

Cisco Multi-Site Configuration Guide for ACI Fabrics, Release 3.3(x)

First Published: 2021-04-01

Last Modified: 2022-03-21

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)

© 2021 Cisco Systems, Inc. All rights reserved.



CONTENTS

CHAPTER 1 **New and Changed Information** 1

New and Changed Information 1

CHAPTER 2 **GUI Overview** 3

Overview 3

Dashboard 4

Application Management > Tenants Page 5

Application Management > Schemas Page 7

Application Management > Policies Page 7

Infrastructure > Sites Page 8

PART I **Application Management** 11

CHAPTER 3 **Tenants** 13

Tenants 13

Adding Tenants 13

CHAPTER 4 **Schemas** 17

Schema Design Considerations 17

Single Schema Deployment 17

Multiple Schemas with Network Separation 18

Multiple Schemas Based On Object Relationships 19

Concurrent Configuration Updates 21

Creating Schemas and Templates 22

Importing Schema Elements From APIC Sites 23

Configuring Application Profiles and EPGs 24

Configuring EPG's Site-Local Properties 26
Configuring VRFs 29
Configuring Bridge Domains 30
Configuring Bridge Domains Site-Local Properties 33
Configuring Contracts and Filters 34
Configuring On-Premises External Connectivity 36
Viewing Schemas 38
Deploying Templates 39
Undeploying Templates 40
Cloning Schemas 41
Cloning Templates 43
Migrating Objects Between Templates 44
Template Deployment View (Beta) 45
Schema Overview and Deployment Visualizer 47
Shadow Objects 50
Hiding Shadow Objects in APIC GUI 54
Configuration Drifts 56

PART II **Operations** 59

CHAPTER 5 **Audit Logs** 61
Audit Logs 61

CHAPTER 6 **Backup and Restore** 63
Configuration Backup and Restore 63
Backup and Restore Guidelines 63
Remote Backups 65
Configuring a Remote Location for Backups 65
Moving Existing Backups to a Remote Location 66
Adding an NFS Share to Store Backups 67
Creating Backups 68
Restoring Backups 68
Downloading Backups 69
Importing Backups 70

CHAPTER 7 **Upgrading Sites** 71

Overview 71

Guidelines and Limitations 73

Downloading Controller and Switch Node Firmware to Sites 73

Upgrading Controllers 76

Upgrading Nodes 78

CHAPTER 8 **Tech Support** 83

Tech Support and System Logs 83

Downloading System Logs 84

Streaming System Logs to External Analyzer 84

PART III **Infrastructure Management** 89

CHAPTER 9 **System Configuration** 91

System Configuration Settings 91

System Alias and Banner 91

Login Attempts and Lockout Time 92

CHAPTER 10 **Upgrading or Downgrading MSO Application** 93

Overview 93

Prerequisites and Guidelines 93

Upgrading MSO Application Using Cisco App Store 95

Upgrading MSO Application Manually 97

CHAPTER 11 **Configuring Cisco ACI Sites** 101

Pod Profile and Policy Group 101

Configuring Fabric Access Policies for All APIC Sites 102

Configuring Fabric Access Global Policies 102

Configuring Fabric Access Interface Policies 103

Configuring Sites That Contain Remote Leaf Switches 105

Remote Leaf Guidelines and Limitations 105

Configuring Routable Subnets for Remote Leaf Switches 105

Enabling Direct Communication for Remote Leaf Switches 106
Cisco Mini ACI Fabrics 106

CHAPTER 12 Adding and Deleting Sites 109
Cisco MSO and APIC Interoperability Support 109
Adding Cisco ACI Sites 111
Removing Sites 113
Cross Launch to Fabric Controllers 114

CHAPTER 13 Configuring Infra General Settings 115
Infra Configuration Dashboard 115
Configuring Infra: General Settings 117

CHAPTER 14 Configuring Infra for Cisco APIC Sites 119
Refreshing Site Connectivity Information 119
Configuring Infra: On-Premises Site Settings 120
Configuring Infra: Pod Settings 122
Configuring Infra: Spine Switches 122

CHAPTER 15 Configuring Infra for Cisco Cloud APIC Sites 125
Refreshing Cloud Site Connectivity Information 125
Configuring Infra: Cloud Site Settings 125

CHAPTER 16 Deploying Infra Configuration for ACI Sites 129
Deploying Infra Configuration 129
Enabling Connectivity Between On-Premises and Cloud Sites 130

CHAPTER 17 CloudSec Encryption 135
Cisco ACI CloudSec Encryption 135
Requirements and Guidelines 136
CloudSec Encryption Terminology 138
CloudSec Encryption and Decryption Handling 139
CloudSec Encryption Key Allocation and Distribution 140

Configuring Cisco APIC for CloudSec Encryption 142
Configuring Cisco APIC for CloudSec Encryption Using GUI 143
Configuring Cisco APIC for CloudSec Encryption Using NX-OS Style CLI 143
Configuring Cisco APIC for CloudSec Encryption Using REST API 144
Enabling CloudSec Encryption Using Multi-Site Orchestrator GUI 145
Rekey Process During Spine Switch Maintenance 146
Disabling and Re-Enabling Re-Key Process Using NX-OS Style CLI 146
Disabling and Re-Enabling Re-Key Process Using REST API 147

PART IV **Features and Use Cases** 149

CHAPTER 18 **DHCP Relay** 151
DHCP Relay Policy 151
Guidelines and Limitations 151
Creating DHCP Relay Policies 152
Creating DHCP Option Policies 154
Assigning DHCP Policies 155
Creating DHCP Relay Contract 156
Verifying DHCP Relay Policies in APIC 157
Editing or Deleting Existing DHCP Policies 157

CHAPTER 19 **Intersite L3Out** 159
Intersite L3Out Overview 159
Intersite L3Out Guidelines and Limitations 160
Configuring External TEP Pool 161
Creating or Importing Intersite L3Out and VRF 162
Configuring External EPG to Use Intersite L3Out 164
Creating a Contract for Intersite L3Out 166
Use Cases 169
Intersite L3Out for Application EPGs (Intra-VRF) 169
Shared Services with Intersite L3Out for Application EPGs (Inter-VRF) 172
Intersite Transit Routing 175

CHAPTER 20 **Intersite L3Out with PBR** 179

Intersite L3Out with PBR 179
Supported Use Cases 180
Guidelines and Limitations 184
Configuring APIC Sites 184
Configuring External TEP Pool 184
Creating and Configuring L4-L7 Devices and PBR Policies 185
Creating Templates 188
Configuring Service Graph 190
Creating Filter and Contract 192
Creating Application EPG 198
Creating VRF and Bridge Domain for Application EPG 198
Creating Application Profile and EPG 199
Creating L3Out External EPG 201
Creating or Importing Intersite L3Out and VRF 201
Configuring External EPG 203

CHAPTER 21 **Layer 3 Multicast** 205
Layer 3 Multicast 205
Layer 3 Multicast Routing 206
Rendezvous Points 206
Multicast Filtering 207
Layer 3 Multicast Guidelines and Limitations 208
Creating Multicast Route Map Policy 209
Enabling Any-Source Multicast (ASM) Multicast 210
Enabling Source-Specific Multicast (SSM) 212

CHAPTER 22 **EPG Preferred Group** 215
EPG Preferred Groups 215
Configuring EPGs for Preferred Group 216

CHAPTER 23 **QoS Preservation Across IPN** 219
QoS and Global DSCP Policy 219
DSCP Policy Guidelines and Limitations 219
Configuring Global DSCP Policy 220

Set QoS Level for EPGs and Contracts 222

CHAPTER 24 SD-WAN Integration 225

SD-WAN Integration 225

SD-WAN Integration Guidelines and Limitations 226

Adding a vManage Controller 227

Configuring Global DSCP Policy 228

Set QoS Level for EPGs and Contracts 231

CHAPTER 25 Sites Connected via SR-MPLS 233

SR-MPLS and Multi-Site 233

Infra Configuration 235

SR-MPLS Infra Guidelines and Limitations 235

Creating SR-MPLS QoS Policy 237

Creating SR-MPLS Infra L3Out 239

SR-MPLS Tenant Requirements and Guidelines 242

Creating SR-MPLS Route Map Policy 244

Enabling Template for SR-MPLS 246

Creating VRF and SR-MPLS L3Out 246

Configuring Site-Local VRF Settings 247

Configuring Site-Local SR-MPLS L3Out Settings 247

Communication Between EPGs Separated by MPLS Network 248

Deploying Configuration 249

CHAPTER 26 vzAny Contracts 251

vzAny and Multi-Site 251

vzAny and Multi-Site Guidelines and Limitations 252

Create Contract and Filters 254

Configure vzAny to Consume/Provide a Contract 255

Create EPGs to Be Part of the vzAny VRF 255

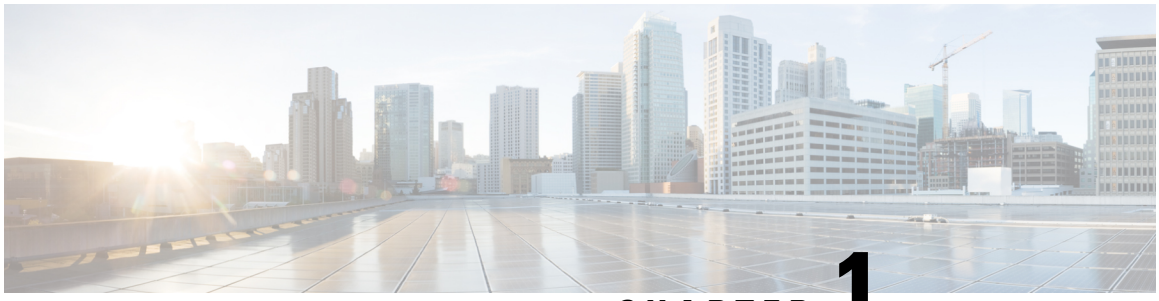
Free Intra-VRF Communication 256

Stretched EPGs 258

Site-Local EPGs 258

Combination of Site-Local and Stretched EPGs 259

Intra-VRF Intersite L3Out 260
Many-to-One Communication 261
Provider EPG Within vzAny VRF 263
Provider EPG In Its Own VRF 263



CHAPTER 1

New and Changed Information

- New and Changed Information, on page 1

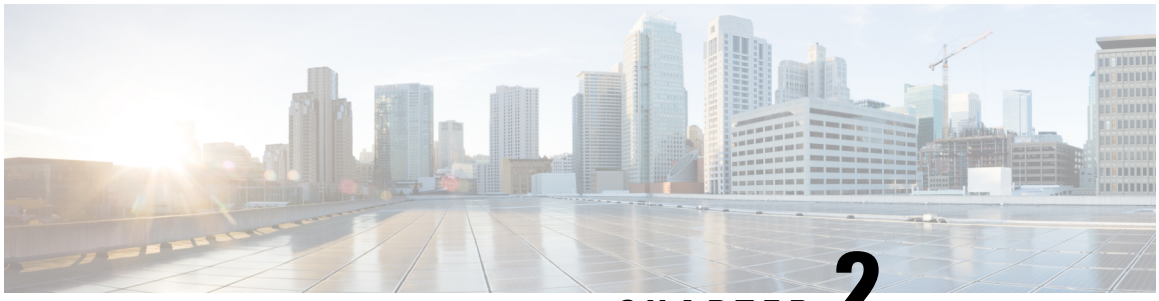
New and Changed Information

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
3.3(1)	First release of this document.	--

New and Changed Information



CHAPTER 2

GUI Overview

- Overview, on page 3
- Dashboard, on page 4
- Application Management > Tenants Page, on page 5
- Application Management > Schemas Page, on page 7
- Application Management > Policies Page, on page 7
- Infrastructure > Sites Page, on page 8

Overview

The Multi-Site Orchestrator (MSO) GUI is a browser-based graphical interface for configuring and monitoring your Cisco APIC, Cloud APIC, and DCNM deployments.

The GUI is arranged according to function. For example, the **Dashboard** page contains an overview of your fabrics and their health, the **Sites** page provides information on each site and allows you to add sites, the **Schemas** page allows you to create and configure schemas, and so on. The functionality of each MSO GUI page is described in the following sections.

The top of each page shows the controller status indicating how many controllers are operational, the **Get Started** menu icon, the **Settings** icon, and the **User** icon.

The **Get Started** menu provides easy access to a number of common tasks you may want to perform, such as adding sites or schemas, configuring specific policies, or performing administrative tasks.

The **Settings** icon allows you to access overview information about your Multi-Site Orchestrator, such as the currently running version, what's new in the current release, API documentation, and system status:

- The **About MSO** link displays information about the version of the Multi-Site Orchestrator currently installed.
- The **What's New in This Release** link displays a short summary of the new features in your release, as well as links to the rest of the Multi-Site Orchestrator documentation.
- The **API Docs** link gives you access to the set of Swagger API object and method references. Using the Swagger API is described in more detail in the *Cisco Multi-Site REST API Configuration Guide*.
- The **System Status** link provides you with the status and health of all running services that are used by the MSO.

The **User** icon allows you to view information about the currently logged in user, preferences and bookmarks. It also allows you to log out of the Orchestrator GUI.



Note Starting with Release 3.2(1), user management has moved to the common user and authentication management in the Nexus Dashboard where your MSO application is running.

- The **Preferences** link allows you to change a few GUI options.
- The **Bookmarks** link opens the list of all the bookmarked schemas you save while using the Orchestrator. You can bookmark a schema by clicking the bookmark icon in the top right corner of the screen while viewing or editing the schema.

When working with fabric objects, a **Display Name** field is used throughout the Orchestrator's GUI whenever the objects are shown. You can specify a display name when creating the objects, however due to object naming requirements on the site controllers, any invalid characters are removed and the resulting **Internal Name** is used when pushing the objects to sites. The **Internal Name** that will be used when creating the tenant is typically displayed below the **Display Name** text box.

Dashboard

The Multi-Site Orchestrator dashboard displays the list of all of your site implementations in addition to their current functionality and health.

The **Dashboard** has the following functional areas:

- **Site Status**: The site status table lists your sites according to name and location. The table also indicates the current health status for your implementation according to a descriptive color code.
  - The Controller State column indicates the number of controllers available and running. You can have a maximum number of 3 controllers in your Multi-Site implementation. For example, if one out of the 3 controller is down it is represented as 2/3.
  - The Connectivity column provides an operational status of the BGP sessions and the dataplane unicast and multicast tunnels that are connected to the peer sites for each site in the dashboard.

    When one or more BGP sessions or tunnels fail to establish, Multi-Site provides the information about which exact local spines and remote spines failed to establish the BGP session or the tunnel. Multi-Site should be enabled in the site in the infrastructure configuration, for the BGP sessions and the dataplane unicast and multicast tunnels to be established to the peer sites.

    BGP Sessions

    - When the BGP peering type is full-mesh in **Infra**-> **General Settings**, the spine node in a site with the BGP peering enabled will establish the BGP sessions to all the spine nodes with the BGP peering enabled in all the peer sites.
    - When the BGP peering type is route-reflector in **Infra**-> **General Settings**, the spine node in a site with both BGP peering enabled and route-reflector enabled, will establish the BGP sessions to all the spine nodes with the BGP peering enabled in all the peer sites. In the route-reflector mode, at least the local spine node or the remote spine node or both should have the route-reflector enabled. Otherwise, the BGP session is not established between them.

- If the local and the remote ASNs are different, then it is eBGP. Therefore, the sessions between those sites are always full-mesh, irrespective of the BGP peering type and the route-reflector configuration.

Unicast and Multicast Tunnels: A spine node in a site that is connected to ISN and has infrastructure configuration, will establish a tunnel to all the spine nodes that are connected to ISN in the peer sites.

The color codes indicate the following conditions:

- **Critical** (red)
- **Major** (orange)
- **Minor** (yellow)
- **Warning** (green)

The numbers in the color indicator columns indicate the number of faults per site.

- **Schema Health**: provides a listing of your schemas with locales and health.
  - You can click the magnifying glass icon and enter a schema name to search for a subject schema.
  - You can click the + sign to start adding a new schema to your site.
  - You can click the site locale in the **Schema Health** table to view the schema details and status for a template.

    The **Schema Health** table provides a heat map type of display; that is, the health of the subject schema is displayed according to color. Schemas that span two columns (i.e, locales) indicate a stretched condition.
    - Click the color highlighted table cell to further discover what policies are incorporated into the subject schema. On the schema details page, you can click the arrow to go into the schema builder and update the policy details in the subject schema.
    - The color coded slider enables you to select a range for identifying schemas whose health require further review. For example, you can adjust the slider value to between 80 and 100. Then all of your schema implementations that fall within that specific range are displayed on the accompanying Schema Health table.

Application Management > Tenants Page

The Multi-Site **Tenants** page lists all of the tenants that comprise your implementation.

The table on the **Tenants** page displays the following:

- **Tenant Name**
- **Assigned to Sites**
- **Assigned to Users**
- **Assigned to Schemas**

- **Actions**

The features and functionality on this page include the following:

- **Name**: click a tenant name to access the **Tenant Details** settings page. On the **Tenant Details** page you can edit or update the following sections:
  - **General Settings**: change the Display Name and Description as required.
  - **Associated Sites**: view the sites associated with the subject tenant.
  - **Associated Users**: view the users associated with the subject tenant - you can associate a user with the subject tenant by checking the empty box next to the user name.
- **Associated Schemas**: click the **Associated Schema** listing to view the schemas associated with the subject tenant.
- **Actions**: click the **Actions** listing to edit the subject tenant's details sites or to create a new network mapping.



Note You can delete the Tenant object by selecting **Delete** on the **Actions** drop down menu.

- **Add Tenant:** click **Add Tenant** button to add an existing tenant to your implementation. On the proceeding Tenant Details page, you can add the tenant name, description, security domain, and associated users.

Audit Logs

Click the **Audit Log** icon next to the **Add Schema** tab to list the log details for the Schemas page. The **Audit Logs: Tenant List** page is displayed.

The table on the page displays the following details:

- **Date**
- **Action**
- **Details**
- **User**

Click the **Most Recent** tab to select the audit logs during a particular time period. For example, when you select the range from November 10, 2019 to February 14, 2020 and click **Apply**, the audit log details for this time period are displayed on the **Audit Logs** page.

Click the **Filter** icon next to the **Most Recent** tab to filter the log details using the following criteria:

- **User**: Select one username or all users and click **Apply** to filter the log details using the username.
- **Action**: Select the action, for example, created, updated, or deleted, and click **Apply** to filter the log details according to the action.

For more information, see the Tenants, on page 13 chapter.

Application Management > Schemas Page

The **Schemas** page lists all schemas that are associated with your deployment.

Use the magnifying glass and associated field to search for a specific schema. Use schemas to configure or import tenant policies, including the VRF, application profile with EPGs, filters and contracts, bridge domains, and external EPGs.

The Schemas table shows the following information:

- **Name**: click the schema name to view or update the settings for the subject schema.
- **Templates**: displays the name of the template that is used for the schema. Templates are analogous to profiles in the ACI context, which group policies. You can create templates for stretched objects or site-specific objects.
- **Tenants**: displays the name of the tenant that is used for the subject schema.
- **Actions**: click the **Action** field with the associated schema to either edit or delete the subject schema.

You can use the **Add Schema** button to add a new schema, which is described in more details in later sections of this document.

Audit Logs

Click the **Audit Log** icon next to the **Add Schema** tab to list the log details for the Schemas page. The **Audit Logs: Schemas List** page is displayed.

The table on the page displays the following details:

- **Date**
- **Action**
- **Details**
- **User**

Click the **Most Recent** tab to select the audit logs during a particular time period. For example, when you select the range from November 10, 2019 to February 14, 2020 and click **Apply**, the audit log details for this time period are displayed on the **Audit Logs** page.

Click the **Filter** icon next to the **Most Recent** tab to filter the log details using the following criteria:

- **User**: Select one username or all users and click **Apply** to filter the log details using the username.
- **Action**: Select the action, for example, created, updated, or deleted, and click **Apply** to filter the log details according to the action.

Application Management > Policies Page

The Multi-Site Orchestrator **Policies** page displays all policies you have configured for your fabrics.

The **Policies** page contains a table of all policies along with the summary of their types, tenants they're associated with, descriptions, and usage. You can use this page to add new policies or edit existing ones.

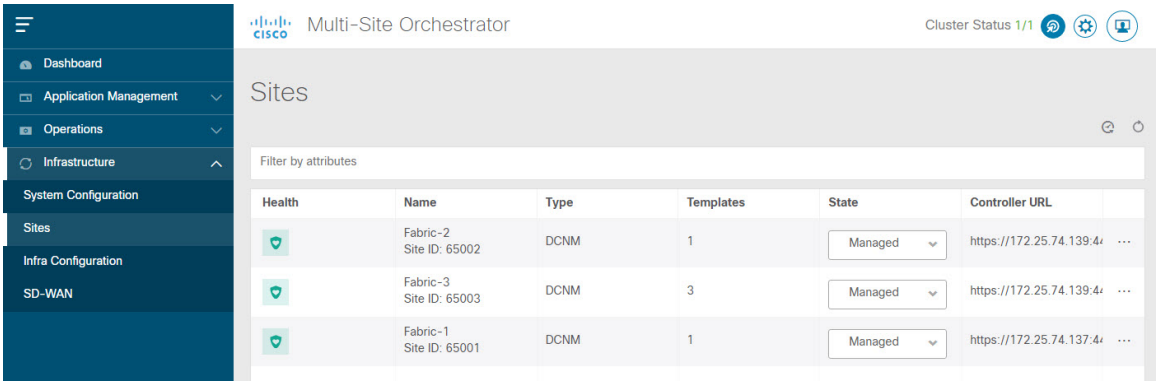
You can configure the following policies:

- DHCP Policy, as described in the DHCP Relay, on page 151 chapter
- MPLS QoS Policy, as described in the Sites Connected via SR-MPLS, on page 233 chapter.
- Route Map Policy, as described in the Sites Connected via SR-MPLS, on page 233 chapter.
- Multicast Route Map Policy, as described in the Layer 3 Multicast, on page 205 chapter.

Infrastructure > Sites Page

The MSO **Infrastructure** > **Sites** page displays all of the sites in your implementation, for example:

Figure 1: Multi-Site Sites Page



The **Sites** page includes the following:

- Site **Health** indicates the status of the site's overall health according to the following color coded identifiers:
  - **Critical** (red)
  - **Major** (orange)
  - **Minor** (yellow)
  - **Warning** (green)
- Site **Name** shows the display name of the site as you defined it when adding the site.
- Site **Type** displays the fabric type, for example `ACI` or `DCNM`.
- The **Templates** column indicates the number of templates associated with the site.
- The **State** column indicates whether or not this particular fabric is managed by MSO.

  You add and manage sites and their properties in the Nexus Dashboard GUI. The MSO **Sites** page displays all the sites available in the Nexus Dashboard GUI and allows you to define which specific sites you want to be managed by the MSO.
- The **Controller URL** column displays the in-band IP address of the site's controller.

- The actions menu (**...**) allows you to import the site's tenants (ACI fabrics only) or open the site's controller UI.

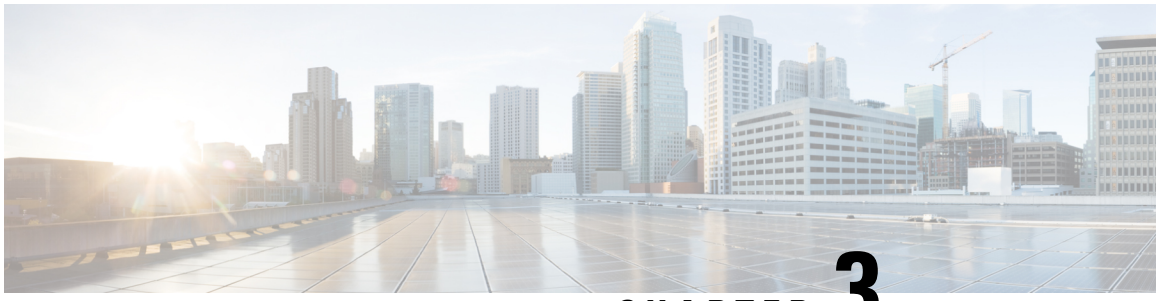
If you click a specific site, a right **Properties** sidebar opens and you can view additional information about the site.



PART I

Application Management

- Tenants, on page 13
- Schemas, on page 17



CHAPTER 3

Tenants

- Tenants, on page 13
- Adding Tenants, on page 13

Tenants

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.

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

Three 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 DCNM 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 Multi-Site Orchestrator to manage Cisco DCNM fabrics, you will use the default `dcnm-default-tn` that is preconfigured for you and allows you to create and manage the following objects:

- VRFs
- Networks

Adding Tenants

This section describes how to add tenants using the Multi-Site Orchestrator GUI.

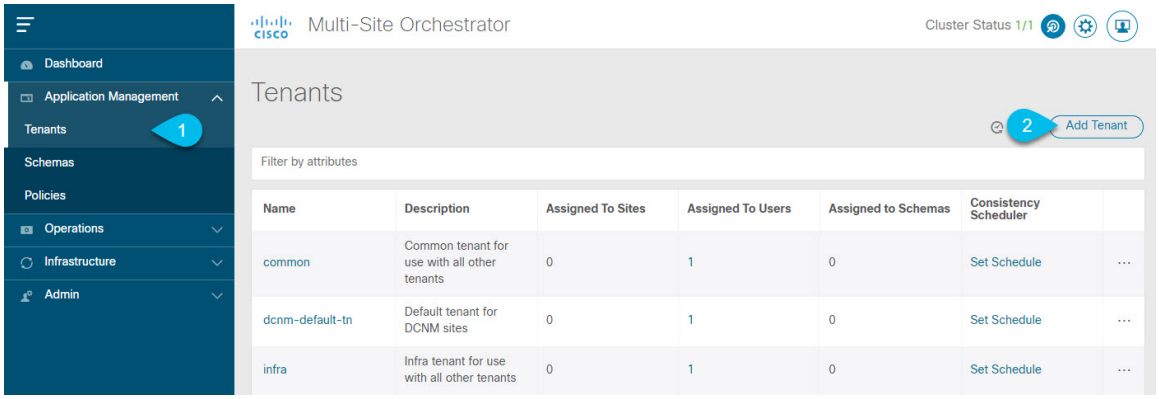
Before you begin

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 the Multi-Site Orchestrator GUI.

Step 2 Add a tenant.



a) In the left navigation menu, select **Application Management** > **Tenants**.

b) In the top right of the main pane, click **Add Tenant**.

The **Add 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** textbox.

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 you want to associate with this tenant and the **Security Domain** to use.

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

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*.

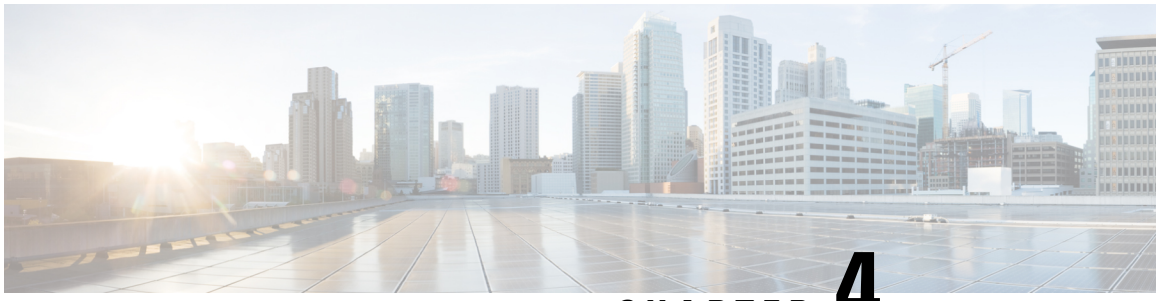
c) In the **Associated Users** section, select the Multi-Site Orchestrator users that are allowed to access the tenant.

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

d) (Optional) Enable consistency checker scheduler.

You can choose to enable regular consistency checks. For more information about the consistency checker feature, see *Cisco Multi-Site Troubleshooting Guide*.

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



CHAPTER 4

Schemas

- Schema Design Considerations, on page 17
- Concurrent Configuration Updates, on page 21
- Creating Schemas and Templates, on page 22
- Deploying Templates, on page 39
- Cloning Schemas, on page 41
- Cloning Templates, on page 43
- Migrating Objects Between Templates, on page 44
- Template Deployment View (Beta), on page 45
- Schema Overview and Deployment Visualizer, on page 47
- Shadow Objects, on page 50
- Configuration Drifts, on page 56

Schema Design Considerations

A schema is a collection of templates, which are used for defining policies, with each template assigned to a specific tenant. There are multiple approaches you can take when it comes to creating schema and template 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 environment.

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 *Cisco Multi-Site Verified Scalability Guides* for your release.

Single Schema 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 adds all VRFs, Bridge Domains, EPGs, Contracts and other elements to that template. You can then create a single application profile or multiple application profiles within the template and deploy it to one or more sites.

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

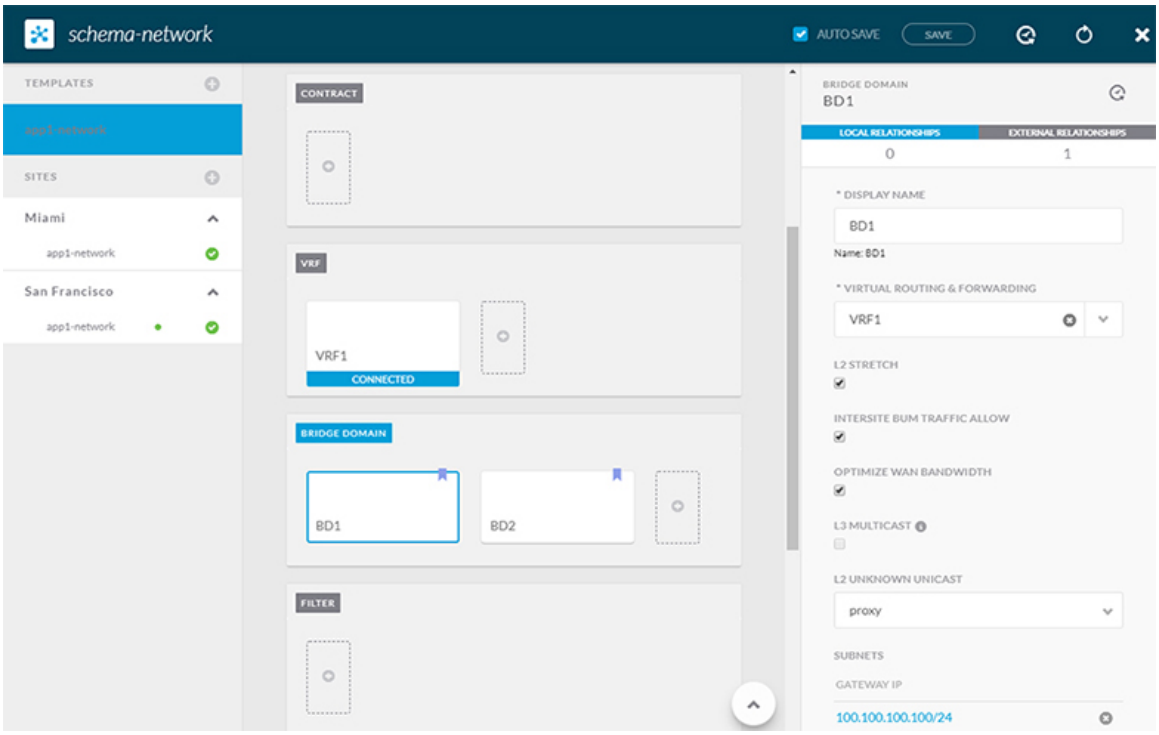
Multiple Schemas with Network Separation

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.

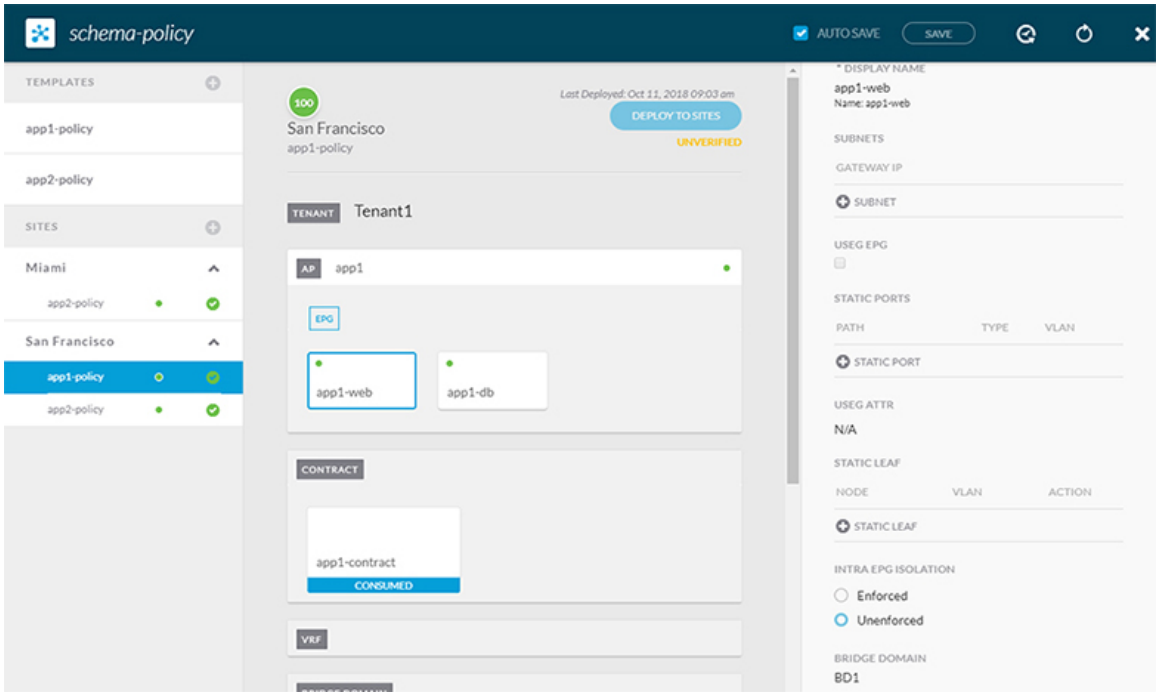
The following figure shows a single networking template configuration with VRF, BD, and subnets configured and deployed to two sites:

Figure 2: Network Schema



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. The following figure shows a policy schema that contains two application templates both of which reference the networking elements in an external schema. One of the applications is local to one site while the other is stretched across two sites:

Figure 3: Policy Schema



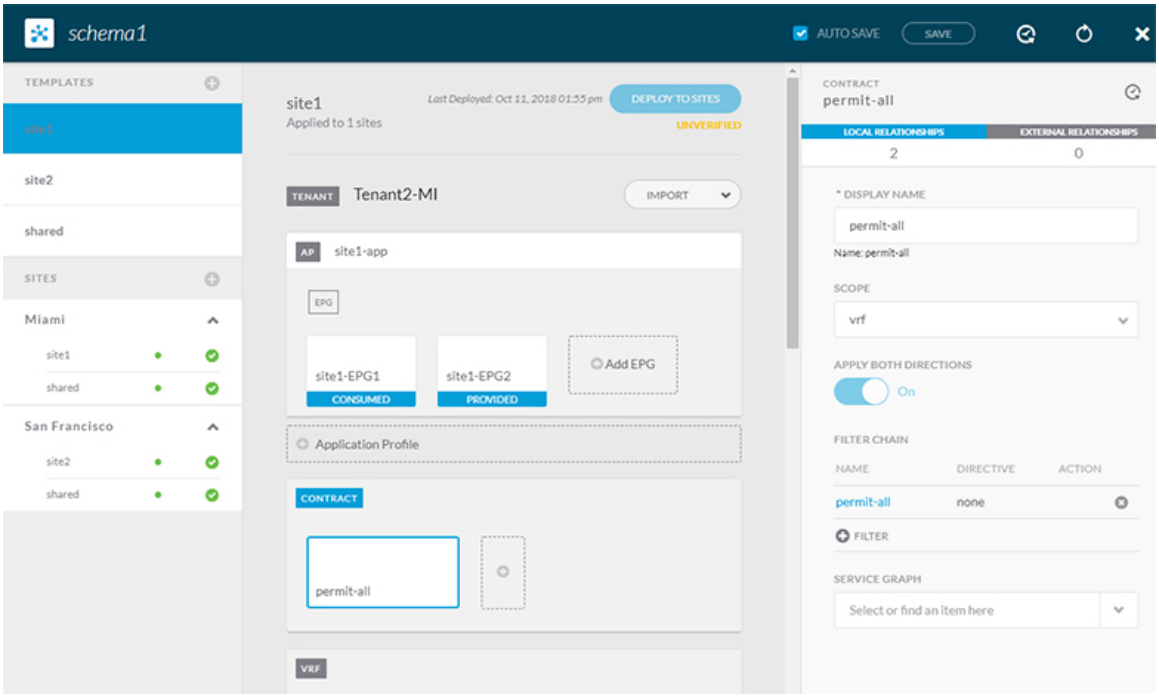
After creating and deploying the policy schemas and templates, the networking objects in the networking schema will display the number of external references by the policy schema elements. The object with external references will also be denoted by the ribbon icon as shown in the "Network Schema" figure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

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 Schemas 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.

Figure 4: One Template per Site



The **site1** template in the above figure contains only the objects that are local to Site1 and the template is deployed to only the Miami site. Similarly, the **site2** template contains only the object relevant to site2 and is deployed to the San Francisco site. Any change made to any object in either of these templates has no effect on the other one. The **shared** template contains any 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.

Concurrent Configuration Updates

The Multi-Site 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 schema 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:



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 MSO 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, MSO 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"
}
```

Creating Schemas and Templates

Before you begin

- You must have an administrative user account with full read/write privileges.
- You must have a user account with read/write tenant policy privileges.

For more information, see the *User Access, Authentication, and Accounting* chapter in the *Cisco APIC Basic Configuration Guide*.

- You must have at least one available tenant that you want to incorporate into your site.

  For more information, refer to Adding Tenants, on page 13.

Procedure

Step 1 Log in to your Nexus Dashboard and open the Multi-Site Orchestrator application.

Step 2 Create a new schema.

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

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

Step 3 Create a template.

a) In the left sidebar under **Templates**, click the + sign to add a new template.
b) In the right sidebar, specify the **Display Name**.
c) (Optional) Provide a **Description**.
d) If you are configuring an SR-MPLS template, enable the **SR-MPLS** knob.

   For detailed information on SR-MPLS templates, see Sites Connected via SR-MPLS, on page 233.

e) From the **Select a Tenant** dropdown, 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 menu. Associating a user account with a tenant is described in Adding Tenants, on page 13.

Step 4 Assign the templates to sites.

You deploy one template at a time, so you need to associate the template with at least one site where you want to deploy the configuration.

a) In the left pane, click the + icon next to Sites
b) In the **Add Sites** window, check the checkbox next to the sites where you want to deploy the template.
c) From the **Assign to Template** dropdown next to each site, select one or more templates.

   While you deploy one template at a time to every site with which it is associated, you can associate multiple templates to a site at once.

d) Click **Save**.

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 MSO, 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 MSO and a new object is imported with the **Include Relations** option enabled, MSO 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.

Beginning with Release 3.1(1), 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 MSO. 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 MSO 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 MSO.

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 created and a tenant assigned to the template, as described in Creating Schemas and Templates, on page 22.

This section also assume 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, mouse over the tile, and click **Add Application Profile**.

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

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, mouse over the **EPGs** tile, and click **Add EPG**.

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

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

Step 4 Add a contract for the EPG.

Creating contracts and filters is described in detail in Configuring Contracts and Filters, on page 34. If you already have a contract created:

a) Click **+ Contract**.

b) On the **Add Contract** dialog, enter the contract name and type.

c) Click **SAVE**.

Step 5 From the **Bridge Domain** dropdown, select the bridge domain for this EPG.

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

Step 6 (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 7 (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** checkbox.
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 8 (Optional) Enable intra-EPG isolation.

By default, endpoints in EPG can freely communicate with each other. If you would like 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 9 (Optional) Enable Layer 3 multicast for the EPG.

For additional information about Layer 3 multicast, see Layer 3 Multicast, on page 205

Step 10 (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, on page 215

Step 11 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 Bridge Domains Site-Local Properties, on page 33

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 24.
- 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 EPG.

Step 2 In the left sidebar, select the template that contains the EPG under the specific site that you want to configure.

Step 3 In the main pane, select the bridge domain.

For most fields, you will see the values you have configured at the template level.

Step 4 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 you want to add.

c) 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.

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 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**.

When using Cisco Multi-Site with this APIC fabric domain (site), indicates that the VRF, EPG, or BD using this subnet are mirrored from another site, which has a relationship to this site through a contract. Do not modify or delete the mirrored objects

f) Click **Save**.

Step 5 Add one or more **Static ports**.

a) Click **+Static Port**.

b) From the **Path Type** dropdown, select the type of port.

c) If configuring a physical interface, select the **Pod** and **Leaf** for the interface.

d) Provide the **Path** for the interface.

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.

For micro-segmentation, when a **Primary MICRO-SEG VLAN** is configured, the **Port Encap VLAN** is configured as an Isolated Secondary VLAN for the Primary VLAN. In this configuration, traffic is forwarded to the leaf that the

host is connected to. Traffic is sent from the host to the leaf using the secondary VLAN and return traffic from the leaf 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**.

Once 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 as soon as the policy is downloaded in the leaf 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 refers to 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 native 802.1p mode, its packets exit that port untagged. When an access port is configured with multiple EPGs, one in native 802.1p mode, and some with VLAN tags, all packets exiting that access port are tagged VLAN 0 for EPG configured in native 802.1p mode and for all other EPGs packets exit with their respective VLAN tags. Note that only one native 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 will 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 6 Add one or more **Static Leaf** nodes.

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

Step 7 Add one or more **Domains** nodes.

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

This is the type of the domain 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 as soon as the policy is downloaded in the leaf 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 pre-provisions the configuration on the switch.

Configuring VRFs

This section describes how to configure 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 22.

Procedure

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

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

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

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

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

Step 2 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 MSO, 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 205.

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

For additional information, see vzAny Contracts, on page 251.

Configuring Bridge Domains

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

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 22.
- You must have the VRF created as described in Configuring VRFs, on page 29

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, mouse over the tile, and click **Add Bridge Domain**.

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

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

Step 3 Configure **On-Premises Properties**.

a) From the **Virtual Routing & Forwarding** dropdown, 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**.

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 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 205.

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

By default, unicast traffic is flooded to all Layer 2 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 as well as the remote leaf switches where the BD is deployed, by selecting Clear Remote MAC Entries. Without this feature, the remote leaf 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 attached with the EPGs associated with this bridge domain. Flooding is not restricted to only M-Router ports of the bridge domain.
- `Optimize 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 attached with the EPGs associated with this bridge domain. Flooding is not restricted to only M-Router ports of the bridge domain.
- `Optimize 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.

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 virutal 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 used only for migration of individual sites to MSO-orchestrated multi-site fabric. Once the migration is completed, these flags can be disabled.

Step 4 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 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 MSO-orchestrated Multi-Site deployments.

d) 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.

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

When using Cisco Multi-Site with this APIC fabric domain (site), indicates that the VRF, EPG, or BD using this subnet are mirrored from another site, which has a relationship to this site through a contract. Do not modify or delete the mirrored objects

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

Enables IGMP Snooping on the subnet

h) Click **Save**.

Step 5 (Optional) Enable **DHCP Policy**.

For additional information, see DHCP Relay, on page 151.

Step 6 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 Domains Site-Local Properties, on page 33

Configuring Bridge Domains 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 30.
- 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 will see the values you have configured at the template level.

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 159 use case.

Step 5 Enable **Host Route**.

This enabled Host Based Routing on the bridge domain, individual host-routes (`/32` prefixes) are advertised from the border leaf switches. Border leaf switches along with the subnet advertise the individual end-point (EP) prefixes. The 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 once the EP is removed from EP database, the route advertise 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

Configuring Contracts and Filters

This section describes how to configure a contract, a filter, and assign the filter to the contract. A filter is similar to an Access Control List (ACL), it is used to filter traffic through contracts 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 will apply it. If the objects that will use the contract are deployed to different sites, we recommend defining the contract in a template associated to multiple sites. However, this is not strictly required and even if the contract and filters are defined only as local objects in Site1, MSO will create those objects in a remote Site2 when a local EPG or external EPG in Site2 needs to consume or provide that contract.

Step 2 Create a filter.

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

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

b) In the right pane, provide the **Display Name** for the filter.
c) (Optional) Provide a **Description**.
d) In the right pane, click **+ 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.

Step 3 Create a filter entry.

a) In the right pane, click **+ 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.

b) Provide the **Name** for the filter.
c) 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 **Save** 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 4 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) 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`

e) 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 will need to provide the filters only once and they will apply for traffic in both directions. If you leave this option disabled, you will need to 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 re-deploy 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 disabled to correctly change the setting in your fabrics.

f) (Optional) From the **Service Graph** dropdown, select a service graph for this contract.
g) (Optional) From the **QoS Level** dropdown, 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 219.

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

Step 5 Assign the filters to the contract

a) In the right pane, scroll down to the **Filter Chain** area and click **+ Filter** to add a filter to the contract.
b) In the **Add Filter Chain** window that opens, select the filter you added in previous step from the **Name** dropdown menu.
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.

d) Click **Save** 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 additional 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.

Configuring On-Premises External Connectivity

Cisco ACI allows you to establish connectivity to the networks outside your on-premises ACI fabric through the border leaf switches. This connectivity is defined using two constructs, L3Out and External EPG, which provide the configuration options necessary to define security and route maps.

This section describes how to create an L3Out and external EPG in the Multi-Site Orchestrator GUI. The Orchestrator then creates the objects on the APIC site where you deploy the template. Keep in mind that when creating an L3Out from the Orchestrator, only the L3Out container object is created in the APIC and you must still perform the full L3Out configuration (such as nodes, interfaces, routing protocols, and so on) directly in the site's APIC.

While in most cases the L3Out will be created directly at the APIC level and then associated to an external EPG that you create in the Orchestrator, it may be useful to create both here in order to directly associate the L3Out to a VRF which you have created from the Orchestrator.

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 22.
- You must have the VRF for the L3Out created as described in Configuring VRFs, on page 29

Procedure

Step 1 Navigate to the Schema and Template you want to edit.

Step 2 Create an L3Out.

a) In the schema edit view, scroll down to the **L3Out** area and click + to add a new L3Out.
b) In the properties pane on the right, provide a display name for the L3Out.
c) From the **Virtual Routing & Forwarding** dropdown, select the VRF you created for this.

Step 3 Create an external EPG.

a) In the schema edit view, scroll down to the **External EPG** area and click + to add a new External EPG.
b) In the properties pane on the right, select **On-Prem** for the site type.

Step 4 Configure external EPG's basic properties.

a) In the right properties sidebar, provide a display name for the External EPG.
b) From the **Virtual Routing & Forwarding** dropdown, select the VRF you created.

This must be the same VRF that you associated with the L3Out.

c) Click **+Contract** to add a contract for the external EPG to communicate with other EPGs.

If you already have the contracts created, you can assign them now. Otherwise, you can come back to this screen to assign any Contracts you plan to create later.

When assigning Contracts:

- If you are associating a contract with the external EPG as provider, choose contracts only from the tenant associated with the external EPG. Do not choose contracts from other tenants.
- If you are associating the contract to the external EPG as consumer, you can choose any available contract.

Step 5 If configuring an external EPG for an on-premises fabric, set **Site Type** to `on-prem` and configure external EPG's on-premises properties.

a) From the **L3Out** dropdown, select the L3Out you created in a previous step.

Note
You can select the L3Out at the template or site-local level. We recommend configuring the L3Out for the external EPG at the site-local level. To do that, select the template in the left sidebar under the site to which it is assigned. Then select the external EPG and associate an L3Out with it.

b) Click **+Subnet** to add a classification subnet.
c) In the **Add Subnet** window, provide the subnet prefix.
d) Select the required **Route Control** options.

You can choose one or more for the following options:

- **Export Route Control** enables a route map that allows external prefixes to get advertised out of the L3Out. These are the prefixes learned from other L3Outs for transit routing use cases.

  If you are adding the `0.0.0.0/0` subnet and enable the export route control option, **Aggregate Export** option becomes available. This allows you to advertise all the external prefixes learned from other L3Outs. If you choose to leave this option disabled, only the default route learned from other L3Out will be advertised out of this L3Out.

- **Import Route Control** configures an ingress route map to give you control over what prefixes you want to import from your L3Out into the fabric. The **Import Route Control** is available only when using BGP as the routing protocol on the L3Out.

  If you are adding the `0.0.0.0/0` subnet and enable the import route control option, **Aggregate Import** option becomes available. This works similar to the export route control case except for ingress routes.

- **Shared Route Control** is used for shared L3Out use case and allows prefixes learned from the external router to be advertised to the other VRF that will use this L3Out.

  If you enable the shared route control option, **Aggregate Shared Routes** option becomes available. Again, this functions similar to previous two aggregate routes options but is available for non-`0.0.0.0/0` subnets.

e) Select the **External EPG Classification** options.

You should check the **External Subnets for External EPG** option for the configured subnet(s) to be able to map external entities to this specific External EPG. This allows you to apply a security policy (contract) between those external networks and endpoints belonging to EPG(s) defined inside the fabric. Enabling this flag for a `0.0.0.0/0` prefix ensures that all the external destinations are considered part of this External EPG.

If you enable the this option, **Shared Security Import** option becomes available, which allows access from the subnet to the endpoints within the fabric.

For both of these options, access is still subject to contract rules.

Step 6 If configuring an external EPG for a cloud fabric, set **Site Type** to `cloud` and configure external EPG's cloud properties.

a) From the **Application Profile** dropdown, select the application profile.
b) Click **+Add Selector** to add a cloud endpoint selector for the EPG.

Step 7 (Optional) If you want to include this external EPG in the preferred group, check the **Include in preferred group** checkbox.

For more info about EPG Preferred Group, see EPG Preferred Groups, on page 215.

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 Multi-Site Orchestrator GUI.

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, 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 Assign the templates to sites.

If you have already assigned the templates to sites, skip this step.

You deploy one template at a time, so you need to associate the template with at least one site where you want to deploy the configuration.

a) In the left pane, click the + icon next to Sites
b) In the **Add Sites** window, check the checkbox next to the sites where you want to deploy the template.
c) From the **Assign to Template** dropdown next to each site, select one or more templates.

While you deploy one template at a time to every site with which it is associated, you can associate multiple templates to a site at once.

d) Click **Save**.

Step 3 In the left sidebar, select the template you want to deploy.

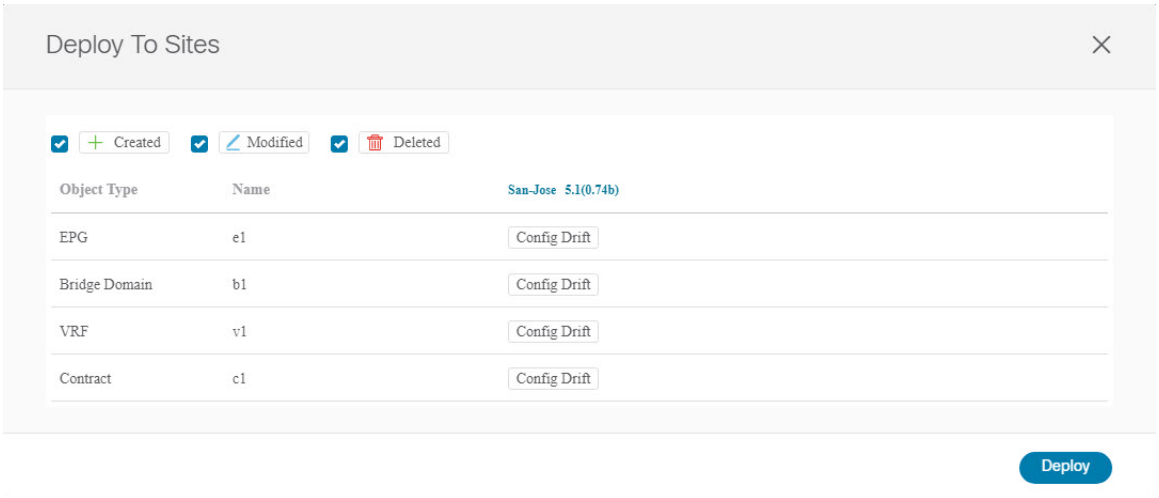
Step 4 In the top right of the template edit view, click **Deploy to sites**.

The **Deploy to Sites** window opens that shows the summary of the objects to be deployed.

Step 5 Click **Deploy** to deploy the template.

- If this is the first time you are deploying this template or you have made changes to a previously deployed template, the **Deploy to Sites** summary will show the configuration difference that will be deployed to sites.

  Note
  You can 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**.



- If you have previously deployed this template but made no changes to it since, the **Deploy to Sites** summary will allow you to re-deploy the entire template only the `Full Template` option and you can simply click **Deploy** to re-deploy the entire template.

Undeploying Templates

This section describes how to undeploy a template from a site.

Before you begin

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

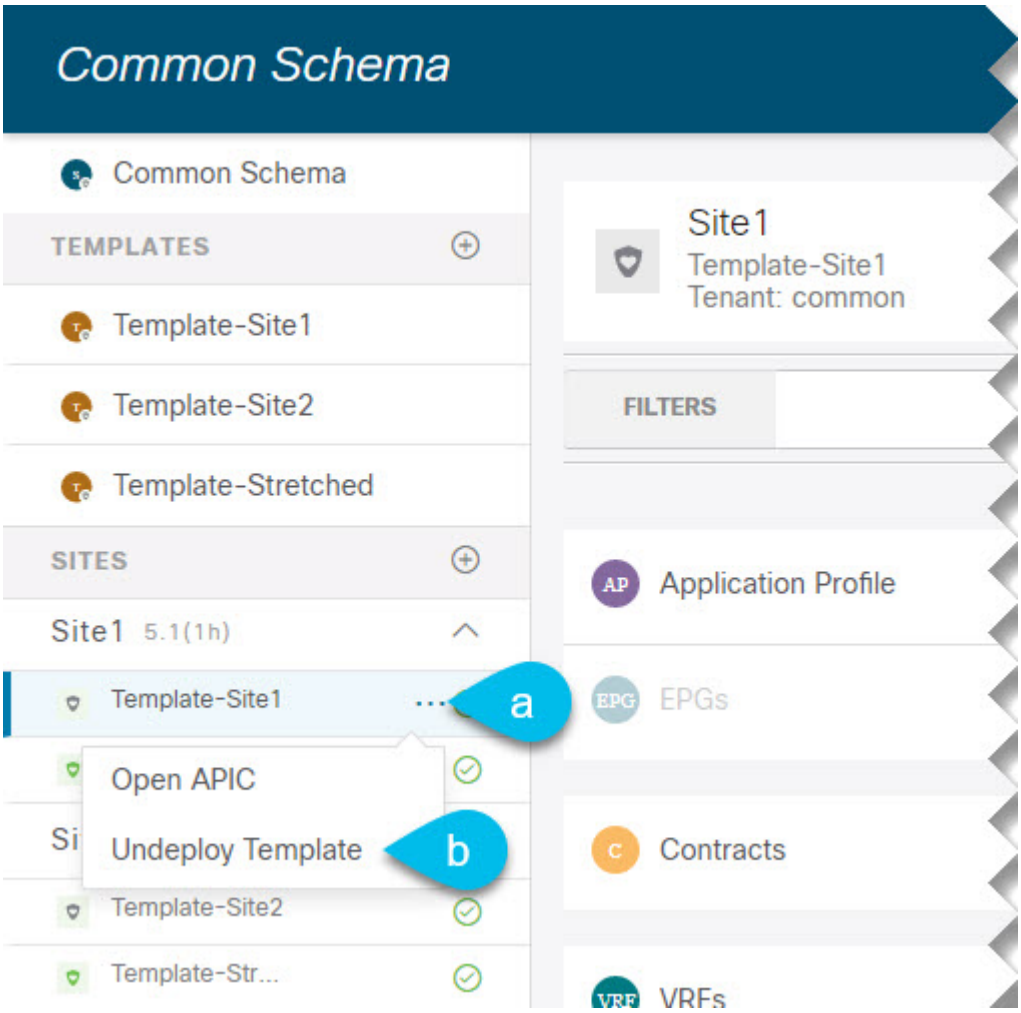
  Undeploying a template that was modified since it was last deployed would create a configuration drift because the set of objects deployed with the template would be different than the set of objects you try to undeploy after making changes to the template.

Procedure

Step 1 Select the schema that contains the template you want to undeploy.

Step 2 In the left sidebar under **SITES**, select the template you want to undeploy.

Step 3 Undeploy the template.



a) Click the **More options** (. . .) menu next to the template.
b) Click **Undeploy Template**.

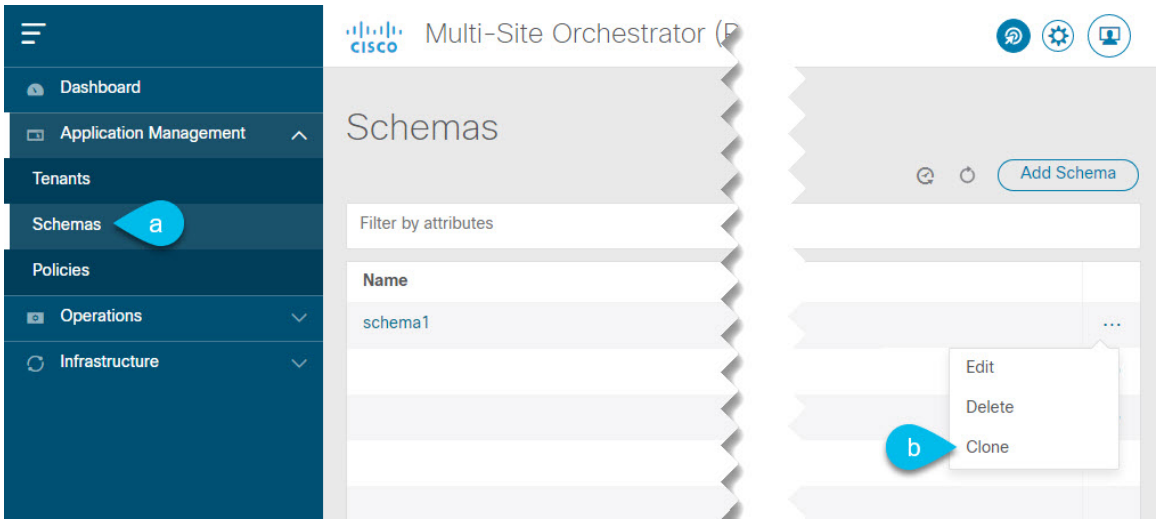
Cloning Schemas

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.

Procedure

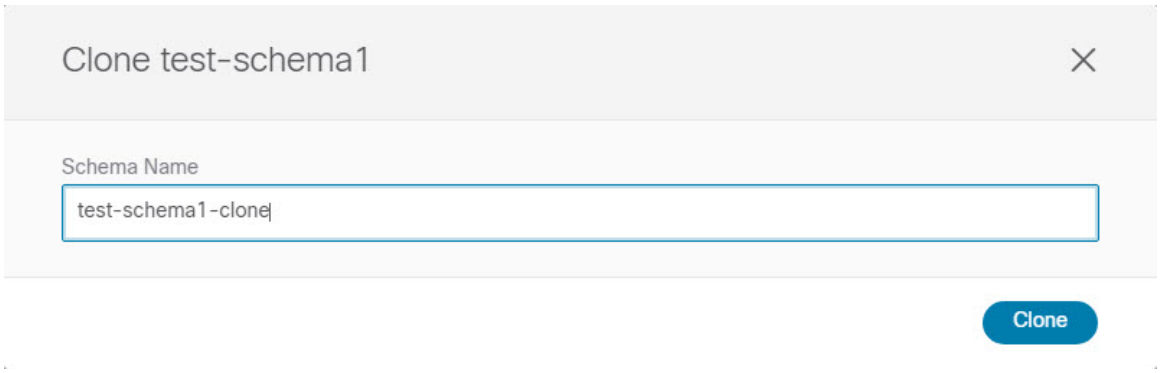
Step 1 Log in to your Multi-Site Orchestrator GUI.

Step 2 Choose the schema to clone..

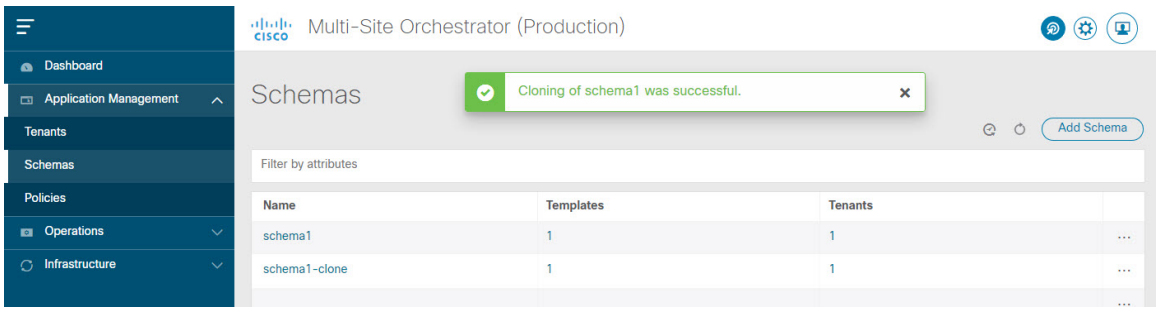


a) From the left navigation menu, select **Application Management** > **Schemas**.
b) From the **Actions** menu next to the name of the schema you want to clone, select **Clone Schema.**

Step 3 Provide the name for the new schema and click **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:

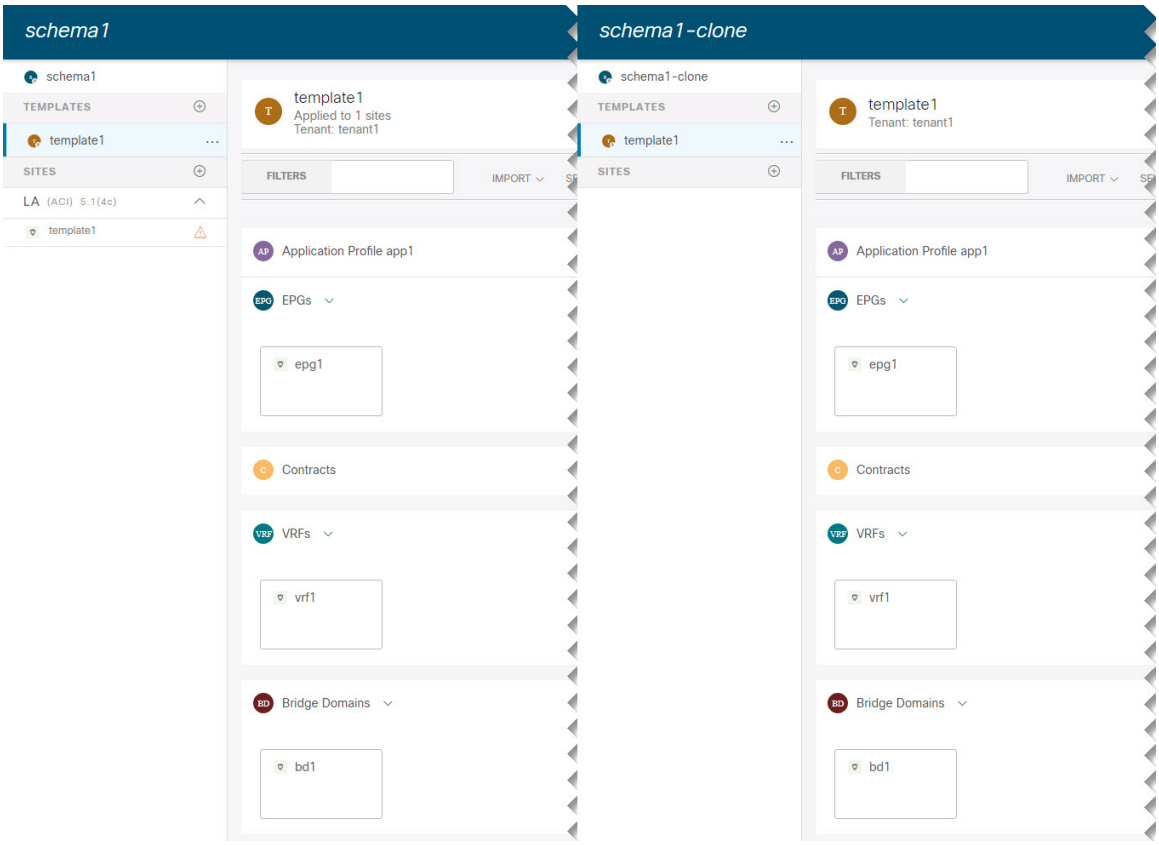


The new schema is created with the exact same templates (and their tenants' association), object, and policy configurations as the original schema.

Note that while the templates, objects, and configurations are copied, the site association is not preserved and you will need to re-associate the template in the cloned schema with any sites where you want to deploy them. Similarly, you will need to 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 the operation completed successfully by comparing the two schemas:



Cloning Templates

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

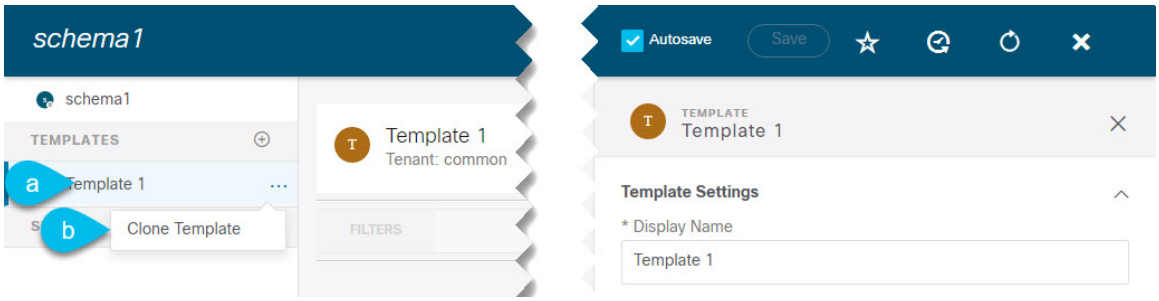
Procedure

Step 1 Log in to your Multi-Site Orchestrator GUI.

Step 2 From the left navigation menu, select **Application Management** > **Schemas**.

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

Step 4 In the Schema view, open the **Clone Template** dialog.



a) Select the Template you want to clone.
b) From the **Actions** menu, select **Clone Template**.

Step 5 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.

If the template you are cloning has cross references to objects in other templates within its schema, then if you choose a different destination schema, the other templates will be cloned as well to preserve object relationships.

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 APIC 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 Multi-Site Orchestrator GUI.

Step 2 From the left navigation menu, select **Schemas**.

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.

Template Deployment View (Beta)

Starting with Release 3.3(1), you can view all objects currently deployed to sites from a specific template. This is a beta feature in this release.

Even though any given template can be deployed, undeployed, updated, and re-deployed any number of times, this API 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

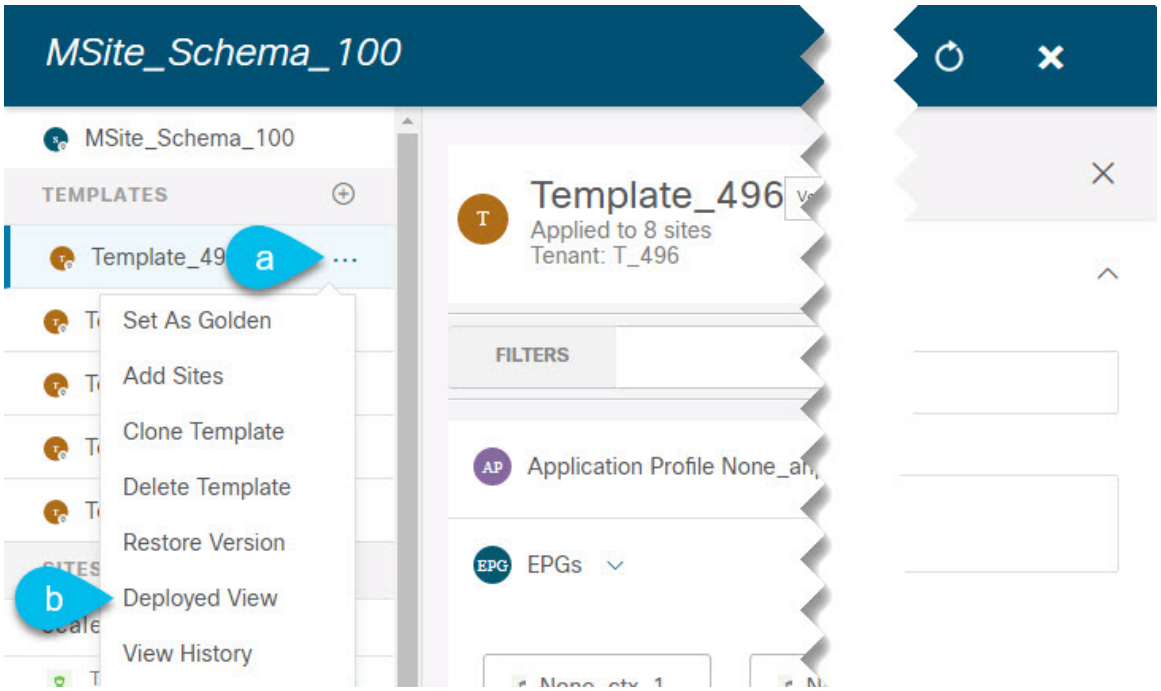
Step 1 Log in to your Multi-Site Orchestrator GUI.

Step 2 From the left navigation menu, select **Application Management** > **Schemas**.

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 **Deployment View** for the template.



a) Click the **Actions** menu next to the template's name.
b) Click **Deployment View**.

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

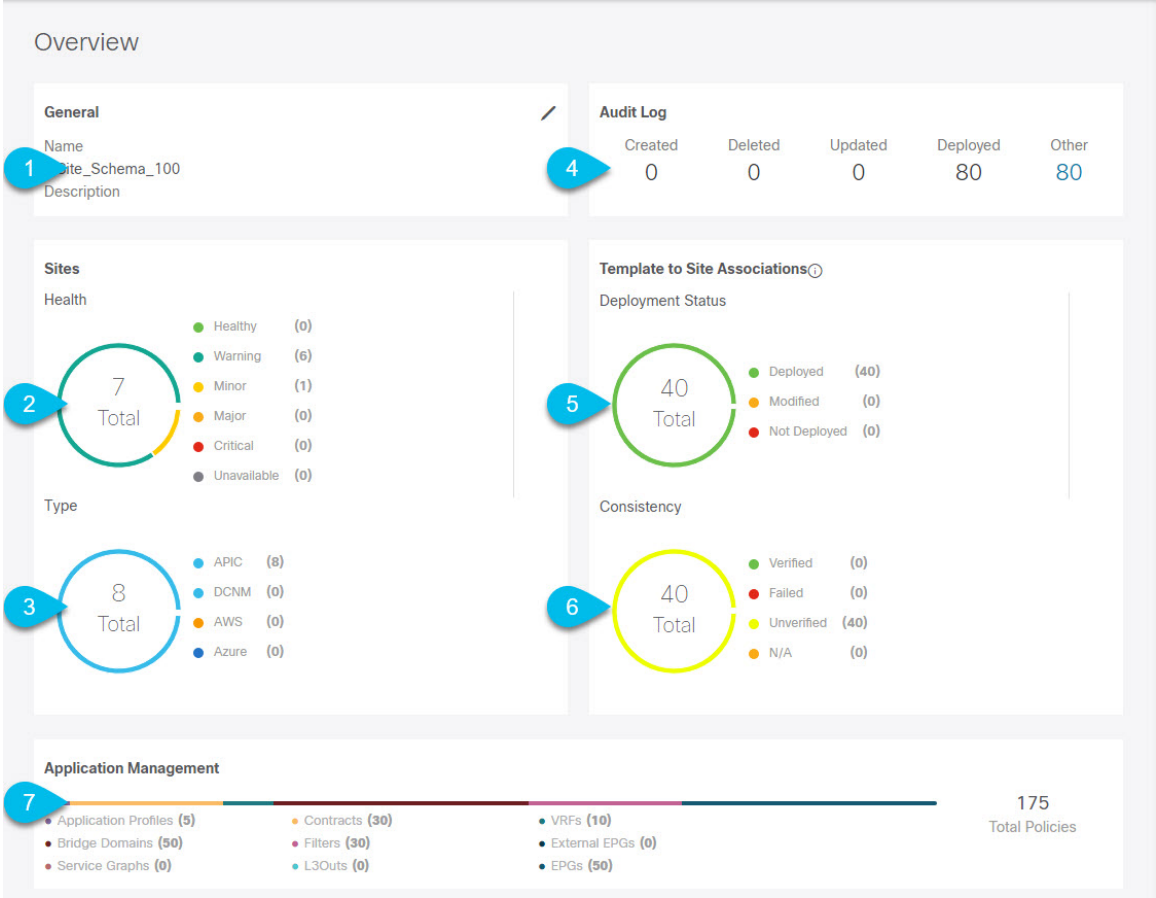
You will be able to see the XML content of all the objects that were pushed to the site when this specific template was last deployed.



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 5: Schema Overview

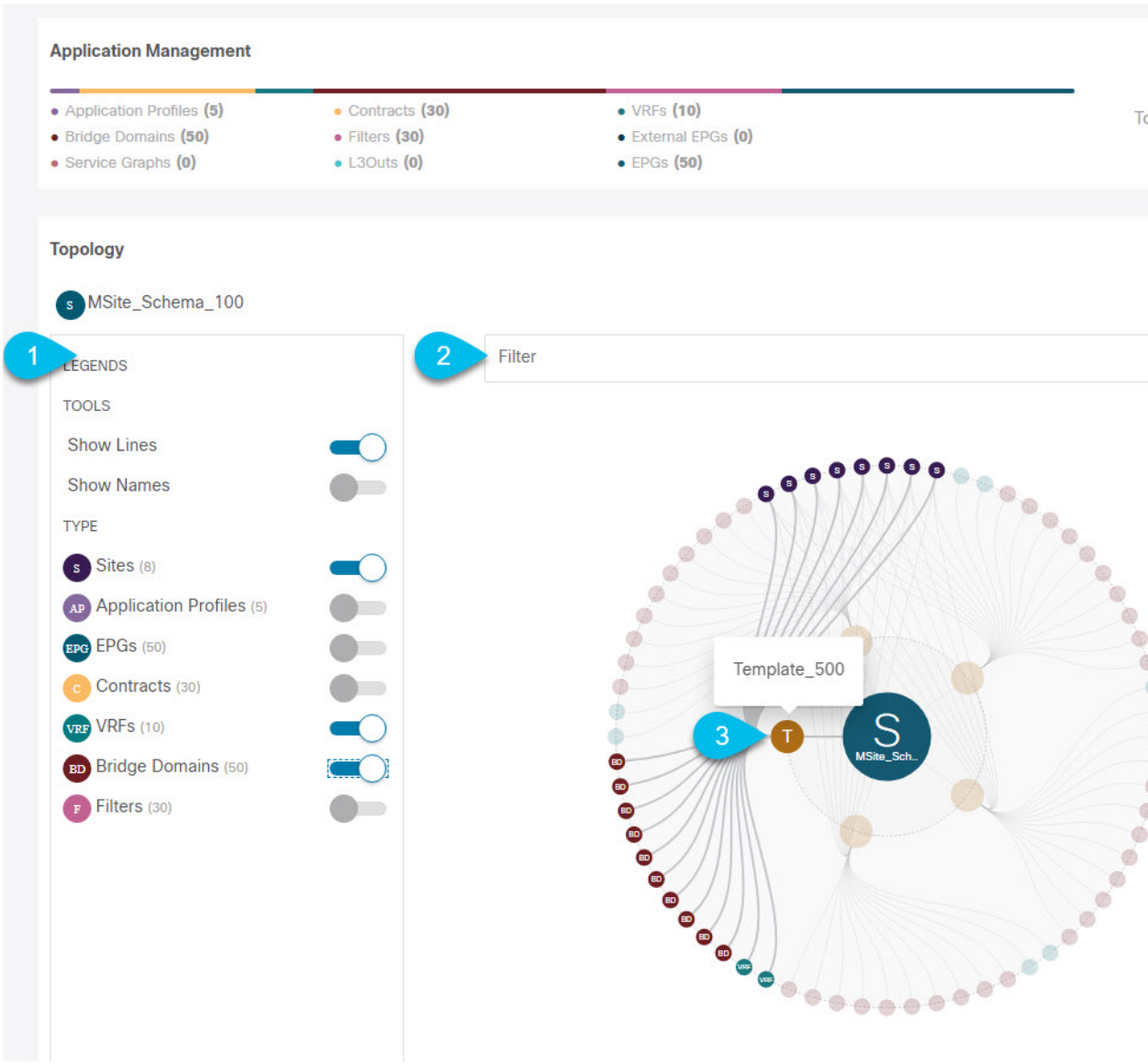


The following details are provided on this page:

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.
4. **Sites** > **Type**—Provides the number of sites associated with the templates in this schema sorted by the site's type.
5. **Template to Site Associations** > **Deployment Status**—Provides the number of templates in this schema that are associated with one or more sites and their deployment status.
6. **Template to Site Associations** > **Consistency**—Provides the number of consistency checks performed on the deployed templates and their status.
7. **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 6: 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.

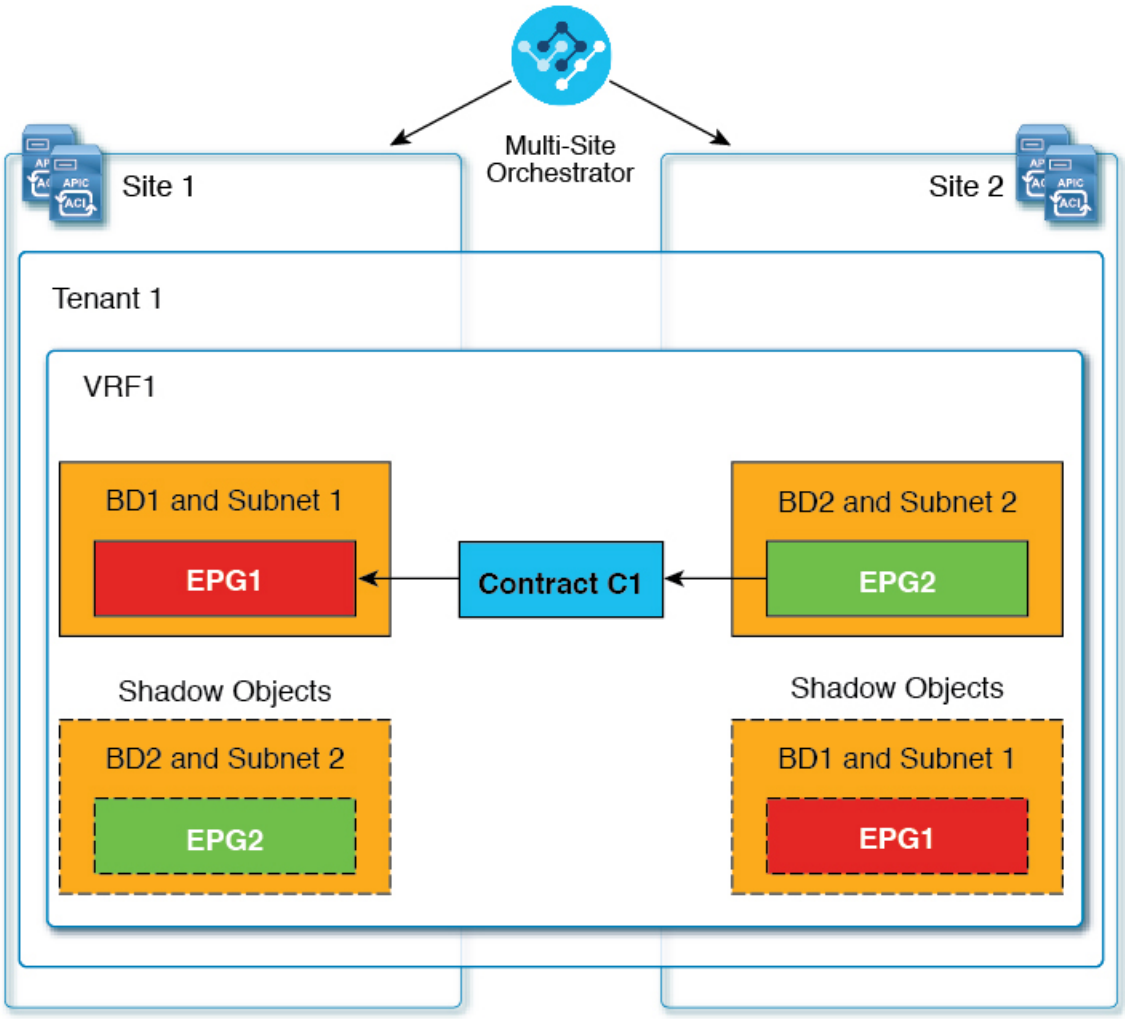
Shadow Objects

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' APICs, 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 deployed as shown in the figure below. They appear with the same names as the ones that were deployed directly to each site.

Figure 7: Basic Shadow EPG



The following objects can be shadowed:

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

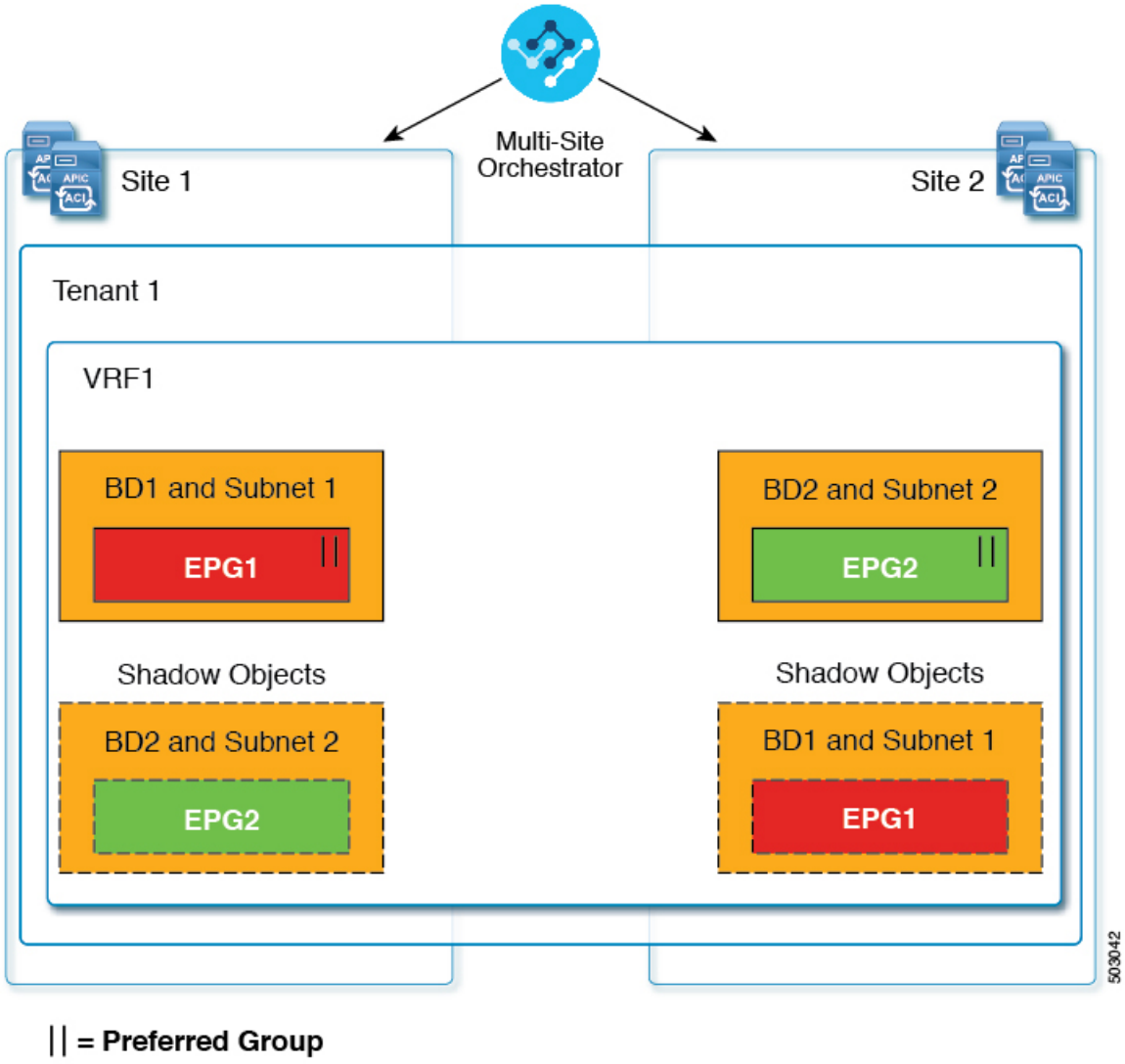
If your fabrics are running APIC Release 5.0(2) or later, when you select a shadow object in the APIC GUI, you will see a `This is a shadow object pushed by MSC to support intersite policies. Do not make any changes or delete this object.` warning at the top of main GUI pane. In addition, shadow

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. You can check for these options as described below:

Other Use Cases with Shadow Objects

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

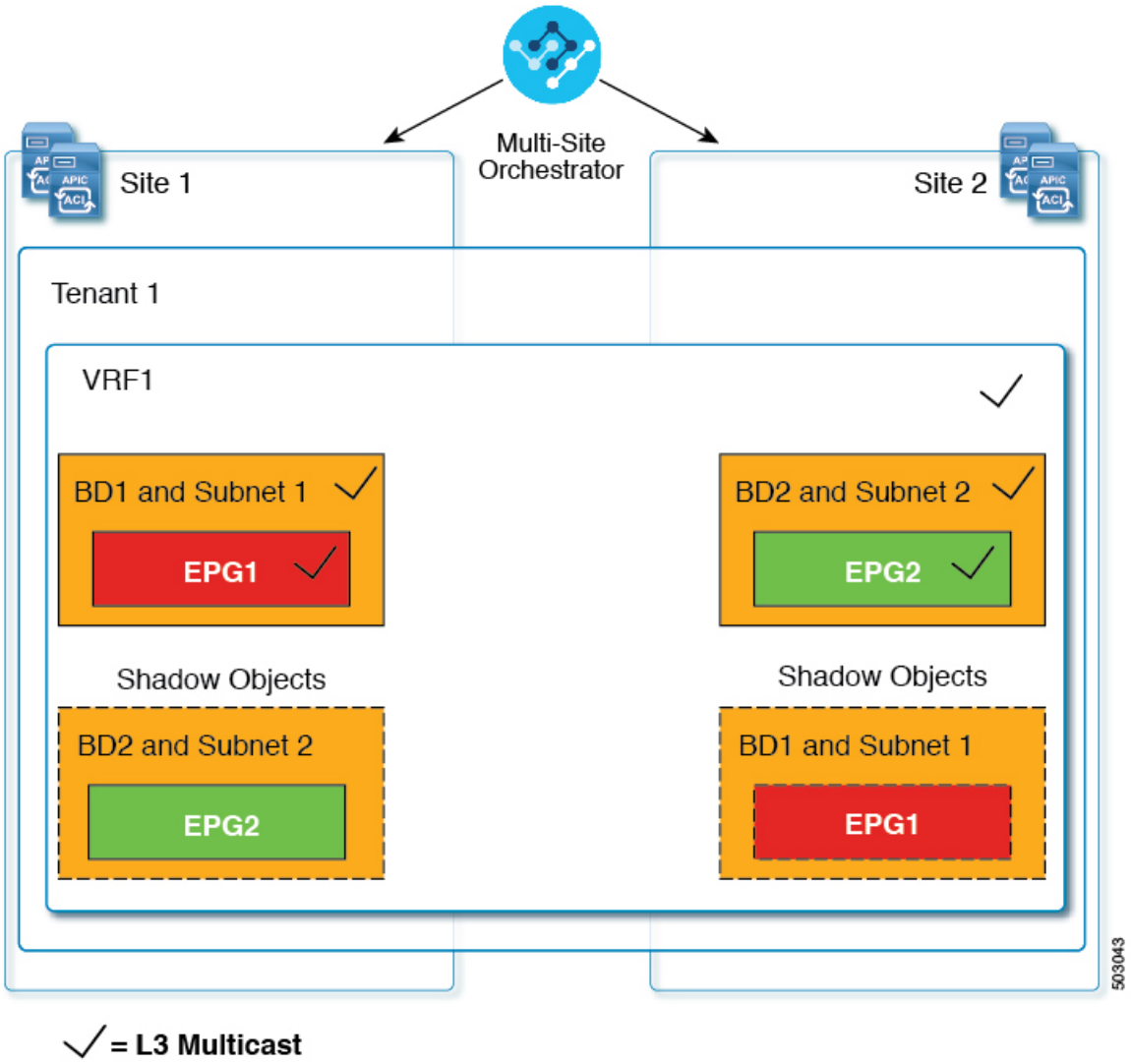
Figure 8: Preferred Group



|| = Preferred Group

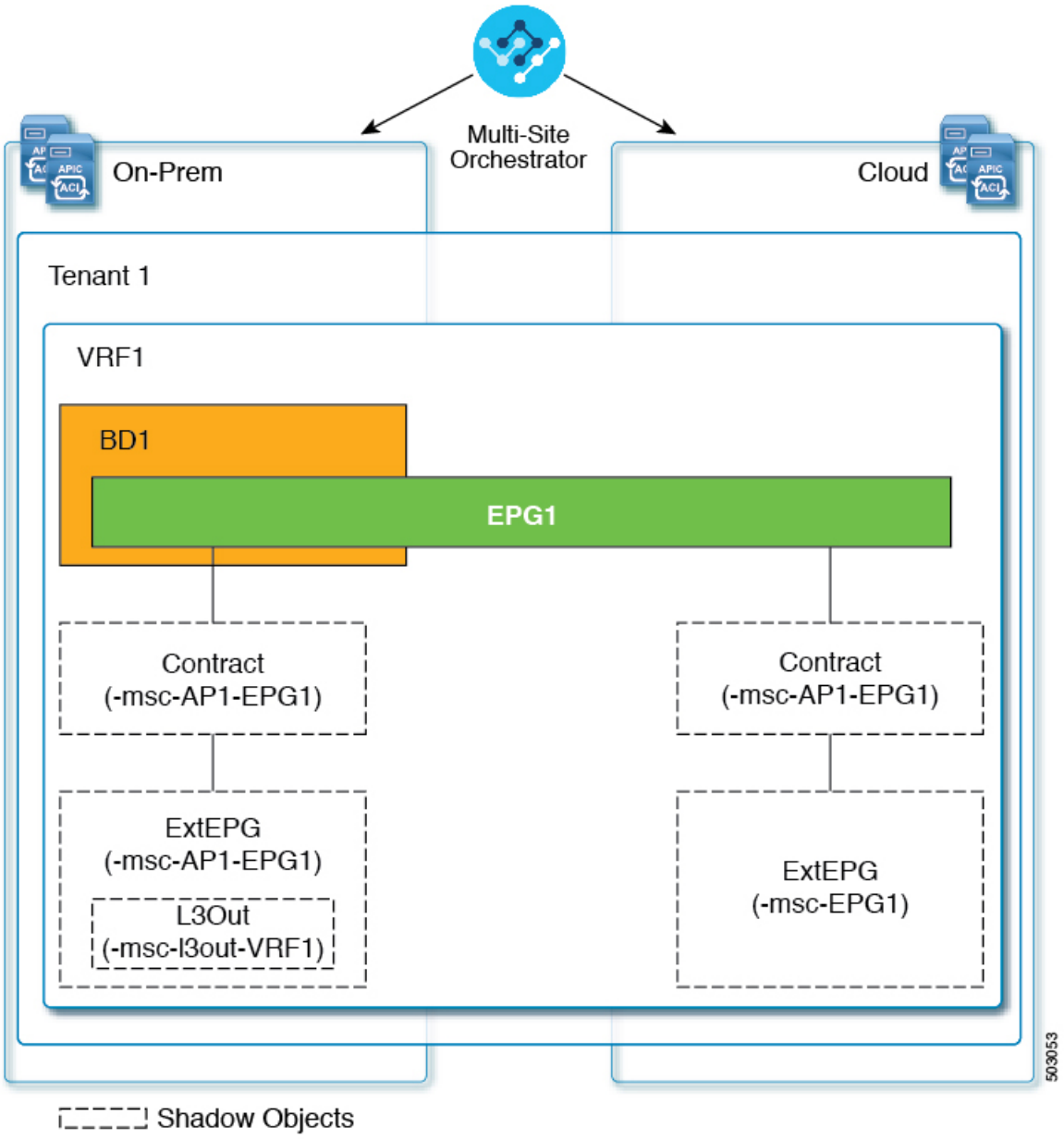
In case of multicast, the shadow objects are created only for EPGs/BDs that have multicast sources connected and the option explicitly configure at the EPG level.

Figure 9: 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 10: Hybrid Cloud



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 Multi-Site Orchestrator in the on-premises site's APIC GUI. Shadow objects in Cloud APIC 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 Multi-Site 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 Multi-Site Orchestrator.
  - When the Multi-Site 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.

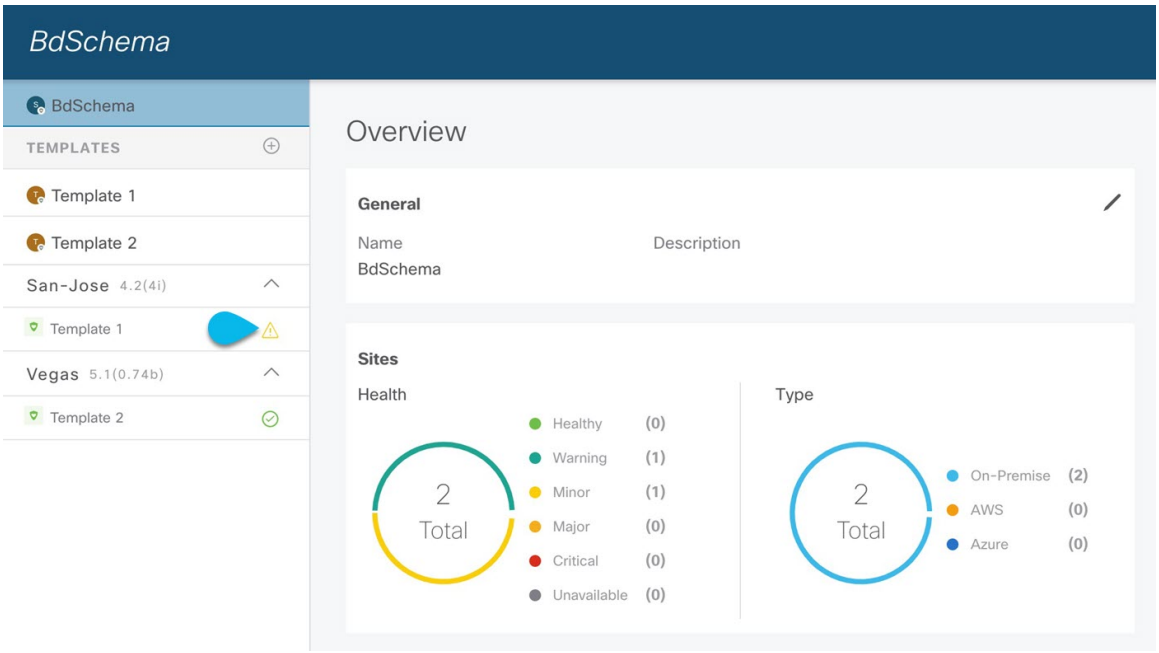
The setting is stored in the user profile and is enable or disabled separately for each user.

Step 4 Repeat the process for any additional APIC sites.

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 Multi-Site Orchestrator. These configuration discrepancies are referred to as **Configuration Drifts** and are indicated by the exclamation point next to the template name in the schema view as shown in the following figure:

Figure 11:



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

- **Configuration is modified in MSO**—when you modify a template in MSO 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 MSO 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.

  - If you want to undo the local changes, simply re-deploy the template from MSO to override any manual changes made at the site level.
  - If you want to preserve the local changes, import the modified objects from the site into the MSO template, save the schema, and re-deploy it back to the site.

- **MSO configuration is restored from backup**—restoring configuration from a backup in MSO 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 since the backup was created, restoring the backup would create a configuration drift.

  - If you want to preserve the changes since the backup, import the modified objects from the site into the MSO template after configuration restore, save the schema, and re-deploy it back to the site.
  - If you do not want to preserve the changes, simply re-deploy the template from MSO to override any configuration discrepancies at the site level.

    Note that this would allow you to override the configuration for objects and objects' properties that can be managed from MSO; it will not remove extra objects and properties configured at the APIC level that are not directly manageable from MSO.

- **MSO 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 in the site's APIC and the values are different from the defaults.
  - If you want to reset the object properties to the default values, simply re-deploy the template.
  - If you want to preserve the already configured custom values for the newly supported object properties, import the modified objects from the site into the MSO template, save the schema, and re-deploy the template back to the site.

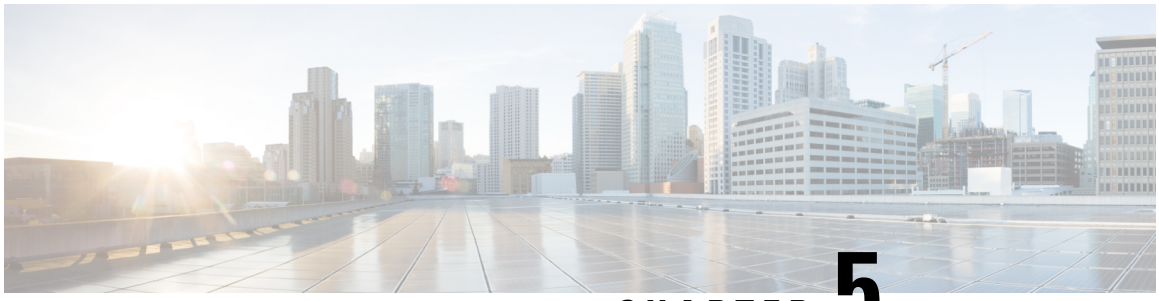
- **MSO is upgraded from an earlier release**—this scenario is similar to the previous one where if new object properties are added in the new release, existing configuration may indicate a drift.
  - If you want to reset the object properties to the default values, simply re-deploy the template.
  - If you want to preserve the already configured custom values for the newly supported object properties, import the modified objects from the site into the MSO template, save the schema, and re-deploy the template back to the site.



PART II

Operations

- Audit Logs, on page 61
- Backup and Restore, on page 63
- Upgrading Sites, on page 71
- Tech Support, on page 83



CHAPTER 5

Audit Logs

- Audit Logs, on page 61

Audit Logs

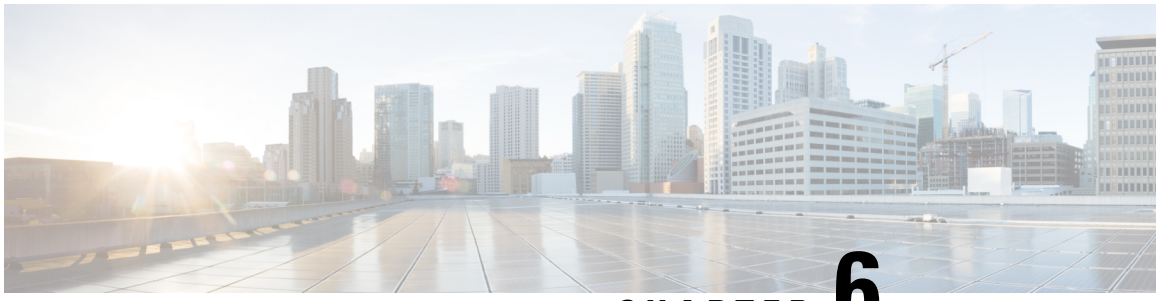
Multi-Site Orchestrator system logging is automatically enabled when you first deploy the Orchestrator cluster and captures the events and faults that occur in the environment.

You can view the Multi-Site Orchestrator logs directly in the GUI by selecting **Operations** > **Audit Logs** from the main navigation menu.

From the **Audit Logs** page, you can click the **Most Recent** 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, Logged in, Logged Out, Login Failed. Select an action and click **Apply** to filter the log details according to the action.



CHAPTER 6

Backup and Restore

- Configuration Backup and Restore, on page 63
- Backup and Restore Guidelines, on page 63
- Remote Backups, on page 65
- Creating Backups, on page 68
- Restoring Backups, on page 68
- Downloading Backups, on page 69
- Importing Backups, on page 70

Configuration Backup and Restore

You can create backups of your Multi-Site 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. We also recommend exporting the backups to an external storage outside of the Multi-Site Orchestrator nodes.

Note Restoring a backup action restores the database on the Multi-Site Orchestrator, but it does not make any changes to the controller databases on each site. Therefore, after you restore the Orchestrator database, you must also re-deploy any existing schemas to avoid potentially mismatching policies between the Multi-Site Orchestrator and each site's controller. For information on specific configuration mismatch scenarios and backup restore procedures related to each one, see Backup and Restore Guidelines, on page 63

Backup and Restore Guidelines

When saving and restoring configuration backups, the following guidelines apply:

- When deploying a new multisite controller cluster and restoring the config from backup, refrain from manually adding the sites. The sites will be added automatically from the backup. Adding them manually can potentially disrupt the existing intersite connectivity.
- Importing and restoring backups created from later releases is not supported.

  For example, if you downgrade your Multi-Site Orchestrator to an earlier release, you cannot restore a backup of the configuration created on a later release.

- Restoring configuration backups created on releases prior to Release 3.2(1) is supported as a one-time step during cluster migration to Nexus Dashboard.

  Subsequent restore of backups from VMware ESX or Application Services Engine deployments are not supported.

- 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 Multi-Site Orchestrator, but it does not make any changes to the controller databases on each site. As such, certain precautions and steps must be taken when restoring a previous configuration to avoid potentially mismatching policies between the Multi-Site Orchestrator and the sites, as described below.

No Configuration Changes Since Backup

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 68.

Sites, Objects, or Policies Created, Modified, or Deleted Since Backup

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 created and deployed since the backup will remain deployed. You will need to manually remove these after restoring the backup to avoid any stale configurations.

  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 defined by those policies.

- The steps required to restore a configuration backup are described in Restoring Backups, on page 68.

- 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, even though none of the configurations will exist in the sites yet. In this case, in order for the configuration to be properly pushed to each site, you will need to re-deploy it to sync the Multi-Site Orchestrator's configuration with the sites.

- If sites that were managed when the backup was created are no longer present in the Nexus Dashboard, the restore will fail.

- If sites' status since the backup has changed (`managed` vs `unmanaged`) but the sites are still present in the Nexus Dashboard, the status will be restored to what it was at the time of backup.

Remote Backups

Multi-Site Orchestrator is deployed as an application in Nexus Dashboard. When you first deploy the cluster, any backups you create are saved to a default location which is located on each node's local disk.

While the backups are available on any one node and can be viewed using the MSO GUI, we recommend exporting all backups to a remote location outside the Nexus Dashboard nodes.

You can configure a remote SCP or SFTP location using the MSO GUI and then exporting the backup files there as described in the following sections.



Note When you create a configuration backup and export it to a remote server, the files are first created on the Orchestrators' local drives, then uploaded to the remote location, and finally deleted from the local storage. There is a limit on the local backups disk space usage, which if reached can prevent remote backups from being created.

Configuring a Remote Location for Backups

This section describes how to configure a remote location in Multi-Site Orchestrator to which you can then export your configuration backups.

Procedure

Step 1 Log in to your Multi-Site Orchestrator.

Step 2 From the left navigation pane, select **Operations** > **Remote Locations**.

Step 3 In the top right of the main window, click **Add Remote Location**.

An **Add New Remote Location** screen appears.

Step 4 Provide the name for the 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 will 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 already 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 used to log in to the remote server.
- `SSH Private Files`—provide the username and the SSH Key/Passphrase pair used to log in to the remote server.

Step 9 Click **Save** to add the remote server.

Moving Existing Backups to a Remote Location

This section describes how to move an existing configuration backup you have created in the Multi-Site Orchestrator GUI from the nodes' local drives to a remote location.

Before you begin

You must have completed the following:

- Created a configuration backup as described in Creating Backups, on page 68.
- Added a remote location for exporting backups as described in Configuring a Remote Location for Backups, on page 65.

Procedure

Step 1 Log in to your Multi-Site Orchestrator.

Step 2 From the left navigation pane, select **Operations** > **Backups & Restore**.

Step 3 Locate the backup you want to export, then click the actions (**...**) icon next to it, then click **Move to remote location**.

A **Move Backup To Remote Location** window opens.

Step 4 From the **Remote Location** dropdown menu, select the remote location.

Step 5 (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 additional subdirectories to the path. However, the directories must be under the default configured path and must have been already created on the remote server.

Adding an NFS Share to Store Backups

This section describes how to add an NFS share to the Multi-Site Orchestrator VMs to store configuration backups.

Note While you can configure a single remote NFS share for your configuration backups, we recommend using the remote backup location feature available in the Orchestrator GUI and described in Configuring a Remote Location for Backups, on page 65 instead.

Procedure

Step 1 Log in directly to your Multi-Site Orchestrator node's VM as the `root` user.

Step 2 Mount the NFS share.

The following command mounts the shared NFS directory to the default Orchestrator backups folder so all future backups are automatically stored to an external storage outside the Orchestrator VMs.

Note
If you have any existing backups in this default directory that you want to save, you must manually move them to a different location before mounting the NFS share. After the share is mounted, any existing files in the mount directory will be hidden from view.

```
# mount <nfs-server-ip>:/<nfs-share-path> /opt/cisco/msc/backups/
```

Step 3 Repeat steps 1 through 2 on each Orchestrator VM.

Because each Orchestrator node can create and store its own backup files, you must mount the same NFS share on all nodes.

Step 4 Update the Docker backup services.

You must run the following Docker update command for the newly mounted file system to be usable by the Orchestrator services. However, since the command updates the services across the cluster, you only need to do this once after mounting the shares on each node.

