Nexus Dashboard Orchestrator Layer 3 Multicast for ACI Fabrics, Release 4.3.x

Table of Contents

Layer 3 Multicast

Cisco Multi-Site Layer 3 multicast is enabled or disabled at three levels, the VRF, the bridge domain (BD), and any EPGs that have multicast sources present.

At the top level, multicast routing must be enabled on the VRF that has any multicast-enabled BDs. On a multicast-enabled VRF, there can be a combination of multicast-enabled BDs and BDs where multicast routing is disabled. Enabling multicast routing on a VRF from the Cisco Nexus Dashboard Orchestrator GUI enables it on the APIC sites where the VRF is stretched.

Once a VRF is enabled for multicast, the individual BDs under that VRF can be enabled for multicast routing. Configuring Layer 3 multicast on a BD enables protocol independent multicast (PIM) routing on that BD. By default, PIM is disabled in all BDs.

If a source belonging to a specific site-local EPG sends multicast traffic to a remote site, the Nexus Dashboard Orchestrator must create a shadow EPG and program the corresponding subnet route(s) on the remote site for the source EPG. In order to limit the configuration changes applied to the remote Top-of-Rack (TOR) switches, you are required to explicitly enable Layer 3 multicast on the local EPGs that have multicast sources present, so that only the configuration necessary for those EPGs is pushed to the remote sites. EPGs with multicast receivers do not require enabling Layer 3 multicast.

Multi-Site supports all of the following Layer 3 multicast source and receiver combinations:

- Multicast sources and receivers inside ACI fabric
- Multicast sources and receivers outside ACI fabric
- Multicast sources inside ACI fabric with external receivers
- Multicast receivers inside ACI fabric with external sources

Layer 3 Multicast Routing

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

- When the multicast source is attached to the ACI fabric as an endpoint (EP) at one site and starts streaming a multicast flow, the specific site's spine switch that is elected as designated forwarder for the source VRF will forward the multicast traffic to all the remote sites where the source's VRF is stretched using Head End Replication (HREP). If there are no receivers in a specific remote site for that specific group. the traffic gets dropped on the receiving spine node. If there is at least a receiver, the traffic is forwarded into the site and reaches all the leaf nodes where the VRF is deployed and at that point is pruned/forwarded based on the group membership information.
- Prior to Cisco ACI Release 5.0(1), the multicast routing solution required external multicast routers to be the Rendezvous Points (RPs) for PIM-SM any-source multicast (ASM) deployments. Each site must point to the same RP address for a given stretched VRF. The RP must be reachable on each site via the site's local L3Out.
- When the source is outside and the receiver is within a fabric, the receiver will pull traffic via site's local L3Out as PIM joins toward RP and source are always sent via site local L3Out.
- Receivers in each site are expected to draw traffic from an external source via the site's local L3Out. As such, traffic received on the L3Out on one site should not be sent to other sites. This is achieved on the spine by pruning multicast traffic from being replicated into HREP tunnels.

  In order to be able to do so, all multicast traffic originated from an external source and received on a local L3Out is remarked with a special DSCP value in the outer VXLAN header. The spines can hence match that specific DSCP value preventing the traffic from being replicated toward the remote sites.

- Traffic originated from a source connected to a site can be sent toward external receivers via a local L3Out or via L3Outs deployed in remote sites. The specific L3Out that is used for this solely depends on which site received the PIM Join for that specific multicast group from the external network.
- When multicast is enabled on a BD and an EPG on the Nexus Dashboard Orchestrator, all of the BD's subnets are programmed in the routing tables of all the leaf switches, including the border leaf nodes (BLs). This enables receivers attached to the leaf switches to determine the reachability of the multicast source in cases where the source BD is not present on the leaf switches. The subnet is advertised to the external network if there is a proper policy configured on the BLs. The /32 host routes are advertised if host-based routing is configured on the BD.

