CHAPTER 1

Virtual Port Channel Operations

This chapter describes the best practices and operational procedures for the virtual port channel (vPC) feature on Cisco Nexus 5000 Series switches that run Cisco NX-OS Release 5.0(2)N2(1) and earlier releases.

This chapter includes the following sections:

Information About vPC Operations

A vPC allows links that are physically connected to two different Cisco Nexus 5000 Series switches to appear as a single port channel to a third switch. The third switch can be a Cisco Nexus 2000 Series Fabric Extender or a switch, server, or any other networking device. A vPC can provide Layer 2 multipath capability which allows you to create redundancy by increasing bandwidth, enabling multiple parallel paths between nodes, and load-balancing traffic where alternative paths exist.

For a quick overview of vPC configurations, see the *Virtual PortChannel Quick Configuration Guide* at the following URL:
http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9670/configuration_guide_c07-543563.html

vPC Consistency Checks

This section includes the following topics:

- Bypassing a vPC Consistency Check When a Peer Link is Lost, page 1-8

Type 1 and Type 2 Consistency Check Parameters

Before a Cisco Nexus 5000 Series switch brings up a vPC, the two Cisco Nexus 5000 Series switches in the same vPC domain exchange configuration information to verify if both switches have compatible configurations for a vPC topology. Depending on the severity of the impact of possible mismatched configurations, some configuration parameters are considered as Type 1 consistency check parameters while others are considered as Type 2.

When a mismatch in Type 1 parameters occur, the following applies:

- If a graceful consistency check is enabled (default), the primary switch keeps the vPC up while the secondary switch brings it down
- If a graceful consistency check is disabled, both peer switches suspend VLANs on the vPC ports.

Note The graceful consistency check is a new feature introduced in Cisco NX-OS Release 5.0(2)N2(1) and is enabled by default. For more details, see the "Graceful Consistency Check" section on page 1-3.

When Type 2 parameters exist, a configuration mismatch generates a warning syslog message. The vPC on the Cisco Nexus 5000 Series switch remains up and running. The global configuration, such as Spanning Tree Protocol (STP), and the interface-level configurations are included in the consistency check.

The **show vpc consistency-parameters global** command lists all global consistency check parameters. Beginning with Cisco NX-OS Release 5.0(2)N1(1), QoS parameters have been downgraded from Type 1 to Type 2.

This example shows how to display all global consistency check parameters:

```
switch# show vpc consistency-parameters global
Legend:
        Type 1 : vPC will be suspended in case of mismatch
Name                        Type  Local Value            Peer Value
-------------               ----  ---------------------- -----------------------
QoS                         2     ([], [3], [], [], [],  ([], [3], [], [], [],
                                  [])                    [])
Network QoS (MTU)           2     (1538, 2240, 0, 0, 0,  (1538, 2240, 0, 0, 0,
                                  0)                     0)
Network Qos (Pause)         2     (T, F, F, F, F, F)     (T, F, F, F, F, F)
Input Queuing (Bandwidth)   2     (50, 50, 0, 0, 0, 0)   (50, 50, 0, 0, 0, 0)
Input Queuing (Absolute     2     (F, F, F, F, F, F)     (F, F, F, F, F, F)
Priority)
Output Queuing (Bandwidth)  2     (50, 50, 0, 0, 0, 0)   (50, 50, 0, 0, 0, 0)
Output Queuing (Absolute    2     (F, F, F, F, F, F)     (F, F, F, F, F, F)
Priority)
STP Mode                    1     MST                    MST
STP Disabled                1     None                   None
STP MST Region Name         1     ""                     ""
STP MST Region Revision     1     0                      0
STP MST Region Instance to  1
 VLAN Mapping
STP Loopguard               1     Disabled               Disabled
STP Bridge Assurance        1     Enabled                Enabled
STP Port Type, Edge         1     Normal, Enabled,       Normal, Enabled,
BPDUFilter, Edge BPDUGuard        Disabled               Disabled
STP MST Simulate PVST       1     Enabled                Enabled
Allowed VLANs               -     1,10,100-101,200-201   1,10,100-101,200-201,2
```

```
                                                        000
Local suspended VLANs       -     -                     -
```

Use the **show vpc consistency-parameters interface port-channel** *number* command to display the interface-level consistency parameters.

This example shows how to display the interface-level consistency parameters:

```
n5k-1# show vpc consistency-parameters interface port-channel 200

    Legend:
        Type 1 : vPC will be suspended in case of mismatch

Name                        Type  Local Value            Peer Value
-------------               ----  ---------------------- -----------------------
STP Port Type               1     Default                Default
STP Port Guard              1     None                   None
STP MST Simulate PVST       1     Default                Default
lag-id                      1     [(7f9b,                [(7f9b,
                                  0-23-4-ee-be-64, 80c8, 0-23-4-ee-be-64, 80c8,
                                   0, 0), (8000,          0, 0), (8000,
                                  0-1e-13-15-7-40, 1, 0, 0-1e-13-15-7-40, 1, 0,
                                   0)]                    0)]
mode                        1     active                 active
Speed                       1     10 Gb/s                10 Gb/s
Duplex                      1     full                   full
Port Mode                   1     trunk                  trunk
Native Vlan                 1     1                      1
Shut Lan                    1     No                     No
Allowed VLANs               -     1-999,1001-3967,4048-4 1-3967,4048-4093
                                  093
```

The Cisco Nexus 5000 Series switch conducts vPC consistency checks when it attempts to bring up a vPC or when you make a configuration change.

