Multi-Site Use Cases

Cisco ACI Multi-Site Service Integration

Starting in Release 1.2(1), there are many variants of service graph deployment in single ACI fabric. For Cisco ACI Multi-Site deployment, not all use cases are supported in current release. Supported uses are mentioned below in the document.

To support the use cases mentioned below, following topology is required for service nodes for all the use cases.

- Each site has individual Active/Standby Service Node pair.
- Layer 4 to Layer 7 devices are in unmanaged mode.
- Policy based redirect (PBR) policies are required to redirect traffic to service nodes.

Supported Use Cases

East-West Intra-VRF/Non-Shared Service

This is the use case for east-west communication between endpoints in the same VRF instance across sites. EPG WEB and EPG APP are deployed across sites and the firewall (FW) needs to be inserted between them. This is a common design for traffic within the application.

Requirements:

- Policy-based redirect (PBR) policies are required on the consumer and provider connectors of the service node.

- Provider EPGs should have the subnet defined under it and should be unique (similar to intra-VRF route leaking). This is to ensure that the first packet does not reach to the provider's leaf switch without the policy being applied, as you want the traffic to select the service node in the consumer site always. The policies are always applied on the consumer's leaf switch. For this purpose, provider's subnet would be leaked to consumer's leaf switch.
- Redirect policies are always applied on the consumer leaf switch, which means the FW device cluster chosen will always be local to the site of the consumer for traffic flows.

The figures below shows the example for traffic from EPG WEB to EPG APP and vice versa.

Figure 1: Across Site Incoming Traffic

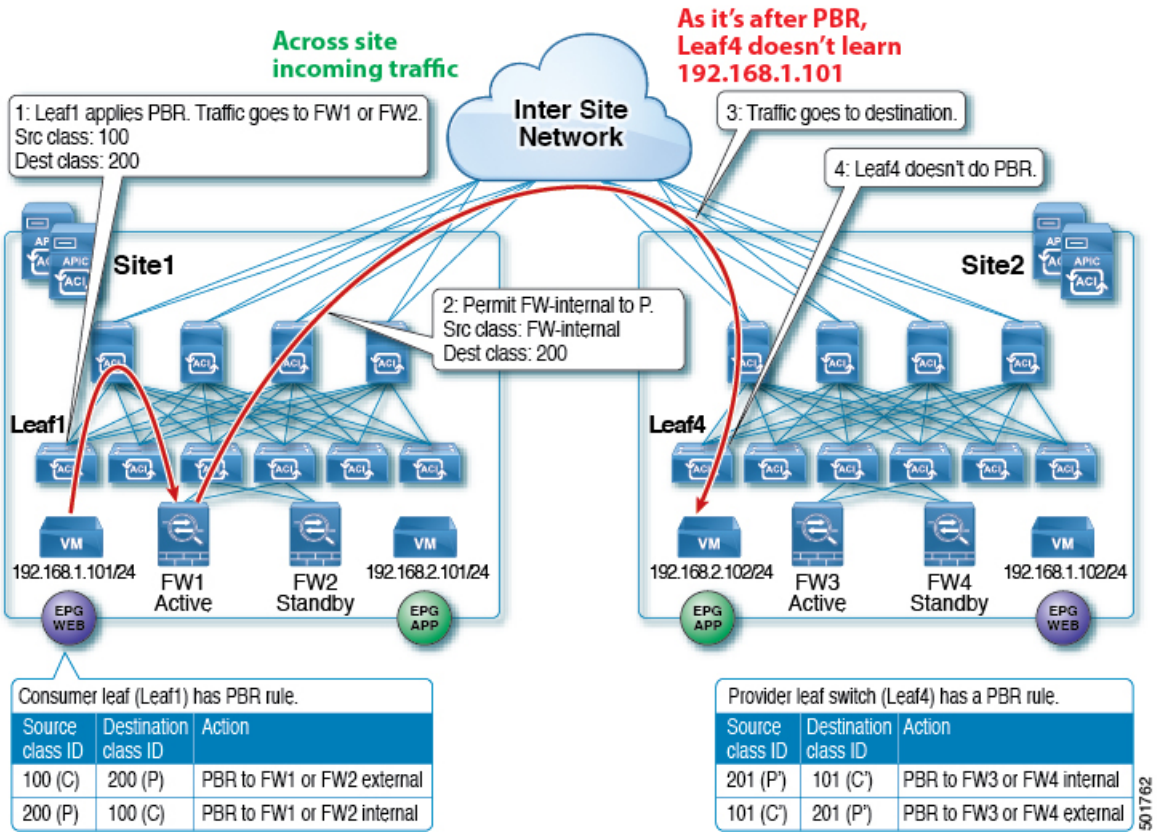
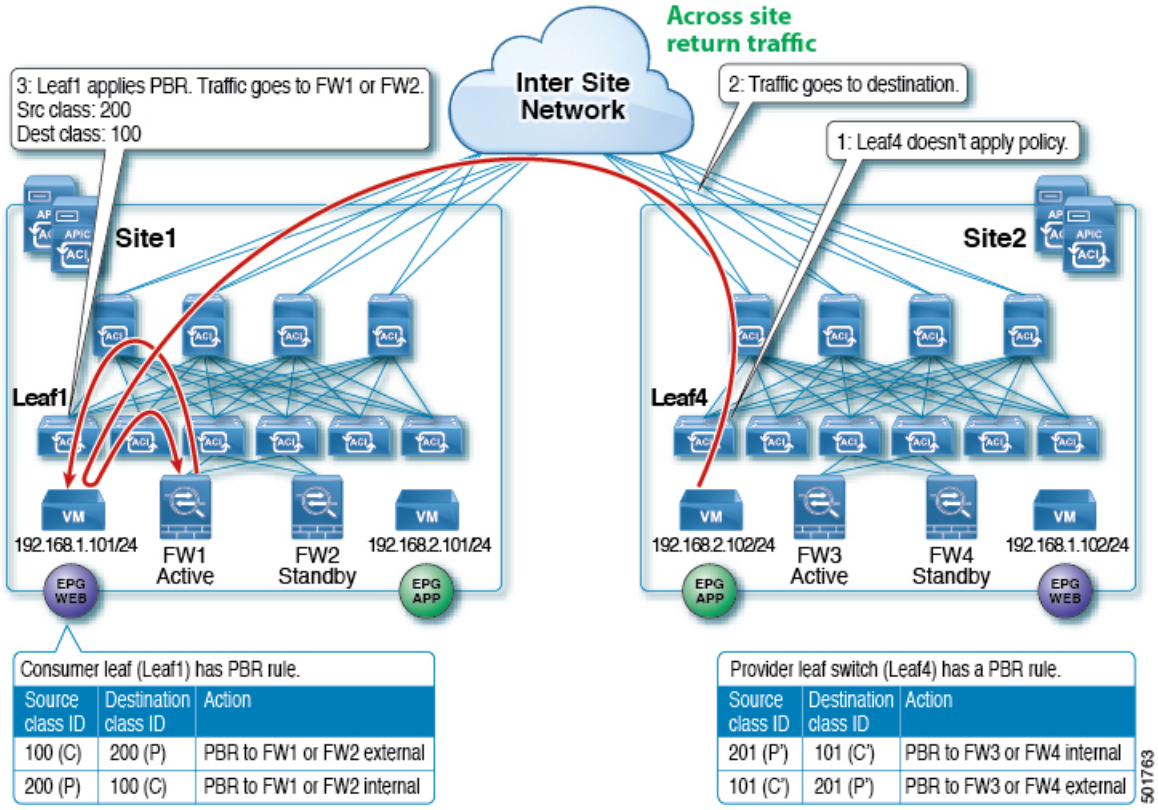


