



NEXUS 7000 – TROUBLESHOOTING Session 2

Presenter: RICHARD MICHAEL

Panelists: SOLOMON SUDHAKAR



AGENDA

➤ VPC/VDC

COMMON ISSUES:

- Loop - Prevention Techniques
- Supported/Un-Supported Topologies
- ISSU Failures/Upgrade
- Module Failures
- Fabric Troubleshooting
- MTS – Memory Leaks
- High CPU
- Layer 2 / Layer 3 Check
- ACL TCAM Troubleshooting
- Proxy Routing – Limitations
- ELAM/Ethalyzer – Troubleshooting



VDC

- What is a VDC?

A Logical partition of one single physical switch where in each partition will behave as a separate switch with one Infrastructure and Kernel shared between them.

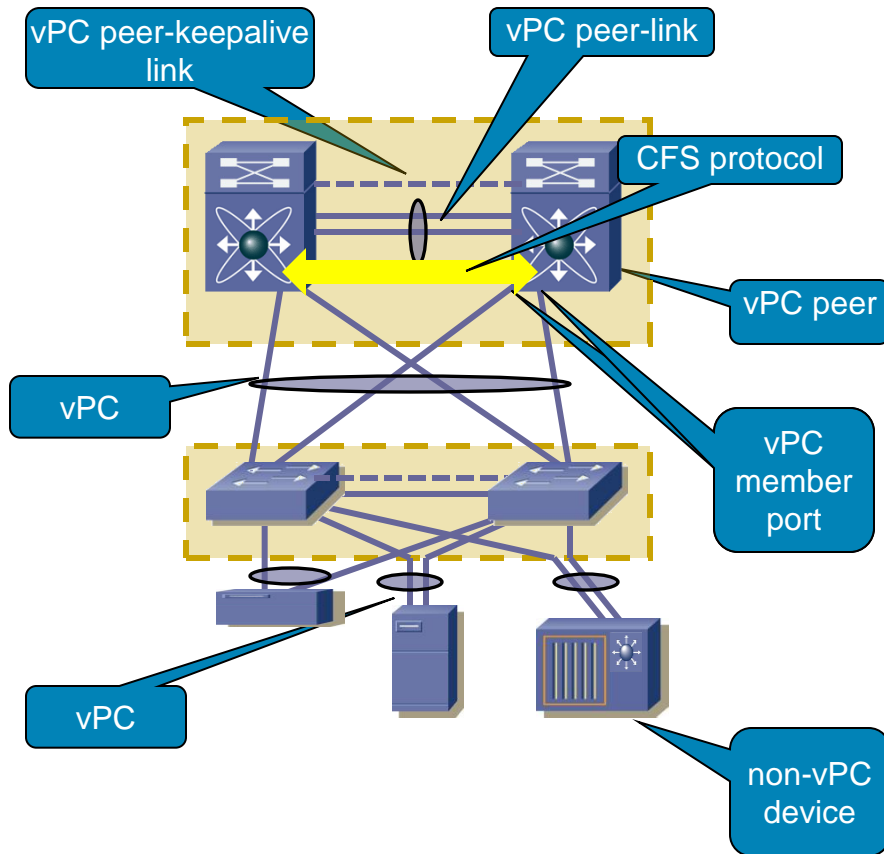
- Supervisor 1 - Four VDCs + 1 Admin VDC*
 - Requires 8GB of RAM
- Supervisor 2 - Four VDCs + 1 Admin VDC (4+1)
- Supervisor 2E - Eight VDCs + 1 Admin VDC (8+1)
- VDCs beyond 4 require additional license

VDC

- http://www.cisco.com/c/en/us/td/docs/switches/data_center/sw/verified_scalability/b_Cisco_Nexus_7000_Series_NX-OS_Verified_Scalability_Guide.html
- BRKDCT-2121 – Cisco Live

Feature Overview & Terminology

vPC Terminology



- **vPC peer** – a vPC switch, one of a pair
- **vPC member port** – one of a set of ports (port channels) that form a vPC
- **vPC** – the combined port channel between the vPC peers and the downstream device
- **vPC peer-link** – Link used to synchronize state between vPC peer devices, must be 10GbE
- **vPC peer-keepalive link** – the keepalive link between vPC peer devices, i.e., backup to the vPC peer-link
- **vPC VLAN** – one of the VLANs carried over the peer-link and used to communicate via vPC with a peer device.
- **non-vPC VLAN** – One of the STP VLANs not carried over the peer-link
- **CFS** – Cisco Fabric Services protocol, used for state synchronization and configuration validation between vPC peer devices

Building a vPC Domain

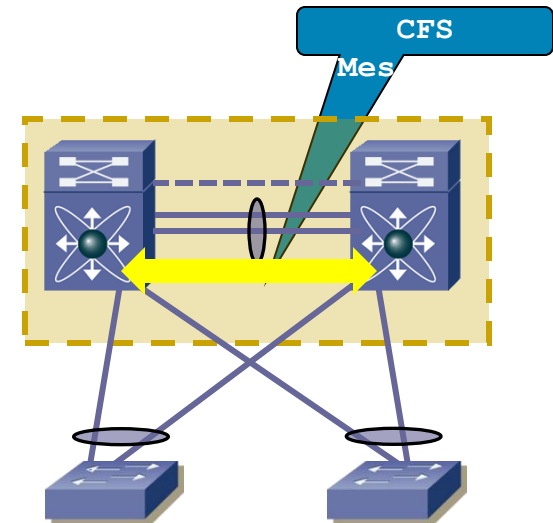
Cisco Fabric Services (CFS)

■ Definition/Uses:

- ✓ Configuration validation/comparison
- ✓ MAC member port synchronization
- ✓ vPC member port status
- ✓ STP Management
- ✓ HSRP and IGMP snooping synchronization
- ✓ vPC status

■ Characteristics:

- ✓ Transparently enabled with vPC features
- ✓ CFS messages encapsulated in standard Ethernet frames delivered between peers exclusively on the peer-link
- ✓ Cisco Fabric Services messages are tagged as CoS=4 for reliable communication.
- ✓ Based on CFS from MDS product development
- ✓ Many years in service, robust protocol