In the interface consistency parameters shown in the above output, all configurations except the Allowed VLANs are considered as Type 1 consistency check parameters. The Allowed VLAN (under the trunk interface) is considered as a Type 2 consistency check parameter. If the Allowed VLAN ranges are different on both VLANs that means that only common VLANs are active and trunked for the vPC while the remaining VLANs are suspended for this port channel.

Graceful Consistency Check

Beginning with Cisco NX-OS Release 5.0(2)N2(1) and later releases, when a Type 1 mismatch occurs, by default, the primary vPC links are not suspended. Instead, the vPC remains up on the primary switch and the Cisco Nexus 5000 Series switch performs Type 1 configurations without completely disrupting the traffic flow. The secondary switch brings down its vPC until the inconsistency is cleared.

However, in Cisco NX-OS Release 5.0(2)N2(1) and earlier releases, this feature is not enabled for dual-homed FEX ports. When Type-1 mismatches occur in this topology, the VLANs are suspended on both switches. The traffic is disrupted on these ports for the duration of the inconsistency.

To minimize disruption, we recommend that you use the configuration synchronization feature for making configuration changes on these ports.

To enable a graceful consistency check, use the **graceful consistency-check** command. Use the **no** form of this command to disable the feature. The graceful consistency check feature is enabled by default.

This example shows how to enable a graceful consistency check:

```
switch(config)# vpc domain 10
```

```
switch(config-vpc-domain)# [no] graceful consistency-check
```

This example shows that the vPC ports are down on a secondary switch when an STP mode mismatch occurs:

```
switch(config)# show vpc brief
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link
vPC domain id                     : 10
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status: failed
Per-vlan consistency status       : success
Configuration consistency reason: vPC type-1 configuration incompatible - STP
                                  Mode inconsistent
Type-2 consistency status         : success
vPC role                          : secondary
Number of vPCs configured         : 2
Peer Gateway                      : Disabled
Dual-active excluded VLANs        : -
Graceful Consistency Check        : Enabled
vPC Peer-link status
---------------------------------------------------------------------
id   Port   Status Active vlans
--   ----   ------ --------------------------------------------------
1    Po1    up     1-10
vPC status
----------------------------------------------------------------------------
id     Port        Status Consistency Reason                     Active vlans
------ ----------- ------ ----------- -------------------------- ------------
20     Po20        down*  failed      Global compat check failed -
30     Po30        down*  failed      Global compat check failed -
```

Global Mismatch

VLANs suspended on Secondary

237955

This example shows that the vPC ports and the VLANs remain up on the primary switch when an STP mode mismatch occurs:

```
switch(config)# sh vpc
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link
vPC domain id                     : 10
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status: failed
Per-vlan consistency status       : success
Configuration consistency reason: vPC type-1 configuration incompatible - STP
                                  Mode inconsistent
Type-2 consistency status         : success
vPC role                          : primary
Number of vPCs configured         : 2
Peer Gateway                      : Disabled
Dual-active excluded VLANs        : -
Graceful Consistency Check        : Enabled
vPC Peer-link status
---------------------------------------------------------------------
id   Port   Status Active vlans
--   ----   ------ --------------------------------------------------
1    Po1    up     1-10
vPC status
----------------------------------------------------------------------------
id     Port        Status Consistency Reason                     Active vlans
------ ----------- ------ ----------- -------------------------- ------------
20     Po20        up     failed      Global compat check failed 1-10
30     Po30        up     failed      Global compat check failed 1-10
```

Global Mismatch

VLANs Up on Primary

237956

This example shows that the vPC ports are down on a secondary switch when an interface-level Type 1 inconsistency occurs:

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

vPC domain id                     : 10
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                          : secondary
Number of vPCs configured         : 2
Peer Gateway                      : Disabled
Dual-active excluded VLANs        : -
Graceful Consistency Check        : Enabled

vPC Peer-link status
---------------------------------------------------------------------
id   Port   Status Active vlans
--   ----   ------ --------------------------------------------------
1    Po1    up     1

vPC status
----------------------------------------------------------------------------
id     Port        Status Consistency Reason                     Active vlans
------ ----------- ------ ----------- -------------------------- -----------
20     Po20        up     success     success                    1

30     Po30        down*  failed      Compatibility check failed -
                                       for port mode
```

VLANs suspended on secondary Interface

237957

This example shows that the vPC ports and the VLANs remain up on the primary switch when an interface-level Type 1 inconsistency occurs:

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

<snip>..

vPC status

id     Port        Status Consistency Reason                     Active vlans
------ ----------- ------ ----------- -------------------------- -----------
20     Po20        up     success     success                    1
30     Po30        up     failed      Compatibility check failed 1
                                       for port mode
```

VLANs Up on Primary Interface

237958

Configuring Per-VLAN Consistency Checks

Beginning with Cisco NX-OS Release 5.0(2)N2(1), the Cisco Nexus 5000 Series switch performs Type-1 consistency checks on a per-VLAN basis when you enable or disable STP on a VLAN. VLANs that do not pass this consistency check are brought down on the primary and secondary switches while other VLANs are not affected.

