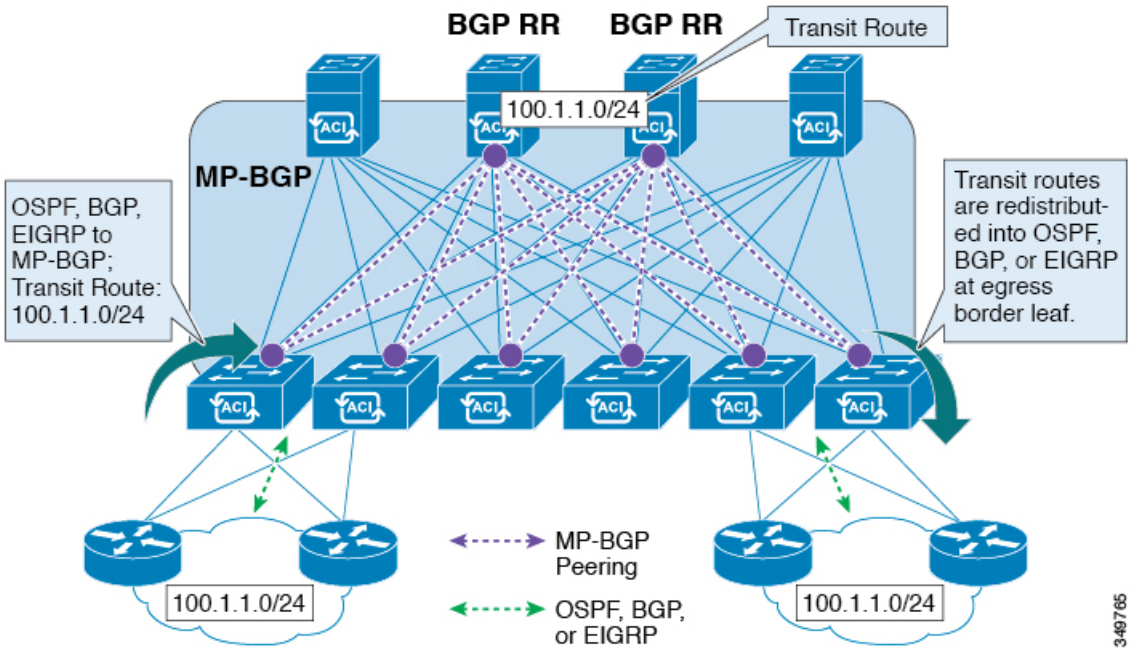
Transit Routing

This chapter contains the following sections:

Transit Routing in the ACI Fabric

The Cisco APIC software supports external Layer 3 connectivity with OSPF (NSSA) and iBGP. The fabric advertises the tenant bridge domain subnets out to the external routers on the External Layer 3 Outside (L3Out) connections. The routes that are learned from the external routers are not advertised to other external routers. The fabric behaves like a stub network that can be used to carry the traffic between the external Layer 3 domains.

Figure 1: Transit Routing in the Fabric



In transit routing, multiple L3Out connections within a single tenant and VRF are supported and the APIC advertises the routes that are learned from one L3Out connection to another L3Out connection. The external Layer 3 domains peer with the fabric on the border leaf switches. The fabric is a transit Multiprotocol-Border Gateway Protocol (MP-BGP) domain between the peers.

The configuration for external L3Out connections is done at the tenant and VRF level. The routes that are learned from the external peers are imported into MP-BGP at the ingress leaf per VRF. The prefixes that are learned from the L3Out connections are exported to the leaf switches only where the tenant VRF is present.

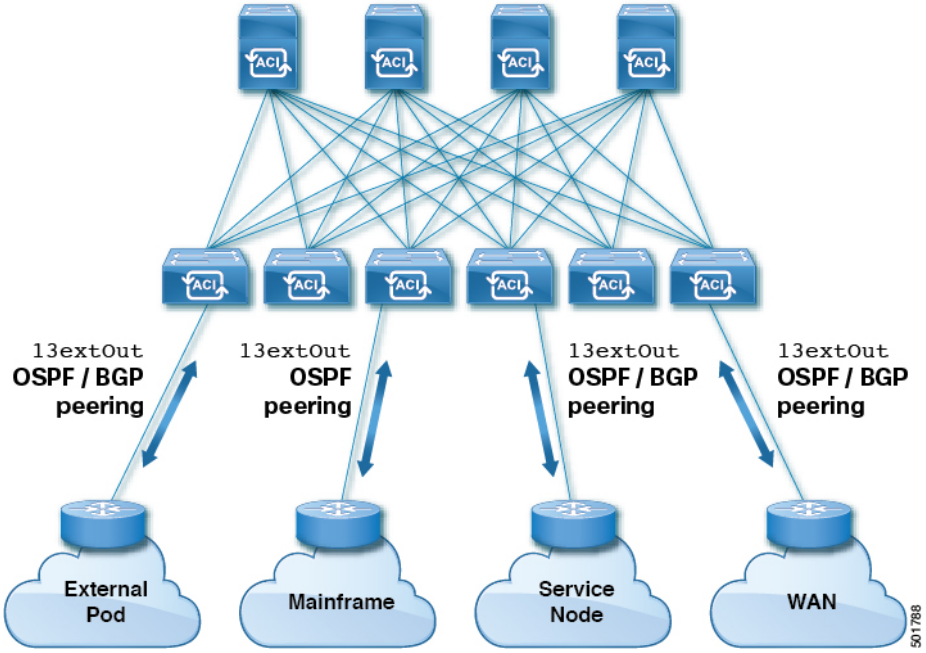
Note For cautions and guidelines for configuring transit routing, see Guidelines for Transit Routing, on page 9

Transit Routing Use Cases

Transit Routing Between Layer 3 Domains

Multiple Layer 3 domains such as external pods, mainframes, service nodes, or WAN routers can peer with the ACI fabric to provide transit functionality between them.

