Schema Management

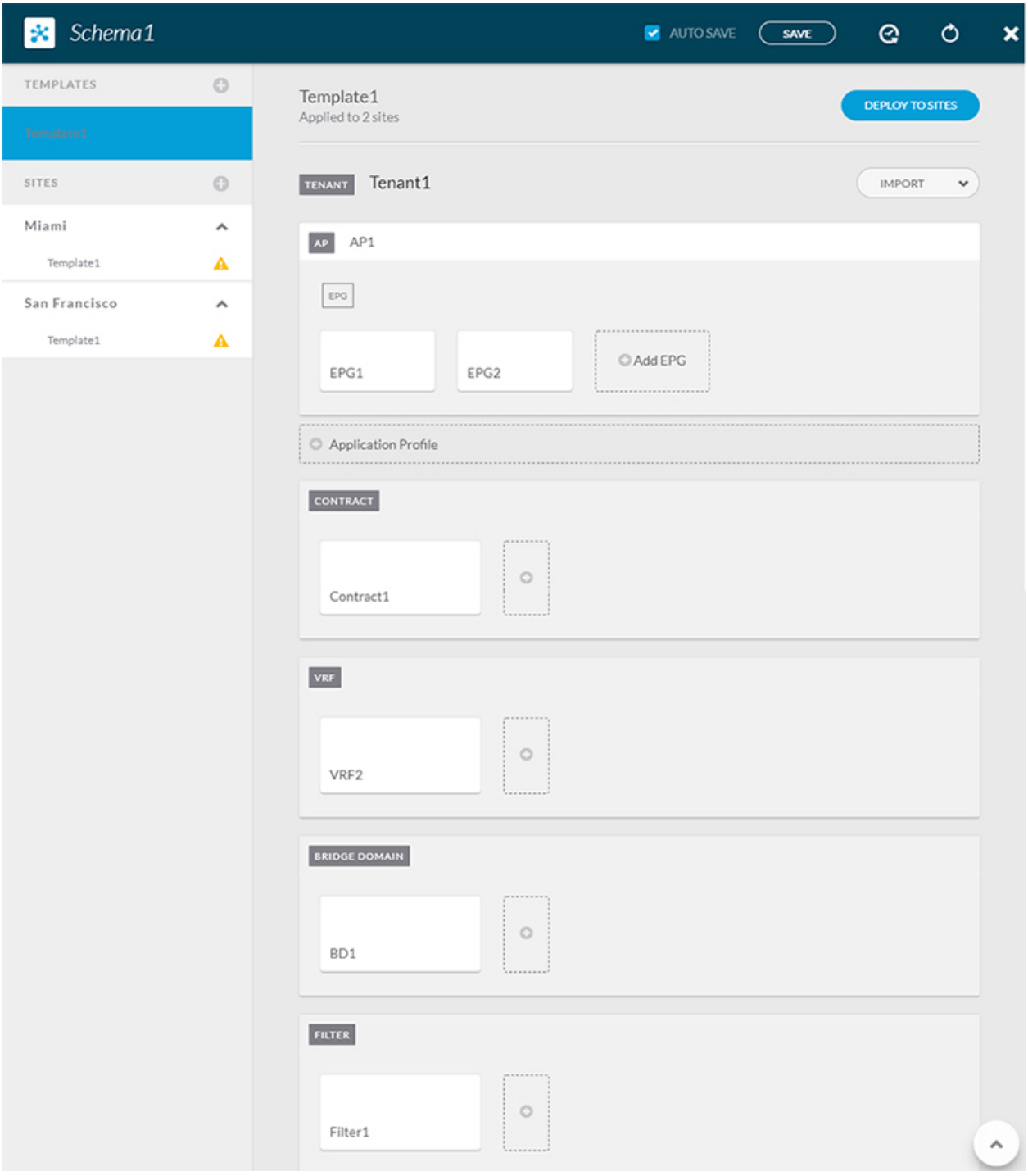
Schema Design Considerations

A schema is a collection of templates, which are used for defining policies. This release of Cisco ACI Multi-Site supports up to 60 schemas with 5 templates and 500 objects per schema. It is important to consider schema design approach based on the deployment use-case.

