



# Floating L3Out Consideration and Restrictions

---

- [Summary of Floating L3Out Deployment Considerations and Restrictions, on page 1](#)

## Summary of Floating L3Out Deployment Considerations and Restrictions

The following list summarizes some of the requirements and limitations of the floating Layer 3 outside network connection (L3Out) feature.

- Floating L3Out with VMware vDS VMM domain considerations include the following:
  - When floating L3Out is deployed, the virtual router must be connected to the port-group created by the VMM domain.
  - A floating L3Out Switch Virtual Interface (SVI) is not programmed on a leaf node if at least a virtual router connected to the created port-group does not exist on one of the locally connected hypervisors.
  - Once a virtual router moves and becomes attached to the port group, the leaf node has SVI programmed. This is the same as a regular endpoint group (EPG) in a Virtual Machine Manager (VMM) domain with on-demand resolution immediacy, and deployment immediacy is immediate.
  - Secondary floating IP with a VMM domain is supported only with Cisco Application Policy Infrastructure Controller (APIC) Release 5.0(1). Secondary address configurations must be removed before downgrading to an earlier release.
- Floating L3Out with physical domains requires Cisco APIC release 5.0(1) or later.
- The functionality to avoid sub-optimal path requiring Next Hop propagation is supported for intra-VRF traffic only. Inter-VRF traffic where the consumer and the provider are in different VRFs is not supported. This consideration is applied to both EPG to external EPG and external EPG to external EPG contracts.



---

**Note** This functionality is only supported for physical domains.

---

- Cisco APIC supports the deployment of floating L3Out across pods that are part of a Cisco Application Centric Infrastructure (ACI) Multi-Pod fabric, but traffic from ACI internal endpoints to the floating L3Out has the following considerations:

- Avoiding sub-optimal flow (next-hop propagation and multi-path) is not supported with Floating L3Out across pods that are part of an ACI Multi-Pod fabric. This consideration is not applicable if the floating L3Out is not deployed across pods.
  - For example, this is fine because the anchor leaf node and the non-anchor leaf node are in the same pod:
    - Anchor leaf node in pod1.
    - Non-anchor leaf node in pod1.
    - Internal endpoint in any pod.
  - For example, this is NOT supported because the anchor leaf node and the non-anchor leaf node are in the different pods:
    - Anchor leaf node in pod1
    - Non-anchor leaf node in pod2.
    - Internal endpoint in any pod.
- Cisco APIC does not support floating L3Out with remote leaf switches even if remote leaf direct is enabled. A remote leaf switch can't be deployed as an anchor leaf or a non-anchor leaf for floating L3Out.
- A non-anchor node should not be configured under a logical node profile.
- Per port VLAN feature is not supported on ports that are part of floating L3Outs.
- If a virtual port channel (vPC) interface is used for external router connectivity and one leaf switch in the vPC pair is an anchor leaf node, the other leaf switch in the same vPC pair must also be an anchor leaf node. A mixture of anchor and non-anchor leaf nodes in the same vPC pair for the same VLAN encapsulation is not supported.
- For a static route, the external router must use the primary or secondary IP address of the anchor leaf switch as the next hop.
- Make sure that the primary IP address and floating primary IP address are part of the same subnet.
- In some situations, you might not be able to achieve a desired result using a single L3Out and you will have to configure two different L3Outs to achieve that result instead (for example, if you want to have both BGP and OSPF to redistribute the routes into the fabric).

For example, consider a situation where you want to have the following configuration:

- An eBGP session that will be established between a border leaf switch's loopback IP address and an external router's loopback IP address
- OSPF configured to exchange the route for the loopback IP address for the border leaf switch and the external router
- OSPF also configured to redistribute additional routes learned from the external nodes into the Cisco ACI fabric

You will not be able to have that configuration using a single L3Out because, in order to get that configuration, you will have to use both OSPF and BGP to redistribute the routes into the fabric, which cannot be done on a single L3Out due to an existing L3Out limitation.

- Be sure to choose the recommended Bidirectional Forwarding Detection (BFD) timers while on the floating L3Out. You can expect traffic loss of 1 to 2 seconds when the virtual router moves from one host to another host within the cluster. We recommend BFD TX/RX intervals of 700 milliseconds or higher.
- A floating L3Out SVI and a non-floating L3Out SVI can exist on the same leaf switch with the same VLAN encapsulation as long as they use the same primary IP address.
- Floating L3Out requires generation-2 leaf switches. Generation-1 switches cannot be configured as an anchor or non-anchor switch. However, you can use a generation-1 switch as a non-border leaf or a compute leaf switch.
- If ARP entry for an external device's IP is manually cleared on the non-anchor leaf node, internal to external traffic path will be sub-optimal even if direct host advertisement is enabled to avoid sub-optimal path. It's because the leaf node stops advertising the external device's IP, whereas the anchor leaf nodes still have the ARP entry which still have the traffic forwarded. To avoid sub-optimal flow, it's recommended to clear the ARP entry on all anchor and non-anchor leaf nodes which will recreate the ARP entry on the leaf nodes where the external IP is connected. So that internal to external traffic path will be back to optimal forwarding. This consideration is applicable to IPv6 adjacency too.
- For Next Hop Propagation, the floating L3Out must be in a physical domain, not in a VMM domain.
- Prior to Cisco Nexus Dashboard Orchestrator (NDO) release 4.1(1), as with other logical interfaces under L3Out, the floating SVI configuration is at the Cisco APIC level and not at the NDO level. However, Cisco NDO can refer an L3Out that contains floating SVI.