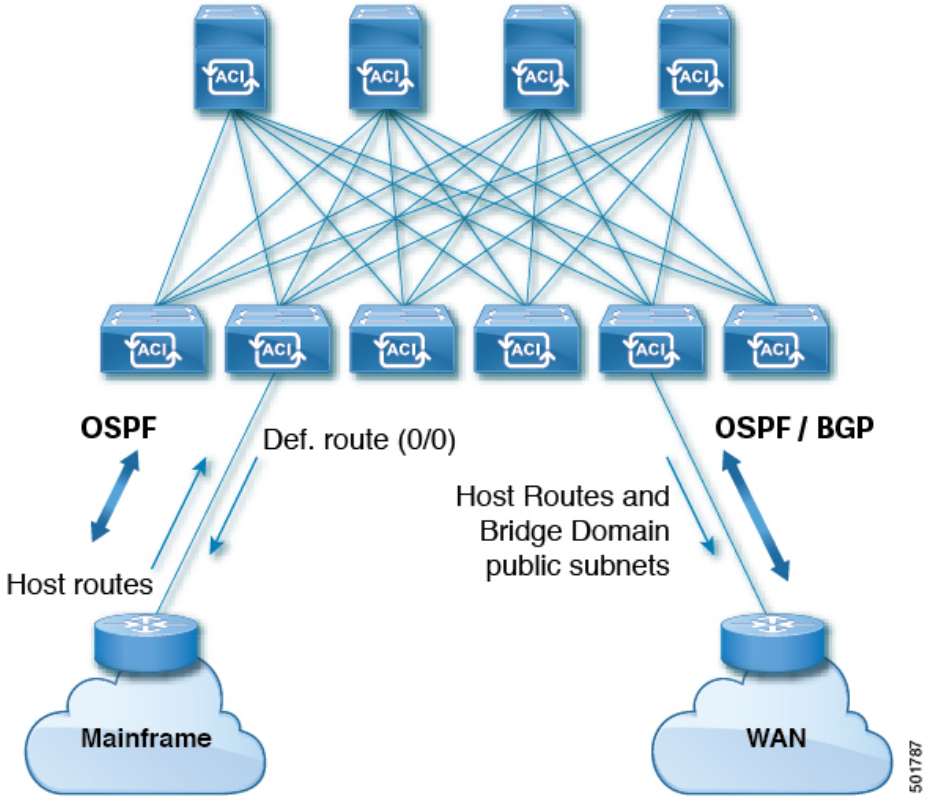
Figure 2: Transit Routing Between Layer 3 Domains



Mainframe Traffic Transiting the ACI Fabric

Mainframes can function as IP servers running standard IP routing protocols that accommodate requirements from Logical Partitions (LPARs) and Virtual IP Addressing (VIPA).

Figure 3: Mainframe Transit Connectivity

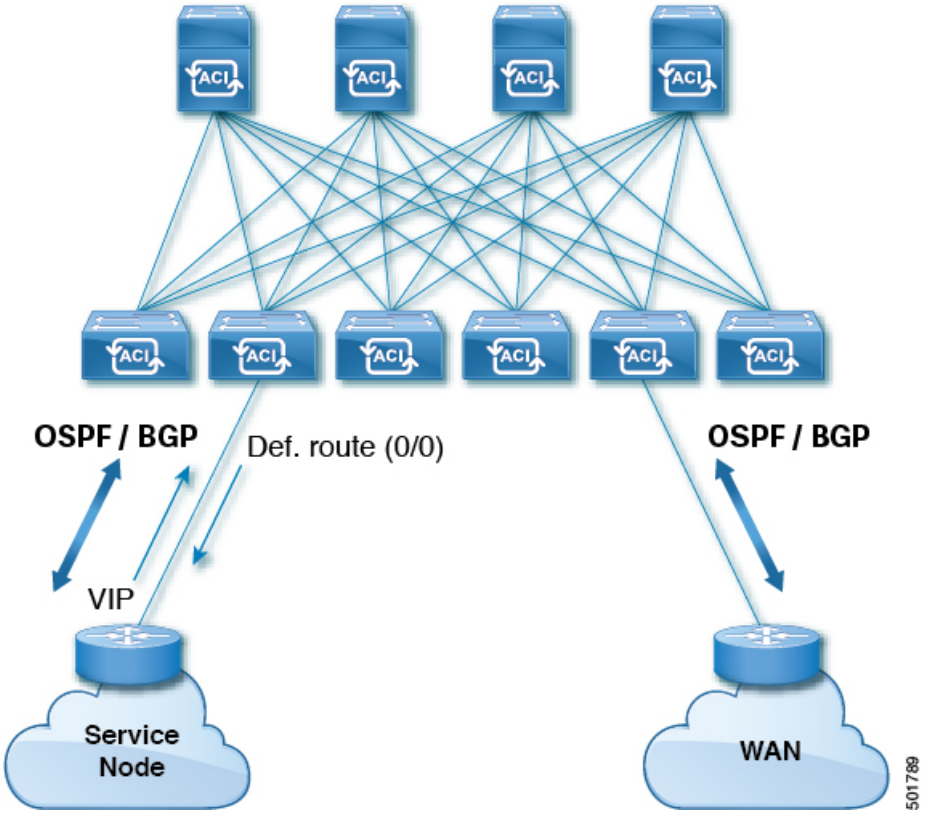


In this topology, mainframes require the ACI fabric to be a transit domain for external connectivity through a WAN router and for east-west traffic within the fabric. They push host routes to the fabric to be redistributed within the fabric and out to external interfaces.

Service Node Transit Connectivity

Service nodes can peer with the ACI fabric to advertise a Virtual IP (VIP) route that is redistributed to an external WAN interface.