Figure 2: Across Site Return Traffic



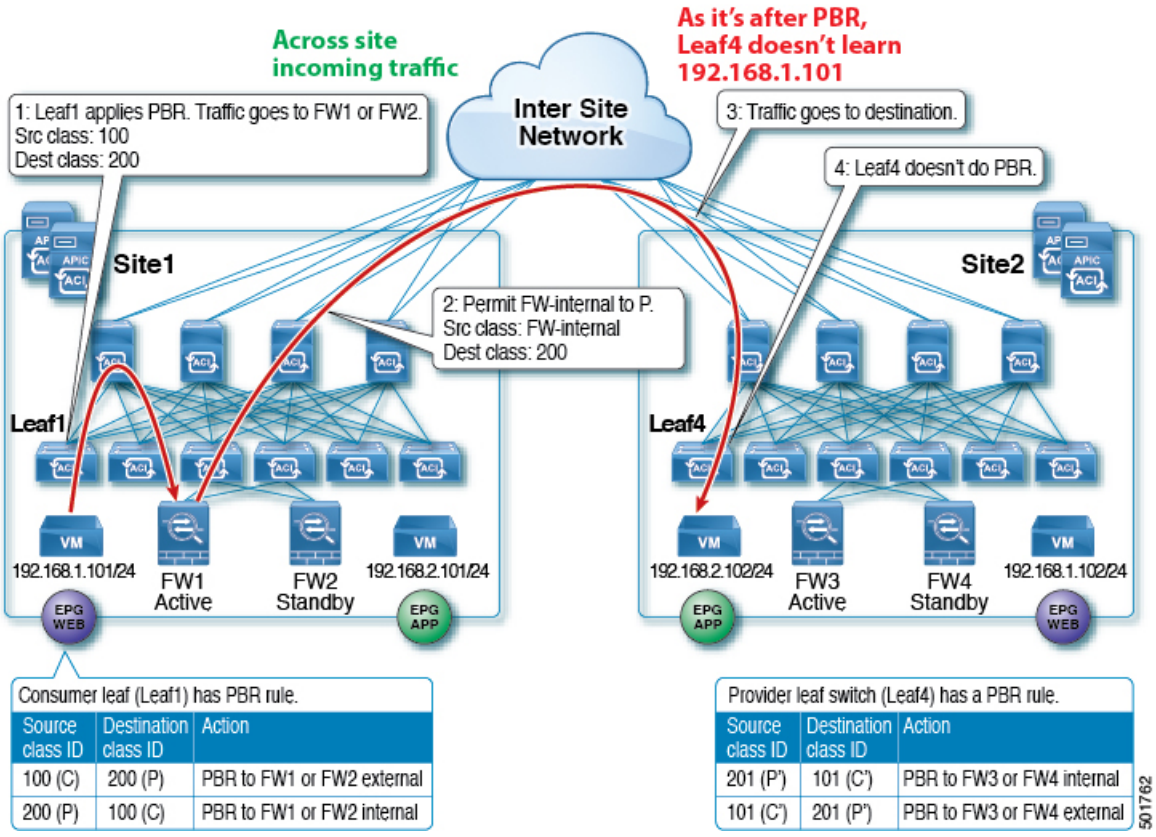
East-West Shared Service

This is the use case for east-west communication between endpoints in the different VRFs across sites. The figures shows the example, EPG WEB in VRF1 and EPG APP in VRF2 are deployed across sites and firewall needs to be inserted between them. This is a common design for the shared service use case. For example, the NFS service is shared to multiple servers in different VRFs.

Requirements:

- PBR policies are required on the consumer and provider connectors of the service node.
- Provider EPGs should have the subnet defined under it and should be unique (similar to intra-VRF route leaking). This is to ensure that first packet does not reach to the provider's leaf switch without the policy being applied, as you want the traffic to select the service node in the consumer site always. The policies are always applied on the consumer's leaf switch. For this purpose, provider's subnet would be leaked to consumer's leaf.
- Redirect policies are always applied on the consumer leaf, which means the firewall (FW) device cluster chosen will always be local to the site of the consumer for traffic flows.

Figure 3: Across Site Incoming Traffic



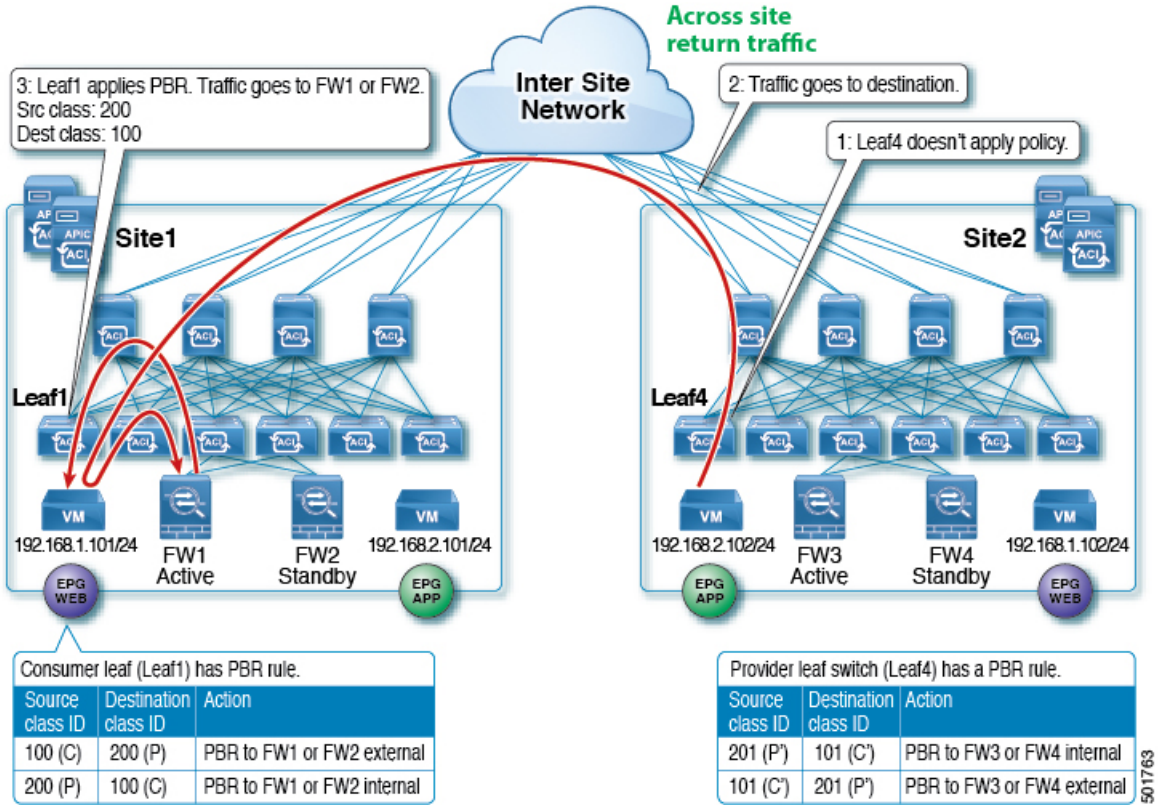
Consumer leaf (Leaf1) has PBR rule.

Source class ID	Destination class ID	Action
100 (C)	200 (P)	PBR to FW1 or FW2 external
200 (P)	100 (C)	PBR to FW1 or FW2 internal

Provider leaf switch (Leaf4) has a PBR rule.

Source class ID	Destination class ID	Action
201 (P')	101 (C')	PBR to FW3 or FW4 internal
101 (C')	201 (P')	PBR to FW3 or FW4 external

Figure 4: Across Site Return Traffic



North-South Intra-VRF/Non-Shared Service

This is the use case for north-south communication between endpoints in the datacenter and outside. Each site has L3out accessing to EPG WEB that is deployed across sites and the firewall needs to be inserted between them. This is a common design for the web front-end that is accessible from client outside of the data center.

Requirements:

- Both L3Out and EPG-WEB are same VRF.
- VRF needs to be in ingress mode (which is by default). In egress mode does not work, because of asymmetric traffic flow.
- PBR policies are needed on both connectors of the service graph.

Figure 5: Intra-VRF, ingress enforcement L3out is the consumer - IncommingTraffic