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

Rendezvous Points

Multicast traffic sources send packets to a multicast address group, with anyone joining that group able to receive the packets. Receivers that want to receive traffic from one or more groups can request to join the group, typically using Internet Group Management Protocol (IGMP). Whenever a receiver joins a group, a multicast distribution tree is created for that group. A Rendezvous Point (RP) is a router in a PIM-SM multicast domain that acts as a shared root for a multicast shared tree.

The typical way to provide a redundant RP function in a network consists in deploying a functionality called Anycast RP, which allows two or more RPs in the network to share the same anycast IP address. This provides for redundancy and load balancing. Should one RP device fails, the other RP can take over without service interruption. Multicast routers can also join the multicast shared tree by connecting to any of the anycast RPs in the network, with PIM join requests being forwarded to the closest RP.

Two types of RP configurations are supported from Nexus Dashboard Orchestrator:

- **Static RP**-If your RP is outside the ACI fabric.
- **Fabric RP**-If the border leaf switches in the ACI fabric will function as the anycast RPs.

Any number of routers can be configured to work as RPs and they can be configured to cover different group ranges. When defining the RP inside the ACI fabric, you can configure which groups the RP covers by creating a route-map policy that contains the list of groups and attaching this policy to the RP when adding it to the VRF. Creating a route map is described in Creating Multicast Route Map Policy, while VRF configuration is described in Enabling Any-Source Multicast (ASM) Multicast.

Both static and fabric RPs require PIM-enabled border leaf switches in the VRF where multicast routing is enabled. L3Out configuration is currently configured locally from the APIC at each site including enabling PIM for the L3Out. Please refer to the *Cisco APIC Layer 3 Networking Configuration Guide* for details on configuration PIM on L3Outs

Multicast Filtering

Multicast filtering is a data plane filtering feature for multicast traffic available starting with Cisco APIC, Release 5.0(1) and Nexus Dashboard Orchestrator, Release 3.0(1).

Cisco APIC supports control plane configurations that can be used to control who can receive multicast feeds and from which sources. In some deployments, it may be desirable to constrain the sending and/or receiving of multicast streams at the data plane level. For example, you may need to allow multicast senders in a LAN to be able to send only to specific multicast groups or to allow receivers to receive multicast from only specific sources.

To configure multicast filtering from the Nexus Dashboard Orchestrator, you create source and destination multicast route maps, each of which contains one or more filter entries based on the multicast traffic's source IP and/or group with an action (Permit or Deny) attached to it. You then enable the filtering on a bridge domain by attaching the route maps to it.

When creating a multicast route map, you can define one or more filter entries. Some entries can be configured with a Permit action and other entries can be configured with a Deny action, all within the same route map. For each entry, you can provide a **Source IP** and a **Group IP** to define the traffic that will match the filter. You must provide at least one of these fields, but can choose to include both. If one of the fields is left blank, it will match all values.

You can enable both multicast source filtering and multicast receiver filtering on the same bridge domain. In this case one bridge domain can provide filtering for both, the source as well as the receivers.

If you do not provide a route map for a BD, the default action is to allow all multicast traffic on the bridge domain. However, if you do select a route map, the default action changes to deny any traffic not explicitly matched to a filter entry in the route map.

Source Filtering

For any multicast sources that are sending traffic on a bridge domain, you can configure a route map policy with one or more source and group IP filters defined. The traffic is then matched against every entry in the route map and one of the following actions takes place:

- If the traffic matches a filter entry with a Permit action in the route map, the bridge domain will allow traffic from that source to that group.
- If the traffic matches a filter entry with a Deny action in the route map, the bridge domain will block traffic from that source to that group.
- If the traffic does not match any entries in the route map, the default Deny action is applied.

Source filter is applied to the bridge domain on the First-Hop Router (FHR), represented by the ACI leaf node where the source is connected. The filter will prevent multicast from being received by receivers in different bridge domains, the same bridge domain, and external receivers.

Destination (Receiver) Filtering

Destination (receiver) filtering does not prevent receivers from joining a multicast group. The multicast traffic is instead allowed or dropped in the data plane based on the source IP and multicast group

combination.

