



Cisco HyperFlex 3.5 Stretched Cluster with Cisco ACI 4.0 Multi-Pod Fabric

Deployment Guide for Cisco HyperFlex 3.5 Stretched Cluster with Cisco ACI 4.0 Multi-Pod Fabric and VMware vSphere 6.5U2

Last Updated: August 30, 2019



About the Cisco Validated Design Program

The Cisco Validated Design (CVD) program consists of systems and solutions designed, tested, and documented to facilitate faster, more reliable, and more predictable customer deployments. For more information, go to:

<http://www.cisco.com/go/designzone>.

ALL DESIGNS, SPECIFICATIONS, STATEMENTS, INFORMATION, AND RECOMMENDATIONS (COLLECTIVELY, "DESIGNS") IN THIS MANUAL ARE PRESENTED "AS IS," WITH ALL FAULTS. CISCO AND ITS SUPPLIERS DISCLAIM ALL WARRANTIES, INCLUDING, WITHOUT LIMITATION, THE WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT OR ARISING FROM A COURSE OF DEALING, USAGE, OR TRADE PRACTICE. IN NO EVENT SHALL CISCO OR ITS SUPPLIERS BE LIABLE FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, INCLUDING, WITHOUT LIMITATION, LOST PROFITS OR LOSS OR DAMAGE TO DATA ARISING OUT OF THE USE OR INABILITY TO USE THE DESIGNS, EVEN IF CISCO OR ITS SUPPLIERS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

THE DESIGNS ARE SUBJECT TO CHANGE WITHOUT NOTICE. USERS ARE SOLELY RESPONSIBLE FOR THEIR APPLICATION OF THE DESIGNS. THE DESIGNS DO NOT CONSTITUTE THE TECHNICAL OR OTHER PROFESSIONAL ADVICE OF CISCO, ITS SUPPLIERS OR PARTNERS. USERS SHOULD CONSULT THEIR OWN TECHNICAL ADVISORS BEFORE IMPLEMENTING THE DESIGNS. RESULTS MAY VARY DEPENDING ON FACTORS NOT TESTED BY CISCO.

CCDE, CCENT, Cisco Eos, Cisco Lumin, Cisco Nexus, Cisco StadiumVision, Cisco TelePresence, Cisco WebEx, the Cisco logo, DCE, and Welcome to the Human Network are trademarks; Changing the Way We Work, Live, Play, and Learn and Cisco Store are service marks; and Access Registrar, Aironet, AsyncOS, Bringing the Meeting To You, Catalyst, CCDA, CCDP, CCIE, CCIP, CCNA, CCNP, CCSP, CCVP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unified Computing System (Cisco UCS), Cisco UCS B-Series Blade Servers, Cisco UCS C-Series Rack Servers, Cisco UCS S-Series Storage Servers, Cisco UCS Manager, Cisco UCS Management Software, Cisco Unified Fabric, Cisco Application Centric Infrastructure, Cisco Nexus 9000 Series, Cisco Nexus 7000 Series, Cisco Prime Data Center Network Manager, Cisco NX-OS Software, Cisco MDS Series, Cisco Unity, Collaboration Without Limitation, EtherFast, EtherSwitch, Event Center, Fast Step, Follow Me Browsing, FormShare, GigaDrive, HomeLink, Internet Quotient, IOS, iPhone, iQuick Study, LightStream, Linksys, MediaTone, MeetingPlace, MeetingPlace Chime Sound, MGX, Networkers, Networking Academy, Network Registrar, PCNow, PIX, PowerPanels, ProConnect, ScriptShare, SenderBase, SMARTnet, Spectrum Expert, StackWise, The Fastest Way to Increase Your Internet Quotient, TransPath, WebEx, and the WebEx logo are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries.

All other trademarks mentioned in this document or website are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (o8ogR)

© 2019 Cisco Systems, Inc. All rights reserved.

Table of Contents

Executive Summary.....	8
Solution Overview	9
Introduction.....	9
Audience	9
Purpose of this Document.....	9
What's New in this Release?	9
Solution Summary	9
Solution Deployment – ACI Fabric (Single Pod).....	13
Deployment Overview	13
Deploy ACI Fabric in Pod-1.....	13
Physical Connectivity	13
Initial Setup of APIC(s) in Pod-1	14
Deploy Spine and Leaf Switches in Pod-1	19
Apply Pod Policies for Pod-1.....	27
Configure DNS (Fabric Wide Setting).....	32
Enable/Review ACI Fabric Settings.....	33
COS Preservation (Fabric Wide Setting).....	34
Enforce Subnet Check for Endpoint Learning (Fabric Wide Setting)	35
Limit IP Learning to Subnet (Bridge-domain, Optional)	36
IP Aging (Fabric Wide Setting).....	37
Endpoint Learning.....	38
Endpoint Dataplane Learning	39
L2 Unknown Unicast	40
Clear Remote MAC Entries	41
Unicast Routing.....	41
ARP Flooding	41
GARP-based Detection.....	42
Jumbo Frames and MTU	43
Pre-configure Access Policies for ACI Fabric	43
Setup Information	44
Deployment Steps.....	44
Solution Deployment – ACI Fabric (To Outside Networks from Pod-1)	55
Deployment Overview	55
Create VLAN Pool for External Routed Domain	55
Configure Domain Type for External Routed Domain.....	57
Create AAEP for External Routed Domain	58

Configure Interfaces to External Routed Domain	60
Configure Tenant Networking for Shared L3Out.....	65
Configure External Routed Networks under Tenant Common.....	67
Create Contracts for External Routed Networks from Tenant (common).....	79
Provide Contracts for External Routed Networks from Tenant (common)	82
Configure External Gateways in the Outside Network.....	84
Solution Deployment – ACI Fabric (Multi-Pod).....	86
Prerequisites.....	86
Topology	86
Deployment Overview	87
Physical Connectivity	87
Deploy Inter-Pod Network (IPN)	87
Setup ACI Fabric for Multi-Pod	88
Setup Pod-2 Spine Switches, Leaf Switches, and APICs	88
Deployment Guidelines.....	88
Deploy Inter-Pod Network	89
Deployment Overview.....	89
Physical Connectivity	89
Configure IPN Devices in Pod-1	91
Configure IPN Devices in Pod-2	93
Setup ACI Fabric for Multi-Pod – Using Configuration Wizard.....	94
Prerequisites	94
Deployment Overview.....	94
Configure Inter-Pod Connectivity	95
Add Physical Pod – Second Pod or Site (Pod-2).....	102
Configure DHCP Relay on IPN Devices.....	106
Configure OSPF Interface Profile for Spines in Pod-2	107
Setup Fabric Access Policies for Spine Switches in Pod-1.....	108
Deployment Overview.....	108
Setup Information	109
Deployment Steps.....	110
Deploy ACI Fabric in Pod-2.....	114
Deployment Overview.....	114
Physical Connectivity	115
Deploy Spine and Leaf Switches in Pod-2	115
Configure NTP for Pod-2 using Out-of-Band Management	126
Update BGP Route Reflector Policy for Pod-2	128
Update Pod Profile to Apply Pod Policies.....	129

Setup Fabric Access Policies for Spine Switches in Pod-2.....	131
Deployment Overview.....	131
Setup Information.....	132
Deployment Steps.....	132
Deploy APICs in Pod-2	135
Prerequisites	135
Deployment Overview.....	135
Verify Pod-2 Switches are Part of the ACI Fabric	136
Initial Setup of Pod-2 APIC	137
Verify Pod-2 APIC is Part of the APIC Cluster	140
Add Pod-2 APIC as DHCP Relay Destination	141
Verify ACI Multi-Pod Fabric Setup	142
Verify OSPF Status on Spine Switches.....	142
Verify MP-BGP EVPN Status on Spine Switches	143
Verify COOP Status on Spine Switches.....	144
Solution Deployment – ACI Fabric (To Outside Networks from Pod-2)	146
Deployment Overview	146
Create VLAN Pool for External Routed Domain	146
Configure Domain Type for External Routed Domain.....	148
Create Attachable Access Entity Profile for External Routed Domain	149
Configure Interfaces to External Routed Domain	151
Configure Tenant Networking for Shared L3Out.....	159
Configure External Routed Networks under Tenant Common.....	159
Create Contracts for External Routed Networks from Tenant (common).....	172
Provide Contracts for External Routed Networks from Tenant (common)	172
Configure External Gateways in the Outside Network.....	173
Solution Deployment – ACI Fabric (To Cisco UCS Domains)	176
Deploy New Leaf Switches for Connectivity to Cisco UCS Domains	176
Topology.....	176
Setup Information	177
ACI Fabric Discovery of Leaf Switches	177
Add Nexus 9000 Series Leaf Switches to the ACI Fabric	178
Setup Out-of-Band Management for New Leaf Switches.....	181
Enable Access Layer Connectivity to Cisco UCS Domains	183
Topology.....	183
Enable 40Gbps Connectivity to Cisco UCS Domain	185
Enable Access Layer Configuration to Cisco UCS Domain	186
Solution Deployment – Setup Cisco UCS Domains	202

Setup Information.....	202
Bring Up UCS Domain with Fabric Interconnects	202
Initial Setup of Cisco UCS Domain	204
Enable Cisco Intersight Cloud-Based Management	207
Prerequisites	207
Cisco Intersight Licensing	207
Setup Information	208
Deployment Steps.....	208
Solution Deployment – Foundational Infrastructure for Cisco HyperFlex	212
Create Foundation Tenant and VRF	212
Configure ACI Fabric for HyperFlex In-Band Management	213
Configure ACI Fabric for HyperFlex Storage Data Traffic on HyperFlex Standard Cluster.....	221
Configure ACI Fabric for HyperFlex vMotion Traffic	227
Solution Deployment – HyperFlex Management Cluster	235
Topology	235
Install HyperFlex Cluster (Management) using Cisco Intersight	235
Prerequisites	235
Setup Information	236
Deployment Steps.....	239
Migrate Virtual Networking to VMware vDS on HyperFlex Management Cluster	258
Setup Information	258
Deployment Steps.....	259
Deploy Virtual Machines – Infrastructure Management	270
Configure ACI Fabric for Infrastructure Management.....	270
Deploy HX Installer Virtual Machine in the HyperFlex Management Cluster	282
Solution Deployment – HyperFlex Application Cluster	286
Topology	286
Deployment Overview	286
Setup Cisco UCS Domain for HyperFlex Stretched Cluster.....	287
Setup ACI Fabric for HyperFlex Stretched Cluster.....	287
Create Static Binding for In-Band Management to HyperFlex Stretched Cluster.....	287
Create Static Binding for vMotion to HyperFlex Stretched Cluster	289
Configure ACI Fabric for Storage Data Traffic on HyperFlex Stretched Cluster	291
Install HyperFlex Stretched Cluster (Applications) using Installer Virtual Machine	296
Prerequisites	296
Setup Information	297
Deployment Steps.....	302
Migrate Virtual Networking to Cisco AVE on HyperFlex Application Cluster.....	331

Setup Information	331
Deployment Overview	332
Deployment Steps.....	333
Deploy Cisco ACI vSphere Plug-in.....	357
Solution Deployment – Onboarding Multi-Tier Applications	366
Deployment Overview	366
Prerequisites.....	366
Configure ACI constructs for Application	366
Create Tenant and VRF for Application	366
Configure Bridge Domains	367
Configure Application Profile.....	367
Configure End Point Groups.....	367
EPG for Web	367
EPG for App	370
Verify Virtual Networking for the Application EPGs.....	373
Configure Contracts.....	374
App-Tier to Web-Tier Contract	374
Web-Tier to Shared L3Out Contract	383
Solution Validation	384
Validated Hardware and Software.....	384
Interoperability	385
Solution Validation	385
Summary	386
References	387
Cisco HyperFlex	387
Cisco UCS	387
Cisco ACI Application Centric Infrastructure (ACI).....	388
Cisco AVE	388
Security	388
Interoperability Matrixes.....	389
About the Author	390
Acknowledgements	390

Executive Summary

Cisco Validated Designs (CVDs) are systems and solutions that are designed, tested, and documented to facilitate and accelerate customer deployments. CVDs incorporate a wide range of technologies, products and best-practices into a portfolio of solutions that address the business needs of our customers.

Cisco Validated Designs based on Cisco HyperFlex deliver a foundational architecture for hyperconverged Virtual Server Infrastructure (VSI). HyperFlex infrastructure, when connected to Cisco Application Centric Infrastructure (ACI) extends the software-defined paradigm into the data center network, to deliver a scalable, application-centric, policy-based infrastructure for Enterprise data centers. For a complete portfolio of HyperFlex and HyperFlex VSI solutions, see: <https://www.cisco.com/c/en/us/solutions/design-zone/data-center-design-guides/data-center-hyperconverged-infrastructure.html>

The **Cisco HyperFlex Stretched Cluster with Cisco ACI Multi-Pod Fabric** solution covered in this document, is a validated reference architecture for building active-active data centers to provide business continuity and disaster avoidance. The solution uses a Cisco HyperFlex *stretched* cluster for the hyperconverged infrastructure in the active-active data centers, and a Cisco ACI Multi-Pod fabric for the data center fabric and for connectivity between the data centers. The data centers can be in geographically separate sites such as a metropolitan area or they can be in the same campus or building. The solution also includes an optional Management cluster to host shared services that multiple tenants require. The Management cluster uses a *standard* HyperFlex that is deployed from the cloud using Cisco Intersight. Cisco Intersight is also used to centrally manage all HyperFlex and UCS infrastructure in the Enterprise.

Cisco Intersight is a subscription-based, cloud service for infrastructure management that simplifies operations by providing pro-active, actionable intelligence for operations. Cisco Intersight provides capabilities such as Cisco Technical Assistance Center (TAC) integration for support and Cisco Hardware Compatibility List (HCL) integration for compliance that Enterprises can leverage for all their Cisco HyperFlex and UCS systems in all locations. Enterprises can also quickly adopt the new features that are being continuously rolled out in Cisco Intersight.

To enable the active-active data center, compute, storage and networking is extended between sites to provide virtual server infrastructure in both sites. In this design, the HyperFlex stretched cluster is stretched across two data centers with equal number of nodes in each site. The nodes connect to an ACI Multi-Pod fabric that provides seamless Layer 2 extension and Layer 3 forwarding between sites, enabling workloads to be placed in either site while also providing workload mobility.

To simplify the deployment of virtualized workloads, ACI integration with VMware vCenter is used in this solution to dynamically orchestrate and manage the virtual networking using either a VMware virtual Distributed Switch (vDS) or Cisco ACI Virtualization Edge (AVE) switch. Cisco AVE is a virtual Leaf that brings the advanced capabilities of an ACI fabric (for example, application policies, micro-segmentation, security) to the virtualization layer.

The **Cisco HyperFlex Stretched Cluster with Cisco ACI Multi-Pod Fabric** CVD consists of the following documents:

- Design Guide: Cisco HyperFlex 3.5 Stretched Cluster with Cisco ACI 4.0 Multi-Pod Fabric Design Guide
- Deployment Guide: Cisco HyperFlex 3.5 Stretched Cluster with Cisco ACI 4.0 Multi-Pod Fabric

This document is the **deployment** guide for the solution. The solution was validated using Cisco HyperFlex 3.5, Cisco Unified Computing System 4.0 (Cisco UCS), Cisco ACI 4.0, and VMware vSphere 6.5.

Solution Overview

Introduction

The **Cisco HyperFlex Stretched Cluster with Cisco ACI Multi-Pod Fabric** solution presented in this document, delivers a hyperconverged Virtual Server Infrastructure solution that provides business continuity and disaster avoidance by extending compute, storage and networking across two geographically dispersed sites. The solution was built and validated using Cisco HyperFlex 3.5 stretched cluster, Cisco Unified Computing System 4.0, Cisco ACI 4.0 Multi-Pod fabric running on Cisco Nexus family of switches and VMware vSphere 6.5U2.

Audience

The audience for this document includes, but is not limited to; sales engineers, field consultants, professional services, IT managers, partner engineers, and customers that are interested in leveraging industry trends towards hyperconvergence and software-defined networking to build agile infrastructures that can be deployed in minutes and keep up with business demands.

Purpose of this Document

This document provides detailed implementation steps for deploying the **Cisco HyperFlex Stretched Cluster with Cisco ACI Multi-Pod Fabric** solution for disaster avoidance using Cisco Hyperflex stretched cluster, Cisco ACI Multi-Pod fabric and VMware vSphere. The solution incorporates technology, product and design best practices to deliver an active-active data center solution. This document is the **deployment** guide for the solution.

What's New in this Release?

The Cisco HyperFlex Stretched Cluster with Cisco ACI Multi-Pod Fabric solution is part of the HyperFlex VSI portfolio of solutions. This release of the Cisco HyperFlex VSI solution delivers a reference architecture for disaster avoidance in Enterprise data centers. This solution adds Cisco HyperFlex stretched cluster and Cisco ACI Multi-Pod fabric to the portfolio of Cisco HyperFlex solutions.

The solution uses two active data centers in two locations to provide disaster avoidance. HyperFlex VSI using a stretched cluster provides the compute, storage and server networking in each location. The HyperFlex VSI connects to an ACI Multi-Pod fabric that provides the network fabric in each data center and the connectivity between the data centers.

The solution includes a standard HyperFlex cluster for management that is deployed from the cloud using Cisco Intersight. The management cluster can be used to manage any virtual server infrastructure in ACI Multi-Pod fabric. The solution also uses the following component versions to validate the design:

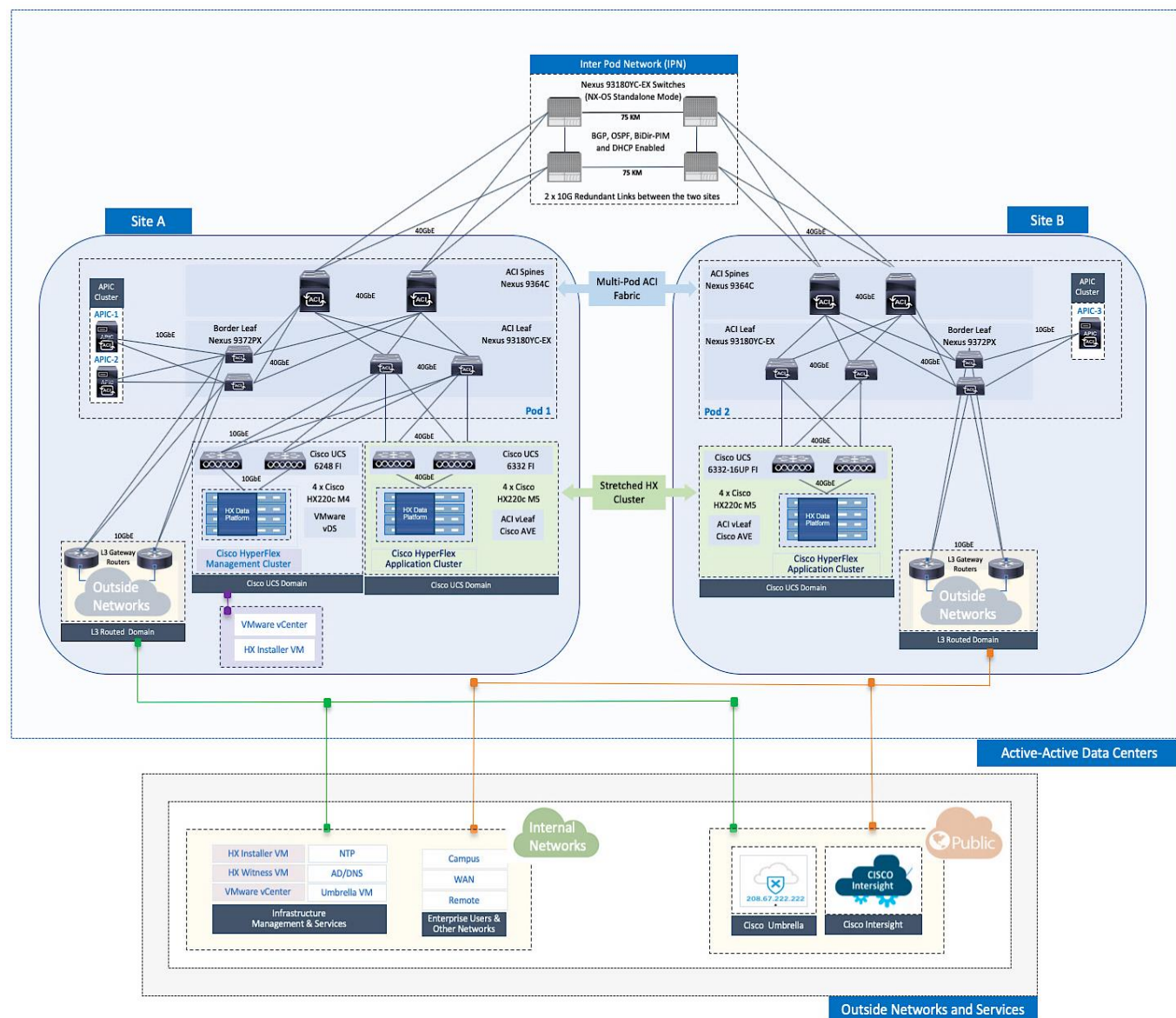
- Cisco HyperFlex 3.5(2e), Cisco UCS Manager 4.0(2d) and Cisco Intersight
- Cisco ACI 4.0(1h), Cisco AVE 2.0(1a) and VMware vDS 6.5.0
- VMware vSphere 6.5U2

Solution Summary

The **Cisco HyperFlex Stretched Cluster with Cisco ACI Multi-Pod Fabric** solution is a validated reference architecture for providing disaster avoidance and business continuity in Enterprise data centers. In the event of a disaster or data center-wide failure, the solution maintains availability of the Virtual Server Infrastructure (VSI) by using an active-active data center architecture .

The end-to-end topology of the solution is shown in [Figure 1](#).

Figure 1 Solution Topology



To provide business continuity and disaster avoidance, the solution uses a HyperFlex *stretched* cluster to extend the hyperconverged infrastructure across two geographically separate sites. Cisco ACI Multi-Pod fabric provides the networking to enable layer 2 extension and layer 3 connectivity between sites. The ACI Multi-Pod fabric consists of two fabrics in this design, one in each site, inter-connected by an **Inter-Pod Network (IPN)**. The fabric in each site is referred to as a **Pod** in the ACI Multi-Pod architecture, where each Pod is deployed as a standard Spine-Leaf architecture. The fabric is managed using a 3-node APIC cluster with two APICs in the first site and a third APIC in the second site. The physical connectivity is based on 40GbE within the Pod and 10GbE or 40GbE to connect to external (outside ACI, IPN) networks and access layer devices (APIC nodes, UCS Fabric Interconnects). A highly-resilient design is used in each Pod to ensure access to networks and services in the event of a link or node failure.

Each Pod also has direct Layer 3 connectivity to outside networks from each site. In this design, all tenants in the ACI fabric share the same Layer 3 connectivity to reach networks outside the ACI fabric. This connectivity is referred to as a **Shared L3Out** in ACI. Networks outside the ACI fabric can either be existing networks (for example, non-ACI data center, campus

or branch) within the Enterprise or networks (for example, Internet, IPN) external to the Enterprise. Shared L3Out connections are defined in either the system-defined **common** Tenant or a user-defined tenant and runs a routing protocol (or static routes) to exchange routes between the ACI and non-ACI portions of the Enterprise network. In this design, the two shared L3Out connections are in the **common** Tenant and use OSPF as the routing protocol.

The two additional tenants used in the solution are:

- **HXV-Foundation** Tenant to provide the infrastructure connectivity and services for bringing up and maintaining the HyperFlex clusters.
- **HXV-App-A** Tenant to provide the connectivity and services for the applications hosted on the HyperFlex clusters

Two types of HyperFlex clusters are used in the solution – a standard HyperFlex cluster for **Management** (optional) and a stretched cluster for **Application** workloads. Both clusters are centrally managed from Cisco Intersight, illustrating the ease and advantages of a cloud-based management tool. The Management cluster is also deployed from the cloud using Cisco Intersight.

ACI manages the virtual networking on the HyperFlex clusters by integrating with VMware vCenter that manages the clusters. Cisco APIC deploys a distributed virtual switch and creates port-groups as necessary to manage the virtual networking. In this design, an APIC-controlled VMware vDS is used in the Management cluster and Cisco AVE in the Application cluster.

Connectivity to the HyperFlex clusters are through two separate pairs of Cisco UCS 6x00 series Fabric Interconnects using multiple 40GbE links in a virtual Port-channel (vPC) configuration for higher bandwidth and resiliency. The HyperFlex nodes in each cluster connect to the Fabric Interconnects using either 10Gb or 40Gb Ethernet. A Cisco UCS domain consists of a pair of UCS Fabric Interconnects, with embedded Cisco UCS manager that manages all servers in that domain. A Cisco UCS domain can support several HyperFlex clusters depending on the port-density on the chosen Fabric Interconnect model. The different Cisco UCS domains are also managed from the cloud using Cisco Intersight. Cisco Intersight offers centralized management of Cisco UCS servers and HyperFlex nodes in all Enterprise locations with enhanced capabilities such as integration with Cisco TAC for simplified support, proactive support through actionable intelligence from telemetry data, compliance check through integration with Cisco Hardware Compatibility List (HCL) and centralized service profiles for policy-based configuration.

The solution was validated in Cisco Labs using the component models deployed in each site – see Table 1 . Other models are supported, provided the software and hardware combinations are supported in Cisco and VMware's hardware compatibility lists. See [Solution](#) Validation section for additional details.

Table 1 Solution Components per Pod

Infrastructure Domain	Component		Comments
Network (ACI MultiPod Fabric)	Pod 1	Pod 2	
	Cisco APIC M2 Server x 2	Cisco APIC M2 Server x 1	APIC Cluster (3-node)
	Cisco Nexus 9364C x 2	Cisco Nexus 9364C x 2	Spine Switches
	Cisco Nexus 93180YC-EX x 2	Cisco Nexus 93180YC-EX x 2	Leaf Switches – To Cisco UCS Domains
	Cisco Nexus 9372PX x 2	Cisco Nexus 9372PX x 2	Leaf Switches – Shared L3Out
	Cisco Nexus 93180YC-EX	Cisco Nexus 93180YC-EX x 2	IPN Routers
Hyperconverged Infrastructure (Cisco HyperFlex Standard & Stretched Clusters)	Pod 1	Pod 2	
	Cisco HX220C-M4S x 4	–	Management Cluster (4-node Standard Cluster)
	Cisco UCS 6248 FI x 2	–	
	Cisco HX220C-M5SX x 4	HX220C-M5SX x 4	Application Cluster (4-4 Stretch Cluster)
	Cisco UCS 6332 UP FI x 2	Cisco UCS 6332 UP FI x 2	
Virtualization Layer	Pod 1	Pod 2	
	VMware vSphere 6.5 U2 EP13	VMware vSphere 6.5 U2 EP13	Hypervisor
	vCenter Server Appliance 6.5 U2e	–	VCSA for Application Cluster; Management Cluster is managed by a VMware vCenter Server outside ACI Fabric
	VMware vDS, Cisco AVE	Cisco AVE	Virtual Switches – VMware vDS used in Management Cluster; Cisco AVE used in Application Cluster
Management & Monitoring	Cisco Intersight, Cisco UCS Manager, Cisco HyperFlex Connector, vCenter Plugins for HyperFlex and Cisco ACI		
Security	Cisco Umbrella (Cloud-based) using On-premise Virtual Appliances		https://umbrella.cisco.com

Solution Deployment – ACI Fabric (Single Pod)

This section provides detailed procedures for deploying a new Cisco ACI fabric. This fabric will serve as the first Pod (Pod 1 in [Figure 1](#)) in the ACI Multi-Pod fabric. The fabric will provide network connectivity for Cisco UCS domains and Cisco HyperFlex clusters that connect to it. In this solution, half of the stretched cluster nodes and the optional Management cluster will connect to this Pod.



The procedures outlined in this section are the same as that for deploying a single ACI fabric.

Deployment Overview

A high-level overview of the steps involved in deploying a single-site ACI fabric is summarized below:

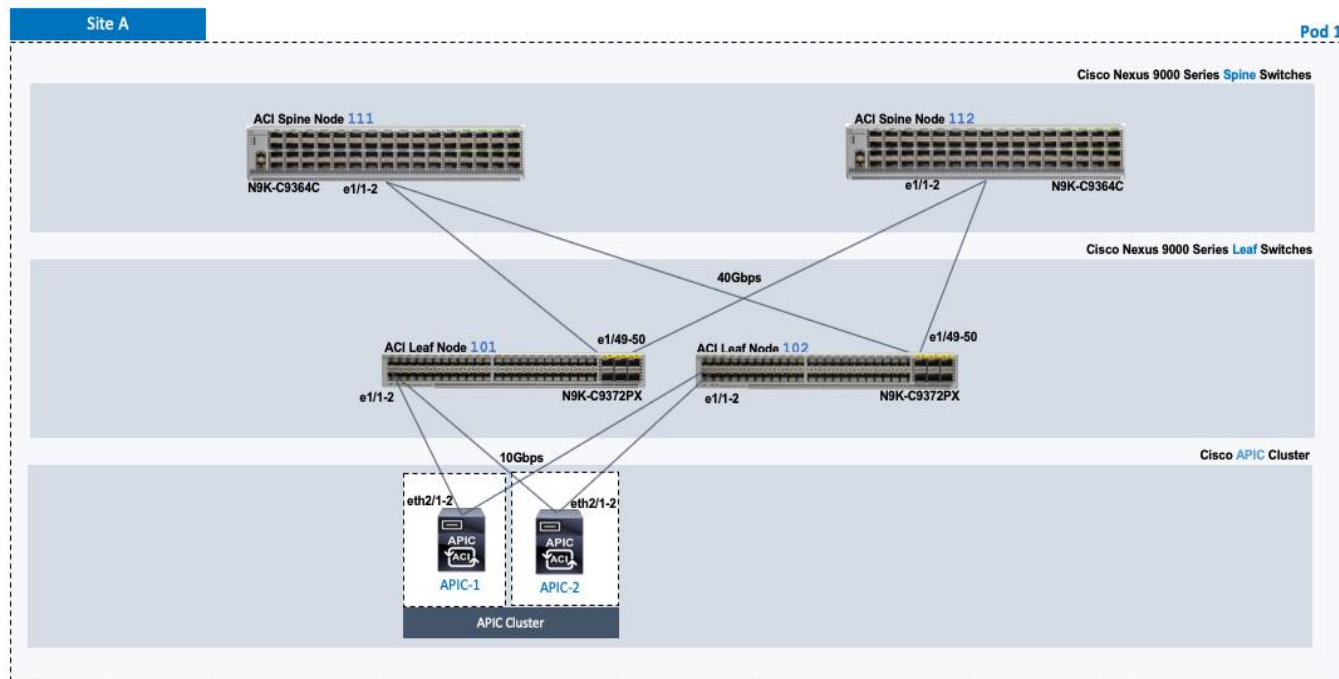
- **Physical Connectivity** - complete the cabling required to connect the devices in Pod-1. An ACI fabric should have a minimum of two Spine switches, two Leaf switches, and three APICs in a cluster. The APICs in an ACI Multi-Pod fabric should be distributed across the different Pods for redundancy. In this design, two APICs are deployed in Pod-1 and one in Pod-2. The Pod-2 APIC will be deployed at a later stage, when Pod-2 fabric is setup. Management connectivity through the Cisco Integrated Management Controller (CIMC) port on the APIC(s) should also be in place. Initial setup of an APIC requires access to the keyboard, video and mouse (KVM) console and this console is accessible through CIMC. Lastly, out-of-band management connectivity to the switches and APICs should also be in place.
- **Initial Setup of APIC(s) in Pod-1** – complete the initial configuration to bring at least one APIC online. In Cisco ACI, all configuration is centralized and managed from the APIC. The Spine and Leaf switches in the fabric are not individually configured – they are configured from the APIC. APIC uses Link Layer Discovery Protocol (LLDP) to discover ACI capable Nexus 9000 series switches in the infrastructure (and other APICs) in the fabric. The newly discovered switches are then added, provisioned and managed from the APIC web GUI.
- **Deploy Spine and Leaf switches in Pod-1.** APICs are connected to Leaf switches. In this design, the APICs are connected to border leaf switches that provide connectivity to networks outside the ACI fabric. APIC(s) can be connected to any leaf switch pair in an ACI fabric. APICs discover other switches in the fabric through LLDP. The discovered switches can then be added to the ACI fabric. The APIC can now manage the switches.
- **Configure Timezone, NTP, BGP Route Reflector function, Fabric Profiles and Access Policies for Pod-1.**
- **Enable/Review ACI Fabric Settings.** These settings include both fabric-wide and bridge-domain specific settings that impact the flow of traffic between endpoints. They are relevant and important to all endpoints in the ACI Multi-Pod fabric.
- **Pre-configure Access Policies for the ACI fabric.** These policies will be used to configure access layer connectivity to endpoints, gateways and other devices connected to the fabric. The policies are used fabric-wide, by all Pods in the ACI Multi-Pod fabric.

Deploy ACI Fabric in Pod-1

Physical Connectivity

Complete the cabling required to deploy an ACI Fabric as shown in [Figure 2](#). Out-of-Band (OOB) management connectivity for all devices and CIMC management for the APICs (not shown below) should also be completed.

Figure 2 Physical Connectivity Details for Pod-1



Initial Setup of APIC(s) in Pod-1

Follow the procedures outlined in this section to do an initial setup and configuration of APIC(s) in Pod-1 that will manage the ACI fabric. In this design, a 3-node APIC cluster is deployed, with two APICs deployed in Pod-1 and a third APIC in Pod-2.

Prerequisites

KVM Console access is necessary to do an initial setup and configuration of new APIC(s). KVM access is available through CIMC Management and therefore access to CIMC Management interface on the APIC server is required.

Setup Information

The initial setup of APICs in Pod-1 requires the information provided in this section.

- CIMC Management IP Address for the APIC(s) being setup
- CIMC log in credentials for the APIC(s) being setup



TEP Address Pool is the APIC TEP pool and should be the same for all APICs in a cluster regardless of their location.

Table 2 Setup Parameters for APICs in Pod-1

APIC	Parameters	Notes	Default Values
Fabric Name	ACI Fabric West		ACI Fabric1
Fabric ID	2	Range: (1-128)	1
Number of Active Controllers	3	Range: (1-9) Minimum # of controllers recommended: 3	3
POD ID	1	Range: (1-254)	1
Standby Controller ?	NO		NO
APIC-X ?	NO		NO
Controller ID(s)	1 2	Range: (1-3) APIC with ID=1 is the 1st controller in the cluster	1
Controller Name(s)	AA11-APIC-M2-WEST-1 AA11-APIC-M2-WEST-2		apic1
TEP Address Pool	10.13.0.0/16	APIC TEP Pool is different from the TEP Pool used by switches; Same pool is used by all APICs in a fabric, including APICs in Pod-2	10.0.0.0/16
Infrastructure VLAN ID	4093	Range: (1-4094)	4093
BD Multicast Address (GIPO)	226.0.0.0/15	GIPO is configured during first APIC setup in Pod-1; Remaining controllers will use this	225.0.0.0/15
OOB Management IP Addresses	172.26.163.121/24 172.26.163.122/24		–
OOB Management Gateway	172.26.163.254		–
OOB Management Speed/Duplex	auto		–
Admin User Password	*****	Password is configured during first APIC setup in Pod-1; Remaining controllers and switches will sync to this	–

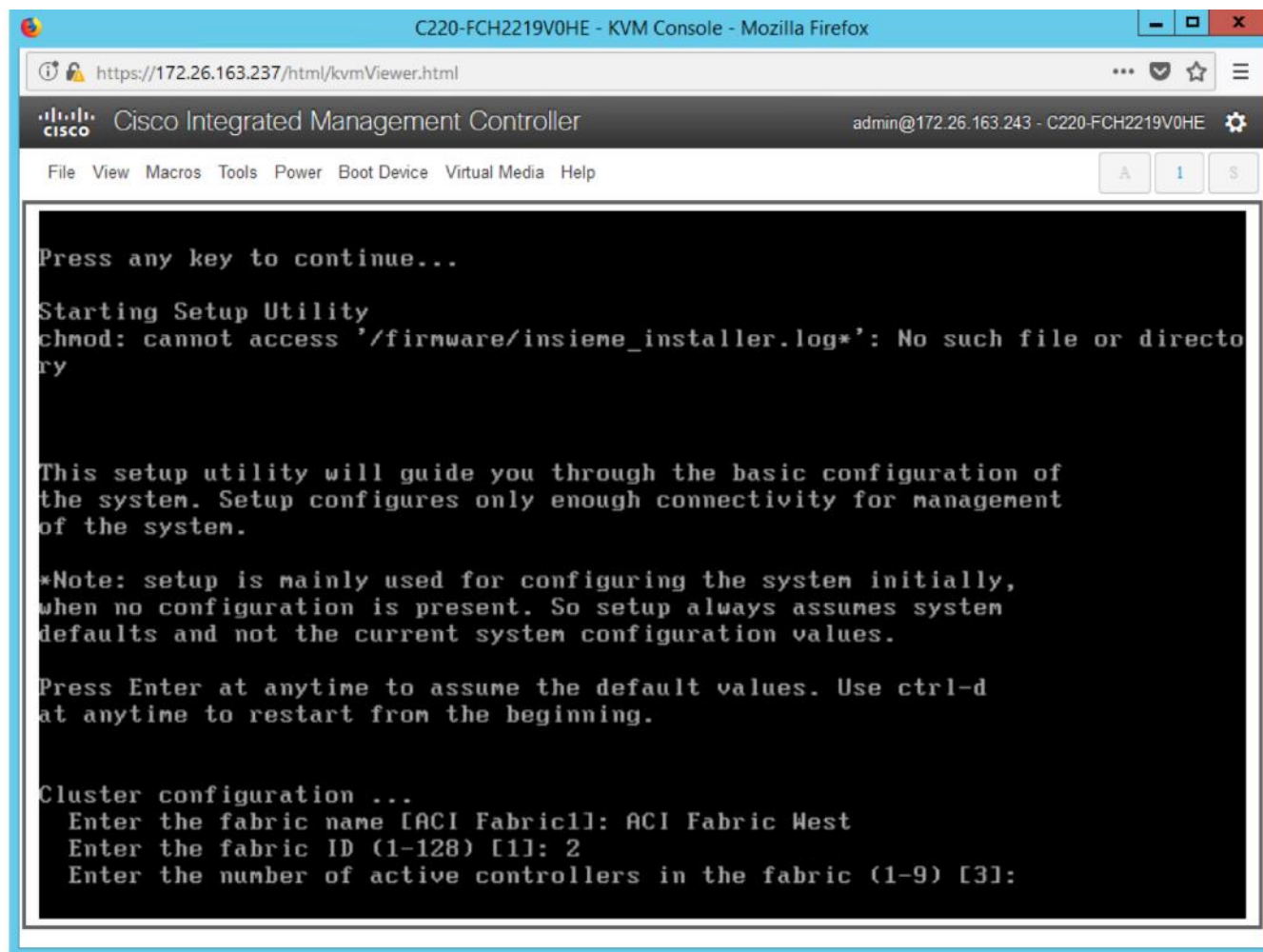
Deployment Steps

To setup new APICs in Pod-1, follow these steps:

1. Use a browser to navigate to the CIMC IP address of the new APIC. Log in using **admin** account.
2. From the top menu, click **Launch KVM**. Select **HTML based KVM** from the drop-down list.
3. When the KVM Application launches, the initial APIC setup screen should be visible. Press any key to start the **Setup Utility**.



If the APIC was previously configured, reset to factory defaults and wipe it clean before proceeding.



4. Use the Setup information provided above to step through the initial APIC configuration as shown below.

```

Cisco Integrated Management Controller
admin@172.26.163.243 - C220-FCH2219V0HE

File View Macros Tools Power Boot Device Virtual Media Help

Press Enter at anytime to assume the default values. Use ctrl-d
at anytime to restart from the beginning.

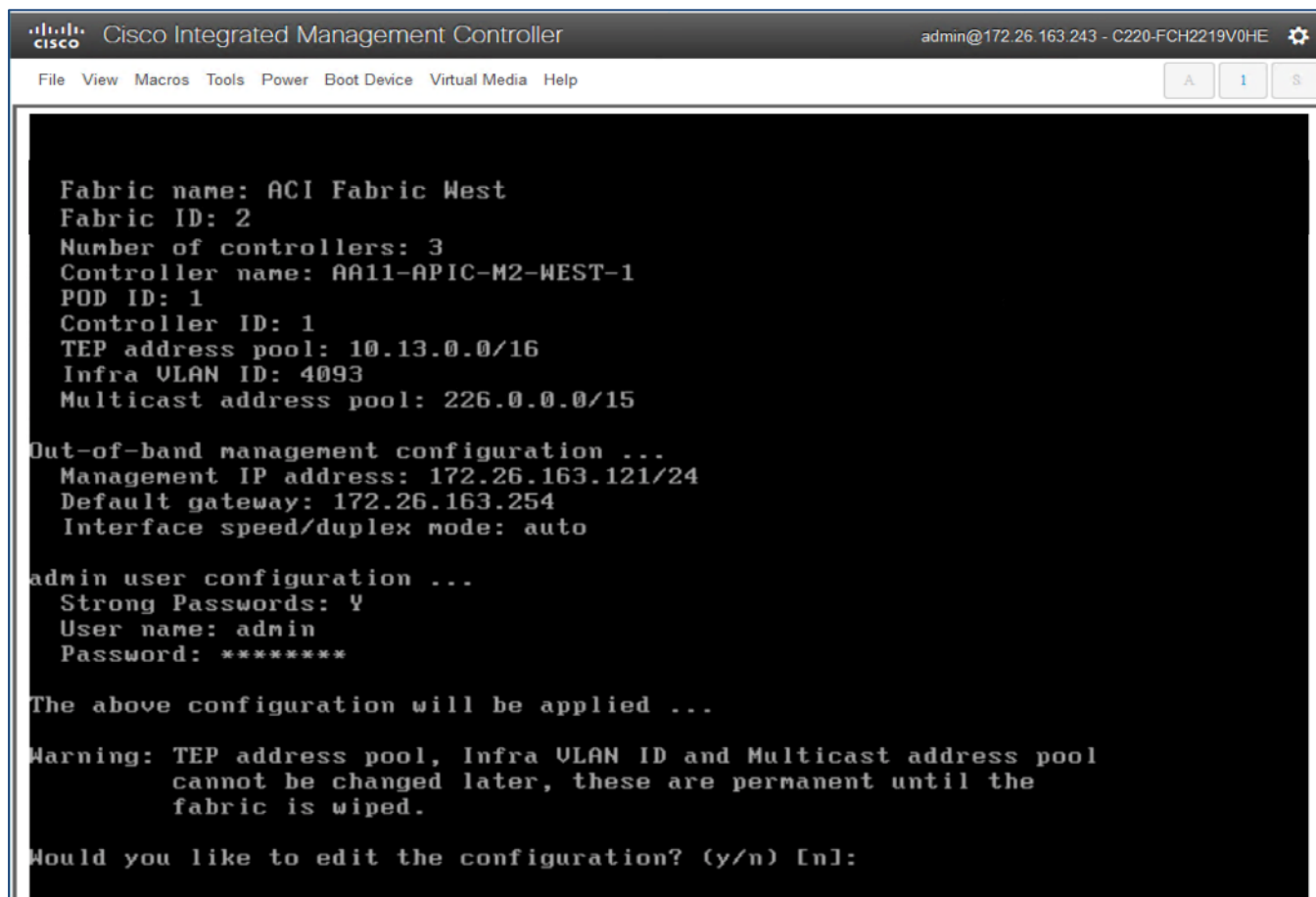
Cluster configuration ...
  Enter the fabric name [ACI Fabric1]: ACI Fabric West
  Enter the fabric ID (1-128) [1]: 2
  Enter the number of active controllers in the fabric (1-9) [3]:
  Enter the POD ID (1-254) [1]:
  Is this a standby controller? [NO]:
  Is this an APIC-X? [NO]:
  Enter the controller ID (1-3) [1]:
  Enter the controller name [apic1]: AA11-APIC-M2-WEST-1
  Enter address pool for TEP addresses [10.0.0.0/16]: 10.13.0.0/16
  Note: The infra VLAN ID should not be used elsewhere in your environment
        and should not overlap with any other reserved VLANs on other platforms.
  Enter the VLAN ID for infra network (1-4094): 4093
  Enter address pool for BD multicast addresses (GIP0) [225.0.0.0/15]: 226.0.0.0
/15

Out-of-band management configuration ...
  Enable IPv6 for Out of Band Mgmt Interface? [N]:
  Enter the IPv4 address [192.168.10.1/24]: 172.26.163.121/24
  Enter the IPv4 address of the default gateway [None]: 172.26.163.254
  Enter the interface speed/duplex mode [auto]: _

admin user configuration ...
  Enable strong passwords? [Y]:
  Enter the password for admin:

```

5. Press **Enter** after the last question (password for admin).



```

Cisco Integrated Management Controller
admin@172.26.163.243 - C220-FCH2219V0HE

File View Macros Tools Power Boot Device Virtual Media Help

Fabric name: ACI Fabric West
Fabric ID: 2
Number of controllers: 3
Controller name: AA11-APIC-M2-WEST-1
POD ID: 1
Controller ID: 1
TEP address pool: 10.13.0.0/16
Infra VLAN ID: 4093
Multicast address pool: 226.0.0.0/15

Out-of-band management configuration ...
Management IP address: 172.26.163.121/24
Default gateway: 172.26.163.254
Interface speed/duplex mode: auto

admin user configuration ...
Strong Passwords: Y
User name: admin
Password: *****

The above configuration will be applied ...

Warning: TEP address pool, Infra VLAN ID and Multicast address pool
cannot be changed later, these are permanent until the
fabric is wiped.

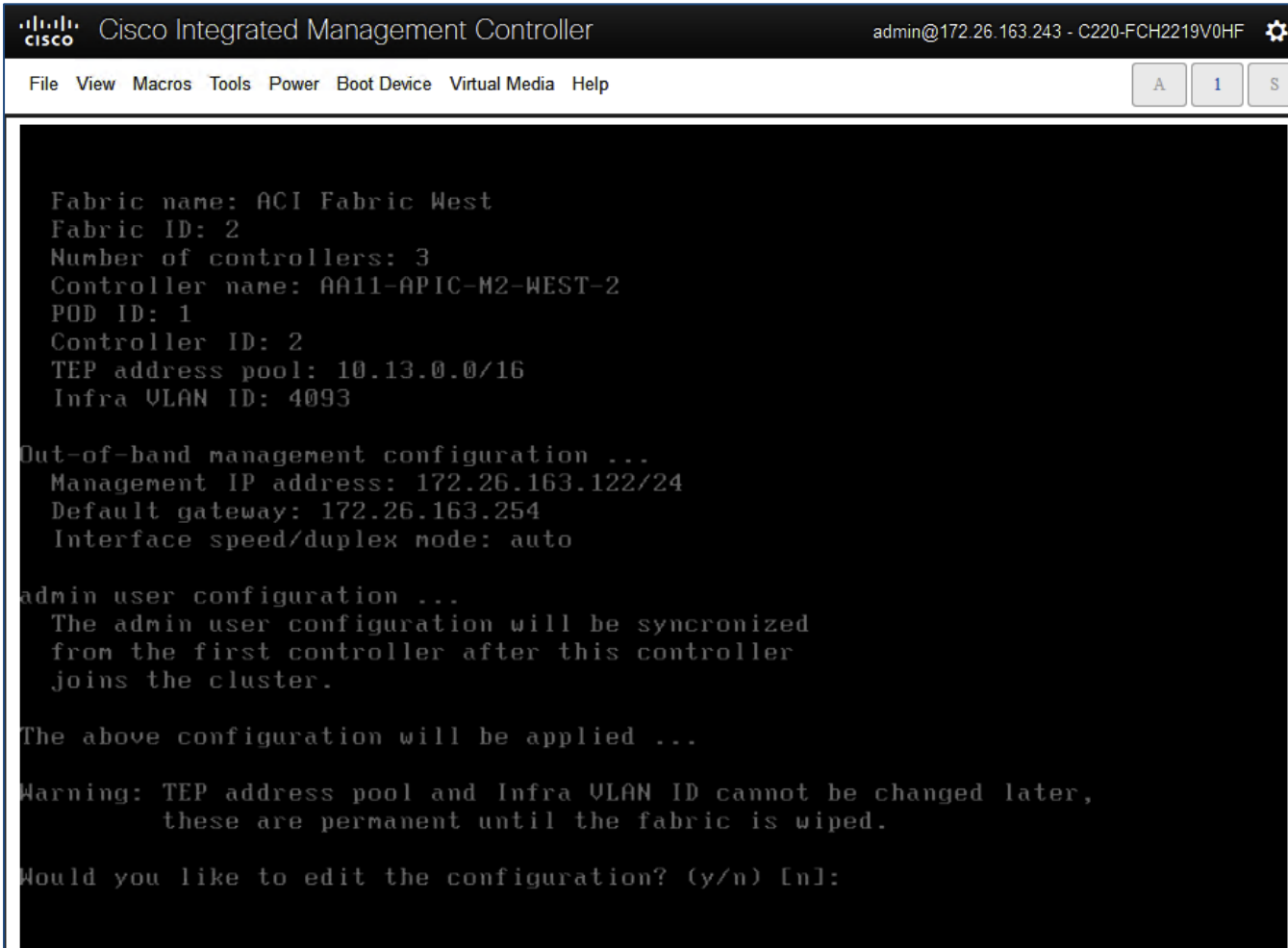
Would you like to edit the configuration? (y/n) [n]:

```

6. Review the configured information. Click **y** if necessary to go back and make changes, otherwise press **Enter** to accept the configuration.
7. Repeat steps 1-6 for the next APIC in Pod-1.



APIC username, password, and BD Multicast Address (GIPO) is configured only once, during the initial setup of APIC-1 or the first controller in the cluster. The remaining controllers and switches sync to the configuration on APIC-1.



```

Cisco Integrated Management Controller
admin@172.26.163.243 - C220-FCH2219V0HF

File View Macros Tools Power Boot Device Virtual Media Help

Fabric name: ACI Fabric West
Fabric ID: 2
Number of controllers: 3
Controller name: AA11-APIC-M2-WEST-2
POD ID: 1
Controller ID: 2
TEP address pool: 10.13.0.0/16
Infra VLAN ID: 4093

Out-of-band management configuration ...
Management IP address: 172.26.163.122/24
Default gateway: 172.26.163.254
Interface speed/duplex mode: auto

admin user configuration ...
The admin user configuration will be synchronized
from the first controller after this controller
joins the cluster.

The above configuration will be applied ...

Warning: TEP address pool and Infra VLAN ID cannot be changed later,
these are permanent until the fabric is wiped.

Would you like to edit the configuration? (y/n) [n]:

```

8. Review the configured information. Click **y** if necessary to go back and make changes, otherwise press **Enter** to accept the configuration.



The third APIC in the cluster is located in Pod-2 - this APIC will be setup at a later stage, after the Multi-Pod IP connectivity (IP Network) is setup between the Pods.

9. The configuration and management of the ACI fabric can be done by navigating to the OOB Management IP address of either APIC. The configuration done from one APIC will be synced to other APICs in the cluster, ensuring a consistent view of the fabric.

Deploy Spine and Leaf Switches in Pod-1

Once an APIC is up and running in Pod-1, it will discover the connected spine and leaf switches through LLDP. Follow the procedures outlined in this section to setup and deploy spine and leaf switches in Pod-1.

Setup Information

The tables below provides the setup information for deploying Spine and Leaf switches in Pod-1.

Table 3 Leaf Switches in Pod-1

Leaf Switches in Pod-1	General	Node ID	Node Names	OOB Management EPG	OOB Management IP	OOB Gateway
	Pod ID: 1	101	AA11-9372PX-WEST-1	default	172.26.163.117/24	172.26.163.254
	Role: Leaf					
	Rack Name (Optional): AA11	102	AA11-9372PX-WEST-2	default	172.26.163.118/24	172.26.163.254

Table 4 Spine Switches in Pod-1

Spine Switches in Pod-1	General	Node ID	Node Names	OOB Management EPG	OOB Management IP	OOB Gateway
	Pod ID: 1	111	AA11-9364C-WEST-1	default	172.26.163.119/24	172.26.163.254
	Role: Spine					
	Rack Name (Optional): AA11	112	AA11-9364C-WEST-2	default	172.26.163.120/24	172.26.163.254

Add Leaf Switches to the ACI Fabric

To add discovered Leaf and Spine switches in Pod-1 to the ACI Fabric, follow these steps:

1. Use a browser to navigate to the APIC GUI using the Out-of-Band (OOB) IP address assigned to the APIC(s) in Pod-1 during the initial setup. Log in using **admin** account.
2. From the top menu, select **Fabric > Inventory**.
3. From the left navigation pane, navigate to **Fabric Membership**.
4. In the right navigation pane, go to the **Nodes Pending Registration** tab.

Serial Number	Pod ID	Node ID	RL TEP Pool	Name	Role	Supported Model	SSL Certificate	Status
SAL1940QAX	1	0	0		leaf	yes	n/a	

5. The newly discovered Leaf Switches will be listed with a Node ID of '0'. You should see at least one of the Leaf switches – the APIC is dual-homed to a pair of Leaf switches. Note that the switch's **Role** is **leaf**.
6. Use the serial numbers to identify the new Leaf switch. Collect the setup information for this switch. Proceed to the next section to configure the newly discovered Leaf switches.

7. Right-click and select **Register**.
8. In the **Register** pop-up window, specify the Pod ID (for example, 1), Node Id (for example, 101), Node Name for example, AA11-9372PX-WEST-1) and Rack Name (for example, AA11).
9. Click **Register**.
10. Click the **Registered Nodes** tab. The newly configured Leaf switch should show up as **Active** after a few minutes.

The screenshot shows the Cisco APIC interface with the **Fabric** tab selected. The **Fabric Membership** section is active, and the **Registered Nodes** tab is selected. The statistics show 1 Leaf, 0 Virtual Leaf, 0 Spine, and 0 Virtual Spine nodes. The table below lists the registered nodes:

Serial Number	Model	Pod ID	Node ID	Name	Role	IP	Status
SAL1940QAAAX	N9K-C9372PX	1	101	AA11-9372PX-WEST-1	leaf	10.13.64.64/32	Active

11. In the right navigation pane, click the **Nodes Pending Registration** tab.
12. Select the **second (-2)** Leaf switch using the serial number. Right-click and select **Register**.
13. In the **Register** pop-up window, specify the Pod ID (for example, 1), Node Id (for example, 102), Node Name for example, AA11-9372PX-WEST-2) and Rack Name (for example, AA11).

The screenshot shows the Cisco APIC interface with the **Fabric** tab selected. The **Fabric Membership** section is active, and the **Nodes Pending Registration** tab is selected. The statistics show 0 Unsupported, 0 Undiscovered, and 0 Unknown nodes. A table below lists the nodes pending registration:

Serial Number	Pod ID	Node ID	RL TEP Pool	Name	Role	Supported Model	SSL Certificate	Status
SAL1940QAEQ	1	0	0		leaf	yes	n/a	

The **Register** pop-up window is open, showing the following fields:

- Serial Number: SAL1940QAEQ
- Pod ID: 1
- Node ID: 102
- RL TEP Pool: 0
- Role: leaf
- Node Name: AA11-9372PX-WEST-2
- Rack Name: AA11 (site: fabric, building: default, fl)

The **Register** button is highlighted.

14. Click **Register**.

15. You should now see the Leaf switches under the **Registered Nodes** tab.
16. Repeat steps 1-14 to add additional leaf switch pairs to the fabric.

Upgrade Firmware on Leaf Switches in Pod-1 (Optional)

To upgrade the firmware on leaf switches in Pod-1, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using **admin** account.
2. From the top menu, navigate to **Admin > Firmware**.
3. Select the tabs for **Infrastructure > Nodes**.
4. Check the **Current Firmware** version column for the newly deployed Leaf switches to verify they are compatible with the APIC version running.
5. If an upgrade is **not** required, proceed to the next section but if an upgrade is required, use the product documentation to upgrade the switches.

Add Spine Switches to the ACI Fabric

To add spine switches to the ACI fabric, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using **admin** account.
2. From the top menu, select **Fabric > Inventory**.
3. From the left navigation pane, navigate to **Fabric Membership**.
4. In the right navigation pane, go to the **Nodes Pending Registration** tab.
5. The newly discovered spine switches will be listed with a **Node ID** of 'o', with **Role** as **spine**.
6. Use the serial numbers to identify the spine switch pair. Collect the information for each switch.
7. Select the **first** (-1) spine switch using the serial number. Right-click and select **Register**.

The screenshot shows the Cisco APIC GUI with the 'Fabric' tab selected in the top navigation bar. The left sidebar shows the 'Inventory' section with 'Fabric Membership' selected. The main content area displays the 'Fabric Membership' page with the 'Nodes Pending Registration' tab active. The page shows three categories: 'Unsupported' (0), 'Undiscovered' (0), and 'Unknown' (0). Below these, a table lists the nodes pending registration:

Serial Number	Pod ID	Node ID	RL TEP Pool	Name	Role	Supported Model	SSL Certificate	Status
FDO22240VHM	1	o	0		spine	yes	n/a	
FDO22240VJ8	1	o	0		spine	yes	n/a	

A context menu is open for the first row (FDO22240VHM), showing options: 'Register', 'Edit Node and Rack Names', and 'Remove From Controller'. The 'Register' option is highlighted.

8. In the **Register** pop-up window, specify the Pod ID (for example, 1), Node Id (for example, 111), Node Name (for example, AA11-9364C-WEST-1) and Rack Name (for example, AA11).

The screenshot shows the Cisco APIC interface with the 'Fabric' tab selected. The 'Fabric Membership' page displays three categories: Registered Nodes (0), Nodes Pending Registration (0), and Unreachable Nodes (0). A 'Register' pop-up window is open, showing the following fields:

- Serial Number: FDO22240VHM
- Pod ID: 1
- Node ID: 111
- RL TEP Pool: 0
- Role: spine
- Node Name: AA11-9364C-WEST-1
- Rack Name: AA11 (site:fabric, building:default, fl)

The 'Register' button is highlighted in blue.

9. Click **Register**.
10. Select the **second** (-2) spine switch using the serial number. Right-click and select **Register**.
11. In the **Register** pop-up window, specify the Pod ID (for example, 1), Node Id (for example, 112), Node Name (for example, AA11-9364C-WEST-2) and Rack Name (for example, AA11).

The screenshot shows the Cisco APIC interface with the 'Fabric' tab selected. The 'Fabric Membership' page displays three categories: Unsupported (0), Undiscovered (0), and Unknown (0). A 'Register' pop-up window is open, showing the following fields:

- Serial Number: FDO22240VJ8
- Pod ID: 1
- Node ID: 112
- RL TEP Pool: 0
- Role: spine
- Node Name: AA11-9364C-WEST-2
- Rack Name: AA11 (site:fabric, building:default, fl)

The 'Register' button is highlighted in blue.

12. Click **Register**.
13. Repeat steps 1-12 to add additional spine switch pairs to the fabric.

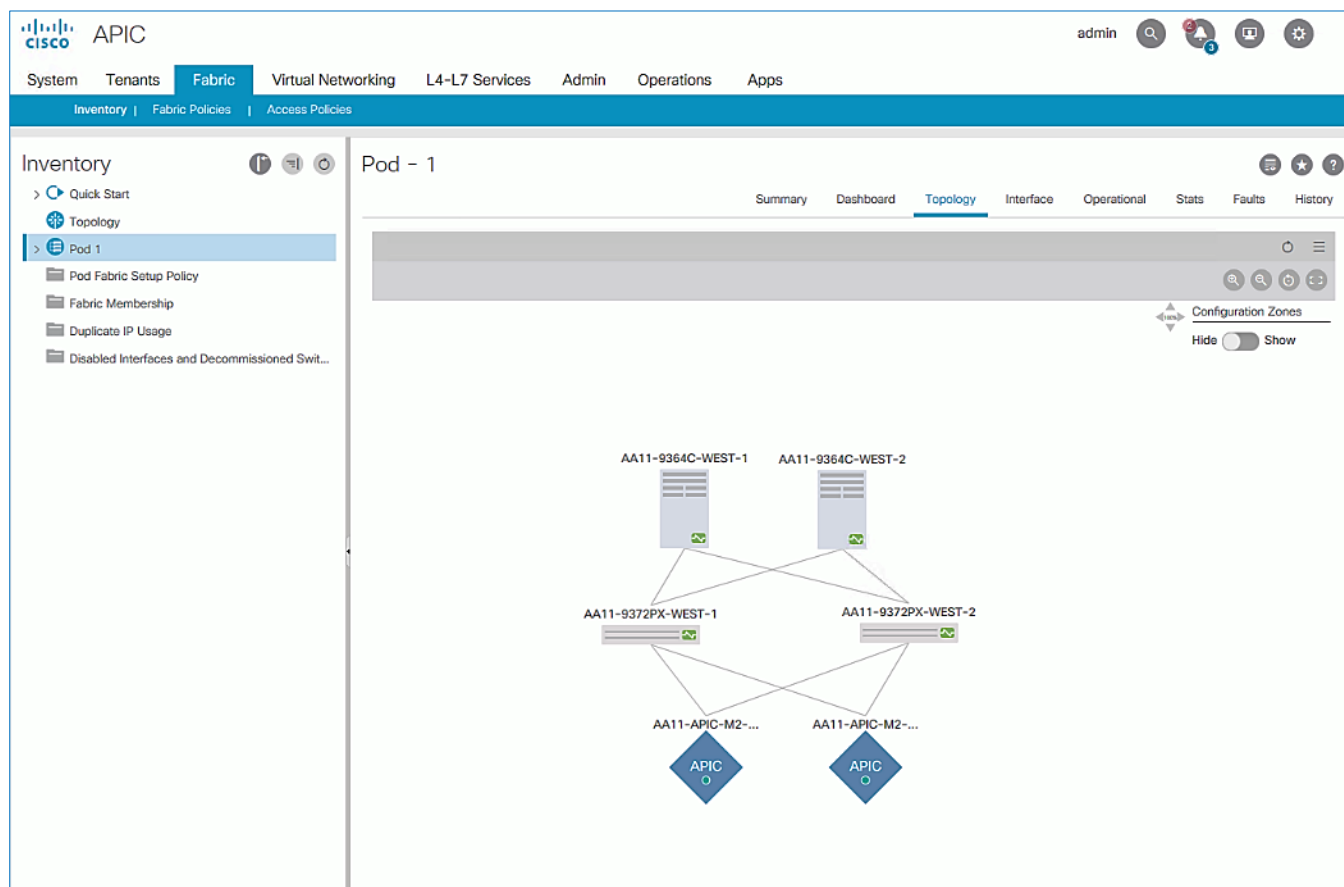
Verify Spine and Leaf Switches are Added to the ACI Fabric

To verify that the spine and leaf switches have been added to the ACI fabric, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using **admin** account.
2. From the top menu, select **Fabric > Inventory**.
3. From the left navigation pane, navigate to **Fabric Membership**.
4. In the right navigation pane, go to the **Registered Nodes** tab.

Serial Number	Model	Pod ID	Node ID	Name	Role	IP	Status
SAL1940QAAX	N9K-C937...	1	101	AA11-9372PX-WEST-1	leaf	10.13.64.64/32	Active
SAL1940QAEG	N9K-C937...	1	102	AA11-9372PX-WEST-2	leaf	10.13.184.66/32	Active
FDO22240VHM	N9K-C936...	1	111	AA11-9364C-WEST-1	spine	10.13.184.64/32	Active
FDO22240VJ8	N9K-C936...	1	112	AA11-9364C-WEST-2	spine	10.13.184.65/32	Active

5. All Spine and Leaf switches are configured and added to the fabric. Note that the APIC has allocated IP addresses from the TEP Pool for Pod-1.
6. From the left navigation pane, select **Topology** to view the fabric topology after all devices have been added to the fabric.



Upgrade Firmware on Spine Switches in Pod-1 (Optional)

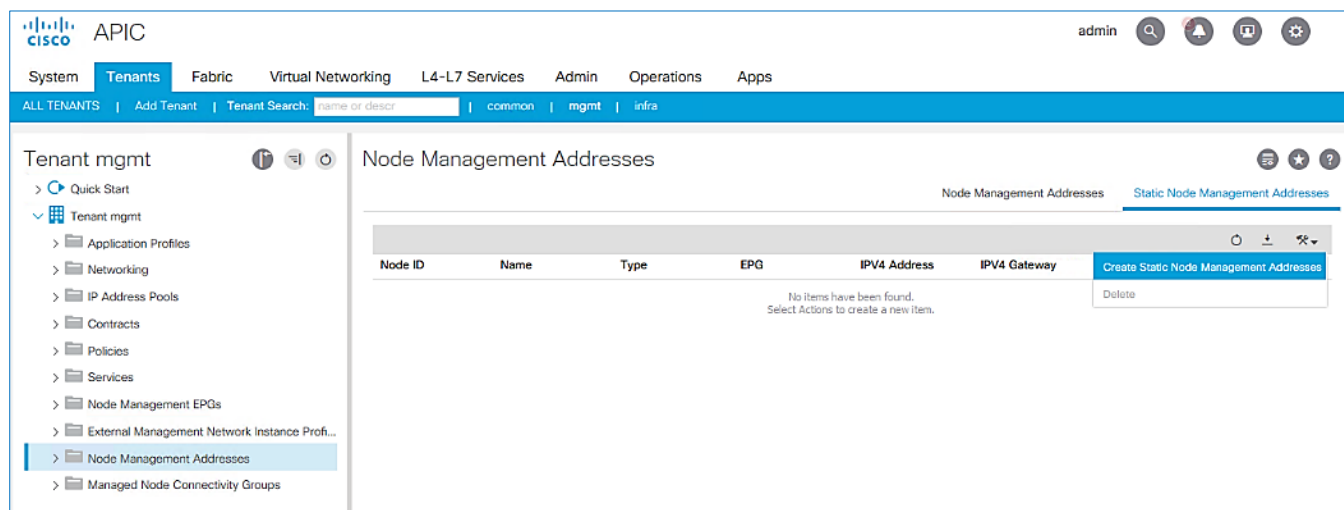
To upgrade the firmware on the spine switches in Pod-1, follow these steps:

1. From the top menu, navigate to **Admin > Firmware**.
2. Select the tabs for **Infrastructure > Nodes**.
3. Check the **Current Firmware** version column for the newly deployed Spine switches to verify they are compatible with the APIC version running.
4. If an upgrade is **not** required, proceed to the next section but if an upgrade is required, use the product documentation to upgrade the switches.

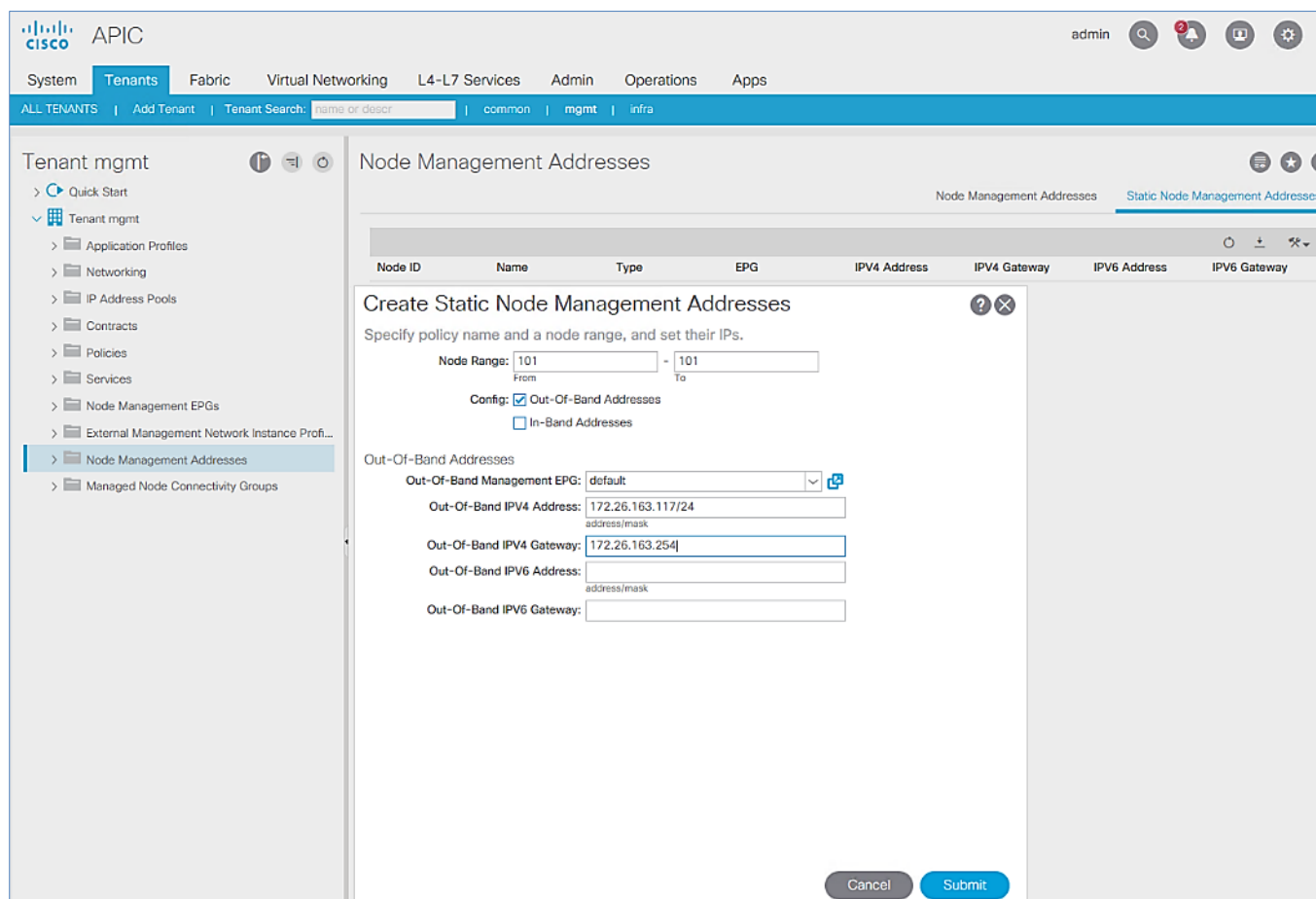
Configure Out-of-Band Management for Switches in Pod-1

To configure Out-of-Band (OOB) Management for Pod-1 Spine and Leaf switches, follow these steps using the setup information provided in Table 3 and Table 4 :

1. Use a browser to navigate to the APIC GUI. Log in using **admin** account.
2. From the top menu, select **Tenants > mgmt**.
3. From the left navigation pane, expand and select **Tenant mgmt > Node Management Addresses**.
4. In the right window pane, select the tab for **Static Node Management Addresses**.
5. Click the arrow next to the Tools icon and select **Create Static Node Management Addresses**.



6. In the Create Static Node Management Addresses pop-up window, specify a **Node Range** (for example, 101), for **Config**: select the box for **Out-of-Band Addresses**.
7. For **Out-of-Band Management EPG**, select **default** from the drop-down list.
8. Specify the **Out-of-Band Management IPv4 Address** for the first node in the specified range.
9. Specify the Out-of-Band Management IPv4 Gateway.



10. Click **Submit** to complete.
11. Click **Yes** in the **Confirm** pop-up window to assign the IP address to the range of nodes specified.
12. Repeat steps 1-11 for the remaining Spine and Leaf switches in Pod-1.

The screenshot shows the Cisco APIC web interface. The top navigation bar includes 'System', 'Tenants', 'Fabric', 'Virtual Networking', 'L4-L7 Services', 'Admin', 'Operations', and 'Apps'. The 'Tenants' tab is selected, and the 'Tenant Search' field is empty. The left sidebar shows the 'Tenant mgmt' menu with 'Static Node Management Addresses' selected. The main content area displays a table of static node management addresses.

Node ID	Name	Type	EPG	IPv4 Address	IPv4 Gateway	IPv6 Address	IPv6 Gateway
pod-1/node-101	AA11-9372PX-WEST-1	Out-Of-Band	default	172.26.163.117...	172.26.163.254	::	::
pod-1/node-102	AA11-9372PX-WEST-2	Out-Of-Band	default	172.26.163.118...	172.26.163.254	::	::
pod-1/node-111	AA11-9364C-WEST-1	Out-Of-Band	default	172.26.163.119...	172.26.163.254	::	::
pod-1/node-112	AA11-9364C-WEST-2	Out-Of-Band	default	172.26.163.120...	172.26.163.254	::	::

13. The switches can now be accessed directly using SSH.
14. To limit access to the ACI Out-of-Band Management IP Addresses, deploy contracts – use the APIC Configuration Guide for the specific steps to enable this. Contracts were not deployed in this setup. You may also need to re-add the APIC Out-of-Band Management IP addresses under **Node Management Addresses** though it was configured during the initial setup of the APIC. Node IDs for the APICs can start from '1' or some other value that is not in the same range as the Node IDs for Spine and Leaf switches.

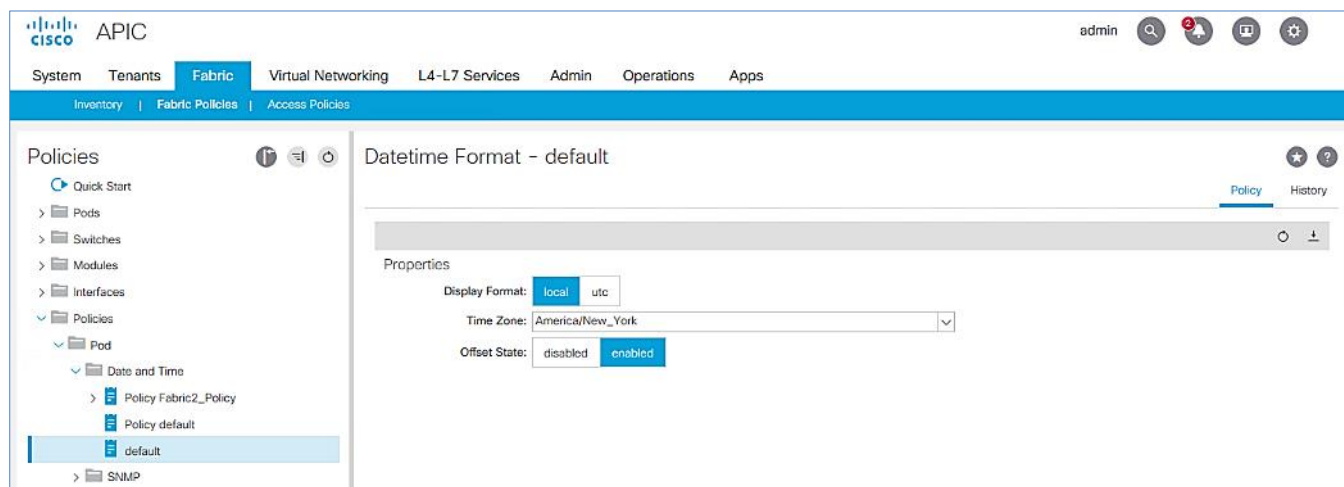
Apply Pod Policies for Pod-1

To apply policies specific to a Pod, complete the procedures outlined in this section.

Configure Time Zone (Fabric Wide Setting)

To configure Time Zone for the ACI fabric, follow these steps:

1. Use a browser to navigate to APIC's Web GUI. Log in using the **admin** account.
2. From the top menu, select **Fabric > Fabric Policies**.
3. In the left navigation pane, expand Policies and select **Policies > Pod > Date and Time > default**.
4. In the right window pane, select the **Time Zone** and verify that **Offset State** is **Enabled**.



5. Click Submit.

Configure NTP for Pod-1

To configure NTP for Pod-1, follow these steps using the setup information provided below:

- **NTP Policy Name:** Pod1-West-NTP_Policy
 - **NTP Server:** 172.26.163.254
 - **Management EPG:** default(Out-of-Band)
1. Use a browser to navigate to the APIC GUI. Log in using **admin** account.
 2. From the top menu, select **Fabric > Fabric Policies**.
 3. From the left navigation pane, navigate to **Policies > Pod > Date and Time**.
 4. Right-click and select Create Date and Time Policy.
 5. In the **Create Date and Time Policy** pop-up window, specify a **Name** for Pod-1's NTP Policy. Verify that the **Administrative State** is **enabled**.

APIC admin

System Tenants **Fabric**

Inventory | **Fabric Policies**

Policies

- Quick Start
- Pods
 - Policy Groups
 - Profiles
- Switches
- Modules
- Interfaces
- Policies
 - Pod
 - Date and Time**
 - SNMP
 - Management Access
 - ISIS Policy default
 - Switch
 - Interface
 - Global
 - Monitoring

Create Date And Time Policy

STEP 1 > Identity

Specify the information about the Date/Time Policy

Name:

Description:

Administrative State: ☐ disabled ☒ enabled

Server State: ☒ disabled ☐ enabled

Authentication State: ☒ disabled ☐ enabled

Previous Cancel Next

- Click **Next**.
- In **Step 2 > NTP Servers**, add NTP server(s) for Pod-1 using the **[+]** to the right of the list of servers.
- In the **Create Providers** pop-up window, specify the Hostname/IP of the NTP server in the **Name** field. If multiple NTP Providers are being created for Pod-1, select the checkbox for **Preferred** when creating the preferred provider. For the **Management EPG**, select **default (Out-of-Band)** from the drop-down list.

APIC admin

System Tenants **Fabric** Virtual Networking L4-L7 Services Admin Operations Apps

Inventory | **Fabric Policies** | Access Policies

Policies

- Quick Start
- Pods
 - Policy Groups
 - Profiles
- Switches
- Modules
- Interfaces
- Policies
 - Pod
 - Date and Time
 - Policy Fabric2_Po
 - Policy Pod1-West**
 - Policy default
 - default
 - SNMP
 - Management Access
 - ISIS Policy default

Create Providers

Specify the information about the NTP Server

Name:

Description:

Preferred: ☒

Minimum Polling Interval:

Maximum Polling Interval:

Management EPG:

Cancel Submit

Show Usage Reset Submit

- Click **OK**.

- Click **Finish**.



The NTP policy is not in effect until it is applied using a Pod Profile – this is covered in an upcoming section.

Update BGP Route Reflector Policy for Pod-1

In an ACI fabric with multiple Spine switches, a pair of Spine switches are selected as Route Reflectors (RR) to redistribute routes from external domains into the fabric. In a Multi-Pod ACI fabric, each Pod has a pair of RR nodes. The procedures in this section will enable RR functionality on Pod-1 Spine switches by updating an existing policy.

Setup Information

- BGP Route-Reflector Policy Name: `default`
- Pod-1 Spine ID(s):** AA11-9364C-WEST-1, AA11-9364C-WEST-2

Deployment Steps

To enable BGP Route Reflector functionality on Spine switches in Pod-1, follow these steps:

- Use a browser to navigate to the APIC GUI. Log in using **admin** account.
- From the top menu, select **System > System Settings**.
- From the left navigation pane, navigate to **BGP Route Reflector**.
- In the right window pane, in the **Route Reflector Nodes** section, click the **[+]** on the right to **Create Route Reflector Node**.
- In the **Create Route Reflector Node** pop-up window, for **Spine Node**, specify the **Node Name** for the first Spine in Pod-1.

The screenshot displays the Cisco APIC GUI. The top navigation bar includes 'System', 'Tenants', 'Fabric', 'Virtual Networking', 'L4-L7 Services', 'Admin', 'Operations', and 'Apps'. The left sidebar shows 'System Settings' with a tree view where 'BGP Route Reflector' is selected. The main content area is titled 'BGP Route Reflector Policy - BGP Route Reflector'. It features a 'Properties' section with fields for 'Name' (default), 'Description' (optional), and 'Autonomous System Number' (201). Below this is a table for 'Route Reflector Nodes' which is currently empty. A 'Create Route Reflector Node' pop-up window is open, prompting the user to 'Specify route reflector node EP id'. In this window, the 'Spine Node' is set to 'AA11-9364C-WEST-1' and the 'Description' is 'Spine-1'. The pop-up window has 'Cancel' and 'Submit' buttons at the bottom.

- Click **Submit**.

7. Repeat steps 1-6 to add a second Spine in Pod-1.
8. You should now see two Spines as Route Reflectors for each Pod in the deployment.

The screenshot shows the Cisco APIC System Settings page. The left sidebar lists various settings, with 'BGP Route Reflector' selected. The main panel displays the 'BGP Route Reflector Policy - BGP Route Reflector' configuration. The 'Properties' section shows the Name as 'default' and the Description as 'optional'. The 'Autonomous System Number' is set to 201. The 'Route Reflector Nodes' table lists two nodes: Pod ID 1, Node ID 111, Node Name AA11-9364C-W..., and Description Spine-1; and Pod ID 1, Node ID 112, Node Name AA11-9364C-W..., and Description Spine-2. The 'External Route Reflector Nodes' section is empty, with a message stating 'No items have been found. Select Actions to create a new item.' The bottom right of the panel has buttons for 'Show Usage', 'Reset', and 'Submit'.

Update Pod Profile to Apply Pod Policies

In ACI, Pod Policies (for example, BGP Route Reflector policy from previous section) are applied through a Pod Profile. A separate Pod Policy Group is used to group policies for each Pod and then they are applied using the Pod Profile. In this design, different NTP servers are used in each Pod. This policy is applied to Pod-1 policy group and then applied to the Pod Profile. A single Pod Profile is used to apply Pod policies for both Pod-1 and Pod-2. This section explains how to apply Pod Policies to Pod-1.

Setup Information

- Pod Policy Group Name for Pod-1: Pod1-West_PPG
- Pod Selector Name for Pod-1: Pod1-West
- Pod Profile: default
- ID for Pod-1: 1
- Names of Pod specific policies to be applied: Pod1-West-NTP_Policy

Deployment Steps

To apply Fabric policies on Spine switches in Pod-1, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using **admin** account.
2. From the top menu, select **Fabric > Fabric Policies**.
3. From the left navigation pane, navigate to **Pods > Policy Groups**.
4. Right-click and select **Create Pod Policy Group**, click **[+]** on the right to create a policy group.
5. In the **Create Pod Policy Group** pop-up window, for the **Name**, specify a Pod Policy Group Name. For the **Date Time Policy**, select the previously created NTP policy for Pod-1. For the remaining policies, select or verify that the default policy is selected from the drop-down list.

6. Click **Submit**.
7. From the left navigation pane, navigate to **Pods > Profiles > Pod Profile default**.
8. In the right window pane, in the **Pod Selectors** section, click the **[+]** to add a Pod Selector.
9. In the newly created row, specify a **Name**. For **Type**, select **Range**. For **Blocks**, specify the Pod Id for Pod-1. For **Policy Group**, select the previously created Policy Group Name for Pod1.
10. Click **Submit** to apply the Fabric Policies to Pod-1.

Configure DNS (Fabric Wide Setting)

To configure Domain Name Server (DNS) for the ACI fabric, follow these steps:

1. Use a browser to navigate to APIC's Web GUI. Log in using the **admin** account.
2. From the top menu, select **Fabric > Fabric Policies**.
3. In the left navigation pane, expand Policies and select **Policies > Global > DNS Profiles > default**.

4. For the **Management EPG**, select the **default (Out-of-Band)** from the drop-down list if the DNS servers are reachable through the out of band management subnet.
5. Use the **[+]** signs to the right of **DNS Providers** and **DNS Domains** to add DNS servers and domains as needed.

Properties

Name: default

Description: optional

Management EPG: select an option

DNS Providers:

Address	Preferred
192.168.160.50	False
192.168.160.51	False
192.168.160.53	False
192.168.160.54	False

DNS Domains:

Name	Default	Description
cisco.com	False	

Buttons: Show Usage, Reset, Submit

Enable/Review ACI Fabric Settings

Customers should consider enabling the following ACI fabric settings after careful evaluation, and if it is appropriate for their environment. In some cases, the settings are recommended and required while in other cases, they are recommended but optional.

- COS Preservation (Fabric Wide)
- Enforce Subnet Check (Fabric Wide, Optional)
- Limit IP Learning to Subnet (Bridge Domain Level, Optional)
- IP Aging (Fabric Wide, Optional)
- Endpoint Learning Features
 - Endpoint Dataplane Learning (Bridge Domain Level, Enabled by default)
 - Layer 2 Unknown Unicast (Bridge Domain Level)
 - Clear Remote MAC Entries (Bridge Domain Level, Optional)

- Unicast Routing (Bridge Domain Level)
- ARP Flooding (Bridge Domain Level)
- GARP Based Detection for EP Move Detection Mode (Bridge Domain Level)
- Jumbo Frames and MTU

The implementation of the above features can vary depending on the generation of ACI leaf switches used in the deployment. Some examples of first and second-generation Cisco ACI leaf switches are provided below - see the Cisco Product documentation for a complete list.

- First-generation Cisco ACI leaf switches models: Nexus 9332PQ, Nexus 9372 (PX, PX-E, TX, TX-E), Nexus 9396 (PX, TX), 93120TX, 93128TX switches
- Second-generation Cisco ACI leaf switches models: Nexus 9300-EX and 9300-FX Series, Nexus 9348GC-FXP, Nexus 9336C-FX2, Nexus 93240YC-FX2 switches.

COS Preservation (Fabric Wide Setting)

Class Of Service (COS) Preservation feature in ACI preserves the COS setting in the traffic received from the endpoints. This feature should be enabled in all HyperFlex deployments to preserve the COS end-to-end across an ACI fabric, including an ACI Multi-Pod fabric.

To enable COS Preservation, follow these steps:



This policy has a fabric-wide impact.

1. Use a browser to navigate to APIC's Web GUI. Log in using the **admin** account.
2. From the top menu, select **Fabric > Access Policies**.
3. In the left navigation pane, select and expand **Policies > Policies > Global**.
4. In the right window plane, select the **QOS Class** tab. For **Preserve QOS**, enable the checkbox for **Dot1p Preserve** is selected.

The screenshot shows the Cisco APIC web interface. The top navigation bar includes 'System', 'Tenants', 'Fabric' (selected), 'Virtual Networking', 'L4-L7 Services', 'Admin', 'Operations', and 'Apps'. Below this, a sub-navigation bar shows 'Inventory', 'Fabric Policies', and 'Access Policies'. The left sidebar lists various policy categories, with 'Global' under 'Policies' selected. The main content area is titled 'Global' and shows the configuration for a 'QOS Class' policy. The 'Properties' section has 'Preserve COS' checked and 'Dot1p Preserve' selected. Below this is a table with columns: Name, Admin State, Priority Flow Contrn, No-Drop-Cos, MTU, Minim Buffer, Congr Algori, Congr Notific, Queu Contr, Queu Limit (bytes), Scheduling Algorithm, and Bandw alloca (in %). The table lists five levels (Level1 to Level5) with their respective configurations. At the bottom, there are buttons for 'Show Usage', 'Reset', and 'Submit'.

Name	Admin State	Priority Flow Contrn	No-Drop-Cos	MTU	Minim Buffer	Congr Algori	Congr Notific	Queu Contr	Queu Limit (bytes)	Scheduling Algorithm	Bandw alloca (in %)
Level1	Enabled	fa...		9216	0	Ta...	Di...	Dy...	15...	Weighted...	20
Level2	Enabled	fa...		9216	0	Ta...	Di...	Dy...	15...	Weighted...	20
Level3 (De...	Enabled	fa...		9216	0	Ta...	Di...	Dy...	15...	Weighted...	20
Level4	Enabled	fa...		9216	0	Ta...	Di...	Dy...	15...	Weighted...	0
Level5	Enabled	fa...		9216	0	Ta...	Di...	Dy...	15...	Weighted...	0

5. Click Submit.

Enforce Subnet Check for Endpoint Learning (Fabric Wide Setting)

Enforce Subnet Check in ACI limits both local and remote IP endpoint learning in a VRF to source IP addresses that belong to one of the bridge domain subnets defined for that VRF. This is a fabric wide policy that impacts data plane learning on all VRFs. Note that for local learning, the source IP address must be in its bridge domain subnet but for remote learning, the source IP just needs to match one of the bridge domain subnets for the VRF.

For subnets outside the VRF, enabling this feature will prevent both MAC and IP addresses from being learned for local endpoints, and IP addresses for remote endpoints. This feature provides a better check than the **Limit IP Learning to Subnet** covered in the next section, which does the subnet check for IP addresses but not for MAC addresses. Also, it does the check only for learning local endpoints and not for remote endpoints. However the **Limit IP Learning to Subnet** feature is more granular in scope as it does the subnet-check on a per bridge domain basis while the **Enforce Subnet Check** does a check against all subnets at the VRF level and is enabled/disabled at the fabric level so it applies to all VRFs in the fabric.

Limiting endpoint learning will reduce ACI fabric resource usage and therefore it is recommended but optional. This feature is disabled by default.

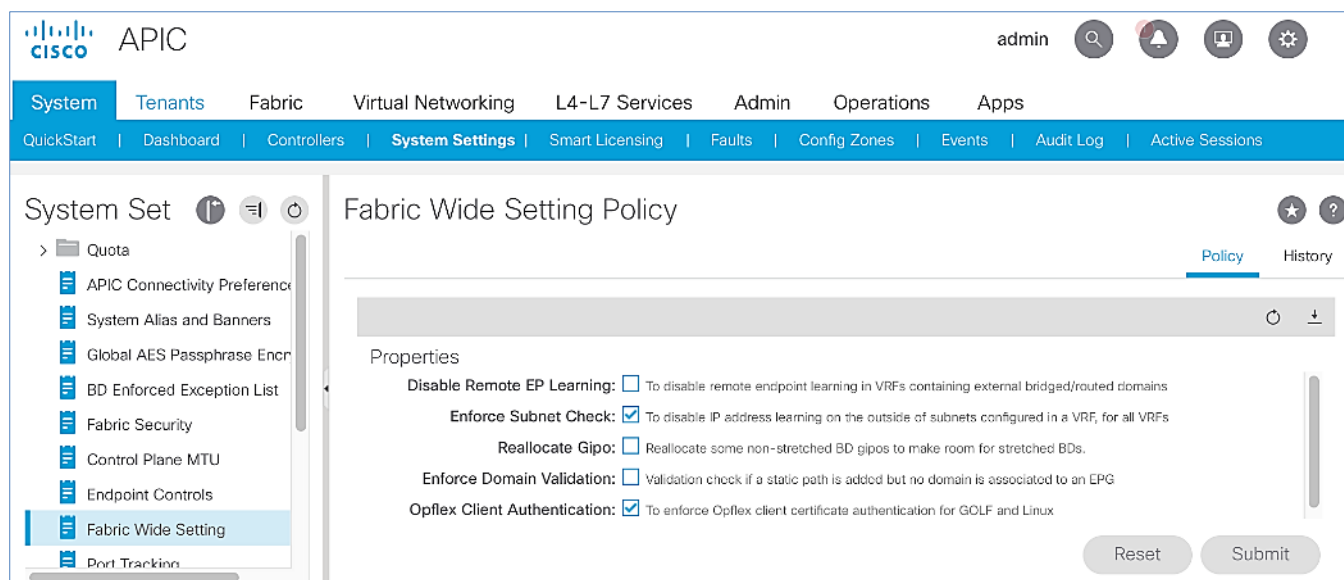
Some guidelines regarding this feature are provided below:

- This feature is available only on second-generation leaf switches. In a mixed environment with first and second-generation leaf switches, the first-generation switches will ignore this feature.
- Enabling this feature will enable it fabric-wide, across all VRFs though the subnet-check is for the subnets in the VRF.

- Available in APIC Releases 2.2(2q) and higher 2.2 releases and in 3.0(2h) and higher. It is not available in 2.3 or 3.0(1x) releases.
- The feature can be enabled/disabled under Fabric > Access Policies > Global Policies > Fabric Wide Setting Policy in earlier releases and under System > System Settings > Fabric Wide Setting in newer releases.

To enable Enforce Subnet Check feature, follow these steps:

1. Use a browser to navigate to APIC's Web GUI. Log in using the **admin** account.
2. From the top menu, select **System > System Settings**.
3. In the left navigation pane, select **Fabric Wide Setting**.
4. In the right window pane, enable check box for **Enforce Subnet Check**.



5. Click **Submit**.

Limit IP Learning to Subnet (Bridge-domain, Optional)

This is a bridge-domain level setting. It is superseded by the Enforced Subnet Check feature in the previous section. This feature changes the default endpoint "IP" address learning behavior of the ACI fabric. Enabling this feature will disable IP address learning on subnets that are not part of the bridge domain subnets and only learn if the source IP address belongs to one of the configured subnets for that bridge domain. A bridge domain can have multiple IP subnets and enabling this feature will limit the IP address learning to the bridge-domain subnets but will not learn addresses for subnets outside the bridge-domain. This feature will also reduce ACI fabric resource usage and therefore it is recommended but optional.

This feature is available as of APIC release 1.1(1j) and enabled by default as of APIC releases 2.3(1e) and 3.0(1k). This feature can be enabled for HyperFlex deployments as shown in the figure below.

Figure 3 Cisco ACI Fabric Settings: Limit IP Learning to Subnet

The screenshot shows the Cisco APIC interface for configuring a Bridge Domain. The left sidebar shows the navigation tree with 'Tenant HXV-Foun' selected. The main panel is titled 'Bridge Domain - HXV-IB-MGMT_BD' and has tabs for 'Summary', 'Policy', 'Operational', 'Stats', 'Health', 'Faults', and 'History'. The 'Policy' tab is active, and the 'General' sub-tab is selected. The 'Properties' section is expanded, showing various configuration options. The 'Limit IP Learning To Subnet' checkbox is checked. Other settings include 'Name: HXV-IB-MGMT_BD', 'Type: regular', 'VLAN: HXV-Foundation_VRF', and 'Resolved VRF: HXV-Foundation/HXV-Foundation_VRF'.

Note the following regarding this feature:

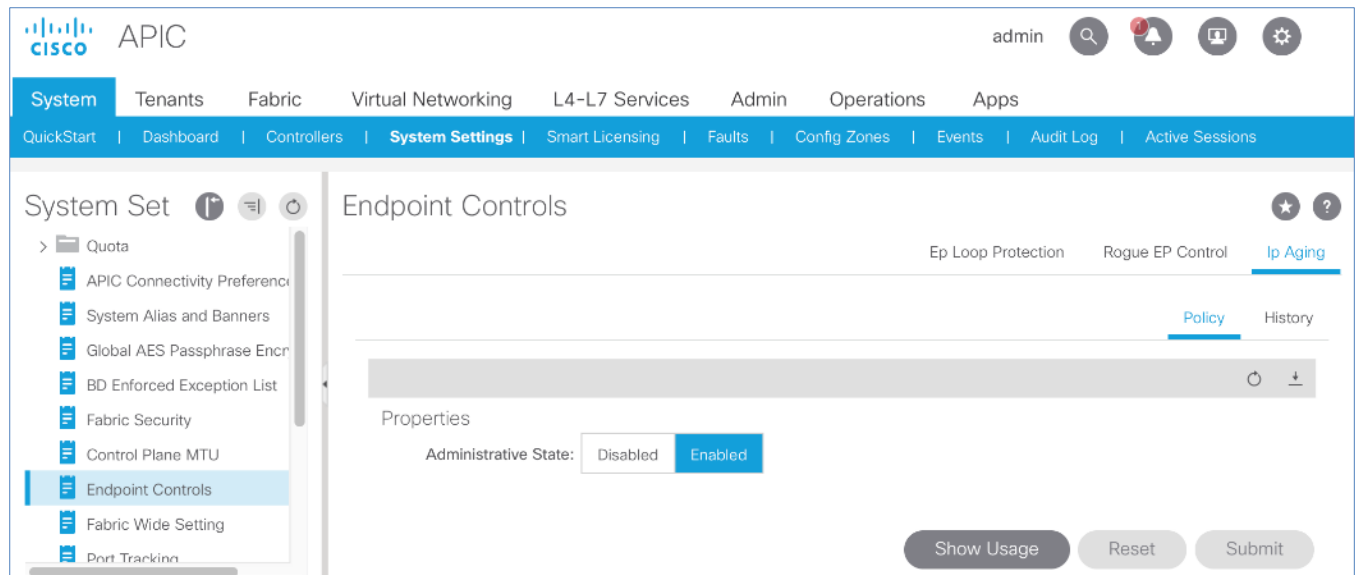
- Available on first and second-generation of ACI leaf switches
- If **Enforce Subnet Checking** is also enabled, it supersedes this feature.
- This feature should be used when subnet-check is for a specific bridge domain (as opposed to all VRF subnets) or when you have an environment with first-generation leaf switches.
- Prior to APIC release 3.0(1k), toggling this feature with **Unicast Routing** enabled could result in an impact of 120s. In prior releases, ACI flushed all endpoints addresses and suspended learning on the bridge domain for 120s. The behavior in 3.0(1k) and later releases is to only flush endpoint IP addresses that are not part of the bridge domain subnets and there is no suspension of address learning.

IP Aging (Fabric Wide Setting)

IP Aging tracks and ages endpoint IP addresses that the fabric has learned, to age out stale entries. This is a fabric wide setting. This feature will also reduce ACI fabric resource usage and therefore it is recommended but optional.

To enable IP aging, follow these steps:

1. Use a browser to navigate to APIC's Web GUI. Log in using the **admin** account.
2. From the top menu, select **System > System Settings**.
3. In the left navigation pane, select **Endpoint Controls**.
4. In the right window pane, select **IP Aging** tab and then **Policy** tab. For **Administrative State**, click **Enabled**.



5. Click **Submit**.

Endpoint Learning

Endpoint learning in ACI is primarily done in hardware from data-plane traffic by examining the incoming traffic, specifically the source MAC and IP address fields in the received traffic. ACI can learn the address (MAC, IP) and location of any endpoint that sends traffic to the fabric. ACI provides several configuration settings (mostly at the bridge-domain level) that impact endpoint learning behavior.

IP vs. MAC Learning

By default, ACI learns the MAC address of all endpoints but for any "IP" learning to occur, **Unicast Routing** must be enabled at the bridge-domain level. **Unicast Routing** enables both Layer 3 forwarding and IP learning in an ACI fabric. The **Endpoint Dataplane Learning** feature is available at the bridge-domain level – see next section.

Silent Hosts

ACI typically learns from data-plane traffic but for silent endpoints that do not send any traffic to the fabric, ACI can also use control plane protocols such as ARP and GARP to do endpoint learning. The behaviour varies depending on whether the Bridge Domain is doing Layer 2 forwarding (**Unicast Routing** disabled) or Layer 3 forwarding (**Unicast Routing** enabled).

For bridge-domains doing Layer 2 forwarding (**Unicast Routing** disabled), **ARP flooding** can be used to learn the location of silent endpoints. **ARP Flooding** enables ACI to learn from the data-plane ARP traffic exchanged between the endpoints. In this scenario, the **L2 Unknown Unicast** option should also be set to "Flood" to prevent ACI from dropping unicast traffic destined to endpoints that it hasn't learned of yet.



APIC GUI automatically enables **ARP Flooding** if **L2 Unknown Unicast** is set to “Flood”. However, regardless of the GUI setting, ARP Flooding is always enabled in hardware when **Unicast Routing** is disabled.

For bridge-domains doing Layer 3 forwarding (**Unicast Routing** enabled), ACI can learn the location of silent or unknown hosts either by generating an ARP request or from data-plane ARP traffic. If IP subnet(s) are configured for the bridge-domain, ACI can generate an ARP request and learn the location of the unknown endpoint from its ARP response (also known as **ARP gleaning**). If **Unicast Routing** is enabled without configuring bridge-domain subnets (not recommended), ACI cannot initiate ARP requests. However, ACI can still learn their location from the data-plane ARP traffic. Though **ARP Flooding** is not necessary in first scenario, it should be enabled so that if the endpoint moves, ACI can learn the new location quickly rather than waiting for ACI to age out the entry for the endpoint. ACI can also detect endpoint moves using GARP by enabling the **GARP-based endpoint move detection** feature.



ARP Flooding must be enabled for GARP-based endpoint move detection feature.

Local vs. Remote Endpoints

Endpoint learning in ACI also depends on whether the endpoints are local or remote endpoints. For a given leaf switch, local endpoints are local to that leaf switch while remote endpoints connect to other leaf switches. Local and remote endpoints are also learned from data-plane traffic. However, unlike local endpoints, ACI typically learns either the MAC or IP address of remote endpoints but not both. The local endpoints information is sent to the Spine switches that maintain the endpoint database but remote endpoints are maintained on the leaf switches. Remote entries are also aged out sooner than local endpoints by default.

As stated earlier, ACI provides several options that impact endpoint learning. These settings are covered in more detail in the upcoming sections.

Endpoint Dataplane Learning

Endpoint Dataplane Learning is bridge-domain level setting that enables/disables “IP” learning in the data-plane. This feature is available as of APIC release 2.0(1m) and it is enabled by default as shown in the figure below.

Figure 4 Cisco ACI Fabric Settings: Endpoint Dataplane Learning

The screenshot displays the Cisco ACI Fabric Settings interface for the Bridge Domain - HXV-IB-MGMT_BD. The left sidebar shows the navigation tree with 'Tenant HXV-Fo' selected. The main content area shows the 'Policy' tab for the bridge domain. The 'General' sub-tab is active, showing various configuration options. The 'L2 Unknown Unicast' setting is set to 'Flood', and 'Endpoint Dataplane Learning' is checked. Other settings include 'Type' set to 'regular', 'Advertise Host Routes' checked, 'Enable Legacy Mode' unchecked, 'Legacy Mode' set to 'No', 'VLAN' set to 'HXV-IB-MGMT_BD', 'VRF' set to 'HXV-IB-MGMT_BD_VRF', 'Resolved VRF' set to 'HXV-IB-MGMT_BD_VRF', 'L3 Unknown Multicast Flooding' set to 'Flood', 'Multi Destination Flooding' set to 'Flood in BD', 'PIM' unchecked, 'IGMP Policy' set to 'select an option', 'ARP Flooding' checked, 'Clear Remote MAC Entries' unchecked, 'Limit IP Learning To Subnet' checked, 'Endpoint Retention Policy' set to 'select a value', and 'IGMP Snoop Policy' set to 'select a value'.

L2 Unknown Unicast

L2 Unknown Unicast is a bridge-domain level setting that specifies how unknown Layer 2 unicast frames should be forwarded within the fabric. This field can be set to "Flood" or "Hardware Proxy" (default) mode. In "Flood mode", the unknown Layer 2 unicast frames are flooded across all ports in the bridge-domain using the bridge-domain specific multicast tree. In "Hardware Proxy" mode, the unknown unicast frames are sent to the spine switch to do a lookup in the endpoint mapping database. However, if the spine has not learned the address of that endpoint, the unicast traffic will be dropped by the fabric. For this reason, if a Layer 2 bridge-domain has **silent** endpoints, the **L2 Unknown Unicast** field should always be set to "Flood".

The default setting for **L2 Unknown Unicast** is "Hardware-Proxy" but in this design, this field is set to "Flood" for deployments that may have **silent** hosts. This feature can be enabled as shown in the figure below.

The screenshot shows the Cisco APIC GUI for creating a bridge domain. The left sidebar lists navigation options: Quick Start, Tenant HXV-Founda, Application Profiles, Networking, Bridge Domains, VRFs, External Bridged Networks, External Routed Networks, Dot1Q Tunnels, Contracts, Policies, and Services. The main area is titled 'Networking - Bridge Domains' and contains a table of existing bridge domains:

Name	Alias	Type	Segment	Multicast Address	Custom MAC Address	L2 Unknown Unicast
HXV-IB-MGMT_BD		regular	15073232	225.1.133.240	00:22:BD:F8:19:FF	Flood

Below the table is the 'Create Bridge Domain' form. It includes a progress bar with three steps: 1. Main, 2. L3 Configurations, and 3. Advanced/Troubleshooting. The form is titled 'Specify Bridge Domain for the VRF' and contains the following fields:

- Name: HXV-STORAGE_BD
- Alias: (empty)
- Description: optional
- Tags: (empty)
- Type: fc regular
- VRF: HXV-Founda_VRF
- Forwarding: Custom
- L2 Unknown Unicast: Flood
- L3 Unknown Multicast Flooding: Flood
- Multi Destination Flooding: Flood in BD
- ARP Flooding: ☒ Enabled
- Clear Remote MAC Entries: ☐
- Endpoint Retention Policy: select a value
- IGMP Snoop Policy: select a value

At the bottom right of the form are buttons for 'Previous', 'Cancel', and 'Next'.

This feature requires **ARP Flooding** to be enabled on the bridge-domain. Customers may also want to enable the **Clear Remote MAC Entries** setting. See upcoming sections for additional information on these two settings.

Clear Remote MAC Entries

This is a bridge-domain level setting that clears the remote Layer 2 MAC addresses on other switches when the corresponding MAC addresses (learnt on a vPC) are deleted from a local switch. The entries are cleared on all remote switches if it is deleted on a local switch. The setting is visible in the GUI when **L2 Unknown Unicast** is set to "Flood". This feature is optional but recommended for deployments that may have **silent** hosts.

Unicast Routing

Unicast Routing setting on the bridge-domain enables both Layer 3 forwarding and "IP" learning in an ACI fabric. The IP endpoint learning is primarily done from the data plane traffic but ACI can also initiate ARP requests to do endpoint learning in the control plane. ACI can originate ARP requests for unknown endpoints if both **Unicast Routing** and bridge-domain subnet is configured. However, ACI cannot generate ARP requests if a subnet is not configured for the bridge-domain, but it can still learn their location from the data-plane ARP traffic if **ARP Flooding** is enabled. In this design, **Unicast Routing** is enabled on HyperFlex bridge-domains **except** for the storage-data bridge-domain.

ARP Flooding

ARP Flooding is used for both Layer 2 (**Unicast Routing** disabled) and Layer 3 bridge-domains (**Unicast Routing** enabled). By default, with **Unicast Routing** enabled, the ACI fabric will treat ARP requests as unicast packets and forward them using

the target IP address in the ARP packets. It will not flood the ARP traffic to all the leaf nodes in the bridge domain. However, the **ARP Flooding** setting provides the ability to change this default behavior and flood the ARP traffic across the fabric to all the leaf nodes in a given bridge domain. See Endpoint Learning section above for other scenarios that require **ARP Flooding**.

ARP Flooding is also required in environments that use Gratuitous ARP (GARP) to indicate an endpoint move. If an endpoint move occurs on the same EPG interface, GARP feature must be enabled in ACI to detect the endpoint move – see GARP based Detection section for more details.

This feature is disabled by default but it is enabled in this design for deployments that may have silent hosts or require GARP. This feature can be enabled as shown in the figure below. Cisco ACI Fabric Settings: ARP Flooding

The screenshot displays the Cisco APIC interface for configuring a Bridge Domain. The left sidebar shows the navigation tree with 'Tenant HXV-' selected, and 'Bridge Domains' expanded. The main panel shows the configuration for 'Bridge Domain - HXV-IB-MGMT_BD'. The 'Policy' tab is active, and the 'General' sub-tab is selected. The 'Properties' section shows the following settings:

- Name: HXV-IB-MGMT_BD
- Alias: (empty)
- Description: optional
- Global Alias: (empty)
- Tags: (empty)
- Type: fc (selected), regular
- Advertise Host Routes: ☒
- Enable Legacy Mode: ☐
- Legacy Mode: No
- VLAN: (empty)
- VRF: HXV-Foundation_VRF
- Resolved VRF: HXV-Foundation/HXV-Foundation_VRF
- L2 Unknown Unicast: Flood (selected), Hardware Proxy
- L3 Unknown Multicast Flooding: Flood (selected), Optimized Flood
- Multi Destination Flooding: Flood in BD (selected), Drop, Flood in Encapsulation
- PIM: ☐
- IGMP Policy: select an option
- ARP Flooding: ☒
- Endpoint Dataplane Learning: ☒
- Clear Remote MAC Entries: ☐
- Limit IP Learning To Subnet: ☒
- Endpoint Retention Policy: select a value
- IGMP Snoop Policy: select a value

At the bottom right, there are buttons for 'Show Usage', 'Reset', and 'Submit'.

GARP-based Detection

Gratuitous ARP (GARP) based detection setting enables ACI to detect an endpoint IP move from one MAC address to another when the new MAC is on the same EPG interface as the old MAC. ACI can detect all other endpoint IP address

moves such as moves between ports, switches, EPGs or bridge-domains but not when it occurs on the same EPG interface. With this feature, ACI can use GARP to learn of an endpoint IP move on the same EPG interface.

This is a bridge-domain level setting that can be enabled as shown in the figure below.

Figure 5 Cisco ACI Fabric Settings: GARP-based Detection

The screenshot displays the Cisco APIC interface for configuring a bridge domain. The left sidebar shows the navigation tree with 'Networking' selected. The main panel is titled 'Networking - Bridge Domains' and contains a table of existing bridge domains. Below this, the 'Create Bridge Domain' wizard is active, showing 'STEP 2 > L3 Configurations'. The configuration options include:

- Unicast Routing:** ☒ Enabled
- ARP Flooding:** ☒ Enabled
- Config BD MAC Address:** ☒ Enabled
- MAC Address:** 00:22:BD:FB:19:FF
- Virtual MAC Address:** not-applicable
- Subnets:** A table with columns: Gateway Address, Scope, Primary IP Address, Subnet Control.
- Endpoint Dataplane Learning:** ☒ Enabled
- Limit IP Learning To Subnet:** ☒ Enabled
- EP Move Detection Mode:** ☒ GARP based detection
- DHCP Labels:** A table with columns: Name, Scope, DHCP Option Policy.
- Associated L3 Outs:** A table with columns: L3 Out.

The 'Next' button is visible at the bottom right of the configuration panel.

Note that **ARP Flooding** must be enabled to use this feature. **GARP-based detection** setting will not be visible on the GUI until **ARP Flooding** is enabled on the bridge domain.

Jumbo Frames and MTU

Traditional switching fabrics typically use a 1500B MTU and must be configured to support Jumbo frames. However, the ACI fabric, by default, uses an MTU of 9150B on core-facing ports of leaf and spine switches and 9000B on access ports of leaf switches. Therefore, no configuration is necessary to support Jumbo frames on an ACI fabric.

Pre-configure Access Policies for ACI Fabric

Fabric Access Policies are policies that are applied to access layer connections, typically on leaf switches. The access layer connections can be to a physical domain or a virtual domain managed by a Virtual Machine Manager (VMM). The physical domains in this design include vPC connections to Cisco UCS/HyperFlex domain and Layer 3 connections to external

networks. Cisco recommends configuring all policies explicitly even when the policies match the defaults to avoid issues in the future as defaults can change in newer releases. Policies can be re-used across the fabric to configure any number of access layer. This section provides the procedures for pre-configuring policies that will be used in upcoming sections of this guide.

Setup Information

The pre-configured policies used in this design are summarized in [Table 5](#).

Table 5 Fabric Access Policies

Access Interface Policies	Policy Name	Purpose
Link Level Policies	40Gbps-Link	Sets link to 40Gbps
	10Gbps-Link	Sets link to 10Gbps
	1Gbps-Link	Sets link to 1Gbps
	Inherit-Link	Inherits the negotiated link speed
CDP Interface Policies	CDP-Enabled	Enables CDP
	CDP-Disabled	Disables CDP
	LLDP-Enabled	Enables LLDP
LLDP Interface Policies	LLDP-Disabled	Disables LLDP
	LACP-Active	Sets LACP Mode
Port Channel Policies	MAC-Pinning-Phy-NIC-Load	Sets MAC Pinning-Physical-NIC-load
	MAC-Pinning	Sets MAC Pinning
	VLAN-Scope-Local	Specifies VLAN Scope as Port Local
Layer 2 Interface Policies	VLAN-Scope-Global	Specifies VLAN Scope as Global
	BPDU-FG-Enabled	Enables BPDU Filter and Guard
Spanning Tree Policies	BPDU-FG-Disabled	Disables BPDU Filter and Guard
	Firewall-Disabled	Disables Firewall

Deployment Steps

To configure all policies from the following location in the GUI, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Access Policies**.
3. From the left navigation pane, select and expand **Policies > Interface**.
4. Create all the policies in [Table 5](#) by following the steps in the next sections.

Create Link Level Policies

To create the link level policies to specify link speeds of 1/10/40-Gbps and other link policies, follow these steps:

1. From the left navigation pane, select **Link Level**. Right-click and select **Create Link Level Policy**.
2. In the **Create Link Level Policy** pop-up window, specify the policy **Name**. For the **Speed**, select **1Gbps** from the drop-down list.

The screenshot displays the Cisco APIC web interface. The top navigation bar includes 'System', 'Tenants', 'Fabric' (selected), 'Virtual Networking', 'L4-L7 Services', 'Admin', 'Operations', and 'Apps'. Below this, a sub-navigation bar shows 'Inventory', 'Fabric Policies', and 'Access Policies'. The left sidebar lists various policy categories under 'Policies', with 'Interface' expanded and 'Link Level' selected. The main content area shows a table titled 'Interface - Link Level' with columns: Name, label, Auto Negotiation, Speed, Link Debounce Interval, Forwarding Error Correction, and Description. A modal dialog box titled 'Create Link Level Policy' is open, prompting the user to 'Specify the Physical Interface Policy Identity'. The dialog contains the following fields and options:

- Name:** 1Gbps-Link
- Description:** optional
- Alias:** (empty field)
- Auto Negotiation:** off (radio button), on (radio button, selected)
- Speed:** 1 Gbps (dropdown menu)
- Link debounce interval (msec):** 100 (dropdown menu)
- Forwarding Error Correction:** CL74-FC-FEC, CL91-RS-FEC, disable-FEC, Inherit (button, highlighted)

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

3. Click **Submit** to complete creating the policy.
4. Repeat steps 1-3 to create a link policies for **10Gbps** and **40Gbps** speeds.
5. Repeat steps 1-3 to create an **inherit** link policy as shown below.

Create Link Level Policy
Specify the Physical Interface Policy Identity

Name:

Description:

Alias:

Auto Negotiation: ☐ ☒

Speed:

Link debounce interval (msec):

Forwarding Error Correction: ☐ ☐ ☐ ☒

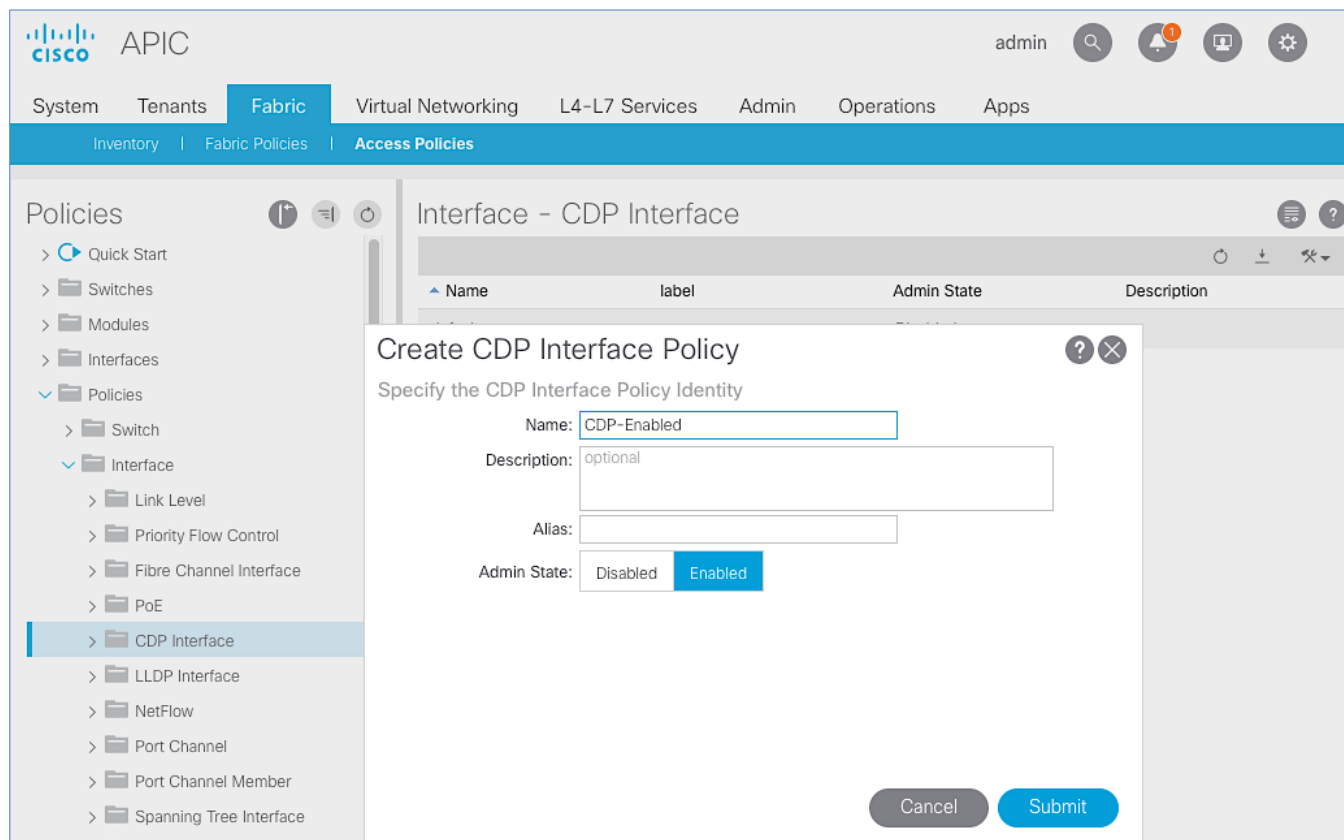
6. Click **Submit** to complete.
7. You should now have the following Link policies in place:

Name	label	Auto Negotiation	Speed	Link Debounce Interval (msec)	Forwarding Error Correction	Description
10Gbps-Link		on	10 Gbps	100	Inherit	
1Gbps-Link		on	1 Gbps	100	Inherit	
40Gbps-Link		on	40 Gbps	100	Inherit	
default		on	inherit	100	Inherit	
Inherit-Link		on	inherit	100	Inherit	

Create CDP Interface Policies

To create CDP interface policies, follow these steps:

1. From the left navigation pane, select **CDP Interface**. Right-click and select **Create CDP Interface Policy**.
2. In the **Create CDP Interface Policy** pop-up window, specify the policy **Name**. For **Admin State**, click **Enabled**.

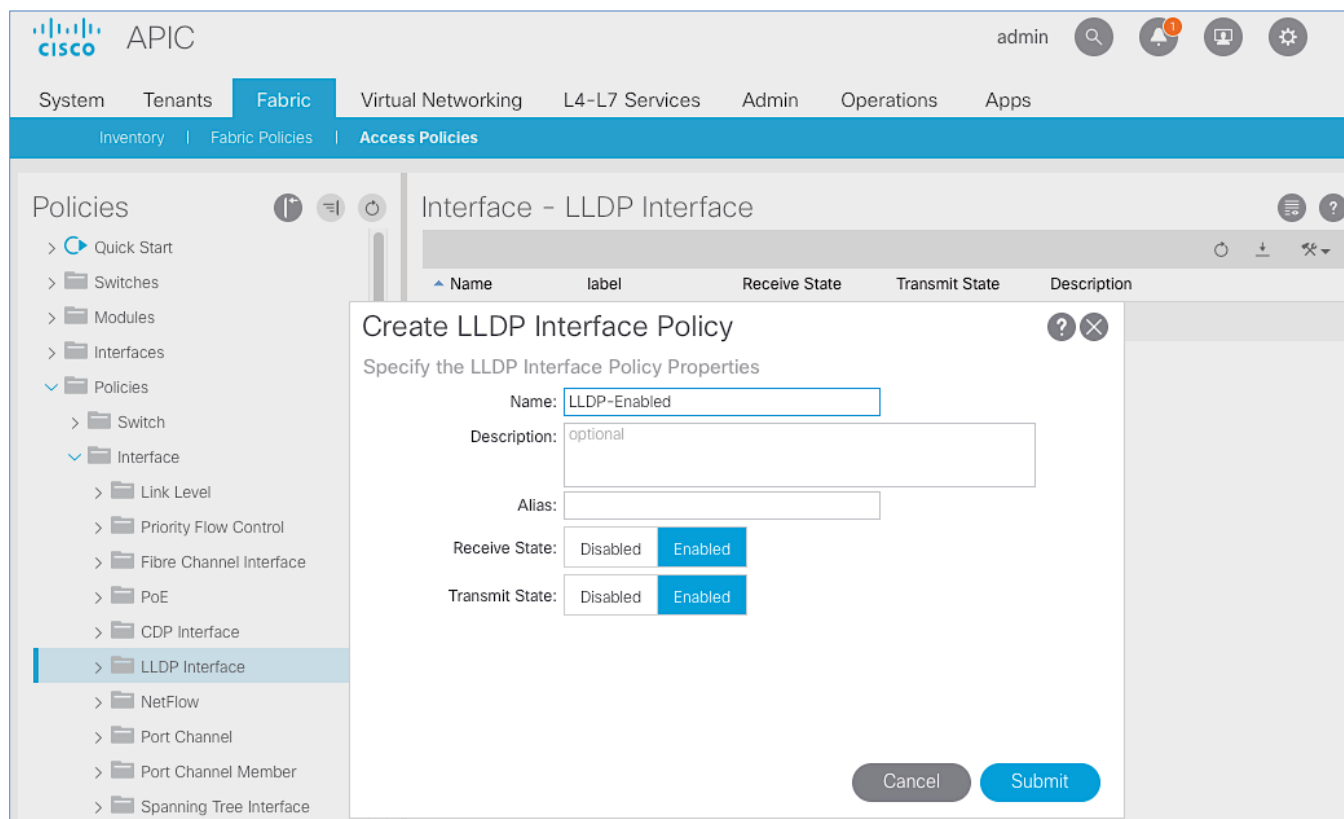


3. Click **Submit** to complete creating the policy.
4. Repeat steps 1-3 to create a policy to disable CDP. The **Admin State** for this policy should be **Disabled**.

Create LLDP Interface Policies

To create LLDP interface policies, follow these steps:

1. From the left navigation pane, select **LLDP Interface**. Right-click and select **Create LLDP Interface Policy**.
2. In the Create LLDP Interface Policy pop-up window, specify the Name. For the Receive and Transmit State, click Enabled.



3. Click **Submit** to complete creating the policy.
4. Repeat steps 1-3 to create a policy to disable LLDP. The **Receive** and **Transmit states** for this policy should be **Disabled**.

Create Port Channel Policies

To create port channel policies, follow these steps:

1. From the left navigation pane, select **Port Channel**. Right-click and select **Create Port Channel Policy**.
2. In the **Create Port Channel Policy** pop-up window, specify the policy **Name**. For the **Mode**, select **LACP-Active** from the drop-down list. Leave everything else as-is.

APIC admin

System | Tenants | **Fabric** | Virtual Networking | L4-L7 Services | Admin | Operations | Apps

Inventory | Fabric Policies | **Access Policies**

Policies

- > Quick Start
- > Switches
- > Modules
- > Interfaces
- > Policies
 - > Switch
 - > Interface
 - > Link Level
 - > Priority Flow Control
 - > Fibre Channel Interface
 - > PoE
 - > CDP Interface
 - > LLDP Interface
 - > NetFlow
 - > **Port Channel**
 - > Port Channel Member
 - > Spanning Tree Interface
 - > Data Plane Policing
 - > MCP Interface
 - > CoPP Interface
 - > L2 Interface
 - > Port Security

Create Port Channel Policy

Specify the Port Channel Policy

Name: LACP-Active

Description: optional

Alias:

Mode: LACP Active
Not Applicable for FC PC

Control: Suspend Individual Port Graceful Convergence Fast Select Hot Standby Ports

Cancel Submit

Last Login Time: 2018-12-11T09:19 UTC-05:00

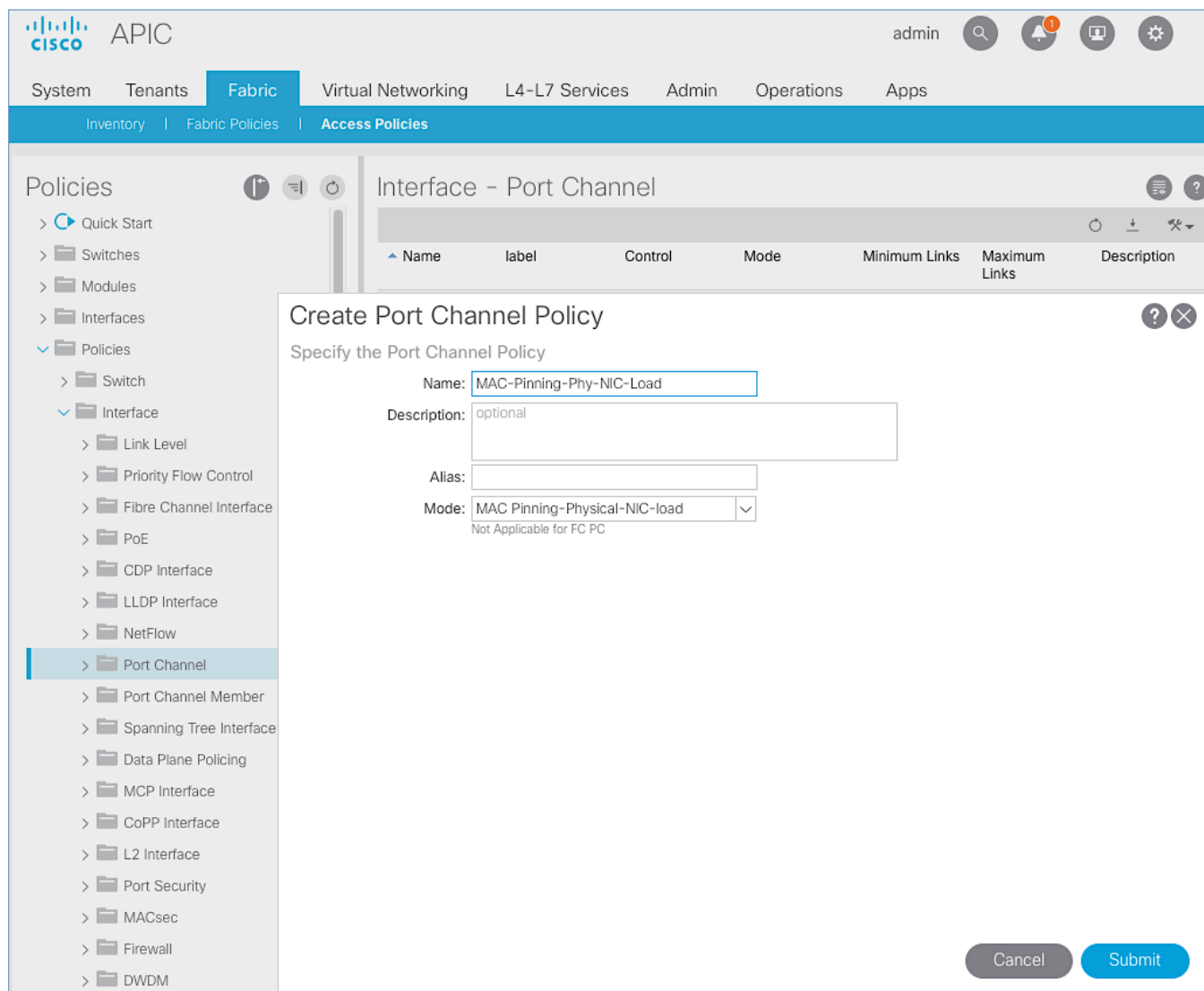
3. Click **Submit** to complete creating the policy.
4. Repeat steps 1-3 to create a port-channel policy for mac-pinning as shown below.

The screenshot displays the Cisco APIC (Application Centric Infrastructure) interface. The top navigation bar includes 'System', 'Tenants', 'Fabric', 'Virtual Networking', 'L4-L7 Services', 'Admin', 'Operations', and 'Apps'. The 'Fabric' tab is active, and the 'Access Policies' sub-tab is selected. On the left, a 'Policies' sidebar lists various policy categories, with 'Port Channel' highlighted under the 'Interface' section. The main content area shows a table titled 'Interface - Port Channel' with columns: Name, label, Control, Mode, Minimum Links, Maximum Links, and Description. A modal dialog box titled 'Create Port Channel Policy' is open, prompting the user to 'Specify the Port Channel Policy'. The dialog contains the following fields:

- Name:** MAC-Pinning
- Description:** optional
- Alias:** (empty field)
- Mode:** MAC Pinning (dropdown menu)

Below the Mode dropdown, a note states: 'Not Applicable for FC PC'. At the bottom right of the dialog are 'Cancel' and 'Submit' buttons.

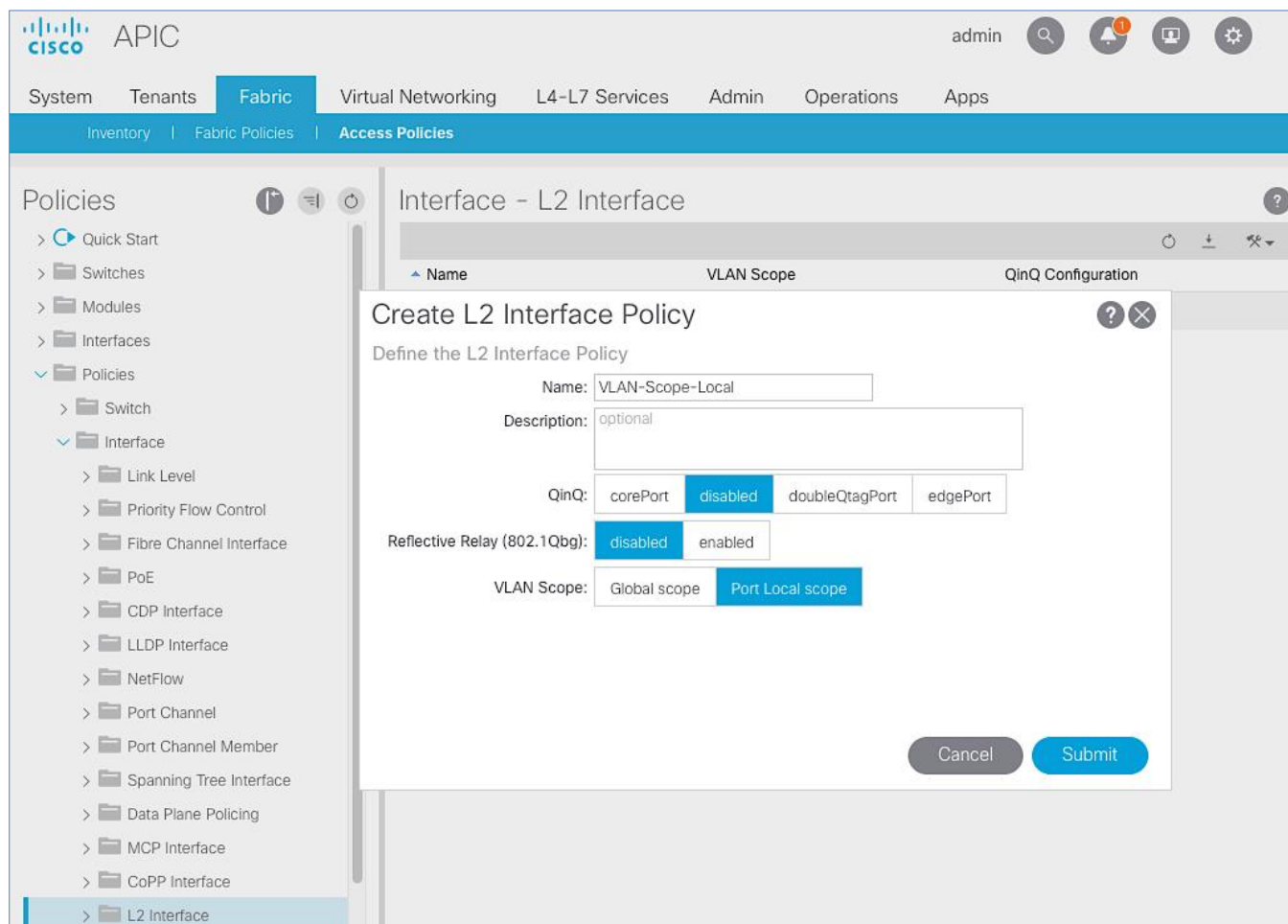
5. Click **Submit** to complete creating the policy.
6. Repeat steps 1-3 to create a policy for mac-pinning based on physical NIC load as shown below.



Create L2 Interface (VLAN Scope) Policies

To create L2 interface policies, follow these steps:

1. From the left navigation pane, select **L2 Interface**. Right-click and select **Create L2 Interface Policy**.
2. In the **Create L2 Interface Policy** pop-up window, specify the policy Name. For **VLAN Scope**, select **Port Local scope**.

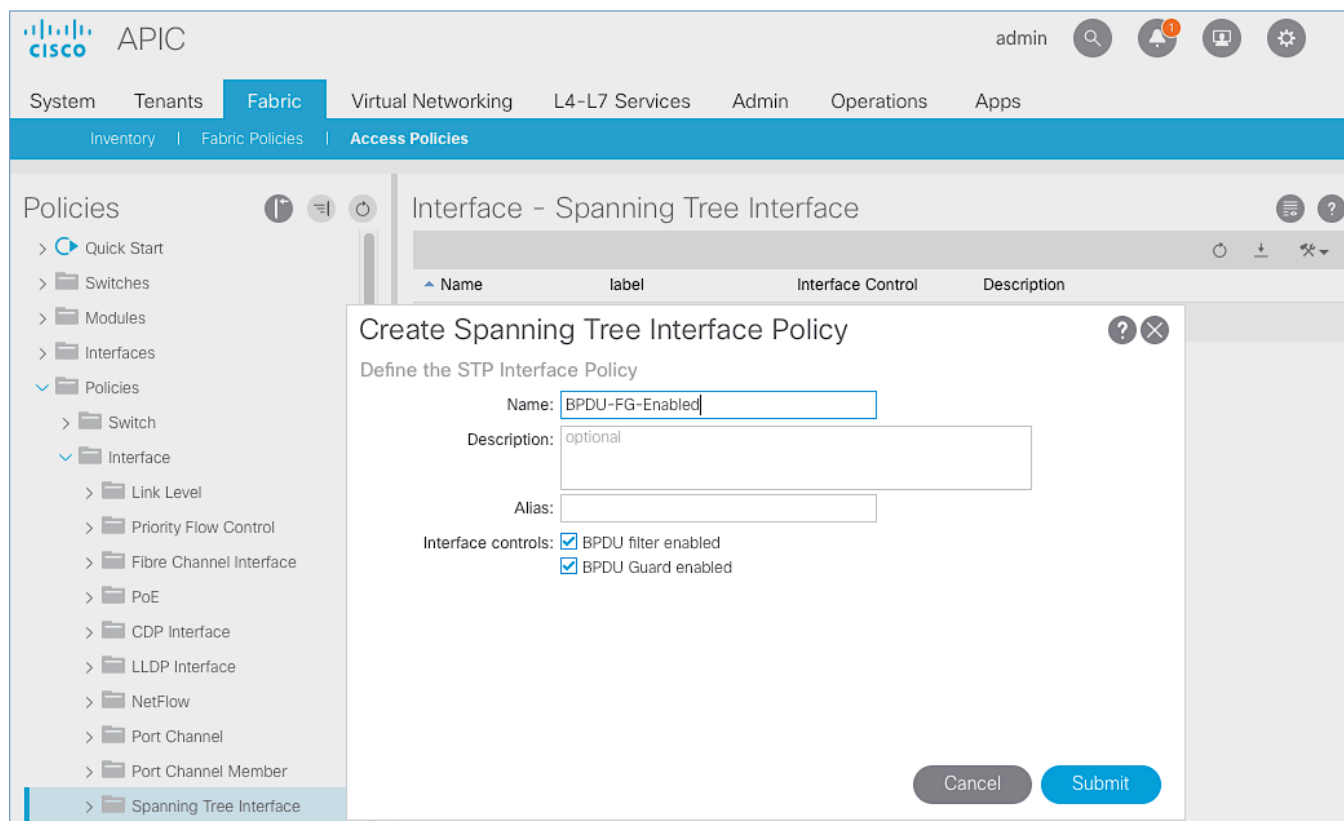


3. Click **Submit** to complete creating the policy.
4. Repeat steps 1-3 to create a L2 Interface policy for VLAN scope global. The **VLAN Scope** for this policy should be **Global scope**.

Create Spanning Tree Interface Policies

To create spanning tree interface policies, follow these steps:

1. From the left navigation pane, select **Spanning Tree Interface**. Right-click and select **Create Spanning Tree Interface Policy**.
2. In the Create Spanning Tree Interface Policy pop-up window, specify the policy Name. For Interface Controls, select checkbox for BPDU Filter enabled and BPDU Guard enabled.



3. Click **Submit** to complete creating the policy.
4. Repeat steps 1-3 to create a policy to disable BPDU Filter and Guard. The **Interface Controls** for this policy should leave both **BPDU filter enabled** and **BPDU Guard enabled** unchecked.

Create Firewall Policy

To create a firewall policy, follow these steps:

1. From the left navigation pane, select **Firewall**. Right-click and select **Create Firewall Policy**.
2. In the **Create Firewall Policy** pop-up window, specify a policy **Name**.
3. For Mode, select **Disabled**. Leave all other values as is.

The screenshot shows the Cisco APIC interface with the 'Fabric' tab selected. A 'Create Firewall Policy' dialog box is open, allowing configuration of a firewall policy. The dialog includes the following fields and options:

- Name:** Firewall-Disabled
- Description:** optional
- Mode:** Disabled (selected), Enabled, Learning
- SysLog Section:**
 - Administrative State:** enabled
 - Included Flows:** Denied flows
 - Polling Interval (seconds):** 60
 - Log Level:** information
 - Dest Group:** select an option

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

4. Click **Submit** to complete creating the policy.

Solution Deployment – ACI Fabric (To Outside Networks from Pod-1)

Complete the steps outlined in this section to deploy shared Layer 3 outside (**Shared L3Out**) connectivity to networks outside the ACI fabric from Pod-1.

Deployment Overview

In this design, the Shared L3Out connection is established in the system-defined **common** Tenant so that it can be used by all tenants in the ACI fabric. Tenants must not use overlapping addresses when connecting to the outside networks using the same shared L3Out connection. The connection uses four 10GbE interfaces between border leaf switches deployed earlier and pair of Nexus 7000 switches. The Nexus 7000 routers serve as the external gateway to the networks outside the fabric. OSPF is utilized as the routing protocol to exchange routes between the two networks. Some highlights of this connectivity are:

- Pair of Nexus 7000 routers are connected to a pair of Nexus ACI leaf switches – using a total of 4 links.
- VLANs are used for connectivity across the 4 links – using a total of 4 VLANs. VLANs are configured on separate sub-interfaces.
- Fabric Access Policies are configured on ACI Leaf switches to connect to the **External Routed** domain (via Nexus 7000s) using VLAN pool (vlans: 311–314).
- A dedicated bridge domain `common-SharedL3Out_BD` and associated dedicated VRF `common-SharedL3Out_VRF` is configured in Tenant **common** for external connectivity.
- The shared Layer 3 Out created in **common** Tenant “provides” an external connectivity contract that can be “consumed” from any tenant.
- The Nexus 7000s are configured to originate and send a default route to the Nexus 9000 leaf switches using OSPF.
- ACI leaf switches advertise tenant subnets back to Nexus 7000 switches.
- In ACI 4.0, ACI leaf switches can also advertise host-routes if it is enabled.

Create VLAN Pool for External Routed Domain

In this section, a VLAN pool is created to enable connectivity to the external networks, outside the ACI fabric. The VLANs in the pool are for the four links that connect ACI Border Leaf switches to the Nexus Gateway routers in the non-ACI portion of the customer’s network.

