Inter-Fabric Transit Routing with PBR for ACI Fabrics, 4.3.1

Table of Contents

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.

Release Version	Feature	Description
Nexus Dashboard 4.3.1	There were no major changes from the previous release.	

Inter-Fabric Transit Routing with PBR

The following sections describe the guidelines, limitations, and configuration steps for the Inter-Fabric Transit Routing with Policy-Based Redirect (PBR) use case in your Multi-Fabric domain.

The following sections apply to the inter-fabric transit routing (L3Out-to-L3Out) with PBR use case only. For information on L3Out-to-EPG inter-fabric communication with PBR, see Inter-Fabric L3Out with PBR instead; and for simple inter-fabric L3Out use cases without PBR, see Inter-Fabric L3Out.

The inter-fabric transit routing with PBR use case described in the following sections is supported for both inter-VRF and intra-VRF scenarios.

Configuration Workflow

The use case described in the following sections is an extension of a basic inter-fabric L3Out PBR use case which is in turn an extension on basic inter-fabric L3Out (without PBR) configuration. To configure this feature:

1. Configure basic external connectivity (L3Out) for each fabric.

   The inter-fabric L3Out with PBR configuration described in the following sections is built on top of existing external connectivity (L3Out) in each fabric. If you have not configured an L3Out in each fabric, create and deploy one as described in External Connectivity (L3Out) before proceeding with the following sections.

2. Create a contract between two external EPGs associated to the L3Outs deployed in different fabrics, as you typically would for the use case **without** PBR.
3. Add service chaining to the previously created contract as described in the following sections, which includes:
   - Creating a Service Device template and assigning it to fabrics.

     The service device template must be assigned to the fabrics for which you want to enable inter-fabric transit routing with PBR.

   - Providing fabric-level configurations for the Service Device template.

     Each fabric can have its own service device configuration including different high-availability models (such as active/active, active/standby, or independent service nodes).

   - Associating the service device you defined to the contract used for the inter-fabric L3Out use case you deployed in the previous step.

Refer to the *ACI Contract Guide* and *ACI PBR White Paper* to understand Cisco ACI contract and PBR terminologies.

Traffic Flow

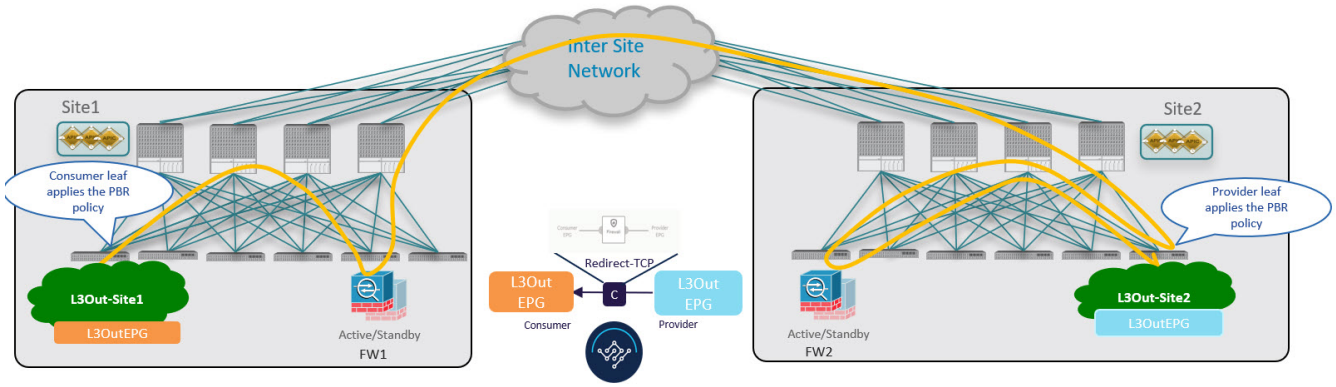
This section summarizes the traffic flow between two external EPGs in different fabrics.

In this case, the traffic flow in both directions is redirected through both firewalls in order to avoid asymmetric traffic flows due to independent FW services deployed in the two fabrics.

Consumer-to-Provider Traffic Flow

Because any IP prefix associated with the destination external EPG for classification purposes is automatically programmed (with its Class-ID) on the consumer leaf switch, the leaf switch can always resolve the class-ID of the destination external EPG and apply the PBR policy redirecting the traffic to the local FW.

After the firewall on the consumer fabric has applied its security policy, the traffic is sent back into the fabric and forwarded across fabrics towards the provider border leaf nodes connecting to the external destination. The border leaf node receiving the traffic originated from Fabric1 applies the PBR policy and redirects the traffic to the local firewall node. After the firewall applies its local security policy, the traffic is sent back to the border leaf nodes, which can now simply forward it toward the external destination.