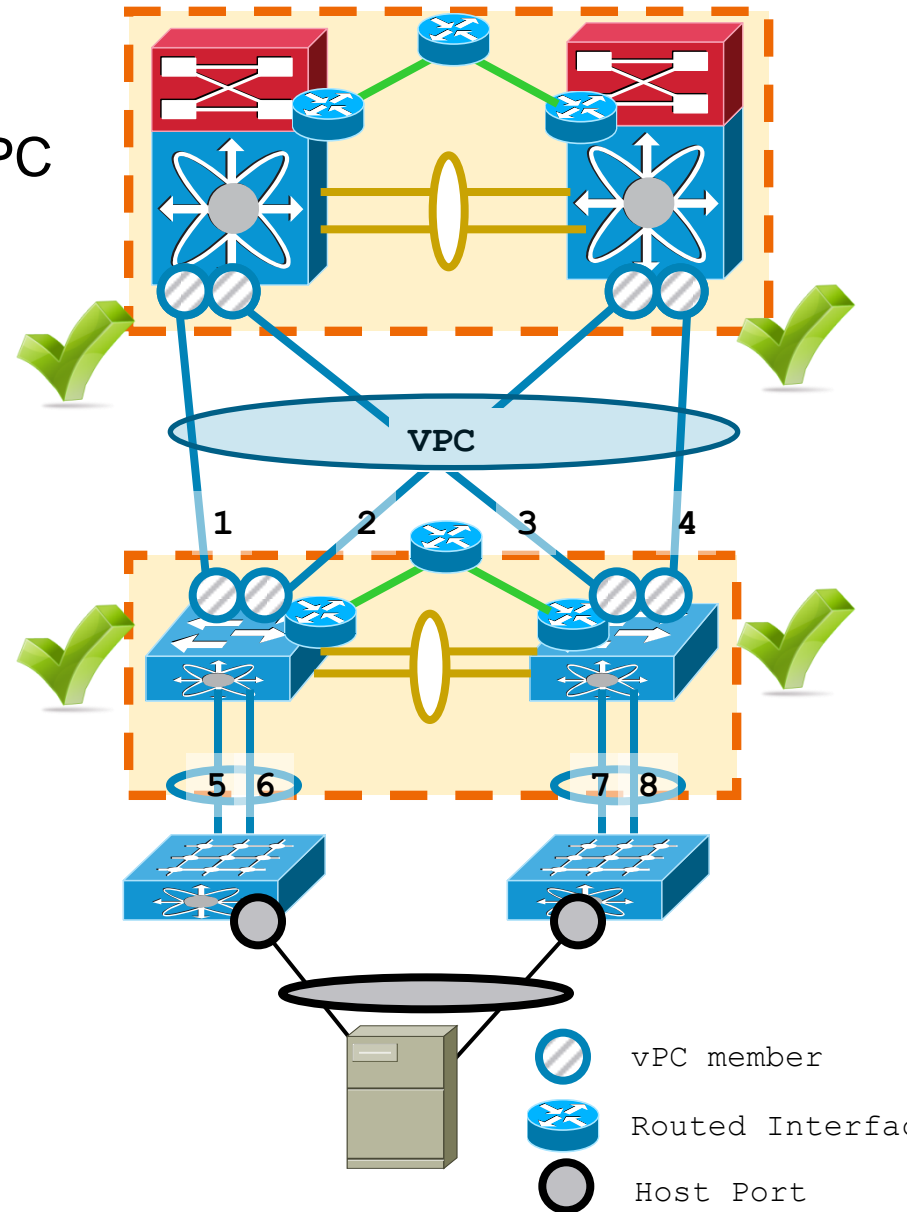


Building a vPC Domain

Configuration Steps

Following steps are needed to build a vPC
(Order **does Matter!**)

- Define domains*
- Establish Peer Keepalive connectivity
- Create a Peer link
- Reuse port-channels and Create vPCs
- *Make Sure Configurations are Consistent*



Configuring vPC

- **Enable the vPC Feature** (Modular NX-OS):

```
(config)# feature vpc
```

- **Example: Configuring vPC domain, starting with the peer-keepalive link and the peer-link on both peers:**

```
(config)# vpc domain 1
```

```
(config-vpc-domain)# peer-keepalive destination x.x.x.x source y.y.y.y vrf management
```

```
(config)# int port-channel 10
```

```
(config-int)# vpc peer-link
```

- **Example: Move any port-channels into appropriate vPC groups:**

```
(config)# int port-channel 20
```

```
(config-int)# vpc 20
```


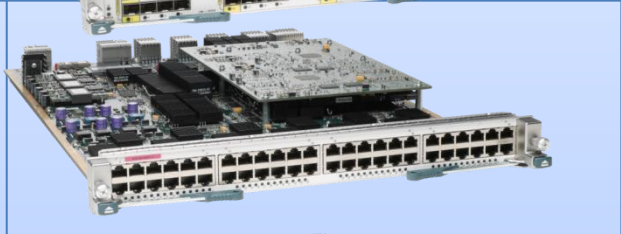
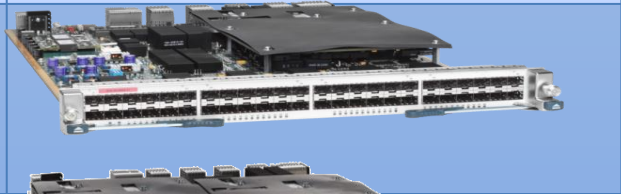


Notes:

- Make sure to leverage LACP
- *domain-id needs to differ* between the N7k vPC and the N5k vPC
- Spanning-Tree root is defined on one of the 2 N7ks
- N5k priorities are unmodified

vPC supported hardware (NEXUS 7000)

vPC support on New modules

- vPC is supported on all existing and new Modules

I/O Module	Picture	vPC Peer-link (10 GE Only)	VPC Interfaces
N7K-M132XP-12		✓	✓
N7K-M148GT-11		✗	✓
N7K-M148GS-11 N7K-M148GS-11L		✗	✓
N7K-M108X2-12L		✓	✓
N7K-F132XP-15		✓	✓

VPC (F1 & M1 LC)

- In mixed chassis, either M1/M1-XL or F1 ports can function as VPC peer link ports
 - Must use ports from same module type on each side of VPC peer link (all M1 or all F1 ports on each side of the VPC peer-link)
 - Recommended to use same module type on both ends of VPC peer link
- If F1 ports form VPC peer link, VPCs with M1/M1-XL ports allowed only if VPC peer link runs in CE (Classical Eth) mode
- If VPC peer link runs in FabricPath mode (to support VPC+ for FabricPath), all VPC ports must also be on F1 modules
- Mixing M1/M1-XL and F1 interfaces in a single port-channel not allowed due to different capabilities
- F1 modules support up to 16 active member ports in a single port-channel
- M1/M1-XL modules support 8 active member ports
- In mixed chassis with an M1/M1-XL channel of 8 ports and an F1 channel of 16 ports:
 - Traffic forwarded from an M1/M1-XL module to the F1 channel will map to no more than 8 of the member ports
 - Traffic forwarded from an F1 module to the M1/M1-XL channel can map to any of the 8 member ports

Building a vPC Domain

VDC Interaction

- vPC works seamlessly in any VDC based environment.
- One vPC domain per VDC is supported, up to the maximum number of VDCs supported in the system.
- It is still necessary to have a separate vPC peer-link and vPC Peer-Keepalive Link infrastructure for each VDC deployed.

Can vPC run between VDCs on the same switch?

- This scenario should technically work, but it is NOT officially supported and has not been extensively tested by our QA team.
- Could be useful for Demo or hands on, but It is **NOT recommended** for production environments. Will consolidate redundant points on the same box with VDCs (e.g. whole aggregation layer on a box) and introduce a single point of failure.
- ISSU will NOT work in this configuration, because the vPC devices can NOT be independently upgraded.

POLLING QUESTION - 1

- Can VPC peer-link can be configured with different Modules like F1 - M1?

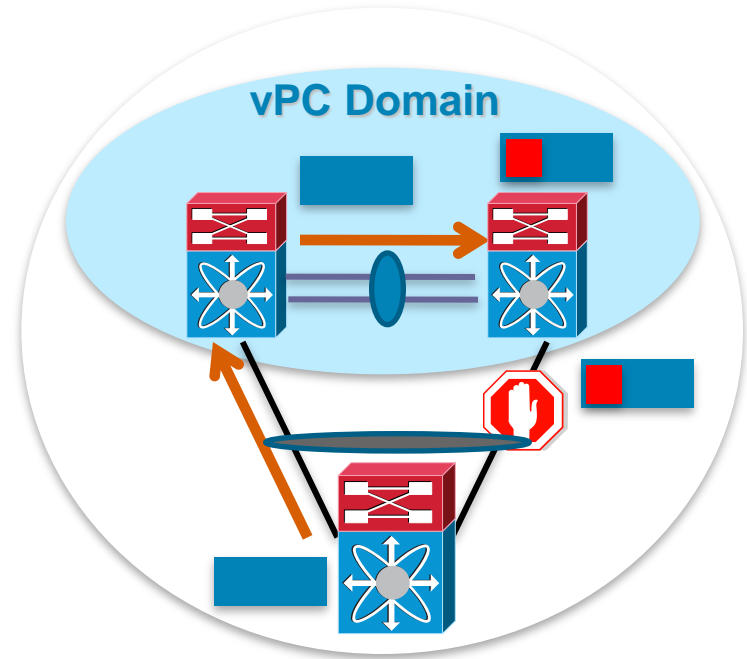
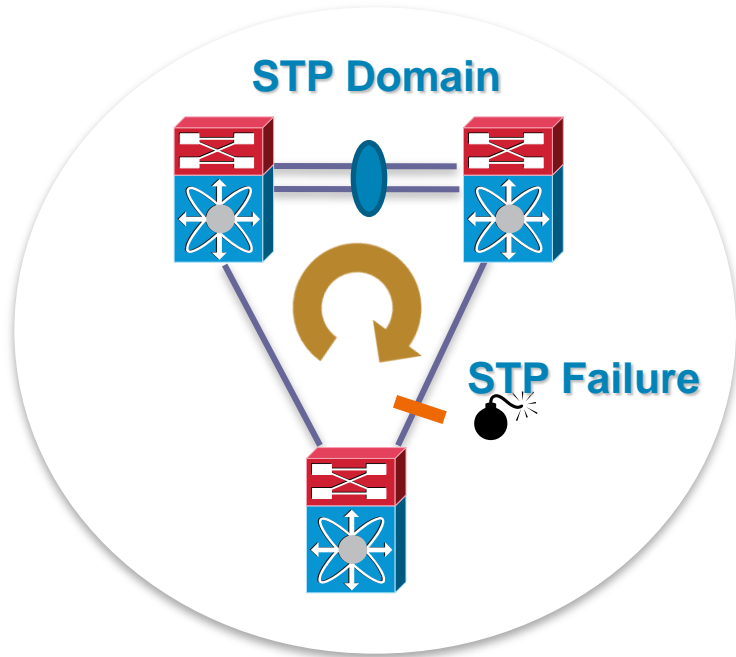
1) Yes, it can be configured with F1 and M1

