Cisco Nexus 3600 Switch NX-OS VXLAN Configuration Guide, Release 10.4(x)

First Published: 2023-08-18

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
http://www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883

CONTENTS

Preface

This preface includes the following sections:

Reference Preface Map here

Preface

Audience

This publication is for network administrators who install, configure, and maintain Cisco Nexus switches.

Document Conventions

Command descriptions use the following conventions:

Convention	Description
bold	Bold text indicates the commands and keywords that you enter literally as shown.
Italic	Italic text indicates arguments for which the user supplies the values.
[x]	Square brackets enclose an optional element (keyword or argument).
[x | y]	Square brackets enclosing keywords or arguments separated by a vertical bar indicate an optional choice.
{x | y}	Braces enclosing keywords or arguments separated by a vertical bar indicate a required choice.
[x {y | z}]	Nested set of square brackets or braces indicate optional or required choices within optional or required elements. Braces and a vertical bar within square brackets indicate a required choice within an optional element.
`variable`	Indicates a variable for which you supply values, in context where italics cannot be used.
string	A nonquoted set of characters. Do not use quotation marks around the string or the string will include the quotation marks.

Examples use the following conventions:

Convention	Description
`screen font`	Terminal sessions and information the switch displays are in screen font.
`boldface screen font`	Information you must enter is in boldface screen font.
italic screen font	Arguments for which you supply values are in italic screen font.
< >	Nonprinting characters, such as passwords, are in angle brackets.
[ ]	Default responses to system prompts are in square brackets.
!, #	An exclamation point (!) or a pound sign (#) at the beginning of a line of code indicates a comment line.

Related Documentation for Cisco Nexus 3600 Platform Switches

The entire Cisco Nexus 3600 platform switch documentation set is available at the following URL:

http://www.cisco.com/c/en/us/support/switches/nexus-3000-series-switches/tsd-products-support-series-home.html

Documentation Feedback

To provide technical feedback on this document, or to report an error or omission, please send your comments to nexus3k-docfeedback@cisco.com. We appreciate your feedback.

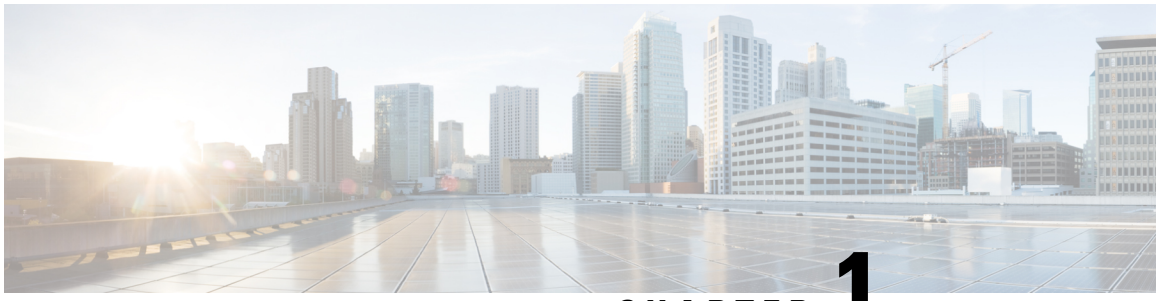
Communications, Services, and Additional Information

- To receive timely, relevant information from Cisco, sign up at Cisco Profile Manager.
- To get the business impact you're looking for with the technologies that matter, visit Cisco Services.
- To submit a service request, visit Cisco Support.
- To discover and browse secure, validated enterprise-class apps, products, solutions and services, visit Cisco Marketplace.
- To obtain general networking, training, and certification titles, visit Cisco Press.
- To find warranty information for a specific product or product family, access Cisco Warranty Finder.

Cisco Bug Search Tool

Cisco Bug Search Tool (BST) is a web-based tool that acts as a gateway to the Cisco bug tracking system that maintains a comprehensive list of defects and vulnerabilities in Cisco products and software. BST provides you with detailed defect information about your products and software.

Communications, Services, and Additional Information



CHAPTER 1

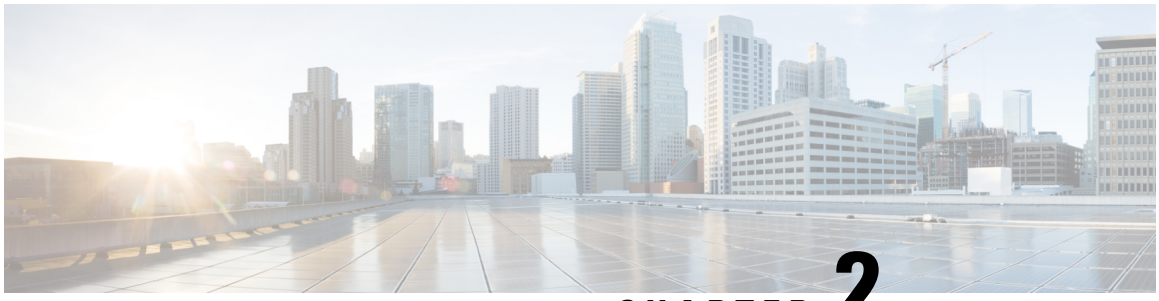
New and Changed Information

New and Changed Information

Table 1: New and Changed Features for Cisco NX-OS Release 10.4(x)

Feature	Description	Changed in Release	Where Documented
NA	No new features added for this release.	10.4(1)F	NA

New and Changed Information



CHAPTER 2

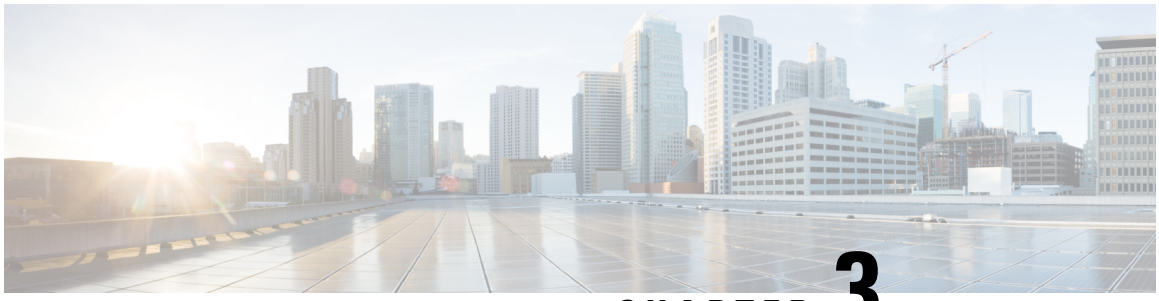
Overview

Licensing Requirements

For a complete explanation of Cisco NX-OS licensing recommendations and how to obtain and apply licenses, see the *Cisco NX-OS Licensing Guide* and the *Cisco NX-OS Licensing Options Guide*.

Supported Platforms

Starting with Cisco NX-OS release 7.0(3)I7(1), use the Nexus Switch Platform Support Matrix to know from which Cisco NX-OS releases various Cisco Nexus 9000 and 3000 switches support a selected feature.



CHAPTER 3

Configuring VXLANs

This chapter contains the following sections:

Overview

VXLAN Overview

The Cisco Nexus 3600 platform switches are designed for a hardware-based Virtual Extensible LAN (VXLAN) function. These switches can extend Layer 2 connectivity across the Layer 3 boundary and integrate between VXLAN and non-VXLAN infrastructures. Virtualized and multitenant data center designs can be shared over a common physical infrastructure.

VXLANs enable you to extend Layer 2 networks across the Layer 3 infrastructure by using MAC-in-UDP encapsulation and tunneling. In addition, you can use a VXLAN to build a multitenant data center by decoupling tenant Layer 2 segments from the shared transport network.

When deployed as a VXLAN gateway, the Cisco Nexus 3600 platform switches can connect VXLAN and classic VLAN segments to create a common forwarding domain so that tenant devices can reside in both environments.

A VXLAN has the following benefits:

- Flexible placement of multitenant segments throughout the data center.

  It extends Layer 2 segments over the underlying shared network infrastructure so that tenant workloads can be placed across physical pods in the data center.

- Higher scalability to address more Layer 2 segments.

  A VXLAN uses a 24-bit segment ID called the VXLAN network identifier (VNID). The VNID allows a maximum of 16 million VXLAN segments to coexist in the same administrative domain. (In comparison, traditional VLANs use a 12-bit segment ID that can support a maximum of 4096 VLANs.)

- Utilization of available network paths in the underlying infrastructure.

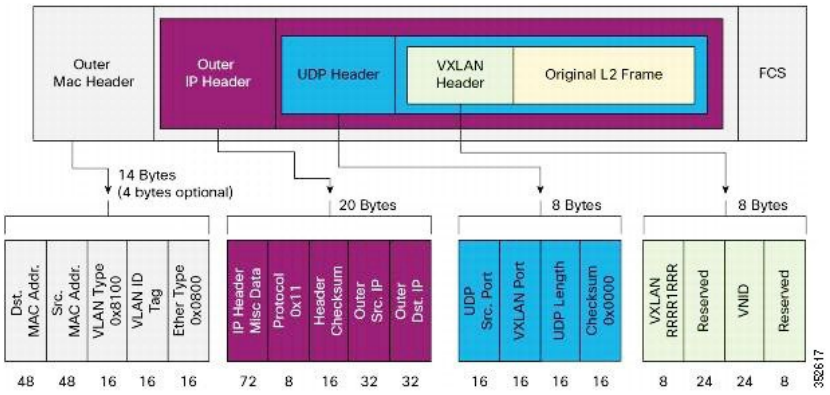
  VXLAN packets are transferred through the underlying network based on its Layer 3 header. It uses equal-cost multipath (ECMP) routing and link aggregation protocols to use all available paths.

VXLAN Encapsulation and Packet Format

A VXLAN is a Layer 2 overlay scheme over a Layer 3 network. It uses MAC-in-UDP encapsulation to extend Layer 2 segments across the data center network. The transport protocol over the physical data center network is IP plus UDP.

A VXLAN defines a MAC-in-UDP encapsulation scheme where the original Layer 2 frame has a VXLAN header added and is then placed in a UDP-IP packet. With this MAC-in-UDP encapsulation, VXLAN tunnels Layer 2 network over the Layer 3 network. The VXLAN packet format is shown in the following figure.

Figure 1: VXLAN Packet Format



A VXLAN uses an 8-byte VXLAN header that consists of a 24-bit VNID and a few reserved bits. The VXLAN header and the original Ethernet frame are in the UDP payload. The 24-bit VNID identifies the Layer 2 segments and maintains Layer 2 isolation between the segments. A VXLAN can support 16 million LAN segments.

VXLAN Tunnel Endpoints

A VXLAN uses VXLAN tunnel endpoint (VTEP) devices to map tenants' end devices to VXLAN segments and to perform VXLAN encapsulation and deencapsulation. Each VTEP device has two types of interfaces:

- Switch port interfaces on the local LAN segment to support local endpoint communication through bridging

- IP interfaces to the transport network where the VXLAN encapsulated frames will be sent

A VTEP device is identified in the IP transport network by using a unique IP address, which is a loopback interface IP address. The VTEP device uses this IP address to encapsulate Ethernet frames and transmits the encapsulated packets to the transport network through the IP interface. A VTEP device learns the remote VTEP IP addresses and the remote MAC address-to-VTEP IP mapping for the VXLAN traffic that it receives.

The VXLAN segments are independent of the underlying network topology; conversely, the underlying IP network between VTEPs is independent of the VXLAN overlay. The IP network routes the encapsulated packets based on the outer IP address header, which has the initiating VTEP as the source IP address and the terminating VTEP or multicast group IP address as the destination IP address.

VXLAN Packet Forwarding Flow

A VXLAN uses stateless tunnels between VTEPs to transmit traffic of the overlay Layer 2 network through the Layer 3 transport network.

ECMP and LACP Load Sharing with VXLANs

Encapsulated VXLAN packets are forwarded between VTEPs based on the native forwarding decisions of the transport network. Most data center transport networks are designed and deployed with multiple redundant paths that take advantage of various multipath load-sharing technologies to distribute traffic loads on all available paths.

A typical VXLAN transport network is an IP-routing network that uses the standard IP equal cost multipath (ECMP) to balance the traffic load among multiple best paths. To avoid out-of-sequence packet forwarding, flow-based ECMP is commonly deployed. An ECMP flow is defined by the source and destination IP addresses and optionally, the source and destination TCP or UDP ports in the IP packet header.

All the VXLAN packet flows between a pair of VTEPs have the same outer source and destination IP addresses, and all VTEP devices must use one identical destination UDP port that can be either the Internet Allocated Numbers Authority (IANA)-allocated UDP port 4789 or a customer-configured port. The only variable element in the ECMP flow definition that can differentiate VXLAN flows from the transport network standpoint is the source UDP port. A similar situation for Link Aggregation Control Protocol (LACP) hashing occurs if the resolved egress interface that is based on the routing and ECMP decision is an LACP port channel. LACP uses the VXLAN outer-packet header for link load-share hashing, which results in the source UDP port being the only element that can uniquely identify a VXLAN flow.

In the Cisco Nexus 3600 platform switches implementation of VXLANs, a hash of the inner frame's header is used as the VXLAN source UDP port. As a result, a VXLAN flow can be unique. The IP address and UDP port combination is in its outer header while the packet traverses the underlay transport network.

Advertising Primary IP Address

On a vPC-enabled leaf or border leaf switch, by default all Layer-3 routes are advertised with the secondary IP address (VIP) of the leaf switch VTEP as the BGP next-hop IP address. Prefix routes and leaf switch generated routes are not synced between vPC leaf switches. Using the VIP as the BGP next-hop for these types of routes can cause traffic to be forwarded to the wrong vPC leaf or border leaf switch and black-holed. The provision to use the primary IP address (PIP) as the next-hop when advertising prefix routes or loopback interface routes in BGP on vPC-enabled leaf or border leaf switches allows users to select the PIP as BGP

next-hop when advertising these types of routes so that traffic will always be forwarded to the right vPC-enabled leaf or border leaf switch.

The configuration command for advertising the PIP is **advertise-pip**.

The following is a sample configuration:

```
switch(config)# router bgp 65536
  address-family 12vpn evpn
    advertise-pip
interface nve 1
    advertise virtual-rmac
```

The **advertise-pip** command lets BGP use the PIP as next-hop when advertising prefix routes or leaf-generated routes if vPC is enabled.

VMAC (virtual-mac) is used with VIP and system MAC is used with PIP when the VIP/PIP feature is enabled.

With the **advertise-pip** and **advertise virtual-rmac** commands enabled, type 5 routes are advertised with PIP and type 2 routes are still advertised with VIP. In addition, VMAC will be used with VIP and system MAC will be used with PIP.

Note The **advertise-pip** and **advertise-virtual-rmac** commands must be enabled and disabled together for this feature to work properly. If you enable or disable one and not the other, it is considered an invalid configuration.

Guidelines and Limitations for VXLANs

VXLAN has the following guidelines and limitations:

- IGMP snooping is supported on VXLAN VLANs.
- VXLAN Layer 2 Gateway functionality is supported.
- VXLAN Flood and Learn functionality is not supported.
- Ensure that the network can accommodate an additional 50 bytes for the VXLAN header.
- Only one Network Virtualization Edge (NVE) interface is supported on a switch.
- Layer 3 VXLAN uplinks are not supported in a nondefault virtual and routing forwarding (VRF) instance.
- Switched Port Analyzer (SPAN) for ports carrying VXLAN-encapsulated traffic is not supported.
- VXLAN with Layer 3 VPN is not supported.
- VXLAN with ingress replication is not supported.
- MLD snooping is not supported on VXLAN VLANs.
- ACLs and QoS policies are not supported on VXLAN VLANs.
- DHCP snooping is not supported on VXLAN VLANs.
- L3VNI's VLAN must be added on the vPC peer-link trunk's allowed VLAN list.

Considerations for VXLAN Deployment

The following are some of the considerations while deploying VXLANs:

- A loopback interface IP is used to uniquely identify a VTEP device in the transport network.
- To establish IP multicast routing in the core, an IP multicast configuration, PIM configuration, and Rendezvous Point (RP) configuration are required.
- You can configure VTEP-to-VTEP unicast reachability through any IGP protocol.
- VXLAN multicast traffic should always use the RPT shared tree.
- An RP for the multicast group on the VTEP is a supported configuration. However, you must configure the RP for the multicast group at the spine layer/upstream device. Because all multicast traffic traverses the RP, it is more efficient to have this traffic directed to a spine layer/upstream device.

Enabling a VXLAN

Enabling VXLANs involves the following:

- Enabling the VXLAN feature
- Enabling VLAN to VN-Segment mapping

Before you begin

Ensure that you have installed the VXLAN Enterprise license.

SUMMARY STEPS

1. switch# **configure terminal**
2. switch(config)# **[no] feature nv overlay**
3. switch (config)# **[no] feature vn-segment-vlan-based**
4. (Optional) switch(config)# **copy running-config startup-config**

DETAILED STEPS

	Command or Action	Purpose
Step 1	switch# **configure terminal**	Enters global configuration mode.
Step 2	switch(config)# **[no] feature nv overlay**	Enables the VXLAN feature.
Step 3	switch (config)# **[no] feature vn-segment-vlan-based**	Configures the global mode for all VXLAN bridge domains. Enables VLAN to VN-Segment mapping. VLAN to VN-Segment mapping is always one-to-one.
Step 4	(Optional) switch(config)# **copy running-config startup-config**	Saves the change persistently through reboots and restarts by copying the running configuration to the startup configuration.

Example

This example shows how to enable a VXLAN and configure VLAN to VN-Segment mapping:

```
switch# configure terminal
switch(config)# feature nv overlay
switch(config)# feature vn-segment-vlan-based
switch(config)# copy running-config startup-config
```

Mapping a VLAN to a VXLAN VNI

SUMMARY STEPS

1. switch# **configure terminal**
2. switch(config)# **vlan** *vlan-id*
3. switch(config-vlan)# **vn-segment** *vnid*

DETAILED STEPS

	Command or Action	Purpose
Step 1	switch# **configure terminal**	Enters global configuration mode.
Step 2	switch(config)# **vlan** *vlan-id*	Specifies a VLAN.
Step 3	switch(config-vlan)# **vn-segment** *vnid*	Specifies the VXLAN virtual network identifier (VNID). The range of values for vnid is 1 to 16777214.

Example

This example shows how to map a VLAN to a VXLAN VNI:

```
switch# configure terminal
switch(config)# vlan 3100
switch(config-vlan)# vn-segment 5000
```

Configuring a Routing Protocol for NVE Unicast Addresses

Configuring a routing protocol for unicast addresses involves the following:

- Configuring a dedicated loopback interface for NVE reachability.
- Configuring the routing protocol network type.
- Specifying the routing protocol instance and area for an interface.
- Enabling PIM sparse mode in case of multicast replication.

Note Open shortest path first (OSPF) is used as the routing protocol in the examples.

SUMMARY STEPS

1. switch# **configure terminal**
2. switch(config)# **interface loopback** *instance*
3. switch(config-if)# **ip address** *ip-address/length*
4. switch(config-if)# **ip ospf network** {**broadcast** | **point-to-point**}
5. switch(config-if)# **ip router ospf** *instance-tag* **area** *area-id*
6. switch(config-if)# **ip pim sparse-mode**

DETAILED STEPS

	Command or Action	Purpose
Step 1	switch# **configure terminal**	Enters global configuration mode.
Step 2	switch(config)# **interface loopback** *instance*	Creates a dedicated loopback interface for the NVE interface. The instance range is from 0 to 1023.
Step 3	switch(config-if)# **ip address** *ip-address/length*	Configures an IP address for this interface.
Step 4	switch(config-if)# **ip ospf network** {**broadcast** | **point-to-point**}	Configures the OSPF network type to a type other than the default for an interface.
Step 5	switch(config-if)# **ip router ospf** *instance-tag* **area** *area-id*	Specifies the OSPF instance and area for an interface.
Step 6	switch(config-if)# **ip pim sparse-mode**	Enables PIM sparse mode on this interface. The default is disabled. Enable the PIM sparse mode in case of multicast replication.

Example

This example shows how to configure a routing protocol for NVE unicast addresses:

```
switch# configure terminal
switch(config)# interface loopback 10
switch(config-if)# ip address 222.2.2.1/32
switch(config-if)# ip ospf network point-to-point
switch(config-if)# ip router ospf 1 area 0.0.0.0
```

Creating and Configuring an NVE Interface

An NVE interface is the overlay interface that initiates and terminates VXLAN tunnels. You can create and configure an NVE (overlay) interface.

SUMMARY STEPS

1. switch# **configure terminal**
2. switch(config)# **interface nve** *instance*
3. switch(config-if-nve)# **source-interface loopback** *instance*

DETAILED STEPS

	Command or Action	Purpose
Step 1	switch# **configure terminal**	Enters global configuration mode.
Step 2	switch(config)# **interface nve** *instance*	Creates a VXLAN overlay interface that initiates and terminates VXLAN tunnels. **Note** Only one NVE interface is allowed on the switch.
Step 3	switch(config-if-nve)# **source-interface loopback** *instance*	Specifies a source interface. The source interface must be a loopback interface that is configured on the switch with a valid /32 IP address. This /32 IP address must be known by the transit routers in the transport network and the remote VTEPs.

Example

This example shows how to create and configure an NVE interface:

```
switch# configure terminal
switch(config)# interface nve 1
switch(config-if-nve)# source-interface loopback 10
```

Configuring a VXLAN VTEP in vPC

SUMMARY STEPS

1. Enter global configuration mode.
2. Enable the vPC feature on the device.
3. Enable the interface VLAN feature on the device.
4. Enable the LACP feature on the device.
5. Enable the PIM feature on the device.
6. Enables the OSPF feature on the device.
7. Define a PIM RP address for the underlay multicast group range.
8. Create the VLAN to be used as a backup link.
9. Carve the TCAM region for the ACL database.
10. Assign the TCAM region for use by a VXLAN.
11. Create the SVI used for the backup routed path over the vPC peer-link.
12. Create primary and secondary IP addresses.

13.
14. Create a vPC domain.
15. Configure the IPv4 address for the remote end of the vPC peer-keepalive link.
16. Enable Peer-Gateway on the vPC domain.
17. Enable Peer-switch on the vPC domain.
18. Enable IP ARP synchronize under the vPC domain to facilitate faster ARP table population following device reload.
19. (Optional) Enable IPv6 nd synchronization under the vPC domain to facilitate faster nd table population following device reload.
20. Create the vPC peer-link port-channel interface and add two member interfaces.
21. Modify the STP hello-time, forward-time, and max-age time.
22. (Optional) Enable the delay restore timer for SVI's.

DETAILED STEPS

Step 1 Enter global configuration mode.

```
switch# configure terminal
```

Step 2 Enable the vPC feature on the device.

```
switch(config)# feature vpc
```

Step 3 Enable the interface VLAN feature on the device.

```
switch(config)# feature interface-vlan
```

Step 4 Enable the LACP feature on the device.

```
switch(config)# feature lacp
```

Step 5 Enable the PIM feature on the device.

```
switch(config)# feature pim
```

Step 6 Enables the OSPF feature on the device.

```
switch(config)# feature ospf
```

Step 7 Define a PIM RP address for the underlay multicast group range.

```
switch(config)# ip pim rp-address 192.168.100.1 group-list 224.0.0/4
```

Step 8 Create the VLAN to be used as a backup link.

```
switch(config)# vlan 10
```

Step 9 Carve the TCAM region for the ACL database.

```
switch(config)# hardware access-list tcam region mac-ifacl 0
```

Note This command is applicable only for the Cisco Nexus 36180YC-R and 3636C-R vPC leaf switches.

Step 10 Assign the TCAM region for use by a VXLAN.

```
switch(config)# hardware access-list tcam region vxlan 10
```

Note This command is applicable only for the Cisco Nexus 36180YC-R and 3636C-R vPC leaf switches.

Step 11 Create the SVI used for the backup routed path over the vPC peer-link.

```
switch(config)# interface vlan 10
switch(config-if)# ip address 10.10.10.1/30
switch(config-if)# ip router ospf UNDERLAY area 0
switch(config-if)# ip pim sparse-mode
switch(config-if)# no ip redirects
switch(config-if)# mtu 9216
```

Step 12 Create primary and secondary IP addresses.

```
switch(config)# interface loopback 0
switch(config-if)# description Control_plane_Loopback
switch(config-if)# ip address x.x.x.x/32
switch(config-if)# ip address y.y.y.y/32 secondary
switch(config-if)# ip router ospf process tag area area id
switch(config-if)# ip pim sparse-mode
switch(config-if)# no shutdown
```

Step 13

```
switch(config)# interface loopback 1
switch(config-if)# description Data_Plane_loopback
switch(config-if)# ip address z.z.z.z/32
switch(config-if)# ip router ospf process tag area area id
switch(config-if)# ip pim sparse-mode
switch(config-if)# no shutdown
```

Step 14 Create a vPC domain.

```
switch(config)# vpc domain 10
```

Step 15 Configure the IPv4 address for the remote end of the vPC peer-keepalive link.

```
switch(config-vpc-domain)# peer-keepalive destination 172.28.x.x
```

Note The system does not form the vPC peer link until you configure a vPC peer-keepalive link

The management ports and VRF are the defaults.

Note We recommend that you configure a separate VRF and use a Layer 3 port from each vPC peer device in that VRF for the vPC peer-keepalive link. For more information about creating and configuring VRFs, see the Cisco Nexus 3600 Series NX-OS Unicast Routing Configuration Guide.

Step 16 Enable Peer-Gateway on the vPC domain.

```
switch(config-vpc-domain)# peer-gateway
```

Note Disable IP redirects on all interface-vlans of this vPC domain for correct operation of this feature.

Step 17 Enable Peer-switch on the vPC domain.

```
switch(config-vpc-domain)# peer-switch
```

Note Disable IP redirects on all interface-vlans of this vPC domain for correct operation of this feature.

Step 18 Enable IP ARP synchronize under the vPC domain to facilitate faster ARP table population following device reload.

```
switch(config-vpc-domain)# ip arp synchronize
```

Step 19 (Optional) Enable IPv6 nd synchronization under the vPC domain to facilitate faster nd table population following device reload.

```
switch(config-vpc-domain)# ipv6 nd synchronize
```

Step 20 Create the vPC peer-link port-channel interface and add two member interfaces.

```
switch(config)# interface port-channel 1
switch(config-if)# switchport
switch(config-if)# switchport mode trunk
switch(config-if)# switchport trunk allowed vlan 1,100-200
switch(config-if)# mtu 9216
switch(config-if)# vpc peer-link
switch(config-if)# no shutdown
switch(config-if)# interface Ethernet 1/1, 1/20
switch(config-if)# switchport
switch(config-if)# mtu 9216
switch(config-if)# channel-group 1 mode active
switch(config-if)# no shutdown
```

Step 21 Modify the STP hello-time, forward-time, and max-age time.

As a best practice, we recommend changing the **hello-time** to four seconds to avoid unnecessary TCN generation when the vPC role change occurs. As a result of changing the **hello-time**, it is also recommended to change the **max-age** and **forward-time** accordingly.

```
switch(config)# spanning-tree vlan 1-3967 hello-time 4
switch(config)# spanning-tree vlan 1-3967 forward-time 30
switch(config)# spanning-tree vlan 1-3967 max-age 40
```

Step 22 (Optional) Enable the delay restore timer for SVI's.

We recommend that you tune this value when the SVI or VNI scale is high. For example, when the SVI count is 1000, we recommend setting the delay restore for interface-vlan to 45 seconds.

```
switch(config-vpc-domain)# delay restore interface-vlan 45
```

Configuring Replication for a VNI

Replication for VXLAN network identifier (VNI) can be configured in one of two ways:

- Multicast replication

Configuring Multicast Replication

Before you begin

- Ensure that the NVE interface is created and configured.
- Ensure that the source interface is specified.

SUMMARY STEPS

1. switch(config-if-nve)# **member vni** {*vnid* **mcast-group** *multicast-group-addr* | *vnid- range* **mcast-group** *start-addr* [*end-addr*]}

DETAILED STEPS

	Command or Action	Purpose
Step 1	switch(config-if-nve)# **member vni** {*vnid* **mcast-group** *multicast-group-addr* | *vnid- range* **mcast-group** *start-addr* [*end-addr*]}	Maps VXLAN VNIs to the NVE interface and assigns a multicast group to the VNIs.

Example

This example shows how to map a VNI to an NVE interface and assign it to a multicast group:

```
switch(config-if-nve)# member vni 5000 mcast-group 225.1.1.1
```

Configuring IGMP Snooping Over VXLAN

Overview of IGMP Snooping Over VXLAN

Starting with Cisco NX-OS Release 7.0(3)F3(4), you can configure IGMP snooping over VXLAN. The configuration of IGMP snooping is same in VXLAN as in configuration of IGMP snooping in regular VLAN domain. For more information on IGMP snooping, see the *Configuring IGMP Snooping* chapter in the Cisco Nexus 3600 NX-OS Multicast Routing Configuration Guide, Release 7.x.

