAN (VNI) on a VTEP
- Broadcast, Unknown Unicast, Multicast (*BUM Traffic*) is flooded to the delivery multicast group for that VNI
- Remote VTEPs part of same multicast group learn host MAC, VNI and source VTEP as the next-hop for the host MAC from flooded traffic
- Unicast packets to the host MAC are sent directly to source VTEP as VxLAN encapsulated packet

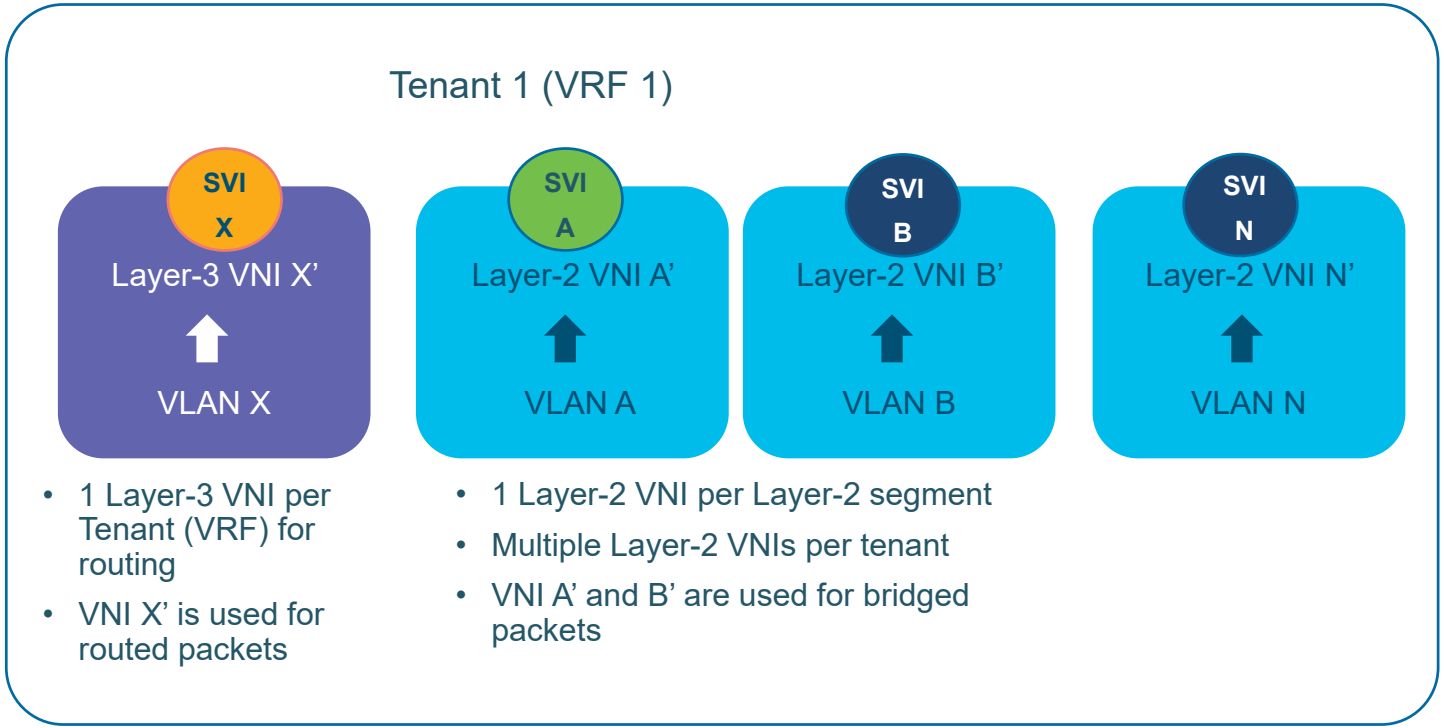
VxLAN Overview

Ingress Replication

- Some customers not comfortable deploying multicast in their core
- With *Ingress Replication* (*IR*), BUM traffic ingress access side is replicated to remote VTEP as unicast
- Static IR VETP tunnel is kept alive as long as the route to the VTEP is available.
- Support multiple VTEPs per VNI and a VTEP in multiple VNIs

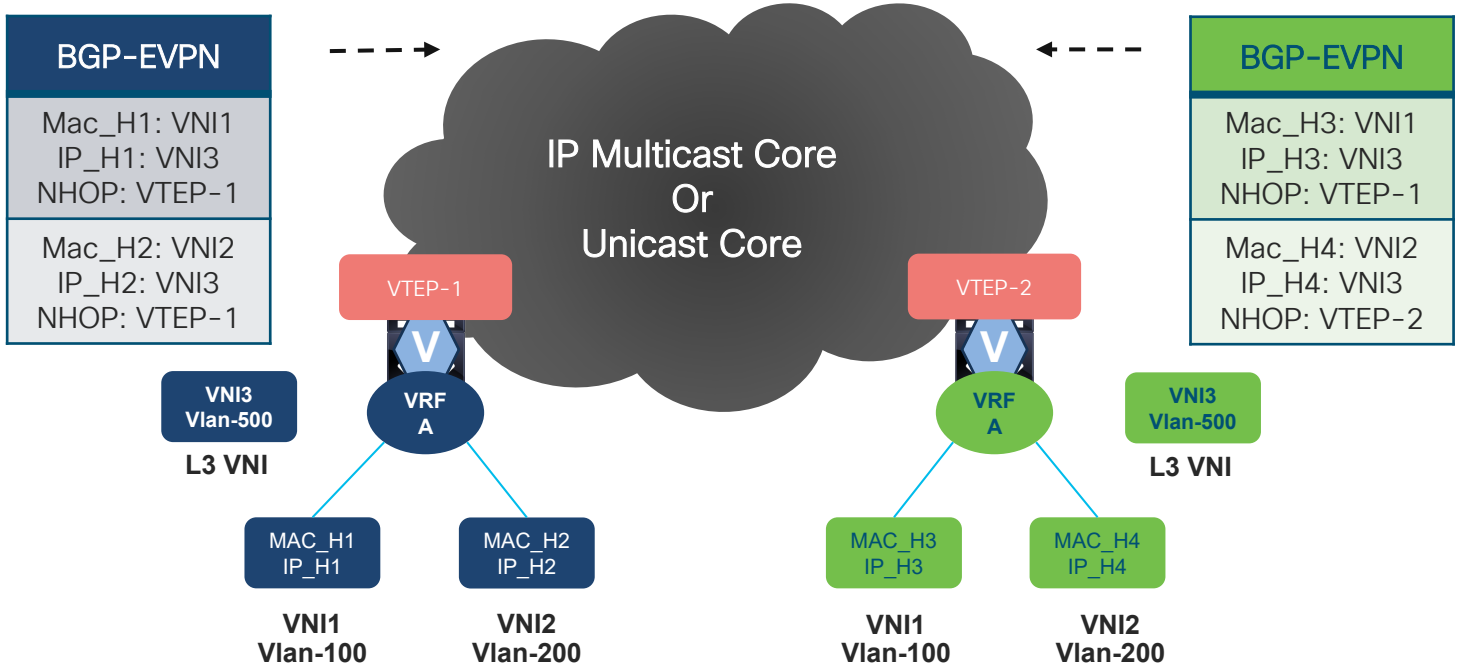
VxLAN Overview

Tenant



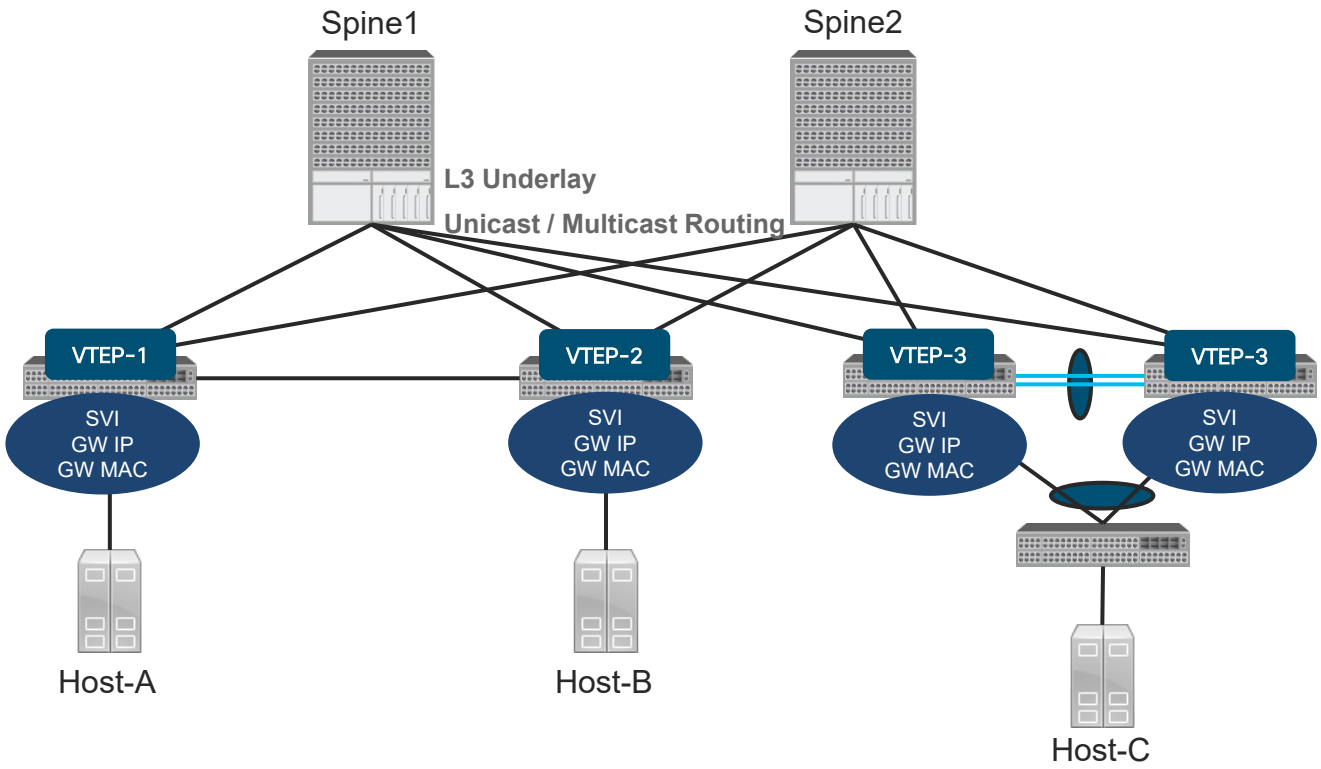
VxLAN Overview

VxLAN EVPN



VxLAN Overview

Distributed Anycast Gateway



VxLAN Overview

Distributed Anycast Gateway – Configuration

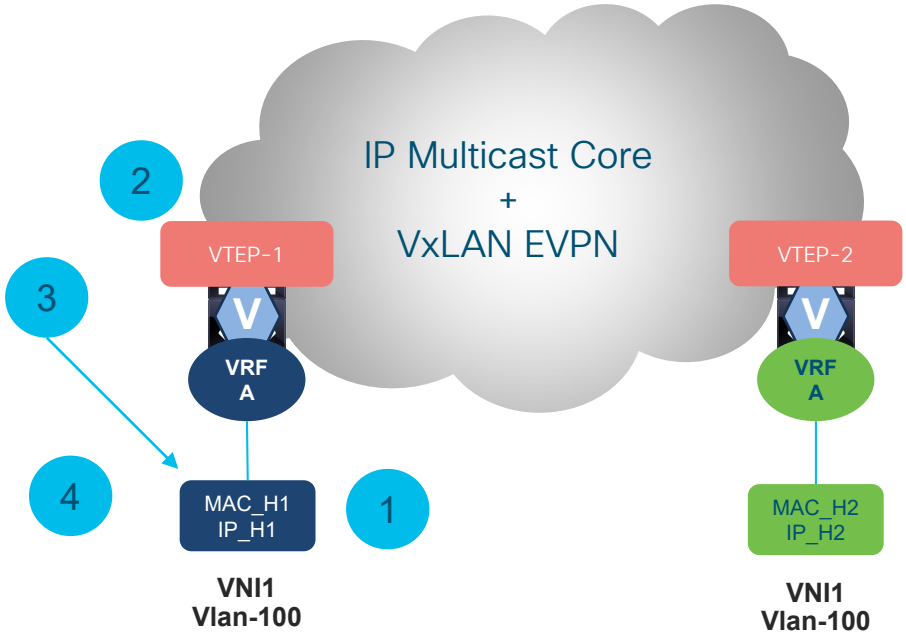
- All VTEPs has same IP address for an L2 VNI
- Anycast Gateway MAC is global to each VTEP for all VNI's for all Tenants
- One virtual MAC / VTEP
- All VTEPs should have same virtual MAC address

```
fabric forwarding anycast-gateway-mac 0001.0001.0001
!
interface Vlan100
no shutdown
vrf context test-evpn-tenant
ip address 172.16.1.254/24
fabric forwarding mode anycast-gateway
```

VxLAN Overview

ARP Suppression

- Hosts send out G-ARP when they come online
- Local leaf node receives G-ARP, creates local ARP cache and advertises to other leaf by BGP as route type 2
- Remote leaf node puts IP-MAC info into remote ARP cache and supresses incoming ARP request for this IP
- If IP info not found in ARP suppression cache table, VTEP floods the ARP request to other VTEPs



