Provisioning Core ACI Fabric Services

This chapter contains the following sections:

- Link Level Policies, on page 1
- Link Flap Policies, on page 3
- Configuring a Fabric Link Flap Policy Using the GUI, on page 3
- Time Synchronization and NTP, on page 4
- Configuring a DHCP Relay Policy, on page 9
- Configuring a DNS Service Policy, on page 14
- Configuring Custom Certificates, on page 18
- Provisioning Fabric Wide System Settings, on page 21
- Provisioning Global Fabric Access Policies, on page 44
- Per Port Policies, on page 47
- Creating a Mis-cabling Protocol Interface Policy Using the GUI (Optional), on page 49

Link Level Policies

You can configure link level policies, which are a type of access policy. A link level policy includes the physical layer (Layer 1) interface configurations, such as auto-negotiation, port speed, and link debounce.

Electromagnetic Interference Retrain

In the 5.2(4) and later releases, the electromagnetic interference (EMI) retrain feature filters any noise on a link due to electromagnetic interference, and retrains the link to avoid a link flap. Enable EMI retrain if your data center environment has a lot of EMI noise.

You can enable EMI retrain when you configure a link level policy by choosing **enable** for the **EMI Retrain** property. This feature is supported only with the Cisco N9K-C93108TC-EX and N9K-C93108TC-FX leaf switches using copper cables.

Configuring a Link Level Policy Using the GUI

Procedure

Step 1 On the menu bar, choose **Fabric** > **Access Policies**.

Step 2 In the **Navigation** pane, choose **Policies** > **Interface** > **Link Level**.

Step 3 Right-click **Link Level** and choose **Create Link Level Policy**.

Step 4 In the **Create Link Level Policy** dialog, fill out the fields as appropriate for your desired configuration.

For **Speed**, we recommend that you choose **inherit**, which is the default value. With this value, the Cisco APIC determines the speed based on the transceiver that is inserted into a switch.

See the tooltips for more information about the fields.

Note
For **Speed**, the **Auto** option is not supported; use the **Inherit** option instead.

Step 5 Click **Submit**.

Port Bring-up Delay

When you configure a link level policy, you can set the **Port bring-up delay (milliseconds)** parameter, which specifies a time in milliseconds that the decision feedback equalizer (DFE) tuning is delayed when a port is coming up. The delay is used to help avoid CRC errors during link bringup when using some third-party adapters. You should set the delay only as required; in most cases, you do not need to set a delay.

Note The **Port bring-up delay (milliseconds)** parameter is not honored on fabric extender (FEX) ports.

Limitations for Link Level Policies

You cannot use link bring up between 50G-SR-S (short-reach) tranceivers on the N9K-X9400-22L line card in the Nexus 9408 ACI-Mode switch and 25G-SR-S tranceivers to the APIC fabric. This is because you configure the 25G speed policy for link bring up and this combination only works between N9K switches and encrypted symetric keys (ESK).

If any link is connected to APIC, the default policy will be applied on the leaf port and the speed policy will be deleted.

This behavior applies to the 10G-GLCT transceiver as well, so you cannot use the 10G-GLCT transceiver for APIC bring up.

Link Debounce Interval

A link level policy applies to the fabric interfaces (leaf switch uplink ports and spine switch ports) at the link level and you use this policy where there might be link noise or transient flaps. In this policy, you can configure

the link debounce interval, which specifies a wait time for the fabric interfaces when there is a flap in the link. After the interval passes, Cisco Application Centric Infrastructure (ACI) checks the link status again to re-confirm the event. If the link is okay at this time, then the interfaces will remain up.

Set the interval in a link level policy using the **Link debounce interval (msec)** field in the Cisco Application Policy Infrastructure Controller (APIC) GUI or the **[no] link debounce time <time>** CLI command. The default debounce interval is 0 ms. We recommend a value of 100 ms, but you should choose a value that is appropriate to your fabric.

Link Flap Policies

Link flap is a situation in which a physical interface on a switch continually goes up and down over a period of time. The cause is usually related to a bad, unsupported, or non-standard cable or Small Form-Factor Pluggable (SFP), or is related to other link synchronization issues, and the cause can be intermittent or permanent.

A link flap policy specifies when to disable a switch port due to link flapping errors. In a link flap policy, you specify maximum number of times that a port of a switch can flap within a specified time span. If the port flaps more than the specified number of times in the specified time span, the port is given the "error-disable" state. The port remains in this state until you perform a manual flap on the port using the Cisco Application Policy Infrastructure Controller (APIC) to disable and enable the port.

Note A link flap policy is not honored on fabric extender (FEX) host interface (HIF) ports nor on leaf switch models without -EX, -FX, -FX2, -GX, or later designations in the product ID.

Configuring a Fabric Link Flap Policy Using the GUI

The following procedure configures a fabric link flap policy using the GUI, which you can then attach to any leaf or spine node interface policy to deploy the fabric link flap policy on the node's fabric ports.

Procedure

Step 1 On the menu bar, choose **Fabric** > **Fabric Policies**.

Step 2 In the **Navigation** pane, choose **Policies** > **Interface** > **Link Flap**.

Step 3 Right-click **Link Flap** and choose **Create Link Flap Policy**.

Step 4 In the **Create Fabric Link Level Policy** dialog, fill out the fields as appropriate for your desired configuration.

See the tooltips and online help for more information about the fields.

Step 5 Click **Submit**.

Time Synchronization and NTP

Within the Cisco Application Centric Infrastructure (ACI) fabric, time synchronization is a crucial capability upon which many of the monitoring, operational, and troubleshooting tasks depend. Clock synchronization is important for proper analysis of traffic flows as well as for correlating debug and fault time stamps across multiple fabric nodes.

An offset present on one or more devices can hamper the ability to properly diagnose and resolve many common operational issues. In addition, clock synchronization allows for the full utilization of the atomic counter capability that is built into the ACI upon which the application health scores depend. Nonexistent or improper configuration of time synchronization does not necessarily trigger a fault or a low health score. You should configure time synchronization before deploying a full fabric or applications so as to enable proper usage of these features. The most widely adapted method for synchronizing a device clock is to use Network Time Protocol (NTP).

Prior to configuring NTP, consider what management IP address scheme is in place within the ACI fabric. There are two options for configuring management of all ACI nodes and Application Policy Infrastructure Controllers (APICs), in-band management and/or out-of-band management. Depending upon which management option is chosen for the fabric, configuration of NTP will vary. Another consideration in deploying time synchronization is where the time source is located. The reliability of the source must be carefully considered when determining if you will use a private internal clock or an external public clock.

In-Band Management NTP

Note See the Adding Management Access section in this guide for information about in-band management access.

- In-Band Management NTP—When an ACI fabric is deployed with in-band management, consider the reachability of the NTP server from within the ACI in-band management network. In-band IP addressing used within the ACI fabric is not reachable from anywhere outside the fabric. To leverage an NTP server external to the fabric with in-band management, construct a policy to enable this communication..

NTP over IPv6

NTP over IPv6 addresses is supported in hostnames and peer addresses. The gai.conf can also be set up to prefer the IPv6 address of a provider or a peer over an IPv4 address. The user can provide a hostname that can be resolved by providing an IP address (both IPv4 or IPv6, depending on the installation or preference).

Configuring NTP Using the GUI

Note There is a risk of hostname resolution failure for hostname based NTP servers if the DNS server used is configured to be reachable over in-band or out-of-band connectivity. If you use a hostname, ensure that the DNS service policy to connect with the DNS providers is configured. Also ensure that the appropriate DNS label is configured for the in-band or out-of-band VRF instances of the management EPG that you chose when you configured the DNS profile policy.

Procedure

Step 1 On the menu bar, choose **Fabric** > **Fabric Policies**.

Step 2 In the **Navigation** pane, choose **Policies** > **Pod** > **Date and Time**.

Step 3 In the **Work** pane, choose **Actions** > **Create Date and Time Policy**.

Step 4 In the **Create Date and Time Policy** dialog box, perform the following actions:

a) Enter a name for the policy to distinguish between the different NTP configurations in your environment..

b) Click **enabled** for the **Authentication State** field and expand the **NTP Client Authentication Keys** table and enter the key information. Click **Update** and **Next**.

c) Click the + sign to specify the NTP server information (provider) to be used.

d) In the **Create Providers** dialog box, enter all relevant information, including the following fields: **Name**, **Description**, **Minimum Polling Intervals**, and **Maximum Polling Intervals**.

- If you are creating multiple providers, check the **Preferred** check box for the most reliable NTP source.
- In the **Management EPG** drop-down list, if the NTP server is reachable by all nodes on the fabric through out-of-band management, choose Out-of-Band. If you have deployed in-band management, see the details about In-Band Management NTP. Click **OK**.

Repeat the steps for each provider that you want to create.

Step 5 In the **Navigation** pane, choose **Pod Policies** > **Policy Groups**.

Step 6 In the **Work** pane, choose **Actions** > **Create Pod Policy Group**.

Step 7 In the **Create Pod Policy Group** dialog box, perform the following actions:

a) Enter a name for the policy group.

b) In the **Date Time Policy** field, from the drop-down list, choose the NTP policy that you created earlier. Click **Submit**.
The pod policy group is created. Alternatively, you can use the default pod policy group.

Step 8 In the **Navigation** pane, choose **Pod Policies** > **Profiles**.

Step 9 In the **Work** pane, double-click the desired pod selector name.

Step 10 In the Properties area, from the **Fabric Policy Group** drop-down list, choose the pod policy group you created. Click **Submit**.

Configuring NTP Using the REST API

Note There is a risk of hostname resolution failure for hostname based NTP servers if the DNS server used is configured to be reachable over in-band or out-of-band connectivity. If you use a hostname, ensure that the DNS service policy to connect with the DNS providers is configured. Also ensure that the appropriate DNS label is configured for the in-band or out-of-band VRF instances of the management EPG that you chose when you configured the DNS profile policy.

Procedure

Step 1 Configure NTP.

Example:

```
POST url: https://APIC-IP/api/node/mo/uni/fabric/time-test.xml

<imdata totalCount="1">
    <datetimePol adminSt="enabled" authSt="disabled" descr="" dn="uni/fabric/time-CiscoNTPPol"
 name="CiscoNTPPol" ownerKey="" ownerTag="">
        <datetimeNtpProv descr="" keyId="0" maxPoll="6" minPoll="4" name="10.10.10.11"
preferred="yes">
            <datetimeRsNtpProvToEpg tDn="uni/tn-mgmt/mgmtp-default/inb-default"/>
        </datetimeNtpProv>
    </datetimePol>
</imdata>
```

Step 2 Add the default Date Time Policy to the pod policy group.

Example:

```
POST url: https://APIC-IP/api/node/mo/uni/fabric/funcprof/podpgrp-calo1/rsTimePol.xml

POST payload: <imdata totalCount="1">
<fabricRsTimePol tnDatetimePolName="CiscoNTPPol">
</fabricRsTimePol>
</imdata>
```

Step 3 Add the pod policy group to the default pod profile.

Example:

```
POST url:
https://APIC-IP/api/node/mo/uni/fabric/podprof-default/pods-default-typ-ALL/rspodPGrp.xml

payload: <imdata totalCount="1">
<fabricRsPodPGrp tDn="uni/fabric/funcprof/podpgrp-calo1" status="created">
</fabricRsPodPGrp>
</imdata>
```

Verifying NTP Operation Using the GUI

Procedure

Step 1 On the menu bar, choose **FABRIC** > **Fabric Policies**.

Step 2 In the **Navigation** pane, choose **Pod Policies** > **Policies** > **Date and Time** > ***ntp_policy*** > ***server_name***.

The *ntp_policy* is the previously created policy. An IPv6 address is supported in the Host Name/IP address field. If you enter a hostname and it has an IPv6 address set, you must implement the priority of IPv6 address over IPv4 address.

Step 3 In the **Work** pane, verify the details of the server.

NTP Server

The NTP server enables client switches to also act as NTP servers to provide NTP time information to downstream clients. When the NTP server is enabled, the NTP daemon on the switch responds with time information to all unicast (IPv4/IPv6) requests from NTP clients. NTP server implementation is complaint to NTP RFCv3. As per the NTP RFC, the server will not maintain any state related to the clients.