Guidelines and Limitations for IGMP Snooping Over VXLAN

See the following guidelines and limitations for IGMP snooping over VXLAN:

- For IGMP snooping over VXLAN, all the guidelines and limitations of VXLAN apply.
- IGMP snooping over VXLAN is not supported on any FEX enabled platforms and FEX ports.

Configuring IGMP Snooping Over VXLAN

SUMMARY STEPS

1. switch(config)#**ip igmp snooping vxlan**
2. switch(config)#**ip igmp snooping disable-nve-static-router-port**

DETAILED STEPS

	Command or Action	Purpose
Step 1	switch(config)#**ip igmp snooping vxlan**	Enables IGMP snooping for VXLAN VLANs. You have to explicitly configure this command to enable snooping for VXLAN VLANs.
Step 2	switch(config)#**ip igmp snooping disable-nve-static-router-port**	Configures IGMP snooping over VXLAN to not include NVE as static mrouter port using this global CLI command. IGMP snooping over VXLAN has the NVE interface as mrouter port by default.

Verifying the VXLAN Configuration

Use one of the following commands to verify the VXLAN configuration, to display the MAC addresses, and to clear the MAC addresses:

Command	Purpose
show nve interface nve *id*	Displays the configuration of an NVE interface.
show nve vni	Displays the VNI that is mapped to an NVE interface.
show nve peers	Displays peers of the NVE interface.
show nve vxlan-params	Displays the VXLAN UDP port configured.
show mac address-table	Displays both VLAN and VXLAN MAC addresses.
clear mac address-table dynamic	Clears all MAC address entries in the MAC address table.

Example

This example shows how to display the configuration of an NVE interface:

```
switch# show nve interface nve 1
Interface: nve1, State: up, encapsulation: VXLAN
Source-interface: loopback10 (primary: 111.1.1.1, secondary: 0.0.0.0)
```

This example shows how to display the VNI that is mapped to an NVE interface for multicast replication:

```
switch# show nve vni
Interface          VNI        Multicast-group   VNI State
-----------------  --------   ---------------   ---------
nve1               5000       225.1.1.1          Up
```

This example shows how to display the VNI that is mapped to an NVE interface for ingress replication:

```
switch# show nve vni
Interface          VNI        Multicast-group   VNI State
```

```
----------------   --------   ---------------   ---------
nve1               5000       0.0.0.0           Up
```

This example shows how to display the peers of an NVE interface:

```
switch# show nve peers
Interface          Peer-IP           Peer-State
----------------   ---------------   -------------
nve1               111.1.1.1          Up
```

This example shows how to display the VXLAN UDP port configured:

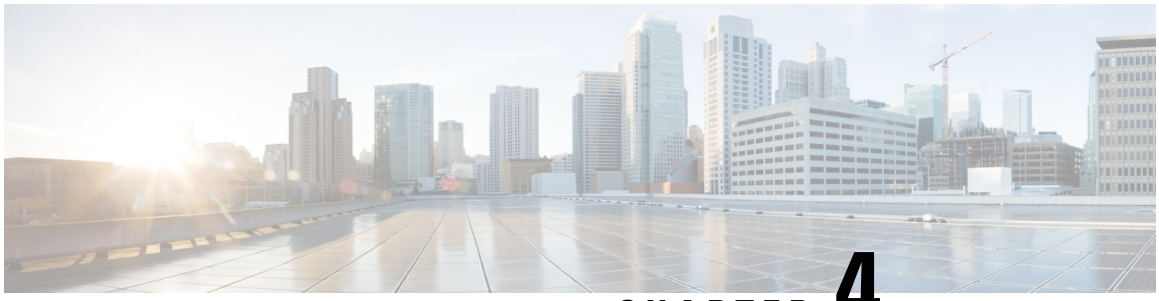
```
switch# show nve vxlan-params
VxLAN Dest. UDP Port: 4789
```

This example shows how to display both VLAN and VXLAN MAC addresses:

```
Added draft comment: hidden contentsswitch# show mac address-table
Legend:
        * - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
        age - seconds since first seen,+ - primary entry using vPC Peer-Link
   VLAN     MAC Address      Type      age     Secure NTFY   Ports/SWID.SSID.LID
---------+-----------------+--------+---------+------+----+------------------
* 109      0000.0410.0902    dynamic   470        F    F  Po2233
* 109      0000.0410.0912    dynamic   470        F    F  Po2233
* 109      0000.0410.0912    dynamic   470        F    F  nve1(1.1.1.200)
* 108      0000.0410.0802    dynamic   470        F    F  Po2233
* 108      0000.0410.0812    dynamic   470        F    F  Po2233
* 107      0000.0410.0702    dynamic   470        F    F  Po2233
* 107      0000.0410.0712    dynamic   470        F    F  Po2233
* 107      0000.0410.0712    dynamic   470        F    F  nve1(1.1.1.200)
* 106      0000.0410.0602    dynamic   470        F    F  Po2233
* 106      0000.0410.0612    dynamic   470        F    F  Po2233
* 105      0000.0410.0502    dynamic   470        F    F  Po2233
* 105      0000.0410.0512    dynamic   470        F    F  Po2233
* 105      0000.0410.0512    dynamic   470        F    F  nve1(1.1.1.200)
* 104      0000.0410.0402    dynamic   470        F    F  Po2233
* 104      0000.0410.0412    dynamic   470        F    F  Po2233
```

This example shows how to clear all MAC address entries in the MAC address table:

```
switch# clear mac address-table dynamic
switch#
```



CHAPTER 4

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:

- SVI and sub-interfaces as core links are not supported along with Layer 2 GW configurations.
- 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.
- The **show** commands with the **internal** keyword are not supported.
- DHCP snooping (Dynamic Host Configuration Protocol snooping) is not supported on VXLAN VLANs.

- SPAN for VXLAN uplink interface is not supported.
- RACLs are not supported on Layer 3 uplinks for VXLAN traffic.
- RACLS and PACLs are not supported for VXLAN VLANs.
- QoS classification is not supported for VXLAN VLANs.
- Uplink ports can be of type Layer 3 interface, sub-interface, or a Layer 3 port-channel interface. However with Layer 2 GW sub-interface uplink ports are not supported.
- 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 does not support co-existence with the MPLS feature.
- VXLAN with Layer 3 VPN is not supported.
- VXLAN with ingress replication is not supported.
- MLD snooping is not supported on VXLAN VLANs.
- DHCP snooping is not supported on VXLAN VLANs.

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.
- 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.
- As a best practice when changing the IP address of a VTEP device, enter the **shut** command on the loopback interface used by the NVE interface and then enter the **no shut** command before changing the IP address.
- Every tenant VRF needs a VRF overlay VLAN and SVI for VXLAN routing.

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 3600 platform 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 3600 platform 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.
  - 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.

BGP EVPN Considerations for VXLAN Deployment

Configuring VXLAN BGP EVPN

Enabling VXLAN

Enable VXLAN and the EVPN.

SUMMARY STEPS

1. **feature vn-segment**
2. **feature nv overlay**
3. **nv overlay evpn**

DETAILED STEPS

	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

SUMMARY STEPS

1. **vlan** *number*
2. **vn-segment** *number*

DETAILED STEPS

	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.

SUMMARY STEPS

1. **vrf context** *vxlan*
2. **vni** *number*
3. **rd auto**
4. **address-family ipv4 unicast**
5. **route-target both auto**
6. **route-target both auto evpn**
7. **address-family ipv6 unicast**
8. **route-target both auto**
9. **route-target both auto evpn**

DETAILED STEPS

	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. 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. Manually configured route targets are required for EBGP.
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.

SUMMARY STEPS

1. **vlan** *number*
2. **interface** *vlan-number*
3. **vrf member** *vxlan-number*
4. **ip address** *address*

DETAILED STEPS

	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

SUMMARY STEPS

1. **vlan** *number*
2. **vn-segment** *number*

DETAILED STEPS

	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.)

SUMMARY STEPS

1. **vrf context** *vxlan*
2. **vni** *number*

DETAILED STEPS

	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

SUMMARY STEPS

1. **fabric forwarding anycast-gateway-mac** *address*
2. **fabric forwarding mode anycast-gateway**

DETAILED STEPS

	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

SUMMARY STEPS

1. **interface** *nve-interface*
2. **host-reachability protocol bgp**
3. **member vni** *vni* **associate-vrf**
4. **member vni** *vni*
5. **mcast-group** *address*

DETAILED STEPS

	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

	Command or Action	Purpose
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	**member vni** *vni*	Add Layer 2 VNIs to the tunnel interface. switch# member vni 900001 associate-vrf
Step 5	**mcast-group** *address*	Configure the mcast group on a per-VNI basis

Configuring BGP on the VTEP

SUMMARY STEPS

1. **router bgp** *number*
2. **router-id** *address*
3. **neighbor** *address* **remote-as** *number*
4. **address-family ipv4 unicast**
5. **address-family l2vpn evpn**
6. (Optional) **Allowas-in**
7. **send-community extended**
8. **vrf** *vrf-name*
9. **address-family ipv4 unicast**
10. **advertise** *l2vpn* **evpn**
11. **address-family ipv6 unicast**
12. **advertise** *l2vpn* **evpn**

DETAILED STEPS

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

	Command or Action	Purpose
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.
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

SUMMARY STEPS

1. **evpn**
2. **vni** *number* **l2**
3. **rd auto**
4. **route-target import auto**
5. **route-target export auto**

DETAILED STEPS

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

SUMMARY STEPS

1. **route-map permitall permit 10**

2. **set ip next-hop unchanged**
3. **router bgp** *autonomous system number*
4. **address-family l2vpn evpn**
5. **retain route-target all**
6. **neighbor** *address* **remote-as** *number*
7. **address-family l2vpn evpn**
8. **disable-peer-as-check**
9. **send-community extended**
10. **route-map permitall out**

DETAILED STEPS

	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.
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]. **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

	Command or Action	Purpose
		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.

Disabling VXLANs

SUMMARY STEPS

1. **configure terminal**
2. **no nv overlay evpn**
3. **no feature vn-segment-vlan-based**
4. **no feature nv overlay**
5. (Optional) **copy running-config startup-config**

DETAILED STEPS

	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 (this means 5 moves in 180 seconds for 5 times) before the switch permanently locks or freezes the duplicate entry.

- 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 (this means 5 moves in 180 seconds for 3 times) before the switch permanently locks or freezes the duplicate entry.

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).

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

Verifying the VXLAN Configuration

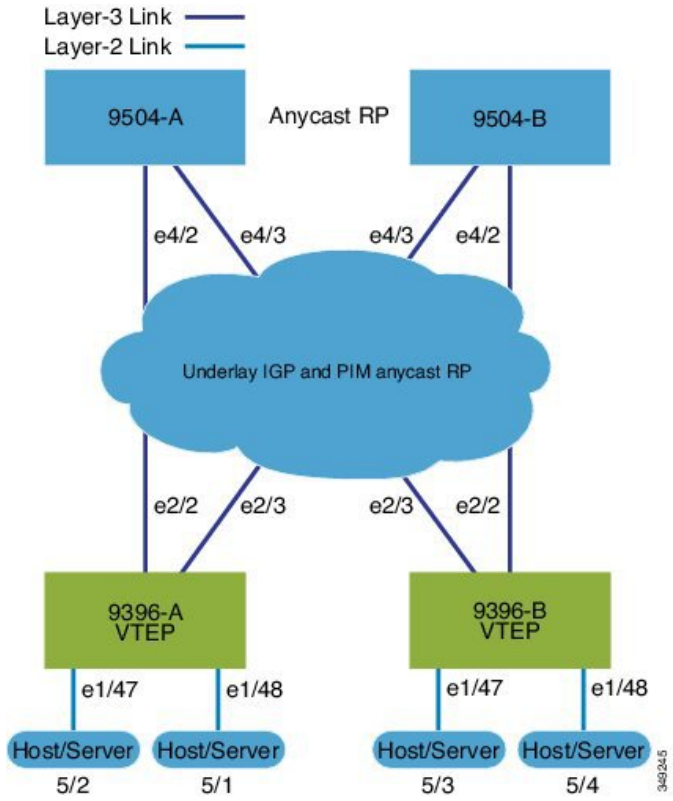
To display the VXLAN configuration information, enter one of the following commands:

Command	Purpose
show tech-support vxlan	Displays related VXLAN tech-support information.
show logging level nve	Displays logging level.
show tech-support nve	Displays related NVE tech-support information.
show tech-support vxlan-evpn	Displays related VXLAN EVPN tech-support information.
show tech-support vxlan platform	Displays VXLAN platform related tech-support information.
show run interface nve	Displays NVE overlay interface configuration.
show nve interface	Displays NVE overlay interface status.
show nve peers	Displays NVE peer status.
show nve peers *peer_IP_address* **interface** *interface_ID* **counters**	Displays per NVE peer statistics.
clear nve peers *peer_IP_address* **interface** *interface_ID* **counters**	Clears per NVE peer statistics.
show nve vni	Displays VXLAN VNI status.
show nve vxlan-params	Displays VXLAN parameters, such as VXLAN destination or UDP port.

Example of VXLAN BGP EVPN (EBGP)

An example of a VXLAN BGP EVPN (EBGP):

Figure 2: 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
```

- Enable OSPF for underlay routing

```
router ospf 1
  log-adjacency-changes detail
```

- 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
  no shutdown

interface Ethernet4/3
  ip address 192.168.3.43/24
  no shutdown
```

- Configure Loopback for local VTEP IP, and BGP

```
interface loopback0
  ip address 20.1.1.1/32
```

- 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 interface-vlan
feature dhcp
```

  - 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
```

  - Configure Loopback for local VTEP IP

```
interface loopback1
  ip address 50.1.1.1/32
```

  - Configure interfaces for Spine-leaf interconnect

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

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

- 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
```

•

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 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 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 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
```

Note The **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
  vni 2001002 l2
```

- 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 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
```

- 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 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
```

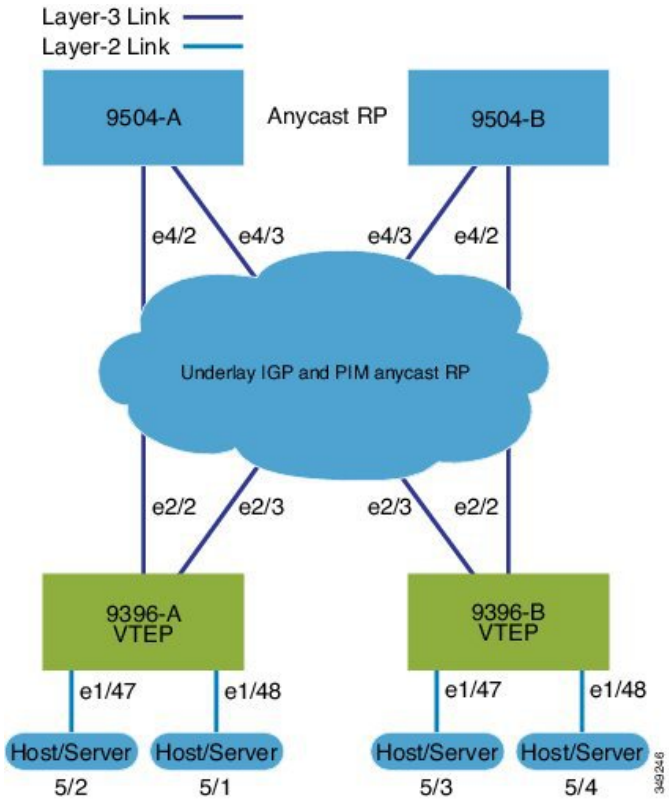
Note The **rd auto** and **route-target** commands are optional unless you want to use them to override the **import** or **export** options.

```
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 3: 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
    ```

  - 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
    ```

  - 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
  no shutdown

interface Ethernet4/3
  ip address 192.168.2.43/24
  ip router ospf 1 area 0.0.0.0
  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 lldp
```

  - 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
  no shutdown

interface Ethernet4/3
  ip address 192.168.3.43/24
  ip router ospf 1 area 0.0.0.0
  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
```

- Configure Loopback for Anycast RP

```
interface loopback1
  ip address 100.1.1.1/32
  ip router ospf 1 area 0.0.0.0
```

- 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 interface-vlan
```

  - 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
```

  - 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
```

```
  no shutdown

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

- 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
```

•

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 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 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 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 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 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
  no shutdown

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

- 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 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**

```
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
40.0.0.1        4    10    8570    8565       60    0    0    5d22h 6
```

- **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 40.0.0.4
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: 40.0.0.2:32868
*>i[2]:[0]:[10001]:[48]:[0000.8816.b645]:[0]:[0.0.0.0]/216
                      40.0.0.2                           100          0 i
*>i[2]:[0]:[10001]:[48]:[0011.0000.0034]:[0]:[0.0.0.0]/216
                      40.0.0.2                           100          0 i
```

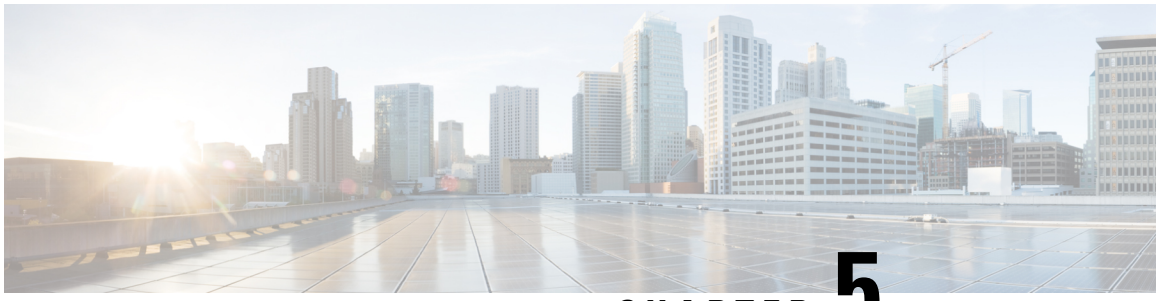
- **show l2route evpn mac all**

```
leaf3# show l2route evpn mac all
Topology    Mac Address    Prod   Next Hop (s)
----------- -------------- ------ ---------------
101         0000.8816.b645 BGP    40.0.0.2
101         0001.0000.0033 Local  Ifindex 4362086
101         0001.0000.0035 Local  Ifindex 4362086
101         0011.0000.0034 BGP    40.0.0.2
```

