



Configure VXLAN BGP EVPN

This chapter contains these sections:

- [VXLAN BGP EVPN, on page 1](#)
- [VXLAN EVPN with Downstream VNI, on page 45](#)
- [EVPN Centralized Gateway, on page 52](#)

VXLAN BGP EVPN

This chapter contains these sections:

VXLAN BGP EVPN

VXLAN BGP EVPN is a data center network overlay protocol suite that

- enables scalable Layer 2 and Layer 3 connectivity between distributed network endpoints
- uses BGP EVPN as the control plane to advertise MAC/IP address bindings, and
- supports multi-tenant network virtualization with enhanced operational flexibility.

VXLAN encapsulates Layer 2 frames in Layer 3 UDP packets, enabling scalable network overlays. BGP EVPN provides a standards-based control plane that supports dynamic endpoint discovery and efficient traffic forwarding.

VXLAN BGP EVPN can be used to interconnect multiple data center sites, providing secure and isolated tenant networks across the infrastructure.

Auto-derived route distinguishers

An auto-derived route distinguisher (rd auto) is a VPN address-mapping mechanism that

- uses a Type 1 encoding format combining a 4-byte BGP Router ID and a 2-byte numbering field
- distinguishes between IP-VRF and MAC-VRF through different numbering schemes, and
- enables unique identification across multiple VRFs.

In Cisco NX-OS, the auto-derived RD uses the IP address of the BGP Router ID (RID) for the 4-byte administrative field and the internal VRF identifier for the 2-byte numbering field (VRF ID). This format is specified in [IETF RFC 4364 section 4.2](#).

The 2-byte numbering field is always derived from the VRF, but results in a different numbering scheme depending on its use for the IP-VRF or the MAC-VRF:

- **IP-VRF:** The 2-byte numbering field for the IP-VRF uses the internal VRF ID, which starts at 1 and increases incrementally. VRF IDs 1 and 2 are reserved for the default VRF and the management VRF, respectively. The first custom-defined IP VRF uses VRF ID 3.

IP-VRF with BGP Router ID 192.0.2.1 and VRF ID 6: RD 192.0.2.1:6

- **MAC-VRF:** The 2-byte numbering field for the MAC-VRF uses the VLAN ID + 32767, which results in 32768 for VLAN ID 1 and incrementing.

MAC-VRF with BGP Router ID 192.0.2.1 and VLAN 20: RD 192.0.2.1:32787

Route-target autos

A route-target (RT) auto is a route-target assignment method that

- derives route-target values automatically based on system parameters
- uses the Type 0 extended community encoding as described in [IETF RFC 4364](#), and
- constructs the route-target using the Autonomous System Number (ASN) and the Service Identifier (VNI).

The auto-derived route-target (using import/export/both auto) is based on the Type 0 encoding format as described in [IETF RFC 4364 section 4.2](#). This encoding allows a 2-byte administrative field and a 4-byte numbering field.

Within Cisco NX-OS, the auto-derived route-target uses the ASN for the 2-byte administrative field. It uses the VNI for the 4-byte numbering field.

In multi-AS environments, route-targets must match the correct ASN portion. You may need to define or rewrite them to ensure compatibility. For more information, see [rewrite-evpn-rt-asn](#).

Examples of an auto-derived Route-Target (RT)

- For 2-byte ASN:
 - IP-VRF within ASN 65001 and L3VNI 50001 - Route-Target 65001:50001
 - MAC-VRF within ASN 65001 and L2VNI 30001 - Route-Target 65001:30001
- For 4-byte ASN:
 - IP-VRF within ASN 65656 and L3VNI 50001 - Route-Target 23456:50001
 - MAC-VRF within ASN 65656 and L2VNI 30001 - Route-Target 23456:30001

When a 4-byte ASN is used, the 2-byte ASN field is set to 23456 (AS_TRANS) as specified in [IETF RFC 6793 section 9](#); this value is registered by IANA as a special-purpose AS number to represent 4-byte ASNs in 2-byte fields.



Note Beginning with Cisco NX-OS Release 9.2(1), auto-derived Route-Target for 4-byte ASN is supported.

Supported features and configuration limits for VXLAN BGP EVPN

VXLAN BGP EVPN has these supported features, platforms, and configuration limits for VXLAN BGP EVPN:

Configuration recommendations

- Switch and port limitations:
 - The VXLAN network identifier (VNID) 16777215 is reserved and should explicitly not be configured.
 - It is recommended to use the **vpc orphan-ports suspend** command for single attached and/or routed devices on a Cisco Nexus 9000 platform switch acting as vPC VTEP.
- Feature limitations:
 - Mobility Sequence number of a locally originated type-2 route (MAC/MAC-IP) can be mismatched between vPC peers, with one VTEP having a sequence number K while other VTEP in the same complex can have the same route with sequence number 0. This does not cause any functional impact and the traffic is not impacted even after the host moves.
 - For SVI-related triggers (such as shut/unshut or PIM enable/disable), a 30-second delay was added, allowing the Multicast FIB (MFIB) Distribution module (MFDM) to clear the hardware table before toggling between L2 and L3 modes or vice versa.
 - You can configure EVPN over segment routing or MPLS. See the [Cisco Nexus 9000 Series NX-OS Label Switching Configuration Guide, Release 9.3\(x\)](#) 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 9000 Series NX-OS Label Switching Configuration Guide, Release 9.3\(x\)](#) for more information.
 - Routing protocol adjacencies using Anycast Gateway SVIs is not supported.
 - When running VXLAN EVPN, any SVI for a VLAN extended over VXLAN must be configured with Anycast Gateway. Any other mode of operation is not supported.
- Command limitations:
 - In a VXLAN EVPN setup, border nodes must be configured with unique route distinguishers, preferably using the **auto rd** command. Not using unique route distinguishers across all border nodes is not supported. The use of unique route distinguishers is strongly recommended for all VTEPs of a fabric.

- Non-Disruptive In Service Software Upgrade (ND-ISSU) is supported on Nexus 9300 with VXLAN enabled. For more information, see [Cisco Nexus 9000 and 3000 Upgrade and ISSU Matrix](#).

Supported platform and releases for VXLAN BGP EVPN

Table 1: VXLAN BGP EVPN support on Cloudscale switches

Release	Platforms
9.3(3)	Cisco Nexus 9300- EX/ FX/GX/FX2/FX3 platform switches. Cisco Nexus 9300-GX switch Cisco Nexus 9504 and 9508 with R-series line cards

Unsupported features

- Switch limitations:
 - VXLAN is not supported on N9K-C92348GC-X switches.
 - On Cisco Nexus 9000 PX/TX/PQ switches configured as VXLAN VTEPs, if any ALE 40G port is used as a VXLAN underlay port, configuring subinterfaces on either this or any other 40G port is not allowed and could lead to VXLAN traffic loss.
- Feature limitations:
 - DHCP snooping (Dynamic Host Configuration Protocol snooping) is not supported on VXLAN VLANs.
 - ACLs are not supported on VXLAN uplink interfaces. VACLs are not supported on VXLAN de-capsulated traffic in egress direction; this applies for the inner traffic coming from network (VXLAN) towards the access (Ethernet).

As a best practice, always use PACLS/VACLs for the access (Ethernet) to the network (VXLAN) direction. See the [Cisco Nexus 9000 Series NX-OS Security Configuration Guide, Release 9.3\(x\)](#) for other guidelines and limitations for the VXLAN ACL feature.
 - The Cisco Nexus 9000 QoS buffer-boost feature is not applicable for VXLAN traffic.

Scale



Note For information about VXLAN BGP EVPN scalability, see the [Cisco Nexus 9000 Series NX-OS Verified Scalability Guide](#).

- In a VXLAN EVPN setup that has 2K VNI scale configuration, the control plane down time may take more than 200 seconds. To avoid potential BGP flap, extend the graceful restart time to 300 seconds.

ARP suppression

- 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.
- The ARP suppression setting must match across the entire fabric. For a specific VNID, all VTEPs must be either configured or not configured.

VXLAN BGP EVPN fabrics with eBGP

- For VXLAN BGP EVPN fabrics with eBGP, the following recommendations are applicable:
 - It is recommended to use loopbacks for the eBGP EVPN peering sessions (overlay control-plane).
 - It is a best practice to use the physical interfaces for eBGP IPv4/IPv6 peering sessions (underlay).
- Only eBGP peering between a VTEP and external nodes (Edge Router, Core Router or VNF) is supported.
 - eBGP peering from the VTEP to the external node using a physical interface or subinterfaces is recommended and it is a best practice (external connectivity).
 - The eBGP peering from the VTEP to the external node can be in the default VRF or in a tenant VRF (external connectivity).
 - The eBGP peering from the VTEP to a external node over VXLAN must be in a tenant VRF and must use the update-source of a loopback interface (peering over VXLAN).
 - Using an SVI for eBGP peering from the VTEP to the External Node requires the VLAN to be local (not VXLAN extended).

NVE interface

- Bind the NVE source-interface to a dedicated loopback interface and do not share this loopback with any function or peerings of Layer-3 protocols. A best practice is to use a dedicated loopback address for the VXLAN VTEP function.
- You must bind NVE to a loopback address that is separate from other loopback addresses that are required by Layer 3 protocols. NVE and other Layer 3 protocols using the same loopback is not supported.
- The NVE source-interface loopback is required to be present in the default VRF.
- During the vPC Border Gateway boot up process the NVE source loopback interface undergoes the hold down timer twice instead of just once. This is a day-1 and expected behavior.
- When SVI is enabled on a VTEP (flood and learn, or EVPN) regardless of ARP suppression, make sure that ARP-ETHER TCAM is carved using the **hardware access-list tcam region arp-ether 256 double-wide** command. This requirement does not apply to Cisco Nexus 9200, 9300-EX, and 9300-FX/FX2 /FX3 and 9300-GX platform switches and Cisco Nexus 9500 platform switches with 9700- EX/ FX line cards.

Supported Templates

- When configuring VXLAN BGP EVPN, only the "System Routing Mode: Default" is applicable for the following hardware platforms:
 - Cisco Nexus 9300 platform switches
 - Cisco Nexus 9300-EX platform switches
 - Cisco Nexus 9300-FX/FX2 /FX3 platform switches
 - Cisco Nexus 9300-GX platform switches
 - Cisco Nexus 9500 platform switches with X9700-EX, and X9700-FX line cards
- Changing the "System Routing Mode" requires a reload of the switch.