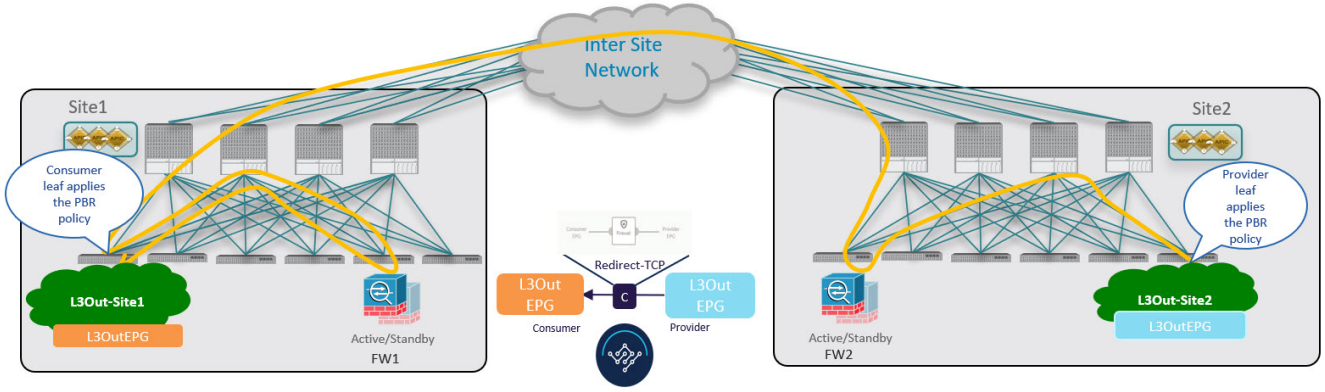


Consumer-to-Provider Traffic Flow

Provider-to-Consumer Traffic Flow

Similarly to the consumer-to-provider, the provider leaf switch can always resolve the class-ID of the destination external EPG and apply the PBR policy redirecting the traffic to the local FW in the other direction as well.

The traffic is then sent to the consumer fabric, where it is steered towards the local firewall before being forwarded to the external destination.



Provider-to-Consumer Traffic Flow

Inter-fabric transit routing with PBR guidelines and limitations

The following guidelines and limitations apply when deploying inter-fabric transit routing in vzAny with PBR with multi-fabric:

- PBR cannot be selected for the **Must not talk to** communication type when creating compliance rules.
- Inter-fabric transit routing with vzAny PBR is supported exclusively for single-node firewall on a one-arm interface.

  This use case is supported for fabrics running Cisco APIC release 6.0(4) or later.

- While you can use your existing Service Graph objects defined in Application templates for these use cases, we recommend using the new service chaining workflow and implicitly creating new service graphs by defining the policies in Service Device templates and associating them to contracts.

  The steps that are described in the following sections use the new Service Device templates to enable the supported use cases but will call out the specific differences when applicable.

  Configuration of Service Graph objects in Application templates will be deprecated in a future release.

- The L3Out VRF can be stretched (for intra-VRF use case) or fabric local (for inter-VRF cases).

  The following sections assume that you already have a VRF and L3Out already configured for each fabric.

  Note that the "Fabric-aware policy enforcement" and "L3 Multicast" options must be enabled for the vzAny VRF to enable the L3Out-toL3Out use cases discussed in this chapter.

  If you do not already have a VRF, you can create one in an Application template as you typically would. VRF configuration is described in detail in the "Configuring VRFs" section in Schemas and Application Templates for ACI Fabrics.

- You must enable the Fabric-aware Policy Enforcement Mode setting on the VRF to enable the new L3Out-toL3Out use cases. Enabling or disabling the Fabric-aware Policy Enforcement Mode option will cause a brief traffic disruption (including the already existing contracts between EPGs) because the zoning rules must be updated on the leaf switches. We recommend that you perform this operation during a maintenance window. Enabling Fabric-aware Policy Enforcement Mode increases TCAM usage on the leaf switches for the existing contracts and contracting permit logging cannot be used in conjunction with this option.
- The service BD to which you want to attach the service device interface must be configured as L2 Stretched (BUM forwarding is optional and should be disabled).

  If you do not already have a service BD, you can create one in an Application template. BD configuration is described in detail in the "Configuring Bridge Domains" section in Schemas and Application Templates for ACI Fabrics.

- The consumer, provider, and the service BDs must be configured in Hardware proxy-mode.
- The vzAny PBR destination must be connected to a stretched service BD, not to an L3Out.
- Only threshold down deny action and sip-dip-protocol hash is supported .
- The PBR destination must be in either the consumer or provider VRF instance. For example see *Cisco Application Centric Infrastructure Policy-Based Redirect Service Graph Design White Paper*.

