



Cisco Nexus Dashboard Orchestrator Configuration Guide for ACI Fabrics, Release 4.2(x)

First Published: 2023-08-22

Last Modified: 2024-01-15

Americas Headquarters

Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
<http://www.cisco.com>
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883

THE SPECIFICATIONS AND INFORMATION REGARDING THE PRODUCTS REFERENCED IN THIS DOCUMENTATION ARE SUBJECT TO CHANGE WITHOUT NOTICE. EXCEPT AS MAY OTHERWISE BE AGREED BY CISCO IN WRITING, ALL STATEMENTS, INFORMATION, AND RECOMMENDATIONS IN THIS DOCUMENTATION ARE PRESENTED WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED.

The Cisco End User License Agreement and any supplemental license terms govern your use of any Cisco software, including this product documentation, and are located at: <https://www.cisco.com/c/en/us/about/legal/cloud-and-software/software-terms.html>. Cisco product warranty information is available at <https://www.cisco.com/c/en/us/products/warranty-listing.html>. US Federal Communications Commission Notices are found here <https://www.cisco.com/c/en/us/products/us-fcc-notice.html>.

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 THIS MANUAL, EVEN IF CISCO OR ITS SUPPLIERS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Any products and features described herein as in development or available at a future date remain in varying stages of development and will be offered on a when-and if-available basis. Any such product or feature roadmaps are subject to change at the sole discretion of Cisco and Cisco will have no liability for delay in the delivery or failure to deliver any products or feature roadmap items that may be set forth in this document.

Any Internet Protocol (IP) addresses and phone numbers used in this document are not intended to be actual addresses and phone numbers. Any examples, command display output, network topology diagrams, and other figures included in the document are shown for illustrative purposes only. Any use of actual IP addresses or phone numbers in illustrative content is unintentional and coincidental.

The documentation set for this product strives to use bias-free language. For the purposes of this documentation set, bias-free is defined as language that does not imply discrimination based on age, disability, gender, racial identity, ethnic identity, sexual orientation, socioeconomic status, and intersectionality. Exceptions may be present in the documentation due to language that is hardcoded in the user interfaces of the product software, language used based on RFP documentation, or language that is used by a referenced third-party product.

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: <https://www.cisco.com/c/en/us/about/legal/trademarks.html>. Third-party trademarks mentioned 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. (1721R)

© 2023–2024 Cisco Systems, Inc. All rights reserved.



CONTENTS

CHAPTER 1	New and Changed Information 1
	New and Changed Information 1
CHAPTER 2	Dashboard, Sites, and GUI Overview 3
	Dashboard 3
	Overview 4
	Operate 5
	Configure 6
	Admin 9
PART I	Application & Fabric Management 11
CHAPTER 3	Templates Overview and Operations 13
	Schema and Template Design Considerations 13
	Concurrent Configuration Updates 17
	Assigning Templates to Sites 19
	Disassociating Template from Sites 19
	Deploying Templates 20
	Undeploying Templates 24
	Bulk Update for Template Objects 25
	Template Versioning 28
	Tagging Templates 29
	Viewing History and Comparing Previous Versions 29
	Reverting Template to Earlier Version 32
	Template Review and Approval 33
	Enabling Template Approval Requirement 33

Create Users with Required Roles	33
Requesting Template Review and Approval	34
Reviewing and Approving Templates	35
Configuration Drifts	36
Reconciling Configuration Drifts in Application Templates	38
Cloning Templates	41
Migrating Objects Between Templates	42
Viewing Currently Deployed Configuration	43
Schema Overview and Deployment Visualizer	44

CHAPTER 4 **Tenants and Tenant Policies Templates** 47

Tenants Overview	47
Creating New Tenants	48
Importing Existing Tenants	49
Creating Tenant Policy Templates	50

CHAPTER 5 **Schemas and Application Templates** 63

Shadow Objects	63
Hiding Shadow Objects in APIC GUI	67
Creating Schemas and Templates	68
Importing Schema Elements From APIC Sites	70
Configuring VRFs	71
Configuring Bridge Domains	72
Configuring Bridge Domain's Site-Local Properties	76
Configuring Application Profiles and EPGs	78
Configuring EPG's Site-Local Properties	80
Configuring Contracts and Filters	84
Viewing Schemas	87
Cloning Schemas	87

CHAPTER 6 **Fabric Management Templates** 89

Fabric Management Templates	89
Creating Fabric Policies	90
Creating Fabric Resources Policies	102

Creating Monitoring Policies 107

PART II Operations 113

CHAPTER 7 Audit Logs 115

Audit Logs 115

CHAPTER 8 Backup and Restore 117

Configuration Backup and Restore Guidelines 117

Configuring Remote Locations for Backups 119

Creating Backups 120

Restoring Backups 120

Exporting (Downloading) Backups 128

Importing Backups to Remote Location 129

Backup Scheduler 130

CHAPTER 9 Upgrading Sites 131

Overview 131

Guidelines and Limitations 133

Downloading Controller and Switch Node Firmware to Sites 133

Upgrading Controllers 136

Upgrading Nodes 138

CHAPTER 10 Tech Support 143

Tech Support and System Logs 143

Downloading System Logs 144

Streaming System Logs to External Analyzer 144

PART III Infrastructure Management 151

CHAPTER 11 System Configuration 153

System Configuration Settings 153

System Alias and Banner 153

CHAPTER 12	Preparing Cisco APIC Sites	155
	Pod Profile and Policy Group	155
	Configuring Fabric Access Policies for All APIC Sites	156
	Configuring Fabric Access Global Policies	156
	Configuring Fabric Access Interface Policies	157
	Configuring Sites That Contain Remote Leaf Switches	159
	Remote Leaf Guidelines and Limitations	159
	Configuring Routable Subnets for Remote Leaf Switches	159
	Enabling Direct Communication for Remote Leaf Switches	160
	Cisco Mini ACI Fabrics	160

CHAPTER 13	Adding and Deleting Sites	163
	Cisco NDO and APIC Interoperability Support	163
	Adding Cisco ACI Sites	165
	Removing Sites	166
	Cross Launch to Fabric Controllers	168

CHAPTER 14	Configuring Infra General Settings	169
	Infra Configuration Dashboard	169
	Partial Mesh Intersite Connectivity	171
	Configuring Infra: General Settings	171

CHAPTER 15	Configuring Infra for Cisco APIC Sites	177
	Refreshing Site Connectivity Information	177
	Configuring Infra: On-Premises Site Settings	178
	Configuring Infra: Pod Settings	181
	Configuring Infra: Spine Switches	181

CHAPTER 16	Configuring Infra for Cisco Cloud Network Controller Sites	185
	Refreshing Cloud Site Connectivity Information	185
	Configuring Infra: Cloud Site Settings	186
	Recovering from Cloud Network Controller Site Downtime	188

CHAPTER 17	Deploying Infra Configuration for ACI Sites	191
	Deploying Infra Configuration	191
	Enabling Connectivity Between On-Premises and Cloud Sites	192
CHAPTER 18	CloudSec Encryption	197
	Cisco ACI CloudSec Encryption	197
	Requirements and Guidelines	198
	CloudSec Encryption Terminology	201
	CloudSec Encryption and Decryption Handling	202
	CloudSec Encryption Key Allocation and Distribution	204
	Configuring Cisco APIC for CloudSec Encryption	206
	Configuring Cisco APIC for CloudSec Encryption Using GUI	207
	Configuring Cisco APIC for CloudSec Encryption Using NX-OS Style CLI	207
	Configuring Cisco APIC for CloudSec Encryption Using REST API	208
	Enabling CloudSec Encryption in Cisco Nexus Dashboard Orchestrator	209
	Verifying CloudSec Configuration on Switches	210
	Rekey Process During Spine Switch Maintenance	212
	Disabling and Re-Enabling Re-Key Process Using NX-OS Style CLI	212
	Disabling and Re-Enabling Re-Key Process Using REST API	213
PART IV	Features and Use Cases	215
CHAPTER 19	DHCP Relay	217
	DHCP Relay Policy	217
	Guidelines and Limitations	217
	Creating DHCP Relay Policies	218
	Creating DHCP Option Policies	220
	Assigning DHCP Policies	221
	Creating DHCP Relay Contract	222
	Verifying DHCP Relay Policies in APIC	223
	Editing or Deleting Existing DHCP Policies	224
CHAPTER 20	EPG Preferred Group	225

EPG Preferred Groups Overview and Limitations	225
Configuring EPGs for Preferred Group	226

CHAPTER 21

External Connectivity (L3Out)	229
L3Out Template Overview	229
Guidelines and Limitations	233
Greenfield Deployment	234
Creating Tenant Policy Template	234
Creating L3Out Template	240
Importing Existing L3Out Configuration	245
Overview of Importing L3Out Configuration	245
Importing Tenant Policy Template Objects	248
Importing L3Out Objects	251
Viewing L3Out Neighbors	256

CHAPTER 22

Intersite L3Out	259
Intersite L3Out Overview	259
Intersite L3Out Guidelines and Limitations	260
Configuring External TEP Pool	261
Configuring External EPG to Use Intersite L3Out	262
Creating a Contract for Intersite L3Out	264
Use Cases	265
Intersite L3Out for Application EPGs (Intra-VRF)	265
Shared Services with Intersite L3Out for Application EPGs (Inter-VRF)	268
Intersite Transit Routing	271

CHAPTER 23

Intersite L3Out with PBR	275
Intersite L3Out with PBR	275
Supported Use Cases	276
Guidelines and Limitations	279
Create Service Device Template	280
Add Service Chaining to Contract	282

CHAPTER 24

Intersite Transit Routing with PBR	283
---	------------

Intersite Transit Routing with PBR	283
Traffic Flow	284
Intersite Transit Routing with PBR Guidelines and Limitations	285
Create Service Device Template	287
Create Contract and Add Service Chaining	293

CHAPTER 25

Layer 3 Multicast	295
Layer 3 Multicast	295
Layer 3 Multicast Routing	296
Rendezvous Points	296
Multicast Filtering	297
Layer 3 Multicast Guidelines and Limitations	298
Creating Multicast Route Map Policy	299
Enabling Any-Source Multicast (ASM) Multicast	301
Enabling Source-Specific Multicast (SSM)	302

CHAPTER 26

QoS Preservation Across IPN	305
QoS and Global DSCP Policy	305
DSCP Policy Guidelines and Limitations	305
Configuring Global DSCP Policy	306
Set QoS Level for EPGs and Contracts	307

CHAPTER 27

SD-Access and ACI Integration	311
Cisco SD-Access and Cisco ACI Integration	311
Macro Segmentation	312
Cisco SD-Access and Cisco ACI Integration Guidelines	314
Onboarding the DNA Center	315
Configuring Connectivity Toward the SD-Access Domain	316
Viewing the Status of the SD-Access to ACI Integration	318
Extending a Virtual Network	320
Mapping or Unmapping a VN to a VRF	321
Configuring Transit Routing	323

CHAPTER 28

SD-WAN Integration	329
---------------------------	------------

SD-WAN Integration	329
SD-WAN Integration Guidelines and Limitations	330
Adding a vManage Controller	331
Configuring Global DSCP Policy	332
Set QoS Level for EPGs and Contracts	334

CHAPTER 29

Multi-Site and SR-MPLS L3Out Handoff	337
Overview and Use Cases	337
SR-MPLS Infra Requirements and Guidelines	340
SR-MPLS Tenant Requirements and Guidelines	342
Greenfield Deployment	344
Creating Custom QoS Policy for SR-MPLS	344
Creating SR-MPLS Infra L3Out	346
Creating SR-MPLS Route Map Policy	349
Creating SR-MPLS Tenant L3Outs in L3Out Templates	351
Configure EPG-to-External-EPG (North-South) Communication	352
Importing Existing SR-MPLSL3Out Configuration	355
Overview of Importing SR-MPLS Configuration	355
Importing Tenant Policies Template Objects	358
Importing SR-MPLS Objects	361

CHAPTER 30

vzAny Contracts	363
vzAny and Multi-Site	363
vzAny and Multi-Site Guidelines and Limitations	364
Create Contract and Filters	366
Configure vzAny to Consume/Provide a Contract	367
Create EPGs to Be Part of the vzAny VRF	367
Free Intra-VRF Communication	368
Stretched EPGs	370
Site-Local EPGs	370
Combination of Site-Local and Stretched EPGs	371
Intra-VRF Intersite L3Out	372
Inter-VRF Intersite L3Out	373
Many-to-One Communication	373

Provider EPG Within vzAny VRF 375

Provider EPG In Its Own VRF 375

CHAPTER 31

vzAny with PBR 377

vzAny with PBR Overview 377

Traffic Flow: Intra-VRF vzAny-to-vzAny 379

Traffic Flow: Intra-VRF vzAny-to-EPG 381

Traffic Flow: Intra-VRF vzAny-to-External-EPG (L3Out EPG) 383

vzAny with PBR Guidelines and Limitations 385

Create Service Device Template 387

Create Application Template 394

Add Service Chaining to Contract 398



CHAPTER 1

New and Changed Information

- [New and Changed Information](#), on page 1

New and Changed Information

The following table provides an overview of the significant changes to the organization and features in this guide from the release the guide was first published to the current release. The table does not provide an exhaustive list of all changes made to the guide.

Table 1: Latest Updates

Release	New Feature or Update	Where Documented
4.2(3)	ACI Multi-Site support for vzAny PBR and L3Out-to-L3Out PBR use cases	vzAny with PBR , on page 377
	New service chaining configuration workflows	Intersite Transit Routing with PBR , on page 283

Release	New Feature or Update	Where Documented
4.2(1)	<p>The following additional template object properties can now be configured directly from Nexus Dashboard Orchestrator for ACI fabrics:</p> <ul style="list-style-type: none">• Annotations for template objects• EPG admin state• VMM Domain customized delimiter, port bindings, NetFlow, MAC address changes, forged transmits, promiscuous mode• Intra EPG contracts• Bridge Domain (BD) endpoint (EP) move detection mode• VRF BD enforcement status• External EPG QoS class and subnet name	



CHAPTER 2

Dashboard, Sites, and GUI Overview

- [Dashboard, on page 3](#)
- [Overview, on page 4](#)
- [Operate, on page 5](#)
- [Configure, on page 6](#)
- [Admin, on page 9](#)

Dashboard

The Cisco Nexus Dashboard Orchestrator (NDO) GUI is a browser-based graphical interface for configuring and monitoring your Cisco APIC, Cloud Network Controller, and NDFC deployments.

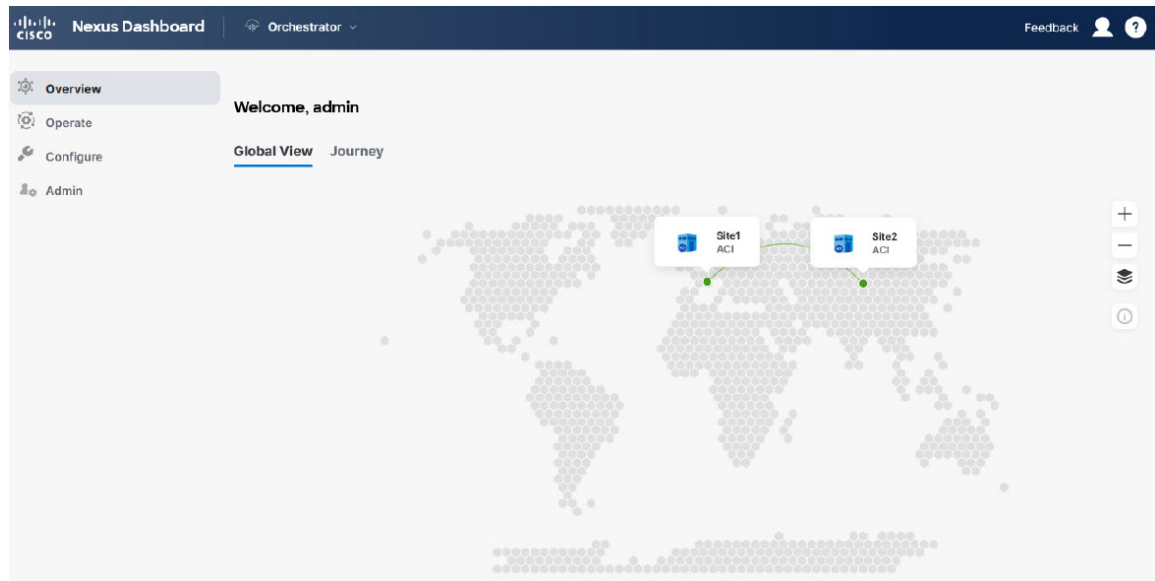
The GUI is arranged according to the functions. For example, the **Overview** page contains a summary of your fabrics, and their health. Toggle between the **Global View** map or **Journey** to view the **Getting Started Map**.

The top navigation bar contains the common Cisco Nexus Dashboard menus, such as the **Nexus Dashboard** home button that allows you to return to the Cisco Nexus Dashboard GUI. You can use **Orchestrator** drop-down list to switch between **Admin-Console** or to **One View**.

User menu has options to configure your user preferences, change password, or sign out. The **Feedback** link is to provide any comments or suggestions about the product, the **?** menu includes help, information about the release and welcome screen.

Functions like **Operate** contain sites and tenants operations, and **Configure** contains site to site connectivity, tenant configurations, and fabric templates. The **Admin** category contains functions like System Configurations, Integrations, and so on. The functionality of each NDO GUI page is described in specific chapters later in this document.

Figure 1: Cisco Nexus Dashboard Orchestrator



Overview

The Cisco Nexus Dashboard Orchestrator **Overview** option displays a **Global View** map of your multi-site implementations in addition to their current functionality and health. The Settings icon allows you to overlay, Site to site Connectivity, Tooltips, and the Group Markers information over your map. You can zoom in or out using the + or - icons to any particular region of the map and then use the **Save Layout** option to save the configuration to the user profile.

The **Sites** page provides general information about each site. If you hover the mouse pointer over any particular site, it animates the site to site connectivity and health status. You can click the sites to see **Site Details** which gives you options like **Refresh** to reload the details and **Launch**, which allows you to open the site directly from the Cisco Nexus Dashboard Orchestrator.

Colors coding denotes the fault severity and referenced in the **Map Legend** icon of the map. Red implies critical, yellow for degraded, and green denotes healthy state. Continuous line implies connectivity and sites or region unreachable by Cisco Nexus Dashboard Orchestrator are grayed out.

The **Overview** section has the following functional information:

- **Audit Logs:** Captures the most recent events and faults that occur in the environment.
- **Fabric Interconnect:** Shows the status of end-to-end interconnect between sites.
- Number of **Sites**, **Tenants**, and **In-Sync / Out-of Sync** Templates. It also displays the **Inventory** status.

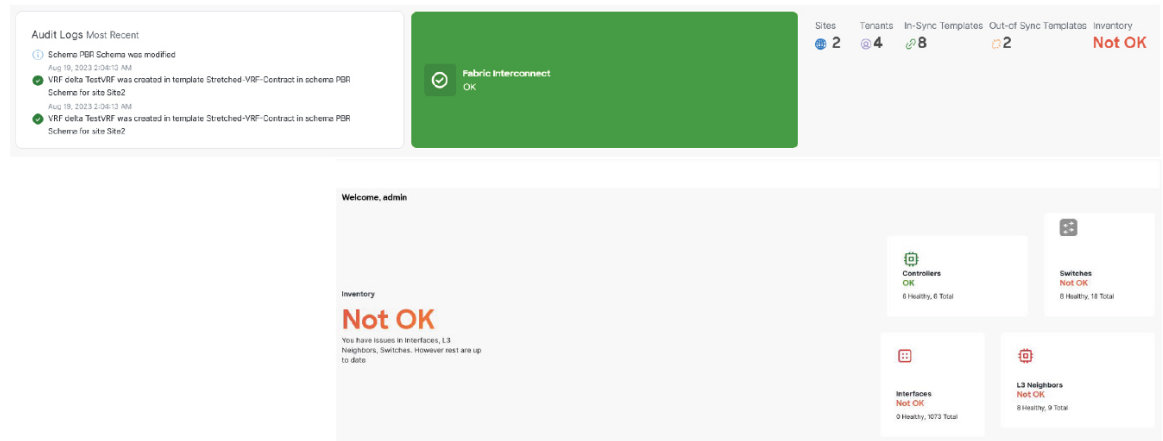
You can toggle between the **Global View** map or **Journey** to access several common tasks, such as adding sites or schemas, configuring specific policies, or performing administrative tasks.

As you are all set in your journey map, you can see a summary of the inventory and status of:

- **Sites:** This page shows general information like site health status, connectivity, and inventory information. You can click **Site Details** to see Operational information about that particular site.

- **Templates:** Shows health and number of templates visualized **By Type**, **By Status**, and **By State**.
- **Tenants:** Shows health and number of tenants visualized **By Policy**, **By Template**, and **By Sites**.
- **Inventory:** Shows inventory health status along with the health status and number of **Controllers**, **Switches**, **Interfaces**, and **L3 Neighbors**.

Figure 2: Overview



Operate

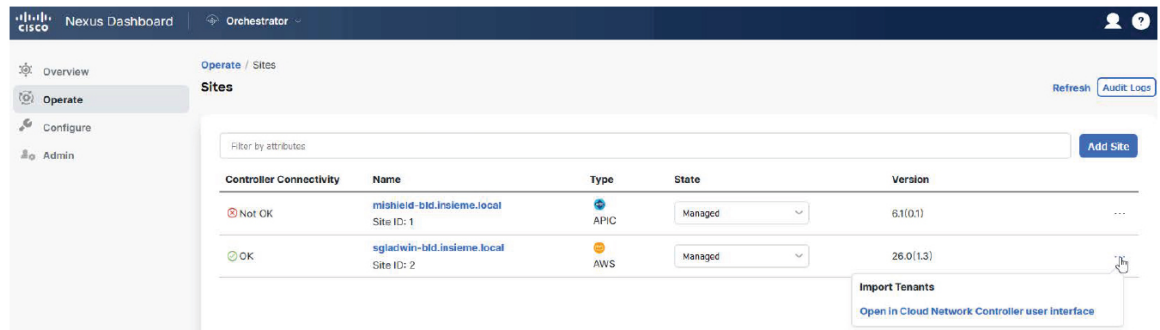
Operate menu allows you to perform operational functions at **Sites** and **Tenants**. **Sites** presents a list of sites in tabular form with operational information. You can filter or sort the table using the attributes like:

- **Controller Connectivity**
- **Name**
- **Type**
- **State**
- **Version**

The three dots in the last column of the table allow you can open the site's UI.

You can add a new site using the **Add Site** button. Click the **Audit Logs** to review audit logs for a set timeframe.

Figure 3: Operate



Configure

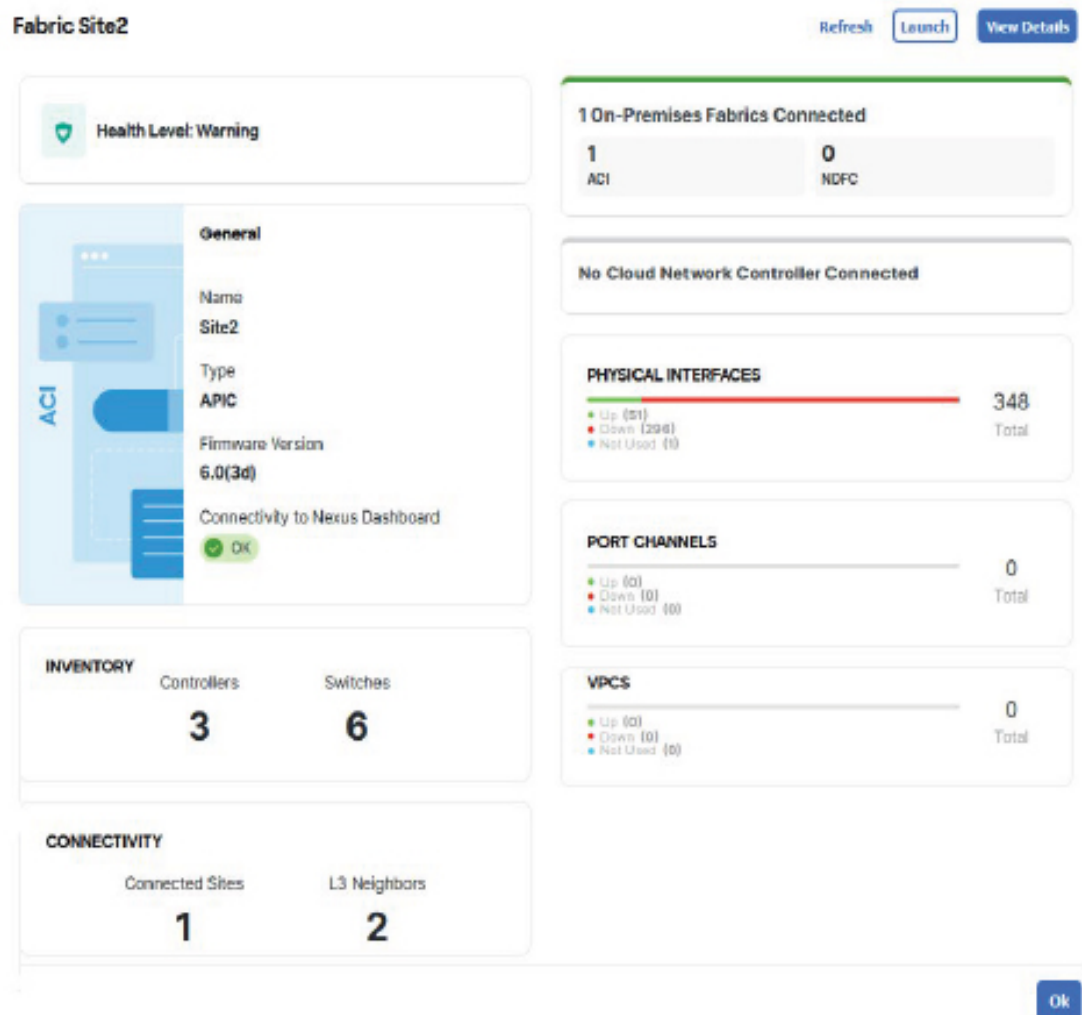
Configure menu allows you to configure **Site to Site Connectivity** and **Tenant or Fabric Templates**.

You can **Configure**:

Site to Site Connectivity

Site to Site Connectivity: Shows a global map view of your Multi-Site implementations in addition to their current functionality and health. The **Sites** page provides general information about each site. If you hover the mouse pointer over any particular site, it animates the site to site connectivity and health status. You can click **View Details** to see the sites details here you get options like **Refresh** to reload the details and **Launch**, which allows you to open the site.

Figure 4: Configure



The Settings icon allows you to overlay the map with information like, Site to site Connectivity, Tooltips, and the Group Markers. You can zoom in or out using the + or - icons to any particular region of the map and then use the **Save Layout** option to save the configuration to the user profile.

Colors coding denotes the fault severity and referenced in the **Map Legend** icon of the map. Red implies critical, yellow for degraded, and green denotes healthy state. Continuous line implies connectivity and sites or region unreachable by Cisco Nexus Dashboard Orchestrator are grayed out.

Current control plane site to site connectivity configuration is displayed under the **General Settings** with fields like:

- **BGP Peering Type**
- **Keep alive interval (Seconds)**
- **Hold Interval (Seconds)**
- **Stale Interval (Seconds)**

- **Graceful Restart**
- **Maximum AS Limit**
- **BGP TTL Between Peers**
- **IANA Assigned Port**

You can configure these options using the **Configure** button. This option allows you to review the **Audit Logs** and also **Deploy** these configurations. The **Sites** tab shows the deployment status of individual sites along with its health status.

Tenant Template

Tenant Template configuration has options to configure templates. The page has options for:

- **Applications:** This tab shows a table of schemas, you can filter or sort the table using the attributes like templates that are deployed, tenants, and policies associated with the schema. The three dots on the last column of the table allow you to **Edit**, **Delete**, and **Clone** the schema. You can add a new schema using the **Add Schema** button. You can select an individual schema for an overview of the schema.
- **L3Out:** This tab shows a tabular deployment status of the L3Out templates. You can filter or sort the table using the attributes like status, name, tenant, sites, and policies. The three dots on the last column of the table allow you to **Edit** or **Delete** the L3Out template. You can add a new L3Out template using the **Create L3Out Template** button. You can select each individual template to get a summary of the template. You can edit or deploy the template using the appropriate buttons under the summary. **Action** menu has options to perform site or template level actions.
- **Monitoring Policies:** Allows you to create and edit monitoring policies templates.
- **Service Device:** This tab shows a tabular deployment status of the service device template. You can filter or sort the table using the attributes like status, name, tenant, sites, and policies. The three dots on the last column of the table allow you to **Edit** or **Delete** the service device template. You can add a new Service Device template using the **Create Service Device Template** button. You can select each individual template to get a summary of the template properties along with each site deployed.
- **Tenant Policies:** This tab shows a tabular deployment status of the tenant policy template. You can filter or sort the table using the attributes like status, name, tenant, sites, and policies. The three dots on the last column of the table allow you to **Edit** or **Delete** the tenant policy template. You can edit or deploy the template using the appropriate buttons under the Template Summary. You can add a new tenant policy template using the **Create Tenant Policy Template** button. **Action** menu has options to perform site or template level actions.

Fabric Template

Fabric Template configuration has options to configure Fabric policy templates. This menu option allows you to configure:

- **Fabric Policies:** This tab shows a table of fabric policies. You can filter or sort the table using the attributes like status, name, sites, and policies. The three dots on the last column of the table allow you to **Edit** or **Delete** the fabric policy template. You can add a new fabric policy template using the **Create Fabric Policy Template** button. You can edit or deploy the template using the appropriate buttons under the Template Summary. **Action** menu has options to perform site or template level actions.

- **Fabric Resource Policies:** This tab shows a table of fabric resource policy templates. You can filter or sort the table using the attributes like status, name, sites, and policies. The three dots on the last column of the table allow you to **Edit** or **Delete** the fabric resource policy template. You can add a new fabric resource policy template using the **Create Fabric Resource Policy Template** button. You can edit or deploy the template using the appropriate buttons under the Template Summary. **Action** menu has options to perform site or template level actions.
- **Monitoring Access Policies:** This tab shows a table of monitoring access policy templates. You can filter or sort the table using the attributes like status, name, sites, and policies. The three dots on the last column of the table allow you to **Edit** or **Delete** the monitoring access policy template. You can add a new monitoring access policy template using the **Create Monitoring Policy Template** button. You can edit or deploy the template using the appropriate buttons under the Template Summary. **Action** menu has options to perform site or template level actions.

Admin

Admin menu allows you to perform administrative functions like system configurations, integrations, software management, tech support, and backups and restore.

Software Management

This option provides the firmware update summary for every **Controllers** and **Nodes** in the fabric summary for each site. The **Overview** tab accounts for status like update completed, downloading, ready to install, installing, not supported, and failed. You can set firmware updates for each site using the **Set Update** button in either of the tabs. Use **Setup Download** in the **Downloads** tab to set up a download firmware update image to the selected sites.

Backup & Restore

Backup and restore menu allows you to upload or create new backup and restore to a remote location. You can schedule backup or restore operations for a remote location using the **No Schedule** button. This menu also has option to create a new remote location.

System Configuration

The system configuration tab allows you to assign System Alias and Banners along with option to assign the severity to the banner. You can enable the **Schema Work Management** by editing the **Change Control** option. You can also view and download the logs in the **Audit Logs** tab.

Integration

Allows you to integrate **SD-WAN** domain controllers and policies and DNAC (Cisco DNA) deployment to the fabric.

Tech Support

Tech support option allows you to capture and view the audit logs or all logs with **External Streaming** option enabled using service like **splunk** or **syslog**. You can add a maximum of 5 servers for this operation. You can download and save system logs to your local system using the download button.

Figure 5: Admin

General Settings

BGP Peering Type full-mesh	Keep Alive Interval (Seconds) 60	Hold Interval (Seconds) 180	BGP TTL Between Peers 16
State Interval (Seconds) 300	Graceful Start True	Maximum AS Limit N/A	IANA Assigned Port False

Site1

Pods 2	Spines 4	ACI Multi-Site On BGP ASN 55550	Cloudsec Encryption Off OSPF Area ID backbone	APIC Site ID 1 OSPF Area Type regular	Overlay Multicast TEIP 192.10.100.100 External Routed Domain main /13-Intersite_RoutedDomain
------------------	--------------------	--	--	--	---

Inter-Site Connections

[Overlay Status](#) [Underlay Status](#)

Site Name	Deployment Status	Operational Status	BGP Evpn Status	Tunnel Status
Site2	N/A	OK	4 ↑ 4 ↓ 0 N/A	16 ↑ 16 ↓ 0

Site2

Pods 1	Spines 2	ACI Multi-Site On BGP ASN 10010	Cloudsec Encryption Off OSPF Area ID backbone	APIC Site ID 2 OSPF Area Type regular	Overlay Multicast TEIP 192.10.100.100 External Routed Domain main /13-Intersite_RoutedDomain
------------------	--------------------	--	--	--	---

Inter-Site Connections

[Overlay Status](#) [Underlay Status](#)

Device	Device Status	Interface Status	Peering Status	BGP Peer
spine-a1	↑ Up	1/63 ↑ Up	OSPF ↑ Up	-
spine-b1	↑ Up	1/63 ↑ Up	OSPF ↑ Up	-



PART I

Application & Fabric Management

- [Templates Overview and Operations, on page 13](#)
- [Tenants and Tenant Policies Templates, on page 47](#)
- [Schemas and Application Templates, on page 63](#)
- [Fabric Management Templates, on page 89](#)



CHAPTER 3

Templates Overview and Operations

- [Schema and Template Design Considerations, on page 13](#)
- [Concurrent Configuration Updates, on page 17](#)
- [Assigning Templates to Sites, on page 19](#)
- [Disassociating Template from Sites, on page 19](#)
- [Deploying Templates, on page 20](#)
- [Undeploying Templates, on page 24](#)
- [Bulk Update for Template Objects, on page 25](#)
- [Template Versioning, on page 28](#)
- [Template Review and Approval, on page 33](#)
- [Configuration Drifts, on page 36](#)
- [Cloning Templates, on page 41](#)
- [Migrating Objects Between Templates, on page 42](#)
- [Viewing Currently Deployed Configuration, on page 43](#)
- [Schema Overview and Deployment Visualizer, on page 44](#)

Schema and Template Design Considerations

Nexus Dashboard Orchestrator provides a number of policy templates that allow you to define one or more policies together and deploy them to one or more sites at the same time. These include Application templates, Tenant Policies templates, Fabric Policies and Fabric Resources Policies templates, and Monitoring templates. A schema is a collection of Application templates, which are used for defining application policies, with each template assigned to a specific tenant; schemas apply to Application templates only. There are multiple approaches you can take when it comes to creating the templates configurations specific to your deployment use case. The following sections describe a few simple design directions you can take when deciding how to define the schemas, templates, and policies in your Multi-Site domain.

Keep in mind that when designing schemas, you must consider the supported scalability limits for the number of schemas, templates, and objects per schema. Detailed information on verified scalability limits is available in the [Nexus Dashboard Orchestrator Verified Scalability Guides](#) for your release.

Application Templates

There are 3 types of schema templates, also known as application templates, available in Nexus Dashboard Orchestrator, each designed for a specific purpose:

- **ACI Multi-Cloud**—Templates used for Cisco ACI on-premises and cloud sites. This template supports two deployment types:

- **Multi-Site** - The template can be associated to a single site (site-local policies) or to multiple sites (stretched policies) and the option should be selected for Multi-Site Network (ISN) or VXLAN intersite communication to allow template and object stretching between multiple sites.
- **Autonomous** - The template can be associated to one or more sites that are operated independently and are not connected through an Inter-Site Network (no intersite VXLAN communication).

Because autonomous sites are by definition isolated and do not have any intersite connectivity, there is no shadow object configuration across sites and no cross-programming of pctxs or VNIDs in the spine switches for intersite traffic flow.

The autonomous templates also allow for significantly higher deployment scale.

The following sections focus primarily on this type of templates.

- **NDFC**—Templates designed for Cisco Nexus Dashboard Fabric Controller (formerly Data Center Network Manager) sites.

This guide describes Nexus Dashboard Orchestrator configurations for on-premises Cisco ACI fabrics. For information on working with Cisco NDFC sites, see the [Cisco Nexus Dashboard Orchestrator Configuration Guide for NDFC Fabrics](#) instead.

- **Cloud Local**—Templates designed for specific Cloud Network Controller use cases, such as Google Cloud site connectivity, and cannot be stretched between multiple sites.

This guide describes Nexus Dashboard Orchestrator configurations for on-premises Cisco ACI fabrics. For information on working with Cloud Network Controller fabrics, see the [Nexus Dashboard Orchestrator use case library](#) instead.

When creating schemas and application templates, you can choose to adopt one of the following simple approaches:

- **Single Template Deployment**

The simplest schema design approach is a single schema, single template deployment. You can create a single schema with a single template within it and add all VRFs, Bridge Domains, EPGs, Contracts and other elements to that template and deploy it to one or more sites.

This simplest approach to Multi-Site schema creation is to create all objects within the same schema and template. However, the supported number of schemas scalability limit may make this approach unsuitable for large scale deployments, which could exceed those limits.

Note also that with this approach all the objects defined in the template become "stretched objects" and all changes made to the template are always simultaneously deployed to all the sites associated to such template.

- **Multiple Templates with Network Separation**

Another approach to schema design is to separate the networking objects from the application policy configuration. Networking objects include VRFs, Bridge Domains, and subnets, while the application policy objects include EPGs, Contracts, Filters, External EPGs, and Service Graphs.

You begin by defining a schema that contains the network elements. You can choose to create a single schema that contains all the network elements or you can split them into multiple schemas based on which applications reference them or which sites the network is stretched to.

You can then define one or more separate schemas which contain each application's policy objects. This new schema can reference the network elements, such as bridge domains, defined in the previous schema.

After creating and deploying the policy schemas and templates, the networking objects in the networking schema will display the number of external references by the policy schema elements. The object with external references will also be denoted by the ribbon icon.

Schemas designed this way provide logical separation of networking objects from the policy objects. However, this creates additional complexity when it comes to keeping track of externally referenced objects in each schema.

• Multiple Templates Based On Object Relationships

When configuring multiple schemas with shared object references, it is important to be careful when making changes to those objects. For instance, making changes to or deleting a shared networking object can impact applications in one or more sites. Because of that, you may choose to create a template around each individual site that contains only the objects used by that site and its applications, including the VRFs, BDs, EPGs, Contracts, and Filters. And create different templates containing the shared objects.

For example, you can create a `Site1` template that contains only the objects that are local to Site1 and the template is deployed to only that site. Similarly, the `Site2` template contains only the object relevant to site2 and is deployed to that site only. Any change made to any object in either of these templates has no effect on the other one. Then you can create a `shared` template which contains objects that are shared between the sites.

You can extend this scenario for an additional site with the following template layout:

- Site 1 template
- Site 2 template
- Site 3 template
- Site 1 and 2 shared template
- Site 1 and 3 shared template
- Site 2 and 3 shared template
- All shared template

Similarly, rather than separating objects based on which site they are deployed to, you can also choose to create schemas and templates based on individual applications instead. This would allow you to easily identify each application profile and map them to schemas and sites as well as easily configure each application as local or stretched across sites.

However, as this could quickly exceed the templates per schema limit (listed in the [Verified Scalability Guide](#) for your release), you would have to create additional schemas to accommodate the multiple combinations. While this creates additional complexity with multiple additional schemas and templates, it provides true separation of objects based on site or application.

Fabric Policy Templates

In addition to the three types of application templates, Release 4.0(1) adds 3 new templates designed for fabric-wide policies:

- **Fabric Policies** templates can be used for managing the following fabric-wide policies:

- VLAN Pool
- Physical Domains
- SyncE Interface Policies
- Interface Settings
- Node Settings
- Pod Settings
- MACsec
- NTP Policies
- PTP Policies
- QoS DSCP Policies
- QoS SR-MPLS Policies
- QoS Class Policies

For additional information, see [Creating Fabric Policies, on page 90](#).

- **Fabric Resources Policies** templates can be used for managing the following fabric-wide policies:

- Physical Interfaces
- Port Channel Interfaces
- Virtual Port Channel Interfaces
- Node Profiles

These templates reference policies are defined in the Fabric Policies templates, so those templates must be created and deployed first. For additional information, see [Creating Fabric Resources Policies, on page 102](#).

- **Monitoring Policy** templates can be used for managing `Tenant SPAN` or `Access SPAN` policies.

For additional information, see [Creating Monitoring Policies, on page 107](#).

Template Design Best Practices

Beginning with Release 4.0(1), Nexus Dashboard Orchestrator validates and enforces a number of best practices when it comes to template design and deployment. Regardless of the type of template you are creating, keep in mind the following:

- All policy objects must be **deployed** in order according to their dependencies.

For example, when creating a bridge domain (BD), you must associate it with a VRF. In this case, the BD has a VRF dependency so the VRF must be deployed to the fabric before or together with the BD. If these two objects are defined in the same template, then the Orchestrator will ensure that during deployment, the VRF is created first and associate it with the bridge domain.

However, if you define these two objects in separate templates and attempt to deploy the template with the BD first, the Orchestrator will return a validation error as the associated VRF is not yet deployed. In this case, you must deploy the VRF template first, followed by the BD template.

- All policy objects must be **undeployed** in order according to their dependencies, or in the opposite order in which they were deployed.

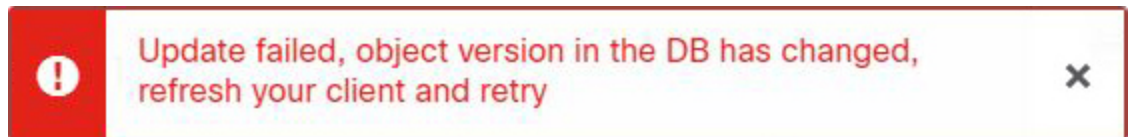
As a corollary to the point above, when you undeploy templates, you must not undeploy objects on which other objects depend. For example, you cannot undeploy a VRF before undeploying the BD with which the VRF is associated.

- No cyclical dependencies are allowed across multiple templates.

Consider a case of a VRF (`vrf1`) associated with a bridge domain (`bd1`), which is in turn associated with an EPG (`epg1`). If you create `vrf1` in `template1` and deploy that template, then create `bd1` in `template2` and deploy that template, there will be no validation errors since the objects are deployed in correct order. However, if you then attempt to create `epg1` in `template1`, it would create a circular dependency between the two template, so the Orchestrator will not allow you to save `template1` addition of the EPG.

Concurrent Configuration Updates

The Nexus Dashboard Orchestrator GUI will ensure that any concurrent updates on the same site or schema object cannot unintentionally overwrite each other. If you attempt to make changes to a site or template that was updated by another user since you opened it, the GUI will reject any subsequent changes you try to make and present a warning requesting you to refresh the object before making additional changes; refreshing the template will lose any edits you made up to that point and you will have to make those changes again:



However, the default REST API functionality was left unchanged in order to preserve backward compatibility with existing applications. In other words, while the UI is always enabled for this protection, you must explicitly enable it for your API calls for NDO to keep track of configuration changes.



Note When enabling this feature, note the following:

- This release supports detection of conflicting configuration changes for Site and Schema objects only.
- Only `PUT` and `PATCH` API calls support the version check feature.
- If you do not explicitly enable the version check parameter in your API calls, NDO will not track any updates internally. And as a result, any configuration updates can be potentially overwritten by both subsequent API calls or GUI users.

To enable the configuration version check, you can pass the `enableVersionCheck=true` parameter to the API call by appending it to the end of the API endpoint you are using, for example:

```
https://<mso-ip-address>/mso/api/v1/schemas/<schema-id>?enableVersionCheck=true
```

Example

We will use a simple example of updating the display name of a template in a schema to show how to use the version check attribute with `PUT` or `PATCH` calls.

First, you would `GET` the schema you want to modify, which will return the current latest version of the schema in the call's response:

```
{
  "id": "601acfed38000070a4ee9ec0",
  "displayName": "Schema1",
  "description": "",
  "templates": [
    {
      "name": "Template1",
      "displayName": "current name",
      [...]
    }
  ],
  "_updateVersion": 12,
  "sites": [...]
}
```

Then you can modify the schema in one of two ways appending `enableVersionCheck=true` to the request URL:



Note You must ensure that the value of the `"_updateVersion"` field in the payload is the same as the value you got in the original schema.

- Using the `PUT` API with the entire updated schema as payload:

```
PUT /v1/schemas/601acfed38000070a4ee9ec0?enableVersionCheck=true
```

```
{
  "id": "601acfed38000070a4ee9ec0",
  "displayName": "Schema1",
  "description": "",
  "templates": [
    {
      "name": "Template1",
      "displayName": "new name",
      [...]
    }
  ],
  "_updateVersion": 12,
  "sites": [...]
}
```

- Using any of the `PATCH` API operations to make a specific change to one of the objects in the schema:

```
PATCH /v1/schemas/601acfed38000070a4ee9ec0?enableVersionCheck=true
```

```
[
  {
    "op": "replace",
    "path": "/templates/Template1/displayName",
    "value": "new name",
    "_updateVersion": 12
  }
]
```

When the request is made, the API will increment the current schema version by 1 (from 12 to 13) and attempt to create the new version of the schema. If the new version does not yet exist, the operation will succeed and the schema will be updated; if another API call (with `enableVersionCheck` enabled) or the UI have modified the schema in the meantime, the operation fails and the API call will return the following response:

```
{
  "code": 400,
  "message": "Update failed, object version in the DB has changed, refresh your client
and retry"
}
```

Assigning Templates to Sites

This section describes how to assign a template to sites.

Before you begin

You must have the schema, template, and any objects you want to deploy to sites already created, as described in previous sections of this document.

Procedure

-
- Step 1** Navigate to the schema that contains one or more templates that you want to deploy.
- Step 2** In the left sidebar, select the template that you want to assign to sites.
- Step 3** In the **Template Summary** view, click **Actions** and choose **Add/ Remove Sites**.
The **Add Sites to <template-name>** window opens.
- Step 4** In the **Add Sites** window, check the checkbox next to the sites where you want to deploy the template.
Note that some sites may not be available for assignment depending on the type of the template you selected and the intersite connectivity between sites:
- If you are assigning a `Cloud Local` template, you will be able to assign it only to a single cloud site.
 - When assigning templates to multiple sites, the intersite connectivity between those sites must be established using BGP-EVPN protocol. If you select a site that has partial mesh connectivity, any site to which there is no intersite connectivity or intersite connectivity is established using BGP-IPv4 will be grayed out and unavailable for assignment.
- Step 5** Click **Ok**.
You deploy one template at a time, so you must associate the template with at least one site before you can deploy it.
-

Disassociating Template from Sites

You can choose to disassociate a template from a site without undeploying it. This allows you to preserve any configuration deployed to the site from NDO while removing the template-site association in the schema. The managed object and policy ownership is transferred from NDO to the site's controller.

Before you begin

- The template and its configuration must already be deployed to a site.

- The template must be deployed to a single site only and not stretched across sites.
- The objects defined in the template must not be deployed as shadow objects in other sites.

Procedure

-
- Step 1** Log in to your Nexus Dashboard Orchestrator GUI.
- Step 2** From the left navigation menu, select **Configure > Tenant Template**.
- Step 3** Under the application tab click on the schema that contains the template you want to disassociate.
- Step 4** In the Schema UI text view drop down menu, select the template under the specific site from which you want to disassociate it.
- Step 5** From the **Actions** menu, select **Disassociate Site**.
- Step 6** In the confirmation window, click **Confirm Action**.
-

Deploying Templates

This section describes how to deploy new or updated policies to the ACI fabrics.

Before you begin

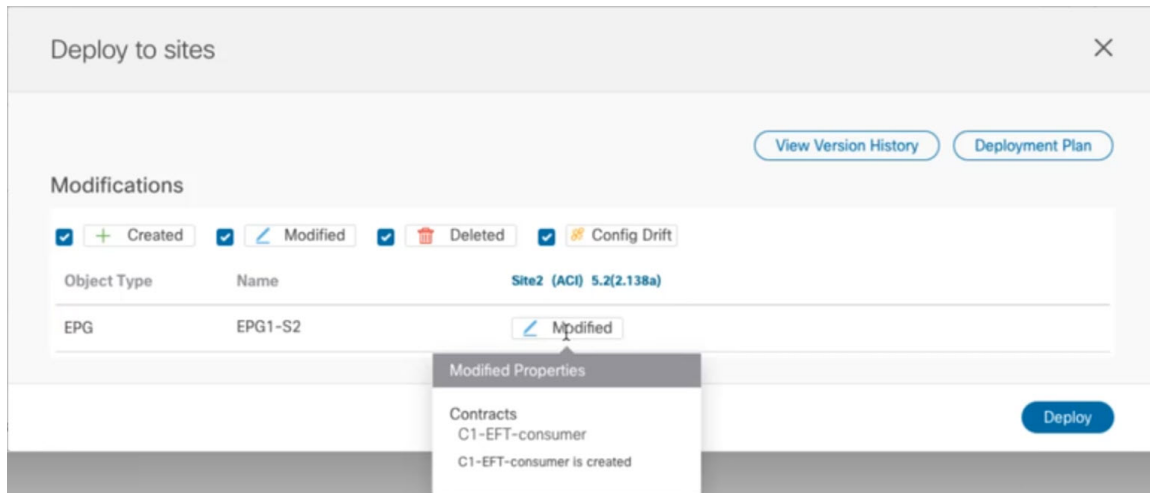
- You must have the schema, template, and any objects you want to deploy to sites already created and the templates assigned to one or more sites, as described in previous sections of this document.
- If template review and approval is enabled, the template must also be already approved by the required number of approvers as described in [Template Review and Approval, on page 33](#).
- Ensure that you understand the required deployment order and object dependencies that are described in [Schema and Template Design Considerations, on page 13](#).

Procedure

-
- Step 1** Navigate to the schema that contains the template that you want to deploy.
- Step 2** From the **View** dropdown menu, select the template you want to deploy.
- Step 3** In the template properties, click **Deploy Template**.
- The **Deploy to sites** window opens that shows the summary of the objects to be deployed.
- Step 4** If you have made changes to your template, review the **Deployment Plan** to verify the new configuration.
- If you have previously deployed this template but made no changes to it since, the **Deploy** summary will indicate that there are no changes and you can choose to re-deploy the entire template. In this case, you can skip this step.
- The **Deploy to sites** window will show you a summary of the configuration differences that will be deployed to sites. The following screenshots show a simple example of adding a `consumer` contract to an existing EPG (`EPG1-S2`) in `Site2`.

Note

In this case, only the difference in configuration is deployed to the sites. If you want to re-deploy the entire template, you must deploy once to sync the differences and then redeploy again to push the entire configuration as described in the previous paragraph.



You can also filter the view using the **Created**, **Modified**, and **Deleted** checkboxes for informational purposes, but keep in mind that all of the changes are still deployed when you click **Deploy**.

Here you can also choose to:

- **View Version History** shows the complete version history and incremental changes made between versions. Additional information about version history is available in [Viewing History and Comparing Previous Versions, on page 29](#).
- Check the **Deployment Plan** to see a visualization and an XML payload of the configuration that will be deployed from this template.

This feature provides better visibility into configuration changes that the Orchestrator will provision to the different fabrics that are part of your Multi-Site domain after you make a change to the template and deploy it to one or more sites.

Unlike earlier releases of the Nexus Dashboard Orchestrator, which still provided a list of specific changes made to the template and site configuration, the Deployment Plan provides full visibility into all the objects that the deployment of the template would provision across the different fabrics. For example, depending on what change you make, shadow objects may be created in multiple sites even if the specific change is applied to only a single site.

Note

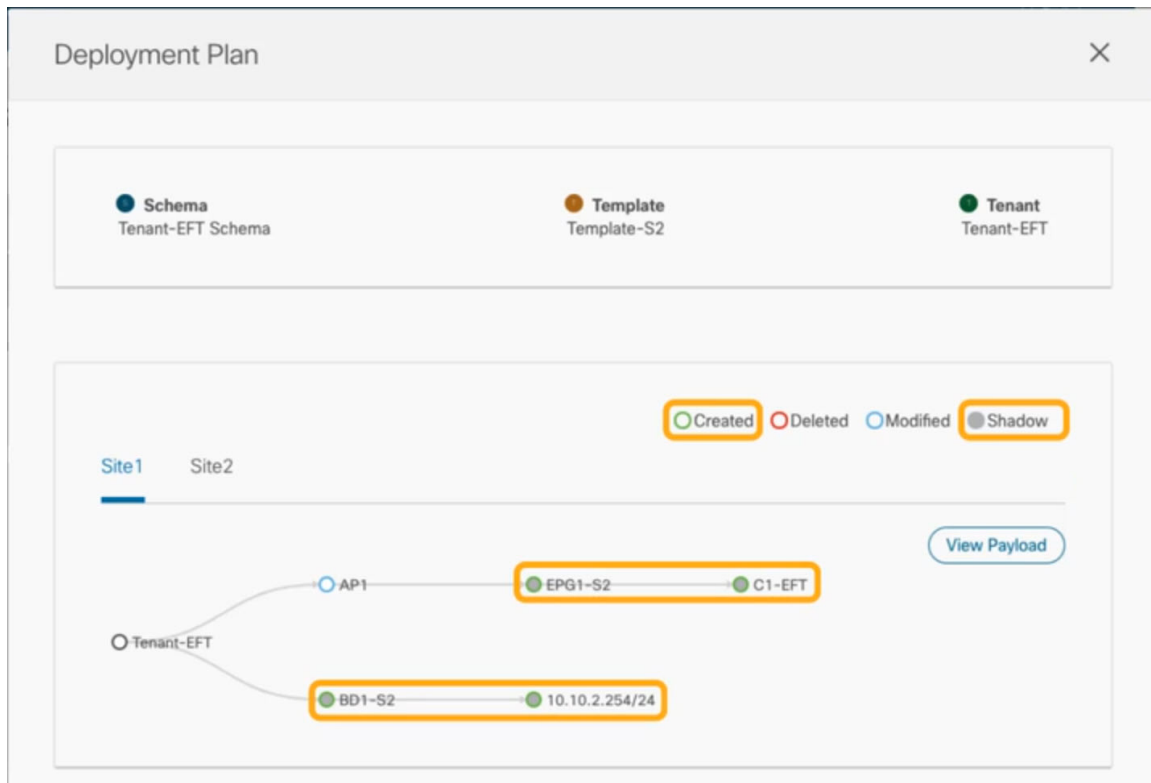
We recommend verifying your changes using the Deployment Plan as described in this step before deploying the template. The visual representation of the configuration changes can help you reduce potential errors from deploying unintended configuration changes.

- Click the **Deployment Plan** button.

Continuing with the same example shown in the previous step, where a consumer contract was added to an existing EPG in Site2, the Deployment Plan allows you to also see that there are additional changes to be deployed to Site1 as a result of the change to Site2.

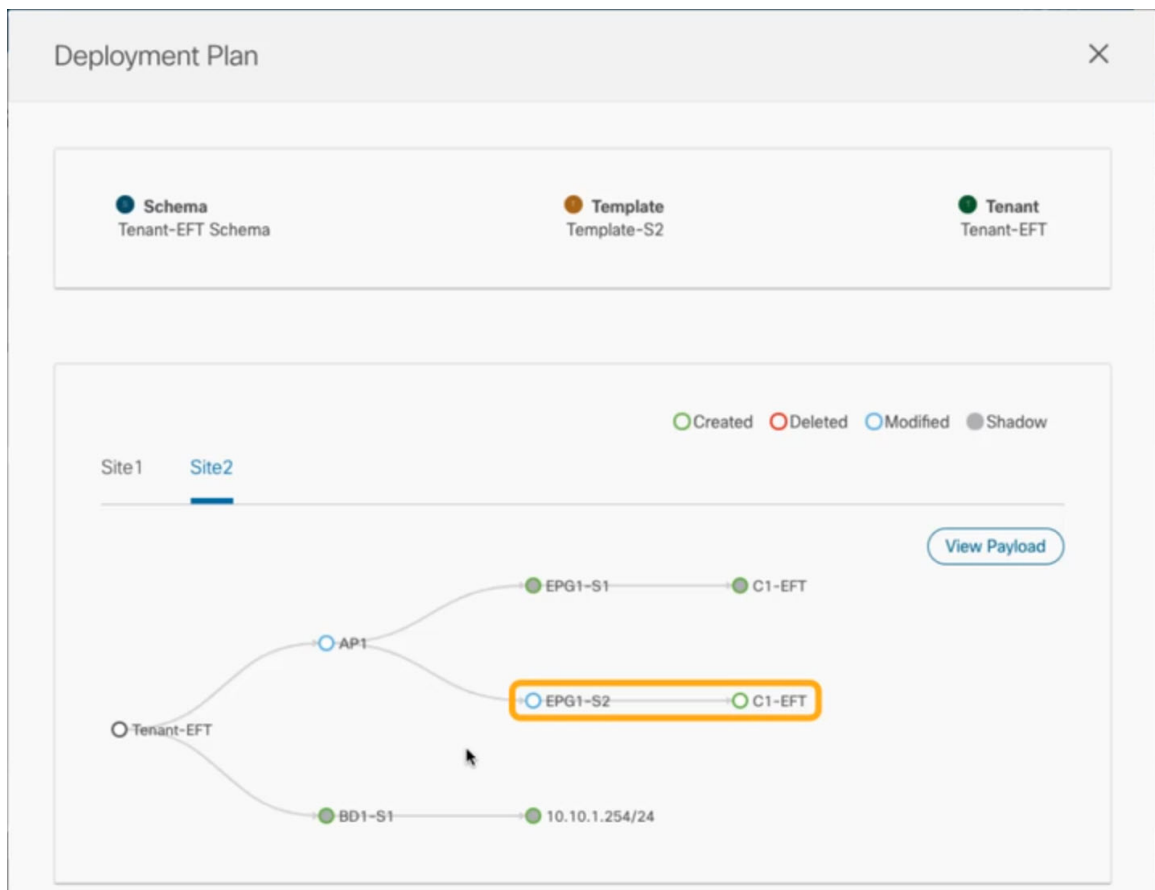
- Verify your changes in the first listed site.

Based on the highlighted legend, you can see that the Orchestrator will create the shadow objects in `Site1` that are required by the contract you added to an EPG in `Site2`.



- c) Repeat the previous substep to verify the changes in other sites

Here you can see the change you made explicitly to the EPG (`EPG1-S2`) in `Site2` when you assigned the contract (`C1-EFT`) to it, as well as the shadow objects for the EPG (`EPG1-S1`) in the other site, which is providing that contract.



- d) (Optional) Click **View Payload** to see the XML payload for each site.

In addition to the visual representation of the new and modified objects, you can also choose to **View Payload** for the changes in each site:



e) After you are done verifying the changes, click the x icon to close the **Deployment Plan** screen.

Step 5 In the **Deploy to sites** window, click **Deploy** to deploy the template.

Undeploying Templates

This section describes how to undeploy a template from a site. Undeploying a template removes all configurations defined in that template from a specific site where the template is deployed.



Note This action removes managed objects (MOs) and their properties from the site's controller and can disrupt the network connectivity that depends on those configurations.

Before you begin

- Ensure that you have not made any changes to the template since you last deployed it.

Undeploying a template that was modified since it was last deployed may create a configuration drift because the set of objects deployed with the template would be different than the set of objects you try to undeploy after making changes to the template.

- If you are undeploying a template that contains VRFs that are used in route leak configurations, the route leaks must be deleted before you can undeploy that template.

Procedure

- Step 1** Select the schema that contains the template you want to undeploy.
- Step 2** From the **View** dropdown, select the template you want to undeploy.
- Step 3** From the **Actions** menu, click **Undeploy template**.

Bulk Update for Template Objects

The bulk update feature allows you to update multiple properties on multiple different objects of the same type within a template at once. For example, you can enforce Infra EPG Isolation on two or more EPGs at the same time, instead of having to modify each object individually. When using this workflow, all selected objects must be of the same type, for example, you cannot choose to update an EPG and a BD simultaneously.

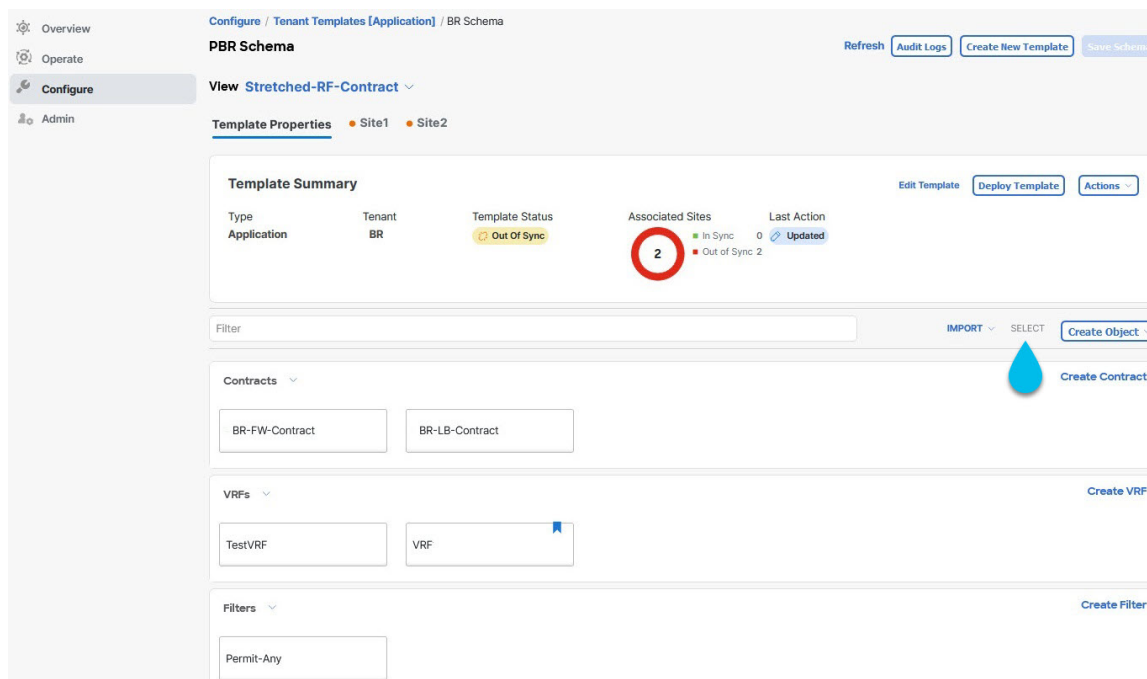
If the selected objects already have different property values configured on them, the update will overwrite those properties with the values you provide. This feature allows you update template-level object properties for on-premises; updating site-local properties and cloud properties are not supported.



Note This feature is supported for Application templates only with Cisco APIC and Cisco NDFC fabrics only; it is not supported for other template types or Cisco Cloud Network Controller sites.

Procedure

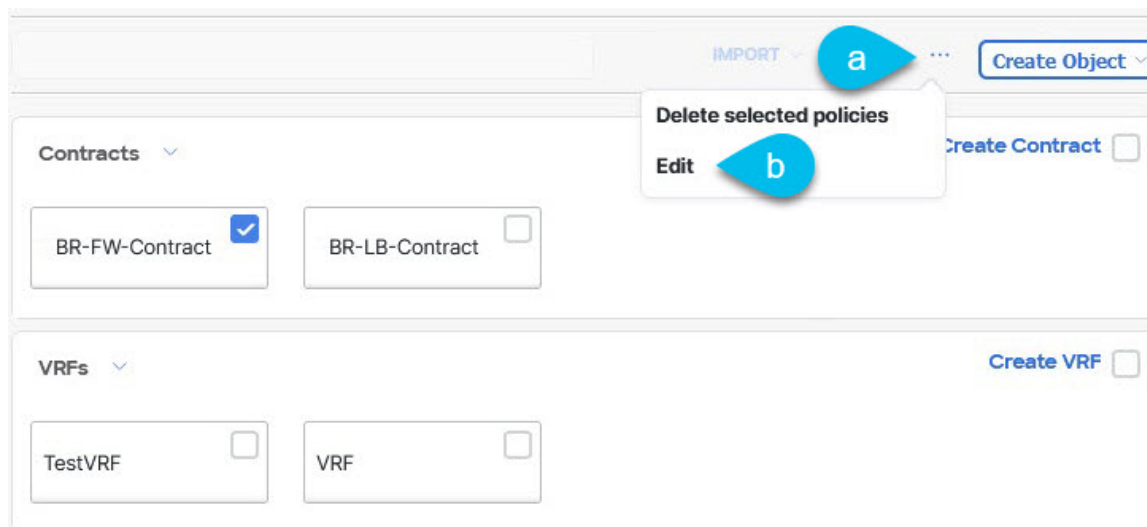
- Step 1** Navigate to the schema and template that contains the objects you want to update.
- Step 2** From the main pane, choose **Select**. It will allow you to choose multiple objects of same type.



Step 3 After selecting all the objects that you want to update.

- Choose “...” right next to the cancel option.
- From the dropdown Choose "Edit".

If you choose objects of different type, you won't see the Edit option in the dropdown.



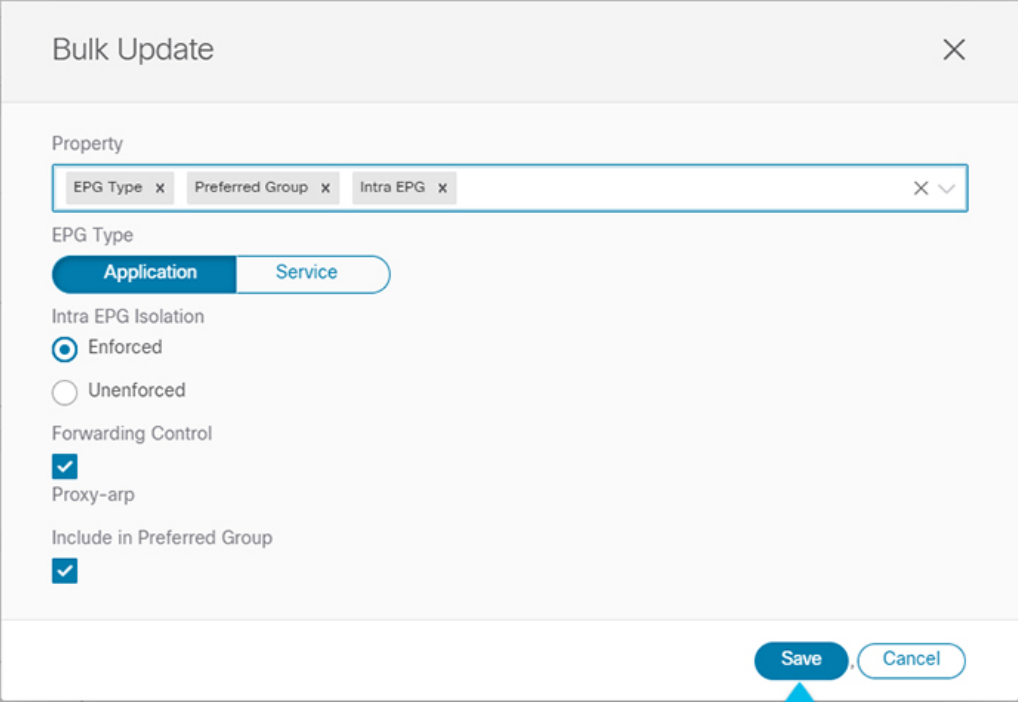
Step 4 After choosing “Edit”, a **Bulk Update** will show up. It will show you a subset of the properties for the objects you selected.

You can **Select Properties** the following properties based on the type of objects you selected.

- EPG:** Bridge Domain, Contract, EPG Type, Infra EPG, Preferred Group.

- b. **Contracts:** Scope, Filter Chain, QOS Level.
- c. **VRF:** IP Data-Plane Learning.
- d. **Bridge Domain:** Virtual Routing and Forwarding, L2 Stretch, L2 Unknown Unicast, Unknown Multicast Flooding, IPv6 Unknown Multicast Flooding, Multi Destination Flooding, DHCP Policies, Unicast Routing.
- e. **External EPG:** Contract, External EPG Type, Preferred Group.

Step 5 After selecting all the fields, you wish to update. Choose “Save” which will implement the bulk update you just made.



The image shows a 'Bulk Update' dialog box in the Cisco Nexus Dashboard Orchestrator. The dialog has a title bar with a close button (X). Below the title bar, there is a 'Property' section with a search bar containing 'EPG Type', 'Preferred Group', and 'Intra EPG'. Below this, there is an 'EPG Type' section with two tabs: 'Application' (selected) and 'Service'. Under the 'Application' tab, there are three settings: 'Intra EPG Isolation' with a radio button set to 'Enforced', 'Forwarding Control' with a checked checkbox, and 'Proxy-arp' with a checked checkbox. At the bottom of the dialog, there are 'Save' and 'Cancel' buttons. A blue teardrop cursor is pointing at the 'Save' button. The background of the dialog shows a list of objects with columns for 'epg2' and 'vrf 2'.

Step 6 As you save the updates, you can see the changes you've made.

The screenshot displays the Cisco Nexus Dashboard Orchestrator (NDO) interface. The top navigation bar includes the Cisco logo, 'Nexus Dashboard', and 'Orchestrator'. The left sidebar contains navigation links: Overview, Operate, Configure (selected), and Admin. The main content area is titled 'Configure / Tenant Templates [Application] / BR Schema' and 'PBR Schema'. It shows 'View Site2' and 'Template Properties' for 'Site2'. A 'Template Summary' table is visible, showing 'Type: Application', 'Tenant: PBR', and 'Template Status: In Sync'. Below this, there are sections for 'Application Profile AP1', 'EPGs' (listing EPG1-S1 and EP-SF-S2), and 'Bridge Domains' (listing BD1-S2 with a 'connected' status). The right panel, titled 'EPG1-S1', shows configuration details for this EPG, including 'Contracts', 'EPG Type' (Application/Service), 'Intra-EPG Contract', 'Properties' (On-Premises Properties), 'Bridge Domain' (BD1-S2), 'Subnets', and 'USeg EPG' (Unenforced).

Template Versioning

A new version of the template is created every time it is saved. From within the NDO UI, you can view the history of all configuration changes for any template along with information about who made the changes and when. You can also compare any of the previous versions to the current version.

New versions are created at the template level, not schema level, which allows you to configure, compare, and roll back each template individually.

Template versions are created and maintained according to the following rules:

- All template versions are either `Deployed` or `Intermediate`.
`Deployed`—versions of the template that have been deployed to sites.
`Intermediate`—versions of the template that have been modified and saved, but not deployed to sites.
- A maximum of 20 `Deployed` and 20 `Intermediate` versions per template can be stored at any given time.
- When a new `Intermediate` version is created that would exceed the 20 version limit, the earliest existing `Intermediate` version is deleted.

- When a template is deployed and a new `Deployed` version is created, all `Intermediate` versions are deleted. If the new `Deployed` version exceeds the 20 version limit, the earliest existing `Deployed` version is deleted.
- Tagging a version `Golden` does not affect the number of stored template versions.
- A template that is tagged `Golden` cannot be deleted.
You must untag the template first before you can delete it.
- When a template is modified and saved or deployed, any versions that exceed the 20 `Deployed` and 20 `Intermediate` scale are removed according to the above rules.
- When upgrading from a release prior to 4.0(1) to release 4.0(1) or later, only the latest versions of templates are preserved.

Tagging Templates

At any point you can choose to tag the current version of the template as "golden", for example for future references to indicate a version that was reviewed, approved, and deployed with a fully validated configuration.

Procedure

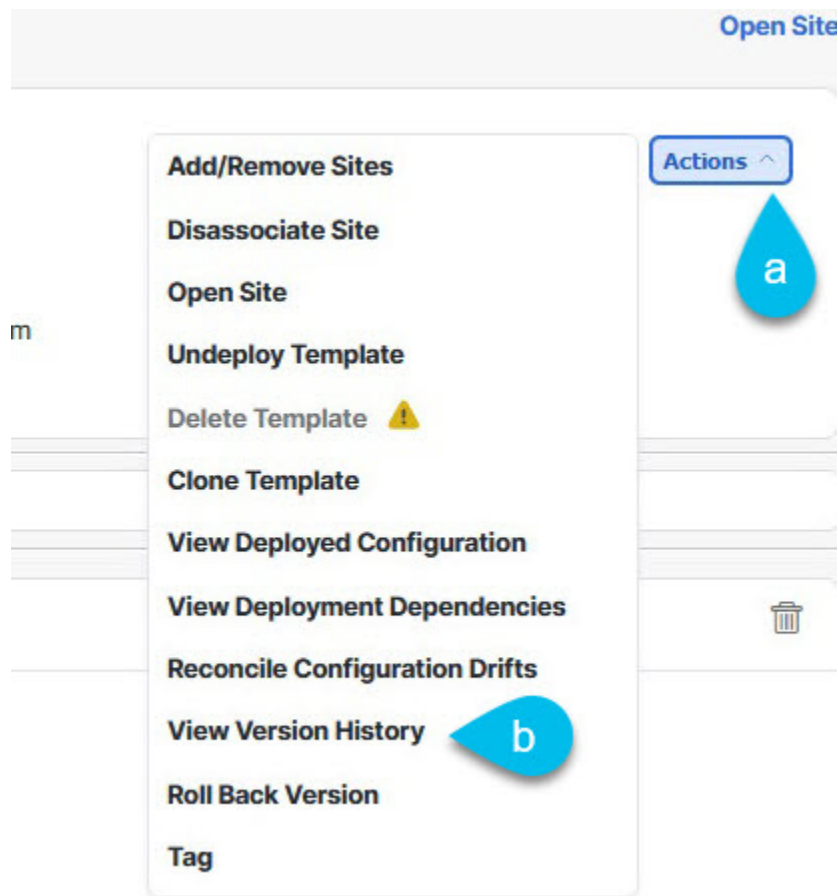
-
- Step 1** Log in to your Nexus Dashboard Orchestrator GUI.
 - Step 2** From the left navigation menu, select **Configure > Schemas**.
 - Step 3** Click the schema that contains the template you want to view.
 - Step 4** In the Schema view, select the template you want to review.
 - Step 5** From the template's actions (...) menu, select **Tag**.
- If the template is already tagged, the option will change to **Un-Tag** and allows you to remove the tag from the current version.
- Any version that was tagged will be indicated by a star icon in the template's version history screen.
-

Viewing History and Comparing Previous Versions

This section describes how to view previous versions for a template and compare them to the current version.

Procedure

-
- Step 1** Log in to your Nexus Dashboard Orchestrator GUI.
 - Step 2** From the left navigation menu, select **Configure > Tenant Template**.
 - Step 3** Click the schema that contains the template you want to view.
 - Step 4** In the Schema view, select the template you want to review.
 - Step 5** From the template's actions (...) menu, select **View Version History**.



Step 6 In the **Version History** window, make the appropriate selections.

Version History

General Information

Schema: PBR Schema Template: Site2 Tenant: PBR

Versions

☐ Golden Versions ☐ Deployed Versions ☐ Pre Reconciled Versions ☐ Post Reconciled Versions

Tag As Golden Delete Versions

Version 6 (Selected): 3 policies | 1 sites

Version 7 (Current): 2 policies | 1 sites

Click to expand

```

"externalEpgs": [
  {
    "externalEpgRef": "/schemas/Site2/externalEpgs/ExtEPG-S2",
    "l3outDn": "",
    "l3outRef": "c98d787f-fa8a-439"
  },
  {
    "contractRelationships": [],
    "description": "",
    "name": "ExtEPG-S2",
    "preferredGroup": false,
    "qosPriority": "unspecified",
    "selectors": [],
    "subnets": [
      {
        "scope": [
          "import-security"
        ]
      }
    ],
    "tagAnnotations": [],
    "vrfRef": "/schemas/64ddddd9btdddd-VRF-Contract/vrfs/VRF1"
  }
],

```

Click to expand

OK

- Enable the **Golden Versions** checkbox to filter the list of previous versions to display only the versions of this template that had been marked as **Golden**.
Tagging a template as "Golden" is described in [Tagging Templates, on page 29](#).
- Enable the **Deployed Versions** checkbox to filter the list of previous versions to display only the versions of this template that had been deployed to sites.
A new template version is created every time the template is changed and the schema is saved. You can choose to only show the versions of the template that were actually deployed to sites at some point.
- Click on a specific version to compare it to the current version.
The version you select is always compared to the current version of the template. Even if you filter the list using the **Golden Versions** or **Deployed Versions** filters, the current version will always be displayed even if it was never deployed or tagged as golden.
- Mouse over the **Edit** icon to see information about who created the version and when.
- Enable the **Pre Reconciled Versions** checkbox to filter the list of previous versions to display only the versions of this template that had been marked as **Reconciled**.
- Enable the **Post Reconciled Versions** checkbox to filter the list of previous versions to display only the versions of this template that had been marked as **Reconciled**.

Step 7 Click **OK** to close the version history window.

Reverting Template to Earlier Version

This section describes how to restore a previous version of the template. When reverting a template, the following rules apply:

- If the target version references objects that are no longer present, restore operation will not be allowed.
- If the target version references sites that are no longer managed by NDO, restore operation will not be allowed.
- If the current version is deployed to one or more sites to which the target version was not deployed, restore operation will not be allowed.

You must first undeploy the current version from those sites before reverting the template.

- If the target version was deployed to one or more sites to which the current version is not deployed, restore operation is allowed.

Procedure

Step 1 Log in to your Nexus Dashboard Orchestrator GUI.

Step 2 From the left navigation menu, select **Configure > Tenant Template**.

Step 3 Click the schema that contains the template you want to view.

Step 4 In the Schema view, select the template you want to review.

Step 5 From the **Actions (...)** menu, select **Rollback Versions**.

Step 6 In the **Rollback** window, select one of the earlier versions to which you want to restore.

You can filter the list of versions using the **Golden Versions**, **Pre Reconciled Versions**, **Post Reconciled Versions**, and **Deployed Versions** checkboxes.

When you select a version, you can compare the template configuration of that version to the current version of the template.

Step 7 Click **Restore** to restore the selected version.

When you restore a previous version, a new version of the template is created with the same configuration as the version you selected in the previous step.

For example, if the latest template version is 3 and you restore version 2, then version 4 is created that is identical to the version 2 configuration. You can verify the restore by browsing to the template version history and comparing the current latest version to the version you had selected during restore, which should be identical.

If template review and approval (change control) is disabled and your account has the correct privileges to deploy templates, you can deploy the version to which you reverted.

However, if change control is enabled, then:

- If you revert to a version that had been previously deployed and your account has the correct privileges to deploy templates, you can immediately deploy the template.

- If you revert to a version that had not been previously deployed or your account does not have the correct privileges to deploy templates, you will need to request template approval before the reverted version can be deployed.

Additional information about review and approval process is available in the [Template Review and Approval](#), on [page 33](#) sections.

Template Review and Approval

Template review and approval (change control) workflow which allows you to set up designated roles for template designers, reviewers and approvers, and template deployers to ensure that the configuration deployments go through a validation process.

From within the NDO UI, a template designer can request review on the template they create. Then reviewers can view the history of all configuration changes for the template along with information about who made the changes and when, at which point they can approve or deny the current version of the template. If the template configuration is denied, the template designer can make any required changes and re-request review; if the template is approved, it can be deployed to the sites by a user with `Deployer` role. Finally, the deployers themselves can deny deployment of an approved template and restart the review process from the beginning.

The workflow is done at the template level, not schema level, which allows you to configure, review, and approve each template individually.

Enabling Template Approval Requirement

Before you can use the review and approval workflow for template configuration and deployment, you must enable the feature in the Nexus Dashboard Orchestrator's system settings.

Procedure

-
- | | |
|---------------|--|
| Step 1 | Log in to your Nexus Dashboard Orchestrator GUI. |
| Step 2 | From the left navigation menu, select Admin > System Configuration . |
| Step 3 | On the Change Control tile, click the Edit icon. |
| Step 4 | In the Change Control window, select Enabled to enable the feature. |
| Step 5 | In the Approvers field, enter the number of unique approvals required before the templates can be deployed. |
| Step 6 | Click Save to save the changes. |
-

Create Users with Required Roles

Before you can use the review and approval workflow for template configuration and deployment, you must create the users with the necessary privileges in the Nexus Dashboard where the NDO service is deployed.

Procedure

Step 1 Log in to your Nexus Dashboard GUI.

Users cannot be created or edited in the NDO GUI, you must log in directly to the Nexus Dashboard cluster where the service is deployed.

Step 2 From the left navigation menu, select **Admin Console > Admin > Users**.

Step 3 Create the required users.

The workflow depends on three distinct user roles: template designer, approver, and deployer. You can assign each role to a different user or combine the roles for the same user; users with `admin` privileges can perform all 3 actions.

There is no `Designer` role predefined on Nexus Dashboard, so the designer duties are assigned to any `Tenant Manager` or `Site Manager` user with write privileges, in addition to the default `Admin` user role:

- `Tenant Manager` should be used when the designer needs to make changes to templates associated only to a specific tenant (or a subset of tenants). In this case, the user should be mapped to the specific tenants.
- `Site Manager` should be used when the designer needs to make changes to templates that belong to different tenants.

In contrast to `Designer` role, there are pre-defined `Approver` and `Deployer` roles on the Nexus Dashboard that can be associated to the users. `Approver` and `Deployer` roles are not bound to specific tenant(s) by design. However, when creating a user role with both designer and approver (or designer and deployer) rights, follow the same guidelines as listed above.

Detailed information about configuring users and their privileges for local or remote Nexus Dashboard users is described in the [Nexus Dashboard User Guide](#).

You must have at least as many unique users with `Approver` role as the minimum number of approvals required, which you configured in [Enabling Template Approval Requirement, on page 33](#).

Note

If you disable the **Change Control Workflow** feature, any `Approver` and `Deployer` users will have read-only access to the Nexus Dashboard Orchestrator.

Requesting Template Review and Approval

This section describes how to request template review and approval.

Before you begin

You must have:

- Enabled the global settings for approval requirement, as described in [Enabling Template Approval Requirement, on page 33](#).
- Created or updated users in Nexus Dashboard with `approver` and `deployer` roles, as described in [Create Users with Required Roles, on page 33](#).
- Created a template with one or more policy configurations and assigned it to one or more sites.

Procedure

-
- Step 1** Log in to your Nexus Dashboard Orchestrator GUI as a user with `Tenant Manager`, `Site Manager`, or `Administrator` role.
- Step 2** If you assigned the `Tenant Manager` role, associate the user with the tenants.
- If you used `Site Manager` or `Administrator` roles, skip this step.
- If you assign the `Tenant Manager` role, you must also associate the user to the specific tenants they will manage.
- From the left navigation menu, select **Operate > Tenants**.
 - Select the tenant which the user will manage.
 - Check the box next to the designer user you created in Nexus Dashboard.
 - Repeat this step for all other tenants the user will manage.
- Step 3** From the left navigation menu, select **Configure > Tenant Template**.
- Step 4** Click the schema that contains the template for which you want to request approval.
- Step 5** In the schema view, select the template.
- Step 6** In the main pane, click **Send for Approval**.
- Note that the **Send for Approval** button will not be available in the following cases:
- The global change control option is not enabled
 - The template has no policy configurations or is not assigned to any sites
 - Your user does not have the right permissions to edit templates
 - The template has already been sent for approval
 - The template was denied by the approver user
-

Reviewing and Approving Templates

This section describes how to request template review and approval.

Before you begin

You must have:

- Enabled the global settings for approval requirement, as described in [Enabling Template Approval Requirement, on page 33](#).
- Created or updated users in Nexus Dashboard with `approver` and `deployer` roles, as described in [Create Users with Required Roles, on page 33](#).
- Created a template with one or more policy configurations and assigned it to one or more sites.
- Had the template approval requested by a schema editor, as described in [Requesting Template Review and Approval, on page 34](#).

Procedure

Step 1 Log in to your Nexus Dashboard Orchestrator GUI as a user with `Approver` or `admin` role.

Step 2 From the left navigation menu, select **Configure > Tenant Template**.

Step 3 Click the schema that contains the template you want to review and approve.

Step 4 In the schema view, select the template.

Step 5 In the main pane, click **Approve**.

If you have already approved or denied the template, you will not see the option until the template designer makes changes and re-sends the template for review again.

Step 6 In the **Approving template** window, review the template and click **Approve**.

The approval screen will display all the changes which the template would deploy to the sites.

You can click **View Version History** to view the complete version history and incremental changes made between versions. Additional information about version history is available in [Viewing History and Comparing Previous Versions, on page 29](#).

You can also click **Deployment Plan** to see a visualization and an XML of the configuration that would be deployed from this template. The functionality of the "Deployment Plan" view is similar to the "Deployed View" for already-deployed templates, which is described in [Viewing Currently Deployed Configuration, on page 43](#).

Configuration Drifts

Occasionally, you may run into a situation where the configuration actually deployed to an APIC domain is different from the configuration defined for that domain in the Nexus Dashboard Orchestrator (NDO). These configuration discrepancies are referred to as **Configuration Drifts** and are indicated by an `Out of Sync` warning next to the site name in the template view page as shown in the following figure:

**Note**

- In certain cases, the template-level notification of a configuration drift shown above may not trigger if the configuration of properties of objects managed by NDO is modified directly in the site's controller. Specifically, addition (and subsequent removal) of the following properties do not show drift notification on NDO:

- Subnets for EPGs or BDs
- Bridge Domain DHCP Labels
- Static Ports configuration for EPGs
- Contract Relationships between EPGs

In these cases, you can still check for configuration drift by manually running drift reconciliation workflow as described in [Reconciling Configuration Drifts in Application Templates, on page 38](#).

- When you deploy a template from NDO, drift notification for objects in that template is disabled for 60 seconds.

Configuration Drift Causes

Configuration drifts can manifest due to a number of different reasons. Specific steps required to resolve a configuration drift depends on its cause. Most common scenarios and their resolutions are outlined below:

- **Configuration is modified in NDO**—when you modify a template in NDO GUI, it will show as configuration drift until you deploy the changes to the sites.

To resolve this type of configuration drift, either deploy the template to apply the changes to the sites or revert the changes in the schema.

- **Configuration is modified directly in the site's APIC**—while the objects deployed from NDO are indicated by a warning icon and text in the site's APIC, an admin user can still make changes to them causing the configuration drift.



Note

Every time an object is modified on APIC, APIC sends a notification to Nexus Dashboard Orchestrator. On receiving the notification, Nexus Dashboard Orchestrator starts a 30 second timer (waiting for further notifications to arrive) and at the expiration of such timer then makes API calls to APIC to retrieve detailed information about the changes made all the objects for which it received a notification. This allows the Nexus Dashboard Orchestrator to display the drift symbol on the UI for all the templates where those objects are defined. The only exception to this behavior is when Nexus Dashboard Orchestrator deploys the configuration for all (or of a subset of) the objects defined in a specific template. In that case, for 60 seconds Nexus Dashboard Orchestrator would ignore any notification received from APIC relative to those specific objects and, as a consequence, it would not be able to display the drift symbol on the UI.

- **NDO configuration is restored from backup**—restoring configuration from a backup in NDO restores only the objects and their state as they were when the backup was created, it does not automatically re-deploy the restored configuration. As such, if there were changes made to the configuration and deployed on APIC since the backup was created, restoring the backup would create a configuration drift.
- **NDO configuration is restored from a backup created on an older release**—if the newer release added support for object properties which were not supported by the earlier release, these properties may cause configuration drift warning. Typically, this happens if the new properties were modified directly in the site's APIC GUI and the values are different from the defaults assumed by the Nexus Dashboard Orchestrator
- **NDO is upgraded from an earlier release**—this scenario is similar to the previous one where if new object properties are added in the new release, existing configuration may indicate a drift.

We recommend that you check for configuration drifts and, if necessary, run the "Reconcile Drift" workflow for templates, to have more visibility into the causes of the drift and be able to reconcile it. This recommendation applies to all the drift scenarios previously described in this section.

Reconciling Configuration Drifts in Application Templates

You can use the drift reconciliation workflow to compare the template's configuration as it is defined in Nexus Dashboard Orchestrator to the configuration rendered in the APIC controllers of the sites that are part of your Multi-Site domain. This provides better visibility into changes that may have been made in Nexus Dashboard Orchestrator or in APIC directly and give you an opportunity to correctly resolve those drifts.



Note Configuration drift reconciliation is supported only for Application templates.

The templates are updated and saved only after you choose **Save** or **Deploy** at the end of the reconciliation workflow. If at any time during the workflow you want to undo the changes you already chose, you can close and re-open the schema to restore the original configurations. You can then re-run the workflow from the start.

Procedure

Step 1 Navigate to the schema that contains the template you want to check for configuration drifts.

Step 2 From the template's **Actions** menu, select **Reconcile Configuration Drift**.

The screenshot shows the Cisco Nexus Dashboard Orchestrator interface. The top navigation bar includes 'Nexus Dashboard' and 'Orchestrator'. The left sidebar has tabs for 'Overview', 'Operate', 'Configure', and 'Admin'. The main content area is titled 'Configure / Tenant Templates [Application] / BR Schema' and 'PBR Schema'. It shows 'Template Properties' for 'Site2' and a 'Template Summary' table with columns: Type (Application), Tenant (BR), Template Status (In Sync), Associated Sites (1 In Sync, 0 Out of Sync), and Last Action (Deployment Successful, Last Deployed: Aug3, 2023 04:53 pm). Below the summary is a 'Filter' section with 'Application Profile AP1', 'EPGs' (EPG-S2, EPG1-S), and 'Bridge Domains' (BD-S). On the right, an 'Actions' menu is open, listing options like 'Add/Remove Sites', 'Disassociate Site', 'Open Site', 'Undeploy Template', 'Delete Template', 'Clone Template', 'View Deployed Configuration', 'View Deployment Dependencies', 'Reconcile Configuration Drifts' (highlighted), 'View Version History', 'Roll Back Version', and 'Tag'.

The **Drift Reconciliation** wizard opens.

Step 3 In the **Drift Reconciliation** screen, compare the template-level configurations for each site and choose the one you want.

Drift Reconciliation for Site1

General Information

Schema	Template	Tenant
Common Schema	Site1	common

1 **Template Properties** 2 **Site Specific Properties**

Template level properties are common across all sites associated to the template. Please select either NDO configuration or one of the sites configuration to apply.

Let's start by selecting a site

APIC Site1 **a**

Great, now choose template level properties between Site1, and NDO

b APIC Site1 ☒ NDO Current Settings ☐

Click to collapse

```

{
  "anps": [],
  "bds": [],
  "contracts": [],
  "description": "",
  "displayName": "Site1",
  "externalEggs": [
    {
      "contractRelationships": [
        {
          "contractRef": "/schemas/C1-Common",
          "relationshipType": "nnc
    
```

```

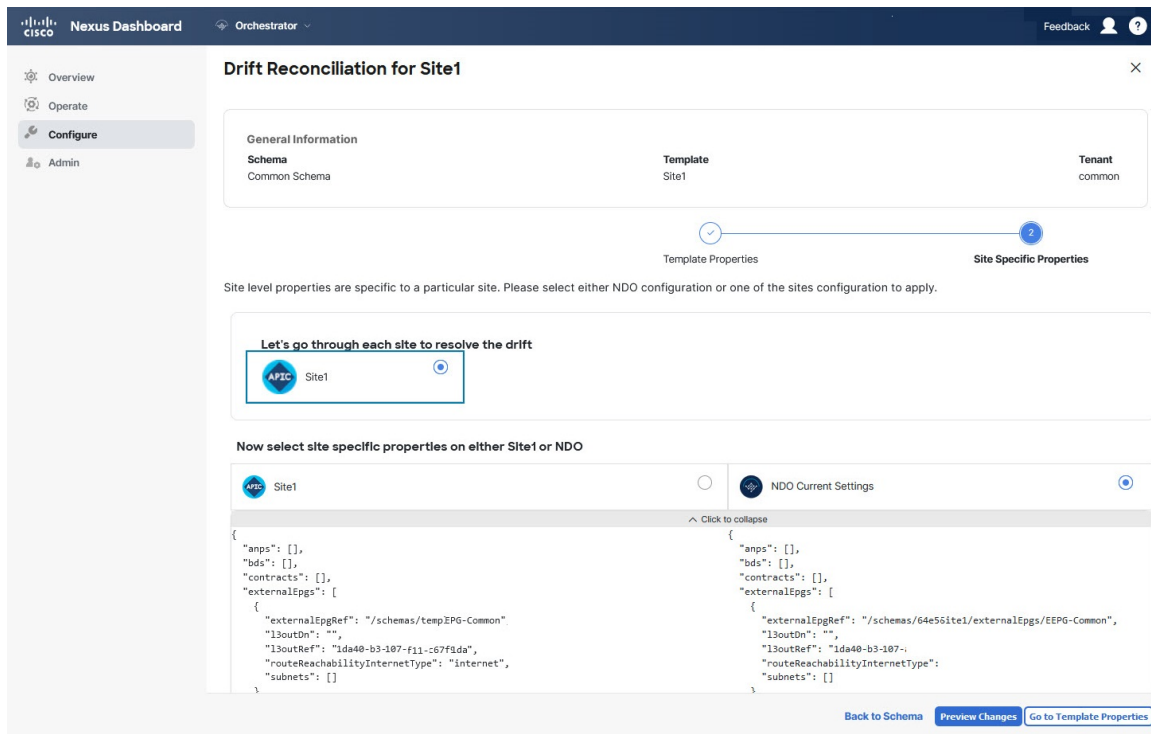
{
  "anps": [],
  "bds": [],
  "contracts": [],
  "description": "",
  "displayName": "Site1",
  "externalEggs": [
    {
      "contractRelationships": [
        {
          "contractRef": "/bn",
          "relationshipType": "nnc
    
```

c

[Back to Schema](#) [Go to Site Specific Properties](#)

Template-level properties are common across all sites associated to the template. You can compare the template level properties defined on Nexus Dashboard Orchestrator with the configuration rendered in each site and decide what should become the new configuration in the Nexus Dashboard Orchestrator template. Selecting the site configuration will modify those properties in the existing Nexus Dashboard Orchestrator template, whereas selecting the Nexus Dashboard Orchestrator configuration will keep the existing Nexus Dashboard Orchestrator template settings as is.

Step 4 Click **Go to Site Specific Properties** to switch to site-level configuration.



You can choose a site to compare that specific site's configuration. Unlike template-level configurations, you can choose either the Nexus Dashboard Orchestrator-defined or actual existing configurations for each site individually to be retained as the template's site-local properties for that site.

Even though in most scenarios you will make the same choice for both template-level and site-level configuration, the drift reconciliation wizard allows you to choose the configuration defined in the site's controller at the "Template Properties" level and the configuration defined in Nexus Dashboard Orchestrator at the "Site Local Properties" level or vice versa.

Step 5 Click **Preview Changes** to verify your choices.

The preview will display full template configuration adjusted based on the choices picked in the **Drift Reconciliation** wizard. You can then click **Deploy to sites** to deploy the configuration and reconcile the drift for that template.

Cloning Templates

This section describes how to create a copy of an existing template using the "Clone Template" feature in the Schema view.

Procedure

- Step 1** Log in to your Nexus Dashboard Orchestrator GUI.
- Step 2** From the left navigation menu, select **Configure > Tenant Template**.
- Step 3** Click the schema that contains the template you want to clone.

Step 4 On the **View** menu, select a template to open it.

Step 5 From the **Actions** menu, select **Clone Template**.

Step 6 Provide the clone destination details.

- a) From the **Destination Schema** dropdown, select the name of the Schema where you want to create the clone of the template.

You can select the same or a different schema to contain the clone of this template. If you want to clone the template into a schema that doesn't already exist, you can create a new schema by typing in the name of the schema and selecting **Create <schema-name>** option from the dropdown.

Note

When cloning across different schemas, the template must not have any objects that reference objects in other templates.

- b) In the **Cloned Template Name** field, provide the name for the new template.
c) Click **Save** to create the clone.

A new template will be created in the destination schema, with the tenant you selected and the exact same object and policy configurations as the original template.

If the destination schema you chose was the same schema as the source template, the schema view will reload and the new template will be displayed in the left sidebar. If you chose a different schema, you can navigate to that schema to see and edit the new template.

Note that while the template objects and configurations are copied, the site association is not preserved and you will need to re-associate the cloned template with any sites where you want to deploy it. Similarly, you will need to provide any site-specific configurations for the template objects after you associate it with the sites.

Migrating Objects Between Templates

This section describes how to move objects between templates or schemas. When moving one or more objects, the following restrictions apply:

- Only EPG and Bridge Domain (BD) objects can be moved between templates.
- Migrating objects to or from Cloud Network Controller sites is not supported.
You can migrate objects between on-premises sites only.
- The source and destination templates can be in the same schema or in different schemas, but the templates must be assigned to the same tenant.
- The destination template must have been created and assigned to at least one site.
- If the destination template is not deployed and has no other objects, the template will be automatically deployed after the objects are migrated.
- Once you initiate one object migration, you cannot perform another migration that involves the same source or target template. The migration is completed when the templates have been deployed to sites.

Procedure

-
- Step 1** Log in to your Nexus Dashboard Orchestrator GUI.
- Step 2** From the left navigation menu, select **Configure > Tenant Template > Applications** to Schemas view.
- Step 3** Click the schema that contains the objects you want to migrate.
- Step 4** In the Schema view, select the Template that contains the objects you want to migrate.
- Step 5** In the top right of the main pane, click **Select**.
This allows you to select one or more objects to migrate.
- Step 6** Click each object that you want to migrate.
Selected objects will display a check mark in their top right corner.
- Step 7** In the top right of the main pane, click the actions (...) icon and choose **Migrate Objects**.
- Step 8** In the **Migrate Objects** window, select the destination Schema and Template where you want to move the objects.
Only the templates with at least one site attached to them will appear in the list. If you don't see your target Template in the dropdown list, cancel the wizard and assign that template to at least one site.
- Step 9** Click **OK** and then **YES** to confirm that you want to move the objects.
The objects will be migrated from the source template to the destination template that you selected. When you deploy your configuration, the objects will be removed from any site where the source Template is deployed and added to the site where the destination template is deployed.
- Step 10** After the migration is completed, redeploy both, the source and the destination, templates.
If the destination template is not deployed and has no other objects, the template will be automatically deployed after the objects are migrated, so you can skip this step.
-

Viewing Currently Deployed Configuration

You can view all objects currently deployed to sites from a specific template. Even though any given template can be deployed, undeployed, updated, and re-deployed any number of times, this feature will show only the final state that resulted from all of those actions. For example, if `Template1` contains only `VRF1` object and is deployed to `Site1`, the API will return only `VRF1` for the template; if you then add `BD1` and redeploy, the API will return both objects, `BD1` and `VRF1`, from this point on.

This information comes from the Orchestrator database, so it does not account for any potential configuration drifts caused by changes done directly in the site's controller.

Procedure

-
- Step 1** Log in to your Nexus Dashboard Orchestrator GUI.
- Step 2** From the left navigation menu, select **Configure > Tenant Template**.

Step 3 Click the schema that contains the template you want to view.

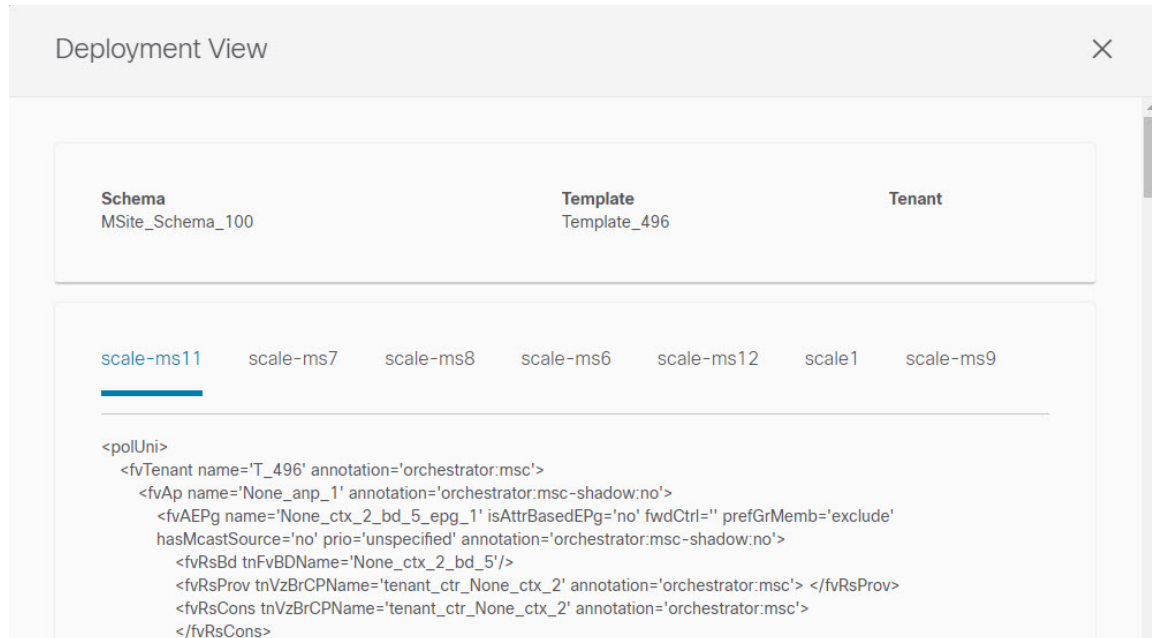
Step 4 In the left sidebar, select the template.

Step 5 Open the **View Deployed Configuration** for the template.

- Click the **Actions** menu next to the template's name.
- Click **Deployed View**.

Step 6 In the **Deployed View** screen, select the site for which you want to view the information.

You will see a graphical representation of the template configuration comparison between what's already deployed to the site and what's defined in the template..



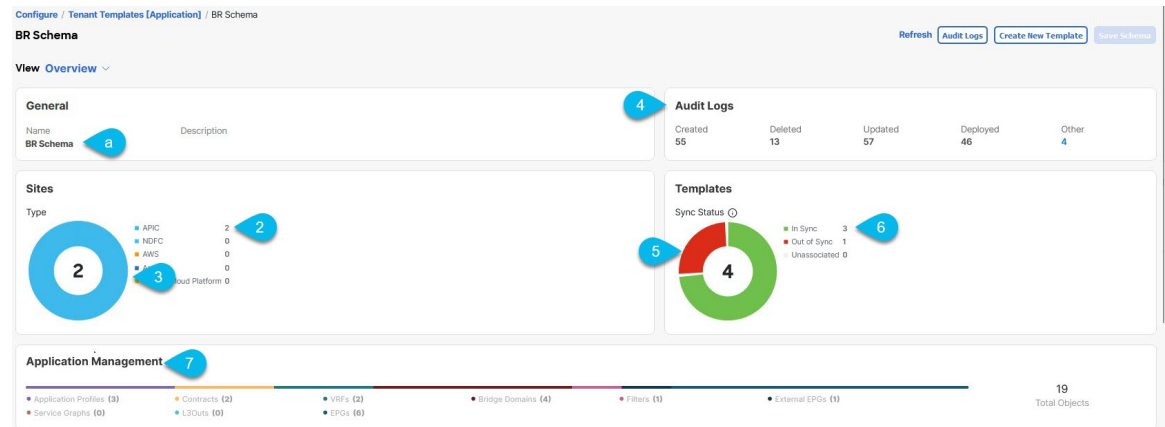
- The color-coded legend indicates which objects would be created, deleted, or modified if you were to deploy the template at this time.

If the latest version of the template is already deployed, the view will not contain any color-coded objects and will simply display the currently deployed configuration.

- You can click on a site name to show configuration for that specific site.
- You can click **View Payload** to see the XML/JSON config of all the objects that are deployed to the selected site.

Schema Overview and Deployment Visualizer

When you open a schema with one or more objects defined and deployed to one or more ACI fabrics, the schema **Overview** page will provide you with a summary of the deployment.

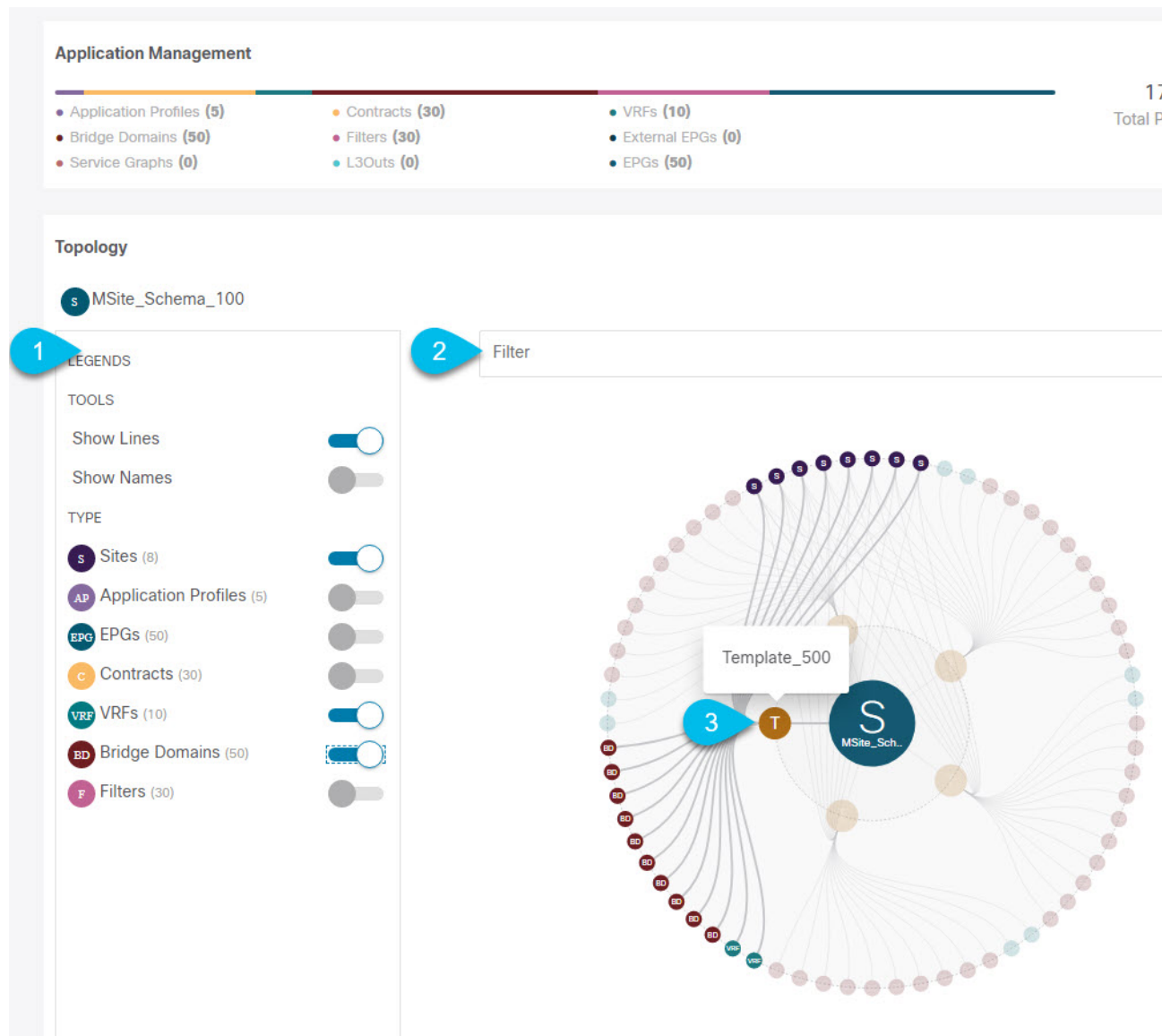
Figure 6: Schema Overview

The following details are provided on this page:

1. **General**—Provides general information of the schema, such the name and description.
2. **Audit Log**—Provides audit log summary of the actions performed on the schema.
3. **Sites > Health**—Provides the number of sites associated with the templates in this schema sorted by the site's health status.
Type—Provides the number of sites associated with the templates in this schema sorted by the site's type.
4. **Template > Sync Status**—Provides the number of templates in this schema that are associated with one or more sites and their deployment status.
Site Associations > Consistency—Provides the number of consistency checks performed on the deployed templates and their status.
5. **Application Management**—Provides a summary of individual objects contained by the templates in this schema.

