



## Preface

---

This preface describes the audience, organization, and conventions of the *Cisco Nexus 5000 and 6000 Series NX-OS Interfaces Operations Guide, Release 6.0(2)N1(1)*. It also provides information on how to obtain related documentation.

This chapter includes the following topics:

- [Audience, page 1](#)
- [Document Conventions, page 1](#)
- [Related Documentation, page 3](#)
- [Obtaining Documentation and Submitting a Service Request, page 3](#)

## Audience

This publication is for experienced network administrators who configure and maintain Cisco NX-OS on Cisco Nexus 5000 Platform switches and Cisco Nexus 5500 Platform switches.

## Document Conventions

Command descriptions use the following conventions:

Convention	Description
<b>bold</b>	Bold text indicates the commands and keywords that you enter literally as shown.
<i>Italic</i>	Italic text indicates arguments for which the user supplies the values.
[x]	Square brackets enclose an optional element(keyword or argument).
[x   y]	Square brackets enclosing keywords or arguments separated by a vertical bar indicate an optional choice.
{x   y}	Braces enclosing keywords or arguments separated by a vertical bar indicate a required choice.

[x {y   z}]	Nested set of square brackets or braces indicate optional or required choices within optional or required elements. Braces and a vertical bar within square brackets indicate a required choice within an optional element.
variable	Indicates a variable for which you supply values, in context where italics cannot be used.
string	A nonquoted set of characters. Do not use quotation marks around the string or the string will include the quotation marks.

Screen examples use the following conventions::

Convention	Description
screen font	Terminal sessions and information the switch displays are in screen font.
<b>boldface screen font</b>	Information you must enter is in boldface screen font.
<i>italic screen font</i>	Arguments for which you supply values are in italic screen font.
< >	Nonprinting characters, such as passwords, are in angle brackets.
[ ]	Default responses to system prompts are in square brackets.
!, #	An exclamation point (!) or a pound sign (#) at the beginning of a line of code indicates a comment line.

**Note**

Means *reader take note*. Notes contain helpful suggestions or references to material not covered in the manual.

**Caution**

Means *reader be careful*. In this situation, you might do something that could result in equipment damage or loss of data.

## Related Documentation

Documentation for Cisco Nexus 5000 Series Switches and Cisco Nexus 2000 Series Fabric Extenders is available at the following URL:

[http://www.cisco.com/en/US/products/ps9670/tsd\\_products\\_support\\_series\\_home.html](http://www.cisco.com/en/US/products/ps9670/tsd_products_support_series_home.html)

## Obtaining Documentation and Submitting a Service Request

For information on obtaining documentation, submitting a service request, and gathering additional information, see *What's New in Cisco Product Documentation* at:

<http://www.cisco.com/en/US/docs/general/whatsnew/whatsnew.html>.

Subscribe to *What's New in Cisco Product Documentation*, which lists all new and revised Cisco technical documentation, as an RSS feed and deliver content directly to your desktop using a reader application. The RSS feeds are a free service.





## New and Changed Information

This chapter provides release-specific information for each new and changed feature in the *Cisco Nexus 5000 and 6000 Series NX-OS Interfaces Operations Guide, Release 6.0(2)N1(1)*. The latest version of this document is available at the following Cisco website:

[http://www.cisco.com/en/US/products/ps9670/products\\_installation\\_and\\_configuration\\_guides\\_list.html](http://www.cisco.com/en/US/products/ps9670/products_installation_and_configuration_guides_list.html)

To check for additional information about Cisco NX-OS Release 5.x, see the *Cisco Nexus 5000 Series Switch NX-OS Release Notes* available at the following Cisco website:

[http://www.cisco.com/en/US/products/ps9670/prod\\_release\\_notes\\_list.html](http://www.cisco.com/en/US/products/ps9670/prod_release_notes_list.html)

**Table 1** lists the new and changed features for the *Cisco Nexus 5000 and 6000 Series NX-OS Interfaces Operations Guide, Release 6.0(2)N1(1)*.

**Table 1** *New and Changed Features for Release 6.0(2)N1(1)*

Feature	Description	Changed in Release	Where Documented
Added new chapter: Cisco Nexus 2000 Series Fabric Extenders	Provides information on Cisco Nexus 2000 Series Fabric Extenders.	6.0(2)N1(1)	<a href="#">Chapter 1, “Cisco Nexus 2000 Series Fabric Extenders”</a>
Added new chapter: QSFP+ Ports Operations and Switch Fabric mode.	Provides information on QSFP+ Ports Operations and Switch Fabric mode.	6.0(2)N1(1)	<a href="#">Chapter 1, “QSFP+ Ports Operations and Switch Fabric Mode”</a>
Enhanced vPC	This feature was introduced.	5.1(3)N1(1)	<a href="#">Chapter 1, “Using Enhanced vPC”</a>
Layer 3 and vPC Topologies	Some Cisco Nexus 5500 Series switch topologies do not work properly when both Layer 3 and vPC are enabled.	5.1(3)N1(1)	<a href="#">Nonfunctional Topologies with Layer 3 and vPC Combined, page 13</a>





# Cisco Nexus 2000 Series Fabric Extenders

This chapter provide information on Cisco Nexus 2000 Series Fabric Extenders (FEXs). This chapter contains the following sections:

- [Loop Prevention Enhancement on FEXs, page 1-1](#)
- [Traffic Storm Control, page 1-2](#)

## Loop Prevention Enhancement on FEXs

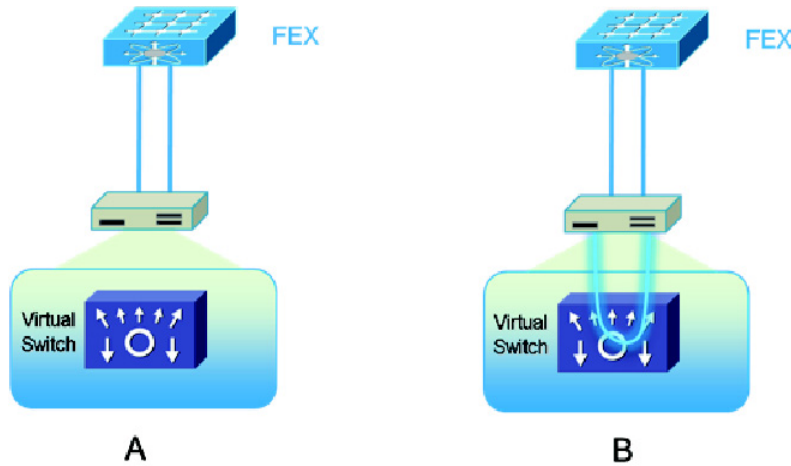
The FEX has two types of interfaces:

- Host Interfaces (HIFs) - Ports that you use to connect to the end host or server device.
- Network Interfaces (NIFs) - Ports that you use to connect to the parent switches. For more information on NIFs , see the [“Network Interface Traffic Storm Control” section on page 1-3](#).

HIF interfaces are always configured as an edge port with Bridge Protocol Data Unit (BPDU) Guard enabled. The BPDU Guard detects loops within a network by sending BPDU messages to the port and putting the port into an error-disabled state. When an HIF experiences a linkup, the BPDU Guard sends out ten BPDUs to prevent any loops within the Layer 2 domain. A BPDU filter is then enabled on the interface and no additional BPDUs are sent out.

You can connect an HIF to a Hypervisor that contains a virtual network. Initially, the virtual network does not bridge the two ports on a server so it does not create a loop. However, a loop might occur if you make any adjustments to the configuration. In this case, the loop prevention mechanism of sending BPDUs at the linkup stage does not work because the links between the FEX and the server stay up. [Figure 1-1](#) shows how a loop might occur within a virtual network.

Figure 1-1 FEX Connection to a Virtual Server



To help with the detection of a loop in this environment, enter the **spanning-tree bpdudfilter disable** command to enable the HIF ports to send out BPDUs. To enable the **spanning-tree bpdudfilter disable** command, you must enter interface configuration mode. This example shows how to enable the feature on the port Ethernet 101/1/10:

```
switch# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
switch(config)# interface Ethernet 101/1/10
switch (config-if)# spanning-tree bpdudfilter disable
switch (config-if)#
```

## Traffic Storm Control

A traffic storm occurs when packets flood the LAN and create excessive traffic and degrade network performance. You can use the traffic storm control feature to prevent LAN ports from being disrupted by a broadcast, multicast, or unicast traffic storm on physical interfaces. Traffic storm control enables you to configure the traffic level to monitor the levels of the incoming broadcast, multicast, and unicast traffic over a 10-microsecond interval. During this interval, the traffic level, which is a percentage of the total available bandwidth of the port, is compared with the traffic storm control level that you configured. When the ingress traffic reaches the traffic storm control level that is configured on that port, traffic storm control drops the traffic until the interval ends.



**Note** By default, Cisco NX-OS takes no corrective action when the traffic exceeds the configured level.

In the Nexus FEX architecture, the traffic storm control feature is implemented using the hardware resource on the Cisco Nexus 5500 and Cisco Nexus 6000 Series switches. For more information on implementing the traffic storm control feature, see the [Cisco Nexus 2000 Fabric Extenders Hardware Installation Guide](#).



**Note** This feature is not supported in the FEX architecture with the Cisco Nexus 5000 switch.

For more information about Fabric Extenders see the [Cisco Nexus 2000 Fabric Extenders Software Configuration Guide](#).



## Network Interface Traffic Storm Control

Network interfaces (NIFs) are the uplinks interfaces on the FEX. The parent switch initiates an NIF traffic storm by monitoring the ingress traffic on the fabric ports. You can use these two methods to connect the FEX to the parent switch:

- Static pinning fabric interface connection.
- EtherChannel fabric interface connection.

With static pinning fabric interface connection, the NIF traffic storm control is configured under the physical fabric interface. The downlink host interfaces on the FEX are pinned to the fabric interfaces in the order that they were initially configured.

This example shows to configure a static pinning fabric interface connection to a switch:

```
switch(config)# interface type slot/port
switch(config-if)# storm-control { broadcast | multicast | unicast } level whole[.decimal]
```

With an EtherChannel fabric interface connection, the NIF traffic storm control feature is configured under the EtherChannel interface, which means that packets are evenly distributed over the links within the EtherChannel. If all the links within the EtherChannel are inside the same port application-specific integrated circuit (ASIC), the port ASIC is programmed with the configured level. If the links spread between different port ASICs, each ASIC is programmed proportionally to the links within the same ASIC. For example, an EtherChannel with three 10 Gigabit Ethernet links is configured with a storm-control level of 80 percent. Two 10 Gigabit Ethernet links are managed by port ASIC A. The third 10 Gigabit Ethernet link is managed by port ASIC B. Port ASIC A is programmed to monitor traffic with a threshold of 16 Gbps on both ports. Port ASIC B is programmed to monitor traffic with a threshold of 8 Gbps for that single port. If the FEX is dual home to two virtual port channel (vPC) peered switches, you must ensure that both vPC peered switches have the same storm-control configuration under the EtherChannel interfaces.

This example shows how to configure an EtherChannel Fabric Interface connection to a switch:

```
switch(config)# interface port-channel number
switch(config-if)# storm-control { broadcast | multicast | unicast } level whole[.decimal]
```





# CHAPTER 1

## QSFP+ Ports Operations and Switch Fabric Mode

This chapter describes the application-specific integrated circuit (ASIC) mapping and configuration of quad small form-factor pluggable plus (QSFP+) ports on the Cisco Nexus 2248PQ Fabric Extender (FEX), the Cisco Nexus 5500 with QSFP+ Generic Expansion Module (N55-M4Q) switch, the Cisco Nexus 6004 switch and the Cisco Nexus 6004 switch fabric mode.

This chapter contains the following sections:

- [Quad Small Form-Factor Pluggable Plus, page 1-1](#)
- [Cisco Nexus 2248PQ Fabric Extender, page 1-1](#)
- [Cisco Nexus 5500 Series with QSFP+ GEM, page 1-2](#)
- [Cisco Nexus 6004 Switch and QSFP+ Ports, page 1-4](#)

## Quad Small Form-Factor Pluggable Plus

The quad small form-factor pluggable plus (QSFP+) is a transceiver module that offers high-density 40 Gigabit Ethernet connectivity options for datacenter and high-performance computing networks. The QSFP+ transceiver module is a hot-swappable, parallel fiber-optical/copper module with four independent optical transmit and receive channels. These channels can terminate in another 40 Gigabit Ethernet QSFP+ transceiver, or the channels can be broken out to four separate 10 Gigabit Ethernet SFP+ transceivers. The QSFP+ transceiver module connects the electrical circuitry of the system with either a copper or an optical external network. For more information about the Cisco 40 Gigabit Ethernet modules, see:

[http://www.cisco.com/en/US/partner/products/ps11708/tsd\\_products\\_support\\_model\\_home.html](http://www.cisco.com/en/US/partner/products/ps11708/tsd_products_support_model_home.html).

## Cisco Nexus 2248PQ Fabric Extender

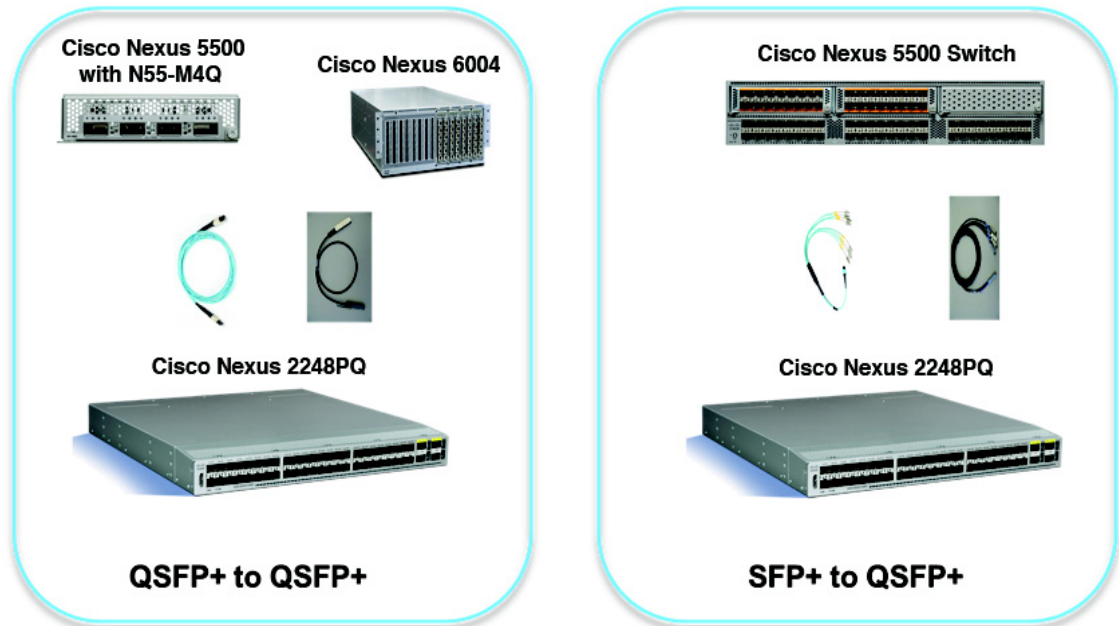
The Cisco Nexus 2248PQ is a 1 rack unit (1 RU) Fabric Extender(FEX) with 48 SFP+ host ports and 4 QSFP+ uplink ports. The host ports can be operated in either 1 Gigabit Ethernet or 10 Gigabit Ethernet mode. The 4 QSFP+ uplink ports can operate only in 4 x10 Gigabit Ethernet mode. The uplink ports do not support the native 40 Gigabit Ethernet mode.

You can connect the Cisco Nexus 2248PQ FEX to its parent switch in the following two ways:

- QSFP+ to QSFP+
- QSFP+ to SFP+

Figure 1-1 shows the cabling options for a Cisco Nexus 2248PQ FEX to Cisco Nexus 5500 and Cisco Nexus 6004 Series switches.

**Figure 1-1** Connection Cable Options for a Cisco Nexus 2248PQ FEX to Cisco Nexus 5500 and Cisco Nexus 6004 Series Switches



**Note**

For the Cisco Nexus 6004 switch, the default mode of the QSFP+ ports is native 40 Gigabit Ethernet. You must configure QSFP+ ports on the Cisco Nexus 6004 that are connected to the Cisco Nexus 2248PQ in 4x10 Gigabit Ethernet mode. See the [Configuring QSFP+ Ports, page 1-5](#) section for more details about the configuration.

## Cisco Nexus 5500 Series with QSFP+ GEM

The QSFP+ GEM (N55-M4Q) is a generic expansion module (GEM) that provides 16x10 Gigabit Ethernet SFP+ in a 4 x QSFP+ form factor. This module does not support native 40 Gigabit Ethernet on the QSFP+ port. This module is supported on all Cisco Nexus 5500 chassis (Cisco Nexus 5548P, Cisco Nexus 5548UP, Cisco Nexus 5596UP, and Cisco Nexus 5596T). This module is suitable for any expansion slot on the Cisco Nexus 5500 Series and you can populate all the expansion slots on the Cisco Nexus 5596 chassis with N55-M4Qs. There is no additional configuration needed to specify the operation mode of the QSFP+ ports because all the ports only operate in 4 x10 Gigabit Ethernet mode. The numbering of the ports uses the following convention(see [Figure 1-2](#)):

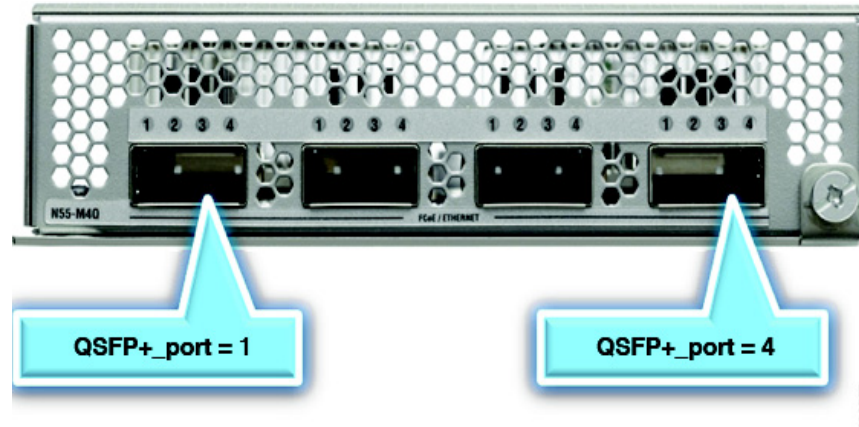
```
interface ethernet [chassis_ID/] slot/ QSFP_port [/subintf_port-no]
```



**Note**

The argument, *chassis\_ID*, is not optional when addressing the host interfaces of a Cisco Nexus 2000 Series Fabric Extender.

