Configuring Fallback Route Groups

Revised: June 21, 2022

New and Changed Information

The following table provides an overview of the significant changes up to this current release. The table does not provide an exhaustive list of all changes or of the new features up to this release.

Table 1: New Features and Changed Information for Floating L3Out

Cisco APIC Release Version	Feature	Description
5.2(4)	First release of this document	--

Overview

This feature provides fast convergence for a destination that is reachable using a primary route and fallback route. You can group multiple next-hops of a route into one fallback route group so that if next-hop of the primary route fails, a Cisco ACI leaf switch can replace the failed primary next-hop with all the next-hops of the group in the hardware table before the control plane convergence with the routing protocol happens. This hardware-based convergence can happen within a second compared to multiple seconds or minutes for convergence through the control plane. In addition, next-hop failures can be detected faster through BFD.

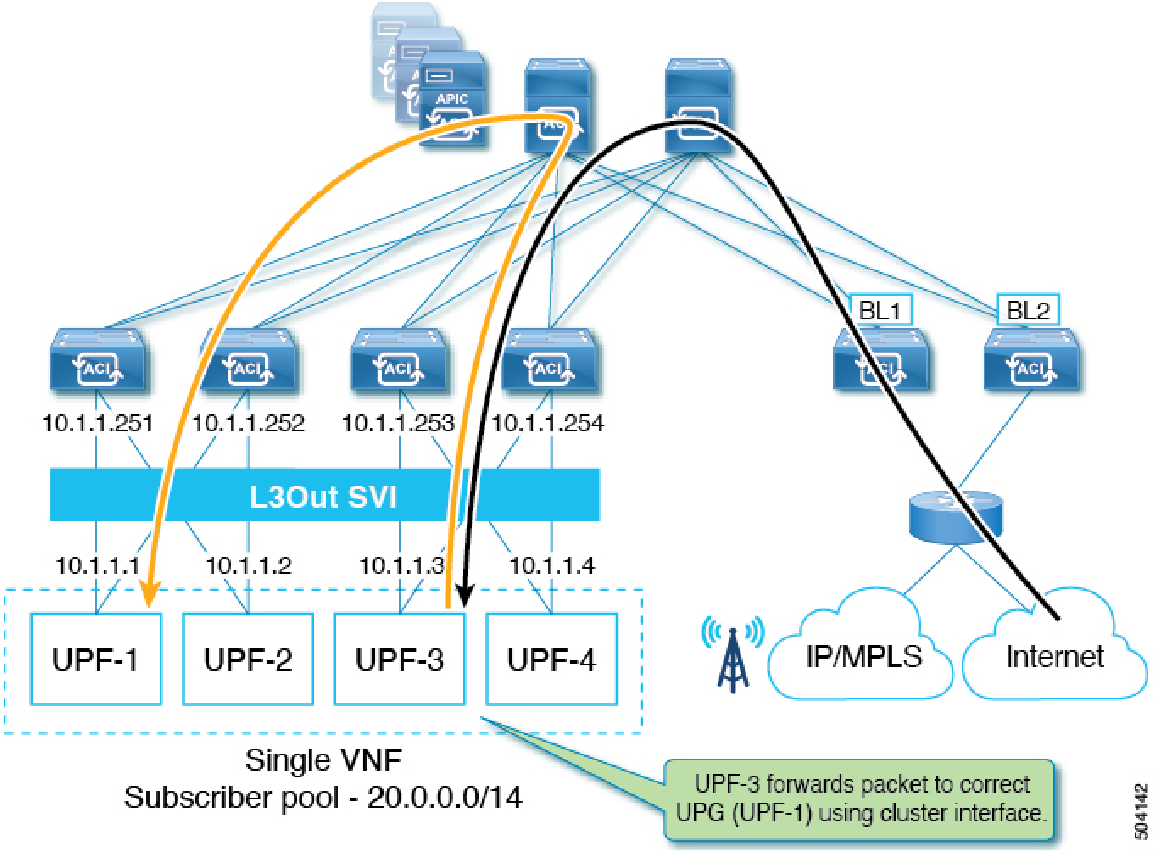
This feature is useful in the Telco packet core deployment and described in detail in the following sections.

Typical Deployment

Note This document assumes that you are adding fallback route configuration to an already deployed and configured fabric, so configuration for L3Outs, VRFs, and other fabric elements are outside the scope of this guide.

