Configuring VXLAN BGP EVPN

This chapter contains the following sections:

Information About VXLAN BGP EVPN

Guidelines and Limitations for VXLAN BGP EVPN

VXLAN BGP EVPN has the following guidelines and limitations:

- Routing between VXLAN VLANs and non-VXLAN VLANs, and Layer 3 interfaces, is not supported on Cisco Nexus 3100-V platform switches. Hence, Cisco Nexus 3100-V platform switches cannot be a border leaf VTEP in a VXLAN EVPN setup.
- You can configure EVPN over segment routing or MPLS. See the Cisco Nexus 3000 Series NX-OS Label Switching Configuration Guide for more information.
- You can use MPLS tunnel encapsulation using the new CLI encapsulation **mpls** command. You can configure the label allocation mode for the EVPN address family. See the Cisco Nexus 3000 Series NX-OS Label Switching Configuration Guide for more information.
- In a VXLAN EVPN setup that has a 2K VNI scale configuration, the control plane down time takes more than 200 seconds. To avoid BGP flap, configure the graceful restart time to 300 seconds.
- SVI and sub-interfaces as core links are not supported in multisite EVPN.
- In a VXLAN EVPN setup, border leaves must use unique route distinguishers, preferably using **auto rd** command. It is not supported to have same route distinguishers in different border leaves.
- ARP suppression is only supported for a VNI if the VTEP hosts the First-Hop Gateway (Distributed Anycast Gateway) for this VNI. The VTEP and the SVI for this VLAN have to be properly configured for the distributed Anycast Gateway operation, for example, global Anycast Gateway MAC address configured and Anycast Gateway feature with the virtual IP address on the SVI.

- When Layer 3 EVPN is configured in Cisco Nexus 3000 platform switches that are based on Broadcom ASIC and these switches are added in the topology with Layer 2 EVPN, the routing for this scenario is not supported. When you configure SVI and Layer 3 EVPN on Cisco Nexus 3000 platform switches based on Broadcom ASIC with Anycast Gateway and when you send the ARP requests from a Layer 2 EVPN device (for example, Cisco Nexus 3000 platform switches, based on a Broadcom ASIC), the Cisco Nexus 3000 platform switches can not be used as a gateway for the ARP requests received on the network ports.
- The **show** commands with the **internal** keyword are not supported.
- DHCP snooping (Dynamic Host Configuration Protocol snooping) is not supported on VXLAN VLANs.
- SPAN TX for VXLAN encapsulated traffic is not supported for the Layer 3 uplink interface.
- RACLs are not supported on Layer 3 uplinks for VXLAN traffic. Egress VACLs support is not available for de-capsulated packets in the network to access direction on the inner payload.

  As a best practice, use PACLs/VACLs for the access to the network direction.
- QoS classification is not supported for VXLAN traffic in the network to access direction on the Layer 3 uplink interface.
- The QoS buffer-boost feature is not applicable for VXLAN traffic.
- VTEP does not support Layer 3 subinterface uplinks that carry VXLAN encapsulated traffic.
- Layer 3 interface uplinks that carry VXLAN encapsulated traffic do not support subinterfaces for non-VxLAN encapsulated traffic.
- Non-VXLAN sub-interface VLANs cannot be shared with VXLAN VLANs.
- Subinterfaces on 40G (ALE) uplink ports are not supported on VXLAN VTEPs.
- Point to multipoint Layer 3 and SVI uplinks are not supported. Since both uplink types can only be enabled point-to-point, they cannot span across more than two switches.
- For EBGP, it is recommended to use a single overlay EBGP EVPN session between loopbacks.
- Bind NVE to a loopback address that is separate from other loopback addresses that are required by Layer 3 protocols. A best practice is to use a dedicated loopback address for VXLAN.
- VXLAN BGP EVPN does not support an NVE interface in a non-default VRF.
- It is recommended to configure a single BGP session over the loopback for an overlay BGP session.
- The VXLAN UDP port number is used for VXLAN encapsulation. For Cisco Nexus NX-OS, the UDP port number is 4789. It complies with IETF standards and is not configurable.
- VXLAN supports In Service Software Upgrade (ISSU).
- VTEP connected to FEX host interface ports is not supported.
- Resilient hashing (port-channel load-balancing resiliency) and VXLAN configurations are not compatible with VTEPs using ALE uplink ports.

Note Resilient hashing is disabled by default.

Note For information about VXLAN BGP EVPN scalability, see the Verified Scalability Guide for your platform.

Notes for EVPN Convergence

The following are notes about EVPN Convergence (7.0(3)I3(1) and later):

- As a best practice, the NVE source loopback should be dedicated to NVE. so that NVE can bring the loopback up or down as needed.
- When vPC has been configured, the loopback stays down until the MCT link comes up.

  Note When **feature vpc** is enabled and there is no VPC configured, the NVE source loopback is in "shutdown" state after an upgrade. In this case, removing **feature vpc** restores the interface to "up" state."

- The NVE underlay (through the source loopback) is kept down until the overlay has converged.
  - When MCT comes up, the source loopback is kept down for an amount of time that is configurable. This approach prevents north-south traffic from coming in until the overlay has converged.
  - When MCT goes down, NVE is kept up for 30 seconds in the event that there is still south-north traffic from vPC legs which have not yet gone down.
- BGP ignores routes from vPC peer. This reduces the number of routes in BGP.

Considerations for VXLAN BGP EVPN Deployment

- A loopback address is required when using the **source-interface config** command. The loopback address represents the local VTEP IP.
- During boot-up of a switch (7.0(3)I2(2) and later), you can use the **source-interface hold-down-time** *hold-down-time* command to suppress advertisement of the NVE loopback address until the overlay has converged. The range for the *hold-down-time* is 0 - 2147483647 seconds. The default is 300 seconds.
- To establish IP multicast routing in the core, IP multicast configuration, PIM configuration, and RP configuration is required.
- VTEP to VTEP unicast reachability can be configured through any IGP/BGP protocol.
- If the anycast gateway feature is enabled for a specific VNI, then the anyway gateway feature must be enabled on all VTEPs that have that VNI configured. Having the anycast gateway feature configured on only some of the VTEPs enabled for a specific VNI is not supported.
- It is a requirement when changing the primary or secondary IP address of the NVE source interfaces to shut the NVE interface before changing the IP address.
- As a best practice, the RP for the multicast group should be configured only on the spine layer. Use the anycast RP for RP load balancing and redundancy.
- Every tenant VRF needs a VRF overlay VLAN and SVI for VXLAN routing.

- When configuring ARP suppression with BGP-EVPN, use the **hardware access-list tcam region arp-ether** *size* **double-wide** command to accommodate ARP in this region. (You must decrease the size of an existing TCAM region before using this command.)

VPC Considerations for VXLAN BGP EVPN Deployment

- The loopback address used by NVE needs to be configured to have a primary IP address and a secondary IP address.

  The secondary IP address is used for all VxLAN traffic that includes multicast and unicast encapsulated traffic.

- Each VPC peer needs to have separate BGP sessions to the spine.
- VPC peers must have identical configurations.
  - Consistent VLAN to VN-segment mapping.
  - Consistent NVE1 binding to the same loopback interface
    - Using the same secondary IP address.
    - Using different primary IP addresses.
  - Consistent VNI to group mapping.
  - The VRF overlay VLAN should be a member of the peer-link port-channel.
- For multicast, the VPC node that receives the (S, G) join from the RP (rendezvous point) becomes the DF (designated forwarder). On the DF node, encap routes are installed for multicast.

  Decap routes are installed based on the election of a decapper from between the VPC primary node and the VPC secondary node. The winner of the decap election is the node with the least cost to the RP. However, if the cost to the RP is the same for both nodes, the VPC primary node is elected.

  The winner of the decap election has the decap mroute installed. The other node does not have a decap route installed.

- On a VPC device, BUM traffic (broadcast, unknown-unicast, and multicast traffic) from hosts is replicated on the peer-link. A copy is made of every native packet and each native packet is sent across the peer-link to service orphan-ports connected to the peer VPC switch.

  To prevent traffic loops in VXLAN networks, native packets ingressing the peer-link cannot be sent to an uplink. However, if the peer switch is the encapper, the copied packet traverses the peer-link and is sent to the uplink.

  Note Each copied packet is sent on a special internal VLAN (VLAN 4041).

- When peer-link is shut, the loopback interface used by NVE on the VPC secondary is brought down and the status is **Admin Shut**. This is done so that the route to the loopback is withdrawn on the upstream and that the upstream can divert all traffic to the VPC primary.

Note Orphans connected to the VPC secondary will experience loss of traffic for the period that the peer-link is shut. This is similar to Layer 2 orphans in a VPC secondary of a traditional VPC setup.

- When peer-link is no-shut, the NVE loopback address is brought up again and the route is advertised upstream, attracting traffic.
- For VPC, the loopback interface has 2 IP addresses: the primary IP address and the secondary IP address.

  The primary IP address is unique and is used by Layer 3 protocols.

  The secondary IP address on loopback is necessary because the interface NVE uses it for the VTEP IP address. The secondary IP address must be same on both vPC peers.
- The VPC peer-gateway feature must be enabled on both peers.