VXLAN uplinks

- Starting from Cisco NX-OS Release 9.3(5), new VXLAN uplink capabilities are introduced:
 - A physical interface in default VRF is supported as VXLAN uplink.
 - A parent interface in default VRF, carrying subinterfaces with VRF and dot1q tags, is supported as VXLAN uplink.
 - A subinterface in any VRF and/or with dot1q tag remains not supported as VXLAN uplink.
 - An SVI in any VRF remains not supported as VXLAN uplink.
 - In vPC with physical peer-link, a SVI can be leveraged as backup underlay, default VRF only between the vPC members (infra-VLAN, system nve infra-vlans).
 - 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.
 - FEX host interfaces remain not supported as VXLAN uplink and cannot have VTEPs connected (BUD node).
- You need to configure the VXLAN uplink with **ip unreachable** in order to enable Path maximum transmission unit (MTU) discovery (PMTUD) in a VXLAN set up. PMTUD prevents fragmentation in the path between two endpoints by dynamically determining the lowest MTU along the path from the packet's source to its destination.
- Cisco Nexus 9500 platform switches with 9700 -EX or -FX line cards support 1G, 10G, 25G, 40G, 100G and 400G for VXLAN uplinks.
- Cisco Nexus 9200 and 9300- EX/ FX/FX2/FX3 and -GX support 1G, 10G, 25G, 40G, 100G and 400G for VXLAN uplinks.

UDP port

- The Cisco Nexus 9000 platform switches use standards conforming UDP port number 4789 for VXLAN encapsulation. This value is not configurable.

- The Cisco Nexus 9200 platform switches with Application Spine Engine (ASE2) have throughput constraints for packet sizes of 99-122 bytes; packet drops might be experienced.

SPAN

The following guidelines and limitations apply to VXLAN/VTEP using BGP EVPN:

- SPAN source or destination is supported on any port.

Gateway functionality

- Gateway functionality for VXLAN to MPLS (LDP), VXLAN to MPLS-SR (Segment Routing) and VXLAN to SRv6 can be operated on the same Cisco Nexus 9000 Series platform.
 - VXLAN to MPLS (LDP) Gateway is supported on the Cisco Nexus 3600-R and the Cisco Nexus 9500 with R-Series line cards.
 - VXLAN to MPLS-SR Gateway is supported on the Cisco Nexus 9300-FX2/FX3/GX and Cisco Nexus 9500 with R-Series line cards.
- VXLAN to SRv6 is supported on the Cisco Nexus 9300-GX platform.

VXLAN and GRE co-existence

- Multiple Tunnel Encapsulations (VXLAN, GRE and/or MPLS, static label or segment routing) can not co-exist on the same Cisco Nexus 9000 Series switch with Network Forwarding Engine (NFE).

ECMP resilient hashing

- Resilient hashing is supported on the following switch platform with a VXLAN VTEP configured:
 - Cisco Nexus 9300- EX/ FX/FX2/FX3/GX support ECMP resilient hashing.
 - Cisco Nexus 9300 with ALE uplink ports does not support resilient hashing.



Note Resilient hashing is disabled by default.

Configure VXLAN BGP EVPN

Refer to these sections for configuring the VXLAN BGP EVPN features.

Enable VXLAN

Use this task when you need to configure VXLAN and EVPN fabric functionality on your device.

Follow these steps to enable VXLAN:

Before you begin

- Ensure you have administrative access to the device.

- Confirm that your device and software release support VXLAN and EVPN features.

Procedure

-
- Step 1** Enter global configuration mode: **configure terminal**
 - Step 2** Enable VLAN-based VXLAN: **feature vn-segment**
 - Step 3** Enable NV overlay functionality: **feature nv overlay**
 - Step 4** Enable VN-Segment for VLANs: **feature vn-segment-vlan-based**
 - Step 5** Enable Switch Virtual Interface (SVI) support: **feature interface-vlan**
 - Step 6** Activate the EVPN control plane for VXLAN: **nv overlay evpn**
-

The device now supports VXLAN and EVPN functionality. To view a configuration example, see the example sections.

Configure VLAN and VXLAN VNI

Use this task to set up VLANs with associated VXLAN VNI mappings, typically required for Layer 2 network segmentation and virtualization on devices participating in EVPN fabrics.

Follow these steps to configure a VLAN and VXLAN VNI:



Note Steps 4 to 7 are optional for configuring the VLAN for VXLAN VNI. These steps are required only if you need a custom route distinguisher or route-target, rather than using auto derivation.

Before you begin

- Ensure you have administrative access to the network device.
- Identify the VLAN and VNI numbers to be used.

Procedure

-
- Step 1** Enter global configuration mode: **configure terminal**
 - Step 2** Specify the VLAN: **vlan *number***
 - Step 3** Map the VLAN to a VXLAN VNI: **vn-segment *number***
This configures Layer 2 VNI under VXLAN VLAN.
 - Step 4** Enter EVPN configuration mode for the VLAN: **evpn**
 - Step 5** Specify the VNI for the EVPN instance: **vni *number* *l2***
 - Step 6** (Optional) Specify the MAC-VRF's route distinguisher: **rd *auto***
 - Step 7** (Optional) Configure the route target for import and export of MAC prefixes: **route-target both {*auto* | *rt*}**

Use `auto` for iBGP. For eBGP or asymmetric VNIs, manually specify the RT. Supported formats for RT include `ASN2:NN`, `ASN4:NN`, or `IPV4:NN`.

The VLAN and VXLAN VNI are configured with the appropriate EVPN settings. You can specify custom RD and RT if needed. For a configuration example, refer to the example sections.

Configure VRF for VXLAN routing

Use this task when setting up a VRF for VXLAN routing on your device. Steps 3–6 are optional unless you require custom route distinguisher or route-target values.

Follow these steps to configure the VRF for VXLAN routing:



Note Step 4 to step 7 are optional for configuring the VRF for VXLAN Routing and are only necessary in case of a custom route distinguisher or route-target requirement (not using auto derivation).

Procedure

- Step 1** Enter configuration mode: **configure terminal**
 - Step 2** Create or select the VRF context: **vrf context** *vrf-name*
 - Step 3** Specify the VNI for the VRF: **vni** *number*
 - Step 4** Specify the IP-VRF's route distinguisher: **rd** **auto**
 - Step 5** Configure the IPv4 or IPv6 unicast address family: **address-family** {**ipv4** | **ipv6**} **unicast**
 - Step 6** Configure the route target for import and export of IPv4 or IPv6 prefixes: **route-target both** {**auto** | *rt*}
- The RT is used for a per-IP-VRF prefix import/export policy. If you enter an RT, these formats are supported: `ASN2:NN`, `ASN4:NN`, or `IPV4:NN`.

Note

Specifying the **auto** option is applicable only for iBGP.

Manually configured route targets are required for eBGP and for asymmetric VNIs.

- Step 7** Set RTs specifically for EVPN: **route-target both** {**auto** | *rt*} **evpn**
- The RT is used for a per-VRF prefix import/export policy. If you enter an RT, these formats are supported: `ASN2:NN`, `ASN4:NN`, or `IPV4:NN`.

Note

Specifying the **auto** option is applicable only for iBGP.

Manually configured route targets are required for eBGP and for asymmetric VNIs.

Configure SVI for core-facing VXLAN routing

You can set up the SVI for Layer 3 VXLAN routing towards the core.

Follow these steps to configure SVI for core-facing VXLAN routing:

Procedure

-
- Step 1** Configure VLAN: **vlan number**
 - Step 2** Map VLAN to VXLAN VNI: **vn-segment number**
 - Step 3** Configure VLAN interface: **interface vlan-number**
 - Step 4** Set the MTU size: **mtu number**
Specify the MTU size in bytes (68-9216).
 - Step 5** Assign the interface to VRF: **vrf member vrf-name**
 - Step 6** Disable sending IP redirect messages for IPv4 and IPv6: **no {ip | ipv6} redirects**
 - Step 7** Enable IPv4-based forwarding: **ip forward**
 - Step 8** Enable IPv6 forwarding: **ipv6 address use-link-local-only**

Note

This command enables the switch to perform IP-based lookup for IPv6 even when the interface VLAN has no configured IPv6 address.

The SVI is now configured for core-facing VXLAN routing, enabling Layer 3 connectivity as required.

Configure SVI for host-facing VXLAN routing

You can configure a host-facing SVI for VXLAN routing.

Follow these steps to configure the SVI for hosts, acting as Distributed Default Gateway.

Before you begin

Ensure you have:

- Administrator access on the switch.
- VXLAN and VRF features are enabled.
- The system is operating on a supported platform and version.

Procedure

-
- Step 1** Configure the distributed gateway virtual MAC address: **fabric forwarding anycast-gateway-mac address**

Note

- One virtual MAC address per VTEP.

- All VTEPs must use the same virtual MAC address for proper anycast gateway operation.

- Step 2** Specify the VLAN ID: **vlan** *number*
- Step 3** Specify the VN-segment identifier for the VLAN: **vn-segment** *number*
- Step 4** Specify the VLAN interface: **interface** *vlan-number*
- Step 5** Assign the interface to the required VRF: **vrf member** *vrf-name*
- Step 6** Specify the SVI IP address: **ip address** *address*
- Step 7** Associate the SVI with anycast gateway functionality under VLAN configuration mode: **fabric forwarding mode anycast-gateway**

The switch interface is configured as a distributed anycast gateway for hosts in the VXLAN overlay.

What to do next

Validate the SVI configuration and ensure host connectivity across the VXLAN network.

Configure the NVE interface and VNIs using multicast

You can enable VXLAN network virtualization and multicast group communication to support bridging and routing for tenant networks.

Follow these steps to configure the NVE interface and VNIs using multicast:

Before you begin

- Enable the NVE feature set on the device.
- Confirm that the required loopback interfaces (for example, loopback1) are configured.
- Ensure multicast routing is enabled in the underlying network.

Procedure

-
- Step 1** Configure the NVE interface: **interface** *nve-interface*
- Step 2** Bind the NVE source-interface to a dedicated loopback interface: **source-interface** *loopback1*
- Step 3** Define BGP as the mechanism for host reachability advertisement: **host-reachability protocol** *bgp*
- Step 4** Configure the multicast group globally for all VNIs on each NVE interface: **global mcast-group** *ip-address* {L2 | L3}

This configuration applies to all Layer 2 and Layer 3 VNIs.

Note

The Layer 3 multicast group is used only for Tenant Routed Multicast (TRM).

- Step 5** Add Layer 2 VNIs to the tunnel interface: **member vni** *vni*
- Step 6** Configure the mcast group on a per-VNI basis: **mcast-group** *ip address*
- Add a multicast group specific to the Layer 2 VNI to override the global configuration.

Note

You can configure ingress replication instead of a multicast group.

Step 7 Add a Layer-3 VNI for each tenant VRF to the overlay: **member vni vni associate-vrf**

Note

Required for VXLAN routing only.

Step 8 Configure the mcast group on a per-VNI basis: **mcast-group address**

Add a multicast group specific to the Layer 3 VNI to override the global configuration.

The NVE interface and associated VNIs are configured with multicast groups for VXLAN, enabling efficient tenant traffic forwarding.

What to do next

Verify the operational status of the NVE interface and VNIs. Ensure that multicast routing is configured correctly in the network.

Configure VXLAN EVPN ingress replication

Configure ingress replication for VXLAN EVPN to enable efficient handling of BUM traffic. The device exchanges VTEP IP addresses using the BGP EVPN control plane.