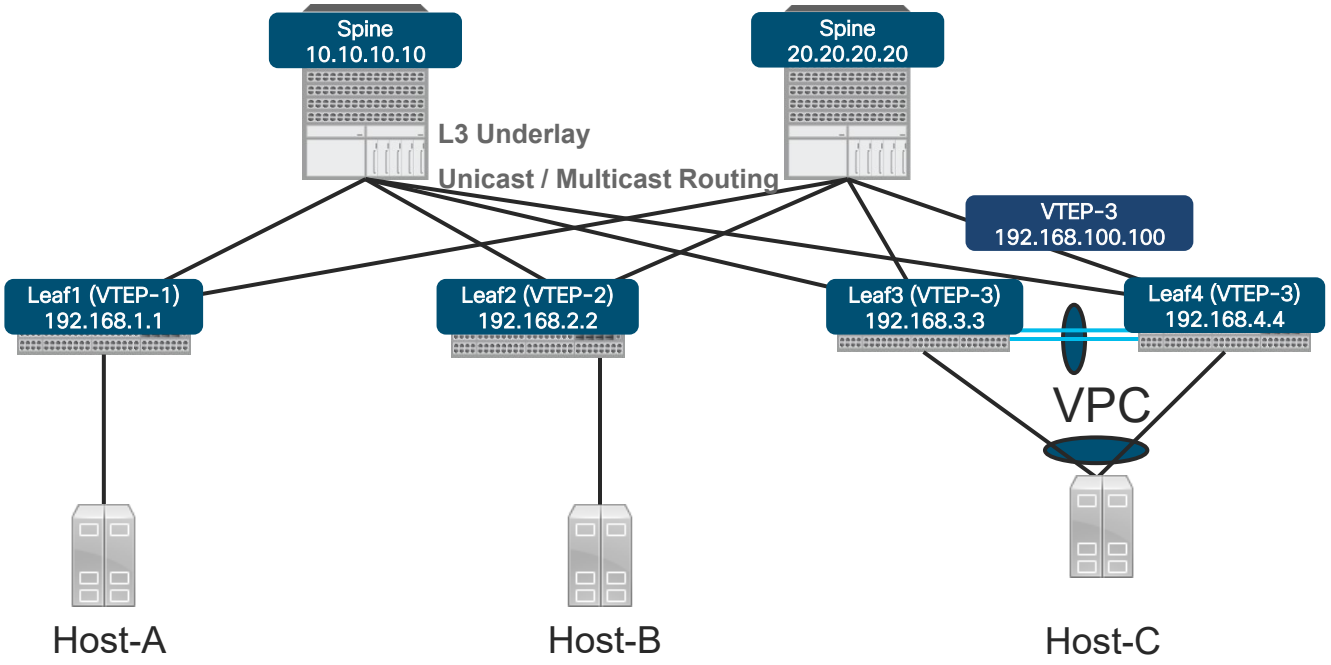
Configuration

VxLAN Configuration

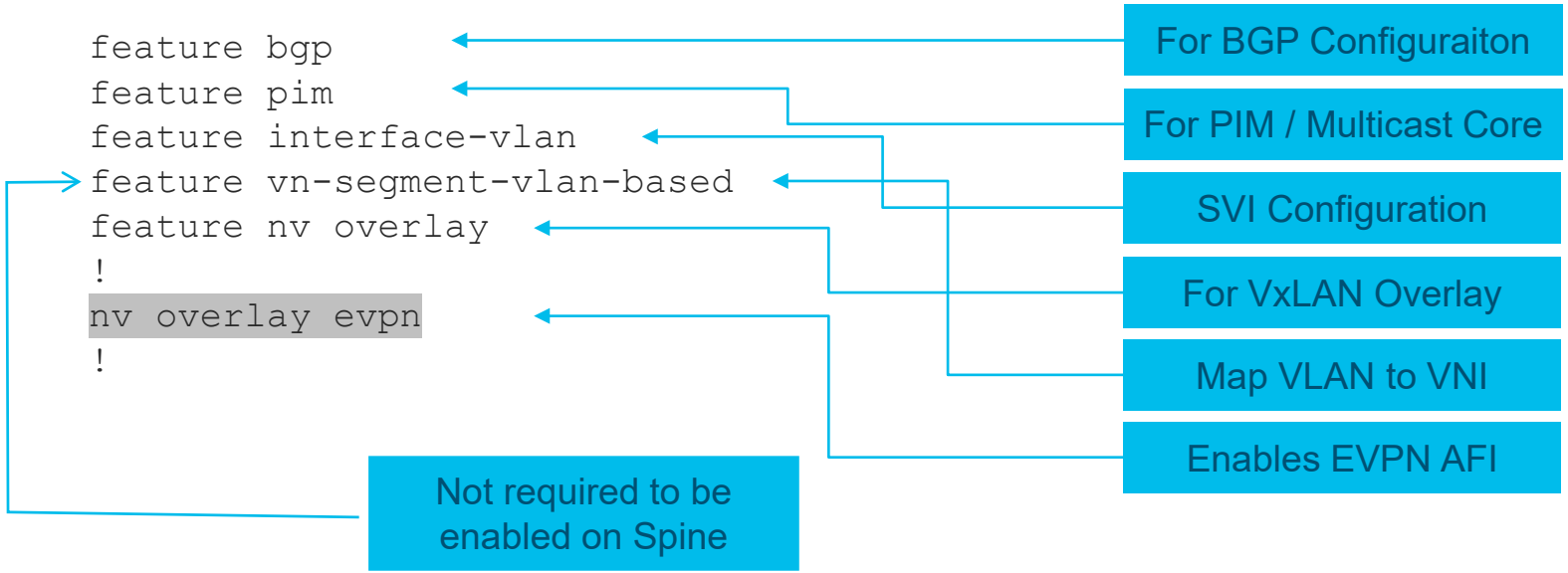
Topology

L2 VNI – 10000, 20000

L3 VNI - 50000



Configuration
Feature Enablement



Underlay Configuration

Leaf

```
router bgp 65000
  router-id 192.168.1.1
  address-family ipv4 unicast
    network 1.1.1.1/32
    network 192.168.1.1/32
  address-family l2vpn evpn
  neighbor 10.1.101.10
    remote-as 65001
    address-family ipv4 unicast
      allowas-in 3
      disable-peer-as-check
  neighbor 10.1.201.20
    remote-as 65001
    address-family ipv4 unicast
      allowas-in 3
      disable-peer-as-check
```

Spine

```
router bgp 65001
  router-id 192.168.10.10
  address-family ipv4 unicast
    network 10.10.10.10/32
    network 192.168.10.10/32
  address-family l2vpn evpn
  nexthop route-map permit-all
  retain route-target-all
  neighbor 10.1.201.1
    remote-as 65000
    address-family ipv4 unicast
      allowas-in 3
      disable-peer-as-check
  neighbor 10.1.202.2
    remote-as 65000
    address-family ipv4 unicast
      allowas-in 3
      disable-peer-as-check
```

Leaf Node Configuration – L2 VNI

```
vlan 100
  vn-segment 10000

! Create L2 VNI

evpn
  vni 10000 l2
    rd 10000:1
    route-target import 10000:1
    route-target export 10000:1
!
interface nve1
  no shutdown
  source-interface loopback0
  host-reachability protocol bgp
  member vni 10000
    ingress-replication protocol bgp
```

```
fabric forwarding anycast-gateway-mac
0001.0001.0001
interface Vlan100
  no shutdown
  vrf member EVPN-TENANT
  ip address 100.1.1.254/24
  fabric forwarding mode anycast-gateway
!
router bgp 65000
neighbor 10.10.10.10
    remote-as 65001
    update-source loopback0
    ebgp-multihop 3
    address-family l2vpn evpn
      allowas-in 3
      disable-peer-as-check
      send-community extended
  vrf EVPN-TENANT
    address-family ipv4 unicast
      advertise l2vpn evpn
```

Leaf Node Configuration – L3 VNI

```
vlan 500
  vn-segment 50000
!
vrf context EVPN-TENANT
  vni 50000
  rd 50000:1
  address-family ipv4 unicast
    route-target import 20000:1
    route-target import 20000:1 evpn
    route-target export 20000:1
    route-target export 20000:1 evpn
!
interface Vlan500
  no shutdown
  vrf member EVPN-TENANT
  ip forward
!
```

```
interface nve1
  no shutdown
  source-interface loopback0
  host-reachability protocol bgp
   member vni 50000 associate-vrf
!
interface loopback200
  vrf member EVPN-TENANT
  ip address 200.1.1.1/32
!
router bgp 65000
  vrf EVPN-TENANT
    address-family ipv4 unicast
      network 200.1.1.1/32
      advertise l2vpn evpn
```

Leaf Node with VPC Configuration

```
vpc domain 10
  peer-switch
  peer-keepalive destination 10.1.34.4 source 10.1.34.3
  delay restore 60
  peer-gateway
  ipv6 nd synchronize
  ip arp synchronize
!
interface loopback0
  ip address 192.168.3.3/32
  ip address 192.168.100.100/32 secondary
```

VTEP IP. The secondary IP is same on both Leaf3 and Leaf 4 running VPC

Backup Routing SVI

```
vlan 5
interface vlan 5
 ip add 10.5.1.1/24

 ip pim sparse-mode
```

Backup Routing SVI
Configured on both vPC peers and part of global routing table.

PIM on backup routing SVI is only required for multicast enabled core, not with IR

Spine Node Configuration

```
interface loopback0
  ip address 192.168.10.10/32
  ip pim sparse-mode
!
interface loopback1
  ip address 192.168.50.50/32
  ip pim sparse-mode
!
ip pim rp-add 192.168.50.50 group-list
239.1.1.0/24
ip pim anycast-rp 192.168.50.50 192.168.10.10
ip pim anycast-rp 192.168.50.50 192.168.20.20
```

```
router bgp 65001
address-family l2vpn evpn
    nexthop route-map permitall
    retain route-target all
  neighbor 1.1.1.1
    remote-as 65000
    update-source loopback0
    ebgp-multihop 3
    address-family l2vpn evpn
      disable-peer-as-check
      send-community extended
      route-map permitall out
  neighbor 2.2.2.2
    remote-as 65000
    update-source loopback0
    ebgp-multihop 3
    address-family l2vpn evpn
      disable-peer-as-check
      send-community extended
      route-map permitall out
!
route-map permitall permit 10
  set ip next-hop unchanged
```

cisco Live!

Leaf Node Configuration – L2 & L3 VNI for IPv6

```
vlan 100
  vn-segment 10000
vlan 200
  vn-segment 20000
evpn
  vni 10000 l2
    rd 10000:1
    route-target import 10000:1
    route-target export 10000:1
!
vrf context EVPN-TENANT
  vni 20000
  rd 20000:1
address-family ipv4 unicast
    route-target both 20000:1
    route-target both 20000:1 evpn
address-family ipv6 unicast
    route-target both 20000:1
    route-target both 20000:1 evpn
!
interface Vlan200
  no shutdown
  vrf member EVPN-TENANT
  ip forward
  ipv6 address use-link-local-only
```

```
fabric forwarding anycast-gateway-mac 0001.0001.0001
interface Vlan100
  no shutdown
  vrf member EVPN-TENANT
  ip address 100.1.1.254/24
  ipv6 address 2001::1/64
  fabric forwarding mode anycast-gateway
!
interface nve1
  no shutdown
  source-interface loopback0
  host-reachability protocol bgp
  member vni 10000
    mcast-group 239.1.1.1
    suppress-arp
member vni 20000 associate-vrf
!
router bgp 100
vrf EVPN-TENANT
    address-family ipv4 unicast
      advertise l2vpn evpn
    address-family ipv6 unicast
      advertise l2vpn evpn
!
vpc domain 10
   ipv6 nd synchronize
```

