Hardware Expansion and Replacement

• Expanding and Shrinking the Fabric, on page 1
• Hardware Diagnostics and Replacement, on page 6

Expanding and Shrinking the Fabric

ACME may decide to expand their ACI fabric as their data center grows, which means adding new leaf and spine switches, and possibly APICs. Generally, spine switches are added for more throughput, and leaf switches are added for more access ports. APICs are added as the number of policies and endpoints increase. Additionally, some switches or APICs may need to be decommissioned. Also, there may be times when you need to replace failed hardware, which is discussed in the Hardware Replacements chapter.

This section will walk through the operations of adding and removing switches and APICs in your existing ACI fabric. This is done the same way for both spine and leaf switches. Adding APICs will also be covered.

Switches

There are two ways switches can be added to ACME's existing fabric: by discovering the switches automatically in the APIC after they have been cabled to the fabric, or by pre-provisioning the switches by adding their serial numbers and later connecting them physically to the fabric when the switches arrive. Both methods have the same outcome: an expanded fabric in the matter of minutes. This section will also cover decommissioning switches.

Add a Connected Switch

The following procedure adds a switch that has already been attached to the fabric:

Note

When replacing spine switches, you must take into account the BGP route reflector function. You must configure at least two spine switches as BGP route reflectors for a Layer 3 Cisco Application Centric Infrastructure (Cisco ACI) fabric. You can make this configuration change under System > System Settings > BGP Route Reflectors, under Route Reflector Nodes. If you replace or remove the spine switches, make the appropriate configuration changes to keep at least one route reflector active while you replace or remove the spine switches, and have at least two active route reflectors after you complete the changes.

For more information about BGP route reflectors, see the Cisco APIC Layer 3 Networking Configuration Guide.
1. In the case of a leaf switch, cable the switch to all of the spine switches. In the case of a spine switch, cable the switch to all of the leaf switches. Ideally, a best-practice Cisco ACI fabric is connected in a full mesh topology with every leaf switch cabled to every spine switch. All devices should connect to the leaf switches, leaf switches should never connect to other leaf switches, and spine switches should never connect to other spines.

2. In the Cisco Application Policy Infrastructure Controller (Cisco APIC) click on **Fabric** at the top of the screen.

3. Click on **Fabric Membership** in the left navigation pane.

4. When the new switch appears, you'll see a node with a serial number but no Node ID or Node Name configured. Double-click the switch and assign a **Node ID** and a **Node Name**. As a best practice, number leaf nodes starting with 101, and spine nodes with 201. Lower numbers are reserved for the Cisco APICs.

5. Optionally, add a **Rack Name** name. This is commonly used to identify the physical location of the switch in the data center.

6. Click **Submit**.

7. Repeat this process for all of the new switches that are connected to the fabric.

**Pre-Provision Switch Before Connection**

Pre-provisioning a switch is a handy operationally proactive step to get the switch registered before it even arrives to your data center. You will need to know the serial number of the switch you will receive to pre-provision. The following steps walk you through switch pre-provisioning for both leaves and spines:

1. On the menu bar, choose **Fabric**.

2. In the Navigation pane, choose **Fabric Membership**.

3. In the Work pane, choose **Actions > Create Fabric Node Member**.

4. In the **Create Fabric Node Member** dialog box, perform the following actions:
   1. In the pop-up window, enter the **serial number** of the switch that will be arriving.
   2. Assign a **Node ID** and a **Switch Name**. As a best practice, number leaf nodes starting with 101, and spine nodes with 201. Lower numbers are reserved for APICs.

5. Click **Submit**.

   **Note:** Repeat this process for all switches you wish to pre-provision.

The new entry in the Fabric Membership window will show **Unsupported** in the **Role** column until the switch is actually connected to the fabric, but the switch will immediately become a member of the fabric once it arrives and is cabled.

To be proactive, you can also pre-provision fabric policies. Fabric policies are covered in the Fabric Connectivity chapter. For more information on pre-provisioning policies, refer to the following white paper:

Decommission Existing Switch

Decommissioning a switch can either remove it from the fabric entirely, or remove a switch temporarily to perform maintenance. Ensure you do not have any devices connected, as the switch will not forward traffic after decommissioning. There are two types of switch decommissioning: Regular, and Remove from Controller.

Regular decommissioning can be used for maintenance and essentially silences the switch from reporting faults and sending SNMP information as a temporary solution, while keeping the switch's node ID and fabric membership. The switch will show up under the Disabled Interfaces and Decommissioned Switches folder in the left navigation pane.