When you enter the **no spanning-tree vlan** *number* command on one peer switch, only the specified VLAN is suspended on both peer switches; the other VLANs remain up.

Note Per-VLAN consistency checks are not dependent on whether graceful consistency checks are enabled.

This example shows the active VLANs before suspending a specified VLAN:

```
switch(config-if)# show vpc brief
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link
<snip>..
--------------------------------------------------------------------
id   Port   Status Active vlans
--   ----   ------ -----------------------------------------------
1    Po1    up     1-10
vPC status
----------------------------------------------------------------------------
id     Port        Status Consistency Reason                     Active vlans
------ ----------- ------ ----------- -------------------------- -----------
20     Po20        up     success     success                    1-10
30     Po30        up     success     success                    1-10
```

All VLANs are up

237959

This example shows that VLAN 5 is suspended but the remaining VLANs are up:

```
switch(config)# no spanning-tree vlan 5

switch(config)# show vpc brief
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link
<snip>..
--------------------------------------------------------------------
id   Port   Status Active vlans
--   ----   ------ -----------------------------------------------
1    Po1    up     1-4,6-10
vPC status
----------------------------------------------------------------------------
id     Port        Status Consistency Reason                     Active vlans
------ ----------- ------ ----------- -------------------------- -----------
20     Po20        up     success     success                    1-4,6-10
30     Po30        up     success     success                    1-4,6-10
```

VLAN 5 is suspended

237960

Identifying Inconsistent vPC Configurations

The **show vpc** command displays the vPC status and the vPC consistency check result for the global consistency check and the interface-specific consistency check.

This example shows the global vPC consistency check failed because of the mismatched Network QoS configuration:

```
n5k-1# sh vpc
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link

vPC domain id                     : 100
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status: failed
Configuration consistency reason: QoSMgr Network QoS configuration incompatible
vPC role                          : secondary
<snip>…
```

237970

You can use the **show vpc consistency-parameters global** command to identify the configuration difference between two vPC peer switches.

This example shows the global consistency check failed because the STP mode was configured differently on the two vPC switches:

```
switch# show vpc consistency-parameters global

    Legend:
        Type 1 : vPC will be suspended in case of mismatch

Name                        Type  Local Value            Peer Value
-------------               ----  ---------------------- -----------------------
QoS                         2     ([], [3], [], [], [],  ([], [3], [], [], [],
                                  [])                    [])
Network QoS (MTU)           2     (1538, 2240, 0, 0, 0,  (1538, 2240, 0, 0, 0,
                                  0)                     0)
Network Qos (Pause)         2     (F, T, F, F, F, F)     (1538, 2240, 0, 0, 0,
                                                         0)
Input Queuing (Bandwidth)   2     (50, 50, 0, 0, 0, 0)   (50, 50, 0, 0, 0, 0)
Input Queuing (Absolute     2     (F, F, F, F, F, F)     (50, 50, 0, 0, 0, 0)
Priority)
Output Queuing (Bandwidth)  2     (50, 50, 0, 0, 0, 0)   (50, 50, 0, 0, 0, 0)
Output Queuing (Absolute    2     (F, F, F, F, F, F)     (50, 50, 0, 0, 0, 0)
Priority)
STP Mode                    1     MST                    Rapid-PVST
STP Disabled                1     None                   None
STP MST Region Name         1     ""                     ""
STP MST Region Revision     1     0                      0
STP MST Region Instance to  1
 VLAN Mapping
STP Loopguard               1     Disabled               Disabled
STP Bridge Assurance        1     Enabled                Enabled
STP Port Type, Edge         1     Normal, Disabled,      Normal, Disabled,
BPDUFilter, Edge BPDUGuard        Disabled               Disabled
STP MST Simulate PVST       1     Enabled                Enabled
Allowed VLANs               -     1-10                   1-2
Local suspended VLANs       -     3-10                   -
```

STP mode Mismatch

237961

You can use the **show vpc** command also shows the vPC consistency check result for each vPC and the reason for the consistency check failure.

This example shows how to display the vPC consistency check status:

```
n5k-1# show vpc
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link

<snip>..
vPC status
----------------------------------------------------------------------------
id     Port          Status Consistency Reason                     Active vlans
------ ------------  ------ ----------- -------------------------- -----------
104    Po104         up     success     success                    3000
200    Po200         up     success     success                    1,101-110,1
                                                                   000,3000
201    Po201         down*  success     success                    -
1002   Po1002        up     success     success                    102-103
1003   Po1003        up     success     success                    1,101,3000
1004   Po1004        up     success     success                    102-103
103424 Eth102/1/1    up     failed      Compatibility check failed 1000
                                        for port mode
103425 Eth102/1/2    down*  failed      Consistency Check Not      -
                                        Performed
103426 Eth102/1/3    down*  failed      Consistency Check Not      -
                                        Performed
```

