Configure and Verify VXLAN with MP-BGP EVPN Control Plane

Contents

Introduction

Prerequisites

- Requirements
- Components Used

Background Information

Configure

- Network Diagram
- Configuration
  - VTEP1
  - VTEP2
  - VTEP3
  - VTEP4
  - SPINE
  - HOST A
  - HOST B
  - HOST C and D

Verify

- Connectivity from Host A to External Host B
- Connectivity from Host A toHOST C (Intra-VNI)
- Connectivity from Host A toHOST D (Inter-VNI)
- Routing Table of Host B (External Peer)
- Verification of Control-Plane

Troubleshoot

Introduction

This document describes the VXLAN configuration using MP-BGP EVPN control-plane.

Prerequisites

Requirements

Cisco recommends that you have knowledge of these topics:

- MPLS Layer 3 VPNs
- MP-BGP would also help.

Components Used

This document is not restricted to specific software and hardware versions.

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

Background Information

This document also shows a sample network scenario and its configuration with relevant outputs for verification and better understanding.

VXLAN is designed to provide network virtualization. It is MAC in UDP encapsulation. Layer 2 infrastructure is extended over Layer 3 underlay network to provide a simplified service which is not dependent on physical and geographical location of servers in the datacenters.

This document describes VXLAN with MP-BGP EVPN Control Plane. This means that BGP protocol is used in overlay infrastructure to send and receive updates.

In traditional network deployments STP was used, this resulted in some uplinks permanently blocking state. In VXLAN design all the uplinks are operational and ECMP is leveraged as underlay infrastructure is IP network.

Discussion of all the details is outside the scope of this document, however some important terminologies are shown below.

VXLAN - Virtual Extensible LAN

MP-BGP - Multiprotocol BGP

EVPN - Ethernet VPN

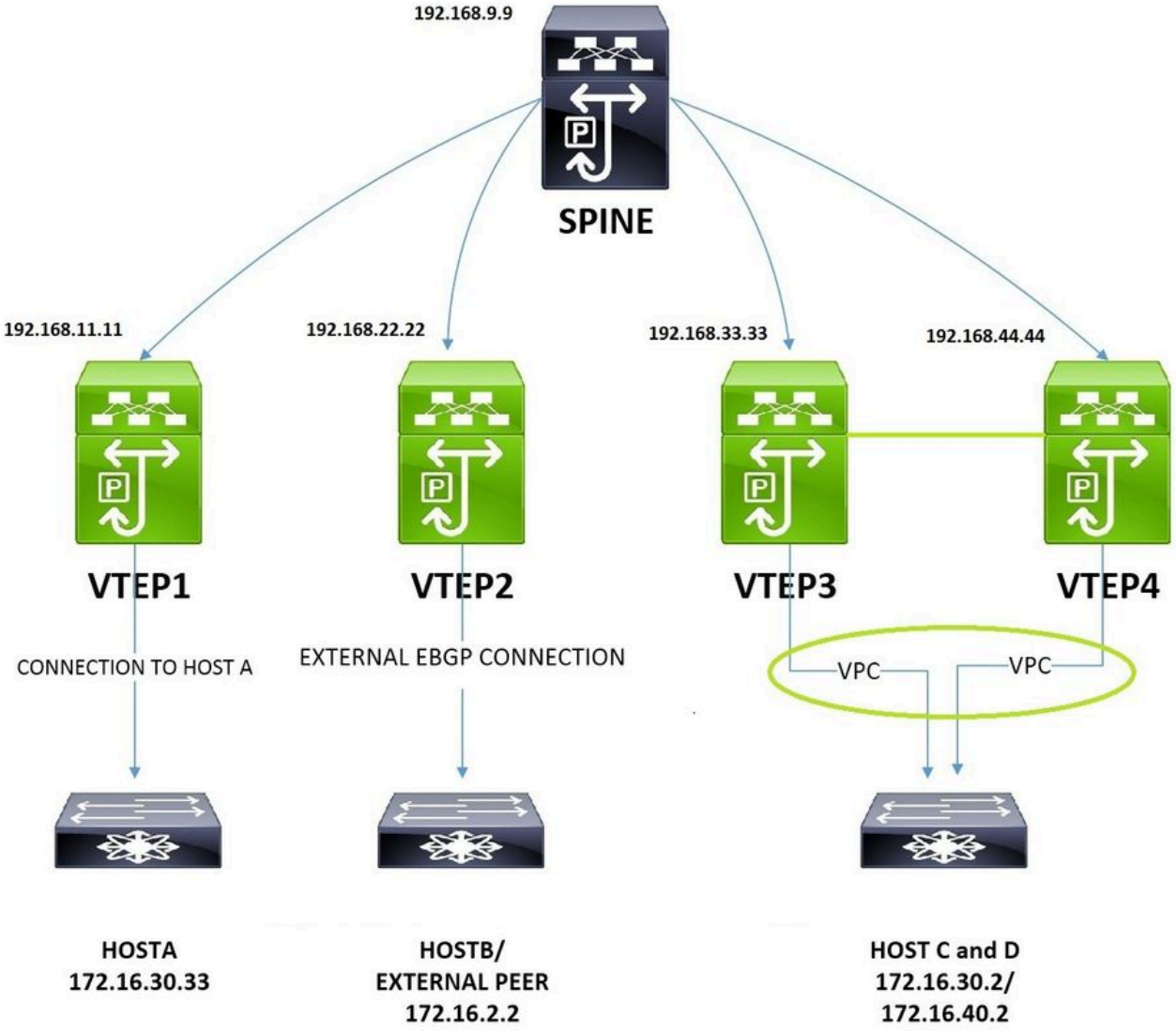
VTEP - Virtual Tunnel End Point. Also known as the leaf. This the place where the packets are encapsulated and decapsulated.