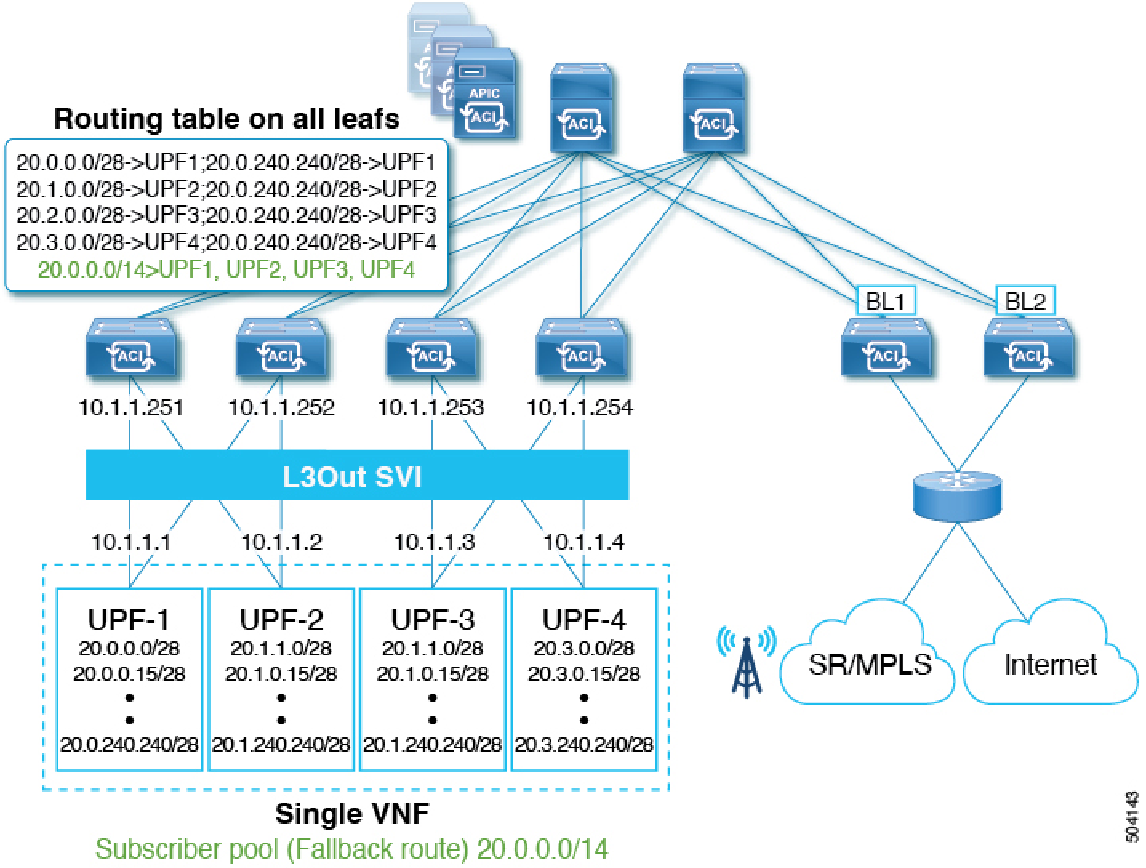
The following diagram illustrates the typical existing deployment, where `UPF-1` through `UPF-4` are all advertising the same subnet subscriber pool `20.0.0.0/14`. Since all of them are advertising the same subnet, the traffic can come to any UPF (for example, `UPF-3`) and then it would have to go through the fabric if that UPF is not the UPF which has the session information (in this case, `UPF-1`).

Figure 1: Typical Topology



In order to mitigate this additional traffic, each UPF is typically configured to advertise a `/28` subnet along with the fallback route (`20.0.0.0/14`), which is the subscriber pool. In other words, there's a common subnet and a smaller subnet to allow the border leaf (BL) to be able to determine the primary subnet and direct the traffic to the correct UPF, as shown in the following figure. Fallback route is used only when the primary route (`/28` subnet) is down.

Figure 2: Typical Topology



The downside of this approach is the high prefix scale and high convergence times due to the control plane having to remove primary route and install fallback route into hardware.

To mitigate these issues, you can update your existing deployment to configure fallback route groups at the VRF level and associate them with an L3Out. Each group will have a fallback route and all the associated next-hops of the UPFs. Each next-hop can be a directly connect interface IPs or a loopback IPs of the UPFs depending on whether eBGP is configured with directly connected interfaces or loopback interfaces.

Benefits of Using Fallback Route Groups

Beginning with Release 5.2(4), you can configure fallback route groups as an alternative to the typical deployment described above to eliminate unnecessary traffic forwarded between UPFs and reduce convergence times in case of a UPF failure.

In this case, multiple UPFs are combined into a fallback route **group**, so if any one UPF fails, the convergence will occur within the hardware itself significantly reducing the convergence times.

Fallback Routes Functionality, Failure Recovery, and Configuration Workflow

This section describes fallback route groups configuration workflow and failover mechanism.

