Hardware Choices

About Hardware Choices

Cisco Application Centric Infrastructure (ACI) offers a variety of hardware platforms. Choose a platform based on the type of physical layer connectivity you need, the amount of ternary content-addressable memory (TCAM) space and buffer space you need, and whether you want to use IP-based classification of workloads into endpoint groups (EPGs).

The following table provides a summary of the hardware options that were available for the Application Policy Infrastructure Controller (APIC) 1.3(2f) release. You should refer to the Cisco product page for the most up-to-date information.

Table 1: ACI Fabric Hardware Options

<table>
<thead>
<tr>
<th></th>
<th>Port Count</th>
<th>Host Ports Type</th>
<th>Use (leaf/spine)</th>
<th>Policy TCAM</th>
<th>IP-based EPGs</th>
</tr>
</thead>
<tbody>
<tr>
<td>9396PX</td>
<td>48 x 1/10-Gigabit ports and 12 x 40-Gigabit ports</td>
<td>10-Gigabit SFP+</td>
<td>Leaf</td>
<td>Regular TCAM with M12PQ</td>
<td>Yes with M6PQ-E</td>
</tr>
<tr>
<td>9396TX</td>
<td>48 x 1/10-Gigabit ports and 12 x 40-Gigabit ports</td>
<td>10GBASE-T</td>
<td>Leaf</td>
<td>Regular TCAM with M12PQ</td>
<td>Yes with M6PQ-E</td>
</tr>
<tr>
<td>Model</td>
<td>Port Count</td>
<td>Host Ports Type</td>
<td>Use (leaf/spine)</td>
<td>Policy TCAM</td>
<td>IP-based EPGs</td>
</tr>
<tr>
<td>----------</td>
<td>------------</td>
<td>-----------------</td>
<td>------------------</td>
<td>-------------</td>
<td>---------------</td>
</tr>
<tr>
<td>93128TX</td>
<td>96 x 1/10-Gigabit ports and 8 x 40-Gigabit ports</td>
<td>10GBASE-T</td>
<td>Leaf</td>
<td>Regular TCAM with M12PQ</td>
<td>Yes with M6PQ-E</td>
</tr>
<tr>
<td>9372PX</td>
<td>48 x 1/10-Gigabit ports and 6 x 40-Gigabit ports</td>
<td>10-Gigabit SFP+</td>
<td>Leaf</td>
<td>Bigger TCAM</td>
<td>No</td>
</tr>
<tr>
<td>9372TX</td>
<td>48 x 1/10-Gigabit ports and 6 x 40-Gigabit ports</td>
<td>10GBASE-T</td>
<td>Leaf</td>
<td>Bigger TCAM</td>
<td>No</td>
</tr>
<tr>
<td>93108TC-EX</td>
<td>96 x 1/10-Gigabit ports and 6 x 100-Gigabit ports</td>
<td>100-Gigabit QSFP28</td>
<td>Leaf</td>
<td>Bigger TCAM</td>
<td>No</td>
</tr>
<tr>
<td>93120TX</td>
<td>96 x 1/10-Gigabit ports and 6 x 40-Gigabit ports</td>
<td>10GBASE-T</td>
<td>Leaf</td>
<td>Bigger TCAM</td>
<td>No</td>
</tr>
<tr>
<td>93180YC-EX</td>
<td>48 x 10/25-Gigabit ports and 6 x 40/100-Gigabit ports</td>
<td>40-Gigabit QSFP28</td>
<td>Leaf</td>
<td>Bigger TCAM</td>
<td>No</td>
</tr>
<tr>
<td>9332PQ</td>
<td>32 x 40-Gigabit ports</td>
<td>40-Gigabit QSFP+</td>
<td>Leaf</td>
<td>Bigger TCAM</td>
<td>No</td>
</tr>
<tr>
<td>9372PX-E</td>
<td>48 x 1/10-Gigabit ports and 6 x 40-Gigabit ports</td>
<td>10-Gigabit SFP+</td>
<td>Leaf</td>
<td>Bigger TCAM</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Port Count</td>
<td>Host Ports Type</td>
<td>Use (leaf/spine)</td>
<td>Policy TCAM</td>
<td>IP-based EPGs</td>
</tr>
<tr>
<td>------------------</td>
<td>------------</td>
<td>----------------</td>
<td>------------------</td>
<td>-------------</td>
<td>---------------</td>
</tr>
<tr>
<td>9372TX-E</td>
<td>48 x 1/10-Gigabit ports and 6 x 40-Gigabit ports</td>
<td>10GBASE-T</td>
<td>Leaf</td>
<td>Bigger TCAM</td>
<td>Yes</td>
</tr>
<tr>
<td>9336PQ</td>
<td>36 x 40-Gigabit ports</td>
<td>40-Gigabit QSFP+</td>
<td>Spine</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>9504</td>
<td>With 9736PQ: 36 x 40-Gigabit ports per linecard</td>
<td>40-Gigabit QSFP+</td>
<td>Spine</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>9508</td>
<td>With 9736PQ: 36 x 40-Gigabit ports per linecard</td>
<td>40-Gigabit QSFP+</td>
<td>Spine</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>9516</td>
<td>With 9736PQ: 36 x 40-Gigabit ports per linecard</td>
<td>40-Gigabit QSFP+</td>
<td>Spine</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

### Expansion Modules

You can choose among three expansion modules according to the switches you are using and your needs:

- **Cisco M12PQ**—Twelve 40-Gbps ports with an additional 40 MB of buffer space and a smaller TCAM compared to the other models. It can be used with the Cisco Nexus 9396PX, 9396TX, and 93128TX switches.