Spine - This is very similar to Route-Reflector in MPLS L3 VPNs. This device takes the updates from one VTEP and passes on it to other VTEP.

VNI - VXLAN Network Identifier. This is mainly used to provide isolation for layer-2 boundaries. This field is 24 bit in length so it overcomes the range limitation of traditional vlans. A VNI in a VTEP is mapped to a traditional vlan. This is discussed later.

Configure

Network Diagram



The image shown is used for configuration and verification aspects. This covers non-vpc, vpc, intra-vni, inter-vni and external connectivity configurations from the perspective of VXLAN infrastructure.

Configuration

VTEP1

<#root>

!

Enabling features

```
nv overlay evpn
feature ospf
feature bgp
feature pim
feature interface-vlan
feature vn-segment-vlan-based
```

```
feature lacp
feature vpc
feature nv overlay

!

fabric forwarding anycast-gateway-mac 0001.0001.0001

!



This is needed for seamless VM mobility across VTEPS,

this configuration is  same on all VTEPS.


ip pim rp-address 192.168.9.9 group-list 224.0.0.0/4

! SPINE is the RP.



!
ip pim ssm range 232.0.0.0/8

!
vlan 1,10,30,40,100,200

!
vlan 10

 !



VLAN 10 is used as layer3 VNI to route Inter-VNI traffic.


 name L3-VNI-VLAN-10
 vn-segment 10000010
 vlan 30

! The Host A resides on Vlan 30, The below command 'maps' vlan 30 with VNID 10000030.


 vn-segment 10000030
!
vrf context EVPN-L3-VNI-VLAN-10

 ! Defining layer3 vrf for Inter-VNI traffic.


 vni 10000010
 rd auto
 address-family ipv4 unicast
 route-target both auto
 route-target both auto evpn

!

interface Vlan10

! Layer3 VNI associated interface vlan does not have an ip address.


 no shutdown
```

```
 vrf member EVPN-L3-VNI-VLAN-10
 ip forward

!

interface Vlan30

 ! Associating the Host A Vlan with layer3 vrf.


 no shutdown
 vrf member EVPN-L3-VNI-VLAN-10
 ip address 172.16.30.1/24
 fabric forwarding mode anycast-gateway

!



This is needed for seamless VM mobility across VTEPS, same on all VTEPS.



!

interface nve1

! Nve is logical interface where VXLAN packets are encapsulated and decapsulated.


 no shutdown
 source-interface loopback2
 host-reachability protocol bgp

! This means BGP control plane is used to exchange updates.


 member vni 10000010 associate-vrf

! associate-vrf is used for for layer3 vni.


 member vni 10000030
 suppress-arp
 mcast-group 239.1.1.10

 ! A vlan or set of vlans mapped to VNI can be given identical multicast address, this is used for contr


!

interface Ethernet1/2

 ! Ospf with PIM is used as Underlay.


 description "Going to Spine"
 no switchport
 ip address 192.168.19.1/24
 ip router ospf UNDERLAY area 0.0.0.0
 ip pim sparse-mode
 no shutdown

!

interface Ethernet1/11
```

```
! Port to Host A.


 switchport mode trunk

!

interface loopback2 !
```

Loopback for BGP Peering.

```
 description "Loopback for "BGP"
 ip address 192.168.11.11/32
 ip router ospf UNDERLAY area 0.0.0.0
 ip pim sparse-mode

!

router ospf UNDERLAY

!
router bgp 65000
 address-family ipv4 unicast
 address-family l2vpn evpn
 neighbor 192.168.9.9
```

!

Peering with SPINE.

```
 remote-as 65000
 update-source loopback2
 address-family ipv4 unicast
 address-family l2vpn evpn
 send-community extended
 vrf EVPN-L3-VNI-VLAN-10
 address-family ipv4 unicast
 advertise l2vpn evpn

!
evpn
 vni 10000030 l2
 rd auto
```

! RD is default calculated as VNI:BGP Router ID

```
 route-target import auto
```

! RT is default calculated as BGP AS:VNI

```
 route-target export auto
```

VTEP2

<#root>

```
!

nv overlay evpn
feature ospf
feature bgp
feature pim
feature interface-vlan
feature vn-segment-vlan-based
feature lacp
feature vpc
feature nv overlay

!

fabric forwarding anycast-gateway-mac 0001.0001.0001

!
ip pim rp-address 192.168.9.9 group-list 224.0.0.0/4

!
ip pim ssm range 232.0.0.0/8
vlan 1,10,30,40,100

!
vlan 10
! This VTEP is dedicated for external connectivity, there is only layer3 VNI config.


 name L3-VNI-VLAN-10
 vn-segment 10000010

!

vrf context EVPN-L3-VNI-VLAN-10
! Defining layer3 vrf for Inter-VNI traffic.


 vni 10000010
 rd auto
 address-family ipv4 unicast
 route-target both auto
 route-target both auto evpn

!

interface Vlan10
! Layer3 VNI associated interface vlan does not have an ip address.


 no shutdown
 vrf member EVPN-L3-VNI-VLAN-10
 ip forward

!

interface Vlan100
! This vlan is used to peer with external EBGP Peer.


 no shutdown
```

```
 vrf member EVPN-L3-VNI-VLAN-10
 ip address 192.168.1.2/24

!

interface nve1
 no shutdown
 source-interface loopback2
 host-reachability protocol bgp
 member vni 10000010 associate-vrf

!

interface Ethernet1/2
```