Setup Information

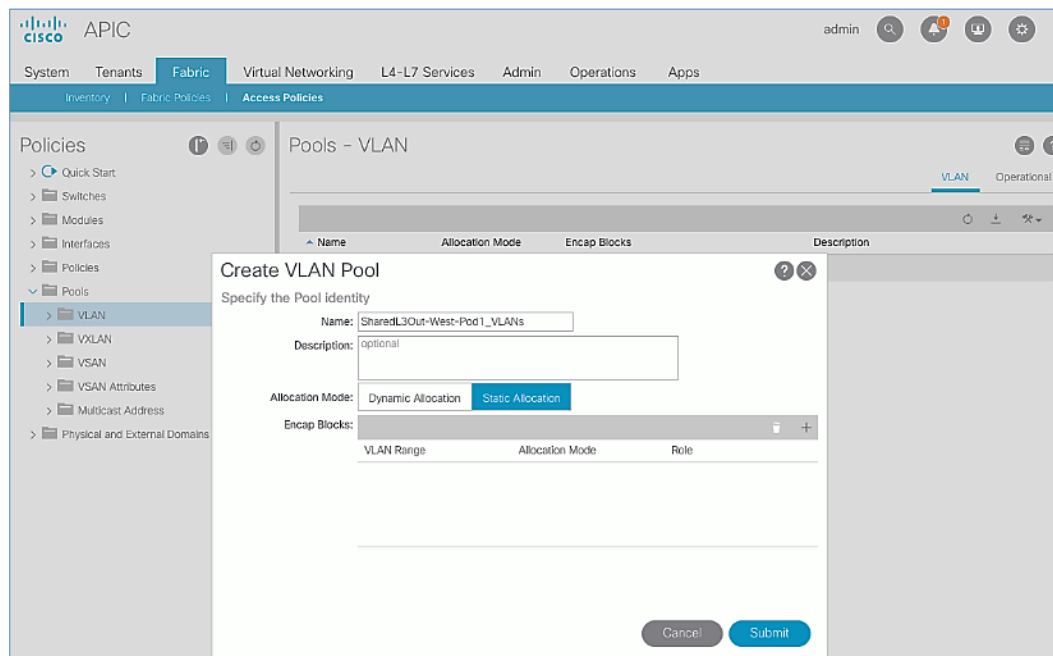
Table 6 VLAN Pool for Shared L3Out in Pod-1

To External Networks Outside ACI – Pod-1	VLAN Pool Name	Leaf Node ID	VLAN ID	Connects To
	SharedL3Out- West-Pod1_VLANS	101	311	1 st L3 Gateway Outside ACI
			312	2 nd L3 Gateway Outside ACI
		102	313	1 st L3 Gateway Outside ACI
			314	2 nd L3 Gateway Outside ACI

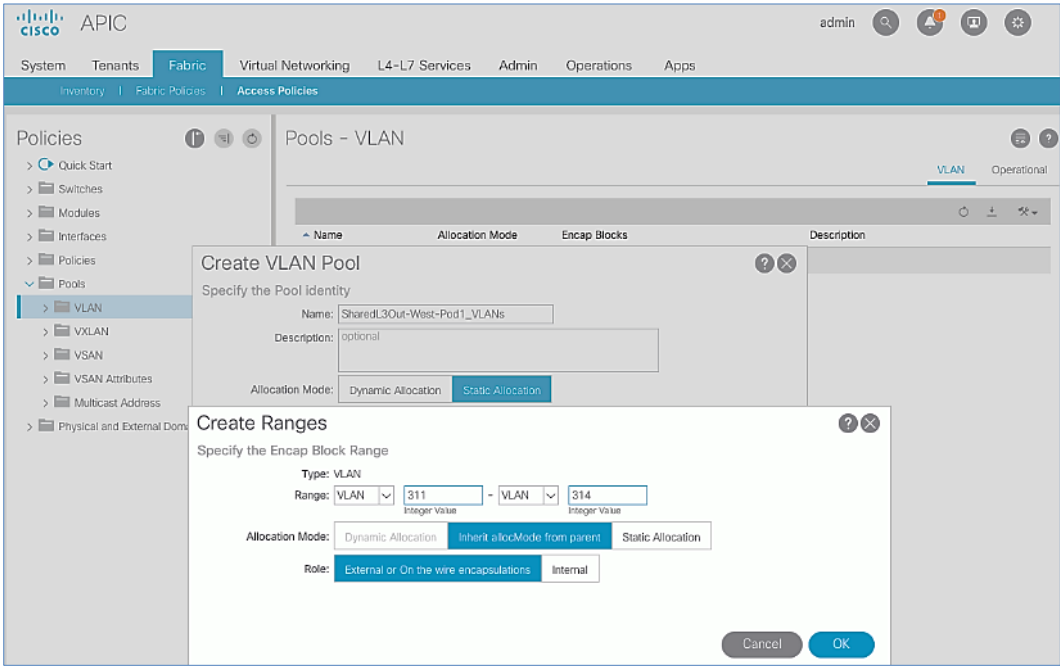
Deployment Steps

To configure a VLAN pool to connect to external gateway routers outside the ACI fabric, follow these steps:

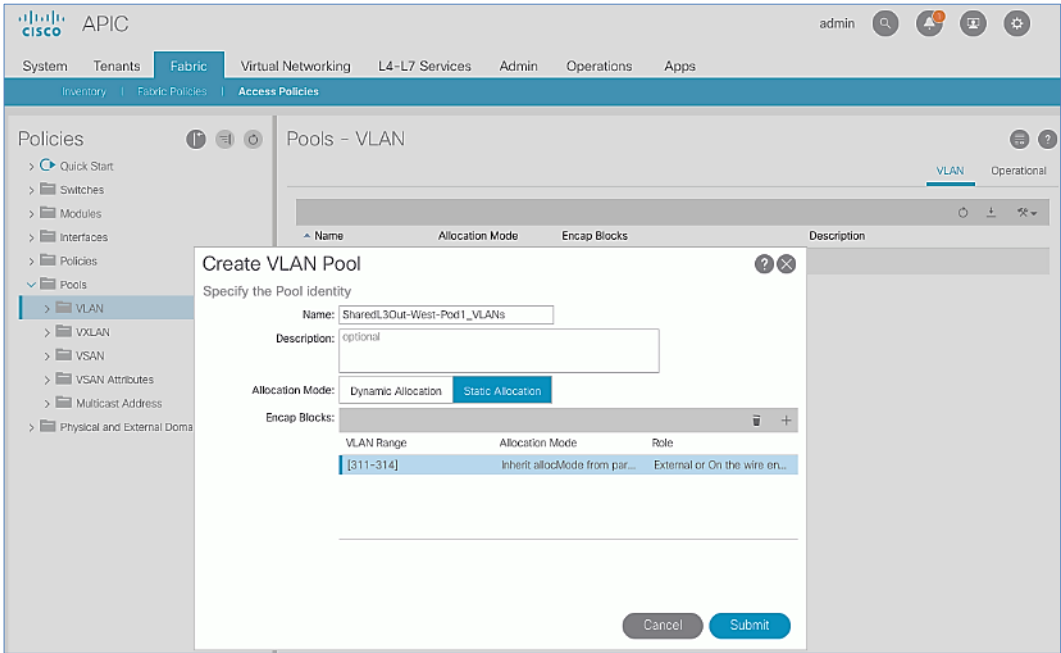
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Access Policies**.
3. From the left navigation pane, expand and select **Pools > VLAN**.
4. Right-click and select **Create VLAN Pool**.
5. In the Create VLAN Pool pop-up window, specify a Name and for Allocation Mode, select Static Allocation.



6. For **Encap Blocks**, use the **[+]** button on the right to add VLANs to the VLAN Pool. In the **Create Ranges** pop-up window, configure the VLANs that need to be configured from the Border Leaf switches to the external gateways outside the ACI fabric. Leave the remaining parameters as-is.



7. Click **OK**. Use the same VLAN ranges on the external gateway routers to connect to the ACI Fabric.



8. Click **Submit** to complete.

Configure Domain Type for External Routed Domain

To configure the domain type for the external domain, follow the procedures outlined in this section.

Setup Information

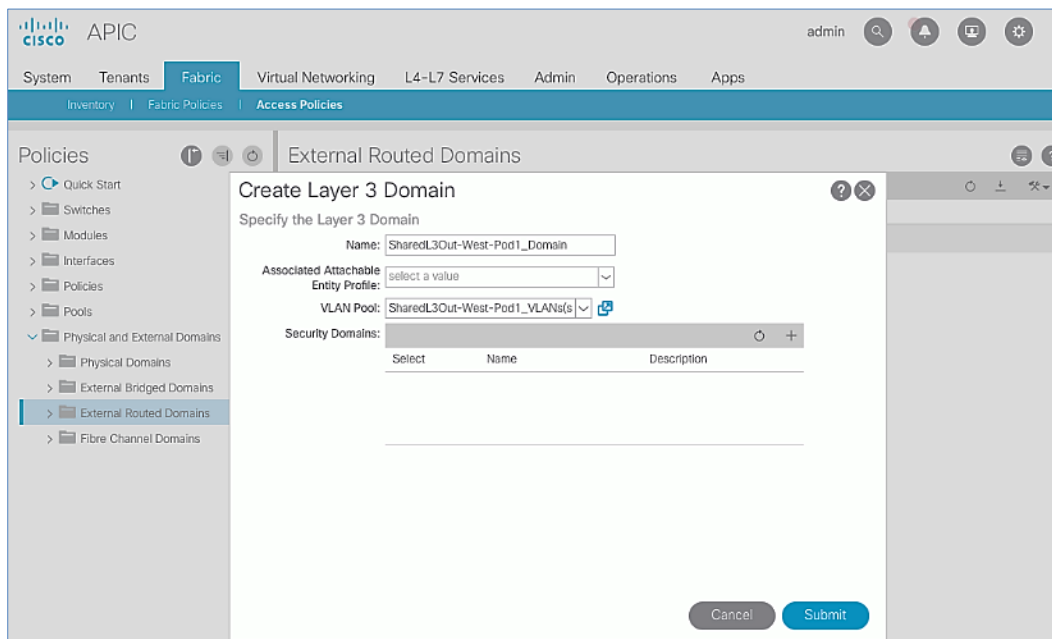
Table 7 Domain Type for Shared L3Out in Pod-1

To External Networks Outside ACI – Pod-1	Domain Name	Domain Type	VLAN Pool Name	Connects To
	SharedL3Out-West-Pod1_Domain	External Routed Domain	SharedL3Out-West-Pod1_VLANS	L3 Gateway Routers Outside ACI

Deployment Steps

To specify the domain type for connecting to external gateway routers outside the ACI fabric, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Access Policies**.
3. From the left navigation pane, expand and select Physical and External Domains > External Routed Domains.
4. Right-click External Routed Domains and select Create Layer 3 Domain.
5. In the **Create Layer 3 Domain** pop-up window, specify a **Name** for the domain. For the **VLAN Pool**, select the previously created VLAN Pool from the drop-down list.



6. Click **Submit** to complete.

Create AAEP for External Routed Domain

To configure Attachable Access Entity Profile (AAEP) for external domain, follow the procedures outlined in this section.

Setup Information

Table 8 Attachable Access Entity Profile (AAEP) for Shared L3Out in Pod-1

To External Networks Outside ACI – Pod-1	AAEP Name	Domain Name	VLAN Pool Name	Connects To
	SharedL3Out-West-Pod1_AAEP	SharedL3Out-West-Pod1_Domain	SharedL3Out-West-Pod1_VLANS	L3 Gateway Routers Outside ACI

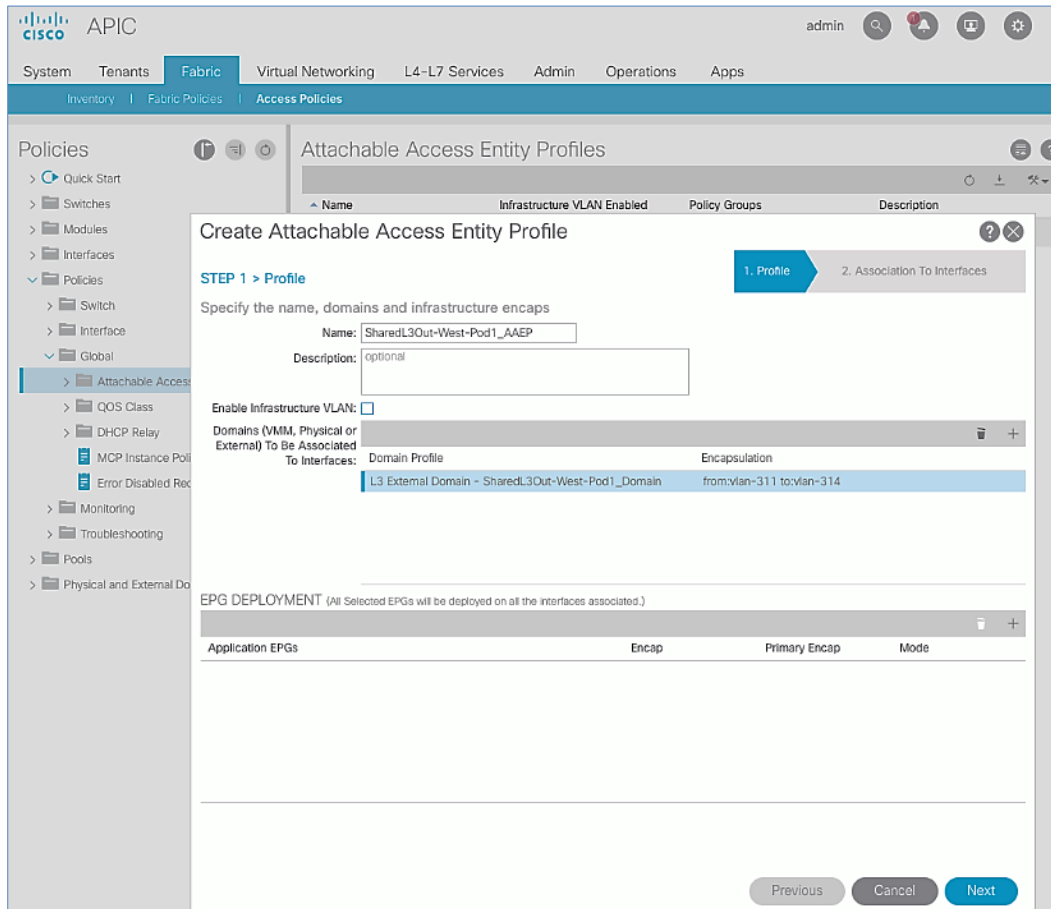
Deployment Steps

To create an Attachable Access Entity Profile (AAEP) to connect to external gateway routers outside the ACI fabric, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Access Policies**.
3. From the left navigation pane, expand and select **Policies > Global > Attachable Access Entity Profiles**.
4. Right-click and select Create Attachable Access Entity Profile.
5. In the Create Attachable Access Entity Profile pop-up window, specify a Name.
6. For the **Domains**, click the **[+]** on the right-side of the window and select the previously created domain from the drop-down list below **Domain Profile**.

The screenshot shows the APIC GUI with the 'Create Attachable Access Entity Profile' dialog box open. The dialog is titled 'Create Attachable Access Entity Profile' and has two steps: '1. Profile' and '2. Association To Interfaces'. In the '1. Profile' step, the 'Name' field is set to 'SharedL3Out-West-Pod1_AAEP'. The 'Description' field is optional. The 'Enable Infrastructure VLAN' checkbox is unchecked. The 'Domains (VMM, Physical or External) To Be Associated To Interfaces' section shows a dropdown menu with 'SharedL3Out-West-Pod1_Domain (L3)' selected. The 'Encapsulation' field is empty. The 'EPG DEPLOYMENT' section is empty. The 'Previous', 'Cancel', and 'Next' buttons are at the bottom.

7. Click **Update**.
8. You should now see the selected domain and the associated VLAN Pool as shown below.



APIC admin

System Tenants **Fabric** Virtual Networking L4-L7 Services Admin Operations Apps

Inventory Fabric Policies **Access Policies**

Policies

- Quick Start
- Switches
- Modules
- Interfaces
- Policies
 - Switch
 - Interface
 - Global
 - Attachable Access**
 - QOS Class
 - DHCP Relay
 - MCP Instance Pol
 - Error Disabled Rec
- Monitoring
- Troubleshooting
- Pools
- Physical and External Do

Attachable Access Entity Profiles

Name	Infrastructure VLAN Enabled	Policy Groups	Description
Create Attachable Access Entity Profile			

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

Specify the name, domains and infrastructure encaps

Name:

Description:

Enable Infrastructure VLAN: ☐

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

Domain Profile	Encapsulation
L3 External Domain - SharedL3Out-West-Pod1_Domain	from:vlan-311 to:vlan-314

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

Application EPGs	Encap	Primary Encap	Mode

Previous Cancel Next

9. Click **Next**. This profile is not associated with any interfaces at this time – they can be associated once the interfaces are configured in the upcoming section.
10. Click **Finish** to complete.

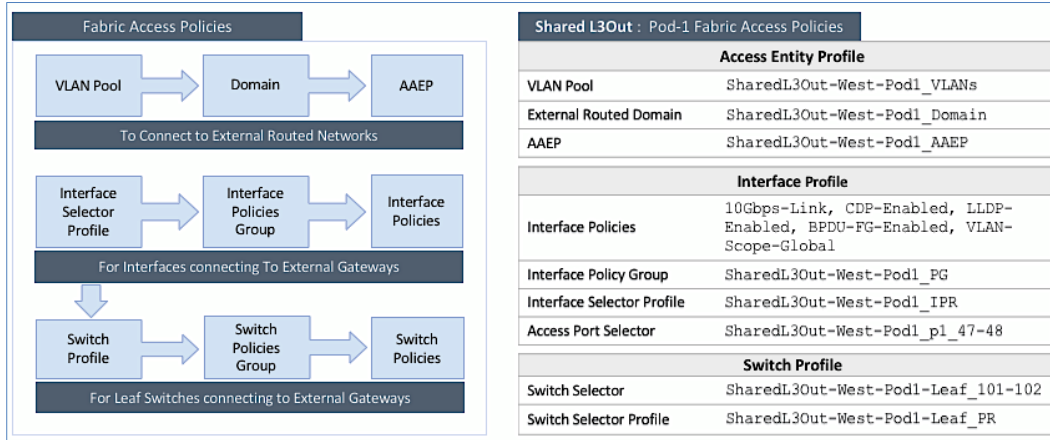
Configure Interfaces to External Routed Domain

To configure interfaces to the external routed domain, follow the procedures outlined in this section.

Setup Information

- Border Leaf switches (Node ID: 101, 102) in Pod-1 connect to External Gateways (Nexus 7000 series switches) using 10Gbps links, on ports 1/47 and 1/48.

Figure 6 Fabric Access Policies for Shared L3Out in Pod-1



Create Interface Policy Group for Interfaces to External Routed Domain

To create an interface policy group to connect to external gateway routers outside the ACI fabric, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Access Policies**.
3. From the left navigation pane, expand and select **Interfaces > Leaf Interfaces > Policy Groups > Leaf Access Port**.
4. Right-click and select **Create Leaf Access Port Policy Group**.
5. In the **Create Leaf Access Port Policy Group** pop-up window, specify a **Name** and select the applicable interface policies from the drop-down list for each field.

Create Leaf Access Port Policy Group

Specify the Policy Group Identity

Name: SharedL3Out-West-Pod1_PG

Description: optional

Link Level Policy: 10Gbps-Link

CDP Policy: CDP-Enabled

MCP Policy: select a value

CoPP Policy: select a value

LLDP Policy: LLDP-Enabled

STP Interface Policy: BPDU-FG-Enabled

Storm Control Interface Policy: select a value

L2 Interface Policy: VLAN-Scope-Global

Port Security Policy: select a value

Egress Data Plane Policing Policy: select a value

Ingress Data Plane Policing Policy: select a value

Monitoring Policy: select a value

Priority Flow Control Policy: select a value

Fibre Channel Interface Policy: select a value

PoE Interface Policy: select a value

Slow Drain Policy: select a value

MACsec Policy: select a value

802.1x Port Authentication Policy: select a value

OTDM Policy: select a value

Cancel Submit

6. For the **Attached Entity Profile**, select the previously created AAEP to external routed domain.

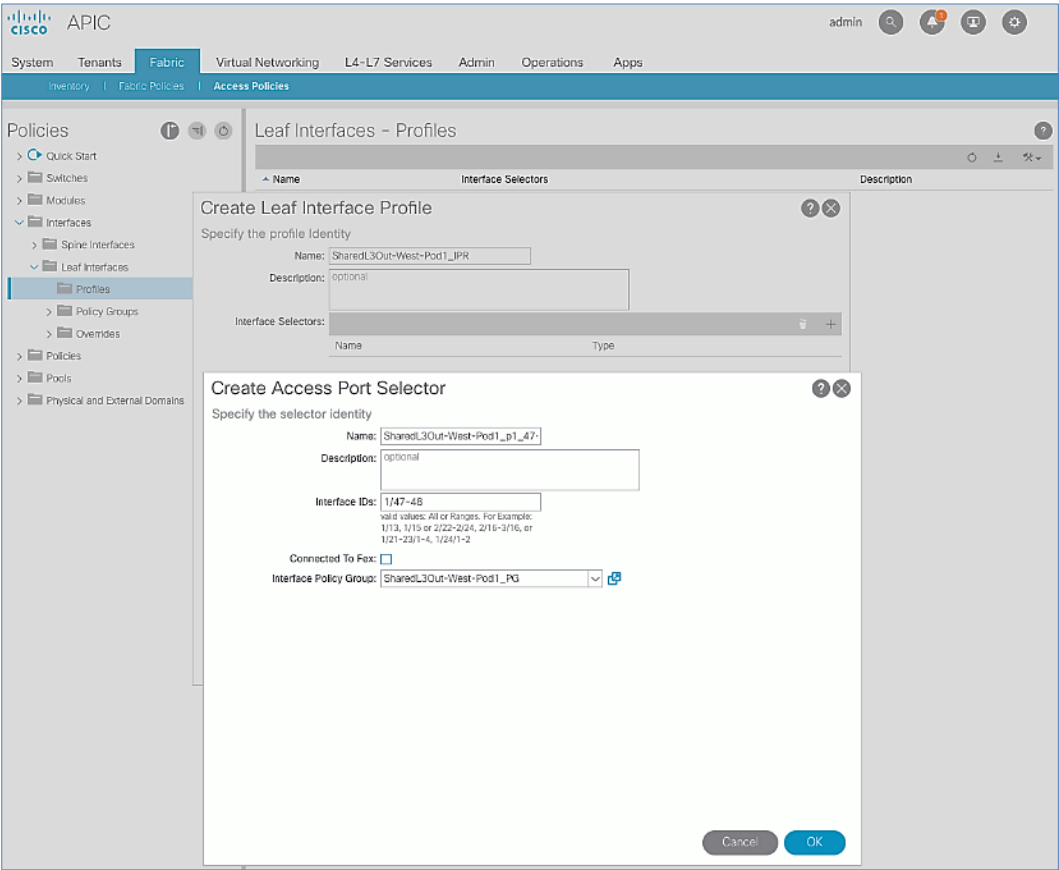
The screenshot shows the Cisco APIC GUI with the 'Create Leaf Access Port Policy Group' dialog open. The dialog is titled 'Policy Groups – Leaf Access Port'. It features a table with columns: Name, Link Level Policy, CDP Policy, LLDP Policy, STP Interface Policy, and Monitoring Policy. Below the table, there are several sections: 'Specify the Policy Group Identity' with dropdowns for Ingress Data Plane Policing Policy, Monitoring Policy, Priority Flow Control Policy, Fibre Channel Interface Policy, PoE Interface Policy, Slow Drain Policy, MACsec Policy, 802.1x Port Authentication Policy, and DWDOM Policy. The 'Attached Entity Profile' is set to 'SharedL3Out-West-Po'. Below this is a 'Connectivity Filters' section with a table for 'Switch IDs' and 'Interfaces'. At the bottom, there is a 'NetFlow Monitor Policies' section with a table for 'NetFlow IP Filter Type' and 'NetFlow Monitor Policy'. The dialog has 'Cancel' and 'Submit' buttons at the bottom right.

7. Click **Submit** to complete.

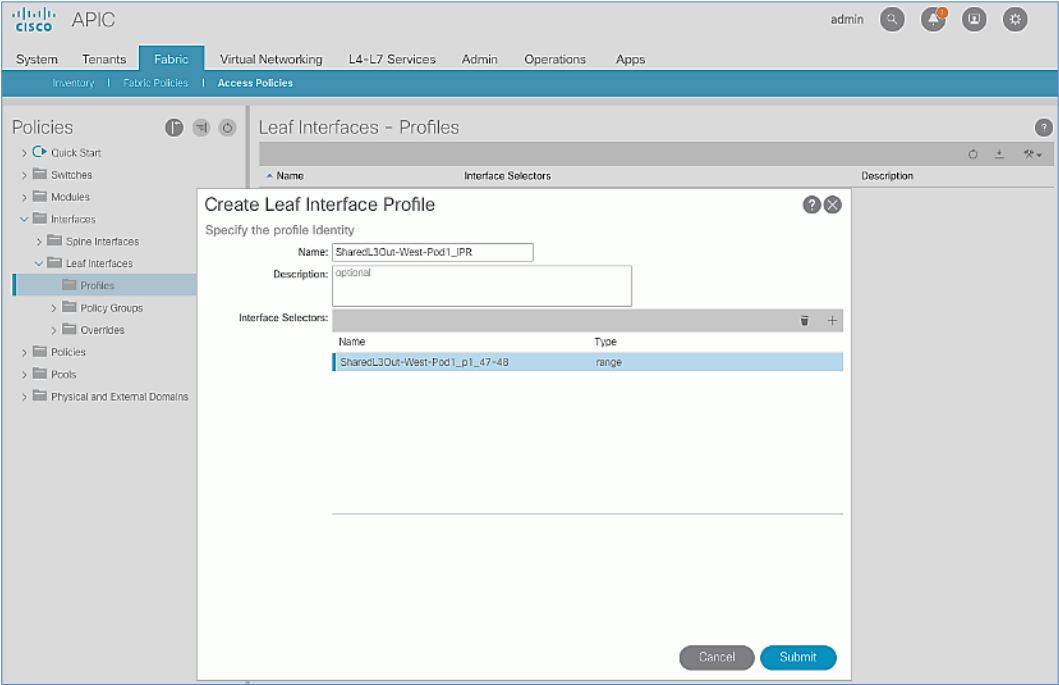
Create Interface Profile for Interfaces to External Routed Domain

To create an interface profile to connect to external gateway routers outside the ACI fabric, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Access Policies**.
3. From the left navigation menu, expand and select **Interfaces > Leaf Interfaces > Profiles**.
4. Right-click and select **Create Leaf Interface Profile**.
5. In the **Create Leaf Interface Profile** pop-up window, specify a **Name**. For **Interface Selectors**, click the **[+]** to select access ports to apply interface policies to. In this case, the interfaces are access ports that connect Border Leaf switches to gateways outside ACI.
6. In the **Create Access Port Selector** pop-up window, specify a selector **Name**. For the **Interface IDs**, specify the access ports connecting to the two external gateways. For the **Interface Policy Group**, select the previously created Policy Group from the drop-down list.



7. Click **OK** to close the **Create Access Port Selector** pop-up window.



8. Click **Submit** to complete.

Create Leaf Switch Profile to External Routed Domain

To create leaf switch profile to connect to external gateway routers outside the ACI fabric, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Access Policies**.
3. From the left navigation menu, expand and select **Switches > Leaf Switches > Profiles**.
4. Right-click and select **Create Leaf Profile**.
5. In the **Create Leaf Profile** pop-up window, specify a profile **Name**. For **Leaf Selectors**, click the **[+]** to select the Leaf switches to apply the policies to. In this case, the Leaf switches are the Border Leaf switches that connect to the gateways outside ACI.
6. Specify a Leaf Selector **Name**. For the **Interface IDs**, specify the access ports connecting to the two external gateways. For **Blocks**, select the Node IDs of the Border Leaf switches from the drop-down list.

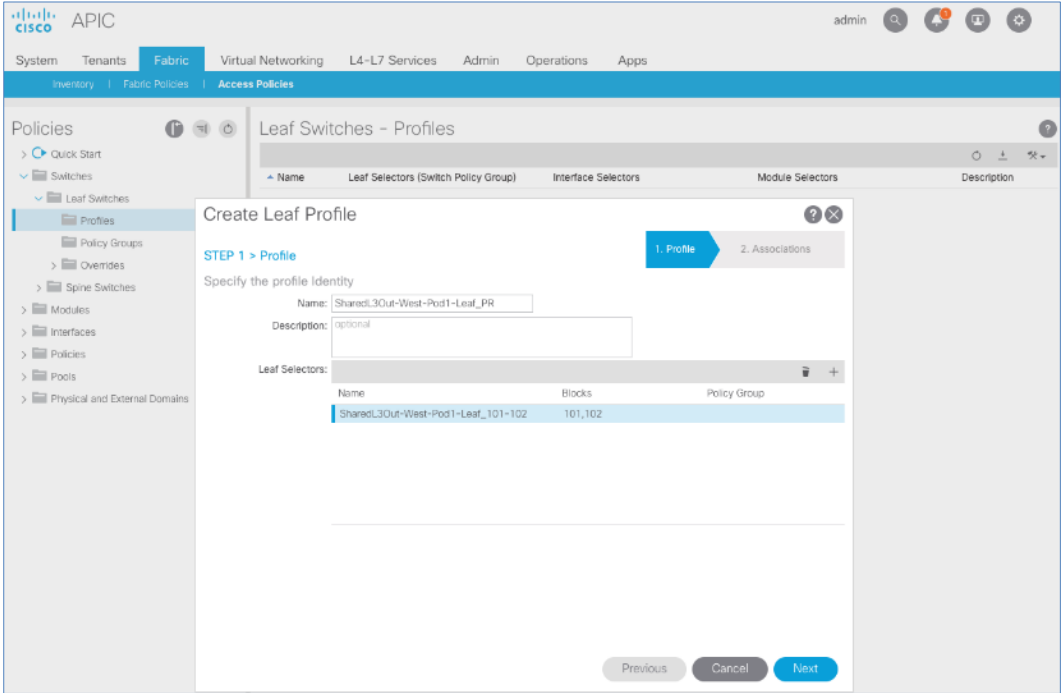
The screenshot shows the Cisco APIC GUI with the 'Create Leaf Profile' dialog box open. The dialog is titled 'Create Leaf Profile' and has two steps: '1. Profile' and '2. Associations'. The current step is '1. Profile'. The 'Specify the profile Identity' section contains the following fields:

- Name:** SharedL3Out-West-Pod1-Leaf_PR
- Description:** optional
- Leaf Selectors:** A table with one entry:

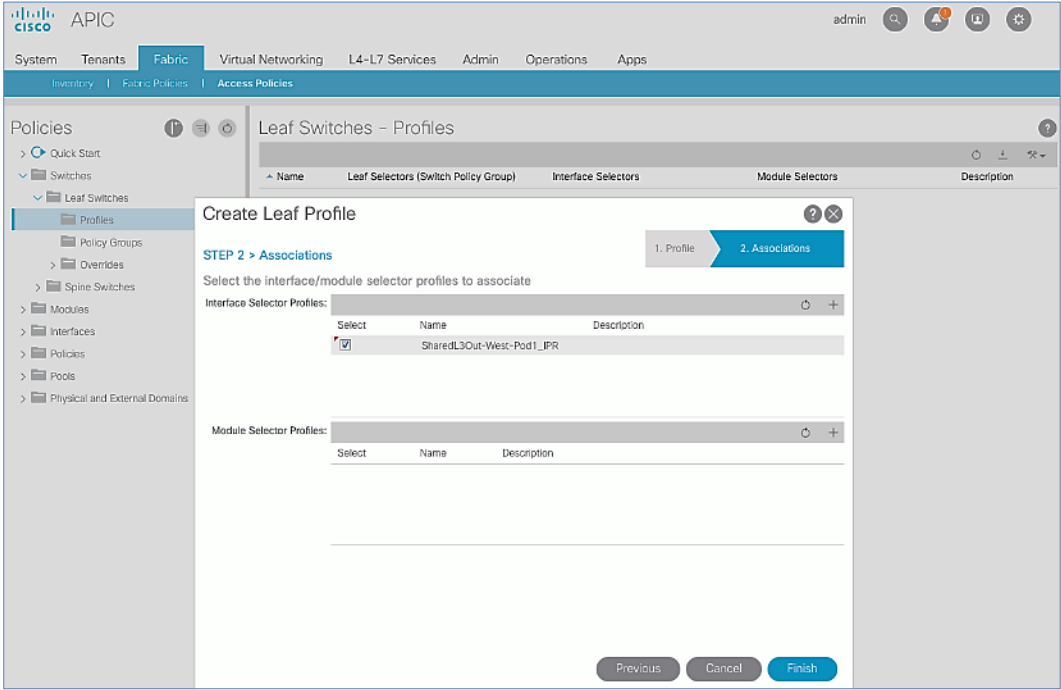
Name	Blocks	Policy Group
SharedL3Out-West-Pod1-Leaf_101-102	101-102	select an option

At the bottom of the dialog, there are buttons for 'Update', 'Cancel', 'Previous', and 'Next'.

7. Click **Update**.



- Click **Next**.
- In the **Associations** window, select the previously created **Interface Selector Profiles** from the list.



- Click **Finish** to complete.

Configure Tenant Networking for Shared L3Out

To configure tenant networking to connect to networks outside the ACI fabric, follow the procedures outlined in this section.

Setup Information

Figure 7 Tenant Networking for Shared L3Out

Shared L3Out	Tenant Name	VRF	Bridge Domain
	common	common-SharedL3Out_VRF	common-SharedL3Out_BD

Deployment Steps

To configure tenant networking for the Shared L3Out for connectivity outside the ACI fabric, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Tenants > common**.
3. From the left navigation pane, select and expand **Tenant common > Networking > VRFs**.
4. Right-click and select **Create VRF**.
5. In the **Create VRF** pop-up window, **STEP 1 > VRF**, specify a **Name** (for example, `common-SharedL3Out_VRF`).
6. Check the box for Create a Bridge Domain.

The screenshot shows the APIC GUI with the 'Create VRF' pop-up window open. The window is titled 'Create VRF' and has two steps: '1. VRF' and '2. Bridge Domain'. The '1. VRF' step is active. The form fields include: Name (common-SharedL3Out_VRF), Alias (empty), Description (optional), Tags (empty), Policy Control Enforcement Preference (Enforced), Policy Control Enforcement Direction (Egress), BD Enforcement Status (unchecked), Endpoint Retention Policy (select a value), Monitoring Policy (select a value), DNS Labels (empty), Route Tag Policy (select a value), IP Data-plane Learning (Disabled), Create A Bridge Domain (checked), Configure BGP Policies (unchecked), and Configure OSPF Policies (unchecked). The 'Next' button is highlighted.

7. Click **Next**.
8. In the **Create VRF** pop-up window, **STEP 2 > Bridge Domain**, specify a **Name** (for example, `common-SharedL3Out_BD`).

- ## Configure External Routed Networks under Tenant Common

Setup Information

Shared L3Out - Pod1	Routed Outside Name	Routed Node Profile	Router IDs	Node IDs	Node Interface Profile	OSPF Policy
	SharedL3Out-West-Pod1_RO OSPF Area 10	SharedL3Out-West-Pod1-Node_PR	13.13.13.1/32	101	SharedL3Out-West-Pod1-Node_IPR	SharedL3Out-West-Pod1-OSPF_Policy
			13.13.13.2/32	102		✓ Point-to-point ✓ MTU ignore)
	Routed Sub-interface	VLAN	Subnet	External Network		
	Eth1/47	311	10.113.1.0/30	Default-Route (0.0.0.0/0)		
	Eth1/48	312	10.113.1.4/30	✓ External Subnets for the External EPG		
	Eth1/47	313	10.113.2.0/30	✓ Shared Route Control Subnet		
	Eth1/48	314	10.113.2.4/30	✓ Shared Security Import Subnet		

To configure the external routed networks under Tenant common, follow these steps:

- 67

4. Right-click and select **Create Routed Outside**.
5. In the **Create Routed Outside** pop-up window, specify a **Name**.
6. Select the check box next to **OSPF**.
7. For the **OSPF Area ID**, enter 0.0.0.10 (should match the external gateway configuration).
8. For the **VRF**, select the previously created VRF from the drop-down list.
9. For the **External Routed Domain**, select the previously created domain from the drop-down list.

The screenshot shows the 'Create Routed Outside' configuration window in the Cisco APIC. The window is divided into two main sections: 'Define the Routed Outside' and 'Nodes and Interfaces Protocol Profiles'.

Define the Routed Outside:

- Aliases:** optional
- Tags:** enter tags separated by comma
- PIM:** ☐ PIM
- Route Control Enforcement:** ☐ Import, ☒ Export
- Target DSCP:** Unspecified
- VRF:** common-Shared_L3Out_VRF
- External Routed Domain:** Shared_L3Out-West-Pod1_Done
- Route Profile for Interleaf:** select a value
- Route Control For Dampening:** Address Family Type, Route Dampening Policy
- Consumer Label:** enter names separated by comma
- OSPF Area ID:** 0.0.0.10
- OSPF Area Control:** ☒ Send redistributed LSAs into NSSA area, ☒ Originate summary LSA, ☐ Suppress forwarding address in translated LSA
- OSPF Area Type:** NSSA area, Regular area, Stub area
- OSPF Area Cost:** 1

Nodes and Interfaces Protocol Profiles:

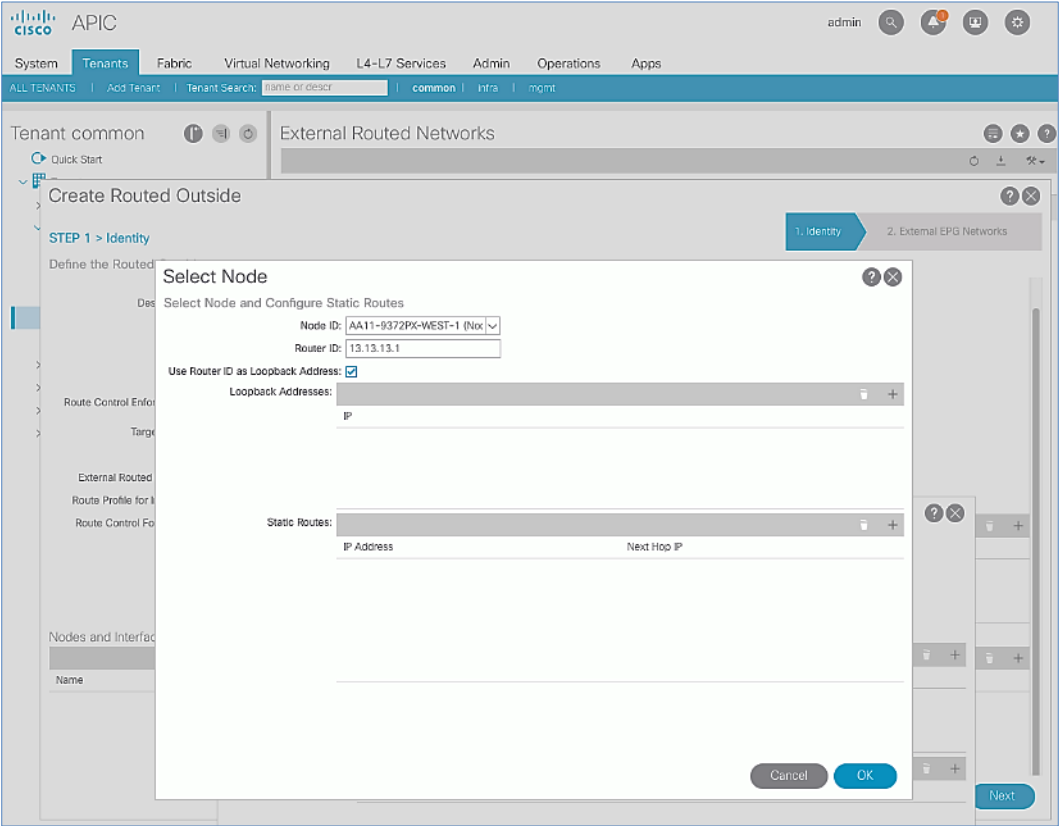
Name	Description	DSCP	Nodes
[Empty table body]			

Buttons at the bottom: Previous, Cancel, Next.

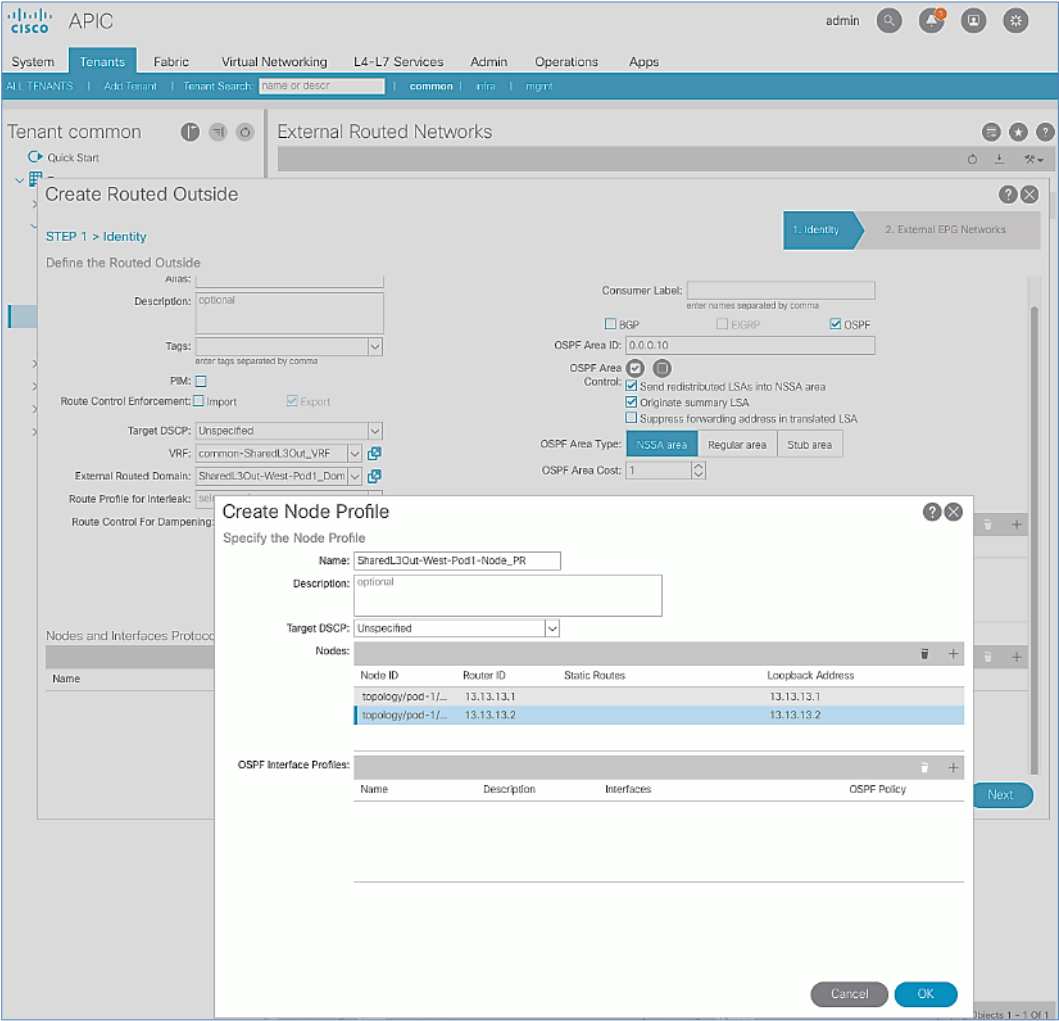
10. For **Nodes and Interfaces Protocol Profiles**, click [+] to add a Node Profile.
11. In the **Create Node Profile** pop-up window, specify a profile **Name**.

The screenshot shows the Cisco APIC interface for configuring a 'Create Routed Outside' network. The main configuration area is titled 'Create Routed Outside' and is in the 'Identity' step. It includes fields for 'ASAS', 'Description', 'Tags', 'PIM', 'Route Control Enforcement' (Import/Export), 'Target DSCP', 'VRF', 'External Routed Domain', 'Route Profile for Interface', and 'Route Control For Dampening'. A 'Create Node Profile' pop-up window is open, showing fields for 'Name', 'Description', 'Target DSCP', and a table for 'Nodes' with columns for 'Node ID', 'Router ID', 'Static Routes', and 'Loopback Address'. The pop-up also has a table for 'OSPF Interface Profiles' with columns for 'Name', 'Description', 'Interfaces', and 'OSPF Policy'.

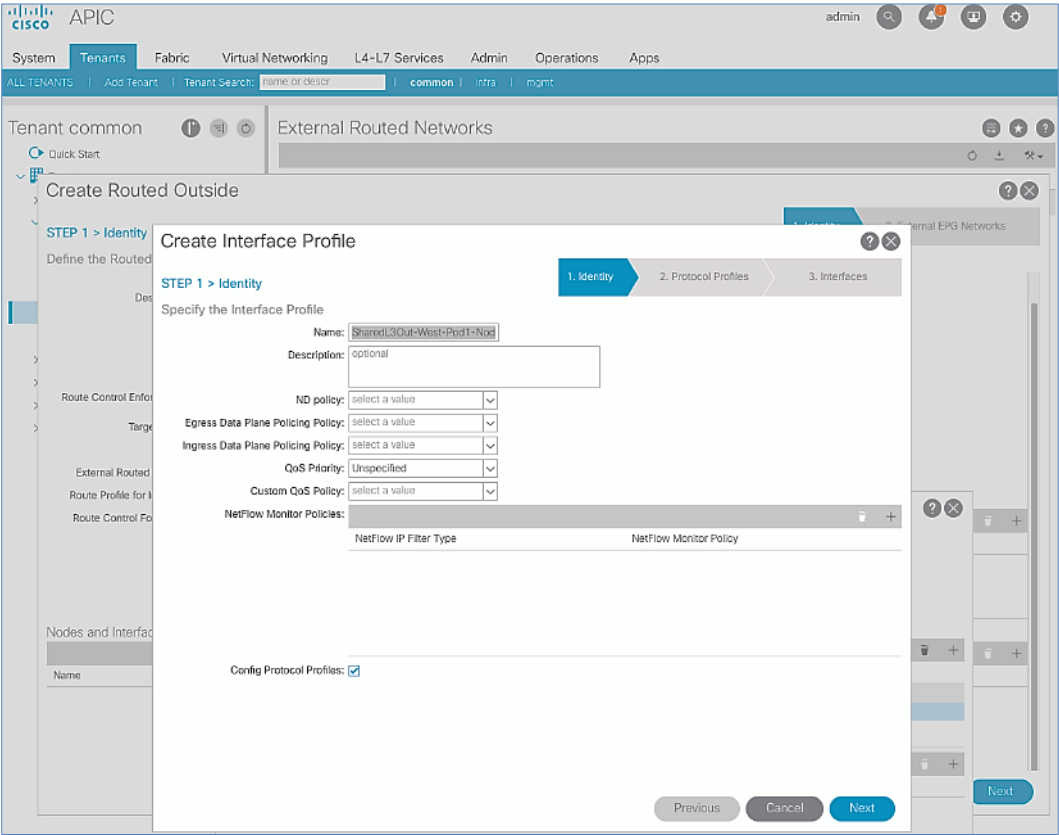
12. For **Nodes**, click **[+]** to add a Node.
13. In the **Select Node** pop-up window, for the **Node ID**, select first Border Leaf switch from the drop-down list. For the **Router ID**, specify the router ID for the first Border Leaf Switch (for example, 13.13.13.1).



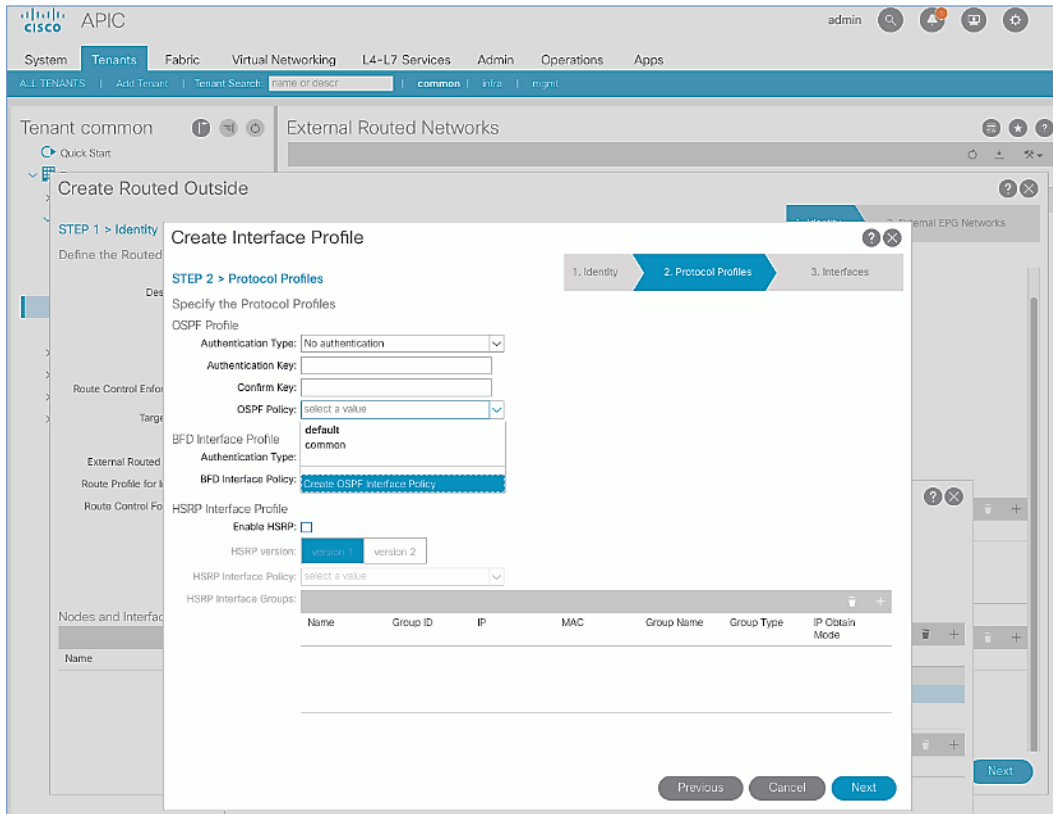
- 14. Click **OK** to complete selecting the Node.
- 15. Repeat steps 1-14 to add the second Border Leaf to the list of Nodes.



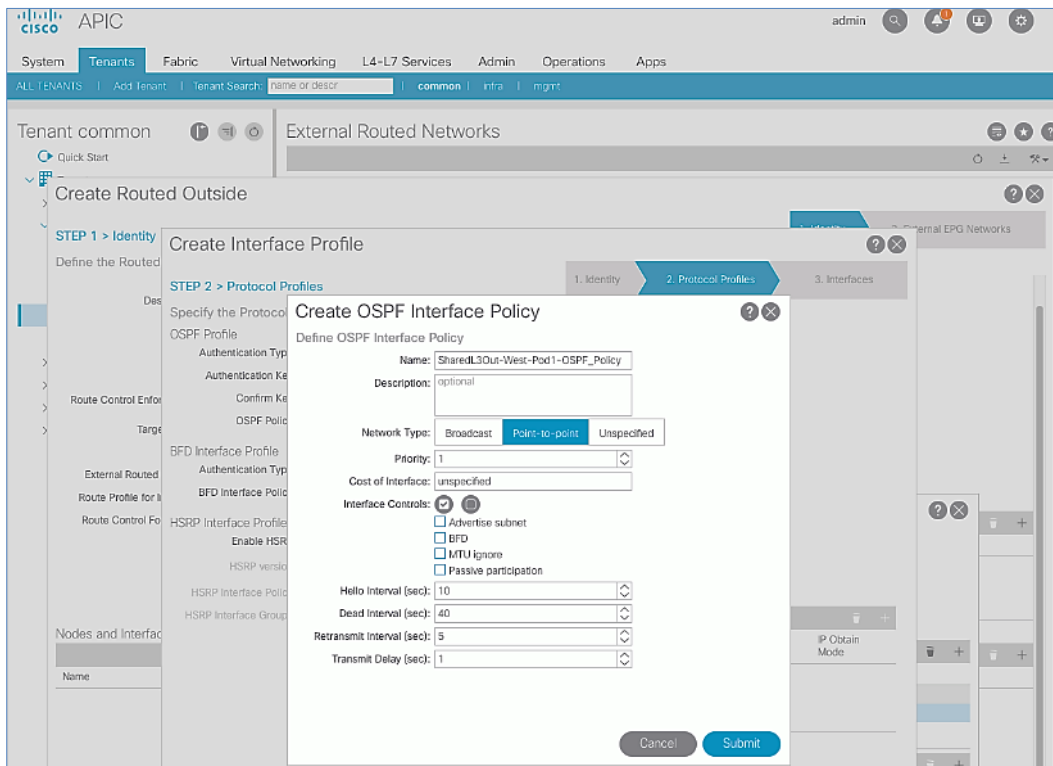
- 16. For **OSPF Interface Profiles**, click **[+]** to add a profile.
- 17. In the **Create Interface Profile** pop-up window, for **Step 1 > Identity**, specify a Name.



- 18. Click **Next**.
- 19. In Step 2 > Protocol Profiles, for the OSPF Policy, use the drop-down list to select Create OSPF Interface Policy.



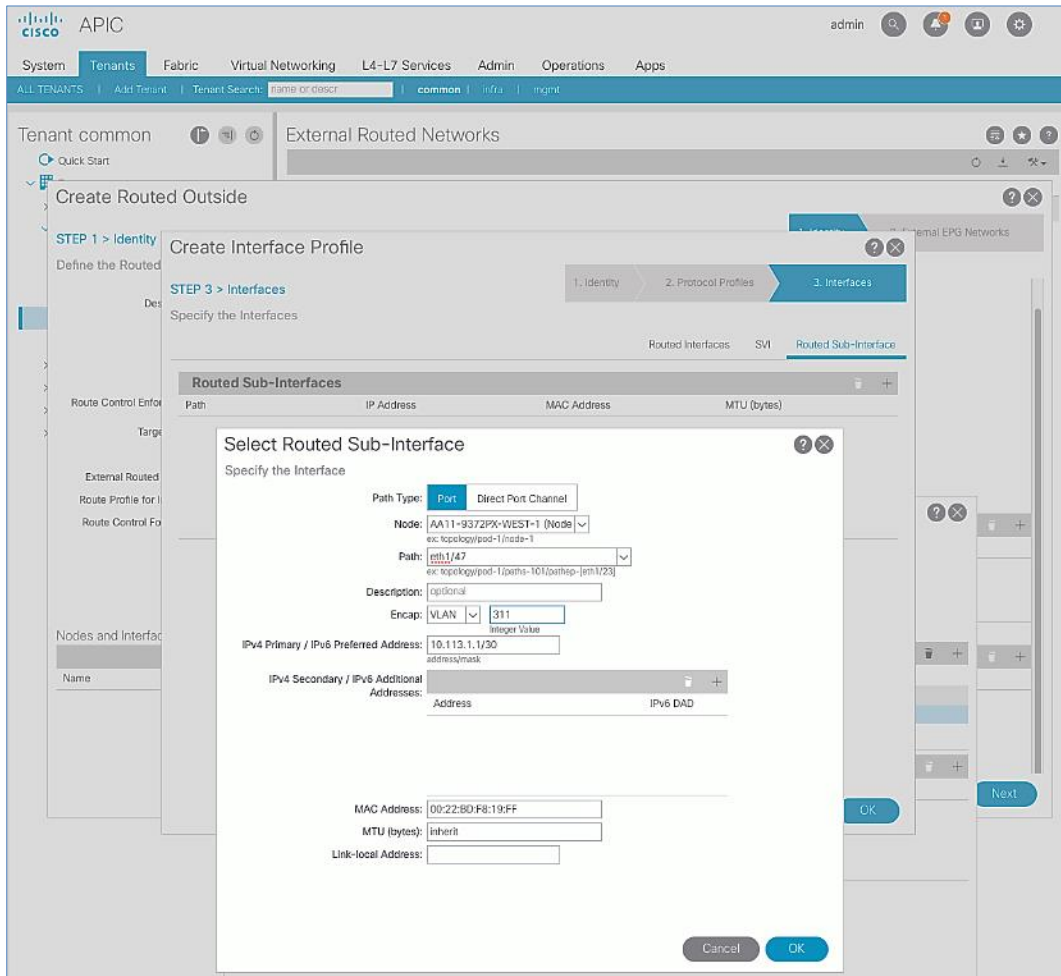
20. In the Create OSPF Interface Policy pop-up window, specify a Name. For Network Type, select Point-to-Point. For Interface Controls, select the checkbox for MTU ignore.



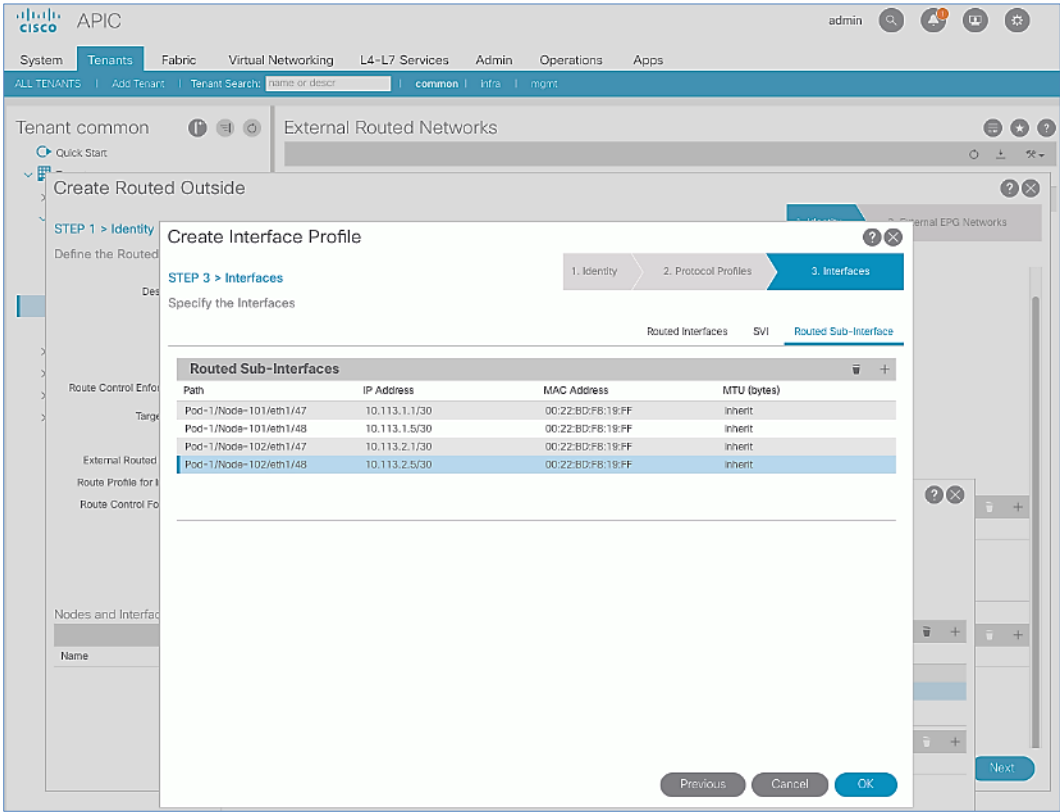
21. Click **Submit** to complete creating the OSPF policy.
22. In the **Create Interface Profile** pop-up window, click **Next**.

The screenshot shows the Cisco APIC interface with the 'Create Interface Profile' pop-up window open. The window is titled 'Create Interface Profile' and shows 'STEP 2 > Protocol Profiles'. It has three tabs: '1. Identity', '2. Protocol Profiles' (active), and '3. Interfaces'. The '2. Protocol Profiles' tab contains sections for 'OSPF Profile', 'BFD Interface Profile', and 'HSRP Interface Profile'. The 'OSPF Profile' section has fields for 'Authentication Type' (No authentication), 'Authentication Key', 'Confirm Key', and 'OSPF Policy' (SharedL3Out-West-Pod1-OSPF_Pv). The 'BFD Interface Profile' section has fields for 'Authentication Type' (No authentication) and 'BFD Interface Policy' (select a value). The 'HSRP Interface Profile' section has a checkbox for 'Enable HSRP', 'HSRP version' (version 1 selected), 'HSRP Interface Policy' (select a value), and 'HSRP Interface Groups' (a table with columns: Name, Group ID, IP, MAC, Group Name, Group Type, IP Obtain Mode). At the bottom are 'Previous', 'Cancel', and 'Next' buttons.

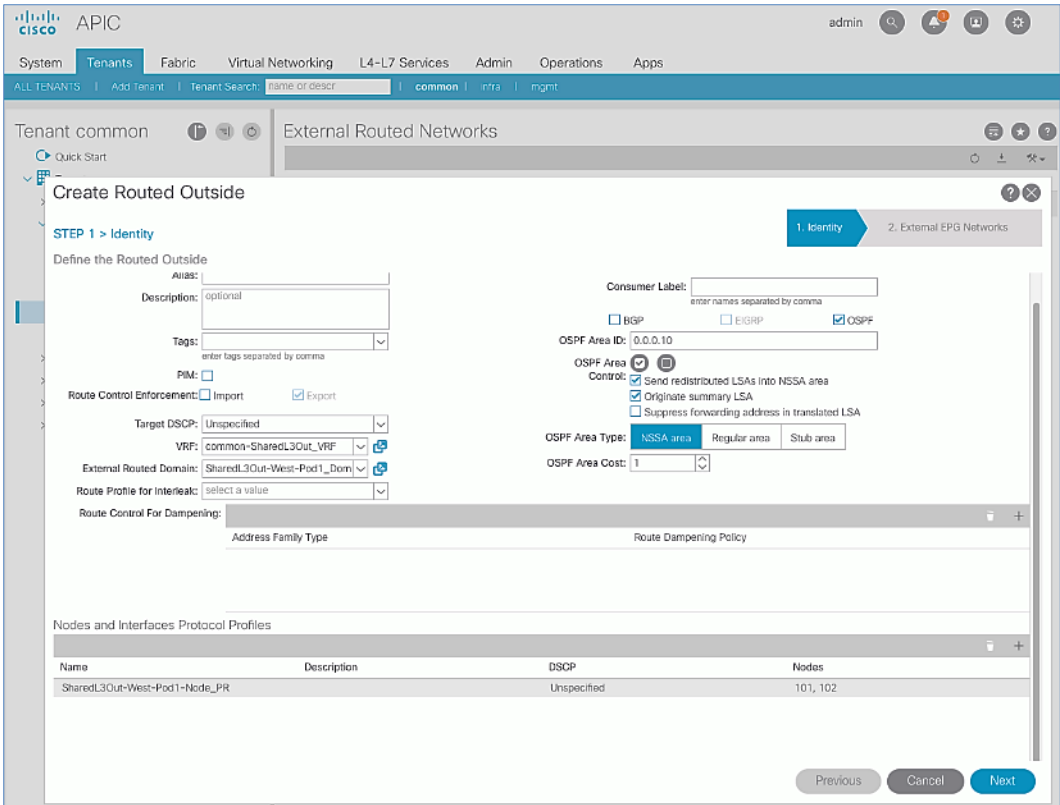
23. For **STEP 3 > Interfaces**, select the tab for **Routed Sub-Interface**. Click **[+]** on the right side of the window to add a routed sub-interface.
24. In the **Select Routed Sub-Interface** pop-up window, for **Node**, select the first Border Leaf. For **Path**, select the interface (for example, 1 / 47) on the first Border Leaf that connects to the first external gateway. For **Encap**, specify the VLAN (for example, 311). For **IPv4 Primary / IPv6 Preferred Address**, specify the address (for example, 10.113.1.1/30).



25. Click **OK** to complete configuring the first routed sub-interface.
26. Repeat steps 1-25 to create the next sub-interface that connects the first Leaf to the second Gateway.
27. Repeat steps 1-25 to create the sub-interfaces on the second Leaf that connects to the two gateways.



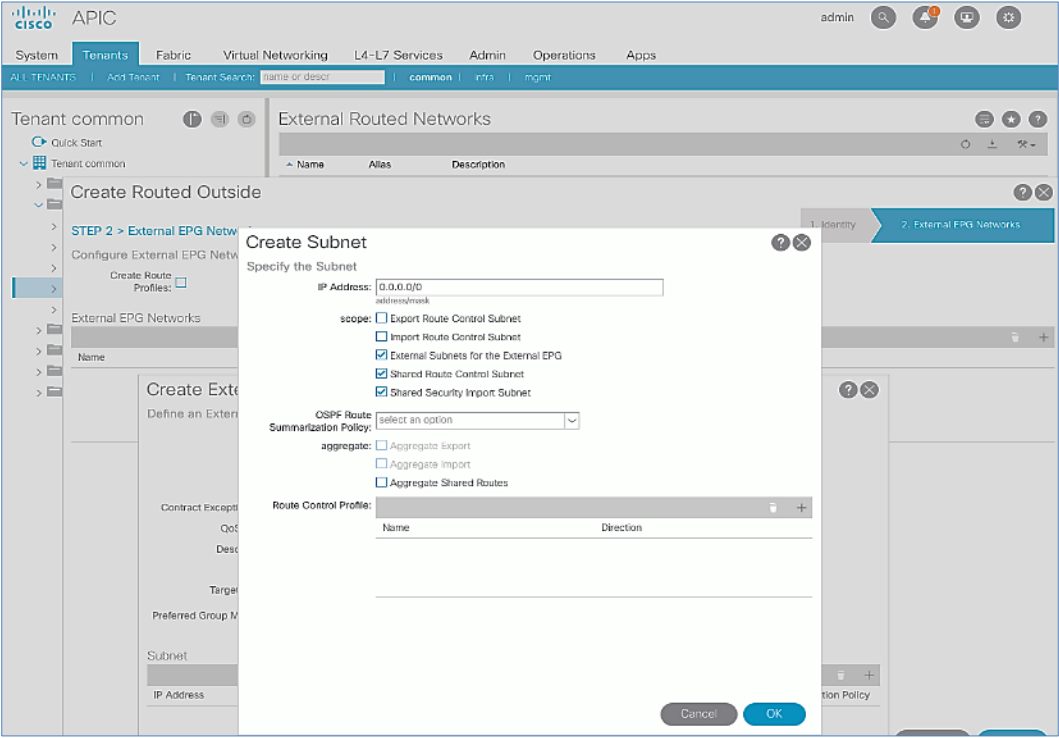
28. Click **OK** to complete creating the Interface Profile.
29. In the **Create Routed Outside** pop-up window, click **Next**.



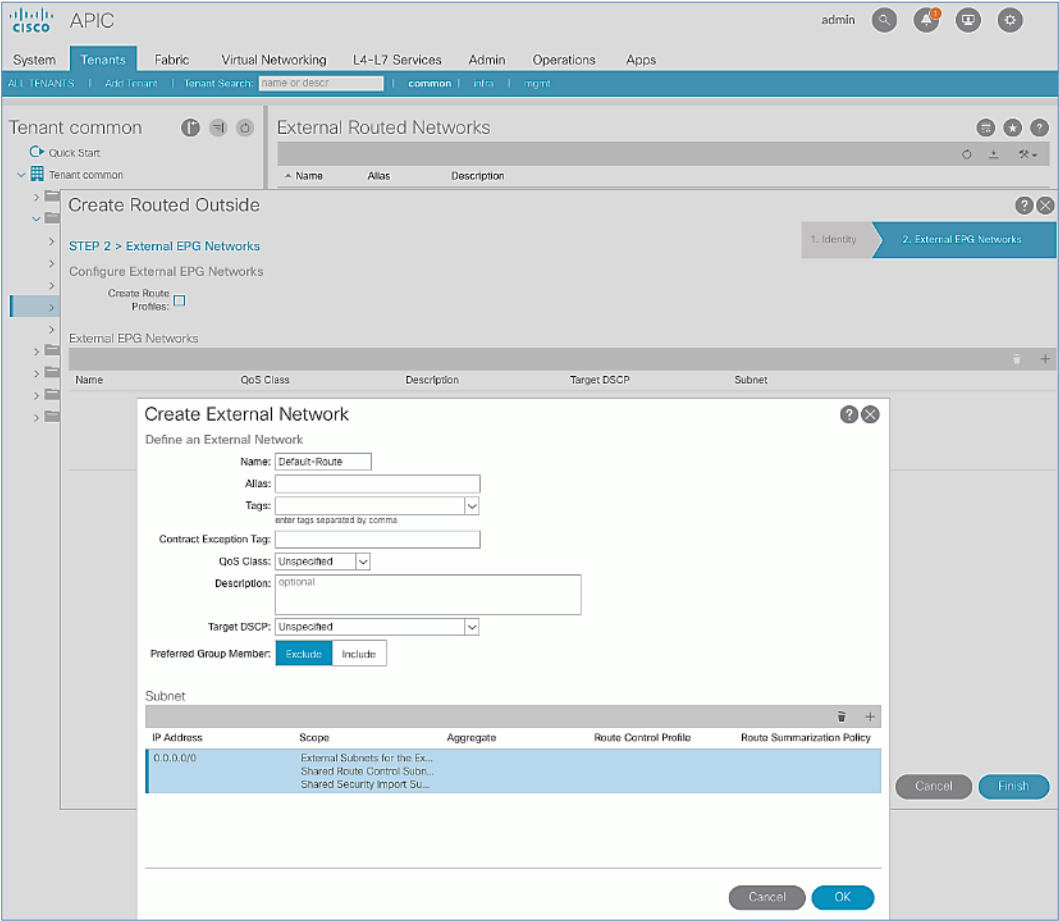
30. In STEP 2 > External EPG Networks, for External EPG Networks, click [+] to add an external network.
31. In the **Created External Network** pop-up window, specify a **Name** (for example, Default-Route).
32. For **Subnet**, click [+] to add a Subnet.

The screenshot displays the Cisco APIC interface. The top navigation bar includes 'System', 'Tenants', 'Fabric', 'Virtual Networking', 'L4-L7 Services', 'Admin', 'Operations', and 'Apps'. The 'Tenants' tab is active, showing a list of tenants with 'common' selected. The left sidebar shows the 'Create Routed Outside' section with 'STEP 2 > External EPG Networks' highlighted. The main content area shows the 'External Routed Networks' table. A 'Create External Network' pop-up window is open, allowing the user to define an external network. The 'Name' field is set to 'Default-Route'. The 'Subnet' table is empty, and the user is prompted to add a subnet by clicking the '+' button. The 'Preferred Group Member' section has 'Exclude' selected. The 'Cancel' and 'OK' buttons are at the bottom of the pop-up window.

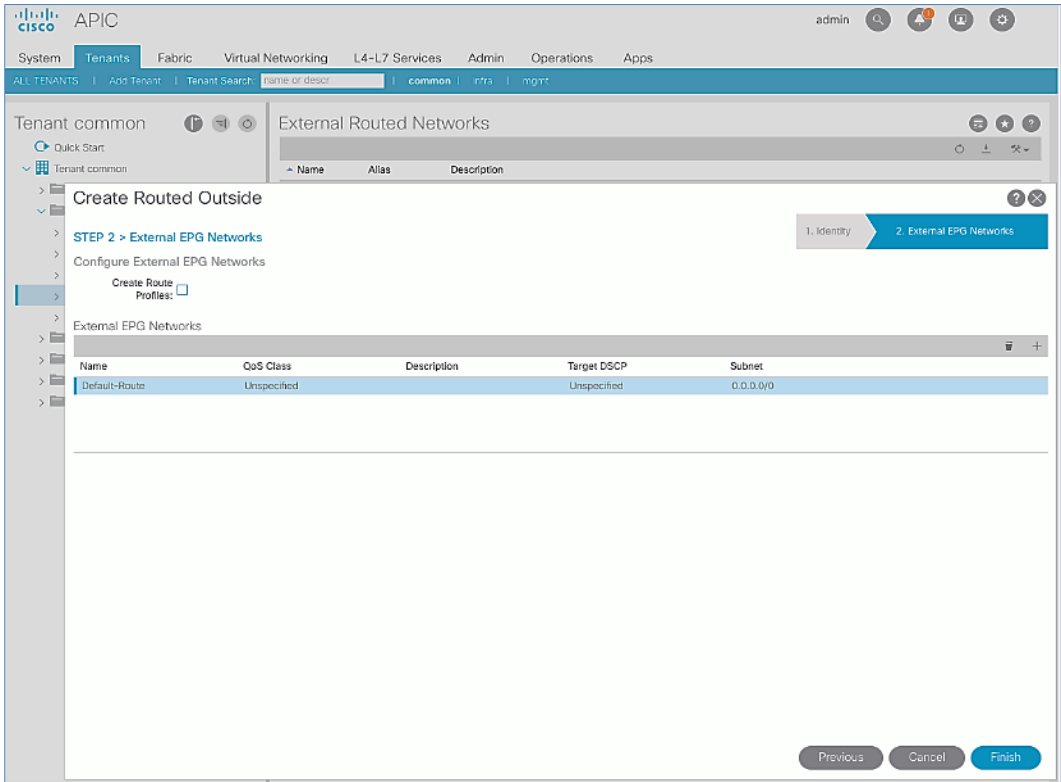
33. In the Create Subnet pop-up window, for the IP Address, enter a route (for example, 0.0.0.0/0). Select the checkboxes for External Subnets for the External EPG, Shared Route Control Subnet, and Shared Security Import Subnet.



34. Click **OK** to complete creating the subnet.



35. Click **OK** again to complete creating the external network.



36. Click **Finish** to complete creating the Routed Outside.

Create Contracts for External Routed Networks from Tenant (common)

To create contracts to access external routed networks, follow the procedures outlined in this section.

Setup Information

Table 10 Contract Created

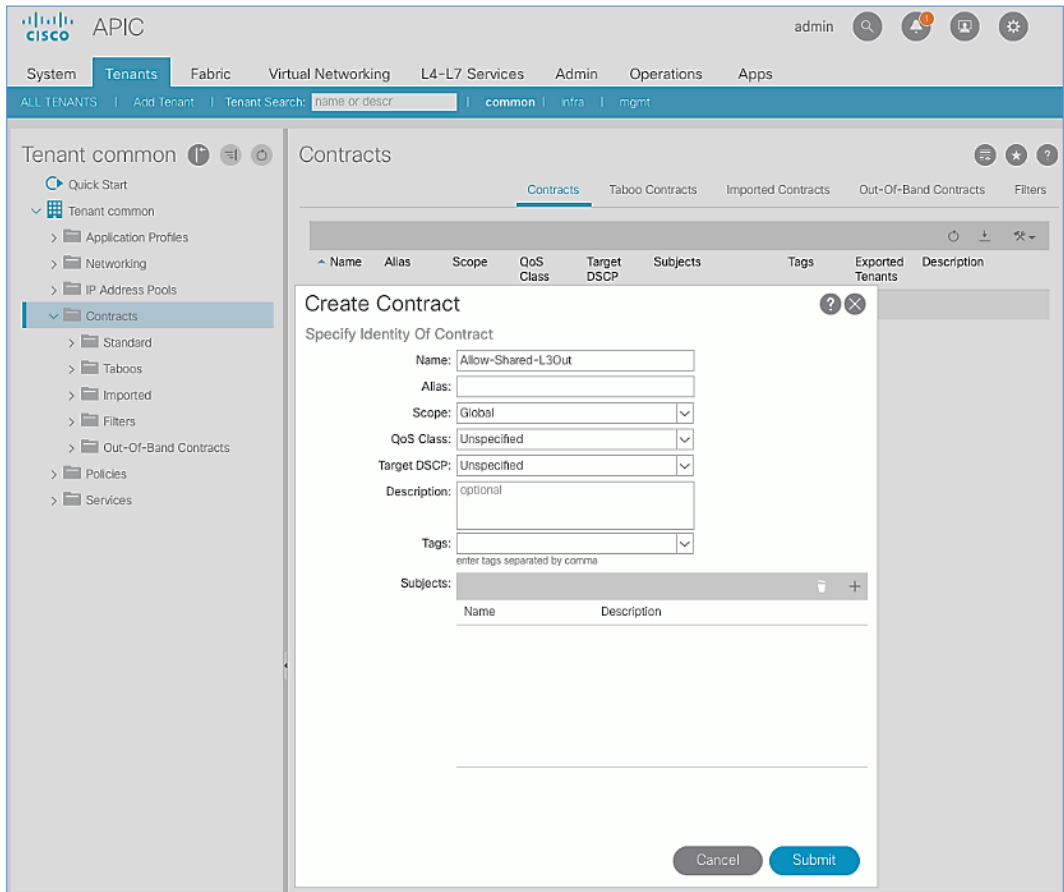
Shared L3Out	Contract	Subject	Filter
	Allow-Shared-L3Out	Allow-Shared-L3Out	common/default ✓ Global Scope

Deployment Steps

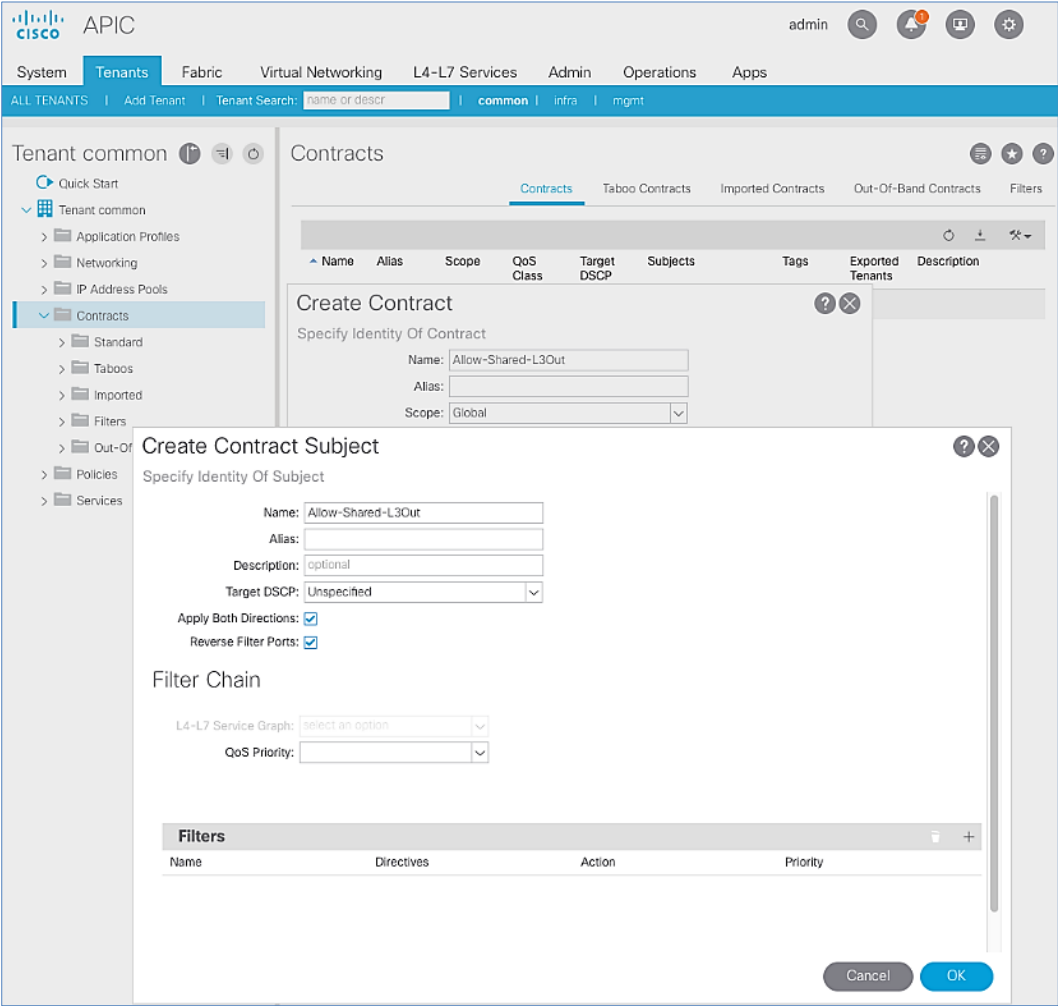
To create contracts for external routed networks from Tenant common, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Tenants > common**.
3. In the left navigation pane, select and expand **Tenant common > Contracts**.
4. Right-click Contracts and select Create Contract.
5. In the **Create Contract** pop-up window, specify a **Name**.

- For **Scope**, select **Global** from the drop-down list to allow the contract to be consumed by all tenants.
- For **Subjects**, click **[+]** on the right side to add a contract subject.



- In the **Create Contract Subject** pop-up window, specify a **Name**.
- For **Filters**, click **[+]** on the right side to add a filter.



10. In the **Filters** section of the window, for **Name**, select default (**common**) from the drop-down list to create a **default** filter for Tenant **common**.

The screenshot shows the Cisco APIC interface with the 'Create Contract Subject' dialog box open. The dialog is titled 'Create Contract Subject' and has a subtitle 'Specify Identity Of Subject'. It contains the following fields and options:

- Name:** Allow-Shared-L3Out
- Alias:** (empty)
- Description:** optional
- Target DSCP:** Unspecified
- Apply Both Directions:** ☒
- Reverse Filter Ports:** ☒
- Filter Chain:**
 - L4-L7 Service Graph:** select an option
 - QoS Priority:** (empty)
- Filters:** A table with columns: Name, Directives, Action, Priority. The table shows a filter named 'common/default' with 'none' directives and 'Permit' action. Below the table, there is a dropdown menu showing a list of filters for the 'common' tenant:
 - arp common
 - default common
 - est common
 - icmp common

Buttons for 'Update', 'Cancel', and 'OK' are visible at the bottom of the dialog.

11. Click **Update**.
12. Click **OK** to complete creating the contract subject.
13. Click **Submit** to complete creating the contract.

Provide Contracts for External Routed Networks from Tenant (common)

To provide contracts to access external routed networks, follow the procedures outlined in this section.

Setup Information

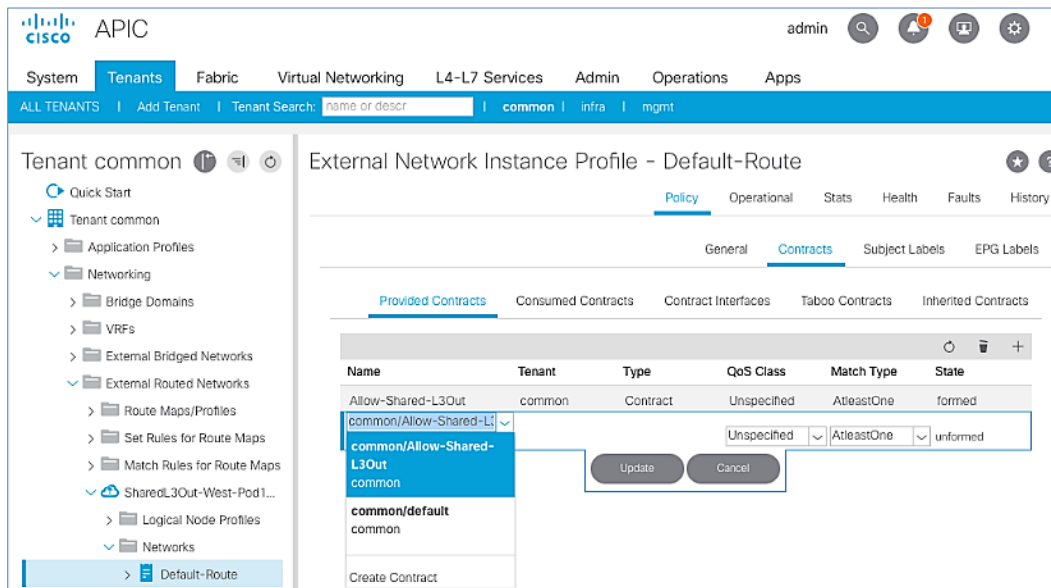
Table 11 External Routed Network Contracts

Shared L3Out	Contract	Subject	Filter
	Allow-Shared-L3Out	Allow-Shared-L3Out	common/default ✓ Global Scope

Deployment Steps

To provide contracts for external routed networks from Tenant common, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Tenants > common**.
3. In the left navigation pane, select and expand **Tenant common > Networking > External Routed Networks**.
4. Select and expand the recently created External Routed Network for SharedL3out or Routed Outside network (for example, SharedL3Out-West-Pod1_RO).
5. Select and expand **Networks**.
6. Select the recently created route (for example, Default-Route).
7. In the right window pane, select the tab for **Policy** and then **Contracts**.
8. Under the **Provided Contracts** tab, click **+** on the right to add a Provided Contract.
9. For **Name**, select the previously created contract (for example, common/Allow-Shared-L3Out) from the drop-down list.



10. Click **Update**.
11. Other Tenants can now 'consume' the Allow-Shared-L3Out contract to route traffic outside the ACI fabric. This deployment example shows a default filter to allow all traffic.



Customers can modify this contract as needed to limit access to specific destinations through the Shared L3Out connection .

Configure External Gateways in the Outside Network

This section provides a sample configuration from the external Layer 3 Gateways routers that connect to Pod-1. The gateways are in the external network and peer using OSPF to two ACI border leaf switches in Pod-1. Nexus 7000 routers are used as External gateway routers in this design but other Cisco models can also be used.



The gateway configuration shown below shows only the relevant portion of the configuration; it is not the complete configuration.

Enable Protocols

The protocols used between the ACI border leaf switches and external gateways have to be explicitly enabled on Nexus platforms used as external gateways in this design. The configuration to enable these protocols are provided below.

Table 12 Protocols Enabled

External Gateway Configuration – Pod-1	AA-West-Enterprise-1 (GW-1)	AA-West-Enterprise-2 (GW-2)
	<pre>feature ospf feature interface-vlan feature lacp feature lldp</pre>	<pre>feature ospf feature interface-vlan feature lacp feature lldp</pre>

Configure OSPF

OSPF is used between the external gateways and ACI border leaf switches to exchange routing between the two domains. The global configuration for OSPF is provided below. Loopback is used as the router IDs for OSPF. Note that interfaces between ACI border leaf switches will be in OSPF Area 10.

Table 13 Routing Protocol Configuration on External Gateways

External Gateway Configuration – Pod-1	AA-West-Enterprise-1 (GW-1)	AA-West-Enterprise-2 (GW-2)
	<pre>interface loopback0 description RID for OSPF ip address 13.13.13.98/32 ip router ospf 10 area 0.0.0.0 router ospf 10 router-id 13.13.13.98 area 0.0.0.10 nssa no-summary no- redistribution default-information-originate</pre>	<pre>interface loopback0 description RID for OSPF ip address 13.13.13.99/32 ip router ospf 10 area 0.0.0.0 router ospf 10 router-id 13.13.13.99 area 0.0.0.10 nssa no-summary no- redistribution default-information-originate</pre>

Configure Interfaces

The interface level configuration for connectivity between external gateways and ACI border leaf switches in Pod-1 is provided below. Note that interfaces to ACI are in OSPF Area 10 while the loopbacks and port-channels between the gateways are in OSPF Area 0.

Table 14 Interface Configuration – To ACI Border Leaf Switches

External Gateway Configuration - Pod-1	AA-West-Enterprise-1 (GW-1)	AA-West-Enterprise-2 (GW-2)
	<pre>interface Ethernet4/16 description To AA11-9372PX-WEST-1:Eth1/47 no shutdown interface Ethernet4/16.311 encapsulation dot1q 311 ip address 10.113.1.2/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.10 no shutdown</pre>	<pre>interface Ethernet4/16 description To AA11-9372PX-WEST-1:Eth1/48 no shutdown interface Ethernet4/16.312 encapsulation dot1q 312 ip address 10.113.1.6/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.10 no shutdown</pre>
	<pre>interface Ethernet4/20 description To AA11-9372PX-WEST-2:Eth1/47 no shutdown interface Ethernet4/20.313 encapsulation dot1q 313 ip address 10.113.2.2/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.10 no shutdown</pre>	<pre>interface Ethernet4/20 description To AA11-9372PX-WEST-2:Eth1/48 no shutdown interface Ethernet4/20.314 encapsulation dot1q 314 ip address 10.113.2.6/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.10 no shutdown</pre>

The configuration on the port-channel with 2x10GbE links that provide direct connectivity between the external gateways is provided below.

Table 15 Interface Configuration – Between External Gateways

External Gateway Configuration - Pod-1	AA-West-Enterprise-1 (GW-1)	AA-West-Enterprise-2 (GW-2)
	<pre>interface port-channel13 description To AA11-7004-2-AA-West-Enterprise-2 ip address 10.113.98.1/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.0 interface Ethernet4/13 description To AA11-7004-2-AA-West-Enterprise-2:Eth4/13 channel-group 13 mode active no shutdown interface Ethernet4/17 description To AA11-7004-2-AA-West-Enterprise-2:Eth4/17 channel-group 13 mode active no shutdown</pre>	<pre>interface port-channel13 description To AA11-7004-1-AA-West-Enterprise-1 ip address 10.113.98.2/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.0 interface Ethernet4/13 description To AA11-7004-1-AA-West-Enterprise-1:Eth4/13 channel-group 13 mode active no shutdown interface Ethernet4/17 description To AA11-7004-1-AA-West-Enterprise-1:Eth4/17 channel-group 13 mode active no shutdown</pre>

Solution Deployment – ACI Fabric (Multi-Pod)

The active-active data centers leverage a Cisco Multi-Pod ACI fabric design to extend the ACI fabric and the stretched cluster across two data centers to provide business continuity in the event of a disaster. The ACI Pods can be in the same data center location or in different geographical sites. This design assumes the two Pods are in two different geographical locations that was validated in the Cisco labs using a 75km fiber spool to interconnect the data centers.

This section provides detailed procedures for setting up a Cisco ACI Multi-Pod Fabric. An Inter-Pod network is first deployed to provide connectivity between data centers, followed by an ACI fabric to provide network connectivity in the second data center. The ACI fabric will serve as the second Pod (Pod-2 in [Figure 1](#)) in the ACI Multi-Pod fabric. In this design, half of the HyperFlex stretched cluster nodes will connect to Pod-1 and the remaining half to Pod-2.



The procedures outlined this section are specific to deploying a Cisco ACI Multi-Pod fabric.

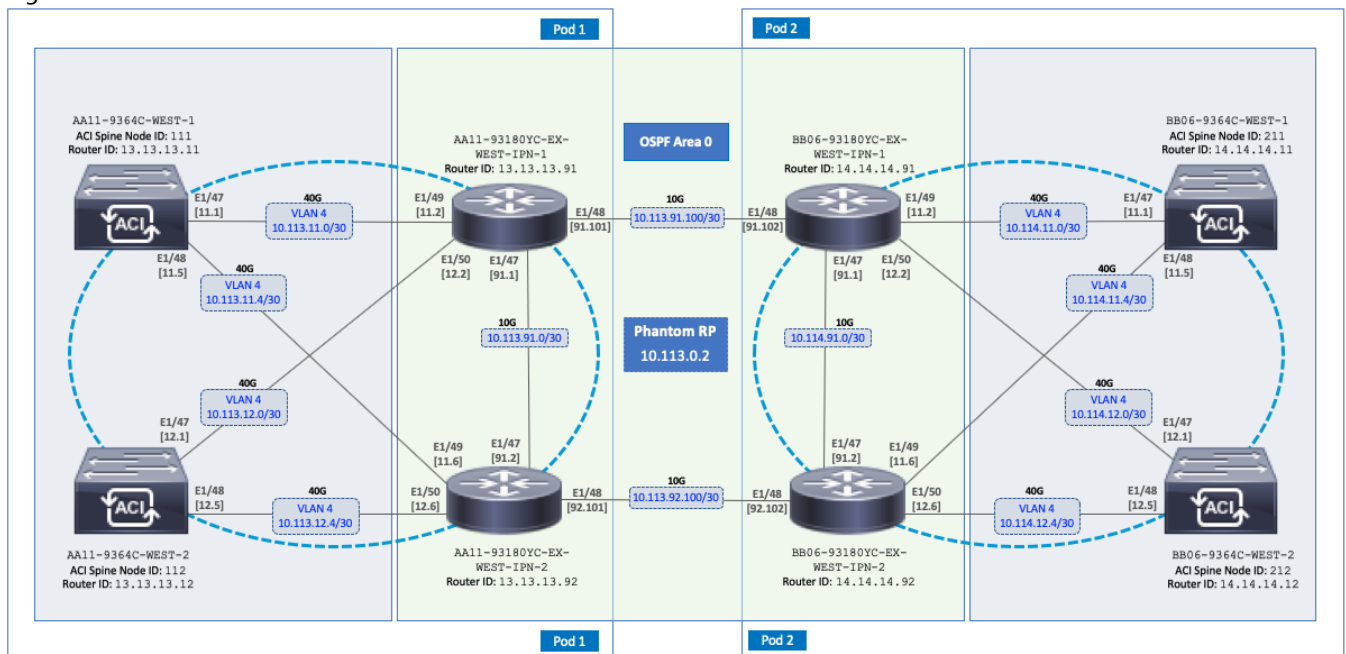
Prerequisites

Before an ACI Multi-Pod fabric can be deployed, the first ACI fabric (or Pod-1) should be up and running with Spine switches, Leaf switches and APICs.

Topology

The figure below shows the connectivity between Pods through the IPN and the connectivity from each Pod to the IPN. The connectivity between IPN devices uses 10GbE but the Spine switches in each Pod connect to the IPN devices using 40GbE links. Multiple nodes and links are used from each Pod to IPN and between IPNs to provide a redundant paths between Pods in the event of failures.

Figure 8 ACI Multi-Pod Fabric



Deployment Overview

A high-level overview of the steps involved in deploying an ACI Multi-Pod fabric is summarized below.

Physical Connectivity

The following are the steps involved to set up the physical connectivity:

- Complete the physical connectivity within the Inter-Pod Network (IPN) to provide connectivity between Pods or sites.
- Deploy Spine switches, Leaf switches and APIC(s) in the second ACI Pod. In this design, the third node in the 3-node APIC cluster is deployed in Pod-2. For discovery and auto-provisioning of the fabric in a new Pod, a Spine switch must have at least one link up to a Leaf switch. Spine switches will learn that a Leaf switch is connected through LLDP, which is enabled by default.
- Complete the physical connectivity to connect Spine switches to the IPN in each Pod. It is not necessary to connect all Spines in a Pod to the IPN. For redundancy, at least two Spines in each Pod should be connected to the IPN. The connected Spine switches will be seen as equal cost paths to that Pod's TEP addresses so connecting more Spine switches to the IPN should increase the number of Equal-Cost Multi-Paths (ECMP) routes for a greater distribution of traffic load.

Deploy Inter-Pod Network (IPN)

The following are the steps involved to deploy the inter-pod network:

- (Optional) Configure a VRF for ACI Multi-Pod traffic on all IPN devices and put the relevant interfaces in the VRF. This isolates the ACI Multi-Pod traffic and protects the ACI underlay network that is now exposed through the IPN. The IPN can be thought of as an extension of the ACI underlay infrastructure in each Pod. The underlay is necessary for establishing VXLAN tunnels between leaf switches and spine switches in each Pod. VXLAN tunnels enable seamless forwarding of Layer 2 and Layer 3 data plane traffic between Pods. The VXLAN overlay is essential for ensuring that the interconnected Pods function as a single ACI fabric.
- Configure Layer 2 encapsulation, Layer 2 protocols (LLDP, CDP), MTU (Jumbo) and IP addressing on relevant interfaces of the IPN devices that provide connectivity within the IPN, and between the IPN and Spines in each Pod. The Spine switches will tag all traffic towards the IPN using VLAN 4. Therefore, IPN devices must be configured for trunking using VLAN 4 on the interfaces connecting to the Spine. Enabling LLDP (preferred) or CDP on IPN interfaces is recommended for determining which ports connect to which devices. Encapsulating traffic in VXLAN adds 50 Bytes of overhead so the IPN must set to an MTU that is at least 50 Bytes higher than the MTU of the traffic being transported across VXLAN in order to prevent fragmentation. For traffic such as HyperFlex storage and vMotion traffic that use jumbo (9000 Byte) MTU, the MTU on the IPN should be the jumbo MTU plus 50 Bytes. MTU used in validation is 9216B as it is a commonly used value for jumbo MTU on many Cisco platforms.
- Enable routing within the IPN and on the connections to Spines to advertise TEP pools between Pods. Each Pod uses a unique TEP pool that must be advertised to the other Pod in order to establish VXLAN Tunnels from one Pod to the other. The Spines in each Pod that connect to the IPN also use Proxy TEP addressing that are also advertised to the other Pods. The proxy TEP addressing enables each Spine to advertise equal cost routes for the Pod subnets to the IPN routers. IPN will use the ECMP to the Spines to distribute traffic to the Pod subnets. Loopback interfaces are used on IPN nodes are used as the router-id for the routing protocol. Currently, OSPFv2 is the only routing protocol supported. Note that underlay infrastructure in an ACI Pod uses ISIS and not OSPF. If the IPN is an extensive L3 network that is already using another routing protocol, it is not necessary to use OSPF everywhere in the IPN – it is only necessary between the Spine switches and IPN devices.
- Enable IP Multicast routing using Bidirectional PIM (BIDIR-PIM) to forward Broadcast, Unknown Unicast and Multicast (BUM) traffic between Pods. This is necessary when endpoints in the same Bridge Domain are distributed

across both Pods, to enable seamless East-West communication between endpoints for multi-destination or non-unicast traffic. BUM traffic is encapsulated in a VXLAN multicast frame to transport it within or between Pods. In an ACI fabric, a multicast traffic within each Bridge Domain is sent to a unique IP multicast group address. The multicast address for the bridge domain is assigned when the bridge domain is first defined in ACI. The address is allocated from a pool of multicast addresses, known as Global IP Outside (GIPO) in ACI. To forward BUM traffic between Pods, the IPN needs to support IP multicast, specifically BIDIR-PIM. In ACI Multi-Pod, when a Bridge Domain is activated within a Pod, an IGMP Join is forwarded to the IPN to receive BUM traffic from remote endpoints in the same Pod. The multicast address pool used for BUM traffic for bridge domains that span the IPN can be the same as the infrastructure GIPO range used within a Pod or different pool can be allocated for this. BIDIR-PIM requires a Rendezvous Point (RP) to be defined. For RP resiliency, a phantom RP can be used. For distributing the RP load,

- Configure DHCP Relay on IPN devices to enable auto-discovery and auto-configuration of Spines and APICs in Pod-2 from Pod-1.

Setup ACI Fabric for Multi-Pod

The following are the steps involved to set up the ACI fabric for Multi-Pod:

- Configure IP connectivity to connect Spine Interfaces to IPN devices in Pod-1.
- Configure Routing Protocols (OSPF, BGP) on the Spine Switches. OSPF will provide IP reachability between Pods, specifically between TEP address pools in each Pod. ACI Fabric will redistribute routes from IS-IS used within each Pod to OSPF and vice-versa. This effectively extends the underlay network (VRF overlay-1 in ACI Fabric) to the IPN. BGP will be used to advertise learned MAC and IP addresses of endpoints and their locations. The endpoint information is maintained on separate Counsel of Oracle Protocol (COOP) database on Spine switches on each Pod. Endpoints learned on each local Pod is advertised across the BGP-EVPN peering between Pods. The peering is directly between Spine switches in the Pods. When multiple Pods are connected across the IPN, BGP route-reflectors can be deployed in the IPN rather than direct peering between Pods.
- Configure External TEP Addresses for Spine switches to use for Spine-to-Spine connections across the IPN.
- Add a second Pod to the ACI fabric.

Setup Pod-2 Spine Switches, Leaf Switches, and APICs

The following are the steps involved to set up the Pod-2 spine switches, leaf switches, and APICs:

- Configure ACI Fabric access policies to enable connectivity from Pod-1 Spines switches to the IPN.
- Configure newly discovered Spine and Leaf switches in Pod-2 from the first Pod.
- Configure ACI Fabric Access Policies to enable connectivity from Pod-2 Spines switches to the IPN.
- Deploy a third APIC in Pod-2 to form a 3-node APIC cluster to manage the fabric.

For additional information about ACI Multi-Pod, see the [References](#) section of this document and the ACI product documentation.

Deployment Guidelines

The following are the deployment guidelines:

- IPN must support an MTU of 50 Bytes higher than the MTU used by the endpoints in the deployment. In this design, the HyperFlex stretched cluster that connects to the ACI Multi-Pod Fabric uses an MTU of 9000 Bytes or Jumbo frames for Storage and vMotion traffic. It is also possible for other (for example, Management, Applications) traffic

in the HyperFlex cluster to use Jumbo frames. In this design, the IPN MTU is set to 9216 Bytes to keep it consistent with the Jumbo MTU on other Cisco platforms.

- ACI Multi-Pod Fabric uses a **VLAN ID of 4** for connectivity between Spine Switches and IPN devices in each Pod. This is system defined and cannot be changed – the IPN devices connecting to the Spines must therefore be configured to use VLAN 4.
- IPN device must support a BIDIR-PIM range of at least /15. First generation Nexus 9000 series switches cannot be used as IPN devices as the ASICs used on these support a max BIDIR-PIM range of /24.
- For auto-discovery and auto-configuration of newly added Spine switches to work, at least one Leaf switch must be online and connected to the Spine switch in the remote Pod. The Spine switch should be able to see the Leaf switch via LLDP.
- A Multi-Pod ACI fabric deployment requires the 239.255.255.240 (System GIPo) to be configured as a BIDIR-PIM range on the IPN devices. This configuration is not required when using the **Infra GIPo as System GIPo** feature. The APIC and switches must be running releases that support this feature.
- Spine switches from each Pod cannot be directly connected to each other – they must go through at least one IPN router/switch.
- It is not necessary to connect all Spines switches in a Pod to the IPN. If possible, connect at least two Spine switches from each Pod to the IPN to provide node redundancy in the event of a Spine switch failure. Traffic is distributed across all the spine switches that are connected to the IPN so more spine switches can be connected to distribute the load even further.

Deploy Inter-Pod Network

This section provides the configuration for deploying Inter-Pod switches that provide Pod-to-Pod connectivity. The IPN is not managed by the APIC. IPN can be thought of as an extension of the ACI underlay network. IPN devices must be enabled for L3 forwarding with VRF Lite (recommended), OSPF, DHCP Relay and BIDIR-PIM. LACP is also required when link bundling is deployed. LLDP is optional but recommended to verify connectivity to peers and ports used for the connection.

Deployment Overview

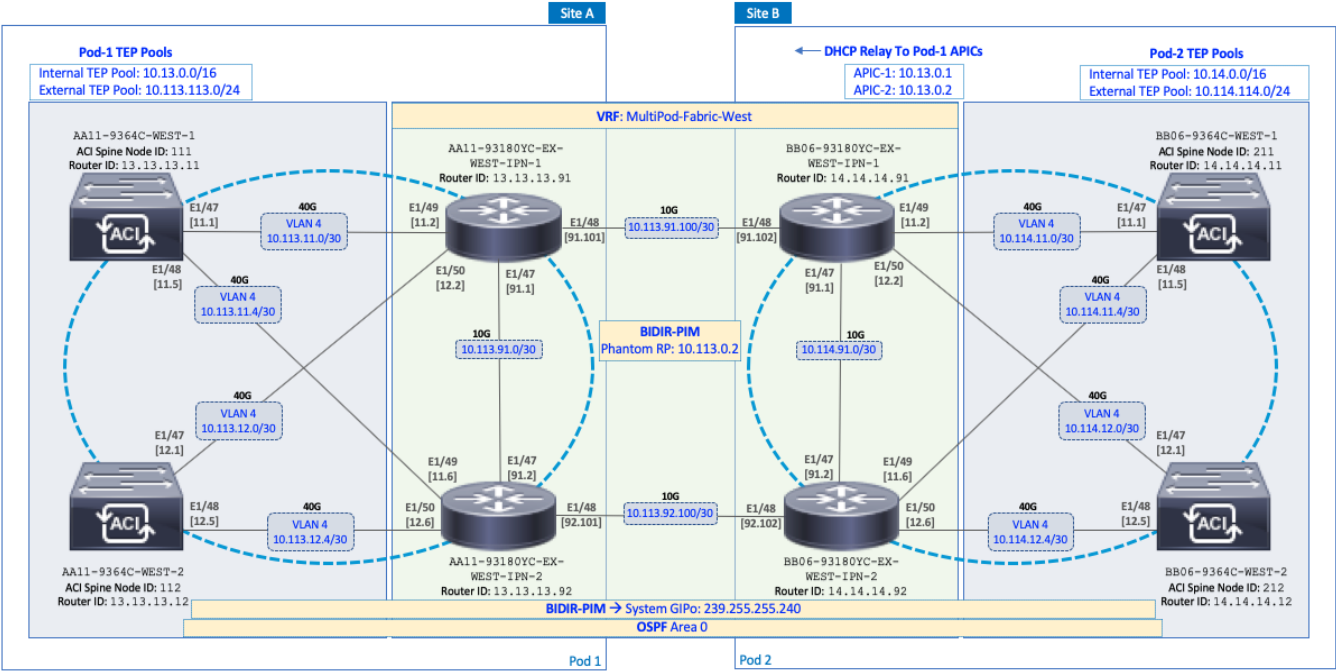
The high-level steps involved in the setting up the Inter-Pod Network is as follows:

- Complete the physical connectivity to connect IPN devices to Spine switches in each Pod and to remote IPN devices in the other Pod.
- Identify and collect the information required to setup the IPN.
- Configure IPN Devices in Pod-1.
- Configure IPN Devices in Pod-2.

Physical Connectivity

[Figure 9](#) illustrates the IPN connectivity between IPN devices and to Spine switches in each Pod. The connectivity between IPN devices uses 10GbE and 40GbE to Spine switches.

Figure 9 Inter-Pod Network Connectivity



Configure IPN Devices in Pod-1

Table 16 Pod-1 IPN Configuration

<pre> switchname AA11-93180YC-EX-WEST-IPN-1 feature ospf feature pim feature lacp feature dhcp feature lldp ntp server 172.26.163.254 service dhcp ip dhcp relay vrf context MultiPod-Fabric-West ip pim rp-address 10.113.0.2 group-list 226.0.0.0/8 bidir ip pim rp-address 10.113.0.2 group-list 239.255.255.240/28 bidir ip pim ssm range 232.0.0.0/8 vrf context management ip route 0.0.0.0/0 172.26.163.254 ... interface Ethernet1/47 description To POD-1:AA11-93180YC-EX-WEST-IPN-2:E1/47 no switchport mtu 9216 vrf member MultiPod-Fabric-West ip address 10.113.91.1/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.0 ip pim sparse-mode no shutdown interface Ethernet1/48 description To POD-2:BB06-93180YC-EX-WEST-IPN-1:E1/48 no switchport mtu 9216 vrf member MultiPod-Fabric-West ip address 10.113.91.101/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.0 ip pim sparse-mode no shutdown interface Ethernet1/49 description To POD-1:AA11-9364C-1:E1/47 no switchport mtu 9216 no shutdown interface Ethernet1/49.4 mtu 9216 encapsulation dot1q 4 vrf member MultiPod-Fabric-West ip address 10.113.11.2/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.0 ip pim sparse-mode no shutdown </pre>	<pre> switchname AA11-93180YC-EX-WEST-IPN-2 feature ospf feature pim feature lacp feature dhcp feature lldp ntp server 172.26.163.254 service dhcp ip dhcp relay vrf context MultiPod-Fabric-West ip pim rp-address 10.113.0.2 group-list 226.0.0.0/8 bidir ip pim rp-address 10.113.0.2 group-list 239.255.255.240/28 bidir ip pim ssm range 232.0.0.0/8 vrf context management ip route 0.0.0.0/0 172.26.163.254 ... interface Ethernet1/47 description To POD-1:AA11-93180YC-EX-WEST-IPN-1:E1/47 no switchport mtu 9216 vrf member MultiPod-Fabric-West ip address 10.113.91.2/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.0 ip pim sparse-mode no shutdown interface Ethernet1/48 description To POD-2:BB06-93180YC-EX-WEST-IPN-2:E1/48 no switchport mtu 9216 vrf member MultiPod-Fabric-West ip address 10.113.92.101/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.0 ip pim sparse-mode no shutdown interface Ethernet1/49 description To POD-1:AA11-9364C-WEST-1:E1/48 no switchport mtu 9216 no shutdown interface Ethernet1/49.4 mtu 9216 encapsulation dot1q 4 vrf member MultiPod-Fabric-West ip address 10.113.11.6/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.0 ip pim sparse-mode no shutdown </pre>
--	---

```

interface Ethernet1/50
  description To POD-1:AA11-9364C-2:E1/47
  no switchport
  mtu 9216
  no shutdown

interface Ethernet1/50.4
  mtu 9216
  encapsulation dot1q 4
  vrf member MultiPod-Fabric-West
  ip address 10.113.12.2/30
  ip ospf network point-to-point
  ip ospf mtu-ignore
  ip router ospf 10 area 0.0.0.0
  ip pim sparse-mode
  no shutdown

...

interface mgmt0
  vrf member management
  ip address 172.26.163.98/24

interface loopback0
  description OSPF Router-ID
  vrf member MultiPod-Fabric-West
  ip address 13.13.13.91/32
  ip router ospf 10 area 0.0.0.0

interface loopback1
  description To BIDIR-PIM Phantom RP
  vrf member MultiPod-Fabric-West
  ip address 10.113.0.1/30
  ip ospf network point-to-point
  ip router ospf 10 area 0.0.0.0
  ip pim sparse-mode

router ospf 10
  vrf MultiPod-Fabric-West
  router-id 13.13.13.91
  log-adjacency-changes

```

```

interface Ethernet1/50
  description To POD-1:AA11-9364C-WEST-2:E1/48
  no switchport
  mtu 9216
  no shutdown

interface Ethernet1/50.4
  mtu 9216
  encapsulation dot1q 4
  vrf member MultiPod-Fabric-West
  ip address 10.113.12.6/30
  ip ospf network point-to-point
  ip ospf mtu-ignore
  ip router ospf 10 area 0.0.0.0
  ip pim sparse-mode
  no shutdown

...

interface mgmt0
  vrf member management
  ip address 172.26.163.99/24

interface loopback0
  description OSPF Router-ID
  vrf member MultiPod-Fabric-West
  ip address 13.13.13.92/32
  ip router ospf 10 area 0.0.0.0

router ospf 10
  vrf MultiPod-Fabric-West
  router-id 13.13.13.92
  log-adjacency-changes

```

Configure IPN Devices in Pod-2

Table 17 Pod-2 IPN Configuration

switchaname BB06-93180YC-EX-WEST-IPN-1	switchaname BB06-93180YC-EX-WEST-IPN-2
<pre> feature ospf feature pim feature lacp feature dhcp feature lldp ntp server 172.26.164.254 service dhcp ip dhcp relay vrf context MultiPod-Fabric-West ip pim rp-address 10.113.0.2 group-list 226.0.0.0/8 bidir ip pim rp-address 10.113.0.2 group-list 239.255.255.240/28 bidir ip pim ssm range 232.0.0.0/8 vrf context management ip route 0.0.0.0/0 172.26.164.254 ... interface Ethernet1/47 description To POD-2:BB06-93180YC-EX-WEST-IPN- 2:E1/47 no switchport mtu 9216 vrf member MultiPod-Fabric-West ip address 10.114.91.1/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.0 ip pim sparse-mode no shutdown interface Ethernet1/48 description To POD-1:AA11-93180YC-EX-WEST-IPN- 1:E1/48 no switchport mtu 9216 vrf member MultiPod-Fabric-West ip address 10.113.91.102/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.0 ip pim sparse-mode no shutdown interface Ethernet1/49 description To POD-2:BB06-9364C-1:E1/47 no switchport mtu 9216 no shutdown interface Ethernet1/49.4 mtu 9216 encapsulation dot1q 4 vrf member MultiPod-Fabric-West ip address 10.114.11.2/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.0 ip pim sparse-mode ip dhcp relay address 10.13.0.1 ip dhcp relay address 10.13.0.2 no shutdown </pre>	<pre> feature ospf feature pim feature lacp feature dhcp feature lldp ntp server 172.26.164.254 service dhcp ip dhcp relay vrf context MultiPod-Fabric-West ip pim rp-address 10.113.0.2 group-list 226.0.0.0/8 bidir ip pim rp-address 10.113.0.2 group-list 239.255.255.240/28 bidir ip pim ssm range 232.0.0.0/8 vrf context management ip route 0.0.0.0/0 172.26.164.254 ... interface Ethernet1/47 description To POD-2:BB06-93180YC-EX-WEST-IPN- 1:E1/47 no switchport mtu 9216 vrf member MultiPod-Fabric-West ip address 10.114.91.2/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.0 ip pim sparse-mode no shutdown interface Ethernet1/48 description To POD-1:AA11-93180YC-EX-WEST-IPN- 2:E1/48 no switchport mtu 9216 vrf member MultiPod-Fabric-West ip address 10.113.92.102/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.0 ip pim sparse-mode no shutdown interface Ethernet1/49 description To POD-2:BB06-9364C-WEST-1:E1/48 no switchport mtu 9216 no shutdown interface Ethernet1/49.4 mtu 9216 encapsulation dot1q 4 vrf member MultiPod-Fabric-West ip address 10.114.11.6/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.0 ip pim sparse-mode ip dhcp relay address 10.13.0.1 ip dhcp relay address 10.13.0.2 no shutdown </pre>

```

interface Ethernet1/50
  description To POD-2:BB06-9364C-2:E1/48
  no switchport
  mtu 9216
  no shutdown

```

```

interface Ethernet1/50.4
  mtu 9216
  encapsulation dot1q 4
  vrf member MultiPod-Fabric-West
  ip address 10.114.12.2/30
  ip ospf network point-to-point
  ip ospf mtu-ignore
  ip router ospf 10 area 0.0.0.0
  ip pim sparse-mode
  ip dhcp relay address 10.13.0.1
  ip dhcp relay address 10.13.0.2
  no shutdown

```

...

```

interface mgmt0
  vrf member management
  ip address 172.26.164.98/24

```

```

interface loopback0
  description OSPF Router-ID
  vrf member MultiPod-Fabric-West
  ip address 14.14.14.91/32
  ip router ospf 10 area 0.0.0.0

```

```

interface loopback1
  description BIDIR-PIM Phantom RP
  vrf member MultiPod-Fabric-West
  ip address 10.113.0.1/29
  ip ospf network point-to-point
  ip router ospf 10 area 0.0.0.0
  ip pim sparse-mode

```

```

router ospf 10
  vrf MultiPod-Fabric-West
  router-id 14.14.14.91
  log-adjacency-changes

```

```

interface Ethernet1/50
  description To POD-2:BB06-9364C-WEST-2:E1/48
  no switchport
  mtu 9216
  no shutdown

```

```

interface Ethernet1/50.4
  mtu 9216
  encapsulation dot1q 4
  vrf member MultiPod-Fabric-West
  ip address 10.114.12.6/30
  ip ospf network point-to-point
  ip ospf mtu-ignore
  ip router ospf 10 area 0.0.0.0
  ip pim sparse-mode
  ip dhcp relay address 10.13.0.1
  ip dhcp relay address 10.13.0.2
  no shutdown

```

...

```

interface mgmt0
  vrf member management
  ip address 172.26.164.99/24

```

```

interface loopback0
  description OSPF Router-ID
  vrf member MultiPod-Fabric-West
  ip address 14.14.14.92/32
  ip router ospf 10 area 0.0.0.0

```

```

router ospf 10
  vrf MultiPod-Fabric-West
  router-id 14.14.14.92
  log-adjacency-changes

```

Setup ACI Fabric for Multi-Pod – Using Configuration Wizard

In APIC Release **4.0(1)** and higher, ACI Multi-Pod can be deployed using a configuration wizard that configures the fabric for Multi-Pod.

Prerequisites

The Inter-Pod network should be setup prior to configuring the ACI fabric for Multi-Pod.

Deployment Overview

Deploying ACI Multi-Pod using the APIC Configuration Wizard consists of the following high-level activities:

- Configure Interpod Connectivity - For connecting the first Pod or site to IPN and setting up Multi-Pod
- Add Physical Pod - For adding a second Pod or site in the Multi-Pod setup

The **Configure Interpod Connectivity** portion of the wizard is for setting up the first Pod or site (Pod-1) for the following:

- **IP Connectivity** from Spines in Pod-1 to the Inter-Pod network. This includes configuring the Spine interfaces that connect to the IPN for IP connectivity. The APIC on the back-end will take the minimal information provided to the wizard, to configure the necessary fabric access policies for connecting devices to the ACI fabric. This includes configuration of interface and switch-level, policies and profiles on the Spines connecting to the IPN.
- **Routing Protocols** to enable IP Routing on the Spines in Pod-1 towards the IPN. This includes OSPF-based underlay network for exchanging routes between the Pods and MP-BGP based overlay network for exchanging endpoint location information using MP-BGP EVPN.
- **External TEP** addressing for Pod-1 to communicate with other Pods or sites. This includes specifying a routable External TEP Pool for the first Pod or site.

The **Add Physical Pod** portion of the wizard is for adding the second Pod or site (Pod-2) and consists of the following:

- **Pod Fabric** information for creating a second Pod. This includes specifying a unique Pod ID and TEP Pool for the new Pod. It also includes parameters for configuring IP connectivity from Spines in Pod-2 to the Inter-Pod network, similar to the information used in Pod-1 for connecting the Spines in Pod-1 to IPN.
- **External TEP** addressing for Pod-2 to communicate with other Pods or sites. This includes specifying a routable External TEP Pool for the second Pod or site.
- Configure **DHCP Relay** on IPN devices in Pod-2 to point to Pod-1 APIC TEP IP Addresses.
- Configure OSPF interface policies for Pod-2 Spine switches that connect to the IPN

The setup information and deployment steps for configuring Interpod connectivity and adding a Physical Pod using the Wizard are covered in the next sections.

Configure Inter-Pod Connectivity

Follow the procedures in this section to configure Inter-Pod connectivity to connect the Spine switches in Pod-1 to IPN and set up ACI Fabric for Multi-Pod.

IP Connectivity

IP Connectivity section of the wizard provides the physical interface and IP configuration on the Spines switches in Pod-1 that connect to IPN devices. The parameters used in this CVD for this portion of the configuration is provided in [Table 18](#) .

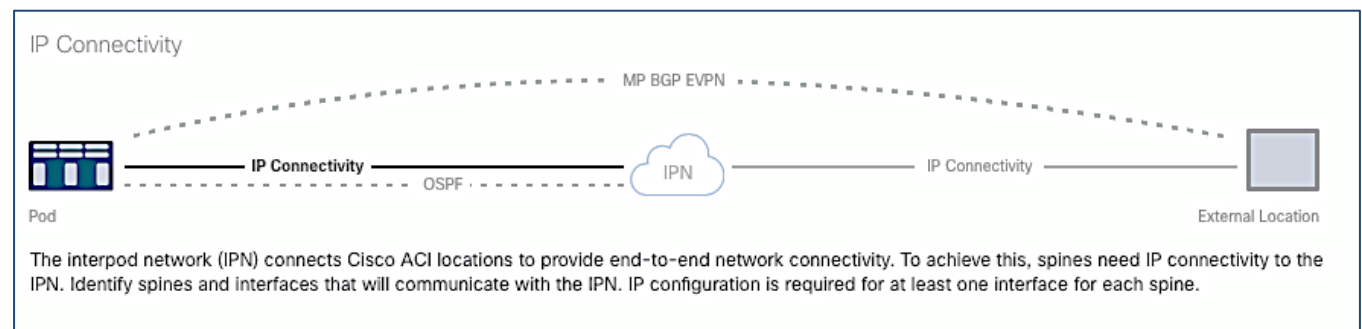


Table 18 IP Connectivity Information for Pod-1

IP Connectivity to IPN	Spine ID	Interfaces	IP Addresses	MTU
	111	E1/47	10.113.11.1/30	9216
		E1/48	10.113.11.5/30	9216
	112	E1/47	10.113.12.1/30	9216
		E1/48	10.113.12.5/30	9216

Routing Protocols

Routing Protocols section of the wizard provides the routing protocol (OSPF, BGP) configuration on the Spine switches in Pod-1 that connect to IPN to enable the OSPF based underlay network and MP-BGP based overlay. The parameters used in this CVD for this portion of the configuration is provided in Table 19 .

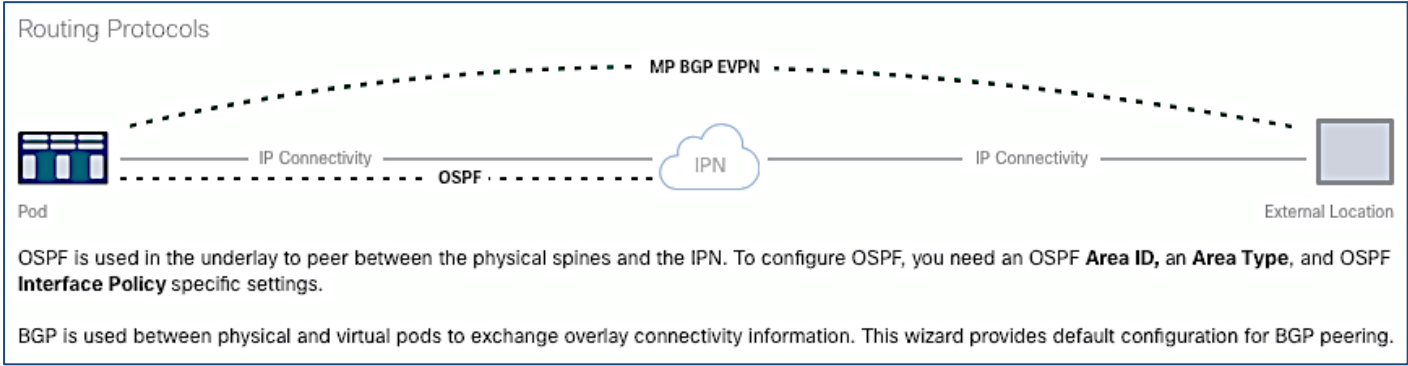


Table 19 Routing Protocols Information for Pod-1

Routing Protocols	OSPF		
	Area ID	0	
	Area Type	Regular	
	Interface Policy	MultiPod-OSPF_IP	Advertise Subnet MTU Ignore
	For remaining parameters	Use Defaults	
	BGP		
	Use Defaults		

External TEP

External TEP section of the wizard provides the addressing configuration on the Spine switches to enabled Pod-to-Pod connectivity across the Inter-Pod network. The parameters used in this CVD for this portion of the configuration is provided in the Table 20 .

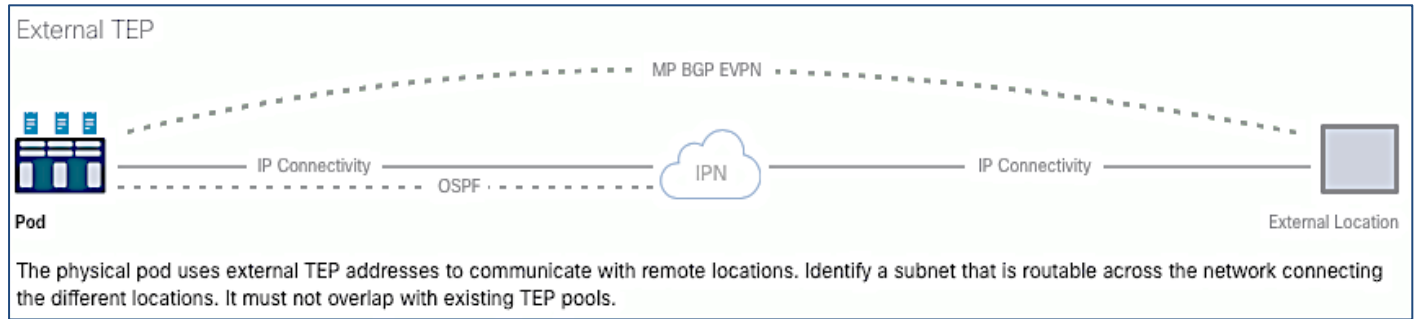


Table 20 External TEP Information for Pod-1

External TEP	POD-1	Addressing
	External TEP Pool	10.113.113.0/24*
	Spine Router ID(s)	13.13.13.11
		13.13.13.12
	Spine Loopback ID(s)	Same as Router IDs

* POD Specific; Can be a smaller pool – see Wizard for addresses allocated

Run Configuration Wizard for IPN Connectivity

To enable IPN connectivity for the Spines in Pod-1, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using **admin** account.
2. From the top navigation menu, select **Fabric > Inventory**.
3. From the left navigation pane, expand and select **Quick Start > Add Pod**.
4. From the right window, click **Add Pod**.
5. In the pop-up window for **Configure Interpod Connectivity** wizard, review the **Overview**. Collect the **Setup Information** for IP Connectivity, Routing Protocols and External TEP. For the parameters used in validating this CVD, see the Setup Information for Pod-1 tables in the previous section. Click **Get Started**.
6. In the **Step 2 > IP Connectivity** window of the wizard, for each Spine switch connecting to IPN devices, specify the Spine ID (for example, 111), interface (for example, e1/47), IP Address (for example, 10.13.11.1/30) and MTU (for example, 9216) that matches the MTU on the interfaces on the IPN devices that these interfaces connect to. ACI Multi-Pod requires a minimum of 9150 bytes but many Cisco devices includes the IP header in the MTU specified and therefore, 9216 bytes is used.
7. Click **[+]** to the right of the MTU to add more interfaces.
8. Click **[+]** to the right of the Spine ID to add more Spine switches.

APIC

System Tenants **Pod 1**

Inventory | Fabric Policy

Inventory

- Quick Start
- Add Remote Leaf
- Add Pod
- Topology
- Pod 1
 - Pod Fabric Setup Policy
 - Fabric Membership
 - Disabled Interfaces and
 - Duplicate IP Usage

Configure Interpod Connectivity

STEP 2 > IP Connectivity

1. Overview 2. IP Connectivity 3. Routing Protocol 4. External TEP

MP BGP EVPN

Pod IP Connectivity OSPF IPN IP Connectivity External Location

IP Connectivity

The interpod network (IPN) connects Cisco ACI locations to provide end-to-end network connectivity. To achieve this, spines need IP connectivity to IPN. Identify each spine by entering its node ID and define the interfaces that are connected to the IPN. Also provide IP configuration for at least one interface for each spine. Multiple interfaces are supported. It is best to have the same MTU set on all spine-to-IPN interfaces.

Spine ID: 111

Interfaces

Interface	IPv4 Address	MTU (bytes)
1/47	10.113.11.1/30	9216
1/48	10.113.11.5/30	9216

Spine ID: 112

Interfaces

Interface	IPv4 Address	MTU (bytes)
1/47	10.113.12.1/30	9216
1/48	10.113.12.5/30	9216

Previous Cancel Next

Last Login Time: 2018-11-18T01:26 UTC

9. Click **Next**.
10. In the **Step 3 > Routing Protocol** window of the wizard, for the Spine switches in Pod-1 connecting to the IPN devices, leave checkbox **Use Defaults** enabled, specify the **Area ID** (for example, 0), **Area Type** (for example, Regular) and for Interface Policy, click the drop-down list and select **Create OSPF Interface Policy**.
11. In the **Create OSPF Interface Policy** pop-up window, specify a **Name** (for example, MultiPod-OSPF_IP) for the interface policy. Specify the OSPF **Network Type** (for example, Point-to-point). For **Interface Controls**, select the checkbox for **Advertise subnet** and **MTU ignore**.

The screenshot shows the Cisco APIC interface for configuring interpod connectivity. The main window is titled "Configure Interpod Connectivity" and is at "STEP 3 > Routing Protocol". A progress bar at the top indicates four steps: 1. Overview, 2. IP Connectivity, 3. Routing Protocol (current), and 4. External TEP. The left sidebar shows the "Inventory" tree with options like "Quick Start", "Add Remote Leaf", "Add Pod", "Topology", "Pod 1", "Pod Fabric Setup Policy", "Fabric Membership", "Disabled Interfaces and", and "Duplicate IP Usage".

The main content area shows a diagram of a multi-pod fabric with a central "MP BGP EVPN" cloud. Below the diagram, there are sections for "Routing Protocols" and "OSPF". The "OSPF" section includes fields for "Area ID: 0", "Area Type: NSSA area", and "Interface Policy: select an option".

The "Create OSPF Interface Policy" modal window is open, titled "Define OSPF Interface Policy". It contains the following fields and options:

- Name:** MultiPod-OSPF_IP
- Description:** optional
- Network Type:** Broadcast, Point-to-point, Unspecified (Point-to-point is selected)
- Priority:** 1
- Cost of Interface:** unspecified
- Interface Controls:**
 - ☒ Advertise subnet
 - ☐ BFD
 - ☒ MTU ignore
 - ☐ Passive participation
- Hello Interval (sec):** 10
- Dead Interval (sec):** 40
- Retransmit Interval (sec):** 5
- Transmit Delay (sec):** 1

At the bottom of the modal are "Cancel" and "Submit" buttons. At the bottom of the main window are "Previous", "Cancel", and "Next" buttons.

12. Click **Submit**.

13. For **BGP**, leave the **Use Defaults** checkbox enabled.

Configure Interpod Connectivity

STEP 3 > Routing Protocol

1. Overview 2. IP Connectivity 3. Routing Protocol 4. External TEP

Diagram: Pod (represented by a server rack icon) connects via IP Connectivity to an IPN (cloud icon), which then connects via IP Connectivity to an External Location (server rack icon). A dashed line labeled 'MP BGP EVPN' connects the Pod and External Location.

Routing Protocols

OSPF is used in the underlay to peer between the physical spines and the IPN. Configure the OSPF **Area ID**, an **Area Type**, and OSPF **Interface Policy**. OSPF interface policy contains OSPF-specific settings like OSPF network type, interface cost, and timers. Configure the OSPF **Authentication Key** and OSPF **Area Cost** by unselecting **Use Defaults**.

BGP is used between physical and virtual pods to exchange overlay connectivity information. Configure **BGP Community**, **Peering Type**, and **Peer Password** by unselecting **Use Defaults**.

OSPF

Use Defaults: ☒

Area ID:

Area Type: ☐ NSSA area ☒ Regular area ☐ Stub area

Interface Policy:

For sub-interfaces

BGP

Use Defaults: ☒

Previous Cancel **Next**

14. Click **Next**.

15. In the **Step 3 > External TEP** section of the wizard, for the Spine switches in Pod-1 connecting to the IPN devices, leave the checkbox **Use Defaults** enabled. Specify the **External TEP Pool** (for example, 10.113.113.0/24) and Router IDs (for example, 13.13.13.11, 13.13.13.12) for the Spines.

APIC

System

Tenants

Inventory

Fabric Policy

Quick Start

Add Remote Leaf

Add Pod

Topology

Pod 1

Pod Fabric Setup Policy

Fabric Membership

Disabled Interfaces and

Duplicate IP Usage

Configure Interpod Connectivity

1. Overview

2. IP Connectivity

3. Routing Protocol

4. External TEP

Pod

MP BGP EVPN

External Location

IPN

IP Connectivity

OSPF

External TEP

The physical pod uses external TEP to communicate with remote locations. Configure a subnet that is routable across the network connecting the different locations. The external TEP pool must not overlap external TEP pools belonging to other pods. The pool size should be between /27 and /22. The pool should be large enough to address all Cisco APICs, all spines, all border leafs, pod-specific TEP addresses, and spine router IDs.

The wizard automatically allocates addresses for pod-specific TEP addresses and spine router IDs from the external TEP pool. Proposed addresses can be modified, but modified addresses must be outside of the external TEP pool.

Use Defaults: ☒

Pod:

Internal TEP Pool:

External TEP Pool:

Data Plane TEP IP:

1

10.13.0.0/16

10.113.113.0/24

10.113.113.1/32

Spine ID:

Router ID:

Loopback Address:

111

13.13.13.11

Leave blank to use Router ID

Spine ID:

Router ID:

Loopback Address:

112

13.13.13.12

Leave blank to use Router ID

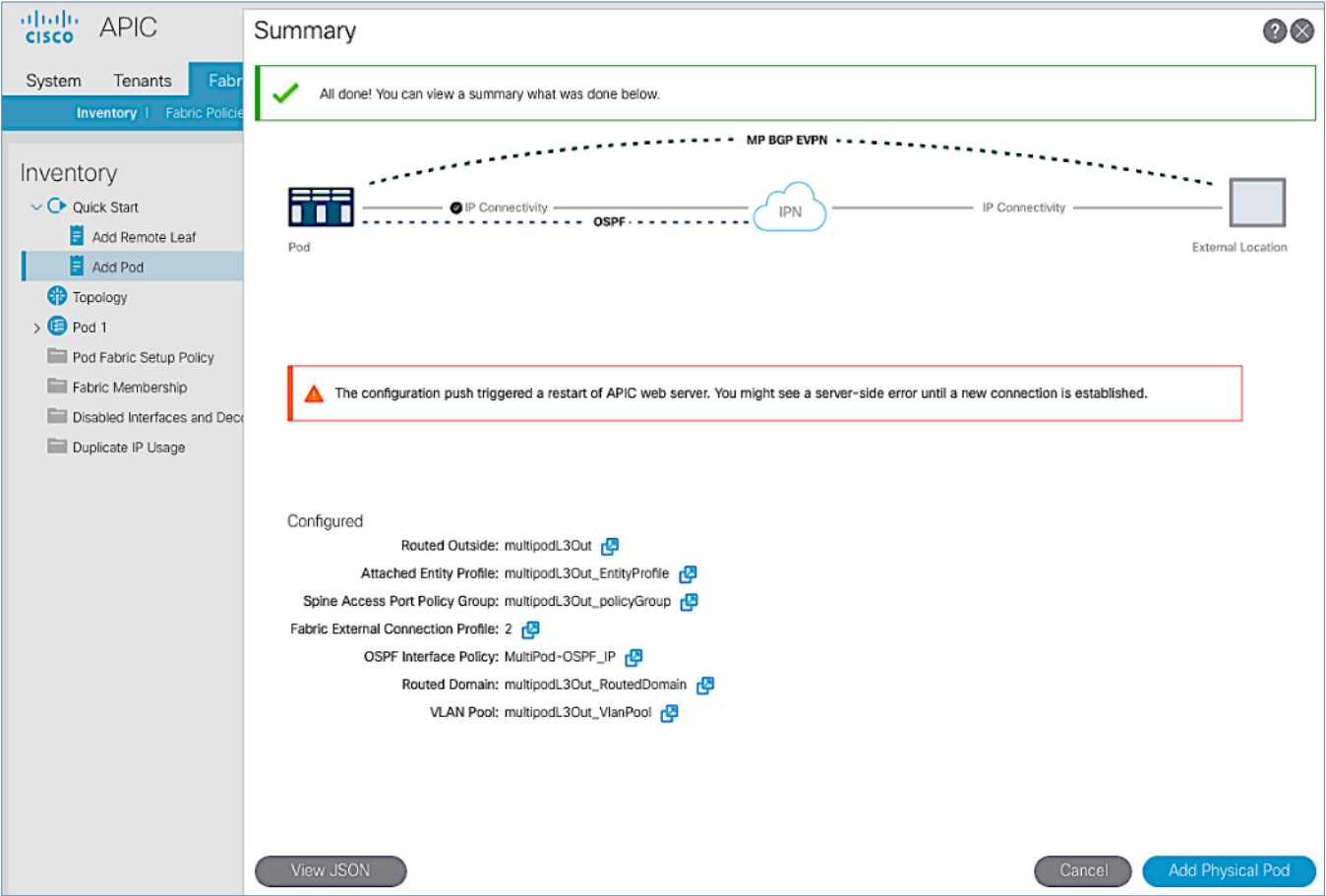
Previous

Cancel

Finish

16. Click **Finish** to complete the Inter-Pod connectivity setup for Spine switches in the first Pod or site (Pod-1).

101



- 17. In the **Summary** window, review the information provided.
- 18. (Optional) Click **View JSON** to save the JSON data for the configuration that was just completed.
- 19. (Optional) Click **Add Physical Pod** to continue to the next stage of the configuration now or come back to this at a later time. See next section for **Adding Physical Pod** deployment steps to add the second Pod or site.

Add Physical Pod – Second Pod or Site (Pod-2)

Table 21 Pod Configuration

		Pod 2
Pod Configuration	Pod Info	Value
	Pod ID	2
	TEP Pool	10.14.0.0/16

Table 22 IP Connectivity

Pod 2				
IP Connectivity to IPN	Spine ID	Interfaces	IP Addresses	MTU
	211	E1/47	10.114.11.1/30	9216
		E1/48	10.114.11.5/30	9216
	212	E1/47	10.114.12.1/30	9216
		E1/48	10.114.12.5/30	9216

Table 23 External TEP

Pod 2		
External TEP	TEP	
	Addressing	
	Internal TEP Pool	10.14.0.0/16
	External TEP Pool	10.114.114.0/24*
	Data Plane TEP IP	10.114.114.1/32

* POD Specific; Can be a smaller pool – see Wizard for addresses allocated

To add the second Pod in the ACI **Multi-Pod** setup, follow the steps below. If continuing immediately from the previous section, click **Add Physical Pod in the last step** and proceed directly to step 5 below.

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Inventory**.
3. From the left navigation pane, expand and select **Quick Start > Add Pod**.
4. From the right window, click **Add Pod**.
5. In the pop-up window for **Add Physical Pod** wizard, review the **Overview**. Collect the **Setup Information** for the new Pod, IP Connectivity and External TEP. For the parameters used in validating this CVD, see the Setup Information for Pod-2 section. Click **Get Started**.
6. In the **Step 2 > Pod Fabric** window of the wizard, for the new Pod, specify a **Pod ID** (for example, 2) and **Pod TEP Pool** (for example, 10.14.0.0/16).



Please make sure TEP Pool subnet is correct and not overlapping.

7. For each Spine switch in Pod-2 connecting to IPN devices, specify the Spine ID (for example, 211), interface (for example, e1/47), IP Address (for example, 10.14.11.1/30) and MTU (for example, 9216) that matches the MTU on the interfaces on the IPN devices that these interfaces connect to. ACI **Multi-Pod** requires a minimum of 9150 bytes but many Cisco devices includes the IP header in the MTU specified and therefore, 9216 bytes is used.
8. Click **[+]** to the right of the MTU field to add more interfaces.
9. Click **[+]** to the right of the Spine ID to add more Spine switches.

APIC

System | Tenants | Fabric | **STEP 2 > Pod Fabric**

Inventory | Fabric Policies

Inventory

- Quick Start
- Add Remote Leaf
- Add Pod
- Topology
- Pod 1
 - Pod Fabric Setup Policy
 - Fabric Membership
 - Disabled Interfaces and Deco
 - Duplicate IP Usage

Add Physical Pod

1. Overview | **2. Pod Fabric** | 3. External TEP

IP Connectivity

Every pod in Cisco ACI needs a pod ID. Choose a unique pod ID.

A pod uses a pool of addresses to allocate IPs for spines, leafs, and virtual leafs. This pool is called a TEP pool, and its addresses are distributed by the Cisco APIC using DHCP. Configure a TEP pool that does not overlap with existing TEP pools.

The interpod network (IPN) connects Cisco ACI locations to provide end-to-end network connectivity. To achieve this, spines need IP connectivity to the IPN.

Identify spines by entering their node IDs. For each spine, define the interfaces that are connected to the IPN and provide IP configuration for at least one interface for each spine. Multiple interfaces are supported. It is best to have the same MTU set on all spine-to-IPN interfaces. Configure the IPN to act as a DHCP relay pointing to Cisco APIC.

Pod Configuration

Pod ID:

Pod TEP Pool: [View existing TEP Pools](#)

Spine ID: + -

Interface	IPv4 Address	MTU (bytes)
<input type="text" value="1/47"/>	<input type="text" value="10.114.11.1/30"/>	<input type="text" value="9216"/> + -
<input type="text" value="1/48"/>	<input type="text" value="10.114.11.5/30"/>	<input type="text" value="9216"/> + -

Spine ID: + -

Interface	IPv4 Address	MTU (bytes)
<input type="text" value="1/47"/>	<input type="text" value="10.114.12.1/30"/>	<input type="text" value="9216"/> + -
<input type="text" value="1/48"/>	<input type="text" value="10.114.12.5/30"/>	<input type="text" value="9216"/> + -

Previous Cancel Next

Last Login Time: 2018-11-18T01:26 UTC-05:00

10. In the **Step 3 > External TEP** window of the wizard, for the Spine switches in Pod-2 connecting to the IPN devices, leave checkbox **Use Defaults** enabled. Specify the **External TEP Pool** for Pod-2 (for example, 10.114.114.0/24) and Router IDs (for example, 14.14.14.11, 14.14.14.12) for the Spines.

APIC

System

Tenants

Fabric

Inventory

Fabric Policies

Quick Start

Add Remote Leaf

Add Pod

Topology

Pod 1

Pod Fabric Setup Policy

Fabric Membership

Disabled Interfaces and Deco

Duplicate IP Usage

Add Physical Pod

STEP 3 > External TEP

1. Overview

2. Pod Fabric

3. External TEP

Pod

MP BGP EVPN

IPN

pPod

IP Connectivity

OSPF

Pod Configuration

External TEP addresses are used by the physical Pod to communicate with remote locations. Configure a subnet that is routable across the network connecting the different locations. The External TEP pool cannot overlap with other Pods internal or external TEP pools. The pool size should be between /27 and /22. The pool should be large enough to address all APICs, all Spines, all Border Leafs, Pod specific TEP addresses and Spine Route IDs. The wizard will automatically allocate addresses for Pod specific TEP addresses and Spine Router IDs from the External TEP pool. Proposed addresses can be modified but modified addresses must be outside of the External TEP pool.

Use Defaults: ☒

Pod:	Internal TEP Pool:	External TEP Pool:	Data Plane TEP IP:
1	10.13.0.0/16	10.113.113.0/24	10.113.113.1/32
2	10.14.0.0/16	<input type="text" value="10.114.114.0/24"/>	<input type="text" value="10.114.114.1/32"/>

Node: 211

Router ID:

Loopback Address:

Leave blank to use Router ID

Node: 212

Router ID:

Loopback Address:

Leave blank to use Router ID

Previous

Cancel

Finish

11. Click **Finish** to complete the Inter-Pod connectivity setup for the Spine switches in the second Pod or site (Pod-2).
12. In the **Summary** window, review the information provided.

Summary

✓ All done! You can view a summary what was done below.

Pod — IP Connectivity — OSPF — IPN — IP Connectivity — pPod

MP BGP EVPN

⚠ The configuration push triggered a restart of APIC web server. You might see a server-side error until a new connection is established.

⚠ OSPF interface profile is needed for OSPF connectivity between Pod to IPN, please configure OSPF interface profile here:
Tenants > infra > Networking > External Routed Network > Node Profile > Interface Profile

Configured

Pod SetupP: 2

Routed Outside: multipodL3Out

The following addresses are the internal TEP addresses of the APIC controllers. Configure the IPN DHCP relay to point to the following address to enable discovery of the Pod components.

APIC Name	Internal TEP Address
AA11-APIC-M2	10.13.0.1
AA11-APIC-M2	10.13.0.2

View JSON OK

13. (Optional) Click **View JSON** to save the JSON data for the configuration that was just completed.
14. Proceed to the next section to configure DHCP relay on Pod-2 IPN devices to point to Pod-1 APIC IP addresses listed in the above **Summary** window.

Configure DHCP Relay on IPN Devices

Per the recommendations from the Configuration Wizard [Summary](#) page in previous section, add DHCP relay statements on Pod-2 IPN devices. DHCP should be relayed to Pod-1 TEP IP Addresses and should match the addresses listed on the Configuration Wizard Summary page. The configuration should be added to the Spine-facing interfaces on Pod-2 IPN devices.



This was completed in the [Deploy Inter-Pod Network](#) section but verify the APIC IP addresses and the interfaces to which it is applied.