The following is not supported for this use case:

- Specific remote leaf switch configurations.

  Specific considerations apply for multi-fabric deployments leveraging remote leaf nodes. Inter-fabric transit routing with PBR is not supported on vzAny PBR and L3Out-to-L3Out for communication between endpoints (consumer or provider) deployed on remote leaf nodes that belongs to different fabrics.

- Each VRF is limited to utilizing only one device in a one-arm configuration for vzAny-to-vzAny, L3OutEPG-to-L3OutEPG, and vzAny-To-L3OutEPG PBR. This restriction is enforced due to special ACL in APIC.
- We must use different firewall VLAN interfaces for redirection for vzAny-to-vzAny/L3OutEPG-to-L3OutEPG, and other use cases such as vzAny-to-EPG, EPG-to-EPG and EPG-to-L3OutEPG if they are in the same VRF.
- When vzAny with PBR is applied to north-south communication for any of the newly supported use cases (vzAny-to-vzAny, vzAny-to-L3OutEPG, L3OutEPG-to-L3OutEPG), ingress traffic optimization needs to be enabled for stretched subnets.
- Only one node service chain with L3 PBR destination is supported.
- Nexus Dashboard does not support chained L4-L7 service devices (for example, firewalls) when all devices attach to the same Bridge Domains (BDs) on both consumer and provider legs. This configuration leads to inaccurate flow path visualization, where the displayed path might show incorrect or unknown nodes and omit intermediate L4-L7 service devices.
- Contract Permit logging is not supported on the VRF that has Fabric-aware Policy Enforcement Mode is enabled, which is required for vzAny PBR and L3OutEPG-to-L3OutEPG PBR.
- Pod-aware vzAny with PBR is not supported.

Enhanced PBR support for communication compliance

The Explorer now visually indicates if communication between Endpoint Groups (EPGs) involves a PBR node. The policy table has been updated to include more details about PBR nodes with APIs. You can create compliance rules to enforce route traffic between EPGs through a PBR device using traffic selectors, and the rule creation APIs have been updated to support this capability.

For compliance anomalies generated by rules, consider the following:

- COMMUNICATION_COMPLIANCE_NOT_VERIFIED: This anomaly occurs if the L4-L7 match criteria selection results in an empty set because it cannot resolve to a deployed service graph instance.
- COMMUNICATION_COMPLIANCE_VIOLATION: This anomaly occurs in three scenarios if L4-L7 criteria are specified:
  - A direct contract exists from EPG1 to EPG2.
  - A redirect rule exists, but not through the intended L4-L7 device.
  - A redirect rule exists, but not for the intended traffic selector.

Creating compliance rules with PBR

The following instructions guide you through the process of creating compliance rules that enforce policies for traffic traversing PBR devices in your inter-fabric environment.

1. Navigate to the **Analyze Hub** page.

   Analyze > Analyze hub > Compliance

2. From the **Select Fabric** drop-down list, choose the fabric.
3. Click **Create compliance rule**.

   In the **Basic information** page, perform the following functions.

   a. **Name**: Enter the compliance rule name.
   b. **Description**: Provide a description of the compliance rule.
   c. **Fabrics**: Choose from the drop-down list, to check from the **Online fabrics** and **Snapshot fabrics**.
   d. **State**: Enabled by default.
4. Click **Next**.

   In the **Settings** page, perform the following functions.

   a. **Compliance rule type**: Choose from the options. **Communication** is chosen by default.
   b. **Communication type**: Choose the type of communication. Options include **Must talk to** is (chosen by default) and **May talk to**.
   c. **From objects**: Choose from the **Object type** drop-down list.

- (Optional) **Matching Criteria**: Click **+Add Criteria** to add the object type.

d. **Enforce communication path**: To route traffic through an L4-L7 device (such as a PBR node), check the **Via L4-L7 device** check box.

- (Optional) **Matching Criteria**: Click **+Add Criteria** to add the object type. Choose **Service graph device name** from the **Object type** drop-down list and enter the ***service graph device name*** in **Value**, as the selection is based on the service graph device name.

e. **To objects**: Choose type from the **Object type** drop-down list.

- (Optional) **Matching Criteria**: Click **+Add Criteria** to add the object type.

f. **Traffic selector**: Click **+Add traffic selector rule**

5. Click **Next**.

   The **Summary** page appears with the information provided.