Consistency check passed but interface is down

Consistency check failed

Consistency check never conducted since port was down

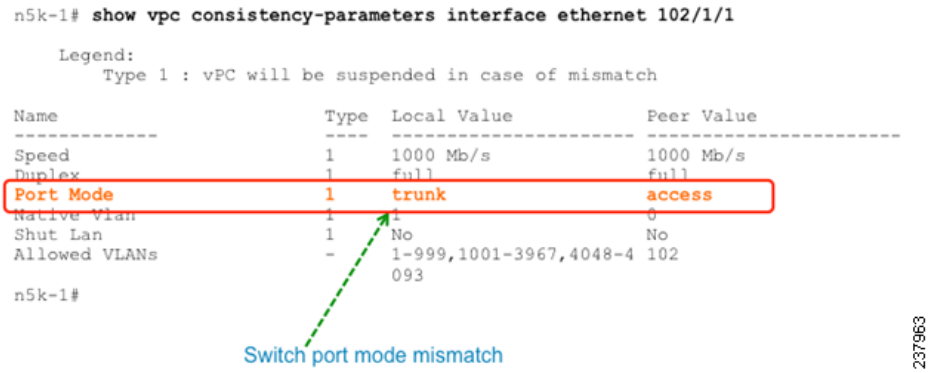
237962

If the consistency check fails, the consistency check is not performed on vPC member ports that are down.

If the consistency check has succeeded and the port is brought down, the consistency check shows that it was successful.

You can use the **show vpc consistency-parameters interface ethernet** *slot/port* command to identify the configuration difference that leads to a consistency check failure for a specified interface or port channel.

This example shows how to display configuration differences that lead to consistency check failures.



Bypassing a vPC Consistency Check When a Peer Link is Lost

The vPC consistency check message is sent by the vPC peer link. The vPC consistency check cannot be performed when the peer link is lost. When the vPC peer link is lost, the operational secondary switch suspends all of its vPC member ports while the vPC member ports remain on the operational primary switch. If the vPC member ports on the primary switch flaps afterwards (for example, when the switch or server that connects to the vPC primary switch is reloaded), the ports remain down due to the vPC consistency check and you cannot add or bring up more vPCs.

Beginning with Cisco NX-OS Release 5.0(2)N2(1), the auto-recovery feature brings up the vPC links when one peer is down. This feature performs two operations:

- If both switches reload, and only one switch boots up, auto-recovery allows that switch to assume the role of the primary switch. The vPC links come up after a configurable period of time if the vPC peer-link and the peer-keepalive fail to become operational within that time. If the peer-link comes up but the peer-keepalive does not come up, both peer switches keep the vPC links down. This feature is similar to the reload restore feature in Cisco NX-OS Release 5.0(2)N1(1) and earlier releases. The reload delay period can range from 240 to 3600 seconds.
- When you disable vPCs on a secondary vPC switch because of a peer-link failure and then the primary vPC switch fails, the secondary switch reenables the vPCs. In this scenario, the vPC waits for three consecutive keepalive failures before recovering the vPC links.

Note The auto-recovery feature in Cisco NX-OS Release 5.0(2)N2(1) and later releases replaces the reload restore feature in Cisco NX-OS Release 5.0(2)N1(1) and earlier releases.

The auto-recovery feature is disabled by default. To enable auto-recovery, enter the **auto-recovery** command in the vPC domain mode.

This example shows how to enable the auto-recovery feature and to set the reload delay period:

```
switch(config)# vpc domain 10
switch(config-vpc-domain)# auto-recovery ?
  <CR>
  reload-delay  Duration to wait after reload to recover vPCs

switch(config-vpc-domain)# auto-recovery reload-delay ?
  <240-3600>  Time-out for restoring vPC links (in seconds)
switch(config-vpc-domain)# auto-recovery reload-delay 240
Warning:
```

```
 Enables restoring of vPCs in a peer-detached state after reload, will wait for 240
seconds (by default) to determine if peer is un-reachable
```

This example shows how to display the status of the auto-recovery feature:

```
switch(config-vpc-domain)# show running-config vpc
!Command: show running-config vpc
!Time: Tue Dec  7 02:38:44 2010

version 5.0(2)N2(1)
feature vpc
vpc domain 10
  peer-keepalive destination 10.193.51.170
  auto-recovery
```

Configuring Changes in vPC Topologies

One of the challenges with vPC topologies is how to make configuration changes with minimum traffic disruption. Due to the consistency check, the configuration made on one vPC switch could potentially lead to consistency check failure and traffic disruption.

Beginning with Cisco NX-OS Release 5.0(2)N2(1), you can use the following procedure to make configuration changes for Type 1 consistency check parameters on a Cisco Nexus 5000 Series switch. We recommend that you perform the following procedure during a maintenance window because it might reduce the vPC bandwidth by half for a short duration.

Note A graceful consistency-check does not apply to dual-homed FEX ports. As a result, both switches keep the port down for the duration of an inconsistency. Using the configuration synchronization feature reduces the duration of the inconsistency.

To make configuration changes for Type 1 consistency-check parameters, follow these steps:

Step 1 Enable graceful consistency-check in a vPC domain.

```
switch# config term
switch(config)# vpc domain 10
switch(config-vpc-domain)# graceful consistency-check
```

Step 2 Enable the configuration synchronization feature on both vPC peer switches.