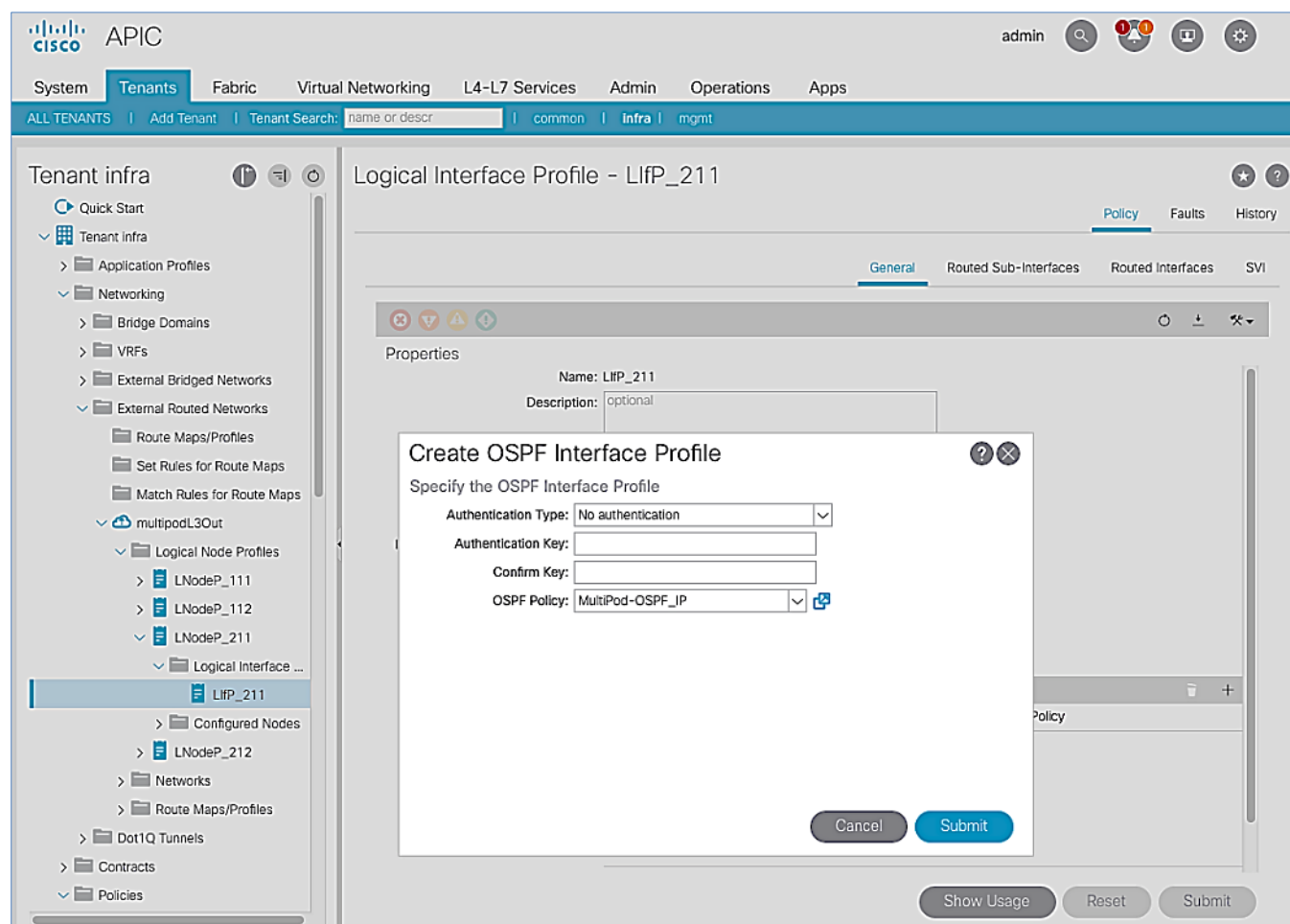
Proceed to the next section to configure the **OSPF Interface Profile** as per the message displayed on the Summary page.

Configure OSPF Interface Profile for Spines in Pod-2

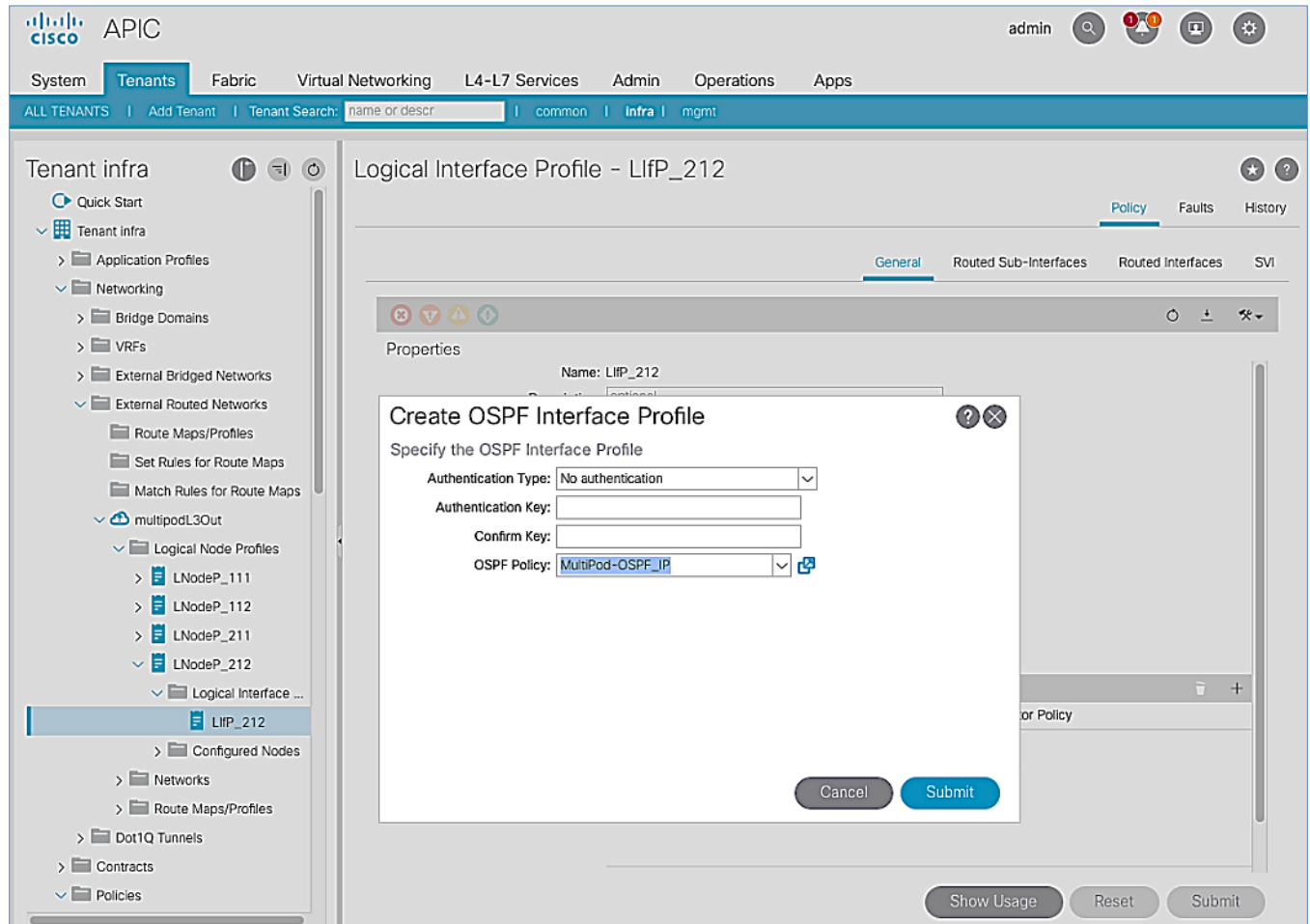
Per [the summary](#) of recommendations at the end of the Configuration Wizard for **Add Pod**, create an OSPF Interface Profile for all Spine switches that will connect to the IPN.

To create the OSPF Interface Profile, follow these steps:

1. From the top navigation menu, select **Tenants > infra**.
2. From the left navigation pane, expand and select Tenant Infra > Networking > External Routed Networks > multipodL3Out > Logical Node Profiles.
3. Select the Node profile (for example, LNodeP_211) for the **first** Pod-2 Spine switch.
4. Expand the Node profile for the selected node and select the profile for that Spine node. Right-click and select **Create OSPF Interface Profile** from the menu.
5. In the pop-up window for **Add Physical Pod** wizard, navigate to **Tenants > Infra** from the top navigation menu.
6. For the **OSPF Policy**, select the previously created policy from the drop-down list.



7. Click **Submit** to complete.
8. Repeat steps 1-7 for the **second** Spine node in Pod-2 as shown below.



9. Click **Submit** to complete.

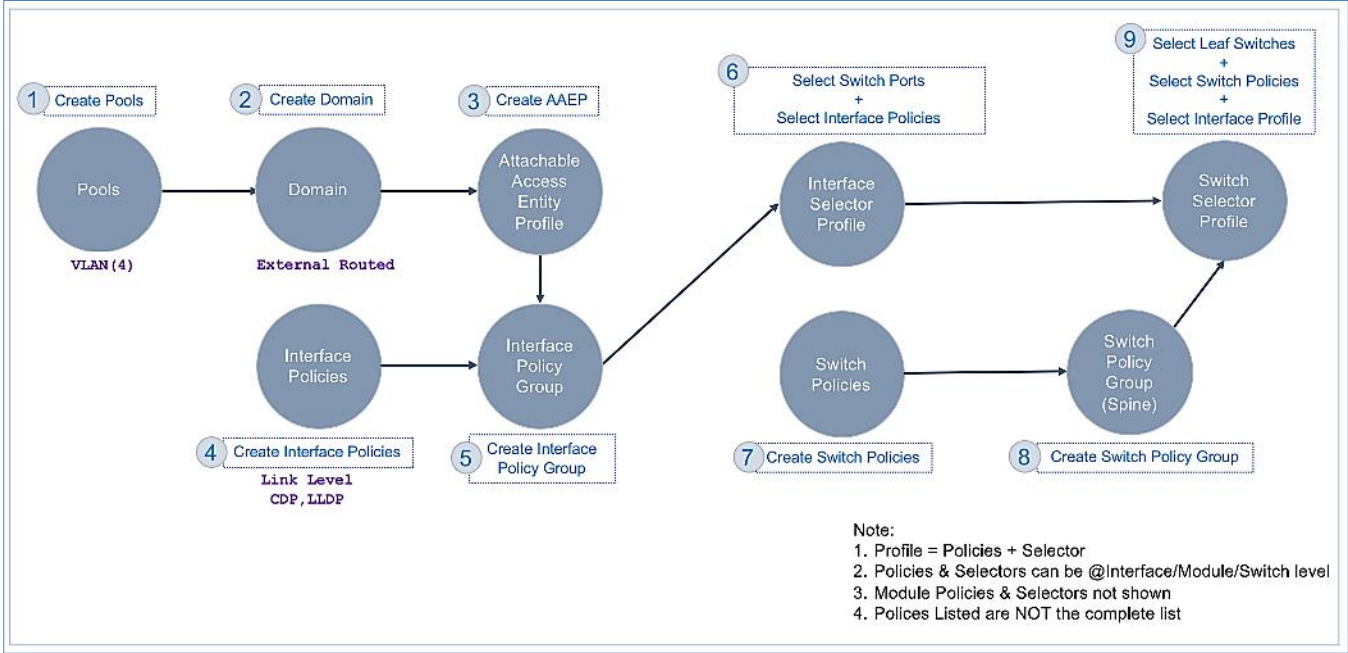
Setup Fabric Access Policies for Spine Switches in Pod-1

In ACI, access policies define the port configuration. In this section, access policies are configured for all interfaces on the spine switches in Pod-1 that connect to the IPN. The access policies enable connectivity between the Spine switches and IPN in Pod-1. The access policies are grouped and applied to specific interfaces and switches using interface and switch profiles respectively.

Deployment Overview

The deployment workflow for configuring Spines to connect to IPN is similar to configuring ACI Leaf switches for connectivity to access layer devices such as Cisco UCS and HyperFlex. The configuration in both cases is done through Fabric Access Policies. The workflow for creating Fabric Access Policies for connecting Spines to IPN devices in Pod-1 is shown in [Figure 10](#).

Figure 10 Fabric Access Policies – For Spine Switch Connectivity to IPN in Pod-1



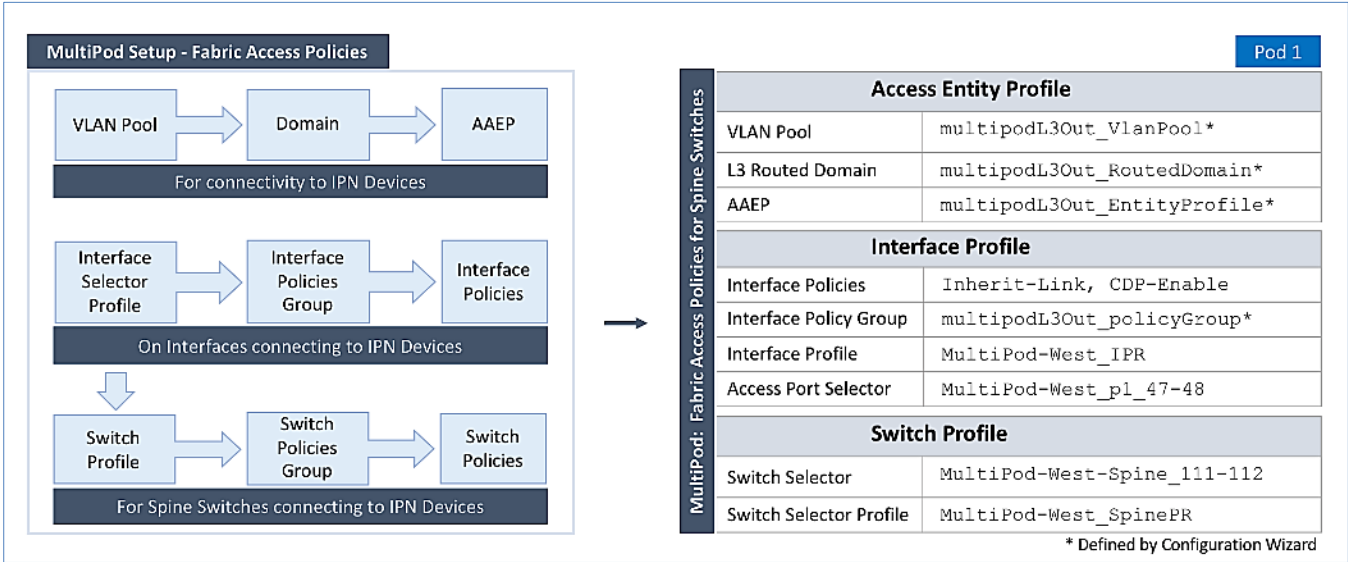
Setup Information

The information for configuring fabric access policies to connect Spine switches in Pod-1 to IPN is provided below.



VLAN Pool, L3 Routed Domain, AAEP, and Interface Policy Group listed below are configured by the Configuration Wizard during the Multi-Pod setup (see section titled Setup ACI Fabric for Multi—Pod – Using Configuration Wizard).

Figure 11 Setup Information – Fabric Access Policies on Pod-1 Spine Switches



Deployment Steps

Complete the procedures outlined in this section to configure access policies on Spine switch interfaces to enable connectivity to IPN in Pod-1. Unlike other access layer connections in this design, the access layer policies here are applied to interfaces on Spine switches and represent fabric-to-fabric connectivity across a L3 network.

Update Interface Policy Group

The interface policy group was created by the APIC configuration wizard as a part of the Multi-Pod setup. In this section, the policy group is updated to include some additional policies. The policies are among the pre-configured Fabric Access Policies completed earlier in the setup.

To update the interface policy group, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top menu, select **Fabric > Access Policies**.
3. From the left navigation pane, select and expand Policies > Interfaces > Spine Interfaces > Policy Groups.
4. Select the previously created policy group (for example, `multipodL3Out_policyGroup`).
5. In the right window pane, for **Link Level Policy**, select the **Inherit-Link** policy that was created earlier. For **CDP Policy**, select **CDP-Enabled**.



Enabling CDP is optional. LLDP should be enabled by default.

The screenshot displays the APIC GUI interface for configuring a policy group. The top navigation bar includes tabs for System, Tenants, Fabric, Virtual Networking, L4-L7 Services, Admin, Operations, and Apps. The 'Fabric' tab is active, and the 'Access Policies' sub-tab is selected. The left-hand navigation pane shows a tree structure under 'Policies' > 'Interfaces' > 'Spine Interfaces' > 'Policy Groups', with 'multipodL3Out_policyGroup' highlighted. The main content area shows the configuration for this policy group. The 'Policy' tab is active, displaying fields for Name, Description, Alias, Link Level Policy (set to 'Inherit-Link'), CDP Policy (set to 'CDP-Enabled'), MACsec Policy (set to 'select a value'), and Attached Entity Profile (set to 'multipodL3Out_EntityProfile'). At the bottom right, there are buttons for 'Show Usage', 'Reset', and 'Submit'.

6. Click Submit and Submit Changes to complete.

Create Interface (Selector) Profile for Spine Connectivity to IPN

The same interface profile can be re-used to configure other access layer connections that share the same interface selectors. In this design, Pod-2 Spine switches connect to the IPN on the same ports as Pod-1 switches and therefore will use this profile.

To create interface (selector) profile for the access layer connections from Spine switches to IPN in Pod-1, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top menu, select **Fabric > Access Policies**.
3. From the left navigation pane, select and expand Policies > Interfaces > Spine Interfaces > Profiles.
4. Right-click Profiles and select **Create Spine Interface Profile**.
5. In the **Create Spine Interface Profile** pop-up window, specify a profile **Name** (for example, `MultiPod-West_IPR`).

The screenshot shows the APIC GUI with the 'Create Spine Interface Profile' window open. The left navigation pane shows 'Policies' expanded, with 'Interfaces' > 'Spine Interfaces' > 'Profiles' selected. The top navigation bar shows 'Fabric' > 'Access Policies'. The 'Create Spine Interface Profile' window has the following fields:

- Name:** MultiPod-West_IPR
- Description:** optional
- Interface Selectors:** A table with columns 'Name' and 'Type'.

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

6. For the **Interface Selectors**, click the **[+]** on the right-side of the window to select access ports connecting to IPN devices. In the **Create Spine Access Port Selector** pop-up window, specify a selector **Name** (for example, `MultiPod-West_p1_47-48`). For the **Interface IDs**, add the ports that connect to IPN devices. For **Interface Policy Group**, select the previously created Interface Policy Group.

APIC admin

System Tenants **Fabric** Virtual Networking L4-L7 Services Admin Operations Apps

Inventory | Fabric Policy

Create Spine Interface Profile

Specify the profile Identity

Create Spine Access Port Selector

Specify the selector identity

Name: MultiPod-West_p1_47-48

Description: optional

Interface IDs: 1/47-48
valid values: All or Ranges. For Example: 1/13,1/15 or 1/22-1/24

Interface Policy Group: multipodL3Out_policyGrou

Cancel OK

- Click **OK** and **Submit** to complete.

Create Switch Profile for Spine connectivity to IPN

To create Switch profile for the access layer connections from Spine switches to IPN in Pod-1, follow these steps:

- Use a browser to navigate to the APIC GUI. Log in using **admin** account.
- From the top menu, select **Fabric > Access Policies**.
- From the left navigation pane, select and expand **Policies > Switches > Spine Switches > Profiles**.
- Right-click and select **Create Spine Profile**.
- In the **Create Spine Profile** pop-up window, specify a profile **Name** (for example, MultiPod-West_SpinePR). For the **Spine Selectors**, click the **[+]** on the right-side of the window to select the Spine switches to apply the interface profile to. Specify a selector **Name** (for example, MultiPod-West-Spine_111-112) and under the **Blocks** column, select the Spine Switch IDs from the drop-down list (for example, 111, 112). Click **Update**. Click **Next**.

APIC

admin

System

Tenants

Fabric

Virtual Networking

L4-L7 Services

Admin

Operations

Apps

Inventory

Fabric Policies

Access Policies

Policies

> Quick Start

> Switches

> Leaf Switches

> Spine Switches

> Profiles

> Policy Groups

> Modules

> Interfaces

> Policies

> Pools

> Physical and External Domains

Create Spine Profile

1. Profile

2. Associations

STEP 1 > Profile

Specify the profile Identity

Name: MultiPod-West_SpinePR

Description: optional

Spine Selectors:

Name	Blocks	Policy Group
MultiPod-West-Spine_111-112	111-112	select an option

Update

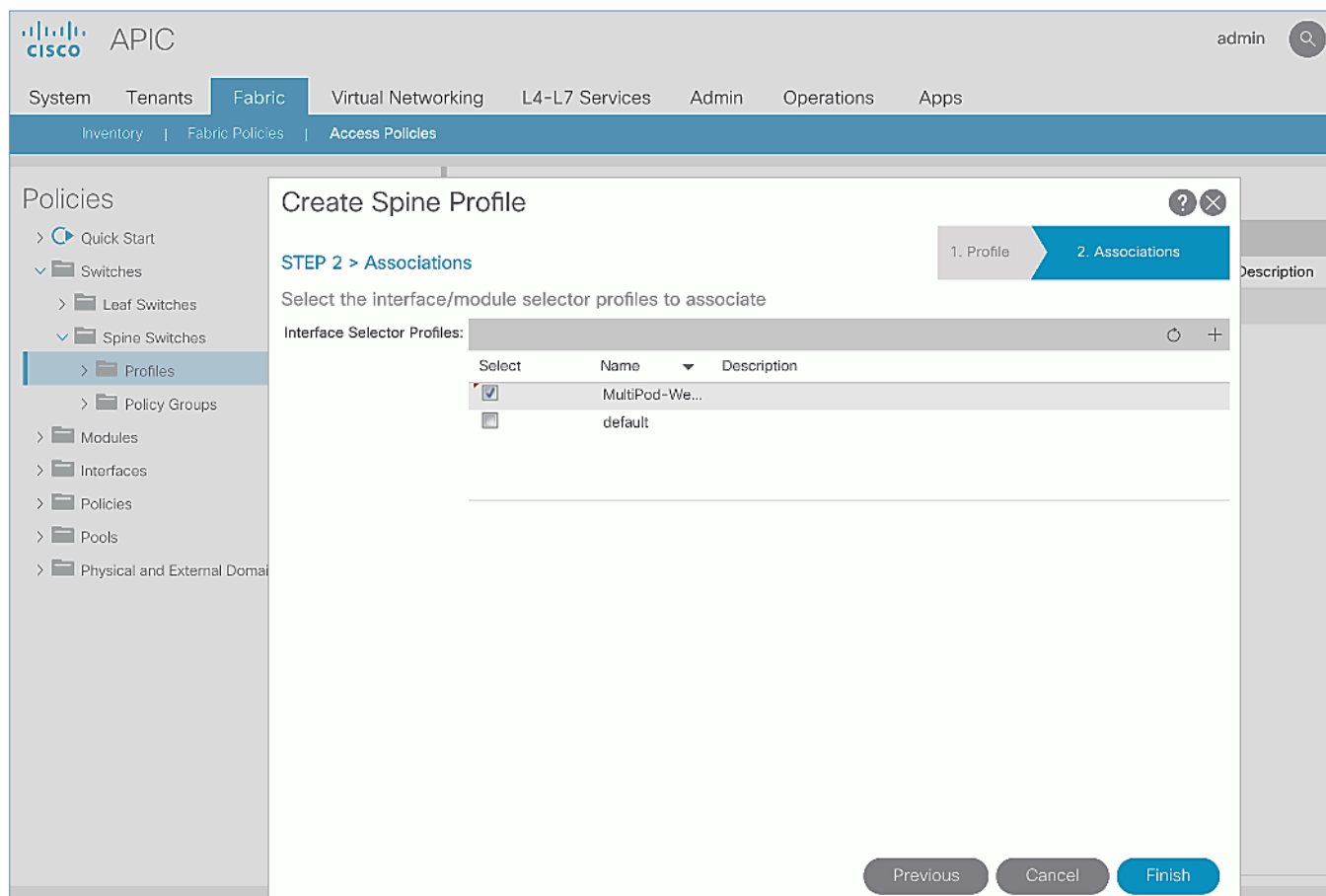
Cancel

Previous

Cancel

Next

6. In the **Step 2 > Associations** window, for the **Interface Selector Profile**, select the previously created Interface Profile.



- Click **Finish** to complete.

Deploy ACI Fabric in Pod-2

Deployment Overview

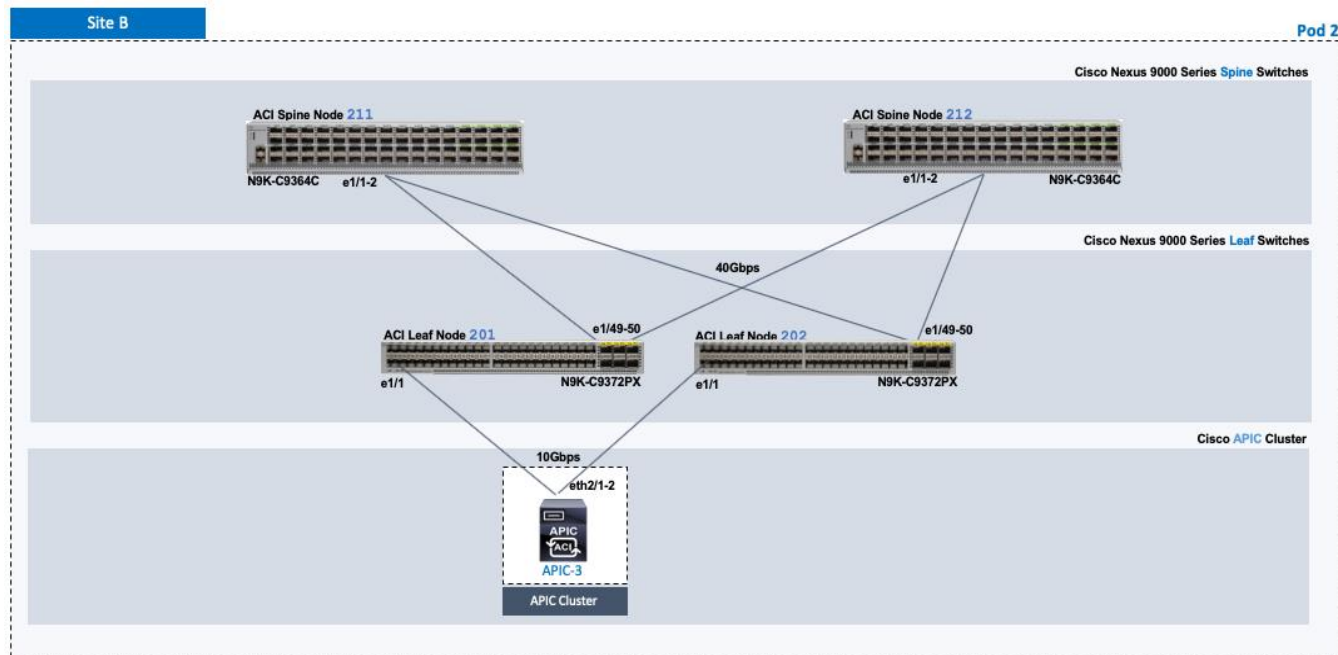
A high-level overview of the steps involved in deploying Pod-2 is summarized below:

- Complete the physical connectivity to connect all the devices in Pod-2. The fabric should have a minimum of two Spine and Leaf switches, and three APICs in a cluster. Since the APIC cluster is part of an ACI Multi-Pod fabric, two APICs are deployed in Pod-1 and one in Pod-2. CIMC management to the APIC in Pod-2 to access the console and out-of-band management connectivity to the switches and APIC should also be in place.
- Deploy Spine and Leaf switches in Pod-1. APICs are connected to the Leaf switches. The leaf switches are also border leaf switches that enable connectivity to networks outside the ACI fabric from Pod-1.
- Setup and configure the third APIC in the cluster. The first two APICs are deployed in Pod-1.
- Configure Out-of-Management (OOB) IP addresses for all switches in Pod-2.
- Configure Pod for NTP, BGP Route Reflector function, Fabric Profiles, and so on.

Physical Connectivity

Complete the cabling required to deploy Pod-2 in the ACI Multi-Pod Fabric as shown in [Figure 12](#). The connectivity for OOB management for all the devices and CIMC management for the third APIC (not shown below) should also be completed.

Figure 12 Physical Connectivity Details for Pod-2



Deploy Spine and Leaf Switches in Pod-2

When the Multi-Pod setup is complete, Pod-2 Spine and Leaf switches should be discoverable by the APIC(s) in the first site. In this section, verify the Spines in Pod-2 are being discovered by the APIC(s) in Pod-1. They will be discovered if the IPN connectivity and Multi-Pod setup is correct. Once discovered, the Spines and Leaf switches are added to the ACI Fabric.

Prerequisites

The following are the prerequisites to deploy the spine and leaf switches in Pod-2:

- Confirm that all Spine and Leaf switches in Pod-2 are running software that is compatible with the APIC release running in the ACI Fabric. Failure to do so can impact the discovery and addition of these switches to the Fabric.
- The Spine switches must be connected to at least one Leaf switch before it can be discovered. The Spine switch must be able to see the Leaf switch via LLDP.

Deployment Overview

The high-level steps for deploying Pod-2 switches to the ACI Fabric are summarized below:

- Discover and add Spine switches in Pod-2
- Discover and add Leaf switches in Pod-2
- Configure Out-of-band Management for Pod-2 switches
- Configure NTP for Pod-2 using Out-of-Band Management
- Update BGP Route Reflector Policy with Pod-2 Spine Switches

Table 24 Leaf Switches in Pod-2

Leaf Switches in Pod-2	General	Node ID	Node Names	OOB Management EPG	OOB Management IP	OOB Gateway
	Pod ID: 2	201	BB06-9372PX-WEST-1	default	172.26.164.117/24	172.26.164.254
	Role: Leaf					
	Rack Name (Optional): BB06	202	BB06-9372PX-WEST-2	default	172.26.164.118/24	172.26.164.254

Table 25 Spine Switches in Pod-2

Spine Switches in Pod-2	General	Node ID	Node Names	OOB Management EPG	OOB Management IP	OOB Gateway
	Pod ID: 2	211	BB06-9364C-WEST-1	default	172.26.164.119/24	172.26.164.254
	Role: Spine					
	Rack Name (Optional): BB06	212	BB06-9364C-WEST-2	default	172.26.164.120/24	172.26.164.254

Verify that APIC Can See the Spine Switches In Pod-2

To verify that APIC can see Leaf and Spine switches in Pod-2 to the ACI Fabric, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using **admin** account.
2. From the top menu, select **Fabric > Inventory**.
3. From the left navigation pane, navigate to **Fabric Membership**.
4. In the right navigation pane, go to the **Nodes Pending Registration** tab.

The screenshot shows the Cisco APIC GUI. The top navigation bar includes 'System', 'Tenants', 'Fabric' (selected), 'Virtual Networking', 'L4-L7 Services', 'Admin', 'Operations', and 'Apps'. The 'Fabric' section has sub-tabs: 'Inventory', 'Fabric Policies', and 'Access Policies'. The left sidebar shows the 'Inventory' section with a tree view including 'Quick Start', 'Topology', 'Pod 2', 'Pod 1', 'Pod Fabric Setup Policy', 'Fabric Membership' (selected), 'Disabled Interfaces and Decommissioned', and 'Duplicate IP Usage'. The main content area is titled 'Fabric Membership' and has tabs for 'Registered Nodes', 'Nodes Pending Registration' (selected), 'Unreachable Nodes', and 'Unmanaged Fabric Nodes'. Below the tabs, there are two large zero counts: '0 Unsupported' and '0 Undiscovered'. At the bottom, a table lists nodes pending registration:

Serial Number	Pod ID	Node ID	RL TEP Pool	Name	Role	Supported Model	SSL Certificate	Status
FDO22182Q9G	1	0	0		spine	yes	n/a	
FDO221914JV	1	0	0		spine	yes	n/a	

5. Confirm that you see all the Spine switches that are directly connected to the IPN devices.

- Identify the spine switches based on their serial numbers and collect the corresponding setup information. Proceed to the next section to configure the Spine switches.

Add Spine Switches in Pod-2 to the ACI Fabric

To add spine switches in Pod-2 to the ACI fabric, follow these steps:

- Use a browser to navigate to the APIC GUI. Log in using **admin** account.
- From the top menu, select **Fabric > Inventory**.
- From the left navigation pane, navigate to **Fabric Membership**.
- In the right navigation pane, go to the **Nodes Pending Registration** tab.
- Identify the Serial number of the Spine switch in Pod-2 that should be configured **first**.
- Select the switch from the list. Right-click and select **Register**.

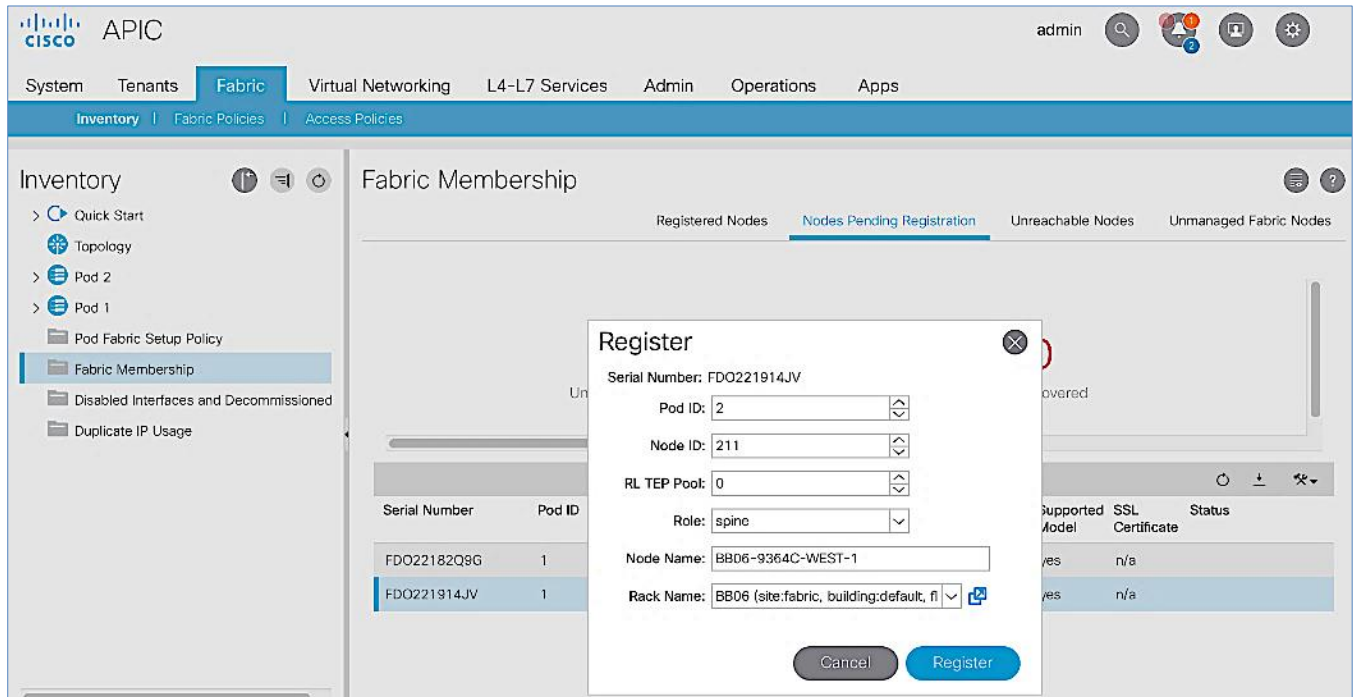
The screenshot shows the Cisco APIC GUI. The top navigation bar includes 'System', 'Tenants', 'Fabric', 'Virtual Networking', 'L4-L7 Services', 'Admin', 'Operations', and 'Apps'. The 'Fabric' tab is selected, and the 'Inventory' sub-tab is active. The left sidebar shows the 'Inventory' section with 'Fabric Membership' selected. The main content area is titled 'Fabric Membership' and has tabs for 'Registered Nodes', 'Nodes Pending Registration' (which is active), 'Unreachable Nodes', and 'Unmanaged Fabric Nodes'. Below the tabs, there are two large red '0' indicators for 'Unsupported' and 'Undiscovered' nodes. A table lists nodes pending registration:

Serial Number	Pod ID	Node ID	RL TEP Pool	Name	Role	Supported Model	SSL Certificate	Status
FDO22182Q9G	1	0	0		spine	yes	n/a	
FDO221914JV	1	0	0		spine	yes	n/a	

A context menu is open over the row for 'FDO221914JV', showing the following options:

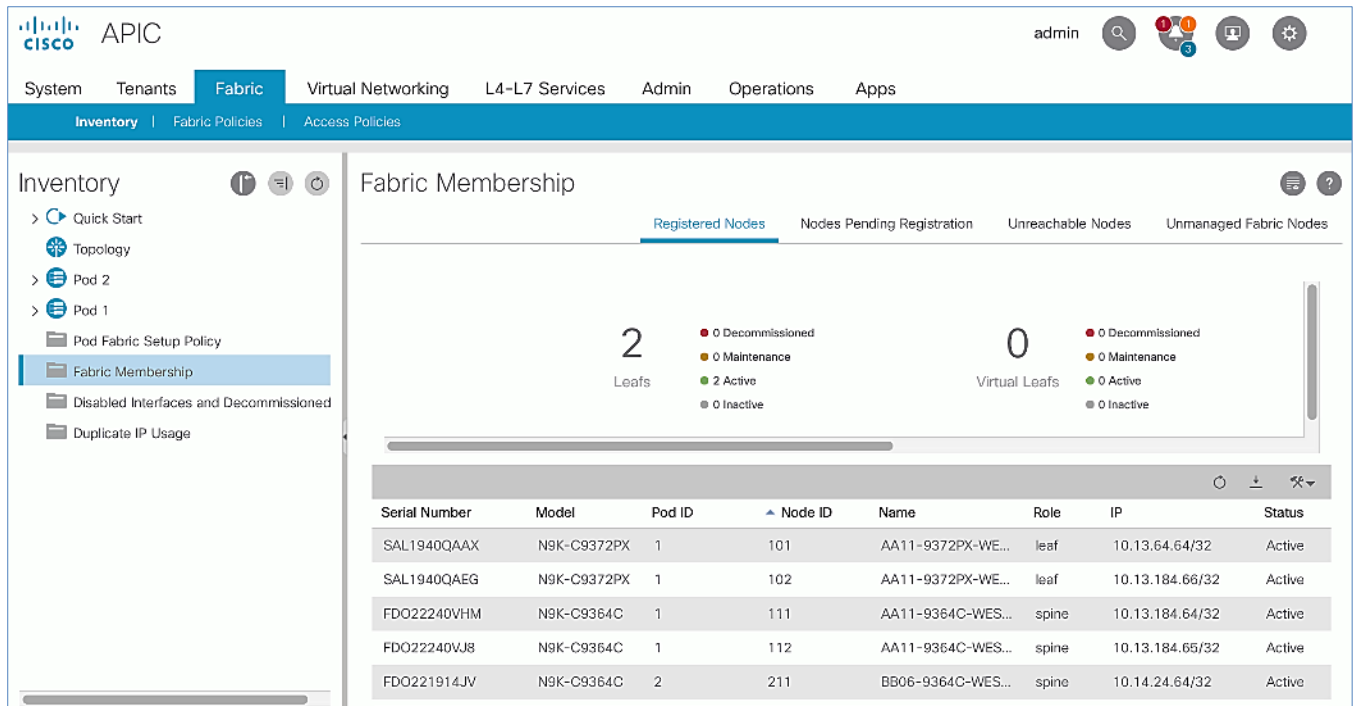
- Register
- Edit Node and Rack Names
- Remove From Controller

- In the **Register** pop-up window, specify the **Pod ID** (for example, 2), **Node Id** (for example, 211), **Node Name** for example, BB06-9364C-WEST-1) and **Rack Name** (for example, BB06).



8. Click **Register**.

9. Click the **Registered Nodes** tab.



10. The newly configured Spine should show up in the registered list. It should transition to **Active** status after a few minutes.

11. In the right navigation pane, go to the **Nodes Pending Registration** tab.

12. You should now see the remaining Spine switches that need to be registered and configured. Note that you will also start to see any discovered Leaf switches that were connected to the Pod-2 Spine. You will configure Leaf switches in the next section after all the Spine switches have been configured.

The screenshot shows the Cisco APIC interface with the 'Fabric' tab selected. The 'Fabric Membership' section is active, showing the 'Nodes Pending Registration' tab. The interface displays two large red zeros for 'Unsupported' and 'Undiscovered' nodes. Below, a table lists three nodes:

Serial Number	Pod ID	Node ID	RL TEP Pool	Name	Role	Supported Model	SSL Certificate	Status
FDO22182Q9G	1	0	0		spine	yes	n/a	
SAL1913CJXR	1	0	0		leaf	yes	n/a	
SAL1914CN42	1	0	0		leaf	yes	n/a	

13. Select the next Spine switch in the list and repeat the above steps to **register** the switch.

The screenshot shows the Cisco APIC interface with the 'Fabric' tab selected. The 'Fabric Membership' section is active, showing the 'Nodes Pending Registration' tab. A 'Register' dialog box is open, displaying the following information:

- Serial Number: FDO22182Q9G
- Pod ID: 2
- Node ID: 212
- RL TEP Pool: 0
- Role: spine
- Node Name: BB06-9364C-WEST-2
- Rack Name: BB06 (site: fabric, building: default, fl)

The dialog box has 'Cancel' and 'Register' buttons at the bottom.

14. Both Pod-2 Spine switches will now show up under the **Registered Nodes** tab.

The screenshot shows the Cisco APIC GUI. The top navigation bar includes 'System', 'Tenants', 'Fabric', 'Virtual Networking', 'L4-L7 Services', 'Admin', 'Operations', and 'Apps'. The 'Fabric' tab is selected, and the 'Inventory' sub-tab is active. The left sidebar shows the 'Inventory' section with 'Fabric Membership' highlighted. The main content area is titled 'Fabric Membership' and shows the 'Registered Nodes' tab. It displays a summary of 2 Active Leaf nodes and 0 Virtual Leaf nodes. Below the summary is a table of registered nodes.

Serial Number	Model	Pod ID	Node ID	Name	Role	IP	Status
SAL1940QAAX	N9K-C9372PX	1	101	AA11-9372PX-WEST-1	leaf	10.13.64.64/32	Active
SAL1940QAEQ	N9K-C9372PX	1	102	AA11-9372PX-WEST-2	leaf	10.13.184.66/32	Active
FDO22240VHM	N9K-C9364C	1	111	AA11-9364C-WEST-1	spine	10.13.184.64/32	Active
FDO22240VJ8	N9K-C9364C	1	112	AA11-9364C-WEST-2	spine	10.13.184.65/32	Active
FDO221914JV	N9K-C9364C	2	211	BB06-9364C-WEST-1	spine	10.14.24.64/32	Active
FDO22182Q9G	N9K-C9364C	2	212	BB06-9364C-WEST-2	spine	10.14.24.65/32	Discoverin

15. In the **Nodes Pending Registration** tab, you should now see all the Leaf switches that were discovered as a result of registering the Spine switches that they connect to.

Upgrade Firmware on Spine Switches in Pod-2 (Optional)

To upgrade the firmware on the spine switches in Pod-2, follow these steps:

1. From the top menu, navigate to **Admin > Firmware**.
2. Select the tabs for **Infrastructure > Nodes**.
3. Check the **Current Firmware** version column for the newly deployed Spine switches to verify they are compatible with the APIC version running.
4. If an upgrade is **not** required, proceed to the next section but if an upgrade is required, use the product documentation to upgrade the switches.

Verify that APIC Can See the Leaf Switches In Pod-2

To verify that APIC can see the leaf switches in Pod-2, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using **admin** account.
2. From the top menu, select **Fabric > Inventory**.
3. From the left navigation pane, navigate to **Fabric Membership**.
4. In the right navigation pane, go to the **Nodes Pending Registration** tab.

The screenshot shows the Cisco APIC GUI. The top navigation bar includes 'System', 'Tenants', 'Fabric', 'Virtual Networking', 'L4-L7 Services', 'Admin', 'Operations', and 'Apps'. The 'Fabric' tab is selected. Below this, there are sub-tabs: 'Inventory', 'Fabric Policies', and 'Access Policies'. The left sidebar shows the 'Inventory' section with a tree view containing 'Quick Start', 'Topology', 'Pod 2', 'Pod 1', 'Pod Fabric Setup Policy', 'Fabric Membership' (selected), 'Disabled Interfaces and Decommissioned', and 'Duplicate IP Usage'. The main content area is titled 'Fabric Membership' and has four tabs: 'Registered Nodes', 'Nodes Pending Registration' (active), 'Unreachable Nodes', and 'Unmanaged Fabric Nodes'. The 'Nodes Pending Registration' tab shows two large '0' counts for 'Unsupported' and 'Undiscovered' nodes. Below this is a table with the following data:

Serial Number	Pod ID	Node ID	RL TEP Pool	Name	Role	Supported Model	SSL Certificate	Status
SAL1913CJXR	1	0	0		leaf	yes	n/a	
SAL1914CN42	1	0	0		leaf	yes	n/a	

- Confirm that you see all the Leaf switches in Pod-2.
- Identify the Leaf switches based on their serial numbers and collect the corresponding setup information. Proceed to the next section to configure the Leaf switches.

Add Leaf Switches in Pod-2 to the ACI Fabric

To add the leaf switches in Pod-2 to the ACI fabric, follow these steps:

- Use a browser to navigate to the APIC GUI. Log in using **admin** account.
- From the top menu, select **Fabric > Inventory**.
- From the left navigation pane, navigate to **Fabric Membership**.
- In the right navigation pane, go to the **Nodes Pending Registration** tab.
- Identify the Serial number of the Leaf switch in Pod-2 that should be configured **first**.
- Select the switch from the list. Right-click and select **Register**.

The screenshot shows the APIC interface with the 'Fabric' tab selected. The 'Fabric Membership' section is active, and the 'Nodes Pending Registration' tab is chosen. The 'Unsupported' and 'Undiscovered' counters both show 0. A table lists nodes, with the first node, SAL1913CJXR, selected. A context menu is open over this node, showing options: 'Register', 'Edit Node and Rack Names', and 'Remove From Controller'.

Serial Number	Pod ID	Node ID	RL TEP Pool	Name	Role	Supported Model	SSL Certificate	Status
SAL1913CJXR			0		leaf	yes	n/a	
SAL1914CN42			0		leaf	yes	n/a	

- In the **Register** pop-up window, specify the Pod ID (for example, 2), Node Id (for example, 201), Node Name for example, BB06-9372PX-WEST-1) and Rack Name (for example, BB06).

The screenshot shows the APIC interface with the 'Fabric Membership' section active. The 'Nodes Pending Registration' tab is chosen. A 'Register' pop-up window is open, displaying the following fields:

- Serial Number: SAL1913CJXR
- Pod ID: 2
- Node ID: 201
- RL TEP Pool: 0
- Role: leaf
- Node Name: BB06-9372PX-WEST-1
- Rack Name: BB06 (site: fabric, building: default, fl)

The 'Register' button is highlighted in blue.

- Click **Register**.
- Click the **Registered Nodes** tab and the newly configured Leaf switch should now show up in the registered list. It will transition to **Active** after a few minutes.

The screenshot shows the APIC interface with the **Fabric** tab selected. The **Fabric Membership** page is displayed, showing the **Registered Nodes** tab. The left navigation pane includes **Inventory** with sub-items like **Quick Start**, **Topology**, **Pod 2**, **Pod 1**, **Pod Fabric Setup Policy**, **Fabric Membership** (selected), **Disabled Interfaces and Decommissioned**, and **Duplicate IP Usage**.

The **Fabric Membership** page displays the following statistics:

- Registered Nodes:** 3 Leafs (0 Decommissioned, 0 Maintenance, 3 Active, 0 Inactive)
- Virtual Leafs:** 0 (0 Decommissioned, 0 Maintenance, 0 Active, 0 Inactive)

The table below lists the registered nodes:

Serial Number	Model	Pod ID	Node ID	Name	Role	IP	Status
SAL1940QAAX	N9K-C9372PX	1	101	AA11-9372PX-WEST-1	leaf	10.13.64.64/32	Active
SAL1940QAEQ	N9K-C9372PX	1	102	AA11-9372PX-WEST-2	leaf	10.13.184.65/32	Active
FDO22240VHM	N9K-C9364C	1	111	AA11-9364C-WEST-1	spine	10.13.184.64/32	Active
FDO22240VJB	N9K-C9364C	1	112	AA11-9364C-WEST-2	spine	10.13.184.65/32	Active
SAL1913CJXR	N9K-C9372PX	2	201	BB06-9372PX-WEST-1	leaf	10.14.32.64/32	Active
FDO221914JV	N9K-C9364C	2	211	BB06-9364C-WEST-1	spine	10.14.24.64/32	Active
FDO22182Q9G	N9K-C9364C	2	212	BB06-9364C-WEST-2	spine	10.14.24.65/32	Active

10. In the right navigation pane, click the **Nodes Pending Registration** tab.

11. Select the next Leaf switch in the list and repeat steps 1-10 to **register** the switch.

The screenshot shows the APIC interface with the **Fabric** tab selected. The **Fabric Membership** page is displayed, showing the **Nodes Pending Registration** tab. The left navigation pane is the same as in the previous screenshot.

The **Fabric Membership** page displays the following statistics:

- Nodes Pending Registration:** 0 Undiscovered

The **Register** dialog box is open, showing the following fields:

- Serial Number: SAL1914CN42
- Pod ID: 2
- Node ID: 202
- RL TEP Pool: 0
- Role: leaf
- Node Name: BB06-9372PX-WEST-2
- Rack Name: BB06 (site: fabric, building: default, fl)

The dialog box has **Cancel** and **Register** buttons.

12. All registered Leaf switches will show up under the **Registered Nodes** tab.

The screenshot shows the APIC GUI with the 'Fabric' tab selected. The 'Fabric Membership' page is displayed, showing a summary of 4 registered leaf nodes and 0 virtual leaf nodes. The nodes are categorized by status: 0 Decommissioned, 0 Maintenance, 4 Active, and 0 Inactive.

Serial Number	Model	Pod ID	Node ID	Name	Role	IP	Status
SAL1940QAAX	N9K-C9372PX	1	101	AA11-9372PX-WEST-1	leaf	10.13.64.64/32	Active
SAL1940QAEQ	N9K-C9372PX	1	102	AA11-9372PX-WEST-2	leaf	10.13.184.66/32	Active
FDO22240VHM	N9K-C9364C	1	111	AA11-9364C-WEST-1	spine	10.13.184.64/32	Active
FDO22240VJ8	N9K-C9364C	1	112	AA11-9364C-WEST-2	spine	10.13.184.65/32	Active
SAL1913CJXR	N9K-C9372PX	2	201	BB06-9372PX-WEST-1	leaf	10.14.32.64/32	Active
SAL1914CN42	N9K-C9372PX	2	202	BB06-9372PX-WEST-2	leaf	10.14.32.65/32	Active
FDO221914JV	N9K-C9364C	2	211	BB06-9364C-WEST-1	spine	10.14.24.64/32	Active
FDO22182Q9G	N9K-C9364C	2	212	BB06-9364C-WEST-2	spine	10.14.24.65/32	Active

Upgrade Firmware on Leaf Switches in Pod-2 (Optional)

To upgrade the firmware on the leaf switches in Pod-2, follow these steps:

1. From the top menu, navigate to **Admin > Firmware**.
2. Select the tabs for **Infrastructure > Nodes**.
3. Check the **Current Firmware** version column for the newly deployed Leaf switches to verify they are compatible with the APIC version running.
4. If an upgrade is **not** required, proceed to the next section but if an upgrade is required, use the product documentation to upgrade the switches.

Configure Out-of-Band Management for Pod-2 Switches

To configure out-of-band Management for Pod-2 Spine and Leaf switches, follow these steps using the setup information in [Table 24](#) and [Table 25](#) :

1. Use a browser to navigate to the APIC GUI. Log in using **admin** account.
2. From the top menu, select **Tenants > mgmt**.
3. From the left navigation pane, expand and select Tenant mgmt > Node Management Addresses > Static Node Management Addresses.
4. Right-click and select Create Static Node Management Addresses.

The screenshot shows the Cisco APIC interface with the 'Tenants' tab selected. The left sidebar shows the 'Tenant mgmt' menu with 'Static Node Management Addresses' highlighted. The main content area displays a table of static node management addresses.

Node ID	Name	Type	EPG	IPv4 Address	IPv4 Gateway	IPv6 Address
pod-1/node-101	AA11-9372PX-WEST-1	Out-Of-Band	default	172.26.163.117/24	172.26.163.254	::
pod-1/node-102	AA11-9372PX-WEST-2	Out-Of-Band	default	172.26.163.118/24	172.26.163.254	::
pod-1/node-111	AA11-9364C-WEST-1	Out-Of-Band	default	172.26.163.119/24	172.26.163.254	::
pod-1/node-112	AA11-9364C-WEST-2	Out-Of-Band	default	172.26.163.120/24	172.26.163.254	::

- In the **Create Static Node Management Addresses** pop-up window, specify a **Node Range** (for example, 201–202), for **Config**: select the box for **Out-of-Band Addresses**.
- For **Out-of-Band Management EPG**, select **default** from the drop-down list.
- Specify the **Out-of-Band Management IPv4 Address** for the first node in the specified range.
- Specify the Out-of-Band Management IPv4 Gateway.

The screenshot shows the 'Create Static Node Management Addresses' pop-up window. The 'Node Range' is set to 201 - 202. The 'Config' section has 'Out-Of-Band Addresses' selected. The 'Out-Of-Band Addresses' section shows the 'Out-Of-Band Management EPG' set to 'default'. The 'Out-Of-Band IPv4 Address' is 172.26.164.117/24, and the 'Out-Of-Band IPv4 Gateway' is 172.26.164.254. The 'Out-Of-Band IPv6 Address' and 'Out-Of-Band IPv6 Gateway' fields are empty. The 'Submit' button is highlighted.

- Click **Submit** to complete.

10. Click **Yes** in the **Confirm** pop-up window to assign the IP address to the range of nodes specified.
11. Repeat steps 1-10 for the remaining Spine and Leaf switches in Pod-2.

The screenshot shows the Cisco APIC GUI with the 'Tenants' tab selected. The left sidebar shows the 'Tenant mgmt' menu with 'Static Node Management Addresses' selected. The main panel displays a table titled 'Static Node Management Addresses' with the following data:

Node ID	Name	Type	EPG	IPv4 Address	IPv4 Gateway	IF A
pod-1/node-101	AA11-9372PX-WEST-1	Out-Of-Band	default	172.26.163.117/24	172.26.163.254	::
pod-1/node-102	AA11-9372PX-WEST-2	Out-Of-Band	default	172.26.163.118/24	172.26.163.254	::
pod-1/node-111	AA11-9364C-WEST-1	Out-Of-Band	default	172.26.163.119/24	172.26.163.254	::
pod-1/node-112	AA11-9364C-WEST-2	Out-Of-Band	default	172.26.163.120/24	172.26.163.254	::
pod-2/node-201	BB06-9372PX-WEST-1	Out-Of-Band	default	172.26.164.117/24	172.26.164.254	::
pod-2/node-202	BB06-9372PX-WEST-2	Out-Of-Band	default	172.26.164.118/24	172.26.164.254	::
pod-2/node-211	BB06-9364C-WEST-1	Out-Of-Band	default	172.26.164.119/24	172.26.164.254	::
pod-2/node-212	BB06-9364C-WEST-2	Out-Of-Band	default	172.26.164.120/24	172.26.164.254	::

The switches can now be accessed directly using SSH.

Configure NTP for Pod-2 using Out-of-Band Management

To configure NTP for Pod-2, follow these steps using the setup information provided below:

- **NTP Policy Name:** Pod2-West-NTP_Policy
 - **NTP Server:** 172.26.164.254
 - **Management EPG:** default (Out-of-Band)
1. Use a browser to navigate to the APIC GUI. Log in using **admin** account.
 2. From the top menu, select **Fabric > Fabric Policies**.
 3. From the left navigation pane, navigate to **Policies > Pod > Date and Time**.
 4. Right-click and select **Create Date and Time Policy**.
 5. In the **Create Date and Time Policy** pop-up window, specify a **Name** for Pod-2's NTP Policy. Verify that the **Administrative State** is **enabled**.

APIC admin

System | Tenants | **Fabric** | Virtual Networking | L4-L7 Services | Admin | Operations | Apps

Inventory | **Fabric Policies** | App

Create Date And Time Policy

STEP 1 > Identity

Specify the information about the Date/Time Policy

Name:

Description:

Administrative State: ☐ ☒

Server State: ☐ ☒

Authentication State: ☐ ☒

Previous Cancel **Next**

Policies

- Quick Start
- Pods
 - Policy Groups
 - Profiles
 - Switches
 - Modules
 - Interfaces
 - Policies
 - Pod
 - Date and Time**
 - Policy Fabric2_Policy
 - Policy Pod1-West-NTP_Pol...
 - Policy default
 - default
 - SNMP
 - Management Access
 - ISIS Policy default

6. Click **Next**.

7. In **Step 2 > NTP Servers**, add NTP server(s) for Pod-2 using the **[+]** to the right of the list of servers.

8. In the **Create Providers** pop-up window, specify the Hostname/IP of the NTP server in the **Name** field. If multiple NTP Providers are being created for Pod-2, select the checkbox for **Preferred** when creating the preferred provider. For the **Management EPG**, select **default (Out-of-Band)** from the drop-down list.

APIC admin

System | Tenants | **Fabric** | Virtual Networking | L4-L7 Services | Admin | Operations | Apps

Inventory | **Fabric Policies** | App

Create Date And Time Policy

STEP 2 > NTP Servers

Create Providers

Specify the information about the NTP Server

Name:

Description:

Preferred: ☒

Minimum Polling Interval:

Maximum Polling Interval:

Management EPG:

Cancel **OK**

Previous Cancel **Finish**

Policies

- Quick Start
- Pods
 - Policy Groups
 - Profiles
 - Switches
 - Modules
 - Interfaces
 - Policies
 - Pod
 - Date and Time**
 - Policy Fabric2_Policy
 - Policy Pod1-West-NTP_Pol...
 - Policy default
 - default
 - SNMP
 - Management Access
 - ISIS Policy default

9. Click **OK**.

10. Click Finish.



The NTP policy is not in effect until it is applied using a Pod Profile.

Update BGP Route Reflector Policy for Pod-2

In an ACI fabric with multiple Spine switches, a pair of Spine switches are configured as Route Reflectors (RR) to redistribute routes from external domains into the fabric. In a Multi-Pod ACI fabric, each Pod has a pair of RR nodes. This section provides enabling the RR functionality on Spine switches in Pod-2.

To enable BGP Route Reflector functionality on Spine switches in Pod-2, follow these steps using the setup information provided below:

- BGP Route-Reflector Policy Name: `default`
 - Pod-2 Spine ID: 211, 212
1. Use a browser to navigate to the APIC GUI. Log in using **admin** account.
 2. From the top menu, select **System > System Settings**.
 3. From the left navigation pane, navigate to **BGP Route Reflector**.
 4. In the right window pane, in the **Route Reflector Nodes** section, click the **[+]** on the right to **Create Route Reflector Node**.
 5. In the **Create Route Reflector Node** pop-up window, for **Spine Node**, specify the **Node ID** (for example, 211) for the first Spine in Pod-2.

The screenshot shows the APIC GUI with the 'System Settings' tab selected. The left navigation pane shows 'BGP Route Reflector' under 'System Settings'. The main pane displays the 'BGP Route Reflector Policy - BGP Route Reflector' configuration page. The 'Properties' section shows the policy name as 'default' and the autonomous system number as '201'. The 'Route Reflector Nodes' table lists one node with Pod ID 1, Node ID 111, and Node Name AA11-9364C-WEST-1. A 'Create Route Reflector Node' pop-up window is open, prompting for the 'Spine Node' (Node ID 211) and 'Description' (Spine-1 in Pod-2).

System Settings

BGP Route Reflector Policy - BGP Route Reflector

Properties

Name: default
 Description: optional
 Autonomous System Number: 201

Route Reflector Nodes:

Pod ID	Node ID	Node Name	Description
1	111	AA11-9364C-WEST-1	Spine-1 In Pod-1

Create Route Reflector Node

Specify route reflector node EP id

Spine Node: 211
 Description: Spine-1 in Pod-2

Cancel Submit

6. Click **Submit**.
7. Repeat steps 1-6 to add second Spine in Pod-2.
8. You should now see two Spines as Route Reflectors for each Pod in the deployment.

System Set **BGP Route Reflector Policy - BGP Route Reflector**

Properties

Name: default

Description: optional

Autonomous System Number: 201

Route Reflector Nodes:

Pod ID	Node ID	Node Name	Description
1	111	AA11-9364C-WEST-1	Spine-1 in Pod-1
1	112	AA11-9364C-WEST-2	Spine-2 in Pod-1
2	211	BB06-9364C-WEST-1	Spine-1 in Pod-2
2	212	BB06-9364C-WEST-2	Spine-2 in Pod-2

External Route Reflector Nodes:

No items have been found.
Select Actions to create a new item.

Show Usage Reset Submit

Update Pod Profile to Apply Pod Policies

In ACI, Pod Policies (for example, BGP Route Reflector policy from previous section) are applied through a Pod Profile. A separate Pod Policy Group is used to group policies for each Pod and then they are applied using the Pod Profile. In this design, different NTP servers are used in each Pod. This policy is applied to Pod-2 policy group and then applied to the Pod Profile. A single Pod Profile is used to apply Pod policies for both Pod-1 and Pod-2. This section explains how to apply Pod Policies to Pod-2.

Setup Information

- Pod Policy Group for Pod-2: Pod2-West_PPG
- Pod Selector Name for Pod-2: Pod2-West
- Pod Profile: default
- ID for Pod-2: 2
- Names of Pod Policies to be applied: Pod2-West-NTP_Policy

Deployment Steps

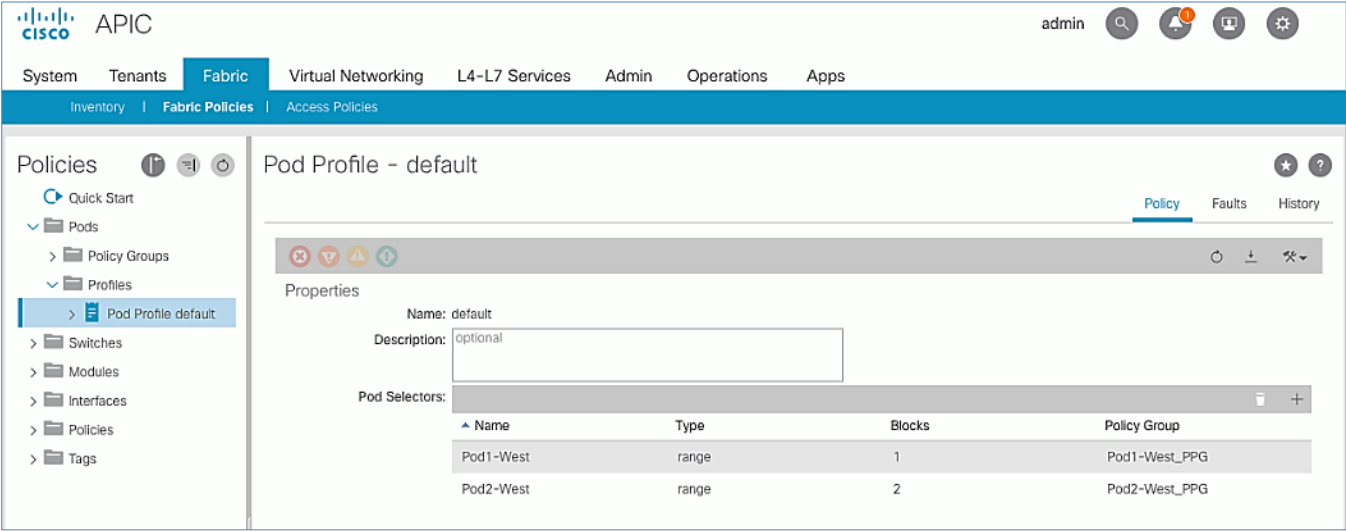
To apply Pod policies on Spine switches in Pod-2, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using **admin** account.
2. From the top menu, select **Fabric > Fabric Policies**.
3. From the left navigation pane, navigate to **Pods > Policy Groups**.
4. Right-click and select **Create Pod Policy Group**, click the **[+]** on the right to **Create Route Reflector Node**.
5. In the **Create Pod Policy Group** pop-up window, for the **Name**, specify a Pod Policy Name (for example, `Pod2-West_PPG`). For the **Date Time Policy**, select the previously created NTP policy for Pod-2 (for example, `Pod2-West-NTP_Policy`). For the different policies, select the `default` policy from the drop-down list, including the BGP Route Reflector Policy that was configured in the previous section.

The screenshot shows the APIC GUI with the 'Create Pod Policy Group' dialog open. The dialog is titled 'Specify the Policy Group properties'. The 'Name' field contains 'Pod2-West_PPG' and the 'Description' field contains 'optional'. The 'Date Time Policy' is set to 'Pod2-West-NTP_Policy', and all other policies (ISIS, COOP Group, BGP Route Reflector, Management Access, SNMP, and MACsec) are set to 'default'. The dialog has 'Cancel' and 'Submit' buttons at the bottom right. The background shows the APIC navigation pane with 'Fabric Policies' selected and a table of existing policies on the right.

Policy Name	Policy Type
Pod2-West-NTP_Policy	NTP Policy
Pod2-West-ISIS_Policy	ISIS Policy
Pod2-West-COOP_Policy	COOP Policy
Pod2-West-BGP_Policy	BGP Policy
Pod2-West-Management_Policy	Management Policy
Pod2-West-SNMP_Policy	SNMP Policy
Pod2-West-MACsec_Policy	MACsec Policy

6. Click **Submit**.
7. From the left navigation pane, navigate to **Pods > Profiles > Pod Profile default**.
8. In the right window pane, in the **Pod Selectors** section, click the **[+]** to add a Pod Selector.
9. In the newly created row, specify a **Name** (for example, `Pod2-West`). For **Type**, select **Range**. For **Blocks**, specify the Pod Id for Pod-2 (for example, `2`). For **Policy Group**, select the previously created Policy Group for Pod2 (for example, `Pod2-West_PPG`).



10. Click **Submit** to apply the Fabric Policies to Pod-2.

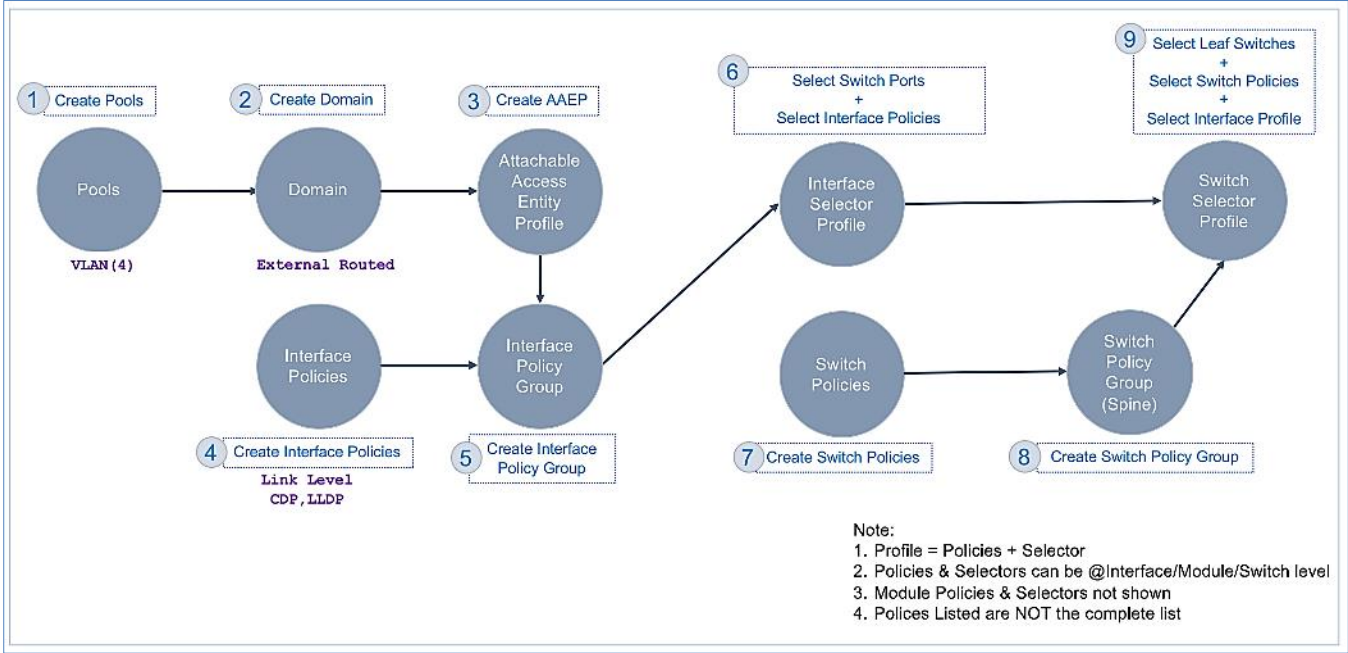
Setup Fabric Access Policies for Spine Switches in Pod-2

In ACI, access policies define the port configuration. In this section, access policies are configured for all interfaces on the spine switches in Pod-2 that connect to the IPN. The access policies enable connectivity between the Spine switches and IPN in Pod-2. The access policies are grouped and applied to specific interfaces and switches using interface and switch profiles respectively.

Deployment Overview

The deployment workflow for configuring Spines to connect to IPN is similar to configuring ACI Leaf switches for connectivity to access layer devices such as Cisco UCS and HyperFlex . The configuration in both cases is done through Fabric Access Policies. The workflow for creating Fabric Access Policies for connecting Spines to IPN devices in Pod-2 is shown in [Figure 13](#).

Figure 13 Fabric Access Policies – For Spine Switch Connectivity to IPN in Pod-2



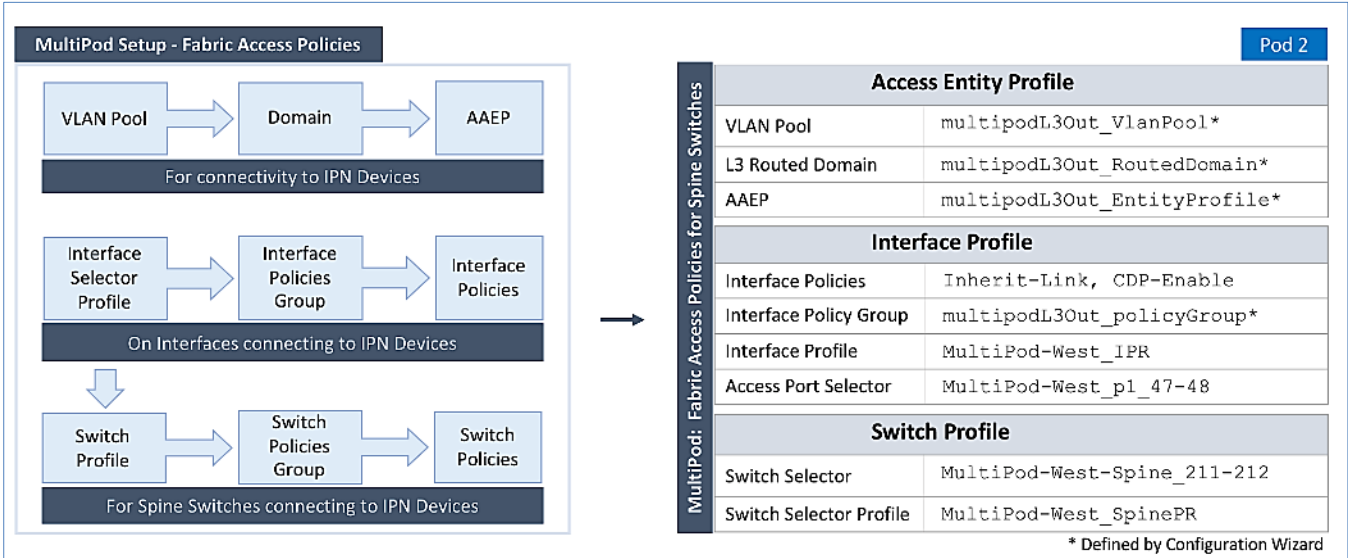
Setup Information

The information for configuring fabric access policies to connect Spine switches to IPN in Pod-2 is provided below.



VLAN Pool, L3 Routed Domain, AAEP and Interface Policy Group listed below are configured by the Configuration Wizard during Multi-Pod setup.

Figure 14 Setup Information – Fabric Access Policies on Pod-2 Spine Switches



Deployment Steps

Follow the procedures outlined in this section to configure access policies on Spine switch interfaces to enable connectivity to IPN in Pod-2. Pod-2 leverages the same interface profile as Pod-1 to enable connectivity to IPN devices in Pod-2. This is

possible because Pod-2 Spine switches connect to the IPN on the same ports and use the same policies as Pod-1 switches in this design, see [Fabric Access Policies configuration in Pod-1](#) for more information.

Update Switch Profile for Spine connectivity to IPN

In this design, the same switch profile is used to configure all Spine switches that connect to the IPN. This is possible because the policies, ports and all other parameters are the same for all Spine switches except that they are all different Spine switches. However, the switch selector profile can be used to select the different switches and apply them to the same switch profile.

To update the switch profile used by Pod-1 Spine switches to include Pod-2 switches, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using **admin** account.
2. From the top menu, select **Fabric > Access Policies**.
3. From the left navigation pane, select and expand **Policies > Switches > Spine Switches > Profiles**.
4. Select the previously created profile (for example, `MultiPod-West_SpinePR`).
5. In the right window pane, for **Spine Selectors**, click the **[+]** on the right-side of the window to add Pod-2 Spine switches to apply the interface profile to. Specify a selector **Name** (for example, `MultiPod-West-Spine_211-212` and under the **Blocks** column, select the Spine Switch IDs from the drop-down list (for example, `211, 212`). Click **Update**. Click **Next**.

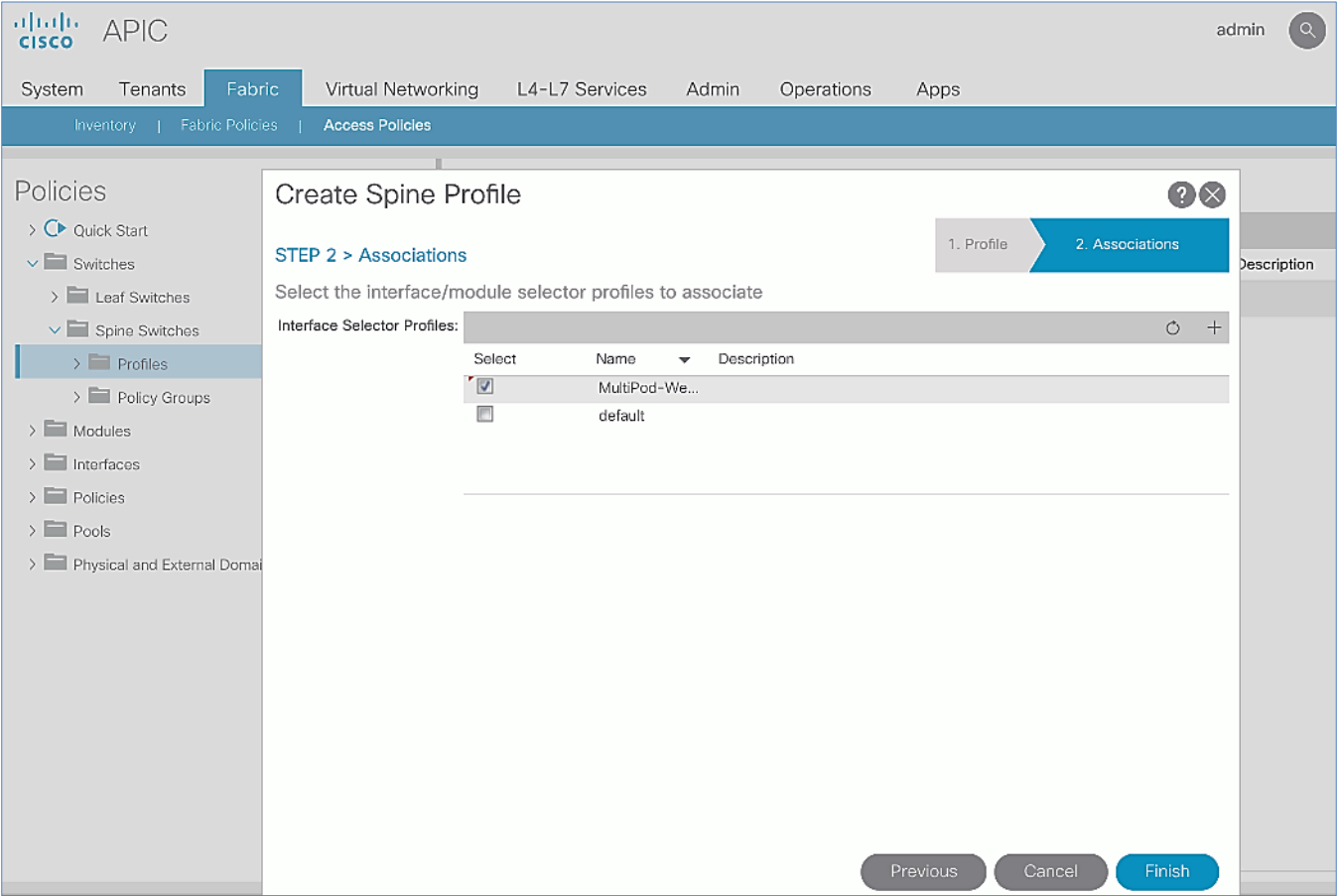
The screenshot shows the Cisco APIC GUI with the 'Create Spine Profile' dialog box open. The dialog is in 'STEP 1 > Profile' and shows the following fields and table:

- Name:** MultiPod-West_SpinePR
- Description:** optional
- Spine Selectors:**

Name	Blocks	Policy Group
MultiPod-West-Spine_211-212	211-212	select an option

Buttons for 'Update', 'Cancel', 'Previous', and 'Next' are visible at the bottom of the dialog.

6. In the **Step 2 > Associations** window, for the **Interface Selector Profile**, select the previously created Interface Profile.



- 7. Click **Finish** to complete.
- 8. Review the switch profile to confirm that Spines in both Pods are selected in the profile.

The screenshot displays the Cisco APIC interface. The top navigation bar includes 'System', 'Tenants', 'Fabric' (selected), 'Virtual Networking', 'L4-L7 Services', 'Admin', 'Operations', and 'Apps'. Below this, a sub-navigation bar shows 'Inventory', 'Fabric Policies', and 'Access Policies'. The left sidebar, titled 'Policies', contains a tree view with categories like 'Quick Start', 'Switches', 'Spine Switches', 'Profiles', 'Policy Groups', 'Modules', 'Interfaces', 'Policies', 'Pools', and 'Physical and External Domains'. The 'MultiPod-West_SpinePR' profile is selected under 'Profiles'. The main content area is titled 'Spine Profile - MultiPod-West_SpinePR' and has tabs for 'Policy', 'Faults', and 'History'. The 'Policy' tab is active, showing a 'Properties' section with 'Name: MultiPod-West_SpinePR' and 'Description: optional'. Below this is a 'Spine Selectors' table:

Name	Blocks	Policy Group
MultiPod-West-Spine_211-212	211-212	
MultiPod-West-Spine_111-112	111-112	

At the bottom of the main panel are three buttons: 'Show Usage', 'Reset', and 'Submit'.

Deploy APICs in Pod-2

This section explains the procedures for deploying an APIC (Pod-2) to the existing APIC (Pod-1) cluster. The new APIC is connected to Pod-2 Leaf switches deployed in the previous section.



For disaster avoidance, at least one APIC should be deployed in Pod-2.

Prerequisites

The following are the prerequisites to deploy APICs in Pod-2:

- All Spine and Leaf switches in Pod-2 should be part of the ACI Fabric and in Active state. APIC should be redundantly connected to an Active Leaf switch pair.
- Pod-2 APIC should run a compatible server firmware version – see APIC release notes for the recommended server firmware. The server firmware version can be seen from the CIMC GUI. See the [Interoperability Matrixes](#) section for the versions used in this CVD.
- APIC in Pod-2 should run the same version of software as other APICs in the cluster APIC cluster. APIC can be upgraded after joining the cluster, but to join the cluster, the software must still be a compatible version.

Deployment Overview

The high-level steps for deploying Pod-2 switches to the ACI Fabric are summarized below:

- Verify that the Pod-2 Spine and Leaf switches are part of the ACI Fabric.
- Complete the initial setup of Pod-2 APIC.
- Verify that the new Pod-2 APIC is part of the APIC cluster
- Add Pod-2 APIC as a destination for DHCP relay on Pod-1 IPN devices.

Verify Pod-2 Switches are Part of the ACI Fabric

Table 26 Pod-2 Switches ACI Fabric Information

Pod-2 Switches	Node ID	Name	Role
	201	BB06-9372PX-WEST-1	Leaf
	202	BB06-9372PX-WEST-2	Leaf
	211	BB06-9364C-WEST-1	Spine
	212	BB06-9364C-WEST-2	Spine

To confirm that the Pod-2 Spine and Leaf switches are part of the ACI Fabric, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using **admin** account.
2. From the top menu, select **Fabric > Inventory**.
3. From the left navigation pane, navigate to **Fabric Membership**.
4. In the right navigation pane, go to the **Registered Nodes** tab.

The screenshot shows the Cisco APIC GUI with the 'Fabric Membership' page selected. The 'Registered Nodes' tab is active, showing a summary of switch counts and a table of registered nodes.

Summary:

- Leaves:** 4 (0 Decommissioned, 0 Maintenance, 4 Active, 0 Inactive)
- Virtual Leaves:** 0 (0 Decommissioned, 0 Maintenance, 0 Active, 0 Inactive)
- Spines:** 4 (0 Decommissioned, 0 Maintenance, 4 Active, 0 Inactive)
- Virtual Spines:** 0 (0 Decommissioned, 0 Maintenance, 0 Active, 0 Inactive)

Registered Nodes Table:

Serial Number	Model	Pod ID	Node ID	Name	Role	IP	Status
SAL1940QAAX	N9K-C9372PX	1	101	AA11-9372PX-WEST-1	leaf	10.13.64.64/32	Active
SAL1940QAEQ	N9K-C9372PX	1	102	AA11-9372PX-WEST-2	leaf	10.13.184.66/32	Active
FDO22240VHM	N9K-C9364C	1	111	AA11-9364C-WEST-1	spine	10.13.184.64/32	Active
FDO22240VJB	N9K-C9364C	1	112	AA11-9364C-WEST-2	spine	10.13.184.65/32	Active
SAL1913CJXR	N9K-C9372PX	2	201	BB06-9372PX-WEST-1	leaf	10.14.24.66/32	Active
SAL1914CN42	N9K-C9372PX	2	202	BB06-9372PX-WEST-2	leaf	10.14.32.64/32	Active
FDO221914JV	N9K-C9364C	2	211	BB06-9364C-WEST-1	spine	10.14.24.64/32	Active
FDO22182Q9G	N9K-C9364C	2	212	BB06-9364C-WEST-2	spine	10.14.24.65/32	Active

5. Confirm that the status is **Active** for all Leaf and Spine switches in Pod-2. For this CVD, the new APIC will be dual-homed to both Leaf switches in Pod-2.

Initial Setup of Pod-2 APIC

Follow the procedures outlined in this section to do an initial setup and configuration of the third APIC in the APIC cluster that will manage the ACI fabric. In this design, two APICs are deployed in Pod-1 and a third APIC in Pod-2.

Prerequisites

KVM Console access is necessary to do an initial setup and configuration of a new APIC. KVM access is available through CIMC Management and therefore access to CIMC Management on the APIC server is required.

Setup Information

The initial setup of APIC in Pod-2 requires the information provided in this section.

- CIMC Management IP Addresses
- CIMC login credentials for the APIC being setup



TEP Address Pool is the APIC TEP pool and should be the same for all APICs in a cluster regardless of their location.



BD Multicast Address (GIPO) is configured only once, during the initial setup of APIC-1. APIC-1 refers to the first controller in the cluster. Remaining controllers and switches sync to the configuration on APIC-1.



APIC username and password is configured only once, during the initial setup of APIC-1 or the first controller in the cluster. Remaining controllers and switches sync to the configuration on APIC-1.

Table 27 Setup Parameters for Pod-2 APIC

APIC	Parameters	Notes	Default Values
Fabric Name	ACI Fabric West		ACI Fabric1
Fabric ID	2	Range: (1-128)	1
Number of Active Controllers	3	Range: (1-9) Minimum # of controllers recommended: 3	3
POD ID	2	Range: (1-254)	1
Standby Controller ?	NO		NO
APIC-X ?	NO		NO
Controller ID	3	Range: (1-3) APIC with ID=1 is the 1st controller in the cluster	1
Controller Name	BB06-APIC-M2-WEST-1		apic1
TEP Address Pool	10.13.0.0/16	APIC TEP Pool is different from the TEP Pool used by switches; Same pool is used by all APICs in a fabric, including APICs in Pod-2	10.0.0.0/16
Infrastructure VLAN ID	4093	Range: (1-4094)	4093
BD Multicast Address (GIPO)	226.0.0.0/15	GIPO is configured during first APIC setup in Pod-1; Remaining controllers will use this	225.0.0.0/15
OOB Management IP	172.26.164.121/24		-
OOB Management Gateway	172.26.164.254		-
OOB Management Speed/Duplex	auto		-
Admin User Password	*****	Password is configured during first APIC setup in Pod-1; Remaining controllers and switches will sync to this	-

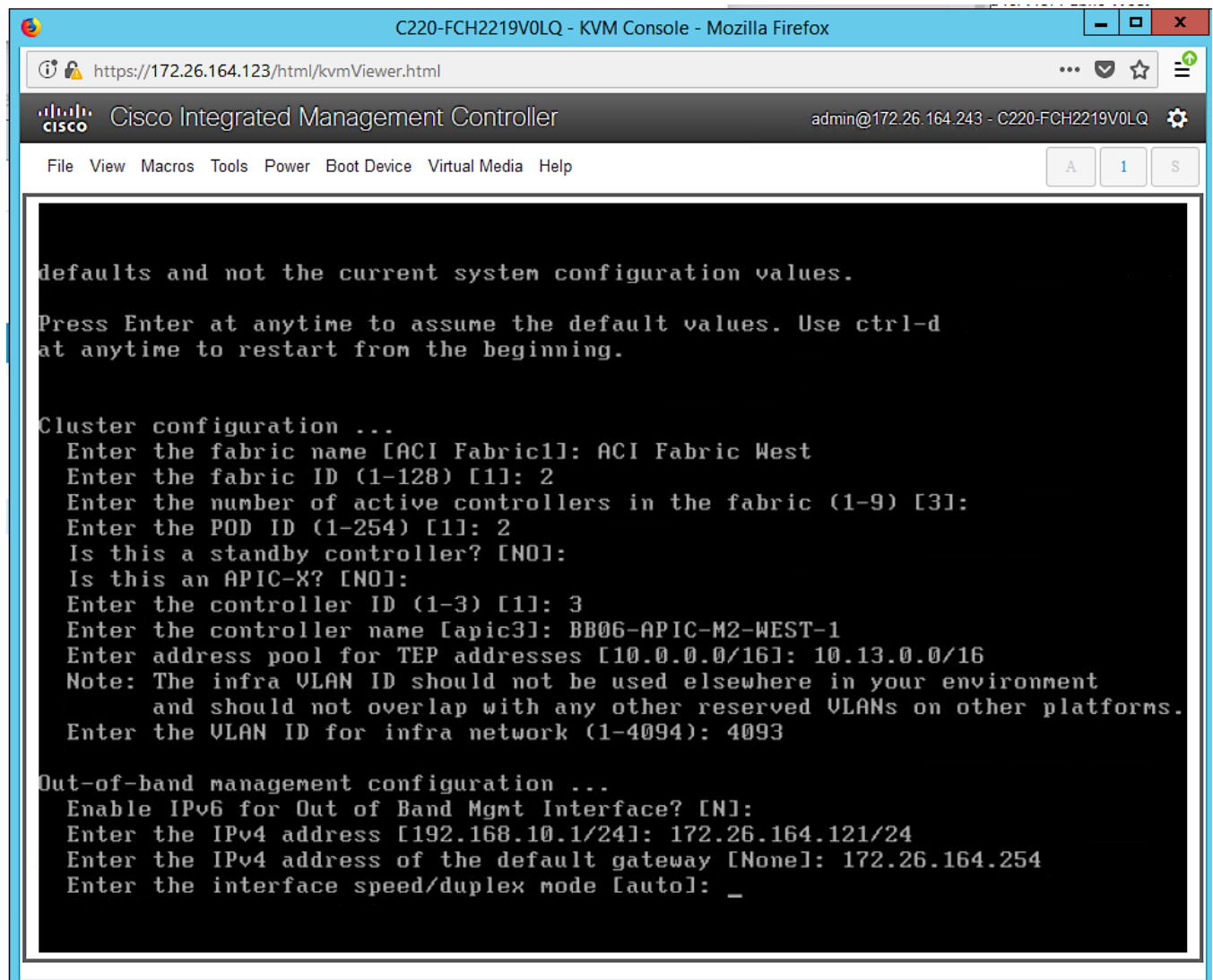
Deployment Steps

To setup a new APIC in Pod-2, follow these steps:

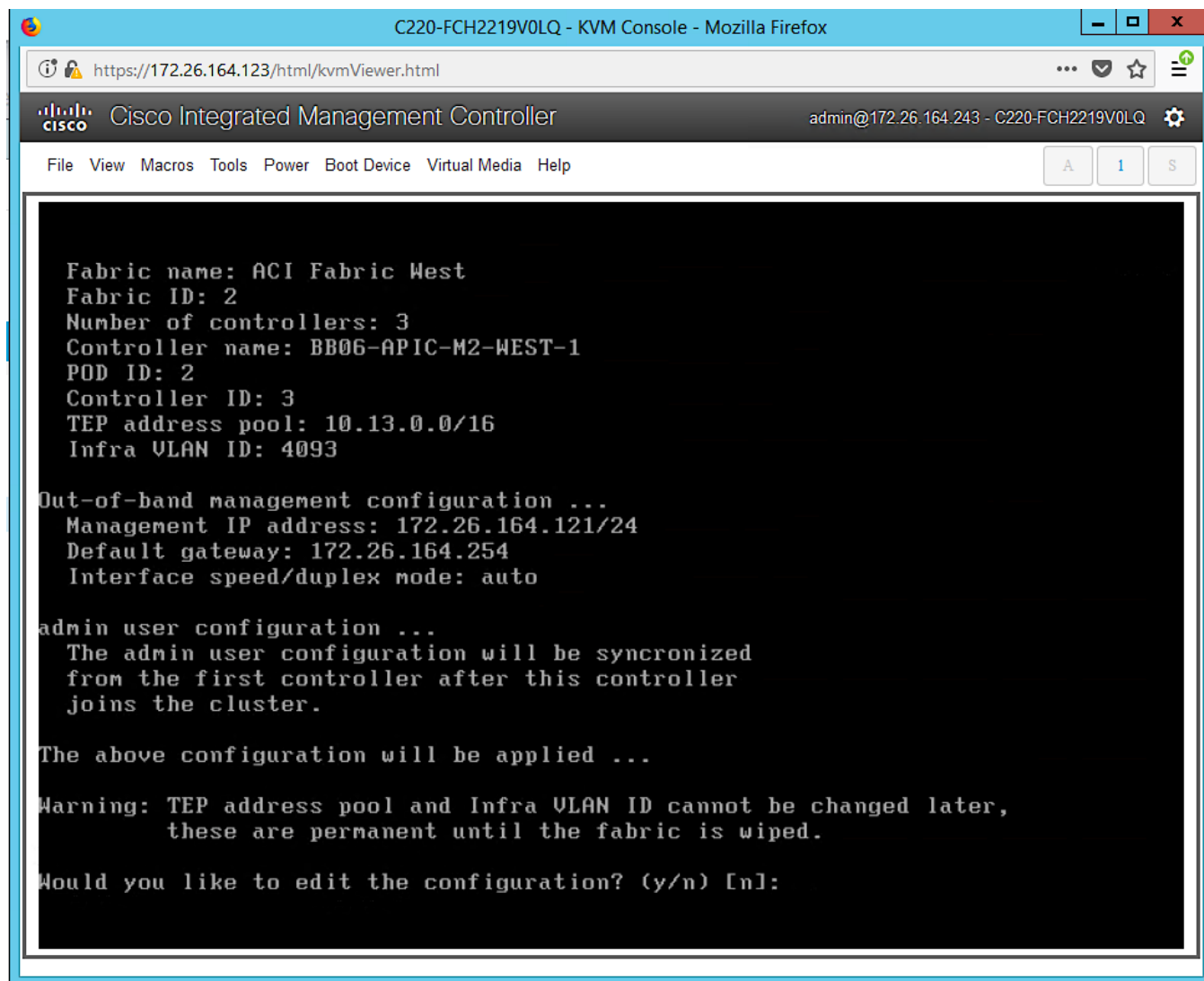
1. Use a browser to navigate to the CIMC IP address of the new APIC. Log in using **admin** account.
2. From the top menu, click **Launch KVM**. Select **HTML based KVM** from the drop-down list.
3. When the KVM Application launches, the initial APIC setup screen should be visible. Press any key to start the **Setup Utility**. Use the Setup information provided above to step through the initial APIC configuration as shown below.



If the APIC was previously configured, reset to factory defaults and wipe it clean before proceeding.



4. Press **Enter** to accept [auto] as the default for the last question.
5. Review the configured information.



6. Click **y** if necessary to go back and make changes, otherwise press **Enter** to accept the configuration.

Verify Pod-2 APIC is Part of the APIC Cluster

To confirm that the Pod-2 APIC was successfully added to the APIC cluster, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using **admin** account.
2. From the top menu, select **System > Controllers**.
3. From the left navigation pane, navigate to **Controllers**.
4. From the left navigation pane, select and expand one of the Pod-1 APICs. Navigate to **Cluster as Seen by Node**.

The screenshot shows the Cisco APIC GUI. The top navigation bar includes tabs for System, Tenants, Fabric, Virtual Networking, L4-L7 Services, Admin, Operations, and Apps. Below this is a secondary navigation bar with links like QuickStart, Dashboard, Controllers, System Settings, Smart Licensing, Faults, Config Zones, Events, Audit Log, and Active Sessions. The left sidebar shows a tree view under 'Controllers' with 'AA11-APIC-M2-WEST-1 (Node-1)' selected, and 'Cluster as Seen by Node' highlighted. The main content area is titled 'Cluster as Seen by Node' and shows the 'APIC Cluster' tab. It displays properties for the ACI Fabric West cluster, including Fabric Name, Target Size, Current Size, Time Difference, and a table of Active Controllers.

ID	Name	IP	Admin State	Operational State	Health State	Failover Status	Serial Number	SSL Certificate
1	AA11-APIC-M2-WEST-1	10.13.0.1	In Service	Available	Fully Fit	idle	FCH2219V...	yes
2	AA11-APIC-M2-WEST-2	10.13.0.2	In Service	Available	Fully Fit	idle	FCH2219V...	yes
3	BB06-APIC-M2-WEST-1	10.13.0.3	In Service	Available	Fully Fit	idle	FCH2219V...	yes

- Verify that the newly deployed Pod-2 APIC is **In Service**, **Available** and **Fully Fit** as shown above.
- Note the TEP IP Address of the newly deployed APIC (for example, 10.13.0.3). This address will be used to configure DHCP Relay on Pod-1 IPN routers to point to the new APIC. For Pod-1 APICs, DHCP relay was configured as a part of the initial IPN configuration.

Add Pod-2 APIC as DHCP Relay Destination

In this section, DHCP Relay is configured on Pod-1 IPN routers to point to the newly deployed APIC in Pod-2. DHCP Relay statements should be configured on the Spine-facing interfaces of Pod-1 IPN routers.

Setup Information

- Pod-2 APIC TEP IP Address: 10.13.0.3

Use the above information to configure DHCP relay on Pod-1 IPN routers to point to the newly deployed APIC in Pod-2.

Configure DHCP Relay for Pod-2 APIC on IPN Devices in Pod-1

POD-1: IPN Router#1

POD-1: IPN Router#2

POD-1: IPN Router#1	POD-1: IPN Router#2
<pre> switchaname AA11-93180YC-EX-WEST-IPN-1 ... interface Ethernet1/49 description To POD-1:AA11-9364C-1:E1/47 no switchport mtu 9216 no shutdown interface Ethernet1/49.4 mtu 9216 encapsulation dot1q 4 vrf member MultiPod-Fabric-West ip address 10.113.11.2/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.0 ip pim sparse-mode ip dhcp relay address 10.13.0.3 no shutdown interface Ethernet1/50 description To POD-1:AA11-9364C-2:E1/47 no switchport mtu 9216 no shutdown interface Ethernet1/50.4 mtu 9216 encapsulation dot1q 4 vrf member MultiPod-Fabric-West ip address 10.113.12.2/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.0 ip pim sparse-mode ip dhcp relay address 10.13.0.3 no shutdown </pre>	<pre> switchaname AA11-93180YC-EX-WEST-IPN-2 ... interface Ethernet1/49 description To POD-1:AA11-9364C-WEST-1:E1/48 no switchport mtu 9216 no shutdown interface Ethernet1/49.4 mtu 9216 encapsulation dot1q 4 vrf member MultiPod-Fabric-West ip address 10.113.11.6/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.0 ip pim sparse-mode ip dhcp relay address 10.13.0.3 no shutdown interface Ethernet1/50 description To POD-1:AA11-9364C-WEST-2:E1/48 no switchport mtu 9216 no shutdown interface Ethernet1/50.4 mtu 9216 encapsulation dot1q 4 vrf member MultiPod-Fabric-West ip address 10.113.12.6/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.0 ip pim sparse-mode ip dhcp relay address 10.13.0.3 no shutdown </pre>

Verify ACI Multi-Pod Fabric Setup

This section provides a few GUI and CLI commands that can be used to verify that the protocols are working correctly before proceeding to the next stage of the deployment.

Verify OSPF Status on Spine Switches

OSPF is running between Spine switches and IPN devices in each Pod. To verify that OSPF is setup and working correctly between Pods, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using **admin** account.
2. From the top menu, select **Fabric > Inventory**.
3. From the left navigation pane, select and expand **Inventory > Pod 1 > (Name_of_Spine_switch_in_Pod_1) > Protocols > OSPF > OSPF for VRF-overlay-1**.

The screenshot displays the Cisco APIC GUI. The top navigation bar includes 'System', 'Tenants', 'Fabric' (selected), 'Virtual Networking', 'L4-L7 Services', 'Admin', 'Operations', and 'Apps'. Below this, the 'Inventory' section is active, showing a tree view of the fabric components. The 'OSPF for VRF-overlay-1' configuration is selected. The main pane shows the 'General' tab for 'OSPF - overlay-1'. A health indicator at the top left shows a score of 100. The 'PROPERTIES' section lists configuration details like Name, Route ID, Distance, Max ECMP, and Bandwidth Reference. The 'STATS' section provides counts for various OSPF metrics. The 'Neighbors' table lists two neighbors with their IP addresses, states (Full), peer IP addresses, and interfaces. At the bottom, the 'Inter Protocol Route Leak Into OSPF' section is partially visible.

4. In the right window pane, under the **General** tab, the top left icon indicates the **Health** for OSPF in VRF `overlay-1`. Confirm that the OSPF **health** is at **100** indicating there are no faults or errors for OSPF. Navigate to the **Neighbors** section and confirm for each IPN neighbor in the same Pod, neighbor state is **Up** and the OSPF **State** is **Full**.
5. Repeat steps 1-4 to verify OSPF on other Spine switches in the Pod that connect to the IPN.
6. You can also verify that OSPF is setup correctly by executing the following commands from CLI. SSH into the Spine switches and log in using the **admin** account.
 - `show ip ospf neighbors vrf overlay-1`
 - `show ip ospf route vrf overlay-1`
 - `show ip route vrf overlay-1`

Verify MP-BGP EVPN Status on Spine Switches

MP-BGP sessions run between Spine switches in each Pod that connect to the IPN. To verify that MP-BGP EVPN is setup and working correctly between Pods, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using **admin** account.
2. From the top menu, select **Fabric > Inventory**.
3. From the left navigation pane, select and expand **Inventory > Pod 1 > (Name_of_Spine_switch_in_Pod_1) > Protocols > BGP > BGP for VRF-overlay-1 > Neighbors**.
4. In the right window pane, select and expand the router ID (for example, `14.14.14.11`) for the peer Spines in Pod-2.

The screenshot shows the Cisco APIC GUI with the following navigation path: **Fabric** > **Inventory** > **Pod 1** > **Protocols** > **BGP** > **BGP for VRF-overlay-1** > **Neighbors**. The **Neighbors** table displays the following data:

Name	State	Neighbor Address Family	Neighbor Address Family Capability	Accepted Paths	Up Since
14.14.14.11	established				2018-11-26T01:4...
Vpnv4 unicast address family		vpn4-ucast	first-eor-rcvd	12	
Vpnv6 unicast address family		vpn6-ucast	first-eor-rcvd	0	
L2Vpn EVpn address family		l2vpn-evpn	first-eor-rcvd	59	
14.14.14.12	established				2018-11-24T07:1...
Vpnv4 unicast address family		vpn4-ucast	first-eor-rcvd	12	
Vpnv6 unicast address family		vpn6-ucast	first-eor-rcvd	0	
L2Vpn EVpn address family		l2vpn-evpn	first-eor-rcvd	59	
10.13.64.64	established				2018-11-24T02:3...
10.13.184.66	established				2018-11-24T02:3...
10.13.64.65	established				2018-12-20T05:2...

- Verify that the State is **Established** and for **L2Vpn EVpn address family**, paths are being learned. Also confirm that the BGP health is at **100** indicating there are no faults or errors for BGP in VRF **overlay-1** by navigating back to **BGP for VRF-overlay-1** in the left navigation pane.
- Repeat steps 1-5 to verify BGP on other Spine switches in the Pod that connect to the IPN.
- You can also verify that MP-BGP EVPN is setup correctly by executing the following commands from CLI. SSH into the Spine switches and log in using the **admin** account.
 - `show bgp l2vpn evpn summary vrf overlay-1`

Verify COOP Status on Spine Switches

Council of Oracles Protocol (COOP) database maintained on Spines in each Pod, is a database of all endpoints learned. This includes endpoints learned from within the Pod as well as the addresses learned through the tunnel between spine switches in different pods. The ETEP used by MP-BGP EVPN will be used by COOP to identify a remote pod's set of anycast addresses.

To verify that COOP database is learning addresses from the remote Pod, follow these steps:

- Use a browser to navigate to the APIC GUI. Log in using **admin** account.
- From the top menu, select **Fabric > Inventory**.
- From the left navigation pane, select and expand **Inventory > Pod 1 > (Name_of_Spine_switch_in_Pod_1) > Protocols > COOP > COOP for VRF-overlay-1**.
- In the right window pane, under the **General** tab, the top left icon indicates the **Health** for COOP in VRF **overlay-1**. Confirm that the COOP health is at **100** indicating there are no faults or errors.
- From the left navigation pane, select and expand **Inventory > Pod 1 > (Name_of_Spine_switch_in_Pod_1) > Protocols > COOP > COOP for VRF-overlay-1 > Endpoint Database**.
- In the right window pane, verify that endpoints from Pod-2 are being learned (for example, `10.1.167.168`).

The screenshot shows the Cisco APIC web interface. The top navigation bar includes 'System', 'Tenants', 'Fabric' (selected), 'Virtual Networking', 'L4-L7 Services', 'Admin', 'Operations', and 'Apps'. Below this, there are tabs for 'Inventory', 'Fabric Policies', and 'Access Policies'. The left sidebar shows the 'Inventory' tree with 'AA11-9364C-WEST-1 (Node-111)' expanded, showing 'Chassis', 'Interfaces', 'Protocols', 'BGP', 'COOP', and 'COOP for VRF-overlay-1'. Under 'COOP for VRF-overlay-1', 'Endpoint Database' is selected. The main panel displays the 'Endpoint Database' table with the following data:

Vrf VnId	Mac	EndPoint IPv4	EndPoint IPv6
3047424	00:0C:29:F4:49:8A	10.1.167.166	
3047424	00:0C:29:02:BA:24	10.1.167.168	
3047424	00:0C:29:B5:20:83		
3047424	00:0C:29:B5:20:79	10.1.167.110, 10.1.167.161	
3047424	00:0C:29:09:07:8F		
3047424	00:0C:29:09:07:85	10.1.167.164	
3047424	00:50:56:A0:79:9F	10.1.167.21	
3047424	00:50:56:A0:BC:1D	10.1.167.22	
16277109	00:50:56:A0:47:01	10.1.144.65	

The bottom of the table shows pagination: Page 3 of 5, Objects Per Page: 15, and Displaying Objects 31 - 45 Of 72.

7. Double-click one endpoint to get additional details. Note that the **Publisher ID** is the ETEP address (for example, 10.114.114.1) of a Spine in Pod-2.
8. Repeat steps 1-7 to verify COOP on other Spine switches in the Pod that connect to the IPN.
9. You can also verify that COOP is functioning correctly by executing the following commands from CLI. SSH into the Spine switches and log in using the **admin** account.
 - `show coop internal info ip-db`

Solution Deployment – ACI Fabric (To Outside Networks from Pod-2)

Complete the steps outlined in this section to deploy shared Layer 3 outside (**Shared L3Out**) connectivity to networks outside the ACI fabric from Pod-2.

Deployment Overview

The Shared L3Out connection is established in the system-defined **common** Tenant as a common resource that can be shared by multiple tenants in the ACI fabric. In this design, the Shared L3Out design in Pod-2 is same as that of Pod-1. For additional details, see the [Shared L3Out deployment](#) section for Pod-1. Some specifics of the Pod-2 deployment are summarized below:

- Pair of Border Leaf switches in Pod-2 connect to a pair of Nexus 7000 routers outside the ACI fabric using 4 x 10GbE links. Nexus 7000 routers serve as a gateway to the networks outside the fabric.
- Routing protocol use to exchange routes between the ACI fabric and networks outside ACI is OSPF
- VLAN tagging is used for connectivity across the 4 links – a total of 4 VLANs for the 4 x 10GbE links. VLANs are configured on separate sub-interfaces.
- Fabric Access Policies are configured on ACI Leaf switches to connect to the **External Routed** domain using VLAN pool (vlans: 315–318).
- Pod-2 uses the same Tenant (**common**), VRF (`common-SharedL3Out_VRF`) and Bridge Domain (`common-SharedL3Out_BD`) as Pod-1 for Shared L3Out.
- The shared L3Out created in **common** Tenant “provides” an external connectivity contract that can be “consumed” from any tenant.
- The Nexus 7000s connected to Pod-2 are configured to originate and send a default route via OSPF to the border leaf switches in Pod-2.
- ACI leaf switches in Pod-2 advertise tenant subnets back to Nexus 7000 switches.
- In ACI 4.0, ACI leaf switches can also advertise host-routes if it is enabled.

Create VLAN Pool for External Routed Domain

In this section, a VLAN pool is created to enable connectivity to the external networks, outside the ACI fabric. The VLANs in the pool are for the four links that connect ACI Border Leaf switches to the Nexus Gateway routers in the non-ACI portion of the customer’s network.

Table 28 VLAN Pool for Shared L3Out in Pod-2

To External Networks Outside ACI – Pod-2	VLAN Pool Name	Leaf Node ID	VLAN ID	Connects To
	SharedL3Out- West-Pod2_VLANS	201	315	1 st L3 Gateway Outside ACI
			316	2 nd L3 Gateway Outside ACI
		202	317	1 st L3 Gateway Outside ACI
			318	2 nd L3 Gateway Outside ACI

To configure a VLAN pool to connect to external gateway routers outside the ACI fabric, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Access Policies**.
3. From the left navigation pane, expand and select **Pools > VLAN**.
4. Right-click and select **Create VLAN Pool**.
5. In the **Create VLAN Pool** pop-up window, specify a **Name** (for example, *SharedL3Out-West-Pod2_VLANS*) and for **Allocation Mode**, select **Static Allocation**.
6. For **Encap Blocks**, use the **[+]** button on the right to add VLANs to the VLAN Pool. In the **Create Ranges** pop-up window, configure the VLANs that need to be configured from the Border Leaf switches to the external gateways outside the ACI fabric. Leave the remaining parameters as is.

The screenshot displays the Cisco APIC GUI. The top navigation bar includes 'System', 'Tenants', 'Fabric', 'Virtual Networking', 'L4-L7 Services', 'Admin', 'Operations', and 'Apps'. The 'Fabric' tab is selected, and the 'Access Policies' sub-tab is active. The left sidebar shows a tree view with 'Policies' expanded, and 'VLAN' selected under 'Pools'. The main area shows a table of VLAN pools:

Name	Allocation Mode	Encap Blocks	Description
multipodL3Out_VlanPool	Dynamic Alloc...	[4]	
SharedL3Out-West-Pod1_...	Static Allocation	[311-314]	

Two pop-up windows are overlaid on the main interface:

Create VLAN Pool
Specify the Pool identity

- Name: SharedL3Out-West-Pod2_VLANS
- Description: optional
- Allocation Mode: **Static Allocation** (selected)
- Encap Blocks: (empty list with a '+' button to add)

Create Ranges
Specify the Encap Block Range

- Type: VLAN
- Range: VLAN [315] - VLAN [318] (Integer Value)
- Allocation Mode: **Inherit allocMode from parent** (selected)
- Role: **External or On the wire encapsulations** (selected)

Buttons for 'Cancel' and 'OK' are visible at the bottom right of the 'Create Ranges' window.

7. Click **OK**. Use the same VLAN ranges on the external gateway routers to connect to the ACI Fabric.
8. Click **Submit** to complete.

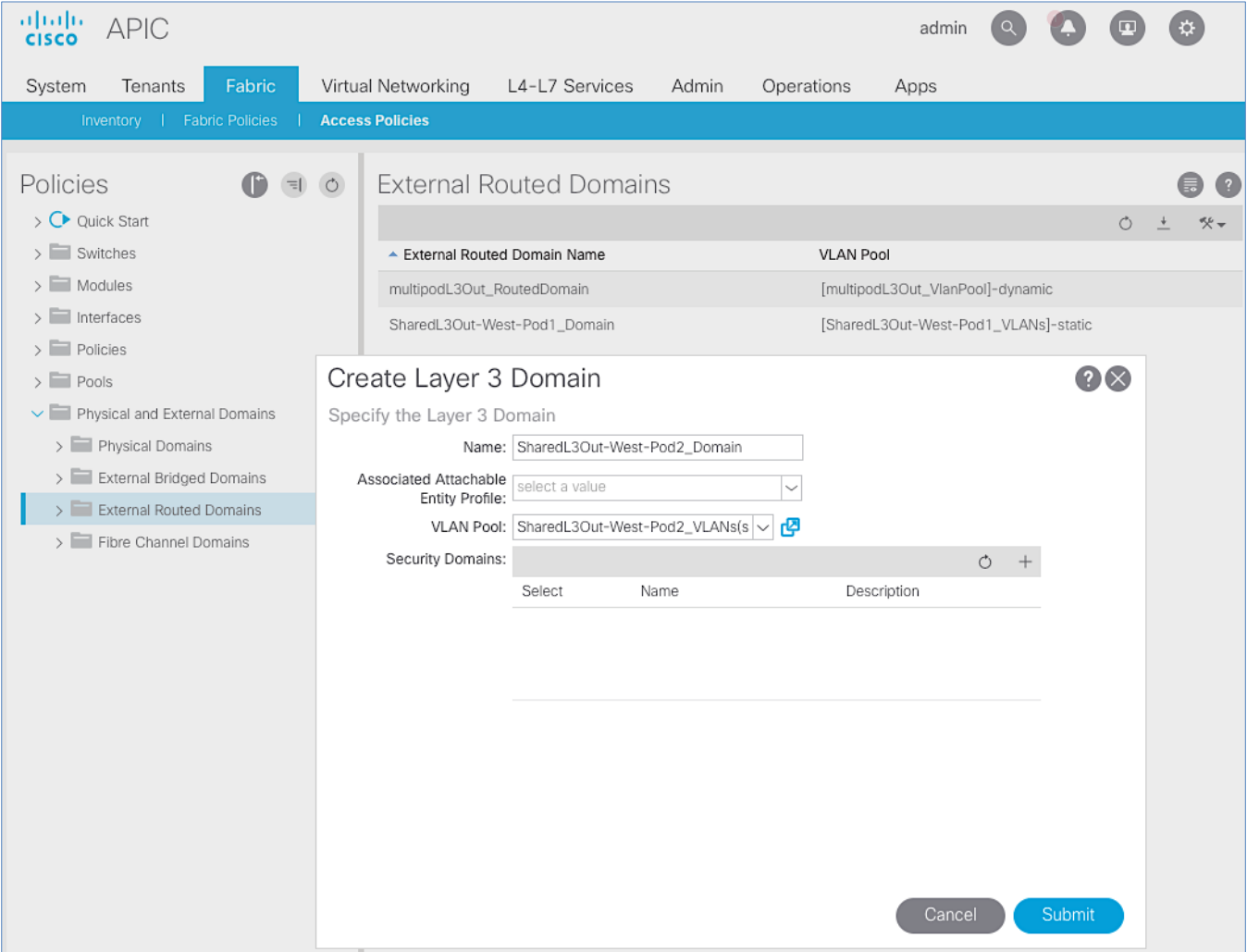
Configure Domain Type for External Routed Domain

Table 29 Domain Type for Shared L3Out in Pod-2

To External Networks Outside ACI – Pod-2	Domain Name	Domain Type	VLAN Pool Name	Connects To
	SharedL3Out- West-Pod2_Domain	External Routed Domain	SharedL3Out- West-Pod2_VLANS	L3 Gateway Routers Outside ACI

To specify the domain type to connect to external gateway routers outside the ACI fabric, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Access Policies**.
3. From the left navigation pane, expand and select Physical and External Domains > External Routed Domains.
4. Right-click External Routed Domains and select **Create Layer 3 Domain**.
5. In the **Create Layer 3 Domain** pop-up window, specify a **Name** for the domain. For the **VLAN Pool**, select the previously created VLAN Pool from the drop-down list.



6. Click **Submit** to complete.

Create Attachable Access Entity Profile for External Routed Domain

Table 30 Attachable Access Entity Profile (AAEP) for Shared L3Out in Pod-2

To External Networks Outside ACI – Pod-2	AAEP Name	Domain Name	VLAN Pool Name	Connects To
	SharedL3Out-West-Pod2_AAEP	SharedL3Out-West-Pod2_Domain	SharedL3Out-West-Pod2_VLANS	L3 Gateway Routers Outside ACI

To create an Attachable Access Entity Profile (AAEP) to connect to external gateway routers outside the ACI fabric, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Access Policies**.
3. From the left navigation pane, expand and select **Policies > Global > Attachable Access Entity Profiles**.

4. Right-click and select **Create Attachable Access Entity Profile**.
5. In the **Create Attachable Access Entity Profile** pop-up window, specify a **Name** (for example, `SharedL3Out-West-Pod2_AAEP`).
6. For the **Domains**, click the **[+]** on the right-side of the window and select the previously created domain from the drop-down list below **Domain Profile**.
7. Click **Update**.
8. You should now see the selected domain and the associated VLAN Pool as shown below.

The screenshot displays the Cisco APIC interface with the **Create Attachable Access Entity Profile** dialog box open. The dialog is in **STEP 1 > Profile** mode. The **Name** field is set to `SharedL3Out-West-Pod2_AAEP` and the **Description** is `optional`. The **Enable Infrastructure VLAN** checkbox is unchecked. Below, the **Domains (VMM, Physical or External) To Be Associated To Interfaces** table shows one entry: `L3 External Domain - SharedL3Out-West-Pod2_Domain` with encapsulation `from:vlan-315 to:vlan-318`. The **EPG DEPLOYMENT** section at the bottom indicates that all selected EPGs will be deployed on all associated interfaces, but no EPGs are currently listed. The **Next** button is highlighted in blue.

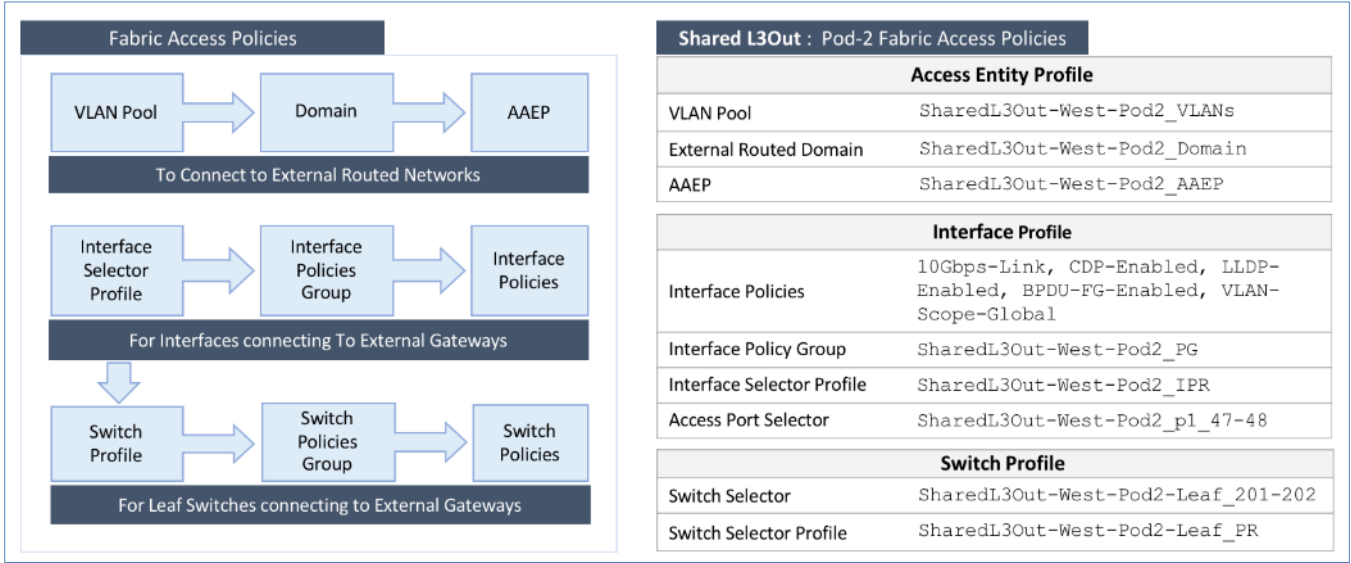
9. Click **Next**. This profile is not associated with any interfaces at this time. They can be associated once the interfaces are configured in an upcoming section.

- 10. Click **Finish** to complete.

Configure Interfaces to External Routed Domain

Border Leaf switches (Node ID: 201, 202) in Pod-2 connect to External Gateways (Nexus 7000 series switches) using 10Gbps links, on ports 1/47 and 1/48.

Figure 15 Fabric Access Policies for Shared L3Out in Pod-2



Create Interface Policy Group for Interfaces to External Routed Domain

To create an interface policy group to connect to external gateways outside the ACI fabric, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Access Policies**.
3. From the left navigation pane, expand and select **Interfaces > Leaf Interfaces > Policy Groups > Leaf Access Port**.
4. Right-click and select **Create Leaf Access Port Policy Group**.
5. In the **Create Leaf Access Port Policy Group** pop-up window, specify a **Name** and select the applicable interface policies from the drop-down list for each field.

The screenshot shows the Cisco APIC interface with the 'Fabric' tab selected. The left sidebar shows a tree view of 'Policies' with 'Leaf Access Port' selected. The main area displays a table of 'Policy Groups - Leaf Access Port'. A modal dialog titled 'Create Leaf Access Port Policy Group' is open, prompting the user to 'Specify the Policy Group identity'. The dialog contains the following fields and options:

- Name:** SharedL3Out-West-Pod2_PG
- Description:** optional
- Link Level Policy:** 10Gbps-Link
- CDP Policy:** CDP-Enabled
- MCP Policy:** select a value
- CoPP Policy:** select a value
- LLDP Policy:** LLDP-Enabled
- STP Interface Policy:** BPDU-FG-Enabled
- Storm Control Interface Policy:** select a value
- L2 Interface Policy:** VLAN-Scope-Global
- Port Security Policy:** select a value
- Egress Data Plane Policing Policy:** select a value
- Ingress Data Plane Policing Policy:** select a value
- Monitoring Policy:** select a value
- Priority Flow Control Policy:** select a value
- Fibre Channel Interface Policy:** select a value
- PoE Interface Policy:** select a value
- Slow Drain Policy:** select a value
- MACsec Policy:** select a value
- 802.1x Port Authentication Policy:** select a value
- DWDM Policy:** select a value

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

- For the **Attached Entity Profile**, select the previously created AAEP to external routed domain.

The screenshot shows the Cisco APIC interface with the 'Fabric' tab selected. The left sidebar shows the 'Policies' section expanded to 'Leaf Access Port'. The main area displays a table titled 'Policy Groups - Leaf Access Port' with the following data:

Name	Link Level Policy	CDP Policy	LLDP Policy	STP Interface Policy	Monitoring Policy
SharedL3Out-West-Pod1_PG	10Gbps-Link	CDP-Enabled	LLDP-Enabled	BPDU-FG-Ena...	

Overlaid on this is the 'Create Leaf Access Port Policy Group' dialog box. It contains the following fields and sections:

- Specify the Policy Group identity**
 - Ingress Data Plane Policing Policy: select a value
 - Monitoring Policy: select a value
 - Priority Flow Control Policy: select a value
 - Fibre Channel Interface Policy: select a value
 - PoE Interface Policy: select a value
 - Slow Drain Policy: select a value
 - MACsec Policy: select a value
 - 802.1x Port Authentication Policy: select a value
 - DWDM Policy: select a value
- Attached Entity Profile: SharedL3Out-West-Poi
- Connectivity Filters:

Switch IDs	Interfaces
- NetFlow Monitor Policies:

NetFlow IP Filter Type	NetFlow Monitor Policy

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

7. Click **Submit** to complete.
8. You should now see the policy groups for both Pods as shown below.

The screenshot shows the Cisco APIC GUI. The top navigation bar includes 'System', 'Tenants', 'Fabric' (selected), 'Virtual Networking', 'L4-L7 Services', 'Admin', 'Operations', and 'Apps'. Below this, a sub-navigation bar shows 'Inventory', 'Fabric Policies', and 'Access Policies'. The left sidebar, titled 'Policies', has a tree view with 'Quick Start', 'Switches', 'Modules', 'Interfaces' (expanded), 'Spine Interfaces', 'Leaf Interfaces' (expanded), 'Profiles', 'Policy Groups' (expanded), and 'Leaf Access Port' (selected). The main panel is titled 'Policy Groups - Leaf Access Port' and contains a table with the following data:

Name	Link Level Policy	CDP Policy	LLDP Policy	STP Interface Policy	Monitoring Policy
SharedL3Out-West-Pod1_PG	10Gbps-Link	CDP-Enabled	LLDP-Enabled	BPDU-FG-Ena...	
SharedL3Out-West-Pod2_PG	10Gbps-Link	CDP-Enabled	LLDP-Enabled	BPDU-FG-Ena...	

Create Interface Profile for Interfaces to External Routed Domain

To create an interface profile to connect to external gateways outside the ACI fabric, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Access Policies**.
3. From the left navigation menu, expand and select **Interfaces > Leaf Interfaces > Profiles**.
4. Right-click and select Create Leaf Interface Profile.
5. In the **Create Leaf Interface Profile** pop-up window, specify a **Name**. For **Interface Selectors**, click the **[+]** to select access ports to apply interface policies to. In this case, the interfaces are access ports that connect Border Leaf switches to gateways outside ACI.
6. In the **Create Access Port Selector** pop-up window, specify a selector **Name**. For the **Interface IDs**, specify the access ports connecting to the two external gateways. For the **Interface Policy Group**, select the previously created Policy Group from the drop-down list.

The screenshot displays the Cisco APIC web interface. The top navigation bar includes 'System', 'Tenants', 'Fabric' (selected), 'Virtual Networking', 'L4-L7 Services', 'Admin', 'Operations', and 'Apps'. Below this, a sub-navigation bar shows 'Inventory', 'Fabric Policies', and 'Access Policies'. The left sidebar lists 'Policies' with a tree view including 'Quick Start', 'Switches', 'Modules', 'Interfaces' (expanded), 'Spine Interfaces', 'Leaf Interfaces' (expanded), 'Profiles' (selected), 'Policy Groups', 'Overrides', 'Policies', 'Pools', and 'Physical and External Domains'. The main content area is titled 'Leaf Interfaces - Profiles' and contains a table with columns 'Name', 'Interface Selectors', and 'Description'. A single entry is visible: 'SharedL3Out-West-Pod1_IPR' with selector '1/47-48'. Two pop-up windows are overlaid on the interface. The top window, 'Create Leaf Interface Profile', prompts for 'Specify the profile Identity' with fields for 'Name' (SharedL3Out-West-Pod2_IPR), 'Description' (optional), and 'Interface Selectors'. The bottom window, 'Create Access Port Selector', prompts for 'Specify the selector identity' with fields for 'Name' (SharedL3Out-West-Pod2_p1_47), 'Description' (optional), 'Interface IDs' (1/47-48), a 'Connected To Fex' checkbox, and an 'Interface Policy Group' dropdown (SharedL3Out-West-Pod2_PG). Both windows have 'Cancel' and 'OK' buttons at the bottom right.

7. Click **OK** to complete and close the **Create Access Port Selector** pop-up window.
8. Click **Submit** to complete and close the **Create Leaf Interface Profile** pop-up window.
9. You should now see the Interface profiles for both Pods as shown below.

The screenshot shows the Cisco APIC GUI. The top navigation bar includes 'System', 'Tenants', 'Fabric' (selected), 'Virtual Networking', 'L4-L7 Services', 'Admin', 'Operations', and 'Apps'. Below this, a sub-navigation bar shows 'Inventory', 'Fabric Policies', and 'Access Policies'. The left sidebar, titled 'Policies', has a tree view with 'Quick Start', 'Switches', 'Modules', 'Interfaces' (expanded), 'Spine Interfaces', 'Leaf Interfaces' (expanded), 'Profiles' (selected), 'Policy Groups', 'Overrides', 'Policies', 'Pools', and 'Physical and External Domains'. The main content area is titled 'Leaf Interfaces - Profiles' and contains a table with the following data:

Name	Interface Selectors	Description
SharedL3Out-West-Pod1_IPR	1/47-48	
SharedL3Out-West-Pod2_IPR	1/47-48	

Create Leaf Switch Profile to External Routed Domain

To create a leaf switch profile to configure connectivity to external gateway routers outside the ACI fabric, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Access Policies**.
3. From the left navigation menu, expand and select **Switches > Leaf Switches > Profiles**.
4. Right-click and select **Create Leaf Profile**.
5. In the **Create Leaf Profile** pop-up window, specify a profile **Name**. For **Leaf Selectors**, click the **[+]** to select the Leaf switches to apply the policies to. In this case, the Leaf switches are the Border Leaf switches that connect to the gateways outside ACI.
6. Specify a Leaf Selector **Name**. For the **Interface IDs**, specify the access ports connecting to the two external gateways. For **Blocks**, select the Node IDs of the Border Leaf switches from the drop-down list. Click **Update**.

APIC

admin

System Tenants **Fabric** Virtual Networking L4-L7 Services Admin Operations Apps

Inventory Fabric Policies **Access Policies**

Policies

> Quick Start

> Switches

> Leaf Switches

> Profiles

> Policy Groups

> Overrides

> Spine Switches

> Modules

> Interfaces

> Policies

> Pools

> Physical and External Domains

Leaf Switches - Profiles

Leaf Selectors (Switch Policy Group)

SharedL3Out-West-Pod1-Leaf_PR 101-102 SharedL3Out-West-Pod1_IPR

Create Leaf Profile

STEP 1 > Profile

1. Profile 2. Associations

Specify the profile Identity

Name:

SharedL3Out-West-Pod2-Leaf_PR

Description:

optional

Leaf Selectors:

Name

Blocks

Policy Group

SharedL3Out-West-Pod2-Leaf_201-202

201,202

Previous

Cancel

Next

7. Click **Next**.
8. In the **Associations** window, select the previously created **Interface Selector Profiles** from the list.

APIC

admin

System

Tenants

Fabric

Virtual Networking

L4-L7 Services

Admin

Operations

Apps

Inventory

Fabric Policies

Access Policies

Policies

Quick Start

Switches

Leaf Switches

Profiles

Policy Groups

Overrides

Spine Switches

Modules

Interfaces

Policies

Pools

Physical and External Domains

Leaf Switches - Profiles

Name	Leaf Selectors (Switch Policy Group)	Interface Selectors	Module Selectors	Description
SharedL3Out-West-Pod1-Leaf_PR	101-102	SharedL3Out-West-Pod1_IPR		

Create Leaf Profile

1. Profile

2. Associations

STEP 2 > Associations

Select the interface/module selector profiles to associate

Interface Selector Profiles:

Select	Name	Description
<input type="checkbox"/>	SharedL3Out-West-Pod1_IPR	
<input checked="" type="checkbox"/>	SharedL3Out-West-Pod2_IPR	

Module Selector Profiles:

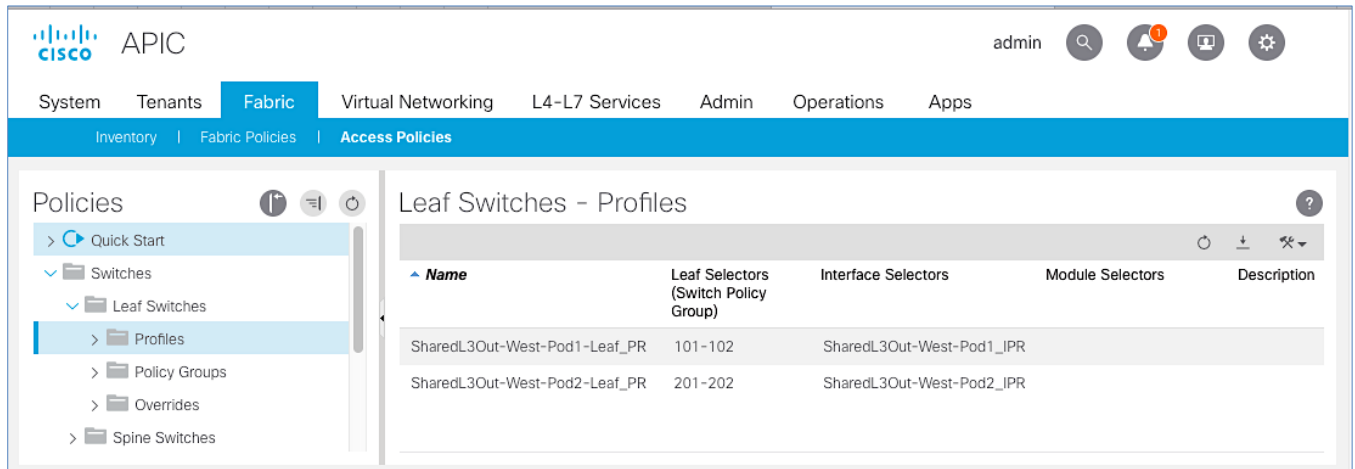
Select	Name	Description
--------	------	-------------

Previous

Cancel

Finish

9. Click **Finish** to complete.
10. You should now see the profiles for both Pods as shown below.



Configure Tenant Networking for Shared L3Out

Pod-2 uses the same Tenant, VRF and Bridge Domain as Pod-1 for Shared L3Out. No additional configuration is therefore necessary to enable Tenant Networking in Pod-2. The figure below shows the Tenant networking for Shared L3Out that was configured during Pod-1 setup. For more information, see [Shared L3Out deployment](#) in the Pod-1 section.

Table 31 Tenant Networking for Shared L3Out

Shared L3Out	Tenant Name	VRF	Bridge Domain
	common	common-SharedL3Out_VRF	common-SharedL3Out_BD

Configure External Routed Networks under Tenant Common

Table 32 Routed Outside – Pod-1

Shared L3Out - Pod-2	Routed Outside Name	Routed Node Profile	Router IDs (/32 Mask)	Node IDs	Node Interface Profile	OSPF Policy
	SharedL3Out-West-Pod2_RO	SharedL3Out-West-Pod2-	14.14.14.1	201	SharedL3Out-West-	SharedL3Out-West-Pod2-OSPF_Policy
	OSPF Area 10 (NSSA)	Node_PR	14.14.14.2	202	Pod2-Node_IPR	✓ Point-to-point ✓ MTU ignore)
	Routed Sub-interface	VLAN	Subnet	External Network		
	Eth1/47	315	10.114.1.0/30	Default-Route (0.0.0.0/0)		
	Eth1/48	316	10.114.1.4/30	✓ External Subnets for the External EPG		
	Eth1/47	317	10.114.2.0/30	✓ Shared Route Control Subnet		
	Eth1/48	318	10.114.2.4/30	✓ Shared Security Import Subnet		

To specify the domain type to connect to external gateway routers outside the ACI fabric, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Tenants > common**.
3. In the left navigation pane, select and expand **Tenant common > Networking > External Routed Networks**.
4. Right-click and select **Create Routed Outside**.

- In the **Create Routed Outside** pop-up window, specify a **Name** (for example, SharedL3Out-West-Pod2_RO). Select the check box next to **OSPF**. For the **OSPF Area ID**, enter 0.0.0.10 (should match the external gateway configuration). For the **VRF**, select the previously created VRF from the drop-down list. For the **External Routed Domain**, select the previously created domain from the drop-down list. For **Nodes and Interfaces Protocol Profiles**, click **[+]** to add a Node Profile.

The screenshot shows the 'Create Routed Outside' configuration window in the Cisco APIC. The window is titled 'Create Routed Outside' and has a progress bar with '1. Identity' and '2. External EPG Networks'. The 'Define the Routed Outside' section includes the following fields and options:

- Name:** SharedL3Out-West-Pod2_RO
- Alias:** (empty)
- Description:** optional
- Tags:** (empty)
- PIM:** ☐
- Route Control Enforcement:** ☐ Import ☒ Export
- Target DSCP:** Unspecified
- VRF:** common-SharedL3Out_VRF
- External Routed Domain:** SharedL3Out-West-Pod2_Dom
- Route Profile for Interleaf:** select a value
- Route Control For Dampening:** (empty)
- Provider Label:** (empty)
- Consumer Label:** (empty)
- OSPF:** ☒ BGP ☐ EIGRP ☒ OSPF
- OSPF Area ID:** 0.0.0.10
- OSPF Area Control:** ☒ Send redistributed LSAs into NSSA area ☒ Originate summary LSA ☐ Suppress forwarding address in translated LSA
- OSPF Area Type:** NSSA area Regular area Stub area
- OSPF Area Cost:** 1

The 'Nodes and Interfaces Protocol Profiles' section at the bottom has a table with columns Name, Description, DSCP, and Nodes, and a '+' button to add a profile.

- In the **Create Node Profile** pop-up window, specify a profile **Name** (for example, SharedL3Out-West-Pod2-Node_PR). For **Nodes**, click **[+]** to add a Node.
- In the **Select Node** pop-up window, for the **Node ID**, select first Border Leaf switch from the drop-down list. For the **Router ID**, specify the router ID for the first Border Leaf Switch (for example, 14.14.14.1). Click **OK** to complete selecting the Node. Repeat to add the second Border Leaf to the list of Nodes. For **OSPF Interface Profiles**, click **[+]** to add a profile.

The screenshot displays the Cisco APIC (Application Policy Infrastructure Controller) interface. The main navigation bar includes tabs for System, Tenants, Fabric, Virtual Networking, L4-L7 Services, Admin, Operations, and Apps. The 'Tenants' tab is active, showing a list of tenants under 'Tenant common'. The 'External Routed Networks' section is visible, with a table listing 'default' and 'SharedL3Out-West-Pod1_RO'.

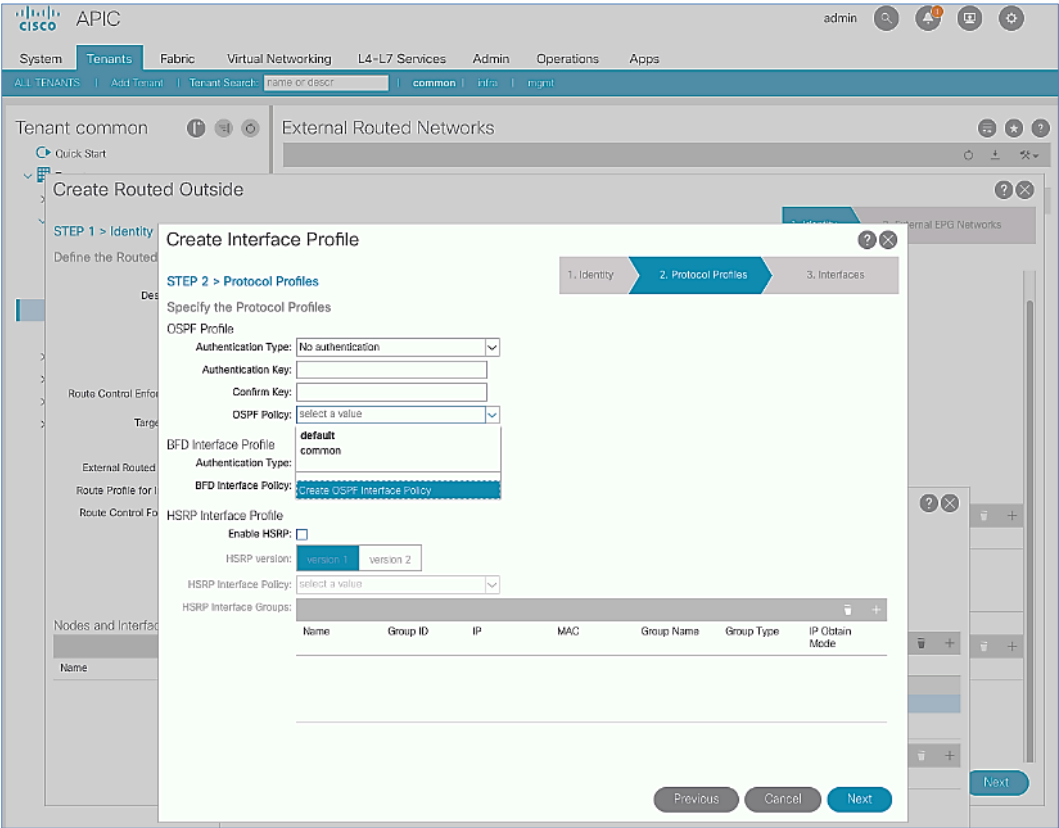
The 'Create Routed Outside' pop-up window is open, showing 'STEP 1 > Identity'. The 'Define the Routed Outside' section includes fields for Name (SharedL3Out-West-Pod2_RO), Alias, Description (optional), and Tags. The 'Route Control Enforcement' section has checkboxes for Import and Export, with 'Export' selected. The 'VRF' is set to 'common-SharedL3Out_VRF'. The 'External Routed Domain' is 'SharedL3'. The 'Route Profile for Interleaf' is 'select a v'. The 'Route Control For Dampening' is 'Add'.

The 'Create Node Profile' pop-up window is also open, showing 'Specify the Node Profile'. The 'Name' is 'SharedL3Out-West-Pod2-Node_PR', and the 'Description' is 'optional'. The 'Target DSCP' is 'Unspecified'. The 'Nodes' table lists the following data:

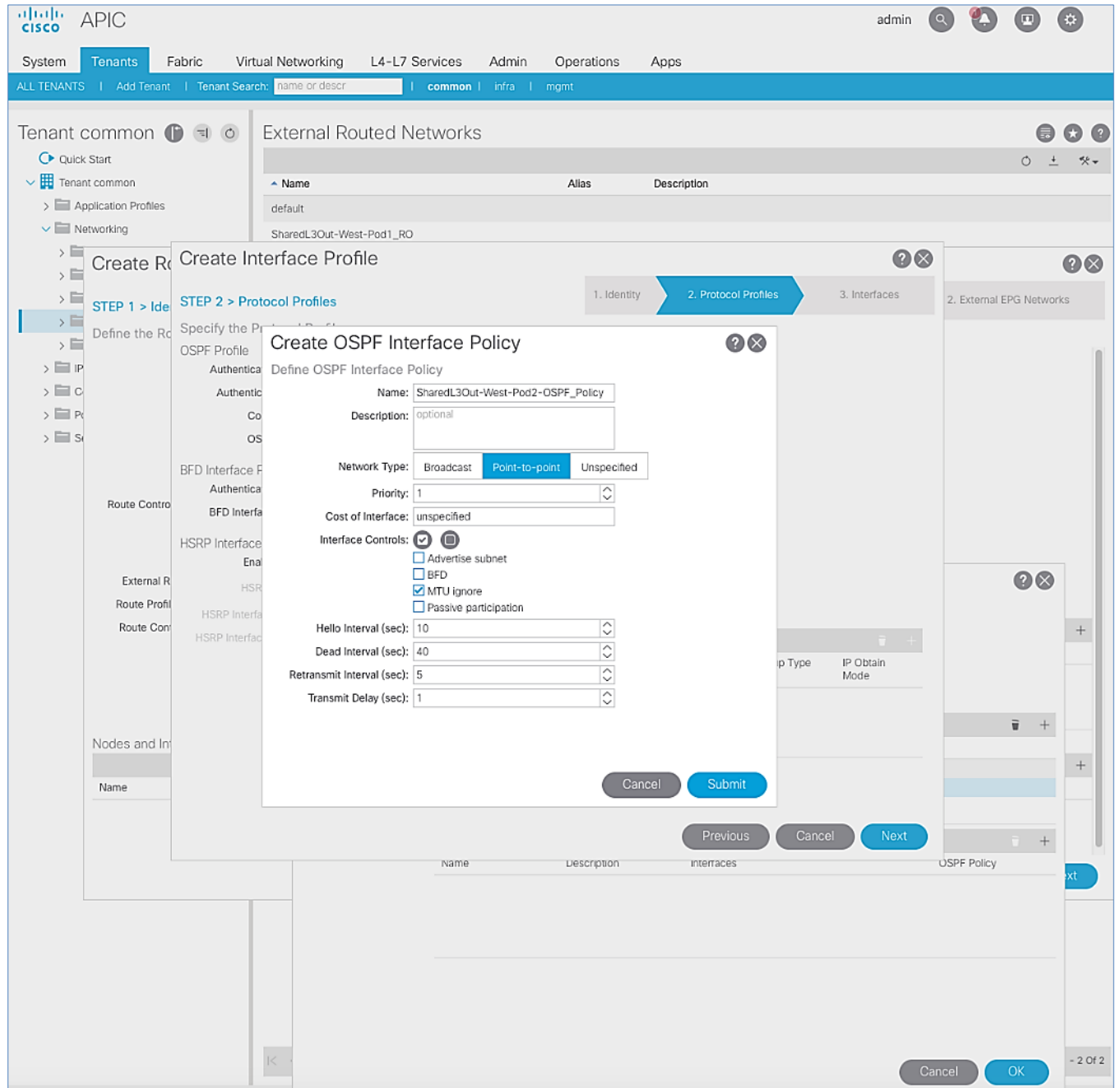
Node ID	Router ID	Static Routes	Loopback Address
topology/pod-2/...	14.14.14.1		14.14.14.1
topology/pod-2/...	14.14.14.2		14.14.14.2

The 'OSPF Interface Profiles' section is empty. The 'OSPF Area ID' is '0.0.0.10'. The 'OSPF Area Type' is 'NSSA area'. The 'OSPF Area Control' section has checkboxes for 'Send redistributed LSAs into NSSA area', 'Originate summary LSA', and 'Suppress forwarding address in translated LSA'. The 'OSPF Area Type' is 'NSSA area'.