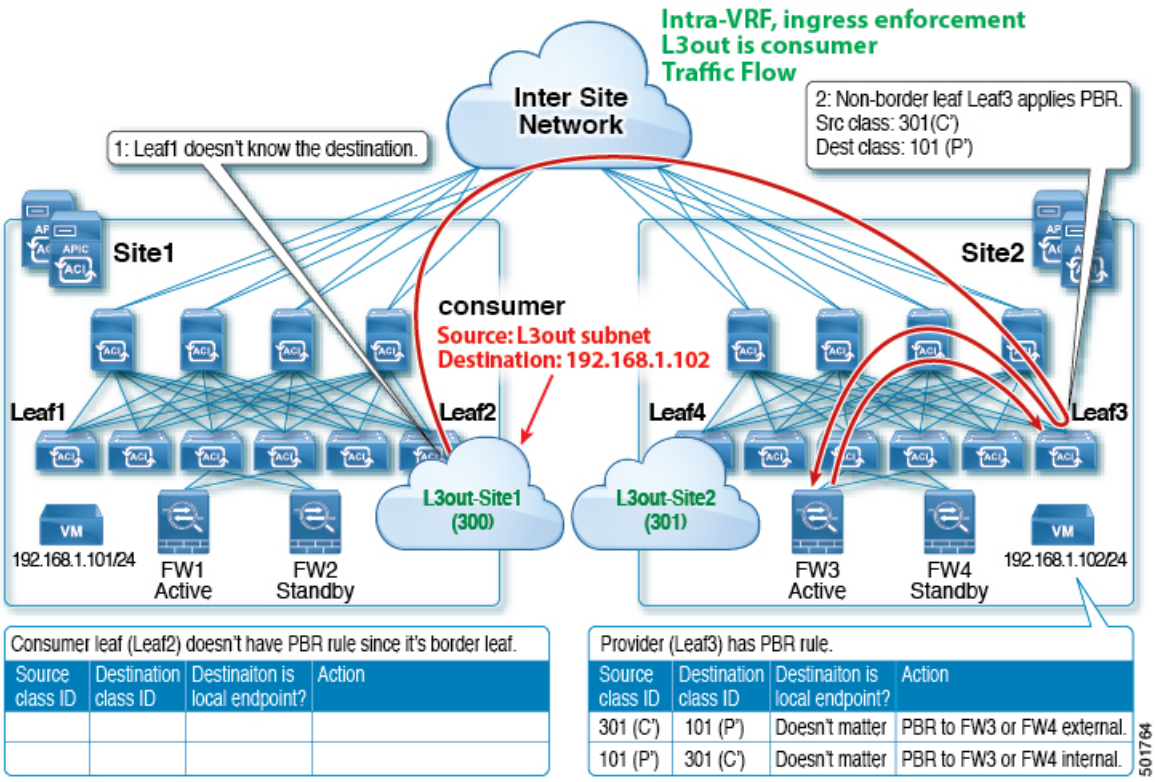
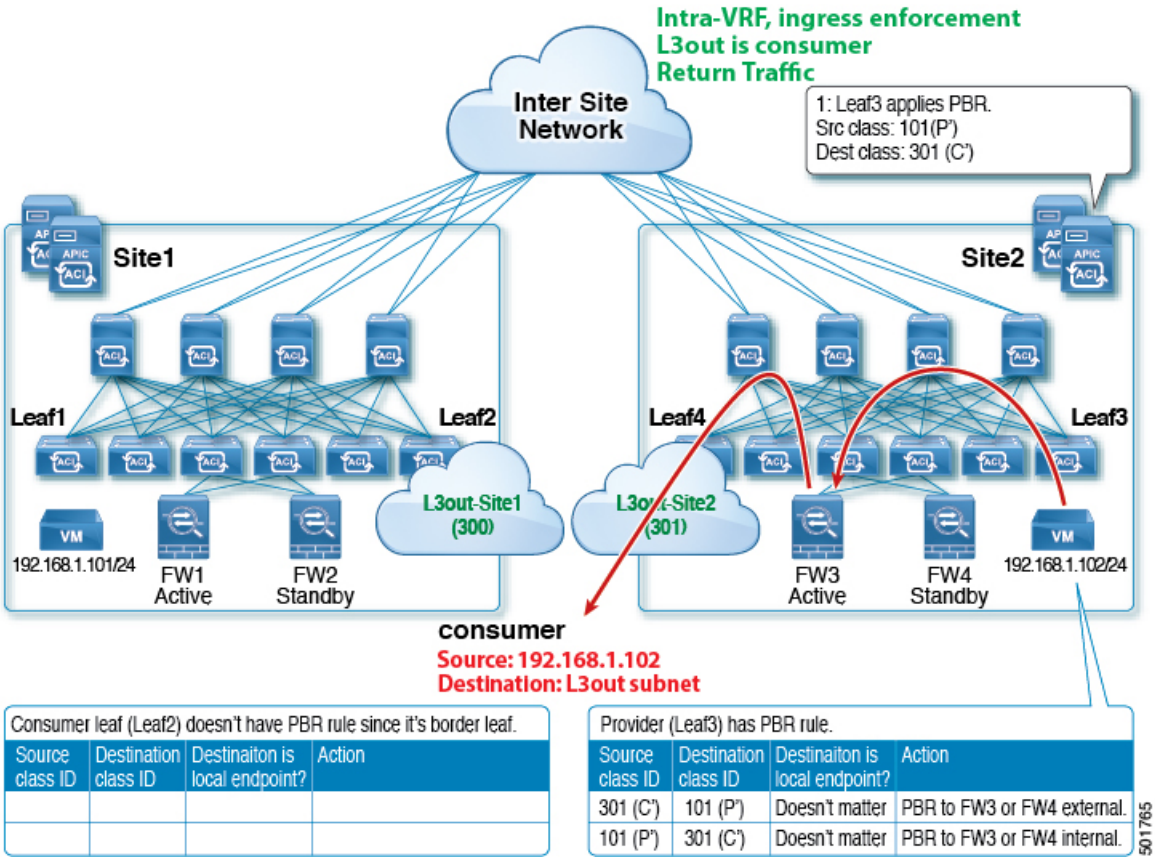


Figure 6: Intra-VRF, ingress enforcement L3out is the consumer - Return Traffic



Limitations

In Multi-Site deployments, when the contract is deployed with a service graph, Scope of the contract should be defined in such a way so that all providers are in the same scope. This creates single graph instance for all the providers.

In Release 1.2(1), Multi-Site does not support handling multiple graph instances created due to provider being in a scope that is bigger than contract scope. While deciding on the contract scope note the following:

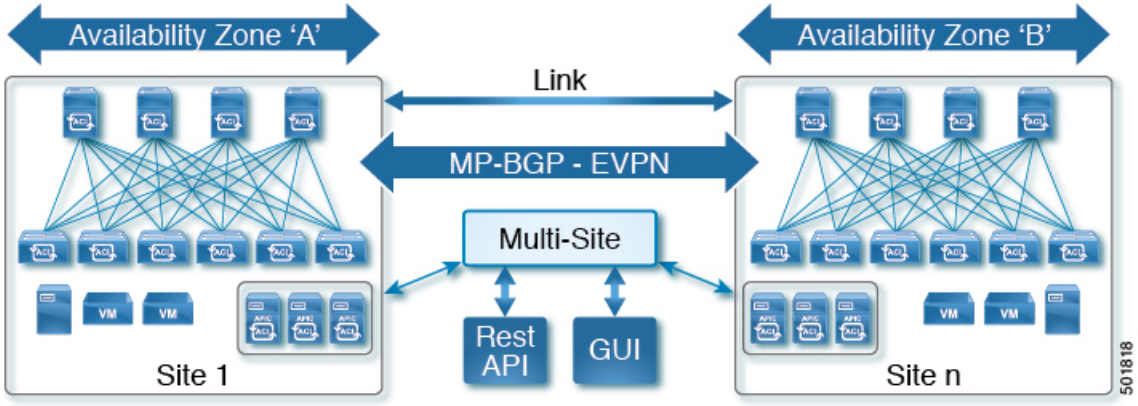
- If all the providers are in a single VRF, contract scope can be set to any available option.
- If providers are in multiple VRFs, but in a single Tenant, scope should be set to “Tenant”.
- If providers are in multiple tenants, this scenario is not supported with a single contract.

To overcome this limitation multiple contracts can be created with the same service graph attached.

Cisco ACI Multi-Site Back-to-Back Spine Connectivity Across Sites Without IPN

This Cisco ACI Multi-Site use case provides support for direct connection between spines of 2 different sites without any IPN between the sites.

Figure 7: Multi-Site Back to Back Spine – Basic Setup



This use case enables:

- Support for direct connection between spines of 2 different sites without any IPN between the sites
- Support for only a single POD per site deployments
- Requires unique fabric names across sites

Design:

- LLDP will detect spine to spine connection and will create a wiring issue on that port
- DHCP relay will not be configured on the link
- When the LLDP detects unique fabric names and when the spines on both sides are discovered, the port will be put back in-service except for the following:
  - ISIS will not be enabled on the link
  - Infra VLAN will not be learned from the link
  - LLDP TLV will be between the sites and will be ignored
- Spine-to-spine link will be treated as external subinterface
- The configuration and data path will be same as a regular Multi-Site set up

Limitations

- With back-to-back connectivity, we recommend that you deploy multiple spines in each site to provide inter-site connectivity. From each of these spines, provide multiple links to each of the spines in each of the other sites.
- In the hybrid case where IPN is also used, we recommend that all sites have to be connected to the IPN in a fault-tolerant fashion to avoid transit situation.
- Only two sites are supported with back-to-back spine.
- No new configuration required in APIC for this use case.

Troubleshooting

This section describes troubleshooting techniques.