```
# docker service update msc_backupservice --force
```

What to do next

If at any point you want to remove the NFS share and go back to storing the backups locally on each VM, simply unmount the directory on each node and run the `docker service update msc_backupservice --force` command again.

Creating Backups

This section describes how to create a new backup of your Multi-Site Orchestrator configuration.

Before you begin

If you want to create the backup using a remote location, you must first add the remote location as described in Configuring a Remote Location for Backups, on page 65.

Procedure

Step 1 Log in to your Multi-Site Orchestrator.

Step 2 Create a new backup.

a) From the left navigation pane, select **Operations** > **Backups & Restore**.

b) In the main window, click **New Backup**.

A **New Backup** window opens.

Step 3 Provide backup information.

a) Provide the **Name** and optional **Notes** for the backup.

The name can contain up to 10 alphanumeric characters, but no spaces or underscores (_).

b) Choose the **Backup Location**.

You can save the backup file locally on the Orchestrator nodes or export it to a remote location.

If you want to save the backup file locally, choose **Local**.

Otherwise, if you want to save the backup file to a remote location, choose **Remote** and provide the following:

- From the **Remote Location** dropdown menu, select the remote location.

  The remote location must be already created as described in Configuring a Remote Location for Backups, on page 65.

- In the **Remote Path**, either leave the default target directory or you can choose to append additional subdirectories to the path. However, the directories must be under the default configured path and must have been already created on the remote server.

c) Click **Save** to create the backup.

Restoring Backups

This section describes how to restore a Multi-Site Orchestrator configuration to a previous state.

Before you begin

Restoring a backup action restores the database on the Multi-Site Orchestrator, but it does not make any changes to the controller databases on each site. Therefore, after you restore the Multi-Site Orchestrator database, you must also re-deploy any existing schemas to avoid potentially mismatching policies between the Multi-Site Orchestrator and sites.

For information on specific configuration mismatch scenarios and recommended restore procedures related to each one, see Backup and Restore Guidelines, on page 63.

Procedure

Step 1 Log in to your Multi-Site 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. Additional context is available in Backup and Restore Guidelines, on page 63.

Step 3 From the left navigation menu, select **Operations** > **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 Multi-Site 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 terminates the current session and the user is logged out.

Note
If your Multi-Site Orchestrator cluster is deployed in Application Services Engine, multiple services are restarted during the configuration restore process. As a result, you may notice an up to 5 minute delay before the restored configuration is properly reflected in the MSO GUI.

Step 6 If necessary, redeploy the configuration.

We recommend you perform this step to sync the restored configuration with the sites. Additional context is available in Backup and Restore Guidelines, on page 63.

Downloading Backups

This section describes how to download you backup from the Multi-Site Orchestrator.

Before you begin

Procedure

Step 1 Log in to your Multi-Site Orchestrator GUI.

Step 2 From the left navigation menu, select **Operations** > **Backups & Restore**.

Step 3 In the main window, click the actions (**...**) icon next to the backup you want to download and select **Download**.

This will download 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

This section describes how to import an existing backup into your Multi-Site Orchestrator.

Before you begin

Procedure

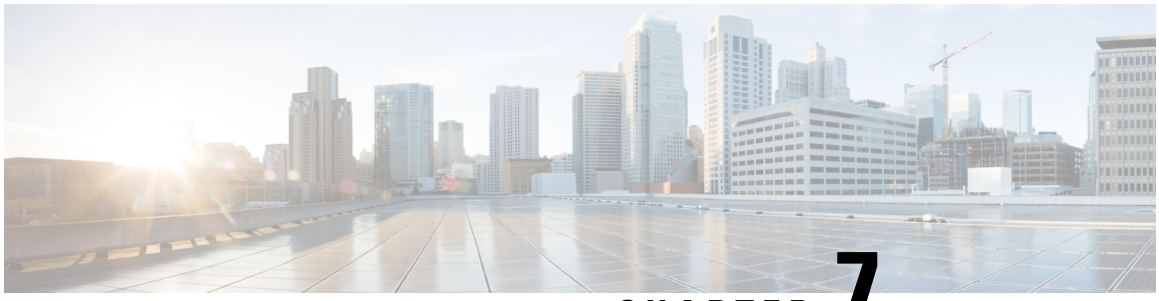
Step 1 Log in to your Multi-Site Orchestrator GUI.

Step 2 From the left navigation menu, select **Operations** > **Backups & Restore**.

Step 3 In the main window, click **Import**.

Step 4 In the **Import from file** window that opens, click **Select File** and choose the backup file you want to import.

Importing a backup will add it to the list of the backups displayed the **Backups** page.



CHAPTER 7

Upgrading Sites

- Overview, on page 71
- Guidelines and Limitations, on page 73
- Downloading Controller and Switch Node Firmware to Sites, on page 73
- Upgrading Controllers, on page 76
- Upgrading Nodes, on page 78

Overview

Note This feature is supported for Cisco APIC sites only. It is not supported for Cisco Cloud APIC or Cisco DCNM fabrics.

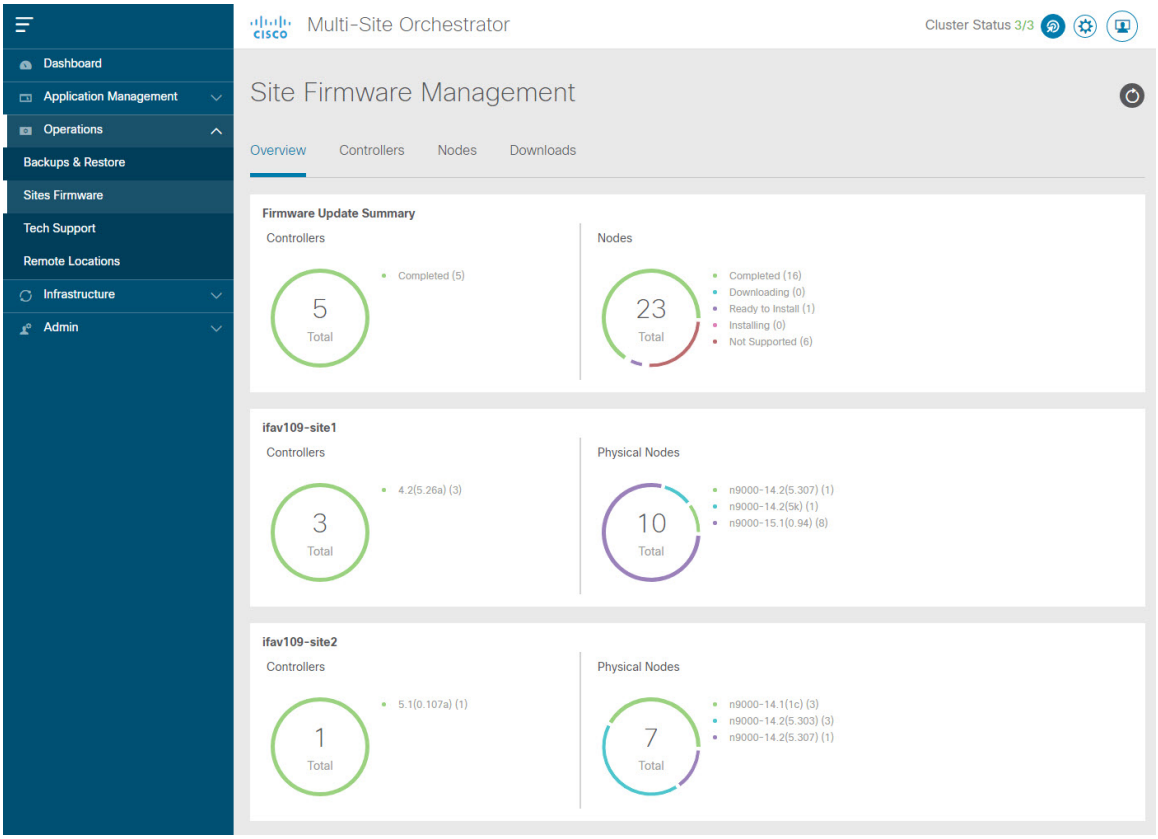
Prior to Release 3.1(1), when you deployed Cisco Multi-Site, each site's APIC clusters and switch nodes software had to be managed individually at the site level. As the number of sites in your Multi-Site domain grew, the release life cycle and upgrades could become complicated as they had to be manually coordinated and managed for release and feature compatibility.

Starting with Release 3.1(1), Cisco Multi-Site 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 **Operations** > **Sites Firmware**. 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 prior to Release 4.2(5).

- **Site-specific information**—additional 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 Multi-Site 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 Multi-Site Orchestrator must be deployed in Cisco Nexus Dashboard.

  The site upgrade feature is not available for MSO deployments in VMware ESX and you need to follow the standard upgrade procedures 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 will be unable to complete. You will not be able to remove or modify the node upgrade policy from MSO during this time as MSO 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 MSO upgrade policy will recognize the change and return a `failed` status. Then you can update the upgrade policy on the MSO to remove the switch and re-run the upgrade.

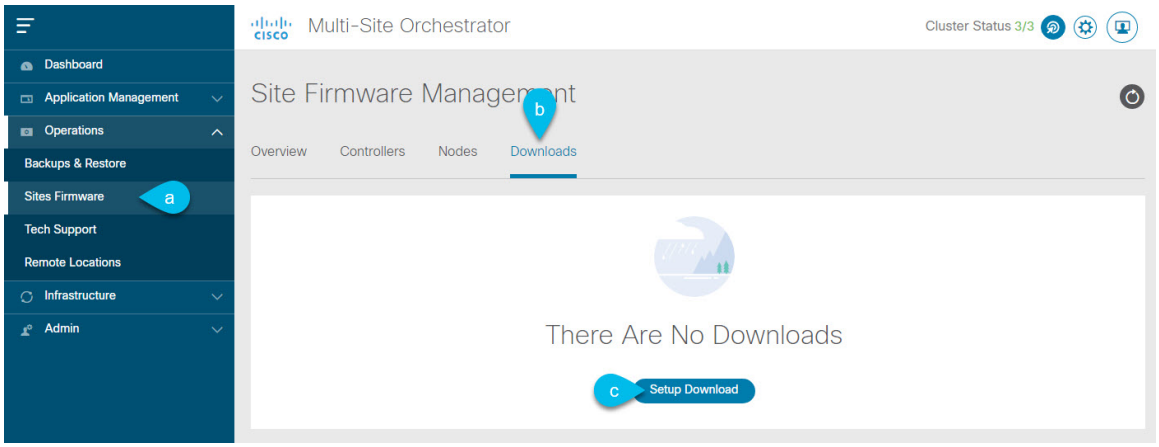
Downloading Controller and Switch Node Firmware to Sites

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 at a later time using the downloaded images.

Procedure

Step 1 Log in to your Multi-Site Orchestrator.

Step 2 Set up firmware download.



a) From the left navigation pane, select **Operations** > **Sites Firmware**.
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.

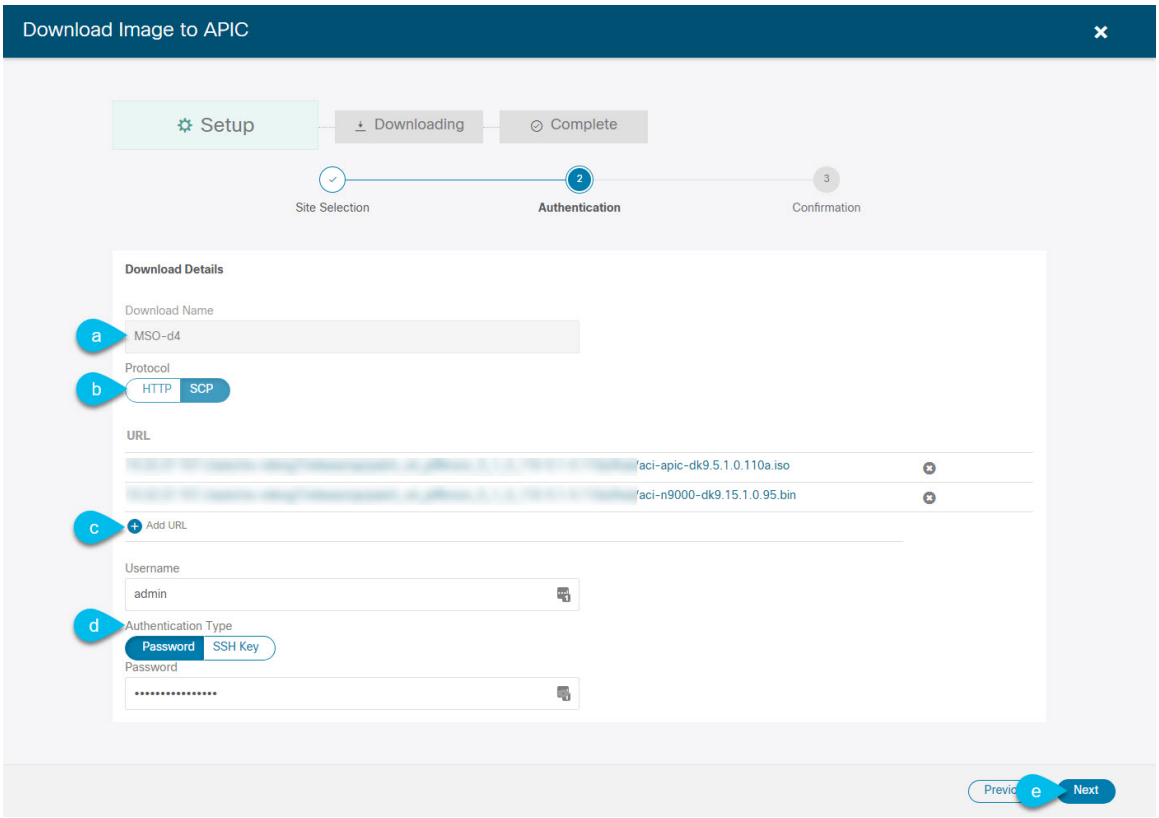
   The **Download Image to APIC** screen opens.

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 via `HTTP` or `SCP`.

c) Click **+ Add URL** to provide location of the image(s).

You can provide both, the APIC and the switch firmware images.

d) If you selected **SCP**, provide the authentication information.

You need to provide the login **Username**, for example `admin`.

Then choose the **Authentication Type**:

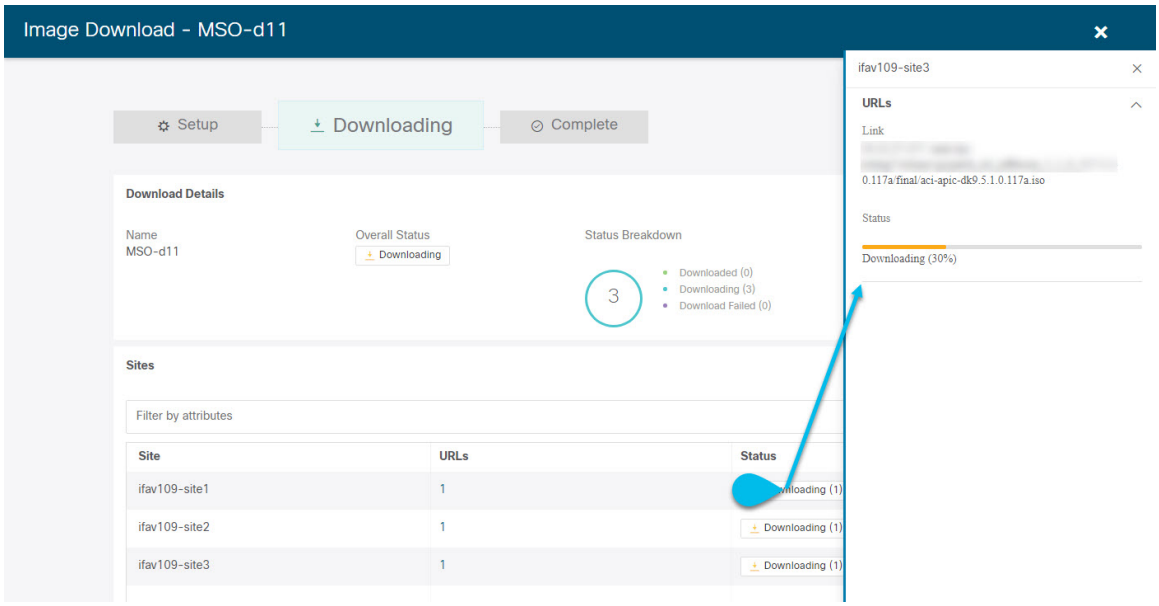
- For **Password** authentication, simply enter the password for the username you provided earlier.
- For **SSH Key** authentication, you will need to enter the **SSH Key** and the **SSH Key Passphrase**.

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 additional details about the progress.



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 you provided in a previous step.

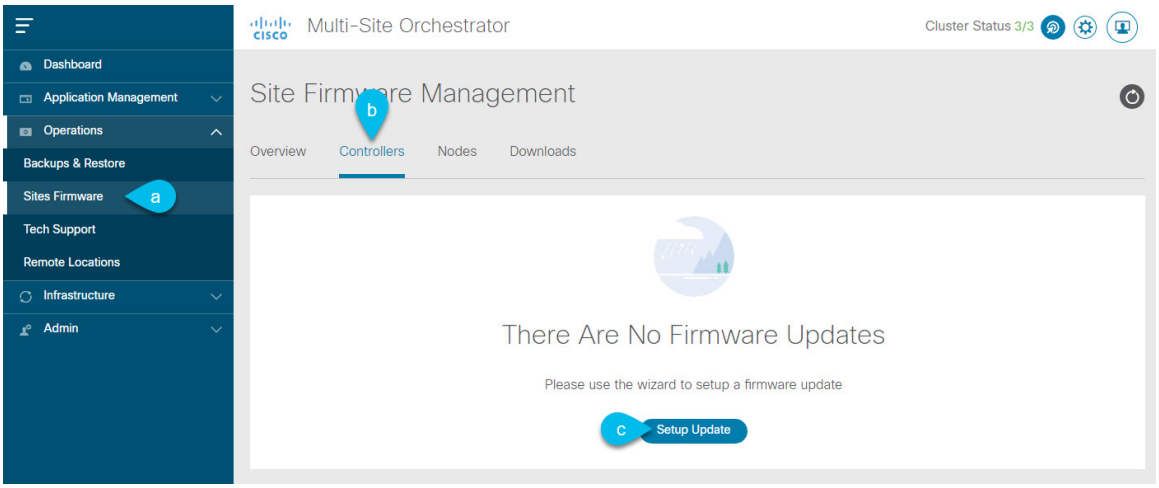
Upgrading Controllers

This section describes how to set up a software upgrade for your sites' APIC clusters.

Procedure

Step 1 Log in to your Multi-Site Orchestrator.

Step 2 Set up APIC cluster upgrade.



a) From the left navigation pane, select **Operations** > **Sites Firmware**.
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 you will be able to use to track the upgrade progress at any time.

b) Click **Select Sites**.

   The **Select Sites** window opens.

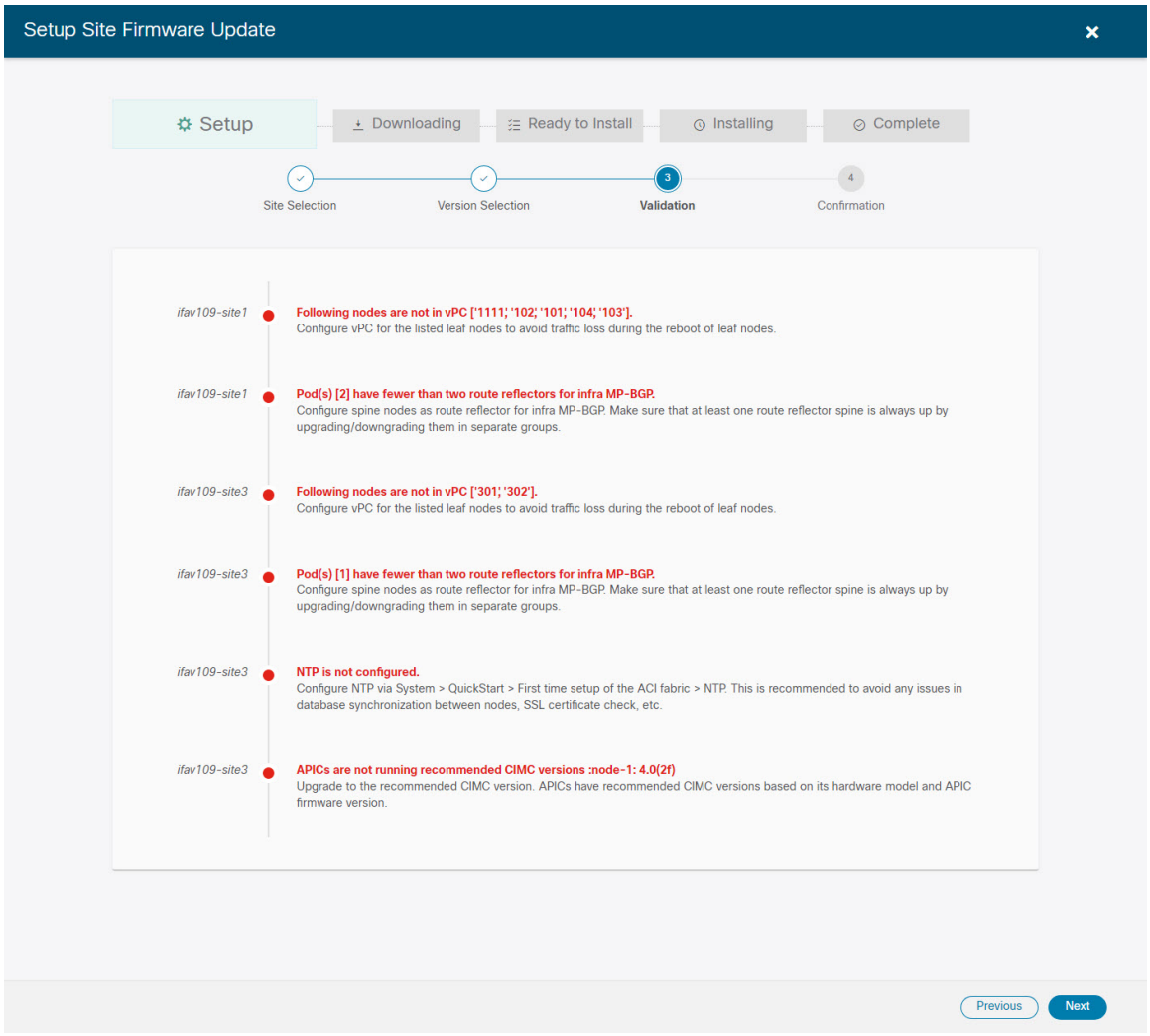
c) In the **Select Sites** window, check one or more sites and click **Add and Close**.
d) Click **Next** to proceed.

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 **Operations** > **Sites Firmware** > **Controllers** 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:



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 MSO to site connectivity is lost during the upgrade process, the GUI will display the last known status of the upgrade prior to loss of connectivity. Once connectivity is re-established, 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.

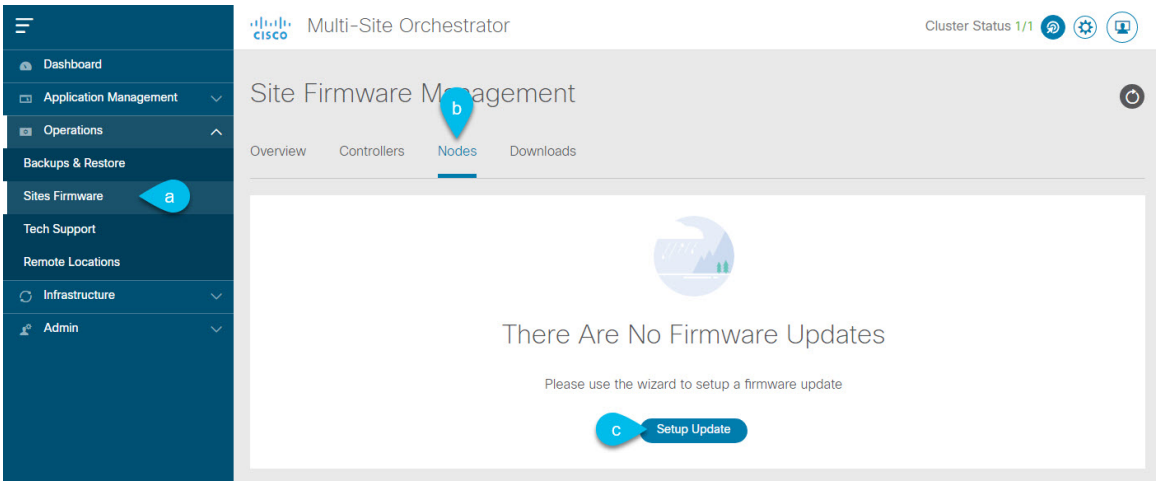
Upgrading Nodes

This section describes how to set up a software upgrade for your sites' switch nodes.

Procedure

Step 1 Log in to your Multi-Site Orchestrator.

Step 2 Set up switch nodes upgrade.



a) From the left navigation pane, select **Operations** > **Sites Firmware**.
b) In the main window, select the **Nodes** 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 Node Firmware Update** screen opens.

Step 3 Provide the upgrade details.

a) Provide the **Name**.

This is the upgrade policy name you will be able to use to track the upgrade progress at any time.

b) Click **Select Nodes**.

The **Select Nodes** window opens.

c) 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 will repeat this step if you want to add switches from other sites.

d) Repeat the previous substep for nodes in other sites
e) Click **Next** to proceed.

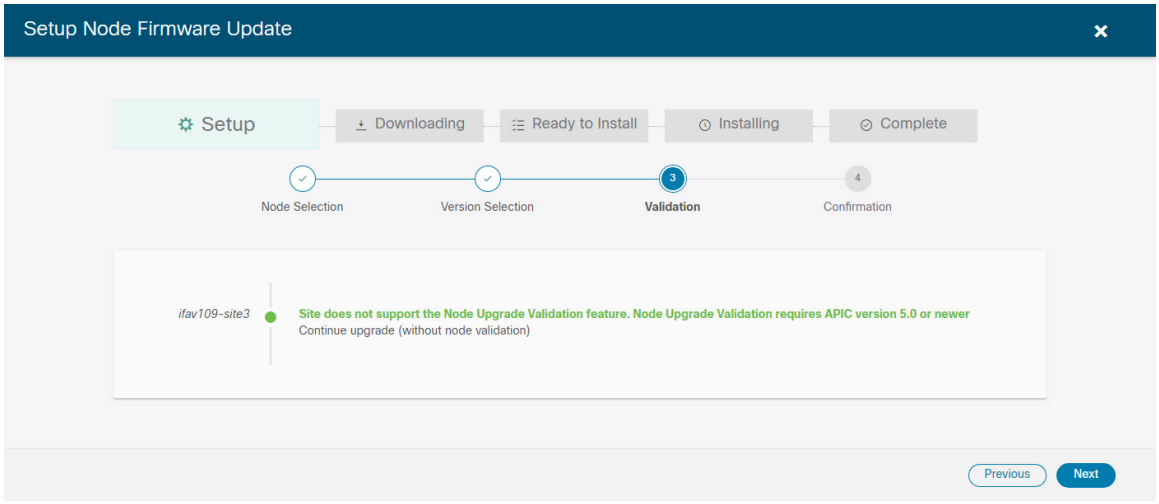
Step 4 In the **Version Selection** screen, select the firmware version and click **Next**.

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 **Operations** > **Sites Firmware** > **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 prior to Release 5.0(1) do not support node validation, so we recommend checking for any switch node faults in the site's APIC prior to starting the upgrade from MSO.



Step 6 In the **Confirmation** screen, review the information and click **Submit**.

This will trigger image to be pre-downloaded 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 additional 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 will have 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.

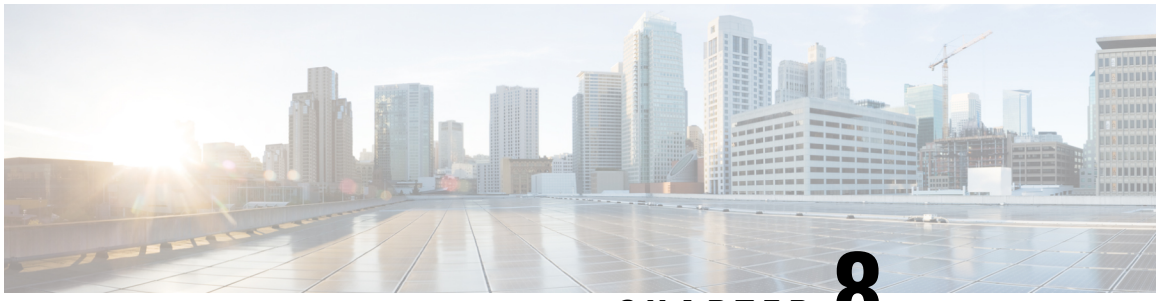
Step 8 Remove any nodes marked as `Failed` from the upgrade.

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 MSO to site connectivity is lost during the upgrade process, the GUI will display the last known status of the upgrade prior to loss of connectivity. Once connectivity is re-established, 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 8

Tech Support

- Tech Support and System Logs, on page 83
- Downloading System Logs, on page 84
- Streaming System Logs to External Analyzer, on page 84

Tech Support and System Logs

Multi-Site 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 additional tools to quickly parse, view, and respond to important events without a delay.

Starting with Release 3.3(1), 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 will 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`—Backup 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
...
```

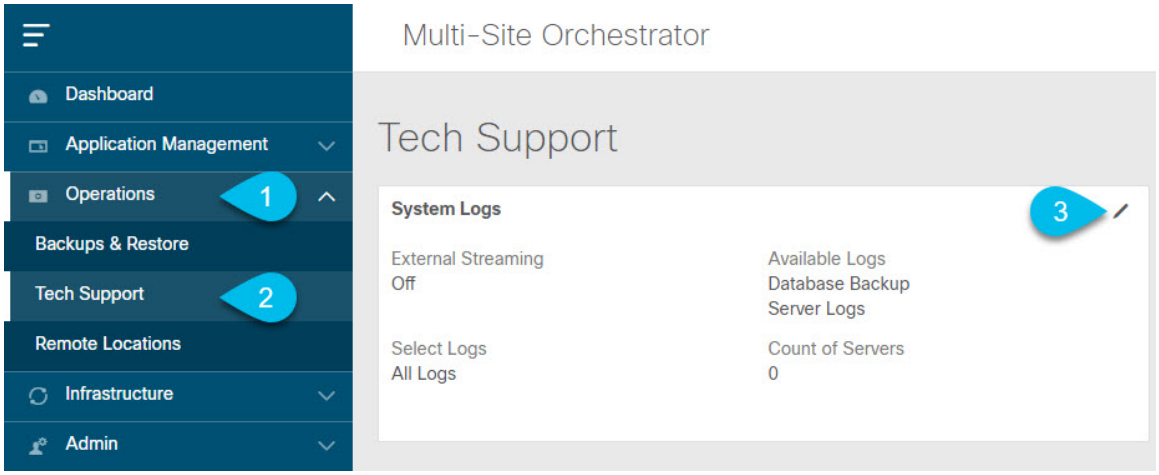
Downloading System Logs

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 Multi-Site Orchestrator.

Procedure

Step 1 Log in to your Multi-Site Orchestrator GUI.

Step 2 Open the **System Logs** screen.



a) In the main menu, select **Operations** > **Tech Support**.

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

Multi-Site 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 Multi-Site 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` only for Multi-Site Orchestrator in Application Services Engine deployments.

- This release supports up to 5 external servers.
- If using Splunk, set up and configure the log analyzer service provider.

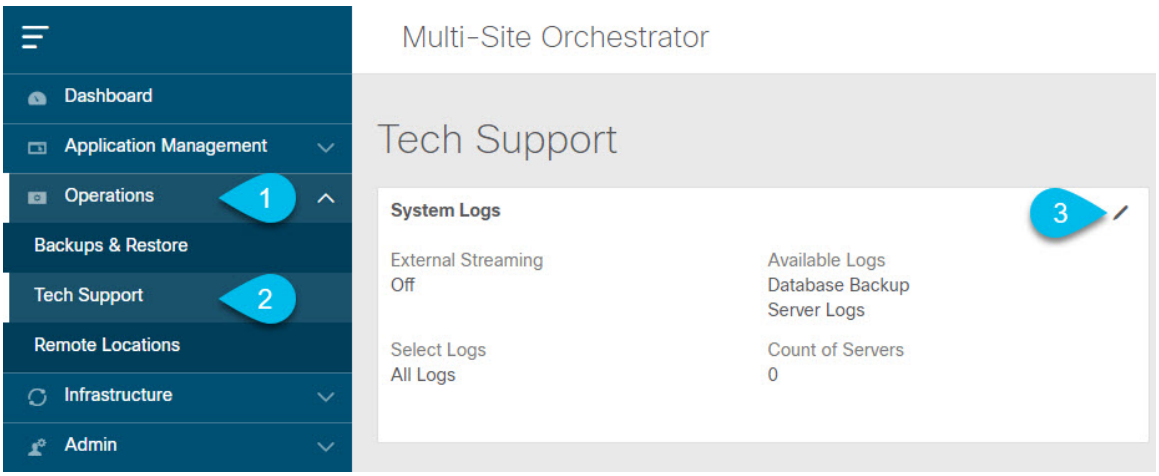
  For detailed instructions on how to configure an external log analyzer, consult its documentation.
- If using Splunk, obtain an authentication token for the service provider.

  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 Multi-Site Orchestrator GUI.

Step 2 Open the **System Logs** screen.



a) In the main menu, select **Operations** > **Tech Support**.
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.



a) Enable the **External Streaming** knob.
b) Choose whether you want to stream **All Logs** or just the **Audit Logs**.
c) 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.



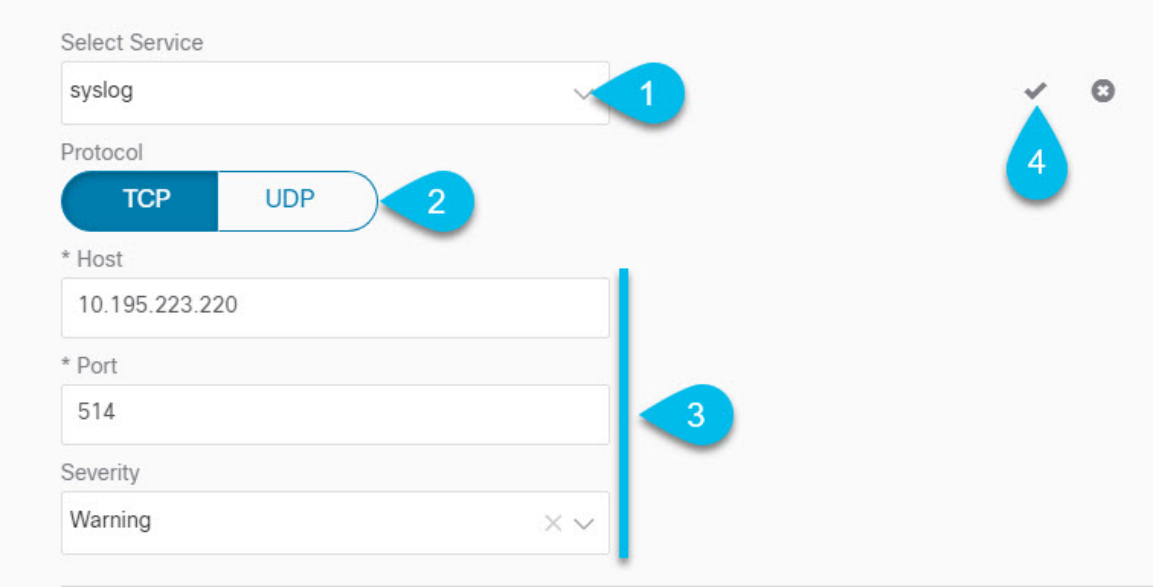
a) Choose `Splunk` for the server type.
b) Choose the protocol.
c) Provide the server name or IP address, port, and the authentication token you obtained from the Splunk service.

   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 checkmark icon to finish adding the server.

Step 5 Add a `syslog` server.

If you do not plan to use `syslog`, skip this step.

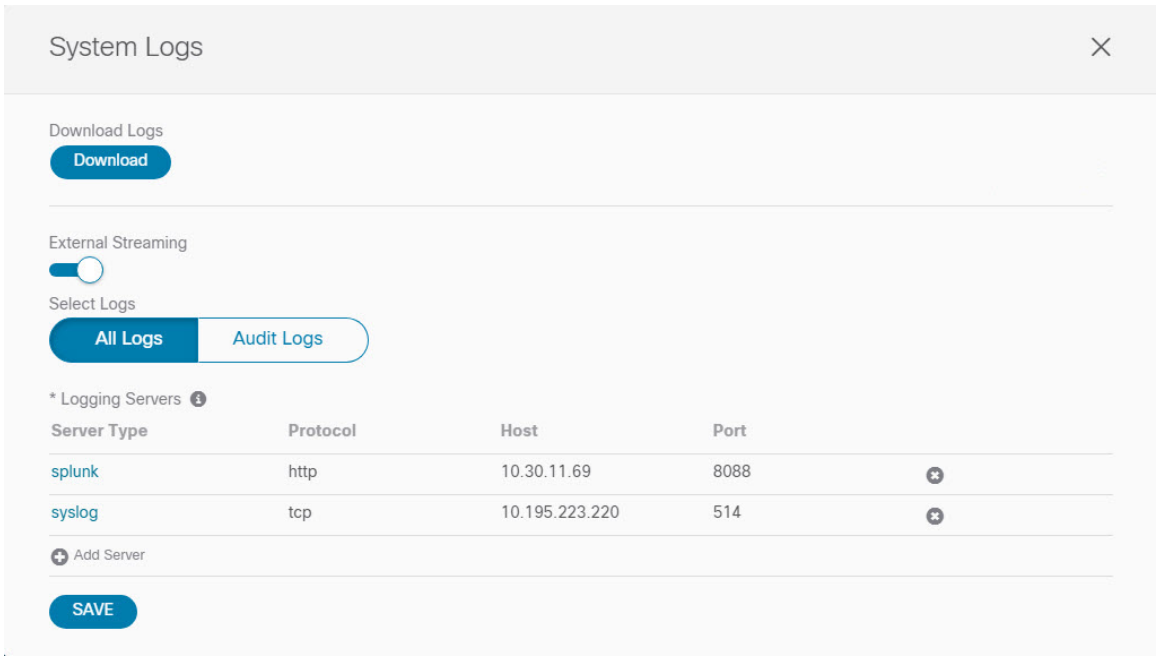


a) Choose `syslog` for the server type.
b) Choose the protocol.
c) Provide the server name or IP address, port number, and the severity level of the log messages to stream.
d) Click the checkmark icon to finish adding the server.

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.

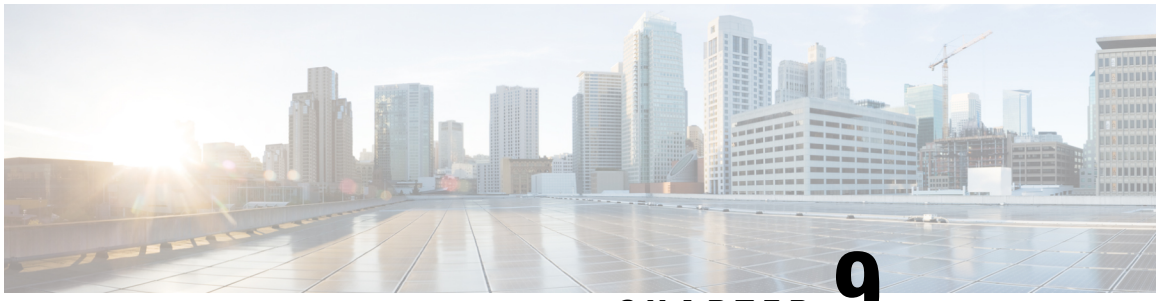

System Logs
Download Logs
Download
External Streaming
Select Logs
All Logs
Audit Logs
* Logging Servers
Server Type
Protocol
Host
Port
splunk
http
10.30.11.69
8088
syslog
tcp
10.195.223.220
514
Add Server
SAVE



PART III

Infrastructure Management

- System Configuration, on page 91
- Upgrading or Downgrading MSO Application, on page 93
- Configuring Cisco ACI Sites, on page 101
- Adding and Deleting Sites, on page 109
- Configuring Infra General Settings, on page 115
- Configuring Infra for Cisco APIC Sites, on page 119
- Configuring Infra for Cisco Cloud APIC Sites, on page 125
- Deploying Infra Configuration for ACI Sites, on page 129
- CloudSec Encryption, on page 135



CHAPTER 9

System Configuration

- System Configuration Settings, on page 91
- System Alias and Banner, on page 91
- Login Attempts and Lockout Time, on page 92

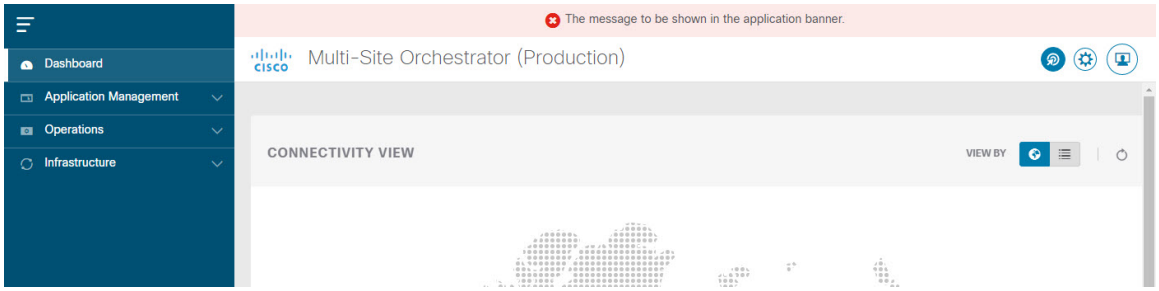
System Configuration Settings

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 Multi-Site 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.

Login Attempts and Lockout Time

When the Orchestrator detects a significant number of failed consecutive login attempts, the user is locked out of the system to prevent unauthorized access. You can configure how failed log in attempts are treated, for example the number of failed attempts before lockout and the length of the lockout.



Note This feature is enabled by default when you first install or upgrade to Release 2.2(1) or later.

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 **Fail Attempts & Lockout Time** area.

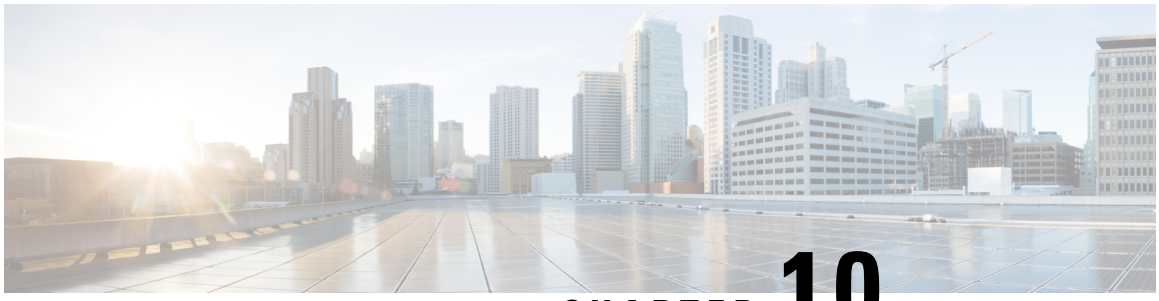
This opens the **Fail Attempts & Lockout Time** settings window.

Step 4 From the **Fail Attempt Settings** dropdown, select the number of attempts before the user is locked out.

Step 5 From the **Lockout Time (Minutes)** dropdown, select the length of the lockout.

This specifies the base lockout duration once it's triggered. The timer is extended up to three times exponentially with every additional consecutive login failure.

Step 6 Click **Save** to save the changes.



CHAPTER 10

Upgrading or Downgrading MSO Application

- Overview, on page 93
- Prerequisites and Guidelines, on page 93
- Upgrading MSO Application Using Cisco App Store, on page 95
- Upgrading MSO Application Manually, on page 97

Overview

The following sections describe how to upgrade or downgrade Cisco Multi-Site Orchestrator, Release 3.2(1) or later that is deployed in Cisco Nexus Dashboard.

If you are running an earlier release deployed in VMware ESX VMs or Cisco Application Services Engine, you must deploy a brand new cluster and then transfer the configuration from your existing cluster, as described in the "Migrating Existing Cluster to Nexus Dashboard" chapter of the *Multi-Site Orchestrator Deployment Guide*.

Prerequisites and Guidelines

Before you upgrade or downgrade your Cisco Multi-Site Orchestrator cluster:

- Stateful upgrades from releases prior to Release 3.2(1) are not supported.

  If you are upgrading from an earlier release, skip the rest of this chapter and follow the instructions described in the "Migrating Existing Cluster to Nexus Dashboard" section of the *Multi-Site Orchestrator Deployment Guide*.

- Ensure that your current Nexus Dashboard cluster is healthy.

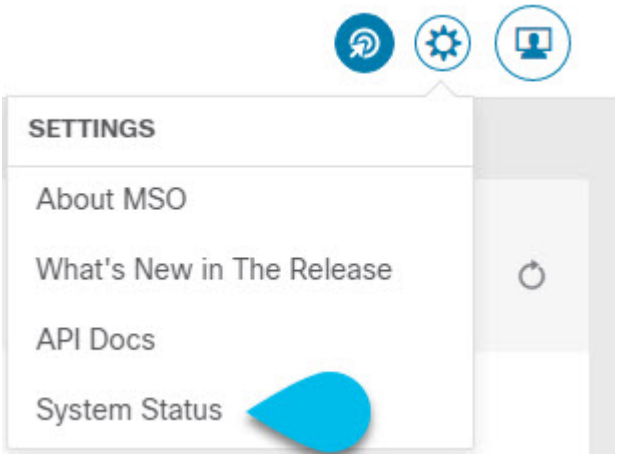
  You can check the Nexus Dashboard cluster health in one of two ways:

  - By logging into your Nexus Dashboard GUI and verifying system status in the **System Overview** page.
  - By logging into any one of the nodes directly as `rescue-user` and running the following command:

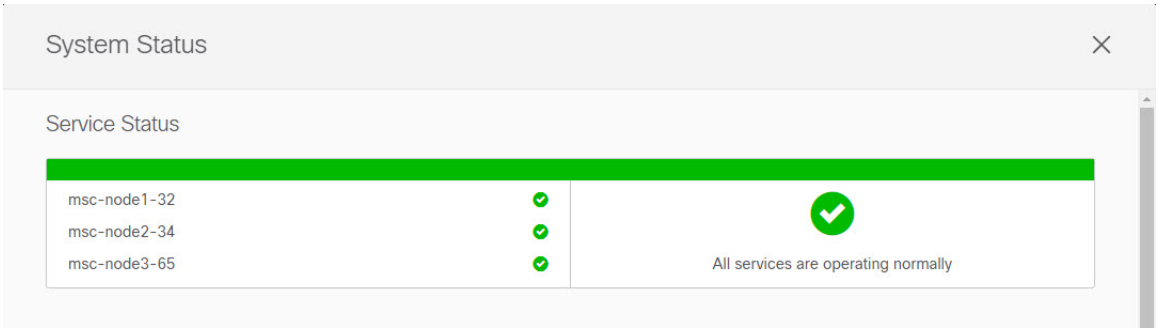
    ```
    # acs health
    All components are healthy
    ```

- Ensure that your current Cisco Multi-Site Orchestrator is running properly.

You can check the status of you Multi-Site Orchestrator application by navigating to **Settings** > **System Status**:



Then ensure that the status of all nodes and services is healthy:



- You can upgrade the MSO application in one of two ways:
  - Using the Nexus Dashboard's App Store, as described in Upgrading MSO Application Using Cisco App Store, on page 95.

    In this case, the Cisco DC App Center must be reachable from the Nexus Dashboard via the Management Network directly or using a proxy configuration. Nexus Dashboard proxy configuration is described in the *Nexus Dashboard User Guide*.

Note The App Store allows you to upgrade to the latest available version of the application only.

  - By manually uploading the new app image, as described in Upgrading MSO Application Manually, on page 97.

    You can use this approach if you are unable to establish the connection to the DC App Center or if you want to upgrade to a version of the application that is not the latest available release.

- If you plan to add and manage new Cloud APIC sites after you upgrade your Multi-Site Orchestrator to release 3.3(1) or later, you must ensure that they are running Cloud APIC release 5.2(1) or later.

On-boarding and managing Cloud APIC sites running earlier releases is not support in Multi-Site Orchestrator 3.3(1).

- Downgrading to releases prior to release 3.3(1) is not supported.

  If you want to downgrade to an earlier release, you must deploy a new Multi-Site Orchestrator cluster on a platform supported by the earlier release, then restore the older configuration backup. Restoring backups created on Release 3.3(1) or later to an older MSO cluster is not supported.

  If you downgrade to an earlier release of Multi-Site Orchestrator, you must also downgrade all Cloud APIC sites to a release prior to Release 5.2(1).

Upgrading MSO Application Using Cisco App Store

This section describes how to upgrade Cisco Multi-Site Orchestrator, Release 3.2(1) or later.

Before you begin

- Ensure that you have completed the prerequisites described in Prerequisites and Guidelines, on page 93.
- Ensure that Cisco DC App Center is reachable from the Nexus Dashboard via the Management Network directly or using a proxy configuration.

  Nexus Dashboard proxy configuration is described in the *Nexus Dashboard User Guide*

Procedure

Step 1 Log in to your Nexus Dashboard..

Step 2 From the left navigation menu, select **Service Catalog**.

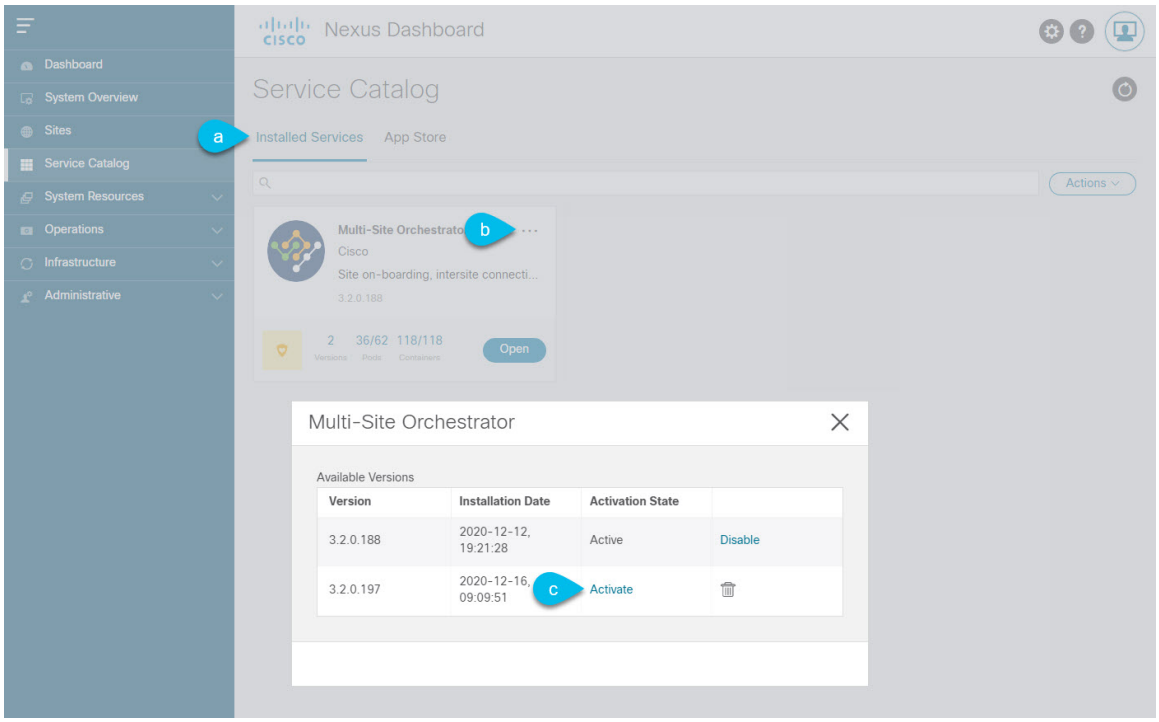
Step 3 Upgrade the application using the App Store.

a) In the **Service Catalog** screen, select the **App Store** tab.
b) In the **Multi-Site Orchestrator** tile, click **Upgrade**.
c) In the License Agreement window that opens, click **Agree and Download**.

Step 4 Wait for the new image to initialize.

It may take up to 20 minutes for the new application image to become available.

Step 5 Activate the new image.



a) In the **Service Catalog** screen, select the **Installed Services** tab.
b) In the top right of the Multi-Site Orchestrator tile, click the menu (. . .) and choose **Available Versions**.
c) In the available versions window, click **Activate** next to the new image.

Note
Do not **Disable** the currently running image before activating the new image. The image activation process will recognize the currently running image and perform the upgrade workflows necessary for the currently running app version.

It may take up to 20 additional minutes for all the application services to start and the GUI to become available. The page will automatically reload when the process is completed.

Step 6 (Optional) Delete the old application image.

You can choose to retain the old application version in case you ever want to downgrade. Or you can delete it as described in this step.

a) In the **Service Catalog** screen, select the **Installed Services** tab.
b) In the top right of the Multi-Site Orchestrator tile, click the menu (. . .) and choose **Available Versions**.
c) In the available versions window, click the delete icon next to the image you want to delete.

It may take up to 20 additional minutes for all the application services to start and the GUI to become available. The page will automatically reload when the process is completed.

Step 7 Launch the app.

To launch the app, simply click **Open** on the application tile in the Nexus Dashboard's **Service Catalog** page.

The single sign-on (SSO) feature allows you to log in to the application using the same credentials as you used for the Nexus Dashboard.

Upgrading MSO Application Manually

This section describes how to upgrade Cisco Multi-Site Orchestrator, Release 3.2(1) or later.

Before you begin

- Ensure that you have completed the prerequisites described in Prerequisites and Guidelines, on page 93.

Procedure

Step 1 Download the target release image.

a) Browse to the Multi-Site Orchestrator page on DC App Center:
https://dcappcenter.cisco.com/nexus-dashboard-orchestrator.html

b) From the **Version** drop-down, choose the version you want to install and click **Download**.

c) Click **Agree and download** to accept the license agreement and download the image.

Step 2 Log in to your Nexus Dashboard..

Step 3 Upload the image to your Nexus Dashboard.

a) From the left navigation menu, select **Service Catalog**.

b) In the Nexus Dashboard's **Service Catalog** screen, select the **Installed Services** tab.

c) From the **Actions** menu in the top right of main pane, select **Upload App**.

d) In the **Upload App** window, choose the location of the image

If you downloaded the application image to your system, choose **Local**.

If you are hosting the image on a server, choose **Remote**.

e) Choose the file.

If you chose **Local** in the previous substep, click **Select File** and select the app image you downloaded.

If you chose **Remote**, provide the full URL to the image file, for example `http://<ip-address>:<port>/<full-path>/cisco-mso-<version>.aci`.

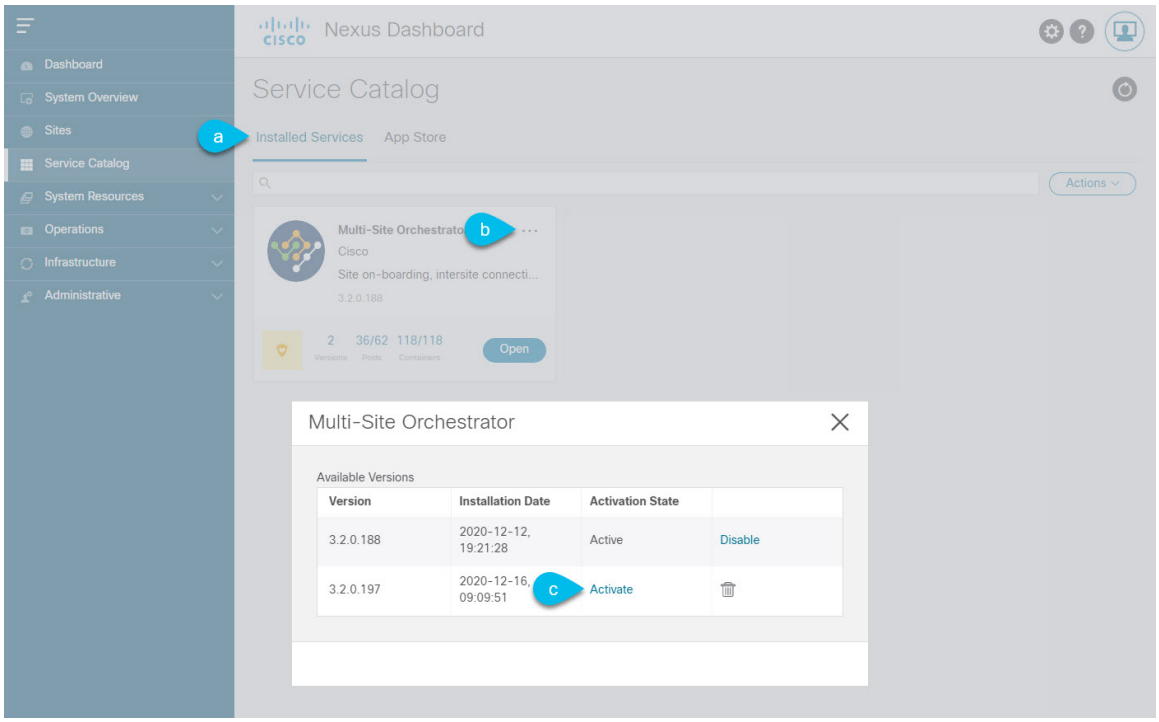
f) Click **Upload** to add the app to the cluster.