Figure 4: Service Node Transit Connectivity

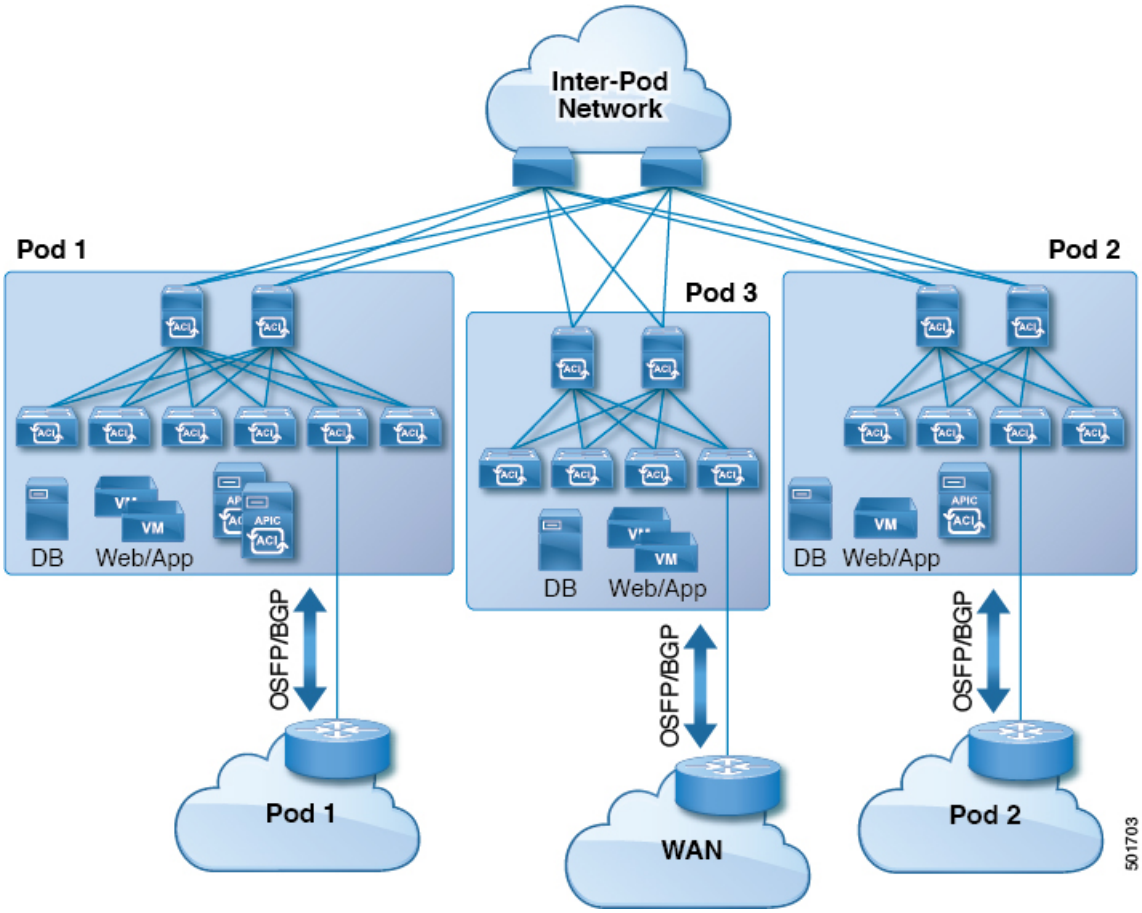


The VIP is the external facing IP address for a particular site or service. A VIP is tied to one or more servers or nodes behind a service node.

Multipod in a Transit-Routed Configuration

In a multipod topology, the fabric acts as a transit for external connectivity and interconnection between multiple pods. Cloud providers can deploy managed resource pods inside a customer datacenter. The demarcation point can be an L3Out with OSPF or BGP peering with the fabric.

Figure 5: Multiple Pods with L3Outs in a Transit-Routed Configuration



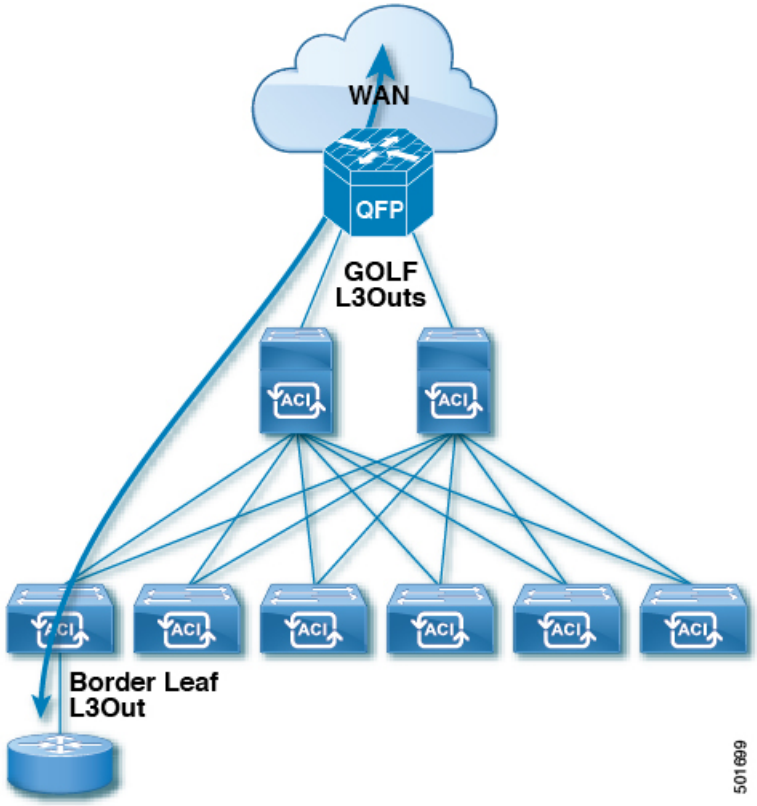
In such scenarios, the policies are administered at the demarcation points and ACI policies need not be imposed.

Layer 4 to Layer 7 route peering is a special use case of the fabric as a transit where the fabric serves as a transit OSPF or BGP domain for multiple pods. You configure route peering to enable OSPF or BGP peering on the Layer 4 to Layer 7 service device so that it can exchange routes with the leaf node to which it is connected. A common use case for route peering is Route Health Injection where the SLB VIP is advertised over OSPF or iBGP to clients outside the fabric. See *L4-L7 Route Peering with Transit Fabric - Configuration Walkthrough* for a configuration walk-through of this scenario.

GOLF in a Transit-Routed Configuration

In APIC, release 2.0 and later, the Cisco ACI supports transit routing with GOLF L3Outs (with BGP and OSPF). For example, the following diagram shows traffic transiting the fabric with GOLF L3Outs and a border leaf L3Out.

Figure 6: GOLF L3Outs and a Border Leaf L3Out in a Transit-Routed Configuration



Supported Transit Combination Matrix

Layer 3 Outside Connection Type	OSPF	iBGP			eBGP			EIGRP v4	EIGRP v6	Static Route
		iBGP over OSPF	iBGP over Static Route	iBGP over Direct Connection	eBGP over OSPF	eBGP over Static Route	eBGP over Direct Connection			
OSPF	Yes	Yes*	Yes	Yes* (tested in APIC release 1.3c)	Yes	Yes	Yes	Yes	Yes* (tested in APIC release 1.2g)	Yes

Layer 3 Outside Connection Type		OSPF	iBGP			eBGP			EIGRP v4	EIGRP v6	Static Route
			iBGP over OSPF	iBGP over Static Route	iBGP over Direct Connection	eBGP over OSPF	eBGP over Static Route	eBGP over Direct Connection			
iBGP	iBGP over OSPF	Yes*	X	X	X	Yes* (tested in APIC release 1.3c)	X	Yes	Yes	X	Yes
	iBGP over Static Route	Yes	X	X	X	Yes* (tested in APIC release 1.2g)	X	Yes* (tested in APIC release 1.2g)	Yes	X	Yes
	iBGP over Direct Connection	Yes	X	X	X	-	X	Yes* (tested in APIC release 1.2g)	Yes	X	Yes
eBGP	eBGP over OSPF	Yes	Yes* (tested in APIC release 1.3c)	Yes* (tested in APIC release 1.3c)	Yes* (tested in APIC release 1.3c)	Yes	Yes* (tested in APIC release 1.3c)	Yes* (tested in APIC release 1.3c)	Yes	X	Yes* (tested in APIC release 1.3c)
	eBGP over Static Route	Yes	X	X	X	Yes* (tested in APIC release 1.2g)	Yes (tested in APIC release 3.0)	Yes* (tested in APIC release 1.2g)	Yes	X	Yes
	eBGP over Direct Connection	Yes	Yes	Yes	Yes* (tested in APIC release 1.3c)	Yes* (tested in APIC release 1.3c)	Yes* (tested in APIC release 1.3c)	Yes	Yes	X	Yes
EIGRPv4		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes (tested in APIC release 1.3c)	X	Yes

Layer 3 Outside Connection Type	OSPF	iBGP			eBGP			EIGRP v4	EIGRP v6	Static Route
		iBGP over OSPF	iBGP over Static Route	iBGP over Direct Connection	eBGP over OSPF	eBGP over Static Route	eBGP over Direct Connection			
EIGRPv6	Yes (tested in APIC release 1.2g)	X	X	X	X	X	X	X	Yes (tested in APIC release 1.3c)	Yes (tested in APIC release 1.2g)
Static Route	Yes	Yes	Yes	Yes	Yes (tested in APIC release 1.3c)	Yes	Yes	Yes	Yes (tested in APIC release 1.2g)	Yes