For details on using the configuration synchronization feature, see the “Configuration Synchronization Operations” chapter.

Step 3 Perform all configuration changes in the switch profile.

```
switch# config sync
switch(config-sync)# switch-profile abc
switch(config-sync-sp)# interface Port-channel 100
switch(config-sync-sp-if)# switchport mode trunk
switch(config-sync-sp-if)# commit
```

When you commit switch profile configurations on the local switch, the configuration is also sent to the vPC peer switch to reduce misconfigurations when changes are made on only one vPC switch and to reduce the downtime because the configuration is applied rapidly. When there is a short mismatch duration, a graceful consistency-check keeps the primary side forwarding traffic.

Note When you are making a configuration change for a Type 2 consistency check parameter, such as Allowed VLAN for trunk ports, you do not need to follow this procedure.

Replacing a Cisco Nexus 5000 Series Switch or Cisco Nexus 2000 Fabric Extender

This section describes how to replace a Cisco Nexus 5000 Series switch or Cisco Nexus 2000 Series Fabric Extender in a vPC topology with minimal disruption.

This section include the following topics:

Replacing a Cisco Nexus 5000 Series Switch

When you replace a Cisco Nexus 5000 Series switch, you must perform the following procedure on the replacement switch to synchronize the configuration with the existing Cisco Nexus 5000 Series switch. The procedure can be done in a hybrid single/dual-homed Fabric Extender vPC topology.

Note Do not connect a peer-link, vPC, or single/dual homed Fabric Extender topology fabric port to the replacement switch.

Note For a vPC+ topology, ensure that you wait for twenty minutes before you replace a vPC+ switch. Otherwise, vPC legs in the primary switch will get suspended due to switch-id conflict.

Before You Begin

- Power up replacement switch with no cables other than mgmt0 and console cable connected to the switch.
- Copy the required Cisco NX-OS kickstart/system files into the switch bootflash.
- If you have a backup of the switch configuration, copy it to new switch bootflash.
- Enable the FEX pre-provisioning feature on the switch in the vPC topology.
- Enable the configuration synchronization feature on the switch and apply all the switch profile configurations except for the sync peer destination IP address.

To replace a Cisco Nexus 5000 Series switch in a vPC topology, follow these steps:

Step 1 Boot the replacement switch.

The new switch comes up without a configuration. Ensure the software version is upgraded to match the existing switch.

Step 2 Enable FEX pre-provisioning for all single or dual homed Fabric Extender modules on the replacement switch.

Note Ensure that you unconfigure the **system default switchport shutdown** command on the replacement switch. Otherwise, when Fabric Extender Modules are coming online on the replacement switch, dual-homed FEX ports on the primary switch will flap causing traffic disruption.

Step 3 Configure the replacement switch as follows:

Note Before you configure the replacement switch using any of the following method, disable the vPC auto-recovery feature on both the vPC peers using the **no auto-recovery** command under the vPC domain. This is to ensure that there is no vPC role change because of the sticky bit feature, when the replacement switch is brought up. vPC auto-recovery feature is enabled by default in Cisco NX-OS release 7.x and later.

- If the running configuration was saved offline, go to Step 4 to Step 10 to apply the configuration.

- If the running configuration was not saved offline, you can obtain it from the peer switch if the configuration synchronization feature is enabled. (Create a switch profile and then go to Step 11).
- If neither condition is met, manually add the configuration and then go to Step 11.

Step 4 Edit the configuration file to remove the sync-peer command if using the configuration synchronization feature.

Step 5 Configure the mgmt0 port IP address and download the configuration file.

Step 6 Copy the saved configuration file to the running configuration.

Step 7 Edit the saved configuration file and delete all commands between the **configure sync** command and the **commit** command, including these two commands.

Step 8 Copy the new, edited configuration file to the running configuration again.

Step 9 Verify that the configuration is correct by entering the **show running-config** command and the **show provision failed-config** *slot* command.

Step 10 If switch profile configuration changes were made on the peer switch while the replacement switch was out of service, apply those configurations in the switch profile and then enter the **commit** command.

Step 11 Shut down all single-homed Fabric Extender vPC host ports.

Step 12 Connect the single-homed Fabric Extender topology fabric ports.

Step 13 Wait for single-homed Fabric Extenders to come online.

Step 14 Ensure the vPC role priority of the existing switch is better than the replacement switch.

Step 15 Connect the vPC peer keepalive link to the peer switch. Ensure that vPC peer keepalive link is operational by entering the **show vpc** command.

Warning **If the auto-recovery feature was not disabled in Step 3, ensure that either the vPC peer-keepalive or the vPC peer-link comes up before the auto-recovery timer expires (default 240 seconds). If this does not happen, the replacement switch will assume the vPC primary role (dual active). If the vPC peer-link is restored in this state, there will be a vPC role change causing vPCs on the peer switch to go down as the switch transitions to vPC secondary role. If required, with just peer-keepalive link operational, reload the replacement switch one more time with all the other interfaces still in the shutdown state.**

Step 16 Ensure that the vPC role field is **none established**. Use the **show vpc** or **show vpc role** command to view the vPC role. If the vPC role field displays **Primary**, then do not proceed with the replacement procedure. Reload the switch to get the vPC role field to **none established**.