6. Click **Add compliance requirement** to create the rules.

Create Service Device Template

Before you begin:

- Ensure that you have read and completed the requirements described in Inter-fabric transit routing with PBR guidelines and limitations.
- You must have created a stretched service bridge domain (BD) to use with the service nodes you will define in this section.

  If you do not already have a service BD, you can create one in an Application template as you typically would. BD configuration is described in detail in the "Configuring Bridge Domains" section in Schemas and Application Templates for ACI Fabrics.

The following steps describe how to create a Service Device template with a service node and its settings which you will use for the inter-fabric transit routing use cases.

1. Navigate to the **Orchestration** page.

   Manage > Orchestration

2. Click **Tenant Templates**.
3. (Optional) Create a tenant policies template and an IP SLA monitoring policy.

   We recommend that you configure an IP SLA policy for traffic redirection as it simplifies the configuration of the PBR policy described in Step 7 below. If you have an IP SLA policy already defined, you can skip this step, otherwise:

   a. Click **Tenant Policies**.
   b. Click **Create Tenant Policy Template**.
   c. In the **Tenant Policies** page's right properties sidebar, provide the **Name** for the template and **Select a Tenant**.
   d. In the **Template Properties** page, choose **Actions > Add/Remove Fabrics** and associate the template with both fabrics.
   e. In the main pane, choose **Create Object > IP SLA Monitoring Policy**.
   f. Provide the **Name** for the policy, and define its settings.
   g. Click **Save** to save the template.
   h. Click **Deploy Template** to deploy it.
4. Create a Service Device template and associate it with a tenant and with the fabrics.
   a. From **Manage > Orchestration > Tenant Templates**, click **Service Device**.
   b. Click **Create Service Device Template**.
   c. In the template properties sidebar that opens, provide the **Name** for the template and **Select a Tenant**.
   d. In the **Template Properties** page, choose **Actions > Add/Remove Fabrics** and associate the template with both fabrics.
   e. Click **Save** to save the template.
5. Create and configure the device cluster.

a. In the **Template Properties** page (template-level configuration), choose **Create Object Service Device Cluster**.

   The device cluster defines the service to which you want to redirect traffic. This release supports redirection to a firewall service node that can be deployed with three different redundancy models: active/standby, active/active, or a cluster of multiple independent nodes. The provisioning for those different options is covered in Step 7 below. Note that you can choose the firewall deployment model at the fabric level and different options can be deployed across different fabrics that are part of the same Multi-Fabric domain.

b. In the **<cluster-name>** sidebar, provide the **Name** for the cluster.

   The **Device Location** and **Device Mode** are pre-populated based on the currently supported use case. **Device Location** should be pre-configured as ACI On-Prem and **Device Mode** as L3.

c. For the **Device Type**, choose Firewall.

d. For **Device Mode**, choose L3.

e. For **Connectivity Mode**, choose One Arm or Two Arm suitable for your deployment. Nexus Dashboard supports service devices connected in both One-Arm and Two-Arm modes.

When changing the device connectivity mode (One-Arm, Two-Arm, or Advanced), interface configuration requirements may differ. If you attempt to modify interface settings while a contract uses the interface, you will receive a warning message, and Nexus Dashboard will restrict the modification to prevent disruptions to existing deployments.

Nexus Dashboard conducts validations for One-Arm and Two-Arm modes. In Advanced mode, no validations are performed; Nexus Dashboard assumes that you are experienced and understand the implications of selecting this mode.

f. Provide the **Interface Name**.

g. For the **Interface Type**, choose BD.

h. Click **Select BD >** to choose the service bridge domain to which you want to attach this device.

   This is the stretched service BD you created as part of the Inter-fabric transit routing with PBR guidelines and limitations, for example FW-external.

i. For the **Redirect** option, choose Yes.

   You must choose to enable redirect for the PBR use case. After choosing Yes, the **IP SLA Monitoring Policy** option becomes available.

j. (Optional) Click **Select IP SLA Monitoring Policy** and choose the IP SLA policy you have created in a previous step.

k. (Optional) In the **Advanced Settings** area, choose **Enable** if you want to provide additional settings for the service cluster.

You can configure the following advanced settings:

- **QoS Policy** - allows you assign a specific QoS level within the ACI fabrics for the redirected traffic.
- **Preferred Group** - specifies whether or not this service device interface is part of the preferred group.
- **Load Balancing Hashing** - allows you to specify the hashing algorithm for PBR load balancing.

You must keep the default value for the vzAny-to-vzAny, vzAny-to-ExtEPG, and ExtEPG-to-ExtEPG use cases as they support only the default configuration. You can change the load balancing hashing for other use cases: EPG-to-EPG, ExtEPG-to-EPG and vzAny-to-EPG.

For additional information, see *ACI Policy-Based Redirect Service Graph Design*.