- **Cisco M6PQ**—Six 40-Gbps ports with additional policy TCAM space. It can be used with the Cisco Nexus 9396PX, 9396TX, and 93128TX switches.

- **Cisco M6PQ-E**—Six 40-Gbps ports with additional policy TCAM space. It can be used with the Cisco Nexus 9396PX, 9396TX, and 93128TX switches and allows you to classify workloads into EPGs based on the IP address of the originating workload.

### Leaf Switches

In the ACI, all workloads connect to leaf switches. The leaf switches used in an ACI fabric are ToR switches. They are divided into four main types based on their hardware:
• Border Leaf—The border leaf switches are ACI leaf switches that provide Layer 2 or Layer 3 external connectivity to outside networks. The border leaf supports routing protocols to exchange routes with external routers, and it also applies and enforces policies for traffic between internal and external endpoints.

• Service Leaf—The service leaf switches are ACI leaf switches that connect to Layer 4-7 service appliances, such as firewall, load balancer, and such. The connectivity between the service leaf and the service appliance can be Layer 2 or Layer 3 depending on design scenarios.

• Compute Leaf—The compute leaf switches are ACI leaf switches that connect to compute systems. The compute leaf supports individual port, port channel, and virtual port channel (vPC) interfaces, based on the nature and requirements of the application or the system. It also applies and enforces policies for traffic to and from local endpoints.

• IP Storage Leaf—The storage leaf switches are ACI leaf switches that connect to IP storage systems. It supports individual port, port channel, and virtual port channel (vPC) interfaces based on the nature and requirements of the application and the system. It also applies and enforces policies for traffic to and from local endpoints.

While it is not a requirement to have dedicated switches to serve certain functions, it is preferred especially for a large data center.

It is easier to standardize configuration templates and enables applications to flexibly tap into any available resources.

For example, a large data center that supports high volume of traffic between the ACI fabric and the core network, might choose to designate two border leaf switches for high availability and scalability considerations.

Spine Switches

The Cisco ACI fabric forwards traffic primarily based on host lookups. A mapping database stores the information about the ToR switch on which each IP address resides. This information is stored in the fabric cards of the spine switches.

The spine switches have several form factors. The models also differ in the number of endpoints that they can hold in the mapping database, which depends on the number of fabric modules installed. Modular switches equipped with six fabric modules can hold the following numbers of endpoints:

• Fixed form-factor Cisco Nexus 9336PQ—Up to 200,000 endpoints

• Modular 4-slot switch—Up to 300,000 endpoints

• Modular 8-slot switch—Up to 600,000 endpoints

• Modular 16-slot switch—Up to 1.2 million endpoints

You can mix spine switches of different types, but the total number of endpoints that the fabric supports is the minimum common denominator. You should stay within the maximum tested limits for the software, which are shown in the Capacity Dashboard in the APIC GUI. At the time of this writing, the maximum number of endpoints that can be used in the fabric is 180,000.

Also keep in mind when choosing the platform:

• Allow for future growth.
• Verify that the features that you want to deploy are supported on the selected platform. For example, the IP-based EPG feature requires the -E, -EX, or later versions of leaf switches.

• Make sure the leaf switch TCAM size is large enough to support the contracts or application rules that will be deployed within the fabric.

• When using two leaf switches for a vPC pair, make sure to use the same switch model to avoid any corner issues.

• Use two or more spine switches for higher bandwidth and for redundant connections to external networks.

Additional References for Hardware Choices

For more information about hardware choices, see:

Leaf Node Categorization

About Leaf Node Categorization

When deploying a Cisco Application Centric Infrastructure (ACI) fabric, you usually delegate specific devices and services to specific leaf nodes. This enables you to understand quickly where issues might be located, given the state of the leaf nodes. This also enables fast diagnosis for node troubleshooting. Typically, the special-purpose leaf categories are as follows:

• Border leaf
• Compute Leaf
• Services Leaf
• Storage Leaf

Prerequisites for Leaf Node Categorization

The following are the prerequisites for leaf node categorization:

• Understand the appliances and devices to be added to the fabric.

• Understand the design to be implemented in the fabric.