Single Schema Deployment

The simplest schema design approach is a single schema deployment. You can create a single schema and adds all VRFs, Bridge Domains, EPGs, Contracts and other elements to it. You can then create a single application profile or multiple application profiles within the schema and deploy it to one or more sites.

Figure 1: Single Schema



This simple approach to Multi-Site schema creation is illustrated in the figure above and allows for all objects to be readily visible within the same schema. However, the 5 templates and 500 objects per schema limit makes this approach unsuitable for large scale deployments, which could exceed those limits.

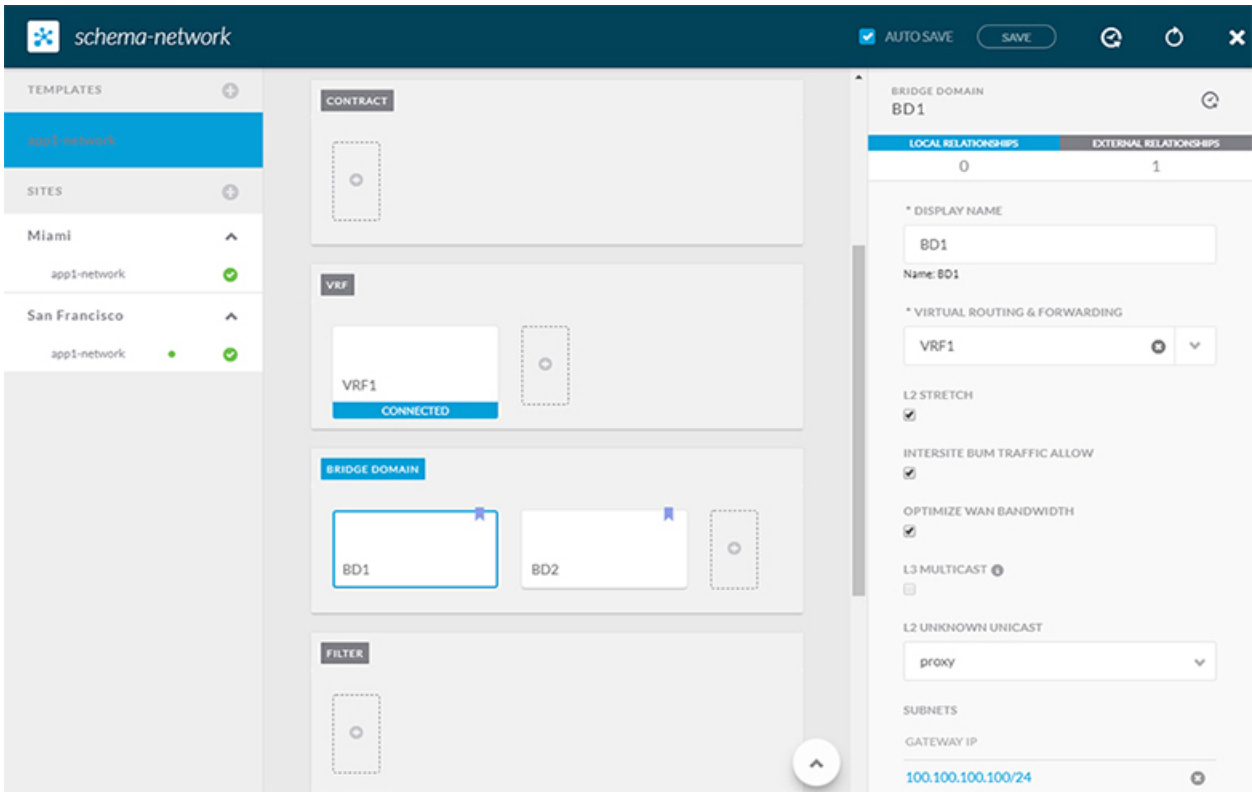
Multiple Schemas with Network Separation

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.