- **Pod Aware Redirection** - can be configured in Multi-Pod configuration if you want to specify the preferred PBR node. When you enable Pod-aware redirection, you can specify the Pod ID and redirection is programmed only in the leaf switches located in the specified Pod.
- **Rewrite Source MAC** - updates the source MAC address if the PBR node uses "source MAC based forwarding" instead of IP based forwarding.

  For additional information, see *ACI Policy-Based Redirect Service Graph Design*.

- **Advanced Tracking Options** - allows you to configure a number of advanced settings for the service node tracking. For additional information, see *Policy-Based Redirect and Threshold Settings for Tracking Service Nodes*.

l. Click **Ok** to save.

Note that after you create the Service Device Cluster, it is highlighted in red in the **Template Properties** (template-level configuration) page. At this point, you have defined redirection to a firewall service, but you must still provide the firewall information and the redirect policy you want to use at the fabric-local level.

6. Provide fabric-local configuration for the Service Device Cluster you created in the previous step.
   a. In the **Service Device Template** screen, choose the **<fabric-name>** tab.
   b. At the fabric level, choose the Service Device Cluster you created.
   c. In the properties sidebar, choose the **Domain Type**.

   You can choose whether the firewall device in this fabric is Physical or VMM (virtual and hosted by a hypervisor that is part of a VMM domain).

   d. Click **Select Domain** to choose the domain to which this firewall device belongs.

   You can choose either a physical or a virtual domain.

   - If you choose a physical domain, provide the following information:
     - **VLAN** - you must provide the VLAN ID used for traffic between the fabric and the

firewall device.

- **Fabric to Device Connectivity** - provide the switch node and interface information for the fabric's connectivity to the firewall device.

- If you choose a VMM domain, provide the additional options:

Nexus Dashboard does not support using the same DVS uplinks for both endpoint VMs and virtual firewall traffic when deploying a virtual firewall within a VMM domain, even with different VLANs. This configuration leads to inaccurate traffic counters and latency measurements for L4-L7 services. To ensure accurate telemetry, virtual firewalls and endpoint VMs must reside in separate VMM domains or use separate DVS uplinks.

  - **Trunking Port** - used to enable tagged traffic for the L4-L7 VM.

    By default, the ACI service graph configuration creates access-mode port groups and attaches them to the vNIC of the L4-L7 VM automatically.

  - **Promiscuous Mode** - required if the L4-L7 virtual appliance must receive traffic destined to a MAC address that is not the vNIC MAC owned by the VM.
  - **VLAN** - optional configuration for VMM domains and will be allocated from the dynamic VLAN pool associated with the domain if not specified.
  - **Enhanced LAG Option** - if you are using enhanced LACP for the port channel between the hypervisor and the fabric.
  - **VM Name** - choose the firewall's VM from the list of all VMs available in this VMM domain and the interface (**VNIC**) used for the firewall traffic.

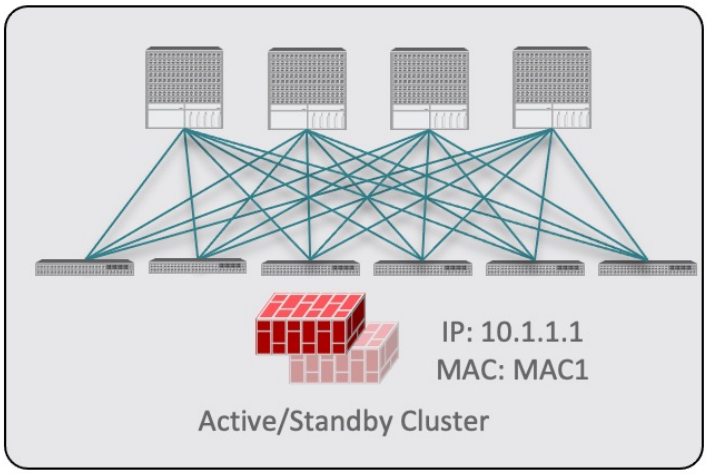
    Depending on the kind of device cluster you are deploying, click **+Add VM information** to provide additional cluster nodes.

7. Provide the FW device information and PBR destination IP addresses.

   As previously mentioned, this release supports 3 deployment options for high-availability FW clusters: active/standby clusters, active/active clusters, and independent active nodes. In all three deployment options, the use of an IP SLA policy (mentioned in Step 3) allows to specify only the IP address of the firewall nodes, and the corresponding MAC address will be automatically discovered.

You can deploy different designs in different fabrics.