- **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  5.1.3.2                       40.0.0.2
102         0011.0000.0034 BGP  5.1.3.2                       40.0.0.2
```



CHAPTER 5

Configuring Tenant Routed Multicast

This chapter contains the following sections:

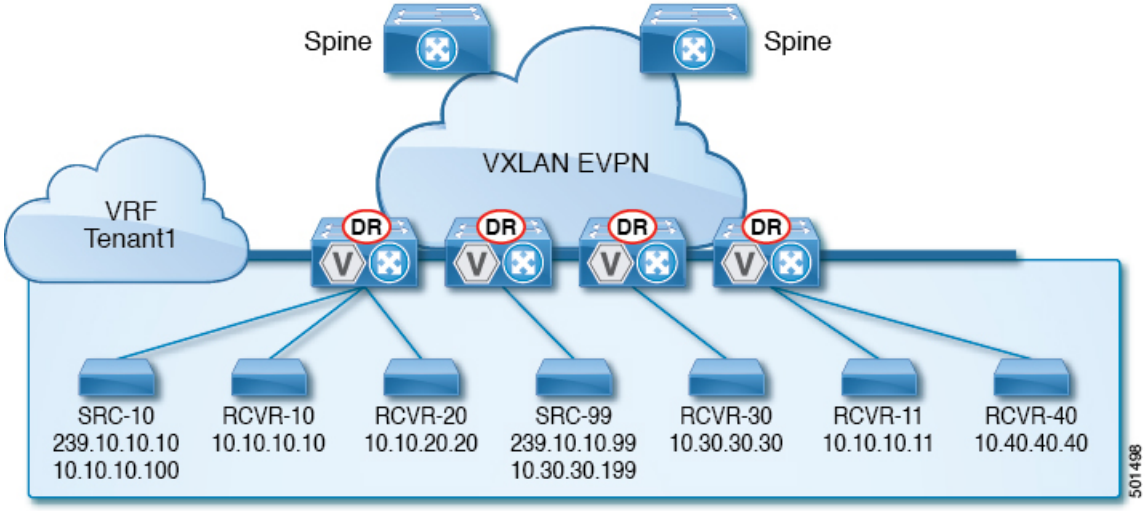
About Tenant Routed Multicast

Tenant Routed Multicast (TRM) enables multicast forwarding on the VXLAN fabric that uses a BGP-based EVPN control plane. TRM provides multi-tenancy aware multicast forwarding between senders and receivers within the same or different subnet local or across VTEPs.

This feature brings the efficiency of multicast delivery to VXLAN overlays. It is based on the standards-based next generation control plane (ngMVPN) described in IETF RFC 6513, 6514. TRM enables the delivery of customer IP multicast traffic in a multitenant fabric, and thus in an efficient and resilient manner. The delivery of TRM improves Layer-3 overlay multicast functionality in our networks.

While BGP EVPN provides the control plane for unicast routing, ngMVPN provides scalable multicast routing functionality. It follows an “always route” approach where every edge device (VTEP) with distributed IP Anycast Gateway for unicast becomes a Designated Router (DR) for Multicast. Bridged multicast forwarding is only present on the edge-devices (VTEP) where IGMP snooping optimizes the multicast forwarding to interested receivers. Every other multicast traffic beyond local delivery is efficiently routed.

Figure 4: VXLAN EVPN TRM



With TRM enabled, multicast forwarding in the underlay is leveraged to replicate VXLAN encapsulated routed multicast traffic. A Default Multicast Distribution Tree (Default-MDT) is built per-VRF. This is an addition to the existing multicast groups for Layer-2 VNI Broadcast, Unknown Unicast, and Layer-2 multicast replication group. The individual multicast group addresses in the overlay are mapped to the respective underlay multicast address for replication and transport. The advantage of using a BGP-based approach allows the VXLAN BGP EVPN fabric with TRM to operate as fully distributed Overlay Rendezvous-Point (RP), with the RP presence on every edge-device (VTEP).

A multicast-enabled data center fabric is typically part of an overall multicast network. Multicast sources, receivers, and multicast rendezvous points, might reside inside the data center but might also be inside the campus or externally reachable via the WAN. TRM allows a seamless integration with existing multicast networks. It can leverage multicast rendezvous points external to the fabric. Furthermore, TRM allows for tenant-aware external connectivity using Layer-3 physical interfaces or subinterfaces.

Guidelines and Limitations for Tenant Routed Multicast

Tenant Routed Multicast (TRM) has the following guidelines and limitations:

- FEX is not supported on Cisco Nexus 3600 platform switches.
- The Guidelines and Limitations for VXLANs, on page 8 also apply to TRM.
- With TRM enabled, SVI as a core link is not supported.
- TRM supports IPv4 multicast only.
- TRM requires an IPv4 multicast-based underlay using PIM Any Source Multicast (ASM) which is also known as sparse mode.
- TRM supports overlay PIM ASM and PIM SSM only. PIM BiDir is not supported in the overlay.
- RP has to be configured either internal or external to the fabric.
- The internal RP must be configured on all TRM-enabled VTEPs including the border nodes.

- The external RP must be external to the border nodes.
- The RP must be configured within the VRF pointing to the external RP IP address (static RP). This ensures that unicast and multicast routing is enabled to reach the external RP in the given VRF.
- TRM supports multiple border nodes. Reachability to an external RP via multiple border leaf switches is supported (ECMP).
- Both PIM and **ip igmp snooping vxlan** must be enabled on the L3 VNI's VLAN in a VXLAN vPC setup.

Guidelines and Limitations for Layer 3 Tenant Routed Multicast

Layer 3 Tenant Routed Multicast (TRM) has the following configuration guidelines and limitations:

- Beginning with Cisco NX-OS Release 9.3(3), Cisco Nexus 3600 platform switches support TRM in Layer 3 mode. This feature is supported on IPv4 overlays only. Layer 2 mode and L2/L3 mixed mode are not supported.

  The Cisco Nexus 3600 platform switches can function as a BL for L3 unicast traffic. For Anycast functionality, the RP can be internal, external, or RP everywhere.
- Beginning with Cisco NX-OS Release 9.3(3), Cisco Nexus 3600 platform switches support TRM with vPC border leafs. The **advertise-pip** and **advertise virtual-rmac** commands must be enabled on the border leafs to support this functionality. For more information, see the "Configuring VIP/PIP" section.
- Well-known local scope multicast (224.0.0.0/24) is excluded from TRM and is bridged.
- When an interface NVE is brought down on the border leaf, the internal overlay RP per VRF must be brought down.
- If one or both VTEPs are a Cisco Nexus 3600 platform switch, the packet TTL is decremented twice, once for routing to the L3 VNI on the source leaf and once for forwarding from the destination L3 VNI to the destination VLAN on the destination leaf.
- Cisco Nexus 3600 platform switches do not support TRM Multi-Site.

Rendezvous Point for Tenant Routed Multicast

With TRM enabled Internal and External RP is supported. The following table displays the first release in which RP positioning is or is not supported.

	RP Internal	**RP External**	**PIM-Based RP Everywhere**
TRM L3 Mode	9.3(3)	9.3(3)	9.3(3)

	RP Internal	**RP External**
TRM L2 Mode	N/A	N/A
TRM L3 Mode	7.0(3)I7(1)	7.0(3)I7(4)

	RP Internal	RP External
TRM L2L3 Mode	7.0(3)I7(1)	N/A

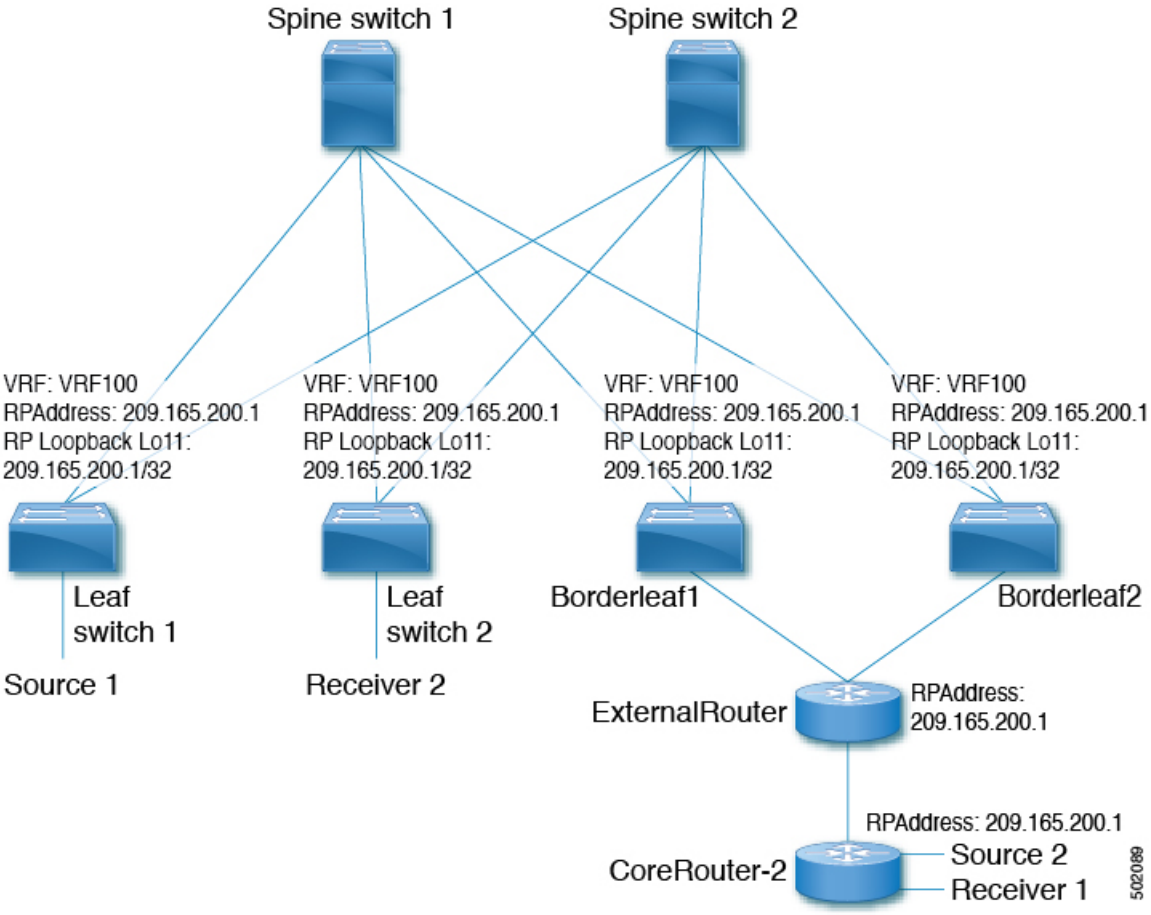
Configuring a Rendezvous Point for Tenant Routed Multicast

For Tenant Routed Multicast, there are two rendezvous point options:

- Configuring a Rendezvous Point Inside the VXLAN Fabric, on page 58
- Configuring an External Rendezvous Point, on page 60

Configuring a Rendezvous Point Inside the VXLAN Fabric

Configure the loopback for the TRM VRFs with the following commands on all devices (VTEP). Ensure it is reachable within EVPN (advertise/redistribute).



SUMMARY STEPS

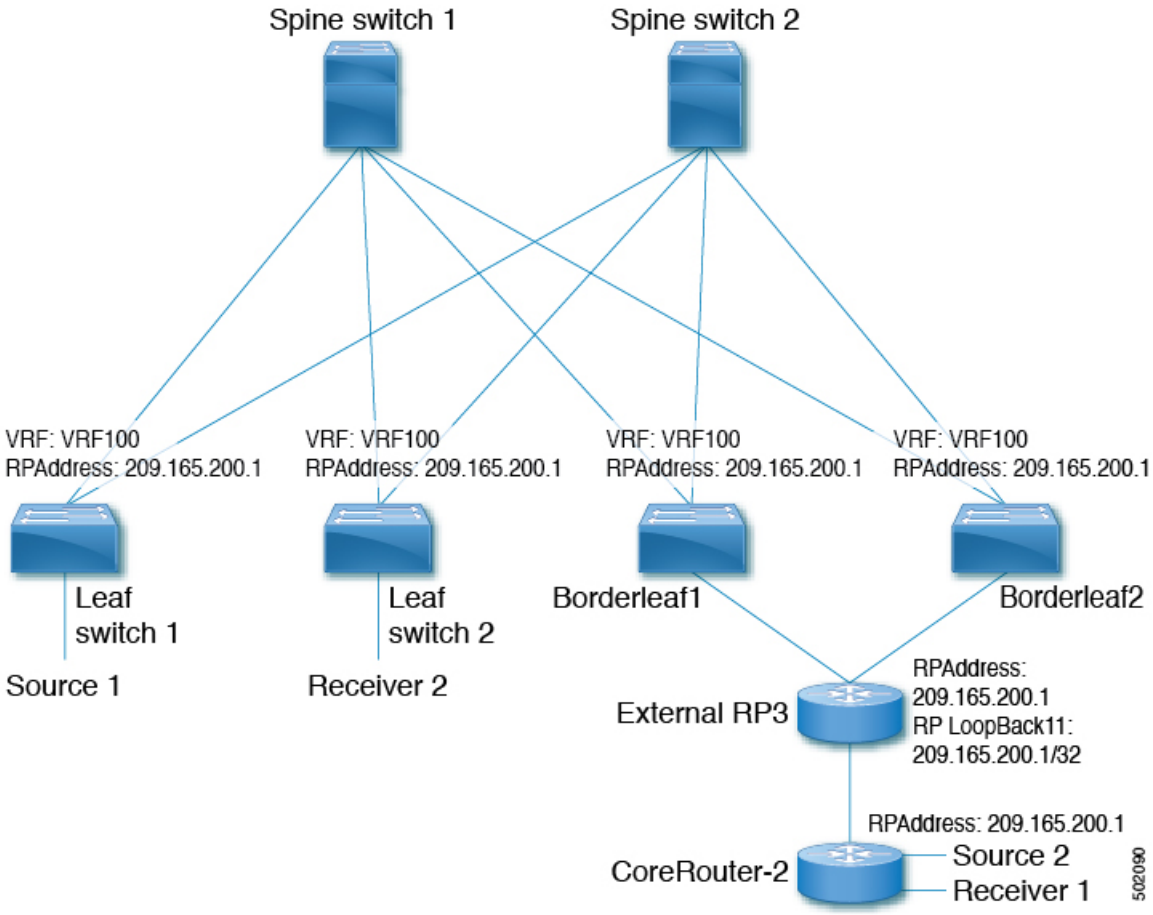
1. **configure terminal**
2. **interface loopback** *loopback_number*
3. **vrf member** *vxlan-number*
4. **ip address** *ip-address*
5. **ip pim sparse-mode**
6. **vrf context** *vrf-name*
7. **ip pim rp-address** *ip-address-of-router* **group-list** *group-range-prefix*

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `switch# configure terminal`	Enters global configuration mode.
Step 2	**interface loopback** *loopback_number* **Example:** `switch(config)# interface loopback 11`	Configure the loopback interface on all TRM-enabled nodes. This enables the rendezvous point inside the fabric.
Step 3	**vrf member** *vxlan-number* **Example:** `switch(config-if)# vrf member vrf100`	Configure VRF name.
Step 4	**ip address** *ip-address* **Example:** `switch(config-if)# ip address 209.165.200.1/32`	Specify IP address.
Step 5	**ip pim sparse-mode** **Example:** `switch(config-if)# ip pim sparse-mode`	Configure sparse-mode PIM on an interface.
Step 6	**vrf context** *vrf-name* **Example:** `switch(config-if)# vrf context vrf100`	Create a VXLAN tenant VRF.
Step 7	**ip pim rp-address** *ip-address-of-router* **group-list** *group-range-prefix* **Example:** `switch(config-vrf# ip pim rp-address 209.165.200.1 group-list 224.0.0.0/4`	The value of the *ip-address-of-router* parameter is that of the RP. The same IP address must be on all the edge devices (VTEPs) for a fully distributed RP.

Configuring an External Rendezvous Point

Configure the external rendezvous point (RP) IP address within the TRM VRFs on all devices (VTEP). In addition, ensure reachability of the external RP within the VRF via the border node. With TRM enabled and an external RP in use, ensure that only one routing path is active. Routing between the TRM fabric and the external RP must be via a single border leaf (non ECMP).



SUMMARY STEPS

1. **configure terminal**
2. **vrf context vrf100**
3. **ip pim rp-address** *ip-address-of-router* **group-list** *group-range-prefix*

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `switch# configure terminal`	Enter configuration mode.

	Command or Action	Purpose
Step 2	**vrf context vrf100** **Example:** `switch(config)# vrf context vrf100`	Enter configuration mode.
Step 3	**ip pim rp-address** *ip-address-of-router* **group-list** *group-range-prefix* **Example:** `switch(config-vrf)# ip pim rp-address 209.165.200.1 group-list 224.0.0.0/4`	The value of the *ip-address-of-router* parameter is that of the RP. The same IP address must be on all of the edge devices (VTEPs) for a fully distributed RP.

Configuring Layer 3 Tenant Routed Multicast

This procedure enables the Tenant Routed Multicast (TRM) feature. TRM operates primarily in the Layer 3 forwarding mode for IP multicast by using BGP MVPN signaling. TRM in Layer 3 mode is the main feature and the only requirement for TRM enabled VXLAN BGP EVPN fabrics. If non-TRM capable edge devices (VTEPs) are present, the Layer 2/Layer 3 mode and Layer 2 mode have to be considered for interop.

To forward multicast between senders and receivers on the Layer 3 cloud and the VXLAN fabric on TRM vPC border leafs, the VIP/PIP configuration must be enabled. For more information, see Configuring VIP/PIP.

Note TRM follows an always-route approach and hence decrements the Time to Live (TTL) of the transported IP multicast traffic.

Before you begin

VXLAN EVPN **feature nv overlay** and **nv overlay evpn** must be configured.

The rendezvous point (RP) must be configured.

Procedure

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `switch# configure terminal`	Enter configuration mode.
Step 2	**feature ngmvpn** **Example:** `switch(config)# feature ngmvpn`	Enables the Next-Generation Multicast VPN (ngMVPN) control plane. New address family commands become available in BGP.
Step 3	**ip igmp snooping vxlan** **Example:** `switch(config)# ip igmp snooping vxlan`	Configure IGMP snooping for VXLAN VLANs.

	Command or Action	Purpose
Step 4	**interface nve1** **Example:** `switch(config)# interface nve 1`	Configure the NVE interface.
Step 5	**member vni** *vni-range* **associate-vrf** **Example:** `switch(config-if-nve)# member vni 200100 associate-vrf`	Configure the Layer 3 virtual network identifier. The range of *vni-range* is from 1 to 16,777,214.
Step 6	**mcast-group** *ip-prefix* **Example:** `switch(config-if-nve-vni)# mcast-group 225.3.3.3`	Builds the default multicast distribution tree for the VRF VNI (Layer 3 VNI). The multicast group is used in the underlay (core) for all multicast routing within the associated Layer 3 VNI (VRF). **Note** We recommend that underlay multicast groups for Layer 2 VNI, default MDT, and data MDT not be shared. Use separate, non-overlapping groups.
Step 7	**exit** **Example:** `switch(config-if-nve-vni)# exit`	Exits command mode.
Step 8	**exit** **Example:** `switch(config-if)# exit`	Exits command mode.
Step 9	**router bgp 100** **Example:** `switch(config)# router bgp 100`	Set autonomous system number.
Step 10	**exit** **Example:** `switch(config-router)# exit`	Exits command mode.
Step 11	**neighbor** *ip-addr* **Example:** `switch(config-router)# neighbor 1.1.1.1`	Configure IP address of the neighbor.
Step 12	**address-family ipv4 mvpn** **Example:** `switch(config-router-neighbor)# address-family ipv4 mvpn`	Configure multicast VPN.

	Command or Action	Purpose
Step 13	**send-community extended** **Example:** `switch(config-router-neighbor-af)# send-community extended`	Enables ngMVPN for address family signalization. The **send community extended** command ensures that extended communities are exchanged for this address family.
Step 14	**exit** **Example:** `switch(config-router-neighbor-af)# exit`	Exits command mode.
Step 15	**exit** **Example:** `switch(config-router)# exit`	Exits command mode.
Step 16	**vrf context** *vrf_name* **Example:** `switch(config-router)#vrf context vrf100`	Configure VRF name.
Step 17	**ip pim rp-address** *ip-address-of-router* **group-list** *group-range-prefix* **Example:** `switch(config-vrf)# ip pim rp-address 209.165.201.1 group-list 226.0.0.0/8`	The value of the *ip-address-of-router* parameter is that of the RP. The same IP address must be on all of the edge devices (VTEPs) for a fully distributed RP. For overlay RP placement options, see the Configuring a Rendezvous Point for Tenant Routed Multicast, on page 58 section.
Step 18	**address-family ipv4 unicast** **Example:** `switch(config-vrf)# address-family ipv4 unicast`	Configure unicast address family.
Step 19	**route-target both auto mvpn** **Example:** `switch(config-vrf-af-ipv4)# route-target both auto mvpn`	Defines the BGP route target that is added as an extended community attribute to the customer multicast (C_Multicast) routes (ngMVPN route type 6 and 7). Auto route targets are constructed by the 2-byte Autonomous System Number (ASN) and Layer 3 VNI.
Step 20	**ip multicast overlay-spt-only** **Example:** `switch(config)# ip multicast overlay-spt-only`	Gratuitously originate (S,A) route when the source is locally connected. The **ip multicast overlay-spt-only** command is enabled by default on all MVPN-enabled switches (typically leaf node).
Step 21	**interface***vlan_id* **Example:** `switch(config)# interface vlan11`	Configures the first-hop gateway (distributed anycast gateway for the Layer 2 VNI. No router PIM peering must ever happen with this interface.
Step 22	**no shutdown** **Example:** `switch(config-if)# no shutdown`	Disables an interface.

	Command or Action	Purpose
Step 23	**vrf member** *vrf-num* **Example:** `switch(config-if)# vrf member vrf100`	Configure VRF name.
Step 24	**ip address** *ip_address* **Example:** `switch(config-if)# ip address 11.1.1.1/24`	Configure IP address.
Step 25	**ip pim sparse-mode** **Example:** `switch(config-if)# ip pim sparse-mode`	Enables IGMP and PIM on the SVI. This is required is multicast sources and/or receivers exist in this VLAN.
Step 26	**fabric forwarding mode anycast-gateway** **Example:** `switch(config-if)# fabric forwarding mode anycast-gateway`	Configure Anycast Gateway Forwarding Mode.
Step 27	**ip pim neighbor-policy NONE*** **Example:** `switch(config-if)# ip pim neighbor-policy NONE*`	Creates an IP PIM neighbor policy to avoid PIM neighborship with PIM routers within the VLAN. The **none** keyword is a configured route map to deny any ipv4 addresses to avoid establishing PIM neighborship policy using anycase IP. **Note** Do not use Distributed Anycast Gateway for PIM Peerings.
Step 28	**exit** **Example:** `switch(config-if)# exit`	Exits command mode.
Step 29	**interface** *vlan_id* **Example:** `switch(config)# interface vlan100`	Configure Layer 3 VNI.
Step 30	**no shutdown** **Example:** `switch(config-if)# no shutdown`	Disable an interface.
Step 31	**vrf member vrf100** **Example:** `switch(config-if)# vrf member vrf100`	Configure VRF name.
Step 32	**ip forward** **Example:** `switch(config-if)# ip forward`	Enable IP forwarding on interface.

	Command or Action	Purpose
Step 33	**ip pim sparse-mode** **Example:** `switch(config-if)# ip pim sparse-mode`	Configure sparse-mode PIM on interface. There is no PIM peering happening in the Layer-3 VNI, but this command must be present for forwarding.

Configuring TRM on the VXLAN EVPN Spine

This procedure enables Tenant Routed Multicast (TRM) on a VXLAN EVPN spine switch.

Before you begin

The VXLAN BGP EVPN spine must be configured. See Configuring BGP for EVPN on the Spine, on page 27.

SUMMARY STEPS

1. **configure terminal**
2. **route-map permitall permit 10**
3. **set ip next-hop unchanged**
4. **exit**
5. **router bgp [autonomous system]** *number*
6. **address-family ipv4 mvpn**
7. **retain route-target all**
8. **neighbor** *ip-address* [**remote-as** *number*]
9. **address-family ipv4 mvpn**
10. **disable-peer-as-check**
11. **rewrite-rt-asn**
12. **send-community extended**
13. **route-reflector-client**
14. **route-map permitall out**

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `switch# configure terminal`	Enter configuration mode.
Step 2	**route-map permitall permit 10** **Example:** `switch(config)# route-map permitall permit 10`	Configure the route-map. **Note** The route-map keeps the next-hop unchanged for EVPN routes • Required for eBGP • Options for iBGP

	Command or Action	Purpose
Step 3	**set ip next-hop unchanged** **Example:** `switch(config-route-map)# set ip next-hop unchanged`	Set next hop address. **Note** The route-map keeps the next-hop unchanged for EVPN routes • Required for eBGP • Options for iBGP
Step 4	**exit** **Example:** `switch(config-route-map)# exit`	Return to exec mode.
Step 5	**router bgp [autonomous system]** *number* **Example:** `switch(config)# router bgp 65002`	Specify BGP.
Step 6	**address-family ipv4 mvpn** **Example:** `switch(config-router)# address-family ipv4 mvpn`	Configure the address family IPv4 MVPN under the BGP.
Step 7	**retain route-target all** **Example:** `switch(config-router-af)# retain route-target all`	Configure retain route-target all under address-family IPv4 MVPN [global]. **Note** Required for eBGP. Allows the spine to retain and advertise all MVPN routes when there are no local VNIs configured with matching import route targets.
Step 8	**neighbor** *ip-address* [**remote-as** *number*] **Example:** `switch(config-router-af)# neighbor 100.100.100.1`	Define neighbor.
Step 9	**address-family ipv4 mvpn** **Example:** `switch(config-router-neighbor)# address-family ipv4 mvpn`	Configure address family IPv4 MVPN under the BGP neighbor.
Step 10	**disable-peer-as-check** **Example:** `switch(config-router-neighbor-af)# 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 11	**rewrite-rt-asn** **Example:**	Normalizes the outgoing route target's AS number to match the remote AS number. Uses the BGP configured neighbors remote AS. The **rewrite-rt-asn** command is required if

	Command or Action	Purpose
	`switch(config-router-neighbor-af)# rewrite-rt-asn`	the route target auto feature is being used to configure EVPN route targets.
Step 12	**send-community extended** **Example:** `switch(config-router-neighbor-af)# send-community extended`	Configures community for BGP neighbors.
Step 13	**route-reflector-client** **Example:** `switch(config-router-neighbor-af)# route-reflector-client`	Configure route reflector. **Note** Required for iBGP with route-reflector.
Step 14	**route-map permitall out** **Example:** `switch(config-router-neighbor-af)# route-map permitall out`	Applies route-map to keep the next-hop unchanged. **Note** Required for eBGP.

Configuring TRM with vPC Support

SUMMARY STEPS

1. **configure terminal**
2. **feature vpc**
3. **feature interface-vlan**
4. **feature lacp**
5. **feature pim**
6. **feature ospf**
7. **ip pim rp-address** *address* **group-list** *range*
8. **vpc domain** *domain-id*
9. **hardware access-list tcam region mac-ifacl**
10. **hardware access-list tcam region vxlan 10**
11. **reload**
12. **peer switch**
13. **peer gateway**
14. **peer-keepalive destination** *ipaddress*
15. **ip arp synchronize**
16. **ipv6 nd synchronize**
17. Create vPC peer-link.
18. **system nve infra-vlans** *range*
19. **vlan** *number*
20. Create the SVI.
21. (Optional) **delay restore interface-vlan** *seconds*

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `switch# configure terminal`	Enter global configuration mode.
Step 2	**feature vpc** **Example:** `switch(config)# feature vpc`	Enables vPCs on the device.
Step 3	**feature interface-vlan** **Example:** `switch(config)# feature interface-vlan`	Enables the interface VLAN feature on the device.
Step 4	**feature lacp** **Example:** `switch(config)# feature lacp`	Enables the LACP feature on the device.
Step 5	**feature pim** **Example:** `switch(config)# feature pim`	Enables the PIM feature on the device.
Step 6	**feature ospf** **Example:** `switch(config)# feature ospf`	Enables the OSPF feature on the device.
Step 7	**ip pim rp-address** *address* **group-list** *range* **Example:** `switch(config)# ip pim rp-address 100.100.100.1 group-list 224.0.0/4`	Defines a PIM RP address for the underlay multicast group range.
Step 8	**vpc domain** *domain-id* **Example:** `switch(config)# vpc domain 1`	Creates a vPC domain on the device and enters vpn-domain configuration mode for configuration purposes. There is no default. The range is 1–1000.
Step 9	**hardware access-list tcam region mac-ifacl** **Example:** `switch(config)# hardware access-list tcam region mac-ifacl 0`	Carves the TCAM region for the ACL database.
Step 10	**hardware access-list tcam region vxlan 10** **Example:** `switch(config)# hardware access-list tcam region vxlan 10`	Assigns the the TCAM region for use by a VXLAN.

	Command or Action	Purpose
Step 11	**reload** **Example:** `switch(config)# reload`	Reloads the switch config for the TCAM assignments to become active.
Step 12	**peer switch** **Example:** `switch(config-vpc-domain)# peer switch`	Defines the peer switch.
Step 13	**peer gateway** **Example:** `switch(config-vpc-domain)# peer gateway`	To enable Layer 3 forwarding for packets that are destined to the gateway MAC address of the virtual port channel (vPC), use the **peer-gateway** command.
Step 14	**peer-keepalive destination** *ipaddress* **Example:** `switch(config-vpc-domain)# peer-keepalive destination 172.28.230.85`	Configures the IPv4 address for the remote end of the vPC peer-keepalive link. **Note** The system does not form the vPC peer link until you configure a vPC peer-keepalive link. The management ports and VRF are the defaults. **Note** We recommend that you configure a separate VRF and use a Layer 3 port from each vPC peer device in that VRF for the vPC peer-keepalive link. For more information about creating and configuring VRFs, see the Cisco Nexus 3600 NX-OS Series Unicast Routing Configuration Guide, Release 9.3(x).
Step 15	**ip arp synchronize** **Example:** `switch(config-vpc-domain)# ip arp synchronize`	Enables IP ARP synchronize under the vPC Domain to facilitate faster ARP table population following device reload.
Step 16	**ipv6 nd synchronize** **Example:** `switch(config-vpc-domain)# ipv6 nd synchronize`	Enables IPv6 and synchronization under the vPC domain to facilitate faster and table population following device reload.

Step 17	Create vPC peer-link.
	Example:
	`switch(config)# interface port-channel 1`
	`switch(config)# switchport`
	`switch(config)# switchport mode trunk`
	`switch(config)# switchport trunk allowed vlan 1,10,100-200`
	`switch(config)# mtu 9216`
	`switch(config)# vpc peer-link`
	`switch(config)# no shut`
		Creates the vPC peer-link port-channel interface and adds two member interfaces to it.

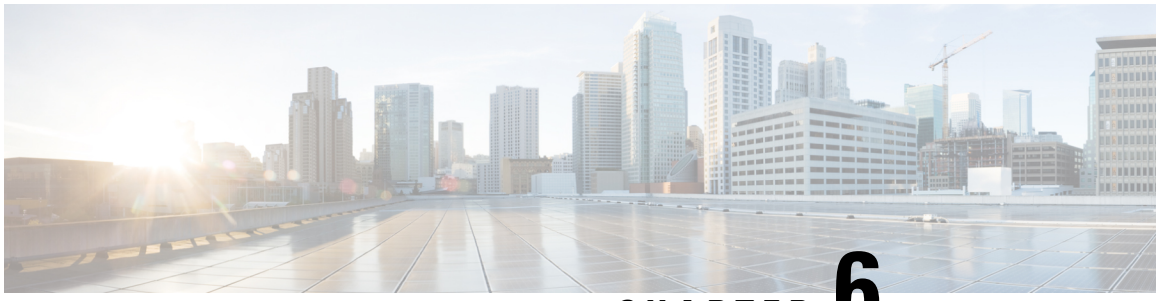
	Command or Action	Purpose

	`switch(config)# interface Ethernet 1/1, 1/21`
	`switch(config)# switchport`
	`switch(config)# mtu 9216`
	`switch(config)# channel-group 1 mode active`
	`switch(config)# no shutdown`	

Step 18	**system nve infra-vlans** *range* **Example:** `switch(config)# system nve infra-vlans 10`	Defines a non-VXLAN enabled VLAN as a backup routed path.
Step 19	**vlan** *number* **Example:** `switch(config)# vlan 10`	Creates the VLAN to be used as an infra-VLAN.

Step 20	Create the SVI.
	Example:
	`switch(config)# interface vlan 10`
	`switch(config)# ip address 10.10.10.1/30`
	`switch(config)# ip router ospf process UNDERLAY area 0`
	`switch(config)# ip pim sparse-mode`
	`switch(config)# no ip redirects`
	`switch(config)# mtu 9216`
	`switch(config)# no shutdown`
		Creates the SVI used for the backup routed path over the vPC peer-link.

Step 21	(Optional) **delay restore interface-vlan** *seconds* **Example:** `switch(config-vpc-domain)# delay restore interface-vlan 45`	Enables the delay restore timer for SVIs. We recommend tuning this value when the SVI/VNI scale is high. For example, when the SCI count is 1000, we recommend that you set the delay restore for **interface-vlan** to 45 seconds.



CHAPTER 6

Configuring External VRF Connectivity and Route Leaking

This chapter contains the following sections:

Configuring External VRF Connectivity

About External Layer-3 Connectivity for VXLAN BGP EVPN Fabrics

A VXLAN BGP EVPN fabric can be extended by using per-VRF IP routing to achieve external connectivity. The approach that is used for the Layer-3 extensions is commonly referred to as VRF Lite, while the functionality itself is more accurately defined as Inter-AS Option A or back-to-back VRF connectivity.

Guidelines and Limitations for External VRF Connectivity and Route Leaking

The following are the guidelines and limitations for External Layer-3 Connectivity for VXLAN BGP EVPN Fabrics:

- Support added for Cisco Nexus 3600 platform switches.
- A physical Layer-3 Interface (Parent-Interface) can be used for external Layer-3 connectivity (ie VRF default).
- The Parent-Interface to multiple Sub-Interfaces can not be used for external Layer-3 connectivity (ie Ethernet1/1 for VRF default). A Sub-Interface can be used instead.
- VTEPs do not support VXLAN encapsulated traffic over Parent-Interfaces if Sub-Interfaces are configured. This is regardless of VRF participation.
- VTEPs do not support VXLAN encapsulated traffic over Sub-Interfaces. This is regardless of VRF participation or IEEE 802.1q encapsulation.
- Mixing Sub-Interfaces for VXLAN and non-VXLAN enabled VLANs is not supported.

Configuring Route Leaking

About Centralized VRF Route-Leaking for VXLAN BGP EVPN Fabrics

VXLAN BGP EVPN uses MP-BGP and its route-policy concept to import and export prefixes. The ability of this very extensive route-policy model allows to leak routes from one VRF to another VRF and vice-versa; any combination of custom VRF or VRF default can be used. VRF route-leaking is a switch-local function at specific to a location in the network, the location where the cross-VRF route-target import/export configuration takes place (leaking point). The forwarding between the different VRFs follows the control-plane, the location of where the configuration for the route-leaking is performed - hence Centralized VRF route-leaking. With the addition of VXLAN BGP EVPN, the leaking point requires to advertise the cross-VRF imported/exported route and advertise them towards the remote VTEPs or External Routers.

The advantage of Centralized VRF route-leaking is that only the VTEP acting as leaking point requires the special capabilities needed, while all other VTEPs in the network are neutral to this function.

Guidelines and Limitations for External VRF Connectivity and Route Leaking

The following are the guidelines and limitations for External Layer-3 Connectivity for VXLAN BGP EVPN Fabrics:

- Support added for Cisco Nexus 3600 platform switches.
- A physical Layer-3 Interface (Parent-Interface) can be used for external Layer-3 connectivity (ie VRF default).
- The Parent-Interface to multiple Sub-Interfaces can not be used for external Layer-3 connectivity (ie Ethernet1/1 for VRF default). A Sub-Interface can be used instead.
- VTEPs do not support VXLAN encapsulated traffic over Parent-Interfaces if Sub-Interfaces are configured. This is regardless of VRF participation.
- VTEPs do not support VXLAN encapsulated traffic over Sub-Interfaces. This is regardless of VRF participation or IEEE 802.1q encapsulation.
- Mixing Sub-Interfaces for VXLAN and non-VXLAN enabled VLANs is not supported.

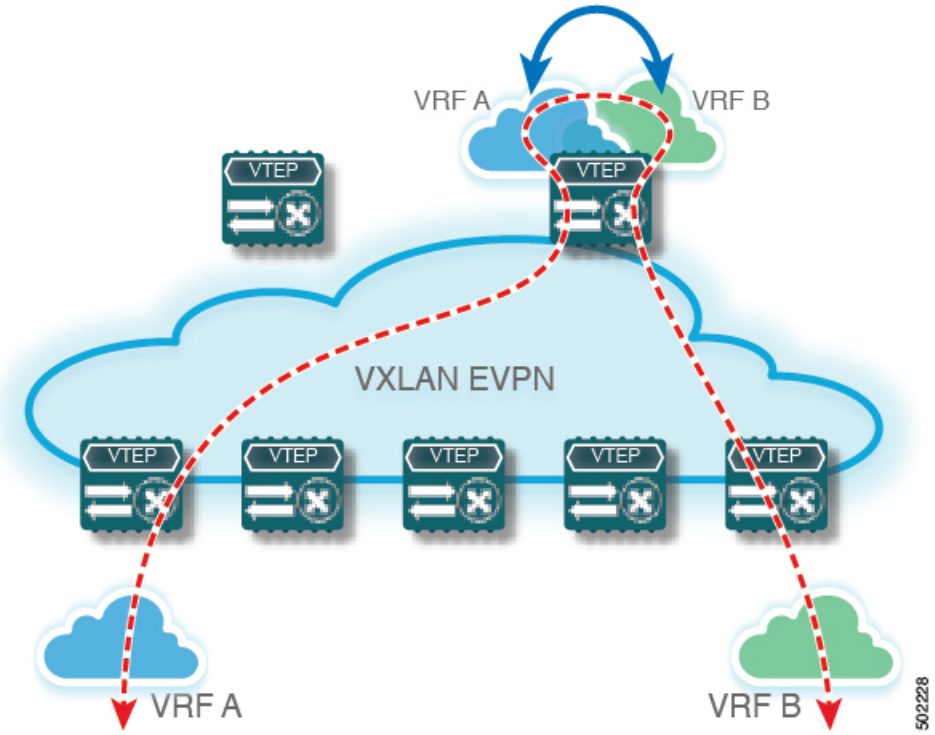
Centralized VRF Route-Leaking Brief - Shared Internet with Custom VRF

Some pointers follow:

- The Shared Internet with VRF route-leaking for VXLAN BGP EVPN fabrics is depicted in the following figure.
- The default-route is made exported from the Shared Internet VRF and re-advertisement within VRF Blue and VRF Red on the Border Node.
- Ensure the default-route in VRF Blue and VRF Red is not leaked to the Shared Internet VRF.
- The less specific prefixes for VRF Blue and VRF Red are exported for the Shared Internet VRF and re-advertised as necessary.

- Configured less specific prefixes (aggregates) that are advertised from the Border Node to the remaining VTEPs to the destination VRF (Blue or Red).
- BGP EVPN does not export prefixes that were previously imported to prevent the occurrence of routing loops.

Figure 5: Centralized VRF Route-Leaking - Shared Internet with Custom VRF



Configuring Centralized VRF Route-leaking - Specific Prefixes between Custom VRF

Configuring VRF Context on the Routing-Block VTEP

This procedure applies equally to IPv6.

SUMMARY STEPS

1. **configure terminal**
2. **vrf context** *vrf-name*
3. **vni** *number*
4. **rd auto**
5. **address-family ipv4 unicast**
6. **route-target both {auto** | *as:vni*}
7. **route-target both {auto** | *as:vni* }**evpn**
8. **route-target import** *rt-from-different-vrf*

9. **route-target import** *rt-from-different-vrf* **evpn**

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal**	Enter global configuration mode.
Step 2	**vrf context** *vrf-name*	Configure the VRF.
Step 3	**vni** *number*	Specify the VNI. The VNI associated with the VRF is often referred to as Layer-3 VNI, L3VNI or L3VPN. The L3VNI is configured as common identifier across the participating VTEPs.
Step 4	**rd auto**	Specify the VRFs Route Distinguisher (RD). The RD uniquely identifies a VTEP within a L3VNI.
Step 5	**address-family ipv4 unicast**	Configure the IPv4 Unicast address-family. Required for IPv4 over VXLAN with IPv4 underlay.
Step 6	**route-target both {auto** | *as:vni*}	Configure the Route Target (RT) for import/export of IPv4 prefixes within the IPv4 unicast address-family The Route Target (RT) is used for a per-VRF prefix import/export policy. If *as:vni* is entered, the value is in the format of ASN:NN, ASN4:NN, or IPv4:NN.
Step 7	**route-target both {auto** | *as:vni* }**evpn**	Configure the Route Target (RT) for import/export of IPv4 prefixes within the IPv4 unicast address-family The Route Target (RT) is used for a per-VRF prefix import/export policy. If *as:vni* is entered, the value is in the format of ASN:NN, ASN4:NN, or IPv4:NN.
Step 8	**route-target import** *rt-from-different-vrf*	Configure the Route Target (RT) for importing IPv4 prefixes from the leaked-from VRF (ie AS:VNI).
Step 9	**route-target import** *rt-from-different-vrf* **evpn**	Configure the Route Target (RT) for importing IPv4 prefixes from the leaked-from VRF (ie AS:VNI).

Configuring the BGP VRF instance on the Routing-Block

This procedure applies equally to IPv6.

SUMMARY STEPS

1. **configure terminal**
2. **router bgp** *autonomous-system number*
3. **vrf** *vrf-name*
4. **address-family ipv4 unicast**
5. **advertise l2vpn evpn**
6. **aggregate-address** *prefix/mask*

7. **maximum-paths ibgp** *number*
8. **maximum-paths** *number*

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal**	Enters global configuration mode.
Step 2	**router bgp** *autonomous-system number*	Configure BGP.
Step 3	**vrf** *vrf-name*	Specify the VRF.
Step 4	**address-family ipv4 unicast**	Configure address family for IPv4
Step 5	**advertise l2vpn evpn**	Enable the advertisement of EVPN routes within IPv4 address-family.
Step 6	**aggregate-address** *prefix/mask*	Create less specific prefix aggregate into the destination VRF.
Step 7	**maximum-paths ibgp** *number*	Enabling equal cost multipathing (ECMP) for iBGP prefixes.
Step 8	**maximum-paths** *number*	Enabling equal cost multipathing (ECMP) for eBGP prefixes

Example - Configuration Centralized VRF Route-Leaking - Specific Prefixes Between Custom VRF

Configuring VXLAN BGP EVPN Routing-Block

The VXLAN BGP EVPN Routing-Block acts as centralized route-leaking point. The leaking configuration is localized such that control-plane leaking and data-path forwarding follow the same path. Most significantly is the VRF configuration of the Routing-Block and the advertisement of the less specific prefixes (aggregates) into the respective destination VRFs.