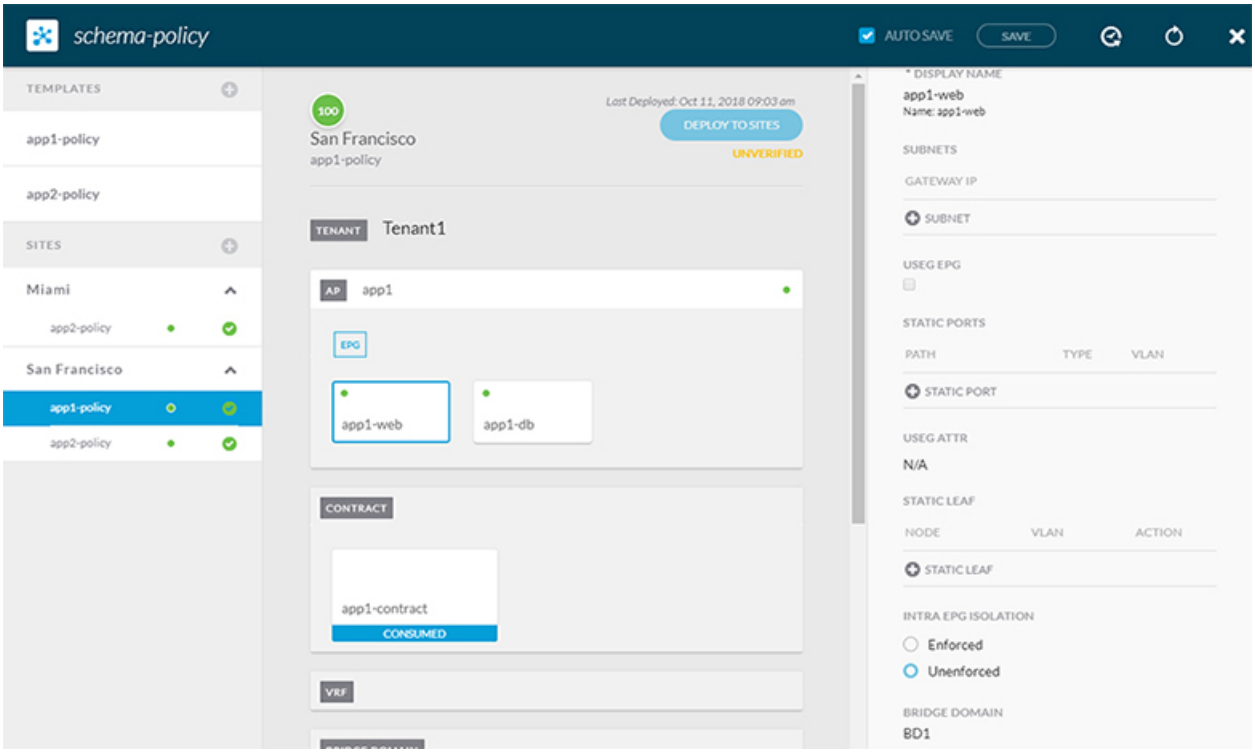
The following figure shows a single networking template configuration with VRF, BD, and subnets configured and deployed to two sites:

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



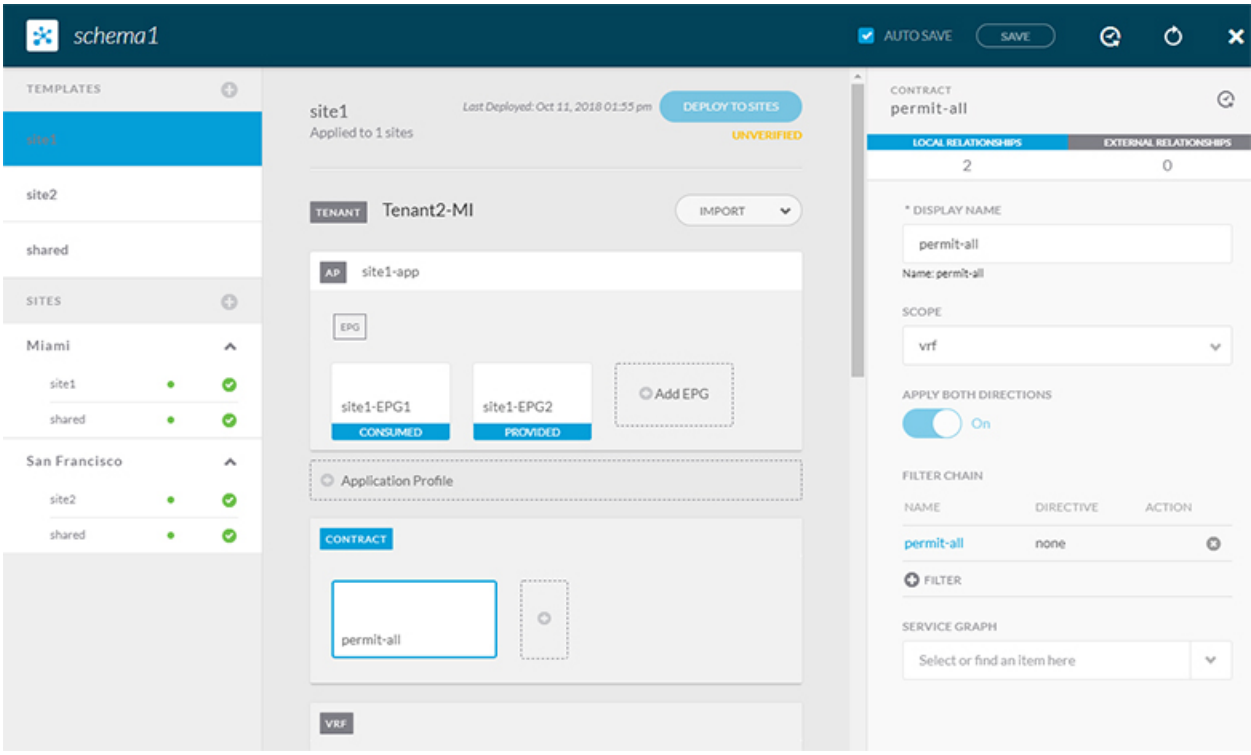
After creating and deploying the policy schemas and templates, the networking objects in the networking schema will display the number of external references by the policy schema elements. The object with external references will also be denoted by the ribbon icon as shown in the Figure 2: Network Schema 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 would exceed the 5 templates per schema limit, 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.

Schema Design for Cisco Cloud APIC Use-Cases

Starting with Release 2.1(1), Cisco ACI Multi-Site supports Cisco Cloud APIC installed in the Amazon Web Services (AWS) cloud as one of the sites. While the sections below outline creation and management of generic schemas, specific use-case scenarios supported with Cloud APIC sites are detailed in the configuration examples available from the following Cloud APIC documentation landing page: https://www.cisco.com/c/en/us/support/cloud-systems-management/cloud-application-policy-infrastructure-controller/tsd-products-support-series-home.html.

Creating a Schema Template

Before you begin

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

Step 1 On the **Schema** page, click the **Add Schema** button.

Step 2 On the **Untitled Schema** page, enter a name for the schema you intend to create.

Step 3 Access the **Select A Tenant** dialog box and select a tenant from the drop-down menu.

Note 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.

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.

Step 1 On the **Schema** page, click the **Import** button.

Step 2 Select the site from which you want to import objects.

Step 3 In the **Import** window that opens, select one or more objects you want to import.

Note The names of the objects imported into the Multi-Site Orchestrator 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.

Configuring an Application Profile

Describes how to configure an Application Profile with EPGs.

Step 1 In the AP field, click **+ Application Profile**.

Step 2 Enter the AP name.

Step 3 Click **+ Add EPG** and enter the EPG name in the Display Name field.

Step 4 Optional. Click **+ Subnet** to add a subnet to your EPG, if appropriate.

You may choose to configure a subnet for the EPG, 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 5 Optional. Select **USEG EPG** if appropriate.

a) Enter the **Name** and **Type** for the **USEG ATTR**.
b) Click **+ USEG ATTRIBUTE** to add USeg attributes if appropriate.
c) On the **Add uSeg Attributes** dialog, enter a **Display Name**, **Description**, and **Attribute Type**.
d) Click **SAVE**.

Step 6 Select either **Enforced** or **Unenforced** for the Intra EPG Isolation field.

Step 7 Choose a bridge domain in the **Bridge Domain** field.

Step 8 Click **+ Contract** to add a contract if appropriate.

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

Configuring a Contract

Provides the procedure for configuring contracts to control EPG communications.

Step 1 Click **+** in the box in the **Contract** pane.

Step 2 Enter a name for the contract in the **Contract** dialog in the **Display Name** field under **Display Name**.

Step 3 Choose a value for **Scope** using the drop-down menu.

Step 4 Click **Apply Both Directions** toggle button to apply the filter specified in the contract to either one direction or both directions.