- In APIC, check if l3extOut is configured for this interface in both sites.
- If there is no reachability between the two site spines, perform the following:
  - Make sure there are no wiring issues, the port is up and switchingSt is enabled:

```
dev-infra1-spine1# cat /mit/sys/lldp/inst/if-[eth1--1]/summary | grep wiringIssues
wiringIssues :
dev-infra1-spine1#
```

  - Make sure the IP address is assigned from the l3extOut configuration and OSPF session is up:

```
IP Interface Status for VRF "overlay-1"
eth1/53.7, Interface status: protocol-up/link-up/admin-up, iod: 63, mode: external
```

- Check the `svc_ifc_policyelem.log*` file in the SPINE that is connected to the other site:

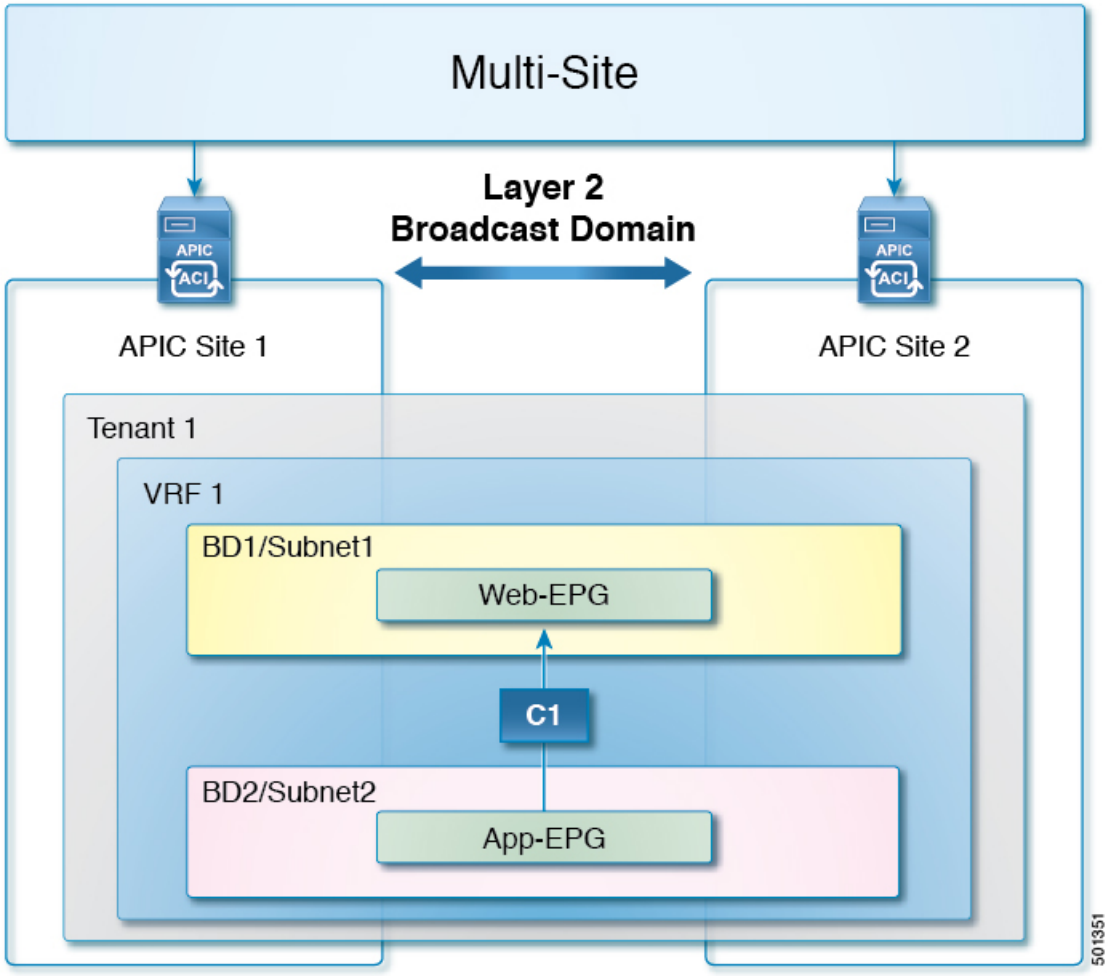
```
$ zgrep "back to back spine ignore wiring check." \
/var/sysmgr/tmp_logs/dme_logs/svc_ifc_policyelem.log*
```

Stretched Bridge Domain with Layer 2 Broadcast Extension

This is the most basic Cisco ACI Multi-Site use case, in which a tenant and VRF are stretched between sites. The EPGs in the VRF (with their bridge domains (BDs) and subnets), as well as their provider and consumer contracts are also stretched between sites.

In this use case, Layer 2 broadcast flooding is enabled across fabrics. Unknown unicast traffic is forwarded across sites leveraging the Head-End Replication (HER) capabilities of the spine nodes that replicate and send the frames to each remote fabric where the Layer 2 BD has been stretched.

Figure 8: Stretched Bridge Domain with Layer 2 Broadcast Extension



This use case enables:

- Same application hierarchy deployed on all sites with common policies. This allows seamlessly deploying workloads belonging to the various EPGs across different fabrics and governing their communication with common and consistent policies.
- Layer 2 clustering
- Live VM migration
- Active/Active high availability between the sites
- Using Service Graphs to push shared applications between sites is not supported.

Prerequisites for this Use Case

- Sites have been added, APIC controllers are active, and communications are established.
- The tenant to be stretched has been created.
- The Multi-Site Site and Tenant Manager account is available

Single profile including the objects in the following table, pushed to multiple sites:

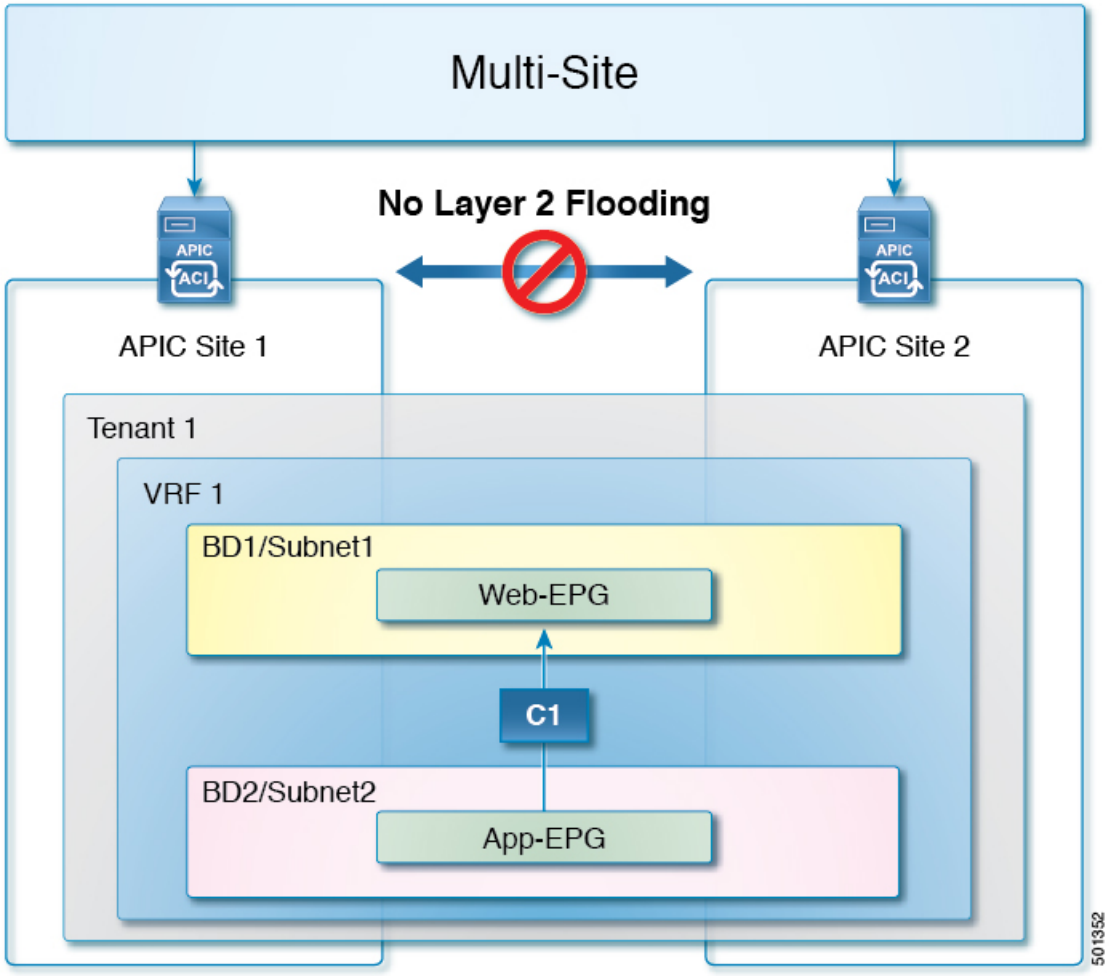
Table 1: Features to be Configured for this Use Case

Configuration	Description	Stretched or Local
Tenant	Imported from APIC or created in Multi-Site	Stretched
Site L3Outs	Configured in the APIC GUI and linked in the stretched tenant and VRF, site-specific templates	Local
VRF	VRF for the tenant	Stretched
Bridge Domain	Layer 2 stretching enabled Layer 2 flooding enabled Subnets to be shared added	Stretched
EPGs	EPGs in the BD	Stretched
Contracts	Include the filters needed to govern EPG communication	Stretched
External EPGs	Network Mappings of Site L3Outs (Cisco ACI Multi-Site, Release 1.0(1) Site Connections of Site L3Outs through External EPGs (Multi-Site, Release 1.0(2)	Local, but linked to other sites

Stretched Bridge Domain with No Layer 2 Broadcast Extension

This Cisco ACI Multi-Site use case is similar to the first use case where a tenant, VRF, and their EPGs (with their bridge domains and subnets) are stretched between sites.