  As a best practice, use peer-switch, peer gateway, ip arp sync, ipv6 nd sync configurations for improved convergence in VPC topologies.

  In addition, increase the STP hello timer to 4 seconds to avoid unnecessary TCN generations when VPC role changes occur.

  The following is an example (best practice) of a VPC configuration:

```
switch# sh ru vpc

version 6.1(2)I3(1)
feature vpc
vpc domain 2
  peer-switch
  peer-keepalive destination 172.29.206.65 source 172.29.206.64
  peer-gateway
  ipv6 nd synchronize
  ip arp synchronize
```

- On a VPC pair, shutting down NVE or NVE loopback on one of the VPC nodes is not a supported configuration. This means that traffic failover on one-side NVE shut or one-side loopback shut is not supported.
- Redundant anycast RPs configured in the network for multicast load-balancing and RP redundancy are supported on VPC VTEP topologies.
- Enabling vpc peer-gateway configuration is mandatory. For peer-gateway functionality, at least one backup routing SVI is required to be enabled across peer-link and also configured with PIM. This provides a backup routing path in the case when VTEP loses complete connectivity to the spine. Remote peer reachability is re-routed over the peer-link in this case.

  The following is an example of SVI with PIM enabled:

```
swithch# sh ru int vlan 2

interface Vlan2
  description special_svi_over_peer-link
  no shutdown
  ip address 30.2.1.1/30
```

```
ip pim sparse-mode
```

Note The SVI must be configured on both VPC peers and requires PIM to be enabled.

- As a best practice when changing the secondary IP address of an anycast VPC VTEP, the NVE interfaces on both the VPC primary and the VPC secondary should be shut before the IP changes are made.
- To provide redundancy and failover of VXLAN traffic when a VTEP loses all of its uplinks to the spine, it is recommended to run a Layer 3 link or an SVI link over the peer-link between VPC peers.
- If DHCP Relay is required in VRF for DHCP clients or if loopback in VRF is required for reachability test on a VPC pair, it is necessary to create a backup SVI per VRF with PIM enabled.

```
switchc# sh ru int vlan 20

interface Vlan20
description backup routing svi for VRF Green
vrf member GREEN
no shutdown
ip address 30.2.10.1/30
```

Network Considerations for VXLAN Deployments

- MTU Size in the Transport Network

  Due to the MAC-to-UDP encapsulation, VXLAN introduces 50-byte overhead to the original frames. Therefore, the maximum transmission unit (MTU) in the transport network needs to be increased by 50 bytes. If the overlays use a 1500-byte MTU, the transport network needs to be configured to accommodate 1550-byte packets at a minimum. Jumbo-frame support in the transport network is required if the overlay applications tend to use larger frame sizes than 1500 bytes.

- ECMP and LACP Hashing Algorithms in the Transport Network

  As described in a previous section, Cisco Nexus 3000 Series Switches introduce a level of entropy in the source UDP port for ECMP and LACP hashing in the transport network. As a way to augment this implementation, the transport network uses an ECMP or LACP hashing algorithm that takes the UDP source port as an input for hashing, which achieves the best load-sharing results for VXLAN encapsulated traffic.

- Multicast Group Scaling

  The VXLAN implementation on Cisco Nexus 3000 Series Switches uses multicast tunnels for broadcast, unknown unicast, and multicast traffic forwarding. Ideally, one VXLAN segment mapping to one IP multicast group is the way to provide the optimal multicast forwarding. It is possible, however, to have multiple VXLAN segments share a single IP multicast group in the core network. VXLAN can support up to 16 million logical Layer 2 segments, using the 24-bit VNID field in the header. With one-to-one mapping between VXLAN segments and IP multicast groups, an increase in the number of VXLAN segments causes a parallel increase in the required multicast address space and the amount of forwarding states on the core network devices. At some point, multicast scalability in the transport network can become a concern. In this case, mapping multiple VXLAN segments to a single multicast group can help conserve multicast control plane resources on the core devices and achieve the desired VXLAN scalability. However, this mapping comes at the cost of suboptimal multicast forwarding. Packets forwarded to the

multicast group for one tenant are now sent to the VTEPs of other tenants that are sharing the same multicast group. This causes inefficient utilization of multicast data plane resources. Therefore, this solution is a trade-off between control plane scalability and data plane efficiency.

Despite the suboptimal multicast replication and forwarding, having multiple-tenant VXLAN networks to share a multicast group does not bring any implications to the Layer 2 isolation between the tenant networks. After receiving an encapsulated packet from the multicast group, a VTEP checks and validates the VNID in the VXLAN header of the packet. The VTEP discards the packet if the VNID is unknown to it. Only when the VNID matches one of the VTEP's local VXLAN VNIDs, does it forward the packet to that VXLAN segment. Other tenant networks will not receive the packet. Thus, the segregation between VXLAN segments is not compromised.

Considerations for the Transport Network

The following are considerations for the configuration of the transport network:

- On the VTEP device:
  - Enable and configure IP multicast.*
  - Create and configure a loopback interface with a /32 IP address.

    (For vPC VTEPs, you must configure primary and secondary /32 IP addresses.)
  - Enable IP multicast on the loopback interface.*
  - Advertise the loopback interface /32 addresses through the routing protocol (static route) that runs in the transport network.
  - Enable IP multicast on the uplink outgoing physical interface.*
- Throughout the transport network:
  - Enable and configure IP multicast.*
- When using SVI uplinks with VXLAN enabled on Cisco Nexus 9200 and 9300-EX platform switches, use the **system nve infra-vlans** command to specify the VLANs that are used for uplink SVI. Failing to specify the VLANs results in traffic loss.

Note

- The **system nve infra-vlans** command specifies VLANs used by all SVI interfaces for uplink and vPC peer-links in VXLAN as infra-VLANs.
- You should not configure certain combinations of infra-VLANs. For example, 2 and 514, 10 and 522, which are 512 apart.

Note * Not required for static ingress replication or BGP EVPN ingress replication.

BGP EVPN Considerations for VXLAN Deployment

Commands for BGP EVPN

The following describes commands to support BGP EVPN VXLAN control planes.

Command	Description
member vni *range* [**associate-vrf**]	Associate VXLAN VNIs (Virtual Network Identifiers) with the NVE interface The attribute **associate- vrf** is used to identify and separate processing VNIs that are associated with a VRF and used for routing. **Note** The VRF and VNI specified with this command must match the configuration of the VNI under the VRF.
show nve vni **show nve vni summary**	Displays information that determine if the VNI is configured for peer and host learning via the control plane or data plane.
show bgp l2vpn evpn **show bgp l2vpn evpn summary**	Displays the Layer 2 VPN EVPN address family.
host-reachability protocol bgp	Specifies BGP as the mechanism for host reachability advertisement.
suppress-arp	Suppresses ARP under Layer 2 VNI.
fabric forwarding anycast-gateway-mac	Configures anycast gateway MAC of the switch.
vrf context	Creates the VRF and enter the VRF mode.
nv overlay evpn	Enables/Disables the Ethernet VPN (EVPN).
router bgp	Configures the Border Gateway Protocol (BGP).

Command	Description
suppress mac-route	Suppresses the BGP MAC route so that BGP only sends the MAC/IP route for a host. Under NVE, the MAC updates for all VNIs are suppressed. **Note** • Receive-side — Suppressing the MAC route depends upon the capability of the remote EVPN peer to derive a MAC route from the MAC/IP route (7.0(3)I2(2) and later). Avoid using the "suppress mac-route" command if devices in the network are running an earlier NX-OS release. • Send-side — Suppressing the MAC route means that the sender has a MAC/IP route. If your configuration has pure-Layer 2 VNIs (such as no corresponding VRF or Layer3-VNI), then there is no corresponding MAC/IP and you should avoid using the "suppress mac-route" command.

Configuring VXLAN BGP EVPN

Enabling VXLAN

Enable VXLAN and the EVPN.

Procedure

	Command or Action	Purpose
Step 1	**feature vn-segment**	Enable VLAN-based VXLAN
Step 2	**feature nv overlay**	Enable VXLAN
Step 3	**nv overlay evpn**	Enable the EVPN control plane for VXLAN.

Configuring VLAN and VXLAN VNI

Procedure

	Command or Action	Purpose
Step 1	**vlan** *number*	Specify VLAN.
Step 2	**vn-segment** *number*	Map VLAN to VXLAN VNI to configure Layer 2 VNI under VXLAN VLAN.

Configuring VRF for VXLAN Routing

Configure the tenant VRF.

Procedure

	Command or Action	Purpose
Step 1	**vrf context** *vxlan*	Configure the VRF.
Step 2	**vni** *number*	Specify VNI.
Step 3	**rd auto**	Specify VRF RD (route distinguisher).
Step 4	**address-family ipv4 unicast**	Configure address family for IPv4.
Step 5	**route-target both auto**	**Note** Specifying the **auto** option is applicable only for IBGP. Manually configured route targets are required for EBGP.
Step 6	**route-target both auto evpn**	**Note** Specifying the **auto** option is applicable only for IBGP. The **auto** option is available beginning with Cisco NX-OS Release 7.0(3)I7(1). Manually configured route targets are required for EBGP.
Step 7	**address-family ipv6 unicast**	Configure address family for IPv6.
Step 8	**route-target both auto**	**Note** Specifying the **auto** option is applicable only for IBGP. The **auto** option is available beginning with Cisco NX-OS Release 7.0(3)I7(1). Manually configured route targets are required for EBGP.

