vPC

Daniel LAZAR
Customer Support Engineer IV
CCIE R&S, DC #36664, CCDE 2016::39

Version 20.17.0508

Agenda

- HW/SW requirements
- Overview and key concepts
- Configuration best practices
- Design best practices
- Troubleshooting
- Tools and relevant tech-support

HW/SW requirements

HW/SW support

1. **Nexus 5500/5600/6000**:
- Carmel ASIC starting with NX-OS 4.1(3)N1(1)
- Bigsur ASIC starting with NX-OS 6.0(2)N2(1)

2. **Nexus 7000/7700**:
- M1 EARL8 ASIC starting with NX-OS 4.1(3)
- F1 Orion ASIC starting with NX-OS 5.1(1)
- F2/F2E Clipper ASIC starting with NX-OS 6.0(1)
- M2 EARL8 ASIC starting with NX-OS 6.1(1)
- F3 Flanker ASIC starting with NX-OS 6.2(6)
- M3 Starlifter ASIC starting with NX-OS 7.3(0)DX(1) and 8.0(1)

HW capabilities

Feature	Nexus 5600/6000	Nexus 7000 F2E	Nexus 7000 F3
Interface speed	1G, 10G, 40G	1G, 10G	1G, 10G, 40G, 100G
L2 throughput	Line rate	Line rate	Line rate
MAC address table	256K (shared with ARP/ND)	16K per SoC	64K per SoC
VLANs	4K	4K	4K
Buffers	25MB per SoC	72MB per module	72MB or 144MB per module
FabricPath	Yes	Yes	Yes
vPC	Yes	Yes	Yes

SW licensing requirements

Nexus 5500/5600/6000:

- No license required (the feature comes with the device).

Nexus 7000/7700:

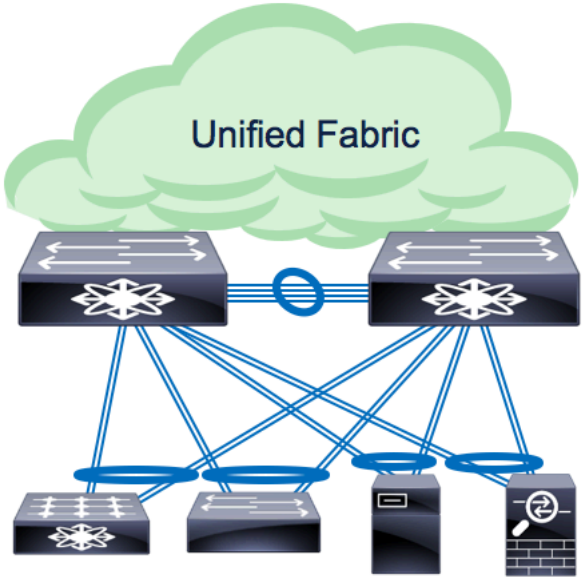
- No license required (the feature comes with the device).

More information **here**.

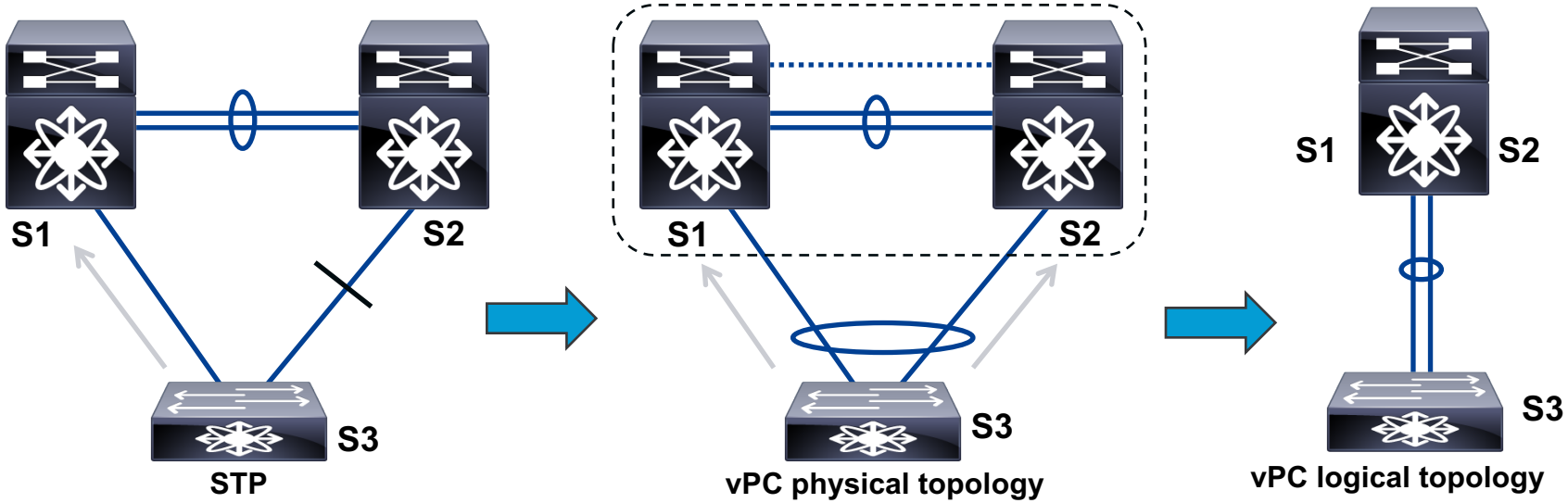
Overview and key concepts

What is vPC and why?

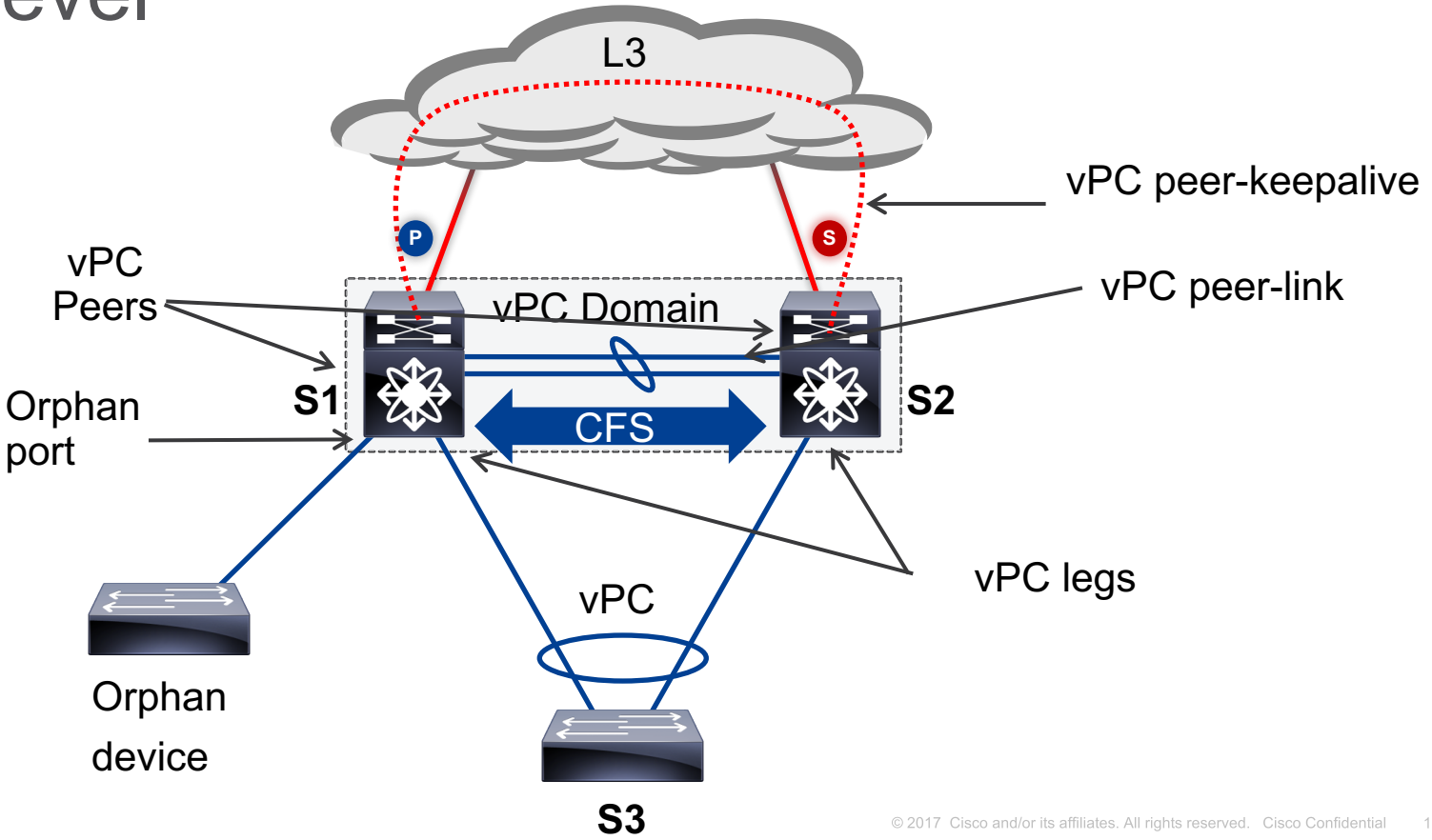
- MC-LAG on Cisco Nexus switches.
- Provides device level redundancy with faster convergence.
- Eliminates STP blocked ports by providing a loop-free topology.
- Better bandwidth utilization.
- Deployed by almost 95% of Cisco Nexus customers.



High level

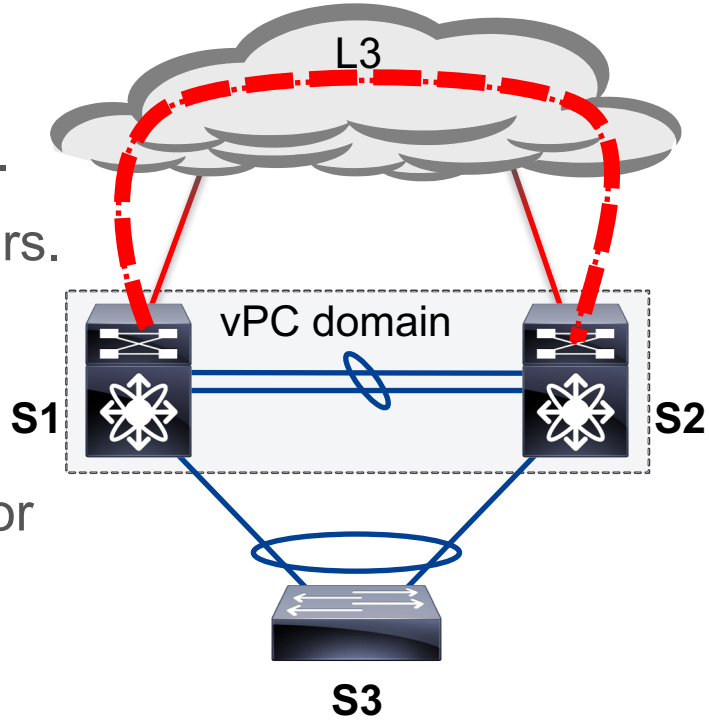


Low level



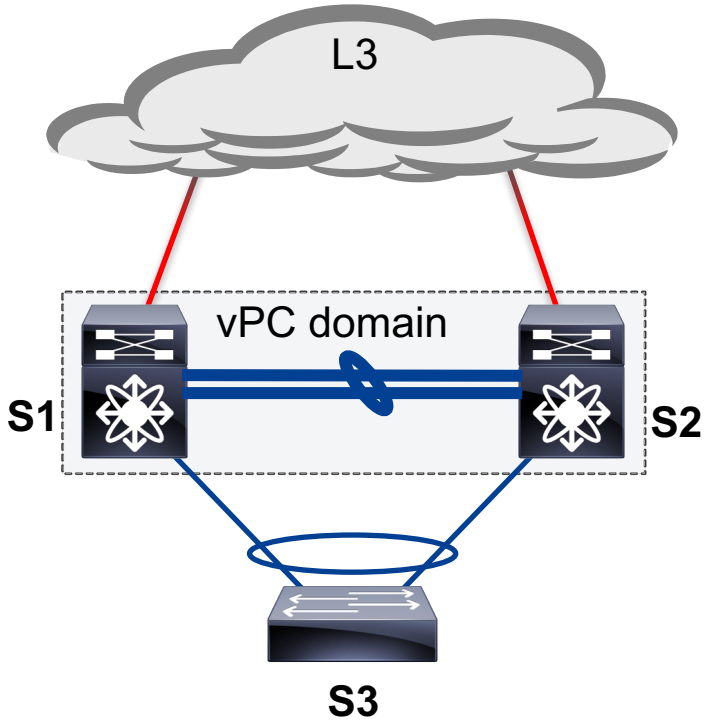
vPC peer-keepalive

- Heartbeat mechanism between vPC peers.
- Requires L3 reachability between vPC peers.
- Heartbeat sent every 1 second by default.
- Uses UDP port 3200.
- Its purpose is to avoid/resolve dual-active or split-brain scenarios.
- Does **not** require a point-to-point link.



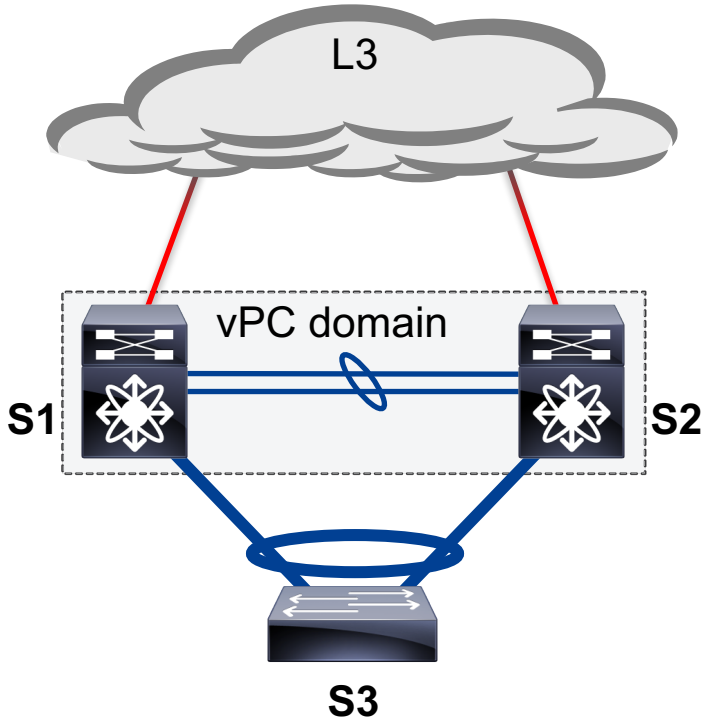
vPC peer-link

- The most important vPC component.
- A port-channel interface that carries:
  - CFS sync messages between peers.
  - STP BPDUs and HSRP Hellos
  - Multicast and orphaned traffic.
- vPC vlans **must** be active on the vPC peer-link in order for the vPCs to be forwarding traffic.