! Ospf and PIM are used in Underlay.

```
 description "Going to Spine"
 no switchport
 ip address 192.168.29.2/24
 ip router ospf UNDERLAY area 0.0.0.0
 ip pim sparse-mode
 no shutdown

!

interface Ethernet1/12
```

! Port to External Peer.

```
 switchport mode trunk

!

interface loopback2
 ip address 192.168.22.22/32
 ip router ospf UNDERLAY area 0.0.0.0
 ip pim sparse-mode

!

router ospf UNDERLAY

!
router bgp 65000
 address-family ipv4 unicast
 address-family l2vpn evpn
 neighbor 192.168.9.9
```

! Peering with SPINE.

```
 remote-as 65000
 update-source loopback2
 address-family ipv4 unicast
 address-family l2vpn evpn
 send-community extended
 vrf EVPN-L3-VNI-VLAN-10
 address-family ipv4 unicast
 advertise l2vpn evpn
 neighbor 192.168.1.1
```

```
! Peering with External Peer, under vrf.


 remote-as 65111
 update-source Vlan100
 address-family ipv4 unicast
```

VTEP3

The configs for VTEP3 and VTEP1 are almost identical. Only difference is VPC and an additional layer2 VNI for vlan 40.

```
<#root>

!

nv overlay evpn
feature ospf
feature bgp
feature pim
feature interface-vlan
feature vn-segment-vlan-based
feature lacp
feature vpc
feature nv overlay

!

fabric forwarding anycast-gateway-mac 0001.0001.0001

!
ip pim rp-address 192.168.9.9 group-list 224.0.0.0/4

!
ip pim ssm range 232.0.0.0/8

!
vlan 1,10,20,30,40

!
vlan 10
 name L3-VNI-VLAN-10
 vn-segment 10000010

!
vlan 30
 vn-segment 10000030

!
vlan 40

!


New host vlan 40.


 vn-segment 10000040
```

```
!

vpc domain 2

! Vpc Configs.


 peer-keepalive destination 10.197.204.103 source 10.197.204.106

!

interface Vlan10
 no shutdown
 vrf member EVPN-L3-VNI-VLAN-10
 ip forward

!

interface Vlan30
 no shutdown
 vrf member EVPN-L3-VNI-VLAN-10
 ip address 172.16.30.1/24

!
 fabric forwarding mode anycast-gateway

!

interface Vlan40
 no shutdown
 vrf member EVPN-L3-VNI-VLAN-10
 ip address 172.16.40.1/24

!
 fabric forwarding mode anycast-gateway

!

interface port-channel2
 switchport mode trunk
 vpc 2

!

interface port-channel34
 switchport mode trunk
 spanning-tree port type network
 vpc peer-link

!

interface nve1
 no shutdown
 source-interface loopback2
 host-reachability protocol bgp
 member vni 10000010 associate-vrf
 member vni 10000030
 suppress-arp
 mcast-group 239.1.1.10
 member vni 10000040

!
```

New layer2 VNI for Vlan 40.

```
 suppress-arp
 mcast-group 239.1.1.20

!

interface Ethernet1/1
```

! Connected to VTEP4.

```
 switchport mode trunk
 channel-group 34 mode active

!

interface Ethernet1/2
 description "going to Spine"
 no switchport
 ip address 192.168.39.3/24
 ip router ospf UNDERLAY area 0.0.0.0
 ip pim sparse-mode
 no shutdown

!

interface Ethernet1/13
```

! Connected to N5K, which simulates Host C and D.

```
 switchport mode trunk
 channel-group 2 mode active

!

interface loopback2
 description "loopback for Bgp"
 ip address 192.168.33.33/32
 ip address 192.168.33.34/32 secondary
```

!

For other VTEPs VTEP3 and VTEP4 look as single entity.

```
 ip router ospf UNDERLAY area 0.0.0.0
```

!

This secondary address is needed in Vpc designs.

```
!

router ospf UNDERLAY

!
router bgp 65000
 address-family ipv4 unicast
```

```
 address-family l2vpn evpn
 neighbor 192.168.9.9 remote-as 100
 remote-as 65000
 update-source loopback2
 address-family ipv4 unicast
 address-family l2vpn evpn
 send-community extended
 vrf EVPN-L3-VNI-VLAN-10
 address-family ipv4 unicast
 advertise l2vpn evpn

!
evpn
 vni 10000030 l2
 rd auto
 route-target import auto
 route-target export auto
 vni 10000040 l2
 rd auto
 route-target import auto
 route-target export auto
```