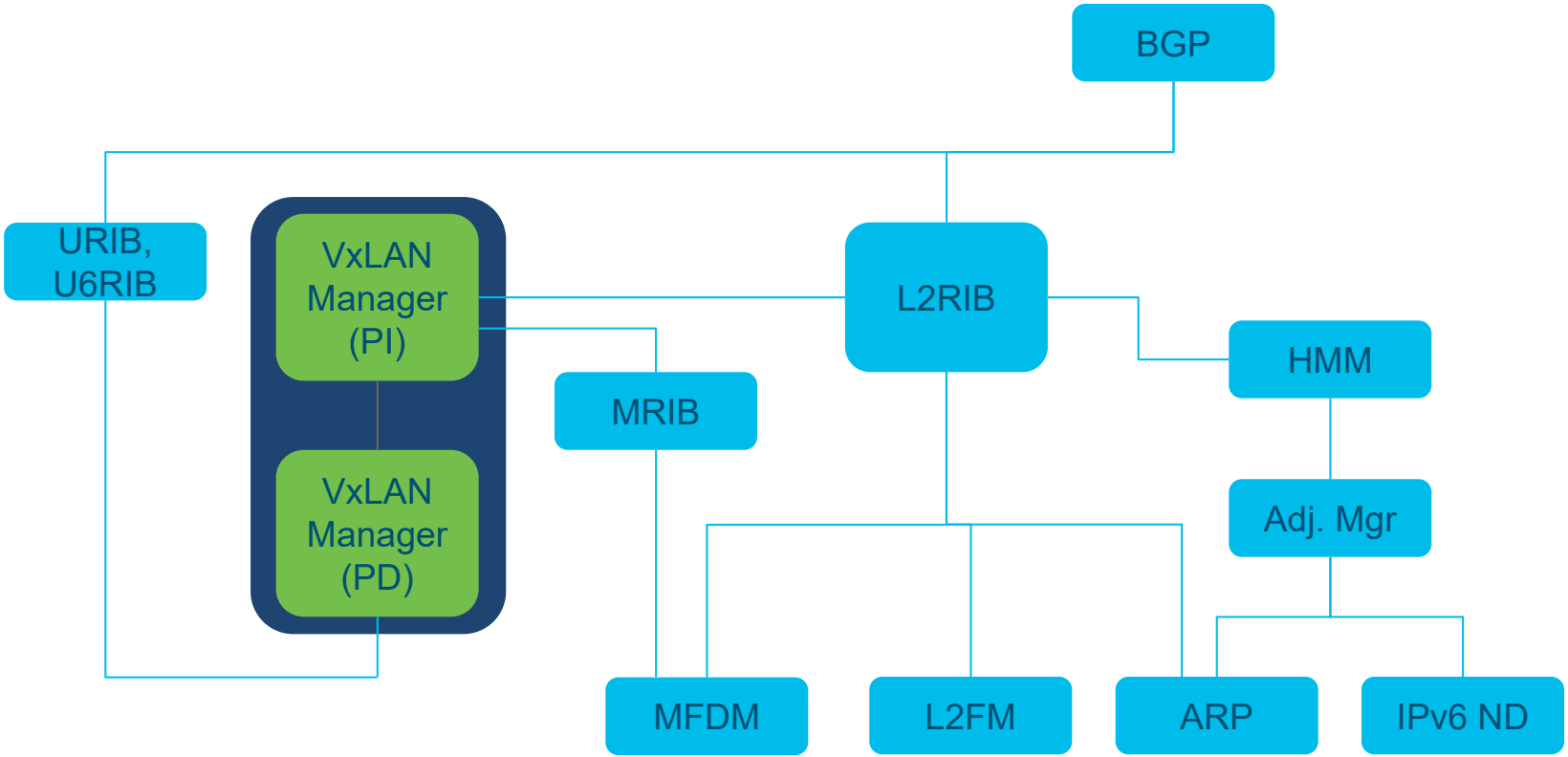
VxLAN EVPN Configuration

Host Learning and Peer Discovery

Host Learning	Data Plane	Control Plane
CORE		
Multicast	Flood and Learn Peer Learning: Data Plane Vlan 100 vn-segment 10000 Interface nve1 Member vni 10000 Mcast-group 239.1.1.1	EVPN-Multicast Peer Learning: BGP-RnH Vlan 100 vn-segment 10000 Interface nve1 host-reachability protocol bgp member vni 10000 Mcast-group 239.1.1.1
Unicast	Static Ingress-Replication Peer Learning - CLI Vlan 150 vn-segment 15000 Interface nve1 member vni 15000 Ingress-replication protocol static peer x.x.x.x	EVPN Ingress-Replication Peer Learning – BGP-IMET Vlan 150 vn-segment 15000 Interface nve1 host-reachability protocol bgp member vni 15000 ingress-replication protocol bgp

VXLAN BGP EVPN Control-Plane Verification

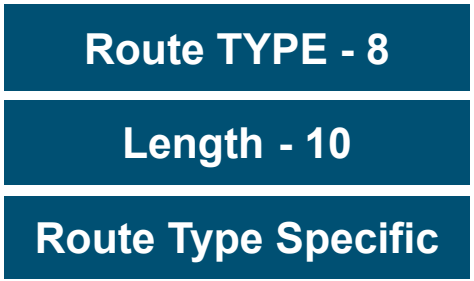
Nexus 9000 VxLAN Architecture



Troubleshooting VxLAN EVPN

EVPN Prefix Types

- BGP EVPN uses 5 different route types for IP prefixes and advertisement
  - Type 1 – Ethernet Auto-Discovery (A-D) route
  - **Type 2 – MAC advertisement route → L2 VNI MAC/MAC-IP**
  - Type 3 – Inclusive Multicast Route → **EVPN IR, Peer Discovery**
  - Type 4 – Ethernet Segment Route
  - **Type 5 – IP Prefix Route → L3 VNI Route**
- Route type 2 or MAC Advertisement route is for MAC and ARP resolution advertisement, MAC or MAC-IP
- Route type 5 or IP Prefix route will be used for the advertisement of prefixes, IP only



Troubleshooting VxLAN EVPN

VxLAN EVPN Prefix Types

```
Leaf1# show bgp l2vpn evpn
   Network            Next Hop            Metric     LocPrf     Weight Path
Route Distinguisher: 10000:1    (L2VNI 10000)
*>l[2]:[0]:[0]:[48]:[523d.e706.ae1b]:[0]:[0.0.0.0]/216
                     192.168.1.1                           100      32768 i
*>l[2]:[0]:[0]:[48]:[523d.e706.ae1b]:[32]:[100.1.1.1]/272
                     192.168.1.1                           100      32768 i
*>l[3]:[0]:[32]:[192.168.1.1]/88
                     192.168.1.1                           100      32768 i
*>e[3]:[0]:[32]:[192.168.2.2]/88
                     192.168.2.2                                        0 65001 65000 i
*>e[3]:[0]:[32]:[192.168.3.3]/88
                     192.168.3.3                                        0 65001 65000 I
Route Distinguisher: 192.168.2.2:32967
* e[3]:[0]:[32]:[192.168.2.2]/88
                     192.168.2.2                                        0 65001 65000 i
*>e                  192.168.2.2                                        0 65001 65000 i

Route Distinguisher: 192.168.1.1:3    (L3VNI 50000)
*>e[5]:[0]:[0]:[32]:[100.100.100.2]:[0.0.0.0]/224
                     192.168.2.2                                        0 65001 65000 i
```

Troubleshooting VxLAN EVPN

NVE Interface

```
Leaf1# show nve interface
Interface: nve1, State: Up, encapsulation: VXLAN
 VPC Capability: VPC-VIP-Only [not-notified]
 Local Router MAC: f40f.1b6f.926f
 Host Learning Mode: Control-Plane
 Source-Interface: loopback0 (primary: 192.168.1.1, secondary: 0.0.0.0)
```

```
Leaf1# show interface nve1
nve1 is up
admin state is up,  Hardware: NVE
  MTU 9216 bytes
  Encapsulation VXLAN
  Auto-mdix is turned off
  RX
    ucast: 40 pkts, 5400 bytes - mcast: 1 pkts, 118 bytes
  TX
    ucast: 54 pkts, 6256 bytes - mcast: 9 pkts, 1026 bytes
```

If NVE Interface status is down, ensure that a no shut is performed on the interface.

Troubleshooting VxLAN EVPN

Local MAC Routes Learning

Mac Learnt on Vlan 100

```
Leaf1#show mac address-table vlan 100
        (T) - True, (F) - False
   VLAN     MAC Address       Type       age     Secure NTFY Ports
---------+-----------------+--------+---------+------+----+-------------
*  100     523d.e706.ae1b    dynamic  0         F      F    Eth1/15
```

```
Leaf1# sh sys inter l2fm event-hist deb | in 523d.e706.ae1b
  [104] l2fm_l2rib_add_delete_local_mac_routes(1095): To L2RIB: topo-id: 100,
macaddr: 523d.e706.ae1b, nhifindx: 0x1a001600 peer_addr 0x1a001600
  [104] l2fm_macdb_insert(6360): slot 0 fe 0 mac 523d.e706.ae1b vlan 100 flags
0x400107 hints 0 E8 NL lc  : if_index 0x1a001600 old_if_index 0
```

Troubleshooting VxLAN EVPN

L2FM installs the MAC in the L2RIB


```
Leaf1#show l2route evpn mac evi 100
Mac Address    Prod   Next Hop (s)
-------------- ------ --------------
523d.e706.ae1b Local  Eth1/15
```

```
Leaf1# show system internal l2rib event-history mac | in 523d.e706.ae1b
[06/01/16 22:31:55.201 UTC 5 9954] Received MAC ROUTE msg: addr: (100, 523d.e706.ae1b) vni: 0
admin_dist: 0 seq_num: 0 rt_flags: L soo: 0 dg_count: 0 res: 0 esi: (F) nh_count: 1
[06/01/16 22:31:55.202 UTC 7 9954] (100,8c60.4f93.5ffc):Mobility check for new rte from prod:
3
[06/01/16 22:31:55.202 UTC 8 9954] (100,523d.e706.ae1b):Current non-del-pending route
local:no, remote:no, linked mac-ip count:1
[06/01/16 22:31:55.202 UTC 9 9954] (523d.e706.ae1b,3):MAC route created with seq num:0,
flags:L (), soo:0, peerid:0
[06/01/16 22:31:55.205 UTC a 9954] (100,523d.e706.ae1b,3):Encoding MAC best route (ADD,
client id 5)
[06/01/16 22:31:55.207 UTC 3 9954] (100,523d.e706.ae1b):Bound MAC-IP(100.1.1.1) to MAC, Total
MAC-IP linked: 1
```

Troubleshooting VxLAN EVPN

L2 VNI, MAC


```
Leaf1#show bgp l2vpn evpn vni-id 10000
   Network            Next Hop            Metric     LocPrf     Weight Path
Route Distinguisher: 10000:1    (L2VNI 10000)
*>l[2]:[0]:[0]:[48]:[523d.e706.ae1b]:[0]:[0.0.0.0]/216
                      192.168.1.1                          100      32768 i
```

```
Leaf1# show bgp internal event-history events | in 523d.e706.ae1b
 2016 Jun  1 22:31:55.205989 bgp 100 [16855]: [16888]: (default) RIB: [L2VPN EVPN
] add prefix 10000:1:[2]:[0]:[0]:[48]:[523d.e706.ae1b]:[0]:[0.0.0.0] (flags 0x1)
 : OK, total 2
2016 Jun  1 22:31:55.205655 bgp 100 [16855]: [16888]: EVT: Received from L2RIB MAC
route: Add topo 10000 mac 523d.e706.ae1b soo 0 seq 0
. . .
```

Troubleshooting VxLAN EVPN

Local MAC Address in BGP L2VPN


```
Leaf1# show bgp l2vpn evpn 523d.e706.ae1b
BGP routing table information for VRF default, address family L2VPN EVPN
Route Distinguisher: 192.168.1.1:32867    (L2VNI 10000)
BGP routing table entry for
[2]:[0]:[0]:[48]:[523d.e706.ae1b]:[0]:[0.0.0.0]/216, version 318
Paths: (1 available, best #1)
Flags: (0x000102) (high32 00000000) on xmit-list, is not in l2rib/evpn
  Advertised path-id 1
  Path type: local, path is valid, is best path
  AS-Path: NONE, path locally originated
    192.168.1.1 (metric 0) from 0.0.0.0 (192.168.1.1)
      Origin IGP, MED not set, localpref 100, weight 32768
      Received label 10000
      Extcommunity: RT:65000:10000 ENCAP:8
  Path-id 1 advertised to peers:
    10.10.10.10          20.20.20.20
```

Troubleshooting VxLAN EVPN

Remote L2 MAC Route Installation via BGP EVPN

BGP L2VPN

```
Leaf2# show bgp l2vpn evpn 523d.e706.ae1b
BGP routing table information for VRF default, address family L2VPN EVPN
Route Distinguisher: 192.168.1.1:32867
BGP routing table entry for [2]:[0]:[0]:[48]:[523d.e706.ae1b]:[0]:[0.0.0.0]/216,
Paths: (2 available, best #1)
Flags: (0x000202) (high32 00000000) on xmit-list, is not in l2rib/evpn, is not in HW

  Advertised path-id 1
  Path type: external, path is valid, is best path
             Imported to 1 destination(s)
  AS-Path: 65001 65000 , path sourced external to AS
    192.168.1.1 (metric 0) from 20.20.20.20 (192.168.20.20)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 10000
      Extcommunity: RT:65000:10000 ENCAP:8

  Path type: external, path is valid, not best reason: newer EBGP path
  AS-Path: 65001 65000 , path sourced external to AS
    192.168.1.1 (metric 0) from 10.10.10.10 (192.168.10.10)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 10000
      Extcommunity: RT:65000:10000 ENCAP:8
```