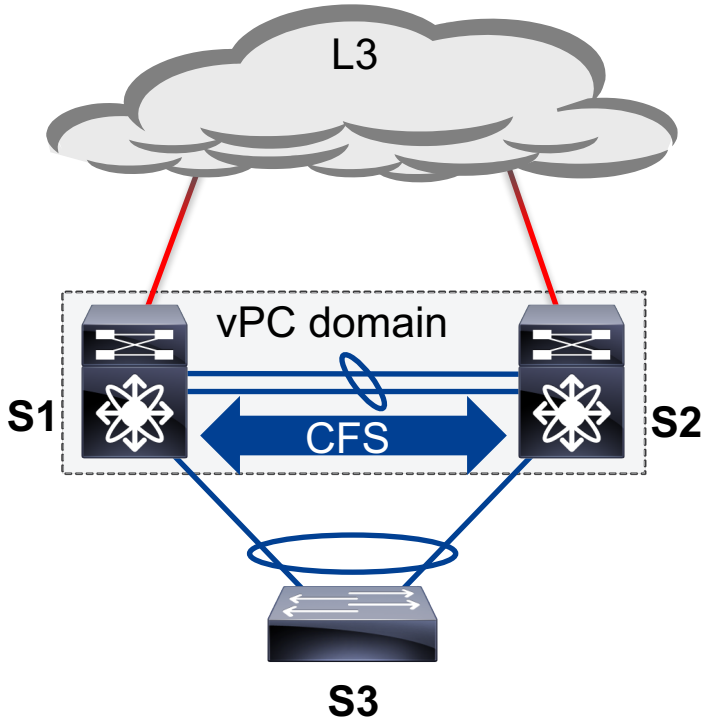
vPC

- The actual port-channel connecting the device in need of redundancy.
- The ports forming the vPC, referred as vPC legs, are split between peers. Non-vPC ports are referred as orphan ports.
- The vlans forwarding on the vPC **must** be active on the vPC peer-link.
- Can only be configured in L2 mode access or L2 mode trunk.



CFS

- Synchronization and consistency checking mechanism needed for maintaining the illusion of a single control-plane from the perspective of other devices.
- Available in two flavors: CFSoE, CFSoIP.
- CFSoE (Ethertype 0x8843/0x8844) is enabled by default and exchanges states about vPC legs, MAC addresses, multicast receivers (IGMP Snooping), etc.



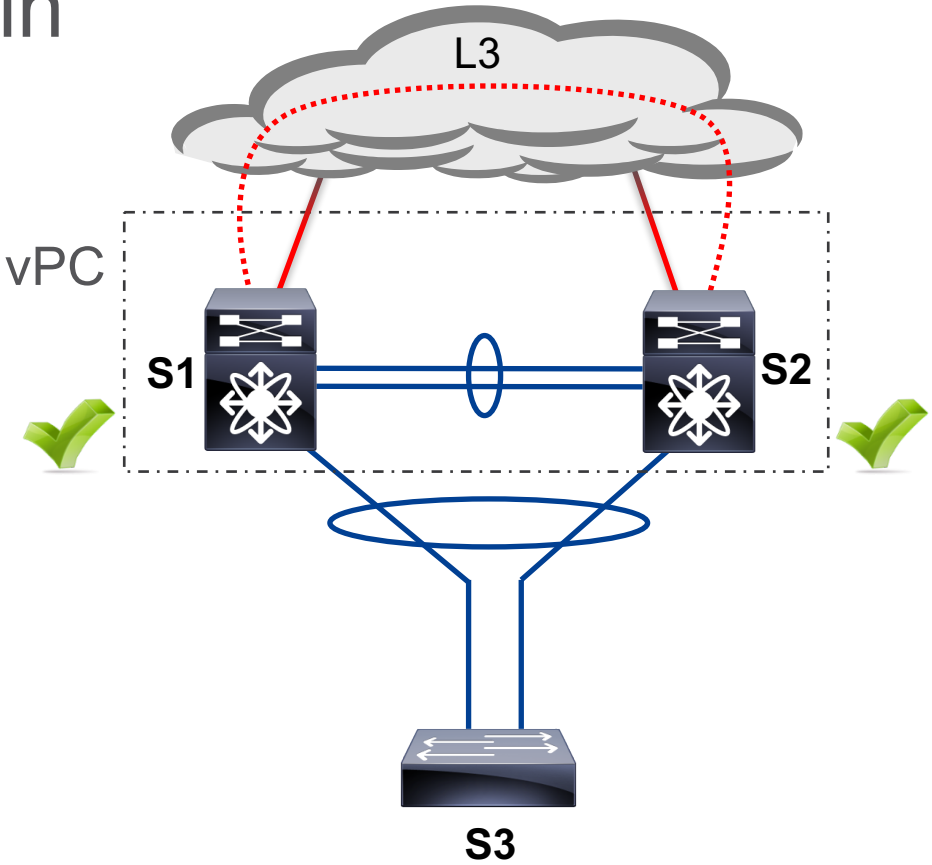
vPC consistency-check

- The vPC feature allows links to be bundled to form a single entity if certain compatibility conditions are met.
- Depending on the severity of the misconfiguration, vPC may either warn the user (for Type-2) or suspend the vPC (for Type-1).
- Type-1 parameters: STP mode, STP global settings, port-channel mode, MTU, etc.
- Type-2 parameters: SVI, ACL, QoS, IGMP snooping, HSRP, etc.

Configuration best practices

Building a vPC domain

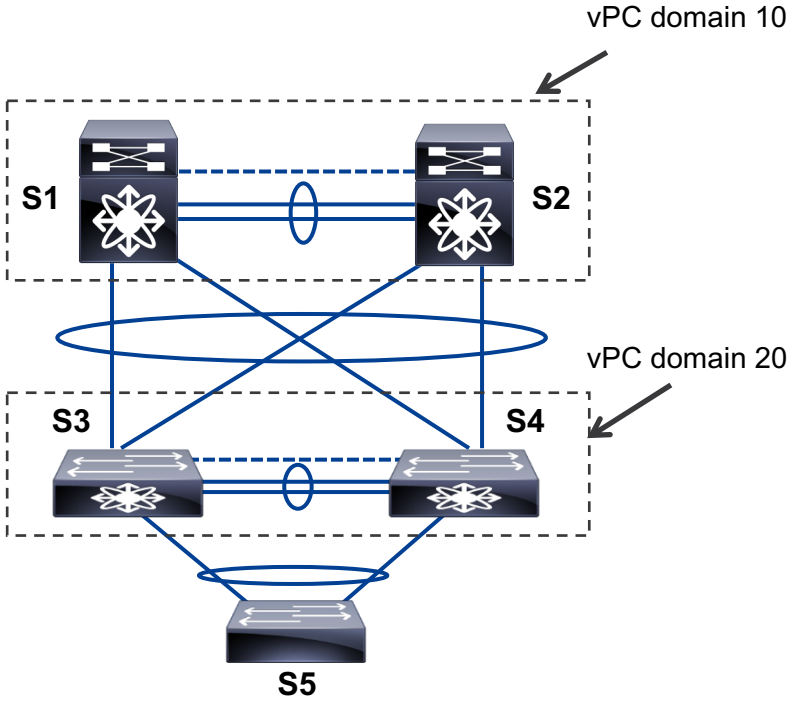
1. Define the vPC domain.
2. Establish connectivity over the vPC peer-keepalive path.
3. Bring up the vPC peer-link.
4. Configure vPCs.
5. Keep configuration consistent.



vPC domain ID

- The vPC peers use the vPC domain ID to automatically assign a system MAC address (representing both vPC peers).
- Within a contiguous layer 2 domain, the vPC domain IDs **must** be unique.

```
n7700-2# show vpc brief | i i domain
vPC domain id                          : 12

n7700-2# show vpc role | i i system-mac
vPC system-mac                 : 00:23:04:ee:be:0c
vPC local system-mac           : 8c:60:4f:e7:f6:43
```

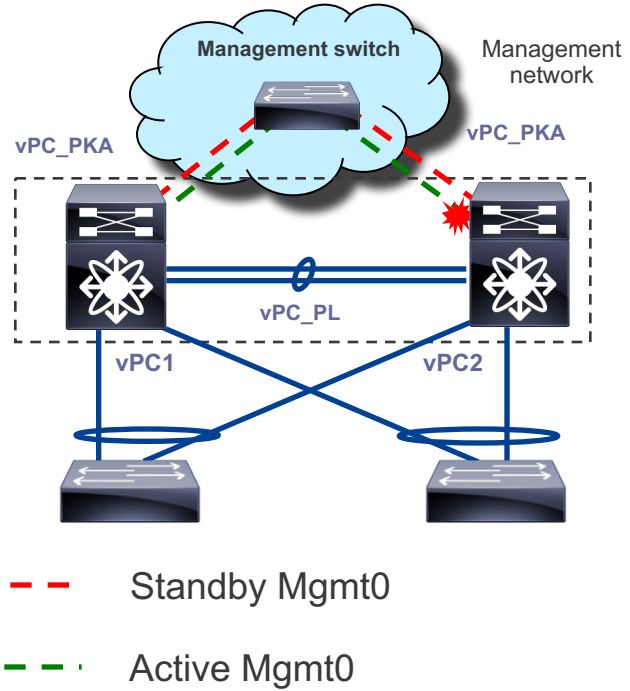


vPC peer-keepalive

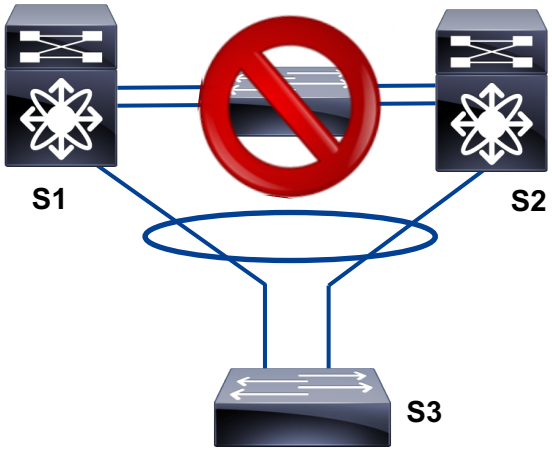
Preference	Nexus 7X00/9500 switches	Nexus 3X00/5X00/6000/9300 switches
1	Dedicated link(s) (1GE/10GE LC)	Mgmt0 interface
2	Mgmt0 interface	Dedicated link(s) (1GE/10GE LC)
3	L3 infrastructure	L3 infrastructure

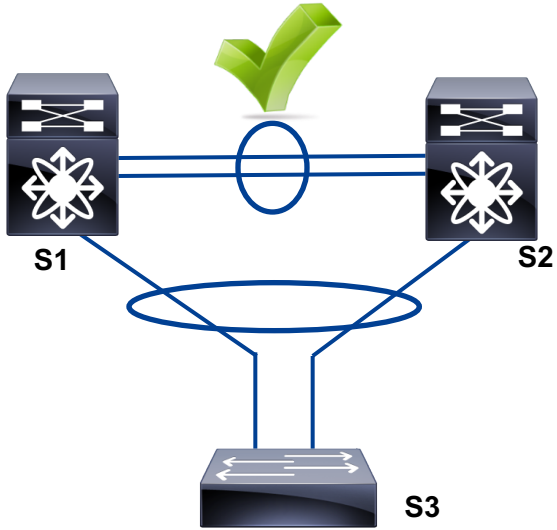
vPC peer-keepalive – Dual supervisors

- Do **not** connect Mgmt0 interfaces back to back between the two switches that have dual supervisors.
- Only one Mgmt0 interface can be active at a given point in time and a switchover may break vPC peer-keepalive connectivity.
- Use the Mgmt0 interface when you have an out-of-band management network.



vPC peer-link

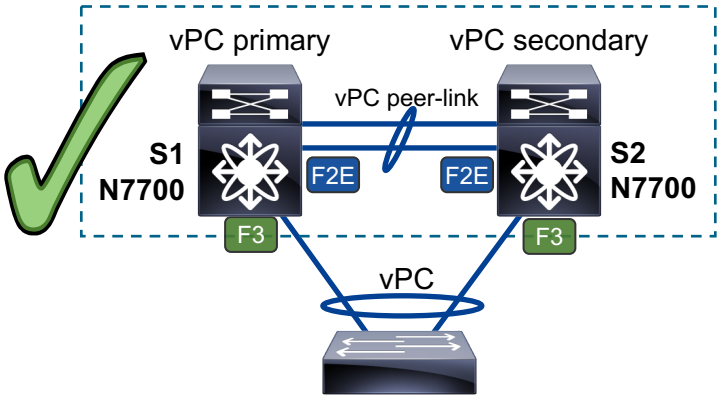


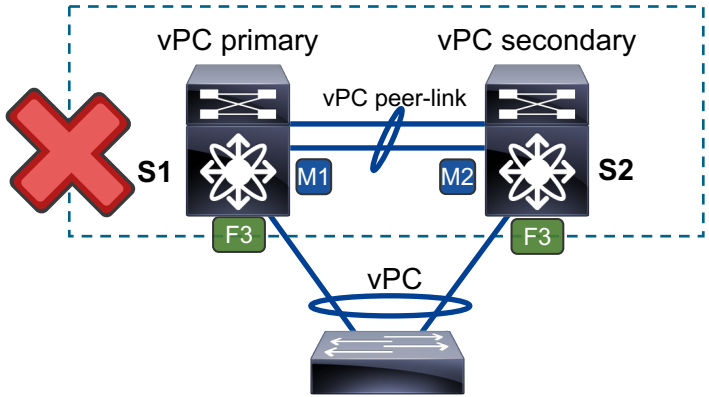


- vPC peer-link **must** be a point-to-point connection.
- vPC peer-link member ports can be 10/40/100GE interfaces.

vPC peer-link – Mixed linecards

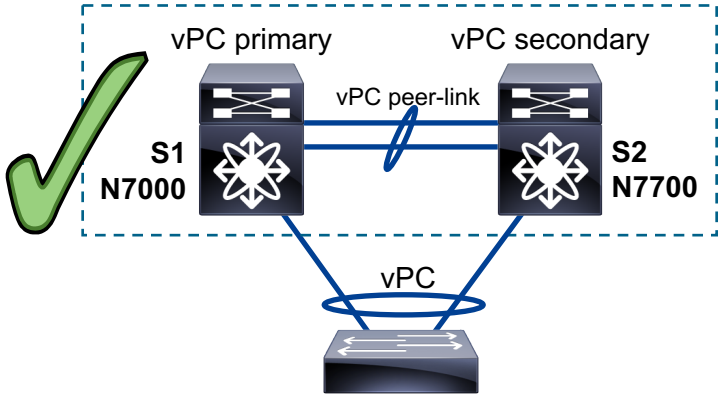
- Use identical line cards on both sides of the vPC peer-link and vPC member ports.

