

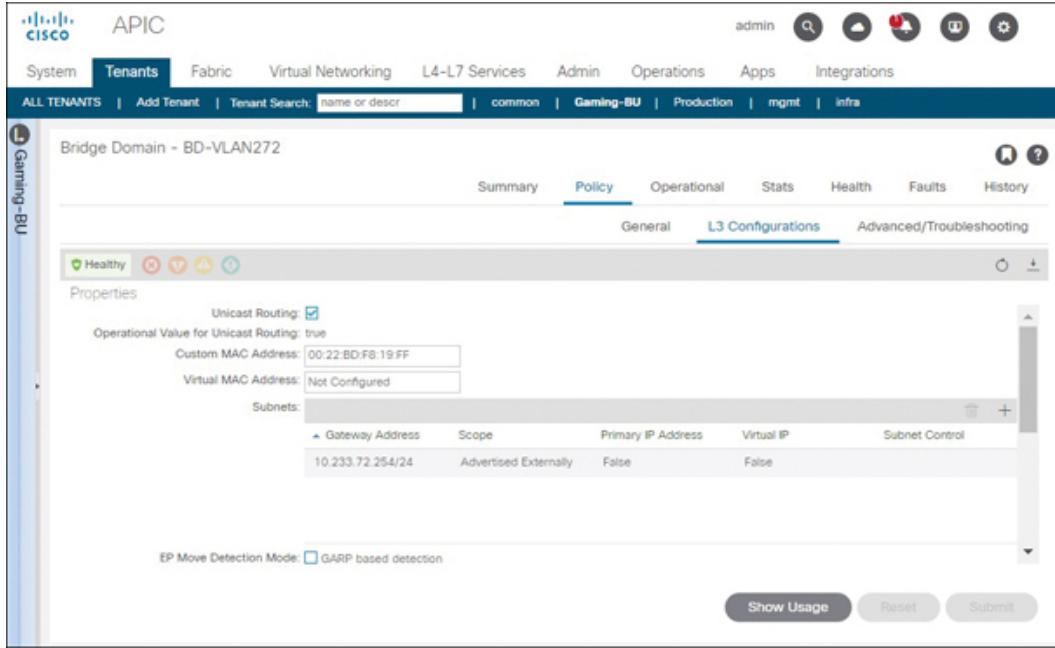
**Figure 10-16** Extending an EPG to Non-ACI Switches

After clicking Submit, navigate to the Operational tab for the EPG. [Figure 10-17](#) shows that ACI has learned five endpoints over the EPG extension. Notice in the figure that MAC addresses have been learned, but no IP addresses have been learned. With the assumption that this is part of a VLAN migration into ACI, this would be normal. The reason ACI has not learned the endpoint IP addresses in this case is that unicast routing has been intentionally disabled on the associated bridge domain to eliminate the possibility of asynchronous routing.

The screenshot shows the Cisco Application Policy Infrastructure Controller (APIC) web interface. The top navigation bar includes tabs for System, Tenants, Fabric, Virtual Networking, L4-L7 Services, Admin, Operations, Apps, and Integrations. The 'Tenants' tab is selected, and the tenant 'Gaming-BU' is chosen. On the left, a sidebar lists tenant components: Quick Start, Application Profiles (including Multiplayer-App1), Application EPGs (listing EPG-Chat-Servers, EPG-Client-VMs, EPG-DB-Servers, EPG-Gaming-Servers, EPG-Login-Servers, EPG-Patch-Servers, EPG-VLAN271, and EPG-VLAN272), uSeg EPGs, Networking (Bridge Domains, VRFs, External Bridged Networks, L3Outs), and L2Outs. The main content area is titled 'EPG - EPG-VLAN271' and shows the 'Operational' tab selected. It displays a table of 'Client End-Points' with columns: End Point, MAC, IP, Learning Source, Interface, and Encap. The table contains five entries, all marked as 'Healthy'. The interface shows the MAC and IP addresses of the endpoints, along with their learned source and interface details. A pagination control at the bottom indicates 'Page 1 Of 1' and 'Objects Per Page: 100', with 'Displaying Objects 1 - 5 Of 5'.

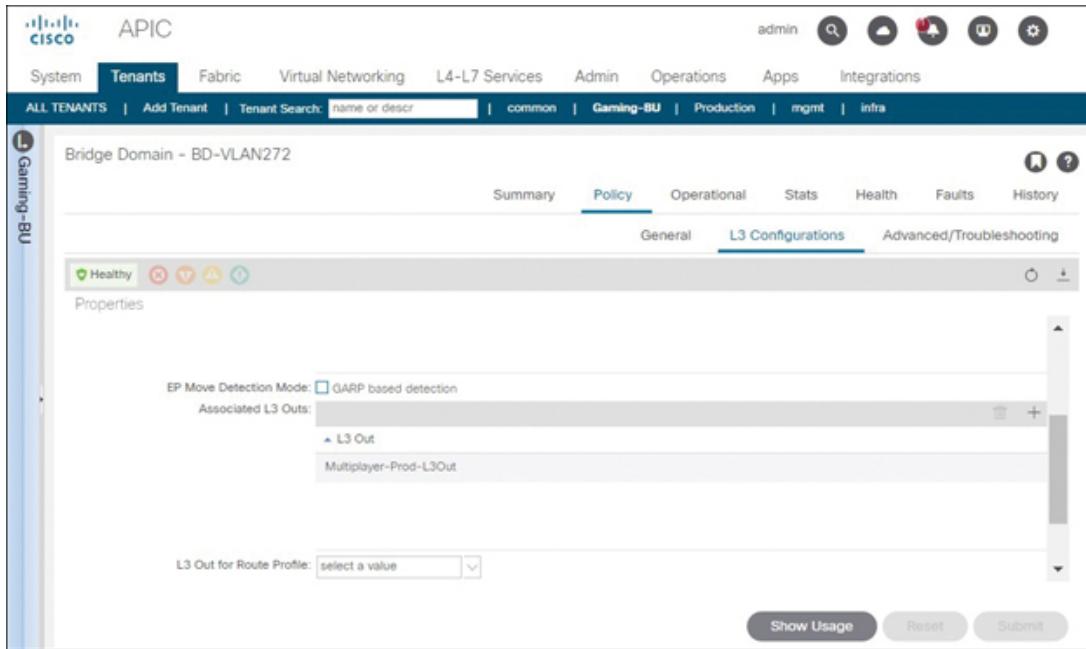
**Figure 10-17** Verifying Endpoint Learning over EPG Extension to Non-ACI Switches

Note that there is no requirement that says that all endpoints in a VLAN need to be moved into ACI before the VLAN default gateway can be migrated into the fabric. [Figure 10-18](#) shows how you can move the default gateway into the fabric by assigning the subnet default gateway to the bridge domain and enabling unicast routing. The default gateway assignment needs to be simultaneously removed from non-ACI devices to reduce the length of the accompanying outage.



**Figure 10-18** Moving the Default Gateway for a VLAN into an ACI Fabric

Once these settings have been applied, ACI learns the endpoint IP addresses and can route traffic to other subnets within the fabric. Notice in [Figure 10-18](#) that the subnet has been defined with the Scope setting Advertise Externally. This implies that the bridge domain is also intended to be advertised out an L3Out. [Figure 10-19](#) shows Multiplayer-Prod-L3Out selected as the L3Out from which the subnet should be advertised to the rest of the enterprise network.



**Figure 10-19** Advertising the Subnet out ACI via an L3Out

### Note

Even though you can move a default gateway into ACI before moving endpoints into the fabric, doing so changes traffic patterns because all cross-subnet traffic between outside endpoints needs to flow over the Layer 2 links between the fabric and non-ACI switches twice. Therefore, you should ensure beforehand that you have sized the Layer 2 connection appropriately.

Although migrating subnets into an ACI fabric is not explicitly within the scope of the DCACI 300-620 exam, the final part of this implementation example illustrates that default gateways can be moved into ACI or out of ACI with relative ease.

[Figure 10-20](#) completes coverage of EPG extensions by showing endpoint information for EPG-VLAN272, which has also been extended over the newly created vPC. This output shows an endpoint off the legacy switch whose IP address has been learned in addition to another endpoint internal to the fabric off node ID 301 port 1/46. Because they are in the same EPG, these endpoints are allowed to communicate with one another without contracts by default. Because the default gateway has been migrated into ACI, however, the

endpoints within the EPG cannot communicate with other EPGs or outside the ACI fabric until proper contracts have been put in place. Also, note in [Figure 10-20](#) that a single MAC address is shown without an IP assignment. In this case, this is the MAC address of a deleted switch virtual interface (SVI) on a switch outside ACI and will eventually time out and be removed from ACI endpoint tables.

End Point	MAC	IP	Learning Source	Interface
EP-00:50:56:B7:75:1D	00:50:56:B7:75:1D	10.233.72.10	learned	Pod-1/Node-301/eth1/46 (learned)
EP-2C:D0:2D:FF:0C:57	2C:D0:2D:FF:0C:57	---	learned	Pod-1/Node-301-302/EPG-Extension-vPC...
EP-AA:AA:AA:AA:AA:AB	AA:AA:AA:AA:AA:AB	10.233.72.1	learned	Pod-1/Node-301-302/EPG-Extension-vPC...

**Figure 10-20** Endpoints Both Within ACI and Outside ACI in a Single EPG

### Note

You may be wondering at this point when it might be an ideal time to change the L2 Unknown Unicast setting from Flood to Hardware Proxy and to disable ARP Flooding for VLANs that have been migrated into ACI. Eliminating flooding, after all, increases stability in the network and reduces the control plane impact of endpoint learning. The ideal time depends! Flooding helps enable non-ACI switches to learn endpoints residing in the fabric because it ensures that non-ACI switches receive all ARP packets. Therefore, this flooding behavior should remain in place until all endpoints within a given VLAN as well as the default gateway have been fully migrated into the ACI fabric.

## Implementing L2Outs

In the previous subsection, you may have noticed that multiple EPG extensions can be implemented by trunking EPGs over a single interface, port channel, or vPC. L2Outs are no different. One configuration difference between EPG extension and bridge domain extension that you do need to remember is that L2Outs do not use physical domains. They require a special type of domain called an *external bridged domain*. Because the implementation of access policies beyond the domain is the same across both Layer 2 extension methods, the implementation of an interface policy group and its assignment to ports is not repeated here. Assume for the following example that an AAEP has been assigned to the interface policy group, and the interface policy group has been assigned to physical ports via an interface profile, but the domain assignment to the AAEP is still outstanding.



To create an L2Out, after configuring access policies, you can navigate to the Tenants view, expose the Networking folder in the tenant, right-click External Bridged Networks, and select Create Bridged Outside. The Create Bridged Outside wizard appears, as shown in [Figure 10-21](#).





**Figure 10-21** The Create Bridged Outside Wizard

In the Create Bridged Outside wizard, enter a descriptive name in the Name field, specify the internal bridge domain to extend in the Bridge Domain field, and enter an encapsulation to use for the Layer 2 EPG in the Encap field. The External Bridge Domain drop-down box allows you to select an existing domain to use for the L2Out or to create a new one. From the options in the drop-down box, select Create External Bridge Domain to open the page shown in [Figure 10-22](#).

**Key Topic**

The screenshot shows the 'Create Layer 2 Domain' configuration window. The 'Name' field is populated with 'L2Out-Domain'. The 'Associated Attachable Entity Profile' dropdown is set to 'L2-to-Legacy-Network-AAEP'. The 'VLAN Pool' dropdown is set to 'BD-Extension-VLANs(static)'. Below these fields is a 'Security Domains' section containing a table with four rows:

Select	Name	Description
<input type="checkbox"/>	Development	
<input type="checkbox"/>	Production	
<input type="checkbox"/>	Sec-Domain	
<input type="checkbox"/>	Infra	

At the bottom right are 'Cancel' and 'Submit' buttons.

**Figure 10-22** Creating an External Bridge Domain

From in [Figure 10-22](#), it is clear that the external bridged domain will be named L2Out-Domain, and it will be bound to the attachable access entity profile (AAEP) created in the “Implementing EPG Extensions” section of this chapter. Note that whatever AAEP you select must be the same AAEP you assign to the interface policy group you use for bridge domain extension. It is best to dedicate a new VLAN pool for bridge domain extension purposes when first implementing this feature. Click Submit to create the external bridged domain.

Back in the Create Bridged Outside wizard, you need to validate that ACI has populated the External Bridged Domain drop-down and that the information in all other required fields is correct. Then you can reference each interface, port channel, and vPC over which the bridge domain should be trunked by selecting the correct path type and path information and clicking Add. [Figure 10-23](#) shows that the full path of the port, port channel, or vPC appears at the bottom of the window. Click Next when you are ready to define the Layer 2 EPG.

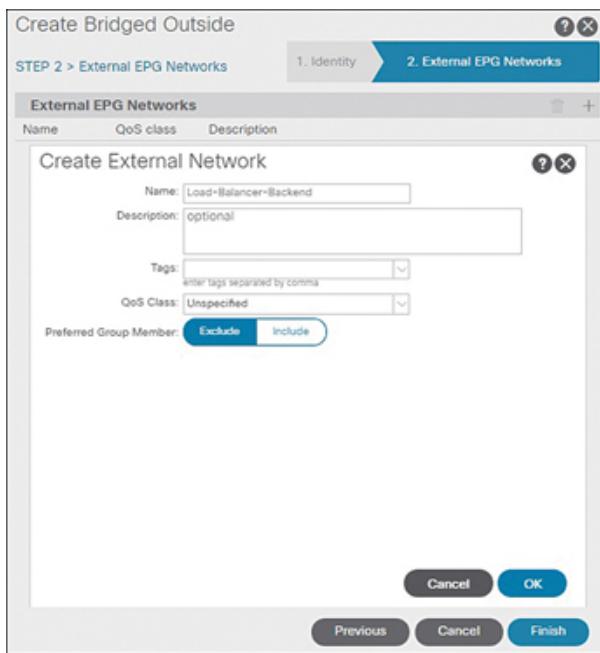




**Figure 10-23** Confirming BD Extension Ports, Port Channels, or vPCs at the Bottom of the Window

On the External EPG Networks page of the Create Bridged Outside wizard, click the + sign to create a Layer 2 EPG. As shown in [Figure 10-24](#), the wizard allows configuration of Preferred Group Member settings at the time of external EPG creation, but it does not allow assignment of contracts. Populate the Name field and click OK.

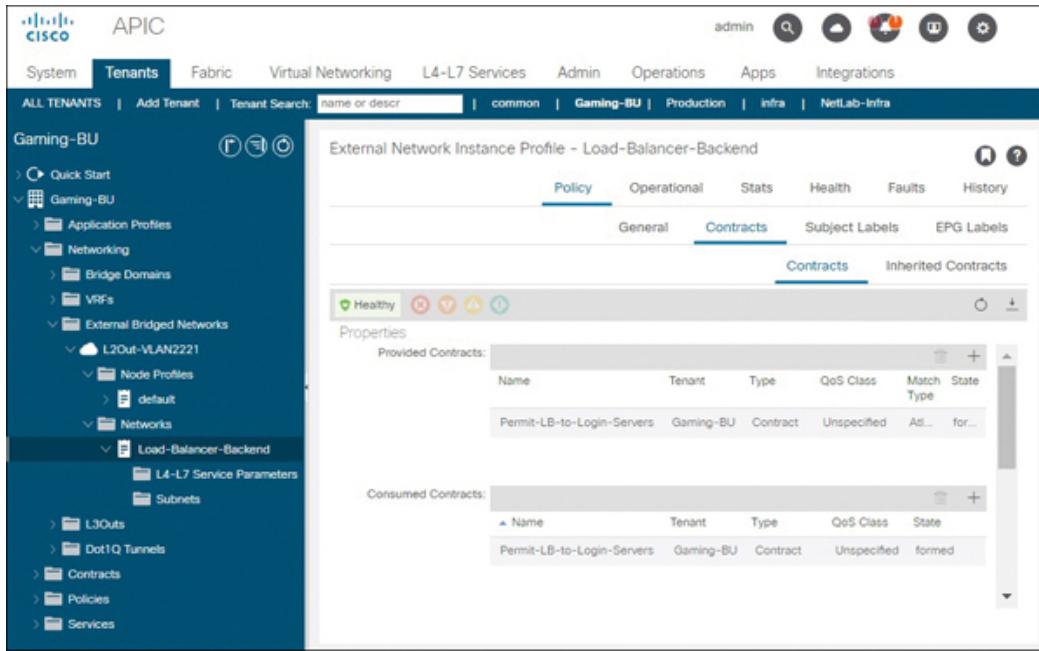
**Key Topic**



**Figure 10-24** Creating a Layer 2 EPG When Implementing Bridge Domain Extension

Once you have configured the Layer 2 EPG (yes, you can only configure one per bridge domain), click Finish to create the bridge extension. [Figure 10-25](#) shows the newly created Layer 2 EPG. In this view, contracts can be assigned in the desired direction(s). Notice that a contract named Permit-LB-to-Login-Servers has been assigned bidirectionally to the Layer 2 EPG. In this case, this contract has a subject that allows any-to-any communication, ensuring that any internal EPGs to which this contract is assigned can communicate with the Layer 2 EPG.

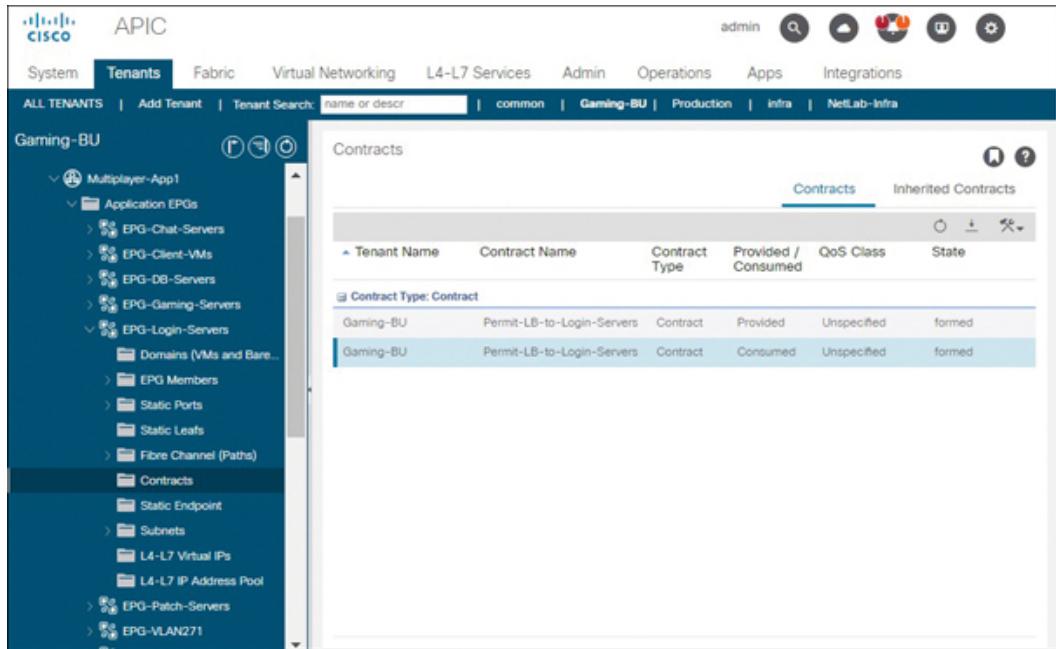




**Figure 10-25 Applying Contracts on a Layer 2 EPG After Implementing BD Extension**

A single internal EPG, EPG-Login-Servers, has been selected for communication with the Layer 2 EPG. [Figure 10-26](#) shows the contract shown earlier applied to this EPG. When the contract is applied as both Provided and Consumed, the communication can be initiated either by endpoints in the Layer 2 EPG or endpoints in the internal EPG.

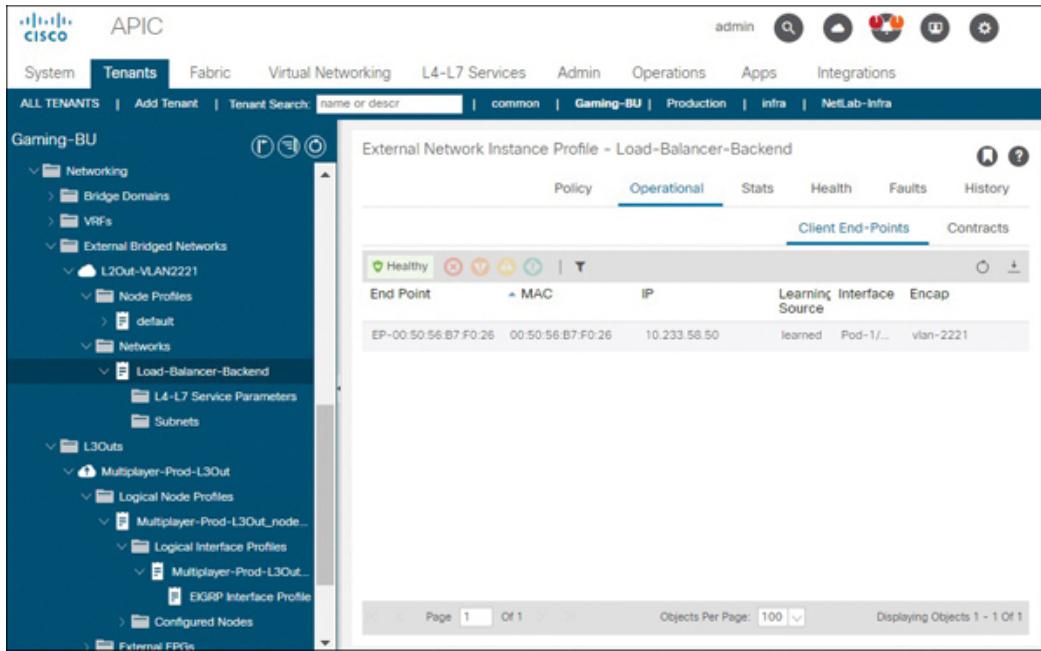
**Key Topic**



**Figure 10-26 Applying Contracts on an Internal EPG to Allow Layer 2 EPG Communication**

As a result of the bridge domain extension, the endpoint depicted in [Figure 10-27](#) and residing behind the L2Out has been learned. Notice that this view does not and should not show endpoints learned in EPG-Login-Servers. With the contract application in place, this endpoint has open communication with any endpoints in EPG-Login-Servers.

**Key Topic**



**Figure 10-27** Verifying Endpoint Learning in a Layer 2 EPG

Be very careful when using any-to-any contracts, especially when applying them in both the consumed and provided directions. In the previous example, if the contract Permit-LB-to-Login-Servers were also applied to a tertiary EPG, it would allow this new EPG to communicate not just with the Layer 2 EPG but also with EPG-Login-Servers. If this is not the intent, and full contract-based communication is expected, either dedicated contracts can be created between each pair of EPGs requiring full communication or more thought needs to be put into the contract design.



### Note

When the Legacy Mode checkbox has been enabled for a bridge domain, the BD can no longer be extended out of the fabric. But why? The reason is that Legacy mode does not allow allocation of an additional encapsulation for any EPGs that are bound to it. This includes the Layer 2 EPG.

# Migrating Overlapping VLANs into ACI

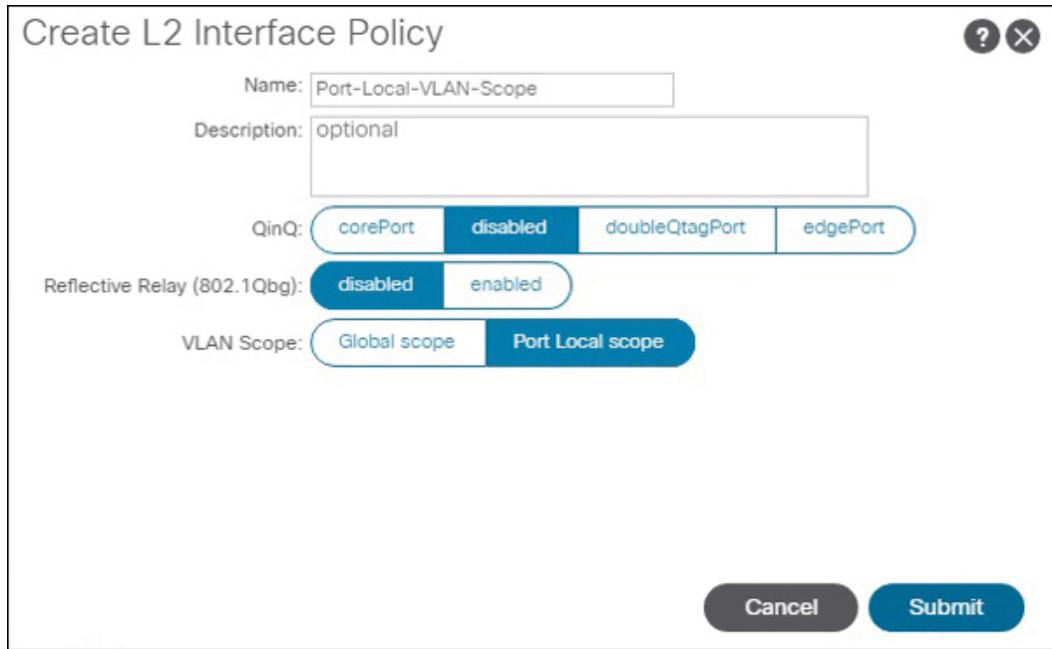


Under ACI default behavior, each VLAN ID used to encapsulate Layer 2 traffic to and from external switches and servers can be mapped to only a single EPG on each leaf switch. If you deploy a second EPG with an encapsulation that has already been used on the leaf, ACI raises a fault and does not encapsulate the traffic for the second EPG. This behavior is governed by the VLAN scope setting, which defaults to Global.

During network migrations into ACI, it sometimes happens that a company has multiple switch blocks that have overlapping VLAN IDs. Suppose that these overlapping VLAN IDs need to be migrated into ACI. Ideally, they would be migrated onto separate Layer 2 extensions terminating on different sets of leaf switches. In such a case, the VLAN overlap would not even matter. But what if both Layer 2 extensions need to terminate on the same leaf switch(es)? In that case, the VLAN Scope setting Port Local Scope can assist.



To implement the Port Local Scope capability, you navigate to **Fabric > Access Policies > Policies > Interface**, right-click L2 Interface, and select Create L2 Interface Policy. Then, as shown in [Figure 10-28](#), you can enter a value in the Name field, set VLAN Scope to Port Local Scope, and click Submit.



**Figure 10-28** Creating an L2 Interface Policy

Once this new L2 interface policy has been assigned to all interface policy groups on which overlapping VLANs will flow, the overlapping VLANs can then be trunked to the intended leaf using EPG extension. Separate VLAN pools and thus separate physical domains are required for EPG extension to successfully deploy this feature.

**Key Topic**

[Example 10-1](#) verifies that EPG-VLAN272 bound to BD-VLAN272 and EPG-Client-VMs bound to BD-Production have both been successfully extended from LEAF102 to outside switches via the encapsulation vlan-272 after deployment of Port Local Scope.

**Example 10-1** Single Encapsulation Across Multiple EPGs with Port Local Scope

[Click here to view code image](#)

```
LEAF102# show vlan extended
```

VLAN Name	Encap	Ports
-----------	-------	-------

```
-----
```

27	Gaming-BU:Multiplayer-App1:EPG-VLAN272	vlan-272	Eth1/6, Po5
28	Gaming-BU:Multiplayer-App1:EPG-Client-VMs	vlan-272	Eth1/33
54	Gaming-BU:BD-VLAN272	vxlan-15040469	Eth1/6, Eth1/45,
61	Gaming-BU:Multiplayer-App1:EPG-VLAN272	vlan-2508	Eth1/46, Po5
65	Gaming-BU:BD-Production	vxlan-15826916	Eth1/33, Eth1/45,
73	Gaming-BU:Multiplayer-App1:EPG-Client-VMs	vlan-3169	Eth1/46

### Note

Use of the Port Local Scope feature does have an impact on leaf scalability. Review the Verified Scalability Guide and Cisco APIC Layer 2 Networking Configuration Guide for your ACI release to understand scalability impact and additional caveats.

## Understanding ACI Interaction with Spanning Tree Protocol



There is not much to know about ACI interaction with Spanning Tree Protocol aside from the fact that ACI does not run Spanning Tree Protocol and does not participate in building the overall Spanning Tree Protocol topology. ACI does, however, flood Spanning Tree Protocol bridge protocol data units (BPDUs) that it receives in an EPG to all other ports with the same VLAN encapsulation within the EPG. From this perspective, ACI acts like a hub for Spanning Tree Protocol.

BPDUs. It provides external switches the data they need to be able to prevent a Layer 2 loop—but that is about all it does.

## Remediating Against Excessive Spanning Tree Protocol TCNs

When a Spanning Tree Protocol topology change occurs, the root bridge sends a special BPDU called a topology change notification (TCN) out its ports. Something as simple as a server-facing port bouncing can trigger a TCN. Because ACI does not process Spanning Tree Protocol packets, the fabric needs to assume that the change may have impacted endpoint locations and must therefore respond by flushing all endpoint table entries in the EPG encapsulation in which the TCN arrived.

If external switches send an excessive number of TCNs into the fabric, this can cause ACI to be constantly flushing endpoint entries. This problem manifests as intermittent packet loss, which can be detrimental to production traffic.

Note that a healthy network should not transmit large numbers of TCNs. This, therefore, is not an ACI problem. Where excessive TCNs occur, it is often the result of a lack of attention to Spanning Tree Protocol optimizations. For instance, a server NIC hardware or driver problem that causes a port to bounce tens of times per second would cause a flurry of TCNs. Preventing servers and other devices that do not bridge traffic from impacting the Spanning Tree Protocol topology would therefore ensure a more stable Spanning Tree Protocol topology with a minimal number of TCNs.

In networks in which excessive TCNs impact production traffic within ACI, the following actions can be taken to remediate the situation:

- Optimize the Spanning Tree Protocol topology by implementing PortFast (**spanning-tree port type edge** or **spanning-tree port type edge trunk** in NX-OS) and BPDU Guard on all switch interfaces that face non-bridging devices such as servers.
- Use a bridge domain extension to non-ACI switches. Because internal EPGs use different encapsulations compared to the Layer 2 EPG configured for bridge domain extension, ACI does not need to flush entries for the internal EPGs if it receives a TCN

on the Layer 2 EPG. This limits the number of entries that need to be relearned.

- Extend EPGs or BDs that need to be extended out of the fabric only for legitimate purposes. When all endpoints and the default gateway for a VLAN have been migrated into ACI, prune the VLAN/EPG/BD off the Layer 2 extension.
- Implement BPDU Filter on non-ACI switch interfaces that face ACI to prevent all BPDUs from entering the fabric in the first place. Note that this should be done only if there is no more than one point of entry (single interface, port channel, or vPC) for external Layer 2 traffic for each VLAN into the ACI fabric. If this is not the case, BPDU Filter could trigger a Layer 2 loop.

## Configuring MST Instance Mappings in ACI

Per-VLAN Spanning Tree (PVST+) and Rapid Per-VLAN Spanning Tree (Rapid PVST+) BPDUs include VLAN tags. ACI can therefore easily identify the encapsulation and EPG associated with the BPDUs and flood the BPDUs out other ports that have the same EPG and encapsulation combination without any further user action.

Multiple Spanning Tree (MST) BPDUs, on the other hand, do not carry VLAN tags and are sent between switches untagged over any interswitch trunk links. Because ACI does not process traffic received on downlink ports unless there is an EPG assigned to it, there is a possibility that MST BPDUs will be dropped on ingress. Furthermore, MST relies on region names, revisions, and instance-to-VLAN mappings to build the Spanning Tree Protocol topology. Because ACI receives MST BPDUs untagged, it has no way to determine which VLAN IDs and therefore EPGs to flush upon TCN receipt.

For the reasons noted here, administrators need to take the following actions to safely connect legacy networks that run MST with ACI:



1. Create a special MST EPG and map it to all ports facing non-ACI switches that run MST. This ensures that ACI does not drop MST BPDUs.

2. Navigate to **Fabric > Access Policies > Policies > Switch > Spanning-Tree > default** and create MST region policies that include the MST region names, MST instance IDs, revision IDs, and relevant VLAN encapsulations. This ensures that ACI knows which EPGs to flush when it receives a TCN and also out of which ports it should forward MST BPDUs.

**Note**

The VLAN Scope setting Port Local Scope is not supported on interfaces configured with MST.

## Understanding Spanning Tree Protocol Link Types

When two Spanning Tree Protocol-speaking switches have a direct connection with one another, the default **spanning-tree link type point-to-point** interface subcommand on IOS and NX-OS switch platforms helps expedite Spanning Tree Protocol convergence. This is because when two switches are the only switches on a link or port aggregation, they are able to use basic proposals and agreements to safely negotiate an immediate transition from the Spanning Tree Protocol blocking state to the forwarding state.

Because ACI can transmit Layer 2 traffic between multiple external switches, an immediate transition to a forwarding state upon receipt of a single agreement can cause Layer 2 loops.



Therefore, when connecting multiple external switches to ACI, you can configure ACI-facing interfaces, port channels, and vPCs with the **spanning-tree link type shared** interface subcommand. This slows down Spanning Tree Protocol convergence if ACI ever becomes a transit point for Spanning Tree Protocol BPDUs between external switches and prevents external switches from negotiating expedited Spanning Tree Protocol state transitions.

## Using MCP to Detect Layer 2 Loops

As discussed in [Chapter 7, “Implementing Access Policies,”](#) MisCabling Protocol (MCP) can detect loops. In response, it can either log an incident or take action by blocking an offending port or link aggregation to break the loop.



While having MCP shut down ports is not always recommended, some customers do implement BPDU Filter on external switch interfaces facing ACI to prevent excessive TCNs from impacting ACI. In such cases, it is important to use MCP to block external Layer 2 connections and break any potential Layer 2 loops if at any point they do occur.

When used, MCP should ideally be enabled on all leaf downlinks, regardless of whether a server or switch connects to the port.

## Exam Preparation Tasks

As mentioned in the section “How to Use This Book” in the Introduction, you have a couple of choices for exam preparation: [Chapter 17, “Final Preparation,”](#) and the exam simulation questions in the Pearson Test Prep Software Online.

## Review All Key Topics

Review the most important topics in this chapter, noted with the Key Topic icon in the outer margin of the page. [Table 10-4](#) lists these key topics and the page number on which each is found.



**Table 10-4** Key Topics for [Chapter 10](#)

Key Topic Element	Description	Page Number
Paragraph	Calls out the most important criteria for network-centric deployments	362
Paragraph	Define vzAny (Any EPG)	364
Table 10-2	Describes the bridge domain settings that should be used when extending Layer 2 out an ACI fabric and some of the logic behind these settings	368
List	Lists the three primary methods for extending Layer 2 outside an ACI fabric	372
Table 10-3	Compares bridge domain extension and EPG extension	375
Paragraph	Describes how to launch the Create Bridged Outside wizard	380
Figure 10-21	Shows the Create Bridged Outside wizard	381

<a href="#">Figure 10-22</a>	Shows how to create an external bridge domain for BD extension	<a href="#">381</a>
<a href="#">Figure 10-23</a>	Emphasizes the need to verify that ports, port channels, or vPCs used to extend a BD have been added as candidates for extension in the wizard	<a href="#">382</a>
<a href="#">Figure 10-24</a>	Shows how to create a Layer 2 EPG when implementing bridge domain extension	<a href="#">383</a>
<a href="#">Figure 10-25</a>	Shows how to apply contracts on a Layer 2 EPG after implementing a BD extension	<a href="#">383</a>
<a href="#">Figure 10-26</a>	Shows how to apply contracts on an internal EPG to allow communication with a Layer 2 EPG	<a href="#">384</a>
<a href="#">Figure 10-27</a>	Shows how to verify endpoint learning in a Layer 2 EPG	<a href="#">384</a>
Paragraph	Describes a common pitfall of any-to-any contract use and how to avoid this pitfall	<a href="#">385</a>
Paragraph	Describes the default VLAN scope setting Global Scope	<a href="#">385</a>
Paragraph	Describes how to implement the Port Local Scope capability	<a href="#">385</a>

Paragraph	Emphasizes the need for separate VLAN pools and physical domains when deploying Port Local Scope	386
Paragraph	Emphasizes the fact that ACI does not run Spanning Tree Protocol or participate in the Spanning Tree Protocol topology and simply transits Spanning Tree Protocol BPDUs	386
List	Calls out two important measures ACI administrators should take if external switches running MST connect to an ACI fabric	388
Paragraph	Describes why it is necessary to configure external switch interfaces that face ACI with the <b>spanning-tree link type shared</b> interface subcommand	388
Paragraph	Describes a use case in which it is sometimes vital to run MCP as an ACI-side loop-detection and mitigation mechanism	388

## Complete Tables and Lists from Memory

Print a copy of [Appendix C, “Memory Tables”](#) (found on the companion website), or at least the section for this chapter, and complete the tables and lists from memory. [Appendix D, “Memory Tables Answer Key”](#) (also on the companion website), includes completed tables and lists you can use to check your work.

## Define Key Terms

Define the following key terms from this chapter and check your answers in the glossary:

Any EPG (vzAny)  
preferred group member  
L2 Unknown Unicast  
L3 Unknown Multicast Flooding  
Multi Destination Flooding  
external bridged network

# **Part IV: Integrations**

# Chapter 11

## Integrating ACI into vSphere Using VDS

This chapter covers the following topics:

**Understanding Networking in VMware vSphere:**

This section provides a primer on networking in ESXi hypervisors.

**Understanding VMM Integration:** This section covers VMM integration concepts and prerequisites for successful integration using a VDS.

**Integrating ACI into vSphere Using VDS:** This section walks through the implementation of the most common and simple form of VMM integration.

This chapter covers the following exam topics:

- 4.1 Implement VMware vCenter DVS integration
- 4.2 Describe resolution immediacy in VMM

APIC controllers can integrate into hypervisor and container environments to extend the benefits of ACI—such as whitelisting and network automation—into virtualized infrastructure.

*vSphere* is an ecosystem of server virtualization products from VMware. ACI has multiple ways to integrate with *vSphere*, the most common of which is by using a *vSphere* distributed switch (VDS).

When APICs integrate with a virtualized environment, they generally do so by integrating with Virtual Machine Manager (VMM). The term VMM can refer to any system or application that manages virtual machines. In the case of *vSphere*, the APICs integrate with the VMM component called vCenter. ACI integrations of this type are called *VMM integrations*.

This chapter addresses the Implementing Cisco Application Centric Infrastructure DCACI 300-620 exam objectives related to VMM integration using the VMware VDS and also builds the context that ACI engineers need to better understand basic networking in *vSphere* environments.

## **“Do I Know This Already?” Quiz**

The “Do I Know This Already?” quiz allows you to assess whether you should read this entire chapter thoroughly or jump to the “Exam Preparation Tasks” section. If you are in doubt about your answers to these questions or your own assessment of your knowledge of the topics, read the entire chapter. [Table 11-1](#) lists the major headings in this chapter and their corresponding “Do I Know This Already?” quiz questions. You can find the answers in [Appendix A, “Answers to the ‘Do I Know This Already?’ Questions.”](#)

**Table 11-1** “Do I Know This Already?” Section-to-Question Mapping

Foundation Topics Section	Questions

Understanding Networking in VMware vSphere	1-3
Understanding VMM Integration	4, 5
Integrating ACI into vSphere Using VDS	6-10

## Caution

The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark that question as wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

- 1.** A vSphere administrator has placed the vCenter instance that manages a VDS on the same VDS. What port binding setting can ensure optimal recoverability of the vSphere environment?
  - a.** Dynamic binding
  - b.** Static binding
  - c.** Default
  - d.** Ephemeral
  
- 2.** Which of the following virtual switch options allows more advanced features but is still owned and managed by VMware?

- a.** vSphere standard switch
  - b.** vSphere distributed switch
  - c.** Nexus 1000v
  - d.** Application Virtual Switch
- 3.** Which of the following vSphere constructs receives and processes system traffic?
- a.** VMkernel adapter
  - b.** Port group
  - c.** Data center
  - d.** vMotion
- 4.** An ACI administrator creates three non-overlapping dynamic VLAN pools and associates a VLAN pool with each of three separate VMM domains. A vSphere administrator notices that port groups across the VMM domains have different VLAN IDs associated with them and asks for feedback on why this is the case. Which responses together are the most accurate responses? (Choose all that apply.)
- a.** The automatic assignment of VLAN IDs increases operational efficiency.
  - b.** Port group VLAN assignments are not important.
  - c.** There is no way to control VLAN assignments when implementing VMM integration.
  - d.** ESXi servers from multiple vCenter instances connect to the same set of leaf switches. Therefore, overlapping VLANs to ensure that VLAN IDs match can cause performance issues.
- 5.** Which objects do vSphere distributed switches tie to in vCenter?

- a.** vCenter
  - b.** Cluster
  - c.** Data center
  - d.** EPG
- 6.** True or false: A vSphere administrator can associate ESXi hosts to an ACI-generated VDS before an ACI administrator configures a VMM domain profile.
- a.** True
  - b.** False
- 7.** Which steps need to be completed before a VM can be reassigned to an ACI-generated VDS? (Choose all that apply.)
- a.** A vSphere administrator needs to add the host on which the VM resides to the VDS.
  - b.** An ACI administrator performs a VMM domain association for the EPG needed to be pushed into ACI.
  - c.** ACI pushes a bridge domain into vCenter.
  - d.** ACI generates a VDS through configuration of a VMM domain profile.
- 8.** Which of the following resolution immediacy and deployment immediacy settings should be used for ESXi VMkernel interfaces with management services enabled, assuming that the VMkernel will reside on an ACI-generated VDS?
- a.** Immediate, Immediate
  - b.** On Demand, On Demand
  - c.** Pre-Provision, Immediate
  - d.** Immediate, On Demand

- 9.** True or false: ACI should not be used to deploy multiple virtual distributed switches that use the same name into a vCenter instance.
- a.** True
  - b.** False
- 10.** True or false: ACI can be used to create multiple LACP port channels on a VDS.
- a.** True
  - b.** False

## Foundation Topics

### Understanding Networking in VMware vSphere

With the movement of servers into hypervisors, the need arises for certain switching functions to be performed within hypervisors. For example, a hypervisor with several virtual machines (VMs) in a particular VLAN needs to know how to forward unicast traffic between the VMs or to a destination on the physical network and how to forward broadcast traffic.

Additional services that virtual switches need to provide include control over network bandwidth allocations for virtual machines and visibility and monitoring functions such as NetFlow.

In the vSphere ecosystem, hypervisors are referred to as *ESXi servers* or *ESXi hosts*. The two traditional flavors of virtual switches within this ecosystem are the vSphere standard switch (vSwitch) and the vSphere distributed switch (VDS).

## Note

VMware has additional virtual switch solutions that can be implemented in vSphere, such as the N-VDS and the NSX-T VDS. Earlier versions of vSphere supported Cisco Nexus 1000v. Cisco also has hypervisor-agnostic solutions that can be deployed in vSphere. These solutions, however, are all beyond the scope of the DCACI 300-620 exam.

## Understanding vSphere Standard Switches

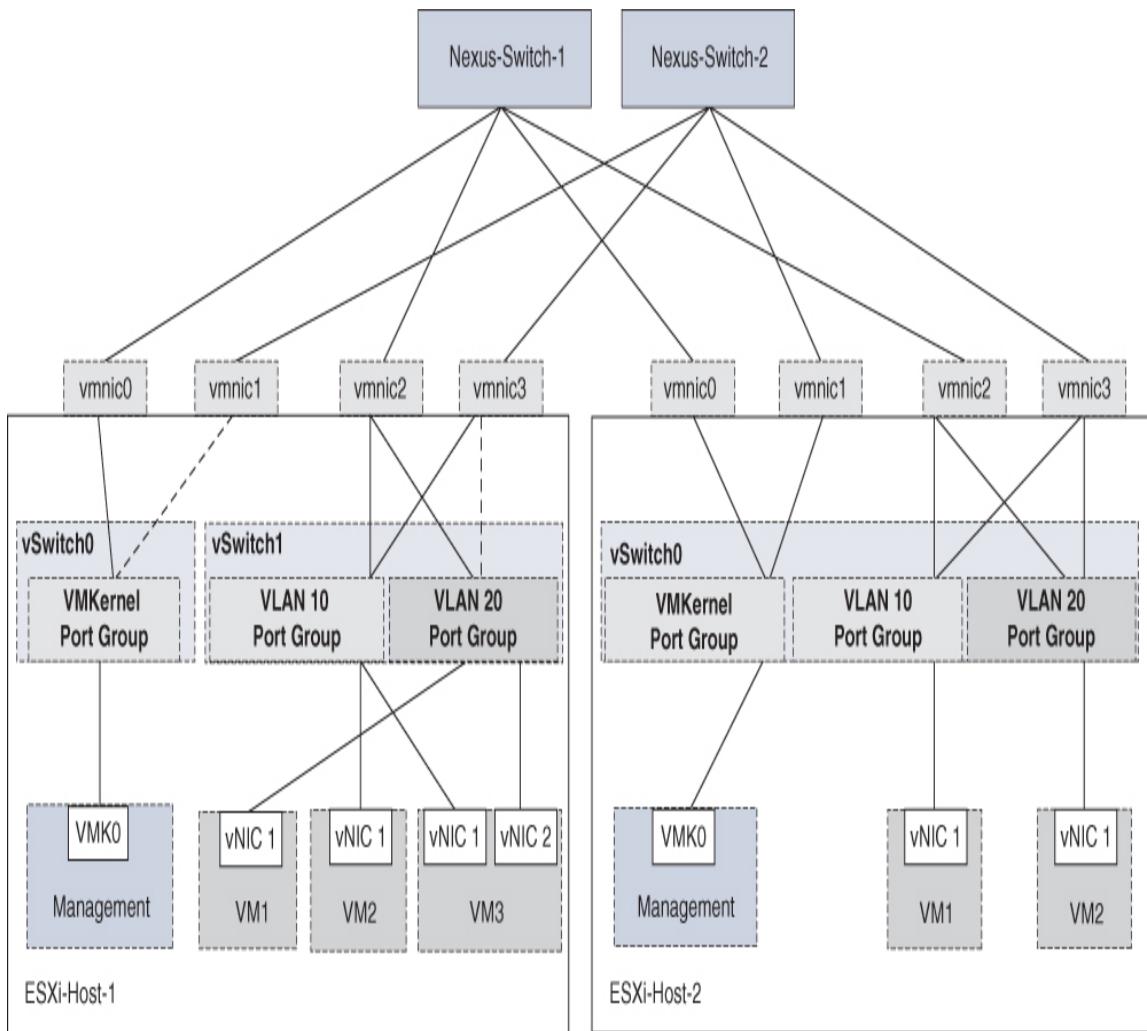
A **vSphere standard switch (vSwitch)** is a basic Layer 2 virtual switch instantiated by and limited in scope to a single hypervisor. It can be used to enable ESXi host IP connectivity with other hosts or endpoints. It can also enable basic connectivity for virtual machines.



When an administrator creates a vSphere standard switch, the system defaults to label such switches with the prefix vSwitch. The term **vmnic** describes a physical network interface card connecting an ESXi server to a physical switch. You can think of vmnic as a string used in the naming of virtual switch uplinks within ESXi hypervisors. To enable network connectivity for VMs, administrators configure virtual network adapters (vNICs) within them and associate those vNICs with port groups. A **port group** is a group of ports with similar policy requirements. For example, a user might create a port group for endpoints in VLAN 10. **VMkernel adapters** are logical interfaces that

enable transmission, receipt, or processing of hypervisor system traffic. Examples of ESXi system traffic include management, vMotion, IP storage, fault tolerance, and vSAN. Like virtual machines, VMkernel adapters need to be associated with port groups.

A lot of the terminology presented in the previous paragraph can be best understood through analysis of [Figure 11-1](#), which provides an overview of the vSphere standard switch architecture using two ESXi hosts. In this example, the ESXi hosts are not part of a vCenter server cluster. VM1 on ESXi-Host-1 is in VLAN 20 because its vNIC has been associated with the VLAN 20 port group. Because VM3 has been associated with the same port group as VM1, ESXi-Host-1 can locally switch traffic between the two virtual machines using vSwitch1. By default, all port groups leverage the interface teaming and failover policies of their parent virtual switch. However, these teaming and failover policies can be overridden at the port group level. For example, [Figure 11-1](#) shows that the port group for VLAN 20 within ESXi-Host-1 has an active path through vmnic2 to a top-of-rack switch named Nexus-Switch-1 and a standby path to Nexus-Switch-2. The standby path becomes active only if vmnic2 fails. Meanwhile, the VLAN 10 port group on the same host has both uplinks always active. Another point worthy of note from [Figure 11-1](#) is that ESXi-Host-1 has two standard switches deployed, while ESXi-Host-2 has only one. The nature of vSphere standard switches and the fact that they can be configured differently across hosts means that configuration drift across ESXi hosts that leverage these types of virtual switches is a serious possibility.



**Figure 11-1** *vSphere Standard Switch Architecture*

### Note

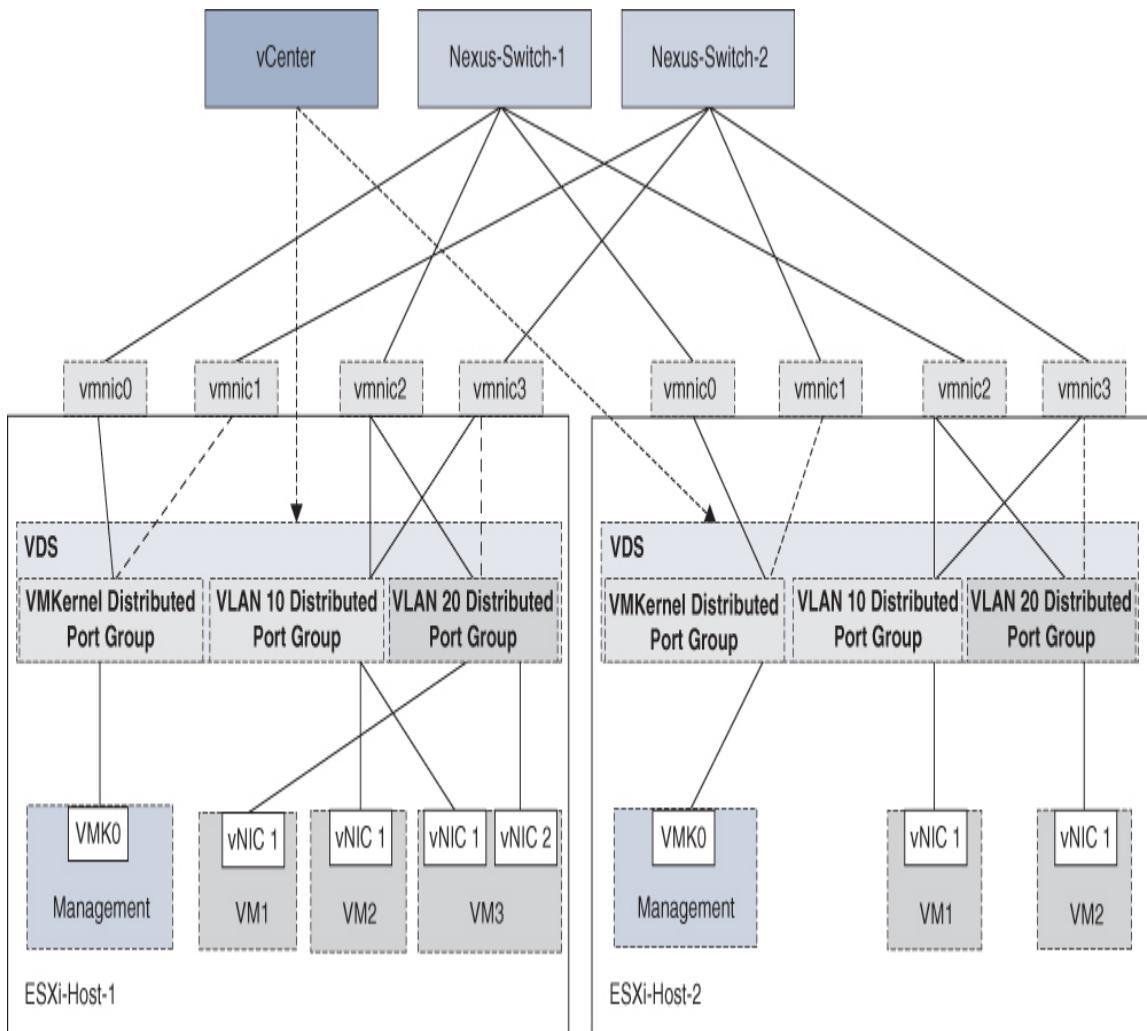
Figure 11-1 may suggest that the number of VLAN IDs in an environment determines the number of port groups. In reality, multiple port groups can be created to accommodate different policies for endpoints within a given VLAN.

## Understanding vSphere Distributed Switches

**Key Topic**

ESXi servers can be standalone hypervisors or can be centrally managed by vCenter. Since the overwhelming majority of ESXi hypervisor deployments are centrally managed, it stands to reason that there should be a way to also manage virtual switches through vCenter. And that is exactly what a vSphere distributed switch does. A **vSphere distributed switch (VDS)** is a virtual switch that is created and centrally managed by vCenter with the switch data plane residing within the ESXi hypervisors. When an administrator updates a setting on a VDS, the change is propagated to all hosts associated with the VDS. This approach to virtual switching eases management and minimizes the possibility of configuration drift across hosts in the environment.

Figure 11-2 presents the conceptual (not literal) architecture of a VDS. Note that the majority of concepts are the same as those covered in the previous section on vSphere standard switches. One notable difference between the two switch types beyond the fact that a VDS is managed by vCenter is that port groups on a VDS are called *distributed port groups*. Another difference reinforced by Figure 11-2 is that a distributed port group whose teaming and failover setting has been modified will have the same setting across all hosts. Finally, one difference not depicted here but that is still significant is that the vSphere distributed switch supports more advanced functionality, such as private VLANs, LACP, and NetFlow.



**Figure 11-2** *vSphere Distributed Switch Architecture*

### Note

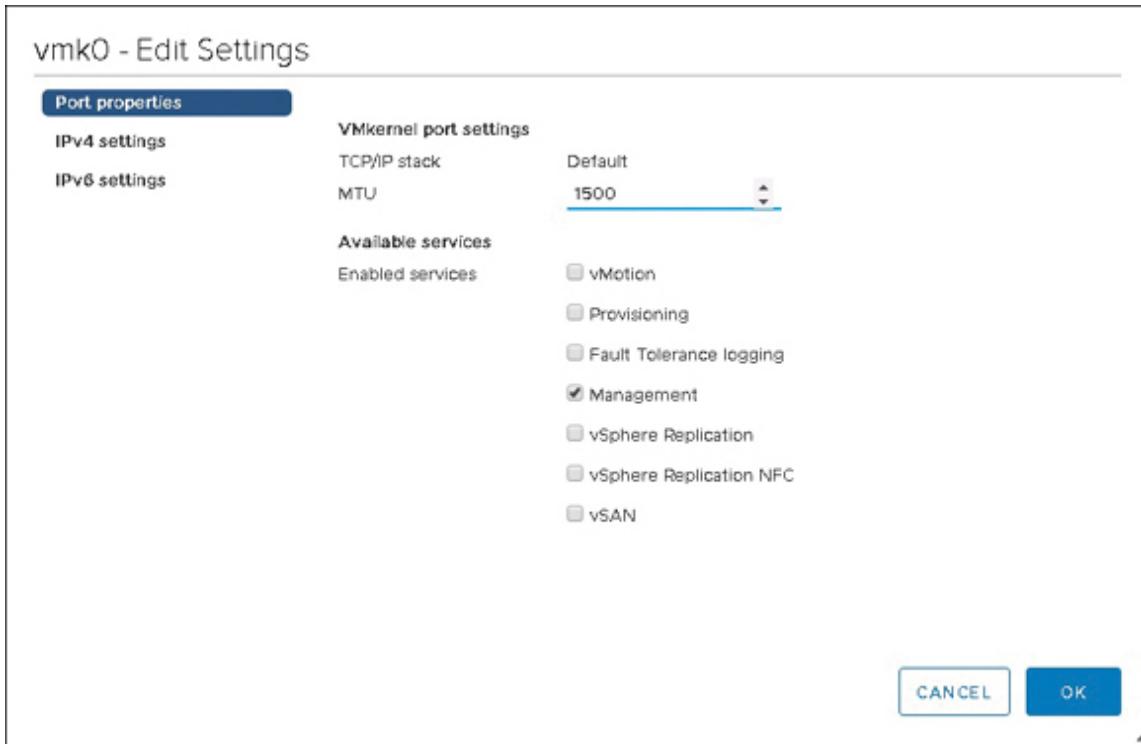
As of the time of writing, the DCACI 300-620 exam blueprint calls out DVS and not VDS as an exam topic. For the purpose of the exam, both terms refer to the same thing.

## Understanding vSphere System Traffic

The DCACI 300-620 exam does not really test your expertise with vSphere. Nonetheless, it seems a reasonable expectation that candidates who push VDS configuration changes from ACI to vCenter understand how to build the minimum ESXi server connectivity necessary for hypervisors to be able to communicate with vCenter. Without basic IP connectivity, ESXi servers have no way to receive updates on VDS configuration changes.



To enable ESXi communication with vCenter, ESXi servers need a VMkernel interface with management services enabled. This typically (but not always) is an interface named vmk0, which is created at the time of host deployment. [Figure 11-3](#) shows the settings of a VMkernel interface used for management of an ESXi host. One of the settings shown is the TCP/IP Stack parameter. If you need multiple VMkernel adapters for various services and different default gateways for each VMkernel adapter, the adapters with such requirements need to be assigned to different TCP/IP stacks. On the left side of the image is a link that allows navigation to the IPv4 Settings page, which is where you allocate an IP address and default gateway to the interface.



**Figure 11-3** A VMkernel Adapter with Management Services Enabled

Aside from the management service, the only other VMkernel service that all (or very close to all) vCenter-managed ESXi environments have enabled is vMotion. With vMotion, servers are able to migrate virtual machines between one another. The vMotion service can be enabled on the same VMkernel interface as management traffic or on a separate VMkernel interface. A detailed analysis of requirements for vMotion is beyond the scope of this book.

### Note

All images that depict vCenter and ESXi settings in this book have been taken with vCenter Release 6.7.

# **Impact of vCenter Failure on Production Traffic**

A failure of vCenter does not impact IP connectivity for endpoints behind vSphere standard switches since these types of virtual switches are local to each hypervisor and have no dependency on vCenter.

A failure of vCenter also does not halt IP connectivity of operational virtual machines and VMkernel interfaces on distributed virtual switches. It does, however, impact certain hypervisor services, such as vMotion and Distributed Resource Scheduler (DRS), that rely on vCenter.

When the vCenter server gets its network access through a VDS that it manages, a chicken-or-egg conundrum has been created that may impact the recoverability of the environment under very specific failure scenarios. An analysis of vSphere virtual port bindings can help illuminate how this type of challenge can be addressed.

## **Understanding Port Bindings in vSphere**

Just like physical switches, distributed switches have the concept of ports—although in the case of distributed switches, the ports are virtual. **Port bindings** determine when virtual machines in a port group get allocated to a virtual port on a VDS. There are three different types of port bindings in vSphere:



- **Static binding:** A virtual port is immediately assigned and reserved for a VM the moment it is assigned to a distributed port group if the distributed port group is configured with a static binding. This virtual port assignment remains valid unless the VM is removed from the distributed port group, even if the VM is migrated between servers. Note that VMs cannot be associated with a distributed port group configured for static bindings unless the vCenter instance managing the relevant VDS is operational. Static binding is the default binding type for new distributed port groups and is ideal for general use.
- **Dynamic binding:** With dynamic bindings configured on a port group, a virtual port is assigned to a virtual machine only after the VM powers on and its vNIC moves to an operational status. This port binding type has been deprecated since ESXi Version 5.0.
- **Ephemeral binding:** If a distributed port group has been configured with an ephemeral binding, a virtual port is created and assigned to a virtual machine by the ESXi host on which the VM resides (not by vCenter) once the VM powers on and its vNIC moves to an operational state.



As a best practice, all general use VMs placed on a VDS should be assigned to port groups configured for static bindings. Critical VMs such as vCenter itself can also be assigned to a distributed port group configured for static bindings. But if vCenter lives on a VDS it itself manages, administrators should consider assigning it and other critical VMs that vCenter relies on to a distributed port group configured for ephemeral bindings. At the very least, a

distributed port group can be created for vCenter purely for recovery operations and can be left without VM assignments.

## Note

In some environments, vCenter and management VMs may be placed on vSphere standard switches. Port groups on vSphere standard switches have ephemeral bindings.

As shown in [Figure 11-4](#), you can configure an ephemeral port group in vSphere by selecting Ephemeral - No Binding as the port binding type. You can access the New Distributed Port Group wizard shown in this figure by right-clicking a VDS under the Network tab in the vSphere web client, selecting Distributed Port Group, and clicking New Distributed Port Group.

New Distributed Port Group

✓ 1 Name and location  
2 Configure settings  
3 Security  
4 Traffic shaping  
5 Teaming and failover  
6 Monitoring  
7 Miscellaneous  
8 Ready to complete

Configure settings  
Set general properties of the new port group.

Port binding: Ephemeral - no binding  
(default)

VLAN  
VLAN type: VLAN  
VLAN ID: 299

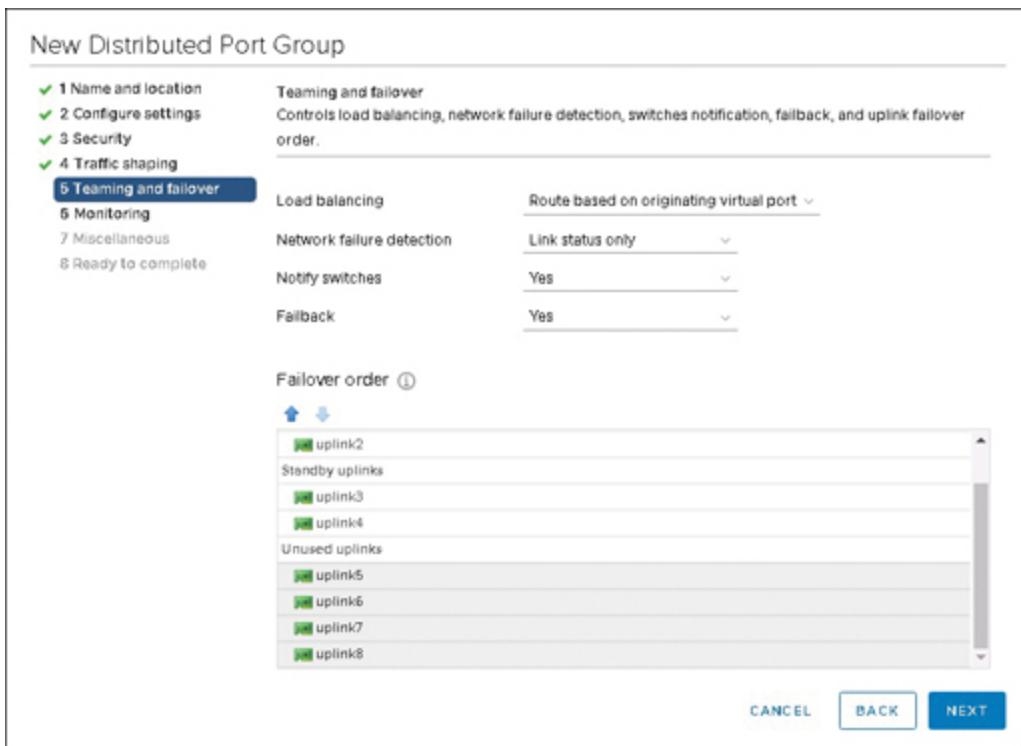
Advanced  
 Customize default policies configuration

CANCEL BACK NEXT

**Figure 11-4** Selecting Ephemeral - No Binding as the Port Binding Type

## Understanding Teaming and Failover Policies

Some of the most important aspects of port group deployments are the teaming and failover policies. [Figure 11-5](#) shows some of the teaming and failover settings available for configuring a distributed port group on a VDS.



**Figure 11-5** Teaming and Failover Settings for Port Groups

**vSphere load balancing** enables you to choose how uplink ports on the virtual switch are used to distribute network traffic originating from virtual machines and VMkernel interfaces. As of the time of this writing, there are five load-balancing methods to choose from:

- **Route based on originating virtual port:** With this load-balancing method, each VM or VMkernel interface is pinned to a single uplink, based on the virtual port assigned to it. If a switch uplink fails, all virtual machines and VMkernel interfaces associated with it are then repinned to an alternate operational uplink. This is the default load-balancing method in vSphere and is preferred in most environments. With this method configured across port groups, physical network switches do not need to configure ports facing ESXi servers in a port channel, even though this load-balancing method can lead to active/active forwarding of traffic across uplinks due to endpoints being pinned to different uplinks.
- **Route based on IP hash:** When you think of the word *hash*, you should automatically think of port channeling. That is exactly what this load-balancing option is. When used on port groups in a vSphere standard switch, this option results in the deployment of a static port channel. LACP port channeling, on the other hand, *is* supported on virtual distributed switches.
- **Route based on source MAC hash:** With this option, the virtual machine MAC address determines the uplink over which network-bound traffic should be forwarded. Compared to routing based on originating virtual port, this load-balancing option is rarely used because it is resource intensive and requires that the virtual switch calculate an uplink for each packet received from virtual machines.
- **Use explicit failover order:** This option is actually not a load-balancing method. When this method is

selected, the virtual switch uses the failover order and sends traffic out the first uplink in the Active adapters list. If all interfaces in the Active adapters list are down, the virtual switch can use uplinks in the Standby adapters list, in order of priority.

- **Route based on physical NIC load:** This is a slight variation of routing based on originating virtual port, in which the virtual switch also tests the load on uplinks every 30 seconds and repins traffic to an alternate uplink if the load exceeds 75% of uplink bandwidth of the physical interface. This load-balancing option is only available with vSphere distributed switches.

Other teaming and failover settings that can be configured for a vSwitch or DVS and that are shown in [Figure 11-5](#) are detailed in [Table 11-2](#).



**Table 11-2** vSphere Teaming and Failover Settings

S	Description
e	
t	
i	
n	
g	

## S Description e t ti n g

N Specifies one of two ways with which a virtual switch can detect network uplink failures. The first, Link Status Only, relies on the operational state of the link. The second option, Beacon Probing, sends probes into the network through its uplinks to see if other uplinks receive the probes. If an uplink is unable to receive the transmitted probes, it is deemed to have suboptimal connectivity.  
F Use of beacon probing is very uncommon because it requires at least three uplinks to allow the virtual switch to effectively determine which of the uplinks has failed.  
The beacon probing option should never be used in combination with port channeling.

## D e t e ct io n

## S Description e t ti n g

N	When set to Yes, prompts the virtual switch repinning a virtual machine to a different uplink following an uplink failure to send a Reverse ARP (RARP) packet into the network in the hope that upstream switches update their CAM tables faster.
W it c h e s	F When all interfaces in the Active Uplinks list have been determined to be non-operational and virtual machines and VMkernel interfaces have been pinned to interfaces in the Standby list, a fallback setting of Yes specifies whether the virtual switch should fall back to interfaces in the Active uplinks list if they become available.

## S Description e t ti n g

F Specifies how traffic should be rerouted when an adapter fails. There are three possible adapter states that can be configured in the teaming and failover page: Active  
v Uplinks, Standby Uplinks, and Unused Uplinks. Active  
e uplinks actively forward traffic, while standby uplinks are  
r only used if all active uplinks fail. An uplink that has been  
O added to the Unused Uplinks list is never used for traffic  
r in a port group, even if both the active and standby  
d uplinks fail.

## Understanding VMM Integration

Why bother with VMM integration? For companies that intend to focus heavily on whitelisting, the splitting of traditional VLANs into EPGs necessitates the creation of more and more port groups within vSphere, thereby increasing the operational burden for vSphere administrators. VMM integration not only minimizes this burden but introduces a level of network policy automation that enables vSphere administrators to delegate the task of vSphere networking to network engineers.

But there is also another argument that can be made in favor of VMM integration. At the end of the day, the vSphere ecosystem relies on the network to get server traffic from point A to point B. If end-to-end connectivity relies on the network in the first place, there really is no reason to define network policy twice (once in the physical network and once in the virtual network). The end results of VMM integration for companies that don't seek to implement whitelisting are less duplication of effort, better end-to-end visibility into connectivity issues, faster policy deployment, and more involvement of network teams in virtual networking.

## **Planning vCenter VMM Integrations**

This section looks at some of the things ACI administrators should take into account when planning VMM integrations.

First, it is important to understand that ACI treats each VMM integration definition as a domain object. This does not necessarily mean that any time ACI integrates with a separate vCenter instance, a separate VMM domain is required. ACI does allow multiple vCenter instances to be grouped under a single VMM domain. Creating a single VMM domain and configuring multiple vCenter instances under the VMM domain makes a lot of sense if all vCenter instances require the same distributed port groups.

For instance, say that you have 10 instances of vCenter that have ESXi hosts in a fabric. Five of these instances are dedicated to production applications, and 5 are dedicated to development purposes. In this case, it might make sense to deploy 2 VMM domains: one for the production environment and one for the development environment. Each VMM domain would have 5 associated vCenter instances. A design like this ensures that when an ACI administrator adds the VMM domain association to an EPG, distributed port

groups are pushed to 5 separate vCenter instances simultaneously. Alternatively, if 5 separate VMM domains had been defined for the production environment, the administrator might find that 5 different VMM domains need to be associated with each EPG that needs to be pushed into the production vCenter instances.

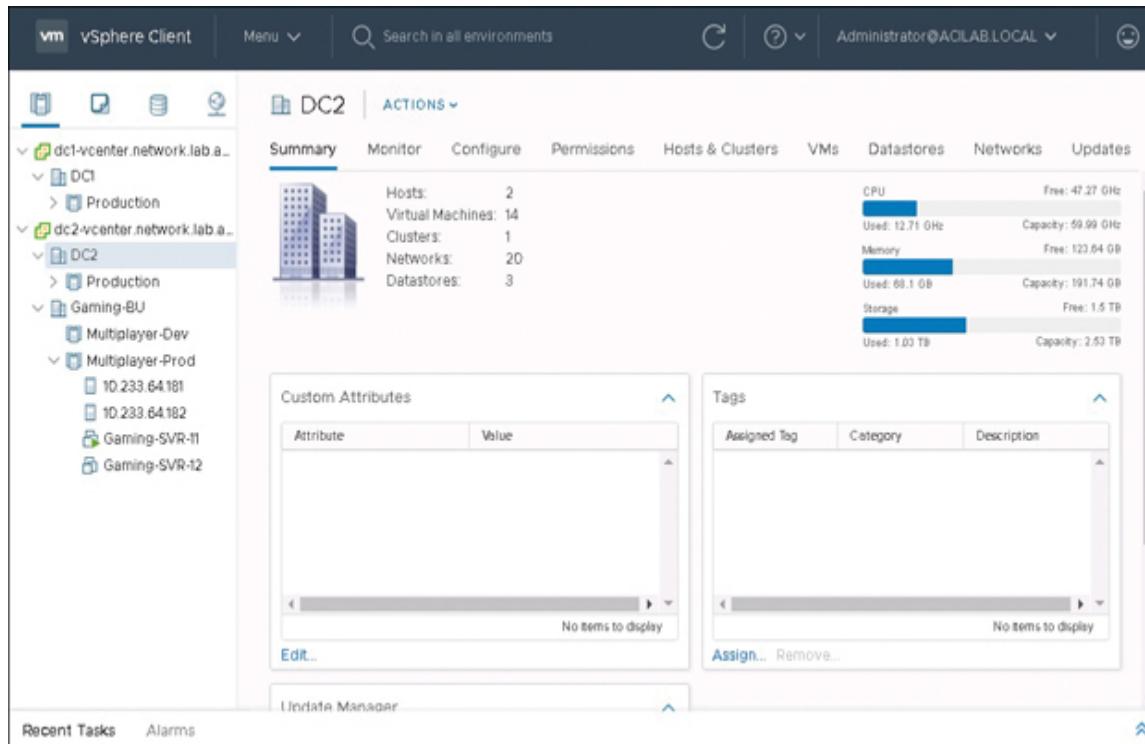
On the topic of domains, remember that an administrator needs to allocate one (and only one) VLAN pool to each VMM domain. In the case of VMM domains, it makes a lot of sense to use dynamic VLAN pools. This approach allows ACI to determine what VLAN ID to allocate to each distributed port group it creates within vSphere and therefore dramatically reduces the time spent implementing network policy. By default, ACI dynamically allocates VLANs to distributed port groups, but this behavior can be modified if administrators need to statically allocate a specific VLAN ID to a distributed port group.

It should be clear by now that a single VMM domain can span multiple vCenter instances, but can a single vCenter instance have multiple VMM domain integrations? The answer to this question is yes! To understand why someone might want to perform multiple VMM integrations into a vCenter instance, let's take a look at the vSphere object hierarchy.



In vSphere, each virtual distributed switch is actually tied directly to a data center object, not to a vCenter instance. A data center object is itself a child object of a vCenter instance. Under each data center object, various clusters of ESXi hosts may exist. [Figure 11-6](#) shows a data center object named DC2, which has been selected by a user. DC2

appears below its parent vCenter instance, dc2-vcenter. A host cluster named Production appears under DC2. A new data center object named Gaming-BU is also shown under dc2-vcenter. Two new clusters and several hosts have also been staged for this new data center object.



**Figure 11-6** Data Center, Cluster, and ESXi Host Hierarchy in vCenter

As noted earlier, a VDS is tied directly to a data center object. If both the DC2 and Gaming-BU data centers shown in [Figure 11-6](#) required access to exactly the same EPGs, on the surface it would seem reasonable to think that you could create two references to the vCenter instance dc2-vcenter, each pointing to separate data centers within a single VMM domain. However, this would not work because vCenter expects each VDS to have a unique name across vCenter. This, therefore, would be a case in which at least two VMM integrations within the vCenter instance would be required in spite of the requirement for access to the same EPGs.

So, if an ACI fabric can integrate into a vCenter instance using multiple VMM domains, is the earlier example in which five production vCenter integrations were made possible using a single VMM domain always accurate? The answer is no. The example assumes that each vCenter instance has only a single data center object that requires VMM integration in the first place.

## What Happens After VDS Deployment?

Once an ACI administrator creates a VMM domain and associates one or more vCenter instances with the VMM domain, ACI leverages vSphere APIs to execute the creation of a VDS in each of the specified vCenter instances.

Just because a VMM domain and therefore a VDS have been created does not mean that EPGs will be available for consumption within vSphere. To create distributed port groups, an ACI administrator needs to first add VMM domain associations to the desired EPGs.

Eventually, a vSphere administrator needs to log in to vCenter and add a number of ESXi hosts to the VDS. The process of adding ESXi hosts to the VDS involves specifying VDS uplinks for each host and possibly migrating VMkernel adapters, or **VM VNICS**, to distributed port groups on the VDS.

Once a vSphere administrator adds a virtual machine or VMkernel interface to an ACI-generated distributed port group, ACI classifies the endpoint as part of the EPG, ensuring that the endpoint inherits the network access specified by any associated contracts.

## Understanding Immediacy Settings

ACI is built for extreme scalability. Some ACI fabrics are home to hundreds (and others thousands) of ESXi servers. For ACI to be able to support such large environments, policies should be deployed where they are needed—not everywhere.

Imagine an environment with 1000 physical servers and 10,000 EPGs spread across 50 leaf switches. Do all 50 of these leaf switches need to know about every single VRF, bridge domain, EPG, and contract detail associated with these 10,000 EPGs? Do all leaf switches need to push all these policies into hardware? The answer to both of these questions is no! To better understand why, let's take a look at resolution immediacy and deployment immediacy.

In ACI ***resolution immediacy*** defines when policies, such as VLANs, VXLAN bindings, contracts, and filters, are downloaded to leaf switches. Three resolution immediacy options are available in ACI:



- **Pre-Provision:** This resolution immediacy option specifically relates to VMM integration. In the context of ACI integrations with vSphere, Pre-Provision prompts policies to be downloaded to leaf switches even if a vCenter instance has not been defined within the VMM domain configuration. Only leaf switches with access policies referencing an AAEP associated with a VMM domain download policies as a result of EPG pre-provisioning.
- **Immediate:** This resolution immediacy option specifies that EPG policies are downloaded to a leaf switch once an ESXi host has been added to a VDS created by the

APIC and the VDS has been verified via LLDP or CDP to be adjacent to the leaf switch.

- **On Demand:** This resolution immediacy option specifies that policies are downloaded to a leaf switch only when an ESXi host has been validated via LLDP or CDP to be attached to a VDS created by the APICs and at least one VM VNIC has been assigned to a distributed port group on the VDS.

### Note

When there is a single layer of intermediary devices between a leaf switch and an APIC-generated VDS and the resolution immediacy settings used are Immediate and On Demand, ACI can sometimes use LLDP and CDP neighborships reported by the VDS as well as data in LLDP and CDP packets arriving on leaf ports to reverse engineer the data path and determine whether policies should be downloaded to the leaf.

Note that the Immediate and On Demand resolution immediacy options both rely on ESXi host VMkernel adapters being operational and ACI trunking the VMkernel management VLAN down to the hypervisors. Otherwise, ACI would never receive LLDP or CDP information from the hypervisor to begin with. This chicken-and-egg headache necessitates use of the resolution immediacy Pre-Provision option for the VMM domain association for the hypervisor management EPG.



Use the resolution immediacy setting Pre-Provision for critical EPGs to which hypervisors require access in order to

reach a fully operational state. This rule also applies to EPGs that are home to critical vSphere servers such as vCenter, LDAP, and any vCenter database servers.

### Note

In small-scale deployments in which hardware resource utilization requirements are never expected to surpass the capabilities of deployed leaf switches, there is little reason not to use the resolution immediacy option Pre-Provision for all EPGs. Also, if hypervisors connect indirectly to ACI leaf switches through an intermediate device and the LLDP and CDP capabilities of the intermediate device are suspect, consider using the resolution immediacy option Pre-Provision.

Whereas resolution immediacy determines whether policies should be downloaded to a leaf, **deployment immediacy** governs whether policies should be pushed into hardware.

When policies, such as VLANs, VXLAN bindings, contracts, and filters, have been downloaded to a leaf switch, deployment immediacy specifies when the policy is actually pushed into the hardware policy content-addressable memory (CAM). There are two deployment immediacy configuration options:



- **Immediate:** This deployment immediacy option specifies that the policy is programmed in the hardware policy CAM as soon as the policy is downloaded in the leaf software.

- **On Demand:** This deployment immediacy option specifies that the policy should be programmed in the hardware policy CAM only after the switch receives a packet through the data path. This setting helps optimize the hardware space.

## Connecting ESXi Servers to the Fabric

It is critical for DCACI candidates to understand that VMM integration does not eliminate the need for ACI access policies. Without access policies, ACI has no way to know to which leaf switches it needs to deploy policy. Based on the load-balancing algorithm configured, ACI downlinks to ESXi servers may be port channels. They may be non-aggregated ports. Whatever the case may be, access policies need to be deployed.

If a leaf downlink connects to a vSphere standard switch or a VDS that will not be managed by ACI, it does not need to have an AAEP association that provides access to a VMM domain. If, on the other hand, a leaf downlink connects to an ESXi uplink that will be associated with an ACI-generated VDS, the VMM domain triggering the creation of the VDS creation needs to have also been associated with the AAEP.

To sum up, it is important to understand that it is distributed port group generation and deployment assignments (and not deployment of access policies) that is automated by VMM integration.

## Configuring Connectivity to ESXi in UCS Domains

Engineers sometimes find the task of selecting optimal load-balancing settings for virtual standard switch or VDS uplinks to be mystifying. The Cisco Unified Computing System (UCS)

B-Series server architecture offers a good Cisco-centric case in point.

Say that an engineer has been asked to provide guidance on virtual switch load balancing prior to deployment of VMM integration. The hypervisors housing the virtual switches reside within a UCS domain. The engineer has studied the basics of UCS B-Series architecture and learns that fabric interconnect links northbound to the network attach to leaf switches via vPCs. He or she also learns about port channeling within the system and may then incorrectly assume that routing based on IP hash should be selected as the virtual switch load-balancing algorithm. What is being misinterpreted in this case is that the port channeling that occurs within UCS domains is between the I/O modules and fabric interconnects, not the server vNICs northbound. In addition, the fabric interconnect vPC connectivity with leaf switches has no direct relationship with the virtual switch load-balancing capabilities.

To avoid these types of pitfalls, engineers should only analyze the port channeling capabilities between hypervisor uplinks and the directly connected device, even if the hypervisor uplinks do not directly connect to ACI leaf switches.

When in doubt, routing based on originating virtual port is often the best load-balancing method for ESXi virtual switches.

## **Integrating ACI into vSphere Using VDS**

Now that we have covered the basics of vSphere networking and VMM integration, let's take a look at how VMM integration works in practice.

# Prerequisites for VMM Integration with vSphere VDS

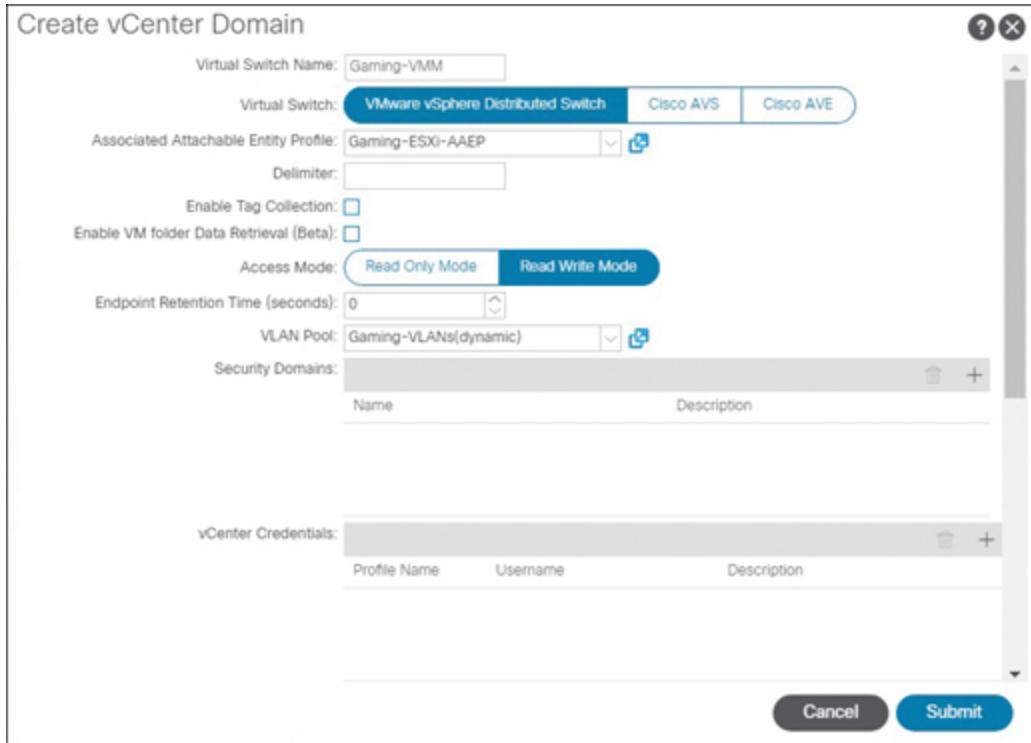
The following items are basic prerequisites for VMM integration with a VMware VDS and should be addressed before you configure VMM integration:

- Either in-band or out-of-band management should have been configured for the APICs.
- The APICs need to be able to reach vCenter from their out-of-band or in-band management connections.
- All leaf switches to which ESXi servers connect should have been discovered and should be fully operational.