Starting from NDO release 4.1(1), the full L3Out configuration, including floating L3Out, is available on NDO. However, the floating L3Out functionalities (such as next-hop propagation, etc.) is available within each fabric.

- When running ESXi on UCS B Series blade switches behind a fabric interconnect, we recommend that you leave "Fabric Failover" disabled and allow the DVS running on ESXi itself to achieve redundancy in the event of a failure. If enabled, the LLDP/CDP packets that Cisco ACI uses for deployment will be seen on the active and standby virtual switch ports (vEths), which could cause constant flapping and deployment issues.
- The following rules apply for a floating L3Out with IPv4 and IPv6 address family:
  - For IPv4 and IPv6 address family for the same L3Out on the same leaf nodes, you need different L3Out logical interface profiles.
  - Physical domain: for the same VLAN encapsulation, you must use the same anchor leaf nodes for IPv4 and IPv6 address family. A different set of leaf nodes cannot be the anchor for an IPv4 and IPv6 address family.
  - VMM domain: Prior to Cisco APIC Release 5.2(4), you must use different VLAN encapsulations for IPv4 and IPv6 address families. Different port groups are created for IPv4 and IPv6 floating SVI interfaces.

Beginning with Cisco APIC release 5.2(4), same encapsulation can be used. One port group is created for IPv4 and IPv6 address families for the same L3Out with a VMM domain. While deploying a floating SVI, if both address families are configured under the L3Out, both floating SVIs for IPv4 and IPv6 are deployed on the leaf nodes.

- If you downgrade from Cisco APIC release 5.2(4) to a previous release, VMM domain dynamic attachments under a floating L3Out with the same SVI encapsulation but different address-families are deleted.

- If you upgrade from pre-5.2(4) to 5.2(4), and want to migrate from different encapsulation for IPv4 or IPv6 to same encapsulation for IPv4 or IPv6, you need to delete the existing logical interface profile, and add it using the same VLAN encapsulation for both IPv4 or IPv6 address families.

- As part of the support for avoiding suboptimal traffic from a Cisco ACI internal endpoint to a floating L3Out that was introduced in release 5.0(1), support was also available for learning the directly attached host routes (external router's IPs in the L3Out SVI subnet) on the anchor and non-anchor leaf switches and redistributing them into the ACI fabric, as described in [Avoiding Suboptimal Traffic From a Cisco ACI Internal Endpoint to a Floating L3Out](#) and [Configuring Direct-Attached Host Route Advertising on L3Out](#).

Prior to release 5.2(1), these attached host routes could be advertised out of the ACI fabric if you configured the export rules to explicitly advertise the attached host routes out of the ACI fabric.

Beginning with release 5.2(1), this behavior has changed and attached host routes are implicitly denied from going out of the ACI fabric.

- The following considerations and restrictions apply to the multi-protocol recursive next hop propagation feature, introduced in release 5.2(1):
  - As of Cisco ACI release 6.0(1), in the case of ACI floating L3Out with next-hop propagation enabled, only one next-hop for a prefix learned via BGP is used for forwarding. If ECMP is needed for the prefix, it's achievable by using ECMP to the next-hop below. This consideration is applicable to both IPv4 and IPv6.
    - BGP sends one primary next-hop for the prefix. (For example, 10.1.0.0/16 via 1.1.1.1).
    - Anchor leaf node redistributes multi-path for the next-hop (For example, 1.1.1.1 via 172.16.1.1 and 172.16.1.2).
  - When next-hop propagation is enabled, it's not recommended to have multiple external routers for control nodes redundancy advertising the same external prefix with the same ECMP paths because of CSCwd28918.
    - For example, this is fine because it uses different next-hop IPs:
      - external router1, that advertises 10.1.0.0/16 via 172.16.1.1 via BGP.
      - external router2, that advertises 10.1.0.0/16 via 172.16.1.2 via BGP.
    - For example, this is NOT recommended because it has only one next-hop IP, not ECMP:
      - external router1, that advertises 10.1.0.0/16 via 1.1.1.1 via BGP.
      - external router2, that advertises 10.1.0.0/16 via 1.1.1.1 via BGP.
  - IPv6 link local addresses are not supported when configuring next hop propagation for recursive route. You must use global addresses for IPv6 in this situation.
  - Redistribution into a Not-So-Stubby Area (NSSA) creates a special type of link-state advertisement (LSA) known as type 7, which can only exist in an NSSA area. Routes should be learned as this type 7 LSA to have next hops as the global address, which means that the L3Out for the forwarding node, with only the OSPF protocol enabled (**l3out-ospf**), must use the **NSSA area** option in the **OSPF Area Type** field.

- If multiple next hops are used for a static route, the route might not be optimal if one of the next hops is down. See [CSCvy10946](#) for more details.
- The following limitations exist when configuring the multi-protocol recursive next hop propagation feature through the CLI:
  - The next-hop unchanged route-map is not supported for BGP peers.
  - The route-profile template doesn't support match rules.