	Command or Action	Purpose
Step 9	**route-target both auto evpn**	**Note** Specifying the **auto** option is applicable only for IBGP. Manually configured route targets are required for EBGP.

Configuring SVI for Hosts for VXLAN Routing

Configure the SVI for hosts.

Procedure

	Command or Action	Purpose
Step 1	**vlan** *number*	Specify VLAN
Step 2	**interface** *vlan-number*	Specify VLAN interface.
Step 3	**vrf member** *vxlan-number*	Configure SVI for host.
Step 4	**ip address** *address*	Specify IP address.

Configuring VRF Overlay VLAN for VXLAN Routing

Procedure

	Command or Action	Purpose
Step 1	**vlan** *number*	Specify VLAN.
Step 2	**vn-segment** *number*	Specify vn-segment.

Configuring VNI Under VRF for VXLAN Routing

Configures a Layer 3 VNI under a VRF overlay VLAN. (A VRF overlay VLAN is a VLAN that is not associated with any server facing ports. All VXLAN VNIs that are mapped to a VRF, need to have their own internal VLANs allocated to it.)

Procedure

	Command or Action	Purpose
Step 1	**vrf context** *vxlan*	Create a VXLAN Tenant VRF
Step 2	**vni** *number*	Configure Layer 3 VNI under VRF.

Configuring Anycast Gateway for VXLAN Routing

Procedure

	Command or Action	Purpose
Step 1	**fabric forwarding anycast-gateway-mac** *address*	Configure distributed gateway virtual MAC address **Note** One virtual MAC per VTEP **Note** All VTEPs should have the same virtual MAC address
Step 2	**fabric forwarding mode anycast-gateway**	Associate SVI with anycast gateway under VLAN configuration mode.

Configuring the NVE Interface and VNIs

Procedure

	Command or Action	Purpose
Step 1	**interface** *nve-interface*	Configure the NVE interface.
Step 2	**host-reachability protocol bgp**	This defines BGP as the mechanism for host reachability advertisement
Step 3	**member vni** *vni* **associate-vrf**	Add Layer-3 VNIs, one per tenant VRF, to the overlay. **Note** Required for VXLAN routing only.
Step 4	**global mcast-group** *ip-address*	
Step 5	**member vni** *vni*	Add Layer 2 VNIs to the tunnel interface.
Step 6	**mcast-group** *address*	Configure the mcast group on a per-VNI basis

Configuring BGP on the VTEP

Procedure

	Command or Action	Purpose
Step 1	**router bgp** *number*	Configure BGP.
Step 2	**router-id** *address*	Specify router address.

	Command or Action	Purpose
Step 3	**neighbor** *address* **remote-as** *number*	Define MP-BGP neighbors. Under each neighbor define l2vpn evpn.
Step 4	**address-family ipv4 unicast**	Configure address family for IPv4.
Step 5	**address-family l2vpn evpn**	Configure address family Layer 2 VPN EVPN under the BGP neighbor. **Note** Address-family ipv4 evpn for vxlan host-based routing
Step 6	(Optional) **Allowas-in**	Allows duplicate AS numbers in the AS path. Configure this parameter on the leaf for eBGP when all leafs are using the same AS, but the spines have a different AS than leafs.
Step 7	**send-community extended**	Configures community for BGP neighbors.
Step 8	**vrf** *vrf-name*	Specify VRF.
Step 9	**address-family ipv4 unicast**	Configure address family for IPv4.
Step 10	**advertise** *l2vpn* **evpn**	Enable advertising EVPN routes. **Note** Beginning with Cisco NX-OS Release 9.2(1), the **advertise l2vpn evpn** command no longer takes effect. To disable advertisement for a VRF toward the EVPN, disable the VNI in NVE by entering the **no member vni** *vni* **associate-vrf** command in interface nve1. The *vni* is the VNI associated with that particular VRF.
Step 11	**address-family ipv6 unicast**	Configure address family for IPv6.
Step 12	**advertise** *l2vpn* **evpn**	Enable advertising EVPN routes.

Configuring RD and Route Targets for VXLAN Bridging

Procedure

	Command or Action	Purpose
Step 1	**evpn**	Configure VRF.
Step 2	**vni** *number* **l2**	**Note** Only Layer 2 VNIs need to be specified.

	Command or Action	Purpose
Step 3	**rd auto**	Define VRF RD (route distinguisher) to configure VRF context.
Step 4	**route-target import auto**	Define VRF Route Target and import policies.
Step 5	**route-target export auto**	Define VRF Route Target and export policies.

Configuring BGP for EVPN on the Spine

Procedure

	Command or Action	Purpose
Step 1	**route-map permitall permit 10**	Configure route-map. **Note** The route-map keeps the next-hop unchanged for EVPN routes. • Required for eBGP. • Optional for iBGP.
Step 2	**set ip next-hop unchanged**	Set next-hop address. **Note** The route-map keeps the next-hop unchanged for EVPN routes. • Required for eBGP. • Optional for iBGP. **Note** When two next hops are enabled, next hop ordering is not maintained. If one of the next hops is a VXLAN next hop and the other next hop is local reachable via FIB/AM/Hmm, the local next hop reachable via FIB/AM/Hmm is always taken irrespective of the order. Directly/locally connected next hops are always given priority over remotely connected next hops.
Step 3	**router bgp** *autonomous system number*	Specify BGP.
Step 4	**address-family l2vpn evpn**	Configure address family Layer 2 VPN EVPN under the BGP neighbor.
Step 5	**retain route-target all**	Configure retain route-target all under address-family Layer 2 VPN EVPN [global].

	Command or Action	Purpose
		Note Required for eBGP. Allows the spine to retain and advertise all EVPN routes when there are no local VNI configured with matching import route targets.
Step 6	**neighbor** *address* **remote-as** *number*	Define neighbor.
Step 7	**address-family l2vpn evpn**	Configure address family Layer 2 VPN EVPN under the BGP neighbor.
Step 8	**disable-peer-as-check**	Disables checking the peer AS number during route advertisement. Configure this parameter on the spine for eBGP when all leafs are using the same AS but the spines have a different AS than leafs. **Note** Required for eBGP.
Step 9	**send-community extended**	Configures community for BGP neighbors.
Step 10	**route-map permitall out**	Applies route-map to keep the next-hop unchanged. **Note** Required for eBGP.

Suppressing ARP

Suppressing ARP includes changing the size of the ACL ternary content addressable memory (TCAM) regions in the hardware.

Procedure

	Command or Action	Purpose
Step 1	**hardware access-list tcam region arp-ether** *size* **double-wide**	Configure TCAM region to suppress ARP. *tcam-size*—TCAM size. The size has to be a multiple of 256. If the size is more than 256, it has to be a multiple of 512. **Note** Reload is required for the TCAM configuration to be in effect.
Step 2	**interface nve 1**	Create the network virtualization endpoint (NVE) interface.
Step 3	**member vni** *vni-id*	Specify VNI ID.
Step 4	**suppress-arp**	Configure to suppress ARP under Layer 2 VNI.

	Command or Action	Purpose
Step 5	**copy running-config start-up-config**	Saves the change persistently through reboots and restarts by copying the running configuration to the startup configuration.

Disabling VXLANs

Procedure

	Command or Action	Purpose
Step 1	**configure terminal**	Enters configuration mode.
Step 2	**no nv overlay evpn**	Disables EVPN control plane.
Step 3	**no feature vn-segment-vlan-based**	Disables the global mode for all VXLAN bridge domains
Step 4	**no feature nv overlay**	Disables the VXLAN feature.
Step 5	(Optional) **copy running-config startup-config**	Saves the change persistently through reboots and restarts by copying the running configuration to the startup configuration.

Duplicate Detection for IP and MAC Addresses

Cisco NX-OS supports duplicate detection for IP and MAC addresses. This enables the detection of duplicate IP or MAC addresses based on the number of moves in a given time-interval (seconds).

The default is 5 moves in 180 seconds. (Default number of moves is 5 moves. Default time-interval is 180 seconds.)