## Configuring a VMM Domain Profile

To create a VMM domain profile, navigate to **Virtual Networking > Inventory > VMM Domains**, right-click VMware, and select Create vCenter Domain.

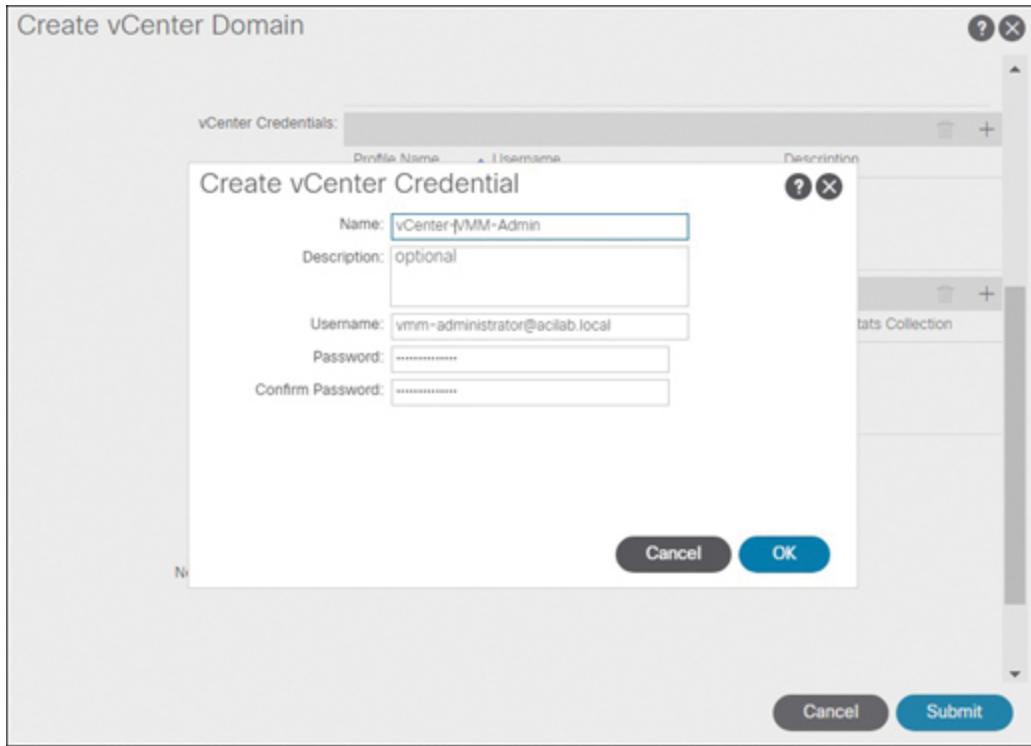
In the Create vCenter Domain wizard, populate the Virtual Switch Name field with a VDS name, select VMware vSphere Distributed Switch as the virtual switch type, create an AAEP or select a pre-created AAEP from the Associated Attachable Entity Profile drop-down box, and select the access mode and VLAN pool. Then click on the + symbol in front of the vCenter Credentials section. [Figure 11-7](#) shows a sample configuration of these parameters. Note that the Access Mode parameter defines whether ACI will be making configuration changes to the VDS. The Read Only Mode setting enables ACI to gain visibility into a VDS it does not manage as well as the hypervisors and VMs associated with it.

**Key Topic**

**Figure 11-7** vCenter Domain Creation Wizard

On the Create vCenter Credential page, shown in [Figure 11-8](#), populate the Username, Password, and Confirm Password fields with credentials that allow ACI to log in to vCenter and conduct networking tasks. Enter a descriptive name describing the credentials or differentiating it from other saved credentials if multiple integrations have been implemented and click OK. Ideally, vCenter credentials used should be for an account that does not expire.

**Key Topic**



**Figure 11-8** The Create vCenter Credentials Page

### Note

VMM integration with vSphere does not require full admin privileges. However, the list of privileges required is long enough that it is not suited for coverage on an exam like the DCACI 300-620 exam. Review the latest Cisco ACI Virtualization Guide to get an idea of the minimum privileges required for the tasks relevant to a given environment.

Back in the Create vCenter Domain page, scroll further down and click the + symbol in front of the section titled vCenter. In the Add vCenter Controller page, shown in [Figure 11-9](#), enter a descriptive name for the vCenter instance being added and provide either the DNS name or IP address of the vCenter instance, select a VDS version from the DVS Version drop-down box, enter the case-sensitive data center

name in the Datacenter field (exactly as it appears in vCenter), select a management EPG if in-band management should be used for APIC connectivity to this vCenter instance, and then select the previously created credentials from the Associated Credential drop-down box and click OK.



Add vCenter Controller

vCenter Controller

Name: dc2-vCenter

Host Name (or IP Address): 10.233.64.50

DVS Version: vCenter Default

Stats Collection:  Disabled  Enabled

Datacenter: Gaming-BU

Management EPG: select an option

Associated Credential: vCenter-VMM-Admin

Cancel OK

The screenshot shows the 'Add vCenter Controller' dialog box. At the top right are help and close buttons. The main area contains fields for 'Name' (dc2-vCenter), 'Host Name (or IP Address)' (10.233.64.50), 'DVS Version' (vCenter Default), 'Stats Collection' (radio button for 'Disabled' is selected), 'Datacenter' (Gaming-BU), 'Management EPG' (a dropdown menu showing 'select an option'), and 'Associated Credential' (vCenter-VMM-Admin). At the bottom are 'Cancel' and 'OK' buttons.

**Figure 11-9** The Add vCenter Controller Page

Back in the Create vCenter Domain page, populate the Number of Uplinks field with the maximum number of uplinks. If no value is entered, a default value of 8 uplinks is assumed. Select a value for Port Channel Mode if port channeling or MAC pinning will be used. In the vSwitch Policy field, select whether LLDP or CDP should be enabled on the VDS, and finally create a NetFlow exporter policy if you want the VDS to send NetFlow data to a NetFlow

collector. Then click Submit to execute the necessary API calls to generate the VDS within the vCenter instances selected, as shown in [Figure 11-10](#).



Create vCenter Domain

vCenter Credentials:

Profile Name	Username	Description
vCenter-VM...	vmm-administrator@acilab.local	

vCenter:

Name	IP	Type	Stats Collection
dc2-vCenter	10.233.64.50	vCenter	Disabled

Number of Uplinks:

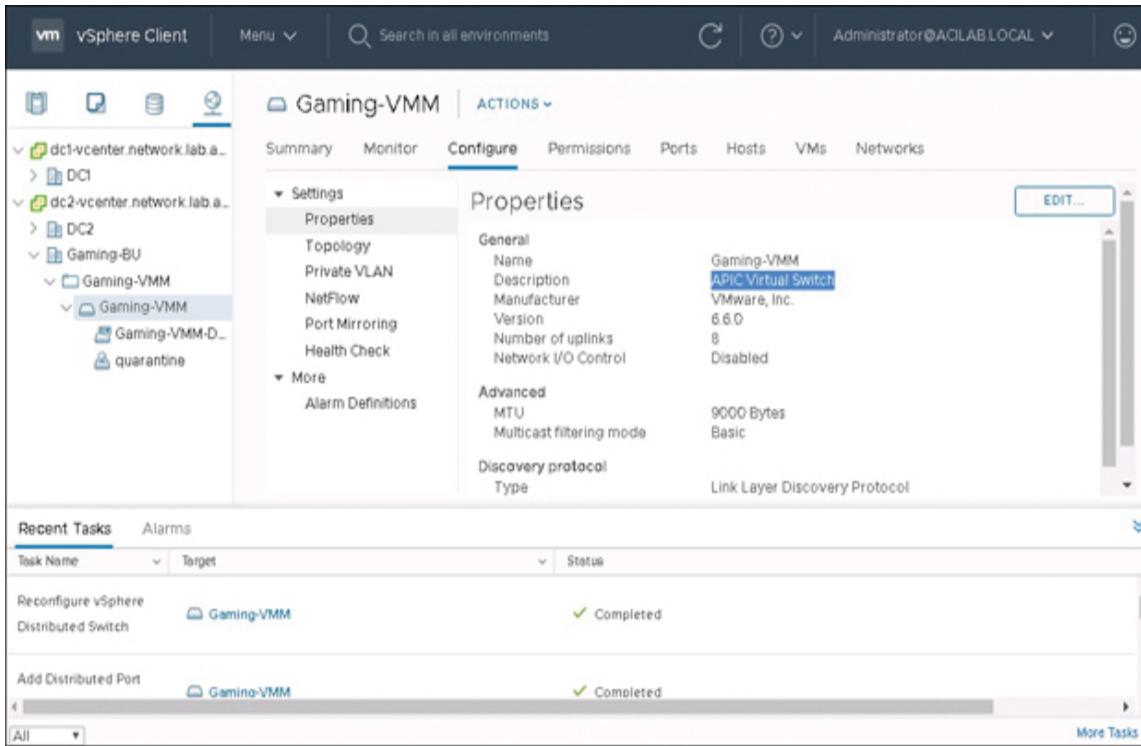
Port Channel Mode:

vSwitch Policy:  CDP  LLDP  Neither

NetFlow Exporter Policy:

**Figure 11-10** vCenter Domain Creation Wizard, Continued

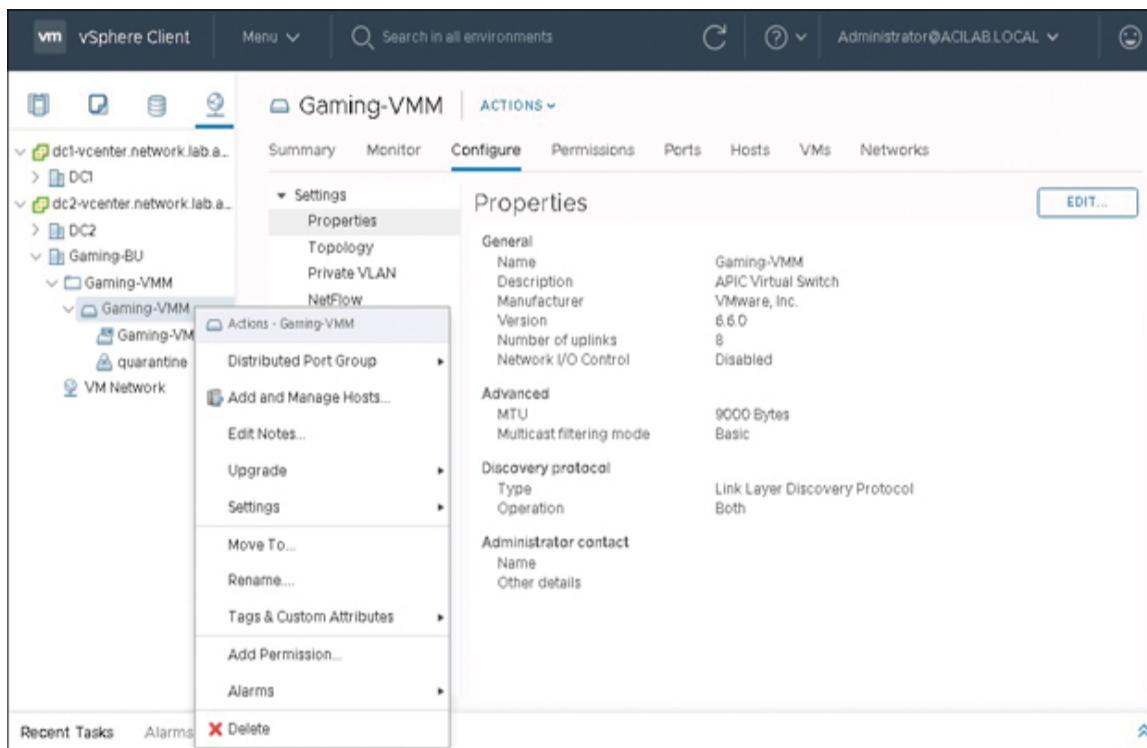
In the Networking tab in each configured vCenter instance, a VDS with the name specified earlier can be validated as having been created, as shown in [Figure 11-11](#). No hypervisors or VMs should be associated with the VDS at this point. By default, the VDS should have two distributed port groups: one called quarantine and the other dedicated to hypervisor uplinks.



**Figure 11-11** Validating VDS Creation in vCenter

## Adding ESXi Hosts to a VDS

Once ACI has generated a VDS in vCenter, you can right-click on the VDS and select Add and Manage Hosts, as shown in [Figure 11-12](#).



**Figure 11-12** Navigating to the Add and Manage Hosts Wizard in vCenter

As indicated in [Figure 11-13](#), select the Add Hosts option and click Next.

## Gaming-VMM - Add and Manage Hosts

✓ 1 Select task

2 Select hosts

3 Manage physical adapters

4 Manage VMkernel adapt...

5 Migrate VM networking

6 Ready to complete

### Select task

Select a task to perform on this distributed switch.

Add hosts

Add new hosts to this distributed switch.

Manage host networking

Manage networking of hosts attached to this distributed switch.

Remove hosts

Remove hosts from this distributed switch.

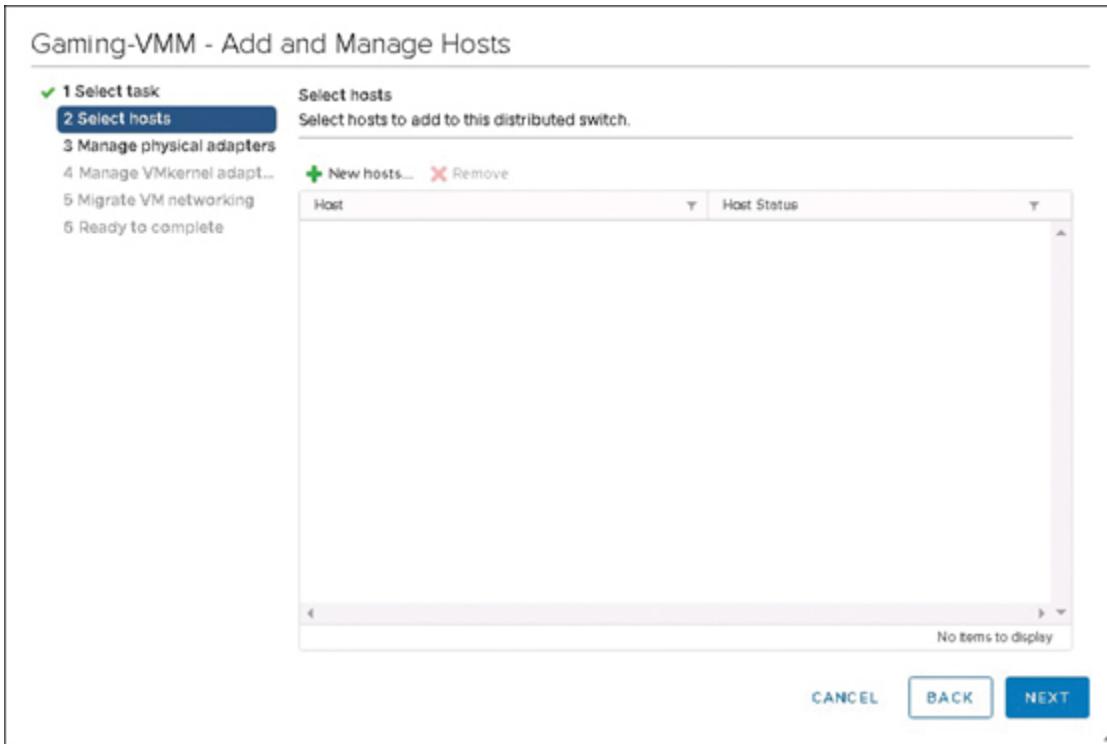
CANCEL

BACK

NEXT

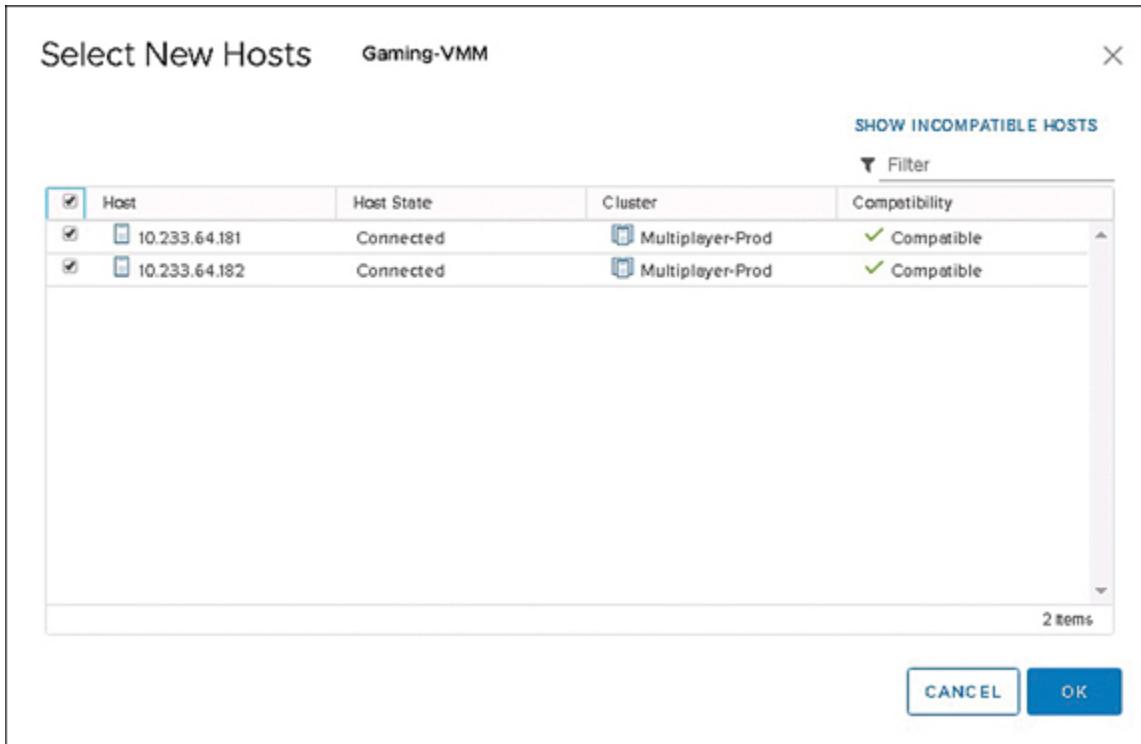
**Figure 11-13** *Selecting Add Hosts*

Click New Hosts to select available hosts to add to the VDS, as shown in [Figure 11-14](#).



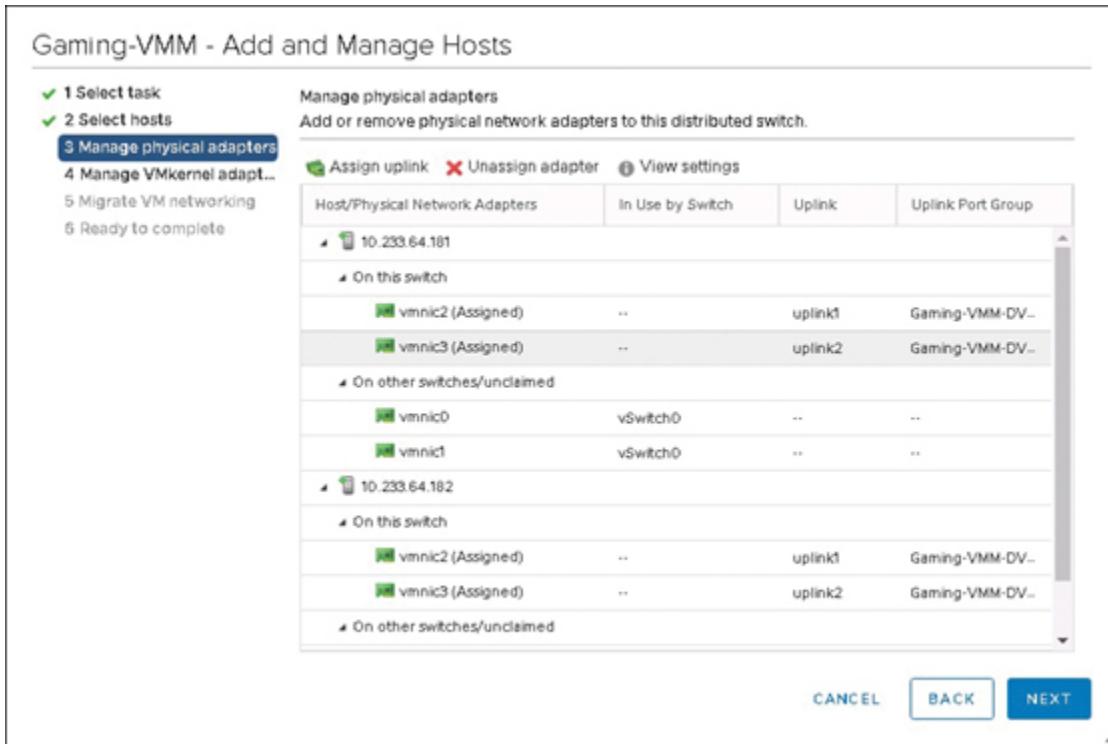
**Figure 11-14** Clicking New Hosts

In the Select New Hosts page, enable the checkbox in front of each ESXi host that should be added to the VDS, as shown in [Figure 11-15](#), and click OK.



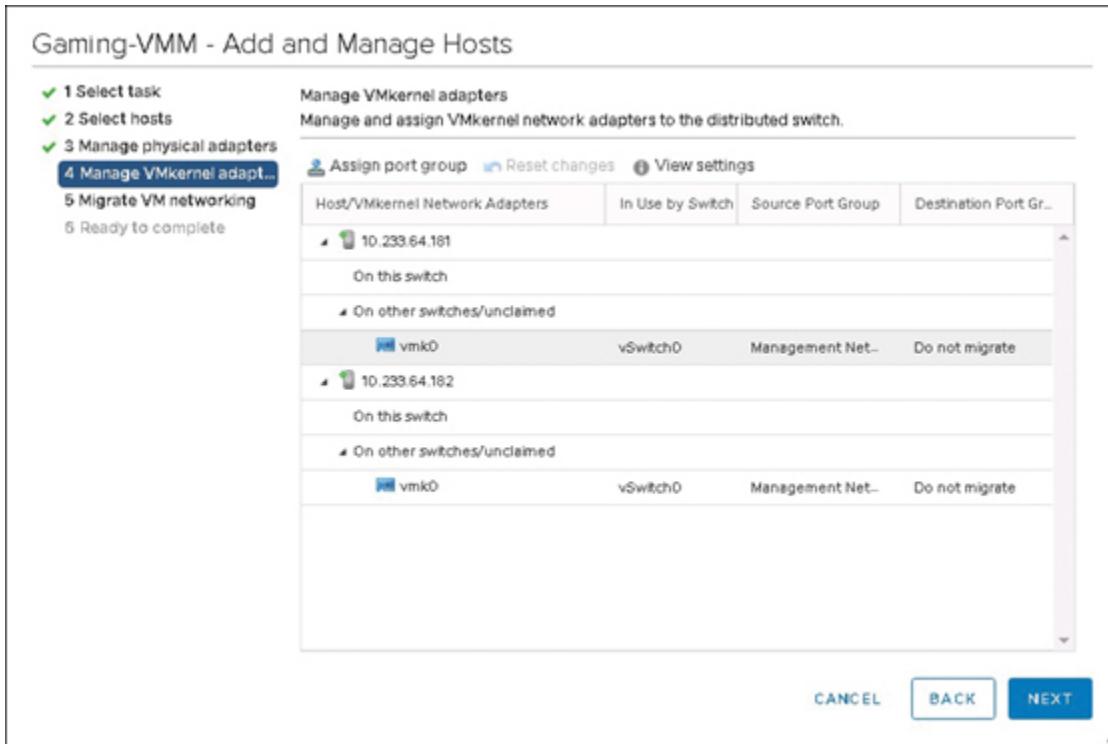
**Figure 11-15** Choosing the Hosts to Add on the Select New Hosts Page

Next, click the Assign Uplink command button to define which vmnics should be assigned to the VDS. Note in [Figure 11-16](#) that vmnic interfaces previously assigned to another virtual switch can be reassigned in this window. Just make sure not to reassign interfaces governing VMkernel connectivity for management or any other critical services. To move on to the next step, click Next.



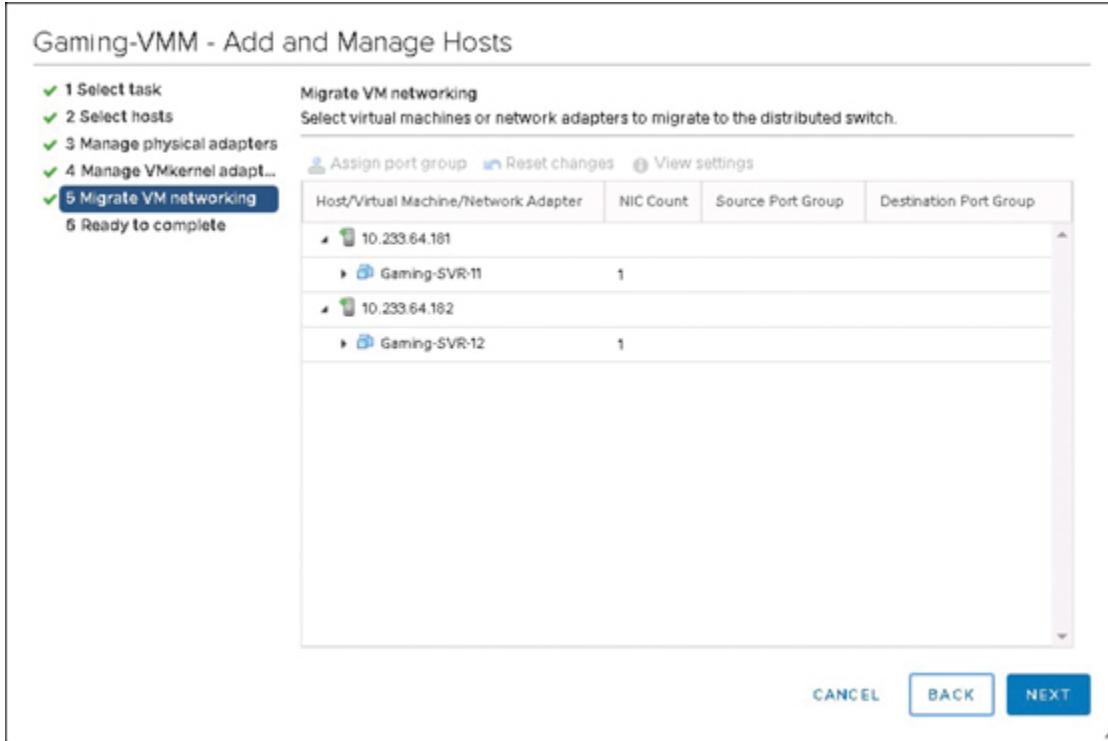
**Figure 11-16** Assigning Uplinks to a VDS

In the Manage VMkernel Adapters page, shown in [Figure 11-17](#), VMkernel interfaces can be migrated to the new VDS. It is sometimes best to first assign uplinks, push port groups, and validate connectivity before migrating VMkernel adapters, but if proper access policies are in place, migration of VMkernel adapters can take place at this stage without issue. When you are ready to move on to the next step, click Next.



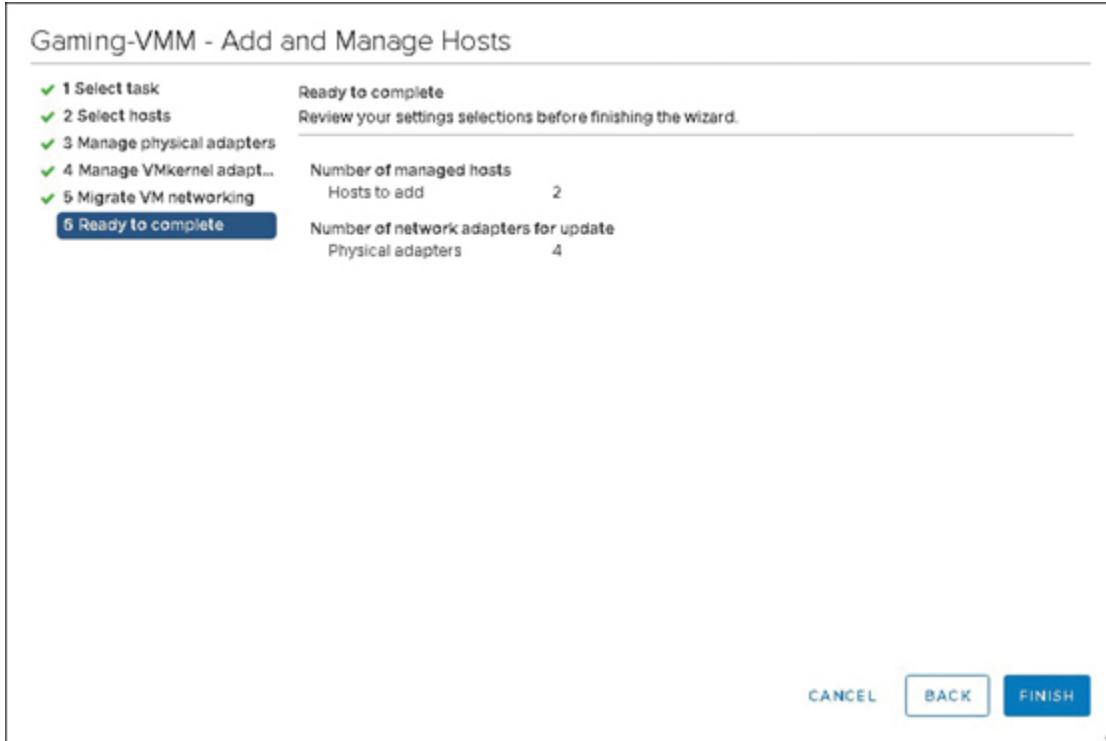
**Figure 11-17** The Manage VMkernel Adapters Page

The Migrate VM Networking page is shown in [Figure 11-18](#). If any VM vNICs should be migrated to VDS distributed port groups, you can select the target port groups. Because ACI has not yet pushed any EPGs into vCenter, you can move on to the next step for now.



**Figure 11-18** The Manage VM Networking Page

Figure 11-19 shows the changes the wizard is making. Click Finish to confirm and add the ESXi hosts to the VDS.

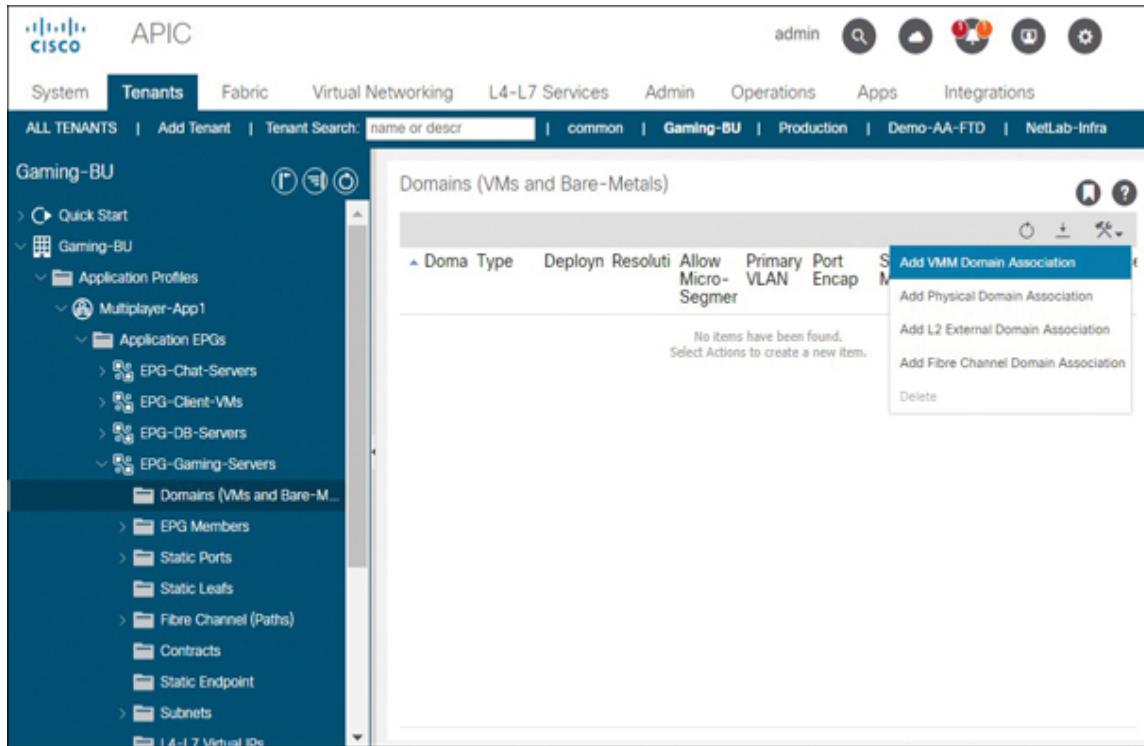


**Figure 11-19** *Confirming the Addition of ESXi Hosts to the VDS*

## Pushing EPGs to vCenter as Distributed Port Groups

Before VM traffic can be placed in EPGs through VMM integration, ACI needs to push desired EPGs into any vCenter instances defined by the VMM domain.

To push an EPG into vCenter, navigate to the desired EPG and expose its subfolders. In the Domains folder, right-click the Tools menu and select Add VMM Domain Association (see [Figure 11-20](#)).



**Figure 11-20** Navigating to the Add VMM Domain Association Page

From the VMM Domain Profile drop-down, select the desired VMM domain with which the EPG should be associated, select Deploy Immediacy and Resolution Immediacy settings, select the VLAN mode (Dynamic indicates dynamic VLAN ID allocation), configure the port binding (the ACI default value for VDS integration is Static Binding, which is suitable for general-use virtual machines), and click Submit. [Figure 11-21](#) shows settings entered to enable the deployment of an EPG named EPG-Gaming-Servers into vCenter as a distributed port group.

**Key Topic**

Add VMM Domain Association

VMM Domain Profile: Gaming-VMM

Deploy Immediacy: Immediate  On Demand

Resolution Immediacy: Immediate  On Demand  Pre-provision

Delimiter:

Enhanced Lag Policy: select an option

Allow Micro-Segmentation:

Untagged VLAN Access:

VLAN Mode: Dynamic  Static

Port Binding: Dynamic Binding  Ephemeral  Default  Static Binding

Netflow: Disable  Enable

Allow Promiscuous: Reject

Forged Transmits: Reject

MAC Changes: Reject

Active Uplinks Order:   
Enter IDs of uplinks separated by comma

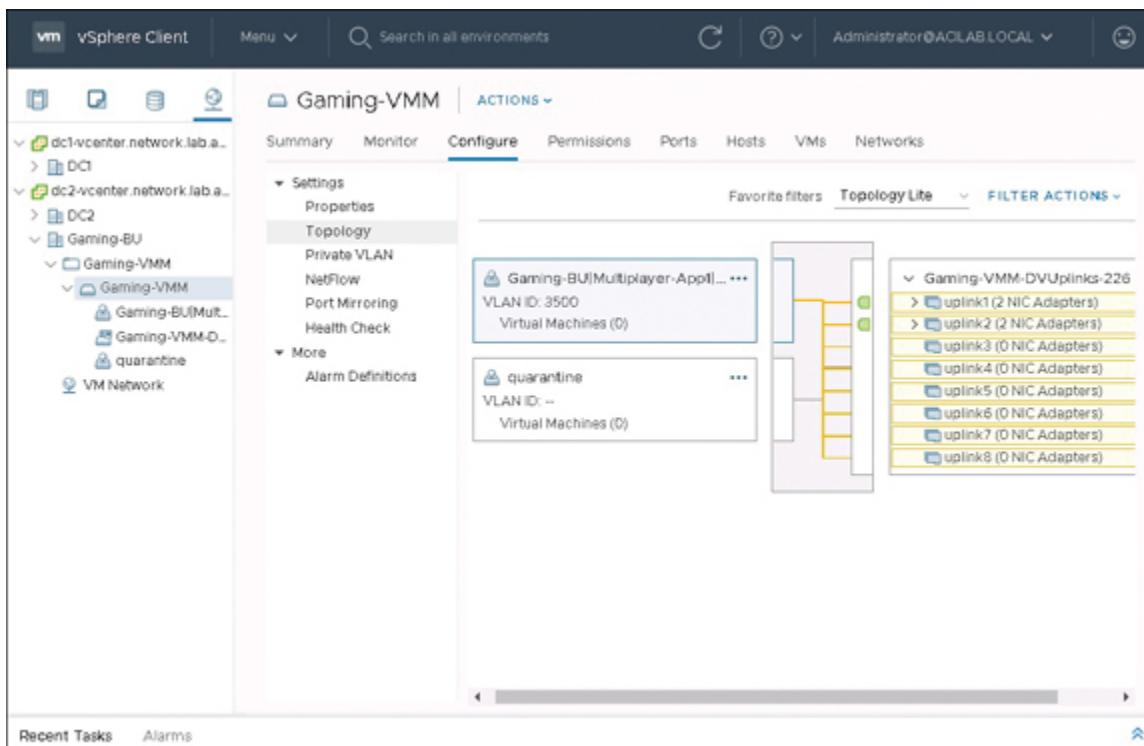
Standby Uplinks:   
Enter IDs of uplinks separated by comma

Custom EPG Name (Beta):

Cancel  Submit

**Figure 11-21** Pushing an EPG with Basic Settings into vCenter

Back in vCenter, you can expect to see a distributed port group created for the EPG. [Figure 11-22](#) shows that the EPG has indeed been pushed to vCenter and uses VLAN ID 3500. By selecting the distributed port group, you can verify the number of uplinks that can potentially be associated with the distributed port group. Because the Active uplinks and Standby uplink settings were left untouched, the first eight possible uplinks have been chosen as candidates for active/active forwarding of traffic over hypervisor uplinks. You may notice that virtual machines have not yet been associated with the distributed port group.



**Figure 11-22** Verifying Distributed Port Group Generation in vCenter

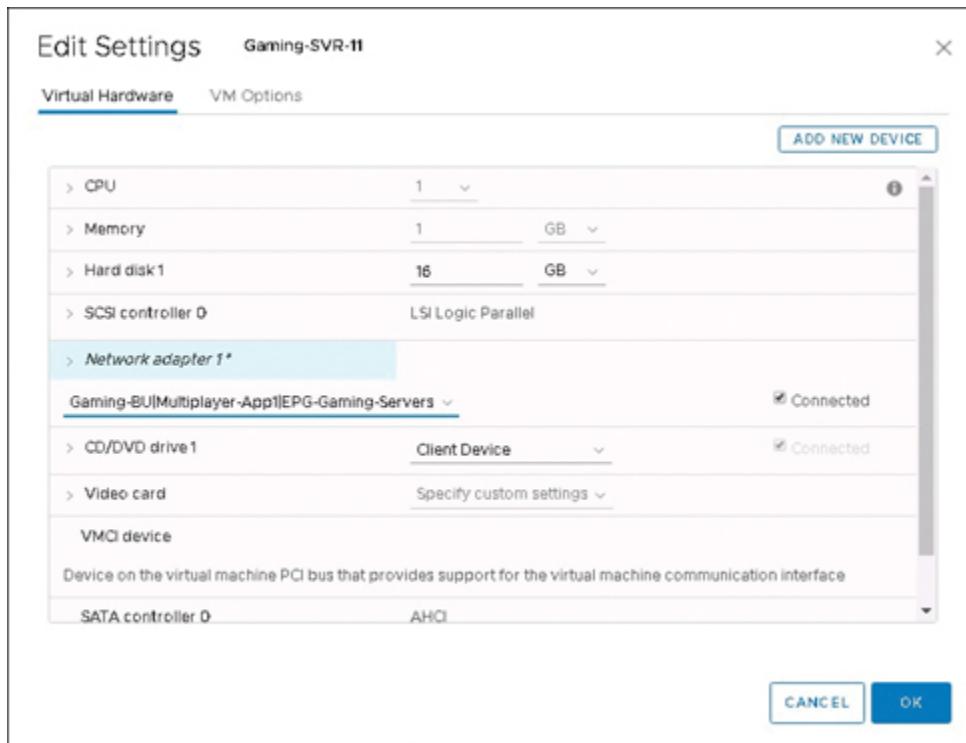
### Note

By default, ACI names the distributed port group using the format *tenant/application/epg*. The character separating the parameters is called a *delimiter* character. In recent versions of ACI code, the naming can be customized using the Custom EPG Name field.

## Assigning VMs to Distributed Port Groups

In vCenter, navigate to the leftmost tab, called Hosts and Clusters, select a VM whose vNIC needs to be reassigned to the new distributed port group, click on the Actions menu, and select Edit Settings. Then, as shown in [Figure 11-23](#),

select the ACI-generated distributed port group under the desired network adapter configuration, and click OK.



**Figure 11-23** Reassigning a VM vNIC to a Distributed Port Group

If the VM has a valid IP address for the associated EPG, it should become available on the network and become visible in ACI. For validation of proper learning, navigate to the EPG, select the Operational tab, and review the list of detected endpoints in the Client End-Points subtab. If troubleshooting is required, first evaluate any faults at the VMM domain and EPG levels.



# Less Common VMM Domain Association Settings

Let's take a look at some more VMM domain association settings by pushing another EPG into vCenter. On the Add VMM Domain Association page shown in [Figure 11-24](#), enter an explicit delimiter character in the Delimiter field. By default, eight uplinks can potentially be active for each distributed port group if the Active Uplinks and Standby Uplinks settings remain unchanged. Selecting the comma-separated values 1,2 for Active Uplinks Order and 3,4 for Standby Uplinks ensures that the remaining uplinks (5 through 8) are unavailable to this particular distributed port group. Finally, select a value for the custom EPG name and click Submit.

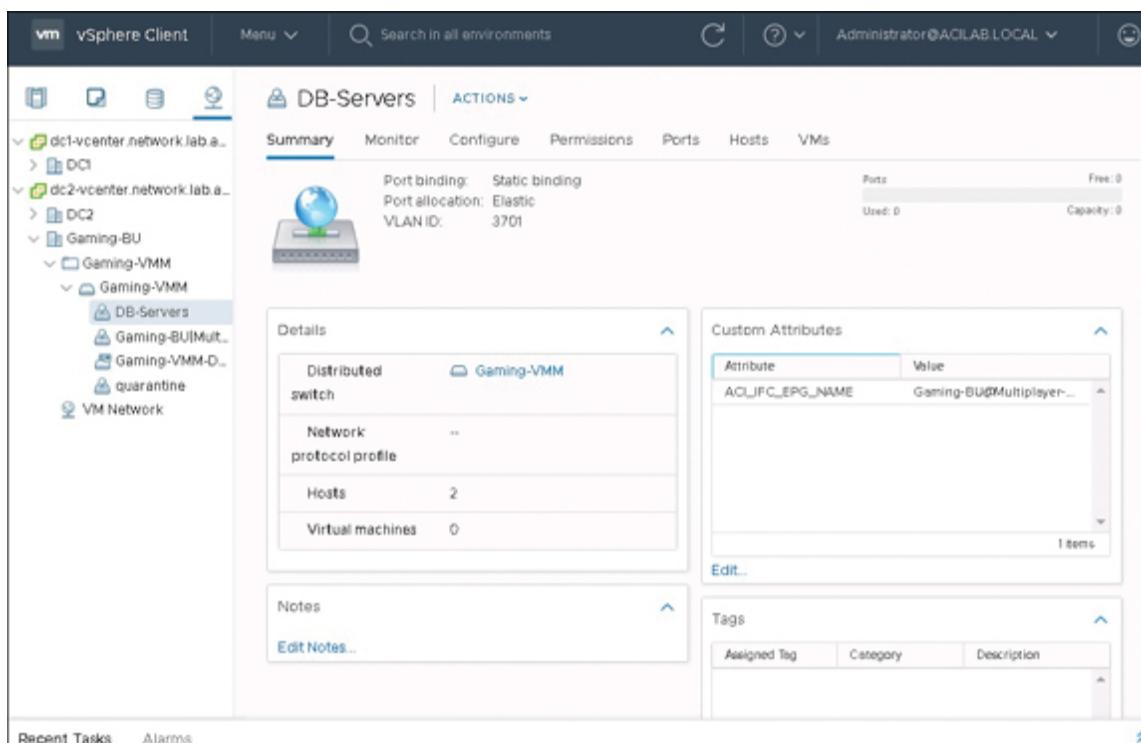
The screenshot shows the 'Add VMM Domain Association' dialog box. Key configuration settings include:

- VMM Domain Profile: Gaming-VMM
- Deploy Immediacy: Immediate
- Resolution Immediacy: Pre-provision
- Delimiter: @
- Enhanced Lag Policy: select an option
- Allow Micro-Segmentation:
- Untagged VLAN Access:
- VLAN Mode: Dynamic
- Port Binding: Default
- Netflow: Enable
- Allow Promiscuous: Reject
- Forged Transmits: Reject
- MAC Changes: Reject
- Active Uplinks Order: 1,2
- Standby Uplinks: 3,4
- Custom EPG Name (Beta): DB-Servers

At the bottom are 'Cancel' and 'Submit' buttons.

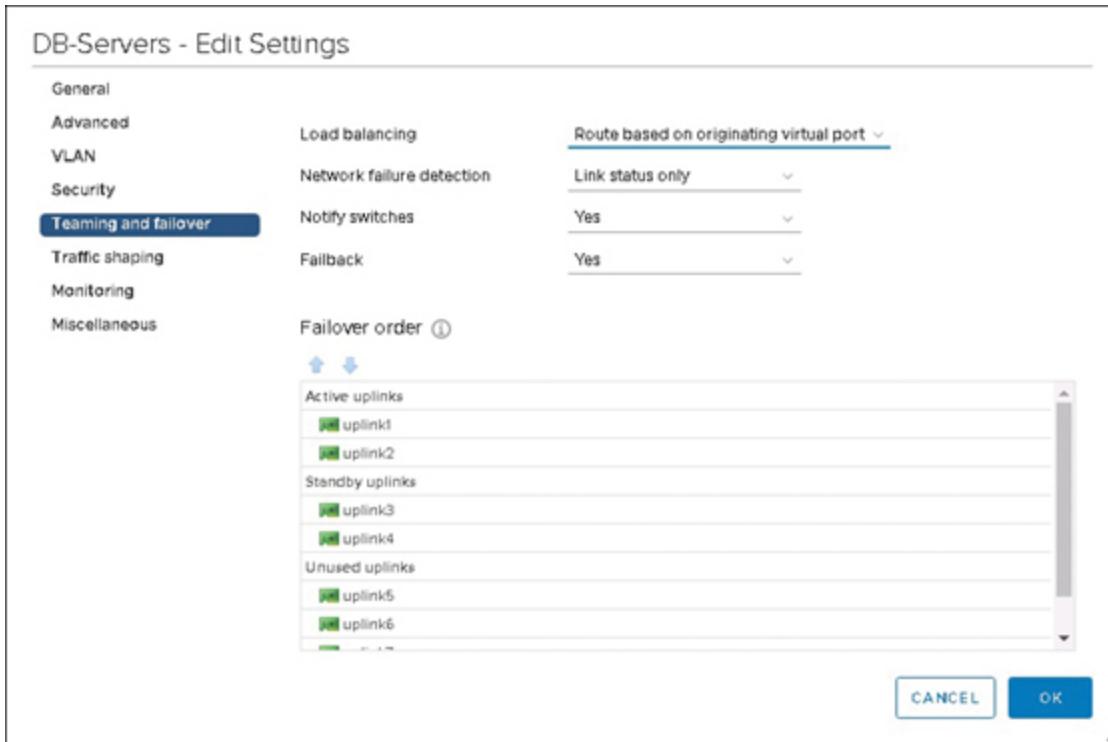
**Figure 11-24** Tweaking VMM Domain Association Settings

[Figure 11-25](#) shows the result of these changes. The resulting distributed port group generated in vCenter has a custom name. The attribute `ACI_IFC_EPG_Name` in the Custom Attributes view, however, continues to show the EPG name (using the new delimiter value) that would have been selected if the Custom EPG Name field had not been populated.



**Figure 11-25** Verifying the Result of Custom EPG Naming and Delimiter Modification

When editing the settings of the distributed port group, it is clear that the first two uplinks have been set as Active and appear in the order entered in ACI. The next two uplinks appear in the Standby uplinks list, and any further uplinks remain unused for this distributed port group (see [Figure 11-26](#)).

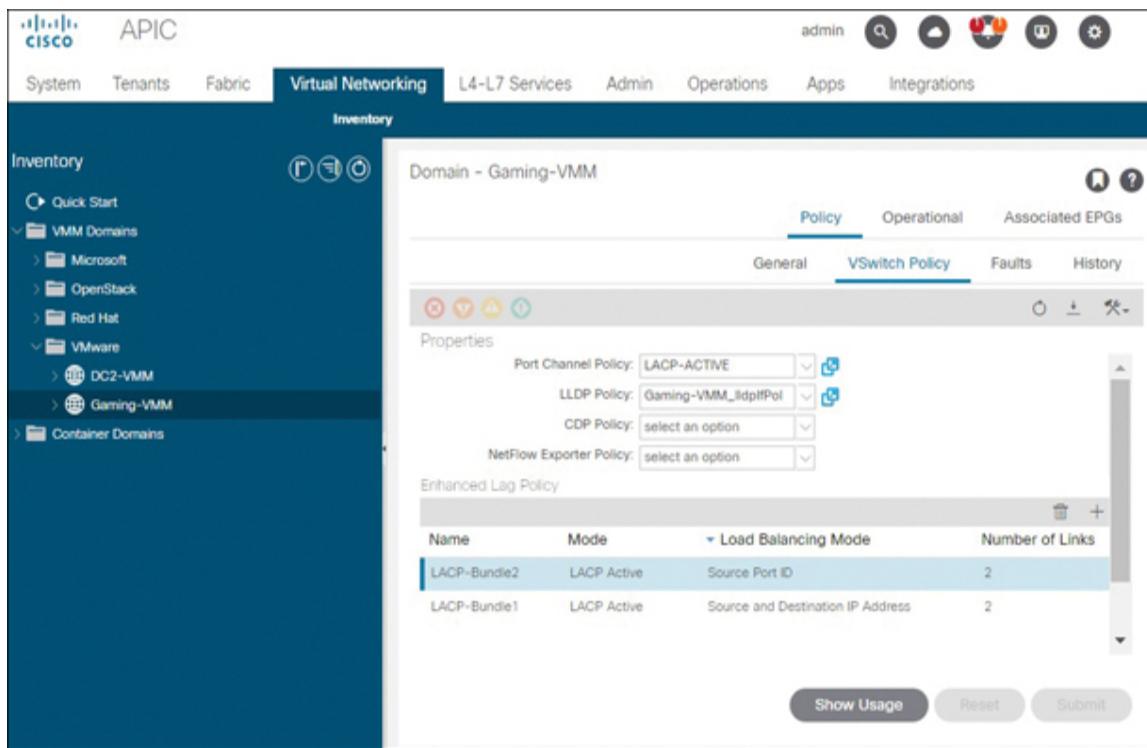


**Figure 11-26** Verifying the Result of Active Uplinks and Standby Uplinks Settings

## Enhanced LACP Policy Support

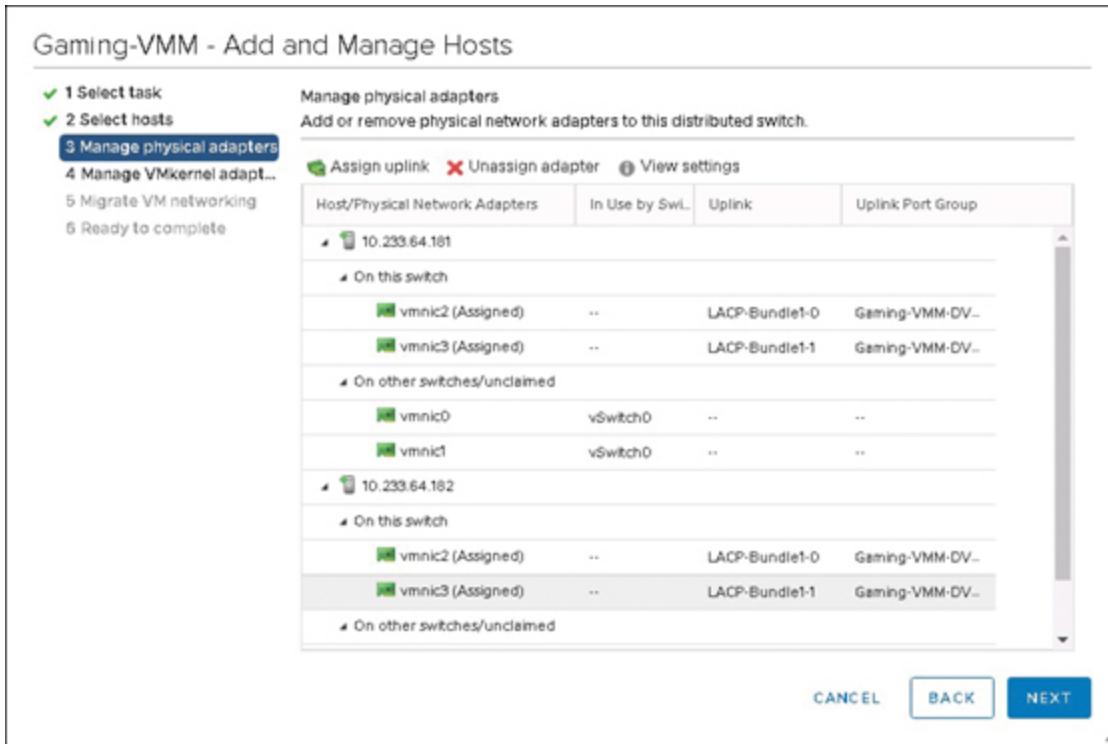
Let's say that you want an ACI-generated VDS to have multiple sets of uplink port channels and want to allow some traffic to flow over one port channel and other traffic to flow over a separate port channel. This capability involves enhanced LACP policy support.

To implement enhanced LACP policy support, navigate to Virtual Networking, select the desired VMM domain, click Policy, and select vSwitch Policy. Then select a port channel policy that enables LACP. Next, define a name and load-balancing mode for each logical LACP port channel that should be created on the VDS. [Figure 11-27](#) shows two sets of LACP uplink port channels being created.



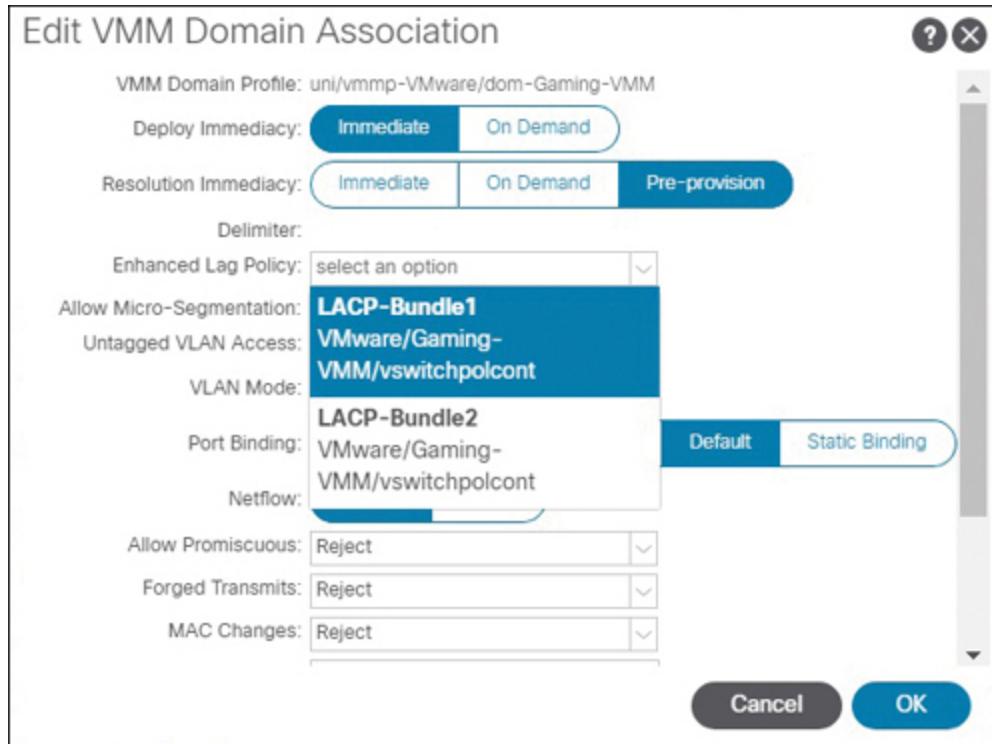
**Figure 11-27** Configuring an Enhanced LAG Policy

This is more likely a policy you would deploy before hosts are added to the VDS. When hosts are being added to the VDS, you can assign hypervisor uplinks to the uplink port groups that resemble the logical port channel name configured earlier. [Figure 11-28](#) shows two ports on each hypervisor being added to the port channel named LACP-Bundle1.



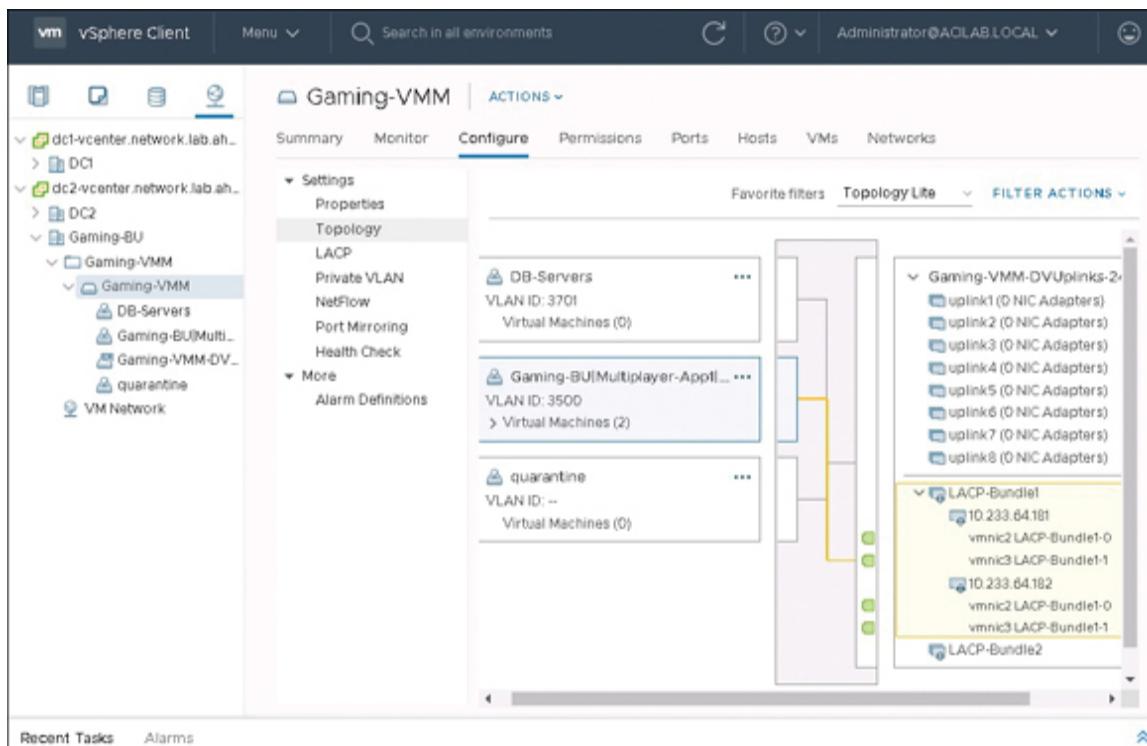
**Figure 11-28** Assigning ESXi Host Uplinks to a Link Aggregation Group

After ESXi hosts have been added to the VDS, navigate to an EPG that has been pushed to vCenter and edit its VMM domain association. [Figure 11-29](#) shows an EPG being associated with an enhanced LAG policy named LACP-Bundle1.



**Figure 11-29** Assigning an Enhanced LAG Policy to an EPG VMM Domain Association

As a result of this change, traffic in the distributed port group begins to flow solely over the selected uplink port channel, as shown in [Figure 11-30](#).



**Figure 11-30** Verifying Distributed Port Group Mapping to Uplinks

## Exam Preparation Tasks

As mentioned in the section “How to Use This Book” in the Introduction, you have a couple of choices for exam preparation: [Chapter 17, “Final Preparation,”](#) and the exam simulation questions in the Pearson Test Prep Software Online.

## Review All Key Topics

Review the most important topics in this chapter, noted with the Key Topic icon in the outer margin of the page. [Table 11-3](#) lists these key topics and the page number on which each is found.



**Table 11-3** Key Topics for Chapter 11

Key Topic Element	Description	Page Number
Paragraph	Defines some basic terminology important to understanding vSphere networking	395
Paragraph	Describes vSphere distributed switches	397
Paragraph	Describes the significance of VMkernel adapters with management services enabled	397
List	Describes port binding types in vSphere and lists the significance of each	400
Paragraph	Details best practices around use of static binding and ephemeral settings	400
List		401

	Lists and details the load-balancing methods available in vSphere	
<b>Table 11-2</b>	Describes vSphere teaming and failover settings	<b>402</b>
Paragraph	Describes the significance of the data center object in vSphere in VDS deployment and therefore VMM integration	<b>404</b>
List	Describes resolution immediacy and details relevant configuration options	<b>405</b>
Paragraph	Describes a key use case for the Pre-Provision resolution immediacy setting	<b>406</b>
List	Describes deployment immediacy and details the two configuration options involving deployment immediacy	<b>406</b>
<b>Figure 11-7</b>	Demonstrates steps to configure a basic VMM domain profile	<b>408</b>
<b>Figure 11-8</b>	Shows how to enter credentials into ACI to allow APICs to generate API calls against vCenter instances	<b>409</b>

<a href="#"><b>Figure 11-9</b></a>	Demonstrates how to enter vCenter access information in a VMM domain profile	<a href="#"><b>410</b></a>
<a href="#"><b>Figure 11-10</b></a>	Demonstrates further steps to configure a basic VMM domain profile	<a href="#"><b>410</b></a>
<a href="#"><b>Figure 11-21</b></a>	Demonstrates how an administrator can push an EPG into vCenter as a distributed port group	<a href="#"><b>416</b></a>
Paragraph	Calls out the location in the ACI GUI where VMM-learned endpoints in an EPG should appear	<a href="#"><b>418</b></a>

## Complete Tables and Lists from Memory

Print a copy of [Appendix C, “Memory Tables”](#) (found on the companion website), or at least the section for this chapter, and complete the tables and lists from memory. [Appendix D, “Memory Tables Answer Key”](#) (also on the companion website), includes completed tables and lists you can use to check your work.

## Define Key Terms

Define the following key terms from this chapter and check your answers in the glossary:

vSphere standard switch (vSwitch)

vmnic

VM vNIC

port group

VMkernel adapter

vSphere distributed switch (VDS)

port binding

vSphere load balancing

resolution immediacy

deployment immediacy

# Chapter 12

## Implementing Service Graphs

This chapter covers the following topics:

**Service Graph Fundamentals:** This section covers service graph theory, including the management models, benefits, and bridge domain configurations.

**Service Graph Implementation Workflow:** This section provides a high-level overview of the steps necessary to deploy service graphs.

**Service Graph Implementation Examples:** This section goes through several service graph deployment examples step by step to solidify the concepts.

This chapter covers the following exam topic:

- 4.3 Implement service graph (managed and unmanaged)

Traditional methods of inserting services devices such as firewalls and load balancers between endpoints often involve architecting the network to ensure that basic routing and switching rules guide traffic through firewalls. It is also typical to have no integration between traditional networks and stateful services devices.

ACI supports all the traditional methods of inserting Layer 4 through Layer 7 services between endpoints, but with service graphs, it also introduces new management models for L4-L7 services. These new management models enable deeper integration between the network and services devices, aligning the network with the ongoing industry shifts toward automation and cross-platform integration.

With service graphs, ACI can also enable selective redirection of traffic to services devices, even if services devices are outside the normal forwarding path.

This chapter does not go into traditional methods of attaching firewalls and load balancers to networks. Rather, it provides implementation guidance on how some of the most commonly deployed traditional services attachment designs can be translated into service graphs.

## **“Do I Know This Already?” Quiz**

The “Do I Know This Already?” quiz allows you to assess whether you should read this entire chapter thoroughly or jump to the “Exam Preparation Tasks” section. If you are in doubt about your answers to these questions or your own assessment of your knowledge of the topics, read the entire chapter. [Table 12-1](#) lists the major headings in this chapter and their corresponding “Do I Know This Already?” quiz questions. You can find the answers in [Appendix A, “Answers to the ‘Do I Know This Already?’ Questions.”](#)

**Table 12-1** “Do I Know This Already?” Section-to-Question Mapping

Foundation Topics Section	Questions

Service Graph Fundamentals	1-5
Service Graph Implementation Workflow	6-8
Service Graph Implementation Examples	9, 10

## Caution

The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark that question as wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

- 1.** Which service insertion model enables ACI to manage the entire configuration of an L4-L7 services device?
  - a.** Manual service insertion
  - b.** Network policy mode
  - c.** Service manager mode
  - d.** Service policy mode
  
- 2.** True or false: A device package should be used only when configuring L4-L7 devices in service policy mode.
  - a.** True
  - b.** False

- 3.** True or false: When deploying a service graph without PBR, ACI effectively steers the desired traffic to the L4-L7 devices, even if they are outside the normal routing and switching path for interesting traffic.
- a.** True  
**b.** False
- 4.** What is the name for the process APICs go through to translate user intentions expressed in service graph configurations into a path through the services devices?
- a.** Rendering  
**b.** Connector configuration  
**c.** Template instantiation  
**d.** Device selection policies
- 5.** When using non-PBR service graphs to deploy a transparent firewall that bridges traffic between endpoints in a single subnet, which deployment mode should be used, and how many bridge domains are needed?
- a.** One-arm mode and one bridge domain  
**b.** GoTo mode and two bridge domains  
**c.** GoThrough mode and two bridge domains  
**d.** GoThrough mode and one bridge domain
- 6.** Which policy associates one or more L4-L7 devices to a service graph template and contract?
- a.** Device package  
**b.** Function profile  
**c.** L4-L7 service graph policy  
**d.** Device selection policy

- 7.** In a service policy mode environment that has granular role-based access control implemented, how can a services administrator dictate that only specified configuration settings be deployed to services appliances?
- a.** Create a function profile and mark parameters of interest with the mandatory attribute.
  - b.** There is no need. L4-L7 devices are not managed by ACI in service policy mode.
  - c.** Prevent any users who may want to change firewall configurations from ACI from having access to the fabric in the first place.
  - d.** Create a function profile, set values for the parameters that should not change, and set the Locked attribute to True.
- 8.** True or false: A service graph template determines which EPGs communicate indirectly through L4-L7 services devices.
- a.** True
  - b.** False
- 9.** True or false: Selecting Virtual as the L4-L7 services device attachment method necessitates VMM integration.
- a.** True
  - b.** False
- 10.** Which of the following items refers to a series of appliance interfaces that require common interface configurations and map directly to physical or virtual interfaces?
- a.** Concrete interfaces
  - b.** Consumer connector interface

- c. Provider connector interface
- d. Cluster interfaces

## Foundation Topics

### Service Graph Fundamentals

Common Layer 4 through Layer 7 services include firewalls, load balancers, traffic inspection appliances, SSL offload functions, and application flow acceleration functions.

Traditionally, inserting these types of services required a highly complicated and manual process of VLAN or VRF stitching between network devices and services appliances. In addition to the fact that deployment of services may necessitate weeks of cross-team coordination and planning, it was difficult to scale services up or down based on load. And, once an application was retired, there was no automated way to remove associated service appliance configurations.

ACI addresses these types of issues by providing customers with a plethora of options in addition to traditional VRF and VLAN stitching methods to automate service insertion tasks based on their comfortability level and technical requirements. This service insertion automation is accomplished through deployment of service graphs.

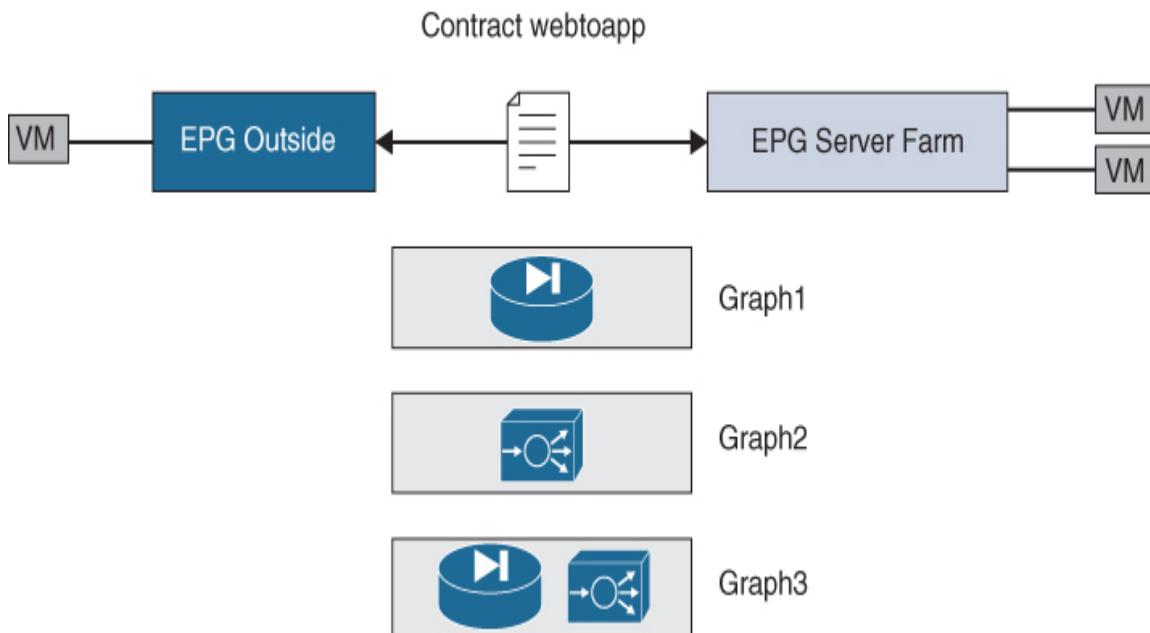
### Service Graphs as Concatenation of Functions

A **service graph** is a variation on the concept of a contract. In ACI, a contract whitelists communication between two endpoint groups (EPGs). This same whitelisting concept can

also be extended to functions such as traffic filtering, traffic load balancing, and SSL offloading. When a contract has been associated with a service graph policy, ACI locates the devices that provide the functions defined by the policy and inserts them into the path. A service graph is, therefore, a concatenation of functions (and not of network devices).



The significance of using contracts as a tool for implementing service functions is not limited to just inserting services devices into the path. Each subject within a contract allows association of a different service graph to matched traffic. [Figure 12-1](#) shows how the allocation of different service graphs to different subjects within a contract can enable very granular control over the flow of traffic through L4-L7 services functions.



**Figure 12-1** *Associating Different Service Graphs to Different Subjects Within a Contract*

Note that not all service graphs are created equally. The phrase “inserting services devices into the path” can be interpreted in two different ways from a traffic flow perspective:

- **Service graph with policy-based redirect (PBR):** The service graph effectively steers traffic to the L4-L7 device, redirecting traffic to the proper services devices even if they are outside the regular routing and switching path between the source and destination EPGs.
- **Service graph without PBR:** The service graph does not steer traffic to the L4-L7 device, but it does dictate that traffic flowing between the source and destination EPGs must pass through the functions identified by the service graph or be dropped. Only traffic that flows through the specified functions and is permitted by intermediate appliances is allowed to reach the intended destination EPG.

Effectively, service graphs without PBR necessitate that normal routing and switching rules steer traffic to services devices. That said, the enforcement of different functions between two EPGs through diversification of contract subjects, as indicated in [Figure 12-1](#), typically requires some form of policy-based routing.

Because PBR is outside the scope of the DCACI 300-620 exam, this chapter focuses on the more basic service graph deployments.

## Service Graph Management Models

In a large company, a network team typically manages the network, a security team manages the firewalls, and application delivery engineers manage load balancers.

While this trend still exists, the desire for agility tends to push IT teams to become more aggregated. For ACI to align with these trends and to begin to bring teams together, Cisco has enabled integrations that use different operational models within the data center.

ACI has three management models for deploying service graphs:



- ***Network policy mode (unmanaged mode)***: This operational model aligns with the traditional approach of network engineers configuring network connectivity to L4-L7 devices without implementing any configurations on the L4-L7 devices themselves.
- ***Service policy mode (managed mode)***: In this operational model, the APICs handle not just the configuration of network connectivity to the L4-L7 devices but also the configuration of the L4-L7 devices. This approach is geared toward end-to-end infrastructure automation.
- ***Service manager mode (hybrid mode)***: With this operational model, the firewall and load balancer administrators define L4-L7 policies using traditional L4-L7 management tools. Network administrators apply a limited set of L4-L7 policies to service graphs. These policies are often limited to interface-level configurations. ACI automates networking to the L4-L7 devices and pushes only basic configurations to the L4-L7 devices.

## Note

ACI can trigger the instantiation of L4-L7 services devices using a feature called *service VM orchestration*. However, it is not very common for ACI to be the platform used to orchestrate the deployment of new services appliances. For ACI to integrate with an L4-L7 device via managed or hybrid mode and be able to push configurations, the services device needs to be bootstrapped with a minimum level of configuration that includes a management interface IP address and default gateway as well as enablement of any programmatic interface that allows the APIC to configure the appliance.

## Understanding Network Policy Mode

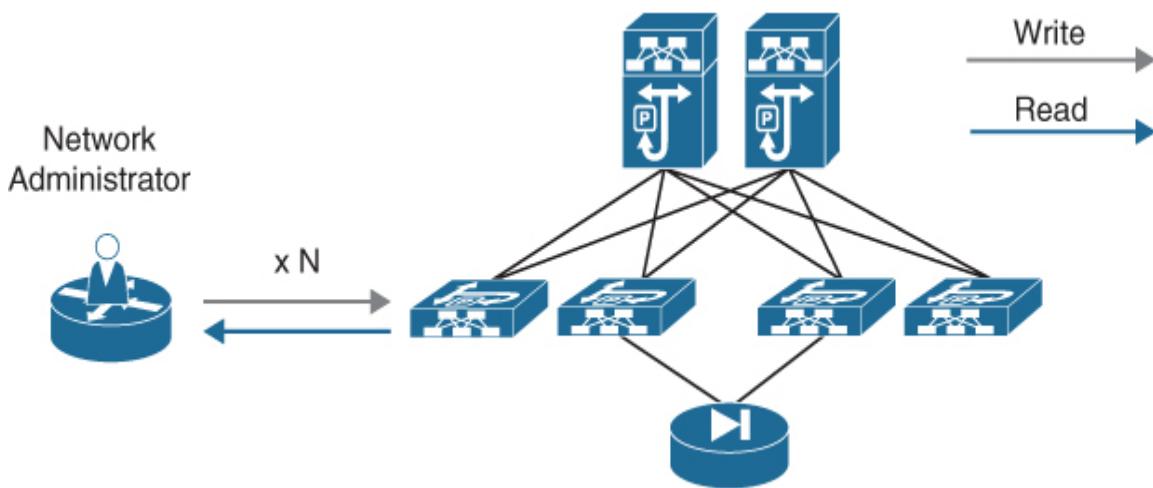
When deploying service graphs in network policy mode, ACI automates the configuration of network connectivity to L4-L7 services appliances. This involves ACI potentially modifying bridge domain settings to bring BDs into compatibility with the associated service graph, creating shadow EPGs, mapping shadow EPGs to physical switch ports, whitelisting communication between shadow EPGs and the pertinent consumer or provider EPG(s), and finally associating the service graph policy with the relevant contract(s). In the event that the L4-L7 device is a virtual machine associated with an ACI-generated VDS, ACI also pushes the shadow EPGs into vCenter as VDS port groups and associates them with services appliance virtual network adapters.



### Note

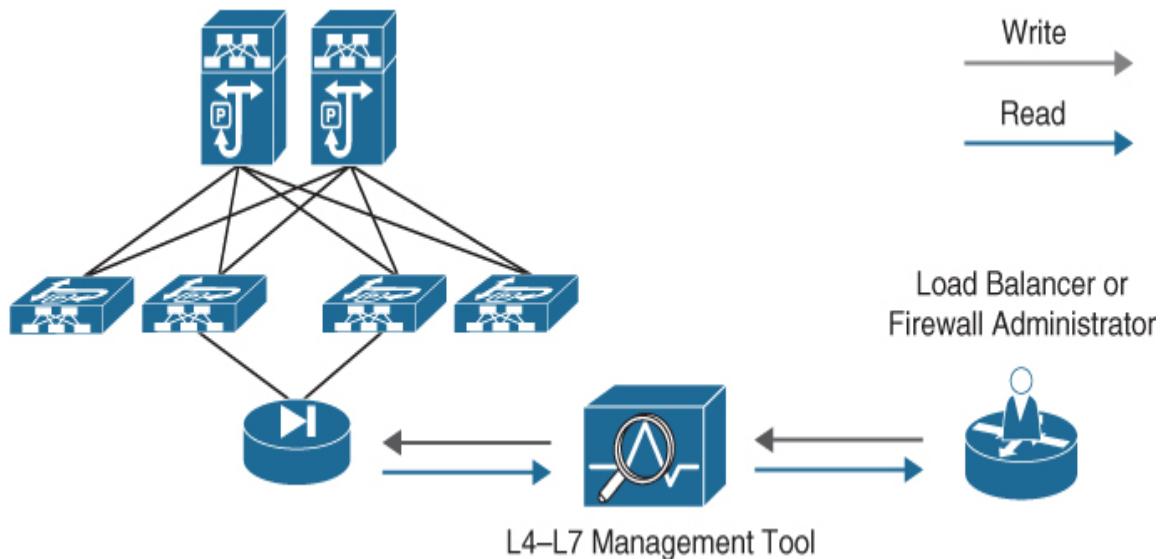
A **shadow EPG** is any form of EPG automatically generated by ACI, typically for the purpose of automating enforcement of contracts between two components. To mandate that traffic flow from a consumer EPG through a services appliance, ACI creates a shadow EPG and assigns it to the consumer-side interface of the L4-L7 services device. A contract is then applied between this shadow EPG and the consumer EPG, and ACI needs to create another shadow EPG to associate with the egress interface of the appliance. These shadow EPGs are effectively placeholders for the application of additional contracts to enforce the desired traffic flow. Service graphs are *not* the only use case for shadow objects in ACI.

To better understand network policy mode, take a look at [Figure 12-2](#), which provides a reminder of the central tenet of network policy mode that network administrators are responsible for configuring the network but not services appliances.



**Figure 12-2 Network Admins Only Configure the Network in Network Policy Mode**

Meanwhile, security and load balancer administrators continue to configure firewalls and load balancers either directly or by using the L4-L7 management tools they have always used, as shown in [Figure 12-3](#).



**Figure 12-3** Services Admins Experience No Change with Network Policy Mode

Network policy mode service graphs are ideal when there is no need for firewalls to be decommissioned and provisioned again periodically (for example, with perimeter firewalls controlling access to the data center), there is no desire to change how teams manage data center infrastructure, or there is no support from an L4-L7 services device vendor for deployment of the device in service policy mode or service manager mode.

## Understanding Service Policy Mode

In environments in which there is a desire for deep integration between all infrastructure components in the data center and IT expects all configurations to be scripted, the APICs can be used as the central point of automating

not just the network but also L4-L7 services device configurations. Full automation of L4-L7 services device configuration through ACI requires that services devices be identified to ACI in service policy mode.

Before ACI can configure a load balancer or firewall, an administrator needs to upload or import a device package for the appliance to the APICs. This is what enables ACI to speak to its APIs.

A *device package* is a zip file that an ACI administrator obtains from an L4-L7 services appliance vendor and imports into the APICs. It includes a description of the device, the functions it performs, and a laundry list of parameters exposed to the APIC. When a device package is uploaded/imported to the APICs, ACI gains a full understanding of what a services device can do, how it connects to the fabric, how to deliver traffic to the device, how to receive return traffic from it, and how to translate user intent to the device.



The Cisco Adaptive Security Appliance (ASA) is an example of a L4-L7 services device that can run in service policy mode and that supports feature-rich device packages.

Device packages designed for service policy mode often expose a large amount of information to ACI, sometimes allowing the L4-L7 device to be configured almost entirely from the APIC.

The main mechanism L4-L7 services device administrators use to define configuration parameters in ACI to feed back to services devices is **function profiles**. Configurations defined in function profiles may be amended at the time of

service graph deployment. When a vendor creates a device package, it defines a number of function profiles corresponding with the various functions provided by the services device. The vendor populates default values for some parameters within function profiles. Administrators can then modify the bulk of the default values to suit their needs.



[Figure 12-4](#) shows the creation of a custom function profile for a device that will operate in service policy mode. Notice that the range of configurations that ACI can deploy to the type of firewall for which this function profile is being created includes not only interface configurations but also access lists and NAT.

[Figure 12-5](#) illustrates the change in the data center management model achieved with service policy mode. Notice that ACI function profiles are the primary avenue for L4-L7 services administrators to dictate policies in service policy mode. Network administrators typically leverage function profiles preconfigured by L4-L7 services administrators when deploying service graphs. With this model, role-based access control is often implemented to enable L4-L7 services administrators to directly log in to APICs and configure function profiles and perhaps service graph templates. Monitoring of L4-L7 services appliances continues to take place within the L4-L7 management tool or services devices themselves. The integration between ACI and L4-L7 services devices also ensures that ACI is able to query services devices for health and to ensure that configurations have not been modified. In service policy mode, ACI may take corrective action and update L4-L7

device configurations if configurations are modified locally on L4-L7 services devices.



Create L4-L7 Services Function Profile

Name: Gaming@U-Function-Profiles  
Description: optional  
Profile Group: ASA-in-Routed-Mode  
Copy Existing Profile Parameters:   
Profile: CISCO-ASA-1.3/WebPolicyForRoutedMode

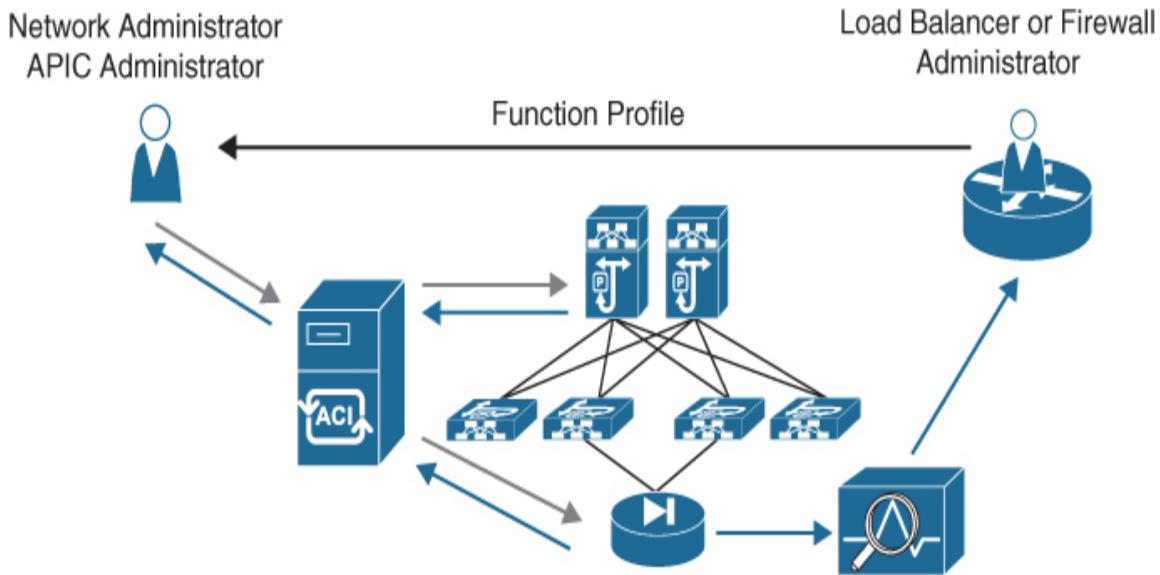
Features and Parameters

Note: In order to automatically apply new values to the parameters of an existing graph instance when users modify function profiles, the name of the top folder must end with "-Default".