VTEP4

```
<#root>

!

nv overlay evpn
feature ospf
feature bgp
feature pim
feature interface-vlan
feature vn-segment-vlan-based
feature lacp
feature vpc
feature nv overlay

!

fabric forwarding anycast-gateway-mac 0001.0001.0001

!
ip pim rp-address 192.168.9.9 group-list 224.0.0.0/4

!
ip pim ssm range 232.0.0.0/8

!
vlan 1,10,20,30,40

!
vlan 10
 name L3-VNI-VLAN-10
 vn-segment 10000010

!
vlan 30
 vn-segment 10000030
```

```
!
vlan 40
 vn-segment 10000040

!
vrf context EVPN-L3-VNI-VLAN-10
 vni 10000010
 rd auto
 address-family ipv4 unicast
 route-target both auto
 route-target both auto evpn

!

interface Vlan10
 no shutdown
 vrf member EVPN-L3-VNI-VLAN-10
 ip forward

!

interface Vlan30
 no shutdown
 vrf member EVPN-L3-VNI-VLAN-10
 ip address 172.16.30.1/24

!
 fabric forwarding mode anycast-gateway

!

interface Vlan40
 no shutdown
 vrf member EVPN-L3-VNI-VLAN-10
 ip address 172.16.40.1/24

!
 fabric forwarding mode anycast-gateway

!

interface port-channel2
 switchport mode trunk
 vpc 2

!

interface port-channel34
 switchport mode trunk
 spanning-tree port type network
 vpc peer-link

!

interface nve1
 no shutdown
 source-interface loopback2
 host-reachability protocol bgp
 member vni 10000010 associate-vrf
 member vni 10000030
 suppress-arp
```

```
 mcast-group 239.1.1.10
 member vni 10000040
 suppress-arp
 mcast-group 239.1.1.20

!

interface Ethernet1/1
```

! Connected to VTEP3.

```
 switchport mode trunk
 channel-group 34 mode active

!

interface Ethernet1/2
 description "going to spine"
 no switchport
 ip address 192.168.49.4/24
 ip router ospf UNDERLAY area 0.0.0.0
 ip pim sparse-mode
 no shutdown

!

interface Ethernet1/13
```

! Connected to N5K, which simulates Host C and D.

```
 switchport mode trunk
 channel-group 2 mode active

!

router ospf UNDERLAY

!
router bgp 65000
 address-family ipv4 unicast
 address-family l2vpn evpn
 neighbor 192.168.9.9 remote-as 100
 remote-as 65000
 update-source loopback2
 address-family ipv4 unicast
 address-family l2vpn evpn
 send-community extended
 vrf EVPN-L3-VNI-VLAN-10
 address-family ipv4 unicast
 advertise l2vpn evpn

!
evpn
 vni 10000030 l2
 rd auto
 route-target import auto
 route-target export auto
 vni 10000040 l2
 rd auto
 route-target import auto
 route-target export auto
```

SPINE

<#root>

```
!

nv overlay evpn
feature ospf
feature bgp
feature pim
feature interface-vlan
feature vn-segment-vlan-based
feature lacp
feature vpc
feature nv overlay

!
ip pim rp-address 192.168.9.9 group-list 224.0.0.0/4

!
ip pim ssm range 232.0.0.0/8

!

interface Ethernet1/1
! To VTEP1.


 ip address 192.168.19.9/24
 ip router ospf UNDERLAY area 0.0.0.0
 ip pim sparse-mode
 no shutdown

!

interface Ethernet1/2
! To VTEP2.


 ip address 192.168.29.9/24
 ip router ospf UNDERLAY area 0.0.0.0
 ip pim sparse-mode
 no shutdown

!

interface Ethernet1/3
! To VTEP3.


 ip address 192.168.39.9/24
 ip router ospf UNDERLAY area 0.0.0.0
 ip pim sparse-mode
 no shutdown

!

interface Ethernet1/4
```

```
! To VTEP4.


 ip address 192.168.49.9/24
 ip router ospf UNDERLAY area 0.0.0.0
 ip pim sparse-mode
 no shutdown

!

interface loopback1

! SPINE is RP(Rendezvous Point).


 ip address 192.168.9.9/32
 ip router ospf UNDERLAY area 0.0.0.0
 ip pim sparse-mode

!

router ospf UNDERLAY

!
router bgp 65000
 log-neighbor-changes
 address-family ipv4 unicast
 address-family l2vpn evpn
 retain route-target all
 template peer VTEP-PEERS
 remote-as 65000
 update-source loopback1
 address-family ipv4 unicast
 send-community both
 route-reflector-client

! Spine treats VTEPs as Route-Reflector Clients.


 address-family l2vpn evpn
 send-community both
 route-reflector-client
 neighbor 192.168.11.11

! VTEP1.


 inherit peer VTEP-PEERS
 neighbor 192.168.22.22

! VTEP2.


 inherit peer VTEP-PEERS
 neighbor 192.168.33.33

! VTEP3.


 inherit peer VTEP-PEERS
 neighbor 192.168.44.44

! VTEP4.


 inherit peer VTEP-PEERS
```

HOST A

Host A is simulated by a 3750 Switch.

<#root>

```
! This port is the uplink to VTEP1.

interface GigabitEthernet1/0/1
 switchport trunk encapsulation dot1q
 switchport mode trunk

!

interface Vlan30
 ip address 172.16.30.33 255.255.255.0

!

! Below the default route to VTEP1.


ip route 0.0.0.0 0.0.0.0 172.16.30.1
```

HOST B

Host B is the external peering device. N5K is used here.