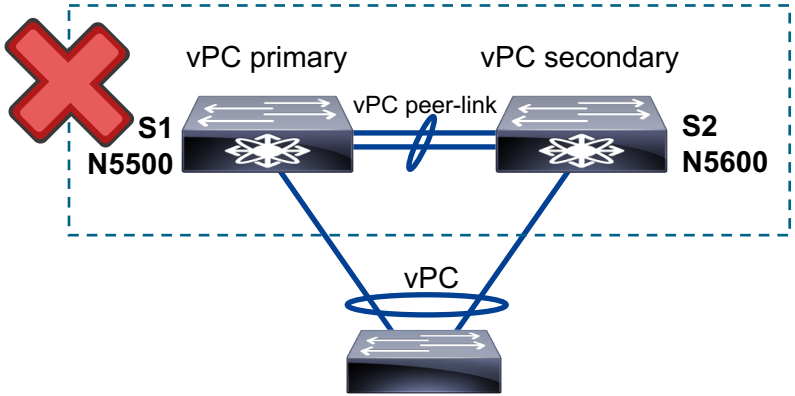




vPC peer-link – Mixed chassis

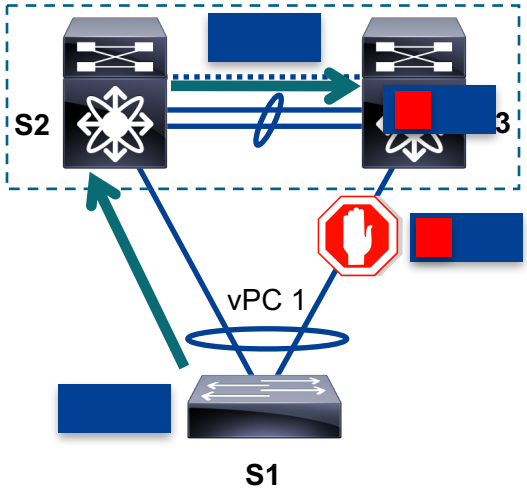
- Nexus 7000 and Nexus 7700 in the same vPC pair is supported.
- Nexus 7000/7700 VDC type **must** match on both vPC peers.
- Nexus 5500 and Nexus 5600 in same vPC pair is **not** supported.





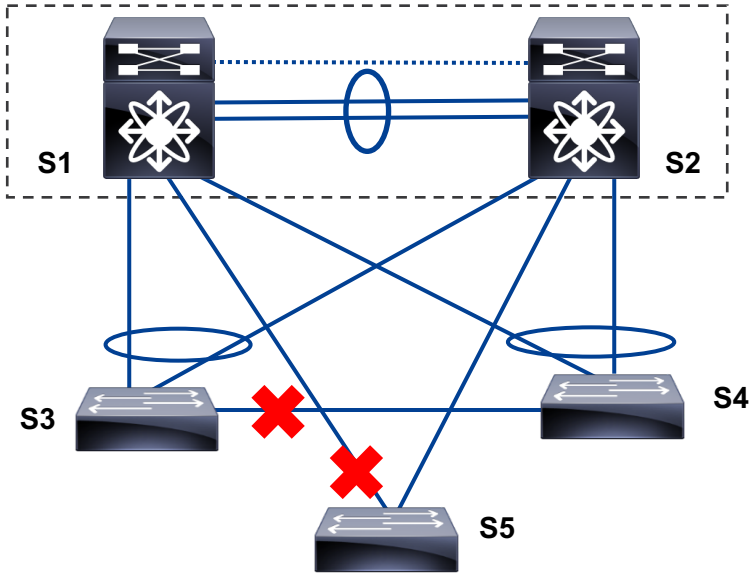
vPC loop avoidance

- Data-plane mechanism for loop prevention.
- The vPC peer forwards the traffic locally whenever possible.
- Traffic coming from a vPC member port, crossing vPC peer-link is **not** allowed to egress any vPC member port.
- Exception of the rule, when a vPC member port goes down.



vPC and STP

- All switches in the L2 domain should run either Rapid-PVST+ or MST.
- Do **not** disable STP for vPC vlans (Type-1 vPC consistency check).
- Configure vPC peers as the STP root in order to reduce convergence time.
- STP is used in vPC environment to prevent loops outside of the vPC domain and misconfiguration.



vPC and STP - vPC peer-switch

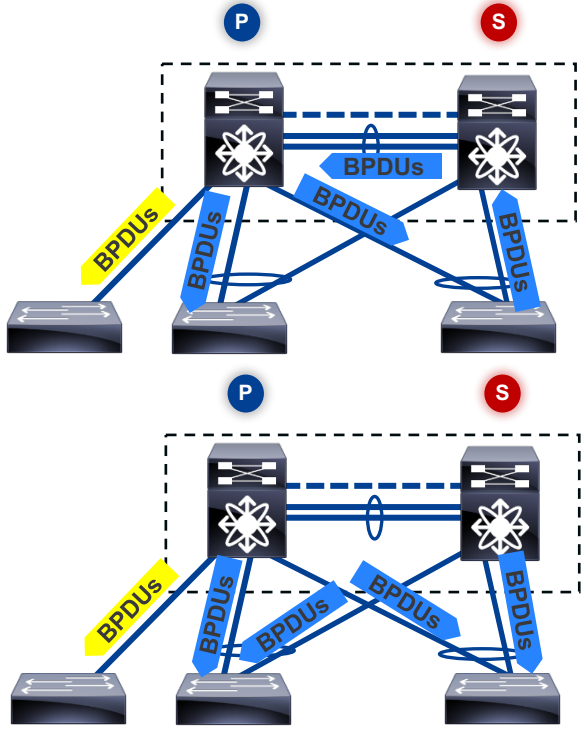
Without vPC peer-switch:

- vPC primary sends BPDUs on designated ports (convergence concern).
- vPC secondary proxies BPDUs to primary.

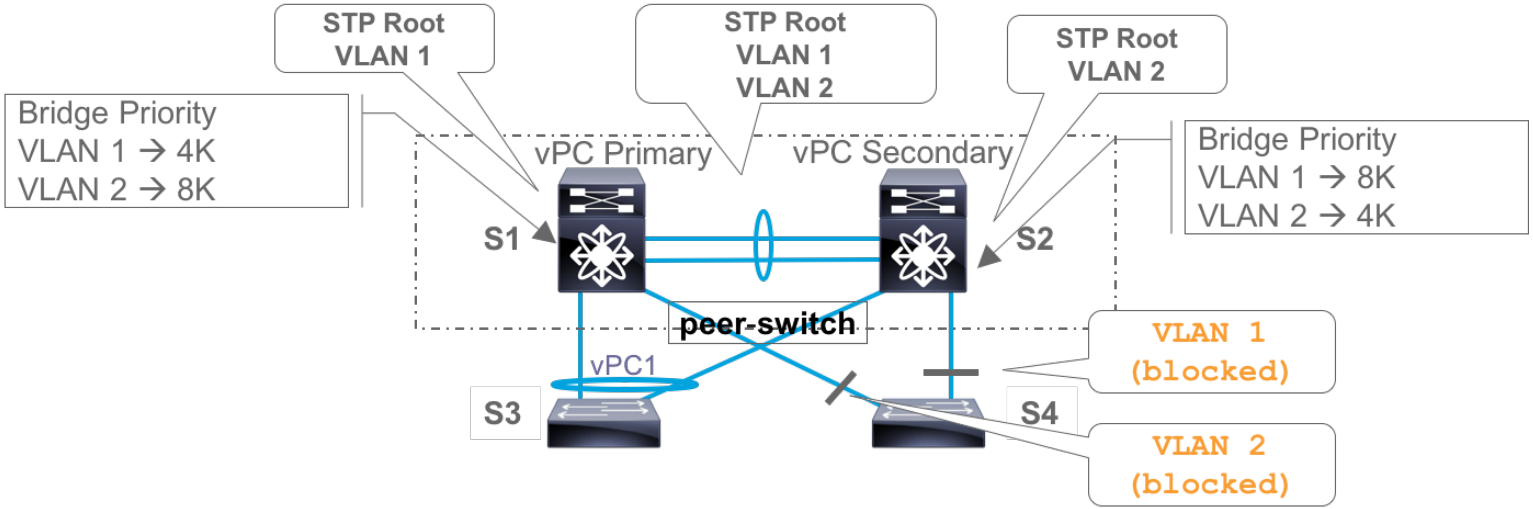
With vPC peer-switch:

- vPC peers appear as a single STP root.
- BPDUs processed by the logical STP root formed by the two vPC peers.

Nexus(config-vpc-domain)# peer-switch



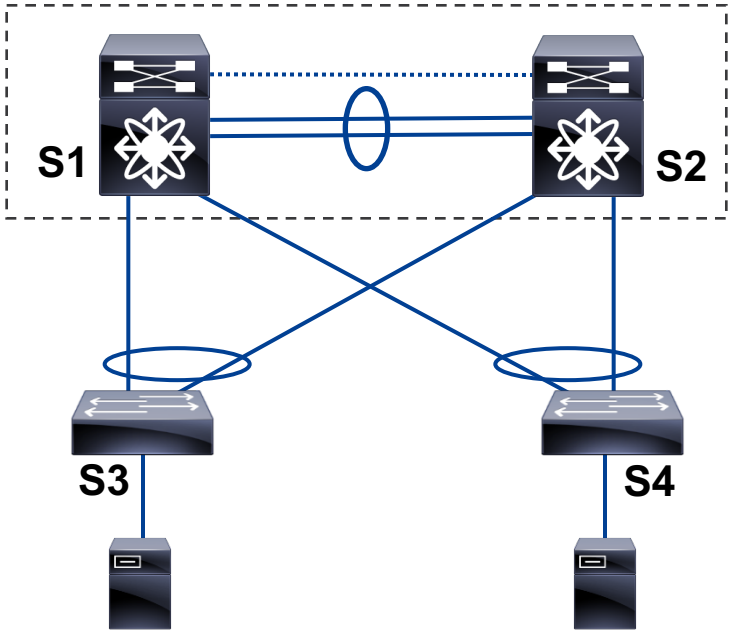
vPC and STP – Hybrid topology



- Needs additional configuration: STP pseudo-information (takes precedence over global STP configuration).

vPC peer-gateway

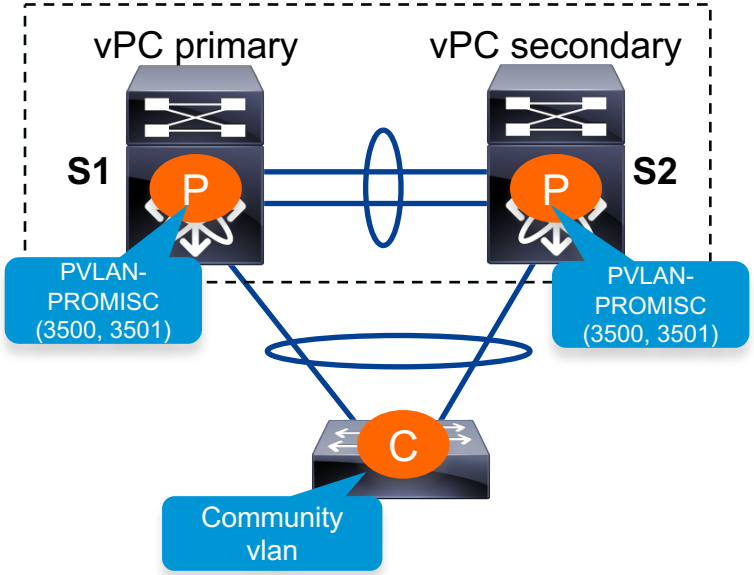
- Allows a vPC switch to route frames addressed to the peer router MAC (which is installed locally with G flag).
- Keeps forwarding of traffic local to the vPC switch and avoids the use of the vPC peer-link.
- Allows interoperability with features of some 3rd party NAS and load-balancer devices.