Features	Basic Parameters	All Parameters																																																
Interfaces																																																		
AccessLists		<table border="1"><thead><tr><th>Folder/Parameter</th><th>Name</th><th>Hint</th><th>Path from Schema</th><th>Value</th><th>Mandatory</th><th>Locked</th><th>Shared</th></tr></thead><tbody><tr><td>Device Config</td><td>Device</td><td></td><td></td><td></td><td>false</td><td>false</td><td></td></tr><tr><td>Access List</td><td>access-list-inbound</td><td></td><td></td><td></td><td>false</td><td></td><td></td></tr><tr><td></td><td>Access Control Entry</td><td>permit-http</td><td></td><td></td><td>false</td><td></td><td></td></tr><tr><td></td><td>Access Control Entry</td><td>permit-https</td><td></td><td></td><td>false</td><td></td><td></td></tr><tr><td></td><td>Access Control Entry</td><td>permit-icmp</td><td></td><td></td><td>false</td><td></td><td></td></tr></tbody></table>	Folder/Parameter	Name	Hint	Path from Schema	Value	Mandatory	Locked	Shared	Device Config	Device				false	false		Access List	access-list-inbound				false				Access Control Entry	permit-http			false				Access Control Entry	permit-https			false				Access Control Entry	permit-icmp			false		
Folder/Parameter	Name	Hint	Path from Schema	Value	Mandatory	Locked	Shared																																											
Device Config	Device				false	false																																												
Access List	access-list-inbound				false																																													
	Access Control Entry	permit-http			false																																													
	Access Control Entry	permit-https			false																																													
	Access Control Entry	permit-icmp			false																																													
NAT																																																		
TrafficSelectionObject																																																		
All																																																		

Cancel Submit

**Figure 12-4** Function Profile for Configuration of L4-L7 Devices in Service Policy Mode



**Figure 12-5** Service Policy Mode Management Model in Review

One difference between service policy mode and traditional service insertion methods is that this mode forces IT teams to configure security and load-balancing rules at the same time the network configurations for the service graph are applied.

A differentiator for this management model is that it automates and manages the lifecycle of security rules, load-balancing rules, and network connectivity and aligns them with one another. When a service graph is deleted, ACI automatically removes relevant configurations within ACI and on associated L4-L7 services devices.

At the same time, a negative aspect of service policy mode is that it creates a dependency between the ACI code versions, L4-L7 services device software versions, and associated device packages.

## Understanding Service Manager Mode

One positive aspect of network policy mode is that it is easy to deploy and has very few associated caveats.

Furthermore, configuration changes at the firewall or load balancer level very seldom bring about the need for a service graph redeployment when network policy mode is used.

Compared to network policy mode, service policy mode is more difficult to operationalize and support unless careful consideration is given to ensuring optimal object reuse.

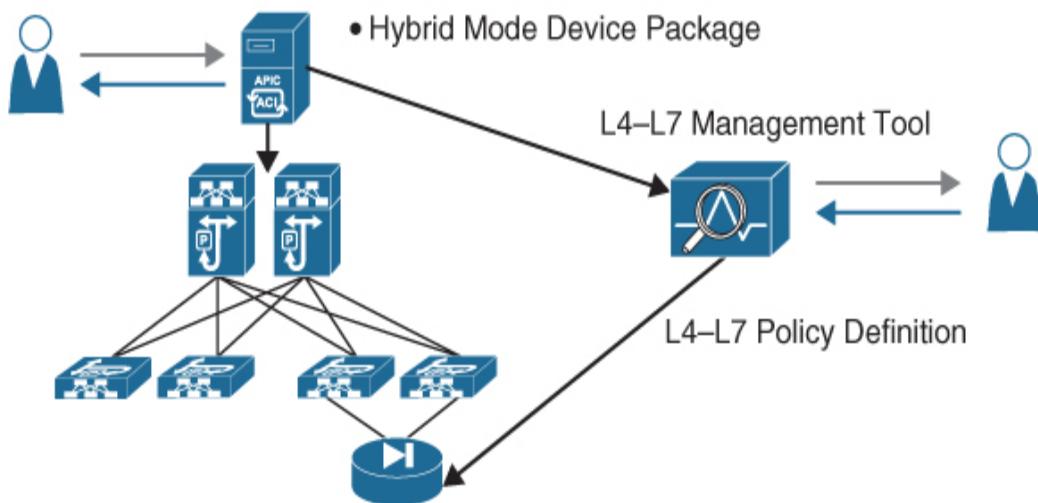
Because service policy mode ties firewall configuration to service graphs, it is common to need to redeploy service graphs for firewall changes to take effect. Also, the sheer number of parameters exposed to APICs through device packages makes the APIC GUI less user-friendly than native firewall and load balancer management applications for implementing configuration changes on L4-L7 services devices.

Service manager mode sits at the intersection between network policy mode and service policy mode and is often acknowledged as the best of both worlds. The idea behind service manager mode is that the APICs should be able to automate network policy and stitch traffic all the way to the L4-L7 services devices but also have the ability to automate certain basic configurations on the L4-L7 devices themselves. The most important facet of this hybrid approach is that function profiles within ACI do not serve as the primary tool for L4-L7 services device configuration. Instead, services administrators deploy the bulk of L4-L7 configurations independently of ACI in the L4-L7 management tool produced by the associated services device vendors.

[Figure 12-6](#) illustrates the management model achieved via service manager mode. Firewall and load balancer administrators define L4-L7 policies within the L4-L7

management tool supported by the L4-L7 device vendor. Example L4-L7 management solutions that support service manager mode include Cisco Firepower Management Center (FMC) and Citrix NetScaler Management and Analytics System (MAS). The service graph deployment process in this case merely references the names of policies created within the L4-L7 management tool to ensure that appliance interfaces as well as ACI endpoints get assigned to the correct policies within the L4-L7 services management tool.

Network Configurations + Reference to L4-L7 Policy



**Figure 12-6** Service Manager Mode Management Model

Even though the diagram omits any reference of function profiles, it is important to understand that service manager mode uses function profiles. However, fewer parameters can actually be configured in ACI through function profiles. [Figure 12-7](#) shows a function profile for FMC-integrated Firepower Threat Defense firewalls that will be deployed in service manager mode. Notice that most of the configuration parameters that the device package allows ACI to deploy to FMC are limited to the configuration of interfaces and the association of pre-created policies to

these interfaces. Now, contrast this with [Figure 12-4](#), shown earlier in the chapter.

Feature	Basic Parameters	All Parameters																																																																								
Interfaces																																																																										
All		<table border="1"><thead><tr><th>Folder/Parameter</th><th>Name</th><th>Value</th><th>Mandatory</th><th>Locked</th><th>Shared</th></tr></thead><tbody><tr><td>Device Config</td><td>Device</td><td></td><td>false</td><td>true</td><td></td></tr><tr><td>Access Policy</td><td>Transparent-FWs</td><td></td><td>false</td><td></td><td></td></tr><tr><td>Access Rules</td><td>Gamer-VMs-to-Backend-Servers</td><td></td><td>false</td><td></td><td></td></tr><tr><td>Destination Interface</td><td>AccDestinationZones</td><td></td><td>false</td><td></td><td></td></tr><tr><td>Source Interface</td><td>AccSourceZones</td><td></td><td>false</td><td></td><td></td></tr><tr><td>Bi-Directional</td><td>Bi-Directional</td><td>true</td><td>false</td><td>false</td><td></td></tr><tr><td>Access Rules</td><td>TEST</td><td></td><td>false</td><td></td><td></td></tr><tr><td>Bridge Group Interface</td><td>BVNI</td><td></td><td>false</td><td>false</td><td></td></tr><tr><td>IPv4 Address Configuration</td><td>IPv4Config</td><td></td><td>false</td><td></td><td></td></tr><tr><td>Use Static IP</td><td>static</td><td></td><td>false</td><td></td><td></td></tr><tr><td>IP Address</td><td>address</td><td>10.68.1.100/24</td><td>true</td><td>false</td><td></td></tr></tbody></table>	Folder/Parameter	Name	Value	Mandatory	Locked	Shared	Device Config	Device		false	true		Access Policy	Transparent-FWs		false			Access Rules	Gamer-VMs-to-Backend-Servers		false			Destination Interface	AccDestinationZones		false			Source Interface	AccSourceZones		false			Bi-Directional	Bi-Directional	true	false	false		Access Rules	TEST		false			Bridge Group Interface	BVNI		false	false		IPv4 Address Configuration	IPv4Config		false			Use Static IP	static		false			IP Address	address	10.68.1.100/24	true	false	
Folder/Parameter	Name	Value	Mandatory	Locked	Shared																																																																					
Device Config	Device		false	true																																																																						
Access Policy	Transparent-FWs		false																																																																							
Access Rules	Gamer-VMs-to-Backend-Servers		false																																																																							
Destination Interface	AccDestinationZones		false																																																																							
Source Interface	AccSourceZones		false																																																																							
Bi-Directional	Bi-Directional	true	false	false																																																																						
Access Rules	TEST		false																																																																							
Bridge Group Interface	BVNI		false	false																																																																						
IPv4 Address Configuration	IPv4Config		false																																																																							
Use Static IP	static		false																																																																							
IP Address	address	10.68.1.100/24	true	false																																																																						

**Figure 12-7** Sample Function Profile for Device in Service Manager Mode

There are three minor implementation differences between service policy mode and service manager mode:

- The APIC administrator needs to import a special device package that supports service manager mode.
- The APIC administrator needs to define a **device manager**, which calls out the IP address, communication protocol, and credentials for APIC access to the L4-L7 management solution.
- When the APIC administrator identifies L4-L7 devices to the fabric, the associated device manager configuration needs to be referenced.

**Note**

Note that the implementation differences outlined are generalizations rather than firm rules.

[Figure 12-8](#) shows a sample configuration of a device manager in ACI. Note that a management EPG needs to be defined only if ACI in-band management is used to access the L4-L7 management application. Device managers can be defined under **Services > L4-L7 > Device Managers** within each tenant.

The screenshot shows the 'Create Device Manager' dialog box. At the top, there are fields for 'Device Manager Name' (set to 'FMC'), 'Management EPG' (set to 'DC2-EPG-VLAN264-ESXIMGMT'), and 'Device Manager Type' (set to 'CISCO-FTDmgr\_FI=1.0'). Below these, the 'Management' section lists a single entry: 'Host' 10.233.64.150 and 'Port' 443. There are 'Delete' and 'Add' buttons next to the list. At the bottom, there are fields for 'Username' (set to 'admin'), 'Password', and 'Confirm Password'. At the very bottom are 'Cancel' and 'Submit' buttons.

**Figure 12-8** Configuration of a Device Manager Within a Tenant

# When to Use Service Graphs

Few concepts in ACI boggle heads and roll eyes like service graphs do. Opponents of service graphs sometimes like to point to the complexities involved with service policy mode and argue that the caveats associated with service graphs far outweigh the benefits. This argument misses the central point that this management model is most useful for creating elastic environments with rapid scale-out capabilities.

Furthermore, unmanaged mode provides benefits over manual service insertion, which makes it very useful in most environments. One such benefit is use of PBR.

Service graphs were never intended to be used in all ACI deployments. Service graphs should be deployed when there are tangible benefits to their use when compared to manual service insertion.

Service graphs offer the following advantages, among others.

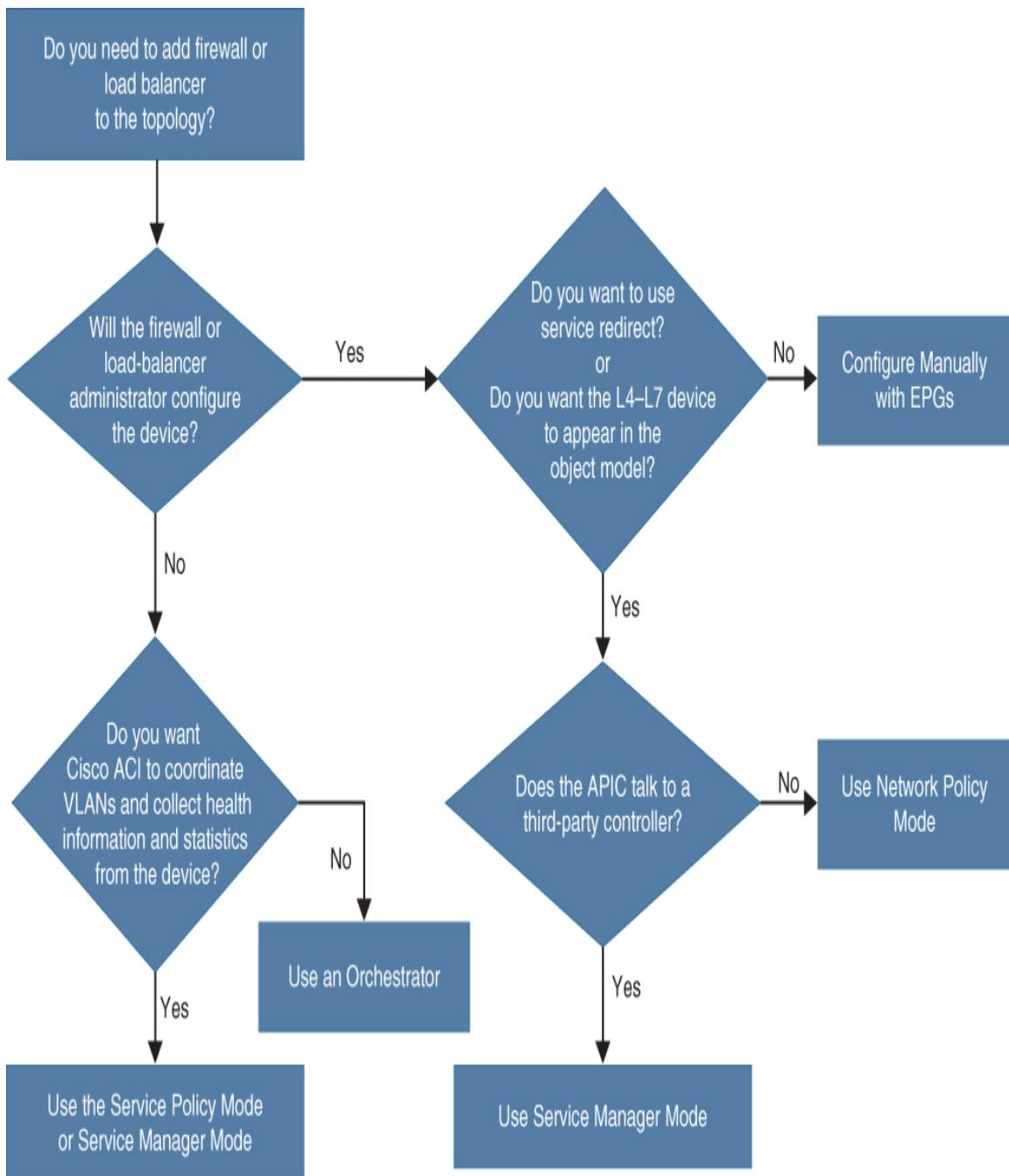
- A service graph with PBR can redirect traffic to L4-L7 devices, eliminating the need for more complex designs.
- Service graphs automatically manage VLAN assignments.
- Service graphs automatically connect virtual network interface cards (vNICs).
- Associated configuration templates can be reused multiple times.
- A service graph provides a logical view of the network and offers an application-related view of services.

- A service graph provides a good model for sharing a device across multiple departments.
- A service graph collects health scores from a device or service.
- A service graph collects statistics from the services devices.
- Service graphs can update firewall ACLs and load balancer server pools automatically using endpoint discovery.

It is only when an IT team sees tangible benefits in service graphs for a given environment that service graphs should be embraced over manual insertion methods. Otherwise, there is a risk that IT may be introducing capabilities that operations teams do not comprehend or that they are reluctant to support.

## **Choosing an L4-L7 Services Integration Method**

Cisco has created the flowchart shown in [Figure 12-9](#) to provide very high-level guidance in deciding how to integrate L4-L7 services, given very general requirements.



**Figure 12-9** Service Graph Decision Flowchart

## Understanding Deployment Modes and the Number of BDs Required

ACI supports the following device deployment modes for non-PBR L4-L7 devices with a service graph:



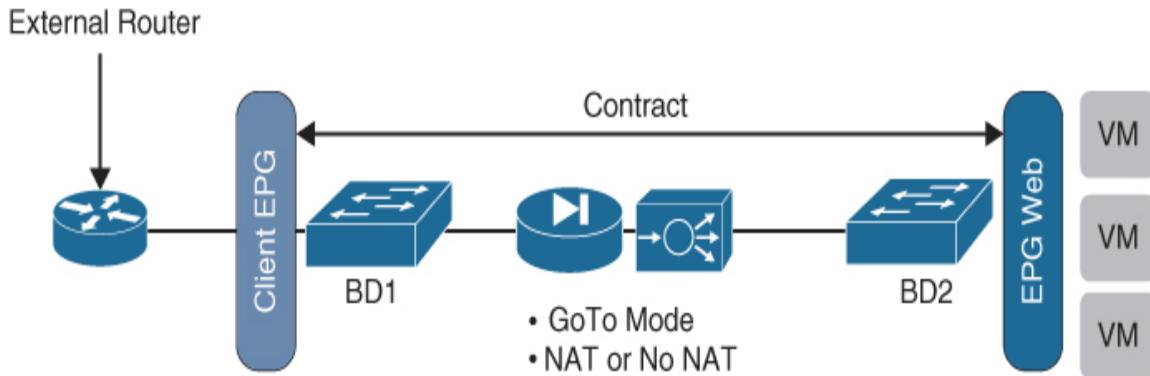
- **GoTo:** In this mode, the default gateway for servers is the L4-L7 device. This mode requires use of separate bridge domains for consumer-side and provider-side interfaces unless PBR is used. This mode is also known as *routed* mode.
- **GoThrough:** This mode requires use of two separate bridge domains. In this mode, the default gateway for the servers is the consumer-side bridge domain. The provider-side bridge domain should *not* be configured to perform routing. The L4-L7 device bridges the consumer-side bridge domain and the provider-side bridge domain. This mode is also known as *transparent* mode, or bridged mode.
- **One-arm:** In this mode, the default gateway for any servers is the server-side bridge domain. The load balancer connects to the fabric using a single bridge domain that serves as the default gateway for the load balancer itself. The services device is inserted into the topology using source NAT (SNAT), which ensures receipt of the return traffic.

## Deploying Service Graphs for Devices in GoTo Mode

Three designs are valid and can be used for deployment of non-PBR service graphs with devices in routed mode:

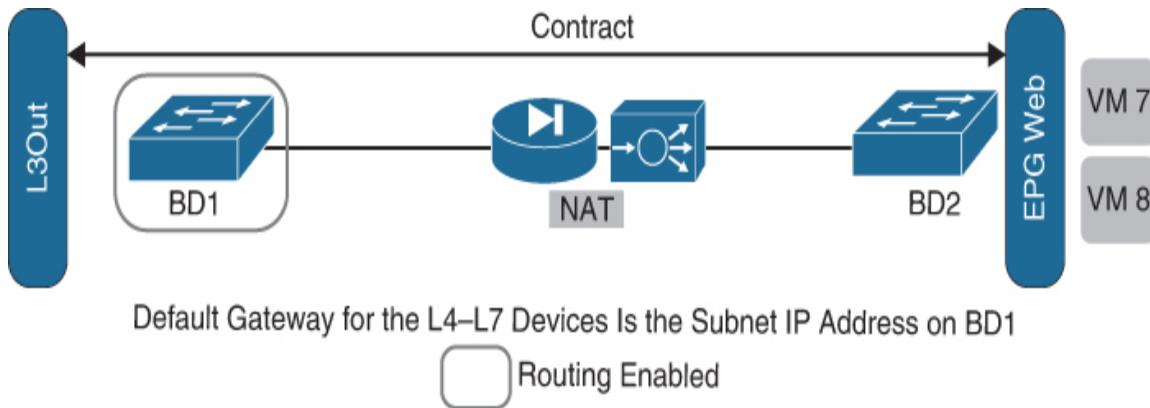
- **Routed mode with outside Layer 2 bridge domain:** In this implementation, ACI may function as a Layer 2 transit for endpoints behind an L4-L7 services device. In this design, it is the job of an external router to direct traffic to the services device. This is a common design for companies that need to move a firewall containing DMZ subnets into ACI but that do not want to rearchitect their DMZ environments until they gain more familiarity with ACI. From an ACI perspective, none of the bridge domains connecting to the L4-L7 services devices or extending the firewall traffic outside the fabric have routing enabled.
- **Routed mode with L3Out and NAT:** In this case, a services device external interface, for instance, may connect to ACI using a bridge domain that has a subnet defined with routing enabled. The services device internal interface connects to another bridge domain. If routing were to be enabled on this internal bridge domain without NAT, there would be no reason for the fabric to send traffic to the L4-L7 device; routing to the destination could happen directly. With NAT enabled, client machines need to go to the services device. Because NAT ranges fall into the external bridge domain subnet range, it is possible to advertise the NAT range outside the fabric over an L3Out using regular bridge domain advertisements.
- **Routed mode with route peering:** This basically involves ACI learning routes behind an L4-L7 services device through an L3Out. (This option is discussed further later in this chapter.)

[Figure 12-10](#) illustrates use of ACI as a Layer 2 network with an external router directing traffic to the L4-L7 services device(s).



**Figure 12-10** *Routed Mode with an Outside Layer 2 Bridge Domain*

[Figure 12-11](#) illustrates the use of NAT and a routable external bridge domain to direct traffic to the services device(s).



**Figure 12-11** *Routed Mode with L3Out and NAT*

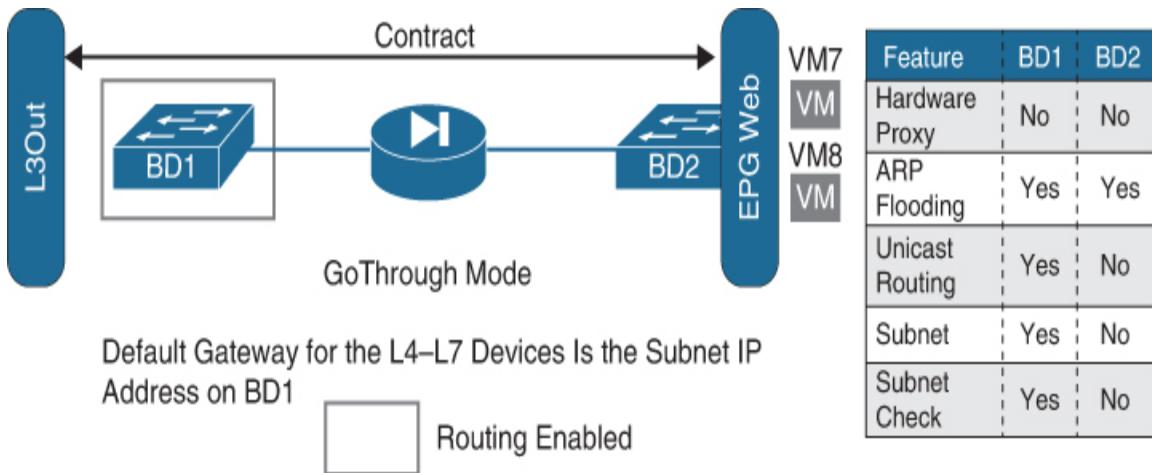
## Deploying Service Graphs for Devices in GoThrough Mode

Two common non-PBR implementations allow transparent devices to leverage service graphs:

**Key Topic**

- **Transparent mode with an outside Layer 2 bridge domain:** In this design, ACI functions as a Layer 2 transit because both the outside and inside of the service graph connect to Layer 2 bridge domains. It is the job of an external routing device to direct traffic to the services device.
- **Transparent mode with L3Out:** In this design, the outside bridge domain of a service graph connects to the outside network through routing provided by the Cisco ACI fabric.

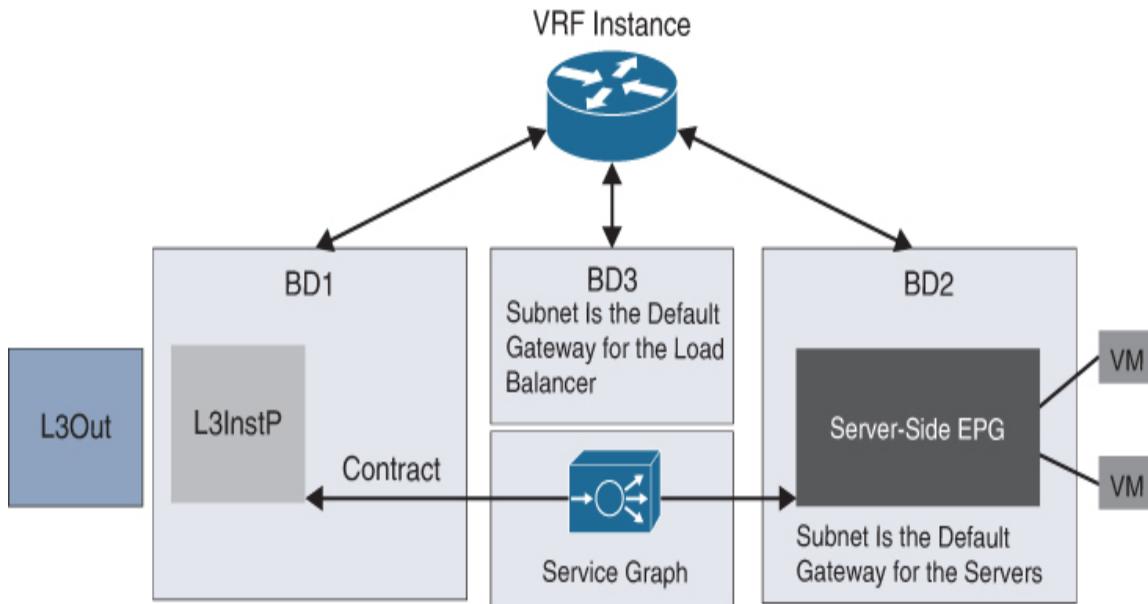
Use of ACI as a Layer 2 transit is typically straightforward. [Figure 12-12](#) shows a transparent mode deployment with an L3Out and required bridge domain settings.



**Figure 12-12** *Transparent Mode Design with L3Out*

## Deploying Service Graphs for One-Arm Load Balancers

[Figure 12-13](#) illustrates a one-arm load balancer deployment using SNAT. This design leverages one bridge domain for ACI to connect to the load balancer, and the server and client sides of the communication are likely to each be in bridge domains of their own.



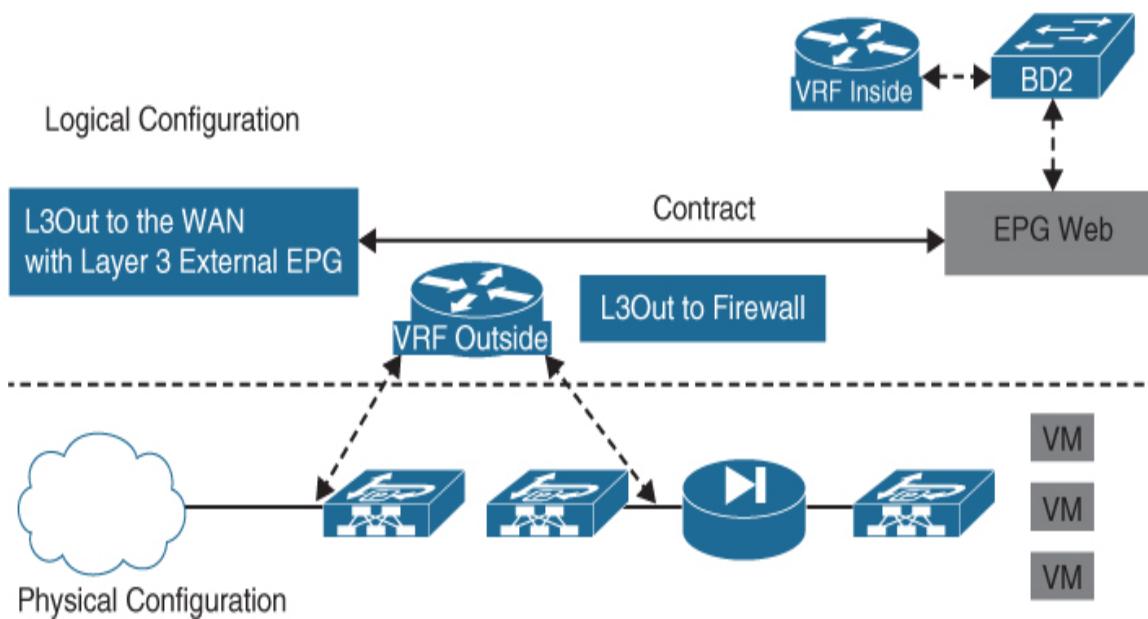
**Figure 12-13** Load Balancer Deployed in One-Arm Mode in ACI

## Understanding Route Peering

Not all of the designs mentioned in the previous sections allow use of ACI anycast default gateways for east-west traffic optimization. From a contract application perspective, the majority of design options covered are less flexible than alternate design options that move the default gateways for all subnets into ACI and attempt to eliminate the need for NAT.

[Figure 12-14](#) shows a VRF sandwich design. This design places each interface of an L4-L7 services device and therefore each associated bridge domain in a different VRF. Because of this, there is no potential for endpoints behind

different subnets to communicate with one another directly in the absence of transit routing. ACI then establishes a separate L3Out with the L4-L7 services device in each VRF. This design is highly recommended when PBR is not an option because it forces traffic through services devices due to the existence of different IP spaces in each VRF. It also enables administrators to move server default gateways in all VRFs into ACI to fully leverage ACI anycast default gateways. Either static routing or dynamic routing can be used to advertise any subnets behind the services device.



**Figure 12-14** Route Peering Through Deployment of an L3Out and Interface in Each VRF

### Note

It is good to understand this design, but is not very likely to be a DCACI 300-620 exam topic because most route peering use cases fall into the gray area of transit routing.

# **Understanding Dynamic Endpoint Attach**

One popular capability with service graphs and a good case in point for harmonious cross-platform integrations within ACI is the *dynamic endpoint attach* feature.

Let's take a look at the benefits of this feature for application delivery controllers as an example. An application delivery engineer defines a virtual server using an IP address and port. The virtual server is typically a frontend IP address provided via DNS to client machines that load balances traffic across multiple backend servers. When integrated with ACI through service manager mode or service policy mode, an EPG can be created to identify the backend servers that correspond with the virtual server. Through this type of integration, ACI is able to communicate endpoint attachment and detachment events to load balancers, thereby dynamically increasing or decreasing the size of the backend server pool without the need for manual configuration or scripting.

The beauty of this feature and this type of integration using service manager mode is that it does not require a drastic management model change but provides many benefits that reduce the need for manual changes in the network.

The dynamic endpoint attach feature can also be used to communicate EPG names and update EPG members on firewalls dynamically to help security administrators in the creation of firewall rules based on EPGs as objects.

# **Understanding Bridge Domain Settings for Service Graphs**

When guiding traffic to or through services devices without PBR, it is often necessary to modify the following settings to accommodate various designs:

- L2 Unknown Unicast
- ARP Flooding
- Unicast Routing
- Subnets

When a device is deployed in routed mode, the goal is to place default gateways on L4-L7 services devices and not in ACI. Therefore, the following BD settings should be used (unless PBR is being used in conjunction with default gateways in the fabric):



- **L2 Unknown Unicast:** Flood
- **ARP Flooding:** Enabled
- **Unicast Routing:** Disabled
- **Subnets:** N/A

Generally speaking, it is highly recommended to use flooding and enable ARP for bridge domains that connect to L4-L7 services devices. Two reasons for this are as follows:

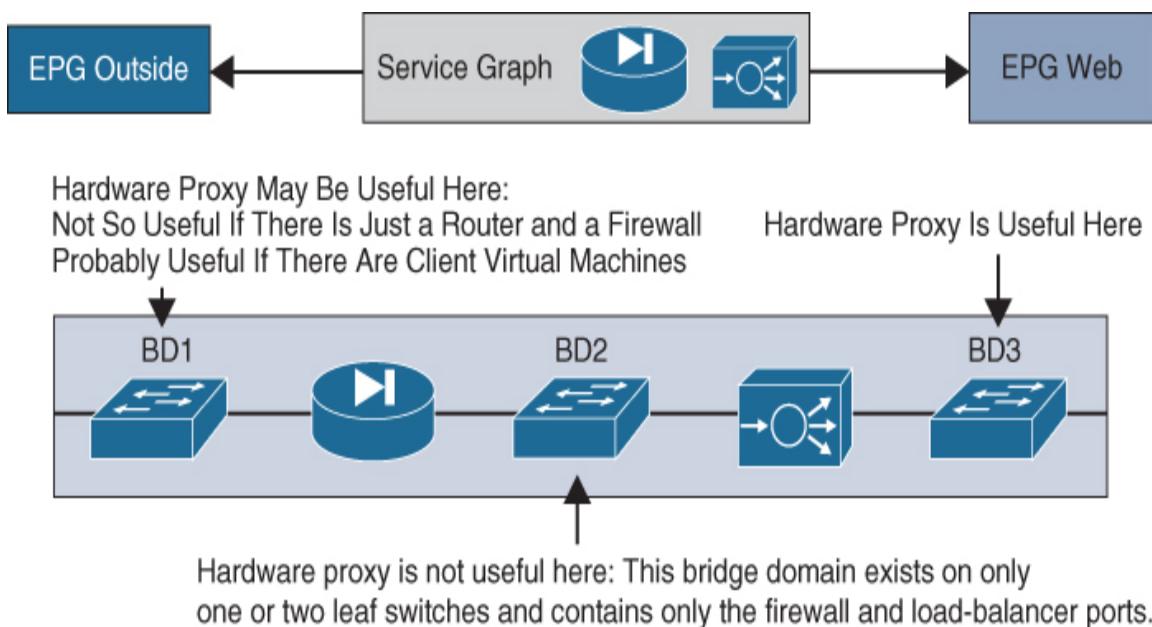
- Some L4-L7 devices in transparent (GoThrough) mode rely on flooding to build the forwarding tables, just like a transparent bridge does.
- When an L4-L7 device fails over, the IP address of that device may or may not change the MAC address as well. If it does change the MAC address, the Gratuitous

ARP (GARP) traffic generated by the L4-L7 device must reach the ARP cache of the adjacent devices. For this to happen, ARP flooding must be enabled.

To accommodate transparent firewall deployments, you should enable ARP flooding and set L2 Unknown Unicast to Flood. If these two settings are not pre-selected, ACI modifies bridge domain settings to conform with these requirements.



In spite of the general recommendations provided on bridge domain settings, there are instances when a bridge domain connecting a service graph can be set to Hardware Proxy to minimize flooding. [Figure 12-15](#) shows an example. Because GoThrough mode does not allow use of Hardware Proxy, we can assume that devices in this figure are in GoTo mode.



**Figure 12-15** Service Chain with Use Case for Hardware Proxy in BD3

This illustration depicts a multimode service chain. A **service chain** is merely a succession of functions. Here, traffic from Outside to Web flows through a firewall and then a load balancer. BD2 is an intermediary bridge domain that connects the two appliances. There is very little to optimize in terms of flooding in BD2 since the only devices in it are the two services devices. Assuming that BD1 connects the firewall to an L3Out, there is also likely to be little benefit in moving away from flooding in this bridge domain. However, if BD3 connects to servers, use of Hardware Proxy can help server performance by minimizing flooding.

## Understanding Service Graph Rendering

Once users define a service graph, the APICs translate the user intentions expressed in the service graph into a path through the services devices. This translation of intent is called *rendering*.



Depending on the service graph management model, the process of rendering can involve configuration of network policies or potentially even deployment of configurations on services appliances.

The end goal of the rendering process is to make sure that the only acceptable path between EPGs is the path defined in the service graph.

## Service Graph Implementation Workflow

The following high-level steps are required to implement service graphs:



- Step 1.** Import device packages (if L4-L7 services devices are deployed in managed or hybrid mode).
- Step 2.** Identify L4-L7 devices to the fabric.
- Step 3.** (Optional) Create custom function profiles.
- Step 4.** Configure a service graph template.
- Step 5.** (Optional) Configure device selection policies.
- Step 6.** Apply the service graph template.
- Step 7.** (Optional) Configure additional service graph parameters.
- Step 8.** Monitor the service graph and devices to confirm proper implementation.

The following sections describe these steps.

## Importing Device Packages

You import device packages into ACI by selecting **L4-L7 Services > Packages**. [Figure 12-16](#) shows the General tab for a sample device package.

The screenshot shows the Cisco Application Policy Infrastructure Controller (APIC) web interface. At the top, there's a navigation bar with links for System, Tenants, Fabric, Virtual Networking, L4-L7 Services (which is selected), Admin, Operations, Apps, and Integration. Below the navigation bar, there are two tabs: Inventory and Packages. The main content area is titled "L4-L7 Service Device Type - CISCO-ASA-1.3". It has a "General" tab selected, along with tabs for Operational, Faults, and History. On the left, there's a sidebar labeled "Packages" with a count of 1. The main panel displays device properties: Vendor (Cisco), Model (ASA), Capabilities (GoThrough, GoTo, L2), Major Version (1.3), Minor Version (12.3), Minimum Required Controller Version (1.1), APIC Controller Version (1.1), and Package Name (device\_script.py). At the bottom right are buttons for Show Usage, Reset, and Submit.

**Figure 12-16** General Information Page for a Device Package Imported into ACI

A device package contains the components described in Table 12-2.



**Table 12-2** Components of a Device Package

**Dev Description**  
ice  
Pac  
kag  
e  
Co  
mp

## one nt

Device XML file that defines the following:

ice  
spe  
cific  
atio  
n

Device properties:

- **Model:** Model of the device
- **Vendor:** Vendor of the device
- **Version:** Software version of the device

Functions provided by the device

	<ul style="list-style-type: none"> <li>▪ Interfaces and network connectivity information for each function</li> <li>▪ Device configuration parameters</li> <li>▪ Configuration parameters for each function</li> </ul>
Device script	A Python script that allows the APICs to interact with the device. The device script maps APIC events to function calls. A device package can contain multiple device scripts. A device script can interface with services devices via REST, SSH, or any similar mechanism.
Function Profile	An L4-L7 configuration template that includes configuration values for deployment to L4-L7 services devices. When a vendor creates a device package, it typically defines a number of function profiles corresponding with the various functions provided by the services device. The vendor populates default values for a number of parameters within each function profile. Administrators can modify most of the default values to suit their requirements.

Dev A configuration file that specifies parameters that are required by a device. This configuration can be shared by one or more service graphs.

I  
con  
figu  
rati  
on  
par  
am  
eter  
s

When an administrator imports a device package to the APICs, various subfolders appear under the installed device package. [Figure 12-17](#) shows the L4-L7 Services Function Profiles page for a sample device package. Function profiles may differ dramatically based on the profiles' purposes. For instance, a function profile used to configure a transparent firewall may include parameters to configure a BVI, while a function profile for a firewall in routed mode may require that IP addresses be assigned to each firewall interface configured.

A vendor may include multiple function profiles for a given use case with only minor differences in the parameters included. (For example, there are multiple routed mode function profiles in [Figure 12-17](#).) This ensures that services administrators have a wide range of templates to choose from when deciding which function profiles are best suited to an environment.

Name	Group Name	Associated Function	Description
WebPolicyForIRBMixedModeIPv4	WebServicePr...	Firewall	Allow any web traffic for ASA in IRB mixed mode ASA for IPv4. Routed mode for IPv6.
WebPolicyForRoutedMode	WebServicePr...	Firewall	Allow any web traffic for ASA in routed mode
WebPolicyForRoutedModeIPv4	WebServicePr...	Firewall	Allow any web traffic for ASA in routed mode for IPv4
WebPolicyForRoutedModeIPv4Cloud	WebServicePr...	Firewall	Allow any web traffic for ASA in routed mode for IPv4
WebPolicyForRoutedModeIPv4HA	WebServicePr...	Firewall	Allow any web traffic for ASA in routed mode for IPv4 and HA
WebPolicyForRoutedModeIPv46Cloud	WebServicePr...	Firewall	Allow any web traffic for ASA in routed mode for IPv4 and IPv6
WebPolicyForRoutedModeIPv4Cluster	WebServicePr...	Firewall	Allow any web traffic for ASA in routed mode for IPv4 in individual cluster
WebPolicyForRoutedModeIPv4OneArmed	WebServicePr...	Firewall	Allow any web traffic for ASA in routed mode for IPv4 of one-armed mode
WebPolicyForRoutedModeIPv6	WebServicePr...	Firewall	Allow any web traffic for ASA in routed mode for IPv6

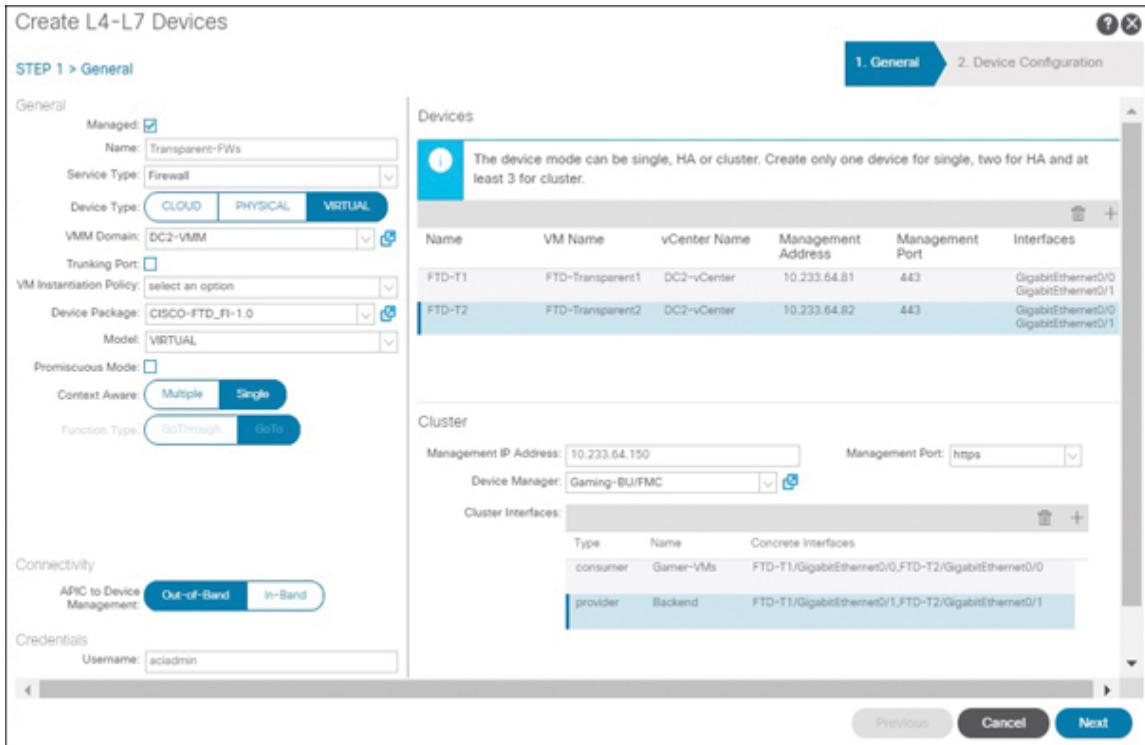
**Figure 12-17** Function Profiles Included in a Sample Device Package

## Identifying L4-L7 Devices to the Fabric

You configure L4-L7 services by opening the tenant where they will be used and navigating to **Services > L4-L7 > Create L4-L7 Devices**. Any services devices that need to be shared across multiple tenants need to be exported to the other tenants. This is true even for L4-L7 services devices defined in the common tenant.

[Figure 12-18](#) shows two L4-L7 devices being identified to ACI as a high-availability pair. A configuration that enables the Managed checkbox and has an L4-L7 service manager defined under the Device Manager pull-down usually points to service manager mode. On the other hand, a device

definition that enables only the Managed checkbox usually suggests a service policy mode deployment.



**Figure 12-18** Configuring a Firewall Pair to ACI in Service Manager Mode

An L4-L7 services appliance that will be used in a service graph template is called a **concrete device**. Services appliances are often deployed as a pair of active/active or active/standby devices. Concrete devices clustered together form a **logical device**. In Figure 12-18, the Devices section defines concrete devices, and the Cluster section maps concrete device interfaces to logical device interfaces.



Most of the configurations you have seen so far in the chapter are straightforward. Table 12-3 describes some of

the items that may require additional information.



**Table 12-3** Configuration Options in the Create L4-L7 Devices Page

Configuration Description
Service Type
ADC
Firewall
Other

Device Type	<p>Possible options available for this parameter include the following:</p> <ul style="list-style-type: none"> <li>■ <b>Physical:</b> Bare-metal servers, physical appliances, or non-VMM integrated virtual environments</li> <li>■ <b>Virtual:</b> A VM integrated using a VMM integration</li> </ul>
VM Instantiation Policy	<p>This configuration option is part of the service VM orchestration solution and enables APICs to trigger an instantiation of a VM template in vCenter.</p>
Context Aware	<p>This option specifies the context-awareness of the device, which can be one of the following:</p> <ul style="list-style-type: none"> <li>■</li> </ul>

	<p><b>Single:</b> The device cluster cannot be shared across multiple tenants of a given type that are hosted on the provider network. You must give the device cluster to a specific tenant for a given user.</p> <ul style="list-style-type: none"> <li>■ <b>Multiple:</b> The device cluster can be shared across multiple tenants of a given type that you are hosting on the provider network. For example, two hosting companies might share the same device.</li> </ul> <p>When defining a load balancer as a Layer 4 to Layer 7 services device, the Context Aware parameter is not used and can be ignored.</p>
Function Type	This option refers to deployment modes, including GoThrough and GoTo for routed firewalls and transparent firewalls, respectively.

## Creating Custom Function Profiles

With a function profile, a services administrator can create a collection of L4-L7 configuration templates that can be reused across service graphs.

To organize customized function profiles, administrators often group these policies for various purposes (for example, organizational hierarchy, purpose, environment).

To create a function grouping, select **Services > L4-L7**, right-click Function Profiles, and select Create L4-L7 Services Function Group.

To create a function profile, right-click the Function Profiles folder and select Create L4-L7 Services Function Profile to launch the page shown in [Figure 12-7](#), earlier in this chapter. Notice that [Figure 12-7](#) shows the checkbox Copy Existing Profile Parameters enabled, and a profile included in a device package is selected from the Profile drop-down box. This indicates that the administrator would like to leverage a built-in function profile from the device package to create a more custom function profile.

On the Create L4-L7 Services Function Profile page, there are three columns to the right:



- **Mandatory:** If a parameter has this attribute set to true, the configuration item is mandatory. This attribute offers a way for vendors or L4-L7 services administrators to ensure entry of a value for certain parameters as a prerequisite for deployment of the service graph. If a parameter has been set as mandatory in a device package, custom function profiles cannot override this mandatory setting. In [Figure 12-7](#), the L4-L7 services administrator mandates entry of an IP address to the BVI1 interface and enters a default value for this attribute.
- **Locked:** L4-L7 parameters used by a service graph can be stored under the provider EPG, bridge domain, application profile, or tenant. When a graph is instantiated, the APIC resolves the needed configuration for a service graph by looking up the

parameters in various places. If Locked has been set to true for a given parameter, parameter values set under the associated provider EPG, bridge domain, application profile, or tenant will be ignored when applying the service graph. If an administrator wants to ensure a specific value is used for a parameter within a function profile at all times, this option should be set to true.

- **Shared:** If this option is set to true, the parameter value in the function profile will be used unless a parameter value is set under a provider EPG, a bridge domain, an application profile, or a tenant. Therefore, setting the value in the Shared column for a parameter to true basically sets the value within the function profile as a modifiable default.

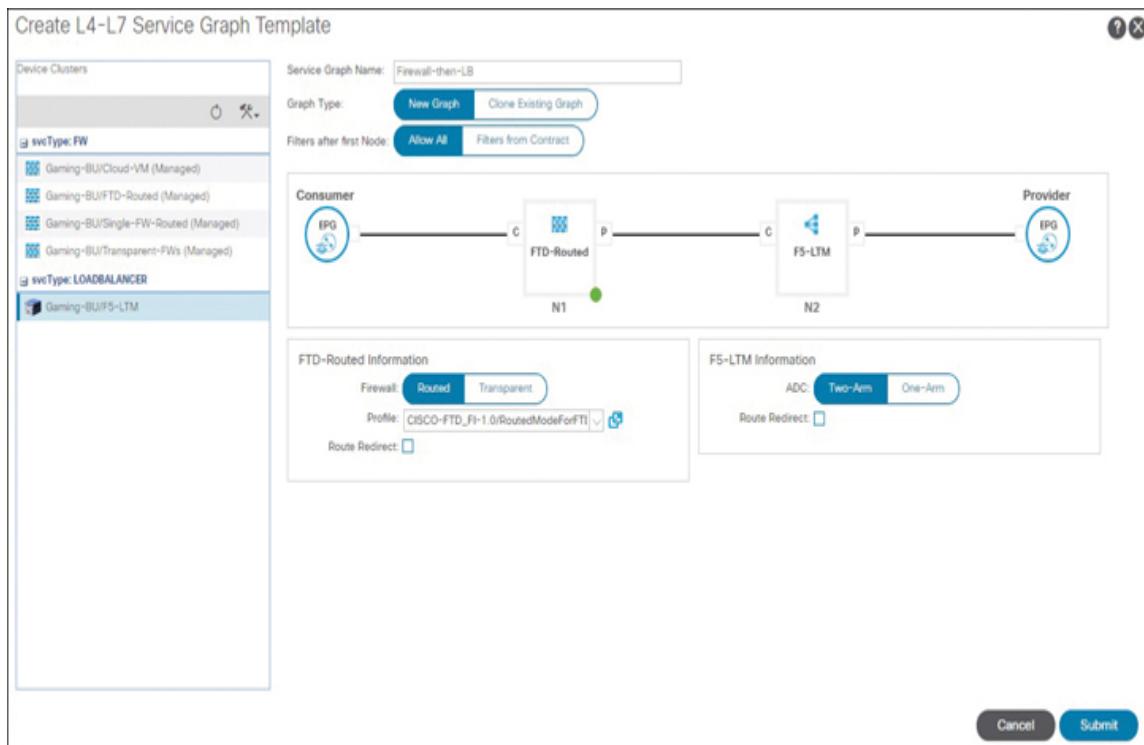
## Configuring a Service Graph Template

A service graph template defines the desired functions through which certain traffic should flow without actually specifying the traffic (EPGs) to which it applies. Service graph templates are intended to be generic, so that they can be ported to multiple contracts between different EPGs when feasible. This generic nature also means that equivalent service graph templates calling the same functions can be deployed in different data centers and rendered with locally available services appliances.



[Figure 12-19](#) shows a service graph template that chains a firewall and load balancer between a consumer and a provider EPG. To create a service graph template like this, you drag select device clusters (logical devices) from the left-hand column into the work pane in the middle of the

screen. Each function is represented as a node. N1 and N2 are *function node* names. Service graph template subconfiguration gets tied to these function node names. Note also the consumer connectors and provider connectors represented by C and P, respectively. These represent connection points for shadow EPGs and application of traffic filters. As shown, service graph templates do not include a dramatic amount of information for each node. Primarily, the configuration includes the firewall or load balancer configuration type, the function profile (Profile) used for deployment of managed or hybrid mode devices, and whether the Route Redirect setting or PBR should be used.



**Figure 12-19** Creating a Service Graph Template

## Configuring Device Selection Policies

As explained earlier in this chapter, service graphs insert L4-L7 devices into the traffic path. Device selection policies

determine which L4-L7 devices are inserted.

In short, a *device selection policy*, or *logical device context*, is a policy that associates or ties one or more L4-L7 devices to a graph template and contract.



Among the benefits of separating device selection from service graphs is that if you have an existing firewall deployed in a graph and you want to replace it, you simply need to define where the new firewall is connected and how it should be managed. Then you reference the newly introduced firewall in the device selection policy, and the associated graph(s) then points to the new firewall. As part of rendering, ACI then configures the new firewall just like the existing one if the firewall has been configured in managed or hybrid mode.

## Applying a Service Graph Template

Creation of a service graph template does not trigger a rendering of service graphs. A service graph template needs to first be instantiated. This can be done by right-clicking a service graph template and selecting Apply L4-L7 Service Graph Template. This launches a wizard that enables users to apply the service graph template to a consumer and a provider EPG by defining a new contract or adding the L4-L7 service policy to an existing contract. This wizard also requests user input regarding consumer connector and provider connector configurations. These connector configurations also include administrators specifying the consumer and provider bridge domains.



What is not evident in the service graph template application wizard is that this same process also prompts ACI to create a device selection policy behind the scenes.

An auto-generated device selection policy is bound to a single contract. For this reason, each time an administrator needs to reuse the service graph template for a new set of EPGs, the service graph template needs to be reapplied using the same wizard.

An alternative to this approach is to manually create a device selection policy that includes Any as the associated contract. Then, any application of the associated L4-L7 services policy in a new contract enables automatic reuse of the service graph template without having to go through the wizard again.

## Configuring Additional Service Graph Parameters

When a graph is instantiated, the APIC resolves the needed configuration for a service graph by looking up the parameters in various places. These parameters are commonly placed at the provider EPG level or the tenant level.

For service graphs that have been instantiated and require configuration updates, it is sometimes easiest to make changes by navigating to **Services > L4-L7 > Service Parameters** in the desired tenant and updating the desired service graph parameters.

# Monitoring Service Graphs and Devices

After a service graph has been instantiated, the best way to monitor it is to navigate to **Services > L4-L7 > Deployed Graph Instances** within the tenant where a service graph has been deployed and find the service graph instance in the list. The state applied means the graph has been applied and is active in the fabric and the services device.

To monitor devices, navigate to **Services > L4-L7 > Deployed Devices** in the tenant. You should be trying to achieve the operational state stable.

Any object directly or indirectly related to service graphs showing fault can be problematic for service graphs. Verify that there are no relevant faults and that the desired data plane forwarding is in place before announcing that the mission has been accomplished.

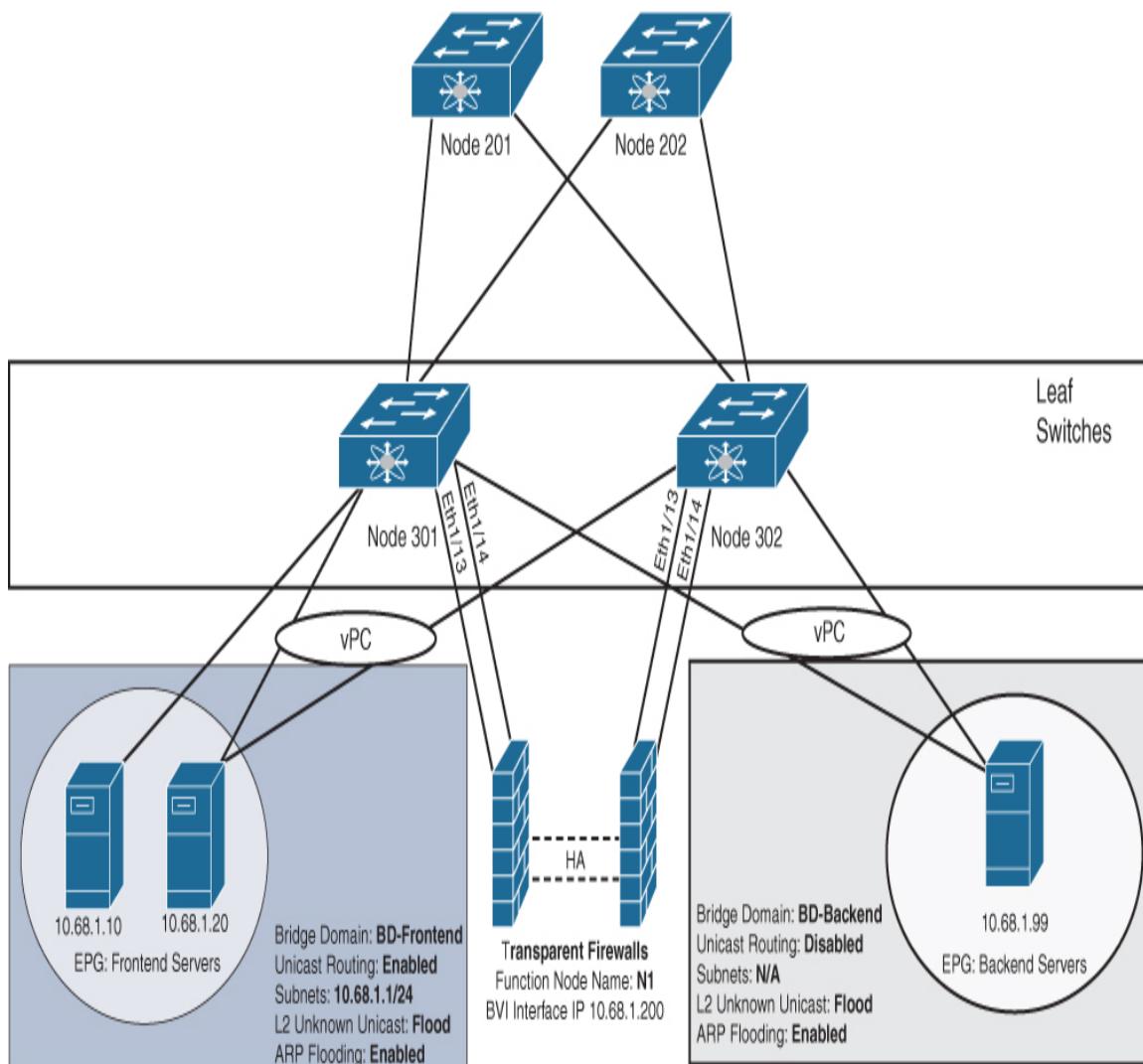
## Service Graph Implementation Examples

It can be difficult to take in all the details of the service graph implementation workflow without an example, so this section solidifies the concepts covered so far in this chapter by showing how to configure some basic service graphs. These examples address a diversity of configuration options while also remaining as straightforward as possible.

## Deploying an Unmanaged Firewall Pair in a Service Graph

In the example shown in [Figure 12-20](#), two ASA transparent firewalls need to be integrated into a tenant named Gaming-

BU to provide application inspection and segment two different EPGs. Base configurations and prohibitive firewall rules have already been put in place on the firewalls, and they have formed a high-availability pair. The firewalls are physical appliances and connect directly to the ACI fabric on ports 1/13 and 1/14 on leafs 301 and 302. The transparent firewalls need to be able to bridge traffic between an EPG called Frontend-Servers and another EPG called Backend-Servers.



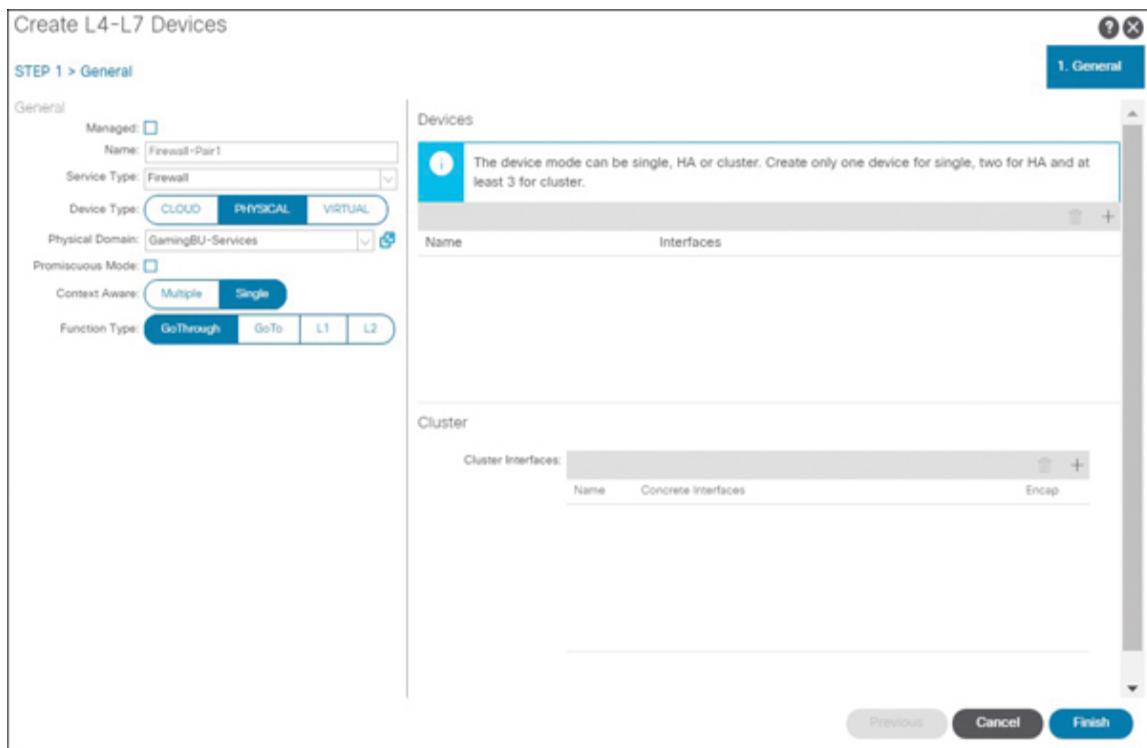
**Figure 12-20 Example for Unmanaged Transparent Firewall Insertion via Service Graphs**

Notice that two bridge domains are needed for this design. The subnet IP 10.68.1.1/24 assigned to BD-Frontend serves as the default gateway for endpoints in both EPGs. Removal of unicast routing capabilities together with the enablement of L2 unknown unicast flooding and ARP flooding on BD-Backend is required for insertion of the firewalls in the natural forwarding path between the EPGs.

Because devices outside the subnet beyond an L3Out may also at some point need access to endpoints in this subnet, one of these bridge domains needs to have unicast routing enabled—but which one? Service graphs that include transparent firewalls trigger faults when the provider-side bridge domain in the service graph enables IP routing. Hence, BD-Frontend has been selected as the routable bridge domain.

Given the requirements, there is no need to import device packages because the firewalls will be deployed in network policy mode.

Next, therefore, you need to identify the L4-L7 devices to the fabric. Open the tenant in which the L4-L7 device will be deployed and navigate to **Services > L4-L7**, right-click Devices, and select Create L4-L7 Devices to launch the wizard shown in [Figure 12-21](#).



**Figure 12-21** Providing General Information About an L4-L7 Device to ACI

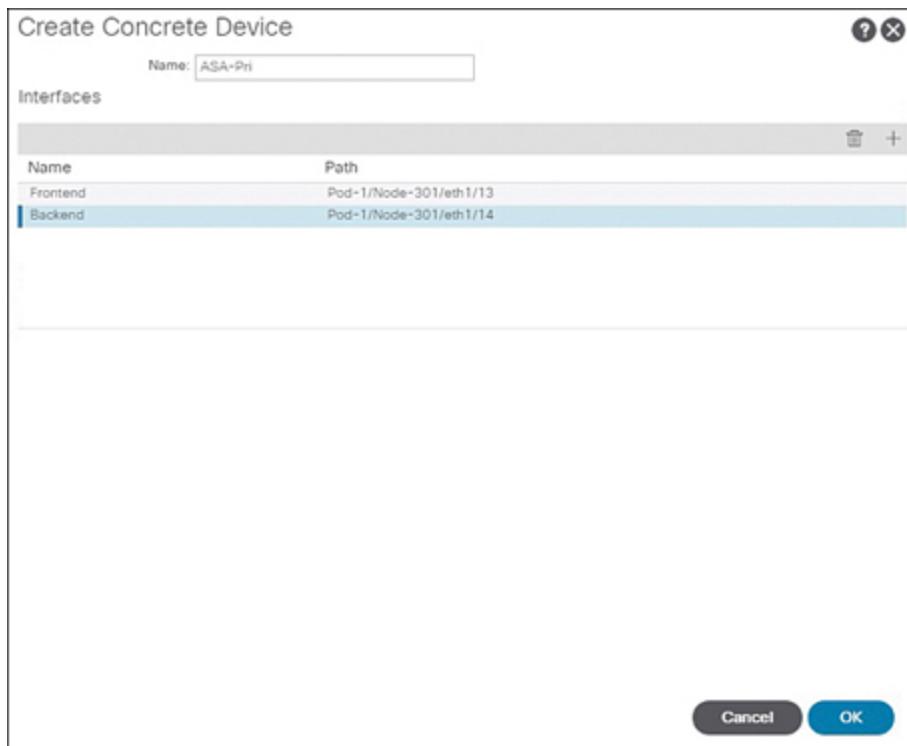
Because the firewalls are physical appliances, select Physical as the device type and select the physical domain to be used. Note that ACI deploys VLANs and stitches traffic for L4-L7 services devices. It is usually best to ensure that the VLAN pool associated with the selected physical domain has a dynamic VLAN range for ACI to choose from. In this example, Single has been selected to communicate to the system that the devices being defined will not be shared across tenants and will have no context awareness. Finally, select GoThrough as the function type to indicate that the devices are transparent firewalls.

### Note

Starting from APIC Release 4.1, PBR can be used with L4-L7 services devices operating in L1 or L2 mode.

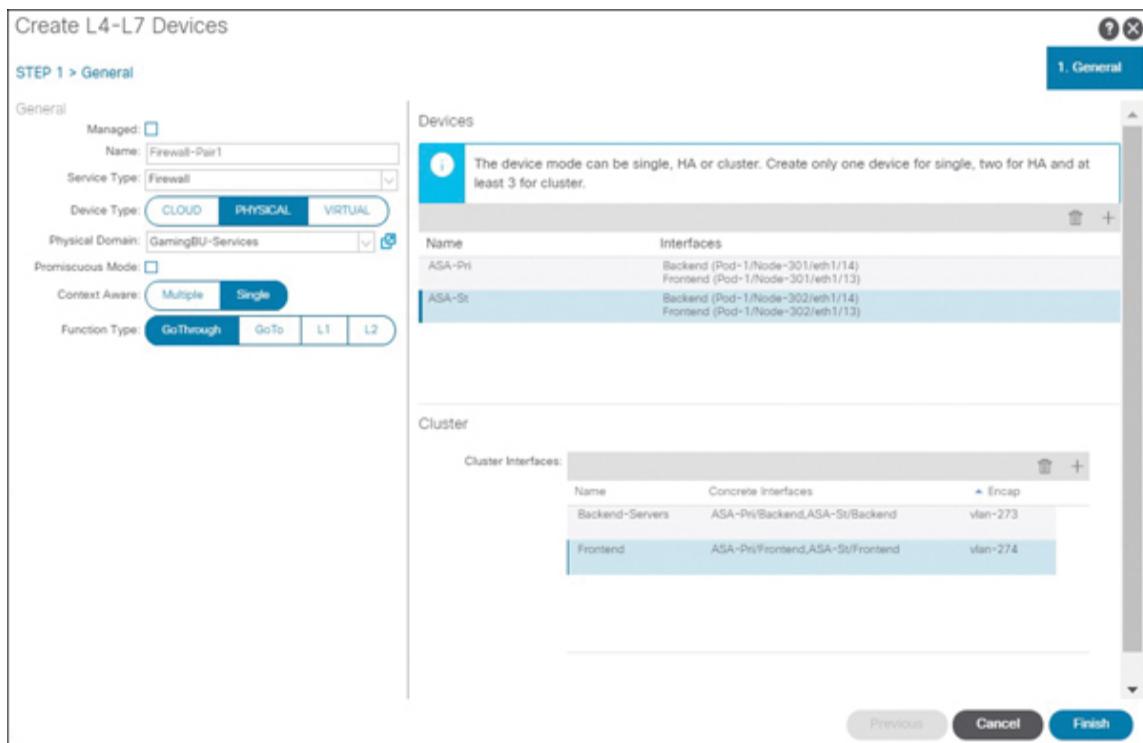
Function type options L1 and L2 shown in [Figure 12-21](#) refer to L1/L2 PBR.

When you are ready to configure concrete devices, click the + symbol in the Devices section to launch the page shown in [Figure 12-22](#). Enter a name in the Name field to identify the first transparent firewall to ACI. Then enter interface names along with path information, one line at a time. Note that [Figure 12-22](#) shows information entered only for the data interfaces because failover and state links are irrelevant when deploying service graphs for unmanaged devices. Click OK when you're finished making selections and then repeat the process to identify the standby firewall as an additional concrete device.



**Figure 12-22** Defining a Concrete Device to an ACI Fabric

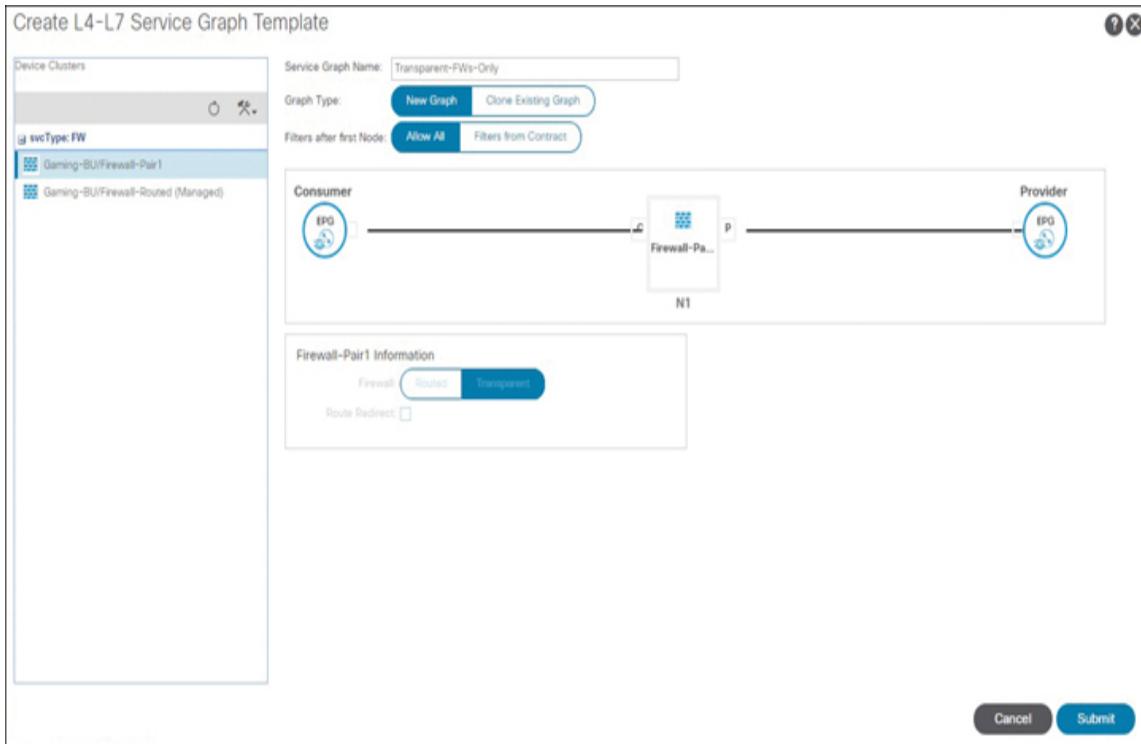
After you define the concrete devices, you need to define cluster interfaces. As opposed to concrete interfaces, cluster or logical interfaces are basically logical mappings to the interfaces on each concrete device. Put in a slightly different way, defining logical interfaces is like putting a label on interfaces that require similar policy assignment. Notice that [Figure 12-23](#) uses the label Backend-Servers as a cluster interface name; as shown earlier, the concrete interface name for the interfaces is Backend. You can see through this example that the names assigned in the Name columns are truly arbitrary. The cluster interfaces need to accurately reference the concrete interface names, but beyond that, concrete interface names are barely used in any other ACI configuration. The cluster interfaces in particular are what other objects reference; therefore, user-friendly names for the cluster interfaces are recommended. Click Finish to execute creation of the L4-L7 device definitions.



**Figure 12-23** *Mapping Cluster Interfaces to Concrete Interfaces on Each Device in a Pair*

With devices defined, it is time to create a service graph template. To do so, navigate to **Services > L4-L7** within the tenant, right-click Service Graph Templates, and select Create Service Graph Template.

[Figure 12-24](#) shows an example with a service graph name and the transparent firewall pair just defined dropped into the work pane. By default, a node is assigned a function node name. This example uses the default function node name N1. This is the name the service graph uses to identify the first function of interest. If additional functions are added to form a service chain, these new functions are assigned different node names. Notice that the wizard has disabled all options for the firewall that has been dropped into the graph. This is because the devices have already been identified as transparent firewalls, and PBR is not an acceptable option for GoThrough devices. Finally, recall that service graph templates are intended to be reusable. Therefore, they focus on flow and functions and not on specific EPGs.

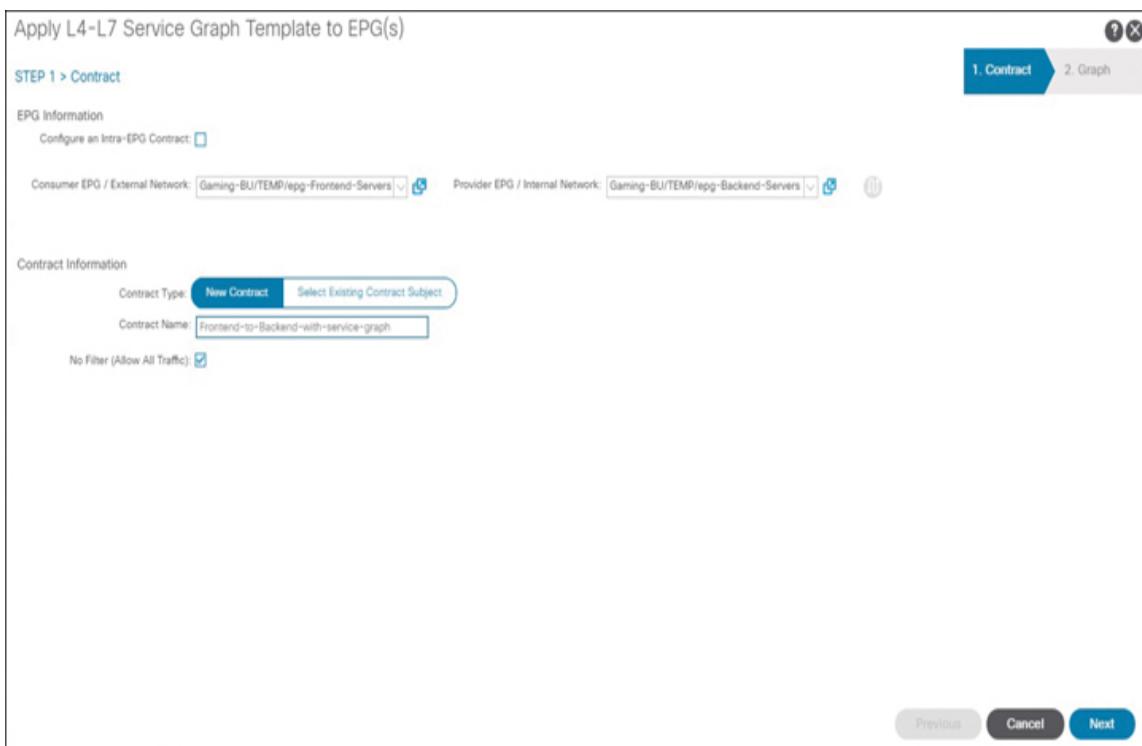


**Figure 12-24** Creating a Service Graph Template

Next, you need to decide whether to configure device selection policies manually. In terms of non-PBR service graphs, this decision often depends on how many contracts need to leverage the service graph template. If a large number of contracts are needed, it is often most practical to define device selection policies with an any contract. Otherwise, the Apply L4-L7 Service Graph Template to EPG(s) wizard automatically creates device selection policies, but the wizard then needs to be run again for each new contract. Because these firewalls are not expected to require any additional contracts, in this case you can apply the service graph template by using the wizard.

Another decision an implementation engineer needs to make is whether to deploy a custom function profile. This decision does not apply to the current deployment because function profiles are specific to managed and hybrid mode devices.

To instantiate a service graph template, right-click it and select Apply L4-L7 Service Graph Template to EPG(s). [Figure 12-25](#) shows the wizard that appears. Identify the client-side EPG or external EPG in the Consumer EPG/External Network drop-down on the left. Select the server-side EPG in the Provider EPG/Internal Network drop-down. Then determine if the service graph policy should be attached to an already existing contract or whether the system should create one. When you ask ACI to create a contract by using the No Filter (Allow All Traffic) checkbox, ACI creates a subject using a filter in the common tenant that matches all traffic. Disable this checkbox if you want to match specific protocols only. Click Next to move on to the next step of the wizard.



**Figure 12-25** Applying a Service Graph Between Two EPGs Using the Wizard

The second page in the wizard, shown in [Figure 12-26](#), is where you configure provider connector and consumer connector interfaces. Based on the EPGs selected in the

previous step, ACI attempts to intelligently populate the BD drop-down boxes. It is not always successful, but in this case, the EPGs are in the same subnet as the firewall interfaces. Therefore, the bridge domains ACI has selected are correct. Ensure that the correct cluster interface has been chosen for both the provider connector and the consumer connector and click Finish.



**Figure 12-26** Configuring Consumer Connector and Provider Connector Interfaces

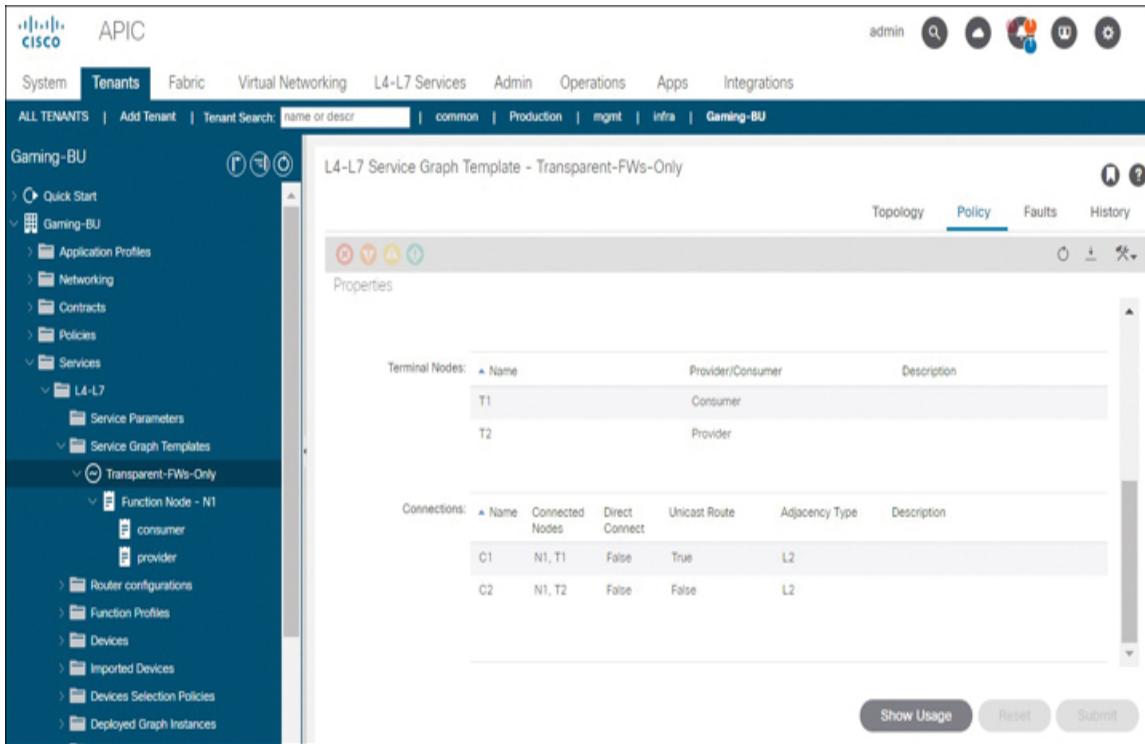
### Note

The L3 Destination (VIP) checkbox identifies nodes that do not require redirection when both PBR and non-PBR nodes are deployed in a service chain. Therefore, this checkbox is of no significance in the non-PBR service graph examples presented in this chapter. In addition, at the time of writing, the only thing that can be

configured within a service EPG policy is preferred group membership.

As a result of instantiating the service graph template, ACI creates a device selection policy and applies a contract whose subject references the L4-L7 service policy to the consumer and provider EPGs selected by the user. ACI is then ready to render the newly created service graph. If any issues are identified during rendering, faults should appear. If you resolve the faults, forwarding should work as expected.

Sometimes there is a need to tell ACI that a given connector interface on a service graph template enables unicast routing. [Figure 12-27](#) shows how to do so. A *terminal node* enables input and output through a service graph. In this case, terminal node T1 has been indicated to be the consumer. This terminal node correlates with connection C1. A connection represents an input or output point for a node. You can see that the Unicast Route parameter is set to True. Note that the Adjacency Type column has not been modified manually. If a connector is associated with a bridge domain that provides an L3Out interface function, however, the adjacency needs to be set to L3. In this example, this setting is insignificant.



**Figure 12-27** Checking Whether Service Graph Template Connector Settings Are Correct

The ultimate test of a service graph is whether traffic is forwarded correctly through the desired functions.