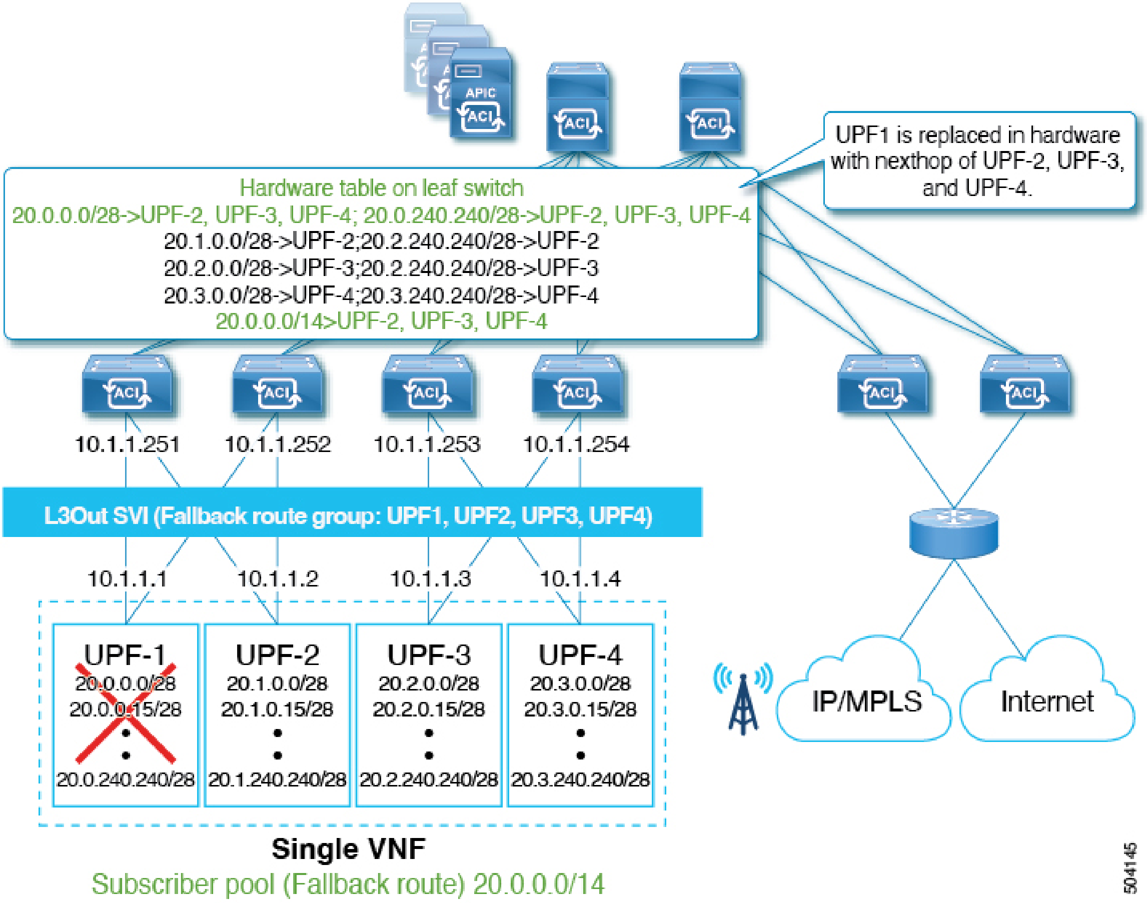
- Initial configuration and route deployment in Cisco ACI hardware:
  - You configure fallback route group at the VRF level and associate it with the L3Out in your APIC GUI.

- The group will have the fallback route and all the associated next-hops of the UPFs
- Each next-hop can be a directly connect interface IPs or a loopback IPs of the UPFs depending on whether eBGP is configured with directly connected interfaces or loopback interfaces
- Fallback route is pointing to all the UPFs in the group as it is advertised by all the UPFs.
- Primary routes are pointing to only one of the UPFs based on the distribution.

- UPF Failure Handling:
  - If a UPF failure is detected by BFD (for example, `UPF-1` in the below illustration), the fabric will automatically replace the failed next-hop with remaining next-hops of the UPFs from fallback route in hardware.
  - All hardware tables are updated with remaining UPF next-hops before control plane convergence by BGP.

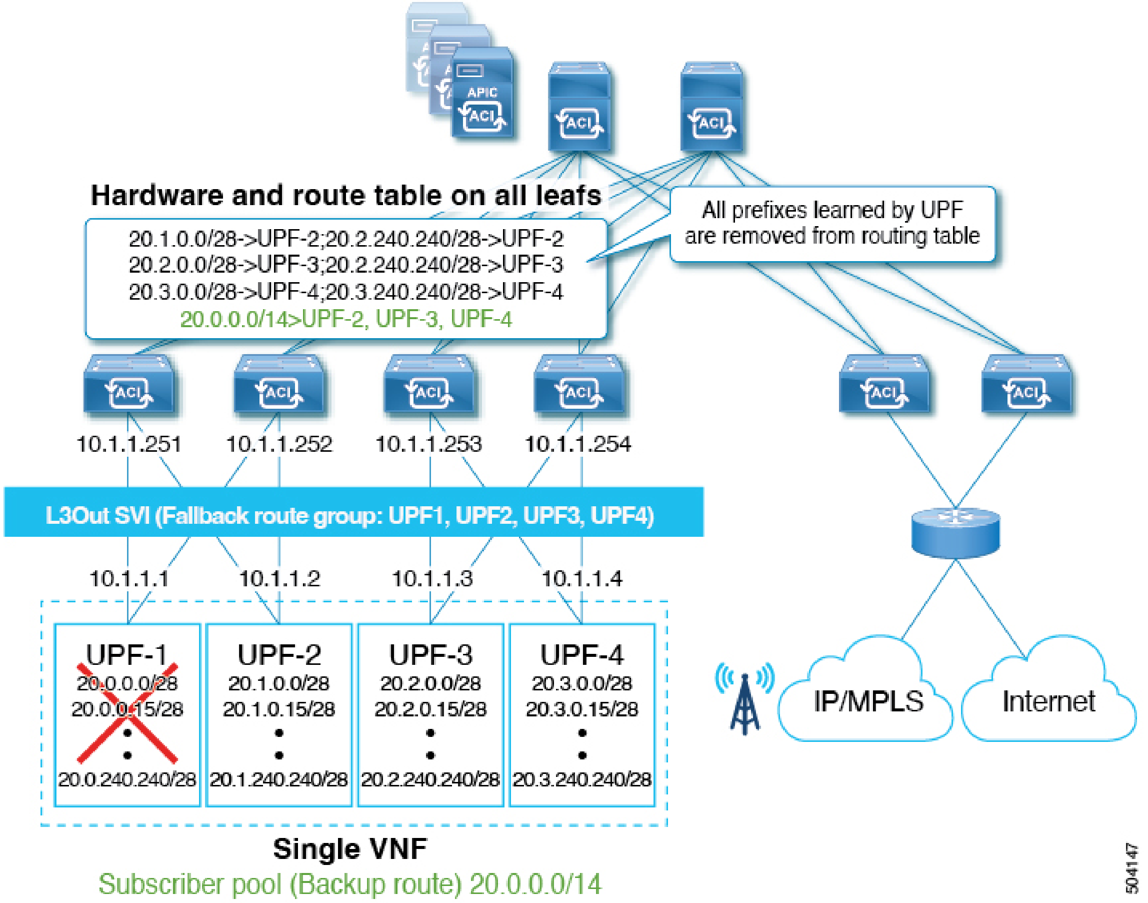
    At this point, route table still has all primary prefixes advertised by BGP.