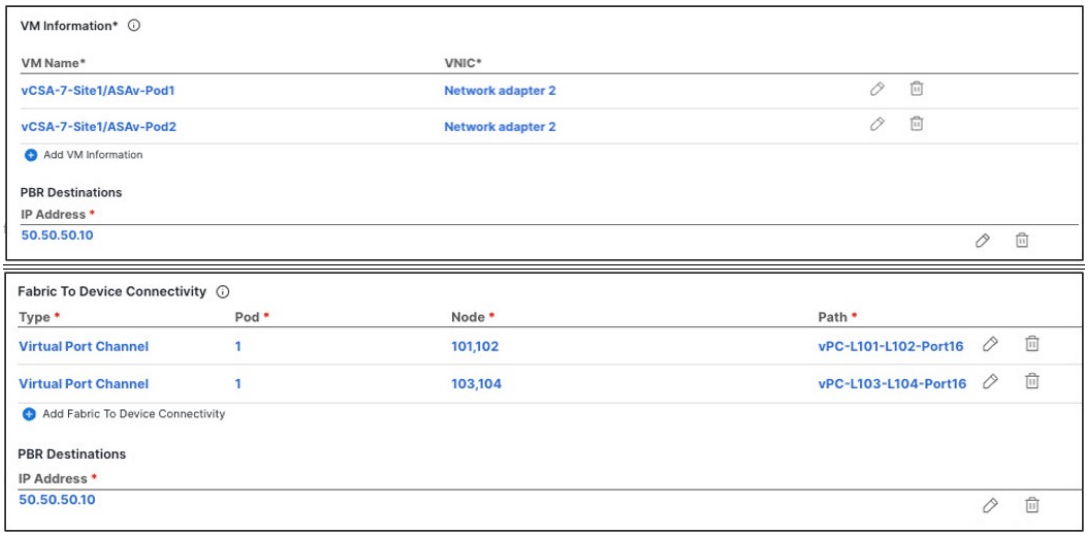
   - Active/standby clusters are identified by a single MAC/IP pair.



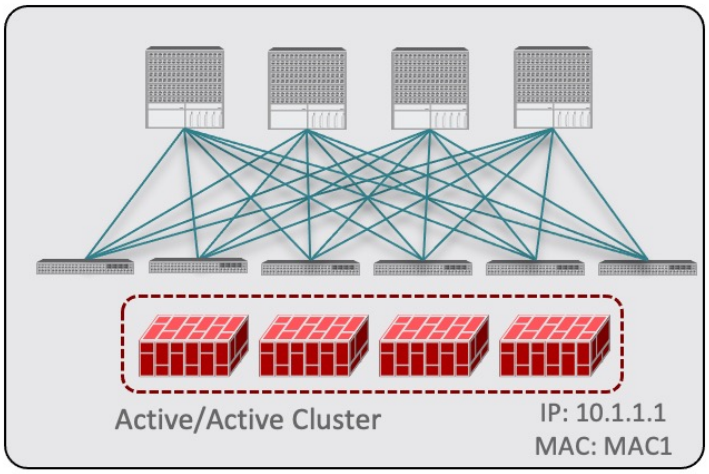
In this case, you need to provide a single PBR destination IP address identifying the active firewall node and also include information about every node in the cluster.

For example, for a 2-nodes active/standby cluster, you would provide the following:

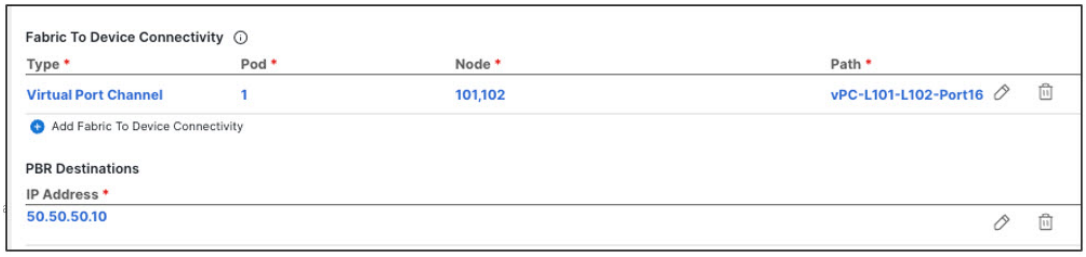
- For a virtual firewall cluster, the VMs representing the active and standby firewall nodes and the IP address of the active firewall as PBR destination.
- For a physical firewall cluster, the interfaces used to connect the active and standby firewall nodes to the leaf switches of the fabric (vPC interfaces in the specific example below) and the IP address of the active firewall as PBR destination.



- Active/active clusters are also identified by a single MAC/IP pair.