Follow these steps to set up VXLAN EVPN with ingress replication:

Before you begin

Ensure you have:

- Enabled VXLAN.
- Configured VLAN and VXLAN VNI.
- Configured BGP on the VTEP.
- Configured RD and Route Targets for VXLAN Bridging.

Procedure

Step 1 Configure the NVE interface: **interface nve-interface**

Step 2 Define BGP as the mechanism for host reachability advertisement: **host-reachability protocol bgp**

Step 3 Enable the VTEP globally for all VNIs: **global ingress-replication protocol bgp**

This configuration allows the device to exchange local and remote VTEP IP addresses. It creates the ingress replication list, enables sending and receiving BUM traffic for the VNI, and overrides the global configuration.

Note

Using ingress-replication protocol bgp avoids the need for any multicast configurations that might have been required for configuring the underlay.

Step 4 Add one Layer-3 VNI per tenant VRF to the overlay: **member vni vni associate-vrf**

Note

This is required only for VXLAN routing.

Step 5 Add Layer 2 VNIs to the tunnel interface: **member vni vni**

Step 6 Enable the VTEP to exchange local and remote VTEP IP addresses for each VNI to create the ingress replication list: **ingress-replication protocol bgp**

This enables sending and receiving BUM traffic for the VNI and overrides the global configuration.

Note

- Instead of ingress replication, you can configure a multicast group.
- Configuring **ingress-replication protocol bgp** eliminates the need for multicast configurations that the underlay might require.

VXLAN EVPN ingress replication is successfully enabled using the BGP control plane, allowing the VTEP to send and receive BUM traffic across all configured VNIs.

What to do next

Verify VXLAN EVPN ingress replication functionality. Check connectivity among VTEPs if needed.

Configure BGP on the VTEP

You can configure BGP peers and enable the EVPN address family for proper VXLAN routing and control plane signaling.

Follow these steps to configure BGP on the VTEP:

Before you begin

- Ensure you have console access to the VTEP.
- Ensure you know the local AS number, router ID, and neighbor addresses.

Procedure

Step 1 Enter BGP router configuration mode: **router bgp number**

Step 2 Specify the router ID: **router-id address**

Step 3 Define each BGP neighbor and its remote AS: **neighbor address remote-as number**

Define the L2VPN EVPN address family under each neighbor.

Step 4 Enter the Layer 2 VPN EVPN address family under BGP: **address-family l2vpn evpn**

Note

Use `address-family ipv4 evpn` for VXLAN host-based routing if required.

- Step 5** (Optional) (If using eBGP and leaf switches share an AS number) Allow duplicate AS numbers in the AS path: **Allowas-in**
- Step 6** Configure extended community sending to neighbors: **send-community extended**
- Step 7** If you use VRFs, specify the VRF context and address family for IPv4 and IPv6:
Example:
vrf vrf-name address-family ipv4 unicast address-family ipv4 unicast
- Step 8** Enable ECMP for EVPN-transported IP prefixes in address families as required: **maximum-paths path {ibgp}**
- Step 9** Save the configuration and validate BGP and EVPN status with appropriate show commands.

BGP is configured on the VTEP, supporting EVPN signaling and proper VXLAN overlay routing.

Configure iBGP for EVPN on the spine

You can set up iBGP on a spine device to enable EVPN functionality across the data center fabric.

Follow these steps to configure iBGP for EVPN on the spine:

Before you begin

- Ensure you have assigned IP addressing and AS numbers for your BGP topology.
- Ensure that leaf and spine devices are reachable.

Procedure

-
- Step 1** Enter BGP configuration mode and specify the spine's autonomous system number: **router bgp autonomous system number**
- Step 2** Define each leaf as a BGP neighbor and specify its remote AS: **neighbor address remote-as number**
- Step 3** Enter the Layer 2 VPN EVPN address-family configuration mode under the BGP neighbor: **address-family l2vpn evpn**
- Step 4** Configure the spine to send extended communities to all EVPN neighbors: **send-community extended**
- Step 5** Configure the spine to act as a route reflector for its neighbors: **route-reflector-client**
- Step 6** (If using eBGP) Configure the spine to retain all route-target information in advertisements: **retain route-target all**
- 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 7** (If using eBGP and AS numbers differ) Configure the spine to disable the peer AS check: **address-family l2vpn evpn**
- Step 8** (Optional) Disable the peer AS number check during route advertisement: **disable-peer-as-check**
- Configure this parameter on the spine for eBGP when all leaves are using the same AS but the spines have a different AS than leaves.

Note

Required for eBGP.

Step 9 (If needed) Apply a route map to ensure the next-hop remains unchanged when advertising to neighbors: **route-map permitall out**

Note

Required for eBGP.

The spine is now configured for iBGP EVPN, enabling proper route reflection and propagation within the fabric.

What to do next

Verify BGP neighbor status and EVPN route propagation using **show bgp l2vpn evpn summary** command.

Configure eBGP for EVPN on the spine

You can set up eBGP on the spine switch to support EVPN by configuring the next-hop attribute, retaining route targets, and specifying neighbor information.

Follow these steps to configure eBGP for EVPN on the spine:

Before you begin

- Confirm that spine and leaf devices are reachable via IP and BGP.
- Collect the autonomous system numbers, neighbor IP addresses, and route-target values.

Procedure

Step 1 Configure a route-map to keep the next-hop unchanged for EVPN routes: **route-map NEXT-HOP-UNCH permit 10**

Step 2 Set next-hop address: **set ip next-hop unchanged**

Note

- When two next hops are enabled, the system does not maintain next hop ordering.
- If one next hop is a VXLAN next hop and the other is a local next hop reachable via FIB/AM/Hmm, the system always selects the local next hop, regardless of the order.
- The system always gives priority to directly or locally connected next hops over remotely connected ones.

Step 3 Specify BGP: **router bgp autonomous system number**

Step 4 Configure address family Layer 2 VPN EVPN under the BGP neighbor: **address-family l2vpn evpn**

Step 5 Configure retain route-target all under address-family Layer 2 VPN EVPN: **retain route-target all**

Note

This is required for eBGP. It allows the spine to retain and advertise all EVPN routes when there are no local VNI's configured with matching import route targets.

- Step 6** Define neighbor: **neighbor** *address* **remote-as** *number*
- Step 7** Configure address family Layer 2 VPN EVPN under the BGP neighbor: **address-family** *l2vpn* **evpn**
- Step 8** Disable the peer AS number check during route advertisement: **disable-peer-as-check**
- If all leaf devices use the same autonomous system and the spines use a different autonomous system, configure this parameter on the spine for eBGP.
- Step 9** Allow extended community attributes and apply the NEXT-HOP-UNCH route-map to the neighbor.
- Example:**
- ```
send-community extended
route-map NEXT-HOP-UNCH out
```
- 

## Configure ARP suppression

ARP suppression optimizes network efficiency by allowing the switch to respond to ARP requests from its local cache instead of broadcasting them across the VLAN. The cache is populated by learning remote host IP or MAC information through BGP EVPN MAC advertisements. If the cache lacks a requested entry, the ARP request is broadcast to detect silent hosts.

Follow these steps to configure ARP suppression:

### Before you begin

- Confirm your switch platform supports ARP suppression.
- Ensure you have access to the device and required privileges.

### Procedure

---

- Step 1** Create the network virtualization endpoint (NVE) interface: **interface** *nve* **1**
- Step 2** Configure ARP suppression globally for all Layer 2 VNIs within the NVE interface: **global suppress-arp**
- Step 3** Specify VNI ID: **member vni** *vni-id*
- Step 4** Configure to suppress ARP under Layer 2 VNI and overrides the global set default: **suppress-arp**
- Step 5** Disable the global ARP suppression setting on a specific VNI: **suppress-arp disable**
- 

## Disable VXLANs

You can completely disable VXLAN features and related EVPN control plane settings on your network device.

Follow these steps to disable VXLANs:

### Before you begin

Ensure you have administrator access to the device.

**Procedure**

- Step 1** Enter configuration mode: **configure terminal**
- Step 2** Disable the EVPN control plane: **no nv overlay evpn**
- Step 3** Disable the global mode for all VXLAN bridge domains: **no feature vn-segment-vlan-based**
- Step 4** Disable the VXLAN overlay feature: **no feature nv overlay**
- Step 5** (Optional) Save the changes persistently through reboots and restarts: **copy running-config startup-config**

## Duplicate host detection mechanisms for IP and MAC addresses

Cisco NX-OS supports duplicate host detection for both IP addresses and MAC addresses to prevent conflicts in complex networks. Detection mechanisms rely on monitoring the number of host moves within set time intervals and automatically act when configured thresholds are met.

### Duplicate IP address detection

The system detects duplicate IP addresses based on the number of detected "moves" within a specified time interval.

- Default threshold: 5 moves within 180 seconds.
- When a host appears simultaneously under two VTEPs,, the host mobility logic detects this event.
- Refresh timeout for simultaneous host detection: 600 milliseconds (IPv4) and default 3 seconds (IPv6).
- After exceeding the move threshold, the switch starts a 30-second hold-down (lock) timer to verify duplication persists (**show fabric forwarding ip local-host-db vrf abc**).
- This hold-down lock can occur 5 times in 24 hours; thereafter, the entry is permanently frozen.
- When an IP address is permanently frozen, a syslog message is generated by HMM.

```
2021 Aug 26 01:08:26 leaf hmm: (vrf-name) [IPv4] Freezing potential duplicate host
192.0.2.30/32, reached recover count (5) threshold
```

**Table 2: Example Commands for Duplicate IP Detection**

| Command                                                                                           | Description                                                                                                                                                                       |
|---------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <pre>switch(config)# fabric forwarding ?   anycast-gateway-mac   dup-host-ip-addr-detection</pre> | <p>Available sub-commands:</p> <ul style="list-style-type: none"> <li>• Anycast gateway MAC of the switch.</li> <li>• To detect duplicate host addresses in n seconds.</li> </ul> |
| <pre>switch(config)# fabric forwarding dup-host-ip-addr-detection ?   &lt;1-1000&gt;</pre>        | <p>The number of host moves allowed in n seconds. The range is 1 to 1000 moves; default is 5 moves.</p>                                                                           |

| Command                                                                                      | Description                                                                                                                       |
|----------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------|
| <pre>switch(config)# fabric forwarding dup-host-ip-addr-detection 100 ? &lt;-36000&gt;</pre> | The duplicate detection timeout in seconds for the number of host moves. The range is 2 to 36000 seconds; default is 180 seconds. |
| <pre>switch(config)# fabric forwarding dup-host-ip-addr-detection 100 10</pre>               | Detects duplicate host addresses (limited to 100 moves) in a period of 10 seconds.                                                |

### Duplicate MAC address detection

MAC address duplication is detected using a similar move-count and time-interval mechanism:

- After 5 moves in 180 seconds, a 30-second lock is applied.
- This can occur up to 3 times in 24 hours before the switch permanently freezes the MAC entry.
- A MAC entry stays frozen as long as matching local and remote entries exist (**show l2rib internal permanently-frozen-list**).
- Unconfiguring these detection commands resets them to default but does not disable freeze functionality.