- connec. = connection
- * = Not supported on the same leaf switch
- X = Unsupported/Untested combinations

Transit Routing Guidelines

Guidelines for Transit Routing

Use the following guidelines when creating and maintaining transit routing connections:

Topic	Caution or Guideline
OSPF/EIGRP Redistribution into ACI Fabric iBGP when Transit Routing across Multiple VRFs - Route Tags	In a transit routing scenario where external routers are used to route between multiple VRFs, and when an entry other than the default route tag (4294967295) is used to identify the policy in different VRFs, there is a risk of routing loops when there's one or more routes withdrawn from a tenant L3Out in OSPF or EIGRP. This is expected behavior. Upon the EIGRP/OSPF redistribution of routes into the ACI fabric, the default iBGP anti-routing loop mechanisms on the border leaf switches either use the specific default route tag 4294967295 or they use the same tag that is assigned in the **Transit Route Tag Policy** field in the **VRF/Policy** page. If you configure a different, specific transit route tag for each VRF, the default anti-routing loop mechanism does not work. In order to avoid this situation, use the same value for the **Transit Route Tag Policy** field across all VRFs. For additional details regarding route-maps and tags usage, see the row for "OSPF or EIGRP in Back to Back Configuration" and other information on route control profile policies in this table. **Note** The route tag policy is configured in the **Create Route Tag Policy** page, which is accessed through the **Transit Route Tag Policy** field in the **VRF/Policy** page: **Tenants** > ***tenant_name*** > **Networking** > **VRFs** > ***VRF_name***

Topic	Caution or Guideline
Transit Routing with a Single L3Out Profile	Before Cisco APIC release 2.3(1f), transit routing was not supported within a single L3Out profile. In Cisco APIC release 2.3(1f) and later, you can configure transit routing with a single L3Out profile, with the following limitations: • If the VRF instance is unenforced, you can use an external subnet (l3extSubnet) of 0.0.0.0/0 to allow traffic between the routers sharing the same Layer 3 EPG. • If the VRF instance is enforced, you cannot use an external default subnet (0.0.0.0/0) to match both source and destination prefixes for traffic within the same Layer 3 EPG. To match all traffic within the same Layer 3 EPG, Cisco APIC supports the following prefixes: • **IPv4** • 0.0.0.0/1—with external subnets for the external EPG • 128.0.0.0/1—with external subnets for the external EPG • 0.0.0.0/0—with import route control subnet, aggregate import • **IPv6** • 0::0/1—with external subnets for the external EPG • 8000::0/1—with external subnets for the external EPG • 0:0/0—with import route control subnet, aggregate import You do not need a contract for intra-Layer 3 EPG forwarding. • Alternatively, you can use a single default subnet (0.0.0.0/0) when combined with a contract that has at least one other EPG (application or external). You cannot use vzAny as a replacement for this EPG. However, you do not need to deploy the other EPG anywhere. As an example, use an application EPG provided contract and a Layer 3 EPG consumed contract (matching 0.0.0.0/0) or an application EPG consumed contract and a Layer 3 EPG provided contract (matching 0.0.0.0/0).
Shared Routes: Differences in Hardware Support	Routes shared between VRFs function correctly on generation 2 switches (Cisco Nexus N9K switches with "EX" or "FX" on the end of the switch model name, or later; for example, N9K-93108TC-EX). On generation 1 switches, however, there may be dropped packets with this configuration, because the physical ternary content-addressable memory (TCAM) tables that store routes do not have enough capacity to fully support route parsing.

Topic	Caution or Guideline
OSPF or EIGRP in Back to Back Configuration	Cisco APIC supports transit routing in export route control policies that are configured on the L3Out. These policies control which transit routes (prefixes) are redistributed into the routing protocols in the L3Out. When these transit routes are redistributed into OSPF or EIGRP, they are tagged 4294967295 to prevent routing loops. The Cisco ACI fabric does not accept routes matching this tag when learned on an OSPF or EIGRP L3Out. However, in the following cases, it is necessary to override this behavior: • When connecting two Cisco ACI fabrics using OSPF or EIGRP. • When connecting two different VRFs in the same Cisco ACI fabric using OSPF or EIGRP. Where an override is required, you must configure the VRF with a different tag policy at the following APIC GUI location: **Tenant** > ***Tenant_name*** > **Policies** > **Protocol** > **Route Tag**. Apply a different tag. In addition to creating the new route-tag policy, update the VRF to use this policy at the following APIC GUI location: **Tenant** > ***Tenant_name*** > **Networking** > **VRFs** > ***Tenant_VRF*** . Apply the route tag policy that you created to the VRF. **Note** When multiple L3Outs or multiple interfaces in the same L3Out are deployed on the same leaf switch and used for transit routing, the routes are advertised within the IGP (not redistributed into the IGP). In this case the route-tag policy does not apply.
Advertising BD Subnets Outside the Fabric	The import and export route control policies only apply to the transit routes (the routes that are learned from other external peers) and the static routes. The subnets internal to the fabric that are configured on the tenant BD subnets are not advertised out using the export policy subnets. The tenant subnets are still permitted using the IP prefix-lists and the route-maps but they are implemented using different configuration steps. See the following configuration steps to advertise the tenant subnets outside the fabric: **1.** Configure the tenant subnet scope as **Public Subnet** in the subnet properties window. **2.** Optional. Set the Subnet Control as **ND RA Prefix** in the subnet properties window. **3.** Associate the tenant bridge domain (BD) with the external Layer 3 Outside (L3Out). **4.** Create contract (provider or consumer) association between the tenant EPG and the external EPG. Setting the BD subnet to Public scope and associating the BD to the L3Out creates an IP prefix-list and the route-map sequence entry on the border leaf for the BD subnet prefix.

Topic	Caution or Guideline
Advertising a Default Route	For external connections to the fabric that only require a default route, there is support for originating a default route for OSPF, EIGRP, and BGP L3Out connections. If a default route is received from an external peer, this route can be redistributed out to another peer following the transit export route control as described earlier in this article. A default route can also be advertised out using a Default Route Leak policy. This policy supports advertising a default route if it is present in the routing table or it always supports advertising a default route. The Default Route Leak policy is configured in the L3Out connection. When creating a Default Route Leak policy, follow these guidelines: • For BGP, the **Always** property is not applicable. • For BGP, when configuring the **Scope** property, choose **Outside**. • For OSPF, the scope value **Context** creates a type-5 LSA while the Scope value **Outside** creates type-7 LSA. Your choice depends on the area type configured in the L3Out. If the area type is **Regular**, set the scope to **Context**. If the area type is **NSSA**, set the scope to **Outside**. • For EIGRP, when choosing the **Scope** property, you must choose **Context**.
MTU	Cisco ACI does not support IP fragmentation. Therefore, when you configure Layer 3 Outside (L3Out) connections to external routers, or multipod connections through an Inter-Pod Network (IPN), it is critical that the MTU is set appropriately on both sides. On some platforms, such as ACI, Cisco NX-OS, and Cisco IOS, the configurable MTU value takes into account the IP headers (resulting in a max packet size to be set as 9216 bytes for ACI and 9000 for NX-OS and IOS). However, other platforms such as IOS-XR configure the MTU value exclusive of packet headers (resulting in a max packet size of 8986 bytes). For the appropriate MTU values for each platform, see the relevant configuration guides. Cisco highly recommends you test the MTU using CLI-based commands. For example, on the Cisco NX-OS CLI, use a command such as `ping 1.1.1.1 df-bit packet-size 9000 source-interface ethernet 1/1`.