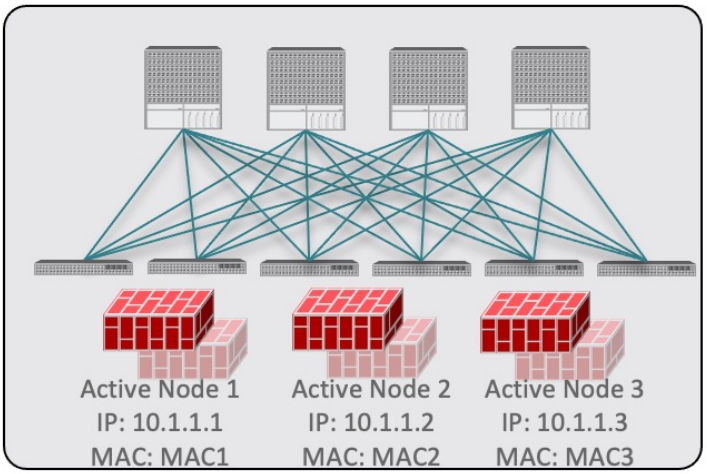


For Cisco Firewalls (ASA or FTD models), the Active/Active cluster with a single MAC/IP pair is only supported for physical form factors, and all the cluster nodes own the same MAC/IP address and must be connected to the same vPC logical connection deployed on a pair of ACI leaf switches. As a result, the figure below shows how a single vPC interface and a single PBR Destination IP address should be configured on Nexus Dashboard, where the MAC address is dynamically discovered when using an IP SLA policy mentioned for the previous use case.



- For independent active nodes configuration, each active node is identified by a unique MAC/IP addresses pair.

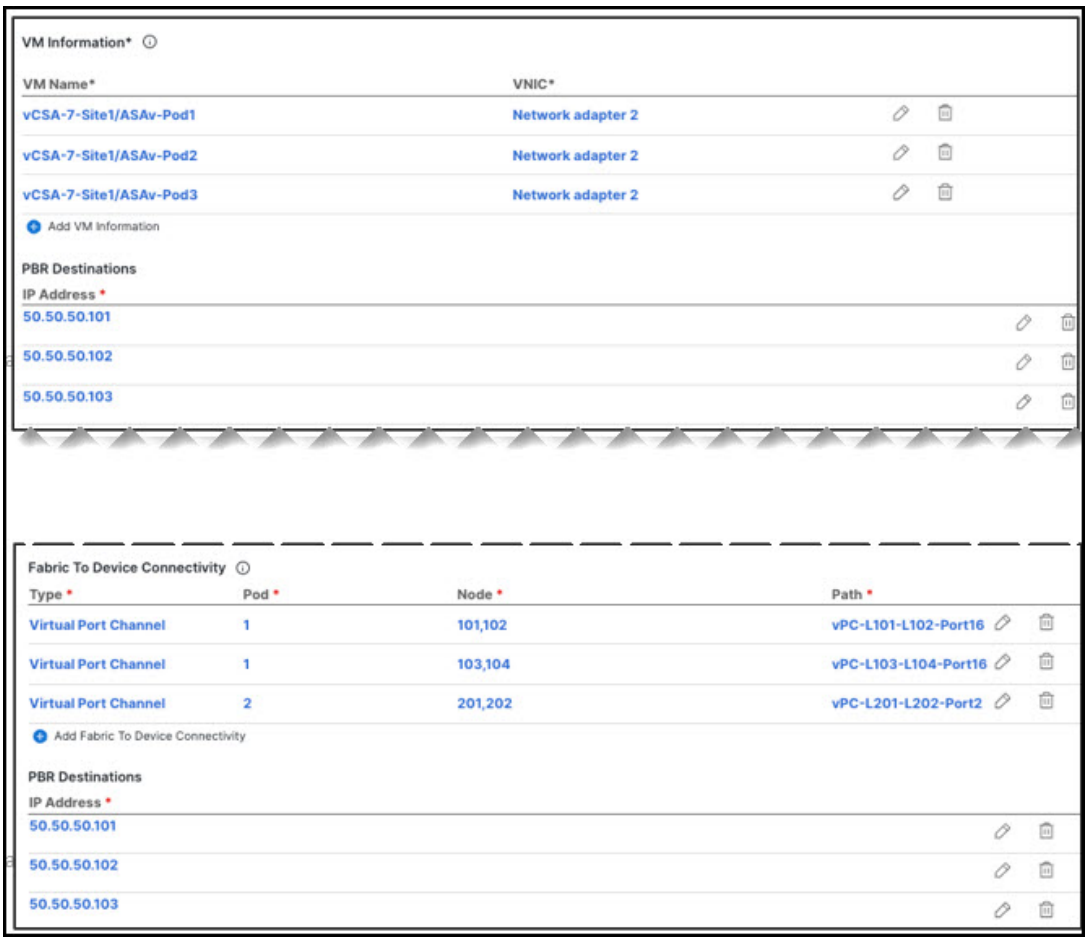
  Note that symmetric PBR ensures that the traffic is handled by the same active node in both directions.