```
vrf context Blue
  vni 51010
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn
    route-target import 65002:51020
    route-target import 65002:51020 evpn
!
vlan 2110
  vn-segment 51010
!
interface Vlan2110
  no shutdown
  mtu 9216
  vrf member Blue
  no ip redirects
  ip forward
!
vrf context Red
  vni 51020
  rd auto
```

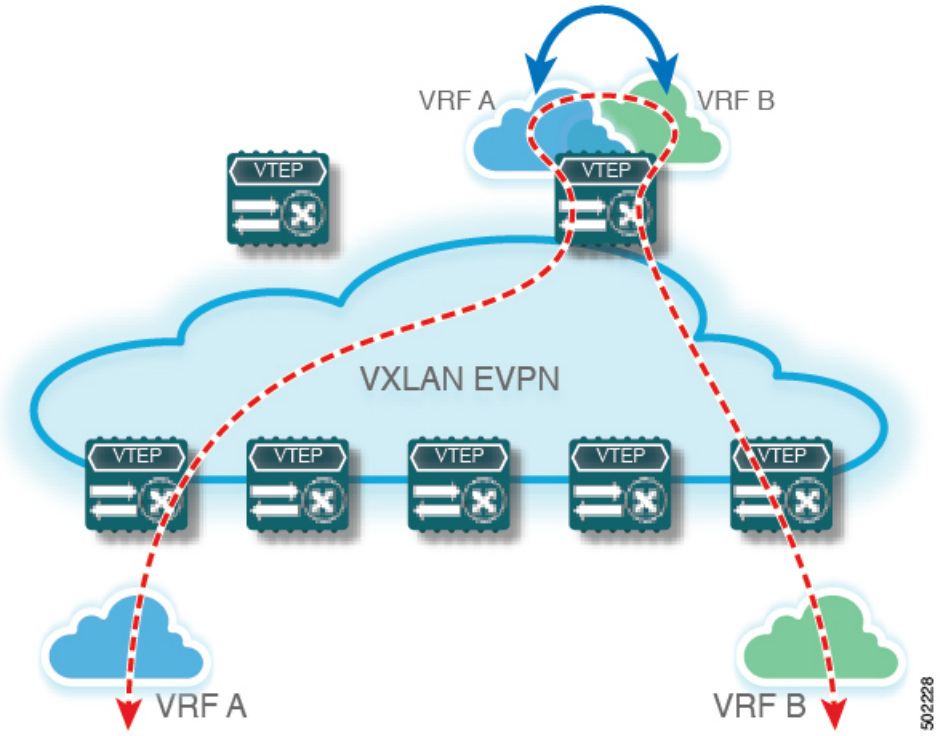
```
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn
    route-target import 65002:51010
    route-target import 65002:51010 evpn
!
vlan 2120
  vn-segment 51020
!
interface Vlan2120
  no shutdown
  mtu 9216
  vrf member Blue
  no ip redirects
  ip forward
!
interface nve1
  no shutdown
  host-reachability protocol bgp
  source-interface loopback1
  member vni 51010 associate-vrf
  member vni 51020 associate-vrf
!
router bgp 65002
  vrf Blue
    address-family ipv4 unicast
      advertise l2vpn evpn
      aggregate-address 10.20.0.0/16
      maximum-paths ibgp 2
      Maximum-paths 2
  vrf Red
    address-family ipv4 unicast
      advertise l2vpn evpn
      aggregate-address 10.10.0.0/16
      maximum-paths ibgp 2
      Maximum-paths 2
```

Centralized VRF Route-Leaking Brief - Shared Internet with Custom VRF

Some pointers follow:

- The Shared Internet with VRF route-leaking for VXLAN BGP EVPN fabrics is depicted in the following figure.
- The default-route is made exported from the Shared Internet VRF and re-advertisement within VRF Blue and VRF Red on the Border Node.
- Ensure the default-route in VRF Blue and VRF Red is not leaked to the Shared Internet VRF.
- The less specific prefixes for VRF Blue and VRF Red are exported for the Shared Internet VRF and re-advertised as necessary.
- Configured less specific prefixes (aggregates) that are advertised from the Border Node to the remaining VTEPs to the destination VRF (Blue or Red).
- BGP EVPN does not export prefixes that were previously imported to prevent the occurrence of routing loops.

Figure 6: Centralized VRF Route-Leaking - Shared Internet with Custom VRF



Configuring Centralized VRF Route-Leaking - Shared Internet with Custom VRF

Configuring Internet VRF on Border Node

This procedure applies equally to IPv6.

SUMMARY STEPS

1. **configure terminal**
2. **vrf context** *vrf-name*
3. **vni** *number*
4. **ip route 0.0.0.0/0** *next-hop*
5. **rd auto**
6. **address-family ipv4 unicast**
7. **route-target both {auto** | *as:vni*}
8. **route-target both** *shared-vrf-rt* **evpn**

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal**	Enter global configuration mode.
Step 2	**vrf context** *vrf-name*	Configure the VRF.

	Command or Action	Purpose
Step 3	**vni** *number*	Specify the VNI. The VNI associated with the VRF is often referred to as Layer-3 VNI, L3VNI or L3VPN. The L3VNI is configured as common identifier across the participating VTEPs.
Step 4	**ip route 0.0.0.0/0** *next-hop*	Configure default-route in shared internet VRF to external router (example).
Step 5	**rd auto**	Specify the VRFs Route Distinguisher (RD). The RD uniquely identifies a VTEP within a L3VNI.
Step 6	**address-family ipv4 unicast**	Configure the IPv4 Unicast address-family. Required for IPv4 over VXLAN with IPv4 underlay.
Step 7	**route-target both {auto** | *as:vni*}	Configure the Route Target (RT) for import/export of EVPN and IPv4 prefixes within the IPv4 unicast address-family.
Step 8	**route-target both** *shared-vrf-rt* **evpn**	Configure a special Route Target (RT) for the import/export of the shared IPv4 prefixes. Additional import/export map for further qualification is supported

Configuring Shared Internet BGP Instance on the Border Node

This procedure applies equally to IPv6.

SUMMARY STEPS

1. **configure terminal**
2. **router bgp** *autonomous-system number*
3. **vrf** *vrf-name*
4. **address-family ipv4 unicast**
5. **advertise l2vpn evpn**
6. **aggregate-address** *prefix/mask*
7. **maximum-paths ibgp** *number*
8. **maximum-paths** *number*

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal**	Enters global configuration mode.
Step 2	**router bgp** *autonomous-system number*	Configure BGP.
Step 3	**vrf** *vrf-name*	Specify the VRF.
Step 4	**address-family ipv4 unicast**	Configure address family for IPv4

	Command or Action	Purpose
Step 5	**advertise l2vpn evpn**	Enable the advertisement of EVPN routes within IPv4 address-family.
Step 6	**aggregate-address** *prefix/mask*	Create less specific prefix aggregate into the destination VRF.
Step 7	**maximum-paths ibgp** *number*	Enabling equal cost multipathing (ECMP) for iBGP prefixes.
Step 8	**maximum-paths** *number*	Enabling equal cost multipathing (ECMP) for eBGP prefixes.

Configuring Custom VRF Context on the Border Node - 1

This procedure applies equally to IPv6.

SUMMARY STEPS

1. **configure terminal**
2. **vrf context** *vrf-name*
3. **vni** *number*
4. **rd auto**
5. **ip route 0.0.0.0/0 Null0**
6. **address-family ipv4 unicast**
7. **route-target both {auto** | *as:vni***}**
8. **route-target both {auto** | *as:vni***} evpn**
9. **import map** *name*

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal**	Enters global configuration mode.
Step 2	**vrf context** *vrf-name*	Configure the VRF.
Step 3	**vni** *number*	Specify the VNI. The VNI associated with the VRF is often referred to as Layer-3 VNI, L3VNI or L3VPN. The L3VNI is configured as the common identifier across the participating VTEPs.
Step 4	**rd auto**	Specify the VRFs Route Distinguisher (RD). The Route Distinguisher (RD) uniquely identifies a VTEP within a L3VNI.
Step 5	**ip route 0.0.0.0/0 Null0**	Configure default-route in common VRF to attract traffic towards Border Node with Shared Internet VRF.
Step 6	**address-family ipv4 unicast**	Configure the IPv4 Unicast address-family.

	Command or Action	Purpose
		Required for IPv4 over VXLAN with IPv4 underlay.
Step 7	**route-target both {auto** | *as:vni***}**	Configure the Route Target (RT) for import/export of IPv4 prefixes within the IPv4 unicast address-family The Route Target (RT) is used for a per-VRF prefix import/export policy. If *as:vni* is entered, the value is in the format of ASN:NN, ASN4:NN, or IPv4:NN.
Step 8	**route-target both {auto** | *as:vni***} evpn**	Configure the Route Target (RT) for import/export of IPv4 prefixes within the IPv4 unicast address-family The Route Target (RT) is used for a per-VRF prefix import/export policy. If *as:vni* is entered, the value is in the format of ASN:NN, ASN4:NN, or IPv4:NN.
Step 9	**import map** *name*	Apply a route-map on routes being imported into this routing table.

Configuring Custom VRF Instance in BGP on the Border Node

This procedure applies equally to IPv6.

SUMMARY STEPS

1. **configure terminal**
2. **router bgp** *autonomous-system-number*
3. **vrf** *vrf-name*
4. **address-family ipv4 unicast**
5. **advertise l2vpn evpn**
6. **network 0.0.0.0/0**
7. **maximum-paths ibgp** *number*
8. **maximum-paths** *number*

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal**	Enters global configuration mode.
Step 2	**router bgp** *autonomous-system-number*	Configure BGP.
Step 3	**vrf** *vrf-name*	Specify the VRF.
Step 4	**address-family ipv4 unicast**	Configure address family for IPv4.
Step 5	**advertise l2vpn evpn**	Enable the advertisement of EVPN routes within IPv4 address-family.
Step 6	**network 0.0.0.0/0**	Creating IPv4 default-route network statement.
Step 7	**maximum-paths ibgp** *number*	Enabling equal cost multipathing (ECMP) for iBGP prefixes.

	Command or Action	Purpose
Step 8	**maximum-paths** *number*	Enabling equal cost multipathing (ECMP) for eBGP prefixes.

Example - Configuration Centralized VRF Route-Leaking - Shared Internet with Custom VRF

An example of Centralized VRF route-leaking with Shared Internet VRF

Configuring VXLAN BGP EVPN Border Node for Shared Internet VRF

The VXLAN BGP EVPN Border Node provides a centralized Shared Internet VRF. The leaking configuration is localized such that control-plane leaking and data-path forwarding following the same path. Most significantly is the VRF configuration of the Border Node and the advertisement of the default-route and less specific prefixes (aggregates) into the respective destination VRFs.

```
vrf context Shared
  vni 51099
  ip route 0.0.0.0/0 10.9.9.1
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn
    route-target both 99:99
    route-target both 99:99 evpn
!
vlan 2199
  vn-segment 51099
!
interface Vlan2199
  no shutdown
  mtu 9216
  vrf member Shared
  no ip redirects
  ip forward
!
ip prefix-list PL_DENY_EXPORT seq 5 permit 0.0.0.0/0
!
route-map RM_DENY_IMPORT deny 10
 match ip address prefix-list PL_DENY_EXPORT
route-map RM_DENY_IMPORT permit 20
!
vrf context Blue
  vni 51010
  ip route 0.0.0.0/0 Null0
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn
    route-target both 99:99
    route-target both 99:99 evpn
    import map RM_DENY_IMPORT
!
vlan 2110
  vn-segment 51010
!
interface Vlan2110
  no shutdown
  mtu 9216
  vrf member Blue
  no ip redirects
```

```
  ip forward
!
vrf context Red
  vni 51020
  ip route 0.0.0.0/0 Null0
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn
    route-target both 99:99
    route-target both 99:99 evpn
    import map RM_DENY_IMPORT
!
vlan 2120
  vn-segment 51020
!
interface Vlan2120
  no shutdown
  mtu 9216
  vrf member Blue
  no ip redirects
  ip forward
!
interface nve1
  no shutdown
  host-reachability protocol bgp
  source-interface loopback1
  member vni 51099 associate-vrf
  member vni 51010 associate-vrf
  member vni 51020 associate-vrf
!
router bgp 65002
  vrf Shared
    address-family ipv4 unicast
      advertise l2vpn evpn
      aggregate-address 10.10.0.0/16
      aggregate-address 10.20.0.0/16
      maximum-paths ibgp 2
      maximum-paths 2
  vrf Blue
    address-family ipv4 unicast
      advertise l2vpn evpn
      network 0.0.0.0/0
      maximum-paths ibgp 2
      maximum-paths 2
  vrf Red
    address-family ipv4 unicast
      advertise l2vpn evpn
      network 0.0.0.0/0
      maximum-paths ibgp 2
      maximum-paths 2
```

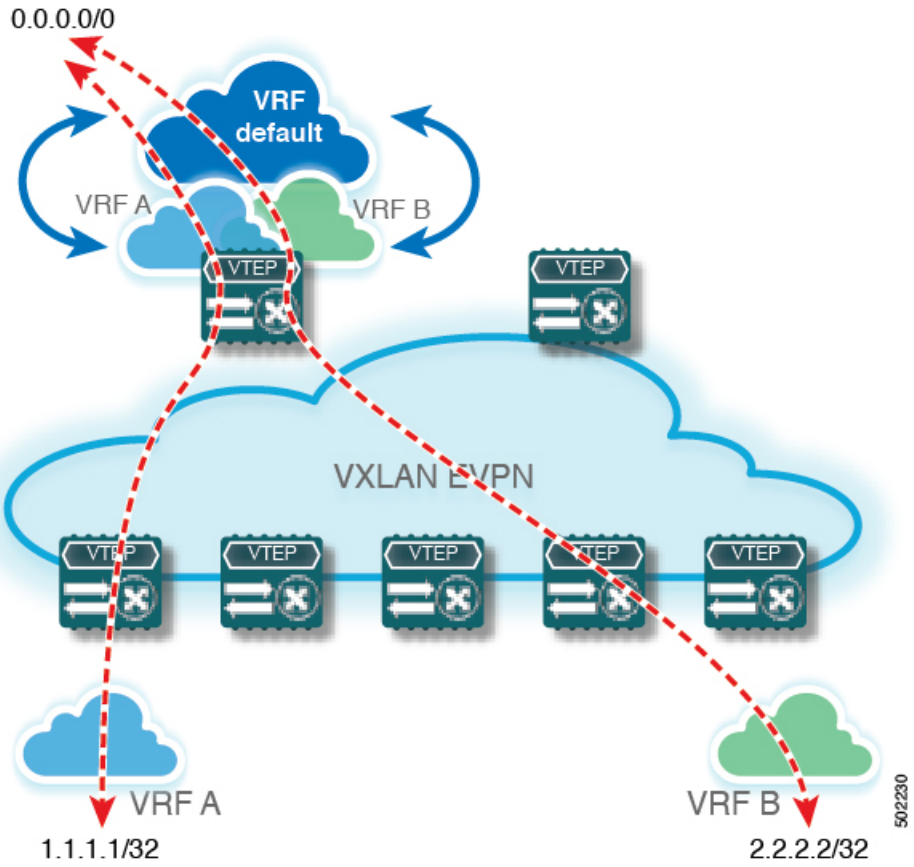
Centralized VRF Route-Leaking Brief - Shared Internet with VRF Default

Some pointers are given below:

- The Shared Internet with VRF route-leaking for VXLAN BGP EVPN fabrics is depicted within Figure 4.
- The default-route is made exported from VRF default and re-advertisement within VRF Blue and VRF Red on the Border Node.
- Ensure the default-route in VRF Blue and VRF Red is not leaked to the Shared Internet VRF

- The less specific prefixes for VRF Blue and VRF Red are exported to VRF default and re-advertised as necessary.
- Configured less specific prefixes (aggregates) that are advertised from the Border Node to the remaining VTEPs to the destination VRF (Blue or Red).
- BGP EVPN does not export prefixes that were previously imported to prevent the occurrence of routing loops.

Figure 7: Centralized VRF Route-Leaking - Shared Internet with VRF Default



Configuring Centralized VRF Route-Leaking - Shared Internet with VRF Default

Configuring VRF Default on Border Node

This procedure applies equally to IPv6.

SUMMARY STEPS

1. **configure terminal**
2. **ip route 0.0.0.0/0** *next-hop*

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal**	Enters global configuration mode.
Step 2	**ip route 0.0.0.0/0** *next-hop*	Configure default-route in VRF default to external router (example)

Configuring BGP Instance for VRF Default on the Border Node

This procedure applies equally to IPv6.

SUMMARY STEPS

1. **configure terminal**
2. **router bgp** *autonomous-system number*
3. **address-family ipv4 unicast**
4. **aggregate-address** *prefix/mask*
5. **maximum-paths** *number*

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal**	Enters global configuration mode.
Step 2	**router bgp** *autonomous-system number*	Configure BGP.
Step 3	**address-family ipv4 unicast**	Configure address family for IPv4.
Step 4	**aggregate-address** *prefix/mask*	Create less specific prefix aggregate in VRF default.
Step 5	**maximum-paths** *number*	Enabling equal cost multipathing (ECMP) for eBGP prefixes.

Configuring Custom VRF on Border Node

This procedure applies equally to IPv6

SUMMARY STEPS

1. **configure terminal**
2. **ip prefix-list** *name* **seq 5 permit 0.0.0.0/0**
3. **route-map** *name* **deny 10**
4. **match ip address prefix-list** *name*
5. **route-map** *name* **permit 20**

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal**	Enters global configuration mode.
Step 2	**ip prefix-list** *name* **seq 5 permit 0.0.0.0/0**	Configure IPv4 prefix-list for default-route filtering.
Step 3	**route-map** *name* **deny 10**	Create route-map with leading deny statement to prevent the default-route of being leaked.
Step 4	**match ip address prefix-list** *name*	Match against the IPv4 prefix-list that contains the default-route.
Step 5	**route-map** *name* **permit 20**	Create route-map with trailing allow statement to advertise non-matching routes via route-leaking.

Configuring Filter for Permitted Prefixes from VRF Default on the Border Node

This procedure applies equally to IPv6.

SUMMARY STEPS

1. **configure terminal**
2. **route-map** *name* **permit 10**

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal**	Enters global configuration mode.
Step 2	**route-map** *name* **permit 10**	Create route-map with allow statement to advertise routes via route-leaking to the customer VRF and subsequently remote VTEPs.

Configuring Custom VRF Context on the Border Node - 2

This procedure applies equally to IPv6.

SUMMARY STEPS

1. **configure terminal**
2. **vrf context** *vrf-name*
3. **vni** *number*
4. **rd auto**
5. **ip route 0.0.0.0/0 Null0**
6. **address-family ipv4 unicast**
7. **route-target both auto** | *AS:VNI*
8. **route-target both auto** | *AS:VNI* **evpn**
9. **route-target both** *shared-vrf-rt*
10. **route-target both** *shared-vrf-rt* **evpn**

11. **import vrf default map** *name*

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal**	Enters global configuration mode.
Step 2	**vrf context** *vrf-name*	Configure the VRF.
Step 3	**vni** *number*	Specify the VNI. The VNI associated with the VRF is often referred to as Layer-3 VNI, L3VNI or L3VPN. The L3VNI is configured as common identifier across the participating VTEPs.
Step 4	**rd auto**	Specify the VRFs Route Distinguisher (RD). The Route Distinguisher (RD) uniquely identifies a VTEP within a L3VNI.
Step 5	**ip route 0.0.0.0/0 Null0**	Configure default-route in common VRF to attract traffic towards Border Node with Shared Internet VRF.
Step 6	**address-family ipv4 unicast**	Configure the IPv4 Unicast address-family. Required for IPv4 over VXLAN with IPv4 underlay.
Step 7	**route-target both auto** | *AS:VNI*	Configure the Route Target (RT) for import/export of EVPN and IPv4 prefixes within the IPv4 unicast address-family.
Step 8	**route-target both auto** | *AS:VNI* **evpn**	Configure the Route Target (RT) for import/export of EVPN and IPv4 prefixes within the IPv4 unicast address-family.
Step 9	**route-target both** *shared-vrf-rt*	Configure a special Route Target (RT) for the import/export of the Shared IPv4 prefixes. Additional import/export map for further qualification is supported
Step 10	**route-target both** *shared-vrf-rt* **evpn**	Configure a special Route Target (RT) for the import/export of the Shared IPv4 prefixes. Additional import/export map for further qualification is supported
Step 11	**import vrf default map** *name*	Permits all routes, from VRF default, from being imported into the custom VRF according to the specific route-map.

Configuring Custom VRF Instance in BGP on the Border Node

This procedure applies equally to IPv6.

SUMMARY STEPS

1. **configure terminal**
2. **router bgp** *autonomous-system-number*
3. **vrf** *vrf-name*
4. **address-family ipv4 unicast**
5. **advertise l2vpn evpn**
6. **network 0.0.0.0/0**
7. **maximum-paths ibgp** *number*
8. **maximum-paths** *number*

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal**	Enters global configuration mode.
Step 2	**router bgp** *autonomous-system-number*	Configure BGP.
Step 3	**vrf** *vrf-name*	Specify the VRF.
Step 4	**address-family ipv4 unicast**	Configure address family for IPv4.
Step 5	**advertise l2vpn evpn**	Enable the advertisement of EVPN routes within IPv4 address-family.
Step 6	**network 0.0.0.0/0**	Creating IPv4 default-route network statement.
Step 7	**maximum-paths ibgp** *number*	Enabling equal cost multipathing (ECMP) for iBGP prefixes.
Step 8	**maximum-paths** *number*	Enabling equal cost multipathing (ECMP) for eBGP prefixes.

Example - Configuration Centralized VRF Route-Leaking - VRF Default with Custom VRF

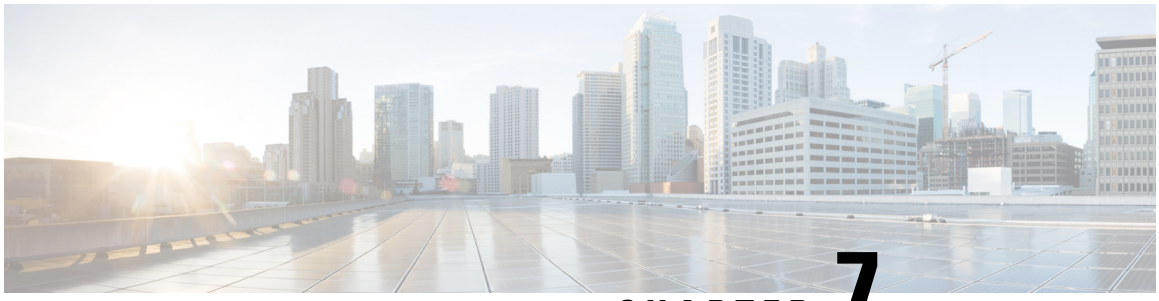
An example of Centralized VRF route-leaking with VRF default

Configuring VXLAN BGP EVPN Border Node for VRF Default

The VXLAN BGP EVPN Border Node provides centralized access to VRF default. The leaking configuration is localized such that control-plane leaking and data-path forwarding following the same path. Most significantly is the VRF configuration of the Border Node and the advertisement of the default-route and less specific prefixes (aggregates) into the respective destination VRFs.