- For IP addresses:
  - After the 5th move within 180 seconds, the switch starts a 30 second lock (hold down timer) before checking to see if the duplication still exists (an effort to prevent an increment of the sequence bit). This 30 second lock can occur 5 times within 24 hours (this means 5 moves in 180 seconds for 5 times) before the switch permanently locks or freezes the duplicate entry. (**show fabric forwarding ip local-host-db vrf abc**)
- For MAC addresses:
  - After the 5th move within 180 seconds, the switch starts a 30 second lock (hold down timer) before checking to see if the duplication still exists (an effort to prevent an increment of the sequence bit). This 30 second lock can occur 3 times within 24 hours (this means 5 moves in 180 seconds for 3 times) before the switch permanently locks or freezes the duplicate entry. (**show l2rib internal permanently-frozen-list**)
- Wherever a MAC address is permanently frozen, a syslog message with written by L2RIB.

```
2017 Jul  5 10:27:34 leaf %$ VDC-1 %$  %USER-2-SYSTEM_MSG: Unfreeze limit (3) hit, MAC
```

```
 0000.0033.3333in topo: 200 is permanently frozen - l2rib
2017 Jul  5 10:27:34 leaf %$ VDC-1 %$  %USER-2-SYSTEM_MSG: Detected duplicate host
0000.0033.3333, topology 200, during Local update, with host located at remote VTEP
1.2.3.4, VNI 2 - l2rib
2017 Jul  5 10:27:34 leaf %$ VDC-1 %$  %USER-2-SYSTEM_MSG: Unfreeze limit (3) hit, MAC
 0000.0033.3334in topo: 200 is permanently frozen - l2rib
2017 Jul  5 10:27:34 leaf %$ VDC-1 %$  %USER-2-SYSTEM_MSG: Detected duplicate host
0000.0033.3334, topology 200, during Local update, with host l
```

The following are example commands to help the configuration of the number of VM moves in a specific time interval (seconds) for duplicate IP-detection:

Command	Description

`switch(config)# fabric forwarding ?`
`anycast-gateway-mac`
`dup-host-ip-addr-detection`
	Available sub-commands:
	• Anycast gateway MAC of the switch.
	• To detect duplicate host addresses in n seconds.

`switch(config)# fabric forwarding dup-host-ip-addr-detection ?` `<1-1000>`	The number of host moves allowed in n seconds. The range is 1 to 1000 moves; default is 5 moves.
`switch(config)# fabric forwarding dup-host-ip-addr-detection 100 ?` `<2-36000>`	The duplicate detection timeout in seconds for the number of host moves. The range is 2 to 36000 seconds; default is 180 seconds.
`switch(config)# fabric forwarding dup-host-ip-addr-detection 100 10`	Detects duplicate host addresses (limited to 100 moves) in a period of 10 seconds.

The following are example commands to help the configuration of the number of VM moves in a specific time interval (seconds) for duplicate MAC-detection:

Command	Description

`switch(config)# l2rib dup-host-mac-detection ?`
`<1-1000>`
`default`
	Available sub-commands for L2RIB:
	• The number of host moves allowed in n seconds. The range is 1 to 1000 moves.
	• Default setting (5 moves in 180 in seconds).

`switch(config)# l2rib dup-host-mac-detection 100 ?` `<2-36000>`	The duplicate detection timeout in seconds for the number of host moves. The range is 2 to 36000 seconds; default is 180 seconds.

Command	Description
`switch(config)# l2rib dup-host-mac-detection 100 10`	Detects duplicate host addresses (limited to 100 moves) in a period of 10 seconds.

Enabling Nuage Controller Interoperability

The following steps enable Nuage controller interoperability.

Procedure

	Command or Action	Purpose
Step 1	**nuage controller interop**	Global command to enable interoperability mode.
Step 2	**router bgp** *number*	Configure BGP.
Step 3	**address-family l2vpn evpn**	Configure address family Layer 2 VPN EVPN under the BGP neighbor.
Step 4	**advertise-system-mac**	Enable Nuage interoperability mode for BGP.
Step 5	**allow-vni-in-ethertag**	Enable Nuage interoperability mode for BGP.
Step 6	**route-map permitall permit 10**	Configure route-map to permit all.
Step 7	**router bgp** *number*	Configure BGP.
Step 8	**vrf** *vrf-name*	Specify tenant VRF.
Step 9	**address-family ipv4 unicast**	Configure address family for IPv4.
Step 10	**advertise l2vpn evpn**	Enable advertising EVPN routes.
Step 11	**redistribute hmm route-map permitall**	Enables advertise host tenant routes as evpn type-5 routes for interoperability.

Example

The following is an example to enable Nuage controller interoperability:

```
/*** Enable interoperability mode at global level. ***/
switch(config)# nuage controller interop

/*** Configure BGP to enable interoperability mode. ***/
switch(config)# router bgp 1001
switch(config-router)# address-family l2vpn evpn
switch(config-router-af)# advertise-system-mac
switch(config-router-af)# allow-vni-in-ethertag

/*** Advertise host tenant routes as evpn type-5 routes for interoperability. ***/
switch(config)# route-map permitall permit 10
switch(config)# router bgp 1001
```

```
switch(config-router)# vrf vni-491830
switch(config-router-vrf)# address-family ipv4 unicast
switch(config-router-vrf-af)# advertise l2vpn evpn
switch(config-router-vrf-af)# redistribute hmm route-map permitall
```

Verifying the VXLAN BGP EVPN Configuration

To display the VXLAN BGP EVPN configuration information, enter one of the following commands:

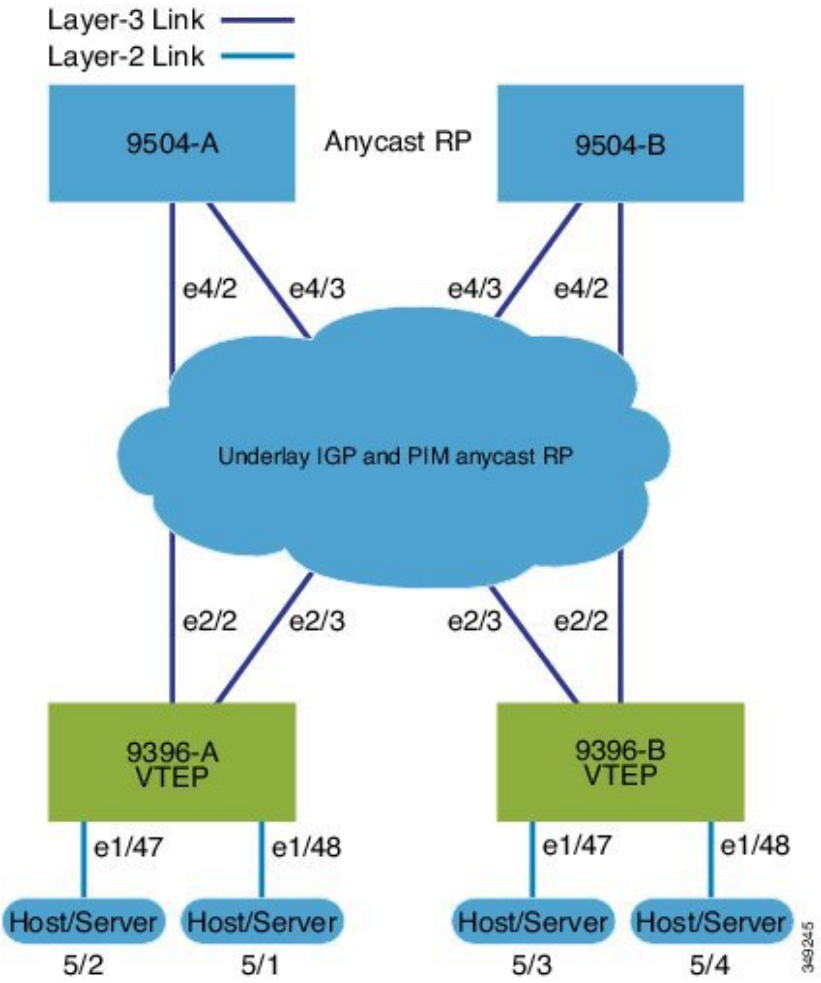
Command	Purpose
show nve vrf	Displays VRFs and associated VNIs
show bgp l2vpn evpn	Displays routing table information.
show ip arp suppression-cache [**detail** | **summary** | **vlan** *vlan* | **statistics**]	Displays ARP suppression information.
show vxlan interface	Displays VXLAN interface status.
show vxlan interface | count	Displays VXLAN VLAN logical port VP count. **Note** A VP is allocated on a per-port per-VLAN basis. The sum of all VPs across all VXLAN-enabled Layer 2 ports gives the total logical port VP count. For example, if there are 10 Layer 2 trunk interfaces, each with 10 VXLAN VLANs, then the total VXLAN VLAN logical port VP count is 10*10 = 100.
show l2route evpn mac [**all** | **evi** *evi* [**bgp** | **local** | **static** | **vxlan** | **arp**]]	Displays Layer 2 route information.
show l2route evpn fl all	Displays all fl routes.
show l2route evpn imet all	Displays all imet routes.
show l2route evpn mac-ip all **show l2route evpn mac-ip all detail**	Displays all MAC IP routes.
show l2route topology	Displays Layer 2 route topology.

Note Although the **show ip bgp** command is available for verifying a BGP configuration, as a best practice, it is preferable to use the **show bgp** command instead.

