Configuring vPC Fabric Peering

This chapter contains these sections:

Information About vPC Fabric Peering

vPC Fabric Peering provides an enhanced dual-homing access solution without the overhead of wasting physical ports for vPC Peer Link.

The following lists the vPC Fabric Peering solution:

- vPC Fabric Peering port-channel with virtual members (tunnels).
- vPC Fabric Peering (tunnel) with removal of the physical peer link requirement.
- vPC Fabric Peering up/down events are triggered based on route updates and fabric up/down.
- Uplink tracking for extended failure coverage.
- vPC Fabric Peering reachability via the routed network, such as the spine.
- Increased resiliency of the vPC control plane over TCP-IP (CFSoIP).
- Data plane traffic over the VXLAN tunnel.
- Communication between vPC member switches uses VXLAN encapsulation.
- Failure of all uplinks on a node result in vPC ports going down on that switch. In that scenario, vPC peer takes up the primary role and forwards the traffic.
- Uplink tracking with state dependency and up/down signalization for vPCs.
- Positive uplink state tracking drives vPC primary role election.
- For border leafs and spines, there is no need for per-VRF peering since network communication uses the fabric.
- Enhance forwarding to orphans hosts by extending the VIP/PIP feature to Type-2 routes.

- Infra-VLAN is not required for vPC fabric peering.

Note The vPC Fabric Peering counts as three VTEPs unlike a normal vPC which counts as one VTEP.

Guidelines and Limitations for vPC Fabric Peering

The following are the vPC Fabric Peering guidelines and limitations:

Configuration recommendations

- vPC Fabric Peering requires TCAM carving of the region **ing-flow-redirect**. TCAM carving requires saving the configuration and reloading the switch prior to using the feature.

Note This requirement only applies to Cisco Nexus 9300-EX, 9300-FX, 9300-FX2, 9300-FX3, and 9364C platform switches.

- When using vPC fabric peering, you cannot create routing over SVIs for such vPC pairs.
- Prior to reconfiguring the vPC Fabric Peering source and destination IP, the vPC domain must be shut down. Once the vPC Fabric Peering source and destination IP have been adjusted, the vPC domain can be enabled (**no shutdown**).
- The source and destination IP supported in **virtual peer-link destination** command are class A, B, and C. Class D and E are not supported for vPC Fabric Peering.
- Immediately after converting from fabric peering to a physical peer link, make the following changes on both peers:
  1. Globally configure a TCAM region using the **hardware access-list tcam region ing-flow-redirect 0** command.
  2. Optionally, allocate the free space to other classes. For more information, see Understand How to Carve Nexus 9000 TCAM Space.
  3. Save the running configuration using the **copy running-config startup-config** command.
  4. Reload the switch.
- The vPC Fabric Peering peer-link is established over the transport network (the spine layer of the fabric). As communication between vPC peers occurs in this manner, control plane information CFS messages used to synchronize port state information, VLAN information, VLAN-to-VNI mapping, host MAC addresses are transmitted over the fabric. CFS messages are marked with the appropriate DSCP value, which should be protected in the transport network. The following example shows a sample QoS configuration on the spine layer of Cisco Nexus 9000 Series switches.

  Classify traffic by matching the DSCP value (DSCP 56 is the default value):

```
class-map type qos match-all CFS
  match dscp 56
```

Set traffic to the qos-group that corresponds with the strict priority queue for the appropriate spine switch. In this example, the switch sends traffic to qos-group 7, which corresponds to the strict priority queue (Queue 7). Note that different Cisco Nexus platforms might have a different queuing structure.

```
policy-map type qos CFS
  class CFS
    Set qos-group 7
```

Assign a classification service policy to all interfaces toward the VTEP (the leaf layer of the network):

```
interface Ethernet 1/1
  service-policy type qos input CFS
```

- Layer 3 Tenant Routed Multicast (TRM) is supported. Layer 2/Layer 3 TRM (Mixed Mode) is not supported.
- If Type-5 routes are used with this feature, the **advertise-pip** command is a mandatory configuration.
- Enhance forwarding to orphan hosts by extending the VIP/PIP feature to Type-2 routes.
- An orphan Type-2 host is advertised using PIP. A vPC Type-2 host is advertised using VIP. This is the default behavior for a Type-2 host.