Figure 3: Hardware-based Convergence


APIC
ACI
UPF1 is replaced in hardware with nexthop of UPF-2, UPF-3, and UPF-4.
Hardware table on leaf switch
20.0.0.0/28->UPF-2, UPF-3, UPF-4; 20.0.240.240/28->UPF-2, UPF-3, UPF-4
20.1.0.0/28->UPF-2;20.2.240.240/28->UPF-2
20.2.0.0/28->UPF-3;20.2.240.240/28->UPF-3
20.3.0.0/28->UPF-4;20.3.240.240/28->UPF-4
20.0.0.0/14>UPF-2, UPF-3, UPF-4
10.1.1.251
10.1.1.252
10.1.1.253
10.1.1.254
L3Out SVI (Fallback route group: UPF1, UPF2, UPF3, UPF4)
10.1.1.1
10.1.1.2
10.1.1.3
10.1.1.4
UPF-1
20.0.0.0/28
20.0.0.15/28
20.0.240.240/28
UPF-2
20.1.0.0/28
20.1.0.15/28
20.1.240.240/28
UPF-3
20.2.0.0/28
20.2.0.15/28
20.2.240.240/28
UPF-4
20.3.0.0/28
20.3.0.15/28
20.3.240.240/28
IP/MPLS
Internet
Single VNF
Subscriber pool (Fallback route) 20.0.0.0/14
504145

  - After Control Plane convergence, BGP removes all the primary prefixes advertised from `UPF-1`.

Figure 4: Control Plane Convergence


APIC
ACI
Hardware and route table on all leafs
20.1.0.0/28->UPF-2;20.2.240.240/28->UPF-2
20.2.0.0/28->UPF-3;20.2.240.240/28->UPF-3
20.3.0.0/28->UPF-4;20.3.240.240/28->UPF-4
20.0.0.0/14>UPF-2, UPF-3, UPF-4
All prefixes learned by UPF
are removed from routing table
10.1.1.251
10.1.1.252
10.1.1.253
10.1.1.254
L3Out SVI (Fallback route group: UPF1, UPF2, UPF3, UPF4)
10.1.1.1
10.1.1.2
10.1.1.3
10.1.1.4
UPF-1
20.0.0.0/28
20.0.0.15/28
20.0.240.240/28
UPF-2
20.1.0.0/28
20.1.0.15/28
20.1.240.240/28
UPF-3
20.2.0.0/28
20.2.0.15/28
20.2.240.240/28
UPF-4
20.3.0.0/28
20.3.0.15/28
20.3.240.240/28
IP/MPLS
Internet
Single VNF
Subscriber pool (Backup route) 20.0.0.0/14
504147

Guidelines and Limitations

When configuring fallback routes, the following guidelines apply:

- This feature is supported starting with Cisco APIC, Release 5.2(4).
- This feature is supported with the following switches:
  - `-FX` and later leaf switches
  - `-EX` and later spine switches as well as the `N9332C` and `N9364C` fixed spine switches.
- The same fallback route is used by all UPFs in the cluster.

  In other words, there is no traffic impact if the fabric forwards traffic to all remaining next-hops (UPFs) in the event of the UPF failure.
- Fallback routes are not supported over floating L3Out.
- Fallback routes are not supported for Multi-Site traffic.

- Fallback routes are supported on L3Out SVI only.
- Stretching SVIs is not supported across local leaf (LL) and remote leaf (RL).

  Note that this is an existing limitation and not a limitation specific to the fallback route group feature.
- If BGP route next-hop is reachable via OSPFv3 route, the OSPFv3 route next-hops must be global IPv6 addresses and not Link Local addresses.

