



Configuring Tenant Routed Multicast

This chapter contains the following sections:

- [About Tenant Routed Multicast, on page 1](#)
- [Guidelines and Limitations for Tenant Routed Multicast, on page 2](#)
- [Guidelines and Limitations for Layer 3 Tenant Routed Multicast, on page 3](#)
- [Rendezvous Point for Tenant Routed Multicast, on page 3](#)
- [Configuring a Rendezvous Point for Tenant Routed Multicast, on page 4](#)
- [Configuring a Rendezvous Point Inside the VXLAN Fabric, on page 4](#)
- [Configuring an External Rendezvous Point, on page 5](#)
- [Configuring Layer 3 Tenant Routed Multicast, on page 7](#)
- [Configuring TRM on the VXLAN EVPN Spine, on page 11](#)
- [Configuring TRM with vPC Support, on page 13](#)

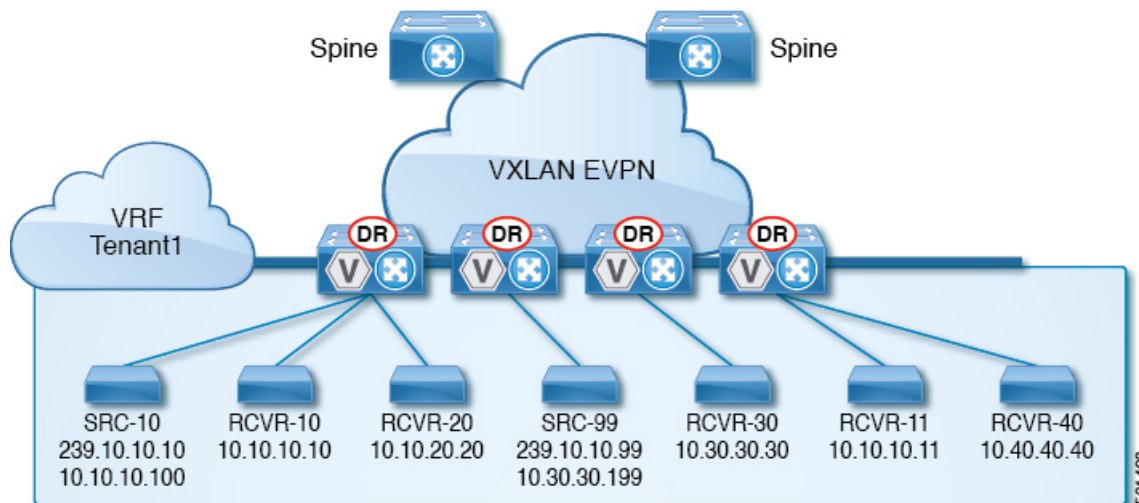
About Tenant Routed Multicast

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 1: 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:

- The [Guidelines and Limitations for VXLANs](#) 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