- **l2rib dup-host-mac-detection**

- **l2rib dup-host-recovery**

- Permanently frozen MAC addresses trigger a syslog message from 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
192.0.2.4, 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 1
```

**Table 3: Example Commands for Duplicate MAC Detection**

| Command                                                                          | Description                                                                                                                                                                                                                       |
|----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <pre>switch(config)# l2rib dup-host-mac-detection ? &lt;1-1000&gt; default</pre> | <p>Available sub-commands for L2RIB:</p> <ul style="list-style-type: none"> <li>• The number of host moves allowed in n seconds. The range is 1 to 1000 moves.</li> <li>• Default setting (5 moves in 180 in seconds).</li> </ul> |
| <pre>switch(config)# l2rib dup-host-mac-detection 100 ? &lt;2-36000&gt;</pre>    | 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                                                                        |
|------------------------------------------------------------------|------------------------------------------------------------------------------------|
| <code>switch(config)# l2rib dup-host-mac-detection 100 10</code> | Detects duplicate host addresses (limited to 100 moves) in a period of 10 seconds. |

## Configure event history size for L2RIB

Adjusting the event history size enables enhanced troubleshooting and logging for L2RIB features on the switch.

Follow these steps to set the event history size for L2RIB.

### Procedure

- 
- Step 1** Enter global configuration mode.
- Example:**
- ```
switch# configure terminal
```
- Step 2** Set the event history size for the L2RIB component.
- Example:**
- ```
switch(config)# l2rib event-history mac size low
```
- Step 3** Clear the set event history size for the L2RIB component.
- Example:**
- ```
switch(config)# clear l2rib event-history mac size low
```
-

Verifying the VXLAN BGP EVPN Configuration

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

Command	Purpose
<code>show nve vrf</code>	Displays VRFs and associated VNIs
<code>show bgp l2vpn evpn</code>	Displays routing table information.
<code>show ip arp suppression-cache [detail summary vlan <i>vlan</i> statistics]</code>	Displays ARP suppression information.
<code>show vxlan interface</code>	Displays VXLAN interface status.

Command	Purpose
<code>show vxlan interface count</code>	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 $10 \times 10 = 100$.
<code>show l2route evpn mac [all evi evi [bgp local static vxlan arp]]</code>	Displays Layer 2 route information.
<code>show l2route evpn fl all</code>	Displays all fl routes.
<code>show l2route evpn imet all</code>	Displays all imet routes.
<code>show l2route evpn mac-ip all</code> <code>show l2route evpn mac-ip all detail</code>	Displays all MAC IP routes.
<code>show l2route topology</code>	Displays Layer 2 route topology.
<code>show l2route evpn ethernet-segment all detail</code>	Displays detailed information about all Ethernet Segment Identifiers (ESIs) in an EVPN (Ethernet VPN) environment.



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.

VXLAN BGP EVPN iBGP topologies

A VXLAN BGP EVPN iBGP topology is a data center network design that

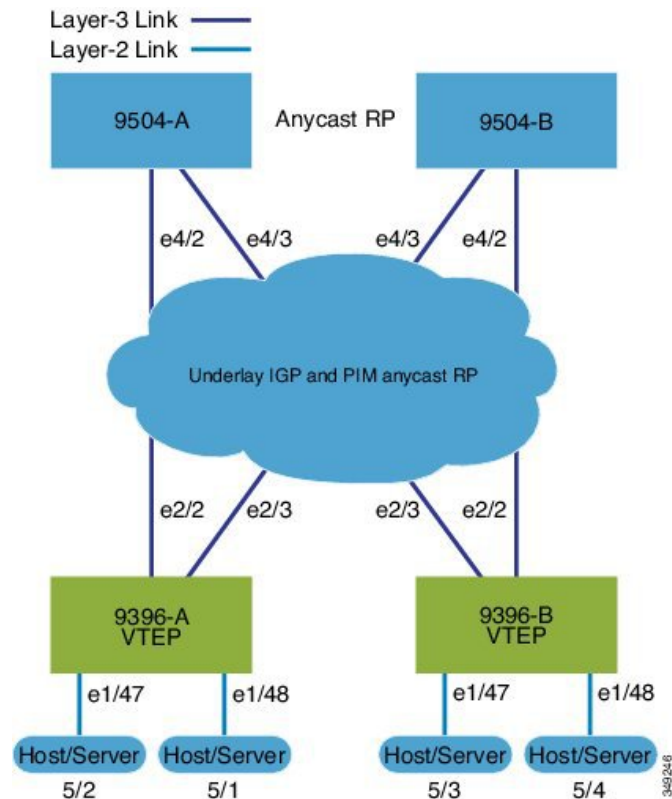
- deploys VXLAN as an overlay technology to provide scalable Layer 2 connectivity over an IP-based Layer 3 fabric
- establishes internal BGP (iBGP) sessions between leaf and spine switches to propagate endpoint reachability information, and
- leverages Ethernet VPN (EVPN) as the control plane to enable efficient MAC and IP address advertisement, host mobility, and multipathing.

VXLAN BGP EVPN iBGP topologies are commonly used in modern data centers to deliver high-performance, scalable, and flexible multi-tenant network environments. In this architecture, all switches participate in the same BGP autonomous system (AS). The switches use iBGP for control plane signaling. EVPN overcomes the traditional limitations of flood-and-learn VXLAN by providing granular control and enhanced route advertisement.

VXLAN BGP EVPN (iBGP) configuration example

The following illustrates a sample VXLAN BGP EVPN iBGP topology and key configuration elements:

Figure 1: VXLAN BGP EVPN Topology (iBGP)



- Spine Switches: Enable the EVPN control plane, OSPF, and PIM. Configure loopback interfaces for router IDs and Anycast-RP. Configure iBGP neighbors among all spine and leaf nodes.
- Leaf Switches: Enable EVPN, OSPF, PIM, interface-VLAN, and VN-segment-VLAN features. Configure loopback interfaces for router ID and VTEP. Define VLAN-to-VNI mapping, enable distributed anycast gateways, and activate NVE (Network Virtualization Edge) interfaces.
- Anycast-RP and ARP suppression: Use Anycast-RP, and configure TCAM regions for ARP suppression across the fabric.
- BGP EVPN configuration: Configure iBGP neighbors (often using loopbacks as update sources) and EVPN address-families to signal MAC/IP information over the VXLAN overlays.
- SVIs and VNIs: Define SVIs for VLANs mapped to VNIs, assigning server and core-facing interfaces.

Spine (9504-A)

- Protocol and feature enablement

- Enable the EVPN control plane

```
nv overlay evpn
```

- Enable the relevant protocols

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

- Loopback configuration

- Configure Loopback for local Router ID, PIM, and BGP

```
interface loopback0
 ip address 192.0.1.1/32
 ip router ospf 1 area 0.0.0.0
 ip pim sparse-mode
```

- Configure Loopback for Anycast RP

```
interface loopback1
 ip address 192.0.2.1/32
 ip router ospf 1 area 0.0.0.0
 ip pim sparse-mode
```

- Configure Anycast RP

```
ip pim rp-address 192.0.2.1 group-list 224.0.0.0/4
ip pim anycast-rp 192.0.2.1 192.0.1.1
ip pim anycast-rp 192.0.2.1 198.51.100.20
```

- 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 192.0.1.1
 neighbor 198.51.100.30 remote-as 65535
 update-source loopback0
 address-family l2vpn evpn
 send-community both
 route-reflector-client
 neighbor 198.51.100.40 remote-as 65535
 update-source loopback0
 address-family l2vpn evpn
 send-community both
 route-reflector-client
```

Spine (9504-B)

• Protocol and feature enablement

- Enable the EVPN control plane

```
nv overlay evpn
```

- Enable the relevant protocols

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

• Loopback configuration

- Configure Loopback for local Router ID, PIM, and BGP

```
interface loopback0
 ip address 198.51.100.20/32
 ip router ospf 1 area 0.0.0.0
 ip pim sparse-mode
```

- Configure Loopback for Anycast RP

```
interface loopback1
 ip address 192.0.2.1/32
 ip router ospf 1 area 0.0.0.0
 ip pim sparse-mode
```

- Configure Anycast RP

```
ip pim rp-address 192.0.2.1 group-list 224.0.0.0/4
ip pim anycast-rp 192.0.2.1 192.0.1.1
ip pim anycast-rp 192.0.2.1 198.51.100.20
```

• Enable OSPF for underlay routing

```
router ospf 1
```

• Configure interfaces for Spine-leaf interconnect

```
interface Ethernet4/2
 ip address 192.168.3.42/24
 ip router ospf 1 area 0.0.0.0
 ip pim sparse-mode
 no shutdown
interface Ethernet4/3
 ip address 192.168.4.43/24
 ip router ospf 1 area 0.0.0.0
 ip pim sparse-mode
 no shutdown
```

• Configure BGP

```
router bgp 65535
 router-id 198.51.100.20
 neighbor 198.51.100.30 remote-as 65535
 update-source loopback0
 address-family l2vpn evpn
 send-community both
 route-reflector client
 neighbor 198.51.100.40 remote-as 65535
```

```

update-source loopback0
address-family l2vpn evpn
  send-community both
  route-reflector client

```

Leaf (9396-A)

- Protocol and feature enablement

- 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

```

- Enable OSPF for underlay routing

```
router ospf 1
```

- Loopback configuration

- Configure Loopback for local Router ID, PIM, and BGP

```

interface loopback0
  ip address 198.51.100.30/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode

```

- Configure Loopback for local VTEP IP

```

interface loopback1
  ip address 198.51.100.33/32
  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
  shutdown

```

- Route Maps for Host SVIs and PIM RP configurations

- Configure route-map to Redistribute Host-SVI (Silent Host)

```
route-map HOST-SVI permit 10
  match tag 54321
```

- Configure PIM RP

```
ip pim rp-address 192.0.2.1 group-list 224.0.0.0/4
```

- VLAN, VRF, and VNI configuration

- Create VLANs

```
vlan 1001-1002
```

- Create overlay VRF VLAN and configure vn-segment

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

- SVI configuration

- Configure Core-facing SVI for VXLAN routing

```
interface vlan101
  no shutdown
  vrf member vxlan-900001
  ip forward
  no ip redirects
  ipv6 address use-link-local-only
  no ipv6 redirects
```

- 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
  rd auto
```



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

```
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 a server facing SVI and enable distributed anycast-gateway.