Figure 1-2 QSFP+ Slot Number on N55-M4Q Module



This example shows how to display the information about the module and interface on the Cisco Nexus 5596 switch:

```
Nexus5596# show module 3
Mod Ports Module-Type Model Status
-----
3 16 N5K QSFP+ Expansion Module N55-M4Q ok

Mod Sw Hw World-Wide-Name(s) (WWN)
-----
3 6.0(2)N1(1) 2.0 --

Mod MAC-Address(es) Serial-Num
-----
3 547f.eed8.b4a0 to 547f.eed8.b4af FOC16417WYT
Nexus5596# show running-config | section Ethernet3
interface Ethernet3/1/1
interface Ethernet3/1/2
interface Ethernet3/1/3
interface Ethernet3/1/4
interface Ethernet3/2/1
interface Ethernet3/2/2
interface Ethernet3/2/3
interface Ethernet3/2/4
interface Ethernet3/3/1
interface Ethernet3/3/2
interface Ethernet3/3/3
interface Ethernet3/3/4
interface Ethernet3/4/1
interface Ethernet3/4/2
interface Ethernet3/4/3
interface Ethernet3/4/4
Nexus5596# show interface e3/1/1
Ethernet3/1/1 is up
Dedicated Interface
Hardware: 1000/10000 Ethernet, address: 547f.eed8.b4a0 (bia 547f.eed8.b4a0)
MTU 1500 bytes, BW 10000000 Kbit, DLY 10 usec
reliability 255/255, txload 1/255, rxload 1/255
Encapsulation ARPA
Port mode is access
full-duplex, 10 Gb/s, media type is 40G
Beacon is turned off
Input flow-control is off, output flow-control is off
Rate mode is dedicated
```

```

Switchport monitor is off
EtherType is 0x8100
Last link flapped 00:00:35
Last clearing of "show interface" counters never
2 interface resets
30 seconds input rate 400 bits/sec, 0 packets/sec
30 seconds output rate 1104 bits/sec, 0 packets/sec
Load-Interval #2: 5 minute (300 seconds)
  input rate 80 bps, 0 pps; output rate 432 bps, 0 pps
RX
  0 unicast packets 68719489350 multicast packets 0 broadcast packets
  12614 input packets 1056826 bytes
  0 jumbo packets 0 storm suppression bytes
  0 runts 0 giants 0 CRC 0 no buffer
  0 input error 0 short frame 0 overrun 0 underrun 0 ignored
  0 watchdog 0 bad etype drop 0 bad proto drop 0 if down drop
  0 input with dribble 0 input discard
  0 Rx pause
TX
  0 unicast packets 660284 multicast packets 0 broadcast packets
  24065 output packets 1919643 bytes
  0 jumbo packets
  0 output errors 0 collision 0 deferred 0 late collision
  0 lost carrier 0 no carrier 0 babble 0 output discard
  0 Tx pause

Nexus5596#

```

## Cisco Nexus 6004 Switch and QSFP+ Ports

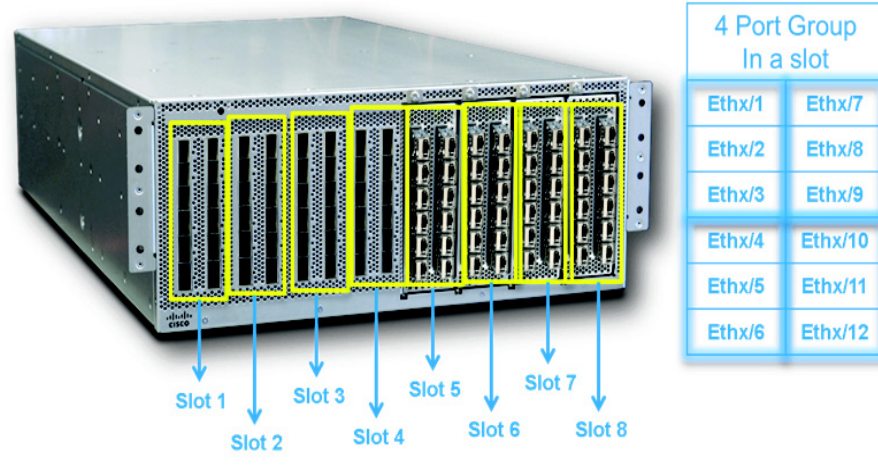
The Cisco Nexus 6000 Series platform uses Cisco's ASIC technology. It provides line rate throughput and low latency performance. The ASICs on the Cisco Nexus 6000 Series platform are unified port controllers (UPCs) and switch fabric (SF). The UPCs provide data-plane processing and the SF cross-connects the UPCs. Each UPC can be operated as either 3 ports of 40G Ethernet or 12 ports of 10G Ethernet.

The Cisco Nexus 6004 switch is a 4 RU unit chassis with 8 vertical slots. The first 4 slots are fixed with 12 QSFP+ ports per slot. The remaining 4 slots are for linecard expansion modules (LEM). A QSFP+ LEM is a 12 QSFP+ port, field-replaceable module that has the same functionality as the fixed slots. These 4 slots provide up to 48 additional 40 Gigabit Ethernet ports. [Figure 1-3](#) shows the Cisco Nexus 6004 switch port numbering and port mapping with UPCs in a slot.

This section includes the following topics:

- [Configuring QSFP+ Ports, page 1-5](#)
- [Cisco Nexus 6004 Switch Fabric Mode, page 1-8](#)

Figure 1-3 Cisco Nexus 6004 Switch Port Numbering and Port Mapping with UPCs in a Slot



396558

## Configuring QSFP+ Ports

QSFP+ ports can operate as either 40 Gigabit Ethernet or 4x10 Gigabit Ethernet modes on the Cisco Nexus 6004 switch. The default mode is 40 Gigabit Ethernet for the fixed QSFP+ ports and QSFP+ ports on a LEM. The port numbering of the QSFP+ port operating in 40 Gigabit Ethernet mode contains two levels. The first level is the slot number. There are eight slots in the Cisco Nexus 6004 switch. The farthest left slot is slot 1 and the farthest right is slot 8. The second level is the QSFP+ port number. There are 16 ports in the built-in slot and 16 ports in the QSFP+ LEM. The QSFP+ port numbering system is from top to bottom and left to right as indicated in Figure 1-3. The QSFP+ port reference on the Cisco Nexus 6004 switch is as follows:

```
interface ethernet [chassis_ID] slot/QSFP_port [/subintf_port-no]
```

To change from the default mode to 4x10 Gigabit Ethernet mode, you have to change it on a per port group basis. A port group is the ports that are managed by the same UPC. The table in Figure 1-3 shows the four available port groups within a given slot. To change from the default mode of 40 Gigabit Ethernet to 4x10 Gigabit Ethernet mode for any port group, follow these steps:

### SUMMARY STEPS

1. **Configure terminal**
2. **interface breakout slot** *slot-number* **port** *port-range* **map** 10g-4x
3. **power off module** *module*
4. **no power off module** *module*
5. (Optional) **show interface brief**

## DETAILED STEPS

	Command	Purpose
Step 1	<b>configure terminal</b>  <b>Example:</b> N6004-TME3# <code>config t</code> N6004-TME3(config)#	Enters global configuration mode.
Step 2	<b>interface breakout slot</b> <i>slot-number port port-range map</i> <b>10g-4x</b>  <b>Example:</b> N6004-TME3(config)# <code>interface</code> <code>breakout slot 2 port 1-3 map</code> <code>10g-4x</code>	Enables you to configure the linecard expansion module in 10g mode.
Step 3	<b>power off module</b> <i>module</i>  <b>Example:</b> N6004-TME3(config)# <code>poweroff</code> <code>module 2</code>	Enables you to power off the module.
Step 4	<b>no power off module</b> <i>module</i>  <b>Example:</b> N6004-TME3(config)# <code>no</code> <code>poweroff module 2</code>	Enables you to return power to the module.
Step 5	(Optional) <b>Show interface brief</b>  <b>Example:</b> N6004-TME3# <code>show interface</code> <code>brief   incl Eth2</code>	(Optional) Displays brief information about the interface.

## Examples

This example shows how to change QSFP+ ports 1 to 3 and 7 to 12 to 4 x10 Gigabit Ethernet mode on slot 2:

```
N6004-TME3# show interface brief | incl Eth2
Eth2/1      1      eth  access down    SFP not inserted      40G(D) --
Eth2/2      1      eth  access down    SFP not inserted      40G(D) --
Eth2/3      1      eth  access down    SFP not inserted      40G(D) --
Eth2/4      1      eth  access down    SFP not inserted      40G(D) --
Eth2/5      1      eth  access down    SFP not inserted      40G(D) --
Eth2/6      1      eth  access down    SFP not inserted      40G(D) --
Eth2/7      1      eth  access down    SFP not inserted      40G(D) --
Eth2/8      1      eth  access down    SFP not inserted      40G(D) --
Eth2/9      1      eth  access down    SFP not inserted      40G(D) --
Eth2/10     1      eth  access down    SFP not inserted      40G(D) --
Eth2/11     1      eth  access down    SFP not inserted      40G(D) --
Eth2/12     1      eth  access down    SFP not inserted      40G(D) --
N6004-TME3# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
N6004-TME3(config)# interface breakout slot 2 port 1-3 map 10g-4x
N6004-TME3(config)# interface breakout slot 2 port 7-12 map 10g-4x
N6004-TME3(config)#
```



This example shows you how to power off a module:

```
N6004-TME3(config)# poweroff module 2
N6004-TME3(config)# 2013 Jan 2 23:30:40 N6004-TME3 %$ VDC-1 %$ %PFMA-2-MOD_REMOVE: Module
2 removed (Serial number FOC16422P28)
N6004-TME3(config)#
```

This example shows you how to power on a module:

```
N6004-TME3(config)# no poweroff module 2
N6004-TME3(config)# show module 2
Mod Ports Module-Type Model Status
-----
2 39 Norcal Ethernet Module N6K-FIXED-LEM ok

Mod Sw Hw World-Wide-Name(s) (WWN)
---
2 6.0(2)N1(1) 1.0 --

Mod MAC-Address(es) Serial-Num
---
2 a44c.11e8.3a10 to a44c.11e8.3a1f FOC16422P28
N6004-TME3(config)# show interface brief | incl Eth2
Eth2/1/1 1 eth access down SFP not inserted 10G(D) --
Eth2/1/2 1 eth access down SFP not inserted 10G(D) --
Eth2/1/3 1 eth access down SFP not inserted 10G(D) --
Eth2/1/4 1 eth access down SFP not inserted 10G(D) --
Eth2/2/1 1 eth access down SFP not inserted 10G(D) --
Eth2/2/2 1 eth access down SFP not inserted 10G(D) --
Eth2/2/3 1 eth access down SFP not inserted 10G(D) --
Eth2/2/4 1 eth access down SFP not inserted 10G(D) --
Eth2/3/1 1 eth access down SFP not inserted 10G(D) --
Eth2/3/2 1 eth access down SFP not inserted 10G(D) --
Eth2/3/3 1 eth access down SFP not inserted 10G(D) --
Eth2/3/4 1 eth access down SFP not inserted 10G(D) --
Eth2/4 1 eth access down SFP not inserted 40G(D) --
Eth2/5 1 eth access down SFP not inserted 40G(D) --
Eth2/6 1 eth access down SFP not inserted 40G(D) --
Eth2/7/1 1 eth access down SFP not inserted 10G(D) --
Eth2/7/2 1 eth access down SFP not inserted 10G(D) --
Eth2/7/3 1 eth access down SFP not inserted 10G(D) --
Eth2/7/4 1 eth access down SFP not inserted 10G(D) --
<SNIP>
Eth2/12/1 1 eth access down SFP not inserted 10G(D) --
Eth2/12/2 1 eth access down SFP not inserted 10G(D) --
Eth2/12/3 1 eth access down SFP not inserted 10G(D) --
Eth2/12/4 1 eth access down SFP not inserted 10G(D) --
N6004-TME3(config)#
```

After the module is back inline, the port naming changes to three levels. The change takes effect only after you have entered the **power off module** and **power on module** commands. The port reference to the 10 Gigabit Ethernet ports changes to the following:

**interface ethernet** [*chassis\_ID*] *slot*/*QSFP\_port*/*breakout\_port* [*/subintf port-no*]

To change the ports configuration back to 40 Gigabit Ethernet mode, use the **no** form of the **interface breakout slot** command as follows:

**no interface breakout slot** *slot-number port port-range map* 10g-4x

After the module is back inline, the port naming returns to two levels. The change takes effect only after you have entered the **power off module** and **power on module** commands. The port reference to the 40 Gigabit Ethernet ports changes back to the following:

```
interface ethernet [chassis_ID] slot/QSFP_port/ [/.subintf_ port-no]
```

## Cisco Nexus 6004 Switch Fabric Mode

Switch fabric mode refers to the single link speed configuration on the Cisco Nexus 6004 switch. The following are the characteristics of switch fabric mode:

- There are two user configurable switch fabric modes: 40 Gigabit Ethernet and 10 Gigabit Ethernet. The default switch fabric mode is 40 Gigabit Ethernet.
- The switch fabric mode is a system-wide configuration.
- Changing the switch fabric mode requires a chassis reload.
- The total bandwidth between UPCs and switch fabric remain the same regardless of the switch fabric mode configuration.
- The switch fabric mode configuration is independent of the QSFP+ port speed configuration (40 Gigabit Ethernet or 4 x10 Gigabit Ethernet).

We recommend that you use the default switch fabric mode of 40 Gigabit Ethernet. You can change to the 10 Gigabit Ethernet mode if most of the front panel ports are operating in 10 Gigabit Ethernet mode and the application desires a lowest latency between 10 Gigabit Ethernet ports. For this software release, when the fabric mode is set to 10 Gigabit Ethernet, an In Service Software Upgrade (ISSU) is disabled. For future software releases, when the fabric mode is set to 10 Gigabit Ethernet, an ISSU will be enabled.

To change the switch fabric mode, use the **fabric-mode** command. You must perform a system reload for the change to take place.

### SUMMARY STEPS

1. **Configure terminal**
2. **fabric-mode {10g | 40g}**
3. (Optional) **copy running-config startup-config**

## DETAILED STEPS

	Command	Purpose
Step 1	<b>configure terminal</b>  <b>Example:</b> N6004-TME3# <code>config t</code> N6004-TME3(config)#	Enters global configuration mode.
Step 2	<b>fabric-mode {10g   40g}</b>  <b>Example:</b> N6004-TME3(config)# <code>fabric-mode 10g</code>	Enables you to set the fabric mode.
Step 3	<b>(Optional) copy running-config startup-config</b>  <b>Example:</b> N6004-TME3(config)# <code>copy running-config startup-config</code>	(Optional) Copies the running configuration to the startup configuration

## EXAMPLES

This example shows how to select a fabric mode of 10 Gigabit Ethernet:

```
switch# configure terminal
N6004-TME3(config)# fabric-mode 10g
Fabric mode set to 10G. Please copy the configuration and reload the switch
N6004-TME3(config)# copy running-config startup-config
```

This example shows how to select a fabric mode of 40 Gigabit Ethernet:

```
switch# configure terminal
N6004-TME3(config)# fabric-mode 40g
Fabric mode set to 10G. Please copy the configuration and reload the switch
N6004-TME3(config)# copy running-config startup-config
```





# 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, page 1-1](#)
- [vPC Consistency Checks, page 1-1](#)
- [Configuring Changes in vPC Topologies, page 1-9](#)
- [Replacing a Cisco Nexus 5000 Series Switch or Cisco Nexus 2000 Fabric Extender, page 1-10](#)
- [vPC Failure Recovery, page 1-14](#)
- [Tracing Traffic Flow in a vPC Topology, page 1-18](#)

## 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](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:

- [Type 1 and Type 2 Consistency Check Parameters, page 1-2](#)
- [Graceful Consistency Check, page 1-3](#)
- [Configuring Per-VLAN Consistency Checks, page 1-5](#)

- [Identifying Inconsistent vPC Configurations, page 1-6](#)
- [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)
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 Priority)    2      (F, F, F, F, F, F)   (F, F, F, F, F, F)
Output Queuing (Bandwidth)           2      (50, 50, 0, 0, 0, 0) (50, 50, 0, 0, 0, 0)
Output Queuing (Absolute Priority)    2      (F, F, F, F, F, F)   (F, F, F, F, F, F)
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 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           1      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, 0-23-4-ee-be-64, 80c8, 0, 0), (8000, 0-1e-13-15-7-40, 1, 0, 0)]	[(7f9b, 0-23-4-ee-be-64, 80c8, 0, 0), (8000, 0-1e-13-15-7-40, 1, 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	-	1-999,1001-3967,4048-4093	1-3967,4048-4093

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

237956

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

<edit>..

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

```

237959

All VLANs are up

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

```

237960

VLAN 5 is suspended

## 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], [], [], [], [], [1], [1], [1], [1])
Network QoS (MTU)                       2      (1538, 2240, 0, 0, 0, 0)
Network QoS (Pause)                     2      (F, T, F, F, F, F)
Input Queuing (Bandwidth)                2      (50, 50, 0, 0, 0, 0)
Input Queuing (Absolute Priority)        2      (F, F, F, F, F, F)
Output Queuing (Bandwidth)               2      (50, 50, 0, 0, 0, 0)
Output Queuing (Absolute Priority)       2      (F, F, F, F, F, F)
STP Mode                                 1      MST
STP Disabled                             1      None
STP MST Region Name                      1      ""
STP MST Region Revision                  1      0
STP MST Region Instance to               1      1
VLAN Mapping
STP Loopguard                           1      Disabled
STP Bridge Assurance                     1      Enabled
STP Port Type, Edge                      1      Normal, Disabled,
BPDUFilter, Edge BPDUGuard              1      Disabled
STP MST Simulate PVST                    1      Enabled
Allowed VLANs                            -      1-10
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
vPC status
id   Port           Status Consistency Reason                               Active vlans
---
104  Po104          up     success success                                           3000
200  Po200          up     success success                                           1,101-110,1000,3000
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.

```
n5k-1# show vpc consistency-parameters interface ethernet 102/1/1
```

```
Legend:
```

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

Name	Type	Local Value	Peer Value
Speed	1	1000 Mb/s	1000 Mb/s
Duplex	1	full	full
Port Mode	1	trunk	access
Native Vlan	1	1	0
Shut Lan	1	No	No
Allowed VLANs	-	1-999,1001-3967,4048-4093	102

Switch port mode mismatch

237963

## 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, page 1-11](#)
- [Replacing a Cisco Nexus 2000 Series Fabric Extender, page 1-13](#)

## 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, page 1-13](#)
- [Replacing a Fabric Extender in a Single-Homed Fabric Extender vPC Topology, page 1-13](#)
- [Installing a New Cisco Nexus 2000 Series Fabric Extender, page 1-14](#)

### 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:

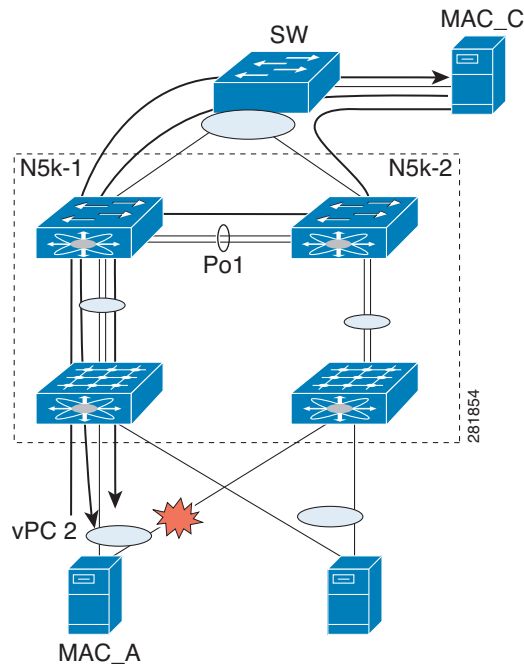
- [vPC Member Port Failure, page 1-14](#)
- [vPC Peer Link Failure, page 1-15](#)
- [vPC Peer Keepalive Link Failure, page 1-16](#)
- [vPC Peer Switch Failure, page 1-17](#)
- [vPC Peer Link Failure Followed by a Peer Keepalive Link Failure, page 1-17](#)
- [vPC Keepalive Link Failure Followed by a Peer Link Failure, page 1-17](#)

## 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

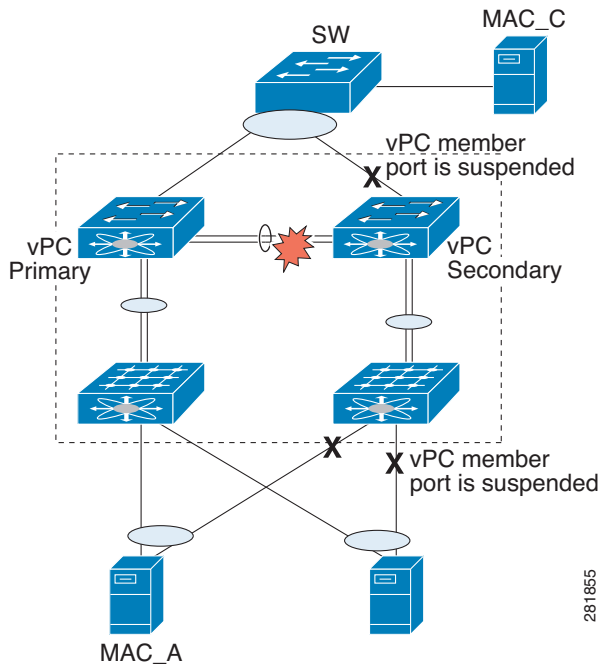


## vPC Peer Link Failure

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 lose 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.