```
ip route 0.0.0.0/0 10.9.9.1
!
ip prefix-list PL_DENY_EXPORT seq 5 permit 0.0.0.0/0
!
route-map permit 10
match ip address prefix-list PL_DENY_EXPORT
route-map RM_DENY_EXPORT permit 20
route-map RM_PERMIT_IMPORT permit 10
!
vrf context Blue
```

```
  vni 51010
  ip route 0.0.0.0/0 Null0
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn
    import vrf default map RM_PERMIT_IMPORT
    export vrf default 100 map RM_DENY_EXPORT allow-vpn
!
vlan 2110
  vn-segment 51010
!
interface Vlan2110
  no shutdown
  mtu 9216
  vrf member Blue
  no ip redirects
  ip forward
!
vrf context Red
  vni 51020
  ip route 0.0.0.0/0 Null0
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn
    import vrf default map RM_PERMIT_IMPORT
    export vrf default 100 map RM_DENY_EXPORT allow-vpn
!
vlan 2120
  vn-segment 51020
!
interface Vlan2120
  no shutdown
  mtu 9216
  vrf member Blue
  no ip redirects
  ip forward
!
interface nve1
  no shutdown
  host-reachability protocol bgp
  source-interface loopback1
  member vni 51010 associate-vrf
  member vni 51020 associate-vrf
!
router bgp 65002
  address-family ipv4 unicast
      aggregate-address 10.10.0.0/16
      aggregate-address 10.20.0.0/16
      maximum-paths 2
      maximum-paths ibgp 2
  vrf Blue
    address-family ipv4 unicast
      advertise l2vpn evpn
      network 0.0.0.0/0
      maximum-paths ibgp 2
      maximum-paths 2
  vrf Red
    address-family ipv4 unicast
      advertise l2vpn evpn
      network 0.0.0.0/0
      maximum-paths ibgp 2
      maximum-paths 2
```



CHAPTER 7

Configuring Seamless Integration of EVPN with L3VPN (MPLS LDP)

This chapter contains the following sections:

Information About Configuring Seamless Integration of EVPN with L3VPN (MPLS LDP)

Data center deployments have adopted VXLAN EVPN for its benefits like EVPN control-plane learning, multitenancy, seamless mobility, redundancy, and easier POD additions. Similarly, the Core is either an LDP-based MPLS L3VPN network or transitioning from traditional an MPLS L3VPN LDP-based underlay to a more sophisticated solution like segment routing (SR). Segment routing is adopted for its benefits like unified IGP and MPLS control planes, simpler traffic engineering methods, easier configuration, and SDN adoption.

With two different technologies, one within the data center and one in the Core, it is natural to handoff from VXLAN to an MPLS-based core at the DCI nodes. These nodes which sit on the edge of the DC domain, interfacing with the Core edge router.

Guidelines and Limitations for Configuring Seamless Integration of EVPN with L3VPN (MPLS LDP)

The following are the guidelines and limitations for Configuring Seamless Integration of EVPN with L3VPN (MPLS LDP):

The following features are supported:

- Layer 3 orphans
- MPLS extended ECMP (enabled by default)

- Beginning with Cisco NX-OS Release 10.3(3)F, Type-6 encryption for MPLS LDP user password is supported on Cisco NX-OS switches.

The following features are not supported:

- Subnet stretches across the DC domain
- vPC
- SVI/Subinterfaces

Configuring Seamless Integration of EVPN with L3VPN (MPLS LDP)

These configuration steps are required on a DCI switch to import and re-originate the routes from a VXLAN domain to an MPLS domain and back to a VXLAN domain.

SUMMARY STEPS

1. **configure terminal**
2. **feature mpls l3vpn**
3. **feature mpls ldp**
4. **nv overlay evpn**
5. **router bgp** *number*
6. **address-family ipv4 unicast**
7. **redistribute direct route-map** *route-map-name*
8. **exit**
9. **address-family l2vpn evpn**
10. **exit**
11. **neighbor** *address* **remote-as** *number*
12. **update-source** *type/id*
13. **ebgp-multihop** *ttl-value*
14. **address-family ipv4 unicast**
15. **send-community extended**
16. **exit**
17. **address-family vpnv4 unicast**
18. **send-community extended**
19. **import l2vpn evpn reoriginate**
20. **neighbor** *address* **remote-as** *number*
21. **address-family ipv4 unicast**
22. **send-community extended**
23. **address-family ipv6 unicast**
24. **send-community extended**
25. **address-family l2vpn evpn**
26. **send-community extended**
27. **import vpn unicast reoriginate**

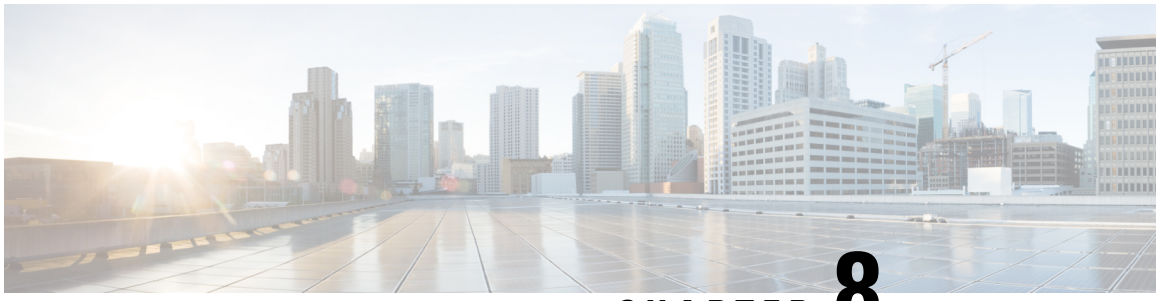
DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `switch# configure terminal`	Enters global configuration mode.
Step 2	**feature mpls l3vpn** **Example:** `switch# feature mpls l3vpn`	Enables the MPLS Layer 3 VPN feature.
Step 3	**feature mpls ldp** **Example:** `switch# feature mpls ldp`	Enables the MPLS Label Distribution Protocol (LDP).
Step 4	**nv overlay evpn** **Example:** `switch(config)# nv overlay evpn`	Enables the EVPN control plane for VXLAN.
Step 5	**router bgp** *number* **Example:** `switch(config)# router bgp 100`	Configures BGP. The value of the *number* argument is from 1 to 4294967295.
Step 6	**address-family ipv4 unicast** **Example:** `switch(config-router)# address-family ipv4 unicast`	Configures the address family for IPv4.
Step 7	**redistribute direct route-map** *route-map-name* **Example:** `switch(config-router-af)# redistribute direct route-map passall`	Configures the directly connected route map.
Step 8	**exit** **Example:** `switch(config-router-af)# exit`	Exits command mode.
Step 9	**address-family l2vpn evpn** **Example:** `switch(config-router)# address-family l2vpn evpn`	Configures the L2VPN address family.
Step 10	**exit** **Example:** `switch(config-router-af)# exit`	Exits command mode.
Step 11	**neighbor** *address* **remote-as** *number* **Example:**	Configures a BGP neighbor. The range of the *number* argument is from 1 to 65535.

	Command or Action	Purpose
	`switch(config-router)# neighbor 108.108.108.108 remote-as 22`	
Step 12	**update-source** *type/id* **Example:** `switch(config-router-neighbor)# update-source loopback100`	Specifies the source of the BGP session and updates.
Step 13	**ebgp-multihop** *ttl-value* **Example:** `switch(config-router-neighbor)# ebgp-multihop 10`	Specifies the multihop TTL for the remote peer. The range of *ttl-value* is from 2 to 255.
Step 14	**address-family ipv4 unicast** **Example:** `switch(config-router-neighbor)# address-family ipv4 unicast`	Configures the unicast sub-address family.
Step 15	**send-community extended** **Example:** `switch(config-router-neighbor-af)# send-community extended`	Configures the community attribute for this neighbor.
Step 16	**exit** **Example:** `switch(config-router-neighbor-af)# exit`	Exits command mode.
Step 17	**address-family vpnv4 unicast** **Example:** `switch(config-router-neighbor)# address-family vpnv4 unicast`	Configures the address family for IPv4.
Step 18	**send-community extended** **Example:** `switch(config-router)# send-community extended`	Sends the extended community attribute.
Step 19	**import l2vpn evpn reoriginate** **Example:** `switch(config-router)# import l2vpn evpn reoriginate`	Reoriginates the route with a new RT.
Step 20	**neighbor** *address* **remote-as** *number* **Example:** `switch(config-router)# neighbor 175.175.175.2 remote-as 1`	Defines the neighbor.
Step 21	**address-family ipv4 unicast** **Example:**	Configures the address family for IPv4.

	Command or Action	Purpose
	`switch(config-router)# address-family ipv4 unicast`	
Step 22	**send-community extended** **Example:** `switch(config-router)# send-community extended`	Configures the community for BGP neighbors.
Step 23	**address-family ipv6 unicast** **Example:** `switch(config-router)# address-family ipv6 unicast`	Configures the IPv6 unicast address family, which is required for IPv6 over VXLAN with an IPv4 underlay.
Step 24	**send-community extended** **Example:** `switch(config-router)# send-community extended`	Configures the community for BGP neighbors.
Step 25	**address-family l2vpn evpn** **Example:** `switch(config-router)# address-family l2vpn evpn`	Configures the L2VPN address family.
Step 26	**send-community extended** **Example:** `switch(config-router)# send-community extended`	Configures the community for BGP neighbors.
Step 27	**import vpn unicast reoriginate** **Example:** `switch(config-router)# import vpn unicast reoriginate`	Reoriginates the route with a new RT.

Configuring Seamless Integration of EVPN with L3VPN (MPLS LDP)



CHAPTER 8

Configuring Seamless Integration of EVPN with L3VPN (MPLS SR)

This chapter contains the following sections:

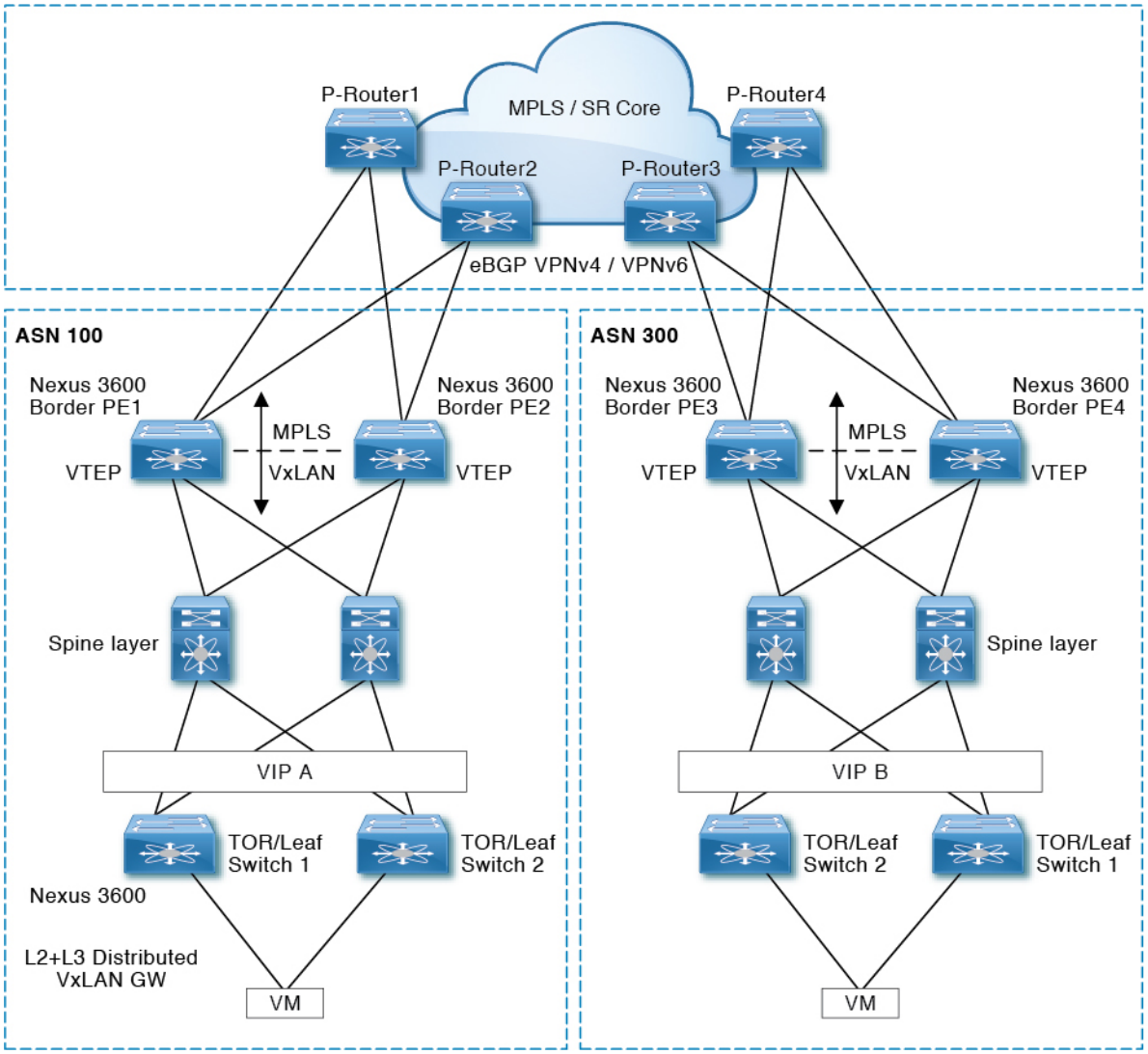
Information About Configuring Seamless Integration of EVPN with L3VPN (MPLS SR)

Data Center (DC) deployments have adopted VXLAN EVPN for its benefits such as EVPN control-plane learning, multitenancy, seamless mobility, redundancy, and easier POD additions. Similarly, the CORE is either an Label Distribution Protocol (LDP)-based MPLS L3VPN network or transitioning from the traditional MPLS L3VPN LDP-based underlay to a more sophisticated solution like Segment Routing (SR). Segment Routing is adopted for its benefits such as:

- Unified IGP and MPLS control planes
- Simpler traffic engineering methods
- Easier configuration
- SDN adoption

With two different technologies, one within the data center (DC) and one in the CORE, there is a natural necessity to handoff from VXLAN to an MPLS-based core at the DCI nodes, which sit on the edge of the DC domain, interfacing with the Core edge router.

Figure 8: Topology Overview



In the previous diagram, two DC pods, each running VXLAN, are being Layer 3 extended over a WAN/Core running MPLS/SR. Another method is classical MPLS L3VPN using LDP. The edge devices in the DC domain (border PE1, PE2, PE3, and PE4) are the DCI nodes doing the handoff between VXLAN and the MPLS-based Core network.

Guidelines and Limitations for Configuring Seamless Integration of EVPN with L3VPN (MPLS SR)

Feature	Cisco Nexus 3600	Comments
VXLAN EVPN to SR-L3VPN	Yes	Extend Layer 3 connectivity between different DC pods Underlay IGP/BGP with SR extensions.
VXLAN EVPN to SR-L3VPN	Yes	Extend Layer 3 connectivity between DC POD running VXLAN and any domain(DC or CORE) running SR.
VXLAN EVPN to MPLS L3VPN (LDP)	Yes	Underlay is LDP.

The following features are supported:

- Layer 3 orphans
- Layer 3 hand-off
- Layer 3 physical interfaces type for core-facing ports
- Per-VRF labels
- LDP
- Segment routing

Note Segment routing and LDP cannot co-exist.

The following features are not supported:

- vPC for redundancy
- Subnet stretches across the DC domain
- SVI/Subinterfaces configured MAC addresses
- Statistics
- SVI toward the MPLS core
- End-to-End Time to Live (TTL) support only in pipe mode for handoff scenario
- End-to-End Explicit Congestion Notification (ECN) for handoff scenario

Configuring Seamless Integration of EVPN with L3VPN (MPLS SR)

The following procedure imports and reoriginates the routes from the VXLAN domain to the MPLS domain and in the other direction.

Before you begin

SUMMARY STEPS

1. **configure terminal**
2. **feature-set mpls**
3. **nv overlay evpn**
4. **feature bgp**
5. **feature mpls l3vpn**
6. **feature mpls segment-routing**
7. **feature interface-vlan**
8. **feature vn-segment-vlan-based**
9. **feature nv overlay**
10. **router bgp** *autonomous-system-number*
11. **address-family ipv4 unicast**
12. **redistribute direct route-map** *route-map-name*
13. **network** *address*
14. **exit**
15. **address-family l2vpn evpn**
16. **neighbor** *address* **remote-as** *number*
17. **update-source** *type/id*
18. **ebgp-multihop** *number*
19. **address-family ipv4 unicast**
20. **send-community extended**
21. **exit**
22. **address-family vpnv4 unicast**
23. **send-community extended**
24. **import l2vpn evpn reoriginate**
25. **neighbor** *address* **remote-as** *number*
26. **address-family ipv4 unicast**
27. **send-community extended**
28. **exit**
29. **address-family ipv6 unicast**
30. **send-community extended**
31. **exit**
32. **address-family l2vpn evpn**
33. **send-community extended**
34. **exit**

35. **import vpn unicast reoriginate**

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `switch# configure terminal`	Enters global configuration mode.
Step 2	**feature-set mpls** **Example:** `switch(config)# feature-set mpls`	Enable MPLS feature set.
Step 3	**nv overlay evpn** **Example:** `switch(config)# nv overlay evpn`	Enable VXLAN.
Step 4	**feature bgp** **Example:** `switch(config)# feature bgp`	Enable BGP.
Step 5	**feature mpls l3vpn** **Example:** `switch(config)# feature mpls l3vpn`	Enable Layer 3 VPN.
Step 6	**feature mpls segment-routing** **Example:** `switch(config)# feature mpls segment-routing`	Enable Segment Routing.
Step 7	**feature interface-vlan** **Example:** `switch(config)# feature interface-vlan`	Enable interface VLAN.
Step 8	**feature vn-segment-vlan-based** **Example:** **Example:** `switch(config)# feature vn-segment-vlan-based`	Enable VLAN based VN segment.
Step 9	**feature nv overlay** **Example:** **Example:** `switch(config)# feature nv overlay`	Enable VXLAN.
Step 10	**router bgp** *autonomous-system-number* **Example:**	Configure BGP. The value of *autonomous-system-number* is from 1 to 4294967295.

	Command or Action	Purpose
	`switch(config)# router bgp 1`	
Step 11	**address-family ipv4 unicast** **Example:** `switch(config-router)# address-family ipv4 unicast`	Configure address family for IPv4.
Step 12	**redistribute direct route-map** *route-map-name* **Example:** `switch(config-router-af)# redistribute direct route-map passall`	Configure redistribution.
Step 13	**network** *address* **Example:** `switch(config-router-af)# network 0.0.0.0/0`	Injects prefixes into handoff BGP along with redistribution.
Step 14	**exit** **Example:** `switch(config-router-af)# exit`	Exit command mode.
Step 15	**address-family l2vpn evpn** **Example:** `switch(config-router)# address-family l2vpn evpn`	Configure L2VPN address family.
Step 16	**neighbor** *address* **remote-as** *number* **Example:** `switch(config-router)# neighbor 108.108.108.108 remote-as 65535`	Define eBGP neighbor IPv4 address and remote Autonomous-System (AS) number.
Step 17	**update-source** *type/id* **Example:** `switch(config-router-af)# update-source loopback100`	Define interface for eBGP peering.
Step 18	**ebgp-multihop** *number* **Example:** `switch(config-router)# ebgp-multihop 10`	Specifies multihop TTL for remote peer. The range of *number* is from 2 to 255.
Step 19	**address-family ipv4 unicast** **Example:** `switch(config-router)# address-family ipv4 unicast`	Configure the address family for IPv4.
Step 20	**send-community extended** **Example:** `switch(config-router-af)# send-community extended`	Configures community for BGP neighbors.

	Command or Action	Purpose
Step 21	**exit** **Example:** `switch(config-router-af)# exit`	Exit command mode.
Step 22	**address-family vpnv4 unicast** **Example:** `switch(config-router)# address-family vpnv4 unicast`	Configure the address family for IPv4.
Step 23	**send-community extended** **Example:** `switch(config-router-af)# send-community extended`	Configures community for BGP neighbors.
Step 24	**import l2vpn evpn reoriginate** **Example:** `switch(config-router)# import l2vpn evpn reoriginate`	Reoriginates the route with new RT. Can be extended to use an optional route-map.
Step 25	**neighbor** *address* **remote-as** *number* **Example:** `switch(config-router)# neighbor 175.175.175.2 remote-as 65535`	Define eBGP neighbor IPv4 address and remote Autonomous-System (AS) number.
Step 26	**address-family ipv4 unicast** **Example:** `switch(config-router)# address-family ipv4 unicast`	Configure the address family for IPv4.
Step 27	**send-community extended** **Example:** `switch(config-router-af)# send-community extended`	Configures community for BGP neighbors.
Step 28	**exit** **Example:** `switch(config-router-af)# exit`	Exit command mode.
Step 29	**address-family ipv6 unicast** **Example:** `switch(config-router)# address-family ipv6 unicast`	Configure the IPv6 unicast address family. This is required for IPv6 over VXLAN with an IPv4 underlay.
Step 30	**send-community extended** **Example:** `switch(config-router-af)# send-community extended`	Configures community for BGP neighbors.
Step 31	**exit** **Example:**	Exit command mode.

	Command or Action	Purpose
	`switch(config-router-af)# exit`	
Step 32	**address-family l2vpn evpn** **Example:** `switch(config-router)# address-family l2vpn evpn`	Configure L2VPN address family.
Step 33	**send-community extended** **Example:** `switch(config-router-af)# send-community extended`	Configures community for BGP neighbors.
Step 34	**exit** **Example:** `switch(config-router-af)# exit`	Exit command mode.
Step 35	**import vpn unicast reoriginate** **Example:** `switch(config-router)# import vpn unicast reoriginate`	Reoriginate the route with new RT. Can be extended to use an optional route-map.

Example Configuration for Configuring Seamless Integration of EVPN with L3VPN (MPLS SR)

The following is a sample CLI configuration that is required to import and reoriginate the routes from the VXLAN domain to the MPLS domain and in the reverse direction.

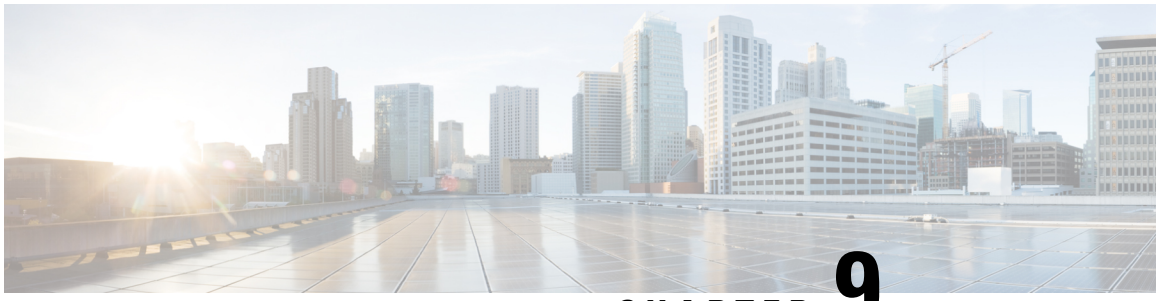
```
switch# sh running-config

!Command: show running-config
!Running configuration last done at: Sat Mar 17 10:00:40 2001
!Time: Sat Mar 17 12:50:12 2001

version 9.2(2) Bios:version 05.22
hardware profile multicast max-limit lpm-entries 0

hostname switch
install feature-set mpls
vdc Scrimshaw id 1
  allow feature-set mpls
  limit-resource vlan minimum 16 maximum 4094
  limit-resource vrf minimum 2 maximum 4096
  limit-resource port-channel minimum 0 maximum 511
  limit-resource u4route-mem minimum 248 maximum 248
  limit-resource u6route-mem minimum 96 maximum 96
  limit-resource m4route-mem minimum 90 maximum 90
  limit-resource m6route-mem minimum 8 maximum 8
feature-set mpls

feature telnet
feature bash-shell
feature sftp-server
nv overlay evpn
```

```
feature ospf
feature bgp
feature mpls l3vpn
feature mpls segment-routing
feature interface-vlan
feature vn-segment-vlan-based
feature bfd
feature nv overlay

no password strength-check
username admin password 5
$5$eEI.wtRs$txfevWxMj/upb/1dJeXy5rNvFYKymzz3Zmc.fpuxTp
1  role network-admin
ip domain-lookup
copp profile strict
snmp-server user admin network-admin auth md5 0x116815e4934ab1f854dce5dd673f33d7
 priv 0x116815e4934ab1f854dce5dd673f33d7 localizedkey
rmon event 1 description FATAL(1) owner PMON@FATAL
rmon event 2 description CRITICAL(2) owner PMON@CRITICAL
rmon event 3 description ERROR(3) owner PMON@ERROR
rmon event 4 description WARNING(4) owner PMON@WARNING
rmon event 5 description INFORMATION(5) owner PMON@INFO

mpls label range 30000 40000 static 6000 8000
vlan 1-2,100,200,555
segment-routing mpls
  global-block 30000 40000
vlan 555
  vn-segment 55500

route-map ALL permit 10
route-map SRmap permit 10
  set label-index 666
route-map ULAY_NETWORK permit 10
  set label-index 600
route-map passall permit 10
vrf context ch5_swap
  ip route 199.1.1.0/24 16.1.1.2
  ip route 200.1.1.0/24 16.1.1.2
vrf context evpn
  vni 55500
  rd auto
  address-family ipv4 unicast
    route-target import 100:55500
    route-target import 100:55500 evpn
    route-target import 6:6000
    route-target export 100:55500
    route-target export 100:55500 evpn
    route-target export 6:6000
  address-family ipv6 unicast
    route-target import 6:6000
    route-target export 6:6000
vrf context management
  ip route 0.0.0.0/0 172.31.144.1
hardware forwarding unicast trace
vlan configuration 2
  ip igmp snooping static-group 225.1.1.1 interface Ethernet1/9

interface Vlan1

interface Vlan555
  no shutdown
  vrf member evpn
```

```
interface nve1
  no shutdown
  host-reachability protocol bgp
  source-interface loopback1
  member vni 55500 associate-vrf

interface Ethernet1/12
  mpls ip forwarding
  no shutdown

interface Ethernet1/13

interface Ethernet1/14
  no shutdown

interface Ethernet1/15
  no shutdown

interface Ethernet1/16
  no shutdown

interface Ethernet1/17
  no shutdown

interface Ethernet1/18

interface Ethernet1/19

interface Ethernet1/20
  no shutdown

interface Ethernet1/21
  ip address 6.2.0.1/24
  mpls ip forwarding
  no shutdown

interface Ethernet1/21.1
  encapsulation dot1q 1211
  vrf member evpn
  ip address 6.22.0.1/24
  no shutdown

interface Ethernet1/21.2
  encapsulation dot1q 1212
  ip address 6.222.0.1/24
  no shutdown

interface Ethernet1/21.3
  encapsulation dot1q 1213
  vrf member ch5_swap
  ip address 16.1.1.1/24
  no shutdown

interface Ethernet1/22
  no shutdown

interface Ethernet1/23
  description underlay
  ip address 6.1.0.1/24
  mpls ip forwarding
  no shutdown

interface Ethernet1/23.1
  encapsulation dot1q 1231
```

```
  vrf member evpn
  ip address 6.11.0.1/23
  no shutdown

interface Ethernet1/24
  no shutdown

interface Ethernet1/25
  no shutdown

interface Ethernet1/26
  description underlay
  ip address 6.0.0.1/24
  mpls ip forwarding
  no shutdown

interface Ethernet1/26.1
  encapsulation dot1q 1261
  ip address 7.0.0.1/24
  no shutdown

interface Ethernet1/27
  no shutdown

interface Ethernet1/28
  no shutdown

interface Ethernet1/29
  no shutdown

interface Ethernet1/30
  no shutdown

interface Ethernet1/31
  ip address 1.31.1.1/24
  no shutdown

interface Ethernet1/32
  no shutdown

interface Ethernet1/33
  ip address 87.87.87.1/24
  ip router ospf 100 area 0.0.0.0
  no shutdown

interface Ethernet1/34
  no shutdown

interface Ethernet1/35
  no shutdown

interface Ethernet1/36
  no shutdown

interface mgmt0
  vrf member management
  ip address 172.31.145.107/21

interface loopback1
  ip address 58.58.58.58/32

interface loopback6
  description used for SR underlay testing
  ip address 6.6.6.1/32
```