Step 17 Connect the vPC peer-link ports to the peer switch. Ensure that the vPC peer link is operational by entering the **show vpc** command.

Step 18 Connect the dual-homed Fabric Extender topology fabric ports.

Step 19 Connect the switch vPC ports.

Step 20 Enter the **no shutdown** command on all single-homed Fabric Extender vPC ports.

Step 21 Verify that the replaced vPC switch and the Fabric Extenders on the replacement switch are online and there is no traffic disruption.

Step 22 If you are using the configuration synchronization feature, add the sync-peer configuration to the switch profile if this wasn't enabled in Step 3.

Step 23 If you are using the configuration synchronization feature, enter the **show switch-profile** *name* **status** command to ensure both switches are synchronized.

Step 24 If vPC auto recovery was disabled, enable auto recovery using the **auto-recovery** command under vPC domain on both switches.

Replacing a Cisco Nexus 2000 Series Fabric Extender

This section describes how to replace a Cisco Nexus 2000 Series Fabric Extender with minimal disruption. This section includes the following topics:

Replacing a Fabric Extender in a Dual-Homed Fabric Extender vPC Topology

Because the hosts behind a Fabric Extender in a dual-homed Fabric Extender vPC topology are by definition singly-connected, traffic disruption will occur for those hosts.

If the replacement Fabric Extender is a different model, the Cisco Nexus 5000 Series switch does not allow you to pre-provision a new type until you disconnect the old Fabric Extender.

To retain the configuration on both Cisco Nexus 5000 Series peer switches in the vPC topology, follow these steps.

Step 1 Save the configuration for the Fabric Extender interfaces to a file.

Step 2 Disconnect the Fabric Extender fabric ports and wait until the Fabric Extender is offline.

Step 3 Pre-provision the slot with the new Fabric Extender model.

Step 4 Modify the configuration file if necessary for the new Fabric Extender if the configurations are incompatible.

Note For vPC ports, this step might affect consistency.

Step 5 Copy the file to the running configuration.

Step 6 Connect the Fabric Extender fabric and host ports and then wait for the Fabric Extender to come online.

Step 7 Verify that all ports are up with the correct configuration.

Replacing a Fabric Extender in a Single-Homed Fabric Extender vPC Topology

If the replacement Fabric Extender is the same model as the original Fabric Extender, then there is no disruption; the configuration on the Fabric Extender interfaces remain unchanged.

If the replacement Fabric Extender is a different model, the Cisco Nexus 5000 Series switch does not allow you to pre-provision a new type until you disconnect the old Fabric Extender.

To replace a Fabric Extender in a single homed Fabric Extender vPC topology, follow the procedure described in "Replacing a Fabric Extender in a Dual-Homed Fabric Extender vPC Topology" section on page 1-13.

Installing a New Cisco Nexus 2000 Series Fabric Extender

With pre-provisioning, you can fully configure the new Fabric Extender before the Fabric Extender is connected to a Cisco Nexus 5000 Series switch.

To install a new Cisco Nexus 2000 Series Fabric Extender, follow these steps:

Step 1 Pre-provision the slot with the Fabric Extender model.

Step 2 Configure the interfaces as though the Fabric Extender is connected.

Step 3 Connect the Fabric Extender and wait for it to come online.

Step 4 Verify that all configurations are applied correctly

Note The switch applies all configurations serially in a best-effort fashion when the Fabric Extender comes online.

vPC Failure Recovery

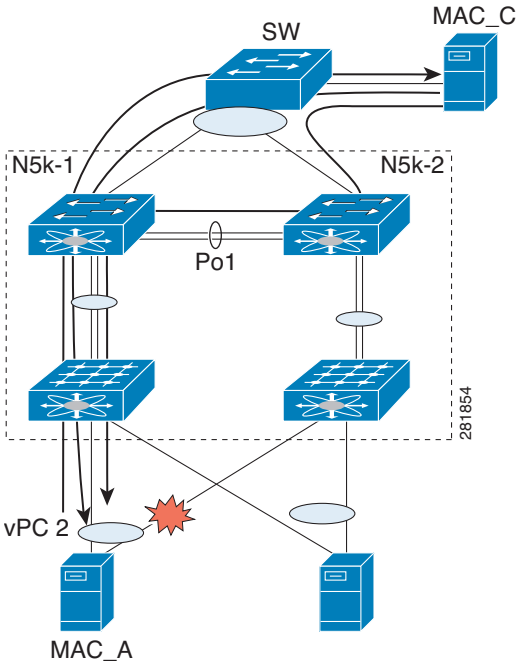
This section describes different vPC failure scenarios and how to recover from them. This section includes the following topics:

vPC Member Port Failure

Figure 1-1 shows the traffic flow when one vPC member port fails. Once the host MAC_A detects a link failure on one of the port-channel members, it redistributes the affected flows to the remaining port channel members. The return flow from MAC_C to MAC_A could take the path of the left- or the right-side Cisco Nexus 5000 Series switch, depending on the port-channel hash algorithm of the top switch. For those flows that traverse the right-side Cisco Nexus 5000 Series switch (the red line), the Cisco Nexus 5000 Series switch passes the traffic to the left-side Cisco Nexus 5000 Series switch, because it no longer has the local connection to host MAC_A. This is one of the scenarios where a vPC peer link is used to carry data traffic.