Example of VXLAN BGP EVPN (EBGP)

An example of a VXLAN BGP EVPN (EBGP):

Figure 1: VXLAN BGP EVPN Topology (EBGP)



EBGP between Spine and Leaf

- Spine (9504-A)
  - Enable the EVPN control plane

    ```
    nv overlay evpn
    ```

  - Enable the relevant protocols

    ```
    feature bgp
    feature pim
    ```

  - Configure Loopback for local VTEP IP, and BGP

```
interface loopback0
  ip address 10.1.1.1/32
  ip pim sparse-mode
```

- Configure Loopback for Anycast RP

```
interface loopback1
  ip address 100.1.1.1/32
  ip pim sparse-mode
```

- Configure Anycast RP

```
ip pim rp-address 100.1.1.1 group-list 225.0.0.0/8
ip pim rp-candidate loopback1 group-list 225.0.0.0/8
ip pim log-neighbor-changes
ip pim ssm range 232.0.0.0/8
ip pim anycast-rp 100.1.1.1 10.1.1.1
ip pim anycast-rp 100.1.1.1 20.1.1.1
```

- Configure route-map used by EBGP for Spine

```
route-map permitall permit 10
  set ip next-hop unchanged
```

- Configure interfaces for Spine-leaf interconnect

```
interface Ethernet4/2
  ip address 192.168.1.42/24
  ip pim sparse-mode
  no shutdown

interface Ethernet4/3
  ip address 192.168.2.43/24
  ip pim sparse-mode
  no shutdown
```

- Configure the BGP overlay for the EVPN address family.

```
router bgp 100
  router-id 10.1.1.1
  address-family l2vpn evpn
    nexthop route-map permitall
    retain route-target all
  neighbor 30.1.1.1 remote-as 200
    update-source loopback0
    ebgp-multihop 3
    address-family l2vpn evpn
      disable-peer-as-check
      send-community extended
      route-map permitall out
  neighbor 40.1.1.1 remote-as 200
    update-source loopback0
    ebgp-multihop 3
    address-family l2vpn evpn
      disable-peer-as-check
      send-community extended
      route-map permitall out
```

- Configure the BGP underlay.

```
neighbor 192.168.1.43 remote-as 200
    address-family ipv4 unicast
      allowas-in
      disable-peer-as-check
```

- Spine (9504-B)
  - Enable the EVPN control plane and the relevant protocols

```
feature telnet
feature nxapi
feature bash-shell
feature scp-server
nv overlay evpn
feature bgp
feature pim
feature lldp
```

  - Configure Anycast RP

```
ip pim rp-address 100.1.1.1 group-list 225.0.0.0/8
ip pim rp-candidate loopback1 group-list 225.0.0.0/8
ip pim log-neighbor-changes
ip pim ssm range 232.0.0.0/8
ip pim anycast-rp 100.1.1.1 10.1.1.1
ip pim anycast-rp 100.1.1.1 20.1.1.1
vlan 1-1002
route-map permitall permit 10
  set ip next-hop unchanged
```

  - Configure interfaces for Spine-leaf interconnect

```
interface Ethernet4/2
  ip address 192.168.4.42/24
  ip pim sparse-mode
  no shutdown

interface Ethernet4/3
  ip address 192.168.3.43/24
  ip pim sparse-mode
  no shutdown
```

  - Configure Loopback for local VTEP IP, and BGP

```
interface loopback0
  ip address 20.1.1.1/32
  ip pim sparse-mode
```

  - Configure Loopback for Anycast RP

```
interface loopback1
  ip address 100.1.1.1/32
  ip pim sparse-mode
```

  - Configure the BGP overlay for the EVPN address family.

```
router bgp 100
```

```
  router-id 20.1.1.1
  address-family l2vpn evpn
    retain route-target all
  neighbor 30.1.1.1 remote-as 200
    update-source loopback0
    ebgp-multihop 3
    address-family l2vpn evpn
      disable-peer-as-check
      send-community extended
      route-map permitall out
  neighbor 40.1.1.1 remote-as 200
    ebgp-multihop 3
    address-family l2vpn evpn
      disable-peer-as-check
      send-community extended
      route-map permitall out
```

- Configure the BGP underlay.

```
neighbor 192.168.1.43 remote-as 200
    address-family ipv4 unicast
      allowas-in
      disable-peer-as-check
```

- Leaf (9396-A)
  - Enable the EVPN control plane

```
nv overlay evpn
```

  - Enable the relevant protocols

```
feature bgp
feature pim
feature interface-vlan
feature dhcp
```

  - Configure DHCP relay for Tenant VRFs

```
service dhcp
ip dhcp relay
ip dhcp relay information option
ip dhcp relay sub-option type cisco
ip dhcp relay information option vpn
```

  - Enable VXLAN with distributed anycast-gateway using BGP EVPN

```
feature vn-segment-vlan-based
feature nv overlay
fabric forwarding anycast-gateway-mac 0000.2222.3333
```

  - Enable PIM RP

```
ip pim rp-address 100.1.1.1 group-list 225.0.0.0/8
```

  - Configure Loopback for BGP

```
interface loopback0
```

```
  ip address 30.1.1.1/32
  ip pim sparse-mode
```

- Configure Loopback for local VTEP IP

```
interface loopback1
  ip address 50.1.1.1/32
  ip pim sparse-mode
```

- Configure interfaces for Spine-leaf interconnect

```
interface Ethernet2/2
  no switchport
  load-interval counter 1 5
  ip address 192.168.1.22/24
  ip pim sparse-mode
  no shutdown

interface Ethernet2/3
  no switchport
  load-interval counter 1 5
  ip address 192.168.3.23/24
  ip pim sparse-mode
  no shutdown
```

- Create the VRF overlay VLAN and configure the vn-segment.

```
vlan 101
  vn-segment 900001
```

- Configure VRF overlay VLAN/SVI for the VRF

```
interface Vlan101
  no shutdown
  vrf member vxlan-900001
  ip forward
```

- Create VLAN and provide mapping to VXLAN

```
vlan 1001
  vn-segment 2001001
vlan 1002
  vn-segment 2001002
```

- Create VRF and configure VNI

```
vrf context vxlan-900001
  vni 900001
```

Note The **rd auto** and **route-target** commands are automatically configured unless one or more are entered as overrides.

```
  rd auto
  address-family ipv4 unicast
    route-target import 65535:101 evpn
```

```
    route-target export 65535:101 evpn
    route-target import 65535:101
    route-target export 65535:101
  address-family ipv6 unicast
    route-target import 65535:101 evpn
    route-target export 65535:101 evpn
    route-target import 65535:101
    route-target export 65535:101
```

- Create server facing SVI and enable distributed anycast-gateway

```
interface Vlan1001
  no shutdown
  vrf member vxlan-900001
  ip address 4.1.1.1/24
  ipv6 address 4:1:0:1::1/64
  fabric forwarding mode anycast-gateway
  ip dhcp relay address 192.168.100.1 use-vrf default

interface Vlan1002
  no shutdown
  vrf member vxlan-900001
  ip address 4.2.2.1/24
  ipv6 address 4:2:0:1::1/64
  fabric forwarding mode anycast-gateway
```

- Configure ACL TCAM region for ARP suppression

```
hardware access-list tcam region arp-ether 256 double-wide
```

- 

Note You can choose either of the following two options for creating the NVE interface. Use the first option for a small number of VNIs. Use the second option to configure a large number of VNIs.

Create the network virtualization endpoint (NVE) interface

Option 1

```
interface nve1
  no shutdown
  source-interface loopback1
  host-reachability protocol bgp
  member vni 900001 associate-vrf
  member vni 2001001
    suppress-arp
    mcast-group 225.4.0.1
  member vni 2001002
    suppress-arp
    mcast-group 225.4.0.1
```

Option 2

```
interface nve1
  no shutdown
  source-interface loopback 1
  host-reachibility protocol bgp
  global suppress-arp
```

```
  global mcast-group 224.1.1.1 L3
  global mcast-group 255.1.1.1 L2
  member vni 10000 associate-vrf
  member vni 10001 associate-vrf
  member vni 10002 associate-vrf
  member vni 10003 associate-vrf
  member vni 10004 associate-vrf
  member vni 10005 associate-vrf
  member vni 20000
  member vni 20001
  member vni 20002
  member vni 20003
  member vni 20004
  member vni 20005
```

- Configure interfaces for hosts/servers.

```
interface Ethernet1/47
  switchport access vlan 1002
interface Ethernet1/48
  switchport access vlan 1001
```

- Configure BGP

```
router bgp 200
router-id 30.1.1.1
  neighbor 10.1.1.1 remote-as 100
    update-source loopback0
    ebgp-multihop 3
      allowas-in
      send-community extended
    address-family l2vpn evpn
      allowas-in
      send-community extended
  neighbor 20.1.1.1 remote-as 100
    update-source loopback0
    ebgp-multihop 3
      allowas-in
      send-community extended
    address-family l2vpn evpn
      allowas-in
      send-community extended
  vrf vxlan-900001

      advertise l2vpn evpn
```

Note The following commands in EVPN mode do not need to be entered.

```
evpn
  vni 2001001 l2
  vni 2001002 l2
```

Note The **rd auto** and **route-target auto** commands are automatically configured unless one or more are entered as overrides.

```
rd auto
    route-target import auto
    route-target export auto


router bgp 200
router-id 30.1.1.1
  neighbor 10.1.1.1 remote-as 100
    update-source loopback0
    ebgp-multihop 3
      allowas-in
      send-community extended
    address-family l2vpn evpn
      allowas-in
      send-community extended
  neighbor 20.1.1.1 remote-as 100
    update-source loopback0
    ebgp-multihop 3
      allowas-in
      send-community extended
    address-family l2vpn evpn
      allowas-in
      send-community extended
  vrf vxlan-900001

      advertise l2vpn evpn
```

Note The following **advertise** command is optional.

```
advertise l2vpn evpn


evpn
  vni 2001001 l2
  vni 2001002 l2
```

Note The following **rd auto** and **route-target** commands are automatically configured unless one or more are entered as overrides.