Similarly to the source filtering, when multicast traffic matches a destination filter, one of the following actions takes place:

- If the traffic matches a filter entry with a Permit action in the route map, the bridge domain will allow the traffic from the multicast group to the receiver.
- If the traffic matches a filter entry with a Deny action in the route map, the bridge domain will block the traffic from the multicast group to the receiver.
- If the traffic does not match any entries in the route map, the default Deny action is applied.

Destination filter is applied to the bridge domain on the Last-Hop Router (LHR), represented by the ACI leaf node where the receiver is connected, so other bridge domains can still receive the multicast traffic.

Layer 3 Multicast Guidelines and Limitations

Up to the current software release, Cisco Nexus Dashboard Orchestrator cannot be used to deploy specific multicast control plane filtering policies, such as IGMP or PIM related policies, on each site. As such you must configure any additional policies required for your use case on each APIC site individually for end-to-end solution to work. For specific information on how to configure those settings on each site, see the *Cisco APIC Layer 3 Networking Configuration Guide*.

You must also ensure that QoS DSCP translation policies in all fabrics are configured consistently. When you create custom QoS policies in ACI fabrics, you can create a mapping between the ACI QoS Levels and the packet header DSCP values for packets ingressing or egressing the fabric. The same ACI QoS Levels must be mapped to the same DSCP values on all sites for the multicast traffic to transit between those sites. For specific information on how to configure those settings on each site, see the *Cisco APIC and QoS*

Multicast Filtering

The following additional guidelines apply if you enable the multicast filtering:

- Multicast filtering is supported only for IPv4.
- You can enable either the multicast source filtering, or the receiver filtering, or both on the same bridge domain.
- If you do not want to have multicast filters on a bridge domain, do not configure a source filter or destination filter route maps on that bridge domain.

  By default, no route maps are associated with a bridge domain, which means that all multicast traffic is allowed. If a route map is associated with a bridge domain, only the permit entries in that route map will be allowed, while all other multicast traffic will be blocked.

  If you attach an empty route map to a bridge domain, route maps assume a deny-all by default, so all sources and groups will be blocked on that bridge domain.

- Multicast filtering is done at the BD level and apply to all EPGs within the BD. As such you cannot configure different filtering policies for different EPGs within the same BD. If you need to apply filtering more granularly at the EPG level, you must configure the EPGs in separate BDs.
- Multicast filtering is intended to be used for Any-Source Multicast (ASM) ranges only. Source-Specific Multicast (SSM) is not supported for source filtering and is supported only for receiver filtering.
- For both, source and receiver filtering, the route map entries are matched based on the specified order of the entry, with lowest number matched first. This means that lower order entries will match first, even if they are not the longest match in the list, and higher order entries will not be considered.

  For example, if you have the following route map for the 192.0.3.1/32 source:

Order	Source IP	Action
1	192.0.0.0/16	Permit
2	192.0.3.0/24	Deny

Even though the second entry (192.0.3.0/24) is a longer match as a source IP, the first entry (192.0.0.0/16) will be matched because of the lower order number.

Creating Multicast Route Map Policy

This section describes how to create a multicast route map policy. You may want to create a route map for one of the following reasons:

- Define a set of filters for multicast source filtering.
- Define a set of filters for multicast destination filtering.
- Define a set of group IPs for a Rendezvous Point (RP).

  When configuring an RP for a VRF, if you do not provide a route map, the RP will be defined for the entire multicast group range (224.0.0.0/4). Alternatively, you can provide a route map with a group or group range that is defined to limit the RP to those groups only.

1. Log in to your Cisco Nexus Dashboard and open the Cisco Nexus Dashboard Orchestrator service.
2. Create a new Tenant Policy.
   a. From the left navigation pane, choose **Configure > Tenanat Template > Tenant Policies**.
   b. On the **Tenant Policy Templates** page, click **Add Tenant Policy Template**.
   c. In the Tenant Policies page's right properties sidebar, provide the **Name** for the tenant.
   d. From the **Select a Tenant** drop-down, choose the tenant with which you want to associate this template.