8. In the **Create Interface Profile** pop-up window, for **Step 1 > Identity**, specify a **Name** (for example, SharedL3Out-West-Pod2-Node_IPR). Click **Next**. In **Step 2 > Protocol Profiles**, for the **OSPF Policy**, use the drop-down list to select **Create OSPF Interface Policy**.



9. In the **Create OSPF Interface Policy** pop-up window, specify a **Name** (for example, SharedL3Out-West-Pod2-OSPF_Policy) . For **Network Type**, select **Point-to-Point**. For **Interface Controls**, select the checkbox for **MTU ignore**.



10. Click **Submit** to complete creating the OSPF policy.
11. In the **Create Interface Profile** pop-up window, for the **OSPF Policy**, the newly created policy should now show up as the policy.

APIC admin

System **Tenants** Fabric Virtual Networking L4-L7 Services Admin Operations Apps

ALL TENANTS | Add Tenant | Tenant Search: name or descr | common | infra | mgmt

Tenant common

Quick Start

Tenant common

Application Profiles

Networking

External Routed Networks

Name Alias Description

default

SharedL3Out-West-Pod1_RO

Create Interface Profile

1. Identity 2. Protocol Profiles 3. Interfaces

STEP 2 > Protocol Profiles

Specify the Protocol Profiles

OSPF Profile

Authentication Type: No authentication

Authentication Key:

Confirm Key:

OSPF Policy: SharedL3Out-West-Pod2-OSPF_Pi

BFD Interface Profile

Authentication Type: No authentication

BFD Interface Policy: select a value

HSRP Interface Profile

Enable HSRP: ☐

HSRP version: version 1 version 2

HSRP Interface Policy: select a value

HSRP Interface Groups:

Name	Group ID	IP	MAC	Group Name	Group Type	IP Obtain Mode
------	----------	----	-----	------------	------------	----------------

Previous Cancel Next

12. Click **Next**.
13. For **STEP 3 > Interfaces**, select the tab for **Routed Sub-Interface**. Click **[+]** on the right side of the window to add a routed sub-interface.
14. In the **Select Routed Sub-Interface** pop-up window, for **Node**, select the first Border Leaf. For **Path**, select the interface (for example, 1 / 47) on the first Border Leaf that connects to the first external gateway. For **Encap**, specify the VLAN (for example, 315). For **IPv4 Primary / IPv6 Preferred Address**, specify the address (for example, 10.114.1.1/30).

The screenshot displays the Cisco APIC interface for configuring a tenant. The 'Create Interface Profile' dialog is open, showing the '3. Interfaces' step. The 'Routed Sub-Interface' tab is selected, and the 'Specify the Interface' section is active. The configuration fields are as follows:

- Path Type:** Port (selected), Direct Port Channel
- Node:** BB06-9372PX-WEST-1 (Node)
- Path:** eth1/47
- Description:** optional
- Encap:** VLAN 315
- IPv4 Primary / IPv6 Preferred Address:** 10.114.1.1/30
- MAC Address:** 00:22:BD:F8:19:FF
- MTU (bytes):** inherit
- Link-local Address:** (empty)

The 'Routed Sub-Interfaces' table is currently empty. The background shows the 'External Routed Networks' table with a row for 'SharedL3Out-West-Pod1_RO'.

15. Click **OK** to complete configuring the first routed sub-interface.
16. In **STEP 3 > Interfaces**, under **Routed Sub-Interface** tab, click **[+]** again to create the next sub-interface that connects the first Border Leaf to the second Gateway.

The screenshot displays the Cisco APIC (Application Policy Infrastructure Controller) interface. The main navigation bar includes tabs for System, Tenants, Fabric, Virtual Networking, L4-L7 Services, Admin, Operations, and Apps. The 'Tenants' tab is active, showing a list of tenants under 'Tenant common'. A sidebar on the left contains a tree view with options like 'Quick Start', 'Tenant common', 'Application Profiles', 'Networking', and 'Route Control'.

The 'Create Interface Profile' dialog box is open, showing 'STEP 3 > Interfaces'. It contains a table of 'Routed Sub-Interfaces' with the following data:

Path	IP Address	MAC Address	MTU (bytes)
Pod-2/Node-201/eth1/47	10.114.1.1/30	00:22:BD:F8:19:FF	inherit

A sub-dialog 'Select Routed Sub-Interface' is also open, allowing for detailed configuration of a sub-interface. The configuration fields include:

- Path Type:** Port (selected), Direct Port Channel
- Node:** BB06-9372PX-WEST-1 (Node)
- Path:** eth1/48
- Description:** optional
- Encap:** VLAN (selected), 316
- IPv4 Primary / IPv6 Preferred Address:** 10.114.1.5/30
- IPv4 Secondary / IPv6 Additional Addresses:** (empty)
- MAC Address:** 00:22:BD:F8:19:FF
- MTU (bytes):** inherit
- Link-local Address:** (empty)

Buttons for 'Cancel' and 'OK' are visible at the bottom of the dialog.

17. Click **OK** to complete configuring the first routed sub-interface.
18. Repeat steps 1-17 to create two more sub-interfaces on the second Border Leaf switch to connect to the two external gateways.

CISCO APIC admin

System **Tenants** Fabric Virtual Networking L4-L7 Services Admin Operations Apps

ALL TENANTS | Add Tenant | Tenant Search: name or descr | common | infra | mgmt

Tenant common

- Quick Start
- Tenant common
 - Application Profiles
 - Networking
 - External Routed Networks

External Routed Networks

Name	Alias	Description
default		
SharedL3Out-West-Pod1_RO		

Create Routed Outside

STEP 1 > Identity

Define the Routed Outside

Name: SharedL3Out-West-Pod2_RO

Alias:

Provider Label: enter names separated by comma

Create Interface Profile

STEP 3 > Interfaces

Specify the Interfaces

Routed Interfaces SVI Routed Sub-Interface

Routed Sub-Interfaces			
Path	IP Address	MAC Address	MTU (bytes)
Pod-2/Node-201/eth1/47	10.114.1.1/30	00:22:BD:F8:19:FF	inherit
Pod-2/Node-201/eth1/48	10.114.1.5/30	00:22:BD:F8:19:FF	inherit
Pod-2/Node-202/eth1/47	10.114.2.1/30	00:22:BD:F8:19:FF	inherit
Pod-2/Node-202/eth1/48	10.114.2.5/30	00:22:BD:F8:19:FF	inherit

Previous Cancel OK

19. Click **OK** to complete the Interface Profile configuration and to close the **Create Interface Profile** pop-up window.

The screenshot shows the Cisco APIC GUI with the 'Tenants' tab selected. The 'Tenant common' section is active, and the 'Create Routed Outside' pop-up window is open. The '1. Identity' step is selected, and the 'Define the Routed Outside' section is visible. The 'Name' field is set to 'SharedL3Out-West-Pod2_RO'. The 'Description' field is set to 'optional'. The 'Tags' field is empty. The 'PIM' checkbox is unchecked. The 'Route Control Enforcement' section has 'Import' unchecked and 'Export' checked. The 'Target DSCP' is set to 'Unspecified'. The 'Provider Label' and 'Consumer Label' fields are empty. The 'OSPF Area ID' is set to '0.0.0.10'. The 'OSPF Area Control' section has 'Send redistributed LSAs into NSSA area' checked, 'Originate summary LSA' checked, and 'Suppress forwarding address in translated LSA' unchecked. The 'OSPF Area Type' is set to 'NSSA area'. The 'Create Node Profile' pop-up window is also open, showing the 'Specify the Node Profile' section. The 'Name' field is set to 'SharedL3Out-West-Pod2-Node_PR'. The 'Description' field is set to 'optional'. The 'Target DSCP' is set to 'Unspecified'. The 'Nodes' table is visible, showing two nodes: 'topology/pod-2/...' with Router ID '14.14.14.1' and Loopback Address '14.14.14.1', and 'topology/pod-2/...' with Router ID '14.14.14.2' and Loopback Address '14.14.14.2'. The 'OSPF Interface Profiles' section is also visible, showing a profile named 'SharedL3Out-West-Pod2-Node_IPR' with interfaces '[eth1/47], [eth1/47], [eth1/48], [eth1/48]' and OSPF Policy 'SharedL3Out-West-Pod2-O...'. The 'Next' button is highlighted.

Create Routed Outside

STEP 1 > Identity

Define the Routed Outside

Name: SharedL3Out-West-Pod2_RO

Alias:

Description: optional

Tags:

PIM: ☐

Route Control Enforcement: ☐ Import ☒ Export

Target DSCP: Unspecified

Provider Label:

Consumer Label:

☐ BGP ☐ EIGRP ☒ OSPF

OSPF Area ID: 0.0.0.10

OSPF Area Control: ☒ Send redistributed LSAs into NSSA area ☒ Originate summary LSA ☐ Suppress forwarding address in translated LSA

OSPF Area Type: NSSA area Regular area Stub area

Create Node Profile

Specify the Node Profile

Name: SharedL3Out-West-Pod2-Node_PR

Description: optional

Target DSCP: Unspecified

Nodes:

Node ID	Router ID	Static Routes	Loopback Address
topology/pod-2/...	14.14.14.1		14.14.14.1
topology/pod-2/...	14.14.14.2		14.14.14.2

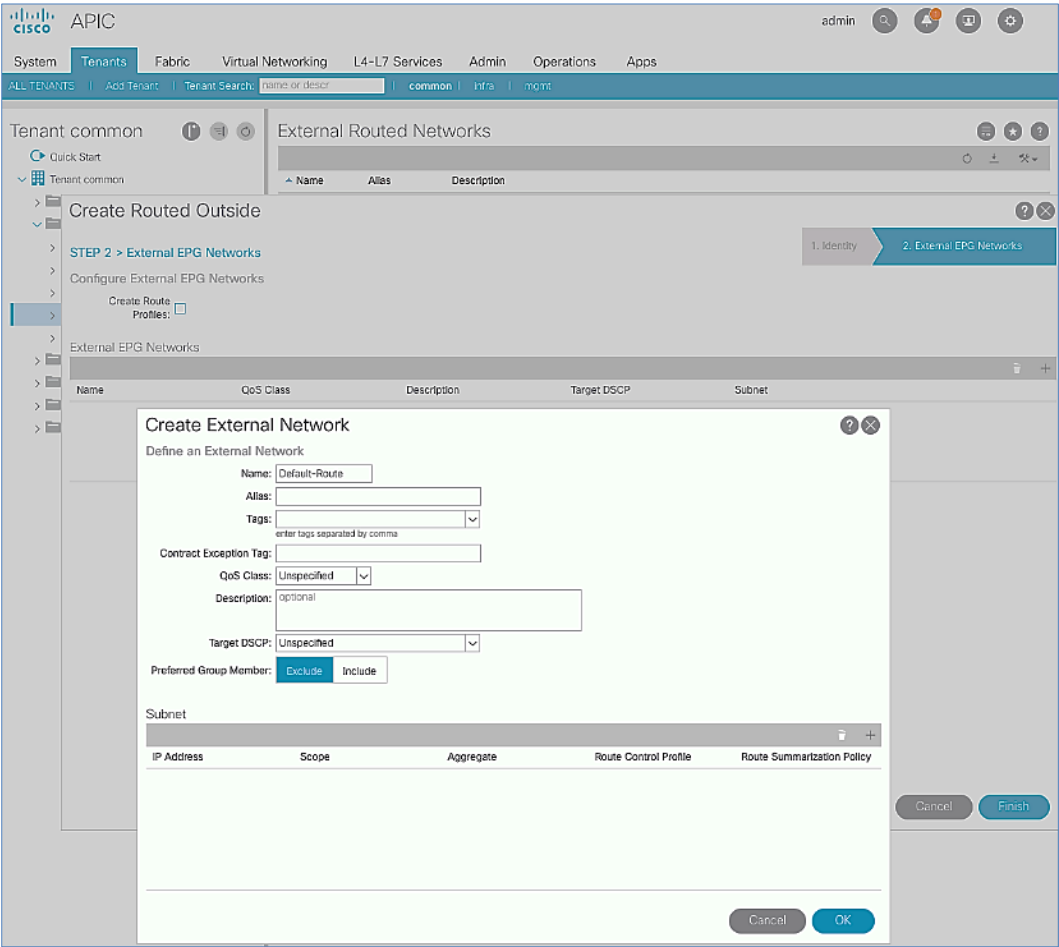
OSPF Interface Profiles:

Name	Description	Interfaces	OSPF Policy
SharedL3Out-West-Pod2-Node_IPR		[eth1/47], [eth1/47], [eth1/48], [eth1/48]	SharedL3Out-West-Pod2-O...

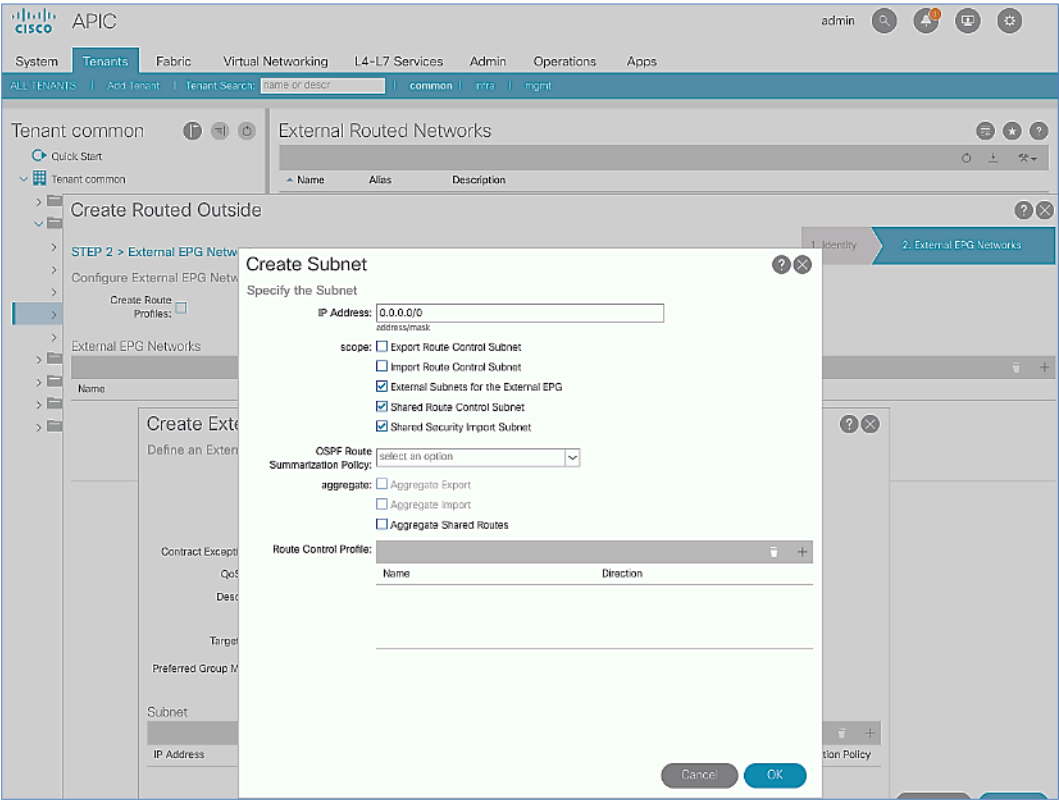
Cancel OK

Displaying Objects 1 - 2 Of 2

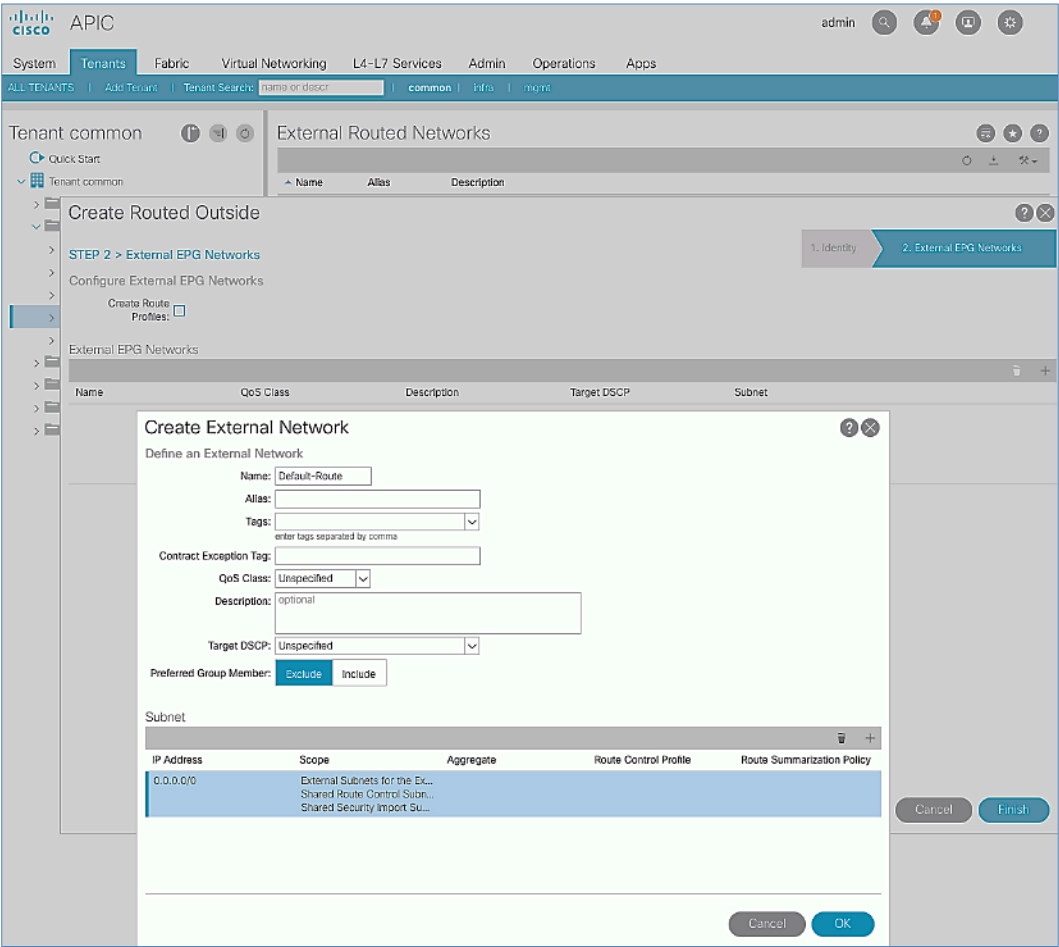
20. Click **OK** to complete the Node Profile configuration and to close the **Create Node Profile** pop-up window.
21. In the **Create Routed Outside** pop-up window, click Next. In **STEP 2 > External EPG Networks**, for **External EPG Networks**, click **[+]** to add an external network.
22. In the **Created External Network** pop-up window, specify a **Name** (for example, Default-Route) . For **Subnet**, click **[+]** to add a Subnet.



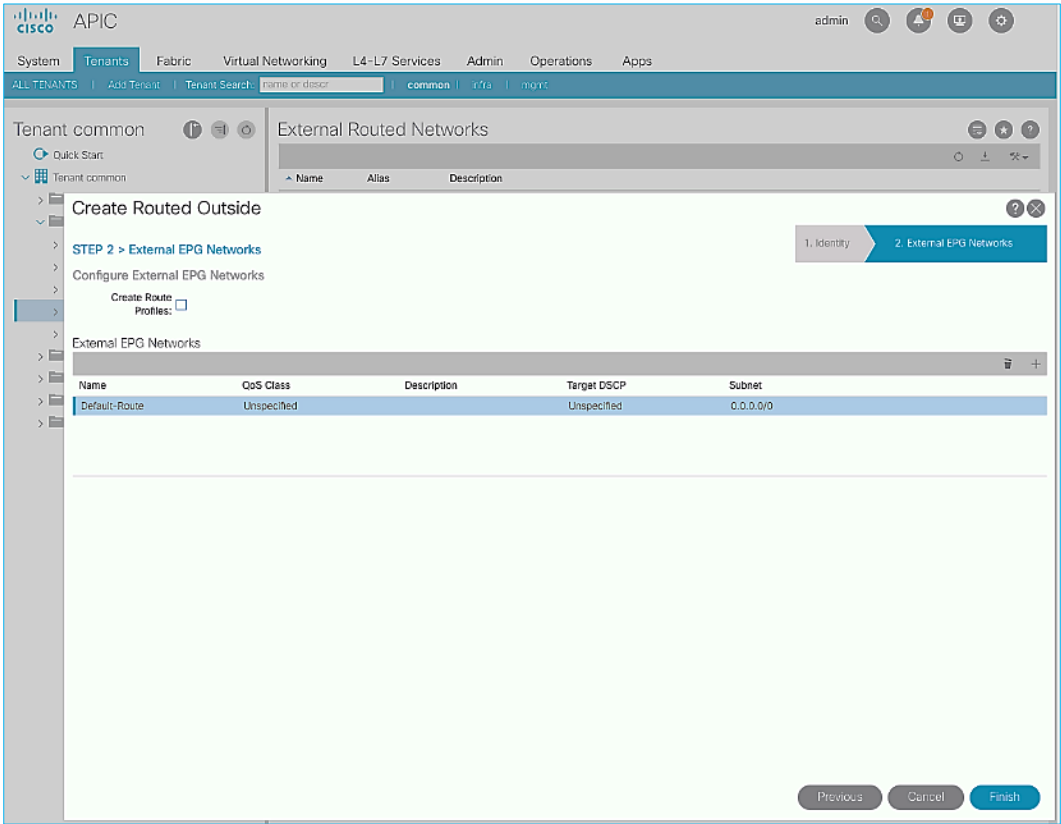
23. In the Create Subnet pop-up window, for the IP Address, enter a route (for example, 0.0.0.0/0) . Select the checkboxes for External Subnets for the External EPG, Shared Route Control Subnet, and Shared Security Import Subnet.



24. Click **OK** to complete creating the subnet and close the **Create Subnet** pop-up window.



25. Click **OK** again to complete creating the external network and close the **Create External Network** pop-up window.



26. Click **Finish** to complete creating the Routed Outside.

Create Contracts for External Routed Networks from Tenant (common)

Table 33 Contracts for External Routed Networks

Shared L3Out	Contract	Subject	Filter
	Allow-Shared-L3Out	Allow-Shared-L3Out	common/default ✓ Global Scope

This contract for external routed networks under Tenant **common** was already created during Pod-1 setup and does not need to be re-created here unless a different contract is being applied to Pod-2.

Provide Contracts for External Routed Networks from Tenant (common)

Table 34 Contracts for External Routed Networks

Shared L3Out	Contract	Subject	Filter
	Allow-Shared-L3Out	Allow-Shared-L3Out	common/default ✓ Global Scope

To provide contracts for external routed networks from Tenant **common**, follow the steps outlined below. The steps are similar to the contract provided in Pod-1 for accessing outside networks using the shared layer 3 connection in Pod-1.

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Tenants > common**.
3. In the left navigation pane, select and expand **Tenant common > Networking > External Routed Networks**.
4. Select and expand the recently created External Routed Network for SharedL3out or Routed Outside network (for example, `SharedL3Out-West-Pod2_RO`).
5. Select and expand **Networks**.
6. Select the recently created route (for example, `Default-Route`).
7. In the right window pane, select the tab for **Policy** and then **Contracts**.
8. Under the **Provided Contracts** tab, click **[+]** on the right to add a Provided Contract.
9. For **Name**, select the previously created contract (for example, `common/Allow-Shared-L3Out`) from the drop-down list.
10. Click **Update**.
11. Other Tenants can now 'consume' the `Allow-Shared-L3Out` contract to route traffic outside the ACI fabric. This deployment example shows a default filter to allow all traffic.



Customers can modify this contract as needed to limit access to specific destinations through the Shared L3Out connection .

Configure External Gateways in the Outside Network

This section provides a sample configuration from the Nexus switches that serve as external Layer 3 Gateways for Pod-2. The gateways are in the external network and peer with ACI border leaf switches in Pod-2 using OSPF. The gateway configuration shown below shows only the relevant portion of the configuration – it is not the complete configuration .

Enable Protocols

The protocols used between the ACI border leaf switches and external gateways have to be explicitly enabled on Nexus platforms used as external gateways in this design. The configuration to enable these protocols are provided below.

Table 35 External Gateways for Pod-2 – Protocols

External Gateway Configuration - Pod-2	BB-West-Enterprise-1 (GW-1)	BB-West-Enterprise-2 (GW-2)
	<code>feature ospf</code>	<code>feature ospf</code>
	<code>feature interface-vlan</code>	<code>feature interface-vlan</code>
	<code>feature lacp</code>	<code>feature lacp</code>
	<code>feature lldp</code>	<code>feature lldp</code>

Configure OSPF

OSPF is used between the external gateways and ACI border leaf switches to exchange routing between the two domains. The global configuration for OSPF is provided below. Loopback is used as the router IDs for OSPF. Note that interfaces between ACI border leaf switches will be in OSPF Area 10.

Table 36 External Gateways for Pod-2 – Protocols

External Gateway Configuration - Pod-2	BB-West-Enterprise-1 (GW-1)	BB-West-Enterprise-2 (GW-2)
	<pre> interface loopback0 description RID for OSPF ip address 14.14.14.98/32 ip router ospf 10 area 0.0.0.0 router ospf 10 router-id 14.14.14.98 area 0.0.0.10 nssa no-summary no- redistribution default-information-originate </pre>	<pre> interface loopback0 description RID for OSPF ip address 14.14.14.99/32 ip router ospf 10 area 0.0.0.0 router ospf 10 router-id 14.14.14.99 area 0.0.0.10 nssa no-summary no- redistribution default-information-originate </pre>

Configure Interfaces

The interface level configuration for connectivity between external gateways and ACI border leaf switches is provided below. Note that interfaces between ACI border leaf switches are in OSPF Area 10 while the loopbacks and port-channel links between the gateways are in OSPF Area 0.

Table 37 Interface Configuration – To ACI Border Leaf Switches

External Gateway Configuration - Pod-2	BB-West-Enterprise-1 (GW-1)	BB-West-Enterprise-2 (GW-2)
	<pre> interface Ethernet4/16 description To BB06-9372PX-WEST-1:Eth1/47 no shutdown interface Ethernet4/16.315 encapsulation dot1q 315 ip address 10.114.1.2/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.10 no shutdown </pre>	<pre> interface Ethernet4/16 description To BB06-9372PX-WEST-1:Eth1/48 no shutdown interface Ethernet4/16.316 encapsulation dot1q 316 ip address 10.114.1.6/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.10 no shutdown </pre>
	<pre> interface Ethernet4/20 description To BB06-9372PX-WEST-2:Eth1/47 no shutdown interface Ethernet4/20.317 encapsulation dot1q 317 ip address 10.114.2.2/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.10 no shutdown </pre>	<pre> interface Ethernet4/20 description To BB06-9372PX-WEST-2:Eth1/48 no shutdown interface Ethernet4/20.318 encapsulation dot1q 318 ip address 10.114.2.6/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.10 no shutdown </pre>

The configuration on the port-channel with 2x10GbE links that provide direct connectivity between the external gateways is provided below.

Table 38 Interface Configuration – Between External Gateways

External Gateway Configuration - Pod-2	BB-West-Enterprise-1 (GW-1)	BB-West-Enterprise-2 (GW-2)
	<pre> interface port-channel14 description To BB02-7004-2-BB-West-Enterprise-2 ip address 10.114.98.1/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.0 </pre>	<pre> interface port-channel14 description To BB02-7004-1-BB-West-Enterprise-1 ip address 10.114.98.2/30 ip ospf network point-to-point ip ospf mtu-ignore ip router ospf 10 area 0.0.0.0 </pre>
	<pre> interface Ethernet4/13 description To BB02-7004-2-BB-West-Enterprise-2:Eth4/13 channel-group 14 mode active no shutdown </pre>	<pre> interface Ethernet4/13 description To BB02-7004-1-BB-West-Enterprise-1:Eth4/13 channel-group 14 mode active no shutdown </pre>
	<pre> interface Ethernet4/17 description To BB02-7004-2-BB-West-Enterprise-2:Eth4/17 channel-group 14 mode active no shutdown </pre>	<pre> interface Ethernet4/17 description To BB02-7004-1-BB-West-Enterprise-1:Eth4/17 channel-group 14 mode active no shutdown </pre>

Solution Deployment – ACI Fabric (To Cisco UCS Domains)

This section provides detailed procedures for configuring the ACI fabric to connect to Cisco UCS domains in the access layer. The access layer setup will enable network connectivity for Cisco HyperFlex clusters that connect to the Cisco UCS domains in each data center or Pod.



The procedures outlined in this section are the same as that for a single ACI fabric except that there are two pairs of leaf switches (one for each Pod) physically located in different data centers.

Deploy New Leaf Switches for Connectivity to Cisco UCS Domains

Leaf switches provide access to the ACI fabric. In this design, dedicated leaf switches are deployed to connect Cisco HyperFlex and UCS domains to the ACI fabric in both data centers. These leaf switches are separate from the leaf switch pair used for the shared L3Out connectivity to outside networks in each Pod. In ACI, new ACI-capable switches are automatically discovered using Link Layer Discovery Protocol (LLDP). The discovered switches are then added, provisioned and managed from the APIC web GUI. All configuration is centralized and managed through the APIC – there is no individual configuration of the Spine and Leaf switches.

In this section, the procedure for discovering and provisioning new leaf switch pairs in each Pod for connecting to Cisco HyperFlex and UCS domains will be explained.

Topology

Figure 16 ACI Fabric Topology – New Leaf Switches in Pod-1

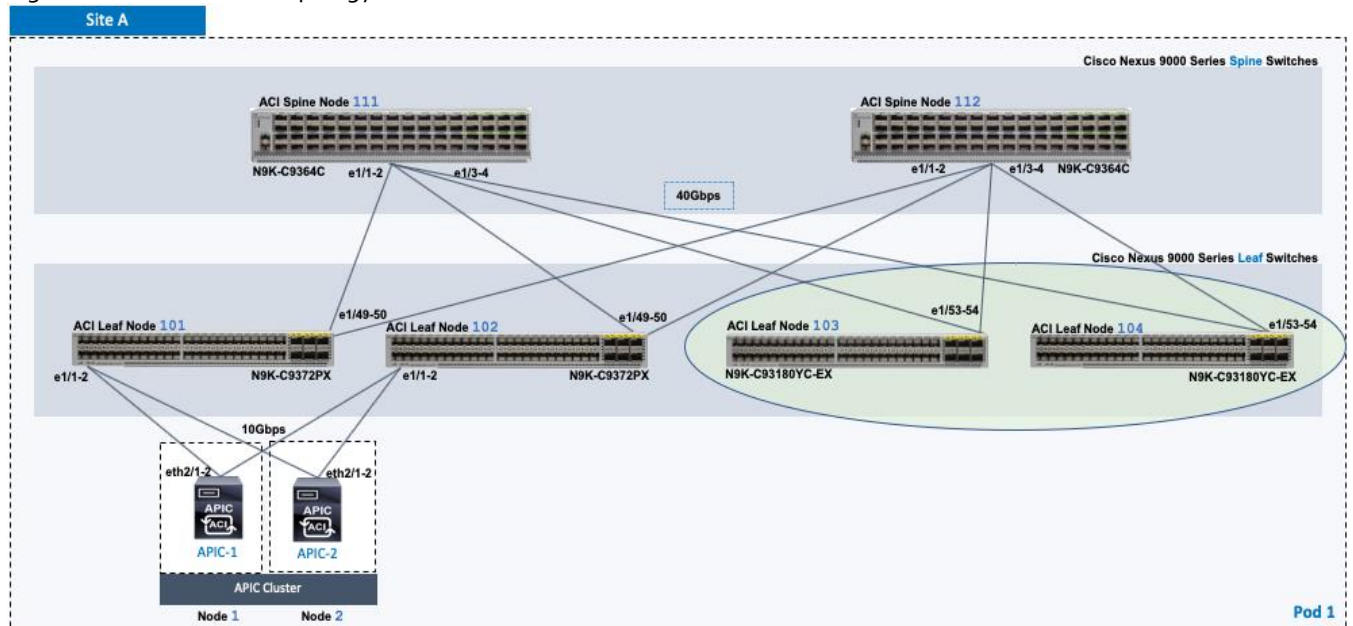
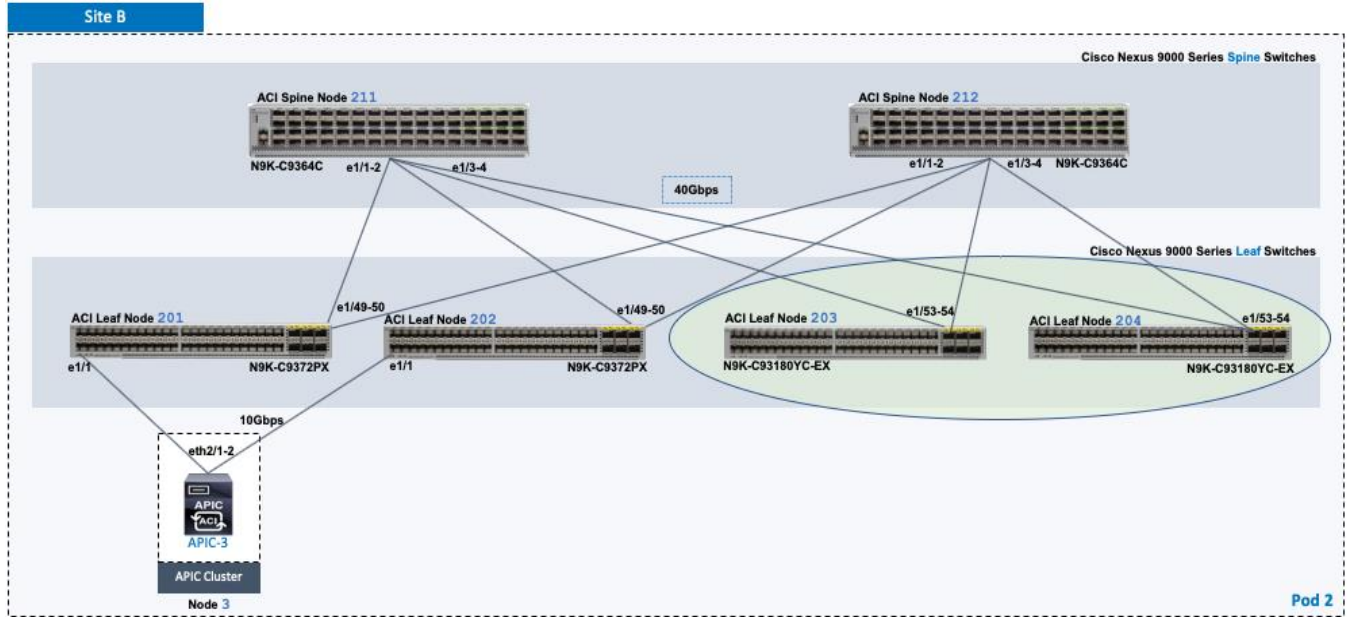


Figure 17 ACI Fabric Topology – New Leaf Switches in Pod-2



Setup Information

Table 39 Pod-1 Leaf Switches - For Connectivity to Cisco UCS and HyperFlex Domains

To Cisco UCS & HyperFlex Domain	General	Node ID	Node Names	OOB Management EPG	OOB Management IP	OOB Gateway
	Pod ID: 1	103	AA07-93180YC-EX-WEST-1	default	172.26.163.37/24	172.26.163.254
	Role: Leaf					
	Rack Name (Optional): AA07	104	AA07-93180YC-EX-WEST-2	default	172.26.163.38/24	172.26.163.254

Table 40 Pod-2 Leaf Switches - For Connectivity to Cisco UCS and HyperFlex Domains

To Cisco UCS & HyperFlex Domain	General	Node ID	Node Names	OOB Management EPG	OOB Management IP	OOB Gateway
	Pod ID: 2	203	BB06-93180YC-EX-WEST-1	default	172.26.164.37/24	172.26.164.254
	Role: Leaf					
	Rack Name (Optional): BB06	204	BB06-93180YC-EX-WEST-2	default	172.26.164.38/24	172.26.164.254

ACI Fabric Discovery of Leaf Switches

ACI automatically discovers new switches (running ACI software) through LLDP when they are connected to the ACI fabric. To verify that the ACI fabric has discovered the two leaf switches deployed in Pod-1/Site A for connecting Cisco UCS and HyperFlex systems, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top menu, select **Fabric > Inventory**.
3. In the left navigation pane, select **Fabric Membership**.

- In the right window pane, select the **Nodes Pending Registration** tab. The newly discovered Leaf Switches will be listed with a Node ID of '0'.

The screenshot shows the Cisco APIC GUI with the 'Fabric' tab selected in the top navigation bar. Under 'Fabric', the 'Inventory' sub-tab is active. In the left-hand 'Inventory' pane, 'Fabric Membership' is selected. The main right-hand pane shows the 'Fabric Membership' section with the 'Nodes Pending Registration' tab highlighted. Above the table, there are three large circular indicators: '0' for 'Unsupported', '0' for 'Undiscovered', and '0' for 'Unknown'. Below these, a table lists the discovered switches.

Serial Number	Pod ID	Node ID	RL TEP Pool	Name	Role	Support Model	SSL Certificate	Status
FDO211304ZX	1	0	0		leaf	yes	n/a	
FDO211314G7	1	0	0		leaf	yes	n/a	

- Note the serial numbers of the newly discovered leaf switches.
- Determine which node will be the -1 and -2 switches in the new leaf switch pair.
- Repeat steps 1-6 for Pod-2/Site-2 leaf switches.

Add Nexus 9000 Series Leaf Switches to the ACI Fabric

To add the newly discovered Nexus 9318oYC-EX leaf switches from the previous step, follow these steps:

- Identify the -1 and -2 switches in the new leaf switch pair based on their physical connectivity into the fabric.
- Determine the serial numbers corresponding to the -1 and -2 switches to map it to the ones collected in the previous step. To find the serial number for a given leaf switch, access its serial console, log in using admin account (no password) and run the command: **show inventory**.
- Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
- From the top menu, select **Fabric > Inventory**.
- In the left navigation pane, select **Fabric Membership**.
- In the right window pane, select the **Nodes Pending Registration** tab. From the list of switches, select the serial number corresponding to the -1 leaf. Right-click and select **Register** from the menu.
- In the **Register** pop-up window, enter the **Pod ID**, **Node ID** and a **Node Name** for the selected Leaf switch.

The screenshot shows the Cisco APIC interface with the 'Fabric' tab selected. The 'Fabric Membership' section is active, displaying the 'Nodes Pending Registration' tab. A modal window titled 'Register' is open, allowing the user to add a new node to the fabric. The modal contains the following fields and values:

- Serial Number: FDO211304ZX
- Pod ID: 1
- Node ID: 103
- RL TEP Pool: 0
- Role: leaf
- Node Name: AA07-93180YC-EX-WEST-1
- Rack Name: AA07 (site:fabric, building:default, fl)

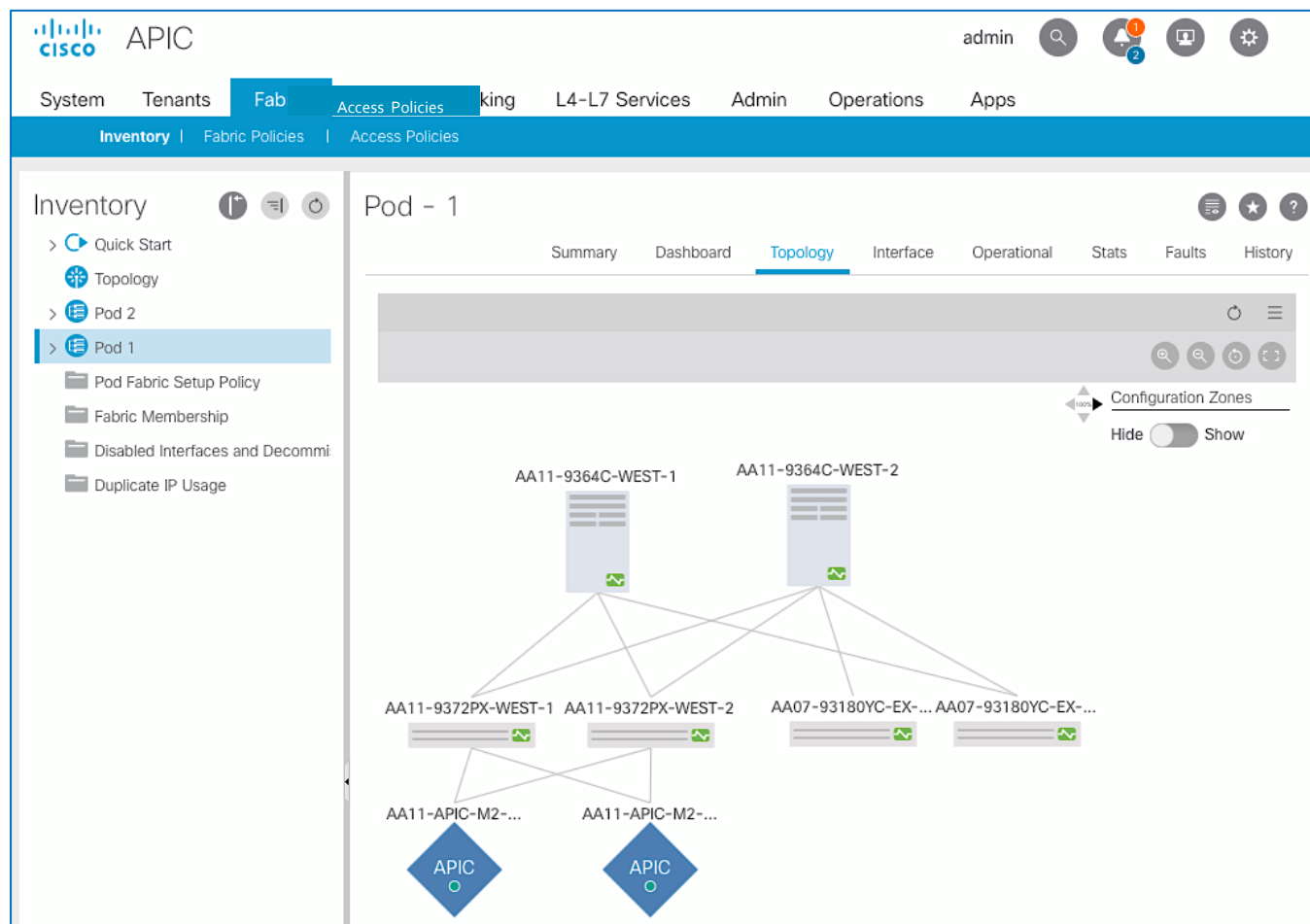
The modal also includes 'Cancel' and 'Register' buttons at the bottom.

8. Click **Register** to complete.
9. Repeat above steps to add the second or -2 Leaf switch to the fabric.
10. Select the tab for **Registered Nodes**. After a few minutes, the newly added switches should transition to a **Status** of **Active**.

The screenshot shows the Cisco APIC interface. The top navigation bar includes 'System', 'Tenants', 'Fabric' (selected), 'Virtual Networking', 'L4-L7 Services', 'Admin', 'Operations', and 'Apps'. Below this, there are tabs for 'Inventory', 'Fabric Policies', and 'Access Policies'. The left sidebar shows the 'Inventory' section with a tree view including 'Quick Start', 'Topology', 'Pod 2', 'Pod 1', 'Pod Fabric Setup Policy', 'Fabric Membership' (selected), 'Disabled Interfaces and Decomm...', and 'Duplicate IP Usage'. The main panel is titled 'Fabric Membership' and has tabs for 'Registered Nodes' (selected), 'Nodes Pending Registration', 'Unreachable Nodes', and 'Unmanaged Fabric Nodes'. The 'Registered Nodes' tab shows a summary: 6 Leafs (0 Decommissioned, 0 Maintenance, 6 Active, 0 Inactive) and 0 Virtual Leafs (0 Decommissioned, 0 Maintenance, 0 Active, 0 Inactive). Below the summary is a table of nodes.

Serial Number	Model	Pod ID	Node ID	Name	Role	IP	Status
SAL1940QA...	N9K-C9372PX	1	101	AA11-9372PX-WEST-1	leaf	10.13.64.64...	Active
SAL1940QA...	N9K-C9372PX	1	102	AA11-9372PX-WEST-2	leaf	10.13.184.6...	Active
FDO211304...	N9K-C93180YC-EX	1	103	AA07-93180YC-EX-WEST-1	leaf	10.13.184.6...	Active
FDO211314...	N9K-C93180YC-EX	1	104	AA07-93180YC-EX-WEST-2	leaf	10.13.64.65...	Active
FDO22240V...	N9K-C9364C	1	111	AA11-9364C-WEST-1	spine	10.13.184.6...	Active
FDO22240VJ8	N9K-C9364C	1	112	AA11-9364C-WEST-2	spine	10.13.184.6...	Active
SAL1913CJ...	N9K-C9372PX	2	201	BB06-9372PX-WEST-1	leaf	10.14.24.66...	Active
SAL1914CN...	N9K-C9372PX	2	202	BB06-9372PX-WEST-2	leaf	10.14.32.64...	Active
FDO221914...	N9K-C9364C	2	211	BB06-9364C-WEST-1	spine	10.14.24.64...	Active
FDO22182Q...	N9K-C9364C	2	212	BB06-9364C-WEST-2	spine	10.14.24.65...	Active

11. From the left navigation menu, navigate to the Pod (for example, **Pod 1**) that the Ngk switches were added to.
12. From the right-window pane, select the **Topology** tab to confirm the newly added switches are part of the Pod topology.



13. Repeat steps 1-12 using setup information for Pod-2/Site-2 leaf switches.

Setup Out-of-Band Management for New Leaf Switches

To enable out-of-band (OOB) management access to the switches in the ACI fabric, ACI provides a pre-defined **mgmt** Tenant. To enable OOB connectivity to the new leaf switches, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top menu, select **Tenants > mgmt**.
3. From the left navigation menu, select and expand Tenant mgmt > Node Management Addresses > Static Node Management Addresses.
4. Right-click Static Node Management Addresses and select **Create Static Node Management Addresses**.
5. In the **Create Static Node Management Addresses** pop-up window, enter the **node ID range** for the new leaf switches (in this case, **103-104**) and select the checkbox for **Out-Of-Band Addresses**. For the **Out-Of-Band Management EPG**, select **'default'** from the drop-down list. For **Out-Of-Band IPv4 Address** and **Out-Of-Band IPv4 Gateway**, specify the IP addresses for OOB management and Gateway.



Consecutive IP addresses will be assigned for the range of nodes so only a starting IP address needs to be specified.

The screenshot shows the APIC interface with the 'Tenants' tab selected. On the left, the 'Tenant mgmt' sidebar is visible, with 'Static Node Management Address' selected under 'Node Management Addresses'. The main panel displays the 'Static Node Management Addresses' configuration dialog. The dialog has a title bar with a question mark and a close button. Below the title bar, it says 'Specify policy name and a node range, and set their IPs.' The 'Node Range' is set to '103' (From) and '104' (To). The 'Config' section has 'Out-Of-Band Addresses' checked and 'In-Band Addresses' unchecked. Under 'Out-Of-Band Addresses', the 'Out-Of-Band Management EPG' is set to 'default'. The 'Out-Of-Band IPv4 Address' is '172.26.163.37/24' (address/mask). The 'Out-Of-Band IPv4 Gateway' is '172.26.163.254'. The 'Out-Of-Band IPv6 Address' and 'Out-Of-Band IPv6 Gateway' fields are empty. At the bottom of the dialog are 'Cancel' and 'Submit' buttons. In the background, a table of 'Static Node Management Addresses' is visible, showing columns for Node ID, Name, Type, EPG, IPv4 Address, IPv4 Gateway, IPv6 Address, and IPv6 Gateway. The table contains 10 rows of data.

Node ID	Name	Type	EPG	IPv4 Address	IPv4 Gateway	IPv6 Address	IPv6 Gateway
pod-1/node-101	AA11-9372PX-WEST-1	Out-Of-Band	default	172.26.163.117/24	172.26.163.254	::	::
pod-1/node-102	AA11-9372PX-WEST-2	Out-Of-Band	default	172.26.163.118/24	172.26.163.254	::	::
pod-1/node-111	AA11-9364C-WEST-1	Out-Of-Band	default	172.26.163.119/24	172.26.163.254	::	::
pod-1/node-112	AA11-9364C-WEST-2	Out-Of-Band	default	172.26.163.120/24	172.26.163.254	::	::
pod-2/node-201	BB06-9372PX-WEST-1	Out-Of-Band	default	172.26.164.117/24	172.26.164.254	::	::
pod-2/node-202	BB06-9372PX-WEST-2	Out-Of-Band	default	172.26.164.118/24	172.26.164.254	::	::
pod-2/node-211	BB06-9364C-WEST-1	Out-Of-Band	default	172.26.164.119/24	172.26.164.254	::	::
pod-2/node-212	BB06-9364C-WEST-2	Out-Of-Band	default	172.26.164.120/24	172.26.164.254	::	::
pod-1/node-103	AA07-93180YC-EX-WEST-1	Out-Of-Band	default	172.26.163.37/24	172.26.163.1	::	::
pod-1/node-104	AA07-93180YC-EX-WEST-2	Out-Of-Band	default	172.26.163.38/24	172.26.163.1	::	::

6. Click **Submit** and then click **Yes** to proceed with assigning the IP addresses.

The screenshot shows the APIC interface after the configuration has been submitted. The 'Static Node Management Addresses' table is now populated with 10 rows of data. The table has columns for Node ID, Name, Type, EPG, IPv4 Address, IPv4 Gateway, IPv6 Address, and IPv6 Gateway. The data is as follows:

Node ID	Name	Type	EPG	IPv4 Address	IPv4 Gateway	IPv6 Address	IPv6 Gateway
pod-1/node-101	AA11-9372PX-WEST-1	Out-Of-Band	default	172.26.163.117/24	172.26.163.254	::	::
pod-1/node-102	AA11-9372PX-WEST-2	Out-Of-Band	default	172.26.163.118/24	172.26.163.254	::	::
pod-1/node-111	AA11-9364C-WEST-1	Out-Of-Band	default	172.26.163.119/24	172.26.163.254	::	::
pod-1/node-112	AA11-9364C-WEST-2	Out-Of-Band	default	172.26.163.120/24	172.26.163.254	::	::
pod-2/node-201	BB06-9372PX-WEST-1	Out-Of-Band	default	172.26.164.117/24	172.26.164.254	::	::
pod-2/node-202	BB06-9372PX-WEST-2	Out-Of-Band	default	172.26.164.118/24	172.26.164.254	::	::
pod-2/node-211	BB06-9364C-WEST-1	Out-Of-Band	default	172.26.164.119/24	172.26.164.254	::	::
pod-2/node-212	BB06-9364C-WEST-2	Out-Of-Band	default	172.26.164.120/24	172.26.164.254	::	::
pod-1/node-103	AA07-93180YC-EX-WEST-1	Out-Of-Band	default	172.26.163.37/24	172.26.163.1	::	::
pod-1/node-104	AA07-93180YC-EX-WEST-2	Out-Of-Band	default	172.26.163.38/24	172.26.163.1	::	::

7. The newly added leaf switches should now be listed and reachable for OOB Management through SSH.
8. Repeat steps 1-7 using setup information for Pod-2/Site-2 leaf switches.

Enable Access Layer Connectivity to Cisco UCS Domains

To use the compute and storage resources provided by a Cisco HyperFlex cluster, the HyperFlex system must first be deployed on Cisco HX-series servers connected to Cisco UCS Fabric Interconnects. Cisco HyperFlex system can be deployed either:

- From the Cloud using Cisco Intersight or
- Using a HyperFlex installer virtual machine deployed in an existing virtualization environment

However, before a HyperFlex system can be deployed, the ACI fabric must provide connectivity from the HyperFlex installer to the HyperFlex nodes connected to Cisco UCS Fabric Interconnects in the Cisco UCS domain. To enable this end-to-end connectivity, the ACI fabric requires:

- Connectivity to the HyperFlex installer (Intersight or Installer VM) and other infrastructure services and networks required to complete the installation. This connectivity was provided by the Shared L3Out – see previous section.
- Physical connectivity to the Cisco UCS domain, consisting of a pair of UCS Fabric Interconnects. The HyperFlex servers are dual-homed to the Fabric Interconnects. A single UCS domain can support multiple HyperFlex clusters. In this design, a separate UCS domain is used for each HyperFlex cluster, and two for HyperFlex stretched cluster.
- Access layer configuration (or Fabric Access Policies) to enable connectivity to the Cisco UCS domain from the ACI fabric.

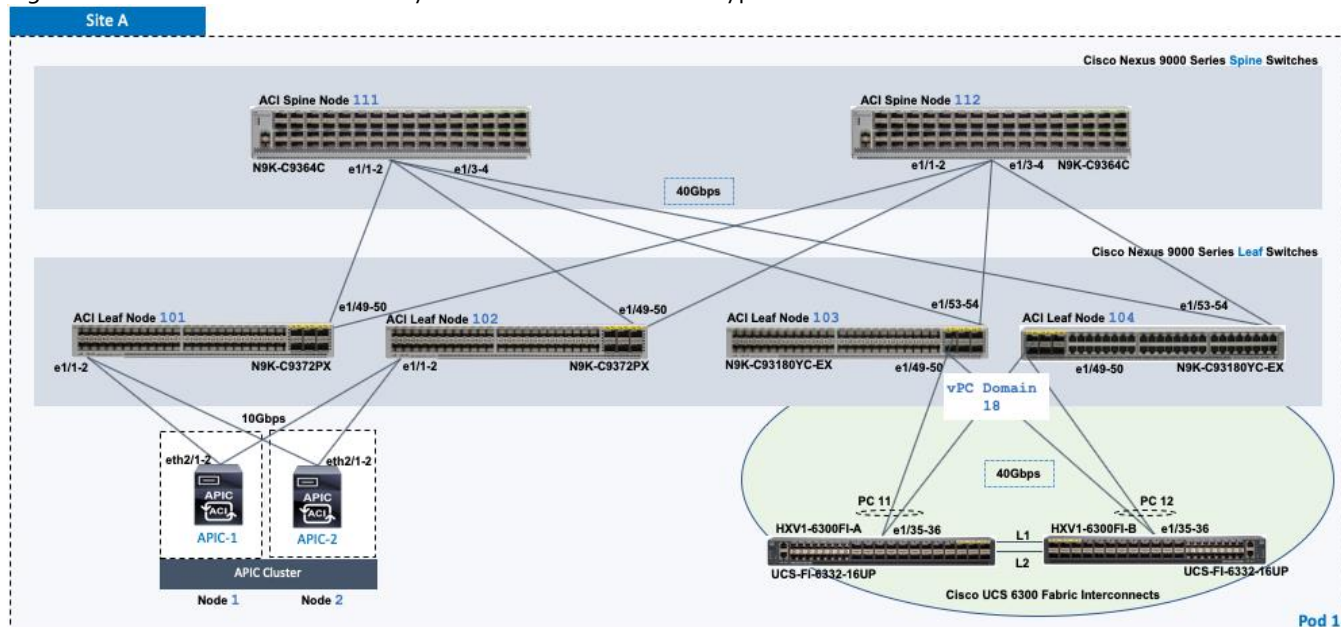
In this section, the ACI fabric configuration to enable connectivity to the Cisco UCS domains is provided. The physical connectivity between ACI Leaf switches and UCS domain is assumed to be in place but configuration to enable 40GbE connectivity (if needed) is done in this section. Two virtual Port Channels (vPCs) are established from the newly deployed Leaf switches (from previous section) to each Cisco UCS Fabric Interconnect (FI-A, FI-B) pair where a Cisco HyperFlex cluster will be deployed. The corresponding configuration in the UCS domains is covered in an upcoming section.

The procedures in this section will configure the ACI fabric to connect to the three UCS domains deployed in this solution.

Topology

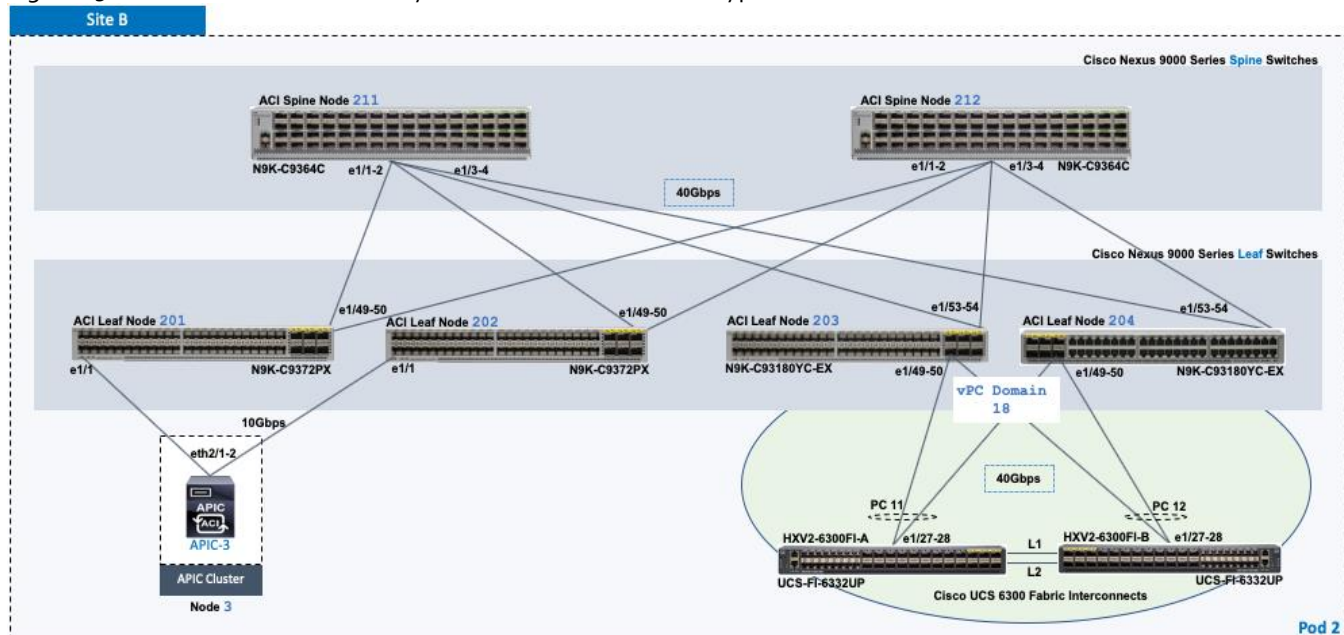
The ACI Fabric topology to connect to the UCS domain for the HyperFlex stretched cluster in Pod-1 is shown in [Figure 18](#).

Figure 18 ACI Fabric – Connectivity to Cisco UCS Domain for HyperFlex Stretched Cluster in Pod-1



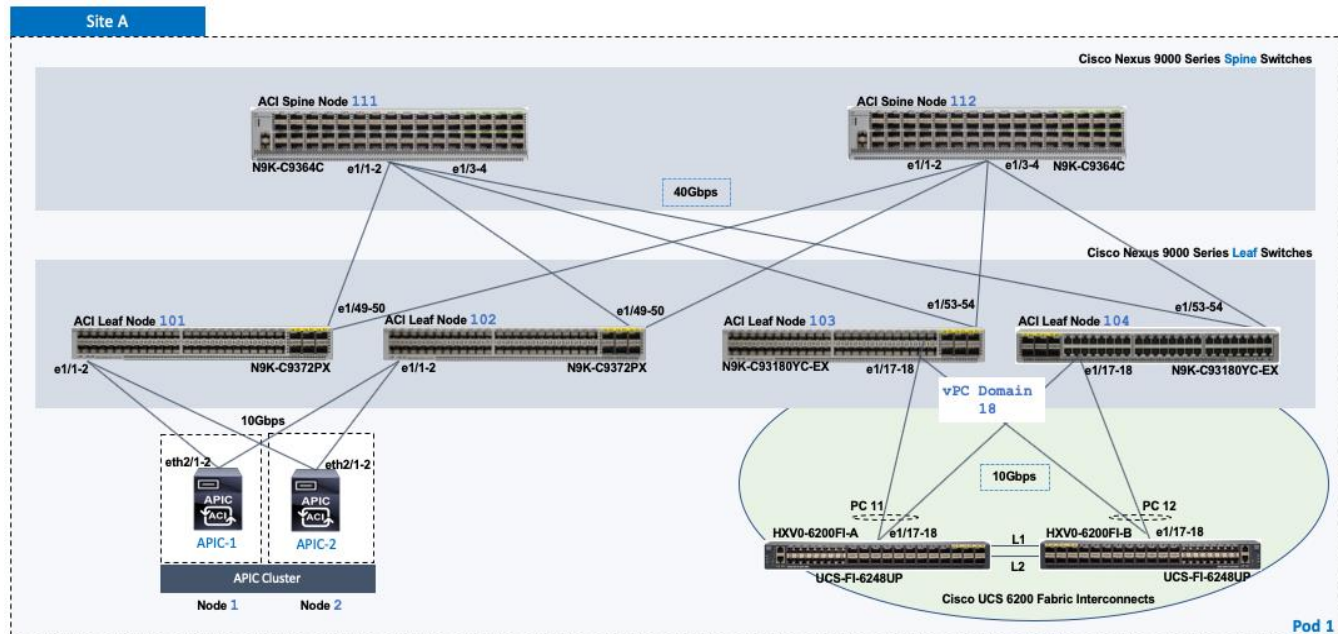
The ACI Fabric topology in Pod-2 to connect to the UCS domain for the HyperFlex stretched cluster in Pod-2 is shown in Figure 19.

Figure 19 ACI Fabric – Connectivity to Cisco UCS Domain for HyperFlex Stretched Cluster in Pod-2



The ACI Fabric topology to connect to the UCS domain for the HyperFlex standard cluster in Pod-1 is shown in Figure 20.

Figure 20 ACI Fabric – Connectivity to Cisco UCS Domain for HyperFlex Standard Cluster in Pod-1



Enable 40Gbps Connectivity to Cisco UCS Domain

In this design, the Cisco UCS domain consisting of Cisco UCS 6300 Series Fabric Interconnects are connected to the Leaf switches using 40Gbps links. 10Gbps links can also be used if needed. The 40Gbps ports for the Nexus leaf switch model used in this design are configured as **Uplink** ports by default. To re-configure these ports as **Downlink** ports, follow these steps:



The changes in this section will require a reload of the Leaf switches.

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Inventory**.
3. From the left navigation pane, select the Pod and the first Leaf switch that connects to the UCS Domain (FI-A, FI-B).
4. In the right window pane, select the **Interface** tab.
5. Under **Mode**, select **Configuration** from the drop-down list.
6. Select the port that connects to the **first** Fabric Interconnect (FI-A).
7. From the menu above the ports, select **Downlink**.
8. In the **Configure Uplink/Downlink Interface** pop-up window, click Submit.
9. Repeat the above steps for the port that connects to the **second** Fabric Interconnect (FI-B).
10. In the **Configure Uplink/Downlink Interface** pop-up window, click **Submit and Reload Switch** to reload the switch so that the changes take effect.
11. Repeat steps 1-10 for the second Leaf switch that connects to the Cisco UCS domain (FI-A, FI-B).
12. Repeat steps 1-11 for Pod-2/Site-2 leaf switches that connect to UCS domain (FI-A, FI-B) in Pod-2.

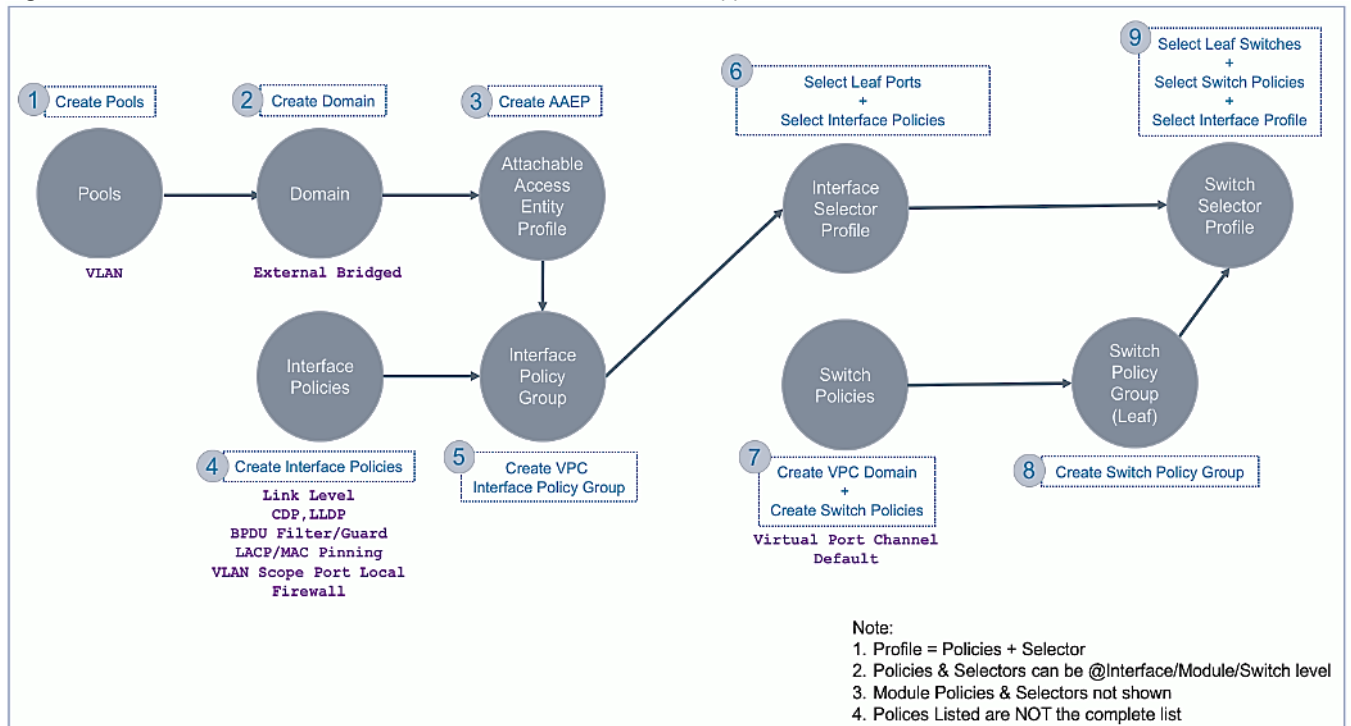
Enable Access Layer Configuration to Cisco UCS Domain

The ACI fabric uses Fabric Access Policies and Profiles for the access layer configuration. [Figure 10](#) shows the deployment workflow for configuring Fabric Access Policies on leaf switches. To create vPCs from the newly deployed ACI leaf switches to the UCS Fabric Interconnects where the HyperFlex cluster will be deployed, use the deployment workflow to complete the steps outlined below.

Deployment Workflow

The following workflow will configure the access ports on a leaf switch pair and create the vPCs to the UCS Domain (FI-A, FI-B).

Figure 21 Fabric Access Policies – To Cisco UCS Domain and HyperFlex Cluster



Create VLAN Pool for Cisco UCS Domain

The VLAN Pool defines all the VLANs that will be used in the Cisco UCS domain. In the ACI Fabric, the VLAN pool is created and associated with the access layer connection to the UCS and Hyperflex domain. When traffic is received from the VLANs in the pool, ACI fabric will use the VLAN tag to map it to an EPG for further forwarding decisions for that traffic. A single UCS domain can support multiple UCS servers and HyperFlex clusters; the VLAN pool should include the VLANs for all servers reachable through the access ports to UCS fabric Interconnects being configured.

The VLANs used in this design are listed in [Table 41](#). The VLAN Names are part of the UCS domain and HyperFlex setup. They are not used in the ACI fabric but the corresponding VLANs are created in the ACI fabric. The VLANs listed only includes the minimal VLANs required for HyperFlex installation. Application or Tenant VLANs are not added at this point in the configuration.

Table 41 VLAN Pool – To Cisco UCS Domain and HyperFlex Cluster

VPC to UCS 6300 FIs	VLAN Pool Name	Allocation Mode	VLAN	VLAN Name	Description
	HXV-UCS_VLANS	Static	118	hxxv-inband-mgmt	Management (InBand) Network for ESXi Hypervisor and Storage Controller VM (SCVM) on HX nodes
			3018	hxxv-vmotion	HX vMotion Network
			3218	hxxv1-storage-data	HX Storage Data Network – a unique VLAN should be used for each HX cluster deployed



The HyperFlex standard cluster uses the same VLAN pool but adds a unique storage-data vlan for the HyperFlex standard cluster. The management and vMotion VLANs are shared by the standard and stretched clusters.

To configure VLAN pools for the Cisco UCS domain where the Cisco HX Cluster is deployed, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Access Policies**.
3. From the left navigation pane, expand and select **Pools > VLAN**.
4. Right-click **VLAN** and select **Create VLAN Pool**.
5. In the Create VLAN Pool pop-up window, specify a **Name**. For Allocation Mode, select **Static Allocation**.

The screenshot displays the Cisco APIC GUI. The top navigation bar includes 'System', 'Tenants', 'Fabric', 'Virtual Networking', 'L4-L7 Services', 'Admin', 'Operations', and 'Apps'. The 'Fabric' tab is active, and the 'Access Policies' sub-tab is selected. The left-hand 'Policies' pane shows a tree structure with 'VLAN' selected under 'Pools'. The main area shows 'Pools - VLAN' with a table of existing pools:

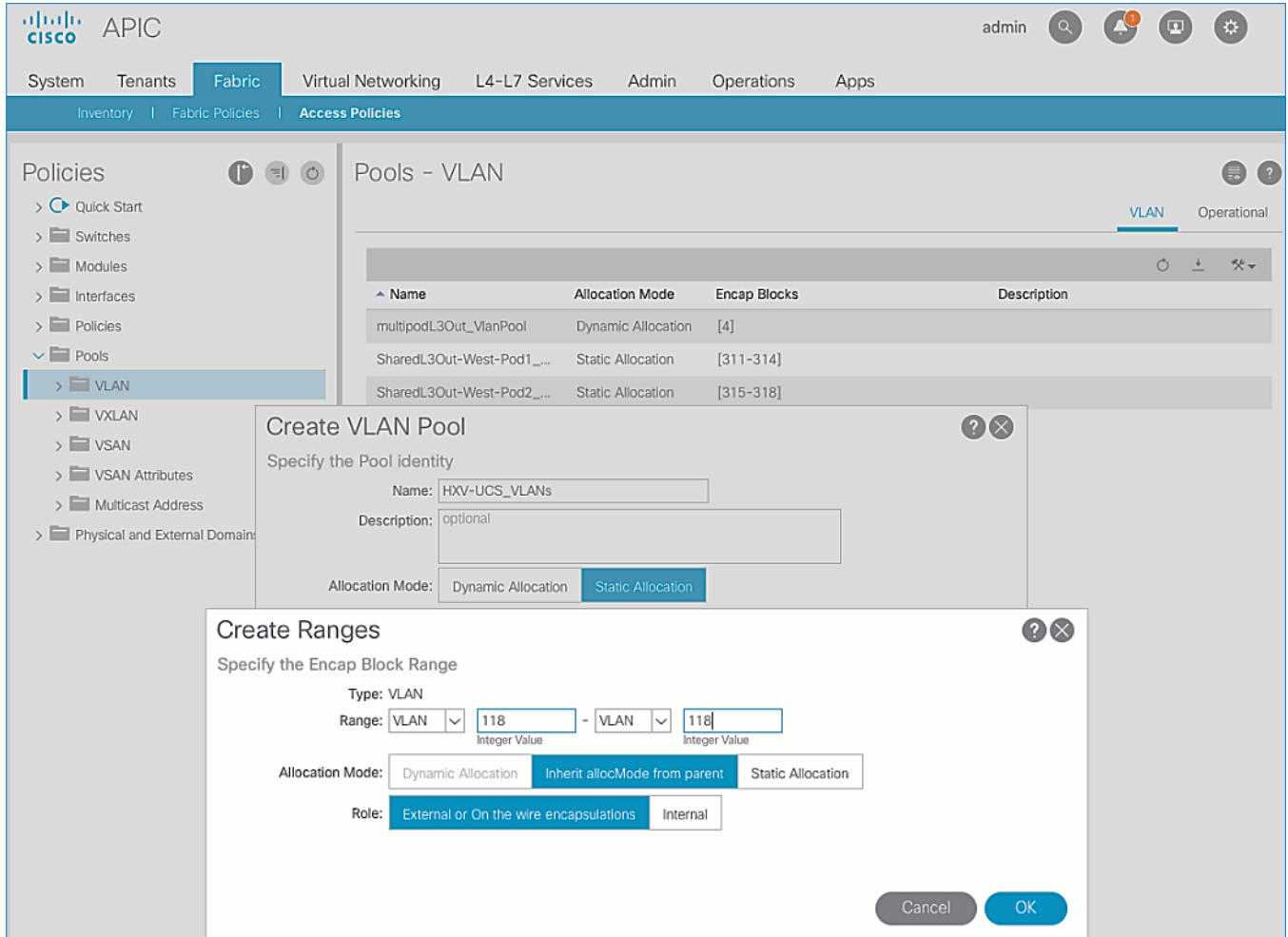
Name	Allocation Mode	Encap Blocks	Description
multipodL3Out_VlanPool	Dynamic Allocation	[4]	
SharedL3Out-West-Pod1_...	Static Allocation	[311-314]	
SharedL3Out-West-Pod2_...	Static Allocation	[315-318]	

In the foreground, the 'Create VLAN Pool' dialog box is open. It contains the following fields and options:

- Name:** HXV-UCS_VLANS
- Description:** optional
- Allocation Mode:** Dynamic Allocation (radio button), **Static Allocation** (radio button, selected)
- Encap Blocks:** A table with columns 'VLAN Range', 'Allocation Mode', and 'Role'. It includes a trash icon and a plus sign to add new blocks.

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

6. For **Encap Blocks**, use the **[+]** button on the right to add VLANs to the VLAN Pool. In the **Create Ranges** pop-up window, configure the VLANs that need to be trunked from the Cisco UCS FIs to the ACI Fabric. Leave the remaining parameters as is. Additional VLANs can be added later as needed.

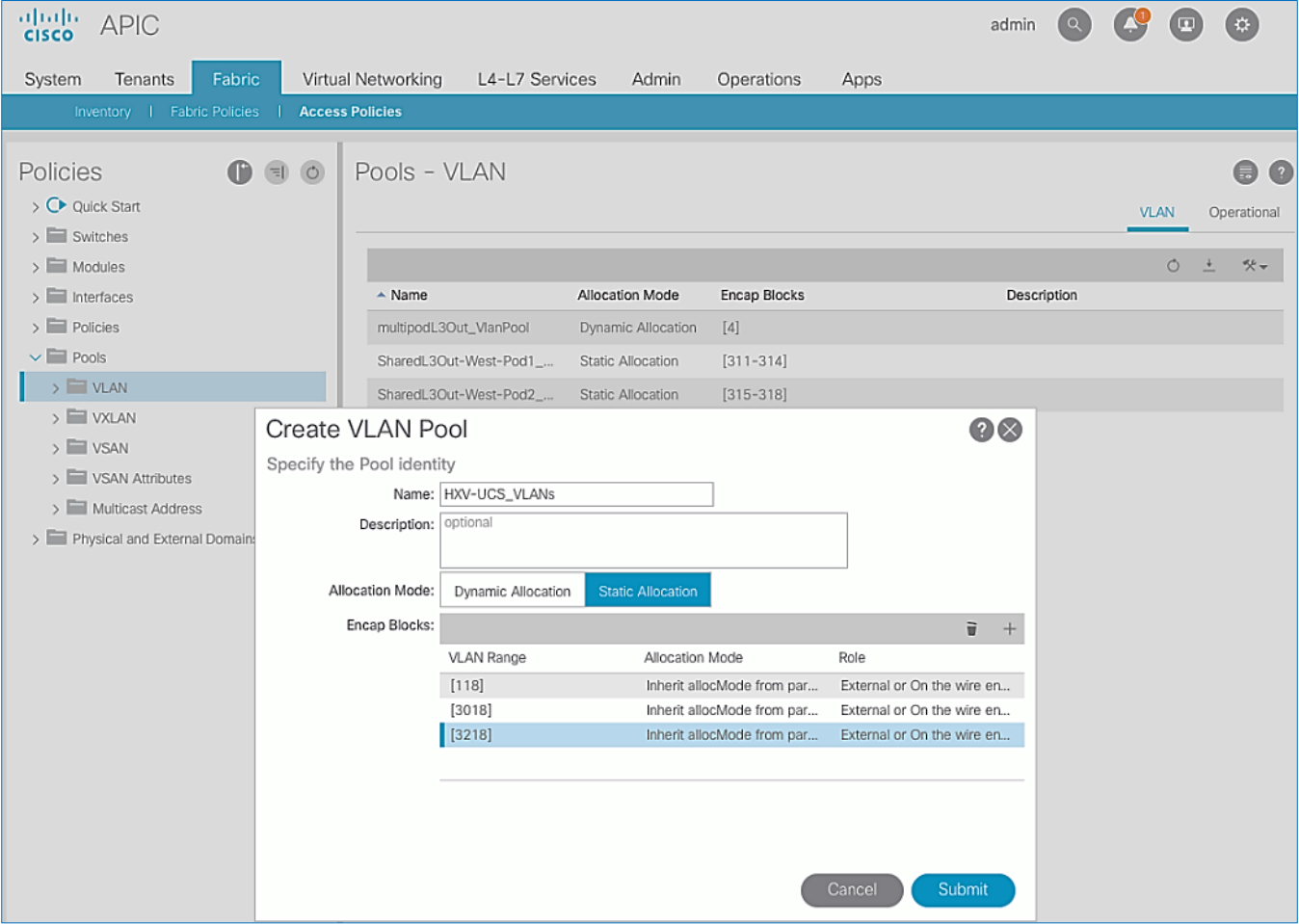


The screenshot shows the Cisco APIC interface with the 'Pools - VLAN' configuration page. The 'Create Ranges' pop-up window is open, showing the 'Specify the Encap Block Range' section. The 'Type' is set to 'VLAN', the 'Range' is 'VLAN 118 - VLAN 118', the 'Allocation Mode' is 'Inherit allocMode from parent', and the 'Role' is 'External or On the wire encapsulations'.

7. Repeat steps 1-6 for the remaining VLANs that need to be added to the VLAN Pool for the UCS Domain – see table in Setup Information above. The same VLANs need to be added to the corresponding Cisco UCS FIs in the UCS domain, on the uplinks from the FIs to the ACI fabric. For HyperFlex environment, the installation process will take care of adding this.



The HX storage data VLANs should be unique (recommended) to each HyperFlex cluster. However, they should still be trunked on the uplinks to the ACI Fabric to handle failure situations where different hosts are forwarding on different UCS fabrics (FI-A, FI-B).



- 8. Click **Submit** to complete.
- 9. Repeat steps 1-6 to add storage-data vlan (VLAN 3118) for the HyperFlex standard cluster.

Create Domain Type for Cisco UCS Domain

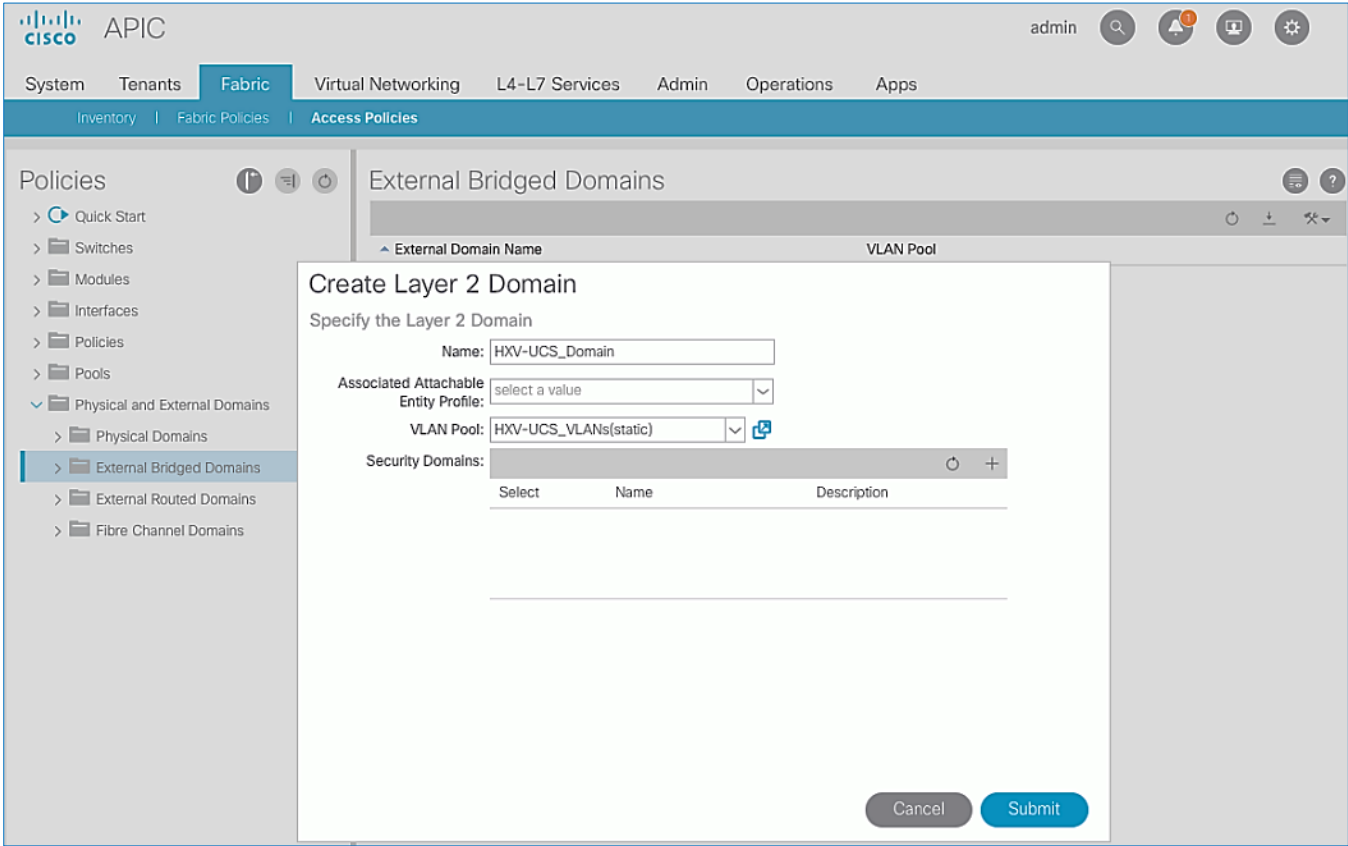
Table 42 External Domain – To Cisco UCS Domain and HyperFlex Cluster

VPC to UCS 6300 Fls	Domain Name	Domain Type	VLAN Pool Name	Connects To
	HXV-UCS_Domain	External Bridged Domain	HXV-UCS_VLANS	Cisco UCS Domain

To configure the domain type for the access layer connection to the Cisco UCS domain where the HyperFlex Cluster is deployed, follow these steps:

- 1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
- 2. From the top navigation menu, select **Fabric > Access Policies**.
- 3. From the left navigation pane, expand and select Physical and External Domains > External Bridged Domains.
- 4. Right-click External Bridged Domains and select **Create Layer 2 Domain**.

5. In the **Create Layer 2 Domain** pop-up window, specify a **Name** and select the previously created **VLAN Pool** from the drop-down list.



6. Click **Submit** to complete.

Create Attachable Access Entity Profile for Cisco UCS Domain

Table 43 Attachable Access Entity Profile – To Cisco UCS Domain and HyperFlex Cluster

vPC to UCS 6300 Fls	AAEP Name	Domain Name	VLAN Pool Name	Connects To
	HXV-UCS_AAEP	HXV-UCS_Domain	HXV-UCS_VLANs	Cisco UCS Domain

To create an Attachable Access Entity Profile (AAEP) for the access layer connection to the Cisco UCS domain where the HyperFlex Cluster is deployed, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Access Policies**.
3. From the left navigation pane, expand and select Policies > Global > Attachable Access Entity Profile.
4. Right-click Attachable Access Entity Profile and select Create Attachable Access Entity Profile.
5. In the Create Attachable Access Entity Profile pop-up window, specify a Name.
6. For the **Domains**, click the [+] on the right-side of the window to add a domain. For the **Domain Profile**, select the previously created domain from the drop-down list.

Attachable Access Entity Profiles

Name	Infrastructure VLAN Enabled	Policy Groups	Description
default	true		
multipodL3Out_EntityProfile	false	multipodL3Out_policyGroup	
SharedL3Out-West-Pod1_AAEP	false	SharedL3Out-West-Pod1_PG	
SharedL3Out-West-Pod2_AAEP	false	SharedL3Out-West-Pod2_PG	

Create Attachable Access Entity Profile

STEP 1 > Profile

Specify the name, domains and infrastructure encaps

Name:

Description:

Enable Infrastructure VLAN: ☐

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

-
-
-
-
-
-

EPG DEPLOYMENT (All Selected)

Application EPGs

-
-
-

Encapsulation

Primary Encap Mode

Previous Cancel Next

- Click **Update**. Click **Next**. Association to interfaces will be done in an upcoming step. Click **Finish**.

Create Interface Policies for the vPC Interfaces to Cisco UCS Domain

Interface policies are pre-configured Policies that can be applied to interfaces that connect to the UCS domain are part of the pre-configured Fabric Access Policies that was covered in a previous section. The pre-configured policies can be used for any access layer connections by grouping the policies into a policy group and applying it to the relevant interfaces. Proceed to next section to create a policy group for the UCS domain.

Table 44 Interface Policies – To Cisco UCS Domain for HyperFlex Stretched Cluster

vPC to UCS 6300 FIs	Interface Policy Name	Description
	40Gbps-Link	Configures link for 40Gbps
	CDP-Enabled	Enables CDP
	LLDP-Enabled	Enables LLDP
	BPDU-FG-Enabled	Enables BPDU Guard
	VLAN-Scope-Local	Configures VLAN Scope to be Local
	LACP-Active	Enables LACP

Table 45 Interface Policies – To Cisco UCS Domain for HyperFlex Standard Cluster

vPC to UCS 6300 FIs	Interface Policy Name	Description
	10Gbps-Link	Configures link for 10Gbps
	CDP-Enabled	Enables CDP
	LLDP-Enabled	Enables LLDP
	BPDU-FG-Enabled	Enables BPDU Guard
	VLAN-Scope-Local	Configures VLAN Scope to be Local
	LACP-Active	Enables LACP

Create Interface Policy Group for the vPC Interfaces to Cisco UCS Domain

Table 46 Interface Policy Group – To Cisco UCS Domain for HyperFlex Stretched Cluster

vPC to UCS 6300 FIs	Interface Policy Group Name	Interface Policy Name	Associated AAEP
	HXV-UCS-6300FI-A_IPG	40Gbps-Link	HXV-UCS_AAEP
		CDP-Enabled	
	HXV-UCS-6300FI-B_IPG	LLDP-Enabled	
		BPDU-FG-Enabled	
		VLAN-Scope-Local	
		LACP-Active	

Table 47 Interface Policy Group – To Cisco UCS Domain for HyperFlex Standard Cluster

vPC to UCS 6200 FIs	Interface Policy Group Name	Interface Policy Name	Associated AAEP
	HXV-UCS-6200FI-A_IPG	10Gbps-Link	HXV-UCS_AAEP
		CDP-Enabled	
	HXV-UCS-6200FI-B_IPG	LLDP-Enabled	
		BPDU-FG-Enabled	
		VLAN-Scope-Local	
		LACP-Active	



Two Interface Policy Groups are necessary to create the separate vPCs to each FI in the UCS domain though interfaces to all Fabric Interconnects use the same policies in this design.

To create an interface policy group to apply to the access ports that connect to the Cisco UCS domain where the HyperFlex Cluster is deployed, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Access Policies**.
3. From the left navigation pane, expand and select **Interfaces > Leaf Interfaces > Policy Groups > VPC Interface**.
4. Right-click VPC Interface and select **Create VPC Interface Policy Group**.
5. In the **Create VPC Interface Policy Group** pop-up window, specify a **Name** and select the relevant pre-configured policies for the UCS domain from the drop-down list for each field. For the **Attached Entity Profile**, select the previously created AAEP to Cisco UCS Domain.

The screenshot displays the APIC GUI with the 'Create VPC Interface Policy Group' dialog box open. The dialog is titled 'Create VPC Interface Policy Group' and contains the following fields and values:

- Name:** HXV-UCS-6300FI-A.JPG
- Description:** optional
- Link Level Policy:** 40Gbps-Link
- CDP Policy:** CDP-Enabled
- MCP Policy:** select a value
- CoPP Policy:** select a value
- LLDP Policy:** LLDP-Enabled
- STP Interface Policy:** BPDU-FG-Enabled
- L2 Interface Policy:** VLAN-Scope-Local
- Port Security Policy:** select a value
- Egress Data Plane Policing Policy:** select a value
- Ingress Data Plane Policing Policy:** select a value
- Priority Flow Control Policy:** select a value
- Fibre Channel Interface Policy:** select a value
- Slow Drain Policy:** select a value
- MACsec Policy:** select a value
- Attached Entity Profile:** HXV-UCS_AAEP
- Port Channel Policy:** LACP-Active

The 'Connectivity Filters' section at the bottom shows two tabs: 'Switch IDs' and 'Interfaces'. The 'Submit' button is highlighted in blue.

6. Click **Submit** to complete.
7. Repeat steps 1-6 using setup information to create an interface policy group for the vPC to the second Fabric Interconnect in the pair.

8. Repeat steps 1-7 using setup information to create interface policy groups for the vPCs to UCS domain (FI-A, FI-B) for HyperFlex standard cluster.

Create Leaf Interface Profile for the vPC Interfaces to Cisco UCS Domain

Table 48 Interface Profile – To Cisco UCS Domain for HyperFlex Stretched Cluster

vPC to UCS 6300 FIs	Leaf Interface Profile Name	Access Port Selector	Interface Policy Group
	HXV-UCS-6300FI_IPR	HXV-UCS_p1_49	HXV-UCS-6300FI-A_IPG
		HXV-UCS_p1_50	HXV-UCS-6300FI-B_IPG

Table 49 Interface Profile – To Cisco UCS Domain for HyperFlex Standard Cluster

vPC to 6200 FIs	Leaf Interface Profile Name	Access Port Selector	Interface Policy Group
	HXV-UCS-6200FI_IPR	HXV-UCS_p1_17	HXV-UCS-6200FI-A_IPG
		HXV-UCS_p1_18	HXV-UCS-6200FI-B_IPG



Two **Access Port Selectors** and **Interface Policy Groups** are necessary to create the separate vPCs to each Fabric Interconnect in the UCS domain though the interfaces use the same interface policies in this design.

To create an interface profile to configure the access ports that connect to the Cisco UCS domain where the HyperFlex Cluster is deployed, follow these steps:

1. Use a browser to navigate to the APIC GUI. Login using the **admin** account.
2. From the top navigation menu, select **Fabric > Access Policies**.
3. From the left navigation pane, expand and select **Interfaces > Leaf Interfaces > Profiles**.
4. Right-click Profiles and select **Create Leaf Interface Profile**.
5. In the **Create Leaf Interface Profile** pop-up window, specify a profile **Name** and for **Interface Selectors**, click the **[+]** to select access ports connecting the Leaf switches to the UCS domain. In the **Create Access Port Selector** pop-up window, specify a selector **Name**, for the **Interface IDs**, select the access port going from the leaf switch to the first Fabric Interconnect. For the **Interface Policy Group**, select the previously configured policy group from the drop-down list for the first Fabric Interconnect.

The screenshot shows the Cisco APIC interface with the 'Fabric' tab selected. The left sidebar shows the 'Policies' menu with 'Leaf Interfaces' > 'Profiles' selected. The main area displays 'Leaf Interfaces - Profiles' with a table of interface selectors. Two dialog boxes are open: 'Create Leaf Interface Profile' and 'Create Access Port Selector'.

Create Leaf Interface Profile

Specify the profile identity

Name: HXV-UCS-6300FI_IPR

Description: optional

Interface Selectors:

Name	Type
SharedL3Out-West-Pod1_IPR	1/47-48
SharedL3Out-West-Pod2_IPR	1/47-48

Create Access Port Selector

Specify the selector identity

Name: HXV-UCS_p1_49

Description: optional

Interface IDs: 1/49

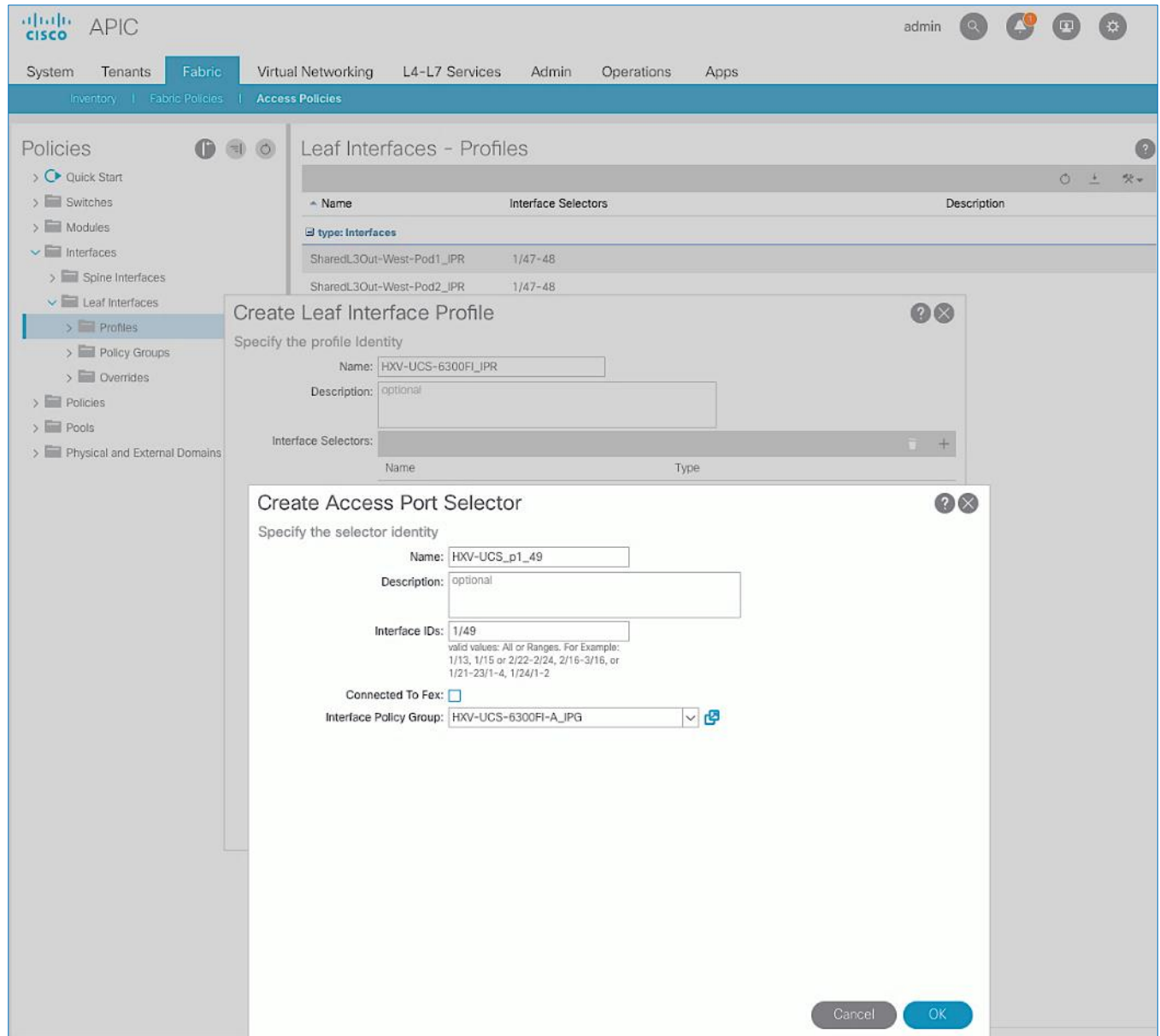
valid values: All or Ranges. For Example:
1/13, 1/15 or 2/22-2/24, 2/16-3/16, or
1/21-23/1-4, 1/24/1-2

Connected To Fex: ☐

Interface Policy Group: HXV-UCS-6300FI-A_IPG

Buttons: Cancel, OK

6. Click **OK**.
7. Repeat steps 1-6 to create a second **Access Port Selector** for the vPC to the second Fabric Interconnect in the Cisco UCS domain by clicking the **[+]** to add more **Interface Selectors** for the same Interface Profile.



8. Verify that all vPC interfaces to UCS have been added and are listed in the **Interface Selectors** section.
9. Click **Submit** to complete.
10. Repeat steps 1-9 using setup information to configure Interface profile for ports going to UCS domain for HyperFlex standard cluster.

Create Switch Policies for the vPC Interfaces to Cisco UCS Domain

Table 50 Switch Policies – vPC to Cisco UCS Domain for HyperFlex Stretched Cluster in Pod-1

vPC to UCS 6300 FIs	Pod 1			
	Switch Policy Name	VPC Explicit Protection Group	vPC Domain ID	Node ID
	Virtual Port Channel default	HXV-UCS-Leaf_103-104_VPC_ExPG	18	103, 104

Table 51 Switch Policies – vPC to Cisco UCS Domain for HyperFlex Stretched Cluster in Pod-2

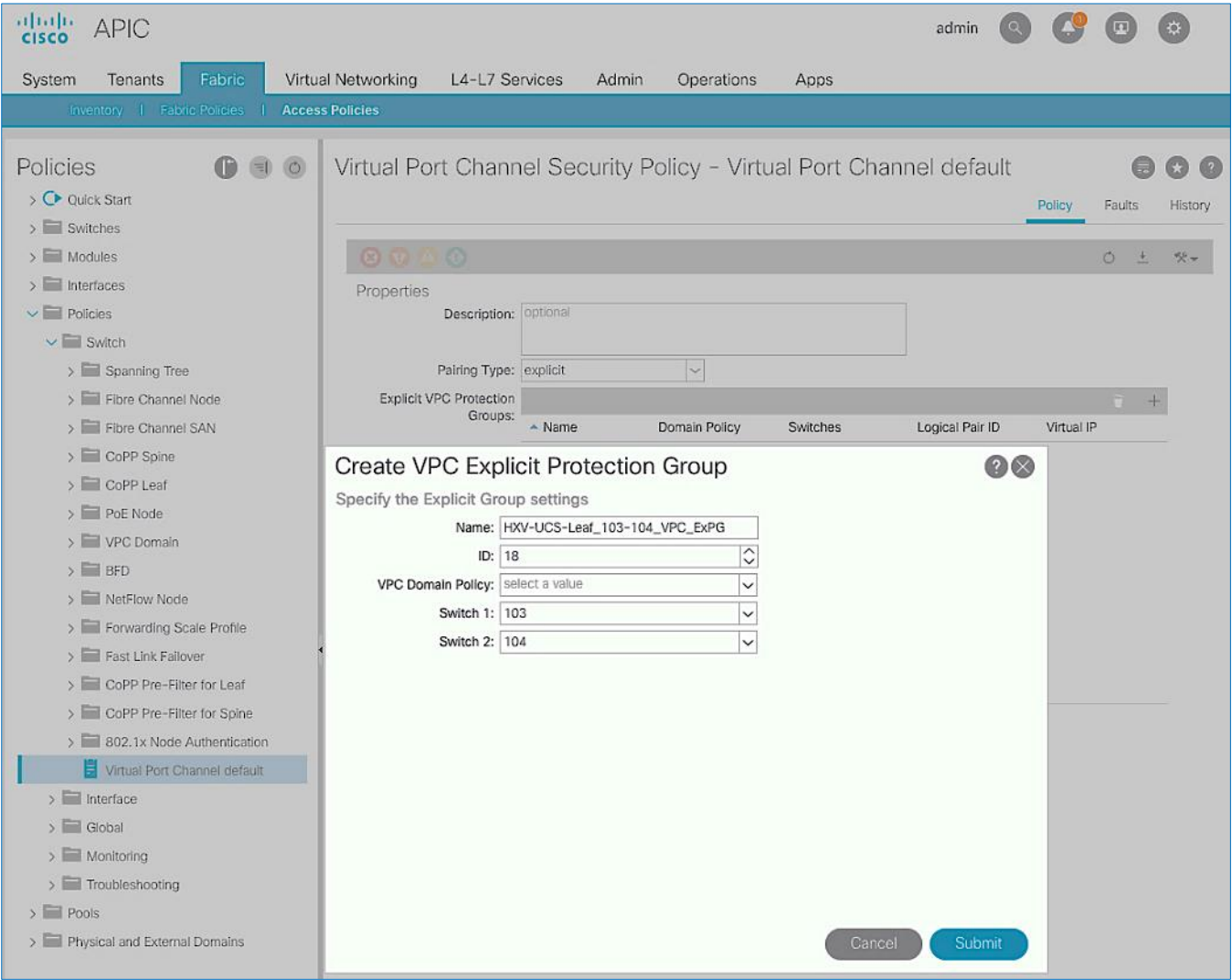
vPC to UCS 6300 FIs	Pod 2			
	Switch Policy Name	VPC Explicit Protection Group	vPC Domain ID	Node ID
	Virtual Port Channel default	HXV-UCS-Leaf_203-204_VPC_ExPG	18	203, 204

Table 52 Switch Policies – vPC to Cisco UCS Domain for HyperFlex Standard Cluster in Pod-1

vPC to UCS 6200 FIs	Pod 1			
	Switch Policy Name	VPC Explicit Protection Group	vPC Domain ID	Node ID
	Virtual Port Channel default	HXV-UCS-Leaf_103-104_VPC_ExPG	18	103, 104

To create leaf switch policies to apply to the vPC interfaces that connect to the Cisco UCS domain where the HyperFlex Cluster is deployed, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Access Policies**.
3. From the left navigation pane, expand and select **Policies > Switch > Virtual Port Channel default**.
4. Right-click Virtual Port Channel default and select Create VPC Explicit Protection Group.
5. In the **Create VPC Explicit Protection Group** pop-up window, specify a **Name** and for the **ID**, provide the vPC Domain ID for the Leaf pair. For **Switch 1** and **Switch 2**, select the Node IDs of the leaf pair from the list.



- 6. Click **Submit** to complete.
- 7. Repeat steps 1-6 using setup information to create switch policies for Pod-2/Site-2 Leaf switches to connect to UCS domain for HyperFlex stretched cluster.
- 8. Repeat steps 1-6 using setup information to create switch policies for Pod-1/Site-1 Leaf switches to connect to UCS domain for HyperFlex standard cluster.

Create Leaf Switch Profile

Table 53 Switch Profile – To Cisco UCS Domain for HyperFlex Stretched Cluster in Pod-1

Pod 1		
VPC to UCS 6300 FIs	Leaf Profile Name	Leaf Interface Profile
	Leaf Selectors	
	HXV-UCS-Leaf_103-104_IPR	HXV-UCS-6300FI_IPR

Table 54 Switch Profile – To Cisco UCS Domain for HyperFlex Stretched Cluster in Pod-2

VPC to UCS 6300 FIs	Pod 2		
	Leaf Profile Name	Leaf Selectors	Leaf Interface Profile
	HXV-UCS-Leaf_203-204_IPR	HXV-UCS-Leaf_203-204	HXV-UCS-6300FI_IPR

Table 55 Switch Profile – To Cisco UCS Domain for HyperFlex Standard Cluster in Pod-1

VPC to UCS 6200 FIs	Pod 1		
	Leaf Profile Name	Leaf Selectors	Leaf Interface Profile
	HXV-UCS-Leaf_103-104_IPR	HXV-UCS-Leaf_103-104	HXV-UCS-6200FI_IPR

To create a switch profile to configure the leaf switches that connect to the Cisco UCS domain where the HyperFlex Cluster is deployed, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Access Policies**.
3. From the left navigation pane, expand and select **Switches > Leaf Switches > Profiles**.
4. Right-click Profiles and select Create Leaf Profile.
5. In the **Create Leaf Profile** pop-up window, specify a profile **Name**. For **Leaf Selectors**, click the **[+]** on the right to select the leaf switches to apply the policies to. For **Name**, specify a name for the Leaf Switch Pair. For **Blocks**, select Node IDs for the Leaf Switch pair that connects to the Cisco UCS Domain.

APIC admin

System Tenants **Fabric** Virtual Networking L4-L7 Services Admin Operations Apps

Inventory Fabric Policies Access Policies

Policies

- Quick Start
- Switches
 - Leaf Switches
 - Profiles**
 - Policy Groups
 - Overrides
 - Spine Switches
 - Modules
 - Interfaces
 - Policies
 - Pools
 - Physical and External Domains

Leaf Switches - Profiles

Name	Leaf Selectors (Switch Policy Group)	Interface Selectors	Module Selectors	Description
SharedL3Out-West-Pod1-Leaf_PR	101-102	SharedL3Out-West-Pod1_IPR		
SharedL3Out-West-Pod2-Leaf_PR	201-202	SharedL3Out-West-Pod2_IPR		

Create Leaf Profile

STEP 1 > Profile

1. Profile 2. Associations

Specify the profile Identity

Name: HXV-UCS-Leaf_103-104_IPR

Description: optional

Leaf Selectors:

Name	Blocks	Policy Group
HXV-UCS-Leaf_103-104	103-104	select an option

Select Pod: All Pods

ID	Name
101	AA11-...
102	AA11-...
<input checked="" type="checkbox"/> 103	AA07-...
<input checked="" type="checkbox"/> 104	AA07-...
201	BB06-9...
202	BB06-9...
203	BB06-9...
204	BB06-9...

Previous Cancel Next

- 6. Click **Update**. Click **Next**.
- 7. In the **STEP 2 > Associations** window, for **Interface Selector Profiles**, select the previously created profile from the list.

The screenshot shows the Cisco APIC interface with the 'Fabric' tab selected. The left sidebar shows the 'Policies' menu with 'Leaf Switches' > 'Profiles' selected. The main area displays the 'Leaf Switches - Profiles' table:

Name	Leaf Selectors (Switch Policy Group)	Interface Selectors	Module Selectors	Description
SharedL3Out-West-Pod1-Leaf_PR	101-102	SharedL3Out-West-Pod1_IPR		
SharedL3Out-West-Pod2-Leaf_PR	201-202	SharedL3Out-West-Pod2_IPR		

Overlaid on this is the 'Create Leaf Profile' dialog box, currently on 'STEP 2 > Associations'. It has two tabs: '1. Profile' and '2. Associations' (active). The instruction says 'Select the interface/module selector profiles to associate'. Under 'Interface Selector Profiles', there is a table:

Select	Name	Description
<input checked="" type="checkbox"/>	HXV-UCS-6300FI_IPR	
<input type="checkbox"/>	SharedL3Out-West-P...	
<input type="checkbox"/>	SharedL3Out-West-P...	

Under 'Module Selector Profiles', there is an empty table with headers 'Select', 'Name', and 'Description'. At the bottom of the dialog are 'Previous', 'Cancel', and 'Finish' buttons.

8. Click **Finish** to complete.
9. Repeat steps 1-8 using setup information to create a switch profile for Pod-2/Site-2 Leaf switches to connect to UCS domain for HyperFlex stretched cluster.
10. Repeat steps 1-8 using setup information to create a switch profile for Pod-1/Site-1 Leaf switches to connect to UCS domain for HyperFlex standard cluster.

Solution Deployment – Setup Cisco UCS Domains

This section covers the setup of a **new** Cisco UCS domain for connecting HyperFlex clusters. In this design, multiple UCS domains are used, two for the HyperFlex stretched cluster (for Applications) and one for the HyperFlex standard cluster (for Management). The same procedures are used for bringing up all three UCS domains in this design. This section also provides detailed procedures for connecting each UCS domain to Cisco Intersight.



Repeat the procedures in this section for each UCS domain in the solution, using the corresponding setup information for that UCS domain.

Setup Information

This section provides the setup information for deploying the three UCS domains in this solution.

Table 56 UCS Domain Setup Information

Pod 1					
UCS 6300 FIs	System Name	Hostname	Management IP	Gateway	Other
	HXV1-6300-FI	HXV1-6300FI-A	192.168.167.205/24	192.168.167.254	Cluster IP: 192.168.167.204
		HXV1-6300FI-B	192.168.167.206/24		DNS Server: 10.99.167.244
					Domain Name: hxv.com
Pod 2					
UCS 6300 FIs	System Name	Hostname	Management IP	Gateway	Other
	HXV2-6300-FI	HXV2-6300FI-A	192.168.167.208/24	192.168.167.254	Cluster IP: 192.168.167.207
		HXV2-6300FI-B	192.168.167.209/24		DNS Server: 10.99.167.244
					Domain Name: hxv.com
Pod 1					
UCS 6200 FIs	System Name	Hostname	Management IP	Gateway	Other
	HXV0-6200-FI	HXV0-6200FI-A	192.168.167.202/24	192.168.167.254	Cluster IP: 192.168.167.201
		HXV0-6200FI-B	192.168.167.203/24		DNS Server: 10.99.167.244
					Domain Name: hxv.com

Bring Up UCS Domain with Fabric Interconnects

This section explains the setup of a new Cisco Unified Computing System (Cisco UCS) domain for use in a HyperFlex environment. The process does an initial setup of a new pair of Cisco UCS Fabric Interconnects that will be used to connect and deploy HyperFlex systems. Use the setup information to deploy the UCS domain.

Cisco UCS Fabric Interconnect A (FI-A)

To start the configuration of the FI-A, connect to the console of the fabric interconnect and step through the Basic System Configuration Dialogue:

```
---- Basic System Configuration Dialog ----
This setup utility will guide you through the basic configuration of
the system. Only minimal configuration including IP connectivity to
the Fabric interconnect and its clustering mode is performed through these steps.
```


Type Ctrl-C at any time to abort configuration and reboot system.
 To back track or make modifications to already entered values,
 complete input till end of section and answer no when prompted
 to apply configuration.

Enter the configuration method. (console/gui) ? console
 Enter the setup mode; setup newly or restore from backup. (setup/restore) ? setup
 You have chosen to setup a new Fabric interconnect. Continue? (y/n): y

Enforce strong password? (y/n) [y]:
 Enter the password for "admin":
 Confirm the password for "admin":
 Is this Fabric interconnect part of a cluster(select 'no' for standalone)? (yes/no) [n]: yes
 Enter the switch fabric (A/B) []: A
 Enter the system name: HXV1-6300-FI
 Physical Switch Mgmt0 IP address : 192.168.167.205
 Physical Switch Mgmt0 IPv4 netmask : 255.255.255.0
 IPv4 address of the default gateway : 192.168.167.254
 Cluster IPv4 address : 192.168.167.204
 Configure the DNS Server IP address? (yes/no) [n]: yes
 DNS IP address : 10.99.167.244
 Configure the default domain name? (yes/no) [n]: yes
 Default domain name : hxv.com
 Join centralized management environment (UCS Central)? (yes/no) [n]:

Following configurations will be applied:

Switch Fabric=A
 System Name=HXV1-6300-FI
 Enforced Strong Password=yes
 Physical Switch Mgmt0 IP Address=192.168.167.205
 Physical Switch Mgmt0 IP Netmask=255.255.255.0
 Default Gateway=192.168.167.254
 Ipv6 value=0
 DNS Server=10.99.167.244
 Domain Name=hxv.com
 Cluster Enabled=yes
 Cluster IP Address=192.168.167.204

NOTE: Cluster IP will be configured only after both Fabric Interconnects are initialized. UCSM will be functional only after peer FI is configured in clustering mode.

Apply and save the configuration (select 'no' if you want to re-enter)? (yes/no): yes
 Applying configuration. Please wait.

Configuration file - Ok

Cisco UCS 6300 Series Fabric Interconnect
 HXV1-6300-FI-A login:

Cisco UCS Fabric Interconnect B (FI-B)

Continue the configuration of Fabric Interconnect B (FI-B) from the console.

Enter the configuration method. (console/gui) ? console
 Installer has detected the presence of a peer Fabric interconnect. This Fabric interconnect will be added
 to the cluster. Continue (y/n) ? y

Enter the admin password of the peer Fabric interconnect:
 Connecting to peer Fabric interconnect... done
 Retrieving config from peer Fabric interconnect... done
 Peer Fabric interconnect Mgmt0 IPv4 Address: 192.168.167.205
 Peer Fabric interconnect Mgmt0 IPv4 Netmask: 255.255.255.0
 Cluster IPv4 address : 192.168.167.204
 Peer FI is IPv4 Cluster enabled. Please Provide Local Fabric Interconnect Mgmt0 IPv4 Address

```
Physical Switch Mgmt0 IP address : 192.168.167.206
```

```
Apply and save the configuration (select 'no' if you want to re-enter)? (yes/no): yes
Applying configuration. Please wait.
```

```
Wed Jul 11 02:23:14 UTC 2018
Configuration file - Ok
```

```
Cisco UCS 6300 Series Fabric Interconnect
HXV1-6300-FI-B login:
```

Initial Setup of Cisco UCS Domain

Log into Cisco UCS Manager

To log into the Cisco Unified Computing System (UCS) environment, follow these steps:

1. Use a browser to navigate to the Cluster IP of the Cisco UCS Fabric Interconnects.
2. Click the **Launch UCS Manager** to launch Cisco UCS Manager.
3. Click **Login** to log in to Cisco UCS Manager using the **admin** account.
4. If prompted to accept security certificates, accept as necessary.

Upgrade Cisco UCS Manager Software to Version 4.0(1c)

This document is based on Cisco UCS 4.0(1c) release of software for Cisco UCS infrastructure and HyperFlex nodes. To upgrade the Cisco UCS Manager software, the Cisco UCS Fabric Interconnect firmware and the server firmware bundles to version 4.0(1c) refer to the following Cisco UCS Manager Firmware Management Guide:

https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/ucs-manager/GUI-User-Guides/Firmware-Mgmt/4-o/b_UCSM_GUI_Firmware_Management_Guide_4-o.pdf.

Configure Cisco UCS Call Home and Anonymous Reporting (Optional)

It is highly recommended by Cisco to configure Call Home in Cisco UCS Manager. Configuring Call Home will accelerate resolution of support cases. To configure Call Home, follow these steps:

To configure Call Home, follow these steps:

1. Use a browser to navigate to the UCS Manager GUI. Log in using the **admin** account.
2. From the left navigation pane, select the **Admin** icon.
3. Select **All > Communication Management > Call Home**.
4. In the **General** Tab, change the **State** to **On**.
5. Use the other tabs to set **Call Home Policies** and other preferences, including **Anonymous Reporting** which enables data to be sent to Cisco for implementing enhancements and improvements in future releases and products.

Configure NTP

To synchronize the Cisco UCS environment to the NTP servers in the Nexus switches, follow these steps:

1. Use a browser to navigate to the UCS Manager GUI. Log in using the **admin** account.
2. From the left navigation menu, select the **Admin** icon.
3. From the left navigation pane, expand and select **All > Time Zone Management > Timezone**.
4. In the right window pane, for **Time Zone**, select the appropriate time zone from the drop-down list.

5. In the **NTP Servers** section, Click **[+] Add** to add NTP servers.
6. In the **Add NTP Server** pop-up window, specify the NTP server to use.
7. Click **OK** and **Save Changes** to accept.

Configure Uplink Ports on Each FI – To Nexus Leaf Switches in ACI Fabric

The Ethernet ports on Cisco UCS Fabric Interconnects can be configured in different modes depending on what is connected to them. The ports can be configured as Network Uplinks, Server ports, Appliance ports, and so on. By default, all ports are unconfigured.

To configure FI ports as network uplink ports to connect to the upstream network (in this case, ACI Fabric), follow these steps:

1. Use a browser to navigate to the Cisco UCS Manager GUI. Log in using the **admin** account.
2. From the left navigation menu, select the **Equipment** icon.
3. From the left navigation pane, expand and select **All > Equipment > Fabric Interconnects > Fabric Interconnect A > Fixed Module (or Expansion Module as appropriate) > Ethernet Ports**.
4. In the right window pane, select the uplink port and right-click to select **Enable** to enable the port and then re-select to select **Configure as Uplink Port**.
5. Click **Yes** and **OK** to confirm.
6. Repeat above steps for the next uplink port that connects to the ACI fabric from the same FI.
7. Navigate to **All > Equipment > Fabric Interconnects > Fabric Interconnect A > Fixed Module (or Expansion Module as appropriate) > Ethernet Ports**.
8. In the right window pane, select the uplink port and right-click to select **Enable** to enable the port and then re-select to select **Configure as Uplink Port**.
9. Click **Yes** and **OK** to confirm.
10. Repeat above steps for the next uplink port that connects to the ACI fabric from the same FI.
11. Verify that all ports are now **Network** ports with an **Overall Status** of **Up**.

Bundle Uplink Ports on each FI – To Nexus Leaf Switches in ACI Fabric

The uplink ports on each FI are bundled into a port-channel. The ports are connected to different Nexus Leaf switches in the ACI fabric. The leaf switches are part of a vPC domain, with a vPC to each FI – see **Solution Deployment - ACI Fabric** section of this document for the corresponding leaf switch configuration to this Fabric Interconnect pair.

To configure the uplink networks ports into a port-channel follow these steps on each FI:

1. Use a browser to navigate to the Cisco UCS Manager GUI. Log in using the **admin** account.
2. From the left navigation menu, select the **LAN** icon.
3. From the left navigation pane, expand and select **All > LAN > LAN Cloud > Fabric A**.
4. Right-click **Fabric A** and select **Create Port Channel** from the list.
5. In the **Create Port Channel** wizard, in the **Set Port Channel Name** section, for **ID**, specify a unique Port-Channel ID for this port-channel and for **Name**, specify a name for this port-channel. Click **Next**.
6. In the **Add Ports** section, select the uplink ports from the **Ports** table and use the **>>** to add them to the **Ports in the port channel** table to add them to port-channel. Click **Finish** and **OK** to complete.
7. Repeat steps 1-6 for **Fabric B** to create a port-channel to the Nexus Leaf switches, using the Fabric B uplink ports.

8. Verify the port channel is up and running on both Fabric Interconnects, with **Active** members.

Configuration of Server Ports – To HyperFlex Servers

The Ethernet ports on Cisco UCS Fabric Interconnects that connect to the rack-mount servers, or to the blade server chassis must be defined as server ports. When a server port comes online, a discovery process starts on the connected rack-mount server or chassis. During discovery, hardware inventories are collected, along with their current firmware revisions.

Rack-mount servers and blade chassis are automatically numbered in Cisco UCS Manager in the order which they are first discovered. For this reason, it is important to configure the server ports sequentially in the order you wish the physical servers and/or chassis to appear within Cisco UCS Manager.

Auto-Discovery of Server Ports (Option 1)

To enable servers to be discovered automatically when rack and blade servers are connected to server ports on the Cisco UCS Fabric Interconnects, follow these steps:

1. In Cisco UCS Manager, click the **Equipment** icon on left-navigation pane.
2. Navigate to **All > Equipment**. In the right window pane, click the tab for **Policies > Port Auto-Discovery Policy**.
3. Under **Properties**, set the Auto Configure Server Port to **Enabled**.
4. Click **Save Changes** and **OK** to complete.

Manual Configuration of Server Ports (Option 2)

To manually define the server ports and have control over the numbering of the servers, follow these steps:

1. In Cisco UCS Manager, from the left navigation menu, click the **Equipment** icon.
2. Navigate to **All > Equipment > Fabric Interconnects > Fabric Interconnect A > Fixed Module (or Expansion Module as appropriate) > Ethernet Ports**.
3. In the right-window pane, select the first port. Right-click and select **Configure as Server Port**.
4. Click **Yes** and **OK** to confirm.
5. Navigate to **All > Equipment > Fabric Interconnects > Fabric Interconnect A > Fixed Module (or Expansion Module as appropriate) > Ethernet Ports**.
6. In the right-window pane, select the matching port from Fabric Interconnect A. Right-click and select **Configure as Server Port**.
7. Click **Yes** and **OK** to confirm.
8. Repeat the above steps for the remaining ports that connect to servers.
9. Verify that all ports connected to chassis, Cisco FEX and rack servers are configured as Server Ports.

Modify Chassis Discovery Policy – For Blade Servers Only (Optional)

If the Cisco HyperFlex system uses Cisco UCS server blades in a Cisco UCS 5108 blade server chassis as compute-only nodes in an extended HyperFlex cluster design, then chassis discovery policy must be configured. The Chassis Discovery policy defines the number of links between the Fabric Interconnect and the Cisco UCS Fabric Extenders on the blade server chassis. These links determine the uplink bandwidth from the chassis to FI and must be connected and active, before the chassis will be discovered. The Link Grouping Preference setting specifies if the links will operate independently, or if Cisco UCS Manager will automatically combine them into port-channels. The number of links and the port types available on the Fabric Extender and Fabric Interconnect models will determine the uplink bandwidth. Cisco best practices recommends using link grouping (port-channeling). For 10 GbE connections Cisco recommends 4 links per side, and for 40 GbE connections Cisco recommends 2 links per side.

To modify the chassis discovery policy when using a Cisco UCS B-series chassis with HyperFlex, follow these steps:

1. Use a browser to navigate to the UCS Manager GUI. Log in using the **admin** account.
2. From the left navigation menu, select the **Equipment** icon.
3. From the left navigation pane, select **All > Equipment**.
4. In the right window pane, click-on the **Policies** tab.
5. Under the **Global Policies** tab, set the **Chassis/FEX Discovery Policy** (for **Action**) to match the minimum number of uplink ports that are cabled between the fabric extenders on the chassis and the fabric interconnects.
6. Set the Link Grouping Preference to Port Channel.
7. Click **Save Changes** and **OK** to complete.

Enable Cisco Intersight Cloud-Based Management

Cisco Intersight can be used to centrally manage all UCS domains and servers regardless of their physical location. Cisco Intersight can also be used to install a new HyperFlex cluster connected to Fabric Interconnects in a Cisco UCS domain. However, Cisco Intersight currently does not support the install of HyperFlex stretched clusters. Therefore, in this design, all Cisco UCS domains and HyperFlex systems are managed from Cisco Intersight but only the management HyperFlex cluster is installed using Cisco Intersight.

In this section, you will connect a Cisco UCS domain to Cisco Intersight to enable cloud-based management of the environment. This procedure is followed for all Cisco UCS domains in the design. The installation of a standard HyperFlex cluster using Cisco Intersight is covered in the next section.

Prerequisites

The prerequisites for setting up access to Cisco Intersight are as follows.

- An account on cisco.com.
- A valid Cisco Intersight account. This can be created by navigating to <https://intersight.com> and following the instructions for creating an account. The account creation requires at least one device to be registered in Intersight and requires Device ID and Claim ID information from the device. See Collecting Information From Cisco UCS Domain for an example of how to get Device ID and Claim ID from Cisco UCS Fabric Interconnect devices.
- Valid License on Cisco Intersight – see Cisco Intersight Licensing section below for more information.
- Cisco UCS Fabric Interconnects must have access to Cisco Intersight. In this design, the reachability is through an out-of-band network in the existing infrastructure, and not through the Cisco ACI Multi-Pod fabric.
- Cisco UCS Fabric Interconnects must be able to do a DNS lookup to access Cisco Intersight.
- Device Connectors on Fabric Interconnects must be able to resolve svc.ucs-connect.com.
- Allow outbound HTTPS connections (port 443) initiated from the Device Connectors on Fabric Interconnects to Cisco Intersight. HTTP Proxy is supported.

Cisco Intersight Licensing

Cisco Intersight is offered in two editions:

- Base license which is free to use, and offers a large variety of monitoring, inventory and reporting features.

- Essentials license, at an added cost but provides advanced monitoring, server policy and profile configuration, firmware management, virtual KVM features, and more. A 90-day trial of the Essentials license is available for use as an evaluation period.

New features and capabilities will be added to the different licensing tiers over time.

Setup Information

To setup access to Cisco Intersight, the following information must be collected from the Cisco UCS Domain. The deployment steps below will show how to collect this information.

- Device ID
- Claim Code

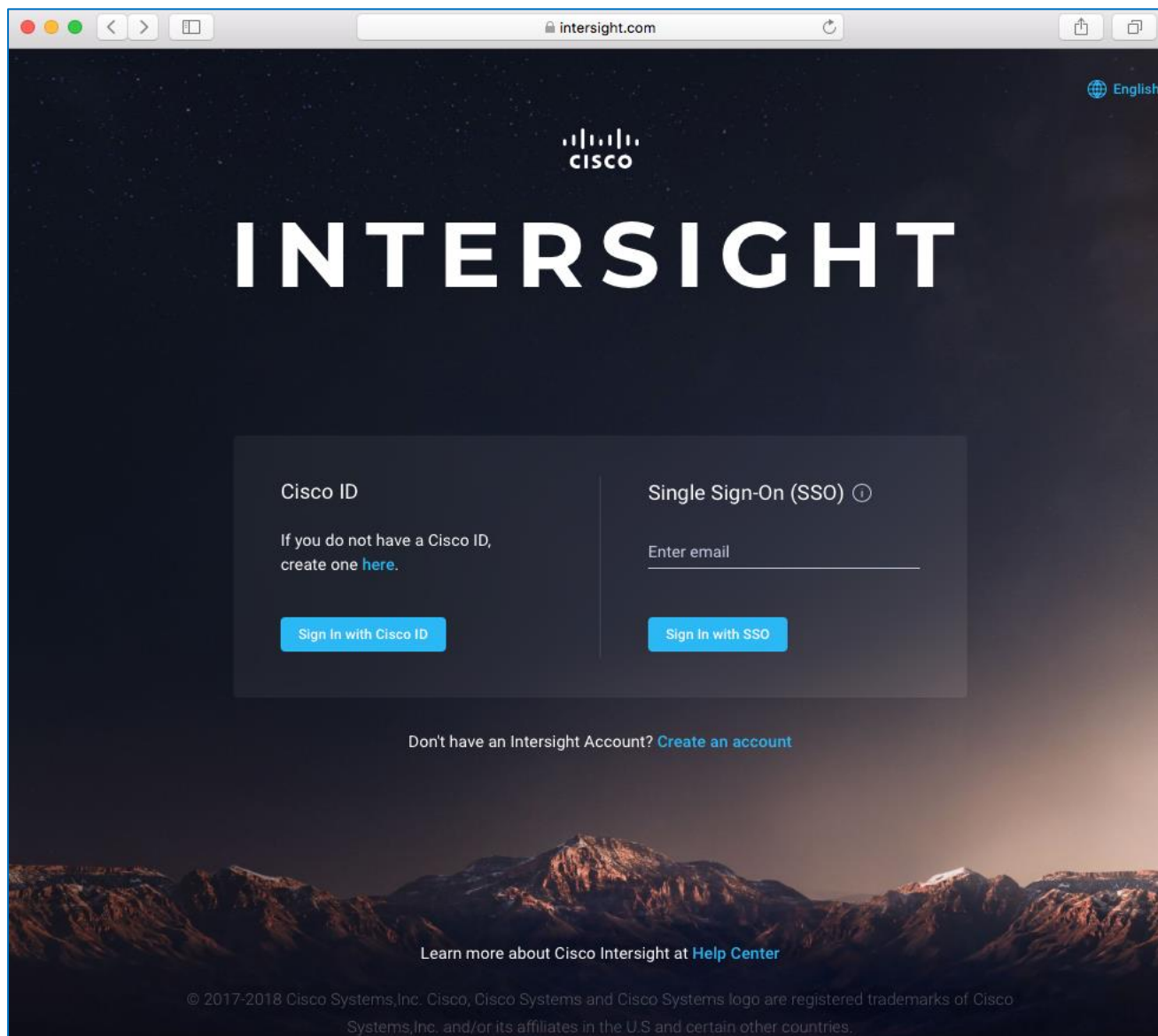
Deployment Steps

To setup access to Cisco Intersight from a Cisco UCS domain, follow these steps:

Connect to Cisco Intersight

To connect and access Cisco Intersight, follow these steps:

1. Use a web browser to navigate to Cisco Intersight at <https://intersight.com/>.

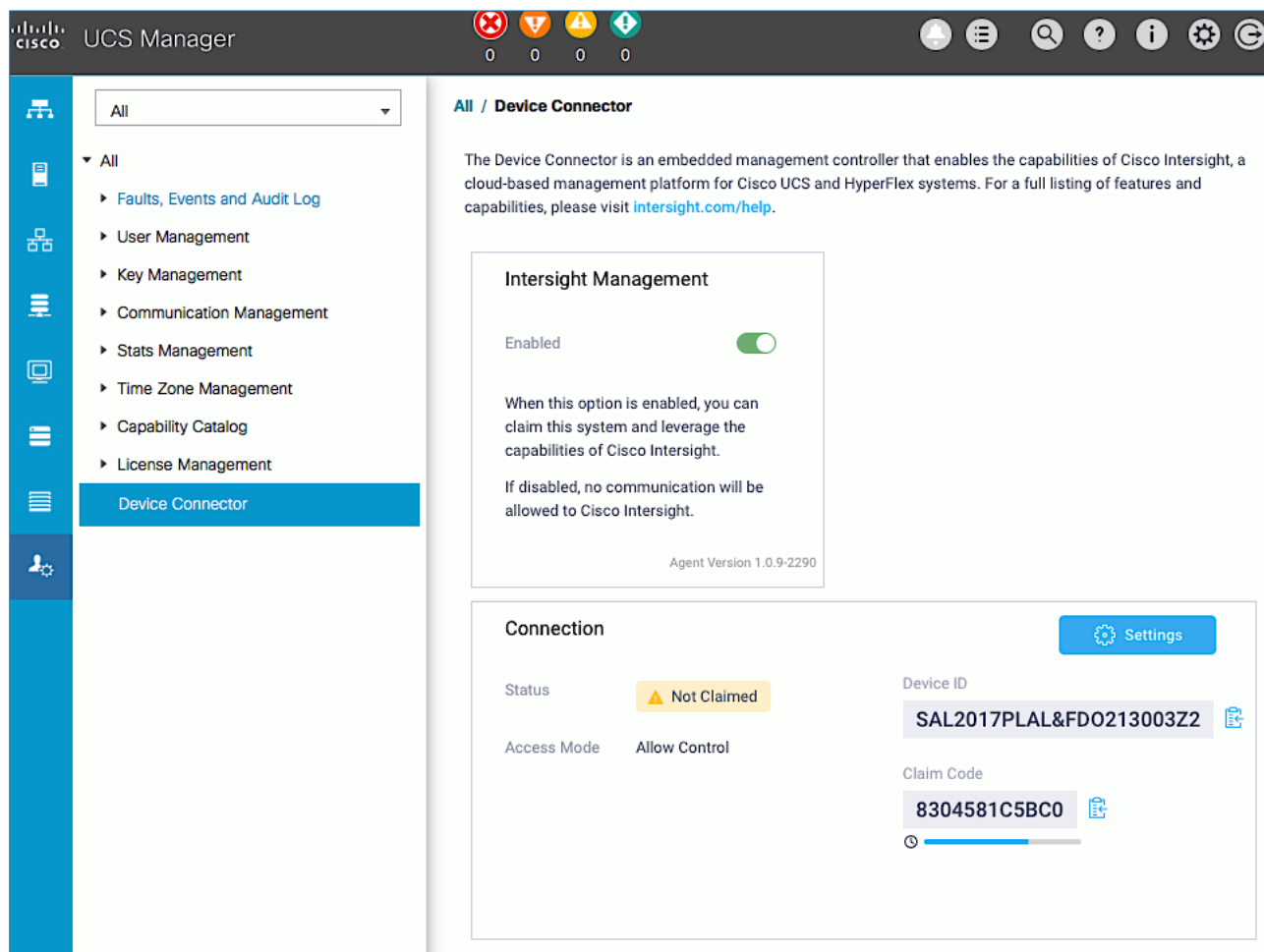


2. Log in with a valid cisco.com account or single sign-on using your corporate authentication.

Collect Information from UCS Domain

To collect information from Cisco UCS Fabric Interconnects to setup access to Cisco Intersight, follow these steps:

1. Use a web browser to navigate to the Cisco UCS Manager GUI. Log in using the **admin** account.
2. From the left navigation menu, select the **Admin** icon.
3. From the left navigation pane, select **All > Device Connector**.
4. In the right window pane, for **Intersight Management**, click **Enabled** to enable Intersight management.



5. From the **Connection** section, copy the **Device ID** and **Claim ID** information. This information will be required to add this device to Cisco Intersight.
6. (Optional) Click **Settings** to change **Access Mode** and to configure **HTTPS Proxy**.

Add Cisco UCS Domain to Cisco Intersight

To add Cisco UCS Fabric Interconnects to Cisco Intersight to manage the UCS domain, follow these steps:

1. From Cisco Intersight, in the left navigation menu, select **Devices**.
2. Click the **Claim a New Device** button in the top right-hand corner.
3. In the **Claim a New Device** pop-up window, paste the **Device ID** and **Claim Code** collected in the previous section.

Claim a New Device

To claim your device, you must have the Device ID and Claim Code.

Device ID *

SAL2017PLAL&FDO213003Z2

Claim Code *

0E04CAD37F3A

Cancel Claim

4. Click **Claim**.
5. On Cisco Intersight, the newly added UCS domain should now have a **Status** of **Connected**.
6. On Cisco UCS Manager, the **Device Connector** should now have a **Status** of **Claimed**.

Add Additional Cisco UCS Domains and Servers to Cisco Intersight

Repeat the procedures in the previous sub-sections to add more UCS domains and servers to Cisco Intersight. The UCS domains in this design that are managed by Cisco Intersight are shown below.

Devices

Claim a New Device

Search 6 items found 3 per page 1 of 2

Claimed By	Name	Status	Type	Device IP	Device ID
asharma@cisco...	HXV1-6300-FI	Connected	UCS Domain	192.168.167.204	SAL2017PLAL & ...
asharma@cisco...	HXV2-6300-FI	Connected	UCS Domain	192.168.167.207	FDO22062U3S & ...
asharma@cisco...	HXV0-6200-FI	Connected	UCS Domain	192.168.167.201	SSI191106FL & ...

1 of 2

Solution Deployment – Foundational Infrastructure for Cisco HyperFlex

In this section, you will create the foundational infrastructure within ACI that will provide the necessary connectivity to the UCS domains and HyperFlex systems in each Pod. This connectivity must be in place before the initial install and deployment of a HyperFlex cluster. The foundation infrastructure provides the following:

- In-Band Management Access to all ESXi hosts in the HX cluster. This is required not only to manage the HX nodes from VMware vCenter but also for the initial HyperFlex install. The same network will be used to access HyperFlex controller VMs deployed on the ESXi hosts.
- Storage Data Connectivity for storage data traffic in the HyperFlex **standard** cluster for **Management**. This includes ESXi hosts accessing datastores on Management storage cluster but also for storage traffic between nodes in the cluster. The storage data connectivity for the HyperFlex **stretched** cluster (**Applications** cluster) is discussed in the [Solution Deployment – HyperFlex Application Cluster](#) section.
- vMotion Network to enable vMotion across the ACI fabric.

Create Foundation Tenant and VRF

To create a Foundation Tenant and VRF for the HyperFlex foundational infrastructure within ACI, follow these steps using the setup information provided below. The same Tenant and VRF will be used by all HyperFlex clusters for foundational infrastructure connectivity across the ACI Multi-Pod fabric.

- **Tenant:** HXV-Foundation
 - **VRF:** HXV-Foundation_VRF
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 2. From the top navigation menu, select **Tenants > Add Tenant**.
 3. In the **Create Tenant** pop-up window, specify a **Name** (for example, HXV-Foundation).
 4. For the **VRF Name**, enter a name for the only VRF in this Tenant (for example, HXV-Foundation_VRF)
 5. Check the box for "Take me to this tenant when I click finish."

APIC admin

System **Tenants** Fabric Virtual Networking L4-L7 Services Admin Operations Apps

ALL TENANTS | Add Tenant | Tenant Search: name or descr | common | mgmt | infra

All Tenants

Name	Alias	Description	Bridge Domains	VRFs	EPGs	Health Score
comm						100
infra						100
mgmt						100

Create Tenant

Specify tenant details

Name:

Alias:

Description:

Tags:
enter tags separated by comma

GUID:

Provider	GUID	Account Name

Monitoring Policy:

Security Domains:

Name	Description

VRF Name:

☒ Take me to this tenant when I click finish

6. Click **Submit** to complete.

Configure ACI Fabric for HyperFlex In-Band Management

This section provides the ACI fabric configuration to support in-band management through the fabric.

Create Bridge Domain for In-Band Management

To create a Bridge Domain for In-Band Management of HyperFlex nodes, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
- **VRF:** HXV-Foundation_VRF
- **Bridge Domain:** HXV-IB-MGMT_BD

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.
3. From the left navigation pane, expand and select **Tenant HXV-Foundation > Networking > Bridge Domains**.
4. Right-click Bridge Domains and select **Create Bridge Domain**.
5. In the **Create Bridge Domain** wizard, for **Name**, specify a name for the bridge domain. For **VRF**, select the previously created VRF from the drop-down list. For **Forwarding**, select **Custom** from the drop-down list. For **L2 Unknown Unicast**, select **Flood** from the drop-down list. The checkbox for **ARP Flooding** should now show up and be enabled.

Create Bridge Domain

STEP 1 > Main

Specify Bridge Domain for the VRF

Name: HXV-IB-MGMT_BD

Alias:

Description: optional

Tags:

enter tags separated by comma

Type: fc regular

Advertise Host Routes: ☒

VRF: HXV-Foundation_VRF

Forwarding: Custom

L2 Unknown Unicast: Flood

L3 Unknown Multicast Flooding: Flood

Multi Destination Flooding: Flood in BD

ARP Flooding: ☒ Enabled

Clear Remote MAC Entries: ☐

Endpoint Retention Policy: select a value

This policy only applies to local L2 L3 and remote L3 entries

IGMP Snoop Policy: select a value

Previous Cancel Next

6. Click **Next**.
7. In the **L3 Configurations** section, for **EP Move Detection Mode**, select the checkbox to enable **GARP based detection** if needed. See [Review/Enable ACI Fabric Settings](#) section for more details on when to enable this feature.

Create Bridge Domain

STEP 2 > L3 Configurations

Specify Bridge Domain for the VRF

Unicast Routing: ☒ Enabled
 ARP Flooding: ☒ Enabled
 Config BD MAC Address: ☒
 MAC Address: 00:22:BD:F8:19:FF
 Virtual MAC Address: not-applicable
 Subnets:

Gateway Address	Scope	Primary IP Address	Subnet Control

Endpoint Dataplane Learning: ☒
 Limit IP Learning To Subnet: ☒
 EP Move Detection Mode: ☒ GARP based detection
 DHCP Labels:

Name	Scope	DHCP Option Policy

Associated L3 Outs:

L3 Out

Previous Cancel Next

8. Click **Next**. Skip the **Advanced/Troubleshooting** section. Click **Finish** to complete.

Create Application Profile for In-Band Management

To create an application profile for In-Band Management, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
 - **Application Profile:** HXV-IB-MGMT_AP
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 2. From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.
 3. From the left navigation pane, right-click **Tenant HXV-Foundation** and select **Create Application Profile**.
 4. In the **Create Application Profile** pop-up window, for **Name**, specify a name for the Application Profile.

The screenshot shows the Cisco APIC GUI with the 'Create Application Profile' dialog open for the 'Tenant HXV-Foundation'. The dialog is titled 'Create Application Profile' and has a subtitle 'Specify Tenant Application Profile'. It contains the following fields:

- Name:** HXV-IB-MGMT_AP
- Alias:** (empty)
- Description:** optional
- Tags:** (empty, with a dropdown arrow and a note 'enter tags separated by comma')
- Monitoring Policy:** select a value

Below the fields is a section for 'EPGs' with a table. The table has the following columns: Name, Alias, BD, Domain, Switching Mode, Static Path, Static Path VLAN, Provided Contract, and Consumed Contract. The table is currently empty.

At the bottom right of the dialog are two buttons: 'Cancel' and 'Submit'.

5. Click **Submit** to complete

Create EPG for In-Band Management and Associate with Bridge Domain

To create an EPG for In-Band Management, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
 - **Application Profile:** HXV-IB-MGMT_AP
 - **Bridge Domain:** HXV-IB-MGMT_BD
 - **EPG:** HXV-IB-MGMT_EPG
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 2. From the top navigation menu, select **Tenants** > HXV-Foundation. If you do not see this tenant in the top navigation menu, select **Tenants** > **ALL TENANTS** and double-click on HXV-Foundation.
 3. From the left navigation pane, select and expand **Tenant HXV-Foundation** > **Application Profiles** > HXV-IB-MGMT_AP.

4. Right-click HXV-IB-MGMT_AP and select **Create Application EPG**.
5. In the **Create Application EPG** pop-up window, for **Name**, specify a name for the EPG. For **Bridge Domain**, select the previously created Bridge Domain.

The screenshot shows the 'Create Application EPG' window in the Cisco APIC. The window is titled 'Create Application EPG' and has a '1. Identity' tab. The 'STEP 1 > Identity' section is active, showing the 'Specify the EPG Identity' form. The form includes the following fields and options:

- Name:** HXV-IB-MGMT_EPG
- Alias:** (empty)
- Description:** optional
- Tags:** (empty dropdown)
- Contract Exception Tag:** (empty)
- QoS class:** Unspecified
- Custom QoS:** select a value
- Data-Plane Policer:** select a value
- Intra EPG Isolation:** Enforced (selected), Unenforced
- Preferred Group Member:** Exclude (selected), Include
- Flood on Encapsulation:** Disabled (selected), Enabled
- Bridge Domain:** HXV-IB-MGMT_BD (selected)
- Monitoring Policy:** select a value
- FHS Trust Control Policy:** select a value
- Shutdown EPG:** ☐
- Associate to VM Domain Profiles:** ☐
- Statically Link with Leaves/Paths:** ☐
- EPG Contract Master:** (empty)

The 'Finish' button is highlighted in blue at the bottom right of the dialog.