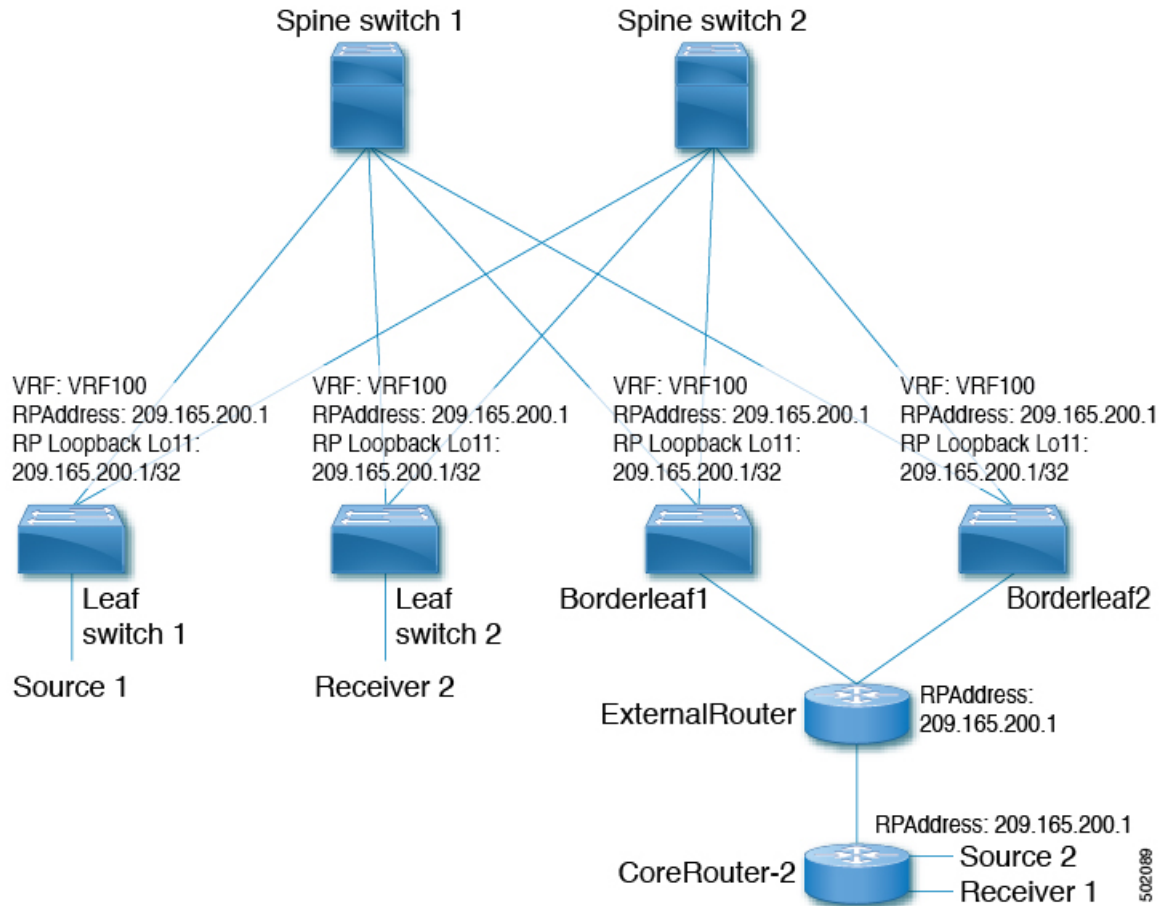
With TRM enabled Internal and External RP is supported. The following table displays the first release in which RP positioning is or is not supported.

	RP Internal	RP External
TRM L2 Mode	N/A	N/A
TRM L3 Mode	7.0(3)I7(1)	7.0(3)I7(4)
TRM L2L3 Mode	7.0(3)I7(1)	N/A

Configuring a Rendezvous Point for Tenant Routed Multicast

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

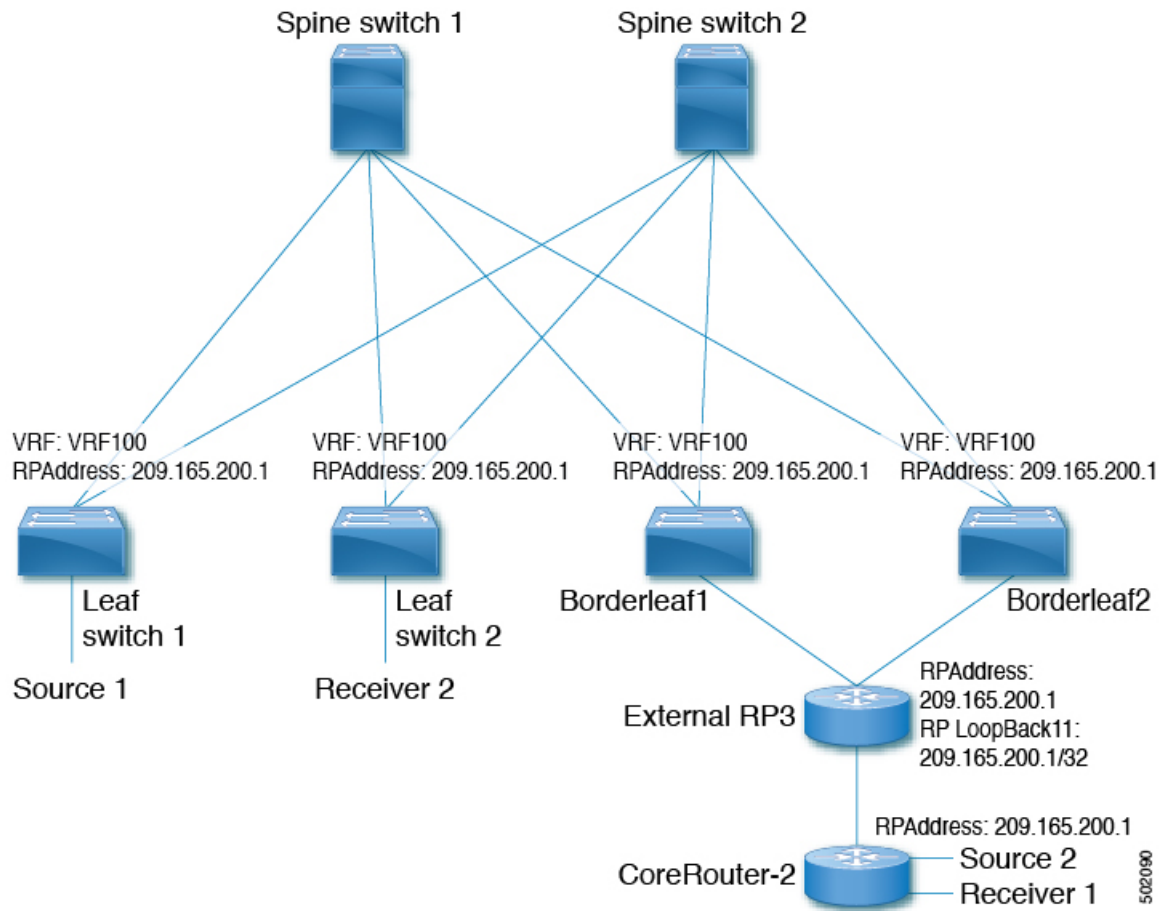
1. **configure terminal**
2. **interface loopback** *loopback_number*
3. **vrf member** *vxlan-number*
4. **ip address** *ip-address*
5. **ip pim sparse-mode**
6. **vrf context** *vrf-name*
7. **ip pim rp-address** *ip-address-of-router group-list group-range-prefix*

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 <i>ip-address-of-router</i> 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# <code>configure terminal</code>	Enter configuration mode.
Step 2	vrf context vrf100 Example: switch(config)# <code>vrf context vrf100</code>	Enter configuration mode.

	Command or Action	Purpose
Step 3	ip pim rp-address <i>ip-address-of-router</i> group-list <i>group-range-prefix</i> Example: <pre>switch(config-vrf)# ip pim rp-address 209.165.200.1 group-list 224.0.0.0/4</pre>	The value of the <i>ip-address-of-router</i> 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: <pre>switch# configure terminal</pre>	Enter configuration mode.
Step 2	feature ngmvpn Example: <pre>switch(config)# feature ngmvpn</pre>	Enables the Next-Generation Multicast VPN (ngMVPN) control plane. New address family commands become available in BGP.
Step 3	ip igmp snooping vxlan Example: <pre>switch(config)# ip igmp snooping vxlan</pre>	Configure IGMP snooping for VXLAN VLANs.
Step 4	interface nve1 Example: <pre>switch(config)# interface nve 1</pre>	Configure the NVE interface.

	Command or Action	Purpose
Step 5	member vni <i>vni-range</i> associate-vrf Example: <pre>switch(config-if-nve)# member vni 200100 associate-vrf</pre>	Configure the Layer 3 virtual network identifier. The range of <i>vni-range</i> is from 1 to 16,777,214.
Step 6	mcast-group <i>ip-prefix</i> Example: <pre>switch(config-if-nve-vni)# mcast-group 225.3.3.3</pre>	<p>Builds the default multicast distribution tree for the VRF VNI (Layer 3 VNI).</p> <p>The multicast group is used in the underlay (core) for all multicast routing within the associated Layer 3 VNI (VRF).</p> <p>Note We recommend that underlay multicast groups for Layer 2 VNI, default MDT, and data MDT not be shared. Use separate, non-overlapping groups.</p>
Step 7	exit Example: <pre>switch(config-if-nve-vni)# exit</pre>	Exits command mode.
Step 8	exit Example: <pre>switch(config-if)# exit</pre>	Exits command mode.
Step 9	router bgp 100 Example: <pre>switch(config)# router bgp 100</pre>	Set autonomous system number.
Step 10	exit Example: <pre>switch(config-router)# exit</pre>	Exits command mode.
Step 11	neighbor <i>ip-addr</i> Example: <pre>switch(config-router)# neighbor 1.1.1.1</pre>	Configure IP address of the neighbor.
Step 12	address-family ipv4 mvpn Example: <pre>switch(config-router-neighbor)# address-family ipv4 mvpn</pre>	Configure multicast VPN.
Step 13	send-community extended Example: <pre>switch(config-router-neighbor-af)# send-community extended</pre>	Enables ngMVPN for address family signalization. The send community extended command ensures that extended communities are exchanged for this address family.

	Command or Action	Purpose
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 <i>vrf_name</i> Example: switch(config-router) # vrf context vrf100	Configure VRF name.
Step 17	ip pim rp-address <i>ip-address-of-router</i> group-list <i>group-range-prefix</i> Example: switch(config-vrf) # ip pim rp-address 209.165.201.1 group-list 226.0.0.0/8	The value of the <i>ip-address-of-router</i> 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 4 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 <i>vlan_id</i> 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.
Step 23	vrf member <i>vrf-num</i> Example: switch(config-if) # vrf member vrf100	Configure VRF name.

	Command or Action	Purpose
Step 24	ip address <i>ip_address</i> 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 if 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 <i>vlan_id</i> 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 <i>vrf100</i> 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.
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](#).

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: <pre>switch# configure terminal</pre>	Enter configuration mode.
Step 2	route-map permitall permit 10 Example: <pre>switch(config)# route-map permitall permit 10</pre>	Configure the route-map. Note The route-map keeps the next-hop unchanged for EVPN routes <ul style="list-style-type: none"> • Required for eBGP • Options for iBGP
Step 3	set ip next-hop unchanged Example: <pre>switch(config-route-map)# set ip next-hop unchanged</pre>	Set next hop address. Note The route-map keeps the next-hop unchanged for EVPN routes <ul style="list-style-type: none"> • Required for eBGP • Options for iBGP

	Command or Action	Purpose
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: switch(config-router-neighbor-af) # rewrite-rt-asn	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 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.

	Command or Action	Purpose
Step 13	route-reflector-client Example: <pre>switch(config-router-neighbor-af) # route-reflector-client</pre>	Configure route reflector. Note Required for iBGP with route-reflector.
Step 14	route-map permitall out Example: <pre>switch(config-router-neighbor-af) # route-map permitall out</pre>	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 team region mac-ifacl**
10. **hardware access-list team 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: <pre>switch# configure terminal</pre>	Enter global configuration mode.

	Command or Action	Purpose
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 <i>address</i> group-list <i>range</i> 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 <i>domain-id</i> 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.
Step 11	reload Example: switch(config)# reload	Reloads the switch config for the TCAM assignments to become active.
Step 12	peer switch Example:	Defines the peer switch.

	Command or Action	Purpose
	<code>switch(config-vpc-domain) # peer switch</code>	
Step 13	peer gateway Example: <code>switch(config-vpc-domain) # peer gateway</code>	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: <code>switch(config-vpc-domain) # peer-keepalive destination 172.28.230.85</code>	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: <code>switch(config-vpc-domain) # ip arp synchronize</code>	Enables IP ARP synchronize under the vPC Domain to facilitate faster ARP table population following device reload.
Step 16	ipv6 nd synchronize Example: <code>switch(config-vpc-domain) # ipv6 nd synchronize</code>	Enables IPv6 and synchronization under the vPC domain to facilitate faster and table population following device reload.
Step 17	Create vPC peer-link. Example: <code>switch(config) # interface port-channel 1</code> <code>switch(config) # switchport</code> <code>switch(config) # switchport mode trunk</code> <code>switch(config) # switchport trunk allowed vlan 1,10,100-200</code> <code>switch(config) # mtu 9216</code> <code>switch(config) # vpc peer-link</code> <code>switch(config) # no shut</code> <code>switch(config) # interface Ethernet 1/1, 1/21</code> <code>switch(config) # switchport</code> <code>switch(config) # mtu 9216</code> <code>switch(config) # channel-group 1 mode active</code> <code>switch(config) # no shutdown</code>	Creates the vPC peer-link port-channel interface and adds two member interfaces to it.

	Command or Action	Purpose
Step 18	system nve infra-vlans <i>range</i> Example: switch(config)# system nve infra-vlans 10	Defines a non-VXLAN enabled VLAN as a backup routed path.
Step 19	vlan <i>number</i> 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 <i>seconds</i> 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.