  To advertise an orphan Type-5 route using PIP, you need to advertise PIP under BGP.
- For orphan ports, it is highly recommended to configure **vpc orphan-port suspend** command on both vPC nodes, to avoid traffic disruption during NVE failure scenarios.
- Traffic from remote VTEP to orphan hosts would land on the actual node which has the orphans. Bouncing of the traffic is avoided.

Note When the vPC leg is down, vPC hosts are still advertised with the VIP IP.

- ARP behavior differs between vPC fabric peering and physical peer links for orphan ports. vPC fabric peering does not sync ARP entries as it does with physical peer links. An orphan Type-2 host is advertised using the PIP in vPC Fabric peering.

Supported feature, release and platforms

- Cisco Nexus 9332C, 9364C, and 9300-EX/FX/FXP/FX2/FX3/GX/GX2 platform switches support vPC Fabric Peering. Cisco Nexus 9200 and 9500 platform switches do not support vPC Fabric Peering.

Note For Cisco Nexus 9300-EX switches, mixed-mode multicast and ingress replication are not supported. VNIs must be configured with either multicast or IR underlay, but not both.

- Beginning with Cisco NX-OS Release 10.1(1), FEX Support is provided with vMCT for IPv4 underlay on Cisco Nexus 9300-EX/FX/FX2/FX3 platform switches.

- Beginning with Cisco NX-OS Release 10.2(2)F, FEX Support is provided with vMCT for IPv4 underlay on Cisco Nexus 9300-GX platform switches.
- Beginning with Cisco NX-OS Release 10.1(1), vPC Fabric Peering supports FEX in Straight Through and Active-Active (dual home) modes in N9K-C9336C-FX2-E, N9K-C93108TC-EX, N9K-C93108TC-FX,N9K-C93180YC-EX, N9K-C93180YC-FX, N9K-C93216TC-FX2, N9K-C93240YC-FX2, N9K-C93360YC-FX2, N9K-C9336C-FX2, N9K-C93180YC-FX3, N9K-C93180YC-FX3S platform switches.

  Refer to *Cisco Nexus 2000 Series NX-OS Fabric Extender Configuration Guide for Cisco Nexus 9000 Series Switches* for details on FEX (Straight Through and Active-Active modes).
- Beginning with Cisco NX-OS Release 10.2(3)F, ND-ISSU and LXC-ISSU are supported with vMCT for IPv4 underlay on Cisco Nexus 9300-EX/FX/FXP/FX2/FX3/GX/GX2 ToR switches.

Unsupported features

- The vPC Fabric Peering domain is not supported in the role of a Multi-Site vPC BGW.
- VTEPs behind vPC ports are not supported. This means that virtual peer-link peers cannot act as a transit node for the VTEPs behind the vPC ports.
- SVI and sub-interface uplinks are not supported.

Configuring vPC Fabric Peering

Ensure the vPC Fabric Peering DSCP value is consistent on both vPC member switches. Ensure that the corresponding QoS policy matches the vPC Fabric Peering DSCP marking.

All VLANs that require communication traversing the vPC Fabric Peering must have a VXLAN enabled (vn-segment); this includes the native VLAN.

Note For MSTP, VLAN 1 must be extended across vPC Fabric Peering if the peer-link and vPC legs have the default native VLAN configuration. This behavior can be achieved by extending VLAN 1 over VXLAN (vn-segment). If the peer-link and vPC legs have non-default native VLANs, those VLANs must be extended across vPC Fabric Peering by associating the VLANs with VXLAN (vn-segment).

Use the **show vpc virtual-peerlink vlan consistency** command for verification of the existing VLAN-to-VXLAN mapping used for vPC Fabric Peering.

peer-keepalive command for vPC Fabric Peering is supported with one of the following configurations:

- Management interface
- Dedicated Layer 3 link in default or non-default VRF
- Loopback interface reachable using the spine.

Configuring Features

Example uses OSPF as the underlay routing protocol.

```
configure terminal
nv overlay evpn
feature ospf
feature bgp
feature pim
feature interface-vlan
feature vn-segment-vlan-based
feature vpc

feature nv overlay
```

vPC Configuration

> **Note** To change the vPC Fabric Peering source or destination IP, the vPC domain must be shutdown prior to modification. The vPC domain can be returned to operation after the modifying by using the **no shutdown** command.