The **Topology** tile allows you to create a topology visualizer by selecting one or more objects to be displayed by the diagram as shown in the following figure.

Figure 7: Deployment Visualizer



1. **Legend**—Allows you to choose which policy objects to display in the topology diagram below.
2. **Filter**—Allows you to filter the displayed objects based on their names.
3. **Topology Diagram**—Provides visual representation of the policies configured in all of the Schema's templates that are assigned to sites.

You can choose which objects you want to display using the **Configuration Options** above.

You can also mouse over an objects to highlight all of its dependencies.

Finally, you can click on any object in the diagram to zoom in to see only its relationships with other objects. For example, clicking a Template will display all objects within that specific template only.



CHAPTER 4

Tenants and Tenant Policies Templates

- [Tenants Overview, on page 47](#)
- [Creating New Tenants, on page 48](#)
- [Importing Existing Tenants, on page 49](#)
- [Creating Tenant Policy Templates, on page 50](#)

Tenants Overview

A tenant is a logical container for application policies that enable an administrator to exercise domain-based access control. A tenant represents a unit of isolation from a policy perspective, but it does not represent a private network. Tenants can represent a customer in a service provider setting, an organization or domain in an enterprise setting, or just a convenient grouping of policies.

Three default tenants are pre-configured for you:

- `common`—A special tenant with the purpose of providing "common" services to other tenants in ACI fabrics. Global reuse is a core principle in the common tenant. Some examples of common services include shared L3Outs, DNS, DHCP, Active Directory, and shared private networks or bridge domains.
- `dcnm-default-tn`—A special tenant with the purpose of providing configuration for Cisco NDFC fabrics.
- `infra`—The Infrastructure tenant that is used for all internal fabric communications, such as tunnels and policy deployment. This includes switch to switch and switch to APIC communications. The `infra` tenant does not get exposed to the user space (tenants) and it has its own private network space and bridge domains. Fabric discovery, image management, and DHCP for fabric functions are all handled within this tenant.

When using Nexus Dashboard Orchestrator to manage Cisco NDFC fabrics, you will always use the default `dcnm-default-tn` tenant.



Note Nexus Dashboard Orchestrator cannot manage the APIC's `mgmt` tenant, so importing the tenant from APIC or creating a new tenant called `mgmt` in NDO it is not allowed.

To manage tenants, you must have either `Power User` or `Site and Tenant Manager` read-write role.

Tenant Policies Templates

Release 4.0(1) adds Tenant Policies templates, which allow you to configure the following tenant-wide policies:

- Route Policies for Multicast
- Route Map Policies for Route Control
- Custom QoS Policies
- DHCP Relay Policies
- DHCP Option Policies
- IGMP Interface Policies
- IGMP Snooping Policies
- MLD Snooping Policies

For additional information, see [Creating Tenant Policy Templates, on page 50](#).

Creating New Tenants

This section describes how to add a new tenant using the Cisco Nexus Dashboard Orchestrator GUI. If you want to import one or more existing tenants from your fabrics, follow the steps that are described in [Importing Existing Tenants, on page 49](#) instead.

Before you begin

You must have a user with either `Power User` or `Site Manager` read/write role to create and manage tenants.

Procedure

Step 1 Log in to your Cisco Nexus Dashboard and open the Cisco Nexus Dashboard Orchestrator service.

Step 2 Create a new tenant.

- a) From the left navigation pane, choose **Operate > Tenants**.
- b) In the top right of the main pane, click **Create Tenant**.

The **Create Tenant** screen opens.

Step 3 Provide tenant details.

- a) Provide the **Display Name** and optional **Description**.

The tenant's **Display Name** is used throughout the Orchestrator's GUI whenever the tenant is shown. However, due to object naming requirements on the APIC, any invalid characters are removed and the resulting **Internal Name** is used when pushing the tenant to sites. The **Internal Name** that will be used when creating the tenant is displayed below the **Display Name** text box.

Note

You can change the **Display Name** of the tenant at any time, but the **Internal Name** cannot be changed after the tenant is created.

- b) In the **Associated Sites** section, check all the sites that you want to associate with this tenant.

Only the selected sites are available for any templates using this tenant.

- c) (Optional) For each selected site, click the **Edit** button next to its name and choose one or more security domains.

A restricted security domain allows a fabric administrator to prevent a group of users, such as Tenant A, from viewing or modifying any objects that are created by a group of users in a different security domain, such as Tenant B, when users in both groups have the same assigned privileges. For example, a tenant administrator in Tenant A's restricted security domain will not be able to see policies, profiles, or users configured in Tenant B's security domain. Unless Tenant B's security domain is also restricted, Tenant B can see policies, profiles, or users configured in Tenant A.

Note

A user will always have read-only visibility to system-created configurations for which the user has proper privileges. A user in a restricted security domain can be given a broad level of privileges within that domain without the concern that the user could inadvertently affect another tenant's physical environment.

Security domains are created using the APIC GUI and can be assigned to various APIC policies and user accounts to control their access. For more information, see the *Cisco APIC Basic Configuration Guide*.

- d) In the **Associated Users** section, select the Cisco Nexus Dashboard Orchestrator users that are allowed to access the tenant.

Only the selected users are able to use this tenant when creating templates.

Step 4 Click **Save** to finish adding the tenant.

Importing Existing Tenants

This section describes how to import one or more existing tenants. If you want to create a new tenant using Cisco Nexus Dashboard Orchestrator, follow the steps that are described in [Creating New Tenants](#), on page 48 instead.

Before you begin

You must have a user with either `Power User` or `Site Manager` read/write role to create and manage tenants.

Procedure

Step 1 Log in to your Cisco Nexus Dashboard and open the Cisco Nexus Dashboard Orchestrator service.

Step 2 In the left navigation menu, click **Operate > Sites**.

Step 3 Locate the site from which you want to import the tenants, click the three dots to get actions (...) menu, and choose **Import Tenants**.

You can import tenants from one site at a time.

Step 4 In the **Import Tenants** dialog, select one or more tenants to import and click **Ok**.

The selected tenants will be imported into the Cisco Nexus Dashboard Orchestrator and show in the **Operate > Tenants** page.

Step 5 Repeat these steps to import tenants from any other sites.

Creating Tenant Policy Templates

This section describes how to create one or more tenant policy templates. Tenant policy templates allow you to create and configure the following policies:

- Route Map Policies for Multicast
- Route Map Policies for Route Control
- Custom QoS Policies
- DHCP Relay Policies
- DHCP Option Policies
- IGMP Interface Policies
- MLD Snooping Policies
- L3Out Node Routing Policies
- L3Out Interface Routing Policies
- BGP Peer Prefix Policies
- IP SLA Monitoring Policies
- IP SLA Track Lists

Procedure

Step 1 Log in to your Cisco Nexus Dashboard and open the Cisco Nexus Dashboard Orchestrator service.

Step 2 Create a new Tenant Policy template.

- a) From the left navigation pane, choose **Configure > Tenant Templates > Tenant Policies**.
- b) On the **Tenant Policy Templates** page, click **Create Tenant Policy Template**.
- c) In the **Tenant Policies** page's right properties sidebar, provide the **Name** for the template.
- d) From the **Select a Tenant** drop-down, choose the tenant with which you want to associate this template.

All the policies that you create in this template as described in the following steps will be associated with the selected tenant and deployed to it when you push the template to a specific site.

By default, the new template is empty, so you must add one or more tenant policies as described in the following steps. You don't have to create every policy available in the template – you can define one or more policies of each type to deploy along with this template. If you don't want to create a specific policy, simply skip the step that describes it.

Step 3 Assign the template to one or more sites.

The process for assigning Tenant Policy templates to sites is identical to how you assign application templates to sites.

- a) In the **Template Properties** view, click **Actions** and choose **Sites Association**.
The **Associate Sites to <template-name>** window opens.
- b) In the **Associate Sites** window, check the check box next to the sites where you want to deploy the template.
Note that only the on-premises ACI sites support tenant policy templates and will be available for assignment.
- c) Click **Ok** to save.

Step 4

Create a Route Map Policy for Multicast.

This policy is part of the overarching Layer 3 Multicast use case. You can use the information in this section as a reference, but we recommend following the full set of steps that are described in the [Layer 3 Multicast, on page 295](#) chapter of the *Features and Use Cases* section of this document.

- a) From the **+Create Object** dropdown, select **Route Map Policy for Multicast**.
- b) In the right properties sidebar, provide the **Name** for the policy.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) Click **+Add Route Map for Multicast Entries** and provide the route map information.

For each route map, you must create one or more route map entries. Each entry is a rule that defines an action based on one or more matching criteria based on the following information:

- **Order** – Order is used to determine the order in which the rules are evaluated.
- **Group IP, Src IP, and RP IP** – You can use the same multicast route map policy UI for two different use cases—To configure a set of filters for multicast traffic or to restrict a rendezvous point configuration to a specific set of multicast groups. Depending on which use case you're configuring, you must fill some of the fields in this screen:
 - For multicast filtering, you can use the **Source IP** and the **Group IP** fields to define the filter. You must provide at least one of these fields, but can choose to include both. If one of the fields is left blank, it matches all values.

The Group IP range must be between 224.0.0.0 and 239.255.255.255 with a netmask between /4 and /32. You must provide the subnet mask.

The **RP IP** (Rendezvous Point IP) is not used for multicast filtering route maps, so leave this field blank.
 - For Rendezvous Point configuration, you can use the **Group IP** field to define the multicast groups for the RP.

The Group IP range must be between 224.0.0.0 and 239.255.255.255 with a netmask between /4 and /32. You must provide the subnet mask.

For a Rendezvous Point configuration, the **RP IP** is configured as part of the RP configuration. If a route-map is used for group filtering it is not necessary to configure an RP IP address in the route-map. In this case, leave the **RP IP** and **Source IP** fields empty.

- **Action** – Action defines the action to perform, either **Permit** or **Deny** the traffic, if a match is found.

- e) Click the check mark icon to save the entry.
- f) Repeat the previous substeps to create any additional route map entries for the same policy.
- g) Click **Save** to save the policy and return to the template page.
- h) Repeat this step to create any additional Route Map for Multicast policies.

Step 5

Create a Route Map Policy for Route Control.

This policy is part of the overarching L3Out and SR-MPLS L3Out use cases. You can use the information in this section as a reference, but we recommend following the full set of steps that are described in the [External Connectivity \(L3Out\), on page 229](#) and [Multi-Site and SR-MPLS L3Out Handoff, on page 337](#) chapters of the *Features and Use Cases* section of this document.

- a) From the **+Create Object** drop-down, select **Route Map Policy for Route Control**.
- b) In the right properties sidebar, provide the **Name** for the policy.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) Click **+Add Entry** and provide the route map information.

For each route map, you must create one or more context entries. Each entry is a rule that defines an action based on one or more matching criteria based on the following information:

- **Context Order** – Context order is used to determine the order in which contexts are evaluated. The value must be in the 0–9 range.
- **Context Action** – Context action defines the action to perform (`permit` or `deny`) if a match is found. If the same value is used for multiple contexts, they are evaluated one in the order in which they are defined.

When the context order and action are defined, choose how you want to match the context:

- Click **+Create Attribute** to specify the action that will be taken should the context match.

You can choose one of the following actions:

- Set Community
- Set Route Tag
- Set Dampening
- Set Weight
- Set Next Hop
- Set Preference
- Set Metric
- Set Metric Type
- Set AS Path
- Set Additional Community

After you have configured the attribute, click **Save**.

- If you want to associate the action that you defined with an IP address or prefix, click **Add IP Address**.

In the **Prefix** field, provide the IP address prefix. Both IPv4 and IPv6 prefixes are supported, for example, `2003:1:1a5:1a5::/64` or `205.205.0.0/16`.

If you want to aggregate IPs in a specific range, check the **Aggregate** check box and provide the range. For example, you can specify `0.0.0.0/0` prefix to match any IP or you can specify `10.0.0.0/8` prefix to match any `10.x.x.x` addresses.

- If you want to associate the action that you defined with community lists, click **Add Community**.

In the **Community** field, provide the community string. For example, `regular:as2-nn2:200:300`.

Then choose the **Scope**: `Transitive` means that the community will be propagated across eBGP peering (across autonomous systems) while `Non-Transitive` means the community will not be propagated.

Note

You must specify an **IP address** or a **Community** string to match a specific prefix (even if you do not provide a **Set** attribute) because it defines the prefixes that must be announced out of the L3Out. This can be either BDs' subnets or transit routes learned from other L3Outs.

- e) Repeat the previous substeps to create any additional route map entries for the same policy.
- f) Click **Save** to save the policy and return to the template page.
- g) Repeat this step to create any additional Route Map for Route Control policies.

Step 6

Create a Custom QoS Policy.

You can create a custom QoS policy in Cisco APIC to classify ingressing traffic based on its DSCP or CoS values and associate it to a QoS priority level (QoS user class) to properly handle it inside the ACI fabric. Classification is supported only if the DSCP values are present in the IP header or the CoS values are present in the Ethernet header of ingressing traffic. Also, the custom QoS policy can be used to modify the DSCP or CoS values in the header of ingressing traffic.

As an example, custom QoS policies allow you to classify traffic coming into the ACI fabric traffic from devices that mark the traffic based only on the CoS value, such as Layer-2 packets which do not have an IP header.

For detailed information about QoS functionality in ACI fabrics, see [Cisco APIC and QoS](#).

- a) From the **+Create Object** drop-down, select **Custom QoS Policy**.
- b) In the right properties sidebar, provide the **Name** for the policy.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) Click **+Add DSCP Mappings** and provide the required information.

The DSCP-mapping configuration allows you to associate ingressing traffic, whose DSCP value is within the range that is specified in the mapping, to the specified QoS priority level (class). It also allows you to set the DSCP or CoS values of the ingressing traffic, so that those values can be retained when the traffic egresses the fabric.

Note

Retaining the target CoS value for egress traffic requires the configuration of the "Preserve CoS" policy, which is part of the NDO Fabric policies.

If the "DSCP Target" or "Target CoS" values are set as part of both the DSCP Mapping and CoS Mapping, the values that are specified in the DSCP Mapping have precedence.

For each mapping, you can specify the following fields:

- **DSCP From** – The start of the DSCP range.
- **DSCP To** – The end of the DSCP range.
- **DSCP Target** – The DSCP value to set on ingressing traffic that will be retained for egressing traffic.
- **Target CoS** – The CoS value to set on ingressing traffic that will be retained for egressing traffic when "Preserve CoS" is enabled.
- **Priority** – The QoS priority class to which the traffic will be assigned.

After you provide the mappings, click the check mark icon to save. Then you can click **+Add DSCP Mappings** to provide extra mappings within the same policy.

- e) Click **Add** to save the policy and return to the template page.

- f) Click **+Add CoS Mappings** and provide the required information.

The DSCP-mapping configuration allows you to associate ingress traffic (whose DSCP value is within the range that is specified in the mapping) to the specified QoS priority level (class). It also allows you to set the DSCP or CoS values of the ingress traffic, so that those values can be retained when the traffic egresses the fabric.

Note

Retaining the target CoS value for egress traffic requires the configuration of the "Preserve CoS" policy in the NDO Fabric policies.

In addition, if the "DSCP Target" or "Target CoS" values are set as part of both the DSCP Mapping and CoS Mapping, the values that are specified in the DSCP Mapping have precedence.

For each mapping, you can specify the following fields:

- **Dot1P From** – The start of the CoS range.
- **Dot1P To** – The end of the CoS range.
- **DSCP Target** – The DSCP value to set on ingress traffic that will be retained for egressing traffic.
- **Target CoS** – The CoS value to set on ingress traffic that will be retained for egressing traffic when "Preserve CoS" is enabled.
- **Priority** – The QoS priority class to which the traffic will be assigned.

After you provide the mappings, click the check mark icon to save. Then you can click **+Add Cos Mappings** to provide extra mappings within the same policy.

- g) Click **Add** to save the policy and return to the template page.
- h) Repeat this step to create any additional Route Map for Route Control policies.

Step 7

Create a DHCP Relay Policy.

This policy is part of the overarching DHCP Relay use case. You can use the information in this section as a reference, but we recommend following the full set of steps that are described in the [DHCP Relay, on page 217](#) chapter of the *Features and Use Cases* section of this document.

- a) From the **+Create Object** drop-down, select **DHCP Relay Policy**.
- b) In the right properties sidebar, provide the **Name** for the policy.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) Click **Add Provider** to configure the DHCP server to which you want to relay the DHCP requests originated by the endpoints.
- e) Select the provider type.

When adding a relay policy, you can choose one of the following two types:

- **Application EPG**—Specifies the application EPG that includes the DHCP server to which you want to relay the DHCP requests.
- **L3 External Network**—Specifies the External EPG associated to the L3Out that is used to access the network external to the fabric where the DHCP server is connected.

Note

You can select any EPG or external EPG that has been created in the Orchestrator and assigned to the tenant you specified, even if you have not yet deployed it to sites. If you select an EPG that hasn't been deployed, you can still complete the DHCP relay configuration, but you need to deploy the EPG before the relay is available for use.

- f) Click **Select an Application EPG** or **Select an External EPG** (based on the provider type you selected) and choose the provider EPG.
- g) In the **DHCP Server Address** field, provide the IP address of the DHCP server.
- h) Enable the **DHCP Server VRF Preference** option if necessary.

This feature was introduced in Cisco APIC release 5.2(4). For more information on the use cases where it is required see the [Cisco APIC Basic Configuration Guide](#).

- i) Click **OK** to save the provider information.
- j) Repeat the previous substeps for any additional providers in the same DHCP Relay policy.
- k) Repeat this step to create any additional DHCP Relay policies.

Step 8

Create a DHCP Option Policy.

This policy is part of the overarching DHCP Relay use case. You can use the information in this section as a reference, but we recommend following the full set of steps that are described in the [DHCP Relay, on page 217](#) chapter of the *Features and Use Cases* section of this document.

- a) From the **+Create Object** drop-down, select **DHCP Option Policy**.
- b) In the right properties sidebar, provide the **Name** for the policy.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) Click **Add Option**.
- e) Provide option details.

For each DHCP option, provide the following:

- **Name** – While not technically required, we recommend using the same name for the option as listed in [RFC 2132](#).

For example, `Name Server`.

- **Id** – Provide the value if the option requires one.

For example, a list of name servers available to the client for the Name Server option.

- **Data** – Provide the value if the option requires one.

For example, a list of name servers available to the client for the Name Server option.

- f) Click **OK** to save.
- g) Repeat the previous substeps for any additional options in the same DHCP Option policy.
- h) Repeat this step to create any additional DHCP Option policies.

Step 9

Create an IGMP Interface Policy.

IGMP snooping examines IP multicast traffic within a bridge domain to discover the ports where interested receivers reside. Using the port information, IGMP snooping can reduce bandwidth consumption in a multiaccess bridge domain environment to avoid flooding the entire bridge domain.

For detailed information on IGMP snooping in ACI fabrics, see the "IGMP Snooping" chapter of the [Cisco APIC Layer 3 Networking Configuration Guide](#) for your release.

- a) From the **+Create Object** drop-down, select **IGMP Interface Policy**.
- b) In the right properties sidebar, provide the **Name** for the policy.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) Provide policy details.

- **Allow Version 3 ASM** – Allow accepting IGMP version 3 source-specific reports for multicast groups outside of the SSM range. When this feature is enabled, the switch creates an (S,G) mroute entry if it receives an IGMP version 3 report that includes both the group and source even if the group is outside of the configured SSM range. This feature is not required if hosts send (*, G) reports outside of the SSM range, or send (S,G) reports for the SSM range.
- **Fast Leave** – Option that minimizes the leave latency of IGMPv2 group memberships on a given IGMP interface because the device does not send group-specific queries. When Fast Leave is enabled, the device removes the group entry from the multicast routing table immediately upon receiving a leave message for the group. The default is disabled.

Use this only when there is only one receiver behind the BD/interface for a given group.

- **Report Link Local Groups** – Enables sending reports for groups in 224.0.0.0/24. Reports are always sent for nonlink local groups. By default, reports are not sent for link local groups.
- **IGMP Version** – IGMP version that is enabled on the bridge domain or interface. The IGMP version can be 2 or 3. The default is 2.
- **Advanced Settings** – Click the arrow next to this section to expand.
 - **Group Timeout** – Group membership interval that must pass before the router decides that no members of a group or source exist on the network. Values range 3–65,535 seconds. The default is 260 seconds.
 - **Query Interval** – Sets the frequency at which the software sends IGMP host query messages. Values can range 1–18,000 seconds. The default is 125 seconds.
 - **Query Response Interval** – Sets the response time that is advertised in IGMP queries. Values can range 1–25 seconds. The default is 10 seconds.
 - **Last Member Count** – Sets the number of times that the software sends an IGMP query in response to a host leave message. Values can range 1–5. The default is 2.
 - **Last Member Response Time** – Sets the query interval waited after sending membership reports before the software deletes the group state. Values can range 1–25 seconds. The default is 1 second.
 - **Startup Query Count** – Configures snooping for several queries that are sent at startup when you do not enable Protocol Independent Multicast because multicast traffic does not need to be routed. Values can range 1–10. The default is 2 messages.
 - **Startup Query Interval** – This configures the IGMP snooping query interval at startup. The range is from 1 second to 18,000 seconds. The default is 125 seconds.
 - **Querier Timeout** – Sets the query timeout that the software uses when deciding to take over as the querier. Values can range 1–65,535 seconds. The default is 255 seconds.
 - **Robustness Variable** – Sets the robustness variable. You can use a larger value for a lossy network. Values can range 1–7. The default is 2.
 - **State Limit Route Map** – Used with Reserved Multicast Entries feature.
The route map policy must be already created as described in Step 2.
 - **Report Policy Route Map** – Access policy for IGMP reports that is based on a route-map policy. IGMP group reports will only be selected for groups that are allowed by the route-map.
The route map policy must be already created as described in Step 2.

- **Static Report Route Map** – Statically binds a multicast group to the outgoing interface, which is handled by the switch hardware. If you specify only the group address, the (*, G) state is created. If you specify the source address, the (S, G) state is created. You can specify a route-map policy name that lists the group prefixes, group ranges, and source prefixes. A source tree is built for the (S, G) state only if you enable IGMPv3.

The route map policy must be already created as described in Step 2.

- **Maximum Multicast Entries** – Limit the mroute states for the BD or interface that are created by IGMP reports. Default is disabled and no limit is enforced. Valid range is 1-4294967295.

- e) Repeat this step to create any additional IGMP Interface policies.

Step 10

Create an MLD Snooping Policy.

Multicast Listener Discovery (MLD) snooping enables the efficient distribution of IPv6 multicast traffic between hosts and routers. It is a Layer 2 feature that restricts IPv6 multicast traffic within a bridge domain to a subset of ports that have sent or received MLD queries or reports. In this way, MLD snooping provides the benefit of conserving the bandwidth on those segments of the network where no node has expressed interest in receiving the multicast traffic. This reduces the bandwidth usage instead of flooding the bridge domain, and also helps hosts and routers save unwanted packet processing.

For detailed information on MLD snooping in ACI fabrics, see the "MLD Snooping" chapter of the [Cisco APIC Layer 3 Networking Configuration Guide](#) for your release.

- From the **+Create Object** drop-down, select **MLD Snooping Policy**.
- In the right properties sidebar, provide the **Name** for the policy.
- (Optional) Click **Add Description** and provide a description for the policy.
- Provide policy details.

- **Admin State** – Enables or disables the MLD snooping feature.

- **Fast Leave Control** – Allows you to turn on or off the fast-leave feature on a per bridge domain basis. This applies to MLDv2 hosts and is used on ports that are known to have only one host doing MLD behind that port.

Default is `disabled`.

- **Querier Control** – Enables or disables MLD snooping querier processing. MLD snooping querier supports the MLD snooping in a bridge domain where PIM and MLD are not configured because the multicast traffic does not need to be routed.

Default is `disabled`.

- **Querier Version** – Allows you to choose the querier version.

Default is `Version2`.

- **Advanced Settings** – Click the arrow next to this section to expand.

- **Query Interval** – Sets the frequency at which the software sends MLD host query messages. Values can range 1–18,000 seconds.

The default is 125 seconds.

- **Query Response Interval** – Sets the response time that is advertised in MLD queries. Values can range 1–25 seconds.

The default is 10 seconds.

- **Last Member Query Interval** – Sets the query response time after sending membership reports before the software deletes the group state. Values can range 1–25 seconds.

The default is 1 second.

- **Start Query Count** – Configures snooping for several queries that are sent at startup when you do not enable PIM because multicast traffic does not need to be routed. Values can range 1–10 .

The default is 2.

- **Start Query Interval** – Configures a snooping query interval at startup when you do not enable PIM because multicast traffic does not need to be routed. Values can range 1–18,000 seconds.

The default is 31 seconds.

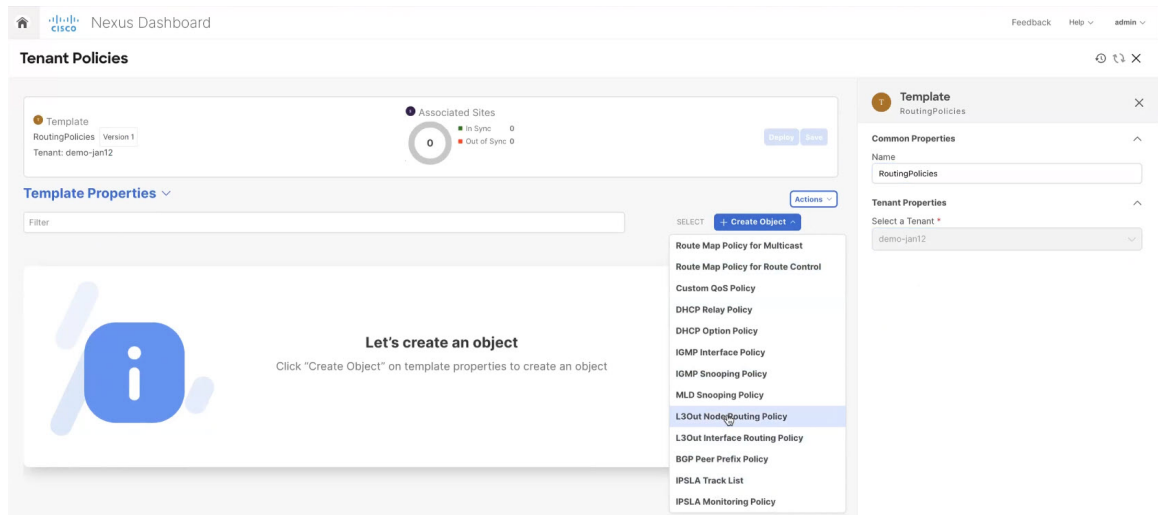
- e) Repeat this step to create any additional MLD Snooping policies.

Step 11

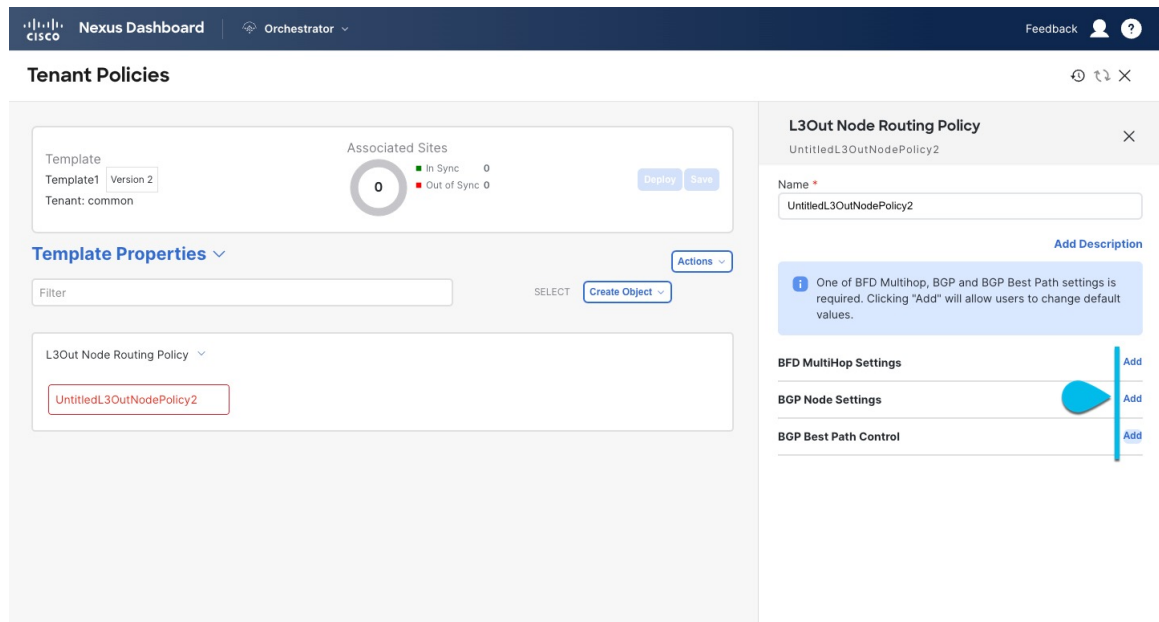
Create an L3Out Node Routing Policy.

This policy is part of the overarching L3Out and SR-MPLS L3Out configuration use case. You can use the information in this section as a reference, but we recommend following the full set of steps that are described in the [External Connectivity \(L3Out\)](#), on page 229 chapter of the *Features and Use Cases* section of this document.

- a) In the main pane, choose **Create Object > L3Out Node Routing Policy**.



- b) Provide the **Name** for the policy, and **Add** at least one of the **BFD MultiHop Settings**, **BGP Node Settings**, or **BGP Best Path Control** options.



- **BFD MultiHop Settings** – provides forwarding failure detection for destinations with more than one hop.

In this case, a MultiHop session is created between the source and destination instead of the interface like in single-hop scenarios.

Note

BFD MultiHop configuration requires Cisco APIC release 5.0(1) or later.

- **BGP Node Settings** – allows you to configure BGP protocol timer and sessions settings for BGP adjacencies between BGP peers.
- **BGP Best Path Control** – enables `as-path multipath-relax`, which allows load-balancing between multiple paths that are received from different BGP ASN.

Step 12 Create an L3Out Interface Routing Policy.

This policy is part of the overarching L3Out and SR-MPLS L3Out configuration use case. You can use the information in this section as a reference, but we recommend following the full set of steps that are described in the [External Connectivity \(L3Out\)](#), on page 229 chapter of the *Features and Use Cases* section of this document.

- In the main pane, choose **Create Object > L3Out Interface Routing Policy**.
- Provide the **Name** for the policy, and define the **BFD Settings**, **BFD Multi-Hop Settings**, and **OSPF Interface Settings**.

- **BFD Settings** – specifies BFD parameters for BFD sessions established between devices on interfaces that are directly connected.

When multiple protocols are enabled between a pair of routers, each protocol has its own link failure detection mechanism, which may have different timeouts. BFD provides a consistent timeout for all protocols to allow consistent and predictable convergence times.

- **BFD MultiHop Settings** – specifies BFD parameters for BFD sessions established between devices on interfaces that are not directly connected.

You can configure these settings at the node level as mentioned in the "Tenant Policy Template: Node Routing Group Policy" section above, in which case the interfaces inherit those settings, or you can overwrite the node-level settings for individual interfaces in the Interface Routing group policy.

Note

BFD multi-hop configuration requires Cisco APIC release 5.0(1) or later.

- **OSPF Interface Settings** – allows you to configure interface-level settings such as OSPF network type, priority, cost, intervals and controls.

Note

This policy must be created when deploying an L3Out with OSPF.

Step 13 Create a BGP Peer Prefix Policy.

This policy is part of the overarching L3Out and SR-MPLS L3Out configuration use case. You can use the information in this section as a reference, but we recommend following the full set of steps that are described in the [External Connectivity \(L3Out\)](#), on page 229 chapter of the *Features and Use Cases* section of this document.

- In the main pane, choose **Create Object** > **BGP Peer Prefix Policy**.
- Provide the **Name** for the policy, and define the **Max Number of Prefixes** and the **Action** to take if the number is exceeded.

The following actions are available:

- **Log**

- **Reject**
- **Restart**
- **Shutdown**

Step 14 Create an IP SLA Monitoring Policy.

This policy is part of the overarching L3Out and SR-MPLS L3Out configuration use case. You can use the information in this section as a reference, but we recommend following the full set of steps that are described in the [External Connectivity \(L3Out\)](#), on page 229 chapter of the *Features and Use Cases* section of this document.

- In the main pane, choose **Create Object > IP SLA Monitoring Policy**.
- Provide the **Name** for the policy, and define its settings.

Note

If you choose `HTTP` for the **SLA Type**, your fabric must be running Cisco APIC release 5.1(3) or later.

Step 15 Create an IP SLA Track List.

This policy is part of the overarching L3Out and SR-MPLS L3Out configuration use case. You can use the information in this section as a reference, but we recommend following the full set of steps that are described in the [External Connectivity \(L3Out\)](#), on page 229 chapter of the *Features and Use Cases* section of this document.

- In the main pane, choose **Create Object > IP SLA Track List**.
- Provide the **Name** for the policy.
- Choose the **Type**.

The definition of a route being available or not available can be based on `Threshold Percentage` or `Threshold Weight`.

- Click **+Add Track List to Track Member Relation** to add one or more track members to this track list.

Note

You must select a bridge domain or an L3Out to associate with the track member. If you do not already have the bridge domain (BD) or L3Out that is created, you can skip adding a track member, save the policy without assigning one, and come back to it after you have created the BD or L3Out.

- In the **Add Track List to Track Member Relation** dialog, provide the **Destination IP**, **Scope Type**, and choose the **IP SLA Monitoring Policy**.

The scope for the track list can be either bridge domain or L3Out. The IP SLA Monitoring policy is the one you created in the previous step.

Step 16 Click **Save** to save the changes you've made to the template.

Note

When you save (or deploy) the template to one or more sites, the Orchestrator will verify that the specified nodes or interfaces are valid for the sites and will return an error.

Step 17 Click **Deploy** to deploy the template to the associated sites.

The process for deploying tenant policy templates is identical to how you deploy application templates.

If you have previously deployed this template but made no changes to it since, the **Deploy** summary indicates that there are no changes, and you can choose to redeploy the entire template. In this case, you can skip this step.

Otherwise, the **Deploy to sites** window shows you a summary of the configuration differences that will be deployed to sites. Note that in this case only the difference in configuration is deployed to the sites. If you want to redeploy the entire template, you must deploy when to sync the differences, and then redeploy again to push the entire configuration as described in the previous paragraph.



CHAPTER 5

Schemas and Application Templates

- [Shadow Objects, on page 63](#)
- [Creating Schemas and Templates, on page 68](#)
- [Cloning Schemas, on page 87](#)

Shadow Objects

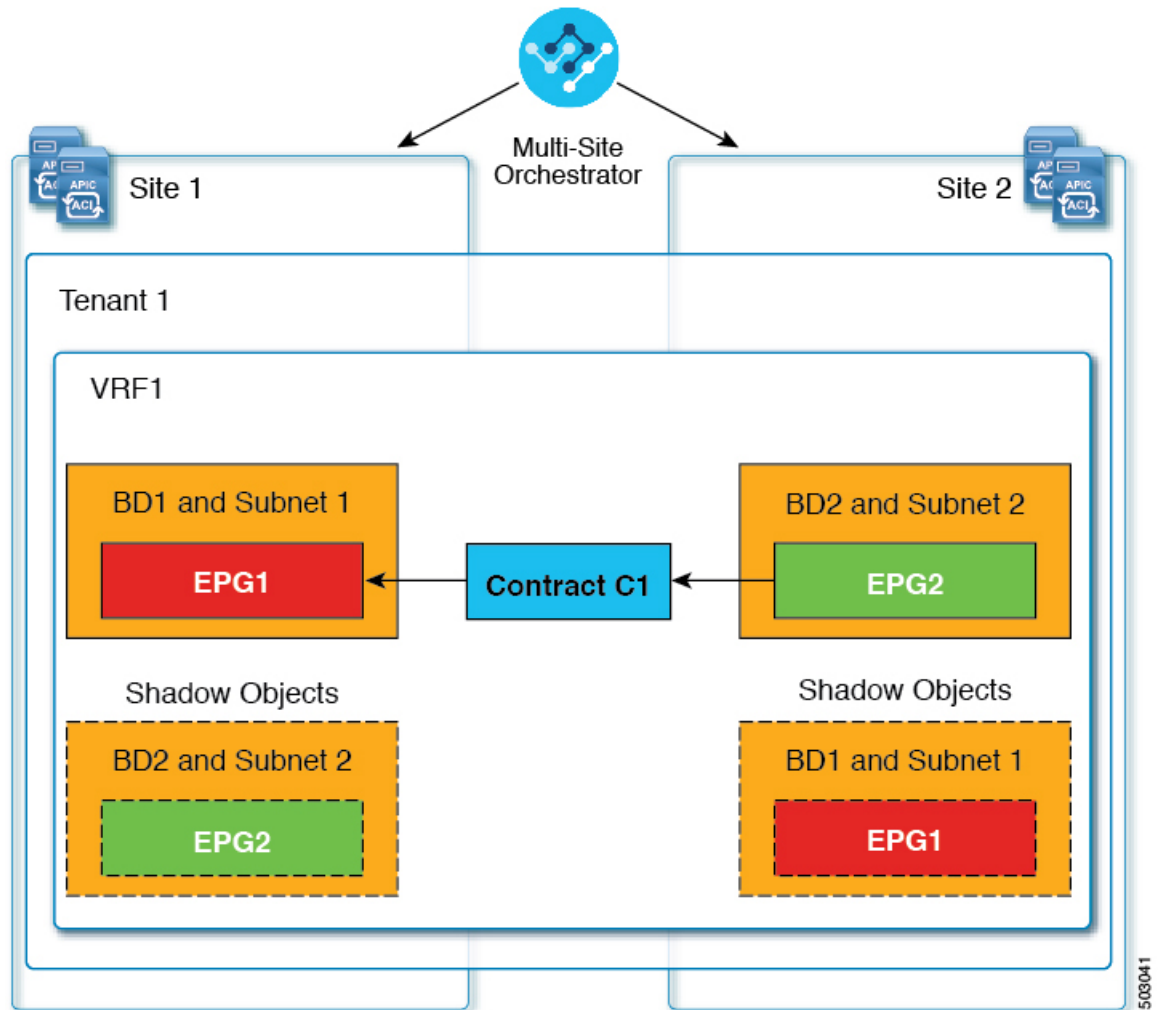
When a contract exists between site-local EPGs in stretched VRF or in Shared Services use-cases where provider and consumer are in different VRFs and communicate through Tenant contracts, the EPGs and bridge domains (BDs) are mirrored on the remote sites. The mirrored objects appear as if they are deployed in each of these sites' controllers, while only actually being deployed in one of the sites. These mirrored objects are called "shadow" objects.



Note Shadow objects should not be removed using the APIC GUI.

For example, if a tenant and VRF are stretched between Site1 and Site2, provider EPG and its bridge domain are deployed in Site2 only, and consumer EPG and its domain are deployed in Site1 only, then corresponding shadow bridge domains and EPGs will be deployed as shown in the figure below. They appear with the same names as the ones that were deployed directly to each site.

Figure 8: Basic Shadow EPG



The following objects can be shadowed:

- VRFs
- Bridge Domains (BDs)
- L3Outs
- External EPGs
- Application Profiles
- Application EPGs
- Contracts (Hybrid Cloud deployments)

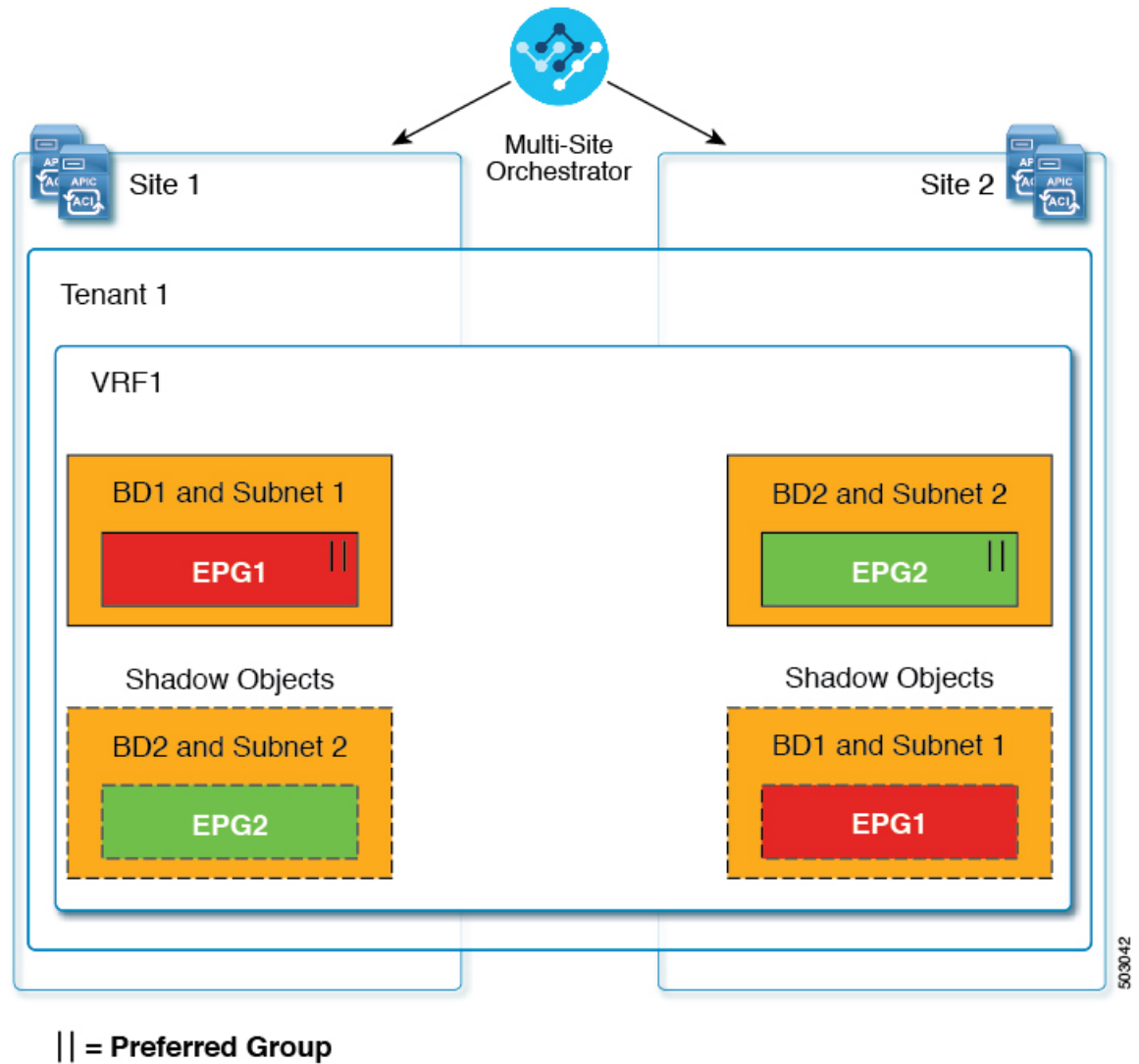
If your fabrics are running APIC Release 5.0(2) or later, when you select a shadow object in the APIC GUI, you will see a **This is a shadow object pushed by MSC to support intersite policies. Do not make any changes or delete this object.** warning at the top of main GUI pane. In addition, shadow

EPGs that are not part of a VMM domain will not have static ports, while shadow BDs will have **No Default SVI Gateway** option enabled in the APIC GUI.

Other Use Cases with Shadow Objects

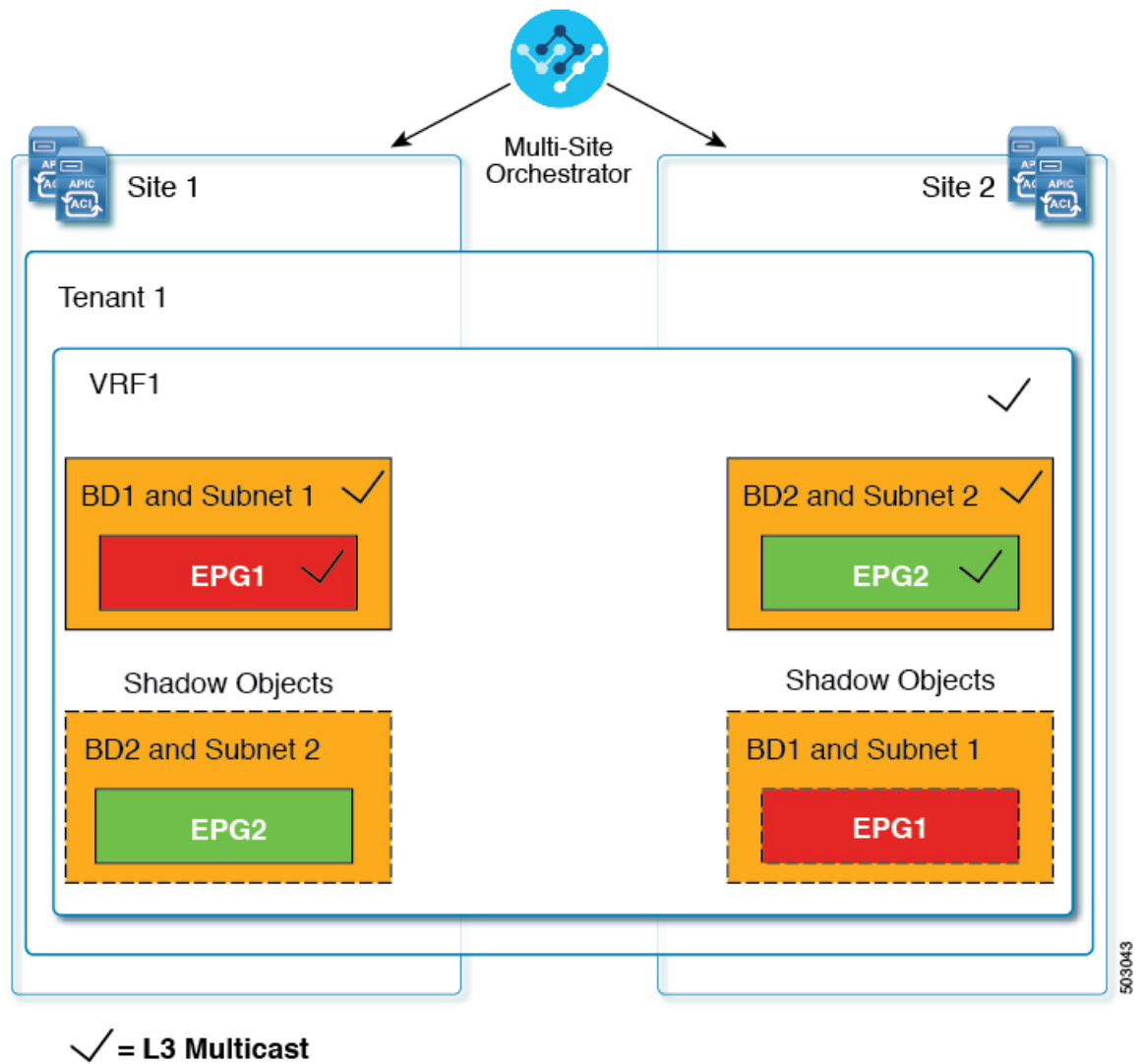
Shadow objects are also created in a number of other use cases, such as Preferred Group, vzAny, and Layer 3 Multicast, and hybrid cloud, as shown in the figures below.

Figure 9: Preferred Group



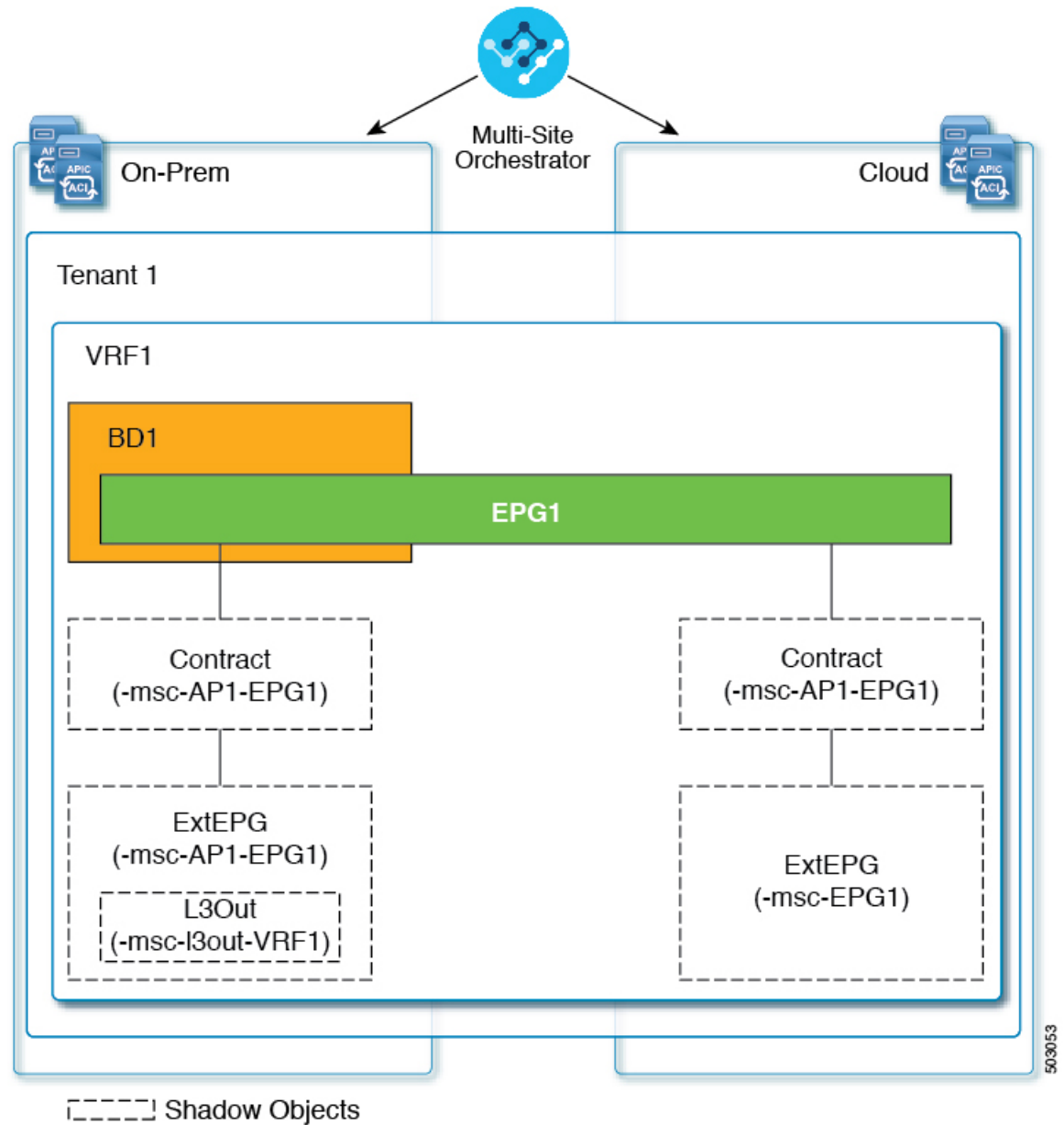
In case of multicast, the shadow objects are created only for EPGs/BDs that have multicast sources connected and the option explicitly configured at the EPG level.

Figure 10: L3 Multicast



In case of hybrid cloud deployments, even stretched objects will create shadow objects where implicit contracts exist. For example, in the following case where an EPG is stretched between an on-premises and cloud sites, shadow external EPGs are created in each site with implicit shadow contracts between the stretched EPG and the shadow external EPGs.

Figure 11: Hybrid Cloud



Starting with Cisco APIC, Release 5.2(3), shadow objects are indicated by a unique icon in the Cisco APIC GUI. Regular Orchestrator-created objects are shown with a green cloud symbol, whereas the shadow objects will have a gray cloud icon.

Hiding Shadow Objects in APIC GUI

Starting with APIC Release 5.0(2), you can choose to show or hide the shadow objects created by the Nexus Dashboard Orchestrator in the on-premises site's APIC GUI. Shadow objects in Cloud Network Controller are always hidden.

If you want to hide shadow objects from the GUI, keep the following in mind:

- This option cannot be set globally from the Orchestrator and must be set directly in each site's APIC as described in this section.
- The option to show shadow objects is turned off by default for all new APIC Release 5.0(2) installations and upgrades, so previously visible objects may become hidden.
- Hiding shadow objects relies on a flag set by the Nexus Dashboard Orchestrator specifically for this feature, which is enabled from Orchestrator Release 3.0(2) and later:
 - If shadow objects are deployed by an earlier Orchestrator version, they will not have the required tag and will always be visible in the APIC GUI.
 - If shadow objects are deployed by Orchestrator version 3.0(2) or later, they will have the tag and can be hidden or shown using the APIC GUI setting.
 - We recommend upgrading each fabric to APIC Release 5.0(2) before upgrading the Nexus Dashboard Orchestrator.

When the Nexus Dashboard Orchestrator is upgraded to Release 3.0(2), any objects deployed to sites running APIC Release 5.0(2) or later will be tagged with appropriate tags and can be shown or hidden using the APIC GUI without having to re-deploy them.

If you upgrade the Orchestrator before the fabric's APIC, the site's objects will not be tagged and you will need to manually re-deploy the configuration after the fabric is upgraded for the flag to be set.

- If you ever downgrade your fabric to a release prior to Release 5.0(2), the shadow objects will no longer be hidden and you may see a different icon for them in the APIC GUI.

Procedure

-
- Step 1** Log in to the site's APIC.
- Step 2** In the top right corner, click the **Manage my profile** icon and choose **Settings**.
- Step 3** In the **Application Settings** window, enable or disable the **Show Hidden Policies** checkbox.
The setting is stored in the user profile and is enable or disabled separately for each user.
- Step 4** Repeat the process for any additional APIC sites.
-

Creating Schemas and Templates

Before you begin

- You must have at least one available tenant that you want to incorporate into your site.
For more information, see [Tenants and Tenant Policies Templates, on page 47](#).

Procedure

Step 1 Log in to your Cisco Nexus Dashboard and open the Cisco Nexus Dashboard Orchestrator service.

Step 2 Create a new schema.

- a) From the left navigation pane, choose **Configure > Tenant Template**.
- b) On the Schemas page, click **Add Schema**.
- c) In the schema creation dialog, provide the **Name** and optional description for the schema and click **Add**.

By default, the new schema is empty, so you must add one or more templates.

Step 3 Create a template.

- a) In the schema page, click **Create New Template**.
- b) In the **Select a Template type** window, choose **ACI Multi-Cloud** and click **Add**.
 - **ACI Multi-Cloud**—Templates that are used for Cisco ACI on-premises and cloud sites, which allow template and object stretching between multiple sites. This template supports two deployment types:
 - **Multi-Site** - The template can be associated to a single site (site-local policies) or to multiple sites (stretched policies) and the option should be selected for Multi-Site Network (ISN) or VXLAN intersite communication to allow template and object stretching between multiple sites.
 - **Autonomous** - The template can be associated to one or more sites that are operated independently and are not connected through an intersite Network (no intersite VXLAN communication).

Because autonomous sites are by definition that is isolated and do not have any intersite connectivity, there is no shadow object configuration across sites and no cross-programming of pctxs or VNIDs in the spine switches for intersite traffic flow.

The autonomous templates also allow for higher deployment scale.

The following sections focus primarily on this type of templates.

- **NDFC**—Templates designed for Cisco Nexus Dashboard Fabric Controller (formerly Data Center Network Manager) sites.

This guide described Cisco Nexus Dashboard Orchestrator configurations for on-premises Cisco ACI fabrics. For information on working with Cisco NDFC sites, see the [Cisco Nexus Dashboard Orchestrator Configuration Guide for NDFC Fabrics](#) instead.
- **Cloud Local**—Templates designed for specific Cloud Network Controller use cases, such as Google Cloud site connectivity, and cannot be stretched between multiple sites.

This guide describes Cisco Nexus Dashboard Orchestrator configurations for on-premises Cisco ACI fabrics. For information on working with Cloud Network Controller fabrics, see the [Cisco Nexus Dashboard Orchestrator use case library](#) instead.

- c) In the right sidebar, provide the **Display Name** for the template.
- d) (Optional) Provide a **Description**.
- e) From the **Select a Tenant** drop-down, select the Tenant for this template.

Keep in mind, the user account you're using to create a new schema must be associated with the tenant you are trying to add to it, otherwise the tenant will not be available in the drop-down list. Associating a user account with a tenant is described in [Tenants and Tenant Policies Templates, on page 47](#).

- f) In the template view page, click **Save**.

You must save the template after this initial configuration for extra options (such as site association) to become available.

- g) Repeat this step to create any additional templates.

For more information on schema and template design, see [Schema and Template Design Considerations, on page 13](#).

Step 4 Assign the templates to sites.

You deploy fabric configuration by deploying one template at a time to one or more sites. So you must associate the template with at least one site where you want to deploy the configuration.

- In the template view page, click **Actions** and choose **Add/Remove Sites**.
- In the **Add/Remove Sites <template>** dialog, select one or more sites where you want to deploy the template and click **Ok**.

What to do next

After you have created a schema and one or more templates, you can proceed with editing the templates as described in the following sections of this document based on your specific use cases. After you finish defining configurations, you can deploy the templates as described in [Deploying Templates, on page 20](#).

Importing Schema Elements From APIC Sites

You can create new objects and push them out to one or more sites or you can import existing site-local objects and manage them using the Multi-Site Orchestrator. This section describes how to import one or more existing objects, while creating new objects is described later on in this document.

When importing policies from APIC into NDO, the common practice is to import some objects, such as VRFs or contracts, into a stretched template and other objects, such as non-stretched EPGs or BDs, into site-local templates.

Prior to Release 3.1(1), importing an object into a site-local template that referenced another object that is part of a stretched template presented certain challenges, for example:

- If a referenced object already exists in NDO and a new object is imported with the **Include Relations** option enabled, NDO would throw an error when trying to deploy the site-local template because of object duplication since the referenced object already existed.
- However, not importing the referenced object (**Include Relations** option disabled) would require an administrator to perform manual mapping with the referenced object after the import.

When importing an object into a site-local template that has references with another object that is part of a different template (in the same or a different schema), the references are automatically resolved by NDO. In such cases, the **Import Relations** option will be grayed-out in the UI for the object that is being imported and a warning tooltip will provide additional info, such as: *[Referenced Object] already exists in [Template]. Existing relations are imported by default.* While such objects are imported with their relations by default, you can change the references once the import operation is completed, for example by re-mapping a BD to a different VRF. The new behavior applies to all configuration objects that can be imported.

To import one or more objects from sites:

Procedure

-
- Step 1** Open the **Schema** where you want to import objects.
- Step 2** In the left sidebar, select the **Template** where you want to import objects.
- Step 3** In the main pane click the **Import** button and select the **Site** from which you want to import.
- Step 4** In the **Import from <site-name>** window that opens, select one or more objects.

Note

The names of the objects imported into NDO must be unique across all sites. Importing different objects with duplicate names will cause a schema validation error and the import to fail. If you want to import objects that have the same name, you must first rename them.

- Step 5** (Optional) Enable the **Import Relations** knob to import all related objects.
- For example, when importing a BD, enabling the **Import Relations** knob will import the associated VRF as well.

Note

As described previously, the **Import Relations** knob will be enabled by default and cannot be disabled for objects whose related objects already exist in NDO.

- Step 6** Click **Import**.
-

Configuring VRFs

This section describes how to create a VRF.

Before you begin

You must have the schema and template created and a tenant assigned to the template, as described in [Creating Schemas and Templates, on page 68](#).

Procedure

-
- Step 1** Select the schema and template where you want to create the VRF.
- Step 2** Create the VRF.
- a) In the main pane, select **Create Object > VRF**.
- Alternatively, you can scroll down to the **VRFs** area and click **Create VRF**.
- b) In the properties pane, provide the **Display Name** for the VRF.
 - c) (Optional) Provide a **Description**.

- Step 3** (Optional) Add one or more **Annotations**.

This allows you to add arbitrary `key:value` pairs of metadata to an object as annotations (`tagAnnotation`). Annotations are provided for any custom purposes you may require, such as descriptions, markers for personal scripting or API calls, or flags for monitoring tools or orchestration applications such as your Nexus Dashboard Orchestrator. Because APIC

ignores these annotations and merely stores them with other object data, there are no format or content restrictions imposed by APIC.

Step 4 Configure the **On-Premises Properties** for the VRF.

a) Specify **Policy Control Enforcement Preference**.

Note that you cannot change the Policy Control Enforcement for newly created VRFs and the setting is locked to the `enforced` mode.

However, you can use this to transition any VRF that you import from an APIC site that is configured as `unenforced` to the `enforced` mode after importing it. A typical use case is for brown field deployments where existing VRFs must be converted to `enforced` mode to support stretching them between sites. Once you have transitioned an imported VRF from `unenforced` to `enforced` in NDO, you will not be able to make further changes to this field.

- `Enforced`—Security rules (contracts) will be enforced.
- `Unenforced`—Security rules (contracts) will not be enforced.

b) (Optional) Enable **IP Data-Plane Learning**.

Defines if IP addresses are learned through data-plane packets for the VRF.

When disabled, IP addresses are not learned from the data-plane packets. Local and remote MAC addresses are still learned, but local IP addresses are not learned from data packets.

Regardless of whether this parameter is enabled or disabled, local IP addresses can still be learned from ARP, GARP, and ND.

c) (Optional) Enable **L3 Multicast** for the VRF.

For additional information, see [Layer 3 Multicast, on page 295](#).

d) (Optional) Enable **vzAny** for the VRF.

For additional information, see [vzAny Contracts, on page 363](#).

e) (Optional) Enable **Preferred Group** for the VRF.

For additional information, see [EPG Preferred Groups Overview and Limitations, on page 225](#)

f) (Optional) Enable **BD Enforcement Status** for the VRF.

By default, servers from an EPG of a given bridge domain can ping the SVI (subnet) of another bridge domain. If you wish to constrain a host to be able to ping only the SVI of the bridge domain to which it belongs, you can enable this BD Enforcement Status option configuration on the VRF. This blocks ICMP, TCP, and UDP traffic to the subnet IP address of bridge domains that are different from the one to which the server belongs.

Configuring Bridge Domains

This section describes how to configure a Bridge Domain (BD).

Before you begin

- You must have the schema and template that is created and a tenant that is assigned to the template, as described in [Creating Schemas and Templates, on page 68](#).

- You must have the VRF created as described in [Configuring VRFs, on page 71](#).

Procedure

Step 1 Select the schema and template where you want to create the bridge domain.

Step 2 Create a bridge domain.

- a) In the main pane, select **+Create Object > Bridge Domain**.

Alternatively, you can scroll down to the **Bridge Domains** area and click **Create Bridge Domain**.

- b) In the properties pane, provide the **Display Name** for the bridge domain.

- c) (Optional) Provide a **Description**.

Step 3 (Optional) Add one or more **Annotations**.

This allows you to add arbitrary `key:value` pairs of metadata to an object as annotations (`tagAnnotation`). Annotations are provided for any custom purposes that you may require, such as descriptions, markers for personal scripting or API calls, or flags for monitoring tools or orchestration applications such as your Cisco Nexus Dashboard Orchestrator. Because APIC ignores these annotations and merely stores them with other object data, there are no format or content restrictions that are imposed by APIC.

Step 4 Configure **On-Premises Properties**.

- a) From the **Virtual Routing & Forwarding** drop-down, select the VRF for this BD.

- b) (Optional) Enable **L2 Stretch**.

- c) (Optional) Enable **Intersite BUM Traffic Allow**.

This option becomes available if you enabled **L2 Stretch**.

- d) (Optional) Enable **Optimized WAN Bandwidth**.

This option becomes available if you enabled **L2 Stretch**.

- e) (Optional) Enable **Unicast Routing**.

If this setting is enabled and a subnet address is configured, the fabric provides the default gateway function and routes the traffic. Enabling unicast routing also instructs the mapping database to learn the endpoint IP-to-VTEP mapping for this bridge domain. The IP learning is not dependent upon having a subnet that is configured under the bridge domain.

- f) (Optional) Enable **L3 Multicast** for the BD.

For additional information about Layer 3 multicast, see [Layer 3 Multicast, on page 295](#).

- g) (Optional) Choose **L2 Unknown Unicast** mode.

By default, unicast traffic is flooded to all Layer two-ports. If enabled, unicast traffic flooding is blocked at a specific port, only permitting egress traffic with MAC addresses that are known to exist on the port. The method can be `Flood` or `Hardware Proxy`.

When the BD has L2 Unknown Unicast set to Flood, if an endpoint is deleted the system deletes it from both the local leaf switches and the remote leaf switches where the BD is deployed, by selecting Clear Remote MAC Entries. Without this feature, the remote leaf switch continues to have this endpoint learned until the timer expires.

Note

Modifying the L2 Unknown Unicast setting causes traffic to bounce (go down and up) on interfaces to devices attached to EPGs associated with this bridge domain.

h) (Optional) Choose **Unknown Multicast Flooding** mode.

This is applicable for IPv4 unknown multicast traffic and is the node forwarding parameter for Layer 3 unknown multicast destinations.

- **Flood** (default)—Unknown IPv4 multicast traffic is flooded on all front panel ports that are attached with the EPGs associated with this bridge domain. Flooding is not restricted to only M-Router ports of the bridge domain.
- **Optimized Flood**—Send the data only to M-router ports in the bridge domain.

i) (Optional) Choose **IPv6 Unknown Multicast Flooding** mode.

This is applicable for IPv6 unknown multicast traffic and is the node forwarding parameter for Layer 3 unknown multicast destinations.

- **Flood** (default)—Unknown IPv6 multicast traffic is flooded on all front panel ports that are attached with the EPGs associated with this bridge domain. Flooding is not restricted to only M-Router ports of the bridge domain.
- **Optimized Flood**—Send the data only to M-router ports in the bridge domain.

j) (Optional) Choose **Multi-Destination Flooding** mode.

The multiple destination forwarding method for Layer 2 multicast and broadcast traffic.

- **Flood in BD**—Sends the data to all ports on the same bridge domain.
- **Drop**—Drops Packet. Never sends the data to any other ports.
- **Flood in Encapsulation**—Send the data to all the EPG ports with the same VLAN within the bridge domain, except for the protocol packets which are flooded to the entire bridge domain.

Note

This mode is supported only when the **L2 Stretch** option is disabled and is not supported for BDs that are stretched across sites.

k) (Optional) Enable **ARP Flooding**.

Enables ARP flooding, so that the Layer 2 broadcast domain maps IP addresses to the MAC addresses. If flooding is disabled, unicast routing will be performed on the target IP address.

Enables ARP flooding, so that ARP request will be flooded inside the Layer 2 broadcast domain. If the BD is stretched across sites, enabling ARP flooding is only possible with enabling **Intersite BUM Traffic Allow**. When ARP flooding is disabled, the leaf switch receiving the ARP request from a locally connected endpoint forwards it directly to the remote leaf switch where the target endpoint of the ARP request is connected (if the IP for the remote endpoint is known in the endpoint table) or to the spines (if the IP for the remote endpoint is not known in the endpoint table).

If you set the **L2 Unknown Unicast** mode to **Flood**, the **ARP Flooding** cannot be disabled. If the **L2 Unknown Unicast** mode is set to **Hardware Proxy**, ARP flooding can be enabled or disabled.

l) (Optional) Provide **Virtual MAC Address**.

The BD virtual MAC address and the subnet virtual IP address must be the same for all ACI fabrics for that bridge domain. Multiple bridge domains can be configured to communicate across connected ACI fabrics. The virtual MAC address and the virtual IP address can be shared across bridge domains.

Note

Virtual MAC along with virtual IP subnet should be used only for migration of individual sites to NDO-managed multi-site fabric. When the migration is completed, these flags can be disabled.

Step 5

Add one or more **Subnets** for the BD.

- a) Click **+Add Subnet**.

An **Add New Subnet** window opens.

- b) Enter the subnet's **Gateway IP** address and a **Description** for the subnet that you want to add.
c) If necessary, enable **Treat as virtual IP address** option.

This option along with the **Virtual MAC Address** on the BD can be used for migration scenarios from individual Common Pervasive Gateway configuration to NDO-managed Multi-Site deployments.

- d) Select the **Scope** for the subnet.

The network visibility of the subnet.

- **Private to VRF**—Prevents the subnet from being announced over L3Out toward an external network domain.
- **Advertised Externally**—The subnet can be announced through L3Out toward an external network domain.

- e) (Optional) Enable **Shared Between VRFs**.

Shared between VRFs—The subnet can be shared with and exported to multiple contexts (VRFs) in the same tenant or across tenants as part of a shared service. An example of a shared service is a routed connection to an EPG present in another context (VRF) in a different tenant. This enables traffic to pass in both directions across contexts (VRFs). An EPG that provides a shared service must have its subnet that is configured under that EPG (not under a bridge domain), and its scope must be set to advertised externally, and shared between VRFs.

Shared subnets must be unique across the contexts (VRF) involved in the communication. When a subnet under an EPG provides a Layer 3 external network shared service, such a subnet must be globally unique within the entire ACI fabric.

- f) Leave the **No Default SVI Gateway** option unchecked.

Enabling this option means that only the proxy route (subnet route to spine proxy) is programmed on the leaf switches and no SVI is created, which means SVI cannot be used as the gateway.

We recommend that SVI is created by the BD subnet as the gateway and the **No Default SVI Gateway** option is enabled on the EPG instead because EPG subnets should only be used for route leaking.

- g) (Optional) Enable **Querier** option.

Enables IGMP Snooping on the subnet

- h) (Optional) Enable **Primary** option to designate the subnet as primary.

There can be one primary IPv4 subnet and one primary IPv6 subnet.

- i) Click **Save**.

Step 6

(Optional) Enable **EP Move Detection Mode**.

Uses the information that is received with a Gratuitous Address Resolution Protocol (GARP) packet to update the endpoint table when a specific IP address that was previously associated to one MAC address (`mac-a`) gets associated to a different MAC address (`mac-b`). This applies to the specific scenario where the move occurs on the same interface.

Although Cisco ACI can detect MAC and IP address movement between leaf switch ports, leaf switches, bridge domains, and EPGs, it does not detect the movement of an IP address to a new MAC address if the new MAC address is from the same interface and same EPG as the old MAC address.

When the GARP-based detection option is enabled, Cisco ACI triggers an endpoint move based on GARP packets if the move occurs on the same interface and same EPG. If a GARP packet comes from the same interface and same EPG, then endpoint learning is triggered only when Unicast Routing, ARP Flooding, and “GARP based detection” are all enabled for the bridge domain.

Step 7 (Optional) Add an **IGMP Interface Policy**.

You can configure several Tenant Policy templates and associate them with policy objects. For more information, see [Creating Tenant Policy Templates, on page 50](#).

Step 8 (Optional) Add an **IGMP Snoop Policy**.

You can configure several Tenant Policy templates and associate them with policy objects. For more information, see [Creating Tenant Policy Templates, on page 50](#).

Step 9 (Optional) Add an **MLD Snoop Policy**.

You can configure several Tenant Policy templates and associate them with policy objects. For more information, see [Creating Tenant Policy Templates, on page 50](#).

Step 10 (Optional) Add a **DHCP Policy**.

For additional information, see [DHCP Relay, on page 217](#).

Step 11 Configure the bridge domain's site-local properties as necessary.

In addition to the template-level configurations, you can also define one or more site-local properties for the bridge domain, as described in [Configuring Bridge Domain's Site-Local Properties, on page 76](#)

Configuring Bridge Domain's Site-Local Properties

In addition to the template-level properties you typically configure for the object when you create it in a template, you can also define one or more properties that are specific to each site to which you assign the template.

When you deploy the object to more than 1 site, the same template-level configurations are deployed to all sites, while the site-local configurations are deployed to those specific sites only.

Before you begin

You must have:

- Created the bridge domain and configured its template-level properties, as described in [Configuring Bridge Domains, on page 72](#).
- Assigned the template that contains the bridge domain to one or more sites.

Procedure

-
- Step 1** Open the schema that contains the template with the bridge domain.
- Step 2** In the left sidebar, select the template that contains the bridge domain under the specific site that you want to configure.
- Step 3** In the main pane, select the bridge domain.
- For most fields, you see the values that you have configured at the template level, which you cannot edit here.
- Step 4** Click **+L3Out** to add an L3Out.
- This is required to advertise the BD subnet out of the remote L3Out and ensure that inbound traffic to the BD can be maintained even if the local L3Out failed. In this case, you would also need to configure the subnet with the `Advertised Externally` flag. For more information, see the [Intersite L3Out, on page 259](#) use case.
- Step 5** Enable **Host Route**.
- This enables Host-Based Routing on the bridge domain. When this knob is enabled, the border leaf switches will also advertise individual endpoint (EP) host-routes (/32 or /128 prefixes) along with the subnet. The host-route information is advertised only if the host is connected to the local Pod. If the EP is moved away from the local Pod or when the EP is removed from EP database, the route advertisement is then withdrawn.
- Step 6** If necessary, change the **SVI MAC Address**.
- The SVI MAC addresses must be unique per site, when virtual MAC and virtual IP are enabled for Common Pervasive Gateway (CPG) scenario. This field can also be used when CPG is not enabled, which will change the default router MAC of the BD.
- Step 7** Add one or more **Subnets** for the BD.
- The concept is the same as adding subnets to the BD at the template level, except the subnets will be configured for the bridge domain on this specific site only.
- Click **+Add Subnet**.
An **Add New Subnet** window opens.
 - Enter the subnet's **Gateway IP** address and a **Description** for the subnet that you want to add.
 - Select the **Scope** for the subnet.
The network visibility of the subnet.
 - `Private to VRF`—The subnet applies only within its tenant.
 - `Advertised Externally`—The subnet can be exported to a routed connection.
 - (Optional) Enable **Shared Between VRFs**.

`Shared between VRFs`—The subnet can be shared with and exported to multiple contexts (VRFs) in the same tenant or across tenants as part of a shared service. An example of a shared service is a routed connection to an EPG present in another context (VRF) in a different tenant. This enables traffic to pass in both directions across contexts (VRFs). An EPG that provides a shared service must have its subnet that is configured under that EPG (not under a bridge domain), and its scope must be set to advertised externally, and shared between VRFs.

Shared subnets must be unique across the contexts (VRF) involved in the communication. When a subnet under an EPG provides a Layer 3 external network shared service, such a subnet must be globally unique within the entire ACI fabric.

- e) (Optional) Enable **No Default SVI Gateway**.

Enabling this option means that only the proxy route (subnet route to spine proxy) is programmed on the leaf switches and no SVI is created, which means SVI cannot be used as the gateway.

We recommend that SVI is created by the BD subnet as the gateway and the **No Default SVI Gateway** option is enabled on the EPG instead because EPG subnets should only be used for route leaking.

- f) (Optional) Enable **Querier**.

Enables IGMP Snooping on the subnet

- g) (Optional) Enable **Primary** option to designate the subnet as primary.

There can be one primary IPv4 subnet and one primary IPv6 subnet.

- h) Click **Save**.
-

Configuring Application Profiles and EPGs

This section describes how to configure an Application Profile and an EPG.

Before you begin

You must have the schema and template that is created and a tenant that is assigned to the template, as described in [Creating Schemas and Templates, on page 68](#).

This section also assumes you have a Contract and a Bridge Domain created.

Procedure

Step 1 Select the schema and template where you want to create the application profile.

Step 2 Create an application profile.

- a) In the main pane, select **+Create Object > Application Profile**.

Alternatively, you can scroll down to the **Application Profile** area and click **Create Application Profile**.

- b) In the right pane, provide the **Display Name** for the application profile.

You can create application profiles with the same name in different templates without any conflicts. You cannot however create other objects (such as VRFs, BDs, EPGs) with the same name in different templates if they will be deployed to the same site and tenant.

- c) (Optional) Provide a **Description**.

Step 3 Create an EPG.

- a) In the main pane, select **+Create Object > EPG**, then select the application profile where you want to create the EPG.

Alternatively, you can scroll down to the specific **Application Profile** area and click **Create EPG**.

- b) In the right pane, provide the **Display Name** for the EPG.

- c) (Optional) Provide a **Description**.

Step 4 (Optional) Add one or more **Annotations** for the EPG.

This allows you to add arbitrary `key:value` pairs of metadata to an object as annotations (`tagAnnotation`). Annotations are provided for any custom purposes that you may require, such as descriptions, markers for personal scripting or API calls, or flags for monitoring tools or orchestration applications such as your Cisco Nexus Dashboard Orchestrator. Because APIC ignores these annotations and merely stores them with other object data, there are no format or content restrictions that are imposed by APIC.

Step 5 Add a **Contract** for the EPG.

Creating contracts and filters is described in detail in [Configuring Contracts and Filters, on page 84](#). If you already have a contract that is created:

- a) Click **Add Contract**.
- b) On the **Add Contract** dialog, enter the contract name and type.
- c) Click **SAVE**.

Step 6 (Optional) Add an **Intra-EPG Contract** for the EPG.

By default, communication between endpoints in an EPG is open, unless you enable Intra-EPG isolation under the EPG policy configuration.

With an intra-EPG contract, you can specify which traffic is allowed within an EPG based on protocol, ports, and other options specified by the contract's filters.

- a) In the **Intra-EPG Contract** area, click **Add Contract**.
- b) On the **Add Contract** dialog, enter the contract name and type.
- c) Click **SAVE**.

Step 7 From the **Bridge Domain** drop-down, select the bridge domain for this EPG.

If you are configuring an on-premises EPG, you must associate it with a bridge domain.

Step 8 (Optional) Click + **Subnet** to add a subnet to your EPG.

You may choose to configure a subnet on the EPG level rather than the bridge domain level, for example for a VRF route-leaking use-case.

- a) On the **Add Subnet** dialog, enter the **Gateway IP** address and a description for the subnet you plan to add.
- b) In the **Scope** field select either **Private to VRF** or **Advertised Externally**.
- c) Click the check box for **Shared Between VRFs** if appropriate.
- d) Click the check box for **No Default SVI Gateway** if appropriate.
- e) Click **OK**.

Step 9 (Optional) Enable microsegmentation.

If you are configuring a microsegmentation EPG (uSeg), you must provide one or more uSeg attributes for matching endpoints to the EPG.

- a) Check the **uSeg EPG** check box.
- b) Click **+uSeg Attribute**.
- c) Provide the **Name** and **Type** for the uSeg attribute.
- d) Based on the attribute type you have selected, provide the attribute details.

For example, if you have selected `MAC` for the attribute type, provide the MAC address to identify an endpoint in this EPG.

- e) Click **SAVE**.

Step 10 (Optional) Enable intra-EPG isolation.

By default, endpoints in EPG can freely communicate with each other. If you want to isolate the endpoints from each other, set the isolation mode to **Enforced**.

intra-EPG endpoint isolation policies provide full isolation for virtual or physical endpoints; no communication is allowed between endpoints in an EPG that is operating with isolation enforced. Isolation-enforced EPGs reduce the number of EPG encapsulations required when many clients access a Common Service but are not allowed to communicate with each other.

Step 11 (Optional) Enable Layer 3 multicast for the EPG.

For additional information about Layer 3 multicast, see [Layer 3 Multicast, on page 295](#)

Step 12 (Optional) Enable preferred group membership for the EPG.

The Preferred Group feature allows you to include multiple EPGs within a single VRF to allow full communication between them with no need for contracts to be created. For additional information about EPG preferred group, see [EPG Preferred Groups Overview and Limitations, on page 225](#).

Step 13 Configure the EPG's site-local properties as necessary.

In addition to the template-level configurations, you can also define one or more site-local properties for the EPG, as described in [Configuring EPG's Site-Local Properties, on page 80](#).

Configuring EPG's Site-Local Properties

In addition to the template-level properties you typically configure for the object when you create it in a template, you can also define one or more properties that are specific to each site to which you assign the template.

When you deploy the object to more than 1 site, the same template-level configurations are deployed to all sites, while the site-local configurations are deployed to those specific sites only.

Before you begin

You must have:

- Created the application profile and EPG and configured the template-level properties, as described in [Configuring Application Profiles and EPGs, on page 78](#).
- Assigned the template that contains the EPG to one or more sites.

Procedure

Step 1 Open the schema that contains the template with the EPG.

Step 2 From the **View <Overview>** drop-down in the schema view, select the template that contains the EPG.

Step 3 In the template view's main pane, click the **<site-name>** tab to select site-specific properties for the template.

Step 4 In the main pane, click the EPG for which you want to update site-local properties.

This opens the EPG's properties pane. For most fields, you see the values you have configured at the template level, which you cannot edit here.

Step 5 Choose the **EPG Admin State**.

This field is available only if the EPG belongs to a tenant other than `infra` or `mgmt`.

When the EPG is in shutdown mode, the ACI policy configuration that is related to the EPG is removed from all the switches in the site. While the EPG still exists in the ACI Data Store, it is in inactive mode.

Step 6 Add one or more **Subnets** for the EPG.

- a) Click **+Add Subnet**.

An **Add New Subnet** window opens.

- b) Enter the subnet's **Gateway IP** address and a description for the subnet that you want to add.
c) Select the **Scope** for the subnet.

The network visibility of the subnet.

- `Private to VRF`—Prevents the subnet from being announced through L3Out toward an external network domain.
- `Advertised Externally`—The subnet can be announced through L3Out toward an external network domain.

- d) (Optional) Enable **Shared Between VRFs**.

`Shared between VRFs`—The subnet can be shared with and exported to multiple contexts (VRFs) in the same tenant or across tenants as part of a shared service. An example of a shared service is a routed connection to an EPG present in another context (VRF) in a different tenant. This enables traffic to pass in both directions across contexts (VRFs). An EPG that provides a shared service must have its subnet that is configured under the BD (not under the EPG), and its scope must be set to advertised externally, and shared between VRFs.

Shared subnets must be unique across the contexts (VRF) involved in the communication. When a subnet under an EPG provides a Layer 3 external network shared service, such a subnet must be globally unique within the entire ACI fabric.

- e) (Optional) Enable **No Default SVI Gateway**.

Enabling this option means that only the proxy route (subnet route to spine proxy) is programmed on the leaf switches and no SVI is created, which means SVI cannot be used as the gateway.

We recommend enabling this option on the EPG subnets, which should only be used for route leaking and leaving this option disabled on the BD subnets so that the SVI can be used as a gateway.

- f) Click **Ok** to save.

Step 7 Add one or more **Static ports**.

- a) Click **+Static Port**.
b) From the **Path Type** drop-down, select the type of port.
c) If configuring a physical interface, select the **Pod**
d) Choose whether you want to configure a single port or a range of ports.

For the interface configuration, you have an option to do it either by entering a single **Leaf** and a **Path** or by entering a range of **Leaf** for example, 120-125 and **Path** eg1/17-20. You will also have an option to enter a range of **Leaf** and associate it with one single **Path**, or enter a range of **Path** for one single **Leaf**.

However, after the configuration it will still be displayed as individual ports in the UI and will require individual changes for any future updates.

- e) Select the **Port Encap VLAN**.

When manually configuring the port encap on a domain for an EPG, the VLAN ID must belong to a static VLAN block within a dynamic VLAN pool.

If EPG is enabled for microsegmentation at the template level, when a **Primary MICRO-SEG VLAN** is configured, the **Port Encap VLAN** is configured as an Isolated Secondary VLAN for the Primary VLAN. Traffic is sent from the host to the leaf switch using the secondary VLAN and return traffic from the leaf switch to the host is sent using the primary VLAN.

- f) (Optional) Select the **Primary MICRO-SEG VLAN**.

The VLAN identifier for microsegmentation.

- g) (Optional) Select the **Deployment Immediacy**.

When policies are downloaded to the leaf nodes, deployment immediacy can specify when the policy is pushed into the hardware policy CAM:

- **Immediate**—Specifies that the policy is programmed in the hardware policy CAM when the policy is downloaded in the leaf switch software.
- **On Demand**—Specifies that the policy is programmed in the hardware policy CAM only when the first packet is received through the data path. This process helps to optimize the hardware space.

- h) (Optional) Select the **Mode**.

The mode of the static association with the path. EPG tagging sees configuring a static path under an EPG:

- **Trunk**—The default deployment mode. Select this mode if the traffic from the host is tagged with a VLAN ID.
- **Access (802.1p)**—Select this mode if the traffic from the host is tagged with a 802.1p tag. When an access port is configured with a single EPG in built-in 802.1p mode, its packets exit that port untagged. When an access port is configured with multiple EPGs, one in built-in 802.1p mode, and some with VLAN tags, all packets exiting that access port are tagged VLAN 0 for EPG configured in built-in 802.1p mode and for all other EPGs packets exit with their respective VLAN tags. Only one built-in 802.1p EPG is allowed per access port.
- **Access (Untagged)**—Select this mode if the traffic from the host is untagged (without VLAN ID). When a leaf switch is configured for an EPG to be untagged, for every port, this EPG uses, the packets exit the switch untagged. Note that when an EPG is deployed as untagged, do not deploy that EPG as tagged on other ports of the same switch.

Step 8 Add one or more **Static Leaf** nodes.

- a) Click **+Static Leaf**.
- b) From the **Leaf** drop-down, select the leaf node that you want to add.
- c) (Optional) In the **VLAN** field, provide the VLAN ID for tagged traffic.

Step 9 Add one or more **Domains**.

- a) Click **+Domain**.
- b) Select the **Domain Association Type**.

This is the type of the domain that you are adding:

- **VMM**
- **Fibre Channel**
- **L2 External**

- L3 External
- Physical

- c) Select the **Domain Profile** name.
- d) Select the **Deployment Immediacy**.

Deployment immediacy can specify when the policy is pushed:

- **Immediate**—Specifies that the policy is programmed in the hardware policy CAM when the policy is downloaded in the leaf switch software.
- **On Demand**—Specifies that the policy is programmed in the hardware policy CAM only when the first packet is received through the data path. This process helps to optimize the hardware space.

- e) Select the **Resolution Immediacy**.

Specifies whether policies are resolved immediately or when needed. The options are:

- **Immediate**—Specifies that EPG policies are pushed to the leaf switch nodes upon hypervisor attachment to the VMware vSphere Distributed Switch (VDS). LLDP or OpFlex permissions are used to resolve the hypervisor to leaf node attachments.
- **On Demand**—Specifies that EPG policies are pushed to the leaf switch nodes only when a hypervisor is attached to VDS and a VM is placed in the port group (EPG).
- **Pre-provision**—Specifies that EPG policies are pushed to the leaf switch nodes even before a hypervisor is attached to the VDS. The download preprovisions the configuration on the switch.

- f) For VMM domains, configure extra settings.

Beginning with release 4.2(1), you can configure several extra properties for VMM domains directly from your Cisco Nexus Dashboard Orchestrator.

- **Port Bindings** – You can choose one of the following options:

- Dynamic Binding
- Ephemeral
- Default
- Static Binding

For additional information about port binding, see the "Cisco ACI with VMware VDS Integration" chapter of the [Cisco ACI Virtualization Guide](#)

- **Netflow** – choose whether you want to enable NetFlow for the VMM domain.
- **Promiscuous Mode** – specifies whether to allow or reject unicast traffic that is not destined to the MAC addresses of the virtual machines attached to the trunk port group.
- **MAC Address Changes** – specifies whether to allow or reject MAC address changes for the network adapter within the VM.
- **Forged Transmits** – specifies whether to allow or reject forged transmits.

A forged transmit occurs when a network adapter starts sending out traffic that identifies itself as something else. This security policy compares the effective address of the virtual network adapter and the source address inside an 802.3 Ethernet frame generated by the virtual machine to ensure that they match.

- **Custom EPG Name** – allows you to provide a custom name for the EPG associated with this VMM domain.

When you associate an EPG to a VMM domain, APIC automatically creates a VMware vCenter port group or a Microsoft VM network.), it is easier to manage the port groups or VM networks because you now have the option of giving the EPG a custom name

Configuring Contracts and Filters

This section describes how to configure a contract, a filter, and assigns the filter to the contract. A filter is similar to an Access Control List (ACL), it is used to filter traffic through contracts that are associated to EPGs.

Procedure

Step 1 Select the schema and template where you want to create contract and filter.

You can create the contract in the same or different template as the objects (EPGs and external EPGs) to which you apply it. If the objects that use the contract are deployed to different sites, we recommend defining the contract in a template that is associated to multiple sites. However, this is not strictly required and even if the contract and filters are defined only as local objects in Site 1, NDO creates those objects in a remote Site 2 when a local EPG or external EPG in Site 2 must consume or provide that contract.

Step 2 Create a filter.

- In the main pane, select **+Create Object > Filter**.

Alternatively, you can scroll down to the **Filters** area, mouse over the tile, and click **Add Filter**.

- In the right pane, provide the **Display Name** for the filter.
- (Optional) Provide a **Description**.

Step 3 (Optional) Add one or more **Annotations**.

This allows you to add arbitrary `key:value` pairs of metadata to an object as annotations (`tagAnnotation`). Annotations are provided for any custom purposes that you may require, such as descriptions, markers for personal scripting or API calls, or flags for monitoring tools or orchestration applications such as your Cisco Nexus Dashboard Orchestrator. Because APIC ignores these annotations and merely stores them with other object data, there are no format or content restrictions that are imposed by APIC.

Step 4 Create a filter entry.

- In the right pane, click **+ Add Entry**.

The filter entry is a combination of network traffic classification properties. You can specify one or more options as described in the following step.

- Provide the **Name** for the filter.
- Choose the **Ether Type**.

For example, `ip`.

- d) Choose the **IP Protocol**.

For example, `icmp`.

- e) Choose the **Destination Port Range From** and **Destination Port Range To**.

The start and end of the destination ports range. You can define a single port by specifying the same value in both fields or you can define a range of ports from 0 to 65535. You can also choose to specify one of the server types instead of specific port numbers, for example `http`.

- f) Enable **Match only fragments** option.

When enabled, the rule applies to any IP fragment with an offset that is greater than 0 (all IP fragments except the first). When disabled, the rule will not apply to IP fragments with an offset greater than 0 because TCP/UDP port information can only be checked in initial fragments.

- g) Enable **Stateful** option.

When this option is enabled, any traffic coming from the provider back to the consumer will always have to have the `ACK` bit set in the packet or else the packets will be dropped.

- h) Specify **ARP flag** (Address Resolution Protocol).

The **ARP Flag** is used when creating a specific filter for ARP and allows you to specify ARP request or ARP reply.

- i) Choose the **Source Port Range From** and **Source Port Range To**.

The start and end of the source ports range. You can define a single port by specifying the same value in both fields or you can define a range of ports from 0 to 65535. You can also choose to specify one of the server types instead of specific port numbers, for example `http`.

- j) Specify **TCP session rules**.

TCP session rules are used when creating a filter for TCP traffic and allow you to configure `stateful` ACL behavior.

- k) Click **Ok** to save the filter.

- l) Repeat this step to create any additional filter entries for this filter.

You can create and assign multiple filter entries for each filter.

Step 5 Create a contract.

- a) In the main pane, select **+Create Object > Contract**.

Alternatively, you can scroll down to the **Contract** area, mouse over the tile, and click **Add Contract**.

- b) In the right pane, provide the **Display Name** for the contract.
- c) (Optional) Provide a **Description**.
- d) (Optional) Add one or more **Annotations**.

This allows you to add arbitrary `key:value` pairs of metadata to an object as annotations (`tagAnnotation`). Annotations are provided for any custom purposes that you may require, such as descriptions, markers for personal scripting or API calls, or flags for monitoring tools or orchestration applications such as your Cisco Nexus Dashboard Orchestrator. Because APIC ignores these annotations and merely stores them with other object data, there are no format or content restrictions that are imposed by APIC.

- e) Select the appropriate **Scope** for the contract.

Contract scope limits the contract's accessibility; the contract will not be applied to any consumer EPG outside the scope of the provider EPG:

- Application Profile
- VRF
- Tenant
- Global

- f) Toggle the **Apply both directions** knob if you want the same filter to apply for both consumer-to-provider and provider-to-consumer directions.

If you enable this option, you must provide the filters only when and they apply for traffic in both directions. If you leave this option disabled, you must provide two sets of filter chains, one for each direction.

Note

If you create and deploy a contract with **Apply both directions** enabled, you cannot simply disable the option and redeploy for the change to apply. To disable this option on an already deployed contract, you must delete the contract, deploy the template, then re-create the contract with the option that is disabled to correctly change the setting in your fabrics.

- g) (Optional) From the **Service Graph** drop-down, select a service graph for this contract.
 h) (Optional) From the **QoS Level** drop-down, select a value for this contract.

This value specifies the ACI QoS Level that will be assigned to the traffic using this contract. For more information, see [QoS Preservation Across IPN, on page 305](#).

If you leave this at `Unspecified`, the default QoS Level 3 is applied to the traffic.

Step 6 Assign the filters to the contract.

- a) In the main pane for template, select a contract. In the right pane, scroll down to the **Filter Chain** area and click + **Add Filter** to add a filter to the contract.
 b) In the **Add Filter Chain** window that opens, select the filter that you added in previous step from the **Name** drop-down list.
 c) Select the **Action** for the filter.

When adding filters, you can choose whether to permit or deny traffic that matches the filter criteria. For `deny` filters, you can set the priority of the filter to one of four levels: `default`, `low`, `medium`, or `high`; the `permit` filters always have the default priority. For more information on ACI contracts and filters, see [Cisco ACI Contract Guide](#).

- d) Click **Ok** to add the filter to the contract.
 e) If you disabled the `Apply both directions` option on the contract, repeat this step for the other filter chain.
 f) (Optional) You can create and assign multiple Filters to each Contract.

If you want to create extra filter for the same contract:

- Repeat Step 2 and Step 3 to create another filter along with its filter entries.
- Then repeat this step to assign the new filter to this Contract.

Viewing Schemas

After you have created one or more schemas, they are displayed both on the Dashboard and the Schemas page.

You can use the functionality available on these two pages to monitor the usage and the health of your schemas when they are deployed. You can also access and edit specific areas of the implemented schema policies using the Cisco Nexus Dashboard Orchestrator GUI.

Cloning Schemas

This section describes how to create a copy of an existing schema and all its templates using the "Clone Schema" feature in the **Schemas** screen.

**Note**

If you clone a template and attempt to deploy it to the same site with a different configuration, the deployment may fail due to a duplicate name error. Changing the object name in the cloned template only updates the display name. It does not alter the database record, which causes the deployment to fail in this scenario.

Procedure

Step 1 Log in to your Cisco Nexus Dashboard Orchestrator GUI.

Step 2 Choose the schema to clone.

- a) From the left navigation menu, select **Configure > Tenant Template**.
- b) From the action menu (...) menu next to the name of the schema you want to clone, select **Clone**.

Step 3 Provide the name for the new schema and click **Clone**.

Clone test-schema1

Schema Name

test-schema1-clone

Clone

After you click **Clone**, the UI will display `Cloning of <schema-name> was successful.` message and the new schema will be listed in the **Schemas** screen.

The new schema is created with the exact same templates (and their tenants' association), object, and policy configurations as the original schema.

Note that while the templates, objects, and configurations are copied, the site association is not preserved and you must reassociate the template in the cloned schema with any sites where you want to deploy them. Similarly, you must provide any site-specific configurations for the template objects after you associate it with the sites.

Step 4 (Optional) Verify that the schema and all its templates were copied.

You can verify that the operation completed successfully by comparing the two schemas.



CHAPTER 6

Fabric Management Templates

- [Fabric Management Templates](#), on page 89
- [Creating Fabric Policies](#), on page 90
- [Creating Fabric Resources Policies](#), on page 102
- [Creating Monitoring Policies](#), on page 107

Fabric Management Templates

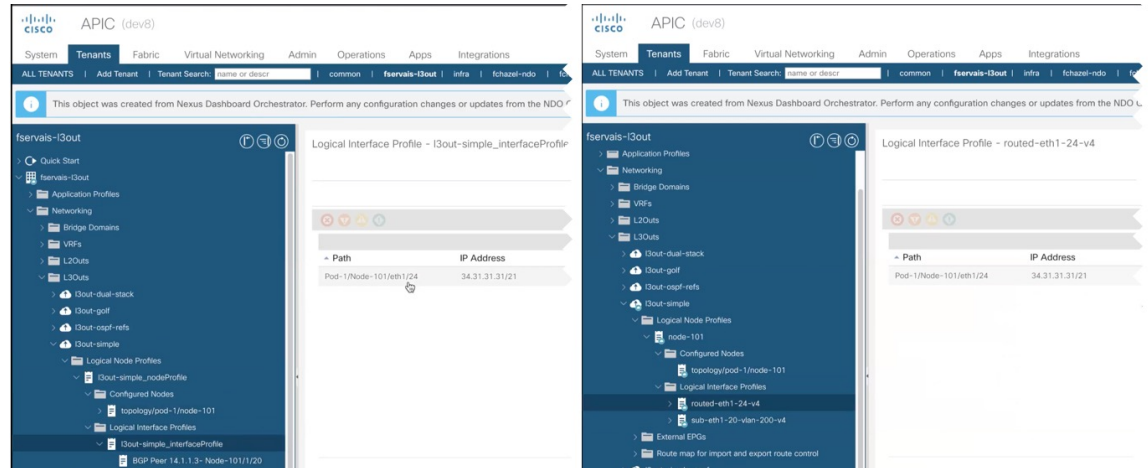
Beginning with release 4.0(1), Cisco Nexus Dashboard Orchestrator allows you to configure several fabric policies, fabric resource policies, and monitoring policies. Similarly to how you create objects and define configuration for VRFs, BDs, or EPGs using [Schemas and Application Templates](#), these new policies are defined in their respective template types. The following sections describe the policies that you can now configure directly from NDO and the steps that are required to do so.

While the objects you define in the Application templates map one-to-one to the same managed objects (MOs) in the site's APIC, the new template types group some objects and policies into logical containers. In these cases, after you define all the required policies within the same logical container in one of the new template types in NDO, individual policies are still created in the APIC when you deploy that configuration from the Orchestrator. For example, separate policies are created for the nodes, interfaces, and even IP address types, so providing IPv4 and IPv6 IP addresses for a single L3Out interface creates two separate interface profiles in the APIC:

The screenshot displays the NDO configuration interface. On the left is a navigation tree for a site named 'demo-jan12'. The tree is expanded to show 'Logical Interface Profiles' under 'L3Outs', with 'routed-eth1-8-v4' selected. The main panel on the right is titled 'Logical Interface Profile - routed-eth1-8-v4'. It has tabs for 'Policy', 'Faults', and 'History'. The 'Policy' tab is active, showing sub-tabs: 'General', 'Routed Sub-Interfaces', 'Routed Interfaces', 'SVI', and 'Floating SVI'. The 'Routed Sub-Interfaces' sub-tab is selected. Below the sub-tabs is a table with columns: 'Path', 'IP Address', 'Secor IP Address', 'MAC Address', 'MTU (bytes)', 'Encap', 'PTP', and 'MultiF Direct'. The table is currently empty, with a message below it stating 'No items have been found. Select Actions to create a new item.' At the bottom right of the panel are three buttons: 'Show Usage', 'Reset', and 'Submit'.



Note Because of how NDO maintains these logical containers with multiple individual policies, it also enforces specific best practices for the policy model in the APIC during template deployment. This can result in a scenario where if you import some previously existing configuration from the APIC into one of the new templates, edit the configuration, and then redeploy, the old MOs are removed and new ones are created with NDO-specific hierarchy, which may cause a brief (up to 1 second) traffic interruption:



This happens only if the imported objects are modified and redeployed. If you simply import the configuration and immediately redeploy it without any changes, NDO will simply take ownership of the MOs in the APIC but will not remove or recreate them.

Creating Fabric Policies

This section describes how to create one or more fabric policy templates. Fabric policy templates allow you to create and configure the following fabric policies:

- VLAN Pool
- Physical Domains
- L3 Domains
- SyncE Interface Policies
- Interface Settings
- Node Settings
- Pod Settings
- MACsec
- NTP Policies
- PTP Policies
- QoS DSCP Policies

- QoS SR-MPLS Policies
- QoS Class Policies
- MCP Global Policies

When creating Fabric Policy templates policies, consider the following:

- Fabric Policy templates do not need to be associated to any tenant, but must be mapped to at least one site to be deployed.
- The configuration of those policies is only possible at the template level and not at the specific site level.
- Undeploying a Fabric Policy template would result in preservation of the associated policies on APIC. In other words, the configuration of those policies on APIC won't be reverted to the default values or to the values that were configured on the APIC before the Orchestrator began managing them.

Procedure

Step 1 Log in to your Cisco Nexus Dashboard and open the Cisco Nexus Dashboard Orchestrator service.

Step 2 Create a new Fabric Policy template.

- From the left navigation pane, choose **Configure > Fabric Templates**.
- On the **Fabric Policy Templates** page, click **Create Fabric Policy Template**.
- In the **Fabric Policies** page's right properties sidebar, provide the **Name** for the template.

By default, the new template is empty, so you must add one or more fabric policies as described in the following steps. You don't have to create every policy available in the template – you can define one or more policies of each type to deploy along with this template. If you don't want to create a specific policy, simply skip the step that describes it.

Step 3 Assign the template to one or more sites.

The process for assigning Tenant Policy templates to sites is identical to how you assign application templates to sites.

- In the **Template Properties** view, click **Actions** and choose **Add/Remove Sites**.

The **Associate Sites to <template-name>** window opens.

- In the **Associate Sites** window, check the check box next to the sites where you want to deploy the template.

Only the on-premises ACI sites support tenant policy templates and will be available for assignment.

- Click **Ok** to save.

Step 4 Create a VLAN Pool.

A VLAN pool specifies the VLAN IDs or ranges that are used for VLAN encapsulation that the physical or VMM domains consume.

- From the **+Create Object** drop-down, select **VLAN Pool**.
- In the right properties sidebar, provide the **Name** for the policy.
- (Optional) Click **Add Description** and provide a description for the policy.
- Click **+Add VLAN Range**, provide the range, and click the check mark icon to save it.
- Repeat the previous substep to create any additional VLAN ranges within the same policy.
- Repeat this step to create any additional VLAN pools.

Step 5 Create a Physical Domain.

Physical domain profiles are typically used for bare metal server attachment and management access. A domain is configured to be associated with a VLAN pool. EPGs are then configured to use the VLANs associated with a domain.

- a) From the **+Create Object** drop-down, select **Physical Domain**.
- b) In the right properties sidebar, provide the **Name** for the domain.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) Click **Select a VLAN Pool Policy** and choose one of the VLAN pools for this domain.

The VLAN pool must be already created as described in Step 3.

- e) Repeat this step to create any additional Physical Domains.

Step 6 Create an L3 Domain.

An L3 Domain profile is a policy for managing the physical infrastructure, such as ports and VLANs that can be used to connect the ACI fabric to a Layer 3 routed outside network.

- a) From the **+Create Object** drop-down, select **L3 Domains**.
- b) In the right properties sidebar, provide the **Name** for the domain.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) (Optional) Click **Select a VLAN Pool Policy** and choose one of the VLAN pools for this domain.

If you plan to use point-to-point routed interfaces, no VLAN Pool is necessary and you can skip this step.

However, if you configure subinterfaces or SVIs, you must provide the required VLANs by adding a VLAN Pool. In this case, the VLAN pool must be already created as described in Step 3.

- e) Repeat this step to create any additional L3 Domains.

Step 7 Create a SyncE Interface Policy.

With Ethernet equipment gradually replacing Synchronous Optical Networking (SONET) and Synchronous Digital Hierarchy (SDH) equipment in service-provider networks, frequency synchronization is required to provide high-quality clock synchronization over Ethernet ports. Frequency or timing synchronization is the ability to distribute precision frequency around a network. Synchronous Ethernet (SyncE) provides the required synchronization at the physical level. In SyncE, Ethernet links are synchronized by timing their bit clocks from high-quality, stratum-1-traceable clock signals in the same manner as SONET/SDH.

For detailed information on SyncE in ACI fabrics, see the "[Synchronous Ethernet \(SyncE\)](#)" chapter of the *Cisco APIC System Management Configuration Guide* for your release.

- a) From the **+Create Object** drop-down, select **SyncE Interface Policy**.
- b) In the right properties sidebar, provide the **Name** for the policy.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) Provide policy details.

- **Admin State** – Enables or disables the policy.

Default is `disabled`.

- **Sync State Msg** – If unchecked, disables sending ESMC packets and also ignores any received ESMC packets.

- **Selection Input** – Enables configuration of the priority of the frequency source on an interface.

- **Src Priority** – The priority of the frequency source on an interface. This value is used in the clock-selection algorithm to choose between two sources that have the same QL.