EVPN BGP Route Type 2 Fields

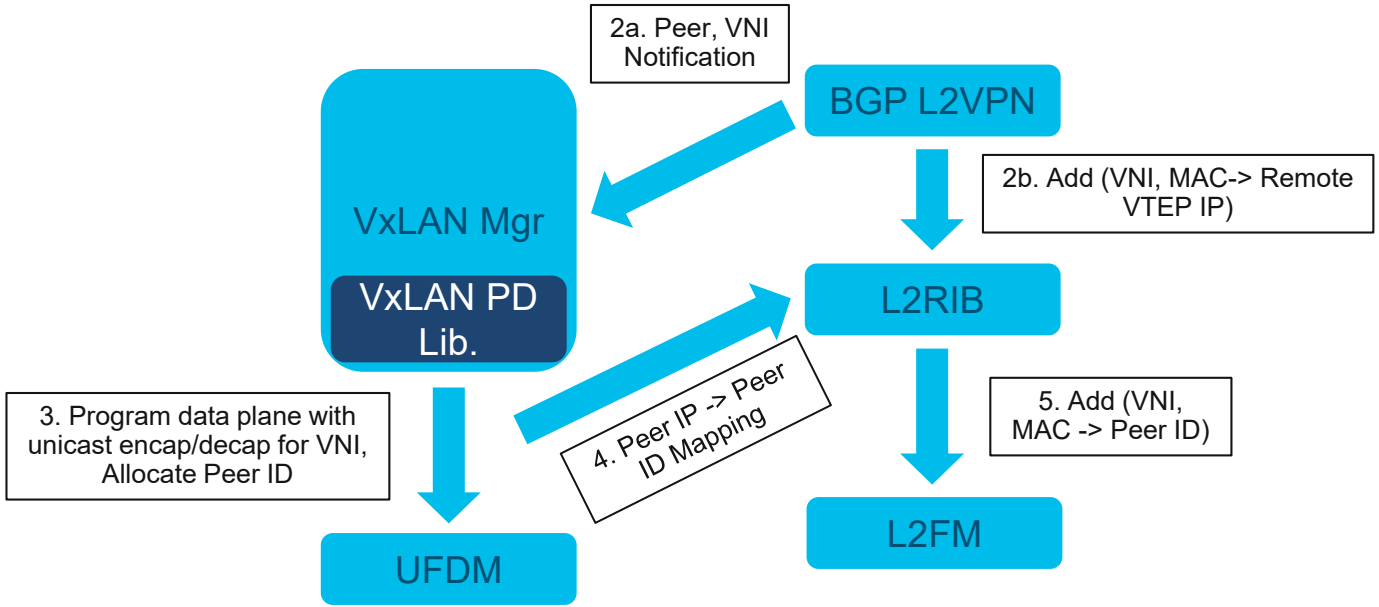
- Ethernet Tag ID, MAC Address Length, MAC Address, IP Address Length, and IP Address fields are considered to be part of the prefix in the NLRI.
- Ethernet Segment Identifier, MPLS Label1, and MPLS Label2 are treated as route attributes, not part of the "route". Both the IP and MAC address lengths are in bits.

```
Leaf1#show bgp l2vpn evpn 523d.e706.ae1b
BGP routing table information for VRF default, address family L2VPN
EVPN
Route Distinguisher: 10000:1     (L2VNI 10000)
BGP routing table entry for
[2]:[0]:[0]:[48]:[8c60.4f93.5ffc]:[0]:[0.0.0.0]/216, version 8
Paths: (1 available, best #1)
Flags: (0x00010a) on xmit-list, is not in l2rib/evpn

  Advertised path-id 1
  Path type: local, path is valid, is best path, no labeled nexthop
  AS-Path: NONE, path locally originated
    192.168.1.1 (metric 0) from 0.0.0.0 (192.168.1.1)
      Origin IGP, MED not set, localpref 100, weight 32768
      Received label 10000
      Extcommunity:  RT:10000:1
```

- Route Distinguisher
- Ethernet Segment ID – 10 byte
- Ethernet Tag ID – 4 byte
- MAC Address Length – 1 byte
- MAC Address – 6 byte
- IP Address Length – 1 byte
- IP Address – 0, 4, 16 byte
- MPLS Label 1 – 3 byte, L2VNI
- MPLS Label 2 – 3 byte L3VNI

Troubleshooting VxLAN EVPN

Remote L2 MAC Route Installation with BGP EPVN (Flow)



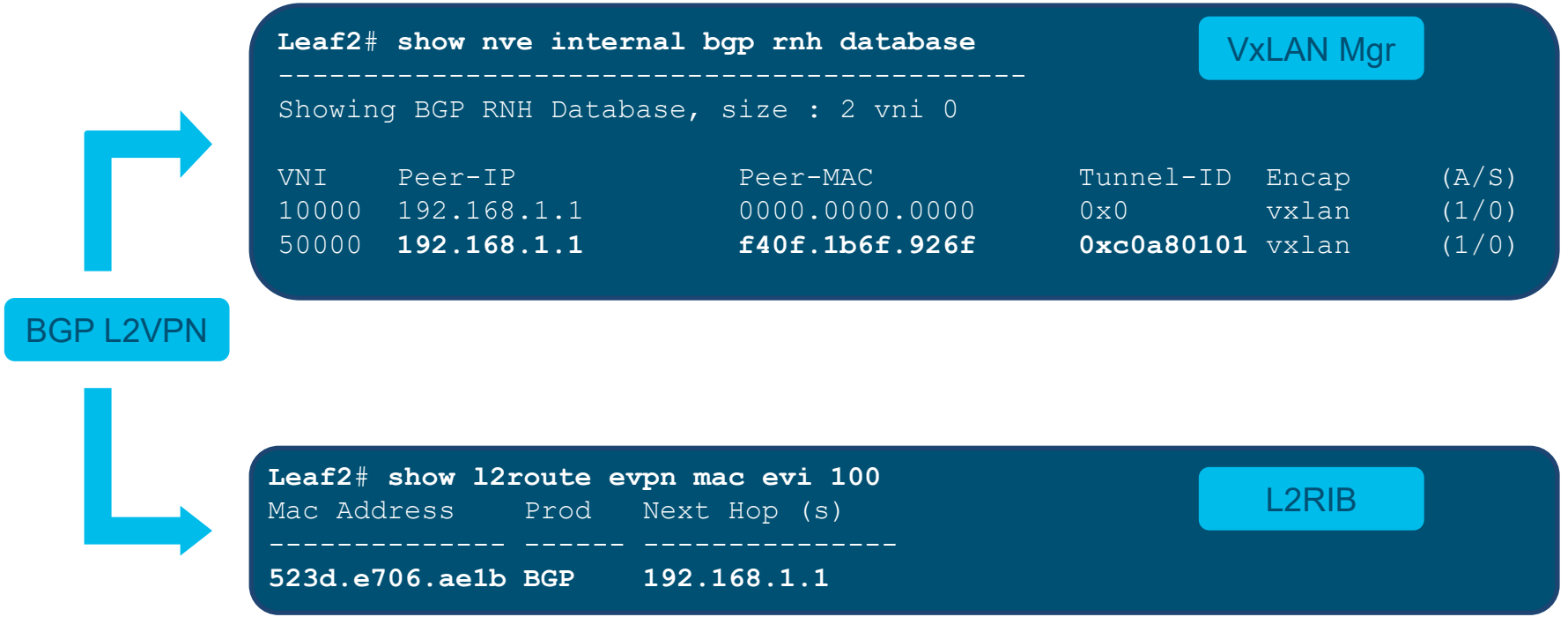
Troubleshooting VxLAN EVPN

BGP to L2RIB

```
Leaf2# show bgp internal event-history events | in 523d.e706.ae1b
2016 Jun  2 02:53:14.844179 bgp 100 [9878]: [9890]: (default) IMP: bgp_tbl_ctx_import:
1812: [L2VPN EVPN] Importing
10000:1:[2]:[0]:[0]:[48]:[523d.e706.ae1b]:[0]:[0.0.0.0]/112 to RD 10000:1
2016 Jun  2 02:53:14.844167 bgp 100 [9878]: [9890]: (default) IMP: bgp_vrf_import:
2740: vrf default 10000:1:[2]:[0]:[0]:[48]:[523d.e706.ae1b]:[0]:[0.0.0.0]/112 result 1
2016 Jun  2 02:53:14.844130 bgp 100 [9878]: [9890]: (default) RIB: [L2VPN EVPN]: Send
to L2RIB 10000:1:[2]:[0]:[0]:[48]:[523d.e706.ae1b]:[0]:[0.0.0.0]/112 via 192.168.1.1
Add 1 EVPN MAC routes succeeded
```

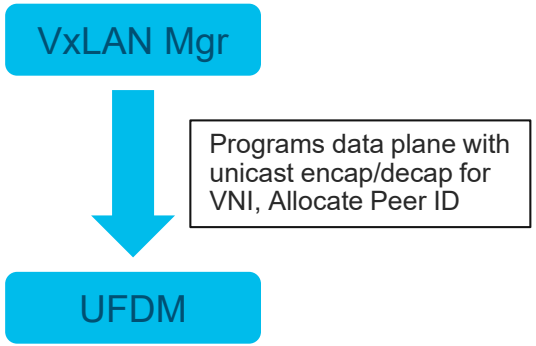
Troubleshooting VxLAN EVPN

Remote L2 MAC Route Installation via BGP EVPN



VxLAN Manager

```
Leaf2# show forwarding nve l3 peers
NVE cleanup transaction-id 0
tunnel_id  Peer_id  Peer_address    Interface     rmac      origin   state      del count
-------------------------------------------------------------------------------------
0xc0a80101   1      192.168.1.1       nve1    f40f.1b6f.926f  NVE    merge-done no    1
```




```
Leaf2# show nve peers detail
Details of nve Peers:
---------------------------------------
Peer-Ip: 192.168.1.1
    NVE Interface        : nve1
    Peer State           : Up
    Peer Uptime          : 01:27:30
    Router-Mac           : f40f.1b6f.926f
    Peer First VNI       : 20000
    Time since Create    : 01:27:30
    Configured VNIs      : 10000,20000
    Provision State      : add-complete
    Route-Update         : Yes
    Peer Flags           : RmacL2Rib, TunnelPD, DisableLearn
    Learnt CP VNIs       : 10000,20000,50000
    Peer-ifindex-resp    : Yes
```

Hardware Programmed

Troubleshooting VxLAN EVPN

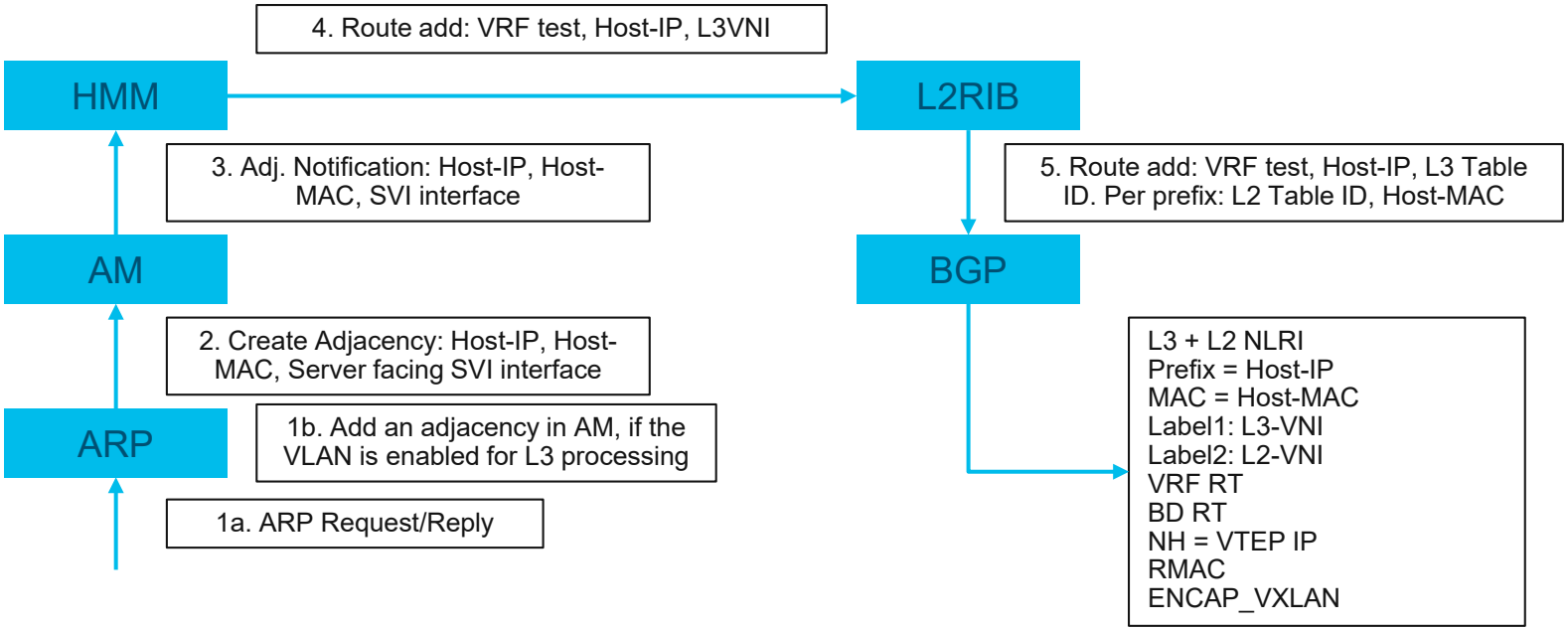
L2FM Verification

```
Leaf2# show system internal l2fm debugs | in 523d.e706.ae1b
    [104] l2fm_macdb_insert(6327): slot 32 fe 0 mac 8c60.4f1b.e43c vlan 100 flags 0x7
hints 0 E8 NL lc  : if_index 0x49080001 old_if_index 0
    [104] l2fm_l2rib_mac_update(21832): Add L2RIB remote mac 523d.e706.ae1b
    [104] l2fm_process_l2rib_remote_route_update(405): Type: 2 Len: 152 Seq: 0, del: 0
(Prod: 5) Flags: Ctrl=3 Rt=0, mac 8c60.4f1b.e43c topo_id 100
```

```
Leaf2# show mac address-table vlan 100
VLAN     MAC Address      Type      age     Secure NTFY Ports
---------+----------------+--------+---------+------+----+------------------
-
*  100     8c60.4f19.51fc   dynamic  0         F      F    Eth1/13
*  100     523d.e706.ae1b   dynamic  0         F      F    nve1(192.168.1.1)
```

Troubleshooting VxLAN EVPN

Host IP and Host MAC local route



Troubleshoot VxLAN EVPN

ARP from Host and ARP -> AM

```
Leaf1# show ip arp vrf EVPN-TENANT
IP ARP Table for context EVPN-TENANT
Total number of entries: 1
Address         Age       MAC Address     Interface
100.1.1.1       00:10:47 523d.e706.ae1b  Vlan100
```