Note The following EVPN mode commands are optional.

```
             evpn
  vni 2001001 l2
```

```
    rd auto
    route-target import auto
    route-target export auto
  vni 2001002 l2
    rd auto
    route-target import auto
    route-target export auto
```

- Leaf (9396-B)
  - Enable the EVPN control plane functionality and the relevant protocols

```
feature telnet
feature nxapi
feature bash-shell
feature scp-server
nv overlay evpn
feature bgp
feature pim
feature interface-vlan
feature vn-segment-vlan-based
feature lldp
feature nv overlay
```

  - Enable VxLAN with distributed anycast-gateway using BGP EVPN

```
fabric forwarding anycast-gateway-mac 0000.2222.3333
```

  - Create the VRF overlay VLAN and configure the vn-segment

```
vlan 1-1002
vlan 101
  vn-segment 900001
```

  - Create VLAN and provide mapping to VXLAN

```
vlan 1001
  vn-segment 2001001
vlan 1002
  vn-segment 2001002
```

  - Create VRF and configure VNI

```
vrf context vxlan-900001
  vni 900001
```

Note The following commands are automatically configured unless one or more are entered as overrides.

```
rd auto
  address-family ipv4 unicast
    route-target import 65535:101 evpn
    route-target export 65535:101 evpn
    route-target import 65535:101
    route-target export 65535:101
```

```
  address-family ipv6 unicast
    route-target import 65535:101 evpn
    route-target export 65535:101 evpn
    route-target import 65535:101 evpn
    route-target export 65535:101 evpn
```

- Configure ACL TCAM region for ARP suppression

```
hardware access-list tcam region arp-ether 256 double-wide
```

- Configure internal control VLAN/SVI for the VRF

```
interface Vlan1

interface Vlan101
  no shutdown
  vrf member vxlan-900001
```

- Create server facing SVI and enable distributed anycast-gateway

```
interface Vlan1001
  no shutdown
  vrf member vxlan-900001
  ip address 4.1.1.1/24
  ipv6 address 4:1:0:1::1/64
  fabric forwarding mode anycast-gateway

interface Vlan1002
  no shutdown
  vrf member vxlan-900001
  ip address 4.2.2.1/24
  ipv6 address 4:2:0:1::1/64
  fabric forwarding mode anycast-gateway
```

- Create the network virtualization endpoint (NVE) interface

Note You can choose either of the following two procedures for creating the NVE interface. Use Option 1 for a small number of VNIs. Use Option 2 to configure a large number of VNIs.

Option 1

```
interface nve1
  no shutdown
  source-interface loopback1
  host-reachability protocol bgp
  member vni 10000 associate-vrf
  mcast-group 224.1.1.1
  member vni 10001 associate-vrf
  mcast-group 224.1.1.1
  member vni20000
  suppress-arp
  mcast-group 225.1.1.1
  member vni 20001
  suppress-arp
  mcast-group 225.1.1.1
```

Option 2

```
interface nve1
  no shutdown
  source-interface loopback 1
  host-reachibility protocol bgp
  global suppress-arp
  global mcast-group 224.1.1.1 L3
  global mcast-group 255.1.1.1 L2
  member vni 10000 associate-vrf
  member vni 10001 associate-vrf
  member vni 10002 associate-vrf
  member vni 10003 associate-vrf
  member vni 10004 associate-vrf
  member vni 10005 associate-vrf
  member vni 20000
  member vni 20001
  member vni 20002
  member vni 20003
  member vni 20004
  member vni 20005
```

- Configure interfaces for hosts/servers

```
interface Ethernet1/47
  switchport access vlan 1002

interface Ethernet1/48
  switchport access vlan 1001
```

- Configure interfaces for Spine-leaf interconnect

```
interface Ethernet2/1

interface Ethernet2/2
  no switchport
  load-interval counter 1 5
  ip address 192.168.4.22/24
  ip pim sparse-mode
  no shutdown

interface Ethernet2/3
  no switchport
  load-interval counter 1 5
  ip address 192.168.2.23/24
  ip pim sparse-mode
  no shutdown
```

- Configure Loopback for BGP

```
interface loopback0
  ip address 40.1.1.1/32
  ip pim sparse-mode
```

- Configure Loopback for local VTEP IP

```
interface loopback1
  ip address 51.1.1.1/32
  ip pim sparse-mode
```

• Configure BGP

```
router bgp 200
router-id 40.1.1.1
  neighbor 10.1.1.1 remote-as 100
    update-source loopback0
    ebgp-multihop 3
      allowas-in
      send-community extended
    address-family l2vpn
      allowas-in
      send-community extended
  neighbor 20.1.1.1 remote-as 100
    update-source loopback0
    ebgp-multihop 3
      allowas-in
      send-community extended
    address-family l2vpn
      allowas-in
      send-community extended
  vrf vxlan-900001
```

Note The following **advertise** command is optional.

```
advertise l2vpn evpn
```

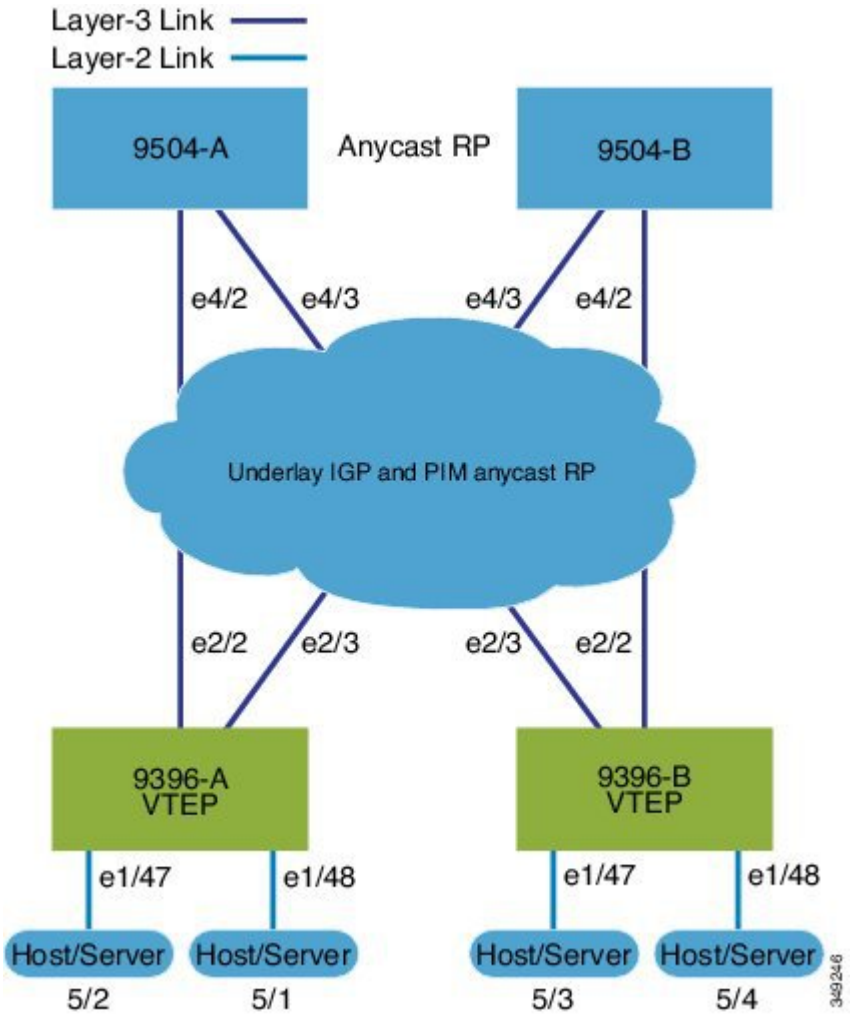
Note The **rd auto** and **route-target** commands are optional unless you want to use them to override the **import** or **export** options.

```
            evpn
  vni 2001001 l2
    rd auto
    route-target import auto
    route-target export auto
  vni 2001002 l2
    rd auto
    route-target import auto
    route-target export auto
```

Example of VXLAN BGP EVPN (IBGP)

An example of a VXLAN BGP EVPN (IBGP):

Figure 2: VXLAN BGP EVPN Topology (IBGP)



IBGP between Spine and Leaf

- Spine (9504-A)
  - Enable the EVPN control plane

    ```
    nv overlay evpn
    ```

  - Enable the relevant protocols

    ```
    feature ospf
    feature bgp
    feature pim
    ```

  - Configure Loopback for local VTEP IP, and BGP

    ```
    interface loopback0
      ip address 10.1.1.1/32
      ip router ospf 1 area 0.0.0.0
      ip pim sparse-mode
    ```

- Configure Loopback for Anycast RP

```
interface loopback1
  ip address 100.1.1.1/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
```