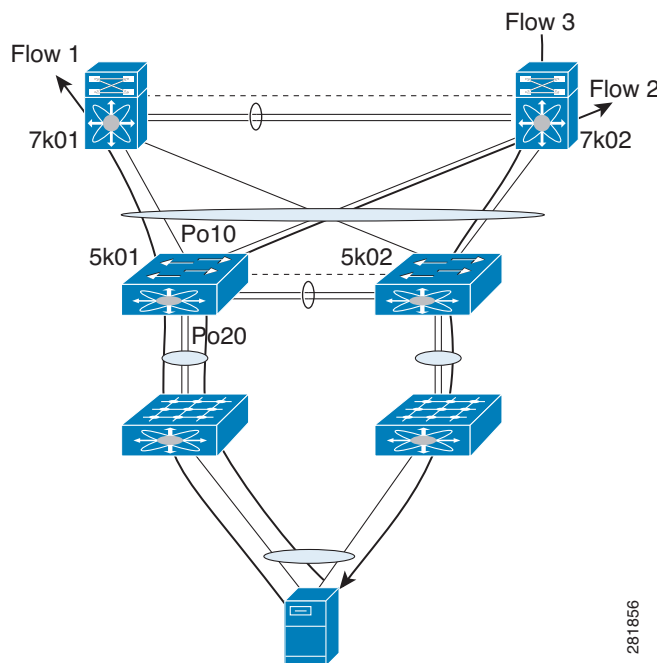
We recommend that you have a well-planned network design that includes spreading peer links and keepalive links to multiple ASICs or multiple modules and different cabling routes for keepalive and peer links to avoid a double failure.

## 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.







# Using Layer 3 and vPC on the Cisco Nexus 5500 Series Device

This chapter describes virtual port channel (vPC) operations when Layer 3 routing features are enabled on the Cisco Nexus 5500 Series device.

This chapter includes the following sections:

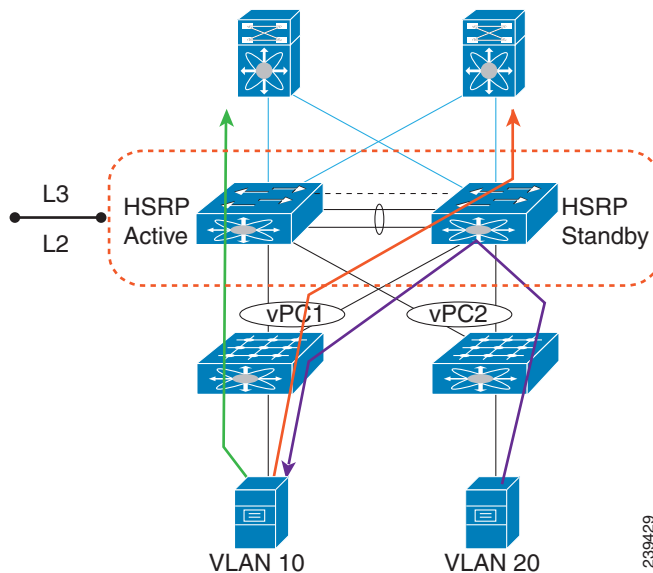
- [vPC and First Hop Redundancy Protocol, page 1-1](#)
- [ARP Processing with vPC, page 1-2](#)
- [Layer 3 Forwarding for Packets to a Peer Switch MAC Address, page 1-2](#)
- [Improved Convergence with a vPC Topology and Layer 3 Routing, page 1-3](#)
- [vPC Peer Link Failure, page 1-4](#)
- [Layer 3 Module Failure, page 1-4](#)
- [Connecting to a Router in a vPC Topology, page 1-5](#)
- [Dedicated VRF For a Keepalive Interface, page 1-6](#)
- [vPC Consistency Check for Layer 3 Parameters, page 1-8](#)
- [Multicast Interaction in a vPC Topology, page 1-8](#)
- [Faster Convergence with the Prebuilt Source Tree, page 1-9](#)
- [Using a vPC Switch as a Designated Router \(PIM DR\), page 1-11](#)
- [Software Upgrade and Downgrade Impact, page 1-17](#)
- [Nonfunctional Topologies with Layer 3 and vPC Combined, page 1-13](#)

## vPC and First Hop Redundancy Protocol

When you use a Cisco Nexus 5548 switch or Cisco Nexus 5596UP switch as a default gateway for hosts, you can deploy the First Hop Redundancy Protocol (FHRP) to provide default gateway redundancy. Beginning with Cisco NX-OS Release 5.0(3)N1(1b), an active FHRP peer and a standby peer can perform Layer 3 forwarding when you enable vPC. This optimization improves bandwidth, avoids sending the Layer 3 traffic over the vPC peer link, and requires no configuration or protocol change. Only the FHRP active peer answers ARP requests. Because both active and standby FHRP peers can forward Layer 3 traffic, you do not need to configure an aggressive timer for FHRP to provide faster failover and convergence time if an active FHRP peer fails.

Figure 1-1 shows that the Layer 3 traffic that originated from the host and is destined to a host several hops away can be routed by both the Host Standby Router Protocol (HSRP) active and the HSRP standby switch.

Figure 1-1 vPC and FHRP



## ARP Processing with vPC

When the host connects to a Cisco Nexus 5500 Platform switch and Cisco Nexus 2000 Fabric Extenders in a vPC topology, the host can send an ARP request to the FHRP standby peer due to a hashing algorithm. The ARP request that is received by the standby peer is forwarded to the active peer and the active peer can answer it with an ARP reply.

Similarly, when traffic is moving from north to south, such as when one Cisco Nexus 5500 Platform switch sends an ARP request to a host, the ARP reply might be sent to another switch. In such a case, the ARP reply is forwarded as a Layer 2 frame to the Cisco Nexus 5500 Platform switch that originated the ARP request.

As of Cisco NX-OS Release 5.0(3)N1(1b), ARP synchronization does not occur between two Cisco Nexus 5500 Platform switches. The two switches resolve and maintain their ARP table independently. When one vPC peer switch is reloaded, the switch needs to resolve the ARP by sending ARP requests to the hosts.

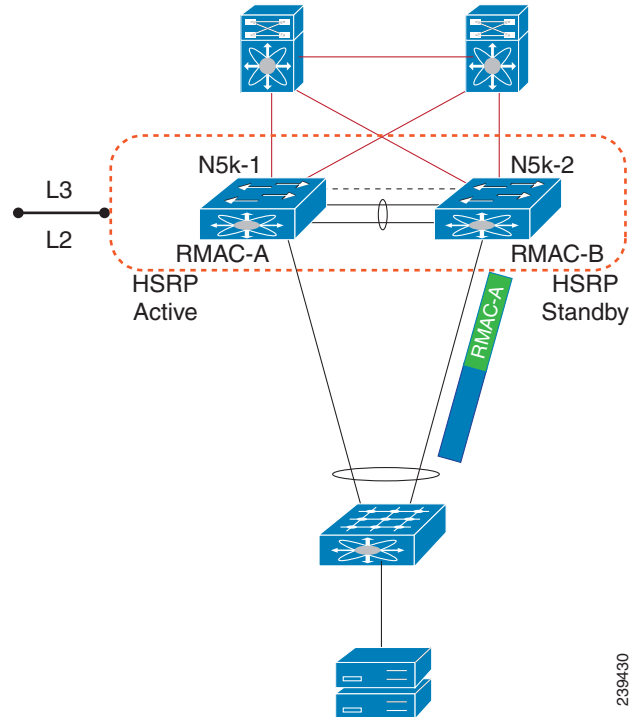
## Layer 3 Forwarding for Packets to a Peer Switch MAC Address

Typically, a router performs a Layer 3 route table lookup and Layer 3 forwarding when the destination MAC in the Ethernet frame matches its own MAC address. Otherwise, the packets are switched (if Layer 2 functionality is enabled) or dropped. In a topology with Layer 3 and vPC enabled, a vPC peer switch could receive IP packets with the peer's MAC address as the destination MAC rather than the virtual MAC address (when FHRP is enabled) or its own MAC address. In this scenario, a Cisco Nexus 5500 Platform switch can forward the traffic to the peer using a peer link and the peer switch performs the Layer 3 forwarding.

The above scenario often happens with some filers. In the case of filers, they may achieve improved load balance and better performance by forwarding traffic to the Burnt-in-Address (BIA) of the routers instead of the HSRP MAC.

Figure 1-2 shows that when the NAS filer sends out packets with N5k-1's MAC RMAC-A as the destination MAC, the packets can be sent over to the N5k-2 switch due to the port channel hashing.

Figure 1-2 vPC and Peer-Gateway



Beginning in Cisco NX-OS Release 5.0(3)N1(1b), you can use the **peer-gateway** command to allow Cisco Nexus 5500 Platform switches to perform Layer 3 forwarding if the destination MAC of the incoming packet is the MAC of its vPC peer switch. The **peer-gateway** command avoids forwarding such packets to the vPC peer link.



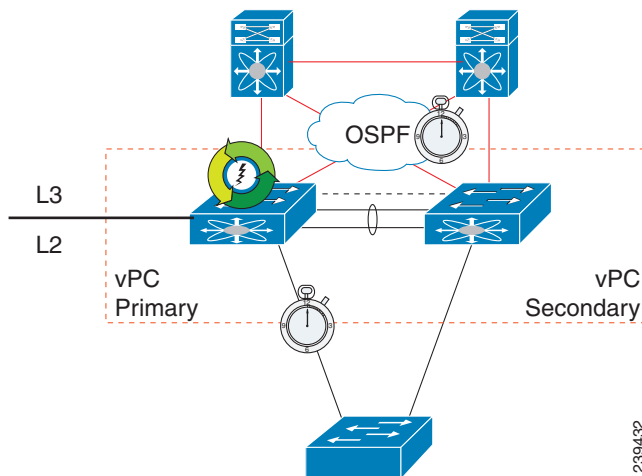
**Note**

You must configure the **peer-gateway** command on both vPC peer switches.

## Improved Convergence with a vPC Topology and Layer 3 Routing

Beginning in Cisco NX-OS Release 5.0(3)N1(1b), a delay timer was introduced to avoid the situation where a vPC member port is brought up before the Layer 3 is converged. For example, when one Cisco Nexus 5500 Platform switch is reloaded, the switch starts to receive traffic from hosts once the vPC member ports are up. A delay might occur before the switch establishes a routing protocol adjacency and learns all routes. During this period of the time, received traffic is dropped due to the lack of a route-to-destination address. Figure 1-3 shows an example of where the delay can be used to avoid black hole traffic when a Cisco Nexus 5000 Platform switch is configured for Layer 3 with vPC.

Figure 1-3 vPC Delay Restore



The delay restore feature allows you to configure a timed delay before vPC member ports are brought online. The delay allows the switch to learn all routes, to bring up the vPC member ports, and to forward traffic from hosts. The following example shows how to configure a timed delay of 120 seconds:

```
layer3-switch(config-vpc-domain)# delay restore ?
<1-3600> Delay in bringing up the vPC links (in seconds)
layer3-switch(config-vpc-domain)# delay restore 120
layer3-switch(config-vpc-domain)#
```

## vPC Peer Link Failure

In addition to suspending vPC member ports, the vPC secondary switch also suspends its switched virtual interface (SVIs) when a vPC peer link is lost. When this occurs, the vPC secondary switch stops advertising the local subnets, which prevents traffic blackholing.

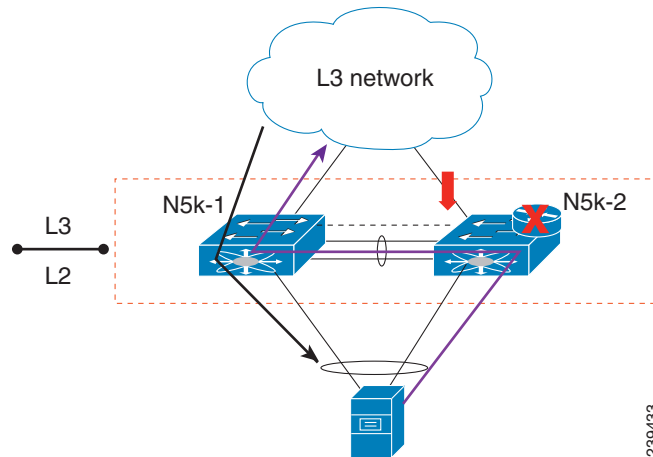
## Layer 3 Module Failure

When a Layer 3 module fails on a Cisco Nexus 5500 Platform switch all Layer 3 interfaces are suspended, including Layer 3 port channel and SVI interfaces. As a result, the Layer 3 routing table on the neighboring routers is updated which results in the north to south traffic to be directed towards the peer Nexus 5500 Platform switch. The Layer 2 interfaces, including the Layer 2 port channel and out-of-band management interfaces, remain up.

In a non-vPC topology, when the Layer 3 and SVI interfaces are down, the redundant Cisco Nexus 5500 Platform switch becomes the active peer for all FHRP groups and it continues to forward traffic.

In a vPC topology, although the SVI interfaces are suspended, the vPC member ports are still up on the Cisco Nexus 5500 Platform switch. Even if the switch has a faulty Layer 3 module, Layer 2 traffic forwarding continues.

Figure 1-4 shows a topology where the Layer 3 module on N5k-2 fails. In this scenario, the Layer 3 connection toward the Layer 3 network and all SVI interfaces are suspended. However, the traffic from the hosts can still be sent to N5k-2 depending on the hash results. With the failure of the Layer 3 module, N5k-2 functions as a Layer 2 switch. It forwards the traffic to N5k-1, which forwards the traffic to the Layer 3 network. The return traffic is sent to N5k-1, which sends the traffic directly to the hosts.

**Figure 1-4 Layer 3 Module Failure****Note**

Only the Layer 3 traffic needs to cross the peer link. The VLAN traffic is switched by N5k-2 locally.

The peer gateway is disabled on both vPC switches if the Layer 3 module fails on one switch.

For topologies with in-band management, the failure of a Layer 3 module means that the connectivity to the management network and the management system is also lost.

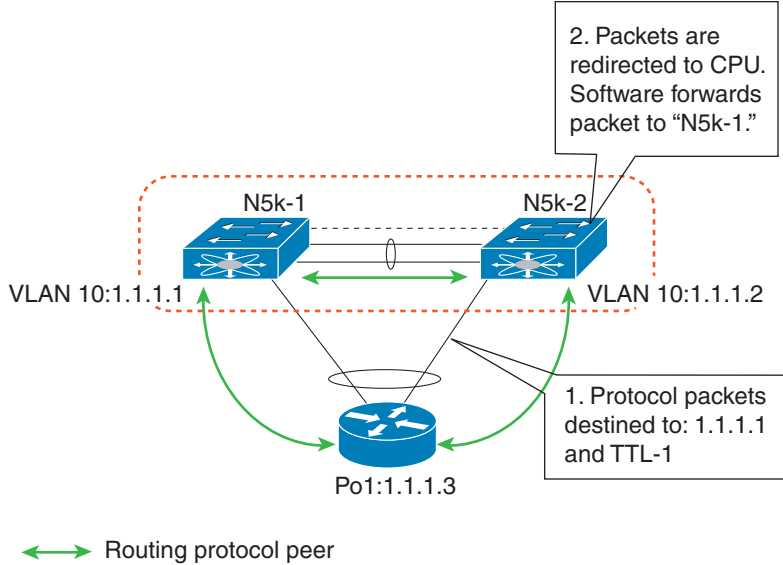
## Connecting to a Router in a vPC Topology

When you connect a router to a pair of Cisco Nexus 5500 Platform switches in a vPC topology and enable routing, traffic forwarding may result in suboptimal traffic paths crossing the peer link similar to the situation described in the [“Layer 3 Forwarding for Packets to a Peer Switch MAC Address”](#) section on page 1-2. We recommend that you use Layer 3 links for connections between the router and the Nexus 5500 switch, instead of a port channel with an IP address.

**Note**

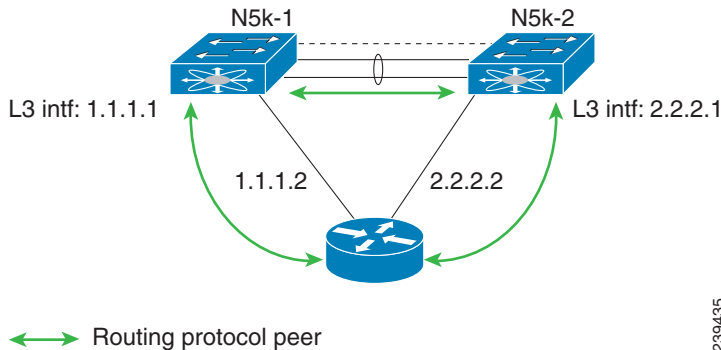
You cannot have a link for non-vPC traffic in parallel with a vPC topology. This can cause errors with the traffic forwarding logic resulting in duplicate or missed packets.

[Figure 1-5](#) illustrates the topology that is not recommended. In this topology, control protocol packets may be hashed by the port channel to the wrong Cisco Nexus 5500 Platform switch, which would then forward the control packets to the correct routing peer (1.1.1.1) in the picture.

**Figure 1-5** Control Traffic Forwarding in a vPC Topology

In this topology, we recommend that you use Layer 3 interfaces instead of vPC interfaces to connect routers to Cisco Nexus 5500 Platform switches whenever possible.

Figure 1-6, shows the recommended topology for connectivity of routers to a vPC domain. The router connects with Layer 3 interfaces 1.1.1.2 and 2.2.2.2 to the two vPC peers and these interfaces are not part of a vPC port channel.