Values can range from 1 (highest priority) to 254 (lowest priority). The default value is 100.

Can be configured only if **Selection Input** is enabled.

- **Wait To Restore** – The wait-to-restore time, in minutes, is the amount of time after the interface comes up before it is used for frequency synchronization on an interface. Values can range 0–12 minutes. The default value is 5.

Can be configured only if **Selection Input** is enabled.

- Repeat this step to create any additional SyncE Interface policies.

Step 8

Create an Interface Settings policy.

If you want to configure SyncE or MACsec for this interface, you must have those policies that are already created as described in corresponding steps.

An interface settings policy allows you to define a set of common interface settings which you can later deploy to one or more ports on one or more switches for consistent configuration across them.

- From the **+Create Object** drop-down, select **Interface Settings**.
- Choose the **Type** of the interface that you are configuring.
- In the right properties sidebar, provide the **Name** for the policy.
- (Optional) Click **Add Description** and provide a description for the policy.
- Provide policy details.
 - **Speed** – The data transfer rate for the port. This should match the destination to which the port is linked. The speed can be changed only for certain ports, and not all speeds are available on all systems. For more information, see the Hardware Installation Guide for your specific switch node.

- **Auto-Negotiation** – Enables autonegotiation for port.

- **VLAN Scope** – The Layer 2 Interface VLAN scope.

Global scope – Sets the VLAN encapsulation value to map only to a single EPG per leaf switch.

Port Local scope – Allows allocation of separate (port, VLAN) translation entries in both ingress and egress directions. This configuration is not valid when the EPGs belong to a single bridge domain.

- **CDP Admin State** – Enables Cisco Discovery Protocol (CDP) on the interface.
- **LLDP** – Enables Link Layer Discovery Protocol (LLDP) on the interface
- **MCP Admin State** – Enables mis-cabling protocol (MCP) on the interface.
- **Domains** – Choose one or more domains with which you want to associate this interface policy.

Specifying a domain is not mandatory, the interface policy can be created and deployed to sites without an associated domain.

- **Advanced Settings** – Click the arrow next to this section to expand.
 - **SyncE** – If you want to have a SyncE policy that is defined and want to assign it to this interface settings policy, select it from the drop-down.
 - **Debounce Interval** – The port debounce time is the amount of time that an interface waits to Notify the supervisor of a link going down. During this time, the interface waits to see if the link comes back up.
 - **Bring Up Delay** – Specifies a time in milliseconds that the decision feedback equalizer (DFE) tuning is delayed when a port is coming up. The delay is used to help avoid CRC errors during link bringup when using some third-party adapters.

You should set the delay only as required; usually, you do not need to set a delay.

- **FEC** – Forwarding Error Correction (FEC) is a method of obtaining error control in data transmission over an unreliable or noisy channel in which the source (transmitter) encodes the data in a redundant way using Error Correcting Code and the destination (receiver) recognizes it and corrects the errors without needing a retransmission.
- **QinQ** – Enables mapping double-tagged VLAN traffic ingressing on a regular interface, computer, or vPC to an EPG. When this feature is enabled and double-tagged traffic enters the network for an EPG, both tags are processed individually in the fabric and restored to double-tags when egressing the ACI switch. Ingressing single-tagged and untagged traffic is dropped.
- **Reflective Relay** – Forwards all traffic to an external switch, which then applies policy and sends the traffic back to the destination or target VM on the server as needed. There is no local switching. For broadcast or multicast traffic, reflective relay provides packet replication to each VM locally on the server.

One benefit of reflective relay is that it leverages the external switch for switching features and management capabilities, freeing server resources to support the VMs. Reflective relay also allows policies that you configure on the Cisco APIC to apply to traffic between the VMs on the same server.

In the Cisco ACI, you can enable reflective relay, which allows traffic to turn back out of the same port it came in on. You can enable reflective relay on individual ports, port channels, or virtual port channels as a Layer 2 interface policy.

The default value is disabled.

- **LLDP Transmit State** – Allows Link Layer Discovery Protocol (LLDP) packets to be sent from the interface.

LLDP Receive/Transmit State flags can be configured only if LLDP is globally enabled in the Interface policy.

- **LLDP Receive State** – Allows LLDP packets to be received by the interface.
- **BPDU Filter** – Bridge Protocol Data Unit (BPDU) filter filters out any BPDUs on the port.
BPDU Filter prevents both inbound and outbound BPDUs – the received BPDUs are dropped and no BPDUs are sent out.
- **BPDU Guard** – BPDU guard prevents the port from receiving BPDUs; if any BPDUs are received on the port, the port is put into `errdisable` mode.
- **LLFC Transmit State** – Allows Link Level Flow Control (LLFC) packets to be sent from the interface.
- **LLFC Receive State** – Allows LLFC packets to be received by the interface.
- **Access MACsec Policy** – If you want to have an access MACsec policy that is defined and want to assign it to this interface settings policy, select it from the drop-down.

f) Repeat this step to create any additional Interface Settings policies.

Step 9

Create a Node Settings policy.

A node settings policy allows you to define a set of common node settings which you can later deploy to one or more switches for consistent configuration across them.

In this release, the Node Settings policy supports enabling SyncE and PTP functionalities.

a) From the **+Create Object** drop-down, select **Node Settings**.

- b) In the right properties sidebar, provide the **Name** for the policy.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) If you want to deploy SyncE configuration to the nodes, enable **SyncE** and provide the settings.

For detailed information about SyncE, see the "Synchronous Ethernet (SyncE)" chapter of the [Cisco APIC System Management Configuration Guide](#) for your release.

- **Admin State** – Enables or disables the policy.
- **Quality Level Option** – Specifies the accuracy of the clock. This information is transmitted across the network using SSMs carried by ESMC and is used to determine the best available source to which the devices in the system can synchronize.

- e) If you want to deploy PTP configuration to the nodes, enable **PTP** and provide the settings.

For detailed information about PTP, see the "Precision Time Protocol" chapter of the [Cisco APIC System Management Configuration Guide](#) for your release.

- f) Repeat this step to create any additional Node Settings policies.

Step 10

Create a Pod Settings policy.

Before you can create a Pod settings policy, you must have an NTP policy that is already created for it as described in the corresponding step.

If you want to configure a Pod-wide MACsec policy, you must have a MACsec policy already created as described in the corresponding steps.

A Pod settings policy allows you to define a set of common Pod settings which you can later deploy to one or more Pods in your fabric for consistent configuration across them.

- a) From the **+Create Object** drop-down, select **Pod Settings**.
- b) In the right properties sidebar, provide the **Name** for the policy.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) Click **Select a NTP Policy** and choose the NTP policy.
- e) From the **Fabric MACsec Policy** drop-down, select the MACsec policy.
- f) Repeat this step to create any additional Pod Settings policies.

Step 11

Create a MACsec policy.

MACsec provides MAC-layer encryption over wired networks by using out-of-band methods for encryption keying. The MACsec Key Agreement (MKA) Protocol provides the required session keys and manages the required encryption keys.

For detailed information on MACsec in ACI fabrics, see the "MACsec" chapter of the [Cisco APIC System Management Configuration Guide](#) for your release.

- a) From the **+Create Object** drop-down, select **MACsec**.
- b) In the right properties sidebar, provide the **Name** for the policy.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) Provide policy details.

- **Type** – Defines the type of the interfaces that this policy will be applied to.

All the links on a spine switch are considered to be fabric links. However, if a spine switch link is used for IPN connectivity, then this link will be treated as an access link. This means that a MACsec access policy must be used to deploy MACsec on these links.

- **Admin State** – Enables or disables the policy.
- **Cipher Suite** – When selecting the cipher suite AES 128 or AES 256 without Extended Packet Numbering (XPN), you must explicitly specify the Security Association Key (SAK) expiry time. Leaving the SAK expiry time value at the default ("disabled") can cause interfaces to go out of service randomly.
- **Window Size** – A replay window is necessary to support the use of MACsec over provider networks that reorder frames. Frames within the window can be received out of order, but are not replay that is protected. The default window size is 64. The replay window size can be configured in the range of 0 to 232–1 if you use the Cisco APIC GUI or CLI. If you use an XPN cipher suite, the maximum replay window size is 230–1, and if you configure a higher window size, the window size gets restricted to 230–1. If you change the cipher suite to a non-XPN cipher suite, then there is no restriction and the configured window size is used.
- **Security Policy** – APIC MACsec supports two security modes. The MACsec `Must-Secure` only allows encrypted traffic on the link while the `Should-Secure` allows both clear and encrypted traffic on the link. For example, a port can turn on MACsec in `Must-Secure` mode before its peer has received its keychain resulting in the link going down. To address this issue, the recommendation is to deploy MACsec in `Should-Secure` mode and when all the links are up then change the security mode to `Must-Secure`.

Note

Before deploying MACsec in `Must-Secure` mode, the keychain must be deployed on the affected interface or the interface will go down.

- **SAK Expiry Time** – When selecting the cipher suite AES 128 or AES 256 without Extended Packet Numbering (XPN), you must explicitly specify the Security Association Key (SAK) expiry time. Leaving the SAK expiry time value at the default can cause interfaces to go out of service randomly.
- **Key Name** – Allows you to create a MACsec key. APIC is responsible for the MACsec keychain distribution to all the nodes in a Pod or to particular ports on a node.
 - Click **+Add MACsec Key**.
 - Provide the **Key Name**.
 - Provide the pre-shared key in the **PSK** field.
 - In the **Start Time** field, provide the date for the key to become valid.
 - In the **End Time** field, provide the date for the key to expire.
 - Click **Ok** to save the key.
 - Repeat the steps for any additional keys that you want to provide.

e) Repeat this step to create any additional MACsec policies.

Step 12

Create an NTP Settings Policy.

Within the ACI fabric, time synchronization is a crucial capability upon which many of the monitoring, operational, and troubleshooting tasks depend. Clock synchronization is important for proper analysis of traffic flows and for correlating debug and fault timestamps across multiple fabric nodes.

For detailed information on NTP in ACI fabrics, see the "Provisioning Core ACI Cisco Fabric Services" chapter of the [Cisco APIC Basic Configuration Guide](#) for your release.

- From the **+Create Object** drop-down, select **NTP Settings**.
- In the right properties sidebar, provide the **Name** for the policy.

- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) Provide policy details.

- Click **+Add Key** to provide the NTP client authentication key.
- **Advanced Settings** – Click the arrow next to this section to expand.

- **Admin State** – Enables or disables the NTP policy.

- **Server State** – Enables the ACI leaf switches to act as NTP servers to provide NTP information to downstream clients.

When enabled, the downstream clients can use the in-band/out-of-band management IP address of the leaf switch to which they are connected as the NTP server.

- **Master Mode** – Enables the designated NTP server to provide undisciplined local clock time to downstream clients with a configured stratum number. For example, a leaf switch that is acting as the NTP server can provide undisciplined local clock time to leaf switches acting as clients. This is only applicable when the server clock is undisciplined.

- **Stratum** – Specifies the stratum level from which NTP clients get their time synchronized.

If the **Server State** option is enabled and clients that are connected to an ACI leaf switch are configured to use the switch's management IP address as NTP server, they receive NTP information with stratum+1.

The range is 1–14.

- **Authentication State** – Enables certificate-based authentication.

If you enable this option, you must provide the key using the **+Add Key** option above.

- Click **+Add Provider** to specify the NTP server information.

In the **Add Provider** window that opens, you must provide the server's **Host name / IP Address**, the name of the **Management EPG**, and the **Management EPG Type**.

Note

The management EPG with the specific type that you choose must be already configured in the APICs of the sites with which this template is associated.

If you are creating multiple providers, check the **Preferred** option for the most reliable NTP source.

- e) Repeat this step to create any additional NTP Settings policies.

Step 13

Create a PTP Settings policy.

The Precision Time Protocol (PTP) is a time synchronization protocol for nodes that are distributed across a network. With PTP, you can synchronize distributed clocks with an accuracy of less than 1 microsecond using Ethernet networks. PTP's accuracy comes from the hardware support for PTP in the ACI fabric spine and leaf switches.

For detailed information on PTP in ACI fabrics, see the "[Precision Time Protocol](#)" chapter of the *Cisco APIC System Management Configuration Guide* for your release.

- a) From the **+Create Object** drop-down, select **PTP Settings**.
- b) In the right properties sidebar, provide the **Name** for the policy.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) Provide policy details.

- **Admin State** – Enables or disables the policy.

- **Global Priority 1** – Specifies a value that is used when advertising this clock. Priority 1 overrides the default criteria (such as clock quality and clock class) for best primary clock selection.

Valid values are from 0 to 255. The default is 128. Lower values take precedence.

- **Global Priority 2** – Specifies a value that is used when advertising this clock. Priority 2 is used as a tie-breaker between two devices that are otherwise equally matched in the default criteria.

Valid values are 0–255 . The default is 128. Lower values take precedence.

- **Global Domain** – Specifies the PTP domain number. Although multiple PTP domains are not supported in Cisco ACI, you can still change the domain number in use. The same value is used on all leaf switch and spine switches.

Valid values are 0–128. The default is 0.

- **Fabric Profile Template** – Specifies the PTP profile that defines the default values for the interval settings below. Profiles are used to define various parameters that are optimized for different use cases of PTP. Some of those parameters include, but not limited to, the appropriate range of PTP message intervals and the PTP transport protocols. A PTP profile is defined by many organizations/standards in different industries.

- **AES67-2015**: AES67-2015, which is the standard for audio over Ethernet and audio over IP interoperability.
- **Default**: IEEE 1588-2008, which is the default PTP profile for clock synchronization.
- **SMPTE-2059-2**: SMPTE ST2059-2015, which is the standard for video over IP.
- **Telecom-8275-1**: ITU-T G.8275.1, which is the standard recommendation for telecommunications with Full Timing Support.

Full Timing Support is the term that is defined by ITU to describe a telecommunication network that can provide devices with PTP G.8275.1 profile on every hop. G.8275.2, which is not supported by ACI, is for Partial Timing Support that may have devices in the path that do not support PTP.

- **Fabric Announce Interval** – Specifies the logarithm of the mean interval in seconds with base 2 for a primary port to send announce messages. The range depends on the chosen profile.
- **Fabric Sync Interval** – Specifies the logarithm of the mean interval in seconds with base 2 for a primary port to send synchronization messages. The range and the default depend on the chosen PTP profile.
- **Fabric Delay Interval** – Specifies the logarithm of the mean interval in seconds with base 2 for a slave port to send delay request messages. The range depends on the chosen PTP profile.
- **Fabric Announce Timeout** – Specifies the number of announce messages that the system waits before the PTP announce that message is considered expired. The range and the default depend on the chosen PTP profile.
- **Advanced Settings** – Click the arrow next to this section to expand.

how are profiles added here different from the one selected above?

1. Click **+Add Profile** to add a PTP profile. Profiles are used to define various parameters that are optimized for different use cases of PTP. Some of those parameters include, but not limited to, the appropriate range of PTP message intervals and the PTP transport protocols. A PTP profile is defined by many organizations/standards in different industries.
2. In the **Add Profile** dialog, provide the **Name**.
3. From the **Profile Template** drop-down, select one of the available profiles.

Detailed information about the profiles is available in the [Cisco APIC System Management Configuration Guide](#).

4. Update the default profile values as required by your specific use case.

- e) Repeat this step to create any additional PTP Settings policies.

Step 14

Create a QoS DSCP policy.

This policy is part of the overarching QoS preservation across IPN use case. You can use the information in this section as a reference, but we recommend following the full set of steps described in the [QoS Preservation Across IPN, on page 305](#) chapter of the *Features and Use Cases* section of this document.

- a) From the **+Create Object** drop-down, select **QoS DSCP**.
- b) In the right properties sidebar, provide the **Name** for the policy.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) Provide policy details.

- **Admin State** – Enables or disables the policy.
- **Advanced Settings** – Click the arrow next to this section to expand.

Choose the DSCP value for each ACI QoS level. Each drop-down contains the default list of available DSCP values. You must choose a unique DSCP value for each level.

- e) Repeat this step to create any additional QoS DSCP policies.

Typically, we recommend applying this policy consistently across all sites that are part of your Multi-Site domain.

Step 15

Create a QoS SR-MPLS policy.

This policy is part of the overarching SR-MPLS use case. You can use the information in this section as a reference, but we recommend following the full set of steps that are described in the [Multi-Site and SR-MPLS L3Out Handoff, on page 337](#) chapter of the *Features and Use Cases* section of this document.

- a) From the **+Create Object** drop-down, select **QoS SR-MPLS**.
- b) In the right properties sidebar, provide the **Name** for the policy.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) Click **+Add Ingress Rule** to add an ingress QoS translation rule.

These rules are applied for traffic that is ingressing the ACI border leaf switches from an MPLS network and are used to map incoming packet's experimental bits (EXP) values to ACI QoS levels, as well as to set differentiated services code point (DSCP) or CoS values for the original traffic. To ensure that the specified CoS value is used for traffic egressing an ACI leaf node, you must also configure CoS preservation feature as part of the "QoS Class Policies".

If a custom policy is not defined or not matched, default QoS Level (`Level 3`) is assigned.

1. In the **Match EXP From** and **Match EXP To** fields, specify the EXP range of the ingressing MPLS packet you want to match.
2. From the **Queuing Priority** drop-down, select the ACI QoS Level to map.

This is the QoS Level that you want to assign for the traffic within ACI fabric, which ACI uses to prioritize the traffic within the fabric. The options range from `Level 1` to `Level 6`. The default value is `Level 3`. If you do not make a selection in this field, the traffic will automatically be assigned a `Level 3` priority.

3. From the **Set DSCP** drop-down, select the DSCP value to assign to the traffic when it will be sent out of the destination ACI leaf switch.

The DSCP value specified is set in the original traffic that is received from the external network, so it will be reexposed only when the traffic is VXLAN decapsulated on the destination ACI leaf node.

If you set the value to *Unspecified*, the original DSCP value of the packet will be retained.

4. From the **Set CoS** drop-down, select the CoS value to assign to the traffic when it will be sent out of the destination ACI leaf switch.

The CoS value that is specified is set for traffic egressing the destination ACI leaf switch. This requires CoS preservation to be enabled.

If you set the value to *Unspecified*, the original CoS value of the packet will be retained, but only if the CoS preservation option is enabled in the fabric. For more information about CoS preservation, see [Cisco APIC and QoS](#).

5. Click the check mark icon to save the rule.
6. Repeat these steps for any additional ingress QoS policy rules within the same policy.

- e) Click **Add Egress Rule** to add an egress QoS translation rule.

These rules are applied on the border leaf switches for the traffic that is leaving the ACI fabric through an MPLS L3Out and are used to match the DSCP value of the packet and, if a match is found, set the MPLS EXP and CoS values based on the configured policy.

If a custom policy is not defined or not matched, the default EXP value of 0 is marked on all labels. EXP values are marked in both, default and custom policy scenarios, and are done on all MPLS labels in the packet.

Custom MPLS egress policy can override existing EPG, L3Out, and Contract QoS policies.

1. Using the **Match DSCP From** and **Match DSCP To** dropdowns, specify the DSCP range that you want to match for assigning the egressing MPLS packet's priority.
2. From the **Set MPLS EXP** drop-down, select the EXP value that you want to assign to the egressing MPLS packet.
3. From the **Set CoS** drop-down, select the CoS value that you want to assign to the egressing MPLS packet.
4. Click the check mark icon to save the rule.
5. Repeat these steps for any additional egress QoS policy rules.

- f) Repeat this step to create any additional QoS SR-MPLS policies.

Step 16

Create a QoS Class Policies policy.

Cisco ACI provides several user-configurable QoS levels. Cisco APIC, Release 4.0(1) and latter supports 6 user-configurable QoS levels, while earlier releases supported 3. This step describes how to configure specific settings for each of these levels using the Cisco Nexus Dashboard Orchestrator.

For detailed information about QoS functionality in ACI fabrics, see [Cisco APIC and QoS](#).

The most common use case for these policies is to enable CoS preservation for traffic coming into your ACI fabric.

- a) From the **+Create Object** drop-down, select **QoS Class Policies**.
- b) In the right properties sidebar, provide the **Name** for the policy.
- c) (Optional) Click **Add Description** and provide a description for the policy.

- d) If necessary, enable **Preserve CoS**.

When traffic enters the ACI fabric, based on the configured QoS policy, each packet can be mapped to an ACI QoS level. These QoS levels are then stored in the CoS field and DE bit of the packet's outer header while the original headers are discarded. If you want to preserve the original CoS values of the ingressing packets and restore it when the packet leaf switches the fabric, you can use this setting to enable the 802.1p Class of Service (CoS) preservation.

- e) Click **+Add Level** to define configuration details for a specific QoS class.

The **Add QoS Level Configuration** windows opens.

- f) In the **Add QoS Level Configuration** window, choose the **QoS Level** you want to configure and provide the configuration details.

- **MTU** – The maximum transmission unit to be used for packets of this QoS class.
- **Minimum Buffer** – The minimum number of reserved buffers. The number can be from 0 through 3.
The default value is 0.
- **Congestion Algorithm** – The congestion algorithm used for this QoS Level.
- **Scheduling Algorithm** – The scheduling algorithm used for this QoS Level.
- **Bandwidth Allocated** – The percentage of total bandwidth allocated to this QoS Level. The value can be from 0 through 100.
The default value is 20.
- **PFC Admin State** – The administrative state of the Priority Flow Control policy that is applied to FCoE traffic.
- **Admin State** – Enables or disables the policy.
- **No Drop Cos** – The CoS level to impose no drop FCoE packet handling even in case of FCoE traffic congestion.
- **PFC Scope** – The Priority Flow Control (PFC) scope. `Fabric-wide PFC` for the entire fabric or `IntraTor PFC` for the spine switches only.

- g) Repeat this step to create any additional QoS Class policies.

Step 17

Create an MCP Global Policy.

The mis-cabling protocol (MCP) is designed to handle misconfiguration that Link Layer Discovery Protocol (LLDP) and Spanning Tree Protocol (STP) are unable to detect. MCP uses a Layer 2 packet to detect and disable ports that form loops in the external infrastructure. You can use MCP packets to detect loops that involve the leaf switches and raise faults and events in the fabric when that occurs. MCP can be enabled globally or per-interface. By default, MCP is disabled globally and is enabled on each port, however you must enable MCP globally for it to function.

Note

If you configure and deploy an MCP global policy to one or more fabrics and then undeploy the template, the policy remains in the sites.

- a) From the **+Create Object** drop-down, select **MCP Global Policy**.

Only a single MCP Global policy can be created.

- b) In the right properties sidebar, provide the **Name** for the policy.
c) (Optional) Click **Add Description** and provide a description for the policy.
d) Enable **Admin State** to enable the policy.

- e) Enable **MCP PDU per VLAN**.

This enables MCP to send packets on a per-EPG basis. If this option is disabled, the packets will only be sent on untagged EPGs which allows detecting loops in the native VLAN only.

- f) If you have enabled the **Admin State**, provide a **Key** to uniquely identify the MCP packets within the fabric.
g) If necessary, update the **Loop Detect Multiplication Factor** value.

This specifies the number of MCP packets that will be received by the ACI fabric before the Loop Protection Action occurs.

- h) (Optional) Modify the additional MCP settings.

- **Initial Delay Time** – the time before MCP starts taking action. From the system start until the Initial Delay Timer timeout, MCP will only create a syslog entry if a loop is detected.
- **Transmission Frequency** – the transmission frequency of the MCP packets.

Step 18 Click **Save** to save the changes you've made to the template.

Step 19 Click **Deploy** to deploy the template to the associated sites.

The process for deploying tenant policy templates is identical to how you deploy application templates.

If you have previously deployed this template but made no changes to it since, the **Deploy** summary indicates that there are no changes, and you can choose to redeploy the entire template. In this case, you can skip this step.

Otherwise, the **Deploy to sites** window shows you a summary of the configuration differences that will be deployed to sites. Note that in this case only the difference in configuration is deployed to the sites. If you want to redeploy the entire template, you must deploy once to sync the differences, and then redeploy again to push the entire configuration as described in the previous paragraph.

Creating Fabric Resources Policies

This section describes how to create one or more fabric resources templates. Fabric resources templates allow you to create and configure the following:

- Physical Interfaces
- Port Channel Interfaces
- Virtual Port Channel Interfaces
- Node Profiles
- Pod Profiles
- FEX Device

Before you begin

- Most of the fabric resource policies require one or more fabric policies, so you must have those fabric policies that are already defined as described in [Creating Fabric Policies, on page 90](#).

For example, when creating an interface policy (physical, port channel, or virtual port channel), you must have an interface settings policy already created.

- The templates that contain the fabric policies that are required for the fabric resources policies must be deployed before any of the fabric resource policy templates.
- The Fabric Resource Policies templates do not need to be associated to any tenant, but must be mapped to at least one site to be deployed.
- In a typical deployment, we recommend that a separate Fabric Resource Policies template is associated to each site that is part of the Multi-Site domain.

In this case, we also recommend provisioning the configuration of the associated policies always at the global template level and not at the site level.

Procedure

Step 1 Log in to your Cisco Nexus Dashboard and open the Cisco Nexus Dashboard Orchestrator service.

Step 2 Create a new Fabric Resources Policy template.

- a) From the left navigation pane, choose **Configure > Fabric Template**.
- b) On the **Fabric Resource Templates** page, click **Add Fabric Resource Template**.
- c) In the **Resource Policies** page's right properties sidebar, provide the **Name** for the template.

By default, the new template is empty, so you must add one or more fabric policies as described in the following steps. You don't have to create every policy available in the template – you can define one or more policies of each type to deploy along with this template. If you don't want to create a specific policy, simply skip the step that describes it.

Step 3 Assign the template to one or more sites.

The process for assigning Tenant Policy templates to sites is identical to how you assign application templates to sites.

- a) In the **Template Properties** view, click **Actions** and choose **Add/Remove Sites**.

The **Associate Sites to <template-name>** window opens.

- b) In the **Associate Sites** window, check the check box next to the sites where you want to deploy the template.

Only the on-premises ACI sites support tenant policy templates and will be available for assignment.

- c) Click **Ok** to save.

Step 4 Create a Physical Interfaces policy.

Before you can create a physical interfaces policy, you must have an Interface Settings (**Physical**) policy that is already created for it as described in [Creating Fabric Policies, on page 90](#).

- a) From the **+Create Object** drop-down, select **Physical Interface**.
- b) In the right properties sidebar, provide the **Name** for the policy.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) In the **Nodes** field, provide one or more node IDs where this physical interface policy will be deployed.

The configuration of the nodes policy can be also done at the site-local view of the template. In that case, the site level configuration would override the global template level configuration. As previously mentioned, in the specific scenario where a different template is created and associated to each site that is part of the Multi-Site domain, we recommend that you configure the nodes policy only at the global template level.

For example, 101, 102, 103.

- e) In the **Interfaces** field, provide the interface names where the policy will be deployed.

For example, 1/1, 1/2-4, 1/5.

- f) Choose whether the interface is a **Physical** or a **Breakout** interface.
- g) If you are configuring a **Physical** interface, click **Select Physical Policy** and choose the interface settings policy that you created for this.

The interface settings that are defined in the interface settings policy will be applied to the interfaces (1/1, 1/2-4, 1/5) on the nodes (101, 102, 103) you provided in the previous substeps.

- h) If you are configuring a **Breakout** interface, choose the **Breakout Mode** for it.

This release supports 4x10G, 4x25G, and 4x100G modes.

- i) Repeat this step to create any additional Physical Interfaces policies.

A different policy could be needed, for example, when a unique set of physical interfaces should be configured on each node. In that case, you would define a unique Physical Interfaces policy for each specific node.

Step 5

Create a Port Channel Interfaces policy.

Before you can create a Port Channel interfaces policy, you must have an Interface Settings (PC/VPC) policy that is already created for it as described in [Creating Fabric Policies, on page 90](#).

- a) From the **+Create Object** drop-down, select **Port Channel Interface**.
- b) In the right properties sidebar, provide the **Name** for the policy.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) In the **Node** field, provide the node ID of the switch where this physical interface policy will be deployed.

The configuration of the nodes policy can be also done at the site-local view of the template. In that case, the site level configuration would override the global template level configuration. As previously mentioned, in the specific scenario where a different template is created and associated to each site that is part of the Multi-Site domain, we recommend that you configure the nodes policy only at the global template level.

For example, 104.

- e) In the **Interfaces** field, provide the interface names of the interfaces that are part of the port channel.

For example, 1/6, 1/7.

- f) Click **No selected PC/VPC Policy** and choose the interface settings policy that you created for this.

The Port Channel settings that are defined in the interface settings policy will be applied to the interfaces (1/6, 1/7) on the node (104) you provided in the previous substeps.

- g) Repeat this step to create any additional Port Channel Interfaces policies.

A different policy could be needed, for example, when a unique set of Port Channel interfaces should be configured on each node. In that case, you would define a unique Port Channel Interfaces policy for each specific node.

Step 6

Create a Virtual Port Channel Interfaces policy.

Before you can create a Virtual Port Channel interfaces policy, you must have an Interface Settings (PC/VPC) policy that is already created for it as described in [Creating Fabric Policies, on page 90](#).

- a) From the **+Create Object** drop-down, select **Virtual Port Channel Interface**.
- b) In the right properties sidebar, provide the **Name** for the policy.

- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) In the **Node 1** field, provide the node ID of the first switch that contains the interfaces that are part of the virtual port channel.
For example, 105.
- e) In the **Interfaces on Node 1** field, provide the interfaces on the first switch.
For example, 1/8, 1/9.
- f) In the **Node 2** field, provide the node ID of the second switch that contains the interfaces that are part of the virtual port channel.

The configuration of the nodes policy can be also done at the site-local view of the template. In that case, the site level configuration would override the global template level configuration. As previously mentioned, in the specific scenario where a different template is created and associated to each site that is part of the Multi-Site domain, we recommend that you configure the nodes policy only at the global template level.
For example, 106.
- g) In the **Interfaces on Node 2** field, provide the interfaces on the second switch.
For example, 1/8, 1/9.
- h) Click **No selected PC/VPC Policy** and choose the interface settings policy that you created for this.

The Port Channel settings that are defined in the interface settings policy will be applied to the interfaces on the nodes you provided in the previous substeps.
- i) Repeat this step to create any additional Virtual Port Channel Interfaces policies.

Step 7

Create a Node Profiles policy.

Before you can create a node profile policy, you must have a node settings policy that is already created for it as described in [Creating Fabric Policies, on page 90](#).

In this release, a node settings policy can be used to enable SyncE or PTP functionalities.

- a) From the **+Create Object** drop-down, select **Node Profile**.
- b) In the right properties sidebar, provide the **Name** for the policy.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) In the **Nodes** field, provide the node IDs of the switches where you want to deploy this node profile policy.

The configuration of the nodes policy can be also done at the site-local view of the template. In that case, the site level configuration would override the global template level configuration. As previously mentioned, in the specific scenario where a different template is created and associated to each site that is part of the Multi-Site domain, we recommend that you configure the nodes policy only at the global template level.

- e) Click **No selected Node Policy** and choose the node settings policy that you created for this.

The node settings that are defined in the node settings policy will be applied to all nodes you provided in the previous substep.

Only a single node settings policy can be referenced in a given node profile. This means that if you want to enable both SyncE and PTP policies for a given node (or set of nodes), the corresponding node settings policy with both functionalities concurrently enabled must be created (as part of a Fabric Policies template) and referenced in the node profile.

- f) Repeat this step to create any additional Node Profile policies.

Only a single Node Profile policy can be associated to a given node (or set of nodes).

Step 8 Create a Pod Profiles policy.

Before you can create a Pod profile policy, you must have a Pod settings policy that is already created for it as described in [Creating Fabric Policies, on page 90](#). In this release, a Pod settings policy can be used to enable the NTP functionality.

- From the **+Create Object** drop-down, select **Pod Profile**.
- In the right properties sidebar, provide the **Name** for the policy.
- (Optional) Click **Add Description** and provide a description for the policy.
- From the **Type** drop-down, select whether you want the policy to apply to **All** pods or a **Range** of pods.
- If you chose **Range** for the **Type**, provide the range of Pods to which to apply this policy.
- Click **No selected Pod Policy** and choose the Pod settings policy that you created for this.

The Pod settings that are defined in the Pod settings policy will be applied to all nodes you provided in the previous substep.

- Repeat this step to create any additional Pod Profile policies.

Only a single Pod Profile policy can be associated to a given Pod (or set of Pods).

Step 9 Create a FEX Device policy.

- From the **+Create Object** drop-down, select **FEX Device**.
- In the right properties sidebar, provide the **Name** for the policy.
- (Optional) Click **Add Description** and provide a description for the policy.
- Provide one or more **Nodes** (switches) that connect to the FEX device.

Currently, only straight-through connections between a FEX and a parent leaf switch are supported, so each FEX should be associated only to a single parent switch.

However, the FEX device policy allows you to specify multiple nodes, for example:

FEX Devices ×

UntitledFexDevice1

Common Properties ^

Name *

UntitledFexDevice1

[Add Description](#)

Nodes

101,102

Interfaces *

1/34

FEX Device ID *

101

The above configuration means that there are two FEX devices, one connected to leaf switch 101 and another connected to leaf switch 102 with both devices having FEX ID 101. Since FEX ID is limited to the leaf switch scope, FEX devices that are connected to different leaf switches can have the same IDs.

- e) Provide one or more **Interfaces** that connect to the FEX device.
- f) Provide the **FEX Device ID**.
- g) Repeat this step to create any additional FEX Device policies.

Step 10 Click **Save** to save the changes you've made to the template.

Note

When you save (or deploy) the template to one or more sites, the Orchestrator will verify that the specified nodes or interfaces are valid for the sites and will return an error if that is not the case.

Step 11 Click **Deploy** to deploy the template to the associated sites.

The process for deploying tenant policy templates is identical to how you deploy application templates.

If you have previously deployed this template but made no changes to it since, the **Deploy** summary indicates that there are no changes, and you can choose to redeploy the entire template. In this case, you can skip this step.

Otherwise, the **Deploy to sites** window shows you a summary of the configuration differences that will be deployed to sites. In this case only the difference in configuration is deployed to the sites. If you want to redeploy the entire template, you must deploy when to sync the differences, and then redeploy again to push the entire configuration as described in the previous paragraph.

Creating Monitoring Policies

This section describes how to create one or more SPAN session policies using the Monitoring Policy templates.

Procedure

Step 1 Log in to your Cisco Nexus Dashboard and open the Cisco Nexus Dashboard Orchestrator service.

Step 2 Create a new Tenant Policy.

- a) From the left navigation pane, choose **Configure > Policy Templates**.
- b) On the **Monitoring Policy Templates** tab, click **Create Monitoring Policy Template**.
- c) Select the SPAN session type for this template.

You can choose one of the following:

- **Tenant** – this type of SPAN sessions are referred to as ERSPAN sessions and allows you to configure an EPG belonging to the specified Tenant anywhere in the fabric as the SPAN session `source` and another EPG belonging to the same or to a different tenant as the `destination`.
- **Access** – allows you to configure one of the following two scenarios:
 - Access ports, port-channels, and vPC as source and destination as physical/port-channel interface, in which case the `source` and `destination` interfaces must be on the same switch.

- Access ports, port-channels, and vPC as source and destination as EPG, in which case it is an ERSPAN session allowing the SPAN destination to be connected anywhere in the fabric.

- If you chose `Tenant` as the session type, choose a **Tenant** with which to associate the monitoring policy.
- Choose the **Site** with which to associate the monitoring policy.
- In the **Monitoring Policy** page's right properties sidebar, provide the **Name** for the template.

By default, the new template is empty, so you must add one or more fabric policies as described in the following steps.

Step 3 Assign the template to one or more sites.

The process for assigning Tenant Policy templates to sites is identical to how you assign application templates to sites.

- In the **Template Properties** view, click **Actions** and choose **Add/Remove Sites**.

The **Associate Sites to <template-name>** window opens.

- In the **Associate Sites** window, check the check box next to the sites where you want to deploy the template.

Only the on-premises ACI sites support tenant policy templates and will be available for assignment.

- Click **OK** to save.

Step 4 Create a SPAN Session policy for a `Tenant` type template.

If you picked `Access` for the template type, use the next step instead.

- From the **+Create Object** drop-down, select **SPAN Session**.
- In the right properties sidebar, provide the **Name** for the policy.
- (Optional) Click **Add Description** and provide a description for the policy.
- Enable the **Admin State** check box.

If admin state is set to `disabled`, no data is sent to the configured monitor.

- Click **+Add Source** and provide the SPAN source information.

For the source information, provide the following:

- **Name**

- **Direction** – the SPAN source packet direction, which can be one of the following:

- **Both** – Replicate and forward packets that are incoming to the source and outgoing from the source.
- **Incoming** – Replicate and forward packets that are incoming to the source.
- **Outgoing** – Replicate and forward packets that are outgoing from the source.

- **Source EPG** – The source of the SPAN traffic.

For `Tenant` type templates, the source is always an EPG.

Click **OK** to save the source. You can then click **+Add Source** to provide extra sources if necessary.

- From the **Destination Group** section, provide the **Tenant**, **Destination EPG**, and **Destination IP Address** to which the replicated packets will be forwarded.

IPv4 and IPv6 IP addresses are supported in this field. However, you must not mix IPv4 for the **Destination IP** and IPv6 for the **Source IP Prefix** or conversely.

- g) Provide the **Source IP Prefix**.

If a specific IP address is configured, all ERSPAN traffic will be sourced from that IP (for example, for all the ACI leaf switches sourcing the ERSPAN traffic). If instead a prefix is configured, each ACI leaf switch will be assigned a unique IP that is part of that prefix to source ERSPAN traffic. This could be useful on the destination switch to distinguish the origin of the ERSPAN traffic.

- h) Choose the **SPAN Version**.

- i) (Optional) If necessary, configure the **Advanced Settings**.

- **Enforce SPAN Version** – When enabled, enforces the chosen SPAN Version.

If `Enabled`, the SPAN session uses the specified SPAN version if the hardware supports it. Otherwise, the session will fail.

If `Disabled` and Version 2 is specified but is not supported by the hardware, then Version 1 is used.

- **Flow ID** – The identifier of the ERSPAN packet.

When packets are copied and sent through ERSPAN, the packets are encapsulated with ERSPAN header. The flow ID is the number in the ERSPAN header to identify by which ERSPAN session these packets were copied.

The range is from 1 to 1023. The default is 1.

- **TTL** – The Time to Live (TTL) or hop limit in the 1-255 hops range; if set to zero, then no TTL is specified. The default is 64 hops.

- **DSCP** – DSCP value set in the IP header of the ERSPAN packet.

- **MTU** – The maximum transmission unit of the ERSPAN-generated packets.

The range is from 64 to 9216. The default is 1518.

For ERSPAN, the real MTU received by the destination device will be larger than the configured MTU because the ERSPAN encapsulation is added. For ERSPAN version 2, an extra 46 bytes are added. For ERSPAN version 1, an extra 34 bytes are added. As a result, with the default MTU of 1518, the end device would actually need to support 1564 (1518 + 36) for version 2 or 1552 (1518 + 34) for version 1.

If the captured frame is larger than the configured MTU, then the frame is truncated to the MTU length when the frame is replicated. The packet/frame payload would be incomplete, but the headers should still be intact for analysis.

- j) Repeat this step to create any additional Tenant SPAN Session policies.

Step 5

Create a SPAN Session policy for an `Access` type template.

If you picked `Tenant` for the template type, use the previous step instead.

- From the **+Create Object** drop-down, select **SPAN Session**.
- In the right properties sidebar, provide the **Name** for the policy.
- (Optional) Click **Add Description** and provide a description for the policy.
- Enable the **Admin State** checkbox.

If admin state is set to `disabled`, no data is sent to the configured monitor.

- Click **+Add Source** and provide the SPAN source information.

For the source information, provide the following:

- **Name**

- Click **+Add Access Path** to add one or more paths on the leaf switch. The following paths are supported:

- Port
- Port Channel
- Virtual Port Channel
- VPC Component PC

You can use the `VPC Component PC` option if you want to configure a vPC as a source and a physical/port-channel interface as destination. Since for this use case all the interfaces must be on the same switch, you must not choose the vPC as a source and must select the `VPC Component PC` option representing the interfaces of that vPC on the same switch where the destination is connected. In other words, you need to create a second SPAN session for the second switch that is part of the vPC domain so that the traffic can be spanned for the source interfaces which are part of the vPC on that switch towards a local destination.

- **Direction** – the SPAN source packet direction, which can be one of the following:
 - **Both** – Replicate and forward packets that are incoming to the source and outgoing from the source.
 - **Incoming** – Replicate and forward packets that are incoming to the source.
 - **Outgoing** – Replicate and forward packets that are outgoing from the source.

- Click **+Add Filter** to provide SPAN traffic filtering information.

Traffic filtering is optional and if no filters are specified then all traffic will be spanned.

You can enable filtering based on the following attributes:

- **Src IP Prefix**
- **Src Port From**
- **Src Port To**
- **Dst IP Prefix**
- **Dst Port From**
- **Dst Port To**
- **IP Protocol**

- **SPAN Drop Packets** – Allows SPAN to capture some of dropped packets that are not captured by the regular SPAN, but its limited to packets dropped as "forward drops".

If `Enabled`, allows for spanning of **only** dropped packets and not any traffic that was not dropped.

If `Disabled`, SPAN captures **only** the traffic that was not dropped.

The default value is `Disabled`.

- **EPG Filter** – If **SPAN Drop Packets** is disabled, you can filter the source packets based on the EPG they are coming from. To enable the filter, choose **EPG** and then select the specific EPG from the **Source EPG** drop-down.

The traffic sent and received on the previously configured source interfaces will be spanned only if it belongs to the specified EPG.

Click **OK** to save the source. You can then click **+Add Source** to provide extra sources if necessary.

f) Choose the **Destination Type**.

The replicated packets can be forwarded to either an EPG or a specific access interface. In the first case, an ERSPAN session is created to send the spanned traffic to a destination connected anywhere in the fabric; in the second case, the destination must be connected to physical/port-channel interfaces on the same switch as the source interfaces.

g) If you chose **EPG** for the **Destination Type**, provide the following information:

- **Tenant, Destination EPG**, and **Destination IP Address** to which the replicated packets will be forwarded.

IPv4 or IPv6 IP addresses are supported in this field. However, you must not mix IPv4 for the **Destination IP** and IPv6 for the **Source IP Prefix** or conversely.

- **Source IP Prefix** – the base IP address of the IP subnet of the source packets.

- **SPAN Version**

- (Optional) **Advanced Settings**

- **Enforce SPAN Version** – When enabled, enforces the chosen SPAN Version.

If **Enabled**, the SPAN session will use the specified SPAN version if the hardware supports it. Otherwise, the session fails.

If **Disabled** and Version 2 is specified but is not supported by the hardware, then Version 1 is used.

- **Flow ID** – The identifier of the ERSPAN packet.

When packets are copied and sent through ERSPAN, the packets are encapsulated with ERSPAN header. The flow ID is the number in the ERSPAN header to identify by which ERSPAN session these packets were copied.

The range is 1–1023. The default is 1.

- **TTL** – The Time to Live (TTL) or hop limit in the 1–255 hops range; if set to zero, then no TTL is specified. The default is 64 hops.

- **DSCP** – DSCP value set in the IP header of the ERSPAN packet.

- **MTU** – The MTU of the ERSPAN-generated packets.

The range is 64–9216. The default is 1518.

For ERSPAN, the real MTU received by the destination device will be larger than the configured MTU because the ERSPAN encapsulation is added. For ERSPAN version 2, an extra 46 bytes are added. For ERSPAN version 1, an extra 34 bytes are added. As a result, with the default MTU of 1518, the end device would actually need to support 1564 (1518 + 36) for version 2 or 1552 (1518 + 34) for version 1.

If the captured frame is larger than the configured MTU, then the frame is truncated to the MTU length when the frame is replicated. The packet/frame payload would be incomplete, but the headers should still be intact for analysis.

h) Otherwise, if you chose **Access Interface** for the **Destination Type**, provide the following information instead:

- **Path Type** – type of the interface, can be **Port** or **Port Channel**.
- For **Port** interfaces, select the **Node** and **Path**.
- For **Port Channel** interfaces, select the name of the port channel.

- **MTU** – The MTU of the ERSPAN-generated packets.

The range is 64–9216. The default is 1518.

For ERSPAN, the real MTU received by the destination device will be larger than the configured MTU because the ERSPAN encapsulation is added. For ERSPAN version 2, an extra 46 bytes are added. For ERSPAN version 1, an extra 34 bytes are added. As a result, with the default MTU of 1518, the end device would actually need to support 1564 (1518 + 36) for version 2 or 1552 (1518 + 34) for version 1.

If the captured frame is larger than the configured MTU, then the frame is truncated to the MTU length when the frame is replicated. The packet/frame payload would be incomplete, but the headers should still be intact for analysis.

- Repeat this step to create any additional Access SPAN Session policies.

Step 6 Click **Save** to save the changes you've made to the template.

Step 7 Click **Deploy** to deploy the template to the associated sites.

The process for deploying tenant policy templates is identical to how you deploy application templates.

If you have previously deployed this template but made no changes to it since, the **Deploy** summary indicates that there are no changes, and you can choose to redeploy the entire template. In this case, you can skip this step.

Otherwise, the **Deploy to sites** window shows you a summary of the configuration differences that will be deployed to sites. Note that in this case only the difference in configuration is deployed to the sites. If you want to redeploy the entire template, you must deploy when to sync the differences, and then redeploy again to push the entire configuration as described in the previous paragraph.



PART II

Operations

- [Audit Logs, on page 115](#)
- [Backup and Restore, on page 117](#)
- [Upgrading Sites, on page 131](#)
- [Tech Support, on page 143](#)



Audit Logs

- [Audit Logs](#), on page 115

Audit Logs

Cisco Nexus Dashboard Orchestrator system logging is automatically enabled when you first deploy the Orchestrator cluster and captures the events and faults that occur in the environment.

You can view the Cisco Nexus Dashboard Orchestrator logs directly in the GUI by selecting **Admin > System Configuration > > Audit Logs** from the main navigation menu.

From the **Audit Logs** page, you can click the **Time Frame** (shown as a range of dates) field to select a specific time period for which you want to see the logs. For example, when you select the range from November 14, 2019 to November 17, 2019 and click **Apply**, the audit log details for this time period are displayed on the **Audit Logs** page.

You can also click the **Filter** icon to filter the log details using the following criteria:

- **User:** Select this option to filter the audit logs by the user type, then click **Apply** to apply the filter.
- **Type:** Select this option to filter the audit logs by the policy types (for example, site, user, template) and click **Apply**.
- **Action:** Select this option to filter the audit logs by an action. The available actions are Created, Updated, Deleted, Added, Removed, Associated, Disassociated, Deployed, Undeployed, Downloaded, Uploaded, Restored, signed in, Logged Out, sign-in Failed. Select an action and click **Apply** to filter the log details according to the action.



CHAPTER 8

Backup and Restore

- [Configuration Backup and Restore Guidelines, on page 117](#)
- [Configuring Remote Locations for Backups, on page 119](#)
- [Creating Backups, on page 120](#)
- [Restoring Backups, on page 120](#)
- [Exporting \(Downloading\) Backups, on page 128](#)
- [Importing Backups to Remote Location, on page 129](#)
- [Backup Scheduler, on page 130](#)

Configuration Backup and Restore Guidelines

You can create backups of your Cisco Nexus Dashboard Orchestrator configuration that can facilitate in recovering from Orchestrator failures or cluster restarts. We recommend creating a backup of the configuration before every upgrade or downgrade of your Orchestrator and after every configuration change or deployment. The backups are always created on a remote server (not Cisco Nexus Dashboard cluster), which is defined in the Cisco Nexus Dashboard Orchestrator as described in the following sections.

When creating configuration backups, the following guidelines apply:

- Importing and restoring backups that are created from later releases is not supported.
For example, if you downgrade your Cisco Nexus Dashboard Orchestrator to an earlier release, you cannot restore a backup of the configuration that is created on a later release.
- Restoring configuration backups created on releases before Release 4.0(1) is supported only during the initial upgrade to this release.
If you want to upgrade from a release before release 4.0(1) to this release, see the "Upgrading NDO Service in Cisco Nexus Dashboard" chapter in the [Cisco Nexus Dashboard Orchestrator Deployment Guide](#).
- When saving a backup, the configuration is saved in the same state in which it was deployed. When restoring a backup, any policies that were deployed will show as `deployed`, while any policies that were not deployed will remain in the `undeployed` state.
- Restoring a backup action restores the database on the Cisco Nexus Dashboard Orchestrator, but it does not make any changes to the controller (such as APIC, Cloud Network Controller, or NDFC) databases on each site.

We recommend that after you restore the Orchestrator database you resolve any configuration drifts that may appear in the templates, as described in the "Configuration Drifts" section of this guide, and then redeploy the existing templates to avoid potentially mismatching policies between the Cisco Nexus Dashboard Orchestrator and each site's controller.

- When you create a configuration backup, the files are first created on the Orchestrator's local drives, then uploaded to the remote location, and finally deleted from the local storage. If there is not enough local disk space, the backup fails.
- If you have a backup scheduler that is enabled to take local backups before upgrading to Release 4.0(1) or later, it will be disabled after the upgrade.

After the upgrade, you must readd any remote locations you had set up and then re-enable backup scheduler.

- Deleting a backup using the UI also deletes the backup files from the remote location.

When restoring configuration backups, the following guidelines apply:

- If there have been no policy changes between when the backup was created and when it is being restored, no additional considerations are required and you can simply restore the configuration as described in [Restoring Backups, on page 120](#).
- If any configuration changes took place between the time when the configuration backup was created and the time it is being restored, consider the following:
 - Restoring a backup will not modify any objects, policies, or configurations on the sites. Any new objects or policies that are created and deployed since the backup will remain deployed.

We recommend that after you restore the Orchestrator database you resolve any configuration drifts that may appear in the templates, as described in "Configuration Drifts" section of this guide, and then redeploy the existing templates to avoid potentially mismatching policies between the Cisco Nexus Dashboard Orchestrator and each site's controller.

Alternatively, you can choose to undeploy all policies first, which will avoid any potential stale objects after the configuration is restored from backup. However, this would cause a disruption in traffic or services that are defined by those policies.

- The steps required to restore a configuration backup are described in [Restoring Backups, on page 120](#).
- If the configuration backup you restored was saved before it was deployed to the sites, it will be restored in the `undeployed` state and you can simply deploy it to the sites as necessary.
- If the configuration backup you restored was saved when the configuration was already deployed, it will be restored in the `deployed` state, although none of the configurations exist in the sites yet.

In this case, resolve any configuration drifts that may appear in the templates, as described in the "Configuration Drifts" section of this guide and redeploy the templates to sync the Cisco Nexus Dashboard Orchestrator's configuration with the sites.

- If sites that were managed when the backup was created are no longer present in the Cisco Nexus Dashboard, the restore fails.
- If sites' status since the backup has changed (`managed` vs `unmanaged`) but the sites are still present in the Cisco Nexus Dashboard, the status will be restored to what it was at the time of backup.

Configuring Remote Locations for Backups

This section describes how to configure a remote location in Cisco Nexus Dashboard Orchestrator to which you can then export your configuration backups.

Procedure

Step 1 Log in to your Cisco Nexus Dashboard and open the Cisco Nexus Dashboard Orchestrator service.

Step 2 From the left navigation pane, select **Admin > Backup and Restore > Remote Locations** tab.

Step 3 In the top right of the main window, click **Create Remote Location**.

An **Create New Remote Location** screen appears.

Step 4 Provide the name for the remote location and an optional description.

Two protocols are currently supported for remote export of configuration backups:

- SCP
- SFTP

Note

SCP is supported for non-Windows servers only. If your remote location is a Windows server, you must use the SFTP protocol.

Step 5 Specify the host name or IP address of the remote server.

Based on your **Protocol** selection, the server you specify must allow SCP or SFTP connections.

Step 6 Provide the full path to a directory on the remote server where you save the backups.

The path must start with a slash (/) characters and must not contain periods (.) or backslashes (\). For example, */backups/multisite*.

Note

The directory must exist on the remote server.

Step 7 Specify the port used to connect to the remote server.

By default, port is set to 22.

Step 8 Specify the authentication type used when connecting to the remote server.

You can configure one of the following two authentication methods:

- **Password**—Provide the username and password that is used to sign in to the remote server.
- **SSH Private Files**—provide the username and the SSH Key/Passphrase pair that is used to sign in to the remote server.

Step 9 Click **Save** to add the remote server.

Creating Backups

This section describes how to create a new backup of your Cisco Nexus Dashboard Orchestrator configuration.

Before you begin

You must first add the remote location as described in [Configuring Remote Locations for Backups, on page 119](#).

Procedure

Step 1 Log in to your Cisco Nexus Dashboard Orchestrator.

Step 2 Backup existing deployment configuration.

- a) From the left navigation pane, select **Admin > Backups & Restore**.
- b) In the main window, click **Create New Backup**.

A **New Backup** window opens.

- c) Provide the backup information.

- In the **Name** field, provide the name for the backup file.

The name can contain up to 10 alphanumeric characters, but no spaces or underscores (_).

- From the **Remote Location** drop-down, select a remote location that you have configured for storing backups.
- (Optional) In the **Remote Path**, provide the specific directory on the remote server where to save the backup.
The directory that you specify must exist.

- d) Click **Save** to create the backup.
-

Restoring Backups

This section describes how to restore a Cisco Nexus Dashboard Orchestrator configuration to a previous state.

Before you begin

- You must have configured a remote location for storing your NDO backups, as described in [Configuring Remote Locations for Backups, on page 119](#).
- Ensure that the backup you want to restore is on the remote location server or import the backup into the remote location, as described in [Importing Backups to Remote Location, on page 129](#).



Note Restoring a backup action restores the database on the Cisco Nexus Dashboard Orchestrator, but it does not make any changes to the controller (such as APIC, Cloud Network Controller, or NDFC) databases on each site.

We recommend that after you restore the Orchestrator database you resolve any configuration drifts that may appear in the templates, as described in the "Configuration Drifts" section of this guide, and then redeploy the existing templates to avoid potentially mismatching policies between the Cisco Nexus Dashboard Orchestrator and each site's controller.

For information on specific configuration mismatch scenarios and recommended restore procedures that are related to each one, see [Configuration Backup and Restore Guidelines, on page 117](#).

Procedure

Step 1 Log in to your Cisco Nexus Dashboard Orchestrator GUI.

Step 2 If necessary, undeploy existing policies.

We recommend you perform this step if new objects or policies were added to the configuration between when the backup was created and current configuration. Extra context is available in [Configuration Backup and Restore Guidelines, on page 117](#).

Step 3 From the left navigation menu, select **Admin > Backups & Restore**.

Step 4 In the main window, click the actions (...) icon next to the backup you want to restore and select **Rollback to this backup**.

If the version of the selected backup is different from the running Cisco Nexus Dashboard Orchestrator version, the rollback could cause a removal of the features that are not present in the backup version.

Step 5 Click **Yes** to confirm that you want to restore the backup you selected.

If you click **Yes**, the system stops the current session and the user is logged out.

Note

Multiple services are restarted during the configuration restore process. As a result, you may notice an up to 10-minute delay before the restored configuration is properly reflected in the NDO GUI.

Step 6 Using REST API calls, validate if NDO-managed policy IDs match across fabrics.

a) Log into your NDO through the REST API:

```
POST https://{ndo-ip}/login
```

Request Body:

```
{"userName": "<username>", "userPasswd": "<password>", "domain": "<domain>"}
```

For example:

```
POST https://192.168.0.1/login
```

Request Body:

```
{"userName": "admin", "userPasswd": "MyNDOPassword", "domain": "DefaultAuth"}
```

b) Post a **GET** policy-report API with `validate=true` for all tenants managed by this NDO.

Request:

```
GET https://{ndo-ip}/mso/api/v1/policy-report?tenants=<comma separated list of
tenants...>&validate=true
```

For example:

```
GET https://192.168.0.1/api/v1/policy-report?tenants=T-policyreport,T-migration&validate=true
```

The following is an example response:

```
{
  "sites": [
    {
      "id": "66903a7f5c1ced5eebed97ec",
      "apicId": "1",
      "name": "fabric-1"
    },
    {
      "id": "668eeb235c1ced5eebed97eb",
      "apicId": "2",
      "name": "fabric-2"
    },
    {
      "id": "60309f0b11000059b8d16c9d",
      "apicId": "3",
      "name": "fabric-3"
    }
  ],
  "policies": {
    "uni/tn-T-migration/BD-BD-singlesite": [
      {
        "apicId": "1",
        "vnid": "16056263",
        "peerContexts": [],
        "remoteMappings": {},
        "importRemoteIds": [],
        "exportFlag": false
      }
    ],
    "uni/tn-T-migration/ctx-VRF-singlesite": [
      {
        "apicId": "1",
        "vnid": "2490368",
        "ctxPcTag": "16386",
        "peerContexts": [],
        "remoteMappings": {},
        "importRemoteIds": [],
        "exportFlag": false
      }
    ],
    "uni/tn-T-policyreport/ctx-VRF-validation": [
      {
        "apicId": "1",
        "vnid": "2719745",
        "ctxPcTag": "16386",
        "peerContexts": [],
        "remoteMappings": {
          "11": "3112961, 49153"
        },
        "importRemoteIds": [],
        "exportFlag": false
      }
    ],
    "uni/tn-tenant1/BD-BD1": [
```

```

    {
      "apicId": "2",
      "vnid": "15925212",
      "peerContexts": [],
      "remoteMappings": {
        "1": "15728629",
        "3": "15761391"
      },
      "importRemoteIds": [],
      "exportFlag": false
    },
    {
      "apicId": "1",
      "vnid": "15859694",
      "peerContexts": [],
      "remoteMappings": {
        "2": "15925212",
        "3": "15761391"
      },
      "importRemoteIds": [],
      "exportFlag": false
    },
    {
      "apicId": "3",
      "vnid": "16711549",
      "peerContexts": [],
      "remoteMappings": {
        "1": "15728629",
        "2": "15925212"
      },
      "importRemoteIds": [],
      "exportFlag": false
    }
  ],
  "validation": {
    "uni/tn-T-policyreport/ctx-VRF-validation": [
      "Fabric id 1 and 2 do not have the same site list for this DN. (1: [1 11], 11: [])",

      "Fabric id 1 does not have the correct number of remote mappings for this DN."
    ],
    "uni/tn-tenant1/BD-BD1": [
      "Site id 2 and 1 have pcTag mismatch for this DN. (2: {2 15925212 [] map[1:15728629 3:15761391 4:15794150] [] false}, 1: 15728629)",
      "Site id 2 and 3 have pcTag mismatch for this DN. (2: {2 15925212 [] map[1:15728629 3:15761391 4:15794150] [] false}, 3: 15761391)",
      "Site id 1 and 3 have pcTag mismatch for this DN. (1: {1 15859694 [] map[2:15925212 3:15761391 4:15794150] [] false}, 3: 15761391)",
      "Site id 3 and 1 have pcTag mismatch for this DN. (3: {3 16711549 [] map[1:15728629 2:15925212 4:15794150] [] false}, 1: 15728629)"
    ]
  }
}

```

c) Determine if any NDO-managed policy IDs have a mismatch across fabrics.

- If you do *not* see validation responses similar to the ones highlighted in the example responses above, then all NDO-managed policies exist across the fabrics as expected and their IDs match (the "validation" json holds information about policies that have mismatches).

You do not have to take any further action in this case; the recovery process is complete.

- If you do see a validation response similar to either of the ones highlighted in the example responses above, then there are NDO-managed policy IDs that have a mismatch across the fabrics.

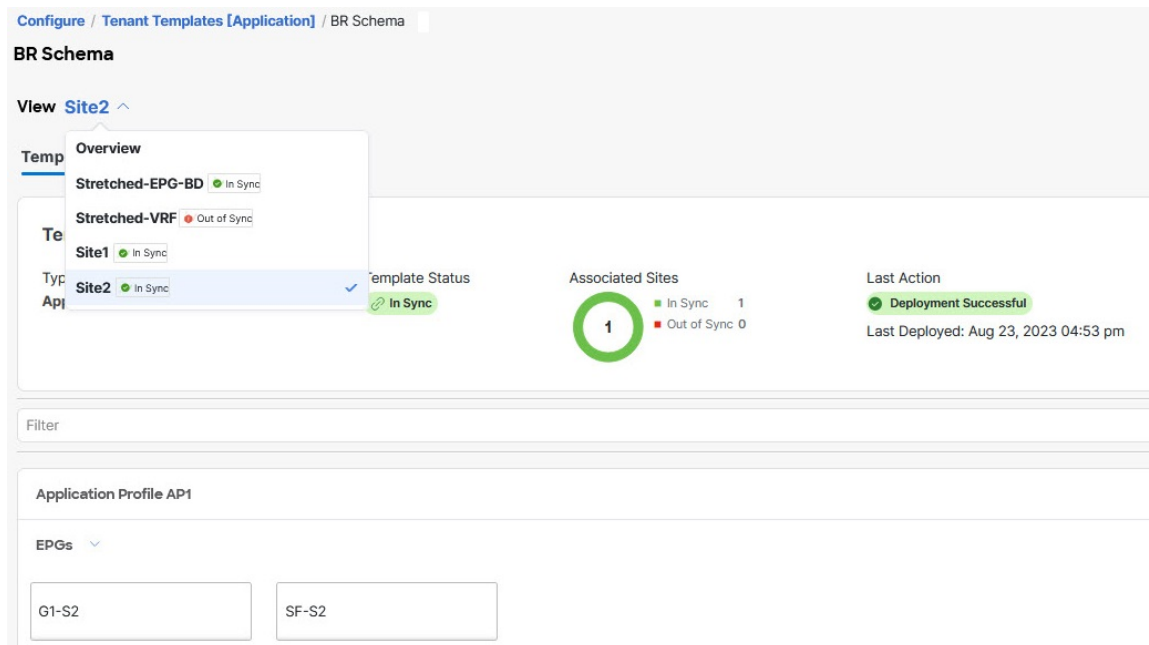
Step 7

If there are NDO-managed policy IDs that have a mismatch across the fabrics, check if any templates contain configuration drifts.

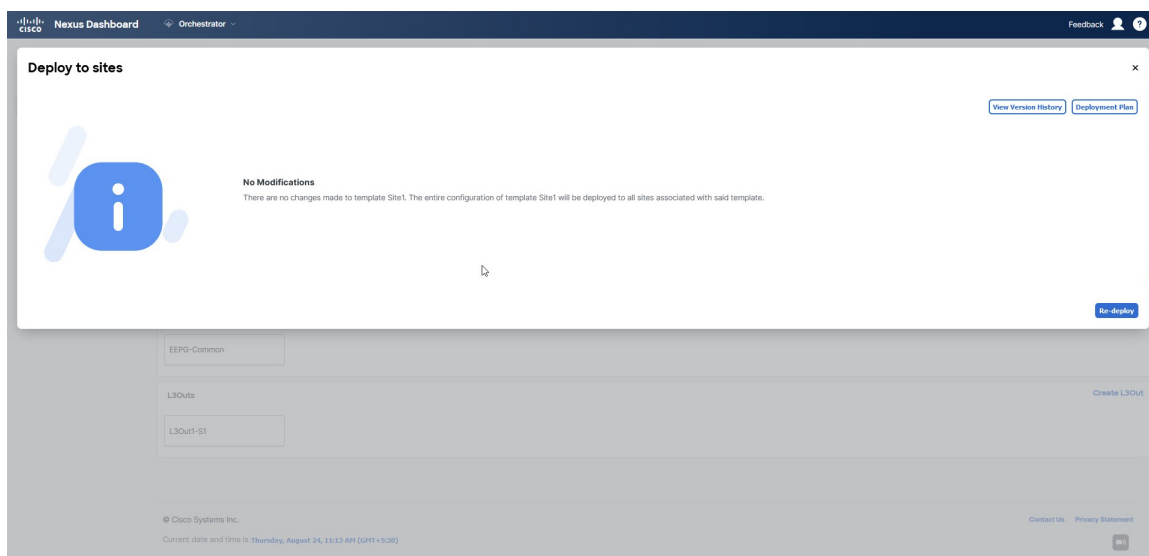
You repeat the following steps for every schema and template in your deployment.

You can check for configuration drifts in one of the following two ways:

- Check the template deployment status icon for each site to which the template is assigned:



- Select the template and click **Deploy template** to bring up the configuration comparison screen to check which objects contain configuration drifts:



Step 8

Determine your next course of action based on the result of the check on the configuration drifts.

- If any template contains a configuration drift, continue to [Step 9, on page 125](#) to resolve the conflicts.
- If no template contains a configuration drift, skip to [Step 10, on page 127](#) to perform a full redeployment of the templates that hold the appropriate policies, where the NDO-managed policy IDs have a mismatch across the fabrics.

Step 9

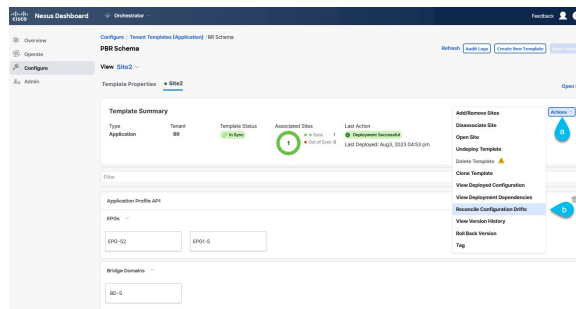
If any template contains a configuration drift, resolve the conflicts.

For more information about configuration drifts, check the "Configuration Drifts" chapter in the [Cisco Nexus Dashboard Orchestrator Configuration Guide for ACI Fabrics](#).

- a) Close the template deployment dialog to return to the Schema view.

Deploying any templates at this point would push the values in the Orchestrator database and overwrite any existing settings in the fabrics.

- b) From the template's **Actions** menu, select **Reconcile Configuration Drifts**.



The **Drift Reconciliation** wizard opens.

- c) In the **Drift Reconciliation** screen, compare the template-level configurations for each site and choose the one you want.

Version History

General Information

Schema	Template	Tenant
PBR Schema	Site2	PBR

Versions

☐ Golden Versions ☒ Deployed Versions ☐ Pre Reconciled Versions ☐ Post Reconciled Versions

3 4 5 6 7

Tag As Golden Delete Versions

Version 6 (Selected) 3 policies | 1 sites

Version 7 (Current) 2 policies | 1 sites

```

"externalEpgs": [
  {
    "externalEpgRef": "/schemas/Site2/externalEpgs/ExtEPG-S2",
    "l3outDn": "",
    "l3outRef": "c98d787f-fa8a-439"
  },
],
"externalEpgs": [
  {
    "contractRelationships": [],
    "description": "",
    "name": "ExtEPG-S2",
    "preferredGroup": false,
    "qosPriority": "unspecified",
    "selectors": [],
    "subnets": [
      {
        "scope": [
          "import-security"
        ]
      }
    ],
    "tagAnnotations": [],
    "vrfRef": "/schemas/64ddddd9btdddd-VRF-Contract/vrfs/VRF1"
  },
],

```

Click to expand

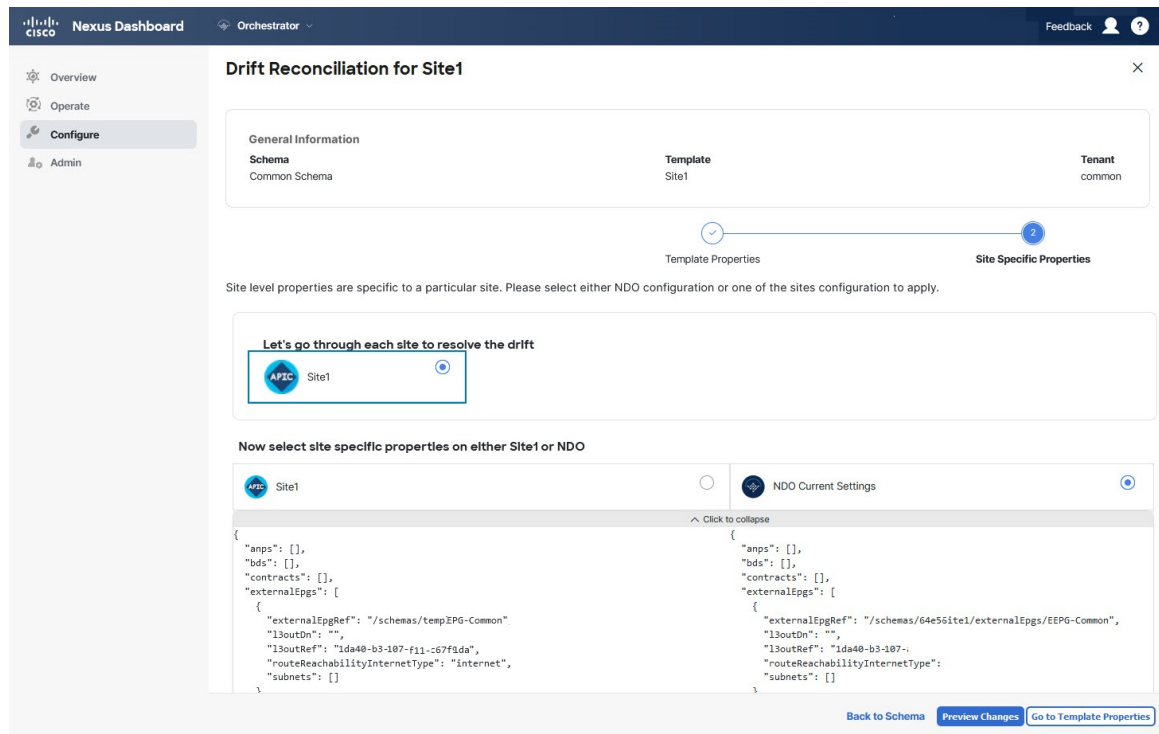
Click to expand

Click to expand

OK

Template-level properties are common across all sites that are associated to the template. You can compare the template level properties that are defined on Cisco Nexus Dashboard Orchestrator with the configuration that is rendered in each site and decide what should become the new configuration in the Cisco Nexus Dashboard Orchestrator template. Selecting the site configuration modifies those properties in the existing Cisco Nexus Dashboard Orchestrator template, whereas selecting the Cisco Nexus Dashboard Orchestrator configuration keeps the existing Cisco Nexus Dashboard Orchestrator template settings as is.

- d) Click **Go to Site Specific Properties** to switch to site-level configuration.



You can choose a site to compare that specific site's configuration. Unlike template-level configurations, you can choose either the Cisco Nexus Dashboard Orchestrator-defined or actual existing configurations for each site individually to be retained as the template's site-local properties for that site.

Although in most scenarios you make the same choice for both template-level and site-level configuration, the drift reconciliation wizard allows you to choose the configuration defined in the site's controller at the "Template Properties" level and the configuration that is defined in Cisco Nexus Dashboard Orchestrator at the "Site Local Properties" level or conversely.

- e) Click **Preview Changes** to verify your choices.

The preview displays full template configuration adjusted based on the choices that are picked in the **Drift Reconciliation** wizard. You can then click **Deploy to sites** to deploy the configuration and reconcile the drift for that template.

Step 10 Perform a full redeployment of the template.

You will perform a full deployment of a template in either of these situations:

- You found one or more templates that contain a configuration drift, and you resolved the conflicts using the instructions in [Step 9, on page 125](#). After all configuration drifts are resolved and there are no changes that are shown in the **Deploy to sites** dialog for the template, perform full redeployment of the template.

or

- No template contained a configuration drift but you found NDO-managed policy IDs that have a mismatch across the fabrics in [Step 6, on page 121](#). The policy names with their tenant names are listed in the validation report in that step. In this case, you must find the templates that hold these policies and redeploy those templates. Redeploying each of those templates pushes the IDs across the fabrics again and resolves the mismatches.

Once you have redeployed all of the necessary templates and the GET policy report API does not find any mismatches, then the recovery procedure is complete.

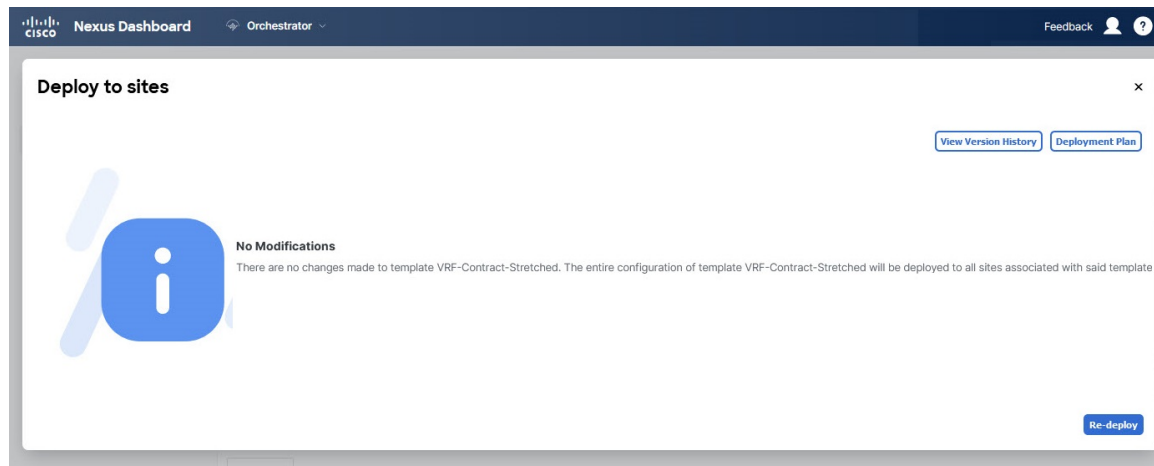
Note

We recommend that you do not redeploy templates if you are running on NDO release 4.2.3j or lower versions because of a known defect (CSCwj99109). This defect is resolved in NDO release 4.2.3k and later.

Note

Due to database transformations, you must perform a full redeployment of each template.

Ensure that the **Deploy to sites** dialog contains no changes as shown in the following figure, then click **Deploy** to redeploy complete configuration:



Step 11 Repeat the above steps for every schema and template in your Cisco Nexus Dashboard Orchestrator.

Step 12 Check audit logs to verify that all templates have been redeployed.

You can view the audit signs in the **Operations** tab.

Audit Logs page and confirm that all templates show as `Redeployed` to ensure that full redeployment successfully completed.

Exporting (Downloading) Backups

This section describes how to download the backup from the Cisco Nexus Dashboard Orchestrator.

Before you begin

Procedure

Step 1 Log in to your Cisco Nexus Dashboard Orchestrator GUI.

Step 2 From the left navigation menu, select **Admin > Backups & Restore**.

- Step 3** In the main window, click the actions (...) icon next to the backup you want to download and select **Download**. This downloads the backup file in `msc-backups-<timestamp>.tar.gz` format to your system. You can then extract the file to view its contents.
-

Importing Backups to Remote Location

This section describes how to upload an existing configuration backup you have previously downloaded and import it into one of the remote locations that are configured in your Cisco Nexus Dashboard Orchestrator.

Before you begin

You must have completed the following:

- Created and downloaded a configuration backup as described in [Creating Backups, on page 120](#) and [Exporting \(Downloading\) Backups, on page 128](#).
If your backup is already on a remote location, for example, if it was created on release 3.4(1) or later, you can download it to your local computer and upload it to a different remote location.
- Added a remote location for backups as described in [Configuring Remote Locations for Backups, on page 119](#).

Procedure

- Step 1** Log in to your Cisco Nexus Dashboard Orchestrator.
- Step 2** From the left navigation pane, select **Admin > Backups & Restore**.
- Step 3** In the main pane, click **Upload**.
- Step 4** In the **Upload from file** window that opens, click **Select File** and choose the backup file that you want to import. Uploading a backup adds it to the list of the backups displayed the **Backups** page.
- Step 5** From the **Remote Location** drop-down list, select the remote location.
- Step 6** (Optional) Update the remote location path.
The target directory on the remote server, which you configured when creating the remote backup location, will be displayed in the **Remote Path** field.
You can choose to append extra subdirectories to the path. However, the directories must be under the default-configured path and must have been already created on the remote server.
- Step 7** Click **Upload** to import the file.
Importing a backup adds it to the list of the backups displayed the **Backups** page.
Note that although the backups are shown on the NDO UI, they are located on the remote servers only.
-

Backup Scheduler

This section describes how to enable or disable the backup scheduler, which will perform complete configuration backup at regular intervals.

Before you begin

You must have already added a remote location for backups as described in [Configuring Remote Locations for Backups, on page 119](#).

Procedure

Step 1 Log in to your Cisco Nexus Dashboard Orchestrator GUI.

Step 2 From the left navigation menu, select **Admin > Backups & Restore**.

Step 3 In the top right of the main pane, click **No Schedule**.

The **Backup Scheduler Settings** window opens.

Step 4 Set up backup scheduler.

- a) Check the **Enable Scheduler** check box.
- b) In the **Select Starting Date** field, provide the day when you want the scheduler to start.
- c) In the **Select Time** fields, provide the time of day when you want the scheduler to start.
- d) From the **Select Frequency** drop-down, choose how often the backup should be performed.
- e) From the **Remote Location** drop-down, select the location where the backups will be saved.
- f) (Optional) In the **Remote Path** field, update the path on the remote location where the backups will be saved.

The target directory on the remote server, which you configured when creating the remote backup location, will be displayed in the **Remote Path** field.

You can choose to append extra subdirectories to the path. However, the directories must be under the default-configured path and must have been already created on the remote server.

- g) Click **Save** to finish.

Step 5 If you want to disable the backup scheduler, simply uncheck the **Enable Scheduler** check box in the preceding step.



CHAPTER 9

Upgrading Sites

- [Overview, on page 131](#)
- [Guidelines and Limitations, on page 133](#)
- [Downloading Controller and Switch Node Firmware to Sites, on page 133](#)
- [Upgrading Controllers, on page 136](#)
- [Upgrading Nodes, on page 138](#)

Overview



Note This feature is supported for Cisco APIC sites only. It is not supported for Cisco Cloud Network Controller or Cisco NDFC fabrics.

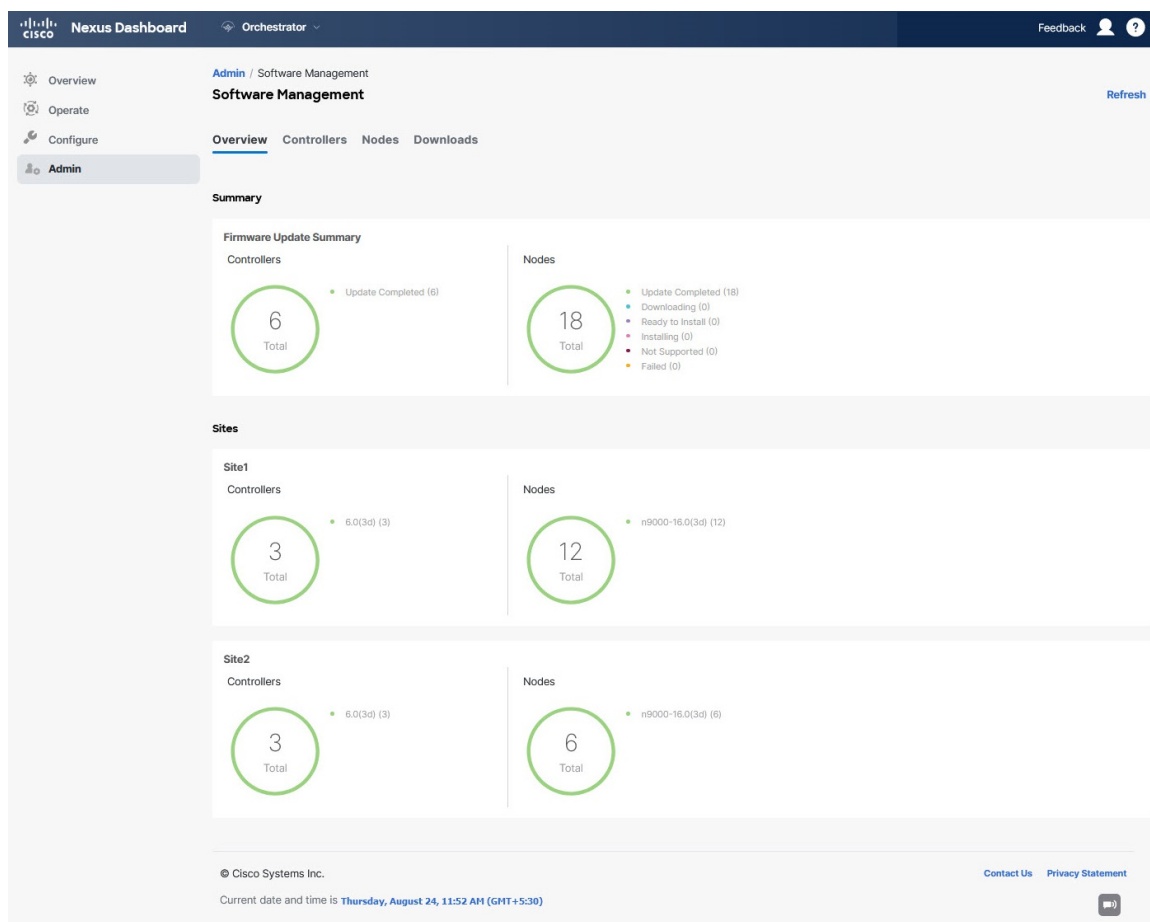
When you deployed Cisco Multi-Site, each site's APIC clusters and switch nodes software had to be managed individually at the site level. As the number of sites in your Multi-Site domain grew, the release life cycle and upgrades could become complicated as they had to be manually coordinated and managed for release and feature compatibility.

Cisco Nexus Dashboard Orchestrator provides a workflow that allows you to manage all sites' software upgrades from a single point eliminating the need for multiple site administrators to manually coordinate software upgrades and giving you insight into any potential issues that could affect the upgrades.

You can access the site upgrades screen by navigating to **Admin > Software Management**. The page contains four tabs, which are described in this and following sections.

The **Overview** tab displays information about the sites in your Multi-Site domain and the firmware versions that are deployed or ready to be deployed. The `Sites Firmware` service polls the sites every 5 minutes for new or changed data such as the latest status of any of the upgrade policies. You can manually trigger a Refresh by clicking the **Refresh** button in the upper right corner of the main pane.

Figure 12: Sites Firmware Overview



The page is divided into the following areas:

- **Firmware Update Summary**—Provides overall summary of the firmware images that are present across all sites in your Multi-Site domain, including the Cisco APIC and the switch firmware.

For each type of image, the specific information includes the number of images in each state:

- **Completed**—The image is currently deployed to the controllers or the switches.
- **Downloading** (for switch nodes only)—The image is being downloaded to the switch nodes.
- **Ready to Install** (for switch nodes only)—The image was successfully downloaded to the switch nodes and is ready to be installed.
- **Installing**—The images currently in the process of being deployed to the controllers or the switch nodes.
- **Not Supported**—The images that do not support remote firmware upgrades, such as releases before Release 4.2(5).
- **Site-specific information**—Extra sections of the page display information about individual sites, which includes the version of the currently deployed software and the number of controllers or nodes.

Guidelines and Limitations

When performing fabric upgrades from the Cisco Nexus Dashboard Orchestrator, the following restrictions apply:

- You must review and follow the guidelines, recommendations, and limitations specific to the Cisco APIC upgrade process described in the [Upgrading and Downgrading the Cisco APIC and Switch Software](#) of the *Cisco APIC Installation, Upgrade, and Downgrade Guide*.
- Your Cisco Nexus Dashboard Orchestrator must be deployed in Cisco Nexus Dashboard.

The site upgrade feature is not available for NDO deployments in VMware ESX and you must follow the standard upgrade procedures that are described in [Cisco APIC Installation, Upgrade, and Downgrade Guide](#).

- The fabrics must be running Cisco APIC, Release 4.2(5) or later.

Fabrics running earlier APIC releases will not be available for selection during the upgrade workflow. Follow the standard upgrade procedures described in [Cisco APIC Installation, Upgrade, and Downgrade Guide](#).

- We recommend coordinating the site upgrades with the site administrators managing those fabrics. You may need access to the controllers or switch nodes to troubleshoot any potential issues should they arise.
- If a fabric switch node goes into an `inactive` state in the middle of the upgrade process, for example due to hardware or power failure, the process is unable to complete. You will not be able to remove or modify the node upgrade policy from NDO during this time as NDO is unable to differentiate whether the node went down or is simply in the middle of a reboot for the upgrade.

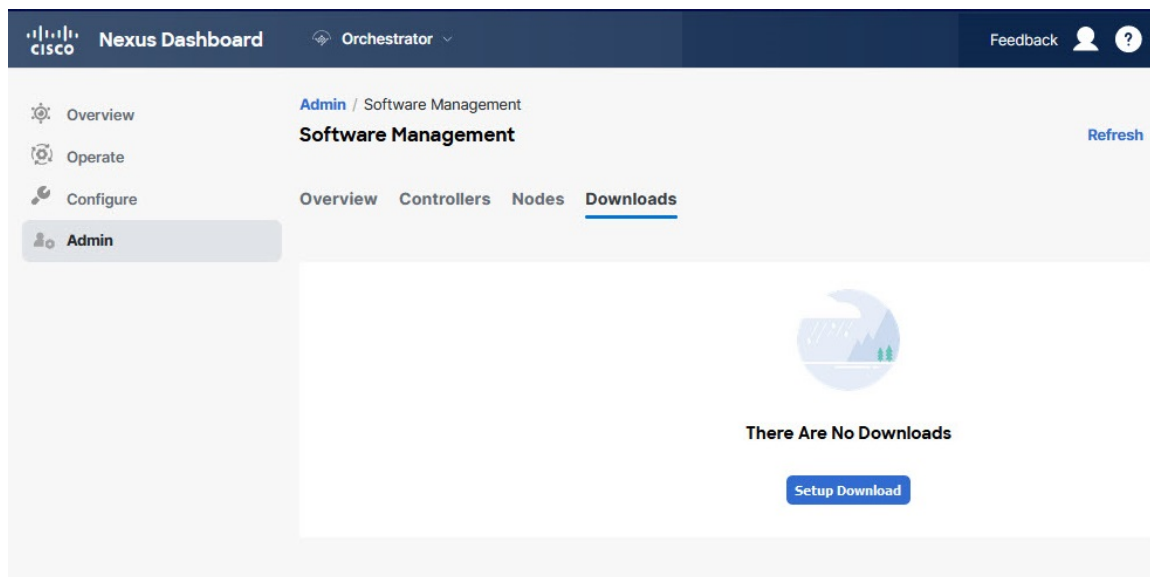
To resolve this issue, you must manually decommission the inactive node from the APIC, at which point the NDO upgrade policy recognizes the change and return a `failed` status. Then you can update the upgrade policy on the NDO to remove the switch and rerun the upgrade.

Downloading Controller and Switch Node Firmware to Sites

You must download the controller and switch software to all the site controllers in your fabrics before performing the upgrade. After you complete the following steps, you will be able to start the upgrade process later using the downloaded images.

Procedure

-
- Step 1** Log in to your Cisco Nexus Dashboard Orchestrator.
- Step 2** Set up firmware download.



- a) From the left navigation pane, select **Admin > Software Management**.
- b) In the main window, select the **Downloads** tab.
- c) Click **Setup Downloads** tab.

If you have previously set up one or more downloads, click the **Setup Downloads** button in the top right of the main pane instead.

The **Download Image to APIC** screen opens.

Step 3 Select the sites.

The image will be downloaded to the Cisco APICs of all the sites you select here.

- a) Click **Select Sites**.
- b) In the **Select Sites** window, check one or more sites and click **Add and Close**.
- c) Click **Next** to proceed.

Step 4 Provide the download details.

- a) Provide the **Name**.

You can provide a descriptive name for tracking the download.

- b) Choose the protocol.

You can choose to download the image through **HTTP** or **SCP**.

- c) Click + **Add URL** to provide location of one or more images.

You can provide both, the APIC and the switch firmware images.

- d) If you selected **SCP**, provide the authentication information.

You must provide the sign-in **Username**, for example `admin`.

Then choose the **Authentication Type**:

- For **Password** authentication, simply enter the password for the username you provided earlier.
- For **SSH Key** authentication, you must enter the **SSH Key** and the **SSH Key Passphrase**.

- e) Click **Next** to proceed.

Step 5

In the confirmation screen, review the information and click **Submit** to proceed.

In the **Downloading** screen that opens, you can view the status of the image download.

You can also click the status, to see extra details about the progress.

Image Download - MSO-d11

Setup | **Downloading** | Complete

Download Details

Name: MSO-d11

Overall Status: **Downloading**

Status Breakdown: 3

- Downloaded (0)
- Downloading (3)
- Download Failed (0)

Sites

Filter by attributes

Site	URLs	Status
ifav109-site1	1	Downloading (1)
ifav109-site2	1	Downloading (1)
ifav109-site3	1	Downloading (1)

ifav109-site3 Details:

URLs: 0.117a final aci-apic-dk9.5.1.0.117a.iso

Status: **Downloading (30%)**

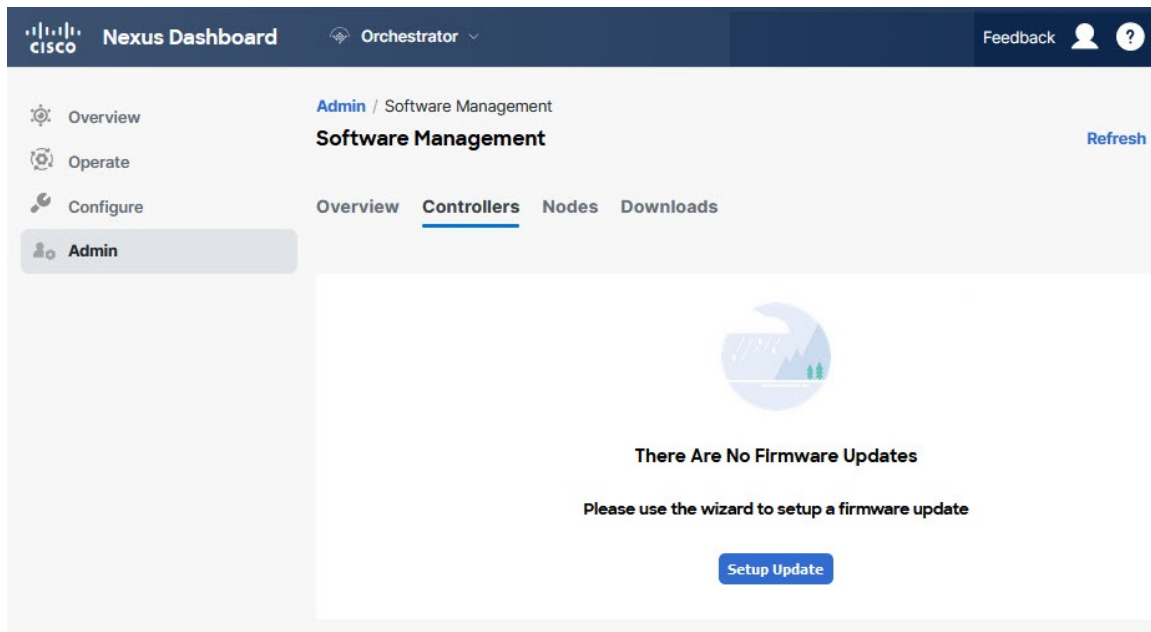
After all downloads complete, you will transition to the **Completed** screen. You do not have to wait at the **Downloading** screen, you can always navigate back to it from the **Downloads** tab by clicking the download name that you provided in a previous step.

Upgrading Controllers

This section describes how to set up a software upgrade for your sites' APIC clusters.

Procedure

- Step 1** Log in to your Cisco Nexus Dashboard Orchestrator.
- Step 2** Set up APIC cluster upgrade.



- a) From the left navigation pane, select **Admin > Software Management**.
- b) In the main window, select the **Controllers** tab.
- c) Click **Setup Update** tab.

If you have previously set up one or more updates, click the **Setup Update** button in the top right of the main pane instead.

The **Setup Site Firmware Update** screen opens.

Step 3 Provide the upgrade details.

- a) Provide the **Name**.

This is the controller upgrade policy name that you will be able to use to track the upgrade progress at any time.

- b) Click **Select Sites**.

The **Select Sites** window opens.

- c) In the **Select Sites** window, check one or more sites and click **Add and Close**.
- d) Click **Next** to proceed.

Step 4 In the **Version Selection** screen, select the firmware version and click **Next**.

The firmware must be downloaded to the sites before it becomes available here. If the download you set up in previous section has completed successfully but the image is still not available here, close the **Setup Site Firmware Update** screen, navigate back to **Admin > Software Management > Overview** tab, and click the **Refresh** button to reload the latest information available for the sites; then restart the upgrade steps.

Step 5 In the **Validation** screen, review the information, then click **Next**.

Ensure that there are no faults and review any additional information that may affect your upgrade:

Setup Site Firmware Update

Setup
Downloading
Ready to Install
Installing
Complete

1
2
3
4

Site Selection
Version Selection
Validation
Confirmation

ifav109-site1

Following nodes are not in vPC ['1111','102','101','104','103'].
Configure vPC for the listed leaf nodes to avoid traffic loss during the reboot of leaf nodes.

ifav109-site1

Pod(s) [2] have fewer than two route reflectors for infra MP-BGP.
Configure spine nodes as route reflector for infra MP-BGP. Make sure that at least one route reflector spine is always up by upgrading/downgrading them in separate groups.

ifav109-site3

Following nodes are not in vPC ['301','302'].
Configure vPC for the listed leaf nodes to avoid traffic loss during the reboot of leaf nodes.

ifav109-site3

Pod(s) [1] have fewer than two route reflectors for infra MP-BGP.
Configure spine nodes as route reflector for infra MP-BGP. Make sure that at least one route reflector spine is always up by upgrading/downgrading them in separate groups.

ifav109-site3

NTP is not configured.
Configure NTP via System > QuickStart > First time setup of the ACI fabric > NTP. This is recommended to avoid any issues in database synchronization between nodes, SSL certificate check, etc.

ifav109-site3

APICs are not running recommended CIMC versions :node-1: 4.0(2f)
Upgrade to the recommended CIMC version. APICs have recommended CIMC versions based on its hardware model and APIC firmware version.

Previous
Next

Step 6 In the **Confirmation** screen, review the information and click **Submit** to start the upgrade.

Step 7 In the **Ready to Install** screen, click **Install** to start the upgrade.

If NDO to site connectivity is lost during the upgrade process, the GUI displays the last known status of the upgrade before loss of connectivity. When connectivity is reestablished, the upgrade status will be refreshed. You can perform a manual Refresh after connectivity loss by clicking the **Refresh** button in the top right of the main pane.

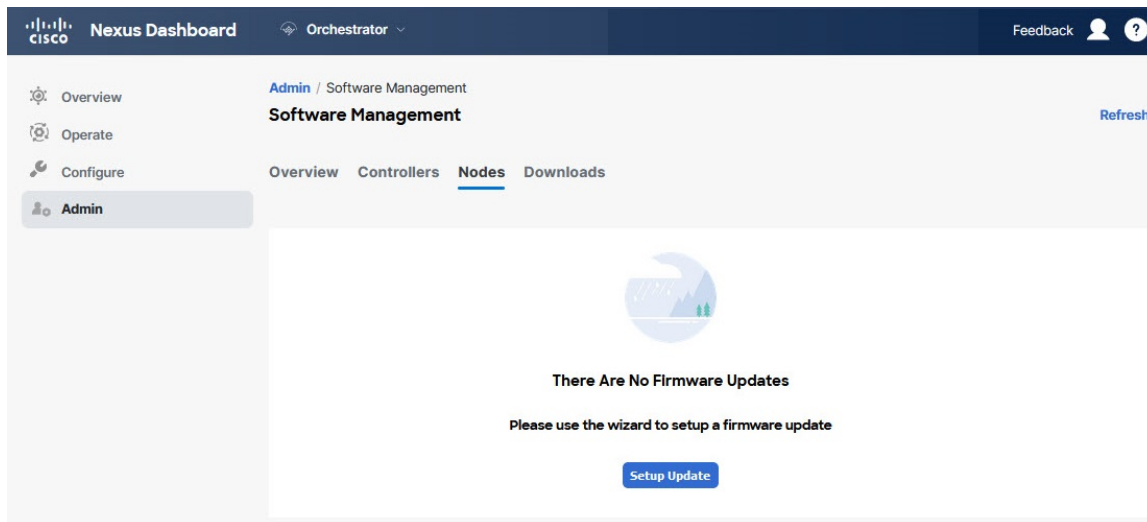
Upgrading Nodes

This section describes how to set up a software upgrade for your sites' switch nodes.

Procedure

Step 1 Log in to your Cisco Nexus Dashboard Orchestrator.

Step 2 Set up switch nodes upgrade.



- From the left navigation pane, select **Admin > Software Management**.
- In the main window, select the **Nodes** tab.
- Click **Setup Update** tab.

If you have previously set up one or more updates, click the **Setup Update** button in the top right of the main pane instead.

The **Setup Node Firmware Update** screen opens.

Step 3 Provide the upgrade details.

- Provide the **Name**.

This is the upgrade policy name that you are able to use to track the upgrade progress at any time.

- Click **Select Nodes**.

The **Select Nodes** window opens.

- Select a site, then select the switch nodes in that site and click **Add and Close**.

You can add switch nodes from a single site at a time. You repeat this step if you want to add switches from other sites.

- Repeat the previous substep for nodes in other sites.

- Click **Next** to proceed.

Step 4 In the **Version Selection** screen, select the firmware version and click **Next**.

The firmware must be downloaded to the sites before it becomes available here. If the download you set up in previous section has completed successfully but the image is still not available here, close the **Setup Site Firmware Update**

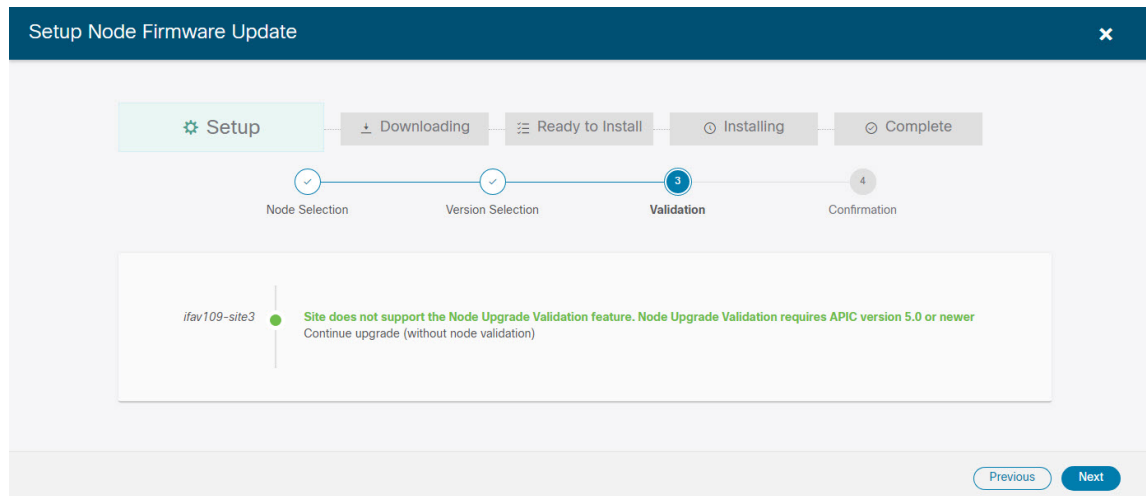
screen, navigate back to **Admin > Software Management > Nodes** tab, and click the **Refresh** button to reload the latest information available for the sites; then restart the upgrade steps.

Step 5 In the **Validation** screen, ensure that there are no faults raised, then click **Next**.

Ensure that there are no faults and review any additional information that may affect your upgrade:

Note

Sites running releases before Release 5.0(1) do not support node validation, so we recommend checking for any switch node faults in the site's APIC before starting the upgrade from NDO.



Step 6 In the **Confirmation** screen, review the information and click **Submit**.

This triggers image to be predownloaded to all the nodes you have selected. After the download completes, the screen will transition to **Ready to Install** and you can proceed to the next step.

Step 7 (Optional) Change **Advanced Settings**.

Note

Review the guidelines, recommendations, and limitations for the Cisco APIC upgrade process described in the [Upgrading and Downgrading the Cisco APIC and Switch Software](#) of the *Cisco APIC Installation, Upgrade, and Downgrade Guide* before making changes to the advanced options.

In the **Ready to Install** screen, you can open the **Advanced Settings** menu for extra options:

- **Ignore Compatibility Check**—By default, the option is set to **No** and compatibility check is enabled and verifies if an upgrade path from the currently running version of the system to a specified newer version is supported.
If you choose to ignore the compatibility check feature, you run the risk of making an unsupported upgrade to your system, which could result in your system going to an unavailable state.
- **Graceful Check**—By default, the option is set to **No** and the upgrade process will not put any of the switches into Graceful Insertion and Removal (GIR) mode before performing the upgrade.
You can choose to enable this option to bring down the node gracefully (using GIR) while performing the upgrade so that the upgrade has reduced traffic loss.
- **Run Mode**—By default, the option is set to **Continue on Failure** and if a node upgrade fails, the process proceeds to the next node. Alternatively, you can set this option to **Pause on Failure** to halt upgrade process if any one of the node upgrades fails.

Step 8 Remove any nodes that are marked as `Failed` from the upgrade.

The upgrade cannot proceed if the upgrade policy contains one or more nodes that failed to download the firmware. You can mouse over the `Failed` status for more information and reason for failure.

To remove the nodes from the upgrade, click **Edit Update Details** link in the **Ready to Install** screen.

Step 9 Click **Install** to start the upgrade.

If NDO to site connectivity is lost during the upgrade process, the GUI displays the last known status of the upgrade before loss of connectivity. When connectivity is reestablished, the upgrade status will be refreshed. You can perform a manual Refresh after connectivity loss by clicking the **Refresh** button in the top right of the main pane.



CHAPTER 10

Tech Support

- [Tech Support and System Logs, on page 143](#)
- [Downloading System Logs, on page 144](#)
- [Streaming System Logs to External Analyzer, on page 144](#)

Tech Support and System Logs

Cisco Nexus Dashboard Orchestrator system logging is automatically enabled when you first deploy the Orchestrator cluster and captures the events and faults that occur in the environment.

You can choose to download the logs at any time or stream them to an external log Analyzer, such as Splunk, if you want to use more tools to quickly parse, view, and respond to important events without a delay.

The tech support logs are split into two parts:

- Original database backup files containing the same information as in prior releases
- JSON-based database backup for ease of readability

Within each backup archive, you find the following contents:

- `x.x.x.x`—One or more files in `x.x.x.x` format for container logs available at the time of the backup.
- `msc-backup-<date>_temp`—Original database backup containing the same information as previous releases.
- `msc-db-json-<date>_temp`—Back up contents in JSON format.

For example:

```
msc_anpEpgRels.json
msc_anpExtEpgRels.json
msc_asyncExecutionStatus.json
msc_audit.json
msc_backup-versions.json
msc_backupRecords.json
msc_ca-cert.json
msc_cloudSecStatus.json
msc_consistency.json
...
```

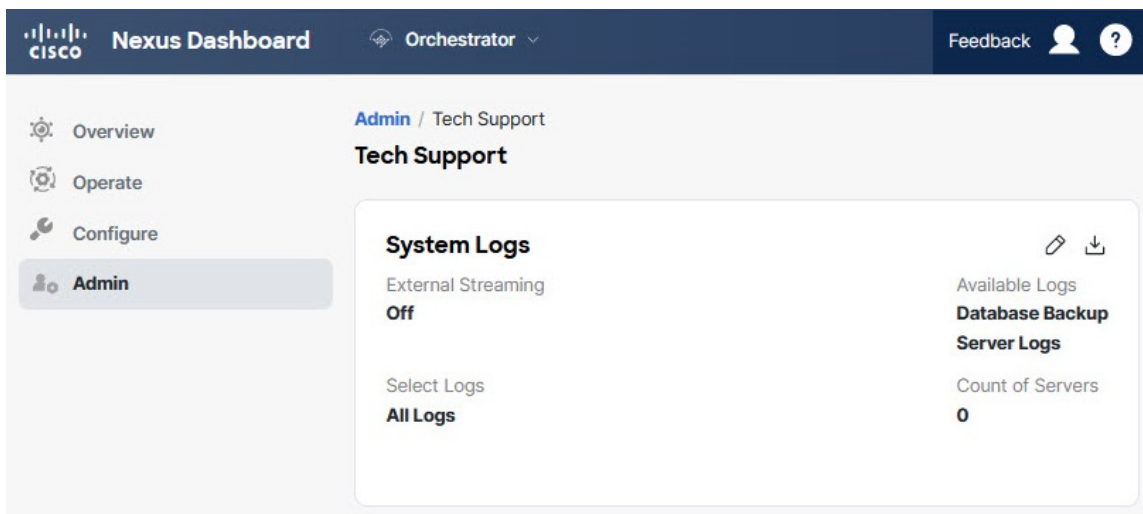
Downloading System Logs

This section describes how to generate a troubleshooting report and infrastructure logs file for all the schemas, sites, tenants, and users that are managed by Cisco Nexus Dashboard Orchestrator.

Procedure

Step 1 Log in to your Cisco Nexus Dashboard Orchestrator GUI.

Step 2 Open the **System Logs** screen.



- a) In the main menu, select **Admin > Software Management**.
- b) In the top-right corner of the **System Logs** frame, click the edit button.

Step 3 Click **Download** download the logs.

An archive will be downloaded to your system. Containing all the information as described in the first section of this chapter.

Streaming System Logs to External Analyzer

Cisco Nexus Dashboard Orchestrator allows you to send the Orchestrator logs to an external log Analyzer tool in real time. By streaming any events as they are generated, you can use the additional tools to quickly parse, view, and respond to important events without a delay.

This section describes how to enable Cisco Nexus Dashboard Orchestrator to stream its logs to an external Analyzer tool, such as Splunk or syslog.

Before you begin

- This release supports only Splunk and `syslog` as external log Analyzer.

- This release supports `syslog` for Cisco Nexus Dashboard Orchestrator in Nexus Dashboard deployments.
- This release supports up to 5 external servers.
- If using Splunk, set up and configure the log Analyzer service provider.

For detailed instructions on how to configure an external log Analyzer, consult its documentation.