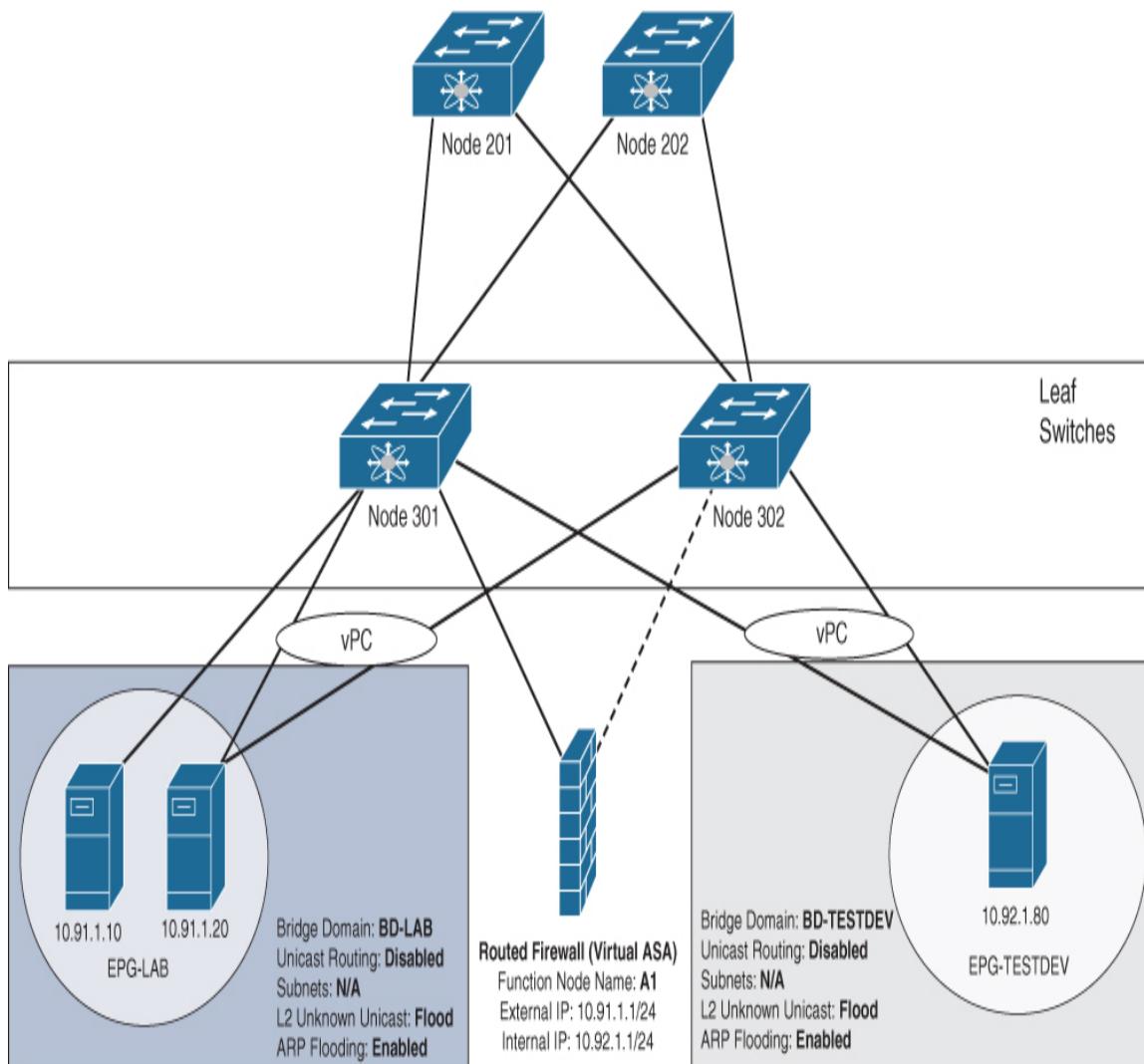
### Note

Another valid design for non-PBR service graphs consisting of transparent firewalls is to have an external router function as the default gateway for the bridged subnet and provide routing functionality.

## Deploying Service Graphs for a Firewall in Managed Mode

The example in this section involves service graphs for managed devices. [Figure 12-28](#) shows the topology used for

this example. In this case, the firewall is a virtual firewall within a VMM-integrated environment. This makes things easier because there is less concern about tracking physical interfaces. Both bridge domains have unicast routing disabled because the firewall serves as default gateway for endpoints within the two bridge domains.



**Figure 12-28 Example for Managed Routed Firewall Insertion via Service Graphs**

A device package has already been imported for the appliance. [Figure 12-29](#) shows creation of the L4-L7 device. The Managed checkbox is selected to signal to the system

that the device should be deployed in service policy or service manager mode. The device package settings determine which of these two modes is supported. When device type Virtual is selected, ACI exposes the VMM Domain drop-down, and you can select the VMM domain where the virtual L4-L7 services device(s) have been deployed. Notice that selection of a device package and model is mandatory during managed mode device definition. Also notice the concrete device portion of the screen. The Interfaces column shows interface identifiers that match those on an ASA virtual firewall. ACI is able to provide valid interface names as configurable options thanks to information it obtained from the device package. When managing L4-L7 devices through ACI, it is important to understand that these same interface names are used for configuration deployment. If incorrect interface names are entered, configuration deployment to the appliance can fail. Click Next to continue.

**General**

- Managed:
- Name: Firewall-Routed
- Service Type: Firewall
- Device Type: **VIRTUAL**
- VMM Domain: DC2-VMM
- Trunking Port:
- VM Instantiation Policy: select an option
- Device Package: CISCO-ASA-1.2
- Mode: ASAv
- Promiscuous Mode:
- Context Aware: **Single**
- Function Type: GoThrough

**Devices**

Name	VM Name	vCenter Name	Management Address	Management Port	Interfaces
ASAv01	ASAv-Routed	DC2-vCenter	10.233.64.90	443	GigabitEthernet0/0 GigabitEthernet0/1

**Cluster**

- Management IP Address: 10.233.64.90
- Management Port: https
- Device Manager: select a value
- Cluster Interfaces:

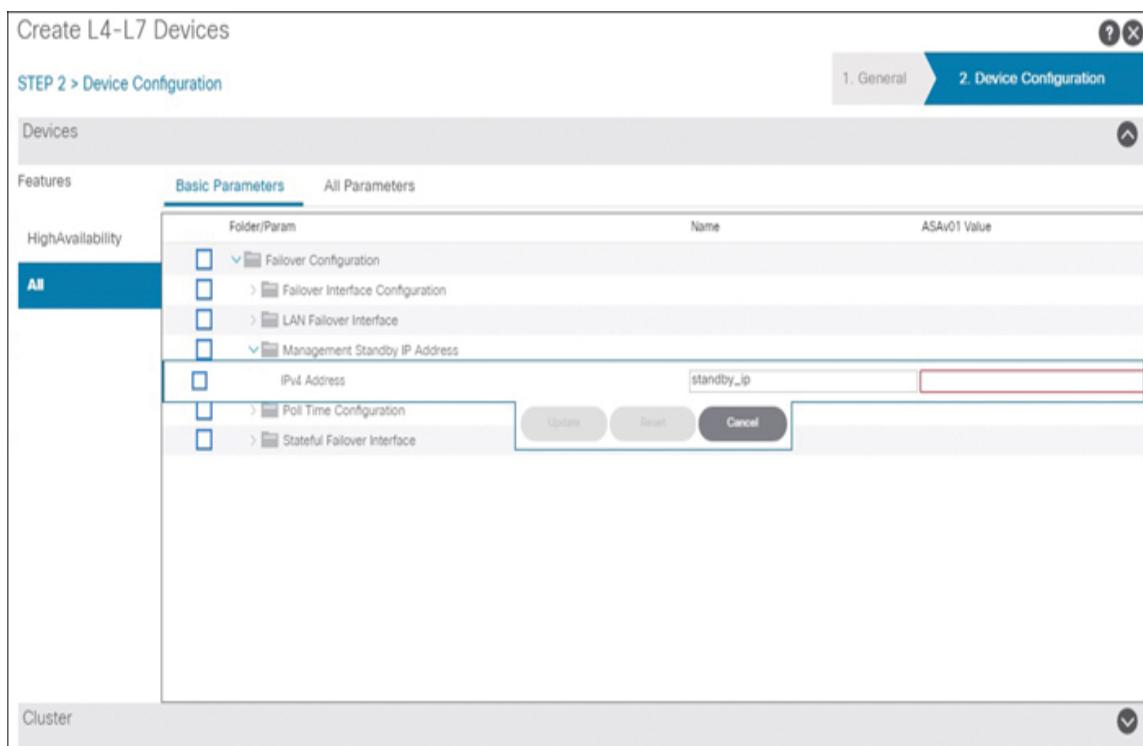
Type	Name	Concrete Interfaces
consumer	Outside	ASAv01/GigabitEthernet0/0
provider	Inside	ASAv01/GigabitEthernet0/1

Buttons: Previous, Cancel, **Next**

**Figure 12-29** Creation of an ACI-Managed L4-L7 Device Definition

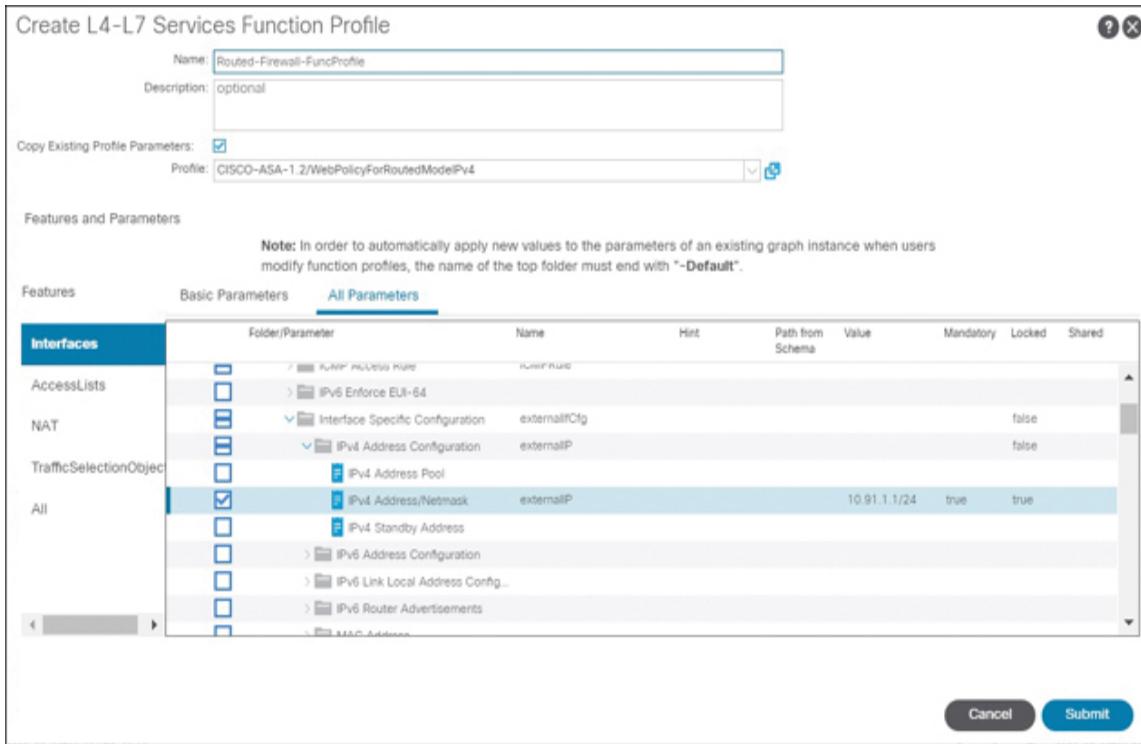
During managed mode device definition, the APIs give you the option to define certain parameters in line with requirements dictated in the selected device package.

[Figure 12-30](#) shows that high availability can be configured during device definition, if desired.



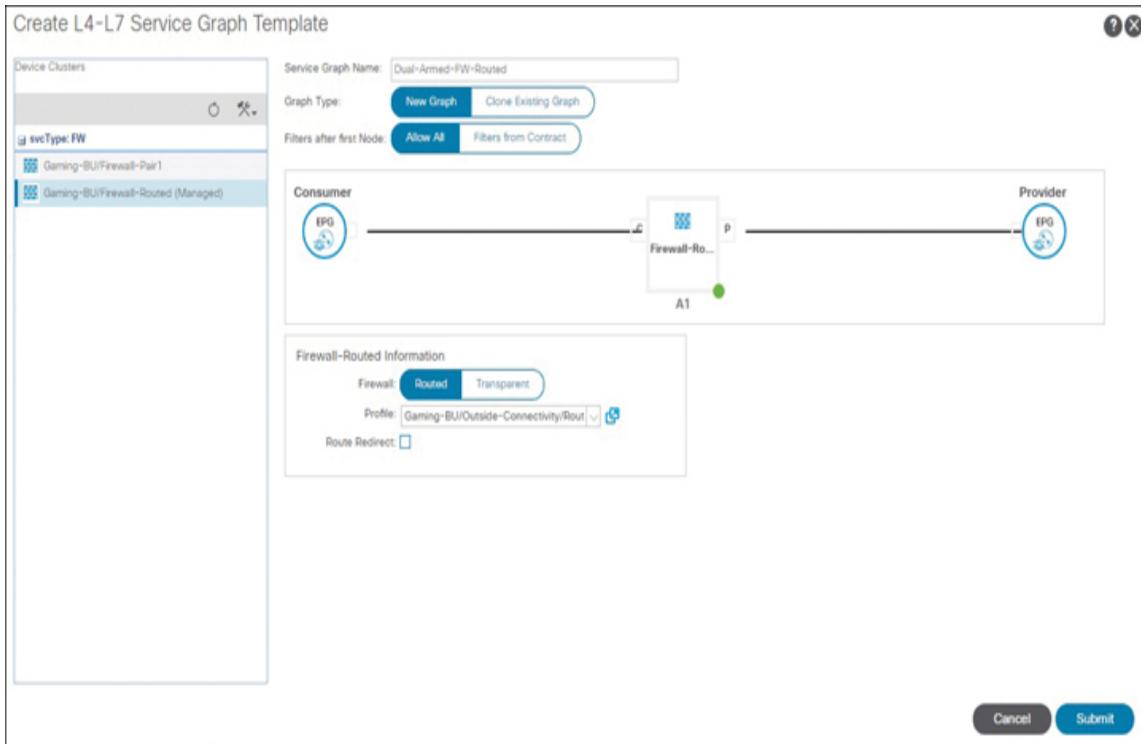
**Figure 12-30** High-Availability Configuration Options If Needed

In parallel with defining L4-L7 devices, a services administrator may want to define a function profile to encourage reuse or prevent other administrators from making certain changes. [Figure 12-31](#) shows configuration of a function profile to avoid repeat entry of critical data. Notice that the Locked attribute for interface IP addresses is toggled to True, to prevent users consuming the function profile from making modifications.



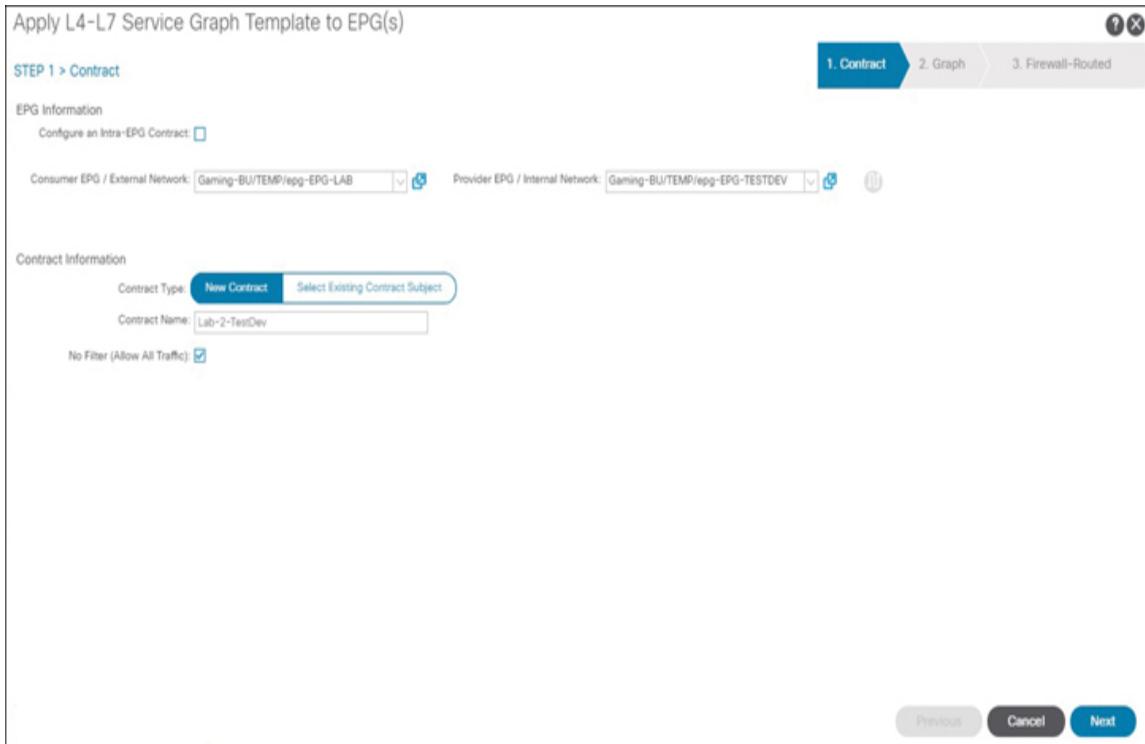
**Figure 12-31** Defining a Function Profile for Configuring an L4-L7 Services Device

Next, a service graph template needs to be created. You can select the newly created function profile from the Profile drop-down box as shown in [Figure 12-32](#).



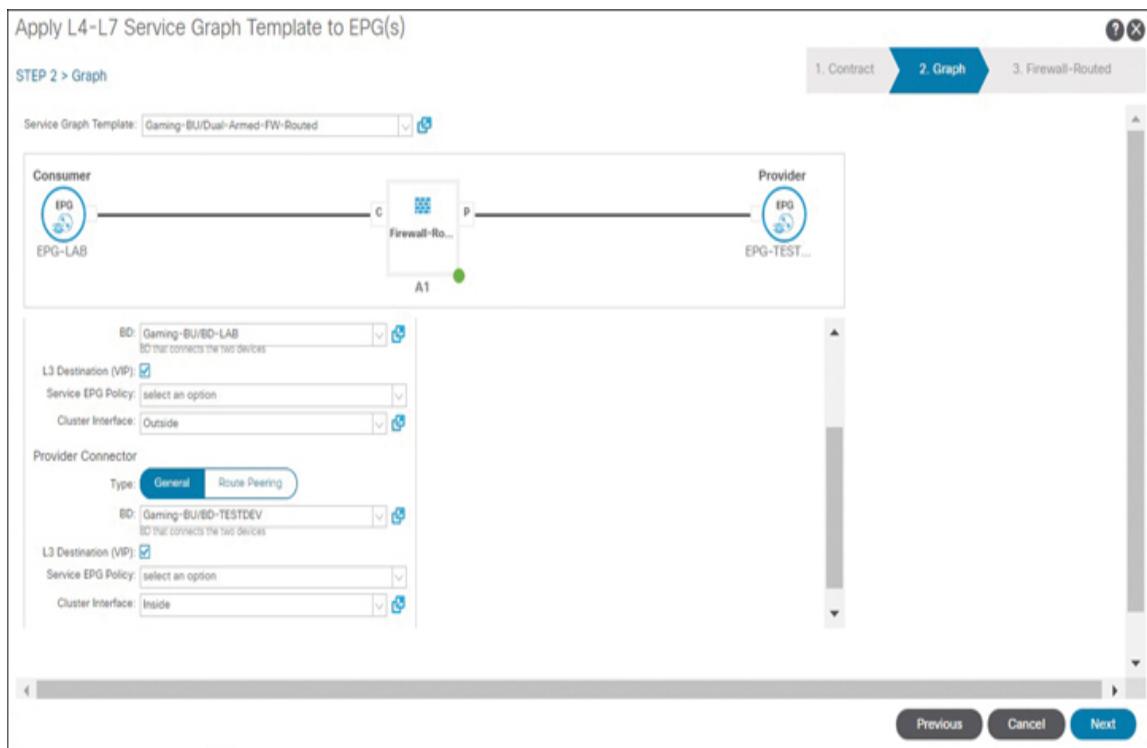
**Figure 12-32** *Associating a Function Profile with a Service Graph Template*

Next, you define the consumer and provider EPGs and create a contract for assignment to the EPGs and allocation of the L4-L7 service graph, as shown in [Figure 12-33](#).



**Figure 12-33** Applying a Contract and Service Graph to Provider and Consumer EPGs

The next step is to configure the consumer connector and the provider connector interfaces for the node. This involves configuring interfaces connecting to the appliance(s) and not configuring ports within the services appliance operation system. The process, shown in [Figure 12-34](#), is very similar to the preceding example.



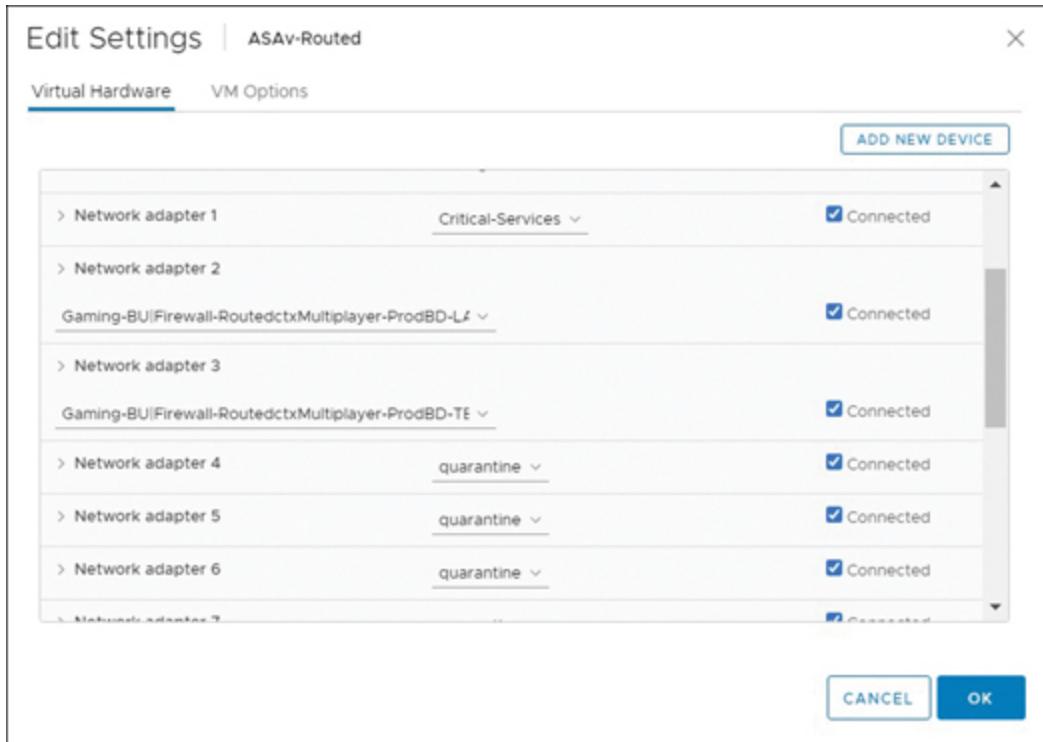
**Figure 12-34** Configuring Consumer and Provider Connector Interfaces

Finally, it is time to define configurations for deployment to the services device(s). Notice that [Figure 12-35](#) suggests that no parameters are required for deployment. This is not because no parameters have been set as mandatory. The reason nothing is shown actually is that all mandatory parameters have already been entered into the associated function profile. Click Finish to have ACI render the service graph.



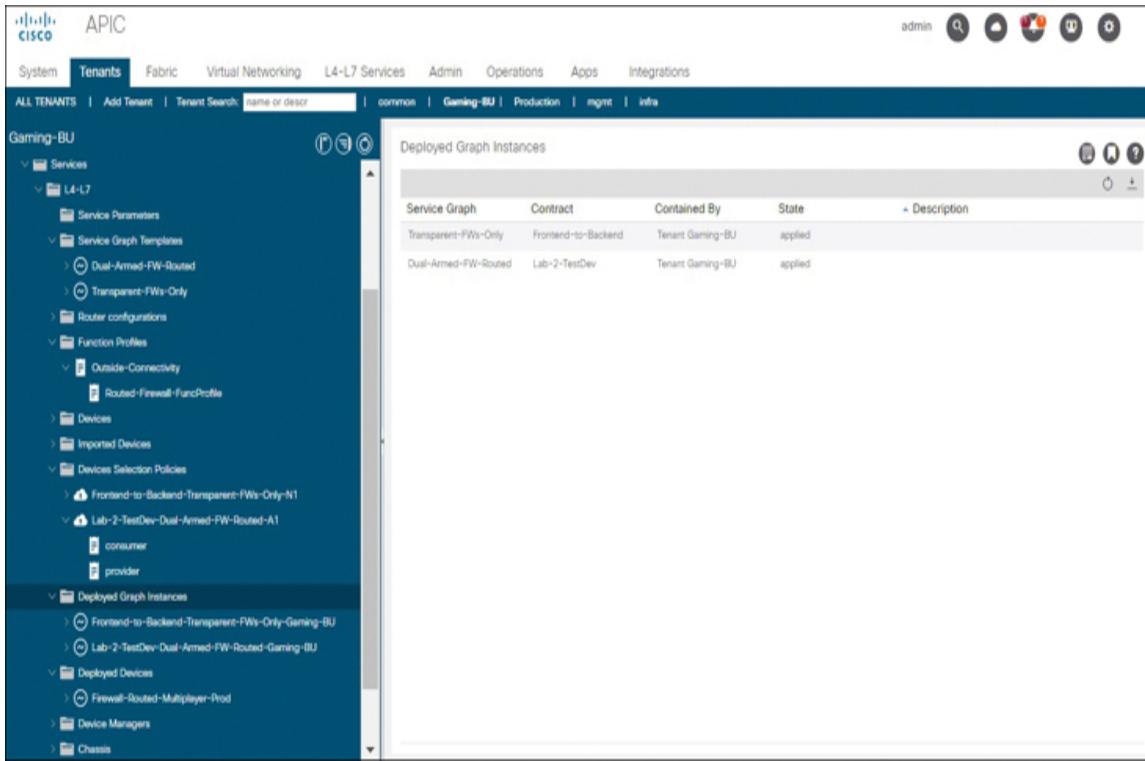
**Figure 12-35** Services Device Configuration Parameters

As part of rendering, ACI pushes port groups into vCenter based on the number of interfaces defined in the device configuration and the graph configuration. These new port groups are shadow EPGs and follow a naming convention that includes the VRF instance name. [Figure 12-36](#) shows two shadow EPGs assigned to the VM interfaces as port groups. This automation takes place without any further user interaction.



**Figure 12-36 ACI Assigns Port Groups to Relevant Network Adapters on a Virtual Appliance**

The status of instantiated service graphs can be reviewed in the Deployed Graph Instances folder, as shown in [Figure 12-37](#). The state applied indicates a successfully rendered graph.



**Figure 12-37 State of Service Graph Instances Within a Tenant**

As a result of the rendering process, ACI has configured the ASA. [Example 12-1](#) shows some lines of interest that are directly related to this configuration.

**Example 12-1 Sample Configuration Deployed by ACI to a Virtual ASA in Managed Mode**

[Click here to view code image](#)

```

interface GigabitEthernet0/0
    nameif externalIf
    security-level 50
    ip address 10.91.1.1 255.255.255.0
!
interface GigabitEthernet0/1
    nameif internalIf
    security-level 100

```

```

ip address 10.92.1.1 255.255.255.0
!
object network web_server
  subnet 10.92.1.0 255.255.255.0
access-list access-list-inbound extended permit tcp any
object web_server eq www
access-list access-list-inbound extended permit tcp any
object web_server eq https
access-group access-list-inbound in interface externalIf

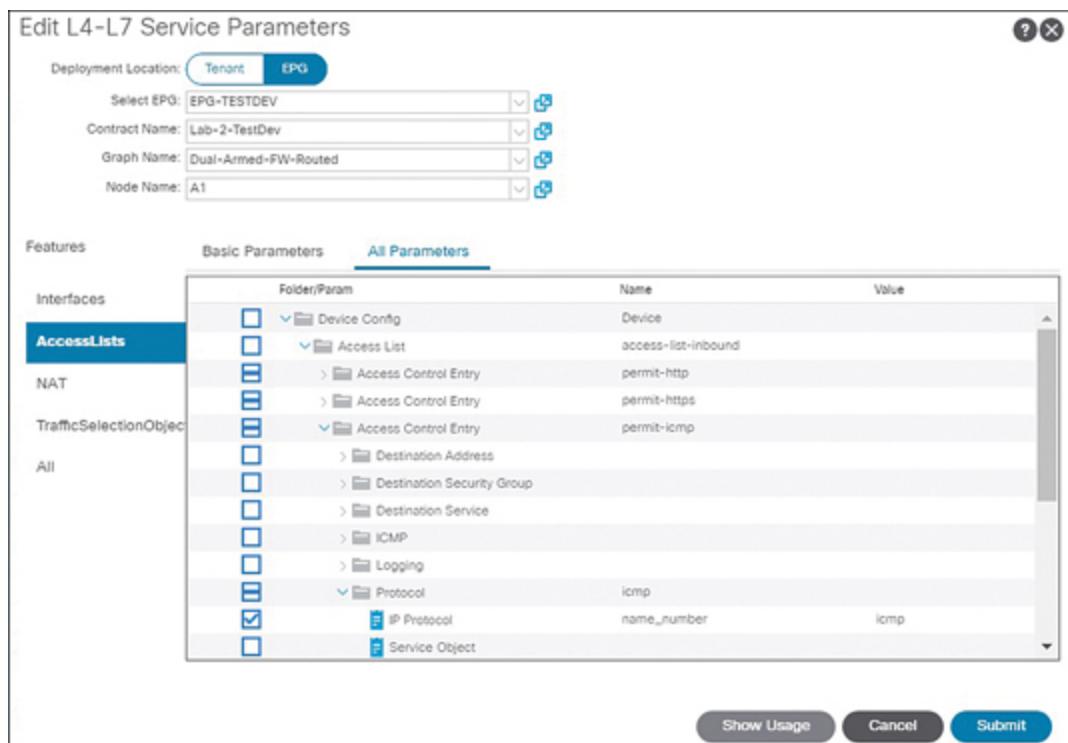
```

Once a graph involving managed devices has been rendered, it is often best to edit the graph parameters if changes need to be made to the services device configuration. As shown in [Figure 12-38](#), you can navigate to **Services > L4-L7 > Services Parameters** within the tenant and click Switch to Edit Mode to modify the deployed service parameters or to add new device configurations.

Meta Folder/Param Key	Contract Name	Deployed Location	Service Graph Name	Service Function Name	Folder/Param Instance Name	Value	Switch To Edit Mode
ExtConfigRefFolder	Lab-2-TestDev	ap-TEMP eppg-...	Dual-Armed-F...	A1		ExtConfig	
IntfConfigRefFolder	Lab-2-TestDev	ap-TEMP eppg-...	Dual-Armed-F...	A1		IntConfig	
<b>AccessList</b>	Lab-2-TestDev	ap-TEMP eppg-...	Dual-Armed-F...	A1		<b>access-list-inb...</b>	<b>Device</b>
Interface	Lab-2-TestDev	ap-TEMP eppg-...	Dual-Armed-F...	A1		externall	
Interface	Lab-2-TestDev	ap-TEMP eppg-...	Dual-Armed-F...	A1		internall	
NetworkObject	Lab-2-TestDev	ap-TEMP eppg-...	Dual-Armed-F...	A1		web_server	

**Figure 12-38** Editing L4-L7 Services Parameters

Device configurations can be applied at the provider EPG level or at the tenant level, as shown in [Figure 12-39](#). Select the relevant configurations from the drop-down boxes to expose current configurations. [Figure 12-39](#) shows configurations being added for ICMP.



**Figure 12-39** Configuring L4-L7 Services Parameters for Deployment to Services Devices

As a result of these configurations, ACI adds the command **access-list access-list-inbound extended permit icmp any any** to the firewall configuration. Although it is simple, this example shows how function profiles and service parameters are used.

## Exam Preparation Tasks

As mentioned in the section “How to Use This Book” in the Introduction, you have a couple of choices for exam

preparation: Chapter 17, “Final Preparation,” and the exam simulation questions in the Pearson Test Prep Software Online.

## Review All Key Topics

Review the most important topics in this chapter, noted with the Key Topic icon in the outer margin of the page. **Table 12-4** lists these key topics and the page number on which each is found.



**Table 12-4** Key Topics for [Chapter 12](#)

Key Topic Element	Description	Page Number
Paragraph	Defines service graphs	<a href="#">427</a>
List	Describes the different management models achieved through service graphs	<a href="#">428</a>

Key Topic Element	Description	Page Number
Paragraph	Describes the scope of configuration automated for services devices deployed in network policy mode	428
Paragraph	Describes device packages	430
Paragraph	Describes function profiles and their use in feeding configuration data to L4-L7 services devices	430
Paragraph	Summarizes some key aspects regarding the deployment of L4-L7 services devices in service policy mode	430
Paragraph	Describes the function of a device manager	433

Key Topic Element	Description	Page Number
List	Explains device deployment modes and the number of bridge domains needed for typical designs using each deployment mode	435
List	Describes service graph designs for GoTo devices	436
List	Describes service graph designs for GoThrough devices	437
List	Lists the most common bridge domain settings for non-PBR service graph connector interfaces	439
Paragraph	Explains the absolute requirement for Layer 2 Unknown Unicast Flooding and ARP Flooding for transparent firewalls	440
Paragraph	Defines rendering	440

Key Topic Element	Description	Page Number
List	Lists the service graph implementation workflow	<a href="#">441</a>
<a href="#">Table 12-2</a>	Describes the components included in a device package	<a href="#">442</a>
<a href="#">Figure 12-18</a>	Provides an example of defining L4-L7 devices for deployment in service manager mode	<a href="#">443</a>
Paragraph	Defines concrete devices and logical devices and their associated interfaces	<a href="#">444</a>
<a href="#">Table 12-3</a>	Describes some configuration options available in the Create L4-L7 Devices page	<a href="#">444</a>
List	Describes the function profile attributes mandatory, locked, and shared	<a href="#">445</a>

Key Topic Element	Description	Page Number
Paragraph	Describes service graph templates	445
Paragraph	Describes device selection policies	446
Paragraph	Describes the purpose of the Apply L4-L7 Service Graph Template wizard	446

## Complete Tables and Lists from Memory

Print a copy of [Appendix C, “Memory Tables”](#) (found on the companion website), or at least the section for this chapter, and complete the tables and lists from memory. [Appendix D, “Memory Tables Answer Key”](#) (also on the companion website), includes completed tables and lists you can use to check your work.

## Define Key Terms

Define the following key terms from this chapter and check your answers in the glossary:

service graph  
network policy mode  
shadow EPG  
service policy mode  
service manager mode  
function profile  
device manager  
service chain  
concrete device  
logical device

# **Part V: Management and Monitoring**

# **Chapter 13**

## **Implementing Management**

**This chapter covers the following topic:**

**Configuring Management in ACI:** This section compares in-band and out-of-band management and covers the deployment of in-band management.

This chapter covers the following exam topic:

- 5.1 Implement out-of-band and in-band

ACI offers two avenues for management and monitoring as well as cross-platform communications such as VMM integration and managed service graphs: in-band management and out-of-band management.

**Chapter 3, “Initializing an ACI Fabric,”** covers the implementation of out-of-band (OOB) management. This chapter revisits out-of-band management to uncover some of the reasons in-band management is sometimes desirable. It also covers the deployment of contracts for management access.

## **“Do I Know This Already?” Quiz**

The “Do I Know This Already?” quiz allows you to assess whether you should read this entire chapter thoroughly or jump to the “Exam Preparation Tasks” section. If you are in doubt about your answers to these questions or your own assessment of your knowledge of the topics, read the entire chapter. [Table 13-1](#) lists the major heading in this chapter and its corresponding “Do I Know This Already?” quiz questions. You can find the answers in [Appendix A, “Answers to the ‘Do I Know This Already?’ Questions.”](#)

**Table 13-1** “Do I Know This Already?” Section-to-Question Mapping

Foundation Topics Section	Questions
Configuring Management in ACI	1-10

### Caution

The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark that question as wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

1. True or false: Changes to ACI access policies cannot directly affect out-of-band management connectivity.
  - a. True

- b.** False
- 2.** True or false: Changes to ACI access policies cannot directly affect in-band management connectivity.
  - a.** True
  - b.** False
- 3.** True or false: One solution for enabling management system connectivity to out-of-band interfaces is to leak data plane traffic to the out-of-band network.
  - a.** True
  - b.** False
- 4.** True or false: When deploying in-band and out-of-band management side by side, Cisco recommends that either static IP addressing or dynamic IP addressing be used for both communication avenues.
  - a.** True
  - b.** False
- 5.** True or false: An administrator can create an L3Out to advertise out-of-band subnets out an ACI fabric.
  - a.** True
  - b.** False
- 6.** Which of the following steps cannot possibly be part of an in-band management deployment process?
  - a.** Assign in-band IP addresses to switches and APICs.
  - b.** Configure a gateway IP address on the inb subnet.
  - c.** Enable NTP under the Fabric Policies menu.
  - d.** Configure access policies and assign them to switch ports.

- 7.** True or false: The configuration of a managed node connectivity group is mandatory when using dynamic IP addressing.
- a.** True
  - b.** False
- 8.** True or false: APICs are VRF aware and have separate routing tables for segmentation of traffic into in-band and out-of-band VRFs.
- a.** True
  - b.** False
- 9.** True or false: All ACI management contracts offer the same features and functionality.
- a.** True
  - b.** False
- 10.** True or false: APIC default route metrics can be modified by using the APIC Connectivity Preferences setting.
- a.** True
  - b.** False

## Foundation Topics

### Configuring Management in ACI

An ACI fabric allows management access in the form of out-of-band management, in-band management, or both. This section helps you gain an understanding of some of the benefits and caveats of each option before going through the implementation of in-band management.

# Understanding Out-of-Band Management Connectivity

When deploying OOB management, network engineers often dedicate a set of low-cost non-ACI copper switches to the out-of-band function and attach all out-of-band links to these switches.

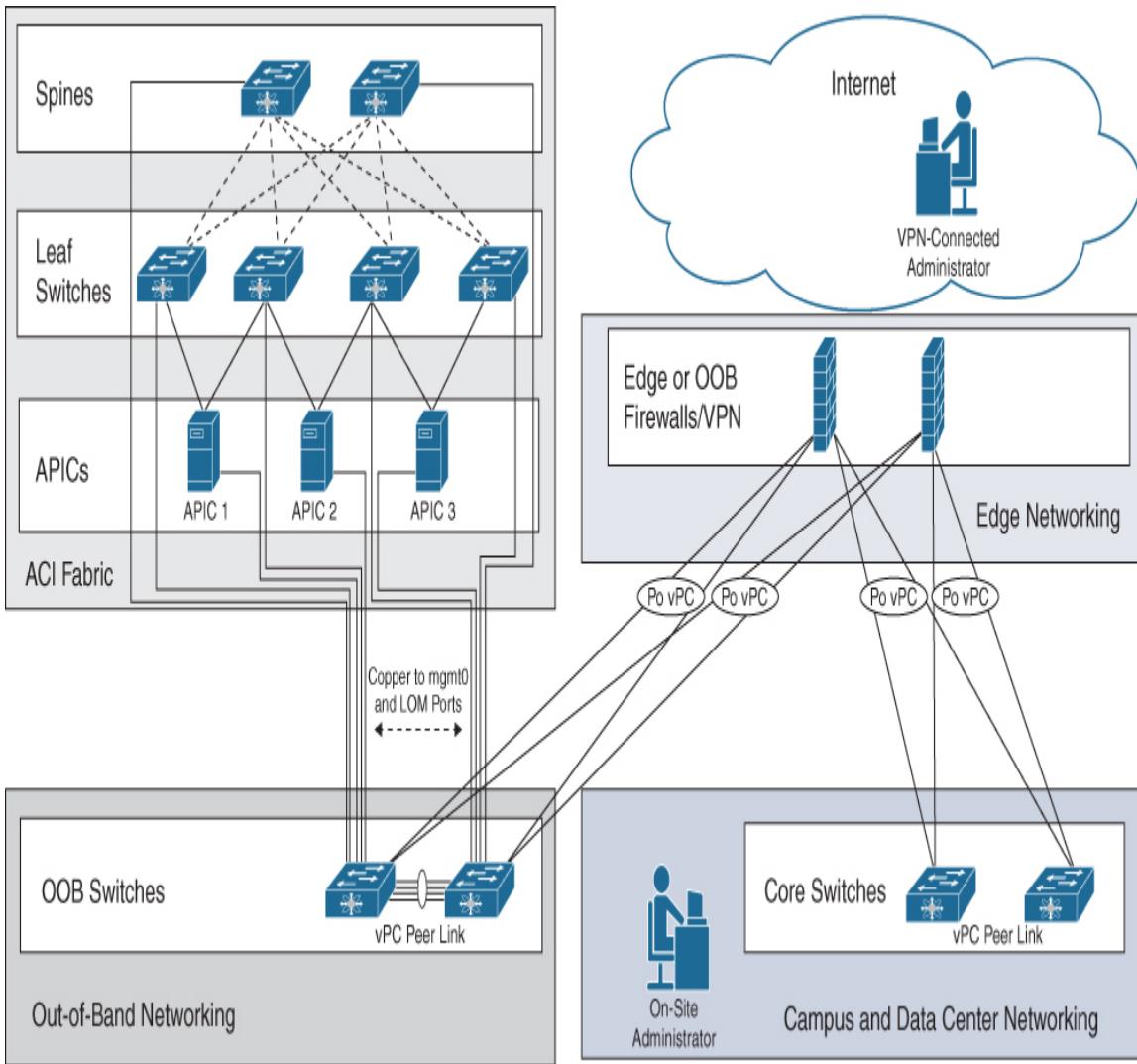
Sometimes terminal servers are also deployed alongside OOB switches to ensure that a misconfigured or inaccessible switch can be remotely restored. The APIC Cisco IMC in a sense functions like a terminal server, enabling KVM access to the APIC operating system. For this reason, terminal servers are not usually necessary for the recoverability of APICs, as long as administrators are able to maintain IP connectivity to the APIC Cisco IMC. This is why administrators commonly allocate APIC Cisco IMC addressing from OOB subnet ranges.



Generally speaking, out-of-band environments are perfect for network recoverability. This is also the case in ACI. In fact, misconfigurations of ACI access policies or fabric policies most likely *cannot* affect core connectivity to OOB interfaces. Furthermore, the design of out-of-band interfaces in ACI is such that ACI does not even handle routing for such interfaces, keeping connectivity into the out-of-band network as simple as possible.

It is very common for network engineers to connect OOB switches directly to firewalls so that administrators can VPN into the network and ensure rapid network recovery in the event of an outage.

**Figure 13-1** shows OOB network deployment for an ACI fabric. Notice that switch mgmt0 interfaces as well as selected APIC LOM ports should connect to the OOB environment. Connecting Cisco IMC interfaces to this network is optional but highly advised.



**Figure 13-1** *ACI OOB Management Connectivity in a Typical Network*

In case an ACI deployment relies solely on OOB management, a mechanism may be needed to ensure that the IP connectivity required for monitoring, VMM integration, and other cross-platform integrations is also available. To

establish such IP connectivity, [Figure 13-1](#) shows that the firewalls to which the out-of-band network connects has a secondary connection back to the data center core switching layer. The core switching layer can then direct traffic to vCenter instances for VMM integration or to management or monitoring stations within the network.

The requirement to establish physical connectivity back to the in-band network in OOB-only ACI deployments illuminates a key aspect of out-of-band management in ACI: ACI does not allow users to leak traffic from the out-of-band VRF into other tenants and VRFs within ACI.



Aside from recoverability, a key use case for out-of-band management is to ensure full management plane and data plane separation. For this reason, there is no configuration available to leak the out-of-band management plane into the data plane within ACI.

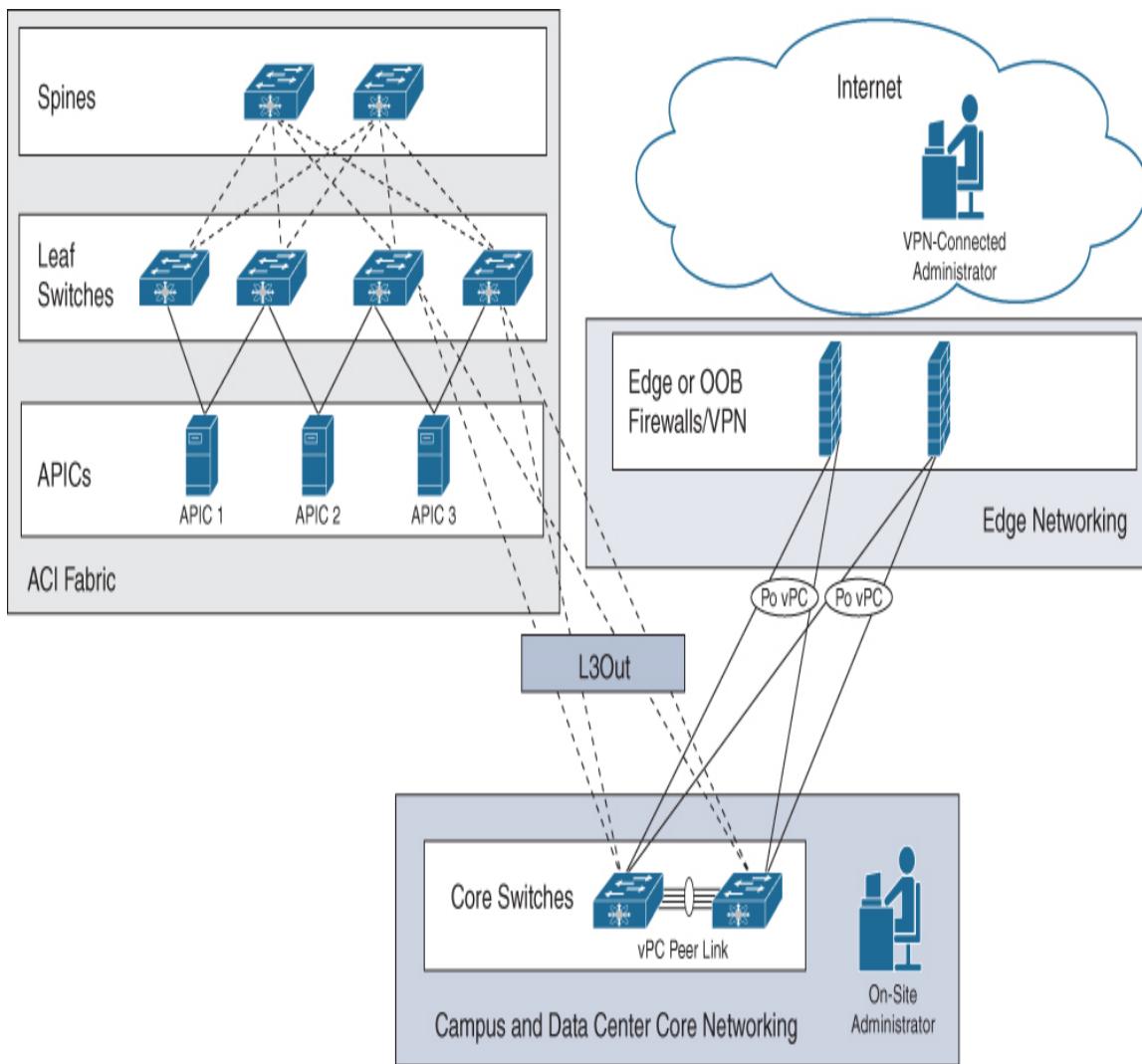
## **Understanding In-Band Management Connectivity**



In stark contrast with out-of-band management, in-band management enables a range of connectivity options, including use of EPG extensions, L2Outs, and L3Outs. Administrators can also add additional virtual network interface cards (vNICs) to VMs that may be used for monitoring and assign them to VMM-integrated port groups for direct IP assignment within the in-band management

subnet. In-band management also allows administrators to leak traffic from the in-band VRF denoted by `inb` into user-created VRFs. If monitoring systems reside within an ACI fabric, direct leaking of traffic to and from the in-band environment can enable monitoring of an ACI fabric to persist even when northbound non-ACI switches suffer an outage.

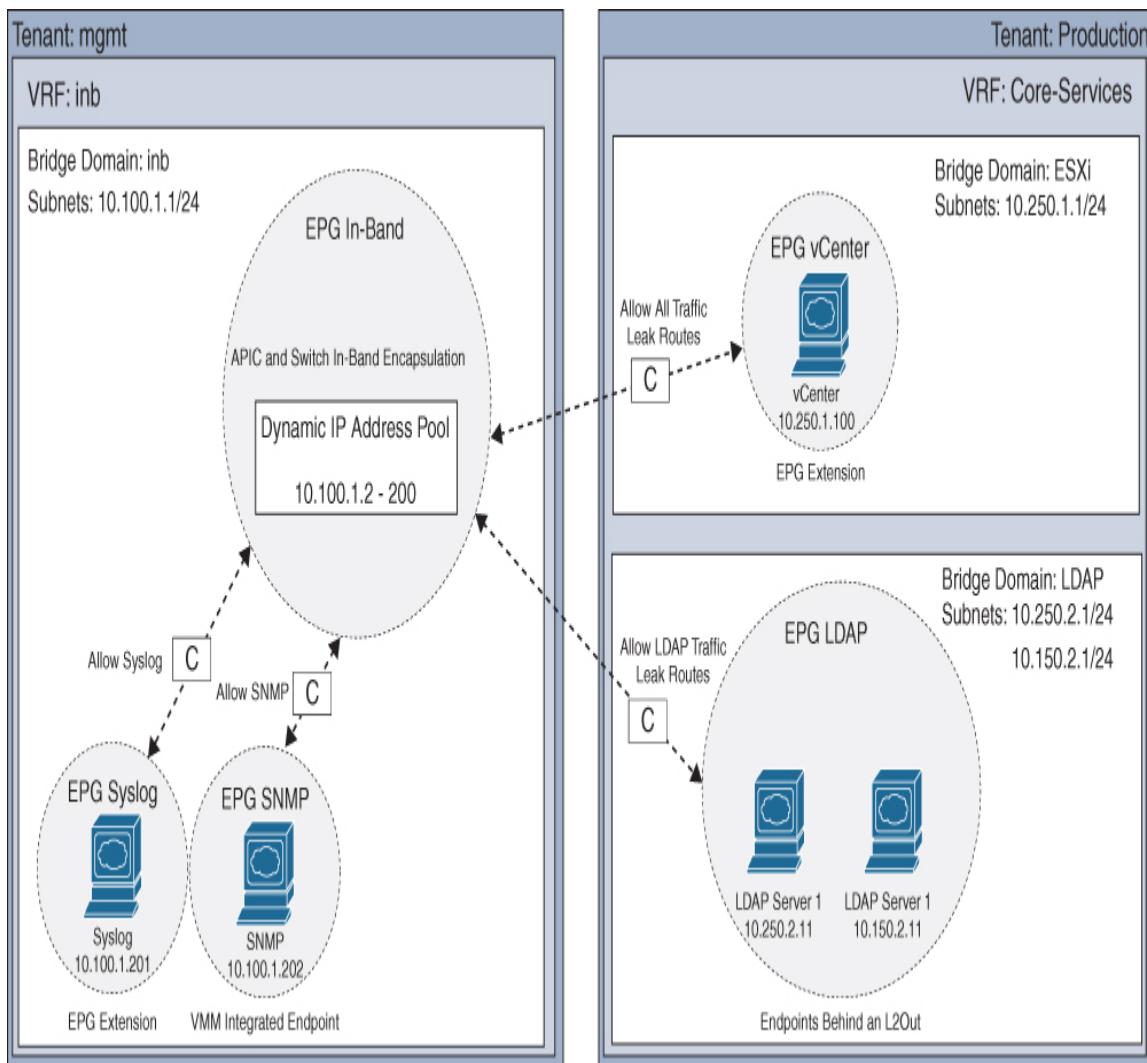
[Figure 13-2](#) demonstrates how an L3Out can be used to advertise an in-band management subnet to the rest of the network. In-band management deployments are likely to incorporate designs that drop connectivity directly onto core switches or a layer that has more direct connectivity to data center interconnects. This puts the in-band network in closer proximity to monitoring systems that may reside in an adjacent data center compared to connectivity models that feature firewall-based segmentation.



**Figure 13-2** *ACI In-Band Management Connectivity in a Typical Network*

Figure 13-3 provides a reminder that L3Outs are not the only connectivity mechanism available to administrators who want to extend in-band connectivity out of an ACI fabric. From a tenancy perspective, the in-band management VRF provides the majority of the bells and whistles of user VRFs. In this example, the out-of-the-box bridge domain `inb` has been configured to serve as the default gateway for in-band traffic via the address `10.100.1.1/24`. The in-band management EPG, shown here generically with the text `EPG In-Band`, has been assigned a dynamic IP address pool that

ranges from 10.100.1.2 to 10.100.1.200. With a dynamic IP address pool assignment, ACI is able to automatically allocate IP addresses to APICs and ACI switches. The in-band management EPG also requires an encapsulation to enable end-to-end in-band connectivity across all ACI switches and APICs. After all, in-band connectivity flows over the same physical interfaces as the infrastructure VLAN and data plane traffic. To enable the desired communication with management stations, administrators can assign contracts to this in-band EPG. To place servers in the mgmt tenant, administrators need to create additional EPGs. The same in-band EPG used for switch and APIC connectivity cannot be reused for this purpose. These new server-mapped EPGs *can* be associated with the inb bridge domain. Hence, a number of syslog and SNMP servers have been depicted in the same subnet as the inb bridge domain. To establish connectivity between these types of server EPGs and the in-band EPG, a contract with the scope VRF is sufficient. For cross-tenant communication through route leaking, the contract scope should be set to Global.



**Figure 13-3 Merging Data Plane Traffic Directly into the In-Band Management VRF**

### Note

The implementation of route leaking is beyond the scope of the Implementing Cisco Application Centric Infrastructure DCACI 300-620 exam and is therefore not covered here. Moreover, this book has already addressed the implementation of EPG extensions and L2Outs, and this chapter does not provide additional coverage.



Although in-band management is powerful, not everything about in-band management in ACI is rosy. In-band management does depend on access policies, tenant-level policies, and sometimes cross-tenant contracts to enable connectivity. You would be correct in assuming that in-band management has greater dependency on fabric configuration and that it is also more susceptible to software defects than OOB management—and therein lies its fundamental challenge.

## **Deploying In-Band and OOB Management Side by Side**

Because ACI out-of-band management is optimized for recoverability and ACI in-band management can enable more direct connectivity into the data plane, a marriage between the two can be a beautiful thing.



In instances in which a company decides to deploy in-band and OOB management alongside each other, Cisco recommends that you not combine static and dynamic IP addressing with one another. Rely on either dynamic IP addressing or static IP addressing for both in-band and out-of-band communication.

## **Configuring In-Band Management**

It is difficult to come up with a specific set of steps that applies to all in-band management deployments. But the

following steps provide an understanding of the high-level thinking necessary to effectively deploy in-band management:



### **Step 1. Configure access policies for APIC in-band interfaces:**

**Access Policies:** When implementing in-band management, it is important to understand that access policies need to be configured and assigned to all switch ports to which APICs connect. At the very least, these access policies need to enable LLDP, trunk the infrastructure VLAN, and allow one additional encapsulation over APIC-facing switch ports.

### **Step 2. Configure the in-band management bridge domain:**

**Bridge Domain:** If ACI will handle routing for the in-band management subnet, ensure at the very least that unicast routing has been enabled and that a default gateway has been configured on the bridge domain. Note that multiple subnets can be assigned to the in-band bridge domain.

### **Step 3. Configure in-band management IP addressing:**

**IP Addressing:** Deploy either static IP addressing or dynamic IP addressing to switches and APICs, based on the intended target state design.

### **Step 4. Optionally extend the in-band network out of the fabric:**

**L3Out Configuration:** If in-band subnets are expected to be advertised across the network, administrators can deploy an L3Out in the mgmt tenant. For basic connectivity requirements, L2Outs and EPG extensions usually suffice.

## **Optionally set up additional connectivity**

**Step 5.within the fabric:** If management and monitoring endpoints reside in the fabric and there is a preference to do route leaking or place new endpoints directly in the inb VRF, this can be easily done at this point to enable communication with in-band ACI addresses.

## **Step 6.Whitelist desired connectivity to and from in-band EPGs:**

**in-band EPGs:** While each ACI node should be able to ping its default gateway within the in-band VRF without issue, all inter-EPG communication requires contract enforcement. Configure and assign the contracts necessary to enable your desired communication.

**Step 7.Evaluate APIC connectivity preferences:** By default, APICs prefer in-band connectivity for outbound traffic unless specific routes are available in the APIC routing table. If, for example, out-of-band connectivity is being decommissioned in favor of in-band management due to lack of out-of-band infrastructure, it makes sense to ensure that once in-band communication has been fully tested, the APIC Connectivity Preferences setting is set to inband.

## **Configuring Access Policies for APIC In-Band Interfaces**

The first step in implementing in-band management is to assign access policies to APIC-facing switch ports. To do so, configuration of a dedicated AAEP is highly advised. In the configuration shown in [Figure 13-4](#), VLAN 260 is the infrastructure VLAN, and VLAN 266 will be used to encapsulate in-band traffic. Notice that a new VLAN pool has

been created for this configuration. The Enable Infrastructure VLAN checkbox should be enabled for this AAEP.



Create Attachable Access Entity Profile

STEP 1 > Profile      1. Profile      2. Association To Interfaces

Name: APIC-Inband-AAEP  
Description: optional

Enable Infrastructure VLAN:

Domains (VMM, Physical or External) To Be Associated To Interfaces:

Domain Profile	Encapsulation
Physical Domain - Inband-Physical	from:vlan-260 to:vlan-260 from:vlan-266 to:vlan-266

EPG DEPLOYMENT (All Selected EPGs will be deployed on all the interfaces associated.)

Application EPGs	Encap	Primary Encap	Mode

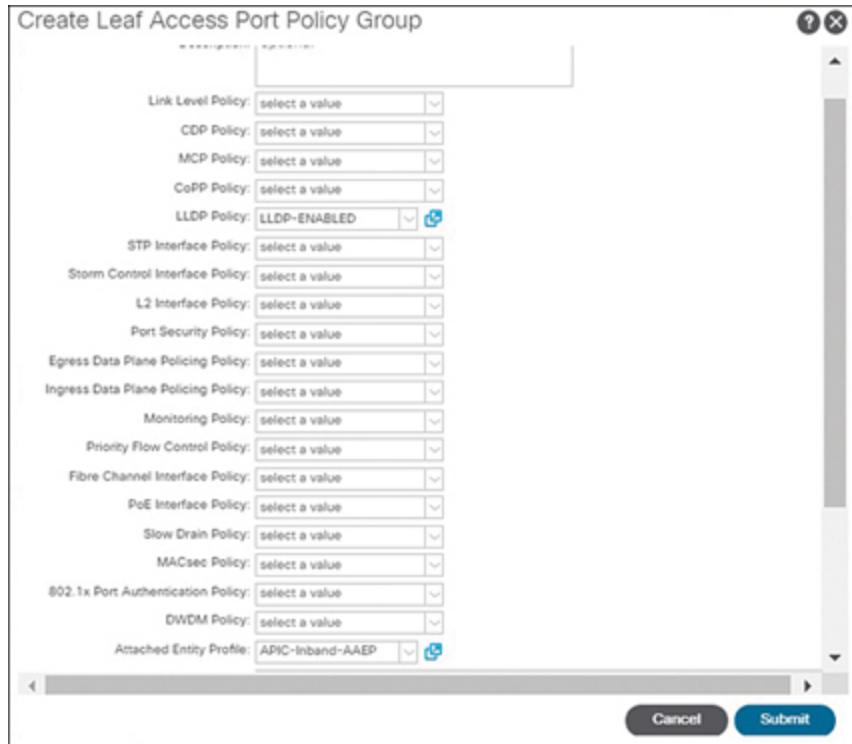
Previous      Cancel      Next

The screenshot shows the 'Create Attachable Access Entity Profile' dialog. The 'Name' field is set to 'APIC-Inband-AAEP'. The 'Enable Infrastructure VLAN' checkbox is checked. In the 'Domains To Be Associated To Interfaces' section, there is one entry for 'Physical Domain - Inband-Physical' with the encapsulation 'from:vlan-260 to:vlan-260' and 'from:vlan-266 to:vlan-266'. The 'EPG DEPLOYMENT' section is currently empty. At the bottom, there are 'Previous', 'Cancel', and 'Next' buttons, with 'Next' being highlighted.

**Figure 13-4** *In-Band Management AAEP Configuration*

Next, an interface policy group that enables LLDP and includes the newly created AAEP needs to be configured. [Figure 13-5](#) indicates that an access port policy group (non-aggregated) should be used for this configuration.





**Figure 13-5** *In-Band Management Interface Policy Group Configuration*

Once the interface policy group has been configured, you can assign it to all switch ports that connect to APICs. [Figure 13-6](#) shows the interface policy group being assigned to ports 1, 2, and 3 on a single-module switch.

Create Access Port Selector

Name: APICs-Inband

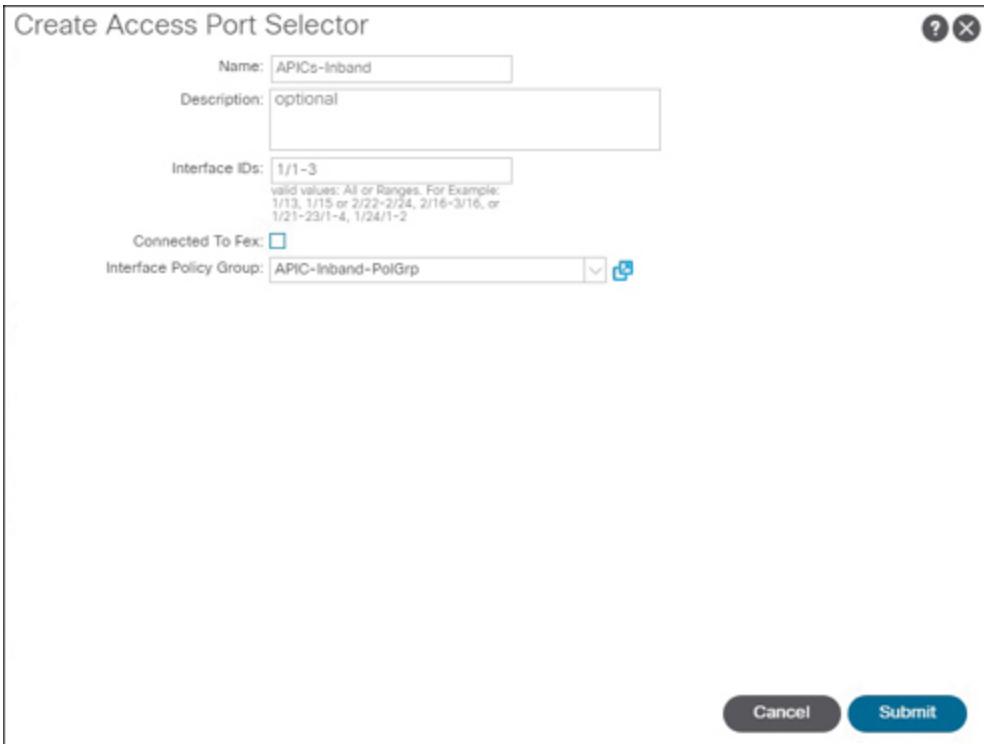
Description: optional

Interface IDs: 1/1-3  
Valid values: All or Ranges. For Example:  
1/13, 1/15 or 2/22-2/24, 2/16-3/16, or  
1/21-23/1-4, 1/24/1-2

Connected To Fex:

Interface Policy Group: APIC-Inband-PolGrp

Cancel Submit



**Figure 13-6** Assigning New Interface Policy Group to Switch Ports

## Configuring the In-Band Management Bridge Domain

An important step in in-band management configuration is to assign a default gateway to the inb bridge domain. [Figure 13-7](#) shows that configuration of the in-band management bridge domain is trivial and like any other bridge domain configuration. If the in-band subnet is not expected to be advertised out the VRF instance, you can keep the Scope setting Private to VRF.

Create Subnet

Gateway IP:  address/mask

Treat as virtual IP address:

Make this IP address primary:

Scope:  Private to VRF  
 Advertised Externally  
 Shared between VRFs

Description:

Subnet Control:  No Default SVI Gateway  
 Querier IP

L3 Out for Route Profile:

Route Profile:

ND RA Prefix policy:

**Figure 13-7** Assigning a Subnet and Gateway for In-Band Management

## Configuring In-Band Management IP Addressing

As noted earlier in this chapter, two options exist for IP addressing. [Figure 13-8](#) shows that the process for configuring static in-band management addresses is almost identical to the process for configuring static out-of-band IP addresses covered in [Chapter 3](#), minus the exception that ACI does not come preconfigured with an in-band management EPG.



Create Static Node Management Addresses

Node Range:  From  To

Config:  Out-Of-Band Addresses  In-Band Addresses

In-Band IP Addresses

In-Band Management EPG:

In-Band IPV4 Address:   
address/mask

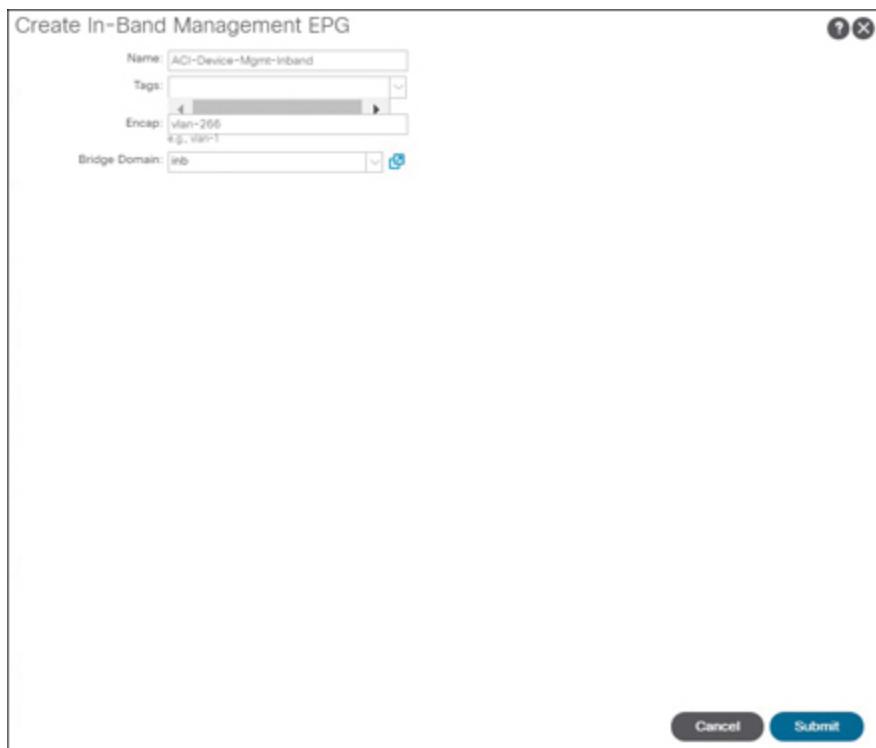
In-Band IPV4 Gateway:

In-Band IPV6 Address:   
address/mask

In-Band IPV6 Gateway:

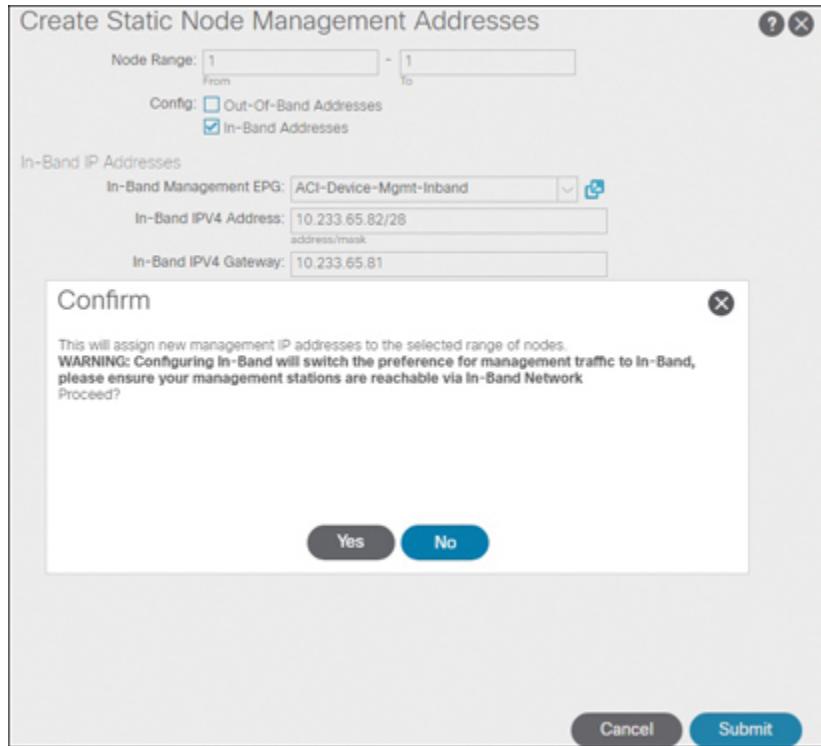
**Figure 13-8** Configuring Static Addresses for In-Band Management

To enable creation of static in-band node management addresses, configure an in-band management EPG either through the pull-down menu on the Create Static Node Management Addresses page or by navigating to Tenants > mgmt > Node Management EPGs. [Figure 13-9](#) shows that the Create In-Band Management EPG page requires assignment of an encapsulation to the in-band EPG.



**Figure 13-9** Configuring an In-Band Management EPG

Creation of an in-band EPG and assignment of IP addresses to it can trigger a warning message, as shown in [Figure 13-10](#), suggesting that a communication outage can be expected. If out-of-band connectivity is fully in place, manually toggling the APIC Connectivity Preferences parameter to ooband can prevent an outage and render the warning invalid.



**Figure 13-10** *Confirming Warning Message*

Static IP assignments, whether in-band or out-of-band, can be validated by navigating to **Tenants > mgmt > Node Management Addresses > Static Node Management Addresses**, as shown in [Figure 13-11](#).

The screenshot shows the Cisco Application Policy Infrastructure Controller (APIC) web interface. At the top, there's a navigation bar with tabs for System, Tenants (which is selected), Fabric, Virtual Networking, L4-L7 Services, Admin, Operations, Apps, and Integrations. Below the tabs, there are links for ALL TENANTS, Add Tenant, and Tenant Search. On the right side of the header are icons for search, cloud, notifications, and settings.

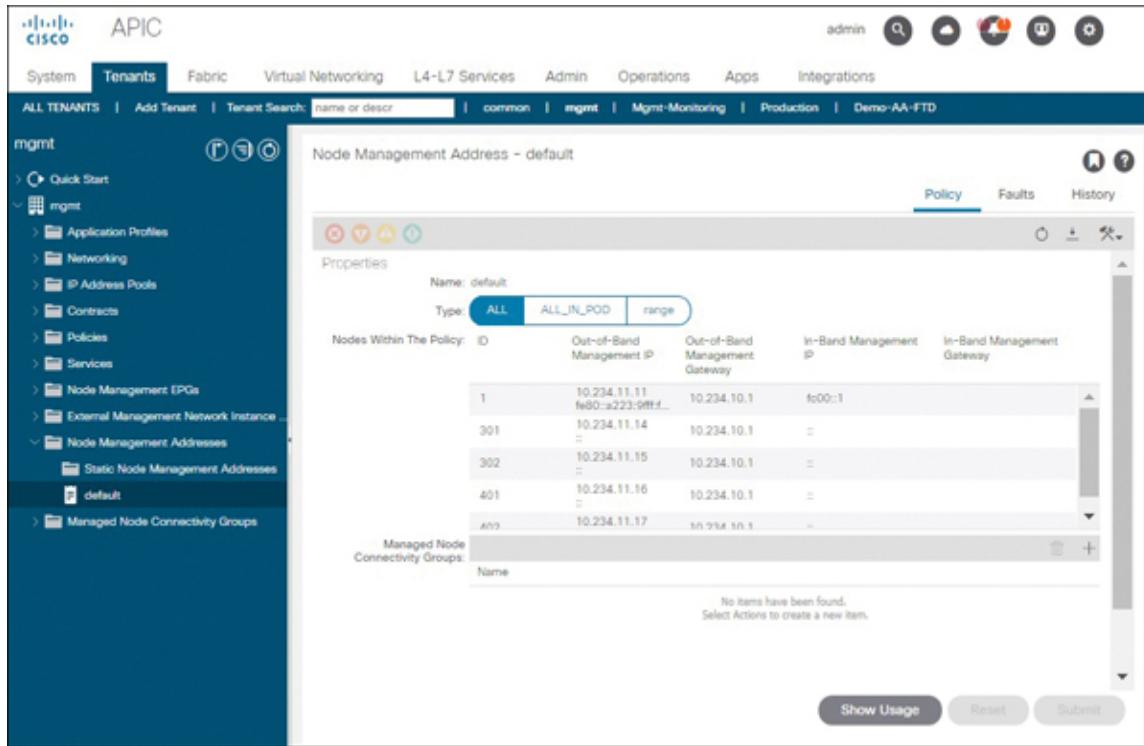
The main content area is titled "Node Management Addresses". There are two tabs: "Node Management Addresses" and "Static Node Management Addresses", with the latter being the active tab. The table below has columns for Node ID, Type, EPG, IPV4 Address, IPV4 Gateway, IPV6 Address, and IPV6 Gateway. The data is as follows:

Node ID	Type	EPG	IPV4 Address	IPV4 Gateway	IPV6 Address	IPV6 Gateway
pod-1/node-1	In-Band	ACI-Device-Mgmt-Inband	10.233.65.82/28	10.233.65.81	..	..
pod-1/node-301	In-Band	ACI-Device-Mgmt-Inband	10.233.65.83/28	10.233.65.81	..	..
pod-1/node-302	In-Band	ACI-Device-Mgmt-Inband	10.233.65.84/28	10.233.65.81	..	..
pod-1/node-401	In-Band	ACI-Device-Mgmt-Inband	10.233.65.85/28	10.233.65.81	..	..
pod-1/node-402	In-Band	ACI-Device-Mgmt-Inband	10.233.65.86/28	10.233.65.81	..	..

**Figure 13-11** Verifying Static IP Address Assignments

Now, let's take a look at how you might go about configuring dynamic IP addressing instead. [Figure 13-12](#) shows the default node management policy. In this example, because all nodes are part of a single fabric and will be placed into a single subnet, you select ALL in the Type field and add all nodes to the default node management policy. The page then displays any configured out-of-band or in-band IP addressing in the fabric. From [Figure 13-12](#), it is clear that all nodes have been assigned static out-of-band addresses.





**Figure 13-12** Configuring the Default Node Management Policy

## Note

If an administrator has deleted the default node management policy that comes with ACI, a new one can be created and automatically gets assigned to all selected nodes.

Assuming that the intention here is to migrate to dynamic IP addressing in both the in-band and out-of-band networks, you would need to navigate to the Tools menu and select Add a Managed Node Connectivity Group. The majority of configuration objects shown in [Figure 13-13](#) should be familiar by now. An IP address pool needs to be configured for dynamic IP addressing to work, so you need to select Create IP Address Pool and configure an IP pool for in-band connectivity.

**Key Topic**

Create Managed Node Connectivity Group

Name: Inband-Nodes

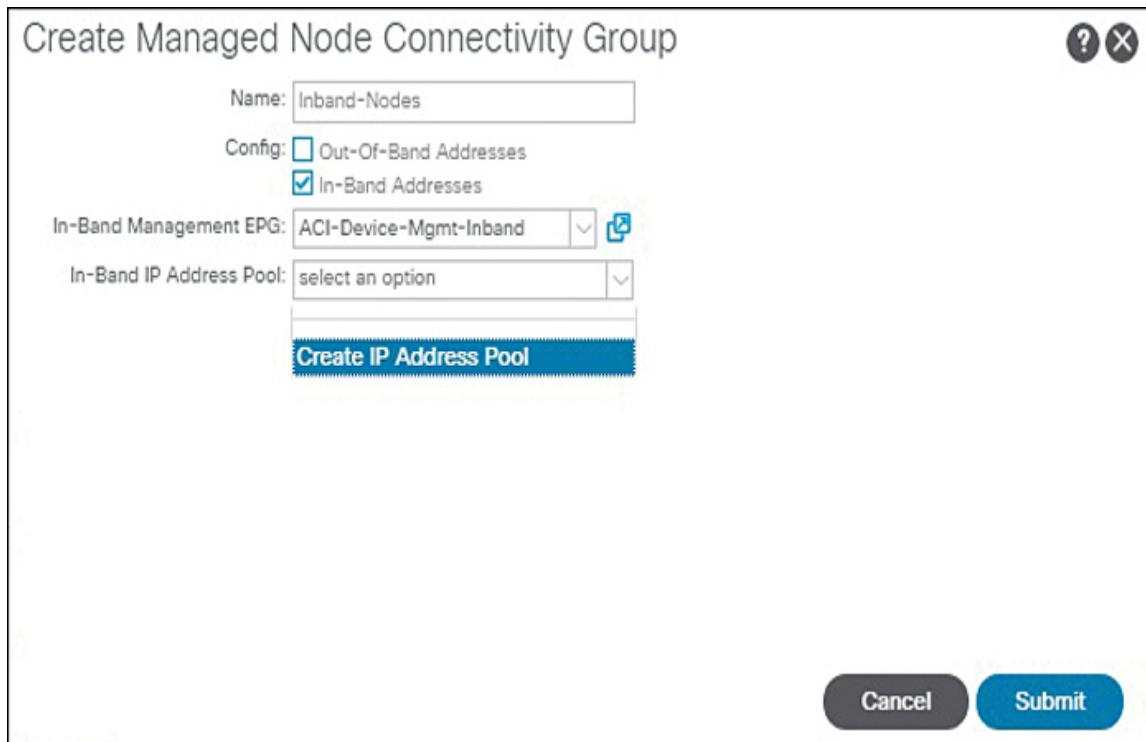
Config:  Out-Of-Band Addresses  In-Band Addresses

In-Band Management EPG: ACI-Device-Mgmt-Inband

In-Band IP Address Pool: select an option

**Create IP Address Pool**

Cancel Submit



**Figure 13-13** Configuring a Node Management Connectivity Group

As indicated in [Figure 13-14](#), the pool configuration enables you to determine the range of addresses to use for ACI nodes that are part of the node management address policy configuration. It also enables you to select the default gateway for nodes assigned to the specified pool. By default, ICMP probes are executed against the specified default gateway to enable default route failover for APICs. This behavior can be changed through the Skip Gateway Validation checkbox. After selecting your desired settings, click Submit on each of the open windows to execute the changes.

**Key Topic**

Create IP Address Pool

Name:

Skip Gateway Validation:

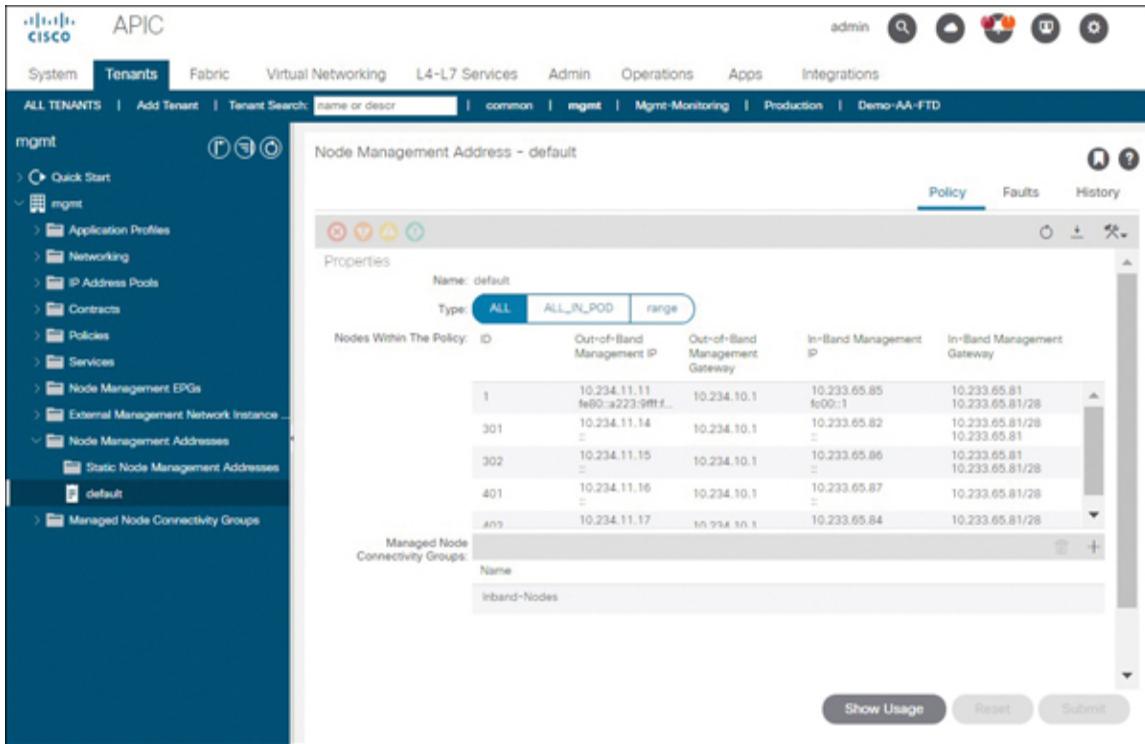
Gateway Address:

Address Ranges:

From	To
10.233.65.82	10.233.65.87

**Figure 13-14** Creating a Dynamic IP Address Pool

Figure 13-15 shows that the node management policy just configured can also be used to verify IP assignment from the dynamic pool.



**Figure 13-15** Verifying Dynamic IP Address Assignments

After in-band IP addresses have been assigned to all nodes within the fabric, the modes should be able to ping one another within the mgmt:inb VRF instance. As explained earlier and indicated in [Example 13-1](#), the APIC is not VRF aware and does not need a VRF reference to ping in-band addresses.

### Example 13-1 Testing Connectivity in the In-Band Network

[Click here to view code image](#)

```
APIC1# ping 10.233.65.81 -c 2
PING 10.233.65.81 (10.233.65.81) 56(84) bytes of data.
64 bytes from 10.233.65.81: icmp_seq=1 ttl=64 time=0.140 ms
64 bytes from 10.233.65.81: icmp_seq=2 ttl=64 time=0.152 ms

--- 10.233.65.81 ping statistics ---
2 packets transmitted, 2 received, 0% packet loss, time
```

```
2010ms
```

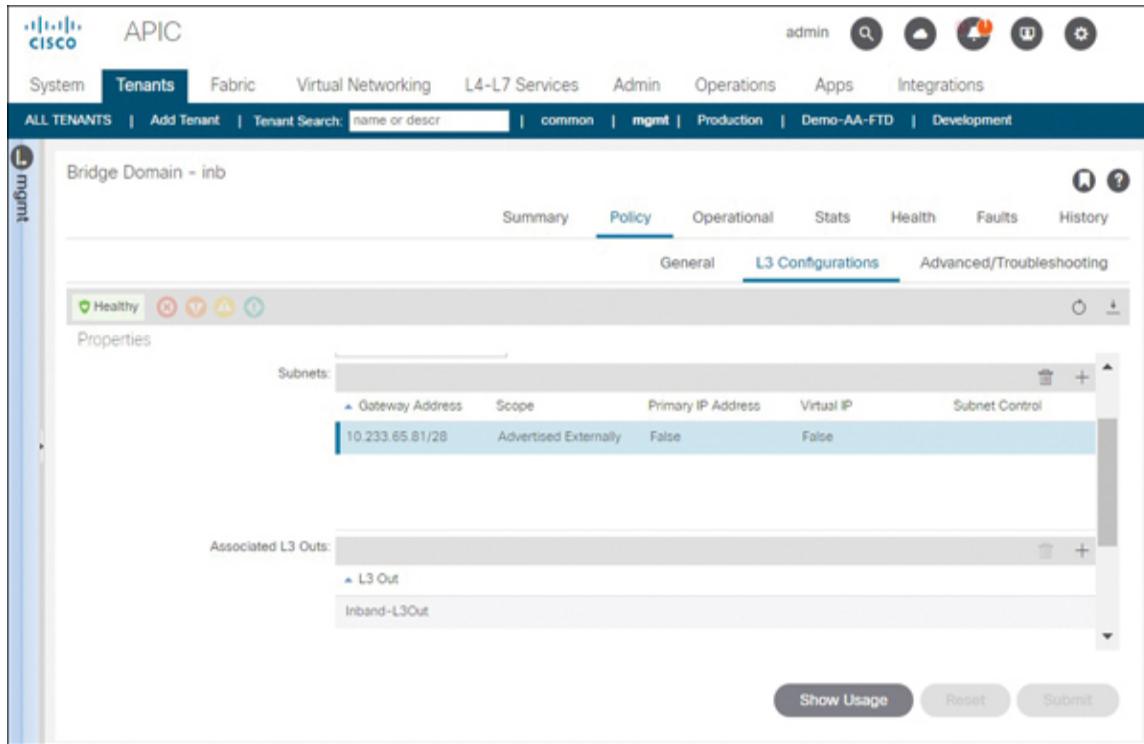
```
rtt min/avg/max/mdev = 0.140/0.155/0.173/0.013 ms
```

### Note

Pinging spine switches from other ACI nodes can sometimes lead to unexpected results. The inability to ping spine nodes should not be interpreted as meaning that the in-band IP addressing configuration has failed.