Guidelines and Limitations for Leaf Node Categorization

Leaf node categorization enables a network operator to easily distinguish the purposes of leaf nodes when they have issues, whether in troubleshooting or further implementation and growth. There is no strict definition of the categories to be used, nor is there a configuration on the Cisco Application Centric Infrastructure (ACI) fabric to enforce these categories. The categories are only a set of labels that are typically used in the ACI fabric.
• Border Leaf—This leaf node is typically connected to L3Outs. L3Outs can serve as a path into the WAN, or into the core of a legacy network.

• Compute Leaf—This leaf node is typically connected to compute resources, whether the resources are physical or virtualized servers.

• Services Leaf—Services within ACI are typically those given by Layer 4 to Layer 7 services. Services include firewalls, load balancers, and intrusion prevention systems. Services do not need to be integrated into ACI through a service graph template to be considered a service; that is a definition from the applications point of view.

• Storage Leaf—This leaf node is typically connected to storage devices for compute resources. This can include iSCSI, NFS, or other Ethernet medium storage devices.

Leaf nodes do not need to be delegated to only one category. Depending on the design, the categories can overlap. For example, a leaf node serving as a border leaf node can also provide compute resources.

**Additional References for Leaf Node Categorization**

For additional information on border leaf switches:


**Fabric Provisioning**

**About Fabric Provisioning**

**Fabric Infrastructure IP Range Recommendations**

When provisioning an Application Policy Infrastructure Controller (APIC), one of the required data points during the setup stage is an IP address range for infrastructure addressing inside of the fabric. This is primarily for the purposes of allocating tunnel endpoint (TEP) addresses. The default value for this range is 10.0.0.0/16. Although technically you can select a range that overlaps with other subnets in the network, you should choose a unique range for this infrastructure range.

Frequently, the infrastructure IP address range must be extended beyond the Cisco Application Centric Infrastructure (ACI) fabric. For example, when the Application Virtual Switch (AVS) is used, a VMK interface is automatically created that uses an address from the infrastructure range as shown in the following figure:
If the infrastructure range overlaps with other subnets elsewhere in the network, routing problems might occur.

The minimum supported subnet size in the recommended three APIC scenario is /22. The number of addresses required depends on a variety of factors, including the number of APICs in your fabric, the number of leaf and spine nodes, the number of AVS instances, and the number of virtual port channels required. To avoid issues with address exhaustion, you should consider allocating a /16 or /17 range if possible.

When considering the preceding requirements, remember that changing either the infrastructure IP address range or the VLAN after initial provisioning is not possible without rebuilding the fabric.

**Fabric Infrastructure VLAN Recommendations**

During fabric provisioning, the system requires a VLAN number to be used as the infrastructure VLAN. This VLAN is used for control communication as a reserved overlay VLAN between the fabric nodes (leaf, spine, and APIC controllers) to bring up the fabric. This VLAN is hard coded on the fabric nodes.

If possible, this VLAN number should be unique within the network. In a scenario where the infrastructure VLAN is extended outside of the ACI fabric (for example, if using Cisco AVS or OpenStack integration with OpFlex), this VLAN might need to traverse other (non-ACI) devices. In that case, be sure that the infrastructure VLAN does not fall within a range that is prohibited on the non-ACI device. The following figure shows an example of the reserved VLAN range within a Cisco Nexus 7000:
In many cases, VLAN 3967 is a good choice for the ACI infrastructure VLAN to avoid the issue outlined in the preceding section.

For more information about fabric infrastructure VLAN recommendations, see the Cisco APIC Getting Started Guide at the following URL:


**Fabric Node ID Recommendation**

The fabric node ID is used to form the fabric membership during fabric initialization. It is also used to configure underlay physical policies, such as access policies and fabric polices, within the fabric. Having a good node ID structure is important to ease the management and operation for the ACI fabric.

Below are general guidelines for configuring fabric node IDs:

- Plan the node ID wisely to allow for further growth and expansion.
- Use different node ID ranges for spine switches and leaf switches. For example, the 100 range for spine switches and the 200 range for leaf switches.
- Using different node ID ranges for leaf switches depends on the use case. For example, if leaf switches are categorized in a different functionality, consider using a different range based on the use. For example, the 200 range for border leaf switches and service leaf switches, the 300 range for compute leaf switches, and the 400 range for storage leaf switches.
Node IDs 1 through 29 are reserved for APICs, which cannot be changed. When APIC redundancy is configured, you should use IDs 1 to 19 for active APICs and IDs 20 to 29 for standby APICs. This allows for expansion of the fabric.

- When a pair of switches is used for the server uplink connectivity using either vPC or active/standby, consider using sequential numbers for the leaf node ID for those switch pairs. For example, node ID 201 for the vPC side A connectivity and node ID 202 for side B. That way, it is easier to configure and easier to manage an upgrade when using maintenance groups.

- If only one ToR switch is deployed, reserve the even leaf ID for future use.

Once the fabric node ID is assigned, the ID is difficult to change unless the fabric nodes (spine and leaf) are decommissioned from the fabric and cleanly rebooted.