The Remove from Controller option will completely remove the switch from the ACI fabric and all APICs. The switch will no longer show up in the fabric membership as a registered node and the infrastructure VTEP IP addresses it was assigned will be removed.

Perform the following steps to decommission a switch from the ACI fabric:

1. On the menu bar, choose Fabric.
2. In the Navigation pane, choose Inventory > Pod 1.
3. Click the switch to decommission in the Navigation pane.
   1. Click the General tab.
   2. Chose the Actions > Decommission.
   3. In the pop-up, choose either Regular or Remove from Controller.
4. Click Submit.

Cisco Nexus 9300 Platform Switches to Cisco Nexus 9300-EX Platform Switches Migration

Use this procedure to migrate Cisco Nexus 9300 platform switches in virtual port channel (vPC) to Cisco Nexus 9300-EX platform switches.

**Procedure**

<table>
<thead>
<tr>
<th>Step 1</th>
<th>Remove the cables from Cisco Nexus 9300 platform switch. Power off the switch.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 2</td>
<td>Log in to Cisco APIC.</td>
</tr>
<tr>
<td>Step 3</td>
<td>Choose Fabric &gt; Inventory &gt; Unreachable Nodes.</td>
</tr>
<tr>
<td></td>
<td>Ensure that the node is unreachable. Make a note of the Node Name and Node ID.</td>
</tr>
<tr>
<td>Step 4</td>
<td>Select the node. From the Actions menu, choose Remove From Controller.</td>
</tr>
<tr>
<td></td>
<td>Wait for 5-10 minutes for the node to be removed from the Cisco APIC.</td>
</tr>
<tr>
<td>Step 5</td>
<td>Monitor the traffic on Cisco Nexus 9300 platform switch. All the traffic should be handled by the other Cisco Nexus 9300 platform switch and there should be minimal or no impact to traffic.</td>
</tr>
<tr>
<td>Step 6</td>
<td>Replace Cisco Nexus 9300 platform switch with Cisco Nexus 9300-EX platform switch.</td>
</tr>
<tr>
<td>Step 7</td>
<td>Power on Cisco Nexus 9300-EX platform switch and connect the cables.</td>
</tr>
<tr>
<td>Step 8</td>
<td>Load the Cisco APIC Release 3.0(1) software on Cisco Nexus 9300-EX platform switch. Boot the switch.</td>
</tr>
<tr>
<td>Step 9</td>
<td>Log in to Cisco APIC.</td>
</tr>
</tbody>
</table>
Step 10  Choose Fabric > Inventory > Fabric Membership.
Verify if the switch is displayed.

Step 11  Assign the Node Name and Node ID from step 3 to Cisco Nexus 9300-EX platform switch.

Step 12  Wait for a few minutes for all the relevant policies to be pushed to the Cisco Nexus 9300-EX platform switch and for the end point synchronization to complete. To verify, choose Operations > Capacity Dashboard. Port channel on this switch is not activated.

Step 13  Remove the cables from the other Cisco Nexus 9300 platform switch. Power off the switch.

Step 14  Repeat the steps 1-12 for the other Cisco Nexus 9300 platform switch.

---

Cisco APICs

Add New APIC

Before making any changes to an Application Policy Infrastructure Controller (APIC) cluster, ensure each APIC in the cluster is fully fit and change the cluster size to reflect the new controller you are adding to the cluster. Perform the following steps to verify cluster health:

1. On the menu bar, choose System > Controllers.
2. In the Navigation pane, choose Controllers.
   1. Expand the first APIC in the folder.
   2. Click the Cluster as Seen by Node folder.
   3. Verify every controller shows Fully Fit under the Health State column.

If any of the APICs are not fully fit, refer to the Cisco APIC Troubleshooting Guide.

Perform the following steps to change the APIC cluster size:

1. On the menu bar, choose System > Controllers.
2. In the Navigation pane, choose Controllers > APIC_Name > Cluster as Seen by Node.
3. In the Work pane, choose Actions > Change Cluster Size.
   1. Change the Target Cluster Administrative Size to reflect the new APICs being added.
      
      Note: A cluster size of two is not permitted as that does not allow for quorum amongst APICs.
   4. Click Submit.

Perform the following steps to add a new APIC to the cluster:

1. Install and stage the APIC by connecting it to the fabric by following the hardware installation guide:
2. On the menu bar, choose System > Controllers.
3. In the Navigation pane, choose Controllers > APIC_Name > Cluster.
1. The APIC controllers are added one by one and displayed in the sequential order starting with $N + 1$ and continuing until the target cluster size is achieved.