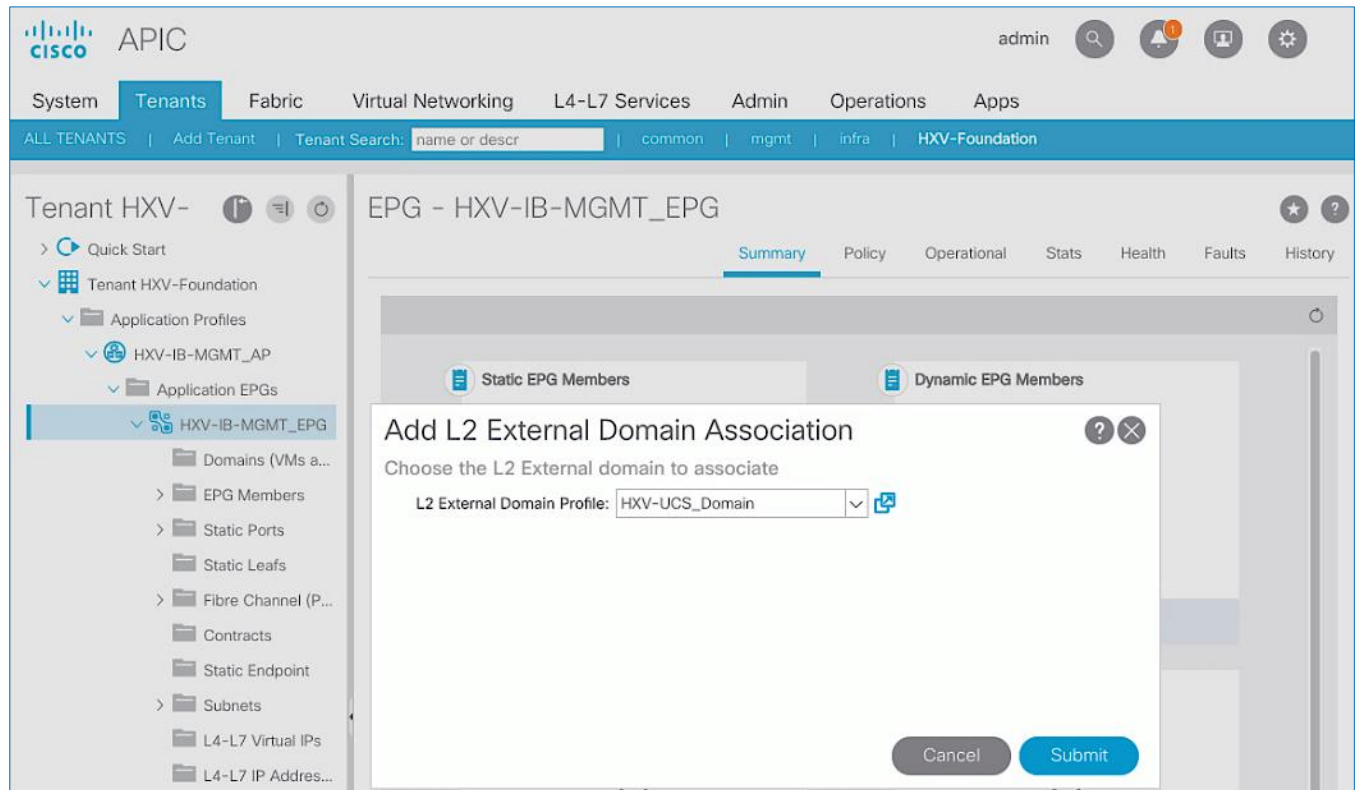
6. Click **Finish**.

Associate EPG with UCS Domain

To associate the In-Band Management EPG with UCS Domain, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
- **Application Profile:** HXV-IB-MGMT_AP
- **Bridge Domain:** HXV-IB-MGMT_BD
- **EPG:** HXV-IB-MGMT_EPG

- **Domain:** HXV-UCS_Domain
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 2. From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.
 3. From the left navigation pane, select and expand **Tenant HXV-Foundation > Application Profiles > HXV-IB-MGMT_AP > Application EPGs > HXV-IB-MGMT_EPG**.
 4. Right-click **HXV-IB-MGMT_EPG** and select **Add L2 External Domain Association**.
 5. In the **Add L2 External Domain Association** pop-up window, select the previously created UCS Domain from the list.



6. Click **Submit**.

Create Static Binding for EPG on vPC Interfaces to UCS Domain

To statically bind the In-Band Management EPG and VLANs to vPC interfaces going to the UCS Domain, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
 - **Application Profile:** HXV-IB-MGMT_AP
 - **EPG:** HXV-IB-MGMT_EPG
 - **Static Paths:** HXV-UCS_6200FI-A_IPG, HXV-UCS_6200FI-B_IPG
 - **VLAN:** 118
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.

- From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.
- From the left navigation pane, select and expand **Tenant HXV-Foundation > Application Profiles > HXV-IB-MGMT_AP > Application EPGs > HXV-IB-MGMT_EPG**.
- Right-click **HXV-IB-MGMT_EPG** and select **Deploy Static EPG on PC, VPC or Interface**.
- In the **Deploy Static EPG on PC, VPC or Interface** pop-up window, for **Path Type**, select **Virtual Port Channel**. For the **Path**, select the vPC to the first UCS Fabric Interconnect from the drop-down list. For the **Port Encap**, leave **VLAN** selected in the drop-down menu and in the box, specify the VLAN ID for the In-Band Management EPG. For the **Deployment Immediacy**, select **Immediate**.

The screenshot shows the Cisco APIC interface with the 'Deploy Static EPG On PC, VPC, Or Interface' dialog box open. The dialog is titled 'EPG - HXV-IB-MGMT_EPG' and has tabs for 'Summary', 'Policy', 'Operational', 'Stats', 'Health', 'Faults', and 'History'. The 'Summary' tab is selected. The dialog contains the following fields and options:

- Path Type:** Port, Direct Port Channel, Virtual Port Channel (selected)
- Path:** HXV-UCS-6200FI-A (selected)
- Port Encap (or Secondary VLAN for Micro-Seg):** VLAN (selected), 118 (Integer Value)
- Deployment Immediacy:** Immediate (selected), On Demand
- Primary VLAN for Micro-Seg:** VLAN (selected), Integer Value
- Mode:** Trunk (selected), Access (802.1P), Access (Untagged)
- IGMP Snoop Static Group:** A table with columns 'Group Address' and 'Source Address'.

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

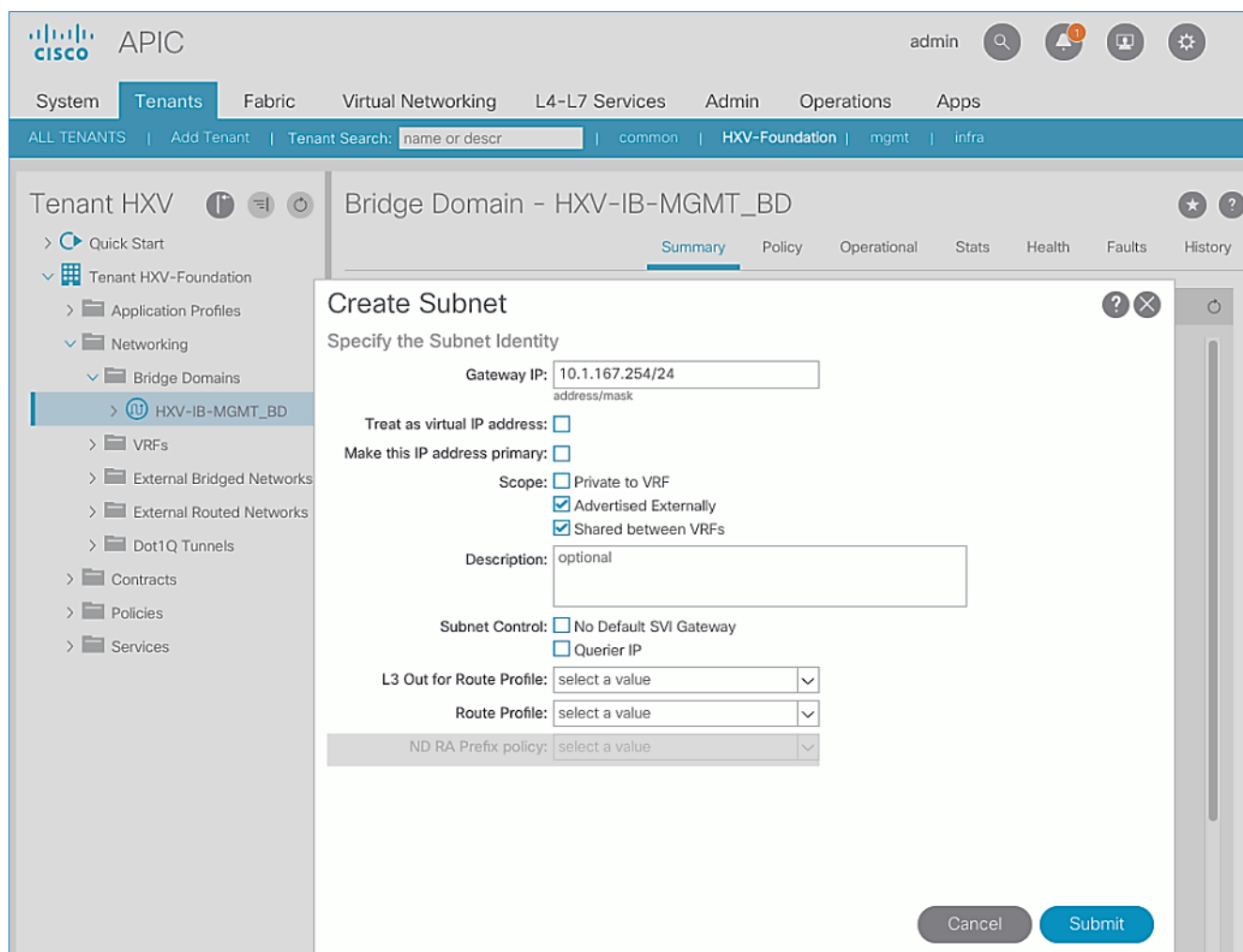
- Click **Submit**.
- Repeat steps 1-6 to bind the EPG to the vPC going to the second UCS Fabric Interconnect.

Configure Subnet Gateway for IB-MGMT EPG

To configure a gateway for In-Band Management, follow these steps using the setup information provided below:

- Tenant:** HXV-Foundation
- Bridge Domain:** HXV-IB-MGMT_BD

- **BD Subnet:** 10.1.167.254
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 2. From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.
 3. From the left navigation pane, select and expand **Tenant HXV-Foundation > Networking > Bridge Domains > HXV-IB-MGMT_BD**.
 4. Right-click **HXV-IB-MGMT_BD** and select **Create Subnet**.
 5. In the **Create Subnet** pop-up window, specify the **Default Gateway IP** and for **Scope**, select **Advertised Externally** and **Shared between VRFs**. Leave everything else as is.



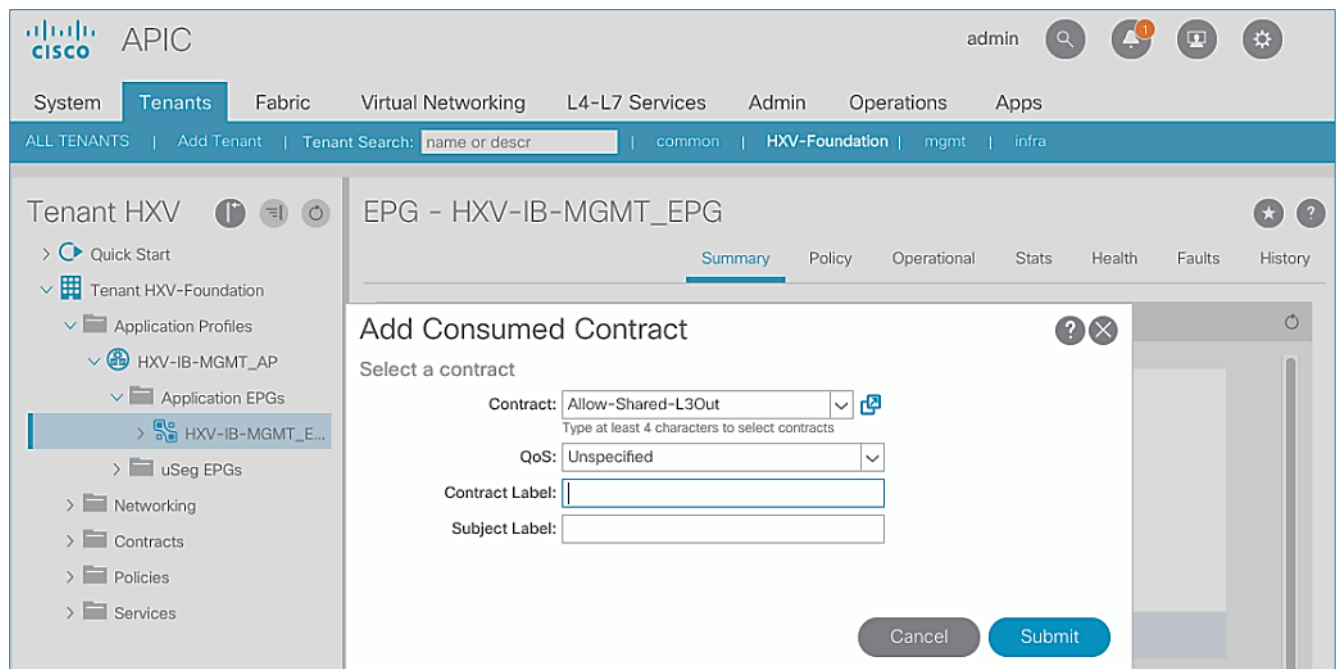
6. Click **Submit**.

Create Contract to Access Outside Networks using Shared L3Out

To enable a contract to access the network and services reachable through the **Shared L3Out** in the **common** Tenant, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
- **Application Profile:** HXV-IB-MGMT_AP

- **EPG:** HXV-IB-MGMT_EPG
 - **Contract:** Allow-Shared-L3Out
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 2. From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.
 3. From the left navigation pane, select and expand **Tenant HXV-Foundation > Application Profiles > HXV-IB-MGMT_AP > Application EPGs > HXV-IB-MGMT_EPG**.
 4. Right-click **HXV-IB-MGMT_EPG** and select **Add Consumed Contract**.
 5. In the **Add Consumed Contract** pop-up window, select the **Allow-Shared-L3Out** contract from the drop-down list.



6. Click **Submit**.

Configure ACI Fabric for HyperFlex Storage Data Traffic on HyperFlex Standard Cluster

This section covers the configuration of the Cisco ACI fabric to enable forwarding of HyperFlex storage data traffic between nodes in HyperFlex **standard** cluster for **Management**. Configuration to enable this traffic through the ACI fabric is required to support failure scenarios where traffic from one UCS Fabric needs to be forwarded to another when different hosts are forwarding on different fabrics (FI-A, FI-B). The storage data connectivity for the HyperFlex **stretched** cluster (**Applications** cluster) is discussed in the [Solution Deployment – HyperFlex Application Cluster](#) section.

Create Bridge Domain for HyperFlex Storage Data Traffic on HyperFlex Standard Cluster

To create a Bridge Domain for Storage data traffic, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
- **VRF:** HXV-Foundation_VRF
- **Bridge Domain:** HXV-Storage_BD

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.
3. From the left navigation pane, expand and select **Tenant HXV-Foundation > Networking > Bridge Domains**.
4. Right-click and select **Create Bridge Domain**.
5. In the **Create Bridge Domain** wizard, for **Name**, specify a name for the bridge domain. For **VRF**, select the previously created VRF from the drop-down list. For **Forwarding**, select **Custom** from the drop-down list. For **L2 Unknown Unicast**, select **Flood** from the drop-down list. The checkbox for **ARP Flooding** should now show up and be enabled.

The screenshot shows the 'Create Bridge Domain' wizard in the Cisco APIC GUI. The left sidebar shows the navigation tree with 'Tenant HXV-Foundation' selected. The main area is titled 'Specify Bridge Domain for the VRF'. The wizard has three steps: 1. Main, 2. L3 Configurations, and 3. Advanced/Troubleshooting. The 'Main' step is active. The form contains the following fields and values:

- Name: HXV-Storage_BD
- Alias: (empty)
- Description: optional
- Tags: (empty)
- Type: fc (selected), regular
- Advertise Host Routes: ☐
- VRF: HXV-Foundation_VRF
- Forwarding: Custom
- L2 Unknown Unicast: Flood
- L3 Unknown Multicast Flooding: Flood
- Multi Destination Flooding: Flood in BD
- ARP Flooding: ☒ Enabled
- Clear Remote MAC Entries: ☐
- Endpoint Retention Policy: select a value
- IGMP Snoop Policy: select a value

At the bottom right, there are three buttons: 'Previous', 'Cancel', and 'Next'.

6. Click **Next**.
7. In the **L3 Configurations** section, for **EP Move Detection Mode**, select the checkbox to enable **GARP based detection** if needed. See [Review/Enable ACI Fabric Settings](#) section for more details on when to enable this feature.
8. Click **Next**. Skip the **Advanced/Troubleshooting** section. Click **Finish** to complete.

Create Application Profile for HyperFlex Storage Data Traffic

To create an application profile for HyperFlex Storage data traffic, follow these steps using the setup information provided below. The same Application profile will be used for storage data by all HyperFlex clusters that connect to the ACI Multi-Pod fabric.

- **Tenant:** HXV-Foundation
- **Application Profile:** HXV-Storage_AP

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.

- From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.
- From the left navigation pane, right-click **Tenant HXV-Foundation** and select **Create Application Profile**.
- In the **Create Application Profile** pop-up window, specify a **Name** the Application Profile.

APIC

admin

System
Tenants
Fabric
Virtual Networking
L4-L7 Services
Admin
Operations
Apps

ALL TENANTS
Add Tenant
Tenant Search:
name or descr
common
HXV-Foundation
mgmt
infra

Tenant HXV
Application Profiles

> Quick Start

> Tenant HXV-Four

> Application Profiles

> Networking

> Contracts

> Policies

> Services

Create Application Profile

Specify Tenant Application Profile

Name: HXV-Storage_AP
Alias:
Description: optional
Tags:
enter tags separated by comma
Monitoring Policy: select a value

EPGs

Name	Alias	BD	Domain	Switching Mode	Static Path	Static Path VLAN	Provided Contract	Consumed Contract

Cancel
Submit

5. Click **Submit** to complete.

Create EPG for HyperFlex Storage on HyperFlex Standard Cluster

To create an EPG for HyperFlex storage data traffic, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
- Application Profile: HXV-Storage_AP
- **Bridge Domain:** HXV-Storage_BD
- **EPG:** HXV-CL0-StorData EPG

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.

- From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.
- From the left navigation pane, select and expand **Tenant HXV-Foundation > Application Profiles > HXV-Storage_AP**.
- Right-click **HXV-Storage_AP** and select **Create Application EPG**.
- In the **Create Application EPG** pop-up window, for **Name**, specify a name for the EPG. For **Bridge Domain**, select the previously created Bridge Domain.

APIC admin

System **Tenants** Fabric Virtual Networking L4-L7 Services Admin Operations Apps

ALL TENANTS | Add Tenant | Tenant Search: name or descr | common | **HXV-Foundation** | mgmt | infra

Create Application EPG

STEP 1 > Identity

Specify the EPG Identity

Name: HXV-CL0-StorData_EPG

Alias:

Description: optional

Tags: enter tags separated by comma

Contract Exception Tag:

QoS class: Unspecified

Custom QoS: select a value

Data-Plane Policer: select a value

Intra EPG Isolation: Enforced **Unenforced**

Preferred Group Member: **Exclude** Include

Flood on Encapsulation: **Disabled** Enabled

Bridge Domain: HXV-Storage_BD

Monitoring Policy: select a value

FHS Trust Control Policy: select a value

Shutdown EPG: ☐

Associate to VM Domain Profiles: ☐

Statically Link with Leaves/Paths: ☐

EPG Contract Master:

Application EPGs

Previous Cancel **Finish**

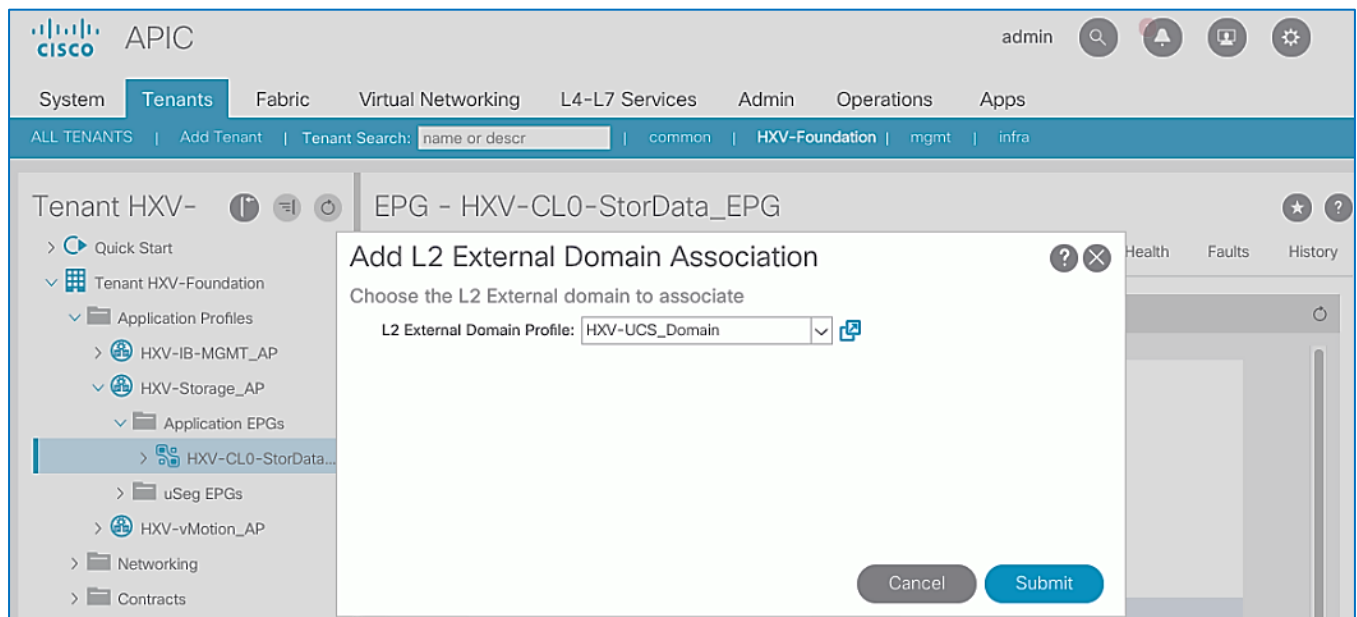
- Click **Finish**.

Associate EPG for Storage Data Traffic with UCS Domain

To associate the HyperFlex Storage EPG with the UCS Domain, follow these steps using the setup information provided below:

- Tenant:** HXV-Foundation

- Application Profile: HXV-Storage_AP
 - **Bridge Domain:** HXV-Storage_BD
 - **EPG:** HXV-CL0-StorData_EPG
 - **Domain:** HXV-UCS_Domain
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 2. From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.
 3. From the left navigation pane, select and expand **Tenant HXV-Foundation > Application Profiles > HXV-Storage_AP > Application EPGs > HXV-CL0-StorData_EPG**.
 4. Right-click HXV-CL0-StorData_EPG and select **Add L2 External Domain Association**.
 5. In the **Add L2 External Domain Association** pop-up window, select the previously created UCS Domain from the list.



6. Click **Submit**.

Create Static Binding for Storage Data Traffic to UCS Domain for HyperFlex Standard Cluster

To statically bind the HyperFlex Storage EPG and VLANs to vPC interfaces going to the UCS Domain that connect to the HyperFlex standard cluster, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
 - Application Profile: HXV-Storage_AP
 - **EPG:** HXV-CL0-StorData_EPG
 - **Static Paths:** HXV-UCS_6200FI-A_IPG, HXV-UCS_6200FI-B_IPG
 - **VLAN:** 3118
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.

- From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.
- From the left navigation pane, select and expand **Tenant HXV-Foundation > Application Profiles > HXV-Storage_AP > Application EPGs > HXV-CL0-StorData_EPG**.
- Right-click **HXV-CL0-StorData_EPG** and select **Deploy Static EPG on PC, VPC or Interface**.
- In the **Deploy Static EPG on PC, VPC or Interface** pop-up window, for **Path Type**, select **Virtual Port Channel**. For the **Path**, select the vPC to the first UCS Fabric Interconnect from the drop-down list. For the **Port Encap**, leave **VLAN** selected in the drop-down menu and in the box, specify the VLAN ID for the In-Band Management EPG. For the **Deployment Immediacy**, select **Immediate**.

The screenshot shows the Cisco APIC interface with the following details:

- Top Navigation:** System, **Tenants**, Fabric, Virtual Networking, L4-L7 Services, Admin, Operations, Apps.
- Tenant Search:** name or descr | common | **HXV-Foundation** | mgmt | infra
- Left Navigation Pane:**
 - Quick Start
 - Tenant HXV-Foundation
 - Application Profiles
 - HXV-IB-MGMT_AP
 - HXV-Storage_AP
 - Application EPGs
 - HXV-CL0-StorData...**
 - uSeg EPGs
 - HXV-vMotion_AP
 - Networking
 - Contracts
 - Policies
 - Services

- EPG - HXV-CL0-StorData_EPG** tabs: Summary (selected), Policy, Operational, Stats, Health, Faults, History.
- Deploy Static EPG On PC, VPC, Or Interface** dialog box:
- Select PC, VPC, or Interface**
- Path Type:** Port, Direct Port Channel, **Virtual Port Channel**
- Path:** HXV-UCS-6200FI-A
- Port Encap (or Secondary VLAN for Micro-Seg):** VLAN, 3118 (Integer Value)
- Deployment Immediacy:** **Immediate**, On Demand
- Primary VLAN for Micro-Seg:** VLAN, (Integer Value)
- Mode:** **Trunk**, Access (802.1P), Access (Untagged)
- IGMP Snoop Static Group:** (Empty table with Group Address and Source Address columns)
- Buttons:** Cancel, Submit

- Click **Submit**.
- Repeat steps 1-6 to bind the EPG to the vPC going to the second UCS Fabric Interconnect.

Configure ACI Fabric for HyperFlex vMotion Traffic

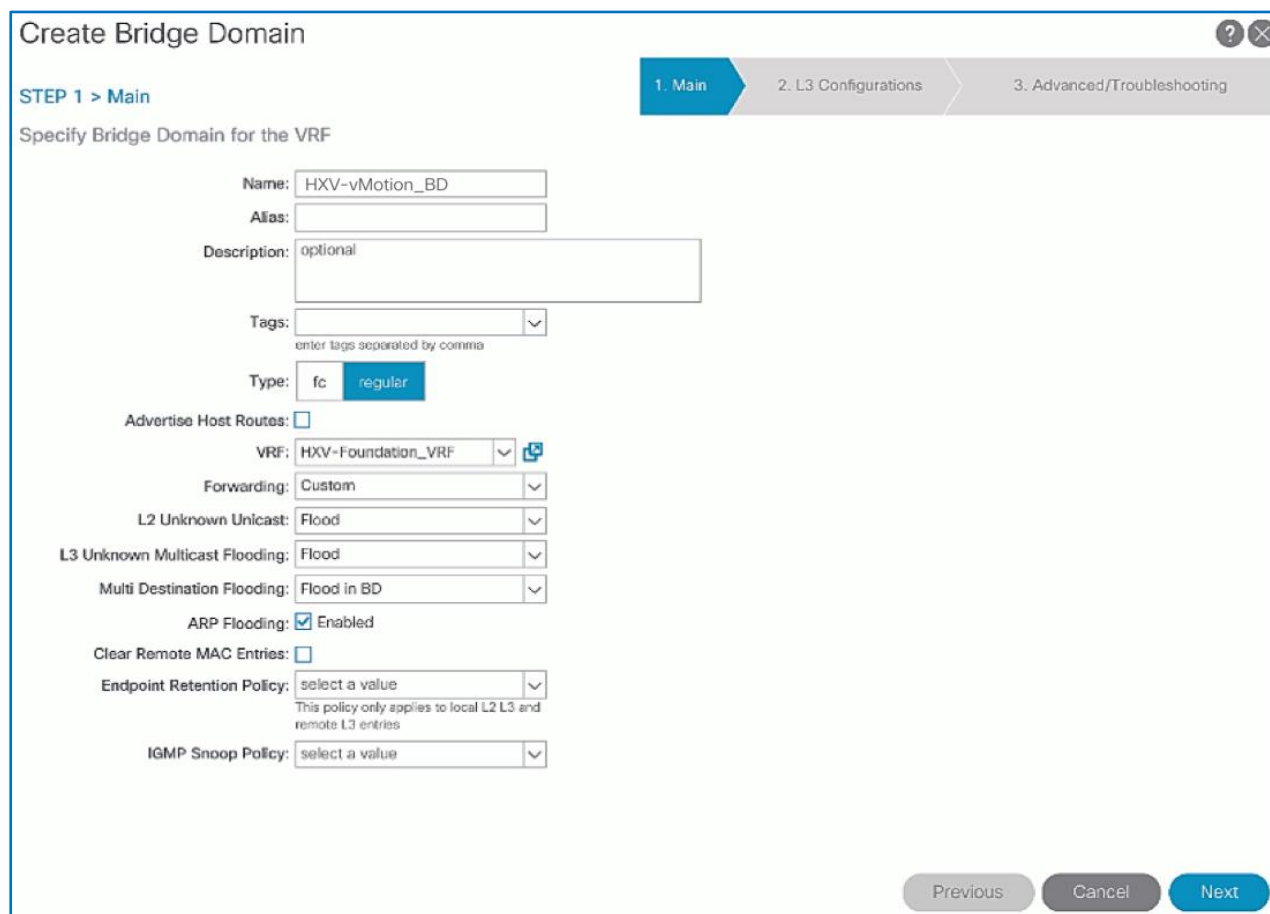
This section details the configuration of the Cisco ACI fabric to enable access to the vMotion network. This is minimally required to support failure scenarios where traffic from one UCS Fabric needs to be forwarded to another when different hosts are forwarding on different fabrics (FI-A, FI-B).

The vMotion network can also be optionally configured with a gateway in the ACI network to enable L3 connectivity to the existing infrastructure. For more information, see the VMware guidelines for L3 vMotion.

Create Bridge Domain for HyperFlex vMotion Traffic

To create a Bridge Domain for HyperFlex vMotion traffic, follow these steps using the setup information provided below:

- Tenant: `HXV-Foundation`
 - VRF: `HXV-Foundation_VRF`
 - Bridge Domain: `HXV-vMotion_BD`
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 2. From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on `HXV-Foundation`.
 3. From the left navigation pane, expand and select Tenant `HXV-Foundation > Networking > Bridge Domains.`
 4. Right-click and select **Create Bridge Domain**.
 5. In the **Create Bridge Domain** wizard, for **Name**, specify a name for the bridge domain. For **VRF**, select the previously created VRF from the drop-down list. For **Forwarding**, select **Custom** from the drop-down list. For **L2 Unknown Unicast**, select **Flood** from the drop-down list. The checkbox for **ARP Flooding** should now show up and be enabled.



6. Click **Next**.
7. In the **L3 Configurations** section, for **EP Move Detection Mode**, select the checkbox to enable **GARP based detection** if needed. See [Review/Enable ACI Fabric Settings](#) section for more details on when to enable this feature..
8. Click **Next**. Skip the **Advanced/Troubleshooting** section. Click **Finish** to complete.

Create Application Profile for HyperFlex vMotion Traffic

To create an application profile for HyperFlex vMotion traffic, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
 - **Application Profile:** HXV-vMotion_AP
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 2. From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.
 3. From the left navigation pane, select **Tenant HXV-Foundation**.
 4. Right-click **Tenant HXV-Foundation** and select **Create Application Profile**.
 5. In the **Create Application Profile** pop-up window, specify a **Name** the Application Profile.

The screenshot shows the Cisco APIC GUI with the 'Create Application Profile' dialog open for Tenant HXV. The dialog is titled 'Specify Tenant Application Profile' and contains the following fields:

- Name:** HXV-vMotion_AP
- Alias:** (empty)
- Description:** optional
- Tags:** (empty dropdown with a note: 'enter tags separated by comma')
- Monitoring Policy:** select a value

Below the fields is a section for EPGs with a table that has the following columns: Name, Alias, BD, Domain, Switching Mode, Static Path, Static Path VLAN, Provided Contract, and Consumed Contract. The table is currently empty.

At the bottom right of the dialog are two buttons: 'Cancel' and 'Submit'.

6. Click **Submit** to complete.

Create EPG for HyperFlex vMotion and Associate with Bridge Domain

To create an EPG for HyperFlex vMotion traffic, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
 - **Application Profile:** HXV-vMotion_AP
 - **Bridge Domain:** HXV-vMotion_BD
 - **EPG:** HXV-vMotion_EPG
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 2. From the top navigation menu, select **Tenants** > HXV-Foundation. If you do not see this tenant in the top navigation menu, select **Tenants** > **ALL TENANTS** and double-click on HXV-Foundation.
 3. From the left navigation pane, select and expand **Tenant HXV-Foundation** > **Application Profiles** > HXV-vMotion_AP.

4. Right-click HXV-vMotion_AP and select **Create Application EPG**.
5. In the **Create Application EPG** pop-up window, for **Name**, specify a name for the EPG. For **Bridge Domain**, select the previously created Bridge Domain.

APIC admin

System **Tenants** Fabric Virtual Networking L4-L7 Services Admin Operations Apps

ALL TENANTS | Ad

Create Application EPG

STEP 1 > Identity

Specify the EPG Identity

1. Identity

Name: HXV-vMotion_EPG

Alias:

Description: optional

Tags: enter tags separated by comma

Contract Exception Tag:

QoS class: Unspecified

Custom QoS: select a value

Data-Plane Policer: select a value

Intra EPG Isolation: Enforced **Unenforced**

Preferred Group Member: **Exclude** Include

Flood on Encapsulation: **Disabled** Enabled

Bridge Domain: HXV-vMotion_BD

Monitoring Policy: select a value

FHS Trust Control Policy: select a value

Shutdown EPG: ☐

Associate to VM Domain Profiles: ☐

Statically Link with Leaves/Paths: ☐

EPG Contract Master:

Application EPGs

Previous Cancel **Finish**

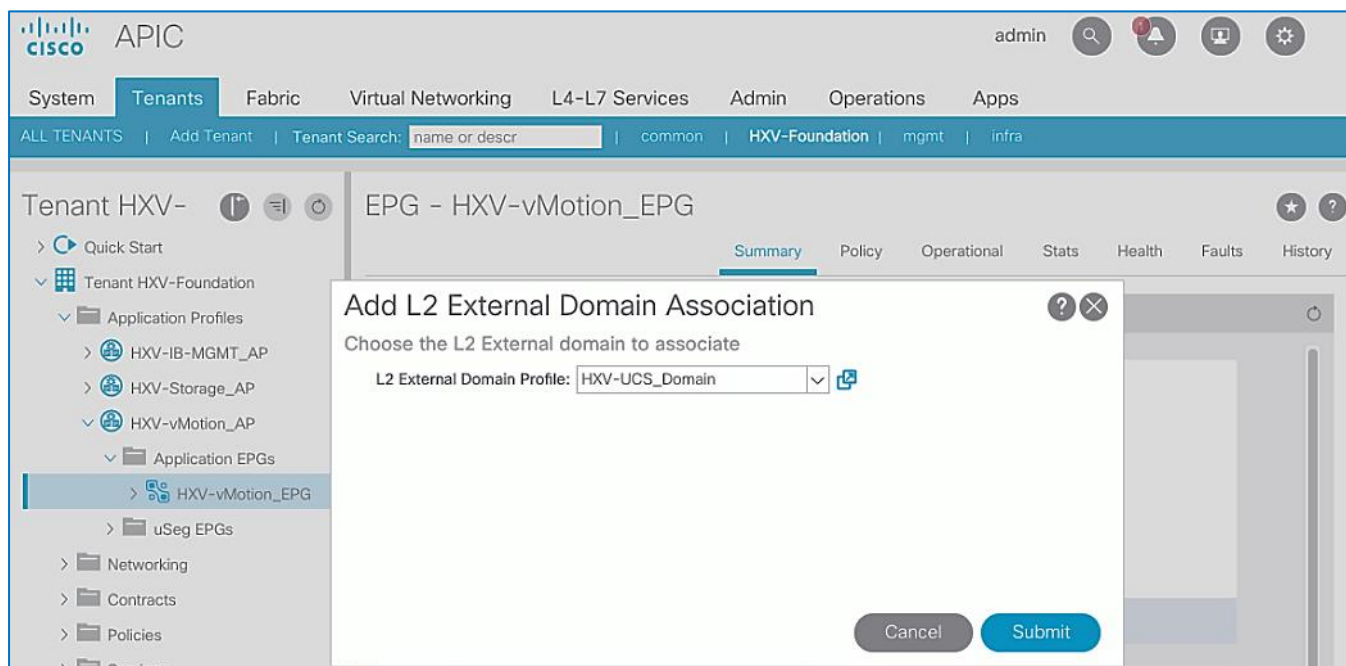
6. Click **Finish**.

Associate EPG with UCS Domain

To associate the HyperFlex vMotion EPG with UCS Domain, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
- **Application Profile:** HXV-vMotion_AP
- **Bridge Domain:** HXV-vMotion_BD
- **EPG:** HXV-vMotion_EPG
- **Domain:** HXV-UCS_Domain

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.
3. From the left navigation pane, select and expand **Tenant HXV-Foundation > Application Profiles > HXV-vMotion_AP > Application EPGs > HXV-vMotion_EPG**.
4. Right-click **HXV-vMotion_EPG** and select **Add L2 External Domain Association**.
5. In the **Add L2 External Domain Association** pop-up window, select the previously created UCS Domain from the list



6. Click **Submit**.

Create Static Binding for EPG on vPC Interfaces to UCS Domain

To statically bind the HyperFlex vMotion EPG and VLANs to vPC interfaces going to the UCS Domain, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
 - **Application Profile:** HXV-vMotion_AP
 - **EPG:** HXV-vMotion_EPG
 - **Static Paths:** HXV-UCS_6200FI-A_IPG, HXV-UCS_6200FI-B_IPG
 - **VLAN:** 3018
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 2. From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.
 3. From the left navigation pane, select and expand **Tenant HXV-Foundation > Application Profiles > HXV-vMotion_AP > Application EPGs > HXV-vMotion_EPG**.

- Right-click HXV-vMotion_EPG and select Deploy Static EPG on PC, VPC or Interface.
- In the **Deploy Static EPG on PC, VPC or Interface** pop-up window, for **Path Type**, select **Virtual Port Channel**. For the **Path**, select the vPC to the first UCS Fabric Interconnect from the drop-down list. For the **Port Encap**, leave **VLAN** selected in the drop-down menu and in the box, specify the VLAN ID for the In-Band Management EPG. For the **Deployment Immediacy**, select **Immediate**.

The screenshot shows the APIC GUI with the 'Deploy Static EPG On PC, VPC, Or Interface' dialog box open. The dialog is titled 'EPG - HXV-vMotion_EPG'. The 'Path Type' is set to 'Virtual Port Channel'. The 'Path' is 'HXV-UCS-6200FI-A'. The 'Port Encap (or Secondary VLAN for Micro-Seg)' is 'VLAN' with a value of '3018'. The 'Deployment Immediacy' is 'Immediate'. The 'Mode' is 'Trunk'. The 'IGMP Snoop Static Group' section is empty. The 'Cancel' and 'Submit' buttons are at the bottom right.

- Click **Submit**.
- Repeat the above steps to bind the EPG to the vPC going to the **second** UCS Fabric Interconnect.

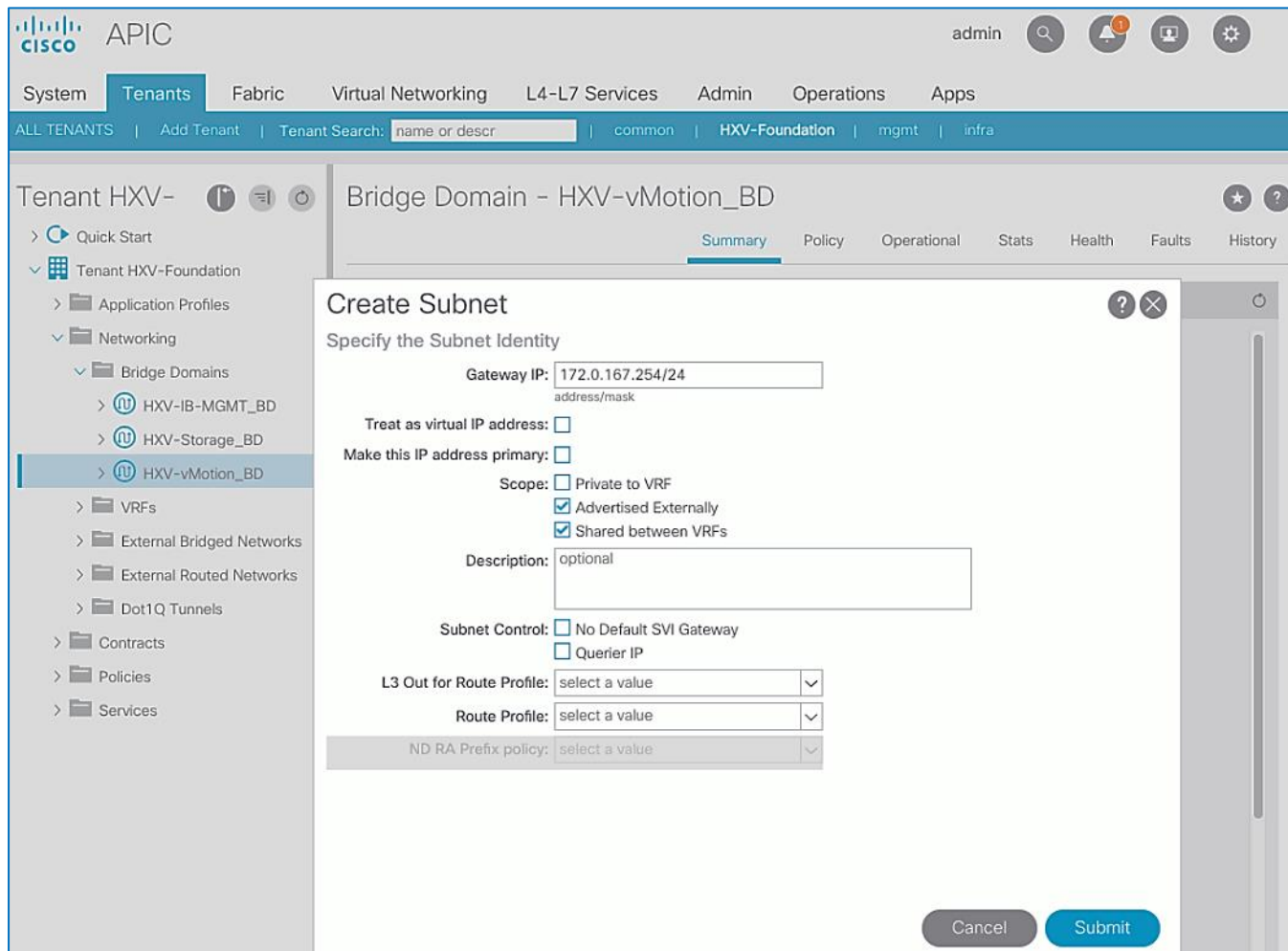
Configure Subnet Gateway for vMotion EPG (Optional)

To configure a gateway for the vMotion EPG, follow these steps using the setup information provided below:

- Tenant:** HXV-Foundation
- Bridge Domain:** HXV-vMotion_BD
- BD Subnet:** 172.0.167.254

- Use a browser to navigate to the APIC GUI. Log in using the **admin** account.

- From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.
- From the left navigation pane, select and expand **Tenant HXV-Foundation > Networking > Bridge Domains > HXV-vMotion_BD**.
- Right-click **HXV-vMotion_BD** and select **Create Subnet**.
- In the **Create Subnet** pop-up window, specify the **Default Gateway IP** and for **Scope**, select **Advertised Externally** and **Shared between VRFs**. Leave everything else as is.



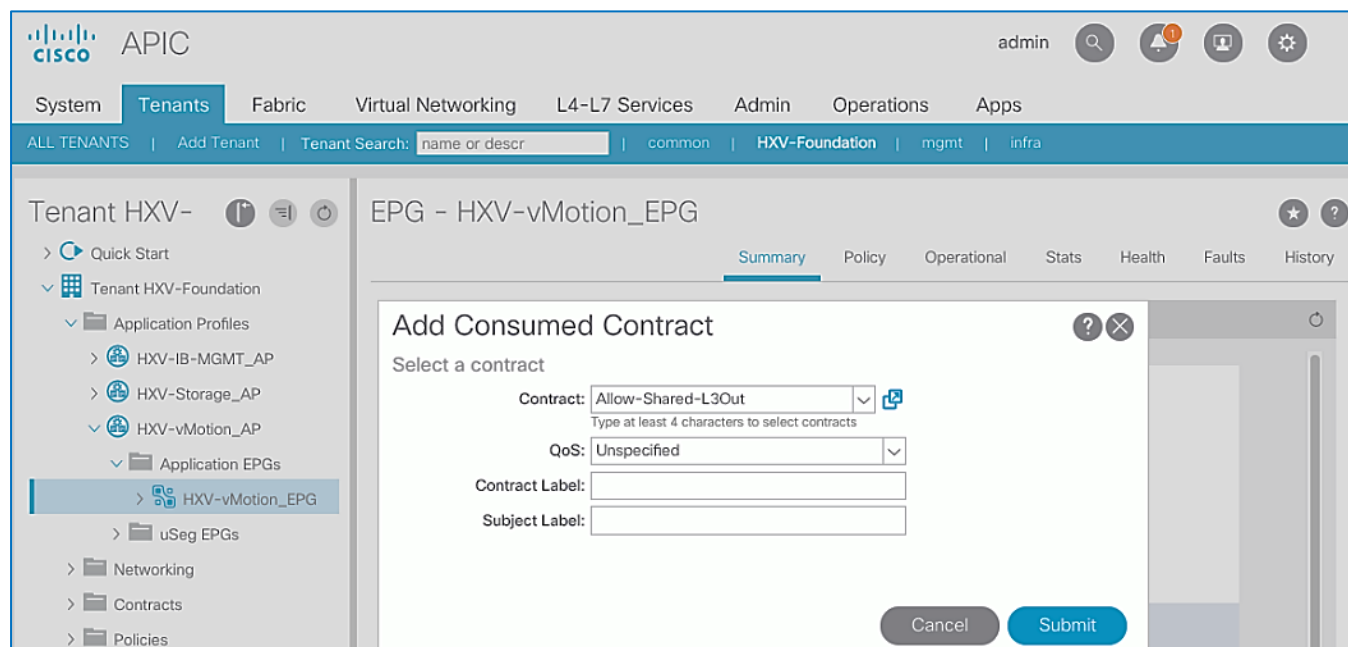
- Click **Submit**.

Create Contract to Access Outside Networks through Shared L3Out (Optional)

To enable a contract to access the network and services reachable via the **Shared L3Out** in the **common** Tenant, follow these steps using the setup information provided below:

- Tenant: HXV-Foundation
- Application Profile: HXV-vMotion_AP
- EPG: HXV-vMotion_EPG
- Contract: Allow-Shared-L3Out

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.
3. From the left navigation pane, select and expand **Tenant HXV-Foundation > Application Profiles > HXV-vMotion_AP > Application EPGs > HXV-vMotion_EPG**.
4. Right-click **HXV-vMotion_EPG** and select **Add Consumed Contract**.
5. In the **Add Consumed Contract** pop-up window, select the **Allow-Shared-L3Out** contract from the drop-down list.



6. Click **Submit**.

Solution Deployment – HyperFlex Management Cluster

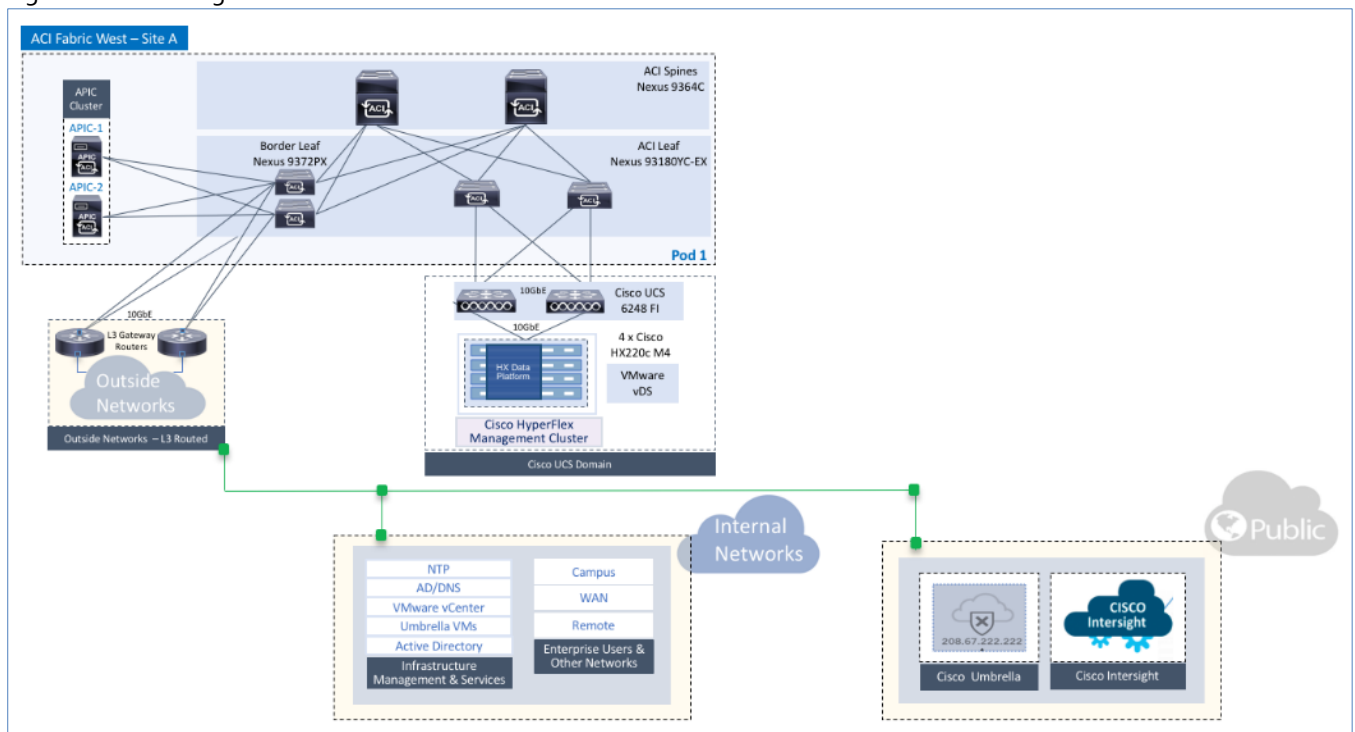
This section provides the detailed procedures for deploying a 4-node **standard** HyperFlex cluster from the cloud using Cisco Intersight. This cluster will serve as an *optional* Management cluster in this design. It will host virtual machines that provide management and infrastructure services to other HyperFlex cluster and Cisco UCS systems that connect to the same ACI Multi-Pod fabric. The cluster can also be installed and deployed using an on-premise HyperFlex Installer virtual machine. VMware vCenter that manages the cluster and other infrastructure services such as Active Directory, DNS, and so on, are located outside the ACI fabric and accessed through the shared L3Out connection from each Pod.



Cisco Intersight currently does not support the installation of HyperFlex stretched clusters. The stretched cluster covered in the next section is deployed using an on-premise HyperFlex Installer virtual machine.

Topology

Figure 22 Management Cluster



Install HyperFlex Cluster (Management) using Cisco Intersight

Cisco Intersight installation will configure Cisco UCS policies, templates, service profiles, and settings, as well as assigning IP addresses to the HX servers that come from the factory with ESXi hypervisor software preinstalled. The installer will deploy the HyperFlex controller virtual machines and software on the nodes, add the nodes to VMware vCenter managing the HX Cluster, and finally create the HyperFlex cluster and distributed filesystem. The above setup is done through a single workflow by providing the necessary information through an Installation wizard on Cisco Intersight.

Prerequisites

The prerequisites for installing a HyperFlex system from Cisco Intersight are as follows:

1. Factory installed HX Controller VM with HX Data Platform version 2.5(1a) or later, must be present on the HX servers. **Intersight deployment is not supported after cluster clean-up is completed.** However, all NEW HX servers may be deployed as-is.
2. Device Connectors on Fabric Interconnects must be able to resolve ***svc.ucs-connect.com***.
3. Allow outbound HTTPS connections (port 443) initiated from the Device Connectors on Fabric Interconnects. HTTP Proxy is supported.
4. Device Connectors (embedded in Fabric Interconnects) must be claimed and connected to Cisco Intersight – see [Enable Cisco Intersight Cloud-based Management](#) section.
5. Controller VM's management interface must be able to resolve ***download.intersight.com***.
6. Allow outbound HTTPS connections (port 443) initiated from Controller virtual machine's management interface. HTTP Proxy is supported.
7. Reachability from Cisco Intersight to the out-of-band management interfaces on Fabric Interconnects that the HyperFlex system being deployed connects to.
8. Reachability from Cisco Intersight to the out-of-band management (CIMC) interfaces on the servers, reachable via the Fabric Interconnects' management interfaces. This network (**ext-mgmt**) should be in the same subnet as the Fabric Interconnect management interfaces.
9. Reachability from Cisco Intersight to the ESXi in-band management interface of the hosts in the HyperFlex cluster being installed.
10. Reachability from Cisco Intersight to the VMware vCenter Server that will manage the HyperFlex cluster(s) being deployed. **Note:** The VMware vCenter Virtual Machine must be hosted on a separate virtualization environment and should not be on the HyperFlex cluster being deployed.
11. Reachability from Cisco Intersight to the DNS server(s) for use by the HyperFlex cluster being installed.
12. Reachability from Cisco Intersight to the NTP server(s) for use by the HyperFlex cluster being installed.
13. ACI Multi-Pod Fabric setup to enable connectivity to HyperFlex cluster networks - ESXi and Storage Controller management, ESXi and Storage Data networks, vMotion and Application VM networks.
14. Reachability from VMware vCenter to ESXi and Storage Controller Management networks.
15. Enable the necessary ports to install HyperFlex from Cisco Intersight. For more information, see Networking Ports section in Appendix A of the HyperFlex Hardening Guide:
https://www.cisco.com/c/dam/en/us/support/docs/hyperconverged-infrastructure/hyperflex-hx-data-platform/HX-Hardening_Guide_v3_5_v12.pdf
16. Review the Pre-installation Checklist for Cisco HX Data Platform:
https://www.cisco.com/c/en/us/td/docs/hyperconverged_systems/HyperFlex_HX_DataPlatformSoftware/HyperFlex_Prereinstall_Checklist/b_HX_Data_Platform_Preinstall_Checklist.html

Setup Information

The setup information used in this design to install a standard HyperFlex cluster from Cisco Intersight is provided below.

Table 57 Cluster Configuration – General

HyperFlex Cluster Configuration - Management		
HyperFlex Cluster Name	HXV-Cluster0	Name used in VMware vCenter and HyperFlex Connect
HX Data Platform Version	3.5 (2e)	Selected from the drop-down list
Type	Cisco HyperFlex with Fabric Interconnect	
Replication Factor (RF)	3	Default

Table 58 Cluster Configuration - Security

	Username	Password
Hypervisor	root	*****
Controller VM	Admin	*****

Table 59 Cluster Configuration – DNS, NTP and Timezone

HyperFlex Cluster Configuration – DNS, NTP and Timezone		
Timezone	America/New_York	
DNS Suffix	hxv.com	
NTP	192.168.167.254	
DNS Servers	10.99.167.244, 10.99.167.245	Cisco Umbrella - On-Premise Virtual Appliances

Table 60 Cluster Configuration – vCenter

HyperFlex Cluster Configuration – VMware vCenter	
vCenter Server FQDN or IP	hxv-vcsa-0.hxv.com (10.99.167.240)
vCenter Username	administrator@hxv.com
vCenter Password	*****
vCenter Datacenter Name	HXV-MGMT
vCenter Single-Sign-On Server	–

Table 61 Cluster Configuration – Storage Configuration

Policy	Enabled	
VDI Optimization	No	Default
Clean up Disk Partitions	No	Default
Logical Availability Zones	No	Default - Recommended for Clusters > 8 nodes

Table 62 Cluster Configuration – IP and Hostname

HyperFlex Cluster Configuration – IP and Hostname	
Hostname Prefix	hxx-cl0-esxi
Management Network Starting IP	10.1.167.101
Management Network Ending IP	10.1.167.104
Management Network Subnet Mask	255.255.255.0
Management Network Gateway	10.1.167.254
Controller VM Management Network Starting IP	10.1.167.151
Controller VM Management Network Ending IP	10.1.167.154
Controller VM Management Network Subnet Mask	255.255.255.0
Controller VM Management Network Gateway	10.1.167.254

Table 63 Cluster Configuration – Cisco UCS Manager Configuration

HyperFlex Cluster Configuration – UCS Manager Configuration	
Server Firmware Version	4.0 (1b)
MAC Prefix Starting Address	00:25:B5:A7
MAC Prefix Ending Address	00:25:B5:A7
KVM Starting IP	192.168.167.101
KVM Ending IP	192.168.167.104
KVM Subnet Mask	255.255.255.0
KVM Gateway	192.168.167.254

Table 64 Cluster Configuration – Network Configuration

Network Type	VLAN Name	VLAN ID
Management Network VLAN Name	hxv-inband-mgmt	118
VM Migration VLAN Name	hxv-vmotion	3018
VM Network VLAN Name	hxv-vm-network	1118
Jumbo Frames	Yes	

Table 65 Cluster Configuration – HyperFlex Storage Network

Network Type	VLAN Name	VLAN ID
HyperFlex Storage Data Network	hxv-cl0-storage-data	3118

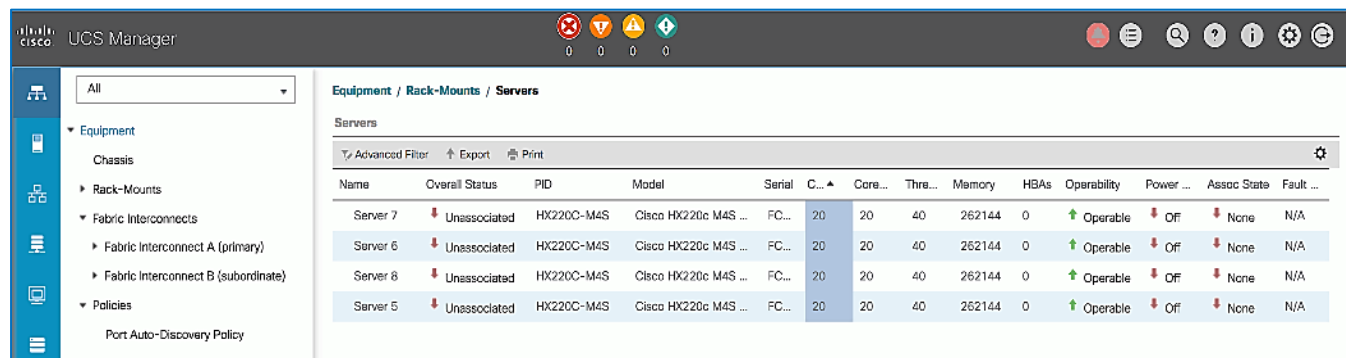
Deployment Steps

To install and deploy a HyperFlex standard cluster for Management from Cisco Intersight, complete the steps outlined in this section.

Verify Server Status before HyperFlex Install

Before starting the HyperFlex installation process that will create the service profiles and associate them with the servers, follow these steps to verify that the servers in the Cisco UCS domain have finished their discovery process and are in the correct state.

1. Use a browser to navigate to the UCS Manager GUI. Log in using the **admin** account.
2. From the left navigation pane, click the **Equipment** icon.
3. Navigate to **All > Equipment**. In the right window pane, click-on the **Servers** tab.



4. For the **Overall Status**, the servers should be in an **Unassociated** state. The servers should also be in an **Operable** state, powered **Off** and have no alerts with no faults or errors.
5. The servers are now ready for installing the HyperFlex Data Platform Software.

Connect to Cisco Intersight

To connect to Cisco Intersight, follow these steps:

1. Use a web browser to navigate to Cisco Intersight at <https://intersight.com/>.
2. Log in using a valid **cisco.com** account or single sign-on with your corporate authentication.

Deploy HyperFlex Cluster using Installation Wizard

To deploy the HyperFlex cluster using the wizard, follow these steps:

1. From Cisco Intersight, use the left navigation menu to select the **Service Profiles** icon.
2. In the right window pane, click the **Create HyperFlex Cluster Profile** button on the top right to open the HyperFlex cluster creation wizard.
3. In the **General** section of the **Create HyperFlex Cluster Profile** wizard, specify a **Name** for the HyperFlex cluster. The same name will be used for the HyperFlex Data Platform cluster and in VMware vCenter. For **HyperFlex Data Platform Version**, select the version from the drop-down list. For **Type**, select **Cisco HyperFlex with Fabric Interconnect**. For **Replication Factor**, select **3** (default) or **2**.


The screenshot shows the 'Create HyperFlex Cluster Profile' wizard in the Cisco Intersight interface. The left sidebar contains navigation links: Dashboards, Servers, HyperFlex Clusters, Fabric Interconnects, Service Profiles (selected), Policies, and Devices. The main panel is titled 'Create HyperFlex Cluster Profile' and shows the 'General' section selected in the left-hand menu. A blue informational banner at the top right states: 'Prior to creating a HyperFlex Cluster profile, ensure that you go through the pre-installation checklist and the detailed HyperFlex installation instructions, [here](#).' The form fields are as follows:

- Name ***: HXV-Cluster0
- HyperFlex Data Platform Version**: 3.5(1a)
- Type**: ☒ Cisco HyperFlex with Fabric Interconnect (selected), ☐ Cisco HyperFlex Edge
- Replication Factor**: ☒ 3 (selected), ☐ 2
- Description**: (empty text field)
- Add Tag**: (empty text field)

At the bottom of the wizard, there are 'Cancel' and 'Next' buttons.

4. Click **Next**.
5. In the **Cluster Configuration** section of the **Create HyperFlex Cluster Profile** wizard, select and expand **Security**. Specify passwords for Hypervisor and Control VM Admin user (**root**).



Note the green check  icon next to Security; this indicates that valid parameters were entered and that a policy was created (name on the top right). The policy is saved under Policies in the left navigation menu and can be individually accessed and edited.

- In the **Cluster Configuration** section of the **Create HyperFlex Cluster Profile** wizard, select and expand **DNS, NTP and Timezone**. For **Timezone**, select the appropriate Timezone from the drop-down list. For **DNS Suffix**, specify the **Domain name** for the cluster. For **DNS Servers**, specify the Domain Name Servers for the environment – use the **[+]** to add multiple servers. For **NTP Servers**, specify a NTP Server for the cluster – use the **[+]** to add multiple servers.

The screenshot shows the 'Create HyperFlex Cluster Profile' wizard in the Cisco Intersight interface. The left sidebar contains navigation links: Dashboards, Servers, HyperFlex Clusters, Fabric Interconnects, Service Profiles, Policies, and Devices. The main panel is titled 'Create HyperFlex Cluster Profile' and shows a progress bar with steps: General, Cluster Configuration (selected), Nodes Assignment, Nodes Configuration, Summary, and Results. The 'Cluster Configuration' section is expanded, showing a list of configuration items: Security, DNS, NTP and Timezone, vCenter (optional), Storage Configuration (optional), Auto Support (optional), and IP & Hostname. The 'vCenter (optional)' item is selected and expanded, showing the following fields:

- vCenter Server FQDN or IP: `hvx-vcsa-0.hvx.com`
- vCenter Username: `administrator@hvx.co`
- vCenter Password: `*****`
- vCenter Datacenter Name: `HXV-MGMT`
- vCenter Single-Sign-On Server: (empty)

At the bottom of the wizard, there are buttons for 'Save & Close', 'Previous', and 'Next'.

- In the **Cluster Configuration** section of the **Create HyperFlex Cluster Profile** wizard, select and expand **vCenter**. Specify the information for the VMware vCenter managing the HX cluster in this section.

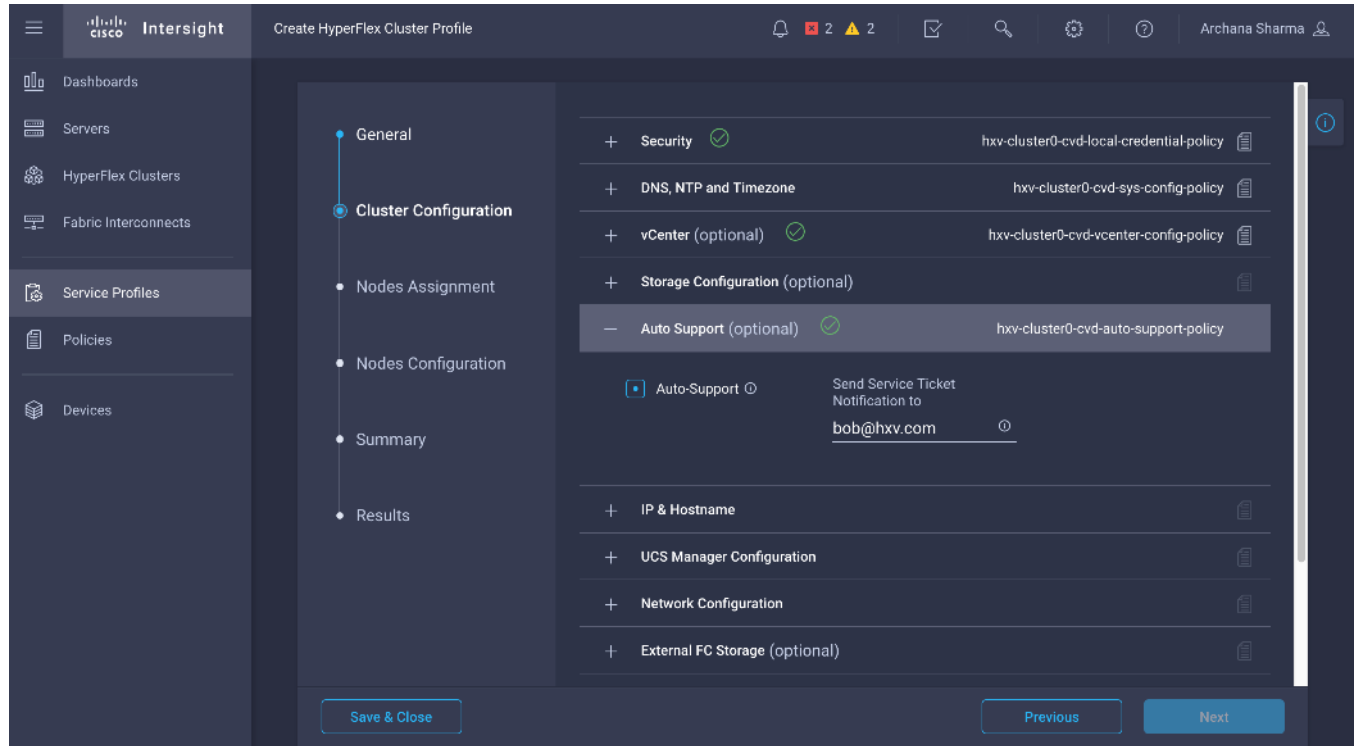
The screenshot shows the 'Create HyperFlex Cluster Profile' wizard in the Cisco Intersight interface. The left sidebar contains navigation links: Dashboards, Servers, HyperFlex Clusters, Fabric Interconnects, Service Profiles, Policies, and Devices. The main panel is titled 'Create HyperFlex Cluster Profile' and shows a progress bar with steps: General, Cluster Configuration (selected), Nodes Assignment, Nodes Configuration, Summary, and Results. The 'Cluster Configuration' section is expanded, showing a list of configuration items: Security, DNS, NTP and Timezone, vCenter (optional), Storage Configuration (optional), Auto Support (optional), IP & Hostname, and UCS Manager Configuration. The 'Storage Configuration (optional)' item is selected and expanded, showing the following fields:

- Storage Configuration: (empty)
- Auto Support: (empty)
- IP & Hostname: (empty)
- UCS Manager Configuration: (empty)

At the bottom of the wizard, there are buttons for 'Save & Close', 'Previous', and 'Next'.

- (Optional) In the Cluster Configuration section of the Create HyperFlex Cluster Profile wizard, select and expand Storage Configuration to specify storage policies such as VDI Optimization, Logical Availability Zones etc.

9. (Optional) In the **Cluster Configuration** section of the **Create HyperFlex Cluster Profile** wizard, select and expand **Auto Support** to specify the email account to send support ticket notifications to.



10. In the **Cluster Configuration** section of the **Create HyperFlex Cluster Profile** wizard, select and expand **IP & Hostname**. For the **Hostname Prefix**, specify a name for the ESXi hosts. For the **Management Network**, specify a starting and ending **IP address**, **subnet mask** and **gateway** for each ESXi host in the cluster. For the **Controller VM Management Network**, specify a starting and ending **Management IP address**, **subnet mask** and **gateway** for the controller virtual machine deployed on each host in the cluster.

Intersight Create HyperFlex Cluster Profile

General

Cluster Configuration

Nodes Assignment

Nodes Configuration

Summary

Results

Save & Close Previous Next

Auto Support (optional) ☒ hxv-cluster0-cvd-auto-support-policy

IP & Hostname ☒ hxv-cluster0-cvd-node-config-policy

Hostname Prefix *
hxv-cl0-esxi

Management Network Starting IP *
10.1.167.101

Management Network Ending IP *
10.1.167.104

Management Network Subnet Mask *
255.255.255.0

Management Network Gateway *
10.1.167.254

Controller VM Management Network Starting IP
10.1.167.151

Controller VM Management Network Ending IP
10.1.167.154

Controller VM Management Network Subnet Mask
255.255.255.0

Controller VM Management Network Gateway
10.1.167.254

11. In the **Cluster Configuration** section of the **Create HyperFlex Cluster Profile** wizard, select and expand **UCS Manager Configuration**. For the **Server Firmware Version**, specify the Cisco UCS Manager version running on the Fabric Interconnects. For the **MAC Prefix**, specify a starting and ending **MAC Prefix range** for the HX nodes. For **KVM** management, specify a starting and ending **IP address**, **subnet mask** and **gateway** for out-of-band management of each HX node in the cluster.

Intersight Create HyperFlex Cluster Profile

General

Cluster Configuration

Nodes Assignment

Nodes Configuration

Summary

Results

Save & Close Previous Next

IP & Hostname ☒ hxv-cluster0-cvd-node-config-policy

UCS Manager Configuration ☒ hxv-cluster0-cvd-ucsm-config-policy

Server Firmware Version *
4.0(1b)

MAC Prefix Starting Address *
00:25:B5:A7

MAC Prefix Ending Address *
00:25:B5:A7

KVM Starting IP *
192.168.167.101

KVM Ending IP *
192.168.167.104

KVM Subnet Mask *
255.255.255.0

KVM Gateway *
192.168.167.254

Network Configuration

External FC Storage (optional)

12. In the **Cluster Configuration** section of the **Create HyperFlex Cluster Profile** wizard, select and expand **Network Configuration**. For the **Management Network VLAN**, specify the **VLAN Name** and **ID** used for in-band ESXi management of HX nodes. For the **VM Migration VLAN**, specify the **VLAN Name** and **VLAN ID** used for vMotion. For the **VM Network VLAN**, specify the **VLAN Name** and **VLAN ID** used for virtual machines hosted on the HX cluster. For **Jumbo Frames**, enable it.

The screenshot shows the 'Create HyperFlex Cluster Profile' wizard in the Cisco Intersight interface. The left sidebar contains navigation options: Dashboards, Servers, HyperFlex Clusters, Fabric Interconnects, Service Profiles, Policies, and Devices. The main panel is titled 'Create HyperFlex Cluster Profile' and shows a progress bar with steps: General, Cluster Configuration (selected), Nodes Assignment, Nodes Configuration, Summary, and Results. The 'Cluster Configuration' section is expanded, showing the 'Network Configuration' tab. The configuration details are as follows:

Network Configuration	
Management Network VLAN Name *	Management Network VLAN ID *
hxx-inband-mgmt	118
VM Migration VLAN Name *	VM Migration VLAN ID *
hxx-vmotion	3018
VM Network VLAN Name *	VM Network VLAN ID *
hxx-vm-network	1118
<input checked="" type="checkbox"/> Jumbo Frames	
+ External FC Storage (optional)	
+ External iSCSI Storage (optional)	
+ Proxy Setting (optional)	
+ HyperFlex Storage Network	

At the bottom of the wizard, there are three buttons: 'Save & Close', 'Previous', and 'Next'.

13. (Optional) In the **Cluster Configuration** section of the **Create HyperFlex Cluster Profile** wizard, select and expand **External FC Storage** if external FC storage is used.
14. (Optional) In the **Cluster Configuration** section of the **Create HyperFlex Cluster Profile** wizard, select and expand **External iSCSI Storage** if external FC storage is used.
15. (Optional) In the **Cluster Configuration** section of the **Create HyperFlex Cluster Profile** wizard, select and expand **Proxy Setting** if proxies are used.
16. In the **Cluster Configuration** section of the **Create HyperFlex Cluster Profile** wizard, select and expand **HyperFlex Storage Network**. For the **Storage Network VLAN**, specify the **VLAN Name** and **ID** used for the storage data network. This network will be accessed by ESXi hosts and Controller virtual machines.

Intersight Create HyperFlex Cluster Profile

Archana Sharma

Cluster Configuration

- General
- Cluster Configuration**
- Nodes Assignment
- Nodes Configuration
- Summary
- Results

Storage Configuration (optional)

Auto Support (optional) ✓ hvx-cluster0-cvd-auto-support-policy

IP & Hostname ✓ hvx-cluster0-cvd-node-config-policy

UCS Manager Configuration ✓ hvx-cluster0-cvd-ucsm-config-policy

Network Configuration ✓ hvx-cluster0-cvd-cluster-network-policy

External FC Storage (optional)

External iSCSI Storage (optional)

Proxy Setting (optional)

HyperFlex Storage Network ✓

Storage Network VLAN Name: hvx-cl0-storage-data

Storage Network VLAN ID: 3118

Save & Close Previous Next

17. Review the Cluster Configuration section of the Create HyperFlex Cluster Profile wizard.

Intersight Create HyperFlex Cluster Profile

Archana Sharma

Cluster Configuration

- General
- Cluster Configuration**
- Nodes Assignment
- Nodes Configuration
- Summary
- Results

Security ✓ hvx-cluster0-cvd-local-credential-policy

DNS, NTP and Timezone hvx-cluster0-cvd-sys-config-policy

vCenter (optional) ✓ hvx-cluster0-cvd-vcenter-config-policy

Storage Configuration (optional)

Auto Support (optional) ✓ hvx-cluster0-cvd-auto-support-policy

IP & Hostname ✓ hvx-cluster0-cvd-node-config-policy

UCS Manager Configuration ✓ hvx-cluster0-cvd-ucsm-config-policy

Network Configuration ✓ hvx-cluster0-cvd-cluster-network-policy

External FC Storage (optional)

External iSCSI Storage (optional)

Proxy Setting (optional)

HyperFlex Storage Network

Save & Close Previous Next

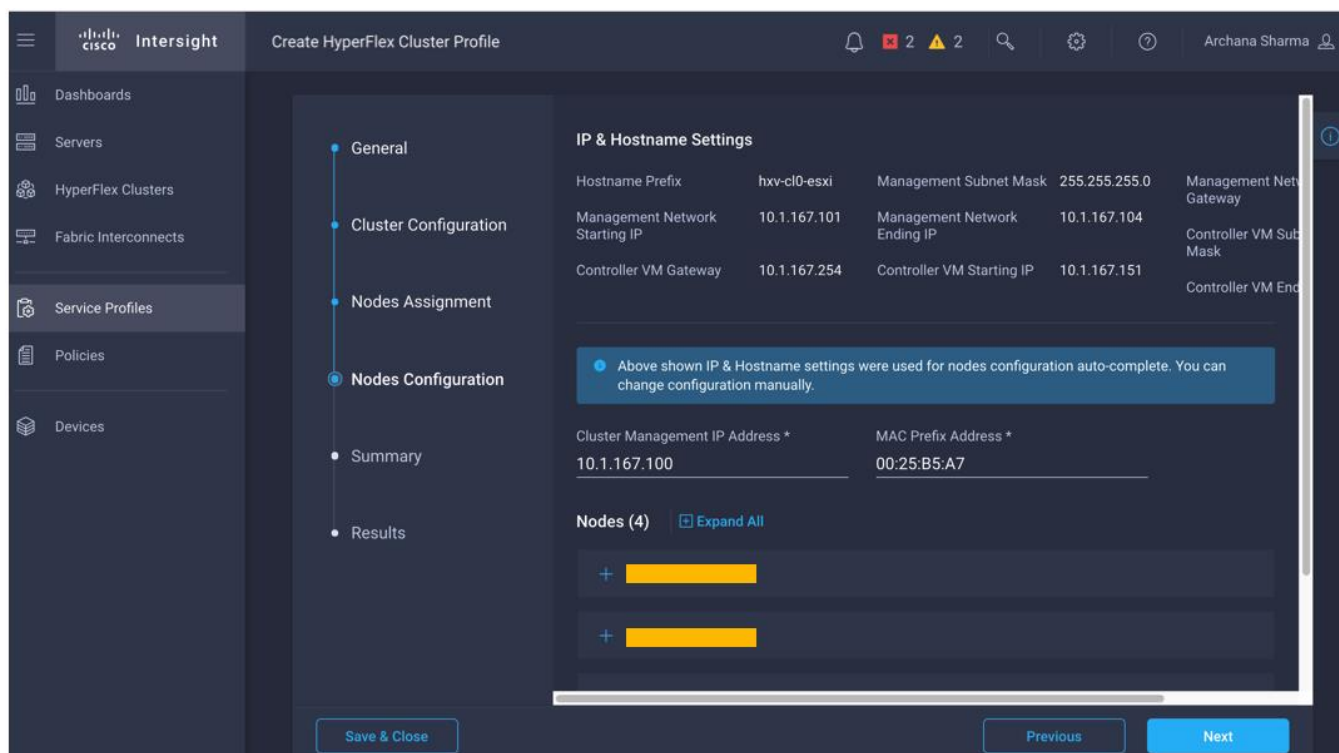
18. Click **Next**.

19. In the **Nodes Assignment** section of the **Create HyperFlex Cluster Profile** wizard, click **Assign Nodes** and select the nodes that should be added to the HX cluster.

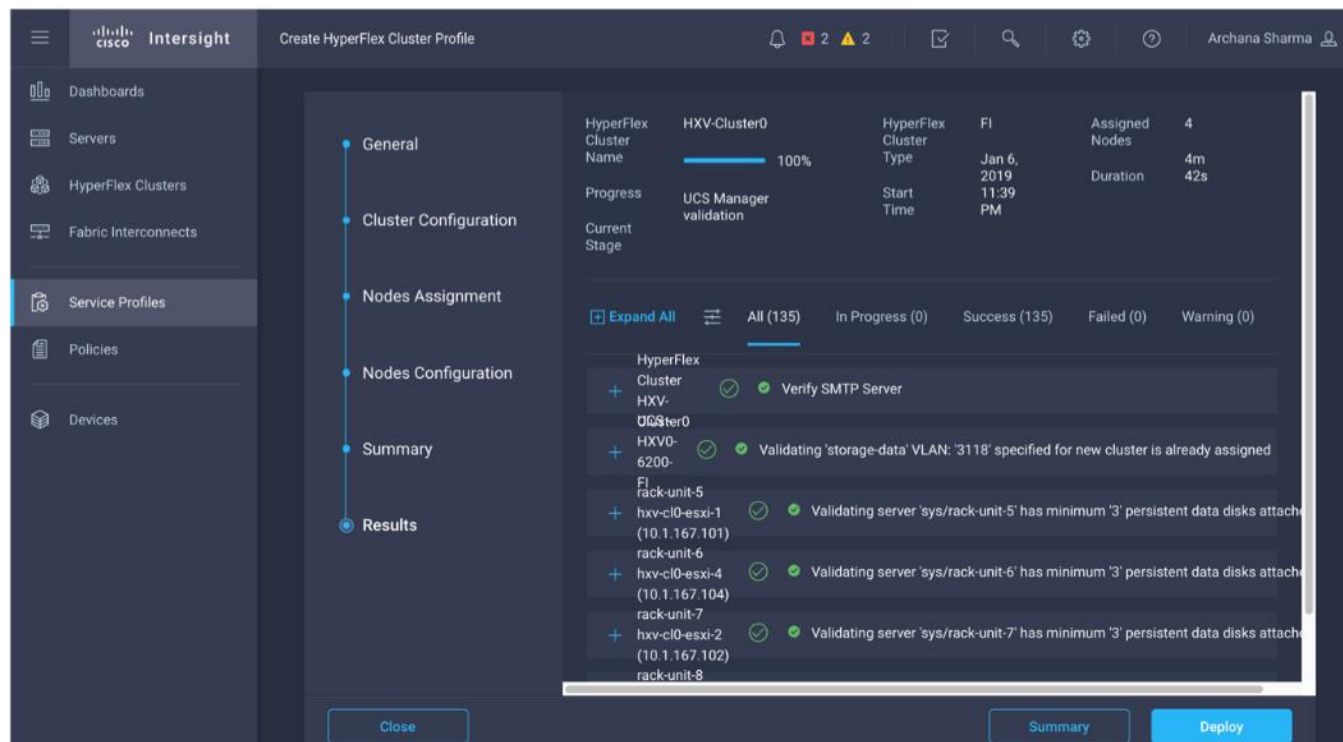
The screenshot shows the Cisco Intersight interface for the 'Create HyperFlex Cluster Profile' wizard. The left sidebar contains navigation links: Dashboards, Servers, HyperFlex Clusters, Fabric Interconnects, Service Profiles (selected), Policies, and Devices. The main content area has a breadcrumb trail: General, Cluster Configuration, Nodes Assignment (selected), Nodes Configuration, Summary, and Results. A blue banner at the top states: 'Cisco HyperFlex Fabric Interconnect cluster allows a minimum of 3 to a maximum of 32 nodes.' Below this, there are radio buttons for 'Assign Nodes' (selected) and 'Assign Nodes Later', and a toggle for 'Show selected(4)'. A table displays 12 items found, with 10 per page. The table has columns: Name, Assign Status, Model, and Serial. Four nodes are selected (marked with a blue square): HXV0-6200-FI-5, HXV0-6200-FI-6, HXV0-6200-FI-7, and HXV0-6200-FI-8, all with an 'Assigned' status. Two nodes are not selected: HXV1-6300-FI-1 and HXV1-6300-FI-2, both with a 'Not Assigned' status. At the bottom, there are buttons for 'Save & Close', 'Previous', and 'Next'.

Name	Assign Status	Model	Serial
HXV0-6200-FI-5	Assigned	HX220C-M4S	FCH1951V07E
HXV0-6200-FI-6	Assigned	HX220C-M4S	FCH1951V06A
HXV0-6200-FI-7	Assigned	HX220C-M4S	FCH1949V2QJ
HXV0-6200-FI-8	Assigned	HX220C-M4S	FCH1951V06J
HXV1-6300-FI-1	Not Assigned	HX220C-M5SX	WZP22060AU8
HXV1-6300-FI-2	Not Assigned	HX220C-M5SX	WZP220607LD

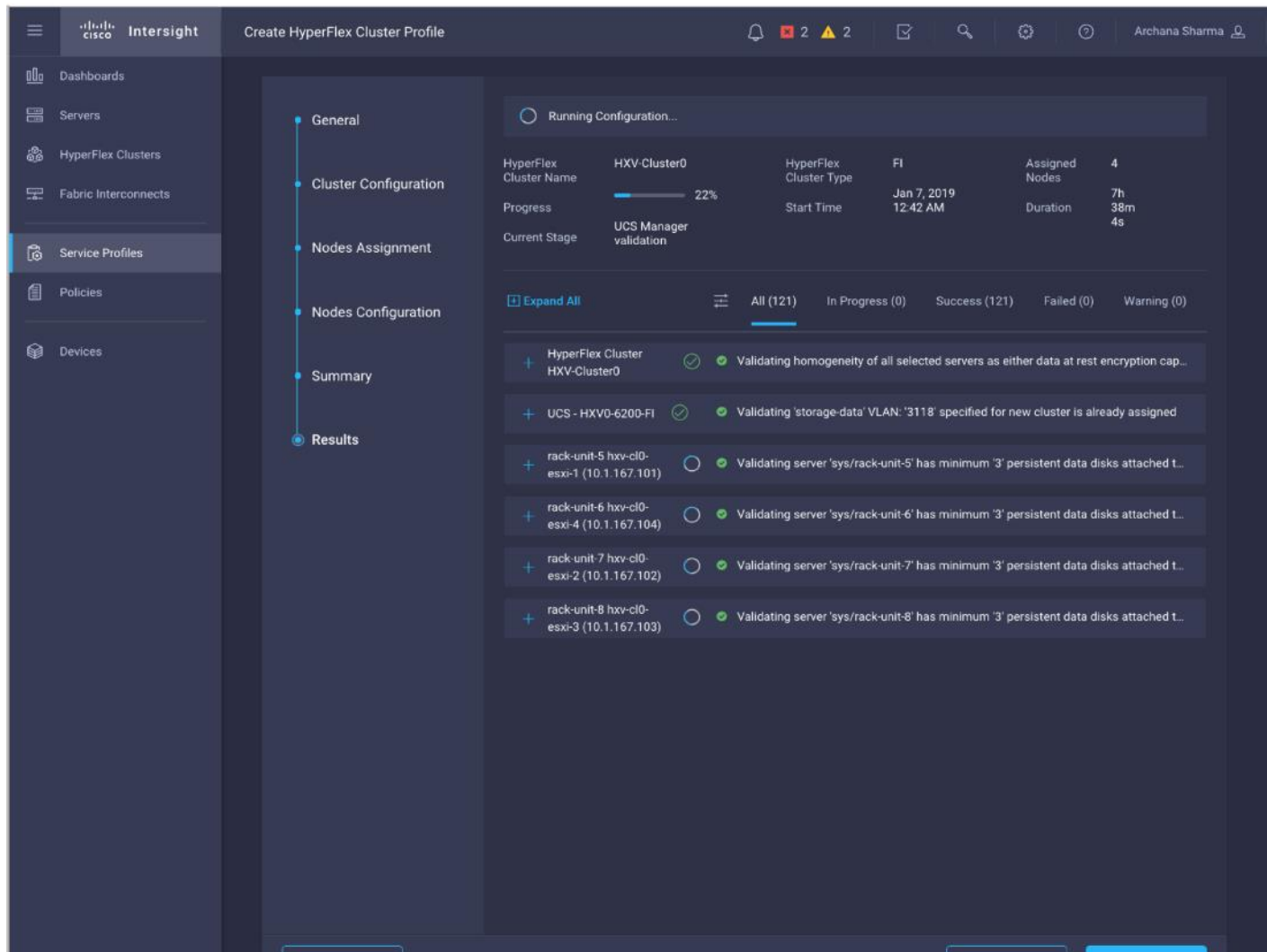
20. Click **Next**.
21. In the **Nodes Configuration** section of the **Create HyperFlex Cluster Profile** wizard, specify the Cluster Management Address.



22. Click **Next**.
23. In the **Summary** section of the **Create HyperFlex Cluster Profile** wizard, review the configuration done so far. Click **Validate** to validate the configuration before deploying it.



24. When the validation completes, click **Deploy** to install and configure the HX system.

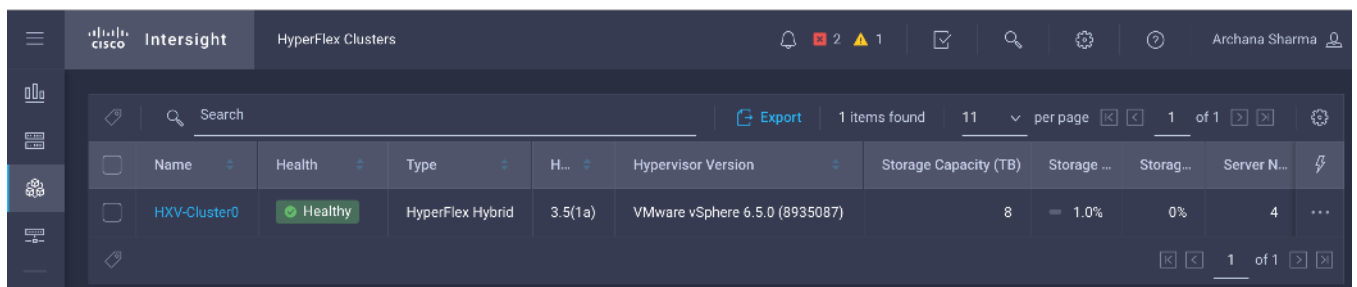


- When the install is complete, proceed to the next section to verify the cluster setup and proceed to the post-installation steps to complete the deployment.

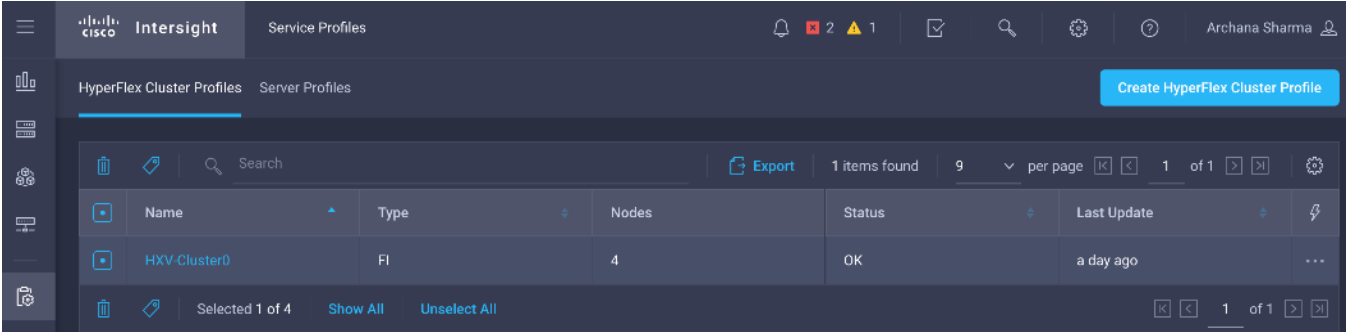
Verify HyperFlex Cluster Installation

To verify that the install was successful from Cisco Intersight, follow these steps:

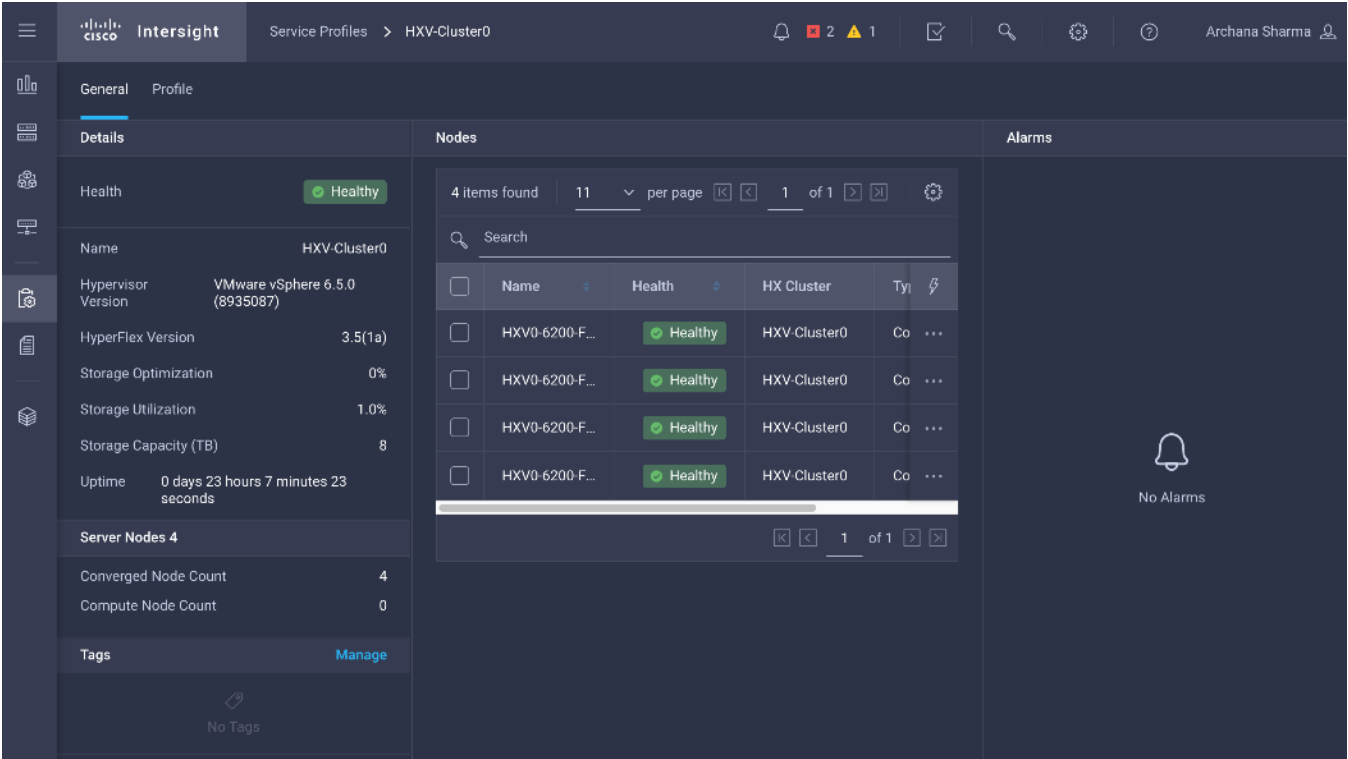
- From Cisco Intersight, use the left navigation menu to select the **HyperFlex Cluster** icon.
- In the right window pane, review the information for the newly deployed HyperFlex cluster.



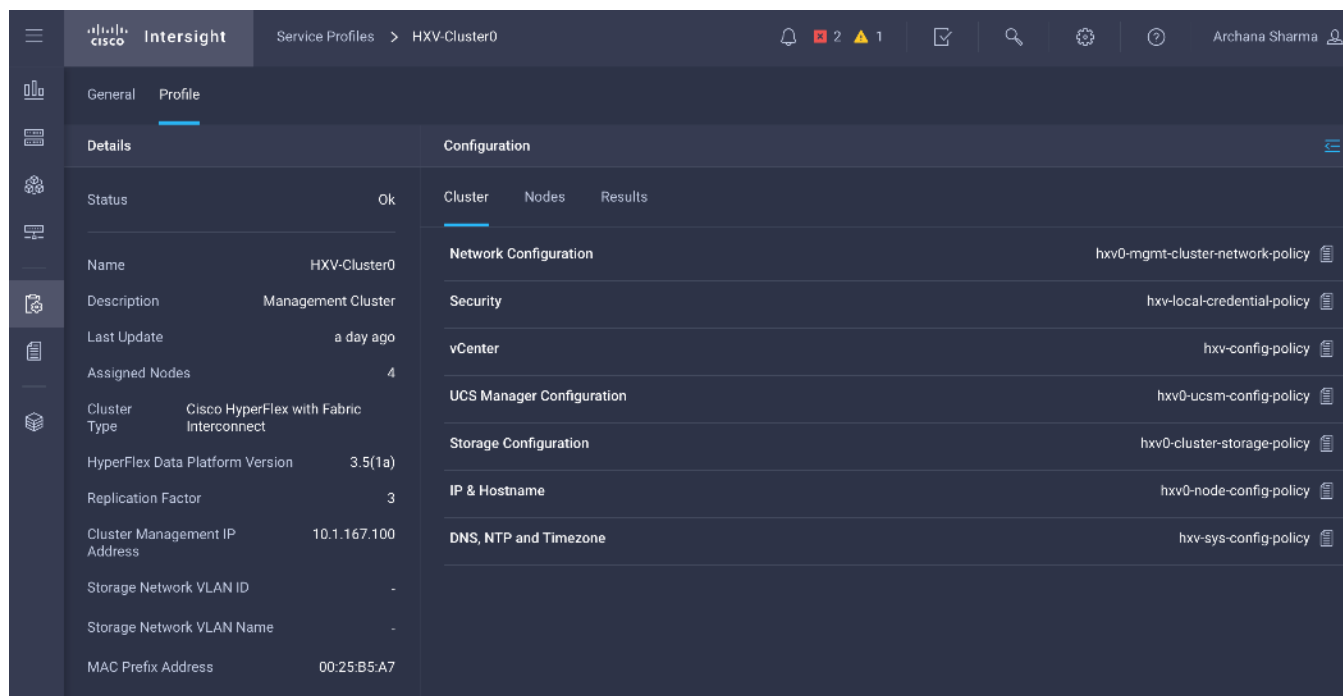
- From the left navigation menu, select the **Service Profiles** icon.




4. In the right window pane, select the Service Profile for the newly deployed HX cluster and double-click the Service Profile to review the information in the **General** tab.



5. Select the **Profile** tab to review additional information about the newly deployed HX cluster.



6. In the **Configuration** section on the right side of the window, under the **Cluster** tab, the individual policies are listed. Click the  icon on the top right to see the details of each policy.
7. Navigate to the **Nodes** tab and **Results** tab for more details on the newly deployed HX cluster.

Complete Post-Installation Tasks

When the installation is complete, additional best-practices and configuration can be implemented using a Cisco provided post-installation script. The script should be run before deploying virtual machine workloads on the cluster. The script is executed from the HyperFlex Controller virtual machine and can do the following:

- License the hosts in VMware vCenter
- Enable HA/DRS on the cluster in VMware vCenter
- Suppress SSH/Shell warnings in VMware vCenter
- Configure vMotion in VMware vCenter
- Enables configuration of additional guest VLANs/port-groups
- Send test Auto Support (ASUP) email if enabled during the install process
- Perform HyperFlex Health check

To run the post-install script to do the above configuration, follow these steps:

1. SSH into a HX Controller VM. Log in using the **admin/root** account.
2. From the Controller VM, run the following command to execute the post-install script:
`/usr/share/springpath/storfs-misc/hx-scripts/post_install.py`
3. Follow the on-screen prompts to complete the post-install configuration.

```

root@SpringpathControllerNWZVFY5XRB:~#
root@SpringpathControllerNWZVFY5XRB:~# /usr/share/springpath/storfs-misc/hx-scripts/post_
install.py
Logging in to controller localhost
HX CVM admin password:
Getting ESX hosts from HX cluster...
vCenter URL: 10.99.167.240
Enter vCenter username (user@domain): administrator@hxv.com
vCenter Password:
Found datacenter HXV-MGMT
Found cluster HXV-Cluster0
Enter ESX root password:

Enter vSphere license key? (y/n) y

1. Add License Key
2. Switch to evaluation mode

Selection: 2
License key on 10.1.167.101 was not Foundation. Skipping license key modification.
License key on 10.1.167.102 was not Foundation. Skipping license key modification.
License key on 10.1.167.103 was not Foundation. Skipping license key modification.
License key on 10.1.167.104 was not Foundation. Skipping license key modification.

Enable HA/DRS on cluster? (y/n) y

Disable SSH warning? (y/n) y

Add vmotion interfaces? (y/n) y
Netmask for vMotion: 255.255.255.0
VLAN ID: (0-4096) 3018
vMotion MTU is set to use jumbo frames (9000 bytes). Do you want to change to 1500 bytes? (y/n) n
vMotion IP for 10.1.167.101: 172.0.167.101
Adding vmotion-3018 to 10.1.167.101
Adding vmkernel to 10.1.167.101
vMotion IP for 10.1.167.102: 172.0.167.102
Adding vmotion-3018 to 10.1.167.102
Adding vmkernel to 10.1.167.102
vMotion IP for 10.1.167.103: 172.0.167.103
Adding vmotion-3018 to 10.1.167.103
Adding vmkernel to 10.1.167.103
vMotion IP for 10.1.167.104: 172.0.167.104
Adding vmotion-3018 to 10.1.167.104
Adding vmkernel to 10.1.167.104

Add VM network VLANs? (y/n) y
Attempting to find UCSM IP
Could not find UCSM IP, enter IP address: 192.168.167.201
UCSM Username: admin
UCSM Password:
HX UCS Sub Organization: HXV-Cluster0
Port Group Name to add (VLAN ID will be appended to the name): hxv-vm-network
VLAN ID: (0-4096) 1218
Adding VLAN 1218 to FI
Adding VLAN 1218 to vm-network-a VNIC template
Adding hxv-vm-network-1218 to 10.1.167.101
Adding hxv-vm-network-1218 to 10.1.167.102
Adding hxv-vm-network-1218 to 10.1.167.103
Adding hxv-vm-network-1218 to 10.1.167.104
Add additional VM network VLANs? (y/n) n

Run health check? (y/n) y

Validating cluster health and configuration...

Cluster Summary:
Version - 3.5.1a-31118
Model - HX220C-M4S
Health - HEALTHY
ASUP enabled - False
root@SpringpathControllerNWZVFY5XRB:~# █

```



Any VLANs created on the HyperFlex cluster and UCSM will need corresponding configuration in the ACI fabric to enable forwarding for that VLAN within the ACI Fabric.

Enable Smart Licensing

HyperFlex 2.5 and later utilizes Cisco Smart Licensing, which communicates with a Cisco Smart Account to validate and check out HyperFlex licenses to the nodes, from the pool of available licenses in the account. At the beginning, Smart Licensing is enabled but the HX storage cluster is unregistered and in a 90-day evaluation period or EVAL MODE. For the HX storage cluster to start reporting license consumption, it must be registered with the Cisco Smart Software Manager (SSM) through a valid Cisco Smart Account. Before beginning, verify that you have a Cisco Smart account, and valid HyperFlex licenses are available to be checked out by your HX cluster.

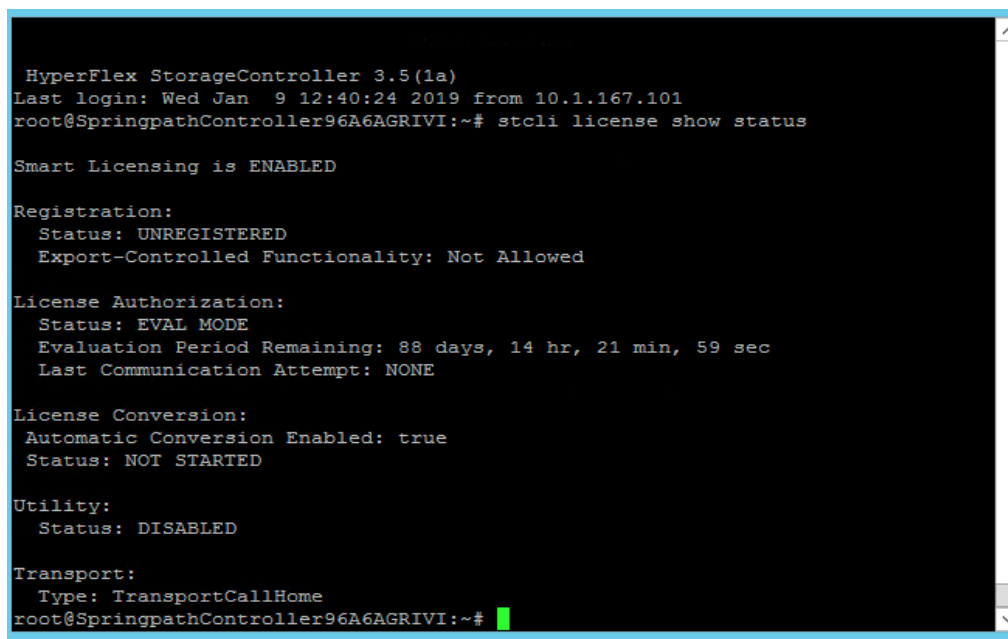
To create a Smart Account, see Cisco Software Central > Request a Smart Account:

<https://webapps.cisco.com/software/company/smartaccounts/home?route=module/accountcreation>.

To activate and configure smart licensing, follow these steps:

1. SSH into a HX Controller VM. Log in using the **admin/root** account.
2. Confirm that your HX storage cluster is in **Smart Licensing mode**.

```
# stcli license show status
```



```
HyperFlex StorageController 3.5(1a)
Last login: Wed Jan  9 12:40:24 2019 from 10.1.167.101
root@SpringpathController96A6AGRIVI:~# stcli license show status

Smart Licensing is ENABLED

Registration:
  Status: UNREGISTERED
  Export-Controlled Functionality: Not Allowed

License Authorization:
  Status: EVAL MODE
  Evaluation Period Remaining: 88 days, 14 hr, 21 min, 59 sec
  Last Communication Attempt: NONE

License Conversion:
  Automatic Conversion Enabled: true
  Status: NOT STARTED

Utility:
  Status: DISABLED

Transport:
  Type: TransportCallHome
root@SpringpathController96A6AGRIVI:~#
```

3. Feedback will show **Smart Licensing is ENABLED, Status: UNREGISTERED**, and the amount of time left during the 90-day evaluation period (in days, hours, minutes, and seconds).
4. Navigate to Cisco Software Central (<https://software.cisco.com/>) and log in to your Smart Account.
5. From Cisco Smart Software Manager, generate a registration token.
6. In the **License** pane, click **Smart Software Licensing** to open Cisco Smart Software Manager.
7. Click **Inventory**.
8. From the virtual account where you want to register your HX storage cluster, click **General**, and then click **New Token**.
9. In the **Create Registration Token** dialog box, add a short Description for the token, enter the number of days you want the token to be active and available to use on other products, and check **Allow export controlled** functionality on the products registered with this token.
10. Click **Create Token**.

11. From the **New ID Token** row, click the **Actions** drop-down list, and click **Copy**.
12. Log into the **controller VM**.
13. Register your HX storage cluster, where **idtoken-string** is the New ID Token from Cisco Smart Software Manager.

```
# stcli license register --idtoken idtoken-string 12.
```
14. Confirm that your HX storage cluster is registered.

```
# stcli license show summary
```
15. The cluster is now licensed and ready for production deployment.

Enable Syslog

To prevent the loss of diagnostic information when a host fails, ESXi logs should be sent to a central location. Logs can be sent to the VMware vCenter server or to a separate syslog server.

To configure syslog on ESXi hosts, follow these steps:



You can also use a multi-exec tool such as **MobaXterm** or **iTerm2** to simultaneously execute the same command on all servers in the cluster.

1. Log into the ESXi host via SSH as the **root** user.
2. Enter the following commands, replacing the IP address in the first command with the IP address of the vCenter or the syslog server that will receive the syslog logs.

```
[root@hxxv-cl0-esxi-1:~] esxcli system syslog config set --loghost='udp://10.99.167.240'
[root@hxxv-cl0-esxi-1:~] esxcli system syslog reload
[root@hxxv-cl0-esxi-1:~] esxcli network firewall ruleset set -r syslog -e true
[root@hxxv-cl0-esxi-1:~] esxcli network firewall refresh
[root@hxxv-cl0-esxi-1:~]
```

3. Repeat steps 1 and 2 for each HX ESXi host.

Manage Cluster using Cisco Intersight

Cisco Intersight provides a centralized dashboard with a single view of all Cisco UCS Domains, HyperFlex clusters and servers regardless of their location. The dashboard elements can be drilled down to get an overview of their health statuses, storage utilization, port counts, and more. For a standard HyperFlex cluster, Cisco Intersight can be used to do the initial install of a cluster as well. New features and capabilities are continually being added over time. Please see the [Cisco Intersight](#) website for the latest information.

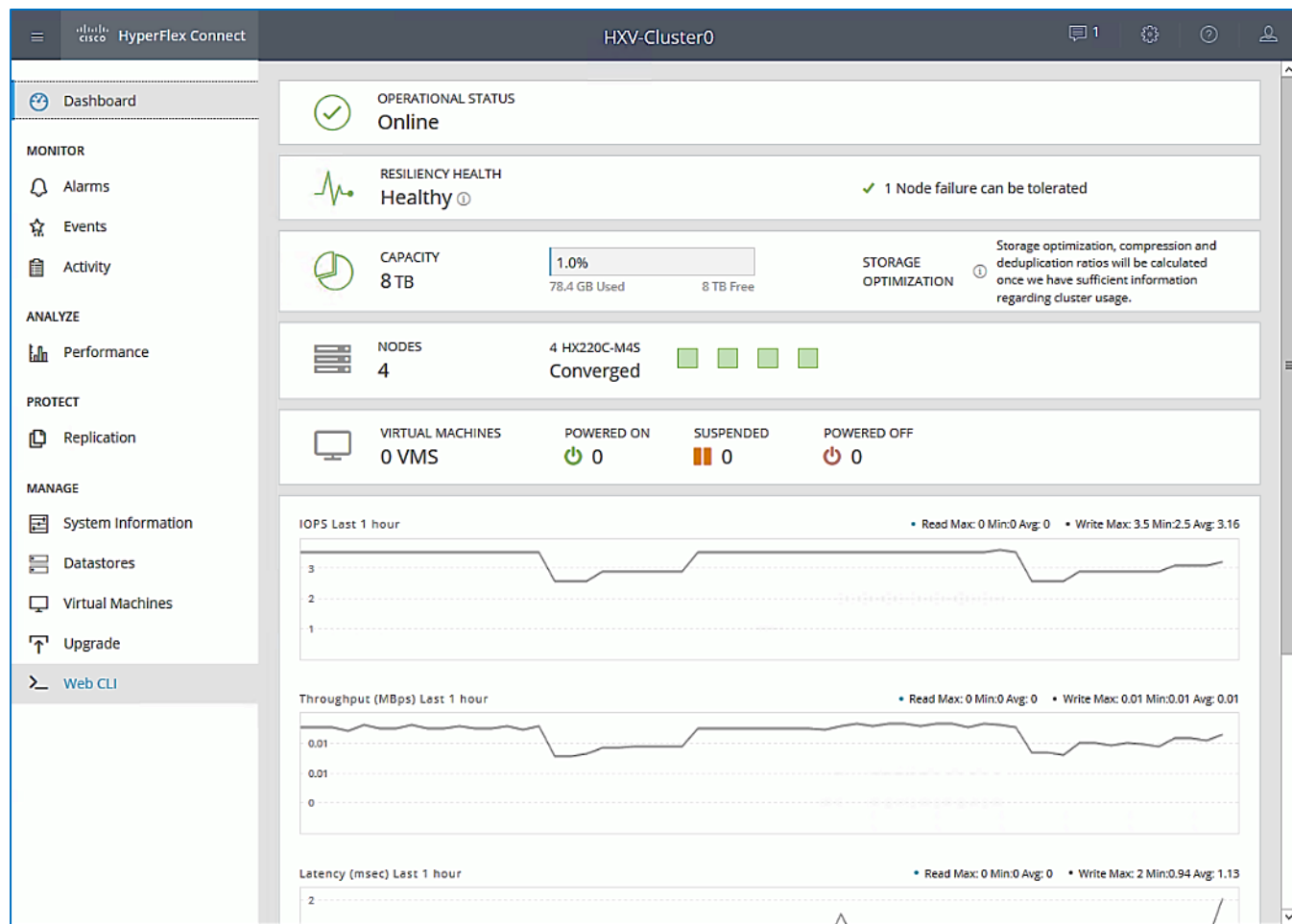
Follow the steps outlined in the [Enable Cisco Intersight Cloud-Based Management](#) section to manage the HyperFlex Cluster from Cisco Intersight.

Manage Cluster using HyperFlex Connect

HyperFlex Connect is an easy to use, powerful primary management tool for managing HyperFlex clusters. HyperFlex Connect is a HTML5 web-based GUI tool that is accessible via the cluster management IP address. It runs on all HX nodes in the cluster for high availability. HyperFlex Connect can be accessed using either pre-defined Local accounts or Role-Based access (RBAC) by integrating authentication with VMware vCenter managing the HyperFlex cluster. With RBAC, you can use VMware credentials either local (for example, administrator@vsphere.local) or Single Sign-On (SSO) credential such as an Active Directory(AD) users defined on vCenter through AD integration.

To manage HyperFlex cluster using HyperFlex Connect, follow these steps:

1. Open a web browser and navigate to the IP address of the HX cluster (for example, <https://10.1.167.100>). Log in using the **admin** account. Log in using the **admin** account. Password should be same as the one specified for the Storage Controller VM during the installation process.



2. The **Dashboard** provides general information about the cluster's operational status, health, Node failure tolerance, Storage Performance and Capacity Details and Cluster Size and individual Node health.

(Optional) Manage Cluster using VMware vCenter (via Plugin)

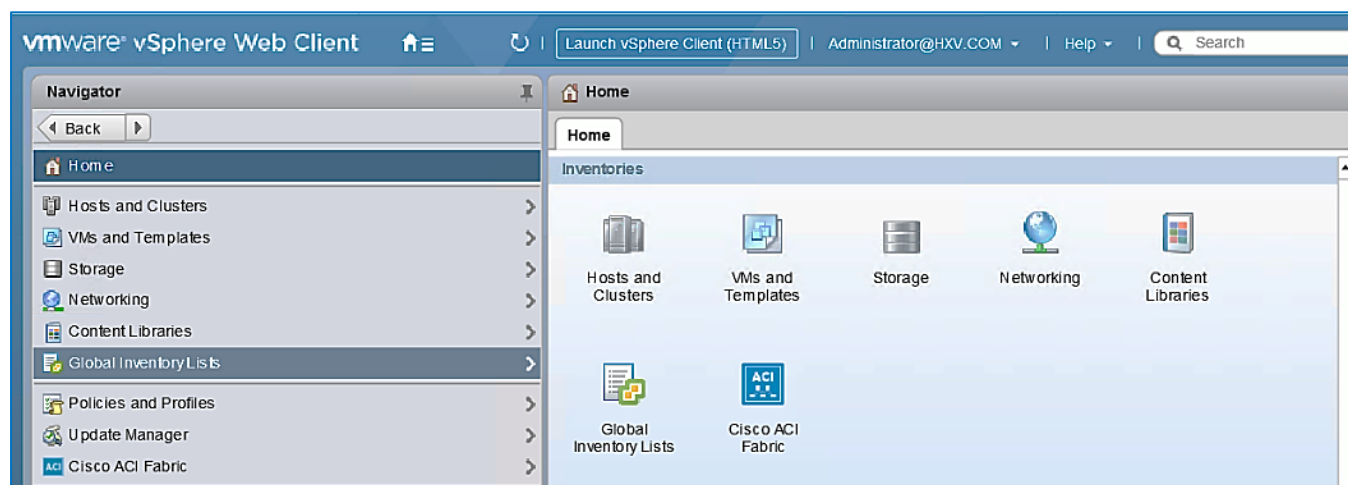
The Cisco HyperFlex vCenter Web Client Plugin can be deployed as a secondary tool to monitor and configure the HyperFlex cluster. The plugin is installed on the specified vCenter server by the HyperFlex installer. The plugin is accessible from vCenter Flash Web Client.



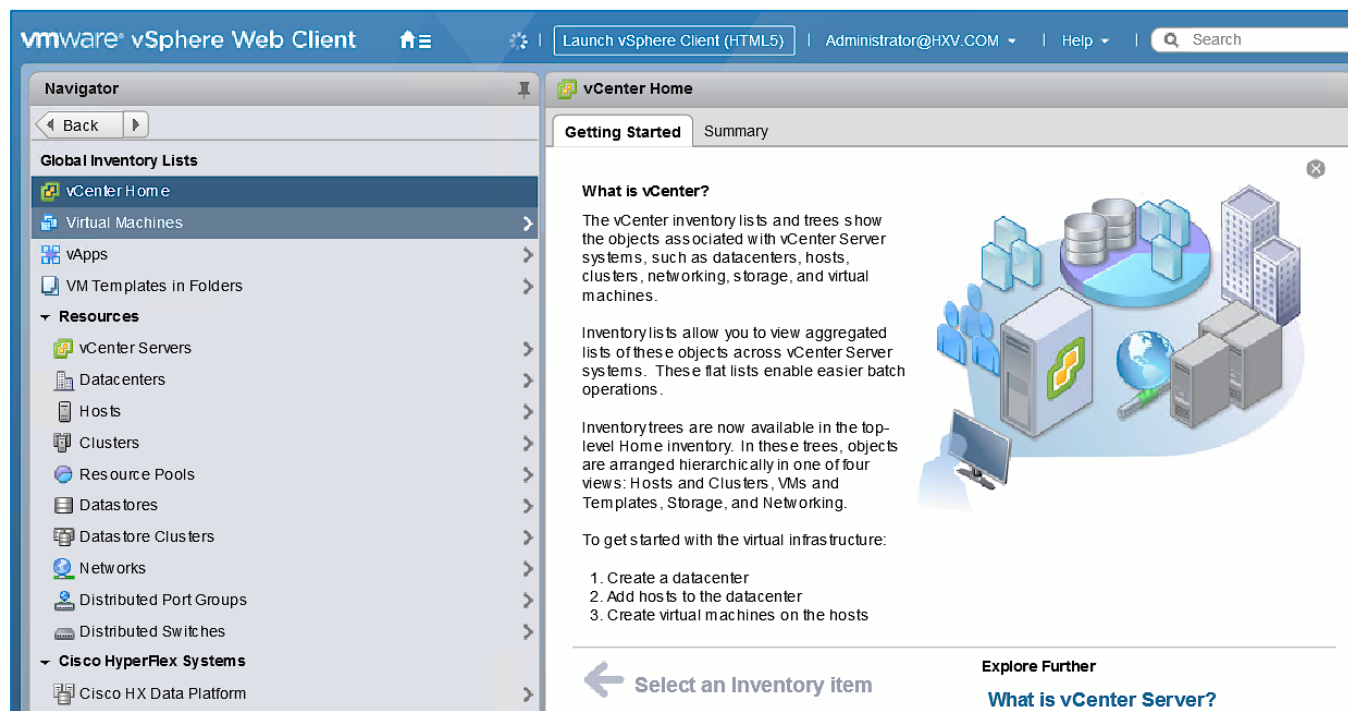
This plugin is not supported in the HTML5 based VMware vSphere Client for vCenter.

To manage the HyperFlex cluster using the vCenter Web Client Plugin for vCenter 6.5, follow these steps:

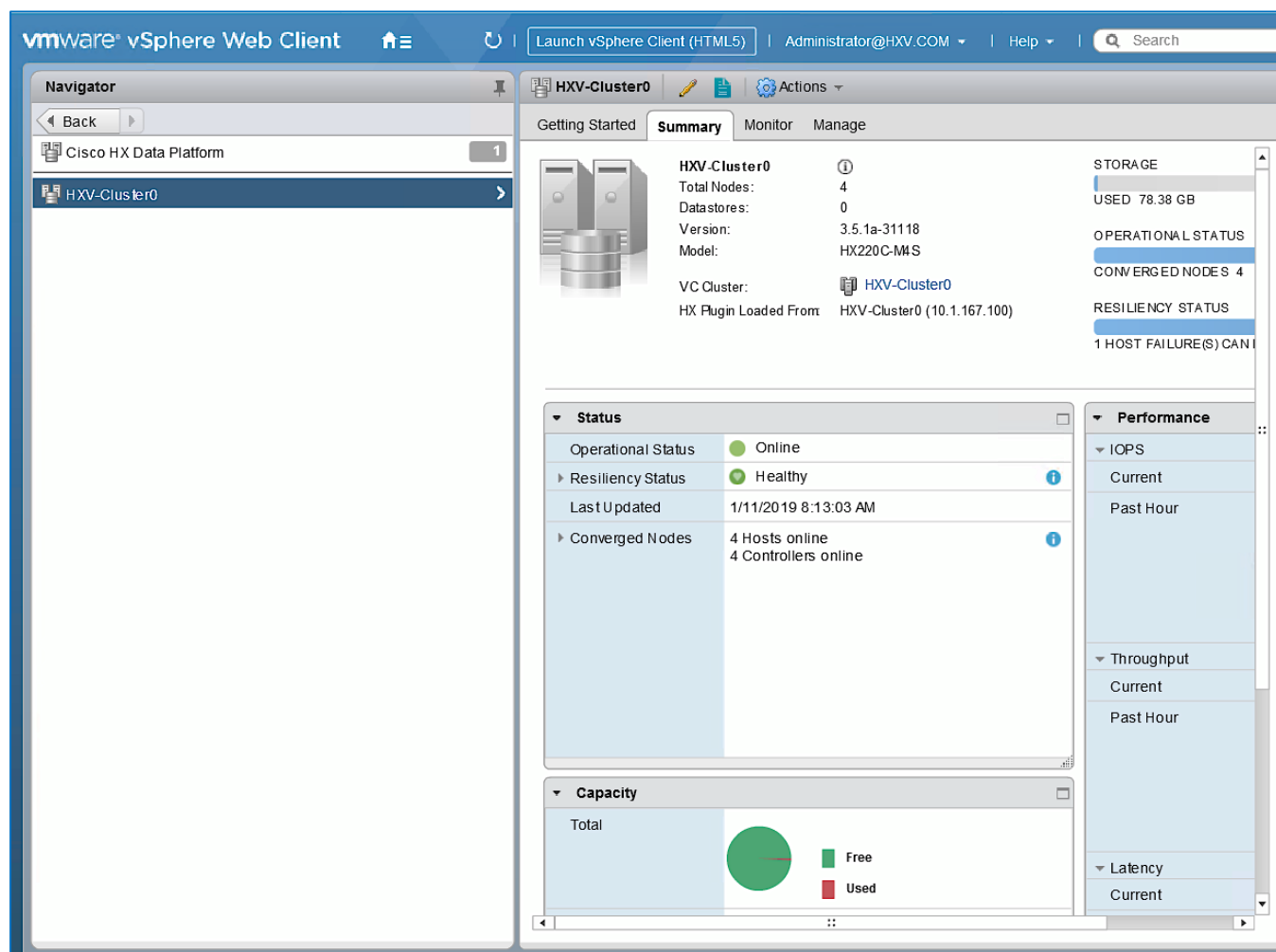
1. Use a browser to navigate and VMware vCenter Web Client. Log in using an **administrator** account.
2. Navigate to the **Home** screen and click **Global Inventory Lists**.



3. In the left navigation pane, click **Cisco HX Data Platform**.



4. In the left navigation pane, click the newly deployed HX cluster (HXV-Cluster0) to manage.



5. Use the **Summary**, **Monitor** or **Manage** tabs in the right-window pane to monitor and manage the cluster status, storage performance and capacity status, create datastores, upgrade cluster and more.

Enable/Disable Auto-Support and Notifications

Auto-Support is enabled if specified during the HyperFlex installation. Auto-Support enables Call Home to automatically send support information to Cisco TAC, and notifications of tickets to the email address specified. If the settings need to be modified, they can be changed in the HyperFlex Connect HTML management webpage.

To change Auto-Support settings, follow these steps:

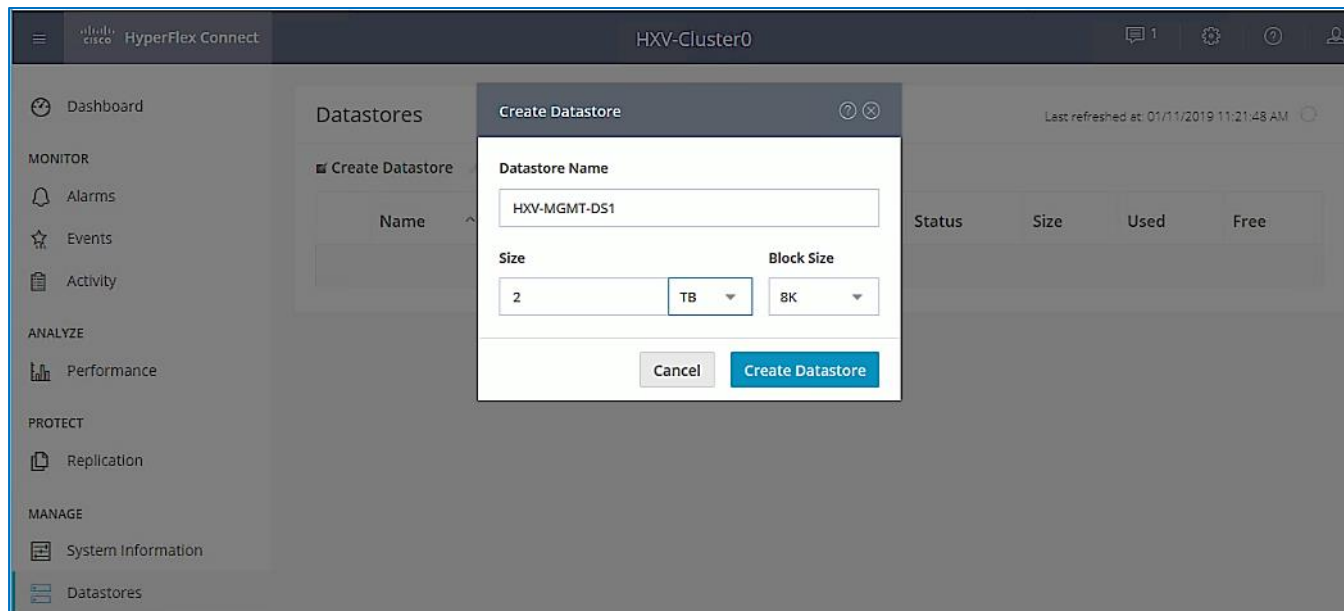
1. Use a browser to navigate to **HyperFlex Connect** using the Management IP of the HX Cluster.
2. Log in using the **admin** account.
3. Click the gear shaped icon in the upper right-hand corner and click **Auto-Support Settings**.
4. Enable or Disable **Auto-Support** as needed. Enter the **email address** to receive notifications for Auto-Support events.
5. Enable or Disable **Remote Support** as needed. Remote support allows Cisco TAC to connect to the HX cluster and accelerate troubleshooting efforts.
6. If a web proxy is used, specify the settings for web proxy. Click **OK**.
7. To enable **Email Notifications**, click the gear shaped icon in top right corner, and click **Notifications Settings**. Enter the outgoing **Mail Server Address** information, the **From Address** and the **Recipient List**. Click **OK**.

Create Datastores for Virtual Machines

This task can be completed by using the vSphere Web Client HX plugin, or by using the HyperFlex Connect HTML management webpage.

To configure a new datastore from HyperFlex Connect, follow these steps:

1. Use a browser to navigate to **HyperFlex Connect** using the Management IP of the HX Cluster.
2. Enter **Login credentials**, either a local credential, or a vCenter RBAC credential with administrative rights. Click **Login**.
3. From the left navigation menu, select **Manage > Datastores**. Click the **Create Datastore** icon at the top.
4. In the **Create Datastore** pop-up window, specify a **Name** and **Size** for the datastore.



5. Click Create Datastore.

Migrate Virtual Networking to VMware vDS on HyperFlex Management Cluster

This section deploys the virtual networking for the virtual machines deployed in the Management cluster. APIC manages the virtual networking for the Management cluster through the VMM integration with VMware vCenter that manages the Management cluster. In this design, the Management uses VMware vDS as the virtual switch for the VM networks though a Cisco AVE could also be used for this. The other networks (Inband Management, Storage Data and vMotion networks) in the Management HyperFlex cluster will remain on the VMware vSwitch as deployed by the HyperFlex Installer. The vCenter that manages the Management HyperFlex cluster is outside the ACI fabric and reachable through the Shared L3Out from the ACI fabric to the existing infrastructure.

Setup Information

The setup information for migrating the default virtual networking from VMware vSwitch to vDS is provided below.

- VLAN Name: HXV0-VMM_VLANS
- VLAN Pool: 1018-1028
- Virtual Switch Name: HXV0-vDS

- Associated Attachable Entity Profile: `HXV-UCS_AAEP`
- VMware vCenter Credentials: Username/Password for the vCenter managing the VMM domain
- VMware vCenter Credentials – Profile Name: `Administrator`
- VMware vCenter Managing the VMM Domain: `hxv-vcsa-0.hxv.com (10.99.167.240)`
- DVS Version: `vCenter Default`
- VMware vCenter Datacenter: `HXV-MGMT`
- Default vSwitch for virtual machine networks: `vswitch-hx-vm-network`
- Uplinks on Default vSwitch for virtual machine Networks: `vmnic2, vmnic6`

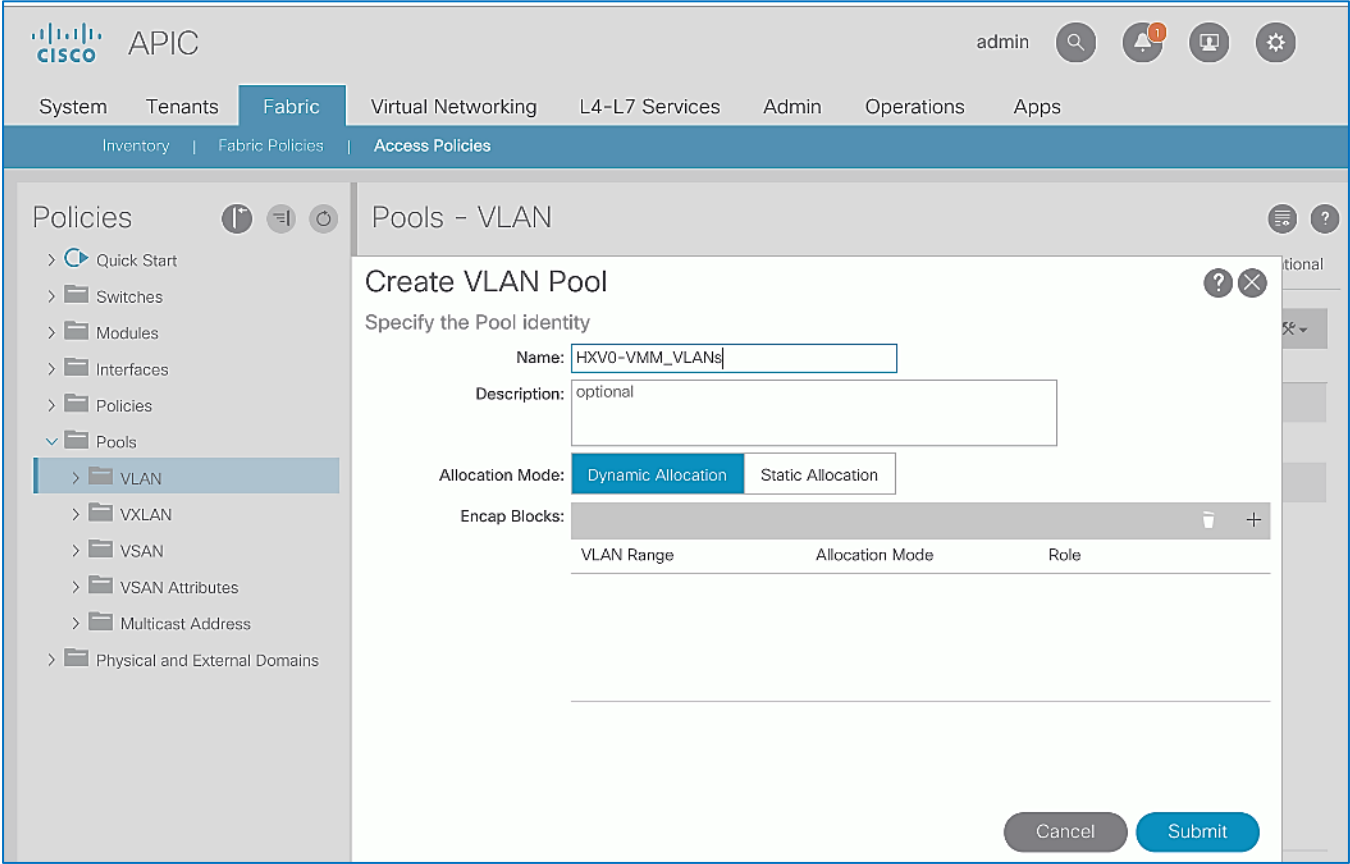
Deployment Steps

To enable APIC-controlled virtual networking for the Management cluster, complete the steps outlined in this section.

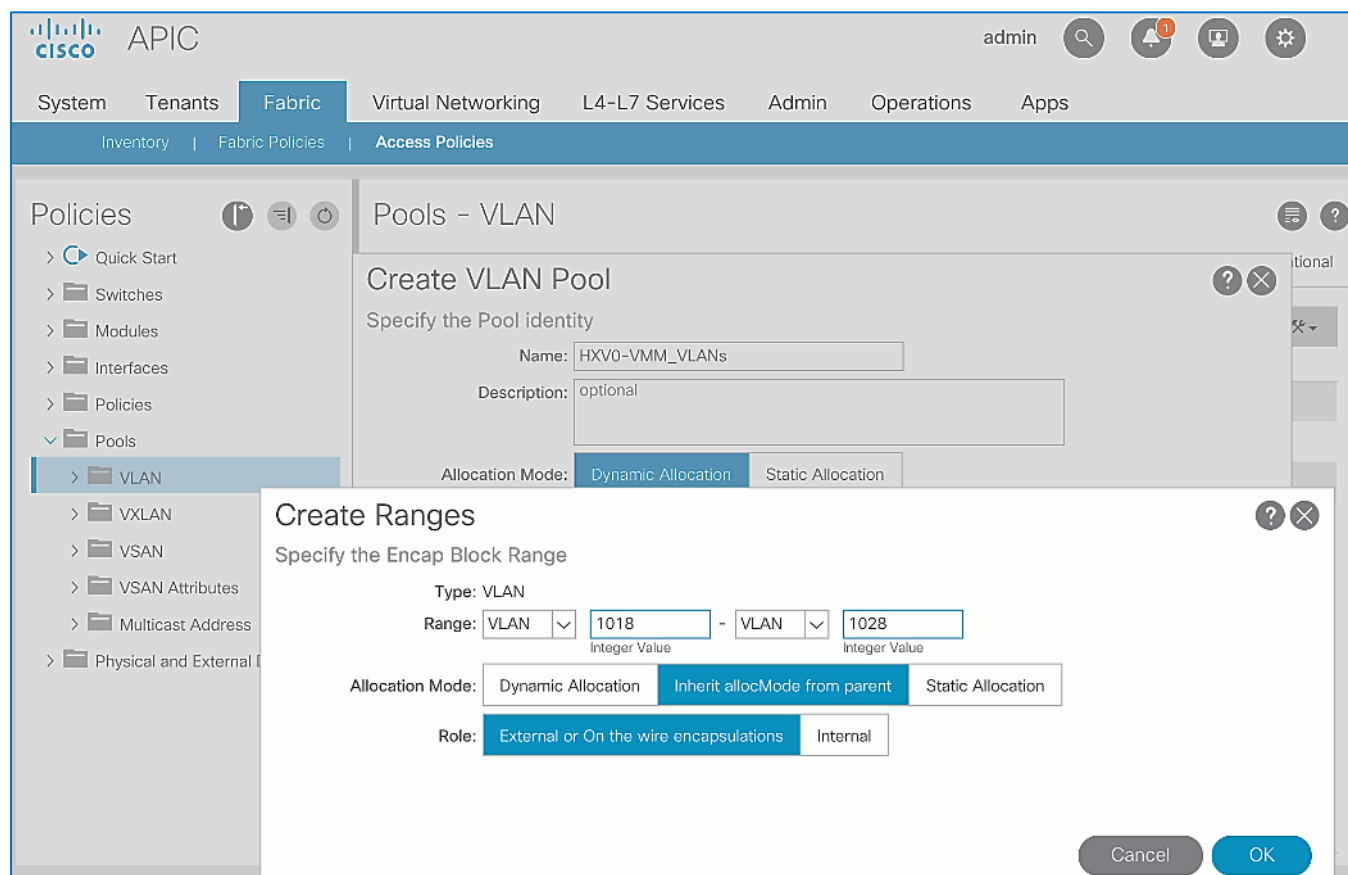
Create VLAN Pool for VMM Domain

To configure VLAN pools for the Management cluster VMM domain, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Fabric > Access Policies**.
3. From the left navigation pane, expand and select **Pools > VLAN**.
4. Right-click **VLAN** and select **Create VLAN Pool**.
5. In the **Create VLAN Pool** pop-up window, specify a **Name** for the pool to be associated with vDS. For **Allocation Mode**, select **Dynamic Allocation**.



- 6. For **Encap Blocks**, click the **[+]** icon on the right side to specify a VLAN range.
- 7. In the **Create Ranges** pop-up window, specify a **VLAN range** for the pool. Leave the other parameters as is.



8. Click **OK** to close the **Create Ranges** pop-up window.
9. Click **Submit** to complete.

Enable VMM Integration for the Management HX Cluster

To enable VMM integration for the Management HyperFlex cluster, follow these steps:

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Virtual Networking**.
3. From the left navigation pane, select **Quick Start**.
4. From the right-window pane, click (VMware hypervisor) Create a vCenter Domain Profile.
5. In the **Create vCenter Domain** window, specify a **Virtual Switch Name** (for example, HXV0-vDS). For the **Virtual Switch**, leave **VMware vSphere Distributed Switch** selected. For the **Associated Attachable Entity Profile**, select the AAEP for the UCS domain (for example, HXV-UCS_AAEP) that the VMM domain is hosted on. For **VLAN Pool**, select the VLAN pool (for example, HXV0-VMM_VLANS) associated with this VMM domain from the drop-down list. Leave all other fields as shown below.

Create vCenter Domain

Specify vCenter domain users and controllers

Virtual Switch Name: HXV0-vDS

Virtual Switch: VMware vSphere Distributed Switch Cisco AVS Cisco AVE

Associated Attachable Entity Profile: HXV-UCS-AAEP

Delimiter:

Access Mode: Read Only Mode Read Write Mode

Endpoint Retention Time (seconds): 1

VLAN Pool: HXV0-VMM_VLANS(dynamic)

Security Domains:

Name	Description
------	-------------

vCenter Credentials:

Profile Name	Username	Description
--------------	----------	-------------

vCenter:

Cancel Submit

6. For **vCenter Credentials**, click the **[+]** icon to the right.
7. In the **Create vCenter Domain** pop-up window, specify a **Name** (for example, Administrator) for the account and specify the credentials (**Username**, **Password**) for the vCenter managing the VMM domain on the Management cluster.



The example provided here uses the Administrator account but an APIC account can be created within the vCenter with the minimum set of privileges. For more information, see the [ACI Virtualization Guide](#) on [cisco.com](#).

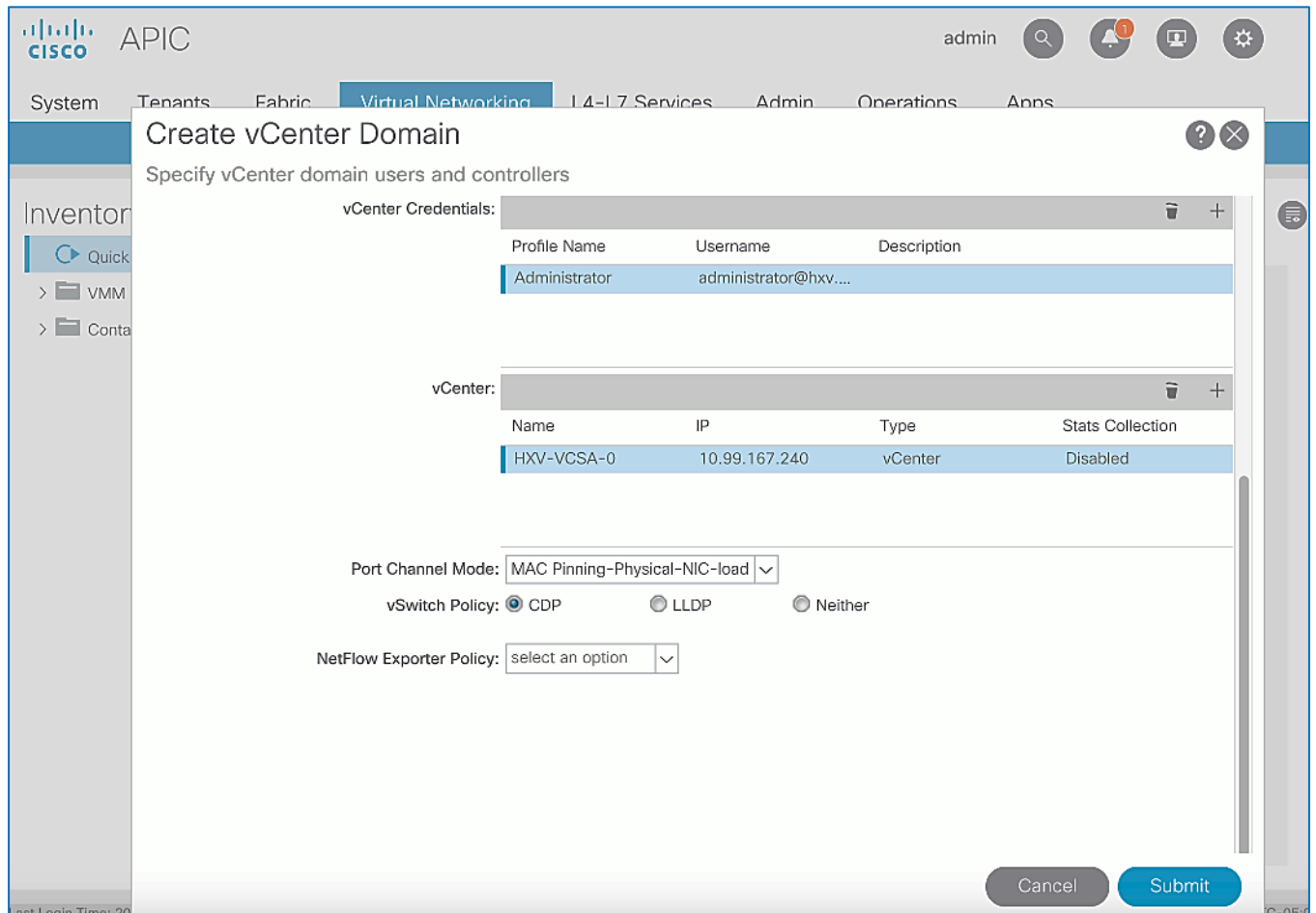
8. Click **OK**.
9. For **vCenter**, click the **[+]** icon on the right.
10. In the **Add vCenter Controller** pop-up window, enter a **Name** for the vCenter. For **IP** address, specify the vCenter IP address. For **DVS Version**, leave it as **vCenter Default**. Set **Stats Collection** to **Enabled**. For **Datacenter**, enter the exact **vCenter Datacenter** name. For **Associated Credential**, select the vCenter credentials created in the last step (Administrator).