- The NTP server enables IP addresses in all tenant VRFs and the in-band/out-of-band management VRFs to serve NTP clients.
- The NTP server responds to incoming NTP requests on both Management VRFs or tenant VRFs, and responds back using the same VRF.
- The NTP server supports both IPv4/IPv6.
- Switches can sync as an IPv4 client and serve as an IPv6 server, and vice versa.
- Switches can sync as an NTP client using either the out-of-band management or in-band management VRF and serve NTP clients from either management VRF or tenant VRF.
- No additional contracts or IP table configurations are required.
- If the switch is synced to the upstream server, then the server will send time info with the stratum number, and increment to its system peer's stratum.
- If the switch clock is undisciplined (not synced to the upstream server), then the server will send time information with stratum 16. Clients will not be able to sync to this server.

By default, NTP server functionality is disabled. It needs to be enabled explicitly by the configuration policy.

Note Clients can use the in-band, out-of-band IP address of the leaf switch as the NTP server IP address. Clients can also use the bridge domain SVI of the EPG of which they are part or any L3Out IP address as the NTP server IP address for clients outside of the fabric.

Fabric switches should not sync to other switches of the same fabric. The fabric switches should always sync to external NTP servers.

Enabling the NTP Server Using the GUI

This section explains how to enable an NTP server when configuring NTP in the APIC GUI.

Procedure

Step 1 On the menu bar, choose **FABRIC** > **Fabric Policies**.

Step 2 In the **Navigation** pane, choose **Pod Policies** > **Policies** .

The **Date and Time** option appears in the **Navigation** pane.

Step 3 From the **Navigation** pane, right-click on **Date and Time** and choose **Create Date and Time Policy**.

The **Create Date and Time Policy** dialog appears in the **Work** pane.

Step 4 In the **Create Date and Time Policy** dialog box, perform the following actions:

a) Enter a name for the policy to distinguish between the different NTP configurations in your environment.

b) For the **Server State** option, click **enabled**.

Server State enables switches to act as NTP servers to provide NTP time information to downstream clients.

Note
To support the server functionality, it is always recommended to have a peer setup for the server. This enables the server to have a consistent time to provide to the clients.

When **Server State** is enabled:

- The NTP server sends time info with a stratum number, an increment to the system peer's stratum number, to switches that are synched to the upstream server.
- The server sends time info with stratum 16 if the switch clock is not synched to the upstream server. Clients are not able to sync to this server.

Note
To support the server functionality, it is always recommended to have a peer setup for the server. The peer setup allows for a consistent time to provide to the clients.

c) For the **Master Mode** option, click **enabled**.

Master Mode enables the designated NTP server to provide undisciplined local clock time to downstream clients with a configured stratum number. For example, a leaf switch that is acting as the NTP server can provide undisciplined local clock time to leaf switches acting as clients.

Note
- **Master Mode** is only applicable when the server clock is undisciplined.
- The default master mode **Stratum Value** is 8.

d) For the **Stratum Value** field, specify the stratum level from which NTP clients will get their time synchronized. The range is from 1 to 14.

e) Click **Next**.

f) Click the **+** sign to specify the NTP server information (provider) to be used.

g) In the **Create Providers** dialog box, enter all relevant information, including the following fields: **Name**, **Description**, **Minimum Polling Intervals**, and **Maximum Polling Intervals**.

- If you are creating multiple providers, check the **Preferred** check box for the most reliable NTP source.
- In the Management EPG drop-down list, if the NTP server is reachable by all nodes on the fabric through out-of-band management, choose Out-of-Band. If you have deployed in-band management, see the details about In-Band Management NTP. Click **OK**.

Repeat the steps for each provider that you want to create.

Step 5 In the **Navigation** pane, choose **Pod Policies** then right-click on **Policy Groups**.

The **Create Pod Policy Group** dialog appears.

Step 6 In the **Work** pane, choose **Actions** > **Create Pod Policy Group**.

Step 7 In the **Create Pod Policy Group** dialog box, perform the following actions:

a) Enter a name for the policy group.

b) In the **Date Time Policy** field, from the drop down list, choose the NTP policy that you created earlier. Click **Submit**.
The pod policy group is created. Alternatively, you can use the default pod policy group.

Step 8 In the **Navigation** pane, choose **Pod Policies** > **Profiles**.

Step 9 In the **Work** pane, double-click the desired pod selector name.

Step 10 In the Properties area, from the **Fabric Policy Group** drop down list, choose the pod policy group you created.

Step 11 Click **Submit**.

Configuring the Datetime Format Using the GUI

This section demonstrates how to configure the datetime format using the Cisco APIC GUI.

Procedure

Step 1 On the menu bar, click **System** > **System Settings** .

Step 2 In the Navigation pane, click **Date and Time**.

Step 3 In the Work pane, choose from the following options:

- **Display Format**—Click **local** to display the date and time in local time, or click **utc** to display the date and time in UTC. The default is **local**.
- **Time Zone**—Click the drop-down arrow to choose the time zone for your domain. The default is **Coordinated Universal Time**.
- **Offset State**—Click **enable** or **disable**. When enabled, the difference between the local time and the reference time is displayed. The default is **enable**.

Configuring a DHCP Relay Policy

A DHCP relay policy may be used when the DHCP client and server are in different subnets.

When a Cisco Application Centric Infrastructure (ACI) fabric acts as a DHCP relay, it inserts the DHCP Option 82 (the DHCP Relay Agent Information Option) in DHCP requests that it proxies on behalf of clients. If a response (DHCP offer) comes back from a DHCP server without Option 82, it is silently dropped by the fabric. Therefore, when the Cisco ACI fabric acts as a DHCP relay, DHCP servers providing IP addresses to compute nodes attached to the Cisco ACI fabric must support Option 82.

Guidelines and Limitations for a DHCP Relay Policy

- A DHCP relay policy created under the infra or common tenant is not available to other tenants when configuring DHCP relay in a bridge domain. For inter-tenant DHCP relay communications, create a global DHCP relay policy, as described in Create a Global DHCP Relay Policy, on page 46.
- The DHCP relay IP address will always be set to the primary SVI IP address.
- You cannot use a DHCP relay policy if the clients and servers are connected through an L3Out EPG. Either the clients or servers (or both) must be in a regular (non-L3Out) EPG.
- When a DHCP relay provider EPG is an L3Out External EPG and is configured in a different VRF instance than the client bridge domain, a DHCP relay packet will be sent sourced from each interface configured under the L3Out. This includes the node router-IDs addresses and all L3Out interface IP addresses.
- When an L3Out SVI or Floating SVI is configured as a DHCP relay agent, (DHCP relay label) the DHCP server/provider must not be connected to the same leaf switches where the L3Out SVI is deployed. For floating SVIs, this includes all anchor and non-anchor switches.
- The DHCP relay option policy only supports option 79, and option 79 is only supported for DHCPv6.
- The incoming DHCPv4 packet uses source port 68 (0x44) and destination port 67 (0x43), with specific source and destination IP addresses (indicating it is a unicast packet). Because the packet has source port 68 and destination port 67, it is subject to sup copying: one copy is sent directly to the server, while another is punted to the DHCP process in ACI. When the DHCP process receives the packet, any DHCPv4 request or response packet with a non-zero gateway IP address (that is not a local address) and is unicast will be forwarded to the destination, even if relay is not configured. As a result, an additional copy of the packet reaches the server.

Configuring a DHCP Server Policy for the APIC Infrastructure Using the GUI

This procedure deploys a DHCP relay policy for an endpoint group (EPG).

Observe the following guidelines and restrictions:

- The port and the encapsulation used by the application Endpoint Group must belong to a physical or VM Manager (VMM) domain. If no such association with a domain is established, the APIC continues to deploy the EPG but raises a fault.
- Cisco APIC supports DHCP relay for both IPv4 and IPv6 tenant subnets. DHCP server addresses can be IPv4 or IPv6. DHCPv6 relay will occur only if IPv6 is enabled on the fabric interface and one or more DHCPv6 relay servers are configured.
- Cisco APIC supports DHCP relay for only the primary IP address pool.

Before you begin

Make sure that Layer 2 or Layer 3 management connectivity is configured.

Procedure

Step 1 On the menu bar, choose **Tenants** > **infra**.

Step 2 In the **Navigation** pane, under **Tenant infra**, expand **Networking** > **Policies** > **Protocol** > **DHCP**.

Step 3 Right-click **Relay Policies** and click **Create DHCP Relay Policy**.

Step 4 In the **Create DHCP Relay Policy** dialog box, perform the following actions:

a) In the **Name** field, enter the DHCP relay profile name (DhcpRelayP).
b) Expand **Providers**. In the **Create DHCP Provider** dialog box, in the **EPG Type** field, click the appropriate radio button depending upon where the DHCP server is connected.
c) In the **Application EPG** area, in the **Tenant** field, from the drop-down list, choose the tenant. (infra)
d) In the **Application Profile** field, from the drop-down list, choose the application. (access)
e) In the **EPG** field, from the drop-down list, choose the EPG. (default)
f) In the **DHCP Server Address** field, enter the IP address for the DHCP server. Click **Update**.
g) Click **Submit**.

The DHCP relay policy is created.

Step 5 In the **Navigation** pane, expand **Networking** > **Bridge Domains** > **default** > **DHCP Relay Labels**.

Step 6 Right-click **DHCP Relay Labels**, and click **Create DHCP Relay Label**.

Step 7 In the **Create DHCP Relay Label** dialog box, perform the following actions:

a) In the **Scope** field, click the tenant radio button.
This action displays, in the **Name** field drop-down list, the DHCP relay policy created earlier.
b) In the **Name** field, from the drop-down list, choose the name of the DHCP policy created (DhcpRelayP) or create a new relay policy by choosing **Create DHCP Relay Policy**.
c) In the **DHCP Option Policy**, select an existing option policy, or create a new one by choosing **Create DHCP Option Policy**.
d) Click **Submit**.

The DHCP server is associated with the bridge domain.

Step 8 In the **Navigation** pane, expand **Networking** > **Bridge Domains** > **default** > **DHCP Relay Labels** to view the DHCP server created.

Configuring Option 79 Using REST API

To configure Option 79 for a DHCP option policy using REST API:

```
POST URL: https://apic-ip-address/api/mo/uni.xml

<dhcpOptionPol dn="uni/tn-dhcp_client/dhcpoptpol-dhcp_option_policy" name="dhcp_option_policy"
 status="">
<dhcpOption data="" id="79" name="option_79"/>
</dhcpOptionPol>
```

Configuring a DHCP Server Policy for the APIC Infrastructure Using the NX-OS Style CLI

- The port and the encapsulation used by the application Endpoint Group must belong to a physical or VM Manager (VMM) domain. If no such association with a domain is established, the APIC continues to deploy the EPG but raises a fault.
- Cisco APIC supports DHCP relay for both IPv4 and IPv6 tenant subnets. DHCP server addresses can be IPv4 or IPv6. DHCPv6 relay will occur only if IPv6 is enabled on the fabric interface and one or more DHCPv6 relay servers are configured.

Before you begin

Ensure that Layer 2 or Layer 3 connectivity is configured to reach the DHCP server address.

Procedure

Configure DHCP server policy settings for the APIC infrastructure traffic.

Example:

DHCP Relay Policy for an Endpoint Group

```
apic1(config)# tenant infra
apic1(config-tenant)# template dhcp relay policy DhcpRelayP
apic1(config-tenant-template-dhcp-relay)# ip address 10.0.0.1 tenant infra application access epg default
apic1(config-tenant-template-dhcp-relay)# exit
apic1(config-tenant)# interface bridge-domain default
apic1(config-tenant-interface)# dhcp relay policy tenant DhcpRelayP
apic1(config-tenant-interface)# exit
```

Example:

DHCP Relay Policy for Layer 3 Outside

```
ifav28-ifc2(config)# tenant dhcpTn
ifav28-ifc2(config-tenant)# template dhcp relay policy DhcpRelayPol
ifav28-ifc2(config-tenant-template-dhcp-relay)# ip address 11.1.1.11 tenant dhcpTn application ap epg serverEpg
ifav28-ifc2(config-tenant-template-dhcp-relay)# exit
ifav28-ifc2(config-tenant)# exit
ifav28-ifc2(config)# leaf 2001
ifav28-ifc2(config-leaf)# interface ethernet 1/4
ifav28-ifc2(config-leaf-if)# no switchport
ifav28-ifc2(config-leaf-if)# vrf member tenant dhcpTn vrf v1
ifav28-ifc2(config-leaf-if)# dhcp relay policy tenant DhcpRelayPol
ifav28-ifc2(config-leaf-if)# exit
```

Configuring a DHCP Server Policy for the APIC Infrastructure Using the REST API

- This task is a prerequisite for users who want to create a vShield domain profile.
- The port and the encapsulation used by the application endpoint group must belong to a physical or VM Manager (VMM) domain. If no such association with a domain is established, the Cisco Application Policy Infrastructure Controller (APIC) continues to deploy the EPG but raises a fault.

- Cisco APIC supports DHCP relay for both IPv4 and IPv6 tenant subnets. DHCP server addresses can be IPv4 or IPv6. DHCPv6 relay will occur only if IPv6 is enabled on the fabric interface and one or more DHCPv6 relay servers are configured.

Before you begin

Make sure that Layer 2 or Layer 3 management connectivity is configured.

Procedure

Configure the Cisco APIC as the DHCP server policy for the infrastructure tenant.

Note
This relay policy will be pushed to all the leaf ports that are connected hypervisors using the attach entity profile configuration. For details about configuring with attach entity profile, see the examples related to creating VMM domain profiles.

Example:

DHCP Relay Policy for EPG

```
<!-- api/policymgr/mo/.xml -->
<polUni>

POST https://apic-ip-address/api/mo/uni.xml


  <fvTenant name="infra">

    <dhcpRelayP name="DhcpRelayP"  owner="tenant">
      <dhcpRsProv tDn="uni/tn-infra/ap-access/epg-default" addr="10.0.0.1" />
    </dhcpRelayP>

    <fvBD name="default">
      <dhcpLbl name="DhcpRelayP" owner="tenant"/>
    </fvBD>

  </fvTenant>
</polUni>
```

Example:

DHCP Relay Policy for Layer 3 Outside

Note
You must specify DHCP Relay label under **l3extLIfP** with an appropriate name and owner.

```
<polUni>
  <fvTenant name="dhcpTn">
    <l3extOut name="Out1" >
      <l3extLNodeP name="NodeP" >
        <l3extLIfP name="Intf1">
           <dhcpLbl name="DhcpRelayPol" owner="tenant" />
       </l3extLIfP>
     </l3extLNodeP>
   </l3extOut>
 </fvTenant>
<polUni>
```

```
POST https://apic-ip-address/api/mo/uni.xml
```

Configuring a DNS Service Policy

A DNS policy is required to connect to external servers, for example AAA, RADIUS, vCenter, and services by hostname. A DNS service policy is a shared policy, so any tenant and VRF that uses this service must be configured with the specific DNS profile label. To configure a DNS policy for the ACI fabric, you must complete the following tasks:

- Ensure that the management EPG is configured for the DNS policy, otherwise this policy will not take into effect on the switches.

Note For the management EPG, only the default DNS policy is supported.

- Create a DNS profile (default) that contains the information about DNS providers and DNS domains.
- Associate the DNS profile (default or another DNS profile) name to a DNS label under the required tenant.

It is possible to configure a per-tenant, per-VRF DNS profile configuration. Additional DNS profiles can be created and applied to specific VRFs of specific tenants using the appropriate DNS label. For example, if you create a DNS profile with a name of acme, you can add a DNS label of acme to the appropriate **Networking** > **VRF** policy configuration in the tenants configuration.

Configuring External Destinations with an In-Band DNS Service Policy

Configure the external destinations for the services as follows:

Source	In-Band Management	Out-of-Band Management	External Server Location
APIC	IP address or Fully Qualified domain name (FQDN)	IP address or FQDN	Anywhere
Leaf switches	IP address	IP address or FQDN **Note** The DNS policy must specify the out-of-band management EPG for reachability of the DNS server.	Anywhere

Source	In-Band Management	Out-of-Band Management	External Server Location
Spine switches	IP address	IP address or FQDN **Note** The DNS policy must specify the out-of-band management EPG for reachability of the DNS server.	Directly connected to a leaf switch

The following is a list of external servers:

- Call Home SMTP server
- Syslog server
- SNMP Trap destination
- Statistics Export destination
- Configuration Export destination
- Techsupport Export destination
- Core Export destination

The recommended guidelines are as follows:

- The external servers must be atatched to the leaf access ports.
- Use in-band connectivity for the leaf switches to avoid extra cabling for the management port.
- Use out-of-band management connectivity for the spine switches. Connect this out-of-band network for spine switches to one of the leaf ports with in-band management virtual routing and forwarding (VRF) so that the spine switches and the leaf switches can reach the same set of external servers.
- Use IP addresses for the external servers.

Dual Stack IPv4 and IPv6 DNS Servers

DNS servers have primary DNS records which can be A records (IPV4) or AAAA records (IPV6). Both A and AAAA records associate domain name with a specific IP address (IPv4 or IPv6).

The ACI fabric can be configured to use reputable public DNS servers that run on IPv4. These servers are able to resolve and respond with A record (IPv4) or AAAA record (IPv6).

In a pure IPv6 environment, the system administrators must use IPv6 DNS servers. The IPv6 DNS servers are enabled by adding them to /etc/resolv.conf.

A more common environment is to have dual-stack IPv4 and IPv6 DNS servers. In the dual-stack case, both IPv4 and IPv6 name servers are listed in /etc/resolv.conf. However, in a dual-stack environment, simply appending the IPv6 DNS servers to the list may cause a large delay in DNS resolutions. This is because the IPv6 protocol takes precedence by default, and it is unable to connect to the IPv4 DNS servers (if they are listed first in /etc/resolv.conf). The solution is to list IPv6 DNS servers ahead of IPv4 DNS servers. Also add “options single-request-reopen” to enable the same socket to be used for both IPv4 and IPv6 lookups.

Here is an example of resolv.conf in dual-stack IPv4 and IPv6 DNS servers where the IPv6 DNS servers are listed first. Also note the “single-request-reopen” option:

```
options single-request-reopen
nameserver 2001:4860:4680::8888
nameserver 2001:4860:4680::8844
nameserver 8.8.8.8
nameserver 8.8.4.4
```

Dual-Stack IPv4 and IPv6 Environment

If the management network in the ACI fabric supports both IPv4 and IPv6, the Linux system application (glibc) will use the IPv6 network by default because getaddrinfo() will return IPv6 first.

Under certain conditions however, an IPv4 address may be preferred over an IPv6 address. The Linux IPv6 stack has a feature which allows an IPv4 address mapped as an IPv6 address using IPv6 mapped IPv4 address (::ffff/96). This allows an IPv6 capable application to use only a single socket to accept or connect both IPv4 and IPv6. This is controlled by the glibc IPv6 selection preference for getaddrinfo() in /etc/gai.conf.

In order to allow glibc to return multiple addresses when using /etc/hosts, “multi on” should be added to the /etc/hosts file. Otherwise, it may return only the first match.

If an application is not aware whether both IPv4 and IPv6 exist, it may not perform fallback attempts using different address families. Such applications may require a fallback implementation.

Policy for Priority of IPv4 or IPv6 in a DNS Profile

The DNS profile supports version preference choices between IPv4 and IPv6. Using the user interface, you can enable your preference. IPv4 is the default.

The following is an example of a policy based configuration using Postman REST API:

```
<?xml version="1.0" encoding="UTF-8”?>
<!— api/node/mo/uni/fabric/dnsp-default.xml —>
<dnsProfile dn="uni/fabric/dnsp-default" IPVerPreference="IPv6" childAction="" descr="" >
</dnsProfile>
```

The gai.conf settings control destination address selection. The file has a label table, precedence table, and an IPv4 scopes table. The changes for prioritizing IPv4 or IPv6 over the other need to go into the precedence table entries. Given below are sample contents of the standard file as it is used in Linux systems for many flavors. A single line of precedence label in the file overrides any default settings.

The following is an example of a gai.conf to prioritize IPv4 over IPv6:

```
# Generated by APIC
label ::1/128       0
label ::/0          1
label 2002::/16     2
label ::/96         3
label ::ffff:0:0/96 4
precedence  ::1/128       50
precedence ::/0           40
precedence 2002::/16      30
precedence ::/96          20
# For APICs prefering IPv4 connections, change the value to 100.
precedence ::ffff:0:0/96  10
```

Configuring a DNS Service Policy to Connect with DNS Providers Using the GUI

Before you begin

Make sure that Layer 2 or Layer 3 management connectivity is configured.

Procedure

Step 1 On the menu bar, choose **FABRIC** > **Fabric Policies**. In the **Navigation** pane, expand **Global Policies** > **DNS Profiles**, and click the default DNS profile.

Step 2 In the **Work** pane, in the **Management EPG** field, from the drop-down list, choose the appropriate management EPG (default (Out-of-Band)).

Step 3 Expand **DNS Providers**, and perform the following actions:

a) In the **Address** field, enter the provider address.

b) In the **Preferred** column, check the check box if you want to have this address as the preferred provider.

You can have only one preferred provider.

c) Click **Update**.

d) (Optional) To add a secondary DNS provider, expand **DNS Providers**, and in the **Address** field, type the provider address. Click **Update**.

Step 4 Expand **DNS Domains**, and perform the following actions:

a) In the **Name** field, enter the domain name (cisco.com).

b) In the **Default** column, check the check box to make this domain the default domain.

You can have only one domain name as the default.

c) Click **Update**.

d) (Optional) To add a secondary DNS domain, expand **DNS Domains**. In the **Address** field, enter the secondary domain name. Click **Update**.

Step 5 Click **Submit**.
The DNS server is configured.

Step 6 On the menu bar, click **TENANTS** > **mgmt**.

Step 7 In the **Navigation** pane, expand **Networking** > **VRF** > **oob**, and click **oob**.

Step 8 In the **Work** pane, under **Properties**, in the **DNS labels** field, enter the appropriate DNS label (default). Click **Submit**.
The DNS profile label is now configured on the tenant and VRF.

Configuring Custom Certificates

Configuring Custom Certificate Guidelines

- Exporting a private key that is used to generate a Certificate Signing Request (CSR) on the Cisco Application Policy Infrastructure Controller (APIC) is not supported. If you want to use the same certificate on multiple servers through a wildcard in the Subject Alternative Name (SAN) field, such as "*cisco.com," by sharing the private key that was used to generate the CSR for the certificate, generate the private key outside of Cisco Application Centric Infrastructure (ACI) fabric and import it to the Cisco ACI fabric.
- You must download and install the public intermediate and root CA certificates before generating a Certificate Signing Request (CSR). Although a root CA Certificate is not technically required to generate a CSR, Cisco requires the root CA certificate before generating the CSR to prevent mismatches between the intended CA authority and the actual one used to sign the CSR. The Cisco APIC verifies that the certificate submitted is signed by the configured CA.
- To use the same public and private keys for a renewed certificate generation, you must satisfy the following guidelines:
  - You must preserve the originating CSR as it contains the public key that pairs with the private key in the key ring.
  - The same CSR used for the originating certificate must be resubmitted for the renewed certificate if you want to re-use the public and private keys on the Cisco APIC.
  - Do not delete the original key ring when using the same public and private keys for the renewed certificate. Deleting the key ring will automatically delete the associated private key used with CSRs.
- Cisco ACI Multi-Site, VCPlugin, VRA, and SCVMM are not supported for certificate-based authentication.
- Only one SSL certificate is allowed per Cisco APIC cluster.
- You must disable certificate-based authentication before downgrading to release 4.0(1) from any later release.
- To terminate the certificate-based authentication session, you must log out and then remove the CAC card.
- The custom certificate configured for the Cisco APIC will be deployed to the leaf and spine switches. If the URL or DN that is used to connect to the fabric node is within the **Subject** or **Subject Alternative Name** field, the fabric node will be covered under the certificate.
- The Cisco APIC GUI can accept a certificate with a maximum size of 4k bytes.
- When a self-signed SSL certificate that you are using for HTTPS access expires, the certificate gets renewed automatically.

Configuring a Custom Certificate for Cisco ACI HTTPS Access Using the GUI

CAUTION: PERFORM THIS TASK ONLY DURING A MAINTENANCE WINDOW AS THERE IS A POTENTIAL FOR DOWNTIME. The downtime affects access to the Cisco Application Policy Infrastructure

Controller (APIC) cluster and switches from external users or systems and not the Cisco APIC to switch connectivity. The NGINX process on the switches will also be impacted, but that will be only for external connectivity and not for the fabric data plane. Access to the Cisco APIC, configuration, management, troubleshooting, and such will be impacted. The NGINX web server running on the Cisco APIC and switches will be restarted during this operation.

Before you begin

Determine from which authority you will obtain the trusted certification so that you can create the appropriate Certificate Authority.

Procedure

Step 1 On the menu bar, choose **Admin** > **AAA**.

Step 2 In the **Navigation** pane, choose **Security**.

Step 3 In the **Work** pane, choose **Public Key Management** > **Certificate Authorities** > **Create Certificate Authority**.

Step 4 In the **Create Certificate Authority** dialog box, in the **Name** field, enter a name for the certificate authority.

Step 5 In the **Certificate Chain** field, copy the intermediate and root certificates for the certificate authority that will sign the Certificate Signing Request (CSR) for the Cisco APIC.

The certificate should be in Base64 encoded X.509 (CER) format. The intermediate certificate is placed before the root CA certificate. It should look similar to the following example:

```
-----BEGIN CERTIFICATE-----
<Intermediate Certificate>
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
<Root CA Certificate>
-----END CERTIFICATE-----
```

Step 6 Click **Submit**.

Step 7 In the **Navigation** pane, choose **Public Key Management** > **Key Rings**.

Step 8 In the **Work** pane, choose **Actions** > **Create Key Ring**.

The key ring enables you to manage a private key (imported from external device or internally generated on APIC), a CSR generated by the private key, and the certificate signed via the CSR.

Step 9 In the **Create Key Ring** dialog box, in the **Name** field, enter a name.

Step 10 In the **Certificate** field, do not add any content if you will generate a CSR using the Cisco APIC through the key ring. Alternately, add the signed certificate content if you already have one that was signed by the CA from the previous steps by generating a private key and CSR outside of the Cisco APIC,

Step 11 In the **Modulus** field, click the radio button for the desired key strength.

Step 12 In the **Certificate Authority** field, from the drop-down list, choose the certificate authority that you created earlier, then click **Submit**.

Step 13 In the **Private Key** field, do not add any content if you will generate a CSR using the Cisco APIC through the key ring. Alternately, add the private key used to generate the CSR for the signed certificate that you entered in step 10.

Note

Do not delete the key ring. Deleting the key ring will automatically delete the associated private key used with CSRs.

If you have not entered the signed certificate and the private key, in the **Work** pane, in the **Key Rings** area, the **Admin State** for the key ring created displays **Started**, waiting for you to generate a CSR. Proceed to step 14.

If you entered both the signed certificate and the private key, in the **Key Rings** area, the **Admin State** for the key ring created displays **Completed**. Proceed to step 23.

Step 14 In the **Navigation** pane, choose **Public Key Management** > **Key Rings** > ***key_ring_name***.

Step 15 In the **Work** pane, choose **Actions** > **Create Certificate Request**.

Step 16 In the **Subject** field, enter the common name (CN) of the CSR.

You can enter the fully qualified domain name (FQDN) of the Cisco APICs using a wildcard, but in a modern certificate, we generally recommend that you enter an identifiable name of the certificate and enter the FQDN of all Cisco APICs in the **Alternate Subject Name** field (also known as the *SAN* – Subject Alternative Name) because many modern browsers expect the FQDN in the SAN field.

Step 17 In the **Alternate Subject Name** field, enter the FQDN of all Cisco APICs, such as "DNS:apic1.example.com,DNS:apic2.example.com,DNS:apic3.example.com" or "DNS:*example.com".

Alternatively, if you want SAN to match an IP address, enter the Cisco APICs' IP addresses with the following format:

```
IP:192.168.2.1
```

You can use DNS names, IPv4 addresses, or a mixture of both in this field. IPv6 addresses are not supported.

Step 18 Fill in the remaining fields as appropriate.

Note
Check the online help information available in the **Create Certificate Request** dialog box for a description of the available parameters.

Step 19 Click **Submit**.

Inside the same key ring, the **Associated Certificate Request** area is now displayed with the **Subject**, **Alternate Subject Name** and other fields you entered in the previous steps along with the new field **Request**, which contains the content of the CSR that is tied to this key ring. Copy the content from the **Request** field to submit the content to the same certificate authority that is tied to this key ring for signing.

Step 20 In the **Navigation** pane, choose **Public Key Management** > **Key Rings** > ***key_ring_name***.

Step 21 In the **Work** pane, in the **Certificate** field, paste the signed certificate that you received from the certificate authority.

Step 22 Click **Submit**.

Note
If the CSR was not signed by the Certificate Authority indicated in the key ring, or if the certificate has MS-DOS line endings, an error message is displayed and the certificate is not accepted. Remove the MS-DOS line endings.

The key is verified, and in the **Work** pane, the **Admin State** changes to **Completed** and is now ready for use in the HTTP policy.

Step 23 On the menu bar, choose **Fabric** > **Fabric Policies**.

Step 24 In the Navigation pane, choose **Pod Policies** > **Policies** > **Management Access** > **default**.

Step 25 In the **Work** pane, in the **Admin Key Ring** drop-down list, choose the desired key ring.

Step 26 (Optional) For Certificate based authentication, in the **Client Certificate TP** drop-down list, choose the previously created Local User policy and click **Enabled** for **Client Certificate Authentication state**.

Step 27 Click **Submit**.
All web servers restart. The certificate is activated, and the non-default key ring is associated with HTTPS access.

What to do next

You must remain aware of the expiration date of the certificate and take action before it expires. To preserve the same key pair for the renewed certificate, you must preserve the CSR as it contains the public key that pairs with the private key in the key ring. Before the certificate expires, the same CSR must be resubmitted. Do not delete or create a new key ring as deleting the key ring will delete the private key stored internally on the Cisco APIC.

Provisioning Fabric Wide System Settings

Configuring APIC In-Band or Out-of-Band Connectivity Preferences

This topic describes how to toggle between in-band and out-of-band connectivity on the APIC server for management access to devices such as authentication servers or SNMP servers external to the ACI fabric. Enabling **inband** executes in-band management connectivity between the APIC server to external devices through leaf switches on the ACI fabric. Enabling **ooband** executes out-of-band management connectivity between the APIC server to external devices through connections external to the ACI fabric.

Before you begin

Configure in-band and out-of-band management networks. For more information, see *Management* in the *Cisco APIC Basic Configuration Guide, Release 3.x*.

Procedure

Step 1 On the menu bar, click **System** > **System Settings**.

Step 2 On the Navigation bar, click **APIC Connectivity Preferences**.

Step 3 To enable the policy, click **inband** or **ooband**.

Step 4 Click **Submit**.

Configure Quota Management Policies

Starting in the Cisco Application Policy Infrastructure Controller (APIC) Release 2.3(1), there are limits on number of objects a tenant admin can configure. This enables the admin to limit the number of managed objects that can be added globally across tenants.

This feature is useful when you want to limit any tenant or group of tenants from exceeding ACI maximums per leaf or per fabric or unfairly consuming a majority of available resources, potentially affecting other tenants on the same fabric.

Procedure

Step 1 On the menu bar, click **System** > **System Settings**.

Step 2 Right-click **Quota** and choose **Create Quota Configuration**..

Step 3 In the **Class** field, choose the object type to limit with the quota.

Step 4 In the **Container Dn** field, enter the distinguished name (DN) that describes the class.

Step 5 In the **Exceed Action** field, choose either **Fail Transaction Action** or **Raise Fault Action**.

Step 6 In the **Max Number** field, enter the maximum number of the managed objects that can be created after which the exceed action will be applied.

Step 7 Click **Submit**.

Create an Enforced BD Exception List

This topic describes how to create a global exception list of subnets which are not subject to an enforced bridge domain. With the Enforced BD feature configured, the endpoints in a subject endpoint group (EPG) can only ping subnet gateways within the associated bridge domain.

The exception IP addresses can ping all of the BD gateways across all of your VRFs.

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.

Before you begin

Create an enforced bridge domain (BD).

Procedure

Step 1 On the menu bar, click **System** > **System Settings**.

Step 2 Click **BD Enforced Exception List**.

Step 3 Click the + on **Exception List**.

Step 4 Add the IP address and network mask for the subnet that can ping any subnet gateway.

Step 5 Repeat to add more subnets that are exceptions to the enforced bridge domain.

Step 6 Click **Submit**.

Create a BGP Route Reflector Policy and Route Reflector Node Endpoints

This topic describes how to create ACI fabric route reflectors, which use multiprotocol BGP (MP-BGP) to distribute external routes within the fabric. To enable route reflectors in the ACI fabric, the fabric administrator must select the spine switches that will be the route reflectors, and provide the autonomous system (AS) number. Once route reflectors are enabled in the ACI fabric, administrators can configure connectivity to external networks.

Before you begin

Required:

- To connect external routers to the ACI fabric, the fabric infrastructure administrator must configure spine nodes as Border Gateway Protocol (BGP) route reflectors.
- For redundancy purposes, more than one spine is configured as a router reflector node (one primary and one secondary reflector).

Procedure

Step 1 To create a BGP Route Reflector policy, perform the following steps:

a) On the menu bar, click **System** > **System Settings**.
b) Click **BGP Route Reflector**.
c) Enter the Autonomous System Number.
d) Click the + on **Route Reflector Nodes**.
e) Enter the spine route reflector node ID endpoint, and click **Submit**.,

Step 2 To create external route reflector node endpoints, perform the following steps:

a) Click the + on **External Route Reflector Nodes**.
b) Choose the spine to serve as external route reflector node endpoint.
c) If this is a site managed by Multi-Site, you can also specify an intersite spine route reflector.
d) Click **Submit**.

Configure a Fabric Wide Control Plane MTU Policy

This topic describes how to create a fabric-wide Control Plane (CP) MTU policy, that sets the global MTU size for control plane packets sent by the nodes (switches) in the fabric.

In a multipod topology, the MTU setting for the fabric external ports must be greater than or equal to the CP MTU value set. Otherwise, the fabric external ports might drop the CP MTU packets.

Note

- If you set the L3Out Interface Profile to inherit the MTU from the IPN, it will be 9150. If you want the MTU to be used across the IPN to be 9216, you must explicitly configure it in the L3Out Interface Profile (at **Tenants** > ***tenant-name*** > **Networking** > **External Routed Networks** > **Create Routed Outside** > **Nodes and Interface Protocol Profiles** > **Create Node Profile** > **Create Interface Profile**).
- Cisco APIC will always establish a TCP connection to fabric switches with an MTU of 1496 bytes (TCP MSS 1456) regardless of the CP-MTU setting. The IPN network for remote pods and remote leaf switches must support at least 1500 byte MTU for fabric discovery.

If you change the IPN or CP MTU, Cisco recommends changing the CP MTU value first, then changing the MTU value on the spine of the remote pod. This reduces the risk of losing connectivity between the pods due to MTU mismatch.

Procedure

Step 1 On the menu bar, click **System** > **System Settings**.
Step 2 Click **Control Plane MTU**.
Step 3 Enter the MTU for fabric ports.
Step 4 Click **Submit**.

Configure Endpoint Loop Protection

The endpoint loop protection policy specifies how loops detected by frequent MAC moves are handled. To configure EP loop protection perform the following steps:

Procedure

Step 1 On the menu bar, click **System** > **System Settings**.
Step 2 Click **Endpoint Controls**.
Step 3 Click the **Ep Loop Protection** tab.
Step 4 To enable the policy, click **Enabled** in the **Administrative State** field.
Step 5 Optional. Set the loop detection interval, which specifies the time to detect a loop. The interval range is from 30 to 300 seconds. The default setting is 60 seconds.
Step 6 Set the loop detection multiplication factor, which is the number of times a single EP moves between ports within the loop detection interval. The range is from 1 to 255. The default is 4.
Step 7 Choose the action to take when detecting a loop.

The action can be:

- **BD Learn Disable**
- **Port Disable**

The default is **Port Disable**.

Step 8 Click **Submit**.

Rogue Endpoint Control Policy

About the Rogue Endpoint Control Policy

A rogue endpoint attacks leaf switches through frequently, repeatedly injecting packets on different leaf switch ports and changing 802.1Q tags (thus, emulating endpoint moves) causing learned class and EPG port changes. Misconfigurations can also cause frequent IP and MAC address changes (moves).

Such rapid movement in the fabric causes significant network instability, high CPU usage, and in rare instances, endpoint mapper (EPM) and EPM client (EPMC) crashes due to significant and prolonged messaging and transaction service (MTS) buffer consumption. Also, such frequent moves may result in the EPM and EPMC logs rolling over very quickly, hampering debugging for unrelated endpoints.

The rogue endpoint control feature addresses this vulnerability by quickly:

- Identifying such rapidly moving MAC and IP endpoints.
- Stopping the movement by temporarily making endpoints static, thus quarantining the endpoint.
- Prior to 3.2(6) release: Keeping the endpoint static for the **Rogue EP Detection Interval** and dropping the traffic to and from the rogue endpoint. After this time expires, deleting the unauthorized MAC or IP address.
- In the 3.2(6) release and later: Keeping the endpoint static for the **Rogue EP Detection Interval** (this feature no longer drops the traffic). After this time expires, deleting the unauthorized MAC or IP address.
- Generating a host tracking packet to enable the system to re-learn the impacted MAC or IP address.
- Raising a fault to enable corrective action.

The rogue endpoint control policy is configured globally and, unlike other loop prevention methods, functions at the level of individual endpoints (IP and MAC addresses). It does not distinguish between local or remote moves; any type of interface change is considered a move in determining if an endpoint should be quarantined.

The rogue endpoint control feature is disabled by default.

Limitations of the Rogue Endpoint Control Policy

The following limitations apply when using a rogue endpoint control policy:

- Changing rogue endpoint control policy parameters will not affect existing rogue endpoints.
- If a rogue endpoint is enabled, loop detection and bridge domain move frequency will not take effect.
- Disabling the rogue endpoint feature clears all rogue endpoints.
- The endpoint mapper (EPM) has value limits for rogue endpoint parameters. If you set the parameter values outside of this range, the Cisco APIC raises a fault for each mismatched parameter.
- The rogue endpoint feature is not supported on remote leaf switches or Cisco ACI Multi-Site.

- You must disable rogue endpoint control before you upgrade to or from Cisco APIC release 4.1.

Configuring the Rogue Endpoint Control Policy Using the GUI

You can configure the **Rogue EP Control** policy for the fabric, to detect and delete unauthorized endpoints, using the Cisco Application Policy Infrastructure Controller (Cisco APIC) GUI. This topic also includes the steps to clear rogue endpoints on a TOR switch, ad-hoc.

The policy options have the following valid and supported values:

- **Rogue EP Detection Interval**—Sets the rogue endpoint detection interval, which specifies the time to detect rogue endpoints. Valid values are from 0 to 65535 seconds. The default is 60.
- **Hold Interval (sec)**—Interval in seconds after the endpoint is declared rogue, where it is kept static so learning is prevented and the traffic to and from the Rogue endpoint is dropped. After this interval, the endpoint is deleted. Valid values are from 1800 to 3600. The default is 1800.
- **Rogue EP Detection Multiplication Factor**—Sets the rogue endpoint detection multiplication factor for determining if an endpoint is unauthorized. If the endpoint moves more times than this number, within the EP detection interval, the endpoint is declared rogue. Valid values are from 2 to 10. The default is 6.

Procedure

Step 1 On the menu bar, click **System** > **System Settings**.

Step 2 On the navigation bar, click **Endpoint Controls** and click the **Rogue EP Control** tab.

Step 3 Set the **Administrative State** to **Enabled**.

Step 4 Optional. Reset the **Rogue EP Detection Interval (sec)**, **Rogue EP Detection Multiplication Factor**, or the **Hold Interval (sec)**.

Step 5 (Optional) To clear rogue endpoints on a TOR switch, perform the following steps:

a) On the Cisco APIC menu bar, click **Fabric** > **Inventory**.
b) On the Navigation bar, expand the Pod and click the leaf switch where you want to clear rogue endpoints.
c) When the leaf switch summary appears in the work pane, right-click the leaf switch name in the Navigation bar, and choose **Clear Rogue Endpoints**.
d) Click **Yes**.

Configure the Rogue Endpoint Control Policy Using the NX-OS Style CLI

You can configure the rogue endpoint control policy for the fabric to detect and delete unauthorized endpoints using the NX-OS style CLI.

Procedure

Step 1 Enter global configuration mode.

Example:

```
apic1# configure
```

Step 2 Enable the global rogue endpoint control policy.

Example:

```
apic1(config)# endpoint rogue-detect enable
```

Step 3 Set the hold interval.

The hold interval is a period of time in seconds after the endpoint is declared rogue that the endpoint is kept static so that learning is prevented, and the traffic to and from the endpoint is dropped. After this interval, the endpoint is deleted. Valid values are from 1800 to 3600 seconds. The default is 1800.

Example:

```
apic1(config)# endpoint rogue-detect hold-interval 1800
```

Step 4 Set the detection interval.

The detection interval is a period of time in seconds during which rogue endpoint control counts the number of moves for an endpoint. If the count during this interval exceeds the value specified by the detection multiplication factor, the endpoint is declared rogue. Valid values are from 0 to 65535 seconds. The default is 60.

Example:

```
apic1(config)# endpoint rogue-detect interval 60
```

Step 5 Set the detection multiplication factor.

If an endpoint moves more times than the value specified by the detection multiplication factor during a period of time specified by the detection interval, the endpoint is declared rogue. Valid values are from 2 to 10. The default is 6.

Example:

```
apic1# endpoint rogue-detect factor 6
```

About the Rogue/COOP Exception List

The rogue/COOP exception list enables you to specify the MAC address of endpoints for which you want to have a higher tolerance for endpoint movement with rogue endpoint control before the endpoints get marked as rogue. Endpoints in the rogue/COOP exception list get marked as rogue only if they move 3,000 or more times within 10 minutes. After an endpoint is marked as rogue, the endpoint is kept static to prevent learning. The rogue endpoint is deleted after 30 seconds.

Beginning with the Cisco Application Policy Infrastructure Controller (APIC) 6.0(3) release, you can create a global rogue/COOP exception list, which excludes a MAC address from rogue endpoint control on all the bridge domains on which the MAC address is discovered, and you can create a rogue/COOP exception list for L3Outs. You can also exclude all MAC addresses for a bridge domain or L3Out. This simplifies creating the exception list when you want to make an exception for every MAC address; you do not need to enter each address individually.

Guidelines and Limitations for the Rogue/COOP Exception List

The following guidelines and limitations apply when using the rogue/COOP exception list:

- The MAC address exception list feature works on Layer 2 bridge domains (bridge domains without IP routing enabled). This is because on a Layer 3 bridge domain (a bridge domain with IP routing enabled),

if there are IP addresses moving along with MAC addresses, the IP addresses would end up being marked rogue first and both the IP and MAC address would then be quarantined.

- For a Layer 3 bridge domain, disable dataplane IP address learning per subnet for specific IP addresses that you want to exclude from rogue endpoint control.

  For information about the dataplane IP address learning per subnet feature, see the *Cisco APIC Layer 3 Networking Configuration Guide*.

- You are expected to have a full understanding of what kind of MAC addresses are being added to this list. You are responsible for ensuring that the MAC addresses in this list do not contribute to excessive moves in the overall fabric or per leaf switch.
- You can add up to 100 MAC addresses to the per-bridge domain exception list, fabric-wide. Beginning in the 6.0(3) release you can also exclude all MAC addresses of a given bridge domain from rogue endpoint control.
- Beginning in the 6.0(3) release, you can add 100 external L3Out MAC addresses to the exception list, fabric-wide. You can also exclude all MAC addresses of a L3Out SVI bridge domain from rogue endpoint control.
- Beginning in the 6.0(3) release, you can add up to 6,000 MAC addresses to the global exception list as long as the total count of endpoints for those MAC addresses is within 6,000. The same MAC address discovered on a number of bridge domains counts as that same number of bridge domains against the total of 6,000 endpoints. For example, if a configured MAC address appears as a discovered MAC endpoint on 10 bridge domains, the MAC address counts as 10 endpoints.
- The exception list exemption on leaf switches applies only when rogue endpoint control is enabled. When rogue endpoint control is disabled, the MAC address exception list is used only by COOP dampening.
- The rogue/COOP exception list can only contain MAC addresses in a bridge domain and not IP addresses in a VRF instance. However, an IP address-only move can still cause the IP address to be marked as rogue if it meets the regular rogue endpoint control criteria.
- To mask out IP address rogue detection and marking based on data path traffic, you can use bridge domain subnet learn disable. Bridge domain subnet learn disable stops Cisco ACI from learning the IP address location at each move.

Configuring the Rogue/COOP Exception List While Creating a Bridge Domain Using the GUI

The following procedure configures the rogue/COOP exception list while you create a bridge domain:

Before you begin

- You must have a tenant for which you will create a bridge domain.
- Rogue endpoint control must be enabled. For the procedure about enabling rogue endpoint control, see Configuring the Rogue Endpoint Control Policy Using the GUI, on page 26.

Procedure

Step 1 In the desired tenant, create a bridge domain. In the menu bar, choose **Tenants** > ***tenant_name***.

Step 2 In the Navigation pane, choose **Networking** > **Bridge Domains**.

Step 3 Right-click **Bridge Domains** and choose **Create Bridge Domain**.

Step 4 In the **Create Bridge Domain** dialog, fill out the fields as desired for **STEP 1 > MAIN** and **STEP 2 > L3 Configurations**.

Step 5 For **STEP 3 > Advanced/Troubleshooting**, for the **Rogue/Coop Exception List**, click the **+**, enter the MAC address of an endpoint that you want to add to the list, and click **Update**.

The format of the MAC address is `AA:BB:CC:DD:EE:FF`.

a) Repeat this step for each endpoint that you want to add to the list.

Step 6 (Optional): Beginning with the Cisco Application Policy Infrastructure Controller (APIC) 6.0(3) release, put a check in the **Enable Rogue Exception MAC for BD** check box to cause rogue endpoint control to ignore all MAC addresses from the bridge domain.

Step 7 Fill out the remaining fields for **STEP 3 > Advanced/Troubleshooting** as desired.

Step 8 Click **Finish**.

Configuring the Rogue/COOP Exception List of an Existing Bridge Domain Using the GUI

The following procedure configures the rogue/COOP exception list of an existing bridge domain.

Before you begin

- You must have a tenant with a bridge domain.
- Rogue endpoint control must be enabled. For the procedure about enabling rogue endpoint control, see Configuring the Rogue Endpoint Control Policy Using the GUI, on page 26.

Procedure

Step 1 In the menu bar, choose **Tenants** > ***tenant_name***.

Step 2 In the Navigation pane, choose **Networking** > **Bridge Domains** > ***bridge_domain_name***.

Step 3 In the Work pane, choose **Policy** > **Advanced/Troubleshooting**.

Step 4 For the **Rogue/Coop Exception List**, click the **+**, enter the MAC address of an endpoint that you want to add to the list, and click **Update**.

The format of the MAC address is `AA:BB:CC:DD:EE:FF`.

a) Repeat this step for each endpoint that you want to add to the list.

Step 5 (Optional): Beginning with the Cisco Application Policy Infrastructure Controller (APIC) 6.0(3) release, put a check in the **Enable Rogue Exception MAC for BD** check box to cause rogue endpoint control to ignore all MAC addresses from the bridge domain.

Step 6 Click **Submit**.

Configuring the Rogue Endpoint Control Exception List on an L3Out SVI Using the GUI

The procedure in this section configures which MAC addresses the rogue endpoint control feature should ignore for a selected switch virtual interface (SVI) of an L3Out.

Before you begin

- You must have a tenant that has an L3Out.
- Rogue endpoint control must be enabled. For the procedure about enabling rogue endpoint control, see Configuring the Rogue Endpoint Control Policy Using the GUI, on page 26

Procedure

Step 1 In the menu bar, choose **Tenants >** > ***tenant_name***.

Step 2 In the Navigation pane, choose **Networking** > **L3Outs** > ***L3Out_name*** > **Logical Node Profiles** > ***node_profile_name*** > **Logical Interface Profiles** > ***interface_profile_name***.

Step 3 In the Work pane, choose **Policy** > **SVI**.

Step 4 In the SVI table, click +.

Step 5 In the **Select SVI** dialog, perform either of the following actions:

- For **Rogue Exception MAC Group**, choose an existing group or create a new group.

  The rogue exception MAC group specifies which MAC addresses the rogue endpoint control feature should ignore for all SVIs that have the same VLAN encapsulation.

- Put a check in the **Exclude all MACs from Rogue EP Control** box.

  Enabling this option causes the rogue endpoint control feature to ignore all MAC addresses for all SVIs that have the same VLAN encapsulation.

Step 6 Fill out the remaining fields as desired.

Step 7 Click **Submit**.

Configuring MAC Address Exceptions for Rogue Endpoint Control for the entire Fabric Using the GUI

Beginning with the 6.0(3) release, at the fabric level, you can configure which MAC addresses rogue endpoint control should ignore. Rogue endpoint control ignores any MAC address that you specify regardless of to which bridge domian or L3Out SVI the MAC address belongs.

Before you begin

Rogue endpoint control must be enabled. For the procedure about enabling rogue endpoint control, see Configuring the Rogue Endpoint Control Policy Using the GUI, on page 26.

Procedure

Step 1 In the menu bar, choose **Fabric** > **Fabric Policies**.

Step 2 In the Navigation pane, choose **Policies** > **Global** > **Fabric Wildcard Rogue Exception**.

Step 3 In the Work pane, in the **Wildcard Rogue Exception** table, click +.

a) Enter a MAC address for rogue endpoint control to ignore, then click **Update**.
b) Repeat this step for each MAC address that you want rogue endpoint control to ignore.

About Max IP Address Flow Control

The 3.2(6) release adds the max IP address flow control feature, which identifies endpoints that are misbehaving and flags them as rogue based on the number of learned IP addresses that are associated with a MAC address. The Cisco Application Centric Infrastructure (ACI) fabric supports a maximum of 4,096 IP addresses on a MAC address. If a leaf switch learns more than 4,096 IP addresses that are associated with a MAC address, then the MAC address and all of the IP addresses are classified as rogue.

After the max IP address flow control feature identifies an endpoint as rogue, the endpoint is quarantined, a fault is raised in the APIC, and there will not be any further learning of new IP Addresses on this endpoint. The quarantine period is 1 hour. If the standard rogue feature is enabled, then the quarantine period is the same as the period configured by the standard rogue configuration.

While the rogue endpoint control policy feature (rogue due to moves) can be configured to be enabled or disabled, the max IP address flow control feature does not require explicit configuration to be enabled.

Prior to this feature, the ACI fabric identified an endpoint as rogue if it kept moving its location for a configurable number of times within a configurable time period. With this feature, the ACI fabric can identify an endpoint as rogue based on the number of moves or if a leaf switch learns more than 4,096 IP addresses on a MAC address.

Configure COOP

About COOP

Council of Oracle Protocol (COOP) is used to communicate the mapping information (location and identity) to the spine proxy. A leaf switch (the "citizen") forwards endpoint address information to the spine switch (the "oracle") using Zero Message Queue (ZMQ). COOP running on the spine nodes will ensure all spine nodes maintain a consistent copy of endpoint address and location information and additionally maintain the distributed hash table (DHT) repository of endpoint identity to location mapping database.

COOP Endpoint Dampening

When malicious or erroneous behavior causes unnecessary endpoint updates, the COOP process can become overwhelmed, preventing the processing of valid endpoint updates. The rogue endpoint detection feature of the leaf switch can prevent many erroneous updates from reaching the spine. In cases where the rogue endpoint detection is inadequate, the COOP process invokes endpoint dampening. To relieve pressure on COOP, the spine asks all leaf switches to ignore updates from the misbehaving endpoint for a specified period. When this occurs, the dampening state of the endpoint is 'Freeze,' and a fault is generated.

Note COOP endpoint dampening is introduced in Cisco APIC Release 4.2(3).

COOP Authentication

COOP data path communication provides high priority to transport using secured connections. To protect COOP messages from malicious traffic injection, the APIC controller and switches support COOP protocol authentication.

The COOP protocol supports two ZMQ authentication modes:

- **Strict mode:** COOP allows MD5 authenticated ZMQ connections only.
- **Compatible mode:** COOP accepts both MD5 authenticated and non-authenticated ZMQ connections for message transportation.

For additional information about COOP authentication, see the *Cisco APIC Security Configuration Guide*.

Viewing COOP Dampened Endpoints Using the GUI

Use this APIC GUI procedure to view all dampened endpoints in a spine node.

Procedure

Step 1 On the menu bar, click **Fabric > Inventory**.

Step 2 In the Navigation pane, expand the pod and the spine node.

Step 3 Expand **Protocols > COOP** and the **COOP** instance.

Step 4 Click **Endpoint Database** to display the endpoints.

Inspect the **Dampened State** column to find the dampened endpoints. The possible states are:

- **Normal:** The endpoint updates are normal.
- **Critical:** Enough updates have been received that the endpoint could be moved into the freeze state.
- **Freeze:** Updates from this endpoint are currently being ignored due to frequent unnecessary updates. A fault has been generated.

Viewing COOP Dampened Endpoints Using the Switch CLI

Use this switch CLI procedure to view all dampened endpoints in a spine or leaf node.

Log in to the spine or leaf switch CLI and enter the following command:

```
show coop internal info repo ep dampening
```

Clearing COOP Dampened Endpoints Using the GUI

Use this APIC GUI procedure to clear and recover all dampened endpoints in a spine or leaf node.

Procedure

Step 1 On the menu bar, click **Fabric** > **Inventory**.

Step 2 In the Navigation pane, expand the pod and the spine or leaf node.

Step 3 Right-click the node and choose **Clear Dampened Endpoints**.

Step 4 Click **Yes** to confirm the action.

Clearing a COOP Dampened Endpoint Using the Switch CLI

Use this procedure to clear and recover a dampened endpoint in a spine or leaf node. The procedure clears a single endpoint whose dampening state is 'Freeze.'

Log in to the spine or leaf switch CLI and enter the following command:

```
clear coop internal info repo ep dampening key <bd> <mac>
```

Disabling COOP Endpoint Dampening Using the REST API

This procedure shows how to disable or enable COOP EP dampening using the APIC REST API.

COOP endpoint dampening is enabled by default, but in some situations it can be necessary to disable it. An example is when you are expecting many IP updates for a single MAC address and ignoring those updates could be disruptive to the network.

Use the following API, setting `disableEpDampening="true"` to disable COOP endpoint dampening.

```
<!-- api/policymgr/mo/.xml -->

<polUni>
    <infraInfra>
        <infraSetPol disableEpDampening="true"></infraSetPol>
    </infraInfra>
</polUni>
```

All nodes in the fabric will disable COOP endpoint dampening and will recover any existing endpoints whose dampened state is 'Freeze.'

Configuring COOP Authentication Using the APIC GUI

Procedure

Step 1 On the menu bar, choose **System** > **System Settings**.

Step 2 In the **Navigation** pane, click on **COOP Group**.

Step 3 In the **Work** pane, under the **Policy Property** area in the **Type** field, choose the desired type from the **Compatible Type** and **Strict Type** options.

Step 4 Click **Submit**.

This completes the COOP authentication policy configuration.

Configuring COOP Authentication Using the Cisco NX-OS-Style CLI

Procedure

Configure the COOP authentication policy using the strict mode option.

Example:

```
apic1# configure
apic1(config)# coop-fabric
apic1(config-coop-fabric)# authentication type ?
compatible  Compatible type
strict      Strict type
apic101-apic1(config-coop-fabric)# authentication type strict
```

Configuring COOP Authentication Using the REST API

Procedure

Configure a COOP authentication policy.

In the example, the strict mode is chosen.

Example:

```
https://172.23.53.xx/api/node/mo/uni/fabric/pol-default.xml

<coopPol type="strict">
</coopPol>
```

Endpoint Listen Policy

About the Endpoint Listen Policy

You can configure an endpoint listen policy to detect untagged traffic that gets sent from anonymous endpoints to Cisco Application Centric Infrastructure (ACI) leaf switches that do not have an enforced policy. By default, if a policy is not deployed for a port, all endpoint traffic gets dropped on that port. If you configure an endpoint listen policy, this policy gets deployed on all leaf switch ports that do not have an existing enforced policy. The endpoint listen policy enables Cisco ACI to detect untagged traffic that arrives on those ports, such that Cisco ACI will know the MAC address or IP address of the anonymous endpoints. This allows the Cisco ACI administrator to decide in which EPG to put those endpoints. The Cisco Application Policy Infrastructure Controller (APIC) GUI displays all detected anonymous endpoints on the **Global Endpoints** configuration screen.

Note The endpoint listen policy is a beta feature. There is no guarantee that this feature will work as intended. Use at your own risk.

Configuring the Endpoint Listen Policy Using the GUI

This procedure configures an endpoint listen policy, which detects untagged traffic that gets sent from anonymous endpoints to Cisco Application Centric Infrastructure (ACI) leaf switches that do not have an enforced policy.

Note The endpoint listen policy is a beta feature. There is no guarantee that this feature will work as intended. Use at your own risk.

Procedure

Step 1 On the menu bar, choose **System** > **System Settings**.

Step 2 In the Navigation pane, choose **Global Endpoints**.

Step 3 In the Work pane, put a check in the **End Point Listen Policy** check box.

Step 4 In the **End Point Listen Encap** drop-down list, choose **VLAN**.

Step 5 In the **End Point Listen Encap** text field, enter the VLAN ID. Valid values are from 1 to 4094. This should be a reserved VLAN encap, such that it can not be used by any EPGs.

Step 6 Click **Submit**.

Configure IP Aging

This topic describes how to enable an IP Aging policy. When enabled, the IP aging policy ages unused IPs on an endpoint.

When the Administrative State is enabled, the IP aging policy sends ARP requests (for IPv4) and neighbor solicitations (for IPv6) to track IPs on endpoints. If no response is given, the policy ages the unused IPs.

Procedure

Step 1 On the menu bar, click **System** > **System Settings**.

Step 2 Click **Endpoint Controls**.

Step 3 Click the **Ip Aging** tab.

Step 4 To enable the policy, click **Enabled** in the **Administrative State** field.

What to do next

Create an End Point Retention policy, which is required, to specify the timer used for tracking IPs on endpoints. Navigate to **Tenants** > ***tenant-name*** > > **Policies** > **Protocol** > **End Point Retention**.

Disable Remote Endpoint Learning

This topic describes how to enable or disable IP end point learning.

The scope of this policy is fabric-wide. After configuration, a policy is pushed to each leaf switch as it comes up.

You should enable this policy in fabrics which include the Cisco Nexus 9000 series switches, 93128 TX, 9396 PX, or 9396 TX switches with the N9K-M12PQ uplink module, after all the nodes have been successfully upgraded to APIC Release 2.2(2x) or higher.

After any of the following configuration changes, you may need to manually flush previously learned IP endpoints:

- Remote IP endpoint learning is disabled
- The VRF is configured for ingress policy enforcement
- At least one Layer 3 interface exists in the VRF

To manually flush previously learned IP endpoints, enter the following command on both VPC peers: `vsh -c "clear system internal epm endpoint vrf <vrf-name> remote"`

To enable or disable IP end point learning, perform the following steps:

Procedure

Step 1 On the menu bar, click **System** > **System Settings**.

Step 2 Click **Fabric Wide Setting**.

Step 3 Click the check box on **Disable Remote EP Learn**.

Step 4 Click **Submit**.

Globally Enforce Subnet Checks

This topic describes how to enable or disable subnet checking. When enabled, IP address learning is disabled outside of subnets configured in a VRF, for all other VRFs.

The scope of this policy is fabric-wide. After configuration, a policy is pushed to each leaf switch as it comes up.

Procedure

Step 1 On the menu bar, click **System** > **System Settings**.

Step 2 Click **Fabric Wide Setting**.

Step 3 Click the check box on **Enforce Subnet Check**.

Step 4 Click **Submit**.

Reallocate a GIPo

This topic describes how to enable reallocating GIPos on non-stretched bridge domains to make room for stretched bridge domains.

With the introduction of Cisco ACI Multi-Site, there was a need to change the GIPo allocation scheme to provide the following benefits:

1. Minimize the number of bridge domains that have the same GIPo.
2. GIPos that are assigned to Cisco ACI Multi-Site stretched bridge domains do not overlap with GIPos that are assigned to non-stretched bridge domains.

To achieve this allocation, Cisco ACI introduced different pools whose sizes change based on amount of stretched and non-stretched bridge domains.

For a fresh installation of Cisco ACI, the Cisco APIC guarantees that both #1 and #2 are accomplished. During a Cisco ACI upgrade from a release prior to 2.3(1), the old schema is maintained to avoid fabric disruption because the existing GIPos might already be used for non-stretched bridge domains. As a result, Cisco ACI cannot guarantee that #2 is accomplished.

Enabling the **Reallocate GIPo** knob in Cisco APIC's fabric-wide setting policy causes Cisco APIC to re-allocate GIPos and use the newer allocation scheme. Enabling the knob is a one-time operation. Afterward, the GIPos will not overlap. This knob is relevant only in a Cisco ACI Multi-Site Orchestrator deployment if you upgrade from a release that is earlier than 2.3(1) to the 3.0(1) release or later.

The scope of this policy is fabric-wide. After configuration, a policy is pushed to each leaf switch as it comes up.

Procedure

Step 1 On the menu bar, click **System** > **System Settings**.

Step 2 Click **Fabric Wide Setting**.

Step 3 Click the check box on **Reallocate Gipo**.

Step 4 Click **Submit**.

Globally Enforce Domain Validation

This topic describes how to enforce domain validation. When enabled, a validation check is performed when a static path is added, to determine if no domain is associated with an EPG.

The scope of this policy is fabric-wide. After configuration, a policy is pushed to each leaf switch as it comes up.

Procedure

Step 1 On the menu bar, click **System** > **System Settings**.

Step 2 Click **Fabric Wide Setting**.

Step 3 Click the check box on **Enforce Domain Validation**.

Step 4 Click **Submit**.

Enable OpFlex Client Authentication

This topic describes how to enable OpFlex client authentication for GOLF and Linux.

To deploy GOLF or Linux Opflex clients in an environment where the identity of the client cannot be guaranteed by the network, you can dynamically validate the client's identity based on a client certificate.

Note When you enable certificate enforcement, connectivity with any GOLF or Linux Opflex client that does not support client authentication is disabled.

The scope of this policy is fabric-wide. After configuration, a policy is pushed to each leaf switch as it comes up.

Procedure

Step 1 On the menu bar, click **System** > **System Settings**.

Step 2 Click **Fabric Wide Setting**.

Step 3 Click the check box on **OpFlex Client Authentication** to enable or disable enforcing client certificate authentication for GOLF and Linux Opflex clients.

Step 4 Click **Submit**.

Fabric Load Balancing

The ACI fabric provides several load balancing options for balancing the traffic among the available uplink links. This topic describes load balancing for leaf to spine switch traffic.

Static hash load balancing is the traditional load balancing mechanism used in networks where each flow is allocated to an uplink based on a hash of its 5-tuple. This load balancing gives a distribution of flows across the available links that is roughly even. Usually, with a large number of flows, the even distribution of flows results in an even distribution of bandwidth as well. However, if a few flows are much larger than the rest, static load balancing might give suboptimal results.

ACI fabric Dynamic Load Balancing (DLB) adjusts the traffic allocations according to congestion levels. It measures the congestion across the available paths and places the flows on the least congested paths, which results in an optimal or near optimal placement of the data.

DLB can be configured to place traffic on the available uplinks using the granularity of flows or flowlets. Flowlets are bursts of packets from a flow that are separated by suitably large gaps in time. If the idle interval between two bursts of packets is larger than the maximum difference in latency among available paths, the second burst (or flowlet) can be sent along a different path than the first without reordering packets. This idle interval is measured with a timer called the flowlet timer. Flowlets provide a higher granular alternative to flows for load balancing without causing packet reordering.

DLB modes of operation are aggressive or conservative. These modes pertain to the timeout value used for the flowlet timer. The aggressive mode flowlet timeout is a relatively small value. This very fine-grained load balancing is optimal for the distribution of traffic, but some packet reordering might occur. However, the overall benefit to application performance is equal to or better than the conservative mode. The conservative mode flowlet timeout is a larger value that guarantees packets are not to be re-ordered. The tradeoff is less granular load balancing because new flowlet opportunities are less frequent. While DLB is not always able to provide the most optimal load balancing, it is never worse than static hash load balancing.

Note Although all Nexus 9000 Series switches have hardware support for DLB, the DLB feature is not enabled in the current software releases for second generation platforms (switches with EX, FX, and FX2 suffixes).

The ACI fabric adjusts traffic when the number of available links changes due to a link going off-line or coming on-line. The fabric redistributes the traffic across the new set of links.

In all modes of load balancing, static or dynamic, the traffic is sent only on those uplinks or paths that meet the criteria for equal cost multipath (ECMP); these paths are equal and the lowest cost from a routing perspective.

Dynamic Packet Prioritization (DPP), while not a load balancing technology, uses some of the same mechanisms as DLB in the switch. DPP configuration is exclusive of DLB. DPP prioritizes short flows higher than long flows; a short flow is less than approximately 15 packets. Because short flows are more sensitive to latency than long ones, DPP can improve overall application performance.

For intra-leaf switch traffic, all DPP-prioritized traffic is marked CoS 0 regardless of a custom QoS configuration. For inter-leaf switch traffic, all DPP-prioritized traffic is marked CoS 3 regardless of a custom QoS configuration.

GPRS tunneling protocol (GTP) is used mainly to deliver data on wireless networks. Cisco Nexus switches are places in Telcom Datacenters. When packets are being sent through Cisco Nexus 9000 switches in a datacenter, traffic needs to be load-balanced based on the GTP header. When the fabric is connected with an external router through link bundling, the traffic is required to be distributed evenly between all bundle members (For example, Layer 2 port channel, Layer 3 ECMP links, Layer 3 port channel, and L3Out on the port channel). GTP traffic load balancing is performed within the fabric as well.

To achieve GTP load balancing, Cisco Nexus 9000 Series switches use 5-tuple load balancing mechanism. The load balancing mechanism takes into account the source IP, destination IP, protocol, Layer 4 resource and destination port (if traffic is TCP or UDP) fields from the packet. In the case of GTP traffic, a limited number of unique values for these fields restrict the equal distribution of traffic load on the tunnel.

To avoid polarization for GTP traffic in load balancing, a tunnel endpoint identifier (TEID) in the GTP header is used instead of a UDP port number. Because the TEID is unique per tunnel, traffic can be evenly load balanced across multiple links in the bundle.

The GTP load balancing feature overrides the source and destination port information with the 32-bit TEID value that is present in GTPU packets.

GTP tunnel load balancing feature adds support for:

- GTP with IPv4/IPv6 transport header on physical interface
- GTPU with UDP port 2152

The ACI fabric default configuration uses a traditional static hash. A static hashing function distributes the traffic between uplinks from the leaf switch to the spine switch. When a link goes down or comes up, traffic on all links is redistributed based on the new number of uplinks.

Leaf/Spine Switch Dynamic Load Balancing Algorithms

The following table provides the default non-configurable algorithms used in leaf/spine switch dynamic load balancing:

Table 1: ACI Leaf/Spine Switch Dynamic Load Balancing

Traffic Type	Hashing Data Points
Leaf/Spine IP unicast	• Source MAC address • Destination MAC address • Source IP address • Destination IP address • Protocol type • Source Layer 4 port • Destination Layer 4 port • Segment ID (VXLAN VNID) or VLAN ID
Leaf/Spine Layer 2	• Source MAC address • Destination MAC address • Segment ID (VXLAN VNID) or VLAN ID

Creating a Load Balancer Policy Using the Cisco APIC GUI

This topic describes how to configure the default Load Balancer policy.

The load balancing policy options balance traffic among the available uplink ports. Static hash load balancing is the traditional load balancing mechanism used in networks where each flow is allocated to an uplink based on a hash of its 5-tuple. This load balancing gives a distribution of flows across the available links that is roughly even. Usually, with a large number of flows, the even distribution of flows results in an even distribution of bandwidth as well. However, if a few flows are much larger than the rest, static load balancing might give suboptimal results.

Procedure

Step 1 On the menu bar, click **System** > **System Settings**.

Step 2 Click **Load Balancer**.

Step 3 Choose the **Dynamic Load Balancing Mode**.

The dynamic load balancer (DLB) mode adjusts the traffic allocations according to congestion levels. It measures the congestion across the available paths and places the flows on the least congested paths, which results in an optimal or near optimal placement of the data. DLB can be configured to place traffic on the available uplinks using the granularity of flows or of flowlets. Flowlets are bursts of packets from a flow that are separated by intervals. The mode can be **Aggressive**, **Conservative**, or **Off** (the default).

Step 4 Enable or disable **Dynamic Packet Prioritization** by choosing **On** or **Off** (the default).

Dynamic Packet Prioritization (DPP) prioritizes short flows higher than long flows; a short flow is less than approximately 15 packets. Short flows are more sensitive to latency than long ones. DPP can improve overall application performance.

Step 5 Choose the Load Balancing Mode. The mode can be **Link Failure** or **Traditional** (the default).

The load balancer administrative state. In all modes of load balancing, static or dynamic, the traffic is sent only on those uplinks or paths that meet the criteria for equal cost multipath (ECMP); these paths are equal and the lowest cost from a routing perspective.

Step 6 Click **Submit**.

Creating a Load Balancer Policy Using the CLI

Creating a Dynamic Load Balancer Policy Using the CLI

There are two dynamic load balancer modes: **dynamic-aggressive** and **dynamic-conservative**. The **dynamic-aggressive** mode enables a shorter flowlet timeout interval, and the **dynamic-conservative** mode enables a longer flowlet timeout interval. For more information about these commands, see the *Cisco APIC NX-OS Style CLI Command Reference*.

This section demonstrates how to configure a dynamic load balancer policy using the CLI.

Procedure

Step 1 To enable aggressive mode dynamic load balancing:

```
apic1# conf t
apic1# (config)# system dynamic-load-balance mode dynamic-aggressive
```

Step 2 To enable conservative mode dynamic load balancing:

```
apic1# conf t
apic1# (config)# system dynamic-load-balance mode dynamic-conservative
```

Creating a Dynamic Packet Prioritization Policy Using the CLI

This section demonstrates how to enable dynamic packet prioritization using the CLI. For more information about this command, see the *Cisco APIC NX-OS Style CLI Command Reference*.

Procedure

Enable dynamic packet prioritization:

```
apic1# conf t
apic1# (config)# system dynamic-load-balance mode packet-prioritization
```

Creating a GTP Load Balancer Policy Using the CLI

This section demonstrates how to create a GTP load balancer policy using the CLI. For more information about this command, see the *Cisco APIC NX-OS Style CLI Command Reference*.

Procedure

Enable dynamic packet prioritization:

```
apic1# conf t
apic1# (config)# ip load-sharing address source_destination gtpu
```

Creating a Load Balancer Policy Using the REST API

This section demonstrates how to enable a DLB, DPP, and a GTP load balancer policy. For a list of all possible property values, see the *Cisco APIC Management Information Model Reference*.

Procedure

To enable a DLB, DPP, and GTP load balancer policy:

```
https://apic-ip-address/api/mo/uni.xml
<polUni>
<fabricInst>
 <lbpPol name="default" hashGtp="yes" pri="on" dlbMode="aggressive">
 </lbpPol>
</fabricInst>
</polUni>
```

Enable a Time Precision Policy

This topic describes how to enable Precision Time Protocol (PTP), a time synchronization protocol for nodes distributed across a network. Its hardware timestamp feature provides greater accuracy than other time synchronization protocols such as Network Time Protocol (NTP).

PTP is a distributed protocol that specifies how real-time PTP clocks in the system synchronize with each other. These clocks are organized into a master-member synchronization hierarchy with the grandmaster clock, the clock at the top of the hierarchy, determining the reference time for the entire system. Synchronization is achieved by exchanging PTP timing messages, with the members using the timing information to adjust their clocks to the time of their master in the hierarchy. PTP operates within a logical scope called a PTP domain.

Procedure

Step 1 On the menu bar, click **System** > **System Settings**.

Step 2 Click **Precision Time Protocol**.

Step 3 Choose **Enabled** or **Disabled**.

If you choose disable PTP, NTP time is used to sync the fabric. If you enable PTP, a spine is automatically chosen as a master to which the entire site gets synced.

Step 4 Click **Submit**.

Enable a Global System GIPo Policy

This topic describes how to use the infra tenant GIPo as the system GIPo.

An ACI multipod deployment requires the 239.255.255.240 system Global IP Outside (GIPo) to be configured on the inter-pod network (IPN) as a PIM BIDIR range. This 239.255.255.240 PIM BIDIR range configuration on the IPN devices can be avoided by using the infra GIPo as System GIPo.

Before you begin

Upgrade all of the switches in the ACI fabric, including the leaf switches and spine switches, to the latest APIC release.

Procedure

Step 1 On the menu bar, click **System** > **System Settings**.

Step 2 Choose **Enabled** or **Disabled** (the default) on **Use Infra GIPo as System GIPo**

Step 3 Click **Submit**.

Configure a Fabric Port Tracking Policy

Uplink failure detection can be enabled in the fabric access fabric port tracking policy. The port tracking policy monitors the status of links between leaf switches and spine switches. When an enabled port tracking policy is triggered, the leaf switches take down all access interfaces on the switch that have EPGs deployed on them. For more information about fabric port tracking, see the *Cisco APIC Layer 2 Networking Configuration Guide*.

Procedure

Step 1 On the menu bar, click **System** > **System Settings**.

Step 2 On the Navigation pane, choose **Port Tracking**.

Step 3 Enable port tracking by setting the **Port tracking state** to **on**.

Step 4 (Optional) Change the **Daily restore timer** value.

Step 5 Configure the **Number of active spine links that triggers port tracking** parameter.

Step 6 Click **Submit**.

Provisioning Global Fabric Access Policies

Create a Global Attachable Access Entity Profile

An Attachable Entity Profile (AEP) represents a group of external entities with similar infrastructure policy requirements. The infrastructure policies consist of physical interface policies that configure various protocol options, such as Cisco Discovery Protocol (CDP), Link Layer Discovery Protocol (LLDP), or Link Aggregation Control Protocol (LACP).

An AEP is required to deploy VLAN pools on leaf switches. Encapsulation blocks (and associated VLANs) are reusable across leaf switches. An AEP implicitly provides the scope of the VLAN pool to the physical infrastructure.

The following AEP requirements and dependencies must be accounted for in various configuration scenarios, including network connectivity, VMM domains, and multipod configuration:

- The AEP defines the range of allowed VLANs but it does not provision them. No traffic flows unless an EPG is deployed on the port. Without defining a VLAN pool in an AEP, a VLAN is not enabled on the leaf port even if an EPG is provisioned.
- A particular VLAN is provisioned or enabled on the leaf port that is based on EPG events either statically binding on a leaf port or based on VM events from external controllers such as VMware vCenter or Microsoft Azure Service Center Virtual Machine Manager (SCVMM).
- Attached entity profiles can be associated directly with application EPGs, which deploy the associated application EPGs to all those ports associated with the attached entity profile. The AEP has a configurable generic function (infraGeneric), which contains a relation to an EPG (infraRsFuncToEpg) that is deployed on all interfaces that are part of the selectors that are associated with the attachable entity profile.

A virtual machine manager (VMM) domain automatically derives physical interface policies from the interface policy groups of an AEP.

Before you begin

Create the tenant, VRF instance, application profiles, and EPGs to associate to the attached entity profile.

Procedure

Step 1 On the menu bar, click **Fabric** > **Access Policies**.

Step 2 On the navigation bar, expand **Policies** and **Global**.

Step 3 Right-click **Attachable Access Entity Profile** and choose **Create Attachable Access Entity Profile**.

Step 4 Enter a name for the policy.

Step 5 Click the + icon on **Domains** table.

Step 6 Enter a physical domain, a previously created physical, Layer 2, Layer 3, or Fibre Channel domain, or create one.

Step 7 Enter the encapsulation for the domain and click **Update**.

Step 8 Click the + icon on the **EPG DEPLOYMENT** table.

Step 9 Enter the tenant, application profile, EPG, encapsulation (such as vlan-1), primary encapsulation (primary encapsulation number) and interface mode (trunk, Access (802.1P, or Access (Untagged).

Step 10 Click **Update**.

Step 11 Click **Next**.

Step 12 Choose the interfaces to associate to the attachable entity profile.

Step 13 Click **Finish**.

Configure the Global QoS Class Policy

The global QoS Class policy can be used to:

- Preserve the CoS priority level, to guarantee that the CoS value in 802.1P packets which enter and transit the ACI fabric is preserved. 802.1P CoS preservation is supported in single pod and multipod topologies. In multipod topologies, CoS Preservation can be used where you want to preserve the QoS priority settings of 802.1P traffic entering POD 1 and egressing out of POD 2, but you are not concerned with preserving the CoS/DSCP settings in interpod network (IPN) traffic between the pods. To preserve CoS/DSCP settings when multipod traffic is transitting an IPN, use a DSCP policy (configured at **Tenants** > **infra** > > **Policies** > **Protocol** > **DSCP class-cos translation policy for L3 traffic**)
- Reset the properties for the default QoS class levels, such as the **MTU**, **Queue Limit**, or **Scheduling Algorithm**.

Procedure

Step 1 On the menu bar, click **Fabric** > **Access Policies**.

Step 2 On the navigation bar, expand **Policies** and **Global**.

Step 3 Click **QOS Class**.

Step 4 To enable 802.1P CoS preservation, click the **Preserve COS** check box.

Step 5 To change the default settings for a QoS class, double-click on it. Enter the new settings and click **Submit**.

Create a Global DHCP Relay Policy

The global DHCP relay policy identifies a DHCP relay server that can be referenced by bridge domains in different tenants across the fabric. With tenant DHCP relay you must configure the DHCP relay policy in each tenant even if the provider EPG is in the common tenant. The global DHCP relay policy differs from a tenant DHCP relay policy in that it can be referenced from any tenant and does not require the configuration of a DHCP relay policy in each tenant.

Procedure

Step 1 On the menu bar, click **Fabric** > **Access Policies**.

Step 2 On the navigation bar, expand **Policies** and **Global**.

Step 3 Right-click **DHCP Relay** and choose **Create DHCP Relay Policy**.

Step 4 Enter a name for the policy.

Step 5 Click the + icon on **Providers**.

Step 6 Choose the EPG type, and for an application EPG, choose the tenant, application profile, and the EPG to be the provider.

Step 7 In the **DHCP Server Address** field, enter the IP address for the server.

Step 8 Click **OK**.

Enable a Global MCP Instance Policy

Enable a global Mis-Cabling Protocol (MCP) instance policy. In the current implementation, only one instance of MCP runs in the system.

Procedure

Step 1 On the menu bar, click **Fabric** > **Access Policies**.

Step 2 On the navigation bar, expand **Policies** and **Global**.

Step 3 Click **MCP Instance Policy default**.

Step 4 Change the **Admin State** to **Enabled**.

Step 5 Set other properties as needed for your fabric.

Step 6 Click **Submit**.

What to do next

Create an Error Disabled Recovery Policy

The error disabled recovery policy specifies the policy for re-enabling a port that was disabled due to one or more pre-defined error conditions.

Procedure

Step 1 On the menu bar, click **Fabric** > **Access Policies**.
Step 2 On the navigation bar, expand **Policies** and **Global**.
Step 3 Click **Error Disabled Recovery Policy**..
Step 4 Double-click on an event to enable it for the recovery policy.
Step 5 Click the check box and click **Update**.
Step 6 Optional. Repeat steps 4 and 5 for more events.
Step 7 Optional. Reset the **Error disable recovery interval (sec)**.
Step 8 Click **Submit.**

Per Port Policies

About Per Port Policies

A per port policy is an implicit policy that you use to configure the interfaces of a leaf switch using the Cisco Application Policy Infrastructure Controller (APIC) GUI. A per port policy is simplified compared to the standard policy-based model, which is useful when you are still learning how to use the Cisco APIC. Due to this simplification, you cannot add new ports to an existing policy. Instead, you can only create new chunks of a policy per interface.

The per port policy pane uses the NX-OS CLI in the background to create implicit and explicit objects. For example, creating a new port channel creates an explicit port channel policy group as well as an implicit override. Making any changes to an explicit policy group will not apply to port until the implicit policy group is removed. We recommend that you do not mix using the CLI and GUI. Use the per port policy wizard to learn about the Cisco Application Centric Infrastructure (ACI) policy model and unconfigure the port from the same wizard when you want to move to advanced use cases for reusable policy configurations.

You can create a per port policy only from the following GUI location:

Fabric > **Inventory** > ***Pod-#*** > ***leaf-switch-name*** > **Interface tab**

Note **Interface tab** refers to the **Interface** tab in the work pane. This is not the **Interfaces** folder in the navigation pane.

Configuring a Per Port Policy Using the GUI

This procedure configures a per port policy using the Cisco Application Policy Infrastructure Controller (APIC) GUI.

Procedure

Step 1 On the menu bar, choose **Fabric** > **Inventory**.

Step 2 In the Navigation pane, choose ***pod-#*** > ***leaf-switch-name***.

Step 3 In the Work pane, choose the **Interface** tab.

Step 4 In the **Mode** drop-down list, choose **Configuration**.

Step 5 Click on one or more of the interface numbers to select those interfaces.

The buttons just under the tabs of the Work pane become active for any of the components that you can configure for the selected interfaces.

Step 6 Click the button for one of the components that you want to configure.

The Work pane displays the properties for that component.

Step 7 Set the properties as desired for the component.

Step 8 Click **Submit**.

Step 9 Configure any additional components for the selected interfaces, or select different interfaces and configure the components.

Validating a Per Port Policy Using the GUI

This procedure instructs you on how to validate a per port policy using the Cisco Application Policy Infrastructure Controller (APIC) GUI.

Before you begin

You must configure the Cisco APIC to show hidden policies. By default, the per port policies are hidden in the Cisco APIC.

Procedure

Step 1 On the menu bar, choose **Fabric** > **Inventory**.

Step 2 In the Navigation pane, choose ***pod-#*** > ***leaf-switch-name***.

Step 3 In the Work pane, choose the **Interface** tab.

Step 4 In the **Mode** drop-down list, choose **Configuration**.

Step 5 Click on an interface numbers to select that interface.

The buttons just under the tabs of the Work pane become active for any of the components that you can configure for the selected interface.

Step 6 Click the button for one of the components for which you want to view the properties.

The Work pane displays the properties for that component.

Step 7 Verify that the properties are set correctly, and change any values that are not correct for your desired configuration.

Step 8 If you made any changes, click **Submit**. Otherwise, click **Cancel**.

Showing the Hidden Policies Using the GUI

By default, some policies, such as per port policies, are hidden in the Cisco Application Policy Infrastructure Controller (APIC). If you want to view these policies, you must configure the Cisco APIC to show hidden policies.

Procedure

Step 1 In the upper right corner of the GUI, choose **Manage My Profile** > **Settings**.

The **Application Settings** dialog opens.

Step 2 Put a check in the **Show Hidden Policies** box.

Step 3 Click **OK**.

Creating a Mis-cabling Protocol Interface Policy Using the GUI (Optional)

The mis-cabling protocol (MCP) was designed to handle misconfigurations that Link Layer Discovery Protocol (LLDP) and Spanning Tree Protocol (STP) are unable to detect. MCP has a Layer 2 packet that it uses, and MCP disables ports that form a loop within the Fabric. Cisco Application Centric Infrastructure (ACI) fabric leaf switches do not participate in spanning tree protocol (STP) and act as hub with respect to STP. The MCP packet is sent, and if the fabric sees that the packet comes back in, then the fabric knows that there is a loop and the fabric takes action based on that event. Faults and events are generated when this happens. MCP can be enabled globally and per-interface. By default, MCP is disabled globally and is enabled on each port. For MCP to work, it must be enabled globally, regardless of the per-interface configuration.

The following procedure creates an MCP interface policy using the GUI.

Procedure

Step 1 On the menu bar, choose **Fabric** > **Access Policies**.

Step 2 In the Navigation pane, choose **Policies** > **Interface** > **MCP Interface**.

Step 3 In the Work pane, choose **Actions** > **Create Mis-cabling Protocol Interface Policy**.

Step 4 In the **Create Mis-cabling Protocol Interface Policy** dialog box, perform the following actions:

a) Enter a name for the policy.
b) (Optional) Enter a description for the policy.
c) For **Admin State**, choose **Enable** to enable the policy, or **Disable** to disable the policy.
d) Choose the **Strict** or **Non-strict** operational mode for MCP.

The following additional fields are displayed when you select **Strict**.

- **Initial Delay Time (sec)**: Time for STP convergence on the external Layer 2 network. Default value is 0 (when STP is disabled on the Layer 2 network). When STP is enabled, based on scale/ topology, the initial delay time range is 45 to 60 seconds for STP to converge.
- **Tranmission Frequency (sec, msec)**: Transmission frequency timer for MCP packets until grace period on each Layer 2 interface. Default value is 500 milliseconds.
- **Grace Period (sec, msec)**: Grace period time during which early loop detection takes place. The port aggressively transmits MCP packets which are used for loop detection. Default grace period value is 3 seconds.

Step 5 Click **Submit**.