2. Verify that the APIC controllers are in operational state, and the health state of each controller is **Fully Fit** from the Cluster folder under the new controller.

---

**Note**

It will take several minutes for the APICs to synchronize and join the new APIC to the cluster. Fabric operation will continue normally.

---

**Decommission an Existing Cisco APIC**

When decommissioning Cisco Application Policy Infrastructure Controllers (Cisco APICs), they must be decommissioned sequentially in reverse order. For example, APIC5 must be decommissioned before APIC4. Again, before making any changes to the Cisco APIC cluster, ensure that each Cisco APIC in the cluster is fully fit with the exception of the faulty Cisco APIC being decommissioned. You cannot decommission a powered on fully fit Cisco APIC.

---

**Note**

If you decommission a Cisco APIC with the intent of replacing or re-adding the Cisco APIC after wiping it (given some direction from Cisco TAC), wait at least 10 minutes between running the decommission command and the recommission command. Failure to do so can result in cluster contention and, in the worst case scenario, a disruption in traffic forwarding.

---

Perform the following steps to decommission a Cisco APIC that needs to be removed from the fabric:

1. On the menu bar, choose **System > Controllers**.

2. In the Navigation pane, choose **Controllers > APIC_Name > Cluster**.

---

**Note**

Select a Cisco APIC that is NOT being decommissioned.

---

3. In the Work pane, choose **Actions > Change Cluster Size**.

   1. Change the **Target Cluster Administrative Size** to reflect the new Cisco APICs being added.

---

**Note**

A cluster size of two is not permitted as that does not allow for quorum amongst Cisco APICs.

---

4. Click **Submit**.

5. In the Navigation pane, choose **Controllers > APIC_Name > Cluster**.

---

**Note**

In the main pane, click the Cisco APIC to be decommissioned.
6. In the Work pane, choose **Actions > Actions > Decommission**.

1. Verify that the Cisco APIC no longer appears in the *Cluster* folder under any of the remaining Cisco APICs.

---

**Hardware Diagnostics and Replacement**

The Cisco Application Centric Infrastructure (Cisco ACI) fabric employs a combination of key software and hardware features that are specifically designed to reduce the mean time between failures (MTBF) and the mean time to repair (MTTR). Regarding hardware, there are several hot-swappable components on both the leaf and spine switches in addition to a few components that are fixed on the chassis. If ACME ever experiences some sort of power surge or sees a component of their switches go bad, the hot-swappable components enable them to replace failed hardware quickly and non-disruptively.

---

**Note**

The procedures for replacing hardware typically expect the new hardware to be the same as the hardware that you are replacing.

Examples of hot-swappable components on both the leaf switches and spine switches include:

- Power supplies
- Fan trays

Examples of hot-swappable components on the spines include:

- Power supplies
- Fan trays
- Supervisors
- System Controller cards
- Linecard modules

Despite significant advances in the above components that reduce the MTBF, there is always the possibility of a failure on a leaf switch either in switch hardware or software, or a combination of the two that necessitates a leaf switch replacement. In such an event, the stateless nature of the Cisco ACI fabric provides significant advantages to administrators from an operations standpoint.

---

**Identify Hardware Failure**

When a hardware failure occurs in the fabric, faults are raised in the system dashboard and are presented to the administrator. For cases where there is a component level failure with redundant components present in the system, syslog messages and SNMP traps are generated.

Examples of hardware events that generate syslog messages and SNMP traps include:

- Linecard failure on a spine switch
- Supervisor failure on a spine switch
- System controller failure on a spine switch
- Power supply or fan failures on a leaf or a spine switch

While Cisco Application Policy Infrastructure Controller (APIC) is a central point of management for the entire fabric, operations teams can leverage their existing NMS tools. Logging messages can be sent to syslog servers, such as Splunk, or SNMP messages can be sent to NMS systems, such as ZenOSS, to provide alerting. The leaf and spine switches in the ACI fabric also support traditional methods of detecting failures, such as SNMP polling at a set interval. If responses are not received from the switch in a certain timeframe, there is a possibility that the hardware has failed.

However, while the leaf and spine switches report SNMP and Syslog messages for component level failures, the APICs themselves do not have the ability to generate alerts using SNMP or syslog. For example a power supply failure on the APIC will not generate an SNMP or syslog message and must be monitored and remediated using the APIC dashboard.

### Resolve Leaf Hardware Failure

As an example of a leaf failure, a Nexus 9396 leaf switch that is a part of the fabric is unreachable, perhaps due to a hardware failure on the uplink modules. You can use the GUI to determine the node health to confirm that the leaf has failed.