- Configure Anycast RP

```
ip pim rp-address 100.1.1.1 group-list 225.0.0.0/8
ip pim rp-candidate loopback1 group-list 225.0.0.0/8
ip pim ssm range 232.0.0.0/8
ip pim anycast-rp 100.1.1.1 10.1.1.1
ip pim anycast-rp 100.1.1.1 20.1.1.1
```

- Enable OSPF for underlay routing

```
router ospf 1
```

- Configure interfaces for Spine-leaf interconnect

```
interface Ethernet4/2
  ip address 192.168.1.42/24
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  no shutdown

interface Ethernet4/3
  ip address 192.168.2.43/24
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  no shutdown
```

- Configure BGP

```
router bgp 65535
router-id 10.1.1.1
  neighbor 30.1.1.1 remote-as 65535
    update-source loopback0
    address-family l2vpn evpn
      send-community both
      route-reflector-client
  neighbor 40.1.1.1 remote-as 65535
    update-source loopback0
    address-family l2vpn evpn
      send-community both
      route-reflector-client
```

- Spine (9504-B)
  - Enable the EVPN control plane and the relevant protocols

```
feature telnet
feature nxapi
feature bash-shell
feature scp-server
nv overlay evpn
feature ospf
feature bgp
```

```
feature pim
feature lldp
```

- Configure Anycast RP

```
ip pim rp-address 100.1.1.1 group-list 225.0.0.0/8
ip pim rp-candidate loopback1 group-list 225.0.0.0/8
ip pim ssm range 232.0.0.0/8
ip pim anycast-rp 100.1.1.1 10.1.1.1
ip pim anycast-rp 100.1.1.1 20.1.1.1
vlan 1-1002
```

- Configure interfaces for Spine-leaf interconnect

```
interface Ethernet4/2
  ip address 192.168.4.42/24
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  no shutdown

interface Ethernet4/3
  ip address 192.168.3.43/24
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  no shutdown
```

- Configure Loopback for local VTEP IP, and BGP

```
interface loopback0
  ip address 20.1.1.1/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
```

- Configure Loopback for Anycast RP

```
interface loopback1
  ip address 100.1.1.1/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
```

- Enable OSPF for underlay routing

```
router ospf 1
```

- Configure BGP

```
router bgp 65535
router-id 20.1.1.1
  neighbor 30.1.1.1 remote-as 65535
    update-source loopback0
    address-family l2vpn evpn
      send-community both
      route-reflector-client
  neighbor 40.1.1.1 remote-as 65535
    update-source loopback0
     address-family l2vpn evpn
      send-community both
      route-reflector-client
```

- Leaf (9396-A)
  - Enable the EVPN control plane

```
nv overlay evpn
```

  - Enable the relevant protocols

```
feature ospf
feature bgp
feature pim
feature interface-vlan
```

  - Enable VxLAN with distributed anycast-gateway using BGP EVPN

```
feature vn-segment-vlan-based
feature nv overlay
fabric forwarding anycast-gateway-mac 0000.2222.3333
```

  - Enabling OSPF for underlay routing

```
router ospf 1
```

  - Configure Loopback for local VTEP IP, and BGP

```
interface loopback0
  ip address 30.1.1.1/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
```

  - Configure interfaces for Spine-leaf interconnect

```
interface Ethernet2/2
  no switchport
  ip address 192.168.1.22/24
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  no shutdown

interface Ethernet2/3
  no switchport
  ip address 192.168.3.23/24
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  no shutdown
```

  - Configure PIM RP

```
ip pim rp-address 100.1.1.1 group-list 225.0.0.0/8
ip pim ssm range 232.0.0.0/8
```

  - Create overlay VRF VLAN and configure vn-segment

```
vlan 101
  vn-segment 900001
```

  - Configure VRF overlay VLAN/SVI for the VRF

```
interface Vlan101
  no shutdown
  vrf member vxlan-900001
```

- Create VLAN and provide mapping to VXLAN

```
vlan 1001
  vn-segment 2001001
vlan 1002
  vn-segment 2001002
```

- Create VRF and configure VNI

```
vrf context vxlan-900001
  vni 900001
```

Note The **rd auto** and **route-target** commands are automatically configured unless one or more are entered as overrides.

```
rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn
  address-family ipv6 unicast
    route-target both auto
    route-target both auto evpn
```

- Create server facing SVI and enable distributed anycast-gateway

```
interface Vlan1001
  no shutdown
  vrf member vxlan-900001
  ip address 4.1.1.1/24
  ipv6 address 4:1:0:1::1/64
  fabric forwarding mode anycast-gateway

interface Vlan1002
  no shutdown
  vrf member vxlan-900001
  ip address 4.2.2.1/24
  ipv6 address 4:2:0:1::1/64
  fabric forwarding mode anycast-gateway
```

- Configure ACL TCAM region for ARP suppression

```
hardware access-list tcam region arp-ether 256 double-wide
```

- 

Note You can choose either of the following two procedures for creating the NVE interfaces. Use the first one for a small number of VNIs. Use the second procedure to configure a large number of VNIs.

Create the network virtualization endpoint (NVE) interface

Option 1

```
interface nve1
  no shutdown
  source-interface loopback0
  host-reachability protocol bgp
  member vni 900001 associate-vrf
  member vni 2001001
    suppress-arp
    mcast-group 225.4.0.1
  member vni 2001002
    suppress-arp
    mcast-group 225.4.0.1
```

Option 2

```
Interface nve1
  source-interface loopback 1
  host-reachability protocol bgp
  global suppress-arp
  global mcast-group 255.1.1.1 L2
  global mcast-group 255.1.1.2 L3
  member vni 10000
  member vni 20000
  member vni 30000
```

- Configure interfaces for hosts/servers

```
interface Ethernet1/47
  switchport access vlan 1002

interface Ethernet1/48
  switchport access vlan 1001
```

- Configure BGP

```
router bgp 65535
router-id 30.1.1.1
  neighbor 10.1.1.1 remote-as 65535
    update-source loopback0
    address-family l2vpn evpn
      send-community both
  neighbor 20.1.1.1 remote-as 65535
    update-source loopback0
    address-family l2vpn evpn
      send-community both
  vrf vxlan-900001
    address-family ipv4 unicast
      advertise l2vpn evpn
```

Note The following commands in EVPN mode do not need to be entered.

```
evpn
  vni 2001001 l2
  vni 2001002 l2
```

Note The **rd auto** and **route-target auto** commands are automatically configured unless one or more are entered as overrides.

```
rd auto
    route-target import auto
    route-target export auto
```

Note The **rd auto** and **route-target** commands are automatically configured unless you want to use them to override the **import** or **export** options.

Note The following EVPN mode commands are optional.

```
            evpn
  vni 2001001 l2
    rd auto
    route-target import auto
    route-target export auto
  vni 2001002 l2
    rd auto
    route-target import auto
    route-target export auto
```

- Leaf (9396-B)
  - Enable the EVPN control plane functionality and the relevant protocols

```
feature telnet
feature nxapi
feature bash-shell
feature scp-server
nv overlay evpn
feature ospf
feature bgp
feature pim
feature interface-vlan
feature vn-segment-vlan-based
feature lldp
feature nv overlay
```

  - Enable VxLAN with distributed anycast-gateway using BGP EVPN

```
fabric forwarding anycast-gateway-mac 0000.2222.3333
```

  - Configure PIM RP

```
ip pim rp-address 100.1.1.1 group-list 225.0.0.0/8
ip pim ssm range 232.0.0.0/8
```

- Create overlay VRF VLAN and configure vn-segment

```
vlan 1-1002
vlan 101
  vn-segment 900001
```

- Create VLAN and provide mapping to VXLAN

```
vlan 1001
  vn-segment 2001001
vlan 1002
  vn-segment 2001002
```

- Create VRF and configure VNI

```
vrf context vxlan-900001
  vni 900001
```

Note The **rd auto** and **route-target** commands are automatically configured unless you want to use them to override the **import** or **export** options.

```
rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn
  address-family ipv6 unicast
    route-target both auto
    route-target both auto evpn
```

- Configure ACL TCAM region for ARP suppression

```
hardware access-list tcam region arp-ether 256 double-wide
```

- Configure internal control VLAN/SVI for the VRF

```
interface Vlan101
  no shutdown
  vrf member vxlan-900001
```

- Create server facing SVI and enable distributed anycast-gateway

```
interface Vlan1001
  no shutdown
  vrf member vxlan-900001
  ip address 4.1.1.1/24
  ipv6 address 4:1:0:1::1/64
  fabric forwarding mode anycast-gateway

interface Vlan1002
  no shutdown
  vrf member vxlan-900001
```

```
  ip address 4.2.2.1/24
  ipv6 address 4:2:0:1::1/64
  fabric forwarding mode anycast-gateway
```

- 

Note You can choose either of the following two command procedures for creating the NVE interfaces. Use Option 1 for a small number of VNIs. Use Option 2 to configure a large number of VNIs.