Figure 9: Stretched Bridge Domain with No Layer 2 Broadcast Extension



However, in this use case, Layer 2 broadcast flooding is localized at each site. Layer 2 broadcast, multicast and unknown unicast traffic is not forwarded across sites over replicated VXLAN tunnels.

This use case enables:

- Control plane overhead is reduced by keeping Layer 2 flooding local
- Inter-site IP mobility for disaster recovery
- "Cold" VM Migration
- Using Service Graphs to push shared applications between sites is not supported.

Prerequisites for this Use Case

- Sites have been added, APIC controllers are active, and communications are established.
- The tenant to be stretched has been created.
- The Multi-Site Site and Tenant Manager account is available

Profile with the objects in the following table, pushed to multiple sites:

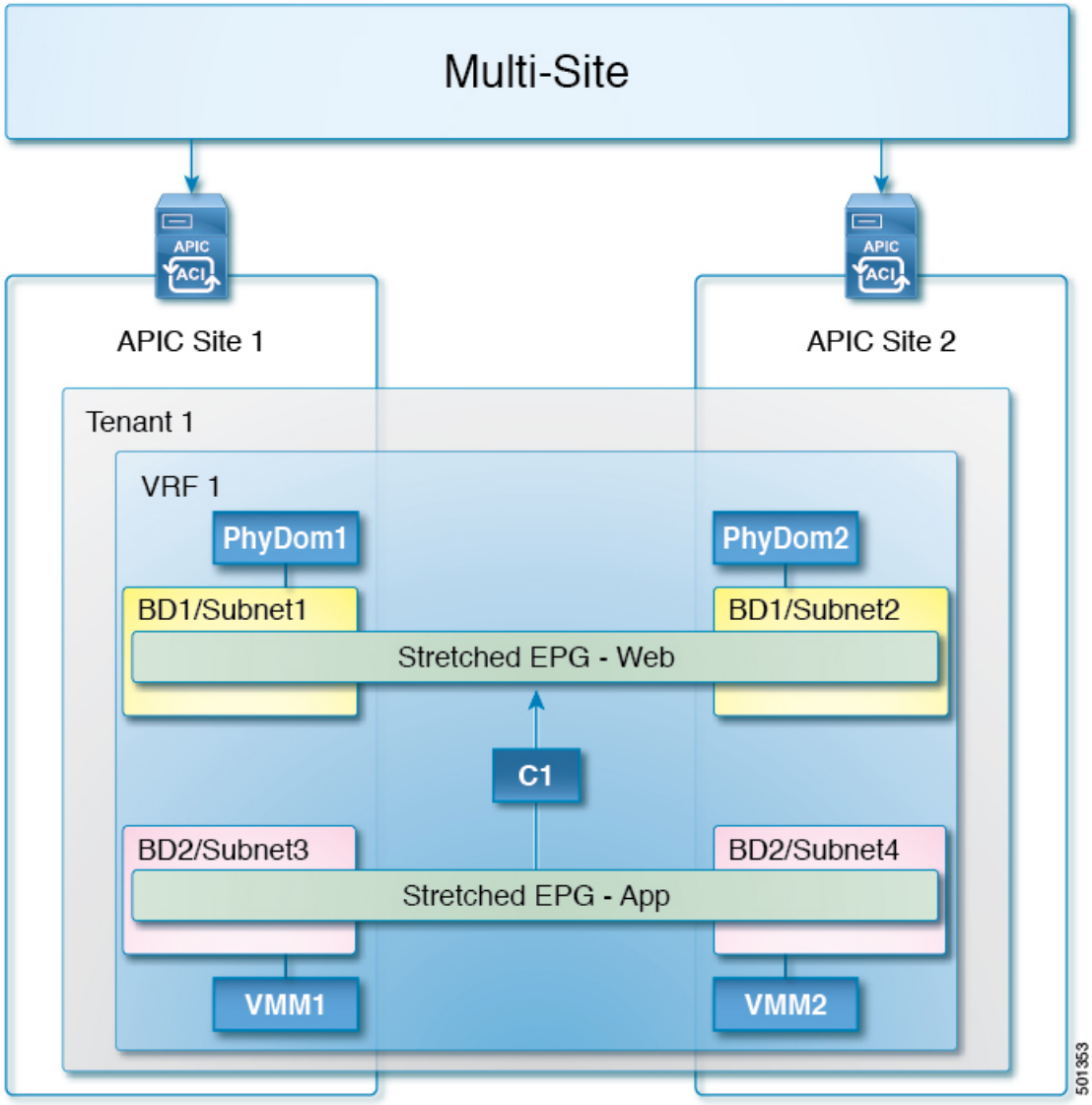
Table 2: Features to Be Configured for this Use Case

Configuration	Description	Stretched or Local
Tenant and VRF	Imported from APIC or created in Multi-Site	Stretched
Site L3Outs	Configured in the APIC GUI and linked in the stretched tenant and VRF, site-specific templates	Local
Bridge Domain	Layer 2 stretching enabled Layer 2 flooding disabled Subnets to be shared added	Stretched
EPGs	All EPGs in the BD	Stretched
Contracts	Include whatever filters and contracts are needed to govern EPG communication	Stretched
External EPGs	Network Mappings of Site L3Outs (Cisco ACI Multi-Site, Release 1.0(1) Site Connections of Site L3Outs through External EPGs (Multi-Site, Release 1.0(2)	Local, but linked to other sites

Stretched EPG Across Sites

This Cisco ACI Multi-Site use case provides endpoint groups (EPGs) stretched across multiple sites. Stretched EPG is defined as an endpoint group that expands across multiple sites where the underlying networking, site local, and bridge domain can be distinct.

Figure 10: Stretched EPG Across Sites



This use case enables Layer 3 forwarding to be used among all sites.

Prerequisites for this Use Case

- Sites have been added, APIC controllers are active, and communications are established.
- The relevant tenants have been created.
- The Multi-Site Site and Tenant Manager account is available
- A physical domain and VMM domain must exist on APIC.

Profiles pushed to single or multiple sites, including the objects in this table:

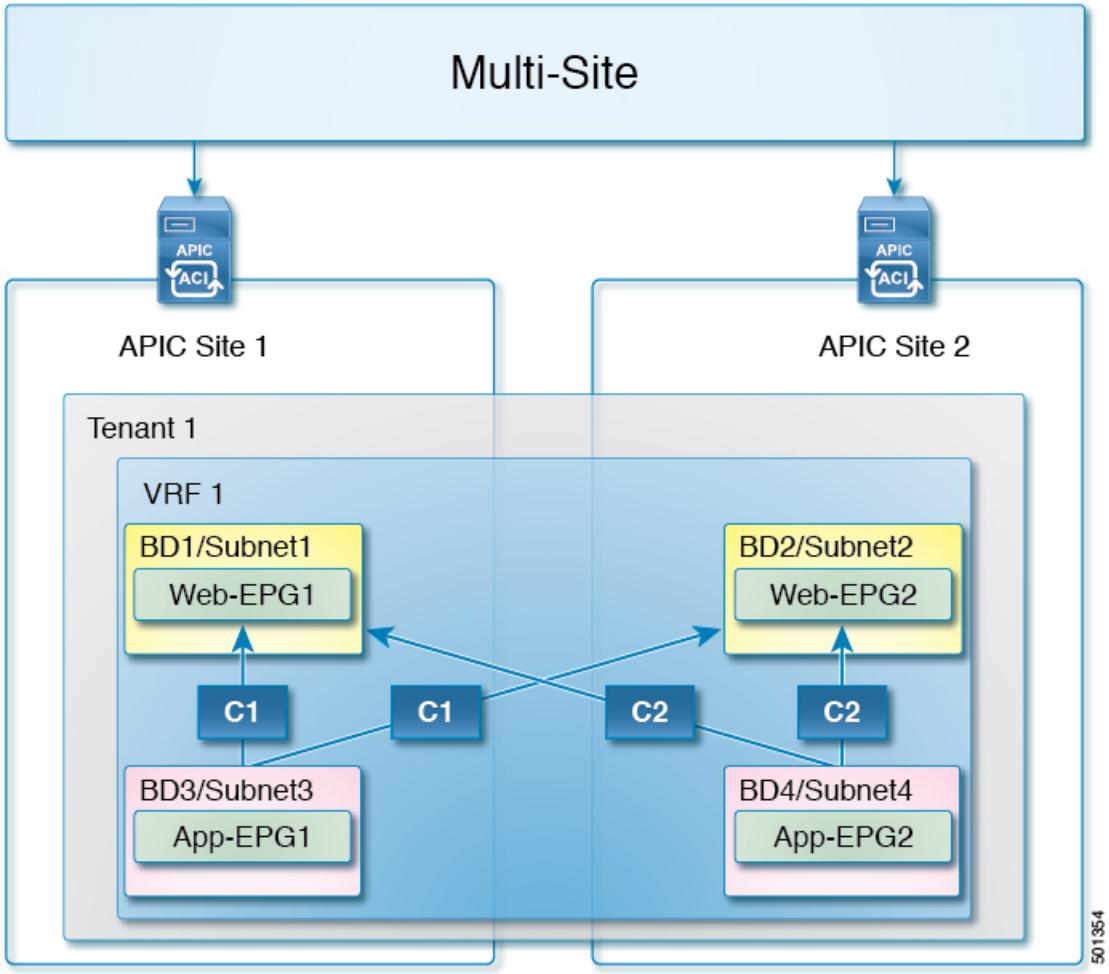
Table 3: Features to be Configured for this Use Case

Configuration	Description	Stretched or Local
Tenant, VRF and EPGs	Imported from APIC or created in Multi-Site.	Stretched
Site L3Outs	Configured in the APIC GUI and linked in the stretched tenant and VRF, site-specific templates	Local
Bridge Domains (DBs)	Layer 2 stretching disabled.	Stretched
Subnets	Unique for each BD on the local site.	Local
Contract	Contracts configured on site where they are provided	Local
External EPGs	Network Mappings of Site L3Outs (Cisco ACI Multi-Site, Release 1.0(1) Site Connections of Site L3Outs through External EPGs (Multi-Site, Release 1.0(2)	Local, but linked to other sites

Stretched VRF with Inter-Site Contracts

This Multi-Site use case provides inter-site communication between endpoints connected to different Bridge Domains (BDs) that are part of the same stretched VRF. VRF Stretching is a convenient way to manage EPGs across sites (and the contracts between them).

Figure 11: VRF Stretching with Inter-site Contracts



In the diagram above, the App-EPGs provide the C1 and C2 contracts across the sites, and the Web-EPGs consume them across the sites.

This use case has the following benefits:

- The tenant and VRF are stretched across sites, but EPGs and their policies (including subnets) are locally defined.
- Because the VRF is stretched between sites, contracts govern cross-site communication between the EPGs. Contracts can be consistently provided and consumed within a site or across sites.
- Traffic is routed within and between sites (with local subnets) and static routing between sites is supported.
- Separate profiles are used to define and push local and stretched objects.
- No Layer 2 stretching and local Layer 2 Broadcast domains.
- “Cold” VM migration, without the capability of preserving the IP address of the migrated endpoints.
- Using Service Graphs to push shared applications between sites are not supported.

Prerequisites for this Use Case

- Sites have been added, APIC controllers are active, and communications are established.
- The tenants to be stretched have been created.
- The Multi-Site Site and Tenant Manager account is available.

Profiles pushed to single or multiple sites, including the objects in this table:

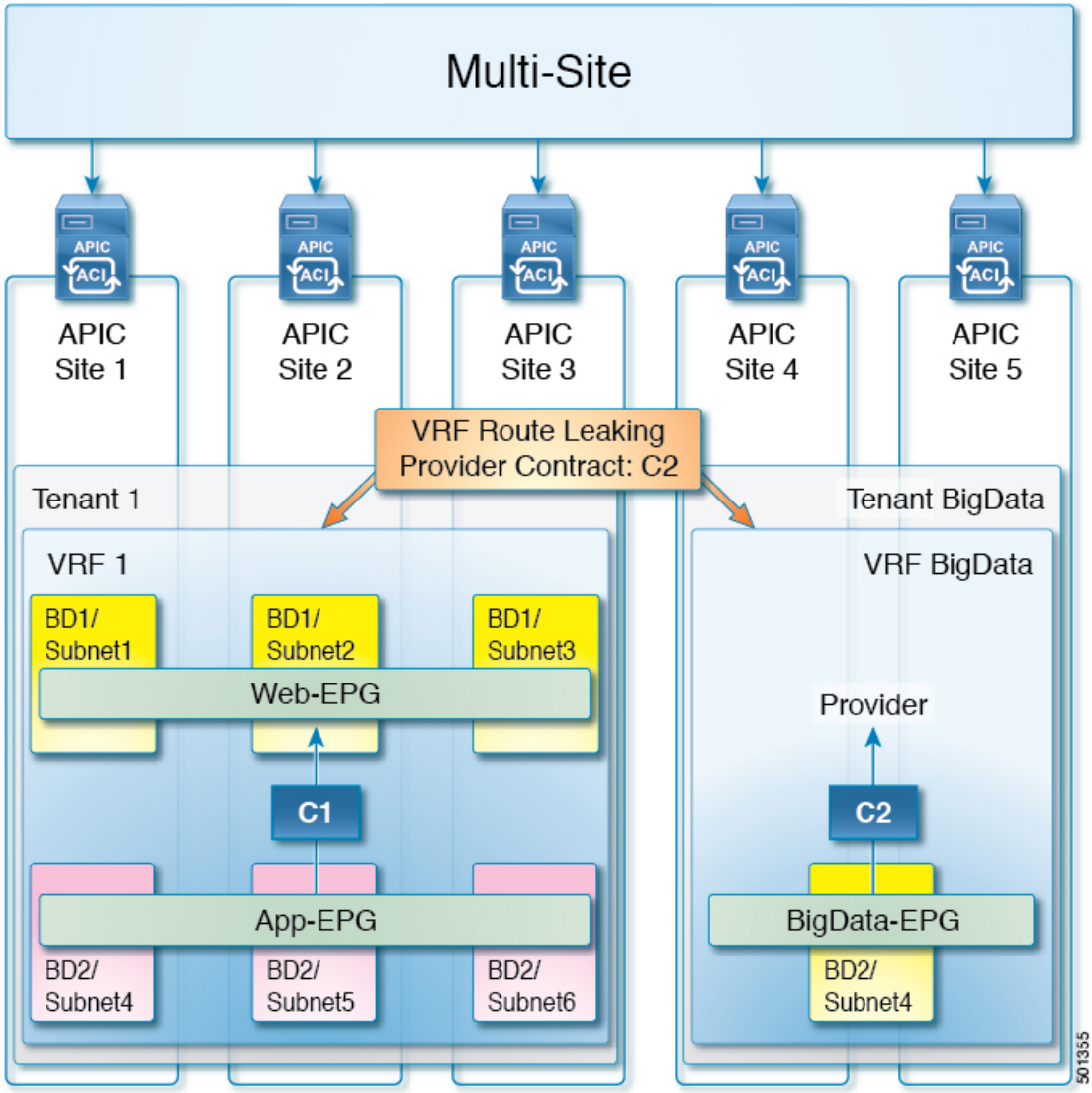
Table 4: Features to be Configured for this Use Case

Configuration	Description	Stretched or Local
Tenant and VRF	Imported from APIC or created in Multi-Site	Stretched
Site L3Outs	Configured in the APIC GUI and linked in the stretched tenant and VRF, site-specific templates	Local
EPGs providing contracts	EPGs for each site that provides services.	Local
EPGs consuming contracts	EPGS that consume the provided contracts, may be in the same site or multiple sites	Local
Bridge Domains for each EPG	Layer 2 stretching disabled Layer 2 flooding disabled	Local
Contracts	Contracts configured on site where they are provided	Local, but shared
External EPGs	Network Mappings of Site L3Outs (Cisco ACI Multi-Site, Release 1.0(1) Site Connections of Site L3Outs through External EPGs (Multi-Site, Release 1.0(2)	Local, but linked to other sites

Shared Services with Stretched Provider EPG

In this use case, the Provider EPGs in one group of sites offer shared services and the EPGs in another group of sites consume the services. All sites have local EPGs and bridge domains.