```
line console
line vty
monitor session 1
  source interface Ethernet1/21 rx
  source interface Ethernet1/23 both
  destination interface sup-eth0

mpls static configuration
  address-family ipv4 unicast
    lsp SL_AGG_BELL
      in-label 6001 allocate policy 88.1.1.0 255.255.255.0
        forward
          path 1 next-hop 6.0.0.2 out-label-stack implicit-null
router ospf 100
  redistribute direct route-map ALL
router bgp 600
  address-family ipv4 unicast
    network 6.6.6.1/32 route-map SRmap
    network 66.1.1.0/24 route-map ULAY_NETWORK
    redistribute direct route-map passall
    maximum-paths 32
    allocate-label all
  neighbor 6.0.0.2
    remote-as 50
    ebgp-multihop 255
    address-family ipv4 labeled-unicast
  neighbor 6.1.0.2
    remote-as 50
    ebgp-multihop 255
    address-family ipv4 labeled-unicast
  neighbor 6.6.6.3
    remote-as 300
    update-source loopback6
    ebgp-multihop 255
    address-family vpnv4 unicast
      send-community
      send-community extended
      next-hop-self
      import l2vpn evpn reoriginate
  neighbor 7.0.0.2
    remote-as 50
    ebgp-multihop 255
    address-family ipv4 labeled-unicast
  neighbor 21.21.21.21
    remote-as 600
    update-source loopback1
    address-family l2vpn evpn
      send-community
      send-community extended
      import vpn unicast reoriginate
  vrf evpn
    address-family ipv4 unicast
      advertise l2vpn evpn
      redistribute direct route-map passall
      redistribute hmm route-map passall
    address-family ipv6 unicast
      redistribute direct route-map passall
```



CHAPTER 9

Configuring Seamless Integration of EVPN (TRM) with MVPN

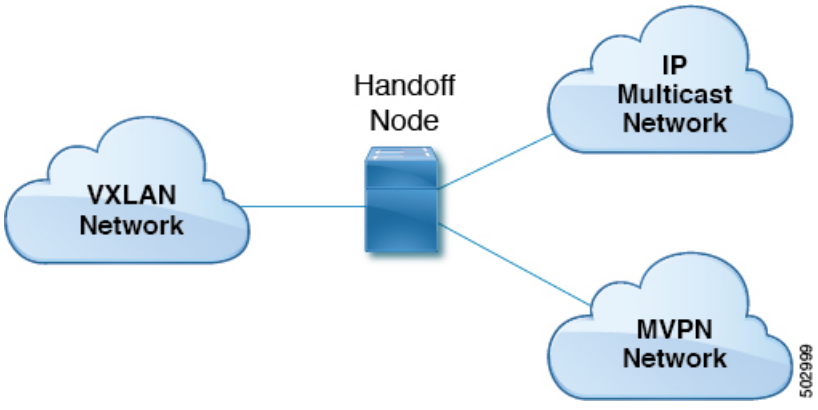
This chapter contains the following sections:

About Seamless Integration of EVPN (TRM) with MVPN (Draft Rosen)

Seamless integration of EVPN (TRM) with MVPN (draft rosen) enables packets to be handed off between a VXLAN network (TRM or TRM Multi-Site) and an MVPN network. To support this feature, VXLAN TRM and MVPN must be supported on a Cisco Nexus device node, the handoff node.

The handoff node is the PE for the MVPN network and the VTEP for the VXLAN network. It connects to the VXLAN, MVPN, and IP multicast networks, as shown in the following figure.

Figure 9: VXLAN - MVPN Handoff Network



Sources and receivers can be in any of the three networks (VXLAN, MVPN, or IP multicast).

All multicast traffic (that is, the tenant traffic from the VXLAN, MVPN, or multicast network) is routed from one domain to another domain. The handoff node acts as the central node. It performs the necessary packet forwarding, encapsulation, and decapsulation to send the traffic to the respective receivers.

Supported RP Positions

The rendezvous point (RP) for the customer (overlay) network can be in any of the three networks (VXLAN, MVPN, or IP multicast).

Table 2: Supported RP Locations

RP Locations	Description
RP in IP network	• The RP can be connected only to the MVPN PE and not to the handoff nodes. • The RP can be connected only to the VXLAN handoff nodes. • The RP can be connected to both the MVPN PE and VXLAN.
RP internal to VXLAN fabric	All VTEPs are RPs inside the VXLAN fabric. All MVPN PEs use the RP configured on the VXLAN fabric.
RP on VXLAN MVPN handoff node	The RP is the VXLAN MVPN handoff node.
RP in MVPN network	The RP is external to the VXLAN network. It's configured on one of the nodes in the MPLS cloud, other than the handoff node.
RP Everywhere (PIM Anycast RP or MSDP-based Anycast RP)	The Anycast RP can be configured on the VXLAN leaf. The RP set can be configured on the handoff node or any MVPN PE.

Guidelines and Limitations for Seamless Integration of EVPN (TRM) with MVPN

This feature has the following guidelines and limitations:

- The handoff node can have local (directly connected) multicast sources or receivers for the customer network.
- Any existing underlay properties, such as ASM/SSM for MVPN or ASM for TRM, are supported on the handoff node.
- The handoff node supports PIM SSM and ASM for the overlay.
- Inter-AS option A is supported on the handoff node toward the IP multicast network.

- The total number of supported MDT source loopback IP addresses and NVE loopback IP addresses is 16. If the number of loopback IP addresses exceeds this limit, traffic drops might occur.
- The following functionality isn't supported for seamless integration of EVPN (TRM) with MVPN:
  - vPC on the handoff node
  - VXLAN ingress replication
  - SVIs and subinterfaces as core-facing interfaces for MVPN
  - Inter-AS options B and C on MVPN nodes
  - PIM SSM as a VXLAN underlay
  - Bidirectional PIM as an underlay or overlay
  - ECMP with a mix of MPLS and IP paths
- Any existing limitations for VXLAN, TRM, and MVPN also apply to seamless integration of EVPN (TRM) with MVPN.

Configuring the Handoff Node for Seamless Integration of EVPN (TRM) with MVPN

This section documents the configurations that are required on the handoff node. Configurations for other nodes (such as VXLAN leafs and spines, MVPN PE, and RS/RR) are the same as in previous releases.

PIM/IGMP Configuration for the Handoff Node

Follow these guidelines when configuring PIM/IGMP for the handoff node:

- Make sure that the Rendezvous Point (RP) is different for TRM and the MVPN underlay, as shown in the following example.

```
ip pim rp-address 90.1.1.100 group-list 225.0.0.0/8 --- TRM Underlay
ip pim rp-address 91.1.1.100 group-list 233.0.0.0/8 --- MVPN Underlay
```

- Use a common RP for overlay multicast traffic.
- The RP can be in static, PIM Anycast, or PIM MSDP mode. The following example shows the RP configuration inside the VRF:

```
vrf context vrfVxLAN5001
  vni 5001
  ip pim rp-address 111.1.1.1 group-list 226.0.0.0/8
  ip pim rp-address 112.2.1.1 group-list 227.0.0.0/8
```

- Enable IGMP snooping for VXLAN traffic using the **ip igmp snooping vxlan** command.
- Enable PIM sparse mode on all source interfaces and interfaces required to carry PIM traffic.

BGP Configuration for the Handoff Node

Follow these guidelines when configuring BGP for the handoff node:

- Add all VXLAN leafs as L2EVPN and TRM neighbors; include the redundant handoff node. If a route reflector is used, add only RR as a neighbor.
- Add all MVPN PEs as VPN neighbors. In MDT mode, add the MVPN PEs as MDT neighbors.
- Import configuration to advertise unicast routes from L2EVPN neighbors to VPN neighbors and vice versa.
- The BGP source identifier can be different or the same as the source interfaces used for the VTEP identifier (configured under the NVE interface)/MVPN PE identifier.

```
feature bgp
address-family ipv4 mdt
address-family ipv4 mvpn

neighbor 2.1.1.1
  address-family ipv4 mvpn
    send-community extended
  address-family l2vpn evpn
    send-community extended
    import vpn unicast reoriginate

neighbor 30.30.30.30
  address-family vpnv4 unicast
    send-community
    send-community extended
    next-hop-self
    import l2vpn evpn reoriginate
  address-family ipv4 mdt
    send-community extended
    no next-hop-third-party
```

- Never use Inter-AS option B between MVPN peers. Instead, configure the **no allocate-label option-b** command under the VPNv4 unicast address family.

```
address-family vpnv4 unicast
    no allocate-label option-b
```

- Set maximum paths should be set in EBGP mode.

```
address-family l2vpn evpn
    maximum-paths 8
vrf vrfVxLAN5001
    address-family ipv4 unicast
        maximum-paths 8
```

- If handoff nodes are deployed in dual mode, use the **route-map** command to avoid advertising prefixes associated with orphan hosts under the VPN address family.

```
ip prefix-list ROUTES_CONNECTED_NON_LOCAL seq 2 premit 15.14.0.15/32

route-map ROUTES_CONNECTED_NON_LOCAL deny
    match ip address prefix-list ROUTES_CONNECTED_NON_LOCAL

neighbor 8.8.8.8
    remote-as 100
    update-source loopback1
    address-family vpnv4 unicast
      send-community
```

```
send-community extended
route-map ROUTES_CONNECTED_NON_LOCAL out
```

VXLAN Configuration for the Handoff Node

Follow these guidelines when configuring VXLAN for the handoff node:

- Enable the following features:

```
feature nv overlay
feature ngmvpn
feature interface-vlan
feature vn-segment-vlan-based
```

- Configure the required L3 VNI:

```
L3VNIs are mapped to tenant VRF.
vlan 2501
  vn-segment 5001 <-- Associate VNI to a VLAN.
```

- Configure the NVE interface:

```
interface nve1
  no shutdown
  host-reachability protocol bgp
  source-interface loopback1 <-- This interface should not be the same as the MVPN
source interface.
  global suppress-arp
member vni 5001 associate-vrf <-- L3VNI
  mcast-group 233.1.1.1 <-- The underlay multicast group for VXLAN should be different
 from the MVPN default/data MDT.
```

- Configure the tenant VRF:

```
vrf context vrfVxLAN5001
  vni 5001 <-- Associate VNI to VRF.
  rd auto
address-family ipv4 unicast
    route-target both auto
    route-target both auto mvpn
    route-target both auto evpn

interface Vlan2501 <-- SVI interface associated with the L3VNI
  no shutdown
  mtu 9216 <-- The overlay header requires 58 byes, so the max tenant traffic is
(Configured MTU - 58).
  vrf member vrfVxLAN5001
  no ip redirects
  ip forward
  ipv6 forward
  no ipv6 redirects
  ip pim sparse-mode <-- PIM is enabled.

interface Vlan2 <-- SVI interface associated with L2 VNI
  no shutdown
  vrf member vrfVxLAN5001
  no ip redirects
  ip address 100.1.1.1/16
  no ipv6 redirects
  ip pim sparse-mode <-- PIM enabled on L2VNI
 fabric forwarding mode anycast-gateway
```

MVPN Configuration for the Handoff Node

Follow these guidelines when configuring MVPN for the handoff node:

- Enable the following features:

```
install feature-set mpls
allow feature-set mpls
feature-set mpls
feature mpls l3vpn
feature mvpn
feature mpls ldp
```

- MPLS LDP Configuration:
  - Enable MPLS LDP (**mpls ip**) on all interfaces that are MPLS links.
  - Do not advertise loopback interfaces used for VXLAN as MPLS prefixes.
    - Configure a prefix list that contains IP addresses that identify the MVPN PE node.

```
ip prefix-list LDP-LOOPBACK seq 51 permit 9.1.1.10/32
ip prefix-list LDP-LOOPBACK seq 52 permit 9.1.2.10/32
```

    - Configure label allocation only for MVPN PE identifiers.

```
mpls ldp configuration
  explicit-null
  advertise-labels for LDP-LOOPBACK
  label allocate global prefix-list LDP-LOOPBACK
```

- Tenant VRF Configuration:
  - For the default MDT mode, make the underlay multicast group the same for all tenant multicast traffic under the VRF.

```
vrf context vrfVxLAN5001
  vni 5001
  mdt default 225.1.100.1
  mdt source loopback100 <-- If the source interface is not configured, the BGP
identifier is used as the source interface.
  mdt asm-use-shared-tree <-- If the underlay is configured in ASM mode
  no mdt enforce-bgp-mdt-safi <-- Enabled by default but should be negated if BGP
MDT should not be used for discovery.
  mdt mtu <mtu-value> <-- Overlay ENCAP Max MTU value
```

  - For the data MDT mode, configure a unique multicast group-set for a subset of or all tenant multicast traffic.

```
mdt data 229.1.100.2/32 immediate-switch
mdt data 232.1.10.4/24 immediate-switch
route-map DATA_MDT_MAP permit 10
  match ip multicast group 237.1.1.1/32
mdt data 235.1.1.1/32 immediate-switch route-map DATA_MDT_MAP
```

- Enable MVPN tunnel statistics.

```
hardware profile mvpn-stats module all
```

CoPP Configuration for the Handoff Node

Both TRM and MVPN are heavily dependent on the control plane. Make sure to set the CoPP policy bandwidth as per the topology.

The following CoPP classes are used for TRM and MVPN traffic:

- **copp-system-p-class-multicast-router** (The default bandwidth is 3000 pps.)
- **copp-system-p-class-l3mc-data** (The default bandwidth is 3000 pps.)
- **copp-system-p-class-l2-default** (The default bandwidth is 50 pps.)
- **copp-class-normal-igmp** (The default bandwidth is 6000 pps.)

The following configuration example shows CoPP policies that can be configured to avoid control packet drops with multicast route scale.

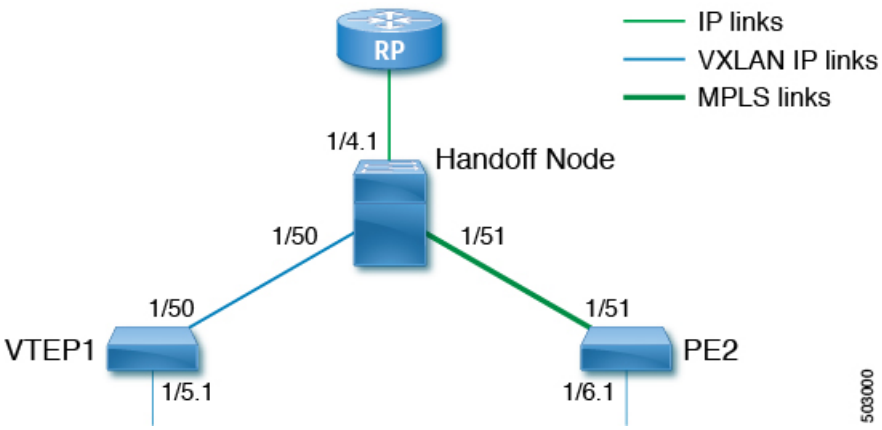
Note The policer values in this example are approximations and might not be optimal for all topologies or traffic patterns. Configure the CoPP policies according to the MVPN/TRM traffic pattern.

```
copp copy profile strict prefix custom
  policy-map type control-plane custom-copp-policy-strict
    class custom-copp-class-normal-igmp
      police cir 6000 pps bc 512 packets conform transmit violate drop
  control-plane
    service-policy input custom-copp-policy-strict

copp copy profile strict prefix custom
  policy-map type control-plane custom-copp-policy-strict
    class custom-copp-class-multicast-router
      police cir 6000 pps bc 512 packets conform transmit violate drop
  control-plane
    service-policy input custom-copp-policy-strict

copp copy profile strict prefix custom
  policy-map type control-plane custom-copp-policy-strict
    class copp-system-p-class-l3mc-data
      police cir 3000 pps bc 512 packets conform transmit violate drop
  control-plane
    service-policy input custom-copp-policy-strict

copp copy profile strict prefix custom
  policy-map type control-plane custom-copp-policy-strict
    class custom-copp-class-l2-default
      police cir 9000 pps bc 512 packets conform transmit violate drop
  control-plane
    service-policy input custom-copp-policy-strict
```

Configuration Example for Seamless Integration of EVPN (TRM) with MVPN

The following figure shows a sample topology with a VXLAN network on the left, an MVPN network on the right, and a centralized handoff node.

Figure 10: Sample Topology for Seamless Integration of EVPN (TRM) with MVPN



The following example show sample configurations for the VTEP, handoff node, and PE in this topology.

Configuration on VTEP1:

```
feature ngmvpn
feature interface-vlan
feature vn-segment-vlan-based
feature nv overlay
feature pim
nv overlay evpn
ip pim rp-address 90.1.1.100 group-list 225.0.0.0/8
ip pim ssm range 232.0.0.0/8

vlan 555
  vn-segment 55500

route-map ALL_ROUTES permit 10
interface nve1
  no shutdown
  host-reachability protocol bgp
  source-interface loopback2
  member vni 55500 associate-vrf
    mcast-group 225.3.3.3

interface loopack1
  ip address 196.196.196.196/32

interface loopback2
  ip address 197.197.197.197/32
  ip pim sparse-mode

feature bgp
router bgp 1
    address-family l2vpn evpn
        maximum-paths 8
```

```
        maximum-paths ibgp 8
    neighbor 2.1.1.2
        remote-as 1
        update-source loopback 1
        address-family ipv4 unicast
          send-community extended
        address-family ipv6 unicast
          send-community extended
        address-family ipv4 mvpn
          send-community extended
        address-family l2vpn evpn
          send-community extended
    vrf vrfVxLAN5023
        address-family ipv4 unicast
          advertise l2vpn evpn
          redistribute direct route-map ALL_ROUTES
          maximum-paths 8
          maximum-paths ibgp 8

vrf context vpn1
  vni 55500
  ip pim rp-address 27.27.27.27 group-list 224.0.0.0/4
  ip pim ssm range 232.0.0.0/8
  ip multicast multipath s-g-hash next-hop-based
rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto mvpn
    route-target both auto evpn

interface Vlan555
  no shutdown
  vrf member vpn1
  ip forward
  ip pim sparse-mode

interface Ethernet 1/50
  ip pim sparse-mode

interface Ethernet1/5.1
  encapsulation dot1q 90
  vrf member vpn1
  ip address 10.11.12.13/24
  ip pim sparse-mode
  no shutdown
```

Configuration on the handoff node:

```
install feature-set mpls
  allow feature-set mpls
feature-set mpls
feature ngmvpn
feature bgp
feature pim
feature mpls l3vpn
feature mvpn
feature mpls ldp
feature interface-vlan
feature vn-segment-vlan-based
feature nv overlay
nv overlay evpn

ip pim rp-address 90.1.1.100 group-list 225.0.0.0/8
ip pim rp-address 91.1.1.100 group-list 232.0.0.0/8
```

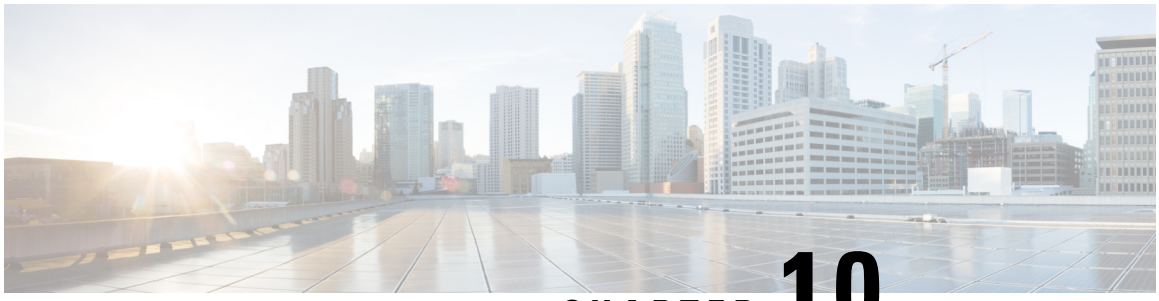
```
interface loopback1
  ip address 90.1.1.100 /32
  ip pim sparse-mode

interface loopback2
  ip address 91.1.1.100 /32
  ip pim sparse-mode

ip prefix-list LDP-LOOPBACK seq 2 permit 20.20.20.20/32
ip prefix-list LDP-LOOPBACK seq 3 permit 30.30.30.30/32
mpls ldp configuration
    advertise-labels for LDP-LOOPBACK
    label allocate label global prefix-list LDP-LOOPBACK

interface Ethernet 1/50
    ip pim sparse-mode

interface Ethernet 1/51
    ip pim sparse-mode
    mpls ip

interface Ethernet1/4.1
  encapsulation dot1q 50
  vrf member vpn1
  ip pim sparse-mode
  no shutdown

interface loopback0
  ip address 20.20.20.20/32
  ip pim sparse-mode

vlan 555
  vn-segment 55500

route-map ALL_ROUTES permit 10

interface nve1
  no shutdown
  host-reachability protocol bgp
  source-interface loopback3
  member vni 55500 associate-vrf
    mcast-group 225.3.3.3

interface loopback3
  ip address 198.198.198.198/32
  ip pim sparse-mode

vrf context vpn1
  vni 55500
  ip pim rp-address 27.27.27.27 group-list 224.0.0.0/4
  ip pim ssm range 232.0.0.0/8
  ip multicast multipath s-g-hash next-hop-based
  mdt default 232.1.1.1
  mdt source loopback 0
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto mvpn
    route-target both auto evpn

interface Vlan555
  no shutdown
  vrf member vpn1
```

```
  ip forward
  ip pim sparse-mode

router bgp 1
    address-family l2vpn evpn
        maximum-paths 8
        maximum-paths ibgp 8
    address-family vpnv4 unicast
        no allocate-label option-b
    address-family ipv4 mdt
    address-family ipv4 mvpn
        maximum-paths 8
        maximum-paths ibgp 8
    neighbor 196.196.196.196
        remote-as 1
        address-family ipv4 unicast
          send-community extended
        address-family ipv6 unicast
          send-community extended
        address-family ipv4 mvpn
          send-community extended
        address-family l2vpn evpn
          send-community extended
          import vpn unicast reoriginate

router bgp 1
    neighbor 30.30.30.30
        remote-as 100
        update-source loopback0
        ebgp-multihop 255
        address-family ipv4 unicast
          send-community extended
        address-family vpnv4 unicast
          send-community
          send-community extended
          next-hop-self
          import l2vpn evpn reoriginate
        address-family ipv4 mdt
          send-community extended
          no next-hop-third-party
```

Configuration on PE2:

```
install feature-set mpls
  allow feature-set mpls
feature-set mpls
feature bgp
feature pim
feature mpls l3vpn
feature mpls ldp
feature interface-vlan

ip pim rp-address 91.1.1.100 group-list 232.0.0.0/8
ip prefix-list LDP-LOOPBACK seq 2 permit 20.20.20.20/32
ip prefix-list LDP-LOOPBACK seq 3 permit 30.30.30.30/32
mpls ldp configuration
    advertise-labels for LDP-LOOPBACK
    label allocate label global prefix-list LDP-LOOPBACK

interface Ethernet 1/51
    ip pim sparse-mode
    mpls ip

interface Ethernet1/6.1
```

```
  encapsulation dot1q 50
  vrf member vpn1
  ip pim sparse-mode
  no shutdown

interface loopback0
  ip address 30.30.30.30/32
  ip pim sparse-mode

vrf context vpn1
  ip pim rp-address 27.27.27.27 group-list 224.0.0.0/4
  ip pim ssm range 232.0.0.0/8
  ip multicast multipath s-g-hash next-hop-based
  mdt default 232.1.1.1
  mdt source loopback 0
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto mvpn
    route-target both auto evpn

router bgp 100
      router-id 30.30.30.30
      address-family vpnv4 unicast
            additional-paths send
            additional-paths receive
            no allocate-label option-b
      neighbor 20.20.20.20
            remote-as 1
            update-source loopback0
            address-family vpnv4 unicast
                send-community
                send-community extended
            address-family ipv4 mdt
                send-community extended
                no next-hop-third-party
```



CHAPTER 10

Configuring vPC Fabric Peering

This chapter contains the following sections:

Information About vPC Fabric Peering

vPC Fabric Peering provides an enhanced dual-homing access solution without the overhead of wasting physical ports for vPC Peer Link. This feature preserves all the characteristics of a traditional vPC.

The following lists the vPC Fabric Peering solution:

- vPC Fabric Peering port-channel with virtual members (tunnels).
- vPC Fabric Peering (tunnel) with removal of the physical peer link requirement.
- vPC Fabric Peering up/down events are triggered based on route updates and fabric up/down.
- Uplink tracking for extended failure coverage.
- vPC Fabric Peering reachability via the routed network, such as the spine.
- Increased resiliency of the vPC control plane over TCP-IP (CFSoIP).
- Data plane traffic over the VXLAN tunnel.
- Communication between vPC member switches uses VXLAN encapsulation.
- Failure of all uplinks on a node result in vPC ports going down on that switch. In that scenario, vPC peer takes up the primary role and forwards the traffic.
- Uplink tracking with state dependency and up/down signalization for vPCs.
- Positive uplink state tracking drives vPC primary role election.
- For border leafs and spines, there is no need for per-VRF peering since network communication uses the fabric.
- Enhance forwarding to orphans hosts by extending the VIP/PIP feature to Type-2 routes.

- Infra-VLAN is not required for vPC fabric peering.

Note The vPC Fabric Peering counts as three VTEPs unlike a normal vPC which counts as one VTEP.

Guidelines and Limitations for vPC Fabric Peering

The following are the vPC Fabric Peering guidelines and limitations:

- Cisco Nexus 9332C, 9364C, and 9300-EX/FX/FXP/FX2/FX3/GX/GX2/H2R platform switches support vPC Fabric Peering. Cisco Nexus 9200 and 9500 platform switches do not support vPC Fabric Peering.

Note For Cisco Nexus 9300-EX switches, mixed-mode multicast and ingress replication are not supported. VNIs must be configured with either multicast or IR underlay, but not both.

- Beginning with Cisco NX-OS Release 10.2(3)F, vPC Fabric Peering is supported on Cisco Nexus C36180YC-R and N3K-C3636C-R platforms. You need to enable TCAM carving in these R-series modules for the vPC Fabric Peering to work.
- The following guidelines and limitations are applicable only to Cisco Nexus C36180YC-R and N3K-C3636C-R platforms:
  - With vPC Fabric Peering enabled, ingress PACL MAC feature is not supported.
  - With vPC Fabric Peering enabled, Layer 2 SPAN based on Layer 2 filters using MAC and CoS values are not supported. Other filters for Layer 2 SPAN are supported.
  - With vPC Fabric Peering enabled, Uni-dimensional scale of Layer 3 host/adjacency will come down to half.
  - On the steady state with all vPC PO's up, BUM traffic from core are received on both the vPC peers. For all flows vPC primary will forward the traffic to vPC PO's.
  - Layer 3 Tenant Routed Multicast (TRM) is not supported.
  - BGP peering behind vPC PO is not supported.
  - IGMP snooping is not supported with vPC Fabric peering
  - DHCP relay agent is not supported.
  - vPC Fabric Peering is not supported on hand off node.
- vPC Fabric Peering requires TCAM carving of region ing-flow-redirect. TCAM carving requires saving the configuration and reloading the switch prior to using the feature. (This requirement does not apply to Cisco Nexus 9300-GX platform switches.)
- Prior to reconfiguring the vPC Fabric Peering source and destination IP, the vPC domain must be shut down. Once the vPC Fabric Peering source and destination IP have been adjusted, the vPC domain can be enabled (**no shutdown**).

- The source and destination IP supported in **virtual peer-link destination** command are class A, B, and C. Class D and E are not supported for vPC Fabric Peering.
- The vPC Fabric Peering peer-link is established over the transport network (the spine layer of the fabric). As communication between vPC peers occurs in this manner, control plane information CFS messages used to synchronize port state information, VLAN information, VLAN-to-VNI mapping, host MAC addresses are transmitted over the fabric. CFS messages are marked with the appropriate DSCP value, which should be protected in the transport network. The following example shows a sample QoS configuration on the spine layer of Cisco Nexus 9000 Series switches.

  Classify traffic by matching the DSCP value (DSCP 56 is the default value):

```
class-map type qos match-all CFS
  match dscp 56
```

  Set traffic to the qos-group that corresponds with the strict priority queue for the appropriate spine switch. In this example, the switch sends traffic to qos-group 7, which corresponds to the strict priority queue (Queue 7). Note that different Cisco Nexus platforms might have a different queuing structure.