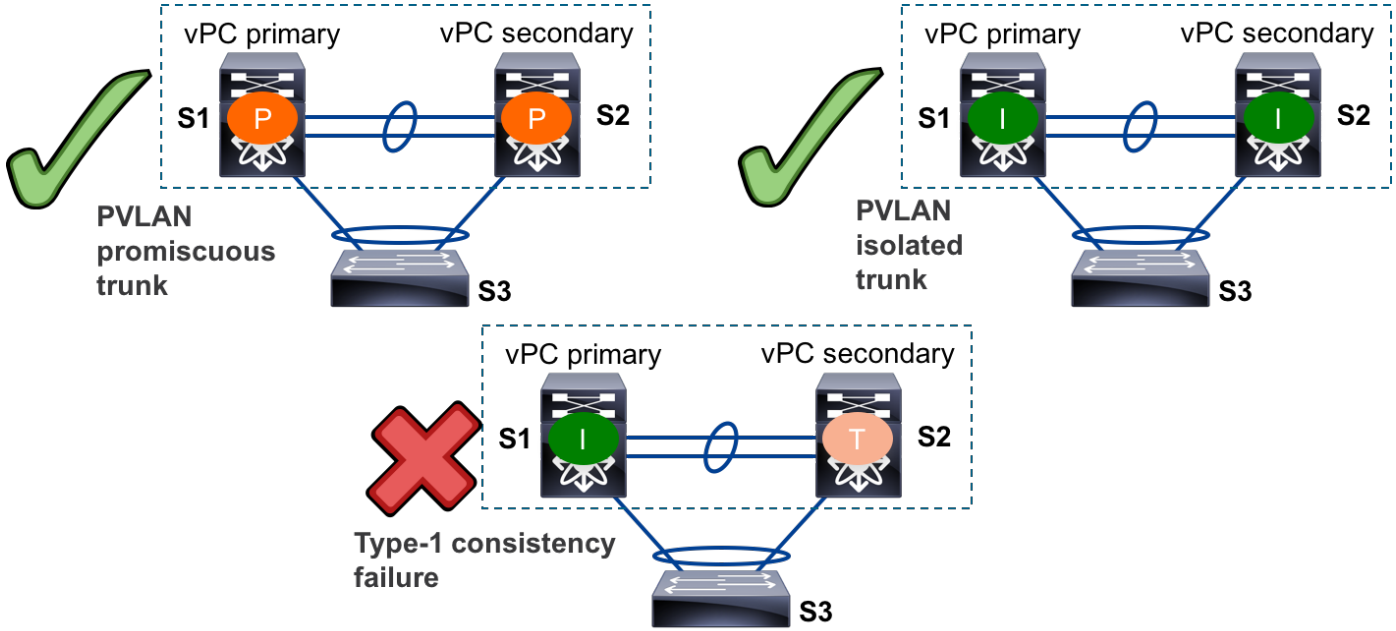
Nexus(config-vpc-domain)# peer-gateway

vPC and PVLAN

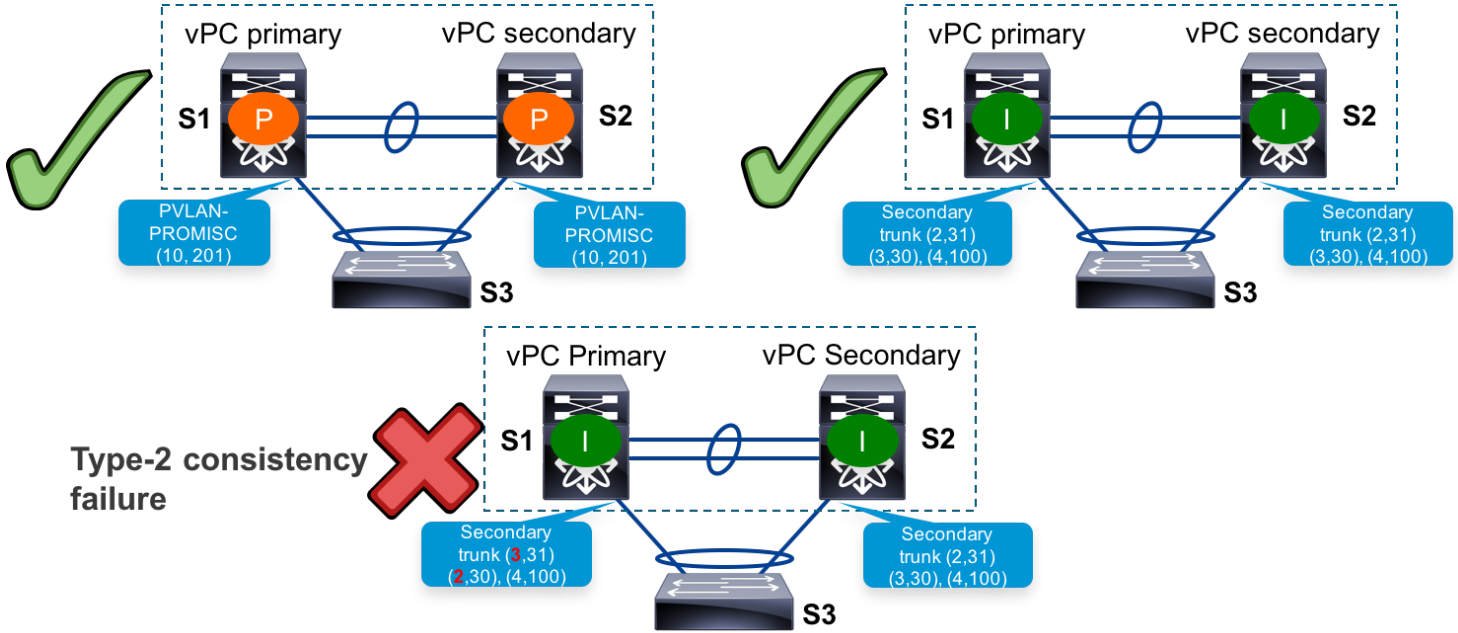
- PVLAN configuration is **not** supported on the vPC peer-link.
- vPC type-1 inconsistency: vPC member port brought down if PVLAN port mode differs between vPC peers.
- vPC type-2 inconsistency: PVLAN will bring down mismatched couples.



vPC and PVLAN – Type-1 inconsistency



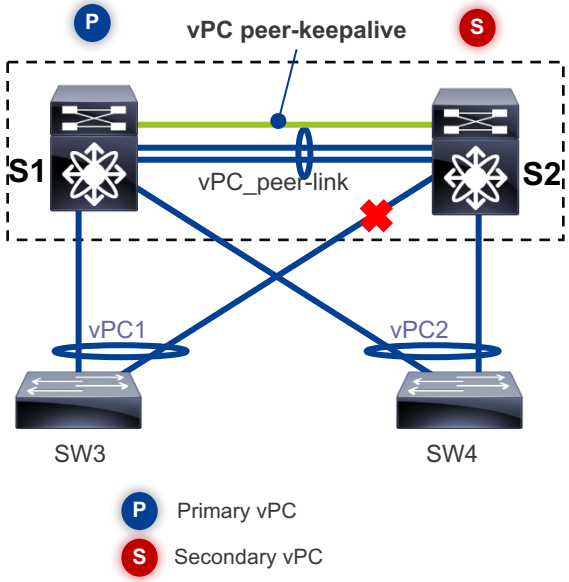
vPC and PVLAN – Type-2 inconsistency



Failure scenarios

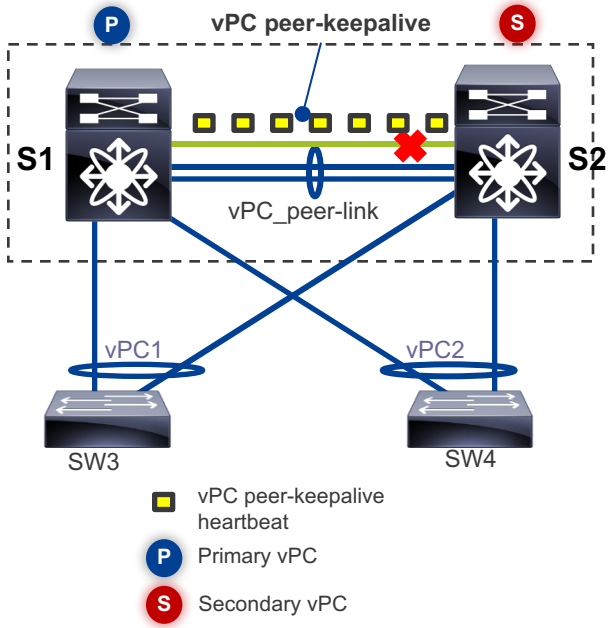
vPC member port failure

- One of the vPC member ports fails.
- No change in role occurs for the vPC primary and vPC secondary switches.
- Forwarding path changes and traffic addressed to the peer will cross the vPC peer-link.
- This is **not** a desirable behavior since the vPC peer-link can be oversubscribed from the bandwidth perspective.



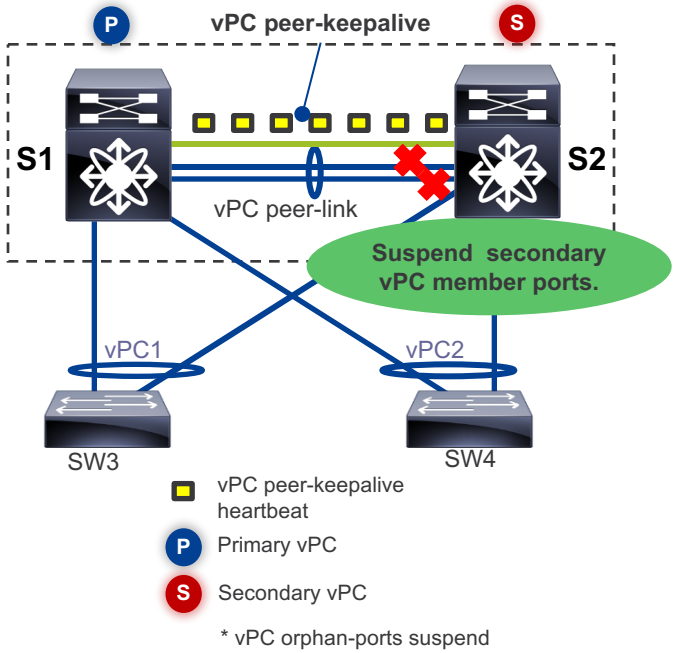
vPC peer-keepalive failure

- The vPC peer-keepalive fails (either at physical link level or at the network level).
- The vPC peer-link remains up.
- No change in role occurs for the vPC primary and vPC secondary switches.
- The status of the other vPC is still known via the vPC peer-link.
- Forwarding path is **not** affected.



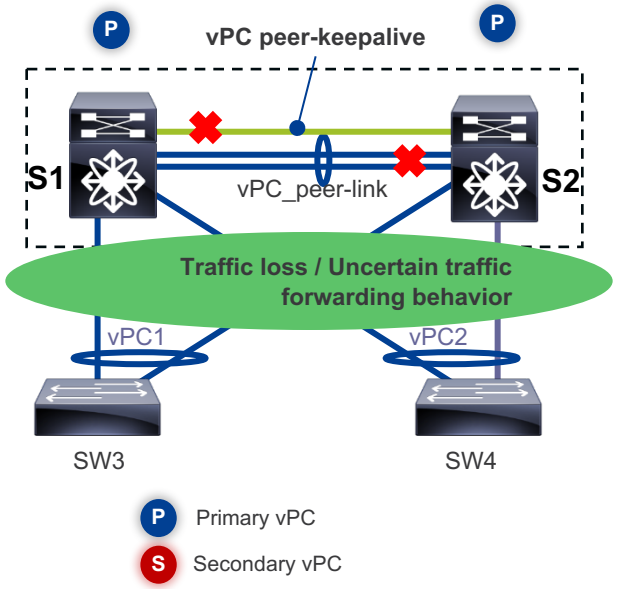
vPC peer-link failure

- The vPC peer-link fails.
- The vPC peer-keepalive remains up.
- No change in role occurs for the vPC primary and vPC secondary switches.
- The status of the other vPC is still known via the vPC peer-keepalive.
- Secondary vPC peer suspends all vPCs.
- Forwarding path changes, the traffic from the devices connected to the secondary vPC peer via orphan ports is blackholed*.



vPC peer-keepalive and peer-link failure

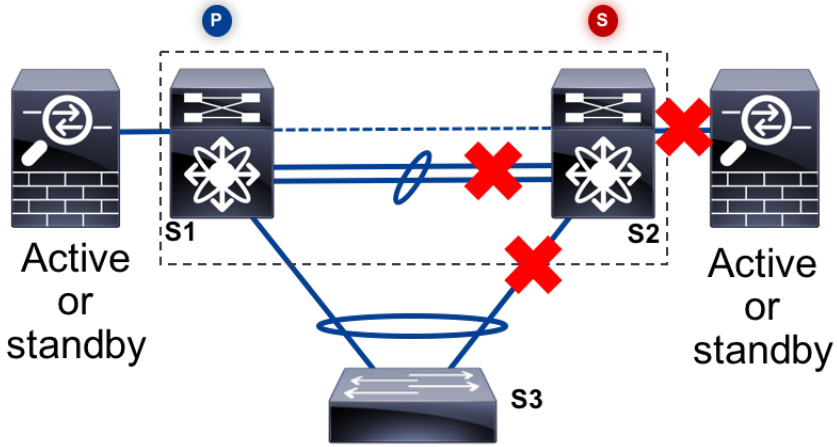
- The vPC peer-keepalive fails, followed by the vPC peer-link failure.
- The role changes for both vPC peers to vPC primary (dual-active/split-brain scenario).
- The result is inconsistent behavior for traffic forwarding on the vPC peers.
- When links are restored, the vPC operational primary (former vPC secondary) keeps the vPC primary role and the former vPC primary becomes vPC operational secondary.



Additional features

vPC orphan-ports suspend

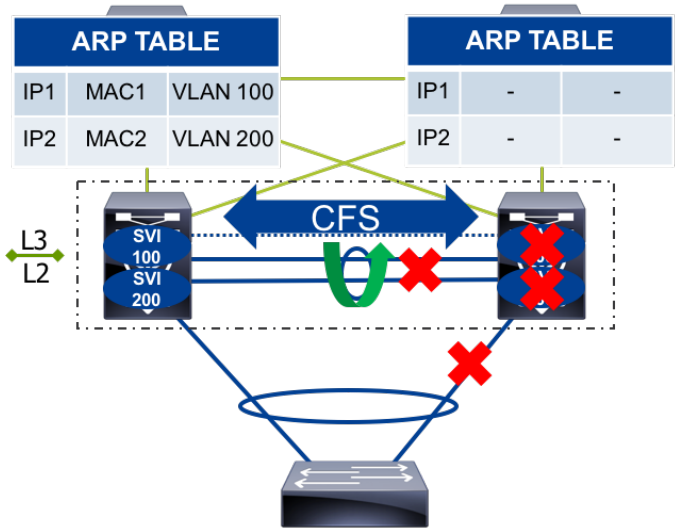
- The traffic from devices single-attached to vPC secondary is blackholed if vPC peer-link fails.
- With this feature, orphan ports are suspended on the vPC secondary upon failure.
- When vPC peer-link is restored, vPC secondary restores the orphan ports.



Nexus(config-if)# vpc orphan-ports suspend

vPC ARP synchronize

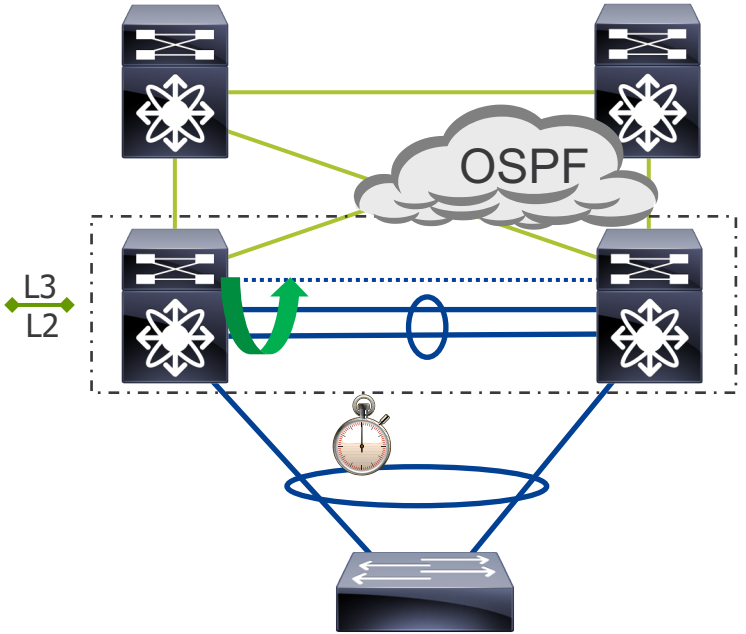
- When the vPC peer goes down or the vPC peer-link goes down, the SVIs are suspended.
- After restoration, the ARP table is empty so traffic gets blackholed.
- With this feature, before bringing up the SVIs, vPC peer devices synchronize ARP table over CFS.
- Reduces convergence time.