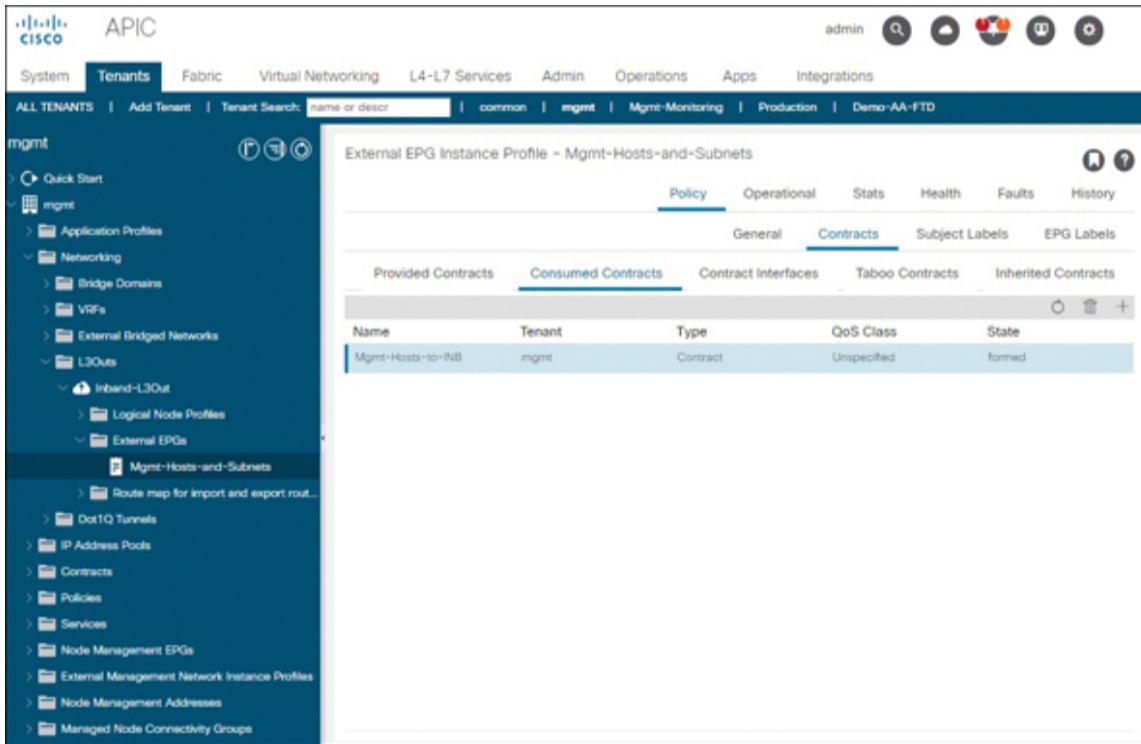
## Optionally Extending the In-Band Network Out of the Fabric

As noted earlier in this chapter, for the most part, the in-band VRF functions just like any other VRF. In anticipation of advertising the in-band management subnets over an L3Out, you can update the `inb` bridge domain scope setting to `Advertised Externally`, as shown in [Figure 13-16](#).



**Figure 13-16** Enabling External Advertisement of the inb Bridge Domain

Figure 13-17 shows that a typical L3Out has been created within the mgmt tenant. Dynamic routing is not necessarily needed, but if it is deployed, remember that neighbor adjacencies do not form until at least one external EPG is created.



**Figure 13-17** Defining at Least One External EPG on L3Out

After adjacencies form over the L3Out, the in-band subnet should be advertised out of the fabric, and ACI should learn routes advertised from outside neighbors.

## Optionally Setting Up Additional Connectivity

At this point, you might decide to establish cross-tenant communications with the in-band subnet or create server EPGs within the in-band environment. There is nothing new to cover on this subject. All the relevant constructs, except for route leaking, function the same as in any other tenant.

## Whitelisting Desired Connectivity to and from an In-Band EPG

Even though end-to-end routing may have been established, ACI still expects contracts to enable management communication. [Figure 13-17](#) shows a contract applied in the consumed direction to the external EPG. Another look at the details of the contract allocation suggests that in-band contracts are very similar to any other contracts. [Figure 13-18](#) shows the same contract from [Figure 13-17](#) configured with the scope VRF and a single subject.

The screenshot shows the 'Create Contract' dialog box. The fields are as follows:

- Name: Mgmt-Hosts-to-INB
- Alias: (empty)
- Scope: VRF
- QoS Class: Unspecified
- Target DSCP: Unspecified
- Description: optional
- Tags: (empty)
- Subjects:
  - Allowed-Mgmt-Protocols-to-INB

At the bottom are 'Cancel' and 'Submit' buttons.

**Figure 13-18** Contract Enabling Access to In-Band Management via Basic Protocols

The subject created for this contract has both the Apply Both Directions and Reverse Filter Ports checkboxes enabled, as shown in [Figure 13-19](#). This ensures that return traffic from ports specified by associated filters from ACI back to any management endpoints is also allowed by the contract. While a review of the filter name implies that in-band management will only be used for SSH and HTTPS access in this particular fabric, you could very easily add

new filters to the depicted subject or, alternatively, add additional protocols to the permitted filter.

The dialog box is titled "Create Contract Subject". It contains three main sections: "General", "Filter Chain", and "Filters".

**General Section:**

- Name: Allowed-Mgmt-Protocols-to-INB
- Alias: (empty)
- Description: optional
- Target DSCP: Unspecified
- Apply Both Directions:
- Reverse Filter Ports:
- Wan SLA Policy: select an option

**Filter Chain Section:**

- L4-L7 Service Graph: select an option
- QoS Priority: (empty)

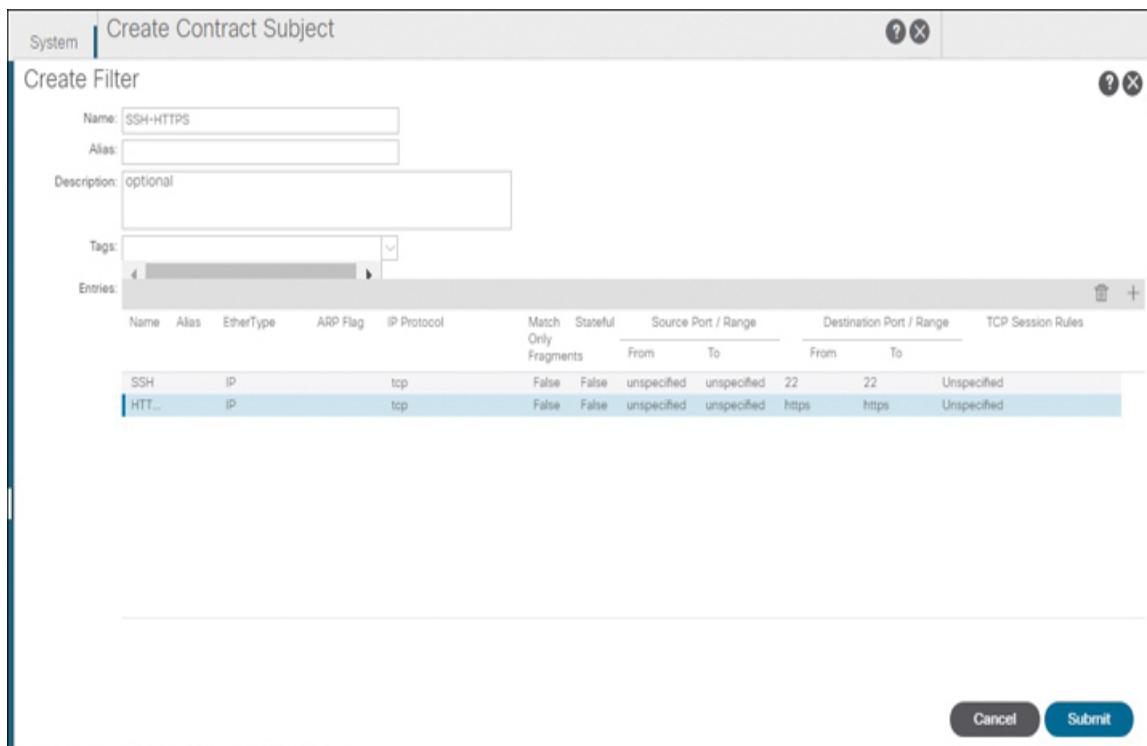
**Filters Section:**

Name	Directives	Action	Priority
mgmt/SSH-HTTPS	none	permit	default

Buttons at the bottom: Cancel, OK.

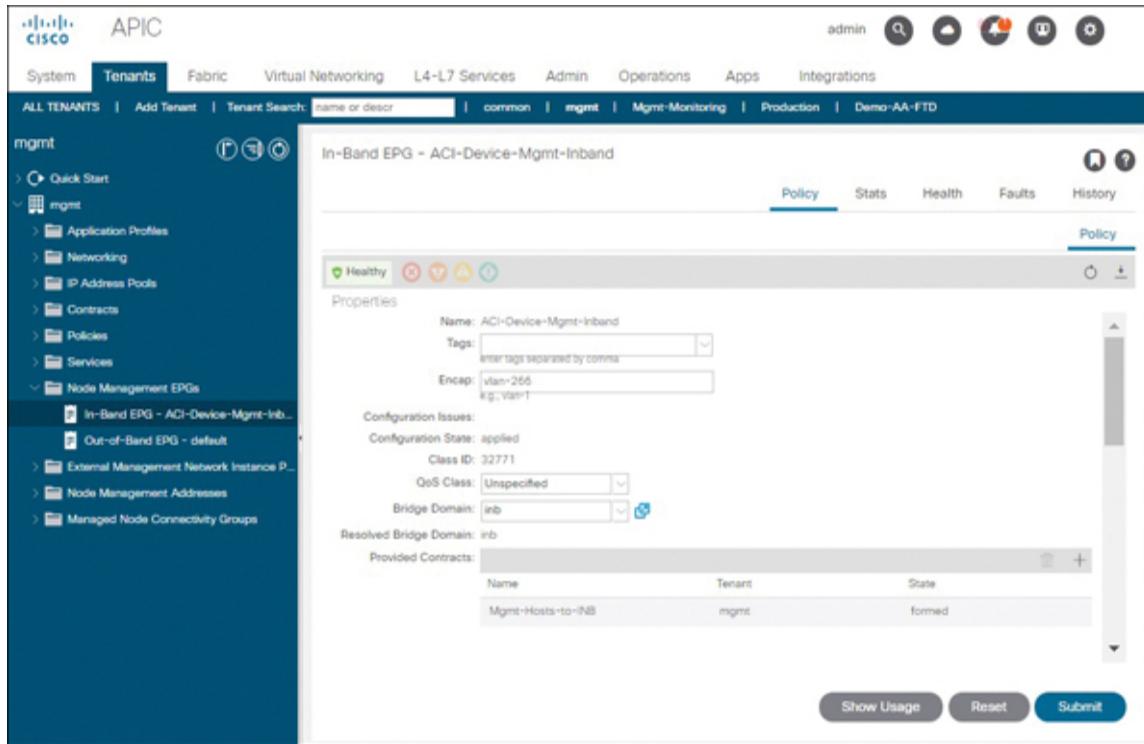
**Figure 13-19** Subject Configuration for a Sample Management Contract

Figure 13-20 shows the filter associated with this contract. This filter classifies SSH and HTTPS traffic as interesting traffic.



**Figure 13-20** Sample Filter That Classifies SSH and HTTPS Traffic

Finally, the configuration cannot be considered complete until the contract has been allocated to the in-band EPG. Notice that [Figure 13-21](#) adds the contract to the in-band EPG in the provided direction. This is because management hosts typically initiate SSH or HTTPS connections.



**Figure 13-21** *Assigning Contracts to In-Band Management EPGs*

## Evaluating APIC Connectivity Preferences



Remember that APICs are essentially servers and not routers. Servers do not segment routing tables via VRF instances. Therefore, a solution is needed to tweak outbound default routing metrics. Under System > System Settings > APIC Connectivity Preferences, you can do just that, ensuring that a specific management connection is preferred unless a more specific route has been added to the APIC routing table.

In [Example 13-2](#), the default route out of the APIC interface called oobmgmt has the lower and more preferred metric, whereas the in-band management connection over bond0.266 has a higher metric. This is because the APIC Connectivity Preferences parameter is set to ooband.

### **Example 13-2 APIC Default Route Metric Lower Toward the OOB Network**

[Click here to view code image](#)

```
APIC1# bash
admin@APIC1:~> route -n
Kernel IP routing table
Destination      Gateway      Genmask      Flags      Metric
Ref    Use    Iface
0.0.0.0          10.234.10.1  0.0.0.0      UG        16
0      0      oobmgmt
0.0.0.0          10.233.65.81  0.0.0.0      UG        32
0      0      bond0.266
10.233.60.0     10.233.60.30  255.255.252.0  UG        0
0      0      bond0.260
10.233.60.30   0.0.0.0       255.255.255.255  UH        0
0      0      bond0.260
10.233.65.80   0.0.0.0       255.255.255.240  U         0
0      0      bond0.266
10.233.65.81   0.0.0.0       255.255.255.255  UH        0
0      0      bond0.266
10.234.10.0     0.0.0.0       255.255.254.0     U         0
0      0      oobmgmt
169.254.1.0     0.0.0.0       255.255.255.0     U         0
0      0      teplo-1
169.254.254.0   0.0.0.0       255.255.255.0     U         0
0      0      lxcbr0
172.17.0.0      0.0.0.0       255.255.0.0      U         0
0      0      docker0
```

After the APIC Connectivity Preferences parameter is set to inband, the APIC updates its routing table to prefer the in-band connection unless a more specific route is available in the routing table (see [Example 13-3](#)).

### **Example 13-3 APIC Default Route Metric Lower in the In-Band Network**

[Click here to view code image](#)

```
admin@APIC1:~> route -n
Kernel IP routing table
Destination      Gateway        Genmask        Flags Metric
Ref    Use   Iface
0.0.0.0          10.233.65.81  0.0.0.0       UG      8
0      0     bond0.266
0.0.0.0          10.234.10.1   0.0.0.0       UG      16
0      0     oobmgmt
10.233.60.0     10.233.60.30  255.255.252.0  UG      0
0      0     bond0.260
10.233.60.30   0.0.0.0       255.255.255.255 UH      0
0      0     bond0.260
10.233.65.80   0.0.0.0       255.255.255.240  U       0
0      0     bond0.266
10.233.65.81   0.0.0.0       255.255.255.255 UH      0
0      0     bond0.266
10.234.10.0    0.0.0.0       255.255.254.0   U       0
0      0     oobmgmt
169.254.1.0    0.0.0.0       255.255.255.0   U       0
0      0     teplo-1
169.254.254.0  0.0.0.0       255.255.255.0   U       0
0      0     lxcbr0
172.17.0.0     0.0.0.0       255.255.0.0    U       0
0      0     docker0
```

---

## Note

In quite a few ACI configurations that reference management and monitoring stations such as SNMP, syslog, and AAA servers, you are asked to enter a management EPG. This is like configuring source interfaces for such services on NX-OS switches. If the APIC Connectivity Preferences parameter has been toggled to inband and an out-of-band management EPG is selected for outbound communications with management and monitoring systems, the selection of the management EPG can be invalidated. On the flipside, if the APIC Connectivity Preferences parameter has been toggled to ooband, an in-band management EPG can be selected in ACI configurations for connectivity toward such servers, as long as the APIC routing tables have routes pointing to the subnets in which these servers reside.

## Out-of-Band Management Contracts in Review

Recall from [Chapter 3](#) that contracts for out-of-band management are defined under a different folder than are in-band contracts. [Figure 13-22](#) shows that OOB contracts are configured under the Out-of-Band Contracts folder. These contracts are not interchangeable with standard contracts used for in-band management.

Create Out-Of-Band Contract

Name: Mgmt-Protocols-to-OOB

Scope: VRF

QoS Class: Unspecified

Description: optional

Subjects:

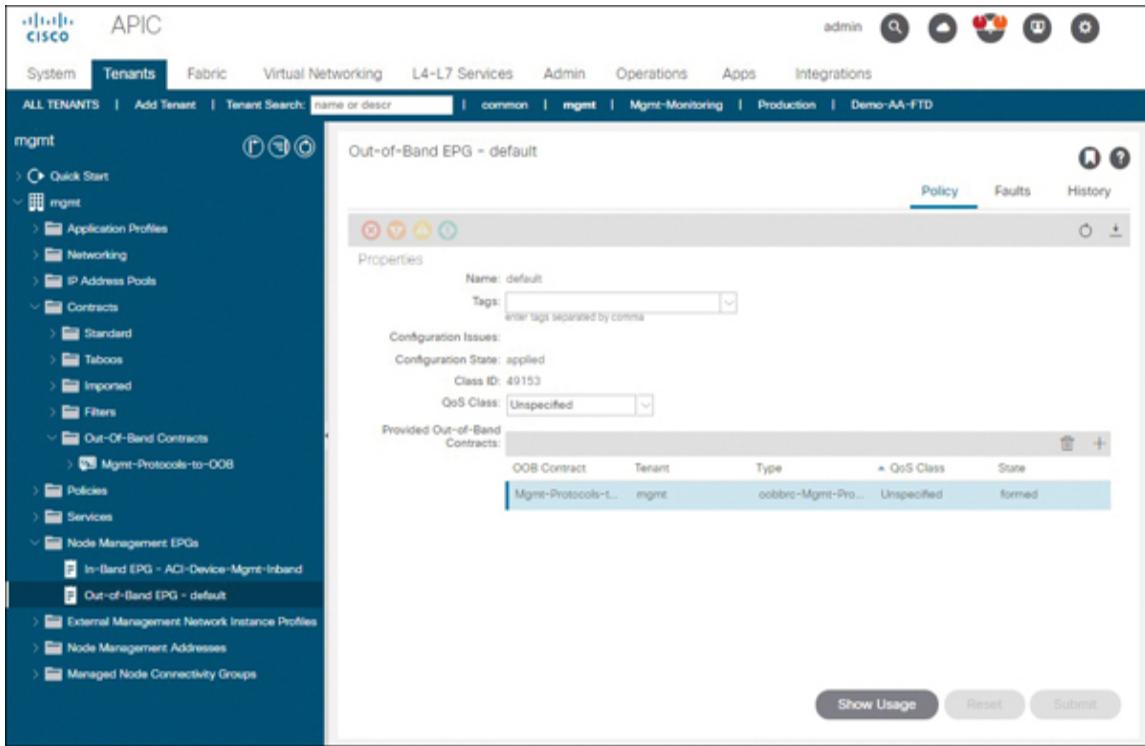
Name	Description
Allowed-Mgmt-Protocols	

Cancel    Submit

The screenshot displays a configuration interface for creating an Out-Of-Band (OOB) contract. The top header reads "Create Out-Of-Band Contract". Below it are four input fields: "Name" (Mgmt-Protocols-to-OOB), "Scope" (VRF), "QoS Class" (Unspecified), and "Description" (optional). A "Subjects" section contains a table with one row, "Allowed-Mgmt-Protocols", under the "Name" column. The "Description" column is empty. At the bottom are two buttons: "Cancel" and "Submit".

**Figure 13-22** Configuration of an Out-of-Band Contract

Just as in the case of in-band connectivity, contracts need to be applied to the relevant management EPG. [Figure 13-23](#) shows a minor difference in the application of contracts to out-of-band EPGs compared to in-band EPGs. Contracts can only be applied to OOB EPGs in the provided direction. Another difference that is not shown, however, is that OOB contracts do not support logging of traffic matching a particular filter chain.



**Figure 13-23 Applying Out-of-Band Contracts Is Possible Only in the Provided Direction**

Finally, another difference is that restricting management access to a subset of subnets over OOB interfaces is configured at the external management network interface profile level. This is the same object to which OOB contracts need to be applied in the consumed direction (see [Figure 13-24](#)). The concept of an external management network interface profile does not exist for in-band connectivity.

Create External Management Network Instance Profile

Name: Mgmt-Hosts-and-Subnets

Tags: enter tags separated by comma

Consumed Out-of-Band Contracts:

Out-of-Band Contract	QoS Class
Mgmt-Protocols-to-OOB	Unspecified

Subnets:

IP
10.0.222.0/24
10.233.65.0/24

Cancel    Submit

The screenshot shows a configuration dialog for creating a network instance profile. At the top, it says 'Create External Management Network Instance Profile'. Below that, there's a 'Name' field containing 'Mgmt-Hosts-and-Subnets' and a 'Tags' field with a placeholder 'enter tags separated by comma'. Under 'Consumed Out-of-Band Contracts', there's a table with one row: 'Out-of-Band Contract' (Mgmt-Protocols-to-OOB) and 'QoS Class' (Unspecified). Below that is a 'Subnets' section with a table showing two IP ranges: '10.0.222.0/24' and '10.233.65.0/24'. At the bottom right are 'Cancel' and 'Submit' buttons.

**Figure 13-24** Object Representing Management Hosts Connecting to OOB Interfaces

Chapter 14, “Monitoring ACI Using Syslog and SNMP,” includes further examples of contracts for the out-of-band network.

## Exam Preparation Tasks

As mentioned in the section “How to Use This Book” in the Introduction, you have a couple of choices for exam preparation: Chapter 17, “Final Preparation,” and the exam simulation questions on the companion website.

## Review All Key Topics

Review the most important topics in this chapter, noted with the Key Topic icon in the outer margin of the page. Table 13-

[2](#) lists these key topics and the page number on which each is found.



**Table 13-2** Key Topics for [Chapter 13](#)

<b>Key Topic Element</b>	<b>Description</b>	<b>Page Number</b>
Paragraph	Describes the fundamental benefit of out-of-band management in ACI	<a href="#">464</a>
Paragraph	Summarizes the key use cases for OOB management in ACI	<a href="#">465</a>
Paragraph	Describes the fundamental benefit of in-band management in ACI	<a href="#">465</a>
Paragraph	Outlines the primary drawback of in-band management in ACI	<a href="#">467</a>
Paragraph	Describes the Cisco recommendation for IP addressing when implementing in-band and	<a href="#">467</a>

	OOB management side by side	
List	Lists the steps necessary to effectively deploy in-band management in ACI	467
Figure 13-4	Illustrates the AAEP configurations required for switch-facing APIC LOM ports in in-band management deployments	468
Figure 13-5	Illustrates the interface policy group configurations needed for switch-facing APIC LOM ports in in-band management deployments	469
Figure 13-8	Demonstrates the configuration of static IP addressing for in-band management	470
Figure 13-12	Shows a configuration example for the default node management policy	472
Figure 13-13	Demonstrates how to configure a node management connectivity group	473
Figure 13-14	Shows how to configure a dynamic IP address pool	473

Paragraph	Describes the significance of the APIC Connectivity Preferences setting	478
-----------	---	-----

## Memory Tables

There are no memory tables or lists in this chapter.

## Define Key Terms

There are no key terms for this chapter.

# Chapter 14

## Monitoring ACI Using Syslog and SNMP

This chapter covers the following topics:

**Understanding System Messages:** This section explains what system messages are and how ACI structures them.

**Forwarding System Messages to Syslog Servers:** This section walks through the process of configuring ACI for system message forwarding to syslog servers.

**Using SNMP in ACI:** This section provides a basic understanding of SNMP as well as SNMP capabilities supported in ACI.

**Configuring ACI for SNMP:** This section explores the process of configuring ACI to respond to SNMP read queries and forward system messages as traps.

This chapter covers the following exam topic:

- 5.2 Utilize syslog and snmp services

[Chapter 4](#), “Exploring ACI,” touches on how ACI provides feedback to users through faults, event logs, health scores, and audit logs. This chapter details how ACI structures faults, events, and other log records into system messages

for forwarding to syslog servers and SNMP managing systems.

This chapter also covers the ACI configurations necessary to allow remote devices to poll ACI switches and APICs via SNMP read queries.

This chapter revisits the MIM and reinforces your understanding of the ACI object model, enabling you to develop highly customized monitoring policies.

## **“Do I Know This Already?” Quiz**

The “Do I Know This Already?” quiz allows you to assess whether you should read this entire chapter thoroughly or jump to the “Exam Preparation Tasks” section. If you are in doubt about your answers to these questions or your own assessment of your knowledge of the topics, read the entire chapter. [Table 14-1](#) lists the major headings in this chapter and their corresponding “Do I Know This Already?” quiz questions. You can find the answers in [Appendix A, “Answers to the ‘Do I Know This Already?’ Questions.”](#)

**Table 14-1** “Do I Know This Already?” Section-to-Question Mapping

Foundation Topics Section	Questions
Understanding System Messages	1, 2
Forwarding System Messages to Syslog Servers	3-6

Using SNMP in ACI	7, 8
Configuring ACI for SNMP	9, 10

## Caution

The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark that question as wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

- 1.** True or false: ACI can only generate system messages for faults.
  - a.** True
  - b.** False
  
- 2.** Which system message format provides the most granularity?
  - a.** Syslog
  - b.** NX-OS
  - c.** ACI
  - d.** Fault
  
- 3.** Which set of monitoring policies needs to be modified to enable syslog forwarding for server port failures, assuming that no custom monitoring policies will be enforced?

- a. Fabric > Fabric Policies > Policies > Monitoring > Common Policies
    - b. Fabric > Fabric Policies > Policies > Monitoring > default
    - c. Fabric > Access Policies > Policies > Monitoring > default
    - d. Tenant > common > Policies > Monitoring > default
- 4. An ACI administrator has been tasked with troubleshooting why a syslog server is not receiving syslog messages from ACI. Which of the following could be possible causes for syslog traffic not arriving on the syslog server? (Choose all that apply.)
  - a. The syslog server's IP address has not been included as an external syslog data collector.
  - b. ACI supports only syslog Version 9.
  - c. A contract permitting all traffic into and out of the management EPG has been assigned, but an explicit contract for syslog has not been created.
  - d. No syslog sources have been defined.
- 5. An administrator has configured a new syslog destination monitoring group and wants to ensure that system messages for a specific tenant are redirected to the new syslog destination group. After configuring a new set of syslog monitoring policies and a syslog source within the monitoring policies, no log forwarding seems to be taking place for problems that occur within the tenant. What went wrong? (Choose the best answer.)
  - a. Custom monitoring policies cannot be configured in ACI.
  - b. The monitoring policies need to be assigned to the tenant in the Policy tab.

- c. A firewall in the path toward the syslog server is blocking port 161.
  - d. A load balancer has been defined as an L4-L7 device within the tenant.
- 6. Which APIC CLI command signals an APIC or ACI switch to generate a system message for transmission toward a syslog server?
  - a. **moquery**
  - b. **grep**
  - c. **logit**
  - d. **syslog**
- 7. An ACI administrator has configured ACI for SNMP and is trying to use MIBs to execute a configuration change in the fabric but is unable to do so. What is the most likely reason for this?
  - a. ACI has been configured for SNMPv3, but an SNMPv3 user has not been configured.
  - b. ACI contract issues need to be resolved.
  - c. A firewall is blocking communication between the SNMP manager and ACI.
  - d. SNMP write commands are not supported in ACI.
- 8. What are feasible reasons an ACI configuration to a specific SNMP manager may be failing or suboptimal? (Choose all that apply.)
  - a. Contracts have been defined on the management EPG, but they are not specific enough.
  - b. The out-of-band EPG is used, but static node management addresses have not been configured.
  - c. The administrator has configured 12 trap destinations.

- d. Client entries in the SNMP policy include the IP address of the SNMP manager but not the subnet default gateway.
- 9.** An ACI administrator is testing an SNMP configuration that has just been deployed. SNMP read queries related to fabric policies and access policies tend to work fine. However, the administrator is unable to perform read queries related to EIGRP for a specific tenant. What is a possible problem?
- a. The SNMP pod policy has not been configured correctly.
  - b. The SNMP contracts associated to the management EPG should be deleted.
  - c. SNMP contexts need to be associated with VRF instances in the tenant.
  - d. SNMP read queries are not supported on ACI leaf switches, and EIGRP runs on switches.
- 10.** With default SNMP manager settings, which IP protocols and ports should be configured in a contract to allow full SNMP read queries and traps?
- a. UDP 161 and TCP 162
  - b. UDP 161 and UDP 514
  - c. UDP 161 and UDP 162
  - d. TCP 161 and UDP 514

## Foundation Topics

### Understanding System Messages

A **system message** is a specially formatted message that typically contains a subset of information about a fault, an

event, or another log record in the fabric. When certain faults and events occur, ACI can forward system messages to the console, external syslog servers, external SNMP servers, call home servers, or, alternatively, log the system message locally on the devices that generate the message.



Both APICs and ACI switches can generate system messages. ACI can structure system messages in a format similar to NX-OS switches, or it can structure system messages in a new ACI-oriented format. The default preferred system message format is the ACI structure because it includes more data.

System messages following the NX-OS format use the following syntax:

[Click here to view code image](#)

timestamp Nexus: FACILITY-SEVERITY-MNEMONIC: Message-text

This syntax includes the following variables and fixed strings:

- **timestamp:** The year, month, date, and time of day of system message generation
- **Nexus:** A fixed string
- **FACILITY:** Two or more uppercase letters that indicate the affected hardware device, protocol, or module of the system software
- **SEVERITY:** A single-digit code from 0 to 7 that reflects the severity of the condition, as outlined in [Table 14-2](#)
- **MNEMONIC:** A code that uniquely identifies the error message

- **Message-text:** A description of the problem encountered or the event that occurred



**Table 14-2 NX-OS System Message Severity Levels**

Severity Level (NX-OS)	Severity Level ITU Level Description (ACI)	—	—
0	Emergency	—	System unusable
1	Alert	Critical	Immediate action required
2	Critical	Major	Critical condition
3	Error	Minor	Error condition
4	Warning	Warning	Warning condition
5	Notification	Cleared	Normal but significant condition
6	Informational	—	Informational message only

7	Debugging	—	Messages that appear during debugging only

Generally, system messages reflecting a lower severity level tend to point to conditions or events that have a higher potential for impacting an ACI fabric or endpoints within the fabric.

[Example 14-1](#) shows a system message generated by an ACI switch. The syslog severity 5 has been logged.

**Example 14-1** *Sample System Message for a Fault in NX-OS Format*

[Click here to view code image](#)

```
2014 Jan 25 21:42:07 Nexus: ETHPORT-5-IF_DOWN_ADMIN_DOWN:
Interface Ethernet3/1 is down (Administratively down))
```

System messages following the ACI format use the following syntax:

[Click here to view code image](#)

```
timestamp host %LOG_LOCALn-severity-SYSTEM_MSG [code][lifecycle state][rule][severity text][DN of affected MO]
```

Message-text

This syntax includes the following variables and fixed strings:

- **timestamp:** The year, month, date, and time of day of system message generation

- **host:** The IP address or hostname of the device that generated the system message
- **%LOG\_LOCALn:** A single-digit code from 0 to 7 that reflects the local facility of the message and is sometimes used to sort received messages
- **severity:** A single-digit code from 1 to 5 following the ITU perceived severity values described in RFC 5674 reflecting the severity of the condition
- **SYSTEM\_MSG:** A fixed string
- **code:** The unique fault or event code associated with the message
- **lifecycle state:** The current state in the fault lifecycle; output is omitted when a message is being generated for an event
- **rule:** The action or condition that caused the event, such as a component failure or a threshold crossing
- **severity text:** The text translation of the numeric severity values (for example, major), as indicated in [Table 14-3](#)
- **DN of affected MO:** The distinguished name (DN) of the managed object (MO) affected by the fault condition or event
- **Message-text:** A description of the problem encountered or the event that occurred



**Table 14-3** ACI System Message Severity Levels

Description
-------------

Lev el	ITU Level (ACI)	
0	—	System is unusable
1	Critical	Immediate action required
2	Major	Critical condition
3	Minor	Error condition
4	Warning	Warning condition
5	Cleared	Normal but significant condition
6	Informational	Informational message only
7	—	Messages that appear during debugging only

[Example 14-2](#) shows a sample system message generated in ACI format for a fault on a switch with node ID 102,

indicating that the APIC named apic1 has lost connectivity to the switch.

**Example 14-2** *Sample ACI-Structured System Message for a Fault*

[Click here to view code image](#)

```
July 22 22:45:28 apic1 %LOG_LOCAL0-2-SYSTEM_MSG [F0110]
[soaking][node-failed]
[critical][topology/pod-1/node-102/fault-F0110]
Node 102 not reachable. unknown
```

[Example 14-3](#) presents a sample system message generated in ACI format for an event logging the transition of a fabric port on node 101 to an optimal status from the view of apic1.

**Example 14-3** *Sample ACI-Structured System Message for an Event*

[Click here to view code image](#)

```
July 22 22:45:27 apic1 %LOG_LOCAL0-6-SYSTEM_MSG [E4208219]
[link-state-change][info]
[subj-[topology/pod-1/lnkcnt-1/lnk-101-1-1-to-1-1-3]/rec-
4294968577]
Link State of Fabric Link is set to ok
```

## Forwarding System Messages to Syslog Servers

Administrators can configure ACI to forward system messages to external syslog servers using the following simple process:



- Step 1.** Apply necessary contracts to allow syslog forwarding.
- Step 2.** Configure syslog monitoring destination groups.
- Step 3.** Configure syslog sources for desired monitoring policies.
- Step 4.** Verify syslog forwarding to desired syslog servers.

These steps do not need to be completed in the order presented. However, it does make sense to create syslog monitoring destination groups first because syslog sources need to reference the syslog monitoring destination groups. Furthermore, there is little reason to verify until the configuration has been completed.

Although the process seems very straightforward, it tends to challenge engineers on their understanding of monitoring policies and the ACI object hierarchy because syslog sources for all desired monitoring policies need to be configured for proper syslog forwarding to take place. Therefore, the following sections provide some details to make these steps clearer.

## Apply Necessary Contracts to Allow Syslog Forwarding

[Chapter 13, “Implementing Management,”](#) outlines the procedure for applying contracts to the out-of-band ACI subnet by limiting the sources for communication to all private IP addresses.

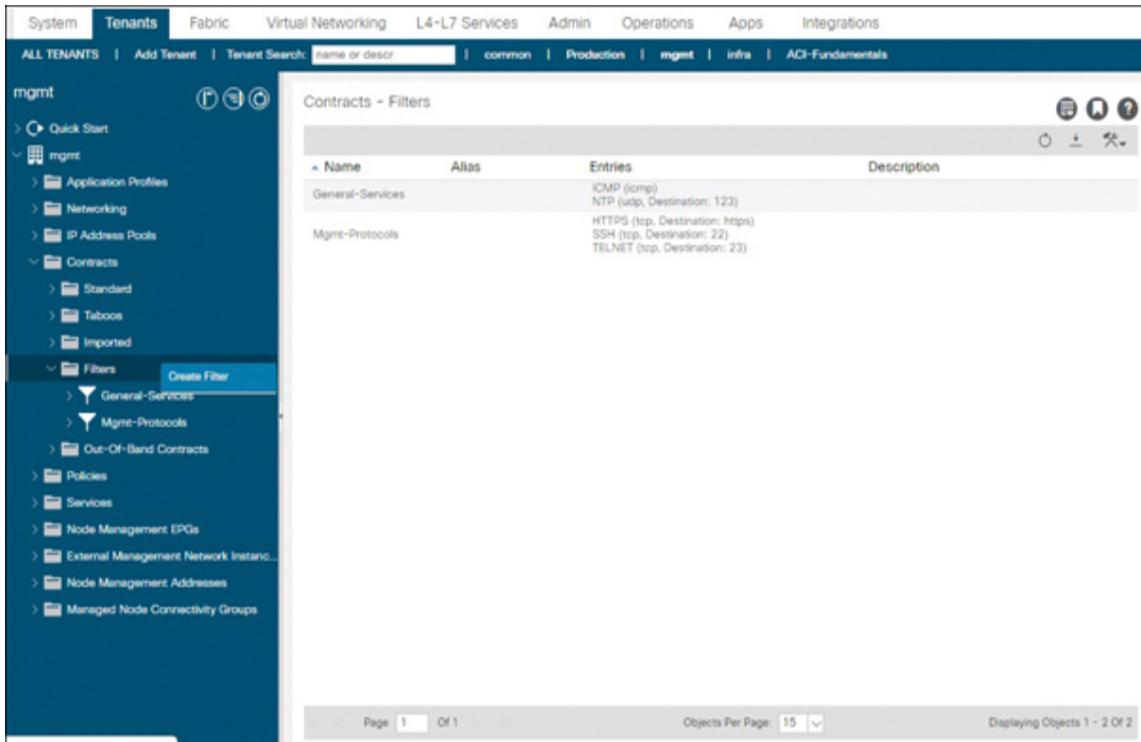
This chapter outlines the process for enabling a contract that allows communication with syslog servers. However,

unlike the procedure described in [Chapter 13](#), this procedure attempts to further lock down access to UDP port 514 by specifying the syslog server IP addresses in a dedicated external management instance profile.

### Note

Some earlier ACI code revisions did not require syslog ports to be opened via contracts for out-of-band communication. However, the application of an out-of-band contract to limit communication has always been considered a Cisco best practice. Where in-band management is used to forward traffic to syslog servers, explicit contract enforcement *is* required.

The first step in creating out-of-band contracts for syslog is to create a filter to classify syslog traffic. [Figure 14-1](#) shows how users can launch the filter creation wizard by navigating to the mgmt tenant, opening the Contracts folder, right-clicking on Filters, and selecting Create Filter.



**Figure 14-1** Launching the Create Filter Wizard from the *mgmt* Tenant

In the Create Filter window, enter a filter name, create a filter entry, and click Submit. [Figure 14-2](#) shows two filters, named syslog-dest and syslog-src, that classify all port 514 traffic in both the source and destination port directions. By default, syslog servers listen on port 514, but they can also communicate outbound via port 514, even though they do not send acknowledgments for client messages.

Create Filter

Name: Syslog

Alias:

Description: optional

Tags:

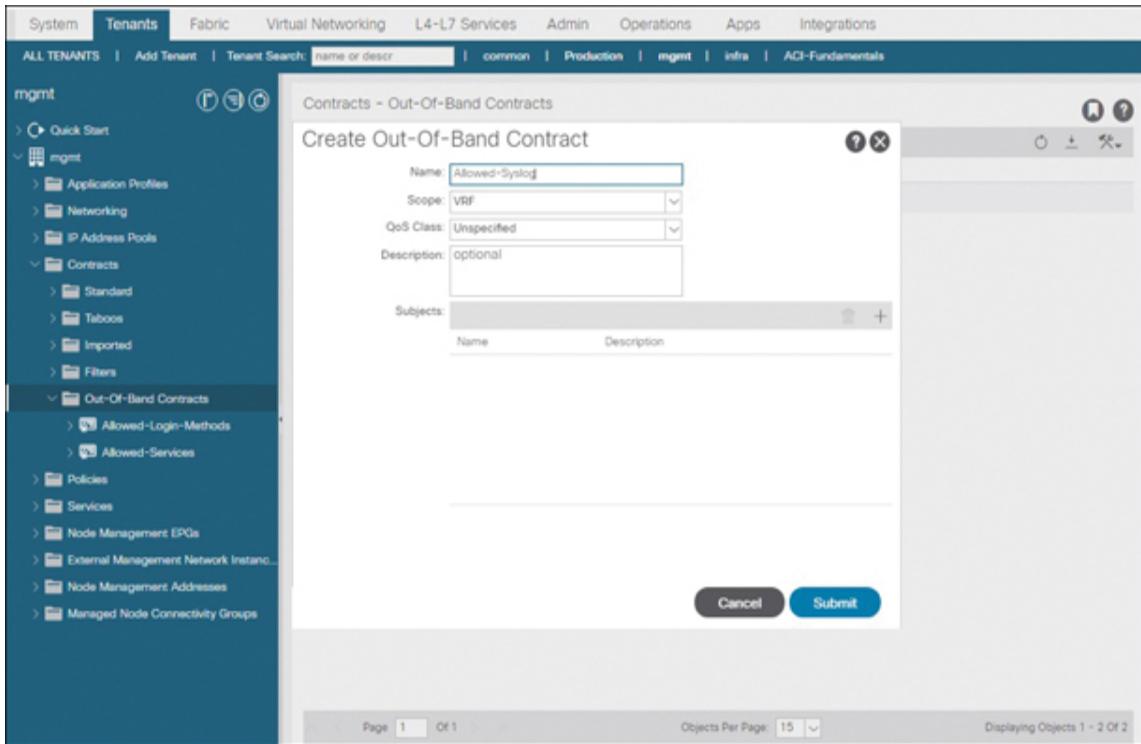
Entries:

Name	Alias	EtherType	ARP Flag	IP Protocol	Match Only Fragments	Stateful	Source Port / Range		Destination Port / Range		TCP Session Rules
							From	To	From	To	
Syslog-dest	IP			udp	False	False	unspecified	unspecified	514	514	
Syslog-src	IP			udp	False	False	514	514	unspecified	unspecified	

Cancel Submit

**Figure 14-2** Creating a Filter for Syslog Traffic

Next, create an out-of-band contract by navigating to Contracts, right-clicking Out-of-Band Contracts, and selecting Create Out-of-Band Contracts. Select a name for the new contract and click on the + sign to create a subject for syslog traffic (see [Figure 14-3](#)).



**Figure 14-3** Launching the Subject Creation Page from the Create Out-of-Band Contract Window

Name the new subject, click the + sign next to Filters, associate the syslog filter with the subject, and then click Update and OK (see [Figure 14-4](#)).

Create Contract Subject

Name: Permit-Syslog

Description: optional

Filter Chain

**Filters**

Name: mgmt/Syslog

Service Graph: select an option

PRIORITY

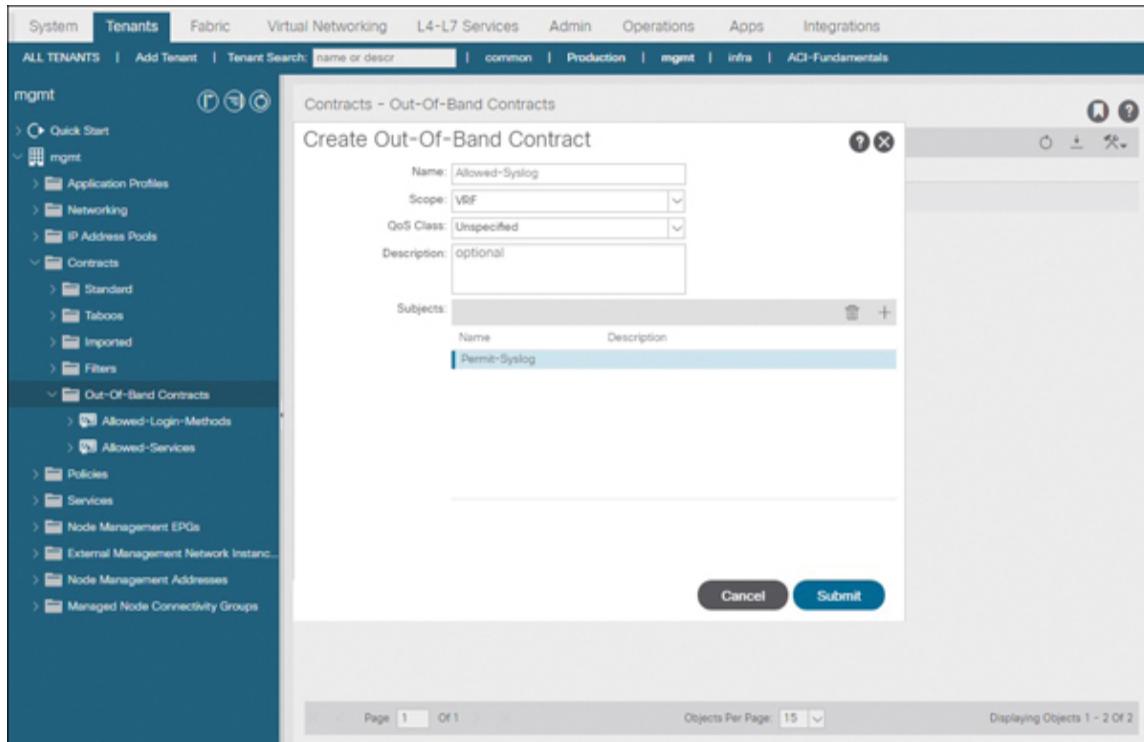
QoS:

Cancel OK

The screenshot shows the 'Create Contract Subject' dialog box. In the 'Name' field, 'Permit-Syslog' is entered. The 'Description' field contains the placeholder 'optional'. Below this, the 'Filter Chain' section is expanded, showing a sub-dialog titled 'Filters' with a single item 'mgmt/Syslog' listed. There are 'Update' and 'Cancel' buttons for this sub-dialog. To the right of the filter chain, there's a 'L4-L7 SERVICE GRAPH' section with a dropdown menu set to 'select an option', and a 'PRIORITY' section with a QoS input field. At the bottom right of the main dialog are 'Cancel' and 'OK' buttons.

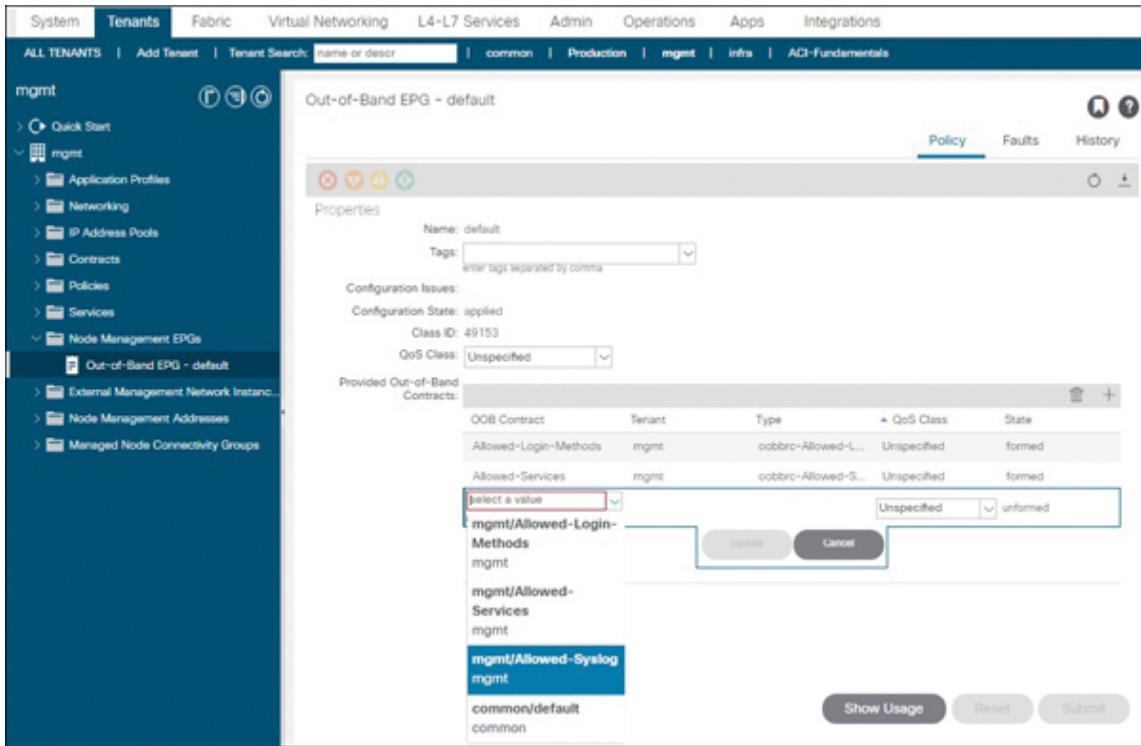
**Figure 14-4** *Associating a Filter with a New Subject*

Once the subject has been created, it appears on the Create Out-of-Band Contract page. Click Submit to finalize creation of the new contract (see [Figure 14-5](#)).



**Figure 14-5** Finalizing the Creation of the Out-of-Band Contract

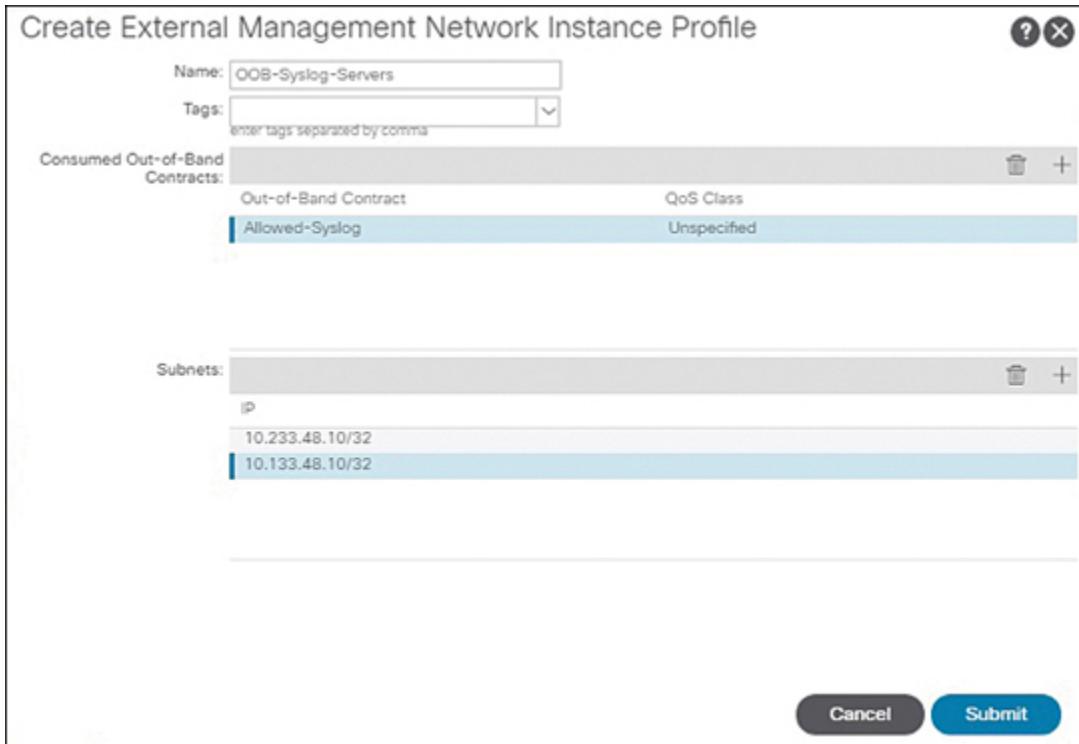
After you create the contract, you can associate the contract with the out-of-band EPG. To do so, navigate to mgmt, click Node Management EPGs, and select Out-of-Band EPG - default. Then select the newly created contract in the Provided Out-of-Band Contracts view and click Update (see [Figure 14-6](#)). Finally, click Submit to update the OOB EPG.



**Figure 14-6** *Associating a Contract with an Out-of-Band EPG*

Next, a new external EPG needs to be created to represent the syslog servers. Remember that creating a specific external EPG for syslog servers is not a requirement, but this section demonstrates how it can be done. In this example, the syslog contract will be consumed by the new external EPG.

To create an external EPG, which is called an external management network instance profile within the management tenant, navigate to mgmt, right-click External Management Network Instance Profiles, and select Create External Management Network Instance Profiles. Enter the external EPG name, the new contract, and host routes for the syslog server, as shown in [Figure 14-7](#), and click Submit to complete the contract creation process.



**Figure 14-7** Creating a New External Management Network Instance Profile

### Note

It is not necessary to create a separate external management network instance profile for each traffic type. A single external EPG can classify all traffic arriving on OOB interfaces. However, if you need different management or monitoring endpoints to have different levels of access to the fabric, this is most feasibly done through specificity and by defining separate EPGs that each classify different sets of external endpoints.

### Note

So how would contract enforcement be different if in-band management were used? The contract itself

would be the same, but it would need to be configured under **mgmt > Contracts > Standard** instead.

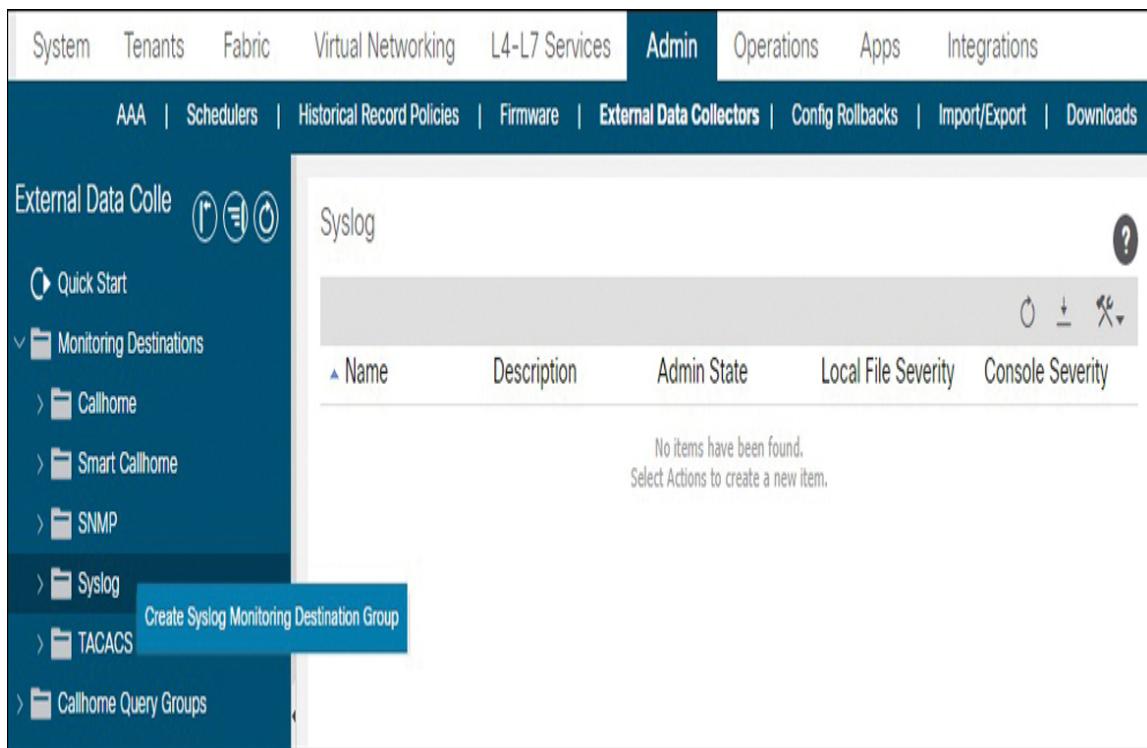
Because in-band management allows contract allocation in both the provider and consumer directions, the contract direction becomes more important. With the contract filters shown in [Figure 14-2](#), the contract would need to be applied on an external EPG associated with an L3Out in the provider direction and on the in-band management EPG in the consumer direction.

## Configuring Syslog Monitoring Destination Groups

Syslog **monitoring destination groups** are used to define and group syslog servers together. The reason for grouping syslog servers together is that a company may want one set of servers to be used for monitoring system messages involving fabric and access policies. A specific department within the company or perhaps a customer or partner may request that system messages related to a specific tenant they manage be forwarded to a dedicated set of servers they also manage.



To configure a syslog monitoring destination group, navigate to the Admin menu, select External Data Collectors, and under Monitoring Destinations, right-click on the Syslog folder and select Create Syslog Monitoring Destination Group, as shown in [Figure 14-8](#).



**Figure 14-8** Navigating to the *Syslog Monitoring Destination Group Wizard*

Enter a name for the syslog monitoring destination group, select the desired system message format, determine the granularity of timestamps and whether they should be precise to the millisecond, and toggle Admin State to enabled. If system messages should also be logged on the device console and within the local file systems, set Admin State for these parameters to enabled and determine the desired message logging severity (see [Figure 14-9](#)). Note that setting console logging severity to a value that generates an exorbitant number of messages is discouraged due to the potential for impact on device CPU utilization.

Create Syslog Monitoring Destination Group

STEP 1 > Profile

1. Profile    2. Remote Destinations

Name: Syslog-Servers

Description: optional

Format: aci  nxos

Show MilliSeconds in Timestamp:

Admin State: enabled

Local File Destination

Admin State: enabled

Severity: Information

Console Destination

Admin State: enabled

Severity: alerts

Previous    Cancel    Next

**Figure 14-9** Configuring a Syslog Monitoring Destination Group, Step 1

Next, create an entry for each syslog server in the destination group, as demonstrated in [Figure 14-10](#). Note that a different system message severity level, destination port, forwarding facility, and source interface can be selected for each syslog server, if desired.

Create Syslog Remote Destination

Host Name/IP:

Name:

Admin State:  disabled  enabled

Severity:

Port:

Forwarding Facility:

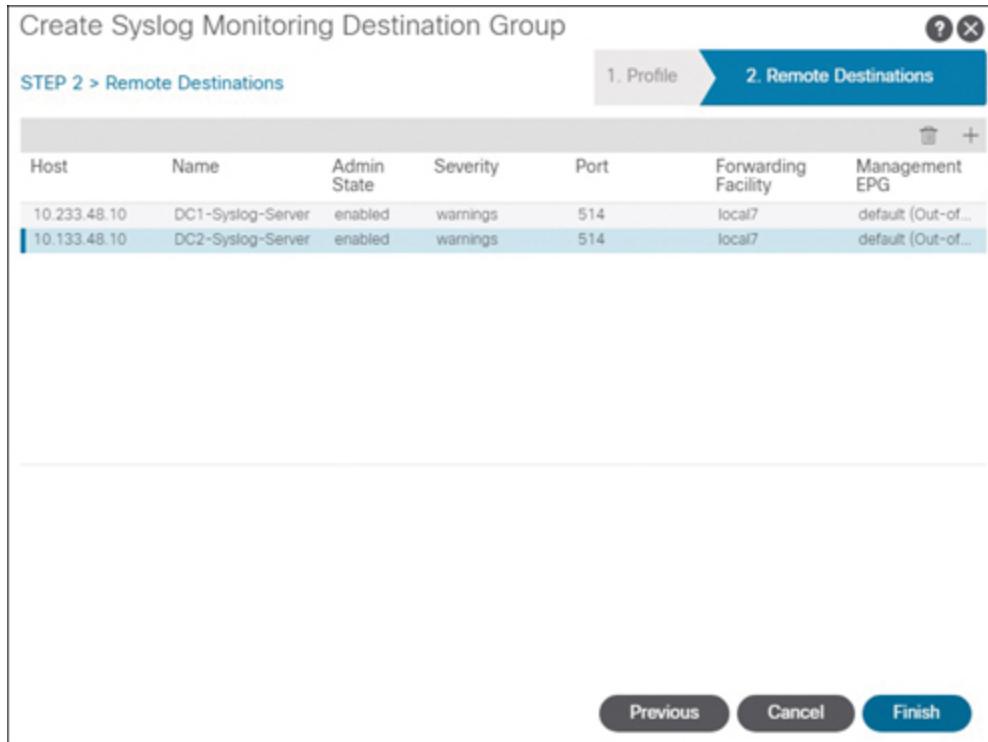
Management EPG:  

Cancel 



**Figure 14-10** Creating a Syslog Remote Destination, Step 2

When all syslog servers that are members of the destination group have been defined, as shown in [Figure 14-11](#), click Finish to move on to configuring syslog sources.



**Figure 14-11** Verifying the Syslog Remote Destinations Configured in a Syslog Monitoring Destination Group

## Configuring Syslog Sources for Desired Monitoring Policies

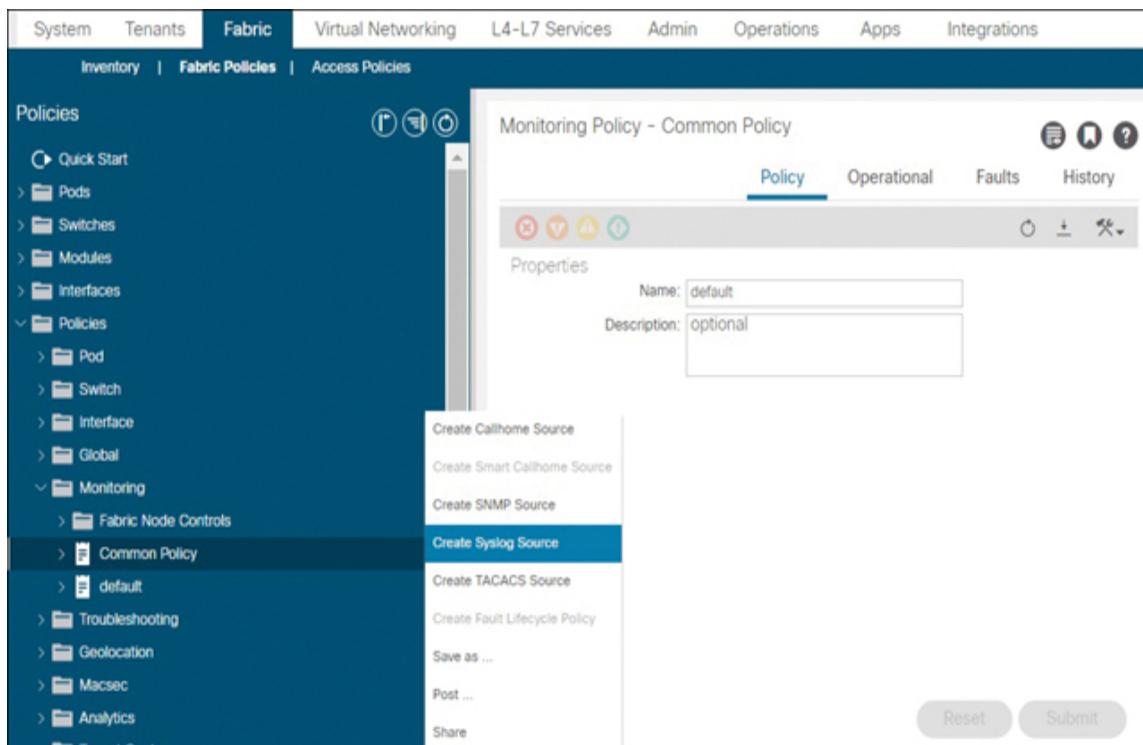
After any planned syslog monitoring destination groups have been configured, monitoring sources need to be configured to point to the destination groups.

As a review, there are four default monitoring groups in ACI. Each addresses a different set of faults, events, and logs. The default monitoring groups can be found in the following locations as of ACI Release 4.2(1):

**Key Topic**

- Fabric > Fabric Policies > Policies > Monitoring > Common Policies
- Fabric > Fabric Policies > Policies > Monitoring > default
- Fabric > Access Policies > Policies > Monitoring > default
- Tenant > common > Policies > Monitoring > default

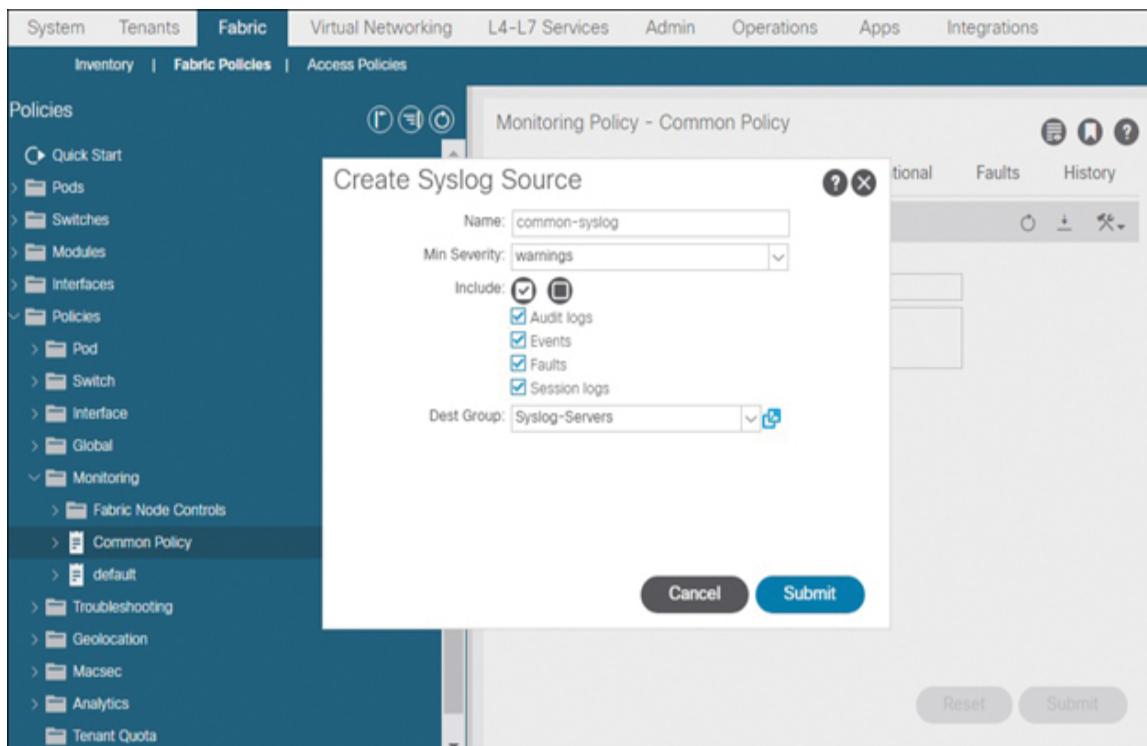
By right-clicking on a desired monitoring policy and selecting Create Syslog Source, as shown in [Figure 14-12](#), you can trigger the launch of the Create Syslog Source wizard.



**Figure 14-12** Launching the Create Syslog Source Wizard

In the Create Syslog Source screen, enter a name for the syslog source object. Then select the types of events, faults,

and log records to convert to system messages by selecting the appropriate Include checkboxes. Determine the minimum message severity level for which ACI should forward messages and then select a syslog monitoring destination group from the Dest Group pull-down. Finally, click Submit. [Figure 14-13](#) shows the creation of a syslog source for the monCommonPol class.



**Figure 14-13** Creating a *Syslog Source*

Remember that even though this book demonstrates steps for the creation of a single syslog source, it is important to create syslog sources for each class of active monitoring policies within the fabric if the intent is to forward syslog messages pertinent to all aspects of the fabric.



Note that when configuring syslog sources in some of the monitoring policies, you may not be able to create the syslog source object by right-clicking the top-level menu. In such cases, you need to navigate to the CallHome/Smart CallHome/SNMP/Syslog/TACACS submenu for the monitoring policy, select Syslog, and then create syslog sources by clicking the + sign on the right side of the screen, as shown in [Figure 14-14](#).

The screenshot shows the Cisco ACI Fabric Policy configuration interface. The top navigation bar includes System, Tenants, Fabric (selected), Virtual Networking, L4-L7 Services, Admin, Operations, Apps, and Integrations. Below the navigation bar, the Fabric Policies tab is selected. The left sidebar lists Policies: Stats Collection Policies, Syslog Message Policies, Callhome/Smart Callhome/Sub, Fault Severity Assignment Policies, Fault Lifecycle Policy, and default (expanded to show Stats Collection Policies, Stats Export Policies, Diagnostics Policies, Callhome/Smart Callhome/Sub, Event Severity Assignment Policies, Fault Severity Assignment Policies, and Fault Lifecycle Policies). The main panel displays the 'Callhome/Smart Callhome/SNMP/Syslog/TACACS' configuration. It includes a Monitoring Object dropdown set to 'ALL', a Source Type dropdown with options Callhome, Smart Callhome, SNMP, and Syslog (Syslog is selected), and a table header for Name, Include, Min Severity, and Destination Group. A note at the bottom states 'No items have been found. Select Actions to create a new item.' A '+' button is located in the bottom right corner of the table area.

**Figure 14-14** Alternative Method for Creating Syslog Sources

Note that the use of the four default monitoring policies in these configuration examples assumes that ACI administrators are not customizing monitoring policies. However, in multitenant environments, monitoring policy customization is very common. This section goes over some common examples of monitoring policy customization—namely, interface-level and tenant-level syslog customizations.

Say that an administrator of a tenant named DCACI wants all system messages relevant to the tenant he or she manages to be forwarded to a dedicated group of syslog servers. A fabric administrator has already created a syslog monitoring destination group named DCACI-syslog that enables syslog forwarding and references the syslog servers managed by the tenant administration team. The next step would be for the tenant administrator to create custom monitoring policies. [Figure 14-15](#) shows the creation of a new monitoring policy called DCACI-monitoring and the association of a child monitoring source named DCACI-syslog-source with the new syslog monitoring destination group.

**Create Syslog Source**

Name:

Min Severity:

Include:    
 Audit logs  
 Events  
 Faults  
 Session logs

Dest Group:



**Figure 14-15** *Associating Syslog Sources with Custom Monitoring Policies*

At this point, you might think that the configuration is complete. Just because a new set of monitoring policies has been created within a tenant, however, does not mean the

tenant or objects within the tenant are associated with it. [Figure 14-16](#) shows how you can verify which monitoring policies are associated with an object. In this figure, the distinguished name of the monitoring policy object associated with the DCACI tenant is still uni/tn-common/monepg-default, which is the set of default monitoring policies of class monEPGPol from the common tenant.

Class or DN or URL	Property	Operation	Value	Run Query																																																		
		==																																																				
1 object found <a href="#">Show URL and response of last query</a>				Empty Properties: <a href="#">Show</a> <a href="#">Hide</a>																																																		
<b>fvTenant</b>																																																						
<table border="1"><thead><tr><th>dn</th><th>&lt; uni/tn-DCACI &gt;</th><th>id</th><th>▲</th><th>▼</th></tr></thead><tbody><tr><td>annotation</td><td></td><td></td><td></td><td></td></tr><tr><td>childAction</td><td></td><td></td><td></td><td></td></tr><tr><td>descr</td><td></td><td></td><td></td><td></td></tr><tr><td>extMngdBy</td><td></td><td></td><td></td><td></td></tr><tr><td>lcOwn</td><td>local</td><td></td><td></td><td></td></tr><tr><td>modTs</td><td>2019-11-11T00:21:54.595+00:00</td><td></td><td></td><td></td></tr><tr><td>monPolDn</td><td>&lt; uni/tn-common/monepg-default &gt;</td><td></td><td></td><td></td></tr><tr><td>name</td><td>DCACI</td><td></td><td></td><td></td></tr><tr><td>nameAlias</td><td></td><td></td><td></td><td></td></tr></tbody></table>				dn	< uni/tn-DCACI >	id	▲	▼	annotation					childAction					descr					extMngdBy					lcOwn	local				modTs	2019-11-11T00:21:54.595+00:00				monPolDn	< uni/tn-common/monepg-default >				name	DCACI				nameAlias					
dn	< uni/tn-DCACI >	id	▲	▼																																																		
annotation																																																						
childAction																																																						
descr																																																						
extMngdBy																																																						
lcOwn	local																																																					
modTs	2019-11-11T00:21:54.595+00:00																																																					
monPolDn	< uni/tn-common/monepg-default >																																																					
name	DCACI																																																					
nameAlias																																																						

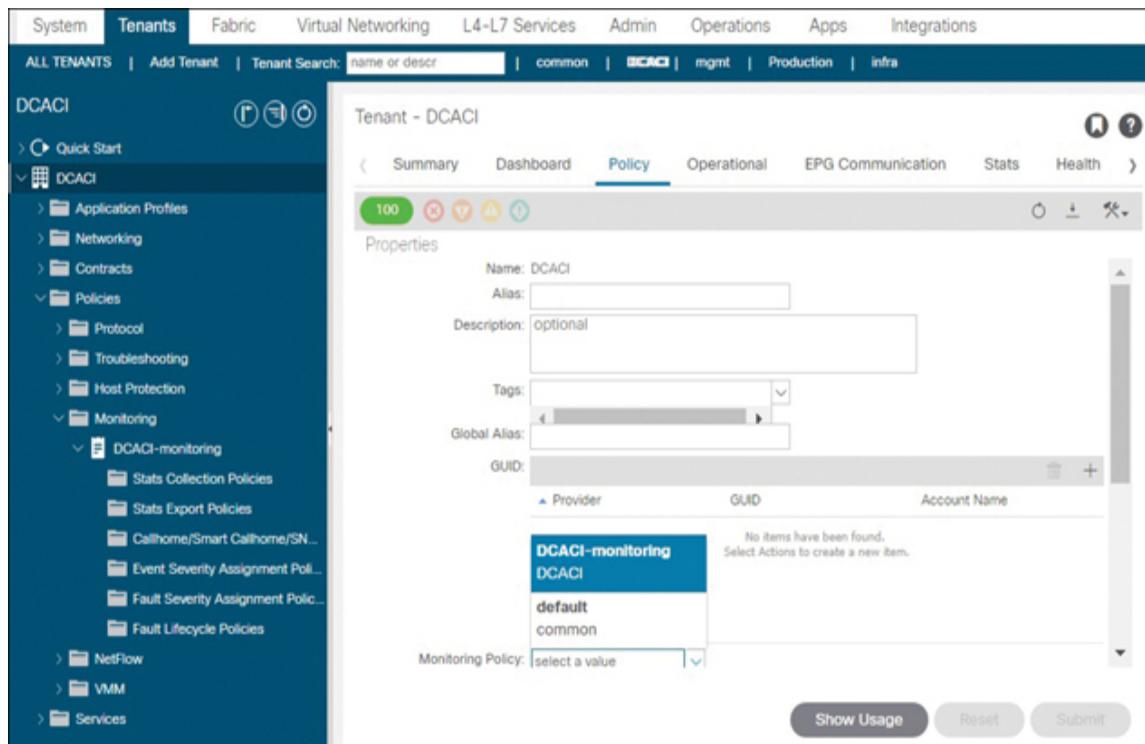
**Figure 14-16** *Checking the Monitoring Policies Associated with an Object*

### Note

[Figure 14-16](#) provides an important example of why administrators should familiarize themselves with the ACI object model. Verification of traditional network configurations is possible using simple CLI commands. In ACI, on the other hand, it can sometimes be difficult

to understand where a configuration went wrong without knowledge of the object model.

To modify the active set of monitoring policies for a tenant, you must navigate to the tenant in question, open the Policy tab, select the new policy from the Monitoring Policy drop-down, and then click Submit (see [Figure 14-17](#)).



**Figure 14-17** Assigning Custom Monitoring Policies to a Tenant

As a result of making this change, the tenant as well as any child objects, such as bridge domains and EPGs, begin to reference the custom monitoring policies through inheritance, as indicated by the updated distinguished name in [Figure 14-18](#).

Class or DN or URL	Property	Operation	Value																																																			
uni/tn-DCACI/ap-DCACI-Apps		==		<button>Run Query</button>																																																		
1 object found <a>Show URL and response of last query</a>																																																						
Empty Properties: <a>Show</a> <a>Hide</a>																																																						
<b>fvAp</b>																																																						
<table border="1"> <thead> <tr> <th>dn</th> <th>&lt; uni/tn-DCACI/ap-DCACI-Apps &gt;</th> <th>■</th> <th>▲</th> <th>▼</th> </tr> </thead> <tbody> <tr> <td>annotation</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>childAction</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>descr</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>extMngdBy</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>lcOwn</td> <td>local</td> <td></td> <td></td> <td></td> </tr> <tr> <td>modTs</td> <td>2019-11-11T00:40:56.798+00:00</td> <td></td> <td></td> <td></td> </tr> <tr> <td>monPolDn</td> <td>&lt; uni/tn-DCACI/monepg-DCACI-monitoring &gt;</td> <td></td> <td></td> <td></td> </tr> <tr> <td>name</td> <td>DCACI-Apps</td> <td></td> <td></td> <td></td> </tr> <tr> <td>nameAlias</td> <td></td> <td></td> <td></td> <td></td> </tr> </tbody> </table>					dn	< uni/tn-DCACI/ap-DCACI-Apps >	■	▲	▼	annotation					childAction					descr					extMngdBy					lcOwn	local				modTs	2019-11-11T00:40:56.798+00:00				monPolDn	< uni/tn-DCACI/monepg-DCACI-monitoring >				name	DCACI-Apps				nameAlias				
dn	< uni/tn-DCACI/ap-DCACI-Apps >	■	▲	▼																																																		
annotation																																																						
childAction																																																						
descr																																																						
extMngdBy																																																						
lcOwn	local																																																					
modTs	2019-11-11T00:40:56.798+00:00																																																					
monPolDn	< uni/tn-DCACI/monepg-DCACI-monitoring >																																																					
name	DCACI-Apps																																																					
nameAlias																																																						

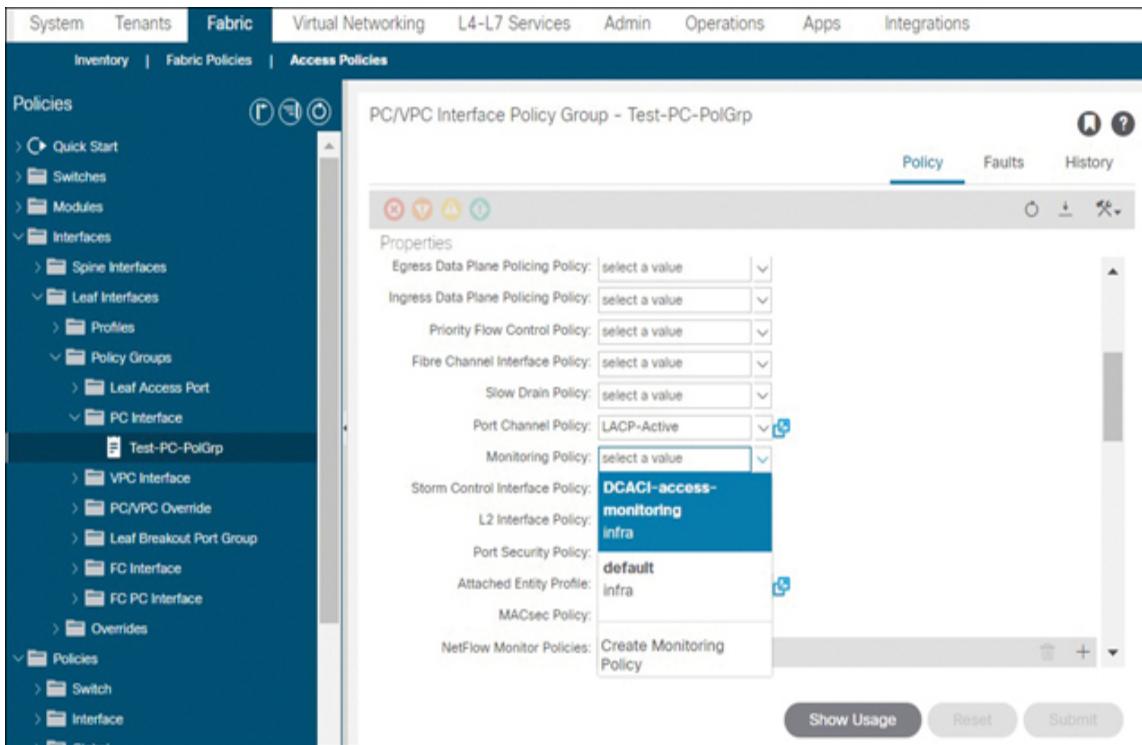
**Figure 14-18** Verifying Assignment of Custom Monitoring Policies

### Note

Certain attributes flow downward in the hierarchical object structure according to the location of objects within the tree.

So how are custom monitoring policies assigned to objects in the fabric policies and access policies view? There is always a possibility that a customer or department might own a physical server or even several ACI switches within a fabric and might therefore expect to be able to receive syslog data pertinent to incidents involving hardware, such as ports bouncing or a switch port failure. In such cases, custom fabric and access monitoring policies need to be defined and associated with the highest-level object in the hierarchy below which the customer needs to monitor.

Through inheritance, the monPolDn attribute of child objects then reflects the custom monitoring policies. [Figure 14-19](#) shows how a custom set of access monitoring policies can be applied to an interface policy group. Note that monitoring policies for switches and modules can also be modified via switch and module policy groups.



**Figure 14-19** Updating Monitoring Policies via Policy Groups

## Verify Syslog Forwarding to Desired Syslog Servers

Configuration of syslog forwarding is often a task that takes place during the initial buildout of an ACI fabric. It is important for administrators not to assume that all desired messages will be forwarded correctly; it is also important that this theory be properly validated.

The first thing to verify is IP connectivity between the ACI out-of-band or in-band management interfaces and the syslog servers.

After verifying IP connectivity, it is a good idea to test whether the syslog service on each node in the fabric can forward traffic properly to each syslog server configured in the syslog monitoring destination groups by using the APIC CLI command **logit severity { severity} dest-grp { destGroup} server { remoteDest} { message} (node { id})**. Example 14-4 shows how a user can prompt the APIC to signal nodes 101 and 102 to generate a test system message using severity level 1 and forward it to each of the configured syslog servers in the Syslog-Servers destination group configured earlier.

#### **Example 14-4 Using *logit* to Test System Message Submission**

[Click here to view code image](#)

```
apic1# logit severity 1 dest-grp Syslog-Servers server  
10.233.48.10 "This is a test"  
    node 101  
apic1# logit severity 1 dest-grp Syslog-Servers server  
10.233.48.10 "This is a test"  
    node 102  
apic1# logit severity 1 dest-grp Syslog-Servers server  
10.133.48.10 "This is a test"  
    node 101  
apic1# logit severity 1 dest-grp Syslog-Servers server  
10.133.48.10 "This is a test"  
    node 102
```

If syslog messages are indeed received by the syslog server, the communication path, the syslog service, and the health

of the syslog server have been sufficiently validated.

After verifying the syslog monitoring destination groups, you can verify the syslog source configuration to ensure that all necessary configurations are in place to forward desired system messages to syslog servers. The most reliable way to verify that syslog sources are in place is to trigger various types of failures within the system and ensure that system messages arrive on syslog servers. Just remember that triggers need to be identified for each monitoring class for which syslog sources have been configured.

One way to verify syslog configurations on all nodes in a fabric is to use the MOQuery tool. To query for all syslog sources (class syslogSrc) configured in a fabric, use the command **moquery -c syslogSrc**. To also verify whether the syslog sources have been assigned to the desired destination groups, you can add the **-x** option with the argument '**rsp-subtree=children**'. This ensures that child objects are also included in the query. [Example 14-5](#) shows how this is possible. In this example, the output has been piped via **grep** to exclude certain output for brevity. This output shows that four syslog sources have been configured: common-syslog, fabric-syslog, access-syslog, and tenant-syslog. These sources correlate to the four out-of-the-box monitoring groups in ACI and are all the syslog sources configured in this specific fabric.