Transit Route Control

A route transit is defined to import traffic through a Layer 3 outside network `L3extOut` profile (`l3extInstP`), where it is to be imported. A different route transit is defined to export traffic through another `l3extInstP` where it is to be exported.

Since multiple `l3extOut` policies can be deployed on a single node or multiple nodes in the fabric, a variety of protocol combinations are supported. Every protocol combination can be deployed on a single node using multiple `l3extOut` policies or multiple nodes using multiple `l3extOut` policies. Deployments of more than two protocols in different `l3extOut` policies in the fabric are supported.

Export route-maps are made up of prefix-list matches. Each prefix-list consists of bridge domain (BD) public subnet prefixes in the VRF and the export prefixes that need to be advertised outside.

Route control policies are defined in an `l3extOut` policy and controlled by properties and relations associated with the `l3extOut`. APIC uses the `enforceRtctrl` property of the `l3extOut` to enforce route control directions. The default is to enforce control on export and allow all on import. Imported and exported routes (`l3extSubnets`), are defined in the `l3extInstP`. The default scope for every route is import. These are the routes and prefixes which form a prefix-based EPG.

All the import routes form the import route map and are used by BGP and OSPF to control import. All the export routes form the export route map used by OSPF and BGP to control export.

Import and export route control policies are defined at different levels. All IPv4 policy levels are supported for IPv6. Extra relations that are defined in the `l3extInstP` and `l3extSubnet` MOs control import.

Default route leak is enabled by defining the `l3extDefaultRouteLeakP` MO under the `l3extOut`.

`l3extDefaultRouteLeakP` can have Virtual Routing and Forwarding (VRF) scope or `L3extOut` scope per area for OSPF and per peer for BGP.

The following set rules provide route control:

- `rtctrlSetPref`
- `rtctrlSetRtMetric`
- `rtctrlSetRtMetricType`

Additional syntax for the `rtctrlSetComm` MO includes the following:

- `no-advertise`
- `no-export`
- `no-peer`

BGP

The ACI fabric supports BGP peering with external routers. BGP peers are associated with an `l3extOut` policy and multiple BGP peers can be configured per `l3extOut`. BGP can be enabled at the `l3extOut` level by defining the `bgpExtP` MO under an `l3extOut`.

Note Although the `l3extOut` policy contains the routing protocol (for example, BGP with its related VRF), the L3Out interface profile contains the necessary BGP interface configuration details. Both are needed to enable BGP.

BGP peer reachability can be through OSPF, EIGRP, a connected interface, static routes, or a loopback. iBGP or eBGP can be used for peering with external routers. The BGP route attributes from the external router are preserved since MP-BGP is used for distributing the external routes in the fabric. BGP enables IPv4 and/or IPv6 address families for the VRF associated with an `l3extOut`. The address family to enable on a switch is determined by the IP address type defined in `bgpPeerP` policies for the `l3extOut`. The policy is optional; if not defined, the default will be used. Policies can be defined for a tenant and used by a VRF that is referenced by name.

You must define at least one peer policy to enable the protocol on each border leaf (BL) switch. A peer policy can be defined in two places:

- Under `l3extRsPathL3OutAtt`—a physical interface is used as the source interface.
- Under `l3extLNodeP`—a loopback interface is used as the source interface.

OSPF

Various host types require OSPF to enable connectivity and provide redundancy. These include mainframe devices, external pods and service nodes that use the ACI fabric as a Layer 3 transit within the fabric and to the WAN. Such external devices peer with the fabric through a nonborder leaf switch running OSPF. Configure the OSPF area as an NSSA (stub) area to enable it to receive a default route and not participate in full-area routing. Typically, existing routing deployments avoid configuration changes, so a stub area configuration is not mandated.

You enable OSPF by configuring an `ospfExtP` managed object under an `l3extOut`. OSPF IP address family versions configured on the BL switch are determined by the address family that is configured in the OSPF interface IP address.

Note Although the `l3extOut` policy contains the routing protocol (for example, OSPF with its related VRF and area ID), the Layer 3 external interface profile contains the necessary OSPF interface details. Both are needed to enable OSPF.

You configure OSPF policies at the VRF level by using the `fvRsCtxToOspfCtxPol` relation, which you can configure per address family. If you do not configured it, default parameters are used.

You configure the OSPF area in the `ospfExtP` managed object, which also exposes IPv6 the required area properties.

Scope and Aggregate Controls for Subnets

The following section describes some scope and aggregate options available when creating a subnet:

Export Route Control Subnet—The control advertises specific transit routes out of the fabric. This is for transit routes only, and it does not control the internal routes or default gateways that are configured on a bridge domain (BD).

Import Route Control Subnet—This control allows routes to be advertised into the fabric with Border Gateway Protocol (BGP) and Open Shortest Path First (OSPF) when Import Route Control Enforcement is configured.

External Subnets for the External EPG (also called Security Import Subnet)—This option does not control the movement of routing information into or out of the fabric. If you want traffic to flow from one external EPG to another external EPG or to an internal EPG, the subnet must be marked with this control. If you do not mark the subnet with this control, then routes learned from one EPG are advertised to the other external EPG, but packets are dropped in the fabric. The drops occur because the APIC operates in a allowed list model where the default behavior is to drop all data plane traffic between EPGs, unless it is explicitly permitted by a contract. The allowed list model applies to external EPGs and application EPGs. When using security policies that have this option configured, you must configure a contract and a security prefix.

Shared Route Control Subnet—Subnets that are learned from shared L3Outs in inter-VRF leaking must be marked with this control before being advertised to other VRFs. Starting with APIC release 2.2(2e), shared L3Outs in different VRFs can communicate with each other using a contract. For more about communication between shared L3Outs in different VRFs, see the *Cisco APIC Layer 3 Networking Configuration Guide*.

Shared Security Import Subnet—This control is the same as External Subnets for the External EPG for Shared L3Out learned routes. If you want traffic to flow from one external EPG to another external EPG or to another internal EPG, the subnet must be marked with this control. If you do not mark the subnet with this control, then routes learned from one EPG are advertised to the other external EPG, but packets are dropped in the fabric.When using security policies that have this option configured, you must configure a contract and a security prefix.