We recommend that you provision enough bandwidth for peer links to accommodate the bandwidth needed for link failure scenarios.

Figure 1-1 vPC Response to a Member Port Failure

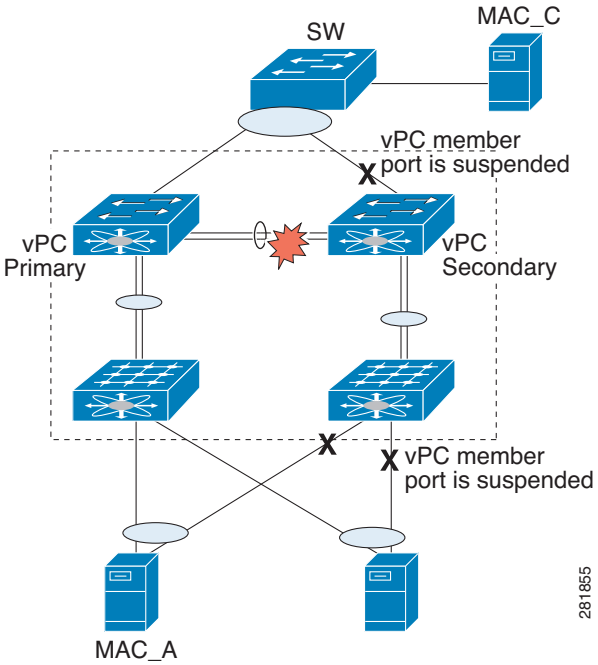


vPC Peer Link Failure

Figure 1-2 shows the vPC response to a peer link failure. In a vPC topology, one vPC peer switch is elected as the vPC primary switch and the other switch is elected as the vPC secondary switch, based on the configured role priority for the switch. In the unlikely scenario where the vPC peer link goes down, the vPC secondary switch shuts down all of its vPC member ports if it can still receive keepalive messages from the vPC primary switch (which indicates that the vPC primary switch is still alive). The vPC primary switch keeps all of its interfaces up. As a result, the hosts or switches that are connected to the Cisco Nexus 5000 Series switch or Cisco Nexus 2000 Series Fabric Extender vPC pair redistributes all the flows to the vPC member ports that are connected to the vPC primary switch.

As a best practice, we recommend that you configure a physical port channel that has at least two 10 Gigabit-Ethernet ports as the vPC peer link.

Figure 1-2 vPC Response to a Peer Link Failure



A vPC consistency check cannot be done when a vPC peer-link is down either due to a link failure or when the peer switch is completely down. In either case, any newly configured vPC does not come up because the vPC consistency check cannot proceed, or the existing vPC remains disabled after the link flaps.

Use the reload restore feature that was introduced in Cisco NX-OS Release 5.0(2)N1(1) to fix this problem. The reload restore feature allows a switch to bypass the vPC consistency check and bring up vPC ports when the peer-link or peer switch fails. The reload restore feature has been replaced with the auto-recovery feature in Cisco NX-OS Release 5.0(2)N2(1).

vPC Peer Keepalive Link Failure

The vPC keepalive link carries the heartbeat message between two vPC peer switches. The failure of the vPC keepalive link alone does not impact the vPC operation or data forwarding. Although it has no impact on data forwarding, we recommend that you fix the keepalive as soon as possible to avoid a double failure scenario that could impact the data traffic.

When both switches come up together (such as after power gets restored following a power outage) and only the mgmt/keepalive link fails, the peers are unreachable. However, all other links, including vPC peer links, are up. In this scenario, reaching the vpc-peers through keepalives are achieved through keepalive links while the primary and secondary role election is established through the vpc-peer link. You must establish the first keepalive for the role election to occur in the case when a switch comes up and the vPC-peer link is up.

When keepalives fail to reach the peer switches, role election does not proceed and the primary or secondary role is not established on either vPC peer switch and all vPC interfaces are kept down on both switches.

Note If this scenario occurs again or if the keepalive link goes down after vPC peers are established, the roles do not change and all vPCs remain up.

vPC Peer Switch Failure

When one peer switch fails, half of the network bandwidth is lost and the remaining vPC switch maintains the network connectivity. If the failure occurs on a primary switch, the secondary switch becomes the primary switch.

When one peer switch fails, the remaining peer switch maintains network connectivity for the vPC until it is reloaded. This situation could happen if both vPC peer switches are reloaded and only one switch comes up or both switches loose power and then the power is restored only on one switch. In either case, since the vPC primary election cannot proceed, the Cisco Nexus 5000 Series switch keeps the vPC ports in suspend mode.

