

ACI Border Gateways

Beginning with Release 6.1(1), a new node type ACI Border Gateways (ACI BGWs) is available as a new feature on Cisco ACI.

Note Procedures in this document describe how to configure ACI Border Gateways by using the GUI and REST API. You cannot configure ACI Border Gateways through the NX-OS style CLI at this time.

- About ACI Border Gateways, on page 1
- Understanding ACI Implementation of ACI Border Gateways, on page 2
- Understanding the ACI Border Gateways Deployment, on page 7
- Guidelines and Limitations for ACI Border Gateways, on page 9
- Discovering ACI Border Gateways, on page 10
- Configuring a VXLAN Infra L3Out Using the GUI, on page 13
- VXLAN Site, on page 21
- Creating Border Gateway Sets Using the GUI, on page 22
- Creating Remote VXLAN Fabrics Using the GUI, on page 22
- Configuring a VXLAN VRF Stretch Using the GUI, on page 24
- Configuring a VXLAN Bridge Domain Stretch Using the GUI, on page 25
- VXLAN Stretched Bridge Domain Selector, on page 26
- VXLAN External Subnet Selector, on page 26
- Creating VXLAN Custom QoS Policy Using the GUI, on page 27

About ACI Border Gateways

With the ACI Border Gateways solution you can now have a seamless extension of virtual routing and forwarding (VRF) instance and a bridge domain between the fabrics.

The ACI Border Gateway is the node that interacts with nodes within a site and with nodes that are external to the site. The ACI Border Gateways feature can be conceptualized as multiple site-local EVPN control planes and IP forwarding domains interconnected via a single common EVPN control and forwarding domain.

The Virtual eXtensible Local Area Network (VXLAN) Ethernet Virtual Private Network (EVPN) Border Gateways is a multi-site solution that interconnects two or more BGP-based EVPN sites or fabrics (overlay domains) in a scalable fashion over an IP-only network. It uses border gateways (BGWs) in anycast mode to interconnect a Cisco ACI side with one or more NX-OS sites and allows new approaches to fabric scaling,

compartmentalization, and using DCI. The BGWs provide the network control boundary that is necessary for traffic enforcement and failure containment functionality.

A site-local EVPN domain consists of EVPN nodes with the same site identifier. BGWs on one hand are also part of the site-specific EVPN domain and on the other hand a part of a common EVPN domain to interconnect with BGWs from other sites. For a given site, these BGWs facilitate site-specific nodes to visualize all other sites to be reachable only via them. This means:

- Site-local bridging domains are interconnected only via BGWs with bridging domains from other sites.
- Site-local routing domains are interconnected only via BGWs with routing domains from other sites.

Understanding ACI Implementation of ACI Border Gateways

ACI implements ACI Border Gateways by using the following ACI components that have been introduced in Cisco APIC.

VXLAN Site

Starting from Cisco APIC 6.1(2), you must configure a site ID. You will not be able to configure the border gateway set policy if you do not have this site ID.

Refer to VXLAN Site, on page 21 to know more about creating a site ID for your VXLAN site..



Note If you have already configured the ACI Border Gateway feature for Cisco APIC 6.1(1), and upgrade to Cisco APIC 6.1(2) without creating a VXLAN site ID a fault is generated for all the stretched VRFs and bridge domains.

ACI Border Gateways Set

These are a set of border gateway nodes that are used to connect to the remote VXLAN EVPN fabrics. These BGW nodes could either be part of an ACI pod or be deployed across different pods when the ACI fabric is a multi-pod fabric. All BGWs within a POD as are assigned the same unique anycast TEP per pod to attract traffic for endpoints within this POD from the remote fabric.

Cisco APIC assigns a unique internal anycast TEP for a border gateway set, which is common across all the pods for a border gateway set. You can configure only one border gateway set.

See Creating Border Gateway Sets Using the GUI, on page 22 for more information.

VXLAN Remote Fabric

In the remote fabric configuration, you will specify the remote non-ACI site's loopback IP address on the remote BGW, which is used to establish the MP-BGP EVPN adjacency. You can associate multiple VXLAN remote fabric policies, one for each remote site, with the same border gateway set.

See Creating Remote VXLAN Fabrics Using the GUI, on page 22 for more information.

VXLAN Infra L3Out

VXLAN Infra L3Out defines the group of border gateway nodes and the associated interfaces for the underlay configuration. Only eBGP is supported as the underlay protocol. Additional BFDs can be enabled on the interfaces for faster failure detection and convergence.

See Configuring a VXLAN Infra L3Out Using the GUI, on page 13 for more information.

VXLAN VRF Stretch

In Cisco APIC 6.1(1), to stretch a user VRF, you configure a user VXLAN L3Out that is associated to a border gateway set. You also associate all the remote fabrics to this L3Out to stretch the VRF to the corresponding non-ACI sites. The VRF that is stretched towards the VXLAN fabric can only be in an unenforced mode.

Starting from Cisco APIC 6.1(2), you need to associate only the border gateway set to stretch the VRF to the corresponding non-ACI sites.

To stretch a VRF for the VXLAN remote fabric, see Configuring a VXLAN VRF Stretch Using the GUI, on page 24 for more information.

VXLAN Bridge Domain Stretch

In Cisco APIC 6.1(1), use the VXLAN bridge domain stretch to stretch a bridge domain to a non-ACI fabric. You can associate multiple remote fabrics to stretch the bridge domain to the corresponding non-ACI sites

Starting from Cisco APIC 6.1(2), you need to associate only the border gateway set to stretch the bridge domain to non-ACI sites.