A new tile will appear with the upload progress bar. Once the image upload is completed, the Nexus Dashboard will recognize the new image as an existing application and add it as a new version.

Step 4 Wait for the new image to initialize.

It may take up to 20 minutes for the new application image to become available.

Step 5 Activate the new image.



a) In the **Service Catalog** screen, select the **Installed Services** tab.
b) In the top right of the Multi-Site Orchestrator tile, click the menu (. . .) and choose **Available Versions**.
c) In the available versions window, click **Activate** next to the new image.

Note
Do not **Disable** the currently running image before activating the new image. The image activation process will recognize the currently running image and perform the upgrade workflows necessary for the currently running app version.

It may take up to 20 additional minutes for all the application services to start and the GUI to become available. The page will automatically reload when the process is completed.

Step 6 (Optional) Delete the old application image.

You can choose to retain the old application version in case you ever want to downgrade. Or you can delete it as described in this step.

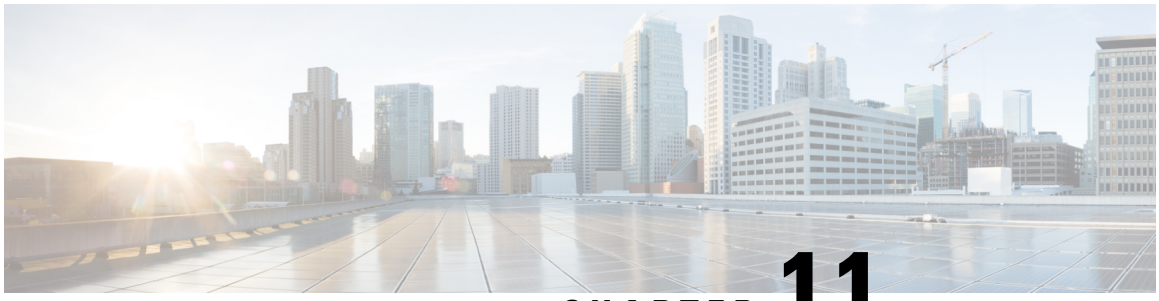
a) In the **Service Catalog** screen, select the **Installed Services** tab.
b) In the top right of the Multi-Site Orchestrator tile, click the menu (. . .) and choose **Available Versions**.
c) In the available versions window, click the delete icon next to the image you want to delete.

It may take up to 20 additional minutes for all the application services to start and the GUI to become available. The page will automatically reload when the process is completed.

Step 7 Launch the app.

To launch the app, simply click **Open** on the application tile in the Nexus Dashboard's **Service Catalog** page.

The single sign-on (SSO) feature allows you to log in to the application using the same credentials as you used for the Nexus Dashboard.



CHAPTER 11

Configuring Cisco ACI Sites

- Pod Profile and Policy Group, on page 101
- Configuring Fabric Access Policies for All APIC Sites, on page 102
- Configuring Sites That Contain Remote Leaf Switches, on page 105
- Cisco Mini ACI Fabrics, on page 106

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 Multi-Site 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 Multi-Site 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 Multi-Site 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 Multi-Site 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-l3`.
- 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 103.

Configuring Fabric Access Interface Policies

This section describes the fabric access interface configurations that must be done for the Multi-Site 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 102.

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

Starting with Release 2.1(2), the Multi-Site architecture supports APIC sites with Remote Leaf switches. The following sections describe guidelines, limitations, and configuration steps required to allow Multi-Site 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 Multi-Site 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 Multi-Site 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 Multi-Site 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 Multi-Site 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.

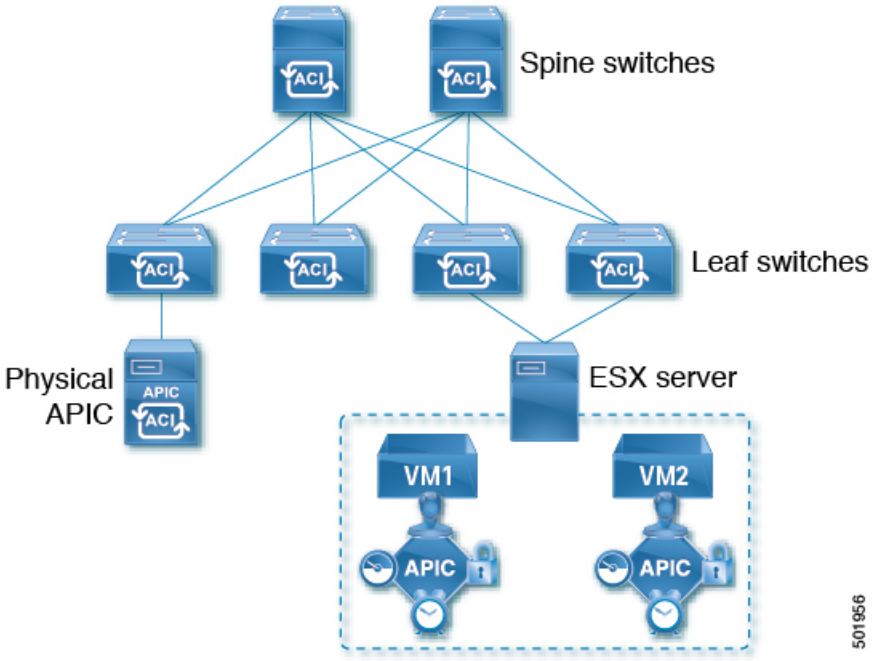
Cisco Mini ACI Fabrics

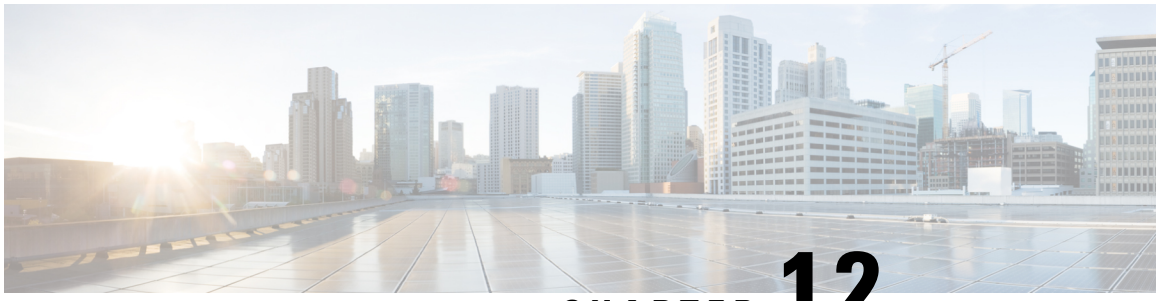
Cisco Multi-Site supports Cisco Mini ACI fabrics as typical on-premises sites without requiring any additional configuration. This section provides a brief overview of Mini ACI fabrics, detailed info on deploying and configuring this type of fabrics is available in *Cisco Mini ACI Fabric and Virtual 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





CHAPTER 12

Adding and Deleting Sites

- Cisco MSO and APIC Interoperability Support, on page 109
- Adding Cisco ACI Sites, on page 111
- Removing Sites, on page 113
- Cross Launch to Fabric Controllers, on page 114

Cisco MSO and APIC Interoperability Support

Cisco Multi-Site Orchestrator (MSO) does not require a specific version of APIC to be running in all sites. The APIC clusters in each site as well as the MSO 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 Multi-Site Orchestrator application is installed. As such, we recommend that you always upgrade to the latest release of the Multi-Site Orchestrator.

However, keep in mind that if you upgrade the MSO before upgrading the APIC clusters in one or more sites, some of the new MSO 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 MSO. 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 Multi-Site 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 APIC site using the Nexus Dashboard GUI and then enable that site to be managed by Multi-Site Orchestrator.

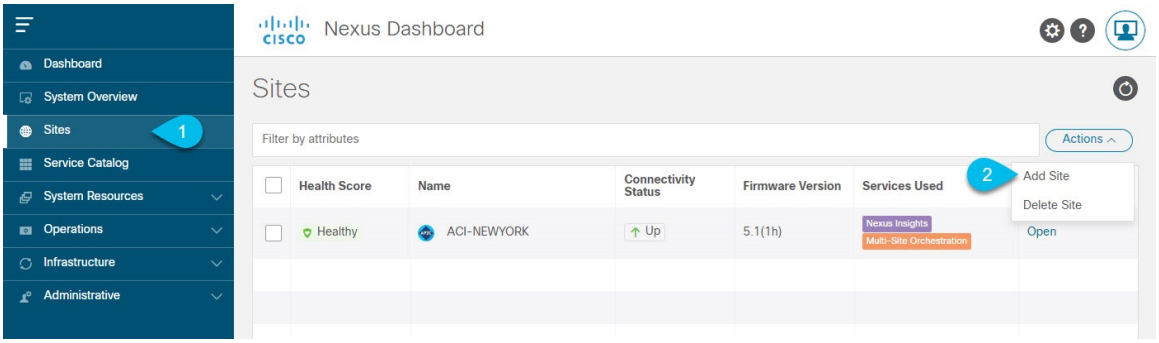
Before you begin

- 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 the site(s) you are adding are running Release 4.2(4) or later.

Procedure

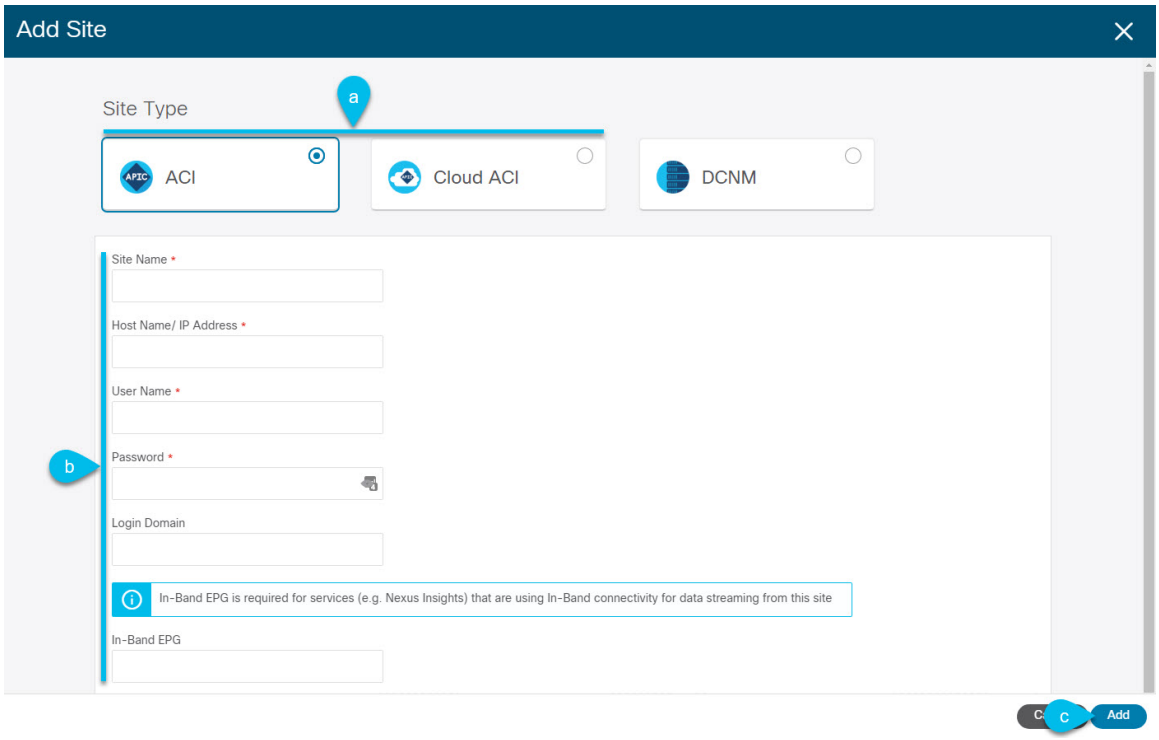
Step 1 Log in to the Nexus Dashboard GUI

Step 2 Add a new site.



a) From the left navigation menu, select **Sites**.
b) In the top right of the main pane, select **Actions** > **Add Site**.

Step 3 Provide site information.



a) For **Site Type**, select **ACI** or **Cloud ACI** depending on the type of ACI fabric you are adding.

b) Provide the controller information.

You need to provide the **Host Name/IP Address**, **User Name**, and **Password.** for the APIC controller currently managing your ACI fabrics.

For on-premises ACI sites managed by Cisco APIC, if you plan to use this site with Day-2 Operations applications such as Nexus Insights, you must also provide the **In-Band EPG** name used to connect the Nexus Dashboard to the fabric you are adding. Otherwise, if you will use this site with Multi-Site Orchestrator only, you can leave this field blank.

c) Click **Add** to finish adding the site.

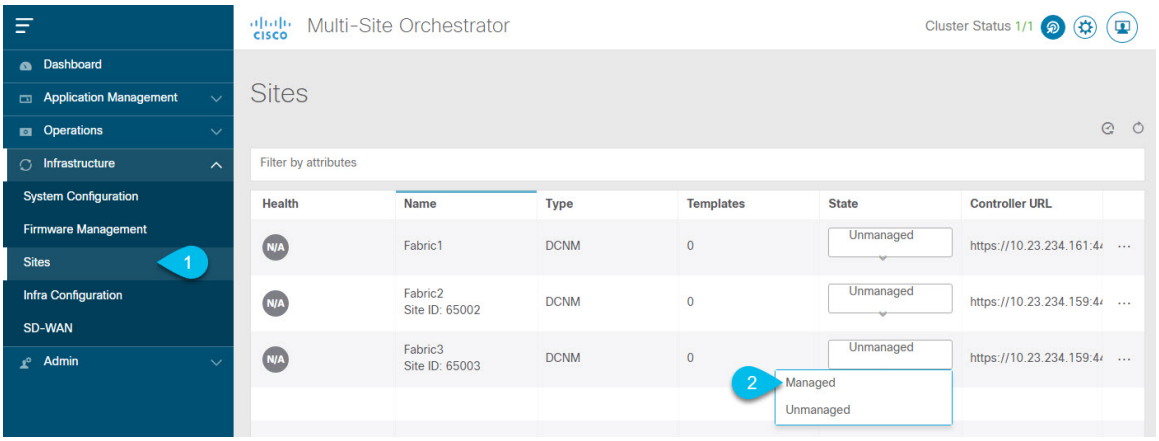
At this time, the sites will be available in the Nexus Dashboard, but you still need to enable them for Multi-Site Orchestrator management as described in the following steps.

Step 4 Repeat the previous steps for any additional ACI sites.

Step 5 From the Nexus Dashboard's **Service Catalog**, open the Multi-Site Orchestrator application.

You will be automatically logged in using the Nexus Dashboard user's credentials.

Step 6 In the Multi-Site Orchestrator GUI, manage the sites.



a) From the left navigation menu, select **Infrastructure** > **Sites**.
b) In the main pane, change the **State** from `Unmanaged` to `Managed` for each fabric that you want the MSO to manage.

Removing Sites

This section describes how to disable site management for one or more sites using the Multi-Site Orchestrator GUI. The sites will remain present in the 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 Multi-Site Orchestrator GUI.

You can open the MSO application from the Nexus Dashboard's **Service Catalog**. You will be automatically logged in using the Nexus Dashboard user's credentials.

Step 2 Remove the site's underlay configuration.

a) From the left navigation menu, select **Infrastructure** > **Infra Configuration**.
b) In the main pane, click **Configure Infra**.
c) In the left sidebar, select the site you want to unmanage.
d) In right bar, under **Overlay Configuration** tab, disable the **ACI Multi-Site** knob.
e) In the right sidebar, select the **Underlay Configuration** tab.
f) Remove all underlay configurations from the site.
g) Click **Deploy** to deploy underlay and overlay configuration changes to the site.

Step 3 In the Multi-Site Orchestrator GUI, disable the sites.

a) From the left navigation menu, select **Infrastructure** > **Sites**.
b) In the main pane, change the **State** from `Managed` to `Unmanaged` for each fabric that you want the MSO to stop managing.

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.

Step 4 Delete the site from 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 Nexus Dashboard as well.

Note
Note that the site must not be currently in use by any of the applications installed in your Nexus Dashboard cluster.

a) From the left navigation menu of the Nexus Dashboard GUI, select **Sites**.
b) Select one or more sites you want to delete.
c) In the top right of the main pane, select **Actions** > **Delete Site**.
d) Provide the site's login information and click **OK**.

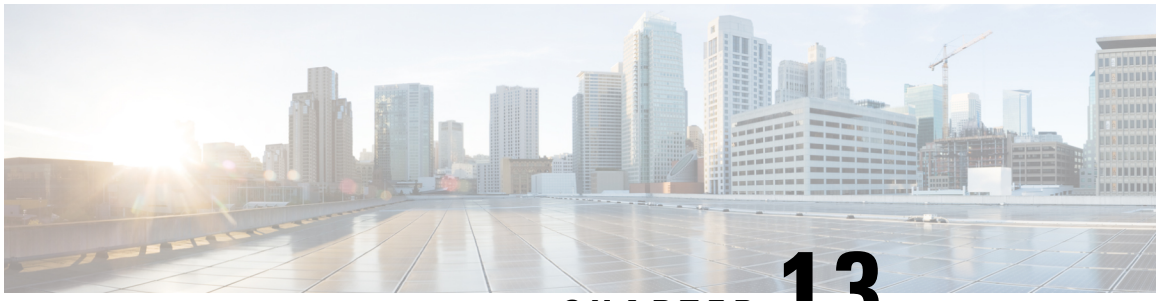
The site will be removed from the Nexus Dashboard.

Cross Launch to Fabric Controllers

Multi-Site Orchestrator currently supports a number of configuration options for each type of fabrics. For many additional configuration options, you may need to log in directly into the fabric's controller.

You can cross launch into the specific site controller's GUI from the MSO's **Infrastucture** > **Sites** screen by selecting the actions (. . .) menu next to the site and clicking **Open in user interface**. Note that cross-launch works with out-of-band (OOB) management IP of the fabric.

If your Nexus Dashboard and the fabric are configured for remote user authentication, you will be logged in automatically into the fabric's controller using the same log in information as the Nexus Dashboard user.



CHAPTER 13

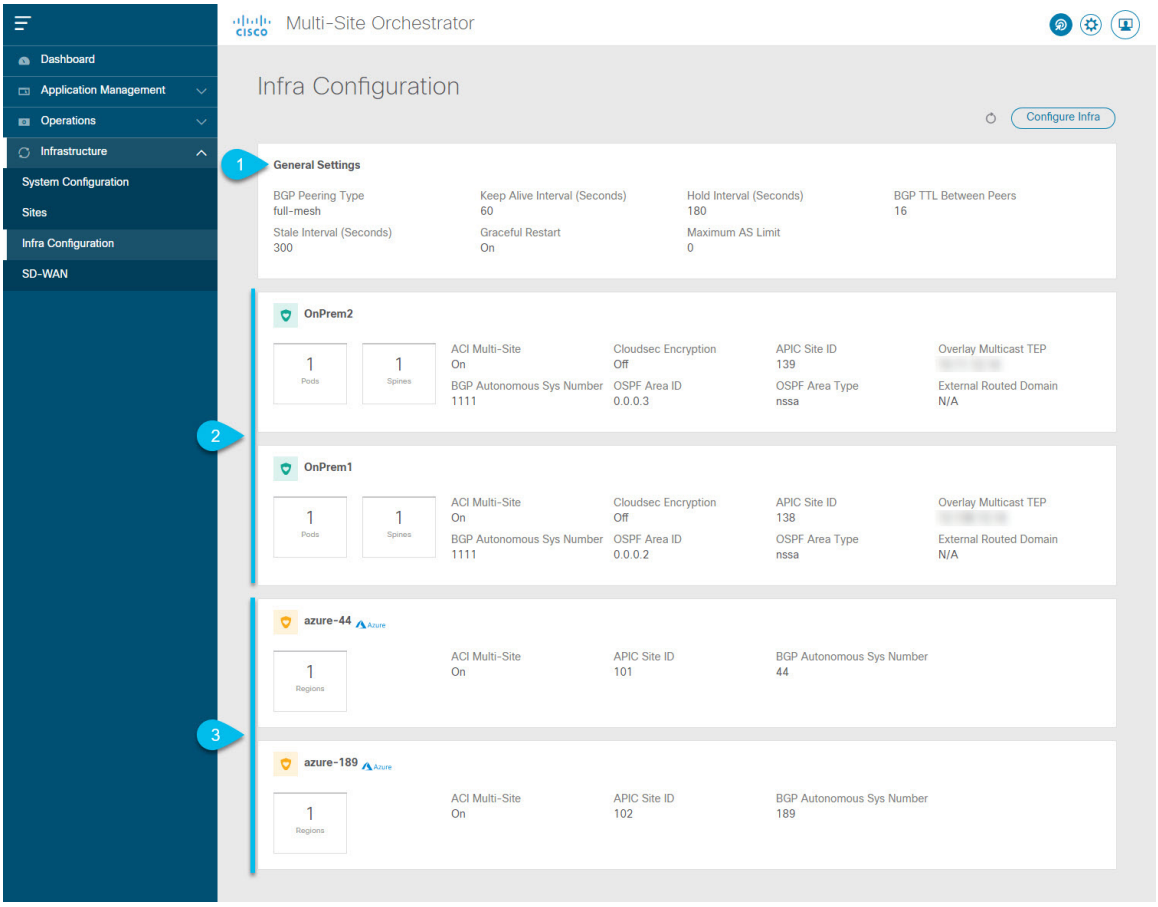
Configuring Infra General Settings

- Infra Configuration Dashboard, on page 115
- Configuring Infra: General Settings, on page 117

Infra Configuration Dashboard

The **Infra Configuration** page displays an overview of all sites and inter-site connectivity in your Multi-Site Orchestrator deployment and contains the following information:

Figure 15: 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 on 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 119.

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 APIC Sites, on page 125.

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 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 Multi-Site Orchestrator fabric connectivity information refresh described in the Refreshing Site Connectivity Information, on page 119 as part of the general Infra configuration procedures.

Configuring Infra: General Settings

This section describes how to configure general Infra settings for all the sites.

Procedure

Step 1 Log in to the Cisco Multi-Site Orchestrator GUI.

Step 2 In the left navigation menu, select **Infrastructure** > **Infra Configuration**.

Step 3 In the main pane, click **Configure Infra**.

Step 4 In the left sidebar, select **General Settings**.

Step 5 Configure **Control Plane BGP**.

a) Select the **Control Plane BGP** tab.

b) Choose **BGP Peering Type**.

- `full-mesh`—All border gateway switches in each site will establish peer connectivity with remote sites' border gateway switches.

  In `full-mesh` configuration, Multi-Site Orchestrator uses the spine switches for ACI managed fabrics and border gateways for DCNM 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 managed by MSO.

  For ACI fabrics, the `route-reflector` option is effective only for fabrics that are part of the same BGP ASN.

c) In the **Keepalive Interval (Seconds)** field, enter the keep alive interval seconds.

We recommend keeping the default value.

d) In the **Hold Interval (Seconds)** field, enter the hold interval seconds.

We recommend keeping the default value.

e) In the **Stale Interval (Seconds)** field, enter stale interval seconds.

We recommend keeping the default value.

f) Choose whether you want to turn on the **Graceful Helper** option.

g) Provide the **Maximum AS Limit**.

We recommend keeping the default value.

h) Provide the **BGP TTL Between Peers**.

We recommend keeping the default value.

The following settings are for cloud sites' inter-site connectivity:

a) Provide the **OSPF Area ID** field.

This is OSPF area ID used by cloud sites for on-premises ISN peering, which you previously configured in the Cloud APIC for inter-site connectivity in earlier Multi-Site Orchestrator releases.

b) 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 used for on-premises connectivity, which you previously configured in the Cloud APIC for inter-site connectivity in earlier Multi-Site 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`.

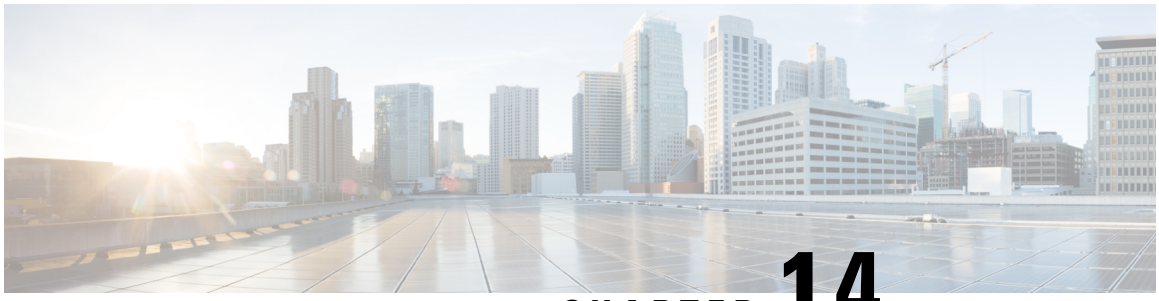
Step 6 Provide the **IPN Devices** information.

When you configure inter-site underlay connectivity between on-premises and cloud sites as described in later sections, you will need to select an on-premises IPN device which will establish 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 120.

a) Select the **IPN Devices** tab.
b) Click **Add IPN Device**.
c) Provide the **Name** and the **IP Address** of the IPN device.

The IP address you provide will be used as the tunnel peer address from the Cloud APIC's CSRs, not the IPN device's management IP address.

d) Click the check mark icon to save the device information.
e) Repeat this step for any additional IPN devices you want to add.



CHAPTER 14

Configuring Infra for Cisco APIC Sites

- Refreshing Site Connectivity Information, on page 119
- Configuring Infra: On-Premises Site Settings, on page 120
- Configuring Infra: Pod Settings, on page 122
- Configuring Infra: Spine Switches, on page 122

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 Multi-Site Orchestrator GUI.

Step 2 In the **Main menu**, select **Infrastructure** > **Infra Configuration**.

Step 3 In the top right of the main **Infra Configuration** view, click the **Configure Infra** button.

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 checkbox, 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 will discover any new or removed spines and all site-related fabric connectivity will be re-imported 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 Multi-Site Orchestrator GUI.

Step 2 In the left navigation menu, select **Infrastructure** > **Infra Configuration**.

Step 3 In the main pane, click **Configure Infra**.

Step 4 In the left pane, under **Sites**, select a specific on-premises site.

Step 5 Provide the **Overlay Configuration**.

a) In the right **<*Site*> Settings** pane, select the **Overlay Configuration** tab.

b) 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.

c) (Optional) Enable the **CloudSec Encryption** knob encryption for the site.

CloudSec Encryption provides inter-site traffic encryption. The "Infrastructure Management" chapter in the *Cisco Multi-Site Configuration Guide* covers this feature in detail.

d) Specify the **Overlay Multicast TEP**.

This address is used for the inter-site 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 Multi-Pod 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.

e) (Optional) From the **External Routed Domain** dropdown, select the domain 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.

f) Specify the **BGP Autonomous System Number**.

g) (Optional) Specify the **BGP Password**.

h) (Optional) Enable **SR-MPLS Connectivity** for the site.

If the site is connected via 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 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 will need to configure additional settings as described in the "Sites Connected via SR-MPLS" chapter of the *Cisco Multi-Site Configuration Guide for ACI Fabrics*.

Step 6 Provide the **Underlay Configuration**.

a) In the right **<*Site*> Settings** pane, select the **Underlay Configuration** tab.

b) Select the **OSPF Area Type** from the dropdown menu.

The OSPF area type can be one of the following:

- `nssa`
- `regular`

c) Configure OSPF settings for the site.

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** dropdown menu, 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`.

Step 7 Configure inter-site connectivity between on-premises and cloud sites.

If you do not need to create inter-site 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 need to provide an IPN device IP address to which the Cloud APIC'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 dropdown, 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 117

c) Configure inter-site 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 APIC Sites, on page 125.

What to do next

While you have configured all the required inter-site connectivity information, it has not been pushed to the sites yet. You need to deploy the configuration as described in Deploying Infra Configuration, on page 129

Configuring Infra: Pod Settings

This section describes how to configure pod-specific settings in each site.

Procedure

Step 1 Log in to the Cisco Multi-Site Orchestrator GUI.

Step 2 In the **Main menu**, click **Sites**.

Step 3 In the **Sites** view, click **Configure Infra**.

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 intersite known unicast traffic.

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 ISN to APIC nodes, spine switches, and border leaf nodes. This is required to enable the intersite L3Out functionality.

External TEP pools previously assigned to the fabric on APIC are automatically inherited by MSO 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.

Procedure

Step 1 Log in to the Cisco Multi-Site Orchestrator GUI.

Step 2 In the **Main menu**, click **Sites**.

Step 3 In the **Sites** view, click **Configure Infra**.

Step 4 In the left pane, under **Sites**, select a specific site.

Step 5 In the main window, 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, enter the following information:

- In the **Ethernet Port ID** field, enter the port ID, for example `1/29`.
- In the **IP Address** field, enter the IP address/netmask.

  MSO creates a sub-interface 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.
- In the **OSPF Policy** field, choose the OSPF policy for the switch that you have configured in Configuring Infra: On-Premises Site Settings, on page 120.

  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`

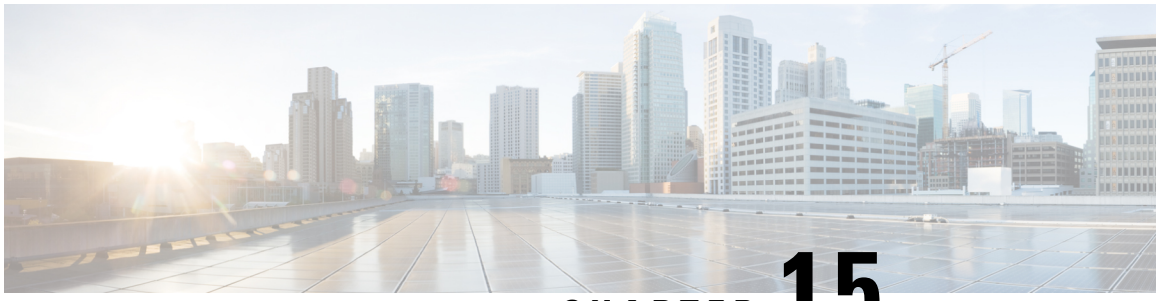
Step 8 Enable **BGP Peering** knob.

In a single Pod fabric with more than two spine switches, BGP peering should only be enabled on a pair (for redundancy) of spine switches called **BGP Speakers**. All other spine switches should have BGP peering disabled and will function as **BGP Forwarders**.

In a Multi-Pod fabric BGP peering should only be enabled on a couple of BGP speaker spine switches, each deployed in a different Pod. All other spines switches should have BGP peering disabled and function as BGP forwarders.

Step 9 In the **BGP-EVPN Router-ID** field, provide the IP address used for BGP-eVPN session between sites.

Step 10 Repeat the procedure for every spine switch.



CHAPTER 15

Configuring Infra for Cisco Cloud APIC Sites

- Refreshing Cloud Site Connectivity Information, on page 125
- Configuring Infra: Cloud Site Settings, on page 125

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 Multi-Site Orchestrator GUI.

Step 2 In the **Main menu**, select **Infrastructure** > **Infra Configuration**.

Step 3 In the top right of the main **Infra Configuration** view, click the **Configure Infra** button.

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 will discover 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 need to 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 APIC sites.

Procedure

Step 1 Log in to the Cisco Multi-Site Orchestrator GUI.

Step 2 In the **Main menu**, select **Infrastructure** > **Infra Configuration**.

Step 3 In the top right of the main pane, click **Configure Infra**.

Step 4 In the left pane, under **Sites**, select a specific cloud site.

Step 5 Provide the **Overlay Configuration**.

a) In the right **<*Site*> Settings** pane, select the **Overlay Configuration** tab.

b) 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.

c) (Optional) Specify the **BGP Password**.

Step 6 Provide the **Underlay Configuration**.

a) In the right **<*Site*> Settings** pane, select the **Underlay Configuration** tab.

b) Click **Add Connectivity**.

c) From the **Site** dropdown, select the site to which you want to establish connectivity.

d) From the **Connection Type** dropdown, choose the type of connection between the sites.

The following options are available:

- `Public Internet`—connectivity between the two sites is established via 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.

e) (Optional) Enable **IPsec**.

The following options are available:

- 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.

f) If IPsec is enabled, choose the **IKE Version** for it.

Internet Key Exchange (IKE) is a protocol used to establish security association for IPsec. You can choose which version of the protocol to use: IKEv1 (`Version 1`) or IKEv2 (`Version 1`) depending on your configuration.

g) Click **Save** to save the inter-site 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 **Underlay Configuration** tab.

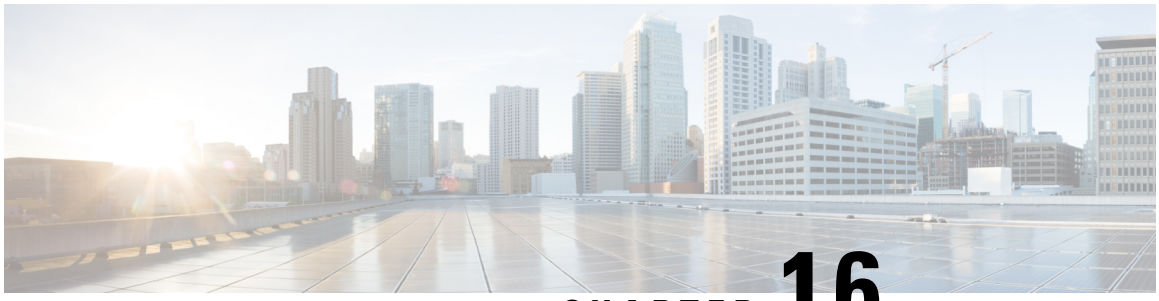
h) Repeat this step to add inter-site connectivity for other sites.

When you establish underlay connectivity from `Site1` to `Site2`, the reverse connetivity is done automatically for you.

However, if you also want to establish inter-site connectivity from `Site1` to `Site3`, you must repeat this step for that site as well.

What to do next

While you have configured all the required inter-site connectivity information, it has not been pushed to the sites yet. You need to deploy the configuration as described in Deploying Infra Configuration, on page 129



CHAPTER 16

Deploying Infra Configuration for ACI Sites

- Deploying Infra Configuration, on page 129
- Enabling Connectivity Between On-Premises and Cloud Sites, on page 130

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 two additional options become available:

- **Deploy & Download IPN Device config files:** Pushes the configuration to both the on-premises APIC site and the Cloud APIC 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 between the Cisco Cloud Services Router (CSR) deployed in your cloud sites and the on-premises IPsec termination device. 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 between the 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 descrbied in Refreshing Site Connectivity Information, on page 119.

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 APIC sites.

By default, the Cisco Cloud APIC 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 Multi-Site Orchestrator as part of the procedures provided in Deploying Infra Configuration, on page 129.

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 Multi-Site 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 Multi-Site Orchestrator, Release 3.3(1) and Cloud APIC, Release 5.2(1). If you are using IKEv1, the IPN configuration file you downloaded form MSO 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 144904725321902286651389219409672714611O
    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

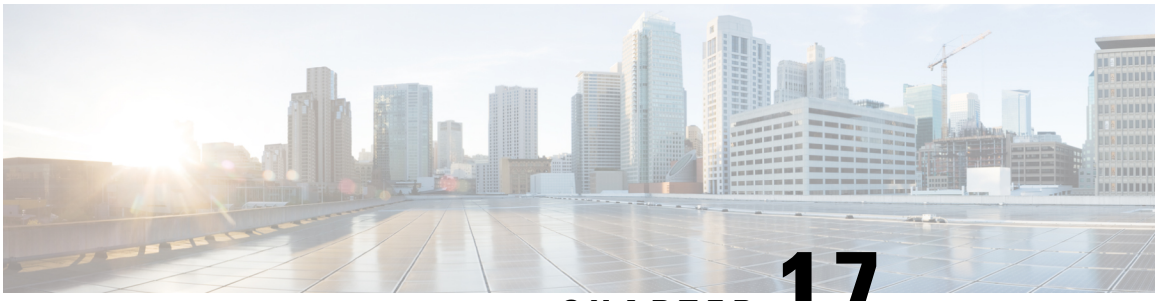
interface tunnel 2001
    ip address 5.5.1.26 255.255.255.252
    ip virtual-reassembly
    ! Please change GigabitEthernet1 to the appropriate interface
    tunnel source GigabitEthernet1
    tunnel destination 52.12.232.0
    tunnel mode ipsec ipv4
    tunnel protection ipsec profile infra:overlay-1-2001
    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 17

CloudSec Encryption

- Cisco ACI CloudSec Encryption, on page 135
- Requirements and Guidelines, on page 136
- CloudSec Encryption Terminology, on page 138
- CloudSec Encryption and Decryption Handling, on page 139
- CloudSec Encryption Key Allocation and Distribution, on page 140
- Configuring Cisco APIC for CloudSec Encryption, on page 142
- Enabling CloudSec Encryption Using Multi-Site Orchestrator GUI, on page 145
- Rekey Process During Spine Switch Maintenance, on page 146

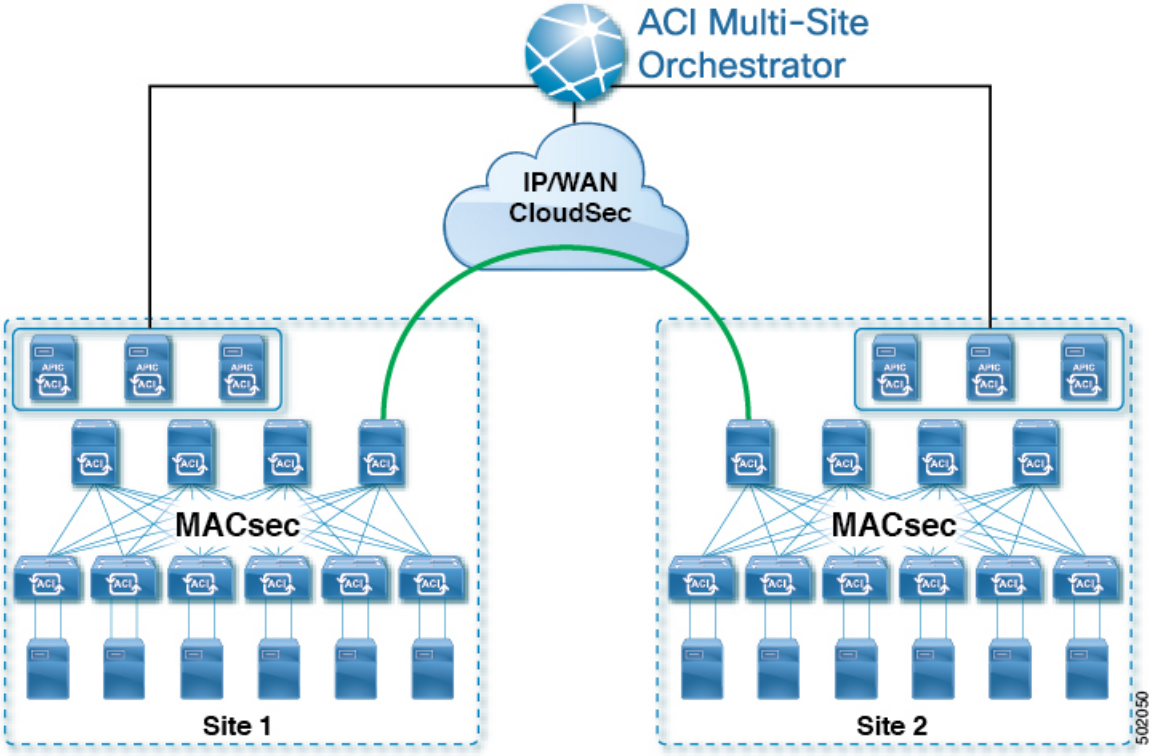
Cisco ACI CloudSec Encryption

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. Multi-Site 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 to provide connectivity between sites. These TEP addresses are configured by the admin on Multi-Site 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 16: 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, 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.

     ```
     ASR1K(config)# platform cloudsec padding-fixup dst-udp-port 9999
     ```

  3. Verify the configuration.

     ```
     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.

- The CloudSec Encryption feature is not supported with the following features:
  - Precision Time Protocol (PTP)
  - Remote Leaf Direct
  - Virtual Pod (vPOD)
  - SDA
  - Intersite L3Out
  - Other routable TEP configurations

Requirements

The CloudSec encryption capability requires the following:

- Cisco ACI spine-leaf architecture with a Cisco APIC cluster for each site
- Cisco ACI Multi-Site 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.

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 datacenter fabric that originates the encrypted VXLAN packets.
- `Downstream site` – The datacenter 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.

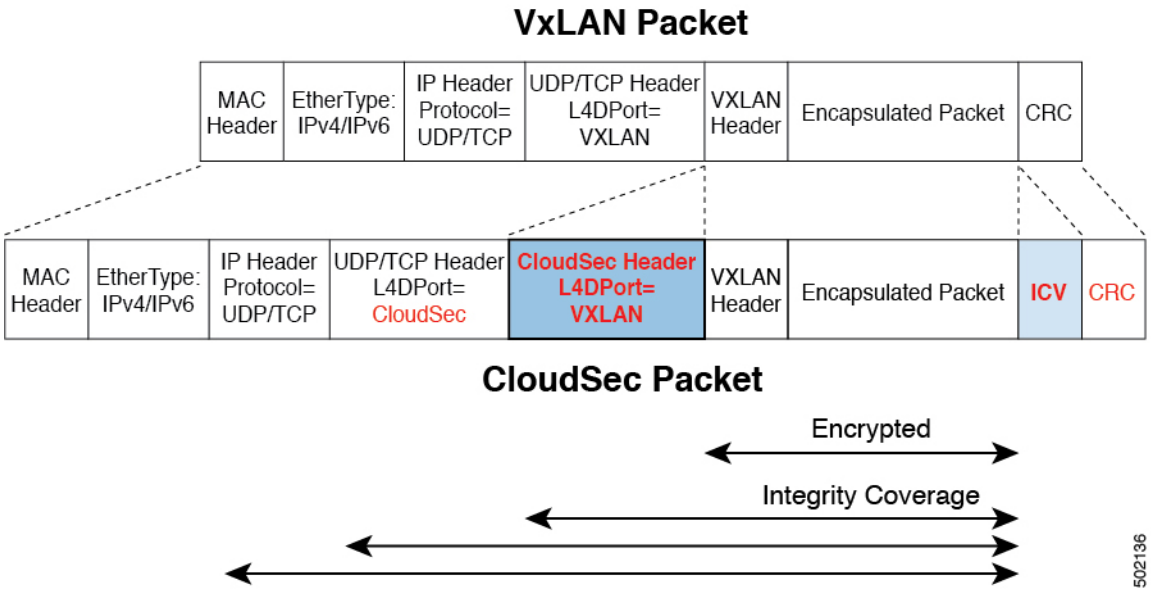
  In ACI only two association number values (`0` and `1`) are used for the two active RX keys and only one association number value (`0` or `1`) is used for the TX Key at any point in time.
- `Pre-shared key (PSK)` – One ore more keys must be configured in the Cisco APIC GUI to be used as a random seed 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 17: CloudSec Packet



Packet Encryption

The following is a high level overview of how CloudSec handles outgoing traffic packets:

- The packets are filtered using the outer IP header and Layer-4 destination port information and matching 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 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 with a Cisco proprietary Layer-4 port number (Port `9999`) indicating that it is a CloudSec packet.
- 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:

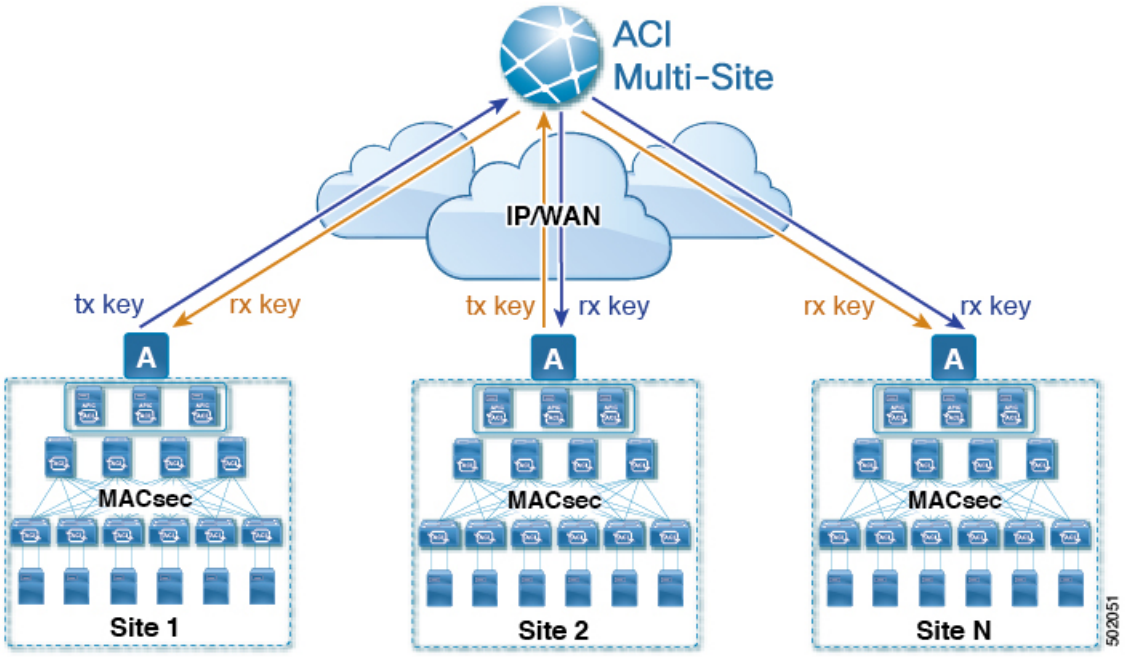
- If the received packet is a CloudSec packet, it is decrypted and the ICV is verified.

  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.
- If the key store returns two or more possible decryption keys, the Association Number (AN) field of the CloudSec header is used to select which key to use.
- If the packet is not a CloudSec packet, the packet is left unchanged.

CloudSec Encryption Key Allocation and Distribution

Initial Key Configuration

Figure 18: 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 Multi-Site Orchestrator (MSO) by the upstream site's Cisco APIC for distribution to downstream remote sites.
- The MSO 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.
- 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 MSO.
- The MSO 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 MSO. 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 146 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 Multi-Site Orchestrator about each key's status for distribution to other sites.

Multi-Site Orchestrator's Role in Key Management

The Multi-Site 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 MSO 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 143
- Using the Cisco APIC NX-OS Style CLI, as described in Configuring Cisco APIC for CloudSec Encryption Using NX-OS Style CLI, on page 143
- Using the Cisco APIC REST API, as described in Configuring Cisco APIC for CloudSec Encryption Using REST API, on page 144

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:

```
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 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:

```
apic1(config-cloudsec)# sakexpirytime <duration>
```

Step 5 Specify one or more Pre-Shared Keys.

In the following command, specify the index of the PSK you're configuring and the PSK string itself.

Example:

```
apic1(config-cloudsec)# pskindex <psk-index>
apic1(config-cloudsec)# pskstring <psk-string>
```

The <*psk-index*> parameter 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.

The <*psk-string*> parameter specifies the actual PSK, which must be a hexadecimal string 64 characters long.

Step 6 (Optional) View the current PSK configuration.

You can view how many PSK are currently configured and their duration using the following command:

Example:

```
apic1(config-cloudsec)# show cloudsec summary
```

Configuring Cisco APIC for CloudSec Encryption Using REST API

This section describes how to configure one or more pre-shared keys (PSK) using the Cisco APIC REST API.

Procedure

Configure PSK expiration time, index, and string.

In the following XML POST, replace:

- The value of **sakExpiryTime** with the expiration time of each PSK.

  This **sakExpiryTime** parameter 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.

- The value of **index** with the index of the PSK you're configuring.

  The **index** parameter 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.

- The value of **pskString** with the index of the PSK you're configuring.

  The **pskString** parameter specifies the actual PSK, which must be a hexadecimal string 64 characters long.

Example:

```
<fvTenant annotation="" descr="" dn="uni/tn-infra" name="infra" nameAlias="" ownerKey="" ownerTag="">

   <cloudsecIfPol descr="cloudsecifp" name="default" sakExpiryTime="10" stopRekey= "false" status=""
 >
        <cloudsecPreSharedKey index="1"
pskString="1234567812345678123456781234567812345678123456781234567812345678" status=""/>
    </cloudsecIfPol>
</fvTenant>
```

Enabling CloudSec Encryption Using Multi-Site Orchestrator GUI

The CloudSec encryption can be enabled or disabled for each site individually. However, the communications between two sites will be 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 Multi-Site Orchestrator, as described in *Cisco Multi-Site Orchestrator Installation and Upgrade Guide*.
- Added each Cisco APIC site to the Multi-Site Orchestrator, as described in *Cisco Multi-Site Configuration Guide*.

Procedure

Step 1 Log in to the Multi-Site Orchestrator.

Step 2 From the left-hand sidebar, select the **Sites** view.

Step 3 Click on the **Configure Infra** button in the top right of the main window.

Step 4 From the left-hand sidebar, select the site for which you want to change the CloudSec configuration.

Step 5 In the right-hand sidebar, toggle the **CloudSec Encryption** setting to enable or disable the CloudSec Encryption feature for the site.

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 146. 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 146. 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="">

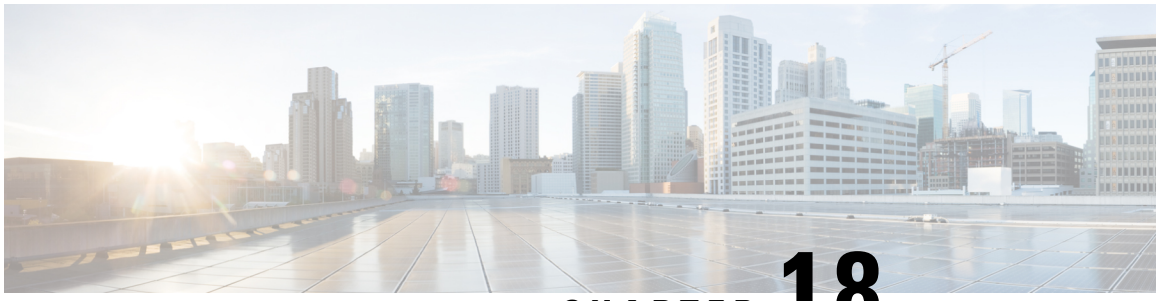
   <cloudsecIfPol descr="cloudsecifp" name="default" sakExpiryTime="10" stopRekey= "false" status=""
 />
</fvTenant>
```



PART IV

Features and Use Cases

- DHCP Relay, on page 151
- Intersite L3Out, on page 159
- Intersite L3Out with PBR, on page 179
- Layer 3 Multicast, on page 205
- EPG Preferred Group, on page 215
- QoS Preservation Across IPN, on page 219
- SD-WAN Integration, on page 225
- Sites Connected via SR-MPLS, on page 233
- vzAny Contracts, on page 251



CHAPTER 18

DHCP Relay

- DHCP Relay Policy, on page 151
- Guidelines and Limitations, on page 151
- Creating DHCP Relay Policies, on page 152
- Creating DHCP Option Policies, on page 154
- Assigning DHCP Policies, on page 155
- Creating DHCP Relay Contract , on page 156
- Verifying DHCP Relay Policies in APIC, on page 157
- Editing or Deleting Existing DHCP Policies, on page 157

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 will need to re-deploy 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 Multi-Site 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 Multi-Site Orchestrator GUI.

Step 2 From the left navigation menu, select **Application Management** > **Policies**.

Step 3 In the top right of the main pane, select **Add Policy** > > **Creating DHCP Policy**.

This opens an **Add DHCP** configuration screen.

Step 4 In the **Name** field, specify the name for the policy.

Step 5 From the **Select Tenant** dropdown, select the tenant that contains the DHCP server.

Step 6 (Optional) In the **Description** field, provide a description for the policy.

Step 7 Select `Relay` for the **Type**.

Step 8 Click **+Provider**.

Step 9 Select the provider type.

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

- `Application EPG`—specifies a specific application EPG that includes the DHCP server you are adding as an endpoint.
- `L3 External Network`—specifies the External EPG associated to the L3Out that is used to access the DHCP server.

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 will need to deploy the EPG before the relay is available for use.

Step 10 From the dropdown menu, pick the EPG or external EPG.

Step 11 In the **DHCP Server Address** field, provide the IP address of the DHCP server.

Step 12 Click **Save** to add the provider.

Step 13 (Optional) Add any additional providers.

Repeat steps 9 through 12 for each additional DHCP server.

Step 14 Click **Save** to save the DHCP relay policy.

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 additional 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 already created in the Multi-Site Orchestrator.
- A DHCP Relay policy created, as described in Creating DHCP Relay Policies, on page 152.

Procedure

Step 1 Log in to your Multi-Site Orchestrator GUI.

Step 2 From the left navigation menu, select **Application Management** > **Policies**.

Step 3 In the top right of the main pane, select **Add Policy** > > **Creating DHCP Policy**.

This opens an **Add DHCP** configuration screen.

Step 4 In the **Name** field, specify the name for the policy.

This is a name for the policy you're creating, not a specific DHCP option name. Each policy can contain multiple DHCP options.

Step 5 From the **Select Tenant** dropdown, select the tenant that contains the DHCP server.

Step 6 (Optional) In the **Description** field, provide a description for the policy.

Step 7 Select `Option` for the **Type**.

Step 8 Click **+Option**.

Step 9 Specify a name of the option.

While not technically required, we recommend using the same name for the option as listed in RFC 2132.

For example, `Name Server`.

Step 10 Specify an ID for the option .

You must provide the option code as listed in RFC 2132.

For example, `5` for Name Server option.

Step 11 Specify the option's 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.

Step 12 Click the check mark next to the **Data** field to save the option.

Step 13 (Optional) Repeat the steps to add any additional options.

Step 14 Click **Save** to save the DHCP option policy.

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 will need to re-deploy the bridge domain it 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 152.
- (Optional) A DHCP option policy, as described in Creating DHCP Option Policies, on page 154.
- The bridge domain to which you will assign the DHCP policy, as described in the Creating Schemas and Templates, on page 22 chapter.

Procedure

Step 1 Log in to your Multi-Site Orchestrator GUI.

Step 2 From the left navigation menu, select **Application Management** > **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 checkbox.

Step 6 From the **DHCP Relay Policy** dropdown, select the DHCP policy you want to assign to this BD.

Step 7 (Optional) From the **DHCP Option Policy** dropdown, select the option policy.