To view the node health score:

1. On the menu bar, choose **Fabric > Inventory**.
2. In the Navigation pane, choose **Pod 1**.

   **Note:** The pod health displays in the Work pane and is zero.

After confirming that the leaf node has failed, you want to remove the failed switch and provision a new switch as part of the fabric. The first step in replacing the failed switch is to get the failed switch's unique ID (node ID). Each node is assigned an ID in the fabric, which is the reference object that allows a replacement switch with a new serial number to inherit the same stateless configuration that was assigned to the old node.

To view the fabric node IDs using the GUI:

1. On the menu bar, choose **Fabric > Inventory**.
2. In the Navigation pane, choose **Fabric Membership**.

You can also use a single REST API call to periodically poll for a full list of nodes that are at or below a certain health level, as shown in the following example:

{{protocol}}://{{apic}}/api/class/topSystem.xml?rsp-subtree-include=health&rspsubtree-filter-le(healthInst.cur,"0")

In the case of a traditional operations model where each switch was managed as an independent entity, the following high-level procedure replaces the switch:

1. Stand up the replacement switch.
2. Load the correct version of code.
3. Attempt to obtain the latest version of configurations from a configuration repository server.
4. Stage the device with the right configuration file and eliminate any errors. For example, update the AAA, NTP, and syslog servers and the ACLs that are associated with each of them.
5. Copy the old configuration over to the switch.
6. Bring up links one by one and verify if data traffic is flowing correctly.
In an ACI fabric, you can take advantage of the stateless nature of the hardware to instantiate the logical configuration profiles. Replacing the node is as simple as decommissioning the switch and recommissioning it.

To decommission and recommission a switch:

1. On the menu bar, choose Fabric > Inventory.
2. In the Navigation pane, expand Pod 1.
3. Right click the failed node and choose Decommission.
4. Replace the failed leaf switch with the new leaf switch.
5. On the menu bar, choose Fabric > Inventory.
7. The new leaf appears with a node ID of 0 and an IP address of 0.0.0.0.
8. In the Work pane, click on the new leaf.
10. When prompted for the node ID, enter the old node’s ID. In most cases, you can also reuse the same leaf name.
11. Click Update.

If the new switch is not operational, the new switch’s name and node ID are different from the name and ID that you entered. You can get the name and ID by viewing the unreachable nodes.

To view the unreachable nodes:

1. On the menu bar, choose Fabric > Inventory.
2. In the Navigation pane, choose Unreachable Nodes.
3. Find the new switch and record its name and node ID.
4. Repeat the "To decommission and recommission a switch" procedure, starting with step 5. When prompted for the name and node ID, enter the information that you recorded in this procedure.

When the new leaf switch is commissioned successfully, the APIC automatically loads the correct version of the firmware into the leaf.

To view which version of the firmware that the APIC will load:

1. On the menu bar, choose Admin > Firmware.
2. In the Navigation pane, choose Fabric Node Firmware > Firmware Groups > All.
   
   **Note:** In the Work pane, you can see the target firmware version, which is automatically set to the latest firmware version.

In addition, by leveraging the stateless object modeling that replaces the traditional running configuration on a device, APIC automatically loads the correct running configuration onto the device, such as AAA, syslog, SNMP, NTP, ACLs, bridge domains, and EPGs.

In the event that the replacement switch runs standalone NX-OS software instead of ACI switch software, you might need to copy the ACI switch software image to the switch in question.

To copy the ACI switch software image to the switch:

1. Connect to the switch console.
2. Set the IP address on the mgmt0 interface to allow connectivity between the switch and the APIC.
3. Enable SCP services:
   
   ```
   # feature scp-server
   ```
4. Copy the firmware image from APIC to the switch:

   # scp -r /firmware/fwrepos/fwrepo/switch_image_name
   admin@switch_ip_address :switch_image_name

5. For dual supervisor systems, ensure that images are copied to the standby supervisor in case of a full chassis replacement by using the command:

   # copy bootflash:aci_image bootflash://sup-standby/

6. Configure the switch not to boot from Cisco NX-OS.

   switch(config)# no boot nxos

7. Save the configuration.

   switch(config)# copy running-config startup-config

8. Boot the active and standby supervisor modules with the ACI image.

   switch(config)# boot aci bootflash:aci-image-name

9. Verify the integrity of the file by displaying the MD5 checksum.

   switch(config)# show file bootflash:aci-image-name md5sum

10. Reload the switch.

    switch(config)# reload

11. Log in to the switch as an administrator.

    Login: admin

12. Verify whether you must install certificates for your device.

    admin@apic1:aci> openssl asn1parse /securedata/ssl/server.crt

13. Look for PRINTABLESTRING in the command output. If "Cisco Manufacturing CA" is listed, the correct certificates are installed. If something else is listed, contact TAC to generate and install the correct certificates for your device.