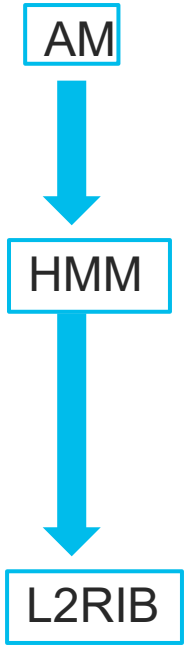
```
Leaf1# show forwarding vrf EVPN-TENANT adjacency
IPv4 adjacency information

next-hop         rewrite info    interface
-------------- --------------- -------------
100.1.1.1       523d.e706.ae1b  Vlan100
```

Troubleshoot VxLAN EVPN

AM -> HMM -> L2RIB

```
Leaf1# show ip route vrf EVPN-TENANT
100.1.1.1/32, ubest/mbest: 1/0, attached
    *via 100.1.1.1, Vlan100, [190/0], 02:41:57, hmm
100.1.1.254/32, ubest/mbest: 1/0, attached
    *via 100.1.1.254, Vlan100, [0/0], 02:59:46, local
```

```
Leaf1# show l2route evpn mac-ip evi 100
Mac Address    Prod Host IP                 Next Hop (s)
-------------- ---- --------------------- --------------
-
523d.e706.ae1b HMM  100.1.1.1                N/A
```



Troubleshooting VxLAN EVPN

L2RIB -> BGP

```
Leaf1# show bgp l2vpn evpn 100.1.1.1
BGP routing table information for VRF default, address family L2VPN EVPN
Route Distinguisher: 10000:1    (L2VNI 10000)
BGP routing table entry for [2]:[0]:[0]:[48]:[523d.e706.ae1b]:[32]:[100.1.1.1]/272,
version 6
Paths: (1 available, best #1)
Flags: (0x00010a) on xmit-list, is not in l2rib/evpn

  Advertised path-id 1
  Path type: local, path is valid, is best path, no labeled nexthop
  AS-Path: NONE, path locally originated
    192.168.1.1 (metric 0) from 0.0.0.0 (192.168.1.1)
      Origin IGP, MED not set, localpref 100, weight 32768
      Received label 10000 50000
      Extcommunity:  RT:10000:1 RT:50000:1

  Path-id 1 advertised to peers:
    10.10.10.10
```

Troubleshooting VxLAN EVPN

Remote Host Prefix - EVPN

```
Leaf2# show bgp l2vpn evpn 100.1.1.1
BGP routing table information for VRF default, address family L2VPN EVPN
Route Distinguisher: 10000:1    (L2VNI 10000)
BGP routing table entry for [2]:[0]:[0]:[48]:[523d.e706.ae1b]:[32]:[100.1.1.1]/272,
version 5
Paths: (1 available, best #1)
Flags: (0x00021a) on xmit-list, is in l2rib/evpn, is not in HW, , is locked
  Advertised path-id 1
  Path type: internal, path is valid, imported same remote RD, is best path, no labeled
nexthop
  AS-Path: NONE, path sourced internal to AS
    192.168.1.1 (metric 5) from 10.10.10.10 (192.168.10.10)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 10000 50000
      Extcommunity:  RT:10000:1 RT:50000:1 ENCAP:8 Router MAC:f40f.1b6f.926f
      Originator: 192.168.1.1 Cluster list: 10.10.10.10

  Path-id 1 not advertised to any peer
. . .
```

Troubleshooting VxLAN EVPN L2-L3 Remote Route Installation

L2RIB and URIB Information

```
Leaf2# show l2route evpn mac-ip evi 100
Mac Address    Prod Host IP                                 Next Hop (s)
-------------- ---- --------------------------------- ---------------
523d.e706.ae1b BGP  100.1.1.1                            192.168.1.1
```

```
Leaf2# show ip route vrf EVPN-TENANT 100.1.1.1
IP Route Table for VRF "EVPN-TENANT"
100.1.1.1/32, ubest/mbest: 1/0
    *via 192.168.1.1%default, [200/0], 04:00:28, bgp-100, internal,
tag 100 (evpn) segid: 50000 tunnelid: 0xc0a80101 encap: VXLAN
```

Troubleshooting VxLAN EVPN

NVE Internal Information for Leaf Nodes without vPC

```
Leaf3# show nve internal platform interface nve 1 detail
Printing Interface ifindex 0x49000001 detail
|======|=========================|===============|===============|=====|=====|
|Intf  |State                    |PriIP          |SecIP          |Vnis |Peers|
|======|=========================|===============|===============|=====|=====|
|nve1  |UP                       |192.168.3.3    |0.0.0.0        |2    |2    |
|======|=========================|===============|===============|=====|=====|
SW_BD/VNIs of interface nve1:
================================================
|======|======|=========================|======|====|======|========
|Sw BD |Vni   |State                    |Intf  |Type|Vrf-ID|Notified
|======|======|=========================|======|====|======|========
|100   |10000 |UP                       |nve1  |CP  |0     |Yes
|200   |20000 |UP                       |nve1  |CP  |3     |Yes
|======|======|=========================|======|====|======|========

Peers of interface nve1:
========================================
Peer_ip: 192.168.1.1
  Peer-ID   : 2
  State     : UP
  Learning  : Disabled
  TunnelID  : 0xc0a80101
  MAC       : f40f.1b6f.926f
  Table-ID  : 0x1
  Encap     : 0x1
```

Troubleshooting VxLAN EVPN

NVE Internal Information for Leaf Nodes with VPC Peers

```
Leaf3# show nve internal platform interface nve 1 detail
Printing Interface ifindex 0x49000001 detail
|======|=========================|===============|===============|=====|=====|
|Intf  |State                    |PriIP          |SecIP          |Vnis |Peers|
|======|=========================|===============|===============|=====|=====|
|nve1  |UP                       |192.168.3.3    |192.168.100.100|2    |2    |
|======|=========================|===============|===============|=====|=====|
SW_BD/VNIs of interface nve1:
================================================
|======|======|=========================|======|====|======|========
|Sw BD |Vni   |State                    |Intf  |Type|Vrf-ID|Notified
|======|======|=========================|======|====|======|========
|100   |10000 |UP                       |nve1  |CP  |0     |Yes
|200   |20000 |UP                       |nve1  |CP  |3     |Yes
|======|======|=========================|======|====|======|========

Peers of interface nve1:
========================================
Peer_ip: 192.168.1.1
  Peer-ID   : 2
  State     : UP
  Learning  : Disabled
  TunnelID  : 0xc0a80101
  MAC       : f40f.1b6f.926f
  Table-ID  : 0x1
  Encap     : 0x1
```

Troubleshooting VxLAN EVPN

NVE Internal Information for Leaf Nodes on Non-VPC Peers

```
Leaf1# show nve internal platform interface nve1 detail
Printing Interface ifindex 0x49000001 detail
|======|=========================|===============|===============|=====|=====|
|Intf  |State                    |PriIP          |SecIP          |Vnis |Peers|
|======|=========================|===============|===============|=====|=====|
|nve1  |UP                       |192.168.1.1    |0.0.0.0        |2    |2    |
|======|=========================|===============|===============|=====|=====|
SW_BD/VNIs of interface nve1:
================================================
|======|======|=========================|======|====|======|========
|Sw BD |Vni   |State                    |Intf  |Type|Vrf-ID|Notified
|======|======|=========================|======|====|======|========
|100   |10000 |UP                       |nve1  |CP  |0     |Yes
|200   |20000 |UP                       |nve1  |CP  |3     |Yes
|======|======|=========================|======|====|======|========

Peers of interface nve1:
==========================================
  Peer_ip: 192.168.100.100
  Peer-ID   : 2
  State     : UP
  Learning  : Disabled
  TunnelID  : 0xc0a86464
  MAC       : 88f0.312b.9e4d
  Table-ID  : 0x1
  Encap     : 0x1
```

Leaf Node with VPC Consistency Check

```
sh vpc consistency-parameters vni

    Legend:
        Type 1 : vPC will be suspended in case of mismatch

Name                        Type  Local Value            Peer Value
-------------               ----  ---------------------- -----------------------
-
Nve1 Adm St, Src Adm St,    1     Up, Up,                Up, Up,
Sec IP, Host Reach, VMAC          192.168.100.100, CP,   192.168.100.100, CP,
Adv, SA,mcast l2, mcast           FALSE, Disabled,       FALSE, Disabled,
l3, IR BGP,MS Adm St, Reo         0.0.0.0, 0.0.0.0,      0.0.0.0, 0.0.0.0,
                                  Disabled, Down,        Disabled, Down,
                                  0.0.0.0                0.0.0.0
Nve1 Vni, Mcast, Mode,      1     10011, 239.0.1.1,      10011, 239.0.1.1,
Type, Flags                       Mcast, L2, None        Mcast, L2, None
Nve1 Vni, Mcast, Mode,      1     10010, 239.0.1.0,      10010, 239.0.1.0,
Type, Flags                       Mcast, L2, None        Mcast, L2, None
Nve1 Vni, Mcast, Mode,      1     50000, 0.0.0.0, n/a,   50000, 0.0.0.0, n/a,
Type, Flags                       L3, L3VNI              L3, L3VNI
Allowed VLANs               -     1,5,9-12,99 ,100       1,5,10-12,99,100
Local suspended VLANs       -     9                      -
Leaf3#
```

Troubleshooting VxLAN EVPN

ARP Suppression Cache

```
Leaf1# show ip arp suppression-cache local
Ip Address      Age      Mac Address    Vlan Physical-ifindex    Flags

100.1.1.1       00:15:48 8c60.4f1b.e43c 100 Ethernet1/12         L
```

```
Leaf1# show ip arp suppression-cache remote
Ip Address      Age      Mac Address    Vlan Physical-ifindex    Flags

100.1.1.2     00:05:19   8c60.4f19.51fc 100 (null)              R
```

```
Leaf1(config)# hardware access-list tcam region arp-ether 256 double-wide
```

Troubleshooting VXLAN EVPN

Uplinks with SVI / Sub-Interfaces

- For SVI based uplinks, define the infra VLANs
  - **system nve infra-vlan** *<svi-vlan>*
- Sub-Interfaces
  - Not supported on ALE links (40G ports) (Documented)
  - Check CCO documentation
- L3 Port-channel supported – Check CCO documentation

https://www.cisco.com/c/en/us/support/docs/switches/nexus-9000-series-switches/214624-configure-system-nve-infra-vlans-in-vxla.html

Troubleshooting BUM Traffic

Troubleshooting BUM Traffic

BUM Traffic over Multicast Core

- BUM Traffic – Broadcast traffic (ARP and other broadcast packets), Multicast traffic from hosts, etc.
- Check the multicast group associated with the L2 VNI
- Get the Source VTEP IP and Router MAC
- Check if ARP Suppression is enabled or not
- Ask the right Questions:
  - Check if the VTEP is a VPC VTEP or Standalone VTEP?
  - Is the issue seen for IPv4 hosts or IPv6 hosts?
  - Know the trigger and understand if the issue is continuously reproducible or not?

Troubleshooting BUM Traffic

Step 1 – Check if the BUM Traffic is making it to the source VTEP

```
LEAF45#  ethanalyzer local interface inband capture-fil "arp" limit-captured-frames 0

Capturing on inband
2018-05-21 14:52:36.289960 20:20:00:00:00:aa -> 54:7f:ee:07:e1:41 ARP 10.150.1.254 is
at 20:20:00:00:00:aa
2018-05-21 14:52:36.295037 54:7f:ee:07:e1:41 -> ff:ff:ff:ff:ff:ff ARP Who has
10.150.1.35?  Tell 10.150.1.36
2018-05-21 14:52:36.295425 2c:54:2d:f6:0f:bc -> 54:7f:ee:07:e1:41 ARP 10.150.1.35 is at
2c:54:2d:f6:0f:bc
2018-05-21 14:52:38.127594 54:7f:ee:07:e1:41 -> ff:ff:ff:ff:ff:ff ARP Who has
10.150.1.38?  Tell 10.150.1.36
```

```
LEAF45#  debug logfile arp
LEAF45#  debug ip arp packet

14:52:38.127774 arp: (context 3) Receiving packet from Vlan1501, logical interface
Vlan1501 physical interface port-channel36, (prty 6) Hrd type 1 Prot type 800 Hrd len 6
Prot len 4 OP 1,  Pkt size 46
14:52:38.127811 arp: Src 547f.ee07.e141/10.150.1.36 Dst ffff.ffff.ffff/10.150.1.38
```

Troubleshooting BUM Traffic

Step 2 – Check the Mroute Entry – Src VTEP

```
LEAF45# show ip mroute 231.1.150.1 10.0.0.204 detail
IP Multicast Routing Table for VRF "default"
Total number of routes: 8
Total number of (*,G) routes: 2
Total number of (S,G) routes: 5
Total number of (*,G-prefix) routes: 1

(10.0.0.204/32, 231.1.150.1/32), uptime: 00:14:01, nve(0) mrib(0) ip(0) pim(1)
  RPF-Source: 10.0.0.204 [0/0]
  Data Created: No
  Received Register stop
  VXLAN Flags
    VXLAN Encap
  VPC Flags
    RPF-Source Forwarder
  Stats: 51/2601 [Packets/Bytes], 27.200  bps
  Stats: Active Flow
  Incoming interface: loopback1, RPF nbr: 10.0.0.204
  Outgoing interface list: (count: 1) (bridge-only: 0)
    Ethernet1/50, uptime: 00:09:52, pim
```