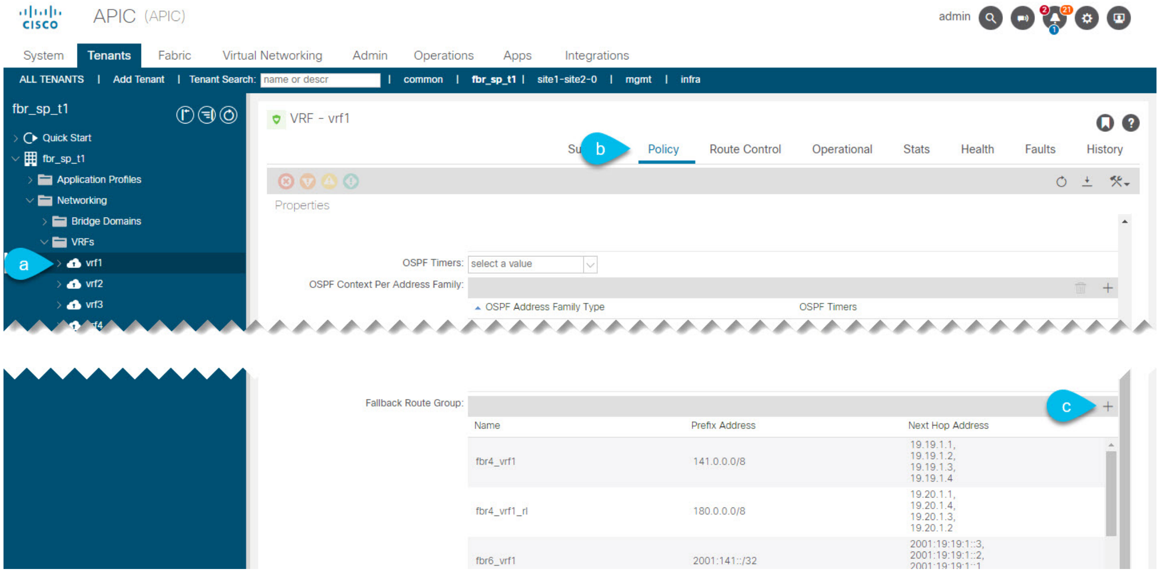
Creating Fallback Route Groups

This section explains how to create a fallback route group.

Procedure

Step 1 From the top navigation bar, choose **Tenants** > ***<tenant_name>***.

Step 2 Add a new fallback route group.



Name	Prefix Address	Next Hop Address
fbr4_vrf1	141.0.0.0/8	19.19.1.1, 19.19.1.2, 19.19.1.3, 19.19.1.4
fbr4_vrf1_rl	180.0.0.0/8	19.20.1.1, 19.20.1.4, 19.20.1.3, 19.20.1.2
fbr6_vrf1	2001:141::/32	2001:19:19:1::3, 2001:19:19:1::2, 2001:19:19:1::1

a) From the left navigation menu, select ***<tenant_name>*** > **Networking** > **VRFs** > ***<vrf_name>***.

b) In the main pane, select the **Policy** tab.

c) Scroll down to the **Fallback Route Group** area and click the plus (+) icon to add a new fallback route group.

   The **Create Fallback Route Group** window opens.

Step 3 In the **Create Fallback Route Group** window, provide the required details.

a) Provide the **Name** for the fallback route group.

b) In the **Fallback Routes** table, click the plus (+) icon to add a new fallback route, enter the route prefix address and click **Update** to save.

   For example, `17.0.0.0/8`. You can repeat this substep if you want to add multiple fallback routes to the same group.

c) In the **Fallback Members** table, click the plus (+) icon to add a next hop address, enter the route prefix address and click **Update** to save.

For example, `1.1.1.1`. You can repeat this substep if you want to add multiple next hop addresses to the same group.

d) Repeat this step to create any additional fallback route groups.

What to do next

After you have created the fallback route group, you must attach it to the L3Out as described in Attaching Fallback Route Policy to L3Out, on page 8

Attaching Fallback Route Policy to L3Out

This section describes how to attach a fallback route policy to an L3Out.

Note When you configure fallback route policy on an L3out, all nodes in the L3Out are configured with the same fallback route policy.

Before you begin

You must have the following:

- An L3Out already created and configured.