      All the policies that you create int his template as described in the following steps will be associated with the selected tenant deployed to it when you push the template to a specific site.

      By default, the new template is empty, so you need to add one or more tenant policies as described in the following steps. You don't have to create every policy available in the template - you can create a template with just a single route map policy for your multicast use case.
3. Create a Route Map Policy for Multicast.
   a. From the **+Create Object** dropdown, select **Route Map Policy for Multicast**.
   b. In the right properties sidebar, provide the **Name** for the policy.
   c. (Optional) Click **Add Description** and provide a description for the policy.
   d. Click **+Add Route Map for Multicast Entries** and provide the route map information.
      For each route map, you must create one or more route map entries. Each entry is a rule that defines an action based on one or more matching criteria based on the following information:
      - **Order** - Order is used to determine the order in which the rules are evaluated.
      - **Group IP**, **Src IP**, and **RP IP** - You can use the same multicast route map policy UI for two different use cases—To configure a set of filters for multicast traffic or to restrict a rendezvous point configuration to a specific set of multicast groups. Depending on which use case you're configuring, you must fill some of the fields in this screen:
        - For multicast filtering, you can use the **Source IP** and the **Group IP** fields to define the filter. You must provide at least one of these fields, but can choose to include

both. If one of the fields is left blank, it matches all values.
The Group IP range must be between 224.0.0.0 and 239.255.255.255 with a netmask between /4 and /32. You must provide the subnet mask.
The **RP IP** (Rendezvous Point IP) is not used for multicast filtering route maps, so leave this field blank.

- For Rendezvous Point configuration, you can use the **Group IP** field to define the multicast groups for the RP.
  The Group IP range must be between 224.0.0.0 and 239.255.255.255 with a netmask between /4 and /32. You must provide the subnet mask.
  For a Rendezvous Point configuration, the **RP IP** is configured as part of the RP configuration. If a route-map is used for group filtering it is not necessary to configure an **RP IP** address in the route-map. In this case, leave the **RP IP** and **Source IP** fields empty.

- **Action** - Action defines the action to perform, either Permit` or Deny the traffic, if a match is found.

e. Click the check mark icon to save the entry.

f. Repeat the previous substeps to create any additional route map entries for the same policy.

g. Click **Save** to save the policy and return to the template page.

h. Repeat this step to create any additional Route Map for Multicast policies.

Enabling Any-Source Multicast (ASM) Multicast

Before you begin:

- Ensure you have read and followed the information described in Layer 3 Multicast Guidelines and Limitations.
- If you plan to enable multicast filtering, create the required multicast route maps, as described in Creating Multicast Route Map Policy.
- Note that the site-local L3Outs must have PIM enabled in the VRF when fabric RP is enabled.

  This is described in Step 6 of the following procedure. Additional information about PIM configuration on an L3Out is available in the *Cisco APIC Layer 3 Networking Configuration Guide*.

The following procedure describes how to enable ASM multicast on VRF, BD, and EPG using the Nexus Dashboard Orchestrator GUI. If you want to enable SSM multicast, follow the steps in Enabling Source-Specific Multicast (SSM) instead.

1. Log in to your Nexus Dashboard Orchestrator.
2. From the left-hand sidebar, select the **Configure > Tenanat Template > Applications > Schemas** view.
3. Click on the Schema you want to change.
4. Enable Layer 3 multicast on a VRF.

   First, you enable Layer 3 multicast on a VRF that is stretched between sites.

   a. Select the VRF for which you want to enable Layer 3 multicast.

   b. In the right properties sidebar, check the **L3 Multicast** checkbox.
5. Add one or more Rendezvous Points (RP).

   a. Select the VRF.

   b. In the right properties sidebar, click **Add Rendezvous Points**.

   c. With the VRF still selected, click **Add Rendezvous Points** in the right sidebar.

   d. In the **Add Rendezvous Points** window, provide the IP address of the RP.

   e. Choose the type of the RP.