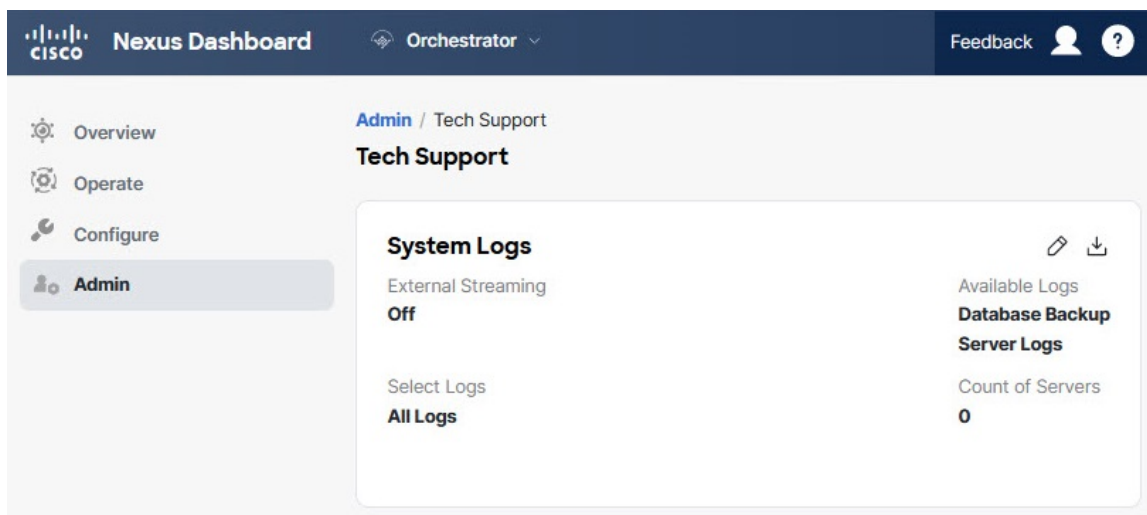
- If using Splunk, obtain an authentication token for the service provider.

Obtaining an authentication token for Splunk service is detailed in the Splunk documentation, but in short, you can get the authentication token by logging into the Splunk server, selecting **Settings > Data Inputs > HTTP Event Collector**, and clicking **New Token**.

Procedure

Step 1 Log in to your Cisco Nexus Dashboard Orchestrator GUI.

Step 2 Open the **Admin > Tech Support > System Logs** screen.



- a) In the main menu, select **Admin > Software Management**.
- b) In the top-right corner of the **System Logs** frame, click the edit button.

Step 3 In the **System Logs** window, enable external streaming and add a server.

System Logs

x

External Streaming



a

Select Logs

All Logs

Audit Logs

b

Logging Servers ⓘ *

c

Server Type	Protocol	Host	Port
+ Add Server			

Save

- Enable the **External Streaming** knob.
- Choose whether you want to stream **All Logs** or just the **Audit Logs**.
- Click **Add Server** to add an external log Analyzer server.

Step 4

Add a Splunk server.

If you do not plan to use Splunk service, skip this step.

Logging Servers ⓘ *

Server Type	Protocol	Host	Port
Select Service			
splunk			
Protocol			
HTTP	HTTPS		
Host *			
Port *			
Token *			
Index ⓘ			
main			

Cancel Save

+ Add Server

Save

d

- a) Choose `Splunk` for the server type.
- b) Choose the protocol.
- c) Provide the server name or IP address, port, and the authentication token you obtained from the Splunk service.

Obtaining an authentication token for Splunk service is detailed in the Splunk documentation, but in short, you can get the authentication token by logging into the Splunk server, selecting **Settings > Data Inputs > HTTP Event Collector**, and clicking **New Token**.

- d) Click the check mark icon to finish adding the server.

Step 5

Add a `syslog` server.

If you do not plan to use `syslog`, skip this step.

System Logs



Logging Servers ⓘ *

Server Type	Protocol	Host	Port
Select Service			
<div>syslog</div> <div>a</div>			
Protocol			
<div>TCP</div> <div>UDP</div> <div>b</div>			
Host *			
<div>10.30.11.69</div>			
Port *			
<div>8088</div> <div>c</div>			
Severity			
<div>Alert</div> <div>d</div>			
TLS			
<input type="checkbox"/>			
<div>Cancel</div> <div>Save</div>			
<div>+ Add Server</div>			

- Choose `syslog` for the server type.
- Choose the protocol.
- Provide the server name or IP address, port number, and the severity level of the log messages to stream.
- Click the check mark icon to finish adding the server.

Step 6 Repeat the steps if you want to add multiple servers.

This release supports up to 5 external servers.

Step 7 Click **Save** to save the changes.

System Logs

×

Download Logs

Download

External Streaming

☐

Select Logs

All Logs

Audit Logs

* Logging Servers ⓘ

Server Type	Protocol	Host	Port	
splunk	http	10.30.11.69	8088	✖
syslog	tcp	10.195.223.220	514	✖

+

 Add Server

SAVE



PART III

Infrastructure Management

- [System Configuration, on page 153](#)
- [Preparing Cisco APIC Sites, on page 155](#)
- [Adding and Deleting Sites, on page 163](#)
- [Configuring Infra General Settings, on page 169](#)
- [Configuring Infra for Cisco APIC Sites, on page 177](#)
- [Configuring Infra for Cisco Cloud Network Controller Sites, on page 185](#)
- [Deploying Infra Configuration for ACI Sites, on page 191](#)
- [CloudSec Encryption, on page 197](#)



System Configuration

- [System Configuration Settings, on page 153](#)
- [System Alias and Banner, on page 153](#)

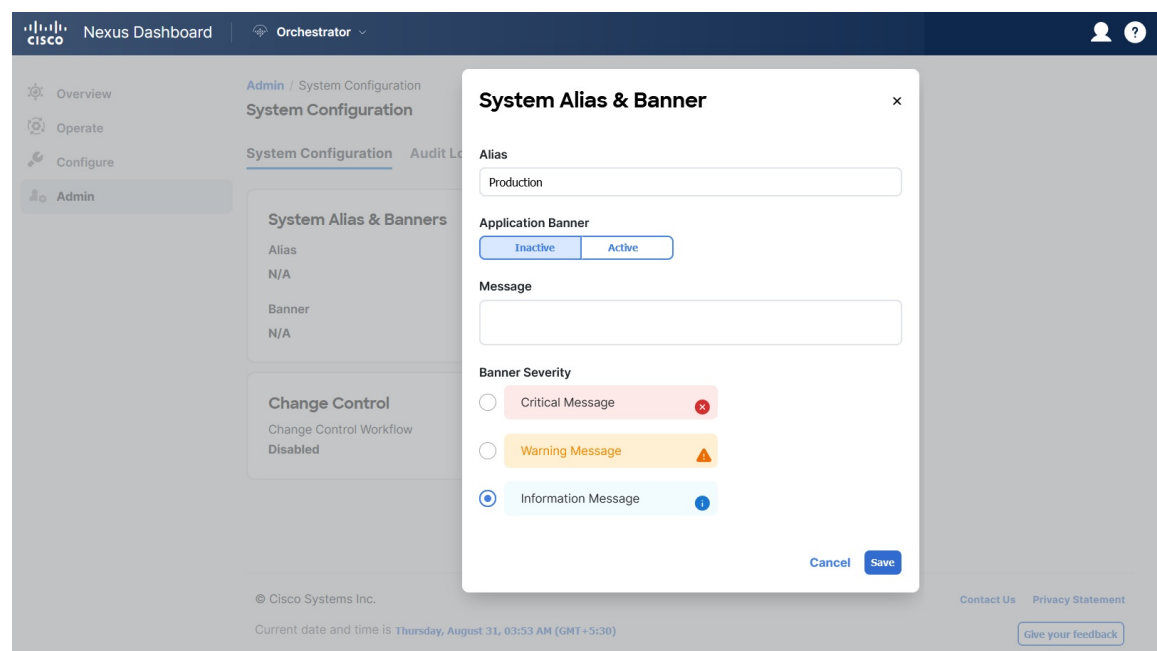
System Configuration Settings

There is a number of global system settings that are available under **Admin > System Configuration**, which you can configure for your Multi-Site Orchestrator as described in the following sections.

System Alias and Banner

This section describes how to configure an alias for your Nexus Dashboard Orchestrator as well as enable a custom GUI-wide banner to be displayed at the top of your screen, as shown in the following figure.

Figure 13: System Banner Display



Procedure

- Step 1** Log in to your Orchestrator.
- Step 2** From the left navigation pane, select **Admin > System Configuration**.
- Step 3** Click the **Edit** icon to the right of the **System Alias & Banners** area.
This opens the **System Alias & Banners** settings window.
- Step 4** In the **Alias** field, specify the system alias.
- Step 5** Choose whether you want to enable the GUI banner.
- Step 6** If you enable the banner, you must provide the message that will be displayed on it.
- Step 7** If you enable the banner, you must choose the severity, or color, for the banner.
- Step 8** Click **Save** to save the changes.
-



CHAPTER 12

Preparing Cisco APIC Sites

- [Pod Profile and Policy Group, on page 155](#)
- [Configuring Fabric Access Policies for All APIC Sites, on page 156](#)
- [Configuring Sites That Contain Remote Leaf Switches, on page 159](#)
- [Cisco Mini ACI Fabrics, on page 160](#)

Pod Profile and Policy Group

In each site's APIC, you must have one Pod profile with a Pod policy group. If your site does not have a Pod policy group you must create one. Typically, these settings will already exist as you will have configured them when you first deployed the fabric.

Procedure

- Step 1** Log in to the site's APIC GUI.
- Step 2** Check that the Pod profile contains a Pod policy group.
Navigate to **Fabric > Fabric Policies > Pods > Profiles > Pod Profile default**.
- Step 3** If necessary, create a Pod policy group.
a) Navigate to **Fabric > Fabric Policies > Pods > Policy Groups**.
b) Right-click **Policy Groups** and select **Create Pod Policy Group**.
c) Enter the appropriate information and click **Submit**.
- Step 4** Assign the new Pod policy group to the default Pod profile.
a) Navigate to **Fabric > Fabric Policies > Pods > Profiles > Pod Profile default**
b) Select the default profile.
c) Choose the new pod policy group and click **Update**.
-

Configuring Fabric Access Policies for All APIC Sites

Before your APIC fabrics can be added to and managed by the Nexus Dashboard Orchestrator, there is a number of fabric-specific access policies that you must configure on each site.

Configuring Fabric Access Global Policies

This section describes the global fabric access policy configurations that must be created for each APIC site before it can be added to and managed by the Nexus Dashboard Orchestrator.

Procedure

Step 1 Log in directly to the site's APIC GUI.

Step 2 From the main navigation menu, select **Fabric > Access Policies**.

You must configure a number of fabric policies before the site can be added to the Nexus Dashboard Orchestrator. From the APIC's perspective, this is something you do just like you would if you were connecting a bare-metal host, where you would configure domains, AEPs, policy groups, and interface selectors; you must configure the same options for connecting the spine switch interfaces to the inter-site network for all the sites that will be part of the same Multi-Site domain.

Step 3 Specify the VLAN pool.

The first thing you configure is the VLAN pool. We use Layer 3 sub-interfaces tagging traffic with VLAN-4 to connect the spine switches to the inter-site network.

- a) In the left navigation tree, browse to **Pools > VLAN**.
- b) Right-click the **VLAN** category and choose **Create VLAN Pool**.

In the **Create VLAN Pool** window, specify the following:

- For the **Name** field, specify the name for the VLAN pool, for example `msite`.
- For **Allocation Mode**, specify `Static Allocation`.
- And for the **Encap Blocks**, specify just the single VLAN 4. You can specify a single VLAN by entering the same number in both **Range** fields.

Step 4 Configure Attachable Access Entity Profiles (AEP).

- a) In the left navigation tree, browse to **Global Policies > Attachable Access Entity Profiles**.
- b) Right-click the **Attachable Access Entity Profiles** category and choose **Create Attachable Access Entity Profiles**.

In the **Create Attachable Access Entity Profiles** window, specify the name for the AEP, for example `msite-aep`.

- c) Click **Next** and **Submit**

No additional changes, such as interfaces, are required.

Step 5 Configure domain.

The domain you configure is what you will select from the Nexus Dashboard Orchestrator when adding this site.

- a) In the left navigation tree, browse to **Physical and External Domains > External Routed Domains**.
- b) Right-click the **External Routed Domains** category and choose **Create Layer 3 Domain**.

In the **Create Layer 3 Domain** window, specify the following:

- For the **Name** field, specify the name the domain, for example `msite-13`.
- For **Associated Attachable Entity Profile**, select the AEP you created in Step 4.
- For the **VLAN Pool**, select the VLAN pool you created in Step 3.

- c) Click **Submit**.

No additional changes, such as security domains, are required.

What to do next

After you have configured the global access policies, you must still add interfaces policies as described in [Configuring Fabric Access Interface Policies, on page 157](#).

Configuring Fabric Access Interface Policies

This section describes the fabric access interface configurations that must be done for the Nexus Dashboard Orchestrator on each APIC site.

Before you begin

You must have configured the global fabric access policies, such as VLAN Pool, AEP, and domain, in the site's APIC, as described in [Configuring Fabric Access Global Policies, on page 156](#).

Procedure

Step 1 Log in directly to the site's APIC GUI.

Step 2 From the main navigation menu, select **Fabric > Access Policies**.

In addition to the VLAN, AEP, and domain you have configured in previous section, you must also create the interface policies for the fabric's spine switch interfaces that connect to the Inter-Site Network (ISN).

Step 3 Configure a spine policy group.

- a) In the left navigation tree, browse to **Interface Policies > Policy Groups > Spine Policy Groups**.
This is similar to how you would add a bare-metal server, except instead of a Leaf Policy Group, you are creating a Spine Policy Group.
- b) Right-click the **Spine Policy Groups** category and choose **Create Spine Access Port Policy Group**.

In the **Create Spine Access Port Policy Group** window, specify the following:

- For the **Name** field, specify the name for the policy group, for example `Spine1-PolGrp`.
- For the **Link Level Policy** field, specify the link policy used between your spine switch and the ISN.

- For **CDP Policy**, choose whether you want to enable CDP.
- For the **Attached Entity Profile**, select the AEP you have configured in previous section, for example `msite-aep`.

c) Click **Submit**.

No additional changes, such as security domains, are required.

Step 4 Configure a spine profile.

- a) In the left navigation tree, browse to **Interface Policies > Profiles > Spine Profiles**.
- b) Right-click the **Spine Profiles** category and choose **Create Spine Interface Profile**.

In the **Create Spine Interface Profile** window, specify the following:

- For the **Name** field, specify the name for the profile, for example `Spine1-ISN`.
- For **Interface Selectors**, click the + sign to add the port on the spine switch that connects to the ISN. Then in the **Create Spine Access Port Selector** window, provide the following:
 - For the **Name** field, specify the name for the port selector, for example `Spine1-ISN`.
 - For the **Interface IDs**, specify the switch port that connects to the ISN, for example `5/32`.
 - For the **Interface Policy Group**, choose the policy group you created in the previous step, for example `Spine1-PolGrp`.

Then click **OK** to save the port selector.

c) Click **Submit** to save the spine interface profile.

Step 5 Configure a spine switch selector policy.

- a) In the left navigation tree, browse to **Switch Policies > Profiles > Spine Profiles**.
- b) Right-click the **Spine Profiles** category and choose **Create Spine Profile**.

In the **Create Spine Profile** window, specify the following:

- For the **Name** field, specify the name for the profile, for example `Spine1`.
- For **Spine Selectors**, click the + to add the spine and provide the following:
 - For the **Name** field, specify the name for the selector, for example `Spine1`.
 - For the **Blocks** field, specify the spine node, for example `201`.

- c) Click **Update** to save the selector.
- d) Click **Next** to proceed to the next screen.
- e) Select the interface profile you have created in the previous step

For example `Spine1-ISN`.

f) Click **Finish** to save the spine profile.

Configuring Sites That Contain Remote Leaf Switches

Multi-Site architecture supports APIC sites with Remote Leaf switches. The following sections describe guidelines, limitations, and configuration steps required to allow Nexus Dashboard Orchestrator to manage these sites.

Remote Leaf Guidelines and Limitations

If you want to add an APIC site with a Remote Leaf to be managed by the Nexus Dashboard Orchestrator, the following restrictions apply:

- You must upgrade your Cisco APIC to Release 4.2(4) or later.
- Only physical Remote Leaf switches are supported in this release
- Only -EX and -FX or later switches are supported as Remote Leaf switches for use with Multi-Site
- Remote Leaf is not supported with back-to-back connected sites without IPN switches
- Remote Leaf switches in one site cannot use another site's L3Out
- Stretching a bridge domain between one site and a Remote Leaf in another site is not supported

You must also perform the following tasks before the site can be added to and managed by the Nexus Dashboard Orchestrator:

- You must enable Remote Leaf direct communication and configure routable subnets directly in the site's APIC, as described in the following sections.
- You must add the routable IP addresses of Cisco APIC nodes in the DHCP-Relay configuration applied on the interfaces of the Layer 3 routers connecting to the Remote Leaf switches.

The routable IP address of each APIC node is listed in the **Routable IP** field of the **System > Controllers > <controller-name>** screen of the APIC GUI.

Configuring Routable Subnets for Remote Leaf Switches

Before you can add a site that contains one or more Remote Leaf switches to the Nexus Dashboard Orchestrator, you must configure routable subnets for the pod with which the Remote Leaf nodes are associated.

Procedure

-
- | | |
|---------------|---|
| Step 1 | Log in directly to the site's APIC GUI. |
| Step 2 | From the menu bar, select Fabric > Inventory . |
| Step 3 | In the Navigation pane, click Pod Fabric Setup Policy . |
| Step 4 | In the main pane, double-click the pod where you want to configure the subnets. |
| Step 5 | In the Routable Subnets area, click the + sign to add a subnet. |
| Step 6 | Enter the IP and Reserve Address Count , set the state to Active or Inactive , then click Update to save the subnet. |

When configuring routable subnets, you must provide a netmask between /22 and /29.

Step 7 Click **Submit** to save the configuration.

Enabling Direct Communication for Remote Leaf Switches

Before you can add a site that contains one or more Remote Leaf switches to the Nexus Dashboard Orchestrator, you must configure direct remote leaf communication for that site. Additional information about remote leaf direct communication feature is available in the *Cisco APIC Layer 3 Networking Configuration Guide*. This section outlines the steps and guidelines specific to the integration with Multi-Site.



Note Once you enable Remote Leaf switch direct communication, the switches will function in the new mode only

Procedure

Step 1 Log in directly to the site's APIC.

Step 2 Enable direct traffic forwarding for Remote Leaf switches.

- a) From the menu bar, navigate to **System > System Settings**.
- b) From the left side bar, select **Fabric Wide Setting**.
- c) Check the **Enable Remote Leaf Direct Traffic Forwarding** checkbox.

Note

You cannot disable this option after you enable it.

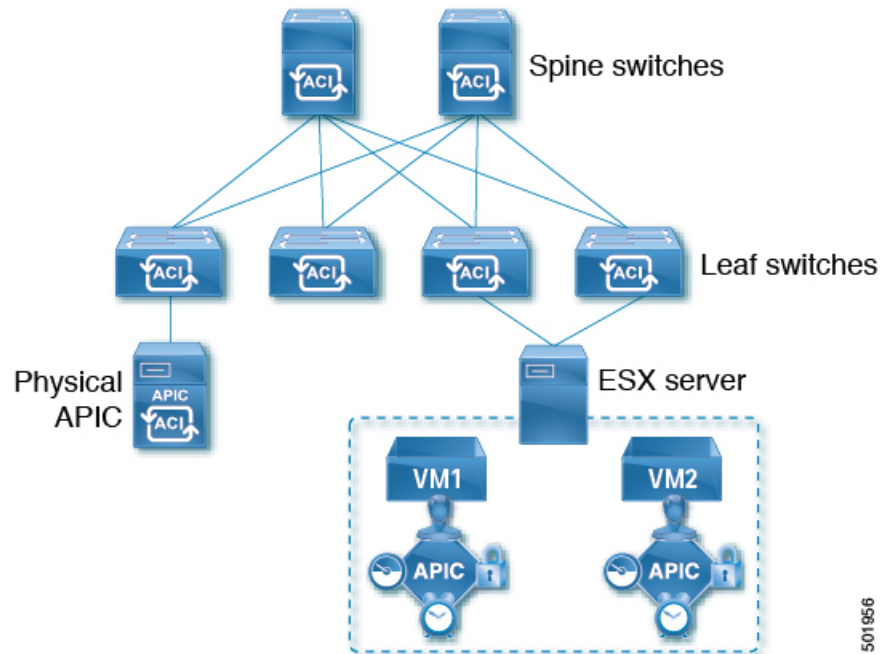
- d) Click **Submit** to save the changes.

Cisco Mini ACI Fabrics

Cisco Multi-Site supports Cisco Mini ACI fabrics as typical on-premises sites without requiring any additional configuration. This section provides a brief overview of Mini ACI fabrics, detailed info on deploying and configuring this type of fabrics is available in [Cisco Mini ACI Fabric and Virtual APICs](#).

Cisco ACI, Release 4.0(1) introduced Mini ACI Fabric for small scale deployment. Mini ACI fabric works with Cisco APIC cluster consisting of one physical APIC and two virtual APICs (vAPIC) running in virtual machines. This reduces the physical footprint and cost of the APIC cluster, allowing ACI fabric to be deployed in scenarios with limited rack space or initial budget, such as a colocation facility or a single-room data center, where a full-scale ACI installations may not be practical due to physical footprint or initial cost.

The following diagram shows an example of a mini Cisco ACI fabric with a physical APIC and two virtual APICs (vAPICs):

Figure 14: Cisco Mini ACI Fabric

501956



CHAPTER 13

Adding and Deleting Sites

- [Cisco NDO and APIC Interoperability Support, on page 163](#)
- [Adding Cisco ACI Sites, on page 165](#)
- [Removing Sites, on page 166](#)
- [Cross Launch to Fabric Controllers, on page 168](#)

Cisco NDO and APIC Interoperability Support

Cisco Nexus Dashboard Orchestrator (NDO) does not require a specific version of APIC to be running in all sites. The APIC clusters in each site as well as the NDO itself can be upgraded independently of each other and run in mixed operation mode as long as the fabric can be on-boarded to the Nexus Dashboard where the Nexus Dashboard Orchestrator service is installed. As such, we recommend that you always upgrade to the latest release of the Nexus Dashboard Orchestrator.

However, keep in mind that if you upgrade the NDO before upgrading the APIC clusters in one or more sites, some of the new NDO features may not yet be supported by an earlier APIC release. In that case a check is performed on each template to ensure that every configured option is supported by the target sites.

The check is performed when you save a template or deploy a template. If the template is already assigned to a site, any unsupported configuration options will not be saved; if the template is not yet assigned, you will be able to assign it to a site, but not be able to save or deploy the schema if it contains configuration unsupported by that site.

In case an unsupported configuration is detected, an error message will show, for example: This APIC site version `<site-version>` is not supported by NDO. The minimum version required for this `<feature>` is `<required-version>` or above.

The following table lists the features and the minimum required APIC release for each one:



Note While some of the following features are supported on earlier Cisco APIC releases, Release 4.2(4) is the earliest release that can be on-boarded to the Nexus Dashboard and managed by this release of Nexus Dashboard Orchestrator.

Feature	Minimum APIC Version
ACI Multi-Pod Support	Release 4.2(4)

Feature	Minimum APIC Version
Service Graphs (L4-L7 Services)	Release 4.2(4)
External EPGs	Release 4.2(4)
ACI Virtual Edge VMM Support	Release 4.2(4)
DHCP Support	Release 4.2(4)
Consistency Checker	Release 4.2(4)
vzAny	Release 4.2(4)
Host Based Routing	Release 4.2(4)
CloudSec Encryption	Release 4.2(4)
Layer 3 Multicast	Release 4.2(4)
MD5 Authentication for OSPF	Release 4.2(4)
EPG Preferred Group	Release 4.2(4)
Intersite L3Out	Release 4.2(4)
EPG QoS Priority	Release 4.2(4)
Contract QoS Priority	Release 4.2(4)
Single Sign-On (SSO)	Release 5.0(1)
Multicast Rendezvous Point (RP) Support	Release 5.0(1)
Transit Gateway (TGW) support for AWS and Azure Sites	Release 5.0(1)
SR-MPLS Support	Release 5.0(1)
Cloud LoadBalancer High Availability Port	Release 5.0(1)
Service Graphs (L4-L7 Services) with UDR	Release 5.0(2)
3rd Party Device Support in Cloud	Release 5.0(2)
Cloud Loadbalancer Target Attach Mode Feature	Release 5.1(1)
Support security and service insertion in Azure for non-ACI networks reachable through Express Route	Release 5.1(1)
CSR Private IP Support	Release 5.1(1)
Extend ACI policy model and automation for Cloud native services in Azure	Release 5.1(1)

Feature	Minimum APIC Version
Flexible segmentation through multiple VRF support within a single VNET for Azure	Release 5.1(1)
Private Link automation for Azure PaaS and third-party services	Release 5.1(1)
Openshift 4.3 IPI on Azure with ACI-CNI	Release 5.1(1)
Cloud Site Underlay Configuration	Release 5.2(1)

Adding Cisco ACI Sites

This section describes how to add a Cisco APIC or Cloud Network Controller site using the Cisco Nexus Dashboard GUI and then enable that site to be managed by Cisco Nexus Dashboard Orchestrator.

Before you begin

- If you are adding on-premises ACI site, you must have completed the site-specific configurations in each site's APIC, as described in previous sections in this chapter.
- You must ensure that one or more sites you are adding are running Release 4.2(4) or later.

Procedure

Step 1 Log in to your Cisco Nexus Dashboard and open the **Admin Console**.

Step 2 From the left navigation menu, choose **Operate** and click **Sites**.

Step 3 Choose **Add Site** and provide site information.

- a) For **Site Type**, select **ACI** or **Cloud Network Controller** depending on the type of ACI fabric you are adding.
- b) Provide the controller information.

- You must provide the **Host Name/IP Address**, **User Name**, and **Password** for the APIC controller currently managing your ACI fabrics.

Note

For APIC fabrics, if you use the site with Cisco Nexus Dashboard Orchestrator service only, you can provide either the in-band or out-of-band IP address of the APIC. If you use the site with Cisco Nexus Dashboard Insights as well, you must provide the in-band IP address.

- For on-premises ACI sites managed by Cisco APIC, if you plan to use this site with Day-2 Operations applications such as Cisco Nexus Insights, you must also provide the **In-Band EPG** name that is used to connect the Cisco Nexus Dashboard to the fabric you are adding. Otherwise, if you use this site with Cisco Nexus Dashboard Orchestrator only, you can leave this field blank.
- For Cloud Network Controller sites, **Enable Proxy** if your cloud site is reachable through a proxy.

Proxy must be already configured in your Cisco Nexus Dashboard's cluster settings. If the proxy is reachable through management network, a static management network route must also be added for the proxy IP address. For more information about proxy and route configuration, see [Nexus Dashboard User Guide](#) for your release.

- c) Click **Save** to finish adding the site.

Currently, the sites are available in the Cisco Nexus Dashboard, but you still must enable them for Cisco Nexus Dashboard Orchestrator management as described in the following steps.

Step 4 Repeat the previous steps for any additional ACI or Cloud Network Controller sites.

Step 5 From the Cisco Nexus Dashboard's **Services** page, open the Cisco Nexus Dashboard Orchestrator service. You are automatically signed in using the Cisco Nexus Dashboard user's credentials.

Step 6 In the Cisco Nexus Dashboard Orchestrator GUI, manage the sites.

- a) From the left navigation menu, select **Sites**.
- b) In the main pane, change the **State** from *Unmanaged* to *Managed* for each fabric that you want the NDO to manage.

When managing the sites, you must provide a unique site ID for each site.

Note

Ensure that ACI site names are limited to 125 characters or less to avoid any issues when enabling orchestration.

Removing Sites

This section describes how to disable site management for one or more sites using the Cisco Nexus Dashboard Orchestrator GUI. The sites remain present in the Cisco Nexus Dashboard.

Before you begin

You must ensure that all templates associated with the site you want to remove are not deployed.

Procedure

Step 1 Open the Cisco Nexus Dashboard Orchestrator GUI.

You can open the NDO service from the Cisco Nexus Dashboard's **Service Catalog**. You are automatically signed in using the Cisco Nexus Dashboard user's credentials.

Step 2 Remove the site from all templates.

You must remove the site from all templates with which it is associated before you can unmanage the site and remove it from your Cisco Nexus Dashboard.

- a) Navigate to **Configure > Tenant Template > Applications**.
- b) Click a **Schema** that contains one or more templates that are associated with the site.
- c) From the **Overview** drop-down, choose a template that's associated with the site that you want to remove.
- d) From the **Actions** drop-down, choose **Add/Remove Sites** and uncheck the site that you want to remove.

This removes configurations that were deployed using this template to this site.

Note

For nonstretched templates, you can choose to preserve the configurations that are deployed by the template to the sites by selecting **Actions > Dissociate Sites** instead. This option allows you to retain configurations that are deployed by NDO but no longer manage those objects from NDO.

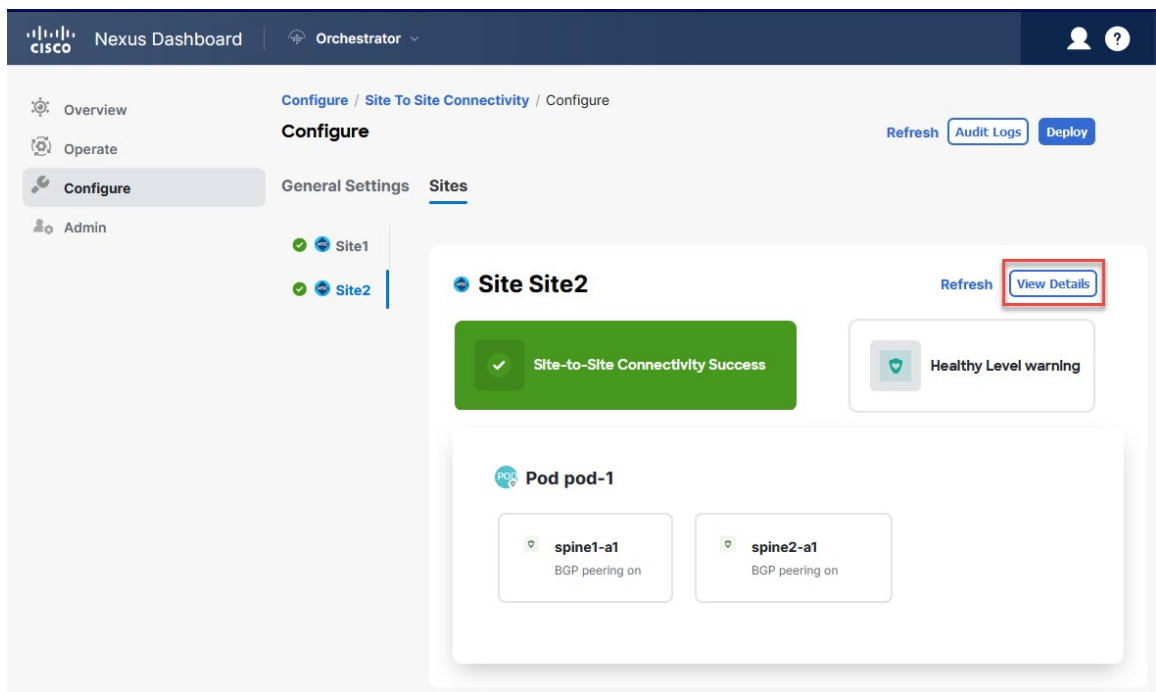
- e) Repeat this step for all templates associated with the site that you want to unmanage in this and all other schemas.

Step 3

Remove the site's underlay configuration.

- From the left navigation menu, select **Configure > Site To Site Connectivity**.
- In the main pane, click **Configure**.
- In the left sidebar, select the site that you want to unmanage.
- Click **View Details** to load site settings.

Figure 15:



- e) In right sidebar's **Inter-Site Connectivity** tab, disable the **Multi-Site** check box.

This disables EVPN peering between this site and other sites.

- f) Click **Deploy** to deploy the changes to the site.

Step 4

In the Cisco Nexus Dashboard Orchestrator GUI, disable the sites.

- From the left navigation menu, select **Sites**.
- In the main pane, change the **State** from **Managed** to **Unmanaged** for the site that you want to unmanage.

Note

If the site is associated with one or more deployed templates, you will not be able to change its state to **Unmanaged** until you undeploy those templates, as described in the previous step.

Step 5 Delete the site from Cisco Nexus Dashboard.

If you no longer want to manage this site or use it with any other applications, you can delete the site from the Cisco Nexus Dashboard as well.

Note

The site must not be currently in use by any of the services that are installed in your Cisco Nexus Dashboard cluster.

- a) In the top navigation bar, click the **Home** icon to return to the Cisco Nexus Dashboard GUI.
- b) From the left navigation menu of the Cisco Nexus Dashboard GUI, select **Operate > Sites**.
- c) Select one or more sites that you want to delete.
- d) In the top right of the main pane, select **Actions > Delete Site**.
- e) Provide the site's sign-in information and click **OK**.

The site will be removed from the Cisco Nexus Dashboard.

Cross Launch to Fabric Controllers

Cisco Nexus Dashboard Orchestrator currently supports several configuration options for each type of fabrics. For many extra configuration options, you may need to sign in directly into the fabric's controller.

You can cross-launch into the specific site controller's GUI from the NDO's **Operate > Sites** screen by selecting the actions (. . .) menu next to the site and clicking **Open in user interface**. Cross-launch works with out-of-band (OOB) management IP of the fabric.

If the same user is configured in Cisco Nexus Dashboard and the fabric, you will be signed in automatically into the fabric's controller using the same log in information as the Cisco Nexus Dashboard user. For consistency, we recommend configuring remote authentication with common users across Cisco Nexus Dashboard and the fabrics.



CHAPTER 14

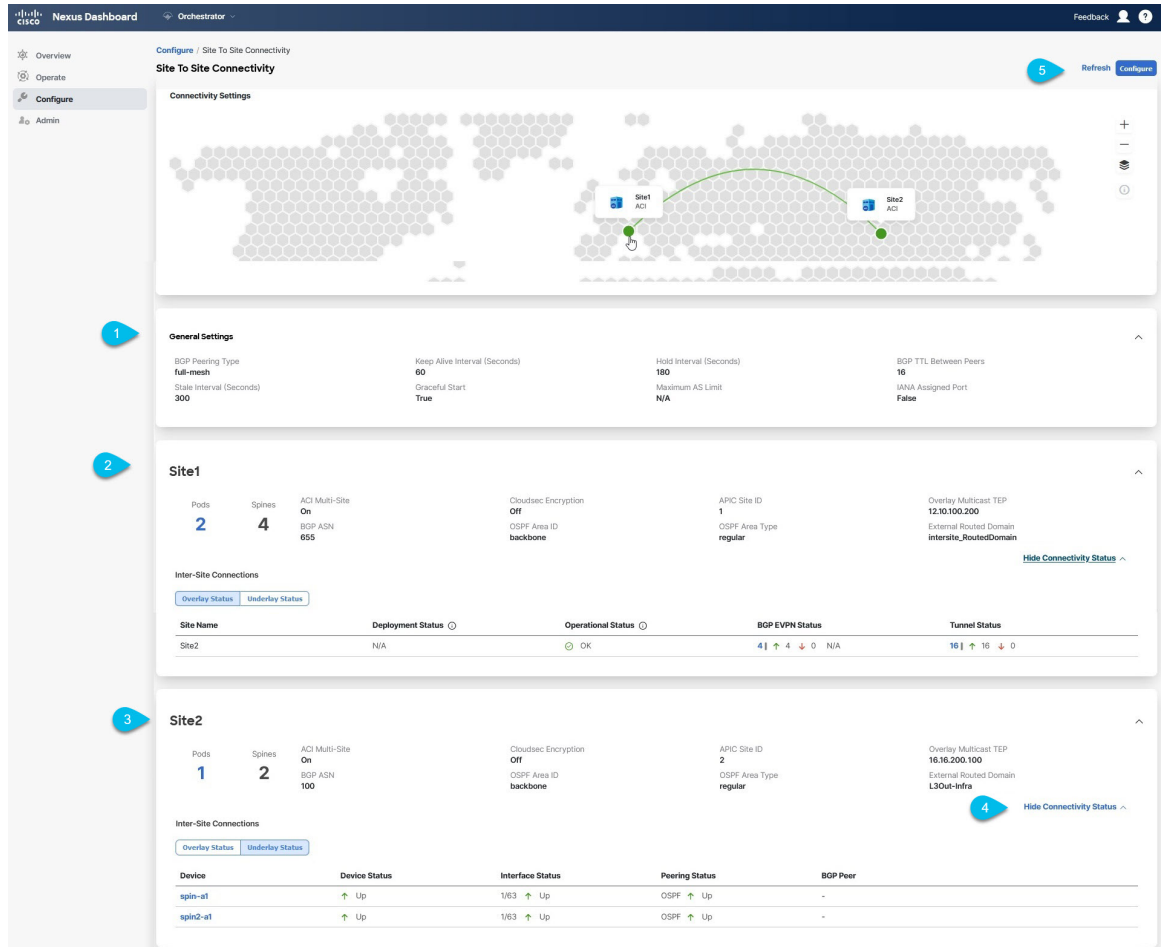
Configuring Infra General Settings

- [Infra Configuration Dashboard, on page 169](#)
- [Partial Mesh Intersite Connectivity, on page 171](#)
- [Configuring Infra: General Settings, on page 171](#)

Infra Configuration Dashboard

The **Config > Site To Site Connectivity** page displays a summary of all sites and intersite connectivity in your Cisco Nexus Dashboard Orchestrator deployment and contains the following information:

Figure 16: Infra Configuration Overview



1. The **General Settings** tile displays information about BGP peering type and its configuration. This is described in detail in the next section.
2. The **On-Premises** tiles display information about every on-premises site that is part of your Multi-Site domain along with their number of Pods and spine switches, OSPF settings, and overlay IPs. You can click the **Pods** tile that displays the number of Pods in the site to show information about the Overlay Unicast TEP addresses of each Pod. This is described in detail in [Configuring Infra for Cisco APIC Sites, on page 177](#).
3. The **Cloud** tiles display information about every cloud site that is part of your Multi-Site domain along with their number of regions and basic site information. This is described in detail in [Configuring Infra for Cisco Cloud Network Controller Sites, on page 185](#).
4. You can click **Show Connectivity Status** to display intersite connectivity details for a specific site.
5. You can use the **Configure** button to navigate to the intersite connectivity configuration, which is described in detail in the following sections.

The following sections describe the steps necessary to configure the general fabric Infra settings. Fabric-specific requirements and procedures are described in the following chapters based on the specific type of fabric that you are managing.

Before you proceed with Infra configuration, you must have configured and added the sites as described in previous sections.

In addition, any infrastructure changes such as adding and removing spine switches or spine node ID changes require a Cisco Nexus Dashboard Orchestrator fabric connectivity information refresh described in the [Refreshing Site Connectivity Information, on page 177](#) as part of the general Infra configuration procedures.

Partial Mesh Intersite Connectivity

In addition to full mesh connectivity where you configure intersite connectivity from every site managed by your Nexus Dashboard Orchestrator to every other site, this release also supports partial mesh configuration. In partial mesh configuration, you can manage sites in standalone mode with no intersite connectivity to any other site or limit the intersite configuration to only a subset of other sites in your Multi-Site domain.

Prior to Nexus Dashboard Orchestrator, Release 3.6(1), you could stretch templates between sites and refer to policies from other templates, which were deployed to other sites, even if the intersite connectivity between those sites was not configured, resulting in intended traffic flow between the sites to not work.

Beginning with release 3.6(1), the Orchestrator will allow you to stretch template and remote reference policies from other templates (deployed on other sites) between two or more sites only if the intersite connectivity between those sites is properly configured and deployed.

When configuring site infra for Cisco APIC and Cisco Cloud Network Controller sites as described in the following sections, for each site you can explicitly choose to which other sites infra connectivity will be established and provide that configuration information only.

Partial Mesh Connectivity Guidelines

When configuring partial mesh connectivity, consider the following guidelines:

- Partial mesh connectivity is supported between two cloud sites or a cloud and on-premises site.

Full mesh connectivity is automatically established between all on-premises sites.

- Partial mesh connectivity is supported using BGP-EVPN or BGP-IPv4 protocols.

Note however that stretching a template is allowed only for sites that are connected using BGP-EVPN protocol. If you are using BGP-IPv4 to connect two or more sites, any template assigned to any of those sites can be deployed to one site only.

Configuring Infra: General Settings

This section describes how to configure general Infra settings for all the sites.



Note

Some of the following settings apply to all sites, while others are required for specific type of sites (for example, Cloud Network Controller sites). Ensure that you complete all the required configurations in infra general settings before proceeding to the site-local settings specific to each site.

Procedure

-
- Step 1** Log in to the Cisco Nexus Dashboard Orchestrator GUI.
- Step 2** In the left navigation menu, select **Configure > Site To Site Connectivity**.
- Step 3** In the main pane, click **Configure**.
- Step 4** In the left sidebar, select **General Settings**.
- Step 5** Provide **Control Plane Configuration**.
- Select the **Control Plane Configuration** tab.
 - Choose **BGP Peering Type**.
 - full-mesh**—All border gateway switches in each site establishes peer connectivity with remote sites' border gateway switches.
In **full-mesh** configuration, Cisco Nexus Dashboard Orchestrator uses the spine switches for ACI-managed fabrics and border gateways for NDFC-managed fabrics.
 - route-reflector**—The route-reflector option allows you to specify one or more control-plane nodes to which each site establishes MP-BGP EVPN sessions. The use of route-reflector nodes avoids creating MP-BGP EVPN full mesh adjacencies between all the sites that are managed by NDO.
For ACI fabrics, the **route-reflector** option is effective only for fabrics that are part of the same BGP ASN.
 - In the **Keepalive Interval (Seconds)** field, enter the keepalive interval seconds.
We recommend keeping the default value.
 - In the **Hold Interval (Seconds)** field, enter the hold interval seconds.
We recommend keeping the default value.
 - In the **Stale Interval (Seconds)** field, enter stale interval seconds.
We recommend keeping the default value.
 - Choose whether you want to turn on the **Graceful Helper** option.
 - Provide the **Maximum AS Limit**.
We recommend keeping the default value.
 - Provide the **BGP TTL Between Peers**.
We recommend keeping the default value.
 - Provide the **OSPF Area ID**.
If you do not have any Cloud Network Controller sites, this field will not be present in the UI.
This is OSPF area ID used by cloud sites for on-premises IPN peering.
 - (Optional) Enable **IANA Assigned Port** for CloudSec encryption.
By default, CloudSec uses a proprietary UDP port. This option allows you to configure CloudSec to use the official IANA-reserved port 8017 for CloudSec encryption between sites.

Note

The IANA-reserved port is supported for Cisco APIC sites running release 5.2(4) or later.

To change this setting, CloudSec must be disabled on all sites. If you want to enable IANA reserved port, but already have CloudSec encryption that is enabled for one or more of your sites, disable CloudSec for all sites, enable **IANA Reserve UDP Port** option, then re-enable CloudSec for the required sites.

For detailed information and steps for configuring CloudSec, see the "CloudSec Encryption" chapter of the [Nexus Dashboard Orchestrator Configuration Guide for ACI Fabrics](#).

Step 6 Provide the **IPN Devices** information.

If you do not plan to configure intersite connectivity between on-premises and cloud sites, you can skip this step.

When you configure intersite underlay connectivity between on-premises and cloud sites as described in later sections, you must select an on-premises IPN device which establishes connectivity to the cloud CSRs. These IPN devices must first be defined here before they are available in the on-premises site configuration screen, which is described in more detail in [Configuring Infra: On-Premises Site Settings, on page 178](#).

- a) Select the **On Premises IPsec Devices** tab.
- b) Click **+Add On-Premises IPsec Device**.
- c) Choose whether the device is **Unmanaged** or **Managed** and provide the device information.

This defines whether the device is directly managed by NDFC:

- For **Unmanaged** IPN devices, simply provide the **Name** and the **IP Address** of the device.

The IP address that you provide will be used as the tunnel peer address from the cloud CSRs, not the IPN device's management IP address.

- For **Managed** IPN devices, choose the NDFC **Site** that contains the device and then the **Device** from that site.

Then choose the **Interface** on the device that is facing the Internet and provide the **Next Hop** IP address, which is the IP address of the gateway that is connecting to the Internet.

- d) Click the check mark icon to save the device information.
- e) Repeat this step for any additional IPN devices that you want to add.

Step 7 Provide the **External Devices** information.

If you do not have any Cloud Network Controller sites, this tab will not be present in the UI.

If you do not have any Cloud Network Controller sites in your Multi-Site domain or you do not plan to configure connectivity between cloud sites and branch routers or other external devices, you can skip this step.

The following steps describe how to provide information about any branch routers or external devices to which you want to configure connectivity from your cloud sites.

- a) Select the **External Devices** tab.

This tab will only be available if you have at least one cloud site in your Multi-Site domain.

- b) Click **Add External Device**.

The **Add External Device** dialogue opens.

- c) Provide the **Name**, **IP Address**, and **BGP Autonomous System Number** for the device.

The IP address that you provide will be used as the tunnel peer address from the Cloud Network Controller's CSRs, not the device's management IP address. The connectivity will be established over public Internet using IPsec.

- d) Click the check mark icon to save the device information.
- e) Repeat this step for any additional IPN devices that you want to add.

After you have added all the external devices, ensure to complete the next step to provide the IPsec tunnel subnet pools from which the internal IP addresses will be allocated for these tunnels.

Step 8 Provide the **IPsec Tunnel Subnet Pools** information.

If you do not have any Cloud Network Controller sites, this tab will not be present in the UI.

There are two types of subnet pools that you can provide here:

- **External Subnet Pool**—Used for connectivity between cloud site CSRs and other sites (cloud or on-premises).

These are large global subnet pools that are managed by Cisco Nexus Dashboard Orchestrator. The Orchestrator creates smaller subnets from these pools and allocates them to sites to be used for intersite IPsec tunnels and external connectivity IPsec tunnels.

You must provide at least one external subnet pool if you want to enable external connectivity from one or more of your cloud sites.

- **Site-Specific Subnet Pool**—Used for connectivity between cloud site CSRs and external devices.

These subnets can be defined when the external connectivity IPsec tunnels must be in a specific range. For example, where a specific subnet is already being used to allocate IP addresses to the external router and you want to continue using those subnets for IPsec tunnels for NDO and cloud sites. These subnets are not managed by the Orchestrator and each subnet is assigned to a site in its entirety to be used locally for external connectivity IPsec tunnels.

If you do not provide any named subnet pools but still configure connectivity between cloud site's CSRs and external devices, the external subnet pool will be used for IP allocation.

Note

The minimum mask length for both subnet pools is /24.

To add one or more **External Subnet Pools**:

- Select the **IPsec Tunnel Subnet Pools** tab.
- In the **External Subnet Pool** area, click **+Add IP Address** to add one or more external subnet pools.

This subnet will be used to address the IPsec tunnel interfaces and loopbacks of the Cloud Routers that are used for on-premises connectivity, which you previously configured in the Cloud Network Controller for intersite connectivity in earlier Cisco Nexus Dashboard Orchestrator releases.

The subnets must not overlap with other on-premises TEP pools, should not begin with 0.x.x.x or 0.0.x.x, and should have a network mask between /16 and /24, for example 30.29.0.0/16.

- Click the check mark icon to save the subnet information.
- Repeat these substeps for any additional subnet pools that you want to add.

To add one or more **Site-Specific Subnet Pools**:

- Select the **IPsec Tunnel Subnet Pools** tab.
- In the **Site-Specific Subnet Pools** area, click **+Add IP Address** to add one or more external subnet pools.

The **Add Named Subnet Pool** dialogue opens.

- Provide the subnet **Name**.

You can use the subnet pool's name to choose the pool from which to allocate the IP addresses later on.

- Click **+Add IP Address** to add one or more subnet pools.

The subnets must have a network mask between /16 and /24 and not begin with 0.x.x.x or 0.0.x.x, for example 30.29.0.0/16.

- e) Click the check mark icon to save the subnet information.

Repeat the steps if you want to add multiple subnets to the same named subnet pool.

- f) Click **Save** to save the named subnet pool.
 - g) Repeat these substeps for any additional named subnet pools that you want to add.
-

What to do next

After you have configured general infra settings, you must still provide additional information for site-specific configurations based on the type of sites (ACI, Cloud Network Controller, or NDFC) you are managing. Follow the instructions described in the following sections to provide site-specific infra configurations.



CHAPTER 15

Configuring Infra for Cisco APIC Sites

- [Refreshing Site Connectivity Information, on page 177](#)
- [Configuring Infra: On-Premises Site Settings, on page 178](#)
- [Configuring Infra: Pod Settings, on page 181](#)
- [Configuring Infra: Spine Switches, on page 181](#)

Refreshing Site Connectivity Information

Any infrastructure changes, such as adding and removing spines or changing spine node IDs, require a multi-site fabric connectivity site Refresh. This section describes how to pull up-to-date connectivity information directly from each site's APIC.

Procedure

-
- | | |
|---------------|--|
| Step 1 | Log in to the Cisco Nexus Dashboard Orchestrator GUI. |
| Step 2 | In the left navigation menu, select Config > Site To Site Connectivity . |
| Step 3 | In the top right of the main pane, click Configure . |
| Step 4 | In the left pane, under Sites , select a specific site. |
| Step 5 | In the main window, click the Refresh button to pull fabric information from the APIC. |
| Step 6 | (Optional) For on-premises sites, in the Confirmation dialog, check the box if you want to remove configuration for decommissioned spine switch nodes.

If you choose to enable this check box, all configuration info for any currently decommissioned spine switches will be removed from the database. |
| Step 7 | Finally, click Yes to confirm and load the connectivity information.

This discovers any new or removed spines and all site-related fabric connectivity will be reimported from the APIC. |
-

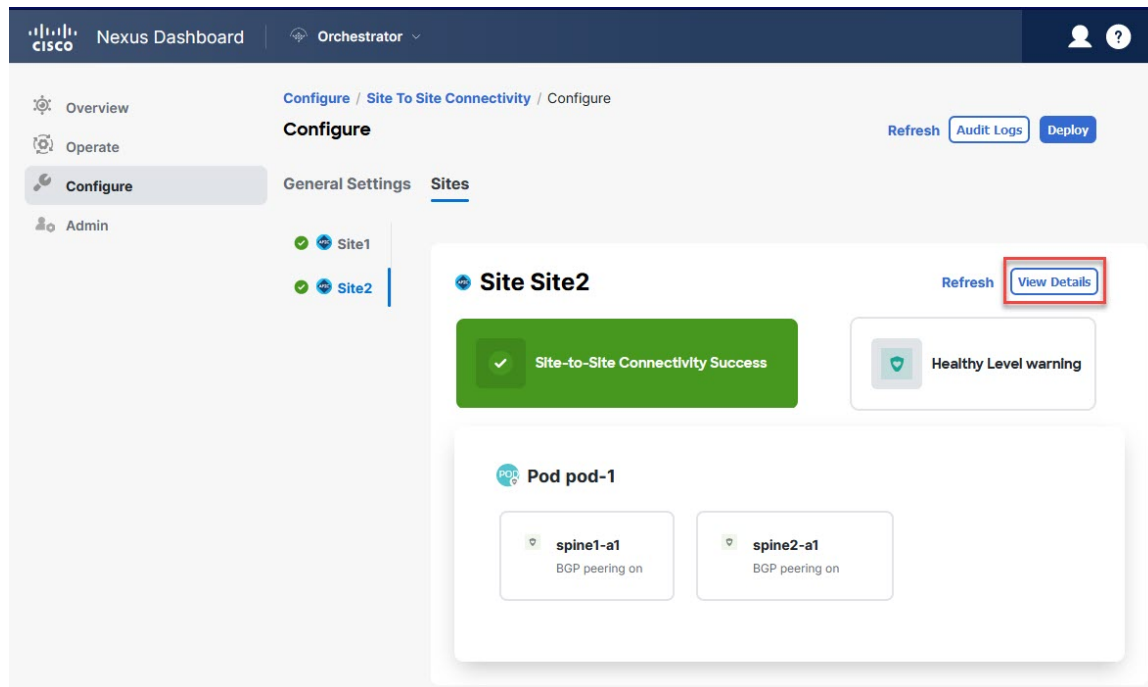
Configuring Infra: On-Premises Site Settings

This section describes how to configure site-specific Infra settings for on-premises sites.

Procedure

- Step 1** Log in to the Cisco Nexus Dashboard Orchestrator GUI.
- Step 2** In the left navigation menu, select **Configure > Site To Site Connectivity**.
- Step 3** In the top right of the main pane, click **Configure**.
- Step 4** In the left pane, under **Sites**, select a specific on-premises site.
- Step 5** Click **View Details** to load site settings.

Figure 17:



- Step 6** Provide the **Inter-Site Connectivity** information.
 - a) In the right **<Site> Settings** pane, enable the **Multi-Site** knob.
This defines whether the overlay connectivity is established between this site and other sites.
 - b) (Optional) Enable the **CloudSec Encryption** knob encryption for the site.
CloudSec Encryption provides intersite traffic encryption. The "Infrastructure Management" chapter in the [Cisco Multi-Site Configuration Guide](#) covers this feature in detail.
 - c) Specify the **Overlay Multicast TEP**.

This address is used for the intersite L2 BUM and L3 multicast traffic. This IP address is deployed on all spine switches that are part of the same fabric, regardless of whether it is a single pod or multipod fabric.

This address should not be taken from the address space of the original fabric's `Infra` TEP pool or from the `0.x.x.x` range.

- d) Specify the **BGP Autonomous System Number**.
- e) (Optional) Specify the **BGP Password**.
- f) Provide the **OSPF Area ID**.

The following settings are required if you are using OSPF protocol for underlay connectivity between the site and the IPN. If you plan to use BGP instead, you can skip this step. BGP underlay configuration is done at the port level, as described in [Configuring Infra: Spine Switches, on page 181](#).

- g) Select the **OSPF Area Type** from the drop-down list.

The following settings are required if you are using OSPF protocol for underlay connectivity between the site and the IPN. If you plan to use BGP instead, you can skip this step. BGP underlay configuration is done at the port level, as described in [Configuring Infra: Spine Switches, on page 181](#).

The OSPF area type can be one of the following:

- `nssa`
- `regular`

- h) Configure OSPF policies for the site.

The following settings are required if you are using OSPF protocol for underlay connectivity between the site and the IPN. If you plan to use BGP instead, you can skip this step. BGP underlay configuration is done at the port level, as described in [Configuring Infra: Spine Switches, on page 181](#).

You can either click an existing policy (for example, `msc-ospf-policy-default`) to modify it or click **+Add Policy** to add a new OSPF policy. Then in the **Add/Update Policy** window, specify the following:

- In the **Policy Name** field, enter the policy name.
- In the **Network Type** field, choose either `broadcast`, `point-to-point`, or `unspecified`.
The default is `broadcast`.
- In the **Priority** field, enter the priority number.
The default is `1`.
- In the **Cost of Interface** field, enter the cost of interface.
The default is `0`.
- From the **Interface Controls** drop-down list, choose one of the following:
 - **advertise-subnet**
 - **bfd**
 - **mtu-ignore**
 - **passive-participation**
- In the **Hello Interval (Seconds)** field, enter the hello interval in seconds.

The default is 10.

- In the **Dead Interval (Seconds)** field, enter the dead interval in seconds.

The default is 40.

- In the **Retransmit Interval (Seconds)** field, enter the retransmit interval in seconds.

The default is 5.

- In the **Transmit Delay (Seconds)** field, enter the transmit delay in seconds.

The default is 1.

- i) (Optional) From the **External Routed Domain** drop-down, select the domain that you want to use.

Choose an external router domain that you have created in the Cisco APIC GUI. For more information, see the *Cisco APIC Layer 3 Networking Configuration Guide* specific to your APIC release.

- j) (Optional) Enable **SDA Connectivity** for the site.

If the site is connected to an SDA network, enable the **SDA Connectivity** knob and provide the **External Routed Domain**, **VLAN Pool**, and **VRF Lite IP Pool Range** information.

If you enable SDA connectivity for the site, you need to configure extra settings as described in the SDA use case chapter of the [Cisco Multi-Site Configuration Guide for ACI Fabrics](#).

- k) (Optional) Enable **SR-MPLS Connectivity** for the site.

If the site is connected through an MPLS network, enable the **SR-MPLS Connectivity** knob and provide the Segment Routing global block (SRGB) range.

The Segment Routing Global Block (SRGB) is the range of label values that are reserved for Segment Routing (SR) in the Label Switching Database (LSD). These values are assigned as segment identifiers (SIDs) to SR-enabled nodes and have global significance throughout the domain.

The default range is 16000–23999.

If you enable MPLS connectivity for the site, you need to configure extra settings as described in the "Sites Connected through SR-MPLS" chapter of the [Cisco Multi-Site Configuration Guide for ACI Fabrics](#).

Step 7 Configure intersite connectivity between on-premises and cloud sites.

If you do not need to create intersite connectivity between on-premises and cloud sites, for example if your deployment contains only cloud or only on-premises sites, skip this step.

When you configure underlay connectivity between on-premises and cloud sites, you must provide an IPN device IP address to which the Cloud Network Controller's CSRs establish a tunnel and then configure the cloud site's infra settings.

- a) Click **+Add IPN Device** to specify an IPN device.
- b) From the drop-down, select one of the IPN devices you defined previously.

The IPN devices must be already defined in the **General Settings > IPN Devices** list, as described in [Configuring Infra: General Settings, on page 171](#)

- c) Configure intersite connectivity for cloud sites.

Any previously configured connectivity from the cloud sites to this on-premises site will be displayed here, but any additional configuration must be done from the cloud site's side as described in [Configuring Infra for Cisco Cloud Network Controller Sites, on page 185](#).

What to do next

While you have configured all the required intersite connectivity information, it has not been pushed to the sites yet. You must deploy the configuration as described in [Deploying Infra Configuration, on page 191](#)

Configuring Infra: Pod Settings

This section describes how to configure Pod-specific settings in each site.

Procedure

-
- | | |
|---------------|--|
| Step 1 | Log in to the Cisco Nexus Dashboard Orchestrator GUI. |
| Step 2 | In the left navigation menu, select Configure > Site To Site Connectivity . |
| Step 3 | In the top right of the main pane, click Configure . |
| Step 4 | In the left pane, under Sites , select a specific site. |
| Step 5 | In the main window, select a Pod. |
| Step 6 | In the right Pod Properties pane, add the Overlay Unicast TEP for the Pod.

This IP address is deployed on all spine switches that are part of the same Pod and used for sourcing and receiving VXLAN encapsulated traffic for Layer2 and Layer3 unicast communication. |
| Step 7 | Click +Add TEP Pool to add an external routable TEP pool.

The external routable TEP pools are used to assign a set of IP addresses that are routable across the IPN to APIC nodes, spine switches, and border leaf nodes. This is required to enable Multi-Site architecture.

External TEP pools previously assigned to the fabric on APIC are automatically inherited by NDO and displayed in the GUI when the fabric is added to the Multi-Site domain. |
| Step 8 | Repeat the procedure for every Pod in the site. |
-

Configuring Infra: Spine Switches

This section describes how to configure spine switches in each site for Cisco Multi-Site. When you configure the spine switches, you are effectively establishing the underlay connectivity between the sites in your Multi-Site domain by configuring connectivity between the spines in each site and the ISN.

Before Release 3.5(1), underlay connectivity was establishing using OSPF protocol. In this release however, you can choose to use OSPF, BGP (IPv4 only), or a mixture of protocols, with some sites using OSPF and some using BGP for intersite underlay connectivity. We recommend configuring either OSPF or BGP and

not both, however if you configure both protocols, BGP will take precedence and OSPF will not be installed in the route table.

Procedure

-
- Step 1** Log in to the Cisco Nexus Dashboard Orchestrator GUI.
- Step 2** In the left navigation menu, select **Config > Site To Site Connectivity**.
- Step 3** In the top right of the main pane, click **Configure**.
- Step 4** In the left pane, under **Sites**, select the specific on-premises site.
- Step 5** In the main pane, select a spine switch within a pod.
- Step 6** In the right **<Spine> Settings** pane, click **+Add Port**.
- Step 7** In the **Add Port** window, provide the underlay connectivity information.

Any port that is already configured directly in APIC for IPN connectivity will be imported and shown in the list. For any new ports you want to configure from NDO, use the following the steps:

a) Provide general information:

- In the **Ethernet Port ID** field, enter the port ID, for example 1/29.
This is the interface which will be used to connect to the IPN.
- In the **IP Address** field, enter the IP address/netmask.
The Orchestrator creates a subinterface with VLAN 4 with the specified IP ADDRESS under the specified PORT.
- In the **MTU** field, enter the MTU. You can specify either `inherit`, which would configure an MTU of 9150B, or choose a value between 576 and 9000.
MTU of the spine port should match MTU on IPN side.

Step 8 Choose the underlay protocol.

a) Enable **OSPF** if you want to use OSPF protocol for underlay connectivity.

If you want to use BGP protocol for underlay connectivity instead, skip this part and provide the information that is required in the next substep.

- Set **OSPF** to `Enabled`.
The OSPF settings become available.
- From the **OSPF Policy** drop-down, select the OSPF policy for the switch that you have configured in [Configuring Infra: On-Premises Site Settings, on page 178](#).
OSPF settings in the OSPF policy you choose should match on IPN side.
- For **OSPF Authentication**, you can pick either `none` or one of the following:
 - MD5
 - Simple
- Set **BGP** to `Disabled`.

- b) Enable **BGP** if you want to use BGP protocol for underlay connectivity.

If you're using OSPF protocol for underlay connectivity and have already configured it in the previous substep, skip this part.

Note

BGP IPv4 underlay is not supported in the following cases:

- If your Multi-Site domain contains one or more Cloud Network Controller sites, in which case you must use the OSPF protocol for intersite underlay connectivity for both On-Prem to On-Prem and On-Prem to cloud sites.
- If you are using GOLF (Layer 3 EVPN services for fabric WAN) for WAN connectivity in any of your fabrics.

In the above cases, you must use OSPF in the Infra L3Out deployed on the spines.

- Set **OSPF** to `Disabled`.

We recommend configuring either OSPF or BGP and not both, however if you configure both protocols, BGP will take precedence and OSPF routes will not be installed in the route table because only EBGp adjacencies with the ISN devices are supported.

- Set **BGP** to `Enabled`.

The BGP settings become available.

- In the **Peer IP** field, provide the IP address of this port's BGP neighbor.
Only IPv4 IP addresses are supported for BGP underlay connectivity.
- In the **Peer AS Number** field, provide the Autonomous System (AS) number of the BGP neighbor.
This release supports only EBGp adjacencies with the ISN devices.
- In the **BGP Password** field, provide the BGP peer password.
- Specify any additional options as required:
 - `Bidirectional Forwarding Detection`—Enables Bidirectional Forwarding Detection (BFD) protocol to detect faults on the physical link this port and the IPN device.
 - `Admin State`—Sets the admin state on the port to enabled.

Step 9 Repeat the procedure for every spine switch and port that connects to the IPN.



CHAPTER 16

Configuring Infra for Cisco Cloud Network Controller Sites

- [Refreshing Cloud Site Connectivity Information, on page 185](#)
- [Configuring Infra: Cloud Site Settings, on page 186](#)
- [Recovering from Cloud Network Controller Site Downtime , on page 188](#)

Refreshing Cloud Site Connectivity Information

Any infrastructure changes, such as CSR and Region addition or removal, require a multi-site fabric connectivity site Refresh. This section describes how to pull up-to-date connectivity information directly from each site's APIC.

Procedure

-
- | | |
|---------------|--|
| Step 1 | Log in to the Cisco Nexus Dashboard Orchestrator GUI. |
| Step 2 | In the left navigation menu, select Config > Site To Site Connectivity . |
| Step 3 | In the top right of the main pane, click Configure . |
| Step 4 | In the left pane, under Sites , select a specific site. |
| Step 5 | In the main window, click the Refresh button to discover any new or changed CSRs and regions. |
| Step 6 | Finally, click Yes to confirm and load the connectivity information.

This discovers any new or removed CSRs and regions. |
| Step 7 | Click Deploy to propagate the cloud site changes to other sites that have connectivity to it.

After you Refresh a cloud site's connectivity and CSRs or regions are added or removed, you must deploy infra configuration so other sites that have underlay connectivity to that cloud site get updated configuration. |
-

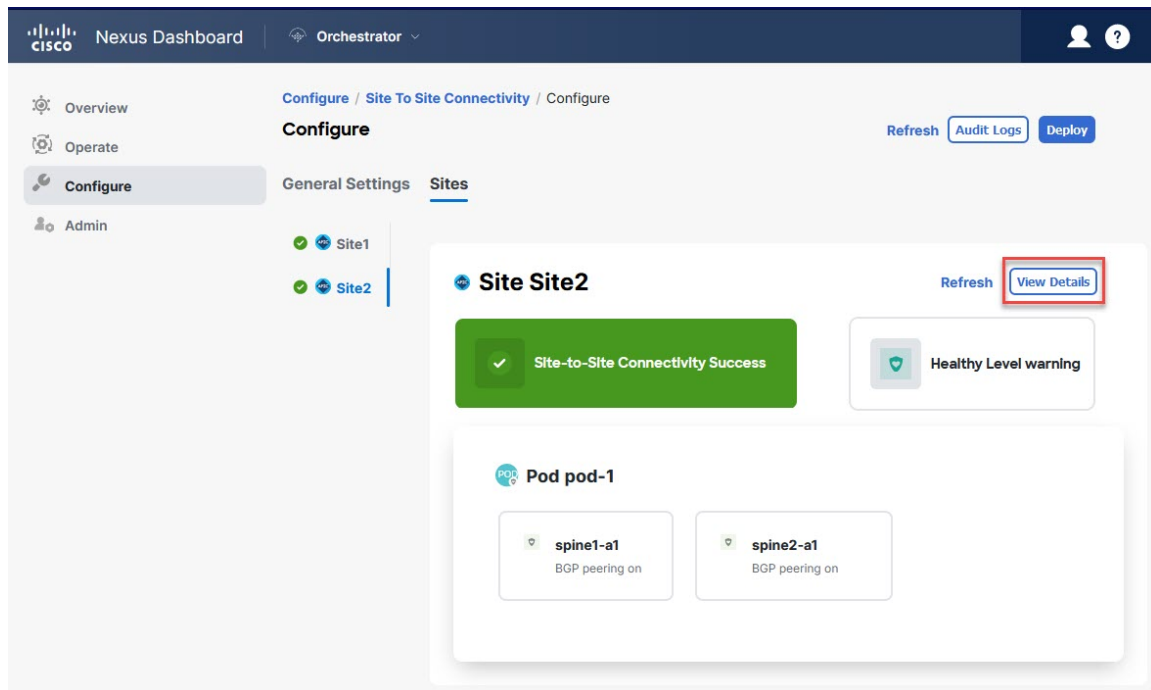
Configuring Infra: Cloud Site Settings

This section describes how to configure site-specific Infra settings for Cloud Network Controller sites.

Procedure

- Step 1** Log in to the Cisco Nexus Dashboard Orchestrator GUI.
- Step 2** In the left navigation menu, select **Config > Site To Site Connectivity**.
- Step 3** In the top right of the main pane, click **Configure**.
- Step 4** In the left pane, under **Sites**, select a specific cloud site.
- Step 5** Click **View Details** to load site settings.

Figure 18:



- Step 6** Provide the general **Inter-Site Connectivity** information.
 - a) In the right **<Site> Settings** pane, select the **Inter-Site Connectivity** tab.
 - b) Enable the **Multi-Site** knob.

This defines whether the overlay connectivity is established between this site and other sites.

The overlay configuration will not be pushed to sites which do not have the underlay intersite connectivity that is established as described in the next step.

- c) (Optional) Specify the **BGP Password**.

- Step 7** Provide site-specific **Inter-Site Connectivity** information.

- a) In the right properties sidebar for the cloud site, click **Add Site**.
The **Add Site** window opens.
- b) Under **Connected to Site**, click **Select a Site** and select the site (for example, `Site2`) to which you want to establish connectivity from the site you are configuring (for example, `Site1`).

When you select the remote site, the **Add Site** window updates to reflect both directions of connectivity: **Site1 > Site2** and **Site2 > Site1**.

- c) In the **Site1 > Site2** area, from the **Connection Type** drop-down, choose the type of connection between the sites.
The following options are available:

- **Public Internet**—Connectivity between the two sites is established through the Internet.
This type is supported between any two cloud sites or between a cloud site and an on-premises site.
- **Private Connection**—Connectivity is established using a private connection between the two sites.
This type is supported between a cloud site and an on-premises site.
- **Cloud Backbone**—Connectivity is established using cloud backbone.
This type is supported between two cloud sites of the same type, such as Azure-to-Azure or AWS-to-AWS.

If you have multiple types of sites (on-premises, AWS, and Azure), different pairs of site can use different connection type.

- d) Choose the **Protocol** that you want to use for connectivity between these two sites.

If using **BGP-EVPN** connectivity, you can optionally enable **IPSec** and choose which version of the Internet Key Exchange (IKE) protocol to use: IKEv1 (Version 1) or IKEv2 (Version 1) depending on your configuration.

- For **Public Internet** connectivity, IPsec is always enabled.
- For **Cloud Backbone** connectivity, IPsec is always disabled.
- For **Private Connection**, you can choose to enable or disable IPsec.

If using **BGP-IPv4** connectivity instead, you must provide an external VRF which will be used for route leaking configuration from the cloud site you are configuring.

After **Site1 > Site2** connectivity information is provided, the **Site2 > Site1** area will reflect the connectivity information in the opposite direction.

- e) Click **Save** to save the intersite connectivity configuration.

When you save connectivity information from `Site1` to `Site2`, the reverse connectivity is automatically created from `Site2` to `Site1`, which you can see by selecting the other site and checking the **Inter-site Connectivity** information in the right sidebar.

- f) Repeat this step to add intersite connectivity for other sites.

When you establish underlay connectivity from `Site1` to `Site2`, the reverse connectivity is done automatically for you.

However, if you also want to establish intersite connectivity from `Site1` to `Site3`, you must repeat this step for that site as well.

Step 8 Provide **External Connectivity** information.

If you do not plan to configure connectivity to external sites or devices that are not managed by NDO, you can skip this step.

Detailed description of an external connectivity use case is available in the [Configuring External Connectivity from Cloud CSRs Using Nexus Dashboard Orchestrator](#) document.

- a) In the right <Site> **Settings** pane, select the **External Connectivity** tab.
- b) Click **Add External Connection**.

The **Add External Connectivity** dialog opens.

- c) From the **VRF** drop-down, select the VRF you want to use for external connectivity.

This is the VRF which will be used to leak the cloud routes. The **Regions** section displays the cloud regions that contain the CSRs to which this configuration be applied.

- d) From the **Name** drop-down in the **External Devices** section, select the external device.

This is the external device that you added in the **General Settings > External Devices** list during general infra configuration and must already be defined as described in [Configuring Infra: General Settings, on page 171](#).

- e) From the **Tunnel IKE Version** drop-down, pick the IKE version that will be used to establish the IPsec tunnel between the cloud site's CSRs and the external device.
- f) (Optional) From the **Tunnel Subnet Pool** drop-down, choose one of the named subnet pools.

Named subnet pools are used to allocate IP addresses for IPsec tunnels between cloud site CSRs and external devices. If you do not provide any **named** subnet pools here, the **external** subnet pool will be used for IP allocation.

Providing a dedicated subnet pool for external device connectivity is useful for cases where a specific subnet is already being used to allocate IP addresses to the external router and you want to continue to use those subnets for IPsec tunnels for NDO and cloud sites.

If you want to provide a specific subnet pool for this connectivity, it must already be created as described in [Configuring Infra: General Settings, on page 171](#).

- g) (Optional) In the **Pre-Shared Key** field, provide the custom keys that you want to use to establish the tunnel.
- h) If necessary, repeat the previous substeps for any additional external devices you want to add for the same external connection (same VRF).
- i) If necessary, repeat this step for any additional external connections (different VRFs).

There's a one-to-one relationship for tunnel endpoints between CSRs and external devices, so while you can create extra external connectivity using different VRFs, you cannot create extra connectivity to the same external devices.

What to do next

While you have configured all the required intersite connectivity information, it has not been pushed to the sites yet. You must deploy the configuration as described in [Deploying Infra Configuration, on page 191](#)

Recovering from Cloud Network Controller Site Downtime

When Cloud Network Controller (formerly Cloud APIC) instance/VM goes down for any reason while still being managed by NDO, you may be unable to undeploy or delete any existing templates associated with that

cloud site. In this case, attempting to forcefully unmanage the site in NDO can cause stale configuration and deployment errors even if the site recovers.

To recover from this:

Procedure

Step 1 Bring up the new Cloud Network Controller sites and reregister the cloud sites.

- a) Log in to NDO.
- b) Open the admin console.
- c) Navigate to the **Operate > Sites** page.
- d) From the action (...) menu next to the site you redeployed, choose **Edit Site**.
- e) Check the "Reregister site" check box.
- f) Provide the new site details.

You must provide the new public IP address of site and sign-in credentials.

- g) Click **Save** to reregister the site.

When the connectivity status of the site shows **UP**, the site IPs in NDO are also updated and the new sites are in 'managed' state.

Step 2 Undeploy the previously deployed templates for each schema.

- a) Log in to NDO.
- b) Navigate to **Configure** and select **Tenant Template > Applications**.
- c) Click a schema with the deployed templates.
- d) From the **Actions** menu next to the **Template Properties**, choose **Undeploy Template** and wait until the template is successfully undeployed.

Step 3 Refresh the site's infra configuration to ensure that the new Cisco Catalyst 8000V switches are added in NDO.

- a) Navigate to **Configure** and select **SiteTo Site Connectivity**.
- b) Click **Configure** at the top right of the screen.
- c) Select the cloud site under the **Sites** panel and click **Refresh**.
- d) Click **Deploy** on the top right of the screen and wait until all sites are successfully deployed.

Step 4 Redeploy all templates associated with this Cloud Network Controller site.

- a) Navigate to **Configure > Tenant Templates** under the **Applications** tab.
 - b) Click a schema with the templates undeployed earlier.
 - c) Click **Deploy to Sites and** wait until the template is deployed.
-



CHAPTER 17

Deploying Infra Configuration for ACI Sites

- [Deploying Infra Configuration, on page 191](#)
- [Enabling Connectivity Between On-Premises and Cloud Sites, on page 192](#)

Deploying Infra Configuration

This section describes how to deploy the Infra configuration to each APIC site.

Procedure

Step 1

In the top right of the main pane, click **Deploy** and choose the appropriate option to deploy the configuration.

If you have configured only on-premises or only cloud sites, simply click **Deploy** to deploy the Infra configuration.

However, if you have both, on-premises and cloud site, the following additional options may be available:

- **Deploy & Download IPN Device Config files:** Pushes the configuration to both the on-premises APIC site and the Cloud Network Controller site and enables the end-to-end interconnect between the on-premises and the cloud sites.
In addition, this option downloads a zip file that contains configuration information that you will use to enable connectivity from the IPN devices to Cisco Cloud Services Router (CSR). A followup screen appears that allows you to select all or some of the configuration files to download.
- **Deploy & Download External Device Config files:** Pushes the configuration to both the Cloud Network Controller sites and enables the end-to-end interconnect between the cloud sites and external devices.
In addition, this option downloads a zip file that contains configuration information that you will use to enable connectivity from external devices to the Cisco Cloud Services Router (CSR) deployed in your cloud sites. A followup screen appears that allows you to select all or some of the configuration files to download.
- **Download IPN Device Config files only:** Downloads a zip file that contains configuration information that you will use to enable connectivity from the IPN devices to Cisco Cloud Services Router (CSR) without deploying the configuration.
- **Download External Device Config files only:** Downloads a zip file that contains configuration information that you will use to enable connectivity from external devices to Cisco Cloud Services Router (CSR) without deploying the configuration.

Step 2

In the confirmation window, click **Yes**.

The Deployment started, refer to left menu for individual site deployment status message will indicate that Infra configuration deployment began and you can verify each site's progress by the icon displayed next to the site's name in the left pane.

What to do next

The Infra overlay and underlay configuration settings are now deployed to all sites' controllers and cloud CSRs. The last remaining step is to configure your IPN devices with the tunnels for cloud CSRs as described in [Refreshing Site Connectivity Information, on page 177](#).

Enabling Connectivity Between On-Premises and Cloud Sites

If you have only on-premises or only cloud sites, you can skip this section.

This section describes how to enable connectivity between on-premises APIC sites and Cloud Network Controller sites.

By default, the Cisco Cloud Network Controller will deploy a pair of redundant Cisco Cloud Services Router 1000vs. The procedures in this section creates two tunnels, one IPsec tunnel from the on-premises IPsec device to each of these Cisco Cloud Services Router 1000vs. If you have multiple on-premises IPsec devices, you will need to configure the same tunnels to the CSRs on each of the on-premises devices.

The following information provides commands for Cisco Cloud Services Router 1000v as your on-premises IPsec termination device. Use similar commands if you are using a different device or platform.

Procedure

Step 1 Gather the necessary information that you will need to enable connectivity between the CSRs deployed in the cloud site and the on-premises IPsec termination device.

You can get the required configuration details using either the **Deploy & Download IPN Device config files** or the **Download IPN Device config files only** option in Nexus Dashboard Orchestrator as part of the procedures provided in [Deploying Infra Configuration, on page 191](#).

Step 2 Log into the on-premises IPsec device.

Step 3 Configure the tunnel for the *first* CSR.

Details for the first CSR are available in the configuration files for the ISN devices you downloaded from the Nexus Dashboard Orchestrator, but the following fields describe the important values for your specific deployment:

- *<first-csr-tunnel-ID>*—unique tunnel ID that you assign to this tunnel.
- *<first-csr-ip-address>*—public IP address of the third network interface of the first CSR.

The destination of the tunnel depends on the type of underlay connectivity:

- The destination of the tunnel is the public IP of the cloud router interface if the underlay is via public internet
- The destination of the tunnel is the private IP of the cloud router interface if the underlay is via private connectivity, such as DX on AWS or ER on Azure

- *<first-csr-preshared-key>*—preshared key of the first CSR.
- *<onprem-device-interface>*—interface that is used for connecting to the Cisco Cloud Services Router 1000v deployed in Amazon Web Services.
- *<onprem-device-ip-address>*—IP address for the *<interface>* interface that is used for connecting to the Cisco Cloud Services Router 1000v deployed in Amazon Web Services.
- *<peer-tunnel-for-onprem-IPsec-to-first-CSR>*—peer tunnel IP address for the on-premises IPsec device to the first cloud CSR.
- *<process-id>* —OSPF process ID.
- *<area-id>*—OSPF area ID.

The following example shows intersite connectivity configuration using the IKEv2 protocol supported starting with Nexus Dashboard Orchestrator, Release 3.3(1) and Cloud Network Controller, Release 5.2(1). If you are using IKEv1, the IPN configuration file you downloaded from NDO may look slightly differently, but the principle remains the same.

```
crypto ikev2 proposal ikev2-proposal-default
  encryption aes-cbc-256 aes-cbc-192 aes-cbc-128
  integrity sha512 sha384 sha256 sha1
  group 24 21 20 19 16 15 14 2
exit

crypto ikev2 policy ikev2-policy-default
  proposal ikev2-proposal-default
exit

crypto ikev2 keyring key-ikev2-infra:overlay-1-<first-csr-tunnel-id>
  peer peer-ikev2-keyring
    address <first-csr-ip-address>
    pre-shared-key <first-csr-preshared-key>
  exit
exit

crypto ikev2 profile ikev2-infra:overlay-1-<first-csr-tunnel-id>
  match address local interface <onprem-device-interface>
  match identity remote address <first-csr-ip-address> 255.255.255.255
  identity local address <onprem-device-ip-address>
  authentication remote pre-share
  authentication local pre-share
  keyring local key-ikev2-infra:overlay-1-<first-csr-tunnel-id>
  lifetime 3600
  dpd 10 5 on-demand
exit

crypto ipsec transform-set infra:overlay-1-<first-csr-tunnel-id> esp-gcm 256
  mode tunnel
exit

crypto ipsec profile infra:overlay-1-<first-csr-tunnel-id>
  set pfs group14
  set ikev2-profile ikev2-infra:overlay-1-<first-csr-tunnel-id>
  set transform-set infra:overlay-1-<first-csr-tunnel-id>
exit

interface tunnel 2001
  ip address <peer-tunnel-for-onprem-IPsec-to-first-CSR> 255.255.255.252
  ip virtual-reassembly
  tunnel source <onprem-device-interface>
  tunnel destination <first-csr-ip-address>
```

```

tunnel mode ipsec ipv4
tunnel protection ipsec profile infra:overlay-1-<first-csr-tunnel-id>
ip mtu 1400
ip tcp adjust-mss 1400
ip ospf <process-id> area <area-id>
no shut
exit

```

Example:

```

crypto ikev2 proposal ikev2-proposal-default
  encryption aes-cbc-256 aes-cbc-192 aes-cbc-128
  integrity sha512 sha384 sha256 sha1
  group 24 21 20 19 16 15 14 2
exit

```

```

crypto ikev2 policy ikev2-policy-default
  proposal ikev2-proposal-default
exit

```

```

crypto ikev2 keyring key-ikev2-infra:overlay-1-2001
  peer peer-ikev2-keyring
    address 52.12.232.0
    pre-shared-key 1449047253219022866513892194096727146110
  exit
exit

```

```

crypto ikev2 profile ikev2-infra:overlay-1-2001
  ! Please change GigabitEthernet1 to the appropriate interface
  match address local interface GigabitEthernet1
  match identity remote address 52.12.232.0 255.255.255.255
  identity local address 128.107.72.62
  authentication remote pre-share
  authentication local pre-share
  keyring local key-ikev2-infra:overlay-1-2001
  lifetime 3600
  dpd 10 5 on-demand
exit

```

```

crypto ipsec transform-set infra:overlay-1-2001 esp-gcm 256
  mode tunnel
exit

```

```

crypto ipsec profile infra:overlay-1-2001
  set pfs group14
  set ikev2-profile ikev2-infra:overlay-1-2001
  set transform-set infra:overlay-1-2001
exit

```

! These tunnel interfaces establish point-to-point connectivity between the on-prem device and the cloud Routers

! The destination of the tunnel depends on the type of underlay connectivity:

! 1) The destination of the tunnel is the public IP of the cloud Router interface if the underlay is via internet

! 2) The destination of the tunnel is the private IP of the cloud Router interface if the underlay is via private connectivity like DX on AWS or ER on Azure

```

interface tunnel 2001
  ip address 5.5.1.26 255.255.255.252
  ip virtual-reassembly
  ! Please change GigabitEthernet1 to the appropriate interface
  tunnel source GigabitEthernet1
  tunnel destination 52.12.232.0
  tunnel mode ipsec ipv4

```

```
tunnel protection ipsec profile infra:overlay-1-2001
ip mtu 1400
ip tcp adjust-mss 1400
! Please update process ID according with your configuration
ip ospf 1 area 0.0.0.1
no shut
exit
```

Step 4 Repeat the previous step for the 2nd and any additional CSRs that you need to configure.

Step 5 Verify that the tunnels are up on your on-premises IPsec device.

Use the following command to display the status. If you do not see that both tunnels are shown as up, verify the information that you entered in the steps in this section to determine where you might have an issue. Do not proceed to the next section until you see that both tunnels are shown as up.

```
ISN_CSR# show ip interface brief | include Tunnel
```

Interface	IP-Address	OK?	Method	Status	Protocol
Tunnel1000	30.29.1.2	YES	manual	up	up
Tunnel1001	30.29.1.4	YES	manual	up	up



CHAPTER 18

CloudSec Encryption

- [Cisco ACI CloudSec Encryption, on page 197](#)
- [Requirements and Guidelines, on page 198](#)
- [CloudSec Encryption Terminology, on page 201](#)
- [CloudSec Encryption and Decryption Handling, on page 202](#)
- [CloudSec Encryption Key Allocation and Distribution, on page 204](#)
- [Configuring Cisco APIC for CloudSec Encryption, on page 206](#)
- [Enabling CloudSec Encryption in Cisco Nexus Dashboard Orchestrator, on page 209](#)
- [Verifying CloudSec Configuration on Switches, on page 210](#)
- [Rekey Process During Spine Switch Maintenance, on page 212](#)

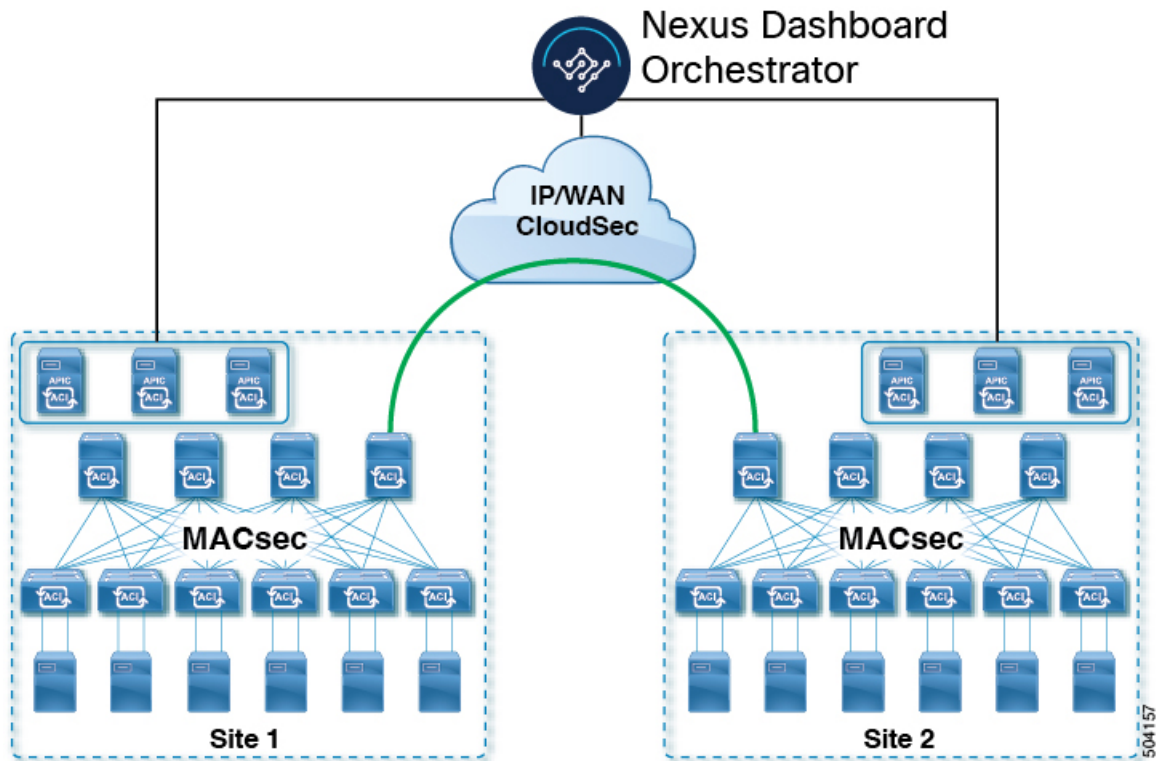
Cisco ACI CloudSec Encryption

As most Cisco ACI deployments are adopting the Multi-Site architecture to address disaster recovery and scale, the current security implementation using MACsec encryption within local site is becoming insufficient to guarantee data security and integrity across multiple sites connected by insecure external IP networks interconnecting separate fabrics. Nexus Dashboard Orchestrator Release 2.0(1) introduces the CloudSec Encryption feature designed to provide inter-site encryption of traffic.

Multi-Site topology uses three tunnel end-point (TEP) IP addresses (Overlay Multicast TEP, Overlay Unicast TEP, and External TEP Pool) to provide connectivity between sites. These TEP addresses are configured by the admin on Nexus Dashboard Orchestrator and pushed down to each site's Cisco APIC, which in turn configures them on the spine switches. These three addresses are used to determine when traffic is destined for a remote site, in which case an encrypted CloudSec tunnel is created between the two spine switches that provide physical connectivity between the two sites through the Inter-Site Network (ISN).

The following figure illustrates the overall encryption approach that combines MACsec for local site traffic and CloudSec for inter-site traffic encryption.

Figure 19: CloudSec Encryption



Requirements and Guidelines

When configuring CloudSec encryption, the following guidelines apply:

- CloudSec has been validated using a Nexus 9000 Inter-Site Network (ISN) infrastructure. If your ISN infrastructure is made up of different devices, or the devices are unknown (such as in the case of circuits purchased from a service provider), it is required that an ASR1K router is the first hop device directly connected to the ACI spine (with a separate pair of ASR1K devices deployed in each site), or the Nexus 9000 ISN network. The ASR1K router with padding-fixup enabled allows the CloudSec traffic to traverse any IP network between the sites.

To configure an ASR1K router:

1. Log in to the device.
2. Configure the UDP ports.

**Note**

- If you're running Release 3.7(1) or later and configure CloudSec to use the IANA-assigned port 8017, specify that port in the following command instead.
- You must have either an Advanced Enterprise or an Advanced IP Services license to run the **platform cloudsec padding-fixup** command on the ASR1K router.

```
ASR1K(config)# platform cloudsec padding-fixup dst-udp-port 9999
```

3. Verify the configuration.

In the following output, ensure that the port you configured in the previous step (8017 or 9999) is shown.

```
ASR1K# show platform software ip rp active cloudsec
CloudSec Debug: disabled
CloudSec UDP destination port: enabled
1st UDP destination port: 9999
2nd UDP destination port: 0
3rd UDP destination port: 0
```

```
ASR1K# show platform software ip fp active cloudsec
CloudSec Debug: disabled
CloudSec UDP destination port: enabled
1st UDP destination port: 9999
2nd UDP destination port: 0
3rd UDP destination port: 0
```

- If one or more spine switches are down when you attempt to disable CloudSec encryption, the disable process will not complete on those switches until the switches are up. This may result in packet drops on the switches when they come back up.

We recommend you ensure that all spine switches in the fabric are up or completely decommissioned before enabling or disabling CloudSec encryption.

- Beginning with Nexus Dashboard Orchestrator, Release 3.7(1), CloudSec encryption can be configured to use the IANA-assigned port.

By default, CloudSec uses a proprietary UDP port. Orchestrator releases 3.7(1) or later can be configured to use the official IANA-reserved port 8017 for CloudSec encryption between sites instead.

**Note**

The IANA-reserved port is supported for Cisco APIC sites running release 5.2(4) or later.

To change this setting, CloudSec must be disabled on all sites. If you want to enable IANA reserved port, but already have CloudSec encryption enabled for one or more of your sites, disable CloudSec for all sites, enable **IANA Reserve UDP Port** option, then re-enable CloudSec for the required sites.

- The CloudSec Encryption feature is not supported with the following features:
 - Precision Time Protocol (PTP)

- Remote Leaf Direct
- Virtual Pod (vPOD)
- SDA
- Remote Leaf or Multi-Pod configurations
- Intersite L3Out, if the sites are running Cisco APIC releases prior to 5.2(4).

CloudSec is supported with intersite L3Out for APIC sites running release 5.2(4) or later.

Requirements

The CloudSec encryption capability requires the following:

- Cisco ACI spine-leaf architecture with a Cisco APIC cluster for each site
- Cisco Nexus Dashboard Orchestrator to manage each site
- One **Advantage** or **Premier** license per each device (leaf only) in the fabric
- An add-on license **ACI-SEC-XF** per device for encryption if the device is a fixed spine
- An add-on license **ACI-SEC-XM** per device for encryption if the device is a modular spine

The following table provides the hardware platforms and the port ranges that are capable of CloudSec encryption.

Hardware Platform	Port Range
N9K-C9364C spine switches	Ports 49-64
N9K-C9332C spine switches	Ports 25-32
N9K-X9736C-FX line cards	Ports 29-36

If CloudSec is enabled for a site, but the encryption is not supported by the ports, a fault is raised with `unsupported-interface` error message.

CloudSec encryption's packet encapsulation is supported if Cisco QSFP-to-SFP Adapters (QSA), such as CVR-QSFP-SFP10G, is used with a supported optic. The full list of supported optics is available from the following link: <https://www.cisco.com/c/en/us/support/interfaces-modules/transceiver-modules/products-device-support-tables-list.html>.

Using IANA-Assigned Port and Orchestrator Downgrades

If you configure CloudSec encryption to use the IANA-assigned port as described in the following sections, there is a number of steps you have to take if you ever downgrade your Orchestrator service to a release prior to Release 3.7(1).

Before you downgrade your Nexus Dashboard Orchestrator to a release where IANA port is not supported:

1. Disable CloudSec encryption for all managed sites.
2. Disable the **IANA Reserved UDP Port** option in infra configuration settings.
3. Re-enable CloudSec encryption for all site where it was previously enabled.

4. Downgrade the Orchestrator services as you typically would.

CloudSec Encryption Terminology

CloudSec Encryption feature provides a secure upstream symmetric key allocation and distribution method for initial key and rekey requirements between sites. The following terminology is used in this chapter:

- **Upstream device** – The device that adds the CloudSec Encryption header and does the encryption of the VXLAN packet payload on transmission to a remote site using a locally generated symmetric cryptography key.
- **Downstream device** – The device that interprets the CloudSec Encryption header and does the decryption of the VXLAN packet payload on reception using the cryptography key generated by the remote site.
- **Upstream site** – The data center fabric that originates the encrypted VXLAN packets.
- **Downstream site** – The data center fabric that receives the encrypted packets and decrypts them.
- **TX Key** – The cryptography key used to encrypt the clear VXLAN packet payload. In ACI only one TX key can be active for all the remote sites.
- **RX Key** – The cryptography key used to decrypt the encrypted VXLAN packet payload. In ACI two RX keys can be active per remote site.

Two RX keys can be active at the same time because during the rekey process, the downstream sites will keep the old and the new RX keys after the new key deployment is finished for some duration to ensure that out of order packet deliveries with either key can be properly decrypted.

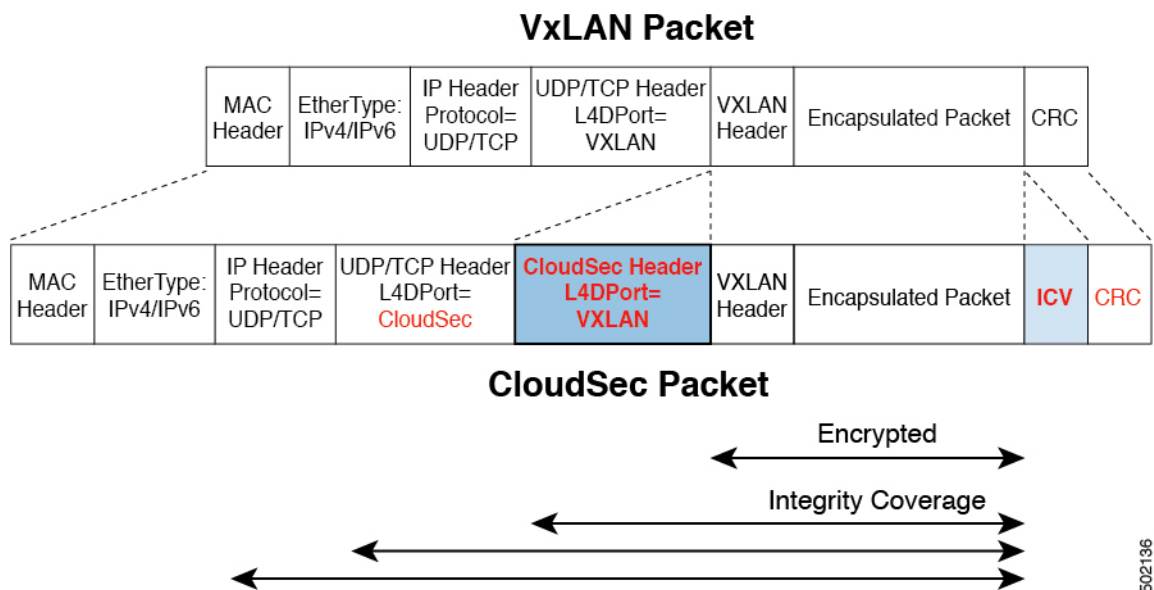
- **Symmetric Keys** – When the same cryptography key is used to encrypt (TX Key) and decrypt (RX Key) a packet stream by the upstream and downstream devices respectively.
 - **Rekey** – The process initiated by the upstream site to replace its old key with a newer key for all downstream sites after the old key expires.
 - **Secure Channel Identifier (SCI)** – A 64-bit identifier that represents a security association between the sites. It is transmitted in encrypted packet in CloudSec header and is used to derive the RX key on the downstream device for packet decryption.
 - **Association Number (AN)** – A 2-bit number (0, 1, 2, 3) that is sent in the CloudSec header of the encrypted packet and is used to derive the key at the downstream device in conjunction with the SCI for decryption. This allows multiple keys to be active at the downstream device to handle out of order packet arrivals with different keys from the same upstream device following a rekey operation.
- In ACI only two association number values (0 and 1) are used for the two active RX keys and only one association number value (0 or 1) is used for the TX key at any point in time.
- **Pre-shared key (PSK)** – One or more keys must be configured in the Cisco APIC GUI to be used as a random seed for generating the CloudSec TX and RX keys. If multiple PSK are configured, each rekey process will use the next PSK in order of their indexes; if no higher index PSK is available, a PSK with the lowest index will be used. Each PSK must be a hexadecimal string 64 characters long. Cisco APIC supports up to 256 pre-shared keys.

CloudSec Encryption and Decryption Handling

In order to provide a fully integrated, simple, and cost-effective solution that addresses both, data security and integrity, starting with Release 2.0(1), Multi-Site provides a CloudSec Encryption feature that allows for complete source-to-destination packet encryption between Multi-Site fabrics.

The following figure shows packet diagram before and after CloudSec encapsulation, followed by descriptions of the encryption and decryption processes:

Figure 20: CloudSec Packet



Packet Encryption

The following is a high level overview of how CloudSec handles outgoing traffic packets:

- The iVXLAN packets are filtered using outer IP header destination address field and Layer 4 destination port information and filtered packets are marked for encryption.
- The offset to use for encryption is calculated according to the fields of the packet. For example, the offset may vary based on whether there is a 802.1q VLAN or if the packet is an IPv4 or IPv6 packet.
The offset is automatically determined and is not visible to the user.
- The encryption keys are programmed in the hardware tables and are looked up from the table using the packet IP header.

Once the packet is marked for encryption, the encryption key is loaded, and the offset from the beginning of the packet where to start the encryption is known, the following additional steps are taken:

- The UDP destination port number is copied from the UDP header into a CloudSec field for recovery when the packet is decrypted.
- The UDP destination port number is overwritten to indicate that it is a CloudSec packet.

In releases prior to 3.7(1), the port is overwritten with a Cisco proprietary Layer-4 port number 9999.

In release 3.7(1) or later where you can configure CloudSec to use the IANA-assigned port 8017, the destination port number used is either 9999 or 8017 depending on whether you enabled this option.

- The UDP length field is updated to reflect the additional bytes that are being added.
- The CloudSec header is inserted directly after the UDP header.
- The Integrity Check Value (ICV) is inserted at the end of the packet, between the payload and the CRC.
- The ICV requires construction of a 128-bit initialization vector. For CloudSec, any use of the source MAC address for ICV purposes is replaced by a programmable value per SCI.
- CRC is updated to reflect the change in the contents of the packet.

Packet Decryption

The way CloudSec handles incoming packets is symmetric to the outgoing packets algorithm described above:

- If the received packet is a CloudSec packet, it is decrypted and the ICV is verified.

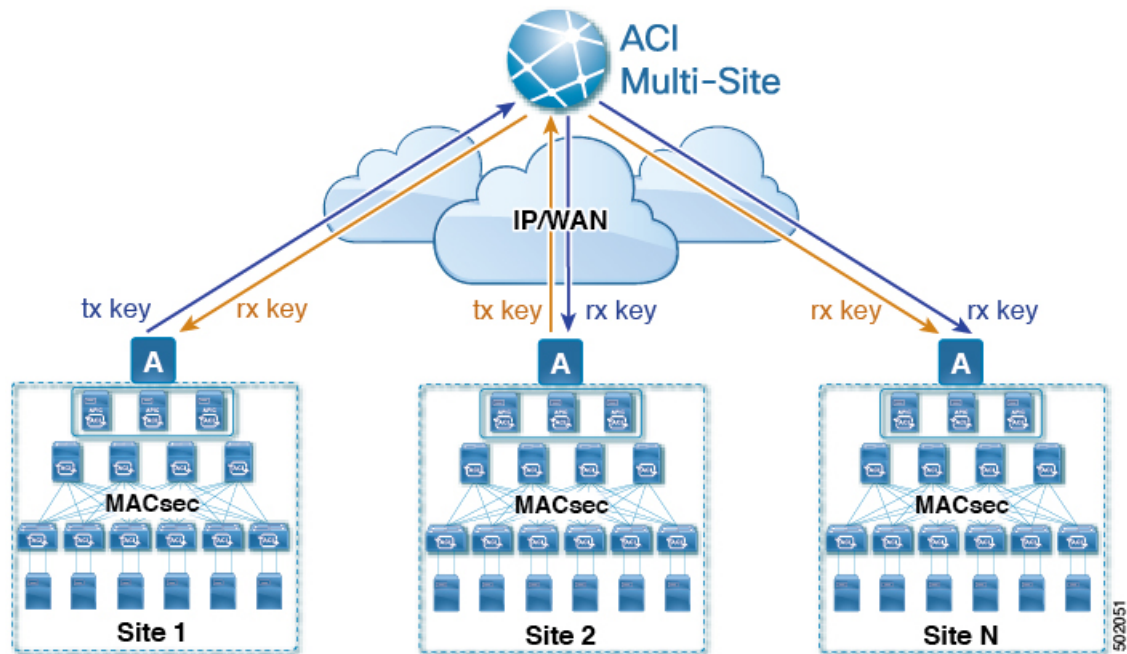
If ICV verification passed, the extra fields are removed, the UDP destination port number is moved from the CloudSec header to the UDP header, the CRC is updated, and the packet is forwarded to destination after decryption and CloudSec header removal. Otherwise the packet is dropped.

- The decryption key is retrieved from the key store using the received CloudSec packet outer IP header source address field, CloudSec header SCI, and AN number fields.
- If the packet is not a CloudSec packet, the packet is left unchanged.

CloudSec Encryption Key Allocation and Distribution

Initial Key Configuration

Figure 21: CloudSec Key Distribution



The following is a high level overview of the CloudSec encryption key initial allocation and distribution process illustrated by the figure above:

- The upstream site's Cisco APIC generates a local symmetric key intended to be used for data encryption of VXLAN packets transmitted from its site. The same key that is used by the upstream site for encryption is used for decryption of the packets on the downstream remote receiving sites.

Every site is an upstream site for the traffic it transmits to other sites. If multiple sites exist, each site generates its own site-to-site key and use that key for encryption before transmitting to the remote site.

- The generated symmetric key is pushed to the Nexus Dashboard Orchestrator (NDO) by the upstream site's Cisco APIC for distribution to downstream remote sites.
- The NDO acts as a message broker and collects the generated symmetric key from the upstream site's Cisco APIC, then distributes it to downstream remote sites' Cisco APICs.

The keys are encrypted using Key Encryption Key (KEK) and are distributed via a TLS-based channel.

- Each downstream site's Cisco APIC configures the received key as RX key on the local spine switches which are intended to receive the traffic from the upstream site that generated the key.
- Each downstream site's Cisco APIC also collects the deployment status of the RX Key from the local spine switches and then pushes it to the NDO.
- The NDO relays the key deployment status from all downstream remote sites back to the upstream site's Cisco APIC.

- The upstream site's Cisco APIC checks if the key deployment status received from all downstream remote sites is successful.
 - If the deployment status received from a downstream device is successful, the upstream site deploys the local symmetric key as its TX key on the spine switches to enable encryption of the VXLAN packets that are sent to the downstream site.
 - If the deployment status received from a downstream device is failed, a fault is raised on the Cisco APIC site where it failed and it is handled based on the "secure mode" setting configured on the NDO. In "must secure" mode the packets are dropped and in the "should secure" mode the packets are sent clear (unencrypted) to the destination site.



Note In current release, the mode is always set to "should secure" and cannot be changed.

Rekey Process

Each generated TX/RX key expires after a set amount of time, by default key expiry time is set to 15 minutes. When the initial set of TX/RX keys expires, a rekey process takes place.

The same general key allocation and distribution flow applies for the rekey process. The rekey process follows the "make before break" rule, in other words all the RX keys on the downstream sites are deployed before the new TX key is deployed on the upstream site. To achieve that, the upstream site will wait for the new RX key deployment status from the downstream sites before it configures the new TX key on the local upstream site's devices.

If any downstream site reports a failure status in deploying the new RX key, the rekey process will be terminated and the old key will remain active. The downstream sites will also keep the old and the new RX keys after the new key deployment is finished for some duration to ensure that out of order packet deliveries with either key can be properly decrypted.



Note Special precautions must be taken in regards to rekey process during spine switch maintenance, see [Rekey Process During Spine Switch Maintenance, on page 212](#) for details.

Rekey Process Failure

In case of any downstream site failing to deploy the new encryption key generated by the rekey process, the new key is discarded and the upstream device will continue to use the previous valid key as TX key. This approach keeps the upstream sites from having to maintain multiple TX keys per set of downstream sites. However, this approach may also result in the rekey process being delayed if the rekey deployment failures continue to occur with any one of the downstream sites. It is expected that the Multi-Site administrator will take action to fix the issue of the key deployment failure for the rekey to succeed.

Cisco APIC's Role in Key Management

The Cisco APIC is responsible for key allocation (both, initial key and rekey distribution), collection of the key deployment status messages from the spine switches, and notification of the Nexus Dashboard Orchestrator about each key's status for distribution to other sites.

Nexus Dashboard Orchestrator's Role in Key Management

The Nexus Dashboard Orchestrator is responsible for collecting the TX keys (both, initial key and subsequent rekeys) from the upstream site and distributing it to all downstream sites for deployment as RX keys. The NDO also collects the RX key deployment status information from the downstream sites and notifies the upstream site in order for it to update the TX key on successful RX key deployment status.