<#root>

```
!

router bgp 65111
 address-family ipv4 unicast

!
 network 172.16.2.2/32

!

Advertsing the external subnet to VXLAN infrastructure.


 neighbor 192.168.1.2 remote-as 65000

!


EBGP Peering with VTEP2.


 address-family ipv4 unicast

!
```

```
interface loopback1
 ip address 172.16.2.2/32

!

interface Ethernet1/19

!

Uplink port to VTEP2.


 switchport mode trunk

!

interface Vlan100
 no shutdown
 ip address 192.168.1.1/24
```

HOST C and D

Host C and D are simulated by Nexus5k , keeping the ip addresses in distinct vrfs.

```
<#root>

!

vrf context vni30

! This vrf simulates the HOST C.


 ip route 0.0.0.0/0 172.16.30.1
vrf context vni40

! This vrf simulates the HOST D.


 ip route 0.0.0.0/0 172.16.40.1

!

interface Vlan30

! Addressing for HOST C.


 no shutdown
vrf member vni30
 ip address 172.16.30.2/24

!

interface Vlan40

! Addressing for HOST D.


 no shutdown
```

```
 vrf member vni40
 ip address 172.16.40.2/24

!

interface Ethernet1/20
```

! Uplink port to VTEP3 in Port-Channel.

```
 switchport mode trunk
 channel-group 2 mode active

!

interface Ethernet1/21
```

! Uplink port to VTEP4 in Port-Channel.

```
 switchport mode trunk
 channel-group 2 mode active< /pre>
```

Verify

Connectivity from Host A to External Host B

```
HOST_A#ping 172.16.2.2
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 172.16.2.2, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/4/9 ms
```

Connectivity from Host A to HOST C (Intra-VNI)

```
HOST_A#ping 172.16.30.2
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 172.16.30.2, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/4/9 ms
```

Connectivity from Host A to HOST D (Inter-VNI)

```
HOST_A#ping 172.16.40.2

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 172.16.40.2, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/2/9 ms
```

Routing Table of Host B (External Peer)

<#root>

```
N5K-5672-1# show ip route bgp
IP Route Table for VRF "default"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

172.16.30.2/32, ubest/mbest: 1/0, pending
```

! Host route for Host C.

```
 *via 192.168.1.2, [20/0], 00:00:22, bgp-65100, external, tag 65000,
172.16.30.33/32, ubest/mbest: 1/0, pending
```

! Host route for Host A.

```
 *via 192.168.1.2, [20/0], 00:00:22, bgp-65100, external, tag 65000,
172.16.40.2/32, ubest/mbest: 1/0, pending
```

! Host route for Host D.

```
 *via 192.168.1.2, [20/0], 00:00:22, bgp-65100, external, tag 65000,
```

It can be observed that the host routes have been successfully advertised to this external BGP peer.

Verification of Control-Plane

- This command shows the mapping of traditional vlans with VNIDs.

```
 VTEP1# show vxlan
Vlan VN-Segment
==== ==========
10 10000010
30 10000030
40 10000040
```

- The next is to check that mac is learned locally on VTEP.

<#root>

```
 VTEP1# show mac address-table vlan 30
Legend:
 * - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
 age - seconds since last seen,+ - primary entry using vPC Peer-Link,
 (T) - True, (F) - False
 VLAN MAC Address Type age Secure NTFY Ports
```

```
---------+----------------+--------+---------+------+----+------------------
```

*****

```
 30 0006.f63f.e3c1 dynamic 0 F F Eth1/11 ! Mac of HOST A
* 30 8c60.4ff2.f541 dynamic 0 F F nve1(192.168.33.34)!
```

Mac of HOST C installed into mac address table, it was learned from BGP.

```
G 30 e00e.da2a.2393 static - F F sup-eth1(R)
```

- Next step is to check that the route is installed into l2rib.

<#root>

```
 VTEP1# show l2route evpn mac evi 30
Mac Address Prod Next Hop (s)
-------------- ------ ---------------
```

0006.f63f.e3c1

```
 Local Eth1/11
```

! Mac of HOST A

i

nstalled into l2rib.

8c60.4ff2.f541

```
 BGP 192.168.33.34
```

! Mac of HOST C installed into l2rib learnt via BGP.

<#root>

```
 VTEP1# show l2route evpn
```

mac-ip

```
 evi 30
Mac Address Prod Host IP Next Hop (s)
-------------- ---- ------------------------------------ ---------------
0006.f63f.e3c1 HMM 172.16.30.33 N/A
```

8c60.4ff2.f541

```
 BGP
```

172.16.30.2

```
 192.168.33.34
```

! Mac+IP of Host C learnt across the Vxlan Fabric.

<#root>

```
VTEP1# show l2route evpn
```

mac-ip

```
 evi 40
Mac Address Prod Host IP Next Hop (s)
-------------- ---- ------------------------------------ ---------------
```

8c60.4ff2.f541

```
 BGP
```

172.16.40.2

```
 192.168.33.34
```

! Mac+IP of Host D learnt across the Vxlan Fabric.

- Next step is to check that l2rib exports the update to l2vpn evpn.

```
 VTEP1# show bgp l2vpn evpn vni-id 10000030
BGP routing table information for VRF default, address family L2VPN EVPN
BGP table version is 31, local router ID is 192.168.11.11
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-i
njected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup
```

<#root>

```
Network Next Hop Metric LocPrf Weight Path
```

Route Distinguisher: 192.168.11.11:32797 (L2VNI 10000030)

```
*>l[2]:[0]:[0]:[48]:[
```

0006.f63f.e3c1

```
]:[0]:[0.0.0.0]/216
```

!