The default setting is **ON**.

Step 5 Click **+ Filter**.

a) On the **Add Filter Chain** dialog, click the **Name** field to choose or find a filter.
b) Optional. Select the available directives in the **Directives** field.
c) Click **SAVE**.

What to do next

After you have configured the contract to your specifications, click **Deploy to Sites**.

Configuring a Bridge Domain

Step 1 Click + in the **Bridge Domain** pane to add a new bridge domain.

Step 2 In the right-hand properties sidebar that opens, provide the following bridge domain details:

- The BD name in the **Display Name** field.
- The VRF in the **Virtual Routing and Forwarding** field.
- If appropriate, check the **L2 STRETCH** checkbox.
- If you enabled **L2 STRETCH**, you can choose to also enable **INTERSITE BUM TRAFFIC ALLOW** checkbox.
- If you did not enable **L2 STRETCH**, you can choose either **proxy** or **flood** for the **L2 UNKNOWN UNICAST** field

Step 3 (Optional) You can choose to add one or more subnets to the bridge domain.

a) Click **+ Subnet**.

An **Add Subnet** window opens.

b) Enter the subnet's **Gateway IP** address and a description for the subnet you want to add.
c) In the **Scope** field, select either **Private to VRF** or **Advertised Externally**.
d) If appropriate, check the **Shared Between VRFs** checkbox.
e) If appropriate, check the **No Default SVI Gateway** checkbox.
f) If appropriate, check the **Querier** checkbox.
g) Click **SAVE**.

Configuring a VRF for the Tenant

Step 1 Click the + in the VRF field.

Step 2 Enter a display name for the VRF in the **Display Name** field.

Configuring a Filter for Contracts

Provides a method to create a filter. A filter is similar to an Access Control List (ACL), used to filter traffic through contracts associated to EPGs.

Step 1 Click the **+** on the Filter object under in the Filter pane.

Step 2 Enter a display name in the **Display Name** field.

Step 3 Click **+ Entry** to provide information for your schema filter on the **Add Entry** display:

a) Enter a name for the schema filter entry in the **Name** field on the **Add Entry** dialog.

b) Optional. Enter a description for the filter in the **Description** field.

c) Enter the details as appropriate to filter EPG communication.

For example, to add an entry allowing HTTPS traffic through a filter, choose **TYPE**: **IP**, **IP PROTOCOL**: **TCP**, and **DESTINATION PORT RANGE FROM** and **DESTINATION PORT RANGE TO**: **https**.

d) Click **SAVE**.

Configuring an External EPG

In this schema field, you configure an **EXTERNAL EPG** for each site, to enable the sites to connect. Alternatively, you may want to configure a single External EPG in a template that is in turn associated to multiple sites.

To configure an external EPG for each site, in a site-specific template perform the following steps:

Before you begin

- Create an L3Out in Cisco APIC on all sites where the tenant and VRF are stretched.
- The VRF for each L3Out must be the same for all sites. Changing the VRF in APIC, after the external EPGs are deployed, resets the L3Out and requires reconfiguring and redeploying the external EPG for the site.

Step 1 Click + to create an **EXTERNAL EPG**.

Step 2 Enter the external EPG name.

Step 3 Add the contracts required for the external EPGs to communicate.

Note 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 4 Click the site-specific template.

Step 5 Click the external EPG.

Step 6 In the external EPG details pane, **L3OUT** field, choose the L3Out on the site to be used for the external EPG.

What to do next

Optional. You can also add a subnet under the external EPG.

Repeat these steps to create an external EPG for each site.

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.

For more information about the functionality of these Multi-Site Orchestrator GUI pages, refer to Overview of the Cisco ACI Multi-Site Orchestrator GUI.

Contract Preferred Groups

Before Release 2.0(2), the Cisco ACI Multi-Site architecture supported communication between EPGs only if a contract was configured between them. If no contract existed, any inter-EPG communication was explicitly disabled by default. Starting with Release 2.0(2), you can include several EPGs in a contract **Preferred Group** which allows full communication between the EPGs in a single VRF.

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 configuration of 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.

Limitations

Inter-site external EPGs are not supported as part of inter-site preferred groups from the Multi-Site Orchestrator. If you want to include a stretched external EPG in a preferred group, you must do so in each site's APIC individually after the external EPG is deployed from the Orchestrator.