Upstream Model

In contrast to other technologies, such as MPLS, that use downstream key allocation, CloudSec's upstream model provides the following advantages:

- The model is simple and operationally easier to deploy in the networks.
- The model is preferred for Multi-Site use cases.
- It provides advantages for multicast traffic as it can use the same key and CloudSec header for each copy of the replicated packet transmitted to multiple destination sites. In downstream model each copy would have to use a different security key for each site during encryption.
- It provides easier troubleshooting in case of failures and better traceability of packets from the source to destination consistently for both, unicast and multicast replicated packets.

Configuring Cisco APIC for CloudSec Encryption

You must configure one or more Pre-Shared Keys (PSK) to be used by the Cisco APIC for generating the CloudSec encryption and decryption keys. The PSK are used as a random seed during the re-key process. If multiple PSK are configured, each re-key process will use the next PSK in order of their indexes; if no higher index PSK is available, a PSK with the lowest index will be used.

Because PSK is used as a seed for encryption key generation, configuring multiple PSK provides additional security by lowering the over-time vulnerability of the generated encryption keys.



Note If no pre-shared key is configured on the Cisco APIC, CloudSec will not be enabled for that site. In that case, turning on CloudSec setting in Multi-Site will raise a fault.

If at any time you wish to refresh a previously added PSK with a new one, simply repeat the procedure as if you were adding a new key, but specify an existing index.

You can configure one or more pre-shared keys in one of three ways:

- Using the Cisco APIC GUI, as described in [Configuring Cisco APIC for CloudSec Encryption Using GUI, on page 207](#)
- Using the Cisco APIC NX-OS Style CLI, as described in [Configuring Cisco APIC for CloudSec Encryption Using NX-OS Style CLI, on page 207](#)
- Using the Cisco APIC REST API, as described in [Configuring Cisco APIC for CloudSec Encryption Using REST API, on page 208](#)

Configuring Cisco APIC for CloudSec Encryption Using GUI

This section describes how to configure one or more pre-shared keys (PSK) using the Cisco APIC GUI.

Procedure

Step 1 Log in to APIC.

Step 2 Navigate to **Tenants > infra > Policies > CloudSec Encryption**

Step 3 Specify the **SA Key Expiry Time**.

This option specifies how long each key is valid (in minutes). Each generated TX/RX key expires after the specified amount of time triggering a re-key process. The expiration time can be between 5 and 1440 minutes.

Step 4 Click the + icon in the **Pre-Shared Keys** table.

Step 5 Specify the **Index** of the pre-shared key you are adding and then the **Pre-Shared Key** itself.

The **Index** field specifies the order in which the pre-shared keys are used. After the last (highest index) key is used, the process will continue with the first (lowest index) key. Cisco APIC supports up to 256 pre-shared keys, so the PSK index value must be between 1 and 256.

Each **Pre-Shared Key** must be a hexadecimal string 64 characters long.

Configuring Cisco APIC for CloudSec Encryption Using NX-OS Style CLI

This section describes how to configure one or more pre-shared keys (PSK) using the Cisco APIC NX-OS Style CLI.

Procedure

Step 1 Log in to the Cisco APIC NX-OS style CLI.

Step 2 Enter configuration mode.

Example:

```
apicl# configure
apicl (config)#
```

Step 3 Enter configuration mode for the default CloudSec profile.

Example:

```
apicl(config)# template cloudsec default
apicl(config-cloudsec)#
```

Step 4 Specify the Pre-Shared Keys (PSK) expiration time.

This option specifies how long each key is valid (in minutes). Each generated TX/RX key expires after the specified amount of time triggering a re-key process. The expiration time can be between 5 and 1440 minutes.

Example:

```
apicl(config-cloudsec)# sakexpirytime <duration>
```



```
</cloudsecIfPol>  
</fvTenant>
```

Enabling CloudSec Encryption in Cisco Nexus Dashboard Orchestrator

The CloudSec encryption can be enabled or disabled for each site individually. However, the communications between two sites are encrypted only if the feature is enabled on both sites.

Before you begin

Before you enable the CloudSec encryption between two or more sites, you must have completed the following tasks:

- Installed and configured the Cisco APIC clusters in multiple sites, as described in *Cisco APIC Installation, Upgrade, and Downgrade Guide*
- Installed and configured Cisco Nexus Dashboard Orchestrator, as described in *Cisco Nexus Dashboard Orchestrator Installation and Upgrade Guide*.
- Added each Cisco APIC site to the Cisco Nexus Dashboard Orchestrator, as described in *Cisco Multi-Site Configuration Guide*.

Procedure

- Step 1** Log in to the Cisco Nexus Dashboard Orchestrator.
- Step 2** From the left navigation menu, select the **Config > Site To Site Connectivity**.
- Step 3** Click the **Configure** button in the top right of the main pane.
- Step 4** (Optional) In the **General Settings** page's **Control Plane Configuration** tab, enable the **IANA Reserved UDP Port** option.

By default, CloudSec uses a proprietary UDP port. This option allows you to configure CloudSec to use the official IANA-reserved port 8017 for CloudSec encryption between sites.

Note

The IANA-reserved port is supported for Cisco APIC sites running release 5.2(4) or later.

To change this setting, CloudSec must be disabled on all sites. If you want to enable IANA reserved port, but already have CloudSec encryption that is enabled for one or more of your sites, disable CloudSec for all sites, enable **IANA Reserve UDP Port** option, then re-enable CloudSec for the required sites.

- Step 5** From the left sidebar, select the site for which you want to change the CloudSec configuration.
- Step 6** In the right sidebar, toggle the **CloudSec Encryption** setting to enable or disable the CloudSec encryption feature for the site.
-

Verifying CloudSec Configuration on Switches

The following command allows you to see the current CloudSec configuration that was deployed to the spine switch after you enable CloudSec encryption from your Nexus Dashboard Orchestrator.

Procedure

Step 1 Log in to your spine switch.

Step 2 Run the `show cloudsec sa interface all` command to show CloudSec configuration.

In the following output, ensure that for each Interface:

- The `Operational Status` value shows `UP`.
- The `Control` value is the same across all interfaces in all CloudSec-enabled sites as it indicates the UDP port currently in use for CloudSec encryption.

The following example shows the default Cisco-proprietary UDP port (`deprecatedUdpPort`). If you configure CloudSec to use the IANA-assigned port 8017, the **Control** field will display `ianaUdpPort` instead.

```
spinel# show cloudsec sa interface all
=====
Interface: Eth1/49.49(0x1a030031) Physical Interface: Eth1/49(0x1a030000)
  Operational Status: UP Retry: Off Control: deprecatedUdpPort
-----
Site-Id: 2 Peer: 200.200.204.0/24 Type: ext-routable-tep-pool Operational Status: UP
Pod-Id: 1
-----
TX Key: ***** Assoc Num: 1 Sci: 0x10002
Cipher Suite: gcm-aes-xpn-256 Rekey Num: 6826 Oper Rekey Num: 6826
Hardware Index: 0 Operational Status: UP Control: NONE
Last Updated: PST 2022-01-11 23:26:37.520-08:00 Retry: Off
Uptime: 11 hours 30 mins 45 secs
-----
Site-Id: 2 Peer: 200.200.202.1/32 Type: msite-unicast-tep Operational Status: UP
-----
TX Key: ***** Assoc Num: 1 Sci: 0x10002
Cipher Suite: gcm-aes-xpn-256 Rekey Num: 6826 Oper Rekey Num: 6826
Hardware Index: 2 Operational Status: UP Control: NONE
Last Updated: PST 2022-01-11 23:26:37.563-08:00 Retry: Off
Uptime: 11 hours 30 mins 45 secs
RX Key: ***** Assoc Num: 1 Sci: 0x20001
Cipher Suite: gcm-aes-xpn-256 Rekey Num: 6826 Oper Rekey Num: 6826
Hardware Index: 3 Operational Status: UP Control: NONE
Last Updated: PST 2022-01-11 23:26:37.442-08:00 Retry: Off
Uptime: 11 hours 30 mins 45 secs
RX Key: ***** Assoc Num: 0 Sci: 0x20001
Cipher Suite: gcm-aes-xpn-256 Rekey Num: 6827 Oper Rekey Num: 6827
Hardware Index: 2 Operational Status: UP Control: NONE
Last Updated: PST 2022-01-11 23:26:37.453-08:00 Retry: Off
Uptime: 11 hours 30 mins 45 secs
-----
Site-Id: 2 Peer: 200.200.201.1/32 Type: msite-multicast-tep Operational Status: UP
-----
TX Key: ***** Assoc Num: 1 Sci: 0x10002
Cipher Suite: gcm-aes-xpn-256 Rekey Num: 6826 Oper Rekey Num: 6826
```

```

Hardware Index: 1 Operational Status: UP Control: NONE
Last Updated: PST 2022-01-11 23:26:37.549-08:00 Retry: Off
Uptime: 11 hours 30 mins 45 secs
RX Key: ***** Assoc Num: 1 Sci: 0x20001
Cipher Suite: gcm-aes-xpn-256 Rekey Num: 6826 Oper Rekey Num: 6826
Hardware Index: 1 Operational Status: UP Control: NONE
Last Updated: PST 2022-01-11 23:26:36.501-08:00 Retry: Off
Uptime: 11 hours 30 mins 46 secs
RX Key: ***** Assoc Num: 0 Sci: 0x20001
Cipher Suite: gcm-aes-xpn-256 Rekey Num: 6827 Oper Rekey Num: 6827
Hardware Index: 0 Operational Status: UP Control: NONE
Last Updated: PST 2022-01-11 23:26:37.495-08:00 Retry: Off
Uptime: 11 hours 30 mins 45 secs

=====
Interface: Eth1/50.50(0x1a031032) Physical Interface: Eth1/50(0x1a031000)
Operational Status: UP Retry: Off Control: deprecatedUdpPort
-----
Site-Id: 2 Peer: 200.200.204.0/24 Type: ext-routable-tep-pool Operational Status: UP
Pod-Id: 1
-----
TX Key: ***** Assoc Num: 1 Sci: 0x10002
Cipher Suite: gcm-aes-xpn-256 Rekey Num: 6826 Oper Rekey Num: 6826
Hardware Index: 1 Operational Status: UP Control: NONE
Last Updated: PST 2022-01-11 23:26:37.577-08:00 Retry: Off
Uptime: 11 hours 30 mins 45 secs
-----
Site-Id: 2 Peer: 200.200.201.1/32 Type: msite-multicast-tep Operational Status: UP
-----
TX Key: ***** Assoc Num: 1 Sci: 0x10002
Cipher Suite: gcm-aes-xpn-256 Rekey Num: 6826 Oper Rekey Num: 6826
Hardware Index: 0 Operational Status: UP Control: NONE
Last Updated: PST 2022-01-11 23:26:37.537-08:00 Retry: Off
Uptime: 11 hours 30 mins 45 secs
RX Key: ***** Assoc Num: 1 Sci: 0x20001
Cipher Suite: gcm-aes-xpn-256 Rekey Num: 6826 Oper Rekey Num: 6826
Hardware Index: 1 Operational Status: UP Control: NONE
Last Updated: PST 2022-01-11 23:26:36.463-08:00 Retry: Off
Uptime: 11 hours 30 mins 46 secs
RX Key: ***** Assoc Num: 0 Sci: 0x20001
Cipher Suite: gcm-aes-xpn-256 Rekey Num: 6827 Oper Rekey Num: 6827
Hardware Index: 0 Operational Status: UP Control: NONE
Last Updated: PST 2022-01-11 23:26:37.416-08:00 Retry: Off
Uptime: 11 hours 30 mins 45 secs

-----
Site-Id: 2 Peer: 200.200.202.1/32 Type: msite-unicast-tep Operational Status: UP
-----
TX Key: ***** Assoc Num: 1 Sci: 0x10002
Cipher Suite: gcm-aes-xpn-256 Rekey Num: 6826 Oper Rekey Num: 6826
Hardware Index: 2 Operational Status: UP Control: NONE
Last Updated: PST 2022-01-11 23:26:37.593-08:00 Retry: Off
Uptime: 11 hours 30 mins 45 secs
RX Key: ***** Assoc Num: 0 Sci: 0x20001
Cipher Suite: gcm-aes-xpn-256 Rekey Num: 6827 Oper Rekey Num: 6827
Hardware Index: 2 Operational Status: UP Control: NONE
Last Updated: PST 2022-01-11 23:26:37.481-08:00 Retry: Off
Uptime: 11 hours 30 mins 45 secs
RX Key: ***** Assoc Num: 1 Sci: 0x20001
Cipher Suite: gcm-aes-xpn-256 Rekey Num: 6826 Oper Rekey Num: 6826
Hardware Index: 3 Operational Status: UP Control: NONE

```

Last Updated: PST 2022-01-11 23:26:37.507-08:00 Retry: Off
Uptime: 11 hours 30 mins 45 secs

Rekey Process During Spine Switch Maintenance

The following is a summary of the CloudSec rekey process during typical maintenance scenarios for the spine switches where the feature is enabled:

- **Normal Decommissioning** – CloudSec rekey process stops automatically whenever a CloudSec-enabled spine switch is decommissioned. Rekey process will not start again until the decommissioned node is commissioned back or the decommissioned node ID is removed from the Cisco APIC
- **Spine Switch Software Upgrade** – CloudSec rekey process stops automatically if a spine switch is reloaded due to software upgrade. Rekey process will resume after the spine switch comes out of reload.
- **Maintenance (GIR mode)** – CloudSec rekey process must be manually stopped using the instructions provided in [Disabling and Re-Enabling Re-Key Process Using NX-OS Style CLI, on page 212](#). Rekey can be enabled back only after the node is ready to forward traffic again.
- **Decommissioning and Removal from Cisco APIC** – CloudSec rekey process must be manually stopped using the instructions provided in [Disabling and Re-Enabling Re-Key Process Using NX-OS Style CLI, on page 212](#). Rekey can be enabled back only after the node is removed from Cisco APIC.

Disabling and Re-Enabling Re-Key Process Using NX-OS Style CLI

It is possible to manually stop and restart the re-key process. You may be required to manually control the re-key process in certain situations, such as switch decommissioning and maintenance. This section describes how to toggle the setting using Cisco APIC NX-OS Style CLI.

Procedure

Step 1 Log in to the Cisco APIC NX-OS style CLI.

Step 2 Enter configuration mode.

Example:

```
apic1# configure
apic1(config)#
```

Step 3 Enter configuration mode for the default CloudSec profile.

Example:

```
apic1(config)# template cloudsec default
apic1(config-cloudsec)#
```

Step 4 Stop or restart the re-key process.

To stop the re-key process:

Example:

```
apic1(config-cloudsec)# stoprekey yes
```

To restart the re-key process:

Example:

```
apic1(config-cloudsec)# stoprekey no
```

Disabling and Re-Enabling Re-Key Process Using REST API

It is possible to manually stop and restart the re-key process. You may be required to manually control the re-key process in certain situations, such as switch decommissioning and maintenance. This section describes how to toggle the setting using Cisco APIC REST API.

Procedure

Step 1 You can disable the rekey process using the following XML message.

Example:

```
<fvTenant annotation="" descr="" dn="uni/tn-infra" name="infra" nameAlias="" ownerKey="" ownerTag="">
  <cloudsecIfPol descr="cloudsecifp" name="default" sakExpiryTime="10" stopRekey= "true" status=""
/>
</fvTenant>
```

Step 2 You can enable the rekey process using the following XML message.

Example:

```
<fvTenant annotation="" descr="" dn="uni/tn-infra" name="infra" nameAlias="" ownerKey="" ownerTag="">
  <cloudsecIfPol descr="cloudsecifp" name="default" sakExpiryTime="10" stopRekey= "false" status=""
/>
</fvTenant>
```



PART IV

Features and Use Cases

- [DHCP Relay, on page 217](#)
- [EPG Preferred Group, on page 225](#)
- [External Connectivity \(L3Out\), on page 229](#)
- [Intersite L3Out, on page 259](#)
- [Intersite L3Out with PBR, on page 275](#)
- [Intersite Transit Routing with PBR, on page 283](#)
- [Layer 3 Multicast, on page 295](#)
- [QoS Preservation Across IPN, on page 305](#)
- [SD-Access and ACI Integration, on page 311](#)
- [SD-WAN Integration, on page 329](#)
- [Multi-Site and SR-MPLS L3Out Handoff, on page 337](#)
- [vzAny Contracts, on page 363](#)
- [vzAny with PBR, on page 377](#)



CHAPTER 19

DHCP Relay

- [DHCP Relay Policy, on page 217](#)
- [Guidelines and Limitations, on page 217](#)
- [Creating DHCP Relay Policies, on page 218](#)
- [Creating DHCP Option Policies, on page 220](#)
- [Assigning DHCP Policies, on page 221](#)
- [Creating DHCP Relay Contract, on page 222](#)
- [Verifying DHCP Relay Policies in APIC, on page 223](#)
- [Editing or Deleting Existing DHCP Policies, on page 224](#)

DHCP Relay Policy

Typically, when your DHCP server is located under an EPG, all the endpoints in that EPG have access to it and can obtain the IP addresses via DHCP. However, in many deployment scenarios, the DHCP server may not exist in the same EPG, BD, or VRF as all the clients that require it. In these cases a DHCP relay can be configured to allow endpoints in one EPG to obtain IP addresses via DHCP from a server that is located in another EPG/BD deployed in a different site or even connected externally to the fabric and reachable via an L3Out connection.

You can create the DHCP `Relay` policy in the Orchestrator GUI to configure the relay. Additionally, you can choose to create a DHCP `Option` policy to configure additional options you can use with the relay policy to provide specific configuration details. For all available DHCP options, refer to [RFC 2132](#).

When creating a DHCP relay policy, you specify an EPG (for example, `epg1`) or external EPG (for example, `ext-epg1`) where the DHCP server resides. After you create the DHCP policy, you associate it with a bridge domain, which in turn is associated with another EPG (for example, `epg2`) allowing the endpoints in that EPG to reach the DHCP server. Finally, you create a contract between the relay EPG (`epg1` or `ext-epg1`) and application EPG (`epg2`) to allow communication. The DHCP policies you create are pushed to the APIC when the bridge domain to which the policy is associated is deployed to a site.

Guidelines and Limitations

The DHCP relay policies are supported with the following caveats:

- DHCP relay policies are supported for fabrics running Cisco APIC Release 4.2(1) or later.
- The DHCP servers must support DHCP Relay Agent Information Option (Option 82).

When an ACI fabric acts as a DHCP relay, it inserts the DHCP Relay Agent Information Option in DHCP requests that it proxies on behalf of clients. If a response (DHCP offer) comes back from a DHCP server without Option 82, it is silently dropped by the fabric.

- DHCP relay policies are supported in user tenants or the `common` tenant only. DHCP policies are not supported for the `infra` or `mgmt` tenants.

When configuring shared resources and services in the ACI fabric, we recommend creating those resources in the `common` tenant, that way they can be used by any user tenant.

- DHCP relay server must be in the same user tenant as the DHCP clients or in the `common` tenant.

The server and the clients cannot be in different user tenants.

- DHCP relay policies can be configured for the primary SVI interface only.

If the bridge domain to which you assign a relay policy contains multiple subnets, the first subnet you add becomes the primary IP address on the SVI interface, while additional subnets are configured as secondary IP addresses. In certain scenarios, such as importing a configuration with a bridge domain with multiple subnets, the primary address on the SVI may change to one of the secondary addresses, which would break the DHCP relay for that bridge domain.

You can use the `show ip interface vrf all` command to verify IP address assignments for the SVI interfaces.

- If you make changes to the DHCP policy after you have assigned it to a bridge domain and deployed the bridge domain to one or more sites, you will need to re-deploy the bridge domain for the DHCP policy changes to be updated on each site's APIC.
- For inter-VRF DHCP relay with the DHCP server reachable via an L3Out, DHCP relay packets must use site-local L3Out to reach the DHCP server. Packets using an L3Out in a different site (Intersite L3Out) to reach the DHCP server is not supported.
- The following DHCP relay configurations are not supported:
 - DHCP relay label on L3Out interfaces
 - Importing existing DHCP policies from APIC.
 - DHCP relay policy configuration in Global Fabric Access Policies is not supported
 - Multiple DHCP servers within the same DHCP relay policy and EPG.

If you configure multiple providers under the same DHCP relay policy, they must be in different EPGs or external EPGs.

Creating DHCP Relay Policies

This section describes how to create a DHCP relay policy.



Note

If you make changes to the DHCP policy after you have assigned it to a bridge domain and deployed the bridge domain to one or more sites, you must redeploy the bridge domain for the DHCP policy changes to update on each site's APIC.

Before you begin

You must have the following:

- A DHCP server set up and configured in your environment.
- If the DHCP server is part of an application EPG, that EPG must be already created in the Cisco Nexus Dashboard Orchestrator.
- If the DHCP server is external to the fabric, the external EPG associated to the L3Out that is used to access the DHCP server must be already created.

Procedure

Step 1 Log in to your Cisco Nexus Dashboard and open the Cisco Nexus Dashboard Orchestrator service.

Step 2 Create a new Tenant Policy.

- a) From the left navigation pane, choose **Config > Tenant Policies**.
- b) On the **Tenant Templates > Tenant Policies** page, click **Add Tenant Policy Template**.
- c) In the Tenant Policies page's right properties sidebar, provide the **Name** for the tenant.
- d) From the **Select a Tenant** drop-down, choose the tenant with which you want to associate this template.

All the policies that you create in this template as described in the following steps will be associated with the selected tenant and deployed to it when you push the template to a specific site.

Step 3 Create a DHCP Relay Policy.

- a) From the **+Create Object** drop-down, select **DHCP Relay Policy**.
- b) In the right properties sidebar, provide the **Name** for the policy.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) Click **Add Provider** to configure the DHCP server to which you want to relay the DHCP requests originated by the endpoints.
- e) Select the provider type.

When adding a relay policy, you can choose one of the following two types:

- **Application EPG**—Specifies the application EPG that includes the DHCP server to which you want to relay the DHCP requests.
- **L3 External Network**—Specifies the External EPG associated to the L3Out that is used to access the network external to the fabric where the DHCP server is connected.

Note

You can select any EPG or external EPG that has been created in the Orchestrator and assigned to the tenant you specified, even if you have not yet deployed it to sites. If you select an EPG that hasn't been deployed, you can still complete the DHCP relay configuration, but you need to deploy the EPG before the relay is available for use.

- f) Click **Select an Application EPG** or **Select an External EPG** (based on the provider type you selected) and choose the provider EPG.
- g) In the **DHCP Server Address** field, provide the IP address of the DHCP server.
- h) Enable the **DHCP Server VRF Preference** option if necessary.

This feature was introduced in Cisco APIC release 5.2(4). For more information on the use cases where it is required see the [Cisco APIC Basic Configuration Guide](#).

- i) Click **OK** to save the provider information.
 - j) Repeat the previous substeps for any additional providers in the same DHCP Relay policy.
 - k) Repeat this step to create any additional DHCP Relay policies.
-

Creating DHCP Option Policies

This section describes how to create a DHCP option policy. DHCP options are appended to the end of the messages that DHCP servers and clients Exchange and can be used to provide extra configuration information to your DHCP server. Each DHCP option has a specific code that you must provide when adding the option policy. For a complete list of DHCP options and codes, see [RFC 2132](#).

Before you begin

You must have the following already configured:

- A DHCP server set up and configured in your environment.
- An EPG that contains the DHCP server that is already created in the Cisco Nexus Dashboard Orchestrator.
- A DHCP Relay policy created, as described in [Creating DHCP Relay Policies, on page 218](#).

Procedure

Step 1 Log in to your Cisco Nexus Dashboard and open the Cisco Nexus Dashboard Orchestrator service.

Step 2 Create a new or update an existing Tenant Policy.

- a) From the left navigation pane, choose **Configure > Tenant Templates > Tenant Policies**.
- b) On the **Tenant Policy Templates** page, select an existing policy or click **Add Tenant Policy Template**.
- c) If creating a new policy, in the Tenant Policies page's right properties sidebar, provide the **Name** for the tenant.
- d) If creating a new policy, from the **Select a Tenant** drop-down, choose the tenant with which you want to associate this template.

All the policies that you create int his template as described in the following steps will be associated with the selected tenant and deployed to it when you push the template to a specific site.

Step 3 Create a DHCP Option Policy.

- a) From the **+Create Object** drop-down, select **DHCP Option Policy**.
- b) In the right properties sidebar, provide the **Name** for the policy.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) Click **Add Option**.
- e) Provide option details.

For each DHCP option, provide the following:

- **Name** – While not technically required, we recommend using the same name for the option as listed in [RFC 2132](#).

For example, `Name Server`.

- **Id** – Provide the value if the option requires one.

For example, a list of name servers available to the client for the Name Server option.

- **Data** – Provide the value if the option requires one.

For example, a list of name servers available to the client for the Name Server option.

- f) Click **OK** to save.
- g) Repeat the previous substeps for any additional options in the same DHCP Option policy.
- h) Repeat this step to create any additional DHCP Option policies.

Assigning DHCP Policies

This section describes how to assign a DHCP policy to a bridge domain.



Note If you make changes to the DHCP policy after you have assigned it to a bridge domain and deployed the bridge domain to one or more sites, you must redeploy the bridge domain for the DHCP policy changes to be updated on each site's APIC.

Before you begin

You must have the following already configured:

- A DHCP relay policy, as described in [Creating DHCP Relay Policies, on page 218](#).
- (Optional) A DHCP option policy, as described in [Creating DHCP Option Policies, on page 220](#).
- The bridge domain to which you assign the DHCP policy, as described in the [Creating Schemas and Templates, on page 68](#) chapter.

Procedure

-
- Step 1** Log in to your Cisco Nexus Dashboard Orchestrator GUI.
 - Step 2** From the left navigation menu, select **Configure > Schemas**.
 - Step 3** Select the schema where the bridge domain is defined.
 - Step 4** Scroll down to the **Bridge Domain** area and select the bridge domain.
 - Step 5** In the right sidebar, scroll down and check the **DHCP Policy** option check box.
 - Step 6** From the **DHCP Relay Policy** drop-down, select the DHCP policy that you want to assign to this BD.
 - Step 7** (Optional) From the **DHCP Option Policy** drop-down, select the option policy.

A DHCP option policy provides extra options to be passed to the DHCP relay. For extra details, see [Creating DHCP Option Policies, on page 220](#).

Step 8 Assign the bridge domain to any EPG that needs access to the DHCP server through the relay.

Creating DHCP Relay Contract

DHCP packets are not filtered by contracts but contracts are required often to propagate routing information within the VRF and across VRFs. Although the DHCP packets are not filtered, it is recommended to configure contracts between the client EPG and the EPG configured as the provider in the DHCP relay policy.

This section describes how to create a contract between the EPG that contains the DHCP server and the EPG that contains endpoints that must use the relay. Although you have already created and assigned the DHCP policy to the bridge domain and the bridge domain to the clients' EPG, you must create and assign the contract to enable programming of routes to allow client to server communication.

Before you begin

You must have the following already configured:

- A DHCP relay policy, as described in [Creating DHCP Relay Policies, on page 218](#).
- (Optional) A DHCP option policy, as described in [Creating DHCP Option Policies, on page 220](#).
- The bridge domain to which you have assigned the DHCP policy, as described in [Assigning DHCP Policies, on page 221](#).

Procedure

Step 1 Log in to your Cisco Nexus Dashboard Orchestrator GUI.

Step 2 From the left navigation menu, select **Configure > Schemas**.

Step 3 Select the schema where you want to create the contract.

Step 4 Create a contract.

DHCP packets are not filtered by the contract so no specific filter is required, but a valid contract should be created and assigned to ensure proper BD and routes deployment.

- a) Scroll down to the **Contracts** area and click + to create a contract.
- b) In the right sidebar, provide the **Display Name** for the contract.
- c) From the **Scope** drop-down, select the appropriate scope.

Because the DHCP server EPG and application EPG must be in the same tenant, you can select one of the following:

- `vrf`, if both EPGs are in the same VRF.
- `tenant`, if the EPGs are in different VRFs.

- d) You can leave the **Apply Both Directions** knob on.

Step 5 Assign the contract to the DHCP relay EPG.

- a) Browse to the template where the EPG is located.
- b) Select the EPG or external EPG where the DHCP server resides.
This is the same EPG that you selected when creating the DHCP relay policy.

- c) In the right sidebar, click **+Contract**.
- d) Select the contract that you created and `provider` for its type.

Step 6 Assign the contract to the application EPG whose endpoints require DHCP relay access.

- a) Browse to the template where the application EPG is located.
- b) Select the application EPG.
- c) In the right sidebar, click **+Contract**.
- d) Select the contract that you created and `consumer` for its type.

Verifying DHCP Relay Policies in APIC

This section describes how to verify that the DHCP relay policies you have created and deployed using the Nexus Dashboard Orchestrator are correctly pushed to each site's APIC. The DHCP policies you create are pushed to the APIC when the bridge domain to which the policy is associated is deployed to a site.

Procedure

Step 1 Log in to the site's APIC GUI.

Step 2 From the top navigation bar, select **Tenants** > **<tenant-name>**.

Select the tenant where you deployed the DHCP policy.

Step 3 Verify that the DHCP relay policy is configured in APIC.

In the left tree view, navigate to **<tenant-name>** > **Policies** > **Protocol** > **DHCP** > **Relay Policies**. Then confirm that the DHCP relay policy you configured has been created.

Step 4 Verify that the DHCP option policy is configured in APIC.

If you have not configured any DHCP option policies, you can skip this step.

In the left tree view, navigate to **<tenant-name>** > **Policies** > **Protocol** > **DHCP** > **Option Policies**. Then confirm that the DHCP option policy you configured has been created.

Step 5 Verify that the DHCP policy is correctly associated with the bridge domain.

In the left tree view, navigate to **<tenant-name>** > **Networking** > **Bridge Domains** > **<bridge-domain-name>** > **DHCP Relay Labels**. Verify that the DHCP policy is also associated with the deployed bridge domain.

Editing or Deleting Existing DHCP Policies

This section describes how to edit or delete a DHCP relay or option policy.

**Note**

- If you make changes to the DHCP policy after you have assigned it to a bridge domain and deployed the bridge domain to one or more sites, you will need to re-deploy it for the DHCP policy changes to update on each site's APIC.
- You cannot delete policies that are associated with one or more bridge domains, you must first unassign the policy from every bridge domain.

Procedure

-
- Step 1** Log in to your Nexus Dashboard Orchestrator GUI.
- Step 2** From the left navigation menu, select **Config > Tenant Templates > Tenant Policies**.
- Step 3** Click the actions menu next to the DHCP policy and select **Edit** or **Delete**.
-



CHAPTER 20

EPG Preferred Group

- [EPG Preferred Groups Overview and Limitations, on page 225](#)
- [Configuring EPGs for Preferred Group, on page 226](#)

EPG Preferred Groups Overview and Limitations

By default, Multi-Site architecture allows communication between EPGs only if a contract is configured between them. If there is no contract between the EPGs, any inter-EPG communication is explicitly disabled. The Preferred Group (PG) feature allows you to specify a set of EPGs that are part of the same VRF to allow full communication between them with no need for contracts to be created.

Preferred Group vs Contracts

There are two types of policy enforcements available for EPGs in a VRF which is stretched to multiple sites with a contract preferred group configured:

- **Included EPGs** – Any EPG that is a member of a preferred group can freely communicate with all other EPGs in the group without any contracts. The communication is based on the `source-any-destination-any-permit` default rule and appropriate Multi-Site translations.
- **Excluded EPGs** – EPGs that are not members of preferred groups continue to require contracts to communicate with each other. Otherwise, the default `source-any-destination-any-deny` rule applies.

The contract preferred group feature allows for greater control and ease of configuring communication between EPGs across sites in a stretched VRF context. If two or more EPGs in the stretched VRF require open communication while others must have only limited communication, you can configure a combination of a contract preferred group and contracts with filters to control the inter-EPG communication. EPGs that are excluded from the preferred group can only communicate with other EPGs if there is a contract in place to override the `source-any-destination-any-deny` default rule.

Stretched vs Shadowed

If EPGs from multiple sites are configured to be part of the same contract preferred group, the Nexus Dashboard Orchestrator creates shadows of each site's EPGs in the other sites in order to correctly translate and program the inter-site connectivity from the EPGs. Contract preferred group policy construct is then applied in each site between a real and shadow EPG for inter-EPG communication.

For example, consider a web-service EPG1 in Site1 and an app-service EPG2 in Site2 added to the contract preferred group. Then if EPG1 wants to access EPG2, it will first be translated to a shadow EPG1 in Site2

and then be able to communicate with EPG2 using the contract preferred group. Appropriate BDs are also stretched or shadowed if the EPG under it is part of a contract preferred group.

VRF Preferred Group Setting

When you configure preferred groups directly in the APIC, you have to explicitly enable the setting on the VRF first before enabling PG membership on individual EPGs. If the PG setting on the VRF is disabled, the EPGs would not be able to communicate without contracts even if they are part of that VRF's preferred group.

**Note**

Beginning with Release 4.0(1), PG configuration in NDO follows the same approach as it does in APIC. In other words, the PG setting on the VRF must be explicitly enabled for the EPGs that are part of that VRF to use the PG configuration.

Nexus Dashboard Orchestrator releases prior to Release 4.0(1) did not allow you to manage the PG setting on VRFs in the GUI, but instead adjusts the setting dynamically as follows:

- If you create and manage the VRF from NDO, NDO will dynamically enable or disable VRF PG value based on whether any EPGs that belong to that VRF are part of the preferred group.

In other words, when you add one or more EPGs to the preferred group, NDO automatically enables the PG setting on the VRF. When you remove the last EPG from the preferred group, NDO disables the VRF flag.

- If you want to permanently enable the PG option on a VRF, you can enable PG on the VRF directly in the APIC first, then import that VRF into NDO.

NDO will preserve the setting and not disable it automatically even if you remove every EPG from the VRF's preferred group.

- If you import the VRF from APIC without first changing the PG setting, NDO will manage the object as if it was created from NDO and overwrite the PG setting dynamically based on EPG membership.

Limitations

The following guidelines and limitations apply when using EPG Preferred Groups:

- Preferred Groups are not supported for intersite L3Out external EPGs.
- EPGs and External EPGs objects in a given VRF must not be configured as part of the Preferred Group if vzAny for that VRF is already consuming or providing a contract.

Configuring EPGs for Preferred Group

This section describes how to enable the Preferred Group (PG) configuration on the VRF and the EPGs.

Before you begin

You must have one or more EPGs added to a schema template.

Procedure

Step 1 Log in to your Cisco Nexus Dashboard and open the Cisco Nexus Dashboard Orchestrator service.

Step 2 From the left navigation pane, choose **Configure > Tenant Template**.

Step 3 Enable PG on the VRF.

- a) Open the Schema that contains the VRF used by the EPGs you want to include in the preferred group.
- b) From the **View Overview** drop-down, select the template that contains the VRF.
- c) Select the VRF.
- d) In the properties sidebar on the right, check the **Preferred Group** check box.

This enabled the PG configuration on that VRF. You will need to enable the PG setting on 2 or more EPGs which you want to be part of the preferred group as described in the next step.

- e) Click **Save** to save the template changes.

Step 4 Configure one or more EPGs to be part of the preferred group.

Note

You must not have a preferred group where some EPGs are managed by Nexus Dashboard Orchestrator and some are managed locally by the APIC.

If you have an existing preferred group in any of the APICs and are planning to import the EPGs from that preferred group into Nexus Dashboard Orchestrator, you must import all EPGs in the group.

- a) If the EPGs you want to include in the preferred group are in a different schema or template, navigate to that template.
- b) Select an EPG.
- c) In the right properties bar, check the **Include in Preferred Group** checkbox.
- d) Click **Save** to save the template changes.

Step 5 (Optional) Verify that all EPGs have been added to the preferred group.

You can view the full list of EPGs that are configured to be part of the preferred group by selecting a VRF and checking the **Preferred Group EPGs** list in the properties sidebar on the right.



CHAPTER 21

External Connectivity (L3Out)

- [L3Out Template Overview, on page 229](#)
- [Guidelines and Limitations, on page 233](#)
- [Greenfield Deployment, on page 234](#)
- [Importing Existing L3Out Configuration, on page 245](#)
- [Viewing L3Out Neighbors, on page 256](#)

L3Out Template Overview

Beginning with release 4.1(1), Nexus Dashboard Orchestrator (NDO) introduced a number of new policies for creating and configuring L3Out for Cisco ACI fabrics, as well as a new template type specifically for IP-based L3Out and SR-MPLS VRF L3Out configurations.

As you may already know, prior releases of NDO provided the ability to create an L3Out object in Application templates that allowed you to create an L3Out and deploy it to your site. However, the actual L3Out configurations had to be done manually by logging in to the sites' controllers (Cisco APIC) and providing the details for each L3Out individually.

With release 4.1(1), the entire configuration of L3Outs and SR-MPLS L3Outs (including nodes, interfaces, and other settings) can be done directly in NDO and deployed to all fabrics in your Multi-Site domain. To achieve this, a new L3Out-specific template type has been added to contain the L3Out and SR-MPLS VRF L3Out configurations. Similar to Application templates, L3Out templates have a one-to-one association with tenants but unlike Application templates, an L3Out template must be associated to a single site only.



Note

The legacy L3Out objects in the Application templates remain functional for backward compatibility. However, if you want to define specific L3Out and SR-MPLS L3Out settings from NDO, you must use the new L3Out template type.

The legacy SR-MPLS VRF L3Out object has been removed from the Application template and all SR-MPLS VRF L3Out configurations must be now done using the L3Out-specific template. The SR-MPLS Infra L3Out configuration is still performed as part of the site connectivity provisioning workflow.

Templates and Policy Objects Dependencies

The following diagram illustrates the template and policy hierarchy across multiple templates that's required for defining a complete L3Out configuration:

- The VRF used by the L3Out and the External EPGs that are associated to the L3Out continue to be defined in the Application templates.
- Node or interface routing policies, BGP peer prefix, and IP SLA policies are now defined in the Tenant Policy template.

These policies are used by the L3Out-specific template and the policies defined in that template as described in the following bullet point.

- For IP-based L3Outs, the template includes the following:
 - Routing Protocol (BGP/OSPF), VRF, L3 Domain and Route Maps for route control.
 - Border leaf switches (nodes) where to deploy the L3Out routing protocol and node-level protocol configurations.
 - Border leaf switch interfaces where to deploy the L3Out routing protocol and interface-level protocol configurations.
 - Node- and interface-level common configuration using Node/Interface Group policies.

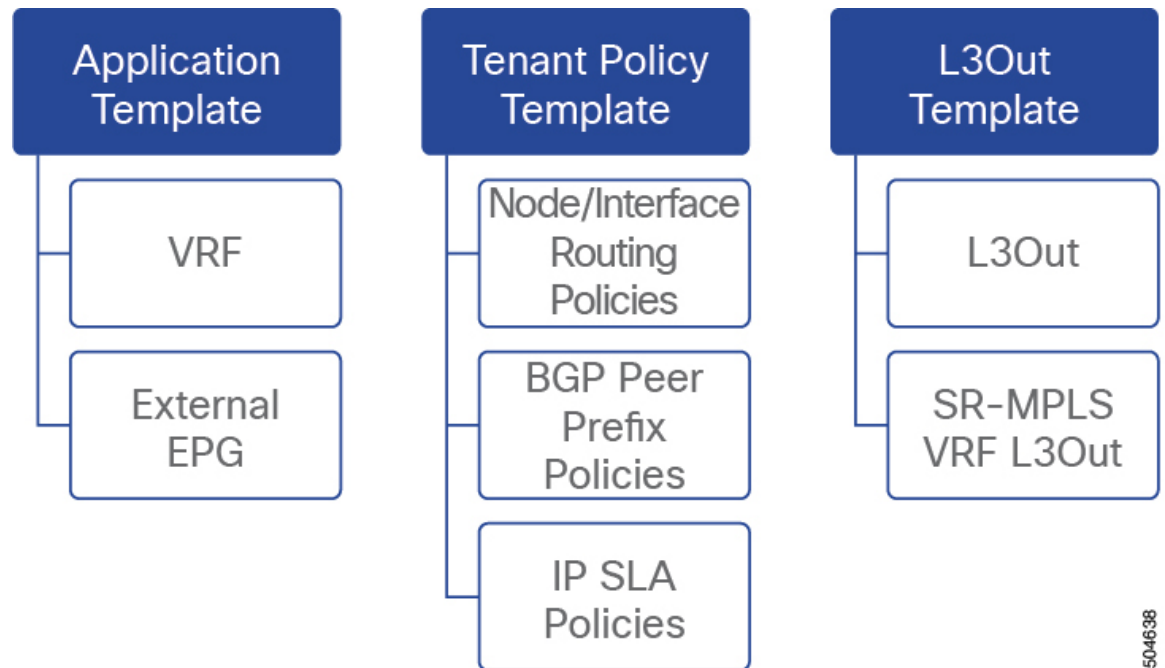
Node Group configuration includes BGP peers for loopback interfaces, BFD multi-hop settings, and association with Node Routing Group Policy described below.

Interface Group configuration includes OSPF and BFD protocol settings and association with Interface Routing Group Policy described below.

These policies consume policies defined in Tenant Policy templates mentioned in the previous bullet point. For example, the node and interface group policies require the node and interface routing policies defined in the Tenant Policy templates

- For SR-MPLS VRF L3Outs, the template allows you to define labels and import/export route maps for route control.

Figure 22: L3Out Templates and Policy Objects



504638

Tenant Policy Template: Node Routing Group Policy

The Node Routing Policy in Tenant Policy template is a set of protocol policies that can be applied at a node or border leaf level and can be used by node group policies in L3Out template. It includes the following 3 settings:

- **BFD MultiHop Settings** – specifies BFD parameters for BFD sessions established between devices on interfaces that are not directly connected.
- **BGP Node Settings** – allows you to configure BGP protocol timer and sessions settings for BGP adjacencies between BGP peers.
- **BGP Best Path Control** – enables `as-path multipath-relax`, which allows load-balancing between multiple paths received from different BGP ASN.

This policy is configured and deployed using Tenant Policy templates and is used by the L3Outs configured in L3Out templates.

Tenant Policy Template: Interface Routing Group Policy

The Interface Routing Policy in Tenant Policy template is a set of policies that can be applied at an interface level and can be used by interface group policies in L3Out template. It includes the following 3 settings:

- **BFD Settings** – specifies BFD parameters for BFD sessions established between devices on interfaces that are directly connected.

When multiple protocols are enabled between a pair of routers, each protocol has its own link failure detection mechanism, which may have different timeouts. BFD provides a consistent timeout for all protocols to allow consistent and predictable convergence times.

- **BFD MultiHop Settings** – specifies BFD parameters for BFD sessions established between devices on interfaces that are not directly connected.

You can configure these settings at the node level as mentioned in the "Tenant Policy Template: Node Routing Group Policy" section above, in which case the interfaces inherit those settings, or you can overwrite the node-level settings for individual interfaces in the Interface Routing group policy.



Note BFD multi-hop configuration requires Cisco APIC release 5.0(1) or later.

- **OSPF Interface Settings** – allows you to configure interface-level settings such as OSPF network type, priority, cost, intervals and controls.



Note This policy must be created when deploying an L3Out with OSPF.

This policy is configured and deployed using Tenant Policy templates and is used by the L3Outs configured in L3Out templates.

Tenant Policy Template: Individual Policies

In addition to the group policies described above, the Tenant Policy templates also contain the following individual policies related to L3Out configuration:

- **BGP peer prefix policy** – defines how many prefixes can be received from a neighbor and what action to take when the number of the allowed prefixes is exceeded.

This policy is configured and deployed using Tenant Policy templates and is used by the L3Outs configured in L3Out templates.

- **IP SLA monitoring policy** – defines the type of probe (ICMP/TCP/HTTP) and respective settings to use for monitoring endpoints. This policy is associated with monitoring probe profiles known as "track members", which represent a network segment to be monitored. You can associate an IP SLA monitoring policy to a track list (which includes multiple track members) and associate this track list to a Static Route for monitoring availability of track list members over the route. In addition, you can associate IP SLA monitoring policy directly to next-hop address of a Static Route for monitoring its availability over the route.



Note IP SLA monitoring policy of HTTP type requires Cisco APIC release 5.1(3) or later.

- **IP SLA track list** - defines the IP addresses to be tracked, IP SLA Monitoring policy (probe frequency and type), and scope (bridge domain or L3Out). IP SLA track list aggregates one or more track members, defines what percentage or weight of track members must be ^{up/down} for the route to be considered available or unavailable. Based on the track list, the available routes remain in the routing table and the unavailable routes are removed until the track list recovers.

This policy is configured and deployed using Tenant Policy templates and is used by the L3Outs configured in L3Out templates. In addition, an IP SLA track list can be configured in the same Tenant Policy template as the monitoring policy and consumed by it.

L3Out Template

The L3Outs defined in L3Out templates allow you to define all the required configurations to enable connectivity from the endpoints inside your ACI fabrics to outside network domains through routing protocols or static routes. The L3Out object in NDO contains settings necessary for the following:

- Learning external routes via routing protocols or static routes.
- Distribution of the learned external routes to other leaf switches.
- Advertisement of ACI internal routes (BD subnets) to the outside networks.
- Advertisement of learned external routes to other L3Outs (transit routing).

When you create an L3Out template and configure L3Out-specific objects and properties as described later in [Creating L3Out Template, on page 240](#), you will:

1. Define a number of common properties, such as the VRF, L3 Domain, and routing protocol (BGP and/or OSPF), for the L3Out.
2. Specify one or more border leaf switches (nodes) and optionally associate each node with a Node Group policy.
3. Specify one or more interfaces on those border leaf switches and optionally associate each interface with an Interface Group policy described above.
4. After you have created an L3Out template and deployed one or more L3Outs (and their associated External EPGs, defined inside Application templates), you can control traffic between the ACI EPGs and external networks using contracts in Application templates as you typically would.

Guidelines and Limitations

The following guidelines apply when using an L3Out template to configure IP-based L3Outs and SR-MPLS VRF L3Outs:

- Similar to Application templates, L3Out templates have a one-to-one association with tenants but unlike Application templates, an L3Out template must be associated to a single site only.
- The legacy L3Out container objects in the Application templates remain functional for backward compatibility.

Note however, if you want to define specific L3Out and SR-MPLS VRF L3Out settings, you must use the L3Out-specific template type. As such, we recommend using the L3Out-specific templates for all new L3Out and SR-MPLS VRF L3Out configurations.

- The legacy SR-MPLS VRF L3Out contain object has been removed from the Application template. All SR-MPLS VRF L3Out configurations must be done using the L3Out-specific template.
- If you want to configure BFD multi-hop settings, your fabric must be running Cisco APIC release 5.0(1) or later.
- If you want to configure an IP SLA monitoring policy of HTTP type, your fabric must be running Cisco APIC release 5.1(3) or later.

Greenfield Deployment

Creating Tenant Policy Template

This section describes how to create a Tenant Policy template and define the L3Out-specific policies, which you will then consume in an L3Out template as described later in this document. For more information about each policy and how it relates to policies and settings in other templates, see [L3Out Template Overview, on page 229](#).



Note If you want to import existing L3Out configurations from a site's APIC, follow the "Importing Existing L3Out Configuration" steps in the following sections of this chapter instead.

Before you begin

- You must have the Cisco Nexus Dashboard Orchestrator service that is installed and enabled.
- You must have the fabrics onboarded to your Cisco Nexus Dashboard and enabled for management in the Orchestrator service.
- Ensure you have read and understood the Templates and Policy Objects dependencies that are described in [L3Out Template Overview, on page 229](#).

Procedure

-
- Step 1** Log in to your Nexus Dashboard Orchestrator.
- Step 2** In the left navigation pane, choose **Configure > Tenant Templates**.
- Step 3** Choose the **Tenant Policies** tab.
- Step 4** In the main pane, click **Create Tenant Policy Template**.
- If you want to update an existing Tenant Policy template instead, simply click its name. This opens the **Tenant Policies** page.
- Step 5** If you created a brand new template, provide the **Name** for the template and **Select a Tenant** with which you want to associate this template.
- Step 6** Associate the template with one or more sites.
- a) In the **Tenant Policies** template view, choose **Actions > Add/Remove Sites**.



b) In the **Associate Sites to <template-name>** dialog, select the sites to which you want to deploy the template.

Step 7

Create a Route Map Policy for Route Control.

While you can associate BDs to the created L3Out (for example, to advertise out the BDs' subnets), we recommend that you create the **Outbound Route Map** for the L3Out instead because it can be used for both BDs' subnets and transit routes received from other L3Outs.

Note

Once an Outbound Route Map is associated to the L3Out, it is no longer possible to advertise out BDs' subnets by associating the BD to the L3Out.

- From the **+Create Object** drop-down, select **Route Map Policy for Route Control**.
- In the right properties sidebar, provide the **Name** for the policy.
- (Optional) Click **Add Description** and provide a description for the policy.
- Click **+Add Entry** and provide the route map information.

For each route map, you must create one or more context entries. Each entry is a rule that defines an action based on one or more matching criteria based on the following information:

- **Context Order** – Context order is used to determine the order in which contexts are evaluated. The value must be in the 0–9 range.
- **Context Action** – Context action defines the action to perform (`permit` or `deny`) if a match is found. If the same value is used for multiple contexts, they are evaluated one in the order in which they are defined.

When the context order and action are defined, choose how you want to match the context:

- Click **+Create Attribute** to specify the action that will be taken should the context match.

You can choose one of the following actions:

- Set Community
- Set Route Tag
- Set Dampening
- Set Weight
- Set Next Hop
- Set Preference
- Set Metric
- Set Metric Type
- Set AS Path

- Set Additional Community

After you have configured the attribute, click **Save**.

- If you want to associate the action that you defined with an IP address or prefix, click **Add IP Address**.

In the **Prefix** field, provide the IP address prefix. Both IPv4 and IPv6 prefixes are supported, for example, 2003:1:1a5:1a5::/64 or 205.205.0.0/16.

If you want to aggregate IPs in a specific range, check the **Aggregate** check box and provide the range. For example, you can specify 0.0.0.0/0 prefix to match any IP or you can specify 10.0.0.0/8 prefix to match any 10.x.x.x addresses.

- If you want to associate the action that you defined with community lists, click **Add Community**.

In the **Community** field, provide the community string. For example, regular:as2-nn2:200:300.

Then choose the **Scope**: *Transitive* means that the community will be propagated across eBGP peering (across autonomous systems) while *Non-Transitive* means the community will not be propagated.

Note

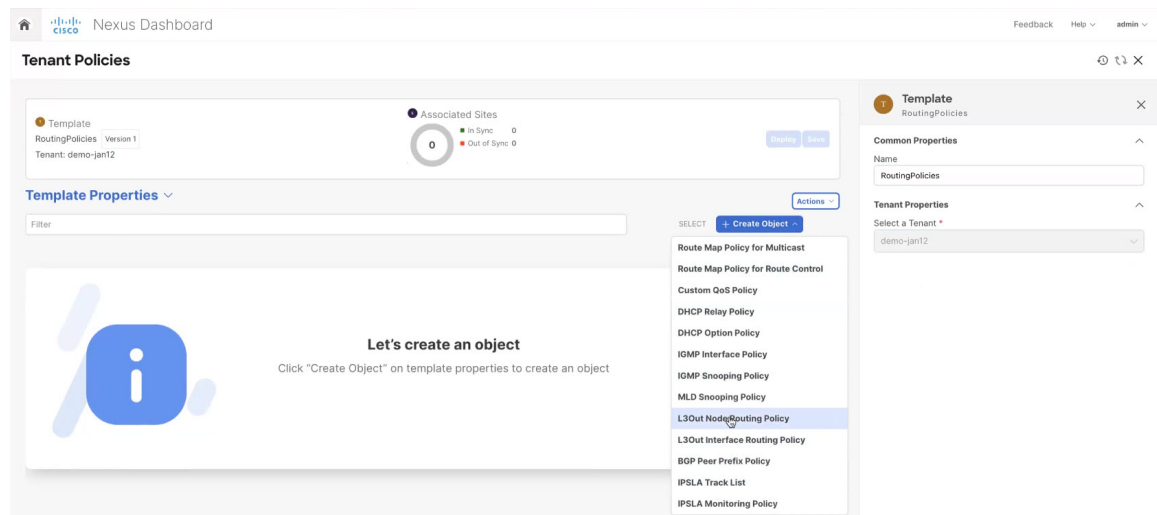
You must specify an **IP address** or a **Community** string to match a specific prefix (even if you do not provide a **Set** attribute) because it defines the prefixes that must be announced out of the L3Out. This can be either BDs' subnets or transit routes learned from other L3Outs.

- Repeat the previous substeps to create any additional route map entries for the same policy.
- Click **Save** to save the policy and return to the template page.
- Repeat this step to create any additional Route Map for Route Control policies.

Step 8

Create an L3Out Node Routing policy.

- In the main pane, choose **Create Object > L3Out Node Routing Policy**.



- Provide the **Name** for the policy, and **Add** at least one of the **BFD MultiHop Settings**, **BGP Node Settings**, or **BGP Best Path Control** options.

The screenshot shows the Cisco Nexus Dashboard Orchestrator interface. The top navigation bar includes the Cisco logo, 'Nexus Dashboard', and 'Orchestrator'. The main section is titled 'Tenant Policies'. On the left, under 'Template Properties', there is a filter input and a 'Create Object' button. A list of templates is shown, including 'L3Out Node Routing Policy' and 'UntitledL3OutNodePolicy2'. On the right, a detailed view of the 'L3Out Node Routing Policy' is shown, including a name field, a description field, and a section for 'BFD MultiHop Settings', 'BGP Node Settings', and 'BGP Best Path Control', each with an 'Add' button.

- **BFD MultiHop Settings** – provides forwarding failure detection for destinations with more than one hop.

In this case, a MultiHop session is created between the source and destination instead of the interface like in single-hop scenarios.

Note

BFD MultiHop configuration requires Cisco APIC release 5.0(1) or later.

- **BGP Node Settings** – allows you to configure BGP protocol timer and sessions settings for BGP adjacencies between BGP peers.
- **BGP Best Path Control** – enables `as-path multipath-relax`, which allows load-balancing between multiple paths that are received from different BGP ASN.

Step 9 Create an L3Out Interface Routing policy.

- In the main pane, choose **Create Object > L3Out Interface Routing Policy**.
- Provide the **Name** for the policy, and define the **BFD Settings**, **BFD Multi-Hop Settings**, and **OSPF Interface Settings**.

- **BFD Settings** – specifies BFD parameters for BFD sessions established between devices on interfaces that are directly connected.

When multiple protocols are enabled between a pair of routers, each protocol has its own link failure detection mechanism, which may have different timeouts. BFD provides a consistent timeout for all protocols to allow consistent and predictable convergence times.

- **BFD MultiHop Settings** – specifies BFD parameters for BFD sessions established between devices on interfaces that are not directly connected.

You can configure these settings at the node level as mentioned in the "Tenant Policy Template: Node Routing Group Policy" section above, in which case the interfaces inherit those settings, or you can overwrite the node-level settings for individual interfaces in the Interface Routing group policy.

Note

BFD multi-hop configuration requires Cisco APIC release 5.0(1) or later.

- **OSPF Interface Settings** – allows you to configure interface-level settings such as OSPF network type, priority, cost, intervals and controls.

Note

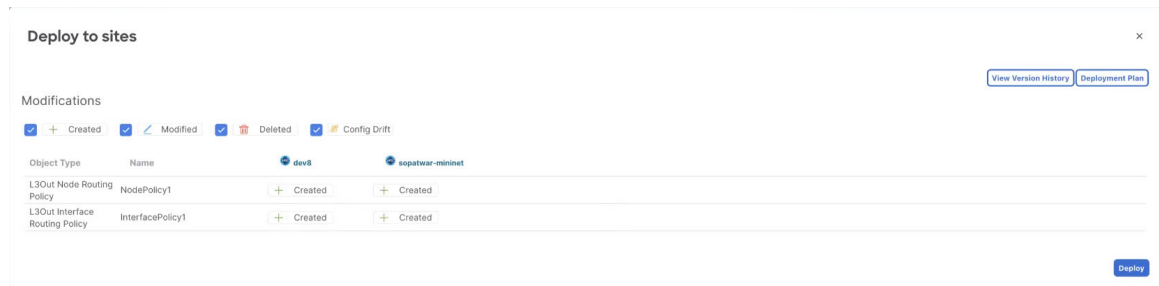
This policy must be created when deploying an L3Out with OSPF.

Step 10 Click **Save** to save the template changes.

Step 11 Deploy the template to site(s).

At this stage, we can deploy the created group policies to the sites and verify them in your APIC as a checkpoint before proceeding with additional configurations.

- In the **Tenant Policies** template view, click **Deploy**.
- In the **Deploy to sites** dialog, confirm the policies being deployed and click **Deploy**.



- c) (Optional) Verify that the policies were deployed correctly.

You can verify that the template was correctly deployed to a site by navigating to the site's APIC, choosing **Tenants > <tenant-name> > Policies > Protocol** and checking the **BFD**, **BGP**, and **OSPF** policies.

Note that while each policy is displayed as a separate object in the APIC GUI, NDO simplifies the configuration workflow by combining them into a single template at the node and interface levels.

Step 12

Create a BGP Peer Prefix policy.

- In the main pane, choose **Create Object > BGP Peer Prefix Policy**.
- Provide the **Name** for the policy, and define the **Max Number of Prefixes** and the **Action** to take if the number is exceeded.

The following actions are available:

- **Log**
- **Reject**
- **Restart**
- **Shutdown**

Step 13

Create an IP SLA Monitoring policy.

- In the main pane, choose **Create Object > IP SLA Monitoring Policy**.
- Provide the **Name** for the policy, and define its settings.

Note

If you choose **HTTP** for the **SLA Type**, your fabric must be running Cisco APIC release 5.1(3) or later.

Step 14

Create an IP SLA Track List.

- In the main pane, choose **Create Object > IP SLA Track List**.
- Provide the **Name** for the policy.
- Choose the **Type**.

The definition of a route being available or not available can be based on **Threshold Percentage** or **Threshold Weight**.

- Click **+Add Track List to Track Member Relation** to add one or more track members to this track list.

Note

You must select a bridge domain or an L3Out to associate with the track member. If you do not already have the bridge domain (BD) or L3Out that is created, you can skip adding a track member, save the policy without assigning one, and come back to it after you have created the BD or L3Out.

- e) In the **Add Track List to Track Member Relation** dialog, provide the **Destination IP**, **Scope Type**, and choose the **IP SLA Monitoring Policy**.

The scope for the track list can be either bridge domain or L3Out. The IP SLA Monitoring policy is the one you created in the previous step.

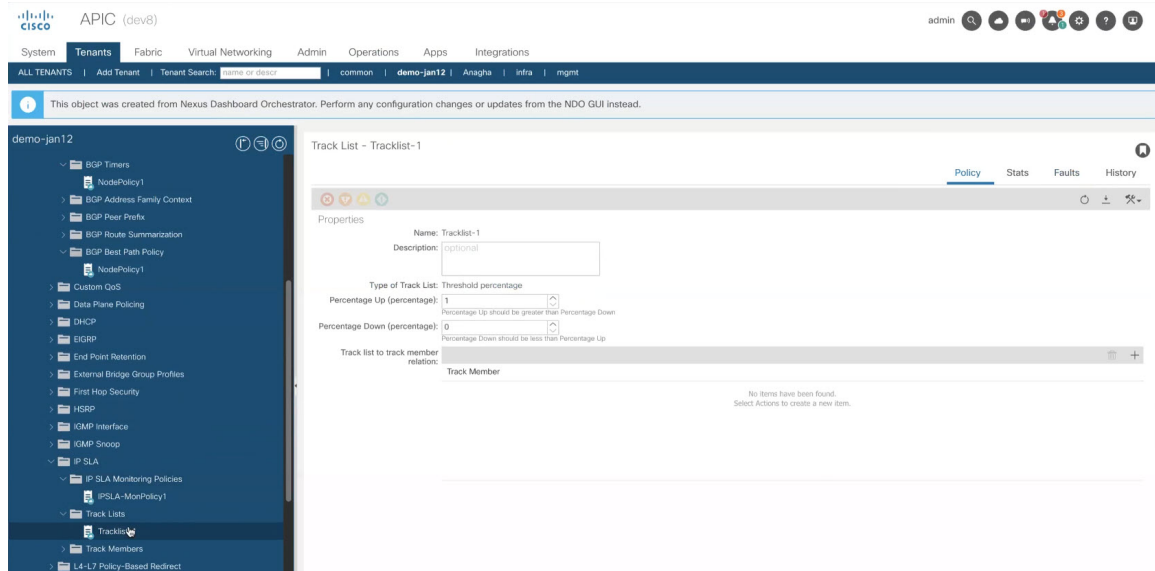
Step 15 Click **Save** to save the template changes.

Step 16 Deploy the template to site(s).

At this stage, we can create another checkpoint by deploying the defined policies to the sites.

- In the **Tenant Policies** template view, click **Deploy**.
- In the **Deploy to sites** dialog, confirm the policies being deployed and click **Deploy**.
- (Optional) Verify that the policies were deployed correctly.

You can verify that the template was correctly deployed to a site by navigating to the site's APIC, choosing **Tenants** > **<tenant-name>** > **Policies** > **Protocol** and checking the **IP SLA** policies, for example:



What to do next

After you've defined the policies in the Tenant Policy template, proceed to [Creating L3Out Template, on page 240](#).

Creating L3Out Template

This section describes how to create an L3Out template and define IP-based L3Out policies, which you will then use with the VRFs and EPGs in an Application template to deploy a complete L3Out configuration to your fabrics. For more information about each policy and how it related to policies and settings in other templates, see [L3Out Template Overview, on page 229](#).

If you are looking to create an SR-MPLS VRF L3Out, see the steps described in [Multi-Site and SR-MPLS L3Out Handoff, on page 337](#).

Before you begin

- You must have created a Template Policy template and defined any policies specific to your deployment scenario, as described in [Creating Tenant Policy Template, on page 234](#).
- Created a VRF that you want to use for the L3Out in one of your Application templates as you typically would.

Procedure

Step 1 In the left navigation pane, choose **Configure > Tenant Templates > L3Out**.

Step 2 In the main pane, click **Create L3Out Template**.

If you want to update an existing L3Out template instead, simply click on its name. This opens the **L3Out Template** page.

Step 3 If you are creating a brand new template, choose the **Tenant** and **Site** with which to associate the template, then click **Save and go to template**.

Each L3Out template is associated with a specific tenant similar to other NDO templates, however it is also assigned to a single site only as L3Out configuration is typically site-specific.

If you want to define L3Out configuration for multiple sites, you must create at least one L3Out template for each site, but you can deploy multiple L3Outs per site/tenant by defining all of them in the same L3Out template. You may have multiple L3Out templates per site as long as they are assigned to different tenants.

Step 4 Provide the **Name** for the template.

Step 5 Create an IP-based L3Out and provide its general configuration.

- a) In the main pane, choose **Create Object > L3Out**.
- b) Provide the **Name** for the L3Out.

Note

We recommend providing unique names for all L3Outs across sites, even if they belong to the same tenant or allow connectivity to the same external resources.

- c) Click **Select VRF>** and choose a VRF to associate with this L3Out.

Note that if you save and deploy the template at this time, the behavior would be identical to what was previously available in NDO release 4.0(x) and earlier. The following steps describe additional settings available in release 4.1(1) and later to allow full L3Out configuration directly from NDO.

- d) Click **Select L3 Domain>** and choose the L3 domain to associate with this L3Out.

The L3 domains can be created directly in the APIC or in NDO using the **Fabric Management > Fabric Policies** page, as described in the [Fabric Management Templates, on page 89](#) chapter.

- e) Select the **Routing Protocols** used by this L3Out.

You can select **BGP** or **OSPF** or both. Alternatively, you can leave both protocols disabled if you plan to use static routing on this L3Out.

If you enable OSPF, you must also provide the **OSPF Area ID** and **OSPF Area Type**.

For both OSPF and BGP:

- Provide **Outbound Route Map** to advertise fabric's BD subnets or prefixes learned from other L3Outs (transit routing) to the outside.

This is the **Route Map Policy for Route Control** that you created in the previous section.

While you can associate BDs to the created L3Out (for example, to advertise out the BDs' subnets), we recommend that you create the **Outbound Route Map** for the L3Out instead because it can be used for both BDs' subnets and transit routes received from other L3Outs.

Note

Once an Outbound Route Map is associated to the L3Out, it is no longer possible to advertise out BDs' subnets by associating the BD to the L3Out.

If an Outbound Route Map is specified here, it must include all prefixes which need to be advertised toward the external network domain. BD subnets configured with BD to L3Out associations and External EPG subnets configured with export route control will not work when this route-map configuration is deployed.

- (Optional) Enable **Import Route Control** to control the external prefixes that should be redistributed inside the fabric.

Step 6

Add one or more border leaf switches (nodes) for the L3Out.

- Click **+Create Node**.
- In the **Create Node** dialog, choose a **Node ID**.
- Provide the **Router ID**.
- Skip the **Node Group Policy** selection.

You can deploy consistent configuration across all nodes by configuring a **Node Group Policy** and applying it to the nodes, as described in Step 9. If you choose to create a common policy, Step 9 will describe how to update the nodes you are adding in this step to use the group policy.

- Choose whether you want to **Use Router ID as Loopback**.
- If you want to define one or more static routes, click **+Add Static Routes**.

For all static routes, you must define an IP address **Prefix** including the network mask using the `ab.cd.ef.gh/xy` format, choose whether you want to **Create a static route to Null0**, and define the **Next Hop** IP address. When providing the next hop IP, you can also choose the **Administrative Distance** and the **Monitoring Policy** which you created in [Creating Tenant Policy Template, on page 234](#).

Here you can also select the **Track Policy**, which you defined in [Creating Tenant Policy Template, on page 234](#).

- Repeat this step for any additional border leaf switches where you want to deploy this L3Out.

Step 7

Add one or more interfaces for the L3Out.

- Click **+Create Interface**.
- Choose the type of interface you want to add.

This release supports the same interface types as the APIC:

- Routed Interface
- Routed Sub-Interface
- SVI
- Floating SVI

You can use the same configuration parameters as you would typically use when configuring an interface directly in the APIC, for example:

Add Routed Interface

Interface Type

Node Id
 dev8-leaf1 (Node-101)

Interface *
 eth1/8

Interface Group Policy

Addresses

Addresses ⓘ
☒ IPv4 Primary Address
 10.1.1.1/24
☒ IPv6 Primary Address
 10::1/64

Secondary Addresses

Address	ND RA PREFIX	IPv6 DAD
+ Add Secondary Address		

MAC Address *
 00:22:BD:F8:19:FF

MTU Bytes ⓘ *
 inherit

L3Out BGP Peers

Peer Address IPv4	Peer Address IPv6
+ Add L3Out BGP Peer	

Advanced Settings

Link Local Address V6 ⓘ

☐ IPv6 DAD

Target DSCP
 Unspecified

PTP Configuration

PTP State

Ok

- c) Repeat this step for any additional interfaces where you want to deploy this L3Out configuration.

Step 8

(Optional) Add one or more node or interface group policies.

While you can configure each node and interface individually as mentioned in the previous two steps, you can also define one or more node or interface group policies and apply a group policy to multiple nodes or interfaces for consistent configuration across them.

- Click **+Create Node/Interface Group Policy**.
- Choose whether you're defining a **Node** or **Interface** group policy and provide a **Name** for it.
- Select the **Node Routing Policy** or **Interface Routing Policy** respectively.

Note

An Interface Group Policy is mandatory when using OSPF on the L3Out.

Those are the policies you created in [Creating Tenant Policy Template, on page 234](#), for example:

- d) Provide any additional node or interface configurations settings as required by your deployment.
Keep in mind that all nodes or interfaces to which you apply this group policy will have exact same configuration as defined in the group policy.
- e) Click **Ok** to save the group policy.
- f) Repeat this step for any additional node or interface group policies for this L3Out.

Step 9

(Optional) Apply a node or interface group policy to one or more nodes/interfaces.

- a) Click on one of the nodes or interfaces you configured for this L3Out.
- b) From the **Node/Interface Group Policy** dropdown, select the group policy you defined in the previous step.

- c) Repeat this step for all nodes and interfaces to which you want to apply the consistent settings defined by the group policies.

Step 10

Click **Save** to save the template changes.

Step 11

Deploy the template to site.

- a) In the **L3Out Template** page, click **Deploy**.

- b) In the **Deploy to sites** dialog, confirm the policies being deployed and click **Deploy**.
- c) (Optional) Verify that the policies were deployed correctly.

You can verify that the template was correctly deployed to a site by navigating to the site's APIC, choosing **Tenants** > *<tenant-name>* > **Networking** > **L3Outs** and checking the L3Out name you provided in NDO.

Note that while you define all of the L3Out configurations in the same template in NDO, separate individual policies are created in the APIC. For example, separate policies are created for the nodes, interfaces, and even IP address types (providing IPv4 and IPv6 IP addresses for a single L3Out interface creates two separate interface profiles) in the APIC.

Importing Existing L3Out Configuration

Overview of Importing L3Out Configuration

Beginning with release 4.1(2), Nexus Dashboard Orchestrator (NDO) supports importing existing L3Out configurations from the APIC sites. The following sections focus on the guidelines and specific steps required to import an L3Out along with its associated policies.

**Note**

If you want to configure and deploy new IP-based L3Out configurations (greenfield deployment), see the earlier sections of this chapter.

If you want to configure or import SR-MPLS VRF L3Out, see the [Multi-Site and SR-MPLS L3Out Handoff, on page 337](#) chapter instead.

This release supports importing the following policies.

- **Route Maps** – may be referenced in the L3Out template's **Outbound Route Map** and **Inbound Route Map** fields to define route import and export policies.
- **L3Out Node Policies:**
 - Nodes configured for an L3Out can be associated to a node group, which in turn can refer to a node routing policy.
 - Node groups can also reference BGP Peer Prefix policy when configuring BGP peers for the nodes.
- **L3Out Interface Policies:**
 - Interfaces configured for an L3Out can be associated to an interface group, which can refer to an interface routing policy and BGP Peer Prefix policy.
 - Interface groups can also reference BGP Peer Prefix policy when configuring BGP peers for the interfaces.
- **BGP Peer Prefix** – can be referenced by the node and interfaces groups for BGP peer configuration on all nodes in the group.

- **IP SLA Monitoring policies and IP SLA Track lists** – can be referenced by the static routes defined for a node.
- **Custom QoS policy** – can be referenced by interface group configuration.

Mapping of Sites' MOs to NDO Objects and Groups

Note that in some cases there is no 1:1 mapping between the managed objects (MOs) created in the site and the policy objects as they are seen on and managed by the Orchestrator. In these cases, when you import an L3Out from APIC, NDO creates NDO-specific logical groups that may contain multiple individual MOs; for example, the following APIC policies are grouped on import:

- The following MOs are grouped into an L3Out Interface Routing policy on NDO:
 - OSPF Interface Policy
 - BFD Policy
 - BFD Multi-Hop Interface Policy
- The following MOs are grouped into an L3Out Node Routing policy on NDO:
 - BGP Timer Policy
 - BGP Best Path Policy
 - BFD Multi-Hop Node Policy



Note If you import an L3Out configuration and then later change one of these policies directly on the APIC, you must re-import the policies in the Tenant Policy template that contains them on NDO.

The following figure shows the **L3Out Node Routing Policy** object in NDO that groups together the 3 policies mentioned above:

Automatic Import of Dependencies

Tenant Policies templates include objects and policies that have local references within the template. For example, an IP SLA track list can contain a list of track members and each track member must refer to a IP SLA monitoring policy. In such cases, importing existing configuration that contains one or more IP SLA track list policies from a site will also automatically import the referenced IP SLA monitoring policy. The import workflow displays additional information about the automatically imported policies when you select an object that has such dependencies:

Importing IP SLA Policies

Typically, IP SLA track members have a Bridge Domain (BD) or an L3Out scope. When you import an IP SLA track list along with its members, NDO will attempt to automatically assign the correct BD or L3Out to those members. However, at the time of the import, the BD or L3Out objects may not yet exist in NDO.

In such cases, NDO still allows you to import the IP SLA track members with a missing scope object reference. To keep track of the correct reference, NDO sets **Scope Type** to `Local Reference` and saves the name of the referenced BD or L3Out in a `scopeDn` property of the IP SLA track member object:

Update Track List to Track Member Relation ×

TrackMember

Destination IP *

10.0.0.1

Scope Type *

BD L3Out Local Reference

Local Reference

demo-tenant/l3out-2

IPSLA Monitoring Policy *

ipslaMonPol-1 ×

Ok

This allows you to save the template that contains the imported IP SLA track members and re-deploy it back to the site, where the `scopeDn` value is used to correctly program the scope reference for the policy.

To import the entire L3Out configuration, you need to import the L3Out objects after you've imported the relevant Tenant policies. So in case where you import the IP SLA track members first, you must manually update their **Scope Type** and reference after you have also imported the associated L3Out. The `scopeDn` and `scopeType=Local Reference` are internal values and can be set only by the configuration import workflow.

References to Policies in Tenant "Common"

Some policies that you import from a site may contain references to policies in tenant `common`. Importing such policies will automatically create a copy of the tenant `common` policy in the Tenant Policies template where the objects are being imported and as a result of that, in the tenant associated with that Tenant Policies template, for example:

- If you import an IP SLA track list that contains a track member which refers to an IP SLA monitoring policy from the `common` tenant, a copy of the tenant `common`'s IP SLA monitoring policy will be created in the Tenant Policies template and the imported track member will reference this newly added IP SLA monitoring policy.
- If you import an L3Out that contains node configuration with a static route which references an IP SLA track list from tenant `common`, a copy of the tenant `common`'s IP SLA track list will be created in the Tenant Policies template.

Unsupported Scenarios

If an L3Out contains one or more configuration options that are currently not supported by NDO, you will not be able to import that L3Out. The following configurations are currently not supported by NDO and will prevent you from importing any L3Out that includes them:

- For IP-based L3Outs:
 - Layer 3 EVPN Services for Fabric WAN (GOLF)
 - Enhanced Interior Gateway Routing Protocol (EIGRP)
 - Fallback Route Group
- For node profiles:
 - Intersite Loopback Addresses
- For interfaces:
 - DHCP Relay
 - SVI/FSVI External Bridge Group Profile
 - VXLAN Encap
- For interface profiles:
 - Internet Group Management Protocol (IGMP)
 - Hot Standby Router Protocol (HSRP) Interfaces
 - DHCP Relay
 - Enhanced Interior Gateway Routing Protocol (EIGRP)
 - Ingress/Egress Data Plane Policies
 - Neighbor Discovery (ND) Policy
 - PIM and PIMv6 Interface Policies
 - NetFlow Monitor Policies

In these cases, the import workflow UI will display an orange exclamation point icon with a message explaining the issue and you will not be able to select that L3Out for import:

Importing Tenant Policy Template Objects

This section describes how to import existing L3Out configuration policies from Cisco APIC into NDO's Tenant Policies template. For more information about each policy and how it relates to policies and settings in other templates, see [Overview of Importing L3Out Configuration, on page 245](#).

Before you begin

- If you want to configure and deploy new L3Out configurations (greenfield deployment), see the earlier sections of this chapter instead.

- You must have the Cisco Nexus Dashboard Orchestrator service that is installed and enabled.
- You must have the fabrics onboarded to your Cisco Nexus Dashboard and enabled for management in the Orchestrator service.
- Ensure you have read and understood the Templates and Policy Objects dependencies that are described in [Overview of Importing L3Out Configuration, on page 245](#).
- Ensure that no changes are made to the tenant policies or the L3Outs which you plan to import into NDO between the time you import the tenant policies as described in this section and when you redeploy the imported L3Out as described in the next section.

There's no drift notification in NDO in case an imported policy that is used by the L3Out is modified directly in the APIC before all policies that are associated with the L3Out are imported and redeployed to be managed by NDO.

Procedure

-
- Step 1** Log in to your Cisco Nexus Dashboard and open the Orchestrator service.
- Step 2** In the left navigation pane, choose **Configure > Tenant Template > > Tenant Policies**.
- Step 3** In the main pane, click **Add Tenant Policy Template**.
- If you want to update an existing Tenant Policy template instead, simply click its name. This opens the **Tenant Policies** page.
- Step 4** If you created a brand new template, provide the **Name** for the template and **Select a Tenant** from which you plan to import configuration.
- Step 5** Associate the template with the site from which you plan to import configuration.
- In the **Tenant Policies** template view, choose **Actions > Sites Association**.
 - In the **Associate Sites to <template-name>** dialog, select the sites to which you want to deploy the template.
- Step 6** Click **Save** to save the template changes.
- Step 7** Import one or more policies into the Tenant Policies template.
- When you choose to import L3Out configuration from a site, the UI shows the list of L3Out policies that can be imported. You may select one or more L3Out policies and import all the provider policies that are used by the L3Out into this Tenant Policy Template.
- In the **Tenant Policies** screen's **Template Properties** view, choose **Import > <site-name>**.
 - In the **Import from <site-name>** dialog, choose one or more L3Outs and click **Import**.
- If there's an L3Out already configured in the site, its associated policies are available for import in the **L3Out Related Tenant Policies** category. When you select an L3Out to import, all policies that are referenced by that L3Out in the site's APIC are imported into the Tenant Policies template you are editing.

- c) Verify that all imported policies are shown in the template and click **Save** to save it.

All policies configured for the L3Out in the site, which you chose to import in the previous step, are added to the Tenant Policies template using the following guidelines:

- Default import route maps are named as `<l3out-name>_imp_<site-id>`.
- Default export route maps are named as `<l3out-name>_exp_<site-id>`.
- Node routing policies are numbered, for example `L3OutNodePolicy1`, `L3OutNodePolicy2`, and so forth.
- Interface routing policies are numbered, for example `L3OutInterfacePolicy1`, `L3OutInterfacePolicy2`, and so forth.

Tenant Policies

- d) If necessary, update the policy names and click **Save** to save the changes.

We recommend keeping the names of the imported policies as they are created. In this case, when you import L3Outs into the L3Out template as described in the next section, the referenced policies will be automatically recognized and configured for the L3Out by NDO.

However, if you have a specific naming convention in your Multi-Site domain, you can update the imported objects' names to follow that convention. In this case, you must manually provide object references during L3Out import in the next section.

Note

For some objects, there is no 1:1 mapping between the managed objects (MOs) created in the site and the policy objects as they are seen on and managed by the Orchestrator. For information about which MOs are combined into logical groups in NDO, see [Overview of Importing L3Out Configuration, on page 245](#).

Step 8 Deploy the template to sites.

After you have imported the policies and saved the template, you must deploy it back to the site.

Note

If the names of the imported objects that are used in NDO do not match the names of those objects in the APIC, NDO does not create new objects in the APIC and simply starts managing the original ones.

However, if you make other changes to a policy object before deploying it back to the site, NDO creates a new object in the APIC.

- a) In the **Tenant Policies** template view, click **Deploy**.
- b) In the **Deploy to sites** dialog, confirm the policies being deployed and click **Deploy**.

What to do next

After you've defined the policies in the Tenant Policy template, proceed to [Importing L3Out Objects, on page 251](#).

Importing L3Out Objects

This section describes how to import an L3Out template from an APIC site into Cisco Nexus Dashboard Orchestrator. For more information about each policy and how it related to policies and settings in other templates, see [Overview of Importing L3Out Configuration, on page 245](#).

Before you begin

- If you want to configure and deploy new L3Out configurations (greenfield deployment), see the earlier sections of this chapter instead.
- You must have created a Template Policy template and imported the policies that are associated with the L3Out you want to import, as described in [Importing Tenant Policy Template Objects, on page 248](#).

Procedure

Step 1 In the left navigation pane, choose **Configure > Tenant Templates > > L3Out**.

Step 2 In the main pane, click **Add L3Out Template**.

If you want to update an existing L3Out template instead, simply click its name. This opens the **L3Out Template** page.

Step 3 If you are creating a brand new template, choose the **Tenant** and **Site** from which you import the L3Out configuration, then click **Save and go to template**.

Each L3Out template is associated with a specific tenant similar to other NDO templates, however it is also assigned to a single site only as L3Out configuration is typically site-specific.

If you want to import L3Out configuration for multiple sites, you must create at least one L3Out template for each site, but you can import multiple L3Outs per site/tenant into the same template or you may choose to have multiple L3Out templates per site as long as they are assigned to different tenants.

Step 4 If you created a brand new template, provide the **Name** for the template and click **Save**.

You must save a brand new template before you can add new or import existing configuration.

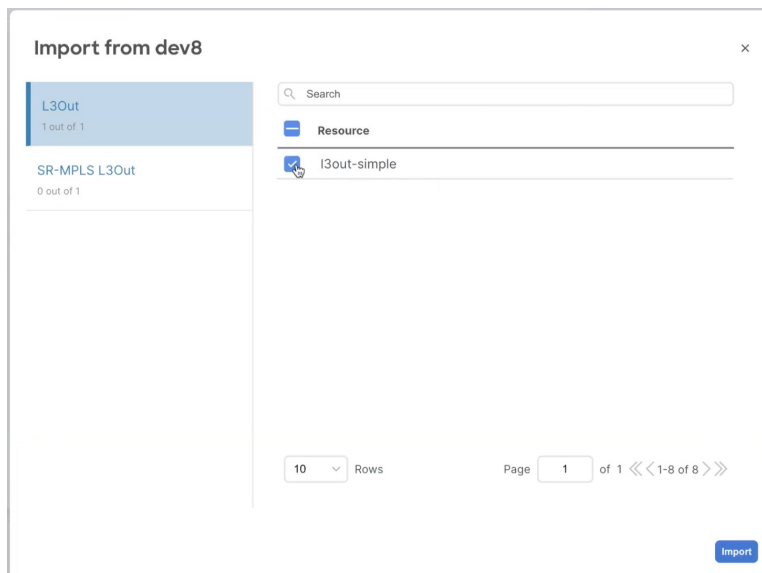
Step 5 Import an L3Out from the site.

a) In the main pane, click **Import**.

b) In the **Import from <site-name>** dialog, select the **L3Out** you want to import and click **Import**.

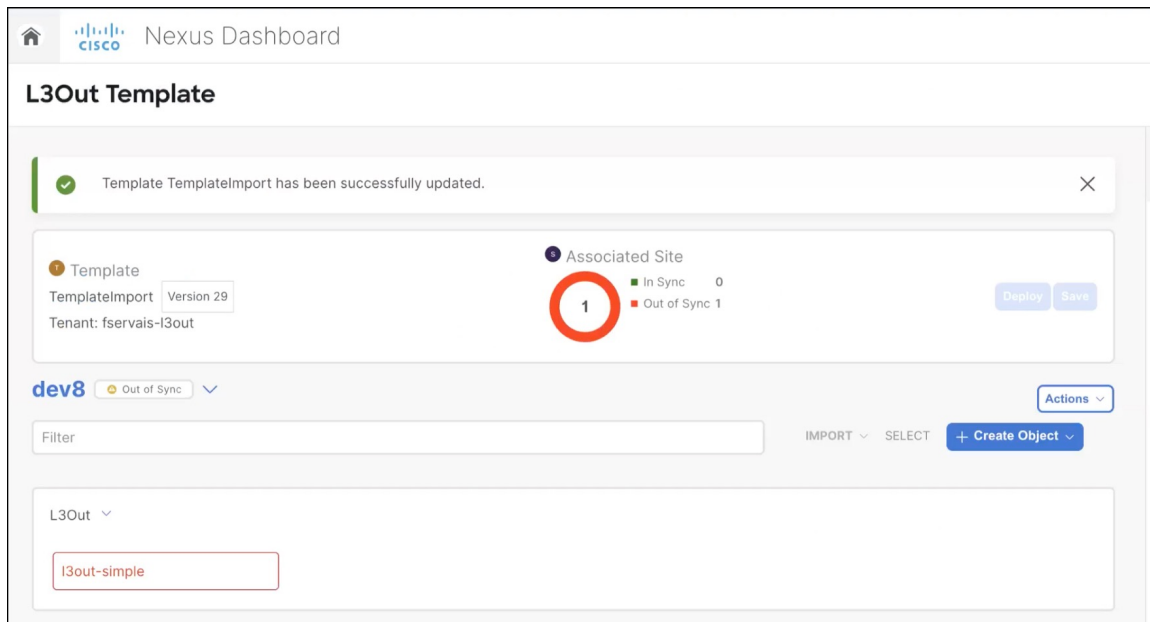
Note

If an L3Out has one or more tenant policy references that are not found in NDO's Tenant Policies templates, you cannot import that L3Out and must first import those references as described in [Importing Tenant Policy Template Objects](#), on page 248.



Step 6 Provide any missing information for the imported L3Out.

When you first import an L3Out, the object in the UI may be shown in red if some of the L3Out settings are not imported and must be provided manually:



For example, if BGP Peer configuration is present on the L3Out, NDO enforces authentication when the L3Out is imported. In this case, you must manually navigate to the authentication settings and either disable **Password Authentication** or provide a valid password:

- Select the imported L3Out.
- Click the setting that shows a warning.

c) Click the setting that shows a warning again.

d) Provide any missing configuration, such as a password.

e) Repeat this step for all other warnings in the template for the imported objects.

Step 7

Click **Save** to save the template changes.

Step 8

If necessary, update any previously imported IP SLA track members that reference the L3Out you imported in the previous step.

If you have imported one or more IP SLA track members in the previous section which reference the L3Out you are importing, you must manually update the track members' scopes and references after you've imported the L3Out. Other details about this behavior are described in [#unique_154 unique_154_Connect_42_sect_IP%20SLA_import](#).

- Ensure that you have saved the L3Out template with the imported L3Out objects.
- Navigate to **Application Management > Tenant Policies**.
- Choose the Tenant Policies template that contains the IP SLA track members.
- Choose the IP SLA Track List policy.
- In the right properties sidebar, click the **Edit** icon next to the Track Member List you want to update.
- In the **Update Track List to Track Member Relation** dialog, update the **Scope Type** and choose the scope object.

The current values are set to `Local Reference` and the name of the referenced object:

Update Track List to Track Member Relation ×

TrackMember

Destination IP *

10.0.0.1

Scope Type *

BD L3Out Local Reference

Local Reference

demo-tenant/l3out-2

IPSLA Monitoring Policy *

ipslaMonPol-1 ×

Ok

You must update the scope type to `L3Out` and then choose the L3Out you imported in the previous step.

- g) Click **Ok** to save the changes.
- h) Click **Save** to save the Tenant Policies template.
- i) Click **Deploy** to redeploy the template to the site.
- j) Return to **Application Management > Tenant Policies** and choose the L3Out template that you were editing in the previous step.

Step 9 Deploy the L3Out template to site.

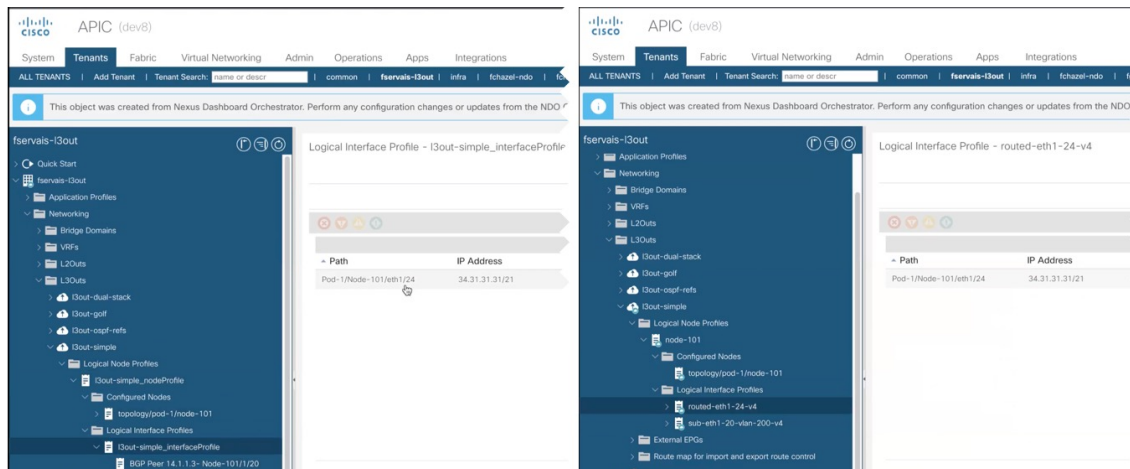
After you have imported the L3Out and saved the template, you must deploy it back to the site.

- a) In the **L3Out Template** page, click **Deploy**.
- b) In the **Deploy to sites** dialog, confirm the policies being deployed and click **Deploy**.
- c) (Optional) Verify that the policies were deployed correctly.

You can verify that the template was correctly deployed to a site by navigating to the site's APIC, choosing **Tenants > <tenant-name> > Networking > L3Outs** and checking that the L3Out name is consistent with the one you imported into the NDO template.

Note

When the configuration is deployed from NDO back to the site, the old MOs are removed and new ones are created with NDO-specific hierarchy, which may cause a brief (up to 1 second) traffic interruption:



Viewing L3Out Neighbors

Beginning with release 4.1(2), Cisco Nexus Dashboard Orchestrator provides a unified view of all L3Outs and their neighbors in your Multi-Site domain. This information provides visibility into operational data reported by the fabrics controller about site-level connectivity and simplifies troubleshooting by showing the various Layer 3 adjacencies (neighbors) for each L3Out.

Procedure

- Step 1** In the left navigation pane, choose **Operate > Sites**.
- Step 2** Click the name of the site for which you want to view L3Out neighbors.
- Step 3** In the site information page, choose **Connectivity > L3 Neighbors**.

The **L3 Neighbors** page provides a unified view of all neighbors based on L3Out configuration for that site. You can **Filter** or sort the page based on each column.

At any time, you can click **Refresh** to pull the latest information from the site's controller.

- Step 4** Click an entry in the **Neighbor** column to view that neighbor's details.

Here, you can view the **Local Switch** information (including its name, IP address, ASN, interface information, and so forth) and **Neighbor Details** (such as its IP address, ASN, route ID, port, and so forth).

For example, the following two figures show sample information for a BGP and OSPF L3Out neighbors:

BGP Neighbor Details							
Local Switch Details							
Name	Local IP	ASN	Interface Type	Interface	Router ID	Port	VRF
F2-P1-Leaf-304	10.110.2.2	65002	Routed Sub-interface	eth1/16	1.1.1.104	36597	L3-Demo:VRF
Authentication							
Disabled							
Neighbor Details							
Neighbor IP	ASN	Router ID	Port	Neighbor Status	Uptime		
10.110.2.3	65111	111.1.1.1	179	↑ Established	1 Weeks, 4 Days		

OSPF Neighbor Details

Local Switch Details

Name	Router ID	Interface Type	MTU	Interface	Encap	Interface IP Address	VRF
F2-P1-Leaf-304	1.1.1.104	SVI	1500	L303-304-VPC11	vlan-802	10.82.1.2	L3-Demo:VRF
OSPF Area							
backbone	Network Type		Interface Controls Enabled				
	Broadcast		-				

Neighbor Details

Neighbor ID	Interface IP Address	Neighbor Status	Uptime
1.1.1.103	10.82.1.1	<div>↑ Full/BDR</div>	1 Weeks

Step 5

If the displayed information is not accurate, verify L3Out configurations.

If the L3Out neighbors are not present in the table view:

- Verify that the L3Out policy is configured in NDO and deployed successfully. The information is displayed only for L3Outs that are configured in NDO.
- Verify that the L3Out neighbors are present in NDO's inventory using the APIs.
 - For BGP: `GET /mso/api/v1/inventorybgpneighbors?status.fabric=<site-id>`
 - For OSPF: `GET /mso/api/v1/inventoryospfneighbors?status.fabric=<site-id>`

If the L3Out neighbors' operational state is not green:

- Verify that the switch interfaces are not in the `shut` state on either of the switches.
- Verify that the protocol settings are configured correctly and there is no mismatch in the peer device configuration.
 - For BGP, check the authentication, eBGP MultiHop TTL, and ASN are configured correctly.
 - For OSPF, check authentication, Area ID, and MTU configurations.



CHAPTER 22

Intersite L3Out

- [Intersite L3Out Overview, on page 259](#)
- [Intersite L3Out Guidelines and Limitations, on page 260](#)
- [Configuring External TEP Pool, on page 261](#)
- [Configuring External EPG to Use Intersite L3Out, on page 262](#)
- [Creating a Contract for Intersite L3Out, on page 264](#)
- [Use Cases, on page 265](#)

Intersite L3Out Overview

NDO enables a number of scenarios in which endpoints located in one site are able to establish connectivity with entities, such as external network, mainframe, or service nodes, reachable through a remote L3Out.

These include the following:

- L3Out across sites—endpoints in an application EPG in one site using an L3Out in another site (both part of the same VRF).
- Intersite transit routing—establishing communication between entities (such as endpoints, network devices, service nodes) connected behind L3Outs deployed in different sites (both L3Outs part of the same VRF).
- Shared services for intersite L3Out—application EPG to remote L3Out or intersite transit routing across VRFs.

The following sections are divided into the generic GUI procedures you can follow to create the objects required to implement intersite L3Out use cases followed by overview and workflows specific to each supported use case scenario.



Note The term "intersite L3Out" refers to the functionality allowing communication to external resources reachable via the L3Out connection of a remote site. However, in this document, the term may also be used to indicate the specific remote L3Out object.

The following sections describe how to configure an intersite L3Out for EPG-to-L3Out use cases without Policy-Based Redirect (PBR). If you want to insert service chaining in the contract between an EPG and a remote L3Out to enable PBR, see [Intersite L3Out with PBR, on page 275](#) instead; and if you want to enable PBR between L3Outs in different sites (transit routing with PBR), see [Intersite Transit Routing with PBR, on page 283](#).

Intersite L3Out Guidelines and Limitations

When configuring intersite L3Out, you must consider the following:

- The steps described in the following sections assume you have L3Out connectivity already configured for your sites.

This includes creating the L3Out template, creating the L3Out object and defining its configuration, and deploying the configuration to the site(s). Detailed information on configuring L3Outs is available in the [External Connectivity \(L3Out\), on page 229](#) chapter.
- Intersite L3Out is supported for IPv4 and IPv6.
- With intersite L3Out, in addition to the BGP eVPN sessions that are always established between sites in Multi-Site topology, MP BGP VPNv4 (or VPNv6) sessions are created to support the intersite L3Out feature.
- You can now associate a bridge domain in one site with the L3Out in another site, however they must both be in the same VRF.

This association is performed at the site-local level and is required to advertise the BD subnet out of the remote L3Out and ensure that inbound traffic to the BD can be maintained even if the local L3Out failed.



Note However, instead of associating a bridge domain, we recommend that you define outbound route-maps for the L3Outs, including the BD prefixes that need to be advertised out.

- The Policy Control Enforcement direction for the VRF associated to the intersite L3Out must be kept configured in the default ingress mode.
- The following scenarios are not supported with intersite L3Out and remote leaf (RL):
 - Transit routing between L3Outs deployed on RL pairs associated to separate sites
 - Endpoints connected to a RL pair associated to a site communicating with the L3Out deployed on the RL pair associated to a remote site
 - Endpoints connected to the local site communicating with the L3Out deployed on the RL pair associated to a remote site

- Endpoints connected to a RL pair associated to a site communicating with the L3Out deployed on a remote site
- The following other features are not supported with intersite L3Out in Multi-Site:
 - Multicast receivers in a site receiving multicast from an external source via another site L3Out. Multicast received in a site from an external source is never sent to other sites. When a receiver in a site receives multicast from an external source it must be received on a local L3Out.
 - An internal multicast source sending multicast to an external receiver with PIM-SM any source multicast (ASM). An internal multicast source must be able to reach an external Rendezvous Point (RP) from a local L3Out
 - GOLF
 - Preferred Groups for External EPG

Configuring External TEP Pool

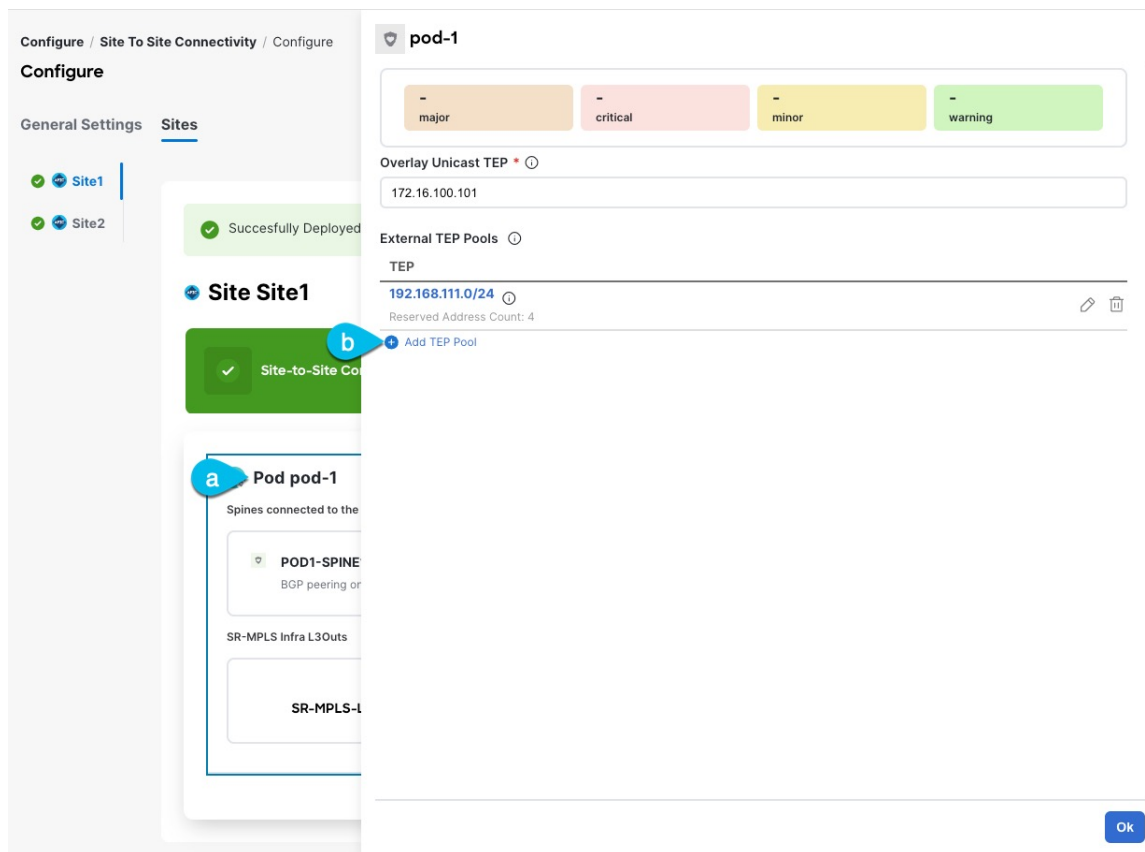
Intersite L3Out requires an external TEP address for the border leaf switches in each pod. If you already have an external TEP pool that is configured, for example for another feature such as Remote leaf switch, the same pool can be used. The existing TEP pool will be inherited by the Cisco Nexus Dashboard Orchestrator and shown in the GUI as part of the infra configuration. Otherwise, you can add a TEP pool in the GUI, as described in this section.



Note Every pod must be assigned a unique TEP pool and it must not overlap with any other TEP pool in the fabric.

Procedure

-
- Step 1** Log in to your Cisco Nexus Dashboard Orchestrator.
 - Step 2** In the left navigation menu, select **Configure > Site To Site Connectivity**.
 - Step 3** In the top right of the main pane, click **Configure**.
 - Step 4** In the main pane, choose the **Sites** tab and then the site for which you want to define an external TEP pool.
 - Step 5** In the main pane, click on the name of the Pod that you want to configure and then click **+Add TEP Pool**.



Step 6 In the **Add TEP Pool** window, specify the **External TEP Pool** that you want to configure for that site and the **Reserved Address Count**.

For the TEP pool, provide the subnet and the subnet mask, for example 192.168.111.0/24.

Note

You must ensure that the TEP pool you are adding does not overlap with any other TEP pools or fabric addresses.

Multiple disjointed TEP pools can be configured , so you are not required to specify a large TEP pool from the beginning.

Step 7 Repeat the process for each site and pod where you plan to use intersite L3Outs.

Configuring External EPG to Use Intersite L3Out

This section describes how to create an external EPG that will be associated to the intersite L3Out. You can then use this external EPG and contracts to configure specific use cases for endpoints in one site to use an L3Out in another site or to configure L3Out-to-L3Out transit routing.

Before you begin

- Ensure that you have read and completed the requirements described in [Intersite L3Out Guidelines and Limitations](#), on page 260.

- You must have created and deployed the L3Outs in each site as part of configuring the site's external connectivity.

Detailed information on configuring L3Outs is available in the [External Connectivity \(L3Out\)](#), on page 229 chapter.

Procedure

Step 1 Select the schema and template where you want to create the external EPG.

If you create the external EPG in a template that is associated to multiple sites, the external EPG will be stretched across all of those sites. This is recommended when the L3Outs defined in those sites provide access to a set of common external resources, for example the WAN.

If you create the external EPG in a template that is associated with a single site, the external EPG will be created in that site only. This is recommended when the L3Out defined in that site provides access to external resources accessible only from that site.

Step 2 Create an external EPG.

- a) In the main pane, select **+Create Object > External EPG**.
- b) Provide the **Display Name** for the external EPG.

For example, `eepg-intersite-l3out`

- c) From the **Virtual Routing & Forwarding** dropdown, select the same VRF you associated with the L3Out.

Step 3 Map the external EPG to the L3Out.

- a) In the template view, select the tab for the site where the external EPG is deployed.
- b) Select the external EPG you created in the previous step.
- c) In the **<external-epg-name> on <site-name>** properties sidebar, choose the L3Out you created from the **L3Out** dropdown.

Note that both the APIC-managed and the Orchestrator-managed L3Outs are available for selection. You can select either the L3Out you have created and deployed from NDO or pick an L3Out that exists in the site's APIC.

Step 4 Configure one or more subnets for the external EPG.

- a) Select the external EPG.
- b) In the right sidebar, click **+Add Subnet**.
- c) In the **Add Subnet** window's **Subnet** field, provide the subnet's network prefix.
- d) (Optional) Provide a descriptive **Name** for the subnet.
- e) Provide any required options for this subnet.

The prefixes and options you configure depend on the specific use cases:

- To classify the inbound traffic as belonging to the external EPG, select the **External Subnets for External EPG** flag for the specified prefix. Depending on the specific use case, this allows you to apply a contract with an internal EPG or with the external network domain reachable via a different L3Out.
- To advertise the external prefixes learned from another L3Out (in the same site or in a remote site) out of this L3Out, select the **Export Route Control** flag for the specified prefix. When specifying the `0.0.0.0/0` prefix, the **Aggregate Export** flag can be selected to advertise all prefixes out of the L3Out; if the **Aggregate Export**

flag is not enabled, only the default route 0.0.0.0/0 would be advertised, if present in the routing table of the border leaf nodes.

Note

However, we recommend that you use the Outbound Route-Map associated to the L3Out to match the external prefixes (in addition to the BDs' subnets) to be advertised out instead.

- To filter out specific routes received from the external network, select the **Import Route Control** flag for the specified prefix. If specifying the 0.0.0.0/0, you can also choose the **Aggregate Import** option.

Note that this is possible only when peering BGP with the external routers.

Note

However, similar to the previous bullet point, we recommend that you use the Outbound Route-Map associated to the L3Out to match the external prefixes (in addition to the BDs' subnets) to be advertised out instead.

- To leak routes to different VRFs, select the **Shared Route Control** and the associated **Aggregate Shared Routes** flags, as well as the **Shared Security Import** flag. These options are required for the specific use case of inter-VRF shared L3Out and inter-VRF intersite transit routing.

Step 5 (Optional) Enable **Include in Preferred Group** if you want the external EPG to be part of the EPG preferred group.

Step 6 (Optional) From the **QoS Level** dropdown, select the QoS level for this external EPG.

For additional information about QoS in ACI fabrics, see [Cisco APIC and QoS](#).

For additional information about configuring QoS Levels in your Nexus Dashboard Orchestrator, see [QoS Preservation Across IPN, on page 305](#).

Creating a Contract for Intersite L3Out

This section describes how to create a filter and a contract you will use to enable communication between an application EPG deployed in a site and the external EPG associated to an L3Out in a different site (intersite L3Out functionality).

Procedure

Step 1 Select the template where you want to create contract and filter.

You can use the same schema and template where you created the VRF and the external EPG or you can choose a different schema and template.

Note

In earlier NDO releases, even if the contract and filters were defined only as local objects in one site (*Site1*), NDO created the corresponding shadow objects in a remote site (*Site2*) when a local EPG or external EPG in *Site2* needed to consume or provide that contract.

This is no longer the case and you must define the contracts and the filters explicitly on all sites where they will be used.

Step 2 Create a filter.

- a) In the main pane, select **+Create Object > Filter**.
- b) Provide the **Display Name** for the filter.
- c) Click **+ Entry** and provide the filter entry information specific to the kind of traffic you want to allow.
- d) Click **Ok** to save the filter.

Step 3

Create a contract

- a) In the main pane, select **+Create Object > Contract**.
- b) In the right pane, provide the **Display Name** for the contract
- c) Select the appropriate **Scope** for the contract.
 - If both intersite L3Out and application EPG are in the same VRF, set the scope to `vrf`.
 - If the intersite L3Out and application EPG are in different VRFs but the VRFs are in the same tenant, set the scope to `tenant`.
 - If the intersite L3Out and application EPG are in different VRFs and the VRFs are in different tenants, set the scope to `global`.
- d) Ensure that the **Apply both directions** option is enabled if you want the same filter to apply for both consumer-to-provider and provider-to-consumer directions.

With this option enabled, you need to provide the filters only once and they will apply for traffic in both directions.
- e) In the **Filter Chain** area, click **Create Filter** and choose the filter you created in the previous step.
- f) Click **Ok** to save the contract.