Configuring TCAM Carving

```
hardware access-list tcam region ing-racl 0
hardware access-list tcam region ing-sup 768
hardware access-list tcam region ing-flow-redirect 512
```

Configuring the vPC Domain

```
vpc domain 100
peer-keepalive destination 192.0.2.1
virtual peer-link destination 192.0.2.100 source 192.0.2.20/32 [dscp <dscp-value>]
Warning: Appropriate TCAM carving must be configured for virtual peer-link vPC
peer-switch
peer-gateway
ip arp synchronize
ipv6 nd synchronize
exit
```

> **Note** The **dscp** keyword in optional. Range is 1 to 63. The default value is 56.

Configuring vPC Fabric Peering Port Channel

No need to configure members for the following port channel.

```
interface port-channel 10
switchport
switchport mode trunk
vpc peer-link
```

`interface loopback0`

> **Note** This loopback is not the NVE source-interface loopback (interface used for the VTEP IP address).

```
interface loopback 0
ip address 192.0.2.20/32
ip router ospf 1 area 0.0.0.0
```

Note You can use the loopback for BGP peering or a dedicated loopback. This lookback must be different that the loopback for peer keep alive.

Configuring the Underlay Interfaces

Both L3 physical and L3 port channels are supported. SVI and sub-interfaces are not supported.

```
router ospf 1
interface Ethernet1/16
port-type fabric
ip address 192.0.2.2/24
ip router ospf 1 area 0.0.0.0
no shutdown
interface Ethernet1/17
port-type fabric
ip address 192.0.2.3/24
ip router ospf 1 area 0.0.0.0
no shutdown
interface Ethernet1/40
port-type fabric
ip address 192.0.2.4/24
ip router ospf 1 area 0.0.0.0
no shutdown
interface Ethernet1/41
port-type fabric
ip address 192.0.2.5/24
ip router ospf 1 area 0.0.0.0
no shutdown
```

Note All ports connected to spines must be port-type fabric.

VXLAN Configuration

Note Configuring **advertise virtual-rmac** (NVE) and **advertise-pip** (BGP) are required steps. For more information, see the Configuring vPC Multi-Homing chapter.

Configuring VLANs and SVI

```
vlan 10
vn-segment 10010
vlan 101
vn-segment 10101

interface Vlan101
no shutdown
mtu 9216
vrf member vxlan-10101
no ip redirects
ip forward
ipv6 address use-link-local-only
no ipv6 redirects
interface vlan10
```

```
no shutdown
mtu 9216
vrf member vxlan-10101
no ip redirects
ip address 192.0.2.102/24
ipv6 address 2001:DB8:0:1::1/64
no ipv6 redirects
fabric forwarding mode anycast-gateway
```

Configuring Virtual Port Channel

```
interface Ethernet1/3
switchport
switchport mode trunk
channel-group 100
no shutdown
exit
interface Ethernet1/39
switchport
switchport mode trunk
channel-group 101
no shutdown
interface Ethernet1/46
switchport
switchport mode trunk
channel-group 102
no shutdown
interface port-channel100
vpc 100
interface port-channel101
vpc 101
interface port-channel102
vpc 102
exit
```

Migrating from vPC to vPC Fabric Peering

This procedure contains the steps to migration from a regular vPC to vPC Fabric Peering.

Any direct Layer 3 link between vPC peers should be used only for peer-keep alive. This link should not be used to advertise paths for vPC Fabric Peering loopbacks.

Note This migration is disruptive.

Before you begin

We recommend that you shut all physical Layer 2 links between the vPC peers before migration. We also recommend that you map VLANs with vn-segment before or after migration.

SUMMARY STEPS

1. **configure terminal**
2. **show vpc**
3. **show port-channel summary**

4. **interface ethernet** *slot/port*
5. **no channel-group**
6. Repeat steps 4 and 5 for each interface.
7. **show running-config vpc**
8. **vpc domain** *domain-id*
9. **virtual peer-link destination** *dest-ip* **source** *source-ip*
10. **interface {ethernet | port-channel}** *value*
11. **port-type fabric**
12. (Optional) **show vpc fabric-ports**
13. **hardware access-list tcam region ing-flow-redirect** *tcam-size*
14. **copy running-config startup-config**
15. **reload**

DETAILED STEPS

Procedure

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `switch# configure terminal`	Enters global configuration mode.
Step 2	**show vpc** **Example:** `switch(config)# show vpc`	Determine the number of members in the port channel.
Step 3	**show port-channel summary** **Example:** `switch(config)# show port-channel summary`	Determine the number of members.
Step 4	**interface ethernet** *slot/port* **Example:** `switch(config)# interface ethernet 1/4`	Specifies the interface you are configuring. **Note** This is the peer link port channel.
Step 5	**no channel-group** **Example:** `switch(config-if)# no channel-group`	Remove vPC peer-link port-channel members. **Note** Disruption occurs following this step.
Step 6	Repeat steps 4 and 5 for each interface. **Example:**	
Step 7	**show running-config vpc** **Example:** `switch(config-if)# show running-config vpc`	Determine the vPC domain.