Mac of Host A in update.

```
 192.168.11.11 100 32768 i
*>i[2]:[0]:[0]:[48]:[8c60.4ff2.f541]:[0]:[0.0.0.0]/216
 192.168.33.34 100 0 i
* i 192.168.33.34 100 0 i
*>l[2]:[0]:[0]:[48]:[
```

0006.f63f.e3c1

```
]:[32]:[
```

172.16.30.33

]/272

!

Mac and IP of Host A in update.

192.168.11.11 100 32768 i
* i[2]:[0]:[0]:[48]:[

8c60.4ff2.f541

]:[32]:[

172.16.30.2

]/272

!

Mac and IP of Host C in update from

Spine.

192.168.33.34 100 0 i
*>i 192.168.33.34 100 0 i

- Next step is to check that the routes are advertised to Spine.

VTEP1# show bgp l2vpn evpn nei 192.168.9.9 advertised-routes

Peer 192.168.9.9 routes for address family L2VPN EVPN:
BGP table version is 31, local router ID is 192.168.11.11
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-i
njected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup

<#root>

Network Next Hop Metric LocPrf Weight Path

Route Distinguisher: 192.168.11.11:32797 (L2VNI 10000030)

*>l[2]:[0]:[0]:[48]:[0006.f63f.e3c1]:[0]:[0.0.0.0]/216
192.168.11.11 100 32768 i
*>l[2]:[0]:[0]:[48]:[

0006.f63f.e3c1

]:[32]:[

172.16.30.33

]/272

! Mac and IP advertised to Spine.

```
 192.168.11.11 100 32768 i
```

- Next step is to check the routes received from Spine.

<#root>

```
 VTEP1# show bgp l2vpn evpn nei 192.168.9.9 routes

Peer 192.168.9.9 routes for address family L2VPN EVPN:
BGP table version is 31, local router ID is 192.168.11.11
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-i
njected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup

Network Next Hop Metric LocPrf Weight Path
Route Distinguisher: 192.168.11.11:32797 (L2VNI 10000030)
*>i[2]:[0]:[0]:[48]:[8c60.4ff2.f541]:[0]:[0.0.0.0]/216
 192.168.33.34 100 0 i
* i 192.168.33.34 100 0 i
* i[2]:[0]:[0]:[48]:[
```

8c60.4ff2.f541

```
]:[32]:[
```

172.16.30.2

```
]/272
```

! This is update from Host C in same VNID.

```
 192.168.33.34 100 0 i
*>i 192.168.33.34 100 0 i

Route Distinguisher: 192.168.11.11:32807 (L2VNI 10000040)
*>i[2]:[0]:[0]:[48]:[8c60.4ff2.f541]:[0]:[0.0.0.0]/216
 192.168.33.34 100 0 i
* i 192.168.33.34 100 0 i
* i[2]:[0]:[0]:[48]:[
```

8c60.4ff2.f541

```
]:[32]:[
```

172.16.40.2

```
]/272
```

!

This is update from Host D in different VNID.

```
 192.168.33.34 100 0 i
*>i 192.168.33.34 100 0 i

Route Distinguisher: 192.168.11.11:3 (L3VNI 10000010)
*>i[2]:[0]:[0]:[48]:[8c60.4ff2.f541]:[32]:[172.16.30.2]/272
 192.168.33.34 100 0 i
```

```
* i 192.168.33.34 100 0 i
*>i[2]:[0]:[0]:[48]:[8c60.4ff2.f541]:[32]:[172.16.40.2]/272
 192.168.33.34 100 0 i
* i 192.168.33.34 100 0 i
*>i[5]:[0]:[0]:[32]:[
```

172.16.2.2

```
]:[0.0.0.0]/224 !
```

! This is update from External Host.

```
 192.168.22.22 100 0 65100 i
```

<#root>

```
VTEP1# show ip bgp vrf
```

EVPN-L3-VNI-VLAN-10

```
BGP routing table information for VRF EVPN-L3-VNI-VLAN-10, address family IPv4 Unicast
BGP table version is 5, local router ID is 192.168.1.254
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-injected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup

Network            Next Hop Metric LocPrf Weight    Path
*>i
```

172.16.2.2/32

```
   192.168.22.22   100      0       65111 i
*>i
```

172.16.30.2/32

```
  192.168.33.34   100      0       i
* i                192.168.33.34   100      0       i
*>i
```

172.16.40.2/32

```
  192.168.33.34   100      0       i
* i                192.168.33.34   100      0       i
```

- In VTEP1, only VNID 10000030 is configured, and it has been verified that mac and ip of Host A are learned locally, and also advertised as evpn route. It has also been observed that the update from Host C has also been received and installed here.

- The nve peers can also be up before the traffic can be forwarded.

<#root>

```
VTEP1# show nve peers
Interface Peer-IP State LearnType Uptime Router-Mac
--------- --------------- ----- --------- -------- -----------------
```

nve1

192.168.22.22

Up CP 01:39:15 0062.ecbf.5325

!

VTEP2

nve1

192.168.33.34

Up CP 01:40:09 f8c2.8823.275f

!

VTEP3 and VTEP4 appear as single entity as both are in Vpc.