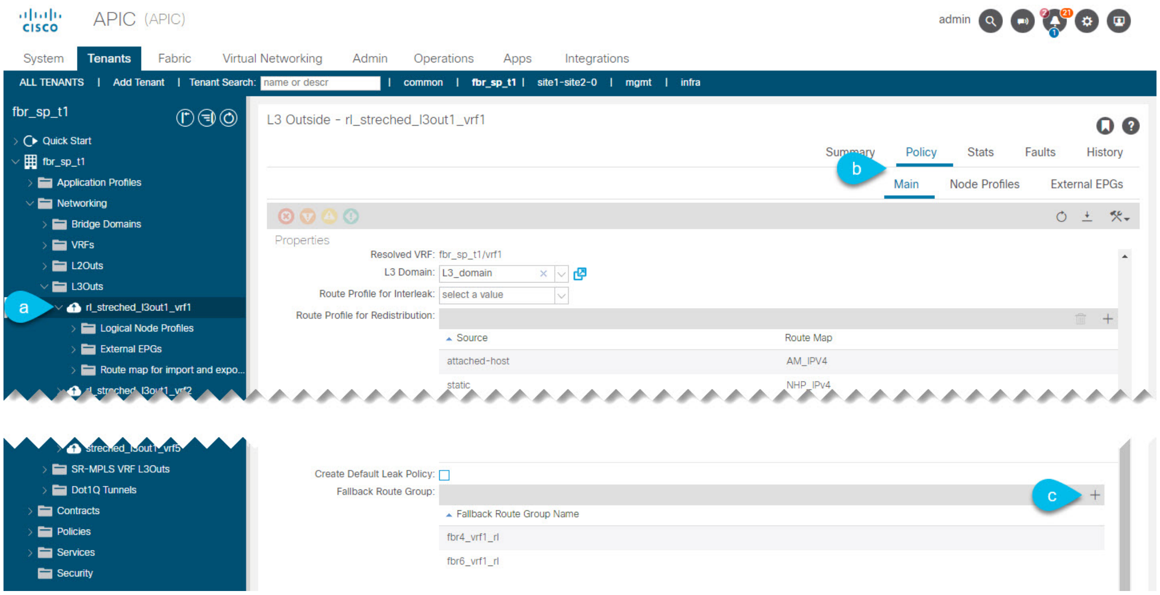
  The L3Out must be in a physical domain, not in a VMM domain.

- A fallback route policy created, as described in Creating Fallback Route Groups, on page 7.

Procedure

Step 1 From the top navigation bar, choose **Tenants** > ***<tenant_name>***.

Step 2 Associate the fallback route group with the L3Out.



L3 Outside - rl_streched_l3out1_vrf1

Resolved VRF: fbr_sp_t1/vrf1

L3 Domain: L3_domain

Route Profile for Interleak: select a value

Route Profile for Redistribution:

Source	Route Map
attached-host	AM_IPV4
static	NHP_IPV4

Create Default Leak Policy: □

Fallback Route Group:

Fallback Route Group Name
fbr4_vrf1_rl
fbr6_vrf1_rl

a) From the left navigation menu, select ***<tenant_name>*** > **Networking** > **L3Outs** > ***<l3out_name>***

b) In the main pane, select the **Policy** > **Main** tab.

c) Scroll down to the **Fallback Route Group** area and click the plus (+) icon.

From the dropdown, select the fallback route group. Then click **Update** to save the changes.

What to do next

After you have created and attached the fallback route policy to the L3Out, you must enable next hop propagation as described in Enabling Next Hop Propagation and Multipath, on page 9.

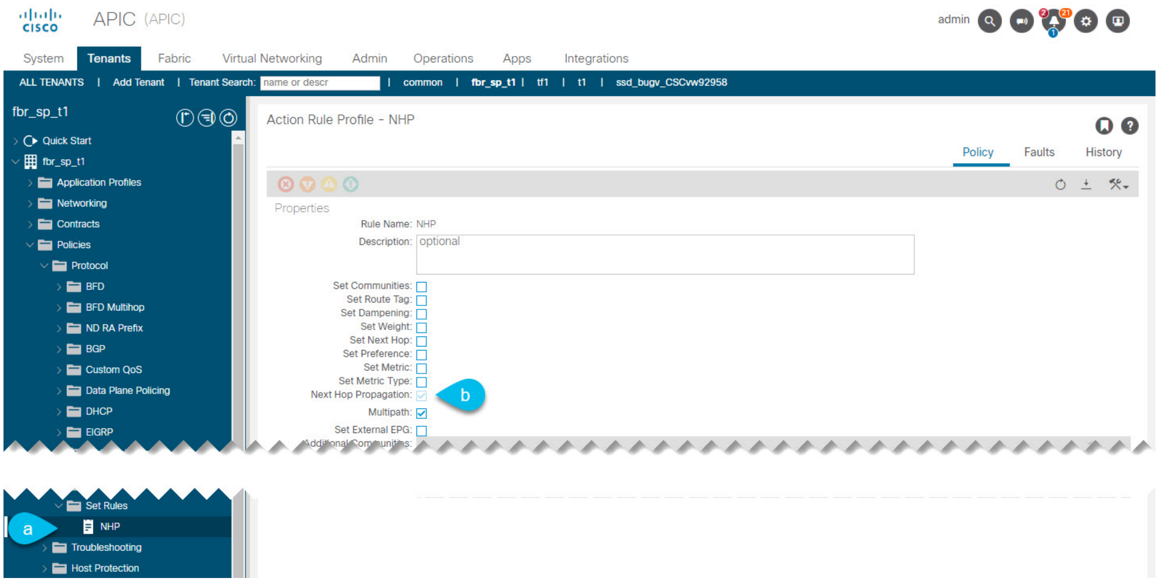
Enabling Next Hop Propagation and Multipath

This section explains to enable next-hop propagation which is required for fallback routes to work as well as multipath if multiple next-hops are present.