Figure 12: Shared Services with Stretched Provider EPG



In the diagram above, Site 4 and Site 5 (with BigData-EPG, in Tenant BigData/VRF BigData), provides shared data services, and the EPGs in Site 1 to Site 3, in Tenant 1/VRF 1, consume the services.

In the Shared Services usecase of Multi-Site, at the VRF boundary routes are leaked between VRFs for routing connectivity and by importing contracts across sites.

This use case has the following benefits:

- Shared services enable communications across VRFs and tenants while preserving the isolation and security policies of the tenants.
- A shared service is supported only with non-overlapping and non-duplicate subnets.
- Each group of sites has a different tenant, VRF, and one or more EPGs stretched across it.
- Site groups can be configured to use Layer 2 Broadcast extensions or to localize Layer 2 flooding.

- Stretched EPGs share the same bridge domain, but the EPGs have subnets that are configured under the EPG, not under the bridge domain.
- The provider contract must be set to global scope.
- VRF route leaking enables communication across the VRFs.
- Using Service Graphs to push shared applications between sites is not supported.

Prerequisites for this Use Case

- Sites have been added, APIC controllers are active, and communications are established.
- The relevant tenants have been created.
- The Multi-Site Site and Tenant Manager account is available

Schemas, with templates, pushed to groups of sites, including the objects in this table:

Table 5: Features to be Configures for this Use Case

Configuration	Description	Stretched or Local
Shared service provider schema, with multiple templates	Shared template, includes the following objects: • Tenant • VRF • Provider Contract with global scope. • EPG with subnet set to **Advertised Externally** and **Shared Between VRFs**. Site-Specific templates, including bridge domains (optionally set for Layer 2 extension) and external EPGs	Stretched (pushed to all sites in the provider group)

Configuration	Description	Stretched or Local
Shared service consumer schema with multiple templates	Shared template, includes the following objects: • Tenant • VRF • EPG with subnet set to **Advertised Externally** and **Shared Between VRFs**. **Note** For the consumer EPGs, the subnets can alternatively be added in the BDs. • Consumer Contract (same name as the provided contract). Site-Specific templates, including bridge domains (optionally set for Layer 2 extension) and external EPGs	Stretched or local
VRF route leaking	Contracts must be configured to enable VRF route leaking.	Configured cross-site

Migration of Cisco ACI Fabric to Cisco ACI Multi-Site

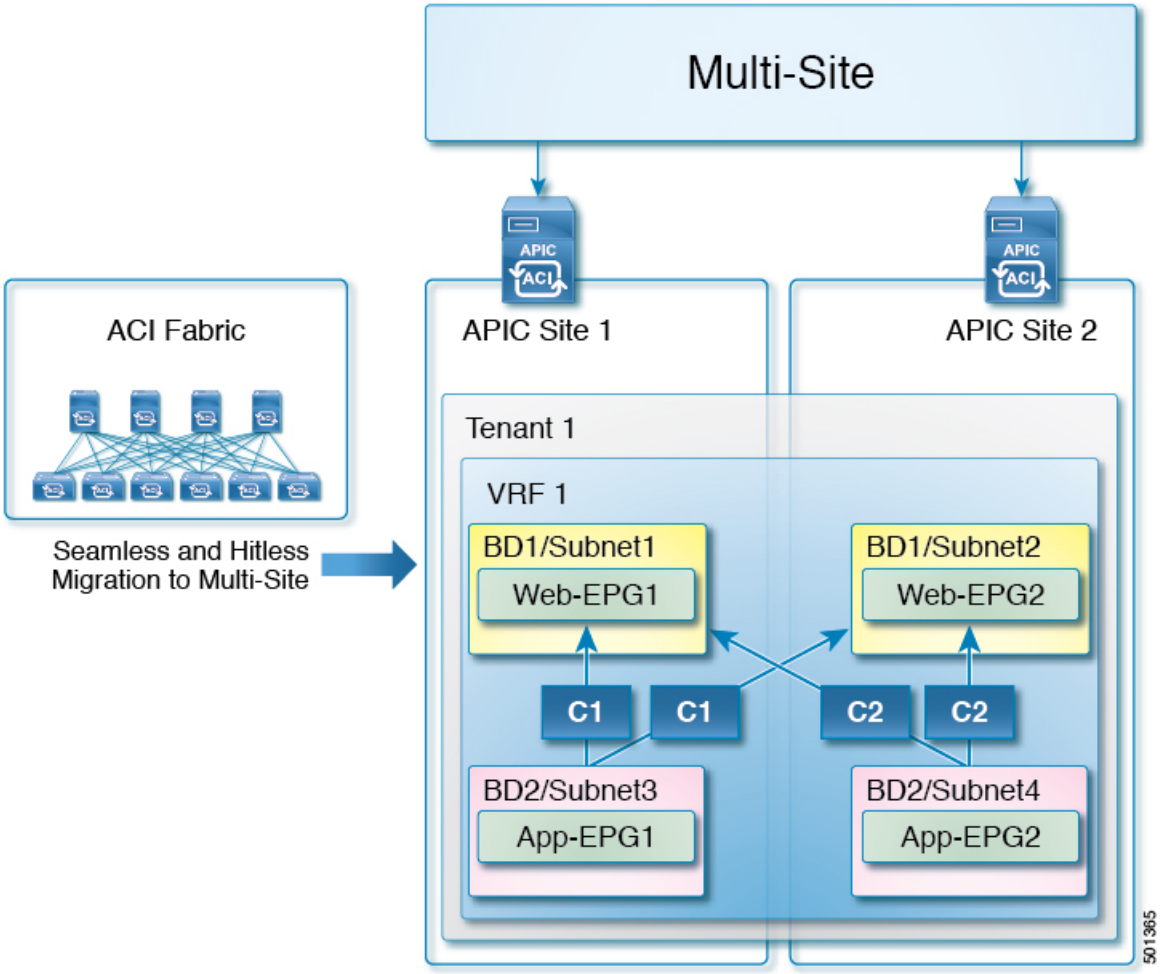
This is a common Cisco ACI Multi-Site use case, in which a tenant is migrated or imported from Cisco ACI fabric to Cisco ACI Multi-Site.

This use case is targeted for Brownfield to Greenfield and Greenfield to Greenfield types of deployments. The Brownfield to Brownfield use case is only supported in this release if both Cisco APIC sites are deployed with the same configuration. Other Brownfield to Brownfield use cases will be deployed in a future Cisco ACI Multi-Site release.

For Brownfield configurations, two scenarios are considered for deployments:

- A single or multiple pod ACI fabric is in place already. You can add another site in a Multi-Site configuration.
- Two ACI fabrics are in place already, the objects (tenants, VRFs, and EPGs) across sites are initially defined with identical names and policies, and they are connected leveraging a traditional L2/L3 DCI solution. You can convert this configuration to Multi-Site as explained in the following configuration diagram:

Figure 13: Migration of Cisco ACI Fabric to Cisco ACI Multi-Site



Setting up Cisco ACI Multi-Site with Multipod-Enabled Fabrics

Starting in release 1.2(1), two use cases add support for setting up Cisco ACI Multi-Site with multipod-enabled fabrics.

Guidelines and limitations for these two use cases:

- Only the following switches will be connected to the IPN/ISN:
  - Cisco Nexus 93180LC-EX, 93180YC-EX, and 93108TC-EX switches.
  - Cisco Nexus 9504, 9408, and 9516 switches with the following line cards:
    - X9736C-EX
    - X97160YC-EX
    - X9732C-EX
    - X9732C-EXM

    - Remove IPN links from old generation spine switches.

- The same IPN/ISN will be used for multipod and Multi-Site.
- In a Cisco ACI Multi-Site deployment, you cannot use an overlapping tunnel endpoints (TEP) pool range and GIPO pool range on the 2 sites using a single IPN/ISN.