Troubleshooting BUM Traffic

Step 3.1 – Check the Mroute Entry – Dest VTEP

```
LEAF43# show ip mroute 231.1.150.1 10.0.0.204 detail
IP Multicast Routing Table for VRF "default"

Total number of routes: 8
Total number of (*,G) routes: 2
Total number of (S,G) routes: 5
Total number of (*,G-prefix) routes: 1

(10.0.0.204/32, 231.1.150.1/32), uptime: 00:03:19, ip(0) mrib(1) pim(0)
  RPF-Source: 10.0.0.204 [3/110]
  Data Created: Yes
  VXLAN Flags
    VXLAN Decap
  Stats: 1/51 [Packets/Bytes], 0.000   bps
  Stats: Inactive Flow
  Incoming interface: Ethernet1/50, RPF nbr: 10.0.0.42
  Outgoing interface list: (count: 1) (bridge-only: 0)
    nve1, uptime: 00:03:19, mrib
```

Troubleshooting BUM Traffic

Step 3.2 – Check the Mroute Entry – Dest VTEP

```
LEAF44# show ip mroute 231.1.150.1 10.0.0.204 detail
IP Multicast Routing Table for VRF "default"
Total number of routes: 8
Total number of (*,G) routes: 2
Total number of (S,G) routes: 5
Total number of (*,G-prefix) routes: 1

(10.0.0.204/32, 231.1.150.1/32), uptime: 00:03:57, ip(0) mrib(1) pim(0)
  RPF-Source: 10.0.0.204 [3/110]
  Data Created: Yes
  VXLAN Flags
    VXLAN Decap
  VPC Flags
    RPF-Source Forwarder
  Stats: 1/51 [Packets/Bytes], 0.000   bps
  Stats: Inactive Flow
  Incoming interface: Ethernet1/50, RPF nbr: 10.0.0.42
  Outgoing interface list: (count: 1) (bridge-only: 0)
    nve1, uptime: 00:03:57, mrib
```

Troubleshooting BUM Traffic

Step 4 – Capturing BUM Traffic in Core

```
LEAF45(config)# monitor session 1
LEAF45(config-monitor)# source interface ethernet 1/50
LEAF45(config-monitor)# destination interface sup-eth 0
LEAF45(config-monitor)# no shut
LEAF45(config-monitor)# end
```

```
LEAF45# ethanalyzer local interface inband capture-filter "host 231.1.150.1" limit-
captured-frames 0

Capturing on inband
2018-05-21 16:21:01.985236   10.0.0.204 -> 231.1.150.1  UDP Source port: 41316
Destination port: 4789
```

Use the **detail** option with **ethanalyzer** to see the whole packet

Demo

Troubleshooting Tenant Routed Multicast (TRM)

Tenant Routed Multicast (TRM)

Overview

- A BGP based solution for allowing multicast routing and snooping over VXLAN EVPN fabric
  - Sources and Receivers are connected to the VTEPs
- Multicast Source and Receiver info is propagated using BGP
  - No PIM/IGMP packets sent to VXLAN fabric from any TRM VTEP
- Modes:
  - L3 Mode
  - L2/L3 Mixed Mode
- Both modes are supported only on N9k – EX or FX platforms
- Supported only with Multicast based core. IR not supported

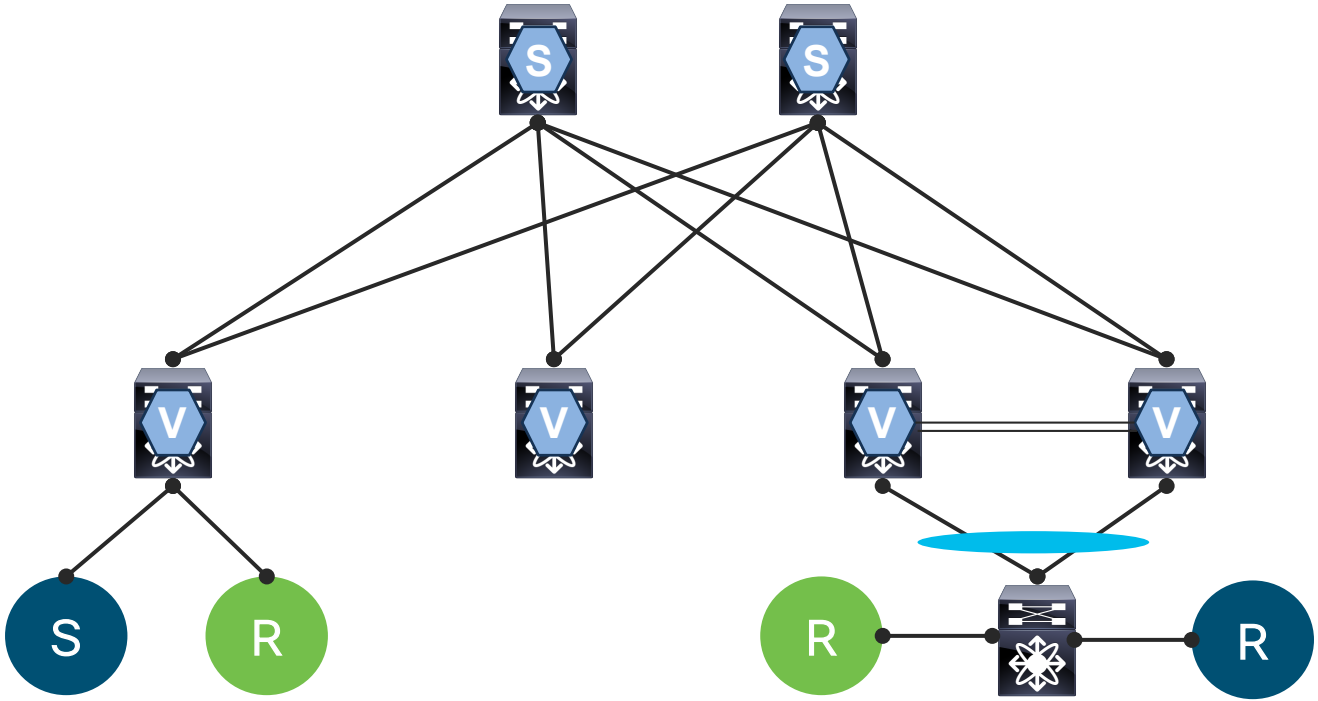
Tenant Routed Multicast (TRM)

Modes and RP

- L3 Mode
  - RP configured on all VTEPs – (**ip multicast overlay-spt-only** command required)
  - RP on selected VPC VTEP – not supported
  - Internal RP – Supported from 7.0(3)I7(1)
  - External RP – Supported from 7.0(3)I7(4) (upcoming release)
- L2/L3 Mode
  - RP on all Distributed-DR – Supported on Tahoe
  - RP on Core – not supported
  - Internal RP - Supported from 7.0(3)I7(1)
  - External RP – Not supported

Tenant Routed Multicast (TRM)

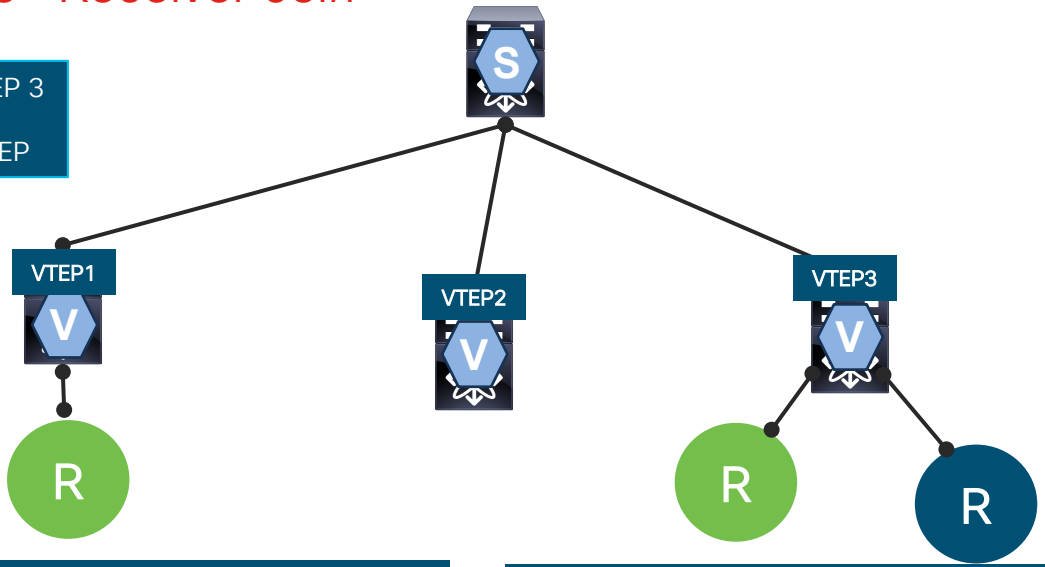
EVPN – L3 Anycast



Tenant Routed Multicast (TRM)

Packet flows –Receiver Join

Receiver on VTEP1 and VTEP 3
Join Group 239.0.0.1
(*, G) is created on both VTEP



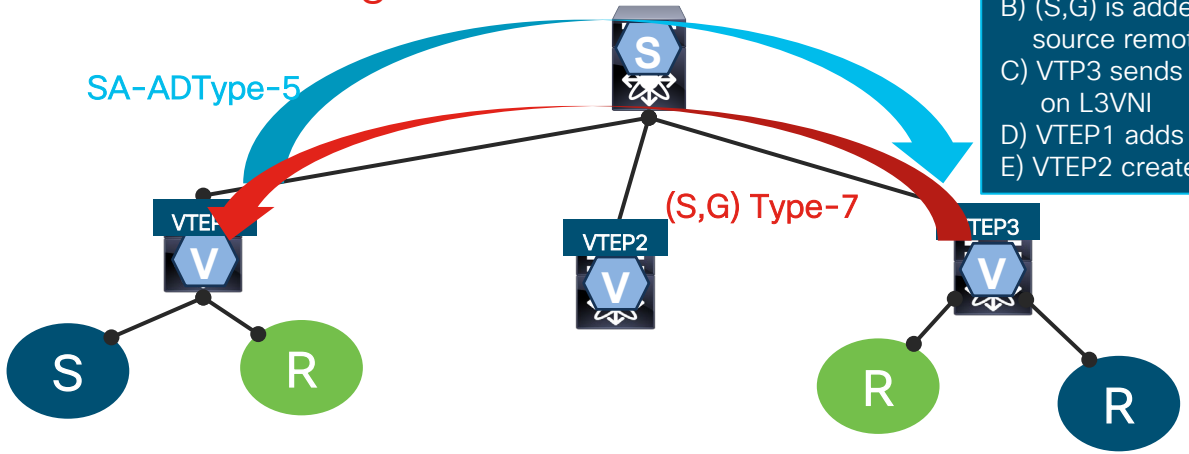
On VTEP1
(*,239.0.0.1) uptime 2d03h static ip pim igmp
Incoming interface loopback1000 , RPF nbr 209.165.200.1
Outgoing Interface list (count 1)
Vlan 11 , uptime 2d 03 h , igmp

On VTEP3
(*,239.0.0.1) uptime 2d03h static ip pim igmp
Incoming interface loopback1000 , RPF nbr 209.165.200.1
Outgoing Interface list (count 2)
Vlan 11 , uptime 2d 03 h , igmp
Vlan 12 , uptime 2d03h, static , igmp

Tenant Routed Multicast (TRM)

Packet flows –Learning Source

A) BGP Sends mvpn SA-AD (type-5) routes to other VTEP
B) (S,G) is added on VTEP3 with receiver local and source remote
C) VTP3 sends (S,G) join (type-7) back to FHR on L3VNI
D) VTEP1 adds L3VNI in oiflist
E) VTEP2 creates (S,G) with emty oif list

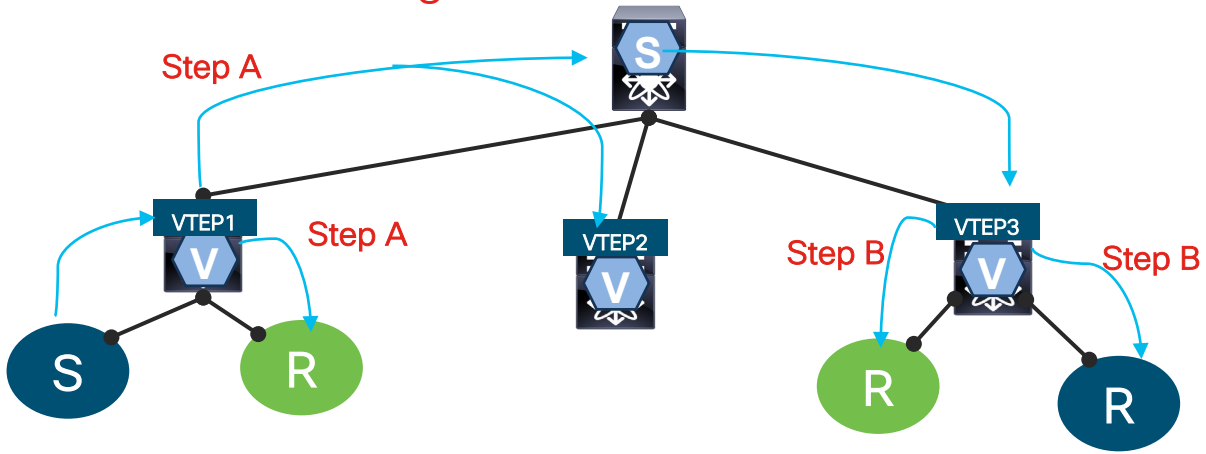


On VTEP1
(12.12.1.55/32,239.0.0.1) uptime 2d03h ip mrib pim ngmvpn
Incoming interface Vlan 12 , RPF nbr 12.12.1.55 internal
Outgoing Interface list (count 3) (Fabric OIF)
Vlan 11 , uptime 2d 03 h , mrib
Vlan 500 uptime 2d03h , ngmvpn Step D