#### **Example 14-5** Using MOQuery to Validate Syslog Source Configurations

[Click here to view code image](#)

```
apic1# moquery -c syslogSrc -x 'rsp-subtree=children' | grep  
-E "name  
|dn|incl|monPolDn|tDn|^$"
```

```
name      : common-syslog
dn        : uni/fabric/moncommon/slsrc-common-syslog
incl     : all,audit,events,faults,session
monPolDn : uni/fabric/moncommon

dn        : uni/fabric/moncommon/slsrc-common-
syslog/rsdestGroup
monPolDn : uni/fabric/moncommon
tDn      : uni/fabric/slgroup-Syslog-Servers

name      : fabric-syslog
dn        : uni/fabric/monfab-default/slsrc-fabric-syslog
incl     : all,audit,events,faults,session
monPolDn : uni/fabric/monfab-default

dn        : uni/fabric/monfab-default/slsrc-fabric-
syslog/rsdestGroup
monPolDn : uni/fabric/monfab-default
tDn      : uni/fabric/slgroup-Syslog-Servers

name      : access-syslog
dn        : uni/infra/moninfra-default/slsrc-access-syslog
incl     : all,audit,events,faults,session
monPolDn : uni/infra/moninfra-default

dn        : uni/infra/moninfra-default/slsrc-access-
syslog/rsdestGroup
monPolDn : uni/infra/moninfra-default
tDn      : uni/fabric/slgroup-Syslog-Servers

name      : tenant-syslog
dn        : uni/tn-common/monepg-default/slsrc-tenant-
syslog
incl     : all,audit,events,faults,session
```

```

monPolDn      : uni/tn-common/monepg-default

dn           : uni/tn-common/monepg-default/slsrc-tenant-
syslog/rsdestGroup
monPolDn      : uni/tn-common/monepg-default
tDn          : uni/fabric/slgroup-Syslog-Servers

```

The MOQuery tool can also display the syslog server destinations that have been configured. The **syslogRemoteDest** class pertains to syslog servers configured in the fabric, as shown in [Example 14-6](#).

#### **Example 14-6** Using MOQuery to Validate Syslog Destination Groups

[Click here to view code image](#)

```

apic1# moquery -c syslogRemoteDest | grep -E
' #|host|adminState|epgDn|severity'
# syslog.RemoteDest
host          : 10.133.48.10
adminState     : enabled
epgDn         : uni/tn-mgmt/mgmtp-default/oob-default
severity       : warnings
# syslog.RemoteDest
host          : 10.233.48.10
adminState     : enabled
epgDn         : uni/tn-mgmt/mgmtp-default/oob-default
severity       : warnings

```

## Using SNMP in ACI

Simple Network Management Protocol (SNMP) allows third-party applications to monitor network devices. The application that performs the monitoring is called an *SNMP*

*manager*; the system being managed is referred to as an *SNMP agent*.

SNMP managers initiate SNMP read queries for SNMP agents to send certain information. If supported, SNMP managers can also send configuration changes to SNMP agents through SNMP write commands.

Whether an SNMP agent supports read queries, write commands, or both, the information that a remote system can request from the agents is defined in an object called a **Management Information Base (MIB)**.

When important system events occur on an SNMP-enabled device, the SNMP agent can send **SNMP notifications** in the form of traps and informs to the SNMP manager. An **SNMP trap** is an unreliable message that does not require an acknowledgment from the SNMP manager. An *inform* is a message that is sent reliably, as it is stored in memory until the SNMP manager issues a response.

Events of interest for SNMP notifications may include a module crashing or an interface going down.

There are three major versions of SNMP. SNMPv1 and SNMPv2c use community strings. SNMPv3 adds encryption and authentication capabilities and is considered the most secure but also generally consumes more CPU and memory resources. SNMPv3 uses a concept called security levels to define the level of security to be enforced. SNMPv3 security level definitions along with keywords used for each security level in ACI are as follows:

- **auth:** Authenticates users but does not encrypt traffic
- **noauth:** Does not authenticate or encrypt traffic and uses a username match

- **priv:** Both authenticates SNMPv3 users and encrypts traffic

The default ports involved in SNMP communication are UDP ports 161 and 162. SNMP agents listen for SNMP manager read queries on port 161. SNMP managers listen for traps on port 162.

## ACI Support for SNMP

ACI switches and APICs run SNMP agents and support SNMPv1, v2, and v3, including both MIBs and notifications (traps). However, ACI does *not* support SNMP write commands.

SNMP support in ACI can be summarized as follows:



- SNMP read queries (Get, Next, Bulk, Walk) are supported by leaf and spine switches and by APICs. Only MIBs specific to ACI are supported.
- SNMP write commands (Set) are not supported in ACI.
- SNMP traps (v1, v2c, and v3) are supported by leaf and spine switches and by APICs. When system messages are sent to an SNMP manager, these messages are in the form of SNMP traps, and so the configuration of SNMP trap forwarding is very similar to the process for forwarding syslog messages.
- ACI supports a maximum of 10 trap receivers. If more than 10 trap receivers are configured, some of them do not receive notifications.

- ACI supports both IPv4 and IPv6 SNMP trap destinations.

### Note

Cisco has specially developed a plethora of MIBs for ACI to enable very granular monitoring. To obtain the list of MIBs supported in ACI, visit the ACI MIB Support List at <https://www.cisco.com/c/dam/en/us/td/docs/Website/dacenter/aci/mib/mib-support.html>.

## ACI SNMP Configuration Caveats

Administrators need to authorize SNMP managers that should be allowed to query ACI via SNMP. To do so, SNMP manager IP addresses need to be configured in a field called Client Entries in an SNMP client group profile. Adding SNMP manager IP addresses to the Client Entries list is like permitting endpoints via an access list. Although a subnet can be used to define the SNMP managers, a broad subnet such as 0.0.0.0/0 should not be used in the Client Entries field.

The SNMP client group profile configuration needs to be applied at the pod level through an SNMP policy. Any SNMP community strings or SNMPv3 user credentials that the SNMP manager will be using to execute read queries against ACI also need to be configured in this same SNMP policy.

An important thing to note about ACI MIBs is that they are divided into two categories:



- **Global scope:** A ***global scope MIB*** is an MIB whose scope is not limited to a specific VRF instance and that touches on broader aspects of the fabric. Examples of MIBs of a global scope are those that request data related to the status of switch power supplies, interface or port channel statuses, CPU utilization, and memory utilization.
- **VRF specific:** A ***VRF-specific MIB*** is an MIB whose scope is limited to a VRF. Examples of VRF-specific MIBs are those involving IP addresses or endpoints residing in a VRF or route peerings out of a specific VRF.

When using VRF-specific MIBs, an extra configuration step is necessary. A VRF-specific SNMP context needs to be created, and one or more SNMP community profiles need to be associated with the SNMP context. The SNMP community then becomes bound to the SNMP context and can only be used for SNMP read queries involving the specific VRF, regardless of whether the SNMP community was previously associated with a global scope at the SNMP pod policy level.

A prerequisite for proper SNMP implementation for APICs is for nodes to have static management IP addresses assigned. The process for configuring static out-of-band addresses is covered in [Chapter 3](#).



## Configuring ACI for SNMP

The following steps are required to configure ACI to respond to SNMP read queries and forward traps:

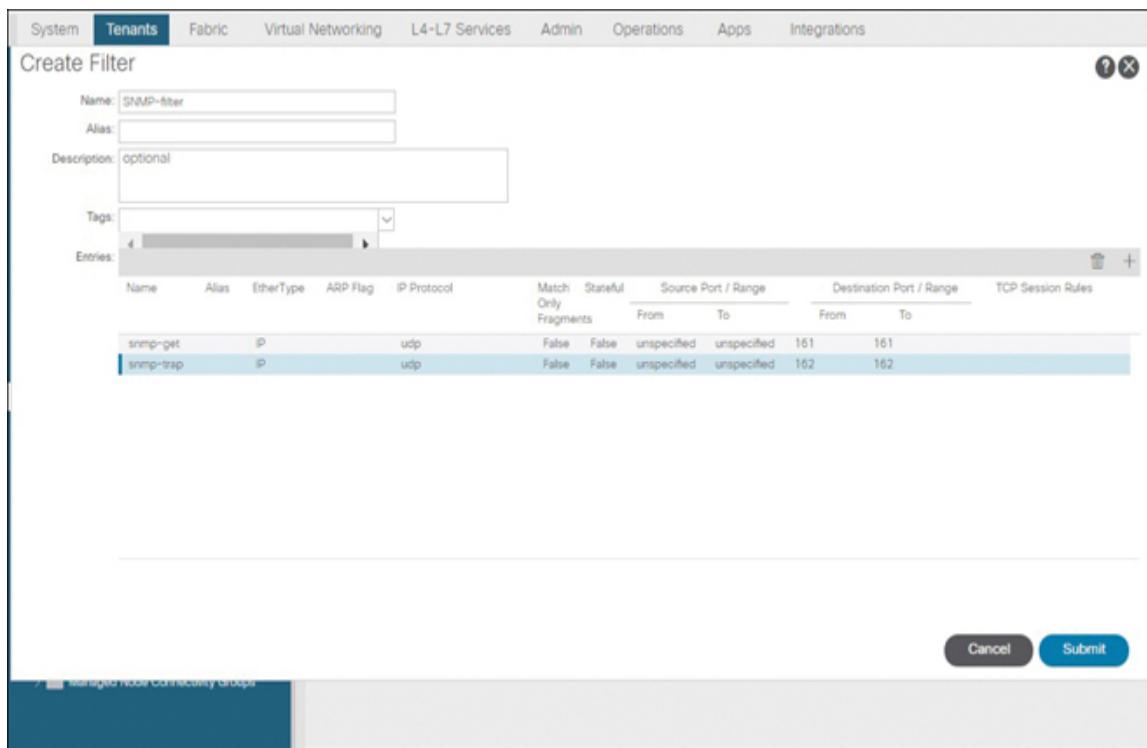


- Step 1.** Apply necessary contracts for SNMP.
- Step 2.** Associate SNMP policy with pod policy.
- Step 3.** (Optional) Associate SNMP contexts with desired VRF instances.
- Step 4.** Configure SNMP monitoring destination groups.
- Step 5.** Configure SNMP sources for all desired monitoring policies.
- Step 6.** Verify SNMP forwarding to desired SNMP servers.

Although these steps generally do not need to be done in order, configuration of monitoring destination groups should take place before SNMP sources are configured. This is because each SNMP source configuration needs to reference an SNMP monitoring destination group. Furthermore, there is little point in verifying SNMP configuration before full implementation. Also, note that these steps assume that either out-of-band or in-band management has been fully set up with static node management addresses. The following sections provide some details to make these steps clearer.

## Apply Necessary Contracts for SNMP

As noted earlier in this chapter, contracts should allow UDP ports 161 and 162 to enable SNMP communication. [Figure 14-20](#) shows filter entries that satisfy requirements for SNMP. These filters need to be associated with a subject that permits traffic forwarding and has the Reverse Filter Ports option enabled. For in-band management, you need to also enable the Apply Both Directions option.



**Figure 14-20 Filters Needed for SNMP Reads and Traps**

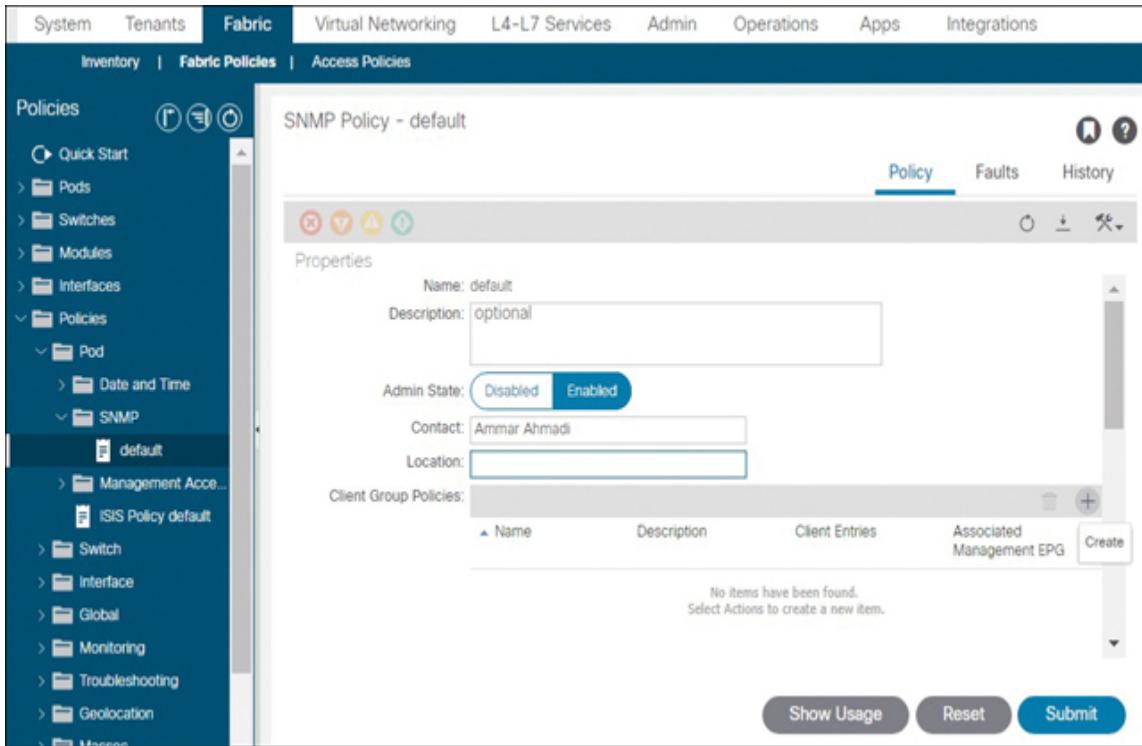
### Note

Under certain conditions, either these filters or contract directionality may need to be tweaked. For example, if in-band management is used and L3Out external EPGs are explicit in the IP addresses allowed to communicate via SNMP, a contract with this same filter can be both consumed and provided by both the external EPG and the in-band EPG to ensure full SNMP communication.

## Associate an SNMP Policy with a Pod Policy

To enable remote SNMP managers to query the fabric, a pod SNMP policy needs to be configured. Navigate to Fabric, click Fabric Policies, double-click Policies, open Pod, double-

click SNMP, and edit the desired SNMP policy. [Figure 14-21](#) shows the default SNMP policy that is already applied to all pods within the fabric being modified.



**Figure 14-21** Filters Needed for SNMP Reads and Traps

Enable SNMP by toggling the Admin State parameter. Then enter the desired contact details and click the + sign in the Client Group Policies portion of the window to create a client SNMP group profile, as shown in [Figure 14-22](#). Associate a management EPG to the client SNMP group profile and enter addresses for the SNMP managers. Finally, click Submit.

Create SNMP Client Group Profile

Name:

Description:

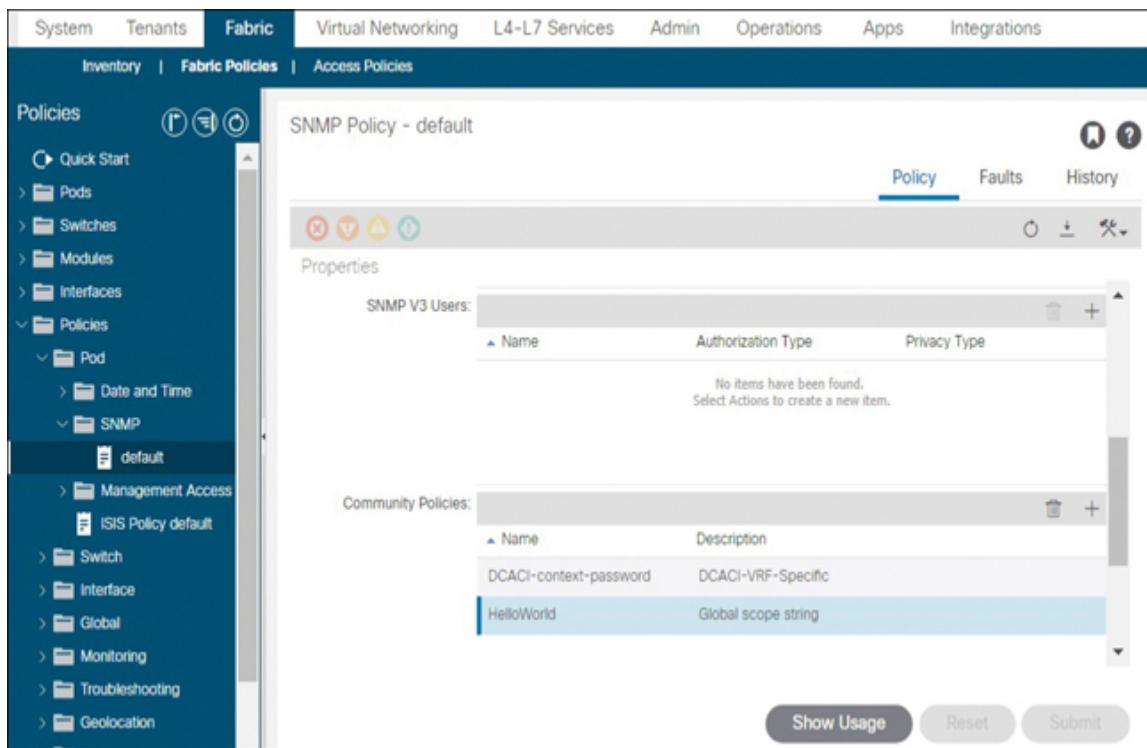
Associated Management EPG:

Client Entries:

Name	Address
DC1-SNMP	10.233.48.10
DC2-SNMP	10.133.48.10

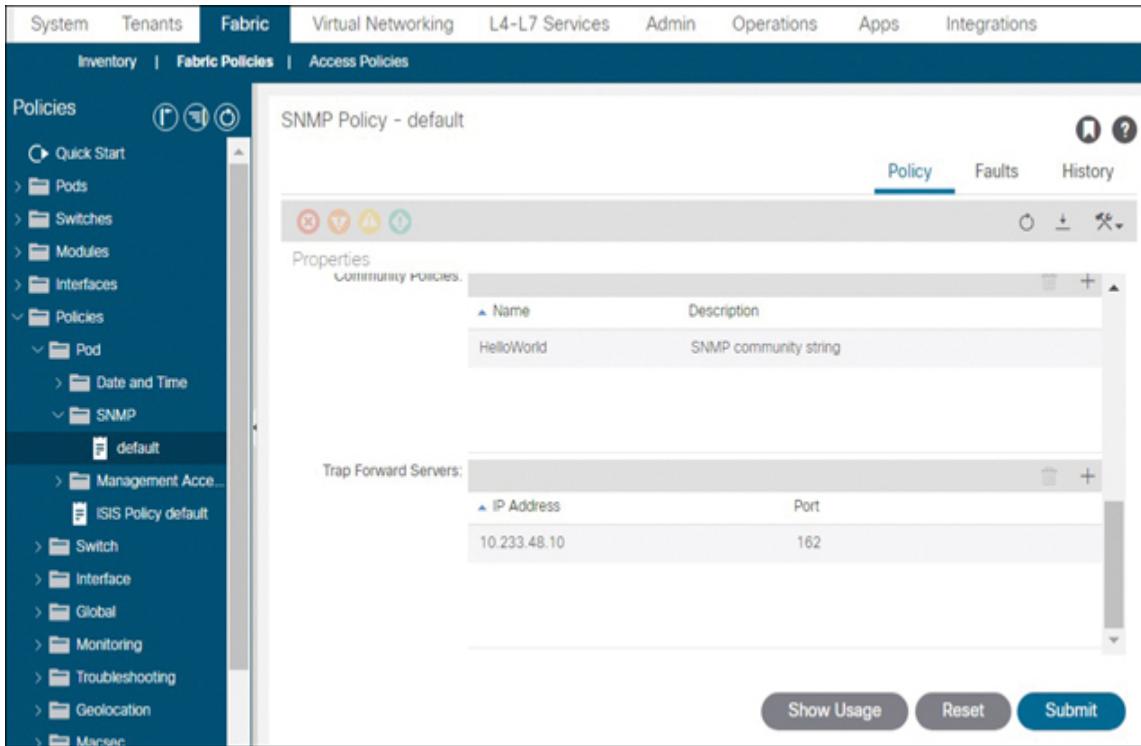
**Figure 14-22** Creating an SNMP Client Group Profile

After defining the client SNMP group profile, scroll down in the SNMP policy and configure either SNMPv3 users or SNMP community strings to allow read queries. [Figure 14-23](#) shows that two SNMP community strings have been defined. One of these community policies will be allocated to an SNMP context that will be bound to a VRF instance called DCACI. Once this binding takes place, the VRF-specific SNMP string cannot be used for queries involving MIBs that are of a global scope.



**Figure 14-23** Creating SNMP Community Policies and/or SNMPv3 Users

Scroll down further within the SNMP policy to define servers that will receive SNMP traps and the port on which these servers listen for traps. [Figure 14-24](#) shows that an SNMP server at 10.233.48.10 has been defined as a trap receiver that listens on port 162. Click Submit to apply the changes to the SNMP policy.

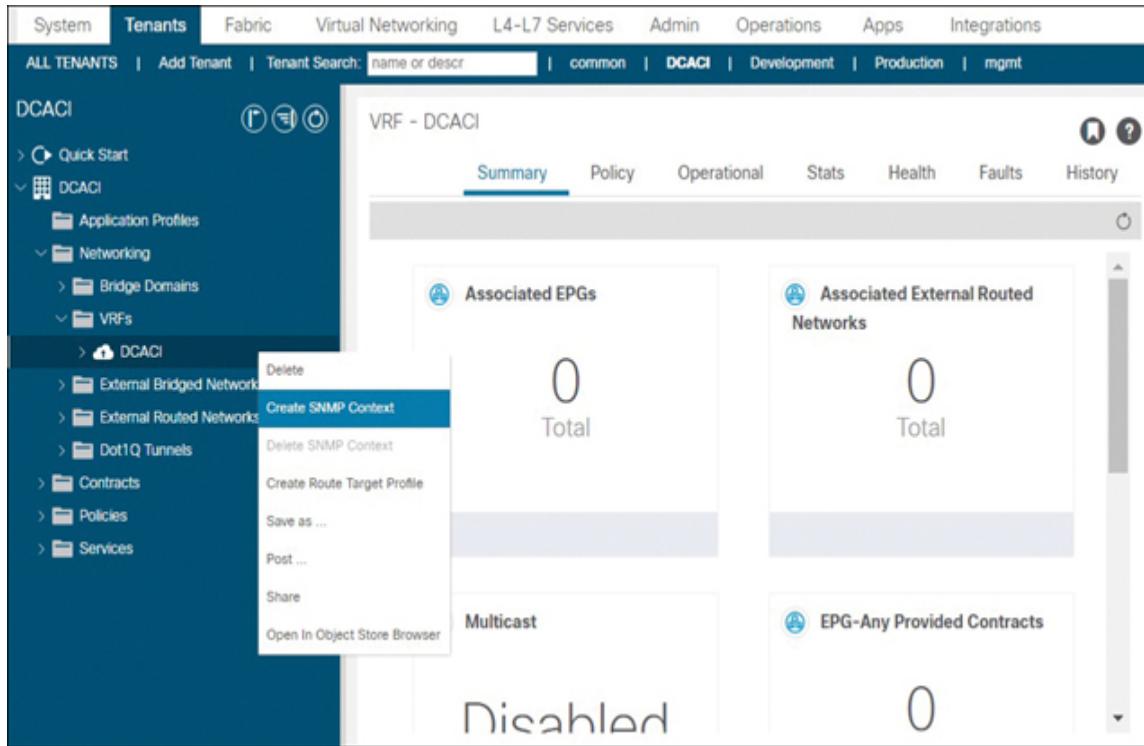


**Figure 14-24** Defining Servers That Will Receive SNMP Traps

Remember that for the SNMP policy to take effect, it has to be assigned to an active pod policy group. Click Show Usage to ensure that the SNMP policy is active. If it is not, navigate to the pod policy group of interest and update the SNMP policy the pod policy group references.

## Associate SNMP Contexts with Desired VRF Instances

To create an SNMP context to enable VRF-specific MIB queries from a remote SNMP manager, navigate to the tenant of interest, right-click the desired VRF instance, and select Create SNMP Context (see [Figure 14-25](#)).



**Figure 14-25** Creating an SNMP Context

In the Create SNMP Context window, enter a context name, define the community string that will be used to query ACI for information specific to the VRF, and click Submit. [Figure 14-26](#) shows that DCACI-context-password is the SNMP community string that will be used for this SNMP context. This string can no longer be used for queries of MIBs of a global scope, even though this string was also defined under the pod SNMP policy.

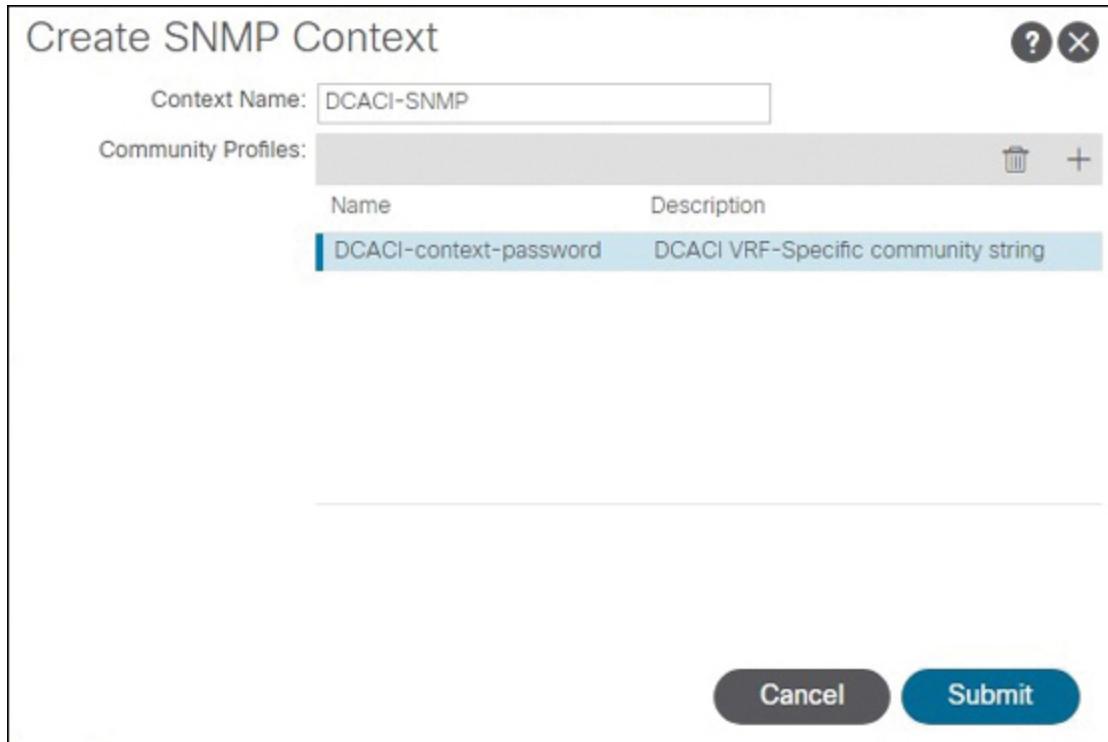
Create SNMP Context

Context Name: DCACI-SNMP

Community Profiles:

Name	Description
DCACI-context-password	DCACI VRF-Specific community string

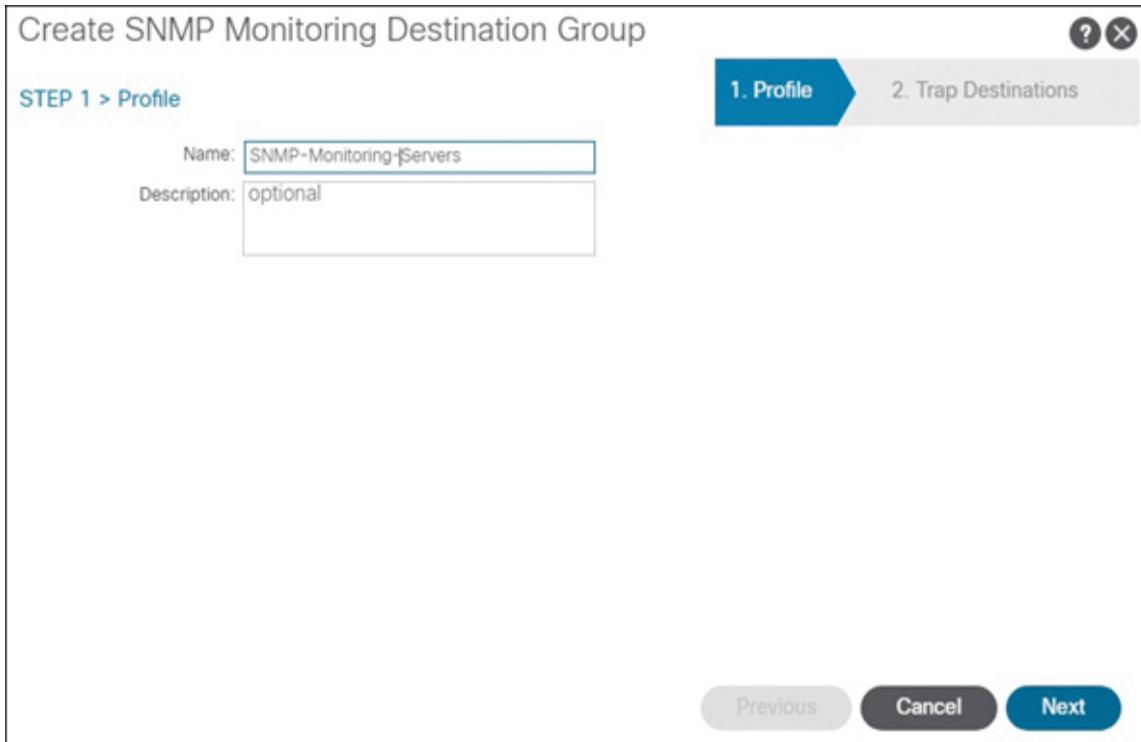
Cancel Submit



**Figure 14-26** Dedicating a Community String to an SNMP Context

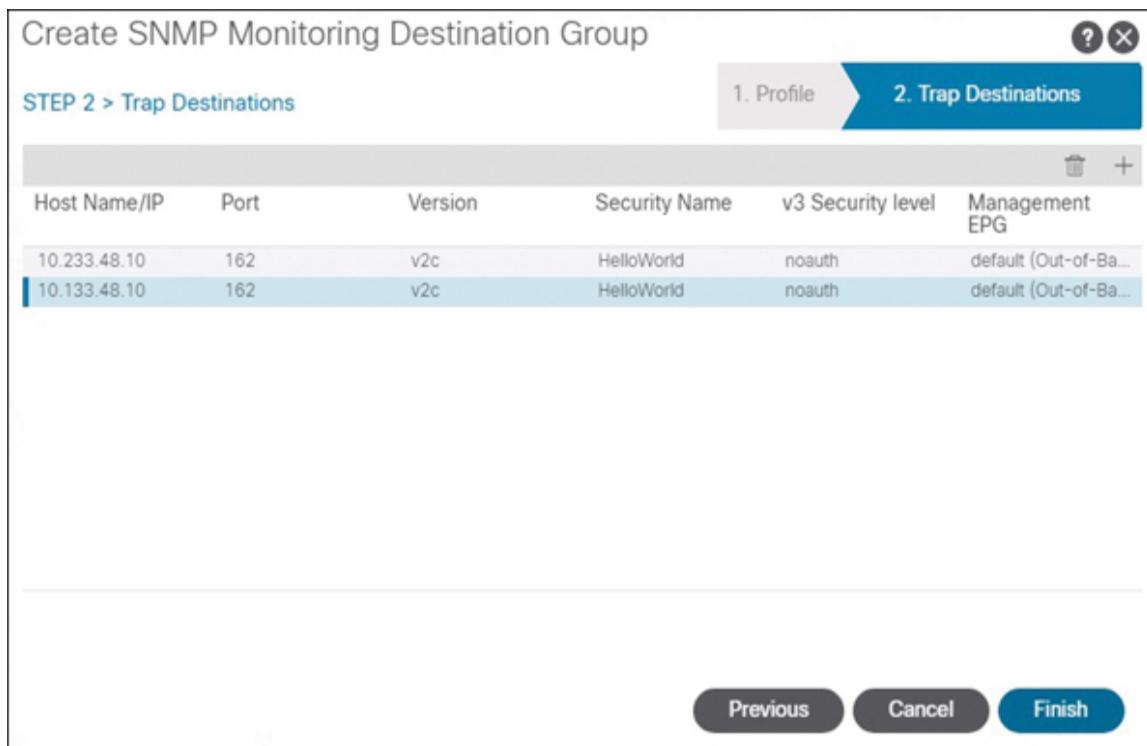
## Configure SNMP Monitoring Destination Groups

Just as with syslog configuration, SNMP trap forwarding requires the definition of SNMP monitoring destination groups. You also need to define the SNMP protocol version used for trap forwarding on a per-server basis. Navigate to Admin, select External Data Collectors, open Monitoring Destinations, right-click SNMP, and select Create SNMP Monitoring Destination Group. [Figure 14-27](#) shows the first page of the wizard, where you select a name for the object and click Next.



**Figure 14-27** Launching the *Create SNMP Monitoring Destination Group Wizard*

Create entries for each SNMP manager that should receive traps. The information required includes destination addresses, ports, SNMP version, and community strings to use for trap forwarding. [Figure 14-28](#) shows that an administrator has configured two servers for trap forwarding. Click Finish to complete the configuration.



**Figure 14-28 Adding Servers to SNMP Monitoring Destination Group for Trap Forwarding**

Note that DNS names can be used in SNMP monitoring destination groups. If SNMPv3 is used, an SNMPv3 security level needs to also be selected. Instead of a community string, SNMPv3 trap forwarding requires that the SNMPv3 username be entered as the community. The GUI sometimes refers to the SNMPv3 username as a security name.

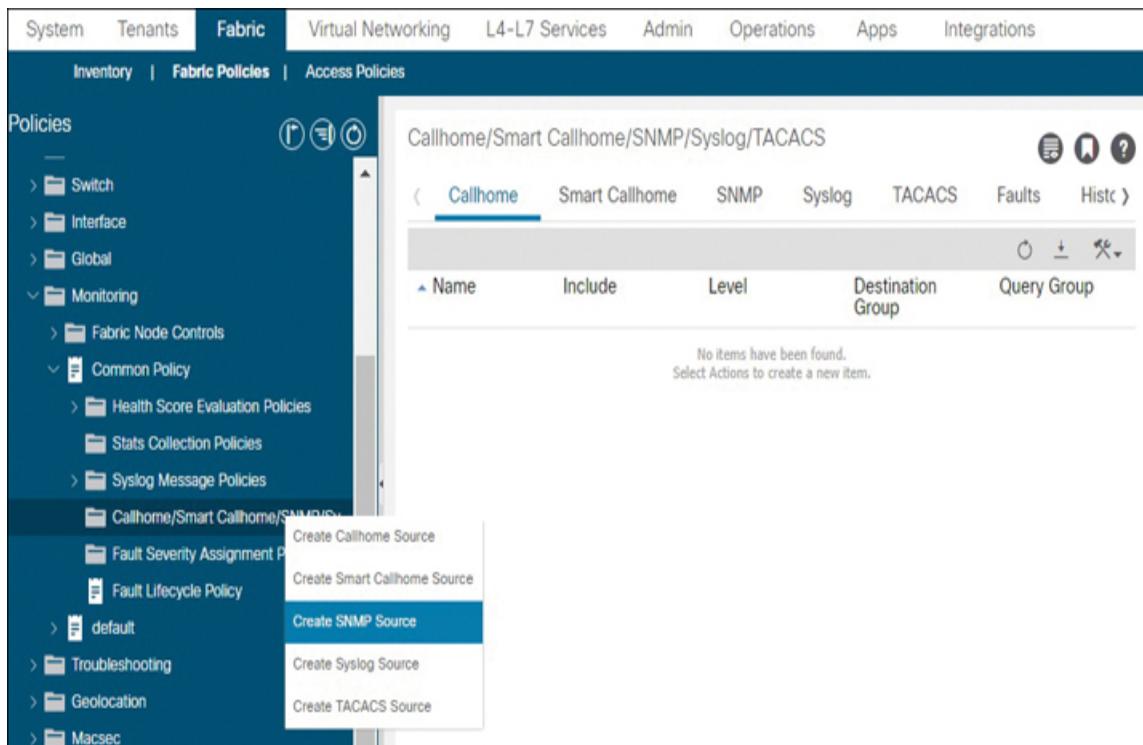
## Configure SNMP Sources for All Desired Monitoring Policies

Much as with syslog, SNMP system message forwarding in the form of traps necessitates that SNMP sources be configured for four default monitoring groups in ACI. Each group addresses a different set of faults, events, and logs. The default monitoring groups can be found in the following locations:

**Key Topic**

- Fabric > Fabric Policies > Policies > Monitoring > Common Policy
- Fabric > Fabric Policies > Policies > Monitoring > default
- Fabric > Access Policies > Policies > Monitoring > default
- Tenant > common > Policies > Monitoring > default

**Figure 14-29** shows the SNMP source creation wizard being launched for the common policy.



**Figure 14-29** Launching the SNMP Source Creation Wizard for the Common Policy

In the wizard, name the SNMP source, map the previously created SNMP monitoring destination group to the SNMP source, and click Submit. [Figure 14-30](#) shows the creation of an SNMP source named common-snmp.

The screenshot shows a dialog box titled "Create SNMP Source". It contains two input fields: "Name" with the value "common-snmp" and "Dest Group" with the value "SNMP-Monitoring-Servers". Below the fields are two buttons: "Cancel" and "Submit". The "Submit" button is highlighted with a blue background.

**Figure 14-30** Creating an SNMP Source

Create SNMP sources for each of the desired monitoring policy classes. Custom SNMP sources can also be used to achieve additional granularity.

## Verify SNMP Forwarding to Desired SNMP Servers

The best way to verify SNMP read queries is to actually execute read queries using the SNMP manager. Command-line tools available on most SNMP managers (for example, `snmpget` and `snmpwalk`) can execute queries against an

SNMP agent. Trap receipt is also best validated on the SNMP manager itself.

[Example 14-7](#) shows how the basic SNMP policy applied at the pod level can be verified using the **show snmp** APIC CLI command.

### **Example 14-7 Verifying Operational Status of SNMP Pod Policy Settings**

[Click here to view code image](#)

```
apic1# show snmp
(...output truncated for brevity...)
Input Statistics:
    34 SNMP packets input
        0 Trap PDUs received
        48 Get-next PDUs
        0 General Errors
        0 Set-request PDUs
        44 Number of requested variables
Output Statistics:
    0 Get-request PDUs generated
    58 Get-responses PDUs generated
    0 Set-requests PDUs generated
    34 SNMP packets output
Other Statistics:
    0 Silent Drops
    0 Proxy Drops
Disabled Authentication Traps Status
  Name          Admin State  Location      Contact
Description
  -----
  -----
  default       enabled      Ammar
  Ahmadi
```

---

Other commands that can be used to verify SNMP configurations using the APIC CLI include **show snmp clientgroups**, **show snmp community**, **show snmp hosts**, and **show snmp users**. The command **show snmp summary** is also very helpful (see [Example 14-8](#)).

### **Example 14-8** Verifying SNMP Configuration Settings

[Click here to view code image](#)

```
apic1# show snmp summary

Active Policy: default, Admin State: enabled

Local SNMP engineID: [Hex]
0x800000098057dbed2fc3d3c45d00000000

-----
Community          Description
-----
HelloWorld         Global scope string
DCACI-context-password DCACI-VRF-Specific

-----
User                Authentication      Privacy
-----

-----
Client-Group        Mgmt-Epg           Clients
-----

-----
SNMP-Servers       default (Out-Of-Band)
10.133.48.10,10.233.48.10
```

Host	Port	Version	Level	SecName
<hr/>				
10.233.48.10	162	v2c	noauth	HelloWorld
10.133.48.10	162	v2c	noauth	HelloWorld

Finally, the MOQuery tool can be used to validate SNMP source configurations. You can query the ACI object hierarchy for SNMP source configurations by using the class snmpSrc. [Example 14-9](#) shows that only a single set of monitoring policies has been configured for the fabric in this case.

### **Example 14-9** Verifying SNMP Source Configurations

[Click here to view code image](#)

```
apic1# moquery -c snmpSrc -x 'rsp-subtree=children'
Total Objects shown: 1
(...output truncated for brevity...)
# snmp.Src
name      : common-snmp
dn        : uni/fabric/moncommon/snmpsrc-common-snmp
incl     : events,faults
monPolDn  : uni/fabric/moncommon
rn       : snmpsrc-common-snmp

# snmp.RsDestGroup
dn        : uni/fabric/moncommon/snmpsrc-common-
snmp/rsdestGroup
monPolDn  : uni/fabric/moncommon
tDn      : uni/fabric/snmpgroup-SNMP-Monitoring-Servers
```

## Exam Preparation Tasks

As mentioned in the section “How to Use This Book” in the Introduction, you have a couple of choices for exam preparation: [Chapter 17](#), “Final Preparation,” and the exam simulation questions in the Pearson Test Prep Software Online.

## Review All Key Topics

Review the most important topics in this chapter, noted with the Key Topic icon in the outer margin of the page. [Table 14-4](#) lists these key topics and the page number on which each is found.



**Table 14-4** Key Topics for [Chapter 14](#)

Key Topic Element	Description	Page Number
Paragraph	Describes system messages	<a href="#">485</a>
<a href="#">Table 14-2</a>	Lists NX-OS system message severity levels	<a href="#">485</a>
<a href="#">Table 14-3</a>	Lists ACI system message severity levels	<a href="#">486</a>

List	Outlines steps for configuring ACI for syslog forwarding	487
Paragraph	Describes syslog monitoring destination groups	492
Paragraph	Revisits the default monitoring policies and where each of them can be found in the GUI	494
Paragraph	Describes the significance of configuring syslog sources for all active monitoring policies	495
List	Outlines the level of support within ACI for SNMP	501
List	Describes types of ACI MIBs from a scope perspective and defines each of them	502
Paragraph	Lists requirement for static node management addresses for SNMP	502
List		502

	Outlines the steps necessary for proper SNMP configuration	
List	Lists the classes of monitoring policies available for SNMP source configuration	508

## Complete Tables and Lists from Memory

There are no memory tables or lists for this chapter.

## Define Key Terms

Define the following key terms from this chapter and check your answers in the glossary:

system message

monitoring destination group

monitoring source

Management Information Base (MIB)

SNMP notification

SNMP trap

global scope MIB

VRF-specific MIB

# Chapter 15

## Implementing AAA and RBAC

This chapter covers the following topics:

**Implementing Role-Based Access Control (RBAC):**

This section addresses the technical tools ACI offers for RBAC enablement.

**Integrating with External AAA Servers:** This section details how ACI integrates with TACACS+, RADIUS, and LDAP servers.

This chapter covers the following exam topic:

- 5.4 Implement AAA and RBAC

This book has placed a lot of emphasis on multitenancy as a key feature of ACI. Multi-tenancy is used to lock down and separate data plane traffic, and it also enables a level of fault isolation from a configuration standpoint.

In the world of multitenancy, management plane lockdown is critical. Where IT is expected to *not* function as a cost center but an enabler and seller of new services to the business, everyone becomes a customer. Tenants then become a powerful tool in the IT arsenal for isolating or restricting customer traffic. At times, the service offered

may be to allow business units to independently deploy new services in the tenants they are assigned. This may require central IT organizations to provide customers a good deal of management access to ACI fabrics, but central IT still needs to be able to restrict access to functions that may enable customers to break ACI fabrics. To sum up, the demands placed on the new world of networking necessitate very granular role-based access control (RBAC). ACI meets such demands head on through its robust RBAC and AAA capabilities.

## **“Do I Know This Already?” Quiz**

The “Do I Know This Already?” quiz allows you to assess whether you should read this entire chapter thoroughly or jump to the “Exam Preparation Tasks” section. If you are in doubt about your answers to these questions or your own assessment of your knowledge of the topics, read the entire chapter. [Table 15-1](#) lists the major headings in this chapter and their corresponding “Do I Know This Already?” quiz questions. You can find the answers in [Appendix A, “Answers to the ‘Do I Know This Already?’ Questions.”](#)

**Table 15-1** “Do I Know This Already?” Section-to-Question Mapping

Foundation Topics Section	Questions
Implementing Role-Based Access Control (RBAC)	1-7
Integrating with External AAA Servers	8-10

## Caution

The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark that question as wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

- 1.** A user has been assigned to a security domain called Production and a domain called Production. The security domain has been mapped to two tenants, named Prod1 and Prod2. In the security domain assignment for the user, the tenant-admin role was selected, and it has access privilege type Write. Which of the following items may the user still be unable to do?
  - a.** Create bridge domains, application profiles, and EPGs in the Prod1 tenant
  - b.** View basic objects within the common tenant
  - c.** Map an EPG in the Prod2 tenant to a port on a leaf
  - d.** Create service graphs in the Prod2 tenant
- 2.** Which of the following roles may be most suitable for a user who needs basic visibility into ACI for the purpose of monitoring and troubleshooting?
  - a.** ops
  - b.** fabric-admin
  - c.** nw-svc-admin

- d.** tenant-admin
- 3.** A user needs full read-only visibility into an ACI fabric. Which predefined security domain can be used to enable such visibility?
- a.** common
  - b.** all
  - c.** infra
  - d.** fabric
- 4.** True or false: RBAC rules can be used to explicitly deny an offending user access to a portion of the ACI object hierarchy.
- a.** True
  - b.** False
- 5.** True or false: A privilege enables access to a particular function within a system. A role is simply a collection of privileges.
- a.** True
  - b.** False
- 6.** When creating an RBAC rule, which three parameters must be entered?
- a.** Object name, domain, and associated user
  - b.** DN, domain, and whether the rule grants write access
  - c.** Object name, domain, and security domain
  - d.** DN, security domain, and whether the rule grants write access
- 7.** Which role is best suited for a user who needs to manage access policies available under the Fabric menu?
- a.** tenant-admin

- b.** read-only
- c.** access-admin
- d.** fabric-admin

**8.** How can users who are successfully authenticated against LDAP be authorized for ACI access? (Choose all that apply.)

- a.** The LDAP server can be configured to assign users specially formatted Cisco AV pairs and return them in the queries ACI runs against the schema.
- b.** ACI cannot use LDAP data for authorization purposes.
- c.** ACI provides users a basic level of read-only access to the common tenant if Remote User Login Policy has been set to Assign Default Role and authorization data is missing.
- d.** ACI can be configured to map user group membership to the desired user access levels.

**9.** What Cisco AV pair authorizes a user to create an EPG in a tenant named Prod?

- a.** shell:domains = all//admin
- b.** shell:domains = all/aaa/tenant-admin
- c.** shell:domains = all/admin/
- d.** domains:shell = admin//all/

**10.** True or false: When Default Authentication Realm has been set to an external AAA service, and all AAA providers are unavailable to service requests or respond to ICMP traffic, users need to re-initialize the fabric to regain management access.

- a.** True
- b.** False

## Foundation Topics

### Implementing Role-Based Access Control (RBAC)

RBAC is a method of restricting and authorizing access to a system based on the roles of individuals within a company. For ACI to grant a user management access to fabric resources, that user must be assigned the following parameters:



- **One or more security domains:** A security domain identifies the portions of the hierarchical object tree the user can access.
- **One or more roles:** A role is a collection of privileges and determines the set of actions the user can take within the scope of the security domains.
- **An access level for each role:** The access level or access privilege type indicates whether the user has read-only or read/write access to carry out the functions associated with the role within the specified subtrees of the ACI object hierarchy.

In addition to local user authentication, ACI supports external user authentication via LDAP, TACACS+, RADIUS, SAML, and RSA.

The majority of configurations related to RBAC can be performed via the GUI in subtabs or submenus of the AAA page by navigating to **Admin > AAA**.

# Understanding Security Domains

ACI objects are designed to enable the templatization of configurations, where possible, to encourage future policy reuse. A security domain is a form of template that enables administrators to identify desired subtrees of the ACI object hierarchy to which one or multiple users should be granted access.



In more technical terms, a **security domain** is a tag that references one or more subtrees in the ACI object hierarchy. An ACI administrator who creates a security domain can then assign it to tenants and domains. For more granular references, an administrator can map an object subtree using the parent distinguished name (DN) to a security domain. A user who is subsequently mapped to the security domain gains a level of access to the referenced subtrees of the object hierarchy.

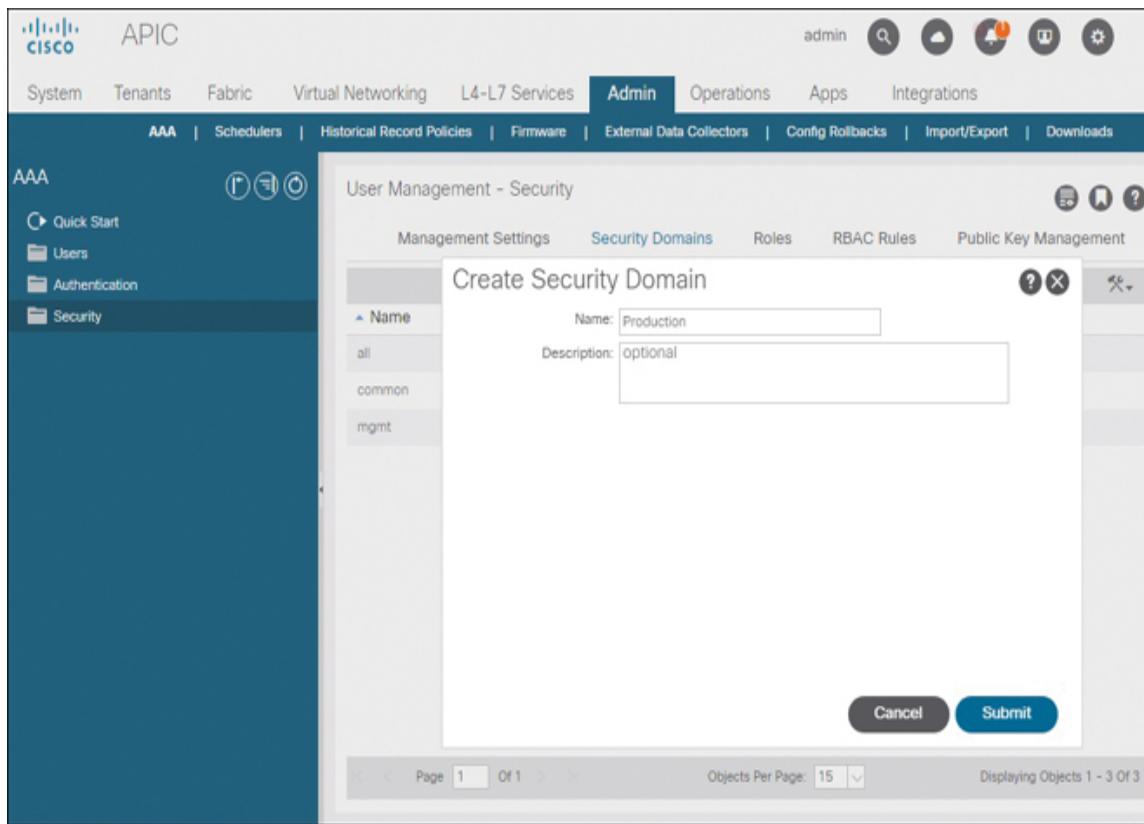
Remember that the level of access to the security domain subtrees that is afforded to a user is determined by the roles and access privilege types for the roles assigned to each individual user. Therefore, it is important to understand that security domain assignment by itself does not enable access rights.

By default, an ACI fabric includes several special predefined security domains that can be used for access assignment to users:



- **all:** Allows access to the entire ACI object hierarchy
- **common:** Allows access to the common tenant
- **mgmt:** Allows access to the management tenant

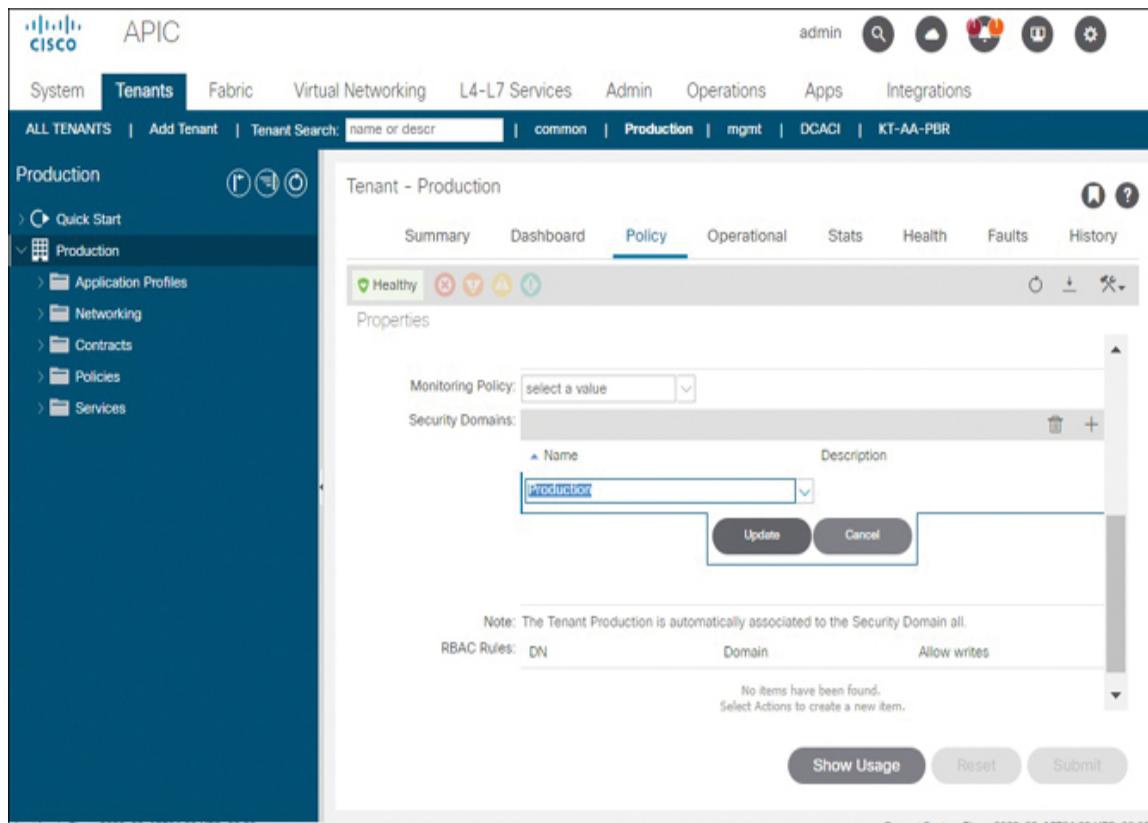
[Figure 15-1](#) shows the creation of a new security domain called Production via the Create Security Domain window, which you can access by navigating to **Admin > AAA > Security > Security Domains** and selecting Create Security Domain from the pull-down menu.



**Figure 15-1** Creating a New Security Domain

After creating a security domain tag, you need to assign the desired tenants, domains, and any other taggable objects to the security domain. [Figure 15-2](#) shows how you can navigate to the Policy subtab of a tenant to add one or more security domains to the tenant. To add a security domain

called Production, you enter the name and click Update to accept the change. In this way, you effectively assign the object tree for the tenant named Production to the security domain called Production.



**Figure 15-2** Assigning a Tenant to a Security Domain

For a user to gain access to the tenant named Production, the user needs to be assigned to the Production security domain using a role and access level that provides the ability to carry out the desired tenant functions.

## Understanding Privileges and Roles



A **privilege** enables access to a particular function within a system. An ACI fabric enforces access privileges at the managed object (MO) level.



A **role** is a collection of privileges. ACI has a set of predefined roles. An ACI administrator can modify the predefined roles or expand on default ACI roles by creating custom roles.

You can explore the set of privileges assigned to each role by navigating to **Admin > AAA > Security > Roles**. Figure 15-3 shows that the fabric-admin role is a collection of 12 privileges.

A screenshot of the Cisco Application Centric Infrastructure (ACI) web interface. The top navigation bar includes tabs for System, Tenants, Fabric, Virtual Networking, L4-L7 Services, Admin (which is selected), Operations, Apps, and Integrations. Below the Admin tab, there are links for AAA, Schedulers, Historical Record Policies, Firmware, External Data Collectors, Config Rollbacks, Import/Export, and Downloads. On the left, a sidebar under the AAA heading shows Quick Start, Users, Authentication, and Security (which is selected). The main content area is titled "User Management - Security" and displays a table of roles. The table has columns for Name, Privileges, and Description. The "Roles" tab is selected. The table shows the following data:

Name	Privileges	Description
admin	admin	
fabric-admin	fabric-connectivity-i1 fabric-connectivity-i2 fabric-connectivity-i3 fabric-connectivity-mgmt fabric-connectivity-util fabric-equipment fabric-protocol-i1 fabric-protocol-i2 fabric-protocol-i3 fabric-protocol-mgmt fabric-protocol-ops fabric-protocol-util	
nw-svc-admin	nw-svc-device nw-svc-devshare nw-svc-policy	
nw-svc-params	nw-svc-params	

At the bottom of the table, there are buttons for Page, Objects Per Page (set to 15), and a message indicating 11 objects are displayed.

### **Figure 15-3 Exploring the Privileges Assigned to Each Role**

The fabric-equipment privilege, for example, is one of the privileges ACI assigns to the fabric-admin role; it is used for atomic counter, diagnostic, and image management policies on leaf and spine switches within ACI.

Another common privilege is tenant-security, which enables users who are assigned the tenant-admin role to create contracts within tenants.

#### **Note**

As of the time of writing, a total of 62 privileges can be found in ACI. It is unlikely that DCACI candidates will be expected to know the function of each and every privilege; however, privilege names are somewhat descriptive of the functions that they enable. DCACI candidates *are* encouraged to create various users with different privileges to better understand the privileges available in ACI.

[Table 15-2](#) describes the predefined ACI roles that can be assigned to users.



**Table 15-2** Predefined Roles in ACI

Role Name	Description
Guest	Provides basic access to the network without any privileges.
Network Operator	Provides access to basic network configuration and monitoring.
System Auditor	Provides access to system logs and audit information.
System Administrator	Provides full administrative access to the system, including configuration and management of network resources.

aaa	Aids in configuring authentication, authorization, accounting, and import/export policies
access-admin	Enables administration and configuration of access policies
admin	Provides full access to an ACI fabric
fabric-admin	Enables administration and configuration of fabricwide settings and also firmware management
nw-service-admin	Allows users to configure L4-L7 network service insertion and orchestration
nw-service-parameter-ms	Grants access to the parameters governing the configuration of external L4-L7 devices

ops	Provides network operator privileges to an ACI fabric to allow for monitoring and troubleshooting functionality in ACI
read-all	Provides read-only visibility into an ACI fabric
tenant-admin-in	When assigned to a limited security domain, allows configuration of most attributes inside a tenant but does not allow changes to fabricwide settings that can potentially impact other tenants
tenant-ext-admin-in	Allows the configuration of external connectivity, such as L3Outs, for ACI tenants; a subset of the tenant-admin role
vm-management-admin-in	Grants access to ACI integrations with virtualization environments such as Microsoft Hyper-V, OpenStack, and VMware vSphere

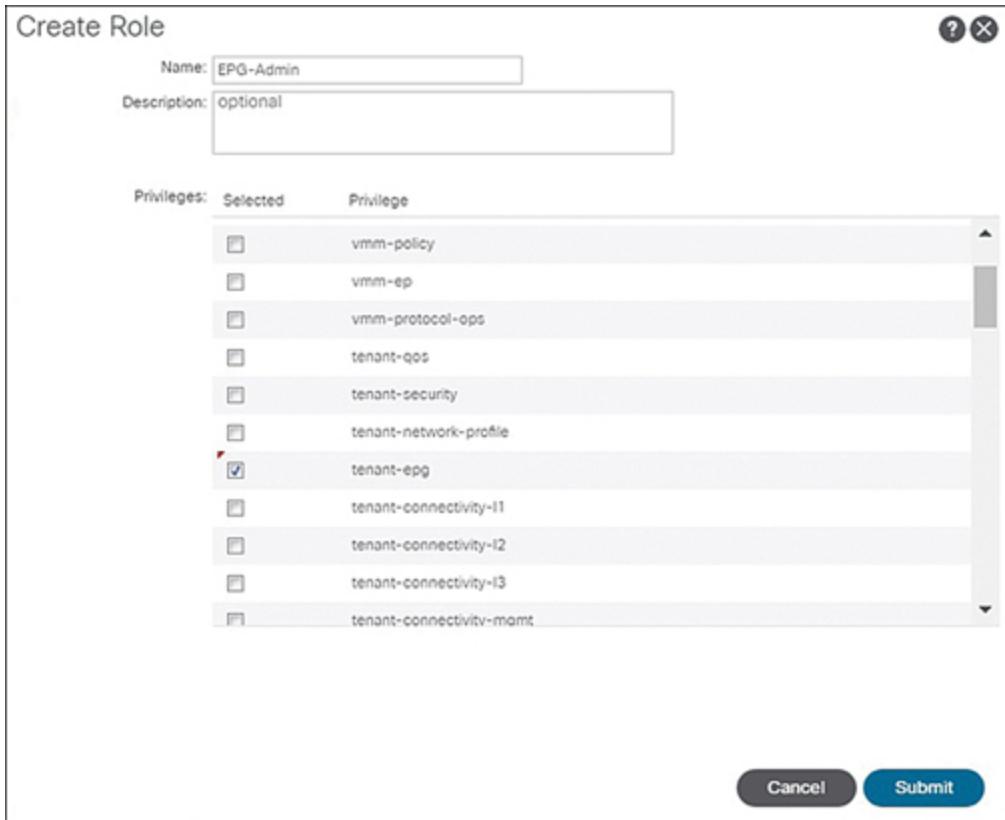
While you cannot create custom privileges, you can define custom roles if the out-of-the-box roles outlined in [Table 15-2](#) do not fit your requirements. [Figure 15-4](#) shows how you can navigate to the Create Role page by going to **Admin >**

**AAA > Security > Roles** and clicking on the Create Role option from the pull-down menu.

The screenshot shows the Cisco Application Policy Infrastructure Controller (APIC) interface. The top navigation bar includes tabs for Admin, Operations, Apps, and Integrations. Below the main navigation, there are links for AAA, Schedulers, Historical Record Policies, Firmware, External Data Collectors, Config Rollbacks, Import/Export, and Downloads. On the left, a sidebar under the AAA heading lists Quick Start, Users, Authentication, and Security. The main content area is titled "User Management - Security" and has tabs for Management Settings, Security Domains, Roles, RBAC Rules, and Public Key Management. The Roles tab is selected, showing a table with columns: Name, Privileges, Description, and a "Create Role" button. The table lists three roles: aaa, access-admin, and admin. The aaa role has many privileges listed under it. The admin role also has a list of privileges. At the bottom of the table, there are pagination controls (Page 1 of 1), an objects per page dropdown set to 15, and a message indicating 11 objects displayed.

**Figure 15-4** Navigating to the Create Role Page

[Figure 15-5](#) show how a new role called EPG-Admin can be created from the Create Role page through assignment of the tenant-EPG privilege. In a role with this privilege, a user can create EPGs, delete them, and bind them to domains from within the tenant view, map them to encapsulations, and assign them to ports (to which the user has visibility).



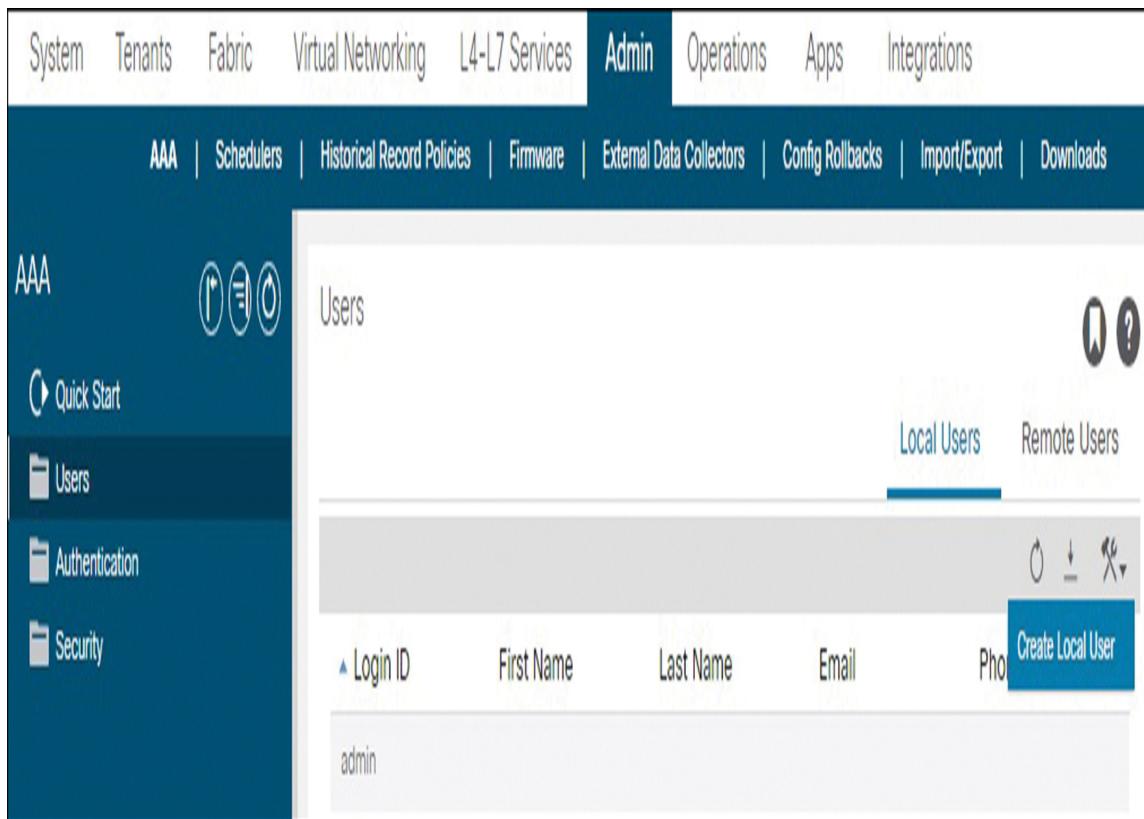
**Figure 15-5** Creating a Custom Role

## Creating Local Users and Assigning Access

Let's say that an IT department has just hired a new engineer named Bob. The IT department has decided that Bob should be allowed access to create and delete EPGs within the tenant named Production but should not be allowed access to modify any other objects within the tenant. After careful consideration, IT has decided that Bob can also be provided complete read-only access to the Production tenant and full access to modify any object within another tenant called Development. Note that these requirements are very basic and do not necessitate the creation of custom RBAC roles. These requirements can be

enforced using the read-all and tenant-admin roles, respectively.

A security domain that grants access to the tenant called Production has already been created (refer to [Figure 15-2](#)), and a custom RBAC role called EPG-Admin has also been created (refer to [Figure 15-5](#)), but a new security domain and a reference to the security domain by the Development tenant is also needed. Once this new security domain is created, the AAA administrator can create a user for Bob and begin to allocate the required access. [Figure 15-6](#) shows how you navigate to **Admin > AAA > Security > Users** and select Create Local User from the pull-down menu to create a user for Bob.



**Figure 15-6** Opening the *Create Local User Wizard*

In the Create Local User window, you select a user ID and password for Bob. [Figure 15-7](#) shows the first page of the Create Local User wizard, where you enter user identity data.

The screenshot shows the 'Create Local User' wizard with the title 'Create Local User' at the top. It is on 'STEP 1 > User Identity'. The 'User Identity' tab is selected, indicated by a blue arrow. There are three tabs in total: '1. User Identity', '2. Security', and '3. Roles'. The 'User Identity' tab has a sub-section titled 'User Certificate Attribute' which is currently empty. Below the certificate attribute are several input fields: 'First Name', 'Last Name', 'Phone', and 'Email', each with a corresponding text input box. There is also a field for 'Description' with the placeholder 'optional'. Under the 'User Identity' section, there are three toggle buttons: 'Account Status' (Active is selected), 'Account Expires' (No is selected), and 'Enable Restricted User' (No is selected). At the bottom of the window are three buttons: 'Previous', 'Cancel', and 'Next'.

**Figure 15-7** Entering User Identity Information in the Create Local User Page

Most of the fields in the User Identity page are self-explanatory. However, the following points are worthy of note:

- **Login ID requirements:** Login IDs can be up to 32 characters long. Letters, numbers, underscores, dashes, and dots are acceptable characters for login IDs.
- **Default ACI password requirements:** By default, the ACI password strength check necessitates that passwords be at least eight characters in length and

have fewer than three consecutive repeated characters. Passwords can contain lowercase or uppercase letters, digits, and symbols. Passwords cannot be those that can be easily predicted using a dictionary attack, cannot be permutations of words like cisco, and cannot be the reverse of the username. You can disable password strength check for non-production use cases by using the **no password pwd-strength-check** APIC CLI command, or you can modify the strength check to reflect corporate password strength policies. To create a custom policy, navigate to **Admin > AAA > Security > Management Settings** and select Create Password Strength Policy from the pull-down menu. If password strength check has been disabled via the APIC CLI, the implementation of a modified password strength policy requires that the password strength check be reenabled first.

- **User Certificate Attribute setting:** This is the client certificate user identity used to enable certificate-based authentication.
- **Account Status setting:** This parameter determines whether the user is active and able to log in to the fabric or is temporarily disabled. The account status field can be toggled whenever there is a need to manually disable or enable an account.
- **Account Expires setting:** By setting the Account Expires option to Yes, you can define an expiration date and time for the user object. ACI does not delete the user account when the expiration date is hit, but it does disable logins into the fabric from the expired user account.

As shown in [Figure 15-8](#), the second page in the Create Local User wizard is where you assign security domains to a

user. Because the requirements for the user Bob involved both the Development tenant and the Production tenant, both security domains should be added here.



Create Local User

STEP 2 > Security

1. User Identity    2. Security    3. Roles

Security Domain:

Name	Description
all	
common	
<input checked="" type="checkbox"/> Development	
mgmt	
<input checked="" type="checkbox"/> Production	

User Certificates:

Name	Expiration Date	State

SSH Keys:

Name	Key

Previous    Cancel    Next

A screenshot of a web-based user creation interface. The title is 'Create Local User'. The current step is 'STEP 2 > Security'. There are three tabs: '1. User Identity', '2. Security' (which is active and highlighted in blue), and '3. Roles'. The 'Security Domain' section contains a table with columns 'Name' and 'Description'. It lists several domains: 'all', 'common', 'Development' (with a checked checkbox), 'mgmt', and 'Production' (with a checked checkbox). Below this are sections for 'User Certificates' and 'SSH Keys', each with a table and a '+' button to add new entries. At the bottom are 'Previous', 'Cancel', and 'Next' buttons.

**Figure 15-8 Assigning a User to Security Domains**

After mapping the user Bob to the two security domains, you need to fulfill the access requirements by assigning Bob the custom EPG-Admin role for the Production security domain using the access privilege type Write. Bob also needs read-only access to the entire tenant named Production. This can be accomplished by mapping the Production security domain to the role named read-all with access privilege type Read. Bob also needs full access to the

Development tenant, which can be accomplished by mapping Bob to the role tenant-admin and the Write access privilege type. These three changes are reflected in [Figure 15-9](#). Click Finish to execute the user creation.



Create Local User

STEP 3 > Roles

1. User Identity    2. Security    3. Roles

Domain Production:

Role Name	Role Privilege Type
EPG-Admin	Write
read-all	Read

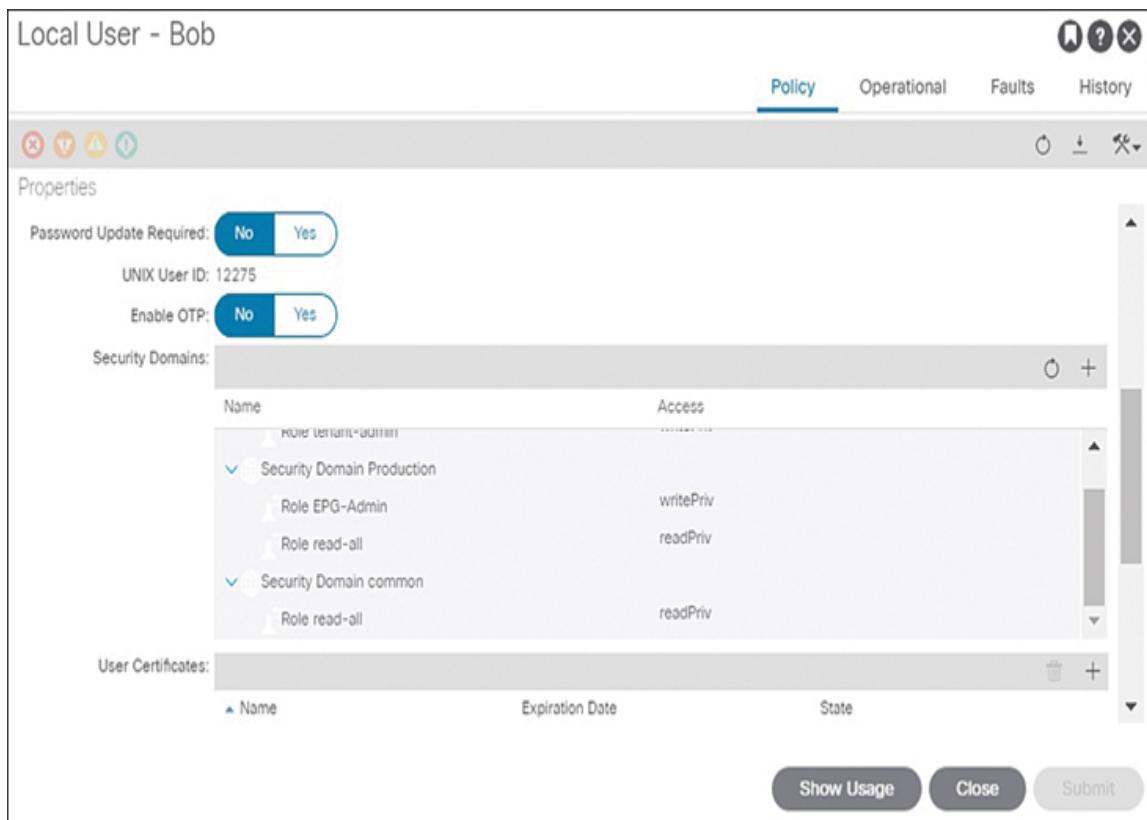
Domain Development:

Role Name	Role Privilege Type
tenant-admin	Write

Previous    Cancel    **Finish**

**Figure 15-9** *Mapping Access Levels to Security Domains for a Local User*

Notice in [Figure 15-10](#) that the user Bob has been assigned to the security domains, roles, and access privilege types specified. However, he has also been assigned read-only access to the common tenant.

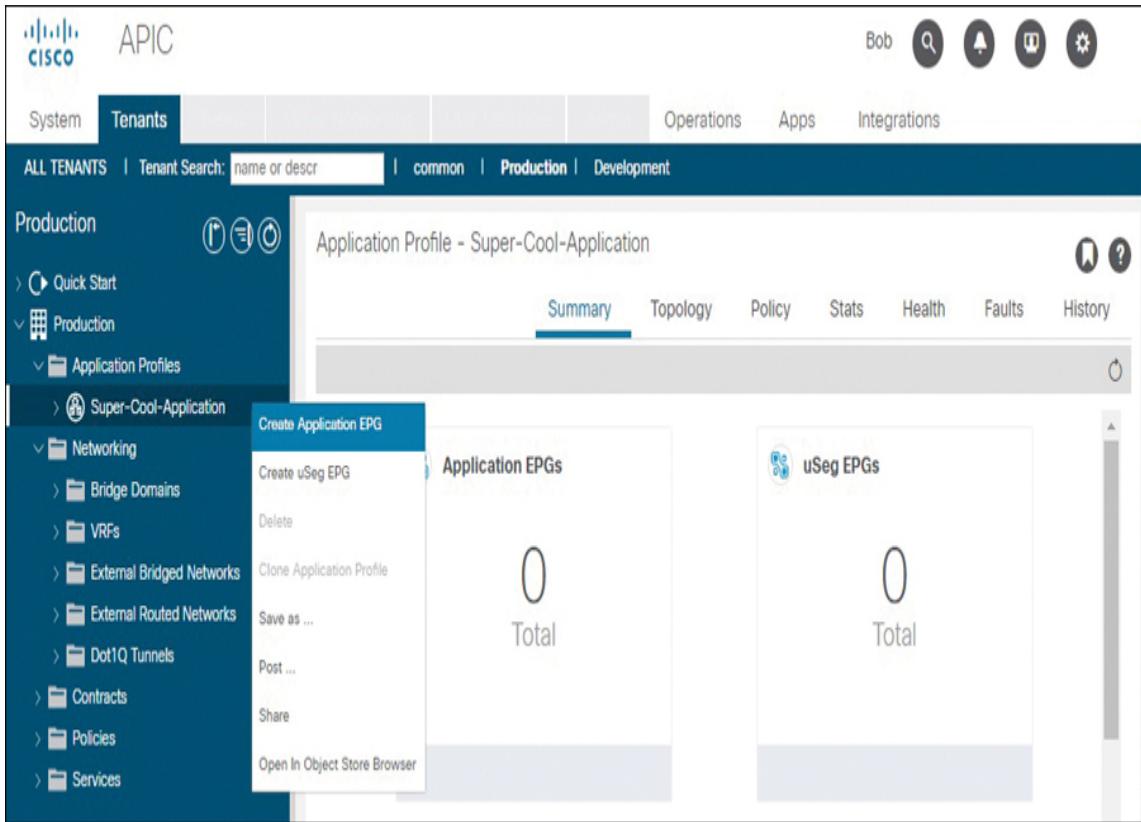


**Figure 15-10** ACI Grants Local Users Read-Only Access to the Common Tenant

**Key Topic**

ACI automatically grants locally created tenant-focused ACI users read-only access to the common tenant because objects within the common tenant can be consumed by all other tenants.

When the user Bob logs in to ACI, he realizes that he does not have visibility into all the menus within the fabric. [Figure 15-11](#) illustrates that Bob does not have access to the Fabric menu, for example. [Figure 15-11](#) also shows that the user Bob can create EPGs within the Production tenant, as indicated by the fact that the Create Application EPG menu option is not grayed out.



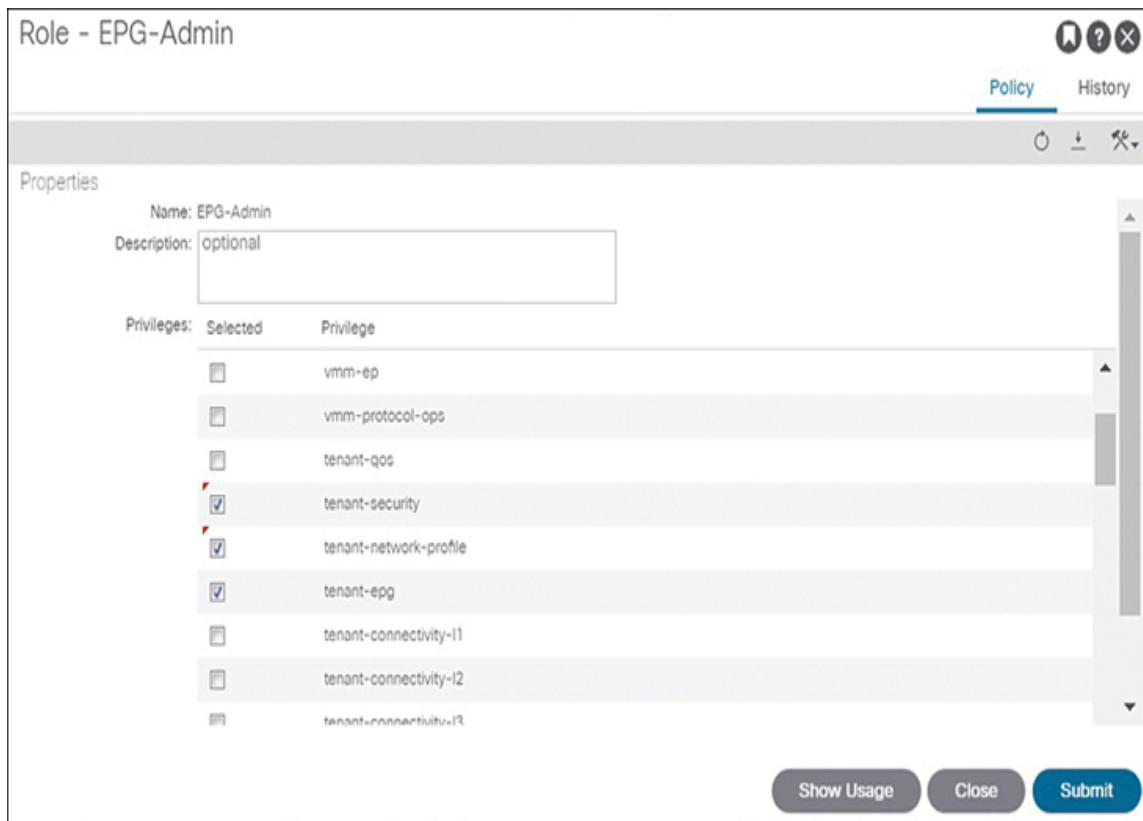
**Figure 15-11** Verifying Limited Visibility in Line with Assigned Access

## Tweaking Roles and User Access

Bob currently has the access he needs. If the organization someday decides, however, that users like Bob need expanded access to the Production tenant, the custom EPG-Admin role can be modified to include a broader number of privileges. This can pose some new problems if not all the users that have already been assigned to the EPG-Admin role are expected to have such elevated privileges. If this is the case and the EPG-Admin role cannot be modified, a new role should be created with elevated privileges, and it should be mapped to Bob and any other relevant users.

[Figure 15-12](#) shows how the IT department could elevate users like Bob to a whitelisting and security administrator

role by expanding the privileges associated with the custom EPG-Admin role to also include the tenant-security and tenant-network-profile privileges. After this change, users like Bob are able to not only create new application profiles but also create filters, filter entries, subjects, and contracts. They are also able to assign the security constructs they create to EPGs within the tenants to which they have been assigned write access via the EPG-Admin role.