Once you have confirmed the that certificate is installed and the switch is in ACI mode, the switch should appear as an unmanaged fabric node when connected to the fabric.

## Resolve APIC Hardware Failure

In this example, you must identify and remediate a hardware failure on one of the APICs in your APIC cluster.

From the GUI of an operational APIC:

1. On the menu bar choose System > Controllers.

2. In the Navigation pane, choose Controllers > apic_name > Cluster.

   **Note:** In the Work pane, you should see the failed APIC in the "Unavailable" operational state.
3. Record the fabric name, target size, node ID of the failed APIC, and the TEP address space. This information is also available through the `acidia avread` command on APIC's CLI.

4. In the Work pane, click the failed APIC to select it.

5. Choose **Actions > Decommission**. The APIC changes to an "Out of Service" admin state.

6. Remove the failed APIC from your rack and install the replacement. The new APIC should boot to the initial setup script.

7. Proceed through the setup script and enter the values of the failed APIC that you recorded in step 3. Failure to configure the APIC with the same settings could result in the fabric entering a partially diverged state.

8. Once the new APIC finishes booting, in the Navigation pane, choose **Controllers > apic_name > Cluster**. You can choose any APIC.

9. In the Work pane, click the new APIC to select it.

10. Choose **Actions > Commission**.

11. The new APIC will receive an IP address, which will be reflected in the APIC GUI. It might take 5 to 10 minutes for this to occur. The new APIC might also cycle between the Available and Unavailable operational states before becoming Fully Fit.

12. On the command line of the new APIC, you can verify that it has joined the fabric by logging in using the credentials that are configured for the rest of the fabric.

---

**Diagnose Equipment Failures**

The ACI fabric provides bootup, runtime and on-demand diagnostics to help assess the hardware health of several sub-systems on each leaf and spine switch.

1. **Boot-up** tests run when switch, card boots up. These are typically ONLY disruptive tests. Comes with default set of tests that can be modified. Deployed via selectors.

2. **Health** (aka On-going) tests run periodically. Can only run non-disruptive tests. Comes with default set of tests that can be modified and are deployed via selectors.

3. **On-Demand** Tests are to be run on specific ports or cards for troubleshooting, there are no defaults, and they can be disruptive.

By default, tests are logically grouped into collections.

To look at the default diagnostic policies, click **Fabric > Fabric Policies > Monitoring Policies > default > diagnostics policy**

In the work pane select the fabric element for which you would like to view the diagnostic monitoring policy.

Test results are viewable by clicking on:

Fabric > Inventory > Pod-1 > Leaf-xx or Spine-xx > Chassis > Supervisor modules > Slot-1

AND

Fabric > Inventory > Pod-1 > Leaf-xx or Spine-xx > Chassis > Line modules > Slot-1

Once there, in the work pane select the **Troubleshooting** tab to view GOLD diagnostic results for the supervisor.

In a modular chassis-based system such as the Cisco Nexus 9500 series switch, diagnostic results are available for all the supervisor, modules, system controller and the fabric modules in the system.
### Guidelines When Replacing a Leaf Switch That is Part of a vPC Pair

When you replace a leaf switch that is part of a vPC pair, the following guidelines apply:

- **Generation 1 switches** are compatible only with other generation 1 switches. You can identify these switch models by the lack of "EX" or "FX" at the end of the switch name. For example, N9K-9312TX.

- **Generation 2 and later switches** can be mixed together in a vPC domain. These switch models can be identified with "EX," "FX," or "FX2" at the end of the switch name. For example N9K-93108TC-EX or N9K-9348GC-FXP.

- When you replace the switch pair, the vPC peer leaf switches must match, and you will have an outage.

The following list provides examples of compatible vPC switch pairs:

- N9K-9312TX and N9K-9312TX
- N9K-93108TC-EX and N9K-9348GC-FXP
- N9K-93180TC-FX and N9K-93180YC-FX
- N9K-93180YC-FX and N9K-93180YC-FX

The following list provides examples of incompatible vPC switch pairs:

- N9K-9312TX and N9K-93108TC-EX
- N9K-9312TX and N9K-93180YC-FX

Alternately, you can change the configuration to break the vPC or introduce the switches as a new vPC pair and migrate the configuration between the nodes.
Guidelines When Replacing a Leaf Switch That is Part of a vPC Pair