A DHCP option policy provides additional options to be passed to the DHCP relay. For additional details see Creating DHCP Option Policies, on page 154.

Step 8 Assign the bridge domain to any EPG that needs access to the DHCP server via the relay.

Creating DHCP Relay Contract

DHCP packets are not filtered by contracts but contracts are required in many cases to propagate routing information within the VRF and across VRFs. Even though 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 need to use the relay. Even though 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 152.
- (Optional) A DHCP option policy, as described in Creating DHCP Option Policies, on page 154.
- The bridge domain to which you have assigned the DHCP policy, as described in Assigning DHCP Policies, on page 155.

Procedure

Step 1 Log in to your Multi-Site Orchestrator GUI.

Step 2 From the left navigation menu, select **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** dropdown, 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 you selected when creating the DHCP relay policy.

c) In the right sidebar, click **+Contract**.

d) Select the contract 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 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 Multi-Site 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

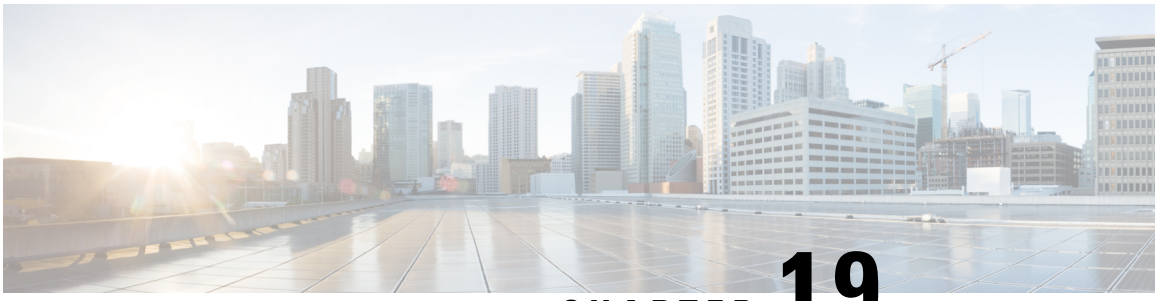
- If you make changes to the DHCP policy after you have assigned it to a bridge domain and deployed the bridge domain to one or more sites, you will need to re-deploy it for the DHCP policy changes to update on each site's APIC.
- You cannot deleted policies that are associated with one or more bridge domains, you must first unassign the policy from every bridge domain.

Procedure

Step 1 Log in to your Multi-Site Orchestrator GUI.

Step 2 From the left navigation menu, select **Policies**.

Step 3 Click the actions menu next to the DHCP policy and select **Edit** or **Delete**.



CHAPTER 19

Intersite L3Out

- Intersite L3Out Overview, on page 159
- Intersite L3Out Guidelines and Limitations, on page 160
- Configuring External TEP Pool, on page 161
- Creating or Importing Intersite L3Out and VRF, on page 162
- Configuring External EPG to Use Intersite L3Out, on page 164
- Creating a Contract for Intersite L3Out, on page 166
- Use Cases, on page 169

Intersite L3Out Overview

Prior to Release 2.2(1), each site managed by the Multi-Site Orchestrator required its own local L3Out configured in order to route traffic out of the fabric, which often resulted in lack of communication between endpoints in one site and a service (such a firewall, server load balancer, or mainframe) connected to the L3Out of another site.

Release 2.2(1) adds a feature that 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.

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.

Intersite L3Out Guidelines and Limitations

When configuring intersite L3Out, you must consider the following:

- 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.
- If you are upgrading from a release prior to Release 2.2(1), any existing External EPG to L3Out association at the site-local level will be preserved. In addition, the Multi-Site Orchestrator will now support creation of an L3Out and associating it with an External EPG at the template level.

  When creating a new L3Out in a schema template and associating it to an existing External EPG:

  - If the L3Out has the **same name** as the L3Out already defined in the APIC, the Orchestrator will take ownership of that L3Out but will not manage the configuration of L3Out node profiles, interface profiles, protocol settings, or route control settings.

Note If the L3Out already exists in APIC, we recommend importing it into Multi-Site Orchestrator along with any associated external EPG instead of creating a new L3Out with the same name from MSO.

    If you then choose to delete this L3Out from the Orchestrator, it will no longer be managed by the Orchestrator, but any previously existing L3Out configuration will be preserved in the APIC.

  - If the L3Out has a **different name** than the APIC defined L3Out the external EPG will be removed from the APIC defined L3Out and added to the L3Out defined in the Orchestrator. If this is the only external EPG under the APIC defined L3Out this can cause the configuration to be removed from the border leaves and can impact traffic.

- If you choose to downgrade to a release prior to Release 2.2(1), the L3Outs created in the Orchestrator MSO will no longer exist in the template so any template-level association between External EPG and L3Out will be removed. In this case, you will need to manually re-configure the External EPG to L3Out association at the site-local level. Any site-local associations will be preserved during the downgrade.
- 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.
- 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 a external TEP address for the border leaf switches in each pod. If you already have an external TEP pool configured, for example for another feature such as Remote Leaf, the same pool can be used. The existing TEP pool will be inherited by the Multi-Site 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 Multi-Site Orchestrator.

Step 2 From the left navigation pane, select **Infrastructure** > **Infra Configuration**.

Step 3 In the top right of the main pane, click **Configure Infra**.

Step 4 In the left sidebar, select the site you want to configure.

Step 5 In the main window, click a pod in the site.

Step 6 In the right sidebar, click **+Add TEP Pool**.

Step 7 In the **Add TEP Pool** window, specify the external TEP pool you want to configure for that site.

Note

You must ensure that the TEP pool you are adding does not overlap with any other TEP pools or fabric addresses.

Step 8 Repeat the process for each site and pod where you plan to use intersite L3Outs.

Creating or Importing Intersite L3Out and VRF

This section describes how to create an L3Out and associate it to a VRF in the Orchestrator GUI, which will then be pushed out to the APIC site, or import an existing L3Out from one of your APIC sites. You will then associate this L3Out with an external EPG and use that external EPG to configure specific intersite L3Out use cases.



Note The VRF you assign to the L3Out can be in any template or schema, but it must be in the same tenant as the L3Out.

Procedure

Step 1 Log in to your Multi-Site Orchestrator.

Step 2 From the left navigation pane, select **Application Management** > **Schemas**.

Step 3 Select the schema and then the template where you want to create or import the VRF and L3Out.

We recommend creating the L3Out in a template that is associated with a single site, in which case the L3Out will be created in that site only.

Alternatively, you can choose to create the L3Out in a template that is associated to multiple sites. In this case the L3Out will be created with the same name across all sites, which may bring some functional restrictions, as explained later in this chapter

Step 4 Create a new VRF and L3Out.

If you want to import an existing L3Out, skip this step.

Note
While you can create the L3Out object in the Orchestrator and push it out to the APIC, the physical configuration of the L3Out must be done in the APIC.

a) Scroll down to the **VRF** area and click the **+** icon to add a new VRF.

If you already have the VRF you plan to use for the L3Out, skip this substep.

In the right sidebar, provide the name for the VRF, for example `vrf-l3out`

b) Scroll down to the **L3Out** area and click the **+** icon to add a new L3Out.

In the right sidebar, provide the required information.

c) Provide the name for the L3Out, for example `l3out-intersite`.

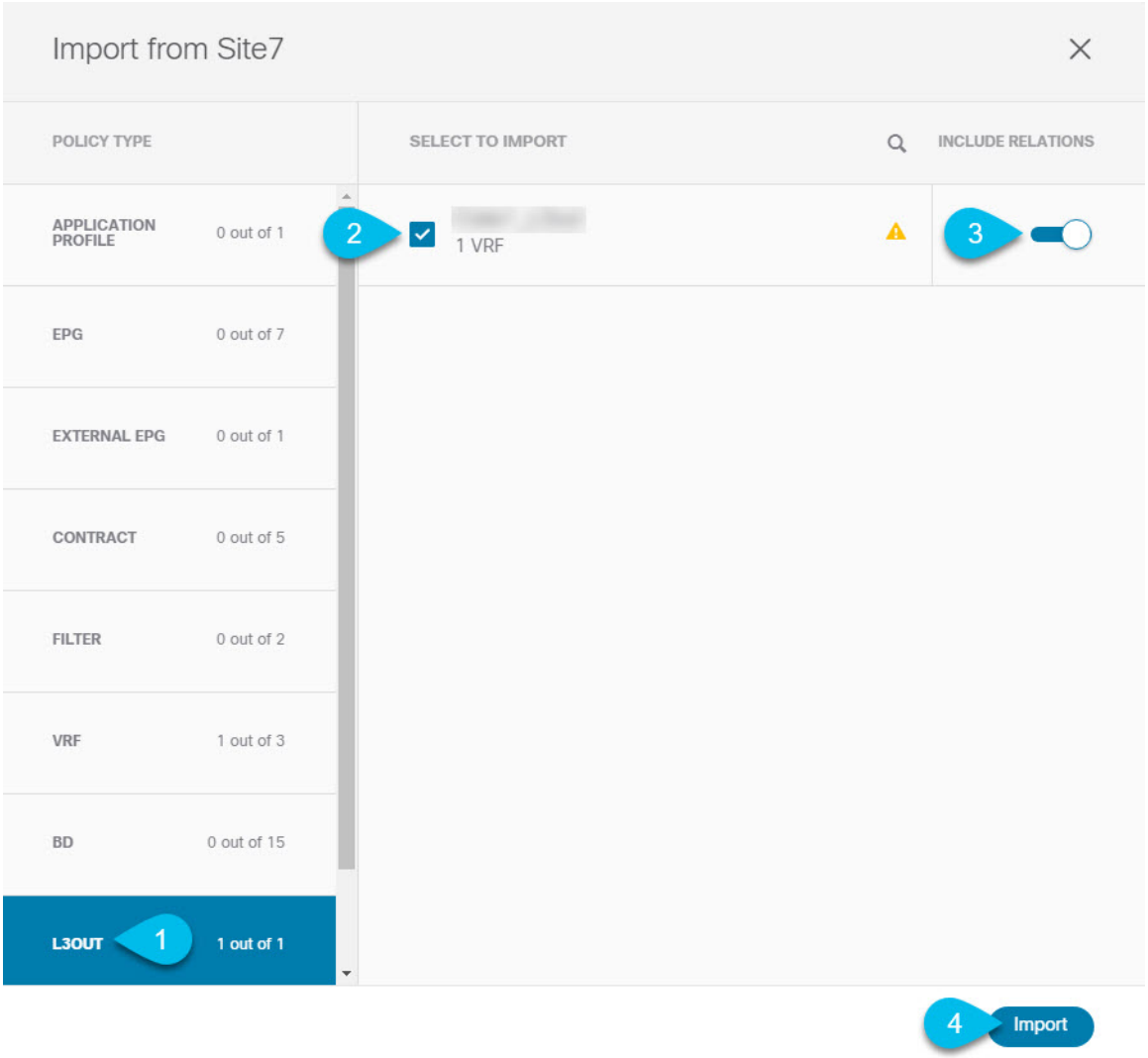
d) From the **Virtual Routing & Forwarding** dropdown, select the VRF.

Select the VRF you created in the first substep or choose a previously existing VRF.

Step 5 Import an existing VRF and L3Out.

If you created a new L3Out in previous step, skip this step.

Click **Import** in the main window pane to open at the



a) At the top of the main template view, click **Import**.
b) Select the site from which you want to import the L3Out.
c) In the import window's **Policy Type** menu, select **L3Out**.
d) Check the L3Out you want to import.

By default, importing the L3Out will also import the corresponding VRF. This may not be desirable when importing the L3Out in a site specific template as you would typically define the VRF in a stretched template associated to multiple sites. In this case, disable the **Include Relations** option before importing the L3Out. In this case, you will also need to re-map the L3Out to the correct VRF after importing it.

e) Click **Import**.

f) If you imported only the L3Out, select it in the template view and associate it to the appropriate VRF.

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.

Before you begin

Create the L3Out and associate it with a VRF as described in Creating or Importing Intersite L3Out and VRF, on page 162.

Procedure

Step 1 Select the 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 created on all of those sites. This is recommended when the external EPG's L3Outs 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 external EPG's L3Out provides access to external resources accessible only from that site.

Step 2 Scroll down to the **External EPG** area and click the + icon to add an external EPG.

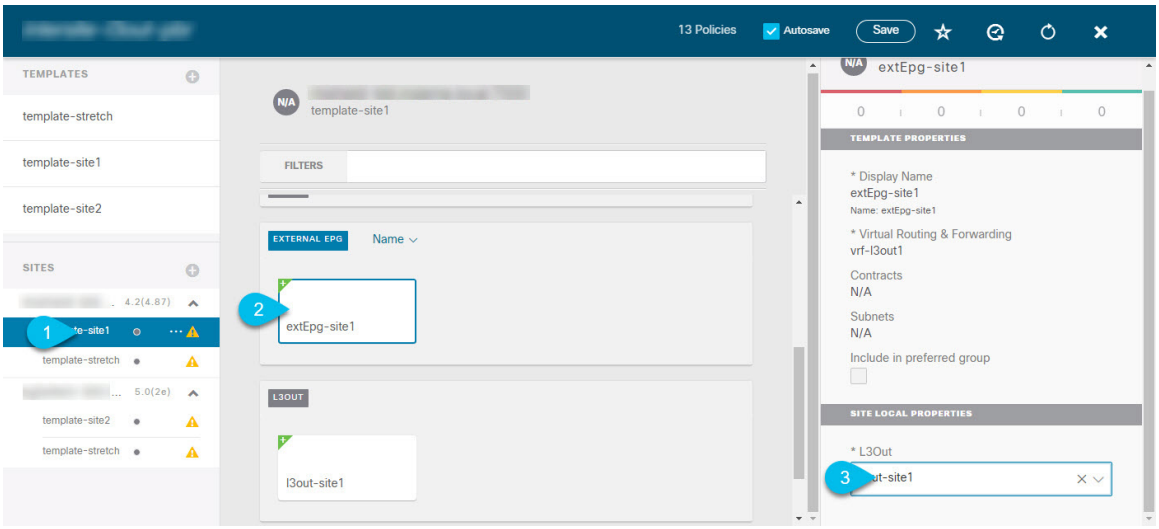
In the right sidebar, provide the required information.

a) Provide the name for the external EPG, for example `eepg-intersite-l3out`.
b) From the **Virtual Routing & Forwarding** dropdown, select the VRF you created and used for the L3Out.

Step 3 Map the external EPG to the L3Out.

You can map the external EPG to an L3Out at the site level or at the template level. We recommend creating the mapping at the site level because commonly each site defines a local L3Out with a unique name so the external EPG can be selectively mapped to each site specific L3Out independent of whether the external EPG itself is stretched.

To associate an L3Out with the external EPG at the site-local level:

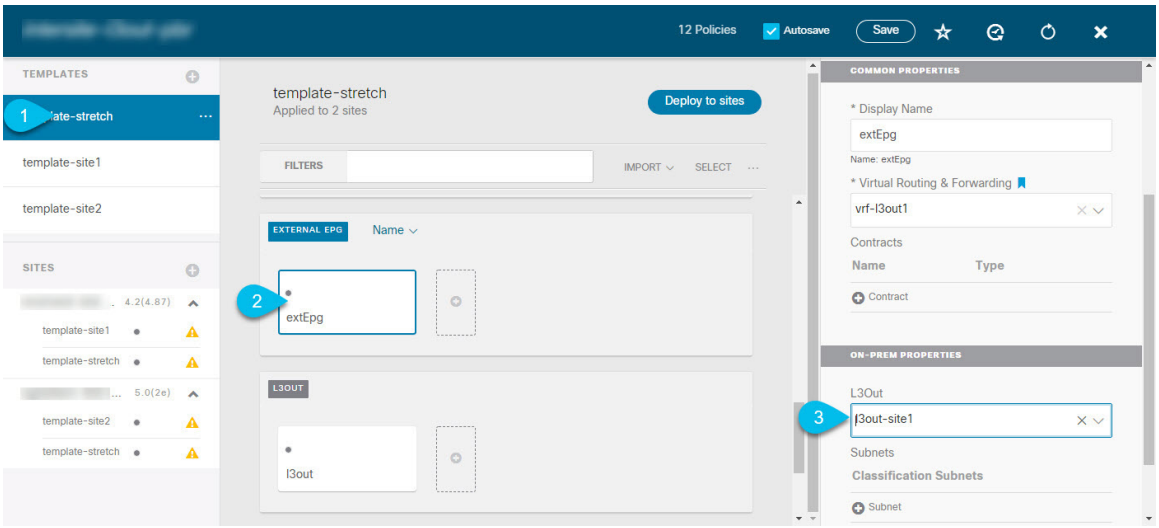


a) In the left sidebar of the schema view, select the site where the external EPG is deployed.
b) Scroll down to the **External EPG** area and select the external EPG.
c) In the right sidebar, scroll down to the **L3Out** dropdown and choose the intersite L3Out you created.

   In this case, both the APIC-managed and the Orchestrator-managed L3Outs will be available for selection. You can select either the L3Out you have created in the previous section specifically for this or pick an L3Out that exists in the site's APIC.

Alternatively, you can map the external EPG to an L3Out at the template level. While this could ease the configuration in deployments where multiple sites have defined the same L3Out name, we do not recommend this approach as it allows less flexibility for the type of connectivity that can be established between the fabrics that are part of the Multi-Site domain and the external routed network. For example, it would not be possible to control where a specific BD's subnets are advertised because mapping the BD to the L3Out would cause the BD subnet to be advertised out of all the L3Outs in all the sites since all the L3Outs have the same name.

To associate an L3Out with the external EPG at the template level:



a) In the left sidebar of the schema view, select the template where the external EPG is located
b) Scroll down to the **External EPG** area and select the external EPG.

c) In the right sidebar, scroll down to the **L3Out** dropdown and choose the intersite L3Out you created.

In addition, it is possible to migrate the configuration of an external EPG initially associated to the L3Out at the template level to a site-level mapping by removing the VRF association on the external EPG, re-associating the external EPG to the same VRF, then mapping the L3Outs at the site level. If this process is completed at once before deploying the template, there would be no traffic impact when pushing the new configuration as no changes are actually applied on the APIC side.

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, provide the classification subnet and the required options.

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 remote 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.
- 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.
- 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.

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

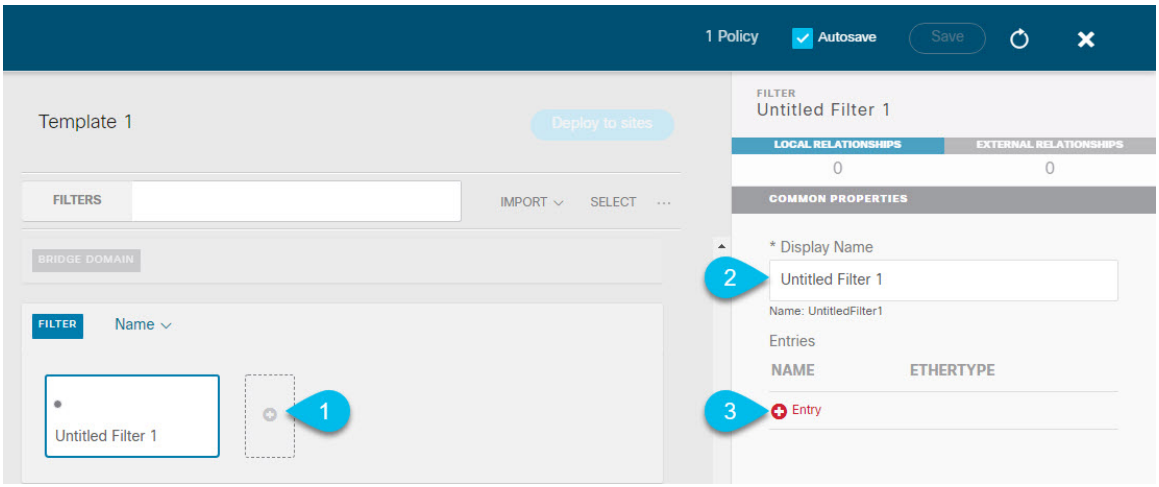
Step 1 Select the template where you want to create contract and filter.

You can use the same schema and template where you created the L3Out, VRF, and the external EPG or you can choose a different schema and template.

Because the contract is applied to objects (EPGs and external EPGs) deployed in different sites, we recommend defining it in a template associated to multiple sites. However, this is not strictly required and even if the contract and filters are

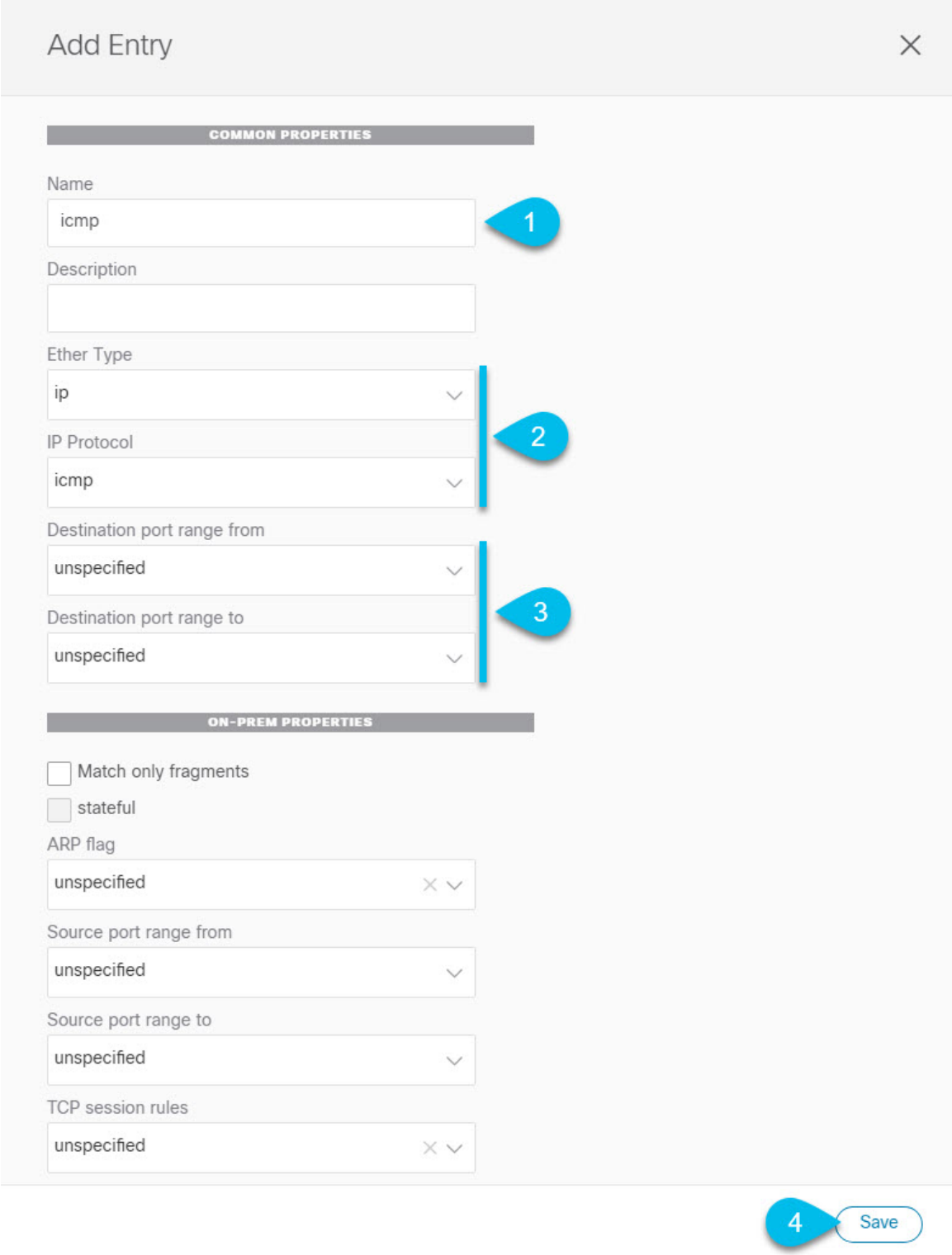
defined only as local objects in Site1, MSO will create the corresponding shadow objects in a remote Site2 when a local EPG or external EPG in Site2 needs to consume or provide that contract.

Step 2 Create a filter.



a) In the middle pane, scroll down to the **Filter** area, then click **+** to create a filter.
b) In the right pane, provide the **Display Name** for the filter.
c) In the right pane, click **+ Entry**.

Step 3 Provide the filter details.



a) Provide the **Name** for the filter.
b) Choose the **Ether Type**.

For example, `ip`.

c) Choose the **IP Protocol**.

For example, `icmp`.

d) Leave other properties unspecified.
e) Click **Save** to save the filter.

Step 4 Create a contract

a) In the middle pane, scroll down to the **Contract** area and click **+** to create a contract.
b) In the right pane, provide the **Display Name** for the contract
c) Select the appropriate **Scope** for the contract.

If you plan to configure different VRFs for the intersite L3Out and application EPG, you must select `tenant` for the scope. Otherwise, if both are in the same VRF, you can set the scope to `vrf`.

d) 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 will need to provide the filters only once and they will apply for traffic in both directions. If you leave this option disabled, you will need to provide two sets of filter chains, one for each direction.

Step 5 Assign the filters to the contract

a) In the right pane, scroll down to the **Filter Chain** area and click **+ Filter** to add a filter to the contract.

If you disabled the `Apply both directions` option, repeat this stem for the other filter chain.

b) In the **Add Filter Chain** window that opens, select the filter you added in previous step from the **Name** dropdown menu.
c) Click **Save** to add the filter to the contract.

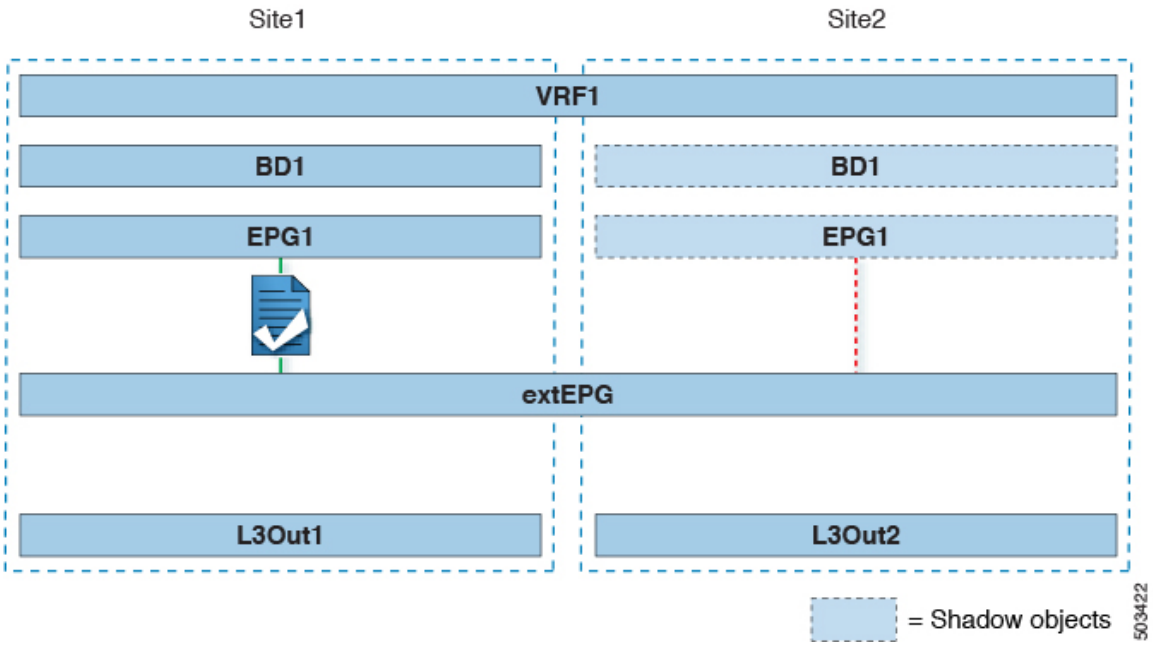
Use Cases

Intersite L3Out for Application EPGs (Intra-VRF)

This section describes the configuration 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).

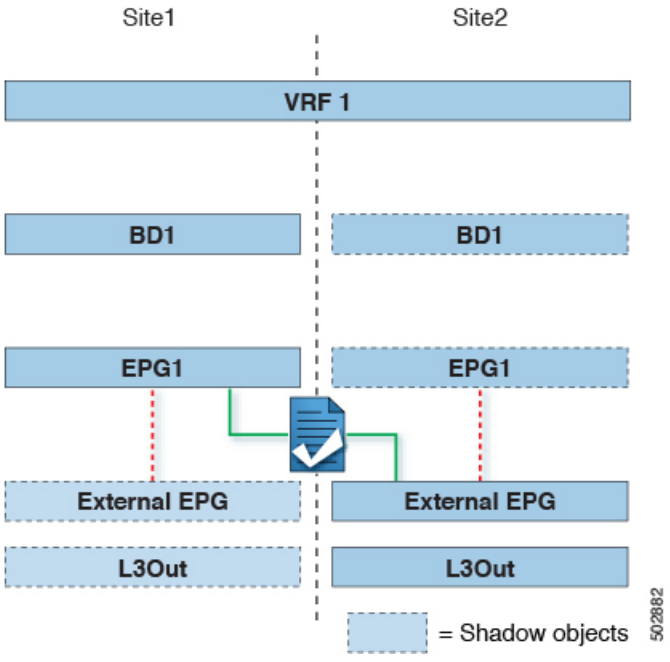
The first figure below shows a stretched external EPG and the associated L3Outs which will be created in both sites. An application EPG (`EPG1`) is created in Site 1 and has a contract with the external EPG. This use case is recommended when the L3Outs in the separate sites provide access to a common 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 19: 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 20: Non-Stretched (Site-Local) External EPG



The following steps describe the configuration 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:

- A schema with three templates.

  Create a template for each site (for example, `template-site1` and `template-site2`) where you will configure the objects unique to that site, such as the application EPG and the L3Outs. In addition, create a separate templates (for example, `template-stretched`) that you will use for the stretched objects, which in this case will be the external EPG.

- The L3Outs in each site, as described in the Creating or Importing Intersite L3Out and VRF, on page 162 section.

  In this use case, a separate L3Out will be imported or created in each site-specific template.

- The external EPG for the intersite L3Out, as described in Configuring External EPG to Use Intersite L3Out, on page 164.

  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.

- The contract you will use between the application EPG and the L3Out external EPG, as described in Creating a Contract for Intersite L3Out, on page 166.

  We recommend creating the contract and the filter in the stretched template (`template-stretched`).

Procedure

Step 1 Log in to your Multi-Site Orchestrator.

Step 2 From the left navigation pane, select **Application Management** > **Schemas**.

Step 3 Select the schema and template for the application EPG and bridge domain.

In this use case, you will associate the template to Site1.

Step 4 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 5 Assign the contract to the application EPG.

a) Select the EPG.
b) In the right sidebar, click **+Contract**.
c) Select the contract you created in previous section and its type.

You can choose whether the application EPG is the `consumer` or the `provider`.

Step 6 Assign the contract to the external EPG mapped to the remote L3Out.

a) Select the `template-stretched` where the external EPG is located.
b) Select the external EPG.
c) In the right sidebar, click **+Contract**.

d) Select the contract 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 7 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 the subnet(s) 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 Site1 and Site2 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 via the remote site's L3Out.

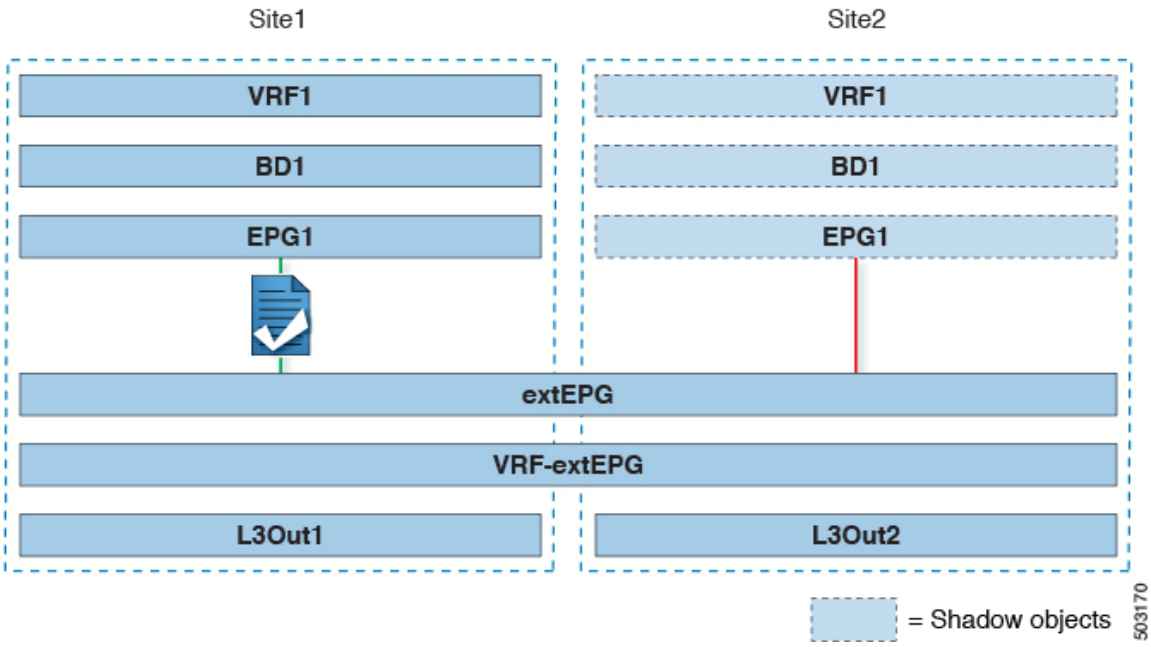
Step 8 Deploy the schema.

Shared Services with Intersite L3Out for Application EPGs (Inter-VRF)

This section describes the configuration 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 21: Stretched External EPG, Site-Local L3Outs and Application EPGs



The following steps describe the configuration required to implement the use case shown in Figure 3.

Before you begin

You need to have the following already configured:

- A schema with three templates.

  Create a template for each site (for example, `template-site1` and `template-site2`) where you will configure the objects unique to that site, such as the application EPGs and the L3Outs. In addition, create a separate templates (for example, `template-stretched`) that you will use for the stretched objects, which in this case will be the external EPG.

- The L3Outs in each site, as described in the Creating or Importing Intersite L3Out and VRF, on page 162 section.

  In this use case, a separate L3Out will be imported or created in each site-specific template.

- The external EPG for the intersite L3Out, as described in Configuring External EPG to Use Intersite L3Out, on page 164.

  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 services use case, you are also required to enable the **Shared Route Control** and the **Shared Security Import** flags for the subnet(s) associated to the external EPG(s) 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 you will use between the application EPG and the L3Out external EPG, as described in Creating a Contract for Intersite L3Out, on page 166.

We recommend creating the contract and the filter in the stretched template (`template-stretched`).

Procedure

Step 1 Log in to your Multi-Site Orchestrator.

Step 2 From the left navigation pane, select **Application Management** > **Schemas**.

Step 3 Select the schema and template for the application EPG and bridge domain.

In this use case, you will associate the template to Site1.

Step 4 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 5 Assign the contract to the application EPG.

a) Select the EPG.
b) In the right sidebar, click **+Contract**.
c) Select the contract 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 need to configure the subnet already defined under the BD also under the EPG in order to leak that route into the L3Out VRF. The same flags 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 6 Assign the contract to the external EPG mapped to the L3Outs.

a) Select the `template-stretched` where the external EPG is located.
b) Select the external EPG.
c) In the right sidebar, click **+Contract**.
d) Select the contract 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 7 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 the subnet(s) 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 Site1 and Site2 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 via the remote site's L3Out.

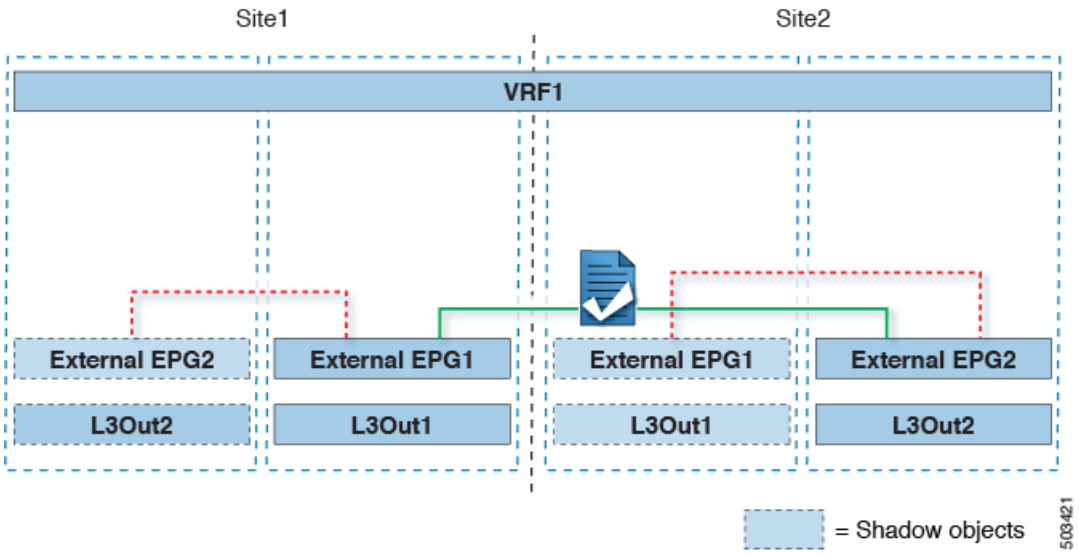
Step 8 Deploy the schema.

Intersite Transit Routing

This section describe the use cases where the Multi-Site domain acts as a distributed router allowing communication between entities (endpoints, network devices, service nodes, etc.) connected behind L3Outs deployed in different sites, a functionality normally referred to as intersite transit routing. The intersite transit routing is supported for intra-VRF as well as inter-VRF use cases.

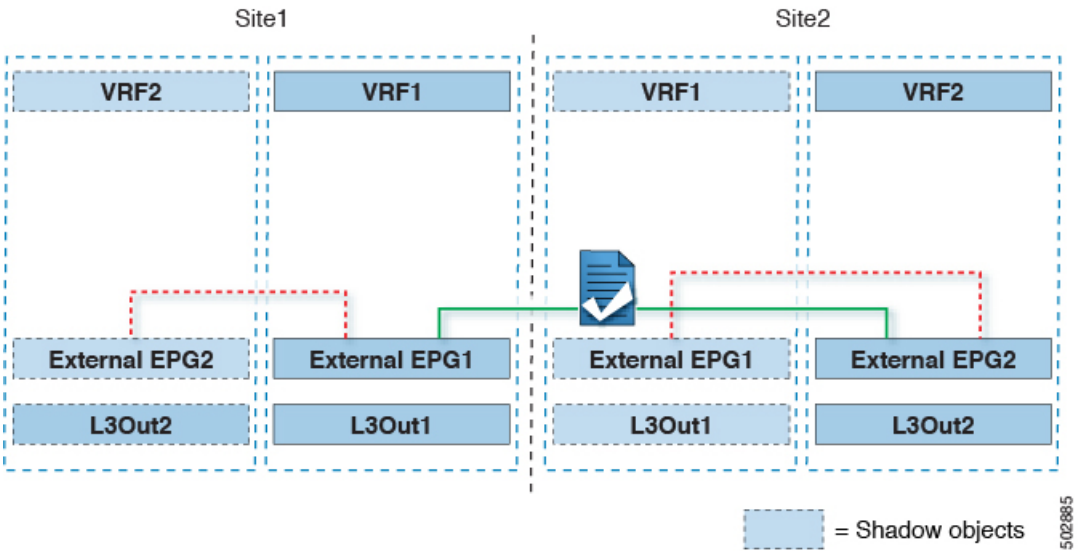
The figure below shows two L3Outs (`L3Out1` and `L3Out2`) configured in different sites. Each L3Out is associated with a respective external EPG (`External EPG1` and `External EPG2`). A contract between the two external EPGs allows communication between entities connected behind two different L3Outs in two different sites.

Figure 22: Intra-VRF Intersite Transit Routing



A similar configuration can be used when each site's L3Outs are in different VRFs.

Figure 23: 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 (obviously not overlapping). A couple of options are hence possible for the configuration of the prefix used for classification:

- Define the same `0.0.0.0/0` prefix on both external EPGs to ensure that inbound traffic received on the border leaf nodes of L3Out1 gets mapped to Ext-EPG1, whereas inbound traffic received on L3Out2 gets mapped to Ext-EPG2. Because the L3Outs are defined in separate fabrics, there are no conflict issues with this configuration.

  The external prefixes received on `L3Out1` must be advertised out of `L3Out2` and vice versa. If you are using `0.0.0.0/0` as classification subnet on both external EPGs, it is sufficient to enable the **Export Route Control** and the **Aggregate Export** flags.

- 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 configured for classification on `External EPG1` must be configured with the **Export Route Control** flag set on `External EPG2` and vice versa.

Note No matter which of the two classification 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 need to have the following already configured:

- A schema with three templates.

  Create a template for each site (for example, `template-site1` and `template-site2`) where you will configure the objects unique to that site, such as the application EPGs and the L3Outs. In addition, create a separate templates (for example, `template-stretched`) that you will use for the stretched objects, which in this case will be the external EPG.

- The L3Outs in each site, as described in the Creating or Importing Intersite L3Out and VRF, on page 162 section.

  In this use case, a separate L3Out will be imported or created in each site-specific template.

- 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 164.

- The contract you will use between the L3Out external EPGs defined in each site, as described in Creating a Contract for Intersite L3Out, on page 166.

  We recommend creating the contract and the filter in the stretched template (`template-stretched`).

Procedure

Step 1 Log in to your Multi-Site Orchestrator.

Step 2 From the left navigation pane, select **Application Management** > **Schemas**.

Step 3 Assign the contract to one of the external EPGs.

a) Select the schema and template where the external EPG is located.
b) Select the external EPG.
c) In the right sidebar, click **+Contract**.
d) Select the contract you created in previous section and its type.

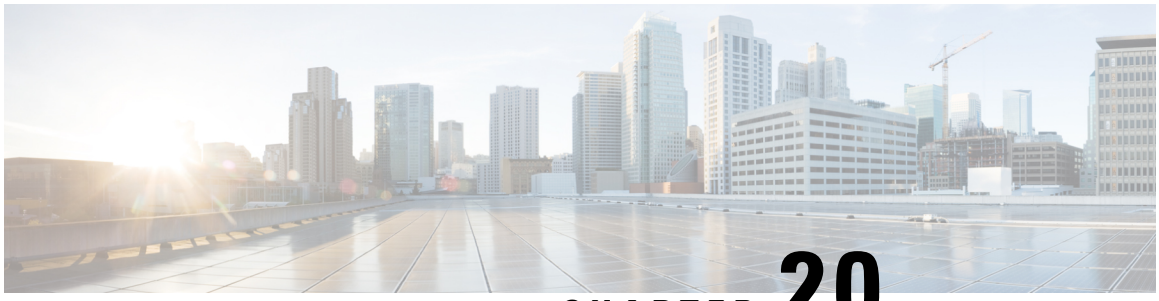
Choose `consumer` or `provider`.

Step 4 Assign the contract to the other external EPG.

a) Select the schema and template where the external EPG is located.
b) Browse to the template where the external EPG is located.
c) Select the external EPG.
d) In the right sidebar, click **+Contract**.
e) Select the contract you created in previous section and its type.

Choose `provider` or `consumer`.

Step 5 Deploy the templates to appropriate sites.



CHAPTER 20

Intersite L3Out with PBR

- Intersite L3Out with PBR, on page 179
- Supported Use Cases, on page 180
- Guidelines and Limitations, on page 184
- Configuring APIC Sites, on page 184
- Creating Templates, on page 188
- Configuring Service Graph, on page 190
- Creating Filter and Contract, on page 192
- Creating Application EPG, on page 198
- Creating L3Out External EPG, on page 201

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 allowed directly.

More in-depth information specific to services graphs and PBR is available in the *Cisco APIC Layer 4 to Layer 7 Services Deployment Guide*

PBR Support in Multi-Site Deployments

Cisco Multi-Site has supported EPG-to-EPG (east-west) and L3Out-to-EPG (north-south) contracts with PBR since Cisco APIC, Release 3.2(1). However, the L3Out-to-EPG across sites (traffic from an external endpoint in `site1` to an endpoint in `site2`) case was supported only if both sites had local L3Outs. The intersite L3Out use cases were limited to the examples and configurations described in the Intersite L3Out, on page 159 chapter. Similarly, the Service Graph integration with PBR but no intersite L3Out is described in great detail in the *Cisco Multi-Site and Service Node Integration White Paper*.

Starting with Cisco APIC, Release 4.2(5), the L3Out-to-EPG with PBR across sites (intersite L3Out) use case has been extended to support cases where the application EPG has no local L3Out or the local L3Out is down.

Supported Use Cases

The following diagrams illustrate the traffic flows between the 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.

The workflow to configure these examples is the same, with the only differences being whether you create the objects in the same or different VRFs (inter-VRF vs intra-VRF) and where you deploy the objects (stretched vs non-stretched):

1. Create the L4-L7 devices directly in the site's APIC, as described in Creating and Configuring L4-L7 Devices and PBR Policies, on page 185.

   You cannot create the devices and PBR policies from the Multi-Site Orchestrator, so you will need to log in to each site's APIC directly to configure those options.

2. Create the required templates, as described in Creating Templates, on page 188.

   We recommend creating a single stretched template that will contain all the objects deployed to all sites. Then an extra template for each site with the objects specific to that site only.

3. Create and configure the service graph, as described in Configuring Service Graph, on page 190.

4. Create the contract and filter you will use for all traffic between the application EPG and the external EPG containing the L3Out in another site, as described in Creating Filter and Contract, on page 192.

5. Create the application EPG with its VRF and bridge domain, as described in Creating Application Profile and EPG, on page 199.

   Depending on whether you plan to stretch the application EPG or not, you will create these objects in different templates. Similarly, you can choose to use the same or different VRFs for the application EPG and the L3Out.

6. Create the L3Out, as described in Creating or Importing Intersite L3Out and VRF, on page 201.

7. Create the external EPG for the L3Out, as described in Configuring External EPG to Use Intersite L3Out, on page 164.

Inter-VRF vs Intra-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.

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.

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 24: Inbound Traffic

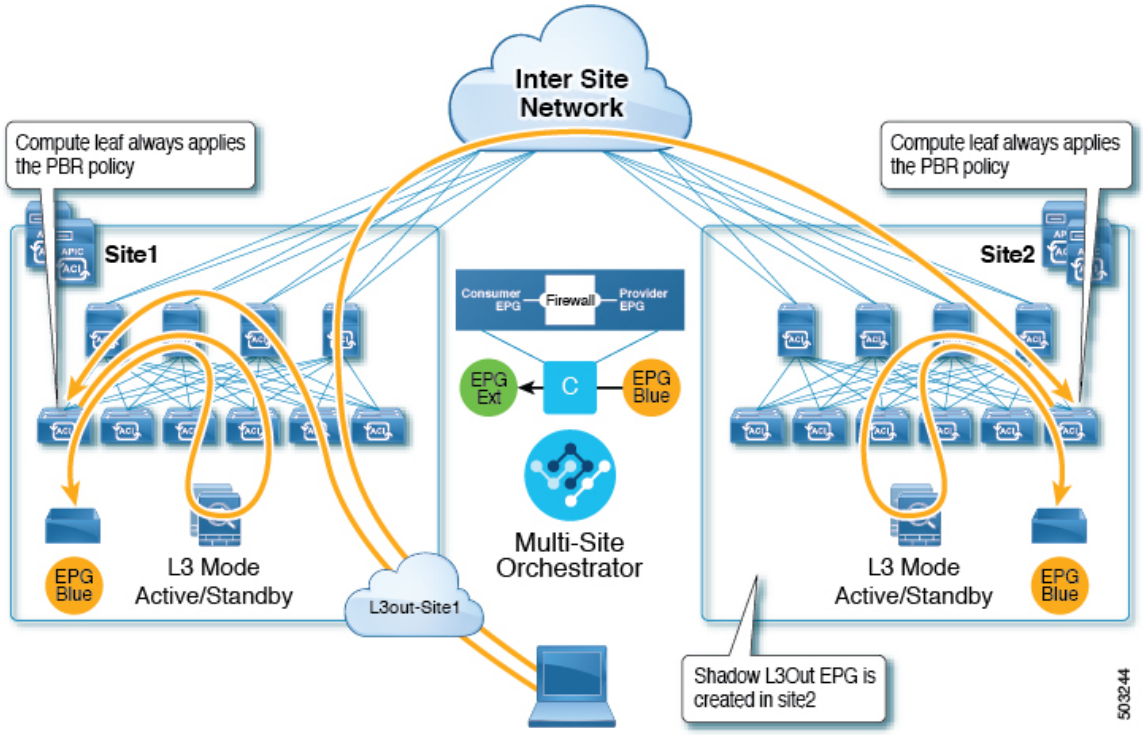
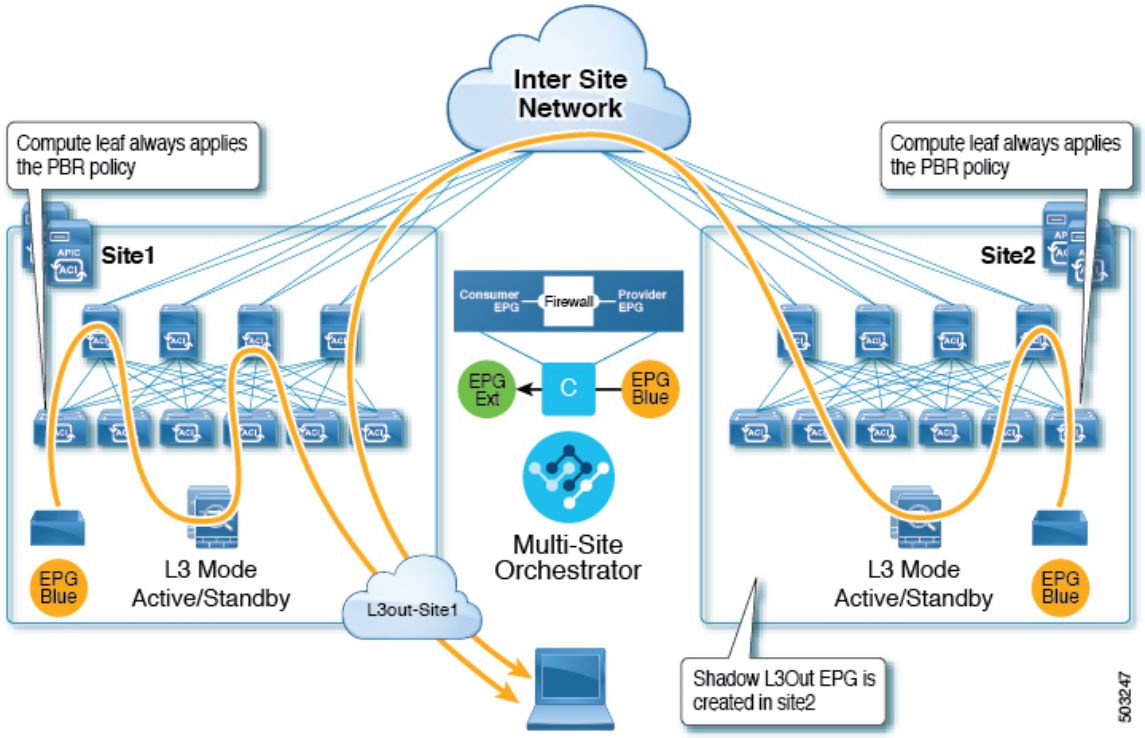


Figure 25: Outbound Traffic



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 26: Inbound Traffic

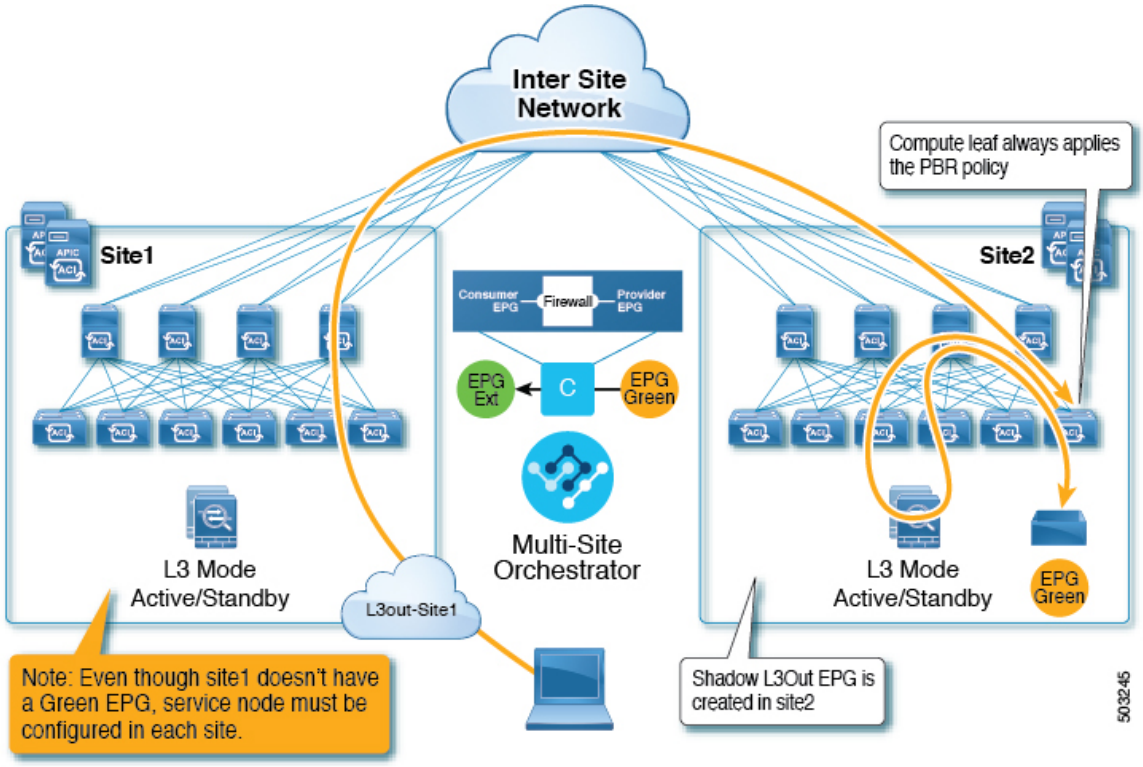
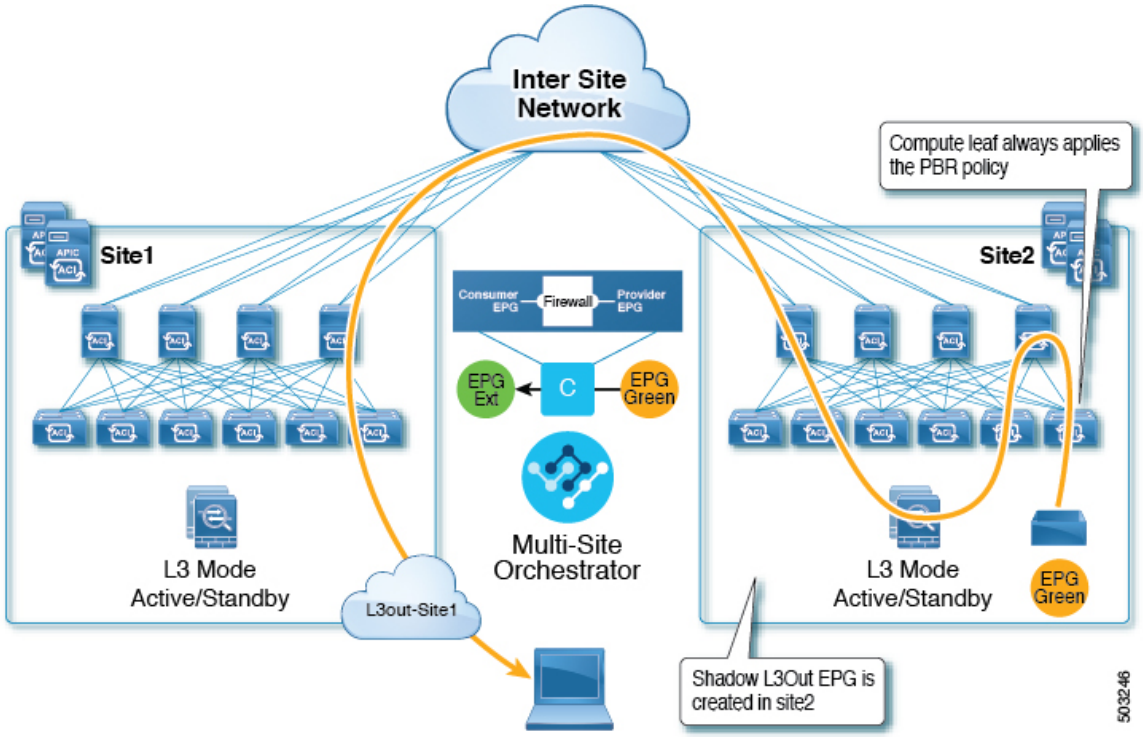


Figure 27: Outbound Traffic



Guidelines and Limitations

When configuring an Intersite L3Out with PBR, the following restrictions apply:

- For intersite L3Out without PBR use cases, see Intersite L3Out, on page 159
- 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 L3Out must be the `provider`.
  - Intra-VRF intersite L3Out with the application EPG as either the `provider` or the `consumer`
  - Intersite transit routing (L3Out-to-L3Out) with PBR is not supported.
- The above use cases are supported for sites running Cisco APIC, Release 4.2(5) or Release 5.1(x). They are not supported for sites running Cisco APIC, Release 5.0(x).
- In all supported cases, the application EPG can be stretched or not stretched.
- Service graph devices must be defined in each site, including the sites that don't have an application EPG that has a PBR contract with an intersite L3Out external EPG.
- Both one-arm and two-arm deployment models 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.
- 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.
- When configuring PBR, destination can be L1, L2, or L3.