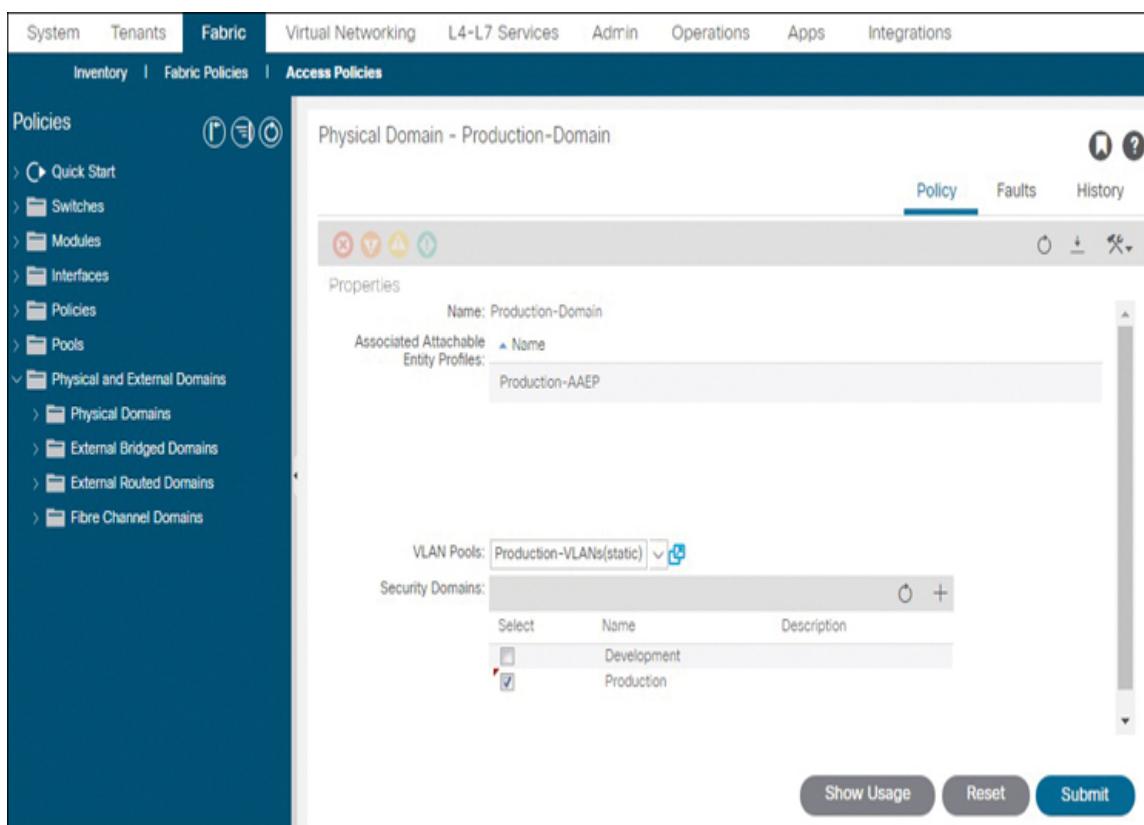


**Figure 15-12** Expanding Privileges Associated with a Role

Let's say that tenant-focused users like Bob are also expected to be able to map EPGs to physical ports and encapsulations in the ACI fabric. There are two obstacles to achieving this level of access. First, Bob and like users have not yet been mapped to a security domain that grants them visibility into the relevant domain subtrees of the ACI

hierarchical object tree. Second, even if they did have visibility to a domain, they would also need to have visibility to the port configurations to be able to select the port to which the EPG should be assigned.

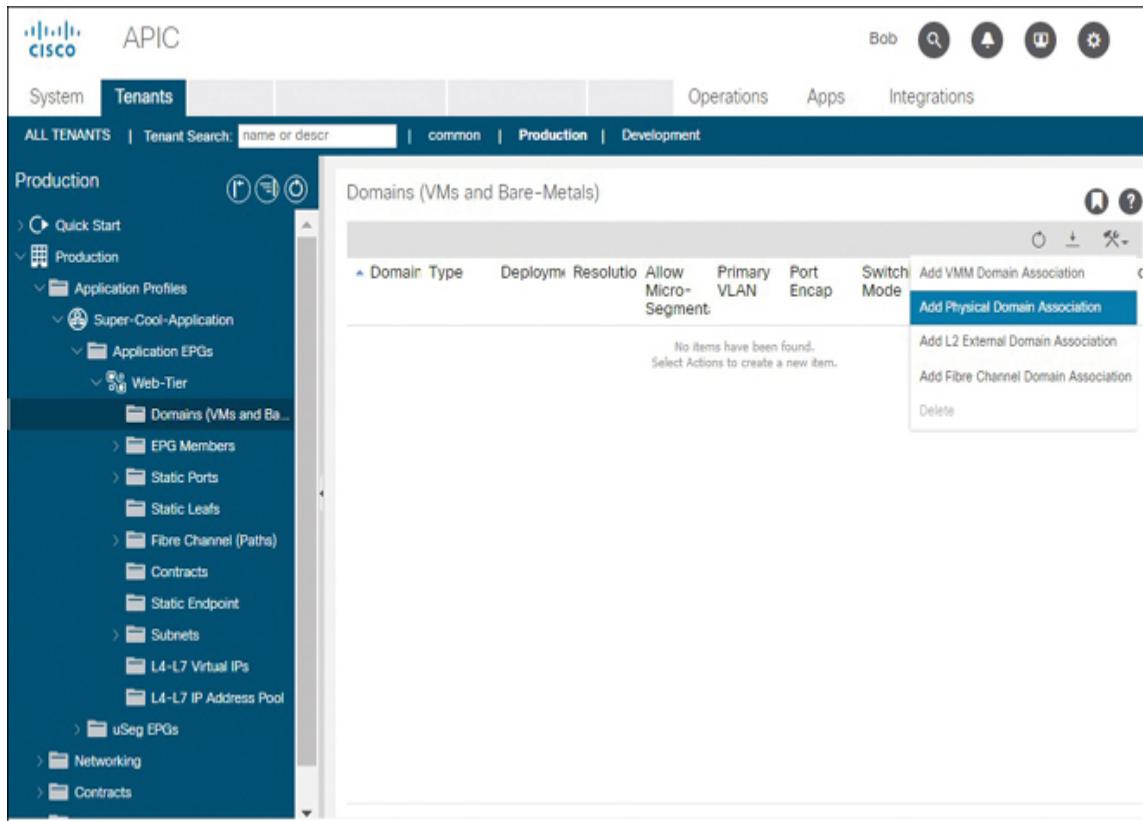
The problem of lack of access to a domain can be easily resolved. In [Figure 15-13](#), the fabric admin user navigates to the domain named Production-Domain and adds a reference to the security domain named Production, to which Bob has already been assigned.



**Figure 15-13** Mapping a Domain to a Security Domain

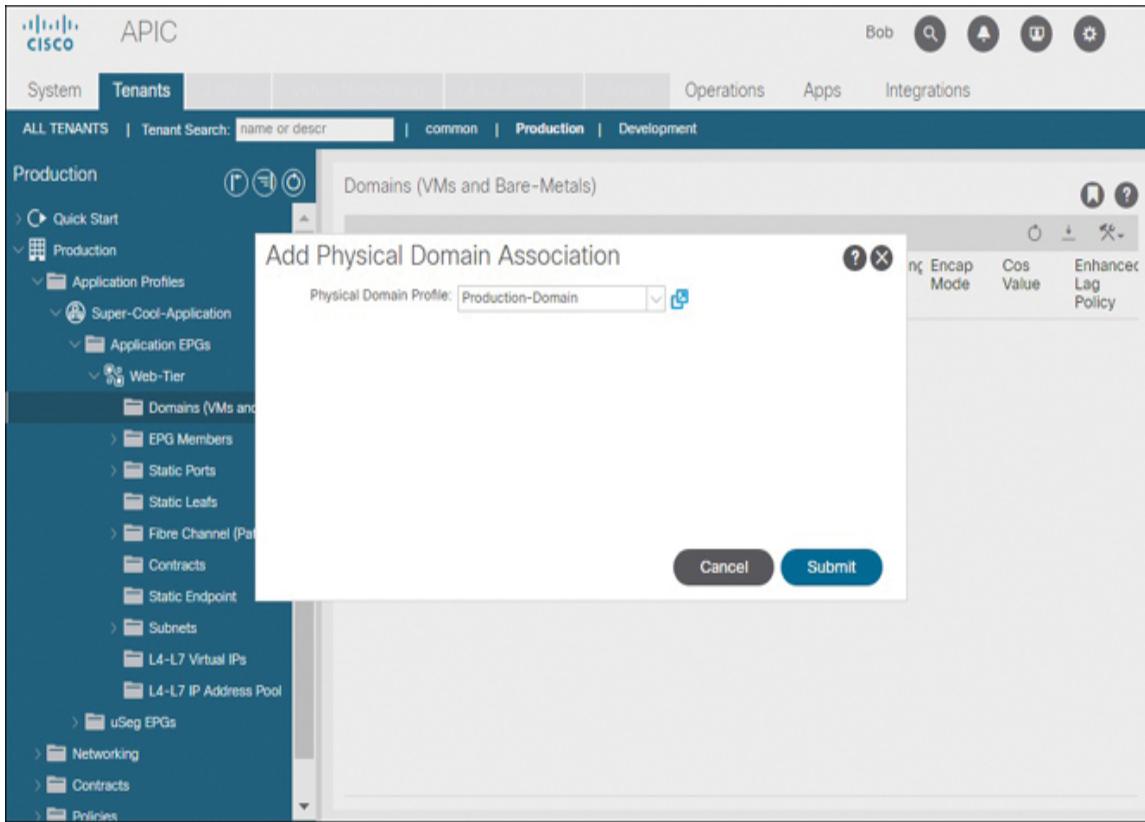
Once the domain named Production-Domain is mapped to the Production security domain, Bob is able to bind EPGs within the tenant named Production to the domain called Production-Domain. To bind an EPG to a domain, Bob navigates to **Tenants > Production > Application**

**Profiles** and then drills down to the relevant application profile, drills down into the relevant EPG, opens up the Domains folder, and selects Add Physical Domain Association from the pull-down menu (see [Figure 15-14](#)).



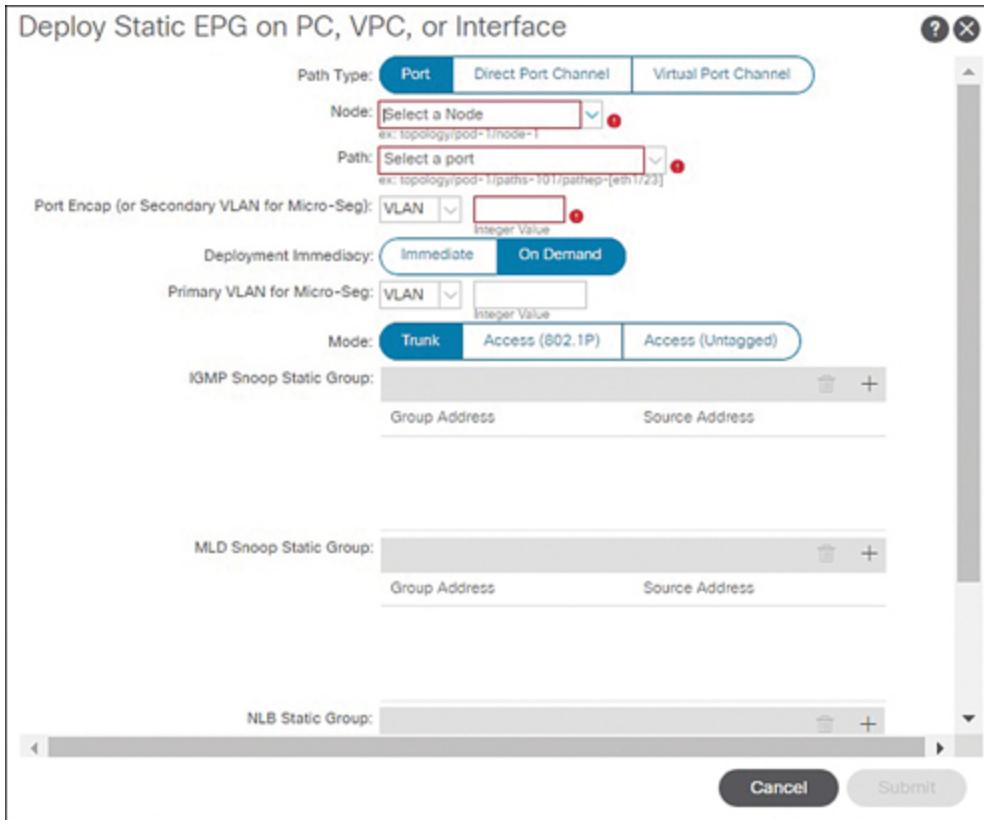
**Figure 15-14** Binding an EPG to a Domain in the Tenant View

The reason Bob selects Add Physical Domain Association and not the other pull-down menu options is that Bob intends to map the EPG to encapsulations for bare-metal servers. The process of enabling Bob to add VMM domain associations to EPGs is similar to the process shown previously. [Figure 15-15](#) shows that Bob now has visibility to the domain called Production-Domain, which he adds and clicks Submit.



**Figure 15-15** Binding a Domain to an EPG

While the first obstacle to allowing Bob to assign EPGs to ports has been resolved, and EPGs like the one called Web-Tier can now be bound to a domain, Bob is still unable to assign the EPGs in the Production tenant to ports. This, you will find, is due to lack of access privileges to the underlying infrastructure. [Figure 15-16](#) demonstrates this challenge and how Bob is unable to select ports to make EPG assignments even though he may be able to select a VLAN encapsulation now that a domain-to-security domain mapping is in place.



**Figure 15-16** Challenges with Lack of Visibility to Underlying Ports

One potential solution to this challenge could be to provide Bob read-only access to the predefined domain called all; however, this may not be an acceptable solution for service providers or in IT environments in which tenant access is granted to partners. Another possible solution is for ACI administrators to create a custom RBAC rule that enables more granular access to the specific portion of the ACI object hierarchy without exposing details of the fabric view to users like Bob.

## Custom RBAC Rules

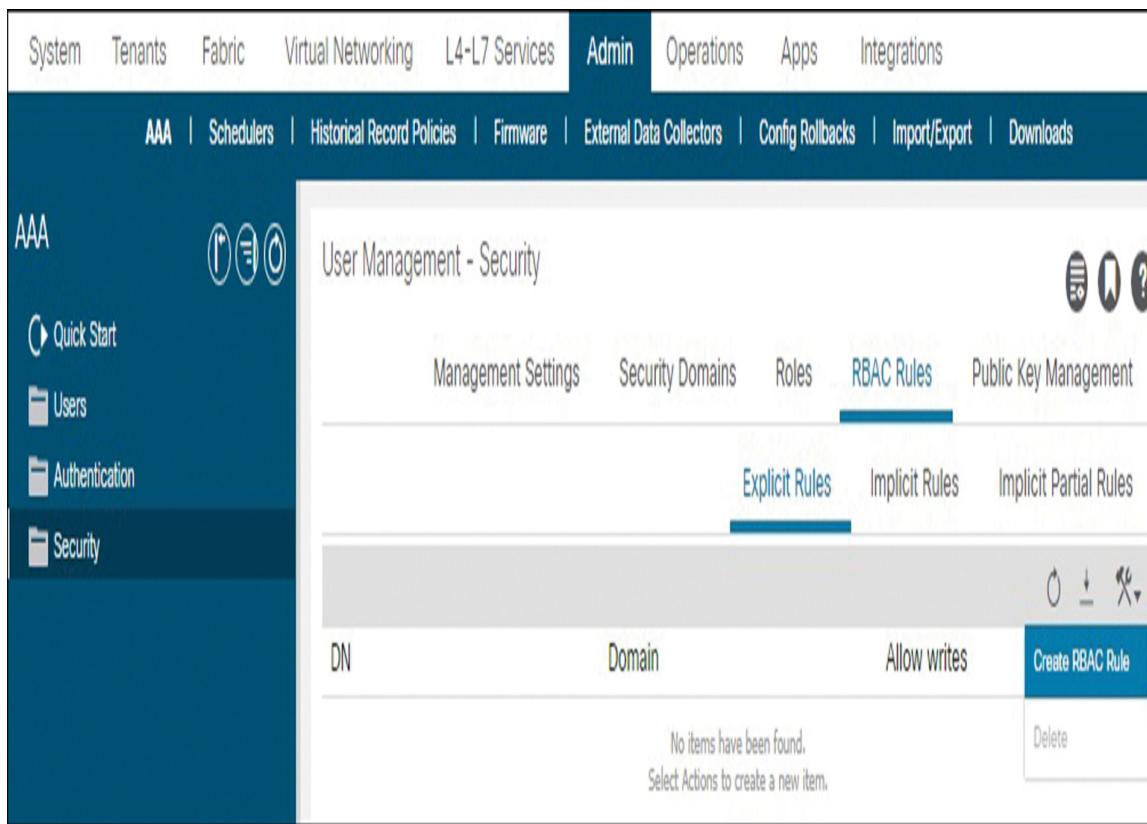
If your organization needs more granular RBAC over resources than is possible via privileges, roles, and security domains, you can create custom RBAC rules.

**Key Topic**

**RBAC rules** allow granular control on top of the existing RBAC framework and conform to the following principles:

- RBAC rules are additive and associative.
- RBAC rules cannot be used to deny access to portions of the ACI object hierarchy.
- Effective use of RBAC rules requires that you gain knowledge of the ACI object hierarchy and learn how to find DNs for any given object.
- When creating an RBAC rule, ACI only validates for DN format and not for the existence of the specified DN. This makes the pre-staging of RBAC prior to full policy deployment a possibility.
- During creation of an RBAC rule, you need to select whether users assigned to relevant security domains will have read-only access or full control of the object subtree.

For Bob and other tenant security administrators to be able to map EPGs to physical ports, an RBAC rule can be created to grant the security domain named Production read-only access to a portion of the ACI object hierarchy called topology. [Figure 15-17](#) shows how you navigate to **Admin > AAA > Security > RBAC Rules > Explicit Rules** and select Create RBAC Rule to open the Create RBAC Rule wizard.



**Figure 15-17** Navigating to the *Create RBAC Rule Wizard*

Figure 15-18 shows how you expose the topology subtree to all users who have access to the Production security domain without allowing such users to make any changes to the topology subtree.

**Key Topic**

Create RBAC Rule

DN: topology

Domain: Production

Allow Writes: No Yes

Cancel Submit

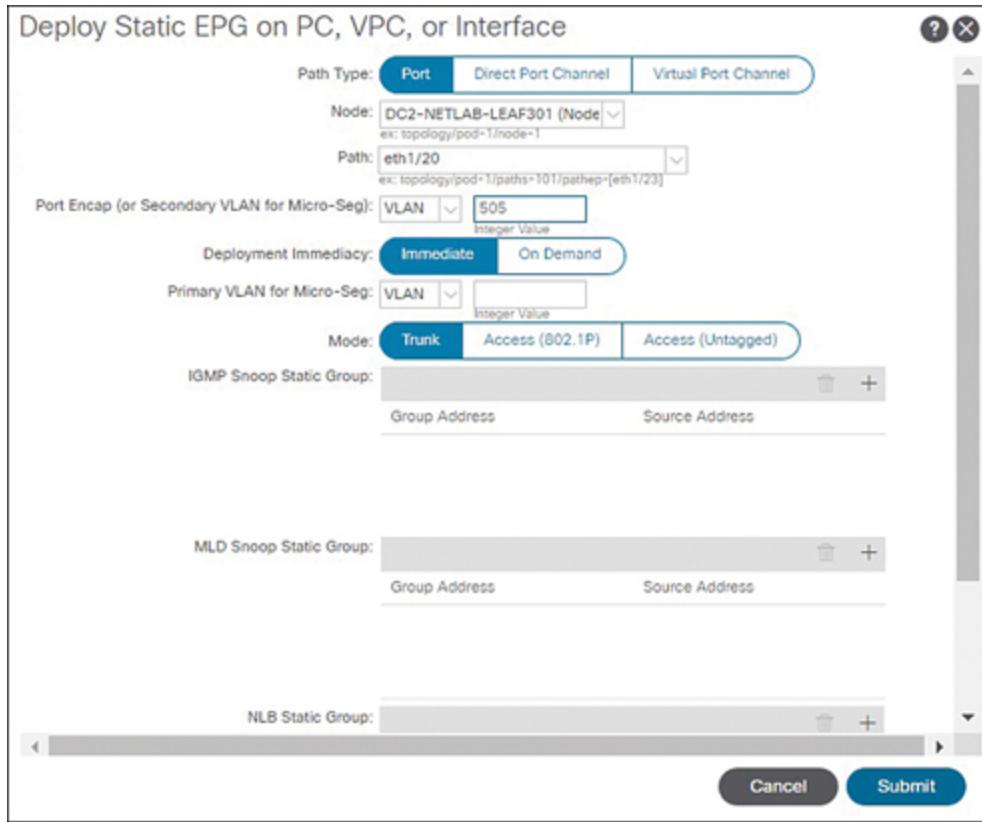
The screenshot shows a 'Create RBAC Rule' dialog box. At the top right are a question mark icon and a close button. Below that are three input fields: 'DN' containing 'topology', 'Domain' set to 'Production' with a dropdown arrow, and 'Allow Writes' with two options: 'No' (selected) and 'Yes'. At the bottom are 'Cancel' and 'Submit' buttons.

**Figure 15-18** Creating a Custom RBAC Rule

The user Bob needs to log out and log back in for access changes to take effect.

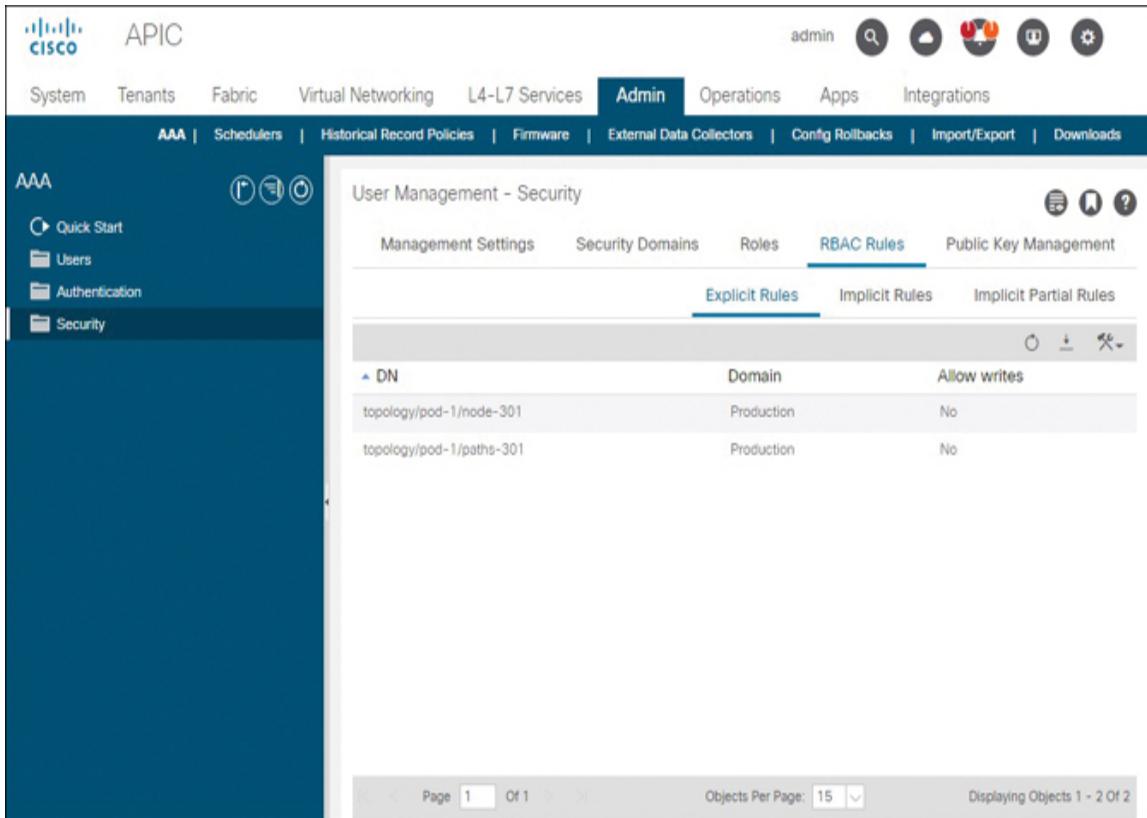
With the custom RBAC rule in place, Bob and other users assigned to the Production security domain should have access to map EPGs within the Production tenant to any physical server or appliance whose port configuration references the AAEP called Production-AAEP, using any encapsulation in the VLAN ID range defined by the VLAN pool called Production-VLANs.

[Figure 15-19](#) shows Bob trunking the EPG called Web-Tier to port 1/20 on a leaf switch identified as Leaf 301, using an encapsulation of VLAN ID 505.



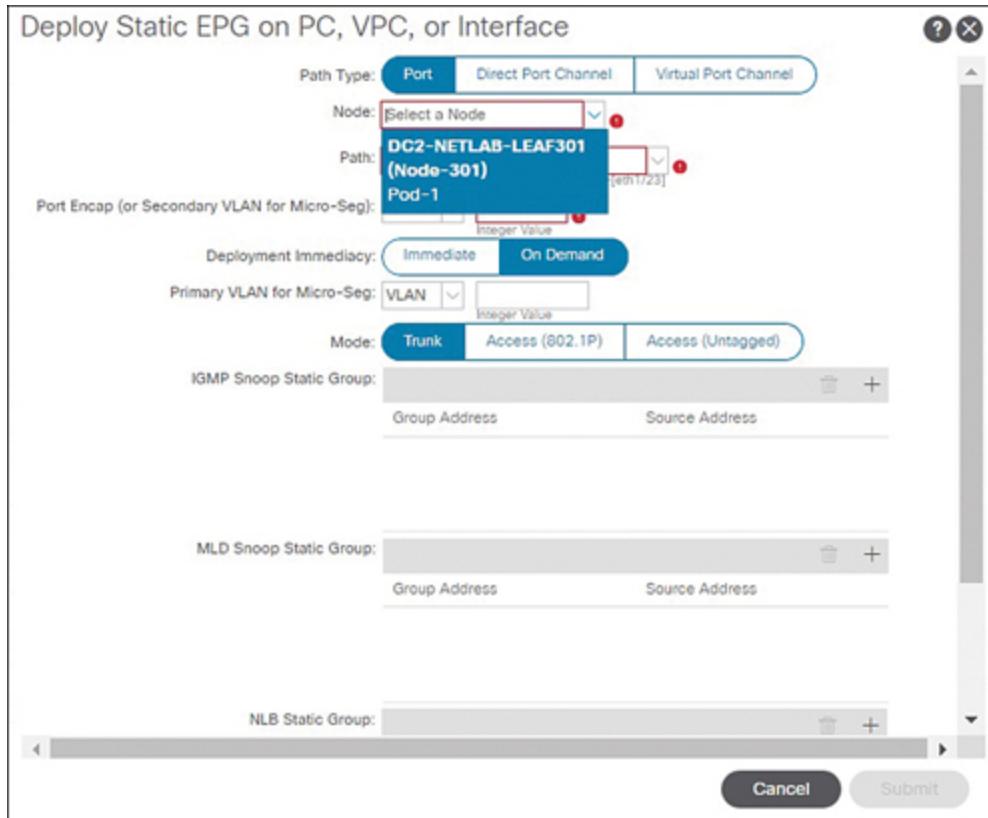
**Figure 15-19** Successfully Mapping an Encapsulation to a Port

If you do not want to grant read-only access to the entire topology hierarchy, you can assign access to a specific subset of fabric nodes. [Figure 15-20](#) demonstrates how alternative custom RBAC rules could be created to assign access to Node 301 and its paths (ports), thereby restricting the access of users like Bob to ports on Leaf 301.



**Figure 15-20** Custom RBAC Rule Allowing Visibility to Leaf 301

As a result of this change, Bob sees only Leaf 301 as an option for mapping encapsulations to underlying switch ports. Other leaf switches, such as Leaf 302, are unavailable to Bob (see [Figure 15-21](#)).



**Figure 15-21** User Visibility Restricted to Leaf 301 and Its Ports

## A Common RBAC Pitfall

Where security policies dictate that users be allocated the minimum level of access necessary to perform their job duties, it is especially important to define RBAC to fulfill the expected granularity requirements.

It is possible to use the predefined roles with security domains to enforce a good level of control, but it is just as important for administrators to fully understand the privileges that are being granted as a result of role assignments and to design RBAC accordingly.

For example, if a AAA administrator decides to allocate a user access to the tenant-admin role using the predefined

security domain all, it may be assumed that the user would only have access to tenant-specific functions due to the name of the role assigned. This assumption would be far from the truth. In reality, the tenant-admin role grants the majority of privileges available in an ACI fabric to the new user. When this is combined with the security domain all, the user has capabilities that come very close to the predefined admin user capabilities in an ACI fabric.

The safest approach to RBAC in environments that demand clear delineation between roles is to start with a restrictive approach and gradually add privileges and visibility to additional subtrees of the ACI object hierarchy when additional requirements are identified.

## Integrating with External AAA Servers

This chapter has so far touched on RBAC deployment using local users. This section addresses how ACI can integrate with external AAA servers via the TACACS+, RADIUS, and LDAP protocols.

### Note

The scope of the Implementing Cisco Application Centric Infrastructure DCACI 300-620 exam is limited to ACI-side configurations only. Knowledge of products such as Cisco Identity Service Engine (ISE) is not tested on the exam. Therefore, this book does not cover the low-level details of ISE configuration for ACI device administration but instead provides a high-level explanation of the ISE configuration process in an effort to avoid confusion around the authentication and authorization process.

# Configuring ACI for TACACS+

ACI configuration for TACACS+ involves three basic steps:



**Step 1.**Create the desired TACACS+ providers.

**Step 2.**If using ACI versions prior to Release 4 or configuring ACI via the APIC CLI, create a TACACS+ provider group.

**Step 3.**Create a TACACS+ login domain.

A TACACS+ provider is a reference to an individual TACACS+ server that details how ACI nodes communicate with the specified server. [Table 15-3](#) describes the configuration options available in a TACACS+ provider definition.



**Table 15-3** Configuration Parameters for TACACS+ Providers

## C Description

o  
n  
f  
i  
g  
u  
r  
a  
t  
i  
o  
n  
P  
a  
r  
a  
m  
e  
t  
e  
r

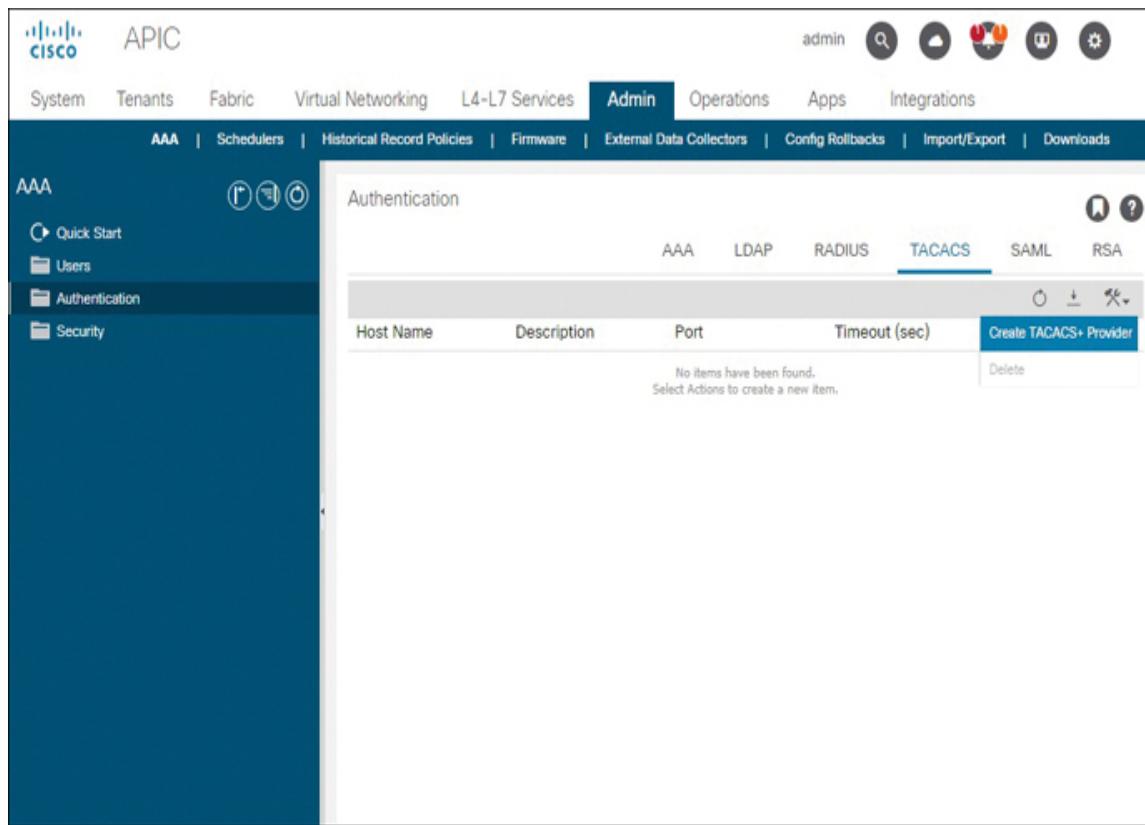
H ACI allows you to reference a TACACS+ server using  
os either its IP address or DNS address.

t  
N  
a  
m  
e  
or  
IP  
A  
d  
dr  
es  
s

Port	This is the TCP port number ACI needs to use to connect to the TACACS+ daemon. The range is from 1 to 65535. The default is 49.
Authorization Protocol	TACACS+ authorization protocols include PAP, MS-CHAP, and CHAP. The default TACACS+ authorization protocol is PAP.
Key	This is the shared secret key that ACI nodes and the TACACS+ server use for encrypting and decrypting traffic between one another.
Timeout	This is the number of seconds the ACI node waits for a response from the TACACS+ provider before timing out. The acceptable range is from 5 to 60 seconds. The default is 5 seconds.

Retries	<p>This is the number of times ACI automatically retries login attempts for a single authentication submission. The acceptable range is from 1 to 5 retries. The default is 1.</p>
Management EPG	<p>This is the management EPG (in-band or out-of-band) from which ACI should source requests to the TACACS+ server.</p>
Server Monitoring	<p>This parameter can be set to either Enabled or Disabled. When it is enabled for a TACACS+ provider, the APICs periodically attempt to execute login attempts against the TACACS+ provider to verify that the TACACS+ service is alive on the server. When enabled, ACI asks administrators to additionally enter a username and password. By default, Server Monitoring is set to Disabled. Server monitoring checks are exclusive to APICs. To enable monitoring checks that leafs and spines are also able to perform, use of the Ping Check option under the AAA Policy view is more common.</p>

To navigate to the TACACS+ provider creation page, click the Admin menu, select AAA, and then select Authentication followed by TACACS. Finally, right-click the Tools menu and select Create TACACS+ Provider, as shown in [Figure 15-22](#).



**Figure 15-22** Navigating to the Create TACACS+ Provider Page

[Figure 15-23](#) shows the creation of a TACACS+ provider that references a server at 10.233.48.60, using mostly default settings.

Create TACACS+ Provider

Host Name (or IP Address):

Description:

Port:

Authorization Protocol:  CHAP  MS-CHAP  PAP

Key:

Confirm Key:

Timeout (sec):

Retries:

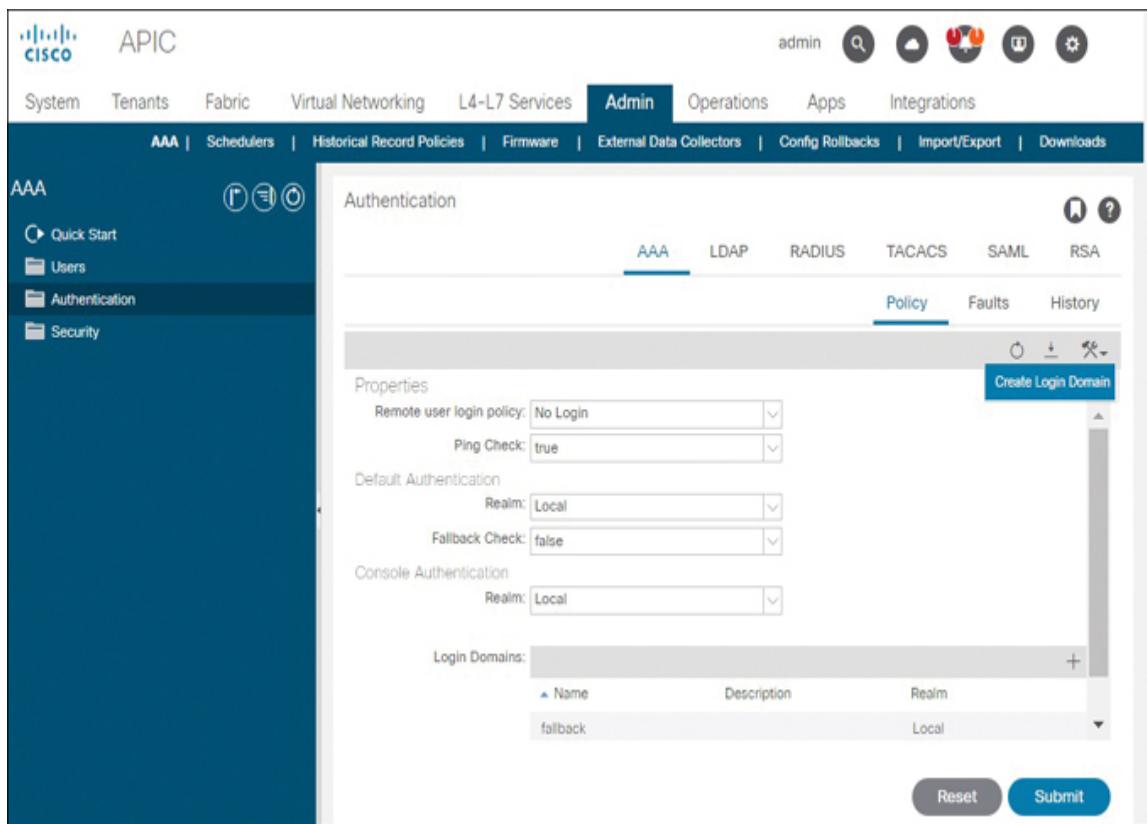
Management EPG:

Server Monitoring:  Disabled  Enabled

**Figure 15-23** Configuring a TACACS+ Provider

In this deployment, the ACI software version does not require creation of a TACACS+ provider group. Therefore, you can move on to step 3 of the process and create a login domain referencing any desired TACACS+ providers. It is only after a TACACS+ login domain is created that users are able to authenticate and authorize against TACACS+ servers.

To create a login domain, navigate to Admin, click AAA, select Authentication, choose the AAA submenu, select Policy, select the Tools menu, and click Create Login Domain, as shown in [Figure 15-24](#).



**Figure 15-24** Navigating to the Create Login Domain Page

On the Create Login Domain page, enter the name that is expected to appear in the login drop-down box, select the desired realm (the AAA method), and select the providers associated with the login domain. You can prioritize the use of a specific provider by using the Priority field. If ACI deems all listed providers as healthy via the Ping Check or Server Monitoring features, ACI attempts to authenticate against the AAA providers in the ranked priority configured. [Figure 15-25](#) shows a login domain configuration that prioritizes the TACACS+ server at 10.233.48.60.

Create Login Domain

Name: TACACS

Realm: TACACS+ ▾

Description: optional

Providers:

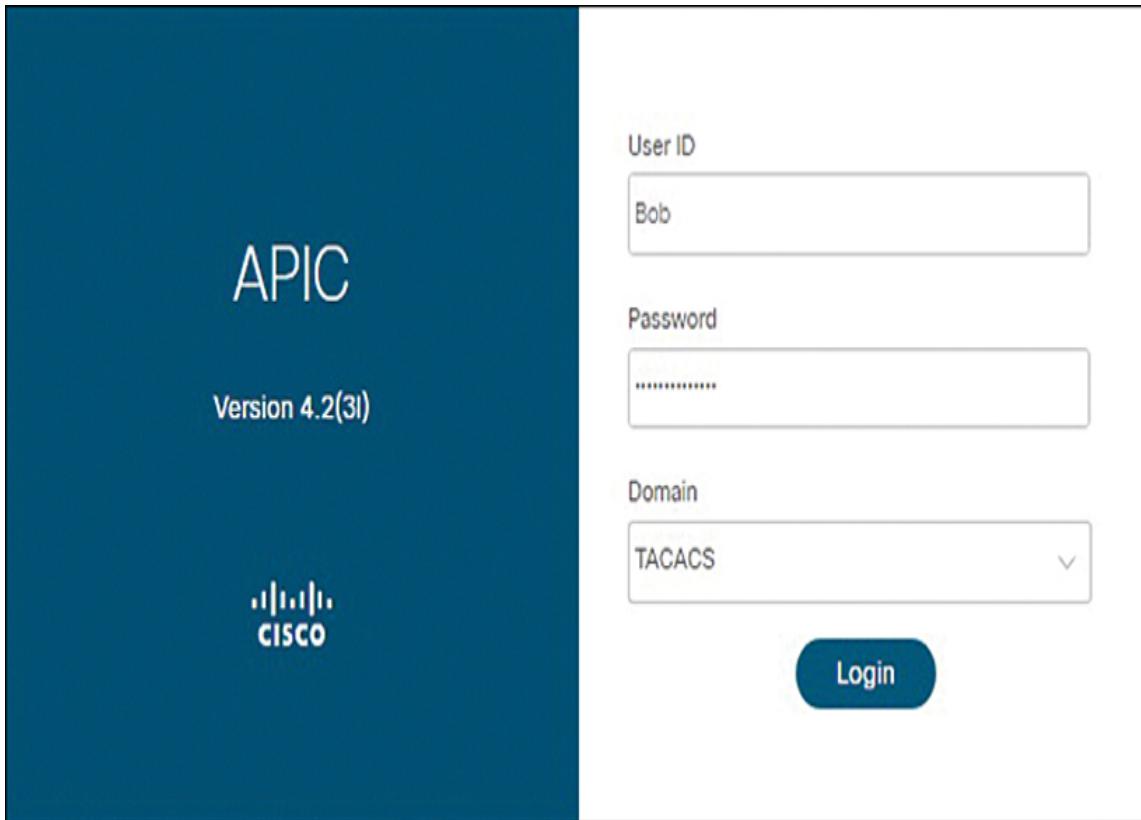
Name	Priority	Description
10.233.48.60	1	
10.233.64.60	2	

Cancel Submit

Name	Priority	Description
10.233.48.60	1	
10.233.64.60	2	

**Figure 15-25** Configuring a TACACS+ Login Domain

Once the TACACS+ login domain is created, users can select it from the Domain drop-down box to authenticate against the configured providers. [Figure 15-26](#) shows the user Bob trying to authenticate against the TACACS+ login domain.



**Figure 15-26** Selecting a Custom Domain During GUI Login

The GUI-based configurations presented in [Figure 15-22](#) through [Figure 15-25](#) can also be done using the APIC CLI. [Example 15-1](#) presents the commands needed.

**Example 15-1** Configuring ACI for TACACS+ via the APIC CLI

[Click here to view code image](#)

```
APIC1(config)# show run
(...output truncated for brevity...)
tacacs-server host "10.233.48.60"
exit
tacacs-server host "10.233.64.60"
exit
aaa group server tacacsplus TACACS
```

```
server 10.233.48.60 priority 1
server 10.233.64.60 priority 2
exit
aaa authentication login domain TACACS
realm tacacs
group TACACS
exit
```

Note in this example that new ACI software still groups TACACS servers together behind the scenes, but the GUI abstracts this step from users. Also, notice that the key values involving passwords do not appear in **show** command output.

## Configuring ISE to Authenticate and Authorize Users for ACI

The high-level process for configuring ISE to allow authentication and authorization of ACI users is as follows:

### Note

In these steps, references to menu items are valid for ISE Release 2.7. Other releases of ISE may require slightly different steps.

**Step 1. Enable Device Admin Service:** For ISE to perform TACACS+ operations, ensure that the Enable Device Admin Service checkbox is enabled under **Administration > System > Deployment**.

**Step 2. Configure AAA clients:** Navigate to **Administration > Network Resources > Network Devices** and create entries for ACI nodes. It is recommended that a specific device

type be defined for ACI nodes under **Administration > Network Resources > Network Device Groups** beforehand. Configuring a specific network device group for ACI nodes enables ISE administrators to group ACI authentication and authorization rules based on a device type condition.

**Step 3. Configure user identity groups:** Think of user identity groups in the context of ACI authorization as a grouping of users with the same expected level of access. The user Bob, for instance, is a junior administrator who is expected to have different levels of access for several different security domains. For example, he might be categorized into a junior-admin user identity group. To fulfill the needs of a hypothetical deployment that requires a group of ACI administrators, a group of ACI network operators, and a group of junior engineers, three user identity groups can be used.

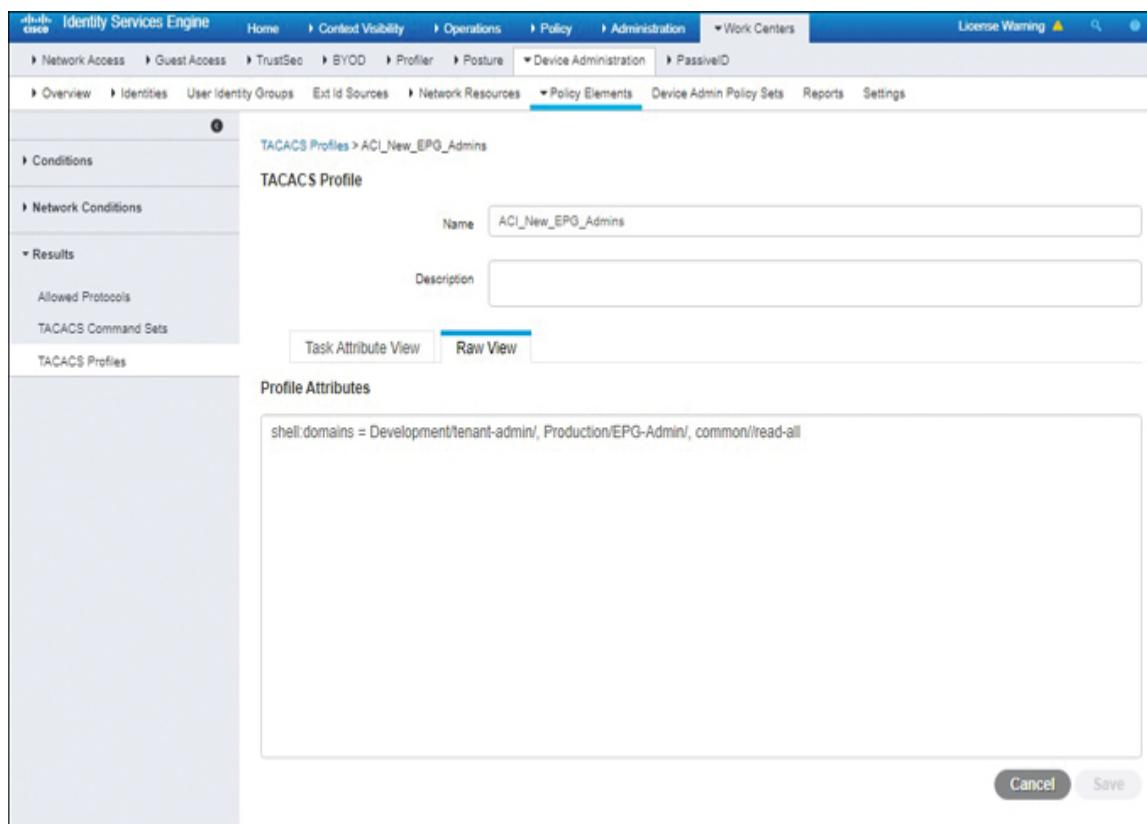
**Step 4. Configure users and associate each of them with a user identity group:** Users can be created locally on ISE servers, or ISE can integrate with external user databases. Each user needs to be assigned to a user identity group so that ISE can map the user to an authorization rule.

**Step 5. Create a TACACS profile for each user identity group:** A TACACS profile includes a custom attribute that specifies the security domains, roles, and privilege types associated with each user identity group. TACACS profiles are sometimes also referred to as shell profiles. A TACACS profile does not directly point to a user identity group. User identity group-to-TACACS profile mappings result from authorization rules.

## **Step 6.Create a device admin policy set for ACI:** An

ACI-specific device admin policy set enables ISE administrators to group authentication and authorization rules for ACI nodes. The authentication condition for the policy set may be the device type. Each authorization rule may be simply map a TACACS profile (shell profile) to a specific user identity group.

Figure 15-27 shows a TACACS profile for ACI that includes the string *shell:domains* followed by a set of three comma-separated security domain assignments. The information entered here has been specially formatted for ACI.



**Figure 15-27 Authorization Data Configuration Through a TACACS Profile**

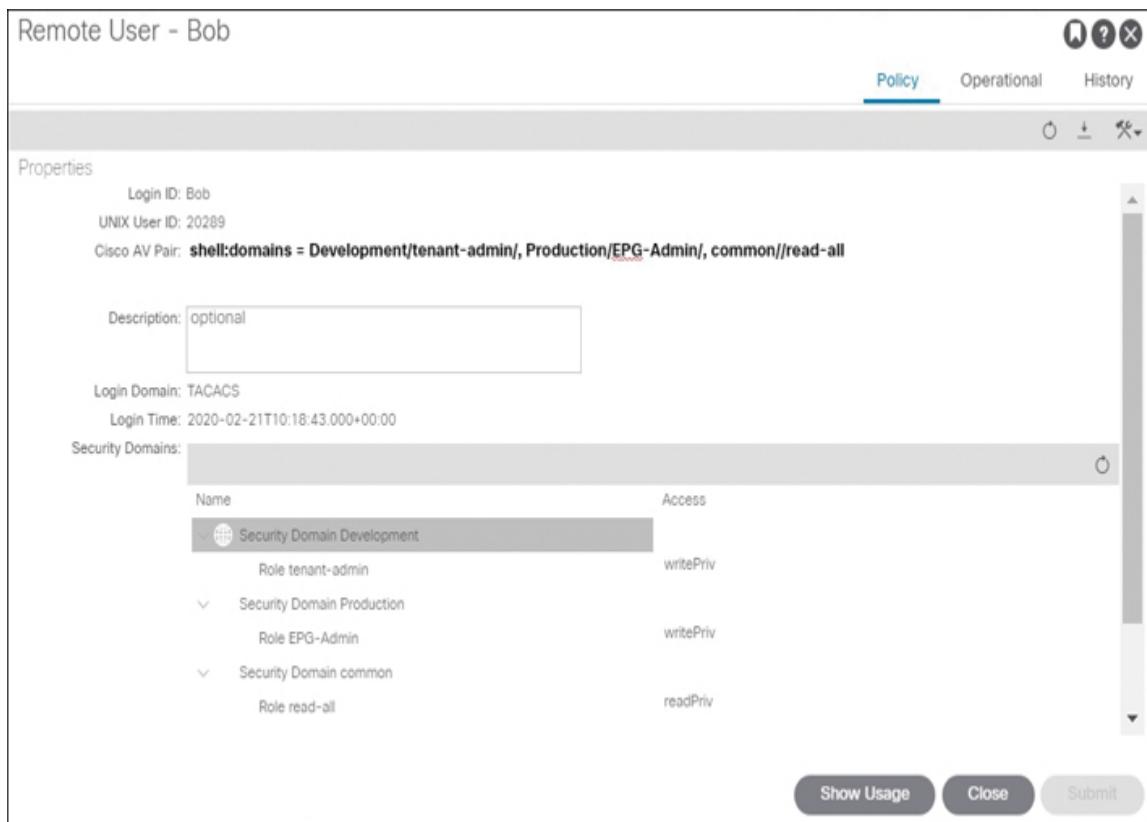
To map the TACACS profiles with users, you can use authorization rules like the ones in [Figure 15-28](#).

The screenshot shows the ISE Policy Sets interface under the 'Policy' tab. A policy set named 'ACI Policy Set' is selected. The conditions for this policy set are defined as 'DEVICE Device Type EQUALS All Device Types/ACI APICs and Switches'. The allowed protocols are set to 'ACI-TACACS'. Below this, there are four authorization policies listed:

- ACI\_New\_EPG\_Admins: Conditions: IdentityGroup Name EQUALS User Identity Groups:ACI\_New\_EPG\_Admins; Results: Command Sets: ACI\_New\_EPG\_Admins, Shell Profiles: ACI\_New\_EPG\_Admins, Hits: 14.
- ACI\_ReadOnly: Conditions: IdentityGroup Name EQUALS User Identity Groups:ACI\_read\_all; Results: Command Sets: ACI\_read\_all, Shell Profiles: ACI\_read\_all, Hits: 0.
- ACI\_Admin: Conditions: IdentityGroup Name EQUALS User Identity Groups:ACI\_admin; Results: Command Sets: ACI\_admin, Shell Profiles: ACI\_admin, Hits: 4.
- Default: Conditions: DenyAllCommands; Results: Command Sets: Deny All Shell Profile, Shell Profiles: Deny All Shell Profile, Hits: 0.

**Figure 15-28** Mapping TACACS Profiles to Users Based on Group Membership

As a result of the ACI and ISE configurations, the user Bob should be able to log in to the fabric by using the TACACS login domain and be assigned the correct Cisco attribute/value (AV) pair. [Figure 15-29](#) verifies that ISE has passed the parameters configured in the desired shell profile to ACI, providing Bob with the access he requires. You can navigate to **Admin > AAA > Users > Remote Users** and double-click the desired user to verify that the proper Cisco AV pair has been allocated to a user.



**Figure 15-29** Verifying Authorization of Externally Configured Users

## Expected Cisco AV Pair Formatting for ACI

For ACI to assign users the correct level of access, the Cisco AV pair returned by the external servers needs to be specially formatted. The following two formats are valid for ACI:

### Note

The Cisco AV pair string is case-sensitive.



Format 1:

[Click here to view code image](#)

```
shell:domains = domainA/writeRole1/writeRole2/writeRole3/  
readRole1/readRole2,
```

```
domainB/writeRole1/writeRole2/writeRole3/readRole1/readRole2
```

Format 2:

[Click here to view code image](#)

```
shell:domains = domainA/writeRole1/writeRole2/writeRole3/  
readRole1/readRole2,
```

```
domainB/writeRole1/writeRole2/writeRole3/  
readRole1/readRole2(userId)
```

Note here that the / character is a separator between write roles and read roles for each login domain and is required even if only one type of role needs to be assigned.

The only difference between the two formats is that the second one also specifies a trailing UNIX user ID. Cisco recommends assigning a unique user ID in the range 16000 to 23999. This is especially important for users who need to log in to switches or APICs via the CLI to create files or run scripts against the system. It ensures that key users have separate home directories and that any files or processes a user creates get associated with the user's unique user ID.

The Cisco AV pair formatting required by ACI can be best understood through examples, such as those presented in [Table 15-4](#).

**Table 15-4** Examples of *shell:domains* Value Assignments

---

<b>Value</b>	<b>User Access</b>
<b>Assigned to shell:domain</b>	
all/read_all	The user gains access to the all security domain via the read_all role, using the privilege type Read. Notice that the privilege type is interpreted to be Read if there are two instances of the / character between the security domain and role.
all/admin/(16005)	The user gains access to the all security domain via the admin role, using privilege type Write. Notice that the privilege type is interpreted to be Write if there is only one instance of the / character between the security domain and role. ACI assigns UNIX user ID 16005 to the user(s).
all/admin	The user gains access to the all security domain via the admin role, using the privilege type Read. Assignment of this value is beneficial when a network operator needs read-only access to the entire fabric but also needs to log in to switches. Logging in to ACI switches via the command line requires the admin role.

<b>Value</b>	<b>User Access</b>
<b>Assigned to shell:domain</b>	
all/aaa/read-all	The user gains access to the all security domain via the aaa role, using the privilege type Write. In addition, the user gains access to the all security domain via the read-all role, using the privilege type Read.
all/fabric-admin access-admin/	The user gains access to the all security domain via both the fabric-admin and access-admin roles, using the privilege type Write.
Development/tenant-admin/, Production/EPG-Admin/, common/read-all	The user gains access to the Development security domain via the tenant-admin role, using the privilege type Write. The user also gains access to the Production security domain via the custom-defined EPG-Admin role, using the privilege type Write. The user also gains access to the common security domain via the read-all role, using the privilege type Read.

The Cisco AV pair formatting expected for authorization is the same regardless of whether TACACS+, RADIUS, or LDAP

is used.

## Configuring ACI for RADIUS

ACI configuration for RADIUS is very similar to ACI configuration for TACACS+ and involves the following three basic steps:



**Step 1.** Create the desired RADIUS providers.

**Step 2.** If using ACI versions prior to Release 4 or configuring via the APIC CLI, create a RADIUS provider group.

**Step 3.** Create a RADIUS login domain.

[Table 15-5](#) describes the configuration options available in a RADIUS provider definition.



**Table 15-5** Configuration Parameters for RADIUS Providers

**C Description**  
**o**  
**n**  
**f**  
**i**  
**g**  
**u**  
**r**  
**a**  
**t**  
**i**  
**o**  
**n**  
**P**  
**a**  
**r**  
**a**  
**m**  
**e**  
**t**  
**e**  
**r**

**H** ACI allows administrators to reference a RADIUS server  
os by using either its IP address or DNS address.  
**t**  
**N**  
**a**  
**m**  
**e**  
**or**  
**IP**  
**A**  
**d**  
**dr**  
**e**  
**s**

A ut hor iz at io n Po rt	This is the service port number for the RADIUS service. The range is from 1 to 65535. The default is 1812.
A ut hor iz at io n Pr ot oc ol	RADIUS authorization protocols include PAP, MS-CHAP, and CHAP. The default RADIUS authorization protocol is PAP.
K e y	This is the shared secret key that ACI nodes and the RADIUS server use for encrypting and decrypting passwords between one another. In contrast with TACACS+, which encrypts the entire payload, RADIUS encrypts passwords only.

Ti m e o ut	This is the number of seconds the ACI node waits for a response from the RADIUS provider before timing out. The acceptable range is from 5 to 60 seconds. The default is 5 seconds.
R et ri es	This is the number of times ACI automatically retries login attempts for a single authentication submission. The acceptable range is from 1 to 5 retries. The default is 1.
M a n a g e m e nt E P G	This is the management EPG (in-band or out-of-band) from which ACI should source requests to the RADIUS server.

S This parameter can be set to either Enabled or Disabled.  
er When it is enabled for a RADIUS provider, the APICs  
v periodically attempt to execute login attempts against  
er the RADIUS provider to verify that the RADIUS service is  
M alive on the server. When enabled, ACI asks  
o administrators to additionally enter a username and  
ni password. By default, Server Monitoring is set to  
to Disabled. Server monitoring checks are exclusive to  
ri APICs. To enable monitoring checks that leafs and spines  
n are also able to perform, use of the Ping Check option  
g under the AAA Policy view is more common.

[Example 15-2](#) presents the commands needed for configuring RADIUS via the APIC CLI.

**Example 15-2 Configuring ACI for RADIUS via the APIC CLI**  
[Click here to view code image](#)

```
APIC1(config)# show run
(...output truncated for brevity...)
    radius-server host "10.233.48.67"
        exit
    aaa group server radius RADIUS
        server 10.233.48.67 priority 1
        exit
    aaa authentication login domain RADIUS
        realm radius
        group RADIUS
        exit
```

Note in this example that new ACI software also groups RADIUS servers together behind the scenes even though the

GUI abstracts this step from users. Again, key values do not appear in **show** command output.

## Configuring ACI for LDAP

ACI configuration for LDAP integration involves the following basic steps:



**Step 1.** Configure the desired LDAP providers.

**Step 2.** Configure LDAP group map rules unless the LDAP providers have been configured to return Cisco AV pairs.

**Step 3.** Configure an LDAP group map.

**Step 4.** Create an LDAP login domain.

[Table 15-6](#) describes the configuration options available in an LDAP provider definition.



**Table 15-6** Configuration Parameters for LDAP Providers

## C Description

o  
n  
fi  
g  
ur  
at  
io  
n  
P  
ar  
a  
m  
et  
er

H ACI allows administrators to reference an LDAP server  
os by using either its IP address or DNS address.

tN  
a  
m  
e  
or  
IP  
Ad  
dr  
es  
s

Port	This is the service port number for the LDAP service. The range is from 1 to 65535. The default is 389.
Bind DN	This is a string referencing an account on an LDAP server that is able to query at least a portion of the LDAP directory hierarchy. This account should ideally be a system account with a non-expiring password. In non-production environments leveraging LDAP servers with anonymous bind capabilities, this field can be left empty. LDAP server administrators need to provide ACI administrators the exact Bind DN string.
Base DN	This is a string referencing the container and subtree under which ACI is able to execute queries for matching users using the bind account. LDAP server administrators need to provide ACI administrators the exact Base DN string.
Password	This is the password of the LDAP account specified in the Bind DN field.

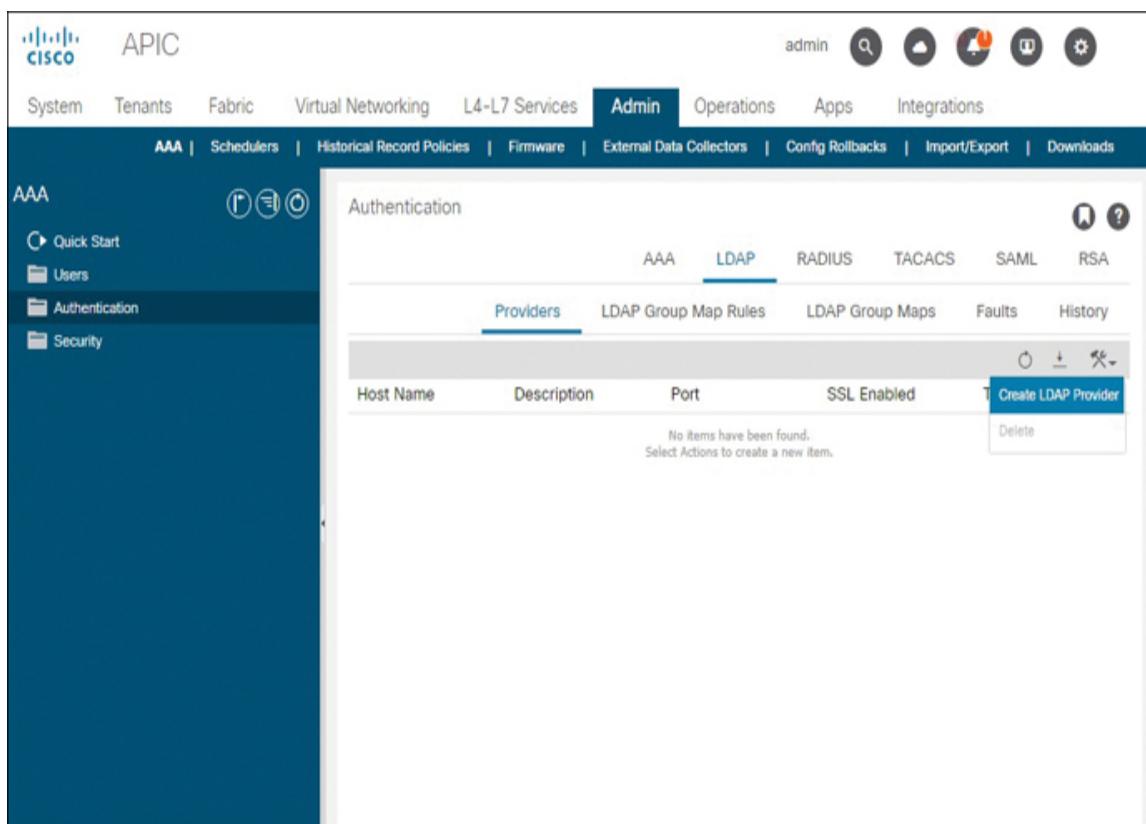
Ti m eo ut	This is the number of seconds ACI nodes wait for a response from the LDAP provider server. The acceptable range is from 5 to 60 seconds. The default is 30 seconds.
Re tri es	This is the number of times ACI automatically retries login attempts for a single authentication submission. The acceptable range is from 1 to 5 retries. The default is 1.
En le SS L	This checkbox allows administrators to enforce SSL-based connections with the LDAP provider.

SS L Ce rti fic at e Va lid ati on Le ve l	<p>Acceptable values for this option are Permissive and Strict. Permissive certificate checking relaxes requirements around certificate validation, making it an ideal option in deployments that use self-signed certificates. Strict certificate validation is ideal in production environments.</p>
At tri bu te	<p>ACI determines the level of user authorization from the Attribute field. The two most common values for this field are memberOf and CiscoAVPair. When using the memberOf option, configuration of LDAP group map rules is mandatory.</p>
Fil te r Ty pe	<p>Filters define how ACI queries an LDAP schema and interprets the existence of a user. The LDAP provider configuration page provides three options for the Filter Type parameter: Default, Microsoft AD, and Custom.</p>

Custom Filter	This is a custom filter value.
Management EPG	This is the management EPG (in-band or out-of-band) from which ACI should source requests to the LDAP server.
Server Monitoring	This parameter can be set to either Enabled or Disabled. When it is enabled for an LDAP provider, the APICs periodically attempt to execute login attempts against the LDAP provider to verify that the LDAP service is alive on the server. When enabled, ACI asks administrators to additionally enter a username and password. By default, Server Monitoring is set to Disabled. Server monitoring checks are exclusive to APICs. To enable monitoring checks that leafs and spines are also able to perform, use of the Ping Check option under the AAA Policy view is more common.

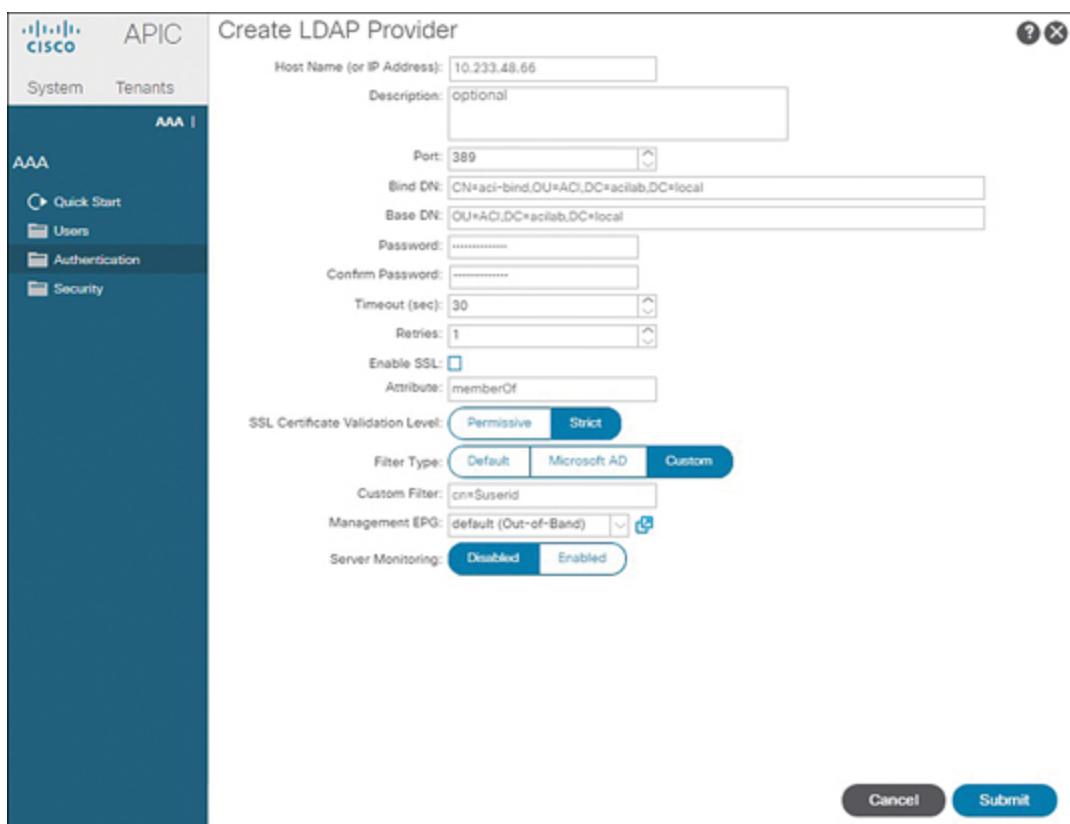
The Filter Type option requires further explanation. If you select the Default filter type in a GUI-based LDAP provider configuration, ACI does not ask users to enter a custom filter. Instead, it uses the string `cn=$userid` as the filter. This is suitable for most LDAP implementations that define users as common name (`cn`) objects. In Microsoft Active Directory implementations, selection of the Microsoft AD filter type prompts ACI to configure `sAMAccountName=$userid` as the filter. For any LDAP schema that requires use of a different filter, select the Custom filter type to expose the Custom Filter field and enter a filter that is more suitable to the LDAP implementation.

Figure 15-30 shows how to open the LDAP provider creation wizard by navigating to **Admin > AAA > Authentication > LDAP > Providers**, clicking the Tools menu, and selecting Create LDAP Provider.



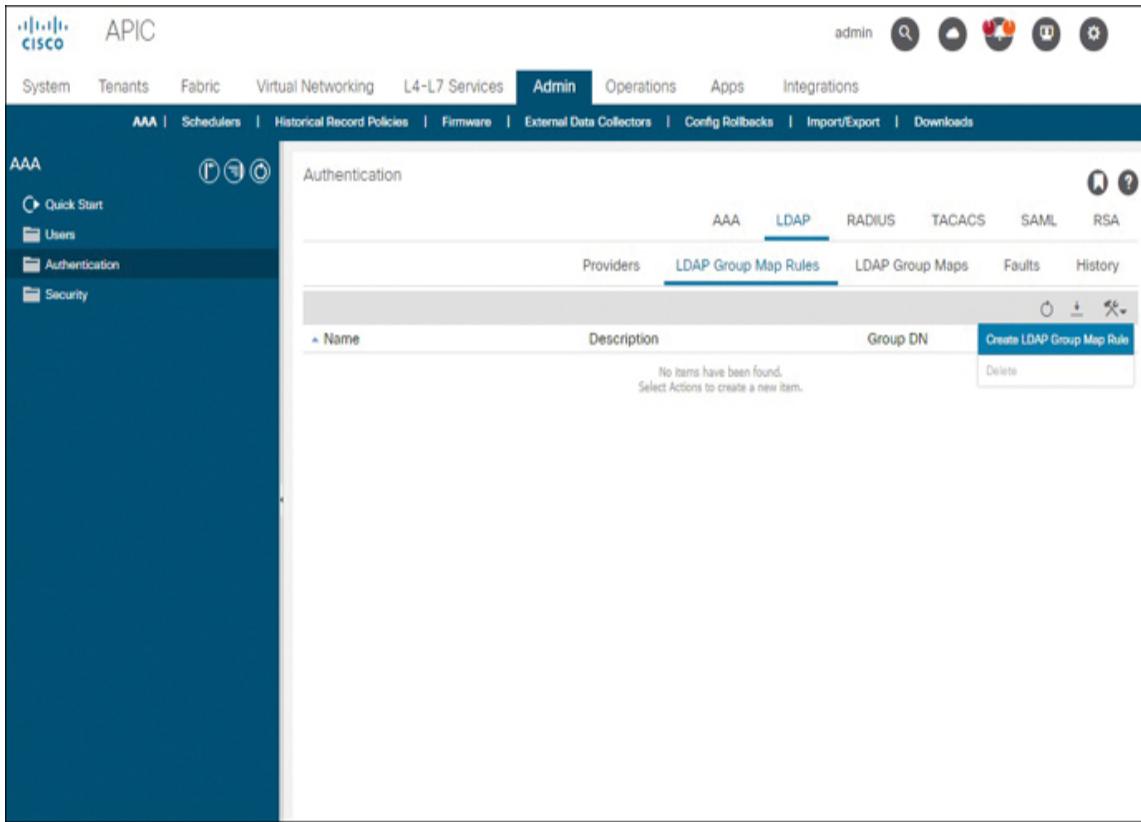
**Figure 15-30** Navigating to the Create LDAP Provider Page

Notice in [Figure 15-31](#) that the memberOf attribute is used in the LDAP configuration. This tells ACI that it should expect the LDAP server to return the LDAP-based group membership of the users instead of a Cisco AV pair. In addition, Filter Type is set to Custom, and the Custom Filter field is set to cn=\$userid in. However, you could instead set Filter Type to Default.



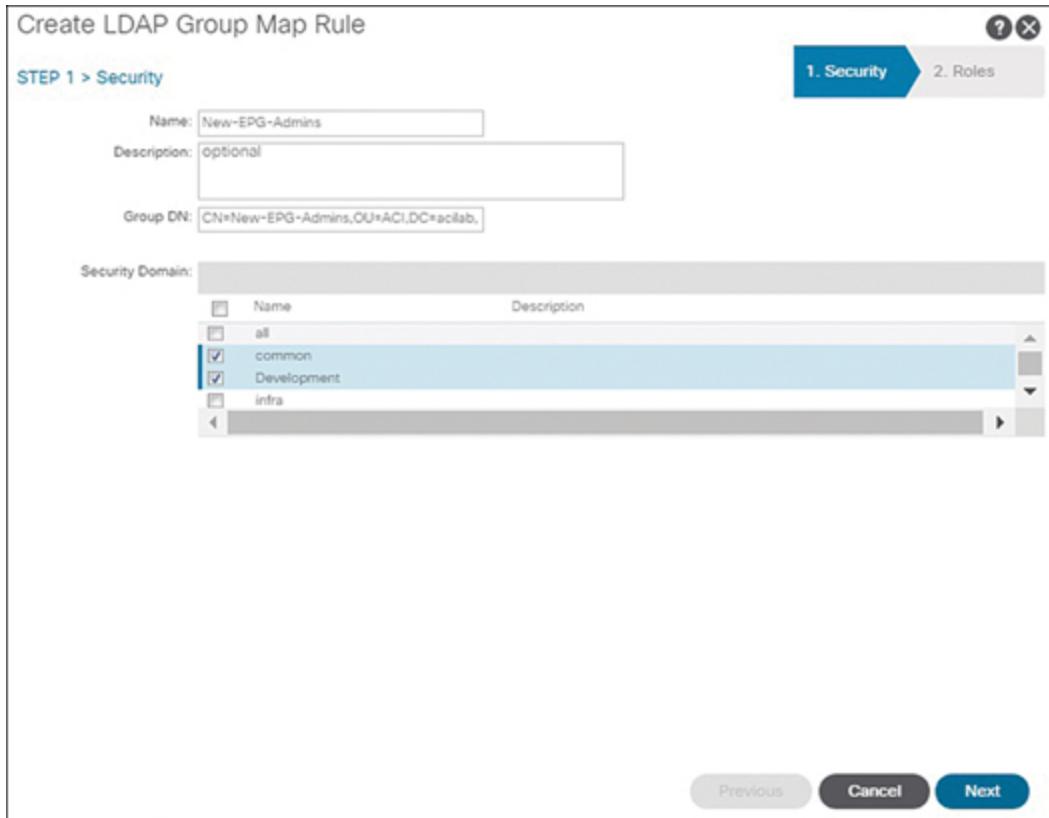
**Figure 15-31** Defining an LDAP Provider

In addition to having LDAP providers defined, ACI needs to know how to map LDAP group memberships with ACI RBAC constructs to determine the level of authorization. To define such mappings, you open the Create LDAP Group Map Rules wizard, as shown in [Figure 15-32](#).



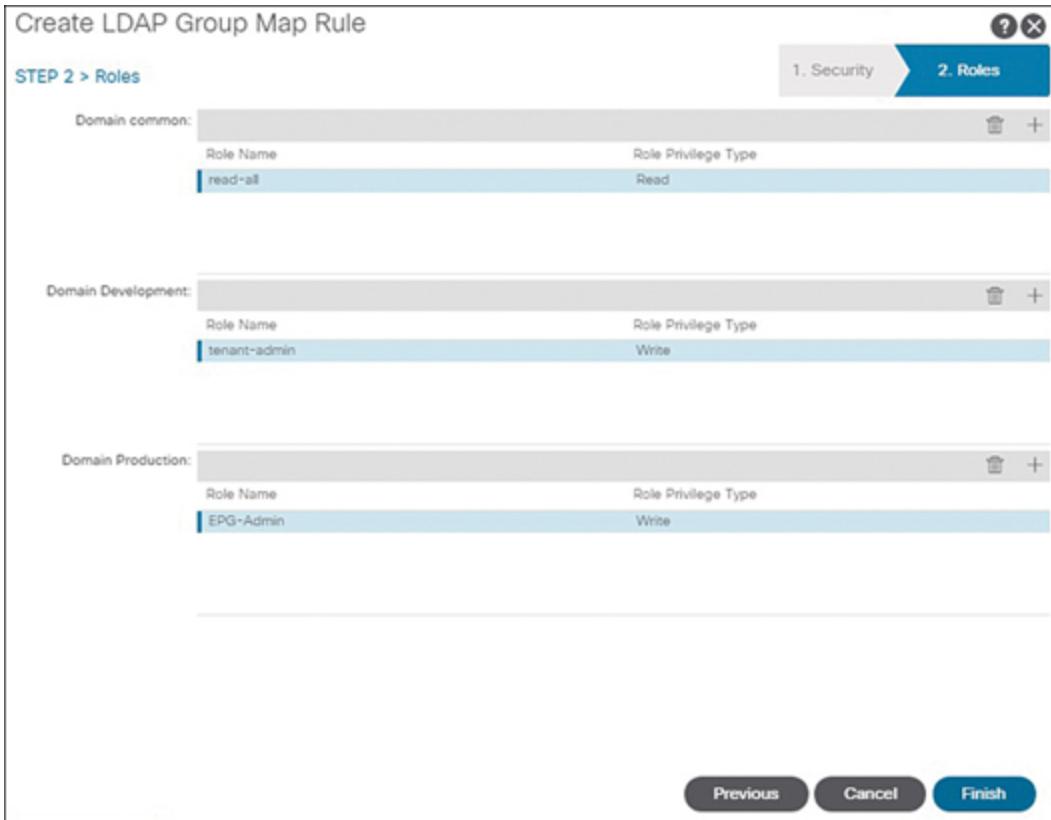
**Figure 15-32** Navigating to the *Create LDAP Group Map Rules Wizard*

The first page in the Create LDAP Group Map Rules wizard asks for the group DN the LDAP servers will return for a given group defined on the LDAP servers. ACI also needs to know what security domains users with the specified group membership should be able to access. [Figure 15-33](#) shows the LDAP Group DN setting and the security domains ACI should correlate with the specified LDAP group.



**Figure 15-33** *The Create LDAP Group Map Rules Wizard, Page 1*

[Figure 15-34](#) shows how to associate roles and privileges with each security domain specified.



**Figure 15-34** *The Create LDAP Group Map Rules Wizard, Page 2*

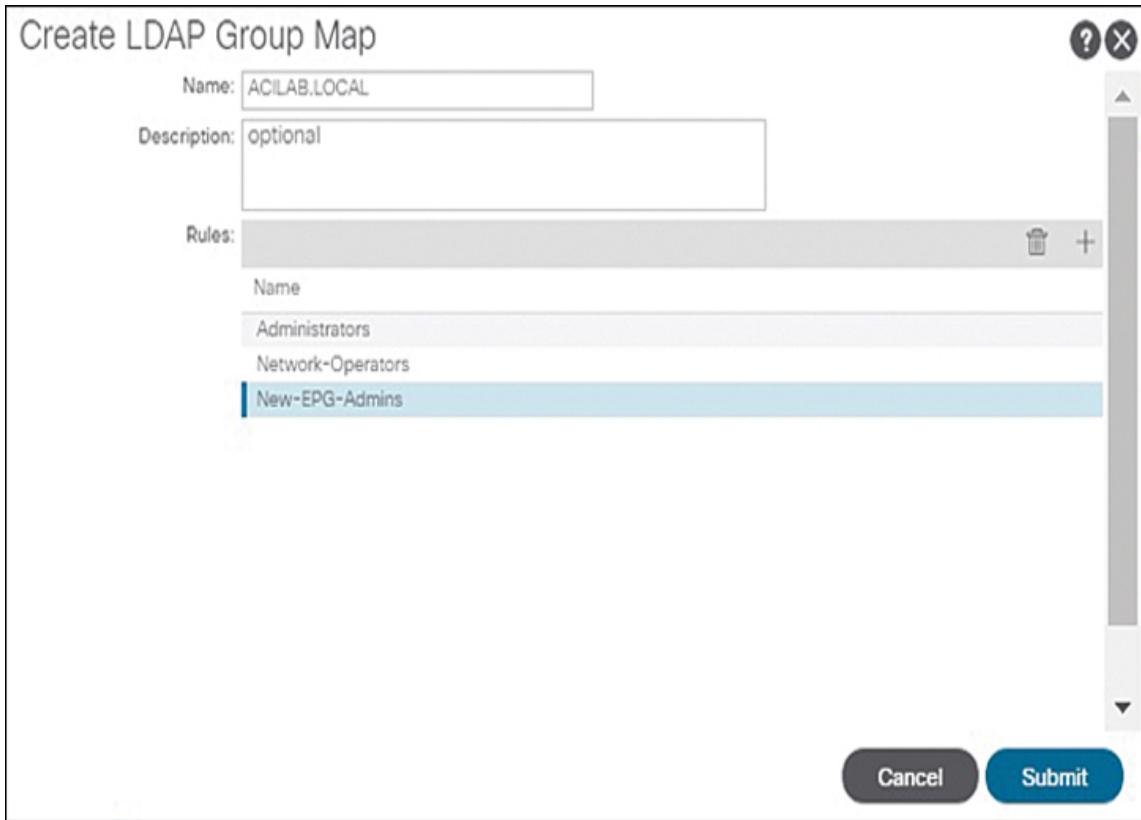
By the end of the process, LDAP group-to-ACI access level mappings should have been configured for each family of ACI users that will be authenticated and authorized via LDAP. [Figure 15-35](#) shows a deployment with three sets of mapping rules.

The screenshot shows the Cisco Application Policy Infrastructure Controller (APIC) web interface. The top navigation bar includes tabs for System, Tenants, Fabric, Virtual Networking, L4-L7 Services, Admin (which is selected), Operations, Apps, and Integrations. The Admin tab has sub-tabs for AAA, Schedulers, Historical Record Policies, Firmware, External Data Collectors, Config Rollbacks, Import/Export, and Downloads. On the left, a sidebar under the AAA section shows Quick Start, Users, Authentication (which is selected), and Security. The main content area is titled 'Authentication' and has tabs for AAA, LDAP (which is selected), RADIUS, TACACS, SAML, and RSA. Under the LDAP tab, there are sub-tabs for Providers, LDAP Group Map Rules (which is selected), LDAP Group Maps, Faults, and History. The 'LDAP Group Map Rules' table lists three entries:

Name	Description	Group DN
Administrators		CN=admin,OU=ACI,DC=acilab,DC=local
Network-Operators		CN=read-all,OU=ACI,DC=acilab,DC=local
New-EPG-Admins		CN=New-EPG-Admins,OU=ACI,DC=acilab,DC=lo...

**Figure 15-35** *LDAP Group Map Rules*

The LDAP mapping rules can then be grouped together into an LDAP group map. When a deployment requires multiple LDAP login domains, administrators can associate a different LDAP group map to each LDAP login domain. [Figure 15-36](#) shows the three LDAP group map rule entries being added to an LDAP group map.



**Figure 15-36** Configuring an LDAP Group Map

Finally, you need to configure an LDAP login domain by selecting the authorization method, the desired LDAP providers, and the associated LDAP group map for user login. [Figure 15-37](#) shows an example of the LDAP login domain configuration. This example uses LDAP group mappings for user authorization. In cases where LDAP can associate Cisco AV pairs with users, the LDAP login domain and provider can leverage CiscoAVPair for authorization, which does not require the configuration of LDAP group map rules.