Nexus(config-vpc-domain)# ip arp synchronize

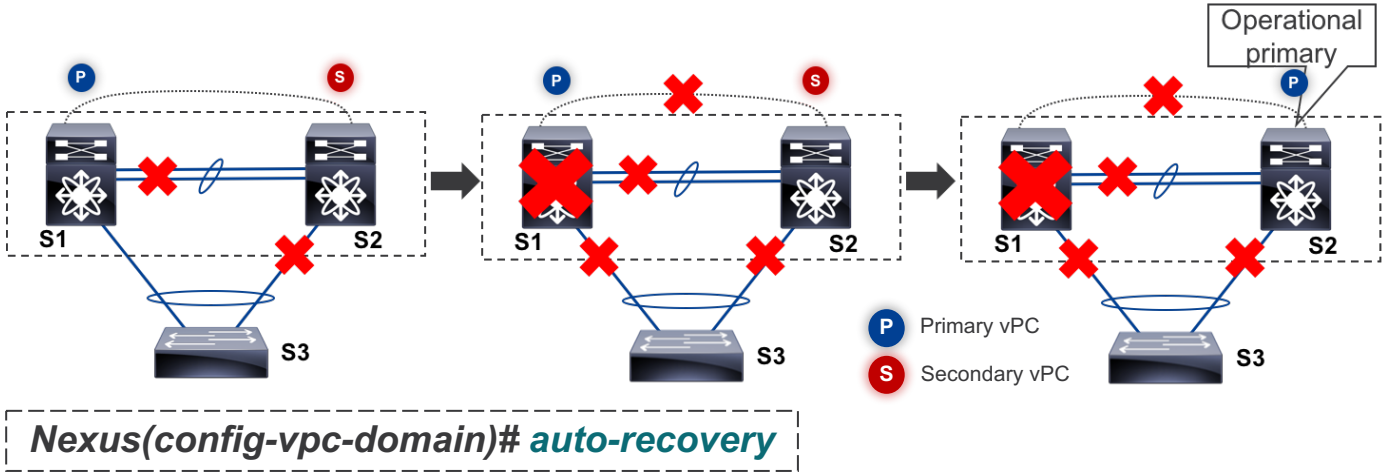
vPC delay restore

- After the vPC peer reloads, the traffic might be blackholed before L3 connectivity is re-established.
- The vPC link bringup can be delayed in order to allow L3 routing protocol convergence.
- The default time is 30 seconds.
- Accommodates for expansion modules or linecards boot time.



Nexus(config-vpc-domain)# delay restore <1-3600 sec>

vPC auto-recovery

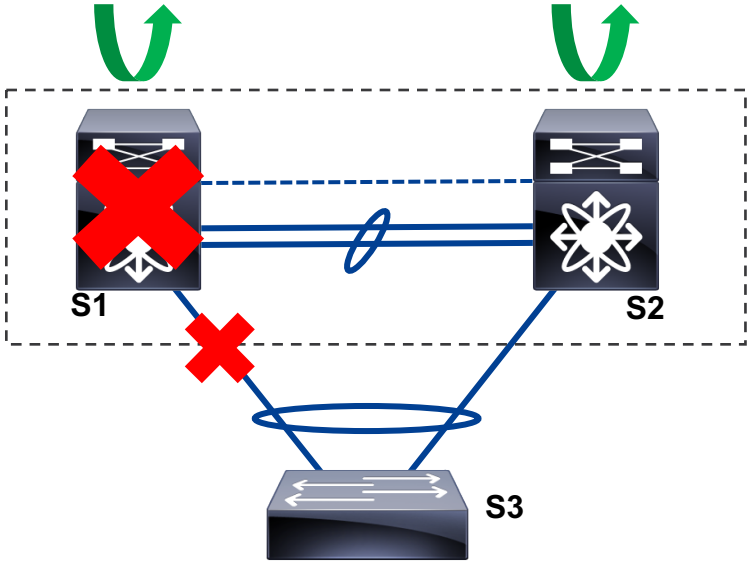


Nexus(config-vpc-domain)# auto-recovery

- Designed to address two scenarios: #1 where the vPC peer-link failure is followed by the vPC primary failure and #2 where both vPC peers reload, but only one of them comes back online.

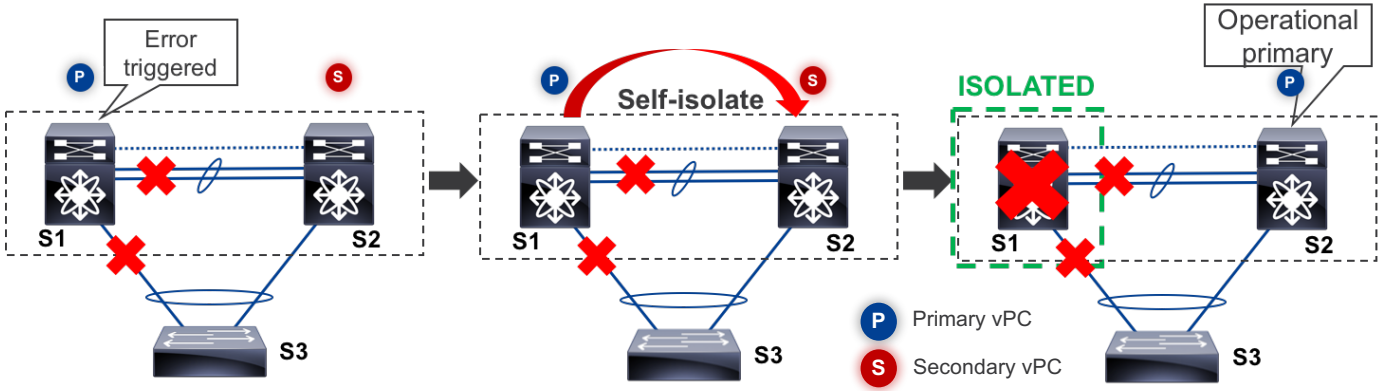
vPC auto-recovery reload delay

- Upon a reload, until the vPC peer-link is re-established between the vPC peers, vPC member ports are suspended.
- vPC auto-recovery reload delay allows “alive” vPC peer to assume primary role after the delay timer expires.
- The value for this timer is 240 seconds by default, but it can be tuned.



Nexus(config-vpc-domain)# auto-recovery reload-delay <240-3600 seconds>

vPC self-isolation



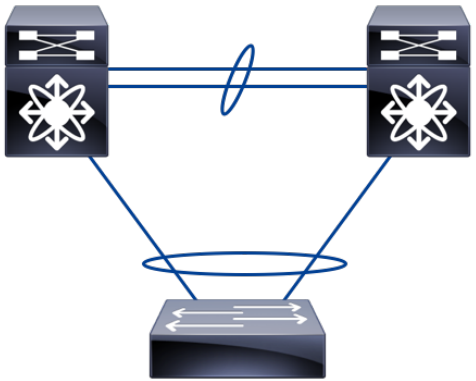
1. Error triggered: all line cards fail or all vlans down on vPC peer-link.
2. S1 sends “self-isolation” message through the vPC peer-keepalive.
3. S2 takes over as vPC operational primary and S1 is isolated.

vPC self-isolation - Configuration

```
vPC domain 100
peer-keepalive destination
10.126.216.44
  peer-gateway
  self-isolation
```

```
Switch# show vPC brief
<snip>
vPC domain id                  : 100
<snip>
vPC role                       : primary
<snip>
Self-isolation                 : Enabled
```

```
2015 Sep 29 22:33:03 S1 %$ VDC-1 %$
%vPC-2-ENTER_SELF_ISOLATION: Local
switch goes into self isolation
state due to all linecards failure.
Please resume failed linecards and
do shut/no shut on peer-link to exit
self-isolation state
```



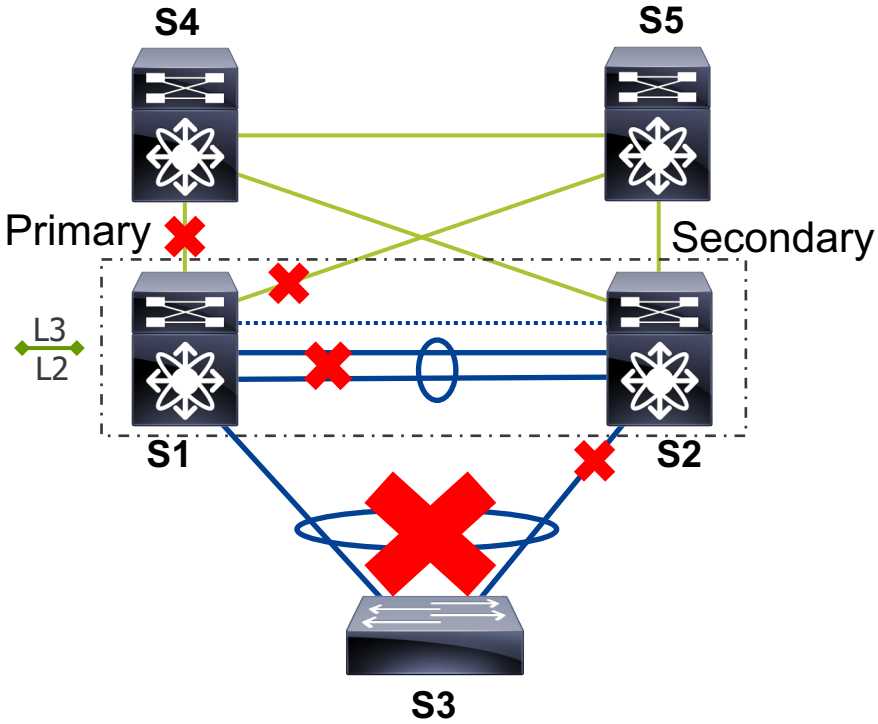
```
vPC domain 100
peer-keepalive destination
10.126.216.41
  peer-gateway
  self-isolation
```

```
Switch# show vPC brief
<snip>
vPC domain id                  : 100
<snip>
vPC role                       : secondary
<snip>
Self-isolation                 : Enabled
```

```
2015 Sep 30 10:33:14 S2 %$ VDC-1 %$
%vPC-2-ENTER_SELF_ISOLATION: Remote
switch goes into self isolation
state due to all linecards failure.
Please resume failed linecards and
do shut/no shut on peer-link to exit
self-isolation state
```

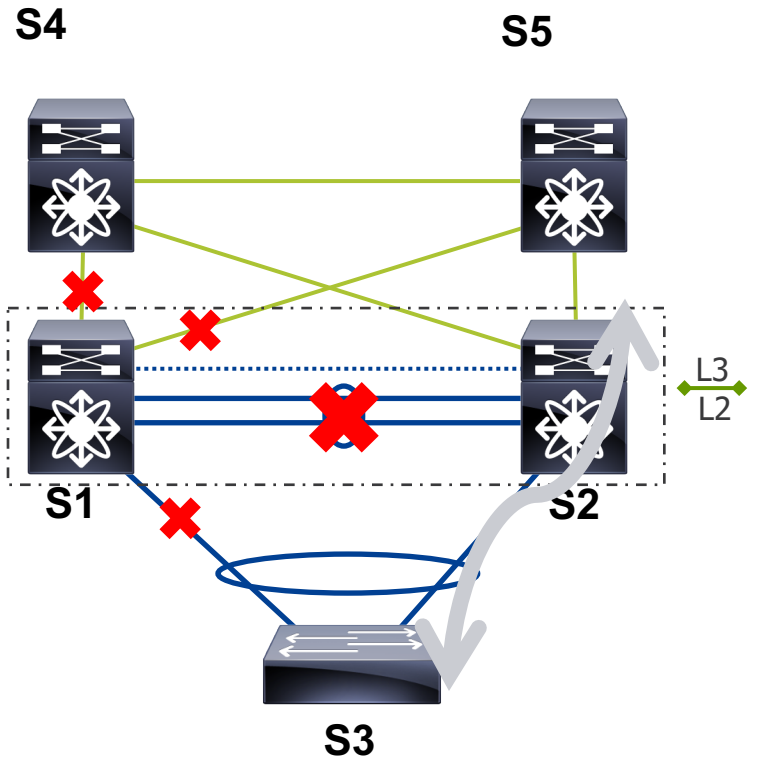
vPC object track

- Scenario: vPC peer-link and uplinks go down on the vPC primary switch.
- vPC secondary shuts down vPCs.
- vPC peer-keepalive is up so vPC auto-recovery does not kick in.
- Traffic is blackholed.
- vPC object track triggers the vPC pair of switches to do a failover.