	Command or Action	Purpose
Step 8	**vpc domain** *domain-id* **Example:** `switch(config-if)# vpc domain 100`	Enter vPC domain configuration mode.
Step 9	**virtual peer-link destination** *dest-ip* **source** *source-ip* **Example:** `switch(config-vpc-domain)# virtual peer-link destination 192.0.2.1 source 192.0.2.100`	Specify the destination and source IP addresses for vPC fabric peering.
Step 10	**interface {ethernet | port-channel}** *value* **Example:** `switch(config-if)# interface Ethernet1/17`	Specifies the L3 underlay interface you are configuring.
Step 11	**port-type fabric** **Example:** `switch(config-if)# port-type fabric`	Configures port-type fabric for underlay interface. **Note** All ports connected to spines must be port-type fabric.
Step 12	(Optional) **show vpc fabric-ports** **Example:** `switch# show vpc fabric-ports`	Displays the fabric ports connected to spine.
Step 13	**hardware access-list tcam region ing-flow-redirect** *tcam-size* **Example:** `switch(config-vpc-domain)# hardware access-list tcam region ing-flow-redirect 512`	Perform TCAM carving. The minimum size for Ingress-Flow-redirect TCAM region size is 512. Also ensure it is configured in multiples of 512.
Step 14	**copy running-config startup-config** **Example:** `switch(config-vpc-domain)# copy running-config startup-config`	Copies the running configuration to the startup configuration.
Step 15	**reload** **Example:** `switch(config-vpc-domain)# reload`	Reboots the switch.

Verifying vPC Fabric Peering Configuration

To display the status for the vPC Fabric Peering configuration, enter one of the following commands:

Table 1: vPC Fabric Peering Verification Commands

Command	Purpose
show vpc fabric-ports	Displays the fabric ports state.
show vpc	Displays information about vPC Fabric Peering mode.
show vpc virtual-peerlink vlan consistency	Displays the VLANs which are not associated with vn-segment.

Example of the show vpc fabric-ports Command

```
switch# show vpc fabric-ports
Number of Fabric port : 9
Number of Fabric port active : 9

Fabric Ports State
-------------------------------------
Ethernet1/9 UP
Ethernet1/19/1 ( port-channel151 ) UP
Ethernet1/19/2 ( port-channel151 ) UP
Ethernet1/19/3 UP
Ethernet1/19/4 UP
Ethernet1/20/1 UP
Ethernet1/20/2 ( port-channel152 ) UP
Ethernet1/20/3 ( port-channel152 ) UP
Ethernet1/20/4 ( port-channel152 ) UP
```

Example of the show vpc Command

```
switch# show vpc
Legend:
                 (*) - local vPC is down, forwarding via vPC peer-link

vPC domain id                     : 3
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status  : success
Per-vlan consistency status       : success
Type-2 consistency status         : success
vPC role                          : primary
Number of vPCs configured         : 1
Peer Gateway                      : Enabled
Dual-active excluded VLANs        : -
Graceful Consistency Check        : Enabled
Auto-recovery status              : Enabled, timer is off.(timeout = 240s)
Delay-restore status              : Timer is off.(timeout = 30s)
Delay-restore SVI status          : Timer is off.(timeout = 10s)
Operational Layer3 Peer-router    : Disabled
Virtual-peerlink mode             : Enabled

vPC Peer-link status
---------------------------------------------------------------------
id    Port   Status Active vlans
--    ----   ------ -------------------------------------------------
1     Po100  up     1,56,98-600,1001-3401,3500-3525

vPC status
----------------------------------------------------------------------------
```

```
Id    Port          Status Consistency Reason                Active vlans
--    ------------  ------ ----------- ------                ---------------
101   Po101         up     success     success               98-99,1001-280
                                                             0

Please check "show vpc consistency-parameters vpc <vpc-num>" for the
consistency reason of down vpc and for type-2 consistency reasons for
any vpc.

ToR_B1#
```

Example of the show vpc virtual-peerlink vlan consistency Command

```
switch# show vpc virtual-peerlink vlan consistency
Following vlans are inconsistent
23
switch#
```