2) No, we should have same compatibility features and we should use only one type of card

3) I don't know but i can try creating what's the harm in trying

AGENDA – Loop Prevention VPC

Data-Plane Loop Avoidance with vPC (1 of 2)



Data-Plane vs. Control-Plane Loop control

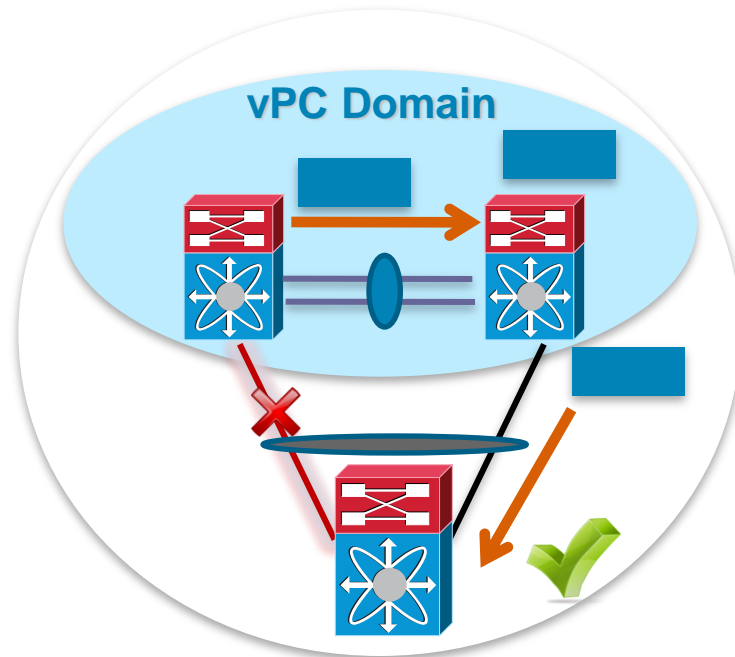
vPC peers can forward all traffic locally

Peer-link does not typically forward data packets (control plane extension)

Traffic on the Peer-link is marked and not allowed to egress on a vPC

Data-Plane Loop Avoidance with vPC (2 of 2)

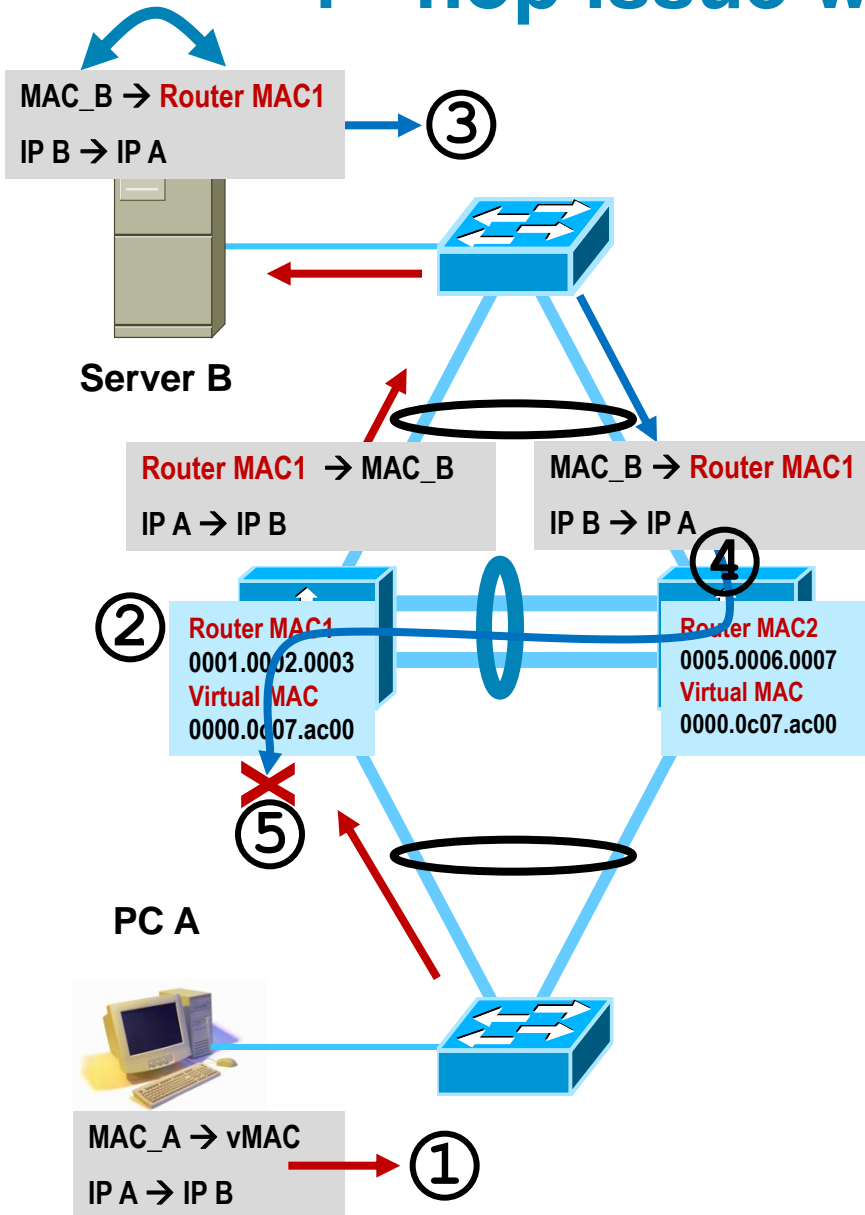
- Exception for single-sided vPC failures
- Peer-link used as Backup path for optimal resiliency



VSL bit

- The VSL bit is set in the DBUS header internal to the Nexus 7k.
- It is not something that is set in the ethernet packet that can be identified. The VSL bit is set on the port ASIC for the port used for the vPC peer link, so if you have Nexus A and Nexus B configured for vPC and a packet leaves Nexus A towards Nexus B, Nexus B will set the VSL bit on the ingress port ASIC. This is not something that would traverse the peer link.
- This mechanism is used for loop prevention within the chassis. The idea being that if the port came in the peer link from the vPC peer, the system makes the assumption that the vPC peer would have forwarded this packet out the vPC-enabled port-channels towards the end device, so the egress vpc interface's port-asic will filter the packet on egress.
- *show sys internal eltm info interface port-channel x*

1st hop issue with some devices



- ① PC A sends a packet to Server B
- ② Left VPC switch will receive the packet and forward it to Server B, note Source MAC of outgoing packet will be that of Router1
- ③ Server B responding to PC A will populate destination MAC from source MAC of received frame (this is wrong, it should use ARP)
- ④ If frame from B→A will be load-balanced to right switch the MAC address of Router1 will point to Peer-Link and this is where the frame will be sent
- ⑤ Left switch will receive the frame from Peer-Link and drop it

Why? Frames received from Peer-Link are never sent out of VPC except those without operational ports on ingress switch (egress port ASICs will drop the frame)