<#root>

```
VTEP1# sh bgp internal nve-peer-vni
PeerAddress VNI VrfID    GatewayMAC     TunnelID Encap EgressVNI F
```

192.168.22.22

10000010 1

0062.ecbf.5325

0xc0a81616

1 0 0

192.168.33.34

10000010 1

0062.ecbf.4e4d

0xc0a82122

```
1    0    0
192.168.33.34 10000010 1 f8c2.8823.275f 0xc0a82122 1    0    0
192.168.33.34 10000030 1 0000.0000.0000 0x0        1    0    0
192.168.33.34 10000040 1 0000.0000.0000 0x0        1    0    0
```

Troubleshoot

- If nve interface is checked, then increments of counters for encapsulation and decapsulation can be seen.

<#root>

```
 VTEP1# show interface nve 1
nve1 is up
admin state is up, Hardware: NVE
 MTU 9216 bytes

 Encapsulation VXLAN
 Auto-mdix is turned off
 RX
 ucast: 133 pkts, 22344 bytes - mcast: 0 pkts, 0 bytes
 TX
 ucast: 134 pkts, 22512 bytes - mcast: 0 pkts, 0 bytes
```

- If a transparent firewall is being used for filtering, make the sure the associated port is allowed.

<#root>

```
VTEP1# show nve vxlan-params
VxLAN Dest. UDP Port: 4789
```

- To check the local VTEP mac address which is used for Inter-VNI routing. Secondary address is seen when the VTEP is in a vpc pair.

<#root>

```
 VTEP1# show nve interface
Interface: nve1, State: Up, encapsulation: VXLAN
 VPC Capability: VPC-VIP-Only [not-notified]

Local Router MAC: e00e.da2a.2393

 Host Learning Mode: Control-Plane

Source-Interface: loopback2 (primary: 192.168.11.11, secondary: 0.0.0.0)
```

- To check the remote VTEPS mac addresses and state of peering.

<#root>

```
VTEP1# sh nve internal platform interface nve1 detail
```

```
Printing Interface ifindex 0x49000001 detail
|======|=========================|===============|===============|=====|=====|
|Intf |State |PriIP |SecIP |Vnis |Peers|
|======|=========================|===============|===============|=====|=====|
|nve1 |UP |192.168.11.11 |
```

0.0.0.0

```
 |3 |2 | !
```

Secondary Ip is 0.0.0.0 because this VTEP is not in vpc

```
|======|=========================|===============|===============|=====|=====|

SW_BD/VNIs of interface nve1:
================================================
|======|======|=========================|======|====|======|========
|Sw BD |Vni |State |Intf |Type|Vrf-ID|Notified
|======|======|=========================|======|====|======|========
|
```

10 |10000010|UP |nve1 |CP |3 |Yes

|30 |10000030|UP |nve1 |CP |0 |Yes

|

40 |10000040|UP |nve1 |CP |0 |Yes

```
|======|======|=========================|======|====|======|========

Peers of interface nve1:
==========================================

Peer_ip: 192.168.22.22
 Peer-ID : 1
 State : UP
 Learning : Disabled
 TunnelID : 0xc0a81616
```

MAC : 0062.ecbf.5325

```
 Table-ID : 0x1
 Encap : 0x1

Peer_ip: 192.168.33.34 ! For both VTEP3 and VTEP4
 Peer-ID : 2
 State : UP
 Learning : Disabled
 TunnelID : 0xc0a82122
```

MAC : 0062.ecbf.4e4d

```
 Table-ID : 0x1
 Encap : 0x1
```

- To check the peering time and VNI information for VTEP Peers.

```
 VTEP1# show nve peer detail
Details of nve Peers:
---------------------------------------

Peer-Ip: 192.168.22.22
 NVE Interface : nve1
 Peer State : Up
 Peer Uptime : 00:22:17

Router-Mac : 0062.ecbf.5325
 Peer First VNI : 10000010
 Time since Create : 00:22:17

Configured VNIs : 10000010,10000030,10000040
 Provision State : add-complete ! Hardware ready for forwarding.
 Route-Update : Yes
 Peer Flags : RmacL2Rib, TunnelPD, DisableLearn
 Learnt CP VNIs : 10000010
 Peer-ifindex-resp : Yes
---------------------------------------
Peer-Ip: 192.168.33.34
 NVE Interface : nve1
 Peer State : Up
 Peer Uptime : 00:22:10
 Router-Mac : 0062.ecbf.4e4d
 Peer First VNI : 10000010
 Time since Create : 00:22:10
 Configured VNIs : 10000010,10000030,10000040
 Provision State : add-complete ! Hardware ready for forwarding.
 Route-Update : Yes
 Peer Flags : RmacL2Rib, TunnelPD, DisableLearn
 Learnt CP VNIs : 10000010,10000030,10000040
 Peer-ifindex-resp : Yes
---------------------------------------
```

- To check how BGP interacts with EVI and internal information is built. Example of vlan 30 mapped to VNI 10000030 is shown here.