**Figure 1-6** Connecting a Router to a vPC Domain Using Layer 3 Interfaces

## Dedicated VRF For a Keepalive Interface

Beginning in Cisco NX-OS Release 5.0(3)N1(1b), the Cisco Nexus 5500 Platform switch supports VRF lite with a Layer 3 module and Enterprise license and you can create a VRF and assign the interface to a VRF. Prior to this release, two VRFs were created by default: the VRF management and VRF default. The management interface(mgmt0) and all SVI interfaces resided in the VRF management and VRF default respectively.

We recommend that you use an out-of-band management interface (mgmt0) as a vPC keepalive interface although you have the option to use the front-panel data port as a vPC keepalive interface. When you choose to use the front panel 10-Gigabit Ethernet port as the vPC keepalive interface, you should create a separate VRF for vPC keepalive packets when Layer 3 is enabled with vPC. This process eliminates the possibility of disrupting the vPC keepalive link by the wrong routes learned by a dynamic routing protocol.

This example shows how to configure a new VRF named `vpc_keepalive` for the vPC keepalive link and how to display the vPC peer keepalive configuration:

```
vrf context vpc_keepalive
interface Ethernet1/31
  switchport access vlan 123
interface Vlan123
  vrf member vpc_keepalive
  ip address 123.1.1.2/30
  no shutdown
vpc domain 1
  peer-keepalive destination 123.1.1.1 source 123.1.1.2 vrf vpc_keepalive
```

```
layer3-switch# show vpc peer-keepalive
```

```
vPC keep-alive status           : peer is alive
--Peer is alive for            : (154477) seconds, (908) msec
--Send status                  : Success
--Last send at                 : 2011.01.14 19:02:50 100 ms
--Sent on interface            : Vlan123
--Receive status               : Success
--Last receive at              : 2011.01.14 19:02:50 103 ms
--Received on interface        : Vlan123
--Last update from peer        : (0) seconds, (524) msec

vPC Keep-alive parameters
--Destination                   : 123.1.1.1
--Keepalive interval            : 1000 msec
--Keepalive timeout             : 5 seconds
--Keepalive hold timeout        : 3 seconds
--Keepalive vrf                 : vpc_keepalive
--Keepalive udp port            : 3200
--Keepalive tos                 : 192
```

The services provided by the Cisco Nexus 5500 Platform switch, such as Ping, SSH, Telnet, and RADIUS, are VRF-aware. You must specify the VRF name in the CLI in order to use the correct routing table.

```
layer3-switch# ping 123.1.1.1 vrf vpc_keepalive
PING 123.1.1.1 (123.1.1.1): 56 data bytes
64 bytes from 123.1.1.1: icmp_seq=0 ttl=254 time=3.234 ms
64 bytes from 123.1.1.1: icmp_seq=1 ttl=254 time=4.931 ms
64 bytes from 123.1.1.1: icmp_seq=2 ttl=254 time=4.965 ms
64 bytes from 123.1.1.1: icmp_seq=3 ttl=254 time=4.971 ms
64 bytes from 123.1.1.1: icmp_seq=4 ttl=254 time=4.915 ms

--- 123.1.1.1 ping statistics ---
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min/avg/max = 3.234/4.603/4.971 ms
```

## vPC Consistency Check for Layer 3 Parameters

In a vPC topology, vPC peer switches run routing protocols independently and they maintain the routing table independently. Consistency checks are not performed to verify that Layer 3 configurations in the vPC domain are configured symmetrically.

For example, if you configure a router ACL (RACL) on one SVI and you do not configure the router on the corresponding SVI on the vPC peer, a syslog message is not displayed. You must configure the RACL on both devices. This is consistent with the operation of independent routing devices.

Similarly, if you configure peer gateway on one vPC peer and you want the same peer gateway configuration on the other vPC peer, you must configure the peer gateway on the vPC peer.

To confirm that a vPC domain is correctly configured for Layer 3 operations, the following configurations must be consistent:

- SVI configurations
- RACLs
- Routing protocol configurations

## Multicast Interaction in a vPC Topology

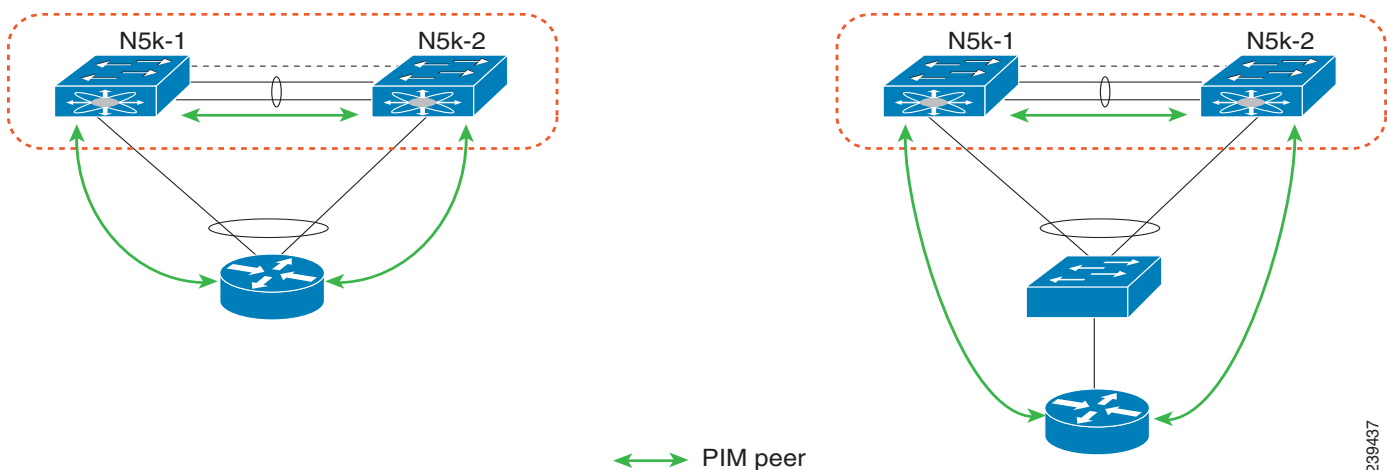
This section includes the following topics:

- [Unsupported Multicast Topology, page 1-8](#)
- [Multicast Routing Table Size, page 1-9](#)

### Unsupported Multicast Topology

Figure 1-7 shows an unsupported multicast topology in a vPC configuration.

Figure 1-7 Unsupported Multicast Topology with a vPC



When a PIM router is connected to Cisco Nexus 5500 Platform switches in a vPC topology, the PIM join messages are received only by one switch. The multicast data might be received by the other switch.



**Note**

Multicast forwarding in this topology does not work.

## Multicast Routing Table Size

When you enable a vPC on a Nexus 5500 Platform switch, one multicast route (\*,G) or (S,G) requires two entries in the routing table; therefore, the multicast routing table size is half the size of what is supported in topologies where vPC is not enabled.

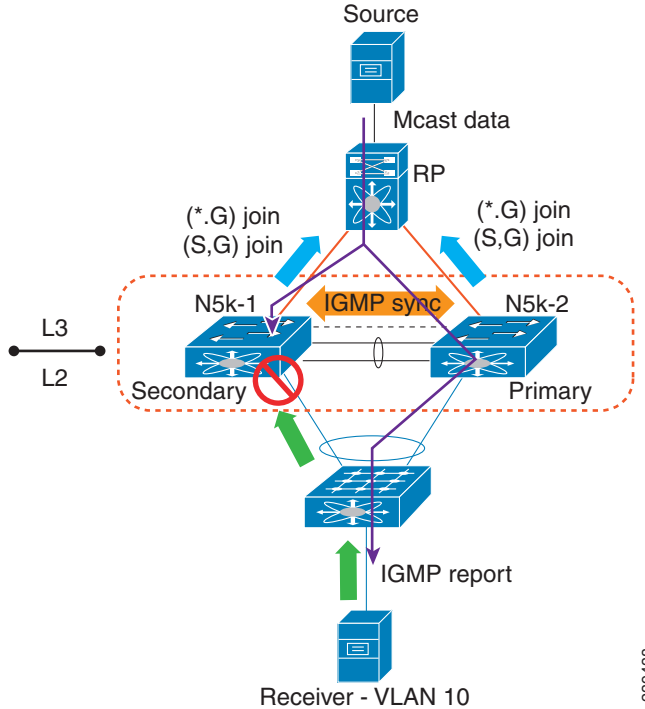
Beginning with Cisco NX-OS Release 5.0(3)N1(1b), the Cisco Nexus 5500 Platform multicast routing table size is 2000 entries in non-vPC topologies and 1000 entries in vPC topologies.

## Faster Convergence with the Prebuilt Source Tree

In a non-vPC topology, only the designated router (DR) can join the source tree. In a vPC topology, when a receiver is connected to a Cisco Nexus 5500 Platform switch or Fabric Extender (FEX) via vPC, both peer switches initiate a PIM (S,G) join toward the source DR. In a topology where both vPC peer switches have equal costs to the source, the vPC primary switch wins the assert and forwards multicast traffic for receivers connected to the Nexus 5500 Platform switch or FEX using the vPC. The vPC secondary switch also joins the source tree and pulls the multicast data. To prevent data duplication, the vPC secondary switch drops the data due to an empty outgoing interface (OIF) list. Once the vPC secondary switch detects the failure of the vPC primary switch, it adds the receiver VLAN to the OIF list and starts to forward the multicast traffic immediately. Because the vPC secondary switch joins the source tree before the failure, it does not need to initiate the (S,G) join and waits for the tree to be built. As a result, it improves the convergence time in the case of a failure with the active multicast traffic forwarder.

[Figure 1-8](#) shows one receiver that is connected to a dual-homed FEX. The source and Rendezvous Point (RP) are in the Layer 3 network. N5k-2, which is the VPC primary switch, is the multicast traffic forwarder for receivers in VLAN 10.

Figure 1-8 vPC Switch as the Receiver Designated Router



This example shows the output of the multicast routing table and VLAN 10 appears in the OIF list of (S,G) entry on N5k-2. N5k-1 joins the source tree but its OIF list remains empty.

```
N5k-1# show ip mroute 224.1.1.1
IP Multicast Routing Table for VRF "default"

(*, 224.1.1.1/32), uptime: 03:03:31, pim ip igmp
  Incoming interface: Ethernet1/6, RPF nbr: 155.1.2.2
  Outgoing interface list: (count: 1)
    Vlan10, uptime: 03:01:16, igmp

(155.1.3.100/32, 224.1.1.1/32), uptime: 02:13:32, ip pim mrrib
  Incoming interface: Ethernet1/6, RPF nbr: 155.1.2.2
  Outgoing interface list: (count: 0)

N5k-2# show ip mroute 224.1.1.1
IP Multicast Routing Table for VRF "default"

(*, 224.1.1.1/32), uptime: 01:48:07, igmp pim ip
  Incoming interface: Ethernet1/6, RPF nbr: 155.1.2.6
  Outgoing interface list: (count: 1)
    Vlan10, uptime: 01:48:07, igmp

(155.1.3.100/32, 224.1.1.1/32), uptime: 01:00:24, ip pim mrrib
  Incoming interface: Ethernet1/6, RPF nbr: 155.1.2.6
  Outgoing interface list: (count: 1)
    Vlan10, uptime: 00:55:14, mrrib
```

The multicast forwarding algorithm applies to all hosts that are connected to the Cisco Nexus 5500 Platform switch or the FEX in a VPC topology, including hosts directly connected to the switch or hosts connected to straight-through FEX topology.

# Using a vPC Switch as a Designated Router (PIM DR)

This section includes the following topics:

- [DR Election and Source Registration, page 1-11](#)
- [Multicast Data Forwarding, page 1-11](#)

## DR Election and Source Registration

In vPC topologies, a DR election occurs based on the DR priority and the IP address. The elected DR is responsible for sending the source registration toward the RP. When multicast traffic from a directly connected source is received by the non-DR peer switch, the peer switch notifies the DR switch using a Cisco Fabric Services (CFS) message about the source and group address. The DR generates source registration packets to the rendezvous point (RP).

## Multicast Data Forwarding

The Cisco Nexus 5500 Platform switch implements a dual-DR mechanism where both vPC peer switches can forward multicast traffic from directly connected sources. The data forwarding rules are as follows:

- The peer switch receives multicast packets from a directly connected source, performs an mroute lookup, and replicates packets for each interface in the OIF list.
- If the OIF is a VLAN trunked over a vPC peer link, one copy is sent over to the peer link for each VLAN that is present in the OIF list. By default, the vPC peer link is considered an mrouter port. Therefore, the multicast packets are sent over to the peer link for each receiving VLAN. You can use the **no ip igmp snooping mrouter vpc-peer link** command to avoid sending multicast traffic over a peer link for each receiver VLAN when there are no orphan ports.

This example shows how to avoid sending the multicast traffic in this scenario:

```
switch-Layer 3-1(config)# no ip igmp snooping mrouter vpc-peer link
Warning: IGMP Snooping mrouter vpc-peer link should be globally disabled on peer VPC
switch as well.
switch-Layer 3-1(config)#
```

With the above CLI configured, the multicast packet is only sent to peer link for VLANs that have orphan ports.

This example shows how to display the list of all orphan ports:

```
switch-Layer 3-1# show vpc orphan-ports
Note:
-----:Going through port database. Please be patient.:-----

VLAN          Orphan Ports
-----
1              Eth1/15
switch-Layer 3-1#
```

**Note**

As of Cisco NX-OS Release 5.0(3)N1(1b), the **no ip igmp snooping mrouter vpc-peer link** command cannot be applied with FEX dual-homed topologies due to a software limitation. The command is used only for interfaces on a Cisco Nexus 5500 Platform switch. This software limitation will be removed in a future software release.

One post-routed multicast packet is sent to a vPC peer link using a reserved VLAN. To configure the reserved VLAN, use the follow commands:

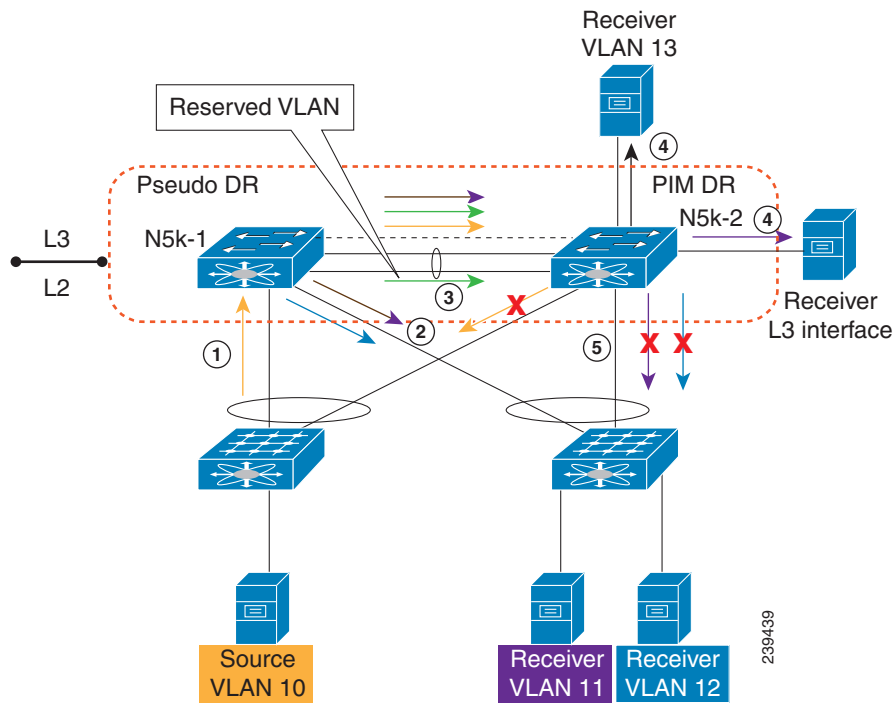
```
switch-Layer 3-1(config)# vpc bind-vrf vrf name vlan VLAN ID
switch-Layer 3-1(config)# vpc bind-vrf default vlan 3000
```

One reserved VLAN is required for each VRF. Without these commands, the receivers in non-vPC VLAN and the receivers connected to Layer 3 interfaces may not receive multicast traffic. The non-vPC VLANs are the VLANs that are not trunked over a peer link.

Multicast traffic that is received over a peer link (with a VLAN ID other than the reserved VLAN ID) is not routed. The multicast traffic is treated as Layer 2 frames that are sent to orphan ports only and not to vPC member ports. The multicast traffic that is received over a peer link with a reserved VLAN ID is routed to a non-vPC VLAN (shown as VLAN 13 in Figure 1-9) and receivers behind the Layer 3 interface. The receivers behind the Layer 3 interface can be hosts directly connected to the Cisco Nexus 5500 Platform switch using Layer 3 interfaces or a router joins the source tree.

Figure 1-9 shows the multicast forwarding rules in a vPC dual-DR topology. In this topology, the source in VLAN 10 and receivers in VLAN 11 and VLAN 12 are the vPC hosts (although in this example they are hosts behind a dual-homed FEX topology where the same rule applies to hosts directly to a Cisco Nexus 5500 Platform switch in a vPC topology). VLAN 13 is a non-vPC VLAN and resides only on N5k-2.

**Figure 1-9 Multicast Data Forwarding**



The forwarding process is as follows:

1. IGMP joins from the hosts are synchronized between the two vPC peer switches. N5k-2 is elected as the PIM DR for VLAN 10. Multicast traffic is sent over to N5k-1.
2. The routing engine of N5k-1 performs an mroute lookup and replicates packets to VLAN 11 and VLAN 12. The data packets for VLAN 11 and VLAN 12 are sent to the FEX which in turn sends packets to the two receivers;
3. By default, the replicated packets are sent to the vPC peer link for the source VLAN as well as each receiver VLAN (VLAN 10, VLAN 11, and VLAN 12) in this example. When you use the **no ip igmp snooping mrouter vpc-peer-link** command, the multicast packets are not sent to the peer link for VLAN 10, VLAN 11, and VLAN 12 because there are no orphan ports. One copy of the packets is sent to the peer link with the reserved VLAN 3000 which was configured using the **vpc bind-vrf default vlan 3000** command.

**Note**

In Cisco NX-OS Release 5.0(3)N1(1b), the **no ip igmp snooping mrouter vpc-peer-link** command cannot be applied with a FEX dual-homed topology.

4. For the multicast traffic received from the peer link, if the VLAN ID is the reserved VLAN ID 3000, the N5k-2 route engine performs a Layer 3 lookup and replicates packets to VLAN 13 (a non-vPC VLAN) and receivers behind Layer 3 interfaces.
5. For the multicast packets received over the peer link, VLAN 10, VLAN 11, and VLAN 12 are dropped by N5k-2 to prevent duplicated packets being sent to the vPC hosts. If any orphan ports are in VLAN 10, VLAN 11, and VLAN 12, the packets are bridged to the orphan ports.