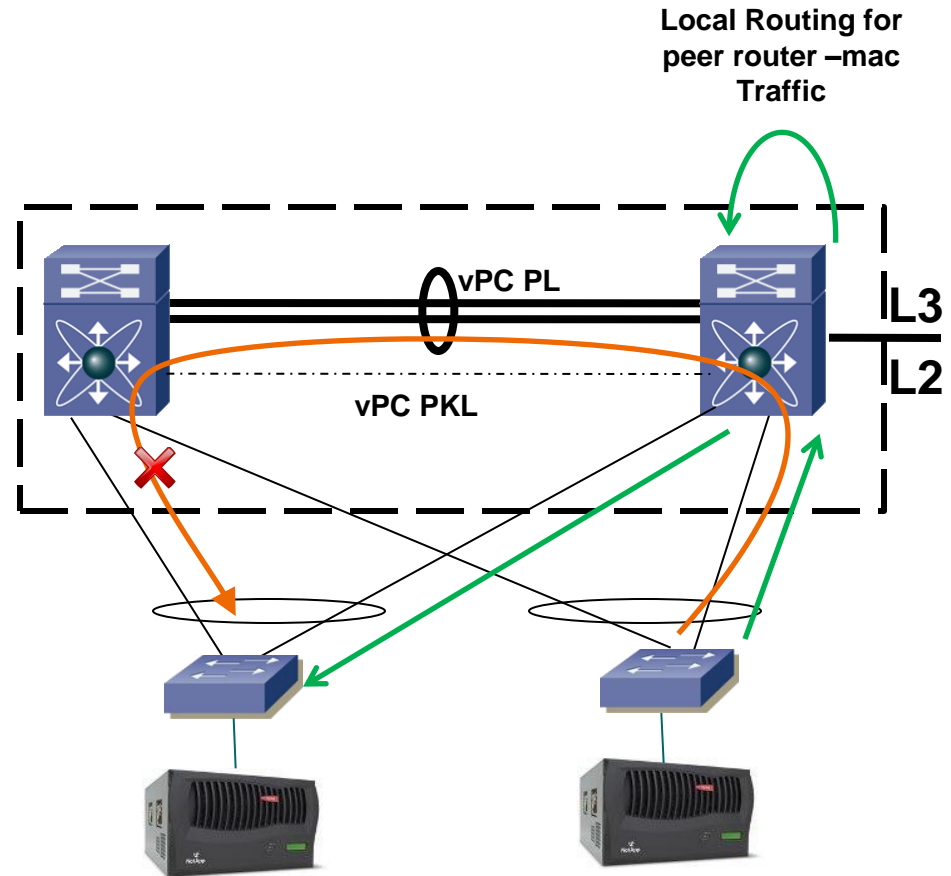
vPC Peer-Gateway for NAS interoperability

Problem/Impact:

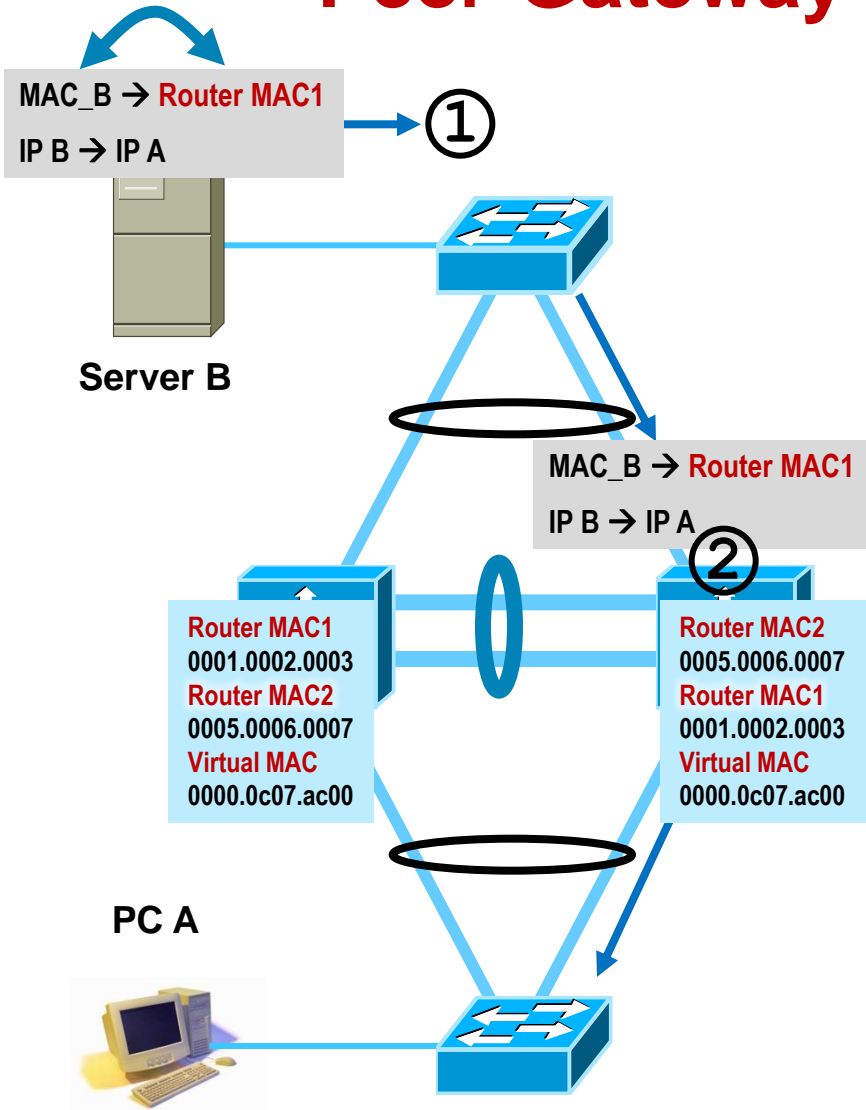
- Lack of interoperability with non RFC compliant features of some NAS devices (i.e. NETAPP Fast-Path or EMC IP-Reflect)
- NAS device may reply to traffic using the MAC address of the sender device rather than the HSRP gateway.
- Packet reaching vPC for the non local Router MAC address are sent across the peer-link and can be dropped if the final destination is behind another vPC.

vPC Peer-Gateway Solution:

- Allows a vPC switch to act as the active gateway for packets addressed to the peer router MAC (CLI command added in the vPC global config)



Peer-Gateway : the workaround

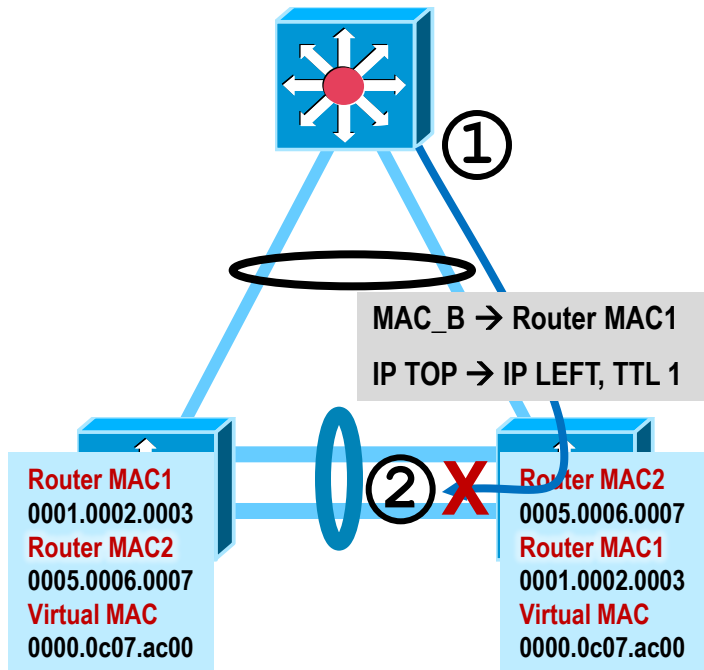


With peer-gateway both peers will install router MACs of each other in L2 table which will allow them to L3 forward traffic destined to either Router MAC

① Server B responding to PC A will populate destination MAC from source MAC of received frame (this is wrong, it should use ARP)

② Right switch will forward packet towards destination

Peer-Gateway : the implications



- ① Top device attempts to establish OSPF adjacency with the **left switch**
- ② If **peer-gateway** is enabled in VPC domain and OSPF unicast packet will be load-balanced to the **right switch**, this packet will be dropped

Why? Right switch will try to L3-switch the unicast packet (because RouterMAC1 is marked as gateway MAC and destination IP is not local)
As packet has TTL==1 it will be dropped

Same applies to any other protocol that uses unicast packets with TTL==1 entering right switch but destined to left switch (or vice versa)