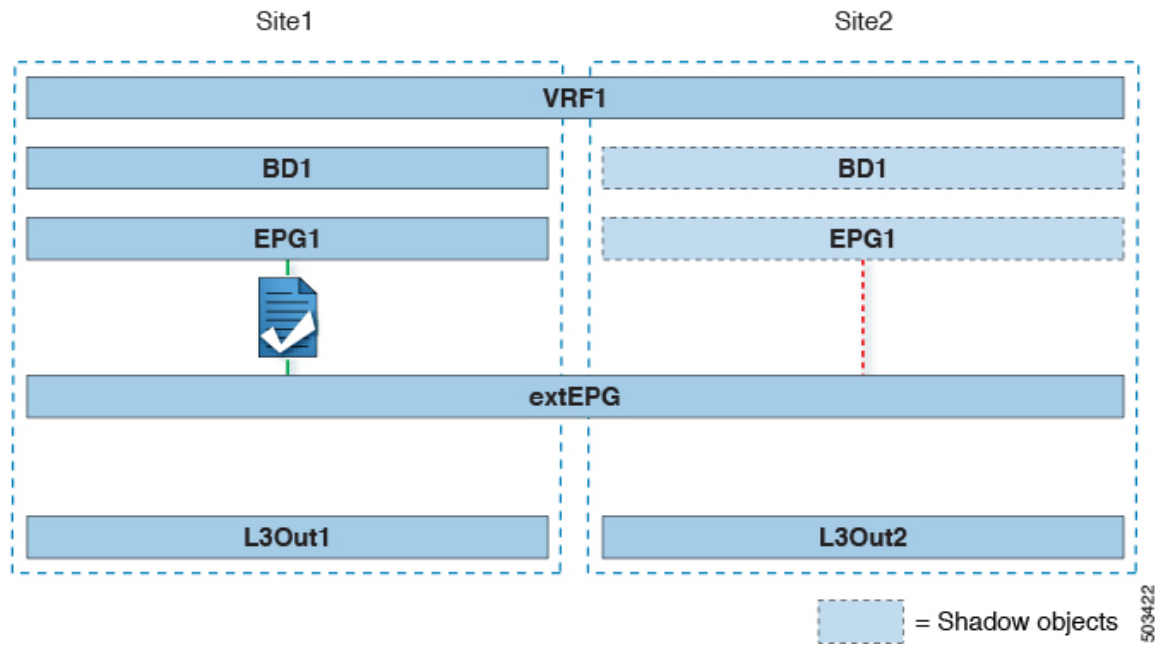
Use Cases

Intersite L3Out for Application EPGs (Intra-VRF)

This section describes the configuration that is required to allow endpoints that are part of an application EPG to communicate with the external network domain reachable through an L3Out deployed in another site but within the same VRF (intra-VRF).

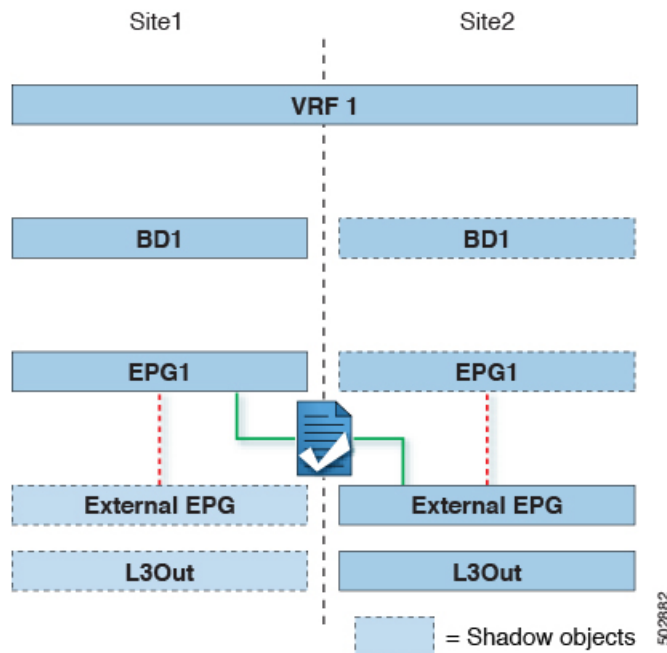
The first figure below shows a stretched external EPG and the associated L3Outs which will be created in both sites. An application EPG (`EPG1`) is created in Site 1 and has a contract with the external EPG. This use case is recommended when the L3Outs in the separate sites provide access to a common set of external resources. It simplifies the policy definition and external traffic classification, while still allowing you to apply route-map policies separately on each L3Out for the independent APIC domains.

Figure 23: Stretched External EPG



The second figure below shows a similar use case but with the external EPG being deployed to only the site where the physical L3Out is located. The application EPG and the contract are configured in the same exact way to allow the traffic flow between the EPG in one site and the physical L3Out in the other.

Figure 24: Non-Stretched (Site-Local) External EPG



The following steps describe the configuration that is required to implement the use case shown in Figure 1, which represents the most common scenario. If you want to deploy the use case shown in Figure 2, you can adapt the procedure with minor changes.

Before you begin

You need to have the following already configured:

- External connectivity (L3Out) in each site, as described in the [External Connectivity \(L3Out\), on page 229](#) section.

In this use case, a separate L3Out will be imported or created in each site-specific template.

- A schema with four templates.

Create a template for each site (for example, `template-site1` and `template-site2`) where you configure the objects unique to those sites, such as the application EPG and the L3Outs.

In addition, create two more stretched templates: one for the stretched EPGs, External EPGs, and BDs and the second one for the VRFs, contracts, and filters.

- The external EPG for the intersite L3Out, as described in [Configuring External EPG to Use Intersite L3Out, on page 262](#).

In this use case, the external EPG is configured as a stretched object that is defined in one of the stretched template (for example, `template-stretched-ext-epg`). Assuming that the external EPG provides access to the entire external address space, we recommend configuring a `0.0.0.0/0` prefix for classification to avoid specifying a long list of more specific prefixes.

- The contract that you will use between the application EPG and the L3Out external EPG, as described in [Creating a Contract for Intersite L3Out, on page 264](#).

We recommend creating the contract and the filter in the second stretched template (for example, `template-stretched-vrf-contract`), which also contains the VRF.

Procedure

-
- Step 1** Log in to your Cisco Nexus Dashboard Orchestrator.
- Step 2** Select the schema and template for the application EPG and bridge domain.
- In this use case, you associate the template to `Site1`.
- Step 3** Configure an application EPG and its bridge domain belonging to the same VRF as the L3Out.
- If you already have an EPG that will use the intersite L3Out, you can skip this step.
- You can create a new or import an existing EPG and bridge domain as you typically would.
- Step 4** Assign the contract to the application EPG.
- Select the EPG.
 - In the right sidebar, click **+Contract**.
 - Select the contract that you created in previous section and its type.
- You can choose whether the application EPG is the `consumer` or the `provider`.
- Step 5** Assign the contract to the external EPG mapped to the remote L3Out.
- Select the `template-stretched` where the external EPG is located.
 - Select the external EPG.
 - In the right sidebar, click **+Contract**.

- d) Select the contract that you created in previous section and its type.

If you chose the application EPG to be the `consumer`, choose `provider` for the external EPG. Otherwise, choose `consumer` for the external EPG.

Step 6 Associate the application EPG's bridge domain with the L3Out.

This enables the BD subnet to be advertised out of the L3Out toward the external network domain. Note that one or more subnets associated to the BD must be configured with the **Advertised Externally** option to be advertised out of the L3Out

- a) In the left sidebar, under **Sites**, select the application EPG's template.
- b) Select the bridge domain associated with the application EPG.
- c) In the right sidebar, click **+L3Out**.
- d) Select the intersite L3Out you created.

For the use case shown in Figure 1, associate the BD to both the L3Outs defined in Site 1 and Site 2 to ensure that the external network can have access to the EPG from both paths. Specific policies can be associated to the L3Out or to the external routers to ensure that a specific L3Out path is normally preferred for inbound traffic. We recommend this when the EPG and BD are local to a site (as in the specific example) to avoid suboptimal inbound traffic path through the remote site's L3Out.

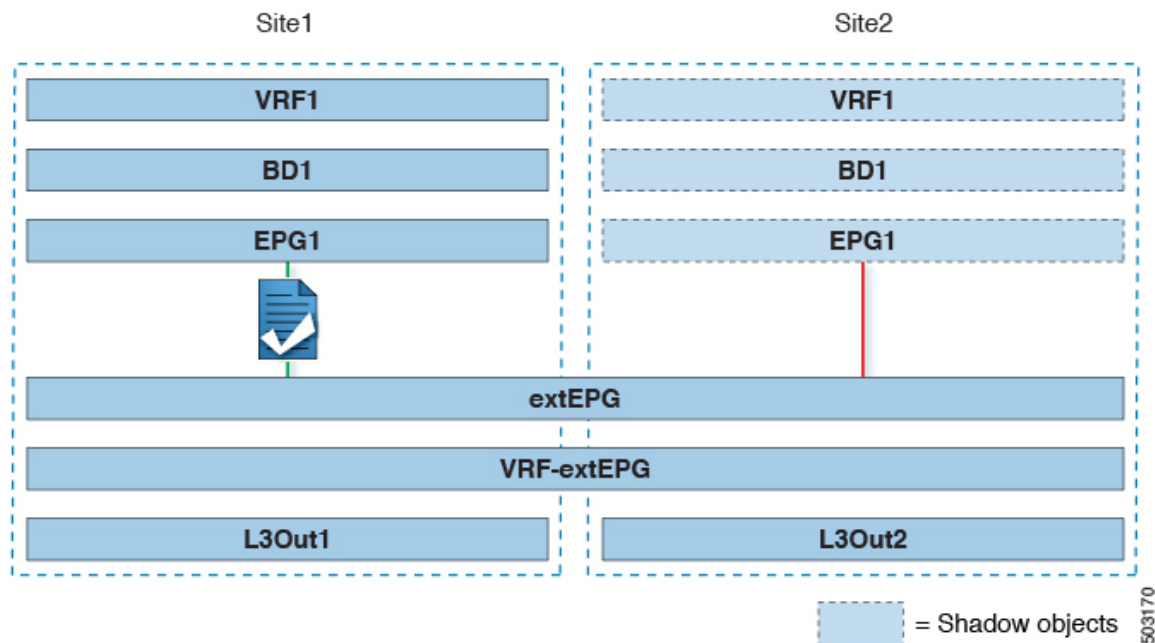
Step 7 Deploy the schema.

Shared Services with Intersite L3Out for Application EPGs (Inter-VRF)

This section describes the configuration that is required to allow endpoints that are part of an application EPG in one VRF to communicate with the external network domain reachable through an L3Out deployed in another site and different VRF, this is also known as "Shared Services".

This scenario is recommended when the L3Outs in separate sites provide access to a common set of external resources. It simplifies the policy definition and external traffic classification, while still allowing you to apply route-map policies separately on each L3Out for the independent APIC domains.

Figure 25: Stretched External EPG, Site-Local L3Outs and Application EPGs



The following steps describe the configuration that is required to implement the use case shown in Figure 3.

Before you begin

You must have the following already configured:

- External connectivity (L3Out) in each site, as described in the [External Connectivity \(L3Out\)](#), on page 229 section.

In this use case, a separate L3Out will be imported or created in each site-specific template.

- A schema with four templates.

Create a template for each site (for example, `template-site1` and `template-site2`) where you configure the objects unique to those sites, such as the application EPG and the L3Outs.

In addition, create two more stretched templates: one for the stretched External EPG and the second one for the contracts, and filters.

- The external EPG for the intersite L3Out, as described in [Configuring External EPG to Use Intersite L3Out](#), on page 262.

In this use case, the external EPG is configured as a stretched object that is defined in the stretched template (`template-stretched`). Assuming that the external EPG provides access to the entire external address space, we recommend configuring a `0.0.0.0/0` prefix for classification to avoid specifying a long list of more specific prefixes.

For this specific shared service, use case, you are also required to enable the **Shared Route Control** and the **Shared Security Import** flags for one or more subnets that are associated to one or more external EPGs of the remote L3Out. If you are using the `0.0.0.0/0` prefix for classification on the external EPG, in addition to the **Shared Route Control** flag, also enable the **Aggregate Shared Routes** flag.

- The contract that you will use between the application EPG and the L3Out external EPG, as described in [Creating a Contract for Intersite L3Out, on page 264](#).

We recommend creating the contract and the filter in the stretched template (`template-stretched`).

Procedure

Step 1 Log in to your Cisco Nexus Dashboard Orchestrator.

Step 2 Select the schema and template for the application EPG and bridge domain.

In this use case, you associate the template to `Site1`.

Step 3 Configure an application EPG and its bridge domain belonging to a separate VRF from the L3Out's.

If you already have an EPG that will use the intersite L3Out, you can skip this step.

You can create a new or import an existing EPG and bridge domain as you typically would.

Step 4 Assign the contract to the application EPG.

- Select the EPG.
- In the right sidebar, click **+Contract**.
- Select the contract that you created in previous section and its type.

You can choose whether the application EPG is the `consumer` or the `provider`.

Note

If the application EPG is configured as `provider`, you must configure the subnet that is already defined under the BD also under the EPG to leak that route into the L3Out VRF. The same flags that are used under the BD for the subnet should also be set under the EPG. In addition to that, for the subnet under the EPG the flag **No default SVI Gateway** should also be enabled, since the default gateway function is enabled at the BD level.

Step 5 Assign the contract to the external EPG mapped to the L3Outs.

- Select the `template-stretched` where the external EPG is located.
- Select the external EPG.
- In the right sidebar, click **+Contract**.
- Select the contract that you created in previous section and its type.

If you chose the application EPG to be the `consumer`, choose `provider` for the external EPG. Otherwise, choose `consumer` for the external EPG.

Step 6 Associate the application EPG's bridge domain with the L3Out.

This enables the BD subnet to be advertised out of the L3Out toward the external network domain. The subnet(s) associated to the BD must be configured with the **Advertised Externally** option to be advertised out of the L3Out

- In the left sidebar, under **Sites**, select the application EPG's template.
- Select the bridge domain associated with the application EPG.
- In the right sidebar, click **+L3Out**.
- Select the intersite L3Out you created.

For the use case shown in Figure 1, associate the BD to both the L3Outs defined in Site 1 and Site 2 to ensure that the external network can have access to the EPG from both paths. Specific policies can be associated to the L3Out

or to the external routers to ensure that a specific L3Out path is normally preferred for inbound traffic. We recommend this when the EPG and BD are local to a site (as in the specific example) to avoid suboptimal inbound traffic path through the remote site's L3Out.

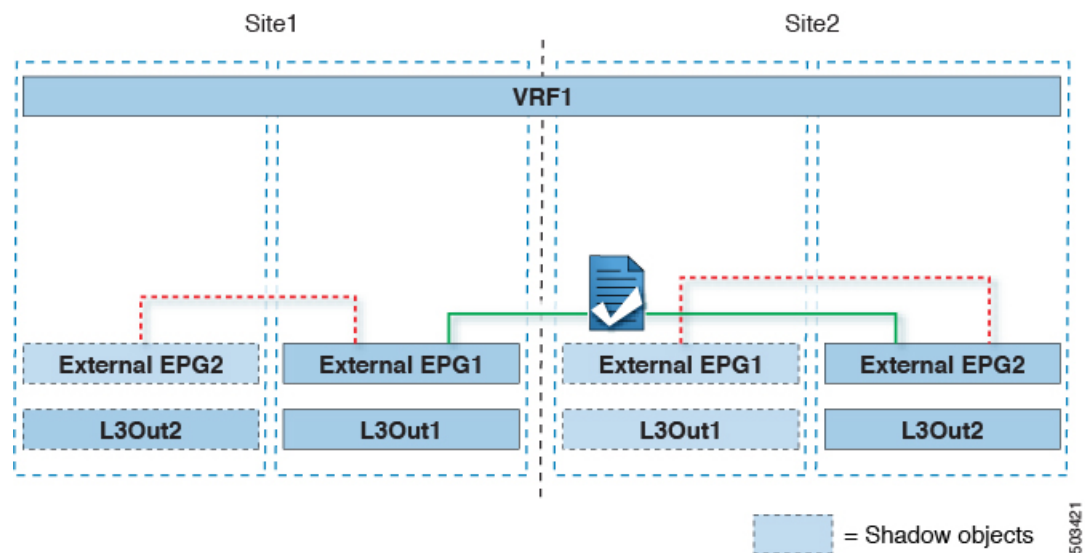
Step 7 Deploy the schema.

Intersite Transit Routing

This section describes the use cases where the Multi-Site domain acts as a distributed router allowing communication between entities (endpoints, network devices, service nodes, and so forth) connected behind L3Outs deployed in different sites, a functionality normally saw as intersite transit routing. The intersite transit routing is supported for intra-VRF and inter-VRF use cases.

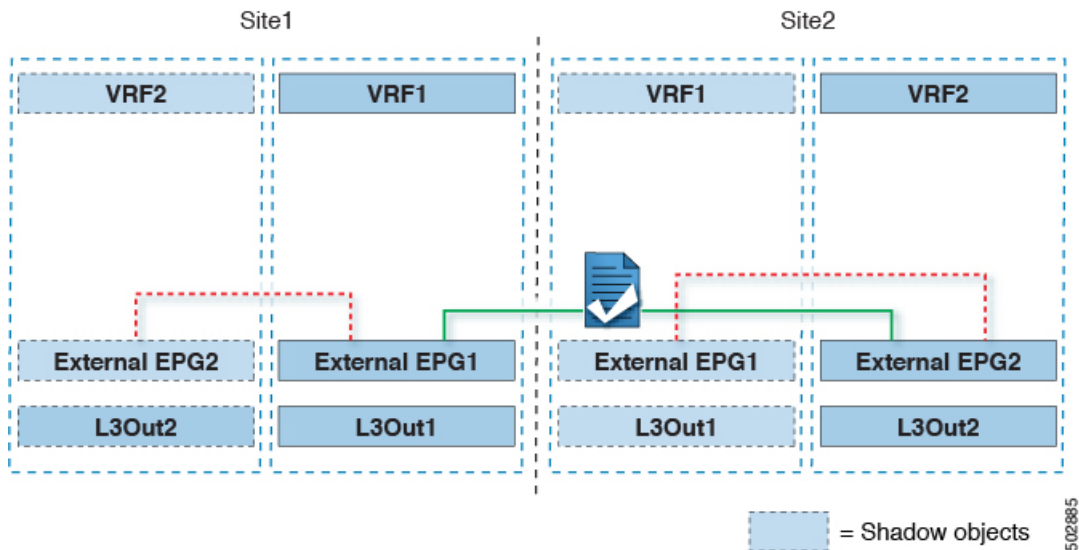
The figure below shows two L3Outs (L3Out1 and L3Out2) configured in different sites. Each L3Out is associated with a respective external EPG (External EPG1 and External EPG2). A contract between the two external EPGs allows communication between entities that are connected behind two different L3Outs in two different sites.

Figure 26: Intra-VRF Intersite Transit Routing



A similar configuration can be used when each site's L3Outs are in different VRFs.

Figure 27: Inter-VRF Intersite Transit Routing



The figures above show the two scenarios where the external EPGs and associated L3Outs are deployed as site-local objects; intersite transit routing can support all the combinations where neither external EPG is stretched, one of them is stretched, or both are stretched between sites.

When deploying intersite transit routing, the assumption is that the different external EPGs defined across sites are providing access to different external address spaces (not overlapping). A couple of options are hence possible for the configuration of the prefix that is used for classification:

- Define the `0.0.0.0/0` prefix for one of the external EPGs and specific prefixes on the other.

The external prefixes that are received on `L3Out1` must be advertised out of `L3Out2` and conversely.

- Define specific prefixes for each external EPG. In this case, you must ensure that the prefixes are not overlapping to avoid a fault from being raised by the site's APIC when the shadow external EPG is created in that site for a contract between the local and remote external EPGs.

When using specific prefixes, the same prefixes that are configured for classification on `External EPG1` must be configured with the **Export Route Control** flag set on `External EPG2` and conversely.



Note No matter which of the two classifications approaches you deploy, for the inter-VRF scenario you must also set the **Shared Route Control** (in addition to **Aggregate Shared Routes** if using `0.0.0.0/0`) and the **Shared Security Import** flags.

Before you begin

You must have the following already configured:

- External connectivity (L3Out) in each site, as described in the [External Connectivity \(L3Out\)](#), on page 229 section.

In this use case, a separate L3Out will be imported or created in each site-specific template.

- A schema with three templates.

Create a template for each site (for example, `template-site1` and `template-site2`) where you configure the objects unique to that site, such as the application EPGs and the L3Outs. In addition, create a separate template (for example, `template-stretched`) that you use for the stretched objects, which in this case will be the external EPG.

- Two different external EPGs for two different L3Outs in different sites. You can use the same procedure to create both external EPGs, as described in [Configuring External EPG to Use Intersite L3Out, on page 262](#).
- The contract that you use between the L3Out external EPGs defined in each site, as described in [Creating a Contract for Intersite L3Out, on page 264](#).

We recommend creating the contract and the filter in the stretched template (`template-stretched`).

Procedure

-
- Step 1** Log in to your Cisco Nexus Dashboard Orchestrator.
- Step 2** From the left navigation pane, select **Configure > Tenant Template > Applications > Schemas**.
- Step 3** Assign the contract to one of the external EPGs.
- Select the schema and template where the external EPG is located.
 - Select the external EPG.
 - In the right sidebar, click **+Contract**.
 - Select the contract that you created in previous section and its type.
Choose `consumer` or `provider`.
- Step 4** Assign the contract to the other external EPG.
- Select the schema and template where the external EPG is located.
 - Browse to the template where the external EPG is located.
 - Select the external EPG.
 - In the right sidebar, click **+Contract**.
 - Select the contract that you created in previous section and its type.
Choose `provider` or `consumer`.
- Step 5** Deploy the templates to appropriate sites.
-



CHAPTER 23

Intersite L3Out with PBR

- [Intersite L3Out with PBR, on page 275](#)
- [Guidelines and Limitations, on page 279](#)
- [Create Service Device Template, on page 280](#)
- [Add Service Chaining to Contract, on page 282](#)

Intersite L3Out with PBR

Cisco Application Centric Infrastructure (ACI) policy-based redirect (PBR) enables traffic redirection for service appliances, such as firewalls or load balancers, and intrusion prevention system (IPS). Typical use cases include provisioning service appliances that can be pooled, tailored to application profiles, scaled easily, and have reduced exposure to service outages. PBR simplifies the insertion of service appliances by using contract between the consumer and provider endpoint groups even if they are all in the same virtual routing and forwarding (VRF) instance.

PBR deployment consists of configuring a route redirect policy and a cluster redirect policy, and creating a service graph template that uses these policies. After the service graph template is deployed, you can attach it to a contract between EPGs so that all traffic following that contract is redirected to the service graph devices based on the PBR policies you have created. Effectively, this allows you to choose which type of traffic between the same two EPGs is redirected to the L4-L7 device, and which is subject to a security policy applied at the fabric level.

More in-depth information specific to services graphs and PBR is available in the [Cisco APIC Layer 4 to Layer 7 Services Deployment Guide](#)

Configuration Workflow

The use cases described in the following sections are an extension of a basic intersite L3Out (without PBR) use case which is in turn an extension on basic external connectivity (L3Out) configuration in each site. The workflow to configure the supported use cases is the same, with the only differences being whether you create the objects in the same or different VRFs (intra-VRF vs inter-VRF) and where you deploy the objects (stretched vs non-stretched).

1. Configure basic external connectivity (L3Out) for each site.

The intersite L3Out with PBR configuration described in the following sections is built on top of existing external connectivity (L3Out) in each site. If you have not configured an L3Out, create and deploy one as described in the [External Connectivity \(L3Out\), on page 229](#) chapter before proceeding with the following sections.

2. Configure an intersite L3Out use case **without** PBR.

We recommend configuring a simple intersite L3Out use case without any policy-based redirection before adding service chaining to it. This is described in detail in the [Intersite L3Out, on page 259](#) chapter.

3. Add service chaining to the L3Out contract as described in the following sections, which includes:

- Adding an external TEP pool for each Pod in each site where intersite L3Out is deployed.
- Creating a Service Device template and assigning it to sites.

The service device template must be assigned to the same sites as the L3Out and application templates that contain other configuration objects.

- Providing site-level configurations for the Service Device template.

Each site can have its own service device configuration including different high-availability models (such as active/active, active/standby, or independent service nodes).

- Associating the service device you defined to the contract used for the basic intersite L3Out use case you deployed in the previous step.

Supported Use Cases

The following diagrams illustrate the traffic flows between an ACI internal endpoint in application EPG and an external endpoint through the L3Out in another site in the supported intersite L3Out with PBR use cases.

Intra-VRF vs Inter-VRF

When creating and configuring the application EPG and the external EPG, you will need to provide a VRF for the application EPG's bridge domain and for the L3Out. You can choose to use the same VRF (intra-VRF) or different VRFs (inter-VRF).

When establishing a contract between the EPGs, you will need to designate one EPG as the provider and the other one as the consumer:

- When both EPGs are in the same VRF, either one can be the consumer or the provider.
- If the EPGs are in different VRFs, the external EPG must be the provider and the application EPG must be the consumer.

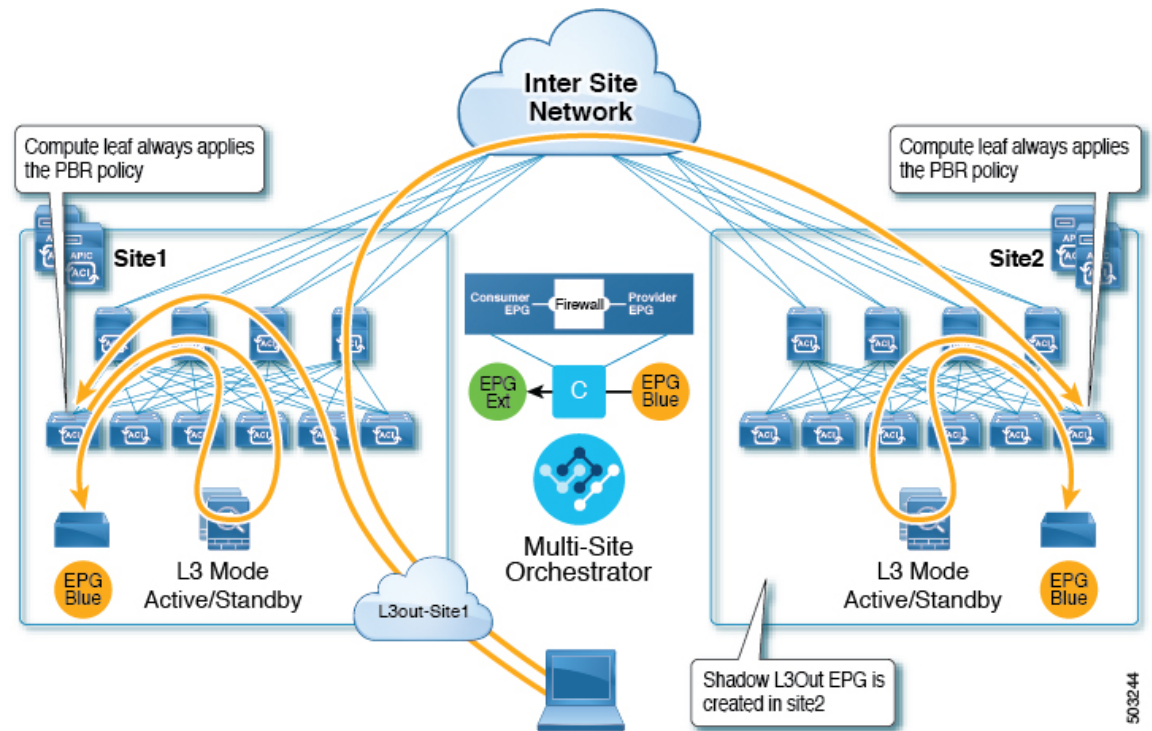
L3Out to Stretched EPG

This use case illustrates a single application EPG that is stretched between two sites and a single L3Out created in only one of the sites. Regardless of whether the application EPG's endpoint is in the same site as the L3Out or the other site, traffic will go through the same L3Out. However, the traffic will always go through the service node that is local to the endpoint's site because for North-South traffic the PBR policy is always applied only on the compute leaf nodes (and not on the border leaf nodes).



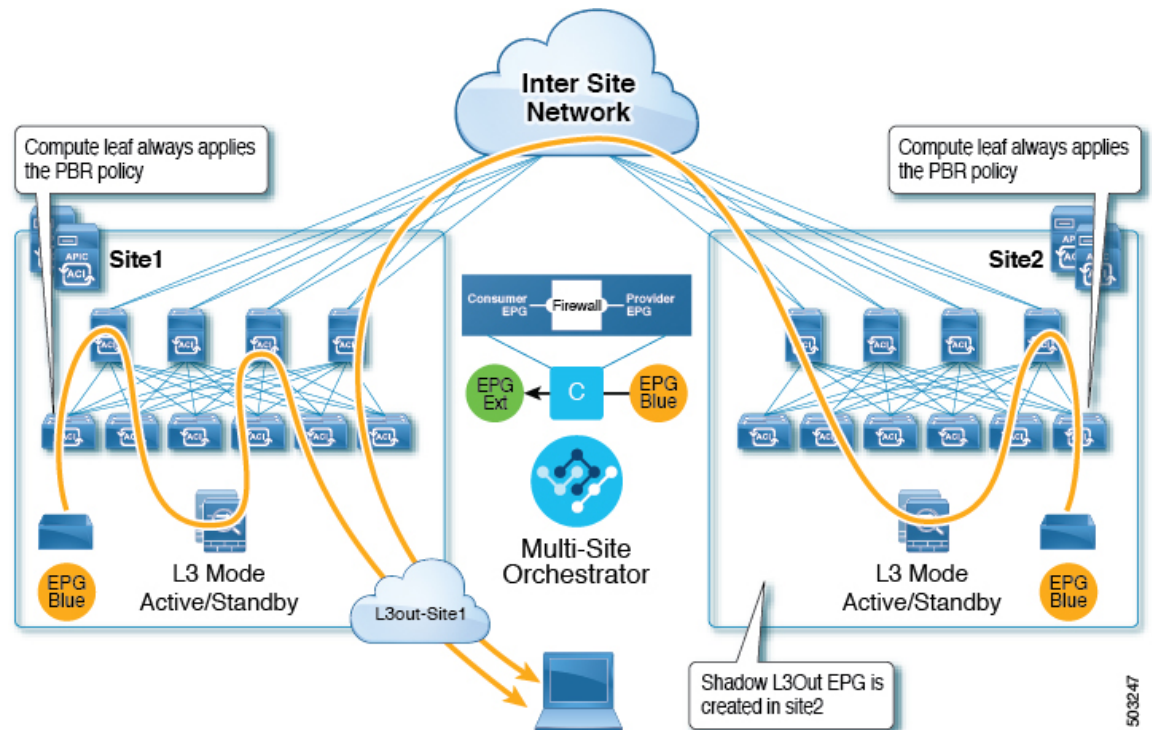
Note The same flow applies in cases when the external EPG is stretched and each site has its own L3Out, but the L3Out in the site where the traffic is originating or is destined to is down.

Figure 28: Inbound Traffic



503244

Figure 29: Outbound Traffic



503247

L3Out to Site-Local EPG

This use case illustrates a site-local application EPG that will use the L3Out in the other site for North-South traffic. Like in the previous example, all traffic will use the EPG's site-local service graph device.



Note The same flow applies in cases where the external EPG is stretched and each site has its own L3Out, but the EPG's local L3Out is down.

Figure 30: Inbound Traffic

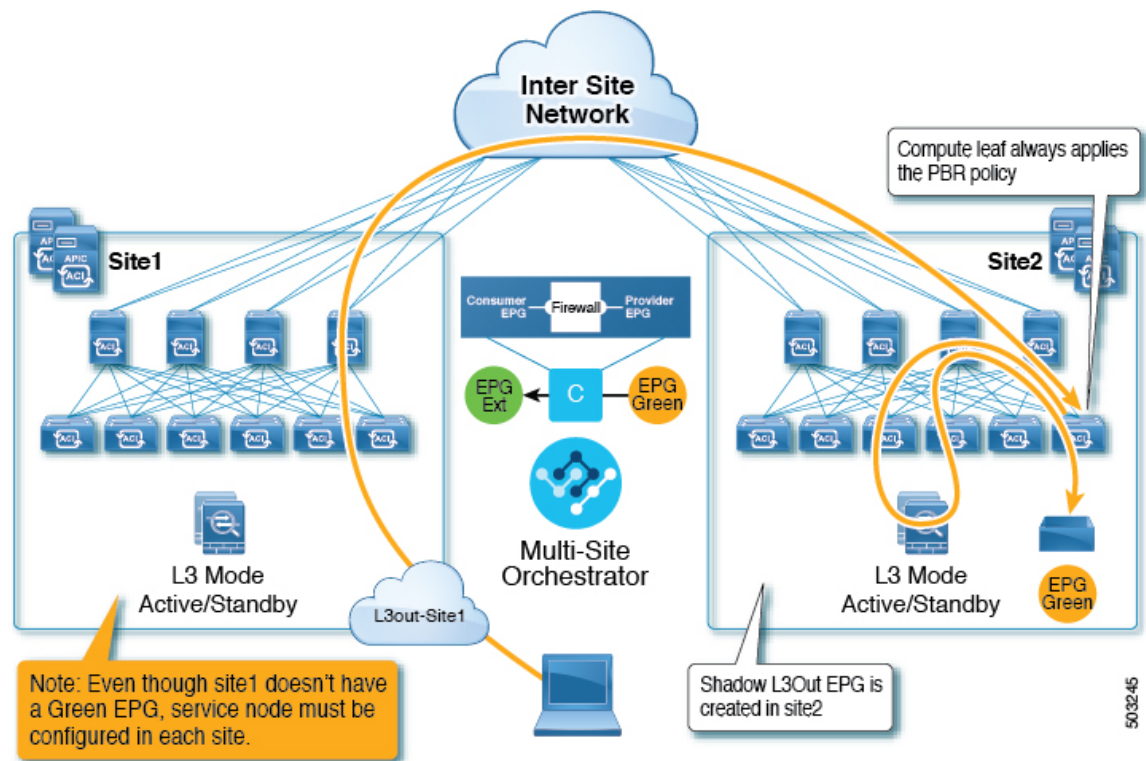
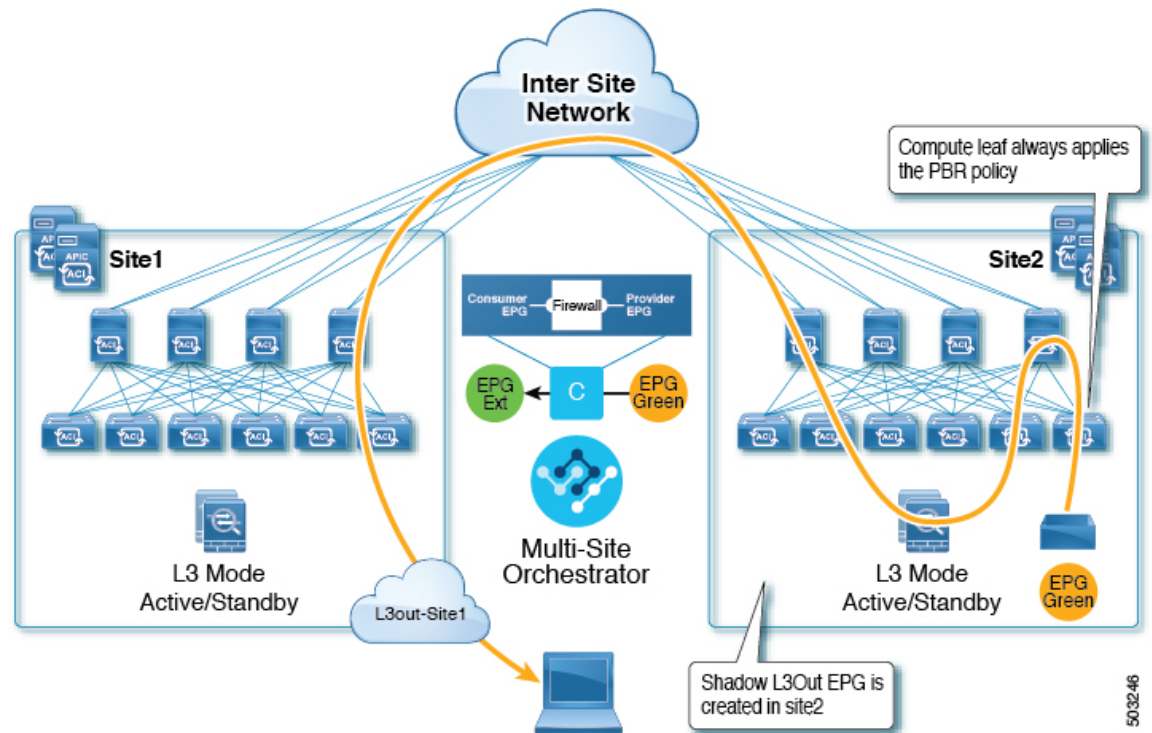


Figure 31: Outbound Traffic



5032-46

Guidelines and Limitations

When configuring an Intersite L3Out with PBR, the following restrictions apply:

- When configuring MultiSite Policy-Based Redirect (PBR) between Endpoint Groups (EPGs), the following features are not supported for specific IP endpoints or host prefixes (/32 for IPv4 and /128 for IPv6):
 - Static Route on a Bridge Domain (NH Reachability)
 - Microsoft Network Load Balancing
 - Anycast MAC
- For intersite L3Out with PBR, the following use cases are supported:
 - Inter-VRF intersite L3Out with the application EPG as the `consumer`.
For inter-VRF contracts, the External EPG associated to the L3Out must be the `provider`.
This use case is supported for sites running Cisco APIC release 4.2(5) and later or release 5.1(x), it is not supported for APIC release 5.0(x).
 - Intra-VRF intersite L3Out with the application EPG as either the `provider` or the `consumer`.
This use case is supported for sites running Cisco APIC release 4.2(5) and later or release 5.1(x), it is not supported for APIC release 5.0(x).
- For EPG-to-L3Out use cases, the application EPG can be stretched or site-local.

- For EPG-to-L3Out use cases, both one-arm and two-arm deployment models are supported; for L3Out-to-L3Out use case, only one-arm firewall devices are supported.

In one-arm deployment, both the inside and outside interfaces of the service graph are connected to the same bridge domain. In two-arm deployments, the service graph interfaces are connected to separate BDs.

- For EPG-to-L3Out use cases, when configuring a load balancer with PBR, the load balancer and the real servers for the virtual IP (VIP) must be in the same site. If PBR is disabled, the load balancer and the real servers can be in different sites.

L3Out-to-L3Out case does not support load balancers.

- You must have the basic use case of intersite L3Out already configured before you insert a service device by enabling service chaining on the contract that is already configured between an L3Out in one site and an EPG in another site or between two L3Outs in different sites.

Detailed instructions on deploying an intersite L3Out without PBR are described in the [Intersite L3Out, on page 259](#) chapter.

Create Service Device Template

- Ensure that you have read and completed the requirements described in [Guidelines and Limitations, on page 279](#).

This section describes how to configure one or more devices for a service graph.

Procedure

-
- Step 1** Log in to the Nexus Dashboard Orchestrator GUI.
- Step 2** From the left navigation pane, select **Configure > Tenant Templates**.
- Step 3** Choose the **Service Device** tab.
- Step 4** Create a Service Device template and associate it with the sites.
- From **Configure > Tenant Templates**, choose the **Service Device** tab.
 - Click **Create Service Device Template**.
 - In the template properties sidebar that opens, provide the **Name** for the template and **Select a Tenant**.
 - In the **Template Properties** page, choose **Actions > Add/Remove Sites** and associate the template with both sites.
 - Click **Save** to save the template.
- Step 5** Create and configure the device cluster.
- In the **Template Properties** page (template-level configuration), choose **Create Object > Service Device Cluster**.
The device cluster defines the service for which you want to redirect traffic.
 - In the **<cluster-name>** sidebar, provide the **Name** for the cluster.
The **Device Location** and **Device Mode** are pre-populated based on the currently supported use case.
 - Choose the **Device Type**.
 - For **Device Mode**, choose **L3**.

- e) Chose the **Connectivity Mode**.

Note

If you are configuring an L3Out-to-L3Out use case, you must use `One Arm`

- f) Provide the **Interface Name**.
g) For the **Interface Type**, choose `BD`.

For `vzAny` use cases, this release supports attaching the service device to a bridge domain only.

- h) Click **Select BD** > to choose the service bridge domain to which you want to attach this device.

This is the stretched service BD you created in the previous section, for example `FW-external`.

- i) For the **Redirect** option, choose `Yes`.

You must choose to enable redirect for the PBR use case. After choosing `Yes`, the **IP SLA Monitoring Policy** option becomes available.

- j) (Optional) Click **Select IP SLA Monitoring Policy** and choose an IP SLA policy if you had created one.
k) (Optional) In the **Advanced Settings** area, choose **Enable** if you want to provide additional settings for the service cluster.

You can configure the following advanced settings:

- **QoS Policy** – allows you assign a specific QoS level within the ACI fabrics for the redirected traffic.
- **Preferred Group** – specifies whether or not this service cluster is part of the preferred group.
- **Load Balancing Hashing** – allows you to specify the hashing algorithm for PBR load balancing.
For additional information, see [ACI Policy-Based Redirect Service Graph Design](#).
- **Pod Aware Redirection** – can be configured in Multi-Pod configuration if you want to specify the preferred PBR node. When you enable Pod-aware redirection, you can specify the Pod ID and redirection is programmed only in the leaf switches located in the specified Pod.
- **Rewrite Source MAC** – updates the source MAC address if the PBR node uses “source MAC based forwarding” instead of IP based forwarding.
For additional information, see [ACI Policy-Based Redirect Service Graph Design](#).
- **Advanced Tracking Options** – allows you to configure a number of advanced settings for the service node tracking. For additional information, see [Policy-Based Redirect and Threshold Settings for Tracking Service Nodes](#)

- l) Click **Ok** to save.

Note that after you create the Service Device Cluster, it is highlighted in red in the **Template Properties** (template-level configuration) page. At this point, you have defined redirection to a firewall service, but you must still provide the firewall information and the redirect policy you want to use at the site-local level.

Add Service Chaining to Contract

After you have deployed the base intersite L3Out use case and the Service Device template, you can add policy-based redirection by adding service chaining to the contract you created between the L3Out and an application EPG or another L3Out.

Procedure

Step 1 Navigate back to the application template where you defined the contract.

Step 2 Select the contract.

Step 3 In the **Service Chaining** area, click **+Service Chaining**.

Step 4 Choose the **Device Type**.

Note

The L3Out-to-L3Out use case supports only one-arm **Firewall** devices. For other intersite L3Out with PBR use cases, you can chain multiple devices.

Step 5 From the **Device** dropdown, choose the FW device cluster you created in the previous step.

Step 6 Ensure that **Consumer Connector Type Redirect** is enabled.

Step 7 Ensure that **Provider Connector Type Redirect** is enabled.

Step 8 Click **Add** to continue.

Step 9 Click **Save** to save the template.

Step 10 Click **Deploy Template** to re-deploy it.



CHAPTER 24

Intersite Transit Routing with PBR

- [Intersite Transit Routing with PBR, on page 283](#)
- [Intersite Transit Routing with PBR Guidelines and Limitations, on page 285](#)
- [Create Service Device Template, on page 287](#)
- [Create Contract and Add Service Chaining, on page 293](#)

Intersite Transit Routing with PBR

The following sections describe the guidelines, limitations, and configuration steps for the Intersite Transit Routing with Policy-Based Redirect (PBR) use case in your Multi-Site domain.



Note The following sections apply to the intersite transit routing (L3Out-to-L3Out) with PBR use case only. For information on L3Out-to-EPG intersite communication with PBR, see the chapter [Intersite L3Out with PBR, on page 275](#) instead; and for simple intersite L3Out use cases without PBR, see [Intersite L3Out, on page 259](#).

The intersite transit routing with PBR use case described in the following sections is supported for both inter-VRF and intra-VRF scenarios.

Configuration Workflow

The use case described in the following sections is an extension of a basic intersite L3Out PBR use case which is in turn an extension on basic intersite L3Out (without PBR) configuration. To configure this feature:

1. Configure basic external connectivity (L3Out) for each site.

The intersite L3Out with PBR configuration described in the following sections is built on top of existing external connectivity (L3Out) in each site. If you have not configured an L3Out in each site, create and deploy one as described in the [External Connectivity \(L3Out\), on page 229](#) chapter before proceeding with the following sections.

2. Create a contract between two external EPGs associated to the L3Outs deployed in different sites, as you typically would for the use case **without** PBR.
3. Add service chaining to the previously created contract as described in the following sections, which includes:
 - Creating a Service Device template and assigning it to sites.

The service device template must be assigned to the sites for which you want to enable intersite transit routing with PBR.

- Providing site-level configurations for the Service Device template.

Each site can have its own service device configuration including different high-availability models (such as active/active, active/standby, or independent service nodes).

- Associating the service device you defined to the contract used for the intersite L3Out use case you deployed in the previous step.



Note Please refer [ACI Contract Guide](#) and [ACI PBR White Paper](#) to understand Cisco ACI contract and PBR terminologies.

Traffic Flow

This section summarizes the traffic flow between two external EPGs in different sites.

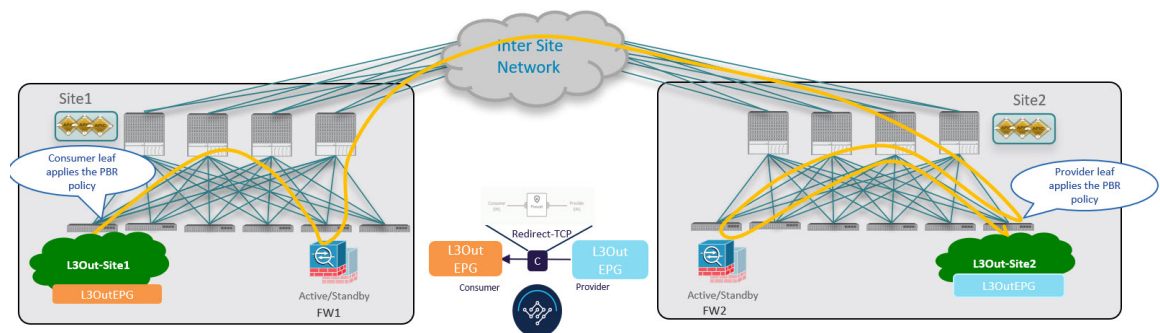


Note In this case, the traffic flow in both directions is redirected through both firewalls in order to avoid asymmetric traffic flows due to independent FW services deployed in the two sites.

Consumer-to-Provider Traffic Flow

Because any IP prefix associated with the destination external EPG for classification purposes is automatically programmed (with its Class-ID) on the consumer leaf switch, the leaf switch can always resolve the class-ID of the destination external EPG and apply the PBR policy redirecting the traffic to the local FW.

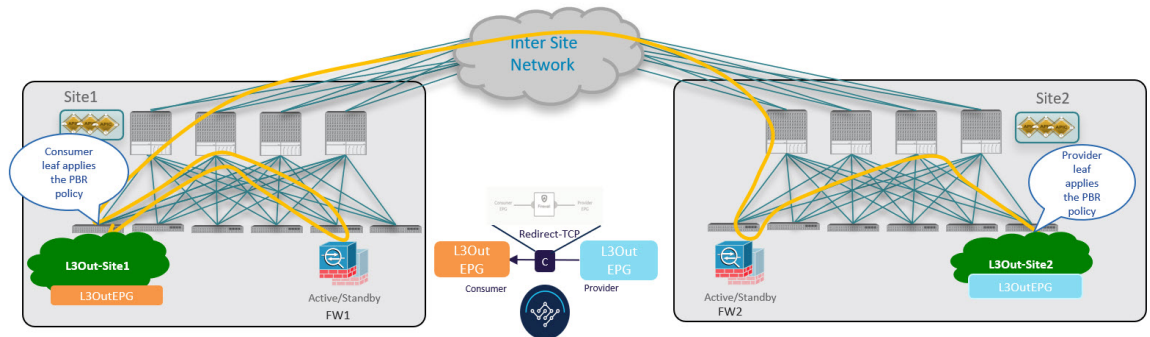
After the firewall on the consumer site has applied its security policy, the traffic is sent back into the fabric and forwarded across sites towards the provider border leaf nodes connecting to the external destination. The border leaf node receiving the traffic originated from Site1 applies the PBR policy and redirects the traffic to the local firewall node. After the firewall applies its local security policy, the traffic is sent back to the border leaf nodes, which can now simply forward it toward the external destination.



Provider-to-Consumer Traffic Flow

Similarly to the consumer-to-provider, the provider leaf switch can always resolve the class-ID of the destination external EPG and apply the PBR policy redirecting the traffic to the local FW in the other direction as well.

The traffic is then sent to the consumer site, where it is steered towards the local firewall before being forwarded to the external destination.



Intersite Transit Routing with PBR Guidelines and Limitations

The following guidelines and limitations apply when deploying intersite transit routing in vAny with PBR with Multi-Site:

- Intersite transit routing with vAny PBR is supported exclusively for single-node firewall on a one-arm interface.



Note This use case is supported for sites running Cisco APIC release 6.0(4) or later.

- While you can use your existing Service Graph objects defined in Application templates for these use cases, we recommend using the new service chaining workflow introduced in release 4.2(3) and implicitly creating new service graphs by defining the policies in Service Device templates and associating them to contracts.

The steps that are described in the following sections use the new Service Device templates to enable the supported use cases but will call out the specific differences when applicable.



Note Configuration of Service Graph objects in Application templates will be deprecated in a future release.

- The L3Out VRF can be stretched (for intra-VRF use case) or site local (for inter-VRF cases).

The following sections assume that you already have a VRF and L3Out already configured for each site.

Note that the "Site-aware policy enforcement" and "L3 Multicast" options must be enabled for the vAny VRF to enable the L3Out-to-L3Out use cases discussed in this chapter.

If you do not already have a VRF, you can create one in an Application template as you typically would. VRF configuration is described in detail in [Configuring VRFs, on page 71](#).

- You must enable the Site-aware Policy Enforcement Mode setting on the VRF to enable the new L3Out-to-L3Out use cases. Enabling or disabling the Site-aware Policy Enforcement Mode option will cause a brief traffic disruption (including the already existing contracts between EPGs) because the zoning rules must be updated on the leaf switches. We recommend that you perform this operation during a maintenance window. Enabling Site-aware Policy Enforcement Mode increases TCAM usage on the leaf switches for the existing contracts and contracting permit logging cannot be used in conjunction with this option.
- The service BD to which you want to attach the service device interface must be configured as L2 Stretched (BUM forwarding is optional and should be disabled).

If you do not already have a service BD, you can create one in an Application template. BD configuration is described in detail in [Configuring Bridge Domains, on page 72](#).

- The consumer, provider, and the service BDs must be configured in Hardware proxy-mode.
- The vzAny PBR destination must be connected to a stretched service BD, not to an L3Out.
- Only threshold down deny action and sip-dip-protocol hash is supported .
- The PBR destination must be in either the consumer or provider VRF instance. For example see [Cisco Application Centric Infrastructure Policy-Based Redirect Service Graph Design White Paper](#).

The following is not supported for this use case:

- Specific Remote leaf switch configurations.

Specific considerations apply for Multi-Site deployments leveraging Remote Leaf nodes. Intersite transit routing with PBR is not supported on vzAny PBR and L3Out-to-L3Out for communication between endpoints (consumer or provider) deployed on remote leaf nodes that belongs to different sites.

- Each VRF is limited to utilizing only one device in a one-arm configuration for vzAny-to-vzAny, L3OutEPG-to-L3OutEPG, and vzAny-To-L3OutEPG PBR. This restriction is enforced due to special ACL in APIC.
- We must use different firewall VLAN interfaces for redirection for vzAny-to-vzAny/L3OutEPG-to-L3OutEPG, and other use cases such as vzAny-to-EPG, EPG-to-EPG and EPG-to-L3OutEPG if they are in the same VRF.
- When vzAny with PBR is applied to north-south communication for any of the newly supported use cases (vzAny-to-vzAny, vzAny-to-L3OutEPG, L3OutEPG-to-L3OutEPG), ingress traffic optimization needs to be enabled for stretched subnets.
- Only one node service chain with L3 PBR destination is supported.
- Contract Permit logging is not supported on the VRF that has Site-aware Policy Enforcement Mode is enabled, which is required for vzAny PBR and L3OutEPG-to-L3OutEPG PBR.
- Pod-aware vzAny with PBR is not supported.

Create Service Device Template

The following steps describe how to create a Service Device template with a service node and its settings which you will use for the intersite transit routing use cases.

Before you begin

- Ensure that you have read and completed the requirements described in [Intersite Transit Routing with PBR Guidelines and Limitations, on page 285](#).
- You must have created a stretched service bridge domain (BD) to use with the service nodes you will define in this section.

If you do not already have a service BD, you can create one in an Application template as you typically would. BD configuration is described in detail in [Configuring Bridge Domains, on page 72](#).

Procedure

Step 1 Log in to the Nexus Dashboard Orchestrator GUI.

Step 2 From the left navigation pane, select **Configure > Tenant Templates**.

Step 3 (Optional) Create a Tenant Policies template and an IP SLA monitoring policy.

We recommend that you configure an IP SLA policy for traffic redirection as it simplifies the configuration of the PBR policy described in Step 7 below. If you have an IP SLA policy already defined, you can skip this step, otherwise:

- a) Choose the **Tenant Policies** tab.
- b) On the **Tenant Policies** page, click **Create Tenant Policy Template**.
- c) In the **Tenant Policies** page's right properties sidebar, provide the **Name** for the template and **Select a Tenant**.
- d) In the **Template Properties** page, choose **Actions > Add/Remove Sites** and associate the template with both sites.
- e) In the main pane, choose **Create Object > IP SLA Monitoring Policy**.
- f) Provide the **Name** for the policy, and define its settings.
- g) Click **Save** to save the template.
- h) Click **Deploy Template** to deploy it.

Step 4 Create a Service Device template and associate it with a tenant and with the sites.

- a) From **Configure > Tenant Templates**, choose the **Service Device** tab.
- b) Click **Create Service Device Template**.
- c) In the template properties sidebar that opens, provide the **Name** for the template and **Select a Tenant**.
- d) In the **Template Properties** page, choose **Actions > Add/Remove Sites** and associate the template with both sites.
- e) Click **Save** to save the template.

Step 5 Create and configure the device cluster.

- a) In the **Template Properties** page (template-level configuration), choose **Create Object > Service Device Cluster**.

The device cluster defines the service to which you want to redirect traffic. This release supports redirection to a firewall service node that can be deployed with three different redundancy models: active/standby, active/active, or a cluster of multiple independent nodes. The provisioning for those different options is covered in Step 7 below. Note that you can choose the firewall deployment model at the site level and different options can be deployed across different fabrics that are part of the same Multi-Site domain.

- b) In the <cluster-name> sidebar, provide the **Name** for the cluster.

The **Device Location** and **Device Mode** are pre-populated based on the currently supported use case. **Device Location** should be pre-configured as `ACI On-Prem` and **Device Mode** as `L3`.

- c) For the **Device Type**, choose `Firewall`.
 d) For **Device Mode**, choose `L3`.
 e) For **Connectivity Mode**, choose `One Arm`.

This release supports a service device connected in one arm mode only.

Note

When changing the device connectivity mode between one arm, two arm and advanced mode, the name of the device interface might change in the process. A warning message will alert the user, and any attempt to modify the interface will be restricted if the interface is currently in use by a contract. If the user wishes to preserve the previously used interface name and avoid disrupting the deployed configuration, they may choose to override the name change during the modification process.

Note

Validations are conducted only for one-arm and two-arm modes. In Advanced mode, no validations are performed, and it is assumed that the user is an expert when choosing this mode.

- f) Provide the **Interface Name**.
 g) For the **Interface Type**, choose `BD`.
 h) Click **Select BD >** to choose the service bridge domain to which you want to attach this device.

This is the stretched service BD you created as part of the [Intersite Transit Routing with PBR Guidelines and Limitations, on page 285](#), for example `FW-external`.

- i) For the **Redirect** option, choose `Yes`.

You must choose to enable redirect for the PBR use case. After choosing `Yes`, the **IP SLA Monitoring Policy** option becomes available.

- j) (Optional) Click **Select IP SLA Monitoring Policy** and choose the IP SLA policy you have created in a previous step.
 k) (Optional) In the **Advanced Settings** area, choose **Enable** if you want to provide additional settings for the service cluster.

You can configure the following advanced settings:

- **QoS Policy** – allows you assign a specific QoS level within the ACI fabrics for the redirected traffic.
- **Preferred Group** – specifies whether or not this service device interface is part of the preferred group.
- **Load Balancing Hashing** – allows you to specify the hashing algorithm for PBR load balancing.

Note

You must keep the default value for the `vzAny-to-vzAny`, `vzAny-to-ExtEPG`, and `ExtEPG-to-ExtEPG` use cases as they support only the default configuration. You can change the load balancing hashing for other use cases: `EPG-to-EPG`, `ExtEPG-to-EPG` and `vzAny-to-EPG`.

For additional information, see [ACI Policy-Based Redirect Service Graph Design](#).

- **Pod Aware Redirection** – can be configured in Multi-Pod configuration if you want to specify the preferred PBR node. When you enable Pod-aware redirection, you can specify the Pod ID and redirection is programmed only in the leaf switches located in the specified Pod.

- **Rewrite Source MAC** – updates the source MAC address if the PBR node uses “source MAC based forwarding” instead of IP based forwarding.

For additional information, see [ACI Policy-Based Redirect Service Graph Design](#).

- **Advanced Tracking Options** – allows you to configure a number of advanced settings for the service node tracking. For additional information, see [Policy-Based Redirect and Threshold Settings for Tracking Service Nodes](#)

- l) Click **Ok** to save.

Note that after you create the Service Device Cluster, it is highlighted in red in the **Template Properties** (template-level configuration) page. At this point, you have defined redirection to a firewall service, but you must still provide the firewall information and the redirect policy you want to use at the site-local level.

Step 6 Provide site-local configuration for the Service Device Cluster you created in the previous step.

- a) In the **Service Device Template** screen, choose the <site-name> tab.
- b) At the site level, choose the Service Device Cluster you created.
- c) In the properties sidebar, choose the **Domain Type**.

You can choose whether the firewall device in this site is `Physical` or `VMM` (virtual and hosted by a hypervisor that is part of a VMM domain).

- d) Click **Select Domain** to choose the domain to which this firewall device belongs.

You can choose either a physical or a virtual domain.

- If you choose a physical domain, provide the following information:
 - **VLAN** – you must provide the VLAN ID used for traffic between the fabric and the firewall device.
 - **Fabric to Device Connectivity** – provide the switch node and interface information for the fabric's connectivity to the firewall device.

- If you choose a VMM domain, provide the additional options:

- **Trunking Port** – used to enable tagged traffic for the L4-L7 VM.

By default, the ACI service graph configuration creates access-mode port groups and attaches them to the vNIC of the L4-L7 VM automatically.

- **Promiscuous Mode** – required if the L4-L7 virtual appliance must receive traffic destined to a MAC address that is not the vNIC MAC owned by the VM.
- **VLAN** – optional configuration for VMM domains and will be allocated from the dynamic VLAN pool associated with the domain if not specified.
- **Enhanced LAG Option** – if you are using enhanced LACP for the port channel between the hypervisor and the fabric.
- **VM Name** – choose the firewall's VM from the list of all VMs available in this VMM domain and the interface (**VNIC**) used for the firewall traffic.

Depending on the kind of device cluster you are deploying, click **+Add VM information** to provide additional cluster nodes.

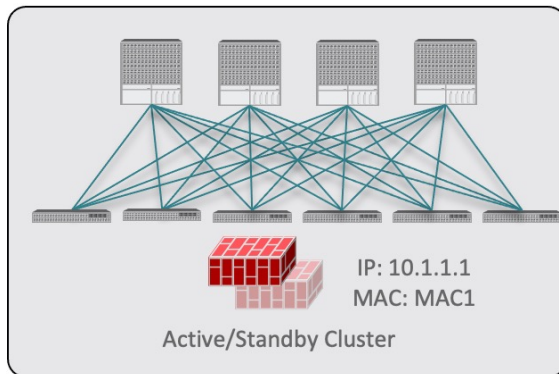
Step 7 Provide the FW device information and PBR destination IP addresses.

As previously mentioned, this release supports 3 deployment options for high-availability FW clusters: active/standby clusters, active/active clusters, and independent active nodes. In all three deployment options, the use of an IP SLA policy (mentioned in Step 3) allows to specify only the IP address of the firewall nodes, and the corresponding MAC address will be automatically discovered.

Note

You can deploy different designs in different sites.

- Active/standby clusters are identified by a single MAC/IP pair.



In this case, you need to provide a single PBR destination IP address identifying the active firewall node and also include information about every node in the cluster.

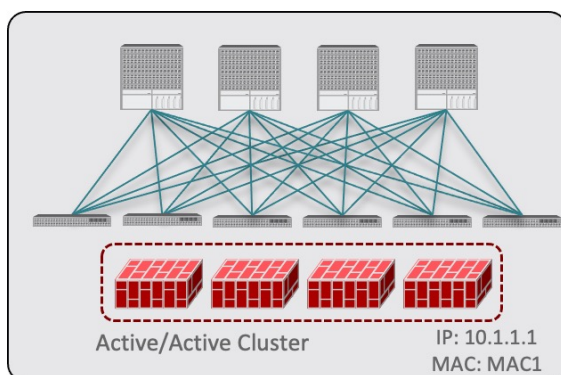
For example, for a 2-nodes active/standby cluster, you would provide the following:

- For a virtual firewall cluster, the VMs representing the active and standby firewall nodes and the IP address of the active firewall as PBR destination.
- For a physical firewall cluster, the interfaces used to connect the active and standby firewall nodes to the leaf switches of the fabric (vPC interfaces in the specific example below) and the IP address of the active firewall as PBR destination.

VM Information*			
VM Name*	VNIC*		
vCSA-7-Site1/ASAv-Pod1	Network adapter 2		
vCSA-7-Site1/ASAv-Pod2	Network adapter 2		
Add VM Information			
PBR Destinations			
IP Address *			
50.50.50.10			

Fabric To Device Connectivity			
Type *	Pod *	Node *	Path *
Virtual Port Channel	1	101,102	vPC-L101-L102-Port16
Virtual Port Channel	1	103,104	vPC-L103-L104-Port16
Add Fabric To Device Connectivity			
PBR Destinations			
IP Address *			
50.50.50.10			

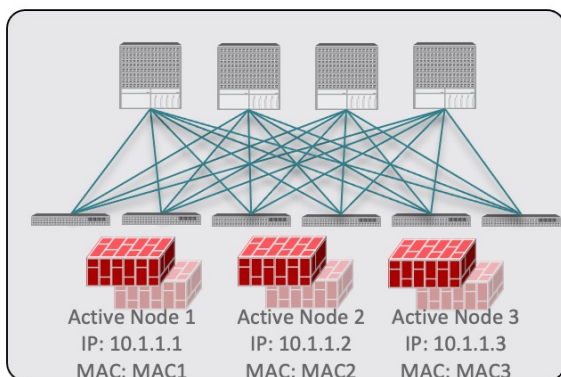
- Active/active clusters are also identified by a single MAC/IP pair.



For Cisco Firewalls (ASA or FTD models), the Active/Active cluster with a single MAC/IP pair is only supported for physical form factors, and all the cluster nodes own the same MAC/IP address and must be connected to the same vPC logical connection deployed on a pair of ACI leaf switches. As a result, the figure below shows how a single vPC interface and a single PBR Destination IP address should be configured on NDO, where the MAC address is dynamically discovered when using an IP SLA policy mentioned for the previous use case.

Fabric To Device Connectivity ⓘ			
Type *	Pod *	Node *	Path *
Virtual Port Channel	1	101,102	vPC-L101-L102-Port16 ⓘ 🗑️
Add Fabric To Device Connectivity			
PBR Destinations			
IP Address *			
50.50.50.10 ⓘ 🗑️			

- For independent active nodes configuration, each active node is identified by a unique MAC/IP addresses pair. Note that symmetric PBR ensures that the traffic is handled by the same active node in both directions.



In this case, you must provide individual PBR Destination IP addresses for each active node as well as each node's information in your NDO configuration.

For example, for a deployment of 3 independent firewall nodes, you would provide the following:

- For a virtual firewall form factor, the VMs representing the 3 firewall nodes and their unique IP addresses as PBR destinations.
- For a physical firewall form factor, the interfaces used to connect each firewall node to the leaf switches of the fabric (vPC interfaces in the specific example below) and the unique IP addresses of each firewall node as PBR destinations.

The screenshot shows two configuration panels. The top panel, 'VM Information', has a table with columns 'VM Name' and 'vNIC'. It lists three VMs: 'vCSA-7-Site1/ASA-vm-Pod1', 'vCSA-7-Site1/ASA-vm-Pod2', and 'vCSA-7-Site1/ASA-vm-Pod3', all with 'Network adapter 2' as the vNIC. Below this is a 'PBR Destinations' section with a table for IP addresses: '50.50.50.101', '50.50.50.102', and '50.50.50.103'. The bottom panel, 'Fabric To Device Connectivity', has a table with columns 'Type', 'Pod', 'Node', and 'Path'. It lists three entries: 'Virtual Port Channel' for Pod 1 with Node 101,102 and Path vPC-L101-L102-Port16; 'Virtual Port Channel' for Pod 1 with Node 103,104 and Path vPC-L103-L104-Port16; and 'Virtual Port Channel' for Pod 2 with Node 201,202 and Path vPC-L201-L202-Port2. Both panels have an 'Add' button and a 'PBR Destinations' section with IP addresses.

- a) Click **Add Fabric To Device Connectivity** (physical domain) or **Add VM Information** (VMM domain).

Depending on whether you selected physical or VMM domain in the previous step, you will specify information for either the firewall VM or the physical fabric connectivity to the firewall device.

For physical domains, provide the Pod, switch node, and the interface information.

For VMM domains, provide the VM name and vNIC information.

- b) Click **Add PBR Destination** to provide the IP address of the interface on the firewall that is connected to the service bridge domain.

Depending on the kind of device cluster you are deploying, you may need to provide one or more PBR destination IP addresses:

Note

This does not provision the IP address on the firewall's interface, but simply configures redirection of traffic toward that IP address. The specific firewall configuration is not deployed from NDO, and you must provision it separately.

- c) Click **Ok** to save the provided configuration.
d) Repeat this step for the other site with which you associated the template.

Step 8

Save and deploy the template.

- a) At the **Service Device Template** level, click **Save** to save the template configuration.
b) Choose the **Template Properties** tab and click **Deploy Template** to push the configuration to the sites.
c) (Optional) Verify that the configuration was created at the site level.

You can verify that the L4-L7 device is configured in the APIC by navigating to **<tenant-name> > Services > L4-L7 > Devices > <cluster-name>** in the APIC GUI. This shows the device cluster along with all the configuration you have provided in the previous steps.

To verify that the PBR policy is now configured on the APIC, navigate to **<tenant-name> > Policies > Protocol > L4-L7 Policy-Based Redirect** and you should see the **<cluster-name>-one-arm** redirect defined with the IP SLA monitoring policy you chose in *Step 5j* and the IP address you provided in *Step 7d*.

What to do next

After you have deployed the service device configuration, create the application template, external EPGs, and a contract with which you will associate the service chaining as described in [Create Contract and Add Service Chaining, on page 293](#).

Create Contract and Add Service Chaining

After you have created and deployed the service device templates and created the application template with the external EPGs for the L3Outs in each site, you can enable intersite transit routing with policy-based redirection by creating a contract between the external EPGs and associating the contract with the service devices you created in a previous section.

Before you begin

- You must have created and deployed external connectivity (L3Out) configuration in each site as described in [External Connectivity \(L3Out\), on page 229](#).
- You must have created and deployed the service device template containing the device configuration as described in [Create Service Device Template, on page 287](#).

Procedure

-
- | | |
|---------------|---|
| Step 1 | Navigate to the application templates where you want to create the external EPGs for the L3Outs and a contract between the external EPGs. Typically you would define the external EPGs in different site local templates associated to each site. |
| Step 2 | Create two external EPGs and associate the L3Outs in each site to the external EPG(s) at the site level.

This is the same process as you typically use when creating external connectivity for a fabric. Detailed information about the L3Out templates and External EPGs is described in External Connectivity (L3Out), on page 229 . |
| Step 3 | Create a contract in a stretched template associated to both sites, as you typically would and associate the contract with both external EPGs.

In this case, one of the external EPGs will be the <code>consumer</code> and the other one is the <code>provider</code> . |
| Step 4 | Choose the contract you created. |
| Step 5 | In the Service Chaining area, click +Service Chaining . |

Note

These steps assume that you have configured a brand new service device for this use case using the new Service Device template workflow introduced in release 4.2(3) as described in [Create Service Device Template, on page 287](#). If you already have a Service Graph defined in an application template, choose `Service Graph` instead and then select the existing service graph. However, keep in mind that the Service Graph option will be deprecated in a future release.

- Step 6** For **Device Type**, choose `Firewall`.
This release supports a service device connected in one arm mode only.
- Step 7** From the **Device** dropdown, choose the FW device cluster you created in the previous step.
- Step 8** Ensure that **Consumer Connector Type Redirect** is enabled.
- Step 9** Ensure that **Provider Connector Type Redirect** is enabled.
- Step 10** Click **Add** to continue.
- Step 11** Click **Save** to save the template.
- Step 12** Click **Deploy Template** to deploy it.
-



CHAPTER 25

Layer 3 Multicast

- [Layer 3 Multicast, on page 295](#)
- [Layer 3 Multicast Routing, on page 296](#)
- [Rendezvous Points, on page 296](#)
- [Multicast Filtering, on page 297](#)
- [Layer 3 Multicast Guidelines and Limitations, on page 298](#)
- [Creating Multicast Route Map Policy, on page 299](#)
- [Enabling Any-Source Multicast \(ASM\) Multicast, on page 301](#)
- [Enabling Source-Specific Multicast \(SSM\), on page 302](#)

Layer 3 Multicast

Cisco Multi-Site Layer 3 multicast is enabled or disabled at three levels, the VRF, the bridge domain (BD), and any EPGs that have multicast sources present.

At the top level, multicast routing must be enabled on the VRF that has any multicast-enabled BDs. On a multicast-enabled VRF, there can be a combination of multicast-enabled BDs and BDs where multicast routing is disabled. Enabling multicast routing on a VRF from the Cisco Nexus Dashboard Orchestrator GUI enables it on the APIC sites where the VRF is stretched.

Once a VRF is enabled for multicast, the individual BDs under that VRF can be enabled for multicast routing. Configuring Layer 3 multicast on a BD enables protocol independent multicast (PIM) routing on that BD. By default, PIM is disabled in all BDs.

If a source belonging to a specific site-local EPG sends multicast traffic to a remote site, the Nexus Dashboard Orchestrator must create a shadow EPG and program the corresponding subnet route(s) on the remote site for the source EPG. In order to limit the configuration changes applied to the remote Top-of-Rack (TOR) switches, you are required to explicitly enable Layer 3 multicast on the local EPGs that have multicast sources present, so that only the configuration necessary for those EPGs is pushed to the remote sites. EPGs with multicast receivers do not require enabling Layer 3 multicast.

Multi-Site supports all of the following Layer 3 multicast source and receiver combinations:

- Multicast sources and receivers inside ACI fabric
- Multicast sources and receivers outside ACI fabric
- Multicast sources inside ACI fabric with external receivers
- Multicast receivers inside ACI fabric with external sources

Layer 3 Multicast Routing

The following is a high level overview of the Layer 3 multicast routing across sites:

- When the multicast source is attached to the ACI fabric as an endpoint (EP) at one site and starts streaming a multicast flow, the specific site's spine switch that is elected as designated forwarder for the source VRF will forward the multicast traffic to all the remote sites where the source's VRF is stretched using Head End Replication (HREP). If there are no receivers in a specific remote site for that specific group, the traffic gets dropped on the receiving spine node. If there is at least a receiver, the traffic is forwarded into the site and reaches all the leaf nodes where the VRF is deployed and at that point is pruned/forwarded based on the group membership information.
- Prior to Cisco ACI Release 5.0(1), the multicast routing solution required external multicast routers to be the Rendezvous Points (RPs) for PIM-SM any-source multicast (ASM) deployments. Each site must point to the same RP address for a given stretched VRF. The RP must be reachable on each site via the site's local L3Out.
- When the source is outside and the receiver is within a fabric, the receiver will pull traffic via site's local L3Out as PIM joins toward RP and source are always sent via site local L3Out.
- Receivers in each site are expected to draw traffic from an external source via the site's local L3Out. As such, traffic received on the L3Out on one site should not be sent to other sites. This is achieved on the spine by pruning multicast traffic from being replicated into HREP tunnels.

In order to be able to do so, all multicast traffic originated from an external source and received on a local L3Out is remarked with a special DSCP value in the outer VXLAN header. The spines can hence match that specific DSCP value preventing the traffic from being replicated toward the remote sites.

- Traffic originated from a source connected to a site can be sent toward external receivers via a local L3Out or via L3Outs deployed in remote sites. The specific L3Out that is used for this solely depends on which site received the PIM Join for that specific multicast group from the external network.
- When multicast is enabled on a BD and an EPG on the Nexus Dashboard Orchestrator, all of the BD's subnets are programmed in the routing tables of all the leaf switches, including the border leaf nodes (BLs). This enables receivers attached to the leaf switches to determine the reachability of the multicast source in cases where the source BD is not present on the leaf switches. The subnet is advertised to the external network if there is a proper policy configured on the BLs. The /32 host routes are advertised if host-based routing is configured on the BD.

For additional information about multicast routing, see the [IP Multicast](#) section of the *Cisco APIC Layer 3 Networking Configuration Guide*.

Rendezvous Points

Multicast traffic sources send packets to a multicast address group, with anyone joining that group able to receive the packets. Receivers that want to receive traffic from one or more groups can request to join the group, typically using Internet Group Management Protocol (IGMP). Whenever a receiver joins a group, a multicast distribution tree is created for that group. A Rendezvous Point (RP) is a router in a PIM-SM multicast domain that acts as a shared root for a multicast shared tree.

The typical way to provide a redundant RP function in a network consists in deploying a functionality called Anycast RP, which allows two or more RPs in the network to share the same anycast IP address. This provides

for redundancy and load balancing. Should one RP device fail, the other RP can take over without service interruption. Multicast routers can also join the multicast shared tree by connecting to any of the anycast RPs in the network, with PIM `join` requests being forwarded to the closest RP.

Two types of RP configurations are supported from Nexus Dashboard Orchestrator:

- **Static RP**—If your RP is outside the ACI fabric.
- **Fabric RP**—If the border leaf switches in the ACI fabric will function as the anycast RPs.

Any number of routers can be configured to work as RPs and they can be configured to cover different group ranges. When defining the RP inside the ACI fabric, you can configure which groups the RP covers by creating a route-map policy that contains the list of groups and attaching this policy to the RP when adding it to the VRF. Creating a route map is described in [Creating Multicast Route Map Policy, on page 299](#), while VRF configuration is described in [Enabling Any-Source Multicast \(ASM\) Multicast, on page 301](#).

Both static and fabric RPs require PIM-enabled border leaf switches in the VRF where multicast routing is enabled. L3Out configuration is currently configured locally from the APIC at each site including enabling PIM for the L3Out. Please refer to the [Cisco APIC Layer 3 Networking Configuration Guide](#) for details on configuration PIM on L3Outs

Multicast Filtering

Multicast filtering is a data plane filtering feature for multicast traffic available starting with Cisco APIC, Release 5.0(1) and Nexus Dashboard Orchestrator, Release 3.0(1).

Cisco APIC supports control plane configurations that can be used to control who can receive multicast feeds and from which sources. In some deployments, it may be desirable to constrain the sending and/or receiving of multicast streams at the data plane level. For example, you may need to allow multicast senders in a LAN to be able to send only to specific multicast groups or to allow receivers to receive multicast from only specific sources.

To configure multicast filtering from the Nexus Dashboard Orchestrator, you create source and destination multicast route maps, each of which contains one or more filter entries based on the multicast traffic's source IP and/or group with an action (`Permit` or `Deny`) attached to it. You then enable the filtering on a bridge domain by attaching the route maps to it.

When creating a multicast route map, you can define one or more filter entries. Some entries can be configured with a `Permit` action and other entries can be configured with a `Deny` action, all within the same route map. For each entry, you can provide a **Source IP** and a **Group IP** to define the traffic that will match the filter. You must provide at least one of these fields, but can choose to include both. If one of the fields is left blank, it will match all values.

You can enable both multicast source filtering and multicast receiver filtering on the same bridge domain. In this case one bridge domain can provide filtering for both, the source as well as the receivers.

If you do not provide a route map for a BD, the default action is to allow all multicast traffic on the bridge domain. However, if you do select a route map, the default action changes to deny any traffic not explicitly matched to a filter entry in the route map.

Source Filtering

For any multicast sources that are sending traffic on a bridge domain, you can configure a route map policy with one or more source and group IP filters defined. The traffic is then matched against every entry in the route map and one of the following actions takes place:

- If the traffic matches a filter entry with a `Permit` action in the route map, the bridge domain will allow traffic from that source to that group.
- If the traffic matches a filter entry with a `Deny` action in the route map, the bridge domain will block traffic from that source to that group.
- If the traffic does not match any entries in the route map, the default `Deny` action is applied.

Source filter is applied to the bridge domain on the First-Hop Router (FHR), represented by the ACI leaf node where the source is connected. The filter will prevent multicast from being received by receivers in different bridge domains, the same bridge domain, and external receivers.

Destination (Receiver) Filtering

Destination (receiver) filtering does not prevent receivers from joining a multicast group. The multicast traffic is instead allowed or dropped in the data plane based on the source IP and multicast group combination.

Similarly to the source filtering, when multicast traffic matches a destination filter, one of the following actions takes place:

- If the traffic matches a filter entry with a `Permit` action in the route map, the bridge domain will allow the traffic from the multicast group to the receiver.
- If the traffic matches a filter entry with a `Deny` action in the route map, the bridge domain will block the traffic from the multicast group to the receiver.
- If the traffic does not match any entries in the route map, the default `Deny` action is applied.

Destination filter is applied to the bridge domain on the Last-Hop Router (LHR), represented by the ACI leaf node where the receiver is connected, so other bridge domains can still receive the multicast traffic.

Layer 3 Multicast Guidelines and Limitations

Up to the current software release, Cisco Nexus Dashboard Orchestrator cannot be used to deploy specific multicast control plane filtering policies, such as IGMP or PIM related policies, on each site. As such you must configure any additional policies required for your use case on each APIC site individually for end-to-end solution to work. For specific information on how to configure those settings on each site, see the [Cisco APIC Layer 3 Networking Configuration Guide](#).

You must also ensure that QoS DSCP translation policies in all fabrics are configured consistently. When you create custom QoS policies in ACI fabrics, you can create a mapping between the ACI QoS Levels and the packet header DSCP values for packets ingressing or egressing the fabric. The same ACI QoS Levels must be mapped to the same DSCP values on all sites for the multicast traffic to transit between those sites. For specific information on how to configure those settings on each site, see the [Cisco APIC and QoS](#)

Multicast Filtering

The following additional guidelines apply if you enable the multicast filtering:

- Multicast filtering is supported only for IPv4.
- You can enable either the multicast source filtering, or the receiver filtering, or both on the same bridge domain.
- If you do not want to have multicast filters on a bridge domain, do not configure a source filter or destination filter route maps on that bridge domain.

By default, no route maps are associated with a bridge domain, which means that all multicast traffic is allowed. If a route map is associated with a bridge domain, only the permit entries in that route map will be allowed, while all other multicast traffic will be blocked.

If you attach an empty route map to a bridge domain, route maps assume a `deny-all` by default, so all sources and groups will be blocked on that bridge domain.

- Multicast filtering is done at the BD level and apply to all EPGs within the BD. As such you cannot configure different filtering policies for different EPGs within the same BD. If you need to apply filtering more granularly at the EPG level, you must configure the EPGs in separate BDs.
- Multicast filtering is intended to be used for Any-Source Multicast (ASM) ranges only. Source-Specific Multicast (SSM) is not supported for source filtering and is supported only for receiver filtering.
- For both, source and receiver filtering, the route map entries are matched based on the specified `order` of the entry, with lowest number matched first. This means that lower order entries will match first, even if they are not the longest match in the list, and higher order entries will not be considered.

For example, if you have the following route map for the `192.0.3.1/32` source:

Order	Source IP	Action
1	192.0.0.0/16	Permit
2	192.0.3.0/24	Deny

Even though the second entry (`192.0.3.0/24`) is a longer match as a source IP, the first entry (`192.0.0.0/16`) will be matched because of the lower order number.

Creating Multicast Route Map Policy

This section describes how to create a multicast route map policy. You may want to create a route map for one of the following reasons:

- Define a set of filters for multicast source filtering.
- Define a set of filters for multicast destination filtering.
- Define a set of group IPs for a Rendezvous Point (RP).

When configuring an RP for a VRF, if you do not provide a route map, the RP will be defined for the entire multicast group range (`224.0.0.0/4`). Alternatively, you can provide a route map with a group or group range that is defined to limit the RP to those groups only.

Procedure

Step 1 Log in to your Cisco Nexus Dashboard and open the Cisco Nexus Dashboard Orchestrator service.

Step 2 Create a new Tenant Policy.

- a) From the left navigation pane, choose **Configure > Tenantat Template > > Tenant Policies**.
- b) On the **Tenant Policy Templates** page, click **Add Tenant Policy Template**.
- c) In the Tenant Policies page's right properties sidebar, provide the **Name** for the tenant.
- d) From the **Select a Tenant** drop-down, choose the tenant with which you want to associate this template.

All the policies that you create int his template as described in the following steps will be associated with the selected tenant deployed to it when you push the template to a specific site.

By default, the new template is empty, so you need to add one or more tenant policies as described in the following steps. You don't have to create every policy available in the template – you can create a template with just a single route map policy for your multicast use case.

Step 3 Create a Route Map Policy for Multicast.

- a) From the **+Create Object** dropdown, select **Route Map Policy for Multicast**.
- b) In the right properties sidebar, provide the **Name** for the policy.
- c) (Optional) Click **Add Description** and provide a description for the policy.
- d) Click **+Add Route Map for Multicast Entries** and provide the route map information.

For each route map, you must create one or more route map entries. Each entry is a rule that defines an action based on one or more matching criteria based on the following information:

- **Order** – Order is used to determine the order in which the rules are evaluated.
- **Group IP, Src IP, and RP IP** – You can use the same multicast route map policy UI for two different use cases—To configure a set of filters for multicast traffic or to restrict a rendezvous point configuration to a specific set of multicast groups. Depending on which use case you're configuring, you must fill some of the fields in this screen:
 - For multicast filtering, you can use the **Source IP** and the **Group IP** fields to define the filter. You must provide at least one of these fields, but can choose to include both. If one of the fields is left blank, it matches all values.

The Group IP range must be between 224.0.0.0 and 239.255.255.255 with a netmask between /4 and /32. You must provide the subnet mask.

The **RP IP** (Rendezvous Point IP) is not used for multicast filtering route maps, so leave this field blank.
 - For Rendezvous Point configuration, you can use the **Group IP** field to define the multicast groups for the RP.

The Group IP range must be between 224.0.0.0 and 239.255.255.255 with a netmask between /4 and /32. You must provide the subnet mask.

For a Rendezvous Point configuration, the **RP IP** is configured as part of the RP configuration. If a route-map is used for group filtering it is not necessary to configure an RP IP address in the route-map. In this case, leave the **RP IP** and **Source IP** fields empty.

- **Action** – Action defines the action to perform, either `Permit` or `Deny` the traffic, if a match is found.

- e) Click the check mark icon to save the entry.
- f) Repeat the previous substeps to create any additional route map entries for the same policy.
- g) Click **Save** to save the policy and return to the template page.
- h) Repeat this step to create any additional Route Map for Multicast policies.

Enabling Any-Source Multicast (ASM) Multicast

The following procedure describes how to enable ASM multicast on VRF, BD, and EPG using the Nexus Dashboard Orchestrator GUI. If you want to enable SSM multicast, follow the steps in [Enabling Source-Specific Multicast \(SSM\)](#), on page 302 instead.

Before you begin

- Ensure you have read and followed the information described in [Layer 3 Multicast Guidelines and Limitations](#), on page 298.
- If you plan to enable multicast filtering, create the required multicast route maps, as described in [Creating Multicast Route Map Policy](#), on page 299.
- Note that the site-local L3Outs must have PIM enabled in the VRF when fabric RP is enabled.

This is described in Step 6 of the following procedure. Additional information about PIM configuration on an L3Out is available in the [Cisco APIC Layer 3 Networking Configuration Guide](#).

Procedure

- Step 1** Log in to your Nexus Dashboard Orchestrator.
- Step 2** From the left-hand sidebar, select the **Configure > Tenant Template > Applications > Schemas** view.
- Step 3** Click on the Schema you want to change.
- Step 4** Enable Layer 3 multicast on a VRF.
First, you enable Layer 3 multicast on a VRF that is stretched between sites.
 - a) Select the VRF for which you want to enable Layer 3 multicast.
 - b) In the right properties sidebar, check the **L3 Multicast** checkbox.
- Step 5** Add one or more Rendezvous Points (RP).
 - a) Select the VRF.
 - b) In the right properties sidebar, click **Add Rendezvous Points**.
 - c) With the VRF still selected, click **Add Rendezvous Points** in the right sidebar.
 - d) In the **Add Rendezvous Points** window, provide the IP address of the RP.
 - e) Choose the type of the RP.
 - **Static RP**—If your RP is outside the ACI fabric.
 - **Fabric RP**—If your RP is inside the ACI fabric.

- f) (Optional) From the **Multicast Route-Map Policy** dropdown, select the route-map policy you configured previously.

By default, the RP IP you provide applies to all multicast groups in the fabric. If you want to restrict the RP to only a specific set of multicast groups, define those groups in a route map policy and select that policy here.

Step 6 Enable PIM on the L3Out.

Both static and fabric RPs require PIM-enabled border leaf switches where multicast routing is enabled. L3Out configuration currently cannot be done from the Nexus Dashboard Orchestrator, so you must ensure that PIM is enabled directly in the site's APIC. Additional information about PIM configuration on an L3Out is available in the [Cisco APIC Layer 3 Networking Configuration Guide](#).

- Log in to your site's Cisco APIC.
- In the top menu, click **Tenants** and select the tenant that contains the L3Out.
- In the left navigation menu, select **Networking > L3Outs > <l3out-name>**.
- In the main pane, choose the **Policy** tab.
- Check the **PIM** options.

Multi-Site supports IPv4 multicast only.

Step 7 Enable Layer 3 multicast on a BD.

Once you have enabled L3 Multicast on a VRF, you can enable L3 multicast on a Bridge Domain (BD) level.

- Select the BD for which you want to enable Layer 3 multicast.
- In the right properties sidebar, check the **L3 Multicast** checkbox.

Step 8 (Optional) If you want to configure multicast filtering, provide the route-maps for source and destination filtering.

- Select the BD.
- In the right properties sidebar, select a **Route-Map Source Filter** and **Route-Map Destination Filter**.

You can choose to enable either the multicast source filtering, or the receiver filtering, or both.

Keep in mind, if you do not select a route map, the default action is to allow all multicast traffic on the bridge domain; however, if you select a route map the default action changes to deny any traffic not explicitly matched to a filter entry in the route map.

Step 9 If your multicast source is in one site and is not stretched to the other sites, enable intersite multicast source option on the EPG.

Once you have enabled L3 Multicast on the BD, you must also enable multicast on the EPGs (part of multicast-enabled BDs) where multicast sources are connected.

- Select the EPG for which you want to enable Layer 3 multicast.
- In the right-hand sidebar, check the **Intersite Multicast Source** checkbox.

Enabling Source-Specific Multicast (SSM)

The following procedure describes how to enable SSM multicast on VRF, BD, and EPG using the Cisco Nexus Dashboard Orchestrator GUI. If you want to enable ASM multicast, follow the steps in [Enabling Any-Source Multicast \(ASM\) Multicast](#), on page 301 instead.

Before you begin

- Ensure you have read and followed the information that is described in [Layer 3 Multicast Guidelines and Limitations, on page 298](#).
- If you plan to enable multicast filtering, create the required multicast route maps, as described in [Creating Multicast Route Map Policy, on page 299](#).
- You must configure IGMPv3 interface policy for the multicast-enabled BDs at the site-local level.
This is described in Step 8 of the following procedure. Additional information is available in the [Cisco APIC Layer 3 Networking Configuration Guide](#).

Procedure

Step 1 Log in to your Cisco Nexus Dashboard Orchestrator.

Step 2 From the left sidebar, select the **Configure > Tenant Template > Application > Schemas** view.

Step 3 Click the Schema that you want to change.

Step 4 Enable Layer 3 multicast on a VRF.

First, you enable Layer 3 multicast on a VRF that is stretched between sites.

- a) Select the VRF for which you want to enable Layer 3 multicast.
- b) In the right properties sidebar, check the **L3 Multicast** check box.

Step 5 (Optional) Configure a custom range for SSM Listeners.

The default SSM range is 232.0.0.0/8, which is automatically configured on the switches in your fabric. If you are using SSM, we recommend configuring your Listeners to join groups in this range, in which case you can skip this step.

If for any reason you do not want to change your listener configuration, you can add extra SSM ranges under the VRF settings by creating a route-map with up to 4 extra ranges. Keep in mind that if you add a new range it becomes an SSM range and cannot be used for ASM at the same time.

Custom SSM range configuration must be done directly in the site's APIC:

- a) Log in to your site's Cisco APIC.
- b) In the top menu, click **Tenants** and select the tenant that contains the VRF.
- c) In the left navigation menu, select **Networking > VRFs > <VRF-name> > Multicast**.
- d) In the main pane, choose the **Pattern Policy** tab.
- e) From the **Route Map** drop-down in the **Source Specific Multicast (SSM)** area, choose an existing route map or click **Create Route Map Policy for Multicast** option to create a new one.

If you select an existing route map, click the icon next to the drop-down to view the route map's details.

In the route map details window or the **Create Route Map Policy for Multicast** window that opens, click + to add an entry. Then configure the Group IP; you must provide only the group IP address to define the new range.

Step 6 (Optional) Enable PIM on the site's L3Out.

If you connect multicast sources or receivers to the external network domain, you must also enable PIM on the site's L3Out. L3Out configuration currently cannot be done from the Cisco Nexus Dashboard Orchestrator, so you must

ensure that PIM is enabled directly in the site's APIC. Additional information about PIM configuration on an L3Out is available in the [Cisco APIC Layer 3 Networking Configuration Guide](#).

- a) Log in to your site's Cisco APIC.
- b) In the top menu, click **Tenants** and select the tenant that contains the L3Out.
- c) In the left navigation menu, select **Networking > L3Outs > <l3out-name>**.
- d) In the main pane, choose the **Policy** tab.
- e) Check the **PIM** options.

Multi-Site supports IPv4 multicast only.

Step 7 Enable Layer 3 multicast on a BD.

When you have enabled L3 Multicast on a VRF, you can enable L3 multicast on a Bridge Domain (BD) level.

- a) Select the BD for which you want to enable Layer 3 multicast.
- b) In the right properties sidebar, check the **L3 Multicast** check box.

Step 8 Enabled IGMPv3 interface policy on the bridge domains where receivers are connected.

Because you are configuring SSM, you must also assign an IGMPv3 interface policy to the BD. By default, when PIM is enabled, IGMP is also automatically enabled on the SVI but the default version is set to IGMPv2. You must explicitly set the IGMP interface policy to IGMPv3. This must be done at the site-local level:

- a) Log in to your site's Cisco APIC.
- b) In the top menu, click **Tenants** and select the tenant that contains the BD.
- c) In the left navigation menu, select **Networking > Bridge Domains > <BD-name>**.
- d) In the main pane, choose the **Policy** tab.
- e) From the **IGMP Policy** drop-down, select the IGMP policy or click **Create IGMP Interface Policy** to create a new one.

If you select an existing policy, click the icon next to the drop-down to view the policy details.

In the policy details window or the **Create Route Map Policy for Multicast** window that opens, ensure that the **Version** field is set to `Version 3`.

Step 9 (Optional) If you want to configure multicast filtering, provide the route-maps for source and destination filtering.

- a) Select the BD.
- b) In the right properties sidebar, select a **Route-Map Source Filter** and **Route-Map Destination Filter**.

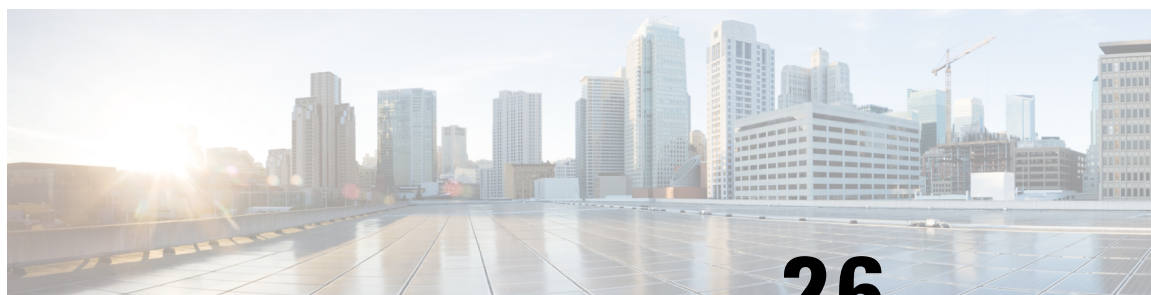
You can choose to enable either the multicast source filtering, or the receiver filtering, or both.

Keep in mind, if you do not select a route map, the default action is to allow all multicast traffic on the bridge domain; however, if you select a route map the default action changes to deny any traffic that is not explicitly matched to a filter entry in the route map.

Step 10 If your multicast source is in one site and is not stretched to the other sites, enable intersite multicast source option on the EPG.

When you have enabled L3 Multicast on the BD, you must also enable multicast on the EPGs (part of multicast-enabled BDs) where multicast sources are connected.

- a) Select the EPG for which you want to enable Layer 3 multicast.
- b) In the right sidebar, check the **Intersite Multicast Source** check box.



CHAPTER 26

QoS Preservation Across IPN

- [QoS and Global DSCP Policy, on page 305](#)
- [DSCP Policy Guidelines and Limitations, on page 305](#)
- [Configuring Global DSCP Policy, on page 306](#)
- [Set QoS Level for EPGs and Contracts, on page 307](#)

QoS and Global DSCP Policy

Cisco ACI Quality of Service (QoS) feature allows you to classify the network traffic in your fabric and then to prioritize and police the traffic flow to help avoid congestion in your network. When traffic is classified within the fabric, it is assigned a QoS Priority Level, which is then used throughout the fabric to provide the most desirable flow of packets through the network.

This release of Nexus Dashboard Orchestrator supports configuration of QoS level based on source EPG or a specific Contract. Additional options are available in each fabric directly. You can find detailed information on ACI QoS in [Cisco APIC and QoS](#).

When traffic is sent and received within the Cisco ACI fabric, the QoS Level is determined based on the CoS value of the VXLAN packet's outer header. In certain use cases, such as multi-pod or remote leaf topologies, the traffic must transit an intersite network, where devices that are not under Cisco APIC's management may modify the CoS values in the packets. In these cases you can preserve the ACI QoS Level between parts of the same fabric or different fabrics by creating a mapping between the Cisco ACI QoS level and the DSCP value within the packet.

DSCP Policy Guidelines and Limitations

When configuring the global DSCP translation policy, the following guidelines apply.



Note

If you plan to use the global DSCP translation policy along with SD-WAN integration, skip this chapter and see the [SD-WAN Integration, on page 329](#) chapter instead for all information including the full list of guidelines and limitations.

- Global DSCP policy is supported for on-premises sites only.
- When defining the global DSCP policy, you must pick a unique value for each QoS Level.

- When assigning QoS level, you can choose to assign it to a specific Contract or an entire EPG.

If multiple QoS levels could apply for any given traffic, only one is applied using the following precedence:

- Contract QoS level: If QoS is enabled in the Contract, the QoS level specified in the contract is used.
- Source EPG QoS level: If QoS level is not specified for the Contract, the QoS level set for the source EPG is used.
- Default QoS level: If no QoS level is specified, the traffic is assigned Level 3 QoS class by default.

Configuring Global DSCP Policy

When traffic is sent and received within a Cisco ACI fabric, it is prioritized based on the ACI QoS Level, which is determined based on the CoS value of the VXLAN packet's outer header. When traffic exits the ACI fabric toward an intersite network, for example in multipod and remote leaf switch topologies, the QoS level is translated into a DSCP value which is included in the outer header of the VXLAN-encapsulated packet.

This section describes how to define the DSCP translation policy for traffic entering or exiting ACI fabric. This is required when traffic must transit through non-ACI networks, where devices that are not under Cisco APIC's management may modify the CoS values in the transiting packets.

Before you begin

- You should be familiar with Quality of Service (QoS) functionality within ACI fabrics.
- QoS is described in more detail in [Cisco APIC and QoS](#).

Procedure

Step 1 Log in to your Cisco Nexus Dashboard and open the Cisco Nexus Dashboard Orchestrator service.

Step 2 Create a new Tenant Policy.

- From the left navigation pane, choose **Configure > Tenant Template > Tenant Policies**.
- On the **Tenant Policy Templates** page, click **Add Tenant Policy Template**.
- In the **Tenant Policies** page's right properties sidebar, provide the **Name** for the template.
- From the **Select a Tenant** drop-down, choose the tenant with which you want to associate this template.

All the policies that you create in this template as described in the following steps will be associated with the selected tenant and deployed to it when you push the template to a specific site.

By default, the new template is empty, so you must add one or more tenant policies as described in the following steps. You don't have to create every policy available in the template – you can define one or more policies of each type to deploy along with this template. If you don't want to create a specific policy, simply skip the step that describes it.

Step 3 Create a QoS DSCP policy.

- From the **+Create Object** drop-down, select **QoS DSCP**.
- In the right properties sidebar, provide the **Name** for the policy.
- (Optional) Click **Add Description** and provide a description for the policy.

d) Provide policy details.

- **Admin State** – Enables or disables the policy.
- **Advanced Settings** – Click the arrow next to this section to expand.

Choose the DSCP value for each ACI QoS level. Each drop-down contains the default list of available DSCP values. You must choose a unique DSCP value for each level.

e) Repeat this step to create any additional QoS DSCP policies.

Typically, we recommend applying this policy consistently across all sites that are part of your Multi-Site domain.

Step 4 Assign the policy to one or more sites.

- a) In the Fabric Policies template view, select **Actions** > **Add/Remove Sites**.
- b) In the **Add Sites to <template>** dialog, select one or more sites for this policy template and click **Ok**.
- c) In the Fabric Policies template view, click **Deploy**.

After you save and deploy, the DSCP policy settings will be pushed to each site. You can verify the configuration by signing in to the site's APIC and navigating to **Tenants** > **infra** > **Policies** > **Protocol** > **DSCP class-CoS translation policy for L3 traffic**.

What to do next

After you have defined the global DSCP policy, you can assign the ACI QoS Levels to EPGs or Contracts as described in [Set QoS Level for EPGs and Contracts, on page 307](#).

Set QoS Level for EPGs and Contracts

This section describes how to choose an ACI QoS level for traffic in your fabrics. You can choose to specify QoS for individual Contracts or entire EPGs.

Before you begin

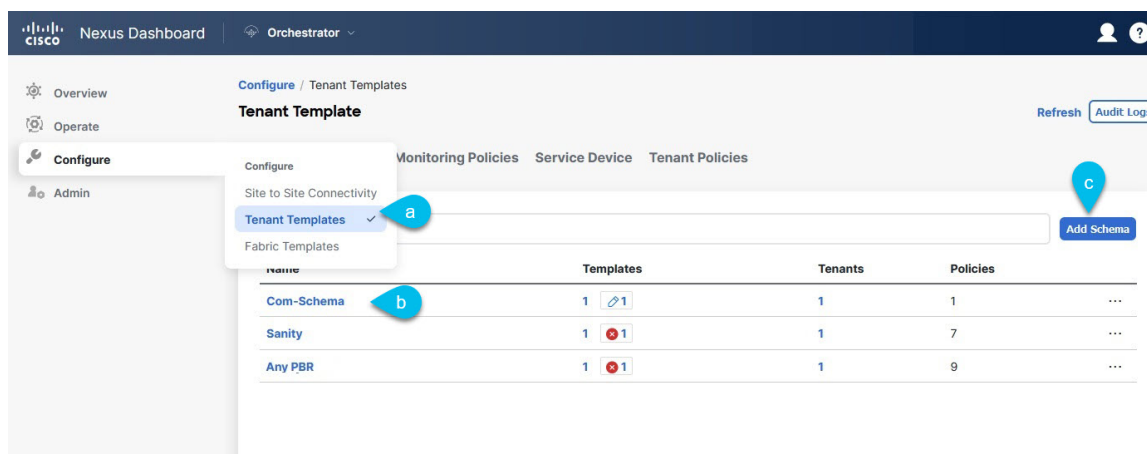
- You must have defined the global DSCP policy, as described in [Configuring Global DSCP Policy, on page 306](#).
- You should be familiar with Quality of Service (QoS) functionality within ACI fabrics. QoS is described in more detail in [Cisco APIC and QoS](#).

Procedure

Step 1 Log in to your Cisco Nexus Dashboard Orchestrator GUI.

Step 2 Choose the Schema that you want to edit.

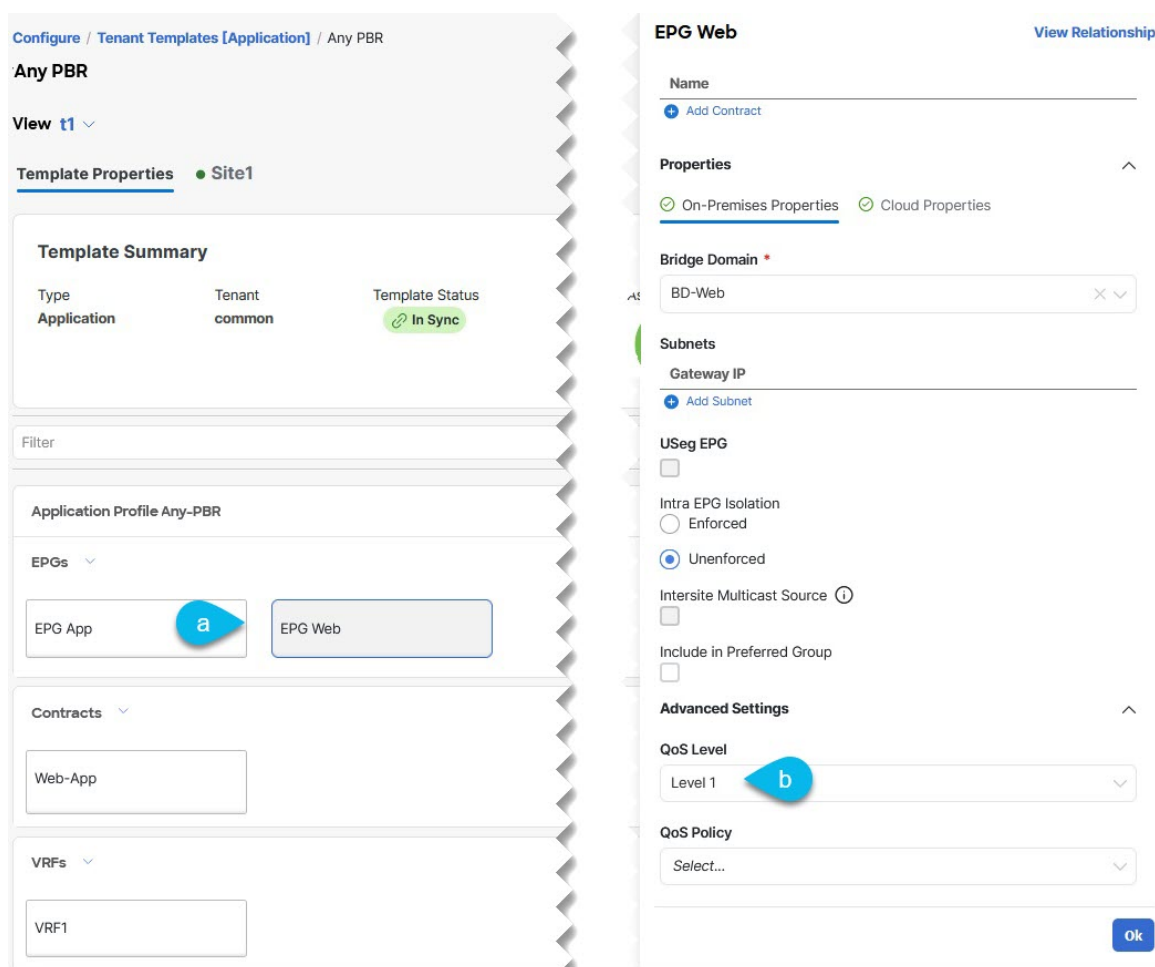
Set QoS Level for EPGs and Contracts



- Navigate to **Configure > Tenant Template > Applications > Schemas > .**
- Click the name of the schema that you want to edit or **Create Schema** to create a new one.

The **Edit Policy** window opens.

Step 3 Pick a QoS Level for an EPG.



- a) In the main pane, scroll down to the **EPG** area and select an EPG or click **Add EPG** to create a new one.
- b) In the right sidebar, scroll down to the **QoS Level** drop-down and choose the QoS Level you want to assign to the EPG.

Step 4 Pick a QoS Level for an EPG.

The screenshot displays the configuration interface for EPGs. On the left, the 'Contracts' section is expanded, showing 'Web-App' selected. A blue callout 'a' points to this selection. Below it, the 'VRFs' section shows 'VRF1' and 'Bridge Domains' section shows 'BD-App', 'BD-Web', and 'FW-extern'. On the right, the 'Filter Chain' for 'Permit-Any' is shown. The 'QoS Level' dropdown is set to 'Level 1', with a blue callout 'b' pointing to it. The 'Target DSCP' is set to 'Unspecified'. An 'Ok' button is visible at the bottom right of the right pane.

- a) In the main pane, scroll down to the **Contract** area and select a Contract or click the + icon to create a new one.
- b) In the right sidebar, scroll down to the **QoS Level** drop-down and choose the QoS Level you want to assign to the Contract.