```
interface vlan1001
  no shutdown
  vrf member vxlan-900001
```

```

ip address 198.51.100.10/24 tag 54321
ipv6 address 2001:db8::1/64 tag 54321
fabric forwarding mode anycast-gateway
interface vlan1002
no shutdown
vrf member vxlan-900001
ip address 198.51.100.1/24 tag 54321
ipv6 address 2001:db8::2/64 tag 54321
fabric forwarding mode anycast-gateway

```

- ARP suppression and NVE interface configurations
 - Configure ACL TCAM region for ARP suppression



Note The **hardware access-list tcam region arp-ether 256 double-wide** command is not needed for Cisco Nexus 9300-EX and 9300-FX/FX2 /FX3 and 9300-GX platform switches.

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

- NVE interface - Basic mode and simplified mode options (if required)



Note You can choose either of these two options for creating the NVE interface. Use Option 1 for a small number of VNIs. Use Option 2 to leverage the simplified configuration mode.

- Option 1:

```

interface nve1
no shutdown
source-interface loopback1
host-reachability protocol bgp
member vni 900001 associate-vrf
member vni 2001001
  mcast-group 239.0.0.1
member vni 2001002
  mcast-group 239.0.0.1

```

- Option 2:

```

interface nve1
source-interface loopback1
host-reachability protocol bgp
global mcast-group 239.0.0.1 L2
member vni 2001001
member vni 2001002
member vni 2001007-2001010

```

- Configure interfaces for hosts/servers

```

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

```

- Configure BGP

```
router bgp 65535
  router-id 198.51.100.30
  neighbor 10.1.1.1 remote-as 65535
    update-source loopback0
    address-family l2vpn evpn
      send-community both
  neighbor 198.51.100.20 remote-as 65535
    update-source loopback0
    address-family l2vpn evpn
      send-community both
  vrf vxlan-900001
    address-family ipv4 unicast
      redistribute direct route-map HOST-SVI
    address-family ipv6 unicast
      redistribute direct route-map HOST-SVI
```



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

```
evpn
  vni 2001001 12
  vni 2001002 12
```



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 commands in EVPN mode do not need to be entered.

```
evpn
  vni 2001001 12
    rd auto
    route-target import auto
    route-target export auto
  vni 2001002 12
    rd auto
    route-target import auto
    route-target export auto
```

Leaf (9396-B)

• Protocol and feature enablement

• 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
```

• Enable OSPF for underlay routing

```
router ospf 1
```

• Loopback configuration

• Configure Loopback for local Router ID, PIM, and BGP

```
interface loopback0
 ip address 198.51.100.40/32
 ip router ospf 1 area 0.0.0.0
 ip pim sparse-mode
```

• Configure Loopback for local VTEP IP

```
interface loopback1
 ip address 192.0.1.1/32
 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 192.168.3.22/24
 ip router ospf 1 area 0.0.0.0
 ip pim sparse-mode
 no shutdown
interface Ethernet2/3
 no switchport
 ip address 192.168.4.23/24
 ip router ospf 1 area 0.0.0.0
 ip pim sparse-mode
 shutdown
```

• Route Maps for Host SVIs and PIM RP configurations

• Configure route-map to Redistribute Host-SVI (Silent Host)

```
route-map HOST-SVI permit 10
 match tag 54321
```

• Configure PIM RP

```
ip pim rp-address 192.0.2.1 group-list 224.0.0.0/4
```

- VLAN, VRF, and VNI configuration

- Create VLANs

```
vlan 1001-1002
```

- Create overlay VRF VLAN and configure vn-segment

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

- SVI configuration

- Configure Core-facing SVI for VXLAN routing

```
interface vlan101
  no shutdown
  vrf member vxlan-900001
  ip forward
  no ip redirects
  ipv6 address use-link-local-only
  no ipv6 redirects
```

- 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
  rd auto
```



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

```
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 198.51.100.10/24
  ipv6 address 2001:db8::1/64
  fabric forwarding mode anycast-gateway
interface vlan1002
  no shutdown
  vrf member vxlan-900001
  ip address 198.51.100.1/24
  ipv6 address 2001:db8::2/64
  fabric forwarding mode anycast-gateway
```

- ARP suppression and interface configuration
 - Configure ACL TCAM region for ARP suppression



Note The **hardware access-list tcam region arp-ether 256 double-wide** command is not needed for Cisco Nexus 9300-EX and 9300-FX/FX2 /FX3 and 9300-GX platform switches.

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



Note You can choose either of these two command procedures for creating the NVE interfaces. Use Option 1 for a small number of VNIs. Use Option 2 to leverage the simplified configuration mode.

- NVE interface - Basic mode and simplified mode options (if required)



Note You can choose either of these two options for creating the NVE interface. Use Option 1 for a small number of VNIs. Use Option 2 to leverage the simplified configuration mode.

- Option 1:

```
interface nve1
  no shutdown
  source-interface loopback1
  host-reachability protocol bgp
  member vni 900001 associate-vrf
  member vni 2001001
    mcast-group 239.0.0.1
  member vni 2001002
    mcast-group 239.0.0.1
```

- Option 2:

```
interface nve1
  source-interface loopback1
  host-reachability protocol bgp
  global mcast-group 239.0.0.1 L2
  member vni 2001001
  member vni 2001002
  member vni 2001007-2001010
```

- Configure interface vlan on Border Gateway (BGW)

```
interface vlan101
  no shutdown
  vrf member evpn-tenant-3103101
  no ip redirects
  ip address 198.51.100.50/16
  ipv6 address 2001:db8::1/48
  no ipv6 redirects
  fabric forwarding mode anycast-gateway
```

- Configure interfaces for hosts/servers

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

- Configure BGP

```
router bgp 65535
  router-id 198.51.100.40
  neighbor 10.1.1.1 remote-as 65535
    update-source loopback0
    address-family l2vpn evpn
      send-community both
  neighbor 198.51.100.20 remote-as 65535
    update-source loopback0
    address-family l2vpn evpn
      send-community both
vrf vxlan-900001
vrf vxlan-900001
  address-family ipv4 unicast
    redistribute direct route-map HOST-SVI
  address-family ipv6 unicast
    redistribute direct route-map HOST-SVI
```



Note

- These commands in EVPN mode do not need to be entered.

```
evpn
  vni 2001001 12
  vni 2001002 12
```

- 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
```

- These commands in EVPN mode do not need to be entered.

```
evpn
  vni 2001001 12
    rd auto
    route-target import auto
    route-target export auto
  vni 2001002 12
    rd auto
    route-target import auto
    route-target export auto
```



Note When you have iBGP session between BGWs and EBGW fabric is used, you need to configure the route-map to make VIP or VIP_R route advertisement with higher AS-PATH when local VIP or VIP_R is down (due to reload or fabric link flap). A sample route-map configuration is provided below. In this example 192.0.2.1 is VIP address and 198.51.100.1 is BGP VIP route's nexthop learned from same BGW site.

```
ip prefix-list vip_ip seq 5 permit 192.0.2.1/32
ip prefix-list vip_route_nh seq 5 permit 198.51.100.1/32
route-map vip_ip permit 5
  match ip address prefix-list vip_ip
  match ip next-hop prefix-list vip_route_nh
  set as-path prepend 5001 5001 5001
route-map vip_ip permit 10
```

VXLAN BGP EVPN eBGP topologies

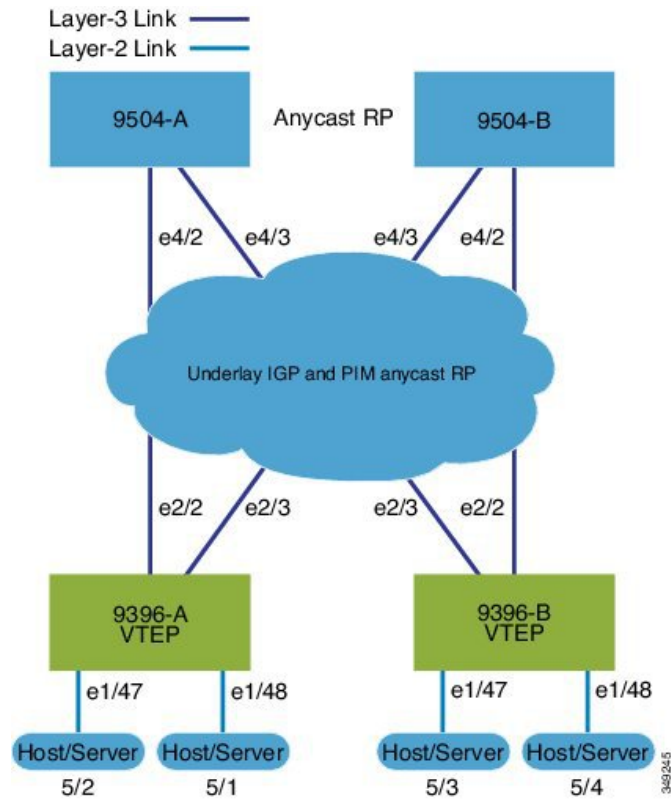
A VXLAN BGP EVPN eBGP topology is a data center overlay network architecture that

- provides scalable Layer 2 and Layer 3 segmentation across a spine-leaf fabric
- leverages BGP EVPN as a control plane for distributing MAC and IP address reachability information, and
- enables automated multi-tenancy, efficient host mobility, and distributed gateway functions.

VXLAN BGP EVPN (eBGP) configuration example

The following example demonstrates a typical VXLAN BGP EVPN deployment using eBGP as the routing protocol between spine and leaf nodes. This design supports scalable, flexible data center environments.

Figure 2: VXLAN BGP EVPN Topology (eBGP)



Spine (9504-A)

- Protocol and feature enablement
 - Enable the EVPN control plane


```
nv overlay evpn
```
 - Enable the relevant features


```
feature bgp
feature pim
```
- Loopback and Anycast RP configuration
 - Configure the loopback for the local router ID, PIM, and BGP.


```
interface loopback0
ip address 198.51.100.1/32 tag 12345
ip pim sparse-mode
```
 - Configure loopback for Anycast RP


```
interface loopback1
ip address 192.0.2.1/32 tag 12345
ip pim sparse-mode
```
 - Configure the Anycast RP.

```

ip pim rp-address 192.0.2.1 group-list 224.0.0.0/4
ip pim anycast-rp 192.0.2.1 198.51.100.1
ip pim anycast-rp 192.0.2.1 198.51.100.2

```

- Route maps

- Configure route-map used by eBGP for Spine

```

route-map NEXT-HOP-UNCH permit 10
set ip next-hop unchanged

```

- Configure route-map to Redistribute Loopback

```

route-map LOOPBACK permit 10
match tag 12345

```

- Interface configuration - Spine-leaf interconnect

- Configure the interfaces for the 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

```

- BGP overlay (EVPN)

- Configure the BGP overlay for the EVPN address family.