Routing protocol peering with devices attached to VPC domain via SVI interface is not supported
Routed interface should be used in this case

vPC Enhancements

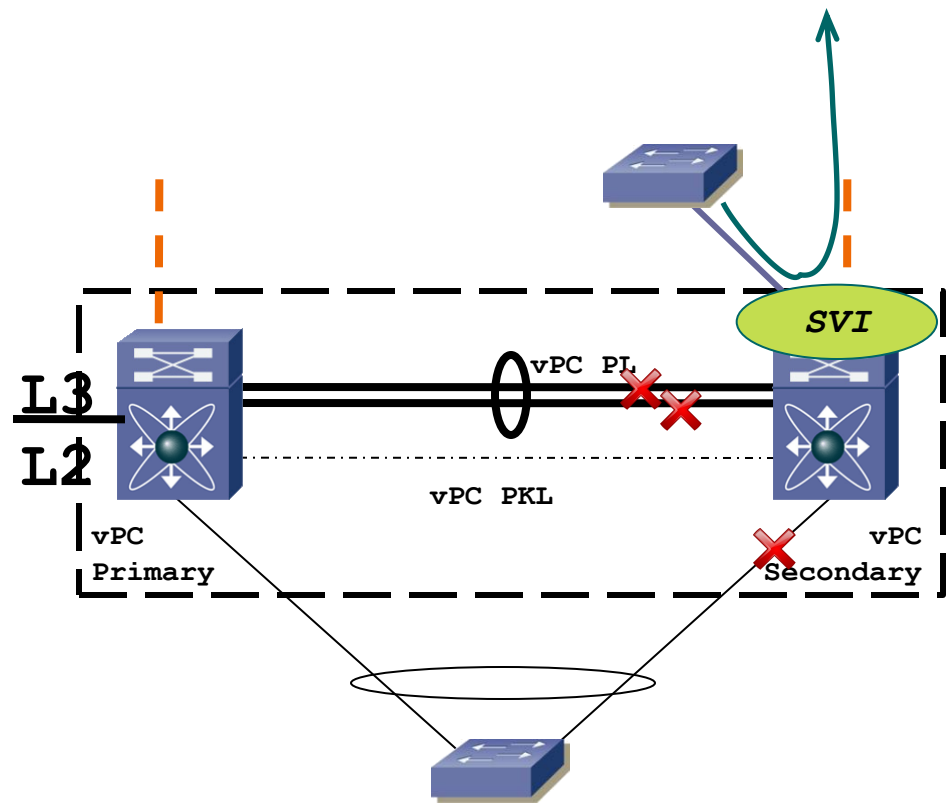
vPC Exclude Interface-VLAN

Problem/Impact:

- When a dual active condition is detected in SVIs and vPCs on the secondary vPC peer are suspended
- Only the primary vPC peer continues data plane and control plane functionalities.

vPC exclude interface-VLAN solution:

- The vPC exclude interface-VLAN feature ensures that a configurable list of SVIs are not suspended on the secondary vPC peer
- Consequently Layer 3 connectivity is maintained even in a dual active condition for a restricted selection of interfaces.



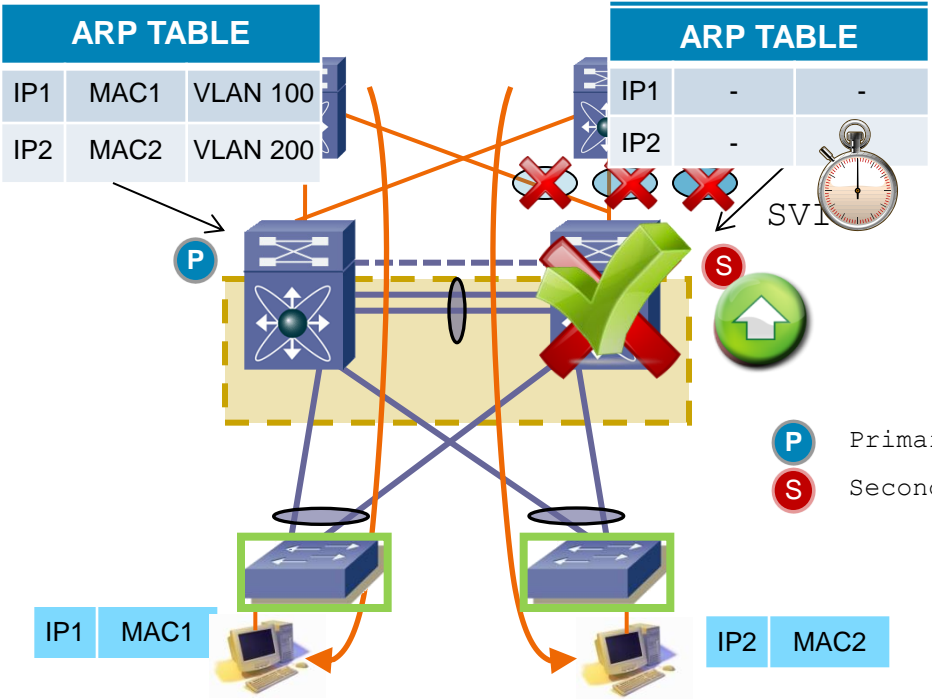
```
N7K (config-vpc-domain)# dual-active exclude interface-vlan ?  
  <1-3967,4048-4093> Set allowed interface vlans
```

Nexus 7000 vPC Update

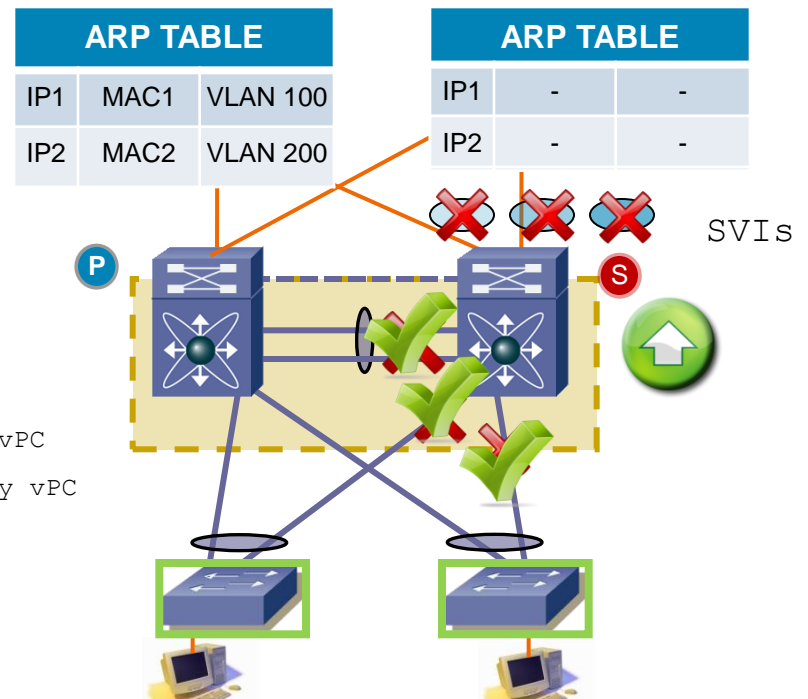
ARP Synchronization



- Unicast traffic to converge impacted when:
 - A Switch in the vPC domain goes offline and comes back
 - Peer-link port-channel fail and recovers, triggering an SVI flap
- Convergence issues due to the delay involved in Address Resolution Protocol (ARP) table restoration