Configuring APIC Sites

Configuring External TEP Pool

Intersite L3Out requires a external TEP address for the border leaf switches in each pod. If you already have an external TEP pool configured, for example for another feature such as Remote Leaf, the same pool can be used. The existing TEP pool will be inherited by the Multi-Site 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 Multi-Site Orchestrator.

Step 2 From the left navigation pane, select **Infrastructure** > **Infra Configuration**.

Step 3 In the top right of the main pane, click **Configure Infra**.

Step 4 In the left sidebar, select the site you want to configure.

Step 5 In the main window, click a pod in the site.

Step 6 In the right sidebar, click **+Add TEP Pool**.

Step 7 In the **Add TEP Pool** window, specify the external TEP pool you want to configure for that site.

Note
You must ensure that the TEP pool you are adding does not overlap with any other TEP pools or fabric addresses.

Step 8 Repeat the process for each site and pod where you plan to use intersite L3Outs.

Creating and Configuring L4-L7 Devices and PBR Policies

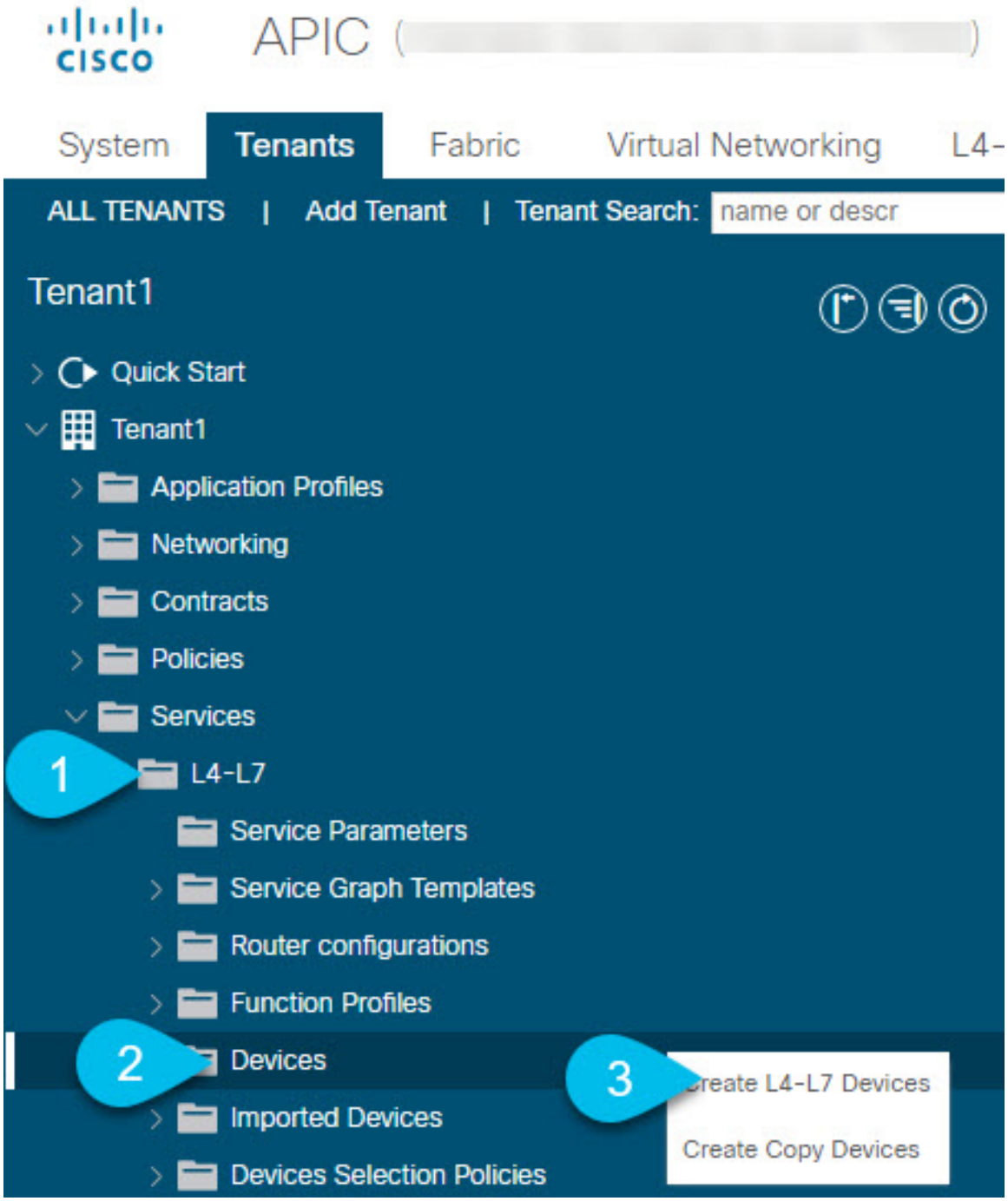
You must create the service graph devices and define the PBR policies directly in each site's APIC.

Procedure

Step 1 Log in to your Cisco APIC.

Step 2 In the top menu bar, click **Tenants**, then select the tenant where you want to create the device.

Step 3 Create an L4-L7 device.

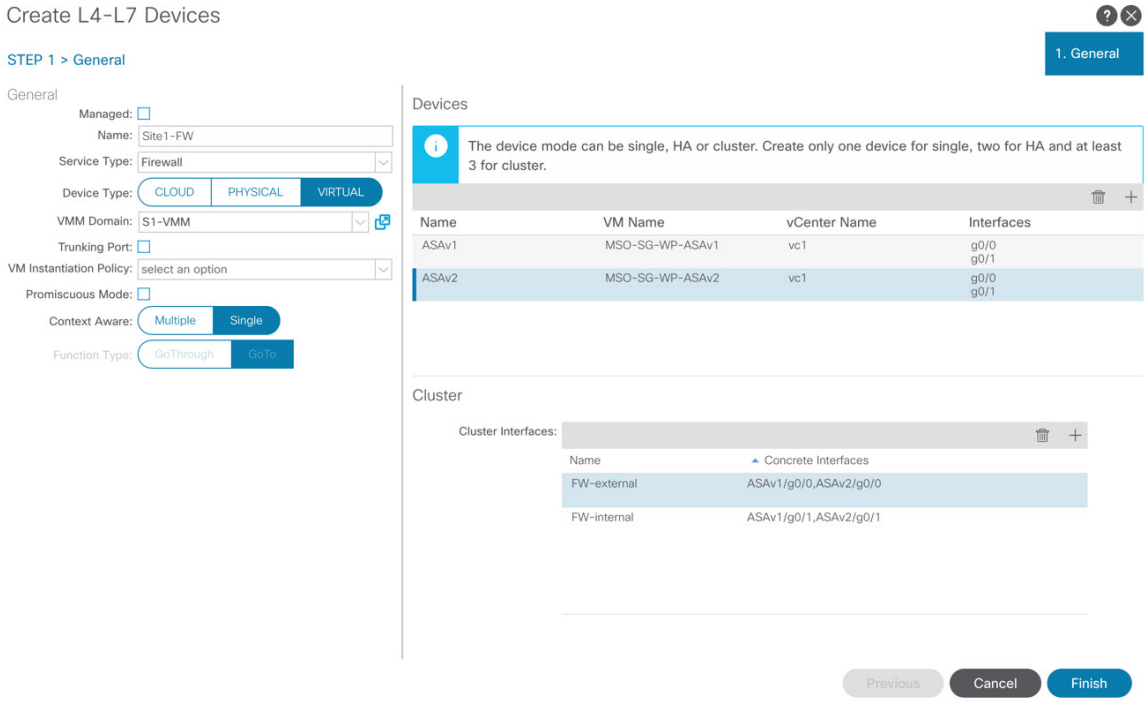


a) In the left sidebar, expand **<tenant-name>** > **Services** > **L4-L7** category.
b) Right-click **Devices** category.
c) Choose **Create L4-L7 Devices**.

The **Create L4-L7 Devices** configuration dialog opens.

Step 4 Configure the L4-L7 device.

The following image shows a sample device configuration. Your configuration settings will depend on the type and purpose of the device.



Step 5 Create a PBR policy.

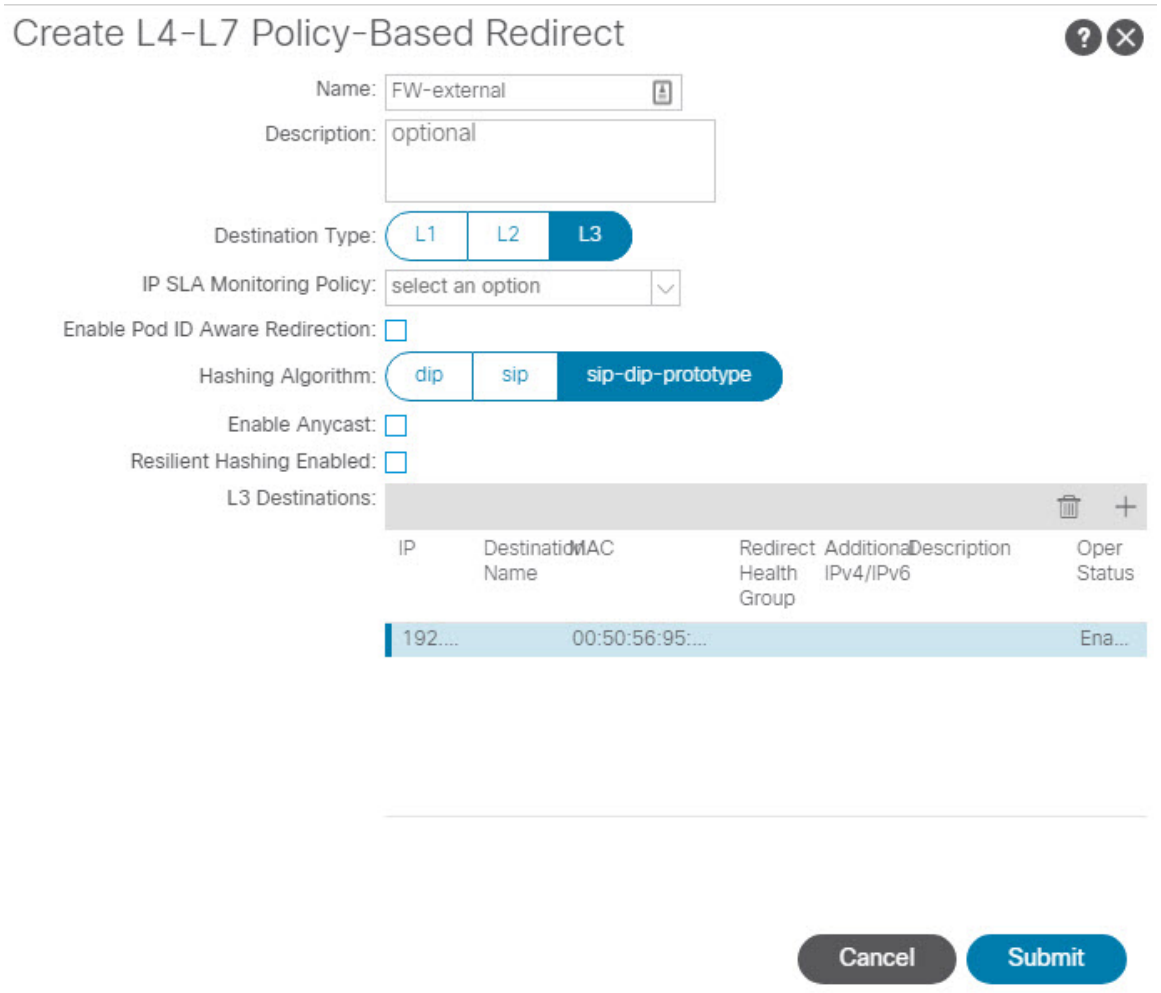
a) In the left sidebar, expand **<tenant-name>** > **Policies** > **Protocol** category.
b) Right-click **L4-L7 Policy-Based Redirect** category.
c) Choose **Create L4-L7 Policy-Based Redirect**.

The **Create L4-L7 Policy-Based Redirect** configuration dialog opens.

Step 6 Configure the PBR policy.

The following image shows a sample PBR policy configuration with destination IP and MAC added.

Your configuration settings will depend on the type and purpose of the device and policy you create. For example, you can configure additional options such as IP-SLA, hashing algorithm, resilient hashing, and so on in the PBR policy.



Step 7 Repeat the previous steps to create the required devices and PBR policies in the other site.

Creating Templates

When creating the schema and template, we recommend separating the templates in the following way:

- A single shared template that will contain all the objects that are stretched between all sites.
- One template per site that will contain the objects you will deploy to that site only.

In this example, we will work with two sites, so we will create a total of three templates: one for each site, plus one stretched.

Before you begin

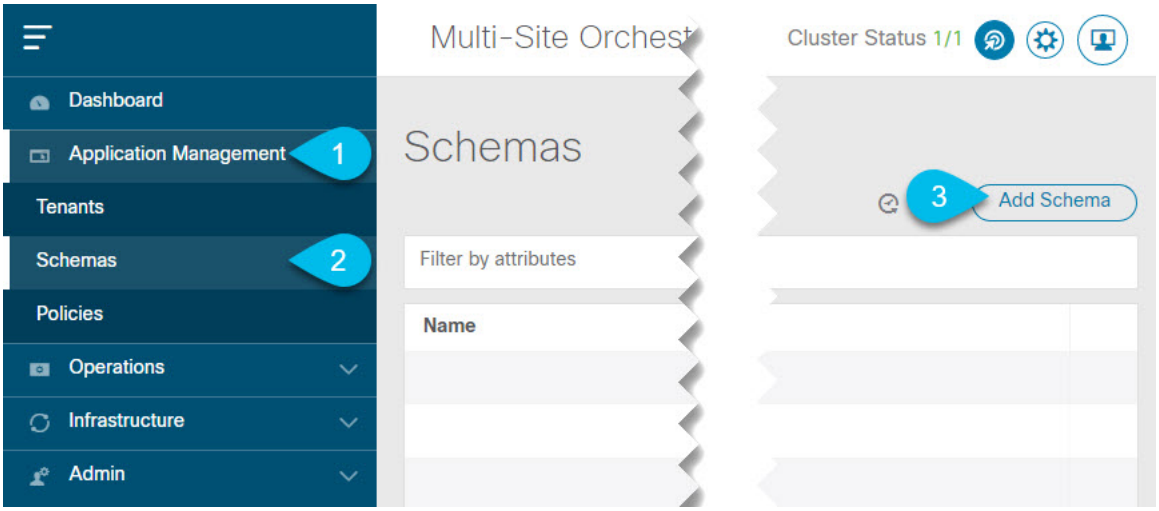
You must have:

- Reviewed the Guidelines and Limitations, on page 184 and completed any prerequisites listed there.
- Finished configuring the individual APIC sites as described in Configuring External TEP Pool, on page 161 and Creating and Configuring L4-L7 Devices and PBR Policies, on page 185.

Procedure

Step 1 Log in to your Cisco Multi-Site Orchestrator GUI.

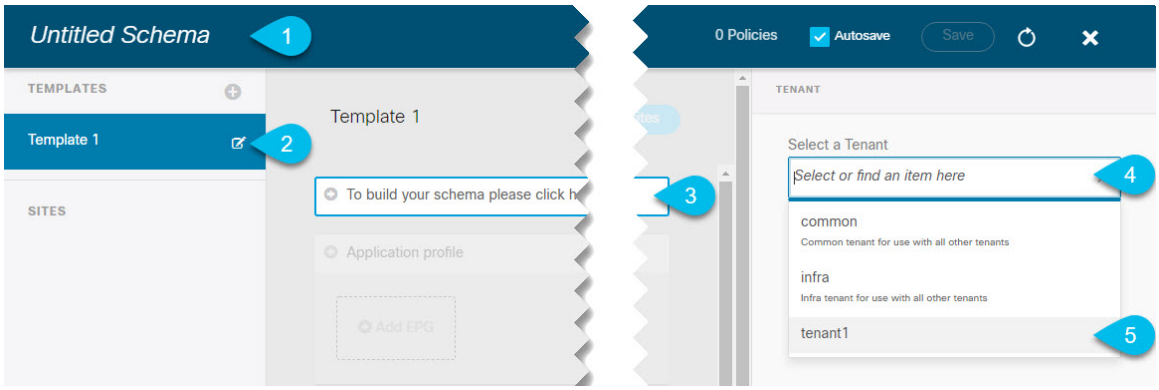
Step 2 Create a new Schema.



a) In the left navigation sidebar, expand the **Application Management** category.
b) Choose **Schemas**.
c) Click **Add Schema** to create a new schema.

The **Edit Schema** window will open.

Step 3 Name the Schema and pick the Tenant.



a) Replace **Untitled Schema** with the name for your schema.

Simply click on the `Untitled Schema` name to edit it.

b) Rename the template.

In the left sidebar, mouse over the template and click the **Edit** icon.

For example, `template-stretched`.

c) In the main pane, click **To build your schema please click here to select a tenant**.
d) In the right sidebar, click the **Select a Tenant** dropdown.
e) Select the tenant.

Step 4 Create any additional templates.

In the left sidebar, click the plus (+) icon next to **Templates** to add the site-specific templates. Then follow the same instructions described in the previous steps to name the templates and pick the tenant.

For example, `template-site1` and `template-site2`.

Configuring Service Graph

You must have:

- Created the L4-L7 devices directly in each site's APIC, as described in Creating and Configuring L4-L7 Devices and PBR Policies, on page 185.
- Created the templates where you will create these objects, as described in Creating Templates, on page 188.

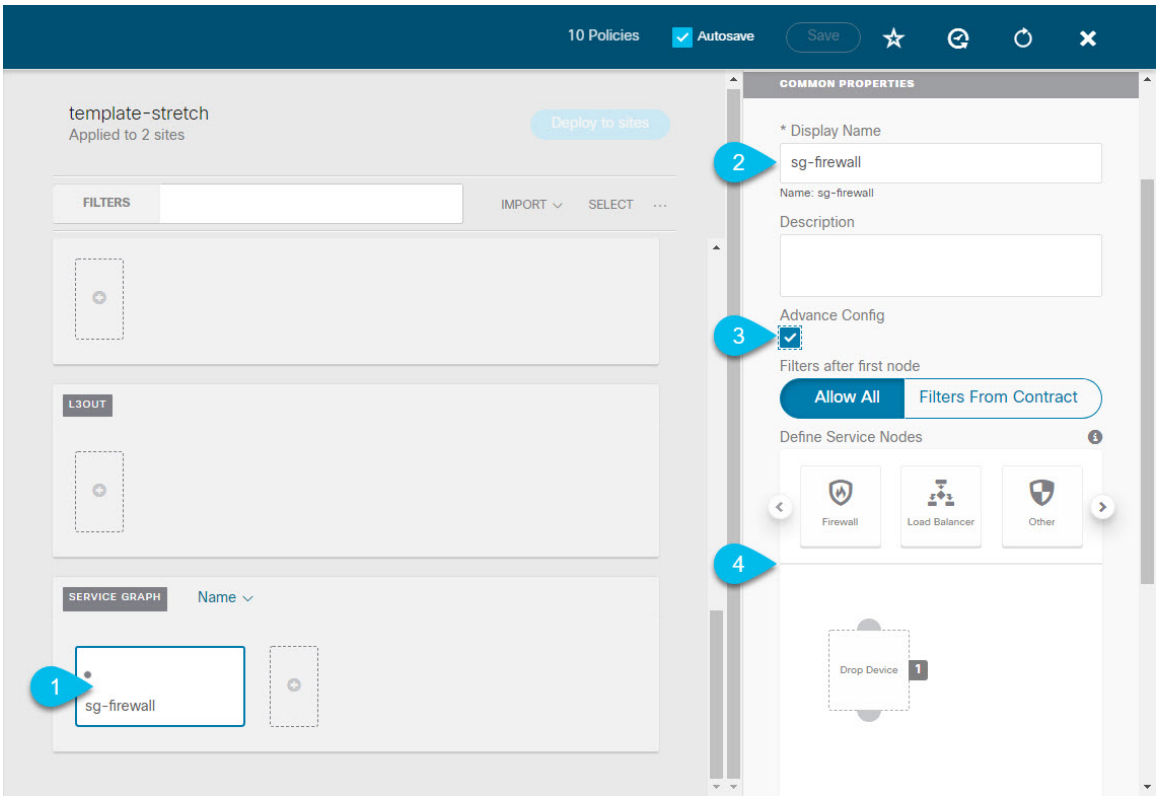
This section describes how to configure one or more devices for a service graph.

Procedure

Step 1 Select the template where you will create the service graph.

You will create a single service graph in the `template-stretch` but configure site-local devices for it as described later in this procedure.

Step 2 Create the Service Graph.



a) In the main pane, scroll down to the **Service Graph** area and click the + sign to create a new one.
b) Provide the **Display Name** for the service graph.
c) (Optional) Check the **Advanced Config** option.

This option allows you to configure whether traffic is restricted or not after the first service graph node. If you do not enable this option, all traffic is allowed after the first service graph node by default.

If you choose to enable the **Advanced Config**, select one of the following two options:

- **Allow All**: Use default (`permit-all`) filter instead of specific filter from contract subject.

  This is the same behavior as with **Advanced Config** disabled.

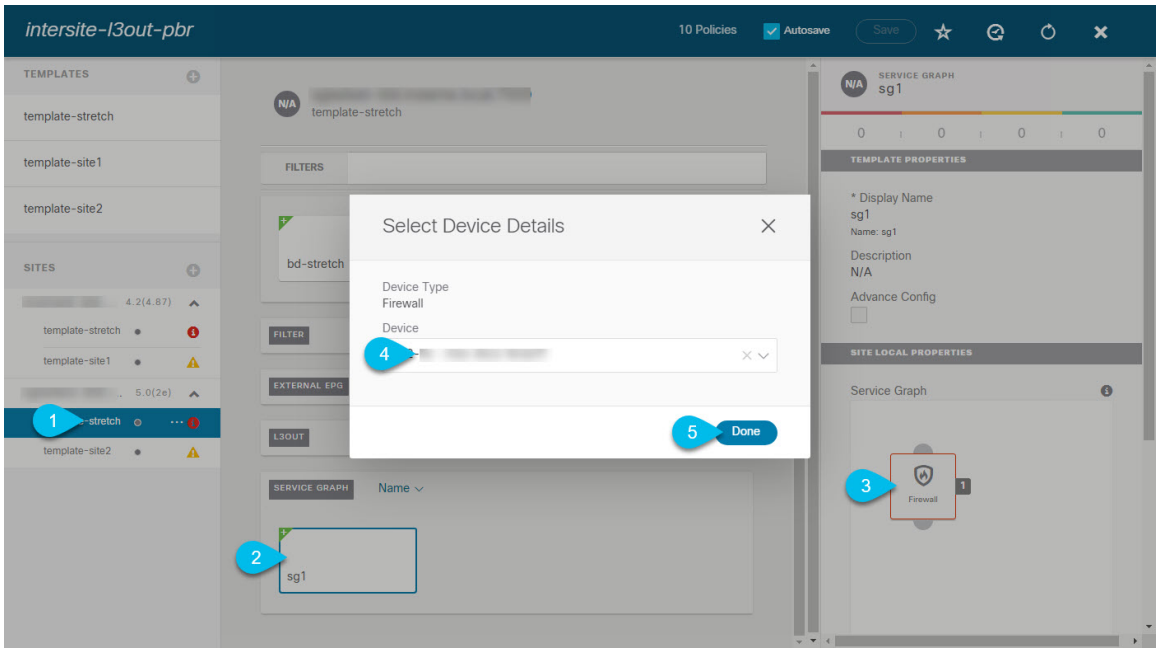
- **Filters From Contract**: Use specific filters from contract subject.

d) In the right sidebar, scroll down to the **Define Service Nodes** area and drag and drop one or more nodes into the **Drop Device** box.

Multi-Site supports up to two nodes per service graph.

Step 3 Configure service graph's site-local devices.

You must perform this step for every site that is part of the Multi-Site domain.



a) From the left sidebar, select one of the sites where you will deploy this service graph.
b) In the main pane, select the service graph you created.
c) In the right sidebar, click on the service graph node.
d) In the **Select Device Details** window, choose the device you have created in the site's APIC.

Creating Filter and Contract

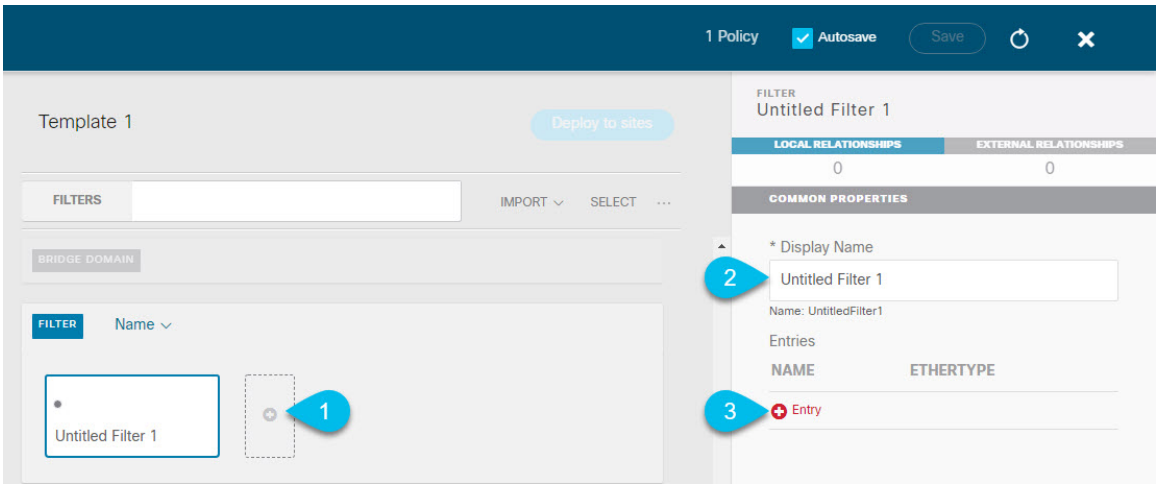
You must have:

- Created the templates where you will create these objects, as described in Creating Templates, on page 188.

This section describes how to create a contract and filters that will be used for the traffic going between the application EPG and the L3Out through the service graph.

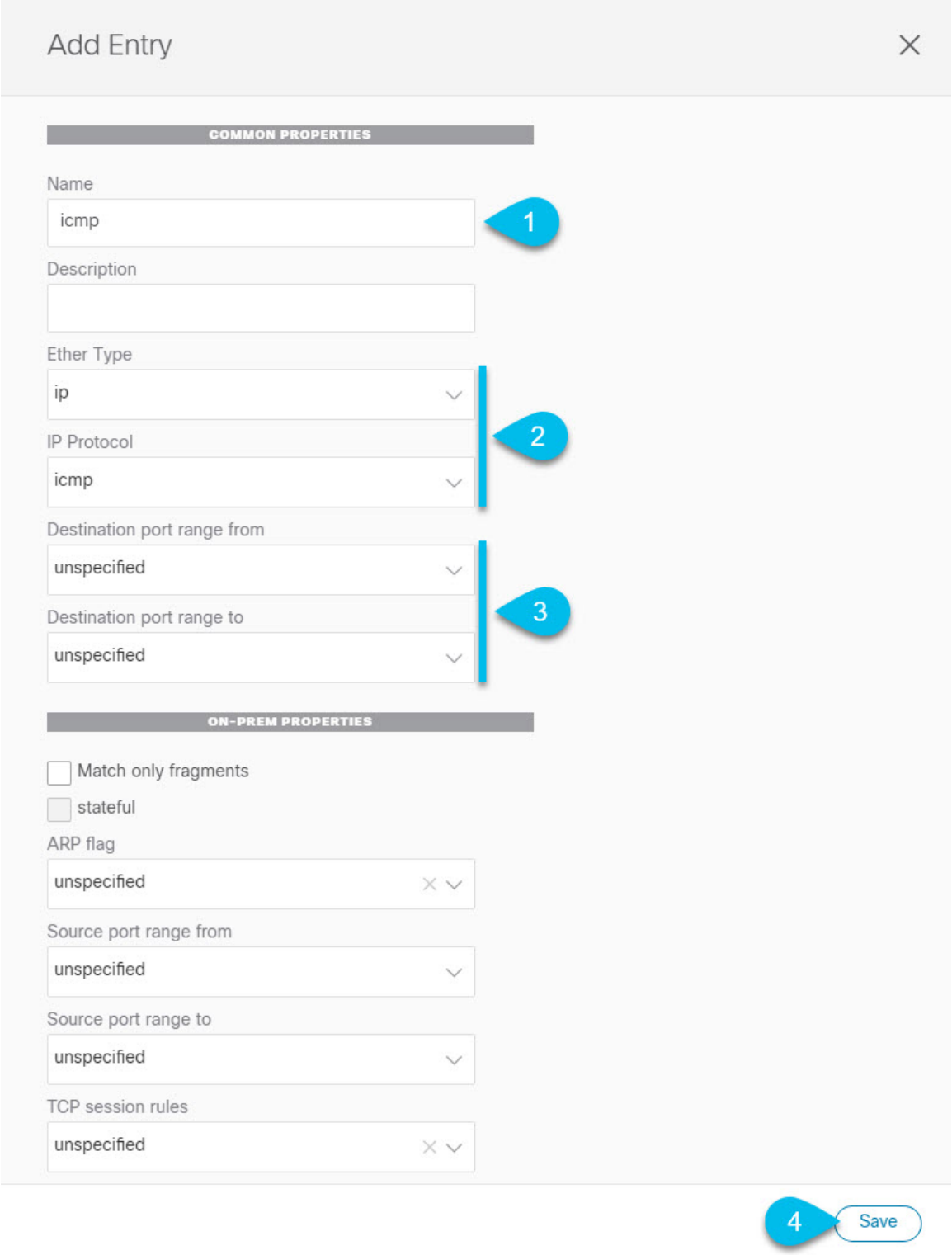
Procedure

Step 1 Create a filter.



a) In the middle pane, scroll down to the **Filter** area, then click + to create a filter.
b) In the right pane, provide the **Display Name** for the filter.
c) In the right pane, click + **Entry**.

Step 2 Provide the filter details.



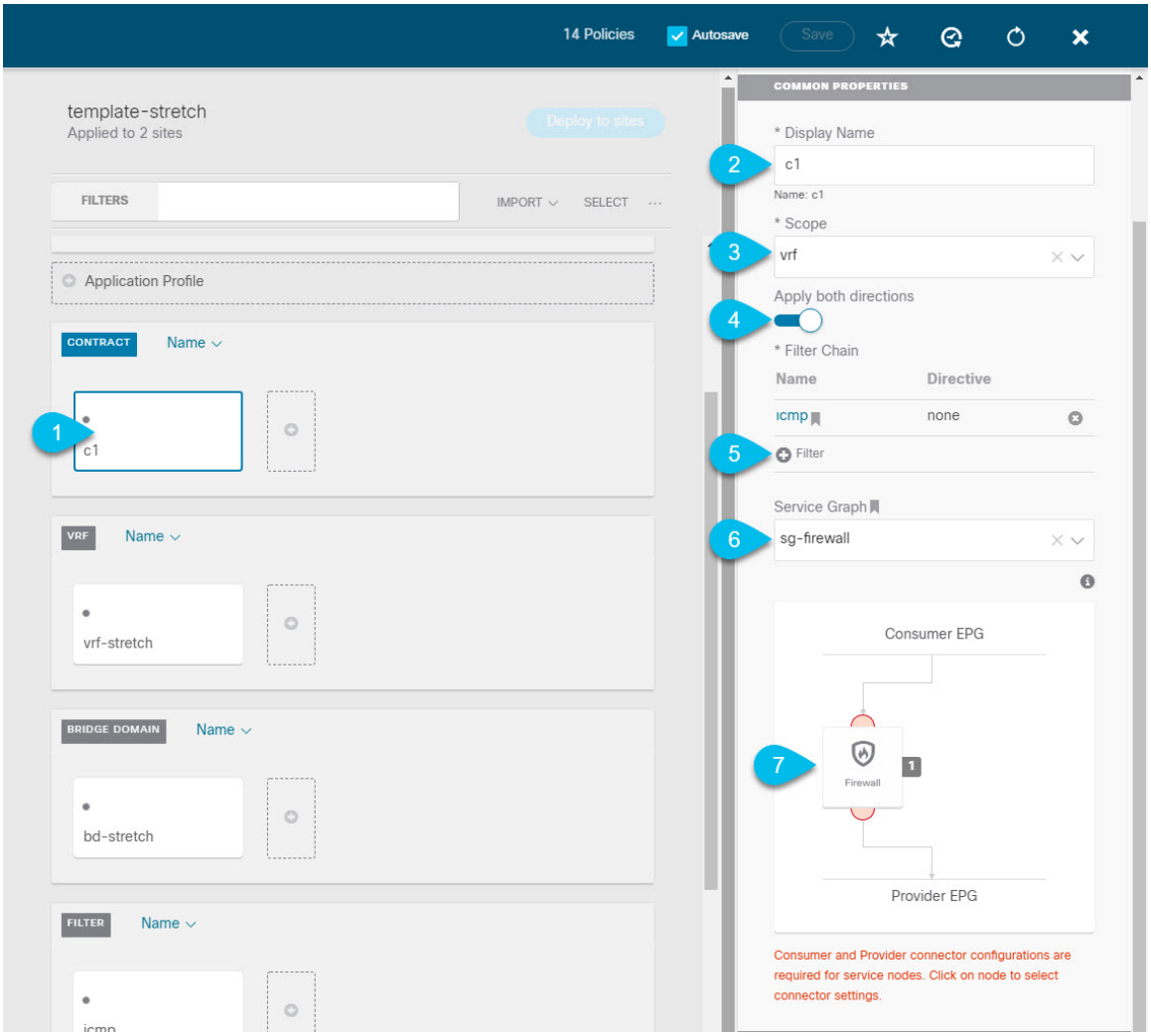
a) Provide the **Name** for the filter.
b) Choose the **Ether Type** and **IP Protocol**.

For example, `ip` and `icmp`.

c) Leave other properties `unspecified`.

d) Click **Save** to save the filter.

Step 3 Create a contract



a) In the middle pane, scroll down to the **Contract** area and click + to create a contract.

b) In the right pane, provide the **Display Name** for the contract

c) From the **Scope** dropdown menu, select the scope of the contract.

If your application EPG and L3Out are in the same VRF, choose `vrf`; otherwise, if you are configure inter-VRF use case, select `tenant`.

d) Ensure that **Apply both directions** is enabled.

This allows you to use the same filter to apply for both consumer-to-provider and provider-to-consumer directions.

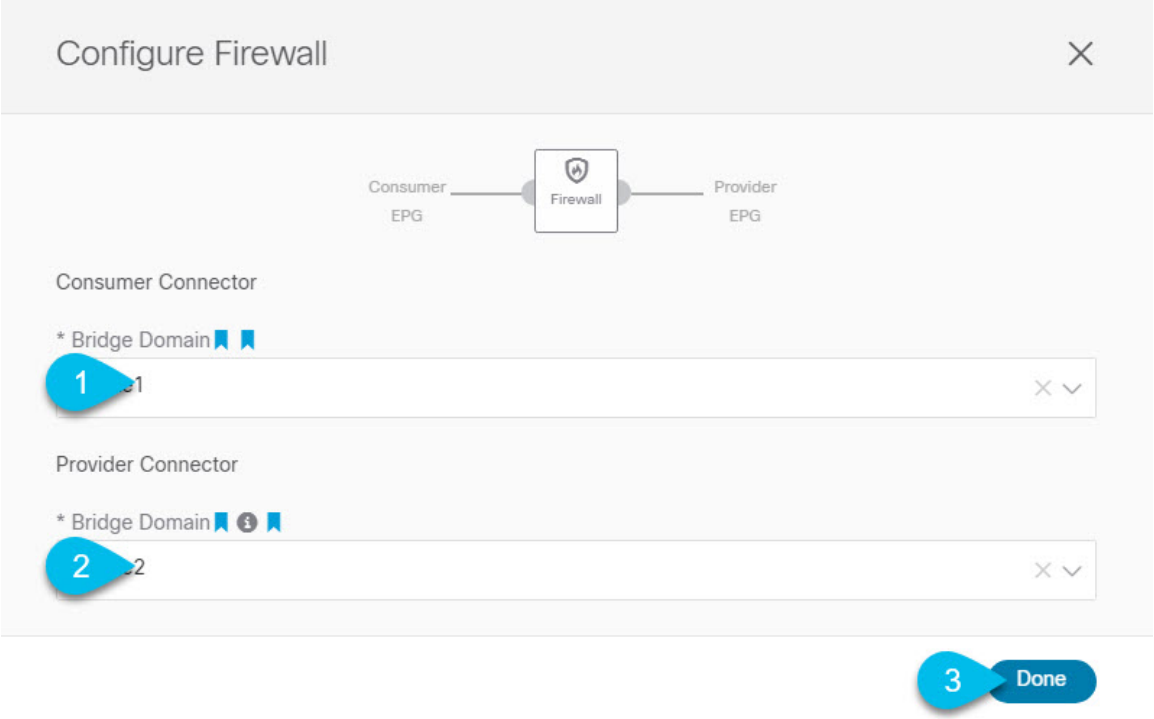
e) In the right pane, scroll down to the **Filter Chain** area and click **+ Filter** to add a filter to the contract.

In the **Add Filter Chain** window that opens, select the filter you added in previous section from the **Name** dropdown menu.

If you disabled the `Apply both directions` option, repeat this stem for the other filter chain.

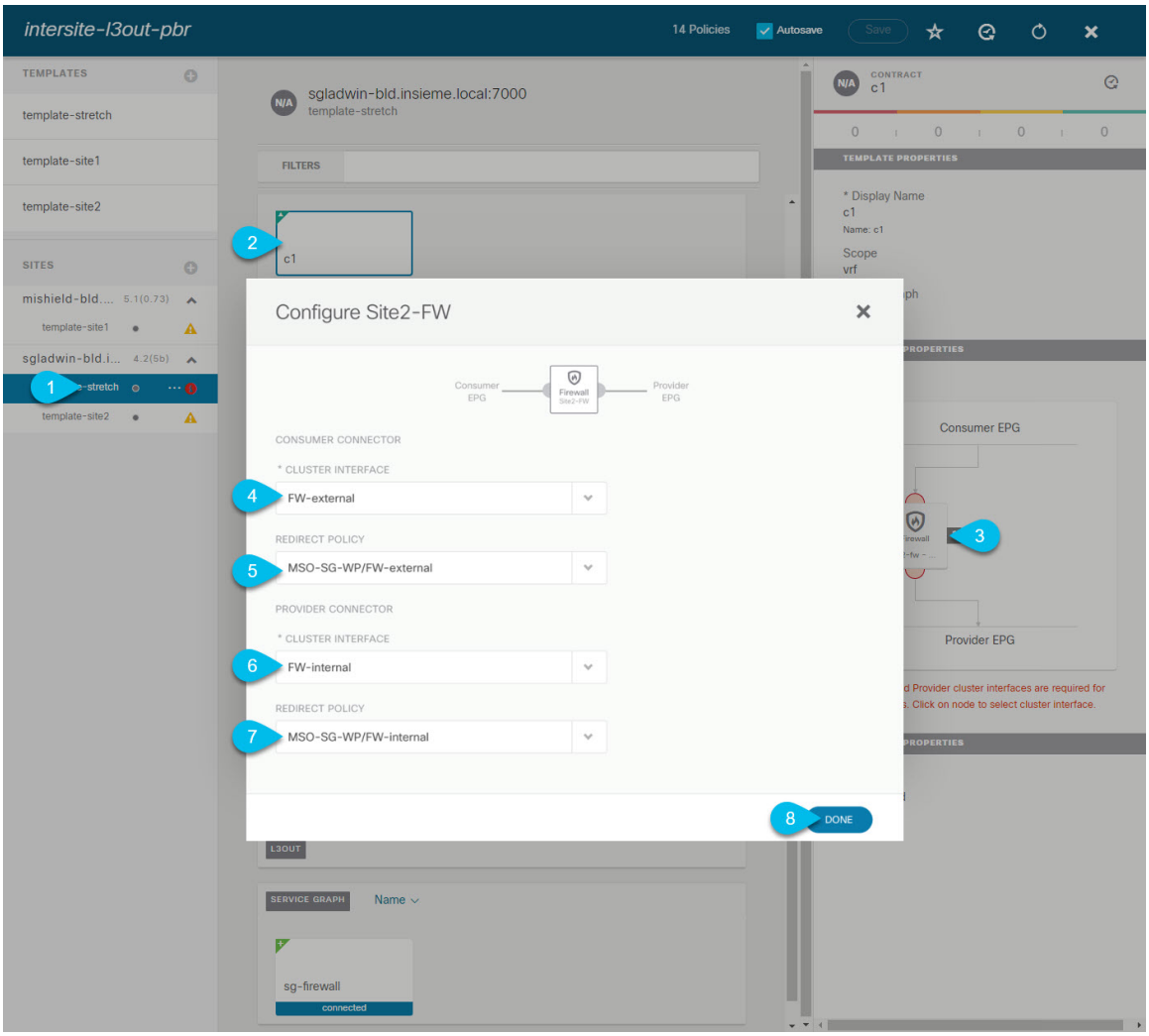
f) From the **Service Graph** dropdown, select the service graph you created in previous section.
g) Click the service graph node to configure its connectors.

Step 4 Select bridge domains for the service graph nodes' connectors.



a) Provide the **Consumer Connector** bridge domain.
b) Provide the **Provider Connector** bridge domain.
c) Click **Done** to save.

Step 5 Configure the contract's site-local properties.



a) In the left sidebar, select the template under a site to which it is assigned.
b) In the main pane, select the contract.
c) In the right sidebar, click a service graph node.
d) Select the **Cluster Interface** for the **Consumer Connector**.
e) Select the **Redirect Policy** for the **Consumer Connector**.
f) Select the **Cluster Interface** for the **Provider Connector**.
g) Select the **Redirect Policy** for the **Provider Connector**.
h) Click **Done** to save the changes.
i) Repeat this step for every site.

Creating Application EPG

Creating VRF and Bridge Domain for Application EPG

This section describes how to create the VRF and bridge domain (BD) for your application EPG.

Before you begin

You must have:

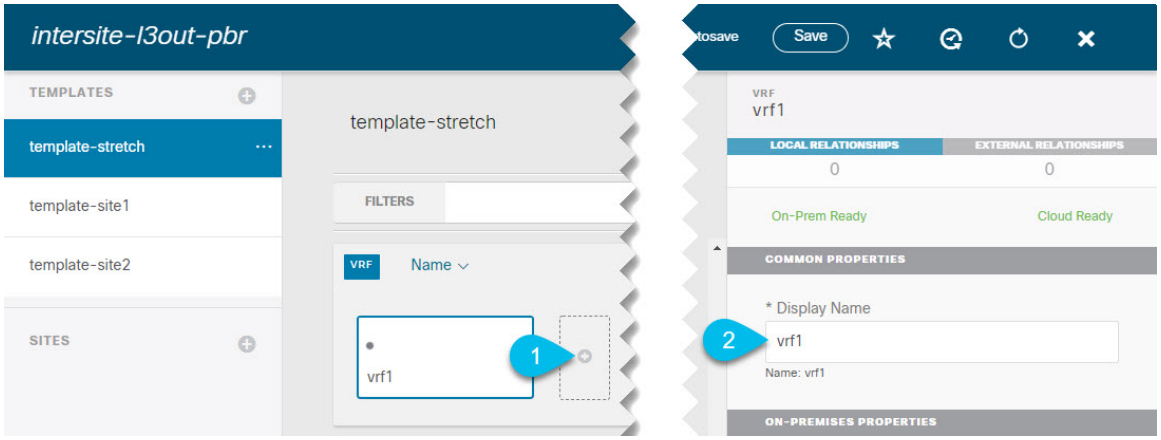
- Created the templates where you will create these objects, as described in Creating Templates, on page 188.

Procedure

Step 1 Select the template where you will create the VRF and BD.

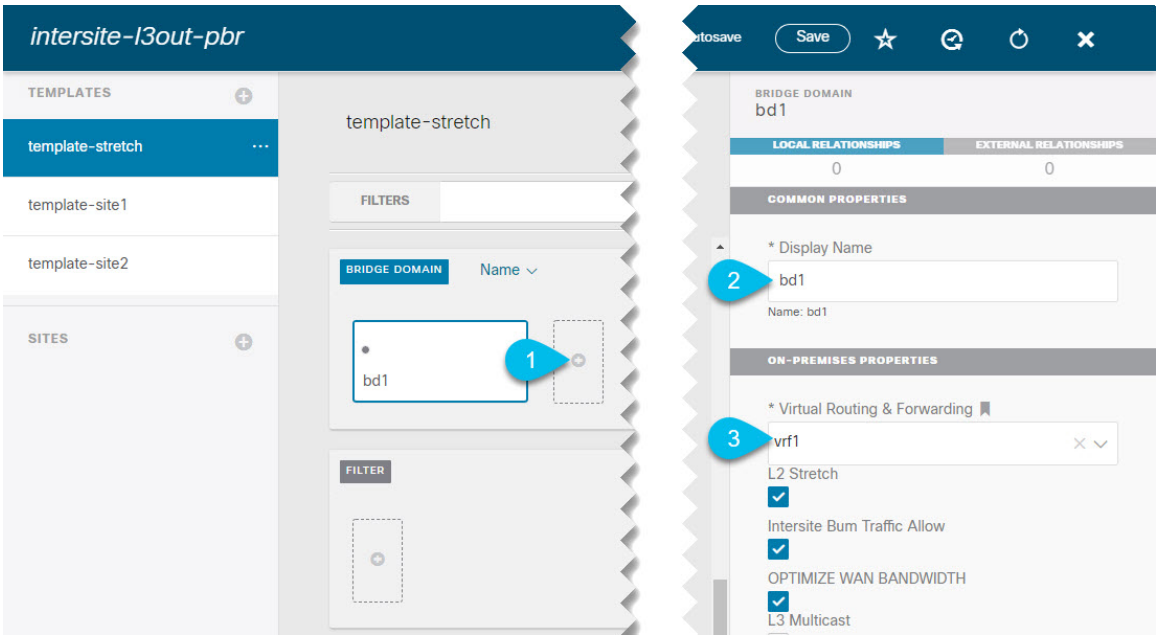
If you are planning to stretch the VRF and BD, select the `template-stretch` template. Otherwise, choose one of the site-specific templates.

Step 2 Create VRF.



a) In the main pane's **VRF** area, click the plus (+) sign to add a VRF.
b) In the right sidebar, provide the **Display Name** for the VRF.
c) Specify other VRF settings as appropriate for your deployment.

Step 3 Create BD.



a) In the main pane's **BD** area, click the plus (+) sign to add a BD.
b) In the right sidebar, provide the **Display Name** for the BD.
c) From the **Virtual Routing & Forwarding** dropdown, select the VRF you created in the previous step.
d) Specify other BD settings as appropriate for your deployment.

Creating Application Profile and EPG

This section describes how to create the application EPG you will later configure to use the intersite L3Out with Service Graph.

Before you begin

You must have:

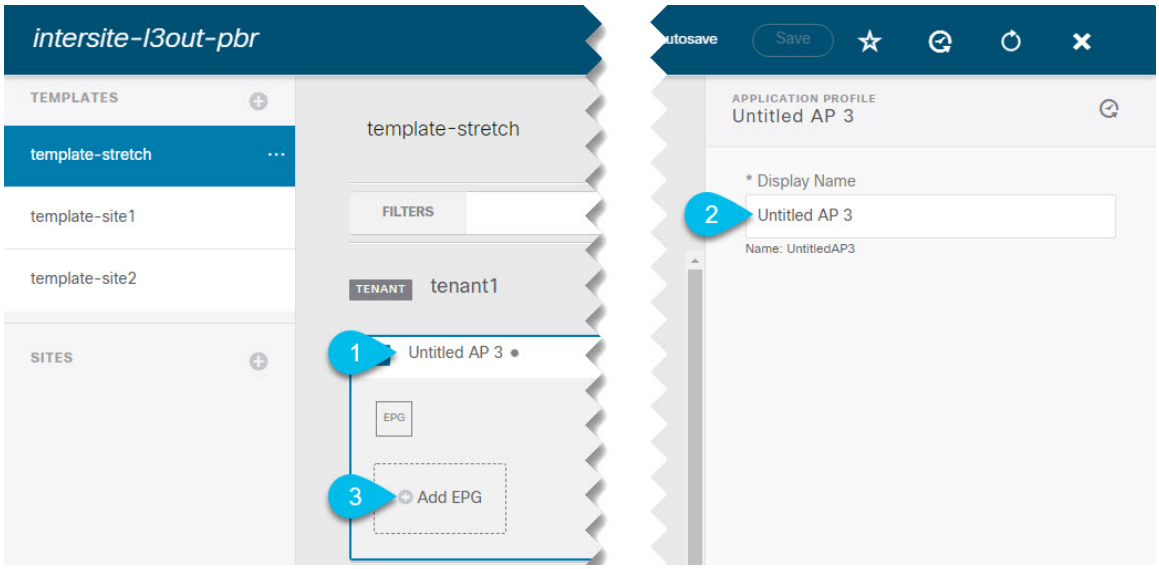
- Created the templates where you will create these objects, as described in Creating Templates, on page 188.
- Created the contract you plan to use for communication between the application EPG and the external EPG, as described in Creating Filter and Contract, on page 192.
- Created the VRF and BD you plan to use for the EPG, as described in Creating VRF and Bridge Domain for Application EPG, on page 198.

Procedure

Step 1 Select the template where you want to create the objects.

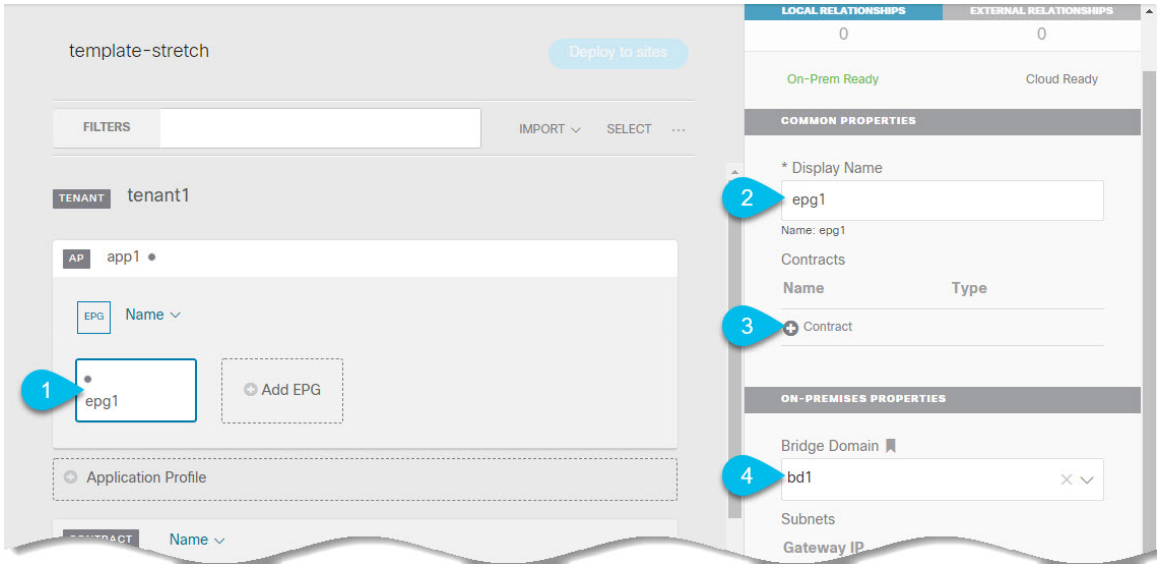
If you plan to stretch the application EPG, create it in the stretched template. If you application EPG is going to be site local, create it in the site-specific template.