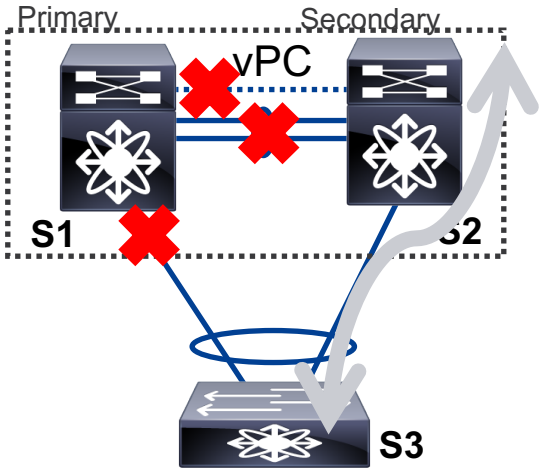
vPC object track - Configuration

```
! Track the vPC peer-link and uplinks
track 1 interface port-channel11 line-protocol
track 2 interface Ethernet1/1 line-protocol
track 3 interface Ethernet1/2 line-protocol

! Combine all tracked objects into one.
track 10 list boolean OR
object 1
object 2
object 3

! If object 10 goes down on the primary vPC
! peer, system will switch over to other vPC
! peer and disable all local vPCs.
vpc domain 1
track 10
```



vPC shutdown

- Isolates a switch from vPC complex for:
  - Debugging.
  - Troubleshooting.
  - Physical isolation.
- Minimal disruption for the traffic flows.
- “no shutdown” brings switch back in vPC.
- Configuration entry is persistent after reload.
- vPC peer-switch for convergence.




```
switch# configure terminal
switch(config)# vpc domain 100
switch(config-vpc)# shutdown
```

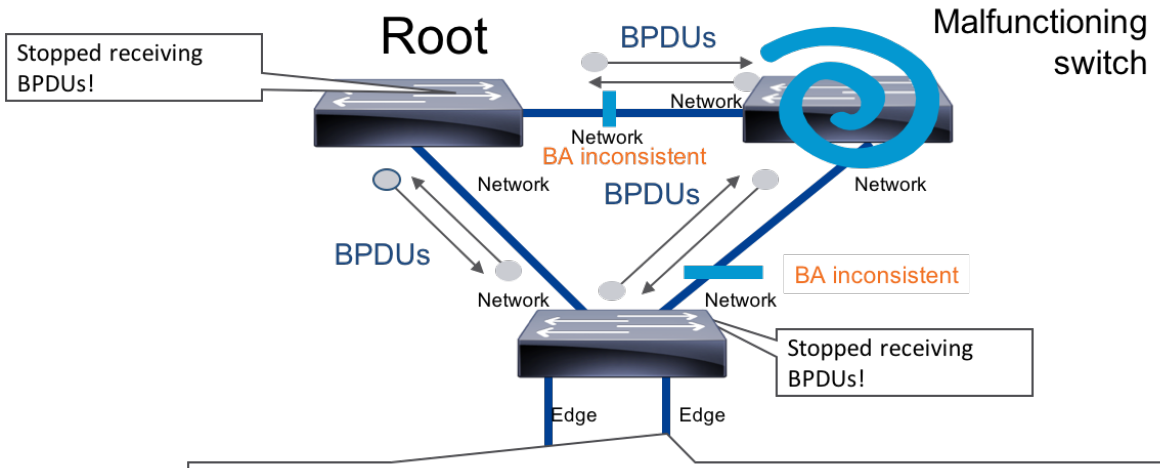
GIR

- Flexible framework providing a comprehensive, systematic method to isolate a Nexus switch node.
- Initial support for:
  - vPC/vPC+
  - IS-IS
  - OSPF
  - EIGRP
  - BGP
  - Interface

Platform	Release
Nexus 5500/5600/6000	NX-OS 7.1
Nexus 7000/7700	NX-OS 7.2

```
Nexus(config)# system mode maintenance
Nexus(config)# system mode normal
```

vPC and STP – Bridge assurance




```
%STP-2-BRIDGE_ASSURANCE_BLOCK: Bridge Assurance blocking port Ethernet2/48
VLAN0700
switch# show spanning vl 700 | in -i bkn
Eth2/48          Altn BKN*4         128.304  Network P2p *BA_Inc
```

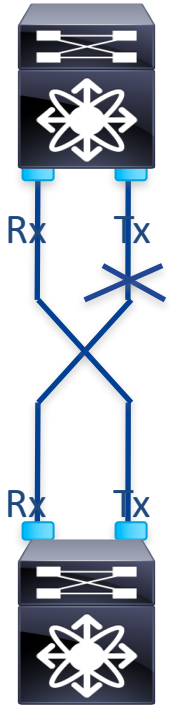
vPC and STP – Bridge assurance

- Turns STP into a bidirectional protocol.
- Ensures that STP fails “closed” rather than “open”.
- All ports with “network” port type send BPDUs regardless of state.
- If a network port stops receiving BPDUs, the port is placed in BA-Inconsistent state (blocked).
- Enabled on vPC peer-link, do **not** enable on vPCs (due to ISSU).

```
%STP-2-BRIDGE_ASSURANCE_BLOCK: Bridge Assurance blocking port Ethernet2/48 VLAN0700.
switch# show spanning vlan 700 | i i bkn
Eth2/48          Desg BKN*4         128.304  Network P2p *BA_Inc
```

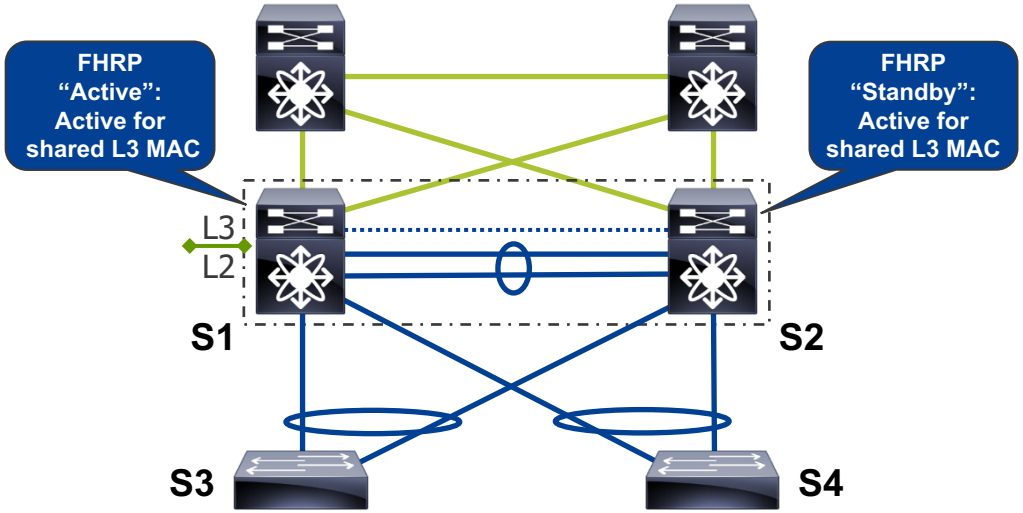
vPC and UDLD

- Lightweight L2 failure detection protocol.
- Designed to detect one-way communication due to physical or soft failure.
- Runs on the individual links, even when the links are bundled in a port-channel.
- Centralized implementation on switching platforms (**SW**).
- UDLD **not** recommended on vPC peer-link (STP BA).
- UDLD **not** recommended on vPC member ports if LACP is enabled. Only UDLD normal mode supported on vPC.



Design best practices

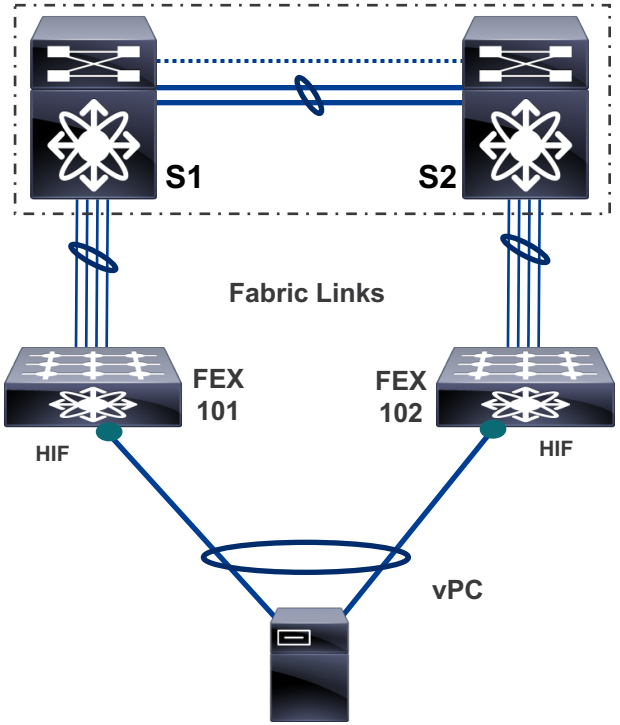
vPC and FHRP



- FHRP in active/active data-plane forwarding and active/standby control-plane with vPC.
- Use default FHRP timers, no need for aggressive timers.

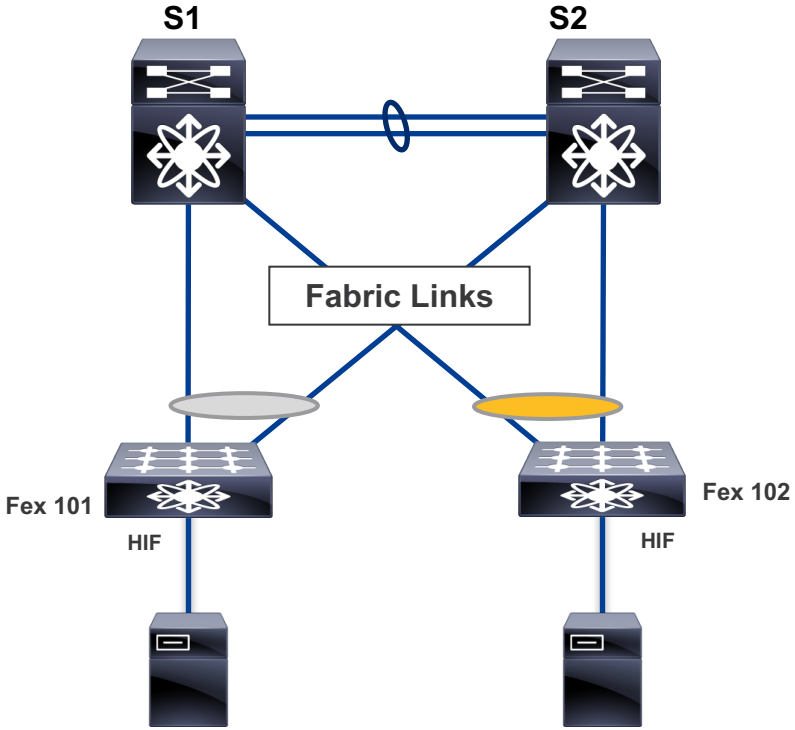
vPC and single-homed FEXs

- Port-channel connectivity from the server to the network.
- Two Nexus parent switches bundled into a vPC pair.
- FEXs single-homed to one of the Nexus parent switches.
- Suited for servers with dual NIC and capable of running port-channel configuration (LACP preferred).



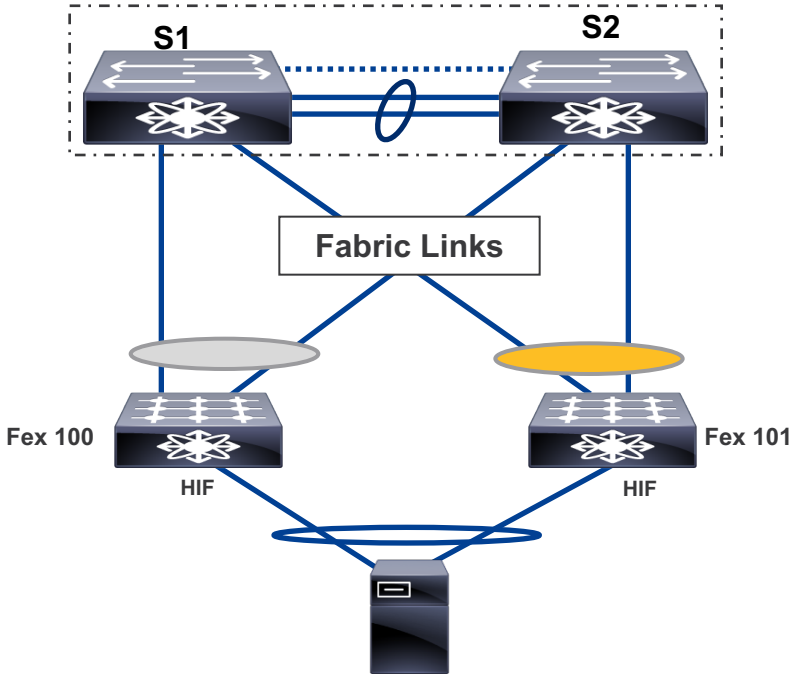
vPC and dual-homed FEXs

- Single interface connectivity from the server to the network.
- Two Nexus parent switches bundled into a vPC pair.
- FEXs dual-homed to both Nexus parent switches.
- Suited for servers with single NIC or dual NIC not having port-channel capability.
- Scale implications.

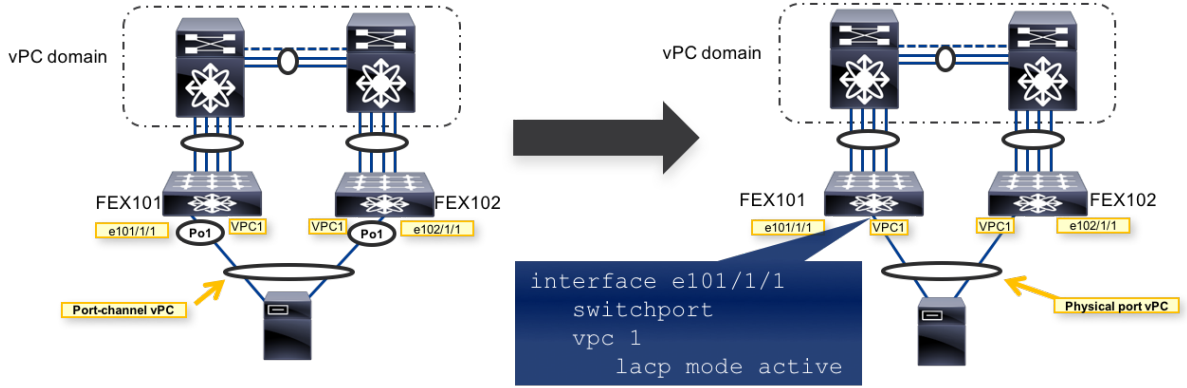


vPC and dual-homed FEXs – Enhanced vPC

- Port-channel connectivity from the server to the network.
- Two Nexus parent switches bundled into a vPC pair.
- FEXs dual-homed to both Nexus parent switches.
- Suited for servers with dual NIC and capable of running port-channel configuration (LACP preferred).
- Not supported on Nexus 7000/7700.



Physical port vPC

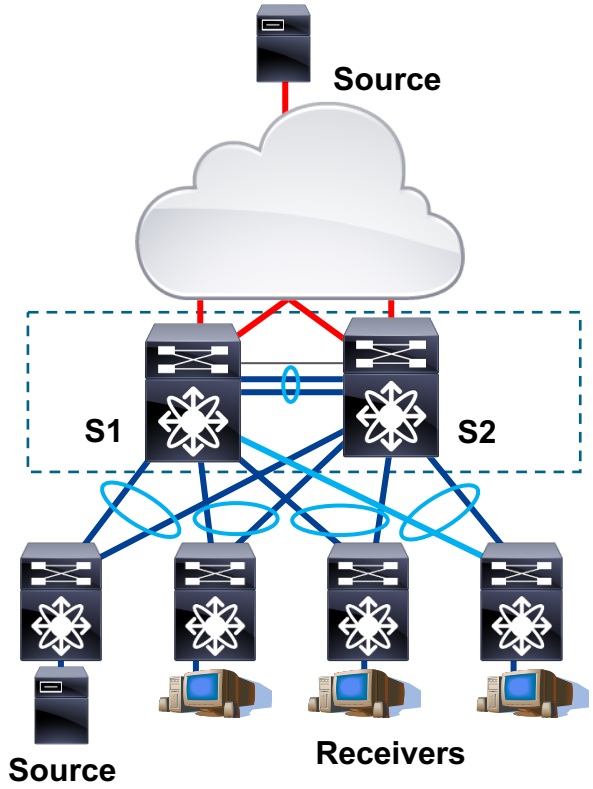


- vPC configuration on a physical L2 port as opposed to a port-channel.
- Improves scaling as separate port-channel interface not created for single-link vPC member port => more than 1000 port vPCs.

vPC and multicast

- vPC supports **only** PIM-ASM.
- vPC uses CFS for IGMP states sync.
- Scenario #1, source in vPC domain:
  - both vPC peers are active forwarders.
  - duplicates avoided via vPC loop-avoidance
- Scenario #2, source in L3 cloud:
  - active forwarder elected based on the unicast metric for the source (election per source, per group).