On VTEP3
(12.12.1.55,239.0.0.1) uptime 05:29:18 mrib PIM ip
Incoming interface Vlan 500 , RPF nbr 209.165.200.1
Outgoing Interface list (count 2)
Vlan 11 , uptime 05:30:18 , mrib
Vlan 12 , uptime 05:31:10 mrib

Tenant Routed Multicast (TRM)

Packet flows –Forwarding



A) VTEP 1 Sends VxLan encaped packet with SIP=VTP1-IP DIP=225.3.3.3 (L3VNI default delivery Group) to the fabric and route packet locally for local receivers
B) VTEP3 decaps packet received from VTEP1 and routes the packet to local receivers in respective Vlans
C) VTEP2 also receives encap packet which it decaps and drop it since it does not have any receivers

Configuring Layer 3 Tenant Routed Multicast

```
feature ngmvpn
Ip igmp snooping vxlan
ip multicast overlay-spt-only
advertise evpn multicast
interface nve1
  no shutdown
  source-interface loopback0
  host-reachability protocol bgp
   member vni 50000 associate-vrf
   mcast-group 225.3.3.3
router bgp 65000
  vrf EVPN-TENANT
    address-family ipv4 unicast
      network 200.1.1.1/32
      advertise l2vpn evpn
      address-family ipv4 mvpn
```

```
interface loopback1000
  vrf member EVPN-TENANT
  ip address 209.165.200.1/32
   Ip pim sparse-mode
Vrf context EVPN-TENANT
Ip pim rp-address 209.165.200.1 group-lost 224.0.0.0/4
interface Vlan500
  no shutdown
  vrf member EVPN-TENANT
  ip forward
 ip pim sparse-mode
Interface Vlan100
  no shutdown
  vrf member EVPN-TENANT
  ip address 100.1.1.254/24
  fabric forwarding mode anycast-gateway
   ip pim neighbor-policy NONE*
```

cisco Live!

Tenant Routed Multicast (TRM)

TRM Verification

```
Leaf1# show fabric multicast vrf all
Fabric Multicast Enabled VRFs
VRF Name                VRF      Vprime         VN-Seg
                        ID       If             ID
default                 1        Null           0
TRM                     4        Vlan500        50000

Leaf1# show fabric multicast globals
Pruning: segment-based
Switch role: leaf
Fabric Control Seg: Null
Peer Fabric Control Address: 0.0.0.0
Advertising vPC RPF routes: Disabled
Created VNI List: -
Fwd Encap: Fabric
Overlay Distributed-DR: FALSE
Overlay spt-only: TRUE
```

Tenant Routed Multicast (TRM)

Show bgp internal mvpn detail

```
Leaf1# show bgp internal mvpn detail
************************************************
NGMVPN feature/server/role:     VxLAN/1/VxLAN
NGMVPN registered:              Yes (Dec  9 00:56:59.297696/never)
NGMVPN TRM mode:                L3 (0x000002)
NGMVPN down: in-prg/up-defer:   0/0
NGMVPN register/failures:       1/0
NGMVPN deregister/failures:     0/0
NGMVPN Convergence sent:        0
NGMVPN local-req sent/skipped:  7/3
NGMVPN local-req sent:          4 (L2VNI)/ 2 (L3VNI)/ 1 (All VNIs)
NGMVPN remote-req rcvd:         0 (L2VNI)/ 0 (L3VNI)/ 1 (All VNIs)
NGMVPN del remote:              0 (L2VNI)/ 0 (L3VNI)/ 0 (All VNIs)
NGMVPN msgs sent/acks rcvd:     9/9
NGMVPN msgs rcvd/acks sent:     24/14
NGMVPN msg err/ack err/drops:   0/0/0
Last xid sent to NGMVPN:        9
Last xid ack by NGMVPN:         9
. . . .
```

Tenant Routed Multicast (TRM)

Show bgp internal mvpn detail (contd...)

```
++++++++++++++++++++++++++++++++++++++++++
BGP MVPN RD Information for 192.168.1.1:4 (0xd625952c)
  VNI ID                          : 50000
  VRF                             : TRM
  Global NGMVPN mode              : L3  (1 L3 VRFs)
  VRF L3 Mode                     : Yes Jun  9 00:57:00.203673
  Enabled                         : Yes
  Delete Pending                  : No
  Cleanup Pending                 : No
  Import Pending                  : No
  Import In Progress              : No
  Created                         : Jun  9 00:56:55.570427
  Enabled At                      : Jun  9 00:56:55.570471
  . . .
  Is Auto RT                      : No
  Config VRF Import RT            : 1
    Import  RT cfg list:            192.168.1.1:1500
  Active VRF Import RT            : 1
  Active VRF Import RT list       : 192.168.1.1:1500
  VRF Import RT chg/chg-pending : 0/0
```

Tenant Routed Multicast (TRM)

Join from Receiver

```
Leaf1# show ip mroute 239.0.0.1 detail vrf TRM
IP Multicast Routing Table for VRF "TRM"

Total number of routes: 3
Total number of (*,G) routes: 1
Total number of (S,G) routes: 1
Total number of (*,G-prefix) routes: 1

(*, 239.0.0.1/32), uptime: 00:00:55, igmp(1) ip(0) pim(0)
  RPF-Source: 99.99.99.99 [0/0]
  RD-RT ext comm Route-Import:
  Data Created: No
  VPC Flags
    RPF-Source Forwarder
  Stats: 0/0 [Packets/Bytes], 0.000   bps
  Stats: Inactive Flow
  Incoming interface: loopback1000, RPF nbr: 209.165.200.1
  Outgoing interface list: (count: 1) (bridge-only: 0)
    Vlan100, uptime: 00:00:55, igmp (vpc-svi)
```

Tenant Routed Multicast (TRM)

FHR VTEP sends SA-AD (Type-5 Route) using BGP to other VTEPs

```
Leaf1# show bgp ipv4 mvpn sa-ad detail vrf TRM
BGP routing table information for VRF default, address family IPv4 MVPN
Route Distinguisher: 192.168.1.1:4    (L3VNI 50000)
BGP routing table entry for [5][10.0.0.1][239.0.0.1]/64, version 34
Paths: (1 available, best #1)
Flags: (0x000002) (high32 00000000) on xmit-list, is not in mvpn

  Advertised path-id 1
  Path type: local, path is valid, is best path
  AS-Path: NONE, path locally originated
    0.0.0.0 (metric 0) from 0.0.0.0 (192.168.1.1)
      Origin IGP, MED not set, localpref 100, weight 32768
      Extcommunity: RT:65000:150000

  Path-id 1 advertised to peers:
    10.10.10.10          20.20.20.20
```

Tenant Routed Multicast (TRM)

LHR / Remote VTEPs build (S, G)

```
Leaf3# show ip mroute 239.0.0.1 10.0.0.1 detail vrf TRM
IP Multicast Routing Table for VRF "TRM"

Total number of routes: 3
Total number of (*,G) routes: 1
Total number of (S,G) routes: 1
Total number of (*,G-prefix) routes: 1

(10.0.0.1/32, 239.0.0.1/32), uptime: 00:01:55, ip(0) mrib(1) pim(0) ngmvpn(0)
  RPF-Source: 10.0.0.1 [0/20]
  RD-RT ext comm Route-Import:  0b c0 a8 01 01 05 dc 00 01 c0 a8 01 01 83 e7
  Data Created: Yes
  Fabric dont age route
  VPC Flags
    RPF-Source Forwarder
  Stats: 1/84 [Packets/Bytes], 0.000   bps
  Stats: Inactive Flow
  Incoming interface: Vlan500, RPF nbr: 192.168.1.1
  Outgoing interface list: (count: 1) (bridge-only: 0)
    Vlan100, uptime: 00:01:55, mrib (vpc-svi)
```

Tenant Routed Multicast (TRM)

VTEP3/4 sends Type-7 back to FHR on L3VNI

```
Leaf3# show bgp ipv4 mvpn route-type 7 detail
BGP routing table information for VRF default, address family IPv4 MVPN
Route Distinguisher: 192.168.1.1:33767    (Local VNI: 50000)
BGP routing table entry for [7][10.0.0.1][239.0.0.1][65000]/96, version 43
Paths: (1 available, best #1)
Flags: (0x000002) (high32 00000000) on xmit-list, is not in mvpn

  Advertised path-id 1
  Path type: local, path is valid, is best path
  AS-Path: NONE, path locally originated
    0.0.0.0 (metric 0) from 0.0.0.0 (192.168.3.3)
      Origin IGP, MED not set, localpref 100, weight 32768
      Extcommunity: RT:192.168.1.1:1500

  Path-id 1 advertised to peers:
    10.10.10.10          20.20.20.20
```

Tenant Routed Multicast (TRM)

VTEP1 Receives the Type-7 route

```
Leaf1# show bgp ipv4 mvpn route-type 7 detail
BGP routing table information for VRF default, address family IPv4 MVPN
Route Distinguisher: 192.168.1.1:4    (L3VNI 50000)
BGP routing table entry for [7][10.0.0.1][239.0.0.1][65000]/96, version 36
Paths: (1 available, best #1)
Flags: (0x00001a) (high32 00000000) on xmit-list, is in mvpn, is not in HW

  Advertised path-id 1
  Path type: external, path is valid, is best path, in rib
             Imported from 192.168.1.1:33767:[7][10.0.0.1][239.0.0.1][65000]/96
  AS-Path: 65001 65000 , path sourced external to AS
    192.168.100.100 (metric 0) from 10.10.10.10 (192.168.10.10)
      Origin IGP, MED not set, localpref 100, weight 0
      Extcommunity: RT:192.168.1.1:500

  Path-id 1 advertised to peers:
```

Tenant Routed Multicast (TRM)

VTEP1 adds L3VNI in the OIF List

```
Leaf1# show ip mroute 239.0.0.1 10.0.0.1 detail vrf TRM
IP Multicast Routing Table for VRF "TRM"

Total number of routes: 3
Total number of (*,G) routes: 1
Total number of (S,G) routes: 1
Total number of (*,G-prefix) routes: 1

(10.0.0.1/32, 239.0.0.1/32), uptime: 00:12:03, ip(0) mrib(1) pim(0) ngmvpn(1)
  RPF-Source: 10.0.0.1 [0/0]
  RD-RT ext comm Route-Import:
  Data Created: Yes
  Received Register stop
  Fabric dont age route
  Stats: 695/58380 [Packets/Bytes], 672.000 bps
  Stats: Active Flow
  Incoming interface: Vlan100, RPF nbr: 10.0.0.1, internal
  Outgoing interface list: (count: 2) (Fabric OIF) (bridge-only: 0)
    Vlan500, uptime: 00:12:02, ngmvpn
    Vlan100, uptime: 00:12:03, mrib, (RPF)
```

Tenant Routed Multicast (TRM)

VTEP3 adds L3VNI as Incoming Interface

```
Leaf3# show ip mroute 239.0.0.1 10.0.0.1 detail vrf TRM
IP Multicast Routing Table for VRF "TRM"

Total number of routes: 3
Total number of (*,G) routes: 1
Total number of (S,G) routes: 1
Total number of (*,G-prefix) routes: 1

(10.0.0.1/32, 239.0.0.1/32), uptime: 00:24:35, ip(0) mrib(1) pim(0) ngmvpn(0)
  RPF-Source: 10.0.0.1 [0/20]
  RD-RT ext comm Route-Import:  0b c0 a8 01 01 05 dc 00 01 c0 a8 01 01 83 e7
  Data Created: Yes
  Fabric dont age route
  VPC Flags
    RPF-Source Forwarder
  Stats: 1/84 [Packets/Bytes], 0.000   bps
  Stats: Inactive Flow
  Incoming interface: Vlan500, RPF nbr: 192.168.1.1
  Outgoing interface list: (count: 1) (bridge-only: 0)
    Vlan100, uptime: 00:24:35, mrib (vpc-svi)
```

Troubleshooting Tools

ELAM Wrapper - Demo

```
debug platform internal tah elam
trigger init
set outer ipv4 src-ip ip-address dst-ip ip-address
start
report [detail]
```

Demo – Consistency Checker (CC)

- `test consistency-checker forwarding [ipv4 unicast | vrf vrf-name] [module slot] [stop]`

- `Show consistency-checker forwarding`

Demo – L2 PI/PD Troubleshooting

Show troubleshoot l2 mac *mac-address* **[vlan** *vlan-id***]**

Demo – L3 PI/PD Troubleshooting

Show troubleshoot l3 [ipv4 | ipv6] *v4/v6 address* **[src-ip** *v4/v6 address*] **[vrf** *vrf-name*]

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.

cisco Live!