```
policy-map type qos CFS
  class CFS
    Set qos-group 7
```

  Assign a classification service policy to all interfaces toward the VTEP (the leaf layer of the network):

```
interface Ethernet 1/1
  service-policy type qos input CFS
```

- Beginning with Cisco NX-OS Release 10.1(1), FEX Support is provided with vMCT for IPv4 underlay on Cisco Nexus 9300-EX/FX/FX2/FX3 platform switches.
- Beginning with Cisco NX-OS Release 10.2(2)F, FEX Support is provided with vMCT for IPv4 underlay on Cisco Nexus 9300-GX platform switches.
- Beginning with Cisco NX-OS Release 10.1(1), vPC Fabric Peering supports FEX in Straight Through and Active-Active (dual home) modes in N9K-C9336C-FX2-E, N9K-C93108TC-EX, N9K-C93108TC-FX,N9K-C93180YC-EX, N9K-C93180YC-FX, N9K-C93216TC-FX2, N9K-C93240YC-FX2, N9K-C93360YC-FX2, N9K-C9336C-FX2, N9K-C93180YC-FX3, N9K-C93180YC-FX3S platform switches.

  Refer to *Cisco Nexus 2000 Series NX-OS Fabric Extender Configuration Guide for Cisco Nexus 9000 Series Switches* for details on FEX (Straight Through and Active-Active modes).
- The vPC Fabric Peering domain is not supported in the role of a Multi-Site vPC BGW.
- Enhance forwarding to orphan hosts by extending the VIP/PIP feature to Type-2 routes.
- Layer 3 Tenant Routed Multicast (TRM) is supported. Layer 2/Layer 3 TRM (Mixed Mode) is not supported.
- If Type-5 routes are used with this feature, the **advertise-pip** command is a mandatory configuration.
- VTEPs behind vPC ports are not supported. This means that virtual peer-link peers cannot act as a transit node for the VTEPs behind the vPC ports.
- SVI and sub-interface uplinks are not supported.

- An orphan Type-2 host is advertised using PIP. A vPC Type-2 host is advertised using VIP. This is the default behavior for a Type-2 host.

  To advertise an orphan Type-5 route using PIP, you need to advertise PIP under BGP.

- Traffic from remote VTEP to orphan hosts would land on the actual node which has the orphans. Bouncing of the traffic is avoided.

Note When the vPC leg is down, vPC hosts are still advertised with the VIP IP.

- Non-disruptive ISSU NX-OS software upgrades are not supported on switches configured with the vPC Fabric Peering feature.
- Beginning with Cisco NX-OS Release 10.2(3)F, ND-ISSU and LXC-ISSU are supported with vMCT for IPv4 underlay on Cisco Nexus 9300-EX/FX/FXP/FX2/FX3/GX/GX2 ToR switches.
- Beginning with Cisco NX-OS Release 10.3(2)F, the vPC Fabric Peering is supported for IPv6 underlay on Cisco Nexus 9300-EX/FX/FXP/FX2/FX3/GX/GX2 ToR switches.
- Beginning with Cisco NX-OS Release 10.4(1)F, the vPC Fabric Peering is supported for IPv6 underlay on Cisco Nexus 9332D-H2R switches.
- Beginning with Cisco NX-OS Release 10.3(2)F, ND-ISSU and LXC-ISSU are supported with vMCT for IPv6 underlay on Cisco Nexus 9300-EX/FX/FXP/FX2/FX3/GX/GX2 ToR switches.
- vMCT for IPv6 underlay does not support attaching FEX to it.
- When converting vPC fabric peering to a physical peer link, make sure to reload the switch.

Configuring vPC Fabric Peering

Ensure the vPC Fabric Peering DSCP value is consistent on both vPC member switches. Ensure that the corresponding QoS policy matches the vPC Fabric Peering DSCP marking.

All VLANs that require communication traversing the vPC Fabric Peering must have a VXLAN enabled (vn-segment); this includes the native VLAN.

Note For MSTP, VLAN 1 must be extended across vPC Fabric Peering if the peer-link and vPC legs have the default native VLAN configuration. This behavior can be achieved by extending VLAN 1 over VXLAN (vn-segment). If the peer-link and vPC legs have non-default native VLANs, those VLANs must be extended across vPC Fabric Peering by associating the VLANs with VXLAN (vn-segment).

Use the **show vpc virtual-peerlink vlan consistency** command for verification of the existing VLAN-to-VXLAN mapping used for vPC Fabric Peering.

peer-keepalive command for vPC Fabric Peering is supported with one of the following configurations:

- Management interface
- Dedicated Layer 3 link in default or non-default VRF

- Loopback interface reachable using the spine.

Configuring Features

Example uses OSPF as the underlay routing protocol.

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

feature nv overlay
```

vPC Configuration

Note To change the vPC Fabric Peering source or destination IP, the vPC domain must be shutdown prior to modification. The vPC domain can be returned to operation after the modifying by using the **no shutdown** command.

Configuring TCAM Carving

```
hardware access-list tcam region ing-racl 0
hardware access-list tcam region ing-sup 768
hardware access-list tcam region ing-flow-redirect 512
```

Note

- When configuring fabric vPC peering, the minimum size for Ingress-Flow-redirect TCAM region size is 512. Also ensure that the TCAM region size is always configured in multiples of 512.
- TCAM carving is not supported on Cisco Nexus 9300-GX/GX2/H2R platform switches.
- Switch reload is required for the TCAM carving to take effect.

Configuring the vPC Domain

For IPv4

```
vpc domain 100
peer-keepalive destination 192.0.2.1
virtual peer-link destination 192.0.2.100 source 192.0.2.20/32 [dscp <dscp-value>]
Warning: Appropriate TCAM carving must be configured for virtual peer-link vPC
peer-switch
peer-gateway
ip arp synchronize
ipv6 nd synchronize
exit
```

For IPv6

```
vpc domain 100
peer-keepalive destination 192:0:2::1
virtual peer-link destination 192:0:2::100 source 192:0:2::20/32 [dscp <dscp-value>]
Warning: Appropriate TCAM carving must be configured for virtual peer-link vPC
```

```
peer-switch
peer-gateway
ipv6 arp synchronize
ipv6 nd synchronize
exit
```

Note The **dscp** keyword in optional. Range is 1 to 63. The default value is 56.

Configuring vPC Fabric Peering Port Channel

No need to configure members for the following port channel.

```
interface port-channel 10
switchport
switchport mode trunk
vpc peer-link
```

```
interface loopback0
```

Note This loopback is not the NVE source-interface loopback (interface used for the VTEP IP address).

For IPv4

```
interface loopback 0
ip address 192.0.2.20/32
ip router ospf 1 area 0.0.0.0
```

For IPv6

```
interface loopback 0
ipv6 address 192:0:2::20/32
ipv6 router ospfv3 1 area 0.0.0.0
```

Note You can use the loopback for BGP peering or a dedicated loopback. This lookback must be different that the loopback for peer keep alive.

Configuring the Underlay Interfaces

Both L3 physical and L3 port channels are supported. SVI and sub-interfaces are not supported.

For IPv4

```
router ospf 1
interface Ethernet1/16
ip address 192.0.2.2/24
ip router ospf 1 area 0.0.0.0
no shutdown
interface Ethernet1/17
port-type fabric
ip address 192.0.2.3/24
ip router ospf 1 area 0.0.0.0
no shutdown
interface Ethernet1/40
port-type fabric
ip address 192.0.2.4/24
```

```
ip router ospf 1 area 0.0.0.0
no shutdown
interface Ethernet1/41
port-type fabric
ip address 192.0.2.5/24
ip router ospf 1 area 0.0.0.0
no shutdown
```

For IPv6

```
router ospfv3 1
interface Ethernet1/16
ipv6 address 192:0:2::2/24
ipv6 router ospfv3 1 area 0.0.0.0
no shutdown
interface Ethernet1/17
port-type fabric
ipv6 address 192:0:2::3/24
ipv6 router ospfv3 1 area 0.0.0.0
no shutdown
interface Ethernet1/40
port-type fabric
ipv6 address 192:0:2::4/24
ipv6 router ospfv3 1 area 0.0.0.0
no shutdown
interface Ethernet1/41
port-type fabric
ipv6 address 192:0:2::5/24
ipv6 router ospfv3 1 area 0.0.0.0
no shutdown
```

Note All ports connected to spines must be port-type fabric.

VXLAN Configuration

Note Configuring **advertise virtual-rmac** (NVE) and **advertise-pip** (BGP) are required steps.

Configuring VLANs and SVI

```
vlan 10
vn-segment 10010
vlan 101
vn-segment 10101
interface Vlan101
no shutdown
mtu 9216
vrf member vxlan-10101
no ip redirects
ip forward
ipv6 address use-link-local-only
no ipv6 redirects
interface vlan10
no shutdown
mtu 9216
vrf member vxlan-10101
no ip redirects
ip address 192.0.2.102/24
ipv6 address 2001:DB8:0:1::1/64
```

```
no ipv6 redirects
fabric forwarding mode anycast-gateway
```

Configuring Virtual Port Channel

```
interface Ethernet1/3
switchport
switchport mode trunk
channel-group 100
no shutdown
exit
interface Ethernet1/39
switchport
switchport mode trunk
channel-group 101
no shutdown
interface Ethernet1/46
switchport
switchport mode trunk
channel-group 102
no shutdown
interface port-channel100
vpc 100
interface port-channel101
vpc 101
interface port-channel102
vpc 102
exit
```

Migrating from vPC to vPC Fabric Peering

This procedure contains the steps to migration from a regular vPC to vPC Fabric Peering.

Any direct Layer 3 link between vPC peers should be used only for peer-keep alive. This link should not be used to advertise paths for vPC Fabric Peering loopbacks.

Note This migration is disruptive.

Before you begin

We recommend that you shut all physical Layer 2 links between the vPC peers before migration. We also recommend that you map VLANs with vn-segment before or after migration.

SUMMARY STEPS

1. **configure terminal**
2. **show vpc**
3. **show port-channel summary**
4. **interface ethernet** *slot/port*
5. **no channel-group**
6. Repeat steps 4 and 5 for each interface.
7. **show running-config vpc**
8. **vpc domain** *domain-id*

9. **virtual peer-link destination** *dest-ip* **source** *source-ip*
10. **interface {ethernet | port-channel}** *value*
11. **port-type fabric**
12. (Optional) **show vpc fabric-ports**
13. **virtual peer-link destination** *dest-ip* | *dest_ipv6* **source** *source-ip* | *source_ipv6* **dhcp** *dhcp_val*
14. **hardware access-list tcam region ing-flow-redirect** *tcam-size*
15. **copy running-config startup-config**
16. **reload**

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `switch# configure terminal`	Enters global configuration mode.
Step 2	**show vpc** **Example:** `switch(config)# show vpc`	Determine the number of members in the port channel.
Step 3	**show port-channel summary** **Example:** `switch(config)# show port-channel summary`	Determine the number of members.
Step 4	**interface ethernet** *slot/port* **Example:** `switch(config)# interface ethernet 1/4`	Specifies the interface you are configuring. **Note** This is the peer link port channel.
Step 5	**no channel-group** **Example:** `switch(config-if)# no channel-group`	Remove vPC peer-link port-channel members. **Note** Disruption occurs following this step.
Step 6	Repeat steps 4 and 5 for each interface. **Example:**	
Step 7	**show running-config vpc** **Example:** `switch(config-if)# show running-config vpc`	Determine the vPC domain.
Step 8	**vpc domain** *domain-id* **Example:** `switch(config-if)# vpc domain 100`	Enter vPC domain configuration mode.
Step 9	**virtual peer-link destination** *dest-ip* **source** *source-ip* **Example:**	Specify the destination and source IP addresses for vPC fabric peering.

	Command or Action	Purpose
	`switch(config-vpc-domain)# virtual peer-link destination 192.0.2.1 source 192.0.2.100`	
Step 10	**interface {ethernet | port-channel}** *value* **Example:** `switch(config-if)# interface Ethernet1/17`	Specifies the L3 underlay interface you are configuring.
Step 11	**port-type fabric** **Example:** `switch(config-if)# port-type fabric`	Configures port-type fabric for underlay interface. **Note** All ports connected to spines must be port-type fabric.
Step 12	(Optional) **show vpc fabric-ports** **Example:** `switch# show vpc fabric-ports`	Displays the fabric ports connected to spine.

Step 13	**virtual peer-link destination** *dest-ip* | *dest_ipv6* **source** *source-ip* | *source_ipv6* **dhcp** *dhcp_val*
	Example:
	For IPv4
	`switch(config-vpc-domain)# virtual peer-link destination 192.0.2.1 source 192.0.2.100 dhcp 56`
	Example:
	For IPv6
	`switch(config-vpc-domain)# virtual peer-link destination 6001:aaa::11 source 6001:aaa::22 dhcp 56`
		Specify the destination and source IPv4/IPv6 addresses for vPC fabric peering.
		Note The IPv4 vPC Fabric peering config works only with the IPv4 VXLAN underlay and the IPv6 vPC Fabric peering config will work only with the IPv6 VXLAN underlay.

Step 14	**hardware access-list tcam region ing-flow-redirect** *tcam-size* **Example:** `switch(config-vpc-domain)# hardware access-list tcam region ing-flow-redirect 512`	Perform TCAM carving. The minimum size for Ingress-Flow-redirect TCAM region size is 512. Also ensure it is configured in multiples of 512.
Step 15	**copy running-config startup-config** **Example:** `switch(config-vpc-domain)# copy running-config startup-config`	Copies the running configuration to the startup configuration.
Step 16	**reload** **Example:** `switch(config-vpc-domain)# reload`	Reboots the switch.

Verifying vPC Fabric Peering Configuration

To display the status for the vPC Fabric Peering configuration, enter one of the following commands:

Table 3: vPC Fabric Peering Verification Commands

Command	Purpose
show vpc fabric-ports	Displays the fabric ports state.
show vpc	Displays information about vPC Fabric Peering mode.
show vpc virtual-peerlink vlan consistency	Displays the VLANs which are not associated with vn-segment.

Example of the show vpc fabric-ports Command

```
switch# show vpc fabric-ports
Number of Fabric port : 9
Number of Fabric port active : 9

Fabric Ports State
-------------------------------------
Ethernet1/9 UP
Ethernet1/19/1 ( port-channel151 ) UP
Ethernet1/19/2 ( port-channel151 ) UP
Ethernet1/19/3 UP
Ethernet1/19/4 UP
Ethernet1/20/1 UP
Ethernet1/20/2 ( port-channel152 ) UP
Ethernet1/20/3 ( port-channel152 ) UP
Ethernet1/20/4 ( port-channel152 ) UP
```

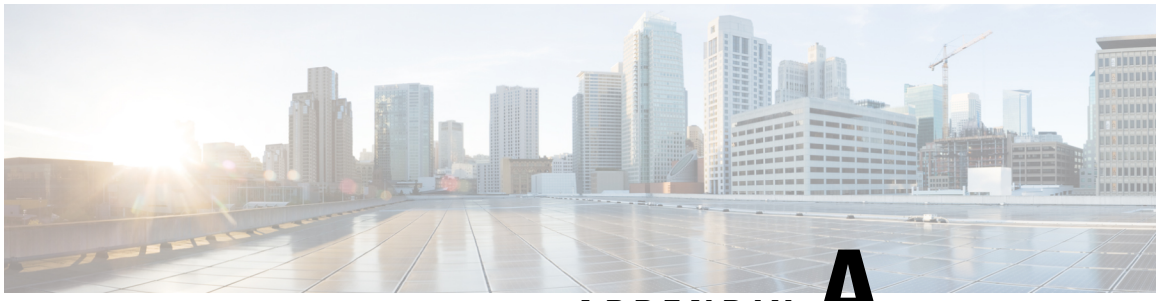
Example of the show vpc Command

```
switch# show vpc
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link

vPC domain id                     : 3
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status  : success
Per-vlan consistency status       : success
Type-2 consistency status         : success
vPC role                          : primary
Number of vPCs configured         : 1
Peer Gateway                      : Enabled
Dual-active excluded VLANs        : -
Graceful Consistency Check        : Enabled
Auto-recovery status              : Enabled, timer is off.(timeout = 240s)
Delay-restore status              : Timer is off.(timeout = 30s)
Delay-restore SVI status          : Timer is off.(timeout = 10s)
Operational Layer3 Peer-router    : Disabled
Virtual-peerlink mode             : Enabled

vPC Peer-link status
```

```
--------------------------------------------------------------------
id     Port   Status Active vlans
--     ----   ------ --------------------------------------------------
1      Po100  up     1,56,98-600,1001-3401,3500-3525

vPC status
----------------------------------------------------------------------------
Id     Port          Status Consistency Reason                Active vlans
--     ------------  ------ ----------- ------                ---------------
101    Po101         up     success     success               98-99,1001-280
                                                              0

Please check "show vpc consistency-parameters vpc <vpc-num>" for the
consistency reason of down vpc and for type-2 consistency reasons for
any vpc.

ToR_B1#
```

Example of the show vpc virtual-peerlink vlan consistency Command

```
switch# show vpc virtual-peerlink vlan consistency
Following vlans are inconsistent
23
switch#
```



APPENDIX A

DHCP Relay in VXLAN BGP EVPN

This appendix contains the following sections:

DHCP Relay in VXLAN BGP EVPN Overview

DHCP relay is supported by VXLAN BGP EVPN and is useful in a multi-tenant VXLAN EVPN deployment to provision DHCP service to EVPN tenant clients.

In a multi-tenant EVPN environment, DHCP relay uses the following sub-options of Option 82:

- Sub-option 151(0x97) - Virtual Subnet Selection

  (Defined in RFC#6607.)

  Used to convey VRF related information to the DHCP server in an MPLS-VPN and VXLAN EVPN multi-tenant environment.

- Sub-option 11(0xb) - Server ID Override

  (Defined in RFC#5107.)

  The server identifier (server ID) override sub-option allows the DHCP relay agent to specify a new value for the server ID option, which is inserted by the DHCP server in the reply packet. This sub-option allows the DHCP relay agent to act as the actual DHCP server such that the renew requests will come to the relay agent rather than the DHCP server directly. The server ID override sub-option contains the incoming interface IP address, which is the IP address on the relay agent that is accessible from the client. Using this information, the DHCP client sends all renew and release request packets to the relay agent. The relay agent adds all of the appropriate sub-options and then forwards the renew and release request packets to the original DHCP server. For this function, Cisco's proprietary implementation is sub-option 152(0x98). You can use the **ip dhcp relay sub-option type cisco** command to manage the function.

- Sub-option 5(0x5) - Link Selection

  (Defined in RFC#3527.)

The link selection sub-option provides a mechanism to separate the subnet/link on which the DHCP client resides from the gateway address (giaddr), which can be used to communicate with the relay agent by the DHCP server. The relay agent will set the sub-option to the correct subscriber subnet and the DHCP server will use that value to assign an IP address rather than the giaddr value. The relay agent will set the giaddr to its own IP address so that DHCP messages are able to be forwarded over the network. For this function, Cisco's proprietary implementation is sub-option 150(0x96). You can use the **ip dhcp relay sub-option type cisco** command to manage the function.

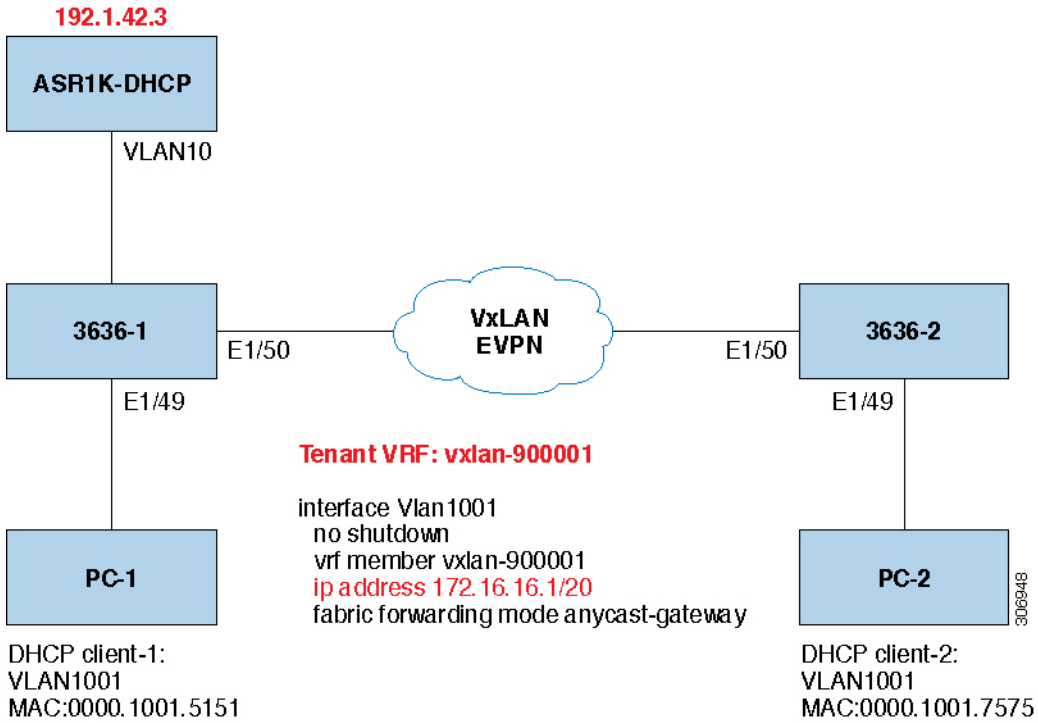
Guidelines and Limitations for DHCP Relay

The following are the guidelines and limitations for DHCP Relay in VXLAN BGP EVPN:

- Beginning in Cisco NX-OS Release 9.2(2), support is added for Cisco Nexus 3636C-R and 36180YC-R.
- IPv6 DHCP is not supported for Cisco Nexus 3636C-R and 36180YC-R switches.

DHCP Relay in VXLAN BGP EVPN Example

Figure 11: Example Topology



Topology characteristics:

- Switches 3636-1 and 3636-2 are VTEPs connected to VXLAN fabric.
- Client1 and client2 are DHCP clients in vlan1001. They belong to tenant VRF vxlan-900001.

- The DHCP server is ASR1K, a router that sits in vlan10.
- DHCP server configuration

```
ip vrf vxlan900001
ip dhcp excluded-address vrf vxlan900001 172.16.16.1 172.16.16.9
ip dhcp pool one
 vrf vxlan900001
 network 172.16.16.0 255.255.240.0
 defaultrouter 172.16.16.1
```

Basic VXLAN BGP EVPN Configuration

- 3636-1

```
version 7.0(3)I1(3)
version 9.2(1)
hostname 3636C-R

nv overlay evpn
feature vn—segment—vlan—based
feature nv overlay

fabric forwarding anycast—gateway—mac 0000.1111.2222

vlan 101
  vn—segment 900001
vlan 1001
  vn—segment 2001001

vrf context vxlan—900001
  vni 900001
  rd auto
  address—family ipv4 unicast
    route—target both auto
    route—target both auto evpn

interface Vlan1O1
  no shutdown
  vrf member vxlan—900001
  ip forward

interface Vlan1OO1
  no shutdown
  vrf member vxlan—900001
  ip address 172.16.16.1/20
  fabric forwarding mode anycast—gateway
```

Note You can choose either of the following two procedures to create NVE interfaces. Use the first option for a small number of VNIs. Use the second option 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


interfaca Ethernetl/49
  switchport mode trunk
  switchport trunk allowed vlan 10,1001
  spanning—tree port type edge trunk

interface Ethernetl/50
  no switchport
  ip address 192.1.33.2/24
  ip router ospf 1 area 0.0.0.0
  ip pire sparse-mode
  no shutdown

interface loopback0
  ip address 1.1.1.1/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse—mode

interface loopbackl
  vrf member vxlan—900001
  ip address 11.11.11.11/32

router bgp 65535
  router—id 1.1.1.1
  log—neighbor-changes
```

```
  neighbor 2.2.2.2 remote—as 65535
    update—source loopback0
    address-family l2vpn evpn
      send-community both
  vrf vxlen—900001
    address—family ipv4 unicast
    network 11.11.11.11/32
    network 192.1.42.0/24
    advertise l2vpn evpn
evpn
  vni 2001001 12
```

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
    route—target import auto
    route—target export auto
```

- 3636-2

```
version 7.0(3)I1(3)
version 9.2(1)
hostname 3636-1

nv overlay evpn
feature vn—segment—vlan—based
feature nv overlay

fabric forwarding anycast—gateway—mac 0000.1111.2222

vlan 101
  vn—segment 900001
vlan 1001
  vn—segment 2001001

vrf context vxlan—900001
  vni 900001
  rd auto
  address—family ipv4 unicast
    route-target both auto
    route—target both auto evpn

interface VianlOl
no shutdown
vrf member vxlan-900001
ip forward

interface Vlanl00l
  no shutdown
  vrf member vxlan—900001
  ip address 172.16.16.1/20
  fabric forwarding mcde anycast—gateway
```

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

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


interface VianlOl
no shutdown
vrf member vxlan-900001
ip forward

interface Vlanl00l
  no shutdown
  vrf member vxlan—900001
  ip address 172.16.16.1/20
  fabric forwarding mcde anycast—gateway
```

Note You can choose either of the following two procedures for creating the NVE interfaces. Use the first option for a small number of VNIs. Use the second option 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


interface Ethernetl/49
  switchport mode trunk
  switchport trunk alluwed vlan 10,1001
  spanning—tree port type edge trunk

interface Ethernetl/50
  no switchport
  ip address 192.1.34.2/24
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  no shutdown

interface loopback0
  ip address 2.2.2.2/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse—mode

interface loopbackl
  vrf member vxlan—900001
  ip address 22.22.22.22/32

router bgp 65535
  router—id 2.2.2.2
  log—neighbor-changes
  neighbor 1.1.1.1 remote—as 65535
    update—source loopback0
    address-family l2vpn evpn
      send-community both
  vrf vxlen—900001
    address—family ipv4 unicast
    network 22.22.22.22/32

    advertise l2vpn evpn
evpn
  vni 2001001 12
```

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

```
rd auto
    route—target import auto
    route—target export auto
```

DHCP Relay on VTEPs

The following are common deployment scenarios:

- Client on tenant VRF and server on Layer 3 default VRF.

- Client on tenant VRF (SVI X) and server on the same tenant VRF (SVI Y).
- Client on tenant VRF (VRF X) and server on different tenant VRF (VRF Y).
- Client on tenant VRF and server on non-default non-VXLAN VRF.

The following sections below move vlan10 to different VRFs to depict different scenarios.

Client on Tenant VRF and Server on Layer 3 Default VRF

Put DHCP server (192.1.42.3) into the default VRF and make sure it is reachable from both 3636-1 and 3636-2 through the default VRF.

```
3636-1# sh run int vl 10

!Command: show running-config interface Vlan10
!Time: Mon Aug 7 07:51:16 2018

version 9.2(1)

interface Vlan10
  no shutdown
  ip address 192.1.42.1/24
  ip router ospf 1 area 0.0.0.0

3636-1# ping 192.1.42.3 cou 1

PING 192.1.42.3 (192.1.42.3): 56 data bytes
64 bytes from 192.1.42.3: icmp_seq=0 ttl=254 time=0.593 ms
- 192.1.42.3 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
roundtrip min/avg/max = 0.593/0.592/0.593 ms

3636-2# ping 192.1.42.3 cou 1
PING 192.1.42.3 (192.1.42.3): 56 data bytes
64 bytes from 192.1.42.3: icmp_seq=0 ttl=252 time=0.609 ms
- 192.1.42.3 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 0.609/0.608/0.609 ms
```

DHCP Relay Configuration

- 3636-1

```
3636-1# sh run dhcp

!Command: show running-config dhcp
!Time: Mon Aug 6 08:26:00 2018

version 9.2(1)
feature dhcp

service dhcp
ip dhcp relay
ip dhcp relay information option
ip dhcp relay information option vpn
ipv6 dhcp relay
```

```
interface Vlan1001
  ip dhcp relay address 192.1.42.3 use—vrf default
```

- 3636-2

```
3636-2# sh run dhcp

!Command: show running—config dhcp
!Time: Mon Aug 6 08:26:16 2018

version 9.2(1)
feature dhcp

service dhcp
ip dhcp relay
ip dhcp relay information option
ip dhcp relay information option vpn
ipv6 dhcp relay

interfaoe Vlan1001
  ip dhcp relay address 192.1.42.3 use—vrf default
```

Debug Output

- The following is a packet dump for DHCP interact sequences.

```
3636-1# ethanalyzer local interface inband display-filter
"udp.srcport==67 or udp.dstport==67" limit-captured frames 0

Capturing on inband
20150824 08:35:25.066530 0.0.0.0 -> 255.255.255.255 DHCP DHCP Discover - Transaction
ID 0x636a38fd
20150824 08:35:25.068141 192.1.42.1 -> 192.1.42.3 DHCP DHCP Discover - Transaction ID
0x636a38fd
20150824 08:35:27.069494 192.1.42.3 -> 192.1.42.1 DHCP DHCP Offer Transaction - ID
0x636a38fd
20150824 08:35:27.071029 172.16.16.1 -> 172.16.16.11 DHCP DHCP Offer Transaction - ID
0x636a38fd
20150824 08:35:27.071488 0.0.0.0 -> 255.255.255.255 DHCP DHCP Request Transaction - ID
 0x636a38fd
20150824 08:35:27.072447 192.1.42.1 -> 192.1.42.3 DHCP DHCP Request Transaction - ID
0x636a38fd
20150824 08:35:27.073008 192.1.42.3 -> 192.1.42.1 DHCP DHCP ACK Transaction - ID
0x636a38fd
20150824 08:35:27.073692 172.16.16.1 -> 172.16.16.11 DHCP DHCP ACK Transaction - ID
0x636a38fd
```

Note Ethanalyzer might not capture all DHCP packets because of inband interpretation issues when you use the filter. You can avoid this by using SPAN.