Device Failure/Restoration



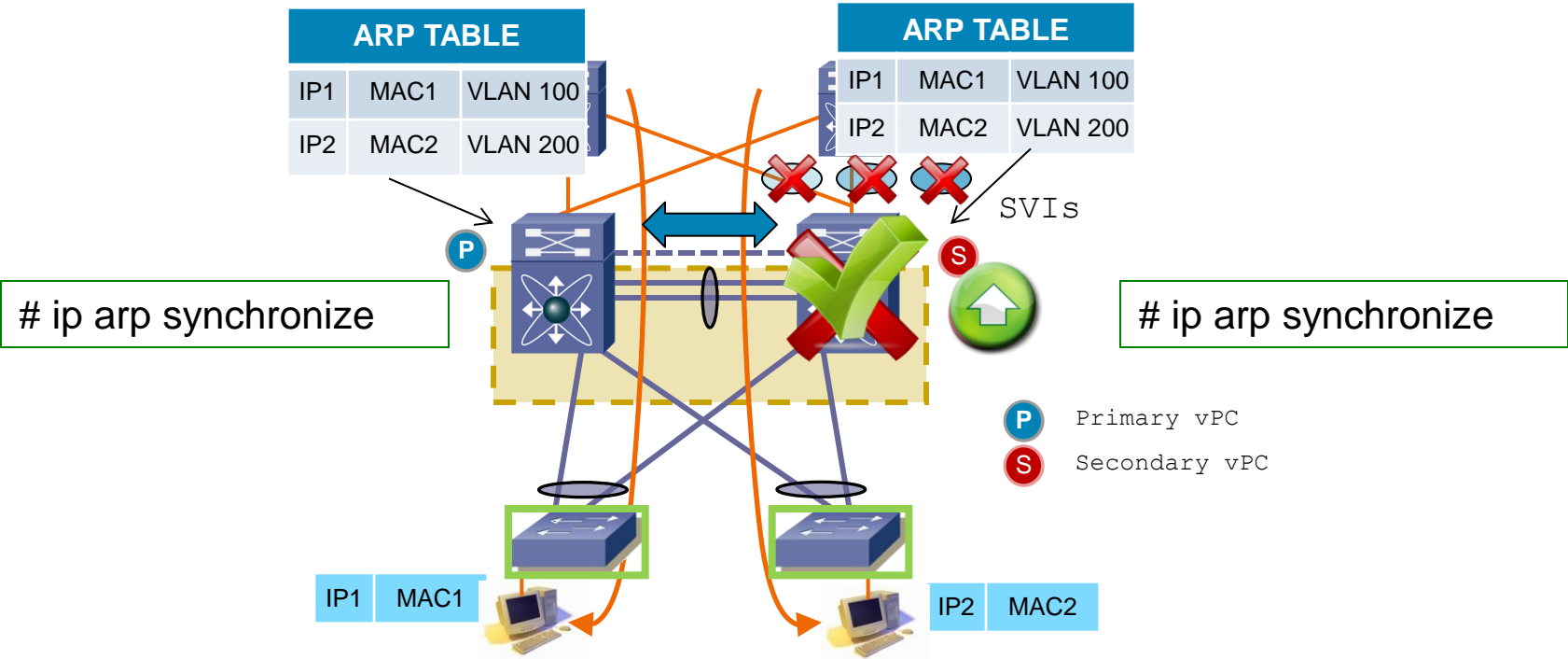
Peer-link Failure/Restoration

Nexus 7000 vPC Update

ARP Synchronization



- After the peer-link comes up perform an ARP bulk sync over CFSoE to the peer switch
- Improve Convergence for Layer 3 flows

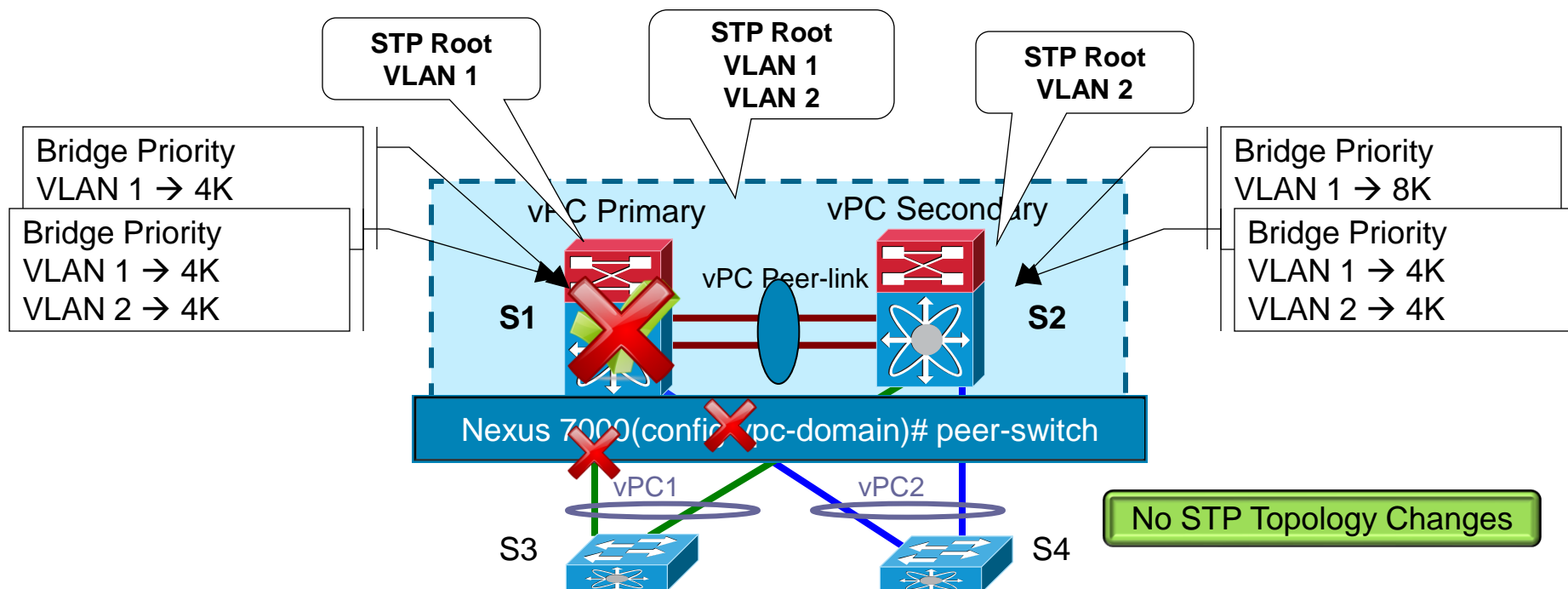


ARP Synchronization Process

vPC peer-switch

Unified STP Root with vPC

http://bock-bock.cisco.com/wiki_file/N7K:tech_resources:vpc/vPC_peer-switch_-_technical_overview.ppt



- vPC peer-switch feature allows a pair of vPC peer devices to appear as a single STP Root in the L2 topology (same bridge-id)
- Simplifies STP configuration by configuring both vPC with the same STP priority
- Eliminates recommendation to pin STP Root to the vPC primary switch
- Improves convergence during vPC primary switch failure/recovery avoiding Rapid-STP Sync
- Supports a hybrid topology of vPC and non-vPC connections by using the spanning-tree pseudo-information

POLLING QUESTION - 2

What is the feature that will make both the N7k's look like a single switch with STP perspective?

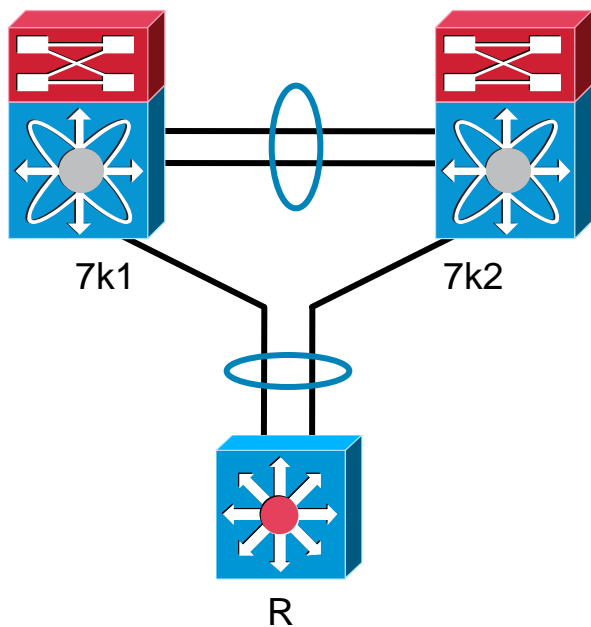
- 1)VPC peer-gateway
- 2)Feature VPC
- 3)Peer-switch
- 4)Auto-recovery
- 5)Arp synchronization