When a tenant is imported from the Cisco APIC GUI, all the objects associated with the tenant are imported in Cisco ACI Multi-Site:

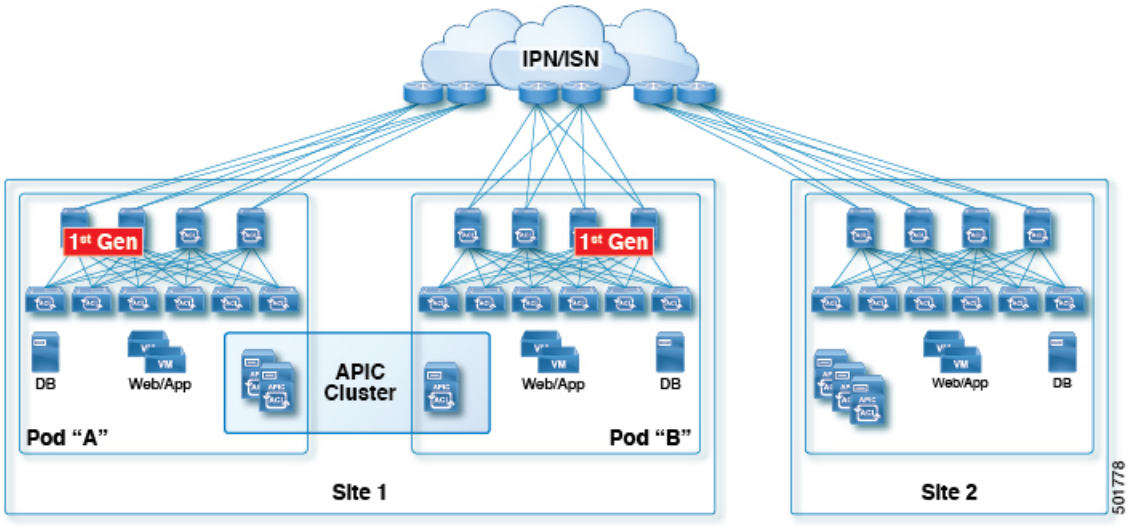
Table 6: Features to be configured for these use cases

Configuration	Description	Stretched or Local
Tenant	Create a tenant in Cisco ACI Multi-Site and import the tenant policies from the Cisco APIC	Stretched
VRF	VRF instance for the tenant	Stretched
Bridge Domain	Layer 2 stretching enabled Layer 2 flooding enabled Subnets to be shared added	Stretched
EPGs	EPGs in the BD	Stretched
Contracts	Include the filters needed to govern EPG communication	Stretched
Site L3Outs	Configured in the Cisco APIC and linked with external EPGs	Local

Adding a Multipod Fabric as a Site on Cisco ACI Multi-Site

This section describes an overview of how to add a multipod fabric as a site on Cisco ACI Multi-Site.

Figure 14: Cisco ACI fabric with multiple PODs as a site in Cisco ACI Multi-Site



Fabric with multiple pods add as a site in Cisco ACI Multi-Site.

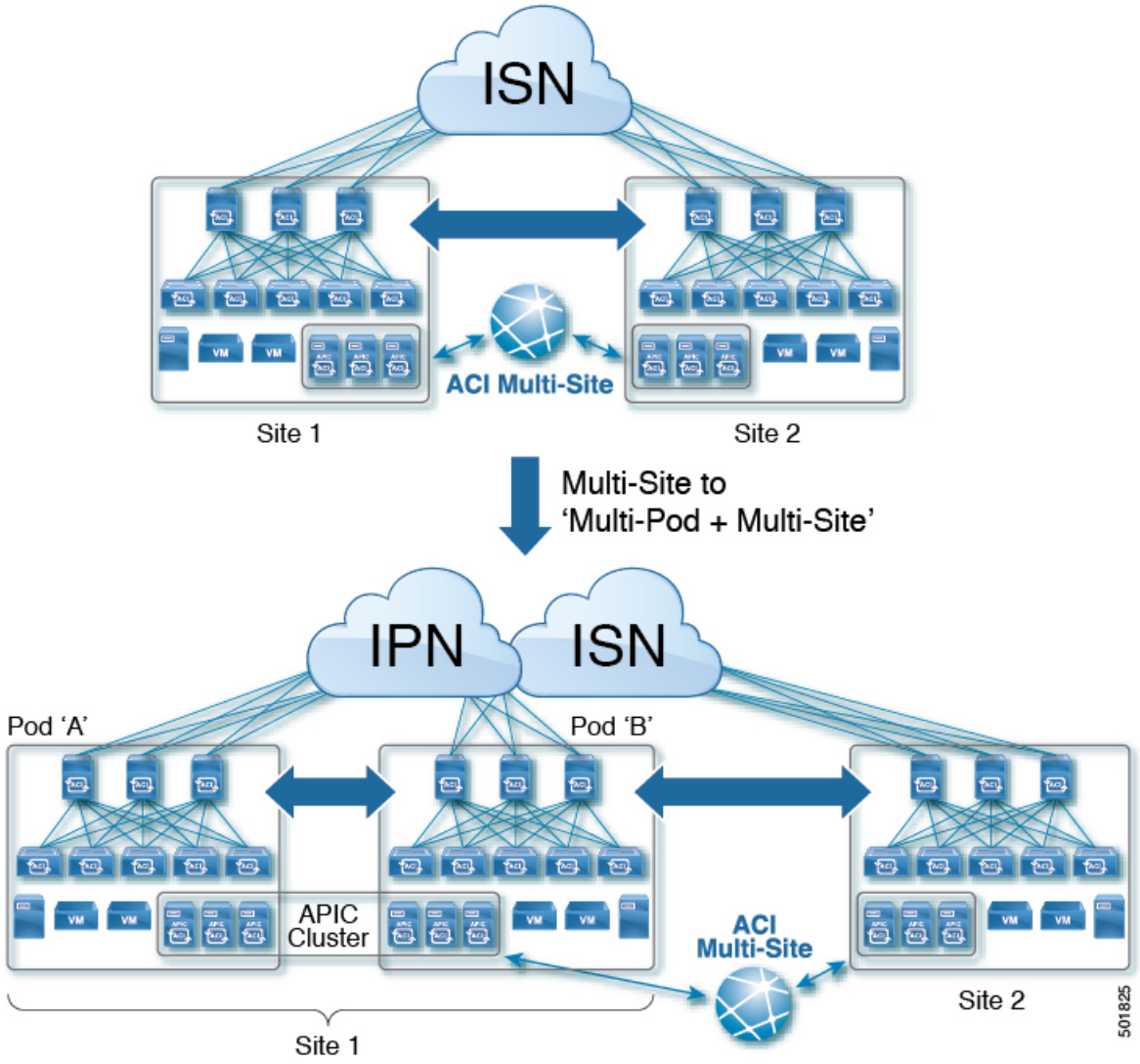
This is an overview of the procedure:

- Add a multipod-enabled fabric as a site in Cisco ACI Multi-Site.
  - Cisco ACI Multi-Site discovers common configurations for Cisco ACI Multi-Site and multipod, such as spine to IPN links configuration, OSPG information, BGP information and auto-populates in the Cisco ACI Multi-Site infra configuration.
- Provide Cisco ACI Multi-Site-specific configurations such as MCAST TEP, MSITE DP-TEP, or MSITE CP-TEP and enable Multi-Site for the site in Cisco ACI Multi-Site infra page.
  - You can also configure for Cisco ACI Multi-Site the same DP-TEP/CP-TEP that you configured for multipod.
- Apply the infra configuration in Cisco ACI Multi-Site.
  - Cisco ACI Multi-Site configures Cisco APIC with Cisco ACI Multi-Site-specific configrations and common configrations for Cisco ACI Multi-Site and multipod, such as spine to IPN links config, OSPF information, and BGP information, and will not configure multipod-specific configration.
  - Cisco ACI Multi-Site uses the same infra L3Out used for multipod to configure Cisco ACI Multi-Site. Cisco ACI Multi-Site determines it based on fabricExtCtrlPeering=yes and fabricExtIntersiteCtrlPeering=yes under l3extInfraNodeP in the infra L3Out.
  - You can configure GOLF for an L3Out that you use with Cisco ACI Multi-Site by using one of the following configurations:
    - One L3Out for Cisco ACI Multi-Site, multipod, and GOLF, and different (zero or more) L3Outs for GOLF.
    - One L3Out for Cisco ACI Multi-Site, multipod and different (zero or more) L3Outs for GOLF.

Converting a Single POD Site to Multipod

This section describes an overview of how to convert a single POD site to multipod.

Figure 15: Coverting a single POD site in Multi-Site to a multipod site



Converting a single POD site to Multipod.

This is an overview of the procedure:

- Use the same spine nodes and uplinks for both communications.
- Use Cisco APIC to configure multipod. Use the same infra L3Out used for Cisco ACI Multi-Site for multipod also.
- You can use the same control-plane (CP) tunnel endpoints (TEP) and data-plane (DP) TEP for both multipod and Cisco ACI Multi-Site, or you can define separate TEPs for multipod and Cisco ACI Multi-Site.

- After configuring Cisco ACI Multi-Site, click on the "refresh" icon in the Cisco ACI Multi-Site infra page to discover the new pods.
- In Cisco ACI Multi-Site, provide Cisco ACI Multi-Site-specific configurations, such as Cisco ACI Multi-Site DP-TEP per pod and Cisco ACI Multi-Site CP-TEP.
- Apply infra.