```

router bgp 100
router-id 198.51.100.1
address-family l2vpn evpn
nexthop route-map NEXT-HOP-UNCH
retain route-target all
neighbor 192.0.2.0 remote-as 200
update-source loopback0
ebgp-multihop 3
address-family l2vpn evpn
send-community both
disable-peer-as-check
route-map NEXT-HOP-UNCH out
neighbor 198.51.100.2 remote-as 200
update-source loopback0
ebgp-multihop 3
address-family l2vpn evpn
send-community both
disable-peer-as-check
route-map NEXT-HOP-UNCH out

```

- BGP underlay (IPv4 unicast)

- Configure BGP underlay for the IPv4 unicast address family.

```

address-family ipv4 unicast
redistribute direct route-map LOOPBACK
neighbor 192.168.1.22 remote-as 200
update-source ethernet4/2
address-family ipv4 unicast
allowas-in
disable-peer-as-check

```

```
neighbor 192.168.2.23 remote-as 200
update-source ethernet4/3
address-family ipv4 unicast
allowas-in
disable-peer-as-check
```

Spine (9504-B)

- Protocol and feature enablement

- Enable the EVPN control plane

```
nv overlay evpn
```

- Enable the relevant protocols

```
feature bgp
feature pim
```

- Loopback and Anycast RP configuration

- Configure the loopback for the local router ID, PIM, and BGP.

```
interface loopback0
ip address 192.0.2.20/32 tag 12345
ip pim sparse-mode
```

- Configure the loopback for Anycast RP.

```
interface loopback1
ip address 192.0.2.1/32 tag 12345
ip pim sparse-mode
```

- Configure the Anycast RP.

```
ip pim rp-address 192.0.2.1 group-list 224.0.0.0/4
ip pim anycast-rp 192.0.2.1 198.51.100.1
ip pim anycast-rp 192.0.2.1 192.0.2.20
```

- Route maps

- Configure route-map used by eBGP for Spine

```
route-map NEXT-HOP-UNCH permit 10
set ip next-hop unchanged
```

- Configure route-map to Redistribute Loopback

```
route-map LOOPBACK permit 10
match tag 12345
```

- Interface configuration - Spine-leaf interconnect

- Configure interfaces for Spine-leaf interconnect

```
interface Ethernet4/2
no switchport
ip address 192.168.3.42/24
ip router ospf 1 area 0.0.0.0
ip pim sparse-mode
no shutdown
interface Ethernet4/3
```

```

no switchport
ip address 192.168.4.43/24
ip router ospf 1 area 0.0.0.0
ip pim sparse-mode
shutdown

```

- BGP overlay (EVPN)

- Configure BGP overlay for the EVPN address family

```

router bgp 100
  router-id 192.0.2.20
  address-family l2vpn evpn
    nexthop route-map NEXT-HOP-UNCH
    retain route-target all
  neighbor 192.0.2.0 remote-as 200
  update-source loopback0
  ebgp-multihop 3
  address-family l2vpn evpn
    send-community both
    disable-peer-as-check
    route-map NEXT-HOP-UNCH out
  neighbor 198.51.100.2 remote-as 200
  update-source loopback0
  ebgp-multihop 3
  address-family l2vpn evpn
    send-community both
    disable-peer-as-check
    route-map NEXT-HOP-UNCH out

```

- BGP underlay (IPv4 unicast)

- Configure the BGP underlay for the IPv4 unicast address family.

```

address-family ipv4 unicast
  redistribute direct route-map LOOPBACK
  neighbor 192.168.3.22 remote-as 200
  update-source ethernet4/2
  address-family ipv4 unicast
    allows-in
    disable-peer-as-check
  neighbor 192.168.4.43 remote-as 200
  update-source ethernet4/3
  address-family ipv4 unicast
    allows-in
    disable-peer-as-check

```

Leaf (9396-A)

- Feature and protocol enablement

- Enable the EVPN control plane and networking features:

```

nv overlay evpn
  feature bgp
  feature pim
  feature interface-vlan
  feature vn-segment-vlan-based
  feature nv overlay
  fabric forwarding anycast-gateway-mac 0000.2222.3333

```

- Routing Protocol Initialization

- Enable OSPF for underlay routing.

```
router ospf 1
```

- Loopback interfaces

- Configure Loopback for local router ID, PIM, and BGP and loopback1 (VTEP)

```
interface loopback0
 ip address 192.0.2.0/32
 ip pim sparse-mode
interface loopback1
 ip address 198.51.100.0/32
 ip pim sparse-mode
```

- Spine-Leaf and host interface configuration

- Set up interfaces for fabric and access

```
interface Ethernet2/2
 no switchport
 ip address 192.168.1.22/24
 ip pim sparse-mode
 no shutdown
interface Ethernet2/3
 no switchport
 ip address 192.168.4.23/24
 ip pim sparse-mode
 shutdown
interface Ethernet1/47
 switchport
 switchport access vlan 1002
interface Ethernet1/48
 switchport
 switchport access vlan 1001
```

- VLAN, VRF, and VNI configuration

- Create VLANs and map them to VXLAN segments.

```
vlan 1001-1002
 vlan 101
  vn-segment 900001
 vlan 1001
  vn-segment 2001001
 vlan 1002
  vn-segment 2001002
```

- Create the VRF and address families.

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



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

- SVI configuration
 - Configure core-facing and server-facing SVIs

```
interface vlan101
  no shutdown
  vrf member vxlan-900001
  ip forward
  no ip redirects
  ipv6 address use-link-local-only
  no ipv6 redirects

interface vlan1001
  no shutdown
  vrf member vxlan-900001
  ip address 192.0.2.10/24 tag 54321
  ipv6 address 2001:DB8:1:1::1/64 tag 54321
  fabric forwarding mode anycast-gateway
interface vlan1002
  no shutdown
  vrf member vxlan-900001
  ip address 198.51.100.20/24 tag 54321
  ipv6 address 2001:DB8:1:1::1/64 tag 54321
  fabric forwarding mode anycast-gateway
```

- PIM RP, ARP suppression, and NVE interface
 - Enable PIM RP and configure ARP suppression region

```
ip pim rp-address 192.0.2.1 group-list 224.0.0.0/4
  hardware access-list tcam region arp-ether 256 double-wide
```



Note The **hardware access-list tcam region arp-ether 256 double-wide** command is not needed for Cisco Nexus 9300-EX and 9300-FX/FX2 /FX3 and 9300-GX platform switches.

- NVE interface - Basic mode and simplified mode options (if required)



Note You can choose either of the following two options for creating the NVE interface. Use Option 1 for a small number of VNIs. Use Option 2 to leverage the simplified configuration mode.

- Option 1:

```
interface nve1
  no shutdown
  source-interface loopback1
  host-reachability protocol bgp
  member vni 900001 associate-vrf
```

```

member vni 2001001
  mcast-group 239.0.0.1
member vni 2001002
  mcast-group 239.0.0.1

```

- Option 2:

```

interface nve1
  source-interface loopback1
  host-reachability protocol bgp
  global mcast-group 239.0.0.1 L2
  member vni 2001001
  member vni 2001002
  member vni 2001007-2001010

```

- Route Maps for Host SVIs and BGP configuration for underlay and overlay

- Configure host SVI redistribution (Silent Host).

```

route-map HOST-SVI permit 10
  match tag 54321

```

- Underlay (IPv4 unicast)

```

router bgp 200
  router-id 192.0.2.0
  address-family ipv4 unicast
    redistribute direct route-map LOOPBACK
  neighbor 192.168.1.42 remote-as 100
  update-source ethernet2/2
  address-family ipv4 unicast
    allowas-in
    disable-peer-as-check
  neighbor 192.168.4.43 remote-as 100
  update-source ethernet2/3
  address-family ipv4 unicast
    allowas-in
    disable-peer-as-check

```

- Overlay (EVPN address family)

```

address-family l2vpn evpn
  nexthop route-map NEXT-HOP-UNCH
  retain route-target all
  neighbor 198.51.100.1 remote-as 100
  update-source loopback0
  ebgp-multihop 3
  address-family l2vpn evpn
    send-community both
    disable-peer-as-check
    route-map NEXT-HOP-UNCH out
  neighbor 192.0.2.20 remote-as 100
  update-source loopback0
  ebgp-multihop 3
  address-family l2vpn evpn
    send-community both
    disable-peer-as-check
    route-map NEXT-HOP-UNCH out
vrf vxlan-900001

```

**Note**

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

```
evpn
vni 2001001 12
vni 2001002 12
```

- 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
```

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

```
evpn
vni 2001001 12
rd auto
route-target import auto
route-target export auto
vni 2001002 12
rd auto
route-target import auto
route-target export auto
```

Leaf (9396-B)

- Enable the control plane and features.

```
nv overlay evpn
feature bgp
feature pim
feature interface-vlan
feature vn-segment-vlan-based
feature nv overlay
fabric forwarding anycast-gateway-mac 0000.2222.3333
```

- Enable OSPF for underlay routing.

```
router ospf 1
```

- Configure interfaces.

- Configure loopback interfaces.

```
interface loopback0
ip address 192.0.4.1/32
ip pim sparse-mode
interface loopback1
ip address 192.0.44.1/32
ip pim sparse-mode
```

- Configure spine-leaf and server/host interfaces

```
interface Ethernet2/2
no switchport
ip address 192.168.3.22/24
ip pim sparse-mode
```

```

no shutdown
interface Ethernet2/3
no switchport
ip address 192.168.2.23/24
ip pim sparse-mode
shutdown
interface Ethernet1/47
switchport
switchport access vlan 1002
interface Ethernet1/48
switchport
switchport access vlan 1001

```

- Configure VLANs and VXLAN segment mapping

```

vlan 1001-1002
vlan 101
vn-segment 900001
vlan 1001
vn-segment 2001001
vlan 1002
vn-segment 2001002

```

- Configure VRF and address families

```

vrf context vxlan-900001
vni 900001
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 SVIs for routing and host connectivity

```

interface vlan101
no shutdown
vrf member vxlan-900001
ip forward
no ip redirects
ipv6 address use-link-local-only
no ipv6 redirects
interface vlan1001
no shutdown
vrf member vxlan-900001
ip address 192.0.2.10/24 tag 54321
ipv6 address 2001:DB8:1:1::1/64 tag 54321
fabric forwarding mode anycast-gateway
interface vlan1002
no shutdown
vrf member vxlan-900001
ip address 198.51.100.20/24 tag 54321
ipv6 address 2001:DB8:1:1::1/64 tag 54321
fabric forwarding mode anycast-gateway

```

- Configure PIM RP, ARP suppression, and NVE interface

- Configure PIM RP, ARP suppression

```

ip pim rp-address 192.0.2.1 group-list 224.0.0.0/4
hardware access-list tcam region arp-ether 256 double-wide

```



Note The **hardware access-list team region arp-ether 256 double-wide** command is not needed for Cisco Nexus 9300-EX and 9300-FX/FX2 /FX3 and 9300-GX platform switches.

- Configure NVE interface - Basic mode and simplified mode options (if required)



Note You can choose either of the following two options for creating the NVE interface. Use Option 1 for a small number of VNIs. Use Option 2 to leverage the simplified configuration mode.

- Option 1:

```
interface nve1
  no shutdown
  source-interface loopback1
  host-reachability protocol bgp
  member vni 900001 associate-vrf
  member vni 2001001
    mcast-group 239.0.0.1
  member vni 2001002
    mcast-group 239.0.0.1
```

- Option 2:

```
interface nve1
  source-interface loopback1
  host-reachability protocol bgp
  global mcast-group 239.0.0.1 L2
  member vni 2001001
  member vni 2001002
  member vni 2001007-2001010
```

- Configure route-maps and BGP for underlay and overlay

```
route-map HOST-SVI permit 10
  match tag 54321

router bgp 200
  router-id 192.0.4.1
  address-family ipv4 unicast
    redistribute direct route-map LOOPBACK
  neighbor 192.168.3.42 remote-as 100
  update-source ethernet2/2
  address-family ipv4 unicast
    allows-in
    disable-peer-as-check
  neighbor 192.168.2.43 remote-as 100
  update-source ethernet2/3
  address-family ipv4 unicast
    allows-in
    disable-peer-as-check

address-family l2vpn evpn
  nexthop route-map NEXT-HOP-UNCH
  retain route-target all
  neighbor 198.51.100.1 remote-as 100
```

```

update-source loopback0
ebgp-multihop 3
address-family l2vpn evpn
  send-community both
  disable-peer-as-check
  route-map NEXT-HOP-UNCH out
neighbor 192.0.2.20 remote-as 100
update-source loopback0
ebgp-multihop 3
address-family l2vpn evpn
  send-community both
  disable-peer-as-check
  route-map NEXT-HOP-UNCH out
vrf vxlan-900001

```

**Note**

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

```

evpn
vni 2001001 l2
vni 2001002 l2

```

- 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

```

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

```

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

```

Sample outputs for EVPN/VXLAN show commands

This topic provides example outputs for various **show** commands used in Nexus switches for EVPN/VXLAN troubleshooting and verification.

This list describes the available **show** commands and their sample outputs:

- **show nve peers**

```

9396-B# show nve peers
Interface Peer-IP           State LearnType Uptime  Router-Mac
-----
nve1      203.0.113.1             Up     CP         00:00:38 6412.2574.9f27

```

- **show nve vni**

```

9396-B# show nve vni
Codes: CP - Control Plane      DP - Data Plane

```

UC - Unconfigured

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]	
nve1	2001002	225.4.0.1	Up	CP	L2	[1002]	

• show vxlan interface



Note The **show vxlan interface** command is not supported for the Cisco Nexus 9300-EX, 9300-FX/FX2/FX3, and 9300-GX platform switches.

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

ARP suppression

• 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
192.0.2.54	00:06:41	0054.0000.0000	1001	Ethernet1/48	L
192.0.2.51	00:20:33	0051.0000.0000	1001	(null)	R
198.51.100.53	00:06:41	0053.0000.0000	1002	Ethernet1/47	L
198.51.100.52	00:20:33	0052.0000.0000	1002	(null)	R

L2 EVPN

• show bgp l2vpn evpn summary

```
leaf3# show bgp l2vpn evpn summary
BGP summary information for VRF default, address family L2VPN EVPN
BGP router identifier 40.0.0.4, local AS number 10
BGP table version is 60, L2VPN EVPN config peers 1, capable peers 1
21 network entries and 21 paths using 2088 bytes of memory
BGP attribute entries [8/1152], BGP AS path entries [0/0]
BGP community entries [0/0], BGP clusterlist entries [1/4]
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
209.165.200.1		4	10	8570	8565	60	0	0	5d22h 6

```
leaf3#
```

• show bgp l2vpn evpn

```
leaf3# show bgp l2vpn evpn
BGP routing table information for VRF default, address family L2VPN EVPN
BGP table version is 60, local router ID is 209.165.200.4
```

```
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
Route Distinguisher: 209.165.200.2:32868
*>i[2]:[0]:[10001]:[48]:[0000.8816.b645]:[0]:[0.0.0.0]/216
209.165.200.2          100             0 i
*>i[2]:[0]:[10001]:[48]:[0011.0000.0034]:[0]:[0.0.0.0]/216
209.165.200.2          100             0 i
```

L2 Route

• show l2route evpn mac all

```
leaf3# show l2route evpn mac all
Topology      Mac Address      Prod    Next Hop (s)
-----
101           0000.8816.b645  BGP    209.165.200.2
101           0001.0000.0033  Local  Ifindex 4362086
101           0001.0000.0035  Local  Ifindex 4362086
101           0011.0000.0034  BGP    209.165.200.2
```

• show l2route evpn mac-ip all