CHAPTER 27

SD-Access and ACI Integration

- [Cisco SD-Access and Cisco ACI Integration, on page 311](#)
- [Macro Segmentation, on page 312](#)
- [Cisco SD-Access and Cisco ACI Integration Guidelines, on page 314](#)
- [Onboarding the DNA Center, on page 315](#)
- [Configuring Connectivity Toward the SD-Access Domain, on page 316](#)
- [Viewing the Status of the SD-Access to ACI Integration, on page 318](#)
- [Extending a Virtual Network, on page 320](#)
- [Mapping or Unmapping a VN to a VRF, on page 321](#)
- [Configuring Transit Routing, on page 323](#)

Cisco SD-Access and Cisco ACI Integration



Note Cisco Nexus Dashboard and Cisco DNAC integration allows for automation of a subset of network connectivity and macro segmentation scenarios across Nexus and campus SD-Access fabric deployments. This integration is under limited availability. Please contact your Cisco representative for additional information.

Cisco Software-Defined Access (SD-Access or SDA) is a solution within the Cisco Digital Network Architecture (DNA), which defines a campus-and-branch architecture that implements Cisco's Intent-Based Networking (IBN) framework. Cisco SD-Access defines a uniform policy-based wired and wireless network fabric that meets business needs with security, automation, and assurance. The Cisco Digital Network Architecture Controller (DNAC), in combination with Cisco Identity Services Engine (ISE), is the unified point of automation and management for the Cisco SD-Access fabric.

Release 3.7(1) of Cisco Nexus Dashboard Orchestrator (NDO) adds support for Cisco SD-Access and Cisco ACI integration. The purpose of SD-Access and ACI integration is to securely connect the campus-and-branch network to the data center network. With Release 3.7(1), NDO can perform the following functions:

- gather network and resource information from both domains
- automatically configure the VRF-Lite inter-domain connection at the ACI side
- provide the configuration of the next-hop device connected to the SD-Access border nodes
- provide cross-domain visibility

Macro Segmentation

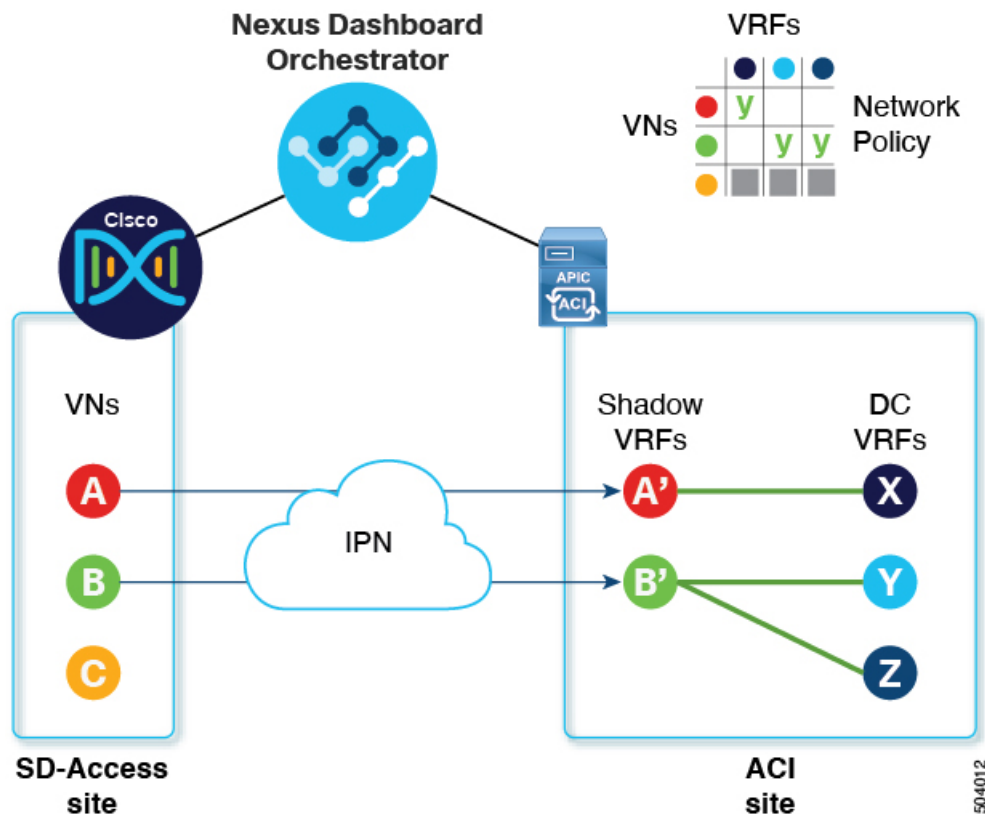
The Cisco SD-Access and Cisco ACI integration feature of Cisco Nexus Dashboard Orchestrator (NDO) allows macro segmentation of network elements between the ACI domain and the SD-Access domain.

In the ACI domain, entities such as EPGs, subnets, and VLANs are grouped as part of Virtual Routing and Forwarding instances (VRFs). When a VRF requires external communication, the VRF is associated with an IP interface (L3Out) on an ACI Border Leaf (BL). In the SD-Access domain, entities such as users, subnets, and IP pools can be grouped as Virtual Networks (VNs). When a VN requires external communication, the VN is associated with an SD-Access Border Node (BN) interface for IP handoff. The border interfaces of the two domains, ACI and SD-Access, can be physically connected through an IP network (IPN), but this basic connection does not provide connectivity between VRFs and VNs. With Cisco Nexus Dashboard Orchestrator Cisco SD-Access and Cisco ACI integration, an administrator can create policies to map (or "stitch") VRFs to VNs.

Macro Segmentation Workflow

A typical Cisco SD-Access and Cisco ACI integration workflow consists of the following steps, which refer to the figure below.

Figure 32: Macro Segmentation using NDO for SD-Access-to-ACI integration



- In an existing SD-Access site, a Cisco Digital Network Architecture Controller (DNAC) administrator has configured a campus fabric in which some entities require external access, such as access to the data center. The DNAC administrator has performed the following tasks:

- created virtual networks (VNs)
- associated IP address pools to those VNs
- configured L3 border nodes and associated interfaces
- created an IP (layer 3) handoff transit network
- configured Layer-3 handoff for those VNs that need external connectivity

Note that these tasks are normal DNAC administrative tasks and no special modification has been made for Cisco SD-Access and Cisco ACI integration.

- An NDO operator logs into and onboards the DNAC, using the DNAC credentials.

In the onboarding process, NDO automatically accesses the DNAC's REST API to query for sites, VNs, and border node devices. Upon discovering these entities, NDO learns which VNs are configured for external connectivity (L3 handoff) and on which Cisco SD-Access border nodes, and it learns their subnets. In the example shown in [Figure 32: Macro Segmentation using NDO for SD-Access-to-ACI integration, on page 312](#), the VNs A and B have been configured for L3 handoff, and these VNs are available to be extended into the ACI site. VN C is not configured for L3 handoff, and will not be available to the ACI site.

NDO continues to periodically query the DNAC for any ongoing configuration changes in the SD-Access fabric.

- The NDO operator configures connectivity between one or more ACI sites to one or more SD-Access sites. This includes specifying ACI site border leaf switches and interfaces, and VLAN and IP pools to be used for VRF-Lite configuration on border leaf interfaces. For directly connected interfaces (no IPN), the VRF-Lite configuration is derived from the configuration provisioned by DNAC for the IP handoff on the SDA border nodes, and VLANs and IP addresses are not taken from these pools.

NDO generates and displays next-hop device configuration for extended SD-Access VNs. This configuration can be applied manually to IPN devices if needed. NDO does not provision the IPN devices.

- The NDO operator extends a VN into the datacenter, making the VN available for connection to VRFs within the ACI domain.

Extending a VN creates an internal representation (a mirrored "shadow VRF") of the VN to represent the VN on the ACI domain. In the example in Figure 1, shadow VRFs A' and B' are automatically created in the ACI site to represent extended SD-Access VNs A and B. These shadow VRFs are stretched across all sites and pods within the ACI domain that require connectivity with the SD-Access domain. NDO automatically creates a schema and template in which these shadow VRFs are configured. The auto-created schema and template appear in NDO, but are read-only. The template is associated to the 'common' tenant and is associated with all 'SDA-Connectivity' enabled sites.

- The NDO operator creates a network policy to map an extended SD-Access VN to a datacenter VRF or VRFs that the VN needs to access. This action is also called "VRF stitching." The datacenter VRFs can be part of different "App tenants", which implies that this integration by design allows you to establish inter-VRF connectivity (a functionality usually referred to as "shared services").

In the example in Figure 1, the network policy shown stitches extended SD-Access VN A (extended as VRF A') to datacenter VRF X, and VN B (extended as VRF B') to datacenter VRFs Y and Z.

As a result of this mapping, a security policy relationship allowing all traffic is automatically established between the external EPG of the L3Out associated to the extended SD-Access VN and the vzAny logical object representing the datacenter VRF. The application of this contract allows free connectivity between

all the subnets of the extended SD-Access VN and all the subnets of the datacenter VRF that have been explicitly configured to be leaked across VRFs.

Cisco SD-Access and Cisco ACI Integration Guidelines

- An ACI site and an SD-Access site can be connected indirectly, through an external IP network (IPN), or directly, with back-to-back connections from ACI border leafs to SD-Access border nodes.
 - If the sites are directly connected, the connectivity between the two domains is configured automatically, including both the control plane and data plane.
 - If the sites are connected using an IPN, the IPN devices must support VRF Lite. NDO and DNAC do not provision the IPN devices, but NDO provides a sample configuration that can be applied to the IPN devices directly connected to the ACI border leafs and to the SD-Access border nodes.
- When multiple sites exist in either domain, note the following guidelines:
 - An SD-Access site can use another SD-Access site (SDA transit) to connect to the ACI sites.
 - When multiple sites exist in the SD-Access (campus) domain, each campus site can connect directly to the datacenter domain (direct peering), or through an intermediate network that could be a generic IP network (IPN), or through another campus site (indirect peering).
 - In a Multi-Site deployment, each ACI fabric that requires direct or indirect connectivity with the SD-Access (campus) domain must deploy a local L3Out connection. If the ACI fabric is a Multi-Pod fabric, the L3Out connection can be deployed only in a pod or a subset of the pods that are part of the same fabric.
- M:N mapping of VNs to VRFs is supported, within the limits described in [Scalability of SD-Access and ACI Integration, on page 315](#).
- M:N mapping of SD-Access sites to ACI sites is supported, within the limits described in [Scalability of SD-Access and ACI Integration, on page 315](#).
- From the DNAC, NDO learns about all SD-Access (campus) VNs and their subnets. When a VN is extended into the ACI site, NDO assumes that all subnets of that extended VN are reachable from ACI border leafs. NDO periodically checks for the presence of these subnets on ACI border leafs. In the **Status** column of the **Integrations > DNAC > Virtual Networks** table for an extended VN, NDO reports the subnets that are not yet reachable.
- By default, when an extended VN is mapped to a DC VRF, the ACI site does not advertise transit routes to the VN. The NDO administrator controls which ACI subnets are leaked into the shadow VRF of the VN as follows:
 - BD subnets that are internal to the ACI VRF are leaked only if the subnets are configured with “Shared between VRFs”.



Note When an SD-Access VN is mapped to multiple ACI VRFs, only non-overlapping prefixes across all mapped ACI VRFs should be configured as “shared between VRFs”.

- External subnets learned from L3Outs configured in the ACI VRF are leaked only if the subnets are configured with “Shared Route Control” and if transit routing is enabled.

For detailed information, see [Configuring Transit Routing, on page 323](#).

- The SD-Access site cannot provide Internet connectivity to the ACI site.
- Automation of IPv6 connectivity is not supported.
- Multicast traffic is not supported between the domains.

Scalability of SD-Access and ACI Integration

- Only a single DNAC can be onboarded to your NDO for SD-Access and ACI integration.
- Multiple SD-Access (campus) sites are supported if managed by a single DNAC.
- Up to 2 ACI sites are supported for peering with SD-Access. Each ACI site can be a single Pod fabric or a Multi-Pod fabric.
- A virtual network (VN) can be mapped to a maximum of 10 ACI VRFs.
- Up to 32 virtual networks (VNs) from the SD-Access domain can be extended into the ACI domain.

Software Compatibility

The minimum software versions that support macro segmentation for SD-Access and ACI integration are listed in the following table.

Product	Supported Product Versions
NDO	3.7 and later releases
ACI	4.2 and later releases
DNAC	2.3.3 and later releases

Onboarding the DNA Center

This section describes how to configure a Cisco Nexus Dashboard Orchestrator (NDO) to sign in to a DNA Center (DNAC). After signing in, NDO can import the SD-Access site configuration information necessary to create a network connection between the SD-Access domain and an ACI domain.

Procedure

-
- Step 1** Log in to your NDO.
- Step 2** From the left navigation pane, select **Admin > Integrations > DNAC**.
- Step 3** In the main pane, click **Add DNAC** to on board a DNA Center.
The **Add DNAC** dialog box opens.
- Step 4** In the **Add DNAC** dialog box, perform the following steps:

- a) Enter a **Name** for the DNA Center.
- b) Enter the URL or IP address of the DNA Center as the **Device IP**.
- c) Enter a **Username** credential for signing in to the DNA Center.

Read-only access is sufficient.

- d) Enter a **Password** credential for signing in to the DNA Center.
- e) Enter the password again in **Confirm Password**.
- f) Click **Add**.

NDO automatically signs in to the DNAC through the REST API and queries for the configuration of virtual networks (VNs) and border node devices in the SD-Access domain that is controlled by the DNAC.

What to do next

- Configure connectivity from the ACI site to the SD-Access site or IPN.
- Create network policies to allow communication between VNs in the DNAC's SD-Access domain and VRFs in the ACI domain.

Configuring Connectivity Toward the SD-Access Domain

This section describes the infrastructure-level configuration that is performed on NDO for Cisco SD-Access to ACI integration. For each ACI fabric, you must select the border leaf nodes and their associated interfaces that provide connectivity toward the Cisco SD-Access domain.

Before you begin

You must have on board the DNA Center.

Procedure

-
- Step 1** Log in to your Cisco Nexus Dashboard Orchestrator.
 - Step 2** From the left navigation pane, select **Admin > Integrations > DNAC**.
 - Step 3** In the main pane, click the **Overview** tab.
A dashboard of DNA Center appears.
 - Step 4** On the right side of the **DNAC Details** box, click the link for **Configuring Connectivity**.
The **Fabric Connectivity Infra** page appears.
 - Step 5** From the left navigation pane, under **Sites**, select the ACI site to be connected.
An **Site Connectivity** pane appears on the right.
 - Step 6** From the **Site Connectivity** pane, scroll down to the **SDA Connectivity** control and set it to **Enabled**.
Several fields appear below the **SDA Connectivity** control. Configure the settings in the following substeps.

- a) From the **External Routed Domain** drop-down list, choose the external routed domain (L3 domain) to be connected. This routed domain must be already defined on APIC.

- b) In the **VLAN Pool** field, enter a range of VLAN numbers.

A VLAN number from this pool will be assigned to the subinterfaces or SVIs when extending a campus VN to the data center. The VLAN pool must be the same as, or a subset of, the VLAN pool that is associated to the external routed domain you selected in the previous step.

If the ACI to SD-Access connection is back-to-back, with no IPN, the VLAN ID is not assigned from this pool. Instead, the VLAN ID is determined by what has been provisioned by DNAC for the IP handoff on the SD-Access border nodes.

- c) Under **VRF Lite IP Pool Ranges**, click the + symbol next to **Add VRF Lite IP Pool Range** and enter an IP subnet in the **IP Address** field.

IP addresses from this subnet will be assigned to the subinterfaces or SVIs when extending a campus VN to the data center.

If the ACI to SD-Access connection is back-to-back, with no IPN, these pools are not used. In this case, the IP addresses for the subinterfaces are determined by what has been provisioned by DNAC for the IP handoff on the SD-Access border nodes.

Step 7

In the center pane that displays the pods of the ACI site, click **Add Leaf Node** under the pod that will connect to the SD-Access site.

A **Select a Leaf** pane appears on the right. Configure the settings in the following substeps.

- a) From the **Leaf Node** drop-down list in the **Select a Leaf** pane, choose the border leaf switch that connects to the SD-Access domain.
- b) In the **Router ID** field, enter the border leaf switch router ID.
- c) Under Interfaces, click the + symbol next to **Add Interface**

The **Add Interface** dialog box appears.

- d) Enter the **Interface ID**.
- e) From the **Interface Type** drop-down list, select either **Sub-Interface** or **SVI**.
- f) Enter the **Remote Autonomous System Number**.

If the ACI to SD-Access connection uses an IPN, this number should match the ASN of the IPN.

If the ACI to SD-Access connection is back-to-back, with no IPN, this number should match the ASN of the SD-Access border nodes.

- g) Click **Save**.

Step 8

In the top bar of the **Fabric Connectivity Infra** page, click **Deploy**.

At this point, the configuration is not yet pushed to APIC. When the first VN is extended, the SD-Access connectivity is configured automatically.

Viewing the Status of the SD-Access to ACI Integration

The **Integrations > DNAC** menu displays details about the integration status and provides an inventory of available virtual networks (VNs).

Overview Tab

The **Overview** tab displays the following information windows:

- **DNAC Details:** Displays the overall status, IP address, and version of the connected DNAC. This window also contains a link to **Configure Connectivity**.
- A summary graphics dashboard for the following resources:
 - **DNAC Enabled Sites:** The number and type of SD-Access sites under management by the DNAC. The supported site types are on-premises, AWS, and NDFC.
 - **Virtual Networks:** The number of available VNs, and how many are extended or not extended.
 - **DC VRFs:** The number of datacenter VRFs available for sharing, and whether they are mapped or unmapped.

Virtual Networks Tab

Click the **Virtual Networks** tab to display details about the VNs.

The top window of the page repeats the summary graphics information from the **Overview** tab.

The **Virtual Networks** window of the page lists the virtual networks (VNs) that have been configured by DNAC for IP handoff on the SD-Access border nodes. A table of VNs displays the following information for each VN:

- **Status:** The current integration status of the VN, along with a color-coded icon indicating the severity of the status. The states are listed in the following table.

Status	Icon color (severity)	Description
Discovered	Green (Normal)	VN is discovered on SDA Border Nodes.
InProgress	Grey (Informational)	Reading the latest status of the VN after a configuration change. This is a temporary state. Tip You can click the Refresh icon in the upper right corner of the page to force an immediate polling of the status.
Success	Green (Normal)	VN is successfully extended.
BGPSessionIssues	Yellow (Warning)	BGP sessions are not established on all interfaces. Check each DC border leaf status for details.
RouteLeakPartial	Yellow (Warning)	VN subnets are partially propagated to the DC border leaf nodes. Check each DC border leaf status for details.

Status	Icon color (severity)	Description
RouteLeakNone	Red (Failure)	VN subnets are not yet propagated to the DC border leaf nodes. Click DC Sites in the VN table to check DC border leaf interfaces for issues.
MapVRFConfigFailure	Red (Failure)	Configuration failed on mapped VRFs. Retry the mapping.
DCSiteConfigFailure	Red (Failure)	VN extension failed on DC sites. Unextend the VN and extend again.

Click the status icon of a VN to display a sidebar containing additional details that can be helpful in troubleshooting warnings and failures.

- **Name:** The name assigned to the VN by the DNAC administrator.
- **Extended:** Indicates whether the VN has been extended.
- **DC Mapped VRFs:** The number of datacenter VRFs to which the VN is mapped. Click this number to open a sidebar displaying the associated schema, template, and tenant of mapped datacenter VRFs.
- **DC Sites:** The number of datacenter sites to which the VN is mapped. Click this number to open a sidebar displaying details of the datacenter sites, including the border leaf interfaces, BGP peering status, and next-hop device information.



Tip For IPN-connected border leaf interfaces, in the sidebar under "Peer Device Configuration", click "Show Details" for a sample configuration of an IPN device connected to this site.

- **Campus Sites:** The number of campus sites associated to this VN. Click this number to open a sidebar displaying details of the campus sites, including the border node interfaces, BGP peering status, and next-hop device information.



Tip For IPN-connected border node interfaces, in the sidebar under "Peer Device Configuration", click "Show Details" for a sample configuration of an IPN device connected to this site.

- **... (actions icon):** Click the icon to access actions for this VN.

The available actions depend on the current status of the VN, but may include the following:

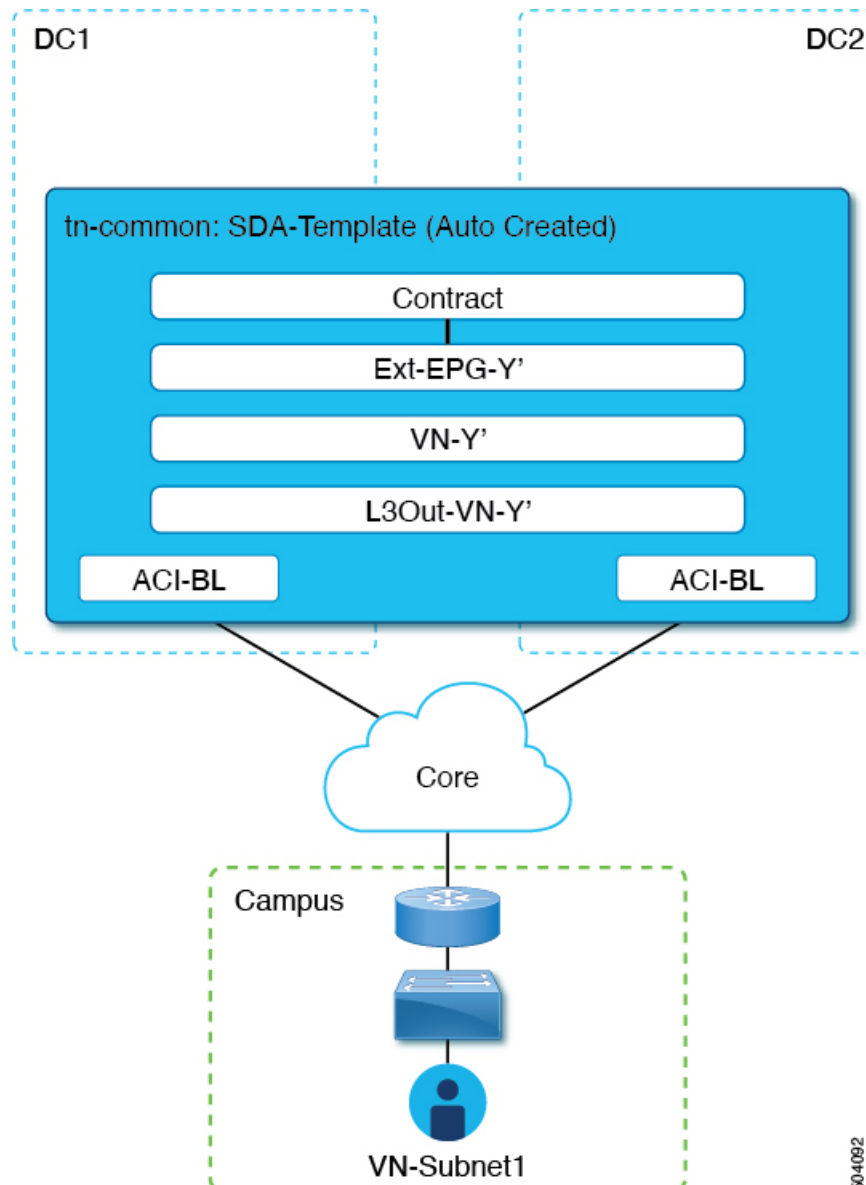
- Extend/Unextend the VN
- Map/Un-Map DC VRFs
- Enable/Disable Transit Route

The **Associated Templates** window of the **Virtual Networks** page appears when you have mapped a campus VN to a datacenter VRF.

Extending a Virtual Network

This section describes how to extend an SD-Access (campus) VN to the ACI (data center) fabrics. This action results in the creation of a VRF (and other associated configuration objects that are shown in [Figure 33: Extending a VN, on page 320](#)) representing the mirrored image of the campus VN on the DC side. The created objects are defined in an autogenerated template that is associated to the 'common' tenant.

Figure 33: Extending a VN



Before you begin

- You must have on board the DNA Center (DNAC).

- You must have configured connectivity to the SD-Access domain at the ACI site level.

Procedure

-
- Step 1** Log in to your Cisco Nexus Dashboard Orchestrator.
- Step 2** From the left navigation pane, select **Admin > Integrations > DNAC**.
- Step 3** In the main pane, click the **Virtual Networks** tab.
- A table of Virtual Networks (VNs) appears, displaying all VNs that have been configured by DNAC for IP handoff on the SD-Access border nodes.
- Step 4** In the row of the VN to be extended, click the actions menu (...) and select **Extend**.
- A dialog box opens, displaying the ACI sites and interfaces to which the VN will be extended. This information reflects the configuration settings in [Configuring Connectivity Toward the SD-Access Domain, on page 316](#).
- If you wish to revoke the extending of the VN later, click the actions menu (...) and select **Unextend**.
- Step 5** In the dialog box, click **Extend**.
- The VN is extended to all ACI sites where SD-Access connectivity is enabled, but it is not yet mapped to any ACI VRFs.
- Step 6**
-

What to do next

Verify the BGP Peering Status of the ACI border leaf switch interfaces:

- If the SD-Access border nodes and the ACI border leafs are connected directly (back-to-back), verify that, because of extending the campus VN, BGP sessions have been established between these devices. In **Admin > Integrations > DNAC > Virtual Networks**, click the **DC Sites** number to open a sidebar that displays details of the ACI border leaf switch interfaces. Check that the BGP Peering Status of the border leaf switch interfaces indicates **Up**.
- If an IPN is deployed between the domains, retrieve the configuration samples to help with configuring the next-hop devices that are directly connected to the SD-Access border nodes and to the ACI border leafs. In **Admin > Integrations > DNAC > Virtual Networks**, click the **DC Sites** number to open a sidebar that displays details of the ACI border leaf switch interfaces. For IPN-connected border leaf switch interfaces, click the "Show Details" link next to Peering Device Configuration to display a sample IPN device configuration. After configuring the IPN devices, check that the BGP Peering Status of the border leaf switch interfaces indicates **Up**.

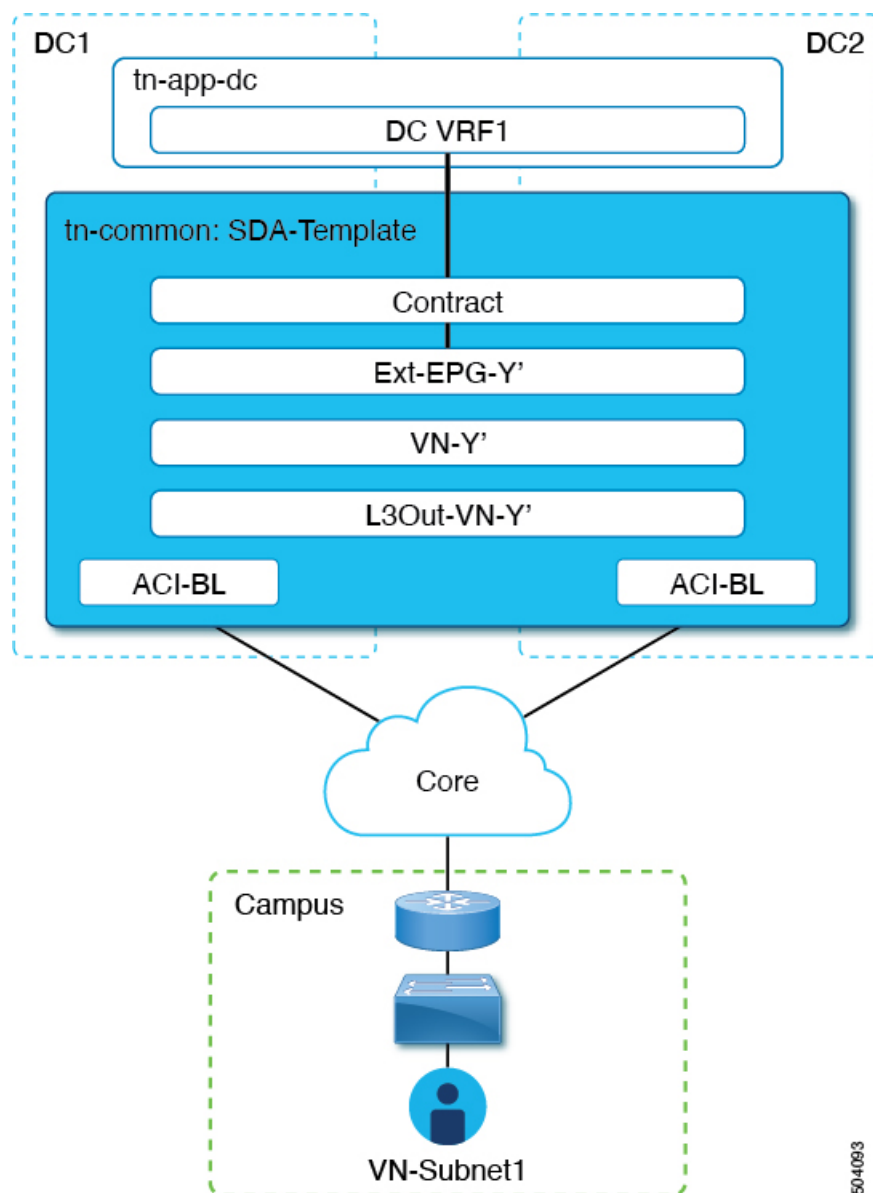
Map the extended VN to one or more ACI VRFs, as described in [Mapping or Unmapping a VN to a VRF, on page 321](#).

Mapping or Unmapping a VN to a VRF

This section describes how to map ("stitch") a virtual network (VN) to one or more data center (DC) VRFs in the ACI fabric. The mapping to a VRF results in the establishment of a contract relationship between the

DC VRF (represented by the "vzAny" object) and the external EPG previously provisioned in the 'common' tenant, as shown in [Figure 34: Mapping to a VRF](#), on page 322:

Figure 34: Mapping to a VRF



Before you begin

You must have extended the VN into the ACI site.

Procedure

Step 1 Log in to your Nexus Dashboard Orchestrator.

- Step 2** From the left navigation pane, select **Admin > Integrations > DNAC**.
- Step 3** In the main pane, click the **Virtual Networks** tab.
- A table of Virtual Networks (VNs) appears, displaying all VNs that have been configured by DNAC for IP handoff on the SD-Access border nodes.
- Step 4** In the row of the VN to be mapped, click the actions menu (...) and select **Map/Un-Map DC VRFs**.
- A **Map/Un-Map DC VRFs** dialog box opens.
- Step 5** In the **Map/Un-Map DC VRFs** dialog box, click the + icon next to **Add Mapped DC VRF**.
- Step 6** From the drop-down list of VRFs, choose a VRF.
- The selected VRF is added to a table that also displays the template for the VRF. Note the template name, as it will be needed in a later step.
- If you wish to map the VN to additional VRFs, click the + icon again to choose additional VRFs from the drop-down list.
- You can also un-map a DC VRF by deleting the existing mapping. To un-map a DC VRF, click the trash icon in the row of the VRF.
- Step 7** Click **Save** and wait until the VN Status has changed to 'Success'.
- Note**
At this point, even if the VN Status indicates 'Success', data connectivity is not yet established between the extended VN and the DC VRF. The mapping operation has modified a template that is associated with the mapped VRF, and you must redeploy the template before connectivity is established. In the **Associated Templates** table under the VN table, the template that is associated with the mapped VRF appears.
- Step 8** In the **Associated Templates** table in the **Admin > Integrations > > DNAC > Virtual Networks** tab, click the link of the template that is associated with the mapped VRF.
- The schema and template page opens.
- Step 9** In the schema and template page, click **Deploy to sites**.
- Step 10** If template review and approval (change control) are enabled, follow the change control workflow to redeploy the template. Otherwise, click **Deploy** to redeploy the template.

What to do next



Note If you have unmapped a DC VRF, no template is displayed in the **Associated Templates** table. However, you must still go to **Configure > Tenant Template > Applications**, select Schemas to redeploy the associated template to remove the vzAny configuration. Otherwise, data plane communication remains enabled.

Configuring Transit Routing

When an extended SD-Access (campus) VN is mapped to an ACI (data center) VRF, any BD subnets of the DC VRF that are configured with “Advertised Externally” and “Shared between VRFs” flags are leaked into

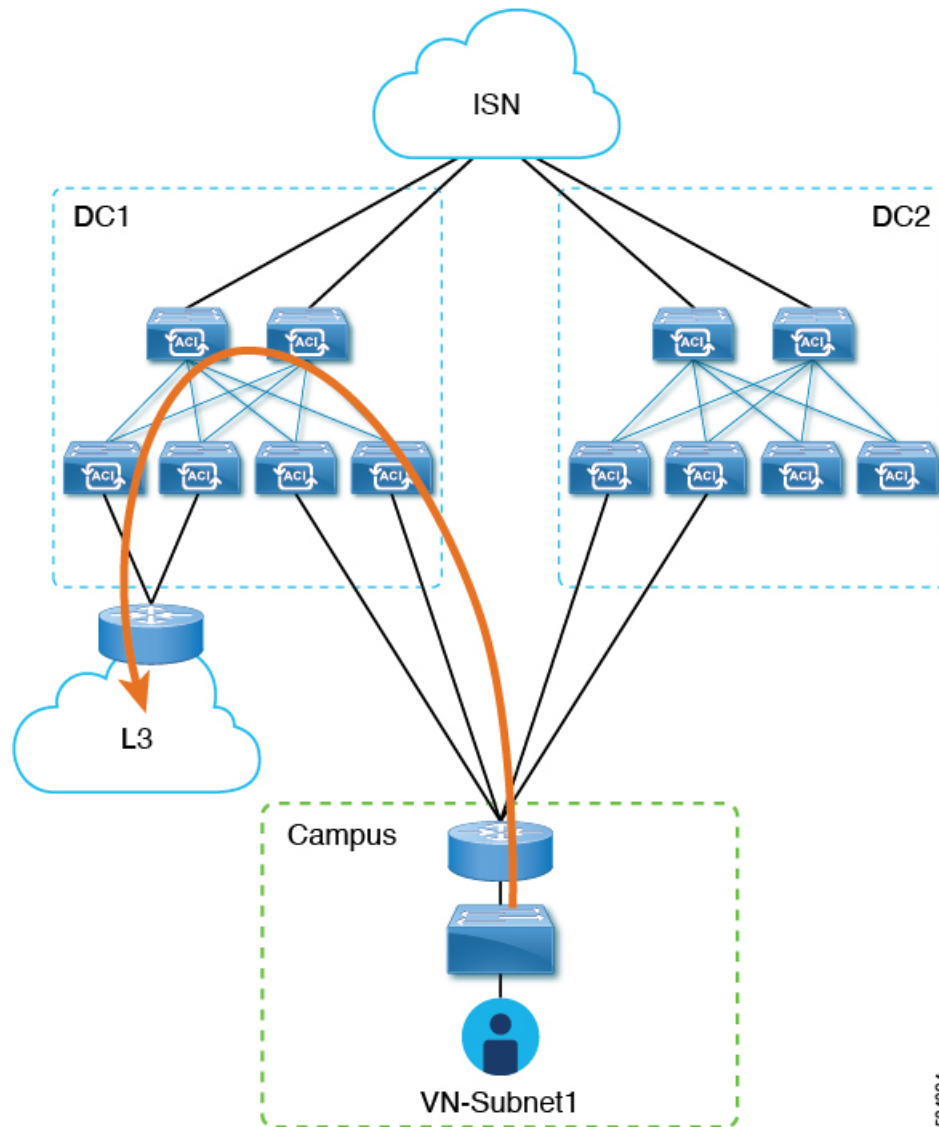
the ‘common’ tenant VRF and then advertised toward the SD-Access domain. This ensures that campus users can gain access to the applications provisioned in the DC VRF.



Note When an SD-Access VN is mapped to multiple ACI VRFs, only nonoverlapping prefixes across all mapped ACI VRFs should be configured as “shared between VRFs”.

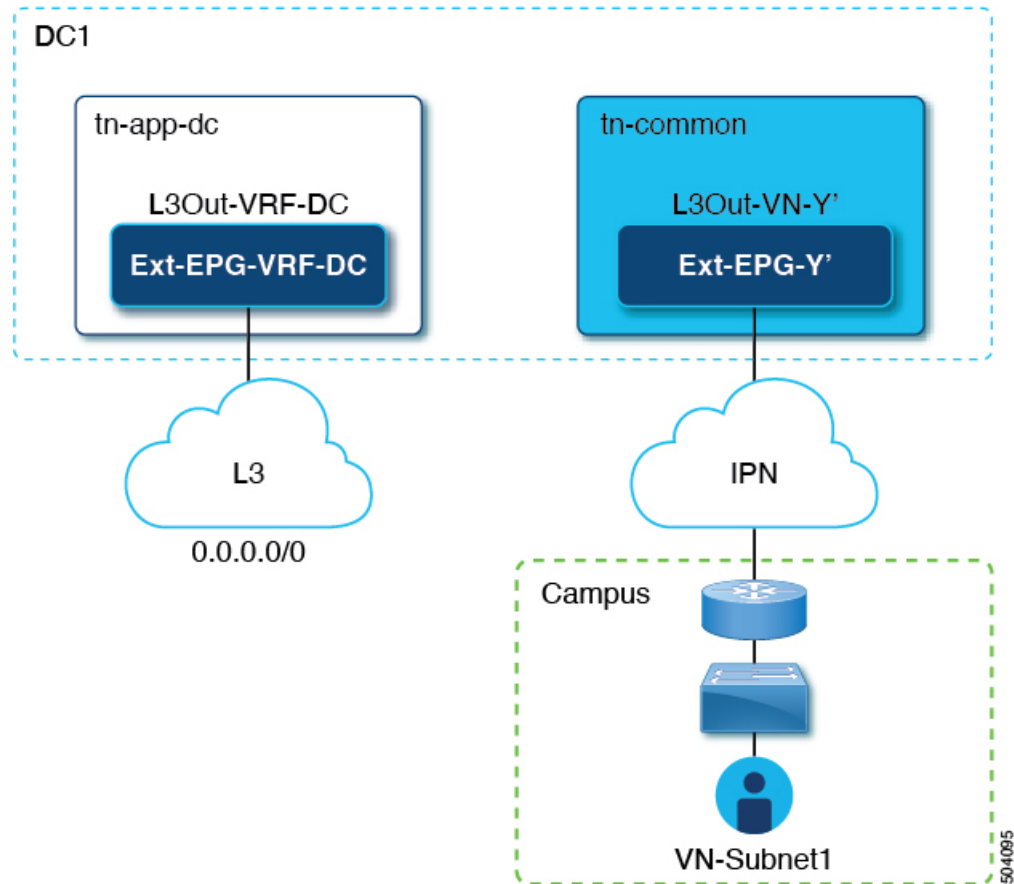
In addition to the advertisement of those BD subnets, there may be a requirement for campus users to access an external L3 network domain using the ACI domain as transit (Figure 35: ACI Domain as Transit, on page 324).

Figure 35: ACI Domain as Transit



In this scenario, an L3Out connection that is associated to the DC VRF (L3Out-DC-VRF) is provisioned for allowing connectivity to the external domain and external routes (a simple 0.0.0.0/0 default in the example in [Figure 36: L3Out Connection, on page 325](#)) are imported in the DC VRF routing table (part of tn-app-dc).

Figure 36: L3Out Connection



To ensure that campus users can connect to the external L3 domain through the data center, the external routes must be leaked to the uncommon VRF so that they can be advertised toward the campus domain through the L3Out connection (L3Out-VN-Y') autogenerated because of the campus VN extension to the DC.

Follow this procedure to enable leaking of the external routes:

Before you begin

You must have mapped an extended campus VN to a data center VRF and established connectivity.

Procedure

- Step 1** Log in to your Cisco Nexus Dashboard Orchestrator.
- Step 2** From the left navigation pane, select **Admin > Integrations > DNAC**.
- Step 3** In the main pane, click the **Virtual Networks** tab.
- Step 4** In the row of a successfully mapped campus VN, click the actions menu (...) and select **Enable Transit Route**.

This configuration ([Figure 37: Export Route Control, on page 326](#)) creates a 0.0.0.0/0 prefix under Ext-EPG-Y', with the following "Route Control" flags set that allows the advertising toward the IPN of all external routes leaked from the tn-app-dc tenant.

Figure 37: Export Route Control

Update Subnet 0.0.0.0/0

Subnet *
0.0.0.0/0

Route Control

☒ Export Route Control

☐ Import Route Control

☐ Shared Route Control

Aggregate

☒ Aggregate Export

External EPG Classification

☒ External Subnets for External EPG

☐ Shared Security Import

To disable transit routing, click the actions menu (...) and select **Disable Transit Route**.

Note

With either setting (enabled or disabled), the campus site has access to shared BD subnets internal to the ACI VRF.

- Step 5** From the left navigation pane, choose **Configure > Tenant Template > Applications > Schemas** and navigate to the template for configuring the data center tenant application.
- Step 6** In the data center tenant application template, configure the flags under the 0.0.0.0/0 prefix that is associated to Ext-EPG-VRF-DC of the DC VRF to be able to leak into uncommon the external routes learned from the Internet ([Figure 38: Shared Route Control, on page 326](#)).

Figure 38: Shared Route Control

Update Subnet 0.0.0.0/0

Subnet *
0.0.0.0/0

Route Control

☐ Export Route Control

☐ Import Route Control

☒ Shared Route Control

Aggregate

☒ Aggregate Shared Routes

External EPG Classification

☒ External Subnets for External EPG

☒ Shared Security Import

Note

The setting that is shown ensures that all the external prefixes that are received on L3Out-VRF-DC are leaked to tm-common and are therefore advertised toward the campus domain. This setting also allows leaking of the 0.0.0.0/0 default route if it is received from the L3 domain. If needed, you can apply a more granular configuration where only a subset of the

external prefixes can be leaked to uncommon. This is achieved by creating specific entries matching those subsets of prefixes and applying to those entries the same flag configuration shown here.

Step 7 In the data center tenant application template, define a specific prefix under Ext-EPG-VRF-DC matching the campus VN subnet (or set of subnets) to be advertised toward the external L3 domain.

In the example shown in [Figure 39: Update Subnet, on page 327](#), this configuration is applied to a specific 192.168.100.0/24 prefix.

Figure 39: Update Subnet

Update Subnet 192.168.100.0/24

Subnet *
192.168.100.0/24

Route Control

☒ Export Route Control

☐ Import Route Control

☐ Shared Route Control

External EPG Classification

☐ External Subnets for External EPG

Note

Creating a separate prefix for a VN subnet provides the most granular level of control for the advertisement of campus VN subnets toward the external L3 domain. If such granular control is not needed, you can set the “Export Route Control” flags associated to the 0.0.0.0/0 prefix instead, which allows sending toward the external domain all the campus VN subnets that have been leaked into tn-app-dc from uncommon.



CHAPTER 28

SD-WAN Integration

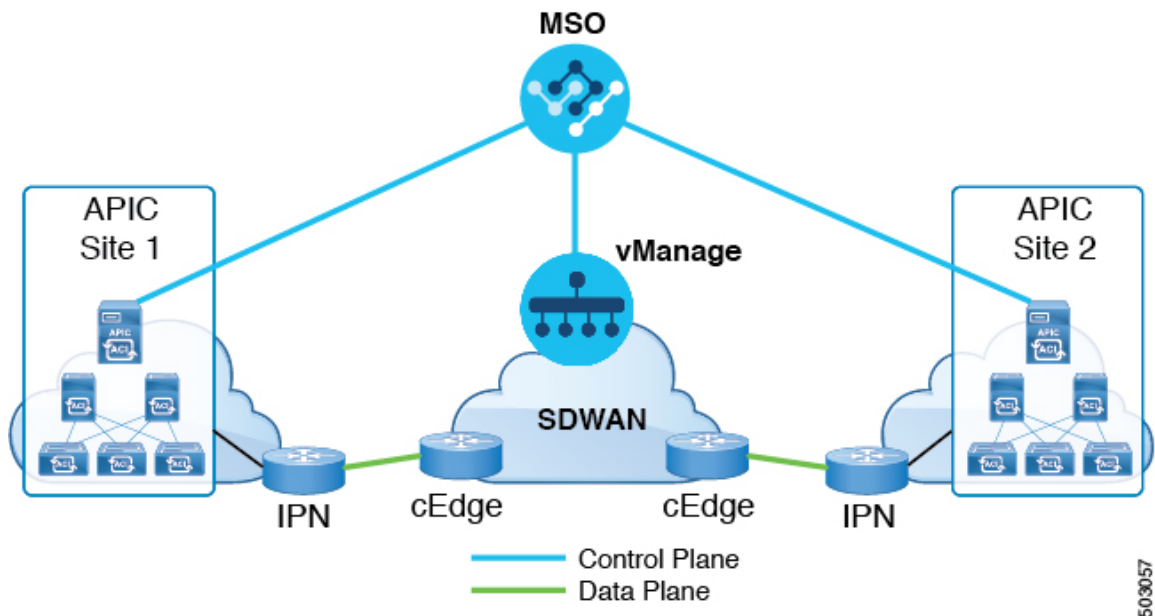
- [SD-WAN Integration, on page 329](#)
- [SD-WAN Integration Guidelines and Limitations, on page 330](#)
- [Adding a vManage Controller, on page 331](#)
- [Configuring Global DSCP Policy, on page 332](#)
- [Set QoS Level for EPGs and Contracts, on page 334](#)

SD-WAN Integration

Cisco Software-Defined Wide Area Network (SD-WAN) is a cloud-delivered overlay WAN architecture connecting branches to datacenter and multicloud environments through a single fabric. Cisco SD-WAN ensures predictable user experience for applications, optimizes SaaS, IaaS and PaaS connections, and offers integrated security either on-premises or in the cloud. Analytics capabilities deliver the visibility and insights necessary for you to isolate and resolve issues promptly and deliver intelligent data analysis for planning and what-if scenarios.

On the dataplane side, SD-WAN deploys an ASR or ISR routers as edge devices (shown as cEdge in the following diagram) with each fabric's spine switches connecting to these edge devices. SD-WAN is managed by a separate controller called vManage, which allows you to define service-level agreement (SLA) policies to determine how each packet's path within SD-WAN is chosen based on its DSCP value.

Figure 40: Multi-Site and SD-WAN Integration



Release 3.0(2) of Cisco Nexus Dashboard Orchestrator adds support for SD-WAN integration. You can configure the NDO to import SLA policies from a vManage controller, assign DSCP values to each SLA policy, and notify the vManage controller of the DSCP-to-SLA mapping. This enables you to apply preconfigured SLA policies to specify the levels of packet loss, jitter, and latency for intersite traffic over SD-WAN. The vManage controller, which is configured as an external device manager that provides SD-WAN capability, chooses the best possible WAN link that meets the loss, jitter, and latency parameters specified in the SLA policy.

Multi-Site SD-WAN integration allows traffic between multiple fabrics to traverse the SD-WAN network while enabling returning traffic from a remote site to retain the ACI QoS level assigned to it. After you register your Cisco NDO to vManage, it imports the SLA policies allowing you to translating the ACI QoS levels to the appropriate DSCP values. NDO then applies DSCP translation policy for traffic transiting SD-WAN to enable quality of service on the returning traffic.

Release 3.0(2) also enables you to assign ACI QoS levels to Contracts and EPGs directly in the NDO GUI. Any time traffic leaves the fabric, its QoS level is translated into a DSCP value, which vManage uses to pick a path for the traffic through SD-WAN.

SD-WAN Integration Guidelines and Limitations

When enabling Multi-Site and SD-WAN integration, the following guidelines apply.

- To enable uniform user QoS Level and DSCP translation for east-west traffic across sites with Multi-Site SD-WAN integration, the spine switches in each fabric must be connected to the SD-WAN edge devices, either directly or via multiple hops.

This is in contrast with the existing implementation of APIC SD-WAN integration for north-south traffic where the leaf switches must be connected to the SD-WAN edge devices.

- Global DSCP policy is supported for on-premises sites only.

- SD-WAN integration is supported for Nexus Dashboard Orchestrator deployments in Cisco Application Services Engine only.

For more information, see the [Deployment Overview](#) chapter in the *Cisco Nexus Dashboard Orchestrator Installation and Upgrade Guide*.

- When defining the global DSCP policy, you must pick a unique value for each QoS Level.
- In addition to existing DSCP policy values, you can import up to four SLA policies from vManage with one of the following values: 41, 42, 43, 45, 47 and 49.
- SLA policies must be already defined in your Cisco vManage.
- When assigning QoS level, you can choose to assign it to a specific Contract or an entire EPG.

If multiple QoS levels could apply for any given traffic, only one is applied using the following precedence:

- Contract QoS level: If QoS is enabled in the Contract, the QoS level specified in the contract is used.
- Source EPG QoS level: If QoS level is not specified for the Contract, the QoS level set for the source EPG is used.
- Default QoS level: If no QoS level is specified, the traffic is assigned Level 3 QoS class by default.

Adding a vManage Controller

This section describes how to add vManage controller to your Cisco Nexus Dashboard Orchestrator to import any configured SLA policies.

Procedure

Step 1 Log in to your Cisco Nexus Dashboard Orchestrator GUI.

Step 2 Add a vManage Controller.

- a) Navigate to **Admin > Integration > SD-WAN**.
- b) Click **Add Domain Controller** name.

The **Add Domain** window opens.

Step 3 Provide the vManage controller information.

In the **Add Domain** window that opens, provide the following details:

- Name of the vManage domain to display in your NDO.
- The device's fully qualified domain name or IP address.
- Username and password that is used to sign in to the vManage controller.

Then click **Add** to save the vManage domain. After the vManage controller information is entered, it can take up to 1 minute before the list of existing SLA policies is displayed in the main pane:

What to do next

Define the global DSCP policy in your Cisco Nexus Dashboard Orchestrator, as described in [Configuring Global DSCP Policy, on page 332](#)

Configuring Global DSCP Policy

When traffic is sent and received within a Cisco ACI fabric, it is prioritized based on the ACI QoS Level, which is determined based on the CoS value of the VXLAN packet's outer header. When traffic exits the ACI fabric from a spine switch toward an intersite network, the QoS level is translated into a DSCP value which is included in the outer header of the VXLAN-encapsulated packet.

This section describes how to define the DSCP translation policy for traffic entering or exiting ACI fabric. This is required when traffic must transit through non-ACI networks, such as between multiple fabrics separated by SD-WAN, where devices that are not under Cisco APIC's management may modify the CoS values in the transiting packets.

Before you begin

- You must have added a vManage controller to your NDO, as described in [Adding a vManage Controller, on page 331](#).
- You should be familiar with Quality of Service (QoS) functionality within ACI fabrics.
QoS is described in more detail in [Cisco APIC and QoS](#).

Procedure

-
- Step 1** Log in to your Cisco Nexus Dashboard Orchestrator GUI.
- Step 2** Open the global DSCP policy configuration screen.
- a) Navigate to **Configure > Tenant Template > > Tenant Policies**.
 - b) Click **Global DSCP Policy** name.
- The **Edit Policy** window opens.
- Step 3** Update the global DSCP policy.

- a) Choose the DSCP value for each ACI QoS level.

Each drop-down contains the default list of available DSCP values and any values imported from the vManage SLA policies, for example `Voice-And-Video SLA (42)`.

- b) Choose the sites where you want to deploy the policy.

We recommend deploying the policy to all sites that are part of the Multi-Site domain in order to achieve a consistent end-to-end QoS behavior.

- c) Choose whether you want to enable the policy on each site when it is deployed.
d) Click **Save & Deploy**.

After you save and deploy, the DSCP policy settings will be pushed to each site. You can verify the configuration by signing in to the site's APIC and navigating to **Tenants > infra > Policies > Protocol > DSCP class-CoS translation policy for L3 traffic**.

What to do next

After you have defined the global DSCP policy, you can assign the ACI QoS Levels to EPGs or Contracts as described in [Set QoS Level for EPGs and Contracts, on page 334](#)

Set QoS Level for EPGs and Contracts

This section describes how to choose an ACI QoS level for traffic in your fabrics. You can choose to specify QoS for individual Contracts or entire EPGs.

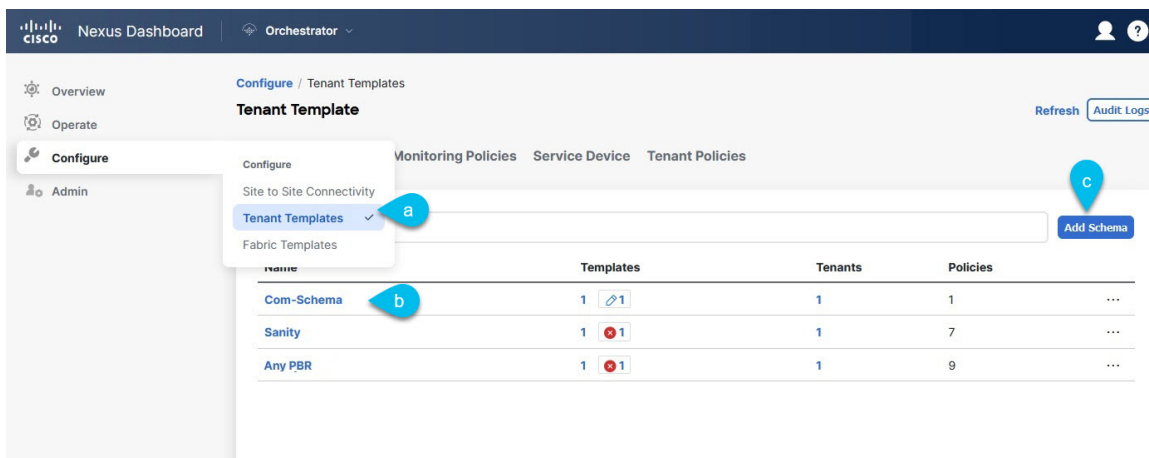
Before you begin

- You must have added a vManage controller to your NDO, as described in [Adding a vManage Controller, on page 331](#).
- You must have defined the global DSCP policy, as described in [Configuring Global DSCP Policy, on page 332](#).
- You should be familiar with Quality of Service (QoS) functionality within ACI fabrics.
QoS is described in more detail in [Cisco APIC and QoS](#).

Procedure

Step 1 Log in to your Cisco Nexus Dashboard Orchestrator GUI.

Step 2 Choose the Schema that you want to edit.



- Navigate to **Configure > Tenant Template > Applications > Schemas**
- Click the name of the schema you want to edit or **Create Schema** to create a new one.

The **Edit Schema** window opens.

Step 3 Pick a QoS Level for an EPG.

The screenshot displays the configuration interface for EPGs and Contracts. The main pane on the left shows the 'Any PBR' configuration page with a 'Template Properties' tab. Under 'Template Summary', it shows 'Type: Application', 'Tenant: common', and 'Template Status: In Sync'. Below this, there are sections for 'Application Profile Any-PBR', 'EPGs', 'Contracts', and 'VRFs'. The 'EPGs' section shows two EPGs: 'EPG App' and 'EPG Web'. A blue callout 'a' points to the 'EPG Web' button. The 'Contracts' section shows 'Web-App'. The 'VRFs' section shows 'VRF1'.

The right pane shows the 'EPG Web' configuration page. It has a 'Name' field and an 'Add Contract' button. Under 'Properties', there are two tabs: 'On-Premises Properties' (selected) and 'Cloud Properties'. The 'Bridge Domain' is set to 'BD-Web'. Under 'Subnets', there is a 'Gateway IP' field and an 'Add Subnet' button. The 'USeg EPG' section has a checkbox. The 'Intra EPG Isolation' section has two radio buttons: 'Enforced' and 'Unenforced' (selected). The 'Intersite Multicast Source' section has a checkbox. The 'Include in Preferred Group' section has a checkbox. The 'Advanced Settings' section has a 'QoS Level' dropdown menu set to 'Level 1' and a 'QoS Policy' dropdown menu set to 'Select...'. A blue callout 'b' points to the 'QoS Level' dropdown. An 'Ok' button is at the bottom right.

- In the main pane, scroll down to the **EPG** area and select an EPG or click **Add EPG** to create a new one.
- In the right sidebar, scroll down to the **QoS Level** drop-down and choose the QoS Level you want to assign to the EPG.

You must choose the QoS level based on the previously configured Global DSCP policy to ensure that intersite traffic from the EPG is treated with the needed SLA across the SD-WAN network.

Step 4 Pick a QoS Level for an EPG.

The screenshot displays the configuration interface for setting QoS levels. On the left, the 'Contracts' section is expanded, showing a list of contracts with 'Web-App' selected (indicated by a blue callout 'a'). Below this, the 'VRFs' section shows 'VRF1' with 'vzAny Enabled'. The 'Bridge Domains' section shows 'BD-App', 'BD-Web', and 'FW-external'. The 'Filters' section is also visible. On the right, the 'Filter Chain' configuration is shown. The 'Name' field is 'Permit-Any'. The 'Directives' section has a '+ Create Filter' button. The 'Properties' section is expanded, showing 'On-Premises Properties' with a green checkmark. The 'QoS Level' dropdown is set to 'Level 1' (indicated by a blue callout 'b'). The 'Target DSCP' dropdown is set to 'Unspecified'. An 'Ok' button is at the bottom right.

- In the main pane, scroll down to the **Contract** area and select a Contract or click the + icon to create a new one.
- In the right sidebar, scroll down to the **QoS Level** drop-down and choose the QoS Level you want to assign to the Contract.

You must choose the QoS level based on the previously configured Global DSCP policy to ensure that intersite traffic between two EPGs is treated with the needed SLA across the SD-WAN network.



CHAPTER 29

Multi-Site and SR-MPLS L3Out Handoff

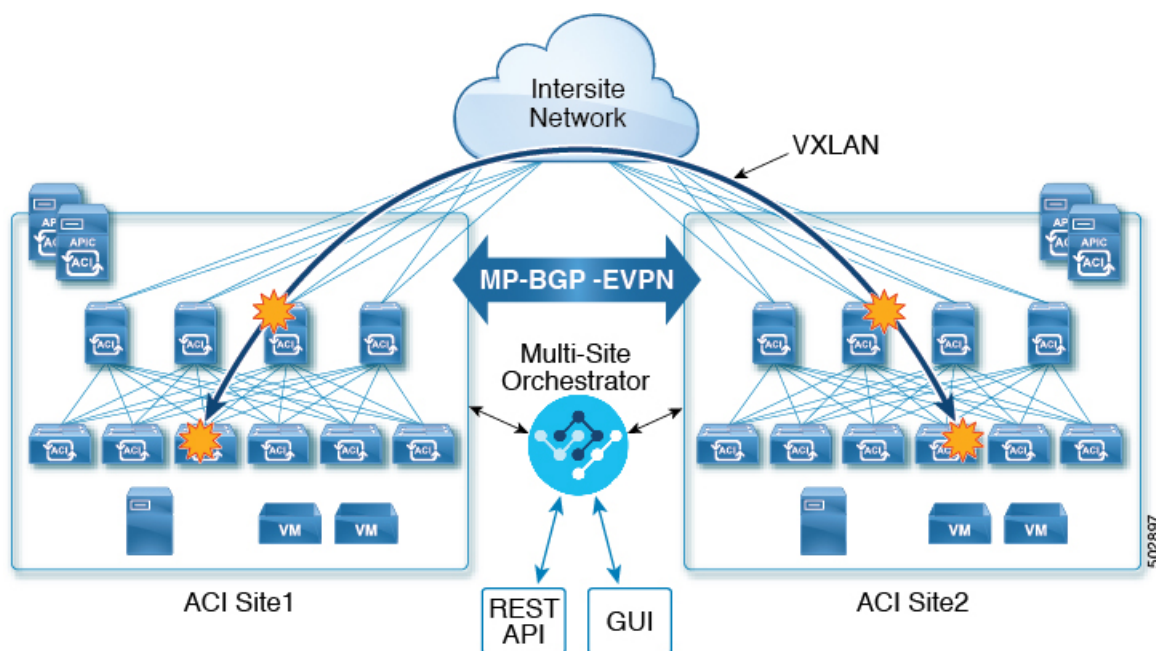
- [Overview and Use Cases, on page 337](#)
- [SR-MPLS Infra Requirements and Guidelines, on page 340](#)
- [SR-MPLS Tenant Requirements and Guidelines, on page 342](#)
- [Greenfield Deployment, on page 344](#)
- [Importing Existing SR-MPLSL3Out Configuration, on page 355](#)

Overview and Use Cases

Starting with Nexus Dashboard Orchestrator release 3.0(1) and APIC Release 5.0(1), the Multi-Site architecture provides better hand-off functionality between ACI border leaf (BL) switches and SR-MPLS networks.

In a typical Multi-Site deployment, traffic between sites is forwarded over an intersite network (ISN) via VXLAN encapsulation:

Figure 41: Multi-Site and ISN

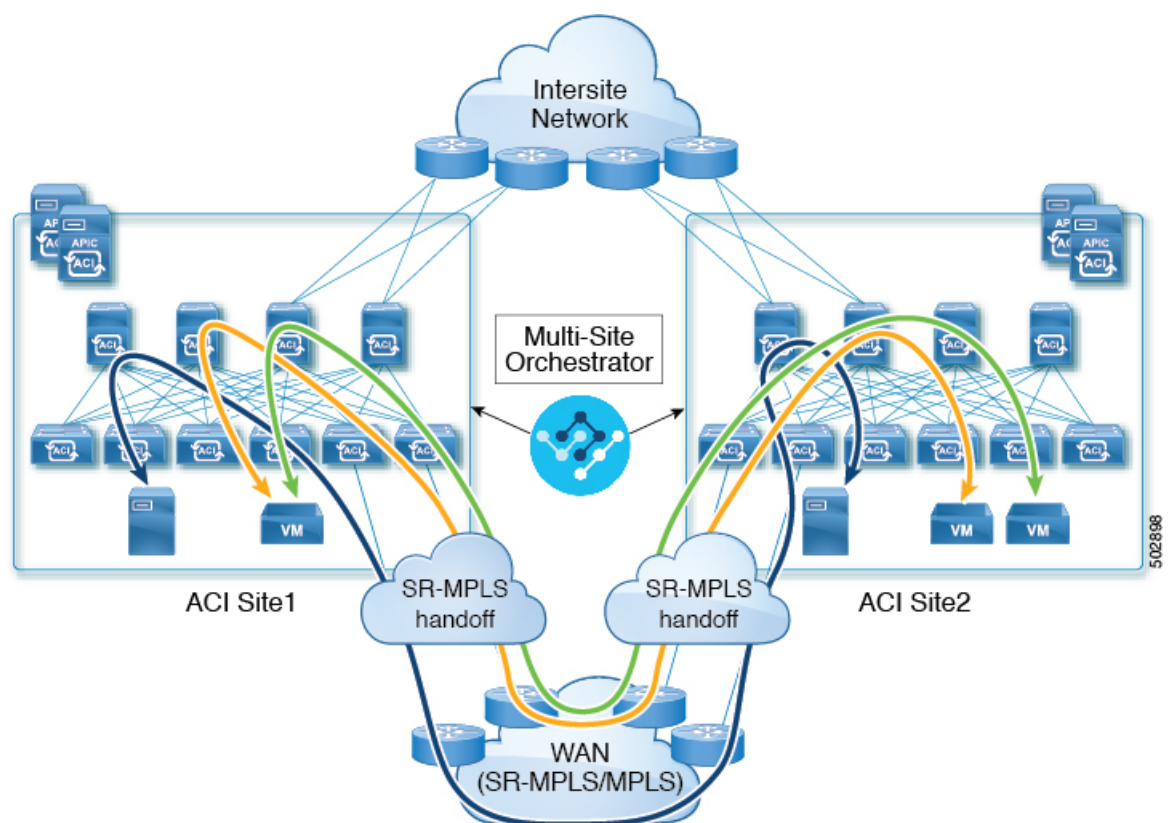


With Release 3.0(1), MPLS network can be used in addition to or instead of the ISN allowing inter-site communication via WAN, as shown in the following figure. In order to force East-West Layer 3 communication to follow the SR-MPLS L3Out data path (instead of the VXLAN data path across the ISN), several restrictions had to be applied to this SR-MPLS hand-off use case:

- The VRF to which the SR-MPLS L3Out belongs must not be stretched across sites.
- Because of the above restriction, every site must deploy one (or more) local SR-MPLS L3Outs for each defined site-local VRF.
- Contracts must not be applied between site-local EPGs belonging to different VRFs.

This forces the communication to follow the SR-MPLS L3Out data path.

Figure 42: Multi-Site and MPLS



Additional Use Cases in NDO Release 4.0(2) and Later

Prior to NDO release 4.0(2), if you wanted to deploy the SR-MPLS use case, you would define a special "SR-MPLS" template that could be associated with only a single site and not stretched across multiple sites. In this case, if you had two sites managed by your Nexus Dashboard Orchestrator and connected via an SR-MPLS network, and you wanted to establish communication between an EPG in `site1` and another EPG in `site2`, you had to deploy two separate SR-MPLS-VRF-L3Outs (one in each site) associated with two separate VRFs and establish contracts between the EPG in each site and that site's SR-MPLS L3Out (instead of directly between the EPGs). In other words, the EPGs' traffic would always use the SR-MPLS data path even for EPG-to-EPG communication across sites without integrating with the traditional Multi-Site data plane for East-West traffic.

Beginning with release 4.0(2), the SR-MPLS L3Outs can function similar to the traditional IP-based L3Outs which allows you to use the SR-MPLS L3Out hand-offs exclusively for North-South connectivity between a site and an external network, while all East-West traffic can be handled in the traditional Multi-Site manner using VXLAN-encapsulated data plane across the ISN. This means that the SR-MPLS hand-offs can now be treated as traditional IP-based hand-offs and the same VRF can deploy a mix of IP and SR-MPLS L3Outs. These changes add support for the following specific use cases:

- Deployment of multiple sites each with their own local SR-MPLS-VRF-L3Outs and intra-VRF traffic using the local L3Out if it is available or a remote SR-MPLS-VRF-L3Out from another site (intersite L3Out).

In this case, the remote SR-MPLS-VRF-L3Out can be used as a simple backup or to reach unique external prefixes received on the remote SR-MPLS-VRF-L3Out. Traffic will transit from a local EPG to the local SR-MPLS-VRF-L3Out and if that path is down or the route is unavailable it can take another site's remote SR-MPLS-VRF-L3Out.

- Similar use cases are supported for shared services, where application EPG in one VRF can use SR-MPLS-VRF-L3Out in a different VRF, either in the local or remote site.

In this case, the EPGs can be in a different tenant as well. For example, Tenant1 in Site1 can contain the application EPGs which will use an SR-MPLS-VRF-L3Out in Tenant2 in Site2.

- Ability to combine IP-based and SR-MPLS hand-offs.

Using SR-MPLS L3Outs (instead of traditional IP-based L3Outs) allows for operational simplification at higher scale by removing the need for the VRF-Lite configuration which requires creation of separate BL nodes, BL logical interfaces, and routing peering for each VRF that needs to be connected to the external network. With the SR-MPLS L3Outs, the logical nodes and logical interfaces are defined once in the infra tenant, together with a single MP-BGP EVPN peering with the external devices. This infra L3Out construct can then be used to provide external connectivity to multiple tenant VRFs and all the VRFs' prefixes are exchanged using the common MP-BGP EVPN control plane.

The following sections describe guidelines, limitations, and configurations specific to managing Schemas that are deployed to sites from the Nexus Dashboard Orchestrator. Detailed information about MPLS hand off, supported individual site topologies (such as remote leaf support), and policy model is available in the [Cisco APIC Layer 3 Networking Configuration Guide](#).

Configuration Workflow

Other sections in this document detail the required configurations, but in short you will go through the following workflow:

- Create an SR-MPLS QoS policy.

SR-MPLS Custom QoS policy defines the priority of the packets coming from an SR-MPLS network while they are inside the ACI fabric based on the incoming MPLS EXP values defined in the MPLS QoS ingress policy. It also marks the CoS and MPLS EXP values of the packets leaving the ACI fabric through an MPLS interface based on IPv4 DSCP values defined in MPLS QoS egress policy.

This is an optional step and if no custom ingress policy is defined, the default QoS Level (`Level3`) is assigned to packets inside the fabric. If no custom egress policy is defined, the default EXP value of 0 will be marked on packets leaving the fabric.

- Create an SR-MPLS Infra L3Out.

This configures an L3Out for traffic leaving a site that is connected to an SR-MPLS network.

The same SR-MPLS Infra L3Out can then be used by multiple SR-MPLS Tenant L3Outs for control and data plane communication with the external network domain.

- Create SR-MPLS route map policy that matches specific Tenant's prefixes.

Route maps are sets of `if-then` rules that enable you to specify which routes are advertised out of the Tenant SR-MPLS L3Out. Route maps also enable you to specify which routes received from the DC-PE routers will be injected into the BGP VPNv4 ACI control plane.

- If you want to deploy a use case similar to releases prior to release 4.0(2), create the VRF, SR-MPLS L3Out, and SR-External EPG for each site connected via an SR-MPLS network and establish a contract within each site between that site's tenant EPG and SR-External EPG.

In this case, all communication from one site will follow the North-South route egressing your Multi-Site domain towards the external SR-MPLS network. If the traffic is destined to an EPG in another site managed by your Orchestrator, it will ingress the other fabric from the external network using that site's SR-MPLS L3Out.

- If you want to use the SR-MPLS L3Outs in the same way as the standard IP-based L3Out exclusively for North-South communication, you can create the VRFs, SR-MPLS L3Outs, EPGs, and contracts as you typically would for all existing EPG-to-EPG communication use cases.

SR-MPLS Infra Requirements and Guidelines

If you want to use your Nexus Dashboard Orchestrator to manage SR-MPLS L3Out hand-offs for an ACI fabric connected to an SR-MPLS network:

- Any changes to the topology, such as node updates, are not reflected in the Orchestrator configuration until site configuration is refreshed, as described in [Refreshing Site Connectivity Information, on page 177](#).
- Any Multi-Site traffic across different sites cannot ingress or egress through Remote Leaf switches.
This limitation is not specific to SR-MPLS use cases and applies to all Multi-Site traffic in general.
- SR-External EPGs that are part of preferred group cannot be the providers of a shared service (inter-VRF) contract.
- Preferred Group is not supported for Intersite SR-MPLS L3Outs.
- vzAny is not supported as a shared service provider.
- VRF that is enabled for Preferred Group cannot be a vzAny consumer.
- We recommend configuring tenant contract objects under dedicated template to avoid circular dependencies with other configuration objects that use the same contracts
- When using SR-MPLS L3Out instead of traditional IP-based L3Outs:
 - Host-based routing advertisement is not supported for bridge domains that are stretched across sites.
 - Tenant Routed Multicast (TRM) is not supported with SR-MPLS L3Outs, so they can only be used for establishing Layer 3 unicast communication with the external network domain.

Supported Hardware

The SR-MPLS hand-off is supported for the following platforms:

- **Border Leaf switches:** The "FX", "FX2", "GX", and "GX2" switch models.
- **Spine switches:**
 - Modular spine switch models with "LC-EX", "LC-FX", and "GX" at the end of the linecard names.
 - The Cisco Nexus 9000 series N9K-C9332C, N9K-C9364C, "-GX", and "-GX2" fixed spine switches.
- **DC-PE routers:**
 - Network Convergence System (NCS) 5500 Series
 - ASR 9000 Series
 - NCS 540 or 560 routers

SR-MPLS Infra L3Out

You will need to create an SR-MPLS Infra L3Out for the fabrics connected to SR-MPLS networks as described in the following sections. When creating an SR-MPLS Infra L3Out, the following restrictions apply:

- Each SR-MPLS Infra L3Out must have a unique name.

The SR-MPLS Infra L3Out allows you to establish the control plane and data plane connectivity between the ACI border leaf switches and the external Provider Edge (PE) devices. SR-MPLS L3Outs that belong to various tenant VRFs can then leverage that Infra L3Out connectivity to establish communication with the external network domain.

- You can have multiple SR-MPLS infra L3Outs connecting to different routing domains, where the same border leaf switch can be in more than one L3Out, and you can have different import and export routing policies for the VRFs toward each routing domain.
- Even though a border leaf switch can be in multiple SR-MPLS infra L3Outs, a border leaf switch/provider edge router combination can only be in one SR-MPLS infra L3Out as there can be only one routing policy for a user VRF/border leaf switch/DC-PE combination.
- If there is a requirement to have SR-MPLS connectivity from multiple pods and remote locations, ensure that you have a different SR-MPLS infra L3Out in each of those pods and remote leaf locations with SR-MPLS connectivity.
- If you have a Multi-Pod or Remote Leaf topology where one of the pods is not connected directly to the SR-MPLS network, that pod's traffic destined for the SR-MPLS network will use standard IPN path to another pod, which has an SR-MPLS L3Out. Then the traffic will use the other pod's SR-MPLS L3Out to reach its destination across SR-MPLS network.

This also can apply to Multi-Site deployments where North-South communication for the endpoints in [Site 1](#) can be established via an SR-MPLS L3Out connection in [Site 2](#).

- Routes from multiple VRFs can be advertised from one SR-MPLS Infra L3Out to provider edge (PE) routers connected to the nodes in this SR-MPLS Infra L3Out.

PE routers can be connected to the border leaf directly or through other provider (P) routers.

- The underlay configuration can be different or can be the same across multiple SR-MPLS Infra L3Outs for one location.

For example, assume the same border leaf switch connects to PE-1 in domain 1 and PE-2 in domain 2, with the underlay connected to another provider router for both. In this case, two SR-MPLS Infra L3Outs will be created: one for PE-1 and one for PE-2. But for the underlay, it's the same BGP peer to the provider router. Import/export route-maps will be set for EVPN session to PE-1 and PE-2 based on the corresponding route profile configuration in the user VRF.

MPLS Custom QoS Policies

Following is the default MPLS QoS behavior:

- All incoming MPLS traffic on the border leaf switch is classified into QoS Level 3 (the default QoS level).
- The border leaf switch will retain the original DSCP values for traffic coming from SR-MPLS without any remarking.
- The border leaf switch will forward packets with the default MPLS EXP (0) to the SR-MPLS network.

Following are the guidelines and limitations for configuring MPLS Custom QoS policies:

- Data Plane Policers (DPP) are not supported at the SR-MPLS L3Out.
- Layer 2 DPP works in the ingress direction on the MPLS interface.
- Layer 2 DPP works in the egress direction on the MPLS interface in the absence of an egress custom MPLS QoS policy.
- VRF level policing is not supported.

SR-MPLS Tenant Requirements and Guidelines

While the Infra MPLS configuration and requirements are described in the Day-0 operations chapter, the following restrictions apply for any user Tenants you will deploy to sites that are connected to SR-MPLS networks.

- In case when traffic between two EPGs in the fabric needs to go through the SR-MPLS network:
 - Contracts must be assigned between each EPG and the SR-External EPG defined on the local Tenant SR-MPLS L3Out.
 - If both EPGs are part of the same ACI fabric but separated by an SR-MPLS network (for example, in Multi-Pod or remote leaf cases), the EPGs must belong to different VRFs and not have a contract between them nor route-leaking configured.
 - If EPGs are in different sites, they can be in the same VRF, but there must **not** be a contract configured directly between the EPGs and any other remote EPG that is part of the same VRF.
- When configuring a route map policy for the SR-MPLS L3Out:
 - Each L3Out must have a single export route map. Optionally, it can also have a single import route map.

- Routing maps associated with any SR-MPLS L3Out must explicitly define all the routes, including bridge domain subnets, which must be advertised out of the SR-MPLS L3Out.
- If you configure a $0.0.0.0/0$ prefix and choose to not aggregate the routes, it will allow the default route only.
However, if you choose to aggregate routes for the $0.0.0.0/0$ prefix, it will allow all routes.
- You can associate any routing policy with any tenant L3Out.
- Beginning with Nexus Dashboard release 4.0(1), transit routing between SR-MPLS networks is supported using the same or different VRFs for fabrics running Cisco APIC release 5.1(1) or later.

Figure 43: Transit Routing Configuration Using Single VRF

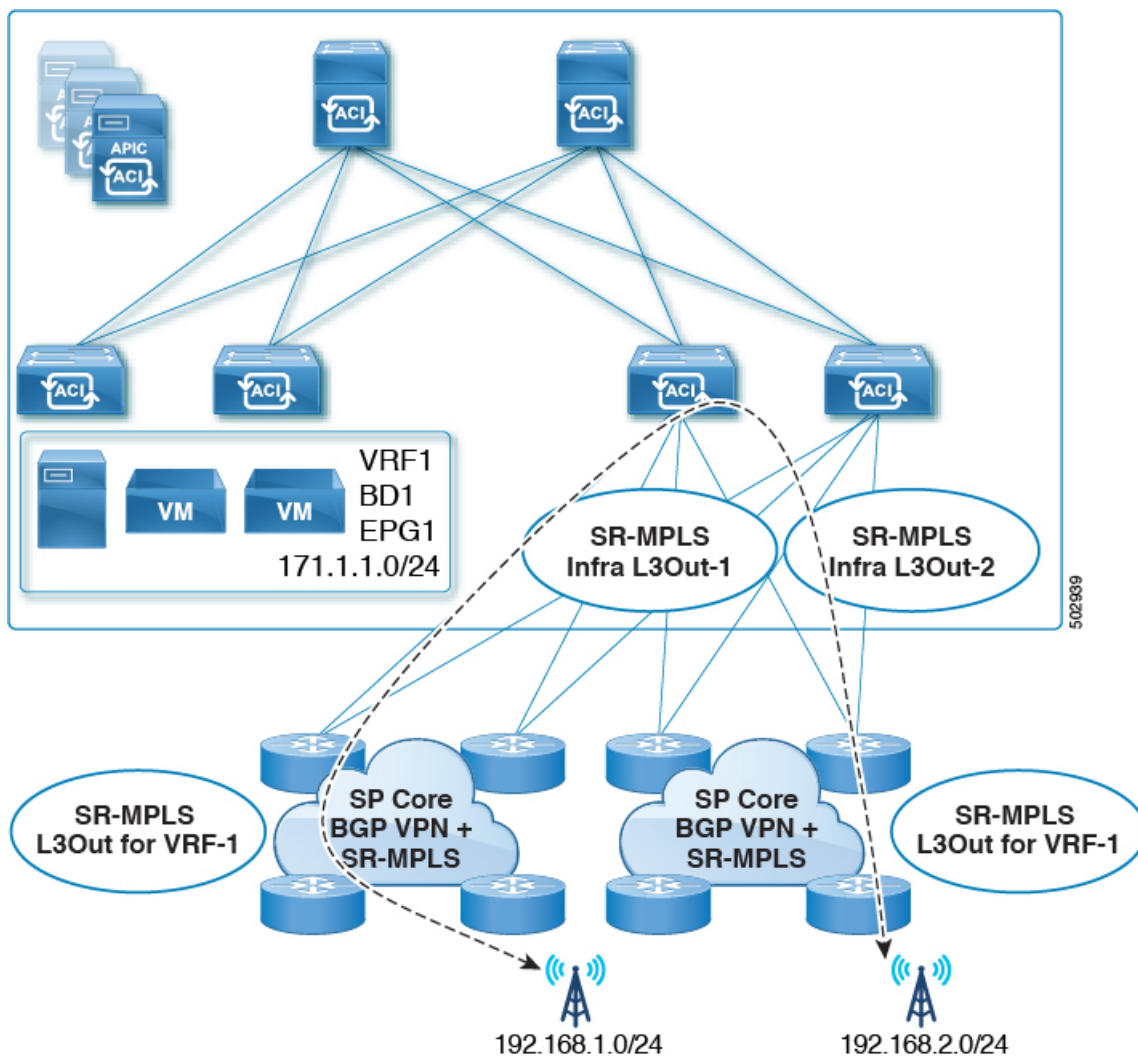
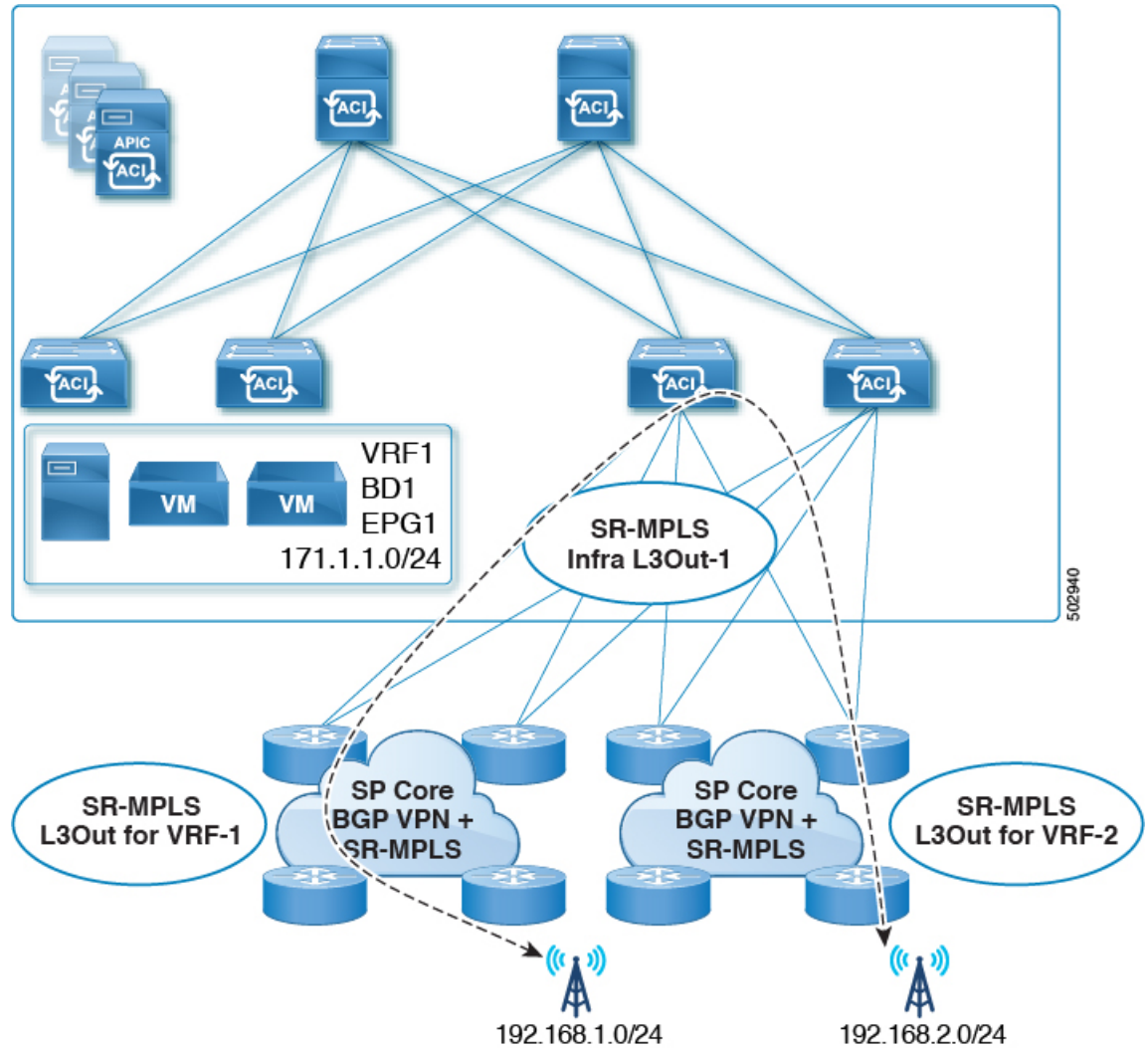


Figure 44: Transit Routing Configuration Using Different VRFs



Prior releases supported transit routing using different VRFs only.

Greenfield Deployment

Creating Custom QoS Policy for SR-MPLS

SR-MPLS Custom QoS policy defines the priority of the packets coming from an SR-MPLS network while they are inside the ACI fabric based on the incoming MPLS EXP values defined in the MPLS QoS ingress policy. It also marks the CoS and MPLS EXP values of the packets leaving the ACI fabric through an MPLS interface based on IPv4 DSCP values defined in MPLS QoS egress policy.

**Note**

Creating custom QoS policy is optional. If no custom ingress policy is defined, the default QoS Level (Level13) is assigned to packets inside the fabric. If no custom egress policy is defined, the default EXP value of 0 will be marked on packets leaving the fabric.

Procedure

Step 1 Log in to your Cisco Nexus Dashboard and open the Cisco Nexus Dashboard Orchestrator service.

Step 2 Create a new Fabric Policy.

- a) From the left navigation pane, choose **Configure > Fabric Template > Fabric Policies**.
- b) On the **Fabric Policy Templates** page, click **Create Fabric Policy Template**.
- c) From the **+Create Object** drop-down, select **QoS SR-MPLS**.
- d) In the right properties sidebar, provide the **Name** for the policy.
- e) (Optional) Click **Add Description** and provide a description for the policy.

Step 3 Click **Add Ingress Rule** to add an ingress QoS translation rule.

These rules are applied for traffic that is ingressing the ACI fabric from an MPLS network and are used to map incoming packet's experimental bits (EXP) values to ACI QoS levels, and set DSCP or CoS values that the packet should have set when forwarded to an endpoint connected to the fabric.

The values are derived at the border leaf switch using a custom QoS translation policy. If a custom policy is not defined or not matched, default QoS Level (Level13) is assigned.

- a) In the **Match EXP From** and **Match EXP To** fields, specify the EXP range of the ingressing MPLS packet you want to match.
- b) From the **Queuing Priority** drop-down, select the ACI QoS Level to map.

This is the QoS Level that you want to assign for the traffic within ACI fabric, which ACI uses to prioritize the traffic within the fabric. The options range from Level 1 to Level 6. The default value is Level13. If you do not make a selection in this field, the traffic will automatically be assigned a Level13 priority.

- c) From the **Set DSCP** drop-down, select the DSCP value to be used when sending the unencapsulated packet to an endpoint connected to the fabric.

The DSCP value that is specified is set in the original traffic that is received from the external network, so it will be reexposed only when the traffic is VXLAN decapsulated on the destination ACI leaf node.

If you set the value to *Unspecified*, the original DSCP value of the packet will be retained.

- d) From the **Set CoS** drop-down, select the CoS value to be used when sending the unencapsulated packet to an endpoint connected to the fabric.

The CoS value that is specified will be reexposed only when the traffic is VXLAN decapsulated on the destination ACI leaf node.

If you set the value to *Unspecified*, the original CoS value of the packet will be retained.

In both of the above cases, the CoS preservation option must be enabled in the fabric. For more information about CoS preservation, see [Cisco APIC and QoS](#).

- e) Click the check mark icon to save the rule.

- f) Repeat this step for any additional ingress QoS policy rules.

Step 4 Click **Add Egress Rule** to add an egress QoS translation rule.

These rules are applied for the traffic that is leaving the ACI fabric through an MPLS L3Out and are used to map the packet's IPv4 DSCP value to the MPLS packet's EXP value and the internal Ethernet frame's CoS value.

The setting of the packet's IPv4 DSCP value is done at the nonborder leaf switch based on existing policies that are used for EPG and L3Out traffic. If a custom policy is not defined or not matched, the default EXP value of 0 is marked on all labels. EXP values are marked in both, default and custom policy scenarios, and are done on all MPLS labels in the packet.

Custom MPLS egress policy can override existing EPG, L3Out, and Contract QoS policies.

- Using the **Match DSCP From** and **Match DSCP To** dropdowns, specify the DSCP range of the ACI fabric packet you want to match for assigning the egressing MPLS packet's priority.
- From the **Set MPLS EXP** drop-down, select the EXP value that you want to assign to the egressing MPLS packet.
- From the **Set CoS** drop-down, select the CoS value that you want to assign to the egressing MPLS packet.
- Click the check mark icon to save the rule.
- Repeat this step for any additional egress QoS policy rules.

Step 5 From the **Actions** menu, select **Add/Remove Sites** and choose the SR-MPLS site with which to associate this template.

Step 6 Click **Save** to save the template policy.

Step 7 Click **Deploy** to deploy the fabric policy to the sites.

What to do next

After you have created the QoS policy, enable MPLS connectivity and configure MPLS L3Out as described in [Creating SR-MPLS Infra L3Out, on page 346](#).

Creating SR-MPLS Infra L3Out

This section describes how to configure SR-MPLS Infra L3Out settings for a site that is connected to an SR-MPLS network.

- The SR-MPLS infra L3Out is configured on the border leaf switch, which is used to set up the underlay BGP-LU and overlay MP-BGP EVPN sessions that are needed for the SR-MPLS hand-off.
- An SR-MPLS infra L3Out will be scoped to a pod or a remote leaf switch site.
- Border leaf switches or remote leaf switches in one SR-MPLS infra L3Out can connect to one or more provider edge (PE) routers in one or more routing domains.
- A pod or remote leaf switch site can have one or more SR-MPLS infra L3Outs.

Before you begin

You must have:

- Added a site that is connected through SR-MPLS network as described in [Adding Cisco ACI Sites, on page 165](#).
- If necessary, created SR-MPLS QoS policy as described in [Creating Custom QoS Policy for SR-MPLS, on page 344](#).

Procedure

- Step 1** Ensure that SR-MPLS Connectivity is enabled for the site.
- In the main navigation menu, select **Configure > Site To Site Connectivity**.
 - In the **Site To Site Connectivity** page, click **Configure**.
 - In the left pane, under **Sites**, select the specific site that is connected through SR-MPLS.
 - In the right **<Site> Settings** pane, enable the **SR-MPLS Connectivity** and provide the SR-MPLS information.
 - The **Segment Routing Global Block (SRGB) Range** is the range of label values that are reserved for Segment Routing (SR) in the Label Switching Database (LSD). The Segment ID (SID) is a unique identifier for a specific segment and is configured on each node for the MPLS transport loopback. The SID index, which you configure later as part of the border leaf switch configuration, is advertised using BGP-LU to the peer router, and the peer router uses the SID index to calculate the local label.

The default range is 16000-23999.
 - The **Domain ID Base** enables the BGP Domain-Path feature. For more information, see [Cisco APIC Layer 3 Networking Configuration Guide](#).
- If you choose to provide a value in this field to enable the Domain-Path feature, ensure that you use a unique value for each SR-MPLS site in your Multi-Site domain, which will be specific to this ACI fabric.
- Step 2** In the main pane, click **+Add SR-MPLS L3Out** within a pod.
- Step 3** In the right **Properties** pane, provide a name for the SR-MPLS L3Out.
- Step 4** (Optional) From the **QoS Policy** drop-down, select a QoS Policy that you created for SR-MPLS traffic. Select the QoS policy that you created in [Creating Custom QoS Policy for SR-MPLS, on page 344](#). Otherwise, if you do not assign a custom QoS policy, the following default values are assigned:
- All incoming MPLS traffic on the border leaf switch is classified into QoS Level 3 (the default QoS level).
 - The border leaf switch does the following:
 - Retains the original DSCP values for traffic coming from SR-MPLS without any remarking.
 - Forwards packets to the MPLS network with the original CoS value of the tenant traffic if the CoS preservation is enabled.
 - Forwards packets with the default MPLS EXP value (0) to the SR-MPLS network.
 - In addition, the border leaf switch does not change the original DSCP values of the tenant traffic coming from the application server while forwarding to the SR network.
- Step 5** From the **L3 Domain** drop-down, select the Layer 3 domain.
- Step 6** Configure settings for border leaf switches and ports that are connected to the SR-MPLS network. You must provide information about the border leaf switches and the interface ports which connect to the SR-MPLS network.
- Click **+Add Leaf** to add a leaf switch.
 - In the **Add Leaf** window, select the leaf switch from the **Leaf Name** drop-down.

- c) In the **SID Index** field, provide a valid segment ID (SID) offset.

When configuring the interface ports later in this section, you are able to choose whether you want to enable segment routing. The SID index is configured on each node for the MPLS transport loopback. The SID index value is advertised using BGP-LU to the peer router, and the peer router uses the SID index to calculate the local label. If you plan to enable segment routing, you must specify the segment ID for this border leaf switch.

If you must update the SID index value, you must first delete it from all SR-MPLS L3Outs in the leaf switch and redeploy the configuration. Then you can update it with the new value, followed by redeploying the new configuration.

- d) Provide the local **Router ID**.

Unique router identifier within the fabric.

- e) Provide the **BGP EVPN Loopback** address.

Note

The BGP EVPN Loopback address must be the same for the selected leaf switch across all SR-MPLS L3Outs in the site.

The BGP-EVPN loopback is used for the BGP-EVPN control plane session. Use this field to configure the MP-BGP EVPN session between the EVPN loopback addresses of the border leaf switch and the DC-PE to advertise the overlay prefixes. The MP-BGP EVPN sessions are established between the BGP-EVPN loopback and the BGP-EVPN remote peer address, which you configure in the "Add Interface" substep below.

While you can use a different IP address for the BGP-EVPN loopback and the MPLS transport loopback, we recommend that you use the same loopback for the BGP-EVPN and the MPLS transport loopback on the ACI border leaf switch.

- f) Provide the **MPLS Transport Loopback** address.

The MPLS transport loopback is used to build the data plane session between the ACI border leaf switch and the DC-PE, where the MPLS transport loopback becomes the next-hop for the prefixes that are advertised from the border leaf switches to the DC-PE routers.

While you can use a different IP address for the BGP-EVPN loopback and the MPLS transport loopback, we recommend that you use the same loopback for the BGP-EVPN and the MPLS transport loopback on the ACI border leaf switch.

- g) Click **Add Interface** to provide switch interface details.

From the **Interface Type** drop-down, select whether it is a Layer 3 physical interface or a port channel interface. If you choose to use a port channel interface, it must have been already created on the APIC.

Then provide the interface, its IP address, and MTU size. If you want to use a subinterface, provide the **VLAN ID** for the subinterface, otherwise leave the VLAN ID field blank.

In the **BGP-Label Unicast Peer IPv4 Address** and **BGP-Label Unicast Remote AS Number**, specify the BGP-LU peer information of the next hop device, which is the device that is connected directly to the interface. The next hop address must be part of the subnet that is configured for the interface.

Choose whether you want to enable an MPLS or an SR-MPLS hand-off.

(Optional) Choose to enable the additional BGP options based on your deployment.

Finally, click the check mark to the right of the **Interface Type** drop-down to save interface port information.

- h) Repeat the previous substep for all interfaces on the switch that connect to the MPLS network.
- i) Click **Save** to save the leaf switch information.

- j) Repeat this step for all leaf switches connected to the MPLS networks.

Step 7

Configure BGP-EVPN connectivity.

You must provide BGP connectivity details for the BGP EVPN connection between the site's border leaf switch (BL) switches and the provider edge (PE) router.

- a) Click **+Add BGP-EVPN Connectivity**.
- b) In the **Add MPLS BGP-EVPN Connectivity** window, provide the details.

For the **MPLS BGP-EVPN Peer IPv4 Address** field, provide the loopback IP address of the DC-PE router, which is not necessarily the device that is connected directly to the border leaf switch.

For the **Remote AS Number**, enter a number that uniquely identifies the neighbor autonomous system of the DC-PE. The Autonomous System Number can be in 4 byte as plain format 1–4294967295. Keep in mind, ACI supports only `asplain` format and not `asdot` or `asdot+` format AS numbers. For more information on ASN formats, see [Explaining 4-Byte Autonomous System \(AS\) ASPLAIN and ASDOT Notation for Cisco IOS](#) document.

For the **TTL** field, specify a number large enough to account for multiple hops between the border leaf switch and the DC-PE router, for example 10. The allowed range is 2–255 hops.

(Optional) Choose to enable the additional BGP options based on your deployment.

- c) Click **Save** to save BGP settings.
- d) Repeat this step to for any additional BGP connections.

Typically, you would be connecting to two DC-PE routers, so provide BGP peer information for both connections.

Step 8

Deploy the changes to sites.

What to do next

After you have enabled and configured MPLS connectivity, you can create and manage Tenants, route maps, and schemas as described in the [Creating SR-MPLS Route Map Policy, on page 349](#).

Creating SR-MPLS Route Map Policy

This section describes how to create a route map policy. Route maps are sets of `if-then` rules that enable you to specify which routes are advertised out of the Tenant SR-MPLS L3Out. Route maps also enable you to specify which routes that are received from the DC-PE routers will be injected into the BGP VPNv4 ACI control plane.

You will use the SR-MPLS route map policy in the next section when defining site-local settings for the Tenant SR-MPLS L3Out.

Procedure**Step 1**

Log in to your Cisco Nexus Dashboard and open the Cisco Nexus Dashboard Orchestrator service.

Step 2

Create a new Tenant Policy.

- a) From the left navigation pane, choose **Configure > Tenanat Template > Tenant Policies**.
- b) On the **Tenant Policy Templates** page, click **Create Tenant Policy Template**.
- c) In the Tenant Policies page's right properties sidebar, provide the **Name** for the tenant.

- d) From the **Select a Tenant** drop-down, choose the tenant with which you want to associate this template.

All the policies that you create in this template as described in the following steps will be associated with the selected tenant and deployed to it when you push the template to one or more sites.

By default, the new template is empty, so you must add one or more tenant policies as described in the following steps. You don't have to create every policy available in the template – you can create a template with just a single route map policy for your SR-MPLS use case.

Step 3 Create a Route Map Policy for Route Control.

- From the **+Create Object** drop-down, select **Route Map Policy for Route Control**.
- In the right properties sidebar, provide the **Name** for the policy.
- (Optional) Click **Add Description** and provide a description for the policy.
- Click **+Add Entry** and provide the route map information.

For each route map, you must create one or more context entries. Each entry is a rule that defines an action based on one or more matching criteria based on the following information:

- **Context Order** – Context order is used to determine the order in which contexts are evaluated. The value must be in the 0–9 range.
- **Context Action** – Context action defines the action to perform (`permit` or `deny`) if a match is found. If the same value is used for multiple contexts, they are evaluated one in the order in which they are defined.

When the context order and action are defined, choose how you want to match the context:

- Click **+Create Attribute** to specify the action that will be taken should the context match.

You can choose one of the following actions:

- Set Community
- Set Route Tag
- Set Dampening
- Set Weight
- Set Next Hop
- Set Preference
- Set Metric
- Set Metric Type
- Set AS Path
- Set Additional Community

After you have configured the attribute, click **Save**.

- If you want to associate the action that you defined with an IP address or prefix, click **Add IP Address**.

In the **Prefix** field, provide the IP address prefix. Both IPv4 and IPv6 prefixes are supported, for example, `2003:1:1a5:1a5::/64` or `205.205.0.0/16`.

If you want to aggregate IPs in a specific range, check the **Aggregate** check box and provide the range. For example, you can specify `0.0.0.0/0` prefix to match any IP or you can specify `10.0.0.0/8` prefix to match any `10.x.x.x` addresses.

- If you want to associate the action that you defined with community lists, click **Add Community**.

In the **Community** field, provide the community string. For example, `regular:as2-nn2:200:300`.

Then choose the **Scope**: *Transitive* means that the community will be propagated across eBGP peering (across autonomous systems) while *Non-Transitive* means the community will not be propagated.

Note

You must specify an **IP address** or a **Community** string to match a specific prefix (even if you do not provide a **Set** attribute) because it defines the prefixes that must be announced out of the L3Out. This can be either BDs' subnets or transit routes learned from other L3Outs.

- e) Repeat the previous substeps to create any additional route map entries for the same policy.
- f) Click **Save** to save the policy and return to the template page.
- g) Repeat this step to create any additional Route Map for Route Control policies.

Step 4 From the **Actions** menu, select **Add/Remove Sites** and choose one or more SR-MPLS sites with which to associate this template.

Step 5 Click **Deploy** to deploy the tenant policy to the sites.

Creating SR-MPLS Tenant L3Outs in L3Out Templates

Beginning with NDO release 4.1(1), L3Out and SR-MPLS L3Out configuration has moved out of application templates and into dedicated L3Out templates. Before you can configure connectivity across an SR-MPLS network, you must create L3Out templates and define the SR-MPLS L3Outs for each site as described in this section.

Procedure

Step 1 Log in to your Nexus Dashboard and open the Nexus Dashboard Orchestrator service.

Step 2 Create a new L3Out template.

- a) From the left navigation pane, choose **Configure > Tenantat Template > > L3Out**.
- b) On the **L3Out Templates** page, click **Create L3Out Template**.
- c) In the **Select a Tenant and Site** dialog, choose the tenant and site with which you want to associate this template, then click **Save and go to template**.

Each L3Out template is associated with a specific tenant similar to other NDO templates, however it is also assigned to a single site only as L3Out configuration is typically site-specific.

If you want to define L3Out configuration for multiple sites, you must create at least one L3Out template for each site, but you can deploy multiple L3Outs per site/tenant by defining all of them in the same L3Out template. You may have multiple L3Out templates per site as long as they are assigned to different tenants.

- d) In the template view, provide the **Name** for the template.

Step 3 Create SR-MPLS L3Out(s).

- a) In the main pane, choose **Create Object > SR-MPLS L3Out**.
- b) Provide the **Name** for the L3Out.

Note

We recommend providing unique names for all SR-MPLS L3Outs across sites, even if they belong to the same tenant or allow connectivity to the same external resources.

- c) Click **Select VRF>** and choose a VRF to associate with this SR-MPLS L3Out.

Note

This step assumes you have a VRF already defined for this SR-MPLS L3Out. If you do not, you can close the template page, define the VRF in an application template as you typically would, and then resume SR-MPLS L3Out creation from this step.

- d) Click **Add SR-MPLS L3Out**.
- e) In the **Add SR-MPLS L3Out** dialog that opens, choose the **SR-MPLS Infra L3Out** that you defined in [Creating SR-MPLS Infra L3Out, on page 346](#).
- f) Click **Add Route Map Policy** and choose the route map policy that you defined in [Creating SR-MPLS Route Map Policy, on page 349](#) and whether it's an **Import** or **Export** policy.

You can repeat this substep if you want to add multiple route map policies to your SR-MPLS L3Out.

- g) Repeat this step for all SR-MPLS L3Outs you want to create for this specific site and tenant.

Step 4 In the template view, click **Deploy** to deploy the template to the site.

Step 5 Repeat this process to create a separate L3Out template for each site with that site's SR-MPLS L3Out(s).

The next section assumes a use case in which two SR-MPLS L3Outs were created in two different sites, for example `mpls-l3out-1` and `mpls-l3out-2`

Configure EPG-to-External-EPG (North-South) Communication

This section describes how to establish North-South communication between an application EPG and an external SR-MPLS network. You can also use this approach to enable EPG-to-EPG communication across sites through the SR-MPLS L3Out data path (leveraging the external SR-MPLS network).

If instead you want to establish EPG-to-EPG intersite connectivity through the VXLAN data plane across the ISN which is supported starting with release 4.0(2), you can simply establish a contract relationship between those EPGs as you typically would.

Procedure

Step 1 Choose the template or create a new one.

You can select the template as you typically would for other ACI fabric use cases:

- a) In the main navigation menu, select **Configure > Tenant Template > Applications > Schemas**.
- b) Select an existing schema or create a new one.
- c) Select an existing template or click **Create New Template** and select `ACI Multi-Cloud` for template type.
- d) Select the tenant for the new template.

- e) (Optional) Enable the **Autonomous** option for the template if you plan to deploy this template only to sites that do not have any intersite connectivity to other sites.

Step 2 Create a VRF.

- a) From the **+Create Object** menu, choose **VRF**.
- b) In the right properties sidebar, provide the name for the VRF.

Step 3 Create an SR-External EPG.

Note

If you assign the template that contains SR-External EPG to multiple sites, the EPG will be stretched to all those sites. In this case, each site must have a local SR-MPLS L3Out or you will not be allowed to deploy that template to all associated sites.

- a) From the **+Create Object** menu, choose **SR-External EPG**.
- b) In the right properties sidebar, provide the name for the external EPG.
- c) From the **Virtual Routing & Forwarding** drop-down, choose the VRF you created in the previous step.
- d) From the **L3Out** drop-down, choose the SR-MPLS L3Out you created in [Creating SR-MPLS Tenant L3Outs in L3Out Templates, on page 351](#).
- e) Click **+Add Subnet** and define a subnet and its route control options as you typically would.

If you want to define multiple subnets, repeat this substep.

Step 4 Assign the template to a single site or to multiple sites, depending on the specific use case you must configure.

Step 5 Select the site-local settings for the template that you are configuring.

In the following few steps, you configure site local settings for the VRF and SR-External EPG you created in the previous steps.

Step 6 Configure site-local settings for the VRF.

You must provide BGP route information for the VRF used by the SR-MPLS L3Out.

- a) In the main pane, scroll down to **VRF** area and select the VRF you created in the previous step.
- b) From the **Address Family** drop-down, select whether it is IPv4 or IPv6 address.
- c) In the **Route Target** field, provide the route string.

Note

Configuration of the import/export route-target values must be consistent with the configuration that is deployed on the DC-PE device and depends on the specific use case being deployed.

For example, `route-target:ipv4-nn2:1.1.1.1:1901`.

- d) From the **Type** drop-down, select whether to import or export the route.
- e) Click **Save** to save the route information.
- f) (Optional) Repeat this step to add any additional BGP route targets.

Step 7 Create and configure an application EPG as you typically would.

Note

The EPG can be in the same or different template and schema.

Step 8 Create a contract between the application EPG and the SR-External EPG.

Step 9 Deploy the configuration.

- a) In the main pane of the **Schemas** view, click **Deploy to Sites**.

- b) In the **Deploy to Sites** window, verify the changes that will be pushed to the site and click **Deploy**.

Note

Starting from release 4.0(2), it is possible to use the EPG-to-SR-External-EPG contracts exclusively for North-South traffic (communication with resources external to the ACI fabrics), similar to the traditional IP-based L3Outs. In that case, EPG-to-EPG intersite communication can be enabled through the VXLAN data path across the ISN by simply creating a contract relationship between those EPGs.

However, if you want to establish EPG-to-EPG (East-West) communication between EPGs in different sites across the external SR-MPLS network, you can do that as outlined in the next step.

Step 10

If you want to use the SR-MPLS L3Out data path for EPG-to-EPG traffic across sites (leveraging the SR-MPLS external network instead of the VXLAN data path across the ISN), you can establish contracts between each site-local EPG and the SR-External EPG associated to the tenant SR-MPLS L3Out.

The SR-External EPG can be deployed as a site-local object in each site or as a stretched object across sites. Note that using the SR-MPLS L3Out data path for EPG-to-EPG traffic across sites is only possible if there are no direct contract relationships between those EPGs or between each EPG and any other remote EPG.

- a) Create two application EPGs as you typically would in templates that are associated to different sites.

For example, `epg1` and `epg2`.

These EPGs can be in the same or different VRFs or Tenants.