The screenshot shows the Cisco APIC interface with the 'Add vCenter Controller' window open. The window title is 'Create vCenter Domain' and the subtitle is 'Specify vCenter domain users and controllers'. The left sidebar shows the 'Inventory' tree with 'VMware' selected. The main form area is titled 'Add vCenter Controller' and 'Specify controller profile'. The form fields are as follows:

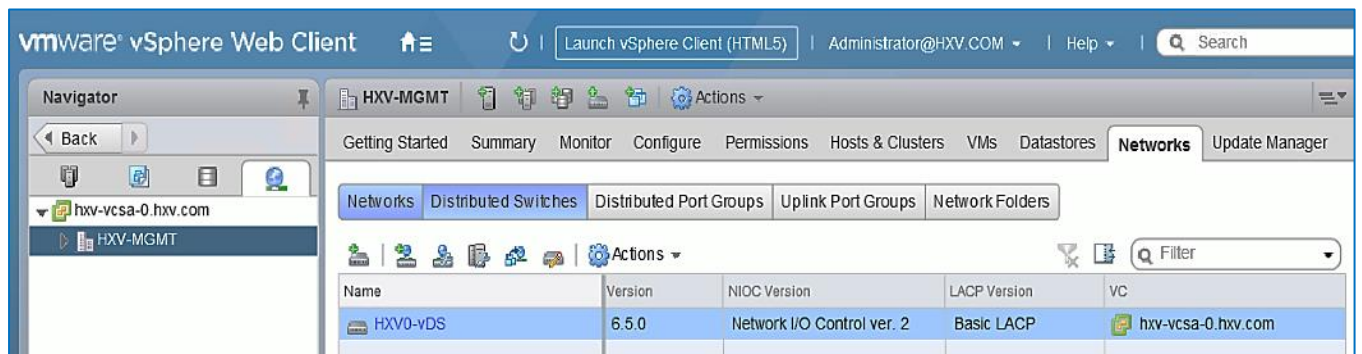
Field	Value
Name	HXV-VCSA-0
Host Name (or IP Address)	10.99.167.240
DVS Version	vCenter Default
Stats Collection	Enabled
Datacenter	HXV-MGMT
Management EPG	select an option
Associated Credential	Administrator

At the bottom right of the window are 'Cancel' and 'OK' buttons.

11. Click **OK**.
12. In the **Create vCenter Domain** Window, select the **MAC Pinning-Physical-NIC-load** as the Port Channel Mode. Select **CDP** for vSwitch Policy.



13. Click **Submit** to create the vDS within the vCenter.
14. Use a browser to navigate to the VMware vCenter server managing the HyperFlex Application cluster. Click the vSphere Web Client of your choice. Log in using an **Administrator** account.
15. Navigate to **Networking**.

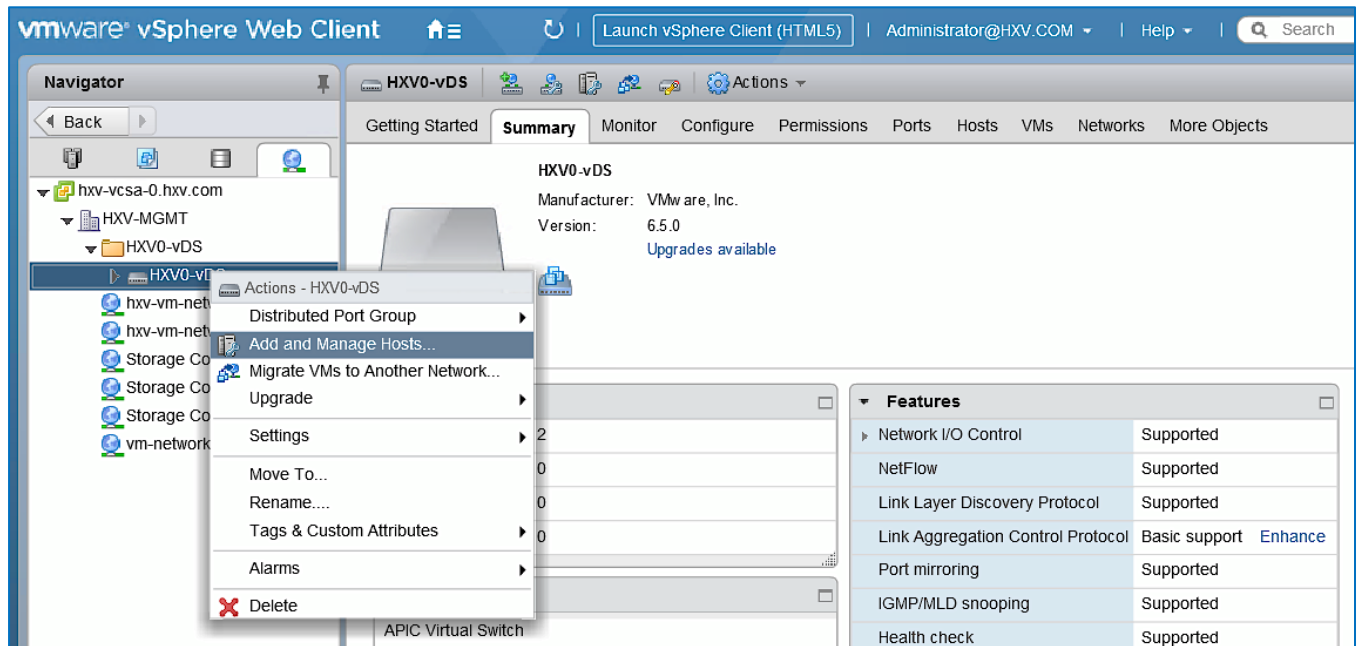


16. Verify that vDS is setup correctly.

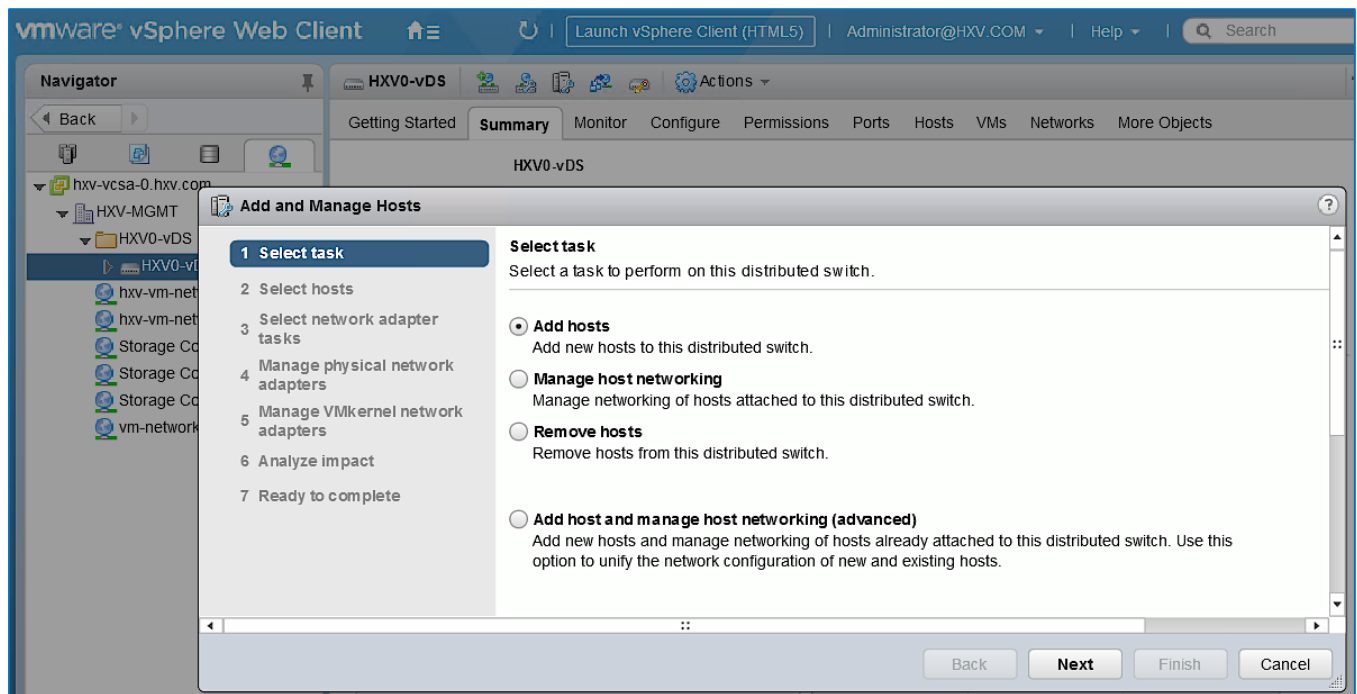
Add HyperFlex ESXi Hosts to VMware vSphere vDS

To add the HyperFlex ESXi Hosts to the newly created vDS, follow these steps:

1. Use a browser to log into the VMware vCenter server managing the HyperFlex Application cluster. Click the vSphere Web Client of your choice. Log in using an **Administrator** account.
2. Navigate to the **Home** screen, select **Networking** in the **Inventories** section.
3. In the left navigation pane, expand the **Datacenter** (for example, **HXV-MGMT**) with the newly deployed vDS (for example, **HXV0-vDS**). Open the vDS folder and select the vDS (for example, **HXV0-vDS**) deployed by the APIC.
4. Right-click the APIC-controlled vDS switch and select **Add and manage hosts**.

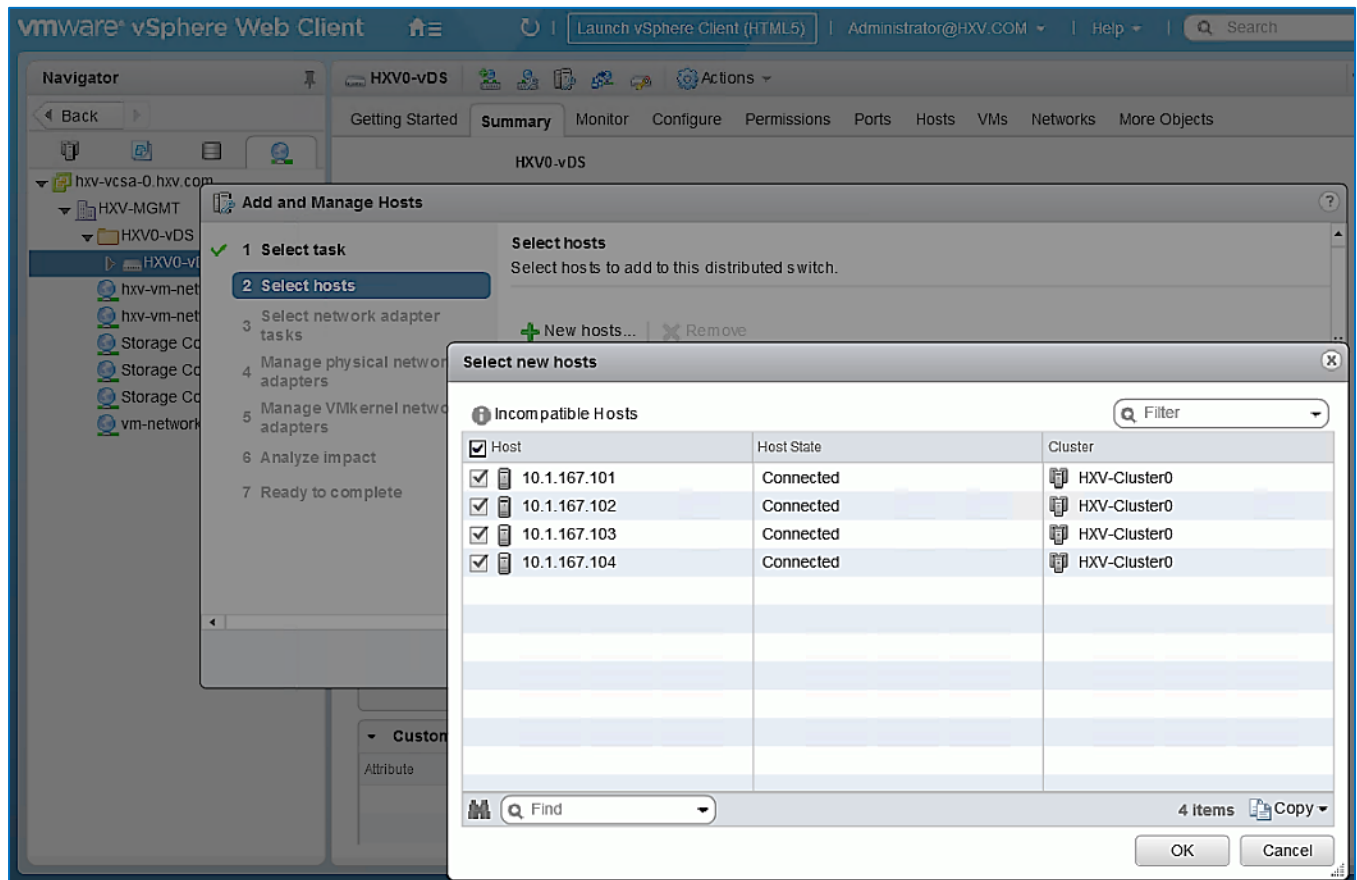


5. In the **Add and Manage Hosts** pop-up window, select the **Add hosts** option. Click **Next**.

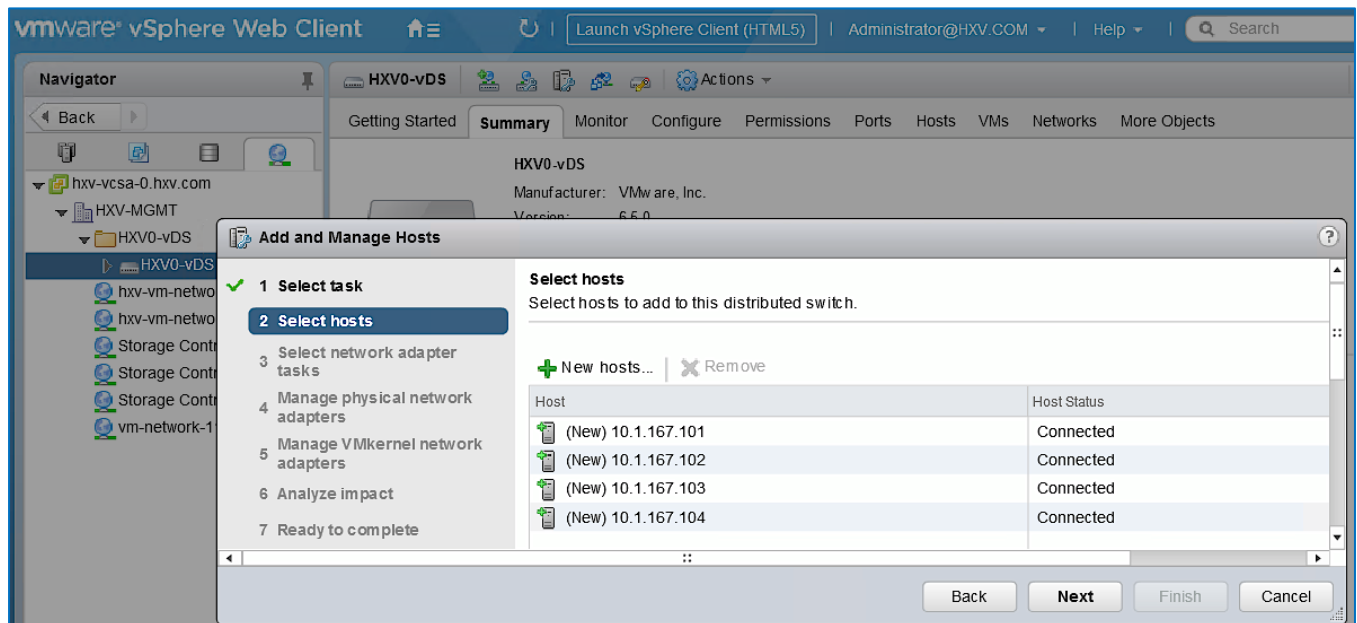


6. In the **Select Hosts** window, click **[+ New host...]** icon at the top to add new host.

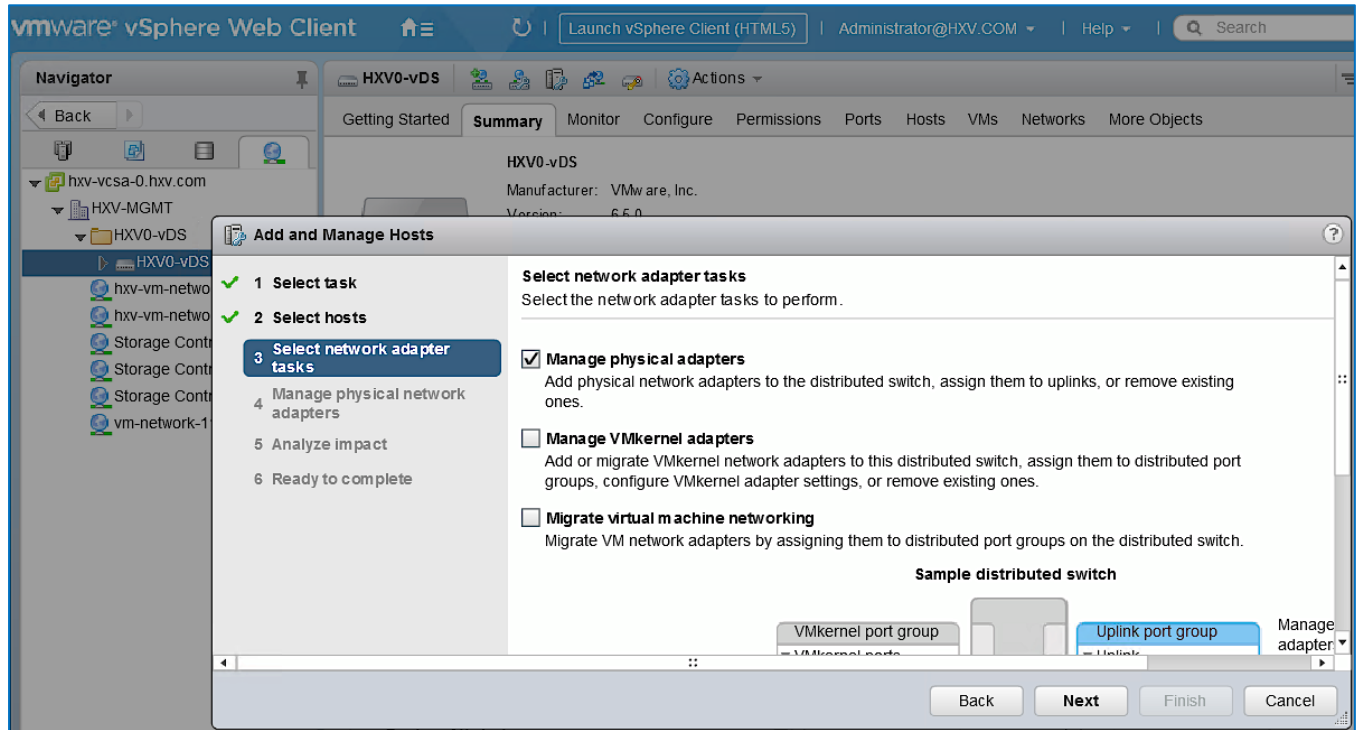
7. In the **Select new hosts** pop-up window, select all hosts in the HX cluster.



8. Click **OK**.

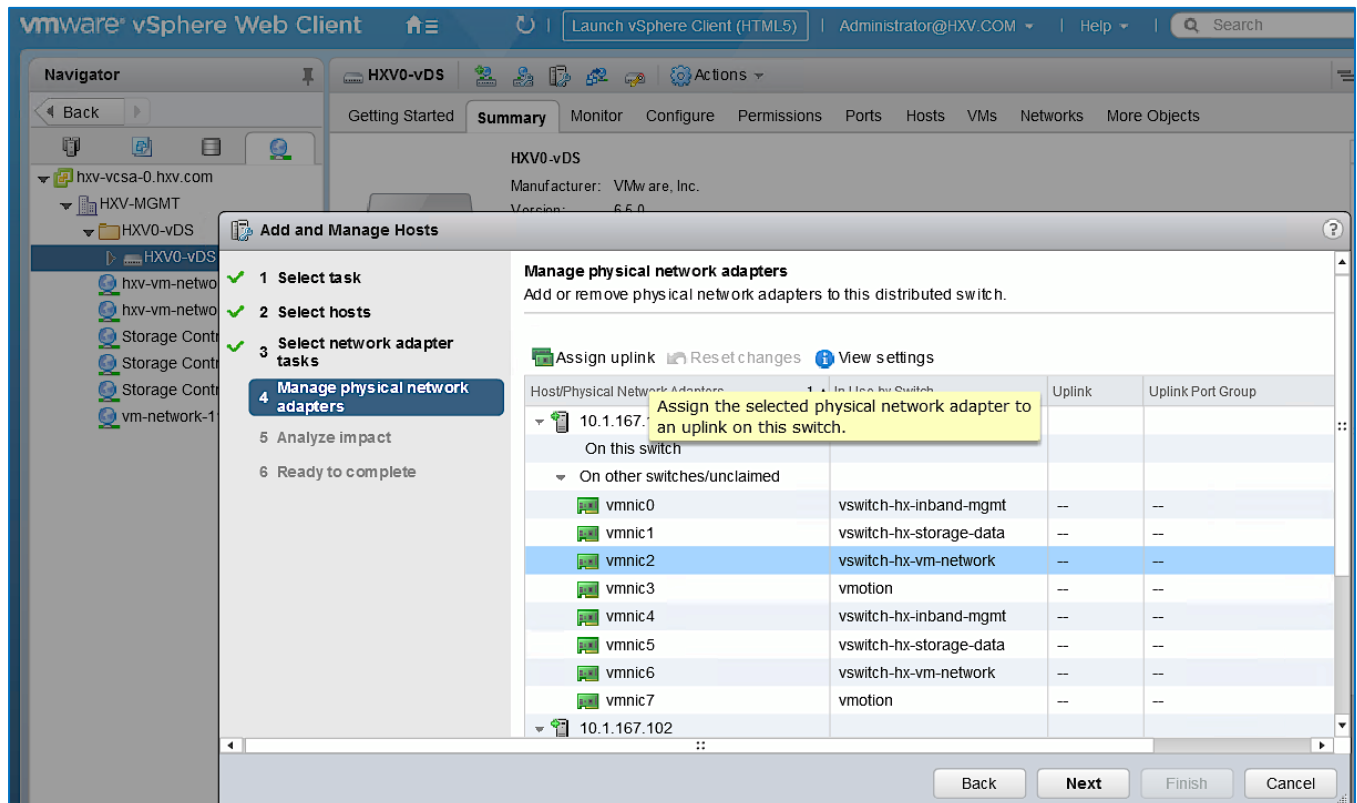


9. Click **Next**.
10. Leave **Manage physical adapters** selected and de-select the other options.

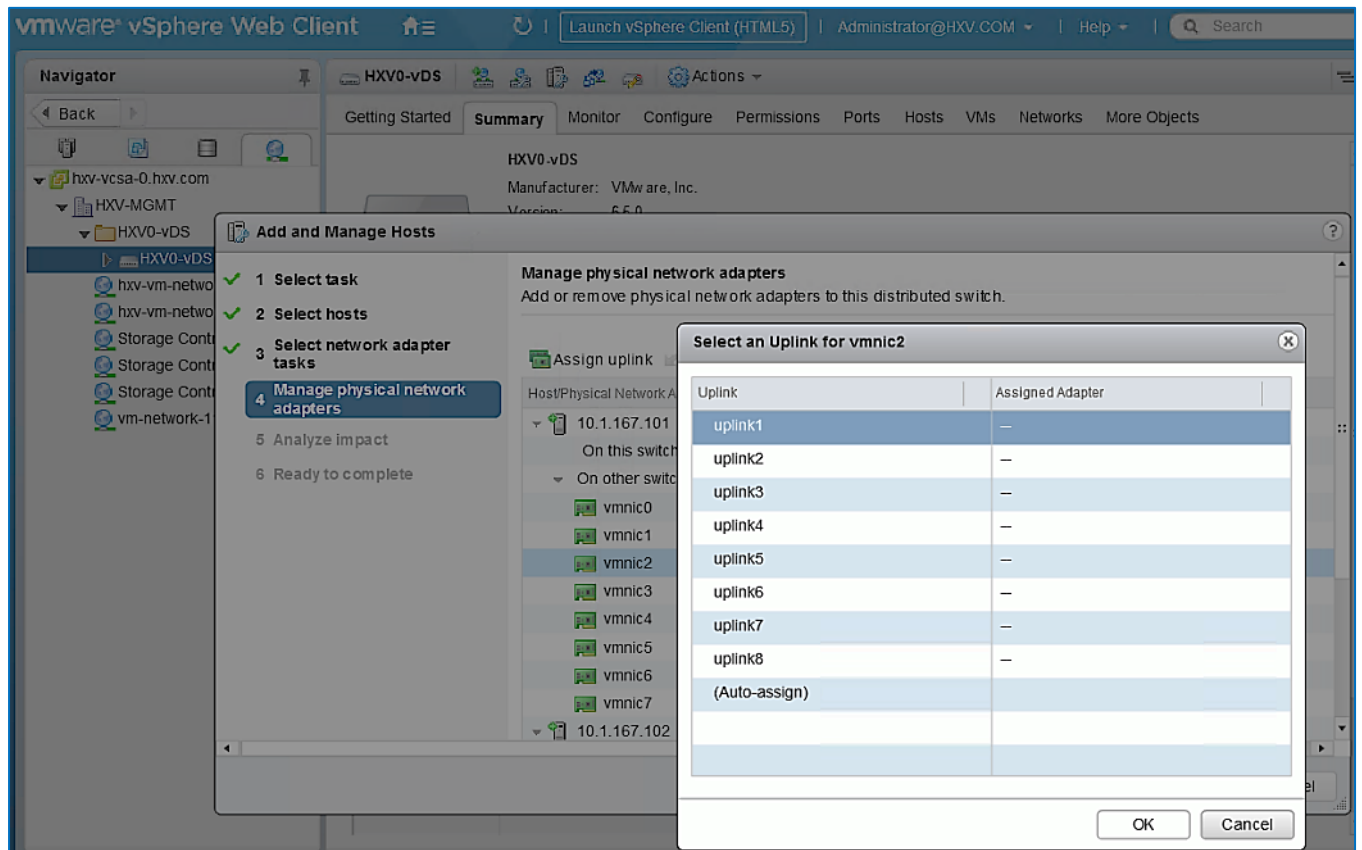


11. Click **Next**.

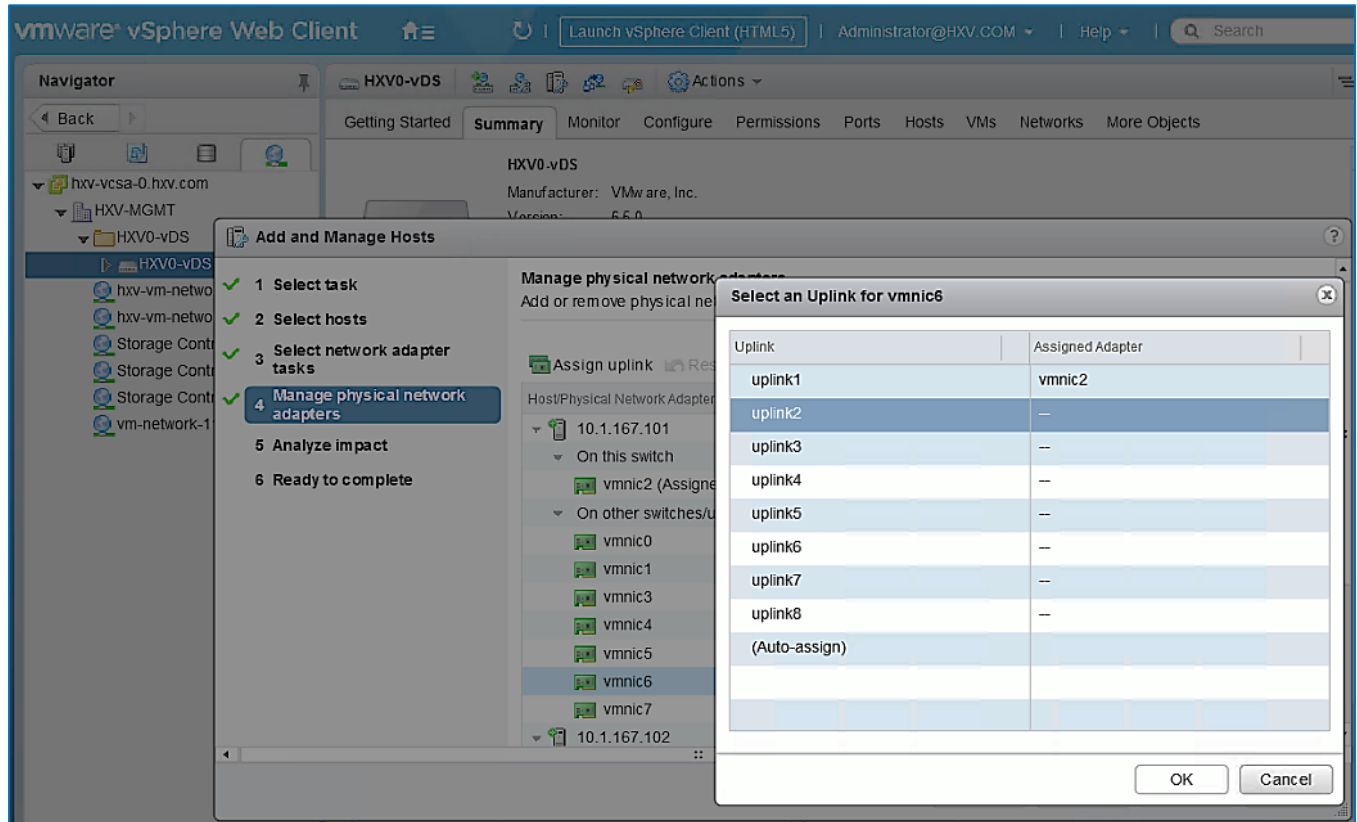
12. In the **Manage physical network adapters** window, for the first host, from the **Host/Physical Network Adapters** column, select the first **vmnic** (for example, **vmnic2**) that currently belongs to the HX VM Network vSwitch (for example, **vswitch-hx-vm-network**). Click the **Assign uplink** icon from the menu.



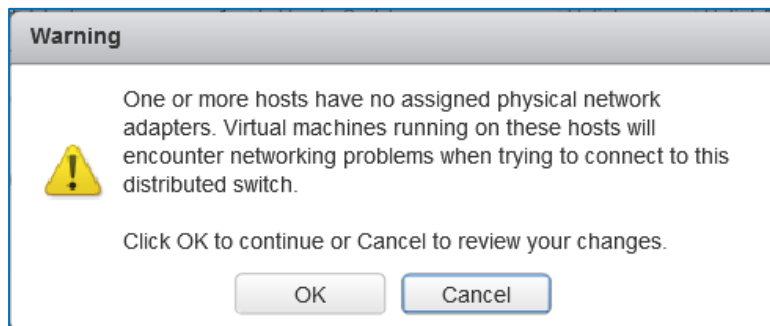
13. In the **Select an Uplink for vmnic** pop-up window, leave **uplink 1** selected.



14. Click **OK**.
15. Repeat steps 1-14 for the second **vmnic** (for example, **vmnic6**) that currently belongs to the HX VM Network vSwitch (for example, **vswitch-hx-vm-network**) – assign it to **uplink2**.



16. Click **OK**.



17. Click **OK** to accept the **Warning**.

18. Repeat steps 1-17 to move uplinks from vSwitch to vDS for all hosts in the cluster. If a server shows no physical adapter available for migration to vDS, exit the wizard. Select the host from left navigation pane and navigate to **Configure > virtual Switches** (under Networking) and select the vSwitch for vm-network (for example, `vswitch-hx-vm-network`) and remove the physical adapters. Once released from the vswitch, the physical adapters for that host can be added to the vDS from the wizard.

19. Click **Next**.

20. In the **Analyze impact** window, click **Next**.

21. Review the settings and click **Finish** to apply.

The management HX cluster is now ready for deploying virtual machines. As EPGs are setup in APIC, the virtual networking will also be setup.

Deploy Virtual Machines – Infrastructure Management

In this design, the Management HyperFlex cluster hosts the infrastructure management virtual machines that manage other virtual server infrastructure on same ACI Multi-Pod fabric. For example, the HyperFlex Installer virtual machine for installing additional HyperFlex clusters in the ACI Fabric and VMware vCenter Appliance(s) are two of the infrastructure services hosted on the Management cluster. The HyperFlex Installer VM will deploy the Applications cluster or the HyperFlex stretched cluster and VMware vCenter will manage the Applications cluster.

The high-level steps for deploying the virtual machines on a HyperFlex cluster connected to a Cisco ACI Multi-Pod fabric are as follows:

- Add VLAN(s) to ACI Fabric for Infrastructure Management Virtual Machines – this is done by adding the VLANs to the VLAN Pool associated with the access layer connection to the Infrastructure Management virtual machines. Ideally, a pool of VLANs should be pre-defined for use by different types of infrastructure and management services rather than adding VLANs one at a time. In this design, VMM integration is enabled between the APIC and the vCenter managing the cluster to dynamically allocate and configure the virtual networking for infrastructure and management virtual machines. The VLAN Pool for use by VMM domain was completed in the [Migrate Virtual Networking on HyperFlex Management Cluster to VMware vDS](#) section. Additional VLANs can be added to the VMM VLAN Pool if needed.
- Define ACI Constructs for Infrastructure Management – this includes specifying the Tenant, VRF, Bridge Domain, Application Profile, EPGs and Contracts so infrastructure virtual machines can be added to the ACI fabric. VMware vCenter and HX Installer virtual machines will be part of the existing `Foundation` Tenant and VRF but a new Application Profile, Bridge Domain and EPG will be created for the HyperFlex Installer and VMware vCenter virtual machines – they can also be deployed in separate EPGs as well. To host additional services such as AD/DNS, Umbrella Virtual Appliances, Monitoring tools etc. new EPGs and Tenants can also be provisioned as needed in the Management cluster.
- Enable contracts to allow communication between Infrastructure EPGs and other components in the network. For example, the Installer virtual machine will need out-of-band management access to Fabric Interconnects and in-band ESXi management access to the HX nodes.
- Deploy the infrastructure virtual machines in the HyperFlex Management cluster.



This section explains the deployment of HyperFlex Installer Virtual Machine in the Management Cluster but not the VMware vCenter install and setup – please refer to the VMware documentation for assistance.

Configure ACI Fabric for Infrastructure Management

This section explains the ACI fabric setup for deploying infrastructure management virtual machines in the Management HyperFlex cluster. The same procedure can be used to bring up other virtual machines on the same cluster.

In this setup, the existing `Foundation` Tenant and VRF used for HyperFlex infrastructure will also be used to host the infrastructure and management virtual machines hosted on the Management cluster. For new Tenants, follow the steps for the `Foundation` Tenant and VRF before doing the configuration in this section.

Create Bridge Domain for Infrastructure Management

To create a Bridge Domain for Infrastructure Management virtual machines in the HyperFlex Management cluster, follow these steps using the setup information provided below:

- **Tenant:** `HXV-Foundation`
- **VRF:** `HXV-Foundation_VRF`

- **Bridge Domain:** HXV-INFRA-MGMT_BD
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 2. From the top navigation menu, select **Tenants > HXV-Foundation**.
 3. From the left navigation pane, expand and select **Tenant HXV-Foundation > Networking > Bridge Domains**.
 4. Right-click Bridge Domains and select **Create Bridge Domain**.
 5. In the **Create Bridge Domain** wizard, for **Name**, specify a name (HXV-INFRA-MGMT_BD) for the bridge domain. For **VRF**, select the previously created VRF (HXV-Foundation_VRF) from the drop-down list. For **Forwarding**, select **Custom** from the drop-down list. For **L2 Unknown Unicast**, select **Flood** from the drop-down list. The checkbox for **ARP Flooding** should now show up and be enabled.

APIC Create Bridge Domain

System | **Tenants** | STEP 1 > Main | 1. Main | 2. L3 Configurations | 3. Advanced/Troubleshooting

ALL TENANTS | Add | Specify Bridge Domain for the VRF

Tenant HXV-F

> Quick Start

> Tenant HXV-Fou

> Application P

> Networking

> Bridge Do

> HXV-C

> HXV-II

> HXV-II

> HXV-S

> HXV-v

> VRFs

> External B

> External R

> Dot1Q Tu

> Contracts

> Policies

> Services

Name: HXV-INFRA-MGMT_BD

Alias:

Description: optional

Tags: enter tags separated by comma

Type: fc regular

Advertise Host Routes: ☐

VRF: HXV-Foundation_VRF

Forwarding: Custom

L2 Unknown Unicast: Flood

L3 Unknown Multicast Flooding: Flood

Multi Destination Flooding: Flood in BD

ARP Flooding: ☒ Enabled

Clear Remote MAC Entries: ☐

Endpoint Retention Policy: select a value
This policy only applies to local L2 L3 and remote L3 entries

IGMP Snoop Policy: select a value

Previous Cancel Next

6. Click **Next**.
7. In the **L3 Configurations** section, for **EP Move Detection Mode**, select the checkbox to enable **GARP based detection** if needed. See [Review/Enable ACI Fabric Settings](#) section for more details on when to enable this feature.
8. Click **Next**. Skip the **Advanced/Troubleshooting** section. Click **Finish** to complete.

Configure Subnet Gateway for Infrastructure Management

To configure a gateway for Infrastructure Management virtual machines, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
- **Bridge Domain:** HXV-INFRA-MGMT_BD

- **BD Subnet:** 10.10.167.254

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
2. From the top navigation menu, select **Tenants > HXV-Foundation**.
3. From the left navigation pane, select and expand **Tenant HXV-Foundation > Networking > Bridge Domains > HXV-INFRA-MGMT_BD**.
4. Right-click **HXV-INFRA-MGMT_BD** and select **Create Subnet**.
5. In the **Create Subnet** pop-up window, for the **Gateway IP**, specify the IP address and mask for the gateway. For **Scope**, select **Advertised Externally** and **Shared between VRFs**. Leave everything else as is.

The screenshot shows the APIC GUI with the 'Create Subnet' dialog open. The left sidebar shows the navigation tree with 'HXV-INFRA-MGMT_BD' selected. The dialog box contains the following fields and options:

- Gateway IP:** 10.10.167.254/24 (address/mask)
- Treat as virtual IP address:** ☐
- Make this IP address primary:** ☐
- Scope:**
 - ☐ Private to VRF
 - ☒ Advertised Externally
 - ☒ Shared between VRFs
- Description:** optional
- Subnet Control:**
 - ☐ No Default SVI Gateway
 - ☐ Querier IP
- L3 Out for Route Profile:** select a value
- Route Profile:** select a value
- ND RA Prefix policy:** select a value

Buttons: Cancel, Submit

6. Click **Submit**.

Create Application Profile for In-Band Management

To create an application profile for Infrastructure Management virtual machines in the HyperFlex Management cluster, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
- **Application Profile:** HXV-INFRA-MGMT_AP

1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.

- From the top navigation menu, select **Tenants** > HXV-Foundation.
- From the left navigation pane, select **Tenant** HXV-Foundation.
- Right-click **Tenant** HXV-Foundation and select **Create Application Profile**.
- In the **Create Application Profile** pop-up window, for **Name** (HXV-INFRA-MGMT_AP), specify a name for the Application Profile.

Create Application Profile

Specify Tenant Application Profile

Name: HXV-INFRA-MGMT_AP

Alias:

Description: optional

Tags:
 enter tags separated by comma

Monitoring Policy: select a value

EPGs

Name	Alias	BD	Domain	Switching Mode	Static Path	Static Path VLAN	Provided Contract	Consumed Contract

Cancel Submit

- Click **Submit** to complete

Create EPG for Infrastructure Management and Associate with Bridge Domain

To create an EPG for Infrastructure Management virtual machines in the HyperFlex Management cluster, follow these steps using the setup information provided below:

- Tenant:** HXV-Foundation
 - Application Profile:** HXV-INFRA-MGMT_AP
 - Bridge Domain:** HXV-INFRA-MGMT_BD
 - EPG:** HXV-INFRA-MGMT_EPG
- Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 - From the top navigation menu, select **Tenants** > HXV-Foundation.
 - From the left navigation pane, select and expand **Tenant** HXV-Foundation > **Application Profiles** > HXV-INFRA-MGMT_AP.

4. Right-click HXV-INFRA-MGMT_AP and select **Create Application EPG**.
5. In the **Create Application EPG** pop-up window, for **Name**, specify a name (HXV-INFRA-MGMT_EPG) for the EPG. For **Bridge Domain**, select the previously created Bridge Domain (HXV-INFRA-MGMT_BD).

The screenshot shows the 'Create Application EPG' window in the Cisco APIC GUI. The left sidebar shows the navigation tree with 'Tenant HXV-Foundation' selected. The main area is titled 'STEP 1 > Identity' and contains the following fields and options:

- Alias:** (empty text field)
- Description:** optional (text area)
- Tags:** (dropdown menu, hint: enter tags separated by comma)
- Contract Exception Tag:** (empty text field)
- QoS class:** Unspecified (dropdown menu)
- Custom QoS:** select a value (dropdown menu)
- Data-Plane Policer:** select a value (dropdown menu)
- Intra EPG Isolation:** Enforced / Unenforced (radio buttons, Unenforced is selected)
- Preferred Group Member:** Exclude / Include (radio buttons, Exclude is selected)
- Flood on Encapsulation:** Disabled / Enabled (radio buttons, Disabled is selected)
- Bridge Domain:** HXV-INFRA-MGMT_BD (dropdown menu, HXV-INFRA-MGMT_BD is selected)
- Monitoring Policy:** select a value (dropdown menu)
- FHS Trust Control Policy:** select a value (dropdown menu)
- Shutdown EPG:** ☐
- Associate to VM Domain Profiles:** ☐
- Statically Link with Leaves/Paths:** ☐
- EPG Contract Master:** (empty text field)
- Application EPGs:** (empty list)

At the bottom right, there are three buttons: 'Previous', 'Cancel', and 'Finish'.

6. Click **Finish**.

Associate EPG with VMM Domain – Dynamic Binding

To associate the Infrastructure Management EPG with the VMM Domain, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
 - **Application Profile:** HXV-INFRA-MGMT_AP
 - **Bridge Domain:** HXV-INFRA-MGMT_BD
 - **EPG:** HXV-INFRA-MGMT_EPG
 - **Domain:** HXV0-vDS
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 2. From the top navigation menu, select **Tenants > HXV-Foundation**.

- From the left navigation pane, select and expand **Tenant HXV-Foundation** > **Application Profiles** > HXV-INFRA-MGMT_AP > **Application EPGs** > HXV-INFRA-MGMT_EPG.
- Right-click HXV-INFRA-MGMT_EPG and select **Add VMM Domain Association**.
- In the **Add VMM Domain Association** pop-up window, for the **VMM Domain Profile**, select the previously created VMM Domain from the list. For **Deploy Immediacy**, select **Immediate**. For **Resolution Immediacy**, select **Immediate**.

APIC

System | **Tenants** | Fabric | Virtual Networking | L4-L7 Services | Admin | Operations | Apps

ALL TENANTS | Add Tenant | Tenant Search: name or descr | common | **HXV-Foundation** | infra | mgmt

Tenant HXV-Foundation | EPG - HXV-INFRA-MGMT_EPG

Add VMM Domain Association

Choose the VMM domain to associate

VMM Domain Profile: HXV0-vDS

Deploy Immediacy: **Immediate** | On Demand

Resolution Immediacy: **Immediate** | On Demand | Pre-provision

Delimiter:

Enhanced Lag Policy: select an option

Allow Micro-Segmentation: ☐

VLAN Mode: **Dynamic** | Static

Port Binding: Dynamic Binding | Ephemeral | **Default** | Static Binding

Netflow: **Disable** | Enable

Allow Promiscuous: Reject

Forged Transmits: Reject

MAC Changes: Reject

Cancel Submit

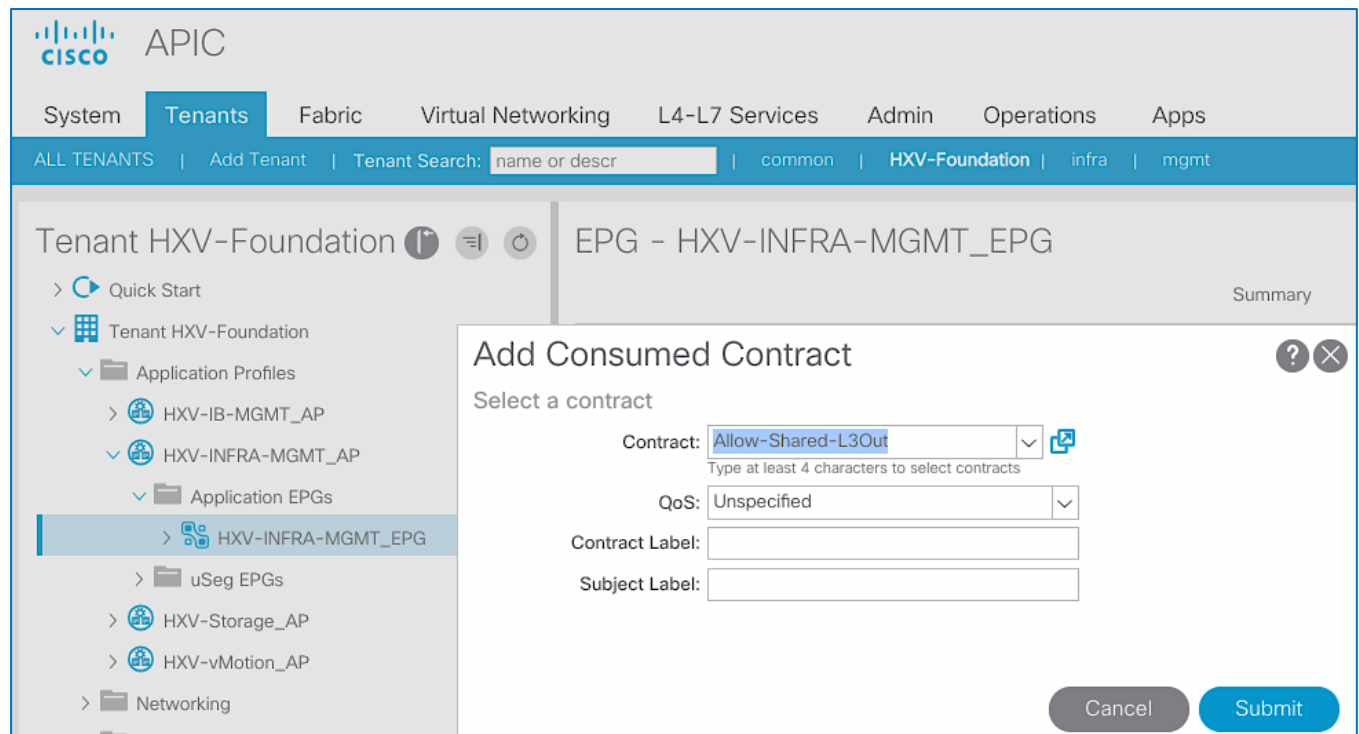
- Click **Submit**.

Enable Contract to Access Outside Networks via Shared L3Out

To access the network and services reachable via the **Shared L3Out** in the **common** Tenant, follow these steps using the setup information provided below:

- Tenant:** HXV-Foundation
- Application Profile:** HXV-INFRA-MGMT_AP

- **EPG:** HXV-INFRA-MGMT_EPG
 - **Consumed Contract:** Allow-Shared-L3Out
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 2. From the top navigation menu, select **Tenants** > HXV-Foundation.
 3. From the left navigation pane, select and expand **Tenant** HXV-Foundation > **Application Profiles** > HXV-INFRA-MGMT_AP > **Application EPGs** > HXV-INFRA-MGMT_EPG.
 4. Right-click HXV-INFRA-MGMT_EPG and select **Add Consumed Contract**.
 5. In the **Add Consumed Contract** pop-up window, select the Allow-Shared-L3Out contract from the drop-down list.



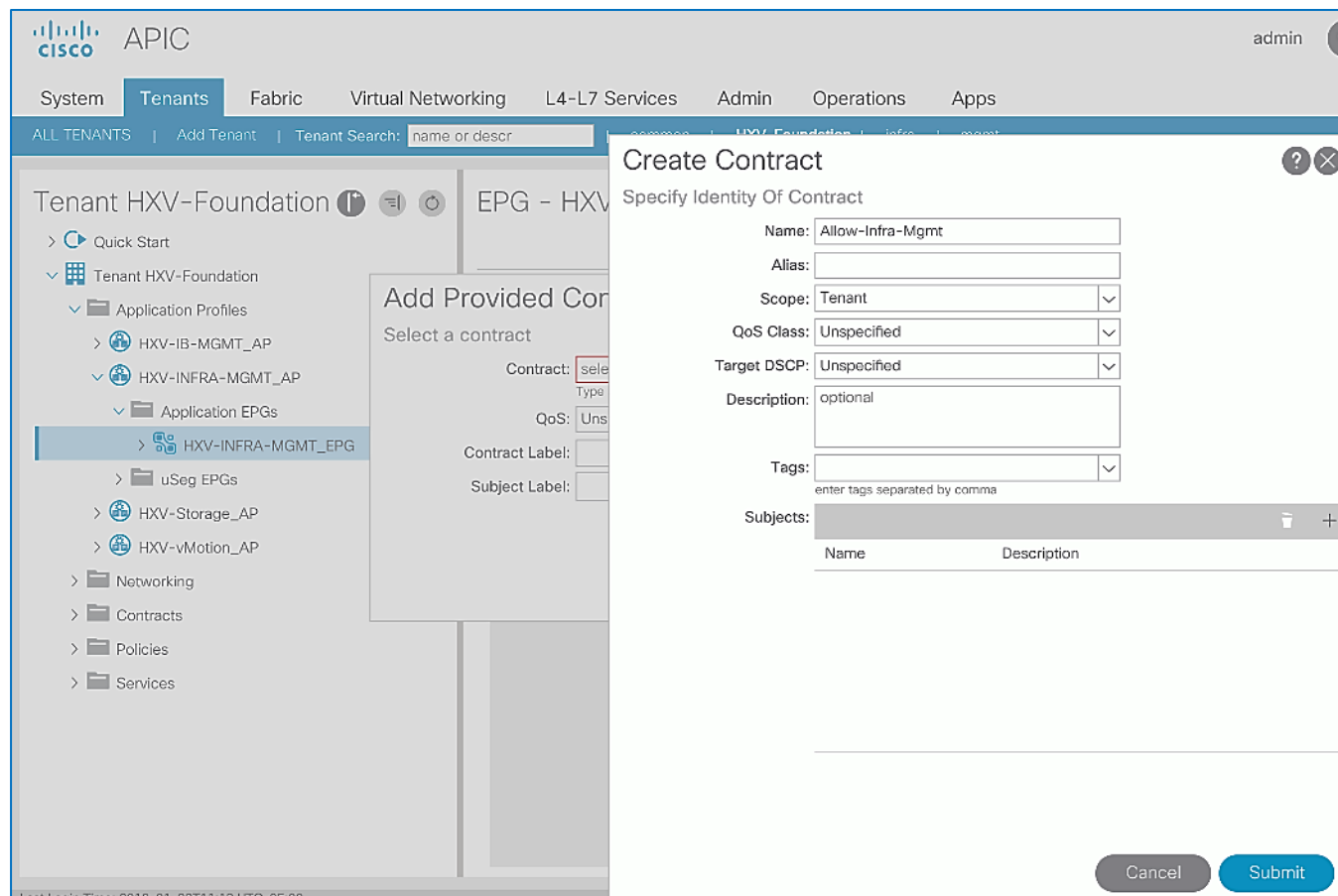
6. Click **Submit**.

Create Contract to Enable Access to Infrastructure Management

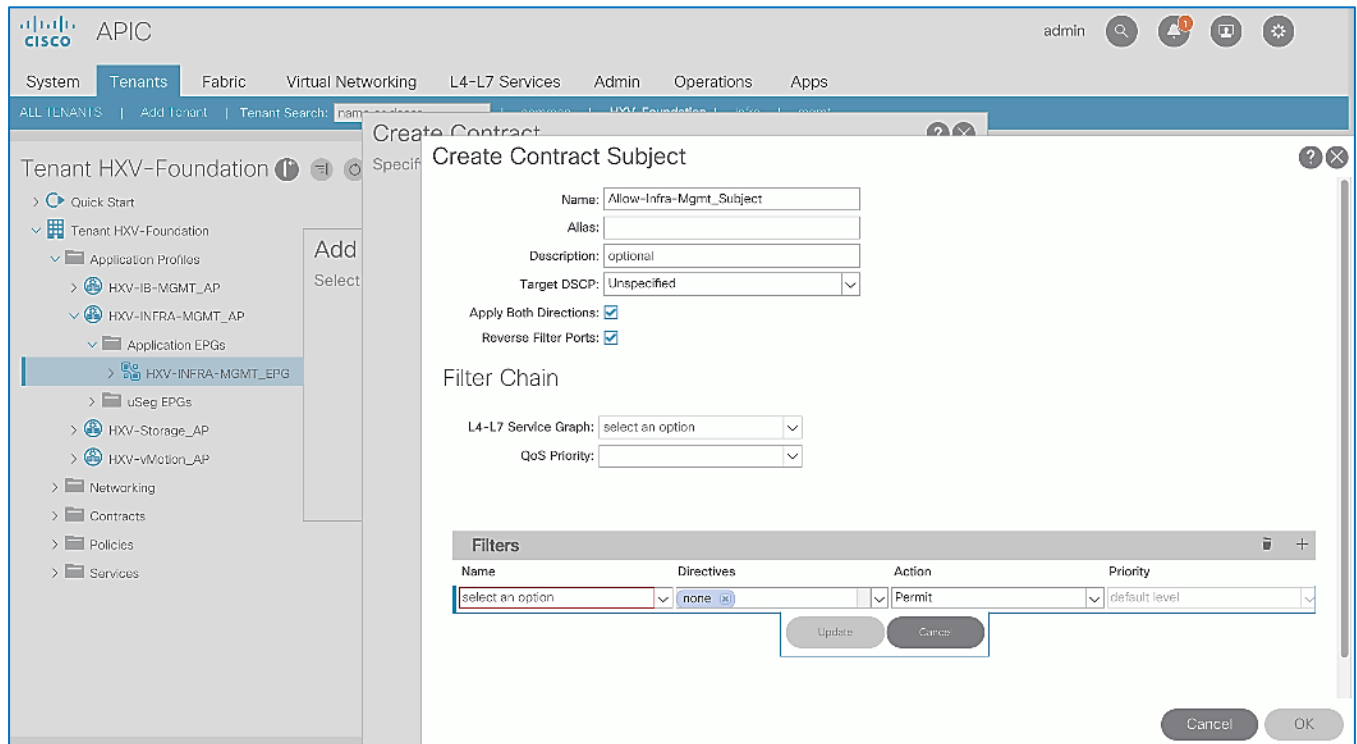
To access the infrastructure and management services hosted in the Management Cluster, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
 - **Application Profile:** HXV-INFRA-MGMT_AP
 - **EPG:** HXV-INFRA-MGMT_EPG
 - **Provided Contract:** Allow-Infra-Mgmt
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 2. From the top navigation menu, select **Tenants** > HXV-Foundation.

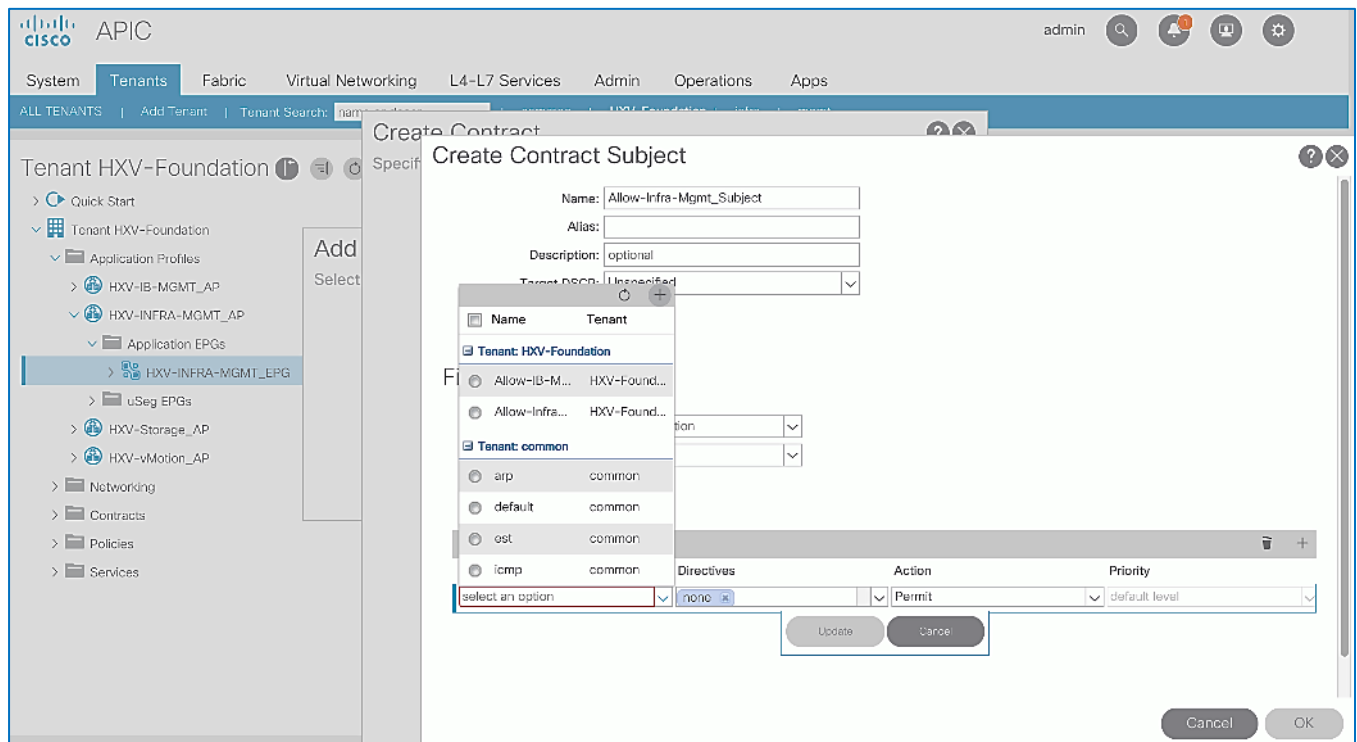
3. From the left navigation pane, select and expand **Tenant HXV-Foundation** > **Application Profiles** > HXV-INFRA-MGMT_AP > **Application EPGs** > HXV-INFRA-MGMT_EPG.
4. Right-click HXV-INFRA-MGMT_EPG and select **Add Provided Contract**.
5. In the **Add Provided Contract** pop-up window, select **Create Contract** from end of the drop-down list.
6. In the **Create Contract** pop-up window, for **Name**, specify a name (Allow-Infra-Mgmt) for the Contract.
7. For **Scope**, select **Tenant** from the drop-down list.



8. For **Subjects**, click **[+]** on the right to add a **Contract Subject**.
9. In the **Create Contract Subject** pop-up window, specify a **Name** (Allow-Infra-Mgmt_Subject) for the subject.
10. For **Filters**, click **[+]** on the right to add a **Contract Filter**.



11. For **Name**, click the down-arrow to see the drop-down list. Click **[+]** to create a Filter.



12. In the **Create Filter** pop-up window, specify a **Name** (Allow-Infra-Mgmt_Filter) for the filter.

13. For **Entries**, click **[+]** to add an Entry.

14. Enter a name (: Allow-All) for the Entry.

15. For the **EtherType**, select **IP** from the drop-down list.

16. Click **Update**.

Create Filter

Specify the Filter Identity

Name:

Alias:

Description:

Tags:

enter tags separated by comma

Entries:

Name	Alias	EtherType	ARP Flag	IP Protocol	Match Only Fragments	Stateful	Source Port / Range	Destination Port / Range	TCP Session Rules
Allow-All		IP		unspecified	False	False			

17. Click **Submit**.

18. Click **Update** in the **Create Contract Subject** pop-up window.

Create Contract Subject

Name:

Alias:

Description:

Target DSCP:

Apply Both Directions: ☒

Reverse Filter Ports: ☒

Filter Chain

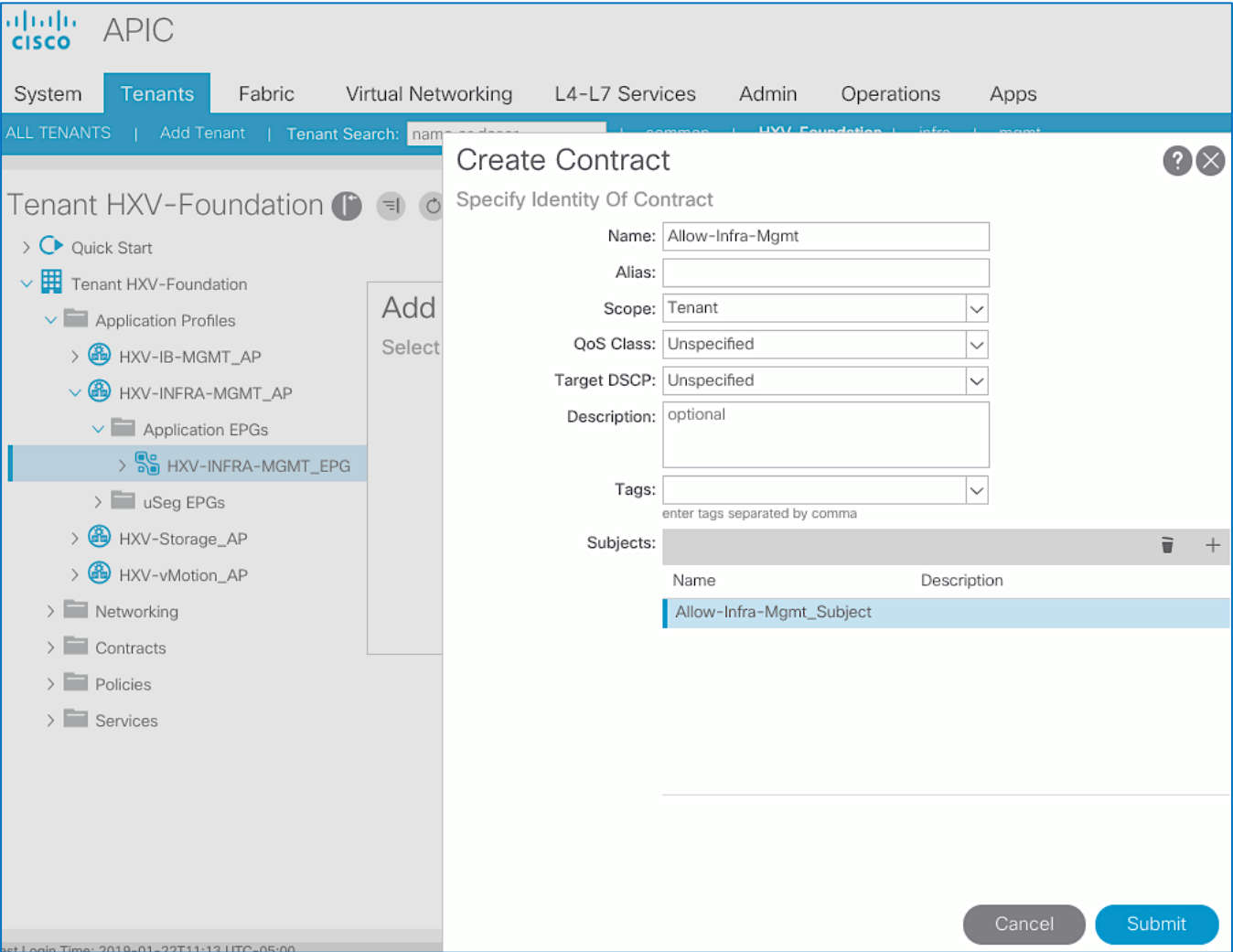
L4-L7 Service Graph:

QoS Priority:

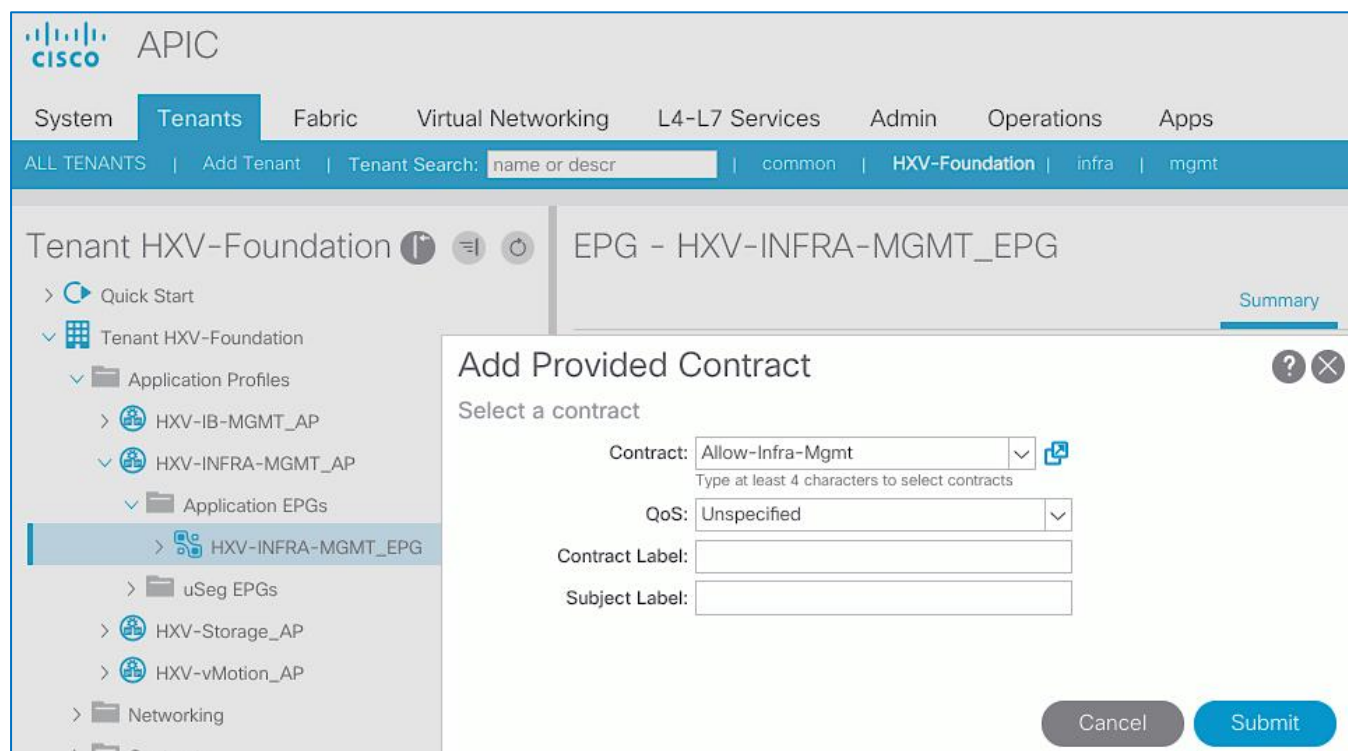
Filters

Name	Directives	Action	Priority
uni/tn-HXV-Foundation/ft-Allow	none	Permit	default level

19. Click **OK** to finish creating the Contract Subject.



20. Click **Submit** to complete creating the Contract.



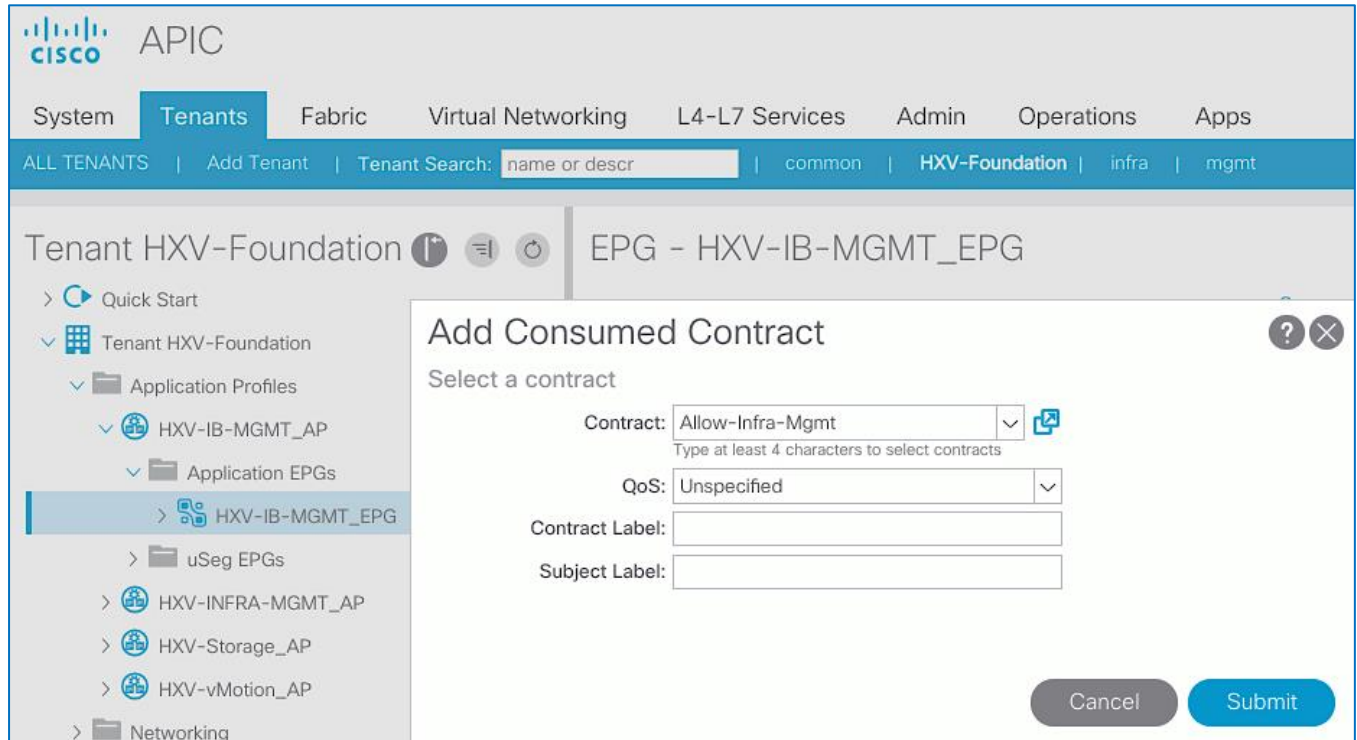
21. Click **Submit** to complete adding the Provided Contract.

This contract can be consumed by other EPGs that need reachability to Infrastructure Management virtual machines.

Enable Access to Infrastructure Management from Foundation Tenant EPGs

To access the network and services reachable via the **Shared L3Out** in the **common** Tenant, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
 - **Application Profile:** HXV-IB-MGMT_AP, HXV-Storage_AP, HXV-vMotion_AP
 - **EPG:** HXV-IB-MGMT_EPG, HXV-CL0-StorData_EPG, HXV-CL0-StorData_EPG, HXV-vMotion_EPG
 - **Consumed Contract:** Allow-Infra-Mgmt
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 2. From the top navigation menu, select **Tenants** > HXV-Foundation.
 3. From the left navigation pane, select and expand **Tenant** HXV-Foundation > **Application Profiles** > HXV-IB-MGMT_AP > **Application EPGs** > HXV-IB-MGMT_EPG.
 4. Right-click HXV-IB-MGMT_EPG and select **Add Consumed Contract**.
 5. In the **Add Consumed Contract** pop-up window, select the Allow-Infra-Mgmt contract from the drop-down list.



6. Click **Submit**. The In-band management network will now be able to access management infrastructure virtual machines.
7. Repeat steps 1-6 on other EPGs that need access to management infrastructure virtual machines.

Deploy HX Installer Virtual Machine in the HyperFlex Management Cluster

This section explains the deployment of HyperFlex Installer virtual machine on the Management HyperFlex cluster. The same procedure can be used for bringing up the other infrastructure and management virtual machines on this cluster.

The Management HyperFlex Cluster is managed by a VMware vCenter virtual machine in an existing network outside the ACI Multi-Pod Fabric, reachable through the Shared L3Out setup between the ACI fabric and the existing (non-ACI) network. The HyperFlex installer, once deployed, can be used to deploy any number of HyperFlex clusters. In this design, the HyperFlex installer deployed in the Management HyperFlex cluster will be used to deploy a HyperFlex stretched cluster for hosting Applications. See the [Install HyperFlex Stretched Cluster](#) section for more details. Both the Management and Application HyperFlex clusters are attached to the same ACI Multi-Pod Fabric.

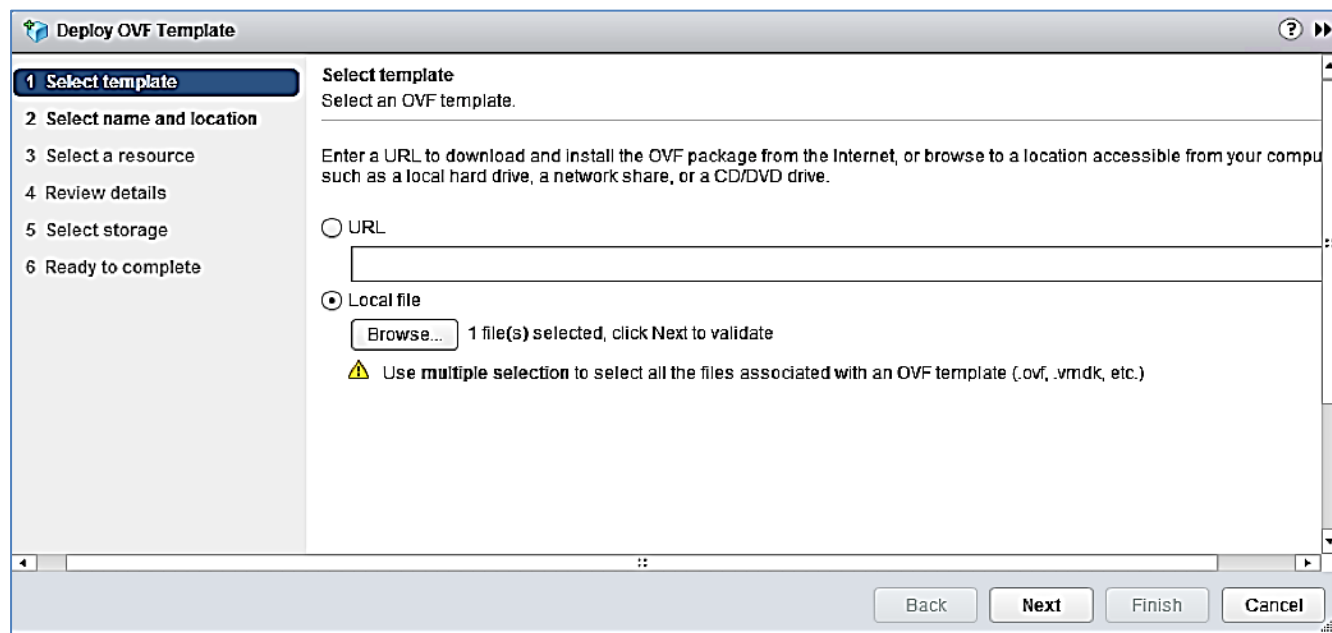
Table 66 Setup Information

VMware vCenter IP Address	10.99.167.240
Installer Virtual Machine	
IP Address	10.10.167.248/24
Gateway	10.10.167.254 (in the ACI Multi-Pod Fabric)

Network	VLAN is dynamically allocated by APIC-managed VMware vDS Port-Group: HXV-Foundation HXV-INFRA-MGMT_AP HXV-INFRA-MGMT_EPG
DNS	10.99.167.244, 10.99.167.245
NTP	192.168.167.254

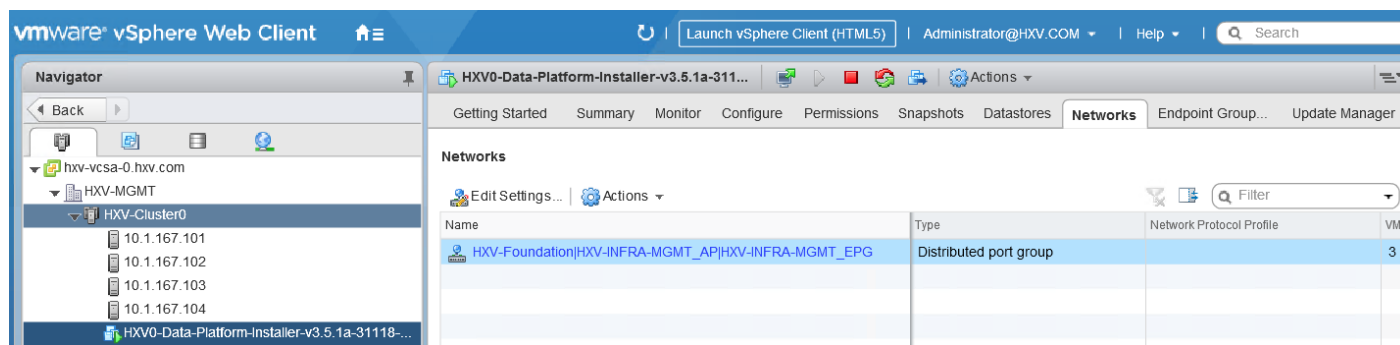
To deploy the HyperFlex installer in the Management HyperFlex Cluster, follow these steps:

1. Use a browser to navigate to the VMware vCenter Server managing the Management cluster. Click the vSphere web client of your choice and log in using an **Administrator** account.
2. From the vSphere Web Client, navigate to **Home > Hosts and Clusters**.
3. From the left navigation pane, select the **Datacenter > Cluster** and right-click to select **Deploy OVF Template....**
4. In the **Deploy OVF Template** wizard, for Select Template, select **Local file** and click the **Browse** button to locate and open the **Cisco-HX-Data-Platform-Installer** OVA file.



5. Click **Next**.
6. For **Select name and location**, specify a name for the virtual machine and select a folder location. Click **Next**.
7. For **Select a resource**, select a host or cluster or resource pool to locate the virtual machine. Click **Next**.
8. Review the details. Click **Next**.
9. For **Select storage**, select a datastore and **Thin provision virtual disk format** for the VM. Click **Next**.
10. For **Select networks**, use the drop-down list in the **Destination Networks** column to specify the network (**HXV-Foundation|HXV-INFRA-MGMT_AP|HXV-INFRA-MGMT_EPG**) the installer VM will communicate on. Click **Next**.
11. For **Customize template**, provide the IP Address, Mask, Gateway, DNS and NTP server info. Click **Next**.
12. Review the settings. Click **Finish**.

13. Power on the virtual machine.



14. From VMware vCenter, console into the installer VM to verify setup. If the HyperFlex installer was deployed using DHCP, the leased IP address can be verified from the console. Login using the default username (`root`) and password (`Cisco123`).

```
Version 3.5(1a)

*****
You can start the installation by visiting
the following URL:

    http://10.10.167.248

*****

HyperFlex-Installer login: _
```

15. Verify the IP address, NTP status, DNS configuration and change the default password as shown below.

```
root@HyperFlex-Installer:~# ifconfig eth0
eth0      Link encap:Ethernet  HWaddr 00:50:56:93:6a:6c
          inet addr:10.10.167.248  Bcast:10.10.167.255  Mask:255.255.255.0
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:14401 errors:0 dropped:0 overruns:0 frame:0
          TX packets:14241 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:10250086 (10.2 MB)  TX bytes:13995621 (13.9 MB)

root@HyperFlex-Installer:~# ping -c 3 10.10.167.254
PING 10.10.167.254 (10.10.167.254) 56(84) bytes of data.
64 bytes from 10.10.167.254: icmp_seq=1 ttl=64 time=0.229 ms
64 bytes from 10.10.167.254: icmp_seq=2 ttl=64 time=0.248 ms
64 bytes from 10.10.167.254: icmp_seq=3 ttl=64 time=0.201 ms

--- 10.10.167.254 ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 1999ms
rtt min/avg/max/mdev = 0.201/0.226/0.248/0.019 ms
root@HyperFlex-Installer:~#
root@HyperFlex-Installer:~# date
Mon Jan 21 17:18:09 UTC 2019
root@HyperFlex-Installer:~#
```

```
root@HyperFlex-Installer:~# nslookup server
Server:          10.99.167.244
Address:         10.99.167.244#53

** server can't find server: NXDOMAIN

root@HyperFlex-Installer:~# passwd
New password:
Retype new password:
passwd: password updated successfully
root@HyperFlex-Installer:~# _
```

The Installer virtual machine is now ready for installing HyperFlex clusters.

Solution Deployment – HyperFlex Application Cluster

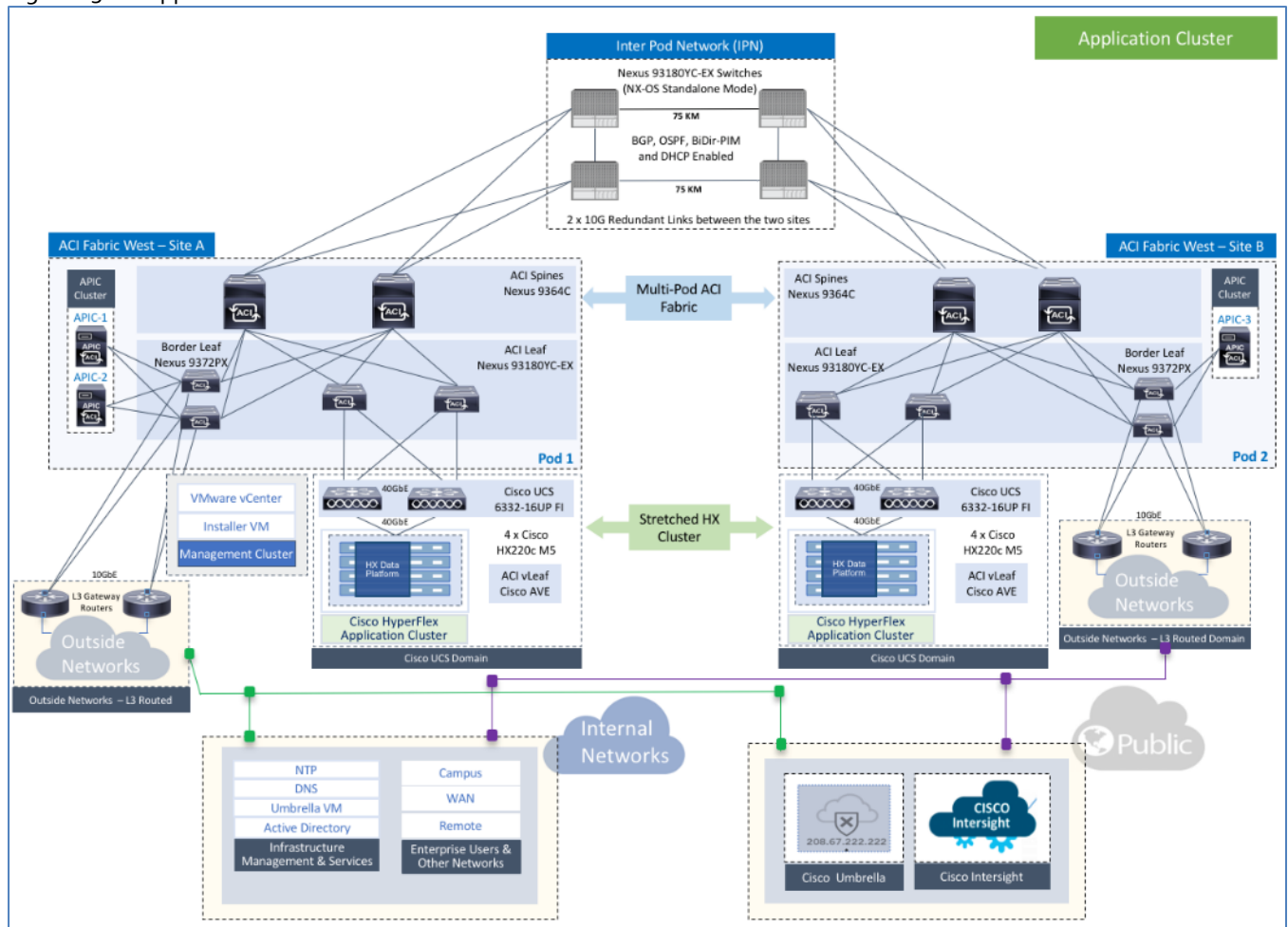
This section provides the detailed procedures for deploying a 8-node HyperFlex **stretched** cluster using an on-premise HyperFlex Installer virtual machine. This cluster will serve as an Application cluster in this design for hosting application virtual machines. The Installer VM and VMware vCenter to install and manage the cluster will be hosted on the Management Cluster. Other infrastructure services such as Active Directory, DNS etc. are located outside the ACI fabric and accessed through the shared L3Out connection from each Pod.



Cisco Intersight currently does not support the install of HyperFlex stretched clusters.

Topology

Figure 23 Application Cluster



Deployment Overview

The high-level steps for deploying an Application HyperFlex cluster in a Cisco ACI Multi-Pod fabric are as follows:

- Setup UCS domain for HyperFlex stretched cluster. This includes deploying two UCS domains, one in each Pod.

- Setup ACI fabric to provide foundational infrastructure connectivity for Cisco HyperFlex clusters. This involves defining the ACI constructs for enabling HyperFlex infrastructure connectivity (Tenant, VRF, Bridge Domain and Application Profile) across the ACI Multi-Pod fabric. The connectivity is necessary to install the HyperFlex stretched cluster.
- Setup ACI fabric for HyperFlex stretched cluster. This includes defining the ACI constructs for enabling HyperFlex infrastructure connectivity (Tenant, VRF, Bridge Domain and Application Profile) necessary to install the HyperFlex stretched cluster using the ACI Multi-Pod fabric for connectivity.
- Install HyperFlex stretched cluster using the HyperFlex installer virtual machine.
- Enable contracts to allow users to access the Application and for communication between different tiers of the application. Also, enable contracts to access the shared L3out for connectivity to outside networks and services.
- Deploy application virtual machines on the Application HyperFlex cluster.
- Add virtual machines to the port-group corresponding to the EPG

Setup Cisco UCS Domain for HyperFlex Stretched Cluster

If it is not already setup, follow the procedures outlined in the [Setup Cisco UCS Domains](#) section to deploy and setup the two Cisco UCS domains for connecting the HyperFlex stretched cluster nodes in Pod-1 and Pod-2.

Setup ACI Fabric for HyperFlex Stretched Cluster

To deploy a HyperFlex cluster in the ACI Fabric, the fabric must provide reachability to the following key infrastructure networks:

- In-Band Management Network for managing ESXi hosts and Storage Controller virtual machines that connect to it.
- Storage Data Network for storage connectivity to ESXi hosts and Storage Controller virtual machines that connect to it. Every HyperFlex cluster should be connected to a dedicated storage data network.
- VMware vMotion Network for virtual machine migration between ESXi hosts that connect to this network.
- Access to Infrastructure Management services – in this design, these services are deployed either in the Management HyperFlex cluster or in an existing network outside the ACI fabric.

The Management and Application HyperFlex clusters in this design will share the same in-band management and vMotion networks. As a result, the ACI fabric setup for these networks in the Management cluster, can be leveraged for the stretched cluster. It is still necessary to configure static bindings from these EPGs to the UCS domains in the stretched cluster in order to enable access to these networks from the stretched cluster.

For storage data traffic, however, a new network must be configured as each HyperFlex cluster requires a dedicated network for storage data. Therefore, the ACI fabric must be configured for a new storage data network.

Create Static Binding for In-Band Management to HyperFlex Stretched Cluster

To statically bind the In-Band Management EPG and VLANs to vPC interfaces going to the UCS Domains in the HyperFlex stretched cluster, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
- **Application Profile:** HXV-IB-MGMT_AP

- **EPG:** HXV-IB-MGMT_EPG
 - **Static Paths:** HXV-UCS_6300FI-A_IPG, HXV-UCS_6300FI-B_IPG
 - **VLAN:** 118
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 2. From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.
 3. From the left navigation pane, select and expand **Tenant HXV-Foundation > Application Profiles > HXV-IB-MGMT_AP > Application EPGs > HXV-IB-MGMT_EPG**.
 4. Right-click **HXV-IB-MGMT_EPG** and select **Deploy Static EPG on PC, VPC or Interface**.
 5. In the **Deploy Static EPG on PC, VPC or Interface** pop-up window, for **Path Type**, select **Virtual Port Channel**. For the **Path**, select the vPC to the first UCS Fabric Interconnect from the drop-down list. For the **Port Encap**, leave **VLAN** selected in the drop-down menu and in the box, specify the VLAN ID for the In-Band Management EPG. For the **Deployment Immediacy**, select **Immediate**.

The screenshot shows the APIC GUI with the 'Deploy Static EPG On PC, VPC, Or Interface' dialog box open. The left navigation pane shows the hierarchy: Tenant HXV-Foundation > Application Profiles > HXV-IB-MGMT_AP > Application EPGs > HXV-IB-MGMT_EPG. The dialog box has the following configuration:

- Path Type:** Virtual Port Channel
- Path:** HXV-UCS-6300FI-A
- Port Encap (or Secondary VLAN for Micro-Seg):** VLAN 118
- Deployment Immediacy:** Immediate
- Primary VLAN for Micro-Seg:** VLAN
- Mode:** Trunk
- IGMP Snoop Static Group:** (Empty table with columns Group Address and Source Address)

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

6. Click **Submit**.
7. Repeat steps 1-6 to bind the EPG to the vPC going to the second UCS Fabric Interconnect in the same UCS domain.
8. Repeat steps 1-6 for the second UCS domain in the HyperFlex stretched cluster. The resulting bindings for this network are as shown below.

The screenshot shows the Cisco APIC GUI with the 'Tenants' tab selected. The left navigation pane shows the hierarchy: Tenant HXV- > Application Profiles > HXV-IB-MGMT_AP > Application EPGs > HXV-IB-MGMT_E... > Static Ports. The main pane displays the 'Static Ports' configuration table.

Path	Primary VLAN for Micro-Seg	Port Encap (or Secondary VLAN for Micro-Seg)	Deployment Immediacy	Mode
Node: Pod-1				
Pod-1/Node-103-104/HXV-UCS-6200FI-A_IPG	unknown	vlan-118	Immediate	Trunk
Pod-1/Node-103-104/HXV-UCS-6200FI-B_IPG	unknown	vlan-118	Immediate	Trunk
Pod-1/Node-103-104/HXV-UCS-6300FI-A_IPG	unknown	vlan-118	Immediate	Trunk
Pod-1/Node-103-104/HXV-UCS-6300FI-B_IPG	unknown	vlan-118	Immediate	Trunk
Node: Pod-2				
Pod-2/Node-203-204/HXV-UCS-6300FI-A_IPG	unknown	vlan-118	Immediate	Trunk
Pod-2/Node-203-204/HXV-UCS-6300FI-B_IPG	unknown	vlan-118	Immediate	Trunk

At the bottom, there is a pagination bar showing 'Page 1 Of 1' and 'Objects Per Page: 15'.

Create Static Binding for vMotion to HyperFlex Stretched Cluster

To statically bind the vMotion EPG and VLANs to vPC interfaces going to the UCS Domains in the HyperFlex stretched cluster, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
 - **Application Profile:** HXV-vMotion_AP
 - **EPG:** HXV-vMotion_EPG
 - **Static Paths:** HXV-UCS_6300FI-A_IPG, HXV-UCS_6300FI-B_IPG
 - **VLAN:** 3018
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 2. From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.
 3. From the left navigation pane, select and expand **Tenant HXV-Foundation > Application Profiles > HXV-vMotion_AP > Application EPGs > HXV-vMotion_EPG**.
 4. Right-click **HXV-vMotion_EPG** and select **Deploy Static EPG on PC, VPC or Interface**.
 5. In the **Deploy Static EPG on PC, VPC or Interface** pop-up window, for **Path Type**, select **Virtual Port Channel**. For the **Path**, select the vPC to the first UCS Fabric Interconnect from the drop-down list. For the **Port Encap**, leave **VLAN** selected in the drop-down menu and in the box, specify the VLAN ID for the In-Band Management EPG. For the **Deployment Immediacy**, select **Immediate**.

Deploy Static EPG On PC, VPC, Or Interface

Select PC, VPC, or Interface

Path Type: ☐ Port ☐ Direct Port Channel ☒ Virtual Port Channel

Path: HXV-UCS-6300FI-A

Port Encap (or Secondary VLAN for Micro-Seg): VLAN Integer Value

Deployment Immediacy: ☒ Immediate ☐ On Demand

Primary VLAN for Micro-Seg: VLAN Integer Value

Mode: ☒ Trunk ☐ Access (802.1P) ☐ Access (Untagged)

IGMP Snoop Static Group:

Group Address	Source Address

Cancel Submit

- Click **Submit**.
- Repeat steps 1-6 to bind the EPG to the vPC going to the second UCS Fabric Interconnect in the same UCS domain.
- Repeat steps 1-6 for the second UCS domain in the HyperFlex stretched cluster. The resulting bindings for this network are as shown below.

Static Ports

Path	Primary VLAN for Micro-Seg	Port Encap (or Secondary VLAN for Micro-Seg)	Deployment Immediacy	Mode
Node: Pod-1				
Pod-1/Node-103-104/HXV-UCS-6200FI-A_IPG	unknown	vlan-3018	Immediate	Trunk
Pod-1/Node-103-104/HXV-UCS-6200FI-B_IPG	unknown	vlan-3018	Immediate	Trunk
Pod-1/Node-103-104/HXV-UCS-6300FI-A_IPG	unknown	vlan-3018	Immediate	Trunk
Pod-1/Node-103-104/HXV-UCS-6300FI-B_IPG	unknown	vlan-3018	Immediate	Trunk
Node: Pod-2				
Pod-2/Node-203-204/HXV-UCS-6300FI-A_IPG	unknown	vlan-3018	Immediate	Trunk
Pod-2/Node-203-204/HXV-UCS-6300FI-B_IPG	unknown	vlan-3018	Immediate	Trunk

Configure ACI Fabric for Storage Data Traffic on HyperFlex Stretched Cluster

This section explains the configuration of the Cisco ACI fabric to enable forwarding of HyperFlex storage data traffic between nodes in the HyperFlex stretched cluster. The nodes in the stretched cluster are distributed across two sites interconnected by a Cisco ACI Multi-Pod fabric. The configuration in this section will enable the forwarding of storage data network traffic across the ACI fabric.

Create Bridge Domain for Storage Data Traffic on HyperFlex Stretched Cluster

To create a Bridge Domain for Storage data traffic, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
 - **VRF:** HXV-Foundation_VRF
 - **Bridge Domain:** HXV-CL1-Storage_BD
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 2. From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.
 3. From the left navigation pane, expand and select Tenant **HXV-Foundation > Networking > Bridge Domains**.
 4. Right-click and select **Create Bridge Domain**.
 5. In the **Create Bridge Domain** wizard, for **Name**, specify a name (**HXV-CL1-Storage_BD**) for the bridge domain. For **VRF**, select the previously created VRF from the drop-down list. For **Forwarding**, select **Custom** from the drop-down list. For **L2 Unknown Unicast**, select **Flood** from the drop-down list. The checkbox for **ARP Flooding** should now show up and be enabled.

APIC Create Bridge Domain

System Tenant **STEP 1 > Main** 1. Main 2. L3 Configurations 3. Advanced/Troubleshooting

Specify Bridge Domain for the VRF

Name: HXV-CL1-Storage_BD

Alias:

Description: optional

Tags: enter tags separated by comma

Type: ☒ fc ☐ regular

Advertise Host Routes: ☐

VRF: HXV-Foundation_VRF

Forwarding: Custom

L2 Unknown Unicast: Flood

L3 Unknown Multicast Flooding: Flood

Multi Destination Flooding: Flood in BD

ARP Flooding: ☒ Enabled

Clear Remote MAC Entries: ☐

Endpoint Retention Policy: select a value
This policy only applies to local L2 L3 and remote L3 entries

IGMP Snoop Policy: select a value

Previous Cancel Next

6. Click **Next**.
7. In the **L3 Configurations** section, for **EP Move Detection Mode**, select the checkbox to enable **GARP based detection** if needed. See [Review/Enable ACI Fabric Settings](#) section for more details on when to enable this feature.

The screenshot shows the 'Create Bridge Domain' configuration page in the Cisco APIC GUI, specifically the 'STEP 2 > L3 Configurations' section. The page is titled 'Specify Bridge Domain for the VRF'. It includes several configuration options:

- Unicast Routing:** ☐ Enabled
- ARP Flooding:** ☒ Enabled
- Config BD MAC Address:** ☒
 - MAC Address:** 00:22:BD:F8:19:FF
 - Virtual MAC Address:** not-applicable
- Subnets:** A table with columns: Gateway Address, Scope, Primary IP Address, Subnet Control.
- Endpoint Dataplane Learning:** ☒
- Limit IP Learning To Subnet:** ☒
- EP Move Detection Mode:** ☒ GARP based detection
- DHCP Labels:** A table with columns: Name, Scope, DHCP Option Policy.
- Associated L3 Outs:** A table with columns: L3 Out.

At the bottom right, there are three buttons: 'Previous', 'Cancel', and 'Next' (highlighted in blue).

8. Click **Next**. Skip the **Advanced/Troubleshooting** section. Click **Finish** to complete.

Create Application Profile for Storage Data Traffic on HyperFlex Stretched Cluster

The application profile for HyperFlex Storage data traffic on the HyperFlex stretched Cluster will use the same profile (HXV-Storage_AP) as that of the Management Cluster. Proceed to the next section to create a separate EPG for the stretched cluster storage traffic.

Create EPG for Storage Data Traffic on HyperFlex Stretched Cluster

To create an EPG for storage data traffic on HyperFlex stretched cluster, follow these steps using the setup information provided below:

- **Tenant:** HXV-Foundation
 - **Application Profile:** HXV-Storage_AP
 - **Bridge Domain:** HXV-CL1-Storage_BD
 - **EPG:** HXV-CL0-StorData_EPG
1. Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 2. From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.

- From the left navigation pane, select and expand **Tenant HXV-Foundation > Application Profiles > HXV-Storage_AP**.
- Right-click HXV-Storage_AP and select **Create Application EPG**.
- In the **Create Application EPG** pop-up window, for **Name**, specify a name for the EPG. For **Bridge Domain**, select the previously created Bridge Domain.

APIC admin

System Tenants

ALL TENANTS | Add Ten

Tenant HXV-Foundation

Quick Start

Tenant HXV-Foundation

Application Profiles

HXV-IB-MGM

HXV-INFRA-M

HXV-Storage_AP

Application Profiles

uSeg EPGs

HXV-vMotion

Networking

Contracts

Policies

Services

Create Application EPG

STEP 1 > Identity

Specify the EPG Identity

Name: HXV-CL1-StorData_EPG

Alias:

Description: optional

Tags: enter tags separated by comma

Contract Exception Tag:

QoS class: Unspecified

Custom QoS: select a value

Data-Plane Policer: select a value

Intra EPG Isolation: Enforced Unenforced

Preferred Group Member: Exclude Include

Flood on Encapsulation: Disabled Enabled

Bridge Domain: HXV-CL1-Storage_BD

Monitoring Policy: select a value

FHS Trust Control Policy: select a value

Shutdown EPG: ☐

Associate to VM Domain Profiles: ☐

Statically Link with Leaves/Paths: ☐

EPG Contract Master:

Application EPGs

Previous Cancel Finish

- Click **Finish**.

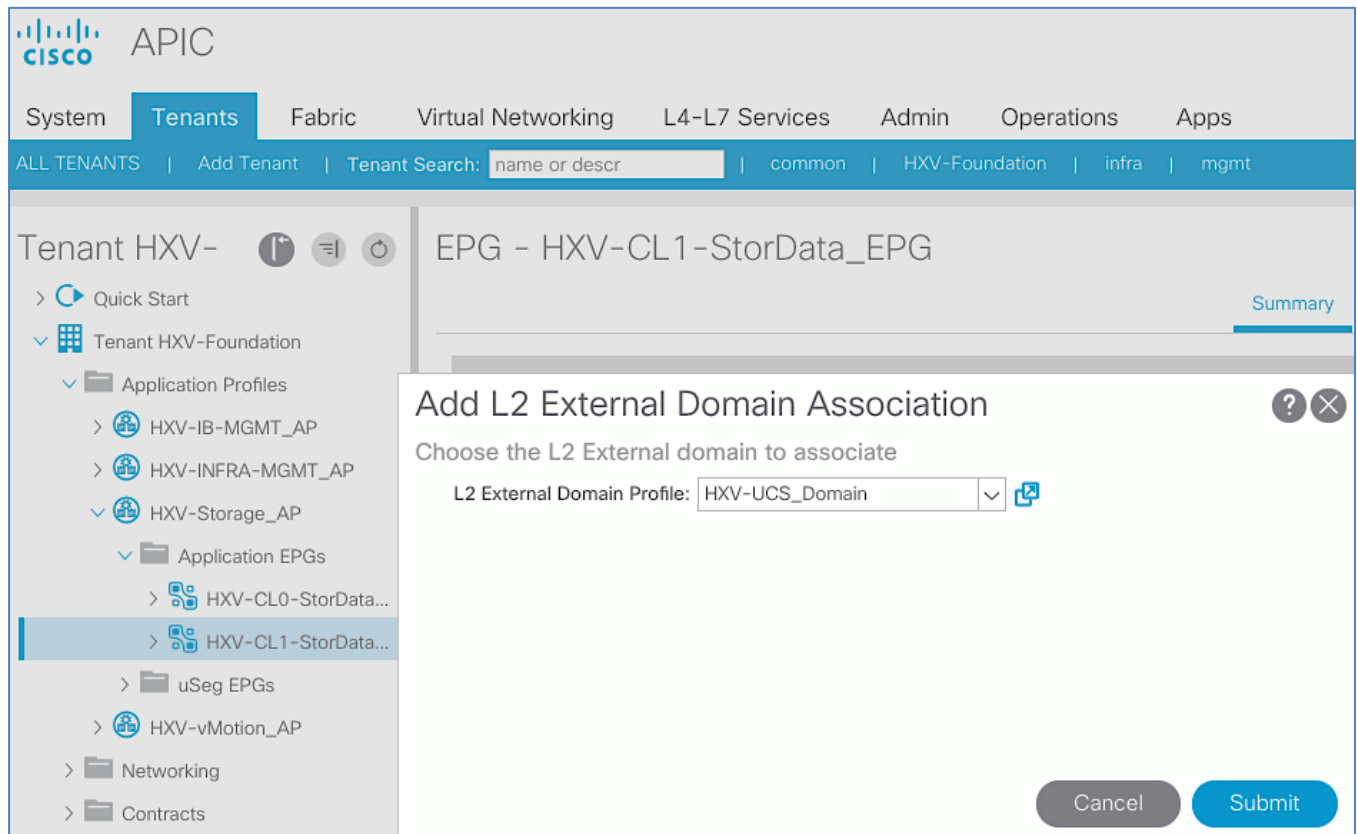
Associate EPG for Storage Data Traffic with UCS Domain

To associate the HyperFlex Storage EPG with UCS Domain, follow these steps using the setup information provided below:

- Tenant:** HXV-Foundation
- Application Profile:** HXV-Storage_AP
- EPG:** HXV-CL1-StorData_EPG
- Domain:** HXV-UCS_Domain

- Use a browser to navigate to the APIC GUI. Log in using the **admin** account.

- From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.
- From the left navigation pane, select and expand **Tenant HXV-Foundation > Application Profiles > HXV-Storage_AP > Application EPGs > HXV-CL1-StorData_EPG**.
- Right-click **HXV-CL1-StorData_EPG** and select **Add L2 External Domain Association**.
- In the **Add L2 External Domain Association** pop-up window, select the previously created UCS Domain from the list.



- Click **Submit**.

Create Static Binding for Storage Data Traffic to UCS Domain for HyperFlex Stretched Cluster

To statically bind the HyperFlex Storage Data EPG and VLANs to vPC interfaces going to the UCS Domains that connect to the HyperFlex stretched cluster, follow these steps using the setup information provided below:

- Tenant:** HXV-Foundation
 - Application Profile:** HXV-Storage_AP
 - EPG:** HXV-CL1-StorData_EPG
 - Static Paths:** HXV-UCS_6300FI-A_IPG, HXV-UCS_6300FI-B_IPG
 - VLAN:** 3218
- Use a browser to navigate to the APIC GUI. Log in using the **admin** account.
 - From the top navigation menu, select **Tenants > HXV-Foundation**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-Foundation**.

- From the left navigation pane, select and expand **Tenant HXV-Foundation > Application Profiles > HXV-Storage_AP > Application EPGs > HXV-CL1-StorData_EPG**.
- Right-click **HXV-CL1-StorData_EPG** and select **Deploy Static EPG on PC, VPC or Interface**.
- In the **Deploy Static EPG on PC, VPC or Interface** pop-up window, for **Path Type**, select **Virtual Port Channel**. For the **Path**, select the vPC to the first UCS Fabric Interconnect from the drop-down list. For the **Port Encap**, leave **VLAN** selected in the drop-down menu and in the box, specify the VLAN ID for the In-Band Management EPG. For the **Deployment Immediacy**, select **Immediate**.

Deploy Static EPG On PC, VPC, Or Interface

Select PC, VPC, or Interface

Path Type: Port Direct Port Channel **Virtual Port Channel**

Path: HXV-UCS-6300FI-A

Port Encap (or Secondary VLAN for Micro-Seg): VLAN 3218
Integer Value

Deployment Immediacy: **Immediate** On Demand

Primary VLAN for Micro-Seg: VLAN
Integer Value

Mode: **Trunk** Access (802.1P) Access (Untagged)

IGMP Snoop Static Group:

Group Address	Source Address

Cancel **Submit**

- Click **Submit**.
- Repeat steps 1-6 to bind the EPG to the vPC going to the second UCS Fabric Interconnect in the same UCS domain.
- Repeat steps 1-6 for the second UCS domain in the HyperFlex stretched cluster. The resulting bindings for this network are as shown below.

The screenshot shows the Cisco APIC interface. The top navigation bar includes tabs for System, Tenants (selected), Fabric, Virtual Networking, L4-L7 Services, Admin, Operations, and Apps. Below the navigation bar, there's a search bar and filters for ALL TENANTS, Add Tenant, and Tenant Search. The left sidebar shows a tree view under Tenant HXV- with options like Quick Start, Tenant HXV-Foundation, Application Profiles, and Application EPGs. The main content area is titled 'Static Ports' and displays a table with columns: Path, Primary VLAN for Micro-Seg, Port Encap (or Secondary VLAN for Micro-Seg), Deployment Immediacy, and Mode. The table lists configurations for Pod-1 and Pod-2, each with two entries for different IP ranges.

Path	Primary VLAN for Micro-Seg	Port Encap (or Secondary VLAN for Micro-Seg)	Deployment Immediacy	Mode
Node: Pod-1				
Pod-1/Node-103-104/HXV-UCS-6300FI-A_IPG	unknown	vlan-3218	Immediate	Trunk
Pod-1/Node-103-104/HXV-UCS-6300FI-B_IPG	unknown	vlan-3218	Immediate	Trunk
Node: Pod-2				
Pod-2/Node-203-204/HXV-UCS-6300FI-A_IPG	unknown	vlan-3218	Immediate	Trunk
Pod-2/Node-203-204/HXV-UCS-6300FI-B_IPG	unknown	vlan-3218	Immediate	Trunk

Install HyperFlex Stretched Cluster (Applications) using Installer Virtual Machine

In this section, the installation of a (4+4) node HyperFlex **stretched** cluster is explained. This cluster is deployed using an on-premise installer. A HyperFlex standard cluster for Management, covered in an earlier section, was installed using Cisco Intersight.



Cisco Intersight currently does not support the installation of HyperFlex stretched clusters.

The HyperFlex stretched cluster in this design is intended for application virtual machines and will be referred to as the Applications Cluster. The Management cluster on the other hand is intended for virtual machines that provide management and other infrastructure services to Application clusters and other HyperFlex clusters attached to the same ACI Multi-Pod fabric.

Similar to Cisco Intersight installation, the HyperFlex installer virtual machine will configure Cisco UCS policies, templates, service profiles, and settings, as well as assigning IP addresses to the HX servers that come from the factory with ESXi hypervisor software preinstalled. The installer will deploy the HyperFlex controller virtual machines and software on the nodes, add the nodes to VMware vCenter managing the HX Cluster, and finally create the HyperFlex cluster and distributed filesystem. The setup is done through a deployment wizard by providing the necessary information.

The deployment of a HyperFlex **stretched** cluster explained in this section consists of the following high-level steps.

- Configure Site 1 (Wizard)
- Configure Site 2 (Wizard)
- Deploy Witness Virtual Machine in a third Site (OVA)
- Create Cluster (Wizard)
- Verify Setup

Prerequisites

The prerequisites necessary for installing a HyperFlex **stretched** cluster from Cisco Intersight is as follows:

1. Reachability from HyperFlex Installer to the out-of-band management interfaces on Fabric Interconnects that the HyperFlex system being deployed connects to. This provides the installer access to Cisco UCS Manager.

2. Reachability from HyperFlex Installer to the out-of-band management (CIMC) interfaces on the servers, reachable via the Fabric Interconnects' management interfaces. This network (**ext-mgmt**) should be in the same subnet as the Fabric Interconnect management interfaces.
3. ACI Multi-Pod Fabric setup to enable connectivity between HyperFlex Installer and infrastructure services necessary for deploying a HyperFlex stretched cluster. This includes access to NTP, AD/DNS, VMware vCenter and Witness Virtual machines. In this design, these services are either in the Management HyperFlex cluster connected to the same ACI Multi-Pod fabric or in an existing non-ACI network that is accessible through the Shared L3Out setup between ACI Multi-Pod fabric and the existing network
4. Reachability from HyperFlex Installer to the ESXi in-band management interface of the hosts in the HyperFlex cluster being installed.
5. Reachability from HyperFlex Installer to the VMware vCenter Server that will manage the HyperFlex cluster(s) being deployed.



The VMware vCenter Virtual Machine must be hosted on a separate virtualization environment and should not be on the HyperFlex cluster being deployed.

6. Reachability from HyperFlex Installer to the DNS server(s) for use by the HyperFlex cluster being installed.
7. Reachability from HyperFlex Installer to the NTP server(s) for use by the HyperFlex cluster being installed.
8. ACI Multi-Pod Fabric setup to enable connectivity to HyperFlex cluster networks - ESXi and Storage Controller management, ESXi and Storage Data networks, vMotion and Application VM networks.
9. Reachability from VMware vCenter to ESXi and Storage Controller Management networks.
10. Enable the necessary ports to install HyperFlex. For more information, see Networking Ports section in Appendix A of the HyperFlex Hardening Guide: https://www.cisco.com/c/dam/en/us/support/docs/hyperconverged-infrastructure/hyperflex-hx-data-platform/HX-Hardening_Guide_v3_5_v12.pdf
11. Review the Pre-installation Checklist for Cisco HX Data Platform: https://www.cisco.com/c/en/us/td/docs/hyperconverged_systems/HyperFlex_HX_DataPlatformSoftware/HyperFlex_Preinstall_Checklist/b_HX_Data_Platform_Preinstall_Checklist.html

Setup Information

The setup information used in this design to install a HyperFlex **stretched** cluster is provided below.

The following are the services in the Management HyperFlex Cluster:

- VMware vCenter VM IP Address: 10.10.167.240
- Installer VM IP Address: 10.10.167.248

The following are the services in the Existing non-ACI Network:

- Witness VM IP Address: 10.99.167.248

Site 1 Information

Table 67 Site 1 – Credentials

HyperFlex Stretched Cluster Install - Credentials	
Cisco UCS Manager > FQDN or IP	192.168.167.204
Cisco UCS Manager > Username/Password	admin/*****
Site Name	Site 1

Table 68 Site 1 – UCSM Configuration

Network Type	VLAN Name	VLAN ID
VLAN for Hypervisor and HyperFlex Management	hxv-inband-mgmt	118
VLAN for VM vMotion	hxv-vmotion	3018
VLAN for HyperFlex storage traffic	hxv-cll-storage-data	3218
VLAN for VM Network	hxv-vm-network	2118

HyperFlex Stretched Cluster Install - Cisco UCSM Configuration	
MAC Pools	
MAC Pool Prefix	00:25:B5:A8
'hx-ext-mgmt' IP Pool for Cisco IMC	
IP Blocks	192.168.167.111-.114
Subnet Mask	255.255.255.0
Gateway	192.168.167.254
Cisco IMC access management	
Out of band	✓
Advanced	
UCS Firmware	4.0(1c)
HyperFlex Cluster Name	HXV-Cluster1
Org Name	HXV-Org1

Table 69 Site 1 – Hypervisor Configuration

HyperFlex Stretched Cluster Install - Cisco UCSM Configuration	
Configure common Hypervisor Settings	
Subnet Mask	255.255.255.0
Gateway	10.1.167.254
DNS Server(s)	10.99.167.244,10.99.167.245
Hypervisor Settings	
Make IP Addresses and Hostnames Sequential	✓
IP Addresses	10.1.167.111-.114
Hostnames	hxv-cl1-esxi-[1-4]
Hypervisor Credentials	
Admin User name	root
Hypervisor Password	*****

Site 2 Information

Table 70 Site 2 – Credentials

HyperFlex Stretched Cluster Install - Credentials	
Cisco UCS Manager > FQDN or IP	192.168.167.207
Cisco UCS Manager > Username/Password	admin/*****
Site Name	Site 2

Table 71 Site 2 – UCSM Configuration

Network Type	VLAN Name	VLAN ID
VLAN for Hypervisor and HyperFlex Management	hxv-inband-mgmt	118
VLAN for VM vMotion	hxv-vmotion	3018
VLAN for HyperFlex storage traffic	hxv-cl1-storage-data	3218
VLAN for VM Network	hxv-vm-network	2118

HyperFlex Stretched Cluster Install - Cisco UCSM Configuration	
MAC Pools	
MAC Pool Prefix	00:25:B5:A9
'hx-ext-mgmt' IP Pool for Cisco IMC	
IP Blocks	192.168.167.115-.118
Subnet Mask	255.255.255.0
Gateway	192.168.167.254
Cisco IMC access management	
Out of band	✓
Advanced	
UCS Firmware	4.0 (1c)
HyperFlex Cluster Name	HXV-Cluster1
Org Name	HXV-Org1

Table 72 Site 2 – Hypervisor Configuration

HyperFlex Stretched Cluster Install - Cisco UCSM Configuration	
Configure common Hypervisor Settings	
Subnet Mask	255.255.255.0
Gateway	10.1.167.254
DNS Server(s)	10.99.167.244,10.99.167.245
Hypervisor Settings	
Make IP Addresses and Hostnames Sequential	✓
IP Addresses	10.1.167.115-.118
Hostnames	hxv-cl1-esxi-[5-8]
Hypervisor Credentials	
Admin User name	root
Hypervisor Password	*****

Cluster Information

Table 73 Cluster – Credentials

UCS Manager			
FQDN or IP	192.168.167.204	192.168.167.207	
Username/Password	admin/*****	admin/*****	
Site Name	Site 1	Site 2	
Org Name	HXV-Org1	HXV-Org1	ORG name can same in different UCS domain/FIs
VMware vCenter			
FQDN or IP	hxv0-vcsa.hxv.com (10.10.167.240)		
Username/Password	administrator@hxv.com/*****		
Hypervisor			
Username/Password	root/*****		Factory Default: Cisco123

Table 74 Cluster – IP Addresses

Hypervisor		Storage Controller VM (SCVM)	
Site 1 – Management IP	10.1.167.111-.114	10.1.167.161-.164	
Site 2 – Management IP	10.1.167.115-.118	10.1.167.165-.168	
Site 1 – Data IP	172.1.167.111-.114	172.1.167.161-.164	
Site 2 – Data IP	172.1.167.115-.118	172.1.167.165-.168	
Cluster	Management	Data	
Cluster IP Address	10.1.167.110	172.1.167.110	
Subnet Mask	255.255.255.0	255.255.255.0	
Gateway	10.1.167.254	–	
Witness			
Witness IP Address	10.99.167.249		Located outside the ACI Fabric in a 3 rd site

Table 75 Cluster Configuration

Cisco HX Cluster			
HyperFlex Cluster Name	HXV-Cluster1		
Replication Factor (RF)	2+2		
Controller VM			
Admin Password	*****		
vCenter Configuration			
vCenter Datacenter	HXV-APP		
vCenter Cluster	HXV-Cluster1		
System Services			
DNS Servers (On-Premise Cisco Umbrella Virtual Appliances)	10.99.167.244, 10.99.167.245		
NTP	192.168.167.254		
DNS Domain Name	hxv.com		
Timezone	America/New_York		
Advanced Networking		VLAN ID	Management vSwitch
Management VLAN Tag – Site 1		118	vswitch-hxv-inband-mgmt
Management VLAN Tag – Site 2		118	
Data VLAN Tag – Site 1		3218	vswitch-hxv-cll-storage-data
Data VLAN Tag – Site 2		3218	
Advanced Configuration			
Jumbo Frames	✓ Enable Jumbo Frames on Data Network		Enabled - Yes
Disk Partitions	❑ Clean Up Disk Partitions		Enabled - No
Virtual Desktop (VDI)	❑ Optimize for VDI Deployment		Enabled - No

Deployment Steps

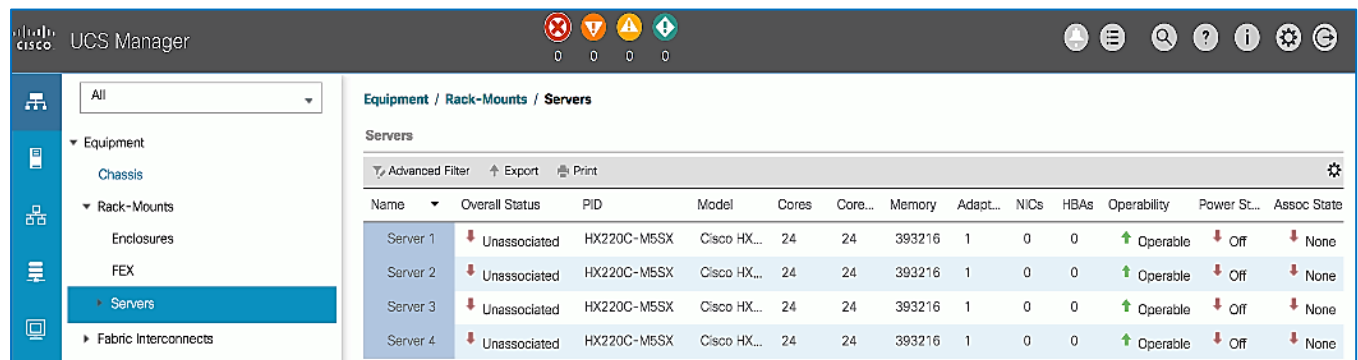
To deploy a HyperFlex **stretched** cluster across two sites interconnected by an ACI Multi-Pod fabric, complete the steps outlined in this section. The HyperFlex servers are connected to a separate pair of Cisco UCS Fabric Interconnects in each site.

Verify Server Status in Site 1 and Site 2 Before HyperFlex Installation

Before starting the HyperFlex installation process that will create the service profiles and associate them with the servers, you must verify that the servers in both Cisco UCS domains have finished their discovery process and are in the correct state.

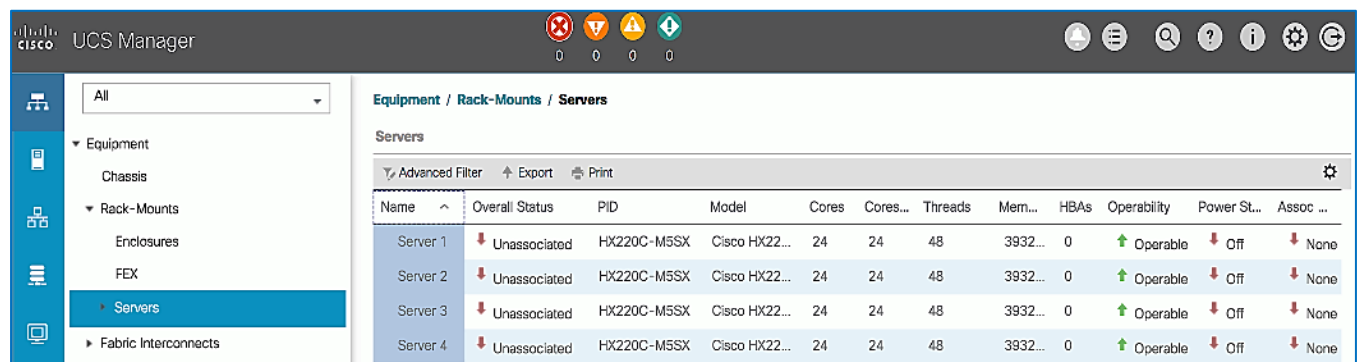
To verify the server status in Site 1 and Site 2, follow these steps:

1. Use a browser to navigate to the Cisco UCS Manager in the first HyperFlex stretched cluster site (**Site 1**). Log in using the **admin** account.
2. From the left navigation pane, click the **Equipment** icon.
3. Navigate to **All > Equipment**. In the right window pane, click the **Servers** tab.



Name	Overall Status	PID	Model	Cores	Core...	Memory	Adapt...	NICs	HBAs	Operability	Power St...	Assoc State
Server 1	Unassociated	HX220C-M5SX	Cisco HX...	24	24	393216	1	0	0	Operable	Off	None
Server 2	Unassociated	HX220C-M5SX	Cisco HX...	24	24	393216	1	0	0	Operable	Off	None
Server 3	Unassociated	HX220C-M5SX	Cisco HX...	24	24	393216	1	0	0	Operable	Off	None
Server 4	Unassociated	HX220C-M5SX	Cisco HX...	24	24	393216	1	0	0	Operable	Off	None

4. For the **Overall Status**, the servers should be in an **Unassociated** state. The servers should also be in an **Operable** state, powered **Off** and have no alerts with no faults or errors.
5. Repeat steps 1-4 for the Hyperflex nodes and Cisco UCS Manager in the **second** HyperFlex stretched cluster site (**Site 2**).



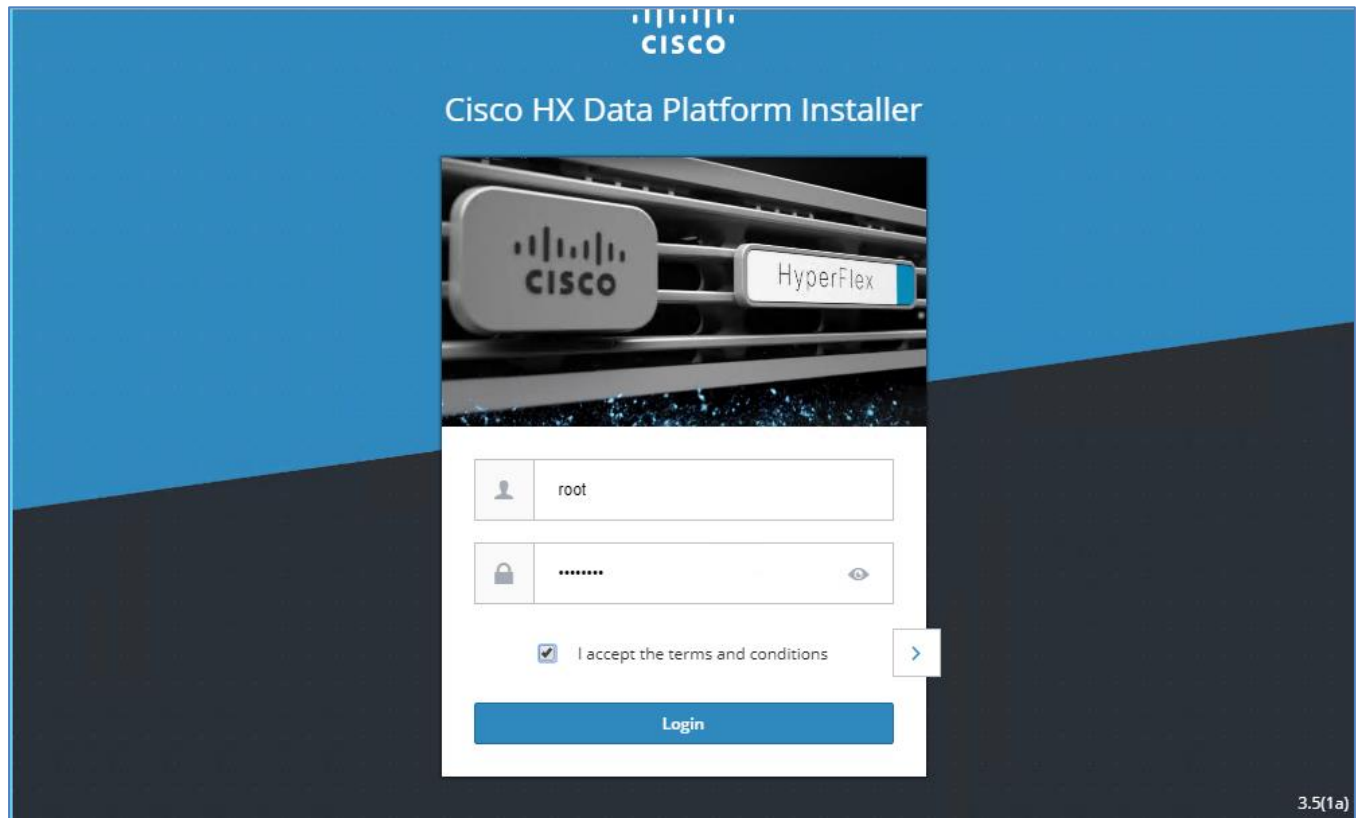
Name	Overall Status	PID	Model	Cores	Cores...	Threads	Mem...	HBAs	Operability	Power St...	Assoc ...
Server 1	Unassociated	HX220C-M5SX	Cisco HX22...	24	24	48	3932...	0	Operable	Off	None
Server 2	Unassociated	HX220C-M5SX	Cisco HX22...	24	24	48	3932...	0	Operable	Off	None
Server 3	Unassociated	HX220C-M5SX	Cisco HX22...	24	24	48	3932...	0	Operable	Off	None
Server 4	Unassociated	HX220C-M5SX	Cisco HX22...	24	24	48	3932...	0	Operable	Off	None

6. The servers in **both** sites are now ready for installing the HyperFlex Data Platform Software.

Access the HyperFlex Installer

To access the HyperFlex installer virtual machine, follow these steps:

1. Use a web browser to navigate to the IP address of the installer virtual machine. Click **accept** or **continue** to bypass any SSL certificate errors.
2. At the login screen, enter the username and password. The default username is: `root`. Password is either the default password (`Cisco123`) or whatever it was changed to after the OVA was deployed. Check the box for accepting terms and conditions. Verify the version of the installer – see lower right-hand corner of the login page.

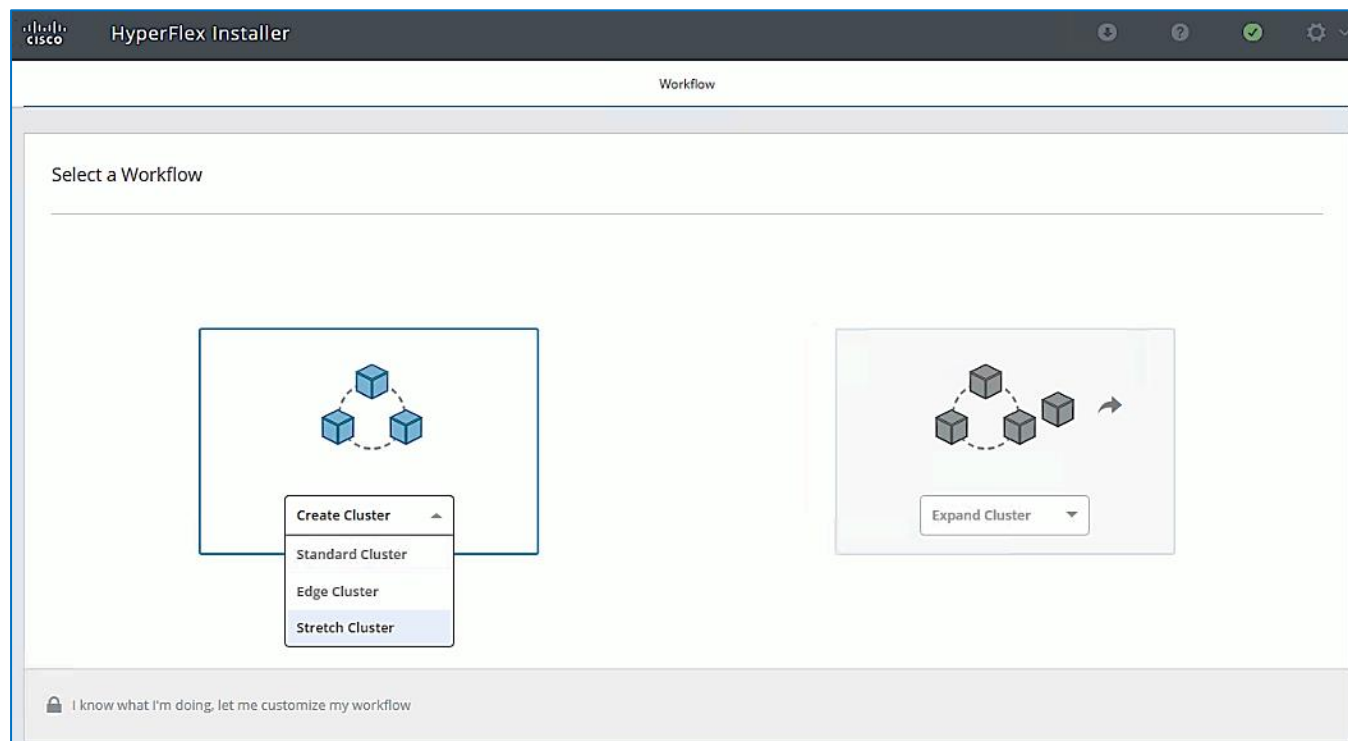


3. Click **Login**.
4. You should now be forwarded to the HyperFlex Installer Workflow page where you can install a new Standard Cluster, Stretch Cluster, Edge Cluster or expand an existing cluster. In this CVD, the installer virtual machine is used to deploy a HyperFlex stretched cluster.

Configure Site 1 from Deployment Wizard

To configure the first site (**Site 1**) in the stretched cluster, follow these steps:

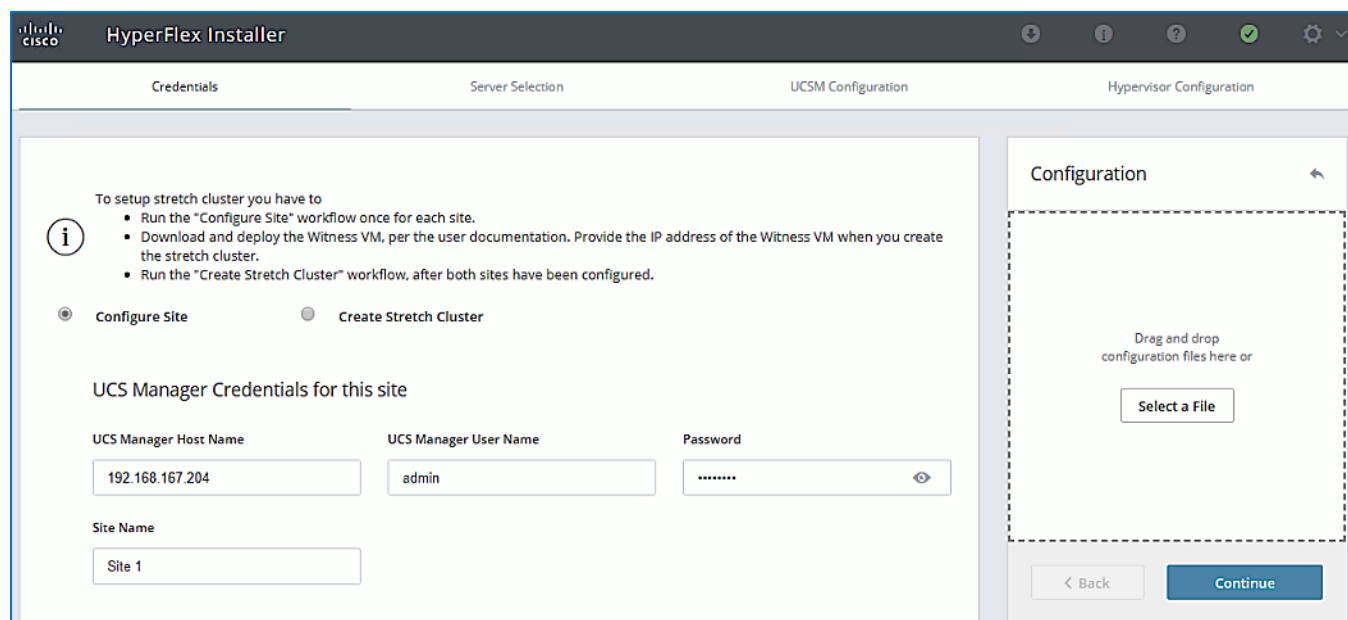
1. From the HyperFlex Installer/Configuration Workflow page, for the **Select a Workflow**, click **Create Cluster** and from the drop-down list, select **Stretch Cluster**.



- In the **Credentials** screen, select the radio button for **Configure Site**. For **Site 1**, specify the **Cisco UCS Manager Hostname** or **IP address**, the log in credentials and the **Site Name** (Site 1). The site name will be the name of the physical site in the Cisco HyperFlex Connect used to manage the cluster.



If you have a JSON configuration file saved from a previous attempt to configure Site 1, you may click Select a File from the box on the right side of the window to select the JSON configuration file and click Use Configuration to populate the fields for configuring this site. The installer does not save passwords.



- Click Continue.

4. In the **Server Selection** screen, select the unassociated servers that should be part of **Site 1** in the stretched cluster.



The Fabric Interconnect ports that connect to HyperFlex servers were enabled in the [Solution Deployment – Setup Cisco UCS Domains](#) section. You can also choose to enable it here by clicking on Configure Server Ports at the top. However, the servers will go through a discovery process that takes a significant amount of time and you will not have control of the server number order.

HyperFlex Installer

Credentials | **Server Selection** | UCSM Configuration | Hypervisor Configuration

Server Selection Configure Server Ports Refresh

Select Nodes for this site.

Unassociated (4) Associated (0)

<input checked="" type="checkbox"/>		Server Name ^	Status	Model	Serial	Actions
<input checked="" type="checkbox"/>		Server 1	unassociated	HX220C-M5SX	WZP22060AU8	none
<input checked="" type="checkbox"/>		Server 2	unassociated	HX220C-M5SX	WZP220607LD	none
<input checked="" type="checkbox"/>		Server 3	unassociated	HX220C-M5SX	WZP22060ATL	none
<input checked="" type="checkbox"/>		Server 4	unassociated	HX220C-M5SX	WZP22060ATU	none

Configuration

Credentials

UCS Manager Host Name 192.168.167.204

UCS Manager User Name admin

Site Name Site 1

< Back Continue

5. Click Continue.
6. In the **UCSM Configuration** screen, specify the UCSM related configuration for **Site 1** as shown below.

HyperFlex Installer

Credentials | Server Selection | **UCSM Configuration** | Hypervisor Configuration

VLAN Configuration

VLAN for Hypervisor and HyperFlex management

VLAN Name: VLAN ID:

VLAN for HyperFlex storage traffic

VLAN Name: VLAN ID:

VLAN for VM vMotion

VLAN Name: VLAN ID:

VLAN for VM Network

VLAN Name: VLAN ID(s):

MAC Pool

MAC Pool Prefix:

'hx-ext-mgmt' IP Pool for Cisco IMC

IP Blocks: Subnet Mask: Gateway:

Cisco IMC access management (Out of band or Inband)

☒ Out of band ☐ In band

> iSCSI Storage

> FC Storage

Advanced

UCS Server Firmware Version: HyperFlex Cluster Name: Org Name:

Configuration Summary

Credentials

UCS Manager Host Name: 192.168.167.204
UCS Manager User Name: admin
Site Name: Site 1

Server Selection

Server 2	WZP220607LD / HX220C-M5SX
Server 4	WZP22060ATU / HX220C-M5SX
Server 1	WZP22060AU8 / HX220C-M5SX
Server 3	WZP22060ATL / HX220C-M5SX

[< Back](#) [Continue](#)

- Enter the **VLAN Names** and **VLAN IDs** that are to be created in Cisco UCS. Multiple VLAN IDs can be specified for the (guest) virtual machine networks.



In this design, the VMware virtual switch that will be created by the Installer for the (guest) virtual machine networks will be migrated to a Cisco ACI controlled Cisco AVE and the VLANs will be dynamically allocated. For this reason, it is not necessary to configure more than one VLAN for the virtual machine network. However, at least one VLAN is required in order to do other configuration for the virtual machine networks such as creating uplink vNICs in Cisco UCS Manager and creating appropriate QoS policies for virtual machine traffic.

8. For the **MAC Pool** prefix, specify the 4th byte (for example: 00:25:B5:**A8**). This prefix must be unique.
9. For the 'hx-ext-mgmt' **IP Pool for Cisco IMC**, specify a unique **IP address** range, **subnet mask** and **gateway** to be used by the CIMC interfaces of the servers in this HX cluster.
10. For the **UCS Firmware Version**, select the version of firmware to be loaded on servers in **Site 1**. The drop-down list shows the versions currently available on Cisco UCS Manager in **Site 1**.
11. For the **HyperFlex Cluster**, for **HyperFlex Cluster Name**, specify a name. For the **Org Name**, specify a unique name. The cluster names in both sites should be the same since both sites are part of a single cluster. The organization name can be the same in both sites of the stretched cluster but only because they're in different UCS domains.



When deploying additional clusters in the same UCS domain, change the VLAN names (even if the VLAN IDs are same), MAC Pool prefix, Cluster Name and Org Name so as to not overwrite the original cluster.

12. Click Continue.
13. In the **Hypervisor Configuration** screen, specify the ESXi Management IP Addresses and Gateway information for the ESXi hosts in **Site 1** as shown below. The default Hypervisor credentials for factory-installed nodes are: `root` with a password of `Cisco123`. The IP addresses will be assigned to the ESXi hosts via Serial over Lan (SoL) from Cisco UCS Manager.