Nexus# show ip pim internal vpc rpf-source vrf default



L3 over vPC

Problem

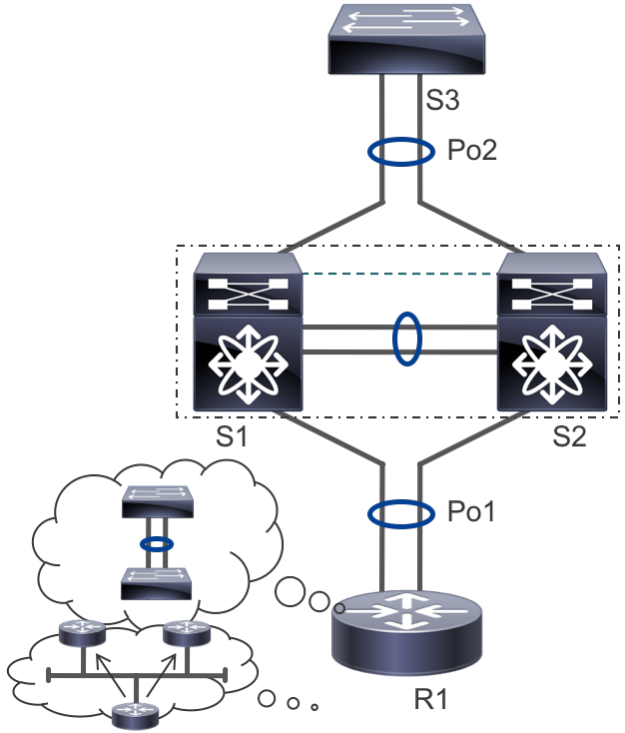
On S1, the frame from R1 to S3 ingresses from the vPC peer-link (from S2) and has to egress over the vPC to S3.

Frame will **only** be forwarded out when received from the vPC peer-link if:

1. Outgoing interface is **not** a vPC.

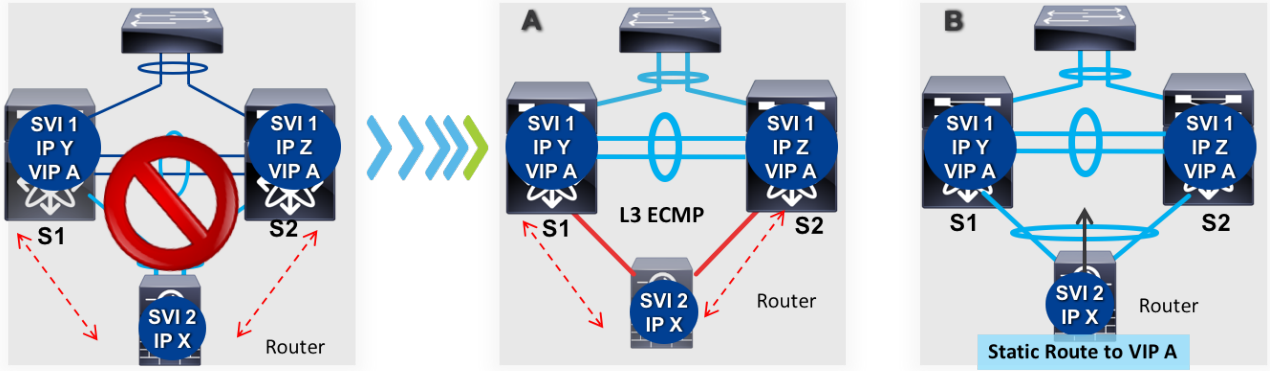
   or

2. Outgoing vPC doesn't have active vPC leg on other vPC peer (S2).



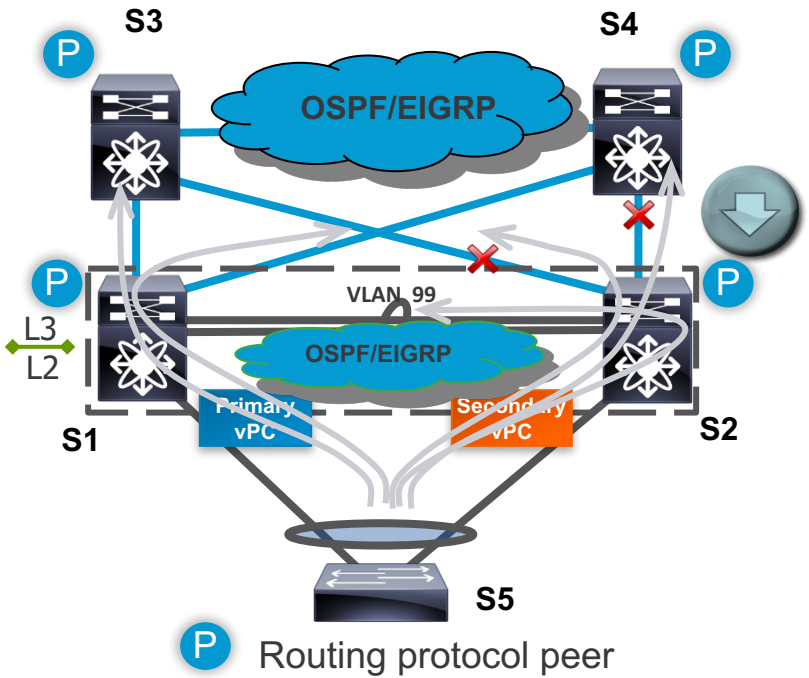
Workaround

- **Not** recommended to attach L3 devices to the vPC domain via a L2 port-channel when routing peering with the vPC peers is needed.
- Common workarounds: individual L3 routed links or static routes with FHRP VIP as next-hop.



Recommendations

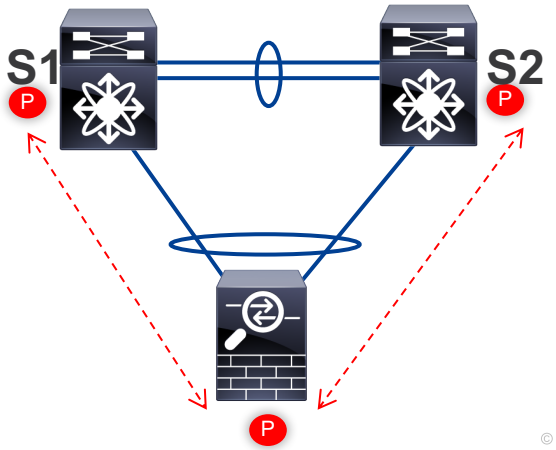
- One point-to-point dynamic routing protocol adjacency between the vPC peers having a L3 backup path to the core in case of uplinks failure.
- Use a separate SVI (vlan used solely for this purpose) for the backup path over the vPC peer-link and configure the other SVIs as passive.



Enhancements

- Dynamic peering between Layer 3 device and vPC peers over vPC.
- Traffic does not get decremented TTL if it traverses the vPC peer-link.
- vPC peer-gateway must be enabled.

```
S1
vpc domain 200
  peer-keepalive destination
10.10.12.42 source 10.10.12.52
  peer-gateway
  layer3 peer-router
```




```
S2
vpc domain 200
  peer-keepalive destination
10.10.12.52 source 10.10.12.42
  peer-gateway
  layer3 peer-router
```

Troubleshooting

Prerequisites

For effective troubleshooting you need to gather as many details as possible for the affected traffic flow:

- L2 (Src_MAC, Dst_MAC, VLAN_ID)
- L3 (Src_IPv4/IPv6, Dst_IPv4/IPv6, Protocol)
- L4 (Src_Port, Dst_Port)
- L7 (Application)

Hint: get also a non-affected flow for states comparison.

vPC

show vpc

```
n5k-1# show vpc
vPC domain id                     : 34
vPC keep-alive status             : peer is alive
vPC fabricpath status             : peer is reachable through fabricpath
Configuration consistency status  : success
vPC role                          : primary
Peer Gateway                      : Enabled
vPC Peer-link status
---------------------------------------------------------------------
id   Port   Status Active vlans
1    Po1    up     1234
---------------------------------------------------------------------
vPC status
id     Port        Status Consistency Reason       Active vlans vPC+ Attrib
103    Po103       up     success     success      1234         DF: Partial,
```

vPCM

show system internal vpcm info

n7700-1# show system internal vpcm info interface port-channel 101

if_index: 0x16000064
Number of Up Vlans: 1, Bitset: 1234
Number of Suspended Vlans: 0, Bitset:
Peer if_index: 0x16000064
Peer state: Up
Number of Up VLANs on peer: 1, Bitset: 1234

n5k-1# show system internal vpcm info interface port-channel 103

if_index: 0x16000066
Number of Up Vlans: 1, Bitset: 1234
Number of Suspended Vlans: 0, Bitset:
Peer state: Up
Number of Up VLANs on peer: 1, Bitset: 1234

vPC consistency

show vpc consistency-parameters

n5k-2# show vpc consistency-parameters global

Name	Type	Local Value	Peer Value
Network QoS (MTU)	2	(1538, 0, 0, 0, 0, 0)	(1538, 0, 0, 0, 0, 0)
STP Mode	1	Rapid-PVST	Rapid-PVST
Allowed VLANs	-	1234	1234
Local suspended VLANs	-	-	-

n5k-2# show vpc consistency-parameters vpc 103

Name	Type	Local Value	Peer Value
STP Port Type	1	Default	Default
mode	1	active	active
Port Mode	1	trunk	trunk
MTU	1	1500	1500
Allowed VLANs	-	1234	1234
Local suspended VLANs	-	-	-

CFS

show cfs peers

show cfs internal ethernet-peer statistics

```
n5k-1# show cfs peers
Switch WWN            IP Address
 20:00:8c:60:4f:c7:7d:40 17.3.13.26                  [Local] n5k-1
 20:00:8c:60:4f:ba:eb:40 17.3.13.30

n5k-2# show cfs internal ethernet-peer statistics | i i "trans|rec" | exc i "\:\ 0" | exc i "\:0”
Number of Segments Transmitted                  : 432
Total Number of Segments Received               : 418
Number of fragmented segments Received             : 14
Number of unfragmented segments Received           : 404
Number of Unreliable segments Transmitted          : 1354
```

Port-channel

show port-channel summary

```
n5k-2# show port-channel summary interface port-channel 103
Flags:  D - Down        P - Up in port-channel (members)
--------------------------------------------------------------------------------
Group Port-       Type     Protocol  Member Ports
      Channel
--------------------------------------------------------------------------------
103   Po103(SU)   Eth      LACP      Eth1/6(P)

n5k-2# show lacp counters interface port-channel 103
                LACPDUs         Marker      Marker Response    LACPDUs
Port            Sent   Recv     Sent   Recv     Sent   Recv      Pkts Err
--------------------------------------------------------------------
port-channel103
Ethernet1/6       2991   2990     0      0        0      0         0
```

STP

show spanning-tree interface

n5k-2# show spanning-tree interface po103
Vlan Role Sts Cost Prio.Nbr Type
VLAN1234 Desg FWD 1 128.4198 (vPC) P2p Peer(STP)

n5k-2# show platform fwm info vlanif 1234 port-channel 103
vlanif vlan 1.1234 if 16000066 stp state: forwarding

n7700-1# slot 1 quoted "show hardware internal mac port 6 table cbl vlan"
| INGRESS |
| Forwarding State | 1234,4032-4035 |
| EGRESS |
| Forwarding State | 1234,4032-4035 |

ETHPM

show interface trunk

n5k-1# show system internal ethpm info interface port-channel 103 | i i vlans
Allowed Vlans: 1234
Operational Vlans: 1234

n5k-1# show interface port-channel 103 trunk
Port Vlans Allowed on Trunk
Po103 1234
Port Vlans Err-disabled on Trunk
Po103 none
Port STP Forwarding
Po103 1234

MAC address table

show mac address-table

show platform fwm info hw-stm

```
n5k-1# show mac address-table address 0015.621c.a942 vlan 1234
Legend:
        * - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
        age - seconds since last seen,+ - primary entry using vPC Peer-Link
   VLAN     MAC Address      Type      age     Secure NTFY   Ports/SWID.SSID.LID
---------+-----------------+--------+---------+------+----+------------------
* 1234     0015.621c.a942    dynamic   1800      F    F  Po103
```

```
n5k-1# show platform fwm info hw-stm | i i vlan|0015.621c.a942
VLAN     MAC Address      Port          loc      misc    cdce
1.1234   0015.621c.a942   Po103          1:9196:0  1:0:1   2.0.22.0.0.b (e:0)
```

FWM

show platform fwm info mac <MAC> <VLAN-ID>

n5k-1# show platform fwm info mac 0015.621c.a942 1234
mac vlan 1.1234 mac 0015.621c.a942: vlan 1.1234
mac vlan 1.1234 mac 0015.621c.a942: learned-on Po103 age 1800 ref_map = 'vlan if'
mac vlan 1.1234 mac 0015.621c.a942: nohit_count 0 hw_programmed 1 mac_clone 0
mac vlan 1.1234 mac 0015.621c.a942: old_if_index 'null'
mac vlan 1.1234 mac 0015.621c.a942: pss_flags 0
mac vlan 1.1234 mac 0015.621c.a942 cfg attrs - not-cli-cfg not-static movable no-drop no-regmac non-netstack-learnt
mac vlan 1.1234 mac 0015.621c.a942: mcec_flags 0x5, mac_info_flags 0, rem_if 0x7e000067, sync_count 0 rcv_count 6
mac vlan 1.1234 mac 0015.621c.a942: CDCE Address b:0:0:22:0:2

FWM – MAC address history

show platform fwm info mac <MAC> <VLAN-ID>

n5k-1# show platform fwm info mac 0015.621c.a942 1234 | i i 18:07:27 prev 2
Mac history (Last 35 operations):
  Operation: Peer sent and local vpc learnt (1)
      (flags: Loc (0x1) mac_info_flags (0x0) if: 0x16000066 hint: 0)
      at Sat May 6 18:07:27 2017

n5k-2# show platform fwm info mac 0015.621c.a942 1234 | i i "18:07:27" p 2
Mac history (Last 35 operations):
  Operation: Mac learned from hw (40)
      (flags: Loc (0x1) mac_info_flags (0x0) if: 0x16000066 hint: 0)
      at Sat May 6 18:07:27 2017
  Operation: Mac sent to peer on local learn (15)
      (flags: Loc (0x1) mac_info_flags (0x0) if: 0x16000066 hint: 0)
      at Sat May 6 18:07:27 2017