Procedure

Step 1 From the top navigation bar, choose **Tenants** > ***<tenant_name>***.

Step 2 Associate the fallback route group with the L3Out.



Rule Name: NHP

Description: optional

Set Communities: □

Set Route Tag: □

Set Dampening: □

Set Weight: □

Set Next Hop: □

Set Preference: □

Set Metric: □

Set Metric Type: □

Next Hop Propagation: ☑

Multipath: ☑

Set External EPG: □

Additional Communities:

a) From the left navigation menu, select ***<tenant_name>*** > **Policies** > **Protocol** > **Set Rules** > ***<rule_name>***.

b) In the main pane, enable the **Next Hop Propagation** and **Multipath** options.

- **Next Hop Propagation**: Select this option to propagate the next hop address advertised by the external BGP peer to infra MP-BGP VPN peers within the fabric.

  You must enable this for fallback route groups feature to work.

- **Multipath**: Select this option to specify if multiple paths (ECMP next hops) need to be picked for redistribution for a particular route when performing a next hop unchanged redistribution. The number of paths used is based on the value that you entered in the **Local Max ECMP** field. Enabling this option also automatically enables the **Next Hop Propagation** option.

What to do next

After you have enabled next-hop propagation, you must configure route redistribution as described in Configuring Attached-Host Route Redistribution, on page 10 or Configuring Static Route Redistribution, on page 11 depending on whether you want to redistribute attached-host or static routes..

Configuring Attached-Host Route Redistribution

In this procedure, you will specify the attached-host routes or subnets connected to the SVIs that need to be redistributed into the MP-BGP fabric.

If you plan to configure static route redistribution, skip this section and follow the steps described in Configuring Static Route Redistribution, on page 11 instead.

Procedure

Step 1 Create a match rule for the route map for route control for the attached-host redistribution policy.

a) From the navigation pane, go to **Tenants** > ***tenant_name*** > **Policies** > **Protocol**.

b) Right-click on **Match Rules** and choose **Create Match Rules for Route Map**.

The **Create Match Rule** dialog appears in the work pane.

c) Enter a name in the **Name** field for this match rule.

For the purposes of this example, we will enter **svi-prefix** as the name of this match rule.

d) Locate the **Match Prefix** summary table and click the **+** to access the **IP**, **Description**, **Aggregate**, **Greater Than Mask** and **Less Than Mask** fields and enter the appropriate values to configure a match rule that matches the L3Out SVI subnet.

e) When finished, click **Submit**.

Step 2 Create a route map for route control for the attached-host redistribution policy.

a) From the navigation pane, go to **Tenants** > ***tenant_name*** > **Policies** > **Protocol**.

b) Right-click on **Route Maps for Route Control** and choose **Create Route Maps for Route Control**.

The **Create Route Maps for Route Control** dialog appears in the work pane.

c) Enter a name for the route map in the **Name** field.

For the purposes of this example, we will enter **attach-pol** as the name of this route map.

d) From the **Contexts** summary table in the **Create Route Maps for Route Control** dialog, click the **+**.

The **Create Route Control Context** window appears.

e) Enter a name for the route control context in the **Name** field.

For the purposes of this example, we will enter **attach-pol-RCC** as the name of this route control context.

f) Click the **Associated Match Rules +** symbol to access the **Rule Name** field and choose the match rule that you created.

g) When finished, click **Submit**.

Step 3 Attach the policy to the L3Out.

a) From the left navigation menu, select ***<tenant_name>*** > **Networking** > **L3Outs** > ***<l3out_name>***

b) In the main pane, select the **Policy** > **Main** tab.

c) Scroll down to the **Route Profile for Redistribution** area and click the plus (+) icon.

d) In the **Source** area, choose **attached-host**.

e) In the **Route Map** area, choose the route map the previous steps.

f) Click **Submit**.

Configuring Static Route Redistribution

In this procedure, you will be configuring the route map for static routes that need to have the next hop propagated into the MP-BGP fabric.

If you plan to configure static route redistribution, skip this section and follow the steps described in Configuring Attached-Host Route Redistribution, on page 10 instead.

Procedure

Step 1 Configure the maximum number of paths for the redistribution of routes in the ACI fabric.

This is a necessary step before you can configure the **Multipath** field.

a) Navigate to **Tenants** > *tenant_name* > **Policies** > **Protocol** > **BGP** > **BGP Address Family Context**.

b) In the **Create BGP Address Family Context Policy** dialog box, perform the following tasks.

1. In the **Name** field, enter a name for the policy.

   For the purposes of this example, we will enter **redistr-mpath** as the name of this match rule.

2. Locate the **Local Max ECMP** field and enter the value to set the maximum number of paths (ECMP next hops) that should be selected when redistributing static routes learned on the border leaf switch into the MP-BGP fabric.

   The default value for this field is 0, which indicates that the **Local Max ECMP** setting is disabled. To enable this setting, enter a value for the maximum number of paths, where the range is from 1 to 16.

   Note Do not select the **Enable Host Route Leak** option in this scenario. This is not supported when configuring multi-protocol recursive next hop propagation.