HyperFlex Installer

Credentials Server Selection UCSM Configuration **Hypervisor Configuration**

Configure common Hypervisor Settings

Subnet Mask: Gateway: DNS Server(s):

Hypervisor Settings

☒ Make IP Addresses and Hostnames Sequential

#	Name	Serial	Static IP Address	Hostname
= <input type="radio"/>	Server 1	WZP22060AU8	<input type="text" value="10.1.167.111"/>	<input type="text" value="hvx-d1-esxi-1"/>
= <input type="radio"/>	Server 2	WZP220607LD	<input type="text" value="10.1.167.112"/>	<input type="text" value="hvx-d1-esxi-2"/>
= <input type="radio"/>	Server 3	WZP22060ATL	<input type="text" value="10.1.167.113"/>	<input type="text" value="hvx-d1-esxi-3"/>
= <input type="radio"/>	Server 4	WZP22060ATU	<input type="text" value="10.1.167.114"/>	<input type="text" value="hvx-d1-esxi-4"/>

Hypervisor Credentials

Admin User name: Hypervisor Password:

Configuration

Credentials

UCS Manager Host Name: 192.168.167.204

UCS Manager User Name: admin

Site Name: Site 1

Admin User name: root

Server Selection

Server 2: WZP220607LD / HX220C-M5SX

Server 4: WZP22060ATU / HX220C-M5SX

Server 1: WZP22060AU8 / HX220C-M5SX

Server 3: WZP22060ATL / HX220C-M5SX

UCSM Configuration

VLAN Name: hvx-inband-mgmt

VLAN ID: 118

VLAN Name: hvx-d1-storage-data

VLAN ID: 3218

VLAN Name: hvx-vmotion

VLAN ID: 3018

VLAN Name: hvx-vm-network

VLAN ID(s): 2118

MAC Pool Prefix: 00:25:B5:A8

IP Blocks: 192.168.167.111-114

Subnet Mask: 255.255.255.0

Gateway: 192.168.167.254

VLAN Name: hx-inband-clmc

UCS Server Firmware Version: 4.0(1c)

HyperFlex Cluster Name: HXV-Cluster1

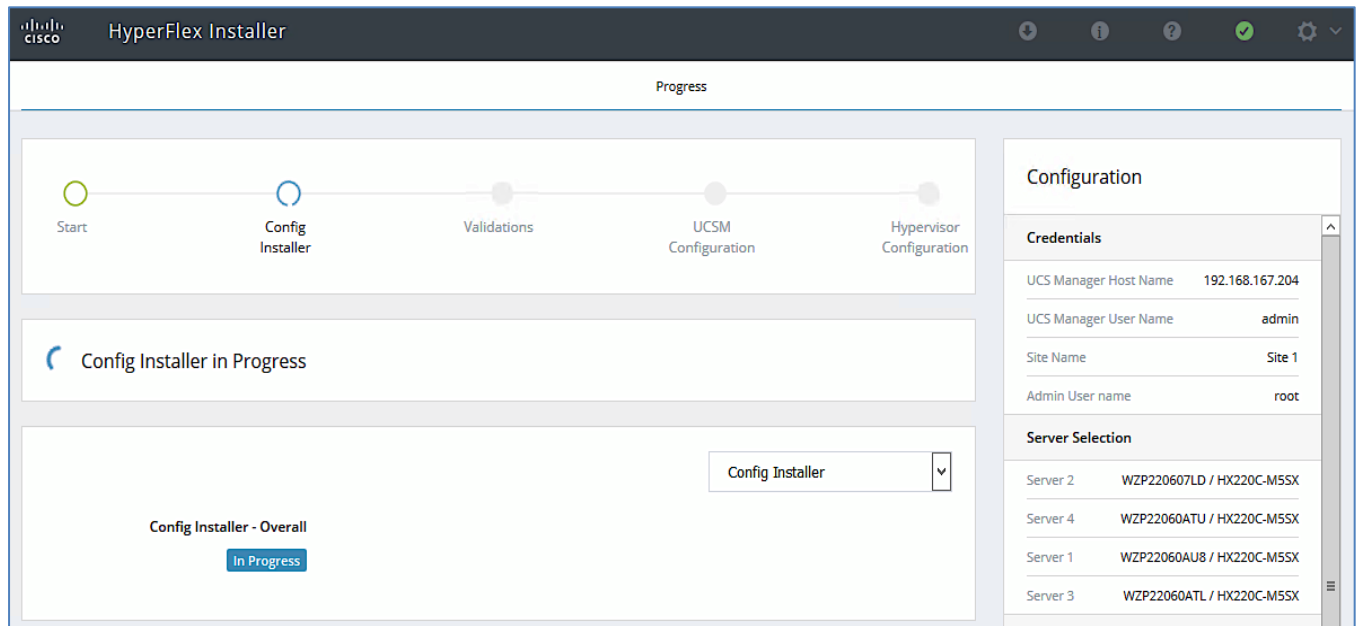
Org Name: HXV-Org1

iSCSI Storage: false

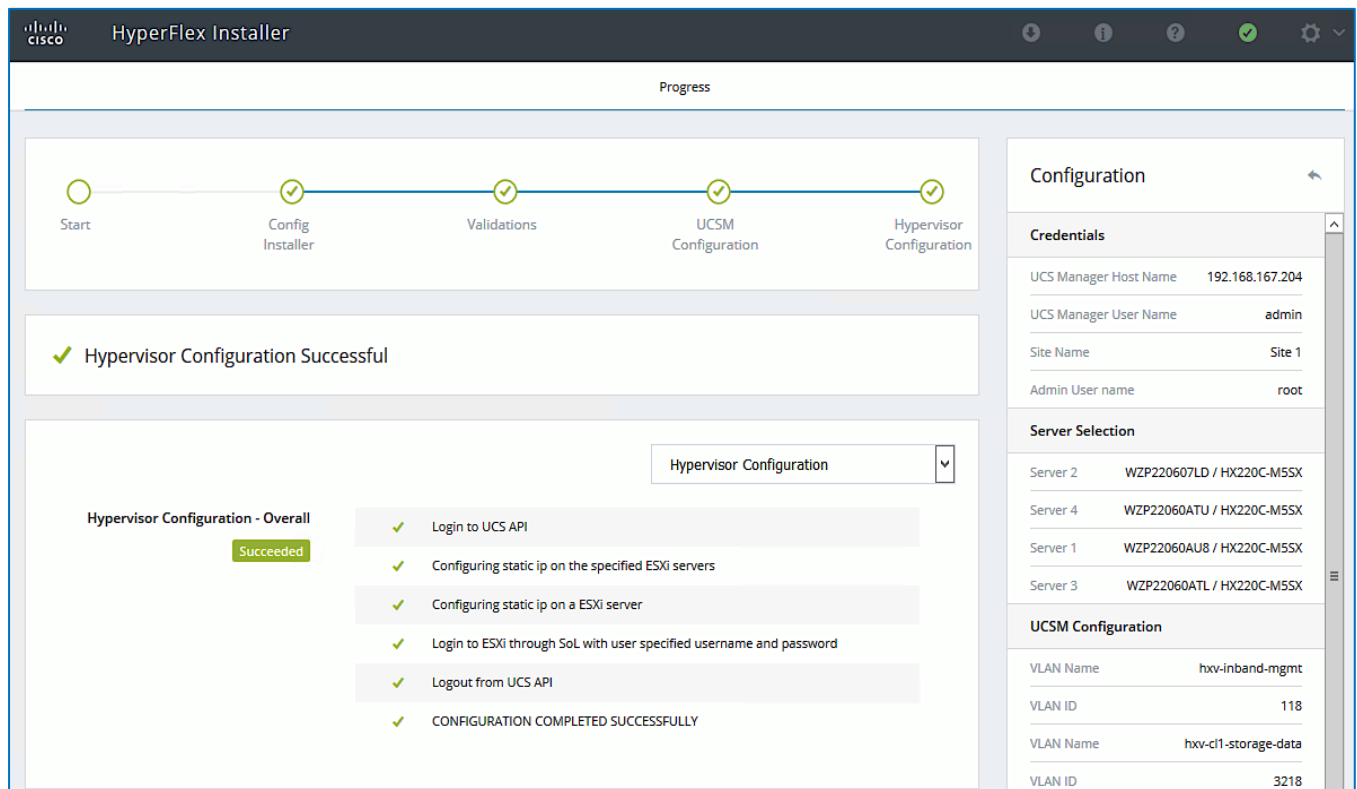
VLAN A Name: hx-ext-storage-iscsi-a

[< Back](#) [Configure Site](#)

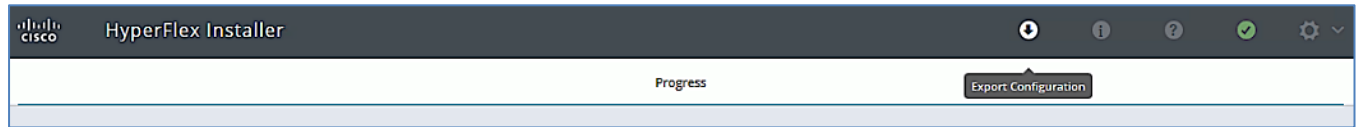
14. Click **Configure Site** to start configuring **Site 1**. The wizard will step through the configuration stages and provide the status for specific configuration completed as shown below.



If the configuration is successful, you will see a screen similar to the one shown below.



- Export the **Site 1** configuration by clicking the down arrow icon in the top right of the screen. Click OK to save the configuration to a JSON file. This file can be used to rebuild the same cluster in the future, and as a record of the configuration options and settings used during the installation.

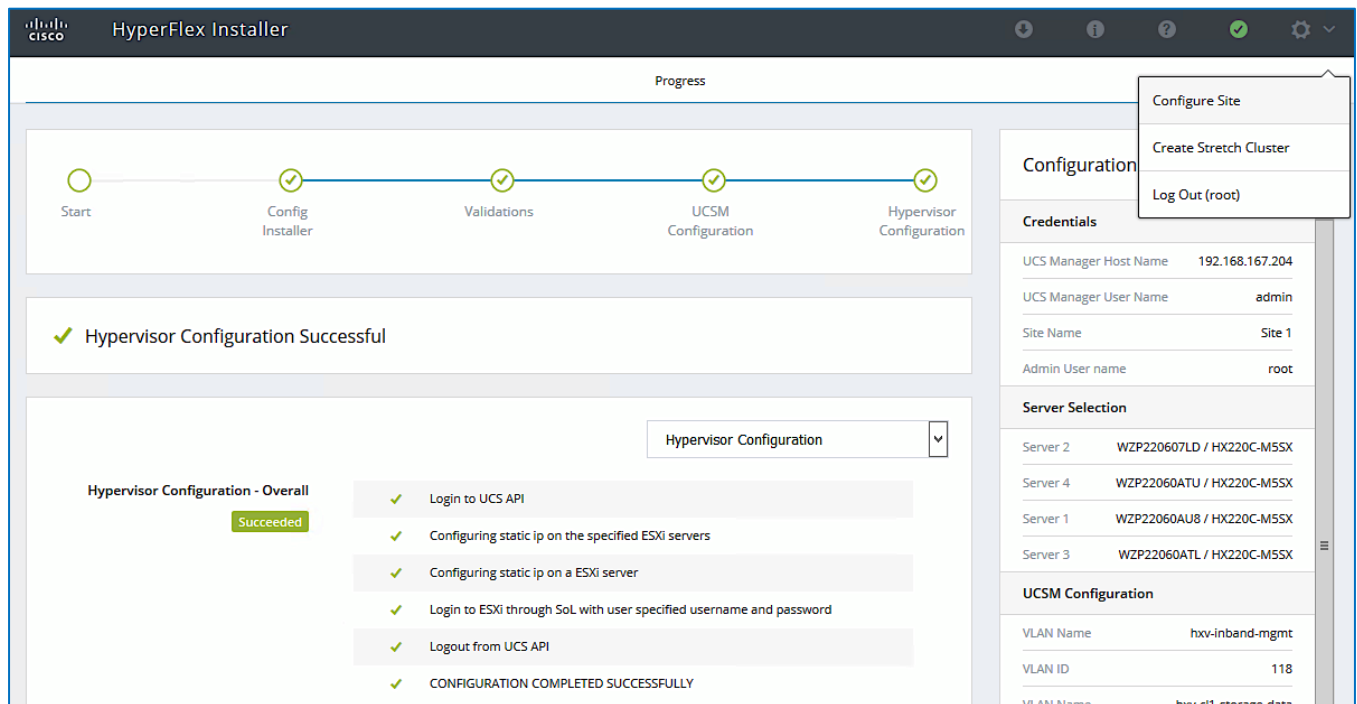


16. Proceed to the next section to **Configure Site 2**.

Configure Site 2 from Deployment Wizard

To configure the second site (**Site 2**) in the stretched cluster, follow these steps:

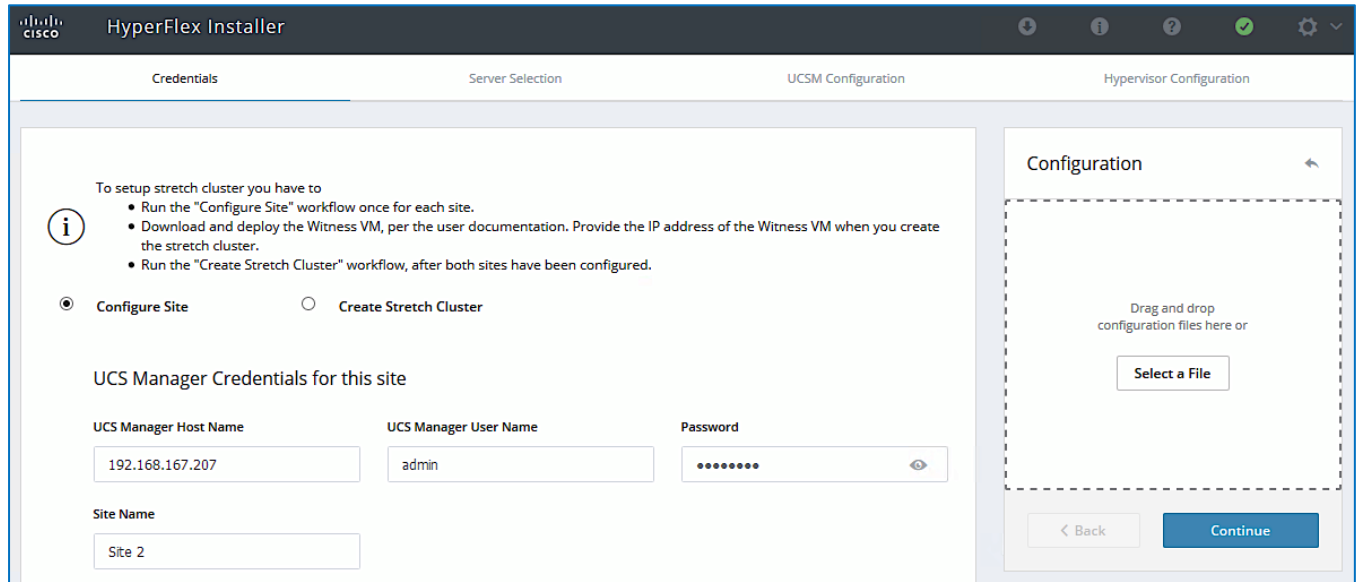
1. From the HyperFlex Installer/Configuration wizard, go to the wheel icon in the top right of the window and select **Configure Site** from the drop-down list.



2. In the **Credentials** screen, select the radio button for **Configure Site**. For **Site 2**, specify the **Cisco UCS Manager Hostname** or IP address, the log in credentials and the **Site Name** (Site 2). The site name will be the name of the physical site in the Cisco HyperFlex Connect used to manage the cluster.



If you have a JSON configuration file saved from a previous attempt to configure Site 2, you may click Select a File from the box on the right side of the window to select the JSON configuration file and click Use Configuration to populate the fields for configuring this site. Installer does not save passwords.



The screenshot shows the 'Credentials' tab of the HyperFlex Installer. It provides instructions for setting up a stretch cluster and offers two options: 'Configure Site' (selected) and 'Create Stretch Cluster'. Below, it prompts for UCS Manager credentials for 'Site 2', including Host Name, User Name, Password, and Site Name.

To setup stretch cluster you have to

- Run the "Configure Site" workflow once for each site.
- Download and deploy the Witness VM, per the user documentation. Provide the IP address of the Witness VM when you create the stretch cluster.
- Run the "Create Stretch Cluster" workflow, after both sites have been configured.

☒ **Configure Site** ☐ **Create Stretch Cluster**

UCS Manager Credentials for this site

UCS Manager Host Name: 192.168.167.207

UCS Manager User Name: admin

Password: [masked]

Site Name: Site 2

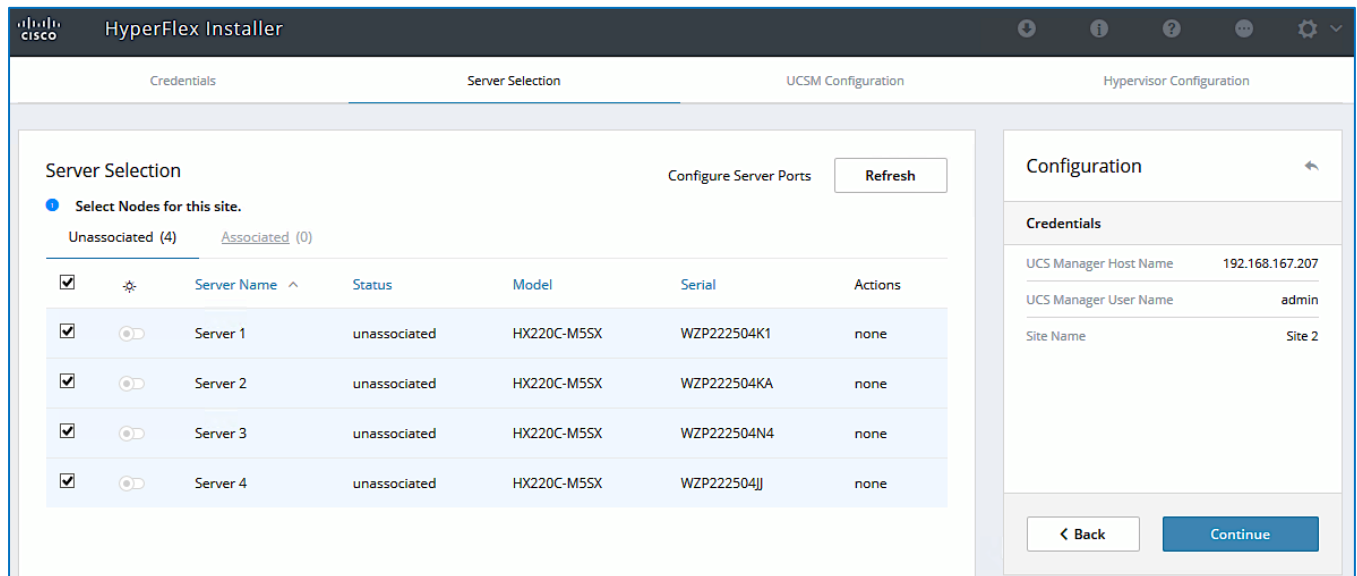
Configuration panel on the right: Drag and drop configuration files here or [Select a File](#)

Buttons: < Back, Continue

3. Click Continue.
4. In the **Server Selection** screen, select the servers that should be part of **Site 2** in the stretched cluster.



The Fabric Interconnect ports that connect to HyperFlex servers were enabled in the [Solution Deployment – Setup Cisco UCS Domains](#) section. You can also choose to enable it here by clicking Configure Server Ports at the top. However, the servers will go through a discovery process that takes a significant amount of time and you will not have control of the server number order.



The screenshot shows the 'Server Selection' tab. It displays a table of unassociated servers for selection. A 'Configure Server Ports' button and a 'Refresh' button are also visible. The right sidebar shows the 'Credentials' section with the same information as the previous screen.

Server Selection

Select Nodes for this site.

Unassociated (4) Associated (0)

		Server Name ^	Status	Model	Serial	Actions
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Server 1	unassociated	HX220C-M5SX	WZP222504K1	none
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Server 2	unassociated	HX220C-M5SX	WZP222504KA	none
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Server 3	unassociated	HX220C-M5SX	WZP222504N4	none
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Server 4	unassociated	HX220C-M5SX	WZP222504JJ	none

Buttons: Configure Server Ports, Refresh

Configuration panel on the right: Credentials

UCS Manager Host Name: 192.168.167.207

UCS Manager User Name: admin

Site Name: Site 2

Buttons: < Back, Continue

5. Click Continue.
6. In the **UCSM Configuration** screen, specify the UCSM related configuration for **Site 2** as shown below.

HyperFlex Installer

Credentials | Server Selection | **UCSM Configuration** | Hypervisor Configuration

VLAN Configuration

VLAN for Hypervisor and HyperFlex management

VLAN Name	VLAN ID
hvx-inband-mgmt	118

VLAN for HyperFlex storage traffic

VLAN Name	VLAN ID
hvx-cl1-storage-data	3218

VLAN for VM vMotion

VLAN Name	VLAN ID
hvx-vmotion	3018

VLAN for VM Network

VLAN Name	VLAN ID(s)
hvx-vm-network	2118

MAC Pool

MAC Pool Prefix

00:25:B5:A9

'hx-ext-mgmt' IP Pool for Cisco IMC

IP Blocks	Subnet Mask	Gateway
192.168.167.115-118	255.255.255.0	192.168.167.254

Cisco IMC access management (Out of band or Inband)

☒ Out of band ☐ In band

> iSCSI Storage

> FC Storage

Advanced

UCS Server Firmware Version	HyperFlex Cluster Name	Org Name
4.0(1c)	HXV-Cluster1	HXV-Org1

Configuration

Credentials

UCS Manager Host Name	192.168.167.207
UCS Manager User Name	admin
Site Name	Site 2

Server Selection

Server 2	WZP222504KA / HX220C-M5SX
Server 3	WZP222504N4 / HX220C-M5SX
Server 1	WZP222504K1 / HX220C-M5SX
Server 4	WZP222504JJ / HX220C-M5SX

[< Back](#) [Continue](#)

- Enter the **VLAN Names** and **VLAN IDs** that are to be created in Cisco UCS. Multiple VLAN IDs can be specified for the (guest) virtual machine networks.



In this design, the VMware virtual switch that will be created by the Installer for the (guest) virtual machine networks will be migrated to a Cisco ACI controlled Cisco AVE and the VLANs will be dynamically allocated. For this reason, it is not necessary to configure more than one VLAN for the virtual machine network. However, at least one VLAN is required in order to do other configuration for the virtual machine networks such as creating uplink vNICs in Cisco UCS Manager and creating appropriate QoS policies for VM traffic.

- For the **MAC Pool** prefix, specify the 4th byte, for example: 00:25:B5:**A9**. This prefix must be unique.

9. For the '**hx-ext-mgmt**' IP Pool for Cisco IMC, specify a **unique IP address** range, **subnet mask** and **gateway** to be used by the CIMC interfaces of the servers in this site.
10. For the **UCS Firmware Version**, select the version of firmware to be loaded on servers in **Site 2**. The drop-down list shows the versions currently available on Cisco UCS Manager in **Site 2**.
11. For the **HyperFlex Cluster**, specify a name. For the **Org Name**, specify a **unique** name. The cluster names in both sites should be the same since both sites are part of a single cluster. The organization name can be the same in both sites of the stretched cluster but only because they're in different UCS domains.



When deploying additional clusters in the same UCS domain, change the VLAN names (even if the VLAN IDs are same), MAC Pool prefix, Cluster Name and Org Name so as to not overwrite the original cluster information.

12. Click Continue.
13. In the **Hypervisor Configuration** screen, specify the ESXi Management IP Addresses and Gateway information for the ESXi hosts in **Site 2** as shown below. The default Hypervisor credentials for factory-installed nodes are: `root` with a password of `Cisco123`. The IP addresses will be assigned to the ESXi hosts via Serial over Lan (SoL) from Cisco UCS Manager.

HyperFlex Installer

Configure common Hypervisor Settings

Subnet Mask: 255.255.255.0 | Gateway: 10.1.167.254 | DNS Server(s): 10.99.167.244, 10.99.167.245

Hypervisor Settings

☒ Make IP Addresses and Hostnames Sequential

Name	Serial	Static IP Address	Hostname
Server 1	WZP222504K1	10.1.167.115	hxv-cl1-esxi-5
Server 2	WZP222504KA	10.1.167.116	hxv-cl1-esxi-6
Server 3	WZP222504N4	10.1.167.117	hxv-cl1-esxi-7
Server 4	WZP222504JJ	10.1.167.118	hxv-cl1-esxi-8

Hypervisor Credentials

Admin User name: root | Hypervisor Password: [masked]

Configuration

Credentials

UCS Manager Host Name: 192.168.167.207
UCS Manager User Name: admin
Site Name: Site 2
Admin User name: root

Server Selection

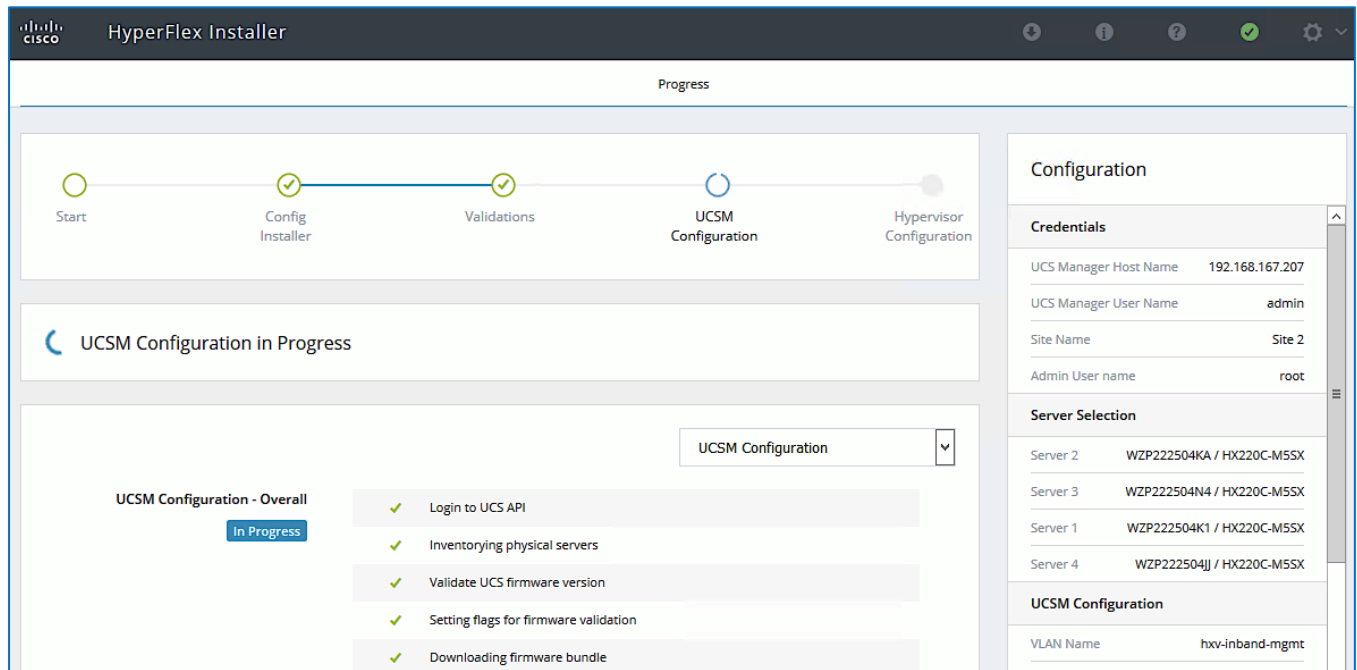
Server 2: WZP222504KA / HX220C-M5SX
Server 3: WZP222504N4 / HX220C-M5SX
Server 1: WZP222504K1 / HX220C-M5SX
Server 4: WZP222504JJ / HX220C-M5SX

UCSM Configuration

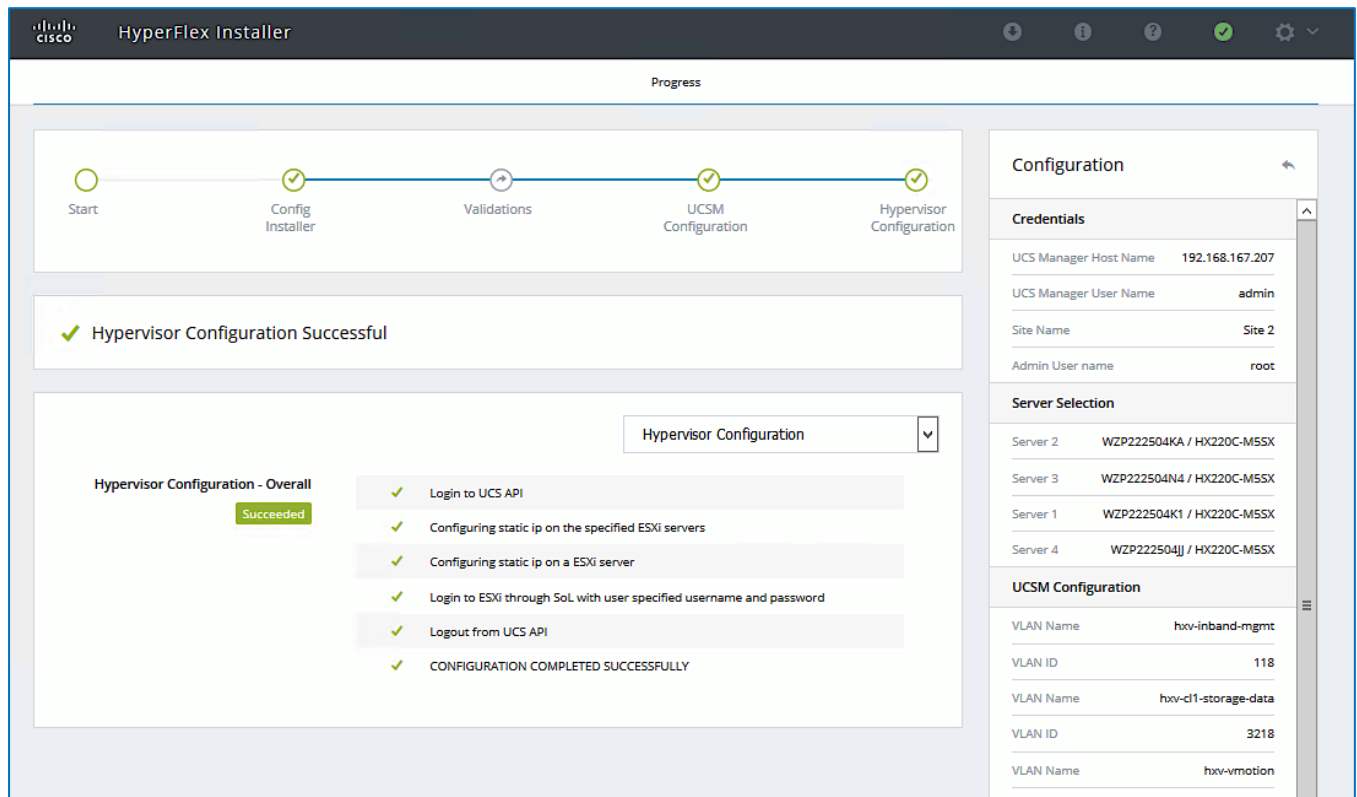
VLAN Name: hxv-inband-mgmt
VLAN ID: 118
VLAN Name: hxv-cl1-storage-data
VLAN ID: 3218
VLAN Name: hxv-vmotion
VLAN ID: 3018
VLAN Name: hxv-vm-network
VLAN ID(s): 2118
MAC Pool Prefix: 00:25:B5:A9

[< Back](#) [Configure Site](#)

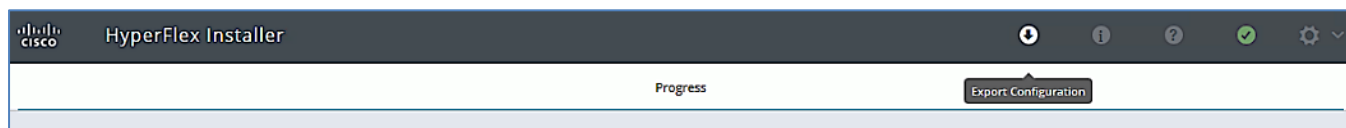
14. Click **Configure Site** to start configuring **Site 2**. The wizard will step through the configuration stages and provide the status for specific configuration completed as shown below.



15. If the configuration is successful, you will see a screen similar to the one below.



16. Export the **Site 2** configuration by clicking the down arrow icon in the top right of the screen. Click OK to save the configuration to a JSON file. This file can be used to rebuild the same cluster in the future, and as a record of the configuration options and settings used during the installation.



17. Proceed to the next section to **Deploy Witness Virtual Machine** at a third site.

Deploy Witness Virtual Machine in a Third Site

To achieve quorum in a HyperFlex stretched cluster, a Witness virtual machine is necessary. The Witness virtual machine should be deployed in a third site and must be reachable from all sites in a HyperFlex stretched cluster. In this design, the Witness virtual machine is deployed in an existing network outside the ACI Multi-Pod Fabric.

Table 76 Setup Information

Witness VM - IP Address/Subnet Mask	10.99.167.249/24
Gateway	10.99.167.254 (outside the ACI Fabric)
DNS	10.99.167.244, 10.99.167.245
NTP	192.168.167.254

To deploy the Witness virtual machine for the HyperFlex stretched cluster, follow these steps:

1. Use a browser to navigate to the VMware vCenter server that will be used to deploy the Witness virtual machine will be deployed.
2. Click the vSphere Web Client of your choice. Log in using an **Administrator** account.
3. From the vSphere Web Client, navigate to **Home > Hosts and Clusters**.
4. From the left navigation pane, select the **Datacenter > Cluster** and right-click to select **Deploy OVF Template....**
5. In the **Deploy OVF Template** wizard, for Select Template, select **Local file** and click the **Browse** button to locate and open the **HyperFlex-Witness-1.0.2.ova** file, click the file and click **Open**. Click **Next**.
6. Modify the name of the virtual machine to be created if desired and click a folder location to place the virtual machine. Click **Next**.
7. Click a specific host or **cluster** to locate the virtual machine. Click **Next**.
8. After the file validation, review the details. Click **Next**.
9. Select a **Thin provision virtual disk format**, and the **datastore** to store the new virtual machine. Click **Next**.
10. Modify the network port group selection from the drop-down list in the **Destination Networks** column, choosing the network the witness VM will communicate on. Click **Next**.
11. Enter the static address settings to be used, fill in the fields for the **Witness Node's IP Address and Mask**, **DNS** server, **Default Gateway**, and **NTP** Server info.

Deploy OVF Template

- ✓ 1 Select an OVF template
- ✓ 2 Select a name and folder
- ✓ 3 Select a compute resource
- ✓ 4 Review details
- ✓ 5 Select storage
- ✓ 6 Select networks
- 7 Customize template**
- 8 Ready to complete

Customize template

Customize the deployment properties of this software solution.

All properties have valid values

Networking Properties	5 settings
Network 1 IP Address	The IP address for this interface. Leave blank if DHCP is desired. <input type="text" value="10.99.167.249"/>
Network 1 Netmask	The netmask or prefix for this interface. Leave blank if DHCP is desired. <input type="text" value="255.255.255.0"/>
Default Gateway	The default gateway address for this VM. Leave blank if DHCP is desired. <input type="text" value="10.99.167.254"/>
DNS	The domain name servers for this VM (comma separated). Leave blank if DHCP is desired. <input type="text" value="10.99.167.244, 10.99.167.2"/>
NTP	NTP servers for this VM (comma separated) to sync time. <input type="text" value="192.168.167.254"/>

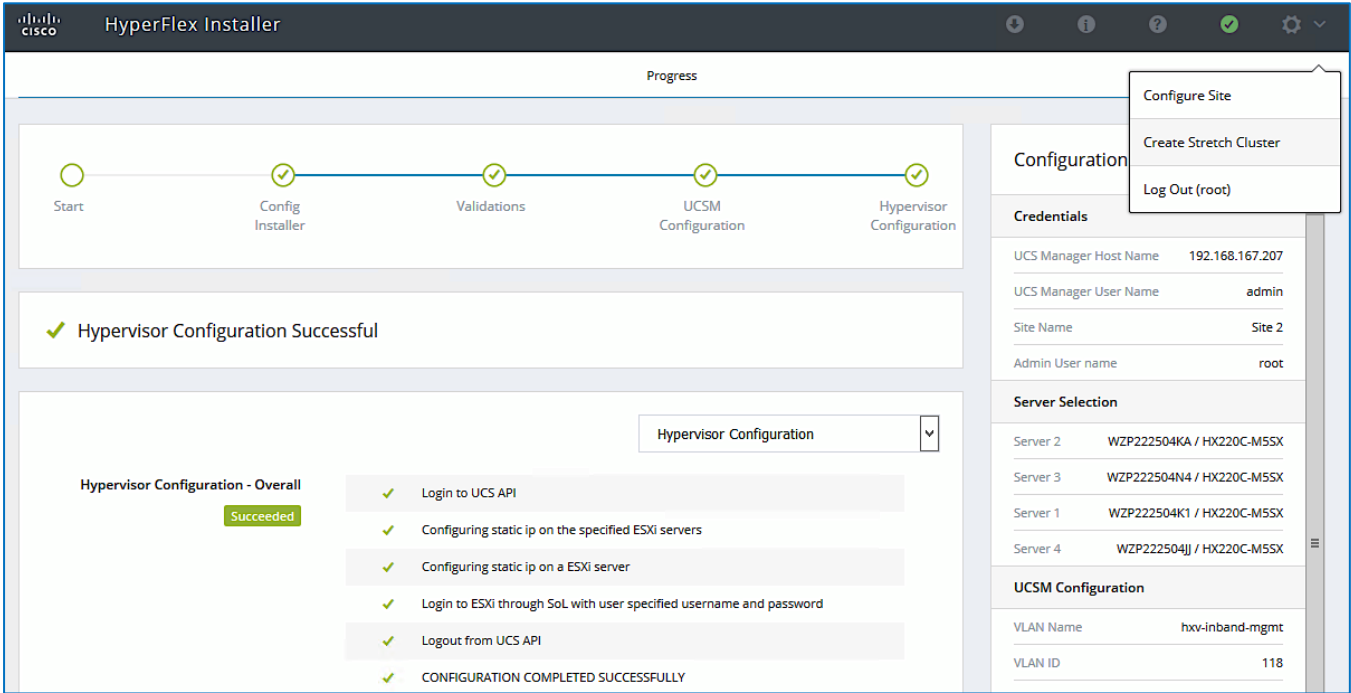
CANCEL
BACK
NEXT

- Click Next.
- Review the final configuration and click **Finish**. The witness VM will take a few minutes to deploy, once it has deployed, power on the new VM.
- Proceed to the next section to create a stretch HyperFlex cluster.

Create Stretch Cluster from Deployment Wizard

To create the stretched cluster using **Site 1** and **Site 2**, follow these steps:

- From the HyperFlex Installer/Configuration Wizard, go to the wheel icon in the top right of the window and select **Create Stretch Cluster** from the drop-down list.



2. In the **Credentials** screen, select the radio button for **Create Stretch Cluster**. For **Site 1** and **Site 2**, specify the **Cisco UCS Manager Credentials** (Hostname or IP address, username and password), **VMware vCenter Credentials** (for the vCenter managing the stretch cluster), and **Hypervisor Credentials** as shown below.



If you have a JSON configuration file saved from a previous attempt for Create Stretch Cluster, you may click Select a File from the box on the right side of the window to select the JSON configuration file and click Use Configuration to populate the fields for configuring this site. The installer does not save passwords.

HyperFlex Installer

Credentials Server Selection IP Addresses Cluster Configuration

Configuration

To setup stretch cluster you have to

- Run the "Configure Site" workflow once for each site.
- Download and deploy the Witness VM, per the user documentation. Provide the IP address of the Witness VM when you create the stretch cluster.
- Run the "Create Stretch Cluster" workflow, after both sites have been configured.

☐ Configure Site ☒ Create Stretch Cluster

UCS Manager Credentials for Site 1

UCS Manager Host Name: 192.168.167.204 User Name: admin Password: [masked]

Site Name: Site 1 Org Name: HXV-Org1

UCS Manager Credentials for Site 2

UCS Manager Host Name: 192.168.167.207 User Name: admin Password: [masked]

Site Name: Site 2 Org Name: HXV-Org1

vCenter Credentials

vCenter Server: hxv0-vcsa.hxv.com User Name: administrator@hxv.com Admin Password: [masked]

Hypervisor Credentials

Admin User name: root

☒ The hypervisor on this node uses the factory default password

☐ You are required to change the factory default password. Enter a new password for the hypervisor

New Password: [masked] Confirm New Password: [masked]

Drag and drop configuration files here or
[Select a File](#)

[< Back](#) [Continue](#)

3. Click Continue.
4. In the **Server Selection** screen, select the servers from **Site 1** and **Site 2** that should be part of the stretched cluster.

HyperFlex Installer

Credentials | **Server Selection** | IP Addresses | Cluster Configuration

Server Selection

Select Nodes for this site.

Associated (8)

<input checked="" type="checkbox"/>		Server Name	Site	Status	Model	Serial	Service Profile	Actions
<input checked="" type="checkbox"/>		Server 2	Site 1	ok	HX220C-M5SX	WZP220607LD	org-root/org-HXV-Org1/ls-rack-unit-2	Actions ▾
<input checked="" type="checkbox"/>		Server 3	Site 1	ok	HX220C-M5SX	WZP22060ATL	org-root/org-HXV-Org1/ls-rack-unit-3	Actions ▾
<input checked="" type="checkbox"/>		Server 1	Site 1	ok	HX220C-M5SX	WZP22060AU8	org-root/org-HXV-Org1/ls-rack-unit-1	Actions ▾
<input checked="" type="checkbox"/>		Server 4	Site 1	ok	HX220C-M5SX	WZP22060ATU	org-root/org-HXV-Org1/ls-rack-unit-4	Actions ▾
<input checked="" type="checkbox"/>		Server 2	Site 2	ok	HX220C-M5SX	WZP222504KA	org-root/org-HXV-Org1/ls-rack-unit-2	Actions ▾
<input checked="" type="checkbox"/>		Server 3	Site 2	ok	HX220C-M5SX	WZP222504N4	org-root/org-HXV-Org1/ls-rack-unit-3	Actions ▾
<input checked="" type="checkbox"/>		Server 1	Site 2	ok	HX220C-M5SX	WZP222504K1	org-root/org-HXV-Org1/ls-rack-unit-1	Actions ▾
<input checked="" type="checkbox"/>		Server 4	Site 2	ok	HX220C-M5SX	WZP222504JJ	org-root/org-HXV-Org1/ls-rack-unit-4	Actions ▾

Configure Server Ports Refresh

Configuration

Credentials

UCS Manager Host Name 1 192.168.167.204

User Name admin

UCS Manager Host Name 2 192.168.167.207

User Name admin

Site Name Site 1

Org Name 1 HXV-Org1

Site Name Site 2

Org Name 2 HXV-Org1

vCenter Server hvx0-vcsa.hvx.com

User Name administrator@hvx.com

Admin User name root

Back Continue

- Click Continue.
- In the **IP Addresses** screen, specify the IP addresses for the cluster (ESXi host and Storage Controller VM's **Management IP Addresses**, ESXi host and Storage Controller VM's **Storage Data Network IP Addresses**, **Cluster IP Addresses** for Management and Storage Data, **Gateway** for Management Subnet and **Witness Node IP Address**) as shown below.



A default gateway is not required for the data network, as those interfaces normally will not communicate with any other hosts or networks, and the subnet can be non-routable.

HyperFlex Installer

Credentials Server Selection **IP Addresses** Cluster Configuration

IP Addresses

☒ Make IP Addresses Sequential

				Management - VLAN		Data - VLAN (FQDN or IP Address)	
It	Name	Site	Hypervisor	Storage Controller	Hypervisor	Storage Controller	
Server 1	Site 1	10.1.167.111	10.1.167.161	72.1.167.111	72.1.167.161		
Server 2	Site 1	10.1.167.112	10.1.167.162	172.1.167.111	172.1.167.161		
Server 3	Site 1	10.1.167.113	10.1.167.163	172.1.167.111	172.1.167.161		
Server 4	Site 1	10.1.167.114	10.1.167.164	172.1.167.111	172.1.167.161		
Server 1	Site 2	10.1.167.115	10.1.167.165	172.1.167.111	172.1.167.161		
Server 3	Site 2	10.1.167.117	10.1.167.167	172.1.167.111	172.1.167.161		
Server 2	Site 2	10.1.167.116	10.1.167.166	172.1.167.111	172.1.167.161		
Server 4	Site 2	10.1.167.118	10.1.167.168	72.1.167.118	72.1.167.168		

	Management	Data
Cluster IP Address	10.1.167.110	172.1.167.110
Subnet Mask	255.255.255.0	255.255.255.0
Gateway	10.1.167.254	
Witness IP	10.99.167.249	

Configuration

Credentials

UCS Manager Host Name 1	192.168.167.204
User Name	admin
UCS Manager Host Name 2	192.168.167.207
User Name	admin
Site Name	Site 1
Org Name 1	HXV-Org1
Site Name	Site 2
Org Name 2	HXV-Org1
vCenter Server	hvx0-vcsa.hvx.com
User Name	administrator@hvx.com
Admin User name	root

Server Selection

Server 2	WZP220607LD / HX220C-M5SX
Server 3	WZP22060ATL / HX220C-M5SX
Server 1	WZP22060AU8 / HX220C-M5SX
Server 4	WZP22060ATU / HX220C-M5SX
Server 2	WZP222504KA / HX220C-M5SX
Server 3	WZP222504N4 / HX220C-M5SX
Server 1	WZP222504K1 / HX220C-M5SX
Server 4	WZP222504JJ / HX220C-M5SX

[< Back](#) [Continue](#)

- Click Continue.
- In the **Cluster Configuration** screen, specify a **name** for the HyperFlex Cluster, the **Replication Factor** to use, Storage Controller VM (SCVM) **Credentials**, VMware vCenter configuration (**Datacenter**, **Cluster**), Services (**DNS**, **NTP**, **Domain Name**, **Timezone**) and Networking (**Management**, **Storage Data**, **Jumbo Frames**, and so on).

HyperFlex Installer

Credentials Server Selection IP Addresses **Cluster Configuration**

Cisco HX Cluster

Cluster Name: Replication Factor:

Controller VM

Create Admin Password: Confirm Admin Password:

vCenter Configuration

vCenter Datacenter Name: vCenter Cluster Name:

System Services

DNS Server(s): NTP Server(s): DNS Domain Name:

Time Zone:

Auto Support

Auto Support: ☒ Enable Connected Services (Recommended) Send service ticket notifications to:

Advanced Networking

Management VLAN Tag - Site 1: Management VLAN Tag - Site 2: Management vSwitch:

Data VLAN Tag - Site 1: Data VLAN Tag - Site 2: Data vSwitch:

Configuration

Credentials

UCS Manager Host Name 1	192.168.167.204
User Name	admin
UCS Manager Host Name 2	192.168.167.207
User Name	admin
Site Name	Site 1
Org Name 1	HXV-Org1
Site Name	Site 2
Org Name 2	HXV-Org1
vCenter Server	hxv0-vcsa.hxv.com
User Name	administrator@hxv.com
Admin User name	root

Server Selection

Server 2	WZP220607LD / HX220C-M5SX
Server 3	WZP22060ATL / HX220C-M5SX
Server 1	WZP22060AU8 / HX220C-M5SX
Server 4	WZP22060ATU / HX220C-M5SX
Server 2	WZP222504KA / HX220C-M5SX
Server 3	WZP222504N4 / HX220C-M5SX
Server 1	WZP222504K1 / HX220C-M5SX
Server 4	WZP222504JJ / HX220C-M5SX

IP Addresses

Cluster Name	HXV-Cluster1
Management Cluster	10.1.167.110
Data Cluster	172.1.167.110
Management Subnet Mask	255.255.255.0
Data Subnet Mask	255.255.255.0
Management Gateway	10.1.167.254
Witness IP	10.99.167.249

Advanced Configuration

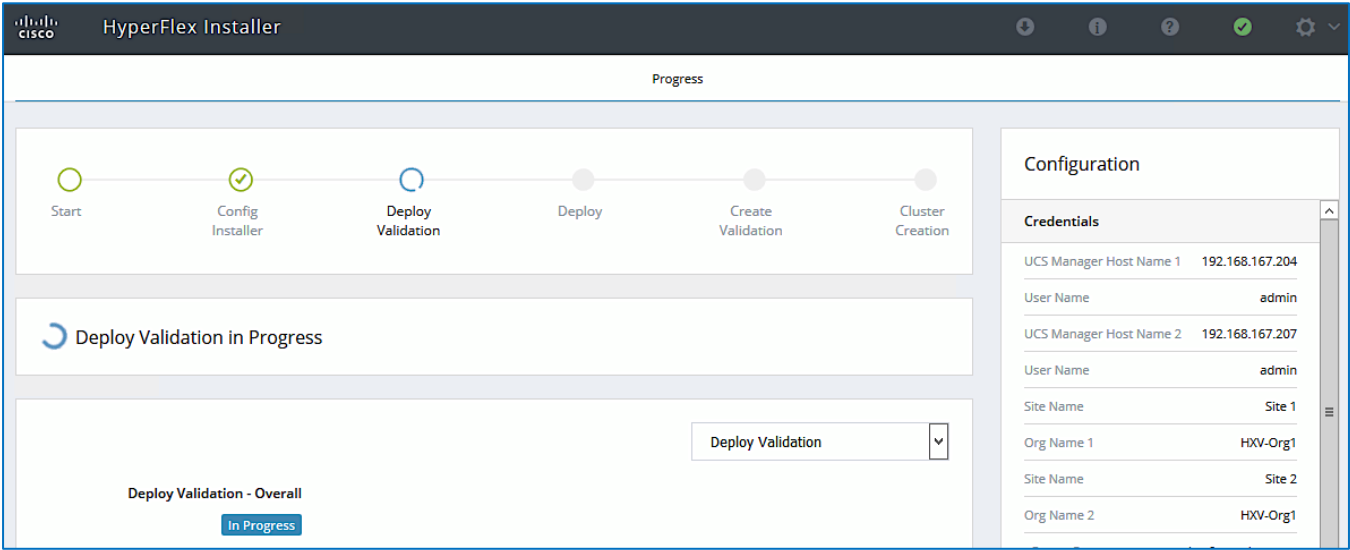
Jumbo Frames
☒ Enable Jumbo Frames on Data Network

Disk Partitions
☐ Clean up disk partitions

Virtual Desktop (VDI)
☐ Optimize for VDI only deployment

vCenter Single-Sign-On Server

- Click **Start** to start the creation of the stretched cluster. The wizard will step through the configuration stages and provide the status for specific configuration completed as shown below.



10. If the configuration is successful, you will see a screen similar to the one below.

The screenshot shows the 'HyperFlex Installer' window with the 'Summary' tab selected. The cluster name is 'HXV-Cluster1' and its status is 'ONLINE' and 'HEALTHY'. The configuration details are as follows:

Property	Value	Property	Value
Version	3.5.1a-31118	vCenter Server	hxxv0-vcsa.hxxv.com
Cluster Management IP Address	10.1.167.110	vCenter Datacenter Name	HXXV-APP
Cluster Data IP Address	172.1.167.110	vCenter Cluster Name	HXXV-Cluster1
Replication Factor	4	DNS Server(s)	10.99.167.245, 10.99.167.244
Available Capacity	12.1 TB	NTP Server(s)	192.168.167.254

Site Info

Property	Value	Property	Value
Name for Site 1	Site 1	Name for Site 2	Site 2
Org Name for Site 1	HXXV-Org1	Org Name for Site 2	HXXV-Org1

Servers

Model	Serial Number	Management Hypervisor	Management Storage Controller	Data Network Hypervisor	Data Network Storage Controller
HX220C-M5SX	WZP22060AU8	10.1.167.111	10.1.167.161	172.1.167.111	172.1.167.161
HX220C-M5SX	WZP220607LD	10.1.167.112	10.1.167.162	172.1.167.112	172.1.167.162
HX220C-M5SX	WZP22060ATL	10.1.167.113	10.1.167.163	172.1.167.113	172.1.167.163
HX220C-M5SX	WZP22060ATU	10.1.167.114	10.1.167.164	172.1.167.114	172.1.167.164
HX220C-M5SX	WZP222504K1	10.1.167.115	10.1.167.165	172.1.167.115	172.1.167.165
HX220C-M5SX	WZP222504KA	10.1.167.116	10.1.167.166	172.1.167.116	172.1.167.166
HX220C-M5SX	WZP222504N4	10.1.167.117	10.1.167.167	172.1.167.117	172.1.167.167
HX220C-M5SX	WZP222504JJ	10.1.167.118	10.1.167.168	172.1.167.118	172.1.167.168

At the bottom right, there are two buttons: 'Back to Workflow Selection' and 'Launch HyperFlex Connect'.

- Export the cluster configuration by clicking the down arrow icon in the top right of the screen. Click OK to save the configuration to a JSON file. This file can be used to rebuild the same cluster in the future, and as a record of the configuration options and settings used during the installation.

The screenshot shows the 'HyperFlex Installer' window with the 'Progress' bar at the bottom. A button labeled 'Export Configuration' is highlighted, indicating the next step in the process.

- Process to the next section to complete the post-installation tasks – run the **post_install** script to create the vMotion interfaces, additional guest virtual machine port groups (optional), and to enable HA and DRS in the cluster.



For stretched clusters, it is very important to review the **DRS Site Affinity rules** to verify that it is setup correctly.

Complete Post-Installation Tasks

When the installation is complete, additional best-practices and configuration can be implemented using a Cisco provided post-install script. The script should be run before deploying virtual machine workloads on the cluster. The script is executed from the Installer virtual machine and can do the following:

- License the hosts in VMware vCenter
- Enable HA/DRS on the cluster in VMware vCenter
- Suppress SSH/Shell warnings in VMware vCenter
- Configure vMotion in VMware vCenter
- Enables configuration of additional guest VLANs/port-groups
- Send test Auto Support (ASUP) email if enabled during the install process
- Perform HyperFlex Health check

To run the post-installation script, follow these steps:

1. SSH into a HyperFlex Installer virtual machine used to deploy the cluster. Log in using the **admin (or root)** account.
2. From the Controller virtual machine, run the command to execute the post-install script: `post_install.py`
3. Follow the on-screen prompts to complete the post-install configuration as shown below.



Any VLANs created on the HyperFlex cluster and UCSM will need a corresponding configuration in the ACI fabric to enable forwarding for that VLAN within the ACI Fabric.

```

root@HyperFlex-Installer:~# cd
root@HyperFlex-Installer:~# post_install
Logging in to controller 10.1.167.110
HX CVM admin password:
Getting ESX hosts from HX cluster...
vCenter URL: 10.10.167.240
Enter vCenter username (user@domain): administrator@hvx.com
vCenter Password:
Found datacenter HXV-APP
Found cluster HXV-Cluster1
Enter ESX root password:

Enter vSphere license key? (y/n) n

Enable HA/DRS on cluster? (y/n) y

Disable SSH warning? (y/n) y

Add vmotion interfaces? (y/n) y
Netmask for vMotion: 255.255.255.0
VLAN ID: (0-4096) 3018
vMotion MTU is set to use jumbo frames (9000 bytes). Do you want to change to 1500 bytes? (y/n) n
vMotion IP for 10.1.167.111: 172.0.167.111
Adding vmotion-3018 to 10.1.167.111
Adding vmkernel to 10.1.167.111
vMotion IP for 10.1.167.112: 172.0.167.112
Adding vmotion-3018 to 10.1.167.112
Adding vmkernel to 10.1.167.112
vMotion IP for 10.1.167.113: 172.0.167.113
Adding vmotion-3018 to 10.1.167.113
Adding vmkernel to 10.1.167.113
vMotion IP for 10.1.167.114: 172.0.167.114
Adding vmotion-3018 to 10.1.167.114
Adding vmkernel to 10.1.167.114
vMotion IP for 10.1.167.115: 172.0.167.115
Adding vmotion-3018 to 10.1.167.115
Adding vmkernel to 10.1.167.115
vMotion IP for 10.1.167.116: 172.0.167.116
Adding vmotion-3018 to 10.1.167.116
Adding vmkernel to 10.1.167.116
vMotion IP for 10.1.167.117: 172.0.167.117
Adding vmotion-3018 to 10.1.167.117
Adding vmkernel to 10.1.167.117
vMotion IP for 10.1.167.118: 172.0.167.118
Adding vmotion-3018 to 10.1.167.118
Adding vmkernel to 10.1.167.118

Add VM network VLANs? (y/n) y
Attempting to find UCSM IP
Site A - UCSM IP: 192.168.167.204
Site A - UCSM Username: admin
Site A - UCSM Password:
Site A - HX UCS Sub Organization: HXV-Org1
Site B - UCSM IP: 192.168.167.207
Site B - UCSM Username: admin
Site B - UCSM Password:
Site B - HX UCS Sub Organization: HXV-Org1
Port Group Name to add (VLAN ID will be appended to the name): hxv-vm-network
VLAN ID: (0-4096) 2218
Adding VLAN 2218 to FI
Adding VLAN 2218 to vm-network-a VNIC template
Adding VLAN 2218 to FI
Adding VLAN 2218 to vm-network-a VNIC template
Adding hxv-vm-network-2218 to 10.1.167.111
Adding hxv-vm-network-2218 to 10.1.167.112
Adding hxv-vm-network-2218 to 10.1.167.113
Adding hxv-vm-network-2218 to 10.1.167.114
Adding hxv-vm-network-2218 to 10.1.167.115
Adding hxv-vm-network-2218 to 10.1.167.116
Adding hxv-vm-network-2218 to 10.1.167.117
Adding hxv-vm-network-2218 to 10.1.167.118
Add additional VM network VLANs? (y/n) n

Run health check? (y/n) y

Validating cluster health and configuration...

Cluster Summary:
Version - 3.5.1a-31118
Model - HX220C-M5SX
Health - HEALTHY
ASUP enabled - False
root@HyperFlex-Installer:~#

```

Enable Smart Licensing for Stretch HyperFlex Cluster

To enable licensing for the newly deployed HyperFlex stretched cluster, follow the procedures outlined in the [Install HyperFlex Management Cluster](#).

Enable Syslog for Stretch HyperFlex Cluster

To prevent the loss of diagnostic information when a host fails, ESXi logs should be sent to a central location. Logs can be sent to the VMware vCenter server or to a separate syslog server.

Use a multi-exec tool (for example, **MobaXterm**) to simultaneously execute the same command on all servers in the cluster as shown below.

To configure syslog on ESXi hosts, follow these steps:

1. Log into the ESXi host through SSH as the **root** user.
2. Enter the following commands, replacing the IP address in the first command with the IP address of the vCenter or the syslog server that will receive the syslog logs.

```

[root@hvx-cl1-esxi-1:~] esxcli system syslog config set --loghost='udp://10.10.167.240'
[root@hvx-cl1-esxi-1:~] esxcli system syslog reload
[root@hvx-cl1-esxi-1:~] esxcli network firewall ruleset set -r syslog -e true
[root@hvx-cl1-esxi-1:~] esxcli network firewall refresh
[root@hvx-cl1-esxi-1:~]

[root@hvx-cl1-esxi-2:~] esxcli system syslog config set --loghost='udp://10.10.167.240'
[root@hvx-cl1-esxi-2:~] esxcli system syslog reload
[root@hvx-cl1-esxi-2:~] esxcli network firewall ruleset set -r syslog -e true
[root@hvx-cl1-esxi-2:~] esxcli network firewall refresh
[root@hvx-cl1-esxi-2:~]

[root@hvx-cl1-esxi-3:~] esxcli system syslog config set --loghost='udp://10.10.167.240'
[root@hvx-cl1-esxi-3:~] esxcli system syslog reload
[root@hvx-cl1-esxi-3:~] esxcli network firewall ruleset set -r syslog -e true
[root@hvx-cl1-esxi-3:~] esxcli network firewall refresh
[root@hvx-cl1-esxi-3:~]

[root@hvx-cl1-esxi-4:~] esxcli system syslog config set --loghost='udp://10.10.167.240'
[root@hvx-cl1-esxi-4:~] esxcli system syslog reload
[root@hvx-cl1-esxi-4:~] esxcli network firewall ruleset set -r syslog -e true
[root@hvx-cl1-esxi-4:~] esxcli network firewall refresh
[root@hvx-cl1-esxi-4:~]

[root@hvx-cl1-esxi-5:~] esxcli system syslog config set --loghost='udp://10.10.167.240'
[root@hvx-cl1-esxi-5:~] esxcli system syslog reload
[root@hvx-cl1-esxi-5:~] esxcli network firewall ruleset set -r syslog -e true
[root@hvx-cl1-esxi-5:~] esxcli network firewall refresh
[root@hvx-cl1-esxi-5:~]

[root@hvx-cl1-esxi-6:~] esxcli system syslog config set --loghost='udp://10.10.167.240'
[root@hvx-cl1-esxi-6:~] esxcli system syslog reload
[root@hvx-cl1-esxi-6:~] esxcli network firewall ruleset set -r syslog -e true
[root@hvx-cl1-esxi-6:~] esxcli network firewall refresh
[root@hvx-cl1-esxi-6:~]

[root@hvx-cl1-esxi-7:~] esxcli system syslog config set --loghost='udp://10.10.167.240'
[root@hvx-cl1-esxi-7:~] esxcli system syslog reload
[root@hvx-cl1-esxi-7:~] esxcli network firewall ruleset set -r syslog -e true
[root@hvx-cl1-esxi-7:~] esxcli network firewall refresh
[root@hvx-cl1-esxi-7:~]

[root@hvx-cl1-esxi-8:~] esxcli system syslog config set --loghost='udp://10.10.167.240'
[root@hvx-cl1-esxi-8:~] esxcli system syslog reload
[root@hvx-cl1-esxi-8:~] esxcli network firewall ruleset set -r syslog -e true
[root@hvx-cl1-esxi-8:~] esxcli network firewall refresh
[root@hvx-cl1-esxi-8:~]

```

Manage Cluster using Cisco Intersight

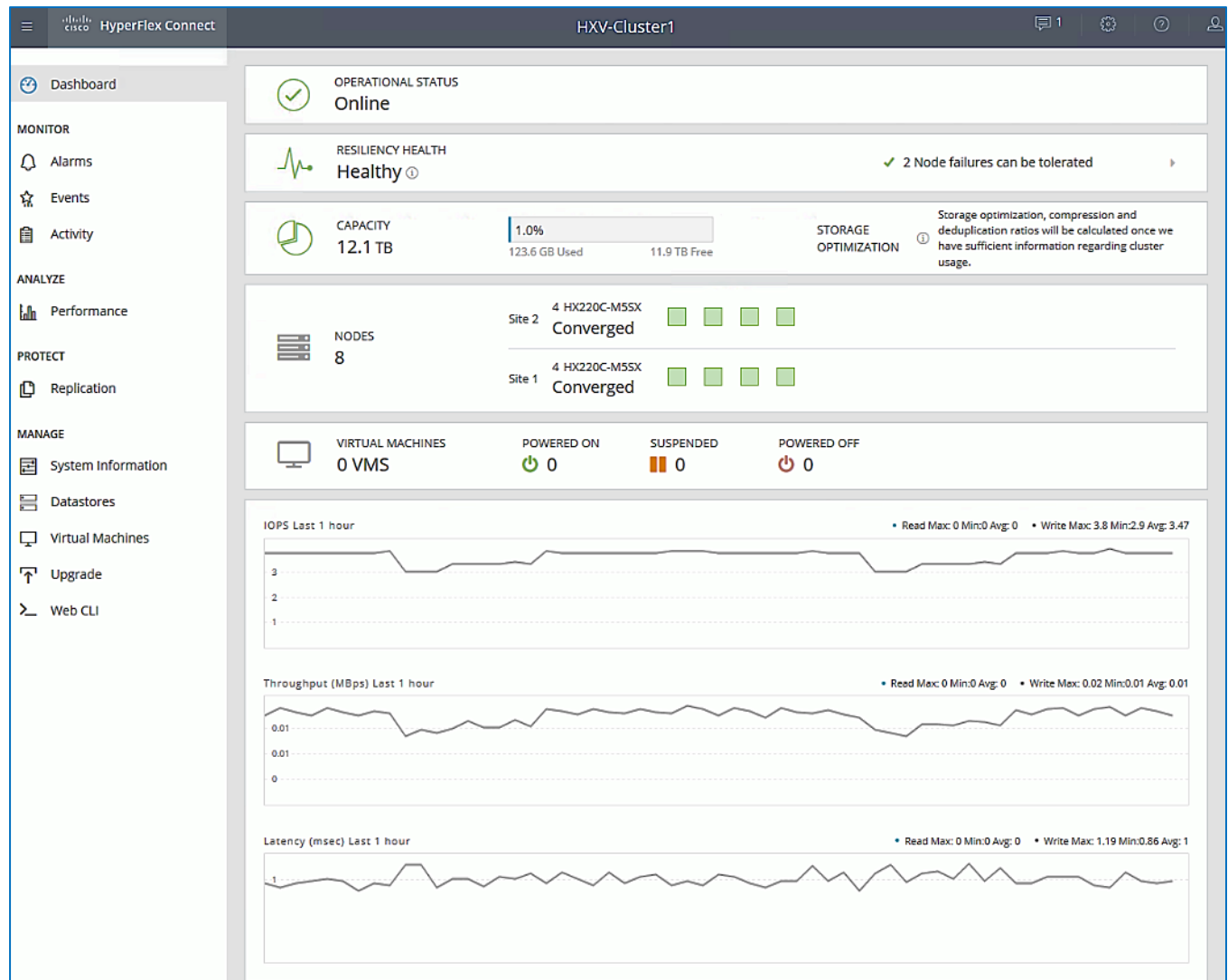
Cisco Intersight provides a centralized dashboard with a single view of all Cisco UCS Domains, HyperFlex clusters and servers regardless of their location. New features and capabilities are continually being added over time. Please see the [Cisco Intersight](#) website for the latest information.

To manage the HyperFlex stretched cluster from Cisco Intersight, follow the procedures outlined in the [Enable Cisco Intersight Cloud-Based Management](#) section.

Manage Cluster using HyperFlex Connect

To manage the HyperFlex stretched cluster using HyperFlex Connect, follow these steps:

1. Open a web browser and navigate to the Management IP address of the HX cluster (for example, <https://10.1.167.110>). Log in using the **admin** account. Password should be same as the one specified for the Storage Controller virtual machine during the installation process.



2. The **Dashboard** provides general information about the cluster's operational status, health, Node failure tolerance, Storage Performance and Capacity Details and Cluster Size and individual Node health.

(Optional) Manage Cluster using VMware vCenter (through Plugin)

The Cisco HyperFlex vCenter Web Client Plugin can be deployed as a secondary tool to monitor and configure the HyperFlex cluster.



This plugin is not supported in the HTML5 based VMware vSphere Client for vCenter.

To manage the HyperFlex cluster using the vCenter Web Client Plugin for vCenter 6.5, follow the procedures outlined in the [Install HyperFlex Management Cluster](#) section of this document.

Enable/Disable Auto-Support and Notifications

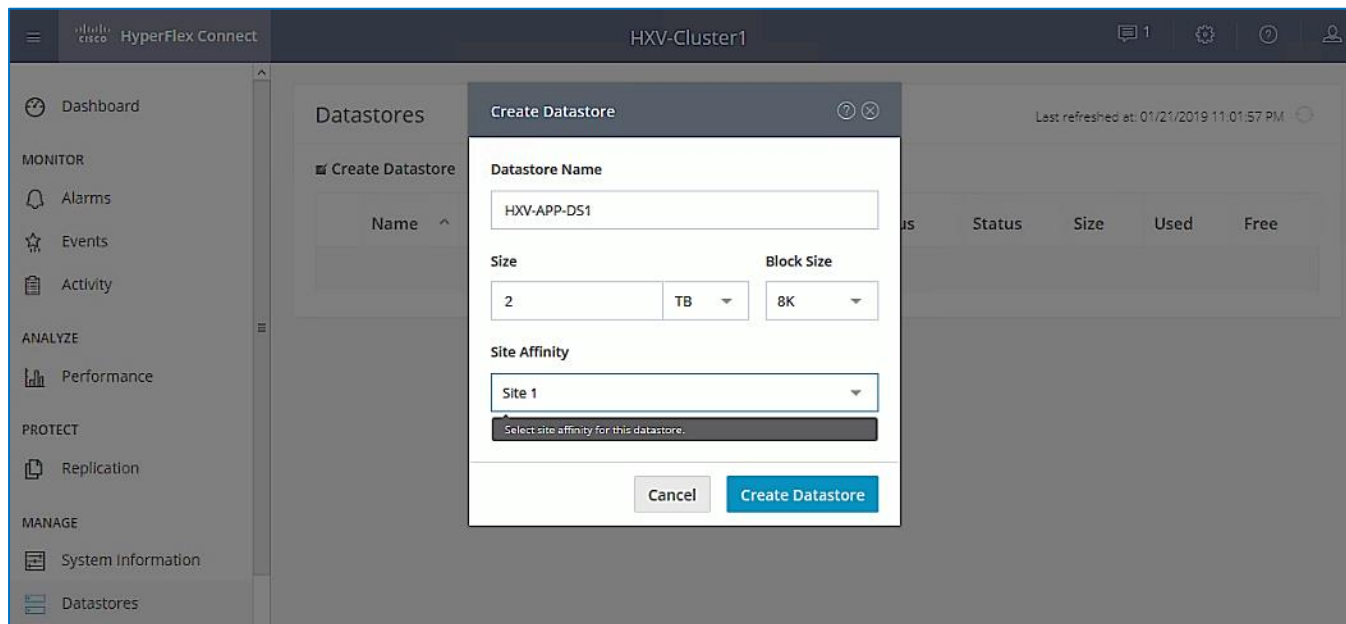
Auto-Support is enabled if specified during the HyperFlex installation. Auto-Support enables Call Home to automatically send support information to Cisco TAC, and notifications of tickets to the email address specified. If the settings need to be modified, they can be changed in the HyperFlex Connect HTML management webpage.

To change Auto-Support settings, follow the procedures outlined in the [Install HyperFlex Management Cluster](#) section of this document.

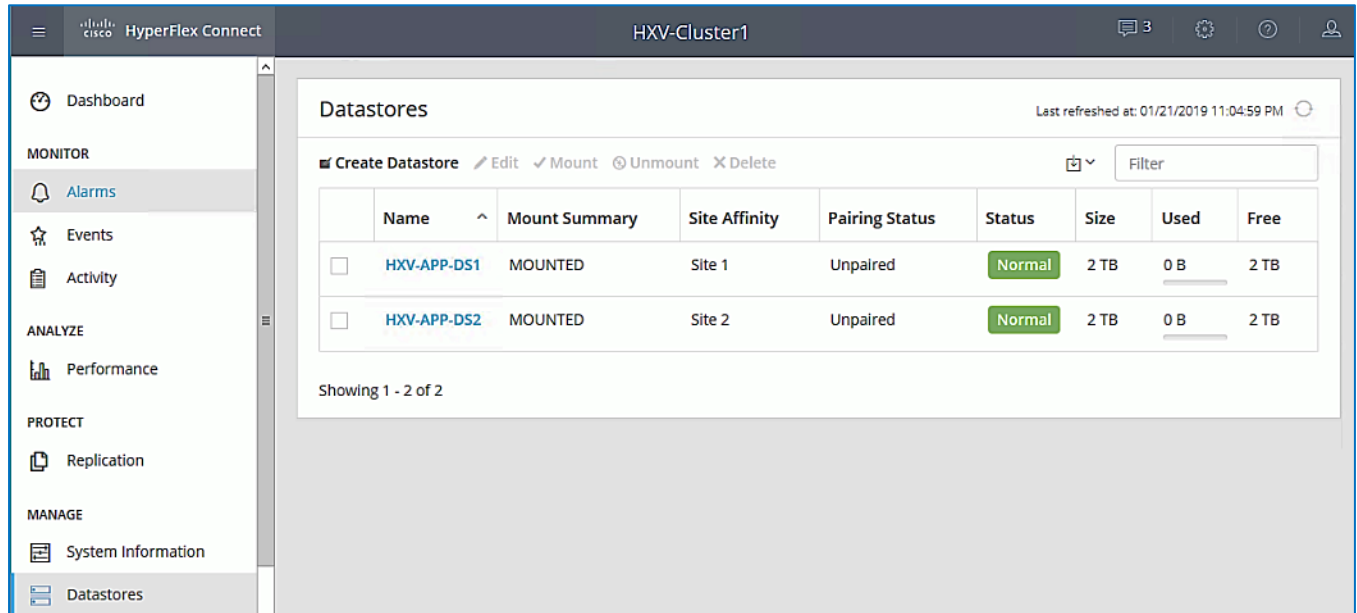
Create Datastores for Virtual Machines with Site Affinity

Datastores created in stretched clusters require a **Site Affinity** setting compared to datastores in standard clusters. Specifying a site association for the datastores ensures that all requests to read data from that datastore will be serviced by the nodes in that specific site, rather than by nodes in the remote site. When deploying Virtual Machines, the virtual machines should be configured to store their virtual disk files in a datastore at the same site as the virtual machine. The placement of the virtual machines using vSphere Dynamic Resource Scheduler (DRS) site affinity rules optimizes the performance in a stretched cluster, by ensuring proximity to the users that consume the services provided by the virtual machine.

To deploy a new datastore from HyperFlex Connect, follow the procedures outlined in the [Install HyperFlex Management Cluster](#) section of this document, however for stretched clusters, the **Site Affinity** needs to be specified as shown below.



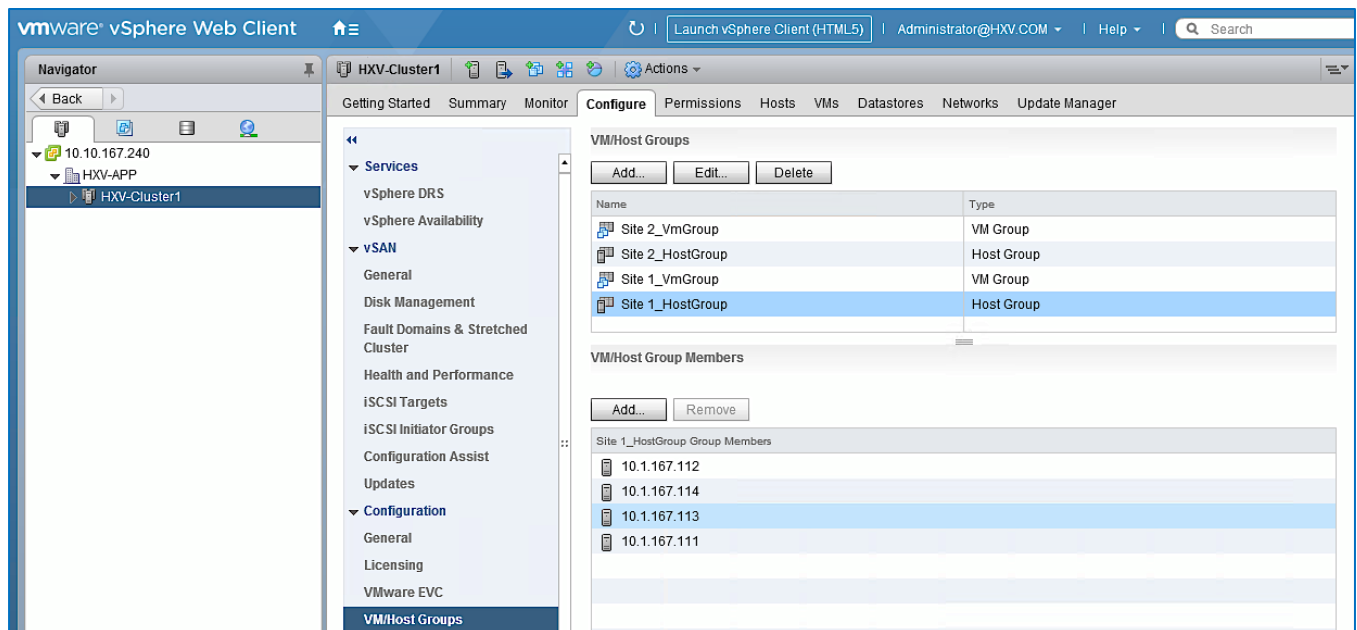
To validate the design, two datastores are created on the stretch cluster with **Site Affinity** to Site 1 (Pod-1) and Site 2 (Pod-2) as shown below.



Configure vSphere DRS with Site Affinity

VMware vSphere Dynamic Resource Scheduler (DRS) must be configured with site affinity rules in order for the stretched cluster to operate in an optimal manner. Virtual machine placement across a stretched cluster uses these site affinity rules, in order to constrain virtual machines to only run on the nodes in their primary site during normal operation. The datastore that stores the virtual machine's virtual disk files will also be associated with the same site. Site affinity rules and groups are automatically created during the installation, and the rules are created in such a manner that the virtual machines are allowed to restart and run in the other site in case of a site failure. When virtual machines are created, they are automatically placed into the virtual machine group associated with the site where they are running. This method helps to balance workloads across all of the nodes in both sites, while retaining the enhanced failover capability of a stretched cluster, in case an entire site was to go offline or otherwise fail.

The automatically created Host Groups and Virtual Machine Groups for each site are shown below.



vSphere High Availability Recommendations

The VMware setup is critical for the operation of a HyperFlex stretched cluster. HyperFlex installation configures many VMware features that a stretched cluster requires such as vSphere HA, DRS, virtual machine and datastore host-groups, site-affinity, etc. In addition, customers should also enable the following vSphere HA settings in VMware vCenter:

- vSphere Availability: vSphere HA should be enabled but keep Proactive HA disabled
- Failure Conditions and responses:
 - Enable Host Monitoring
 - For Host Failure Response, select Restart VMs
 - For Response for Host Isolation, select Power off and restart VMs
 - For Datastore with PDL, select Power off and restart VMs
 - For Datastore with PDL, select Power off and restart VMs (conservative)
 - For VM Monitoring: Customer can enable this if they prefer. It is typically disabled.
- Admission Control: select Cluster resource percentage for Define host failover capacity by
- Datastore Heartbeats: Select Use datastores only from the specified list and select HyperFlex datastores in each site
- Advanced Settings:
 - select False for `das.usedefaultisolationaddress`
 - select an IP address in Site A for `das.isolationaddress0`
 - select an IP address in Site B for `das.isolationaddress1`
- For additional recommendations, see Operating Cisco HyperFlex Data Platform Stretched Clusters white paper in the [References](#) section of this document.

Migrate Virtual Networking to Cisco AVE on HyperFlex Application Cluster

This section configures the virtual networking for the virtual machine networks in the Application cluster. APIC manages the virtual networking through the VMM integration with VMware vCenter that manages the Application HyperFlex cluster. The other networks (Inband Management, Storage Data and vMotion networks) for the Application HyperFlex cluster will remain on the VMware vSwitch as deployed by the Installer. The virtual networking uses Cisco ACI Virtual Edge (AVE) as the virtual switch for the VM networks hosted on the Application cluster. The vCenter that manages Application HyperFlex cluster is hosted on the Management cluster.



VMM integration with Cisco AVE requires VMware Enterprise Plus license, the same as VMware vDS. However, Cisco AVE is an ACI virtual Leaf (vLeaf) that brings advanced ACI capabilities such as micro-segmentation, policies and visibility, VxLAN, distributed firewall and so on, to the virtualization domain.

Setup Information

The setup information for migrating the default virtual networking from VMware vSwitch to Cisco is provided below:

- Associated Attachable Entity Profile: `HXV-UCS_AAEP`

- Infrastructure VLAN for VXLAN: 4093
- VLAN Name in Cisco UCS: Infra-VLAN
- Cisco UCS vNIC Templates for VM Networks: vm-network-a, vm-network-b
- vNIC Template QoS Policy: Gold
- VMware vCenter Managing the VMM Domain: hxv0-vcsa.hxv.com (10.10.167.240)
- Virtual Switch Name: HXV1-AVE
- VLAN Name: HXV1-VMM_VLANS
- VLAN Pool: 1118-1128
- AVE Fabric-Wide Multicast Address: 239.167.10.240
- Pool of Multicast Addresses (one per EPG): HXV1-AVE-MCAST_POOL
- Create Multicast Address Range: 239.167.10.18-.28
- Associated Attachable Entity Profile: HXV-UCS_AAEP
- VMware vCenter Credentials: Username/Password for the vCenter managing the VMM domain
- VMware vCenter Credentials – Profile Name: Administrator
- VMware vCenter Managing the VMM Domain: hxv0-vcsa.hxv.com (10.10.167.240)
- DVS Version: vCenter Default
- VMware vCenter Datacenter: HXV-APP
- Default vSwitch for VM networks: vswitch-hxv-vm-network
- Uplinks on Default vSwitch for VM Networks: vmnic2, vmnic6

Deployment Overview

The high-level steps for deploying Cisco AVE in a VMware vSphere environment with HyperFlex are:

- Verify IP connectivity between Cisco APIC and VMware vCenter.
- Enable Infrastructure VLAN (4093) for VXLAN tunneling between ACI Leaf switch and Cisco AVE. This requires the VLAN to be configured on Cisco UCS Manager and on links (vNICs) to HyperFlex servers. The MTU should be 1600 or higher for VxLAN. The QoS system class should be changed to reflect the MTU change.
- Setup a new VMM domain in ACI for Cisco AVE. Allocate a VLAN pool and Multicast Address Pool for Application EPGs and port-groups to use. The pool should accommodate the number of EPGs published to the VMware vCenter domain. Apply pre-configured policies to the virtual switch in the new VMM domain. Enable statistics collection for the new VMM domain.
- Add HyperFlex ESXi hosts to Cisco AVE.
- Download Cisco AVE OVF file to VMware vCenter.

- Setup networking for deploying Cisco AVE virtual machines to Management network (port-group). Allocate a Management IP Address for each Cisco AVE Virtual Machine – one virtual machine per host. Setup DHCP to allocate IP address to Cisco AVE virtual machine – either using an existing DHCP server or VMware vCenter. VMware vCenter is used in this setup. The addresses should be in a contiguous block when VMware vCenter is the DHCP server.
- Deploy Cisco ACI vSphere plug-in - VMware vCenter 6.0U3 or higher is recommended.
- Deploy Cisco AVE virtual machine on the ESXi Hosts Using the Cisco ACI Plug-In.
- The Cisco AVE environment is now ready for deploying Application EPGs and the corresponding virtual-networking and port-groups will be dynamically deployed by Cisco APIC for Application virtual machines to connect to.

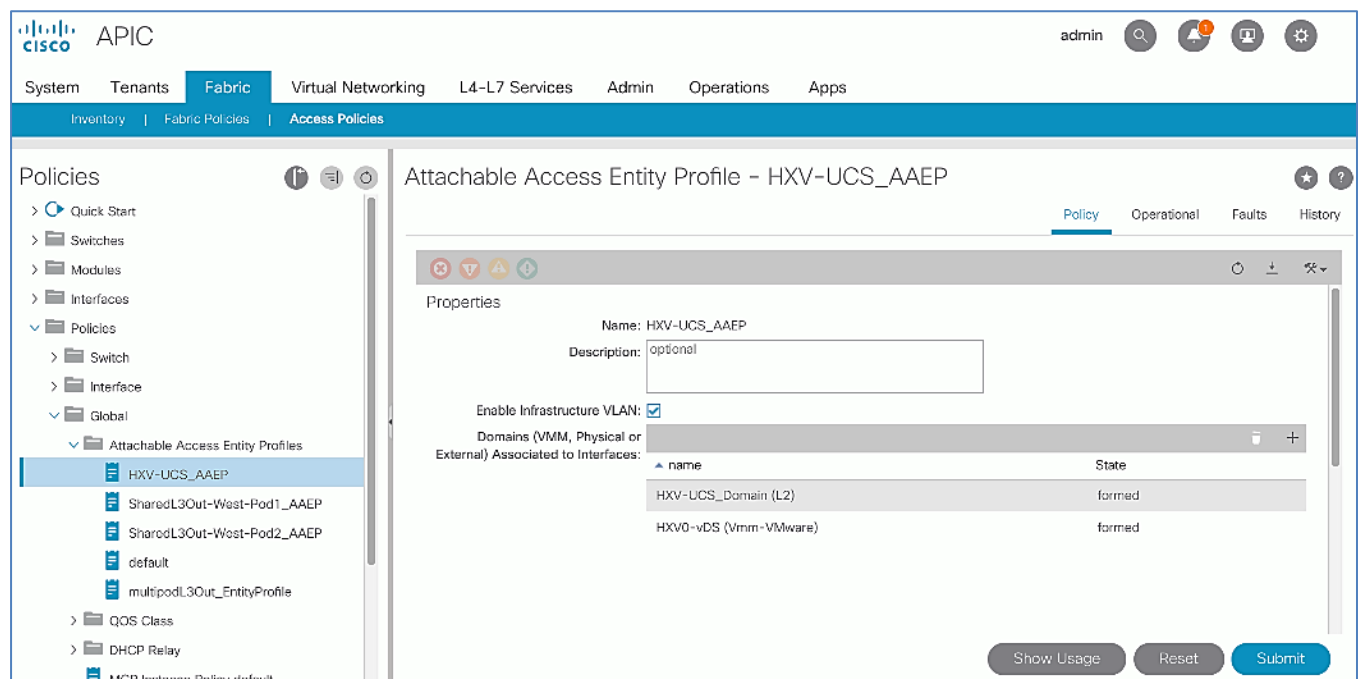
Deployment Steps

To enable APIC-controlled virtual networking for the Management cluster, complete the steps outlined in this section.

Enable Infrastructure VLAN for Cisco AVE in the ACI Fabric

To enable the infrastructure VLAN (4093) in the ACI fabric for VXLAN tunneling between ACI Leaf switch and Cisco AVE, follow these steps:

1. Use a browser to navigate to APIC's Web GUI. Log in with the **admin** account.
2. From the top menu, select **Fabric > Access Policies**.
3. From the left navigation pane, select and expand **Policies > Global > Attachable Access Entity Profiles > HXV-UCS_AAEP**.
4. In the right window pane, select the checkbox for **Enable Infrastructure VLAN**.



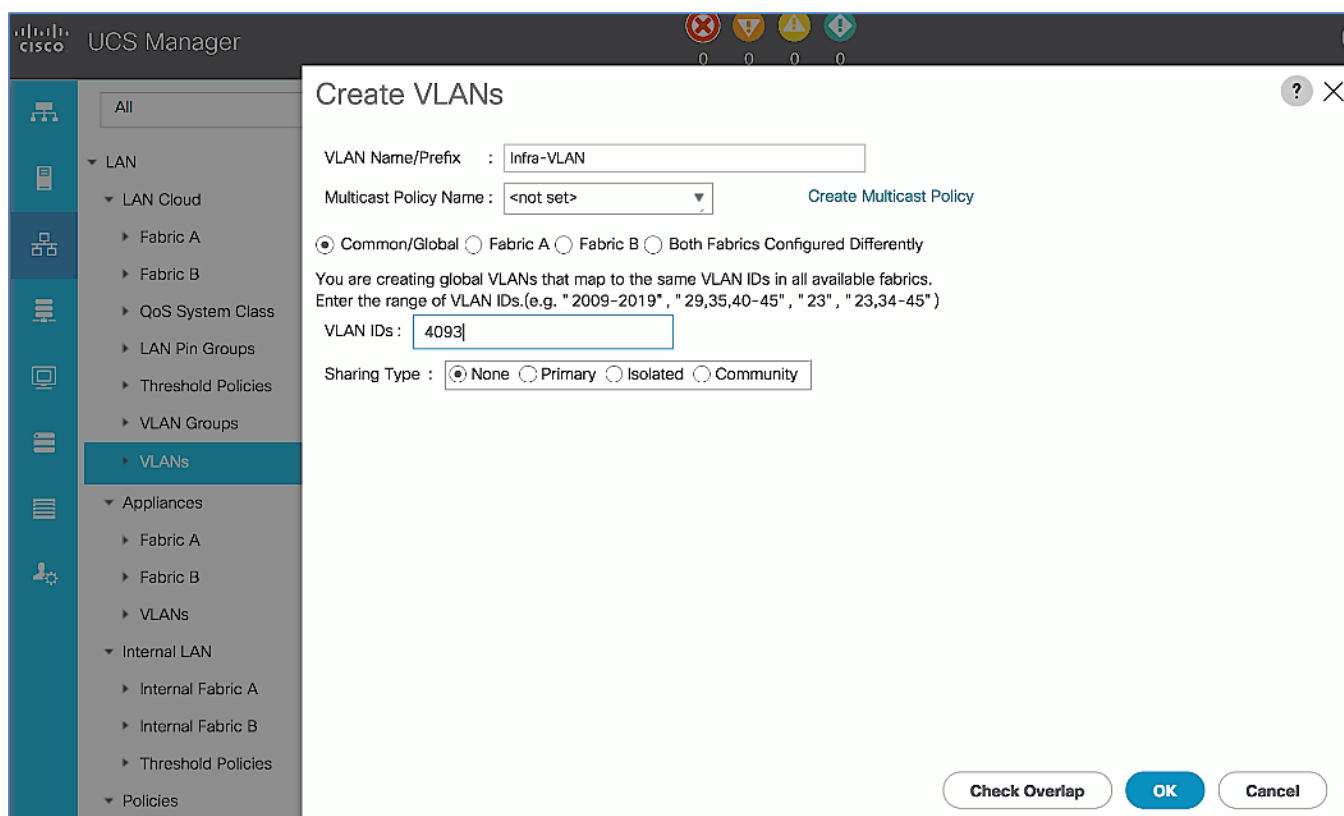
5. Click Submit.

Enable Infrastructure VLAN for Cisco AVE in the Cisco UCS Domain

The infrastructure VLAN is enabled on the two Cisco UCS domains in the HyperFlex stretched cluster in order to deploy Cisco AVE on HyperFlex ESXi hosts in both domains. The VLAN should be enabled on the uplinks to the ACI fabric and on the vNICs to the HyperFlex servers.

To enable the infrastructure VLAN (4093) on the Cisco UCS Domains for VXLAN tunneling between ACI Leaf switch and Cisco AVE, follow these steps:

1. Use a browser to navigate to Cisco UCS Manager Web GUI. Log in using **admin** account.
2. From the left navigation pane, select the **LAN** icon. Select and expand **LAN > LAN Cloud > VLANs**.
3. Right-click **VLANs** and select **Create VLANs**.
4. In the **Create VLANs** pop-up window, specify a **VLAN Name** (for example, *Infra-VLAN*). For **VLAN IDs**, specify 4093 for Infrastructure VLAN for VXLAN. Keep everything else as-is.



5. Click **OK** twice.
6. Repeat steps 1-5 to enable the infrastructure VLAN on the second UCS domain in the HyperFlex stretched cluster.

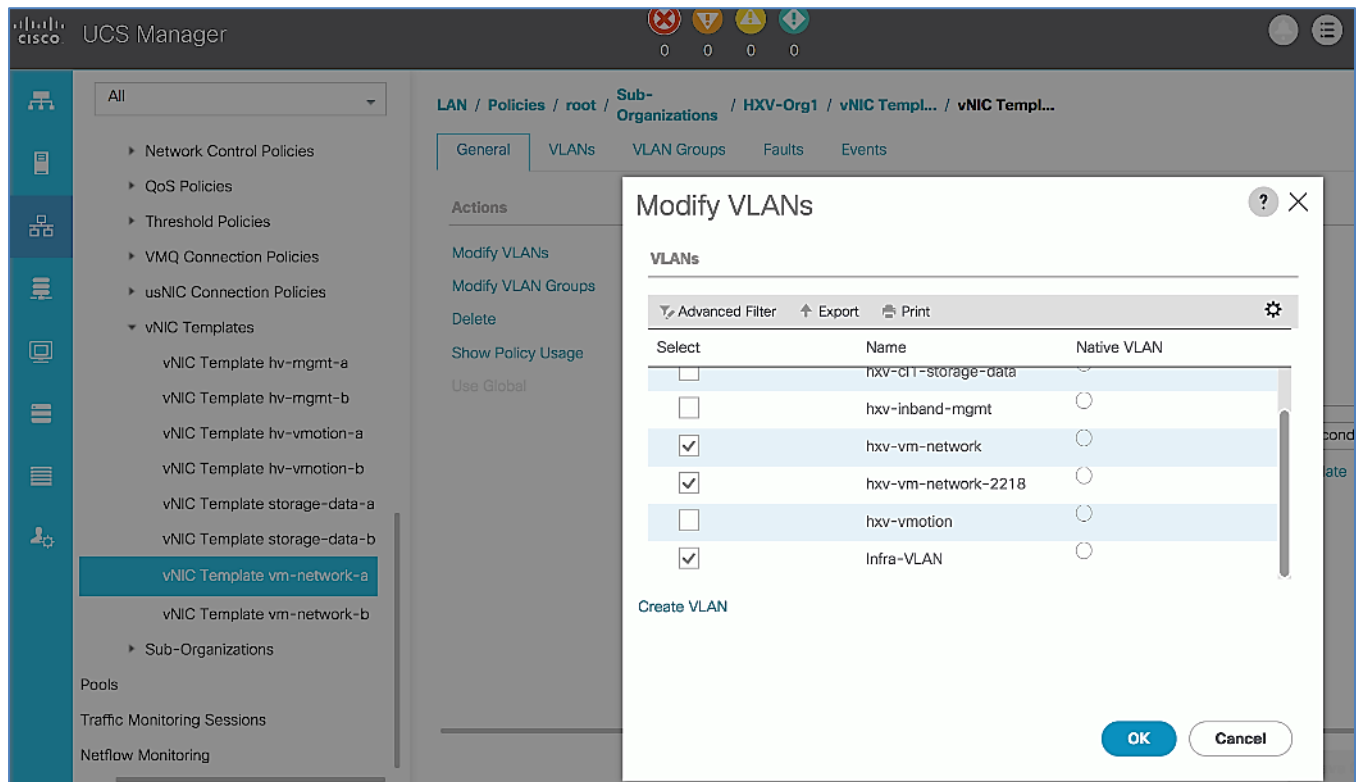
Enable Infrastructure VLAN for Cisco AVE on HyperFlex Server Uplinks

The virtual machine networks are trunked within a VXLAN tunnel to and from the ACI fabric. Dedicated vNICs (*vm-network-a*, *vm-network-b*) are assigned for virtual machine networks through UCS Fabric A (FI-A) and Fabric B (FI-B). The VXLAN VLAN is enabled on the same vNICs.

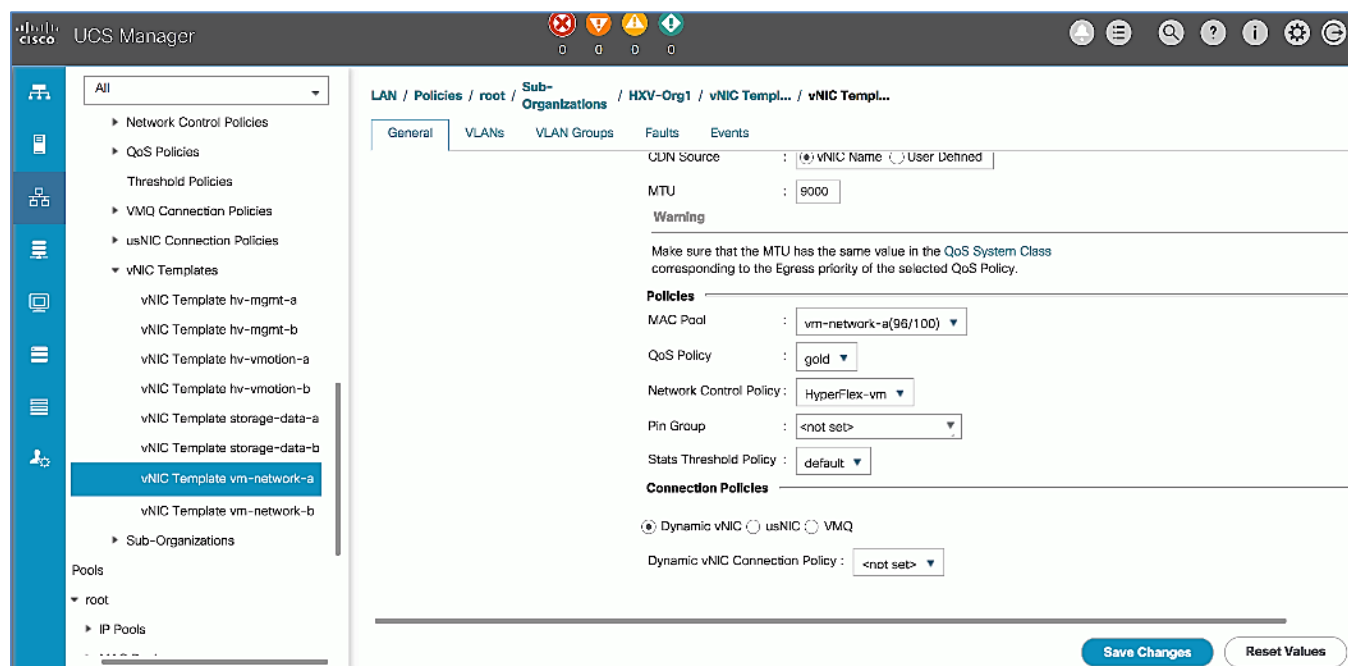
To enable the infrastructure VLAN (4093) on HyperFlex server uplinks for VXLAN tunneling between ACI Leaf switch and Cisco AVE, follow these steps:

1. Use a browser to navigate to Cisco UCS Manager Web GUI. Login using **admin** account.

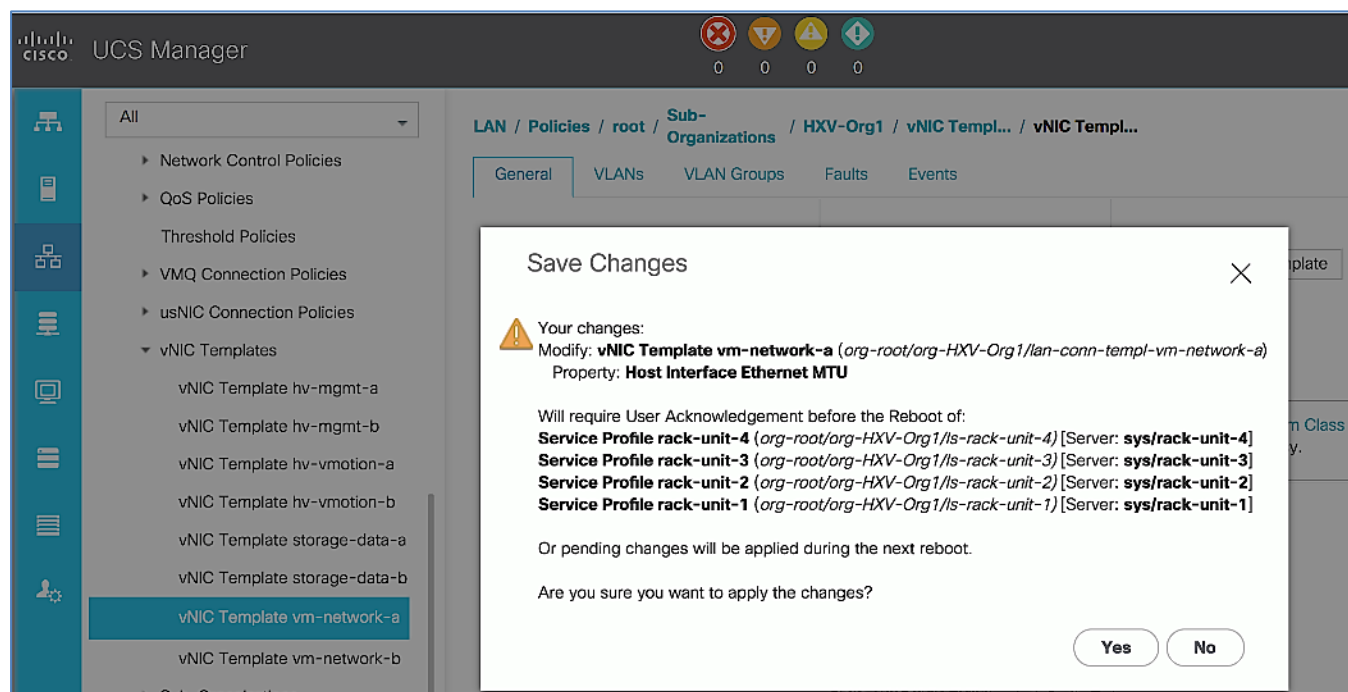
- From the left navigation pane, select the **LAN** icon. Select and expand **LAN > Policies > root > Sub-Organizations > HXV1-Orig1 > vNIC Templates > vm-network-a**.
- In the right window pane, navigate to **General**. Click **Modify VLANs**.
- In the **Modify VLANs** pop-up window, select the check-box for the infrastructure VLAN (Infra-VLAN).



- Click **OK** twice.
- In the **General** tab, scroll down to the **MTU** field and change the MTU. Increase the MTU to something higher than 1600B for VXLAN traffic. In this design, MTU is increased to 9000B to stay consistent with MTU on other vNICs. See the **Warning** below the MTU field; the QoS System Class for the QoS policy (`gold`) used by this vNIC template should be changed to reflect the above MTU setting. For more information, see the following subsection.



7. Click **Save Changes**.
8. In the **Save Changes** pop-up window, select **Yes** to apply changes.



9. In the **Pending Activities** pop-up, click **X** to cancel the pop-up or if window does open, click **Cancel** to exit without acknowledging. **Reboot will be done at a later step**.

These changes will be automatically applied to the second vNIC template (vm-network-b) for virtual machine networks.



A reboot will be done from HyperFlex Connect or VMware vCenter and not from Cisco UCS. The HyperFlex Data Platform plug-in will be used to reboot one host at a time so as to ensure the cluster is healthy before proceeding to the next host.

- Repeat steps 1-9 to enable the infrastructure VLAN and vNIC templates changes on the second UCS domain in the HyperFlex stretched cluster.

Change QoS System Class Policy for the QoS Policy Used by the Virtual Machine vNIC Templates

To change the QoS System Class Policy for the QoS policy used by the VM Network vNIC templates such that the MTUs match at the system level and vNIC level, follow these steps:

- Use a browser to navigate to Cisco UCS Manager Web GUI. Log in using **admin** account.
- From the left navigation pane, select the **LAN** icon. Select and expand **LAN > LAN Cloud > QoS System Class**.
- In the right window pane, find the QoS policy used by the vNIC template above (where the MTU changes were done). Change the MTU to **9216** for the **Gold** policy used by the above vNIC templates.
- Click **Save Changes**.
- In the **Save Changes** pop-up window, select **Yes** to apply changes if the warning is acceptable.
- Click **OK**.
- Repeat steps 1-6 to modify the **QoS System Class** on the second UCS domain in the HyperFlex stretched cluster.

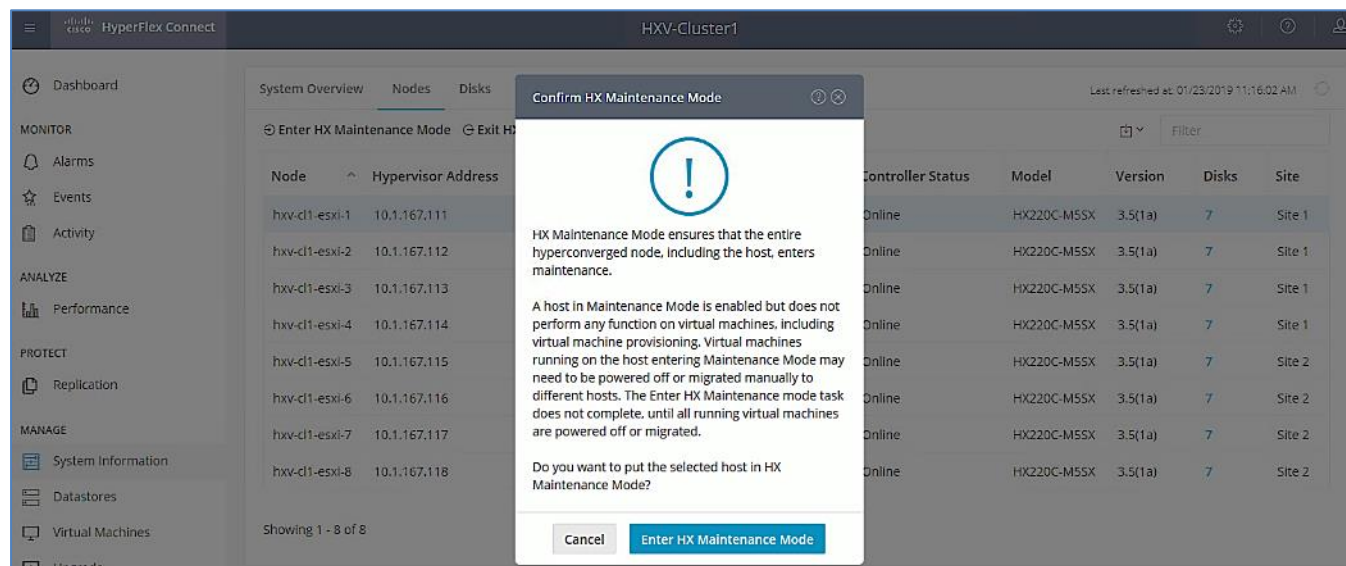
Reboot HyperFlex Hosts from vCenter to Apply Changes – Option 1

To apply the changes made from Cisco UCS Manager to the vNIC template, follow these steps:

- Use a web browser to navigate to HyperFlex Connect. Log in using the **admin** account.
- From the left navigation pane, select **System Information**.
- In the right window pane, select the **Nodes** tab.
- Right-click the first host and from the top menu, click **Enter HX Maintenance Mode**.

Node	Hypervisor Address	Hypervisor Status	Controller Address	Controller Status	Model	Version	Disks	Site
hvx-cl1-esxi-1	10.1.167.111	Online	10.1.167.161	Online	HX220C-MSSX	3.5(1a)	7	Site 1
hvx-cl1-esxi-2	10.1.167.112	Online	10.1.167.162	Online	HX220C-MSSX	3.5(1a)	7	Site 1
hvx-cl1-esxi-3	10.1.167.113	Online	10.1.167.163	Online	HX220C-MSSX	3.5(1a)	7	Site 1
hvx-cl1-esxi-4	10.1.167.114	Online	10.1.167.164	Online	HX220C-MSSX	3.5(1a)	7	Site 1
hvx-cl1-esxi-5	10.1.167.115	Online	10.1.167.165	Online	HX220C-MSSX	3.5(1a)	7	Site 2
hvx-cl1-esxi-6	10.1.167.116	Online	10.1.167.166	Online	HX220C-MSSX	3.5(1a)	7	Site 2
hvx-cl1-esxi-7	10.1.167.117	Online	10.1.167.167	Online	HX220C-MSSX	3.5(1a)	7	Site 2
hvx-cl1-esxi-8	10.1.167.118	Online	10.1.167.168	Online	HX220C-MSSX	3.5(1a)	7	Site 2

- Review the warning in the pop-up window and click **Enter HX Maintenance Mode**.



6. Monitor the Activity Page for the status and once the node is in maintenance mode and the controller virtual machine is powered off, navigate to VMware vCenter to reboot the host.



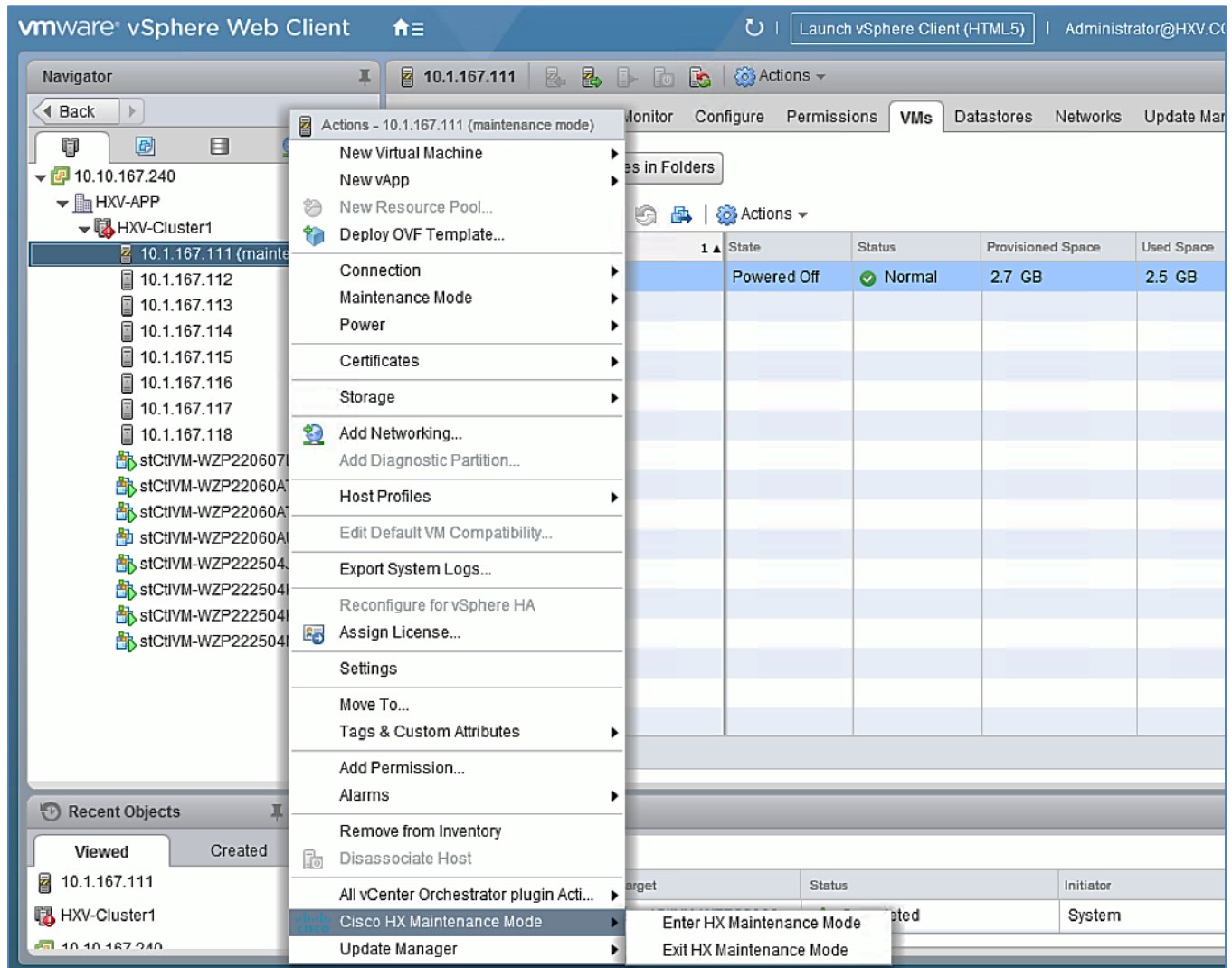
The reboot takes a few minutes and if you can ping the server but it doesn't show up in vCenter, you may need to reconnect by selecting **Connect > Reconnect**.

7. When the host reboots and comes back up, go to HyperFlex Connect and select **System Information > Nodes** and select the Node from the list and click **Exit HX Maintenance Mode** from the top.
8. Monitor the status of the Controller virtual machine on the host from the **Activity** section in the left-navigation pane.
9. When the cluster comes up and is seen as healthy from HyperFlex Connect, repeat above steps for each host in the HX cluster.

Reboot HyperFlex Hosts from VMware vCenter to Apply Changes – Option 2

To apply the changes made to the vNIC template from Cisco UCS Manager, follow these steps:

1. Use a browser to navigate to the VMware vCenter server that will be used to deploy the Witness virtual machine. Click the vSphere Web Client of your choice. Login using an **Administrator** account.
2. Navigate to **Home > Hosts and Clusters** and select the first host in the HX cluster to put in **HX maintenance mode** and reboot to apply the service profile changes done in UCS.
3. Right-click the first host and select **Cisco HX Maintenance Mode > Enter HX Maintenance Mode** from the bottom of the list.



4. Click **OK** to accept changes.
5. When the host is in maintenance mode, right-click the host again and select **Power > Reboot** option. Enter a reason in the pop-up window or click **OK**.
6. When the host reboots and comes back up, right-click the host and select **Cisco HX Maintenance Mode > Exit HX Maintenance Mode**.



The reboot takes a few minutes and if you can ping the server but it doesn't show up in vCenter, you may need to reconnect by selecting **Connect > Reconnect**.

7. When the cluster comes up and is seen as healthy (ideally from HyperFlex Connect), repeat above steps for each host in the HX cluster.

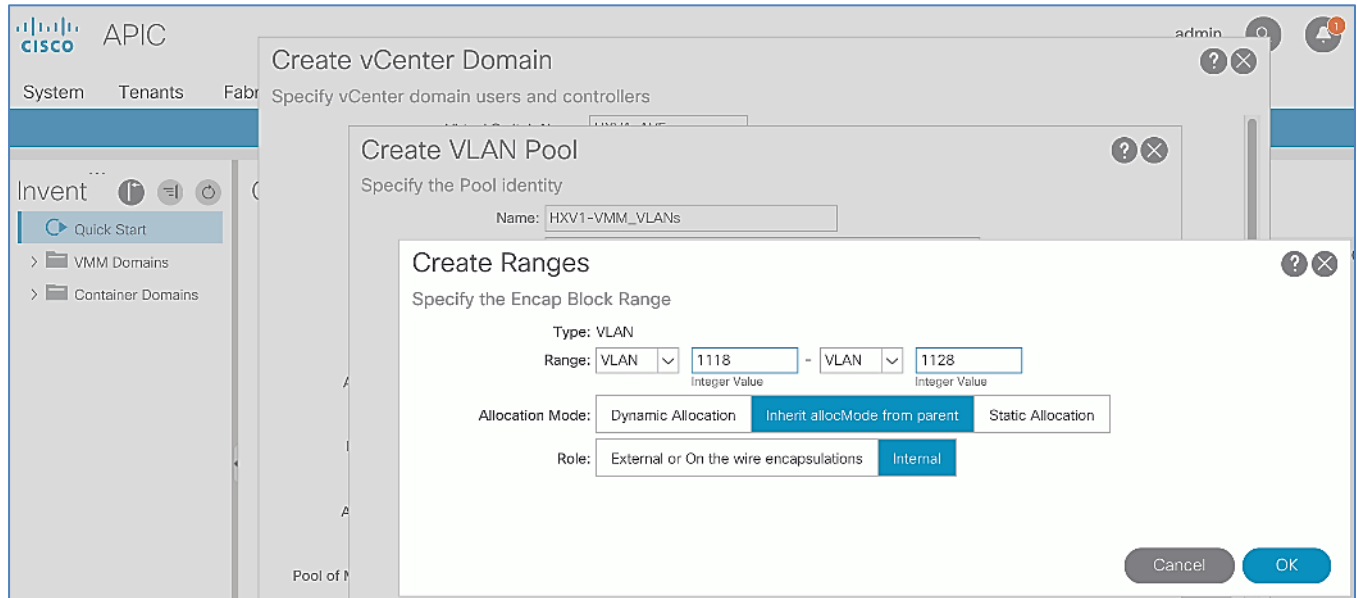
Setup a New VMM Domain in ACI for Cisco AVE

A new VMM domain must be configured Cisco ACI in order to deploy an APIC-controlled Cisco AVE in that domain. The VMM domain will require a VLAN pool and Multicast Address Pool to be allocated for Application EPGs and port-groups. The pool should accommodate the number of EPGs published to the VMware vCenter domain in the form of port-groups. Pre-configured policies and statistics collection are also enabled for the new VMM domain.

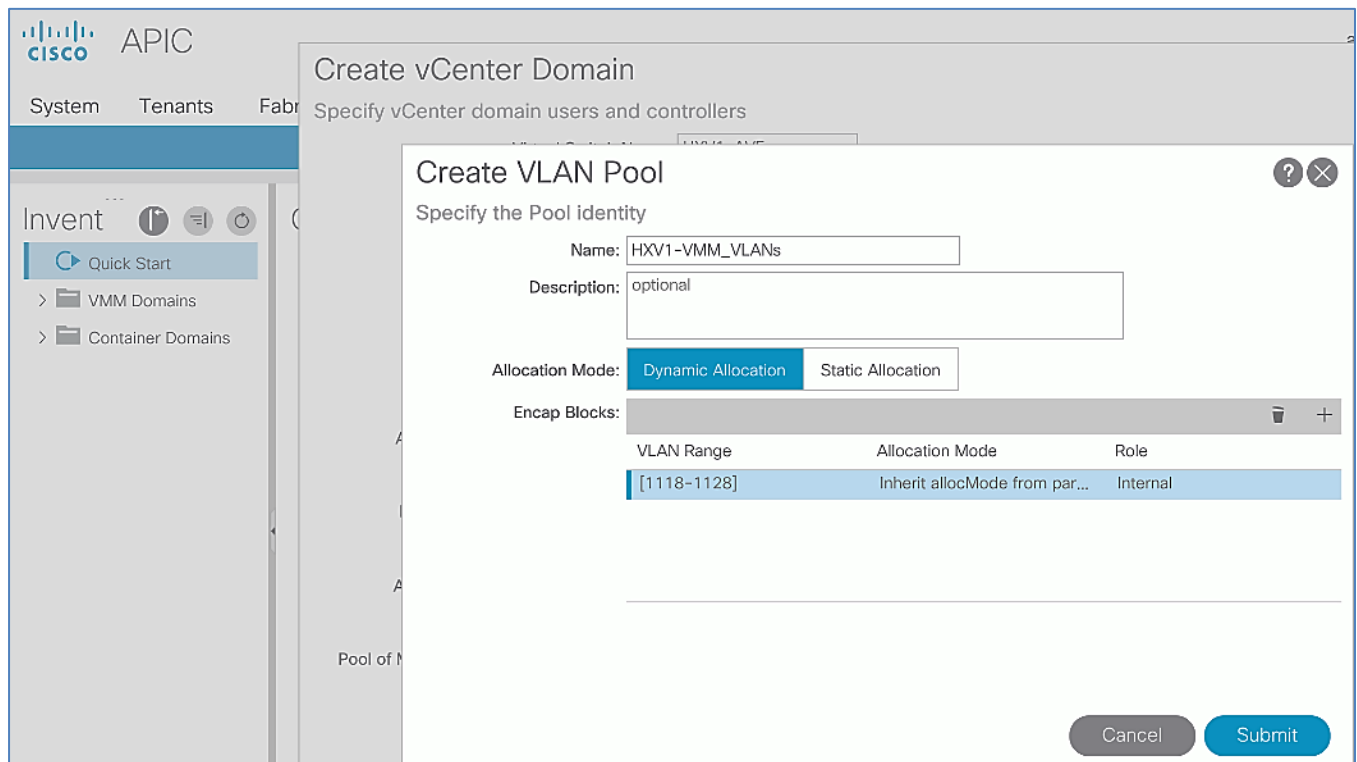
To setup a new VMM domain in ACI where the Cisco AVE will be deployed, follow these steps:

1. Use a browser to navigate to APIC's Web GUI. Log in using the **admin** account.
2. From the top menu, navigate to **Virtual Networking**.
3. From the left navigation pane, select **Quick Start**. From the right-window pane, click **(VMware hypervisor) Create a vCenter Domain Profile**.
4. In the **Create vCenter Domain** window, for the **Virtual Switch Name**, specify a name (for example, HXV1-AVE). This is the name of the Cisco AVE switch in VMware vCenter.
5. For Virtual Switch, select Cisco AVE.
6. For Switching Preference, select **Local Switching**.
7. For Default Encap Mode, select VXLAN Mode.
8. For **Attachable Entity Profile**, select the previously created UCS AAEP (for example, HXV-UCS_AAEP).
9. For the **VLAN Pool**, select **Create VLAN Pool** from the pull-down menu options. In VXLAN mode, the internal VLAN range is used for private VLAN allocations on the distributed virtual switch used Cisco AVE. These VLANs will not be seen outside the ESXi host or on the wire.

10. In the **Create VLAN Pool** pop-up window, specify a **Name** (HXV1-VMM_VLANS) for the pool to be associated with Cisco AVE. For **Allocation Mode**, select **Dynamic Allocation**. For **Encap Blocks**, click the **[+]** icon on the right side.
11. For the **Encap Blocks**, click the **[+]** icon on the right side of the Encap Block section.
12. In the **Create Ranges** pop-up window, for **Range**, specify a VLAN range for virtual machine networks on Cisco AVE. For **Role**, select **Internal**.



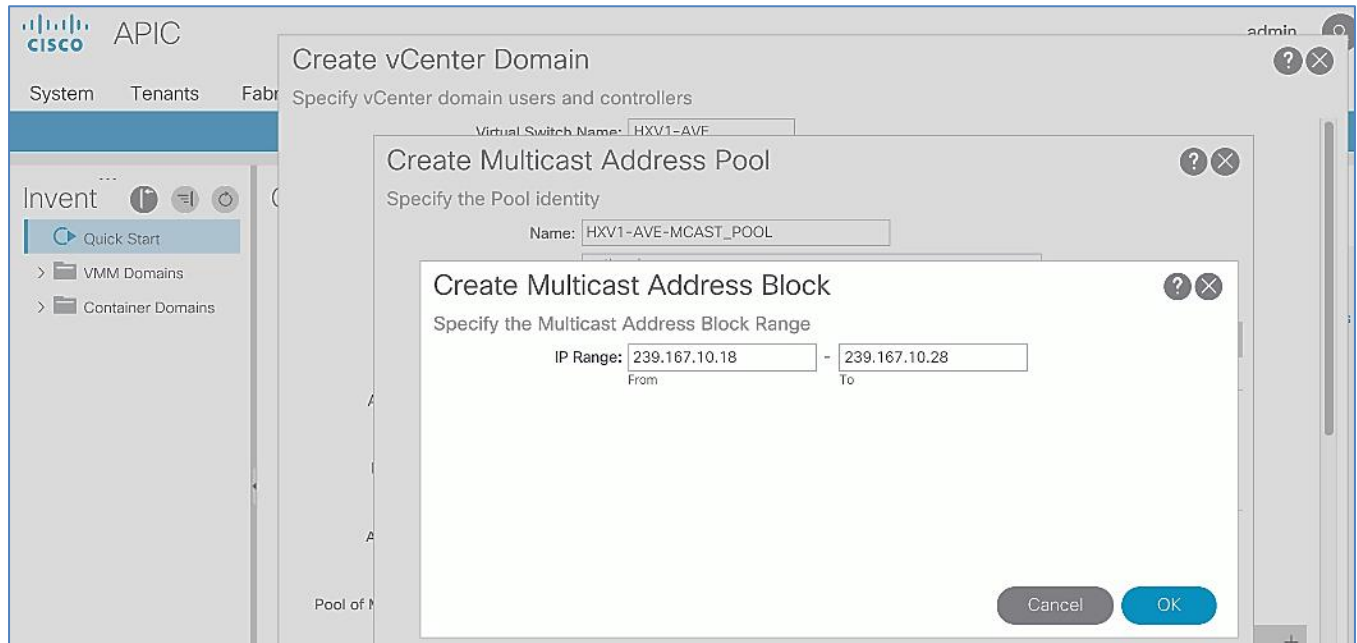
13. Click **OK** to complete the VLAN range configuration and close the **Create Ranges** pop-up window.



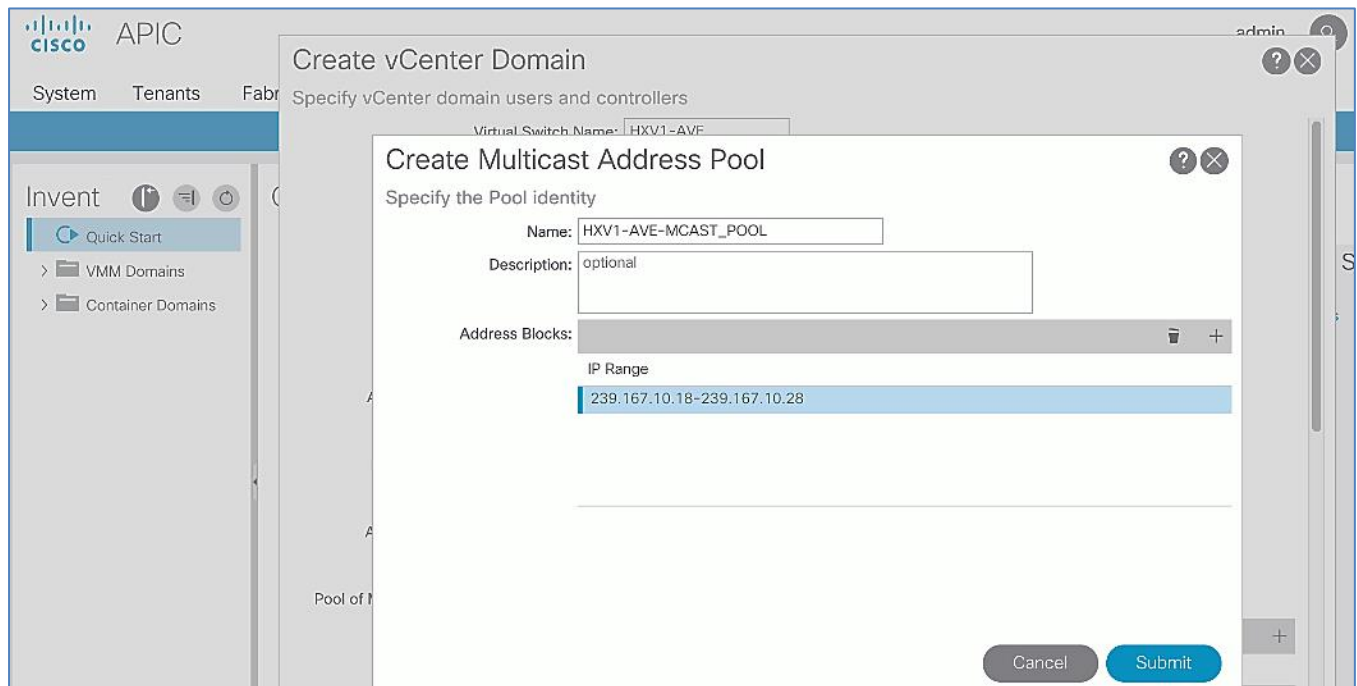
14. Click **Submit** to complete the VLAN pool configuration and close the **Create VLAN Pool** pop-up window.
15. In the **Create vCenter Domain** window, for **AVE Fabric-Wide Multicast Address**, specify an address (for example, 239.167.10.240) outside the multicast address pool defined in the next step.
16. For the **Pool of Multicast Addresses** (one per-EPG), select **Create Multicast Address Pool** from the pull-down menu options.
17. In the **Create Multicast Address Pool** pop-up window, specify a name for the Address Pool (for example, HXV1-AVE-MCAST_POOL).

18. For **Address Blocks**, click the **[+]** on the right side of the Address Blocks section.

19. In the **Create Multicast Address Block** pop-up window, specify a range (239.167.10.18–.28).



20. Click **OK** to create the multicast address block and close the **Create Multicast Address Block** window.



21. Click **Submit** to complete and close the **Create Multicast Address Pool** pop-up window.

Create vCenter Domain

Specify vCenter domain users and controllers

Virtual Switch Name: HXV1-AVE

Virtual Switch: VMware vSphere Distributed Switch | Cisco AVS | **Cisco AVE**

AVE Time Out Time (seconds): 30

AVE HeartBeat Interval (seconds): 5

Host Availability Assurance: ☐

Switching Preference: No Local Switching | **Local Switching**

Default Encap Mode: VLAN mode | **VXLAN mode**

Associated Attachable Entity Profile: HXV-UCS_AAEP

Delimiter:

Endpoint Retention Time (seconds): 0

VLAN Pool: HXV1-VMM_VLANS(dynamic)

AVE Fabric-Wide Multicast Address: 239.167.10.240
Must Use a Multicast Address different from the Pool of Multicast Addresses.

Pool of Multicast Addresses (one per-EPG): HXV1-AVE-MCAST_POOL

Security Domains:

Name	Description
------	-------------

Cancel Submit

22. In the **Create vCenter Domain** window, scroll-down to **vCenter Credentials** and click the **[+]** icon on the right side to add a vCenter Account Profile.
23. In the **Create vCenter Credential** pop-up window, specify a **Name** for the credentials, along with the appropriate account **Username** and **Password**.

Create vCenter Credential

Specify account profile

Name: Administrator

Description: optional

Username: administrator@hvx.com

Password:

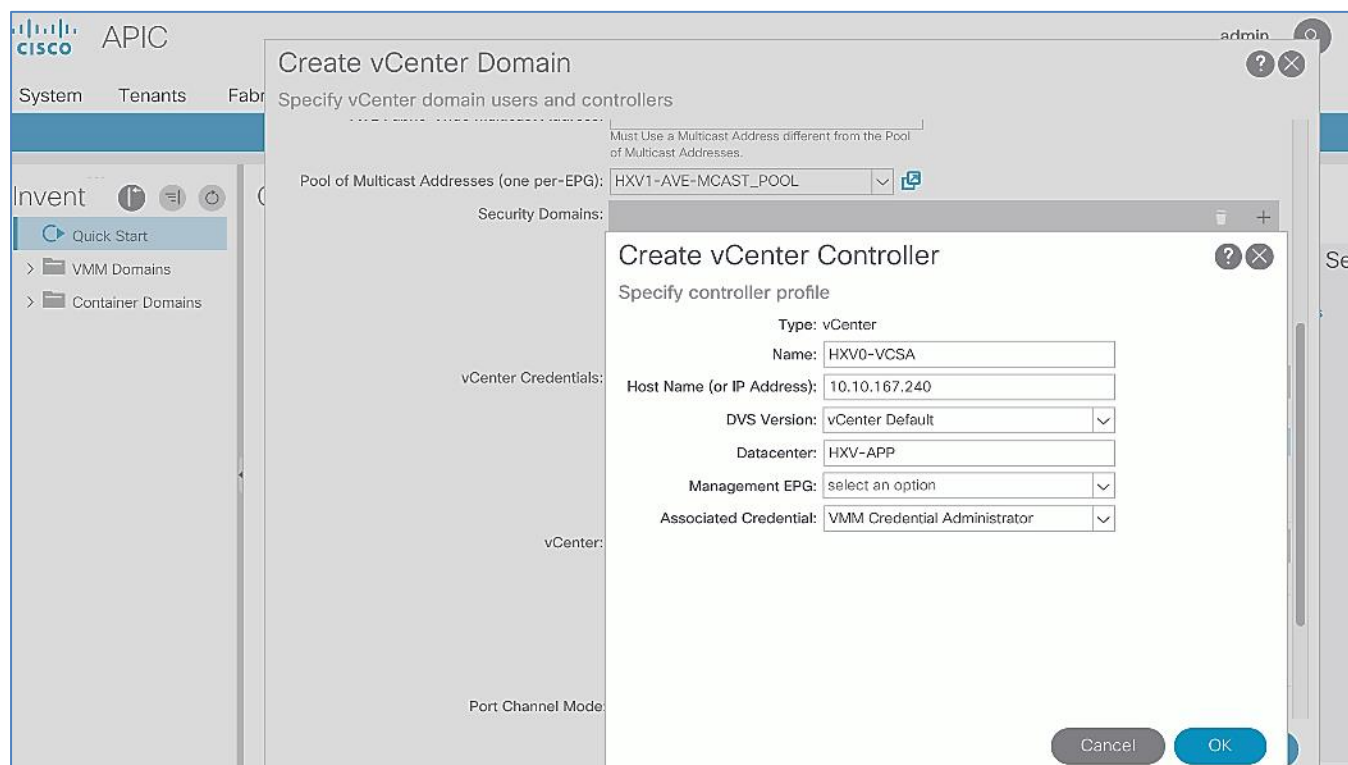
Confirm Password:

Cancel OK

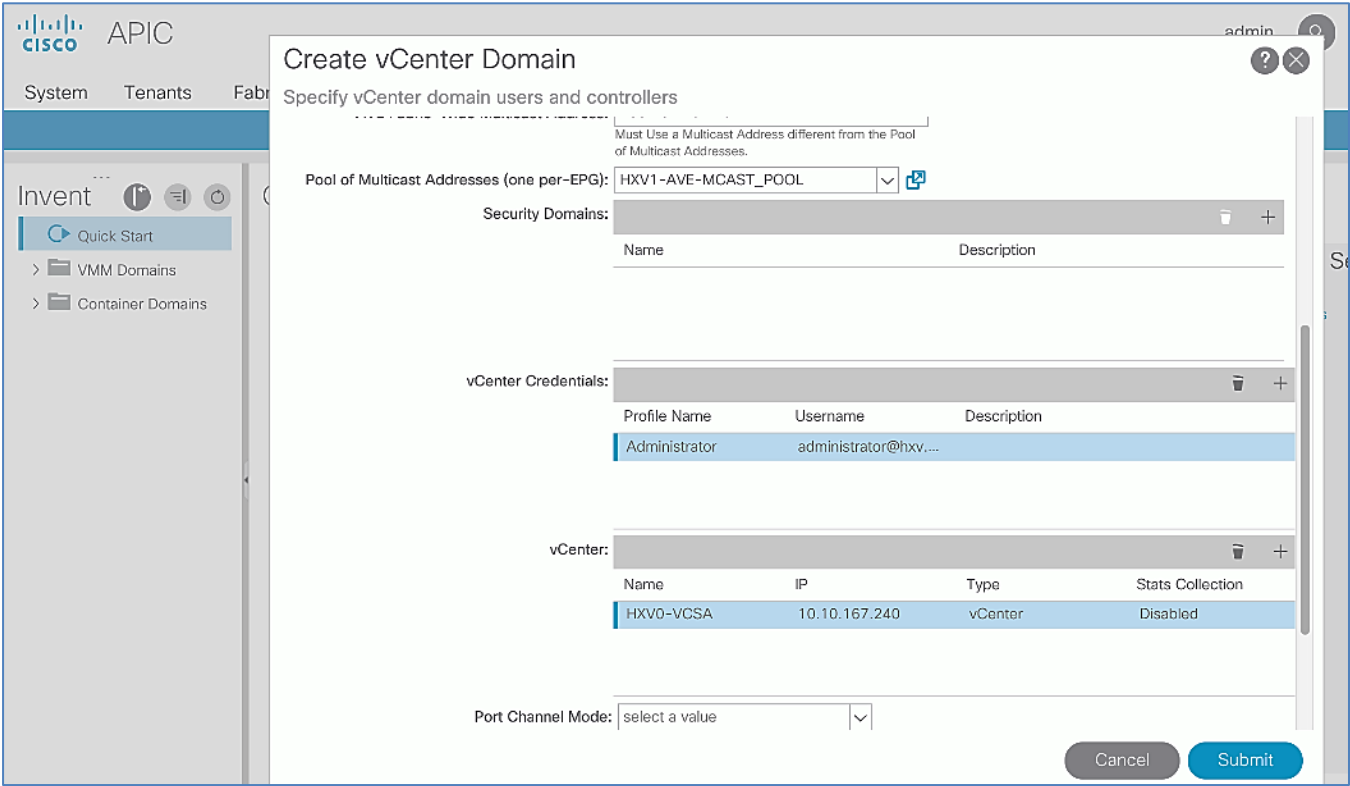


The Administrator account is used in this example, but an APIC account can be created within the vCenter to enable the minimum set of privileges. For more information, see the ACI Virtualization Guide on cisco.com.

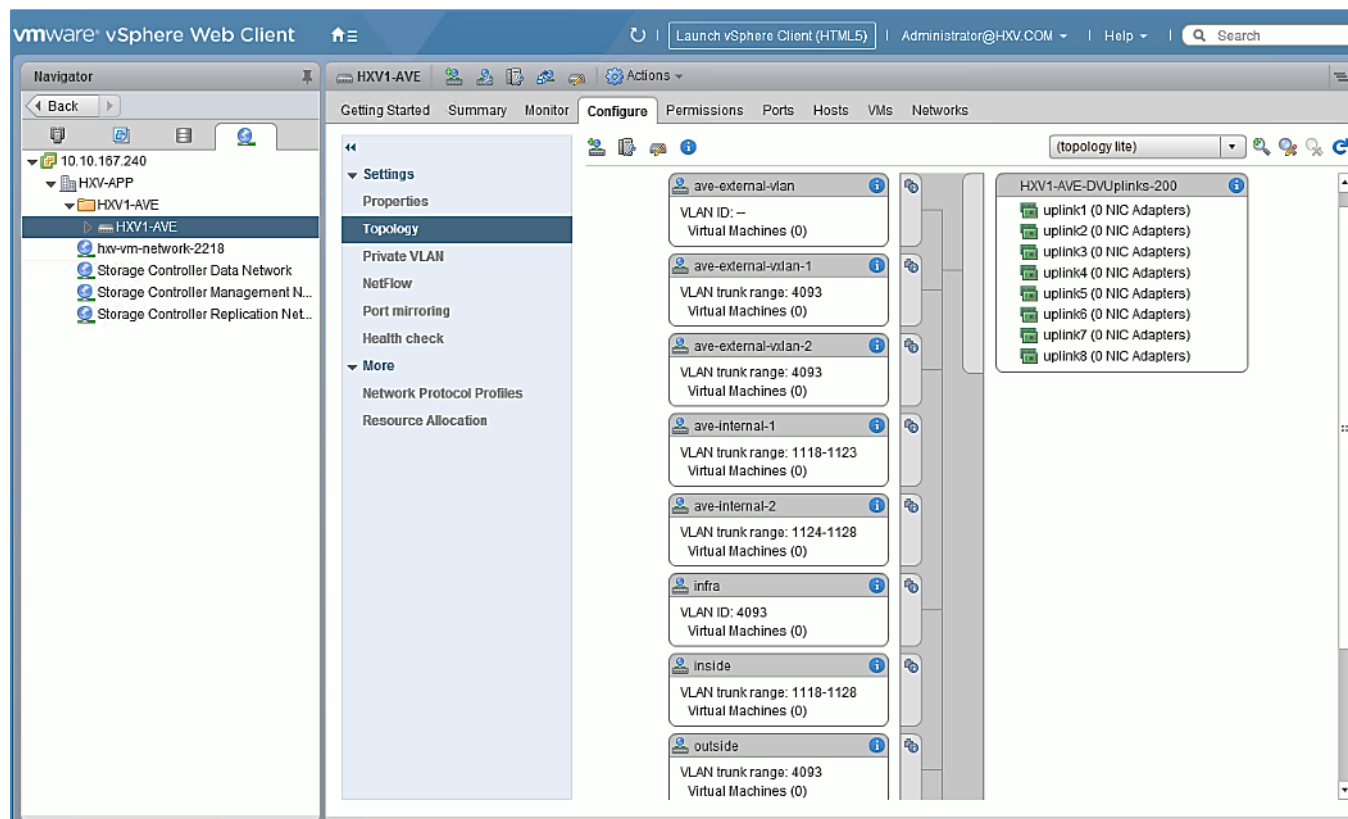
24. Click **OK** to close the **Create vCenter Credential** pop-up window.
25. In the **Create vCenter Domain** window, scroll-down to **vCenter** and click the **[+]** icon on the right side to configure a vCenter Controller.
26. In the **Create vCenter Controller** pop-up window, enter a **Name** (HXV0-VCSA) for the vCenter. For **Host Name (or IP Address)**, enter the vCenter IP or Hostname. For **DVS Version**, leave it as **vCenter Default**. For **Datacenter**, enter the Datacenter name provisioned on the vCenter. Name is case-sensitive. For **Associated Credential**, select the vCenter credentials created in the last step (Administrator).



27. Click **OK** to close the **Create vCenter Controller** pop-up window.



- 28. Leave everything else as is in the **Create vCenter Domain** window and click **Submit**.
- 29. Log into the VMware vCenter server. Click the vSphere Web Client of your choice. Log in as an **Administrator**.
- 30. Navigate to **Networking**.



31. Verify that a Cisco AVE has been created under the Datacenter.

Apply Virtual Switch Policies for Cisco AVE

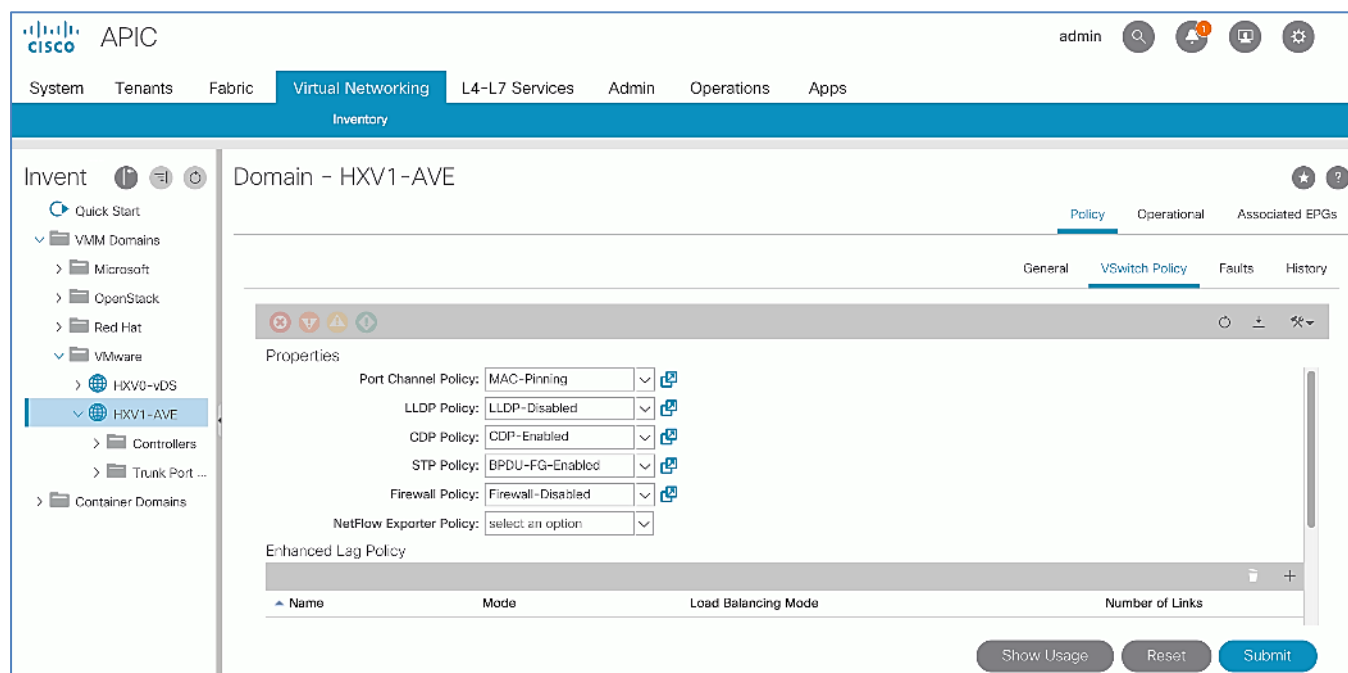
To apply pre-configured policies to VMM domain for Cisco AVE, follow these steps:

1. Navigate to Virtual Networking > Inventory > VMM Domains > VMware.
2. From the left navigation pane, select the newly created VMM Domain (HXV1-AVE).
3. In the right window pane, select the **Policy > vSwitch Policy** tab.
4. **For Port Channel Policy**, select the pre-configured **MAC-Pinning** policy from the pull-down menu options.



If the AVE is connected through a Cisco UCS FI, MAC Pinning-Physical-NIC-Load is not supported.

5. For **LLDP Policy**, select the pre-configured **LLDP-Disabled** policy.
6. For **CDP Policy**, select the pre-configured **CDP-Enabled** policy.
7. For **STP Policy**, select the pre-configured **BPDU-FG-Enabled** policy.
8. For **Firewall Policy**, select the pre-configured **Firewall-Disabled** policy.
9. Leave everything else as is.

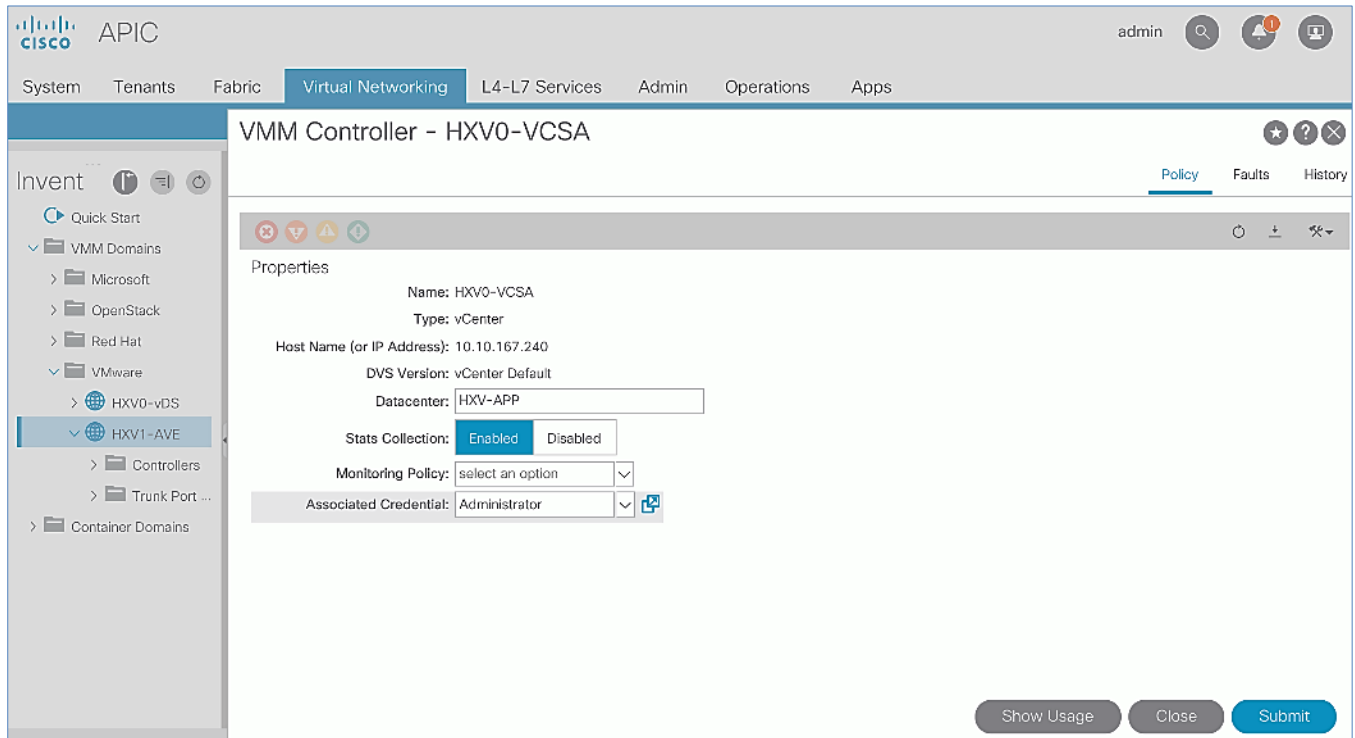


10. Click **Submit** and **Submit Changes** to apply the policy changes and close window.

Enable Statistics on Cisco AVE

To enable statistics collection, follow these steps:

1. Navigate to Virtual Networking > Inventory > VMM Domains > VMware.
2. From the left navigation pane, select the newly created VMM Domain (HXV1-AVE).
3. In the right window pane, select the **Policy > General** tab.
4. In the **vCenter** section, select and double-click the vCenter configured in the previous step.
5. In the **VMM Controller** window, for Stats Collection, select **Enabled**.