- b) Create two separate site-local SR-External EPGs or a single stretched SR-External EPG.

If you are creating separate SR-External EPGs, they can be in the same or different VRFs or Tenants and the same template or different templates depending on the specific deployment scenario.

Note

In contrast with regular External EPGs where you associate an L3Out explicitly, there is only one SR-MPLS L3Out per VRF so when you create the SR-External EPGs, you associate them with the same VRF as you used for your SR-MPLS Tenant L3Outs that you created in [Creating SR-MPLS Tenant L3Outs in L3Out Templates, on page 351](#).

For example, the next step assumes you create `mpls-extepg-1` and `mpls-extepg-2`.

- c) Create a contract that you use to allow traffic between each site local EPG and SR-MPLS L3Out local connection.

You must create and define a filter for the contract as you typically would.

- d) Assign the contracts to the appropriate EPGs.

To allow traffic between the two application EPGs you created, you will actually need to assign the contract twice: when between `epg1` and its `mpls-extepg-1` and then again between `epg2` and its `mpls-extepg-2`. It's possible to have the same SR-External EPG instead of two separate ones if it is stretched across sites.

As an example, if you want `epg1` to provide a service to `epg2`, you would:

- Assign the contract to `epg1` with type `provider`.
- Assign the contract to `mpls-extepg-1` with type `consumer`.
- Assign the contract to `epg2` with type `consumer`.
- Assign the contract to `mpls-extepg-2` with type `provider`.

Importing Existing SR-MPLSL3Out Configuration

Overview of Importing SR-MPLS Configuration

Beginning with release 4.1(2), Nexus Dashboard Orchestrator (NDO) supports importing existing SR-MPLS configurations from the APIC sites. The following sections focus on the steps required



Note If you want to configure and deploy new SR-MPLS configurations (greenfield deployment), see the earlier sections of this chapter.

This release supports importing the following policies.

- **Route Maps** – may be referenced in the L3Out template's **Outbound Route Map** and **Inbound Route Map** fields to define route import and export policies.
- **L3Out Node Routing:**
 - Nodes configured for an L3Out can be associated to a node group, which in turn can refer to a node routing policy.
 - Node groups can also reference BGP Peer Prefix policy when configuring BGP peers for the nodes.
- **L3Out Interface Routing:**
 - Interfaces configured for an L3out can be associated to an interface group, which can refer to an interface routing policy and BGP Peer Prefix policy
 - Interface groups can also reference BGP Peer Prefix policy when configuring BGP peers for the interfaces.
- **BGP Peer Prefix** – can be referenced by the node and interfaces groups for BGP peer configuration on all nodes in the group.
- **IPSLA Monitoring policies and IPSLA Track lists** – can be referenced by the static routes defined for a node

Mapping of Sites' MOs to NDO Objects and Groups

Note that in some cases there is no 1:1 mapping between the managed objects (MOs) created in the site and the policy objects as they are seen on and managed by the Orchestrator. In these cases, when you import an L3Out from APIC, NDO creates imports the MOs using NDO-specific logical groups; for example, the following APIC policies are grouped on import:

- The following MOs are grouped into an L3Out Node Routing policy on NDO:
 - BGP Timer Policy
 - BGP Best Path Policy
 - BFD Multi-Hop Node Policy

The following figure shows the **L3Out Node Routing Policy** object in NDO that groups together the 3 policies mentioned above:

Orchestrator ▼ ?

[Configure](#) / [Tenant Templates](#) [[Tenant Policies](#)] / [vz](#)

Tenant Policies

[Template Properties](#) ● Site2

Template Summary

Type	Tenant common
Tenant Policy Template	

Filter

L3Out Node Routing Policy ▼

L3OutNodePolicy

IPSLA Track List ▼

IPSLATrack

IPSLA Monitoring Policy ▼

vzany:

L3OutNodePolicy

Name * [Add Description](#)

BFD MultiHop Settings [Add](#)

BGP Node Settings [Delete](#)

Graceful Restart Helper ☒ Enabled

Keep Alive Interval (sec)

Hold Interval (sec)

Stale Interval (sec)

Max As Limit

BGP Best Path Control [Delete](#)

AS Path Multipath Relax ☒ Enabled

[Ok](#)

- The following MOs are grouped into an L3Out Interface Routing policy on NDO:
 - OSPF Interface Policy
 - BFD Policy
 - BFD Multi-Hop Interface Policy

Automatic Import of Dependencies

Tenant Policies templates include objects and policies that have local references within the template. For example, an IPSLA track list can contain a list of track members and each track member must refer to a IPSLA monitoring policy. In such cases, importing existing configuration that contains one or more IPSLA track list policies from a site will also automatically import the referenced IPSLA monitoring policy. The import

workflow displays additional information about the automatically imported policies when you select an object that has such dependencies:

The screenshot shows the 'Tenant Policies' page in the NDO interface. The 'Template Summary' section includes the following details:

Type	Tenant	Template Status	Associated Sites	Last Action
Tenant Policy Template	common	In Sync	2 In Sync, 0 Out of Sync	Deployment Successful Last Deployed: Sep 25, 2023 01:16 am

Below the summary, the 'IPSLA Track List' section shows a dropdown menu with 'IPSLATrackList' selected and a 'Create IPSLA Track List' button.

References to Policies in Tenant "Common"

Some policies that you import from a site may contain references to policies in tenant `common`. Importing such policies will automatically create a copy of the tenant `common` policy in the Tenant Policies template where the objects are being imported and as a result of that, in the tenant associated with that Tenant Policies template, for example:

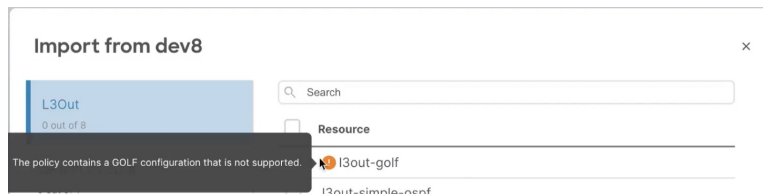
- If you import an IPSLA track list that contains a track member which refers to an IPSLA monitoring policy from the `common` tenant, a copy of the tenant `common`'s IPSLA monitoring policy will be created in the Tenant Policies template and the imported track member will reference this newly added IPSLA monitoring policy.
- If you import an L3Out that contains node configuration with a static route which references an IPSLA track list from tenant `common`, a copy of the tenant `common`'s IPSLA track list will be created in the Tenant Policies template.

Unsupported Scenarios

If an L3Out contains one or more configuration options that are currently not supported by NDO, you will not be able to import that L3Out. The following configurations are currently not supported by NDO and will prevent you from importing any L3Out that includes them:

- GOLF
- EIGRP

In these cases, the import workflow UI will display an orange exclamation point icon with a message explaining the issue and you will not be able to select that L3Out for import:



Importing Tenant Policies Template Objects

This section describes how to import existing SR-MPLS L3Out configuration policies from Cisco APIC into NDO's Tenant Policies template. For more information about each policy and how it relates to policies and settings in other templates, see [Overview of Importing SR-MPLS Configuration, on page 355](#).

Before you begin

- If you want to configure and deploy new SR-MPLS L3Out configurations (greenfield deployment), see the earlier sections of this chapter instead.
- You must have the Cisco Nexus Dashboard Orchestrator service that is installed and enabled.
- You must have the fabrics on board to your Cisco Nexus Dashboard and enabled for management in the Orchestrator service.
- Ensure you have read and understood the Templates and Policy Objects dependencies that are described in [Overview of Importing SR-MPLS Configuration, on page 355](#).

SUMMARY STEPS

1. Log in to your Cisco Nexus Dashboard and open the Orchestrator service.
2. In the left navigation pane, choose **Configure > Tenanat Template > Tenant Policies**.
3. In the main pane, click **Add Tenant Policy Template**.
4. If you created a brand new template, provide the **Name** for the template and **Select a Tenant** from which you plan to import configuration.
5. Associate the template with the site from which you plan to import configuration.
6. Click **Save** to save the template changes.
7. Import one or more policies into the Tenant Policies template.
8. Deploy the template to sites.

DETAILED STEPS

Procedure

-
- Step 1** Log in to your Cisco Nexus Dashboard and open the Orchestrator service.
- Step 2** In the left navigation pane, choose **Configure > Tenanat Template > Tenant Policies**.
- Step 3** In the main pane, click **Add Tenant Policy Template**.

If you want to update an existing Tenant Policy template instead, simply click its name. This opens the **Tenant Policies** page.

Step 4 If you created a brand new template, provide the **Name** for the template and **Select a Tenant** from which you plan to import configuration.

Step 5 Associate the template with the site from which you plan to import configuration.

- a) In the **Tenant Policies** template view, choose **Actions** > **Add/Remove Sites**.
- b) In the **Add Sites to <template-name>** dialog, select the sites to which you want to deploy the template.

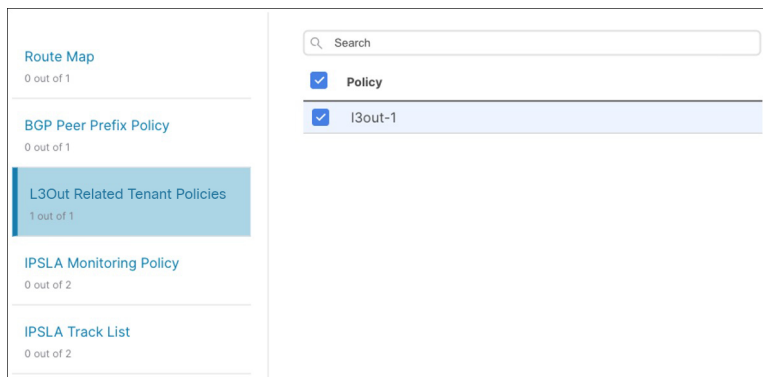
Step 6 Click **Save** to save the template changes.

Step 7 Import one or more policies into the Tenant Policies template.

When you choose to import L3Out configuration from a site, the UI shows the list of L3Out policies that can be imported. You may select one or more L3 out policies and import all the provider policies that are used by the L3 out into this Tenant Policy Template.

- a) In the **Tenant Policies** screen's **Template Properties** view, choose **Import** > **<site-name>**.
- b) In the **Import from <site-name>** dialog, choose one or more L3Outs and click **Import**.

If there's an SR-MPLS L3Out already configured in the site, its associated policies are available for import in the **L3OutSources** category. When you select an L3OutSource to import, all policies that are referenced by that L3Out in the site's APIC are imported into the Tenant Policies template you are editing.



- c) Verify that all imported policies are shown in the template and click **Save** to save it.

All policies configured for the L3Out in the site, which you chose to import in the previous step, are added to the Tenant Policies template using the following guidelines:

- Default import route maps are named as `<l3out-name>_di`.
- Default export route maps are named as `<l3out-name>_de`.
- Node routing policies are named as `<l3out-name>_<node-profile-name>`.
- Interface routing policies are named as `<l3out-name>_<interface-profile-name>`.

Orchestrator

Configure / Tenant Templates [Tenant Policies] / any-pbr

Tenant Policies Refresh Audit Logs Save

Template Properties S1 S2

Template Summary Edit Template Deploy Template Actions

Type	Tenant	Template Status	Associated Sites	Last Action
Tenant Policy Template	common	In Sync	2 In Sync, 0 Out of Sync	Deployment Successful Last Deployed: Sep 25, 2023 01:16 am

Filter IMPORT SELECT Create Object

Route Map Policy for Route Control Create Route Map Policy for Route Control Collapse

I3out-routemaps_exp_7 I3out-routemaps_imp_7 I3out-routemaps3_exp_7 I3out-routemaps3imp_7

Custom QoS Policy Create Custom QoS Policy

DemoCustomQoS

L3Out Node Routing Policy Create L3Out Node Routing Policy

L3OutInterfacePolicy1 L3OutInterfacePolicy2

L3Out Interface Routing Policy Create L3Out Interface Routing Policy

I3out-1_routed-eth1-8-v4

- d) If necessary, update the policy names and click **Save** to save the changes.

We recommend keeping the names of the imported policies as they are created. In this case, when you import L3Outs into the L3Out template as described in the next section, the referenced policies will be automatically recognized and configured for the L3Out by NDO.

However, if you have a specific naming convention in your Multi-Site domain, you can update the imported objects' names to follow that convention. In this case, you must manually provide object references during L3Out import in the next section.

Note

For some objects, there is no 1:1 mapping between the managed objects (MOs) created in the site and the policy objects as they are seen on and managed by the Orchestrator. For information about which MOs are combined into logical groups in NDO, see [Importing Tenant Policies Template Objects, on page 358](#).

Step 8 Deploy the template to sites.

After you have imported the policies and saved the template, you must deploy it back to the site.

- In the **Tenant Policies** template view, click **Deploy**.
- In the **Deploy to sites** dialog, confirm the policies being deployed and click **Deploy**.

What to do next

After you've defined the policies in the Tenant Policy template, proceed to [Importing SR-MPLS Objects, on page 361](#).

Importing SR-MPLS Objects

This section describes how to import an L3Out template from an APIC site into Cisco Nexus Dashboard Orchestrator. For more information about each policy and how it related to policies and settings in other templates, see [Overview of Importing SR-MPLS Configuration, on page 355](#).

Before you begin

- If you want to configure and deploy new L3Out configurations (greenfield deployment), see the earlier sections of this chapter instead.
- You must have created a Template Policy template and imported the policies that are associated with the L3Out you want to import, as described in [Importing Tenant Policies Template Objects, on page 358](#).

SUMMARY STEPS

1. In the left navigation pane, choose **Configure > Tenant Template > L3Out**.
2. In the main pane, click **Create L3Out Template**.
3. If you are creating a brand new template, choose the **Tenant** and **Site** from which you import the L3Out configuration, then click **Save and go to template**.
4. If you created a brand new template, provide the **Name** for the template and click **Save**.
5. Import an SR-MPLS L3Out from the site.
6. Click **Save** to save the template changes.
7. Deploy the template to site.

DETAILED STEPS

Procedure

-
- | | |
|---------------|---|
| Step 1 | In the left navigation pane, choose Configure > Tenant Template > L3Out . |
| Step 2 | <p>In the main pane, click Create L3Out Template.</p> <p>If you want to update an existing L3Out template instead, simply click its name. This opens the L3Out Template page.</p> |
| Step 3 | <p>If you are creating a brand new template, choose the Tenant and Site from which you import the L3Out configuration, then click Save and go to template.</p> <p>Each L3Out template is associated with a specific tenant similar to other NDO templates, however it is also assigned to a single site only as L3Out configuration is typically site-specific.</p> <p>If you want to import SR-MPLS L3Out configuration for multiple sites, you must create at least one L3Out template for each site, but you can import multiple SR-MPLS L3Outs per site/tenant into the same template or you may choose to have multiple SR-MPLS L3Out templates per site as long as they are assigned to different tenants.</p> |
| Step 4 | If you created a brand new template, provide the Name for the template and click Save . |

You must save new templates before you can add new or import existing configuration.

Step 5 Import an SR-MPLS L3Out from the site.

- a) In the main pane, click **Import**.
- b) In the **Import from <site-name>** dialog, select the **SR-MPLS L3Out** you want to import and click **Import**.

Note

Some SR-MPLS L3Outs may be listed with a warning icon. Typically, this means that the associated tenant policies references are not found in NDO Tenant Policies templates and you must first import those references as described in [Importing Tenant Policies Template Objects, on page 358](#).

If you choose to import an L3Out before importing the policies it references and then redeploy that SR-MPLS L3Out to the site, the existing configuration will be removed and the SR-MPLS L3Out will be redeployed from NDO resulting in a loss of any policies that are referenced by the SR-MPLS L3Out that were not imported into NDO.

Step 6 Click **Save** to save the template changes.

Step 7 Deploy the template to site.

After you have imported the L3Out and saved the template, you must deploy it back to the site.

- a) In the **L3Out Template** page, click **Deploy**.
 - b) In the **Deploy to sites** dialog, confirm the policies being deployed and click **Deploy**.
-



CHAPTER 30

vzAny Contracts

- [vzAny and Multi-Site, on page 363](#)
- [vzAny and Multi-Site Guidelines and Limitations, on page 364](#)
- [Create Contract and Filters, on page 366](#)
- [Configure vzAny to Consume/Provide a Contract, on page 367](#)
- [Create EPGs to Be Part of the vzAny VRF, on page 367](#)
- [Free Intra-VRF Communication, on page 368](#)
- [Many-to-One Communication, on page 373](#)

vzAny and Multi-Site

The `vzAny` managed object provides a convenient way of associating all endpoint groups (EPGs) in a Virtual Routing and Forwarding (VRF) instance to one or more contracts, instead of creating a separate contract relation for each EPG.

In the Cisco ACI fabric, EPGs can only communicate with other EPGs according to contract rules. A relationship between an EPG and a contract specifies whether the EPG provides the communications defined by the contract rules, consumes them, or both. By dynamically applying contract rules to all EPGs in a VRF, `vzAny` automates the process of configuring EPG contract relationships. Whenever a new EPG is added to a VRF, `vzAny` contract rules automatically apply. The `vzAny` one-to-all EPG relationship is the most efficient way of applying contract rules to all EPGs in a VRF.



Note External EPGs that are associated with L3Outs and are part of a VRF are also included in the `vzAny` logical group.

Advantages

Policy information in Cisco ACI is programmed in the fabric switches' TCAM tables. TCAM entries are typically specific to each pair of EPGs that are allowed to communicate with each other via a Contract. This means that even if the same contract is re-used, multiple TCAM entries are created for every pair of EPGs.

The size of the policy TCAM table depends on the generation of the switches that you are using. In certain large scale environments it is important to take policy TCAM usage into account and ensure that the limits are not exceeded.

vzAny allows you to combine all EPGs within the same VRF into a single "group" and create a contract relationship with that group rather than individual EPGs within it, while consuming only a single TCAM entry. This saves the time you would otherwise spend creating multiple contract relationships for individual EPGs in the VRF as well as the TCAM space.

Use Cases

There are two typical use cases for vzAny:

- Free communication between EPGs within the same VRF, as described in [Free Intra-VRF Communication, on page 368](#).
- Many-to-one communication allowing all EPGs within the same VRF to consume a shared service from a single EPG, as described in more detail in [Many-to-One Communication, on page 373](#).

vzAny and Multi-Site Guidelines and Limitations

The following guidelines and limitations apply when using vzAny:

- If you plan to enable the vzAny object for a given VRF to provide or consume a contract, the following additional restrictions apply:
 - If vzAny for a given VRF is configured as consumer of a contract c1, the vzAny objects for other VRFs must not be configured as providers of c1.
 - If vzAny for a given VRF is configured as provider of a contract c1, the vzAny objects for other VRFs must not be configured as consumers of c1.
 - If an External EPG part of a given VRF is consuming a contract c1, the vzAny objects for other VRFs must not be configured as providers of c1.
 - If an EPG part of a given VRF is consuming a contract c1, the vzAny objects for other VRFs must not be configured as providers of c1.
 - If vzAny for a given VRF is configured as provider of a contract c1, then EPGs, External EPGs or vzAny objects for other VRFs must not be configured as consumers of c1.
- EPGs and External EPGs objects in a given VRF must not be configured as part of the Preferred Group if vzAny for that VRF is already consuming or providing a contract.
- If any EPG or External EPG objects in a given VRF are deployed in a cloud site, it is not possible to configure vzAny for that VRF to consume or provide a contract
- vzAny is supported with inter-VRF intersite L3Out configurations only when the fabrics are part of the Multi-Site domain running Cisco ACI 5.2(4) release or later.
- vzAny must not consume or provide a contract that is associated with a Service-Graph with PBR.
- vzAny can be configured as provider, consumer or both of a contract for establishing intra-VRF communication.
- vzAny is supported only as a consumer of a shared service but not as a provider.
- We recommend stretching the vzAny VRF to all sites where you plan to deploy EPGs and BDs that use it.

- You can import existing vzAny configurations from an APIC.

**Note**

In certain cases due to an existing issue ([CSCvt47568](#)), if you make changes to the imported configuration before re-deploying it from the Nexus Dashboard Orchestrator, some changes may not get correctly updated in the APICs. To avoid this, re-deploy the configuration immediately after importing but before making any changes to it. After you re-deploy the unchanged config, you will be able to update it as normal.

- vzAny providers and consumers include application EPGs, external EPGs associated to L3Outs, and endpoint groups for in-band or out-of-band access.
- vzAny implicitly creates a 0.0.0.0/0 classification for externally originating traffic, allowing all traffic originating from any external IP subnet. When vzAny is in use for a VRF, it also includes the external EPGs associated to the L3Outs part of that VRF, hence it is equivalent to having created a L3external classification that includes the subnets specified in the VRF itself.
- If an EPG within a VRF is consuming a shared service contract provided by an EPG in a different VRF, the traffic from the EPG of the provider VRF is filtered within the consumer VRF. vzAny is equivalent to a wildcard for the source or destination EPG.

Be careful when you configure a shared service contract between vzAny in the consumer VRF and an EPG1 in a different provider VRF. Since the policy enforcement (filtering) is always performed in the consumer VRF, if the subnet associated to another EPG2 that is part of the provider VRF is leaked into the consumer VRF, then EPG2 will start communicating with consumer EPGs across VRFs even without explicitly providing a contract. Failure to observe this guideline could allow unintended traffic between EPGs across VRFs.

- Configuring a VRF with vzAny as both provider and consumer of a contract using an "allow all" filter, is the same as configuring an unenforced VRF. This implies that all EPGs within that VRF are free to communicate to each other without a contract.
- If the contract scope is application-profile, the vzAny configuration is ignored and filter rules are expanded; CAM utilization is the same as if specific contracts were deployed between each pair of consumer and provider EPGs. In this case, there is no benefit in terms of TCAM space usage.
- In the case of shared services, you must define the provider EPG shared subnet under the EPG in order to properly derive the classification (`pcTag`) of the destination on the consumer (vzAny) side. If you are migrating from a BD-to-BD shared services configuration, where both the consumer and provider subnets are defined under bridge domains, to vzAny acting as a shared service consumer, you must take an extra configuration step where you add the provider subnet to the EPG with the shared flags at minimum. However, since the subnet under the EPG is not needed for connectivity, it is always recommended to check the `No default SVI gateway` flag.

If you add the EPG subnet as a duplicate of the defined BD subnet, ensure that both definitions of the subnet always have the same flags defined. Failure to do so can result in an error.

Create Contract and Filters

When using vzAny, you are essentially creating a single point for a contract relationship, as such you must have a typical contract you will use for any such relationship as well as the filter for the contract.

This section describes how to create a new contract specifically for this purpose. Alternatively, you can choose to import any existing contracts you have configured on each APIC site.

Procedure

Step 1 Log in to the Nexus Dashboard Orchestrator GUI.

Step 2 From the left navigation pane, select **Schemas**.

Step 3 Select the Schema where you want to create your Contract.

If you have an existing Schema you want to update, simply click the Schema's name in the main window pane. Otherwise, if you want to create a new Schema, click the **Add Schema** button and provide the schema information, such as the name and tenant, as you typically would.

Step 4 Create a filter.

- Scrolls down to the **Filter** area and click the + sign to add a new filter.
- Provide the name for the Contract.
- Click **+Entry** to add a filter entry.
- In the **Add Entry** window, provide filter details.

Provide the filter details as you typically would to define the kind of traffic you want to allow.

- Click **SAVE** to add the entry.
- (Optional) If required, create additional filter entries.

Step 5 Create the contract.

- Scrolls down to the **Contract** area and click the + sign to add a new contract.
- Provide the name for the Contract.

For example, `contract-vzany`.

- Choose the scope for the contract

Choose the scope appropriate for your use-case. For example, if you want to enable cross-tenant shared services, you must set the scope to `Global`.

- Choose whether the contract will apply in both directions
- Click **+Filter** to add one or more contract filters.
- In the **Add Filter Chain** window, choose the filter you created in the previous step.
- Click **SAVE** to add the filter.
- (Optional) If required, repeat the procedure to provide additional filters.
- (Optional) If you disabled the **Apply Both Directions** option, provide filters for both, consumer and provider directions.

You have now created the contract you will use with vzAny in the next section.

Configure vzAny to Consume/Provide a Contract

This section describes how to create a vzAny VRF or enable an existing VRF for vzAny.

Before you begin

You must have:

- Created a Contract and one or more Filters to use with vzAny as described in [Create Contract and Filters, on page 366](#).

Procedure

-
- | | |
|---------------|---|
| Step 1 | Log in to the Nexus Dashboard Orchestrator GUI. |
| Step 2 | From the left navigation pane, select Schemas . |
| Step 3 | Select the Schema containing the specific template with the definition of the VRF.

For new configuration, you can create a new schema with the Add Schema button and then define a new template (associated to the tenant of interest) where you can configure VRF. |
| Step 4 | Create or select a VRF.

If you have an existing VRF for which you want to configure vzAny to provide/consume a contract, simply click the VRF in the main window pane. Otherwise, if you want to create a new VRF, scroll down to the VRF area and click the + sign. |
| Step 5 | Select vzAny.

In the right sidebar, check the vzAny checkbox. |
| Step 6 | Select the vzAny contract.

The +Contract option becomes available after you enable the vzAny checkbox.

a) Click +Contract to add the contract
b) Select the contract.

Select the contract you created in Create Contract and Filters, on page 366 .

c) Select the Contract type.

You can choose either <code>consumer</code> or <code>provider</code> for the contract based on your use case. |
-

Create EPGs to Be Part of the vzAny VRF

You can choose to create new or use existing EPGs for your vzAny use cases. There are no explicit vzAny settings on the EPGs and as soon as an EPG is associated to a BD in a VRF, the EPG becomes part of the

vzAny logical group for that VRF (the *vzAny* VRF). If you simply enabled vzAny for an existing VRF with all its EPGs already created and configured, you can skip this section.

Before you begin

You must have:

- Created a Contract and one or more Filters to use with vzAny as described in [Create Contract and Filters, on page 366](#).
- Created the vzAny VRF and assigned the Contract to it as described in [Configure vzAny to Consume/Provide a Contract, on page 367](#).

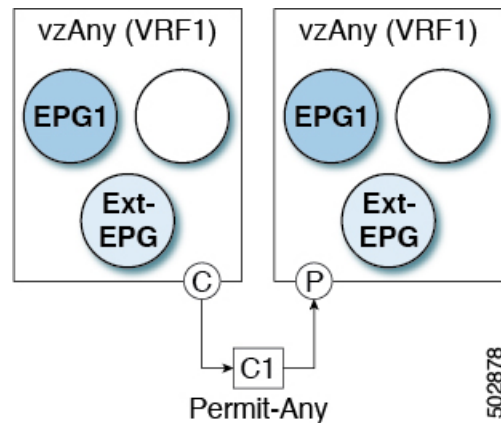
Procedure

-
- Step 1** If you want to create an EPG to be part of the vzAny VRF
- a) Create a BD you will use for your EPG.
 - b) In the BD configuration sidebar's **Virtual Routing & Forwarding** dropdown, select the vzAny VRF you created.
 - c) Create an EPG.
 - d) In the EPG configuration sidebar's **Bridge Domain** dropdown, select the BD you created.
- Step 2** If you want to create an External EPG to be part of the vzAny VRF
- a) Create an external EPG.
 - b) In the External EPG configuration sidebar's **Virtual Routing & Forwarding** dropdown, select the vzAny VRF you created.
-

Free Intra-VRF Communication

This section shows a number of schema examples for unrestricted intra-VRF communication. In all shown scenarios vzAny provides and consumes a contract with a `permit-any` filter. This essentially uses the ACI fabrics for network connectivity only without any policy enforcement and is equivalent to the *VRF Unenforced* option.

Figure 45:



For all the following use cases, you will need to create the same objects and policies summarized below. However, the schema and template design will depend on the number of sites as well as which objects are going to be stretched. The specific sections below contain recommendation on template layout.

Procedure

-
- Step 1** Create a Schema.
- Step 2** Create a common Template used to deploy configuration objects in all the sites (that is *stretched objects*).
- Step 3** Create any additional templates for every combination of sites where EPGs will be deployed .
- If you will deploy a single template to all sites, you can skip this step. The use-case diagrams in the following sections provide template examples.
- Step 4** Within the common Template, create the contract and filters to be consumed/provided by vzAny.
- In this specific use case, the contract should have a single "permit-any" filter rule.
- For specific steps, see [Create Contract and Filters, on page 366](#).
- Step 5** Within the common Template, create a VRF and configure vzAny to consume and provide the previously defined contract with the "permit-any" rule.
- This ensures that free intra-VRF communication can be established.
- For specific steps, see [Configure vzAny to Consume/Provide a Contract, on page 367](#).
- Step 6** Within each site's template, create and configure the EPGs that will be deployed to that site only.
- If you will deploy a single template to all sites, create the EPGs within the same template as the VRF instead. The use-case diagrams in the following sections provide template examples.
- This is described in [Create EPGs to Be Part of the vzAny VRF, on page 367](#).
- Step 7** Assign the common Template to every site.
- Step 8** Assign each template to the appropriate sites.
- Step 9** Deploy the templates.
-

Stretched EPGs

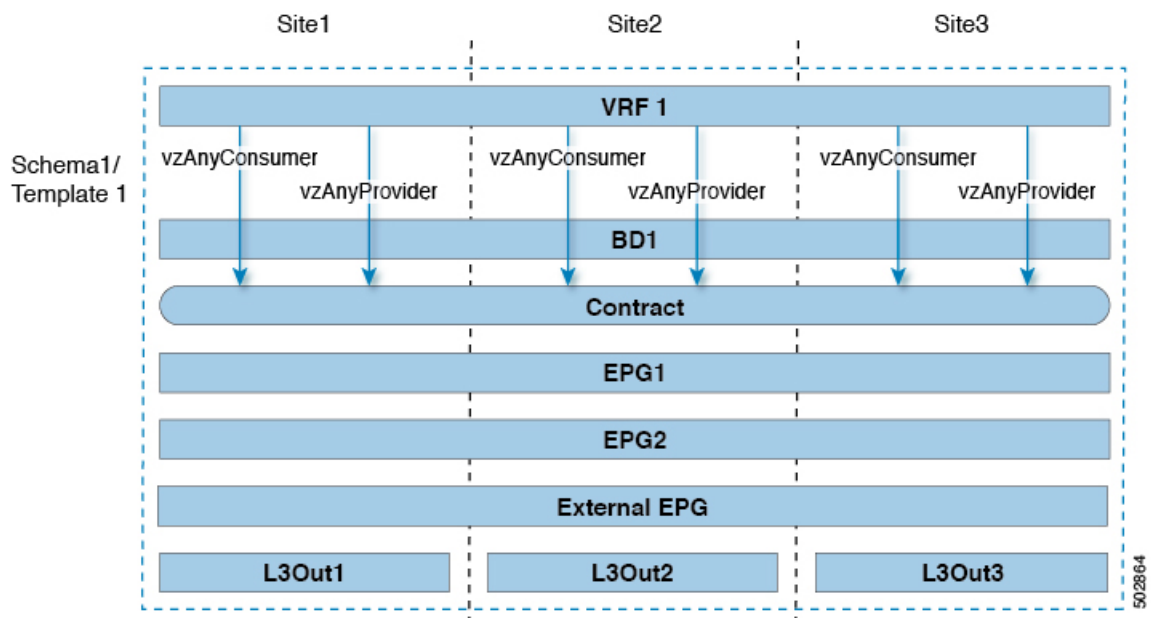
The following example shows intra-VRF communication between EPGs or External EPGs all of which are stretched between sites. In this example EPG1 and EPG2 are mapped to the same BD1, but they could each be part of different BDs as long as both BDs are part of VRF1.

In this case you can create all objects within the same template and then deploy the template to all sites.



Note As a best practice, it is recommended that the L3Out objects should instead remain defined only on Cisco APIC or configured on-site local templates on MSO.

Figure 46:



Site-Local EPGs

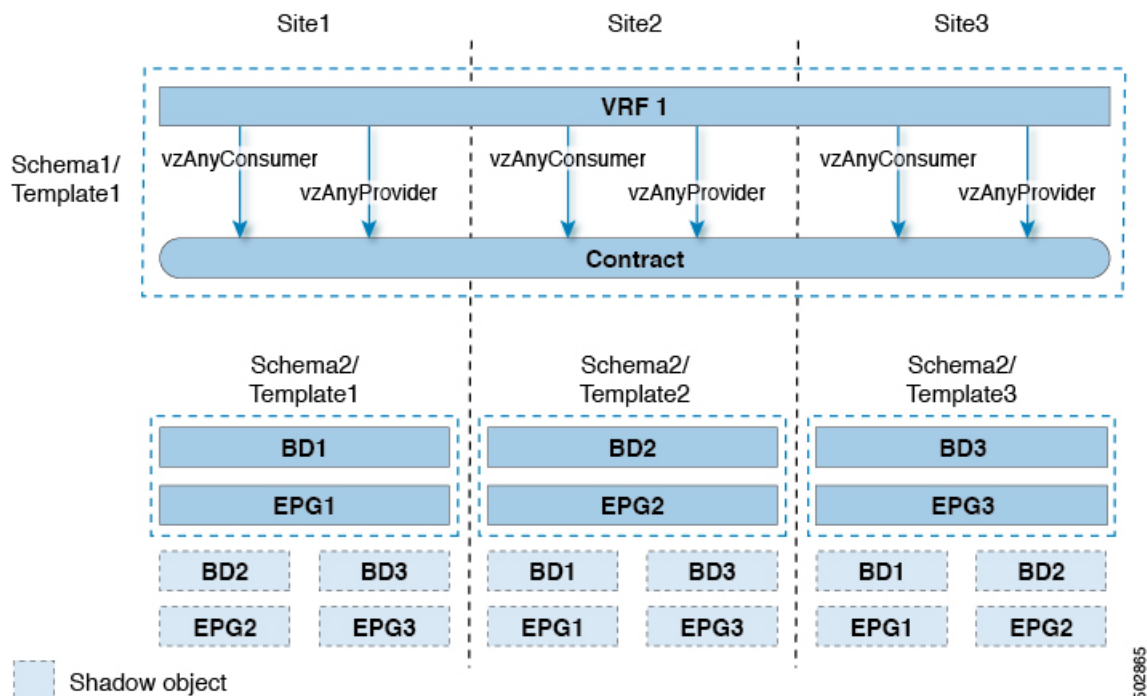
The following example shows intra-VRF communication between EPGs or External EPGs where none of the EPGs are stretched but can still freely communicate with each other since vzAny consumes and provides the "permit-any" contract.

In this case you will need to create multiple templates:

- A single template for the shared objects (VRF, Contract) deployed to every site.
- And a separate template for every site containing the EPG and BD deployed that site.

For the objects that are not stretched, shadow objects are created in other sites.

Figure 47:



502865

Combination of Site-Local and Stretched EPGs

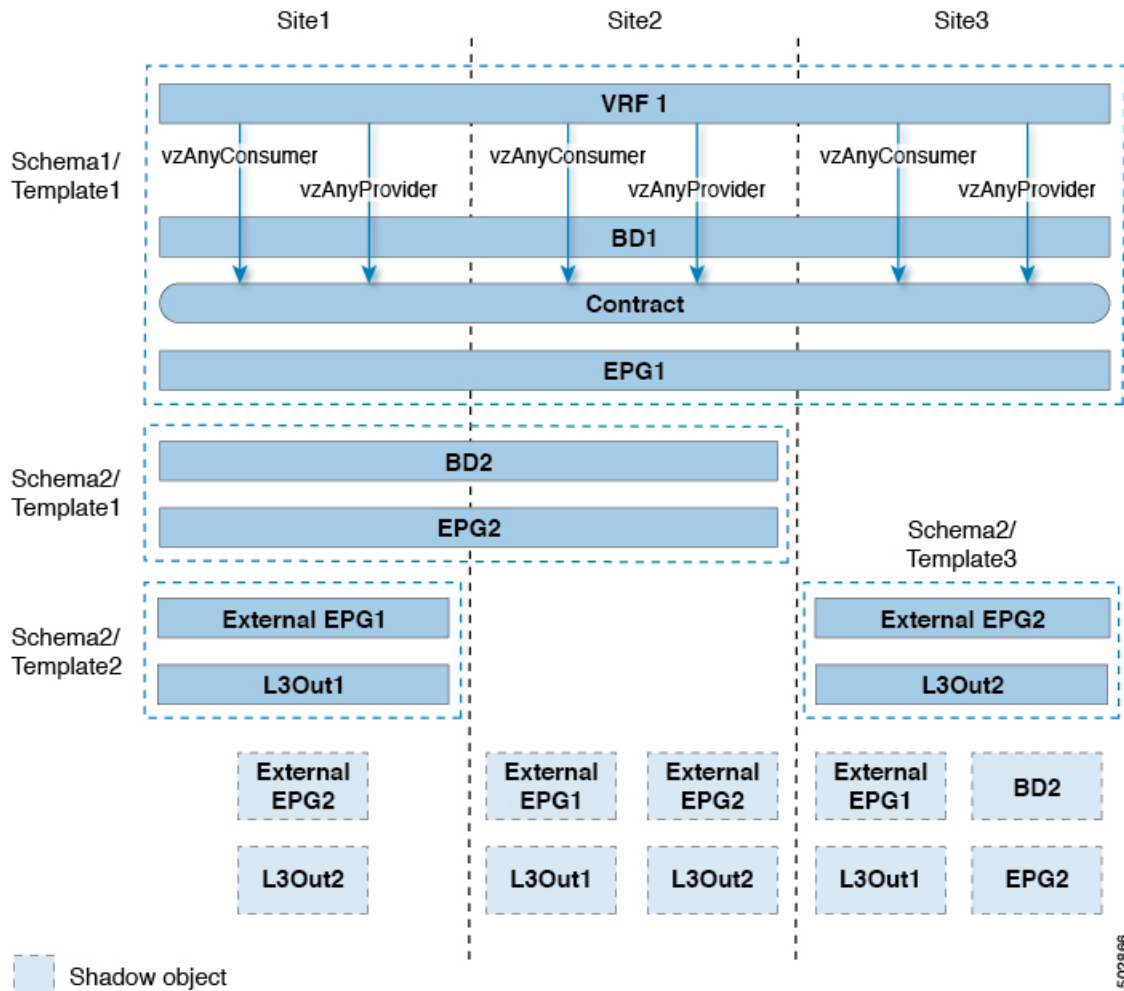
The following example shows intra-VRF communication between EPGs or External EPGs where some EPGs are stretched while others are deployed to a single site only. All EPGs can still freely communicate with each other since *vzAny* consumes and provides the "permit-any" contract.

In this case you will need to create multiple templates:

- A single template for the shared objects (VRF, Contract, BDs) deployed to every site.
- And a separate template for every site combination containing the objects deployed only to those sites.

For the objects that are not stretched, shadow objects are created in other sites.

Figure 48:



Intra-VRF Intersite L3Out

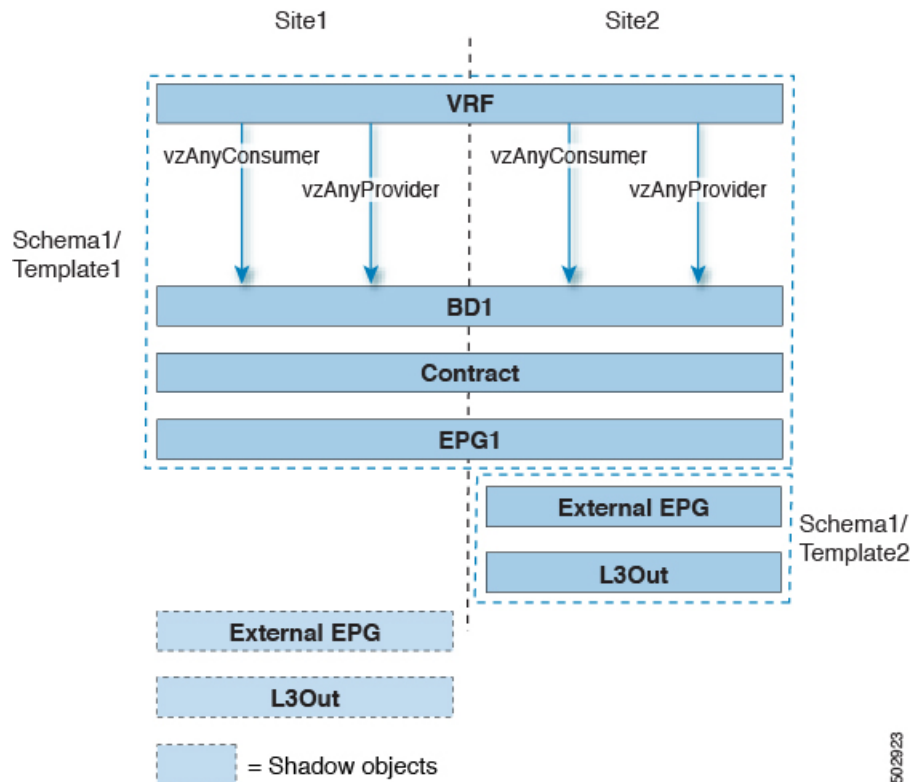
This use case allows you to configure an intersite L3Out for multiple EPGs within a vzAny VRF. When the L3Out's external EPG is in the same VRF, you do not need to explicitly add the provider contract to the external EPG.

Keep in mind, when configuring an intersite L3Out, you must configure a routable TEP pool for each Pod. Additional intersite L3Out details and requirements are described in the [Intersite L3Out Overview](#), on page 259 section.

In this case you will need to create multiple templates:

- A single template for the shared vzAny objects (VRF, Contract, BD) deployed to one or more sites.
- And a separate template for every site combination containing the objects deployed only to those sites.

Figure 49:



Based on the configuration shown in the above figure, endpoints that are part of the stretched EPG1 and connected to Site1 will be able to communicate with the external network domain via the L3Out connection deployed in Site2. The same would apply for endpoints that are part of site-local EPGs deployed in Site1.

Inter-VRF Intersite L3Out

This use case allows you to enable vzAny contracts between a consumer VRF and L3Out external EPGs in a different provider VRF. Multiple EPGs that are part of the vzAny consumer VRF can communicate with a single EPG that is providing a shared service in a provide VRF. vzAny contract acts as a contract for all EPGs in a VRF. Each of the participating VRFs and L3Out external EPG can be stretched across sites.



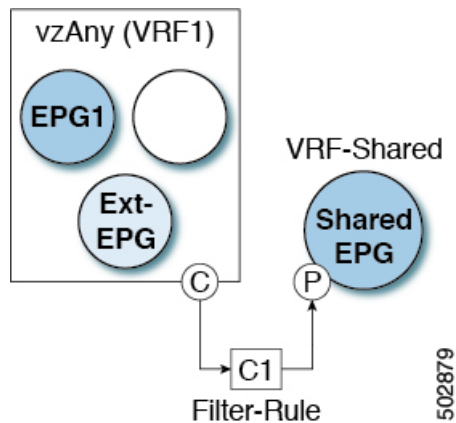
Note A VRF cannot be a vzAny provider.

Many-to-One Communication

The following three sections provide schema examples of multiple EPGs that are part of the same vzAny VRF communicating with a single EPG that is providing a shared service. In this case, the contract can specify one or more filter rules.

The EPG providing shared services can be in a separate VRF (as shown in the figure below) or it can be part of the vzAny VRF.

Figure 50:



For all the following use cases, you will need to create the same objects and policies summarized below. However, the schema and template design will depend on the number of sites as well as which objects are going to be stretched. The specific sections below contain recommendation on template layout.

Procedure

-
- Step 1** Create a Schema.
- Step 2** Create a common Template used to deploy configuration objects in all the sites (that is *stretched objects*).
- Step 3** Create any additional templates for every combination of sites where EPGs will be deployed .
- Step 4** Within the common Template, create the contract and filters to be consumed by vzAny and provided by the EPG offering shared services.
- This is described in [Create Contract and Filters, on page 366](#).
- Step 5** Within the common Template, create a VRF and configure vzAny to consume the previously defined contract.
- This is described in [Configure vzAny to Consume/Provide a Contract, on page 367](#).
- Step 6** Within each site's template, create and configure the EPGs that are part of the vzAny VRF.
- This is described in [Create EPGs to Be Part of the vzAny VRF, on page 367](#).
- Step 7** Create new or configure existing provider EPG or external EPG.
- You create and configure the provider EPG or external EPG as you typically would.
- Step 8** Assign the Contract to the provider EPG.
- In addition to assigning the contract to be consumed by vzAny, you will also need to assign the same contract to the provider EPG.
-

Provider EPG Within vzAny VRF

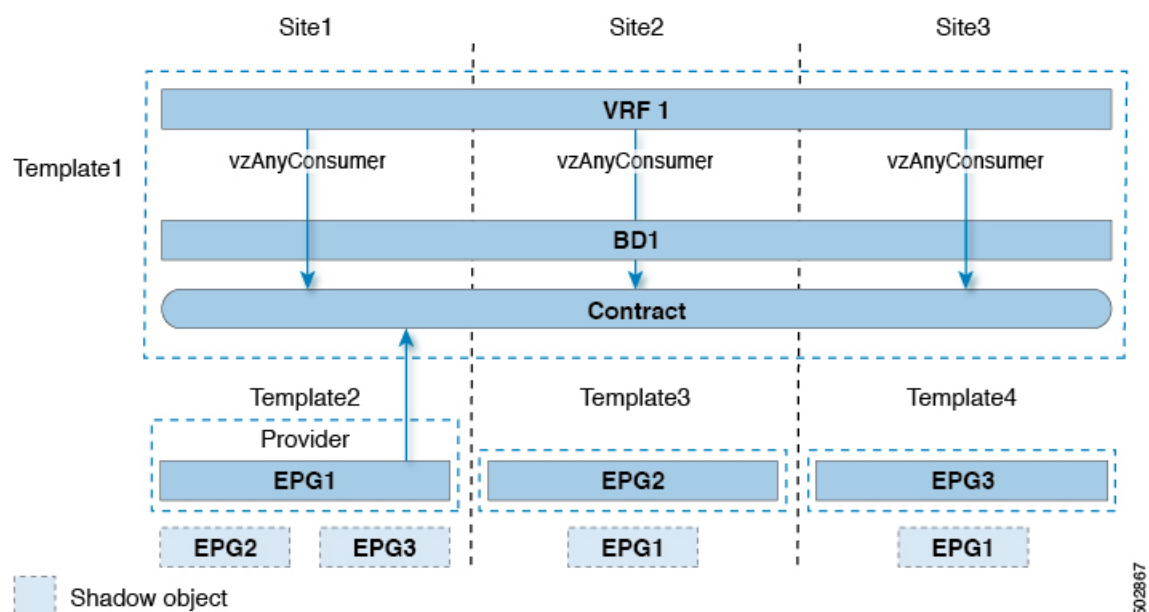
The following example shows intra-VRF communication between a single provider EPG (for example, shared service) and all other EPGs within the same VRF consuming the service.

In this case you will need to create multiple templates:

- A single template for the shared objects (VRF, Contract, BDs) deployed to every site.
- And a separate template for every site combination containing the objects deployed only to those sites.

The following figure shows a single stretched VRF/BD configuration. Alternatively, you can also configure and map a dedicated BD for each EPG, in which case shadow BDs would be deployed in the remote sites.

Figure 51:



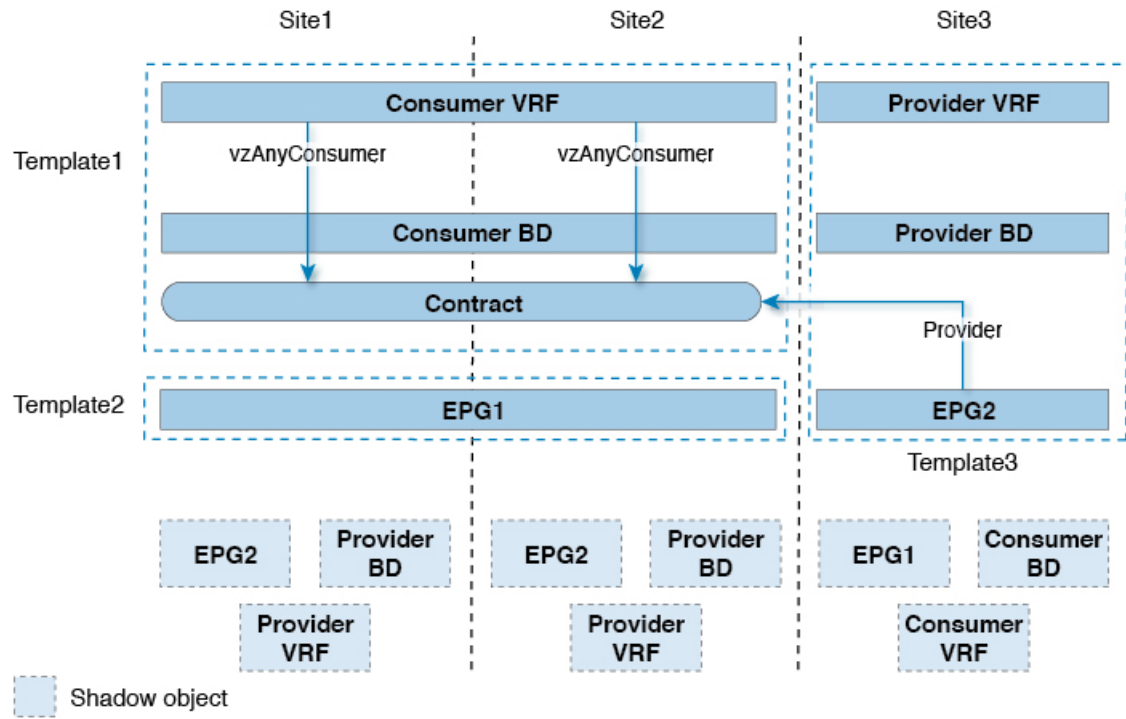
Provider EPG In Its Own VRF

The following example shows communication between a single EPG (for example, shared service provider) in its own VRF and all EPGs within a different, vzAny VRF. The provider EPG can be deployed to the same or a different site as the consumer EPGs in the vzAny VRF.

In this case you will need to create multiple templates:

- A single template for the shared vzAny objects (VRF, Contract, BD) deployed to one or more sites.
- And a separate template for every site combination containing the objects deployed only to those sites.

Figure 52:



502868



CHAPTER 31

vzAny with PBR

- [vzAny with PBR Overview, on page 377](#)
- [vzAny with PBR Guidelines and Limitations, on page 385](#)
- [Create Service Device Template, on page 387](#)
- [Create Application Template, on page 394](#)
- [Add Service Chaining to Contract, on page 398](#)

vzAny with PBR Overview

The following sections provide an overview, requirements and guidelines, and configuration steps for enabling vzAny contracts with Policy-Base Redirects (PBR) in your Multi-Site domain. For an overview of vzAny in general and basic vzAny use cases that do not include PBR, see the [vzAny Contracts, on page 363](#) chapter instead.

Use Cases

Prior to release 4.2(3), the following basic vzAny use cases (without PBR) were supported with Multi-Site, all of which are described in the [vzAny Contracts, on page 363](#) chapter:

- Free communication between EPGs within the same VRF.
- Many-to-one communication allowing all EPGs within the same VRF to consume a shared service from a single EPG that is in the same or different VRF.

Beginning with NDO release 4.2(3), the following additional use cases for vzAny with PBR are supported for ACI fabrics running APIC release 6.0(4) or later, which allow redirecting traffic to a logical firewall service connected in each site in one-arm mode:

- Any intra-VRF communication (vzAny-to-vzAny) between two EPGs or External EPGs within the same VRF.
- Many-to-one communication between all the EPGs in a VRF (vzAny) and a specific EPG that is part of the same VRF.
- Many-to-one communication between all the EPGs in a VRF (vzAny) and a specific External EPG that is part of the same VRF.

General Workflow for Configuring vzAny with PBR

The following sections describe how to create and configure the individual building blocks (such as templates, EPGs, contracts) that are required for all of the vzAny with PBR use cases followed by user-case-specific sections that provide the workflows necessary to put the individual building blocks together for the specific use case you want to configure.

When configuring any of the vzAny with PBR use cases, you will go through the following workflow which includes the new Service Device templates introduced in release 4.2(3) and used to define service graph configurations:

1. Create a Service Device template and associate it to a specific tenant and to all the sites where the configuration is required, which includes:

- (Optional) Referencing an IP SLA policy.

The IP SLA policy must be already defined in a Tenant Policy template associated to the same tenant.

- Creating one or more service node devices in the Service Device template.

Note that when you create a service device configuration, you will need to provide a bridge domain which must already exist in one of the Application templates. The exact BD requirements are listed in the following [vzAny with PBR Guidelines and Limitations](#), on page 385 section.

- Providing site-level configurations for the service node device defined in the Service Device template and deploying it.



Note Beginning with release 4.2(3) and the introduction of Service Device templates, there's no Service Graph object that must be explicitly created in Nexus Dashboard Orchestrator for PBR use cases. NDO implicitly creates the service graph and deploys it in the site's APIC.

2. Complete the configuration for the specific tenant associated to the Service Device template that you just created, which includes:

- Creating a Tenant Application template and assigning it to all sites where the configuration is required.
- Configuring vzAny VRF settings required to enable PBR and a contract.
- Configuring the consumer and provider EPGs.

While the service BD must be stretched across sites, the BDs you use for the EPGs can be stretched or site-local.

3. Associate the service device you created in Step 1 with the vzAny contract you created in Step 2.



Note Please refer [ACI Contract Guide](#) and [ACI PBR White Paper](#) to understand Cisco ACI contract and PBR terminologies.

Traffic Flow: Intra-VRF vzAny-to-vzAny

This section summarizes the traffic flow between two EPGs that are part of the logical vzAny construct for a given VRF in different sites. In this use case, vzAny is both the provider and the consumer of a PBR contract.

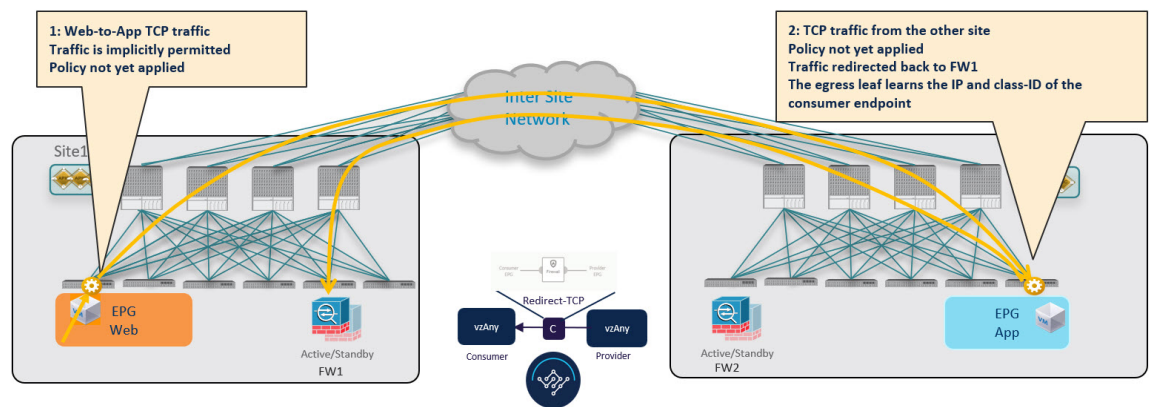


Note In this case, the traffic flow in both directions is redirected through both firewalls in order to avoid asymmetric traffic flows due to independent FW nodes deployed in the two sites.

Initial Consumer-to-Provider Traffic Flow and Conversational Learning

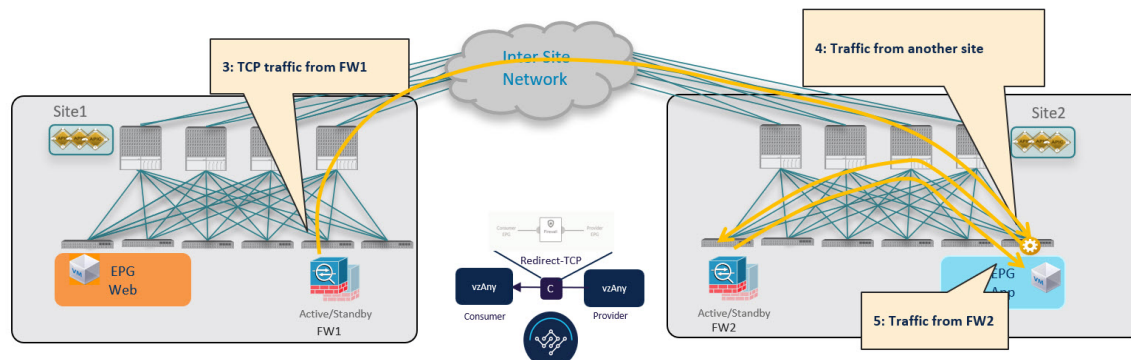
The design principle for redirecting the traffic to the FW service nodes in both the local site and the remote site is that the PBR policy should always be applied on the ingress leaf switch for both directions of the traffic flow. For this to happen, the ingress leaf switch must be aware of the destination's endpoint policy information (Class-ID). The figure below shows an example where communication is initiated from the consumer endpoint, and the ingress (consumer) leaf switch does not yet have the Class-ID information for the destination (provider) endpoint. So the traffic is simply forwarded toward the destination connected to the remote site. This release implements a new logic to support this use case, so that the provider leaf switch that receives the traffic can understand that the flow originated in Site 1 but it has not been sent through the firewall service node connected in that site. As a result, after learning the consumer endpoint information (Class-ID), the provider leaf in Site 2 bounces back the traffic toward the firewall in Site 1.

Figure 53: Conversational Learning



The firewall in Site 1 applies the security policy, then the traffic is forwarded again to the destination leaf switch in Site 2. This leaf is now able to understand that, while the traffic is still coming from Site 1, it now has been sent through the firewall deployed in that site. As a result, the destination leaf switch forwards the packet to its local firewall device for inspection and after that it is delivered to the destination endpoint as shown in the following figure.

Figure 54: Conversational Learning

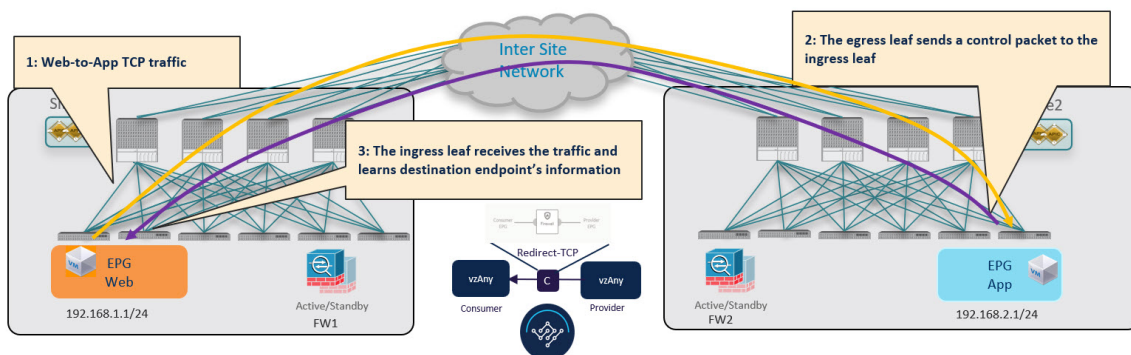


In order to avoid the suboptimal bounce of traffic shown in [Figure 53: Conversational Learning, on page 379](#), the provider leaf switch generates a special control packet and sends it to the consumer leaf switch in Site 1, so that the consumer leaf can learn the provider endpoint's Class-ID information.



Note The same behavior described above for the consumer-to-provider traffic direction applies if the initial flow is established in the provider-to-consumer direction instead.

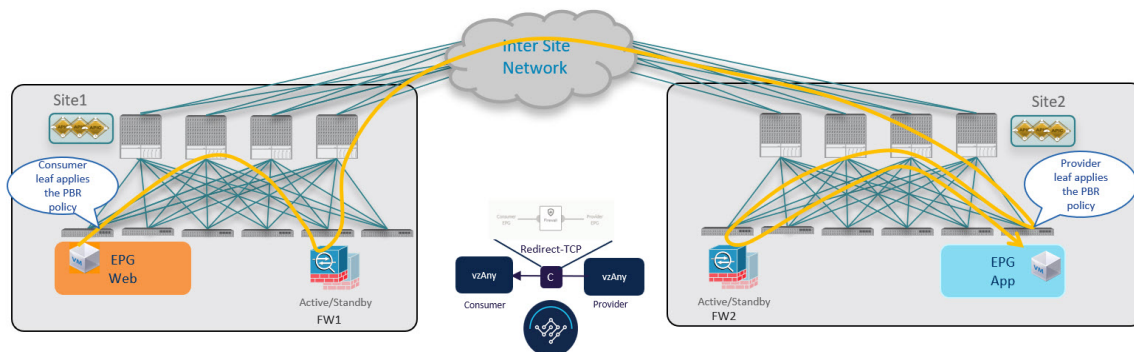
Figure 55: Conversational Learning



Consumer-to-Provider Traffic Flow (at Steady State)

After the consumer leaf switch has learned the provider endpoint information from the conversational learning stage described above, it can apply policy and redirect traffic to its local firewall for all future traffic:

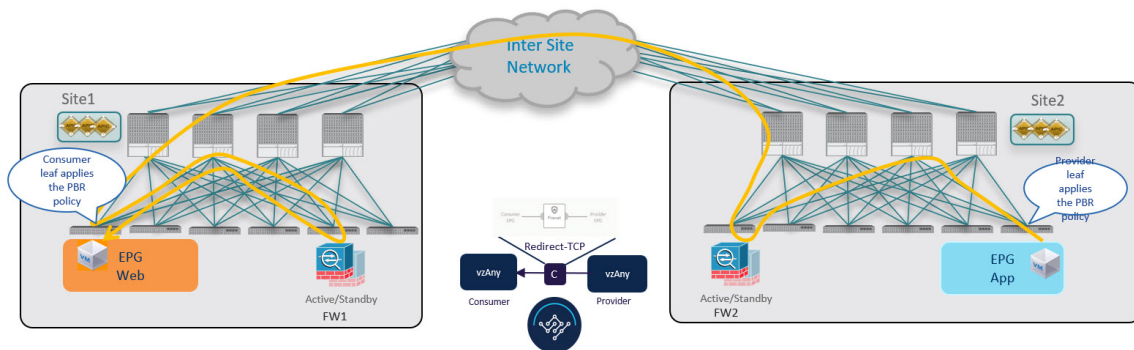
Figure 56: Consumer-to-Provider Traffic Flow



Provider-to-Consumer Traffic Flow (at Steady State)

After the provider leaf switch has learned the consumer endpoint information either from the direct packet shown in [Figure 53: Conversational Learning, on page 379](#) or based on conversational learning, it can apply policy and redirect traffic to its local firewall for all future traffic:

Figure 57: Provider-to-Consumer Traffic Flow



Traffic Flow: Intra-VRF vzAny-to-EPG

This section summarizes the traffic flow between a consumer EPG that is part of the logical vzAny construct for a given VRF and a provider EPG that is part of the same VRF. In this use case, vzAny is the consumer of the PBR contract, whereas a specific EPG is the provider.



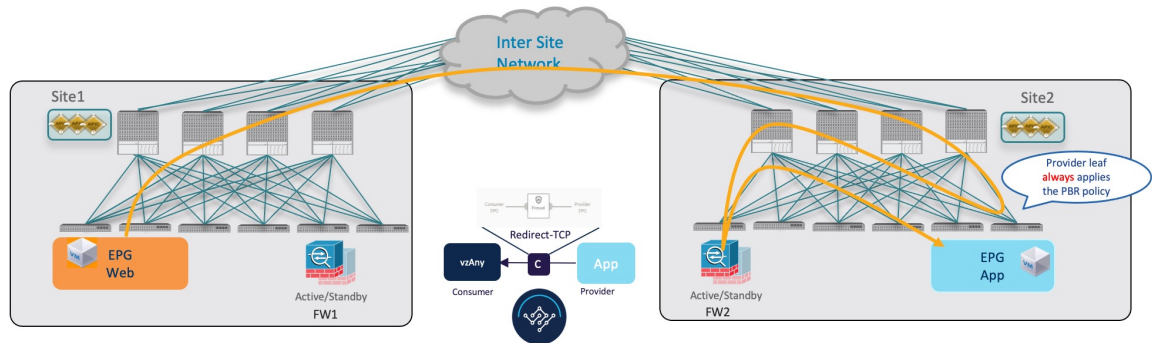
Note Unlike the vzAny-to-vzAny and vzAny-to-L3Out use cases where traffic always flows through the firewall devices in both sites, vzAny-to-EPG always uses only the device in the provider's site.

Consumer-to-Provider Traffic Flow

For the vzAny-to-EPG use case, policy is applied on the provider leaf switch only regardless of the traffic direction. So for consumer-to-provider traffic, the consumer EPG sends traffic directly to the provider EPG's

leaf switch, which learns the consumer endpoint information (Class-ID) and redirects the traffic to its local firewall for inspection:

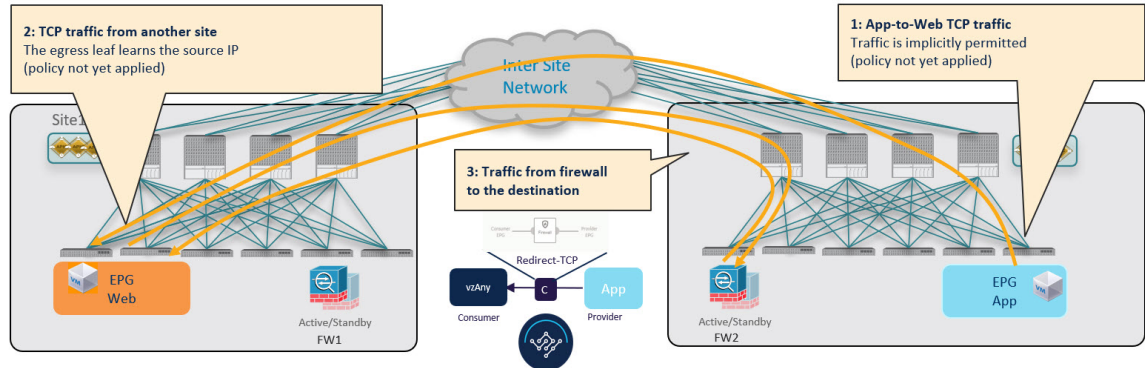
Figure 58: vzAny-to-EPG Consumer-to-Provider Traffic Flow



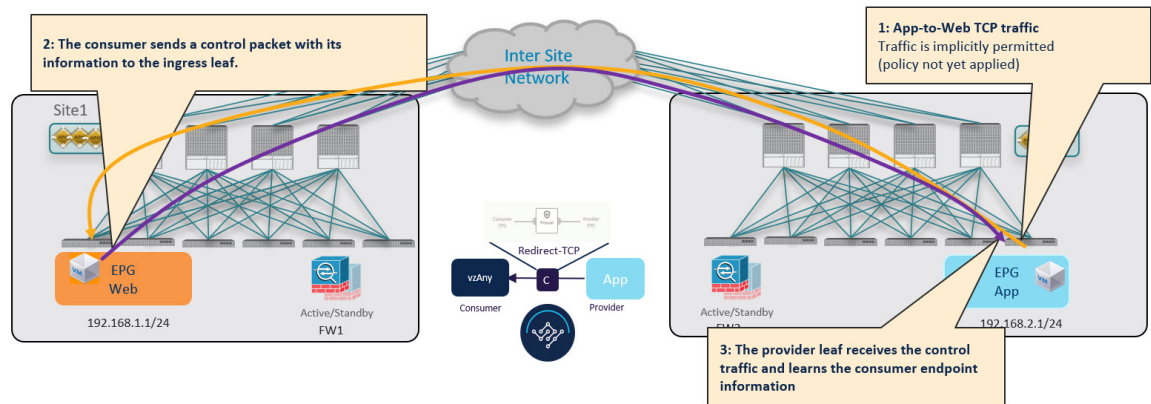
Provider-to-Consumer Traffic Flow (Initial Traffic and Conversational Learning)

If the communication is initiated by the provider endpoint before the provider leaf switch can learn the consumer endpoint information (Class-ID), it cannot apply the policy to redirect traffic to its local firewall, so the traffic is sent across sites to the consumer leaf switch. Because the policy was not applied (indicated by a control bit in the packet), the consumer leaf switch redirects the traffic back to the provider site's firewall for inspection, which finally bounces the traffic back to the consumer endpoint.

Figure 59: vzAny-to-EPG Provider-to-Consumer Traffic Flow (Initial Traffic and Conversational Learning)



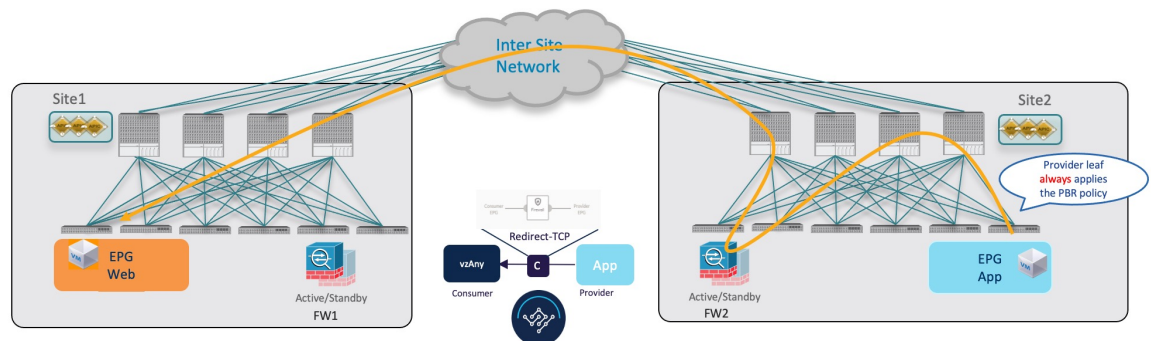
While this suboptimal traffic flow can continue indefinitely, the consumer EPG's leaf switch also sends a separate control packet to the provider leaf switch with consumer endpoint information in order to optimize future traffic and prevent it from bouncing between both sites:



Provider-to-Consumer Traffic Flow (at Steady State)

After the provider leaf switch has learned the consumer endpoint information either from the direct packet originated from the consumer endpoint shown in [Figure 58: vzAny-to-EPG Consumer-to-Provider Traffic Flow, on page 382](#) or based on conversational learning, it can apply policy and redirect traffic to its local firewall for all future traffic:

Figure 60: vzAny-to-EPG Provider-to-Consumer Traffic Flow



Traffic Flow: Intra-VRF vzAny-to-External-EPG (L3Out EPG)

This section summarizes the traffic flow between an EPG that is part of the logical vzAny construct for a given VRF and an external EPG (L3Out EPG) that is part of the same VRF in another site. In this use case, vzAny is the consumer of a vzAny contract, while an External EPG associated to the L3Out is the provider.



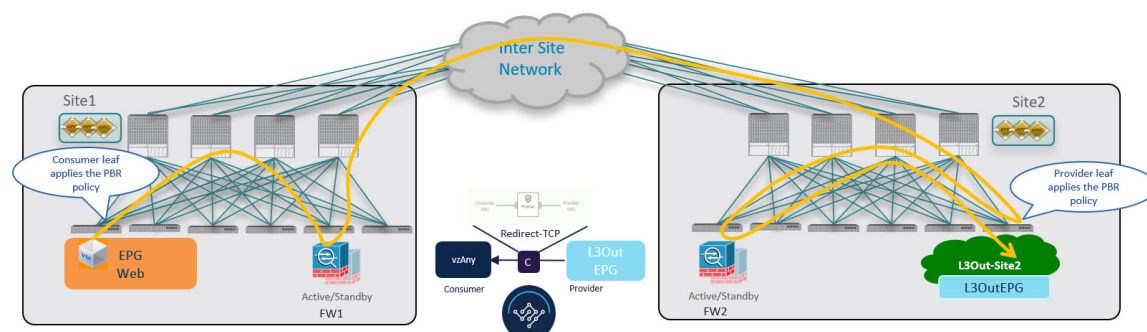
Note In this use case, the traffic is always redirected through firewall devices in both sites.

Consumer-to-Provider Traffic Flow

The ingress leaf switch can always resolve the class ID of the destination external EPG and applies the PBR policy redirecting the traffic to the local FW, so no conversational learning is necessary for traffic in this direction. Because the traffic is received by the provider leaf switch after going through the firewall node in

Site 1, it is not possible for the provider leaf switch to learn the consumer endpoint information (Class-ID) from this data-plane communication.

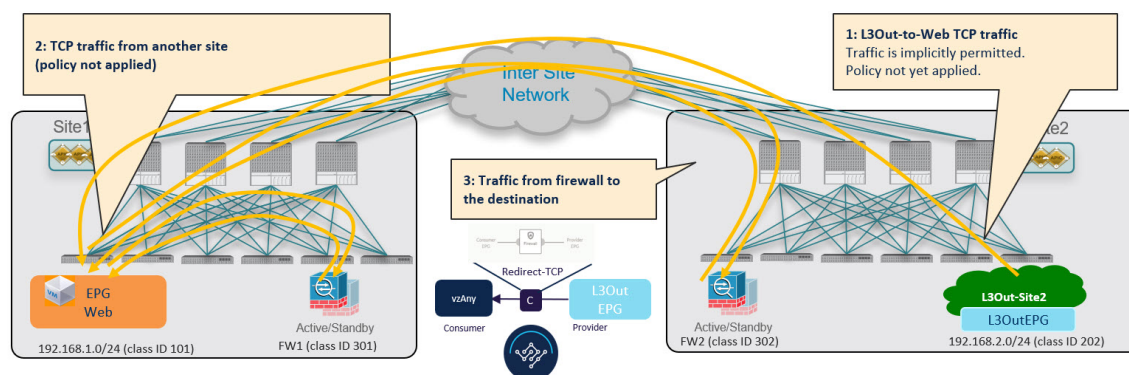
Figure 61: vzAny-to-External EPG Consumer-to-Provider Traffic Flow



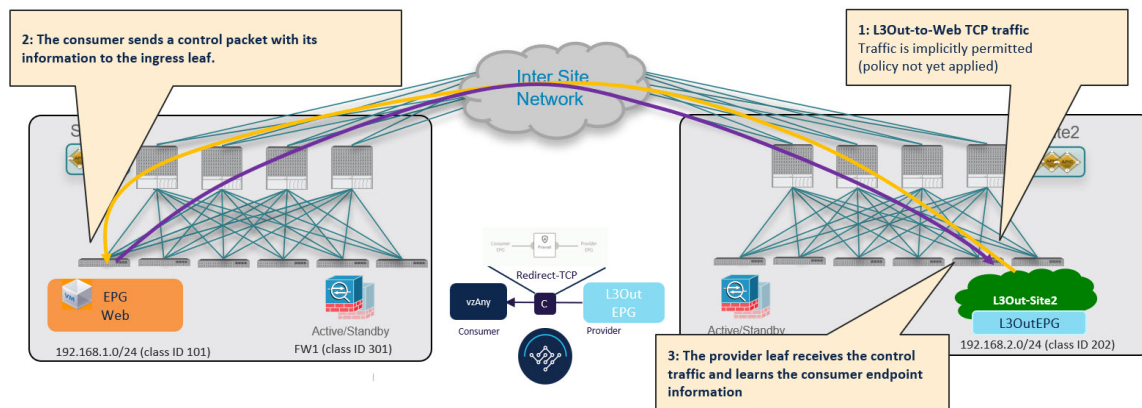
Provider-to-Consumer Traffic Flow (Initial Traffic and Conversational Learning)

Before the provider leaf switch learns the consumer endpoint information, it cannot apply the policy to redirect traffic to its local firewall, so the traffic is sent across sites to the consumer leaf switch. Because the policy was not applied (indicated by a control bit in the packet), the consumer leaf switch redirects the traffic back to the provider site's firewall for inspection, which finally forwards the traffic back to the consumer endpoint.

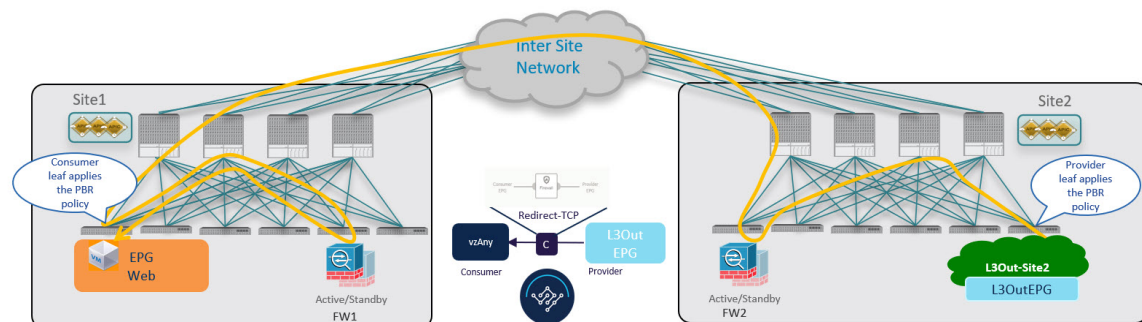
Figure 62: vzAny-to-L3Out Provider-to-Consumer Traffic Flow (Initial Traffic and Conversational Learning)



While this traffic flow can continue indefinitely, the consumer leaf switch also sends a separate control packet to the provider leaf switch with consumer endpoint information in order to optimize future traffic and prevent it from bouncing between both sites:

Figure 63: vzAny-to-L3Out Provider-to-Consumer Traffic Flow (Initial Traffic and Conversational Learning)**Provider-to-Consumer Traffic Flow (at Steady State)**

After the provider leaf switch has learned the consumer endpoint information, it applies the PBR policy to redirect traffic to its local firewall device first, which then sends traffic across sites to the consumer leaf switch, which redirects traffic to the firewall device in its site and then finally to the consumer endpoint.

Figure 64: vzAny-to-L3Out Provider-to-Consumer Traffic Flow

vzAny with PBR Guidelines and Limitations

The following guidelines and limitations apply when using vzAny with PBR in Multi-Site deployments:

**Note**

The following sections apply to the vzAny with PBR use cases only. For information about basic vzAny concepts and use cases, see the [vzAny Contracts](#), on page 363 chapter instead.

- When configuring MultiSite Policy-Based Redirect (PBR) between Endpoint Groups (EPGs), the following features are not supported for specific IP endpoints or host prefixes (/32 for IPv4 and /128 for IPv6):
 - Static Route on a Bridge Domain (NH Reachability)
 - Microsoft Network Load Balancing

- Anycast MAC
- The ACI sites must be running Cisco APIC release 6.0(4) or later.
- This release supports redirecting vzAny traffic to a single-node firewall or to a single-node load-balancer with a single interface attached to the service bridge domain. Single-Node firewall is supported in vzAny-to-vzAny and vzAny-to-L3Out, vzAny-to-EPG, and L3Out-to-L3Out use cases. Single-Node load-balancer is supported in vzAny-to-EPG use case.

This includes the following three use cases for one-arm mode firewall service graphs:

- Any intra-VRF communication (vzAny-to-vzAny) across sites.
- Many-to-one communication between all the EPGs in a VRF (vzAny) and a specific EPG that is part of the same VRF.
- Many-to-one communication between all the EPGs in a VRF (vzAny) and a specific External EPG that is part of the same VRF.

In all of the above cases, conversational endpoint learning is enabled only when vzAny with PBR is configured and is used when IP prefixes under the EPG are not configured. A mix of EPGs with IP prefixes and EPGs without IP prefixes is also supported.

- While you can use your existing Service Graph objects defined in Application templates for these use cases, we recommend using the new service chaining workflows introduced in release 4.2(3) and implicitly creating new service graphs by defining the policies in Service Device templates and associating them to contracts.

The steps described in the following sections use the new Service Device templates to enable the supported use cases but will call out the specific differences when applicable.



Note Configuration of Service Graph objects in Application templates will be deprecated in a future release.

- The vzAny VRF must be stretched across the sites.

Note that the "Site-aware policy enforcement" and "L3 Multicast" options must be enabled for the vzAny VRF to enable the vzAny PBR use cases discussed in this chapter.

The following sections assume that you already have a VRF for which you have or will enable vzAny and which you will use for these use cases.

If you do not already have a VRF, you can create one in an Application template as you typically would. VRF configuration is described in detail in [Configuring VRFs, on page 71](#).

- The service BD to which you want to attach the service device interface must be L2 stretched (BUM forwarding is instead optional and should be disabled).

If you do not already have a service BD, you can create one in an Application template as you typically would. BD configuration is described in detail in [Configuring Bridge Domains, on page 72](#).

- The consumer, provider, and the service BDs must be configured in Hardware proxy-mode.
- The vzAny PBR destination must be connected to a stretched service BD, not to an L3Out.
- Only threshold down deny action and sip-dip-protocol hash is supported .

- The PBR destination node must be in either the consumer or provider VRF instance. For example see [Cisco Application Centric Infrastructure Policy-Based Redirect Service Graph Design White Paper](#).

The following is not supported for vzAny with PBR use case:

- Specific Remote Leaf configurations.
 - Specific considerations apply for Multi-Site deployments leveraging Remote Leaf nodes. Intersite transit routing with PBR is not supported on vzAny PBR and L3Out-to-L3Out for communication between endpoints (consumer or provider) deployed on remote leaf nodes that belongs to different sites.
- Each VRF is limited to utilizing only one device in a one-arm configuration for vzAny-to-vzAny, L3OutEPG-to-L3OutEPG, and vzAny-To-L3OutEPG Policy-Based Routing (PBR). This restriction is enforced due to special ACL in APIC.
- We must use different firewall VLAN interfaces for redirection for vzAny-to-vzAny/L3OutEPG-to-L3OutEPG, and other use cases such as vzAny-to-EPG, EPG-to-EPG and EPG-to-L3OutEPG if they are in the same VRF.
- When vzAny with PBR is applied to north-south communication for any of the newly supported use cases (vzAny-to-vzAny, vzAny-to-L3OutEPG, L3OutEPG-to-L3OutEPG), ingress traffic optimization needs to be enabled for stretched subnets.
- Only one node service chain with L3 PBR destination is supported.
- Contract Permit logging is not supported on the VRF that has Site-aware Policy Enforcement Mode is enabled, which is required for vzAny PBR and L3OutEPG-to-L3OutEPG PBR.
- Pod-aware vzAny with PBR is not supported.

Create Service Device Template

The following steps describe how to create a Service Device template with a service node and its settings which you will use for the vzAny with PBR use cases.

Before you begin

- Ensure that you have read and completed the requirements described in [vzAny with PBR Guidelines and Limitations, on page 385](#).
- You must have created a stretched service bridge domain (BD) to use with the service nodes you will define in this section.

If you do not already have a BD, you can create one in an Application template as you typically would. BD configuration is described in detail in [Configuring Bridge Domains, on page 72](#).

Procedure

-
- Step 1** Log in to the Nexus Dashboard Orchestrator GUI.
- Step 2** From the left navigation pane, select **Configure > Tenant Templates**.

Step 3 (Optional) Create a Tenant Policies template and an IP SLA monitoring policy.

We recommend that you configure an IP SLA policy for traffic redirection as it simplifies the configuration of the PBR policy described in Step 7 below. If you have an IP SLA policy already defined, you can skip this step, otherwise:

- a) Choose the **Tenant Policies** tab.
- b) On the **Tenant Policies** page, click **Create Tenant Policy Template**.
- c) In the **Tenant Policies** page's right properties sidebar, provide the **Name** for the template and **Select a Tenant**.
- d) In the **Template Properties** page, choose **Actions > Add/Remove Sites** and associate the template with both sites.
- e) In the main pane, choose **Create Object > IP SLA Monitoring Policy**.
- f) Provide the **Name** for the policy, and define its settings.
- g) Click **Save** to save the template.
- h) Click **Deploy Template** to deploy it.

Step 4 Create a Service Device template and associate it with a tenant and with the sites.

- a) From **Configure > Tenant Templates**, choose the **Service Device** tab.
- b) Click **Create Service Device Template**.
- c) In the template properties sidebar that opens, provide the **Name** for the template and **Select a Tenant**.
- d) In the **Template Properties** page, choose **Actions > Add/Remove Sites** and associate the template with both sites.
- e) Click **Save** to save the template.

Step 5 Create and configure the device cluster.

- a) In the **Template Properties** page (template-level configuration), choose **Create Object > Service Device Cluster**.

The device cluster defines the service to which you want to redirect traffic. This release supports redirection to a firewall service node that can be deployed with three different redundancy models: active/standby, active/active, or a cluster of multiple independent nodes. The provisioning for those different options is covered in Step 7 below. Note that you can choose the firewall deployment model at the site level and different options can be deployed across different fabrics that are part of the same Multi-Site domain.

- b) In the **<cluster-name>** sidebar, provide the **Name** for the cluster.

The **Device Location** and **Device Mode** are pre-populated based on the currently supported use case. **Device Location** should be pre-configured as **ACI On-Prem** and **Device Mode** as **L3**.

- c) For **Device Type**, choose **Firewall**.

This release supports only firewall devices for the vzAny with PBR use cases.

- d) For **Device Mode**, choose **L3**.

- e) For **Connectivity Mode**, choose **One Arm**.

This release supports only one-arm device for the vzAny with PBR use cases.

Note

When changing the device connectivity mode between one arm, two arm and advanced mode, the name of the device interface might change in the process. A warning message will alert the user, and any attempt to modify the interface will be restricted if the interface is currently in use by a contract. If the user wishes to preserve the previously used interface name and avoid disrupting the deployed configuration, they may choose to override the name change during the modification process.

Note

Validations are conducted only for one-arm and two-arm modes. In Advanced mode, no validations are performed, and it is assumed that the user is an expert when choosing this mode.

- f) Provide the **Interface Name**.
- g) For the **Interface Type**, choose **BD**.

For vzAny with PBR use cases, this release supports attaching the service device to a bridge domain only.

- h) Click **Select BD >** to choose the service bridge domain to which you want to attach this device.
This is the stretched service BD you created as part of the [vzAny with PBR Guidelines and Limitations, on page 385](#), for example `FW-external`.
- i) For the **Redirect** option, choose **Yes**.
You must choose to enable redirect for the PBR use case. After choosing **Yes**, the **IP SLA Monitoring Policy** option becomes available.
- j) (Optional) Click **Select IP SLA Monitoring Policy** and choose the IP SLA policy you have created in a previous step.
- k) (Optional) In the **Advanced Settings** area, choose **Enable** if you want to provide additional settings for the service cluster.

You can configure the following advanced settings:

- **QoS Policy** – allows you assign a specific QoS level within the ACI fabrics for the redirected traffic.
- **Preferred Group** – specifies whether or not this service device interface is part of the preferred group.
Leave this option disabled when configuring a vzAny use case.
- **Load Balancing Hashing** – allows you to specify the hashing algorithm for PBR load balancing.

Note

You must keep the default value for the vzAny-to-vzAny, vzAny-to-ExtEPG, and ExtEPG-to-ExtEPG use cases as they support only the default configuration. You can change the load balancing hashing for other use cases: EPG-to-EPG, ExtEPG-to-EPG and vzAny-to-EPG.

For additional information, see [ACI Policy-Based Redirect Service Graph Design](#).

- **Pod Aware Redirection** – can be configured in Multi-Pod configuration if you want to specify the preferred PBR node. When you enable Pod-aware redirection, you can specify the Pod ID and redirection is programmed only in the leaf switches located in the specified Pod.
- **Rewrite Source MAC** – updates the source MAC address if the PBR node uses “source MAC based forwarding” instead of IP based forwarding.

For additional information, see [ACI Policy-Based Redirect Service Graph Design](#).

- **Advanced Tracking Options** – allows you to configure a number of advanced settings for the service node tracking. For additional information, see [Policy-Based Redirect and Threshold Settings for Tracking Service Nodes](#)

- l) Click **Ok** to save.

Note that after you create the Service Device Cluster, it is highlighted in red in the **Template Properties** (template-level configuration) page. At this point, you have defined redirection to a firewall service, but you must still provide the firewall information and the redirect policy you want to use at the site-local level.

Step 6 Provide site-local configuration for the Service Device Cluster you created in the previous step.

- a) In the **Service Device Template** screen, choose the **<site-name>** tab.

- b) At the site level, choose the Service Device Cluster you created.
- c) In the properties sidebar, choose the **Domain Type**.

You can choose whether the firewall device in this site is *Physical* or *VMM* (virtual and hosted by a hypervisor that is part of a VMM domain).

- d) Click **Select Domain** to choose the domain to which this firewall device belongs.

You can choose either a physical or a virtual domain.

- If you choose a physical domain, provide the following information:

- **VLAN** – you must provide the VLAN ID used for traffic between the fabric and the firewall device.
- **Fabric to Device Connectivity** – provide the switch node and interface information for the fabric's connectivity to the firewall device.

- If you choose a VMM domain, provide the additional options:

- **Trunking Port** – used to enable tagged traffic for the L4-L7 VM.

By default, the ACI service graph configuration creates access-mode port groups and attaches them to the vNIC of the L4-L7 VM automatically.

- **Promiscuous Mode** – required if the L4-L7 virtual appliance must receive traffic destined to a MAC address that is not the vNIC MAC owned by the VM.
- **VLAN** – optional configuration for VMM domains and will be allocated from the dynamic VLAN pool associated with the domain if not specified.
- **Enhanced LAG Option** – if you are using enhanced LACP for the port channel between the hypervisor and the fabric.
- **VM Name** – choose the firewall's VM from the list of all VMs available in this VMM domain and the interface (**VNIC**) used for the firewall traffic.

Depending on the kind of device cluster you are deploying, click **+Add VM information** to provide additional cluster nodes.

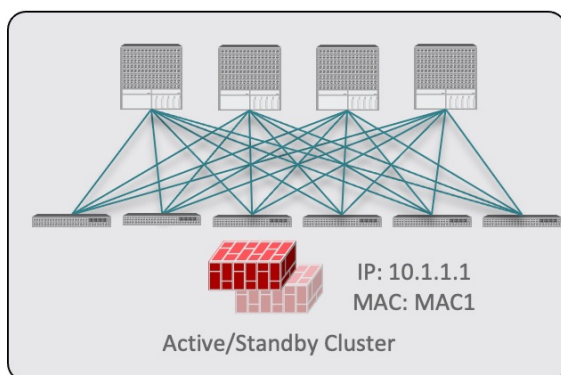
Step 7 Provide the FW device information and PBR destination IP addresses.

As previously mentioned, this release supports 3 deployment options for high-availability FW clusters: active/standby clusters, active/active clusters, and independent active nodes. In all three deployment options, the use of an IP SLA policy (mentioned in Step 3) allows to specify only the IP address of the firewall nodes, and the corresponding MAC address will be automatically discovered.

Note

You can deploy different designs in different sites.

- Active/standby clusters are identified by a single MAC/IP pair.



In this case, you need to provide a single PBR destination IP address identifying the active firewall node and also include information about every node in the cluster.

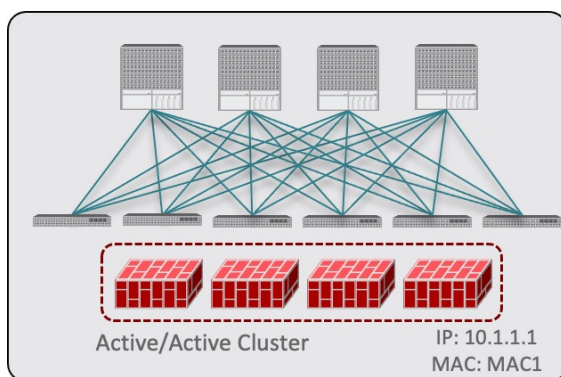
For example, for a 2-nodes active/standby cluster, you would provide the following:

- For a virtual firewall cluster, the VMs representing the active and standby firewall nodes and the IP address of the active firewall as PBR destination.
- For a physical firewall cluster, the interfaces used to connect the active and standby firewall nodes to the leaf switches of the fabric (vPC interfaces in the specific example below) and the IP address of the active firewall as PBR destination.

VM Information* ⓘ			
VM Name*	VNIC*		
vCSA-7-Site1/ASAv-Pod1	Network adapter 2		
vCSA-7-Site1/ASAv-Pod2	Network adapter 2		
Add VM Information			
PBR Destinations			
IP Address *			
50.50.50.10			

Fabric To Device Connectivity ⓘ			
Type *	Pod *	Node *	Path *
Virtual Port Channel	1	101,102	vPC-L101-L102-Port16
Virtual Port Channel	1	103,104	vPC-L103-L104-Port16
Add Fabric To Device Connectivity			
PBR Destinations			
IP Address *			
50.50.50.10			

- Active/active clusters are also identified by a single MAC/IP pair.

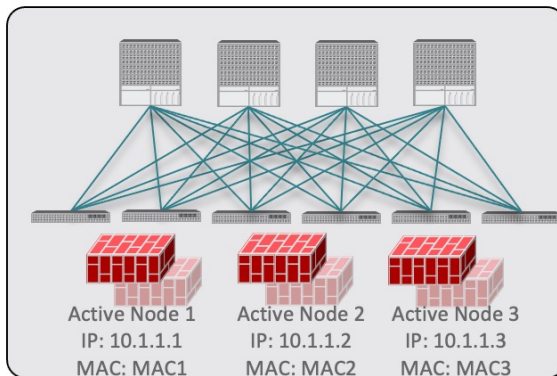


For Cisco Firewalls (ASA or FTD models), the Active/Active cluster is only supported for physical form factors, and all the cluster nodes own the same MAC/IP address and must be connected to the same vPC logical connection deployed on a pair of ACI leaf switches. As a result, the figure below shows how a single vPC interface and a single IP address should be configured on NDO, where the MAC address is dynamically discovered when using an IP SLA policy mentioned for the previous use case.

Type *	Pod *	Node *	Path *
Virtual Port Channel	1	101,102	vPC-L101-L102-Port16
<div> Add Fabric To Device Connectivity </div>			
PBR Destinations			
IP Address * 50.50.50.10			

- For independent active nodes configuration, each active node is identified by a unique MAC/IP addresses pair.

Note that symmetric PBR ensures that the traffic is handled by the same active node in both directions.



In this case, you must provide individual IP addresses for each active node as well as each node's information in your NDO configuration.

For example, for a deployment of 3 independent firewall nodes, you would provide the following:

- For a virtual firewall form factor, the VMs representing the 3 firewall nodes and their unique IP addresses as PBR destinations.
- For a physical firewall form factor, the interfaces used to connect each firewall node to the leaf switches of the fabric (vPC interfaces in the specific example below) and the unique IP addresses of each firewall node as PBR destinations.

The screenshot shows two configuration panels. The top panel, 'VM Information', has a table with columns 'VM Name' and 'vNIC'. It lists three VMs: 'vCSA-7-Site1/ASAv-Pod1', 'vCSA-7-Site1/ASAv-Pod2', and 'vCSA-7-Site1/ASAv-Pod3', all with 'Network adapter 2' as the vNIC. Below this is a 'PBR Destinations' section with a table for IP addresses: '50.50.50.101', '50.50.50.102', and '50.50.50.103'. The bottom panel, 'Fabric To Device Connectivity', has a table with columns 'Type', 'Pod', 'Node', and 'Path'. It lists three entries: 'Virtual Port Channel' for Pod 1 (Node 101,102, Path vPC-L101-L102-Port16), Pod 1 (Node 103,104, Path vPC-L103-L104-Port16), and Pod 2 (Node 201,202, Path vPC-L201-L202-Port16). Both panels have 'Add' buttons and 'PBR Destinations' sections with IP address lists.

- a) Click **Add Fabric To Device Connectivity** (physical domain) or **Add VM Information** (VMM domain).

Depending on whether you selected physical or VMM domain in the previous step, you will specify information for either the firewall VM or the physical fabric connectivity to the firewall device.

For physical domains, provide the Pod, switch node, and the interface information.

For VMM domains, provide the VM name and vNIC information.

- b) Click **Add PBR Destination** to provide the IP address of the interface on the firewall that is connected to the service bridge domain.

Depending on the kind of device cluster you are deploying, you may need to provide one or more PBR destination IP addresses:

Note

This does not provision the IP address on the firewall's interface, but simply configures redirection of traffic toward that IP address. The specific firewall configuration is not deployed from NDO, and you must provision it separately.

- c) Click **Ok** to save the provided configuration.
d) Repeat this step for the other site with which you associated the template.

Step 8

Save and deploy the template.

- a) At the **Service Device Template** level, click **Save** to save the template configuration.
b) Choose the **Template Properties** tab and click **Deploy Template** to push the configuration to the sites.
c) (Optional) Verify that the configuration was created at the site level.

You can verify that the L4-L7 device is configured in the APIC by navigating to **<tenant-name> > Services > L4-L7 > Devices > <cluster-name>** in the APIC GUI. This shows the device cluster along with all the configuration you have provided in the previous steps.

To verify that the PBR policy is now configured on the APIC, navigate to **<tenant-name> > Policies > Protocol > L4-L7 Policy-Based Redirect** and you should see the **<cluster-name>-one-arm** redirect defined with the IP SLA monitoring policy you chose in *Step 8i* and the IP address you provided in *Step 7d*.

What to do next

After you have deployed the service device configuration, create the application template and a contract with which you will associate the service chaining as described in [Create Application Template, on page 394](#).

Create Application Template

The following steps describe how to create tenant template(s) and the configuration objects which you will use for the vzAny with PBR use cases.

Before you begin

- Ensure that you have read and completed the requirements described in [vzAny with PBR Guidelines and Limitations, on page 385](#).
- You must have created a VRF for which you have or will enable vzAny and which you will use for these use cases.

If you do not already have a VRF, you can create one in an Application template as you typically would. VRF configuration is described in detail in [Create Contract and Filters, on page 366](#).

Procedure

-
- Step 1** Log in to the Nexus Dashboard Orchestrator GUI.
- Step 2** From the left navigation pane, select **Configure > Tenant Templates**.
- Step 3** Choose the **Applications** tab.
- Step 4** Choose the Schema where you want to define your configuration.
- If you have an existing Schema you want to update, simply click the Schema's name in the main window pane. Otherwise, if you want to create a new Schema, click the **Add Schema** button and provide the schema information as you typically would.
- Step 5** Choose the Template where you want to define your configuration.
- If you have an existing template you want to update, choose the template in the schema view.

Note

While these steps describe how to create a single application template and stretch all objects across both sites, only the service BD (BD FW-external) must be stretched. The EPG BDs can be configured as stretched or site-local; if you choose to configure site-local BDs for the EPGs, you will need to create additional application templates for those objects and assign them to the specific sites only.

To create a new template:

- a) Click **Create Template**.
- b) In the **Select a Template Type** screen, choose `ACI Multi-Cloud`.
- c) Provide the **Display Name** for the template and **Select a Tenant**.
- d) For the **Deployment Mode**, you can choose either `Multi-Site` or `Autonomous`.

The vzAny with PBR use cases described in this chapter can be deployed for both Multi-Site and autonomous templates. If you choose to create an autonomous template, the redirection policy would apply only for intra-fabric traffic flows.

- e) Click **Continue to Template** to save the information.
- f) Choose **Actions > Add/Remove Sites** and associate the template with the sites.
- g) Repeat these substeps if you want to create additional templates for non-stretched objects.

Step 6

Create a contract.

You will associate a service device previously defined in the Service Device template to this contract to enable the PBR functionality. The contract will then be used (consumed/provided) by vzAny and by EPG/ExtEPG depending on the specific use case to provision.

- a) In the **Template Properties** view, choose **Create Object > Contract** to add a new contract.
- b) Provide the name for the Contract.

For example, `vzAny-to-vzAny`.

- c) From the **Scope** dropdown, choose `VRF`.

You must set the contract's scope to VRF.

- d) Click **+Create Filter** to add one or more contract filters.

For example, you can create a `Permit-IP` contract filter to redirect all traffic.

- e) Skip the **Service Chaining/Service Graph** configuration for now, you will associate a Service Device template to this contract in the next sections.
- f) Define the other contract options as you typically would and click **Ok** to save.

Step 7

Enable the required settings on the VRF.

- a) Select the VRF you want to use for the vzAny with PBR use case.

You can use an existing VRF or create a new one as you typically would.

- b) Enable **vzAny** and **Add Contract** that you created in the previous step.

The contract **Type** depends on the use case you want to configure:

- For any intra-VRF communication (`vzAny-to-vzAny`) use case, assign the contract to the VRF twice – once as `consumer` and again as a `provider`.
- For many-to-one communication between all the EPGs in a VRF (`vzAny`) and specific EPG as part of same VRF, assign the contract as `consumer` if you want the `vzAny` EPGs as `consumer` and specific EPG as `provider`.
- Similarly, for many-to-one communication between all the EPGs in a VRF (`vzAny`) and a specific External EPG that is part of the same VRF, assign the contract as `consumer` if you want the `vzAny` EPGs to consume a service provided by an L3Out External EPG.

- c) Enable **Site-aware Policy Enforcement Mode**

You must enable the **Site-aware Policy Enforcement Mode** setting on the VRF to enable the new vzAny PBR use cases.

Note

Enabling or disabling the **Site-aware Policy Enforcement Mode** option will cause a brief traffic disruption (including the already existing contracts between EPGs) because the zoning rules must be updated on the leaf switches. We recommend that you perform this operation during a maintenance window.

Enabling **Site-aware Policy Enforcement Mode** increases TCAM usage on the leaf switches for the existing contracts and contracting permit logging cannot be used in conjunction with this option.

- d) Enable **L3 Multicast**.

The L3 Multicast option must be enabled for the vzAny VRF to enable the conversational learning functionality described earlier in this chapter.

- e) Click **Ok** to save the changes.

Step 8

Ensure that the service BD is associated with the same VRF as you used for the vzAny contract in the previous step.

Step 9

Create application bridge domains, configured in hardware proxy mode.

Each application EPG that you will create in the next step requires a BD to be associated with it.

- a) In the **Template Properties** view, choose **Create Object > Bridge Domain**.

- b) Provide the name for the BD.

For example, `BD-App`.

- c) From the **Virtual Routing & Forwarding** dropdown, ensure to select the VRF from the previous step.

- d) Define the other BD options as you typically would.

For additional information about all available BD configurations, see [Configuring Bridge Domains, on page 72](#).

- e) Click **Ok** to save the changes.

- f) Repeat this step to create the second BD.

Following the illustration above, use `BD-Web` for the BD's name.

Step 10

Create the EPGs.

In this step, you will configure either two application EPGs or an application EPG and an External EPG depending on your specific use case.

- a) Create an Application Profile by choosing **+Create Object > Application Profile**.

- b) Choose **+Create Object > EPG** and select the application profile you created.

- c) In the properties pane, provide the **Display Name** for the EPG and choose the BD you created for this EPG.

For example, `EPG-App`. For additional information about all available BD configurations, see [Configuring Application Profiles and EPGs, on page 78](#).

- d) Define the other EPG options as you typically would.

For additional information about all available BD configurations, see [Configuring Bridge Domains, on page 72](#).

- e) Click **Ok** to save the changes.

- f) Create a second EPG.

The type of the EPG and its contract configuration depend on the use case you want to configure:

- Any intra-VRF communication (vzAny-to-vzAny).

This is the use case illustrated in [Traffic Flow: Intra-VRF vzAny-to-vzAny, on page 379](#), and you can simply create a second EPG in the same VRF. For example, create `EPG-Web` and assign the `BD-Web` bridge domain to it.

- Many-to-one communication between all the EPGs in a VRF (vzAny) and a specific EPG that part of the same VRF

In this case, create a second EPG within the same VRF but explicitly assign the contract to it as `provider` (the vzAny VRF contract for the specific EPG is assigned as `consumer`).

- Many-to-one communication between all the EPGs in a VRF (vzAny) and a specific External EPG that is part of the same VRF

In this case, you must create an External EPG instead (**+Create Object > External EPG**), associate an L3Out with the external EPG, and then explicitly assign the contract to the External EPG as `provider`.

Step 11 Click **Save Schema** to save the defined configurations.

We recommend not deploying the template until the service chaining has been configured as described in the next section to avoid undesirable communication between endpoints without firewall redirection.

At this stage, you have effectively configured a basic use case for vzAny communication between two EPGs without adding service chaining with PBR:

The screenshot displays the configuration interface for an Application Profile named 'vzAny-PBR'. It is organized into five main sections, each with a dropdown menu and a 'Create' button:

- EPGs**: Contains two input fields, 'EPG App' and 'EPG Web'. A 'Create EPG' button is in the top right.
- Contracts**: Contains one input field, 'vzAny-to-vzAny'. A 'Create Contract' button is in the top right.
- VRFs**: Contains one input field, 'VRF1'. A 'Create VRF' button is in the top right.
- Bridge Domains**: Contains three input fields, 'BD-App', 'BD-Web', and 'FW-external'. A 'Create Bridge Domain' button is in the top right.
- Filters**: Contains one input field, 'Permit-IP'. A 'Create Filter' button is in the top right.

At the top right of the page, there is a 'Create Application Profile' button with a trash icon.

The next section describes how to associate the service device you created in the previous section with the contract you created in the previous step.

What to do next

After you have created the application template and the contract, proceed to associating the service device with the contract, as described in [Add Service Chaining to Contract, on page 398](#).

Add Service Chaining to Contract

After you have created the application and the service device templates, you can add policy-based redirection by associating the contract with the service devices you created in a previous section.

Before you begin

- You must have created and deployed the service device template containing the device configuration as described in [Create Service Device Template, on page 387](#).
- You must have created (but not yet deployed) the application template containing the application bridge domains and EPGs as described in [Create Application Template, on page 394](#).

Procedure

Step 1 Navigate back to the application template that you created in the previous section.

Step 2 Select the contract you created in the previous section.

Step 3 In the **Service Chaining** area, click **+Service Chaining**.

Note

These steps assume that you have configured a brand new service device for this use case using the new Service Device template workflow introduced in release 4.2(3) as described in [Create Service Device Template, on page 387](#). If you already have a Service Graph defined in an application template, choose `Service Graph` instead and then select the existing service graph. However, keep in mind that the Service Graph option will be deprecated in a future release.

Step 4 For **Device Type**, choose `Firewall`.

This release supports one-arm firewall service graphs only.

Step 5 From the **Device** dropdown, choose the FW device cluster you created in the previous step.

Step 6 Ensure that **Consumer Connector Type Redirect** is enabled.

Step 7 Ensure that **Provider Connector Type Redirect** is enabled.

Step 8 Click **Add** to continue.

Step 9 Click **Save** to save the template.

Step 10 Click **Deploy Template** to deploy it.
