

## **Bridging**

This chapter contains the following sections:

- Bridged Interface to an External Router, on page 1
- Bridge Domains and Subnets, on page 2
- Creating a Tenant, VRF, and Bridge Domain Using the GUI, on page 6
- Creating a Tenant, VRF, and Bridge Domain Using the NX-OS Style CLI, on page 7
- Creating a Tenant, VRF, and Bridge Domain Using the REST API, on page 8
- Configuring an Enforced Bridge Domain, on page 9
- Configuring Flood in Encapsulation for All Protocols and Proxy ARP Across Encapsulations, on page
   12

## **Bridged Interface to an External Router**

As shown in the figure below, when the leaf switch interface is configured as a bridged interface, the default gateway for the tenant VNID is the external router.

Fabric

VNID 5

12.1.1.100
12.1.1.101
12.1.1.102
12.1.1.1/24

EP EP EP

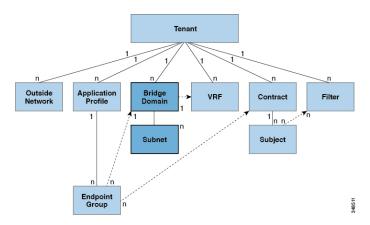
Figure 1: Bridged External Router

The ACI fabric is unaware of the presence of the external router and the APIC statically assigns the leaf switch interface to its EPG.

## **Bridge Domains and Subnets**

A bridge domain (fvBD) represents a Layer 2 forwarding construct within the fabric. The following figure shows the location of bridge domains (BDs) in the management information tree (MIT) and their relation to other objects in the tenant.

Figure 2: Bridge Domains



A BD must be linked to a VRF (also known as a context or private network). With the exception of a Layer 2 VLAN, it must have at least one subnet (fvSubnet) associated with it. The BD defines the unique Layer 2 MAC address space and a Layer 2 flood domain if such flooding is enabled. While a VRF defines a unique IP address space, that address space can consist of multiple subnets. Those subnets are defined in one or more BDs that reference the corresponding VRF.

The options for a subnet under a BD or under an EPG are as follows:

- Public—the subnet can be exported to a routed connection.
- Private—the subnet applies only within its tenant.
- Shared—the subnet can be shared with and exported to multiple VRFs in the same tenant or across tenants as part of a shared service. An example of a shared service is a routed connection to an EPG present in another VRF in a different tenant. This enables traffic to pass in both directions across VRFs. An EPG that provides a shared service must have its subnet configured under that EPG (not under a BD), and its scope must be set to advertised externally, and shared between VRFs.



Note

Shared subnets must be unique across the VRF involved in the communication. When a subnet under an EPG provides a Layer 3 external network shared service, such a subnet must be globally unique within the entire ACI fabric.

BD packet behavior can be controlled in the following ways:

Packet Type	Mode	
ARP	You can enable or disable <b>ARP Flooding</b> ; without flooding, ARP packets are sent with unicast.	
	Note  If the limitIpLearnToSubnets in fvBD is set, endpoint learning is limited to the BD only if the IP address is in a configured subnet of the BD or an EPG subnet that is a shared service provider.	

Packet Type	Mode		
Unknown Unicast	L2 Unknown Unicast, which can be Flood or Hardware Proxy.		
	When the BD has L2 Unknown Unicast set to Flood, if an endpoint is deleted the system deletes it from both the local leaf switches as well as the remote leaf switches where the BD is deployed, by selecting Clear Remote MAC Entries. Without this feature, the remote leaf continues to have this endpoint learned until the timer expires.		
	Modifying the <b>L2 Unknown Unicast</b> setting causes traffic to bounce (go down and up) on interfaces to devices attached to EPGs associated with this bridge domain.		
Unknown IP Multicast	L3 Unknown Multicast Flooding		
	<b>Flood</b> —Packets are flooded on ingress and border leaf switch nodes only. With N9K-93180YC-EX, packets are flooded on all the nodes where a bridge domain is deployed.		
	<b>Optimized</b> —Only 50 bridge domains per leaf are supported. This limitation is not applicable for N9K-93180YC-EX.		
L2 Multicast, Broadcast, Unicast	Multi-Destination Flooding, which can be one of the following:		
	• Flood in BD—flood in bridge domain		
	• Flood in Encapsulation—flood in encapsulation		
	• <b>Drop</b> —drop the packets		



Note

Beginning with Cisco APIC Release 3.1(1), on the Cisco Nexus 9000 series switches (with names ending with EX and FX and onwards), the following protocols can be flooded in encapsulation or flooded in a bridge domain: OSPF/OSPFv3, BGP, EIGRP, CDP, LACP, LLDP, ISIS, IGMP, PIM, ST-BPDU, ARP/GARP, RARP, ND.

Bridge domains can span multiple switches. A bridge domain can contain multiple subnets, but a subnet is contained within a single bridge domain. If the bridge domain (fvBD) limitIPLearnToSubnets property is set to yes, endpoint learning will occur in the bridge domain only if the IP address is within any of the configured subnets for the bridge domain or within an EPG subnet when the EPG is a shared service provider. Subnets can span multiple EPGs; one or more EPGs can be associated with one bridge domain or subnet. In hardware proxy mode, ARP traffic is forwarded to an endpoint in a different bridge domain when that endpoint has been learned as part of the Layer 3 lookup operation.

### **Bridge Domain Options**

A bridge domain can be set to operate in flood mode for unknown unicast frames or in an optimized mode that eliminates flooding for these frames. When operating in flood mode, Layer 2 unknown unicast traffic is flooded over the multicast tree of the bridge domain (GIPo). For the bridge domain to operate in optimized mode you should set it to hardware-proxy. In this case, Layer 2 unknown unicast frames are sent to the spine-proxy anycast VTEP address.



#### Caution

Changing from unknown unicast flooding mode to hw-proxy mode is disruptive to the traffic in the bridge domain.

If IP routing is enabled in the bridge domain, the mapping database learns the IP address of the endpoints in addition to the MAC address.

The **Layer 3 Configurations** tab of the bridge domain panel allows the administrator to configure the following parameters:

- Unicast Routing: If this setting is enabled and a subnet address is configured, the fabric provides the default gateway function and routes the traffic. Enabling unicast routing also instructs the mapping database to learn the endpoint IP-to-VTEP mapping for this bridge domain. The IP learning is not dependent upon having a subnet configured under the bridge domain.
- Subnet Address: This option configures the SVI IP addresses (default gateway) for the bridge domain.
- Limit IP Learning to Subnet: This option is similar to a unicast reverse-forwarding-path check. If this option is selected, the fabric will not learn IP addresses from a subnet other than the one configured on the bridge domain.



#### Caution

Enabling **Limit IP Learning to Subnet** is disruptive to the traffic in the bridge domain.

#### Disabling IP Learning per Bridge Domain

IP learning per bridge domain is disabled when two hosts are connected as active and standby hosts to the Cisco ACI switches. The MAC learning still occurs in the hardware but the IP learning only occurs from the ARP/GARP/ND processes. This functionality allows for flexible deployments, for example, firewalls or local gateways.

See the following guidelines and limitations for disabling IP learning per bridge domain:

- Layer 3 multicast is not supported because the source IP address is not learned to populate the S,G information in the remote top-of-rack (ToR) switches.
- As the DL bit is set in the iVXLAN header, the MAC address is also not learned from the data path in
  the remote TORs. It results in flooding of the unknown unicast traffic from the remote TOR to all TORs
  in the fabric where this BD is deployed. It is recommended to configure the BD in proxy mode to overcome
  this situation if endpoint dataplane learning is disabled.
- ARP should be in flood mode and GARP based detection should be enabled.

- When IP learning is disabled, Layer 3 endpoints are not flushed in the corresponding VRF. It may lead
  to the endpoints pointing to the same TOR forever. To resolve this issue, flush all the remote IP endpoints
  in this VRF on all TORs.
- On Cisco ACI switches with Application Leaf Engine (ALE), the inner MAC address is not learned from the VXLAN packets.
- When dataplane learning is disabled on a BD, the existing local endpoints learned via dataplane in that BD are not flushed. If the data traffic is flowing, the existing local endpoints do not age out.

When IP learning is disabled, you have to enable the **Global Subnet Prefix** check option in **System > System Settings > Fabric Wide Setting > Enforce Subnet Check** in the Online Help.

## Creating a Tenant, VRF, and Bridge Domain Using the GUI

If you have a public subnet when you configure the routed outside, you must associate the bridge domain with the outside configuration.

#### **Procedure**

- Step 1 On the menu bar, choose Tenants > Add Tenant.
- **Step 2** In the **Create Tenant** dialog box, perform the following tasks:
  - a) In the **Name** field, enter a name.
  - b) Click the **Security Domains** + icon to open the **Create Security Domain** dialog box.
  - c) In the Name field, enter a name for the security domain. Click Submit.
  - d) In the Create Tenant dialog box, check the check box for the security domain that you created, and click Submit.
- Step 3 In the Navigation pane, expand Tenant-name > Networking, and in the Work pane, drag the VRF icon to the canvas to open the Create VRF dialog box, and perform the following tasks:
  - a) In the **Name** field, enter a name.
  - b) Click **Submit** to complete the VRF configuration.
- Step 4 In the Networking pane, drag the BD icon to the canvas while connecting it to the VRF icon. In the Create Bridge Domain dialog box that displays, perform the following tasks:
  - a) In the **Name** field, enter a name.
  - b) Click the L3 Configurations tab.
  - c) Expand Subnets to open the Create Subnet dialog box, enter the subnet mask in the Gateway IP field and click OK.
  - d) Click **Submit** to complete bridge domain configuration.
- Step 5 In the Networks pane, drag the L3 icon down to the canvas while connecting it to the VRF icon. In the Create Routed Outside dialog box that displays, perform the following tasks:
  - a) In the Name field, enter a name.
  - b) Expand Nodes And Interfaces Protocol Profiles to open the Create Node Profile dialog box.
  - c) In the **Name** field, enter a name.
  - d) Expand **Nodes** to open the **Select Node** dialog box.
  - e) In the **Node ID** field, choose a node from the drop-down list.

- f) In the **Router ID** field, enter the router ID.
- g) Expand Static Routes to open the Create Static Route dialog box.
- h) In the **Prefix** field, enter the IPv4 or IPv6 address.
- i) Expand **Next Hop Addresses** and in the **Next Hop IP** field, enter the IPv4 or IPv6 address.
- j) In the **Preference** field, enter a number, then click **UPDATE** and then **OK**.
- k) In the **Select Node** dialog box, click **OK**.
- 1) In the Create Node Profile dialog box, click OK.
- m) Check the **BGP**, **OSPF**, or **EIGRP** check boxes if desired, and click **NEXT**. Click **OK** to complete the Layer 3 configuration.

To confirm L3 configuration, in the **Navigation** pane, expand **Networking** > **VRFs**.

# Creating a Tenant, VRF, and Bridge Domain Using the NX-OS Style CLI

This section provides information on how to create tenants, VRFs, and bridge domains.



Note

Before creating the tenant configuration, you must create a VLAN domain using the **vlan-domain** command and assign the ports to it.

#### **Procedure**

Step 1 Create a VLAN domain (which contains a set of VLANs that are allowable in a set of ports) and allocate VLAN inputs, as follows:

#### Example:

In the following example ("exampleCorp"), note that VLANs 50 - 500 are allocated.

```
apic1# configure
apic1(config)# vlan-domain dom_exampleCorp
apic1(config-vlan)# vlan 50-500
apic1(config-vlan)# exit
```

Step 2 Once the VLANs have been allocated, specify the leaf (switch) and interface for which these VLANs can be used. Then, enter "vlan-domain member" and then the name of the domain you just created.

#### **Example:**

In the following example, these VLANs (50 - 500) have been enabled on leaf 101 on interface ethernet 1/2-4 (three ports including 1/2, 1/3, and 1/4). This means that if you are using this interface, you can use VLANS 50-500 on this port for any application that the VLAN can be used for.

```
apicl(config-vlan)# leaf 101
apicl(config-vlan)# interface ethernet 1/2-4
apicl(config-leaf-if)# vlan-domain member dom_exampleCorp
apicl(config-leaf-if)# exit
apicl(config-leaf)# exit
```

**Step 3** Create a tenant in global configuration mode, as shown in the following example:

#### **Example:**

```
apic1(config) # tenant exampleCorp
```

**Step 4** Create a private network (also called VRF) in tenant configuration mode as shown in the following example:

#### **Example:**

```
apicl(config) # tenant exampleCorp
apicl(config-tenant) # vrf context exampleCorp_v1
apicl(config-tenant-vrf) # exit
```

**Step 5** Create a bridge domain (BD) under the tenant, as shown in the following example:

#### **Example:**

```
apicl(config-tenant)# bridge-domain exampleCorp_b1
apicl(config-tenant-bd)# vrf member exampleCorp_v1
apicl(config-tenant-bd)# exit
```

**Note** In this case, the VRF is "exampleCorp\_v1".

**Step 6** Allocate IP addresses for the BD (ip and ipv6), as shown in the following example.

#### **Example:**

```
apicl(config-tenant) # interface bridge-domain exampleCorp_b1
apicl(config-tenant-interface) # ip address 172.1.1.1/24
apicl(config-tenant-interface) # ipv6 address 2001:1:1::1/64
apicl(config-tenant-interface) # exit
```

#### What to do next

The next section describes how to add an application profile, create an application endpoint group (EPG), and associate the EPG to the bridge domain.

#### **Related Topics**

Configuring a VLAN Domain Using the NX-OS Style CLI

## Creating a Tenant, VRF, and Bridge Domain Using the REST API

#### **Procedure**

#### **Step 1** Create a tenant.

#### Example:

```
POST https://apic-ip-address/api/mo/uni.xml
<fvTenant name="ExampleCorp"/>
```

When the POST succeeds, you see the object that you created in the output.

#### **Step 2** Create a VRF and bridge domain.

Note

The Gateway Address can be an IPv4 or an IPv6 address. For more about details IPv6 gateway address, see the related KB article, *KB: Creating a Tenant, VRF, and Bridge Domain with IPv6 Neighbor Discovery*.

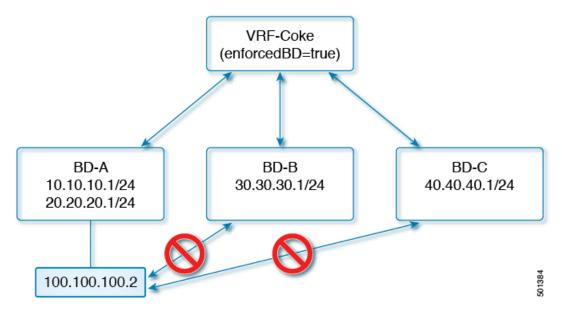
#### **Example:**

**Note** If you have a public subnet when you configure the routed outside, you must associate the bridge domain with the outside configuration.

## **Configuring an Enforced Bridge Domain**

An enforced bridge domain configuration entails creating an endpoint in a subject endpoint group (EPG) that can only ping subnet gateways within the associated bridge domain. With this configuration, you can then create a global exception list of IP addresses that can ping any subnet gateway.

Figure 3: Enforced Bridge Domain





Note

- The exception IP addresses can ping all of the bridge domain gateways across all of your VRF instances.
- A loopback interface configured for an L3Out does not enforce reachability to the IP address that
  is configured for the subject loopback interface.
- When an eBGP peer IP address exists in a different subnet than the subnet of the L3Out interface, the peer subnet must be added to the allowed exception subnets. Otherwise, eBGP traffic is blocked because the source IP address exists in a different subnet than the L3Out interface subnet.
- An enforced bridge domain is not supported with the Management tenant, regardless if the VRF instances are in-band or out-of-band, and any rules to control the traffic to these VRF instances should be configured using regular contracts.

### Configuring an Enforced Bridge Domain Using the NX-OS Style CLI

This section provides information on how to configure your enforced bridge domain using the NX-OS style command line interface (CLI).

#### **Procedure**

#### **Step 1** Create and enable the tenant:

#### **Example:**

In the following example ("cokeVrf") is created and enabled.

```
apicl(config-tenant) # vrf context cokeVrf
apicl(config-tenant-vrf) # bd-enforce enable
apicl(config-tenant-vrf) # exit
apicl(config-tenant) #exit
```

#### **Step 2** Add the subnet to the exception list.

#### Example:

```
apic1(config) #bd-enf-exp-ip add1.2.3.4/24
apic1(config) #exit
```

You can confirm if the enforced bridge domain is operational using the following type of command:

```
apic1# show running-config all | grep bd-enf
bd-enforce enable
bd-enf-exp-ip add 1.2.3.4/24
```

#### Example

The following command removes the subnet from the exception list:

```
apic1(config) # no bd-enf-exp-ip 1.2.3.4/24
apic1(config) #tenant coke
apic1(config-tenant) #vrf context cokeVrf
```

#### What to do next

To disable the enforced bridge domain run the following command:

apic1(config-tenant-vrf)# no bd-enforce enable

## **Configuring an Enforced Bridge Domain Using the REST API**

#### **Procedure**

	Command or Action  Create a tenant.  Example:  POST https://apic-ip-address/api/mo/uni.xml <fvtenant name="ExampleCorp"></fvtenant>		Purpose  When the POST succeeds, you see the object that you created in the output.	
Step 1				
Step 2	POST https://apic-ip-address/api/mo/uni.xml <fvtenant name="ExampleCorp"></fvtenant>	Note	The Gateway Address can be an IPv4 or an IPv6 address. For more about details IPv6 gateway address, see the related KB article, KB: Creating a Tenant, VRF, and Bridge Domain with IPv6 Neighbor Discovery.	

# Configuring Flood in Encapsulation for All Protocols and Proxy ARP Across Encapsulations

Cisco Application Centric Infrastructure (ACI) uses the bridge domain as the Layer 2 broadcast boundary and each bridge domain can include multiple Endpoint Groups (EPGs). You can bind an encapsulation VLAN to the desired EPG to carry user traffic. In some design scenarios, flooding can cross different user VLANs (EPGs) when the EPGs are in the same bridge domain.

Using multiple VLANs in tunnel mode can introduce a few challenges. In a typical deployment using Cisco ACI with a single tunnel, as illustrated in the following figure, there are multiple EPGs under one bridge domain. In this case, certain traffic is flooded within the bridge domain (and thus in all the EPGs), with the risk of MAC learning ambiguities that can cause forwarding errors.

BD
1.1.1.254/24

EPG1

EPG2

VLAN tunnel mode

VC
vNet

Vian 10

Vian 11

H1 H2 H3 H4 88
1.1.1.10 1.1.1.11 1.1.1.20 1.1.1.21 88

Figure 4: Challenges of Cisco ACI with VLAN Tunnel Mode

In this topology, the fabric has a single tunnel network defined that uses one uplink to connect with the Cisco ACI leaf node. Two user VLANs, VLAN 10 and VLAN 11 are carried over this link. The bridge domain domain is set in flooding mode as the servers' gateways are outside the Cisco ACI cloud. ARP negotiations occur in the following process:

- The server sends one ARP broadcast request over the VLAN 10 network.
- The ARP packet travels through the tunnel network to the external server, which records the source MAC address, learned from its downlink.
- The server then forwards the packet out its uplink to the Cisco ACI leaf switch.

- The Cisco ACI fabric sees the ARP broadcast packet entering on access port VLAN 10 and maps it to EPG1.
- Because the bridge domain is set to flood ARP packets, the packet is flooded within the bridge domain and thus to the ports under both EPGs as they are in the same bridge domain.
- The same ARP broadcast packet comes back over the same uplink.
- The external server sees the original source MAC address from this uplink.

Result: the external device has the same MAC address learned from both the downlink port and uplink port within its single MAC forwarding table, causing traffic disruptions.

#### **Recommended Solution**

The **Flood in Encapsulation** option is used to limit flooding traffic inside the bridge domain to a single encapsulation. When two EPGs share the same bridge domain and **Flood in Encapsulation** is enabled, the EPG flooding traffic does not reach the other EPG.

Beginning with Cisco Application Policy Infrastructure Controller (APIC0 release 3.1(1), on the Cisco Nexus 9000 series switches (with names ending with EX and FX and later), all protocols are flooded in encapsulation. Also, when enabling **Flood in Encapsulation** for any inter-VLAN traffic, Proxy ARP ensures that the MAC flap issue does not occur, and it limits all flooding (ARP, GARP, and BUM) to the encapsulation. This applies for all EPGs under the bridge domain where it is enabled.



Note

Before Cisco APIC release 3.1(1), these features are not supported (Proxy ARP and all protocols being included when flooding within encapsulation). In an earlier Cisco APIC release or earlier generation switches (without EX or FX on their names), if you enable **Flood in Encapsulation** it does not function, no informational fault is generated, but Cisco APIC decreases the health score by 1.

The recommended solution is to support multiple EPGs under one bridge domain by adding an external switch. This design with multiple EPGs under one bridge domain with an external switch is illustrated in the following figure.

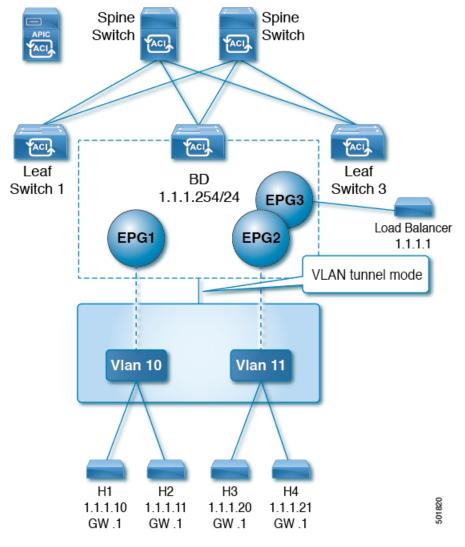


Figure 5: Design with Multiple EPGs Under one Bridge Domain with an External Switch

Within the same bridge domain, some EPGs can be service nodes and other EPGs can have flood in encapsulation configured. A Load Balancer resides on a different EPG. The load balancer receives packets from the EPGs and sends them to the other EPGs (There is no Proxy ARP and flood within encapsulation does not take place).

If you want to add flood in encapsulation only for selective EPGs, using the NX-OS style CLI, enter the **flood-on-encapsulation enable** command under EPGs.

If you want to add flood in encapsulation for all EPGs, you can use the **multi-destination encap-flood** CLI command under the bridge domain.

Using the CLI, flood in encapsulation configured for an EPG takes precedence over flood in encapsulation that is configured for a bridge domain.

When both bridge domains and EPGs are configured, the behavior is described as follows:

Table 1: Behavior When Both Bridge Domains and EPGs Are Configured

Configuration	Behavior
Flood in encapsulation at the EPG and flood in encapsulation at the bridge domain	Flood in encapsulation takes place for the traffic on all VLANs within the bridge domain.
No flood in encapsulation at the EPG and flood in encapsulation at the bridge domain	Flood in encapsulation takes place for the traffic on all VLANs within the bridge domain.
Flood in encapsulation at the EPG and no flood in encapsulation at the bridge domain	Flood in encapsulation takes place for the traffic on that VLAN within the EPG of the bridge domain.
No flood in encapsulation at the EPG and no flood in encapsulation at the bridge domain	Flooding takes place within the entire bridge domain.

#### **Multi-Destination Protocol Traffic**

The EPG/bridge domain level broadcast segmentation is supported for the following network control protocols:

- OSPF
- EIGRP
- CDP
- LACP
- LLDP
- IS-IS
- BGP
- IGMP
- PIM
- STP-BPDU (flooded within EPG)
- ARP/GARP (controlled by ARP Proxy)
- ND

#### Limitations

Here are the limitations for using flood in encapsulation for all protocols:

- Flood in encapsulation does not work in ARP unicast mode.
- Neighbor Solicitation (NS/ND) is not supported for this release.
- You must enable per-port CoPP with flood in encapsulation.
- Flood in encapsulation is supported only in bridge domain in flood mode and ARP in flood mode. Bridge domain spine proxy mode is not supported.
- IPv4 Layer 3 multicast is not supported.

- IPv6 is not supported.
- VM migration to a different VLAN has momentary issues (60 seconds).
- A load balancer acting as a gateway is supported, for example, in one to one communication between VMs and the load balancer in non-proxy mode. No Layer 3 communication is supported. The traffic between VMs and the load balancer is on Layer 2. However, if intra-EPG communication passes through the load balancer, then the load balancer changes the SIP and SMAC; otherwise it can lead to a MAC flap. Therefore, Dynamic Source Routing (DSR) mode is not supported in the load balancer.
- Setting up communication between VMs through a firwall, as a gateway, is not recommended because
  if the VM IP address changes to the gateway IP address instead of the firewall IP address, then the firewall
  can be bypassed.
- Prior releases are not supported (even interoperating between prior and current releases).
- The Proxy ARP and flood in encapsulation features are not supported for VXLAN encapsulation.
- A mixed-mode topology with Application Leaf Engine (ALE) and Application Spine Engine (ASE) is not recommended and is not supported with flood in encapsulation. Enabling them together can prevent QoS priorities from being enforced.
- Flood in encapsulation is not supported with remote leaf switches and Cisco ACI Multi-Site.
- Flood in encapsulation is not supported for Common Pervasive Gateway (CPGW).
- Flood in encapsulation is not supported on EPGs where microsegmentation is configured.
- If you configure the flood in encapsulation on all EPGs of a bridge domain, ensure that you configure the flood in encapsulation on the bridge domain as well.
- IGMP snooping is not supported with flood in encapsulation.
- There is a condition that causes Cisco ACI to flood in the bridge domain (instead of the encapsulation) packets that are received on an EPG that is configured for flood in encapsulation. This happens regardless of whether the administrator configured flood in encapsulation directly on the EPG or on the bridge domain. The condition for this forwarding behavior is if the ingress leaf node has a remote endpoint for the destination MAC address while the egress leaf node does not have a corresponding local endpoint. This can happen due to reasons such as an interface flapping, an endpoint flush due to STP TCN, learning being disabled on the bridge domain due to an excessive amount of moves, and so on.