## Nonfunctional Topologies with Layer 3 and vPC Combined

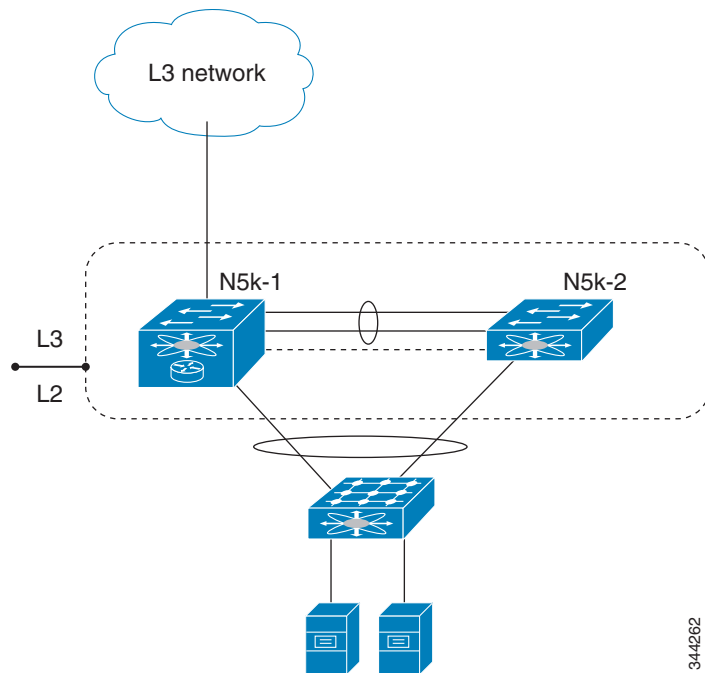
Some Cisco Nexus 5500 Series switch topologies do not work properly when both Layer 3 and vPC are enabled.

### vPC Domain With Layer 3 Enabled on Only One Switch

When two Cisco Nexus 5548/5596 switches are deployed in a vPC domain, both of the switches need to have the same Layer 3 capabilities and the same Layer 3 configuration. The general rule for vPC is that the two devices participating in the vPC domain must have the same functions and capabilities.

[Figure 1-10](#) shows an example of a nonfunctional topology where Layer 3 is enabled on only one switch. When the host sends Layer 3 traffic with the N5k-1 switch's MAC address as the destination MAC address, the traffic could be hashed to the N5k-2 switch. To prevent packet duplication, the traffic received from the peer link is not routed because the assumption is that the peer switch should have routed the traffic and all of the traffic received from the peer link should only be bridged.

**Figure 1-10** Nonfunctional Topology: Layer 3 Enabled on Only One vPC Switch



This mismatched Layer 3 configuration can happen in the following scenarios:

- Only one Cisco Nexus 5000 series switch has a Layer 3 module or only one Cisco Nexus 5000 series switch has the Layer 3 license installed.
- Both Cisco Nexus 5000 switches have a Layer 3 module and the Layer 3 license installed, but only one Cisco Nexus 5000 series switch has SVI configured.
- Both Cisco Nexus 5000 Series switches have an SVI configured, but only one Cisco Nexus 5000 Series switch has the First Hop Redundancy Protocol (FHRP) configured.

In all these scenarios, the traffic forwarding does not work properly. Additionally, we recommend that you have identical configurations for all other Layer 3 parameters, such as Router ACLs (RACLs) and routing protocols.

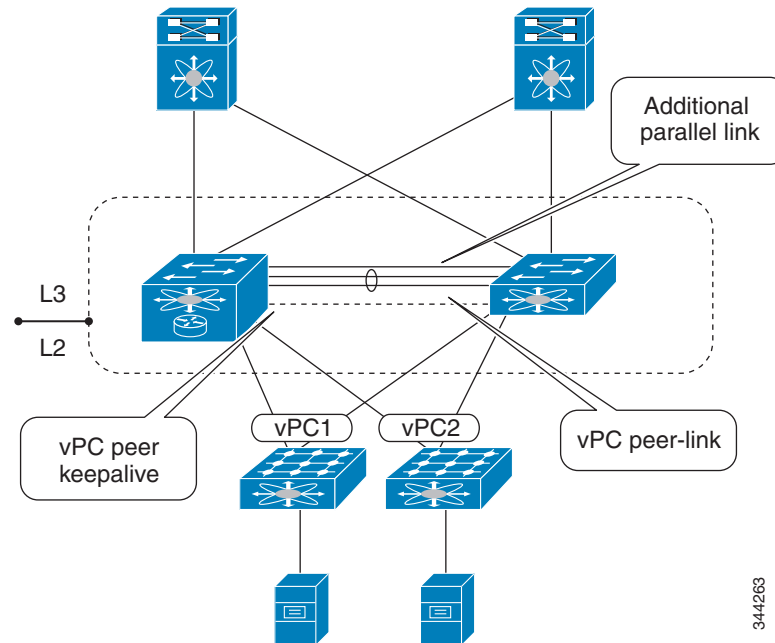
Layer 3 parameters are not part of the vPC consistency check. So, you have to manually verify that the Layer 3 configurations are identical on both Cisco Nexus 5000 Series switches.

## Topology with an Additional Parallel Link Between Two Switches

Figure 1-11 shows the nonfunctional topology where a parallel link in addition to the vPC peer link and the vPC peer-keepalive link between the two switches, and the two switches have Layer 3 enabled. You can have a link between the two switches using the front panel ports for the vPC peer-keepalive link, but this link should only be used to carry vPC-keepalive message.

In some circumstances, you might consider having a separate link between the two vPC switches, either to carry non-vPC VLAN traffic or to form Layer 3 routing protocol peering. While this design is supported on the Cisco Nexus 7000 Series switch, it does not work on the Cisco Nexus 5000 Series switch. With the Cisco Nexus 5000 Series switch, we recommend that you use a vPC peer link for Layer 3 peering to carry both vPC and non-vPC VLAN traffic.

**Figure 1-11 Nonfunctional Topology: Additional Parallel Peer Link Between Switches**

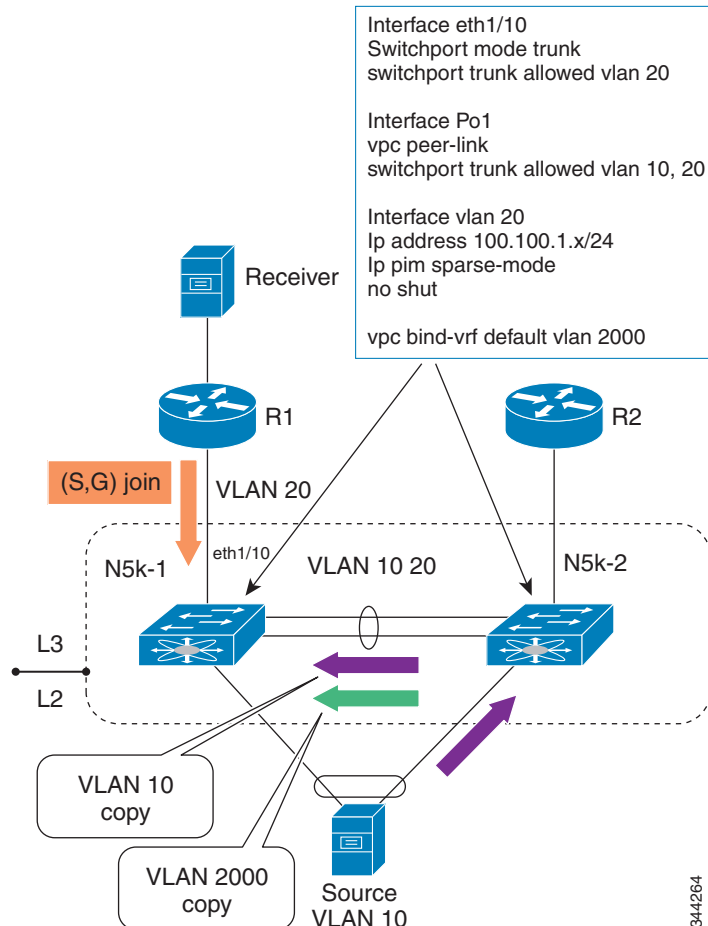


344263

## Connecting a Router Using a VLAN Trunk Port

Figure 1-12 shows the nonfunctional topology where a router is connected to a Cisco Nexus 5000 Series switch using a VLAN trunk port and PIM is enabled for the same VLAN.

Figure 1-12 Nonfunctional Topology: Connecting a Router Using a VLAN Trunk Port



The R1 and R2 routers can be any platform that support both the Layer 2 and Layer 3 functions, such as a Cisco Nexus 7000 Series switch, a Cisco Catalyst 6500 Series switch, or a Cisco Catalyst 4900 Series switch. The intention of this design is to extend a VLAN to all four devices.

From the Layer 3 point of view, the topology has four devices in same VLAN. Let us assume that the PIM (S,G) join message is sent from the R1 router to the N5k-1 switch. It is possible that a source behind the vPC can send traffic to the N5k-2 switch. The N5k-2 switch sends two copies of the multicast packets to the peer link. One copy is for source VLAN 10, and the second copy is for the special VLAN 2000 that is configured with the **vpc bind-vrf default vlan** command. When the N5k-1 switch receives the packet through the peer link for VLAN 10, it only conducts the Layer 2 bridging. That is, it only sends the packet to the orphan ports that reside in VLAN 10. In addition, the N5k-1 switch tries to route the multicast packet for the VLAN 2000 copy that it received from the peer link. In order to prevent packet duplication, the N5k-1 switch only routes the multicast packet to the Layer 3 interface or the non-vPC VLAN. (In this example, VLAN 20 is trunked over the peer-link and is considered to be a vPC VLAN.) Therefore, the N5k-1 switch does not route the multicast packet to the R1 router.

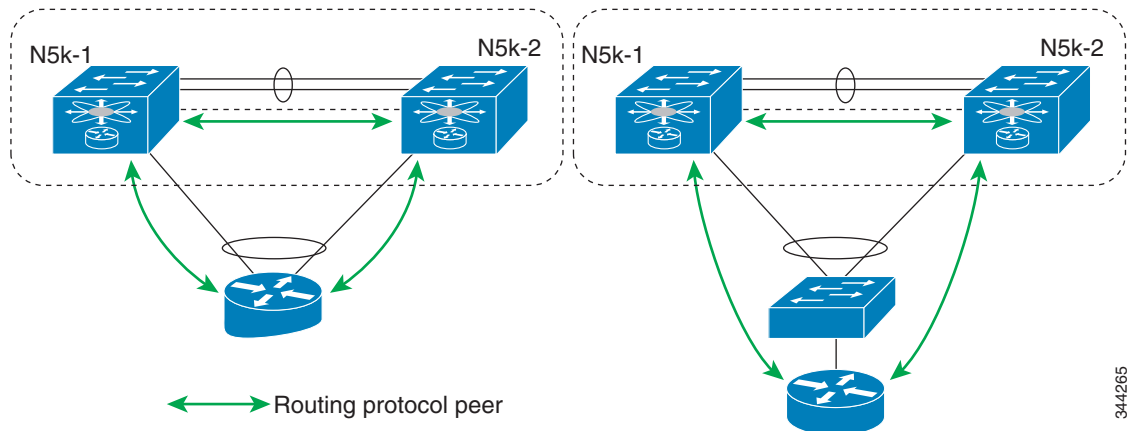
In this scenario, if the intention is to extend the VLAN to all four devices, the alternative design is to not enable Layer 3 on the N5k-1 and N5k-2 switches. Such topology is supported if the vPC is replaced with vPC+, which requires FabricPath. For more information, see the *Cisco Nexus 5000 Series NX-OS FabricPath Operations Guide, Release 5.1(3)N1(1)*.



## Routing Peering Over vPC

Figure 1-13 shows a nonfunctional topology where the dynamic routing protocol is enabled between a router and two Cisco Nexus 5000 Series switches in same vPC domain. The PIM protocol does not work in this topology design, and while the unicast routing protocol allows peering in the vPC, we do not recommend this design. When Enhanced vPC is deployed between a pair of Cisco Nexus 5000 Series switches, the routing peering topology shown in Figure 1-13 is supported. Enhanced vPC requires FabricPath. For more information, see the *Cisco Nexus 5000 Series NX-OS FabricPath Operations Guide, Release 5.1(3)N1(1)*.

Figure 1-13 Nonfunctional Topology: Routing Peering Over vPC



344265

## Software Upgrade and Downgrade Impact

In Cisco NX-OS Release 5.0(3)N1(1b), the Cisco Nexus 5500 Platform switch does not support ISSUs when Layer 3 modules are installed and Layer 3 features are enabled. Use the **install all** command and the **show install all impact** command to determine the impact of the software upgrade and to indicate whether the software upgrade with Layer 3 features enabled will be disruptive and would require a switch and FEX reload.

### show install all impact kickstart

This example shows the output of the **show install all** command:

```
Layer 3-N5548-2# show install all impact kickstart
n5000-uk9-kickstart.5.0.3.N1.0.271.bin.upg system n5000-uk9.5.0.3.N1.0.271.bin.upg

Verifying image bootflash:/n5000-uk9-kickstart.5.0.3.N1.0.271.bin.upg for boot variable
"kickstart".
[#####] 100% -- SUCCESS

Verifying image bootflash:/n5000-uk9.5.0.3.N1.0.271.bin.upg for boot variable "system".
[#####] 100% -- SUCCESS

Verifying image type.
[#####] 50%
[#####] 100% -- SUCCESS
```

```
Extracting "system" version from image bootflash:/n5000-uk9.5.0.3.N1.0.271.bin.upg.
[#####] 100% -- SUCCESS
```

```
Extracting "kickstart" version from image
bootflash:/n5000-uk9-kickstart.5.0.3.N1.0.271.bin.upg.
[#####] 100% -- SUCCESS
```

```
Extracting "bios" version from image bootflash:/n5000-uk9.5.0.3.N1.0.271.bin.upg.
[#####] 100% -- SUCCESS
```

```
Extracting "fexth" version from image bootflash:/n5000-uk9.5.0.3.N1.0.271.bin.upg.
[#####] 100% -- SUCCESS
```

```
Performing module support checks.
[#####] 100% -- SUCCESS
```

```
Notifying services about system upgrade.
[#####] 100% -- SUCCESS
```

Compatibility check is done:

Module	bootable	Impact	Install-type	Reason
1	yes	<b>disruptive</b>	reset	Non-disruptive install not supported if Layer 3 was enabled
100	yes	<b>disruptive</b>	reset	Non-disruptive install not supported if Layer 3 was enabled

Images will be upgraded according to following table:

Module	Image	Running-Version	New-Version	Upg-Required
1	system	5.0(3)N1(1b)	5.0(3u)N1(1u)	yes
1	kickstart	5.0(3)N1(1b)	5.0(3u)N1(1u)	yes
1	bios	v3.4.0(01/13/2011)	v3.4.0(01/13/2011)	no
100	fexth	5.0(3)N1(1b)	5.0(3u)N1(1u)	yes
1	power-seq	v3.0	v3.0	no
2	power-seq	v1.0	v1.0	no
1	uC	v1.0.0.14	v1.0.0.14	no

Layer 3-N5548-2#

You can perform a nondisruptive ISSU from an earlier release to NX-OS Release 5.0(3)N1(1b) when upgrading without Layer 3 features enabled.

## show spanning-tree issu-impact

To verify that the current STP topology is consistent with ISSU requirements, use the **show spanning-tree issu-impact** command to display the STP configuration and whether or not there are potential STP issues.

This example shows how to display information about the STP impact when performing an ISSU:

```
nexus5010# show spanning-tree issu-impact
```

For ISSU to Proceed, Check the Following Criteria :

1. No Topology change must be active in any STP instance
2. Bridge assurance(BA) should not be active on any port (except MCT)
3. There should not be any Non Edge Designated Forwarding port (except MCT)
4. ISSU criteria must be met on the VPC Peer Switch as well

Following are the statistics on this switch

No Active Topology change Found!  
Criteria 1 PASSED !!

No Ports with BA Enabled Found!  
Criteria 2 PASSED!!

No Non-Edge Designated Forwarding Ports Found!  
Criteria 3 PASSED !!

ISSU Can Proceed! Check Peer Switch.

For information on upgrade procedures, see the *Cisco Nexus 5000 Series NX-OS Upgrade and Downgrade Guide*.





# Using Enhanced vPC

---

This chapter provides an overview of Enhanced virtual port channelling(vPC).

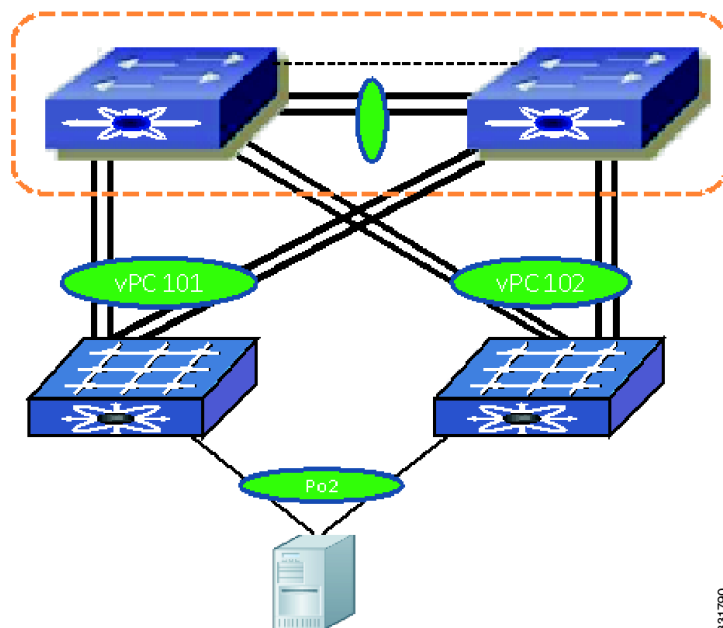
This chapter includes the following sections:

- [Information About Enhanced vPC, page 1-1](#)
- [Enhanced vPC Topology and Scalability, page 1-4](#)
- [Enhanced vPC Scalability, page 1-7](#)
- [Enhanced vPC with FCoE, page 1-8](#)
- [Enhanced vPC Failure Reaction, page 1-10](#)
- [Deploying and Monitoring Enhanced vPC, page 1-11](#)

## Information About Enhanced vPC

Enhanced vPC enables you to support server connectivity with one topology and address requirement for both high availability and high bandwidth. Enhanced vPC is technology that supports the topology that is shown in [Figure 1-1](#), where a Cisco Nexus 2000 Fabric Extender (FEX) is dual-homed to a pair of Cisco Nexus 5500 Series devices while the hosts are also dual-homed to a pair of FEXs using a vPC.