      - **Static RP**-If your RP is outside the ACI fabric.
      - **Fabric RP**-If your RP is inside the ACI fabric.

   f. (Optional) From the **Multicast Route-Map Policy** dropdown, select the route-map policy you configured previously.

      By default, the RP IP you provide applies to all multicast groups in the fabric. If you want to restrict the RP to only a specific set of multicast groups, define those groups in a route map policy and select that policy here.
6. Enable PIM on the L3Out.

Both static and fabric RPs require PIM-enabled border leaf switches where multicast routing is enabled. L3Out configuration currently cannot be done from the Nexus Dashboard Orchestrator, so you must ensure that PIM is enabled directly in the site's APIC. Additional information about PIM configuration on an L3Out is available in the *Cisco APIC Layer 3 Networking Configuration Guide*.

a. Log in to your site's Cisco APIC.

b. In the top menu, click **Tenants** and select the tenant that contains the L3Out.

c. In the left navigation menu, select **Networking > L3Outs > <l3out-name>**.

d. In the main pane, choose the **Policy** tab.

e. Check the **PIM** options.

Multi-Site supports IPv4 multicast only.

7. Enable Layer 3 multicast on a BD.

Once you have enabled L3 Multicast on a VRF, you can enable L3 multicast on a Bridge Domain (BD) level.

a. Select the BD for which you want to enable Layer 3 multicast.

b. In the right properties sidebar, check the **L3 Multicast** checkbox.

8. (Optional) If you want to configure multicast filtering, provide the route-maps for source and destination filtering.

a. Select the BD.

b. In the right properties sidebar, select a **Route-Map Source Filter** and **Route-Map Destination Filter**.

You can choose to enable either the multicast source filtering, or the receiver filtering, or both.

Keep in mind, if you do not select a route map, the default action is to allow all multicast traffic on the bridge domain; however, if you select a route map the default action changes to deny any traffic not explicitly matched to a filter entry in the route map.

9. If your multicast source is in one site and is not stretched to the other sites, enable intersite multicast source option on the EPG.

Once you have enabled L3 Multicast on the BD, you must also enable multicast on the EPGs (part of multicast-enabled BDs) where multicast sources are connected.

a. Select the EPG for which you want to enable Layer 3 multicast.

b. In the right-hand sidebar, check the **Intersite Multicast Source** checkbox.

Enabling Source-Specific Multicast (SSM)

Before you begin:

- Ensure you have read and followed the information that is described in Layer 3 Multicast Guidelines and Limitations.
- If you plan to enable multicast filtering, create the required multicast route maps, as described in Creating Multicast Route Map Policy.
- You must configure IGMPv3 interface policy for the multicast-enabled BDs at the site-local level.

  This is described in Step 8 of the following procedure. Additional information is available in the *Cisco APIC Layer 3 Networking Configuration Guide*.

The following procedure describes how to enable SSM multicast on VRF, BD, and EPG using the Cisco Nexus Dashboard Orchestrator GUI. If you want to enable ASM multicast, follow the steps in Enabling Any-Source Multicast (ASM) Multicast instead.

1. Log in to your Cisco Nexus Dashboard Orchestrator.
2. From the left sidebar, select the **Configure > Tenant Template > Application > Schemas** view.
3. Click the Schema that you want to change.
4. Enable Layer 3 multicast on a VRF.

   First, you enable Layer 3 multicast on a VRF that is stretched between sites.

   a. Select the VRF for which you want to enable Layer 3 multicast.
   b. In the right properties sidebar, check the **L3 Multicast** check box.
5. (Optional) Configure a custom range for SSM Listeners.

   The default SSM range is 232.0.0.0/8, which is automatically configured on the switches in your fabric. If you are using SSM, we recommend configuring your Listeners to join groups in this range, in which case you can skip this step.