- DHCP Discover packet 3636-1 sent to DHCP server.

  giaddr is set to 192.1.42.1 (ip address of vlan10) and suboptions 5/11/151 are set accordingly.

```
Bootp flags: 0x0000 (unicast)
client IP address: 0.0.0.0 (0.0.0.0)
Your (client) IP address: 0.0.0.0 (0.0.0.0)
Next server IP address: 0.0.0.0 (0.0.0.0)
Relay agent IP address: 192.1.42.1 (192.1.42.1)
client MAC address Hughes_01:51:51 (00:00:10:01:51:51)
client hardware address padding: 00000000000000000000
Server host name not given
Boot file name not given
Magic cookie: DHCP
Option: (53) DHCP Message Type
  Length: 1
  DHCP: Discover (1)
Option: (55) Parameter Request List
  Length: 4
  Parameter Request List Item: (1) Subnet Mask
  Parameter Request List Item: (3) Router
  Parameter Request List Item: (58) Renewal Time Value
  Parameter Request List Item: (59) Rebinding Time Value
Option: (61) client identifier
  Length: 7
  Hardware type: Ethernet (0x01)
  Client MAC address: Hughes_01:51:51 (00:00:10:01:51:51)
Option: (82) Agent Information Option
  Length: 47
Option 82 Suboption: (1) Agent Circuit ID
  Length: 10
  Agent Circuit ID: 01080006001e88690030
Option 82 Suboption: (2) Agent Remote ID
  Length: 6
  Agent Remote ID: f8c2882333a5
Option 82 Suboption: (151) VRF name/VPN ID
Option 82 Suboption: (11) Server ID Override
  Length: 4
  Server ID Override: 172.16.16.1 (172.16.16.1)
Option 82 Suboption: (5) Link selection
  Length: 4
  Link selection: 172.16.16.0 (172.16.16.0)
```

```
ASR1K-DHCP# sh ip dhcp bin
Bindings from all pools not associated with VRF:
IP address ClientID/ Lease expiration Type State Interface
         Hardware address/
         User name

Bindings from VRF pool vxlan900001:
IP address ClientID/ Lease expiration Type State Interface
         Hardware address/
         User name
172.16.16.10 0100.0010.0175.75 Aug 25 2015 09:21 AM Automatic Active GigabitEthernet2/1/0
172.16.16.11 0100.0010.0151.51 Aug 25 2015 08:54 AM Automatic Active GigabitEthernet2/1/0

3636-1# sh ip route vrf vxlan900001
IP Route Table for VRF "vxlan900001"
'*' denotes best ucast nexthop
'**' denotes best mcast nexthop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

11.11.11.11/32, ubest/mbest: 2/0, attached
```

```
  *via 11.11.11.11, Lo1, [0/0], 18:31:57, local
  *via 11.11.11.11, Lo1, [0/0], 18:31:57, direct
22.22.22.22/32, ubest/mbest: 1/0
  *via 2.2.2.2%default, [200/0], 18:31:57, bgp65535,internal, tag 65535 (evpn)segid:
900001 tunnelid: 0x2020202
encap: VXLAN

172.16.16.0/20, ubest/mbest: 1/0, attached
 *via 172.16.16.1, Vlan1001, [0/0], 18:31:57, direct
172.16.16.1/32, ubest/mbest: 1/0, attached
 *via 172.16.16.1, Vlan1001, [0/0], 18:31:57, local
172.16.16.10/32, ubest/mbest: 1/0
 *via 2.2.2.2%default, [200/0], 00:00:47, bgp65535,internal, tag 65535 (evpn)segid:
900001 tunnelid: 0x2020202
encap: VXLAN

172.16.16.11/32, ubest/mbest: 1/0, attached
 *via 172.16.16.11, Vlan1001, [190/0], 00:28:10, hmm

3636-1# ping 172.16.16.11 vrf vxlan900001 count 1
PING 172.16.16.11 (172.16.16.11): 56 data bytes
64 bytes from 172.16.16.11: icmp_seq=0 ttl=63 time=0.846 ms
- 172.16.16.11 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 0.846/0.845/0.846 ms

3636-1# ping 172.16.16.10 vrf vxlan900001 count 1
PING 172.16.16.10 (172.16.16.10): 56 data bytes
64 bytes from 172.16.16.10: icmp_seq=0 ttl=62 time=0.874 ms
- 172.16.16.10 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 0.874/0.873/0.874 ms
```

Client on Tenant VRF (SVI X) and Server on the Same Tenant VRF (SVI Y)

Put DHCP server (192.1.42.3) into VRF of vxlan-900001 and make sure it is reachable from both 3636-1 and 3636-2 through VRF of vxlan-900001.

```
3636-1# sh run int vl 10

!Command: show running-config interface Vlan10
!Time: Mon Aug 6 09:10:26 2018

version 9.2(1)

interface Vlan10
  no shutdown
  vrf member vxlan-900001
  ip address 192.1.42.1/24
```

Because 172.16.16.1 is an anycast address for vlan1001 configured on all the VTEPs, we need to pick up a unique address as the DHCP relay packet's source address to make sure the DHCP server can deliver a response to the original DHCP Relay agent. In this scenario, we use loopback1 and we need to make sure loopback1 is reachable from everywhere of VRF vxlan-900001.

```
3636-1# sh run int lo1
```

```
!Command: show running-config interface loopback1
!Time: Mon Aug 6 09:18:53 2018

version 9.2(1)

interface loopback1
  vrf member vxlan-900001
  ip address 11.11.11.11/32

3636-1# ping 192.1.42.3 vrf vxlan900001 source 11.11.11.11 cou 1
PING 192.1.42.3 (192.1.42.3) from 11.11.11.11: 56 data bytes
64 bytes from 192.1.42.3: icmp_seq=0 ttl=254 time=0.575 ms
- 192.1.42.3 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 0.575/0.574/0.575 ms

3636-2# sh run int lo1

!Command: show running-config interface loopback1
!Time: Mon Aug 6 09:19:30 2018

version 9.2(1)

interface loopback1
  vrf member vxlan900001
  ip address 22.22.22.22/32

3636-2# ping 192.1.42.3 vrf vxlan-900001 source 22.22.22.22 cou 1
PING 192.1.42.3 (192.1.42.3) from 22.22.22.22: 56 data bytes
64 bytes from 192.1.42.3: icmp_seq=0 ttl=253 time=0.662 ms
- 192.1.42.3 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 0.662/0.662/0.662 ms
```

DHCP Relay Configuration

- 3636-1

```
3636—1# sh run dhcp

!Command: show running-config dhcp
!Time: Mon Aug 6 08:26:00 2018

version 9.2(1)
feature dhcp

service dhcp
ip dhcp relay
ip dhcp relay information option
I4ip dhcp relay information option vpn
ipv6 dhcp relay

interface Vlanl00l
  ip dhcp relay address 192.1.42.3
  ip dhcp relay source—interface loopback1
```

- 3636-2

```
3636—2# sh run dhcp
```

```
!Command: show running-config dhcp
!Time: Mon Aug 6 08:26:16 2018

version 9.2(1)
feature dhcp

service dhcp
ip dhcp relay
ip dhcp relay information option
ip dhcp relay information option vpn
ipv6 dhcp relay

interface Vlan100l
  ip dhcp relay address 192.1.42.3
  ip dhcp relay source—interface loopback1
```

Debug Output

- The following is a packet dump for DHCP interact sequences.

```
3636-1# ethanalyzer local interface inband display-filter
"udp.srcport==67 or udp.dstport==67" limit-captured frames 0

Capturing on inband
20150824 09:31:38.129393 0.0.0.0 -> 255.255.255.255 DHCP DHCP Discover - Transaction
ID 0x860cd13
20150824 09:31:38.129952 11.11.11.11 -> 192.1.42.3 DHCP DHCP Discover - Transaction ID
 0x860cd13
20150824 09:31:40.130134 192.1.42.3 -> 11.11.11.11 DHCP DHCP Offer - Transaction ID
0x860cd13
20150824 09:31:40.130552 172.16.16.1 -> 172.16.16.11 DHCP DHCP Offer - Transaction ID
0x860cd13
20150824 09:31:40.130990 0.0.0.0 -> 255.255.255.255 DHCP DHCP Request - Transaction ID
 0x860cd13
20150824 09:31:40.131457 11.11.11.11 -> 192.1.42.3 DHCP DHCP Request - Transaction ID
0x860cd13
20150824 09:31:40.132009 192.1.42.3 -> 11.11.11.11 DHCP DHCP ACK - Transaction ID
0x860cd13
20150824 09:31:40.132268 172.16.16.1 -> 172.16.16.11 DHCP DHCP ACK - TransactionID
0x860cd13
```

Note Ethanalyzer might not capture all DHCP packets because of inband interpretation issues when you use the filter. You can avoid this by using SPAN.

- DHCP Discover packet 3636-1 sent to DHCP server.

  giaddr is set to 11.11.11.11(loopback1) and suboptions 5/11/151 are set accordingly.

```
Bootstrap Protocol
  Message type: Boot Request (1)
  Hardware type: Ethernet (0x01)
  Hardware address length: 6
  Hops: 1
  Transaction ID: 0x0860cd13
```

```
  Seconds elapsed: O
  Bootp flags: 0x0000 (unicast)
  Client IP address: 0.0.0.0 (0.0.0.0)
  Your (client) IP address: 0.0.0.0 (0.0.0.0)
  Next server IP address: 0.0.0.0 (0.0.0.0)
  Relay agent iP address: 11.11.11.11 (11.11.11.11)
  Client MAC address: Hughes_01:51:51 (00:00:10:01:51:51)
  Client hardware address padding: 00000000000000000000
  Server host name not given
  Boot file name not given
  Magic cookie: DHCP
  Option: (53) DHCP Message Type
    Length: 1
    DHCP: Discover (1)
  Option: (55) Parameter Request List
  Option: (61) Client Identifier
  Option: (82) Agent Information Option
    Length: 47
  Option 82 suboption: (1) Aqent Circuit ID
  Option 82 suboption: (151) Agent Remote ID
  Option 82 suboption: (11) Server ID Override
    Length: 4
    Server ID override: 172.16.16.1 (172.16.16.1)
  Option 82 suboption: (5) Link selection
    Length: 4
    Link selection: 172.16.16.0 (172.16.16.0)
```

```
ASR1K-DHCP# sh ip dhcp bin
Bindings from all pools not associated with VRF:
IP address ClientID/Lease expiration Type State Interface
         Hardware address/
         User name

Bindings from VRF pool vxlan-900001:
IP address ClientID/Lease expiration Type State Interface
         Hardware address/
         User name

172.16.16.10 0100.0010.0175.75 Aug 25 2015 10:02 AM Automatic Active GigabitEthernet2/1/0
172.16.16.11 0100.0010.0151.51 Aug 25 2015 09:50 AM Automatic Active GigabitEthernet2/1/0

3636-1# sh ip route vrf vxlan-900001
IP Route Table for VRF "vxlan-900001"
'*' denotes best ucast nexthop
'**' denotes best mcast nexthop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

11.11.11.11/32, ubest/mbest: 2/0, attached
  *via 11.11.11.11, Lo1, [0/0], 19:13:56, local
  *via 11.11.11.11, Lo1, [0/0], 19:13:56, direct
22.22.22.22/32, ubest/mbest: 1/0
  *via 2.2.2.2%default, [200/0], 19:13:56, bgp65535,internal, tag 65535 (evpn)segid:
900001 tunnelid: 0x2020202
encap: VXLAN
172.16.16.0/20, ubest/mbest: 1/0, attached
  *via 172.16.16.1, Vlan1001, [0/0], 19:13:56, direct
172.16.16.1/32, ubest/mbest: 1/0, attached
  *via 172.16.16.1, Vlan1001, [0/0], 19:13:56, local
172.16.16.10/32, ubest/mbest: 1/0
  *via 2.2.2.2%default, [200/0], 00:01:27, bgp65535,
internal, tag 65535 (evpn)segid: 900001 tunnelid: 0x2020202
```

```
encap: VXLAN
172.16.16.11/32, ubest/mbest: 1/0, attached
  *via 172.16.16.11, Vlan1001, [190/0], 00:13:56, hmm
192.1.42.0/24, ubest/mbest: 1/0, attached
  *via 192.1.42.1, Vlan10, [0/0], 00:36:08, direct
192.1.42.1/32, ubest/mbest: 1/0, attached
  *via 192.1.42.1, Vlan10, [0/0], 00:36:08, local
9372-1# ping 172.16.16.10 vrf vxlan-900001 cou 1
PING 172.16.16.10 (172.16.16.10): 56 data bytes
64 bytes from 172.16.16.10: icmp_seq=0 ttl=62 time=0.808 ms
- 172.16.16.10 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 0.808/0.808/0.808 ms

3636-1# ping 172.16.16.11 vrf vxlan-900001 cou 1
PING 172.16.16.11 (172.16.16.11): 56 data bytes
64 bytes from 172.16.16.11: icmp_seq=0 ttl=63 time=0.872 ms
- 172.16.16.11 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 0.872/0.871/0.872 ms
```

Client on Tenant VRF (VRF X) and Server on Different Tenant VRF (VRF Y)

The DHCP server is placed into another tenant VRF vxlan-900002 so that DHCP response packets can access the original relay agent. We use loopback2 to avoid any anycast ip address that is used as the source address for the DHCP relay packets.

```
3636-1# sh run int vl 10
!Command: show runningconfig interface Vlan10
!Time: Tue Aug 6 08:48:22 2018

version 9.2(1)
interface Vlan10
  no shutdown
  vrf member vxlan900002
  ip address 192.1.42.1/24

3636-1# sh run int lo2
!Command: show runningconfig interface loopback2
!Time: Tue Aug 7 08:48:57 2018
version 9.2(1)
interface loopback2
  vrf member vxlan900002
  ip address 33.33.33.33/32

3636-2# sh run int lo2
!Command: show runningconfig interface loopback2
!Time: Tue Aug 7 08:48:44 2018
version 9.2(1)
interface loopback2
  vrf member vxlan900002
  ip address 44.44.44.44/32

9372-1# ping 192.1.42.3 vrf vxlan-900002 source 33.33.33.33 cou 1
PING 192.1.42.3 (192.1.42.3) from 33.33.33.33: 56 data bytes
64 bytes from 192.1.42.3: icmp_seq=0 ttl=254 time=0.544 ms
- 192.1.42.3 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 0.544/0.544/0.544 ms
```

```
3636-2# ping 192.1.42.3 vrf vxlan-900002 source 44.44.44.44 count 1
PING 192.1.42.3 (192.1.42.3) from 44.44.44.44: 56 data bytes
64 bytes from 192.1.42.3: icmp_seq=0 ttl=253 time=0.678 ms
- 192.1.42.3 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 0.678/0.678/0.678 ms
```

DHCP Relay Configuration

- 3636-1

```
3636—1# sh run dhcp

!Command: show running—config dhcp
!Time: Mon Aug 6 08:26:00 2018

version 9.2(1)
feature dhcp

service dhcp
ip dhcp relay
ip dhcp relay information option
ip dhcp relay information option vpn
ipv6 dhcp relay

interface VlanlOOl
  ip dhcp relay address 192.1.42.3 use—vrf vxlan—900002
  ip dhcp relay source—interface loopback2
```

- 3636-2

```
!Command: show running-config dhcp
!Time: Mon Aug 6 08:26:16 2018

version 9.2(1)
feature dhcp

service dhcp
ip dhcp relay
ip dhcp relay information option
ip dhcp relay information option vpn
ipv6 dhcp relay

interface VlaniOOl
  ip dhcp relay address 192.1.42.3 use-vrf vxlan—900002
  ip dhcp relay source—interface loopback2
```

Debug Output

- The following is a packet dump for DHCP interact sequences.

```
3636-1# ethanalyzer local interface inband display-filter "udp.srcport==67 or
udp.dstport==67" limit-captured-frames 0
Capturing on inband
20180806 08:59:35.758314 0.0.0.0 -> 255.255.255.255 DHCP DHCP Discover - Transaction
ID 0x3eebccae
```

```
20180806 08:59:35.758878 33.33.33.33 -> 192.1.42.3 DHCP DHCP Discover - Transaction ID
 0x3eebccae
20180806 08:59:37.759560 192.1.42.3 -> 33.33.33.33 DHCP DHCP Offer - Transaction ID
0x3eebccae
20180806 08:59:37.759905 172.16.16.1 -> 172.16.16.11 DHCP DHCP Offer - Transaction ID
0x3eebccae
20180806 08:59:37.760313 0.0.0.0 -> 255.255.255.255 DHCP DHCP Request - Transaction ID
 0x3eebccae
20180806 08:59:37.760733 33.33.33.33 -> 192.1.42.3 DHCP DHCP Request - Transaction ID
0x3eebccae
20180806 08:59:37.761297 192.1.42.3 -> 33.33.33.33 DHCP DHCP ACK - Transaction ID
0x3eebccae
20180806 08:59:37.761554 172.16.16.1 -> 172.16.16.11 DHCP DHCP ACK - Transaction ID
0x3eebccae
```

- DHCP Discover packet 3636-1 sent to DHCP server.

  giaddr is set to 33.33.33.33 (loopback2) and suboptions 5/11/151 are set accordingly.

```
Bootstrap Protocol
  Message type: Boot Request (1)
  Hardware type: Ethernet (0x01)
  Hardware address length: 6
  Hops: 1
  Transaction ID: Ox3eebccae
  Seconds elapsed: O
  Bootp flags: 0x0000 (unicast)
  Client IP address: 0.0.0.0 (0.0.0.0)
  Your (client) IP address: 0.0.0.0 (0.0.0.0)
  Next server IP address: 0.0.0.0 (0.0.0.0)
  Relay agent IP address: 33.33.33.33 (33.33.33.33)
  Client MAC address: i-iughes_01:51:51 (00:00:10:01:51:51)
  Client hardware address padding: 00000000000000000000
  Server host name not given
  Boot file name not given
  Magic cookie: DHCP
  Option: (53) DHCP Message Type
    Length: 1
    DHCP: Discover (1)
  Option: (55) Parameter Request List
  Option: (61) client identifier
  Option: (82) Agent Information option
    Length: 47
  Option 82 Suboption: (1) Agent circuit W
  Option 82 suboption: (2) Agent Remote 10
  Option 82 suboption: (151) VRF name/VPN ID
  Option 82 Suboption: (11) Server ID Override
    Length: 4
    Server ID Override: 172.16.16.1 (172.16.16.1)
  Option 82 Suboption: (5) Link selection
    Length: 4
    Link selection: 172.16.16.0 (172.16.16.0)
```

Client on Tenant VRF and Server on Non-Default Non-VXLAN VRF

The DHCP server is placed into the management VRF and is reachable the through M0 interface. The IP address changes to 10.122.164.147 accordingly.

```
3636-1# sh run int m0
!Command: show running-config interface mgmt0
!Time: Tue Aug 7 09:17:04 2018
version 9.2(1)
interface mgmt0
  vrf member management
  ip address 10.122.165.134/25

3636-1# ping 10.122.164.147 vrf management cou 1
PING 10.122.164.147 (10.122.164.147): 56 data bytes
64 bytes from 10.122.164.147: icmp_seq=0 ttl=251 time=1.024 ms
- 10.122.164.147 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 1.024/1.024/1.024 ms

3636-2# sh run int m0
!Command: show running-config interface mgmt0
!Time: Tue Aug 25 09:17:47 2015
version 7.0(3)I1(3)
interface mgmt0
  vrf member management
  ip address 10.122.165.148/25

3636-2# ping 10.122.164.147 vrf management cou 1
PING 10.122.164.147 (10.122.164.147): 56 data bytes
64 bytes from 10.122.164.147: icmp_seq=0 ttl=251 time=1.03 ms
- 10.122.164.147 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 1.03/1.03/1.03 ms
```

DHCP Relay Configuration

- 3636-1

```
3636—1# sh run dhcp 3636—2# sh run dhcp

!Command: show running-config dhcp
!Time: Mon Aug 6 08:26:00 2018

version 9.2(1)
feature dhcp

service dhcp
ip dhcp relay
ip dhcp relay information option
ip dhcp relay information option vpn
ipv6 dhcp relay

interface VlanlOOl
  ip dhcp relay address 10.122.164.147 use—vrf management
```

- 3636-2

```
3636-2# sh run dhcp
!Command: show running-config dhcp
!Time: Tue Aug 7 09:17:47 2018
```

```
version 9.2(1)
feature dhcp

service dhcp
ip dhcp relay
ip dhop relay information option
ip dhcp relay information option vpn
ipv6 dhcp relay

interface VlanlOOl
ip dhcp relay address 10.122.164.147 use—vrf management
```

Debug Output

- The following is a packet dump for DHCP interact sequences.

```
3636-1# ethanalyzer local interface inband display-filter "udp.srcport==67 or
udp.dstport==67" limit-captured-frames 0
Capturing on inband
20180806 09:30:54.214998 0.0.0.0 -> 255.255.255.255 DHCP DHCP Discover - Transaction
ID 0x28a8606d
20180806 09:30:56.216491 172.16.16.1 -> 172.16.16.11 DHCP DHCP Offer - Transaction ID
0x28a8606d
20180806 09:30:56.216931 0.0.0.0 -> 255.255.255.255 DHCP DHCP Request - Transaction ID
 0x28a8606d
20180806 09:30:56.218426 172.16.16.1 -> 172.16.16.11 DHCP DHCP ACK - Transaction ID
0x28a8606d

3636-1# ethanalyzer local interface mgmt display-filter "ip.src==10.122.164.147 or
ip.dst==10.122.164.147" limit-captured-frames 0
Capturing on mgmt0
20180806 09:30:54.215499 10.122.165.134 -> 10.122.164.147 DHCP DHCP Discover - Transaction
 ID 0x28a8606d
20180806 09:30:56.216137 10.122.164.147 -> 10.122.165.134 DHCP DHCP Offer - Transaction
 ID 0x28a8606d
20180806 09:30:56.217444 10.122.165.134 -> 10.122.164.147 DHCP DHCP Request - Transaction
 ID 0x28a8606d
20180806 09:30:56.218207 10.122.164.147 -> 10.122.165.134 DHCP DHCP ACK - Transaction
ID 0x28a8606d
```

- DHCP Discover packet 3636-1 sent to DHCP server.

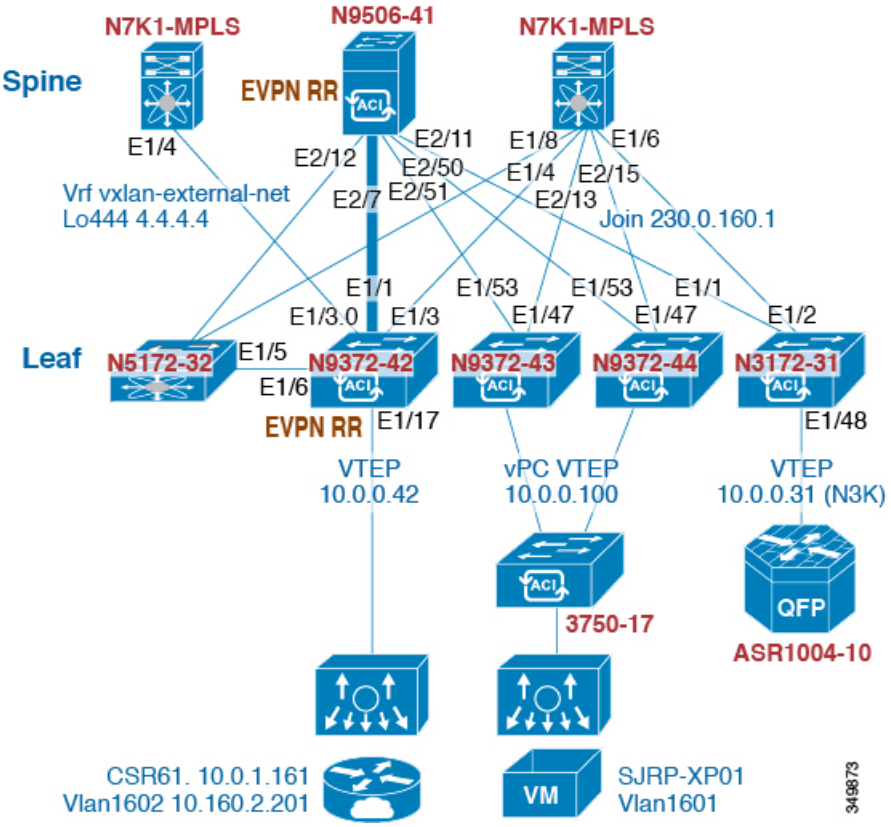
  giaddr is set to 10.122.165.134 (mgmt0) and suboptions 5/11/151 are set accordingly.

```
Bootstrap Protocol
  Message type: Boot Request (1)
  Hardware type: Ethernet (0x01)
  Hardware address length: 6
  Hops: 1
  Transaction ID: 0x28a8606d
  Seconds elapsed: O
  Bootp flags: 0x0000 (Unicast)
  Client IP address: 0.0.0.0 (0.0.0.0)
  Your (client) IP address: 0.0.0.0 (0.0.0.0)
  Next server IP address: 0.0.0.0 (0.0.0.0)
  Relay agent IP address: 10.122.165.134 (10.122.165.134)
  Client MAC address: Hughes_01:51:51 (00:00:10:01:51:51)
  Client hardware address padding: 00000000000000000000
  Server host name not given
  Boot file name not given
  Magic cookie: DHCP
```

```
Option: (53) DHCP Message Type
  Length: 1
  DHCP: Discover (1)
Option: (55) Parameter Request List
Option: (61) Client identifier
Option: (82) Agent Information Option
  Length: 47
  Option 82 Suboption: (1) Agent Circuit ID
  Option 82 Suboption: (2) Aqent Remote ID
  Option 82 Suboption: (151) VRF name/VPN ID
  Option 82 Suboption: (11) Server ID Override
    Length: 4
    Server ID Override: 172.16.16.1 (172.16.16.1)
  Option 82 Suboption: (5) Link selection
    Length: 4
    Link selection: 172.16.16.0 (172.16.16.0)
```

Configuring VPC Peers Example

The following is an example of how to configure routing between VPC peers in the overlay VLAN for a DHCP relay configuration.



- Enable DHCP service.

```
service dhcp
```

- Configure DHCP relay.

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

- Create loopback under VRF where you need DHCP relay service.

```
interface loopback601
  vrf member evpn-tenant-kk1
  ip address 160.1.0.43/32
  ip router ospf 1 area 0     /* Only required for VPC VTEP. */
```

- Advertise LoX into the Layer 3 VRF BGP.

```
Router bgp 2
vrf X
   network 10.1.1.42/32
```

- Configure DHCP relay on the SVI under the VRF.

```
interface Vlan1601
  vrf member evpn-tenant-kk1
  ip address 10.160.1.254/24
  fabric forwarding mode anycast-gateway
  ip dhcp relay address 10.160.2.201
  ip dhcp relay source-interface loopback601
```

- Configure Layer 3 VNI SVI with **ip forward**.

```
interface Vlan1600
  vrf member evpn-tenant-kk1
    ip forward
```

- Create the routing VLAN/SVI forthe VPC VRF.

Note Only required for VPC VTEP.

```
Vlan 1605
interface Vlan1605
  vrf member evpn-tenant-kk1
  ip address 10.160.5.43/24
  ip router ospf 1 area 0.0.0.41
```

- Create the VRF routing.

Note Only required for VPC VTEP.

```
router ospf 1
vrf evpn-tenant-kk1
    router-id 10.160.5.43
```

vPC VTEP DHCP Relay Configuration Example

To address a need to configure a VLAN that is allowed across the MCT/peer-link, such as a vPC VLAN, an SVI can be associated to the VLAN and is created within the tenant VRF. This becomes an underlay peering, with the underlay protocol, such as OSPF, that needs the tenant VRF instantiated under the routing process.

Alternatively, instead of placing the SVI within the routing protocol and instantiate the Tenant-VRF under the routing process, you can use the static routes between the vPC peers across the MCT. This approach ensures that the reply from the server returns to the correct place and each VTEP uses a different loopback interface for the GiAddr.

The following are examples of these configurations:

- Configuration of SVI within underlay routing:

```
/* vPC Peer-1 */

router ospf UNDERLAY
vrf tenant-vrf

interface Vlan2000
  no shutdown
  mtu 9216
  vrf member tenant-vrf
  ip address 192.168.1.1/30
  ip router ospf UNDERLAY area 0.0.0.0


/* vPC Peer-2 */

router ospf UNDERLAY
vrf tenant-vrf

interface Vlan2000
  no shutdown
  mtu 9216
  vrf member tenant-vrf
  ip address 192.168.1.2/30
  ip router ospf UNDERLAY area 0.0.0.0
```

- Configuration of SVI using static routes between vPC peers across the MCT:

```
/* vPC Peer-1 */

interface Vlan2000
  no shutdown
```

```
  mtu 9216
  vrf member tenant-vrf
  ip address 192.168.1.1/30

vrf context tenant-vrf
ip route 192.168.1.2/30 192.168.1.1


/* vPC Peer-2 */

interface Vlan2000
  no shutdown
  mtu 9216
  vrf member tenant-vrf
  ip address 192.168.1.2/30

vrf context tenant-vrf
ip route 192.168.1.1/30 192.168.1.2
```

DHCP Relay in VXLAN BGP EVPN

INDEX

U

V