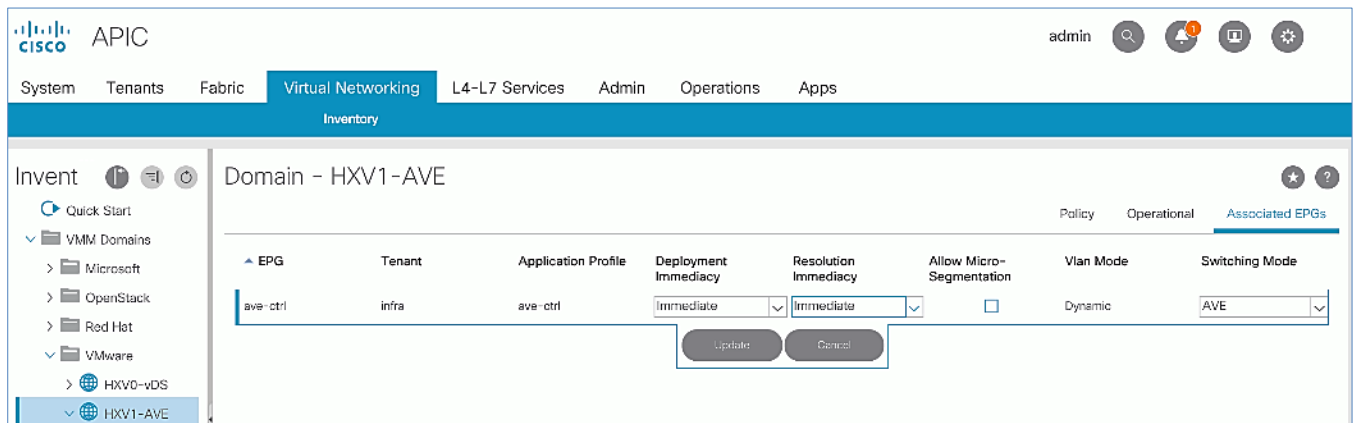


- Click **Submit** and **Submit Changes** to accept the change and close the window.

Configure Deployment and Resolution Immediacy for VLAN Allocation

To enable statistics collection, follow these steps:

- Navigate to Virtual Networking > Inventory > VMM Domains > VMware.
- From the left navigation pane, select the newly created VMM Domain (HXV1-AVE).
- In the right window pane, select the **Associated EPGs** tab.
- Select the ave-ctrl EPG and double-click **Deployment Immediacy** to change it **Immediate**. Double-click **Resolution Immediacy** to change it to **Immediate**.

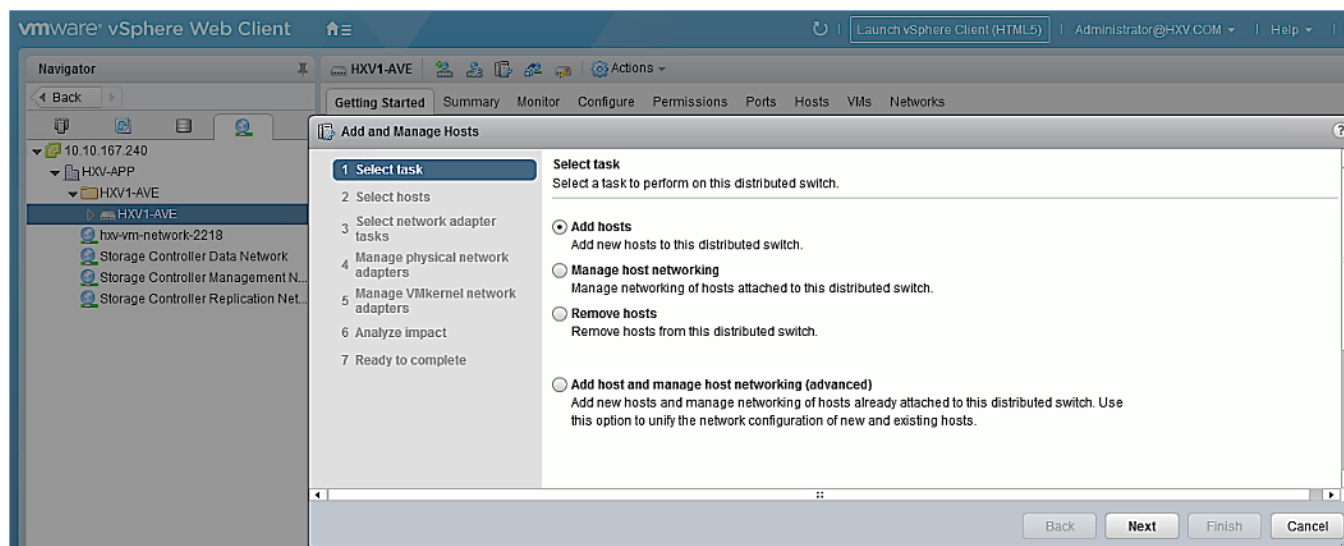


- Click **Update** and **Continue** to apply the changes.

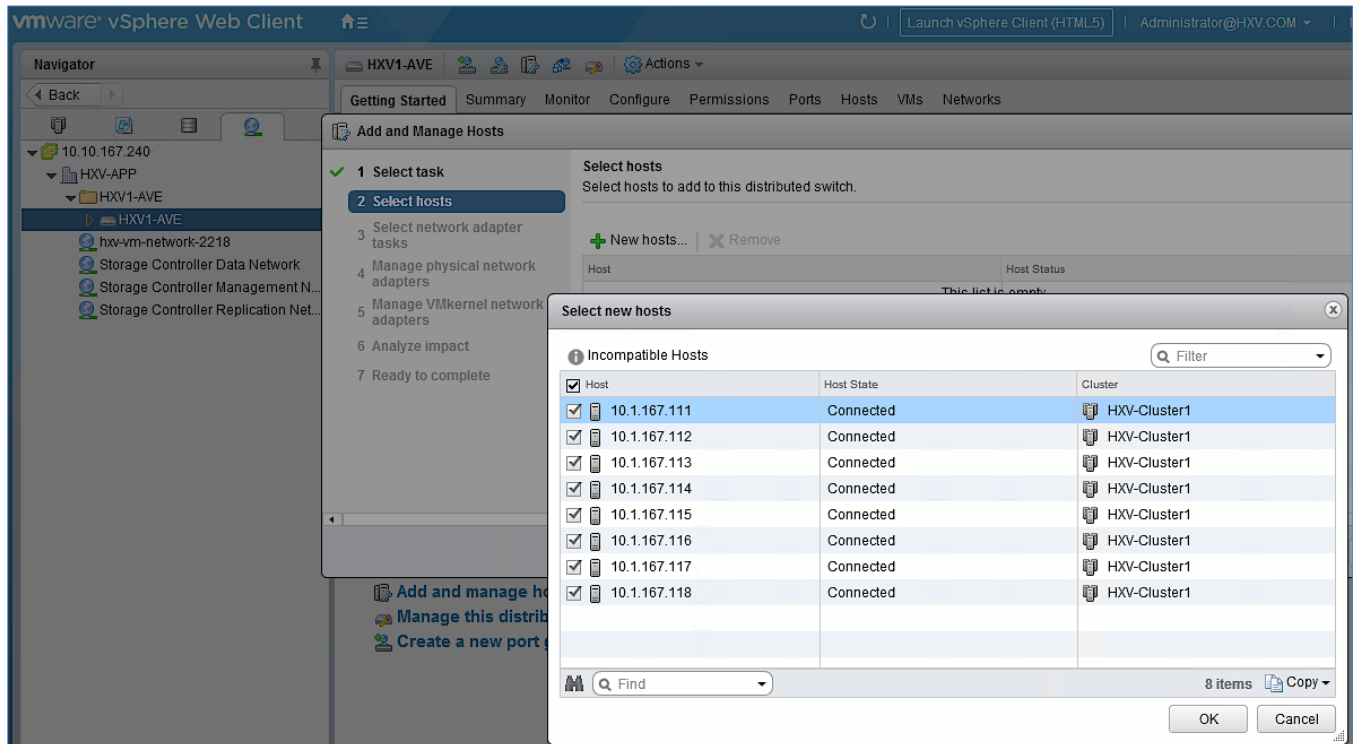
Add HyperFlex ESXi Hosts to Cisco AVE Virtual Switch

To add the HyperFlex ESXi Hosts to the newly created Cisco AVE distributed virtual switch, follow these steps:

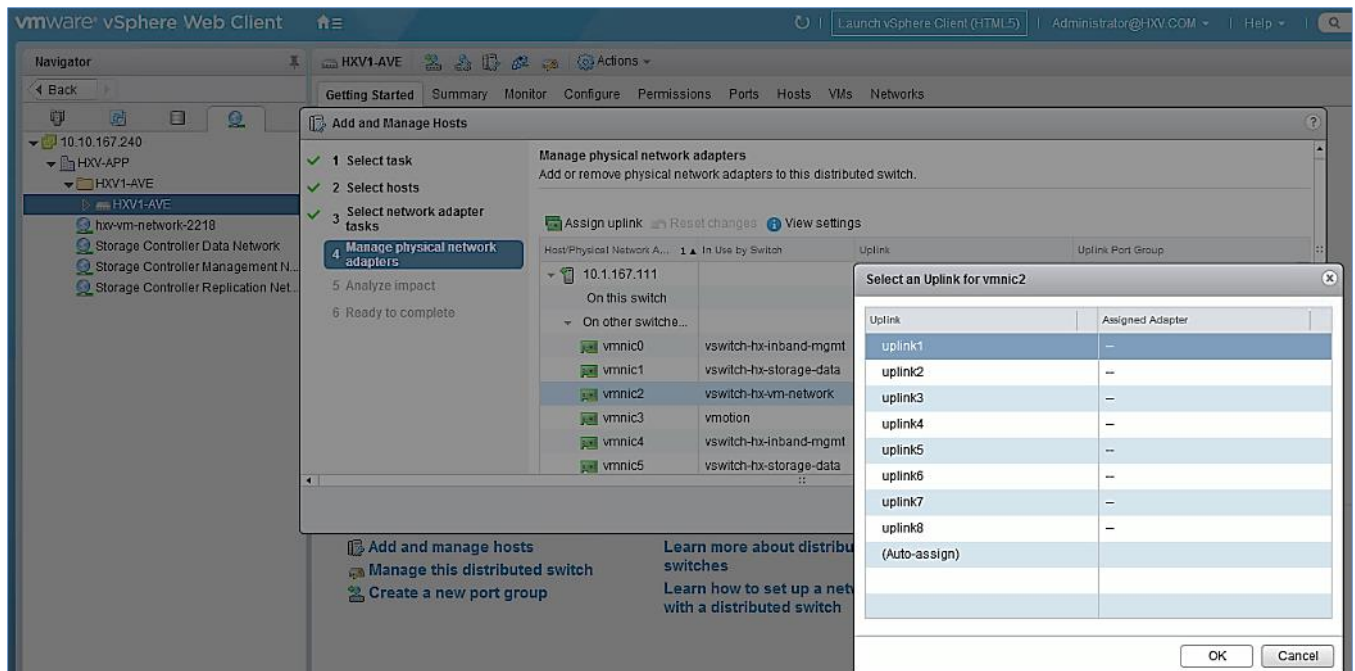
1. Use a browser to navigate to the VMware vCenter server managing the HyperFlex Application cluster. Click the vSphere Web Client of your choice. Log in using an **Administrator** account.
2. From the **Home** screen, select **Networking** in the **Inventories** section.
3. In the left navigation pane, expand the **Datacenter (HXV-APP)** with the newly deployed Cisco AVE. Open the folder and select the APIC deployed Cisco AVE (HXV1-AVE) distributed virtual switch.
4. Right-click the Cisco AVE (HXV1-AVE) distributed virtual switch and select **Add and Manage hosts....**
5. In the **Add and Manage Hosts** wizard, select the **Add hosts** option. Click **Next**.



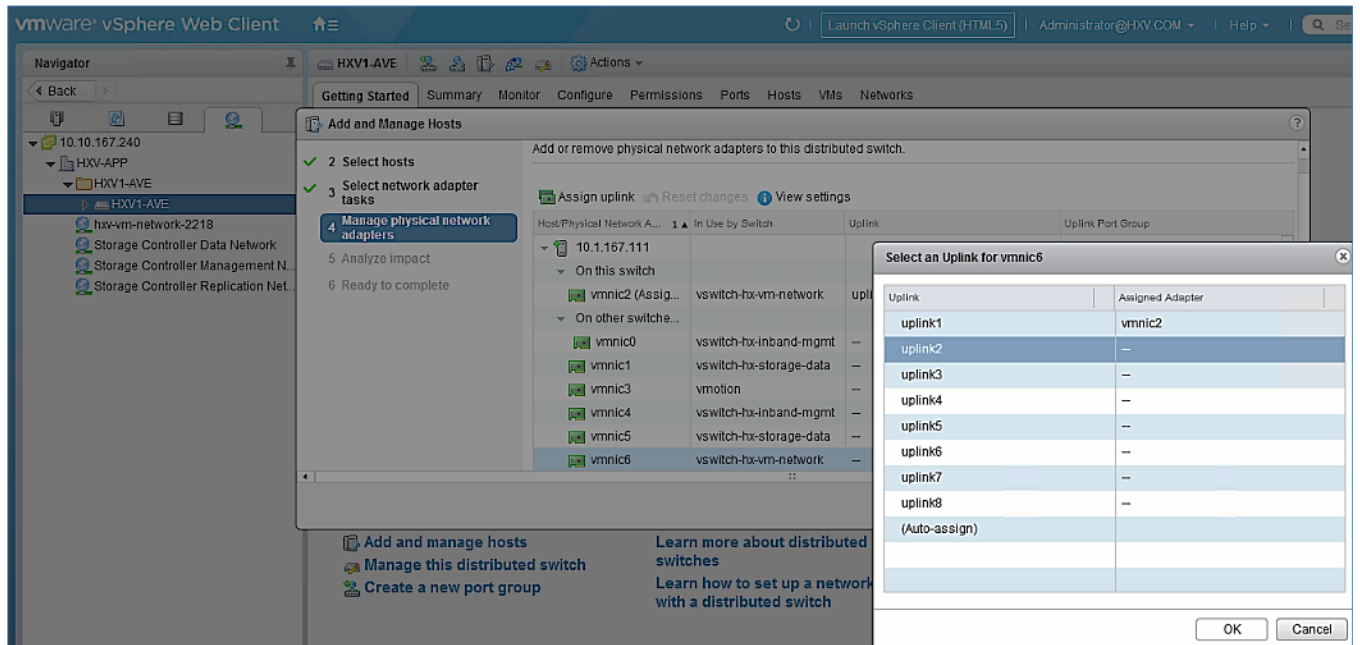
6. In the **Select hosts** screen, click the **[+] New hosts...** to select hosts to add to Cisco AVE switch.
7. In the **Select new hosts** pop-up window, select all the hosts to be added to the Cisco AVE switch.



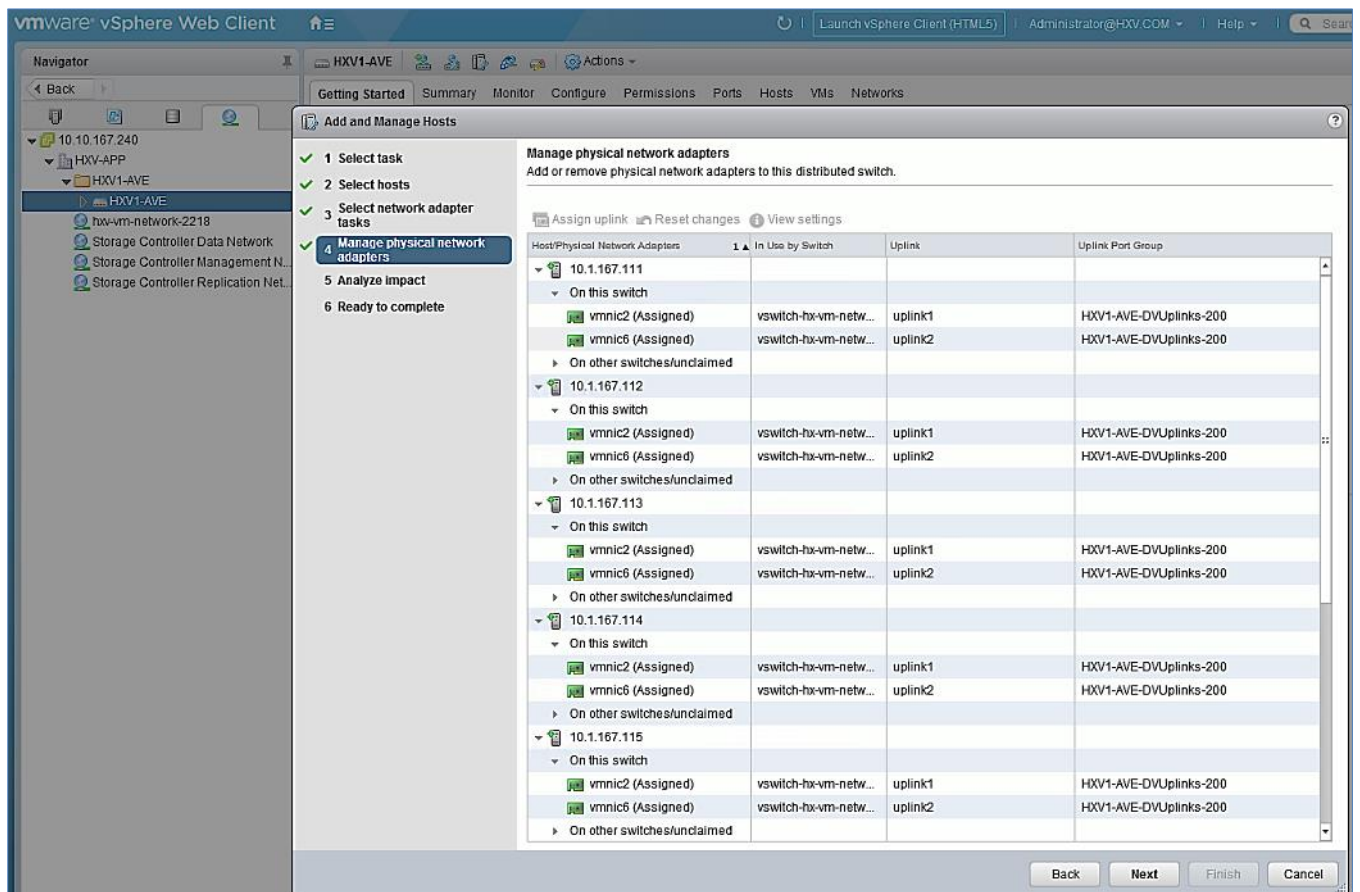
8. Click **OK**. Click **Next**.
9. In the Select network adapter tasks screen, select Manage physical adapters. Click Next.
10. For the first host, under the **Host/Physical Network Adapters** column, select `vmnic2`. Click the **Assign uplink** from the menu above.
11. In the **Select an Uplink for vmnic2** pop-up window, leave **uplink 1** selected and click **OK**.



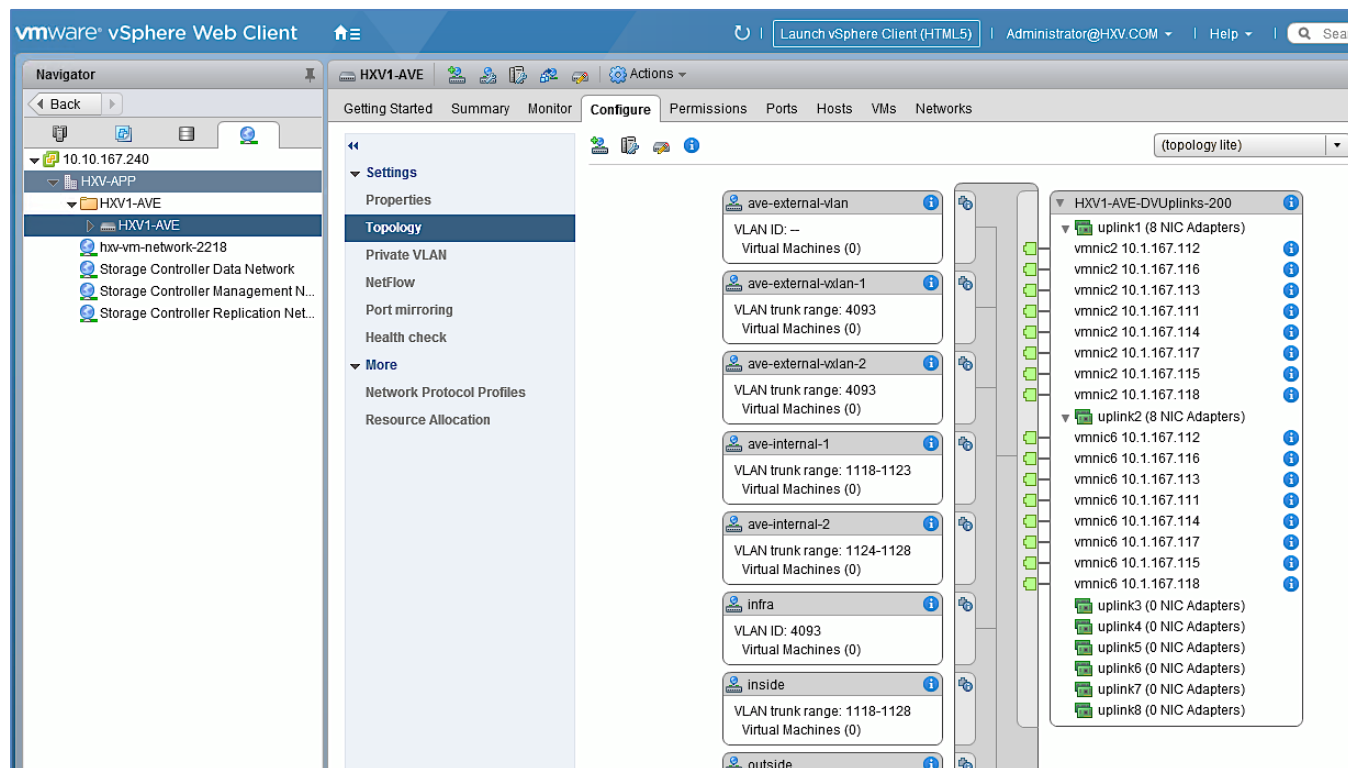
12. Repeat steps 1-11 to assign a second uplink to the same host by selecting `vmnic6` as **uplink2**. Click **Next**.



13. Repeat steps 1-12 to add remaining hosts to migrate hosts and virtual machine network vmnics to Cisco AVE.
14. Scroll down and verify that vmnics for virtual machine networks on each host has been assigned as **uplink1** and **uplink2** on all hosts. Click **Next**.



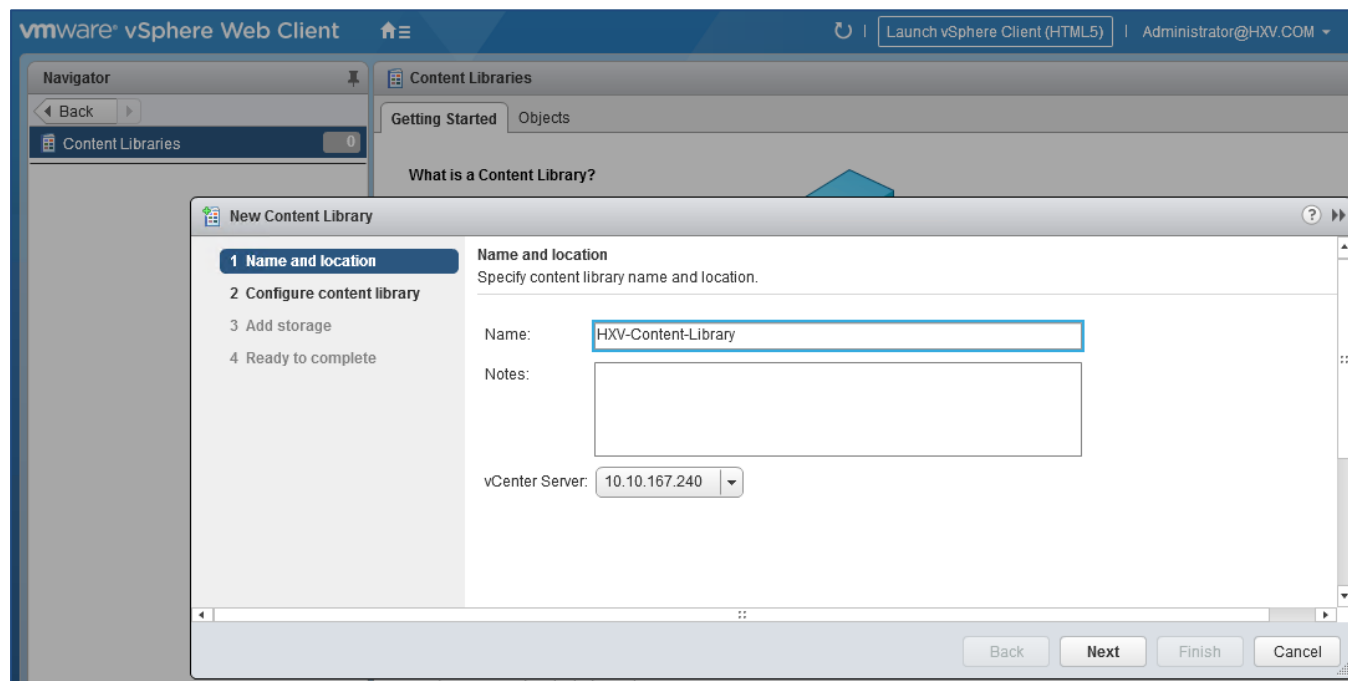
15. In the **Analyze impact** screen, click **Next**.
16. In the **Ready to complete** screen. Click **Finish** to apply.
17. Verify the uplinks have been migrated to Cisco AVE.



Download Cisco AVE OVF File to VMware vCenter

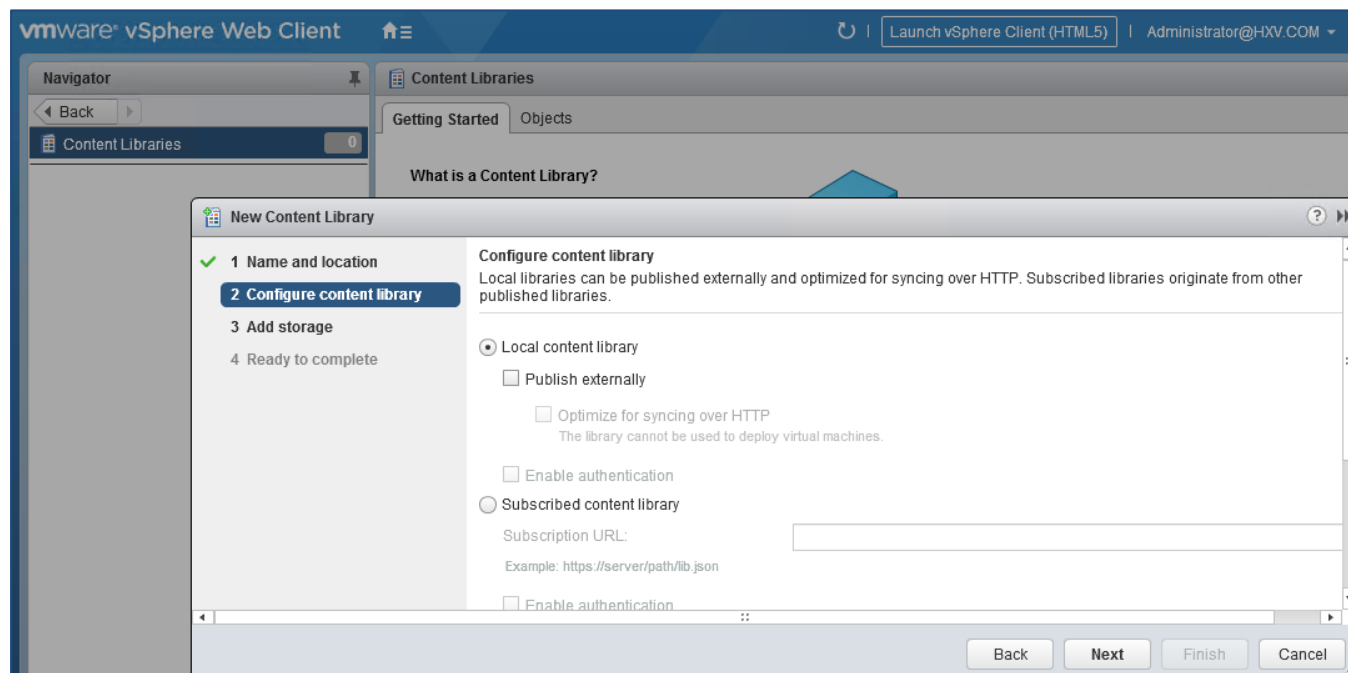
To upload the Cisco AVE OVF to VMware vCenter, follow these steps:

1. Use a browser to navigate to the VMware vCenter Server managing the HyperFlex Application cluster. Click the vSphere web client of your choice and log in using an **Administrator** account.
2. From the **Home** screen, select **Content Libraries** from the **Inventories** section.
3. From the right window pane, select **Create a new content library** or go directly to **Import Item** if one already exists. Click **Next**.
4. In the **New Content Library** pop-up window, for **Name**, specify a name for the new Content Library.



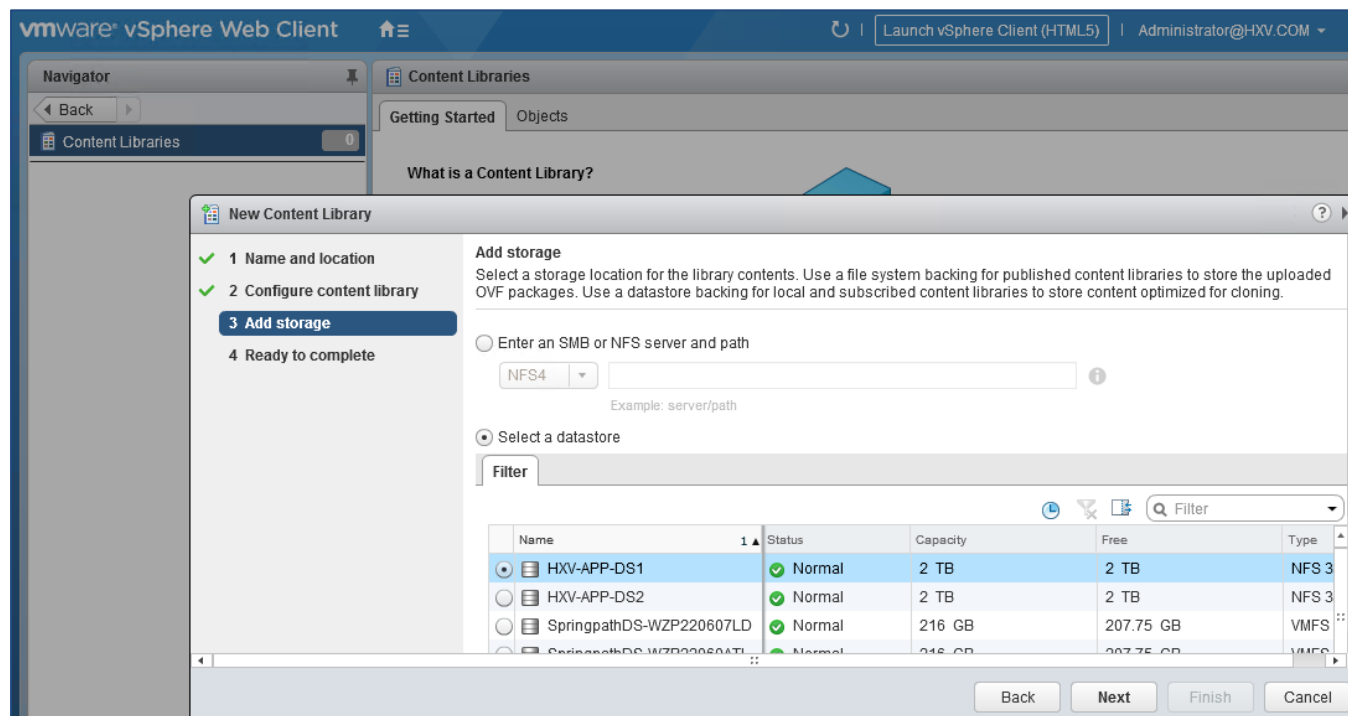
5. Click **Next**.

6. In the Configure Content Library screen, select Local content library.

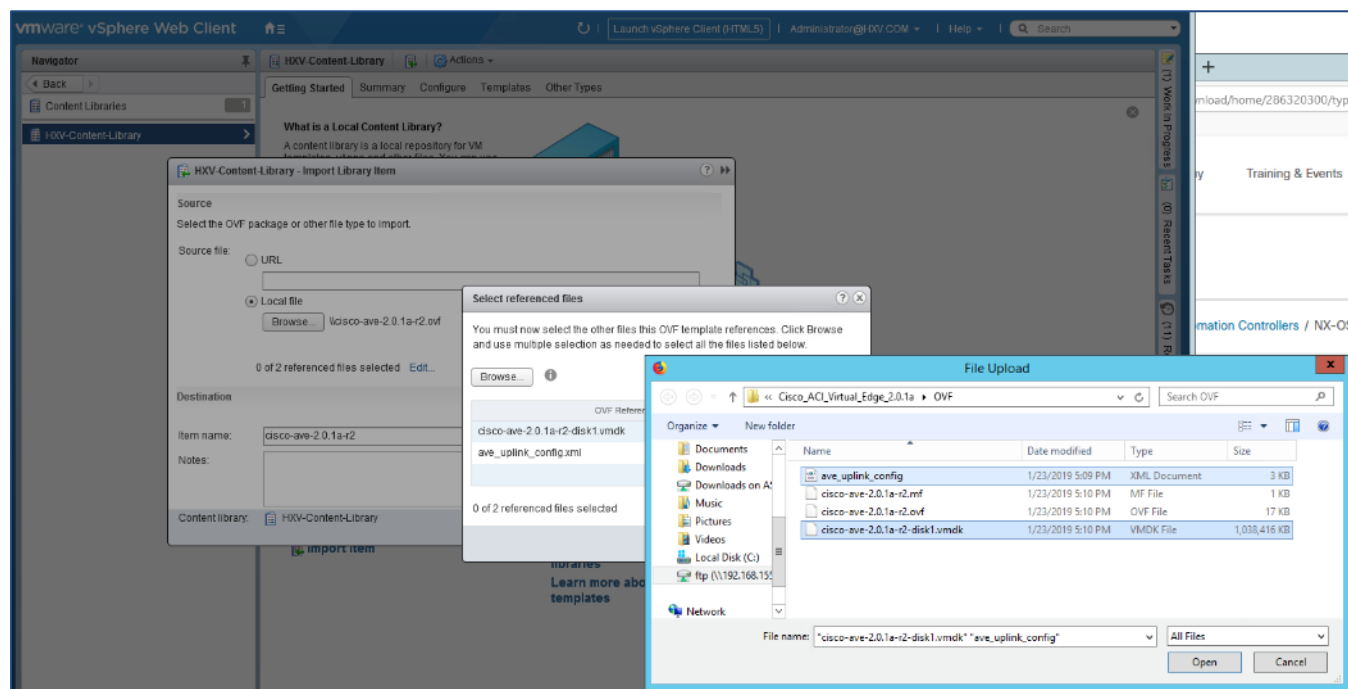


7. Click **Next**.

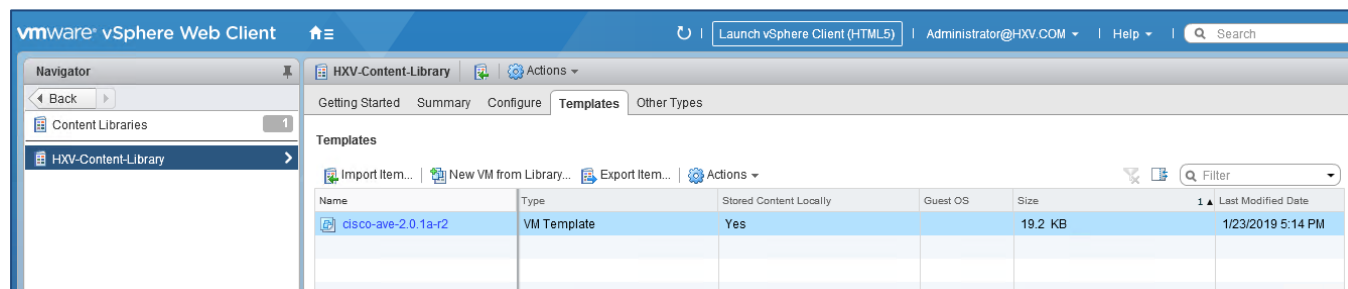
8. In the **Add storage** screen, select the radio button to **Select a datastore**.



9. Click **Next**.
10. Review and click **Finish** to complete.
11. In the left navigation pane, select the newly created Content Library.
12. In the right window, click **Import Item**.
13. In the **Import Item to Content Library** pop-up window, select the source and click **OK**.
14. In the **Select referenced files** pop-up window, select the referenced files needed by clicking **Browse**. Click **OK**.



15. Click **OK**. Click **OK** again to import the OVF file.
16. When the OVF file is uploaded to the content library, it appears in the work pane under the **Templates** tab.



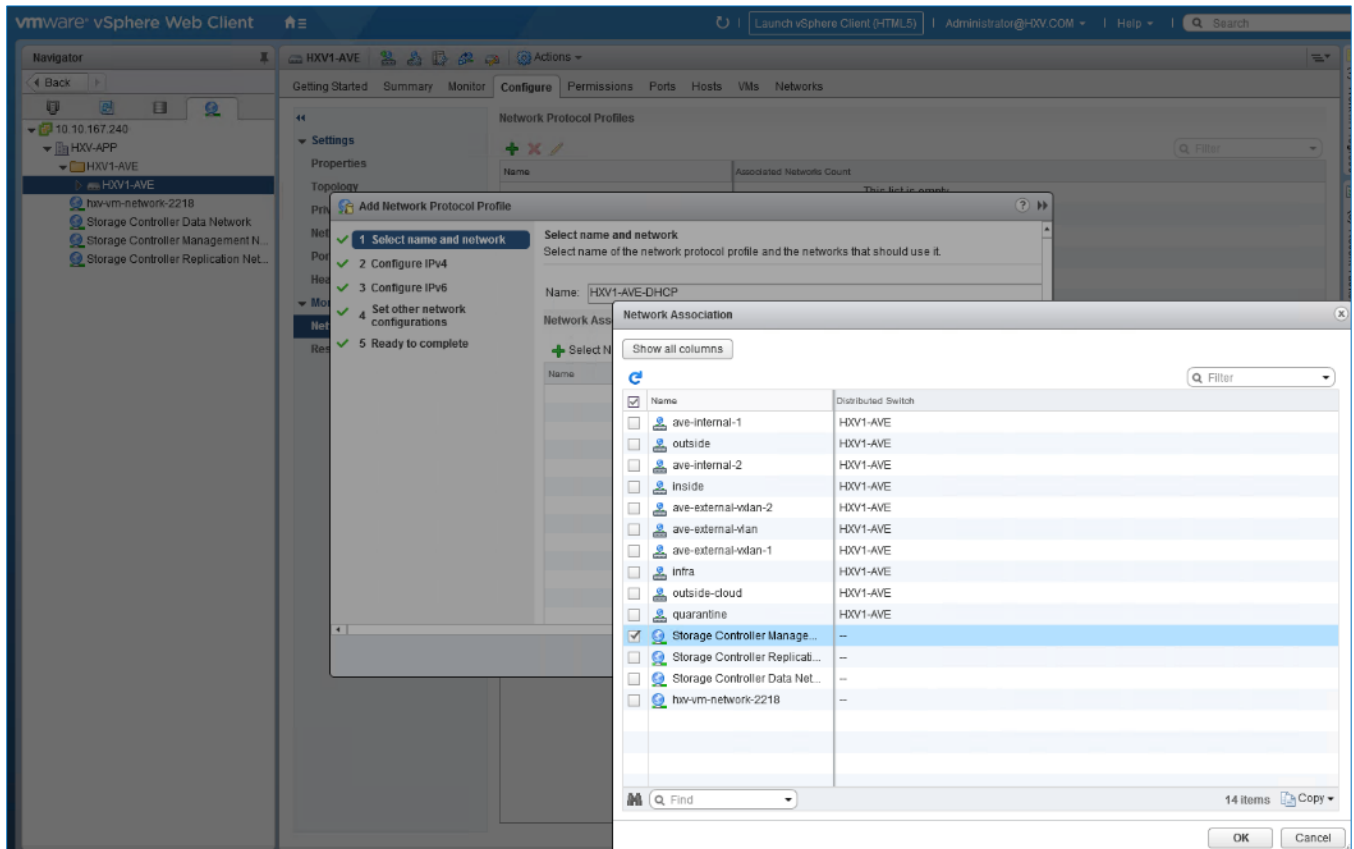
Setup Networking for Cisco AVE Virtual Machines

The networking setup for deploying Cisco AVE Virtual machines are as follows:

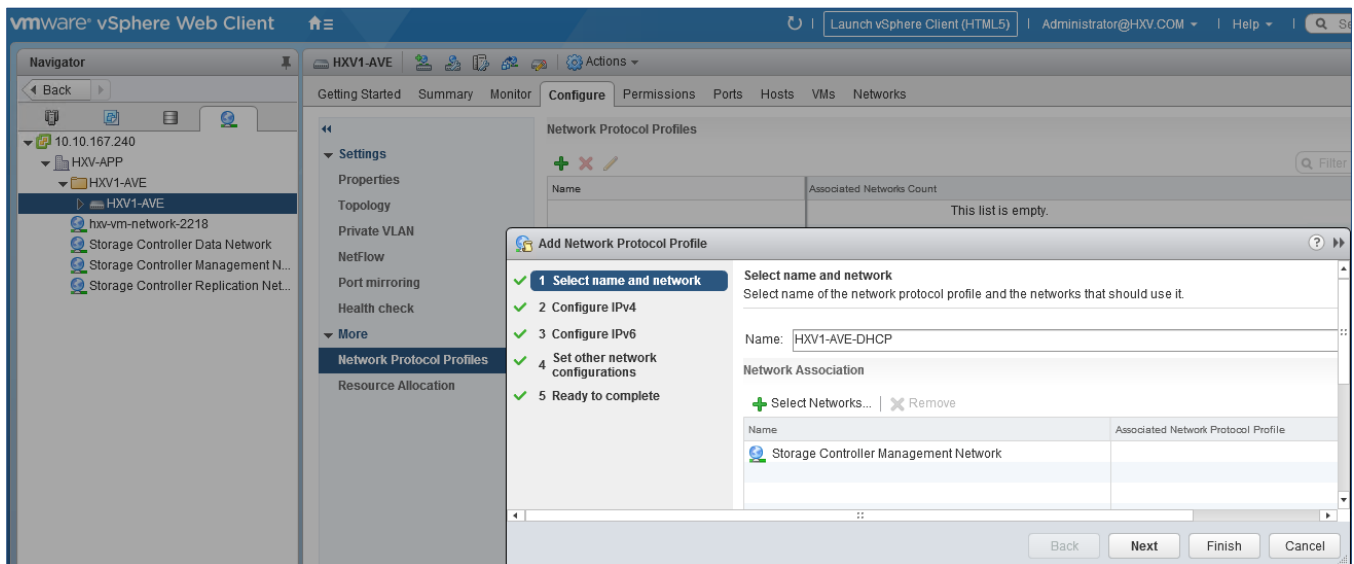
- Allocate a Management IP Address for each Cisco AVE Virtual Machine, one per host. The addresses are allocated via DHCP, either using an existing DHCP server or VMware vCenter.
- Add/Verify Virtual Networking for Cisco AVE virtual machines. The addresses are part of the in-band management network used for (1) ESXi management and (2) HX Storage Controller Management. Cisco AVE management could also be done by creating a separate AVE management network dedicated to AVE hosts with appropriate routing enabled in the ACI fabric.
- Setup VMware vCenter for DHCP to allocate IP addresses to Cisco AVE VM. Verify that there is no IP address overlap with other devices in the same network. The address block allocated for Cisco AVE is: 10.1.167.161 – 10.1.167.164/24.

To setup VMware vCenter for DHCP, follow these steps:

1. Use a browser to navigate to the VMware vCenter Server managing the HyperFlex Application cluster. Click the vSphere web client of your choice and log in using an **Administrator** account.
2. From the **Home** screen, select **Networking** from the **Inventories** section.
3. From the left navigation pane, expand and select the Cisco AVE (HXV1-AVE) virtual distributed switch.
4. In the right window pane, navigate to **Configure > Network Protocol Policies**.
5. Click the **[+]** to add a new network control profile.
6. In the **Add Network Protocol Profile** wizard, specify a **Name** (HXV1-AVE-DHCP) for the profile. In the **Network Associations** section, click **[+] Select Networks**.
7. In the **Network Associations** pop-up window, select the management port-group for Cisco AVE.

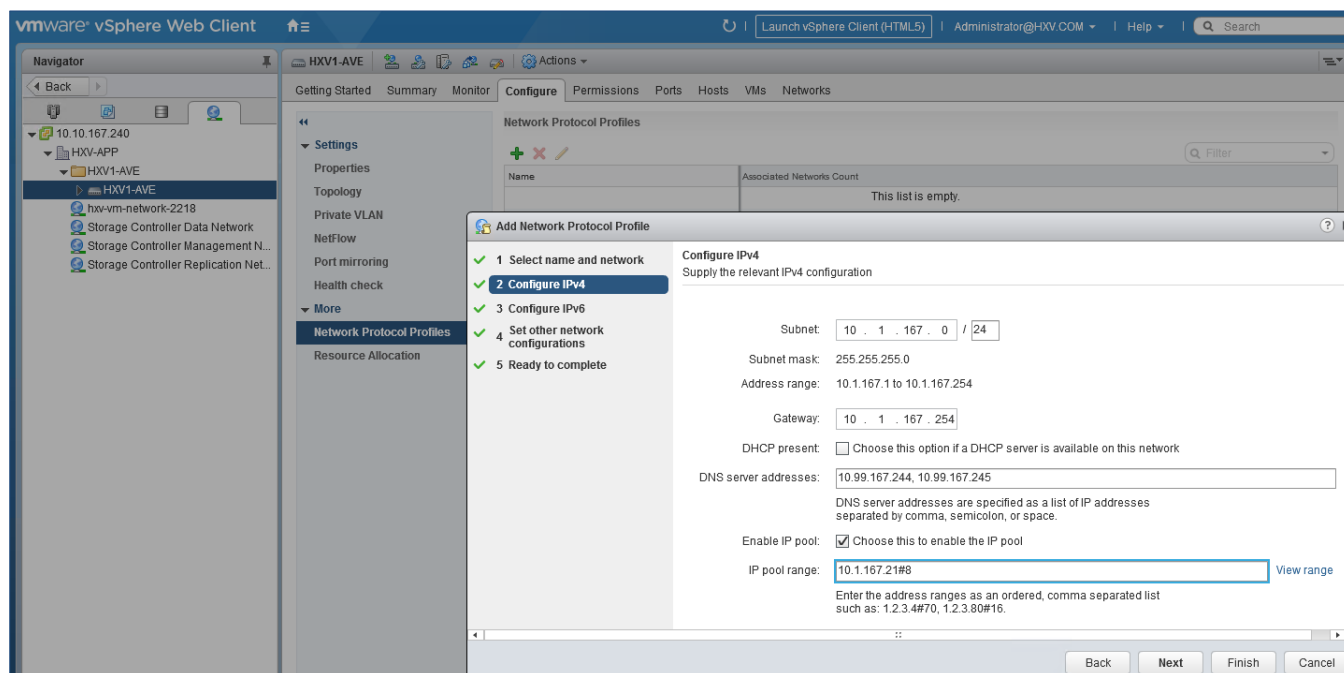


8. Click **OK**.



9. Click **Next**.

10. In the **Configure IPv4** screen, specify the IP Pool Subnet, Gateway and range of IP addresses. Select **Enable IP Pool** check box. Click **View Range** to view the exact IP addresses in the pool.



11. Click **Next**.
12. In the **Configure IPv6** screen, click **Next**.
13. In the **Set other network configurations** screen, enter DNS information. Click **Next**.
14. In the **Ready to complete** screen, review the configuration and click **Finish** to complete.

Deploy Cisco ACI vSphere Plug-in

Cisco AVE can be deployed using Cisco ACI vCenter Plug-in, VMware PowerCLI, Python Script. In this design, Cisco ACI vCenter Plug-in is used. The plug-in also exposes a subset of APIC capabilities to VMware vCenter that are relevant to Virtualization Administrators and provides an interface to manage the ACI Fabric from VMware vCenter.

Prerequisites

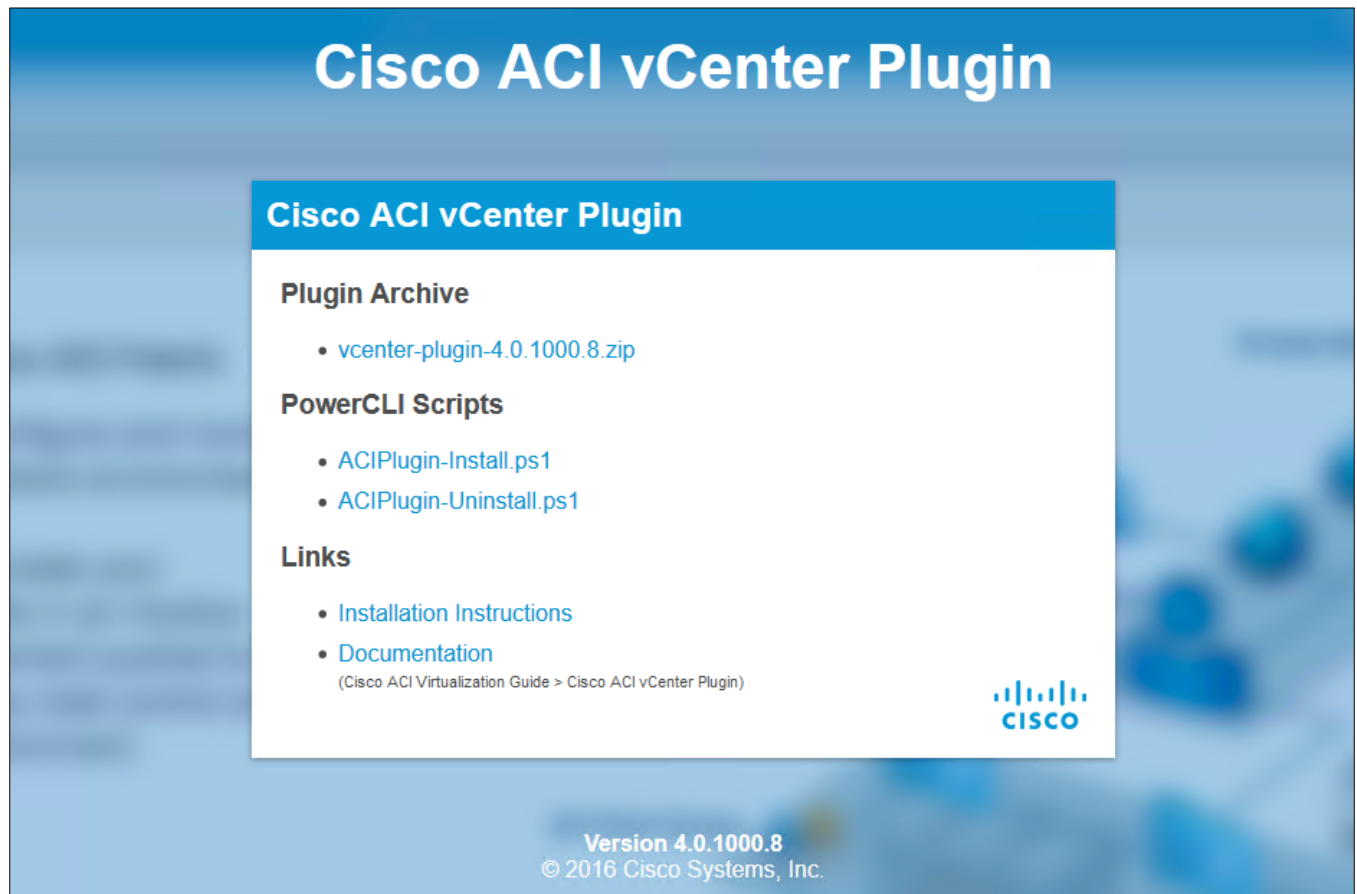
The prerequisites for installing the Cisco ACI Plug-in on VMware vCenter are as follows:

- At least one VMM domain should already exist between the APIC and the vCenter where the plug-in is being installed.
- HTTPS traffic must be allowed between VMware vCenter server and Cisco APIC. vCenter will directly download the plug-in using HTTPS.
- VMware PowerCLI installed on a Windows machine. The PowerShell scripts for installing the plug-in will be executed from the PowerCLI console.
- VMware vCenter 6.0U3 or higher is recommended.

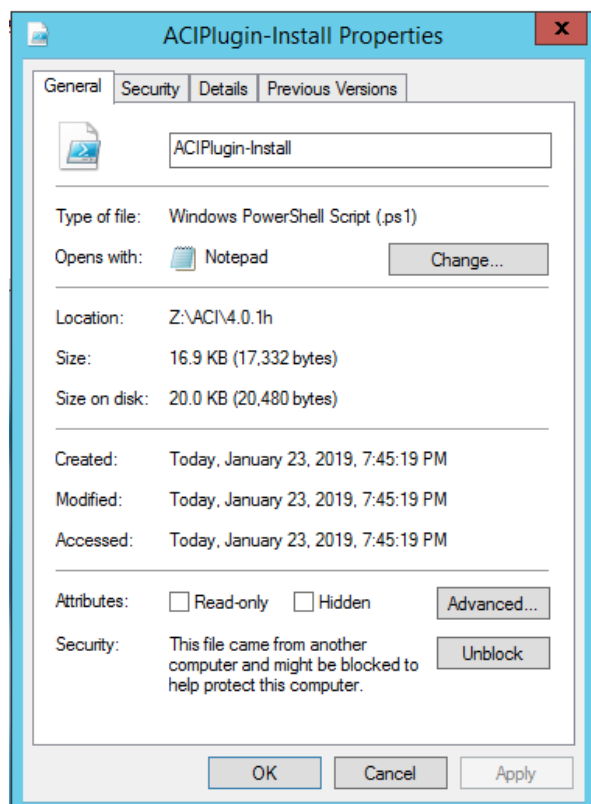
Install Cisco ACI Plug-in

To install the Cisco AC Plug-in for VMware vCenter, follow these steps:

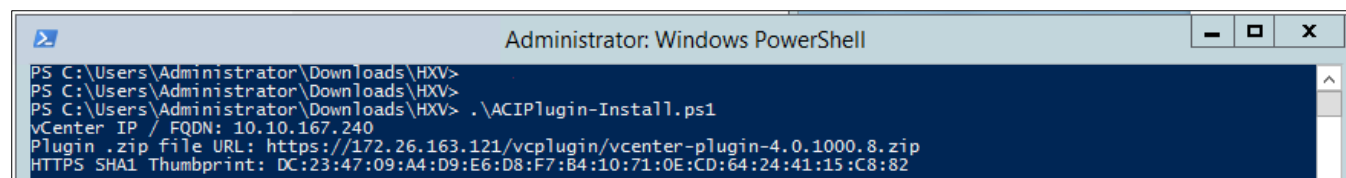
1. Use a web browser to navigate to the APIC at <https://<APIC-IP>/vcplugin>.



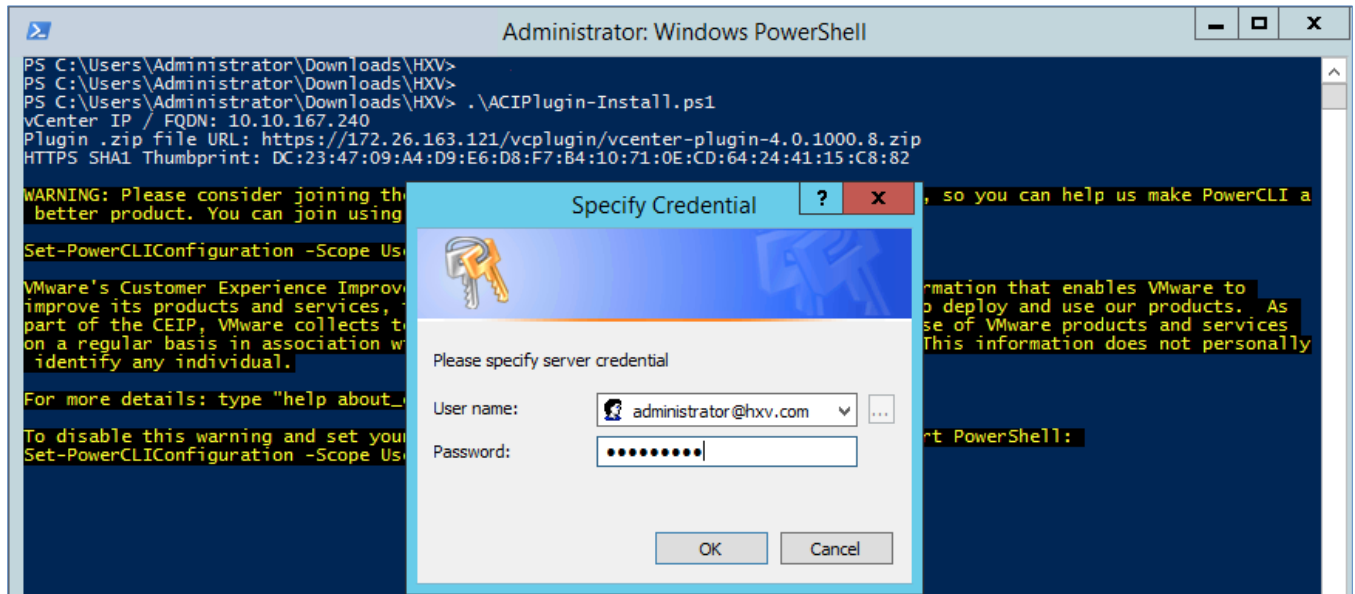
2. Download the **ACIPlugin-Install.ps1** script.
3. Copy the script to the system where it will be executed. The script will be executed from the PowerCLI console.
4. Go to the folder with the installation script.
5. Select the script. Right-click and select **Properties**.
6. In the pop-up window, click **Unblock**.



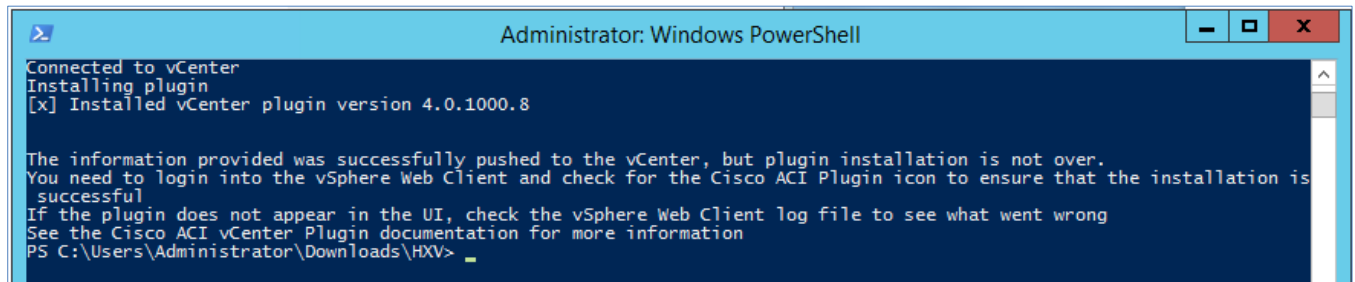
7. Click **Apply**. Click **OK** to close the pop-up.
8. Open the PowerCLI console (**Run as Administrator**) and execute the **ACIPlugin-Install.ps1** script.



9. Enter the address to the vCenter, the http source for the plugin bundle and the HTTPS SHA1 Thumbprint:
https://<APIC-IP-Address-or-Hostname>/vcplugin/vcenter-plugin-3.2.2000.12.zip. To determine the HTTP SHA1 Thumbprint, review the Installation instructions on the same web page as the .zip file; it varies depending on the browser you are using.
10. In the **Specify Credential** pop-window, provide the vCenter Administrator credentials.



11. If the installation was successful, you should see something similar to the following:



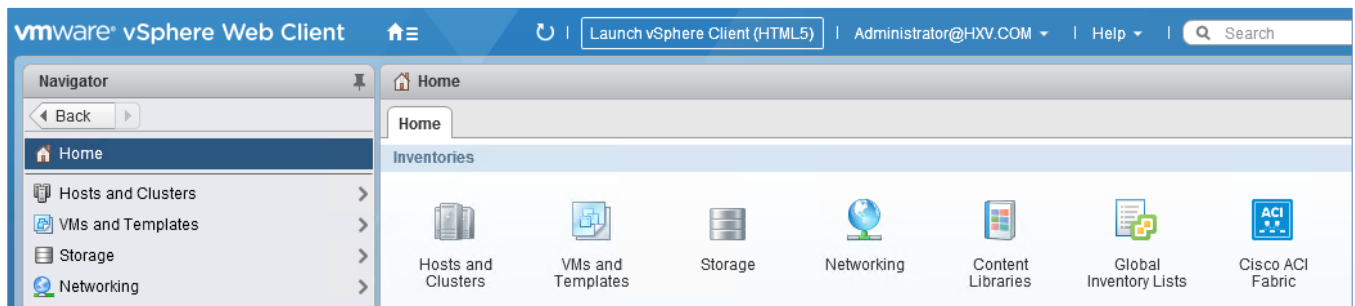
Verify the ACI Plug-in Installation from VMware vCenter

To verify from vCenter that the installation was successful, follow these steps:

1. Use a browser to navigate to the VMware vCenter server managing the HyperFlex Application cluster. Select the vSphere Web Client of your choice. Log in using an Administrator account. Verify that you see the **Cisco ACI Fabric** icon as shown in the figure below.



If you are already logged in, you may need to disconnect and re-login back to see the icon.



2. If you do not see the icon, navigate to the Managed Object Browser for the VMware vCenter (**Error! Hyperlink reference not valid.**) and check the status of the plug-in registration. If it shows a successful registration but you still do not see the icon, a reboot of the vCenter may be necessary.

Home

Managed Object Type: **ManagedObjectReference:ExtensionManager**
Managed Object ID: **ExtensionManager**

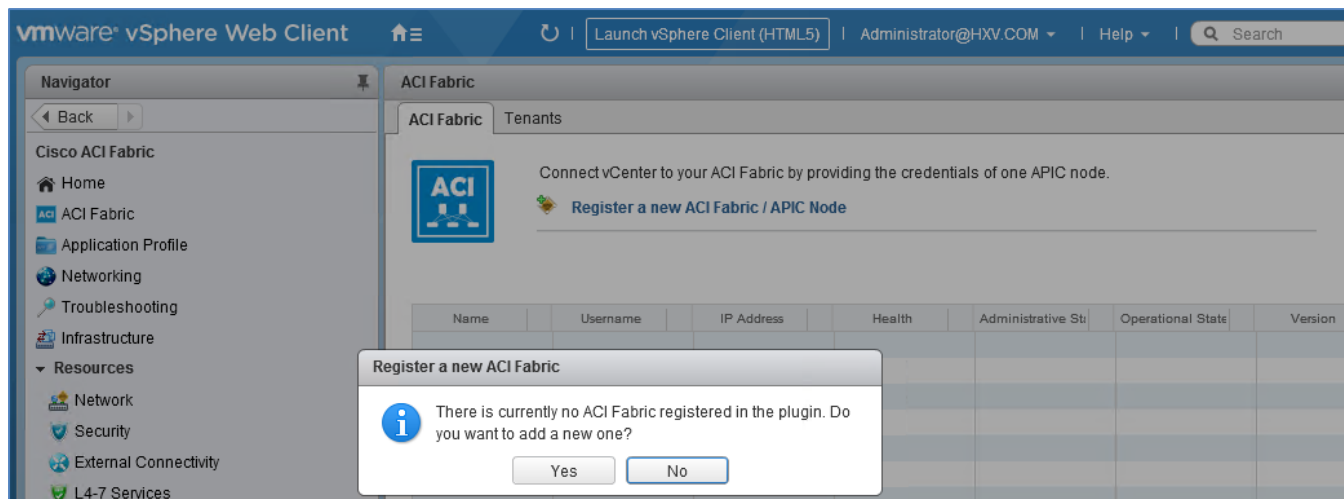
Properties

NAME	TYPE	VALUE	
extensionList	Extension[]	extensionList["com.vmware.vim.sms"]	Extension
		extensionList["com.vmware.vim.vsm"]	Extension
		extensionList["VirtualCenter"]	Extension
		extensionList["com.vmware.vim.stats.report"]	Extension
		extensionList["com.vmware.vim.sps"]	Extension
		extensionList["com.vmware.vim.vcha"]	Extension
		extensionList["hostdiag"]	Extension
		extensionList["com.vmware.vim.ls"]	Extension
		extensionList["com.vmware.vcenter.iso"]	Extension
		extensionList["com.vmware.cl"]	Extension
		extensionList["com.vmware.ovf"]	Extension
		extensionList["com.vmware.vim.eam"]	Extension
		extensionList["com.vmware.rbd"]	Extension
		extensionList["com.vmware.vcIntegrity"]	Extension
		extensionList["com.vmware.vmcam"]	Extension
		extensionList["com.vmware.vsan.health"]	Extension
		extensionList["com.vmware.vrops.install"]	Extension
		extensionList["com.springpath.sysmgmt"]	Extension
		extensionList["com.springpath.sysmgmt.domain-c166"]	Extension
		extensionList["com.cisco.aci.avs.tasks"]	Extension
		extensionList["com.cisco.aciPlugin"]	Extension
(less...)			

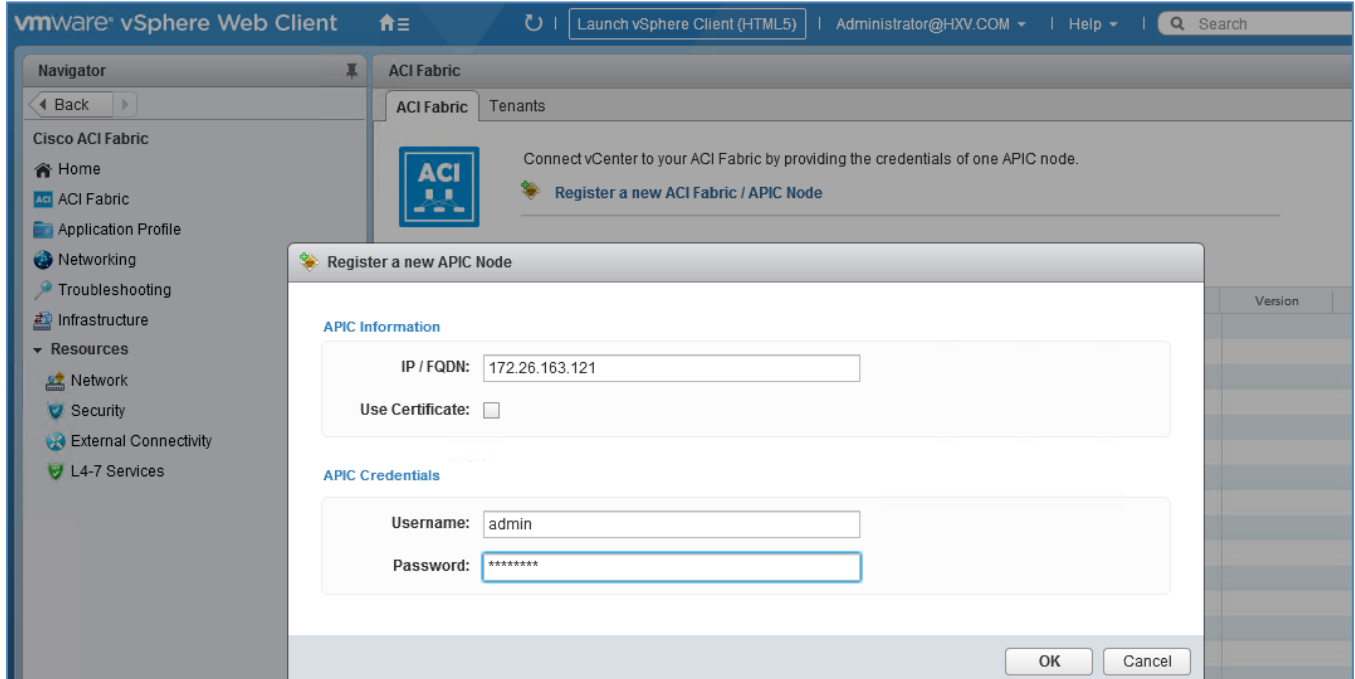
Connect ACI Plug-in to ACI Fabric

To connect the ACI plug-in to the ACI Fabric, follow these steps:

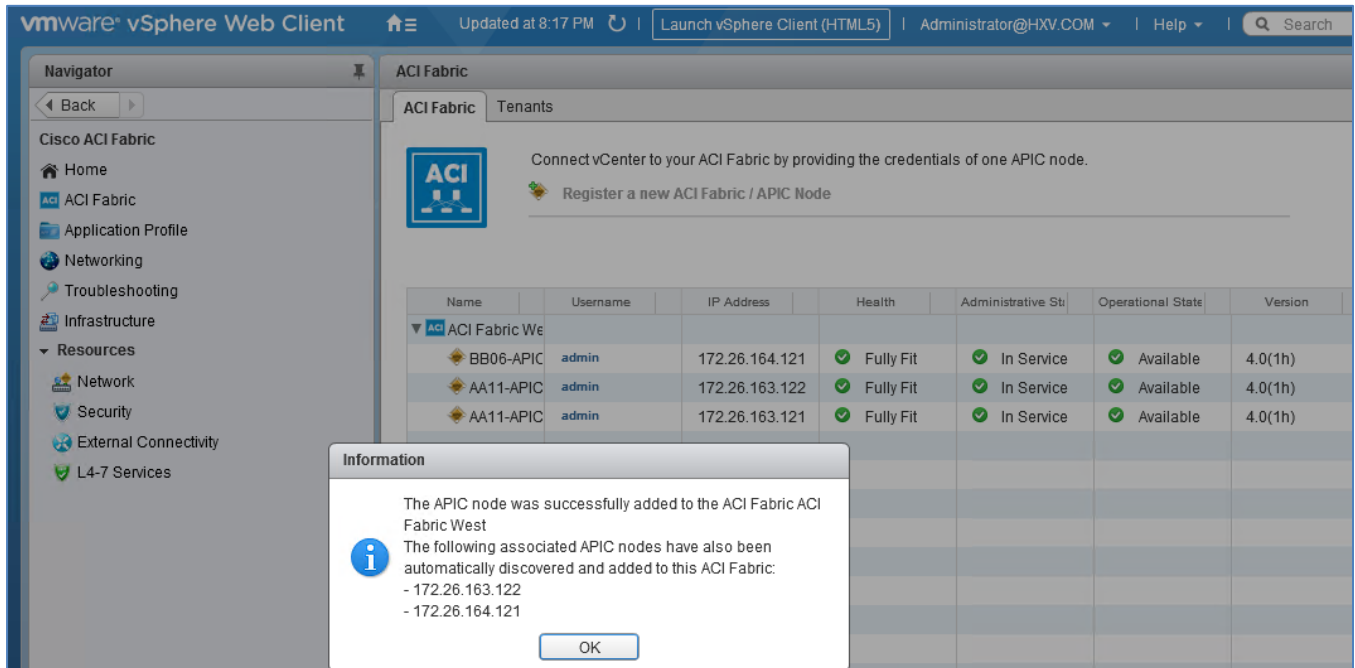
1. Use a browser to navigate to the VMware vCenter server managing the HyperFlex Application cluster. Click the vSphere Web Client of your choice. Log in using an **Administrator** account.
2. From the **Home** menu, click the **Cisco ACI Fabric** icon.
3. From the Getting Started tab, click Connect vSphere to your ACI Fabric.
4. In the **Register a new ACI Fabric** pop-up window, click **Yes** to register a new ACI fabric.



- In the **Register a new APIC Node** pop-up window, specify the IP or hostname of an APIC node. Deselect **Use Certificate**. For the APIC Credentials, specify the **Username** and **Password**.



- Click **OK** to complete the configuration and close the window.
- If the registration was successful, you should see something similar to the following.

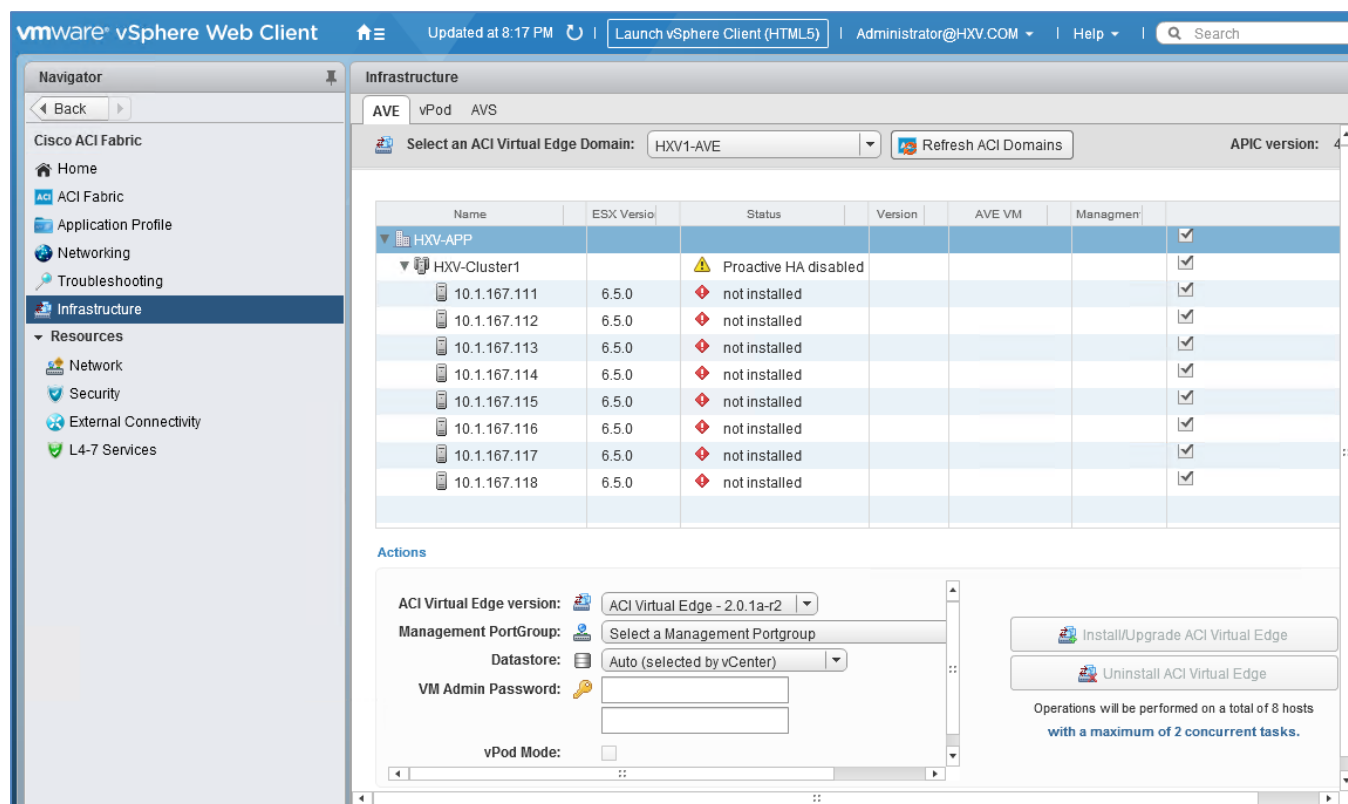


- Click **OK** to close the pop-up window. You should now see the APICs managing the ACI fabric.

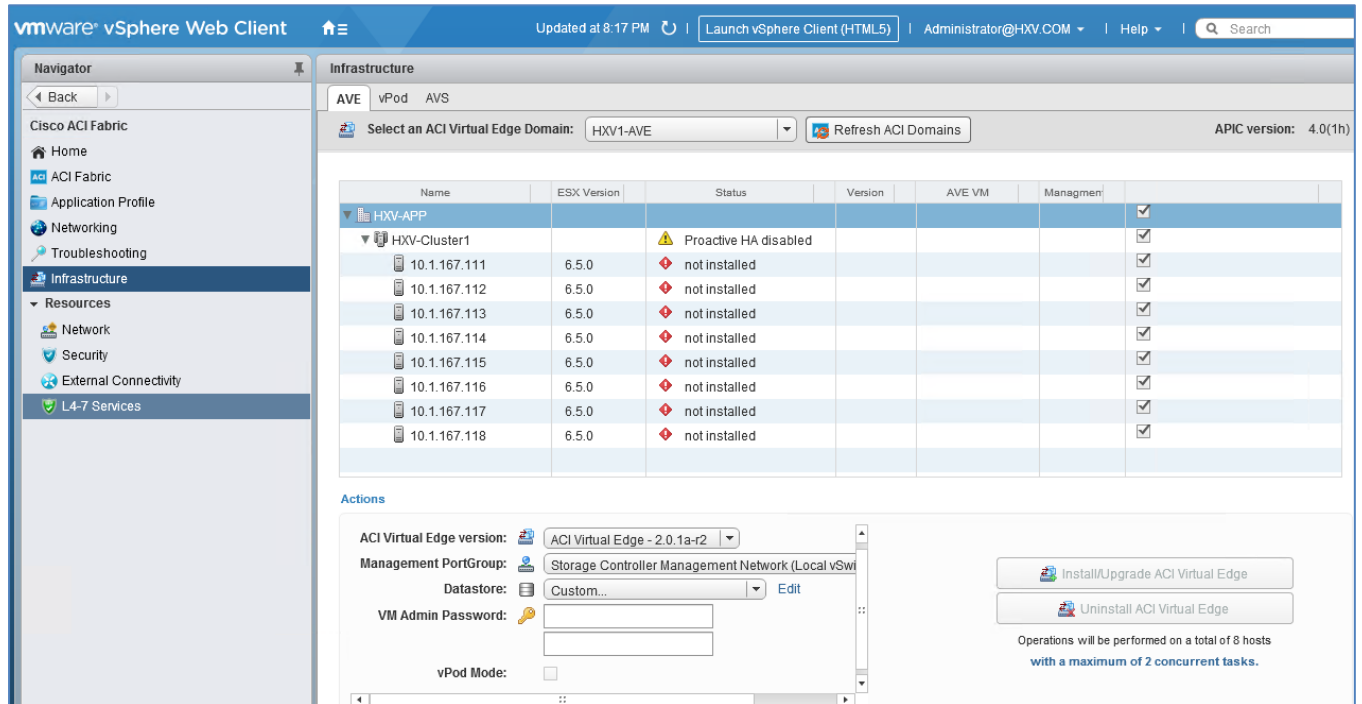
Deploy Cisco AVE Virtual Machine on the ESXi Hosts Using the Cisco ACI Plug-In

To deploy Cisco AVE virtual machine on HyperFlex ESXi hosts, follow these steps:

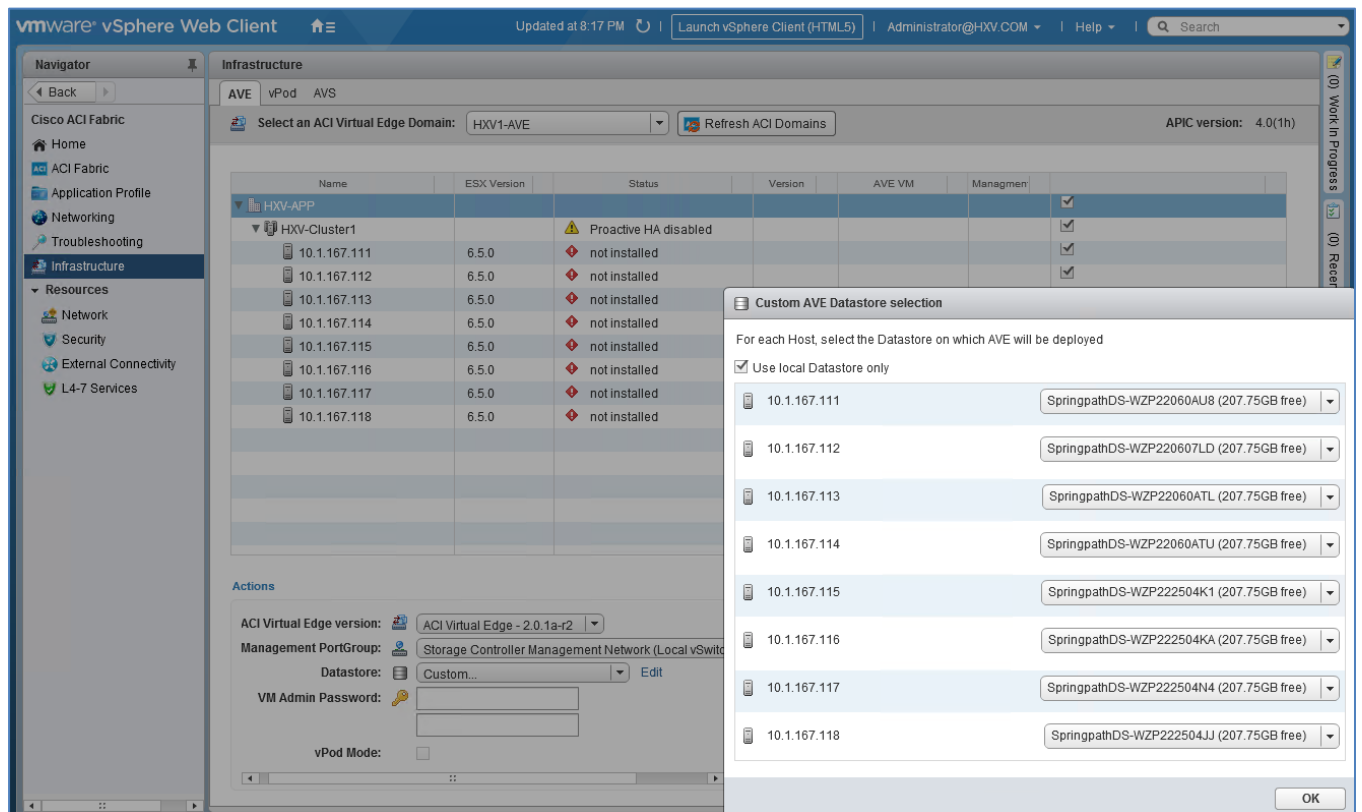
1. Use a browser to navigate to the VMware vCenter Server managing the HyperFlex Application cluster. Select the vSphere web client of your choice and log in using an **Administrator** account.
2. From the **Home** screen, click the **Cisco ACI Fabric** icon in the **Inventories** section.
3. From the left navigation pane, select **Infrastructure**.
4. From the right window pane, select **AVE** tab. Log in with VMware vCenter password.
5. Expand the **datacenter** and **cluster** to select the hosts where Cisco AVE should be deployed. Use the boxes to the right of each host.



6. Scroll down to the bottom of the screen. For **ACI Virtual Edge version**, choose the version to use from the drop-down list. For the **Management PortGroup**, choose the management port group from the drop-down list.



- For the **Datastore**, choose **Custom** from the drop-down list, click **Edit**.
- In the **Custom AVE Datastore selection** pop-up window, select the **Use local datastore only** checkbox, and specify local data store for each Cisco ACI Virtual Edge.

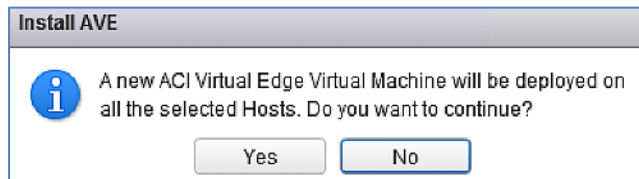




Cisco ACI Virtual Edge installation is supported only on local data stores in the current release.

9. Click **OK**.
10. For the **VM Admin Password** fields, enter a password for the Cisco ACI Virtual Edge virtual machines.
11. Click the Install/Upgrade ACI Virtual Edge button.

12. In the **Install AVE** pop-up window, click **Yes**.



When the install is complete, you should see something similar to the following:

Name	ESX Version	Status	Version	AVE VM	Management IP
HXV-APP					
HXV-Cluster1					
10.1.167.111	6.5.0	online	2.0.1a	cisco-ave_10.1.167.111_HXV1-AVE	10.1.167.27
10.1.167.112	6.5.0	online	2.0.1a	cisco-ave_10.1.167.112_HXV1-AVE	10.1.167.22
10.1.167.113	6.5.0	online	2.0.1a	cisco-ave_10.1.167.113_HXV1-AVE	10.1.167.26
10.1.167.114	6.5.0	online	2.0.1a	cisco-ave_10.1.167.114_HXV1-AVE	10.1.167.23
10.1.167.115	6.5.0	online	2.0.1a	cisco-ave_10.1.167.115_HXV1-AVE	10.1.167.25
10.1.167.116	6.5.0	online	2.0.1a	cisco-ave_10.1.167.116_HXV1-AVE	10.1.167.21
10.1.167.117	6.5.0	online	2.0.1a	cisco-ave_10.1.167.117_HXV1-AVE	10.1.167.28
10.1.167.118	6.5.0	online	2.0.1a	cisco-ave_10.1.167.118_HXV1-AVE	10.1.167.24

Now you are ready to deploy Virtual Machines on the HyperFlex cluster using Cisco AVE virtual leaf switches.

Solution Deployment – Onboarding Multi-Tier Applications

This section provides the detailed procedures for onboarding multi-tier applications onto the Application cluster. Application virtual machines can be deployed in either data center in this active-active data center solution.

Deployment Overview

The high-level steps for deploying multi-tier applications on a Cisco HyperFlex cluster connected to a Cisco ACI Multi-Pod fabric are as follows:

- Define ACI Constructs for the new Application. This includes defining an Application Tenant, VRF, Bridge Domain and an Application Profile.
- Define End Point Groups. A three-tier application could be deployed using three EPGs, for example, Web, App and Database EPGs.
- Enable contracts to allow users to access the Application and for communication between different tiers of the application. Also, enable contracts to access the shared L3out for connectivity to outside networks and services.
- Deploy application virtual machines on the Application HyperFlex cluster.
- Add virtual machines to the port-group corresponding to the EPG.

In this section, a sample two-tier (Web, App) application is deployed in a dedicated tenant **HXV-App-A**. The Web and App Tier will be mapped to corresponding EPGs in the ACI fabric.

Prerequisites

- Integration with Virtual Machine Manager or VMware vCenter for virtual networking should be in place before onboarding applications as outlined in this section. As a part of this integration, a VLAN pool should also be pre-defined for future use. VLANs from the VLAN pool will be assigned to Application EPGs such that when an EPG is defined in ACI, a corresponding port-group is created in the VMM domain. The application virtual machines, when deployed, can now be added to the correct port-group to enable connectivity through the ACI fabric.
- When a VLAN Pool is defined for VMM integration, the VLANs need to be created in the UCS domain hosting the VMM domain. For the Application cluster in this design, the VLANs need to be enabled on the UCS domains that connects the HyperFlex stretched cluster in the different Pods.



If the VLANs (`hxv-vm-network`) were specified during cluster install or as input to the **post-install** script, then they are already created and trunked on the Cisco UCS Fabric Interconnect uplinks, and on the virtual NICs (`vNIC vm-network-a`, `vNIC vm-network-b`) of each HyperFlex node.

Configure ACI constructs for Application

In this section, the ACI constructs (Tenant, VRF, Bridge Domain and Application Profile) for the new Application are setup.

Create Tenant and VRF for Application

To create Tenant and VRF for the application, follow these steps:

1. Use a browser to navigate to APIC's Web GUI. Log in with the **admin** account.

2. From the top menu, select **Tenants > Add Tenant**.
3. In the **Create Tenant** pop-up window, specify a **Name** (for example, HXV-App-A).
4. For the **VRF Name**, enter a name for the only VRF in this Tenant (for example, HXV-App-A_VRF)
5. Leave the checkbox for Take me to this tenant when I click finish checked.
6. Click **Submit** to complete the configuration.

Configure Bridge Domains

At least one bridge domain will need to be created. Insertion and configuration of this firewall is not covered in this document. To create an internal versus an external bridge domain to allow an optional insertion of a firewall between EPGs connecting from the differing bridge domains, follow these steps:

1. Use a browser to navigate to APIC's Web GUI. Login with the **admin** account.
2. From the top menu, select **Tenants > HXV-App-A**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on HXV-App-A .
3. In the left navigation pane, navigate to Tenant **HXV-App-A > Networking > Bridge Domains**
4. Right-click Bridge Domains and select Create Bridge Domain.
5. In the **Create Bridge Domain** pop-up window, for **Name**, specify a name (HXV-App-A-Ext_BD) and for **VRF**, select the previously created VRF (HXV-App-A_VRF).
6. Click **Next** twice and then **Finish** to complete adding the Bridge Domain.
7. Repeat steps 1-6 to add another Bridge Domain HXV-App-A-Int_BD under the same VRF.

Configure Application Profile

To configure the application profile, follow these steps:

1. Use a browser to navigate to APIC's Web GUI. Login with the **admin** account.
2. From the top menu, select **Tenants > HXV-App-A**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on HXV-App-A .
3. In the left navigation pane, navigate to **Tenant HXV-App-A > Application Profiles**.
4. Right-click Application Profiles and select Create Application Profile.
5. In the **Create Application Profile** pop-up window, for **Name**, specify a name (HXV-App-A_AP) .
6. Click **Submit**.

Configure End Point Groups

EPG for Web

To configure end point groups for EPG for web, follow these steps:

1. Use a browser to navigate to APIC's Web GUI. Login with the **admin** account.
2. From the top menu, select **Tenants > HXV-App-A**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on HXV-App-A .
3. In the left navigation pane, navigate to **Tenant HXV-App-A > Application Profiles > HXV-App-A_AP**.

4. Right-click and select **Create Application EPG**.
5. In the **Create Application EPG** pop-up window, for **Name**, specify a name (HXV-A-Web_EPG).
6. For the **Bridge Domain**, select the previously created external Bridge Domain (HXV-App-A-Ext_BD) from the drop-down list.
7. Check the Associate to VM Domain Profiles checkbox.

The screenshot shows the Cisco APIC interface with the 'Create Application EPG' dialog box open. The dialog is titled 'Create Application EPG' and has a 'STEP 1 > Identity' header. The 'Specify the EPG Identity' section contains the following fields and options:

- Name:** HXV-A-Web_EPG
- Alias:** (empty)
- Description:** optional
- Tags:** (empty dropdown)
- Contract Exception Tag:** (empty)
- QoS class:** Unspecified
- Custom QoS:** select a value
- Data-Plane Policer:** select a value
- Intra EPG Isolation:** Enforced (selected), Unenforced
- Preferred Group Member:** Exclude (selected), Include
- Flood on Encapsulation:** Disabled (selected), Enabled
- Bridge Domain:** HXV-App-A-Ext_BD
- Monitoring Policy:** select a value
- FHS Trust Control Policy:** select a value
- Shutdown EPG:** ☐
- Associate to VM Domain Profiles:** ☒
- Statically Link with Leaves/Paths:** ☐
- EPG Contract Master:** (empty)

At the bottom right, there are three buttons: 'Previous', 'Cancel', and 'Next'.

8. Click **Next**.
9. Click **[+]** to Associate VM Domain Profiles.
10. For the **Domain Profile**, select VMware/HXV1-AVE from the drop-down list.
11. Change the Deployment Immediacy and Resolution Immediacy to Immediate.

12. Click **Update**.
13. Click **Finish** to complete the configuration.
14. In the left navigation pane, navigate to newly created EPG (HXV-A-Web_EPG), right-click and select **Create EPG Subnet**.



Cisco recommends configuring subnets at the Bridge domain level when possible. Configuring subnets at the EPG level should be used only in certain situations.

15. For the **Default Gateway IP**, enter a **gateway IP address** and **mask** (for example, 172.19.201.254/24).
16. Since the Web VM Subnet is advertised to networks outside ACI and to **App EPG**, select checkboxes for **Advertised Externally** and **Shared between the VRFs**.

The screenshot shows the Cisco APIC Web GUI. The top navigation bar includes tabs for System, Tenants, Fabric, Virtual Networking, L4-L7 Services, Admin, Operations, and Apps. The 'Tenants' tab is active, and the 'HXV-App-A' tenant is selected. The left navigation pane shows the hierarchy: Tenant HXV- > Application Profiles > HXV-App-A_AP > Application EPGs > HXV-A-Web_EPG. The main content area displays the 'EPG - HXV-A-Web_EPG' configuration page with tabs for Summary, Policy, and Operational. A 'Create EPG Subnet' dialog box is open, prompting the user to 'Specify the Subnet Identity'. The dialog contains the following fields and options:

- Default Gateway IP:** 172.19.201.254/24 (address/mask)
- Treat as virtual IP address:** ☐
- Scope:**
 - ☐ Private to VRF
 - ☒ Advertised Externally
 - ☒ Shared between VRFs
- Description:** optional
- Subnet Control:**
 - ☐ No Default SVI Gateway
 - ☐ Querier IP
- ND RA Prefix policy:** select a value
- Type Behind Subnet:** None (selected), EP Reachability, Anycast MAC

At the bottom right of the dialog are 'Cancel' and 'Submit' buttons. The 'Submit' button is highlighted in blue.

17. Click **Submit**.

EPG for App

To create EPG for App, follow these steps:

1. Use a browser to navigate to APIC's Web GUI. Login with the **admin** account.
2. From the top menu, select **Tenants > HXV-App-A..** If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-App-A**.
3. In the left navigation pane, navigate to **Tenant HXV-App-A > Application Profiles > HXV-App-A_AP**.
4. Right-click and select **Create Application EPG**.
5. Name the EPG **HXV-A-App_EPG**.

6. Leave Intra EPG Isolation as Unenforced.
7. For the **Bridge Domain**, select HXV-App-A-Int_BD from the drop-down list.
8. Check the box next to Associate to VM Domain Profiles.

Create Application EPG

STEP 1 > Identity

Specify the EPG Identity

Name: HXV-A-App_EPG

Alias:

Description: optional

Tags:
 enter tags separated by comma

Contract Exception Tag:

QoS class: Unspecified

Custom QoS: select a value

Data-Plane Policer: select a value

Intra EPG Isolation: Enforced Unenforced

Preferred Group Member: Exclude Include

Flood on Encapsulation: Disabled Enabled

Bridge Domain: HXV-App-A-Int_BD

Monitoring Policy: select a value

FHS Trust Control Policy: select a value

Shutdown EPG: ☐

Associate to VM Domain Profiles: ☒

Statically Link with Leaves/Paths: ☐

EPG Contract Master:
 Application EPGs

Previous Cancel Next

9. Click **Next**.
10. Click [+] to Associate VM Domain Profiles.
11. For the **Domain Profile**, select VMware/HXV1-AVE from the drop-down list.
12. Change the Deployment Immediacy and Resolution Immediacy to Immediate.
13. Click **Update**.

The screenshot shows the Cisco APIC interface for configuring an Application Profile. The left navigation pane is expanded to 'Tenant HXV-' and 'HXV-App-A'. The main content area is titled 'Application Profile - HXV-App-A_AP' and shows 'STEP 2 > Domains'. The sub-header is 'Specify the VM Domain'. Below this is a table of 'Domain Profiles' with the following columns: Domain Profile, Deployment Immediacy, Resolution Immediacy, Delimiter, Encap Mode, Port Encap (or Secondary VLAN for Micro-Seg), Allow Micro-Segmentation, and Switching Mode. The table contains one row: 'HXV1-AVE', 'Immediate', 'Immediate', empty, 'Auto', empty, 'False', and 'AVE'. At the bottom right are 'Previous', 'Cancel', and 'Finish' buttons.

14. Click **Finish** to complete the configuration.
15. In the left navigation pane, select the newly created EPG (HXV-A-App_EPG), right-click and select **Create EPG Subnet**.



Cisco recommends configuring subnets at the Bridge domain level when possible. Configuring subnets at the EPG level should be used only in certain situations.

16. For the **Default Gateway IP**, enter a **gateway IP address** and **mask** (for example, 172.19.202.254/24).
17. Since the App virtual machines only need to communicate with Web VMs EPG, leave the checkbox for **Private to VRF** selected.

The screenshot shows the Cisco APIC interface with the 'Tenants' tab selected. The left sidebar shows the navigation tree for 'Tenant HXV-A', with 'HXV-A-App_EPG' selected. The main panel displays the 'EPG - HXV-A-App_EPG' configuration page. A 'Create EPG Subnet' dialog box is open, prompting the user to 'Specify the Subnet Identity'. The dialog contains the following fields and options:

- Default Gateway IP:** 172.19.202.254/24 (address/mask)
- Treat as virtual IP address:** ☐
- Scope:**
 - ☒ Private to VRF
 - ☐ Advertised Externally
 - ☐ Shared between VRFs
- Description:** optional
- Subnet Control:**
 - ☐ No Default SVI Gateway
 - ☐ Querier IP
- ND RA Prefix policy:** select a value
- Type Behind Subnet:** None (selected), EP Reachability, Anycast MAC

The 'Submit' button is highlighted in blue.

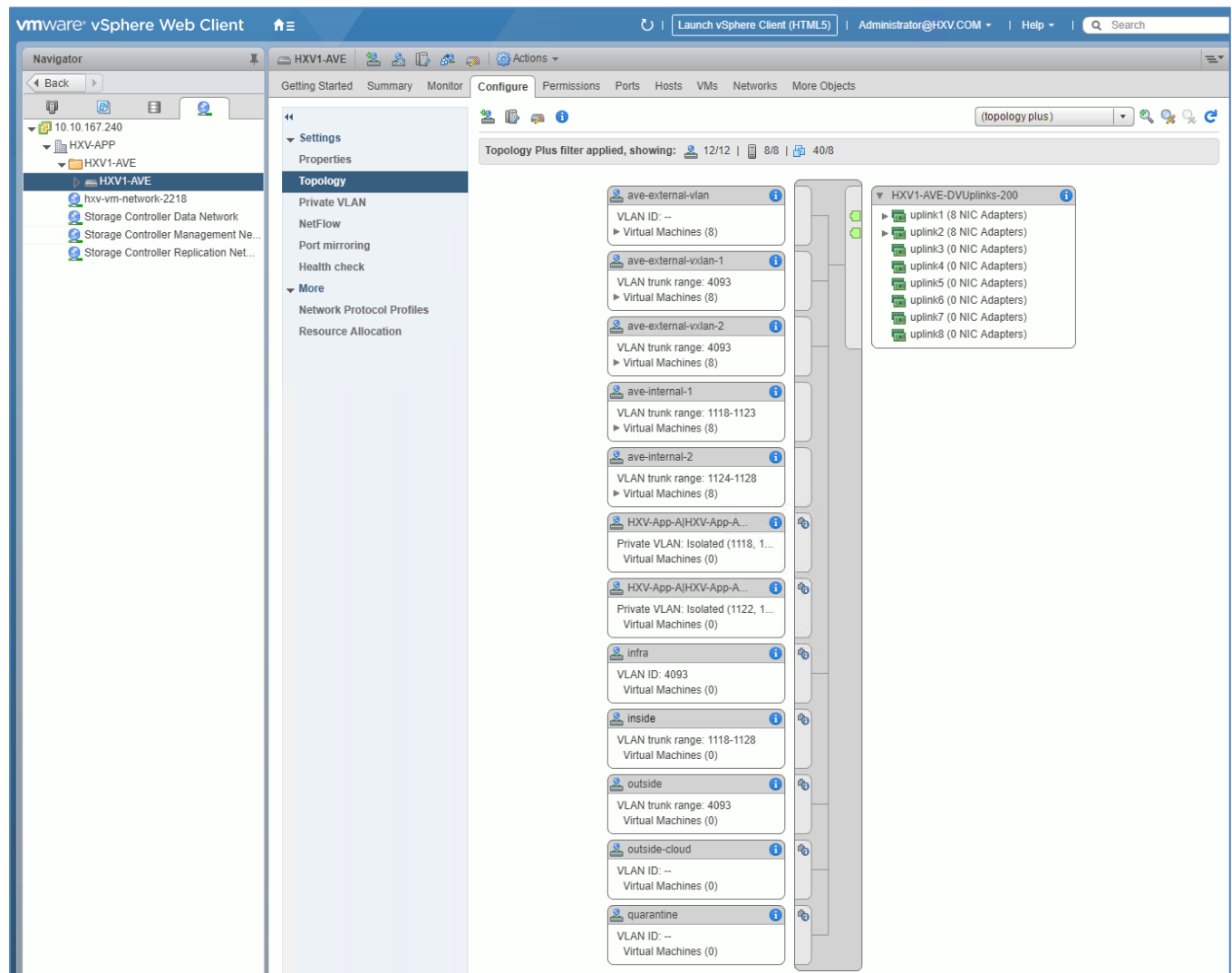
18. Click **Submit**.

Verify Virtual Networking for the Application EPGs

When the two Application EPGs are provisioned in the ACI fabric and associated with a VMM domain (in this case, HXV1 – AVE), you should now see two port-groups corresponding to the EPGs in the Cisco AVE switch. To verify that the port-groups have been created in the VMM domain (VMware vCenter), follow these steps:

1. Use a browser to navigate to the VMware vCenter server managing the HyperFlex Application cluster. Click the vSphere Web Client of your choice. Log in using an **Administrator** account
2. Navigate to the **Home** screen, select **Networking** in the **Inventories** section.

3. In the left navigation pane, expand the **datacenter** folder and **distributed virtual switch** for the ACI VMM domain associated with the EPGs. The distributed virtual switch would've been created by the Cisco APIC when the VMM domain was first created.
4. In the right window pane, navigate to **Configure > Topology**. The port-groups associated with the two EPGs should've been automatically created as shown below. When the VMM domain is associated with a Cisco AVE, two VLANs from the VLAN pool are used for each EPG to create a private VLAN.



5. The application virtual machines can now be deployed and added to these port-groups. However, for connectivity outside the EPG, the necessary contracts need to be provided and consumed between the different EPGs as outlined in the next section.

Configure Contracts

App-Tier to Web-Tier Contract

To enable communication between Web and App tiers of the application, follow these steps:

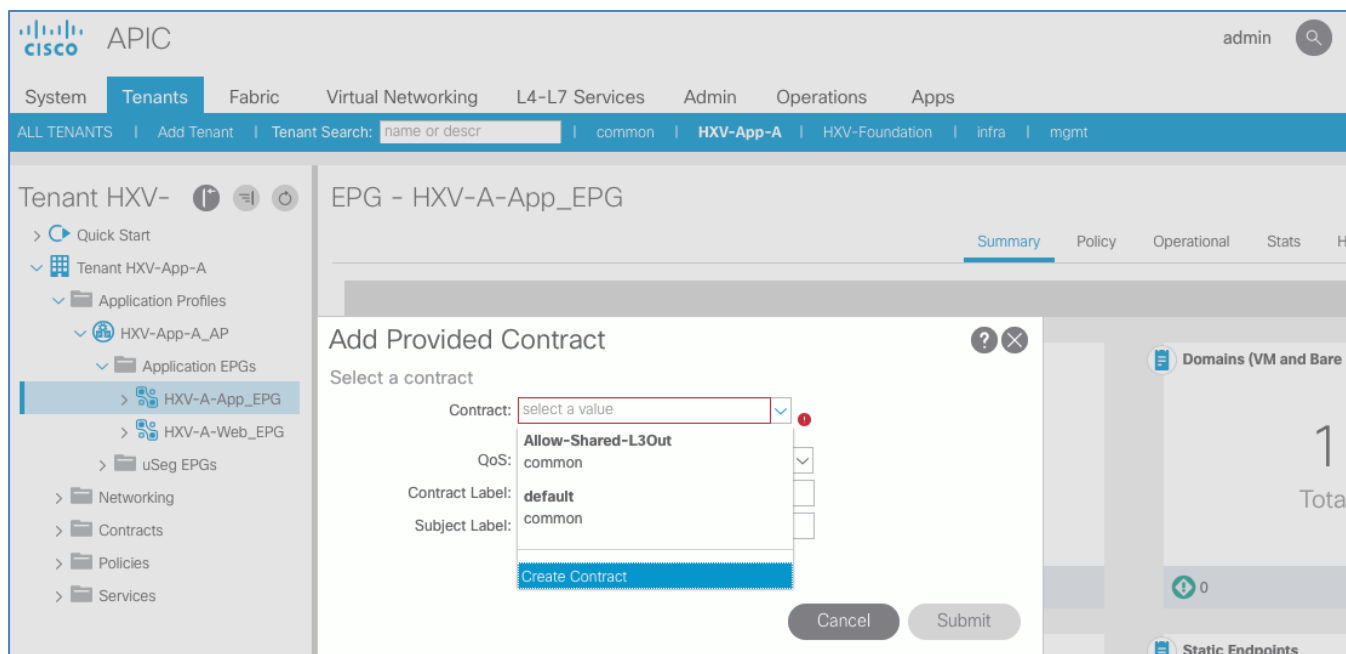


You can use more restrictive contracts to replace the `Allow-Shared-L3Out` contract defined in this example.

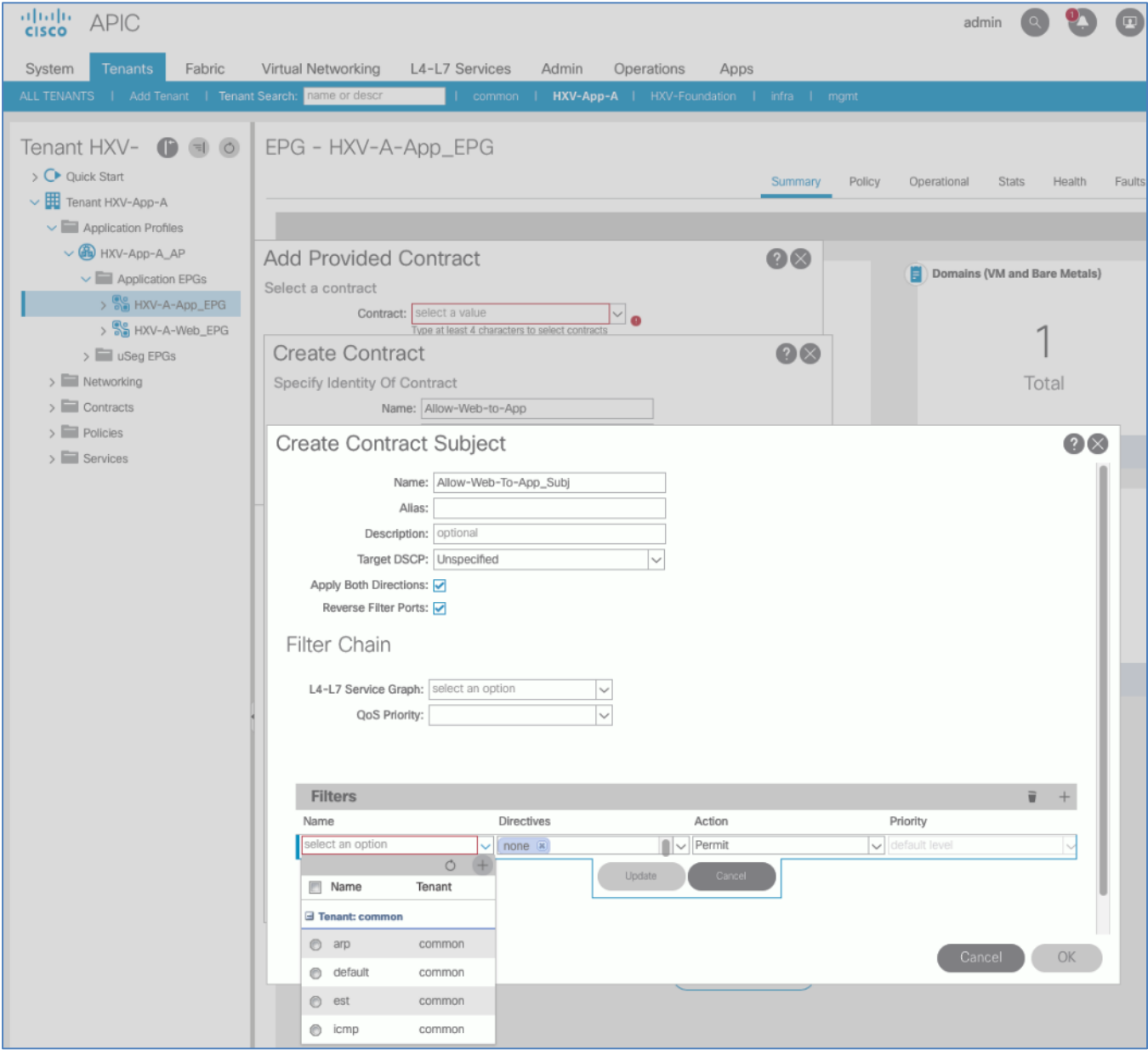
Provided Contract in EPG App-A

To add a Provided Contract in EPG App-A, follow these steps:

1. Use a browser to navigate to APIC's Web GUI. Login with the **admin** account.
2. From the top menu, select **Tenants > HXV-App-A**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on **HXV-App-A**.
3. From the left navigation pane, expand and select **Tenant HXV-App-A > Application Profiles > HXV-App-A_AP > Application EPGs > HXV-A-App_EPG**.
4. Right-click **HXV-A-App_EPG** and select **Add Provided Contract**.
5. In the **Add Provided Contract** pop-up window, for **Contract**, select **Create Contract** from end of the drop-down list.



6. In the **Create Contract** pop-up window, for **Name**, specify a name for the contract (**Allow-Web-to-App**).
7. For **Scope**, select **Tenant** from the drop-down list.
8. For **Subjects**, click **[+]** to add a Contract Subject.
9. In the **Create Contract Subject** pop-up window, specify a **Name** (**Allow-Web-to-App_Subj**) for the subject.
10. For **Filters**, click **[+]** to add a Contract filter.



- 11. Click **[+]** to add a new filter.
- 12. In the **Create Filter** pop-up window, specify a **Name** for the filter: Allow-Web-A-All.

The screenshot shows the Cisco APIC interface with the 'Create Filter' dialog box open. The dialog is titled 'Create Filter' and has a 'Specify the Filter Identity' section. The 'Name' field is filled with 'Allow-Web-A-All'. The 'Alias' field is empty. The 'Description' field is labeled 'optional'. Below the description is a 'Tags' field with a dropdown menu and a note 'enter tags separated by comma'. The 'Entries' section is a table with the following columns: Name, Alias, EtherType, ARP Flag, IP Protocol, Match Only Fragments, Stateful, Source Port / Range (From, To), Destination Port / Range (From, To), and TCP Session Rules. A '+' icon is in the top right of the table. At the bottom right of the dialog are 'Cancel' and 'Submit' buttons.

13. For **Entries**, click **[+]** to add an Entry.
14. Enter a **Name** for the Entry, for example: Allow-All.
15. For the **EtherType**, select **IP** from the drop-down list.

The screenshot shows the Cisco APIC interface with the 'Tenants' tab selected. The left sidebar shows the navigation tree with 'HXV-A-App_EPG' selected. The main content area displays the configuration for 'EPG - HXV-A-App_EPG'. Overlaid on this is the 'Create Filter' pop-up window.

Create Filter

Specify the Filter Identity

Name:

Alias:

Description: optional

Tags:

enter tags separated by comma

Entries:

Name	Alias	EtherType	ARP Flag	IP Protocol	Match Only	Stateful	Source Port / Range	Destination Port / Range	TCP Session Rules	
							From	To	From	To
Allow-AI		IP		Unspecified			Unspecified	Unspecified	Unspecified	Unspecified

Update Cancel

Cancel Submit

16. Click **Update** and **Submit** to finish creating the filter and close the **Create Filter** pop-up window.
17. Click **Update** and **OK** to finish creating the Contract Subject and close the **Create Contract Subject** pop-up window.

APIC

admin

System

Tenants

Fabric

Virtual Networking

L4-L7 Services

Admin

Operations

Apps

ALL TENANTS

Add Tenant

Tenant Search:

common

HXV-App-A

HXV-Foundation

infra

mgmt

Tenant HXV-

Quick Start

Tenant HXV-App-A

Application Profiles

HXV-App-A-AP

Application EPGs

HXV-A-App_EPG

HXV-A-Web_EPG

uSeg EPGs

Networking

Contracts

Policies

Services

EPG - HXV-A-App_EPG

Summary

Policy

Operational

Stats

Health

Fault

Add Provided Contract

Select a contract

Contract:

Type at least 4 characters to select contracts

Create Contract

Specify Identity Of Contract

Name:

Create Contract Subject

Name:

Alias:

Description:

Target DSCP:

Apply Both Directions: ☒

Reverse Filter Ports: ☒

Filter Chain

L4-L7 Service Graph:

QoS Priority:

Filters

Name	Directives	Action	Priority
HXV-App-A/Allow-Web-A-All	<input type="text" value="none"/>	<input type="text" value="Permit"/>	<input type="text" value="default level"/>

Update

Cancel

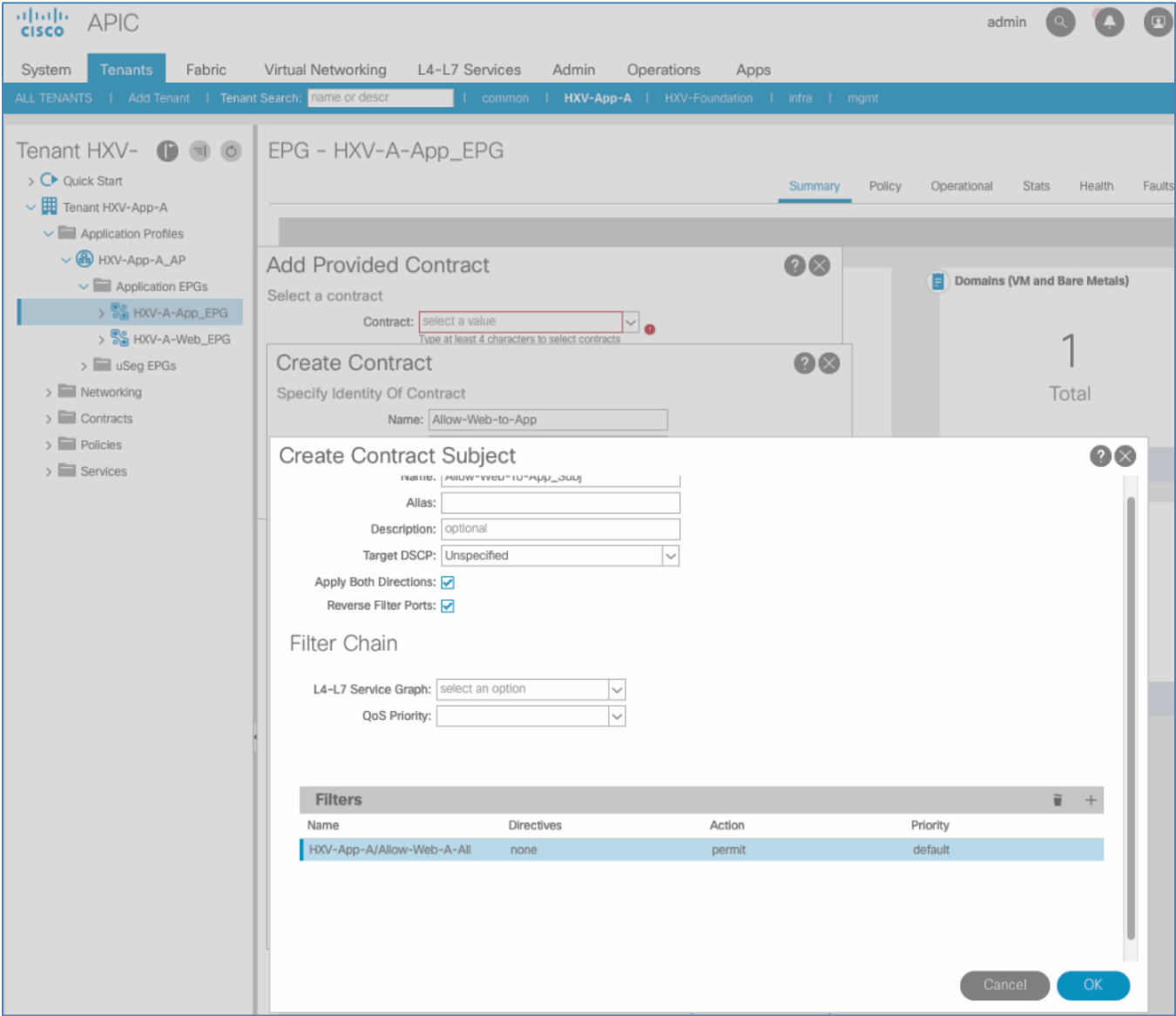
Cancel

OK

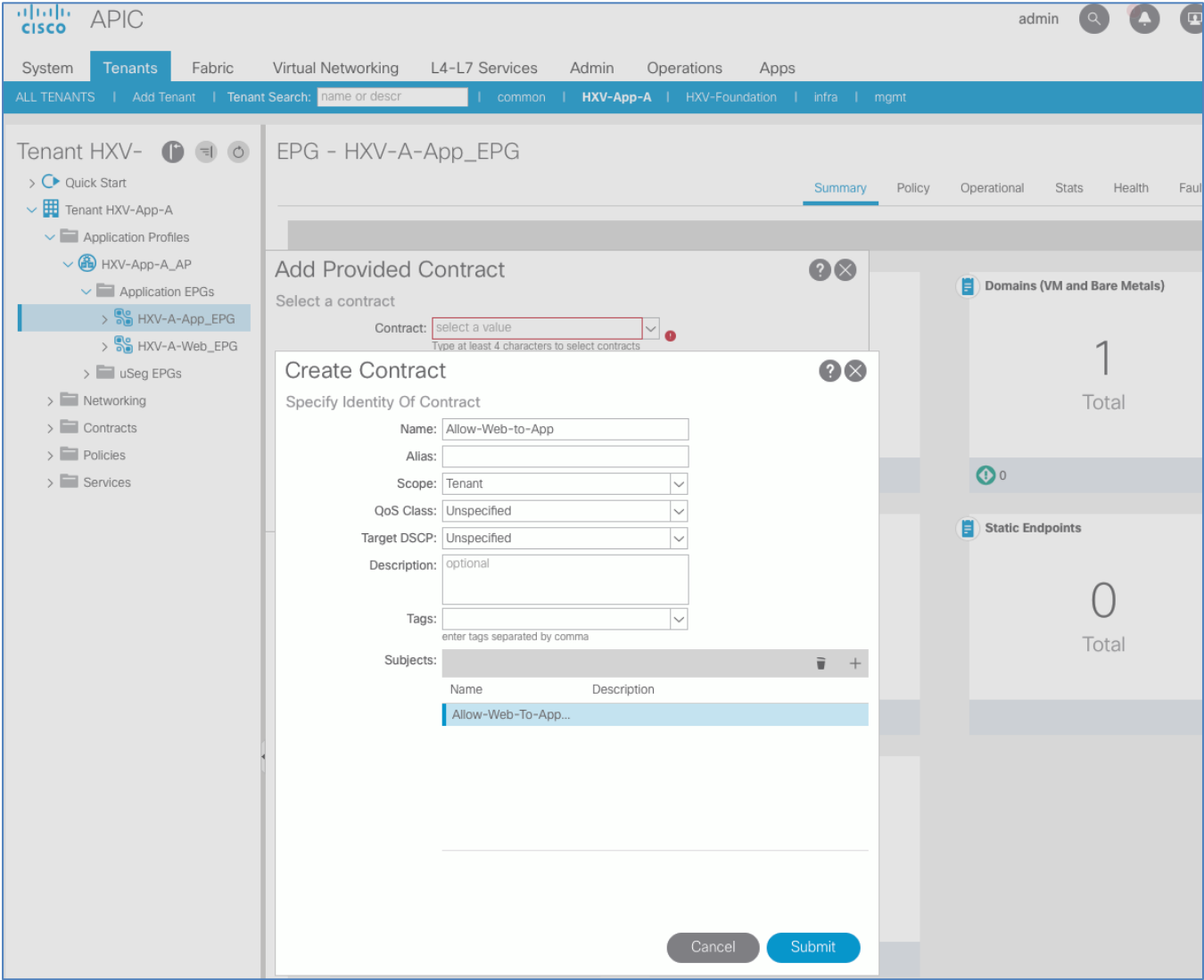
Domains (VM and Bare Metals)

1

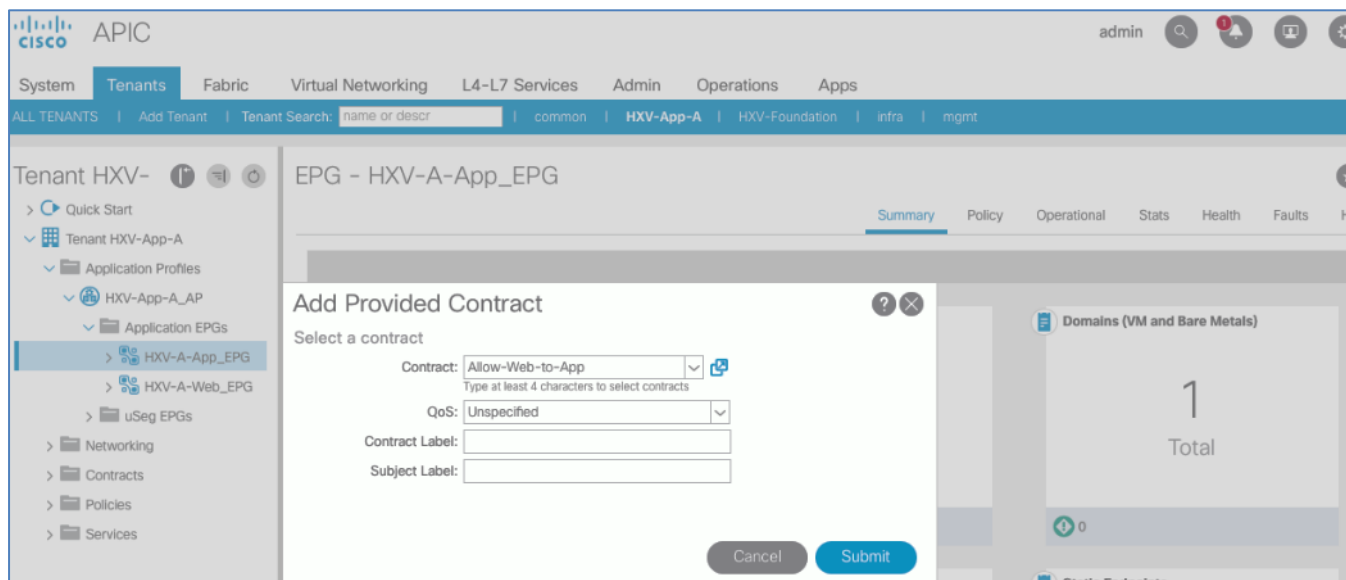
Total



18. Click **Submit** to complete creating the Contract and close the **Create Contract** pop-up window.



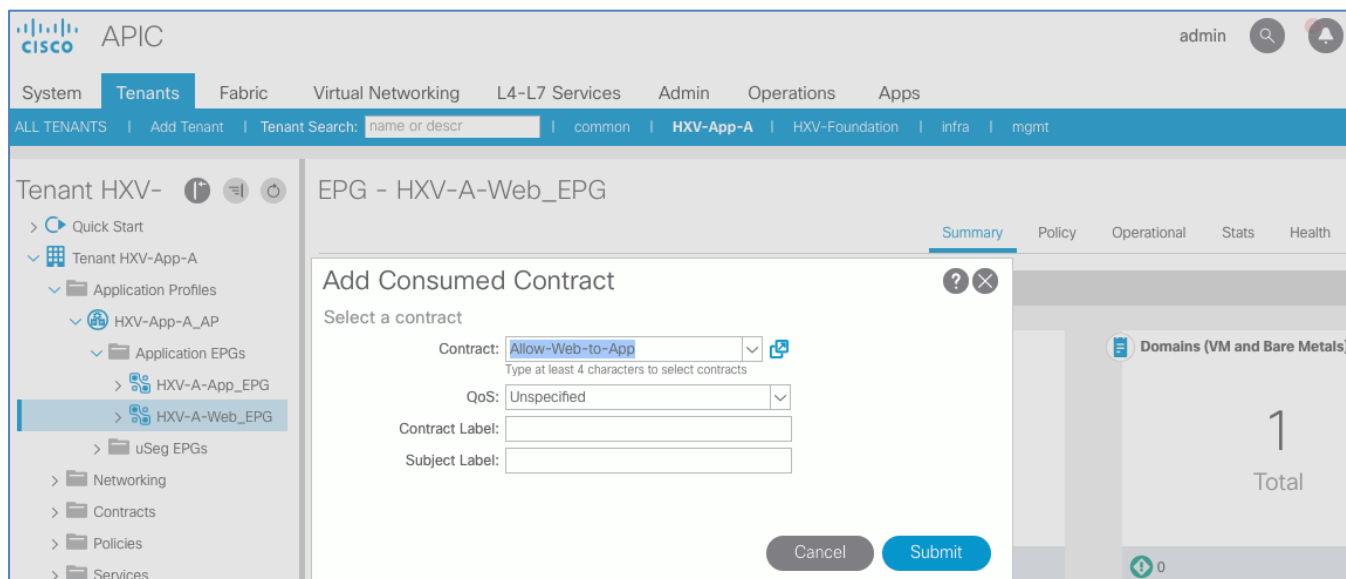
19. Click **Submit** to complete adding the Provided Contract and close the **Add Provided Contract** pop-up window.



Consume Contract in EPG Web-A

To add a Consume Contract in EPG Web-A, follow these steps:

1. Use a browser to navigate to APIC's Web GUI. Login using the **admin** account.
2. From the top menu, select **Tenants** > HXV-App-A. If you do not see this tenant in the top navigation menu, select **Tenants** > **ALL TENANTS** and double-click on HXV-App-A.
3. In the left navigation pane, expand and select **Tenant** HXV-App-A > **Application Profiles** > HXV-App-A_AP > **Application EPGs** > HXV-A-Web_EPG.
4. Right-click and select **Add Consumed Contract**.
5. In the **Add Consumed Contract** pop-up window, select the newly created contract (Allow-Web-to-App) from the drop-down list.

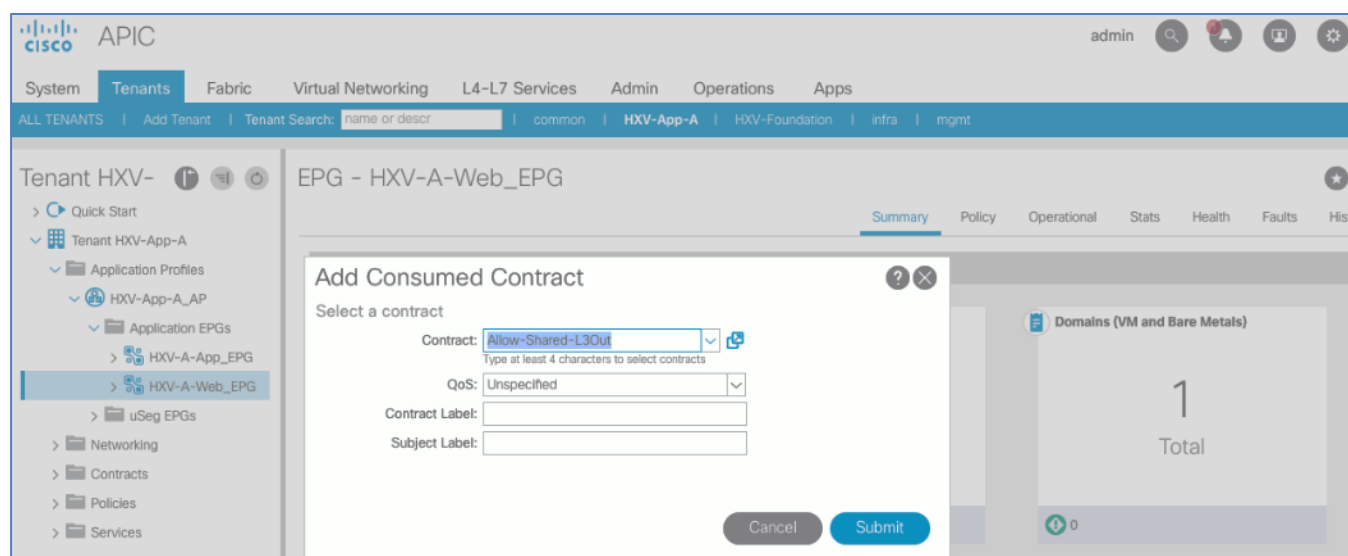


6. Click **Submit** to complete adding the Consumed Contract.

Web-Tier to Shared L3Out Contract

To enable App-A's Web VMs to communicate outside the Fabric, Shared L3 Out contract defined in the Common Tenant will be consumed by the Web EPG. To enable Web virtual machines to outside the fabric, follow these steps:

1. Use a browser to navigate to APIC's Web GUI. Login using the **admin** account.
2. From the top menu, select **Tenants > HXV-App-A**. If you do not see this tenant in the top navigation menu, select **Tenants > ALL TENANTS** and double-click on HXV-App-A.
3. In the left navigation pane, expand and select **Tenant HXV-App-A > Application Profiles > HXV-App-A_AP > Application EPGs > HXV-A-Web_EPG**.
4. Right-click and select **Add Consumed Contract**.
5. In the **Add Consumed Contract** pop-up window, select the shared L3Out contract (common/Allow-Shared-L3Out).



6. Click **Submit** to complete adding the Consumed Contract.

Solution Validation

This section provides a high-level summary of the validation done for this CVD.

Validated Hardware and Software

Table 77 lists the hardware and software versions used during the solution validation. The versions used have been certified within interoperability matrixes supported by Cisco and VMware.

Table 77 Hardware and Software Versions

Infrastructure Domain	Component		Software	Notes
Network (ACI MultiPod Fabric)	Pod 1	Pod 2		
	Cisco APIC M2 Server x 2 (APIC-SERVER-M2)	Cisco APIC M2 Server x 1 (APIC-SERVER-M2)	4.0.1h	3-node APIC Cluster
	Cisco Nexus 9364C x 2 (N9K-C9364C)	Cisco Nexus 9364C x 2 (N9K-C9364C)	aci-n9000-dk9.14.0.1h	ACI Spine Switches
	Cisco Nexus 93180YC-EX x 2 (N9K-C93180YC-EX)	Cisco Nexus 93180YC-EX x 2 (N9K-C93180YC-EX)	aci-n9000-dk9.14.0.1h	ACI Leaf Switches for HyperFlex and UCS Domains
	Cisco Nexus 9372PX x 2 (N9K-C9372PX)	Cisco Nexus 9372PX x 2 (N9K-C9372PX)	aci-n9000-dk9.14.0.1h	ACI Border Leaf Switches for Shared L3Out
	Cisco Nexus 93180YC-EX x 2 (N9K-C93180YC-EX)	Cisco Nexus 93180YC-EX x 2 (N9K-C93180YC-EX)	NX-OS 9.2(1)	IPN router deployed in NX-OS Standalone Mode
Hyperconverged Infrastructure (Cisco HyperFlex Standard & Stretched Clusters)	Witness VM		1.0.4	Deployed in existing infrastructure, outside ACI; Available as an OVA
	Pod 1	Pod 2		
	Cisco HX220c M4S x 4 (HX220C-M4S)	—	3.5(2e) *	<ul style="list-style-type: none"> 4-node Management Cluster (Standard Cluster); Cisco HyperFlex Hybrid M4 Nodes with 10G VIC 1227 (UCSC-MLOM-CSC-02)
	Cisco UCS 6248 FI x 2 (UCS-FI-6248UP)	—	4.0(2d)	1RU 10G Fabric Interconnect with 48 ports
	Cisco HX220C-M5SX x 4 (HX220C-M5SX)	HX220C-M5SX x 4 (HX220C-M5SX)	3.5(2e) *	<ul style="list-style-type: none"> 8-node Application Cluster (4-4 Stretch Cluster); Cisco HyperFlex Hybrid M5 Nodes with 40G VIC 1387 (UCSC-MLOM-C40G-03)
	Cisco UCS 6332 FI x 2 (UCS-FI-6332-16UP)	Cisco UCS 6332 FI x 2 (UCS-FI-6332UP)	4.0(2d)	<ul style="list-style-type: none"> Pod 1 FI: 1RU, 40G FI with 40 ports (24 fixed ports) Pod 2 FI: 1RU, 40G FI with 32 fixed ports
Virtualization	Pod 1	Pod 2		
	VMware vSphere 6.5U2 EP13	VMware vSphere 6.5U2 EP13	6.5U2 EP13	Hypervisor – Custom Cisco Build: 13004031
	VMware vCenter Server Appliance 6.5 U2e	—	6.5U2e	<ul style="list-style-type: none"> VCSA for Application Cluster; Management Cluster is managed by a VCSA outside the ACI Fabric Version: 6.5.0.23100 Build Number 11347054
	VMware vDS, Cisco AVE	Cisco AVE	2.0(1a)	<ul style="list-style-type: none"> Virtual Switches VMware vDS used in Management Cluster Cisco AVE used in Application Cluster
Security	Cisco Umbrella			Cloud-based security for Enterprise; Virtual Appliances(Optional) deployed on-premise: https://umbrella.cisco.com
Management & Monitoring	Cisco UCS Manager		4.0(2d)	Management Cluster is managed by a VMware vCenter Server outside ACI Fabric
	Cisco HyperFlex Connector			Virtual Switches – VMware vDS in Management Cluster and Cisco AVE in Application Cluster
	Cisco Intersight			Cloud-based Management Tool
	Cisco HyperFlex vCenter Plugin		3.0.1.29754	6.5 Flash client – added by HX Installer
	Cisco ACI vCenter Plugin		3.2.2000.12	
Tools	HX Bench, VdBench			Load Generation Tools



This solution was primarily validated using HyperFlex release 3.5(2d), and 3.5(2e) for a subset of test cases.

Interoperability

To use other hardware models or software versions in this design, verify interoperability using the following matrixes. Also, review the release notes for release and product documentation.

- [Cisco UCS and HyperFlex Hardware and Software Interoperability Tool](#)
- [Cisco ACI Recommended Release](#)
- [Cisco ACI Virtualization Compatibility Matrix](#)
- [Cisco APIC and ACI Virtual Edge Support Matrix](#)
- [VMware Compatibility Guide](#)

Solution Validation

The solution was validated for basic data forwarding by deploying virtual machine running VdBench and IOMeter tools. The system was validated for resiliency by failing various aspects of the system under load. Examples of the types of tests executed include:

- Failure and recovery of various links and components between the sites and within each site.
- Failure events triggering vSphere high availability between sites.
- Failure events triggering vMotion between sites.
- All tests were performed under load, using load generation tools. Different IO profiles representative of customer deployments were used.

Summary

The **Cisco HyperFlex Stretched Cluster with Cisco ACI Multi-Pod Fabric** solution for VMware vSphere deployments delivers an active-active data center solution that can span different geographical locations to provide disaster avoidance in Enterprise data centers. In the event of a site failure, Cisco HyperFlex stretched cluster can enable business continuity with no data loss. To interconnect the data centers, Cisco HyperFlex offers is integrated with Cisco ACI Multi-Pod fabric to provide seamless Layer 2 extension and workload mobility between sites. Cisco ACI also offers a software-defined, application-centric, policy-based network architecture that enable applications to be deployed in a simple and secure manner. The ACI Multi-Pod fabric is also centrally and uniformly managed using a single APIC cluster that simplifies the operation of a multi data center solution.

References

Cisco HyperFlex

- Comprehensive Documentation for Cisco HyperFlex: <http://hyperflex.io>
- Comprehensive Documentation Roadmap for Cisco HyperFlex: https://www.cisco.com/c/en/us/td/docs/hyperconverged_systems/HyperFlex_HX_DataPlatformSoftware/HX_Documentation_Roadmap/HX_Series_Doc_Roadmap.html
- Pre-installation Checklist for Cisco HX Data Platform: https://www.cisco.com/c/en/us/td/docs/hyperconverged_systems/HyperFlex_HX_DataPlatformSoftware/HyperFlex_Preinstall_Checklist/b_HX_Data_Platform_Preinstall_Checklist.html
- HyperFlex Hardening Guide: https://www.cisco.com/c/dam/en/us/support/docs/hyperconverged-infrastructure/hyperflex-hx-data-platform/HX-Hardening_Guide_v3_5_v12.pdf
- HyperFlex Installation Guide for Cisco Intersight: https://www.cisco.com/c/en/us/td/docs/hyperconverged_systems/HyperFlex_HX_DataPlatformSoftware/HyperFlex_Installation_Guide_for_Intersight/b_HyperFlex_Installation_Guide_for_Intersight/b_HyperFlex_Installation_Guide_for_Intersight_chapter_011.html
- Operating Cisco HyperFlex HX Data Platform Stretched Clusters: <https://www.cisco.com/c/dam/en/us/products/collateral/hyperconverged-infrastructure/hyperflex-hx-series/operating-hyperflex.pdf>
- Cisco HyperFlex Systems Stretched Cluster Guide, Release 3.5: https://www.cisco.com/c/en/us/td/docs/hyperconverged_systems/HyperFlex_HX_DataPlatformSoftware/HyperFlex_Stretched_Cluster/3_5/b_HyperFlex_Systems_Stretched_Cluster_Guide_3_5.html

Cisco UCS

- Cisco Unified Computing System: <http://www.cisco.com/en/US/products/ps10265/index.html>
- Cisco UCS 6300 Series Fabric Interconnects: <http://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-6300-series-fabric-interconnects/index.html>
- Cisco UCS 5100 Series Blade Server Chassis: <http://www.cisco.com/en/US/products/ps10279/index.html>
- Cisco UCS 2300 Series Fabric Extenders: <https://www.cisco.com/c/en/us/products/collateral/servers-unified-computing/ucs-6300-series-fabric-interconnects/datasheet-c78-675243.html>
- Cisco UCS 2200 Series Fabric Extenders: https://www.cisco.com/c/en/us/products/collateral/servers-unified-computing/ucs-6300-series-fabric-interconnects/data_sheet_c78-675243.html
- Cisco UCS B-Series Blade Servers: <http://www.cisco.com/en/US/partner/products/ps10280/index.html>

- Cisco UCS C-Series Rack Mount Servers:
<http://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-c-series-rack-servers/index.html>
- Cisco UCS VIC Adapters:
http://www.cisco.com/en/US/products/ps10277/prod_module_series_home.html
- Cisco UCS Manager:
<http://www.cisco.com/en/US/products/ps10281/index.html>
- Cisco UCS Manager Plug-in for VMware vSphere Web Client:
http://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/sw/vmware_tools/vCenter/vCenter_Plugin_Release_Notes/2_o/b_vCenter_RN_for_2x.html

Cisco ACI Application Centric Infrastructure (ACI)

- Cisco ACI Infrastructure Best Practices Guide:
https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/1-x/ACI_Best_Practices/b_ACI_Best_Practices.html
- Cisco ACI Infrastructure Release 2.3 Design Guide:
<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-737909.pdf>
- Cisco ACI Multi-Pod Configuration Whitepaper: <https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739714.html>
- Cisco ACI Multi-Pod White Paper:
<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-737855.html>
- Cisco APIC Layer Network Configuration Guide, Release 4.0(1):
https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/4-x/L3-configuration/Cisco-APIC-Layer-3-Networking-Configuration-Guide-401/Cisco-APIC-Layer-3-Networking-Configuration-Guide-401_chapter_010110.html#id_30270
- ACI Switch Command Reference, NX-OS Release 13.X:
https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/3-x/cli/inxos/13x/b_ACI_Switch_Command_Ref_13x.html

Cisco AVE

- Cisco ACI Virtual Edge White paper:
<https://www.cisco.com/c/dam/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-740131.pdf>
- Cisco APIC and ACI Virtual Edge Support Matrix:
<https://www.cisco.com/c/dam/en/us/td/docs/Website/datacenter/aveavsmatrix/index.html>

Security

- Integrating Cisco Umbrella to Cisco HyperFlex and Cisco UCS Solutions:
<https://www.cisco.com/c/dam/en/us/products/collateral/hyperconverged-infrastructure/hyperflex-hx-series/whitepaper-c11-741088.pdf>

Interoperability Matrixes

- Cisco UCS and HyperFlex Hardware Compatibility Matrix: <https://ucshcltool.cloudapps.cisco.com/public/>
- VMware and Cisco Unified Computing System:
<http://www.vmware.com/resources/compatibility>
- Cisco ACI Virtualization Compatibility Matrix:
<https://www.cisco.com/c/dam/en/us/td/docs/Website/datacenter/aci/virtualization/matrix/virtmatrix.html>
- Cisco APIC and ACI Virtual Edge Support Matrix:
<https://www.cisco.com/c/dam/en/us/td/docs/Website/datacenter/aveavsmatrix/index.html>

About the Author

Archana Sharma, Technical Leader, Cisco UCS Data Center Solutions, Cisco Systems Inc.

Archana Sharma is Technical Marketing Engineer with over 20 years of experience at Cisco on a range of technologies that span Data Center, Desktop Virtualization, Collaboration, and other Layer2 and Layer3 technologies. Archana is focused on systems and solutions for Enterprise and Provider deployments, including delivery of Cisco Validated designs for 10 years. Archana is currently working on designing and integrating Cisco UCS-based Converged Infrastructure solutions. Archana holds a CCIE (#3080) in Routing and Switching and a Bachelor's degree in Electrical Engineering from North Carolina State University.

Acknowledgements

For their support and contribution to the design, validation, and creation of this Cisco Validated Design, the author would like to thank:

- Haseeb Niazi, Technical Marketing Engineer, Cisco Systems, Inc.
- Ramesh Isaac, Technical Marketing Engineer, Cisco Systems, Inc.
- Allen Clark, Technical Marketing Engineer, Cisco Systems, Inc.