AGENDA – Supported/Unsupported Topologies

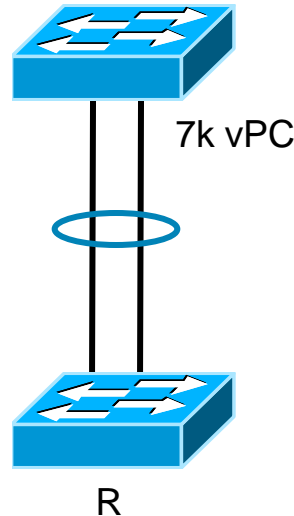
Layer 3 and vPC Interactions

Router Interconnection: different angles

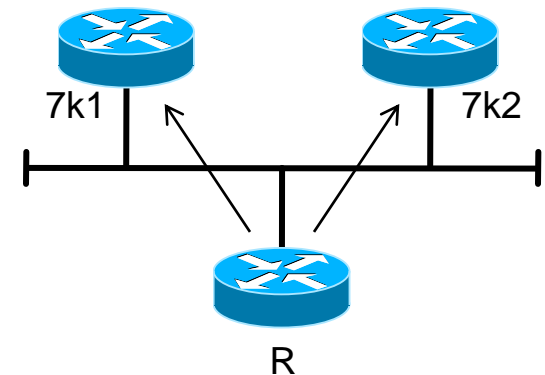
vPC view



Layer 2 topology



Layer 3 topology



R could be any router,
L3 switch or VSS
building a port-channel

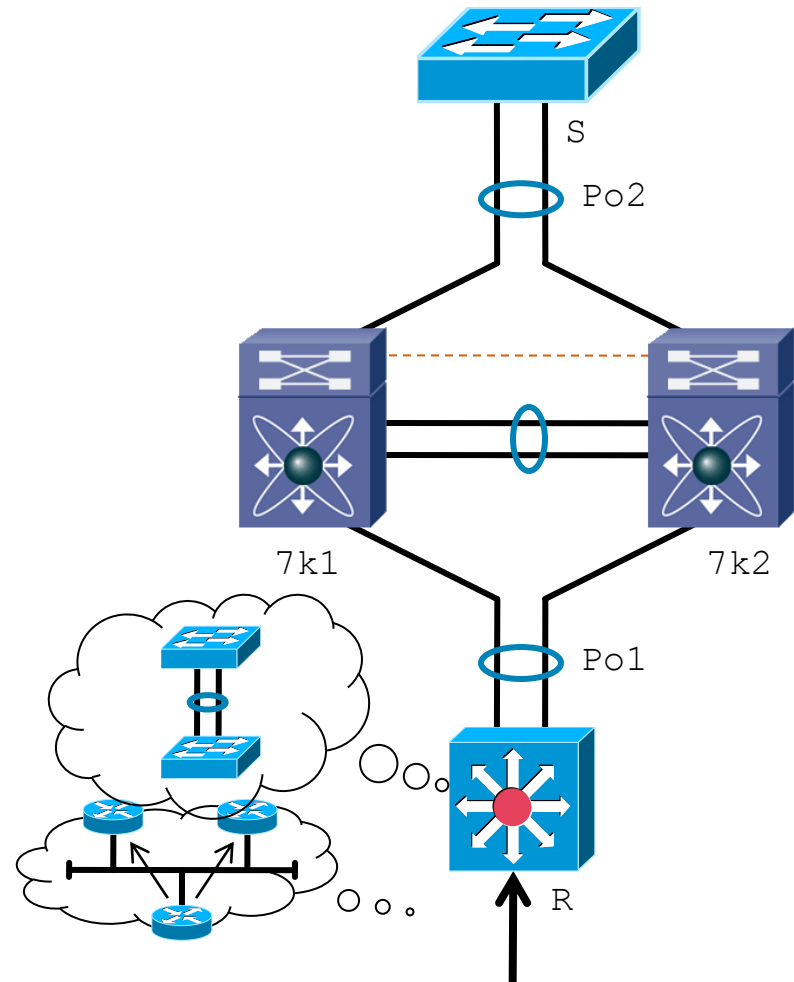
Port-channel looks like
a single L2 pipe.
Hashing will decide
which link to chose

Layer 3 will use ECMP
for northbound traffic

Layer 3 and vPC Interactions

Router Interconnection: Forwarding sequence

- 1) Packet arrives at R
- 2) R does lookup in routing table and sees 2 equal paths going north (to 7k1 & 7k2)
- 3) Assume it chooses 7k1 (ECMP decision)
- 4) R now has rewrite information to which router it needs to go (router MAC 7k1 or 7k2)
- 5) L2 lookup happens and outgoing interface is port-channel 1
- 6) Hashing determines which port-channel member is chosen (say to 7k2)
- 7) Packet is sent to 7k2
- 8) 7k2 sees that it needs to send it over the peer-link to 7k1 based on MAC address



Layer 3 and vPC Interactions

Router Interconnection: Forwarding sequence (continued)

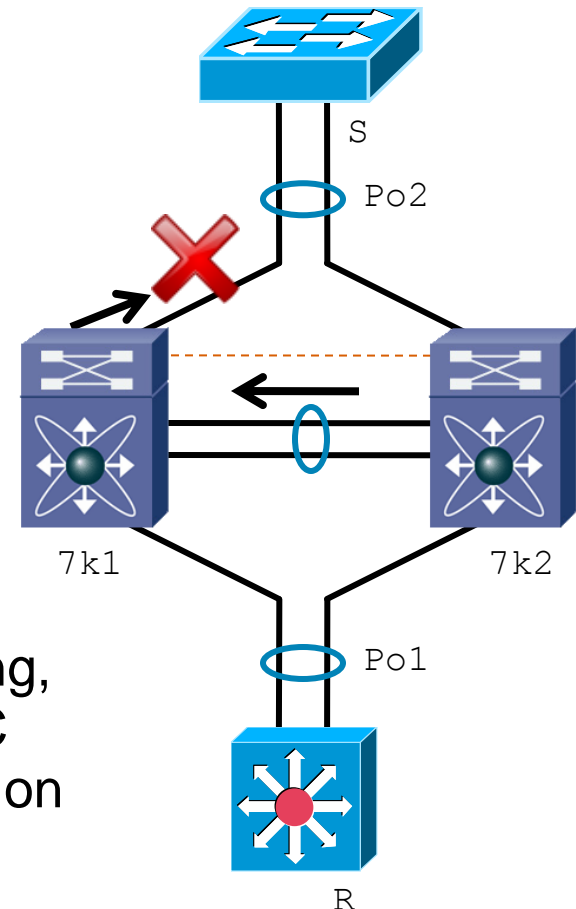
- 9) 7k1 performs lookup and sees that it needs to send to S
- 10) 7k1 performs check if the frame came over peer link & is going out on a vPC.
- 11) Frame will ONLY be forwarded if:

Outgoing interface is NOT a vPC or

Outgoing vPC doesn't have active

Note: interface on other vPC peer (in our example 7k2)