To fix these problems, use the reload restore feature and the auto recovery feature as follows:

In NX-OS Release 5.0(2)N1(1), enter the **reload restore** command:

```
switch(config-vpc-domain)# reload restore <timeout in second>
```

In NX-OS Release 5.0(2)N2(1), enter the **auto-recovery reload-delay** command:

```
switch(config-vpc-domain)# auto-recovery reload-delay ?
  <240-3600>  Time-out for restoring vPC links (in seconds)
```

These commands allow the vPC peer switch to bypass the vPC consistency check and bring up vPC ports after the delay timer expires.

vPC Peer Link Failure Followed by a Peer Keepalive Link Failure

If a peer link failure occurs, the vPC secondary switch checks if the primary switch is alive. The secondary switch suspends its vPC member ports after it confirms that the primary switch is up.

If the vPC primary switch goes down, the vPC secondary switch stops receiving Keepalive messages on the vPC Peer Keepalive link. After three consecutive Keepalive message timeouts, the vPC secondary switch changes its role to be the vPC primary switch and brings up its vPC member ports.

In Cisco NX-OS Release 5.0(2)N2(1), if you enable the auto-recovery feature and if the vPC primary switch goes down, the vPC secondary switch does not receive messages on the vPC peer keepalive link. Then, after three consecutive keepalive timeouts, the vPC secondary switch changes its role to primary and brings up the vPC member ports.

vPC Keepalive Link Failure Followed by a Peer Link Failure

If the vPC keepalive link fails first and then a peer link fails, the vPC secondary switch assumes the primary switch role and keeps its vPC member ports up.

If the peer link and keepalive link fails, there could be a chance that both vPC switches are healthy and the failure occurs because of a connectivity issue between the switches. In this situation, both vPC switches claim the primary switch role and keep the vPC member ports up. This situation is known as a

split-brain scenario. Because the peer link is no longer available, the two vPC switches cannot synchronize the unicast MAC address and the IGMP group and therefore they cannot maintain the complete unicast and multicast forwarding table. This situation is rare.

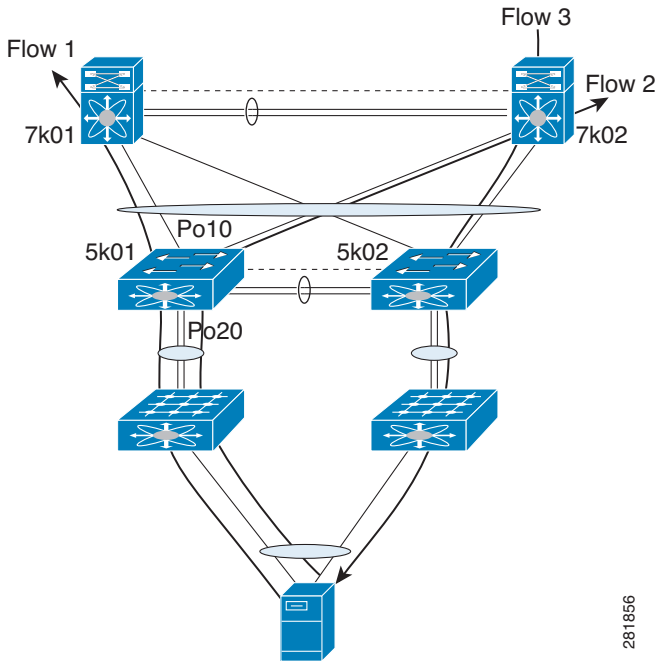
We recommend that you have a well-planned network design that includes spreading peer links and keepalive links to multiple ASICs or multiple modules and different cabling routes for keepalive and peer links to avoid a double failure.

Tracing Traffic Flow in a vPC Topology

This section describes how to trace a traffic flow in a vPC topology that is similar to a port-channel environment.

Figure 1-3 shows that each hop in the network chooses one vPC member port to carry the traffic flow independently.

Figure 1-3 Traffic Flow in a vPC Topology



In this example, for flow 1, the host makes a decision whether the traffic flow is sent to the FEX on left or the right side. The FEX runs its hash algorithm to choose one uplink to carry the flow. The N5k determines if the flow should be sent to N7k1 or N7k2. When the egress port for a traffic flow is a vPC, the vPC switch always prefers to use its own vPC member port to carry the traffic in order to minimize the utilization of peer links.

The Cisco NX-OS and Cisco IOS software includes commands to identify the port channel member that carries a particular flow.

This example assumes that the default hash algorithm is used which is src-mac, dst-mac, src-ip and dst-ip. If the hash algorithm also includes the Layer 4 UDP/TCP port, the port information also needs to be provided in the command. The port channel in the command should be the egress port channel.

```
switch# show port-channel load-balance forwarding-path interface Po3 src-interface
ethernet 1/1 vlan 1 src-mac 0000.0000.1111 src-ip 1.1.1.1 dst-mac 001e.1324.4dc0 dst-ip
2.2.2.2
Missing params will be substituted by 0's.
Load-balance Algorithm on switch: source-dest-ip
crc8_hash: 14   Outgoing port id: Ethernet1/31
Param(s) used to calculate load-balance:
        dst-ip:   2.2.2.2
        src-ip:   1.1.1.1
        dst-mac:  001e.1324.4dc0
        src-mac:  0000.0000.1111
switch#
```

The commands do not show how flows are distributed on the FEX uplink from the FEX to the N5k.

While using the SPAN feature to monitor the traffic flow, the communications between two hosts can be split between two vPC switches. Therefore, you may need to enable SPAN on both vPC switches to obtain a complete trace.