Figure 1-1 Enhanced vPC Topology

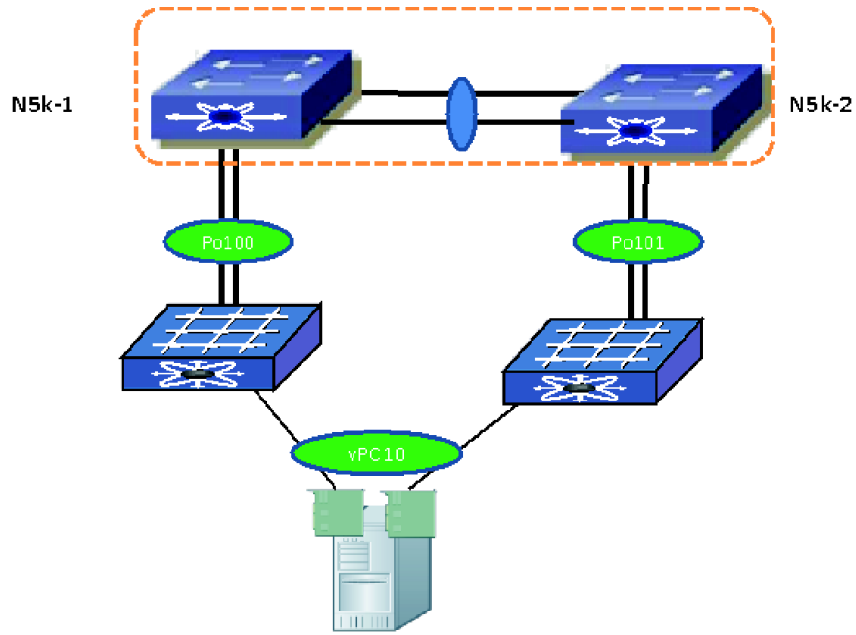


With Enhanced vPC, all available paths from hosts to FEXs and from FEXs to the Cisco Nexus 5500 Series device are active, carry Ethernet traffic, and maximize the available bandwidth. All available paths in the Enhanced vPC topology can carry Ethernet traffic.

With Enhanced vPC, you could choose either a single-homed FEX topology see [Figure 1-2](#) or a dual-homed FEX topology see [Figure 1-3](#) for a dual-homed FEX topology example.

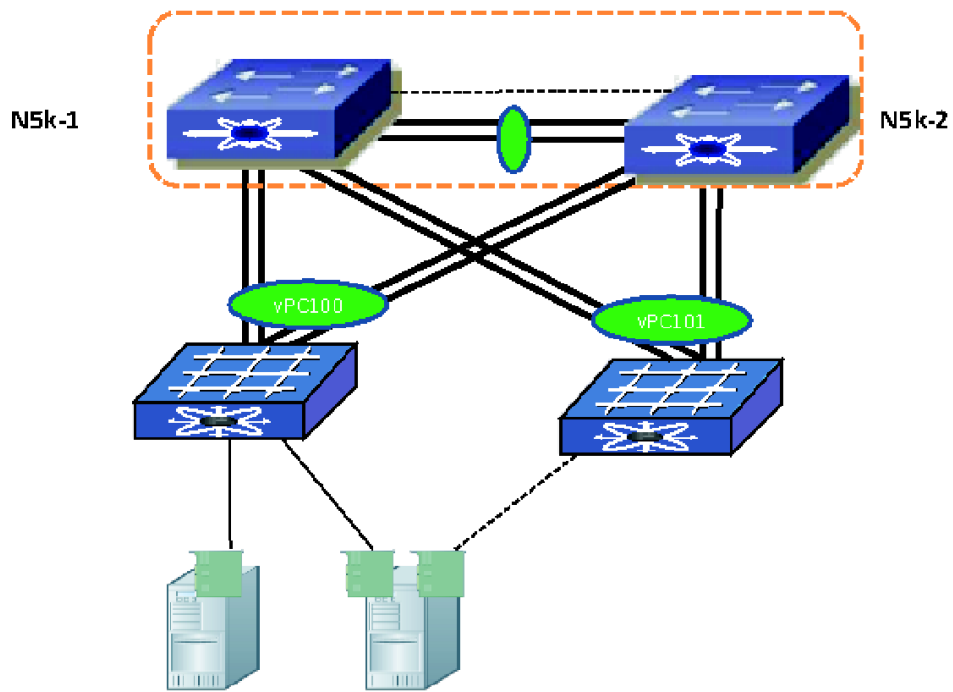
The single-homed FEX topology is well suited for servers with multiple NICs that support 802.3ad port channel. The dual-homed FEX topology is ideal for servers with one NIC, because the failure of one Cisco Nexus 5500 Series device does not bring down the FEX and does not cut the single NIC server out of the network. The Dual-homed FEX topology can also be deployed for servers that have multiple NICs but do not support 802.3ad. Without an Enhanced vPC server, you cannot connect port channels to FEXs when the FEXs are dual-homed to both Cisco Nexus 5500 Series devices.

**Figure 1-2 Single-Homed FEX Topology**



331791

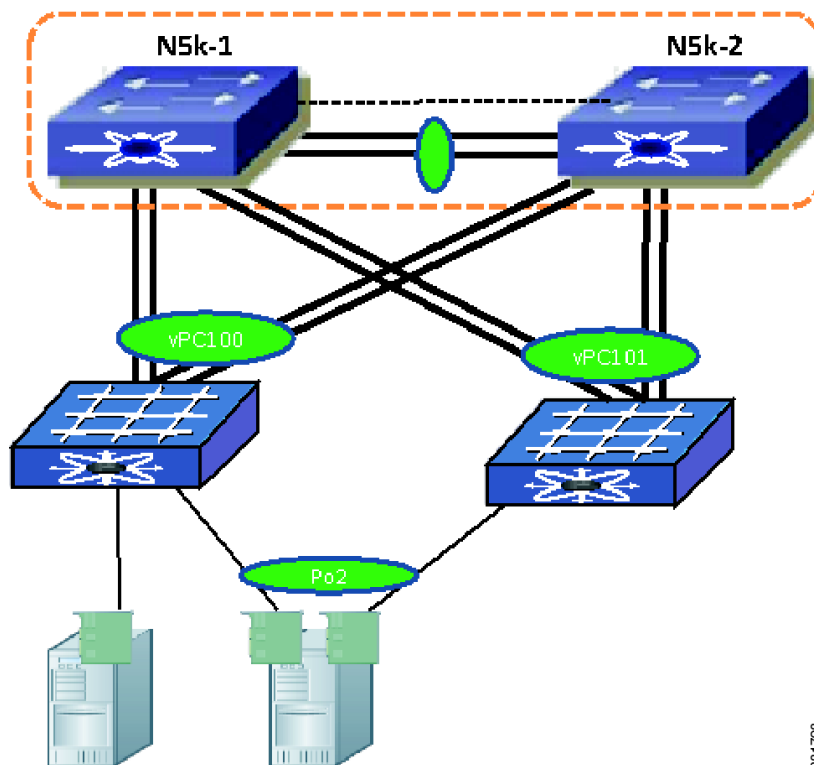
**Figure 1-3 Dual-Homed FEX Topology**



331792

Use Enhanced vPC if you want to connect a dual-homed FEX to a Cisco Nexus 5000 Series device to have better redundancy for a single NIC server and at the same time, want to run port channels between the servers that have multiple NICs and FEXs. See Figure 1-4 for an example of enhanced vPC with a single-homed server and host vPC.

Figure 1-4 Enhanced vPC with Single-Homed Server and Host vPC



## Supported Platform

Cisco NX-OS 5.1(3)N1(1) supports Enhanced vPC on the Cisco Nexus 5548P, Cisco Nexus 5548UP, and Cisco Nexus 5596UP devices. The Cisco Nexus 5010 and Cisco Nexus 5020 devices cannot support Enhanced vPC. Enhanced vPC is implemented on the Cisco Nexus 5500 Series device and has no specific requirements from a FEX. As a result, different types of FEXs can be deployed in an Enhanced vPC topology.

Enhanced vPC is also supported with Layer 3 running on Cisco Nexus 5500 Series devices, but it does not change any Layer 3 CLI or how Layer 3 features are implemented on a Cisco Nexus 5500 Series device.

With the introduction of the Enhanced vPC, the three topologies supported by a Cisco Nexus 5500 Series device and FEX are the single-homed FEX topology, the dual-homed FEX topology, and the Enhanced vPC topology. A hybrid topology is also supported, where the FEX and hosts are connected to the same pair of Cisco Nexus 5500 Series devices.

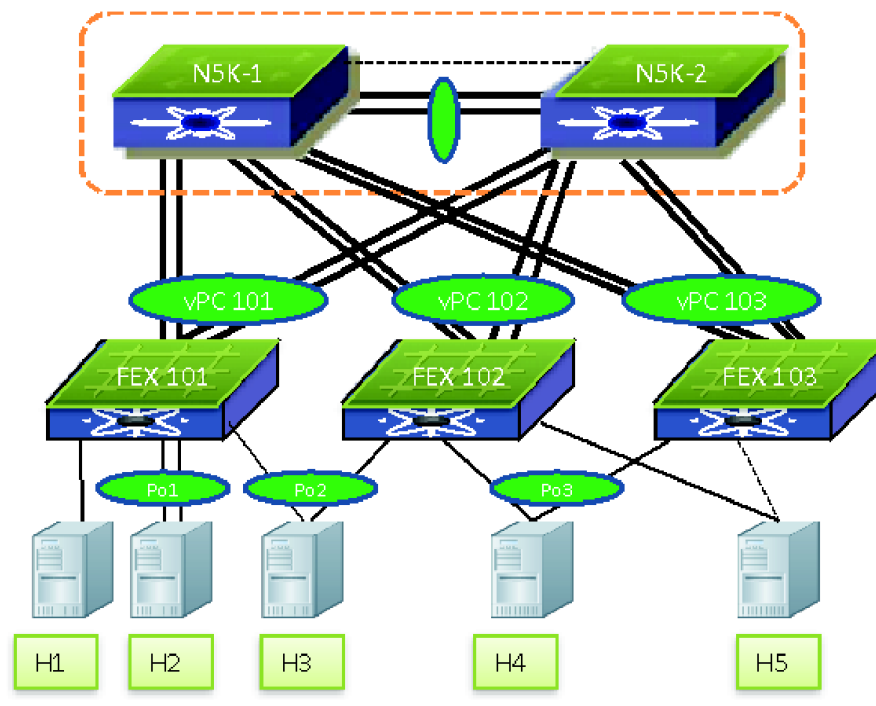
## Enhanced vPC Topology and Scalability

### Supported Enhanced vPC Topology

Figure 1-5 shows various server connections that are supported with an Enhanced vPC topology.



**Figure 1-5** Enhanced vPC with Various Types of Server Connections



The figure shows the following configurations:

- A single-homed server — Provides a failover path when one of the Cisco Nexus 5000 Series devices or the link between a Cisco Nexus 5000 Series device and a FEX fails. (Shown as H1 in the [Figure 1-5](#).)
- A dual-homed server runs a port channel to the same FEX — (Shown as H2 in the [Figure 1-5](#).)
- A dual-homed server runs port channel to two FEXs — Support both static port channel and LACP based port channels. The port channel members can span up to two FEXs. (Shown as H3 in the figure.)
- A dual-homed server runs a port channel to any two FEXs connected to the same pair of Cisco Nexus 5000 Series devices — Enhanced vPC topology supports the port channels between the host and any randomly chosen FEXs connected to the same pair of Cisco Nexus 5000 Series devices. As shown in the [Figure 1-5](#), while host H3 runs a port channel to ports from FEX 101 and FEX 102, host H4 can run a port channel to ports from FEX 102 and FEX 103, which implies that a port channel can span to any two line cards. This configuration is useful when the Cisco Nexus 5000 Series device and the FEX are deployed as EoR devices where all the FEXs are installed in the network. You do not need to track which FEXs are considered as a pair of FEXs for a host vPC connection and they can connect the server to any two FEXs where ports are available.
- A dual-homed server runs the Fiber Channel over Ethernet (FCoE) and a port channel for Ethernet — The Enhanced vPC topology supports FCoE connectivity to servers. The [Figure 1-5](#), hosts H3 and H4 can use CNA to connect to a Cisco Nexus 2232PP to run FCoE for storage traffic and a port channel for Ethernet traffic. See [FEX Uplink Traffic Load, page 1-9](#) for details about how FC traffic is handled and how the FC traffic isolation for [SAN A and SAN B Traffic Isolation, page 1-8](#) are implemented in the Enhanced vPC topology.

331794

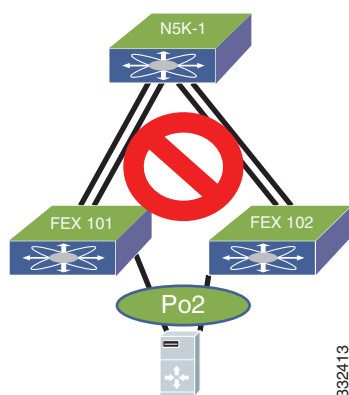
- A dual-homed server runs active/standby NIC teaming — This configuration is similar to the Dual-homed FEX topology. The server can run active or standby NIC and connect to two or more FEXs. See H5 in the [Figure 1-5](#).
- In addition to the various types of connections, the same pair of Cisco Nexus 5000 Series devices can also support the single-homed FEX topology. This kind of combination is referred to as a hybrid topology.

## Unsupported Enhanced vPC Topology

### vPC Between Hosts and a Pair of FEXs that are Connected to a Single Cisco Nexus 5500 Series Device

[Figure 1-6](#) shows unsupported topology where a vPC is between hosts and two FEXs that are connected to one Cisco Nexus 5500 Series device. This topology does not provide a good high availability solution because the server loses the connectivity to the network when the Cisco Nexus 5000 Series device fails.

**Figure 1-6** *Unsupported Topology—Host vPC With One Cisco Nexus 5000 Series Device*

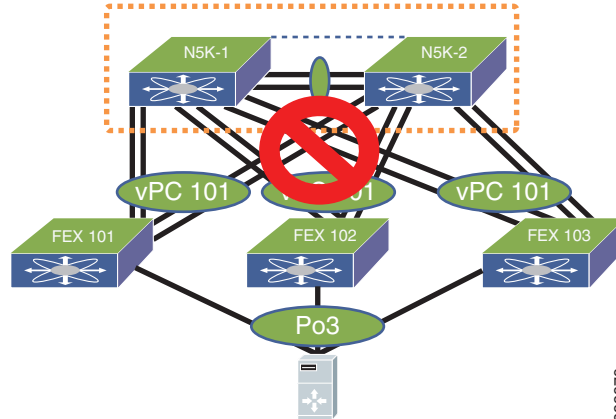


If you need to connect a multi-homing server to a pair of FEXs when there is only one Cisco Nexus 5000 Series device, you have the option to run active or standby NIC teaming from the server.

## Port Channel Between Host and Ports from More Than Two FEXs

With Enhanced vPC, the port channel can be formed among ports from up to two FEXs that are connected to the same pair of Cisco Nexus 5000 Series devices. This topology which is shown in [Figure 1-7](#), does not work and is not supported.

**Figure 1-7** *Unsupported Topology—Host vPC Spans Across More Than Two FEXs*



This topology adds little value in terms of high availability but increases the complexity of cabling and management. The CLI rejects the configuration when it detects the port channel members are from more than two FEXs.

## Enhanced vPC Scalability

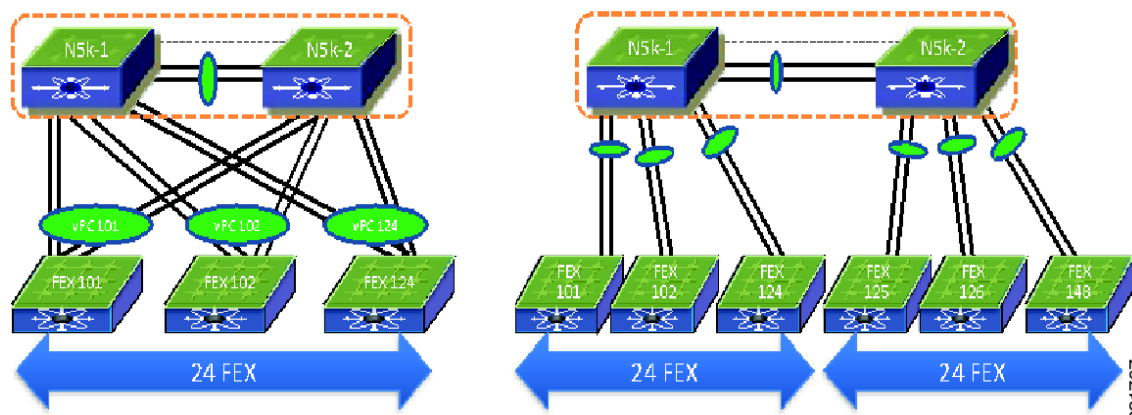
In general, the Enhanced vPC does not change the scalability of a Cisco Nexus 5000 Series device and a FEX. Scalability is similar to the dual-homed FEX topology.

### Total Number of FEXs Per Cisco Nexus 5000 Series Device

From the Cisco NX-OS 5.1(3)N1(1) release and later releases, each Cisco Nexus 5500 Series device can manage and support up to 24 FEXs without Layer 3. With Layer 3, the number of FEXs supported per Cisco Nexus 5500 Series device is 8. With Enhanced vPC and a dual-homed FEX topology each FEX is managed by both Cisco Nexus 5000 Series devices. As a result, one pair of Cisco Nexus 5500 Series devices can support up to 24 FEXs and 16 FEXs for Layer 2 and Layer 3.

There are differences in scalability between the straight-through topology, the dual-homed FEX topology, and the Enhanced vPC topology. In the straight through topology, only one Cisco Nexus 5000 Series device manages each FEX and a pair of Cisco Nexus 5500 Series devices manage up to 48 FEXs. This difference is shown in [Figure 1-8](#) for an Layer 2 scenario.

Figure 1-8 FEX Scalability



Because the total number of FEXs that are supported by a pair of Cisco Nexus 5000 Series devices is different between these two topologies, the FEX straight-through design with more than 24 FEXs per one pair of Cisco Nexus 5000 Series devices cannot migrate to Enhanced vPC topology.

The configurations are as follows:

- Total number of host vPC — With Enhanced vPC, each FEX port can be part of a host vPC. The host vPC does not consume port channel resources on the parent Cisco Nexus 5000 Series device.
- Total number of ports per host vPC — The total number of ports that can be assigned to each host vPC differs with each FEX model.

The Cisco Nexus 2148 device does not support port channels. With the Cisco Nexus 2148 device, the host vPC can have up to two ports with one from each Cisco Nexus 2148.

The Cisco Nexus 2248, Cisco Nexus 2224, Cisco Nexus 2232 and Cisco Nexus 2248TP-E devices support hardware port channels and up to 16 ports in a host vPC with up to 8 ports from each FEX.