3. Click **Submit** after you have updated your entries.

c) Navigate to **Tenants** > *tenant_name* > **Networking** > **VRFs** > *vrf_name*.

d) Review the configuration details of the subject VRF.

e) Locate the **BGP Context Per Address Family** field and, in the **BGP Address Family Type** area, select either **IPv4 unicast address family** or **IPv6 unicast address family**.

f) Access the BGP Address Family Context you created in the **BGP Address Family Context** drop-down list and associate it with the subject VRF.

g) Click **Submit**.

Step 2 Create a match rule for the route map for route control for static redistribution on the L3Out.

a) From the navigation pane, go to **Tenants** > *tenant_name* > **Policies** > **Protocol**.

b) Right-click on **Match Rules** and choose **Create Match Rules for Route Map**.

   The **Create Match Rule** dialog appears in the work pane.

c) Enter a name in the **Name** field for this match rule.

   For the purposes of this example, we will enter **OSPF-NH-static-NH-IPs** as the name of this match rule.

d) Locate the **Match Prefix** summary table and click the **+** to access the **IP**, **Description**, **Aggregate**, **Greater Than Mask** and **Less Than Mask** fields and enter the appropriate values to configure a match rule to match on the next hop/static next hop IP addresses (the vRouters loopback IP addresses).

e) When finished, click **Submit**.

Step 3 Create a route map for route control for static redistribution on the L3Out.

a) From the navigation pane, go to **Tenants** > *tenant_name* > **Policies** > **Protocol**.

b) Right-click on **Route Maps for Route Control** and choose **Create Route Maps for Route Control**.

   The **Create Route Maps for Route Control** dialog appears in the work pane.

c) Enter a name for the route map in the **Name** field.

   For the purposes of this example, we will enter **OSPF-to-BGP-static** as the name of this route map.

d) From the **Contexts** summary table in the **Create Route Maps for Route Control** dialog, click the **+**.

The **Create Route Control Context** window appears.

e) Enter a name for the route control context in the **Name** field.

For the purposes of this example, we will enter **OSPF-to-BGP-static-RCC** as the name of this route control context.

f) Click the **Associated Match Rules +** symbol to access the **Rule Name** field and choose the match rule that you created (for example, **OSPF-NH-static-NH-IPs**).

g) Locate the **Set Rule** drop-down menu and choose the set rule that you created (for example, **NH-Prop-SR_Mpath**).

h) Click **OK** in the **Create Route Control Context** window.

You are returned to the **Create Route Maps for Route Control** window.

i) Click **Submit** in the **Create Route Maps for Route Control** window.

Step 4 Attach the policy to the L3Out.

a) From the left navigation menu, select ***<tenant_name>*** > **Networking** > **L3Outs** > ***<l3out_name>***

b) In the main pane, select the **Policy** > **Main** tab.

c) Scroll down to the **Route Profile for Redistribution** area and click the plus (+) icon.

d) In the **Source** area, choose **static**.

e) In the **Route Map** area, choose the route map the previous steps.

f) Click **Submit**.

Verifying Configuration

This section describes how to verify that the configuration was deployed correctly using the APIC's managed object (MO) browser.

Procedure

Step 1 Open the **Object Store browser**.

You can verify the MOs using the **Object Store browser**, which you can open by clicking the **Help and Tools** icon in the top right of your UI and selecting **Object Store browser** or by navigating directly to `https://<apic-ip>/visore.html`.

Step 2 Verify that the fallback route group concrete MOs were created.

a) In the Object Store browser, search for `ipFBRGroup`.

b) In the list of the result, ensure that group object was created.

You can search for the group name, the corresponding object will have `groupName` field set to the group name you specified when creating it.

c) In the `dn` field of the object, click the right arrow to see its children.

Every prefix (`ipv4FBRMember`) and next-hop address (`ipv4FBRoute`) you added to the group should be defined under the group's object.

Step 3 Verify that the fallback route group resolved MOs were created.

There are two types of resolved MOs—one under `RtdCtxDef` and one under `RtdOutDef`, which contains all the FBR routes and member configured in the VRF. This also shows the fallback route group association configured in the L3Out.

In the case where one FBR group is associated with multiple L3Outs, a reference count MO `fvFBRGroupSrc` is created to track each associated L3Out. The FBR group can be removed only when all `fvFBRGroupSrc` are deleted.

Step 4 Check for any fallback route faults.

You can check any faults raised for the fallback route group in the APIC UI by navigating to the L3Out where the group is attached and selecting the **Faults** tab.

There is a single fault that may be raised for the fallback route feature if a route is not present in the fallback route group.

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

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

IN NO EVENT SHALL CISCO OR ITS SUPPLIERS BE LIABLE FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, INCLUDING, WITHOUT LIMITATION, LOST PROFITS OR LOSS OR DAMAGE TO DATA ARISING OUT OF THE USE OR INABILITY TO USE THIS MANUAL, EVEN IF CISCO OR ITS SUPPLIERS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

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

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

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

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA 95134-1706
USA

Asia Pacific Headquarters
CiscoSystems(USA)Pte.Ltd.
Singapore

Europe Headquarters
CiscoSystemsInternationalBV
Amsterdam,TheNetherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at www.cisco.com/go/offices.