Create the network virtualization endpoint (NVE) interface

Option 1

```
interface nve1
  no shutdown
  source-interface loopback0
  host-reachability protocol bgp
  member vni 900001 associate-vrf
  member vni 2001001
    suppress-arp
    mcast-group 225.4.0.1
  member vni 2001002
    suppress-arp
    mcast-group 225.4.0.1
```

Option 2

```
Interface nve1
  source-interface loopback0
  host-reachability protocol bgp
  global suppress-arp
  global mcast-group 255.4.0.1
  member vni 900001
  member vni 2001001
```

- Configure interfaces for hosts/servers

```
interface Ethernet1/47
  switchport access vlan 1002

interface Ethernet1/48
  switchport access vlan 1001
```

- Configure interfaces for Spine-leaf interconnect

```
interface Ethernet2/1

interface Ethernet2/2
  no switchport
  ip address 192.168.4.22/24
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  no shutdown

interface Ethernet2/3
  no switchport
  ip address 192.168.2.23/24
  ip router ospf 1 area 0.0.0.0
```

```
  ip pim sparse-mode
  no shutdown
```

- Configure Loopback for local VTEP IP, and BGP

```
interface loopback0
  ip address 40.1.1.1/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
```

- Enabling OSPF for underlay routing

```
router ospf 1
```

- Configure BGP

```
router bgp 65535
router-id 40.1.1.1
  neighbor 10.1.1.1 remote-as 65535
    update-source loopback0
    address-family l2vpn evpn
      send-community both
  neighbor 20.1.1.1 remote-as 65535
    update-source loopback0
    address-family l2vpn evpn
      send-community both
  vrf vxlan-900001
    address-family ipv4 unicast
      advertise l2vpn evpn
evpn
  vni 2001001 l2
    rd auto
    route-target import auto
    route-target export auto
  vni 2001002 l2
    rd auto
    route-target import auto
    route-target export auto
```

Note The **rd auto** and **route-target** commands are optional unless you want to use them to override the **import** or **export** options.

```
evpn
  vni 2001001 l2
    rd auto
    route-target import auto
    route-target export auto
  vni 2001002 l2
    rd auto
    route-target import auto
    route-target export auto
```

Example Show Commands

- **show nve peers**

```
9396-B# show nve peers
Interface Peer-IP          Peer-State
--------- ---------------  ----------
nve1      30.1.1.1         Up
```

- **show nve vni**

```
9396-B# show nve vni
Codes: CP - Control Plane        DP - Data Plane
       UC - Unconfigured         SA - Suppress ARP

Interface VNI      Multicast-group   State Mode Type [BD/VRF]      Flags
--------- -------- ----------------- ----- ---- ------------------ -----
nve1      900001   n/a               Up    CP   L3 [vxlan-900001]
nve1      2001001  225.4.0.1         Up    CP   L2 [1001]          SA
nve1      2001002  225.4.0.1         Up    CP   L2 [1002]          SA
```

- **show ip arp suppression-cache detail**

```
9396-B# show ip arp suppression-cache detail

Flags: + - Adjacencies synced via CFSoE
       L - Local Adjacency
       R - Remote Adjacency
       L2 - Learnt over L2 interface

Ip Address      Age      Mac Address    Vlan Physical-ifindex    Flags

4.1.1.54        00:06:41 0054.0000.0000 1001 Ethernet1/48        L
4.1.1.51        00:20:33 0051.0000.0000 1001 (null)              R
4.2.2.53        00:06:41 0053.0000.0000 1002 Ethernet1/47        L
4.2.2.52        00:20:33 0052.0000.0000 1002 (null)              R
```

- **show vxlan interface**

```
9396-B# show vxlan interface
Interface       Vlan    VPL Ifindex     LTL             HW VP
=========       ====    ===========     ===             =====
Eth1/47         1002    0x4c07d22e      0x10000         5697
Eth1/48         1001    0x4c07d02f      0x10001         5698
```

- **show bgp l2vpn evpn summary**

```
9396-B# show bgp l2vpn evpn summary
BGP summary information for VRF default, address family L2VPN EVPN
BGP router identifier 40.1.1.1, local AS number 65535
BGP table version is 27, L2VPN EVPN config peers 2, capable peers 2
14 network entries and 18 paths using 2984 bytes of memory
BGP attribute entries [14/2240], BGP AS path entries [0/0]
```

```
BGP community entries [0/0], BGP clusterlist entries [2/8]

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.1.1.1        4 65535   30199   30194       27    0    0     2w6d 4
20.1.1.1        4 65535   30199   30194       27    0    0     2w6d 4
```

- **show bgp l2vpn evpn**

```
9396-B# show bgp l2vpn evpn
BGP routing table information for VRF default, address family L2VPN EVPN
BGP table version is 27, Local Router ID is 40.1.1.1
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: 30.1.1.1:33768
*>i[2]:[0]:[0]:[48]:[d8b1.9071.e903]:[0]:[0.0.0.0]/216
                      30.1.1.1                           100          0 i
* i                   30.1.1.1                           100          0 i
*>i[2]:[0]:[0]:[48]:[d8b1.9071.e903]:[32]:[4.1.1.12]/272
                      30.1.1.1                           100          0 i
* i                   30.1.1.1                           100          0 i

Route Distinguisher: 30.1.1.1:33769
*>i[2]:[0]:[0]:[48]:[d8b1.9071.e903]:[0]:[0.0.0.0]/216
                      30.1.1.1                           100          0 i
* i                   30.1.1.1                           100          0 i
*>i[2]:[0]:[0]:[48]:[d8b1.9071.e903]:[32]:[4.2.2.11]/272
                      30.1.1.1                           100          0 i
* i                   30.1.1.1                           100          0 i

Route Distinguisher: 40.1.1.1:33768    (L2VNI 2001001)
*>i[2]:[0]:[0]:[48]:[d8b1.9071.e903]:[0]:[0.0.0.0]/216
                      30.1.1.1                           100          0 i
*>l[2]:[0]:[0]:[48]:[f8c2.8890.2a45]:[0]:[0.0.0.0]/216
                      40.1.1.1                           100      32768 i
*>i[2]:[0]:[0]:[48]:[d8b1.9071.e903]:[32]:[4.1.1.12]/272
                      30.1.1.1                           100          0 i
*>l[2]:[0]:[0]:[48]:[f8c2.8890.2a45]:[32]:[4.1.1.122]/272
                      40.1.1.1                           100      32768 i

Route Distinguisher: 40.1.1.1:33769    (L2VNI 2001002)
*>i[2]:[0]:[0]:[48]:[d8b1.9071.e903]:[0]:[0.0.0.0]/216
                      30.1.1.1                           100          0 i
*>l[2]:[0]:[0]:[48]:[f8c2.8890.2a45]:[0]:[0.0.0.0]/216
                      40.1.1.1                           100      32768 i
*>i[2]:[0]:[0]:[48]:[d8b1.9071.e903]:[32]:[4.2.2.11]/272
                      30.1.1.1                           100          0 i
*>l[2]:[0]:[0]:[48]:[f8c2.8890.2a45]:[32]:[4.2.2.111]/272
                      40.1.1.1                           100      32768 i

Route Distinguisher: 40.1.1.1:3    (L3VNI 900001)
*>i[2]:[0]:[0]:[48]:[d8b1.9071.e903]:[32]:[4.1.1.12]/272
                      30.1.1.1                           100          0 i
*>i[2]:[0]:[0]:[48]:[d8b1.9071.e903]:[32]:[4.2.2.11]/272
                      30.1.1.1                           100          0 i
```

- **show l2route evpn mac all**

```
9396-B# show l2route evpn mac all
```

```
Flags -(Rmac):Router MAC (Stt):Static (L):Local (R):Remote (V):vPC link
(Dup):Duplicate (Spl):Split (Rcv):Recv (AD):Auto-Delete (D):Del Pending
(S):Stale (C):Clear, (Ps):Peer Sync (O):Re-Originated (Nho):NH-Override
(Pf):Permanently-Frozen

Topology    Mac Address    Prod   Flags          Seq No     Next-Hops
----------- -------------- ------ -------------- ---------- ----------------
101         6412.2574.9f27 VXLAN  Rmac           0          30.1.1.1
1001        d8b1.9071.e903 BGP    SplRcv         0          30.1.1.1
1001        f8c2.8890.2a45 Local  L,             0          Eth1/48
1002        d8b1.9071.e903 BGP    SplRcv         0          30.1.1.1
1002        f8c2.8890.2a45 Local  L,             0          Eth1/47
```

- **show l2route evpn mac-ip all**

```
9396-B# show l2route evpn mac-ip all
Flags -(Rmac):Router MAC (Stt):Static (L):Local (R):Remote (V):vPC link
(Dup):Duplicate (Spl):Split (Rcv):Recv(D):Del Pending (S):Stale (C):Clear
(Ps):Peer Sync (Ro):Re-Originated
Topology    Mac Address    Prod   Flags          Seq No     Host IP         Next-Hops
----------- -------------- ------ ---------- --------------- ---------------
1001        d8b1.9071.e903 BGP    --             0          4.1.1.12       30.1.1.1
1001        f8c2.8890.2a45 HMM    --             0          4.1.1.122      Local
1002        d8b1.9071.e903 BGP    --             0          4.2.2.11       30.1.1.1
1002        f8c2.8890.2a45 HMM    --             0          4.2.2.111      Local
```