## Enhanced vPC with FCoE

### SAN A and SAN B Traffic Isolation

You can deploy FCoE in an Enhanced vPC topology. Traditionally, the SAN network maintains two fabrics, the SAN A and SAN B. The traffic from side A is isolated from side B. Hosts and storage arrays are attached to host SAN networks for high availability. The FCoE traffic in an Enhanced vPC topology maintain the traffic isolation for two SAN networks. The FCoE command ensures the FCoE traffic from the FEX is only sent to one Cisco Nexus 5000 Series device as follows:

```
N5k-1(config)# fex 101
N5k-1(config-fex)# fcoe
```

```
N5k-2(config)# fex 102
N5k-2(config-fex)# fcoe
```

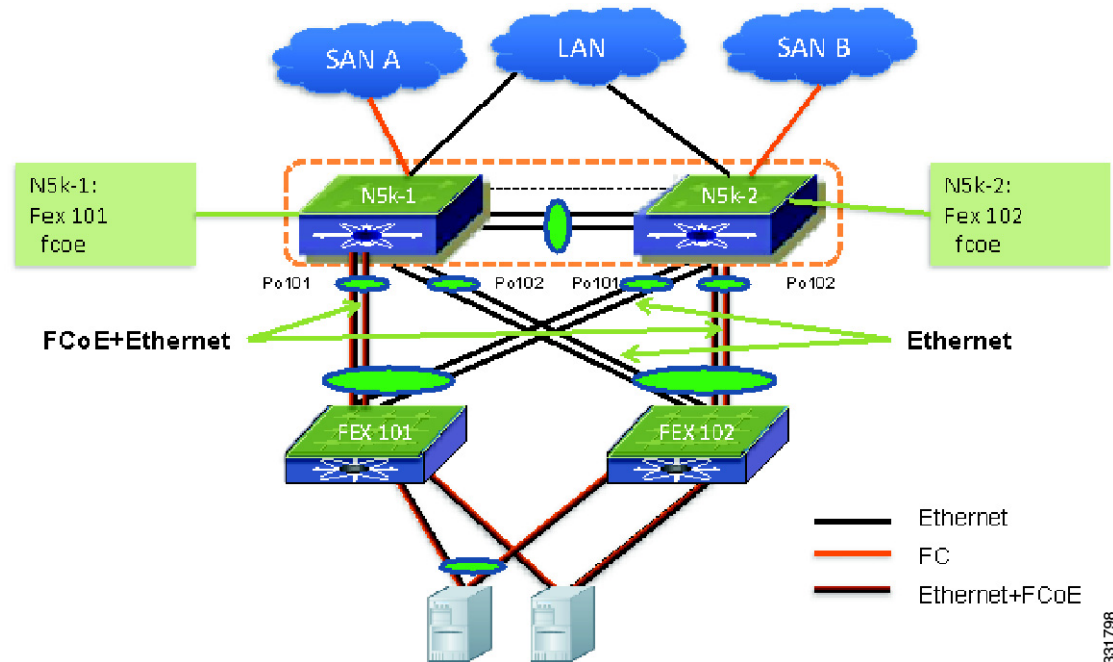
In this configuration, the FCoE traffic from FEX 101 is sent only to N5k-1 and the FCoE traffic from FEX 102 is sent only to N5k-2 although both FEXs are connected to both devices. This is true for the reverse FCoE traffic from a Cisco Nexus 5000 Series device to a FEX where the FCoE traffic is sent only on a FEX. As a result, SAN A and SAN B separation is achieved.

Beginning with Cisco NX-OS 5.1(3)N1(1) release, FCoE is supported for an Enhanced vPC topology and dual-homed FEX topology.

For an Ethernet-only network with Enhanced vPC, you have the option to connect hosts to random selected pairs of FEXs, such as the host H3 and H4 in [Figure 1-9](#). However, this approach is not recommended because it may have hosts connected to the two FEXs that are mapped to the same side of a SAN network.

From the Cisco NX-OS 5.1(3)N1(1) release and later releases, the topology that has four ports in the host vPC (two ports to each FEX) is not supported.

**Figure 1-9 FCoE Traffic Flow with Enhanced vPC**



## FEX Uplink Traffic Load

In the Enhanced vPC topology, the traffic load for the FEX uplinks is not even, because the FCoE traffic from a FEX is forwarded to one Cisco Nexus 5000 Series device for SAN traffic isolation. Po101 between N5k-1 and FEX 101, and Po102 between N5k-2 and FEX 102 carries more traffic than the rest of two the port channels between a FEX and a Cisco Nexus 5000 Series device. You must provision enough bandwidth to prevent the undesirable oversubscription for links that carry both FCoE and Ethernet traffic.

To avoid the imbalance of traffic distribution, we recommend a FEX straight-through topology for FCoE deployments that have a FEX. In an enhanced vPC topology, you can have only up to four 10-Gigabyte Ethernet links for FCoE traffic for each of the Cisco Nexus 2232 devices. However, in a FEX straight-through topology, all eight 10-Gigabyte Ethernet uplinks can carry FCoE traffic.

You can control how the bandwidth of a FEX uplink is shared between the Ethernet and FCoE traffic by configuring an ingress and egress queuing policy. The default QoS template allocates half of the bandwidth to each. For the links that carry both Ethernet and FCoE traffic, each gets half of the

guaranteed bandwidth in case of congestion. When there is congestion, each type of traffic can take all available bandwidth. For links that carry only Ethernet traffic, all 10-Gbps bandwidth is available for Ethernet traffic.

## Enhanced vPC Failure Reaction

### Port Channel Member Port Failure

If one port channel member fails, the traffic flow is moved to the remaining port channel members. If the host loses all its connections to one FEX, the traffic flow is redirected to another flow for both host to network and network to host.

### FEX Failure

If a FEX fails, all flows are moved to the second FEX in the Enhanced vPC topology. For both directions, the traffic does not need to traverse the vPC peer link.

### Cisco Nexus 5000 Series Switch Failure

When one Cisco Nexus 5000 Series device goes down, all FEXs remain connected to the other Cisco Nexus 5000 Series devices. All FEX front panel ports are still operational. All traffic flows continue to be forwarded by all FEXs.

### FEX Uplink Failure

When a FEX loses its uplinks, it shuts down its front panel ports and the traffic is carried by another FEX for Enhanced vPC topology.

### vPC Peer-Link Failure

When the vPC secondary device detects a peer-link failure it checks if the primary device is alive via a peer keepalive link. If the primary device is alive, the secondary device suspends all its vPC member ports. In an Enhanced vPC topology, the vPC secondary device suspends all the interfaces that connects to a FEX. As a result, all FEXs are connected only to a vPC primary device. All FEX host ports are up and running and traffic continues to be distributed to both FEXs.

Because the FEXs are not connected to a vPC secondary device when the peer-link fails, the vPC secondary device cannot carry the FCoE traffic and the secondary device shuts down all the VFC interfaces that are bound to the FEX host ports. The multi pathing software that runs on the host moves all the SAN traffic flow to the remaining VFC interface.

If the secondary device cannot reach the primary device via the keepalive link, the secondary device keeps its vPC member ports up and running.

## vPC Keepalive Failure

The vPC keepalive failure does not have any impact on the vPC and traffic flow. We recommend that you check and restore the keepalive link as soon as possible.

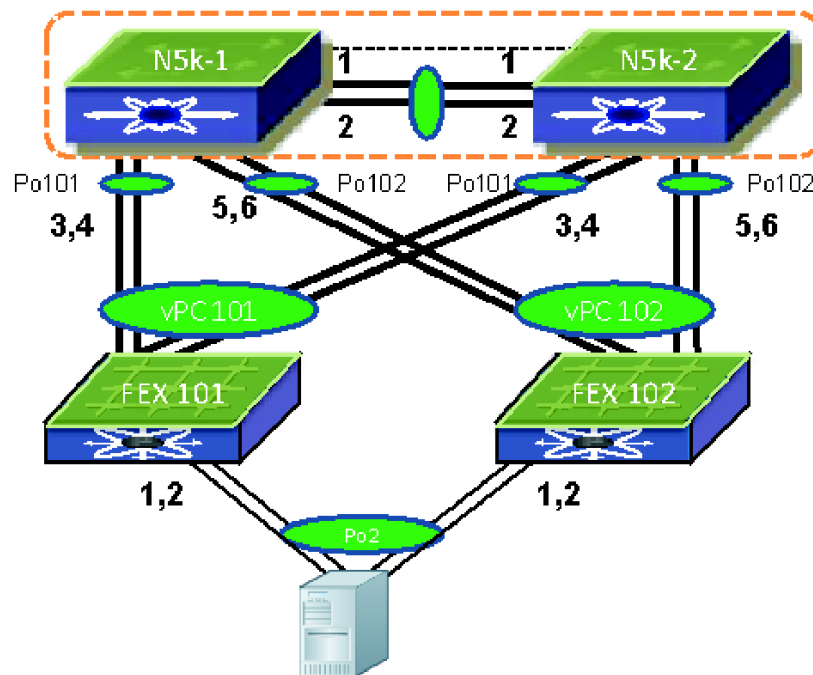
# Deploying and Monitoring Enhanced vPC

## Enhanced vPC Configuration

In the Enhanced vPC topology, the FEXs are virtual line cards and the FEX front panel ports are mapped to the virtual interfaces on a parent Cisco Nexus 5000 Series device. From the CLI perspective, the configuration of Enhanced vPC is the same as a regular port channel with member ports from two FEXs. You do not have to enter the CLI `vpc vpc ID` to create an Enhanced vPC. An example of how to create an Enhanced vPC with topology.

The following procedure uses the topology in [Figure 1-10](#). In the figure, the number next to the line is the interface ID. Assume all the ports are base ports the interface ID 2 represent interface `eth1/2` on the Cisco Nexus 5000 Series device.

**Figure 1-10** Creating an Enhanced vPC Topology



331799

### Configuration on the first Cisco Nexus 5000 Series device:

```
N5k-1(config)# interface eth101/1/, eth101/1/2
N5k-1(config-if)# channel-group 2 mode active
N5k-1(config-if)# interface eth102/1/, eth102/1/2
N5k-1(config-if)# channel-group 2 mode active
```

### Configuration from the second Cisco Nexus 5000 Series device:

```
N5k-2 (config)# interface eth101/1/, eth101/1/2
N5k-2 (config-if)# channel-group 2 mode active
N5k-2 (config-if)# interface eth102/1/, eth102/1/2
N5k-2 (config-if)# channel-group 2 mode active
```

Although the `vPC vPC ID` command is not required, the software assigns an internal vPC ID for each Enhanced vPC. The output of the `show vpc` command displays this internal vPC ID.

---

**Step 1** Enable a vPC and LACP.

```
N5k-1 (config)# feature vpc
N5k-1 (config)# feature lacp
N5k-2 (config)# feature vpc
N5k-2 (config)# feature lacp
```

**Step 2** Create VLANs.

```
N5k-1 (config)# vlan 10-20
N5k-2 (config)# vlan 10-20
```

**Step 3** Assign the vPC domain ID and configure the vPC peer keepalive.

```
N5k-1 (config)# vpc domain 123
N5k-1 (config-vpc)# peer-keepalive destination 172.25.182.100

N5k-2 (config)# vpc domain 123
N5k-2 (config-vpc)# peer-keepalive destination 172.25.182.99
```

**Step 4** Configure the vPC peer-link.

```
N5k-1 (config)# interface eth1/1-2
N5k-1 (config-if)# channel-group 1 mode active
N5k-1 (config-if)# interface Po1
N5k-1 (config-if)# switchport mode trunk
N5k-1 (config-if)# switchport trunk allowed vlan 1, 10-20
N5k-1 (config-if)# vpc peer-link

N5k-2 (config)# interface eth1/1-2
N5k-2 (config-if)# channel-group 1 mode active
N5k-2 (config-if)# interface Po1
N5k-2 (config-if)# switchport mode trunk
N5k-2 (config-if)# switchport trunk allowed vlan 1, 10-20
N5k-2 (config-if)# vpc peer-link
```

**Step 5** Configure FEX 101.

```
N5k-1 (config)# fex 101
N5k-1 (config-fex)# interface eth1/3-4
N5k-1 (config-if)# channel-group 101
N5k-1 (config-if)# interface po101
N5k-1 (config-if)# switchport mode fex-fabric
N5k-1 (config-if)# vpc 101
N5k-1 (config-if)# fex associate 101

N5k-2 (config)# fex 101
N5k-2 (config-fex)# interface eth1/3-4
N5k-2 (config-if)# channel-group 101
N5k-2 (config-if)# interface po101
N5k-2 (config-if)# switchport mode fex-fabric
N5k-2 (config-if)# vpc 101
N5k-2 (config-if)# fex associate 101
```

**Step 6** Configure FEX 102.



```

N5k-1(config)# fex 102
N5k-1(config-fex)# interface eth1/5-6
N5k-1(config-if)# channel-group 102
N5k-1(config-if)# interface po102
N5k-1(config-if)# switchport mode fex-fabric
N5k-1(config-if)# vpc 102
N5k-1(config-if)# fex associate 102

N5k-2(config)# fex 102
N5k-2(config-fex)# interface eth1/5-6
N5k-2(config-if)# channel-group 102
N5k-2(config-if)# interface po102
N5k-2(config-if)# switchport mode fex-fabric
N5k-2(config-if)# vpc 102
N5k-2(config-if)# fex associate 102

```

### Step 7 Create Enhanced vPC.

```

N5k-1(config)# interface eth101/1/1, eth101/1/2
N5k-1(config-if)# channel-group 2 mode active
N5k-1(config-if)# interface eth102/1/1, eth102/1/2
N5k-1(config-if)# channel-group 2 mode active
N5k-1(config-if)# int po2
N5k-1(config-if)# switchport access vlan 10

N5k-2(config)# interface eth101/1/1, eth101/1/2
N5k-2(config-if)# channel-group 2 mode active
N5k-2(config-if)# interface eth102/1/1, eth102/1/2
N5k-2(config-if)# channel-group 2 mode active
N5k-2(config-if)# int po2
N5k-2(config-if)# switchport access vlan 10

```

As shown in the above procedure, the Enhanced vPC configuration is the same configuration as when you configure the host port channel with channel members from the same FEX.

## Enhanced vPC Consistency Checks

Cisco NX-OS checks whether the vPC-related configuration is consistent between the two vPC peer devices to avoid undesired data forwarding behavior. Cisco NX-OS checks both global configuration parameters and interface level configuration parameters. For Enhanced vPC, the consistency check for global configuration parameters remain the same as for a dual-homed FEX topology. For details about the vPC consistency check, see the vPC operation guide:

[http://www.cisco.com/en/US/docs/switches/datacenter/nexus5000/sw/operations/n5k\\_vpc\\_ops.html](http://www.cisco.com/en/US/docs/switches/datacenter/nexus5000/sw/operations/n5k_vpc_ops.html)

## Port Channel ID Checks

Cisco NX-OS requires that the same port channel ID is used on two peer devices for Enhanced vPC. The port channel and its channel members are suspended when different port channel IDs are used for the same FEX ports. The following example shows that the FEX interfaces eth110/1/1 and eth111/1/1 are assigned to different port channels on the two vPC devices. As a result, the two FEX interfaces are suspended on both Cisco Nexus 5000 Series devices and the port channel is not operational.

```
N5596-1# show run int e110/1/1,e111/1/1
```

```
!Command: show running-config interface Ethernet110/1/1, Ethernet111/1/1
```

```

!Time: Sun Aug 28 03:38:23 2011

version 5.1(3)N1(1)

interface Ethernet110/1/1
  channel-group 1002

interface Ethernet111/1/1
  channel-group 1002

N5596-2# show run int e110/1/1,e111/1/1

!Command: show running-config interface Ethernet110/1/1, Ethernet111/1/1
!Time: Mon Aug 29 21:01:20 2011

version 5.1(3)N1(1)

interface Ethernet110/1/1
  hardware N2348TP queue-limit 1024000 rx
  hardware N2348TP queue-limit 1024000
  switchport access vlan 20
  channel-group 1001

interface Ethernet111/1/1
  switchport access vlan 20
  channel-group 1001

N5596-2#

N5596-2# show int e110/1/1
Ethernet110/1/1 is down (suspended by vpc)
  Hardware: 100/1000 Ethernet, address: 7081.0500.2402 (bia 7081.0500.2402)
  MTU 1500 bytes, BW 1000000 Kbit, DLY 10 usec
<snip>

N5596-1# show int e110/1/1
Ethernet110/1/1 is down (suspended by vpc)
  Hardware: 100/1000 Ethernet, address: 7081.0500.2402 (bia 7081.0500.2402)
  MTU 1500 bytes, BW 1000000 Kbit, DLY 10 usec

<snip>

```

## Different Port Channel Members

The port channel is up and operational when there is at least one common port channel member between the two Cisco Nexus 5000 Series devices. The FEX interfaces that are assigned to A port channel on only one Cisco Nexus 5000 Series device are suspended. In the following example, FEX interfaces eth110/1/1 and eth111/1/1, are assigned to Po1001 on N5596-1. However, on N5596-2 only eth110/1/1 is assigned to Po1001. Therefore, only eth110/1/1 becomes the active port channel member and Po1001 is up on both Cisco Nexus 5000 Series devices. FEX interface eth111/1/1 is suspended on both Cisco Nexus 5000 Series devices. With this configuration, the host is connected only to FEX 110.

```

N5k-1(config)# interface eth110/1/1, eth111/1/1
N5k-1(config-if)# channel-group 1001
N5k-1(config-if)# int po1001
N5k-1(config-if)# switchport access vlan 20

N5k-2(config)# interface eth110/1/1
N5k-2(config-if)# channel-group 1001
N5k-2(config-if)# int po1001
N5k-2(config-if)# switchport access vlan 20

```

```

N5596-1(config)# show int e111/1/1
Ethernet111/1/1 is down (suspended by vpc)
  Hardware: 100/1000 Ethernet, address: 7081.0500.2582 (bia 7081.0500.2582)
  MTU 1500 bytes, BW 1000000 Kbit, DLY 10 usec
  reliability 255/255, txload 1/255, rxload 1/255
N5596-1# show port-channel summary
Flags: D - Down          P - Up in port-channel (members)
       I - Individual    H - Hot-standby (LACP only)
       s - Suspended     r - Module-removed
       S - Switched      R - Routed
       U - Up (port-channel)
       M - Not in use. Min-links not met

```

Table 1-1 shows the port channel summary for N5596-1.

**Table 1-1** Port Channel Summary for N5596-1

Group	Port Channel	Type	Protocol	Member	Ports
1	Po1(SU)	Ethernet	LACP	1/1 (P)	E <sup>th</sup> 1/2 (P)
31	Po31(RU)	Ethernet	LACP	1/21 (P)	E <sup>th</sup> 1/22 (P)
101	Po101(SD)	Ethernet	—	1/41 (D)	E <sup>th</sup> 1/42 (D)
102	Po102(SU)	Ethernet	—	1/43 (P)	E <sup>th</sup> 1/44 (P)
103	Po103(SD)	Ethernet	—	1/10 (D)	E <sup>th</sup> 1/11 (D)
110	Po110(SU)	Ethernet	—	1/33 (P)	—
111	Po111(SU)	Ethernet	—	1/35 (P)	—
961	Po961(SD)	Ethernet	—	—	—
1001	Po1001(SU)	Ethernet	—	110/1/1 (P)	E <sup>th</sup> 111/1/1 (D)
2000	Po2000(SD)	Ethernet	—	110/1/3 (D)	E <sup>th</sup> 110/1/5 (D)

```

N5596-1#
N5596-2# show int e111/1/1
Ethernet111/1/1 is down (suspended by vpc)
  Hardware: 100/1000 Ethernet, address: 7081.0500.2582 (bia 7081.0500.2582)
  MTU 1500 bytes, BW 1000000 Kbit, DLY 10 usec
  reliability 255/255, txload 1/255, rxload 1/255
N5596-2# show port-channel summary
Flags: D - Down          P - Up in port-channel (members)
       I - Individual    H - Hot-standby (LACP only)
       s - Suspended     r - Module-removed
       S - Switched      R - Routed
       U - Up (port-channel)
       M - Not in use. Min-links not met

```

Table 1-2 shows the port channel summary for N5596-2.