MAC address table

show mac address-table

show hardware mac address-table

n7700-1# show mac address-table address 64f6.9d26.7073 vlan 1234 hardware-age

```
VLAN     MAC Address      Type      age     Secure NTFY Ports/SWID.SSID.LID
* 1234    64f6.9d26.7073   dynamic   30        F    F  Po101
```

n7700-1# show hardware mac address-table 1 address 64f6.9d26.7073 vlan 1234

```
FE | Valid| PI|  BD  |      MAC       |  Index| Stat| SW  | Modi| Age| Tmr|UC| SWID| SSWID|  LID
     |      |    |       |                |       |  ic   |       | fied  |Byte| Sel|     |        |          |
0    1     1   22    64f6.9d26.7073  0x00e02   0   0x089   0    19    1  1  0x00c  0x00b 0x00e02
```

n7700-1# show system internal pixm info ltl-range start-ltl 0x00e02 ltl-cnt 1

```
LTL    IFIDX       PORT/LTL_TYPE
0x0e02 0x16000064  Po101
```

L2FM

show system internal l2fm macdb|l2dbg

n7700-1# show system internal l2fm info macdb vlan 1234

```
Legend:
        P - VPC/ES peer is also owner, E8 - Earl8 owner, H - Has Hier Mac, ES - Has ES owner, EP - ES peer has entry
        LL - Lazy Learn for MCEC case O - OTV Entry
 VLAN     MAC Address      Ports  Flags    N_Flags   [P E8 H ES EP LL O AM] AgedBmp      Owners [slot.fe.hints]/ES Owners
-------+----------------+---------------+--------+-----------+-----------------------------------------------
 1234    64f6.9d26.7073  Po101     00c.00b.0e02 0x4500103    0x60 [10011000] bm[00.00.1111] [ES101]
```

n7700-1# show system internal l2fm l2dbg macdb address 64f6.9d26.7073 vlan 1234

```
 VLAN: 1234 MAC: 64f6.9d26.7073 FE ID: 0
  Time                  If/swid    Db Op                Src Slot  FE  HW_ADDR Count
    Sun May  7 01:18:40 2017 0x16000064 0  AGE                3    19   0    0
```

L2FM – MAC address history

show system internal l2fm l2dbg macdb

n7700-1# show system internal l2fm l2dbg macdb address 64f6.9d26.7073 vlan 1234
Db: 0-MACDB, 1-GWMACDB, 2-SMACDB, 3-RMDB, 4-SECMACDB
Src: 0-UNKNOWN, 1-L2FM, **2-PEER**, 3-LC, 4-HSRP
VLAN: 1234 MAC: 64f6.9d26.7073

Time	If/swid	Db	Op	**Src**	Slot	FE	HW_
Sun May 7 22:17:18 2017	0x16000064	0	INSERT	**2**	19	0	0

n7700-2# show system internal l2fm l2dbg macdb address 64f6.9d26.7073 vlan 1234
Db: 0-MACDB, 1-GWMACDB, 2-SMACDB, 3-RMDB, 4-SECMACDB
Src: 0-UNKNOWN, 1-L2FM, 2-PEER, **3-LC**, 4-HSRP
VLAN: 1234 MAC: 64f6.9d26.7073

Time	If/swid	Db	Op	**Src**	Slot	FE	HW_
Sun May 7 22:17:18 2017	0x16000064	0	INSERT	**3**	0	2	0

Tools

Ethanalyzer

- Applicable for both Nexus 5600/6000 and 7000/7700 switches.
- Frame capture for control-plane traffic received/sent on/from the supervisor main CPU (does **not** capture data-plane traffic).
- Captures a defined number of frames/packets that meet the filter parameters (by default the limit is 10 frames/packets).
- On Nexus 5600/6000 it can be **only** used with display filters (capture filters are broken at the moment).

See more information here.

Ethanalyzer (continued)

ARP Request frame/packet capture (simple):

N6000# ethanalyzer local interface inbound-low display-filter "arp.src.proto_ipv4 == 10.10.12.1" limit-c 0
2017-01-01 18:15:30.261685 b0:aa:77:49:1d:3c -> ff:ff:ff:ff:ff:ff ARP Who has 10.10.12.2? Tell 10.10.12.1

ARP Reply frame/packet capture (detailed):

N6000# ethanalyzer local interface inbound-hi decode-internal display-filter "arp.src.proto_ipv4 == 10.10.12.1 or arp.dst.proto_ipv4 == 10.10.12.1" limit-captured-frames 0 detail

CDCE SA: sid_lo: bd ssid: 0 lid: d
s_l3if: 0 tr_opt: 0, bd: 15, sup_src_if: 13, sup_dst: 2
802.1Q Virtual LAN, PRI: 0, CFI: 0, ID: 15

lid: d is in hex => 13 (0xd)
sup_src_if: 13 is the ingress interface
ID: 15 is the internal vlan

N6000# show platform fwm info vlan all | i i "int-vlan 15"
vlan 1.12 pd: int-vlan 15 state table idx 18 vacl_label 2048 mbr_bitmap_idx 0, vlan_flags 0x0

N6000# show platform fwm info lif all | i i "local_id 13"
Eth2/2 pd: local_id 13 endnode_id 0 endnode_id_alloced 1 vif_id 0

ELAM

- Applicable for Nexus 5600/6000 and 7000/7700 switches.
- Frame capture for both data-plane and control-plane traffic in Bigsur ASIC (Nexus 5600/6000), Orion ASIC (Nexus 7000 F1), Clipper ASIC (Nexus 7000 F2/F2E), Flanker ASIC (Nexus 7000/7700 F3).
- Captures **only** the first frame/packet that meets the trigger parameters, after the trigger is armed.

See more information here.

ELAM – Nexus 5600/6000 – Ingress LU

Get the ingress Bigsur ASIC number relative to the slot number:

N6000# show platform fwm info pif eth2/2 | i i slot_asic
Eth2/2 pd: slot 1 logical port num 1 slot_asic_num 0 global_asic_num 5 fw_inst 4 phy_fw_inst 1 fc 0

Set a LU trigger for an ingress frame/packet capture:

N6000# elam slot 2 asic bigsur instance 0
N6000(bigsur-elam)# trigger lu ingress ipv4 if source-ipv4-address_ipv4 10.10.12.1 destination-ipv4-address_ipv4 …
N6000(bigsur-elam)# start capture
N6000(bigsur-elam)# show capture lu
ELAM: Nothing captured
N6000(bigsur-elam)# show capture lu
Ingress Interface: Ethernet2/2 IS NOT A PC
N6000(bigsur-elam)# show capture rs
Egress Interface: Ethernet1/4/1 IS NOT A PC
N6000(bigsur-elam)# stop capture

ELAM – Nexus 5600/6000 – Egress LU

Get the egress Bigsur ASIC number relative to the slot number:

N6000# show platform fwm info pif eth1/4/1 | i i slot_asic
Eth1/4/1 pd: slot 0 logical port num 12 slot_asic_num 1 global_asic_num 2 fw_inst 0 phy_fw_inst 0 fc 0

Set a LU trigger for an egress frame/packet capture:

N6000# elam slot 1 asic bigsur instance 1
N6000(bigsur-elam)# trigger lu egress ipv4 if source-ipv4-address_ipv4 10.10.12.1 destination-ipv4-address_ipv4 …
N6000(bigsur-elam)# start capture
N6000(bigsur-elam)# show capture lu
ELAM: Nothing captured
N6000(bigsur-elam)# show capture lu
Egress Interface: Ethernet1/4/1 IS NOT A PC
N6000(bigsur-elam)# stop capture

ELAM – Nexus 5600/6000 – Ingress RS

Get the ingress Bigsur ASIC number relative to the slot number:

N6000# show platform fwm info pif eth2/2 | i i slot_asic
Eth2/2 pd: slot 1 logical port num 1 slot_asic_num 0 global_asic_num 5 fw_inst 4 phy_fw_inst 1 fc 0

Set a RS trigger for an ingress frame/packet capture:

N6000# elam slot 2 asic bigsur instance 0
N6000(bigsur-elam)# trigger rs 1 ingress ipv4 if out_mcast 0x1 out_drop 0x1 source-ipv4-address 10.10.12.1 …
N6000(bigsur-elam)# start capture
N6000(bigsur-elam)# show capture lu
Ingress Interface: Ethernet2/2 IS NOT A PC
N6000(bigsur-elam)# show capture rs
Egress Interface: Ethernet1/4/1 IS NOT A PC
| OUT_MCAST | 1 |
| OUT_DROP | 1 |
N6000(bigsur-elam)# stop capture

N.B. Trigger parameters out_mcast and out_drop are hidden parameters of the command, but can be used as part of the forwarding lookup results.

ELAM – Nexus 7000/7700 F3 - Ingress L2

Get the ingress Flanker ASIC number:

n7700-2# slot 1 quoted "show hardware internal dev-port-map" | i i L2LKP|21
FP port | PHYS | MAC_0 | **L2LKP** | L3LKP | QUEUE |SWICHF
21 2 2 **2** 2 2 0,1

Set a L2 trigger for an ingress frame/packet capture:

n7700-2# attach module 1
module-1# elam asic flanker instance 2
module-1(fln-elam)# layer2
module-1(fln-l2-elam)# trigger dbus ipv4 ingress if source-ipv4-address 10.0.0.101 destination-ipv4-address ...
module-1(fln-l2-elam)# trigger rbus ingress if trig
module-1(fln-l2-elam)# start
module-1(fln-l2-elam)# stat
L2 DBUS: Armed L2 RBUS: Armed
module-1(fln-l2-elam)# stat
L2 DBUS: Triggered L2 RBUS: Triggered

ELAM – Nexus 7000/7700 F3 - Ingress L2

Check the header before the L2 lookup:

module-1(fln-l2-elam)# show dbus | i i "\-ip|-mac" | exc "mim“
source-ipv4-address: 10.0.0.101
destination-ipv4-address: 239.10.0.103
destination-mac-address : 0100.5e0a.0067
source-mac-address : 64f6.9d26.7073

Check the result after the L2 lookup:

module-1(fln-l2-elam)# show rbus | i i di-ltl
di-ltl-index : 0xbfde l3-multicast-di : 0x0

n7700-2# show system internal pixm info ltl-range start-ltl 0xbfde ltl-cnt 1
LTL IFIDX PORT/LTL_TYPE

0xbfde 0x00000004 MCAST_GROUP

PACLs/RACLs

- Applicable for Nexus 5500/5600/6000 and 7000/7700 switches.
- Frame/Packet permit/deny/log with counters for both data-plane and control-plane traffic at the hardware(ASIC) level.
- Can **only** be applied in the ingress direction.
- PACLs are to be applied on L2 ports/port-channels.
- RACLs are to be applied on L3 ports/port-channels.
- The counters are very useful in tracking packet loss throughout the network (need to configure “statistics per-entry” in the ACL).

PACLs/RACLs (continued) - PACL

```
N6000(config)# ip access-list IPv4_PACL_Eth2/2_IN
N6000(config-acl)# 1000 permit ip any any
N6000(config-acl)# 10 permit icmp 10.10.12.1/32 10.10.34.4/32
N6000(config-acl)# statistics per-entry
N6000(config)# ipv6 access-list IPv6_PACL_Eth2/2_IN
N6000(config-ipv6-acl)# 1000 permit ipv6 any any
N6000(config-ipv6-acl)# 10 permit ipv6 fc00:10:10:12::1/128 fc00:10:10:34::4/128
N6000(config-ipv6-acl)# statistics per-entry

N6000(config)# interface eth2/2
N6000(config-if)# ip port access-group IPv4_PACL_Eth2/2_IN in
N6000(config-if)# ipv6 port traffic-filter IPv6_PACL_Eth2/2_IN in

N6000# show ip access-lists IPv4_PACL_Eth2/2_IN ; show ipv6 access-lists IPv6_PACL_Eth2/2_IN
IPV4 ACL IPv4_PACL_Eth2/2_IN
        10 permit icmp 10.10.12.1/32 10.10.34.4/32 [match=5]
IPV6 ACL IPv6_PACL_Eth2/2_IN
        10 permit ipv6 fc00:10:10:12::1/128 fc00:10:10:34::4/128 [match=10]
```

PACLs/RACLs (continued) - RACL

```
N6000(config-if)# ip access-list IPv4_RACL_Vlan12_IN
N6000(config-acl)# 1000 permit ip any any
N6000(config-acl)# 10 permit icmp 10.10.12.1/32 10.10.34.4/32
N6000(config-acl)# statistics per-entry
N6000(config# ipv6 access-list IPv6_RACL_Vlan12_IN
N6000(config-ipv6-acl)# 1000 permit ipv6 any any
N6000(config-ipv6-acl)# 10 permit ipv6 fc00:10:10:12::1/128 fc00:10:10:34::4/128
N6000(config-ipv6-acl)# statistics per-entry

N6000(config)# interface Vlan12
N6000(config-if)# ip access-group IPv4_RACL_Vlan12_IN in
N6000(config-if)# ipv6 traffic-filter IPv6_RACL_Vlan12_IN in

N6000# show ip access-lists IPv4_RACL_Vlan12_IN ; show ipv6 access-lists IPv6_RACL_Vlan12_IN
IPV4 ACL IPv4_RACL_Vlan12_IN
        10 permit icmp 10.10.12.1/32 10.10.34.4/32 [match=5]
IPV6 ACL IPv6_RACL_Vlan12_IN
        10 permit ipv6 fc00:10:10:12::1/128 fc00:10:10:34::4/128 [match=10]
```

Relevant tech-support

Nexus 5600/6000

vPC

show tech-support vpc

CFS

show tech-support cfs

FWM

show tech-support fwm

STP

show tech-support stp

LACP

show tech-support lacp

Nexus 7000/7700

vPC

show tech-support vpc

CFS

show tech-support cfs

L2FM

show tech-support l2fm detail

STP

show tech-support stp

LACP

show tech-support lacp

Nexus 7000/7700 (continued)

PIXM

show tech-support pixm

ELTM

show tech-support eltm

MCM

show tech-support mcm

VLAN

show tech-support vlan

Port-channel

show tech-support port-channel

Nexus 5600/6000 and Nexus 7000/7700

TAC-PAC

show tech-support details

or

tac-pac bootflash:$(SWITCHNAME)-tac-pac.txt.gz


CISCO