Step 2 Create an application profile and EPG.



a) In the main pane, click **+ Application profile**.
b) In the right sidebar, provide the **Display Name** for the profile.
c) In the main pane, click **+Add EPG**.

Step 3 Configure the EPG.



a) In the main pane, select the application EPG.
b) In the right sidebar, provide the **Display Name** for the EPG.
c) Click **+Contract** and select the contract.

Select the contract you have created for the EPG communication and set its type.

If you are using the same VRF for your application EPG and the L3Out external EPG, you can choose either one to be the `consumer` or the `provider`. However, if they are in different VRFs, you must select `consumer` for the application EPG's contract type.

d) From the **Bridge Domain** dropdown, select the BD.
e) Specify other EPG settings as appropriate for your deployment.

Creating L3Out External EPG

Creating or Importing Intersite L3Out and VRF

This section describes how to create an L3Out and associate it to a VRF in the Multi-Site Orchestrator (MSO) GUI, which will then be pushed out to the APIC site, or import an existing L3Out from one of your APIC sites. You will then associate this L3Out with an external EPG and use that external EPG to configure specific intersite L3Out use cases.

Note The VRF you assign to the L3Out can be in any template or schema, but it must be in the same tenant as the L3Out.

Before you begin

You must have:

- Created the templates where you will create these objects, as described in Creating Templates, on page 188.

Procedure

Step 1 Log in to your Multi-Site Orchestrator.

Step 2 From the left navigation pane, select **Application Management** > **Schemas**.

Step 3 Select the schema and then the template where you want to create or import the VRF and L3Out.

If you create the L3Out in a template that is associated to multiple sites, the L3Out will be created on all of those sites. If you create the L3Out in a template that is associated with a single site, the L3Out will be created in that site only.

Step 4 Create a new VRF and L3Out.

If you want to import an existing L3Out, skip this step.

Note
While you can create the L3Out object in the MSO and push it out to the APIC, the physical configuration of the L3Out must be done in the APIC.

a) Scroll down to the **VRF** area and click the + icon to add a new VRF.

In the right sidebar, provide the name for the VRF, for example `vrf-l3out`

b) Scroll down to the **L3Out** area and click the + icon to add a new L3Out.

In the right sidebar, provide the required information.

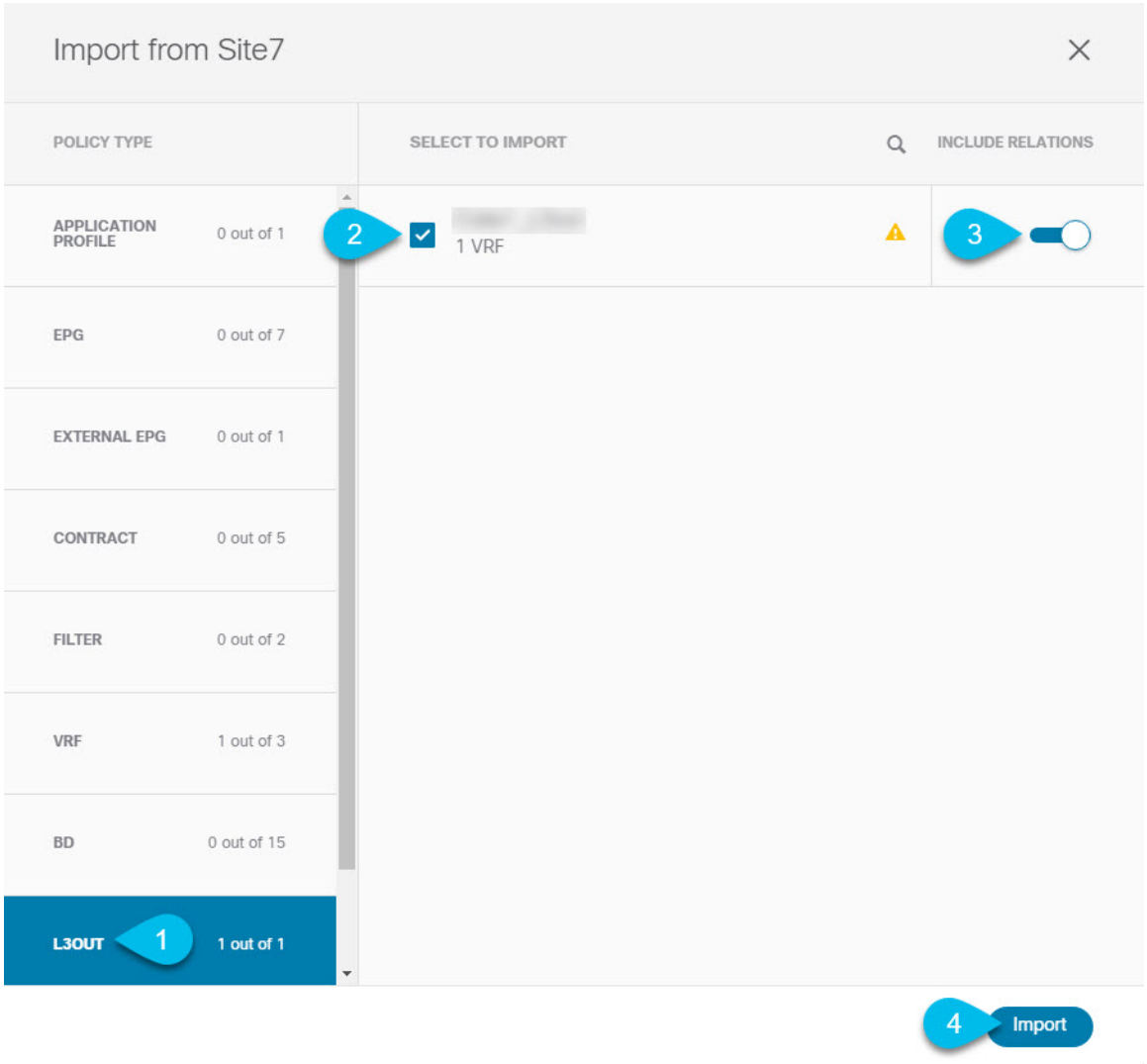
c) Provide the name for the L3Out, for example `l3out-intersite`.

d) From the **Virtual Routing & Forwarding** dropdown, select the VRF you created in the previous step.

Step 5 Import an existing L3Out.

If you created a new L3Out in previous step, skip this step.

At the top of the main template view, click **Import**, then select the site from which you want to import.



a) In the import window's **Policy Type** menu, select **L3Out**.

b) Check the L3Out you want to import.

c) (Optional) If you want to import all objects associated with the L3Out, enable the **Include Relations** knob.

d) Click **Import**.

Configuring External EPG

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.

Before you begin

You must have:

- Created the templates where you will create these objects, as described in Creating Templates, on page 188.
- Created or imported the L3Out and VRF as described in Creating or Importing Intersite L3Out and VRF, on page 201.

Procedure

Step 1 Select the 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 created on all of those sites. 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.

Step 2 Scroll down to the **External EPG** area and click the + icon to add an external EPG.

In the right sidebar, provide the required information.

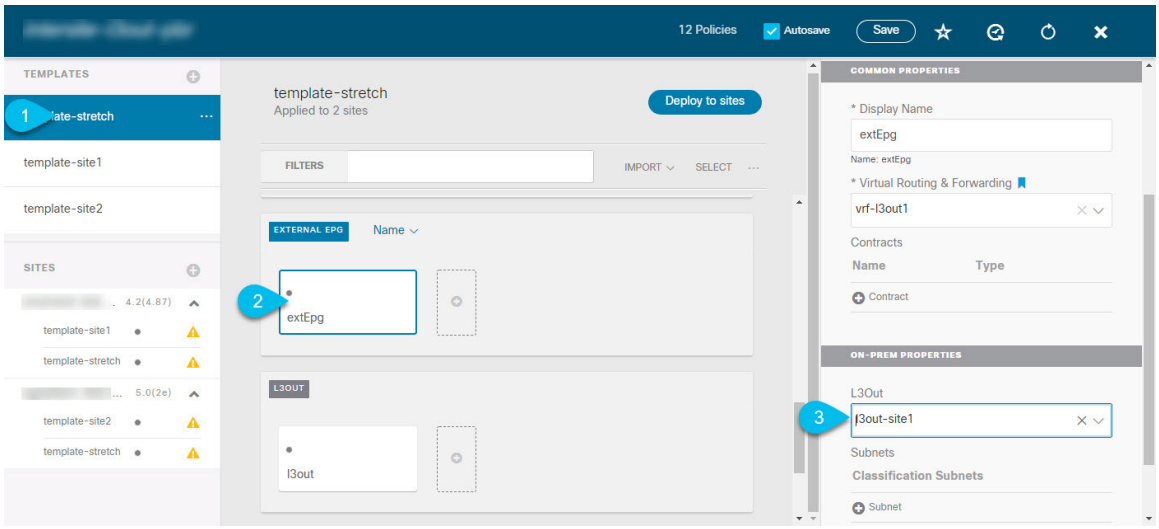
a) Provide the name for the external EPG, for example `extEpg`.
b) From the **Virtual Routing & Forwarding** dropdown, select the VRF you created and used for the L3Out.
c) Click **+Contract** and select the contract.

Select the contract you have created for the EPG communication and set its type.

If you are using the same VRF for your application EPG and the L3Out external EPG, you can choose either one to be the `consumer` or the `provider`. However, if they are in different VRFs, you must select `provider` for the external EPG's contract type.

Step 3 If you want to assign the L3Out at the template level...

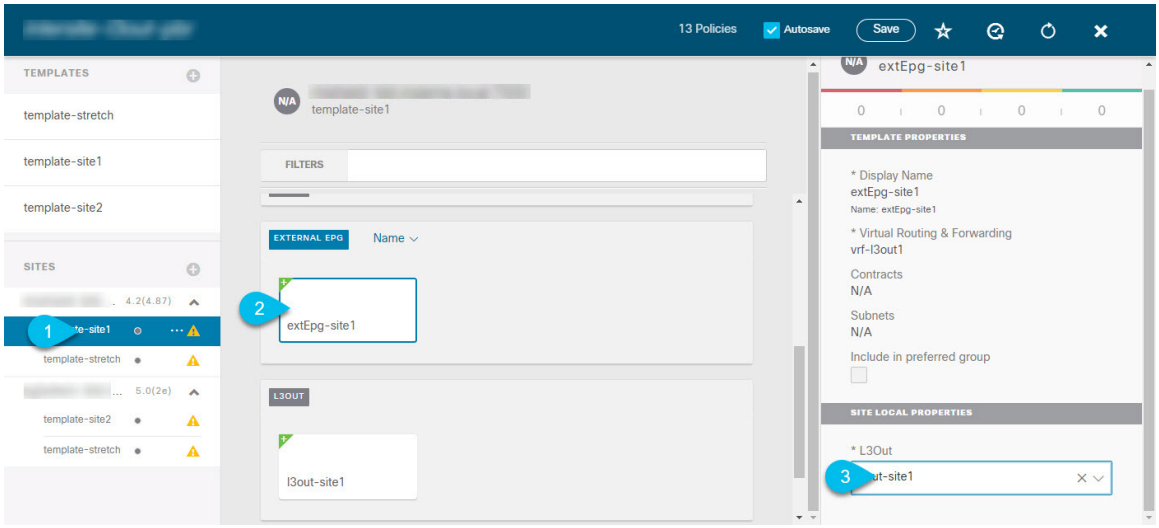
You can choose to configure the L3Out for the external EPG at the template level, in which case, you will not be able to set the L3Outs at the site-local level.



a) In the left sidebar of the schema view, select the template where the external EPG is located
b) Scroll down to the **External EPG** area and select the external EPG.
c) In the right sidebar, scroll down to the **L3Out** dropdown and choose the intersite L3Out you created.

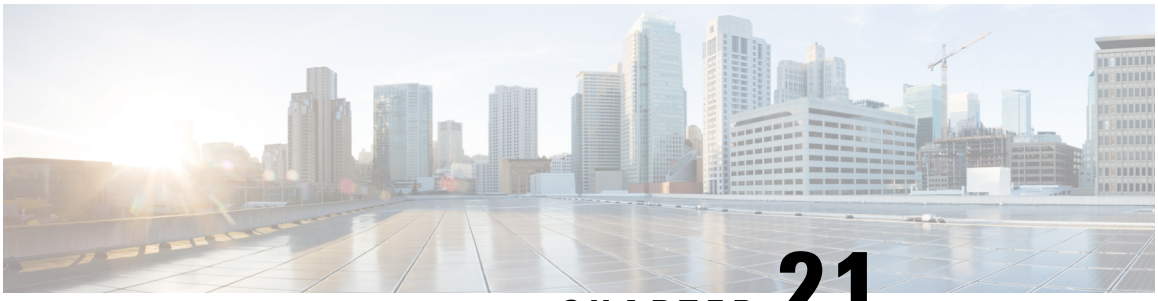
Step 4 If you want to assign the L3Out at the site local level...

Alternatively, you can choose to associate an L3Out with the external EPG at the site-local level.



a) In the left sidebar of the schema view, select the site where the external EPG is deployed.
b) Scroll down to the **External EPG** area and select the external EPG.
c) In the right sidebar, scroll down to the **L3Out** dropdown and choose the intersite L3Out you created.

In this case, both the APIC-managed and the Orchestrator-managed L3Outs will be available for selection. You can select either the L3Out you have created in the previous section specifically for this or pick an L3Out that exists in the site's APIC.



CHAPTER 21

Layer 3 Multicast

- Layer 3 Multicast, on page 205
- Layer 3 Multicast Routing, on page 206
- Rendezvous Points, on page 206
- Multicast Filtering, on page 207
- Layer 3 Multicast Guidelines and Limitations, on page 208
- Creating Multicast Route Map Policy, on page 209
- Enabling Any-Source Multicast (ASM) Multicast, on page 210
- Enabling Source-Specific Multicast (SSM), on page 212

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 Multi-Site 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 Multi-Site 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 Multi-Site 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 fails, 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 Multi-Site 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 209, while VRF configuration is described in Enabling Any-Source Multicast (ASM) Multicast, on page 210.

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 Multi-Site 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 Multi-Site 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 Multi-Site 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 defined to limit the RP to those groups only.

Procedure

Step 1 Log in to the Multi-Site Orchestrator GUI.

Step 2 In the **Main menu**, select **Application Management** > **Policies**.

Step 3 In the main pane, select **Add Policy** > **Create Multicast Route-Map Policy**.

Step 4 In the **Add Multicast Route-Map Policy** screen, select a Tenant and provide the name for the policy.

Step 5 Under **Route-Map Entry Order**, click **Add Route-Map Entry** to add an entry.

a) Provide the **Order** and **Action**.

Each entry is a rule that defines an action based on one or more matching criteria.

Order is used to determine the order in which the rules are evaluated.

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

b) Provide the **Group IP**, **Source IP**, and **RP IP** information as required.

As mentioned at the start of this section, 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 only need to 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 will match all values.

  The Group IP range must be between `224.0.0.0` and `239.255.255.255` with a netmask between `/8` 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 `/8` and `/32`. You must provide the subnet mask.

  For 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.

c) Click **Save** to save the entry.

Step 6 (Optional) Repeat the previous step if you want to add multiple entries to the same route policy.

Step 7 Click **Save** to save the route map policy.

Enabling Any-Source Multicast (ASM) Multicast

The following procedure describes how to enable ASM multicast on VRF, BD, and EPG using the Multi-Site Orchestrator GUI. If you want to enable SSM multicast, follow the steps in Enabling Source-Specific Multicast (SSM), on page 212 instead.

Before you begin

- Ensure you have read and followed the information described in Layer 3 Multicast Guidelines and Limitations, on page 208.
- If you plan to enable multicast filtering, create the required multicast route maps, as described in Creating Multicast Route Map Policy, on page 209.
- 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 Multi-Site Orchestrator.

Step 2 From the left-hand sidebar, select the **Application Management** > **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 Multi-Site 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.

Once 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** checkbox.

Step 8 (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 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.

a) Select the EPG for which you want to enable Layer 3 multicast.
b) 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 Multi-Site Orchestrator GUI. If you want to enable ASM multicast, follow the steps in Enabling Any-Source Multicast (ASM) Multicast, on page 210 instead.

Before you begin

- Ensure you have read and followed the information described in Layer 3 Multicast Guidelines and Limitations, on page 208.
- If you plan to enable multicast filtering, create the required multicast route maps, as described in Creating Multicast Route Map Policy, on page 209.
- Note that you need to 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 Multi-Site Orchestrator.

Step 2 From the left-hand sidebar, select the **Application Management** > **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 (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 additional SSM ranges under the VRF settings by creating a route-map with up to 4 additional ranges. Keep in mind that if you add a new range it will become 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** dropdown 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 dropdown 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 need to 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 and/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 Multi-Site 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.

Once 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** checkbox.

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** dropdown, 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 dropdown 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**.

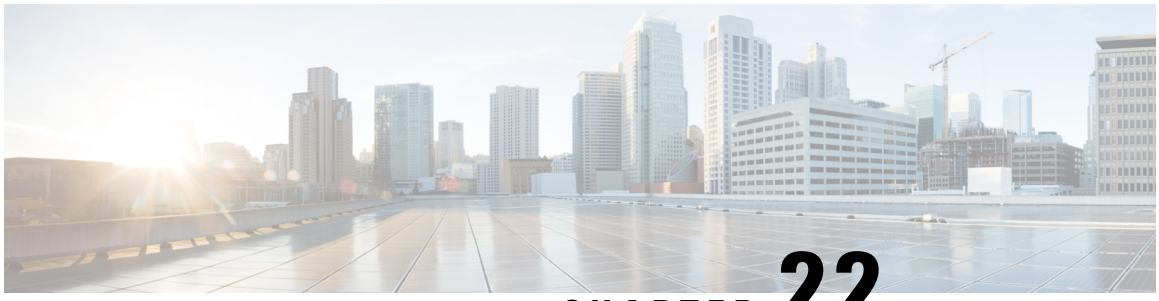
You can choose to enable either the multicast source filtering, or the receiver filtering, or both.

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 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.

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.

a) Select the EPG for which you want to enable Layer 3 multicast.
b) In the right-hand sidebar, check the **Intersite Multicast Source** checkbox.



CHAPTER 22

EPG Preferred Group

- EPG Preferred Groups, on page 215
- Configuring EPGs for Preferred Group, on page 216

EPG Preferred Groups

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 Multi-Site 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.

On the contrary, Multi-Site Orchestrator does 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 MSO, MSO 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, MSO automatically enables the PG setting on the VRF. When you remove the last EPG from the preferred group, MSO 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 MSO.

  MSO 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, MSO will manage the object as if it was created from MSO and overwrite the PG setting dynamically based on EPG membership.

Limitations

Preferred Groups are not supported for intersite L3Out external EPGs.

Configuring EPGs for Preferred Group

Before you begin

You must have one or more EPGs added to a schema template.

Procedure

Step 1 Log in to your Multi-Site Orchestrator.

Step 2 From the left navigation pane, select the **Schemas** view.

Step 3 Click the Schema that you want to change.

Step 4 Configure one or more EPGs in the schema to be part of the preferred group.

Note
If you have an existing preferred group in any of the APICs and are planning to import the EPGs from that preferred group into Multi-Site Orchestrator, you must import all EPGs in the group. You must not have a preferred group where some EPGs are managed by the Multi-Site Orchestrator and some are managed by the local APIC.

To add or remove a single EPG:

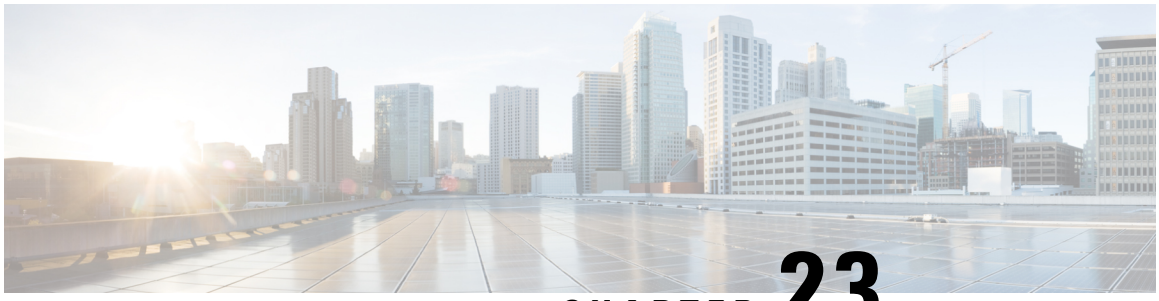
a) Select an EPG.
b) In the right properties bar, check or uncheck the **Include in Preferred Group** checkbox.
c) Click **SAVE** in the top right corner of the main window.

To add or remove multiple EPGs at once:

a) Click **SELECT** in the top-right corner of the **Application Profile** tab.
b) Select one or more EPGs by clicking on each one or click **Select All** to select all EPGs.
c) Click **...** in the top-right corner of the **Application Profile** tab and choose **Add EPGs to Preferred Group** or **Remove EPGs from Preferred Group**.
d) Click **SAVE** in the top right corner of the main window.

What to do next

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 23

QoS Preservation Across IPN

- QoS and Global DSCP Policy, on page 219
- DSCP Policy Guidelines and Limitations, on page 219
- Configuring Global DSCP Policy, on page 220
- Set QoS Level for EPGs and Contracts, on page 222

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 Multi-Site 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 225 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 towards an intersite network, for example in multi-pod and remote leaf 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

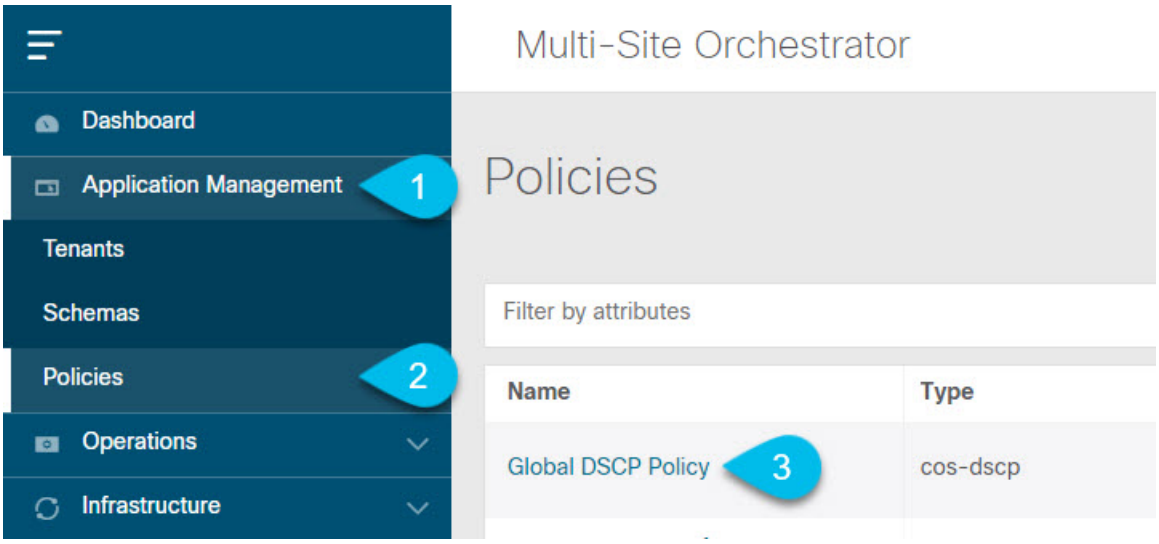
- You should be familiar with Quality of Service (QoS) functionality within ACI fabrics.

  QoS is described in more detail in *Cisco APIC and QoS*.

Procedure

Step 1 Log in to your Cisco Multi-Site Orchestrator GUI.

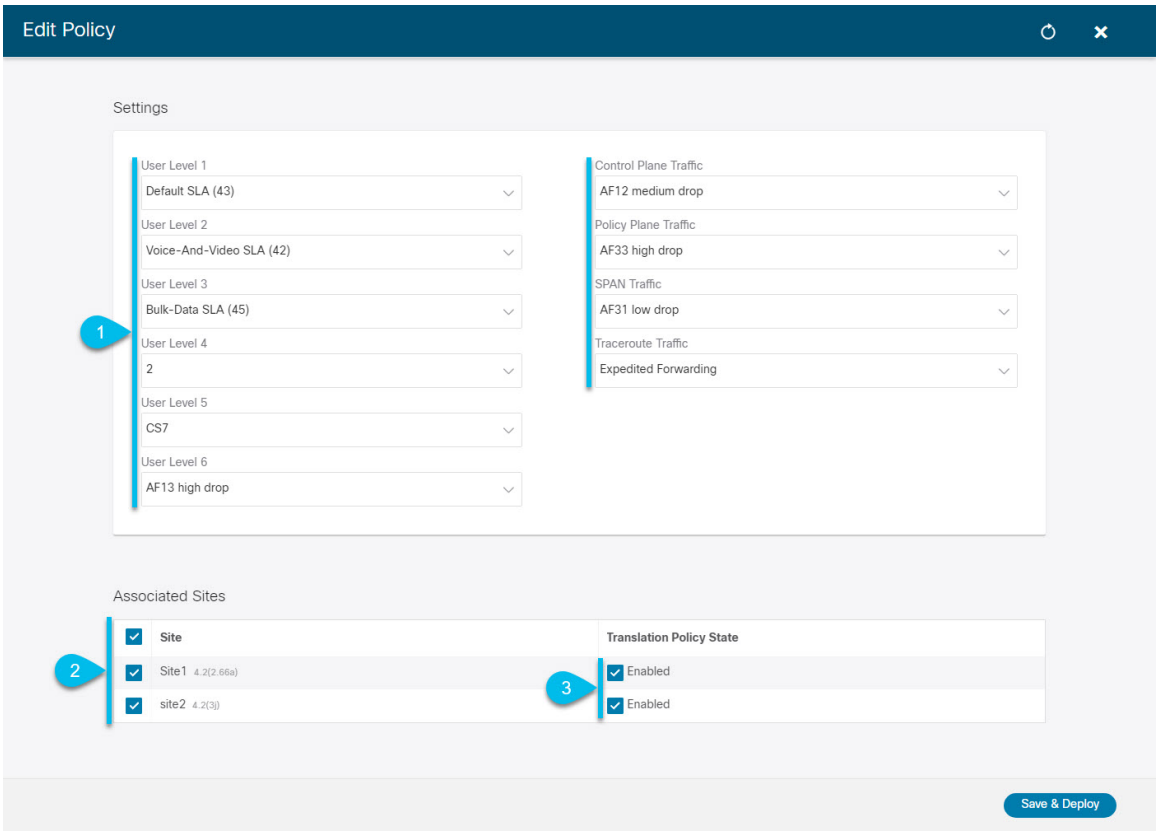
Step 2 Open the global DSCP policy configuration screen.



a) Navigate to **Application Management** > **Policies**.
b) Click **Global DSCP Policy** name.

The **Edit Policy** window will open.

Step 3 Update the global DSCP policy.



a) Choose the DSCP value for each ACI QoS level.

Each dropdown contains the default list of available DSCP values. You must choose a unique DSCP value for each level.

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 logging 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 222.

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

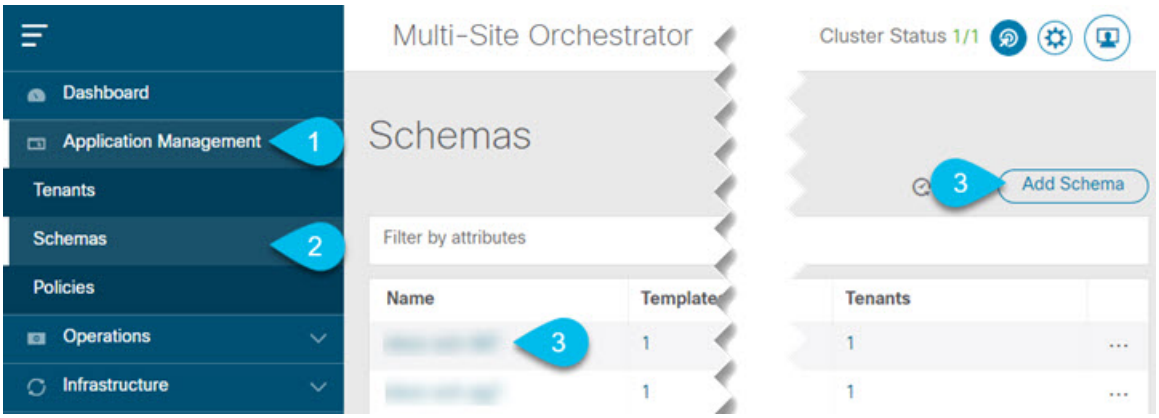
- You must have defined the global DSCP policy, as described in Configuring Global DSCP Policy, on page 220.
- You should be familiar with Quality of Service (QoS) functionality within ACI fabrics.

  QoS is described in more detail in *Cisco APIC and QoS*.

Procedure

Step 1 Log in to your Cisco Multi-Site Orchestrator GUI.

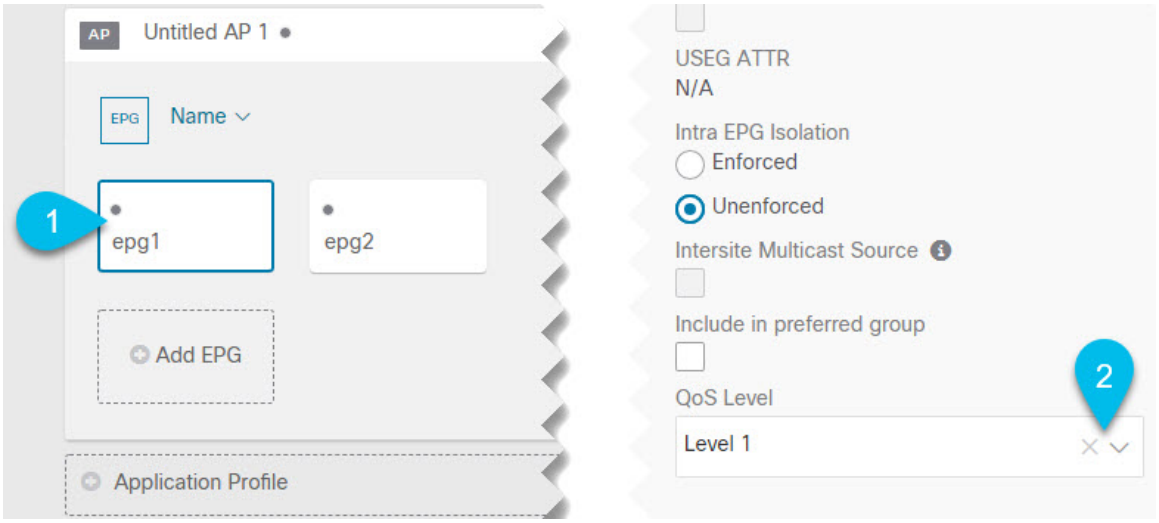
Step 2 Choose the Schema you want to edit.



a) Navigate to **Application Management** > **Schemas** > **.**
b) Click the name of the schema you want to edit or **Add Schema** to create a new one.

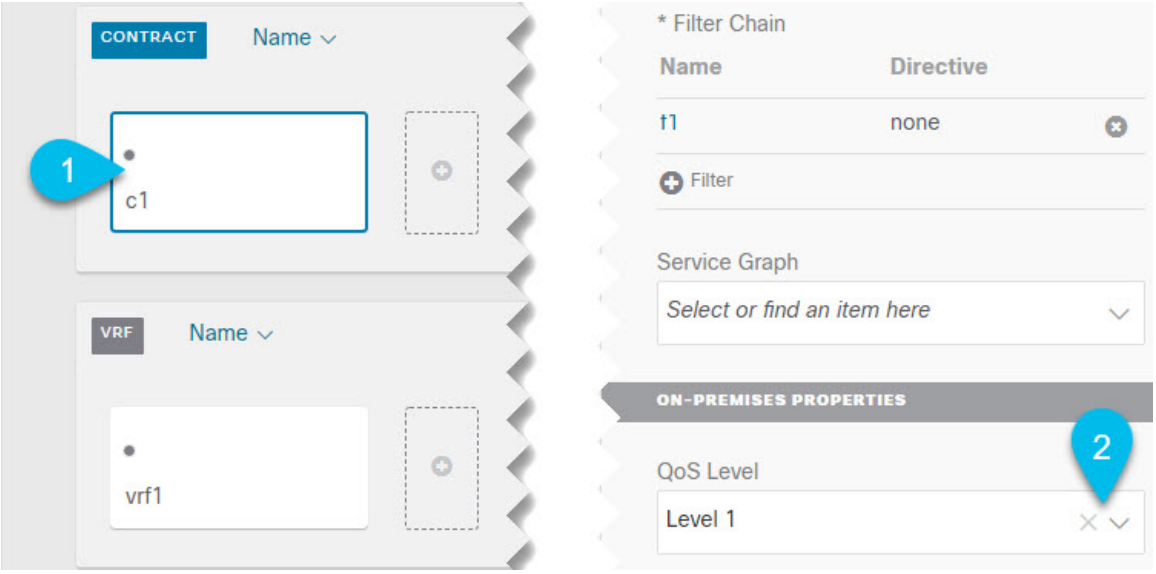
The **Edit Policy** window will open.

Step 3 Pick a QoS Level for an EPG

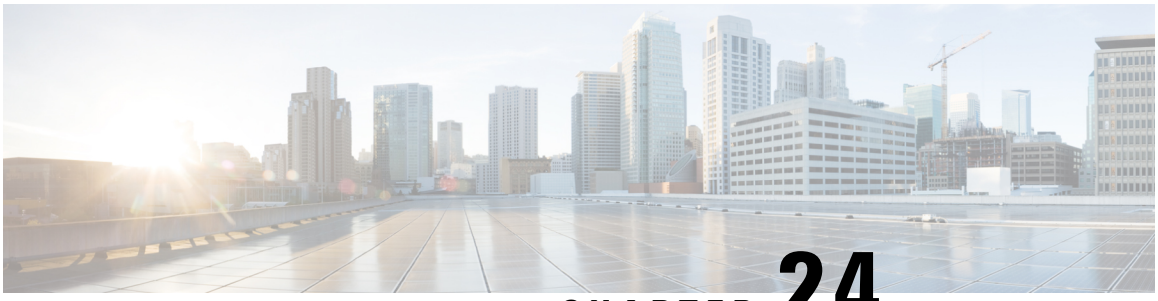


a) In the main pane, scroll down to the **EPG** area and select an EPG or click **Add EPG** to create a new one.
b) In the right sidebar, scroll down to the **QoS Level** dropdown and choose the QoS Level you want to assign to the EPG.

Step 4 Pick a QoS Level for an EPG



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** dropdown and choose the QoS Level you want to assign to the Contract.



CHAPTER 24

SD-WAN Integration

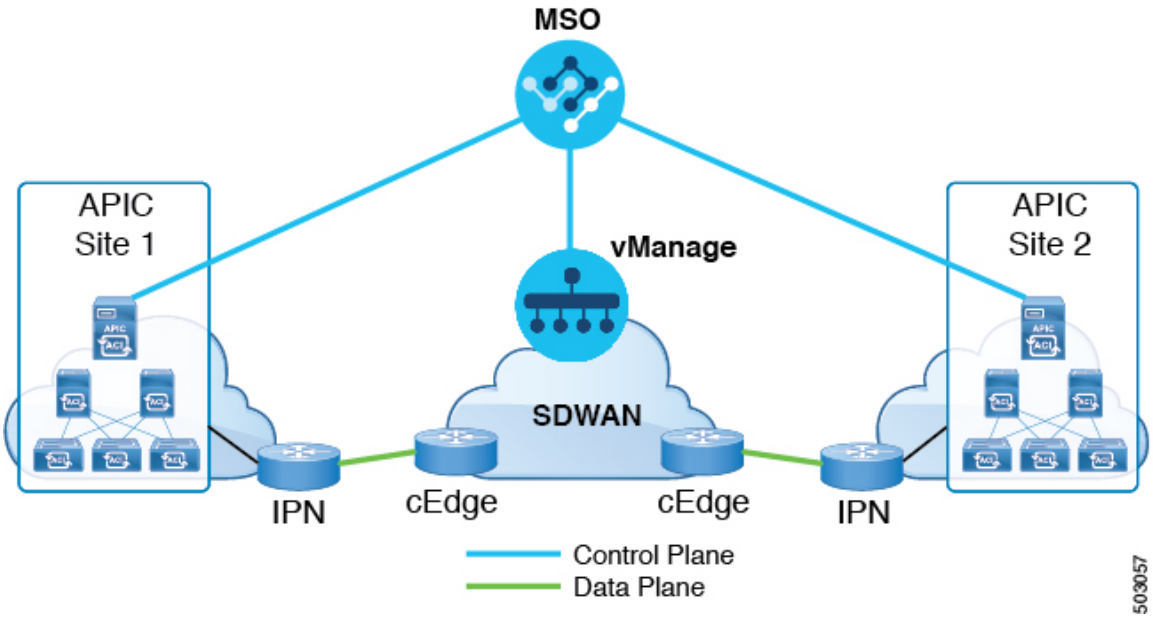
- SD-WAN Integration, on page 225
- SD-WAN Integration Guidelines and Limitations, on page 226
- Adding a vManage Controller, on page 227
- Configuring Global DSCP Policy, on page 228
- Set QoS Level for EPGs and Contracts, on page 231

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 28: Multi-Site and SD-WAN Integration



Release 3.0(2) of Cisco Multi-Site Orchestrator adds support for SD-WAN integration. You can configure the MSO 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 MSO to vManage, it imports the SLA policies allowing you to translating the ACI QoS levels to the appropriate DSCP values. MSO 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 MSO 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 Muilti-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 Multi-Site Orchestrator deployments in Cisco Application Services Engine only.

  For more information, see the Deployment Overview chapter in the *Cisco Multi-Site 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.

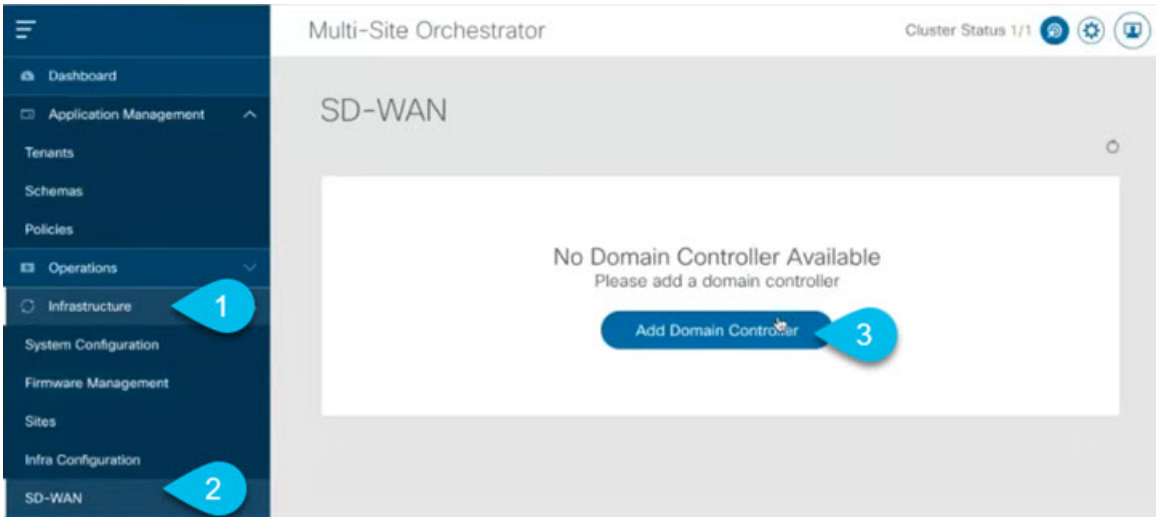
Adding a vManage Controller

This section describes how to add vManage controller to your Multi-Site Orchestrator in order to import any configured SLA policies.

Procedure

Step 1 Log in to your Cisco Multi-Site Orchestrator GUI.

Step 2 Add a vManage Controller.



a) Navigate to **Infrastructure** > **SD-WAN**.
b) Click **Add Domain Controller** name.

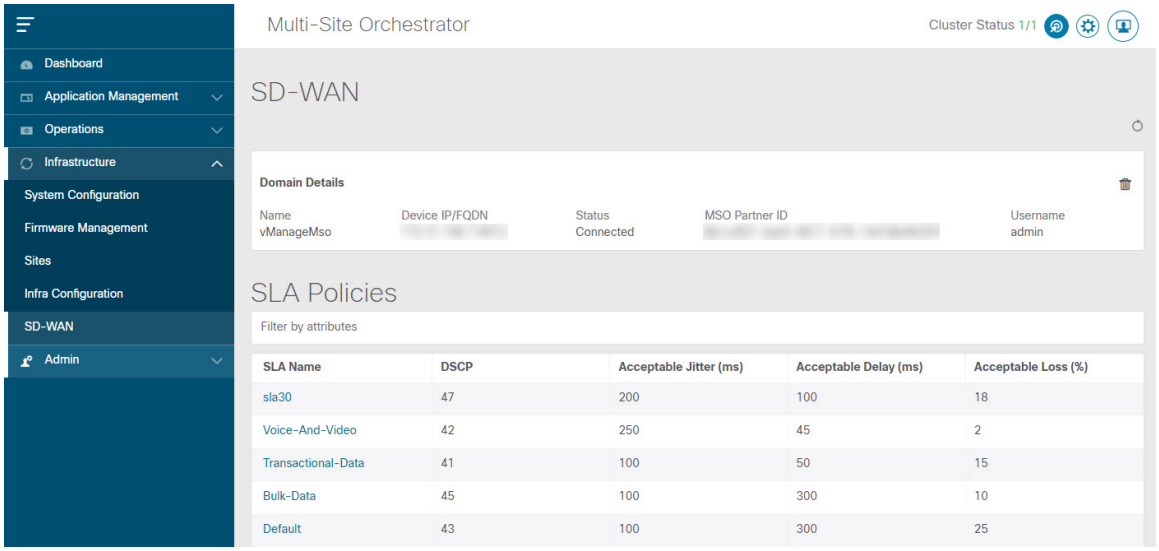
The **Add Domain** window will open.

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 MSO.
- The device's fully qualified domain name or IP address.
- Username and password used to log 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 one min before the list of existing SLA policies is displayed in the main pane:



What to do next

Define the global DSCP policy in your Multi-Site Orchestrator, as described in Configuring Global DSCP Policy, on page 228

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 towards 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

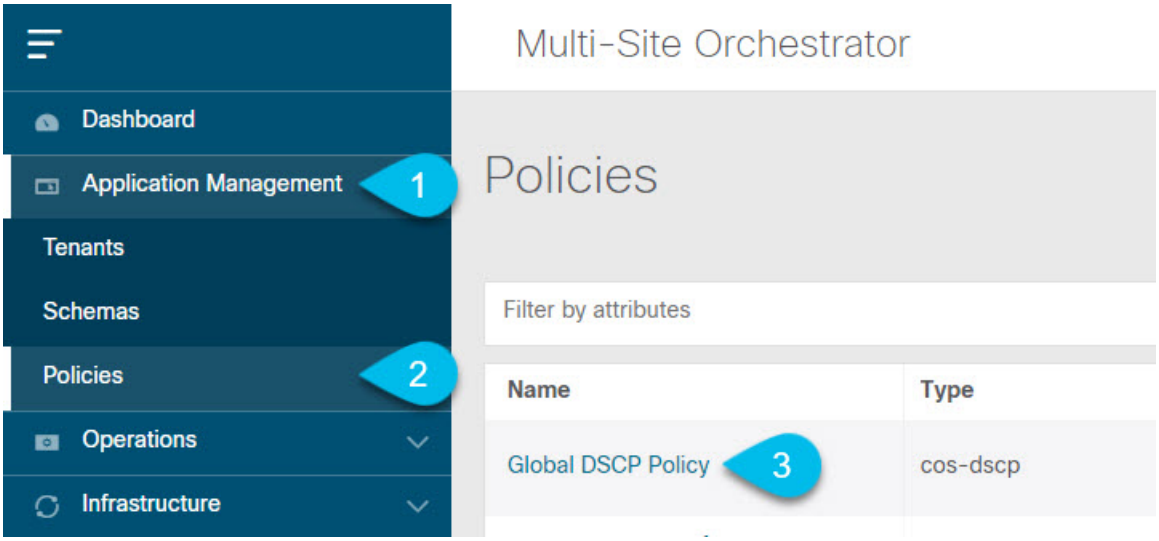
- You must have added a vManage controller to your MSO, as described in Adding a vManage Controller, on page 227.
- You should be familiar with Quality of Service (QoS) functionality within ACI fabrics.

  QoS is described in more detail in *Cisco APIC and QoS*.

Procedure

Step 1 Log in to your Cisco Multi-Site Orchestrator GUI.

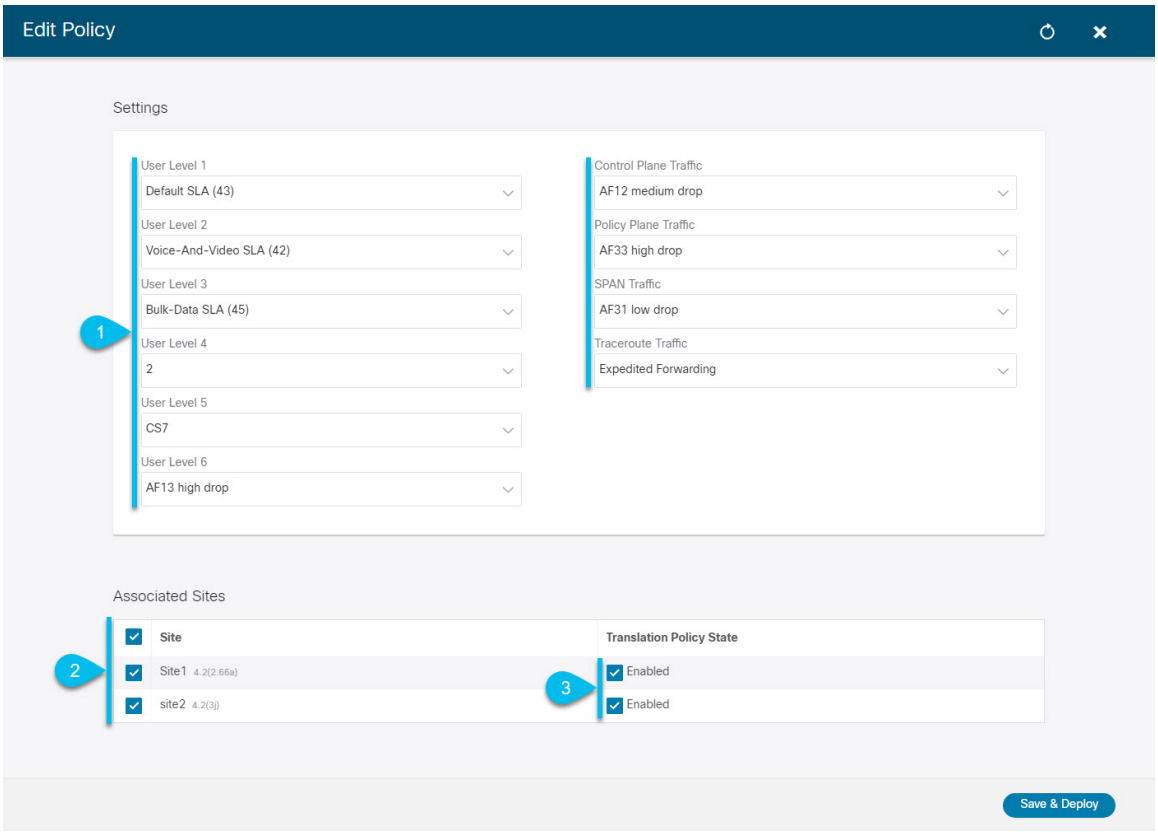
Step 2 Open the global DSCP policy configuration screen.



a) Navigate to **Application Management** > **Policies**.
b) Click **Global DSCP Policy** name.

The **Edit Policy** window will open.

Step 3 Update the global DSCP policy.



a) Choose the DSCP value for each ACI QoS level.

   Each dropdown contains the default list of available DSCP values as well as 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 logging 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 231

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

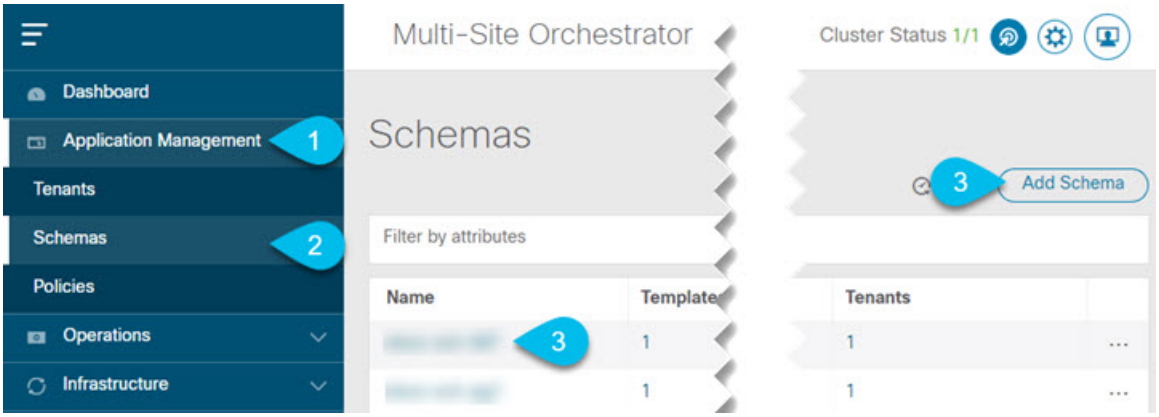
- You must have added a vManage controller to your MSO, as described in Adding a vManage Controller, on page 227.
- You must have defined the global DSCP policy, as described in Configuring Global DSCP Policy, on page 228.
- You should be familiar with Quality of Service (QoS) functionality within ACI fabrics.

  QoS is described in more detail in *Cisco APIC and QoS*.

Procedure

Step 1 Log in to your Cisco Multi-Site Orchestrator GUI.

Step 2 Choose the Schema you want to edit.

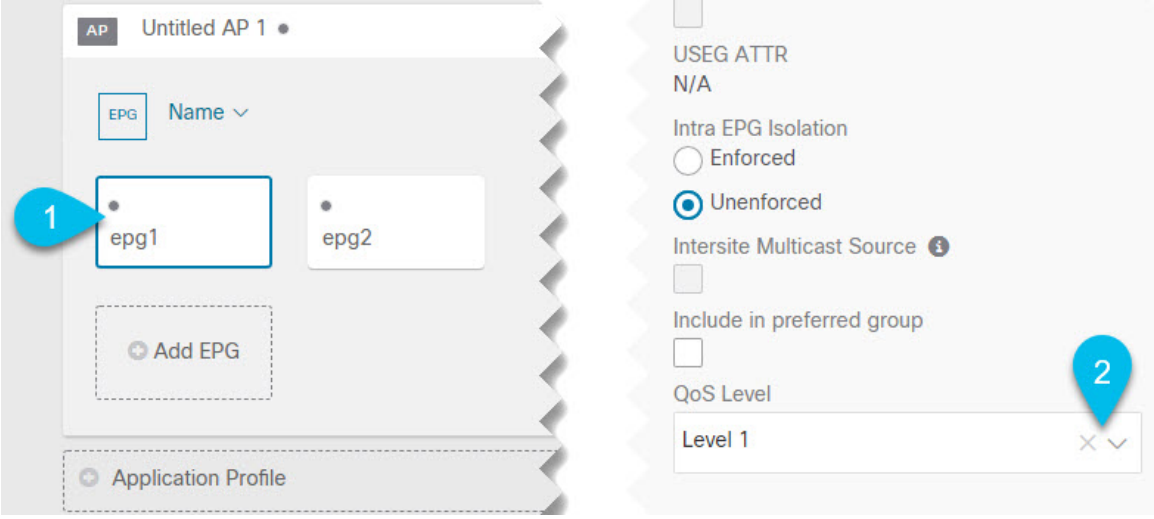


a) Navigate to **Application Management** > **Schemas** > **.**

b) Click the name of the schema you want to edit or **Add Schema** to create a new one.

The **Edit Schema** window will open.

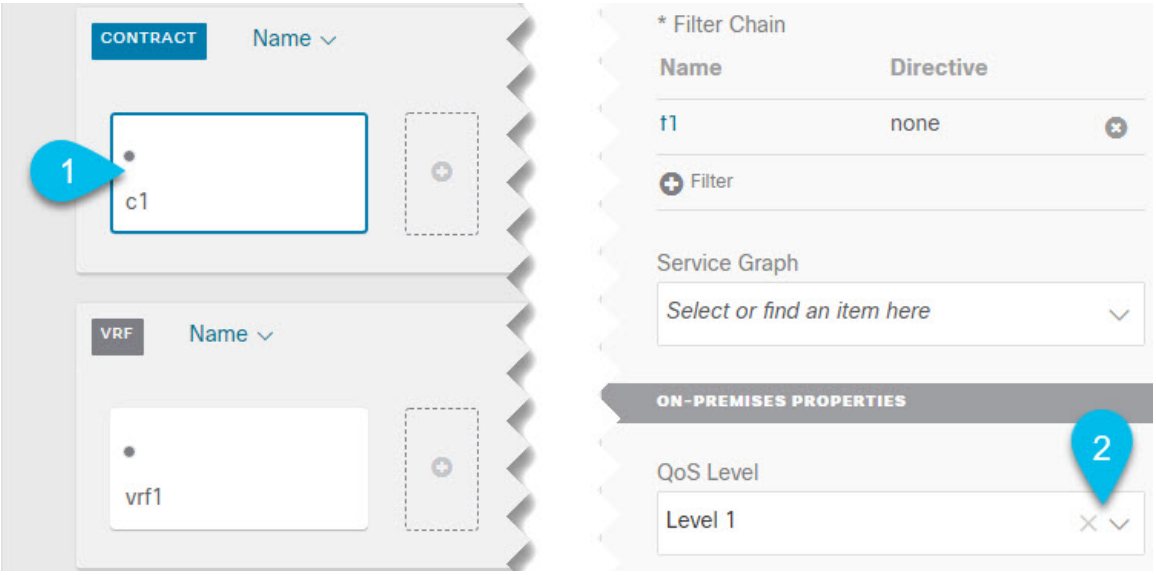
Step 3 Pick a QoS Level for an EPG



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** dropdown and choose the QoS Level you want to assign to the EPG.