**Table 1-2** Port Channel Summary for N5596-2

Group	Port Channel	Type	Protocol	Member	Ports
1	Po1(SU)	Ethernet	LACP	1/1 (P)	1/2 (P)
31	Po31(SD)	Ethernet	—	—	—
32	Po32(RU)	Ethernet	LACP	1/21 (P)	1/22 (P)
101	Po101(SU)	Ethernet	—	1/41 (D)	1/42 (D)
102	Po102(SD)	Ethernet	—	—	—
110	Po110(SU)	Ethernet	—	1/33 (P)	—
111	Po111(SU)	Ethernet	—	1/35 (P)	—
1001	Po1001(SU)	Ethernet	—	110/1/1 (P)	—

## Global vPC Consistency Check

The configuration that is subject to a global vPC consistency check is displayed as follows:

```
N5596-1# show vpc consistency-parameters global
```

Legend:

Type 1 : vPC will be suspended in case of mismatch

Table 1-3 shows the vPC consistency check.

**Table 1-3 vPC Consistency Check**

Name	Type	Local Value	Peer Value
QoS	2	([], [], [], [], [], [])	([], [], [4], [], [], [])
Network QoS (MTU)	2	(9216, 0, 0, 0, 0, 0)	(1538, 0, 1538, 0, 0, 0)
Network QoS (Pause)	2	(F, F, F, F, F, F)	(F, F, F, F, F, F)
Input Queuing (Bandwidth)	2	(100, 0, 0, 0, 0, 0)	(100, 0, 0, 0, 0, 0)
Input Queuing (Absolute Priority)	2	(F, F, F, F, F, F)	(F, F, F, F, F, F)
Output Queuing (Bandwidth)	2	(100, 0, 0, 0, 0, 0)	(100, 0, 0, 0, 0, 0)
Output Queuing (Absolute Priority)	2	(F, F, F, F, F, F)	(F, F, F, F, F, F)
STP Mode	1	MST	Rapid-PVST
STP Disabled	1	VLANs 123	None
STP MST Region Name	1	""	""
STP MST Region Revision	1	0	0

Name	Type	Local Value	Peer Value
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	—	1,10,20,58-61,100-102, 1,10,20,58-61,100,1000	
N5596-1#			

For a type 2 consistency check parameter, a warning message displays to remind you to have identical configurations on both vPC devices. The vPC and vPC member ports is up and running on both Cisco Nexus 5000 Series devices. Starting in the Cisco NX-OS 5.0(2)N2(1) release, an enhancement called Graceful consistency check was introduced to improve the resilience for vPC. With this feature, the mismatch of the type 1 consistency check parameter on a vPC secondary device, suspends its vPC member port. The vPC primary device keeps the vPC member ports operational to avoid the complete loss of connectivity for a network device behind a vPC.

The graceful consistency check feature does not work for the dual-homed FEX topology and Enhanced vPC topology. With the mismatch of the type 1 consistency check parameters, both vPC devices suspend their vPC member ports for a dual-homed FEX and Enhanced vPC. For example, the output shown above indicates that the STP mode is configured differently on two Cisco Nexus 5596 devices. All the interfaces from dual-homed FEXs are suspended due to the mismatch of the type-1 configuration.

The vPC graceful consistency check works for the FEX straight through topology.

## Port Channel Interface Level Configuration Checks

Under the port channel for Enhanced vPC configuration the two important parameters are the port mode (access vs. trunk) and the allowed VLAN for trunk mode.

The applicable configuration parameters that are subject to interface level consistency check are displayed with CLI **show vpc consistency check interface**.

```
N5K# show vpc consistency-parameters interface po1000
```

Legend:

Type 1 : vPC will be suspended in case of mismatch

Table 1-4 shows the port channel interface level configuration checks.

**Table 1-4 Port Channel Interface Level Configuration Checks**

Name	Type	Local Value	Peer Value
mode	1	on	on
Speed	1	1000 Mb/s	1000 Mb/s
Duplex	1	full	full
Port Mode	1	access	trunk
MTU	1	1500	1500
Admin port mode	1		
Shut Lan	1	No	No
vPC+ Switch-id	1	3000	3000
Allowed VLANs	-	10	1-57,61-3967,40 48-4093
Local suspended VLANs	-	10	-
N5596-1#			

## FCoE Configuration with Enhanced vPC

Prior to the Cisco NX-OS 5.1(3)N1(1) release, when an Ethernet interface was a port channel member, the VFC interface could only be bound to a port channel.

This example shows how to configure FCoE with Enhanced vPC:

```
interface eth100/1/1
Channe-group 1001

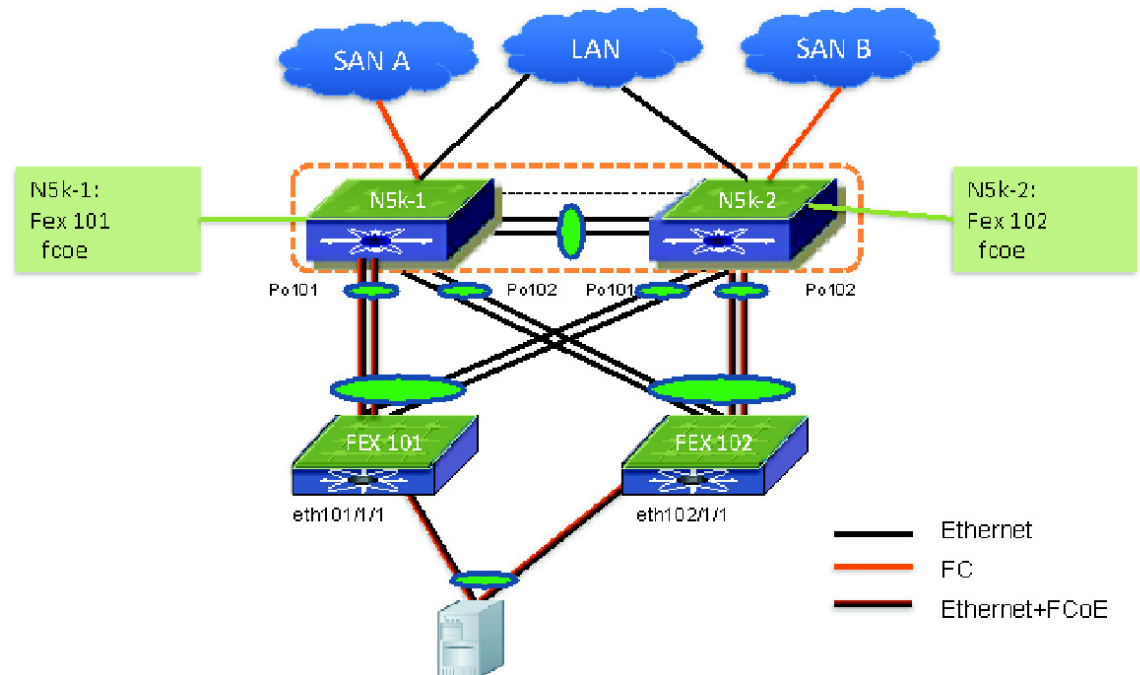
Interface Po1001
Switchport mode trunk
Switchport trunk allowed vlan 1, 100,2000

Interface vfc 1
bind interface Po1001
```

This configuration model requires that there is only one member for the port channel. It cannot be supported for Enhanced vPC where the port channel has at least two interfaces. Starting from the Cisco NX-OS 5.1(3)N1(1) release and later releases the VFC interface can be bound to physical interfaces.

Figure 1-11 shows an FCoE configuration topology.

Figure 1-11 FCoE Configuration Example Topology



This example shows how to configure the topology that is shown in Figure 1-11:

```
N5k-1(config)# fex 101
N5k-1(config-fex)# fcoe
N5k-1(config-fex)# interface vfc1
N5k-1(config-if)# bind interface eth101/1/1

N5k-2(config)# fex 102
N5k-2(config-fex)# fcoe
N5k-2(config-fex)# interface vfc1
N5k-2(config-if)# bind interface eth102/1/1
```

The FCoE portion of the configuration on the two vPC devices are different for the SAN traffic isolation and they are not subject to vPC consistency check.

You can enter the **fcoe** command to create an association between a FEX and a Cisco Nexus 5000 Series device for FCoE traffic. This command allows you to specify which Cisco Nexus 5000 Series device the FCoE traffic should be forwarded to from the FEX.

The same FEX cannot be associated to both the Cisco Nexus 5000 Series devices. The configuration shown in the following examples shows that with the same FEX that is associated with both Cisco Nexus 5000 Series device for FCoE is rejected.

```
N5k-1(config)# fex 101
N5k-1(config-fex)# fcoe
N5k-2(config)# fex 101
N5k-2(config-fex)# fcoe
```

On a Cisco Nexus 5000 Series device, the VFC can only be bound to a FEX interface if the FEX where the interface resides has already been associated to the Cisco Nexus 5000 Series device from an FCoE point of view. In the following example, FEX 101 is associated with N5k-1. When you try to bind a VFC interface to an interface from the FEX 102, the commands are rejected.


```
N5k-1(config)# fex 101
```

```
N5k-1(config-fex)# fcoe
N5k-1(config-fex)# interface vfc1
N5k-1(config-if)# bind interface eth102/1/1
```

The FEX association to the Cisco Nexus 5000 Series device has to be configured before any VFC interfaces can be created for an Enhanced vPC topology.

## Software Upgrade with Enhanced vPC

The Enhanced vPC topology does not change the Cisco NX-OS software upgrade procedure. It has an identical upgrade procedure as the dual-home FEX topology. It supports ISSU if the ISSU conditions are met. For the detailed software upgrade/downgrade procedure see the following URL:

 [www.cisco.com/en/US/docs/switches/datacenter/nexus5000/sw/upgrade/503\\_N1\\_1/n5k\\_upgrade\\_downgrade\\_503.html](http://www.cisco.com/en/US/docs/switches/datacenter/nexus5000/sw/upgrade/503_N1_1/n5k_upgrade_downgrade_503.html)

## Monitoring the Traffic in Enhanced vPC

The vPC topology presents a challenge for traffic monitoring because each vPC device carries half of the traffic flow. Prior to ERSPAN, you had to configure local SPAN on both vPC devices to monitor all flows sent to and received from the vPC. The procedure required that you combined the packets trace from two SPAN destination ports to get a complete view.

From the Cisco NX-OS 5.1(3)N1(1) release and later releases, the Cisco Nexus 5000 Series devices support ERSPAN source session. With ERSPAN, you can monitor and capture all the flows for the same vPC from one sniffer. The following example shows how to capture all the traffic flows from a host behind an Enhanced vPC.

**Note**

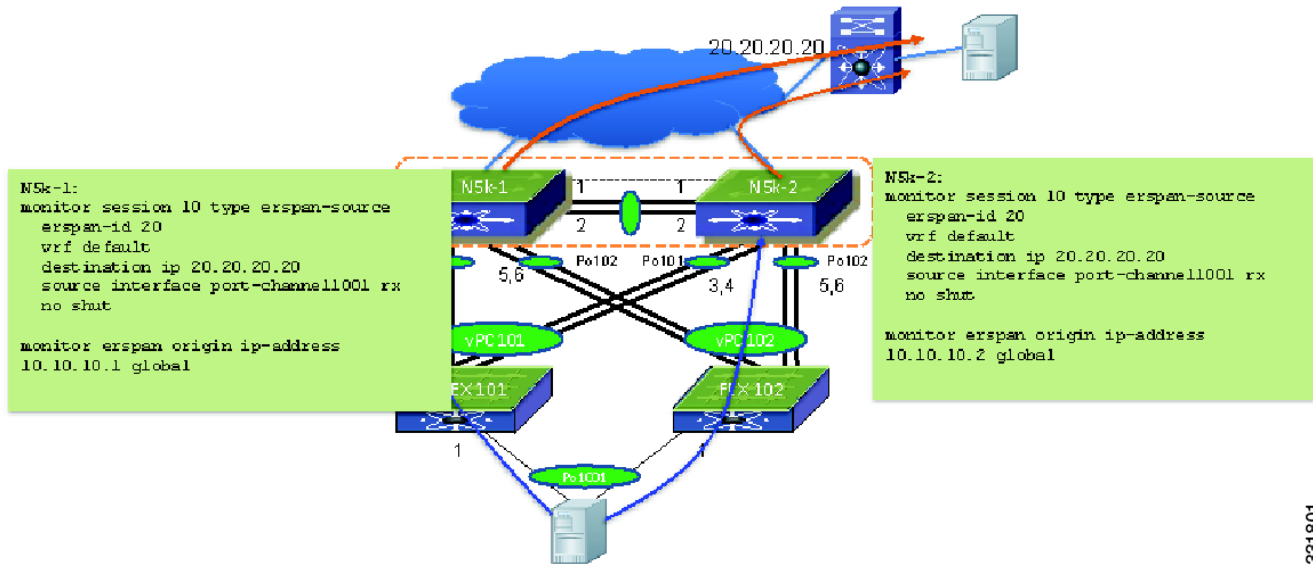
---

The Cisco Nexus 5000 Series device supports only ERSPAN source sessions not the ERSPAN destination sessions. The platforms that support the ERSPAN destination session are the Cisco Nexus 7000 Series devices, the Cisco Catalyst 6500 Series devices, and the Cisco Nexus 1010 NAM.

---



Figure 1-12 Monitoring the Traffic in Enhanced vPC



331801

In the Enhanced vPC topology, a unicast and multicast flow can be sent to either one of two FEXs involved in the host vPC. The following example shows how to identify which path the traffic flow is forwarded to. In the following example po1001 is an Enhanced vPC host port channel consists of eth110/1/1 and eth111/1/1. The unicast flow from 30.30.1.2 to 30.30.3.2 is sent to FEX 111. Enter the command to determine which FEX interface from N5k-1 to FEX 111 is carrying the flow.

```
N5596-1# show port-channel load-balance forwarding-path interface po1001 vlan 10 src-ip 30.30.1.2 dst-ip 30.30.3.2 src-mac 0000.0100.1100 dst-mac 0000.0000.0b00
```

```

Missing params will be substituted by 0's.
Load-balance Algorithm on FEX: source-dest-ip
crc8_hash: Not Used      Outgoing port id: Ethernet111/1/1
Param(s) used to calculate load-balance (Unknown unicast, multicast and broadcast packets):
  dst-mac: 0000.0000.0b00
  vlan id: 10

```

This example shows FEX 110 and interface eth110/1/1 carries the multicast flow for Po1001:

```
N5596-1# show port-channel load-balance forwarding-path interface po1001 vlan 10 src-mac 0000.0100.1100 dst-mac 0100.5e01.010a src-ip 30.30.1.2 dst-ip 224.1.1.10
```

```

Missing params will be substituted by 0's.
Load-balance Algorithm on FEX: source-dest-ip
crc8_hash: Not Used      Outgoing port id: Ethernet110/1/1
Param(s) used to calculate load-balance (Unknown unicast, multicast and broadcast packets):
  dst-ip: 224.1.1.10
  vlan id: 10
N5596-1#

```





---

## Numerics

- 10 Gigabit-Ethernet
  - peer link ports [1-14](#)

---

## A

- ARP processing with vPC [2-2](#)
- auto-recovery
  - about [1-8](#)
  - replacing reload restore [1-8](#)
  - status [1-9](#)

---

## C

- Cisco 6004 port numbering and port mapping with UPCs in a slot [1-5](#)
- Cisco Nexus 2000 Series Fabric Extender
  - installing a new Fabric Extender [1-13](#)
  - replacing in a dual-homed vPC topology [1-12](#)
  - replacing in a single-homed vPC topology [1-13](#)
  - replacing in a vPC topology [1-12](#)
- Cisco Nexus 2248PQ Fabric Extender [1-1](#)
- Cisco Nexus 5000 Series switch
  - replacing in a vPC topology [1-11](#)
- Cisco Nexus 5500 Series and QSFP+ GEM [1-2](#)
- Cisco Nexus 6004 Switch Fabric Mode [1-7](#)
- Configuring QSFP+ ports as 4x 10GE or 40GE Ports [1-5](#)
- connecting to a router in a vPC topology [2-3](#)
- consistency check
  - bypassing when a peer link is lost [1-8](#)
  - failure [1-7](#)
    - configuration differences that lead to [1-7](#)

- status [1-7](#)
- successful [1-7](#)

- consistency checks
  - configuring per-VLAN [1-5](#)
- control traffic forwarding in a vPC topology [2-6](#)

---

## D

- dedicated VRF [2-7](#)
- delay restore [2-4](#)
- delay timer [2-4](#)
- designated router [2-10](#)
  - CFS message [2-11](#)
  - elected [2-11](#)
  - priority [2-11](#)
- DR election
  - see designated router [2-11](#)

---

## F

- faster convergence
  - in vPC topology [2-9](#)
- FHRP. See also First Hop Redundancy Protocol
- First Hop Redundancy Protocol [2-1](#)

---

## G

- graceful consistency check [1-2](#)
  - about [1-3](#)

---

## I

- improved convergence [2-4](#)

## ISSUs

- not supported [2-17](#)
- supported [2-18](#)

**K**

## keepalive interface

- dedicated VRF for a [2-7](#)

## keepalive link

- failure followed by a peer link failure [1-16](#)

**L**

## Layer 3

- and ISSUs [2-17](#)
- connecting to a router in a vPC topology [2-6](#)
- improved convergence with a vPC topology [2-4](#)
- module failure [2-5](#)
- recommendation for connections between a router and switch [2-6](#)
- source and Rendezvous Point (RP) [2-10](#)
- vPC consistency check [2-8](#)

**M**

## multicast

- data forwarding [2-11](#)
- forwarding algorithm [2-11](#)
- forwarding process [2-13](#)
- forwarding rules [2-12](#)
- routing table size [2-9](#)
- unsupported topology in vPC configurations [2-9](#)

## multicast routing table

- example of switch output [2-10](#)

## multicast traffic

- not routed [2-12](#)

**N**

- new and changed features (table) [2-9](#)

**P**

- peer-gateway command [2-4](#)

## peer link

- failure followed by a peer keepalive link failure [1-16](#)

## peer links

- bandwidth [1-14](#)
- failure [1-14](#)

## peer switch

- failure [1-16](#)

PIM router [2-9](#)

## prebuilt source tree

- faster convergence [2-9](#)

**Q**

- QSFP+ slot number on N55-M4Q [1-3](#)

- Quad Small Form-Factor Pluggable Plus (QSFP+) [1-1](#)

**R**

- reload delay period [1-8](#)

reload restore [1-8](#)

- bypassing the vPC consistency check [1-15](#)

Rendezvous Point (RP) [2-10](#)

- routing table size [2-9](#)

**S**

## STP

- mode mismatch example [1-4](#)
- Type 1 consistency checks [1-5](#)

---

**T**

## traffic flow

- tracing in a vPC topology [1-17](#)

## Type 1

- interface-level inconsistency [1-4, 1-5](#)

## Type 2

- parameter mismatch [1-2](#)

---

**U**

- unsupported multicast topology [2-9](#)

---

**V**

## VLAN

- consistency checks [1-5](#)

## vPC

- consistency checks [1-1](#)

- identifying inconsistent configurations [1-6](#)

- member port failure [1-13](#)

- peer keepalive link failure [1-15](#)

- traffic flow [1-17](#)

  - diagram [1-17](#)

- unsupported multicast topology [2-9](#)

- vPC and peer-gateway [2-3](#)

- vPC failure scenarios [1-13](#)

## vPC operations

- about [1-1](#)

- vPC peer link failure [2-5](#)

## vPC topologies

- configuration changes [1-9](#)

## vPC topology

- multicast interaction [2-8](#)

## VRF

- services that are recognized [2-8](#)