   If for any reason you do not want to change your listener configuration, you can add extra SSM ranges under the VRF settings by creating a route-map with up to 4 extra ranges. Keep in mind that if you add a new range it becomes an SSM range and cannot be used for ASM at the same time.

   Custom SSM range configuration must be done directly in the site's APIC:

   a. Log in to your site's Cisco APIC.
   b. In the top menu, click **Tenants** and select the tenant that contains the VRF.
   c. In the left navigation menu, select **Networking > VRFs > <VRF-name> > Multicast**.
   d. In the main pane, choose the **Pattern Policy** tab.
   e. From the **Route Map** drop-down in the **Source Specific Multicast (SSM)** area, choose an existing route map or click **Create Route Map Policy for Multicast** option to create a new one.

      If you select an existing route map, click the icon next to the drop-down to view the route

map's details.

In the route map details window or the **Create Route Map Policy for Multicast** window that opens, click **+** to add an entry. Then configure the Group IP; you must provide only the group IP address to define the new range.

6. (Optional) Enable PIM on the site's L3Out.

   If you connect multicast sources or receivers to the external network domain, you must also enable PIM on the site's L3Out. L3Out configuration currently cannot be done from the Cisco Nexus Dashboard Orchestrator, so you must ensure that PIM is enabled directly in the site's APIC. Additional information about PIM configuration on an L3Out is available in the *Cisco APIC Layer 3 Networking Configuration Guide*.

   a. Log in to your site's Cisco APIC.
   b. In the top menu, click **Tenants** and select the tenant that contains the L3Out.
   c. In the left navigation menu, select **Networking > L3Outs > <l3out-name>**.
   d. In the main pane, choose the **Policy** tab.
   e. Check the **PIM** options.

      Multi-Site supports IPv4 multicast only.

7. Enable Layer 3 multicast on a BD.

   When you have enabled L3 Multicast on a VRF, you can enable L3 multicast on a Bridge Domain (BD) level.

   a. Select the BD for which you want to enable Layer 3 multicast.
   b. In the right properties sidebar, check the **L3 Multicast** check box.

8. Enabled IGMPv3 interface policy on the bridge domains where receivers are connected.

   Because you are configuring SSM, you must also assign an IGMPv3 interface policy to the BD. By default, when PIM is enabled, IGMP is also automatically enabled on the SVI but the default version is set to IGMPv2. You must explicitly set the IGMP interface policy to IGMPv3. This must be done at the site-local level:

   a. Log in to your site's Cisco APIC.
   b. In the top menu, click **Tenants** and select the tenant that contains the BD.
   c. In the left navigation menu, select **Networking > Bridge Domains > <BD-name>**.
   d. In the main pane, choose the **Policy** tab.
   e. From the **IGMP Policy** drop-down, select the IGMP policy or click **Create IGMP Interface Policy** to create a new one.

      If you select an existing policy, click the icon next to the drop-down to view the policy details.

      In the policy details window or the **Create Route Map Policy for Multicast** window that opens, ensure that the **Version** field is set to Version 3.

9. (Optional) If you want to configure multicast filtering, provide the route-maps for source and

destination filtering.

a. Select the BD.

b. In the right properties sidebar, select a **Route-Map Source Filter** and **Route-Map Destination Filter**.

   You can choose to enable either the multicast source filtering, or the receiver filtering, or both.

   Keep in mind, if you do not select a route map, the default action is to allow all multicast traffic on the bridge domain; however, if you select a route map the default action changes to deny any traffic that is not explicitly matched to a filter entry in the route map.

10. If your multicast source is in one site and is not stretched to the other sites, enable intersite multicast source option on the EPG.

    When you have enabled L3 Multicast on the BD, you must also enable multicast on the EPGs (part of multicast-enabled BDs) where multicast sources are connected.

    a. Select the EPG for which you want to enable Layer 3 multicast.

    b. In the right sidebar, check the **Intersite Multicast Source** check box.

First Published: 2024-03-01
Last Modified: 2024-03-01

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
http://www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883