Continue your education

Demos in the Cisco campus

Walk-in labs

Meet the engineer 1:1 meetings

Related sessions

Thank you

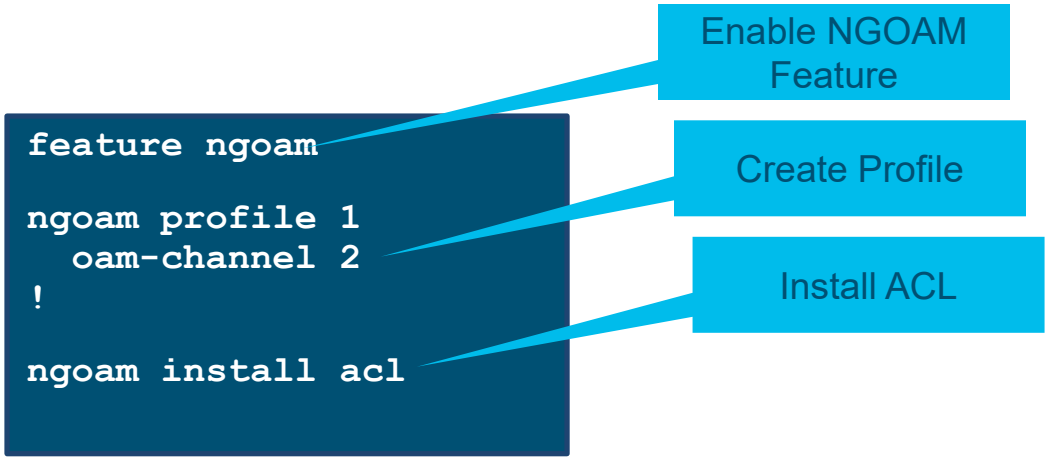
You make **possible**

Troubleshooting Tools – Extra Slides

Troubleshooting Tools

VXLAN OAM

- VXLAN OAM feature introduced in 7.0(3)I5(2) – NGOAM
- Need a feature for Path verification and Path tracking with Telemetry data
- Similar to Fabric Path OAM
- Provides 3 features
  - VxLAN Ping
  - VxLAN Traceroute
  - VxLAN PathTrace



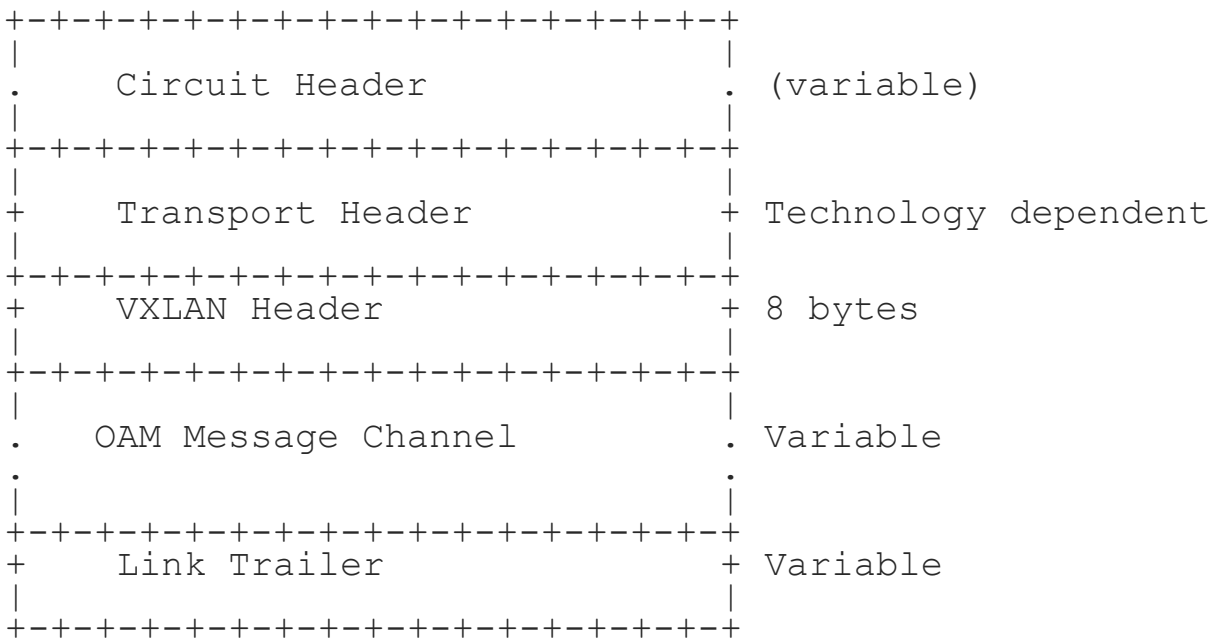
VXLAN OAM

Usability

- Helps diagnose underlay / overlay reachability of VMs
- Covers exact path as Data Packet
- Path verification for all ECMP paths in Overlay
- Path tracking – Exact path host traffic takes in overlay and underlay network
- Layer 2 – Traceroute / Ping to VM host from Leaf
- Layer 3 – Traceroute / Ping to Vm host from Leaf
- Flexible OAM channel supporting multiple drafts
  - Tissa draft – nvo3

VXLAN OAM

VXLAN OAM Message



VXLAN OAM

VXLAN PING

- VxLAN Ping checks connectivity to the destination, where the destination can be VM's IP address or routed loopback addresses on the leaf switch
- Since there are multiple paths, only one path is followed based on the flow parameters
- Ping for both VM / Host MAC and IP
- Default ping support – Ping based on just destination address and VNI segment
- Allow users to specify flow parameters such as UDP port, destination and source address
  - This helps VxLAN ping follow the specific path the unicast ping will take to reach the destination

VxLAN OAM

PING NVE MAC

```
Leaf1# ping nve mac 0050.56b3.bcef 200 port-channel 101 profile 1 vni 20000
verbose

Codes: '!' - success, 'Q' - request not sent, '.' - timeout,
'D' - Destination Unreachable, 'X' - unknown return code,
'm' - malformed request(parameter problem)
,'c' - Corrupted Data/Test, '#' - Duplicate response

Sender handle: 21

! sport 51932 size 39,Reply from 192.168.100.100,time = 5 ms
! sport 51932 size 39,Reply from 192.168.100.100,time = 4 ms
! sport 51932 size 39,Reply from 192.168.100.100,time = 4 ms
! sport 51932 size 39,Reply from 192.168.100.100,time = 4 ms
! sport 51932 size 39,Reply from 192.168.100.100,time = 4 ms
```

VxLAN OAM

VxLAN Traceroute

- VxLAN Traceroute – Used to trace the path a packet takes between source and destination
- Only one path is traced based on the given flow parameters
  - Trace will show uni-directional path the packet takes to the destination, but the return path may be different
- Should be able to trace VTEPs, access switch and end-host. For the default mode, the user should be able to trace the tunnel endpoint IP address and the segment ID
- Actual path taken by a packet is dependent on all the L2/L3/L4 header fields and network topology at the time the packet is sent
  - Users can specify the flow parameters

VxLAN OAM

VxLAN PathTrace

- Similar to Traceroute, but uses Nvo3 channel
  - Carries additional ingress / egress and load information of the path
  - Provides additional information if the device supports nvo3 channel else its same as traceroute
- Actual path taken by a packet is dependent on all the L2/L3/L4 header fields and network topology at the time the packet is sent
  - Users can specify the flow parameters

```
Leaf1# pathtrace nve mac 0050.56b3.bcef 200 port-channel 101 vni 20000
<snip>
Path trace Request to peer ip 192.168.100.100 source ip 192.168.99.99
Sender handle: 35Hop   Code   ReplyIP   IngressI/f  EgressI/f   State
================================================
1 !Reply from 10.101.1.10, Eth2/1  Eth1/17  UP / UP
2 !Reply from 192.168.100.100, Eth1/17  Unknown  UP / DOWN
```

Troubleshooting Tools

ELAM

- Embedded Logic Analyzer module (ELAM) – tool used to capture a packet processed by a Cisco ASIC
- Depending on the N9k platform,
  - ELAM on NS ASIC
  - ELAM on TAHOE ASIC
- Useful in scenario's where packet forwarding is impacted
- Can perform capture for raw packet from the host and even VxLAN encapsulated packet towards the VxLAN Core
- Useful for Cisco TAC and Cisco Engineering for understanding the cause of packet loss or impacted forwarding

ELAM on N9k Platform

Northstar (NS) ASIC

- NS datapath is divided into ingress and egress pipelines
  - 2 ELAM's present in each pipeline (Input & Output)
  - Packets can be captured based on either input select lines or output select lines but not both
- Input Select Lines
  - 3 → Outerl2-outerl3-outerl4
  - 4 → Innerl2-innerl3-inner l4
  - 5 → Outerl2-innerl2
  - 6 → Outerl3-innerl3
  - 7 → Outerl4-innerl4
- Output Select Lines
  - 0 → Pktrw
  - 5 → Sideband

ELAM on N9k Platform

NS ELAM Steps

- **Attach module X**
- **Debug platform internal ns elam asic [0 | 1]**
- **Trigger [init | reset] [ingress | egress] in-select [3 – 7] out-select [0, 5]**
  - Ingress – For capturing packets from front panel to Fabric
  - Egress – For capturing packets from Fabric to Front Panel
- **Set outer [ipv4 | l2 | l4] . . . .**
- **Start**
- **Status {Can be Armed / Triggered}**
- **report**

ELAM on N9k Platform

Verify Port-Asic, Slice and Src-Id

```
N9K# show hardware internal tah interface e1/49
<snip>
 IfIndex: 436232192
 DstIndex: 5952
 IfType: 26
 Asic: 0
 AsicPort: 56
 SrcId: 48 <<
 Slice: 1
 PortOnSlice: 24
```

ELAM on N9k Platform

TAHOE ELAM Steps

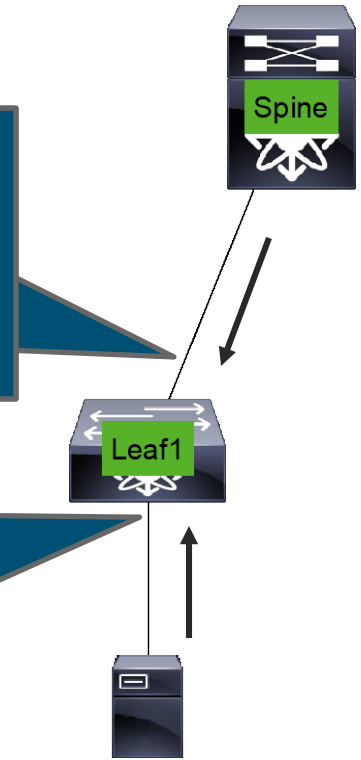
- **Attach module X**
- **Debug platform internal tah elam asic [0 | 1]**
- **Trigger [init | reset] asic [num] slice [num] lu-a2d [0 | 1] in-select [3-7] out-select [0-5] use-src-id [src-id]**
  - Lu-a2d 0 – used for reverse ELAM, where trigger is based on result
  - Lu-a2d 1 – used for ELAM where trigger is based on packet attributes
- **Set outer [ipv4 | l2 | l4] . . . .**
- **Start**
- **Status {Can be Armed / Triggered}**
- **Report [detail]**

ELAM on N9k Platform

TAHOE ELAM

```
att mod 1
   debug platform internal tah elam asic 0
    trigger init asic 0 slice 0 lu-a2d 1 in-select 7 out-select 0 use-src-id 52
      reset
      set inner ipv4 src_ip 100.1.1.1 dst_ip 100.1.1.5
      start
      report
```

```
att mod 1
  debug platform internal tah elam asic 0
    trigger init asic 0 slice 0 lu 1 in-select 6 out-select 0
    reset
      set outer ipv4 src_ip 100.1.1.1 dst_ip 100.1.1.5
      start
      report
```



ELAM on N9k Platform

Partial Output

- Dot1Q Header

```
module-1(TAH-elam-insel6)# report detail | grep pr_lu_vec_l2v.qtag0
GBL_C++: [MSG] pr_lu_vec_l2v.qtag0_vld: 0x1 << dot1q yes? 0x1
GBL_C++: [MSG] pr_lu_vec_l2v.qtag0_cos: 0x0
GBL_C++: [MSG] pr_lu_vec_l2v.qtag0_de: 0x0
GBL_C++: [MSG] pr_lu_vec_l2v.qtag0_vlan: 0x64 << VL 100
```

- VLAN

```
module-1(TAH-elam-insel6)# report detail | grep -1 fpx_lookup_vec.lkup.macsakey.key.fid
GBL_C++: [MSG] fpx_lookup_vec.lkup.macsakey.key.vld: 0x1
GBL_C++: [MSG] fpx_lookup_vec.lkup.macsakey.key.fid_type: 0x0
GBL_C++: [MSG] fpx_lookup_vec.lkup.macsakey.key.fid_vld: 0x0
GBL_C++: [MSG] fpx_lookup_vec.lkup.macsakey.key.fid: 0x64 << dec 0xa = VL 10
GBL_C++: [MSG] fpx_lookup_vec.lkup.macsakey.key.mac: 0xFEC80E2715
```

ELAM on N9k Platform

Partial Output

- Src & Dst IP

```
module-1(TAH-elam-insel6)# report detail | grep vec_l3v.ip.*a
 GBL_C++: [MSG] pr_lu_vec_l3v.ip.da: 0x00000000000000000064010101 << 100.1.1.1
 GBL_C++: [MSG] pr_lu_vec_l3v.ip.sa: 0x00000000000000000064010105 << 100.1.1.5
```

- Src MAC

```
module-1(TAH-elam-insel6)# report detail | grep -i fpx_lookup_vec.lkup.macsakey.key.mac
 GBL_C++: [MSG] fpx_lookup_vec.lkup.macsakey.key.mac: 0xFEC80E2715 << 00fe.c80e.2715
 GBL_C++: [MSG] fpx_lookup_vec.lkup.macsakey.key.mac: 0xFEC80E2715
 GBL_C++: [MSG] fpx_lookup_vec.lkup.macsakey.key.mac: 0xFEC80E2715
 GBL_C++: [MSG] fpx_lookup_vec.lkup.macsakey.key.mac: 0xFEC80E2715
 GBL_C++: [MSG] fpx_lookup_vec.lkup.macsakey.key.mac: 0xFEC80E2715
```

You make **possible**