Create Login Domain

Name:

Realm:

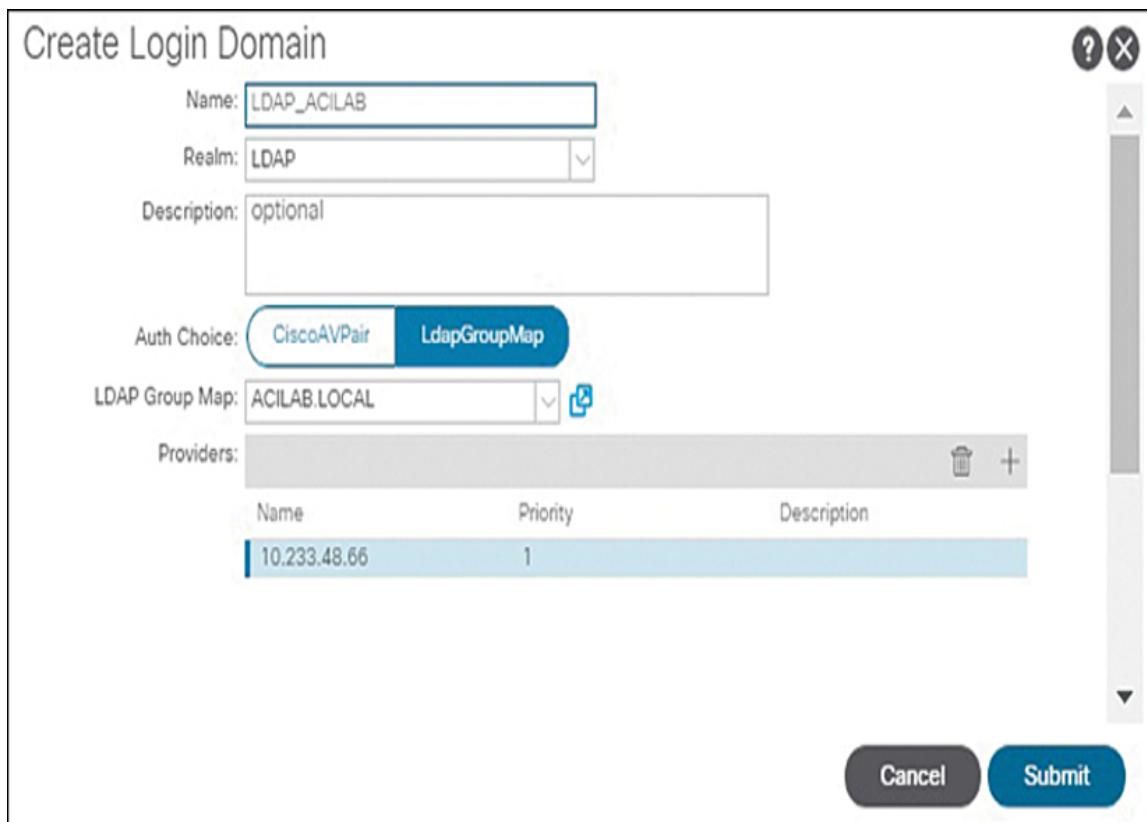
Description: optional

Auth Choice:  CiscoAVPair  LdapGroupMap

LDAP Group Map:  

Providers:

Name	Priority	Description
10.233.48.66	1	



**Figure 15-37** Configuring an LDAP Login Domain

When a user logs in to the fabric using an LDAP login domain, the level of access granted to the user can be validated in the Remote User screen. Regardless of the configured authorization method, the Remote User screen shows the associated Cisco AV pair, as shown in [Figure 15-38](#).

Properties

Login ID: Bob  
UNIX User ID: 20289  
Cisco AV Pair: shell:domains = common//read-all,Production/EPG-Admin/,Development/tenant-admin/

Description: optional

Login Domain: LDAP ACI LAB  
Login Time: 2020-02-22T00:16:11.000+00:00

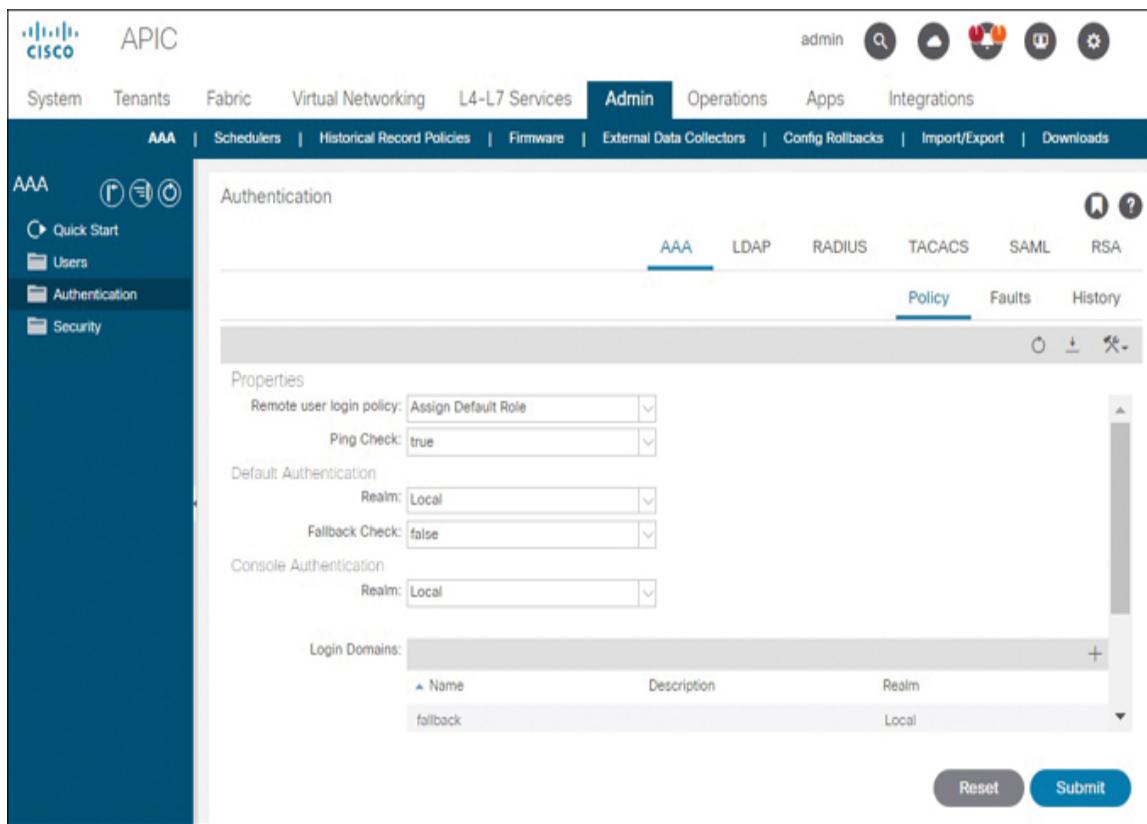
Name	Access
Security Domain Development	writePriv
Role tenant-admin	
Security Domain Production	writePriv
Role EPG-Admin	
Security Domain common	readPriv
Role read-all	

Show Usage   Close   Submit

**Figure 15-38** Verifying the Level of Access for LDAP-Authenticated Users

## AAA Authentication Policy Settings

There are several global AAA configuration settings that users can tweak at **Admin > AAA > Authentication > AAA > Policy**. [Figure 15-39](#) shows these settings.



**Figure 15-39** Configuration Options on the AAA Policy Page

Table 15-7 describes these settings and the drop-down list values users can choose for each setting.



**Table 15-7** Settings on the AAA Policy Page

## S Description et ti n g

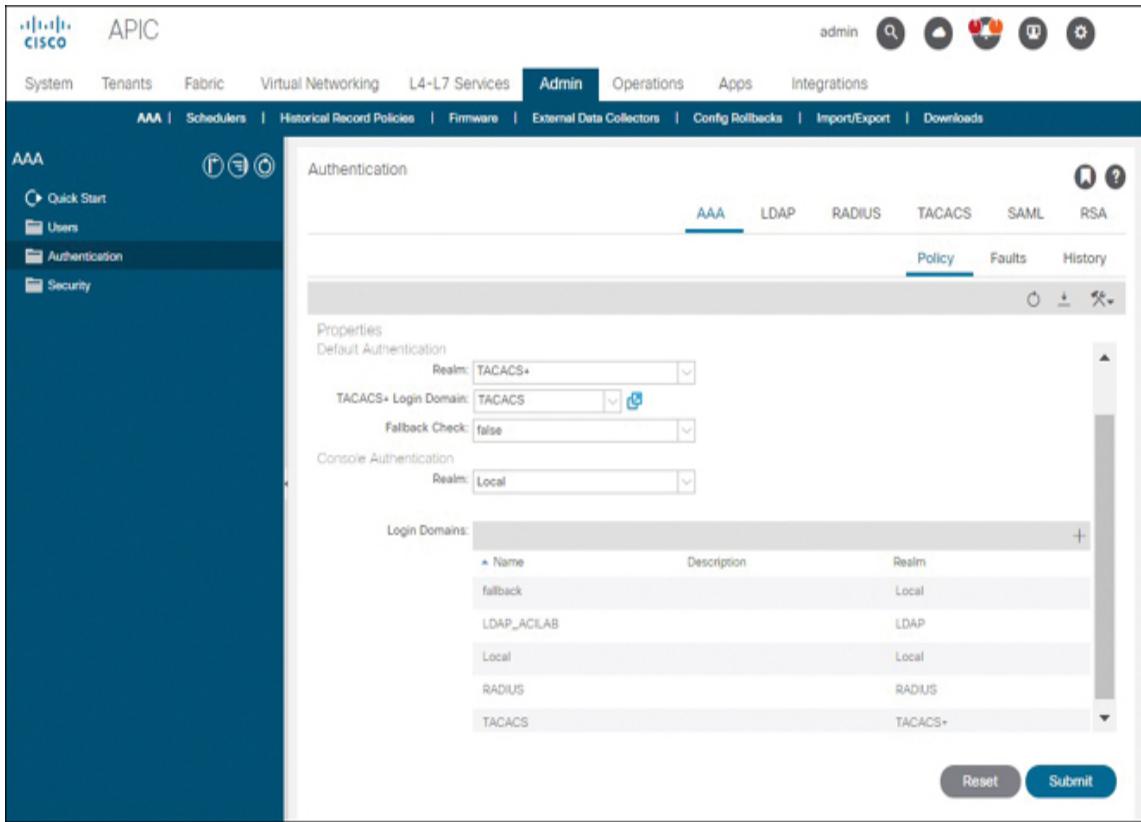
Re	There are two acceptable values for this setting. Assignm Default Role ensures that any remotely authenticated ot user with a bad or missing Cisco AV pair gets assigned e to the common security domain using the read-all role Us and the privilege type Read. No Login, on the other er hand, ensures that such users cannot log in to the Lo fabric. The default value for this setting is No Login.
Pi	ACI performs ICMP health checks against AAA providers. ng The Ping Check setting can be set to either True or False. C ACI continues to run ICMP health checks against all AAA he providers, regardless of the value chosen. If Ping Check ck is set to True, ACI removes inaccessible providers from the authentication process and authenticates against operational providers instead. If ICMP traffic cannot reach AAA servers due to firewall rules, the Ping Check setting should be set to False to ensure that ACI continues to authenticate against all servers, regardless of the result of ICMP health checks.

<p>D ef au It A ut he nt ic ati on Re al m</p>	<p>This setting governs the login domain ACI uses when a user does not select a domain when logging in to the fabric. It also determines the login domain used when a user attempts to log in to the fabric by using the DefaultAuth domain. You can select Local, LDAP, RADIUS, TACACS+, RSA, or SAML for Default Authentication Realm. Local is selected by default. When you select a setting other than Local, an additional Login Domain field appears, allowing selection of the intended domain.</p>
	<p>Fa llb ac k C he ck</p> <p>ACI comes with a preconfigured login domain called fallback that is set to local authentication by default. The Fallback Check setting, which can be set to True or False, enables or disables reliance on the AAA provider ICMP health check for activation of the fallback domain. If fallback check is set to True and the configured AAA providers respond to ICMP traffic but are unable to authenticate users, the fallback login domain will be unavailable, and users may remain locked out of the fabric. For this reason, Fallback Check is often kept at its default value of False.</p>

C This parameter allows users to specify the authentication method for console logins to ACI nodes.  
on so By default, it is set to Local. Other valid options are  
le A setting besides Local, ACI exposes an additional field for  
ut you to specify the login domain.

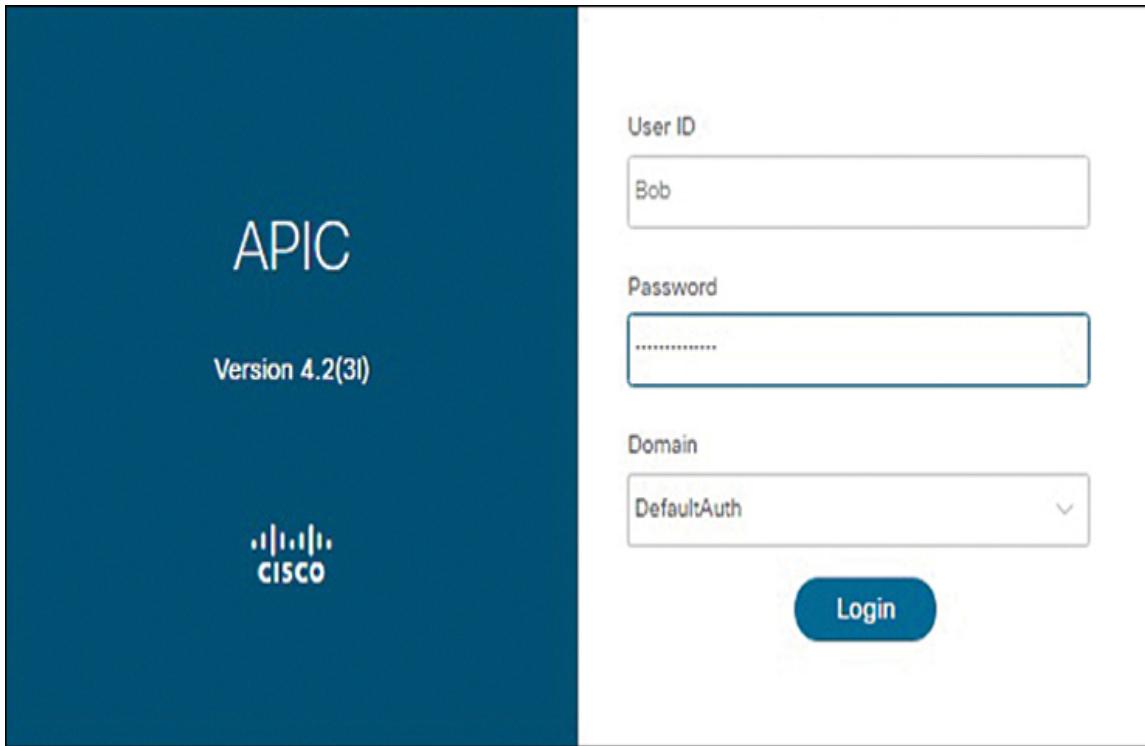
he  
nt  
ic  
ati  
on  
Re  
al  
m

**Figure 15-40** shows Default Authentication Realm set to TACACS+ and a login domain called TACACS. Fallback Check has been kept at its default value of False. The figure also shows all the available login domains in a particular fabric.



**Figure 15-40** *Login Domains and Global AAA Settings Configured in a Fabric*

As a result of these settings, the login domain DefaultAuth, as shown in [Figure 15-41](#), references the login domain called TACACS. Remember from [Figure 15-40](#) that there is no explicit domain called DefaultAuth, as this refers to any domain set for Default Authentication Realm.



**Figure 15-41** DefaultAuth Domain Shown in ACI Login Page

Before changing the value of the Default Authentication Realm parameter, ensure that the AAA providers associated with the preferred login domain are available and that those providers authorize users as expected.

## Regaining Access to the Fabric via Fallback Domain

To use the fallback domain to regain access to a fabric due to AAA connectivity or configuration issues, use the syntax **apic:fallback||username** in the GUI User ID field. Use the syntax **apic#fallback||username** when logging in via the CLI.

## Exam Preparation Tasks

As mentioned in the section “How to Use This Book” in the Introduction, you have a couple of choices for exam preparation: [Chapter 17](#), “Final Preparation,” and the exam simulation questions on the companion website.

## Review All Key Topics

Review the most important topics in this chapter, noted with the Key Topic icon in the outer margin of the page. [Table 15-8](#) lists these key topics and the page number on which each is found.



**Table 15-8** Key Topics for [Chapter 15](#)

Key Topic Element	Description	Page Number
List	Lists the elements associated with users that collectively determine the level of management access a user has to an ACI fabric	<a href="#">516</a>
Paragraph	Defines security domains	<a href="#">517</a>

Key Topic Element	Description	Page Number
List	Lists the predefined out-of-the-box security domains	517
Paragraph	Defines privileges in ACI	519
Paragraph	Defines roles in ACI	519
Table 15-2	Lists and describes predefined roles in ACI	520
Figure 15-8	Demonstrates how to assign a local user to security domains	523
Figure 15-9	Demonstrates how to map roles and role privilege types to a local user	524

Key Topic Element	Description	Page Number
Paragraph	Explains that all users created in ACI gain read-only access to the common tenant	525
Paragraph/List Item	Defines RBAC rules and describes their characteristics	528
Figure 15-18	Shows how an RBAC rule can be created	529
List	Explains the high-level steps involved in configuring ACI for integration with TACACS+	532
Table 15-3	Describes the settings available for TACACS+ provider configurations	532
Paragraph	Documents the two Cisco AV pair formats ACI accepts from external AAA servers	539

Key Topic Element	Description	Page Number
List	Explains the high-level steps involved in configuring ACI for integration with RADIUS	540
Table 15-5	Describes the settings available for RADIUS provider configurations	540
List	Explains the high-level steps involved in configuring ACI for integration with LDAP	541
Table 15-6	Describes the settings available for LDAP provider configurations	542
Table 15-7	Describes the global AAA settings that administrators can tweak	548

## Complete Tables and Lists from Memory

Print a copy of [Appendix C, “Memory Tables”](#) (found on the companion website), or at least the section for this chapter, and complete the tables and lists from memory. [Appendix D](#),

[“Memory Tables Answer Key”](#) (also on the companion website), includes completed tables and lists you can use to check your work.

## Define Key Terms

Define the following key terms from this chapter and check your answers in the glossary:

[security domain](#)  
[privilege](#)  
[role](#)  
[RBAC rule](#)

# **Part VI: Operations**

# Chapter 16

## ACI Anywhere

This chapter covers the following topics:

**ACI Multi-Site Fundamentals:** This section reviews ACI Multi-Site terminology and describes, at a high level, how ACI Multi-Site works.

**Building Primary and Disaster Recovery Data Centers with ACI:** This section explains why ACI Multi-Site shines in primary and DR data center designs.

**Building Active/Active Data Centers with ACI:** This section contrasts the use of ACI Multi-Pod and the use of ACI Multi-Site when designing active/active data centers.

**Extending ACI to Remote Locations and Public Clouds:** This section describes some additional use cases for ACI Multi-Pod and Multi-Site.

This chapter covers the following exam topics:

- 6.1 Describe multipod
- 6.2 Describe multisite

This book has looked closely at many of the details of ACI constructs and single-pod fabric deployments. Modern data centers, however, often need to extend beyond the confines

of single-pod fabrics and into other environments. The network and whitelisting policies that have been defined for applications also need to extend into other environments. Enter ACI Anywhere.

ACI Anywhere is a marketing term that highlights not just the evolution of ACI but also all the places Cisco has been taking ACI. This term encompasses a host of solutions, including ACI Multi-Pod, ACI Multi-Site, ACI Multicloud, Remote Leaf, vPod, and other solutions yet to come. The common theme that underpins all these solutions is that together they extend the ACI operational framework across data centers, across remote locations, across bare-metal clouds, and even into public cloud environments. Together, ACI Anywhere solutions transform ACI into a true hybrid cloud solution.

Although those interested in implementation guidance need to search outside this book, this chapter shines light on some of the key use cases for ACI Multi-Pod and ACI Multi-Site. Which one best fits into your data center strategy? As you will see in this chapter, sometimes the answer can be both.

## **“Do I Know This Already?” Quiz**

The “Do I Know This Already?” quiz allows you to assess whether you should read this entire chapter thoroughly or jump to the “Exam Preparation Tasks” section. If you are in doubt about your answers to these questions or your own assessment of your knowledge of the topics, read the entire chapter. [Table 16-1](#) lists the major headings in this chapter and their corresponding “Do I Know This Already?” quiz questions. You can find the answers in [Appendix A, “Answers to the ‘Do I Know This Already?’ Questions.”](#)

**Table 16-1** “Do I Know This Already?” Section-to-Question Mapping

Foundation Topics Section	Questions
ACI Multi-Site Fundamentals	1-5
Building Primary and Disaster Recovery Data Centers with ACI	6, 7
Building Active/Active Data Centers with ACI	8, 9
Extending ACI to Remote Locations and Public Clouds	10

### Caution

The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark that question as wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

- 1.** Which protocol does ACI Multi-Site use to advertise information about an endpoint in a stretched subnet to an adjacent fabric?

  - a.** COOP
  - b.** OSPF
  - c.** MP-BGP EVPN
  - d.** EIGRP
- 2.** Which protocol establishes the underlay for ISN connectivity in an ACI Multi-Site deployment?

  - a.** COOP
  - b.** OSPF
  - c.** MP-BGP EVPN
  - d.** EIGRP
- 3.** True or false: Administrators should refrain from configuring access policies locally within a fabric after integrating the fabric into ACI Multi-Site.

  - a.** True
  - b.** False
- 4.** True or false: A schema defines a unit of change within ACI Multi-Site.

  - a.** True
  - b.** False
- 5.** How do you make application traffic flow over the ISN between two ACI fabrics that have been integrated into ACI Multi-Site?

  - a.** Put in place contracts between the desired EPGs and ensure that the route is known in both the source and destination VRF instances.

- b.** Navigate to Configure Infra in ACI Multi-Site and ensure that all configuration parameters are correct.
  - c.** ACI Multi-Pod needs to also be deployed for application communication to function.
  - d.** ACI Multi-Site is only used for stretching EPGs, not for allowing communication.
- 6.** What feature can ACI Multi-Pod and ACI Multi-Site use to optimize ingress routing for stretched subnets into data centers over the WAN?
  - a.** Contracts
  - b.** Access policies
  - c.** COOP
  - d.** Host-based routing
- 7.** True or false: ACI Multi-Site can enable IP mobility across data centers without the need to flood broadcast, unknown unicast, and multicast traffic.
  - a.** True
  - b.** False
- 8.** A company needs to perform VMM integration, conduct cross-data center vMotion, and leverage vSphere DRS in a multi-data center design. Which solution best fits these requirements?
  - a.** ACI Multi-Pod
  - b.** ACI Multi-Site
  - c.** Remote leaf
  - d.** ACI Multi-Tier
- 9.** A company wants to deploy active/active firewall clustering across two data centers. Which solution supports this requirement?

- a.** ACI Multi-Pod
- b.** ACI Multi-Site
- c.** Remote leaf
- d.** ACI Multi-Tier

**10.** A company wants to integrate its on-premises ACI fabrics with public cloud environments. Which of the following ACI solutions can support such an integration? What solution, if any, needs to be deployed in the cloud to make this integration work?

- a.** ACI Multi-Pod and Cisco ASR 1000
- b.** ACI Multi-Site and Cisco Cloud APIC
- c.** ACI Multi-Pod and Cisco Cloud APIC
- d.** ACI Multi-Site and Cisco ASR 1000

## Foundation Topics

### ACI Multi-Site Fundamentals

ACI Anywhere extends the confines of the data center network to anywhere an organization owns data. ACI Multi-Site is often the glue that enables such extension. Therefore, we address ACI Multi-Site first.

[Chapter 2, “Understanding ACI Hardware and Topologies,”](#) covers what a valid ACI Multi-Site topology looks like as well as the reasons second-generation spines are required for deploying ACI Multi-Site. But beyond hardware and topologies, there are other questions about these two solutions that beg answers. For example, what protocols do ACI Multi-Site fabrics use to interconnect? What new terminology does ACI Multi-Site introduce? What new

constructs are required to integrate multiple ACI fabrics? Finally, how do you choose between ACI Multi-Pod and ACI Multi-Site for a given use case? To develop the context needed to answer these questions, a basic understanding of ACI Multi-Site fundamentals is required.

## Interconnecting ACI Fabrics with ACI Multi-Site

As mentioned earlier in this book, ACI Multi-Site requires the deployment of an intersite network (ISN) between sites that supports OSPF on the last-hop routers facing the spines in each site. OSPF establishes the underlay for cross-site communication.

To configure ACI to form OSPF adjacencies with the ISN, you need to configure access policies on spines locally within each fabric for ports that connect to the ISN. In addition, you need to define APIC addresses for each fabric as sites within the Multi-Site Orchestrator (MSO) cluster. Then you can use the MSO cluster to deploy a special L3Out in the infra tenant. This L3Out is very similar to L3Outs that enable IPN functionality in ACI Multi-Pod deployments, but you should refrain from manually modifying ACI Multi-Site L3Outs.

[Table 16-2](#) describes some of the cross-site connectivity concepts related to ACI Multi-Site.

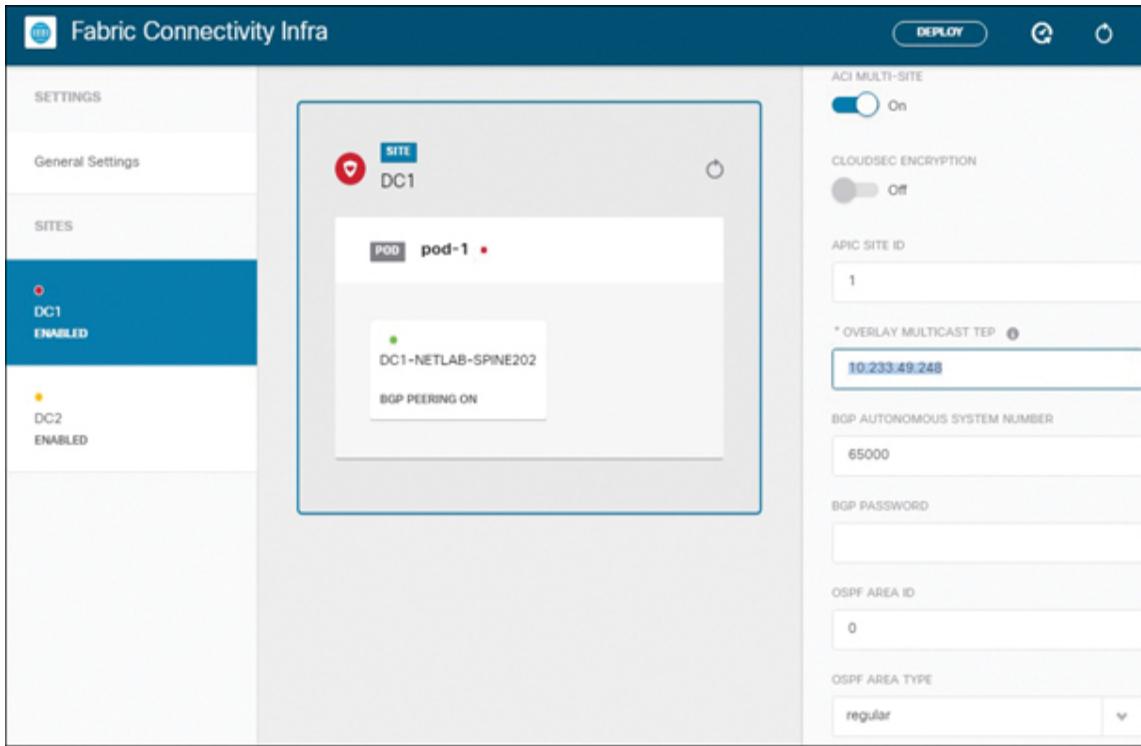


**Table 16-2** ACI Multi-Site Cross-Site Connectivity Configuration Concepts

Term	Definition
<b>Site</b>	Each independent ACI fabric. To be able to add a site and deploy policy to it, MSO needs to know the name to allocate to the specified fabric, the URLs of APICs within the fabric, and the site ID.
<b>Site ID</b>	A unique numeric identifier for each fabric. Once selected, the site ID cannot be changed. Note that this setting has no relationship to the fabric ID value that administrators need to enter during fabric initialization.
<b>Overlay multicast TEP</b>	A single anycast tunnel endpoint configured within each fabric (site) for ingress replication of cross-site data plane BUM traffic.
BGP autonomous system number (ASN)	The BGP ASN configured within the fabric or site for route reflection. MSO pulls the BGP ASN configured in each site when users define sites in MSO.

Term	Definition
<b>Overlay unicast TEP</b>	A single anycast TEP address assigned to each pod within each fabric for forwarding of cross-site unicast traffic.
<b>BGP EVPN router ID</b>	An ID for each spine that has been enabled for ACI Multi-Site forwarding that enables route peering across sites.
<b>Cloud Sec encryption</b>	Encryption used when traffic is egressing a fabric through Multi-Site spines and is destined for spines in another fabric and needs to be encrypted. Note that not all spines and not all spine ports support CloudSec. This is a VTEP-to-VTEP encryption feature.

Figure 16-1 shows the Fabric Connectivity Infra page, which can be accessed in current MSO versions at Sites > Configure Infra. This screenshot depicts an ACI Multi-Site deployment with two sites defined: DC1 and DC2. In this figure, you can see the connectivity details at the site (fabric) level for the DC1 fabric. Recall from Table 16-2 that some configuration parameters, such as overlay multicast TEP, are configured at the fabric level. Users can further drill down to pod-specific configurations within each site or to spine-specific configurations.



**Figure 16-1** Configuring ACI Multi-Site for Connectivity to the ISN

With the proper ACI Multi-Site L3Outs in place, spines within each fabric that have been enabled for Multi-Site functionality can establish MP-BGP EVPN adjacencies with one another. ACI fabrics can then advertise endpoint and other routing information to each other. ACI uses Type 2 EVPN updates to communicate endpoint information to other fabrics. The COOP database on spines in each site can easily differentiate between remote fabric endpoints and endpoints within the local fabric by referencing the overlay unicast TEP of the target site for all remote fabric endpoint entries. A fabric never needs to update remote fabrics of endpoint moves that stay local to the fabric. If an endpoint moves between fabrics, EVPN Type 2 update messages are used to synchronize endpoint data within all relevant fabrics.

The establishment of Layer 3 connectivity between fabrics does not by itself mean that user traffic will start to flow over the ISN. ACI Multi-Site constructs need to be used to explicitly enable cross-site communication through contracts or preferred group settings. For this to happen, administrators configure intersite policies using schemas, templates, and other ACI Multi-Site constructs.

## New ACI Multi-Site Constructs and Configuration Concepts

Table 16-3 lists a number of new terms that need to be understood to effectively make use of ACI Multi-Site.



**Table 16-3** New ACI Multi-Site Constructs and Concepts

Ter Description	m
<b>Sc he ma</b>	A collection of configuration templates and the assignment of each template to sites that have been defined in the Multi-Site deployment. A schema can cover policies defined for a single tenant or policies for multiple tenants.
<b>Te mp</b>	A child of a schema that contains configuration objects that are either shared between sites or site specific. Each template gets associated with a single tenant.

<b>late</b>	
<b>Str etc he d</b>	Objects, such as tenants, VRF instances, EPGs, bridge domains, subnets, or contracts, that are deployed to multiple sites.
<b>Te mp lat e co nfo rmi ty</b>	A feature of ACI Multi-Site that runs checks to validate that configurations under a template pushed to multiple sites by the MSO have not been altered within a given fabric by administrators. When templates are stretched across sites, their configuration details are shared and standardized across sites.
<b>Int ers ite L3 Ou t</b>	In current MSO releases, a feature that enables endpoints located in one site to use a remote L3Out to connect to entities in an external network. Early versions of ACI Multi-Site required locally configured L3Outs in each ACI fabric to route traffic out of the fabric.
<b>Imp ort</b>	A process through which the majority of tenant objects within a production ACI fabric are brought into ACI Multi-Site.

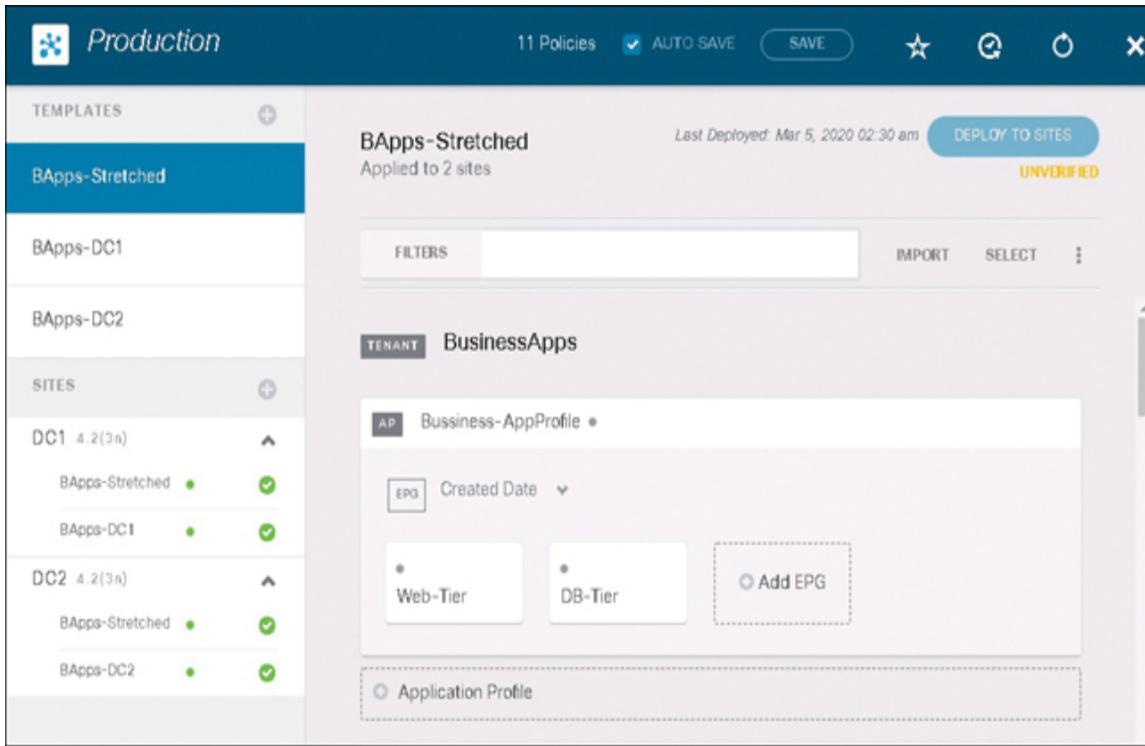
# **Locally Governed Versus MSO-Governed Configurations**

Some configurations are governed by MSO, and others continue to be governed by local fabrics.

ACI Multi-Site does *not* make any changes to access policies or fabric policies. ACI Multi-Site is strictly concerned with objects under the Tenants menu. Furthermore, even under this menu, ACI Multi-Site does not currently deal with the creation of certain objects, such as local L3Outs and L4-L7 device configurations. In addition, there is no requirement to import all tenants into ACI Multi-Site. It is okay to have a mixed deployment in which some tenants remain site local and others are imported into ACI Multi-Site.

## **Schemas and Templates in Practice**

It may not be vital for DCACI candidates to fully understand all the concepts in this chapter, but understanding schemas and templates can definitely help Implementing Cisco Application Centric Infrastructure DCACI 300-620 exam candidates get an idea of the power and flexibility of ACI Multi-Site. [Figure 16-2](#) shows a schema called Production, under which three templates have been created. In this case, all three of these templates have been associated with a stretched tenant called BusinessApps. The template names are locally significant to the MSO. One of the templates, BAApps-Stretched, houses all stretched objects, including a VRF, and can be seen to have been associated with both DC1 and DC2 fabrics. The template named BAApps-DC1 contains objects for the tenant that are local to DC1, and BAApps-DC2 manages objects specific to DC2.



**Figure 16-2** A View of a Schema, Templates, and Sites in ACI Multi-Site

The schema could combine templates from multiple tenants. A schema is not much more than a container for grouping templates and related policies together. Templates, on the other hand, are very significant. A template defines the unit or scope of a configuration change. When you navigate to the BApps-Stretched template view and select Deploy to Sites, any changes in the template from the last configuration push are simultaneously submitted to all sites associated with the template. If you make changes locally within the sites to any objects pertinent to the template, those configuration changes are modified to reflect the current configuration state of the MSO template.

## Building Primary and Disaster Recovery Data Centers with ACI

Business requirements such as business continuity and disaster recovery sometimes necessitate the deployment of separate data center fabrics. These fabrics still need to be interconnected. But which ACI solution is ideal for disaster recovery data centers?

While ACI Multi-Pod can be used for primary and disaster recovery data center deployments, the assumption in disaster recovery sometimes is that infrastructure within each data center needs to be completely separate, and the data centers need to have zero dependencies on one another. Because each ACI Multi-Pod deployment uses a single APIC cluster and some configuration changes can potentially impact all pods within a deployment, ACI Multi-Pod can sometimes be ruled out as a potential option.

In general, ACI Multi-Site has some strengths in this area that make it a more valid solution for enabling disaster recovery.

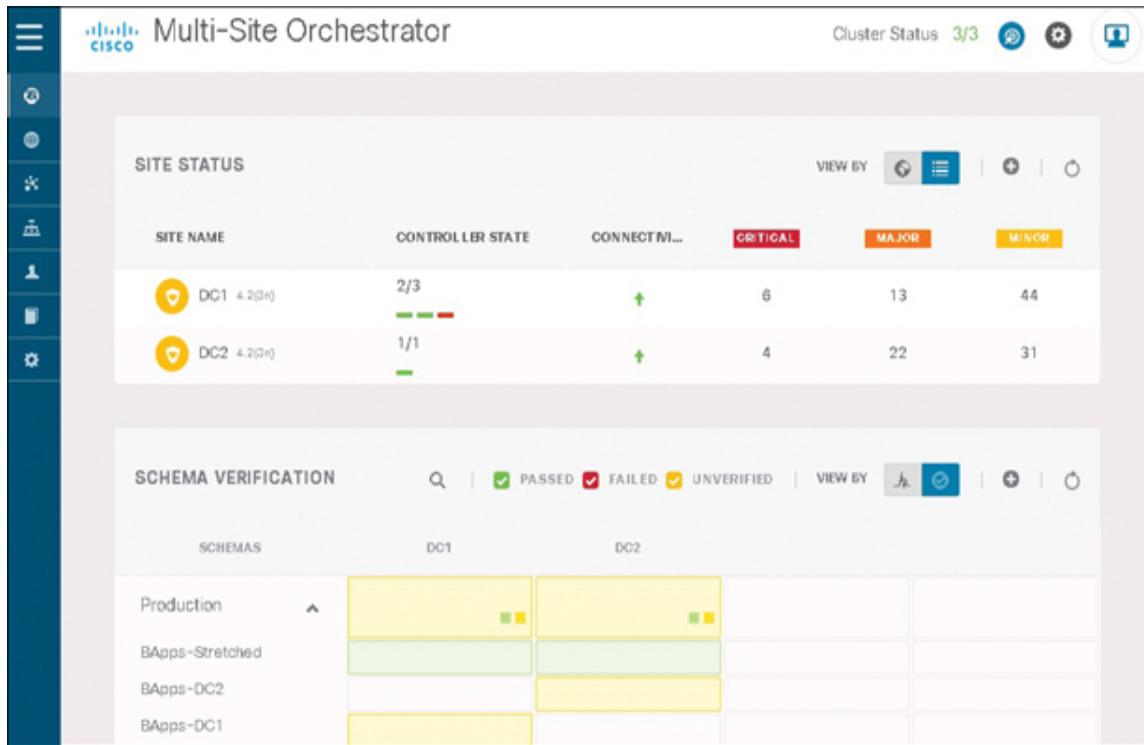
## **Centralized Orchestration and Management of Multiple Fabrics**

One important benefit of ACI Multi-Site is that MSO is yet another management tool in a network engineer's toolbelt. MSO can help validate the operational state of multiple fabrics. It allows you to audit whether there is significant configuration drift between sites and whether changes have been made that may have broken configuration conformity.

[Figure 16-3](#), for example, shows the status of sites integrated into ACI Multi-Site, faults that administrators should evaluate, and the results of schema and template verifications performed by MSO.

### **Note**

Although MSO is a management tool, it is only used in ACI Multi-Site deployments. It is not meant to be used for general management of single-fabric or single-pod ACI deployments.

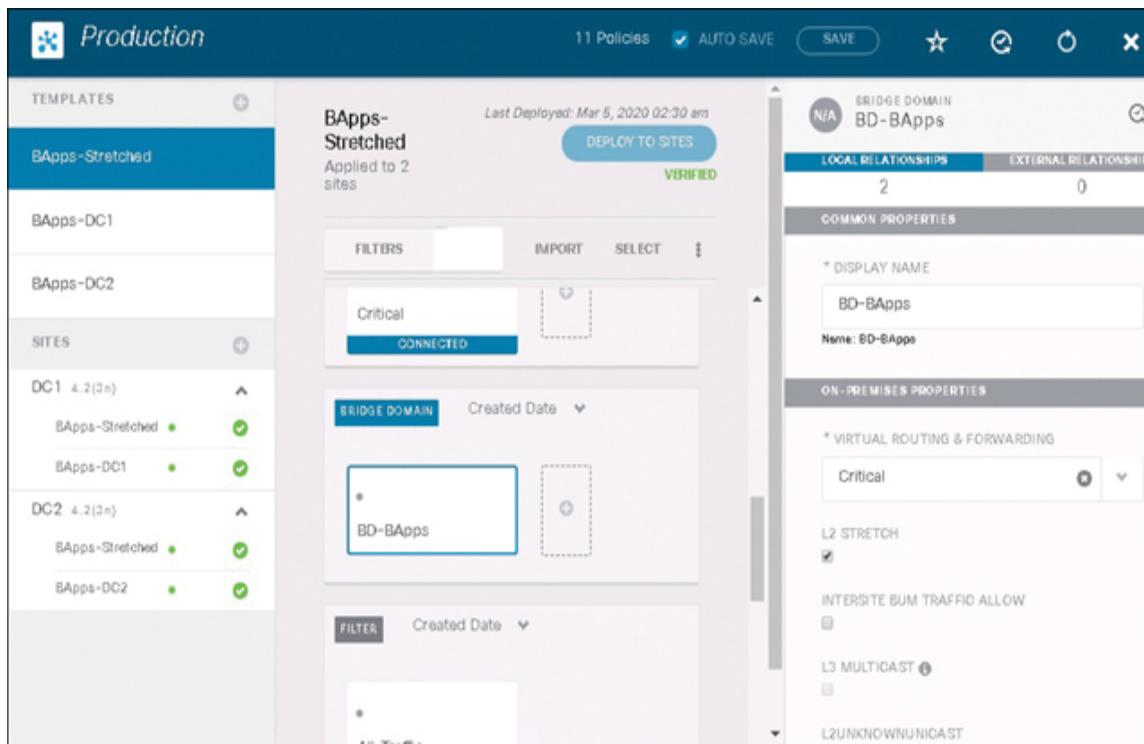


**Figure 16-3** Fault Management Across Multiple Data Centers with ACI Multi-Site

These centralized orchestration and management capabilities are important for managing primary and disaster recovery data centers because they ensure that policy is in sync and that the disaster recovery data center is always ready or can be made ready quickly if disaster strikes.

## Tweaking Broadcast and Stretch Settings on a Per-BD Basis

Traditional cross-site VLAN stretching solutions do not offer a lot of granularity in terms of how traffic is stretched across sites. Solutions such as Cisco Overlay Transport Virtualization (OTV) block a few select protocols, but data center solutions today generally lack the configuration knobs needed to have different policies at an individual subnet level. With ACI, however, stretch settings can be tweaked for any given requirement. [Figure 16-4](#) illustrates how a bridge domain named BD-BApps can be configured as a stretched subnet without the need for intersite broadcast, unknown unicast, and multicast (BUM) forwarding to take place. Settings such as Intersite BUM Traffic Allow ensure that subnets can be stretched without flooding remote data centers with packets that can adversely impact performance and reliability. This is a critical requirement for a lot of disaster recovery data centers today.



**Figure 16-4** Preventing BUM Traffic Across an ISN for a Single Bridge Domain

Important per-bridge domain tweaks to consider when designing primary and disaster recovery data centers include the following:

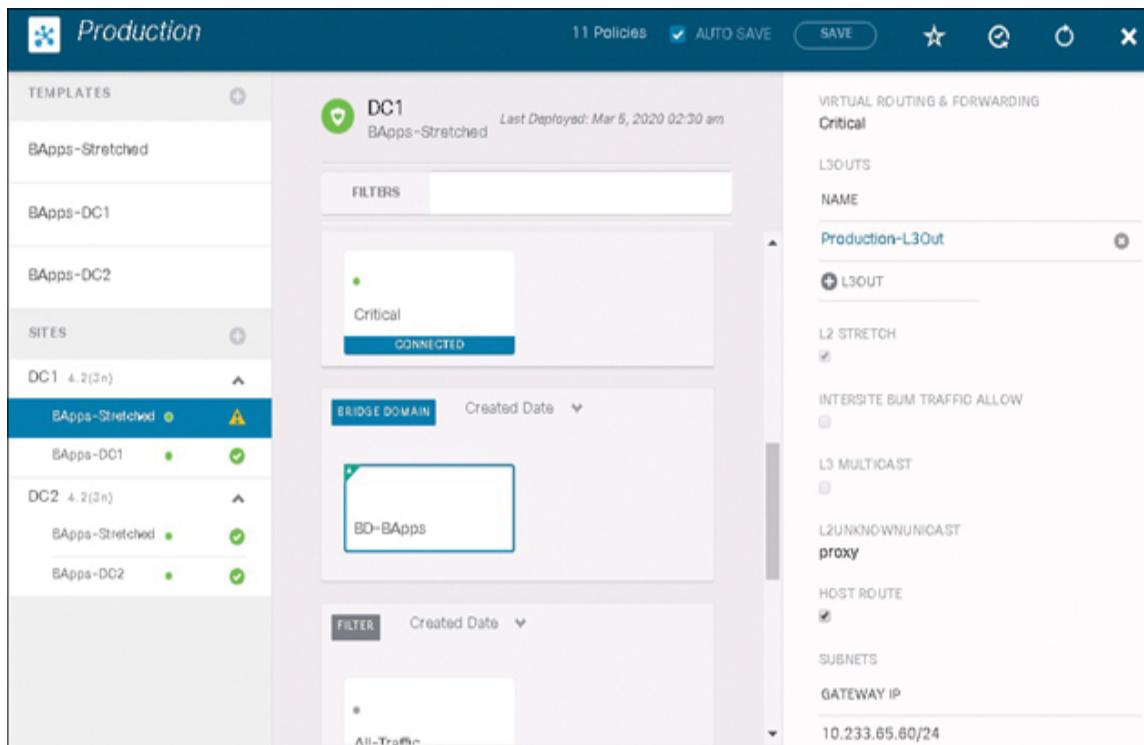


- **Layer 3 only across sites:** For this type of scenario, the L2 Stretch checkbox is disabled. If the bridge domain is placed in a stretched template, different Layer 3 subnets can be assigned to the bridge domain in each site. This conforms closely with a lot of primary/disaster recovery data center designs.
- **IP mobility without BUM flooding:** This involves the ability to perform vMotion across sites without cross-site flooding by enabling the L2 Stretch setting and disabling the Intersite BUM Traffic Allow setting.
- **Layer 2 flooding across sites:** Where flooding across sites may be a requirement for a subnet (for example, some type of cross-site clustered service), the L2 Stretch and Intersite BUM Traffic Allow settings can both be selected.

## Cross-Data Center Ingress Routing Optimizations

To direct traffic to the right data center over the WAN, ACI enables the advertisement of host routes out L3Outs. This is also supported with ACI Multi-Site. [Figure 16-5](#) shows how the Host Route feature is configured at the individual site level. This enables some interesting designs. For example, an administrator can enable host-based routing out only the disaster recovery data center, while the subnet itself is only advertised out the primary data center. A route map can

then be used to prevent advertisement of the subnet itself from the disaster recovery data center. If, at any point, endpoints within the subnet need to be migrated to the disaster recovery data center, routing tables converge to prefer the more specific host routes for endpoints that have been moved to the disaster recovery fabric. This provides an easy way to ensure that traffic is sent to the right data center while keeping enterprise routing tables manageable.



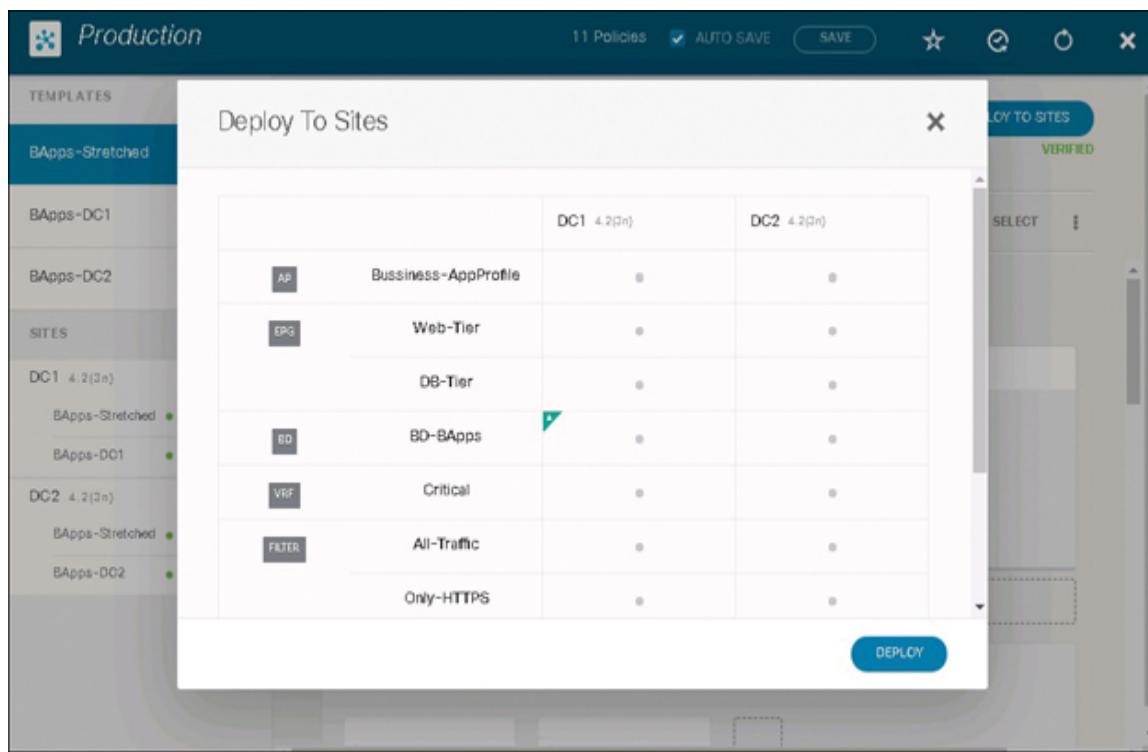
**Figure 16-5** Using Host-Based Routing for Ingress Traffic Optimization in ACI Multi-Site

## Simultaneous or Independent Policy Deployment to Sites

One top-of-mind issue for a lot of engineers who design and support primary and disaster recovery data centers is to guarantee that configuration changes can be made

independently to data centers, so that if a change goes bad, both data centers are not affected.

While templates and tenants do allow for some control over the scope of changes, verifications within MSO also confirm which objects will be created, deleted, or modified as a result of a change. [Figure 16-6](#) shows a relatively safe change in which the only object being modified is a bridge domain named BD-BApps. MSO also confirms that the changes will only be applied to DC1.



**Figure 16-6** Confirming Configuration Changes Prior to Deployment to Sites

Note that ACI Multi-Site stretched templates currently make configuration changes simultaneously across multiple sites. However, new features are being introduced to enable configurations impacting multiple data centers to be pushed to one fabric at a time. This feature will make ACI Multi-Site

a powerful orchestration tool for primary and disaster recovery data centers.

## **Building Active/Active Data Centers with ACI**

Whereas primary and disaster recovery environments are typically loosely coupled data centers with limited stretched subnets, active/active designs involve tightly coupled data centers and sometimes involve stretching all subnets between data centers. Active/active data centers are usually deployed short distances from one another to accommodate the hitless migration of workloads from one data center to another. The focus with active/active data centers is disaster avoidance and prevention.

There are multiple definitions for active/active data centers today. Some engineers identify deployment of active/active firewall clustering or breaking up active/standby firewall pairs across the data centers as a requirement for acknowledging a design as being active/active. The thought process is that minimal or no loss of firewall state information should occur in such data centers upon the loss of one of the data centers. Other engineers believe data centers that rely on global server load balancing and placement of active workloads in each data center are active/active designs. With this definition, loosely coupled primary and disaster recovery data centers that have no stretched subnets or any ability to proactively migrate workloads could also be called active/active data centers.

This book defines active/active a bit in between these two definitions. For a network design to be considered active/active, all or the majority of subnets should exist in both data centers, and the network should be able to direct traffic to the intended firewalls and other stateful services

appliances at all times—even if an entire data center is lost. This definition does not necessarily require that active/active clustering be used or that TCP transaction state never be lost. However, workloads should be able to easily migrate from one data center to another with ease for disaster avoidance and prevention. Using this definition, ACI Multi-Pod often wins against ACI Multi-Site as the ideal solution for active/active design requirements. (This is not meant to imply that ACI Multi-Site cannot be used for active/active data center design.)

## **VMM Integrations Applicable to Multiple Data Centers**

One of the reasons ACI Multi-Pod is ideal in active/active designs is that ACI Multi-Pod has a single APIC cluster spread across sites and can therefore have a single VMM integration for each vCenter data center object, even if the ESXi hosts within the data center object reside in both data centers. With ACI Multi-Site, however, separate APIC clusters in each data center also mean that each data center has to perform VMM integration separately with each vCenter instance.

Using ACI Multi-Site to implement separate integrations in each data center does not rule out cross-vCenter vMotion capabilities; however, it does make some features, such as cross-site vSphere high availability, fault tolerance, and distributed resource scheduling, impossible when VMM integration is performed. This is because separate DVS instances are deployed to each data center.

In many active/active designs deployed across short distances, such as in stretched metro cluster designs, the idea is to use advanced features such as vSphere high availability, fault tolerance, and distributed resource

scheduling to avoid disasters in the first place. And, in these designs, ACI Multi-Pod often becomes the solution of choice.

## **Stateful Services Integration in ACI Multi-Pod and Multi-Site**

There are three primary stateful services integration models that can be used to attach devices directly to ACI Multi-Pod or Multi-Site fabrics:



- **Each active/standby appliance pair deployed across data centers:** In this connectivity model, each pair of active/standby appliances is broken up and attached to the fabric in separate data centers. This is a common design in data centers where there is no sensitivity to asymmetric traffic flows because one appliance will always be actively forwarding traffic while the other will be in a ready state, waiting to take over in the case of an outage. ACI Multi-Pod fully supports this attachment model, but ACI Multi-Site supports this model only when the firewalls are running in L3 or L2 mode and function as the default gateway for endpoints.
- **Active/active stateful services clustering across data centers:** In this model, stateful services appliances are deployed in an active/active cluster across sites. This attachment model requires that ACI be able to learn the same endpoint data (MAC address and IP information) in both sites at the same time. While this connectivity model is fully supported with ACI

Multi-Pod today, ACI Multi-Site does not support active/active clustering as of the time of this writing.

- **Independent active/standby appliances in each data center:** When there is no requirement for proactive migration of endpoints between data centers, engineers can deploy separate firewall pairs in each data center and use either PBR or host-based routing to direct traffic to the right firewall pair. Both ACI Multi-Pod and ACI Multi-Site support use of independent active/standby stateful services appliances in each data center. However, use of independent firewall pairs in each data center is often seen more as a use case for ACI Multi-Site deployments due to the loss of firewall state information during failures or during cross-site workload migrations.

Of these three attachment models, the first two are more in line with the objectives of active/active data centers because they are able to ensure preservation of firewall state information during failures. This is why ACI Multi-Pod is often the platform of choice for active/active data center designs. This should not be understood as a rule but as general guidance. This also does not mean that ACI Multi-Site will never be needed as a tool to enable disaster recovery capabilities for the active/active pods. Furthermore, there is also the issue of product development. As ACI Multi-Site continues to mature, it may one day support a broader number of stateful services connectivity models.

## Extending ACI to Remote Locations and Public Clouds

The increasing popularity of public cloud offerings and the desire to place workloads closer to users mean that

engineers need to be more flexible about the changing boundaries of data centers. ACI Anywhere has options to extend data center boundaries to meet these types of new requirements without the need to sacrifice security, stability, or agility.

## **Extending ACI into Public Clouds with ACI Multi-Site**

Companies that embrace ACI within the data center can integrate public cloud environments such as AWS and Azure into ACI Multi-Site and manage these environments as additional sites within the overall deployment. This integration requires use of an additional product called the Cisco Cloud APIC, which translates ACI Multi-Site constructs into the relevant public cloud constructs and deploys policy into the public cloud by using cloud-native APIs.

The benefits of this approach include consistent network policies and security posture across clouds and on-premises data centers; secure automated connectivity and network stitching, centralized orchestration, visibility, and monitoring; and seamless workload migration across environments.

This solution requires the use of the Cisco Cloud Service Router (CSR) 1000V.

## **Extending ACI into Bare-Metal Clouds with vPod**

In addition to extension into popular public clouds, ACI can extend into bare-metal clouds by using the vPod solution. At the time of this writing, vPod cannot be configured as an independent site in ACI Multi-Site and needs to be

configured as a pod within a larger ACI Multi-Pod deployment.

The vPod solution consists of virtual leaf (vLeaf), virtual spine (vSpine), and ACI Virtual Edge (AVE). The vPod solution can also be deployed in data centers that are too small to justify the cost of an ACI fabric deployment.

## **Integrating Remote Sites into ACI Using Remote Leaf Switches**

Remote leaf switches significantly ease application migrations for the purpose of data center exits. They are also great for permanent use at locations where small numbers of servers need to be housed. They are also very easy to deploy.

One aspect of remote leafs, however, is failover. Engineers deploying remote leaf switches associate remote leafs with a pod. If remote leafs are deployed in ACI Multi-Pod environments that span multiple data centers, they can potentially fail over to another data center if there is an outage within the original data center or pod to which the leafs were associated.

ACI remote leaf does not support failover across sites within Multi-Site deployments as of the time of writing. Cross-data center failover is currently a remote leaf use case for ACI Multi-Pod only.

### **Exam Preparation Tasks**

As mentioned in the section “How to Use This Book” in the Introduction, you have a couple of choices for exam preparation: [Chapter 17, “Final Preparation,”](#) and the exam simulation questions on the companion website.

# Review All Key Topics

Review the most important topics in this chapter, noted with the Key Topic icon in the outer margin of the page. [Table 16-4](#) lists these key topics and the page number on which each is found.



**Table 16-4** Key Topics for [Chapter 16](#)

Key Topic Element	Description	Page Number
<a href="#">Table 16-2</a>	Defines some key ACI Multi-Site cross-site configuration parameters	<a href="#">555</a>
<a href="#">Table 16-3</a>	Defines a number of new constructs and concepts that are unique to ACI Multi-Site	<a href="#">557</a>
List	Lists some key tweaks that can be made at the bridge domain level using the MSO to make ACI Multi-Site ideal for primary/DR data centers	<a href="#">560</a>
List		<a href="#">563</a>

	Describes the supported stateful services attachment models for ACI Multi-Pod and ACI Multi-Site	
--	--	--

## Memory Tables

There are no memory tables or lists in this chapter.

## Define Key Terms

Define the following key terms from this chapter and check your answers in the glossary:

site  
site ID  
overlay multicast TEP  
overlay unicast TEP  
BGP EVPN router ID  
CloudSec encryption schema  
template  
stretched  
template conformity  
intersite L3Out

# **Part VII: Final Preparation**

# Chapter 17

## Final Preparation

The first 16 chapters of this book cover the technologies, protocols, design concepts, and considerations required to be prepared to pass the Implementing Cisco Application Centric Infrastructure DCACI 300-620 exam. While these chapters supply the detailed information, most people need more preparation than simply reading the first 16 chapters of this book. This chapter details a set of tools and a study plan to help you complete your preparation for the exam.

This short chapter has several main sections. The first section helps you get ready to take the exam, and the second section lists exam preparation tools that can be useful at this point in the study process. The final section provides a suggested study plan you can use now that you have completed all the earlier chapters in this book.

### Getting Ready

Here are some important tips to keep in mind to ensure that you are ready for this rewarding exam:

- **Build and use a study tracker:** Consider using the exam objectives shown in this chapter to build a study tracker for yourself. Such a tracker can help ensure that you have not missed anything and that you are confident for your exam. As a matter of fact, this book

offers a sample study planner as [Appendix E](#) on the companion website.

- **Think about your time budget for questions on the exam:** When you do the math, you will see that, on average, you have one minute per question. While this does not sound like a lot of time, keep in mind that many of the questions will be very straightforward, and you will take 15 to 30 seconds on those. This leaves you extra time for other questions on the exam.
- **Watch the clock:** Check in on the time remaining periodically as you are taking the exam. You might even find that you can slow down pretty dramatically if you have built up a nice block of extra time.
- **Get some earplugs:** The testing center might provide earplugs but get some just in case and bring them along. There might be other test takers in the center with you, and you do not want to be distracted by their screams. I personally have no issue blocking out the sounds around me, so I never worry about this, but I know it is an issue for some.
- **Plan your travel time:** Give yourself extra time to find the center and get checked in. Be sure to arrive early. As you test more at a particular center, you can certainly start cutting it closer time-wise.
- **Get rest:** Most students report that getting plenty of rest the night before the exam boosts their success. All-night cram sessions are not typically successful.
- **Bring in valuables but get ready to lock them up:** The testing center will take your phone, your smartwatch, your wallet, and other such items and will provide a secure place for them.
- **Take notes:** You will be given note-taking implements and should not be afraid to use them. I always jot down

any questions I struggle with on the exam. I then memorize them at the end of the test by reading my notes over and over again. I always make sure I have a pen and paper in the car, and I write down the issues in my car just after the exam. When I get home—with a pass or fail—I research those items!

## **Tools for Final Preparation**

This section lists some information about the available tools and how to access the tools.

### **Pearson Cert Practice Test Engine and Questions on the Website**

Register this book to get access to the Pearson IT Certification test engine (software that displays and grades a set of exam-realistic multiple-choice questions). Using the Pearson Cert Practice Test Engine, you can either study by going through the questions in Study mode or take a simulated (timed) exam.

The Pearson Test Prep practice test software comes with two full practice exams. These practice tests are available to you either online or as an offline Windows application. To access the practice exams that were developed with this book, please see the instructions in the card inserted in the sleeve in the back of the book. This card includes a unique access code that enables you to activate your exams in the Pearson Test Prep software.

### **Accessing the Pearson Test Prep Software Online**

The online version of this software can be used on any device with a browser and connectivity to the Internet including desktop machines, tablets, and smartphones. To start using your practice exams online, simply follow these steps:

**Step 1.** Go to <http://www.PearsonTestPrep.com>.

**Step 2.** Select **Pearson IT Certification** as your product group.

**Step 3.** Enter your email and password for your account. If you don't have an account on PearsonITCertification.com or CiscoPress.com, you need to establish one by going to PearsonITCertification.com/join.

**Step 4.** In the My Products tab, click the **Activate New Product** button.

**Step 5.** Enter the access code printed on the insert card in the back of your book to activate your product. The product is then listed in your My Products page.

**Step 6.** Click the **Exams** button to launch the exam settings screen and start the exam.

## **Accessing the Pearson Test Prep Software Offline**

If you wish to study offline, you can download and install the Windows version of the Pearson Test Prep software. You can find a download link for this software on the book's companion website, or you can just enter this link in your browser:

<http://www.pearsonitcertification.com/content/downloads/pcpt/engine.zip>

To access the book's companion website and the software, simply follow these steps:

- Step 1.** Register your book by going to PearsonITCertification.com/register and entering the ISBN: **9780136602668**.
- Step 2.** Respond to the challenge questions.
- Step 3.** Go to your account page and select the **Registered Products** tab.
- Step 4.** Click on the **Access Bonus Content** link under the product listing.
- Step 5.** Click the **Install Pearson Test Prep Desktop Version** link in the Practice Exams section of the page to download the software.
- Step 6.** When the software finishes downloading, unzip all the files onto your computer.
- Step 7.** Double-click the application file to start the installation and follow the onscreen instructions to complete the registration.
- Step 8.** When the installation is complete, launch the application and click the **Activate Exam** button on the My Products tab.
- Step 9.** Click the **Activate a Product** button in the Activate Product Wizard.
- Step 10.** Enter the unique access code from the card in the sleeve in the back of your book and click the **Activate** button.

**Step 11.** Click **Next** and then click the **Finish** button to download the exam data to your application.

**Step 12.** You can now start using the practice exams by selecting the product and clicking the **Open Exam** button to open the exam settings screen.

Note that the offline and online versions sync together, so saved exams and grade results recorded on one version will be available to you in the other version as well.

## Customizing Your Exams

When you are in the exam settings screen, you can choose to take exams in one of three modes:

- Study mode
- Practice Exam mode
- Flash Card mode

Study mode allows you to fully customize an exam and review answers as you are taking the exam. This is typically the mode you use first to assess your knowledge and identify information gaps. Practice Exam mode locks certain customization options in order to present a realistic exam experience. Use this mode when you are preparing to test your exam readiness. Flash Card mode strips out the answers and presents you with only the question stem. This mode is great for late-stage preparation, when you really want to challenge yourself to provide answers without the benefit of seeing multiple-choice options. This mode does not provide the detailed score reports that the other two modes provide, so it is not the best mode for helping you identify knowledge gaps.

In addition to these three modes, you will be able to select the source of your questions. You can choose to take exams that cover all of the chapters, or you can narrow your selection to just a single chapter or the chapters that make up specific parts in the book. All chapters are selected by default. If you want to narrow your focus to individual chapters, simply deselect all the chapters and then select only those on which you wish to focus in the Objectives area.

You can also select the exam banks on which to focus. Each exam bank comes complete with full exams of questions that cover topics in every chapter. You can have the test engine serve up exams from all four banks or just from one individual bank by selecting the desired banks in the exam bank area.

There are several other customizations you can make to your exam from the exam settings screen, such as the time allowed for taking the exam, the number of questions served up, whether to randomize questions and answers, whether to show the number of correct answers for multiple-answer questions, and whether to serve up only specific types of questions. You can also create custom test banks by selecting only questions that you have marked or questions on which you have added notes.

## **Updating Your Exams**

If you are using the online version of the Pearson Test Prep software, you should always have access to the latest version of the software as well as the exam data. If you are using the Windows desktop version, every time you launch the software, it will check to see if there are any updates to your exam data and automatically download any changes that have been made since the last time you used the

software. This requires that you be connected to the Internet at the time you launch the software.

Sometimes, due to a number of factors, the exam data might not fully download when you activate your exam. If you find that figures or exhibits are missing, you might need to manually update your exams.

To update a particular exam you have already activated and downloaded, simply select the Tools tab and click the Update Products button. Again, this is only an issue with the desktop Windows application.

If you wish to check for updates to the Windows desktop version of the Pearson Test Prep exam engine software, simply select the Tools tab and click the Update Application button. Doing so allows you to ensure that you are running the latest version of the software engine.

## Premium Edition

In addition to the free practice exam provided on the website, you can purchase additional exams with expanded functionality directly from Pearson IT Certification. The Premium Edition of this title contains an additional two full practice exams and an eBook (in both PDF and ePUB format). In addition, the Premium Edition title also has remediation for each question to the specific part of the eBook that relates to that question.

Because you have purchased the print version of this title, you can purchase the Premium Edition at a deep discount. There is a coupon code in the book sleeve that contains a one-time-use code and instructions for where you can purchase the Premium Edition.

To view the premium edition product page, go to [www.informit.com/title/9780136602668](http://www.informit.com/title/9780136602668).

## Suggested Plan for Final Review/Study

This section lists a suggested study plan from the point at which you finish reading through [Chapter 16](#) until you take the DCACI 300-620 exam. You can ignore this plan, use it as is, or take suggestions from it.

The plan involves four steps:

### **Step 1. Review Key Topics and “Do I Know This Already?” (DIKTA) Questions:**

**already?**: You can use the table that lists the key topics in each chapter or just flip the pages looking for key topics. Also, reviewing the DIKTA? questions from the beginning of the chapter can be helpful for review.

### **Step 2. Complete Memory Tables:** Open [Appendix C](#),

“[Memory Tables](#)” (found on the companion website), and print the entire appendix or print the tables by major part. Then complete the tables.

### **Step 3. Gain hands-on experience:** Nothing can replace real experience with an ACI fabric. The best way to gain hands-on experience is to administer an ACI fabric or use a lab environment.

### **Step 4. Use the Pearson Test Prep Practice Test engine to practice:**

The Pearson Test Prep practice test engine enables you to study using a bank of unique exam-realistic questions available only with this book.

## Summary

The tools and suggestions listed in this chapter have been designed with one goal in mind: to help you develop the skills required to pass the DCACI 300-620 exam. This book has been developed from the beginning to not just tell you the facts but to also help you learn how to apply the facts. No matter what your experience level leading up to when you take the exam, it is our hope that the broad range of preparation tools, and even the structure of the book, will help you pass the exam with ease. We hope you do well on the exam.

# Appendix A

## Answers to the “Do I Know This Already?” Questions

### Chapter 1

- 1.** A, C, D. Level of expertise, number of managed endpoints, and difficulty of information correlation across devices all contribute to management complexity. However, open standards *do not* contribute to management complexity, even though different vendor interpretations of open standards do contribute to network management complexity.
- 2.** A. 12 header bits are used in the definition of VLAN IDs. Out of the resulting 4096 VLAN IDs, 2 are reserved and cannot be used for data traffic, resulting in 4094 usable VLAN IDs.
- 3.** B. Firewalls in traditional networks are primarily used for securing north-south traffic flows.
- 4.** B. Internally, ACI uses VXLAN. VLANs are primarily used to classify inbound traffic into EPGs and also to enable trunking to servers.
- 5.** D. Multi-Site Orchestrator (MSO) is used to orchestrate configurations across multiple ACI fabrics.
- 6.** A. Stateless networking involves using node IDs as the identities of switches and APICs and allows a device to

be decommissioned and another device commissioned with minimal changes to the network.

7. A. Blacklisting is the practice of allowing all traffic except that which is denied through security mechanisms like access lists. Blacklisting is a feature of traditional switching and routing solutions.
8. B. ACI has been built around policy reuse, which enables companies that do not have a strong desire for automation to still achieve better agility.
9. C. In ACI multipod, each pod forms its own control plane. While each site in an ACI multisite deployment also forms separate control planes, each site is also considered a distinct fabric.
10. D. Microsegmentation can be accomplished within a single tenant and is not by itself a technical driver for deploying multiple tenants.

## Chapter 2

1. B. If new spines cannot be deployed at the remote site, then ACI Multi-Pod and Multi-Site are not valid options. Furthermore, ACI Multi-Tier generally assumes that cables will be run directly between spines and leaf switches or, alternatively, between Tier 1 and Tier 2 leaf switches without the use of an ISN or IPN. ACI Remote Leaf, on the other hand, can be deployed at remote locations through an IPN and therefore does not require dedicated cross-site circuits or fiber.
2. D. The requirement for Multicast PIM-Bidir is unique to ACI Multi-Pod. ACI Multi-Site, on the other hand, uses the ingress replication function of the spine nodes in the source site to replicate BUM traffic to all remote

sites to which a given cross-site BUM-enabled bridge domain has been stretched.

- 3.** D. APIC-to-leaf connectivity as well as leaf-to-spine connectivity have been valid connectivity options since the first shipments of ACI hardware. With ACI Multi-Tier leaf-to-leaf cabling is acceptable as long as the cabling connects a Tier 1 leaf to a Tier 2 leaf. As of the time of this writing, there are no use cases for spine-to-spine connectivity. ACI disables spine-to-spine connections when it detects them.
- 4.** A, C. Ingress replication of BUM traffic and namespace normalization are both functionalities that are enabled by second-generation spine hardware. IP fragmentation is not supported in ACI. Multicast forwarding via PIM-Bidir applies to ACI Multi-Pod and not ACI Multi-Site.
- 5.** A. Border leaf switches provide Layer 2 and Layer 3 connectivity to outside networks, meaning connectivity to non-ACI switches and routers.
- 6.** B. At the time of this book's publication, a cluster of at least four L3 APICs is necessary to scale to 200 leaf switches. Sharding is a result of the evolution of what is called horizontal partitioning of databases. In an ACI deployment that has three or more APICs, there will always be three shards for each attribute in the APIC database. Standby APICs do not actively synchronize data with active APICs. They are passive players in an ACI fabric that need to be manually activated to replace a failed APIC and restore the APIC quorum.
- 7.** B, C. Nexus 93180YC-EX is a cloud-scale leaf switch. Nexus 9364C is a second-generation spine and supports ACI Multi-Site. Nexus 9736C-FX is a line card that can be populated into a Nexus 9500 Series spine

and supports ACI Multi-Site. The Nexus 9396PX is a first-generation leaf switch.

- 8.** A, C. The Nexus 9336PQ is a first-generation spine. It does not support the namespace normalization required for ACI Multi-Site support. It does, however, support 40 Gbps connectivity. Therefore, an upgrade from Nexus 9336PQ would not help increase bandwidth unless 100 Gbps leaf-to-spine connectivity is the desired target. This upgrade would allow enablement of CloudSec. ACI Multi-Pod is supported in first-generation hardware.
- 9.** A. All interfaces with speeds of 10 Gbps or higher on leaf switches whose model numbers end with FX support MACsec.
- 10.** D. The Nexus 9348GC-FXP is well positioned for low-bandwidth 100 Mbps and 1 Gigabit Ethernet use cases that rely on RJ-45 transceiverless connectivity.

## Chapter 3

- 1.** D. The first three options listed are spine switches and therefore do not allow APIC connectivity. The Nexus 93180YC-FX, on the other hand, is a leaf, and in-band APIC ports can therefore be connected to it to allow fabric discovery.
- 2.** A, C. If the infrastructure VLAN ID or fabric ID needs to change, a fabric rebuild is required. However, APIC OOB IP addresses can be changed after initialization. The active or standby status of a controller may necessitate that an individual APIC be re-initialized with a standby APIC node ID, but it does not necessitate the re-initialization of an entire fabric.

- 3.** B. Following the establishment of encrypted IFM communication channels, APICs are able to push configuration changes to ACI switches. It is at this point that a switch is considered to have become activated and fully operational.
- 4.** B, C. A seed leaf needs to be activated first before a spine can be activated. In-band APIC interfaces function in an active/standby mode. If the APIC interface that is active is not connected to an ACI leaf switch, the discovery will fail. For an APIC to add a switch to the Fabric Membership tab to allow administrators to authorize the switch to join the fabric, the APIC needs to see LLDP packets from the seed leaf and must also be able to process DHCP Discover packets from it. APICs cannot possibly form a cluster until at least a seed leaf has been discovered.
- 5.** B. All options listed except B have minimal impact on the fabric discovery process. However, the network mode should remain set to Dedicated. Otherwise, unexpected connectivity problems can occur.
- 6.** D. Out of the options listed, Proxy-TEP is the only option that exclusively references spine switches.
- 7.** C. Atomic Replace allows the replacement of all current settings with configurations from a file. Atomic Merge and Best Effort Merge both enable merging configurations with currently configured settings. There is no such thing as a Best Effort Replace.
- 8.** B, C, D. ACI supports forwarding of backups and restoration of backups via FTP, SFTP, and SCP.
- 9.** D. Using four upgrade groups—two for spines and two for leafs with node ID-based separation into groups based on odd and even node IDs—tends to enable the most resilient setup. Option A results in fabricwide

outages. Option B can result in zero downtime, but a number of spines and leafs are under upgrade simultaneously. Option C with random node assignment can easily lead to downtime.

- 10.** A. By tying the export operation to a scheduler, an administrator can have ACI perform automated scheduled backups.

## Chapter 4

- 1.** C. Logging in to the CLI via a non-default login domain requires the special login syntax of `apic# domain\username`.
- 2.** B. The APICs are the brains of an ACI network. For the most part, the switch CLI only allows use of **show** and **debug** commands that enable verification that desired configurations have been pushed to a device and also verification of device function.
- 3.** B. An administrator can enable or disable supported management protocols or change the ports associated with the enabled management access methods by editing the management access policy associated with active pod policy groups.
- 4.** C. Access policies primarily govern the configuration and operation of non-fabric (access) ports, including parameters such as link speed and other port-level configurations, including LLDP and LACP.
- 5.** C. Fabric policies govern configurations that apply more holistically at the switch or pod level.
- 6.** A, D. MOQuery is a CLI-based tool used for querying the ACI object hierarchy, and Visore is the GUI-based equivalent.

- 7.** B. The Raised state almost always suggests the existence of an active problem in the network. Faults in the Raised state remain in this state until the underlying condition is resolved.
- 8.** A, D. Answer A is correct because acknowledgment of a fault whose underlying condition has been resolved leads to the deletion of the fault. Answer D is also correct. The default amount of time faults are retained after their underlying conditions are resolved is 3600 seconds, and this interval is called the retention interval. Note that while events are immutable and serve as the permanent record of occurrences in the system, faults follow a lifecycle.
- 9.** D. The Common policy has a global fabricwide scope and deals with monitoring of objects such as the APIC controllers and fabric nodes. The policies configured in this class are also used when there is no corresponding policy under the more specific infra or tenant scopes.
- 10.** B. Health scores are meant to provide a means for periodic reporting of the health and operational status of ACI objects, tenants, pods, or entire fabrics.

## Chapter 5

- 1.** C. Fabric APICs push policy to leaf switches via in-band connections into the infra tenant. The infra tenant forms both the underlay and overlay for intra-fabric communication. The infra tenant does not get exposed to the user space (tenants), and it has its own private network space and bridge domains.
- 2.** C. The common tenant is a special tenant whose purpose is to provide common services to other tenants in an ACI fabric. The common tenant enables

the creation of shared L3Outs, shared bridge domains, and shared VRF instances.

3. D. An endpoint group is a grouping of physical or virtual network endpoints that reside within a single bridge domain and have similar policy requirements. ACI can classify both IP-based and non-IP-based endpoints into EPGs. An EPG does not necessarily define a broadcast domain, and it is bridge domains that best correlate to broadcast domains.
4. B. An application profile is a container that allows EPGs to be grouped according to their relationship with one another to enable easier configuration and auditing of relevant policies or to make policy reuse possible.
5. D: The command **show ip route** needs to be followed up with the name of the intended tenant and the VRF. Hence, the command **show ip route CCNP:DCACI** displays the routes in a VRF named DCACI, which resides in a tenant named CCNP.
6. C. A filter matches interesting traffic and consists of a series of filter entries. A subject determines the action to be taken on traffic that is matched, and a contract is directionally applied to EPGs to determine the traffic flows to which the contract applies.
7. B. The setting Intra EPG Isolation allows an administrator to enforce traffic isolation between endpoints that reside in an EPG. This feature is configured at the EPG level.
8. D. Where a client is requesting a service from another system, the client is a consumer of the server-side service, and the server is the provider of the service.
9. A. An external EPG is a special type of EPG that represents endpoints outside an ACI fabric, such as

user laptops, campus IoT devices, or Internet users. External EPGs are applied at the VRF instance level on an object called an L3Out.

- 10.** A. There is an  $n:1$  relationship between filters and tenants, meaning that large numbers of filters can be created in each tenant.

## Chapter 6

- 1.** D. A physical domain governs the attachment of bare-metal servers and appliances that need static VLAN allocations.
- 2.** B. Even if an administrator assigns a VLAN pool to an AAEP and assigns the AAEP to ports, the VLANs in the pool do *not* get activated on the switch ports. The VLAN pool, in itself, just defines the potential list of VLAN IDs that a tenant administrator can use to map EPGs to switch ports. For traffic to flow over VLANs, one or more EPGs still need to be mapped to ports either by a tenant administrator or via direct mappings on the AAEP, which is called an AAEP EPG.
- 3.** A. VMM domains allow both static and dynamic mapping of EPGs to VLANs for virtualized environments. The use of dynamic mapping of EPGs is recommended.
- 4.** D. A tenant administrator who wants to map an EPG to ports and encapsulations should first bind the EPG to a domain. The domain association by the tenant administrator acts as authorization that the endpoints within the specified EPG are indeed meant to connect to the fabric via the method specified by the domain. The domains that are bound to the EPG also indicate

which VLAN IDs are acceptable encapsulations for the EPG.

5. B. An EPG can be bound to multiple domains, but the domains should ideally not reference overlapping VLAN pools. This is especially important if there are leaf switches that have the EPG mapped to an encapsulation using more than one of the domains. EPGs can be bound statically to physical ports for bare-metal connectivity. EPGs are not directly associated with VRF instances. The construct that does get mapped directly to a VRF is called a bridge domain.
6. C. ACI uses MisCabling Protocol (MCP) for loop prevention.
7. B. When port channel interface policy groups are reused on a switch, ACI interprets the reuse as an intention to add the new ports to the previous port channel bundle. Interface policy groups involving aggregations should generally not be reused.
8. A. Interface policy groups pertaining to non-aggregated ports can be reused—without caveats.
9. A. A switch can be assigned more than one interface profile.
10. C. A switch profile is a logical object that represents one or more physical switches from a configuration standpoint and directly points to node IDs. The use of switch profiles provides part of the basis for stateless networking in ACI.

## Chapter 7

1. C. The out-of-service status appears when access policies have been successfully deployed for a port but

there are no EPG mappings to the port, and therefore the port is not actively forwarding traffic.

- 2.** A. vPC domains or vPC explicit protection groups are configured at Fabric > Access Policies > Policies > Switch > Virtual Port Channel default. Another confusing location, Fabric > Access Policies > Policies > Switch > VPC Domain, is where vPC peer dead intervals are configured.
- 3.** B. Leaf access port policy groups can only be used for individual (non-aggregated) ports. To configure a port channel, configure a port channel interface policy group instead and associate it with leaf switch ports.
- 4.** B. Switches encapsulate traffic destined to an endpoint behind a vPC and send it to the destination loopback 1 IP address, which is common across both vPC peers.
- 5.** B. Interface policies can typically enable CDP without the need for overrides.
- 6.** A. This is an accurate statement.
- 7.** A. This is an accurate statement.
- 8.** B, C, D. MCP needs to be enabled globally and at the interface level. When MCP PDUs need to also be sent out on a tagged VLAN, Enable MCP PDU per VLAN options needs to be checked.
- 9.** B. This is true only if some sub-ports are left unused.
- 10.** B. By default, ACI attempts only to preserve DSCP settings.

## Chapter 8

- 1.** C. When an endpoint attaches to a leaf, the endpoint is considered local to the leaf. A leaf that learns an endpoint through local attachment is the most significant source of truth for the endpoint information in ACI, and therefore the leaf learns both the MAC address and any associated IP addresses for the endpoint.
- 2.** B. When a bridge domain does not have unicast routing enabled and a default gateway to leaf switches and BD L2 unknown unicast has been set to hardware proxy, ARP gleaning cannot take place, and traffic destined toward the silent host is dropped at the spine. Consequently, BD L2 unknown unicast should always be set to flood when an L2 BD houses a silent host.
- 3.** B. An endpoint in ACI is defined as a single MAC address and all IP addresses associated with it. If an IP address is flapping between MAC addresses, ACI detects a MAC duplication. This can impact endpoint learning and consequently lead to disruption of traffic toward and from the endpoint. If the issue is transient, it could indicate that an appliance failover has taken place and likely does not indicate an issue. If, on the other hand, the problem is non-transient, it should be investigated. Note that answers A and C are incorrect because a MAC address can have multiple IP address associations.
- 4.** A. The leaf to which an endpoint attaches propagates the endpoint information to the spines and thus is the holder of the single source of truth in the fabric.
- 5.** A. This is a large part of the definition of hardware proxy forwarding.
- 6.** B. With hardware proxy, ACI still needs ARP, but it is able to unicast ARP traffic from endpoints unless ARP