Aggregate Export, Aggregate Import, and Aggregate Shared Routes—This option adds 32 in front of the 0.0.0.0/0 prefix. Currently, you can only aggregate the 0.0.0.0/0 prefix for the import/export route control subnet. If the 0.0.0.0/0 prefix is aggregated, no route control profile can be applied to the 0.0.0.0/0 network.

Aggregate Shared Route—This option is available for any prefix that is marked as Shared Route Control Subnet.

Route Control Profile—The ACI fabric also supports the route-map set clauses for the routes that are advertised into and out of the fabric. The route-map set rules are configured with the Route Control Profile policies and the Action Rule Profiles.

Property	OSPF	EIGRP	BGP	Comments
Set Community	X	X	Yes	Supports regular and extended communities.
Route Tag	Yes	Yes	X	Supported only for BD subnets. Transit prefixes are always assigned the tag 4294967295.
Preference	X	X	Yes	Sets BGP local preference.
Metric	Yes	X	Yes	Sets MED for BGP and changes the metric for EIGRP, but you cannot specify the EIGRP composite metric.
Metric Type	Yes	X	X	OSPF Type-1 and OSPF Type-2.

Route Control Profile Policies

The ACI fabric also supports the route-map set clauses for the routes that are advertised into and out of the fabric. The route-map set rules are configured with the Route Control Profile policies and the Action Rule Profiles.

ACI supports the following set options:

Table 1: Action Rule Profile Properties (route-map set clauses)

Property	OSPF	EIGRP	BGP	Comments
Set Community			Yes	Supports regular and extended communities.
Set Additional Community			Yes	Supports regular and extended communities.
Route Tag	Yes	Yes		Supported only for BD subnets. Transit prefixes are always assigned the tag 4294967295.
Preference			Yes	Sets BGP local preference.
Metric	Yes		Yes	Sets MED for BGP. Will change the metric for EIGRP but you cannot specify the EIGRP composite metric.
Metric Type	Yes			OSPF Type-1 and OSPF Type-2.

The Route Profile Polices are created under the Layer 3 Outside connection. A Route Control Policy can be referenced by the following objects:

- Tenant BD Subnet
- Tenant BD
- External EPG
- External EPG import/export subnet

Here is an example of using Import Route Control for BGP and setting the local preference for an external route learned from two different Layer 3 Outsides. The Layer 3 Outside connection for the external connection to AS300 is configured with the Import Route Control enforcement. An action rule profile is configured to set the local preference to 200 in the Action Rule Profile for Local Preference window.

The Layer 3 Outside connection External EPG is configured with a 0.0.0.0/0 import aggregate policy to allow all the routes. This is necessary because the import route control is enforced but any prefixes should not be blocked. The import route control is enforced to allow setting the local preference. Another import subnet 151.0.1.0/24 is added with a Route Profile that references the Action Rule Profile in the External EPG settings for Route Control Profile window.

Use the **show ip bgp vrf overlay-1** command to display the MP-BGP table. The MP-BGP table on the spine displays the prefix 151.0.1.0/24 with local preference 200 and a next hop of the border leaf for the BGP 300 Layer 3 Outside connection.

There are two special route control profiles—default-import and default-export. If the user configures using the names default-import and default-export, then the route control profile is automatically applied at the Layer3 outside level for both import and export. The default-import and default-export route control profiles cannot be configured using the 0.0.0.0/0 aggregate.

A route control profile is applied in the following sequential order for fabric routes:

1. Tenant BD subnet
2. Tenant BD
3. Layer3 outside

The route control profile is applied in the following sequential order for transit routes:

1. External EPG prefix
2. External EPG
3. Layer3 outside

Security Import Policies

The policies discussed in the documentation have dealt with the exchange of the routing information into and out of the ACI fabric and the methods that are used to control and tag the routes. The fabric operates in a allowed list model in which the default behavior is to drop all dataplane traffic between the endpoint groups unless it is explicitly permitted by a contract. This allowed list model applies to the external EPGs and the tenant EPGs.

There are some differences in how the security policies are configured and how they are implemented for the transit traffic compared to the tenant traffic.

Transit Security Policies

- Uses prefix filtering.
- Starting with Release 2.0(1m), support for Ethertype, protocol, L4 port, and TCP flag filters is available.
- Implemented with the security import subnets (prefixes) and the contracts that are configured under the external EPG.

Tenant EPG Security Policies

- Do not use prefix filtering.
- Support Ethertype, protocol, L4 port, and TCP flag filters.
- Supported for tenant EPGs ⟷ EPGs and tenant EPGs ⟷ External EPGs.

If there are no contracts between the external prefix-based EPGs, the traffic is dropped. To allow traffic between two external EPGs, you must configure a contract and a security prefix. As only prefix filtering is supported, the default filter can be used in the contract.

External L3Out Connection Contracts

The union of prefixes for L3Out connections is programmed on all the leaf nodes where the L3Out connections are deployed. When more than two L3Out connections are deployed, the use of the aggregate rule 0.0.0.0/0 can allow traffic to flow between L3Out connections that do not have a contract.

You configure the provider and consumer contract associations and the security import subnets in the L3Out Instance Profile (instP).

When security import subnets are configured and the aggragate rule, 0.0.0.0/0, is supported, the security import subnets follow the ACL type rules. The security import subnet rule 10.0.0.0/8 matches all the addresses from 10.0.0.0 to 10.255.255.255. It is not required to configure an exact prefix match for the prefixes to be permitted by the route control subnets.

Be careful when configuring the security import subnets if more than two L3Out connections are configured in the same VRF, due to the union of the rules.

Transit traffic flowing into and out of the same L3Out is dropped by policies when configured with the 0.0.0.0/0 security import subnet. This behavior is true for dynamic or static routing. To prevent this behavior, define more specific subnets.

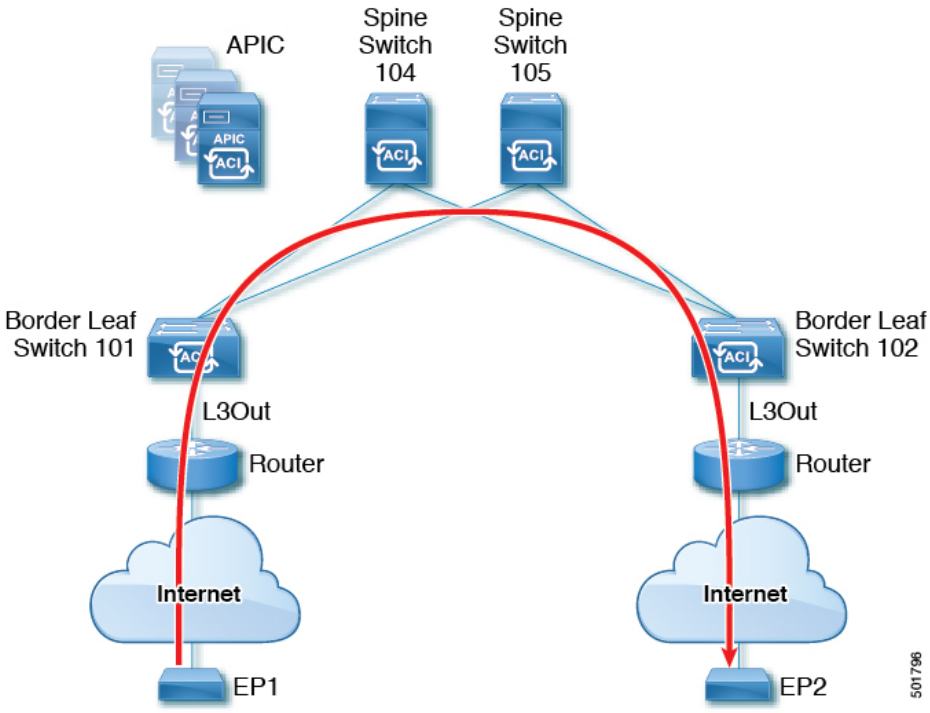
Configuring Transit Routing

Transit Routing Overview

This topic provides a typical example of how to configure Transit Routing when using Cisco APIC.

