



Configuring VXLAN EVPN Multi-Site

This chapter contains the following sections:

- [About VXLAN EVPN Multi-Site, on page 1](#)
- [Guidelines and Limitations for VXLAN EVPN Multi-Site, on page 2](#)
- [Enabling VXLAN EVPN Multi-Site, on page 3](#)
- [Configuring VNI Dual Mode, on page 4](#)
- [Configuring Fabric/DCI Link Tracking, on page 5](#)
- [Configuring Fabric External Neighbors, on page 6](#)

About VXLAN EVPN Multi-Site

The VXLAN EVPN Multi-Site solution uses border gateways is either anycast or virtual port channel configuration in the data plane to terminate and interconnect overly domains.

The border gateways provide the network control boundary that is necessary for traffic enforcement and failure containment functionality.

In the control plane, BGP sessions between the border gateways rewrite the next hop information of EVPN routes and re-originate them. VXLAN Tunnel Endpoints (VTEPs) are only aware of their overlay domain internal neighbors including the border gateways. All routes external to the fabric have a next hop on the border gateways for Layer 2 and Layer 3 traffic.

The VXLAN EVPN Multi-Site feature is a solution to interconnect two or more BGP-based Ethernet VPN (EVPN) site's fabrics in a scalable fashion over an IP-only network.

The Border Gateway (BG) is the node that interacts with nodes within a site and with nodes that are external to the site. For example, in a leaf-spine data center fabric, it can be a leaf, a spine, or a separate device acting as a gateway to interconnect the sites.

The VXLAN EVPN Multi-Site feature can be conceptualized as multiple site-local EVPN control planes and IP forwarding domains interconnected via a single common EVPN control and IP forwarding domain. Every EVPN node is identified with a unique site-scope identifier. A site-local EVPN domain consists of EVPN nodes with the same site identifier. Border Gateways on one hand are also part of site-specific EVPN domain and on the other hand a part of a common EVPN domain to interconnect with Border Gateways from other sites. For a given site, these Border Gateways facilitate site-specific nodes to visualize all other sites to be reachable only via them. This would mean:

- Site-local bridging domains are interconnected only via Border Gateways with bridging domains from other sites.

- Site-local routing domains are interconnected only via Border Gateways with routing domains from other sites.
- Site-local flood domains are interconnected only via Border Gateways with flood domains from other sites.

Selective Advertisement is defined as the configuration of the per-tenant information on the border gateway. Specifically, this means IP-VRF or MAC-VRF (EVPN Instance). In cases where External Connectivity (VRF-lite) and EVPN Multi-Site co-exist on the same border gateway, the advertisements are always enabled.

Guidelines and Limitations for VXLAN EVPN Multi-Site

VXLAN EVPN Multi-Site has the following configuration guidelines and limitations:

- Beginning with Cisco NX-OS Release 7.0(3)I7(3), support for VXLAN EVPN Multi-Site functionality on the Cisco Nexus N9K-C9336C-FX and N9K-C93240YC-FX2 is added. N9K-C9348GC-FXP does not support VXLAN EVPN Multi-Site functionality.
- Beginning with Cisco NX-OS Release 7.0(3)I7(2), VXLAN EVPN Multi-Site and Tenant Routed Multicast (TRM) is supported between source and receivers deployed in the same site.
- Beginning with Cisco NX-OS Release 7.0(3)I7(2), the Multi-Site border gateway allows the co-existence of Multi-Site extensions (Layer 2 unicast/multicast and Layer 3 unicast) as well as Layer 3 unicast and multicast external connectivity.
- The following switches support VXLAN EVPN Multi-Site:
 - Cisco Nexus 9300-EX, 9300-FX, and 9500 platform switches with X9700-EX line cards, beginning with Cisco NX-OS Release 7.0(3)I7(1)



Note The Cisco Nexus 9348GC-FXP switch does not support VXLAN EVPN Multi-Site functionality.

- Cisco Nexus 9396C switch and Cisco Nexus 9500 platform switches with X9700-FX line cards, beginning with Cisco Nexus NX-OS Release 7.0(3)I7(2)
- Cisco Nexus 9336C-FX2 switch, beginning with Cisco Nexus NX-OS Release 7.0(3)I7(3)
- The number of border gateways per site is limited to four.
- Border Gateways (BGWs) in a vPC topology are not supported.
- Support for Multicast Flood Domain between inter-site/fabric border gateways is not supported.
- Multicast Underlay between sites is not supported.
- PIM is not supported on multisite VXLAN DCI links.
- iBGP EVPN Peering between border gateways of different fabrics/sites is not supported.
- The **peer-type fabric-external** command configuration is required only for VXLAN Multi-site BGWs (this command must not be used when peering with non-Cisco equipment).



Note The **peer-type fabric-external** command configuration is not required for pseudo BGWs.

- If different Anycast Gateway MAC addresses are configured across sites, ARP suppression must be enabled for all VLANs that have been extended.
- Bind NVE to a loopback address that is separate from loopback addresses that are required by Layer 3 protocols. A best practice is to use a dedicated loopback address for the NVE source interface (PIP VTEP) and Multi-Site source interface (anycast and virtual IP VTEP).

Enabling VXLAN EVPN Multi-Site

This procedure enables the VXLAN EVPN Multi-Site feature. Multi-Site is enabled on the border gateways only. The site-id must be the same on all border gateways in the fabric/site.

Procedure

	Command or Action	Purpose
Step 1	configure terminal Example: <code>switch# configure terminal</code>	Enters global configuration mode.
Step 2	evpn multisite border-gateway <i>ms-id</i> Example: <code>switch(config)# evpn multisite border-gateway 100</code>	Configure the site ID for a site/fabric. The range of values for <i>ms-id</i> is 1 to 2,814,749,767,110,655. The <i>ms-id</i> must be the same in all border gateways within the same fabric/site.
Step 3	interface nve 1 Example: <code>switch(config-evpn-msite-bgw)# interface nve 1</code>	Creates a VXLAN overlay interface that terminates VXLAN tunnels. Note Only 1 NVE interface is allowed on the switch
Step 4	source-interface loopback <i>src-if</i> Example: <code>switch(config-if-nve)# source-interface loopback 0</code>	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 transient devices in the transport network and the remote VTEPs. This is accomplished by advertising it through a dynamic routing protocol in the transport network.
Step 5	host-reachability protocol bgp Example:	Defines BGP as the mechanism for host reachability advertisement.

	Command or Action	Purpose
	<code>switch(config-if-nve)# host-reachability protocol bgp</code>	
Step 6	multisite border-gateway interface loopback <i>vi-num</i> Example: <code>switch(config-if-nve)# multisite border-gateway interface loopback 100</code>	Defines the loopback interface used for the border gateway virtual IP address (VIP). The border-gateway 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 transient devices in the transport network and the remote VTEPs. This is accomplished by advertising it through a dynamic routing protocol in the transport network. This loopback must be different than the source interface loopback. The range of <i>vi-num</i> is from 0 to 1023.
Step 7	no shutdown Example: <code>switch(config-if-nve)# no shutdown</code>	Negate shutdown command.
Step 8	exit Example: <code>switch(config-if-nve)# exit</code>	Exits the NVE configuration mode.
Step 9	interface loopback <i>loopback_number</i> Example: <code>switch(config)# interface loopback 0</code>	Configure the loopback interface.
Step 10	ip address <i>ip-address</i> Example: <code>switch(config-if)# ip address 198.0.2.0/32</code>	Configures the IP address for the loopback interface.

Configuring VNI Dual Mode

This procedure describes the configuration of BUM traffic domain for a given VLAN. Support exists for using multicast or ingress replication inside the fabric/site and Ingress replication across different fabrics/sites.



Note If you have multiple VRFs and only one is extended to ALL leaf switches, you can add a dummy loopback to that one extended VRF and advertise through BGP. Otherwise, you'll need to check how many VRFs are extended and to which switches, and then add a dummy loopback to the respective VRFs and advertise them as well. Therefore, use the **advertise-pip** command to prevent potential user errors in the future.

For more information about configuring the mcast-group (or ingress-replication protocol bgp) for a large number of VNIs, see [Example of VXLAN BGP EVPN \(EBGP\)](#).

Procedure

	Command or Action	Purpose
Step 1	configure terminal Example: switch# configure terminal	Enters global configuration mode.
Step 2	interface nve 1 Example: switch(config)# interface nve 1	Creates a VXLAN overlay interface that terminates VXLAN tunnels. Note Only one NVE interface is allowed on the switch.
Step 3	member vni vni-range Example: switch(config-if-nve)# member vni 200	Configure the virtual network identifier (VNI). The range for <i>vni-range</i> is from 1 to 16,777,214. The value of <i>vni-range</i> can be a single value like 5000 or a range like 5001-5008. Note Enter one of the Step 4 or Step 5 commands.
Step 4	mcast-group ip-addr Example: switch(config-if-nve-vni)# mcast-group 255.0.4.1	Configure the NVE Multicast group IP prefix within the fabric.
Step 5	ingress-replication protocol bgp Example: switch(config-if-nve-vni)# ingress-replication protocol bgp	Enables BGP EVPN with ingress replication for the VNI within the fabric.
Step 6	multisite ingress-replication Example: switch(config-if-nve-vni)# multisite ingress-replication	Defines the Multi-Site BUM replication method. Per-VNI knob for extending Layer 2 VNI.

Configuring Fabric/DCI Link Tracking

This procedure describes the configuration to track all DCI facing interfaces and site internal/fabric facing interfaces. Tracking is mandatory and is used to disable re-origination of EVPN routes either from or to a site if all the DCI/fabric links go down.

Procedure

	Command or Action	Purpose
Step 1	configure terminal Example:	Enters global configuration mode.

	Command or Action	Purpose
	<code>switch# configure terminal</code>	
Step 2	interface ethernet <i>port</i> Example: <code>switch(config)# interface ethernet1/1</code>	Enters interface configuration mode for DCI interface. Note Enter one of the following commands in Step 3 or Step 4.
Step 3	evpn multisite dci-tracking Example: <code>switch(config-if)# evpn multisite dci-tracking</code>	Configure DCI interface tracking.
Step 4	interface ethernet <i>port</i> Example: <code>switch(config)# interface ethernet1/2</code>	Enters interface configuration mode for fabric interface.
Step 5	evpn multisite fabric-tracking Example: <code>switch(config-if)# evpn multisite fabric-tracking</code>	Enters interface configuration mode for fabric interface.
Step 6	ip address <i>ip-addr</i> Example: <code>switch(config-if)# ip address 192.1.1.1</code>	Configure IP features.
Step 7	no shutdown Example: <code>switch(config-if)# no shutdown</code>	Negate shutdown command.

Configuring Fabric External Neighbors

This procedure describes the configuration of Fabric External/DCI Neighbors for communication to other site/fabric border gateways.

Procedure

	Command or Action	Purpose
Step 1	configure terminal Example: <code>switch# configure terminal</code>	Enters global configuration mode.
Step 2	router bgp <i>as-num</i> Example:	Configure the autonomous system number. The range for <i>as-num</i> is from 1 to 4,294,967,295.

	Command or Action	Purpose
	<code>switch(config)# router bgp 100</code>	
Step 3	neighbor <i>ip-addr</i> Example: <code>switch(config-router)# neighbor 100.0.0.1</code>	Configure a BGP neighbor.
Step 4	peer-type fabric-external Example: <code>switch(config-router-neighbor)# peer-type fabric-external</code>	<p>Enables the next hop rewrite for multi-site. Defines site external BGP neighbors for EVPN exchange. The default for peer-type is fabric-internal.</p> <p>Note The peer-type fabric-external command is required only for VXLAN Multi-Site border gateways. It is not required for pseudo border gateways.</p>
Step 5	address-family l2vpn evpn Example: <code>switch(config-router-neighbor)# address-family l2vpn evpn</code>	
Step 6	rewrite-evpn-rt-asn Example: <code>switch(config-router-neighbor)# rewrite-evpn-rt-asn</code>	Rewrites the route target (RT) information to simplify the MAC-VRF and IP-VRF configuration. BGP receives a route, and as it processes the RT attributes, it checks if the AS value matches the peer AS that is sending that route and replaces it. Specifically, this command changes the incoming route target's AS number to match the BGP-configured neighbor's remote AS number. You can see the modified RT value in the receiver router.