To stretch a bridge domain for the VXLAN remote fabric, see Configuring a VXLAN Bridge Domain Stretch Using the GUI, on page 25 for more information.

VXLAN QoS Policy

You can use custom VXLAN QoS policies to define how traffic coming from the ACI BGWs is prioritized within the ACI fabric. You can also use these policies to re-mark the traffic when it leaves the ACI fabric via the ACI BGW. The custom QoS policy is divided into an ingress QoS policy and an egress QoS policy.

• **Ingress rules:** As part of the ingress VXLAN policy, you can define how the traffic is treated inside the fabric (queuing priority). You can match on the inner dscp range and define the cos and dscp values that should be set in the inner header.



• Egress rules: As part of the egress VXLAN policy, you can control the values that needs to be marked in the outer dscp and cos fields. These values will be matched with the inner dscp values and the outer dscp and cos values will set accordingly. If you do not specify any values, the outer dscp and cos values are set to the default value of zero.



• Forwards Packets: Replace the default VXLAN value (0) to the VXLAN network with the ol_dscp value.

In the event where cos preservation is enabled, the **ol_dscp** value will be the encoded value corresponding to a combination of QoS Level and the **cos** value of the packet when it entered the fabric. This is sent out along with the preserved **cos** value when the packet is exiting the fabric. If you do want to enable cos, it is advisable to have explicit **ol_dscp** remarking enabled via egress VXLAN QoS rules.

VRF in Enforced Mode

Starting with Cisco APIC 6.1(2), the VRFs can be configured in an enforced mode. The endpoints and prefixes that are advertised from the remote VXLAN EVPN fabrics can be classified into endpoint groups that are represented through Endpoint Security Group objects (ESG). To enable communication between endpoints that belong to different ESGs, you need to setup policies represented through contracts.

The following sections discuss the various building blocs that are required to configure a VRF in enforced mode.

Endpoint Security Groups

Endpoint Security Groups (ESGs) are a network security component in Cisco ACI. An ESG is a security construct that has certain match criteria to define which endpoint belongs to the ESG, and uses contracts or policies to define the security criteria. This match criteria is called the ESG selectors that are based on attributes.

ESGs can only communicate with other ESGs according to the contract rules. The administrator uses a contract to select the types of traffic that can pass between ESGs, including the protocols and ports that are allowed.

You must classify ACI endpoints into ESGs by using one of the existing selectors such as EPG selector, IP subnet selector, or so on to associate a contract with the other ESGs to communicate with the remote VXLAN classified end points within an ESG.

Contracts

Contracts are the Cisco ACI equivalent of access control lists (ACLs). Endpoint Security Groups (ESG)s can only communicate with other ESGs according to the contract rules. You can use a contract to select the types of traffic that can pass between ESGs, including the protocols and ports allowed. An ESG can be a provider, consumer, or both provider and consumer of a contract, and can consume multiple contracts simultaneously. ESGs can also be part of a preferred group so that multiple ESGs can talk freely with other ESGs that are part of the preferred group.

EVPN VXLAN Selectors

Selectors are configured under each ESG with a variety of matching criteria to classify endpoints to the ESG. Starting with Cisco APIC 6.1(2), two new selectors have been added to classify remote VXLAN endpoints.

For more information on the existing selectors that are available on Cisco APIC, refer to Endpoint Security Groups section of the Cisco APIC Security Configuration Guide, Release 6.1(x).

These supported selectors are only applicable for remote VXLAN endpoints:

VXLAN Stretched Bridge Domain Selectors

Use this selector to classify all the L2 MAC addresses from the remote VXLAN fabrics into a corresponding ESG. This selector can be configured only for bridge domains that are VXLAN stretched. The endpoints from all the remote fabrics belonging to this bridge domain use the same ESG for classification.

See VXLAN Stretched Bridge Domain Selector, on page 26 for more information.

VXLAN External Subnet Selector

Use this selector to classify all the L3Out prefixes from a remote VXLAN fabric into a corresponding ESG. You cannot have the same prefix configured under an external subnet selector and an external EPG selector under the L3Out. If you have an overlap, the longest prefix match determines the classification of the prefix. You cannot configure a default prefix to match all the prefixes under the external subnet selector. A specific prefix configuration is the preferred approach. As a workaround, **0.0.0/1** or **128.0.0.0/1** can be used if the **Catch All** entry is required.

See VXLAN External Subnet Selector, on page 26 for more information.

Classification of L3 Host Policy

Classification of L3 endpoints (connected hosts) from the remote VXLAN sites use the existing IP subnet selectors. To classify the L3 endpoints, use the L3 subnet that is configured under bridge domain in a remote VXLAN site under IP subnet selector.

You can classify specific L2 MAC addresses and specific L3 IP addresses from the remote VXLAN fabric into an ESG by using the existing MAC tag selector or the IP tag selector.

VXLAN EVPN Route-Maps

Starting from Cisco APIC 6.1(2), the ACI Border Gateway feature also supports VRF level route-maps that can be configured on the stretched VRFs. These Route-maps are applicable for all the remote fabrics that are

associated to the border gateway set. The route-map set rules are configured with the route control profile policies and the action rule profiles.

Use the Configuring a VXLAN VRF Stretch Using the GUI, on page 24 to specify the outbound and inbound route-maps.

For more information on how to configure a route-map, see Configuring Route Control Policy in VRF Using the GUI.



Note This is an optional configuration. However, if you do not configure the import route-maps, all the routes that you receive from remote fabric are accepted and if you do not configure the export route-maps, to all the remote fabrics associated to border gateway set.

Following are the list of match and set clauses that are supported by both the inbound route-map and the outbound route-map:

- Supported Match Clauses
 - IP Prefix List
 - AS-Path
 - Community
 - Extended Community (match on color extended community is not supported)
 - Regex Community
 - Regex Extended Community

Supported Set Clauses

- Community
- Extended Community
- Weight
- Preference
- Metric

Service Graphs

A service graph is a sequence of Layer 4 to Layer 7 services functions, Layer 4 to Layer 7 services devices, or copy devices and their associated configuration. The service graph must be associated with a contract to be "rendered"—or configured—on the Layer 4 to Layer 7 services device or copy device, and on the fabric. While configuring the policy application, ensure that you configure the service devices with the ACI fabric.

Understanding the ACI Border Gateways Deployment

The following figure shows the deployment for the ACI Border Gateways in Cisco ACI.



- VXLAN is used as the overlay technology to encapsulate the data packets and tunnel the traffic over the Layer 3 network.
- VXLAN handoff is through a node role called border gateways via the VXLAN tunnels.
- L2/L3 VXLAN connectivity between Cisco ACI pods that are part of the same fabric is achieved via the spine-to-spine data path, through the IPN.



- · Cisco ACI BGWs must be locally present in each pod.
- For each bridge domain extended across domains, a specific Cisco ACI BGW is elected as the designated forwarder across all the BGWs in all the pods. The designated forwarder BGW will then send and receive flood traffic for that BD with the external domains.
- Different sets of VRFs or bridge domains can be extended between each Cisco ACI fabric and the VXLAN EVPN domain.
- No VXLAN EVPN connectivity between Cisco ACI BGWs of different Cisco ACI fabrics.



Guidelines and Limitations for ACI Border Gateways

Following are the guidelines and limitations for the ACI border gateways feature.

- Hardware support for ACI border gateways are leaf platforms with FX and above with 32GB of RAM.
- Dedicated leaf nodes for the ACI border gateway functionality. Coexistence with border leaf functions (L3Outs) in a border gateway is planned for a future release.
- L2 Multicast traffic forwarded as BUM.
- In Cisco APIC 6.1(1), an unenforced VRF is required on the Cisco ACI fabric for VRFs that need to be stretched.
- Support for a single ACI fabric (can be Multi-Pod).
- In Cisco APIC 6.1(1), you must specify unique values for the ACI border gateway node ID and the NX-OS site ID.



Note This limitation is no longer applicable in Cisco APIC 6.1(2).

Starting with Cisco APIC 6.1(2), the ACI fabric needs to specify a unique site ID for the EVPN domain, which is defined by **vxlanSiteId** configuration.

VNIs for a bridge domain or VRF needs to be symmetrical across Cisco ACI and the NX-OS fabric. As
you cannot control the VNIDs that are assigned by Cisco APIC to the VRFs or bridge domains, initial
support is only available for the VRFs or bridge domains that are stretched from Cisco ACI to VXLAN
EVPN domain to ensure that matching VNIDs can be configured on the remote VXLAN EVPN fabrics.

A future release will introduce support for the namespace normalization function on the Cisco ACI border gateways to ensure asymmetric deployments can also be deployed (for example to be able to stretch VRFs or bridge domains from the VXLAN EVPN domain to ACI).

 We recommend that you isolate VRF in IPN and VXLAN-ISN when the same node is used for both ACI Multi pod or ACI Multi-site and VXLAN inter site. The VXLAN inter site routes should not advertise to ACI spine via Cisco ACI IPN or ISN network. • In Cisco APIC 6.1(1), NX-OS uses the non-VLAN based L3 VNI configuration. This is referred as *new way of VRF configuration*. This is applicable only to the VRFs that are stretched between the ACI and the VXLAN EVPN domains.

Starting with Cisco APIC 6.1(2), the NXOS L3VNI configuration can either be configured with vlan/SVI or without vlan/SVI.

- The ACI border gateway feature requires that all nodes in the fabric be running on Cisco APIC 6.1 (1).
- You must select the same set of spine nodes as the internal route-reflector and the mpod-spine in the given Pod.
- The following features are not supported in this release:
 - SPAN with ACL
 - ERSPAN destination can be an endpoint or prefix that is local to the ACI fabric
 - Multi-site EVPN deployment can either be in full-mesh mode or route-server mode. To integrate with Cisco ACI, it can only done in full-mesh EVPN mode between Cisco ACI and the NX-OS fabric. Hence, the route-server model is not supported
 - The VRF or bridge domains stretched to VXLAN sites should not be deployed on remote leaf switches

Cisco ACI fabric with border gateways can be part of the ACI multisite domain. But the VRFs or bridge domains that are stretched towards the VXLAN EVPN domain cannot be stretched to other ACI Multisite fabrics and vice-versa. Also, the VXLAN stretched VRF or bridge domains cannot be stretched or deployed to the remote leafs

• In Cisco APIC 6.1(1), there is no support for ingress or egress route-maps in EVPN peers in ACI. Any route-filtering can be done only on the remote NX-OS fabric BGW

Starting from Cisco APIC 6.1(2), support has been introduced for route-maps. You can control the routes that can be imported or exported to the remote EVPN fabrics

Inbound route-maps are applied on Type-5 and the IP portion of the Type-2 routes. Type 2 MAC routes are not impacted and are imported irrespective of the IP import status.

Outbound route-maps are only applied to Type-5 routes.

- · IGMP snooping and L3 Multicast traffic is not supported across domains
- Inter-VRFs (shared services) are not supported
- For a VRF stretch, only the L4-L7 service graph is supported with L3/Go-to mode service devices that are deployed on Cisco ACI fabric
- When an ACI fabric interconnects with a policy-aware remote VXLAN fabric, any policy or class
 details from the remote VXLAN fabric are ignored by the ACI border gateway nodes. Similarly,
 the ACI fabric also doesn't advertise its policy or class information to the remote VXLAN fabric.

Discovering ACI Border Gateways

To register a node type as a border-gateway, complete the following steps:

Before you begin

You must register each leaf node with the node type border-gateway for it to be displayed as an ACI border gateway.



Note You cannot register a spine with the node type border-gateway. The discovery will be blocked.

Procedure

- **Step 1** To pre-configure the node registration policy, if you are already aware of the serial number:
 - a) Navigate to Fabric > Inventory > Fabric Membership > Registered Node tab.
 - b) In the Work pane, click Actions > Create Fabric Node Member and complete the following steps.

Figure 1: Discovery of ACI Border Gateways

Create Fabric N	Node Member 🛛 🛛 🔊
Pod ID:	1 🗸 🗸
Serial Number:	12345
Node ID:	101
Switch Name:	VXLANSwitch
Node Type:	Leaf Spine Unknown
Is Remote:	
Is Tier-2 Leaf:	
Is Border Gateway:	
VPC Pair:	101 🗸
VPC Domain ID:	23
4	
	Cancel Submit

- a) In the **Pod ID** field, choose the pod ID from the drop down menu.
- b) In the Serial Number field, enter the serial number for the leaf switch.
- c) In the Node ID field, assign a node ID to the leaf switch.
- d) In the **Switch Name** field, assign a name to the leaf switch.
- e) In the **Node Type** field, select Leaf as the node type.
- f) Put a check in Is Border Gateway check box to register the leaf as a node type.
- g) Click Submit.
- **Step 2** To configure the node based on the DHCP discovery:
 - a) Navigate to Fabric > Inventory > Fabric Membership > Nodes Pending Registration tab.

b) In the **Work** pane, right click the serial number of the newly discovered leaf, click **Register** and complete the following steps.

Figure 2: Discovery of ACI Border Gateways

Register	\bigotimes
Serial Number: S	SAL2028T5PW
Pod ID:	1 🗸 🗸
Node ID:	501
Node Name:	boder-gw1
Role:	Border-Gateway
Rack Name:	select 🗸
	Cancel Register
In the Pod ID field, cho In the Node ID field, as In the Node Name field In the Role field, select	ose the pod ID from the drop down menu. sign a node ID to the leaf switch. l, assign a name to the leaf switch. Border Gateway as the role type

- e) (Optional) In the **Rack Name**, specify the rack name.
- f) Click Register.

a)
b)
c)
d)

What to do next

Create Border Gateway Sets by using the procedures provided in Creating Border Gateway Sets Using the GUI, on page 22

Configuring a VXLAN Infra L3Out Using the GUI

The VXLAN infra L3Out configuration allows you to select the ACI Border Gateway nodes and interfaces to establish EBGP underlay adjacencies with the external network devices. This is required to exchange underlay reachability information with the remote NX-OS Border Gateways and establish the overlay EVPN adjacencies with them.

You will configure the following pieces when configuring the VXLAN infra L3Out:

- Configure the ACI Border Gateway Set. Refer to Creating Border Gateway Sets Using the GUI, on page 22.
- Configure the remote VXLAN fabric. Refer to Creating Remote VXLAN Fabrics Using the GUI, on page 22.
- Nodes
 - Only border gateways are allowed to be configured as nodes in the VXLAN infra L3Out.
 - Each VXLAN infra L3Out can have border gateways from multiple pods that are part of the same ACI multi-pod fabric.
 - The border gateway can either be configured in a single VXLAN infra L3Out or multiple VXLAN infra L3Outs.
 - When you configure a node profile, you can configure the Router ID and the loopback interface underneath the node. The loopback interface is the control plane TEP on a BGW, which is used for the BGP EVPN peering with the VXLAN gateway on the remote fabric.

Interfaces

- Supported types of interfaces are:
 - Routed interface or sub-interface
- You will also configure the underlay BGP peer policy in the interfaces tab in the VXLAN infra L3Out. This is the basic underlay configuration that is needed to bring the BGP underlay to exchange the loopback address to a connected device.

QoS rules

- You can configure the VXLAN ingress rule and VXLAN egress rule through the VXLAN QoS policy in the VXLAN Infra L3Out. Refer to Creating VXLAN Custom QoS Policy Using the GUI, on page 27 for more information.
- If you do not create a VXLAN QoS policy, any ingressing VXLAN traffic is assigned the default QoS level.

You will also configure the underlay and overlay through the VXLAN Infra L3Out:

- Underlay: BGP peer IP configuration as part of the interface configuration.
- Overlay: BGP EVPN remote configuration is part of the remote fabric configuration.

Before you begin

Ensure that you have registered the leaf node as a new node type *border-gateway* for it to be displayed as a VXLAN EVPN border gateway. Refer to Discovering ACI Border Gateways, on page 10 for more information.

Procedure

 Step 1
 Navigate to Tenants > infra > Networking > VXLAN L3Outs.

 Step 2
 Right-click on VXLAN L3Outs and choose Create VXLAN L3Out.

 The Connectivity window appears.

Create VXLAN Infra L3Out



Connectivity

The creation of an VXLAN Infra Layer 3 outside (L3Out) is required to enable the VXLAN handoff from ACI.

Underlay Configuration

The underlay Configuration used is BGP.

Name:	I3out	Layer 3 Domain:		I3out_domain	
VXLAN Custom QoS Policy:	default	~ 🗳			

Step 3 In the **Connectivity** window, enter the necessary information.

a) In the Name field, enter a name for the VXLAN Infra L3Out.

This will be the name for the policy controlling connectivity to the outside. The name can be up to 64 alphanumeric characters.

Note

You cannot change this name after the object has been saved.

b) (Optional) In the VXLAN Custom QoS Policy field, choose an existing QoS policy or choose Create VXLAN Custom QoS Policy to create a new QoS policy.

For more information on creating a new QoS policy, see Creating VXLAN Custom QoS Policy Using the GUI, on page 27.

c) Click Next.

The Nodes and Interfaces window appears.

Create VXLAN Infra L3Out

1. Connectivit

Nodes and Interfaces

Select the Border Gateway nodes (BGWs) for the VXLAN configuration.

Node Profile Name:	3out_nodePr	ofile					
Interface Profile Name:	3out_interfac	ceProfile					
BFD Interface Policy:	select a value)		\sim			
Interface Types							
Layer 3:	Interface	Sub-In	terface				
Layer 2:	Port	Direct Port	Channel				
Nodes							
Node ID		Router ID		Loopback	Address		
nxbgp6-leaf1 (Node	-101) 🗸	8.8.8.8		8.8.8.8		\blacksquare $+$	Hide Interfaces
Interface	MTU	(bytes)	IPv4 Addr	ess	Peer IPv4 A	ddress	Remote ASN
eth1/5	~ 900	0	25.1.1.1/24		25.1.1.2		200
Ex: eth1/1 or topology/po 1/paths-101/pathep-[eth	d- 1/23]		address/ma	sk			

Step 4 In the Nodes and Interfaces window, enter the necessary information to configure the border gateway nodes and interfaces.

a) In the **Node Profile Name** and **Interface Profile Name** fields, determine if you want to use the default naming convention for the node profile and interface profile names.

The default node profile name is L3Out-name_nodeProfile, and the default interface profile name is L3Out-name_interfaceProfile, where L3Out-name is the name that you entered in the Name field in the Connectivity page. Change the profile names in these fields, if necessary.

- b) (Optional) In the **BFD Interface Policy** field, choose an existing BFD interface policy or choose **Create BFD Interface Policy** to create a new BFD interface policy.
- c) In the Interface Types area, make the necessary selections in the Layer 3 and Layer 2 fields.

The options are:

• Layer 3:

- **Interface**: Choose this option to configure a Layer 3 interface to connect the border leaf switch to the external router.
- **Sub-Interface**: Choose this option to configure a Layer 3 sub-interface to connect the border leaf switch to the external router.

• Layer 2:

- Port Layer 2 can either be a port or a port channel. Cisco APIC 6.1(1) only supports port.
- d) From the Node ID field drop-down menu, choose the border gateway node for the VXLAN infra L3Out..

You might see the following warning message appear on your screen, describing how to configure the router ID.

The leaf switch 103 has a Operational Router ID 3.3.3.3 which is used for MP-BGP sessions running between this leaf and spines. User can still configure a different Route ID than 3.3.3.3 but will flap the MP-BGP sessions which are already running on this leaf.

- If you do not have a router ID already configured for this node, go to 4.e, on page 18 for instructions on configuring a router ID for this node.
- If you have a router ID already configured for this node (for example, if you had configured MP-BGP route reflectors previously).

Use the same router ID for the VXLAN configuration: The same router ID must be used across the VXLAN infra L3Out configuration. This is the recommended option. Make a note of the router ID displayed in this warning to use in the next step, 4.e, on page 18 for instructions on configuring a router ID for this node.

e) In the **Router ID** field, enter a unique router ID (the IPv4 address) for the border leaf switch part of the infra L3Out.

The router ID must be unique across all border leaf switches and the non-ACI fabric BGWs.

As described in 4.d, on page 18, if a router ID has already been configured on this node, you have several options:

- If you want to use the same router ID for the VXLAN configuration, enter the router ID that was displayed in the warning message in 4.d, on page 18.
- You must configure the same router ID across all infra L3Outs for a given node.
- f) Enter an IP address in the Loopback field. This is the routable control plane TEP address which is used for EVPN peering, It is advertised via the underlay protocol.
- g) In the the Interface field, choose a port from the drop-down list.

- h) If you selected **Sub-Interface** in the Layer 3 area above, the **VLAN Encap** field appears. Enter the encapsulation used for the layer 3 outside profile.
- i) In the MTU (bytes) field, enter the maximum transmit unit of the external network.

Acceptable entries in this field are from 576-9216. To inherit the value, enter inherit in this field.

j) In the IPv4 Address field, enter an IP address for the eBGP underlay configuration.

This is the IP address assigned to the Layer 3 interface/sub-interface that you configured in the previous step.

k) In the Peer IPv4 Address field, enter the eBGP underlay unicast peer IP address.

This is the interface's IP address of the router directly connected to the border leaf switch.

- 1) In the Remote ASN field, enter the BGP Autonomous System Number of the directly-connected router.
- m) Determine if you want to configure additional interfaces for this node for the VXLAN infra L3Out.
 - If you do not want to configure additional interfaces for this node for this VXLAN infra L3Out, skip to 4.0, on page 19.
 - If you want to configure additional interfaces for this node for this VXLAN infra L3Out, click + in the **Interfaces** area to bring up the same options for another interface for this node.

Note

If you want to delete the information that you entered for an interface for this node, or if you want to delete an interface row that you added by accident, click the trash can icon for the interface row that you want to delete.

n) Determine if you want to configure additional border gateways for this VXLAN infra L3Out.

• If you do not want to configure additional border gateways for this VXLAN infra L3Out, skip to 4.0, on page 19.

• If you want to configure additional border gateways for this VXLAN infra L3Out, click + in the **Nodes** area to bring up the same options for another node.

Note

If you want to delete the information that you entered for a node, or if you want to delete a node row that you added by accident, click the trash can icon for the node row that you want to delete.

o) Click Next.

The **Policy Configuration** window appears.

Create VXLAN Infra L3Out

					1. Connec
order Gateway Set					
Policy:					
Bgwset	· @				
amote VXI AN Eabric					
ANOLE VALAIN FADIL					
Remote Fabric Name:					
Fabric1	-				
Remote Data Plane TEP Address					
5.1.1.1/32	-				
IPv4 address/mask					
Remote EVPN Peer Address	Remote AS		TTL		
40.1.1.1	250	\bigcirc	5	\bigcirc	- $+$
IPv4 address		-			

Step 5 In the **Policy Configuration** window, enter the necessary information to configure the border gateway nodes and interfaces.

- a) In the Border Gateway Set field, determine if you want to use an existing border gateway set or create a new border gateway set.
- b) Check the Configure VXLAN Remote Fabrics and configure the following fields:
 - 1. In the **Remote VXLAN Fabric** field, specify an existing remote VXLAN fabric or click + to create a new remote VXLAN fabric.
 - 2. In the Remote EVPN Peer Address field, specify the remote EVPN address.
 - 3. In the **Remote AS** field, enter the BGP autonomous system number of the BGP ASN of the remote NX-OS BGW node to configure the remote AS for each remote fabric peer.
 - 4. In the **TTL** field, enter the connection time to live (TTL). The value must be greater than 1.
- Step 6 Click Finish to complete the necessary configurations in the Create VXLAN Infra L3Out wizard.

What to do next

Configure an VXLAN VRF Stretch using the procedures provided in Configuring a VXLAN VRF Stretch Using the GUI, on page 24.

VXLAN Site

Starting from Cisco APIC 6.1(2), you must configure a site ID. You will not be able to configure the border gateway set policy if you do not have this site ID.



Note

If you have already configured the ACI Border Gateway feature for Cisco APIC 6.1(1), and upgrade to Cisco APIC 6.1(2) without creating a VXLAN site ID a fault is generated for all the stretched VRFs and bridge domains.

Procedure

Step 1 Step 2	From the top menu bar, navigate to Tenants > infra >Policies >VXLAN Gateway >VXLAN Site . Right-click on VXLAN Site and select Create VXLAN Site . The Create VXLAN Site window appears.
Step 3	In the Name field, enter a name for your VXLAN site.
Step 4	In the ID field, enter a unique site ID for the VXLAN Site.
Step 5	(Optional) In the Description field, enter a description for the VXLAN Site.
Step 6	Click Submit.

Creating Border Gateway Sets Using the GUI

To create border gateway sets, complete the following procedure:

Before you begin

This policy assigns a data plane TEP for border gateways in each POD, which is used to communicate with remote non-ACI fabrics. This is the external anycast TEP for the POD. Cisco APIC also allocates one internal anycast TEP for all the border gateways within the fabric.

Procedure

Step 1	From the top menu bar, navigate to Tenants > infra >Policies >VXLAN Gateway >Border Gateway Sets.
Step 2	On the Border Gateway Set work pane, click Actions > Create Border Gateway Set Policy.
Step 3	In the VXLAN Site ID field, enter a unique site ID for the Border Gateway Set Policy.
Step 4	In the Name field, assign a name to the Border Gateway Set Policy.
Step 5	In the External Data Plane IP field, enter the address for each POD. Click + to enter the POD ID and the Address.
Step 6	Click Submit .

What to do next

Create Remote VXLAN fabrics by using the procedures provided in Creating Remote VXLAN Fabrics Using the GUI, on page 22.

Creating Remote VXLAN Fabrics Using the GUI

To create remote VXLAN fabrics, complete the following procedure:

Before you begin

This policy represents a unique remote non-ACI fabric and the configuration specific to this fabric. The remote fabric policy provides the control plane peering connectivity on the associated border gateway set for a remote fabric.

Procedure

- Step 1 From the top menu bar, navigate to Tenants > infra >Policies >VXLAN Gateway > Remote VXLAN Fabrics.
- Step 2 On the Remote VXLAN Fabrics work pane, click Actions > Create Remote VXLAN Fabric.
- **Step 3** In the Name field, assign a name to the remote VXLAN fabric.
- **Step 4** To enter the Peer IP Address and its associated TTL, Click + in the Remote EVPN Peers section, and complete the following steps in the **Create Remote EVPN Peer** dialog box:

Note

For a infra peer TTL, you must specify a value greater than 1.

- a) **Peer Address**: Enter the peer IP address. This is the loopback IP address of the remote NX-OS BGW device, which is used to establish the EVPN control-plane adjacency.
- b) (Optional) In the **Description** field, enter descriptive information about the remote EVPN policy.
- c) **Remote ASN**: Enter a number that uniquely identifies the neighbor autonomous system. The Autonomous System Number can be in 4-byte as plain format from 1 to 4294967295.

Note

ACI does not support asdot or asdot+ format AS numbers.

- d) In the Admin State field, select Enabled or Disabled to enable or disable Remote EVPN Peer for this particular policy.
- e) In the **BGP Controls** field, check the desired controls.

The peer controls specify which Border Gateway Protocol (BGP) attributes are sent to a peer. The peer control options are:

- Allow Self AS: Enables the autonomous number check on itself. This allows BGP peer to inject updates if the same AS number is being used.
- Disable Peer AS Check: Disables the peer autonomous number check. When the check box is checked, if the advertising router finds the AS number of the receiver in the AS path, it will not send the route to the receiver.
- f) In the **Peer Type** field, the VXLAN BGW Connectivity is already selected.
- g) (Optional) In the Password and Confirm Password field, enter the administrative password.
- h) In the **TTL** field, enter the connection time to live (TTL).

The range is from 2 to 255 hops.

i) In the **BGP Peer Prefix Policy** field, select an existing peer prefix policy or create a new one.

The peer prefix policy defines how many prefixes can be received from a neighbor and the action to take when the number of allowed prefixes is exceeded. This feature is commonly used for external BGP peers, but can also be applied to internal BGP peers.

j) In the Local-AS Number Config field, choose the local Autonomous System Number (ASN) configuration.

When you configure the local ASN in the Cisco ACI fabric, the Cisco ACI BGWs still derive the bridge domain and VRF route targets by using the fabric ASN value. If the peer-ASN value differs from the ASN value in the received route targets with EVPN routes, the EVPN route targets rewrite will not work on the remote VXLAN fabric BGWs. To resolve this, you must manually configure the route targets to match the Cisco ACI derived route targets based on the fabric ASN value for both the bridge domain and the VRF.

Using a local AS number rather than the Global AS permits the routing devices in the associated network to appear to belong to the former AS. The configuration can be:

 no-Prepend+replace-as+dual-as—Does not allow prepending on local AS and is replaced with both AS numbers.

Note

You can prepend one or more autonomous system (AS) numbers at the beginning of an AS path. The AS numbers are added at the beginning of the path after the actual AS number from which the route originates has been added to the path. Prepending an AS path makes a shorter AS path look longer and therefore less preferable to BGP.

• no-prepend—Does not allow prepending on local AS.

- no options—Does not allow alteration of local AS.
- no-Prepend+replace-as—Does not allow prepending on local AS and is replaces AS number.
- k) In the Local-AS Number field, choose the desired value.

Optionally required for the local autonomous system feature for eBGP peers. The local Autonomous System Number can be in 4-byte asplain format from 1 to 4294967295.

Note

ACI does not support asdot or asdot+ format AS numbers.

- l) Click OK.
- **Step 5** To enter the **Associated Border Gateway Set**, select an existing border gateway set from the drop down list or click + in the Associated Border Gateway Set box and select an existing border gateway set.

Step 6 Click Submit.

What to do next

Configure an VXLAN Infra L3Out by using the procedures provided in the Configuring a VXLAN Infra L3Out Using the GUI, on page 13 section.

Configuring a VXLAN VRF Stretch Using the GUI

Using the procedure in this section, you can stretch tenant VRF(s) between the ACI and VXLAN EVPN domains. This ensures that routed communications for the tenants between those domains can happen by leveraging the VXLAN data-plane encapsulation. Some specific deployment considerations when stretching a tenant VRF are:

- User tenant VRFs that are stretched are associated to a BGW set, which are associated to the VXLAN infra L3Out.
- Only one VXLAN VRF L3Out is supported on each VRF. This is used to stretch the VRF towards a BGW.

Before you begin

- Review the Guidelines and Limitations for ACI Border Gateways, on page 9.
- Configure the VXLAN Gateway infra L3Out using the procedures provided in Configuring a VXLAN Infra L3Out Using the GUI.

Procedure

Step 1Navigate to the Tenants > Networking > VXLAN Stretch.Step 2Right-click on VXLAN Stretch and select Create VXLAN VRF Stretch.

The Create VXLAN VRF Stretch window appears.

Step 3 In the **VRF** field, select an existing VRF or click **Create VRF** to create a new VRF with the following steps:

- a) In the Name field, enter a name for the VRF.
- b) In the Alias field, enter an alias name for the VRF.
- c) (Optional) In the **Description**field, enter a description of the VRF.
- d) In the Policy Control Enforcement Preference field, choose Unenforced.
- e) In the Policy Control Enforcement Direction field, choose Ingress.
- f) In the **OSPF Timers** field, from the drop down list, choose the OSPF timer policy that you want to associate with this specific VRF (default or Create OSPF Timers Policy).
- g) In the **Monitoring Policy** field, from the drop down list, choose the Monitoring policy that you want to associate with this specific VRF.
- h) Click Submit.
- **Step 4** In the **Border Gateway Set** field, select an existing border gateway set or click **Create Border Gateway Set** to create a new border gateway set.
- **Step 5** Navigate to the **Configured Remote VNI** area and, in the **Configured Remote VNI** area, complete the following procedure.
 - a) In the Remote Fabric Name field, the option All select a remote fabric name.
 - b) In the Remote VNI field, the option Symmetric is already selected.

Note

Starting from Cisco APIC 6.1(2), you do not have to specify the **Remote Fabric Name** or the **Remote VNI** as these options have been preselected.

Step 6 In the **Outbound** field, specify an outbound route map to control the routes that are advertised to the NXOS site.

For more information on how to create Route Maps, refer to Route Control Profile Policies.

Step 7 In the **Inbound** field, specify the inbound route map to control the routes that are imported from the NXOS fabric.

For more information on how to create Route Maps, refer to Route Control Profile Policies.

Step 8 Click Submit.

What to do next

Configure a VXLAN bridge domain stretch using the procedures provided in Configuring a VXLAN Bridge Domain Stretch Using the GUI, on page 25.

Configuring a VXLAN Bridge Domain Stretch Using the GUI

Using the procedure in this section, you can stretch tenant bridge domain (s) between the ACI and VXLAN EVPN domains. This ensures that bridged communications for the tenants between those domains can happen by leveraging VXLAN data-plane encapsulation

Before you begin

• Review the Guidelines and Limitations for ACI Border Gateways, on page 9.

• Configure the VXLAN Gateway infra L3Out using the procedures provided in Configuring a VXLAN Infra L3Out Using the GUI.

Procedure

Step 1 Step 2	Navigate to Tenants > Networking > VXLAN Stretch . Right-click on VXLAN Stretch and select Create VXLAN BD Stretch .
	The Create VXLAN BD Stretch window appears.
Step 3	In the Bridge Domain field, select an existing bridge domain or click Create Bridge Domain to create a new bridge domain.
Step 4	In the Border Gateway Set field, select an existing border gateway set. As mentioned on the text box in the GUI, ensure that L2 Unknown Unicast is set to flood for the bridge domain that is stretched.
Step 5	Navigate to the Configured Remote VNI area and, in the Configured Remote VNI area, complete the following procedure.
	a) In the Remote Fabric Name field, the option All is already selected as this release supports all remote fabrics.
	b) In the Remote VNI field, the option Symmetric is already selected as this release only supports symmetric namespaces in the ACI and VXLAN EVPN domains.
Step 6	Click Submit .

VXLAN Stretched Bridge Domain Selector

Use this procedure to create a VXLAN stretched bridge domain selector .

Procedure

Step 1	On the menu bar, choose Tenants and select the applicable Tenant.
Step 2	In the Navigation pane, expand tenant_name > Application Profiles > application_profile_name > Endpoint Security Groups > esg_name > Selectors.
Step 3	Right click VXLAN BD Selector and select Create a VXLAN BD Selector.
Step 4	In the Create a VXLAN BD Selector dialog box, enter the following information:
	a) Bridge Domain: From the drop down, select the stretched bridge domain to be mapped.

- b) Description: (Optional) A description of the selector.
- c) Click Submit.

VXLAN External Subnet Selector

Use this procedure to create a VXLAN external subnet selector .

Procedure

Step 1	On the menu bar, choose Tenants and select the applicable Tenant.
Step 2	In the Navigation pane, expand tenant_name > Application Profiles > application_profile_name > Endpoint Security C roups > asg_name > Selectors
Step 3	Right click VXLAN External Subnet Selectors and select Create a VXLAN External Subnet Selector.
Step 4	In the Create a VXLAN External Subnet Selector dialog box, enter the following information:
	a) IP : Specify the IP prefix to be matched.
	1) Descriptions (Outline) A description of the selector

- b) **Description**: (Optional) A description of the selector.
- c) Click Submit.

Creating VXLAN Custom QoS Policy Using the GUI

VXLAN custom QoS policy defines the priority of the packets coming from a VXLAN EVPN fabric while they are inside the ACI fabric based on the incoming values defined in the VXLAN QoS ingress policy. These COS/DSCP values are set in the inner header. It also marks the COS and DSCP values of the packets leaving the ACI fabric toward a remote VXLAN EVPN fabric based on IPv4 DSCP values that are defined in VXLAN QoS egress policy. If no custom egress policy is defined, the outer dscp and cos values are set to the default value of zero before leaving the ACI fabric.

Procedure

Step 1	Fre	om the top menu bar, navigate to Tenants > infra > Networking > VXLAN L3Outs.
Step 2	Ri	ght-click on VXLAN L3Outs and choose Create VXLAN L3Out.
Step 3	In	the Connectivity window, enter the necessary information.
Step 4	In to	the VXLAN Custom QoS Policy field, choose an existing QoS policy or choose Create VXLAN Custom QoS Policy create a new QoS policy.
Step 5	In cre	the Create VXLAN Custom QoS Policy window that opens, provide the name and description of the policy you're eating.
Step 6	In	the VXLAN Ingress Rule area, click + to add an ingress QoS translation rule.
	Da if a	ta traffic coming into the border gateway connected to the ACI fabric will be checked for the inner DSCP value and a match is found, the traffic is classified into an ACI QoS Level and marked with appropriate COS and DSCP values.
	a)	In the Priority field, select the priority for the ingress rule.
		This is the QoS level you want to assign for the traffic within ACI fabric, which ACI uses to prioritize the traffic within the fabric. The options range from Level 1 to Level 6. The default value is Level 3. If you do not make a selection in this field, the traffic will automatically be assigned a Level 3 priority.
	b)	Using the DSCP Range From and DSCP Range To dropdowns, specify the DSCP range of the ingressing VXLAN packet that you want to match.
	c)	Use the Target DSCP to select the inner DSCP value to assign to the packet when it's inside the ACI fabric.

d) In the Target COS field, select the COS value to assign to the packet when it's inside the ACI fabric.

The COS value specified is set in the original traffic received from the external network, so it will be re-exposed only when the traffic is VXLAN decapsulated on the destination ACI leaf node.

The default is Unspecified, which means that the original COS value of the packet will be retained, but only if the COS preservation option is enabled in the fabric.

- e) Click Update to save the ingress rule.
- f) Repeat this step for any additional ingress QOS policy rules.
- **Step 7** In the **VXLAN Egress Rule** area, click + to add an egress QOS translation rule.
 - a) Using the **DSCP Range From** and **DSCP Range To** dropdowns, specify the DSCP range of the ACI fabric packet you want to match for assigning the egressing VXLAN's packet's priority.
 - b) From the **Target Overlay DSCP** dropdown, select the outer dscp value that you want to assign to the egressing VXLAN packet.
 - c) From the Target COS dropdown, select the outer COS value that you want to assign to the egressing VXLAN packet.
 - d) Click Update to save the ingress rule.
 - e) Repeat this step for any additional egress QoS policy rules.
- **Step 8** Click **OK** to complete the creation of the custom VXLAN QoS Policy.