```
leaf3# show l2route evpn mac-ip all
Topology ID Mac Address      Prod Host IP          Next Hop (s)
-----
101           0011.0000.0034  BGP  10.1.3.2            209.165.200.2
102           0011.0000.0034  BGP  10.1.3.2            209.165.200.2
```

VXLAN EVPN with Downstream VNI

This chapter contains these sections:

VXLAN EVPN with downstream VNIs

A VXLAN EVPN with downstream VNIs is a network feature that

- enables asymmetric VNI communication across nodes in a VXLAN EVPN network
- provides customers access to a common shared service outside of their domain (tenant VRF), and
- supports communication between isolated VXLAN EVPN sites that have different sets of VNIs.

Cisco NX-OS Release 9.3(5) introduces the VXLAN EVPN with downstream VNI feature. In earlier releases, VNI configuration required consistency across all nodes in the VXLAN EVPN network to enable communication.

Asymmetric VNIs

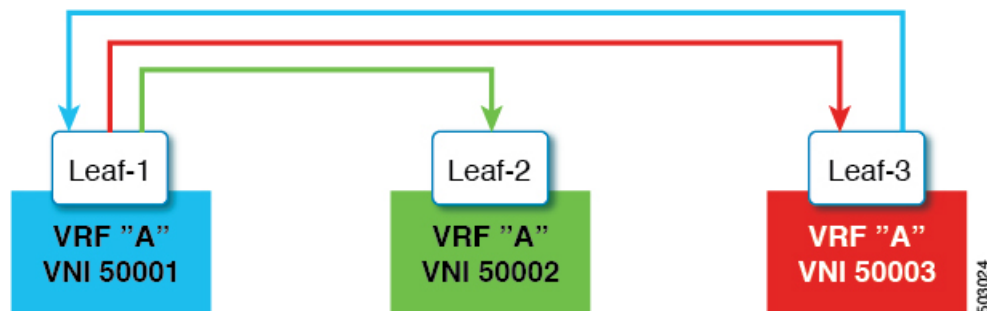
Asymmetric VNIs are a VXLAN configuration method that

- assign different VXLAN network identifiers (VNIs) for the same IP VRF or MAC VRF across multiple VTEPs

- enable routing and bridging in networks with downstream VNI support, and
- allow flexible segmentation by accommodating non-uniform VNI assignments on different endpoints.

VXLAN EVPN with downstream VNI supports asymmetric VNI allocation. In the following topology, all three VTEPs have different VNIs configured for the same IP VRF or MAC VRF.

Figure 3: Asymmetric VNIs



Shared services VRFs

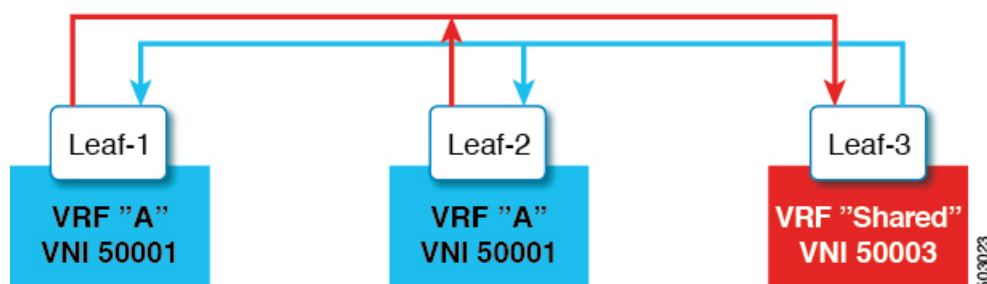
A shared services VRF is a virtual routing and forwarding instance that

- allows importing multiple L3VRFs into a single local L3VRF,
- supports disparate values of downstream L3VNIs on a per-peer basis, and
- enables shared services VRFs to provide services (such as DNS) to multiple tenant VRFs.

For example, a DNS server needs to serve multiple hosts in a data center regardless of the tenant VRFs on which the hosts sit. The DNS server is attached to a shared services VRF, which is attached to an L3VNI. To access this server from any of the tenant VRFs, the switches must import the routes from the shared services VRF to the tenant VRF, even though the L3VNI associated to the shared services VRF is different from the L3VNI associated to the tenant VRF.

In the following illustration, Tenant VRF A in Leaf-1 can communicate with Tenant VRF A in Leaf-2. However, Tenant VRF A requires access to a shared service sitting behind Leaf-3.

Figure 4: Shared Services VRFs



Multi-site deployments with asymmetric VNIs

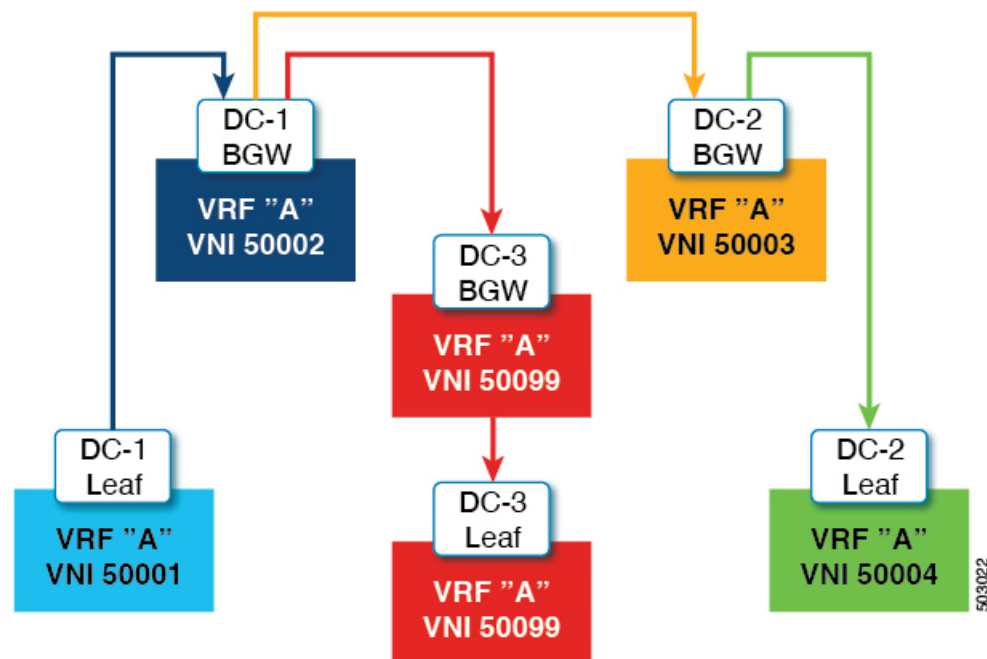
Multi-site deployments with asymmetric VNIs are VXLAN EVPN network architectures that

- connect data center sites using different sets of VNIs within their internal networks,
- enable seamless communication between sites by translating and stitching asymmetric VNIs at the border gateways, and
- support integration between sites that use both symmetric and asymmetric VNI allocations.

Asymmetric VNI mapping occurs when sites assign VNIs independently, which creates mismatches across the network. Border gateways translate these mismatches so that traffic can flow between sites. This translation allows organizations to connect multiple data centers, even if internal VNI assignments are different.

In the illustrated example, DC-1 and DC-2 are asymmetric sites with different VNI sets, while DC-3 is a symmetric site with consistent VNI mapping. The border gateways between the sites facilitate communication by adjusting VNI assignments as needed.

Figure 5: Multi-Site with Asymmetric VNIs



Supported features for VXLAN EVPN with downstream VNI

VXLAN EVPN with downstream VNI has the following guidelines and limitations.

VXLAN EVPN downstream VNI is supported on various Cisco Nexus platforms. This support enables flexible network segmentation and advanced encryption features.

- Cisco Nexus 9332C, 9364C, 9300-EX, and 9300-FX/FX2/FXP platform switches and Cisco Nexus 9500 platform switches with -EX/FX line cards support VXLAN EVPN with downstream VNI.

Unsupported features

- VXLAN EVPN with downstream VNI is currently not supported with these feature combinations:
 - VXLAN static tunnels
 - TRM and TRM with Multi-Site
 - CloudSec VXLAN EVPN Tunnel Encryption
 - ESI-based multihoming
 - Seamless integration of EVPN with L3VPN (MPLS SR)
 - VXLAN policy-based routing (PBR)
 - IPv6 underlay
 - DSVNI with IPv6 underlay
- iBGP sessions between vPC peer nodes in a VRF are not supported.
- VXLAN consistency checker is not supported for VXLAN EVPN with downstream VNI.

ISSU

- For existing centralized VRF route leaking deployments, a brief traffic loss might occur during ISSU to Cisco NX-OS Release 9.3(5) or later.
- For successful downgrade from Cisco NX-OS Release 9.3(5) to a prior release, ensure that the asymmetric VNI configuration has been removed. Downstream VNI is not supported before Cisco NX-OS Release 9.3(5) and hence traffic forwarding would be impacted.

Configuration and operational guidelines

- Make sure that you configure L2VNI SVI on Anycast BGW to enable DSVNI MAC-IP Layer 3 label translation in a multisite environment. The functionality of DSVNI is limited for reoriginated routes, which requires an association between L2VNI and VRF. You can associate using the VRF member command in L2VNI SVI.
- Downstream VNI is configured based on route-target export and import. The following conditions must be met to leverage Downstream VNI:
 - Downstream VNI requires the usage of different VRF (MAC-VRF or IP-VRF), each VRF must have a different VNI (Asymmetric VNI).
 - To import routes of a foreign VRF (MAC-VRF or IP-VRF) the appropriate route-target for the import into the local VRF must be configured.
 - The configuration of only auto-derived route-targets will not result in downstream VNI.
 - The export of VRF prefixes can be done by static or auto-derived route-target configuration.
 - The import of a foreign VRF's auto-derived route-target is supported.
 - The import of a foreign VRFs statically configured route-target is supported.
- Downstream VNI is supported for the following underlay constellations:

- For downstream VNI with Layer-3 VNI, the underlay can be ingress replication or multicast based.
- For downstream VNI with Layer-2 VNI, the underlay must be in ingress replication. Multicast based underlay is not supported with downstream VNI of Layer-2 VNIs.
- Downstream VNI requires consistent configuration:
 - All multi-site Border Gateway (BGW) in a site must have a consistent configuration.
 - All vPC members in a vPC domain must have consistent configuration.
- The usage of downstream VNI with multi-site requires all BGW across all sites to run at least Cisco NX-OS Release 9.3(5).
- Layer-3 VNIs (IP-VRF) can be flexibly mapped between VNIs per peer.
 - VNI 50001 on VTEP1 can perform symmetric VNI with VNI 50001 and asymmetric VNI with VNI 50002 on VTEP2 at the same time.
 - VNI 50001 on VTEP1 can perform asymmetric VNI with VNI 50002 on VTEP2 and VNI 50003 on VTEP3.
 - VNI 50001 on VTEP1 can perform asymmetric VNI with VNI 50002 and VNI5003 on VTEP2 at the same time.
- Layer-2 VNIs (MAC-VRF) can only be mapped to one VNI per peer.
 - VNI 30001 on VTEP1 can perform asymmetric VNI with VNI 30002 on VTEP2 and VNI 30003 on VTEP3.
 - VNI 30001 on VTEP1 cannot perform asymmetric VNI with VNI 30002 and VNI 3003 on VTEP2 at the same time.
- BGP peering across VXLAN and Downstream VNI support the following constellations:
 - BGP peering between symmetric VNI is supported by using loopbacks.
 - BGP peering between asymmetric VNI is supported if the VNIs are in a direct message relationship. A loopback from VNI 50001 (on VTEP1) can peer with a loopback in VNI 50002 (on VTEP2).
 - BGP peering between asymmetric VNI is supported if the VNIs are in a direct message relationship but on different VTEPs. A loopback from VNI 50001 (on VTEP1) can peer with a loopback in VNI 50002 (on VTEP2 and VTEP3).
 - BGP peering between asymmetric VNI is not supported if the VNIs are in a 1:N relationship. A loopback in VNI 50001 (VTEP1) can't peer with a loopback in VNI 50002 (VTEP2) and VNI 50003 (VTEP3) at the same time.

Verification commands for VXLAN EVPN with downstream VNIs

These commands provide verification and troubleshooting information for VXLAN EVPN configurations that use downstream VNIs. Each command displays operational data, and sample outputs help clarify the information.

Table 4: Command Reference for VXLAN EVPN with Downstream VNI

Command	Purpose
show bgp evi l2-evi	Displays the VRF associated with an L2VNI.
show forwarding adjacency nve platform	Displays both symmetric and asymmetric NVE adjacencies with the corresponding DestInfoIndex.
show forwarding route vrfvrf	Displays the egress VNI or downstream VNI for each next-hop.
show ip route detail vrfvrf	Displays the egress VNI or downstream VNI for each next-hop.
show l2route evpn mac-ip all detail	Displays labeled next-hops that are present in the remote MAC routes.
show l2route evpn imet all detail	Displays the egress VNI associated with the remote peer.
show nve peers control-plane-vni peer-ipip-address	Displays the egress VNI or downstream VNI for each NVE adjacency.

Sample outputs

- **show bgp evi l2-evi:**

```
switch# show bgp evi 100
-----
L2VNI ID           : 100 (L2-100)
RD                 : 3.3.3.3:32867
Secondary RD      : 1:100
Prefixes (local/total) : 1/6
Created           : Jun 23 22:35:13.368170
Last Oper Up/Down : Jun 23 22:35:13.369005 / never
Enabled          : Yes
Associated IP-VRF : vni100
Active Export RT list :
    100:100
Active Import RT list :
    100:100
```

- **show forwarding adjacency nve platform:**

```
switch# show forwarding adjacency nve platform
slot 1
=====
IPv4 NVE adjacency information

next_hop:192.0.2.12 interface:nve1 (0x49000001) table_id:1
Peer_id:0x49080002 dst_addr:192.0.2.12 src_addr:192.0.2.13 RefCt:1 PBRct:0
Flags:0x440800
cp : TRUE, DCI peer: FALSE is_anycast_ip FALSE dsvni peer: FALSE
HH:0x7a13f DstInfoIndex:0x3002
tunnel init: unit-0:0x3 unit-1:0x0

next_hop:192.0.2.12 interface:nve1 (0x49000001) table_id:1
Peer_id:0x49080002 dst_addr:192.0.2.12 src_addr:192.0.2.13 RefCt:1 PBRct:0
```

```
Flags:0x10440800
cp : TRUE, DCI peer: FALSE is_anycast_ip FALSE dsvni peer: TRUE
  HH:0x7a142 DstInfoIndex:0x3ffd
  tunnel init: unit-0:0x6 unit-1:0x0
...
```

• **show forwarding route vrf***vrf*:

```
switch# show forwarding route vrf vrf1000
```

```
slot 1
=====
```

```
IPv4 routes for table vrf1000/base
```

Prefix	Next-hop	Interface	Labels	Partial Install
10.1.1.11/32	192.0.2.12	nve1	dsvni: 301000	
10.1.1.20/32	198.51.100.123	nve1	dsvni: 301000	
10.1.1.21/32	192.0.2.30	nve1	dsvni: 301000	
10.1.1.30/32	10.1.1.30	Vlan10		

• **show ip route detail vrf***vrf*:

```
switch# show ip route detail vrf default
```

```
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>
```

```
193.0.1.0/24, ubest/mbest: 4/0
  *via 192.0.2.2, Eth1/1, [100/0], 00:00:05, urib_dt6-client1 segid: 6544, tunnelid:
  0x7b9 encap: VXLAN

  *via 192.0.2.3, Eth1/1, [100/0], 00:00:05, urib_dt6-client1 segid: 6545,
  (Asymmetric) tunnelid: 0x7ba encap: VXLAN

  *via 192.0.2.4, Eth1/1, [100/0], 00:00:05, urib_dt6-client1 segid: 6546,
  (Asymmetric) tunnelid: 0x7bb encap: VXLAN
```

• **show l2route evpn mac-ip all detail:**

```
switch# 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 (Orp):Orphan
Topology Mac Address      Host IP    Prod   Flags Seq No  Next-Hops
-----
5          0000.0005.1301 1.3.13.1  BGP   --    0      192.0.2.1 (Label: 2000005)
5          0000.0005.1401 1.3.14.1  BGP   --    0      192.0.2.145 (Label: 2000005)
```

• **show l2route evpn imet all detail:**

```
switch# show l2route evpn imet all
```

```
Flags- (F): Originated From Fabric, (W): Originated from WAN
```

```
Topology ID VNI          Prod IP Addr      Flags
```

```

-----
3          2000003    BGP  192.0.2.1    -
3          2000003    BGP  192.0.31.1   -
3          2000003    BGP  192.0.32.1   -
3          2000003    BGP  192.0.2.145  -

```

- **show nve peers control-plane-vni:**

```

switch# show nve peers control-plane-vni peer-ip 203.0.113.1
Peer      VNI      Learn-Source Gateway-MAC      Peer-type  Egress-VNI SW-BD  State
-----
203.0.113.1 2000003 BGP              f40f.1b6f.f8db  FAB        3000003  3005
peer-vni-add-complete

```

EVPN Centralized Gateway

This chapter contains these sections: