



## Deploying Cisco Virtual APIC Using VMware vCenter

[New and Changed Information](#) 2

[Overview](#) 2

[Modes of Deployment](#) 2

[Guidelines and Limitations for Deploying Virtual APIC on ESXi](#) 5

[Deploying a virtual APIC Using VMware vCenter](#) 6

[Creating an ACI Network with a Layer 3 Connected APIC Cluster](#) 9

[Migrating from a Physical APIC to a Virtual APIC](#) 19

Revised: April 19, 2024

## New and Changed Information

The following table provides an overview of the significant changes up to this current release. The table does not provide an exhaustive list of all changes or of the new features up to this release.

Cisco APIC Release Version	Feature
6.0(2)	Support for deploying virtual APIC using VMware vCenter.

## Overview

Beginning with Cisco APIC release 6.0(2), you can deploy a cluster wherein all the APICs in the cluster are virtual APICs. You can deploy a virtual APIC on AWS using the CloudFormation template, or a virtual APIC on an ESXi host using the OVF template in VMware Center. The virtual APIC can be deployed on an existing server on the customer premises.

This document provides details about deploying a virtual APIC using ESXi; for details about deploying virtual APIC using AWS, see the [Deploying Cisco Virtual APIC Using AWS](#) document.

## Modes of Deployment

Two modes of deployment are supported:

- Layer 2—the ESXi host, on which the virtual APIC(s) are deployed, is directly connected to the leaf switches of the ACI fabric. The ESXi hosts hosting the APICs can be connected to the ACI leaf switches with active-standby uplinks or active-active uplinks using LACP port channels.
- Layer 3—the ESXi host, on which the virtual APIC(s) are deployed, is remotely attached to the ACI fabric, via an external network.

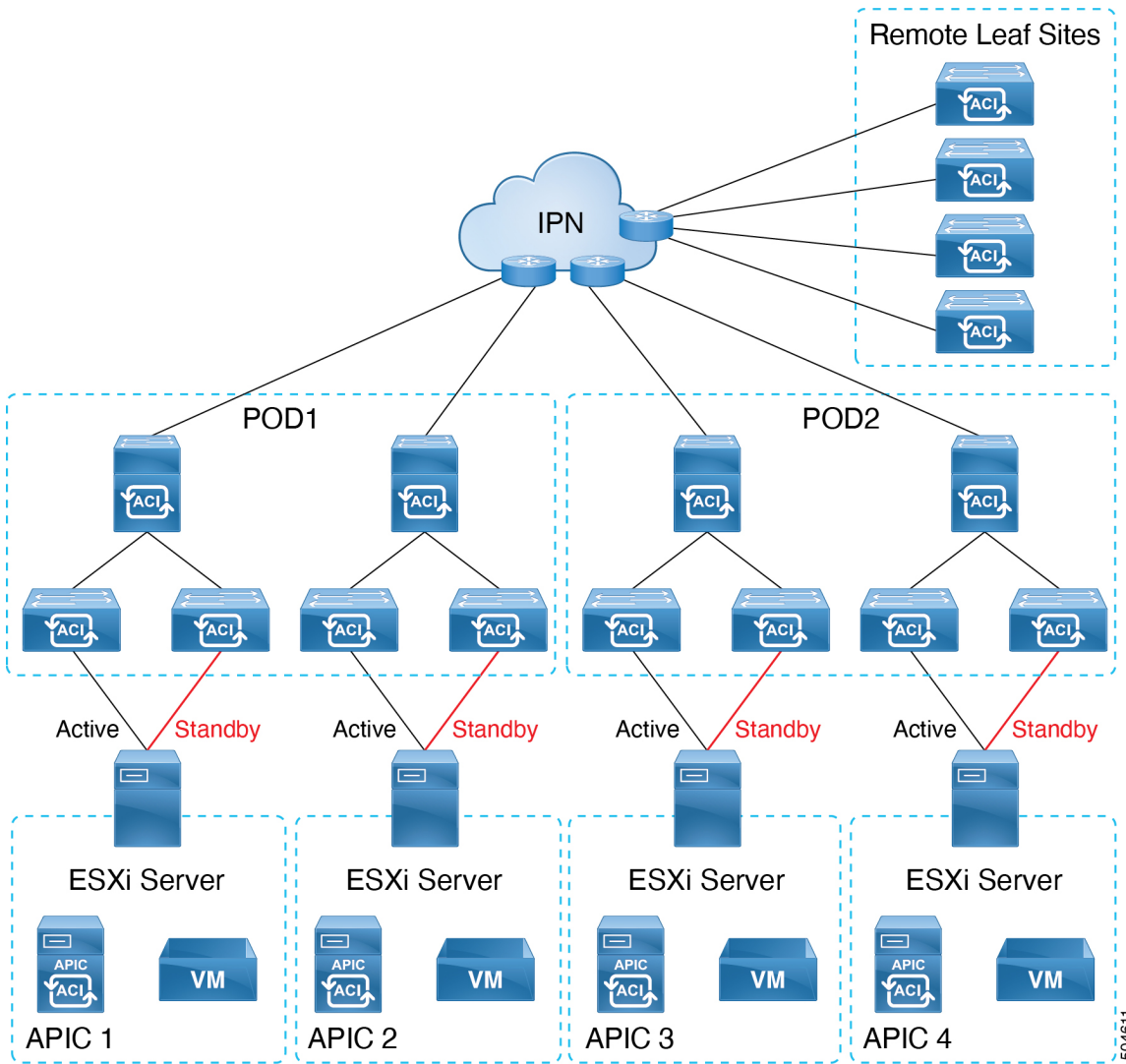
### ESXi Host is directly connected to the ACI Fabric

In this mode, referred to as Layer 2 connected, wherein the ESXi host is directly connected to the leaf switches of the ACI fabric. The ESXi host(s) hosting APICs in Layer 2 mode can be connected in active-active or active-standby mode. However, the ESXi host hosting APIC 1 should only use active-standby mode.

#### Virtual APICs with active-standby ESXi uplinks

As shown in the image below, the APIC VMs are hosted on ESXi hosts which are directly connected to the leaf switch(es) of the ACI fabric. One of the ESXi host uplinks connected to the leaf switch is in active mode, the other is in standby mode

Figure 1: Virtual APICs Directly Connected to the ACI Fabric (with Active-Standby ESXi Uplinks)



### Virtual APICs with active-active ESXi uplinks

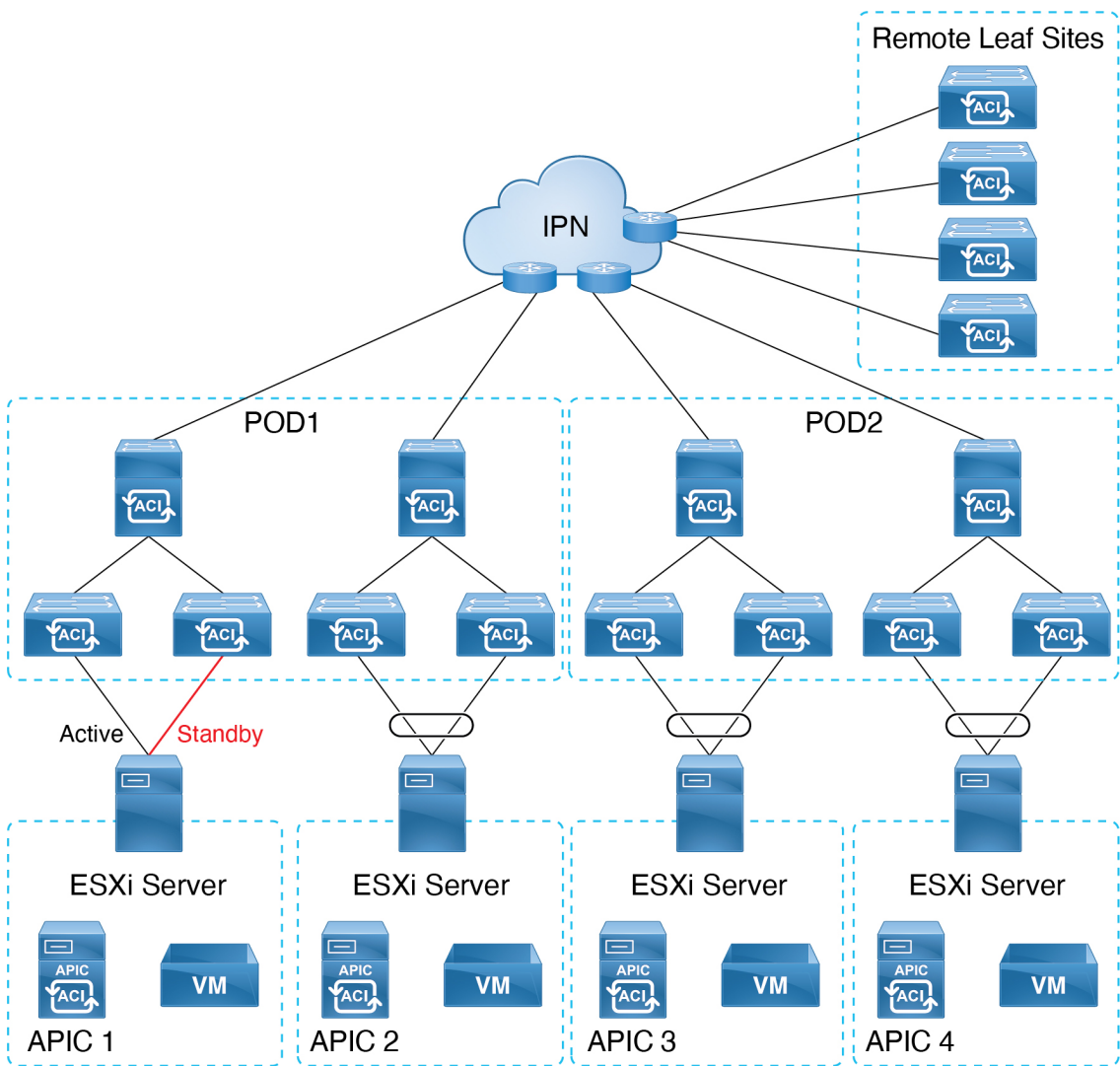
As shown in the image below, the APIC VMs are hosted on ESXi hosts which are directly connected to the leaf switch of the ACI fabric. Both uplinks of the ESXi host are connected to the leaf switch as an LACP port channel.



**Note** LACP requires vSphere Distributed Virtual Switch (DVS). LACP is not supported on vSphere Standard Switch.

APIC 1 of the cluster should be connected with active-standby ESXi uplinks. The other APICs of the cluster, APIC 2 to N (where N is the cluster size) can be connected with active-active ESXi uplinks.

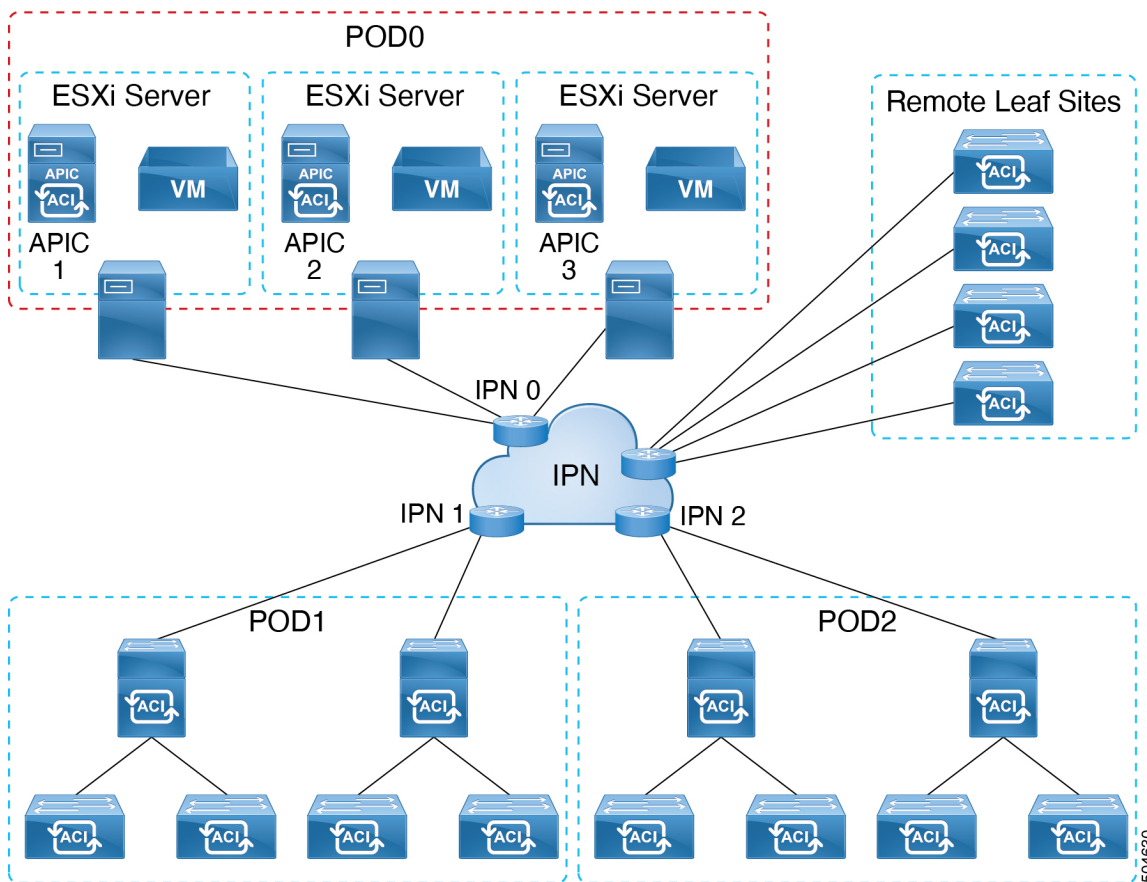
Figure 2: Virtual APICs Directly Connected to the ACI Fabric (with Active-Active ESXi Uplinks)



## ESXi Host is remotely attached to the ACI Fabric

As shown in the image below, the APICs VMs are hosted on the ESXi hosts which are connected to an external network and remotely attached to the ACI fabric through IPN.

Figure 3: Virtual APICs Remotely Attached to the ACI Fabric



## Guidelines and Limitations for Deploying Virtual APIC on ESXi

The following guidelines and limitations apply for deploying virtual APIC on ESXi:

- Although multiple virtual APICs per ESXi host are supported; for high availability, a single virtual APIC per ESXi host is recommended.
- Fabric switches must be running ACI release 6.0(2) or later. Fabric switches running versions earlier than 6.0(2) can be automatically upgraded to the 6.0(2) release during fabric discovery using Auto Firmware Update.
- No support for mixed mode, that is, all the APICs in each cluster must either be of the same type (layer 2 or layer 3). Virtual APIC on Layer 3 ESXi cannot form a cluster with Virtual APIC on Layer 3 AWS or physical APIC on Layer 3.
- Before deploying Virtual APIC on ESXi, ensure all the ESXi host clocks are synchronized using NTP.
- Virtual APICs deployed using ESXi cannot be migrated to AWS, and vice-versa (deployed on AWS to ESXi). For the supported migration scenario, see the [Migrating from a Physical APIC to a Virtual APIC](#) section.
- After you have deployed a virtual APIC (with release 6.0(2)) using ESXi, you cannot downgrade it to a release prior to Cisco APIC release 6.0(2).
- No support for standby APIC.

In the case of a physical APIC cluster, the standby APIC is automatically updated with firmware updates to ensure it is the same firmware version as the active cluster. This enables APIC replacement in the event of a failure. However, in the case of a virtual APIC cluster, the user can create an instance of the APIC with the same version as needed. See the [Replacing APIC Controllers](#) section in the Cisco APIC Getting Started Guide.

- Cluster and fabric security is provided using self-signed certificates.
- For VMM deployment in the same DVS used for hosting the virtual APICs, enable CDP, and disable LLDP on Cisco APIC GUI. Navigate to **Virtual Networking > VMware > DVS**.
- ESXi hosts hosting virtual APIC VMs must be directly connected to the ACI leaf switches. For a vAPIC directly connected to the fabric, it needs to consume LLDP traffic to/from the leafs. Hence, LLDP is disabled on the vDS, to avoid interception. To operate the vDS for VMM, enable CDP. See the Networking Prerequisites section in the [Deploying a virtual APIC Using VMware vCenter, on page 6](#) procedure for details.

ESXi hosts connected over an intermediate switch, including UCS Fabric Interconnects, are not supported.

- Support only for pre-packaged apps. The supported pre-packaged apps are:
  - Network Insights Cloud Connector
  - Pre-Upgrade Validator
- Virtual APIC cluster is supported in all Multi-Pod, Remote Leaf, and Multi-Site topologies. Refer the [Verified Scalability Guide for Cisco APIC](#) for scalability limitations.
- Virtual APIC supports VMware vMotion.
- Virtual APIC does not support VMware vSphere High Availability (HA) nor VMware vSphere Fault Tolerance (FT).

## Deploying a virtual APIC Using VMware vCenter

Use this procedure for deploying a virtual APIC on an ESXi host using the VMware vCenter GUI.

The initial deployment is done by using the `.ova` image (which is available under Software Downloads on Cisco.com). After the cluster is brought up (either using the GUI or using APIs), and it is *fully-fit*, the APICs can be upgraded to future versions using the `.iso` image.

### Before you begin

#### Networking Prerequisites

Ensure that the following network-related prerequisites are met before deploying the virtual APIC on an ESXi:

- Log in to the VMware vCenter GUI and edit the Distributed Virtual Switch (DVS) settings as follows:
  - Disable LLDP and/or enable CDP, based on the following scenarios:
    - Configure the discovery protocol type to either CDP or (disabled) when the DVS is used without the VMM integration. This configuration will ensure that LLDP packets are passed through the distributed virtual switch to the APIC VM.
    - Configure the discovery protocol type as CDP when the DVS is used with VMM integration.



---

**Note** LLDP is not supported on the Standard Virtual Switch. Standard Virtual Switch supports CDP.

---

- Configure VLANs.

Configure the New Distributed Port Group for virtual APICs. Enable VLAN trunking and set the VLAN range. Add the following VLANs to the VLAN trunk range:

- Infra VLAN
- Inband Management VLAN (if Inband management is configured)
- VLAN 0 (VLAN 0 is required for LLDP forwarding LLDP packets. LLDP packets sent from the leaf switch will be untagged)

- Add ESXi hosts to DVS.

Assign the physical adapter vmnics that connect to the fabric leaf(s) to the DVS uplinks, for example Uplink1 and Uplink2.

- In addition to the ones mentioned above, the following prerequisites are applicable when the virtual APICs are connected with active-active ESXi uplinks (except APIC 1, which will always be connected with active-standby ESXi uplinks):
  - If active-active is used for ESXi hosts hosting APICs 2 to N (*N* is the cluster size), create a LAG at the DVS assigning the vmnics connected to the leaf switches as the LAG members. Add the LAG to the active-active port group.




---

**Note** Active-active and active-standby uplinks cannot be configured in the same DVS port group. Using both active-active and active-standby configurations will require two port groups.

---

- Connect ESXi Uplinks to the ACI leaf(s).

- (for remotely attached virtual APICs) Enable trunking and set the VLAN range to include Layer 3 APIC infra VLANs and inband management VLANs used for connecting to the routed IP Network (IPN).

### Virtual Machine Prerequisites

Ensure that the ESXi hosts can support the VM specifications mentioned in the table below.

It is recommended to have 600GB (including the swap space) of storage.

CPU	16 vCPU of 3 GHz or Higher
Memory	96 GB of RAM
Storage	<ul style="list-style-type: none"> <li>• Disk 1: SSD or NVMe – 120GB (root disk)</li> <li>• Disk 2: SSD or NVMe – 360GB (empty data disk)</li> <li>• I/O latency of 20ms</li> </ul>
Network	<p>Two Interfaces.</p> <ul style="list-style-type: none"> <li>• OOB 1Gbps or higher</li> <li>• Infra / Inband 10Gbps or higher</li> </ul> <p>Latency tolerance between virtual APICs is up to 50 ms.</p>





---

**Note** Virtual APIC cluster supports the medium CPU, hard drive, and memory configuration (up to 1,200 edge ports).

---

## Procedure

---

**Step 1** Select **Deploy OVF template**. There are six steps (screens) to be completed. Enter the following details in each of these screens:

- a) In the **Select an OVF template** field, select the URL from where the .ova file needs to be picked up from. You can download the relevant .ova file from the software downloads page and save it in the required location. You can also select a file from your local machine.
- b) Click **Next**.
- c) In the **Select a name and folder** screen, enter the **Virtual Machine Name**. This is the identification of the virtual APIC that will be deployed.

You will also need to select the folders from the **Select** location in the **Virtual Machine** pane.

- d) In the **Select a compute resource** screen, select the ESXi on which you want to deploy the virtual APIC.
- e) Click **Next**.
- f) In the **Review details** page, verify the details and click **Next**.
- g) In the **Select storage** screen, select the datastore you want to use for the deployment, and for the **Select virtual disk format** field, select the recommended **Thick Provision Lazy Zeroed** option from the drop-down list. This reserves the required amount of disk space for the VM.
- h) Click **Next**.
- i) In the **Select networks** screen, select the Destination Network information from the drop-down list.
  - **OOB Network** - Provide the port group which connects to the Out of Band (OOB) management network.
  - **Infra Network** - Provide the port group that is used for connecting to the leaf switch(es), for Layer 2 (directly attached to the ACI fabric) or provide the port group that is used for connecting to the IPN in case of Layer 3 (remotely attached to the ACI fabric).
- j) Click **Next**.
- k) In the **Customize template** screen, enter the following:
  - **Admin password** which you will be using to log in to the created virtual APIC.
  - **OOB IP Address** (with subnet) and the **OOB Management Gateway**. Make a note of this address as this is the IP address (IPv4 or IPv6) which you will later need to access the **APIC Cluster Bringup** screen.
- l) Click **Next**.
- m) Review the details in the **Ready to Complete** screen. Click **Finish**.

Check the VMs tab, the newly created virtual APIC appears under the selected ESXi.

- n) Click the virtual APIC and the details are displayed on the right. Power on the VM.
- o) To check the status of the deployment, navigate to the **Monitor** tab, and select **Tasks**.

**Step 2** Repeat the above step based on the number of nodes you want in a cluster. For example, if you are building a 3-node cluster, you need to run the step three times. Ensure that all the VMs are powered on before proceeding to the bootstrapping



and cluster bringup procedure. See the detailed [Bringing up the Cisco APIC Cluster Using the GUI](#) procedure in the *Cisco APIC Getting Started Guide*. Use the IP address of APIC 1 to access the APIC Cluster Bringup GUI.

---

### What to do next

If you want to change the APICs that are connected with active-standby ESXi uplinks to active-active ESXi uplinks (except for APIC 1), use the procedure detailed below.

1. After APIC1 is *fully fit*, register all the fabric nodes in APIC1 using the node registration policy in the Cisco APIC GUI.  
Navigate to **Fabric > Inventory > Fabric Membership** for node registration.
2. Connect the port-group with active-active ESXi uplinks (LAG) for APICs 2 to *N* in the DVS using the VMware vCenter GUI.
3. Configure VPC policy group for the remaining virtual APIC connected interfaces using the LACP policy.  
In the Port Channel Policy page of the Cisco APIC GUI, delete the **Suspend Individual Port** option in the **Control** field and select the mode as LACP active.  
Navigate to **Fabric > Access Policies > Policies > Interface > Port Channel**.
4. Verify all APICs are *Fully fit* (Cisco APIC GUI).  
Navigate to **System > Dashboard**. In the **Controller Status** pane (bottom right), the **Health State** column should display *Fully fit*.

## Creating an ACI Network with a Layer 3 Connected APIC Cluster

The procedures in this section establish the connectivity between the created virtual APIC cluster, and the remote ACI fabric.

The layer 3 connected APIC cluster is able to discover the fabric nodes using DHCP relay and an OSPF or BGP underlay provided by the IPN.

The following list outlines the steps to deploy an ACI network with a layer 3 connected virtual APIC cluster:

### Procedure

---

- |               |  |
|---------------|--|
| <b>Step 1</b> | Configure the IPN as described in these procedures — <a href="#">Provisioning the Fabric-Facing IPN Device, on page 11</a> and <a href="#">Provisioning the APIC Cluster-Facing IPN Device, on page 10</a> . |
| <b>Step 2</b> | Bring up the APIC cluster using the <i>Bringing up the APIC Cluster</i> procedure described in the <i>Cisco APIC Getting Started Guide</i> .   |
| <b>Step 3</b> | Configure a layer 3 connection for the APIC cluster to communicate over the IPN with the fabric pod. See <a href="#">Preparing Connectivity to the Fabric Pod, on page 13</a> .                              |
| <b>Step 4</b> | Bring up the fabric pod. The fabric will be discovered by the APIC cluster over the layer 3 connection as described in <a href="#">Summary of Fabric Discovery and Registration, on page 15</a> .            |
- 

### What to do next

You can connect additional fabric pods and remote leaf sites to the layer 3 connected APIC cluster in a similar manner.

## Guidelines and Restrictions for Deploying APIC Cluster Connectivity to the Fabric Over a Layer 3 Network

When deploying a virtual layer 3-connected APIC cluster, follow these guidelines and limitations.

- Ensure that the APIC connected port-group is configured for the APIC Infra Layer 3 Network VLAN. If in-band management is used, the port-group should be configured as a trunk allowing both the Infra Layer 3 Network VLAN and the in-band management VLAN.
- All APIC cluster sizes are supported in a layer 3 connected APIC pod.
- APICs in a layer 3 connected APIC pod cannot form a cluster with APICs within the fabric pod. In this topology, there should be no APICs in the fabric pod.
- The layer 3 connected APICs can be in the same subnet or in different subnets. In case of same subnet, configure "no ip redirects" in APIC-connected IPN interface.
- The layer 3 connected APICs can be geographically distributed from each other provided that the latency between APICs and with the fabric pod does not exceed 50 milliseconds round-trip time (RTT), which translates approximately to a geographical distance of up to 2,500 miles.
- Although any device that can meet the IPN network requirements can be used as an IPN device, we recommend to deploy, when possible, switches of the Cisco Nexus 9300 Cloud Scale family. These are the devices most commonly found in production and also the devices more frequently validated in Cisco internal testing. For further information about IPN device requirements, see "Inter-Pod Connectivity Deployment Considerations" in the [ACI Multi-Pod White Paper](#).
- The APIC subnets must be advertised to the spines as either OSPF or BGP routes. Both OSPF and BGP are supported as underlay protocols between the APIC nodes and the IPN devices
- As all control plane traffic between the APIC cluster and the fabric pod traverses the IPN, we recommend configuring QoS for this traffic. See the [Configuring QoS for the Layer 3 Connected APIC Cluster](#) section in this guide.
- APIC Cluster Connectivity to the Fabric Over a Layer 3 Network does not support the following:
  - ACI CNI for Kubernetes (Redhat Openshift, SUSE/Rancher RKE, Upstream Kubernetes on Ubuntu)
  - ACI ML2 for Openstack (Redhat Openstack, Canonical Openstack)
- APIC Cluster Connectivity to the Fabric Over a Layer 3 Network supports strict mode. In strict mode, you must approve the controller explicitly.

## Provisioning the APIC Cluster-Facing IPN Device

This section describes the configuration of the IPN device connected to Pod 0, the APIC cluster pod. With reference to the topology for Layer 3 in the [ESXi Host is remotely attached to the ACI Fabric , on page 4](#) section, the cluster-facing IPN device is shown as IPN0. As a recommended practice, the IPN0 comprises two devices for redundancy. The fabric interface of each APIC is dual-homed to the two devices. In the following configuration example, two Cisco Nexus 9000 series switches (IPN0a and IPN0b) are configured with the following choices:

- VLAN 1500 is used as interface VLAN for the APICs.
- The switch interfaces are configured as layer 2 trunk ports. As an alternative, the interfaces could be access ports if the APIC fabric interface is configured to use VLAN 0 during APIC setup.
- Both switches are configured using HSRP to share a single IP address that serves as the APIC subnet default gateway address.

- APIC subnets are advertised to the spines using OSPF as the underlay protocol. As an alternative, a BGP underlay could be deployed.

# Example configuration of IPN0a:

```
interface Vlan1500
  no shutdown
  vrf member IPN
  ip address 172.16.0.252/24
  ip ospf passive-interface
  ip router ospf 1 area 0.0.0.0
  hsrp version 2
  hsrp 1500
    ip 172.16.0.1

interface Ethernet1/1
  switchport mode trunk
  switchport trunk allowed vlan 1500
  spanning-tree port type edge trunk
```

# Example configuration of IPN0b:

```
interface Vlan1500
  no shutdown
  vrf member IPN
  ip address 172.16.0.253/24
  ip ospf passive-interface
  ip router ospf 1 area 0.0.0.0
  hsrp version 2
  hsrp 1500
    ip 172.16.0.1

interface Ethernet1/1
  switchport mode trunk
  switchport trunk allowed vlan 1500
  spanning-tree port type edge trunk
```

## Provisioning the Fabric-Facing IPN Device

This section describes the configuration of the MPod IPN, which is the IPN device connected to a fabric pod. The IPN is not managed by the APIC. It must be preconfigured with the following information:

- Configure the interfaces connected to the spines of the fabric pod. Use Layer 3 sub-interfaces tagging traffic with VLAN-4 and increase the MTU at least 50 bytes above the maximum MTU required for inter-site control plane and data plane traffic.
- Enable OSPF (or BGP) on the sub-interface specifying the OSPF process and area ID.
- Enable DHCP Relay on the IPN interfaces connected to spines.
- Enable PIM.
- Add bridge domain GIPo range as PIM Bidirectional (**bidir**) group range (default is 225.0.0.0/15).  
A group in **bidir** mode has only shared tree forwarding capabilities.
- Add 239.255.255.240/28 as PIM **bidir** group range.
- Enable PIM on the interfaces connected to all spines.



---

**Note** Multicast is not required for a single pod fabric with a layer 3-connected APIC cluster, but it is required between pods in a multi-pod fabric.

---



---

**Note** When deploying PIM **bidir**, at any given time it is only possible to have a single active RP (Rendezvous Point) for a given multicast group range. RP redundancy is hence achieved by leveraging a **Phantom RP** configuration. Because multicast source information is no longer available in Bidir, the Anycast or MSDP mechanism used to provide redundancy in sparse-mode is not an option for **bidir**.

---

The following switch configuration example is for a switch deployed as the MPod IPN. The DHCP relay configuration allows the fabric to be discovered by the APIC cluster. The deployment of a dedicated VRF in the IPN for inter-pod connectivity is optional, but is a best practice recommendation. As an alternative, you can use a global routing domain.

```
Example: OSPF as the underlay protocol
feature dhcp
feature pim
service dhcp
ip dhcp relay

# Create a new VRF.
vrf context overlay-1
  ip pim rp-address 12.1.1.1 group-list 225.0.0.0/15 bidir
  ip pim rp-address 12.1.1.1 group-list 239.255.255.240/28 bidir

interface Ethernet1/54.4    #spine connected interface
  mtu 9150
  encapsulation dot1q 4
  vrf member overlay-1
  ip address 192.168.0.1/30
  ip ospf network point-to-point
  ip router ospf infra area 0.0.0.0
  ip dhcp relay address 172.16.0.2 #infra address of APIC 1
  ip dhcp relay address 172.16.0.3 #infra address of APIC 2
  ip dhcp relay address 172.16.0.4 #infra address of APIC 3
  no shutdown

interface loopback29
  vrf member overlay-1
  ip address 12.1.1.2/30

router ospf infra
  vrf overlay-1
  router-id 29.29.29.29
```

```
Example: BGP as the underlay protocol
router bgp 65010
  vrf IPN
    neighbor 192.168.0.2 remote-as 65001
    address-family ipv4 unicast
      disable-peer-as-check
```

In the BGP configuration, the `disable-peer-as-check` command is needed for multi-pod because each pod uses the same ASN.

## Preparing Connectivity to the Fabric Pod

Before bringing up the fabric pod (Pod 1), you first must pre-configure the layer 3-connected APIC cluster (Pod 0) for connectivity through the IPN to a spine in the fabric pod. This is necessary for automatic fabric discovery.

### Before you begin

- If the layer 3 connected virtual APIC cluster is deployed in a separate security zone from the fabric, configure the firewall to allow any necessary protocols and ports.
- Configure the inter-pod network (IPN) device that is connected to the fabric pod spines.
- Configure a fabric external routing profile.
- Configure an OSPF interface policy if you are using OSPF as the underlay protocol.

### Procedure

---

- Step 1** Log in to one of the APICs in the layer 3 connected cluster.
- Step 2** Choose **Fabric > Inventory > Pod Fabric Setup Policy**.
- Step 3** In the work pane, click the + symbol in the **Pod Fabric Setup Policy** page.  
The **Set Up Pod TEP Pool** dialog box opens.
- Step 4** In the **Set Up Pod TEP Pool** dialog box, complete the following steps:
- a) Using the **Pod ID** selector, choose Pod 1.
  - b) In the **TEP Pool** field, enter the TEP pool of the fabric pod.
  - c) Click **Submit**.
- Step 5** In the navigation pane, expand **Quick Start** and click **Add Pod**.
- Step 6** In the work pane, click **Add Pod**.
- Step 7** In the **Configure Interpod Connectivity STEP 1 > Overview** panel, review the tasks that are required to configure interpod network (IPN) connectivity, and then click **Get Started**.
- Step 8** In the **Configure Interpod Connectivity STEP 2 > IP Connectivity** dialog box, complete the following steps:
- a) If you see a **Name** field in an **L3 Outside Configuration** area, choose an existing fabric external routing profile from the **Name** drop-down list.
  - b) Using the **Spine ID** field, enter the Node ID allocated for the spine switch that will be the initial spine to communicate with APIC 1 in Pod 0.
  - c) In the **Interfaces** area, in the **Interface** field, enter the spine switch interface (slot and port) used to connect to the IPN.  
Click the + (plus) icon to add more interfaces.
  - d) In the **IPv4 Address** field, enter the IPv4 gateway address and network mask for the interface.
  - e) From the **MTU (bytes)** drop-down list, choose a value for the maximum transmit unit of the external network.  
The MTU should be 9150 (the default). This value should also be configured on the IPN interface.
  - f) Click **Next**.
- Step 9** In the **Configure Interpod Connectivity STEP 3 > Routing Protocols** dialog box, in the **OSPF** area, complete the following steps to configure OSPF for the spine to IPN interface:

- a) Leave the **Use Defaults** checked or uncheck it.

When the **Use Defaults** check box is checked, the GUI conceals the optional fields for configuring Open Shortest Path (OSPF). When it is unchecked, it displays all the fields. The check box is checked by default.

- b) In the **Area ID** field, enter the OSPF area ID.
- c) In the **Area Type** area, choose an OSPF area type.

You can choose **NSSA area** or **Regular area** (the default). **Stub area** is not supported.

- d) (Optional) With the **Area Cost** selector, choose an appropriate OSPF area cost value.
- e) From the **Interface Policy** drop-down list, choose or configure an OSPF interface policy.

You can choose an existing policy, or you can create one with the **Create OSPF Interface Policy** dialog box. An example is shown in the following table:

**Table 1: OSPF Interface Policy Example**

Property	Setting
<b>Name</b>	ospflfPol
<b>Network Type</b>	Point-to-point
<b>Priority</b>	1
<b>Cost of Interface</b>	unspecified
<b>Interface Controls</b>	none checked
<b>Hello Interval (sec)</b>	10
<b>Dead Interval (sec)</b>	40
<b>Retransmit Interval (sec)</b>	5

**Step 10**

In the **Configure Interpod Connectivity STEP 3 > Routing Protocols** dialog box, in the **BGP** area, leave the **Use Defaults** checked or uncheck it.

The **Use Defaults** check box is checked by default. When the check box is checked, the GUI conceals the fields for configuring Border Gateway Protocol (BGP). When it is unchecked, it displays all the fields. If you uncheck the box, configure the following steps:

- a) Leave the **Use Defaults** checked or uncheck it.
- b) In the **Community** field, enter the community name.

We recommend that you use the default community name. If you use a different name, follow the same format as the default.

- c) In the **Peering Type** field, choose either **Full Mesh** or **Route Reflector** for the route peering type.

If you choose **Route Reflector** in the **Peering Type** field and you later want to remove the spine switch from the controller, you must first disable **Route Reflector** in the *BGP Route Reflector* page. Not doing so results in an error.

To disable a route reflector, right-click on the appropriate route reflector in the **Route Reflector Nodes** area in the **BGP Route Reflector** page and select **Delete**. See the section "Configuring an MP-BGP Route Reflector Using the GUI" in the chapter "MP-BGP Route Reflectors" in the *Cisco APIC Layer 3 Networking Configuration Guide*.

- d) In the **Peer Password**, field, enter the BGP peer password. In the **Confirm Password** field, reenter the BGP peer password.
- e) In the **Route Reflector Nodes** area, click the + (plus) icon to add nodes.

For redundancy purposes, more than one spine is configured as a route reflector node: one primary reflector and one secondary reflector. It is best practice to deploy at least one external route reflector per pod for redundancy purposes.

The **External Route Reflector Nodes** fields appear only if you chose **Route Reflector** as the peering type.

Also, enter the **Peer Address** and **Remote AS** to use for BGP underlay.

**Step 11** Click **Next**.

**Step 12** In the **Configure Interpod Connectivity STEP 4 > External TEP** dialog box, complete the following steps:

- a) Leave the **Use Defaults** checked or uncheck it.

The **Use Defaults** check box is checked by default. When the check box is checked, the GUI conceals the optional fields for configuring the external TEP pool. When it is unchecked, it displays all the fields.

- b) Note the nonconfigurable values in the **Pod** and **Internal TEP Pool** fields.
- c) In the **External TEP Pool** field, enter the external TEP pool for the physical pod.

The external TEP pool must not overlap the internal TEP pool or external TEP pools belonging to other pods.

- d) In the **Data Plane TEP IP** field, accept the default, which is generated when you configure the **External TEP Pool**; if you enter another address, it must be outside of the external TEP pool.
- e) Enter a **Router ID** for the spine node.

**Step 13** Click **Next**.

The **Summary** panel appears, displaying a list of policies created by this wizard. You can change the names of these policies here.

**Step 14** Click **Finish**.

---

## What to do next

Monitor the discovery and registration of the fabric nodes by APIC, as summarized in the following section.

## Summary of Fabric Discovery and Registration

The following is a summary of the switch registration and discovery process.

- The IPN, acting as a DHCP relay agent, forwards DHCP requests from the spine to the APIC.
- Spine switch is now visible to the APIC and appears under **Nodes Pending Registration** in the **Fabric Membership** screen. Navigate to **Fabric > Inventory > Fabric Membership**.
- Register the spine switch on the APIC .
- The APIC allocates an IP address for the spine interface to the IPN and a TEP IP from the TEP pool configured for the fabric pod containing the spine. At this point, the spine joins the ACI fabric.



- The spine advertises the TEP subnet to the IPN through OSPF or BGP so that the IPN learns this subnet.
- The spine acts as a DHCP relay agent for its connected leaf switches, and forwards the requests to the APIC.
- The leaf switches are then visible to the APIC and appear in the section of **Nodes Pending Registration** in **Fabric > Inventory > Fabric Membership**.
- You must manually register these leaf switches on the APIC.
- After a leaf switch is registered, the APIC forwards the TEP IP address and DHCP configuration information to the leaf, which then joins the ACI fabric. Through this discovery process, the layer 3-connected APIC cluster discovers all switches in the fabric pod.

## Configuring QoS for the Layer 3 Connected APIC Cluster

As all traffic between the APIC cluster and the fabric traverses the IPN, we recommend configuring QoS for this traffic. Specific recommendations are as follows:




---

**Note** The configuration examples in this section are for a Cisco Nexus 9000 series switch which can be used as an IPN device. The configuration may differ when using a different platform.

---

- Enable a DSCP class CoS translation policy.
- Retain end-to-end DSCP values for all IPN traffic. When the DSCP class CoS translation policy is enabled, the spine switches will set the DSCP value in packets sent to the IPN per this policy. Traffic destined to the APICs will use the Policy Plane class. Inter-pod control plane traffic (that is, OSPF and MP-BGP packets) will use the Control Plane class. Ensure that the Policy Plane class is configured for Expedited Forwarding (EF) and the Control Plane class is configured for CS4. The IPN network can be configured to prioritize these two classes to ensure that the policy plane and control plane remain stable during times of congestion.

The following example shows the configuration for setting QoS for policy plane traffic on a Cisco Nexus 9000 series switch used as an IPN device.

```
interface Vlan1500
  no shutdown
  vrf member overlay-1
  ip address 172.16.0.2/24
  ip ospf passive-interface
  ip router ospf 1 area 0.0.0.0
  hsrp version 2
  hsrp 100
  ip 172.16.0.1

ip access-list APIC_Cluster
  10 permit ip 172.16.0.0/24 any

class-map type qos match-all APIC-Class
  match access-group name APIC_Cluster
policy-map type qos APIC_Policy
  class APIC-Class
    set dscp 46

interface Ethernet1/1
  switchport mode trunk
  switchport access vlan 1500
  service-policy type qos input APIC_Policy
```

```

interface Ethernet1/2
  switchport mode trunk
  switchport access vlan 1500
  service-policy type qos input APIC_Policy

interface Ethernet1/3
  switchport mode trunk
  switchport access vlan 1500
  service-policy type qos input APIC_Policy

```

- We also recommend that you rate limit non-policy-plane traffic to the APIC to prevent drops on policy plane traffic. Deploy a policer on the cluster-facing IPN to limit traffic with a DSCP value other than 46 (EF) to 4 Gbps with burst of 60 Mbps. The following example shows a rate limiting configuration on a Cisco Nexus 9000 series switch used as an IPN device.

```

class-map type qos match-all no_rate_limit
  match dscp 46

policy-map type qos APIC_Rate_Limit
  class no_rate_limit
  class class-default
    police cir 4 gbps bc 7500 kbytes conform transmit violate drop

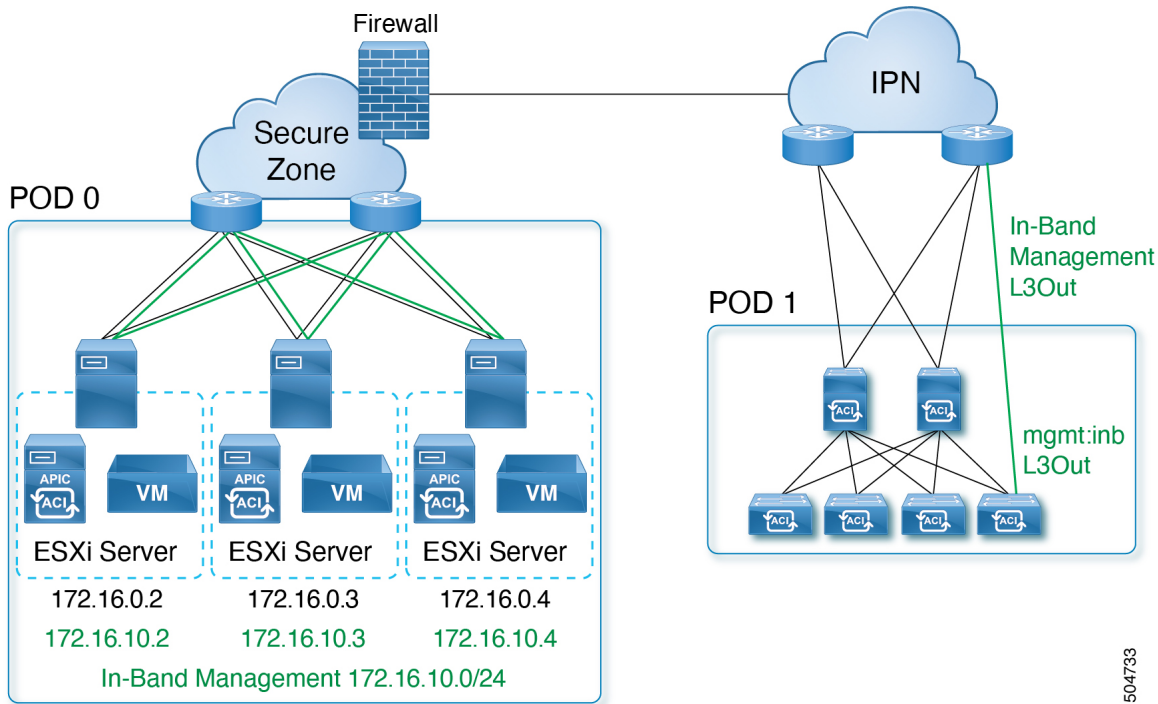
interface Ethernet1/1-3
  switchport mode trunk
  switchport access vlan 1500
  service-policy type qos input APIC_Policy
  service-policy type qos output APIC_Rate_Limit

```

## Configuring In-band Management

When you deploy in-band management with a layer 3 connected APIC cluster, the in-band management VRF (mgmt:inb) is not routed through the spines to the IPN. Connectivity to the in-band management interfaces on the APIC must be routed from an L3Out configured from one of the border leaf switches. This L3Out should be configured in the **mgmt** tenant for the in-band management VRF.

Figure 4: In-band Management



504733

Use this procedure to configure in-band management for the layer 3 connected APIC cluster.

### Before you begin

Configure the in-band VLAN in the APIC-connected port-group in DVS (VMware VCenter).

### Procedure

**Step 1** Configure the APIC upstream switches.

Example:

```
interface Vlan100
  no shutdown
  vrf member IPN
  ip address 172.16.0.252/24
  ip ospf passive-interface
  ip router ospf 1 area 0.0.0.0
  hsrp version 2
  hsrp 100
  ip 172.16.0.1

interface Vlan101
  no shutdown
  vrf member IPN
  ip address 172.16.10.252/24
  ip ospf passive-interface
  ip router ospf 1 area 0.0.0.0
```

```
hsrp version 2
hsrp 101
ip 172.16.10.1
```

You can route the in-band management network using the IPN VRF or you can choose a different VRF.

- Step 2** Configure the APICs for in-band management using the normal in-band management configuration procedure.
- Configure the Static Node Management Addresses for each APIC.
  - Configure the In-Band EPG. For the encapsulation VLAN, you can use any VLAN except for the infrastructure VLAN ID specified during the initial cluster bringup.

For additional information about configuring in-band management, see the "Static In-band Management" chapter of the [Cisco APIC and Static Management Access](#) tech note.

- Step 3** Configure the APIC interfaces on the upstream switches.

The APIC-connected interface must be a trunk interface. If the APICs were initialized with an infra VLAN other than 0, you can configure the interface as in the following example.

```
interface Ethernet1/1
switchport mode trunk
switchport trunk allowed vlan 100-101
```

**Note** You cannot change the APIC infra VLAN ID after the cluster is formed.

---

## Migrating from a Physical APIC to a Virtual APIC



**Attention** Migration from physical to virtual APICs requires erasing the configuration from all fabric nodes and restoring the configuration from backup. All fabric nodes will be reinitialized. This is a disruptive procedure.

It is recommended to contact Cisco TAC for this procedure.

---

Use this procedure to migrate a physical APIC (layer 2, wherein the APIC is directly connected to the ACI fabric) to a virtual APIC (layer 2, ESXi host is directly connected to the ACI fabric). The following procedure also has details for migrating a physical APIC (layer 3, wherein the APIC is remotely attached to the ACI fabric) to a virtual APIC (layer 3, ESXi host is remotely attached to the ACI fabric).

### Before you begin

#### Prerequisites

- Upgrade the physical APIC and all the nodes (spine and leaf switches) to 6.0.2.
- Take a back-up of all the existing configurations.
- Shutdown the physical APIC nodes.
- Clean reload all the fabric nodes. To clean reboot the fabric nodes, log in to each fabric node and execute the following:

```
setup-clean-config.sh
reload
```

## Guidelines and Limitations

- Ensure that the cluster sizes match, that is, migration from a 5-node physical cluster to a 3-node virtual cluster is not supported.
- Ensure both the clusters are healthy and the new cluster can support the required number of APICs.
- The infra VLAN configured in the APICs of the new cluster should be the same as the imported configuration. If there is a mismatch, the infra EPG will not be imported.

## Procedure

---

**Step 1** In the VMware vCenter GUI, create a DVS under the network folder (use DVS Switch version compatible with the ESXi version). Navigate to **Select DataCenter > Actions > New Folder > New Network Folder**.

An existing DVS created for VMM domain integration can be used if the discovery protocol used is CDP (not LLDP).

**Step 2** (only for layer 2) Disable LLDP on the DVS uplinks (VMware vCenter).

**Step 3** Add ESXi hosts. See below for layer 2 and layer 3 details:

- (for layer 2) Add ESXi hosts to the DVS and map the interfaces connected to the leaf nodes (Uplink1, Uplink2 in VMware vCenter).
- (for layer 3) Add ESXi hosts to the DVS and map the interfaces connected to the routed IPN.

**Step 4** Create **InfraPortGroup** on DVS. See below for layer 2 and layer 3 details:

- (for layer 2) Enable trunking with Uplink1 in active and Uplink2 in standby. Enable VLAN 0, inband VLAN, and infra VLAN (Eg: 0, 10, 4093).
- (for layer 3) Enable trunking and set the VLAN range to include inband, interface VLAN used for connecting IPN.

**Step 5** Deploy virtual APIC VMs in the ESXi hosts. Use the same IP addresses as the physical APIC nodes and select the OOB and infra port groups, accordingly.

**Step 6** Power on the virtual APIC VMs. After the VMs are up, proceed to the APIC Cluster Bringup GUI. See the detailed [Bringing up the Cisco APIC Cluster Using the GUI](#) procedure in the *Cisco APIC Getting Started Guide*.

**Step 7** Import the configuration to APIC 1 and verify all fabric nodes are discovered with status *Active*.

Navigate to **Fabric > Inventory > Fabric Membership**. In the **Registered Nodes** tab, check if all the fabric nodes are in *Active* state.

**Step 8** Verify APIC 2 ( to  $N$  , where  $N$  is the cluster size) joins the cluster and the cluster is healthy.

Navigate to **System > Dashboard** . In the **Controller Status** pane (bottom right), the **Health State** column should display *Fully fit* .

---

## What to do next

Traffic verification:

1. Check **Operations Dashboard** on the Cisco APIC GUI, and see the programmed policies.
2. Check for any faults. Navigate to **System > Faults**.







**Americas Headquarters**  
Cisco Systems, Inc.  
San Jose, CA 95134-1706  
USA

**Asia Pacific Headquarters**  
CiscoSystems(USA)Pte.Ltd.  
Singapore

**Europe Headquarters**  
CiscoSystemsInternationalBV  
Amsterdam,TheNetherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at [www.cisco.com/go/offices](http://www.cisco.com/go/offices).