The examples in this chapter use the following topology:

Figure 7:



In the examples in this chapter, the Cisco ACI fabric has 2 leaf switches and two spine switches, that are controlled by an APIC cluster. The border leaf switches 101 and 102 have L3Outs on them providing connections to two routers and thus to the Internet. The goal of this example is to enable traffic to flow from EP 1 to EP 2 on the Internet into and out of the fabric through the two L3Outs.

In this example, the tenant that is associated with both L3Outs is `t1`, with VRF `v1`.

Before configuring the L3Outs, configure the nodes, ports, functional profiles, AEPs, and a Layer 3 domain. You must also configure the spine switches 104 and 105 as BGP route reflectors.

Configuring transit routing includes defining the following components:

1. Tenant and VRF
2. Node and interface on leaf 101 and leaf 102
3. Primary routing protocol on each L3Out (used to exchange routes between border leaf switch and external routers; in this example, BGP)
4. Connectivity routing protocol on each L3Out (provides reachability information for the primary protocol; in this example, OSPF)
5. Two external EPGs
6. One route map
7. At least one filter and one contract
8. Associate the contract with the external EPGs

Note For transit routing cautions and guidelines, see Guidelines for Transit Routing, on page 9.

The following table lists the names that are used in the examples in this chapter:

Property	Names for L3Out1 on Node 101	Names for L3Out2 on Node 102
Tenant	`t1`	`t1`
VRF	`v1`	`v1`
Node	`nodep1` with router ID 11.11.11.103	`nodep2` with router ID 22.22.22.203
OSPF Interface	`ifp1` at `eth/1/3`	`ifp2` at `eth/1/3`
BGP peer address	`15.15.15.2/24`	`25.25.25.2/24`
External EPG	`extnw1` at `192.168.1.0/24`	`extnw2` at `192.168.2.0/24`
Route map	`rp1` with `ctx1` and route destination `192.168.1.0/24`	`rp2` with `ctx2` and route destination `192.168.2.0/24`
Filter	`http-filter`	`http-filter`
Contract	`httpCtrct` provided by `extnw1`	`httpCtrct` consumed by `extnw2`

Configure Transit Routing Using the GUI

These steps describe how to configure transit routing for a tenant. This example deploys two L3Outs, in one VRF, on two border leaf switches, that are connected to separate routers.

Except for the step to create the tenant and VRF, perform these steps twice, to create the two L3Outs under the same tenant and VRF.

For sample names, see Transit Routing in the ACI Fabric, on page 1.

Before you begin

- Configure an L3 Domain and Fabric Access Policies for interfaces that are used in the L3Out (AAEP, VLAN pool, Interface selectors).
- Configure a BGP Route Reflector Policy for the fabric infra MPBGP.

Procedure

Step 1 To create the tenant and VRF, on the menu bar, choose **Tenants** > **Add Tenant** and in the **Create Tenant** dialog box, perform the following tasks:

a) In the **Name** field, enter the tenant name.
b) In the **VRF Name** field, enter the VRF name.
c) Click **Submit**.

Note After this step, perform the steps twice to create two L3Outs in the same tenant and VRF for transit routing.

Step 2 To start creating the L3Out, on the **Navigation** pane, expand **Tenant** and **Networking**, then right-click **L3Outs** and choose **Create L3Out**.

The **Create L3Out** wizard appears. The following steps provide the steps for an example L3Out configuration using the **Create L3Out** wizard.

Step 3 Enter the necessary information in the **Identity** window of the **Create L3Out** wizard.

a) In the **Name** field, enter a name for the L3Out.
b) From the **VRF** drop-down list, choose the VRF.
c) From the **L3 Domain** drop-down list, choose the external routed domain that you previously created.
d) In the area with the routing protocol check boxes, check the desired protocols (BGP, OSPF, or EIGRP).

For the example in this chapter, choose **BGP** and **OSPF**.

Depending on the protocols you choose, enter the properties that must be set.

e) Enter the OSPF details, if you enabled OSPF.

For the example in this chapter, use the OSPF area **0** and type **Regular area**.

f) Click **Next** to move to the **Nodes and Interfaces** window.

Step 4 Enter the necessary information in the **Nodes and Interfaces** window of the **Create L3Out** wizard.

a) Determine if you want to use the default naming convention.

In the **Use Defaults** field, check if you want to use the default node profile name and interface profile names:

- The default node profile name is *L3Out-name*_nodeProfile, where *L3Out-name* is the name that you entered in the **Name** field in the **Identity** page.
- The default interface profile name is *L3Out-name*_interfaceProfile, where *L3Out-name* is the name that you entered in the **Name** field in the **Identity** page.

b) In the **Interface Types** area, make the necessary selections in the Layer 3 and Layer 2 fields.

The options are:

- Layer 3:
  - **Routed**: Select this option to configure a Layer 3 route to the port channels.

    When selecting this option, the Layer 3 route can be to either physical ports or direct port channels, which are selected in the **Layer 2** field in this page.
  - **Routed Sub**: Select this option to configure a Layer 3 sub-interface route to the port channels.

    When selecting this option, the Layer 3 sub-interface route can be to either physical ports or direct port channels, which are selected in the **Layer 2** field in this page.
  - **SVI**: Select this option to configure a Switch Virtual Interface (SVI), which is used to provide connectivity between the ACI leaf switch and a router.

    SVI can have members that are physical ports, direct port channels, or virtual port channels, which are selected in the **Layer 2** field in this page.
  - **Floating SVI**: Select this option to configure floating L3Out.

    Floating L3Out enables you to configure an L3Out that allows a virtual router to move from under one leaf switch to another. The feature saves you from having to configure multiple L3Out interfaces to maintain routing when VMs move from one host to another.
- Layer 2: (not available if you select Virtual SVI in the Layer 3 area)
  - Port
  - Virtual Port Channel (available if you select **SVI** in the Layer 3 area)
  - Direct Port Channel

c) From the **Node ID** field drop-down menu, choose the node for the L3Out.