In this case, you must provide individual PBR Destination IP addresses for each active node as well as each node's information in your Nexus Dashboard configuration.

For example, for a deployment of 3 independent firewall nodes, you would provide the following:

- For a virtual firewall form factor, the VMs representing the 3 firewall nodes and their unique IP addresses as PBR destinations.
- For a physical firewall form factor, the interfaces used to connect each firewall node to the leaf switches of the fabric (vPC interfaces in the specific example below) and the unique IP addresses of each firewall node as PBR destinations.



a. Click **Add Fabric To Device Connectivity** (physical domain) or **Add VM Information** (VMM domain).

   Depending on whether you selected physical or VMM domain in the previous step, you will specify information for either the firewall VM or the physical fabric connectivity to the firewall device.

   For physical domains, provide the Pod, switch node, and the interface information.

   For VMM domains, provide the VM name and vNIC information.

b. Click **Add PBR Destination** to provide the IP address of the interface on the firewall that is connected to the service bridge domain.

   Depending on the kind of device cluster you are deploying, you may need to provide one or more PBR destination IP addresses:

   This does not provision the IP address on the firewall's interface, but simply configures redirection of traffic toward that IP address. The specific firewall configuration is not deployed from Nexus Dashboard, and you must provision it separately.

c. Click **Ok** to save the provided configuration.

d. Repeat this step for the other fabric with which you associated the template.

8. Save and deploy the template.

   a. At the **Service Device Template** level, click **Save** to save the template configuration.

   b. Choose the **Template Properties** tab and click **Deploy Template** to push the configuration to the fabrics.

   c. (Optional) Verify that the configuration was created at the fabric level.

      You can verify that the L4-L7 device is configured in the APIC by navigating to **<tenant-name> > Services > L4-L7 > Devices > <cluster-name>** in the APIC GUI. This shows the device cluster along with all the configuration you have provided in the previous steps.

      To verify that the PBR policy is now configured on the APIC, navigate to **<tenant-name> > Policies > Protocol > L4-L7 Policy-Based Redirect** and you should see the **<cluster-name>-one-arm** redirect defined with the IP SLA monitoring policy you chose in *Step 5j* and the IP address you provided in *Step 7d*.

What to do next:

After you have deployed the service device configuration, create the application template, external EPGs, and a contract with which you will associate the service chaining as described in Create Contract and Add Service Chaining.

Create Contract and Add Service Chaining

Before you begin:

- You must have created and deployed external connectivity (L3Out) configuration in each fabric as described in External Connectivity (L3Out).
- You must have created and deployed the service device template containing the device configuration as described in Create Service Device Template.

After you have created and deployed the service device templates and created the application template with the external EPGs for the L3Outs in each fabric, you can enable inter-fabric transit routing with policy-based redirection by creating a contract between the external EPGs and associating the contract with the service devices you created in a previous section.

1. Navigate to the application templates where you want to create the external EPGs for the L3Outs and a contract between the external EPGs. Typically you would define the external EPGs in different fabric local templates associated to each fabric.
2. Create two external EPGs and associate the L3Outs in each fabric to the external EPG(s) at the fabric level.

   This is the same process as you typically use when creating external connectivity for a fabric. Detailed information about the L3Out templates and External EPGs is described in External Connectivity (L3Out).

3. Create a contract in a stretched template associated to both fabrics, as you typically would and associate the contact with both external EPGs.

   In this case, one of the external EPGs will be the consumer and the other one is the provider.

4. Choose the contract you created.
5. In the **Service Chaining** area, click **+Service Chaining**.

   These steps assume that you have configured a brand new service device for this use case using the Service Device template workflow as described in Create Service Device Template. If you already have a Service Graph defined in an application template, choose Service Graph instead and then select the existing service graph. However, keep in mind that the Service Graph option will be deprecated in a future release.

6. For **Device Type**, choose Firewall.

   This release supports a service device connected in one arm mode only.

7. From the **Device** dropdown, choose the FW device cluster you created in the previous step.
8. Ensure that **Consumer Connector Type Redirect** is enabled.
9. Ensure that **Provider Connector Type Redirect** is enabled.
10. Click **Add** to continue.
11. Click **Save** to save the template.
12. Click **Deploy Template** to deploy it.

Copyright

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706

USA
https://www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883