Configuring EPGs for Preferred Group

Before you begin

You must have one or more EPGs added to a schema template.

Step 1 Log in to your Cisco ACI 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.

Layer 3 Multicast

Note Layer 3 Multicast across sites is a limited availability feature in Multi-Site 2.0(1). If you plan to enable this feature in your production environment, please consult Cisco for deployment planning and validation.

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 routing (PIM) on that BD. By default, multicast is disabled in all BDs.

When an EPG sends multicast traffic to a remote site where it is not stretched, the MSC creates a shadow EPG on the remote site for each such EPG. This could potentially result in an increased amount of configuration changes, such as subnet routes, being pushed to the remote TORs. To alleviate this, Layer 3 multicast has to also be enabled on the individual EPGs which have multicast sources present, so that the configuration necessary for only 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

The following is a high level overview of the Layer 3 multicast routing across sites:

- When the multicast source is attached to ACI fabric as End Point (EP) at one site, that site's spine switch will send the multicast traffic to other sites where the source's VRF is instantiated using the Head End Replication (HREP). The multicast traffic will be sent over to other sites where VRF is stretched and multicast traffic will be pruned/forwarded at egress leaf switches based on the group membership.
- The multicast routing solution requires external multicast router to be the Rendezvous Point (RP). 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 source outside the fabric via the site's local L3Out. As such, traffic coming in on L3Out on one site should not be sent to other sites. This is achieved on the spine by pruning multicast traffic from replicating into HREP tunnels.

- All multicast traffic ingressing a TOR's L3out bridge domain from external router is remarked with a special DCSP value in the outer VXLAN header. On the Spine, that DSCP value is matched to prune all multicast traffic from replicating HREP copies into the ISN network
- Traffic sent from one site can be sent out of any site's L3Out.
- When multicast is enabled on a BD and an EPG from the Multi-Site Orchestrator, all of the BD's subnets are injected into all leaf switches, including the border leaf (BL). This enable 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 if there is a policy configured on the BL. The /32 host routes are advertised if host-based routing is configured on the BD. The BD's subnets and host routes are advertised if the L3Out policy allows a large subnet range including 0/0 and multicast is enabled on the EPG.

For additional information about multicast routing, see the IP Multicast section of the *Cisco APIC Layer 3 Networking Configuration Guide*.

Enabling Layer 3 Multicast

The following procedure describes how to enable Layer 3 multicast on VRF, BD, and EPG using the Cisco ACI Multi-Site Orchestrator GUI.

Before you begin

Cisco ACI Multi-Site Orchestrator cannot create the required local policies on each site, as such you must configure IGMP related policies, PIM related policies, route-maps, RPs, and L3Outs 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* .

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

Step 2 From the left-hand sidebar, select the **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-hand sidebar, check the **L3 Multicast** checkbox.

Step 5 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-hand sidebar, check the **L3 Multicast** checkbox.

Step 6 Enable Layer 3 multicast on an EPG.

Once you have enabled L3 Multicast on the BD, you can select EPGs which have multicast sources. You can only do that if the EPG is part of multicast-enabled BD and VRF.

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.

Shadow EPGs and BDs

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. These 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" EPGs or BDs.

For example, if the provider site group tenant and VRF are stretched across Site 1 and Site 2, and the consumer site group tenant and VRF are stretched across Site 3 and Site 4, in the APIC GUI at Site 1, Site 2, Site 3, and Site 4, you can see both tenants and their policies. They appear with the same names as the ones that were deployed directly to each site.

You can distinguish these shadow EPGs and BDs in the APIC GUI as described below:

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

Step 1 To identify a shadow EPG in a pair of EPGs with the same name, in the APIC GUI, navigate to **Tenants** > ***tenant-name*** > **Application Profiles** > ***ap-name*** > **Application EPGs** > ***epg-name*** > **Static Ports**.

A shadow EPG has no path to the static port.

Step 2 To identify a shadow BD from a pair of BDs with the same name, in the APIC GUI, navigate to **Tenants** > ***tenant-name*** > **Networking** > **Bridge Domains** > ***bd-name*** > **Subnets** > ***subnet-name*** .

The subnet for a shadow BD has **No Default SVI Gateway** enabled.