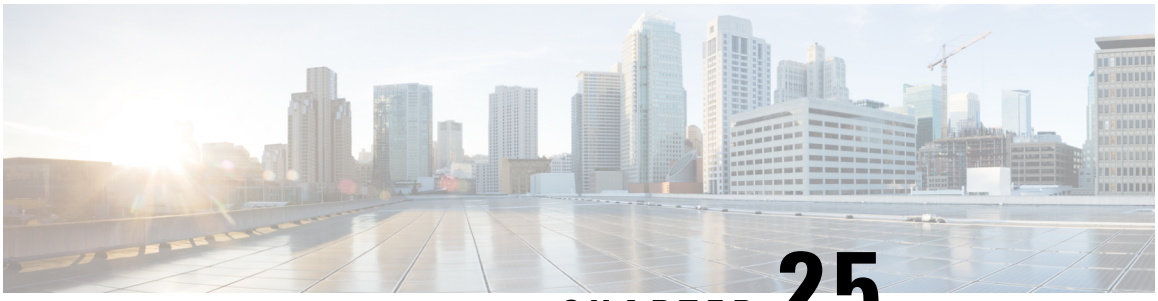
You must choose the QoS level based on the previously configured Global DSCP policy to ensure that intersite traffic from the EPG is treated with the desired SLA across the SD-WAN network.

Step 4 Pick a QoS Level for an EPG



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** dropdown 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 desired SLA across the SD-WAN network.



CHAPTER 25

Sites Connected via SR-MPLS

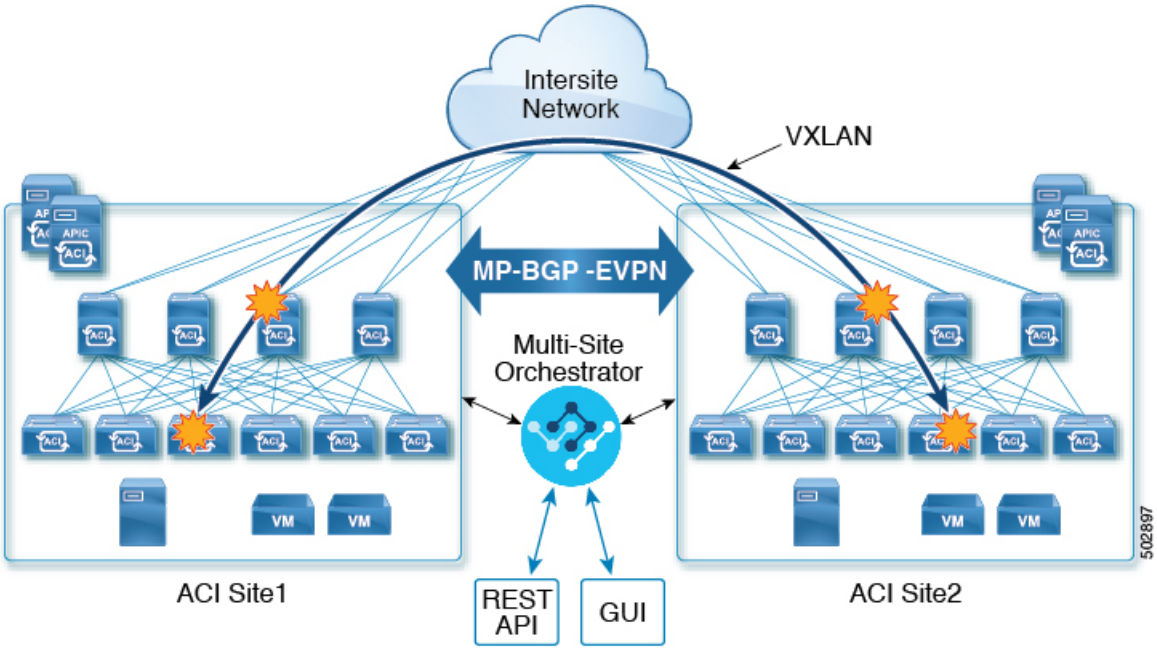
- SR-MPLS and Multi-Site, on page 233
- Infra Configuration, on page 235
- SR-MPLS Tenant Requirements and Guidelines, on page 242
- Creating SR-MPLS Route Map Policy, on page 244
- Enabling Template for SR-MPLS, on page 246
- Creating VRF and SR-MPLS L3Out, on page 246
- Configuring Site-Local VRF Settings, on page 247
- Configuring Site-Local SR-MPLS L3Out Settings, on page 247
- Communication Between EPGs Separated by MPLS Network, on page 248
- Deploying Configuration, on page 249

SR-MPLS and Multi-Site

Starting with Multi-Site Orchestrator, Release 3.0(1) and APIC Release 5.0(1), the Multi-Site architecture supports APIC sites connected via MPLS networks.

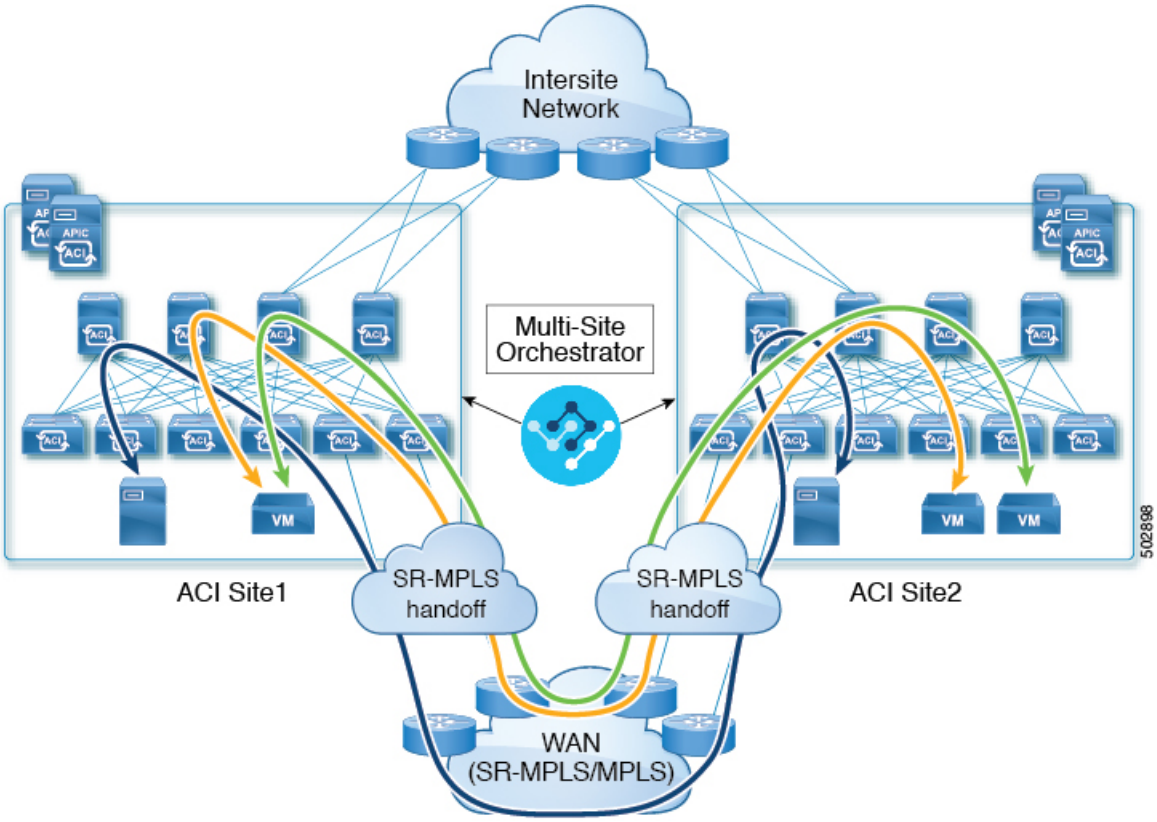
In a typical Multi-Site deployment, traffic between sites is forwarded over an intersite network (ISN) via VXLAN encapsulation:

Figure 29: 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:

Figure 30: Multi-Site and MPLS



The following sections describe guidelines, limitations, and configurations specific to managing Schemas that are deployed to these sites from the Multi-Site 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*.

Infra Configuration

SR-MPLS Infra Guidelines and Limitations

If you want to add an APIC site that is connected to an SR-MPLS network to be managed by the Multi-Site Orchestrator, keep the following in mind:

- 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 119.
- Objects and policies deployed to a site that is connected to an SR-MPLS network cannot be stretched to other sites.

  When you create a template and specify a Tenant, you will need to enable the `SR-MPLS` option on the tenant. You will then be able to map that template only to a single ACI site.

- Tenants deployed to a site that is connected via an SR-MPLS network will have a set of unique configuration options specifically for SR-MPLS configuration. For information about configuring tenants, see Tenants, on page 13.

Supported Hardware

The SR-MPLS connectivity is supported for the following platforms:

- **Border Leaf switches**: The "FX", "FX2", and "GX" 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 and N9K-C9364C 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.
- 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.
- 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.

Guidelines and Limitations for 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.

Creating SR-MPLS QoS Policy

This section describes how to configure SR-MPLS QoS policy for a site that is connected via an MPLS network. If you have no such sites, you can skip this section.

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.

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.

Procedure

Step 1 Log in to the Multi-Site Orchestrator GUI.

Step 2 In the **Main menu**, select **Application Management** > **Policies**.

Step 3 In the main pane, select **Add Policy** > **Create QoS Policy**.

Step 4 In the **Add QoS Policy** screen, provide the name for the policy.

Step 5 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, as well as to set differentiated services code point (DSCP) values in the VXLAN header for the packet while it's inside the ACI fabric.

The values are derived at the border leaf using a custom QoS translation policy. The original DSCP values for traffic coming from SR-MPLS without any remarking. If a custom policy is not defined or not matched, default QoS Level (`Level3`) 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** dropdown, select the ACI QoS Level to map.

   This is the QoS Level you want to assign for the traffic within ACI fabric, which ACI uses to prioritize the traffic within the fabric.. The options range from Level1 to Level6. The default value is `Level3`. If you do not make a selection in this field, the traffic will automatically be assigned a `Level3` priority.

c) From the **Set DSCP** dropdown, select the DSCP value to assign to the packet when it's inside the ACI fabric.

   The DSCP value specified is set in the original traffic received from the external network, so it will be re-exposed 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** dropdown, select the CoS value to assign to the packet when it's inside the ACI fabric.

   The CoS value specified is set in the original traffic received from the external network, so it will be re-exposed 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, but only if the CoS preservation option is enabled in the fabric. For more information about CoS preservation, see *Cisco APIC and QoS*.

e) Click the checkmark icon to save the rule.
f) Repeat this step for any additional ingress QoS policy rules.

Step 6 Click **Add Egress Rule** to add an egress QoS translation rule.

These rules are applied for the traffic that is leaving the ACI fabric via an MPLS L3Out and are used to map the packet's IPv4 DSCP value to the MPLS packet's EXP value as well as the internal ethernet frame's CoS value.

Classification is done at the non-border leaf switch based on existing policies 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

a) 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.
b) From the **Set MPLS EXP** dropdown, select the EXP value you want to assign to the egressing MPLS packet.
c) From the **Set CoS** dropdown, select the CoS value you want to assign to the egressing MPLS packet.
d) Click the checkmark icon to save the rule.
e) Repeat this step for any additional egress QoS policy rules.

Step 7 Click **Save** to save the QoS policy.

What to do next

After you have created the QoS policy, enable MPLS connectivity and configure MPLS L3Out as described in #unique_179.

Creating SR-MPLS Infra L3Out

This section describes how to configure SR-MPLS 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 handoff.
- 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 via SR-MPLS network as described in Adding Cisco ACI Sites, on page 111.
- If necessary, created SR-MPLS QoS policy as described in Creating SR-MPLS QoS Policy, on page 237.

Procedure

Step 1 Log in to the Cisco Multi-Site Orchestrator GUI.

Step 2 Ensure that SR-MPLS Connectivity is enabled for the site.

a) In the main navigation menu, select **Infrastructure** > **Infra Configuration**.
b) In the **Infra Configuration** view, click **Configure Infra**.
c) In the left pane, under **Sites**, select a specific site.
d) In the right **<*Site*> Settings** pane, enable the **SR-MPLS Connectivity** knob and provide the Segment Routing global block (SRGB) range

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.

The Segment Routing Global Block (SRGB) is the range of label values reserved for Segment Routing (SR) in the Label Switching Database (LSD). 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.

The default range is `16000-23999`.

Step 3 In the main pane, click **+Add SR-MPLS L3Out** within a pod.

Step 4 In the right **Properties** pane, provide a name for the SR-MPLS L3Out.

Step 5 (Optional) From the **QoS Policy** dropdown, select a QoS Policy you created for SR-MPLS traffic.

Select the QoS policy you created in Creating SR-MPLS QoS Policy, on page 237.

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 6 From the **L3 Domain** dropdown, select the Layer 3 domain.

Step 7 Configure BGP settings.

You must provide BGP connectivity details for the BGP EVPN connection between the site's border leaf (BL) switches and the provider edge (PE) router.

a) Click **+Add BGP Connectivity**.

b) In the **Add BGP 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 connected directly to the border leaf.

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 from 1 to 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 and the DC-PE router, for example `10`. The allowed range `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 Configure settings for border leaf switches and ports connected to the SR-MPLS network.

You need to provide information about the border leaf switches as well as the interface ports which connect to the SR-MPLS network.

a) Click **+Add Leaf** to add a leaf switch.

b) In the **Add Leaf** window, select the leaf switch from the **Leaf Name** dropdown.

c) Provide a valid segment ID (SID) offset.

When configuring the interface ports later in this section, you will be 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.

- The value must be within the SRGB range you configured earlier.
- The value must be the same for the selected leaf switch across all SR-MPLS L3Outs in the site.
- The same value cannot be used for more than one leaf across all sites.
- If you need to update the value, you must first delete it from all SR-MPLS L3Outs in the leaf and re-deploy the configuration. Then you can update it with the new value, followed by re-deploying the new configuration.

d) Provide the local **Router ID**.

Unique router identifier within the fabric.

e) Provide the **BGP EVPN Loopback** address.

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 loopbacks of the border leaf switch and the DC-PE to advertise the overlay prefixes. The MP-BGP EVPN sessions are established between the BP-EVPN loopback and the BGP-EVPN remote peer address (configured in the **MPLS BGP-EVPN Peer IPv4 Address** field in the **BGP Connectivity** step before).

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 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** dropdown, select whether it is a typical interface or a port channel. 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 sub-interface, 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 connected directly to the interface. The next hop address must be part of the subnet configured for the interface.

Choose whether you want to enable segment routing (SR) MPLS.

(Optional) Choose to enable the additional BGP options based on your deployment.

Finally, click the checkmark to the right of the **Interface Type** dropdown to save interface port information.

h) Repeat the previous sub-step for all interfaces on the switch that connect to the MPLS network.

i) Click **Save** to save the leaf switch information.

Step 9 Repeat the previous step for all leaf switches connected to MPLS networks.

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 *Multi-SiteConfiguration Guide, Release 3.0(x)*.

SR-MPLS Tenant Requirements and Guidelines

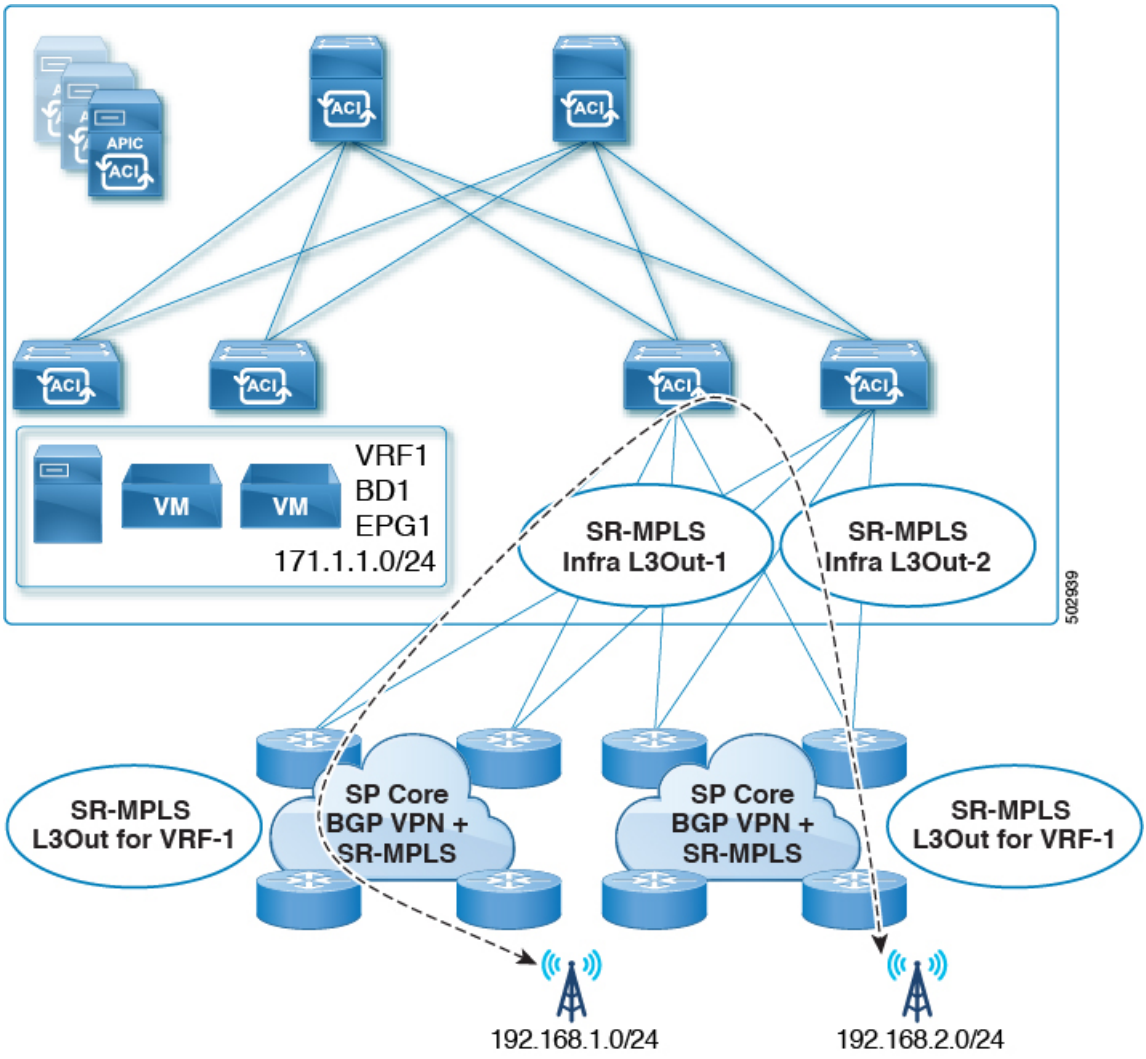
While the Infra MPLS configuration and requirements are described in the Day-0 operations chapter, the following restrictions apply for any user Tenants you will deploy to sites that are connected to SR-MPLS networks.

- You must have created and configured the SR-MPLS Infra L3Outs, including the QoS policies, as described in the Day-0 operations chapter.
- 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 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 them.

    Keep in mind, if the EPGs are in different sites, each EPG must be deployed to a single site only. Stretching EPGs between sites is not supported when using SR-MPLS L3Outs.
- 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` prefixe and choose to not aggregate the routes, it will allow the default route only.

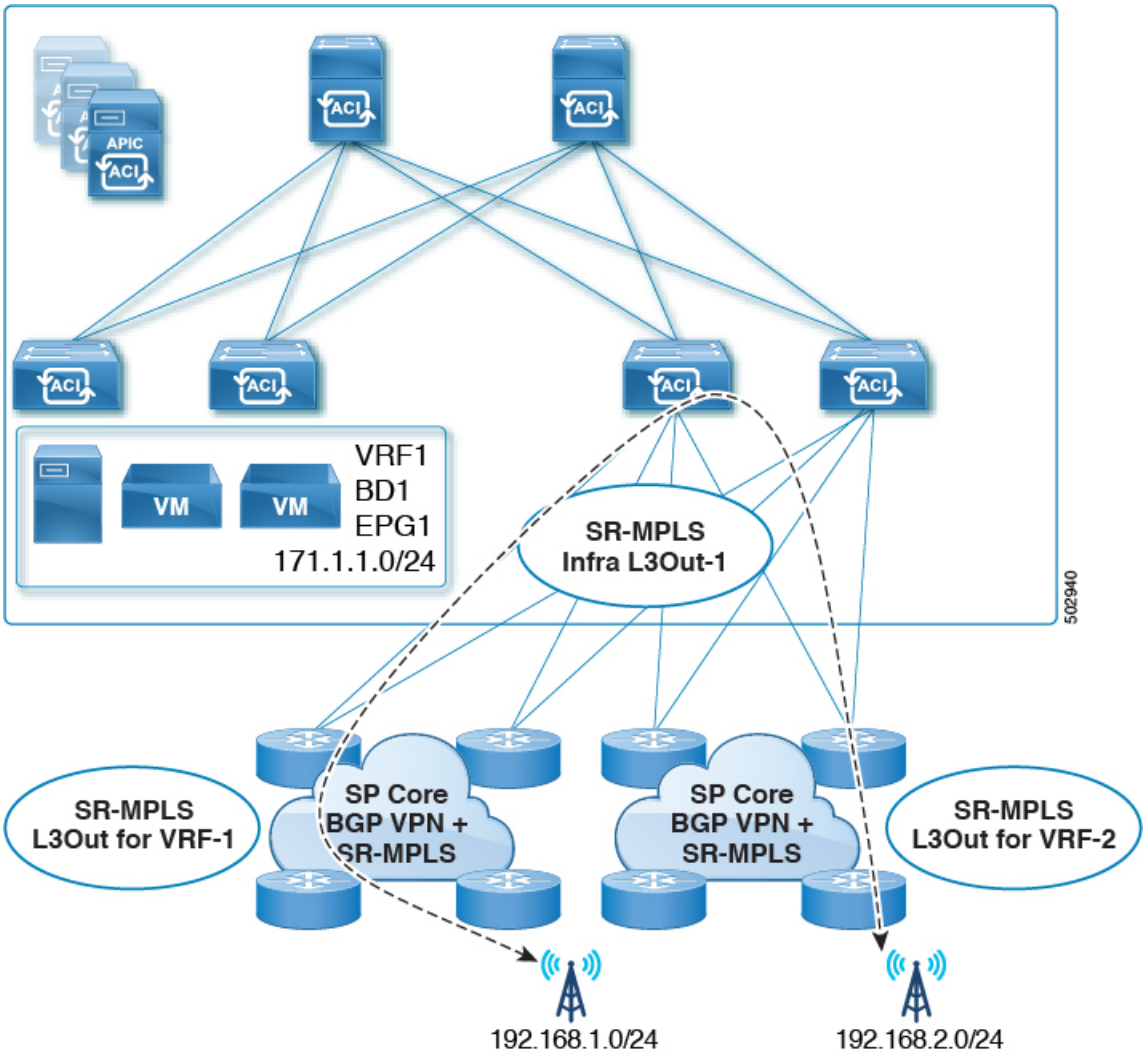
    However, if you choose to aggregate routes `0` through `32` for the `0.0.0.0/0` prefix, it will allow all routes.
  - You can associate any routing policy with any tenant L3Out.
- Transit routing is supported, but with some restrictions:
  - Transit routing between two SR-MPLS networks **using the same VRF** is not supported. The following figure shows an example of this unsupported configuration.

Figure 31: Unsupported Transit Routing Configuration Using Single VRF



- Transit routing between two SR-MPLS networks **using different VRFs** is supported. The following figure shows an example of this supported configuration.

Figure 32: Supported Transit Routing Configuration Using Different VRFs



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 received from the DC-PE routers will be injected into the BGP VPNv4 ACI control plane.

If you have no sites connected to MPLS networks, you can skip this section.

Procedure

Step 1 Log in to the Multi-Site Orchestrator GUI.

Step 2 In the **Main menu**, select **Application Management** > **Policies**.

Step 3 In the main pane, select **Add Policy** > **Create Route Map Policy**.

Step 4 In the **Add Route Map Policy** screen, select a Tenant and provide the name for the policy.

Step 5 Click **Add Entry** under **Route-Map Entry Order** to add a route map entry.

a) Provide the **Context Order** and **Context Action**.

Each context is a rule that defines an action based on one or more matching criteria.

Context order is used to determine the order in which contexts are evaluated. The value must be in the `0-9` range.

Action defines the action to perform (`permit` or `deny`) if a match is found.

b) If you want to match an action based on an IP addres 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** checkbox and provide the range. For example, you can specify `0.0.0.0/0` prefix and choose to aggregate routes `0` through `32`.

c) If you want to match an action based on community lists, click **Add Community**.

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

Then choose the **Scope**.

d) Click **+Add Action** 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 Weight`
- `Set Next Hop`
- `Set Preference`
- `Set Metric`
- `Set Metric Type`

After you have configured the action, click the checkmark icon to save the action.

e) (Optional) You can repeat the previous substeps to specify multiple match criteria and actions within the same Context entry.

f) Click **Save** to save the Context entry.

Step 6 (Optional) Repeat the previous step if you want to add multiple entries to the same route policy.

Step 7 Click **Save** to save the route map policy.

Enabling Template for SR-MPLS

There is a number of template configuration settings that are unique when deploying them to sites connected via MPLS. Enabling SR-MPLS for a Tenant restricts and filters certain configurations that are not available for MPLS sites while bringing in additional configurations only available for such sites.

Before you can update MPLS-specific settings, you must enable the **SR-MPLS** knob in the template's Tenant properties.

Procedure

Step 1 Log in to the Multi-Site Orchestrator GUI.

Step 2 In the main navigation menu, select **Application Management** > **Schemas**.

Step 3 Create a new or select an existing Schema where you will configure SR-MPLS Tenant.

Step 4 Select the Tenant.

If you created a new Schema, choose a Tenant as you typically would. Otherwise click an existing Template in the left sidebar.

Step 5 In the right sidebar **Template** properties, enable **SR-MPLS** knob.

Creating VRF and SR-MPLS L3Out

This section describes how to create the VRF, tenant SR-MPLS L3Out, and External EPG you will use to configure communication between application EPGs separated by an MPLS network.

Before you begin

You must have:

- Created a template and enabled SR-MPLS for its tenant, as described in Enabling Template for SR-MPLS, on page 246.

Procedure

Step 1 Select the template.

Step 2 Create a VRF.

a) In the main pane, scroll down to the **VRF** area and click the + sign to add a VRF.
b) In the right properties sidebar, provide the name for the VRF.

Step 3 Create an SR-MPLS L3Out.

a) In the main pane, scroll down to the **SR-MPLS L3Out** area and click the + sign to add an L3Out.
b) In the right properties sidebar, provide the name for the L3Out.

c) From the **Virtual Routing & Forwarding** dropdown, select the same VRF you selected for the external EPG in the previous step.

Step 4 Create an external EPG.

a) In the main pane, scroll down to the **External EPG** area and click the + sign to add an external EPG.
b) In the right properties sidebar, provide the name for the external EPG.
c) From the **Virtual Routing & Forwarding** dropdown, select the VRF you created in the previous step.

Configuring Site-Local VRF Settings

You must provide BGP route information for the VRF used by the SR-MPLS L3Out.

Before you begin

You must have:

- Created a template and enabled SR-MPLS for its tenant, as described in Enabling Template for SR-MPLS, on page 246.
- Created a VRF and SR-MPLS L3Out, as described in Creating VRF and SR-MPLS L3Out, on page 246.
- Added the template to an MPLS site.

Procedure

Step 1 Select the schema that contains your template.
Step 2 In the left sidebar of the schema view under **Sites**, select the template to edit its site-local properties.
Step 3 In the main pane, scroll down to **VRF** area and select the VRF.
Step 4 In the right properties sidebar, click **+Add BGP Route Target Address**.
Step 5 Configure the BGP settings.

a) From the **Address Family** dropdown, select whether it is IPv4 or IPv6 address.
b) In the **Route Target** field, provide the route string.

For example, `route-target:ipv4-nn2:1.1.1.1:1901`.

c) From the **Type** dropdown, select whether to import or export the route.
d) Click **Save** to save the route information.

Step 6 (Optional) Repeat the previous step to add any additional BGP route targets.

Configuring Site-Local SR-MPLS L3Out Settings

Similar to how you configure site-local L3Out properties for typical external EPGs, you need to provide SR-MPLS L3Out details for external EPGs deployed to sites connected via MPLS.

Before you begin

You must have:

- Created a template and enabled SR-MPLS for its tenant, as described in Enabling Template for SR-MPLS, on page 246.
- Created a VRF and SR-MPLS L3Out, as described in Creating VRF and SR-MPLS L3Out, on page 246.
- Configured the VRF's site-local properties, as described in Configuring Site-Local VRF Settings, on page 247.
- Added the template to an MPLS site.

Procedure

Step 1 Select the schema that contains your template.

Step 2 In the left sidebar of the schema view under **Sites**, select the template to edit its site-local properties.

Step 3 In the main pane, scroll down to **SR-MPLS L3Out** area and select the MPLS L3Out.

Step 4 In the right properties sidebar, click **+Add SR-MPLS Location**.

Step 5 Configure the SR-MPLS Location settings.

a) From the **SR-MPLS Location** dropdown, select the Infra SR-MPLS L3Out you created when configuring Infra for that site.

b) Under **External EPGs** section, select an external EPG from the dropdown and click the checkmark icon to add it.

You can add multiple external EPGs.

c) Under **Route Map Policy** section, select a route map policy you created in previous section from the dropdown, specify whether you want to import or export the routes, then click the checkmark icon to add it.

You must configure a single export route map policy. Optionally, you can configure an additional import route map policy.

d) Click **Save** to add the location to the MPLS L3Out.

Step 6 (Optional) Repeat the previous step to add any additional SR-MPLS Locations for your SR-MPLS L3Out.

Communication Between EPGs Separated by MPLS Network

Typically, if you wanted to establish communication between two EPGs, you would simply assign the same contract to both EPGs with one EPG being the provider and the other one a consumer.

However, if the two EPGs are separated by an MPLS network, the traffic has to go through each EPG's MPLS L3Out and you establish the contracts between each EPG and its MPLS L3Out instead. This behavior is the same whether the EPGs are deployed to different sites or within the same fabric but separated by an SR-MPLS network, such as in Multi-Pod or Remote Leaf cases.

Before you begin

You must have:

- Added one or more sites connected to MPLS network(s) to the Orchestrator.
- Configured Infra MPLS settings, as described in "Day-0 Operations" chapter.
- Created a schema, added a Tenant, and enabled the Tenant for SR-MPLS, as described in Enabling Template for SR-MPLS, on page 246.

Procedure

Step 1 Log in to the Multi-Site Orchestrator GUI.

Step 2 Create two application EPGs as you typically would.

For example, `epg1` and `epg2`.

Step 3 Create two separate external EPGs

These EPGs can be part of the same template or different templates depending on the specific deployment scenario.

For example, `mpls-extepg-1` and `mpls-extepg-2`

Step 4 Configure two separate Tenant SR-MPLS L3Outs.

For example, `mpls-l3out-1` and `mpls-l3out-2`

For each Tenant SR-MPLS, configure the VRF, route map policies, and external EPGs as described in Configuring Site-Local VRF Settings, on page 247 and Configuring Site-Local SR-MPLS L3Out Settings, on page 247.

Step 5 Create a contract you will use to allow traffic between the two application EPGs you created in Step 2.

You will need to create and define a filter for the contract just as you typically would.

Step 6 Assign the contracts to the appropriate EPGs.

In order to allow traffic between the two application EPGs you created, you will actually need to assign the contract twice: once between `epg1` and its `mpls-l3out-1` and then again between `epg2` and its `mpls-l3out-2`.

As an example, if you want `epg1` to provide a service to `epg2`, you would:

a) Assign the contract to `epg1` with type `provider`.
b) Assign the contract to `mpls-l3out-1` with type `consumer`.
c) Assign the contract to `epg2` with type `consumer`.
d) Assign the contract to `mpls-l3out-2` with type `provider`.

Deploying Configuration

You can deploy the configuration Template to an MPLS site as you typically would, with one exception: because you cannot stretch objects and policies between MPLS site and another site, you can only select a single site when deploying the template.

Procedure

Step 1 Add the site to which you want to deploy the template.

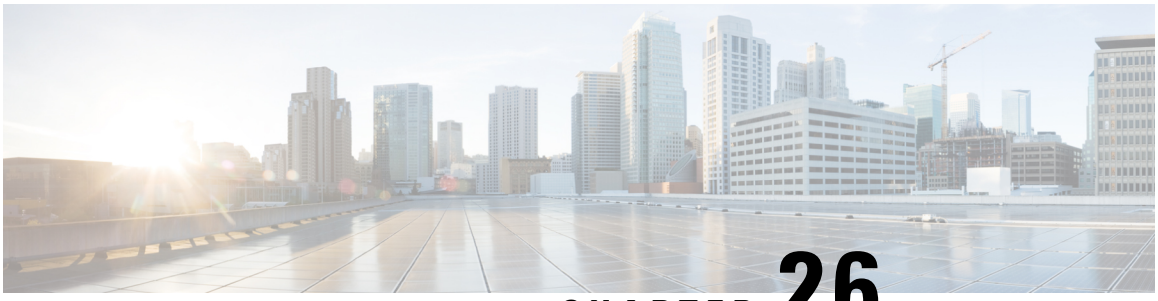
a) In the left sidebar of the **Schema** view under **Sites**, click the **+** icon.
b) In the **Add Sites** window, select the site where you want to deploy the Template.

You can only select a single site if your template is MPLS-enabled.

c) From the **Assign to Template** dropdown, select one or more Template you have created in this Schema.
d) Click **Save** to add the site.

Step 2 Deploy the configuration

a) In the main pain 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**.



CHAPTER 26

vzAny Contracts

- vzAny and Multi-Site, on page 251
- vzAny and Multi-Site Guidelines and Limitations, on page 252
- Create Contract and Filters, on page 254
- Configure vzAny to Consume/Provide a Contract, on page 255
- Create EPGs to Be Part of the vzAny VRF, on page 255
- Free Intra-VRF Communication, on page 256
- Many-to-One Communication, on page 261

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 256.
- 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 261.

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 Multi-Site 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 Multi-Site 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.

a) Scrolls down to the **Filter** area and click the + sign to add a new filter.
b) Provide the name for the Contract.
c) Click **+Entry** to add a filter entry.
d) 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.

e) Click **SAVE** to add the entry.
f) (Optional) If required, create additional filter entries.

Step 5 Create the contract.

a) Scrolls down to the **Contract** area and click the + sign to add a new contract.
b) Provide the name for the Contract.

For example, `contract-vzany`.

c) 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`.

d) Choose whether the contract will apply in both directions
e) Click **+Filter** to add one or more contract filters.
f) In the **Add Filter Chain** window, choose the filter you created in the previous step.
g) Click **SAVE** to add the filter.
h) (Optional) If required, repeat the procedure to provide additional filters.
i) (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 254.

Procedure

Step 1 Log in to the Multi-Site 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 254.

c) Select the Contract type.

You can choose either `consumer` or `provider` 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 254.
- Created the vzAny VRF and assigned the Contract to it as described in Configure vzAny to Consume/Provide a Contract, on page 255.

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.

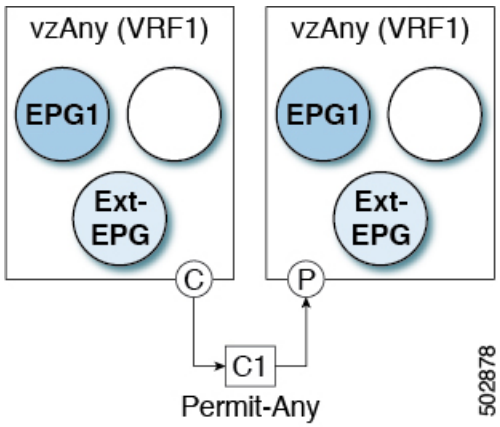
Step 2 If you want to create an External EPG to be part of the vzAny VRF

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 33:



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 254.

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 255.

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 255.

Step 7 Assign the common Template to every site.

Step 8 Assign each template to the appropriate sites.

Step 9 Deploy the templates.

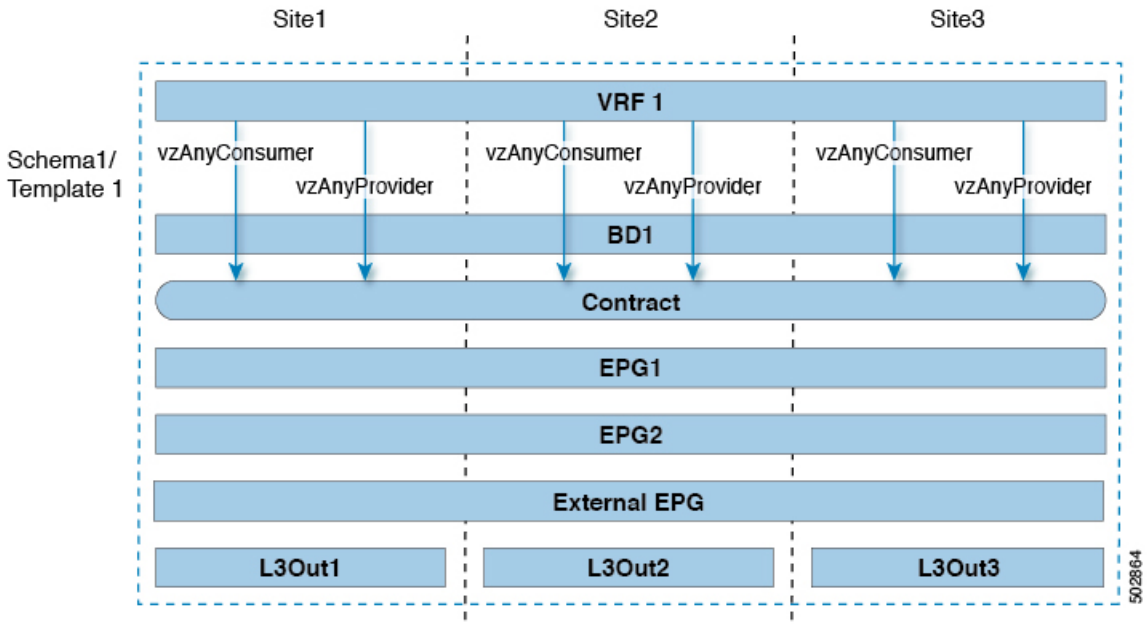
Stretched EPGs

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 34:



Site-Local EPGs

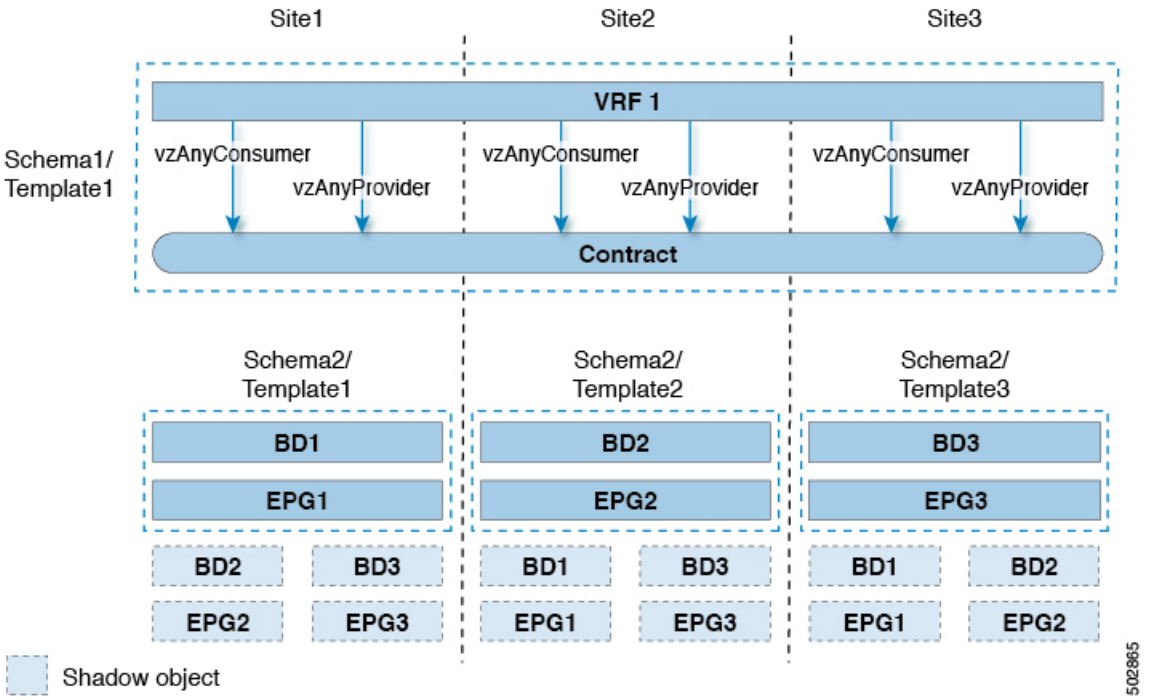
The following example shows intra-VRF communication between EPGs or External EPGs where none of the EPGs are stretched but can still freely communicate with each other since vzAny consumes and provides the "permit-any" contract.

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 35:



Combination of Site-Local and Stretched EPGs

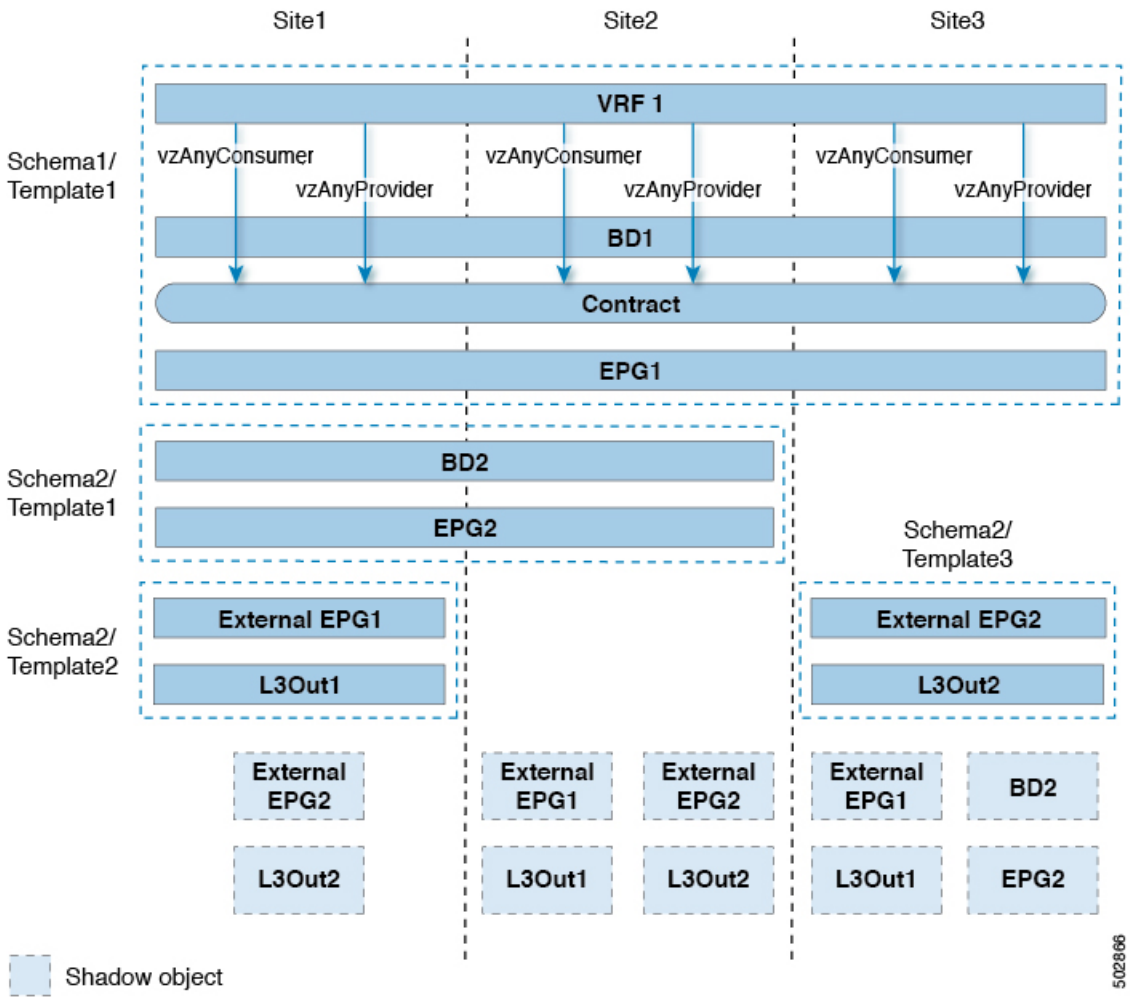
The following example shows intra-VRF communication between EPGs or External EPGs where some EPGs are stretched while others are deployed to a single site only. All EPGs can still freely communicate with each other since vzAny consumes and provides the "permit-any" contract.

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 36:



Intra-VRF Intersite L3Out

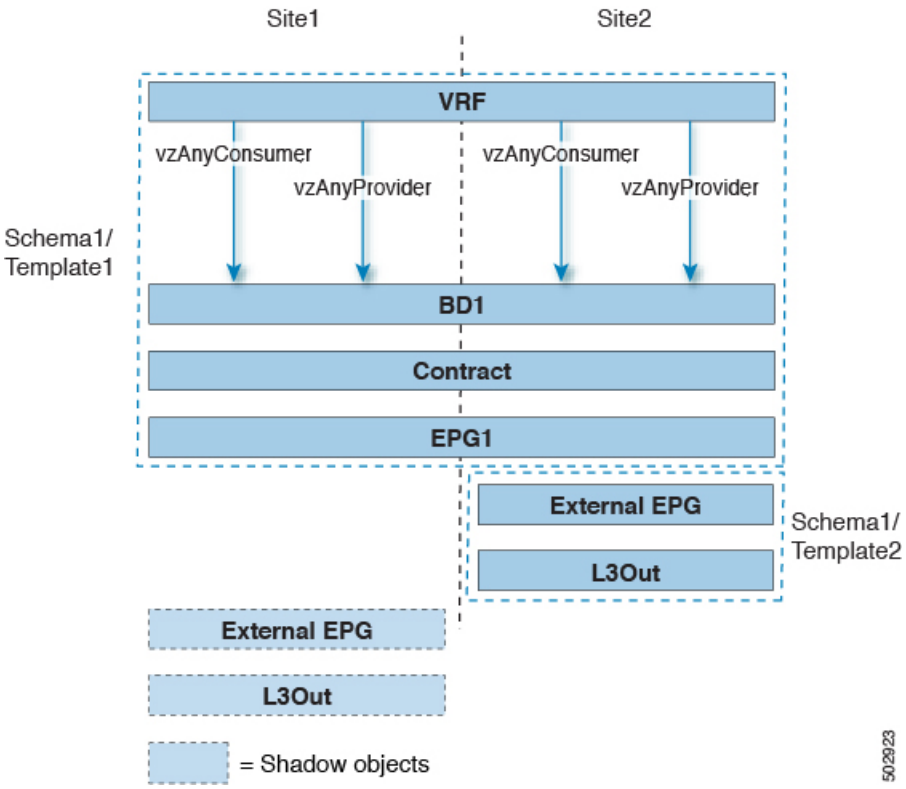
This use case allows you to configure an intersite L3Out for multiple EPGs within a vzAny VRF. When the L3Out's external EPG is in the same VRF, you do not need to explicitly add the provider contract to the external EPG.

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 159 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 37:



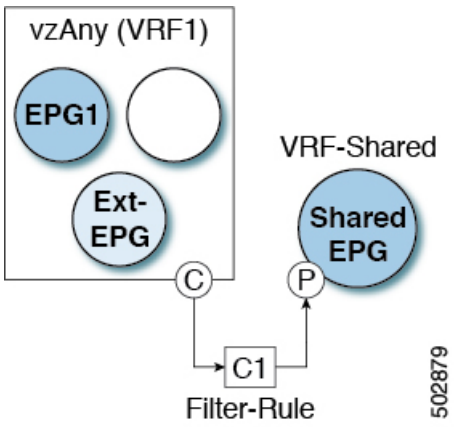
Based on the configuration shown in the above figure, endpoints that are part of the stretched EPG1 and connected to Site1 will be able to communicate with the external network domain via the L3Out connection deployed in Site2. The same would apply for endpoints that are part of site-local EPGs deployed in Site1.

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 38:



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. Specific sections 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 254.

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 255.

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 255.

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

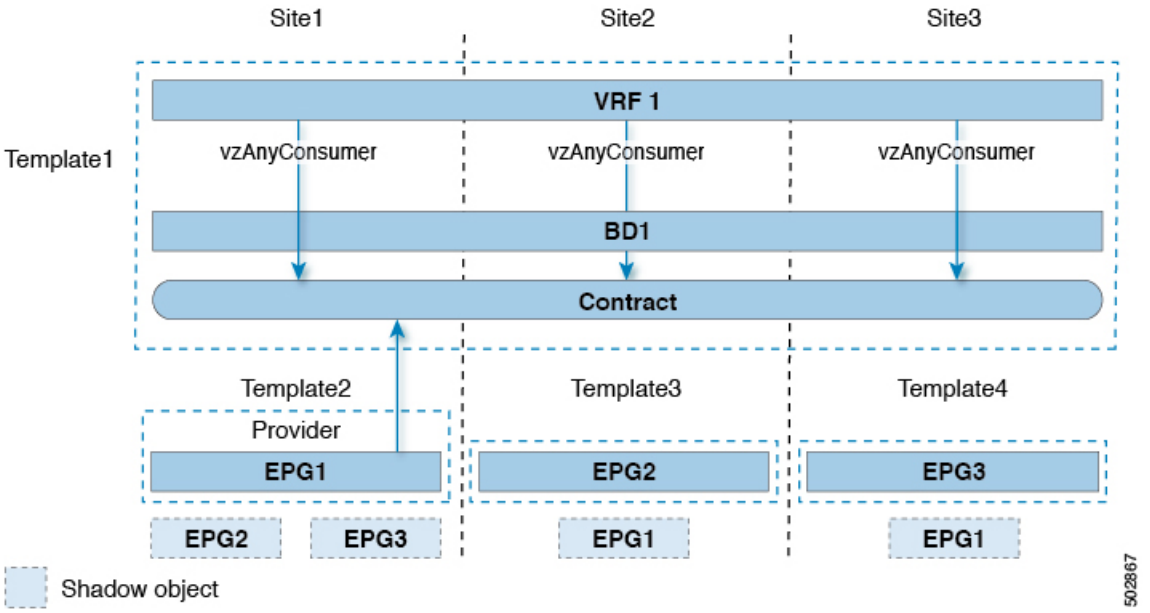
The following example shows intra-VRF communication between a single provider EPG (for example, shared service) and all other EPGs within the same VRF consuming the service.

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 39:



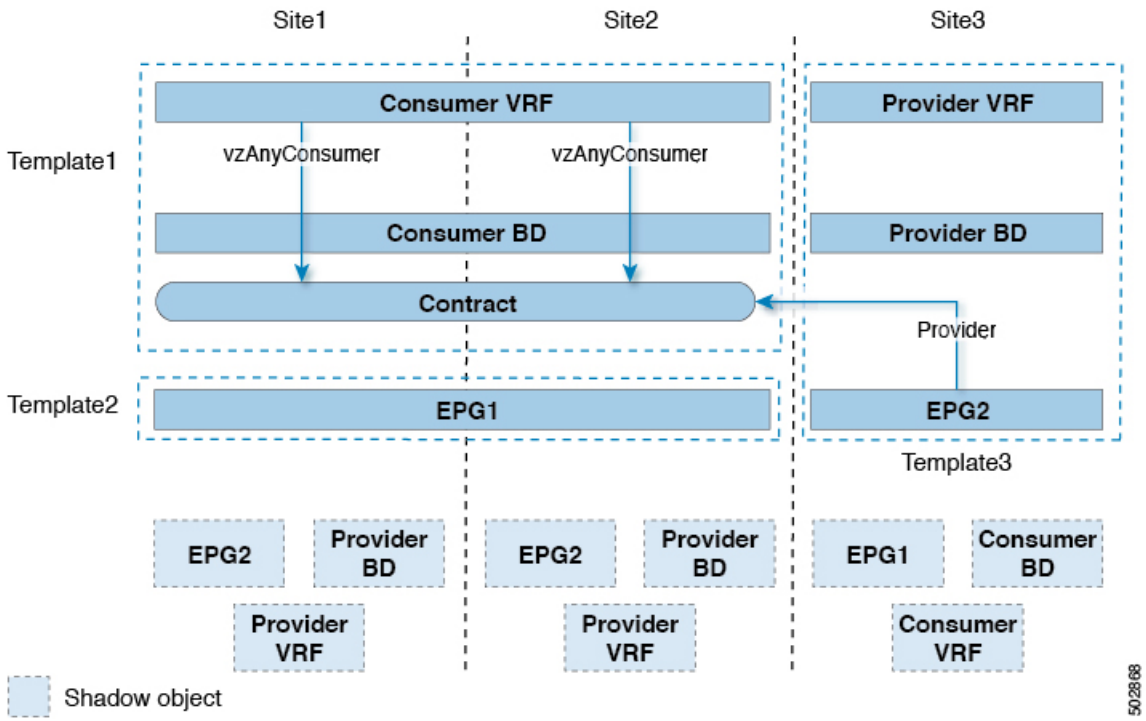
Provider EPG In Its Own VRF

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 40:


Site1
Site2
Site3
Consumer VRF
Provider VRF
Template1
vzAnyConsumer
vzAnyConsumer
Consumer BD
Provider BD
Contract
Provider
Template2
EPG1
EPG2
Template3
EPG2
Provider BD
EPG2
Provider BD
EPG1
Consumer BD
Provider VRF
Provider VRF
Consumer VRF
Shadow object
502868