<#root>

VTEP1#

sh bgp internal evi 10000030

```
**************************************************
L2RIB bound / VNI Req to L2RIB : Yes / 1
L2VNI Adds / Dels / ALL Dels from L2RIB : 4 / 3 / 1
 First L2VNI Add/Del : Dec 17 19:07:41.680736 / Dec 17 19:10:48.455562
 Last L2VNI Add/Del : Dec 17 19:11:13.916893 / Dec 17 19:10:48.455792
L3VNI Adds / Dels from L2RIB : 2 / 0 / 1
 First L3VNI Add/Del : Dec 17 19:07:41.681313 / never
 Last L3VNI Add/Del : Dec 17 19:11:11.838315 / never
 First/Last All VNI Del : Dec 17 19:10:48.455542 / Dec 17 19:10:48.455543
ALL VNI Del from L2RIB state (cleanup status) : All VNI Not Start (0x000006)
```

```
All VNI down loop count : 0
L2RIB is up/registered/local-req: 1/1
L2RIB down: in-prg/up-defer: 0/0
L2RIB register/failures: 1/0
L2RIB deregister/failures: 0/0
L2RIB flow control (#enabled/#disabled): Disabled (0/0)
**************************************************
```

BGP L2VPN/EVPN RD Information for 192.168.11.11:32797

```
 L2VNI ID : 10000030 (evi_10000030)
 #Prefixes Local/BRIB : 2 / 4
 #Paths L3VPN->EVPN/EVPN->L3VPN : 129 / 0
**************************************************
==========================================
BGP Configured VNI Information:
 evi_cfg : 0xd87786c8
```

VNI ID (Index) : 10000030

```
 (1)
```

RD : 192.168.11.11:32797

Export RTs : 1

```
 ExportRT cfg list:
 65000:10000030 (auto)
 Import RTs : 1
 ImportRT cfg list:
 65000:10000030 (auto)
 Topo Id : 30
```

VTEP IP : 192.168.11.11

```
 VTEP VPC IP : 0.0.0.0
 Encap Type : 8
 Refcount : #00000003
```

Enabled : Yes ! If this is no then check the NVE interface config for this VNID

```
 Delete Pending : No
 Creation Req : No
 Future RD : NULL
 evi_ctx : 0xd86e554c
 RD/Import RT/Export RT : Yes(Auto)/Yes/Yes
 MAC First Add/Del : Dec 17 19:11:12.45086 / never
 MAC Last Add/Del : Dec 17 19:11:12.45086 / never
 MAC IP First Add/Del : Dec 17 19:11:12.54976 / never
 MAC IP Last Add/Del : Dec 17 19:11:12.54977 / never
 IMET First Add/Del : never / never
 IMET Last Add/Del : never / never

==========================================
++++++++++++++++++++++++++++++++++++++++++
BGP VNI Information for evi_10000030 (0xd86e554c)
 L2VNI ID : 10000030 (evi_10000030)
 RD (rdinfo) : 192.168.11.11:32797 (0xd8811eb0)
 Prefixes (local/total) : 2/4
```

```
 Created : Dec 17 19:11:12.37640
 Last Oper Up/Down : Dec 17 19:11:12.37827 / never
 Enabled : Yes
 Delete pending : 0
 Stale : No
 Import pending : 0
 Import in progress : 0
 Encap : VxLAN
 Topo Id : 30
 VTEP IP : 192.168.11.11
 VTEP VPC IP : 0.0.0.0
Router-MAC : 0000.0000.0000
 Active Export RTs : 1

 Active Export RT list : 65000:10000030


 Config Export RTs : 1
 ExportRT cfg list:
 65000:10000030 (auto)
 Export RT chg/chg-pending : 0/0
 Active Import RTs : 1

 Active Import RT list : 65000:10000030


 Config Import RTs : 1
 ImportRT cfg list:
 65000:10000030 (auto)
 Import RT chg/chg-pending : 0/0
 IMET Reg/Unreg from L2RIB : 2/0
 MAC Reg/Unreg from L2RIB : 2/0
 MAC IP Reg/Unreg from L2RIB : 2/0
 IMET Add/Del from L2RIB : 0/0
 MAC Add/Del from L2RIB : 1/0
 MAC IP Add/Del from L2RIB : 1/0
 IMET Dnld/Wdraw to L2RIB : 0/0
 MAC Dnld/Wdraw to L2RIB : 1/0
 MAC IP Dnld/Wdraw to L2RIB : 1/0
```

- When an update is received, irrespective of the fact of it being an Inter-VNI or Intra-VNI update, make sure that correct Route Targets (RTs) are being received and VTEP receiving the update has relevant configs. An update from VTEP3 coming via SPINE is analysed here for RT consistency. The local state of RT and RD for VTEP1 has been shown in these outputs.
  <#root>

  ```
  SPINE# show bgp l2vpn evpn 172.16.30.2 !

  Update from Spine


  BGP routing table information for VRF default, address family L2VPN EVPN
  Route Distinguisher: 3.3.3.3:32797
  BGP routing table entry for [2]:[0]:[0]:[48]:[8c60.4ff2.f541]:[32]:[172.16.30.2]
  /272, version 25
  Paths: (1 available, best #1)
  Flags: (0x000202) on xmit-list, is not in l2rib/evpn, is not in HW,

  Advertised path-id 1
   Path type: internal, path is valid, is best path, remote nh not installed, no
  labeled nexthop
  ```

```
 AS-Path: NONE, path sourced internal to AS
 192.168.33.34 (metric 5) from 192.168.33.33 (3.3.3.3)
 Origin IGP, MED not set, localpref 100, weight 0
 Received label 10000030 1000001

 Extcommunity: RT:65000:10000010 RT:65000:10000030 SOO:192.168.33.34:0 ENC


AP:8 Router MAC:0062.ecbf.4e4d

Path-id 1 advertised to peers:

 192.168.11.11 192.168.22.22 192.168.44.44
```