For the topology in these examples, use node 103.

d) In the **Router ID** field, enter the router ID (IPv4 or IPv6 address for the router that is connected to the L3Out).

e) (Optional) You can configure another IP address for a loopback address, if necessary.

The **Loopback Address** field is automatically populated with the same entry that you provide in the **Router ID** field. This is the equivalent of the **Use Router ID for Loopback Address** option in previous builds. Enter a different IP address for a loopback address, if you don't want to use route ID for the loopback address, or leave this field empty if you do not want to use the router ID for the loopback address.

f) Enter necessary additional information in the **Nodes and Interfaces** window.

The fields shown in this window varies, depending on the options that you select in the **Layer 3** and **Layer 2** areas.

g) When you have entered the remaining additional information in the **Nodes and Interfaces** window, click **Next**.

The **Protocols** window appears.

Step 5 Enter the necessary information in the **Protocols** window of the **Create L3Out** wizard.

Because you BGP and OSPF as the protocols for this example, the following steps provide information for those fields.

a) In the **BGP Loopback Policies** and **BGP Interface Policies** areas, enter the following information:

- **Peer Address**: Enter the peer IP address
- **EBGP Multihop TTL**: Enter the connection time to live (TTL). The range is from 1 to 255 hops; if zero, no TTL is specified. The default is zero.
- **Remote ASN**: Enter a number that uniquely identifies the neighbor autonomous system. The Autonomous System Number can be in 4-byte as plain format from 1 to 4294967295.

Note ACI does not support asdot or asdot+ format AS numbers.

b) In the **OSPF** area, choose the default OSPF policy, a previously created OSPF policy, or **Create OSPF Interface Policy**.

c) Click **Next**.

The **External EPG** window appears.

Step 6 Enter the necessary information in the **External EPG** window of the **Create L3Out** wizard.

a) In the **Name** field, enter a name for the external network.

b) In the **Provided Contract** field, enter the name of a provided contract.

c) In the **Consumed Contract** field, enter the name of a consumed contract.

d) In the **Default EPG for all external networks** field, uncheck if you don't want to advertise all the transit routes out of this L3Out connection.

The Subnets area appears if you uncheck this box. Specify the desired subnets and controls as described in the following steps.

e) Click the **+** icon to expand **Subnet**, then perform the following actions in the **Create Subnet** dialog box.

f) In the **IP address** field, enter the IP address and network mask for the external network.

g) In the **Name** field, enter the name of the subnet.

h) In the **Scope** field, check the appropriate check boxes to control the import and export of prefixes for the L3Out.

Note For more information about the scope options, see the online help for this **Create Subnet** panel.

i) (Optional) Click the check box for **Export Route Control Subnet**.

The **BGP Route Summarization Policy** field now becomes available.

j) In the **BGP Route Summarization Policy** field, from the drop-down list, choose an existing route summarization policy or create a new one as desired.

The type of route summarization policy depends on the routing protocols that are enabled for the L3Out.

k) Click **OK** when you have completed the necessary configurations in the **Create Subnet** window.
l) (Optional) Repeat to add more subnets.
m) Click **Finish** to complete the necessary configurations in the **Create L3Out** wizard.

Step 7 Navigate to the L3Out that you just created, then right-click on the L3Out and select **Create Route map for import and export route control**.

Step 8 In the **Create Route map for import and export route control** window, perform the following actions:

a) In the **Name** field, enter the route map name.
b) Choose the **Type**.

For this example, leave the default, **Match Prefix AND Routing Policy**.

c) Click the **+** icon to expand **Contexts** and create a route context for the route map.
d) Enter the order and name of the profile context.
e) Choose **Deny** or **Permit** for the action to be performed in this context.
f) (Optional) In the **Set Rule** field, choose **Create Set Rules for a Route Map**.

Enter the name for the set rules, click the objects to be used in the rules, and click **Finish**.

g) In the **Match Rule** field, choose **Create Match Rule for a Route Map**.
h) Enter the name for the match rule and enter the **Match Regex Community Terms**, **Match Community Terms**, or **Match Prefix** to match in the rule.
i) When you have finished filling in the fields in the **Create Match Rule** window, click **Submit**.
j) In the **Create Route Control Context** dialog box, click **OK**.
k) In the **Create Route map for import and export route control** dialog box, click **Submit**.

Step 9 In the Navigation pane, expand **L3Outs** > ***L3Out_name*** > **External EPGs** > ***externalEPG_name*** , and perform the following actions:

a) Click the **+** icon to expand **Route Control Profile**.
b) In the **Name** field, choose the route control profile that you previously created from the drop-down list.
c) In the **Direction** field, choose **Route Export Policy**.
d) Click **Update**.

Step 10 Navigate to the L3Out that you just created, then right-click on the L3Out and select **Create Route map for import and export route control**.

Step 11 In the **Create Route map for import and export route control** window, perform the following actions.

Note To set attributes for BGP, OSPF, or EIGRP for received routes, create a default-import route control profile, with the appropriate set actions and no match actions.

a) In the **Name** field, choose **default-import**.
b) In the **Type** field, you must select **Match Routing Policy Only**
c) In the **Create Route map for import and export route control** dialog box, click **Submit**.

Step 12 To enable communications between the EPGs consuming the L3Out, create at least one filter and contract, using the following steps:

a) In the Navigation pane, under the tenant consuming the L3Out, expand **Contracts**.
b) Right-click **Filters** and choose **Create Filter**.
c) In the **Name** field, enter a filter name.

A filter is essentially an Access Control List (ACL).

d) Click the + icon to expand **Entries**, and add a filter entry.
e) Add the Entry details.

For example, for a simple web filter, set criteria such as the following:

- **EtherType—IP**
- **IP Protocol—tcp**
- **Destination Port Range From—Unspecified**
- **Destination Port Range To** to **https**

f) Click **Update**.
g) In the **Create Filter** dialog box, click **Submit**.

Step 13 To add a contract, use the following steps:

a) Under **Contracts**, right-click **Standard** and choose **Create Contract**.
b) Enter the name of the contract.
c) Click the + icon to expand **Subjects** to add a subject to the contract.
d) Enter a name for the subject.
e) Click the + icon to expand **Filters** and choose the filter that you previously created from the drop-down list.
f) Click **Update**.
g) In the **Create Contract Subject** dialog box, click **OK**.
h) In the **Create Contract** dialog box, click **Submit**.

Step 14 Associate the EPGs for the L3Out with the contract, with the following steps:

The first L3 external EPG (`extnw1`) is the provider of the contract and the second L3 external EPG (`extnw2`) is the consumer.

a) To associate the contract to the L3 external EPG, as the provider, under the tenant, click **Networking**, expand **L3Outs**, and expand the L3Out.
b) Expand **External EPGs**, click the L3 external EPG, and click the **Contracts** tab.
c) Click the the + icon to expand **Provided Contracts**.

For the second L3 external EPG, click the + icon to expand Consumed Contracts.

d) In the **Name** field, choose the contract that you previously created from the list.
e) Click **Update**.
f) Click **Submit**.