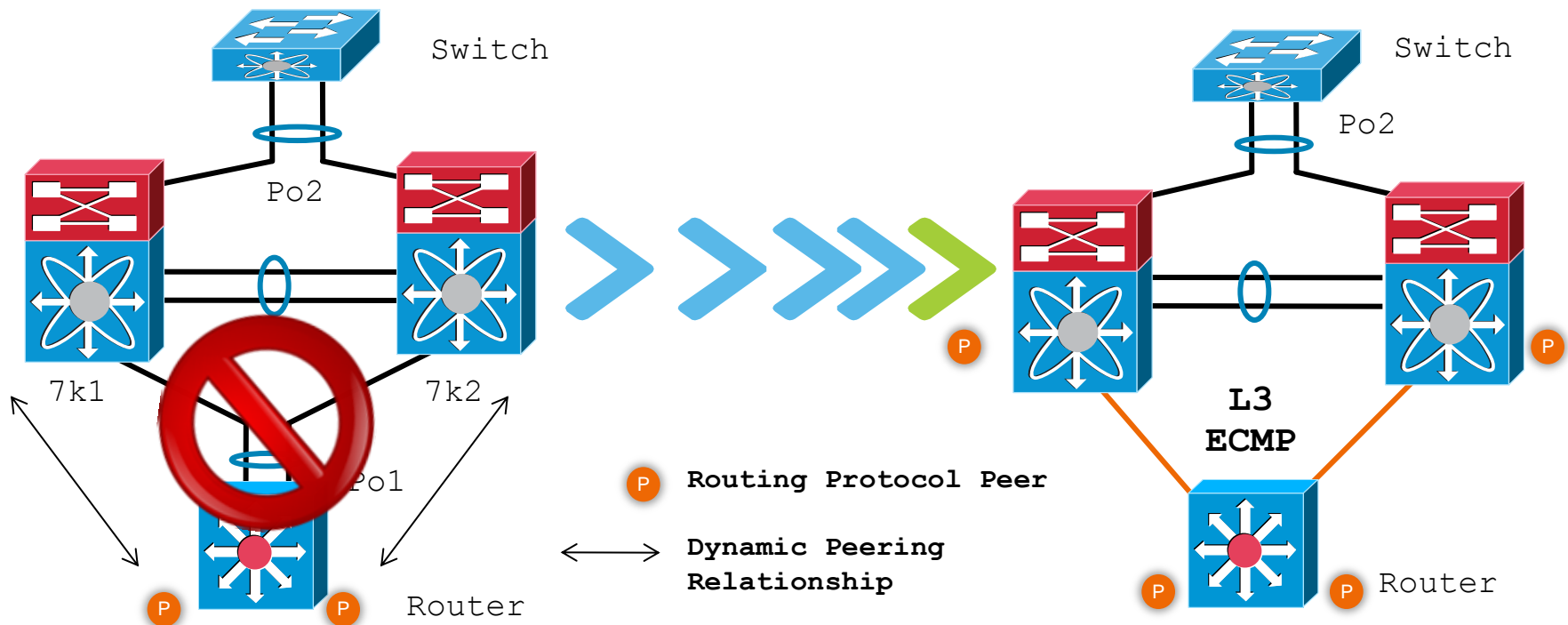
- Use of Peer-Gateway allows data-path forwarding, routing/forwarding traffic for the peer-router MAC locally, but **does NOT Enable** Dynamic Routing on vPC VLANs
- Even with Peer-Gateway Routing protocols (e.g. OSPF) are **still broken** due to TTL expiry when traversing in transit the peer vPC Router device.



Layer 3 and vPC Designs

Layer 3 and vPC Design

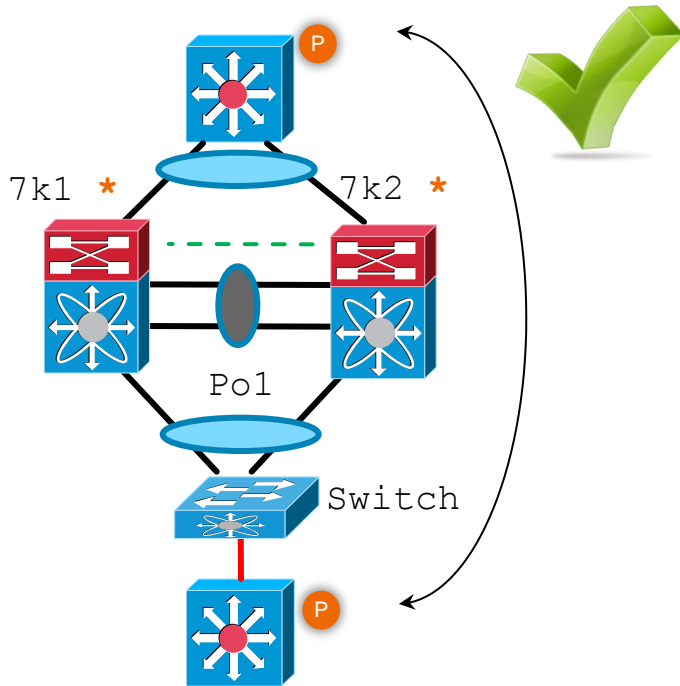
- Use L3 links to hook up routers and peer with a vPC domain
- Don't use L2 port channel to attach routers to a vPC domain unless you statically route to HSRP address
- If both, routed and bridged traffic is required, use individual L3 links for routed traffic and L2 port-channel for bridged traffic



Layer 3 and vPC Designs

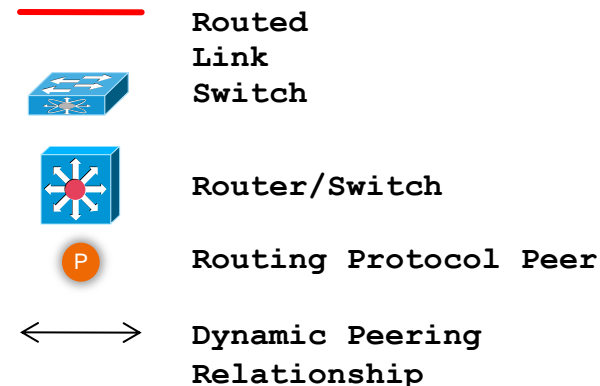
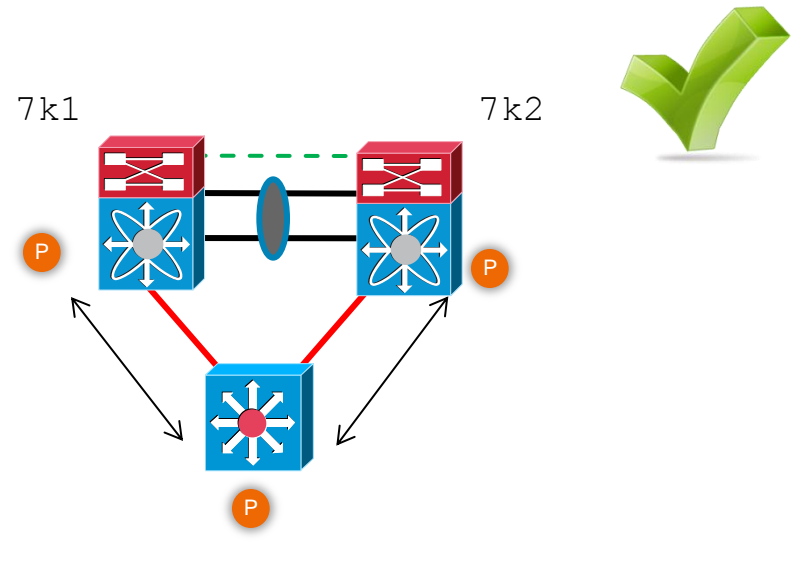
Layer 3 and vPC Interactions: Supported Designs

1. Peering between Routers



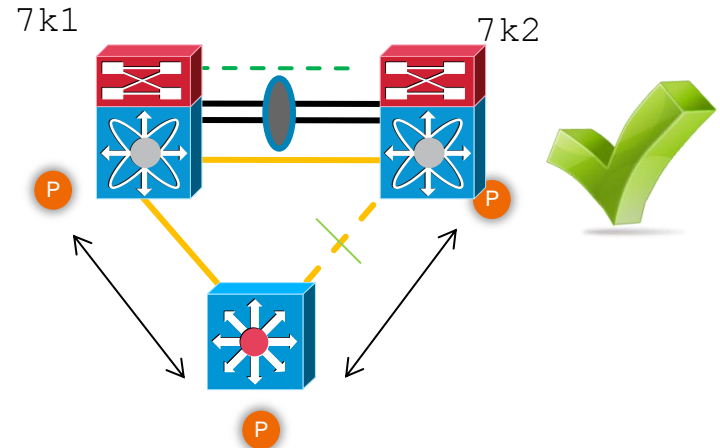
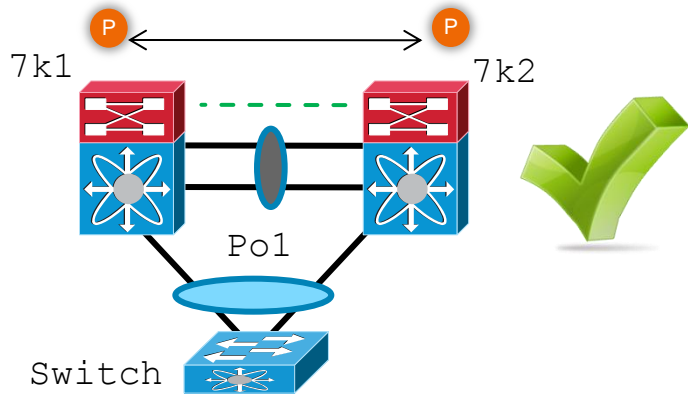
* Nexus 7000 configured for L2 Transport only

2. Peering with an external Router on Routed ports inter-connection



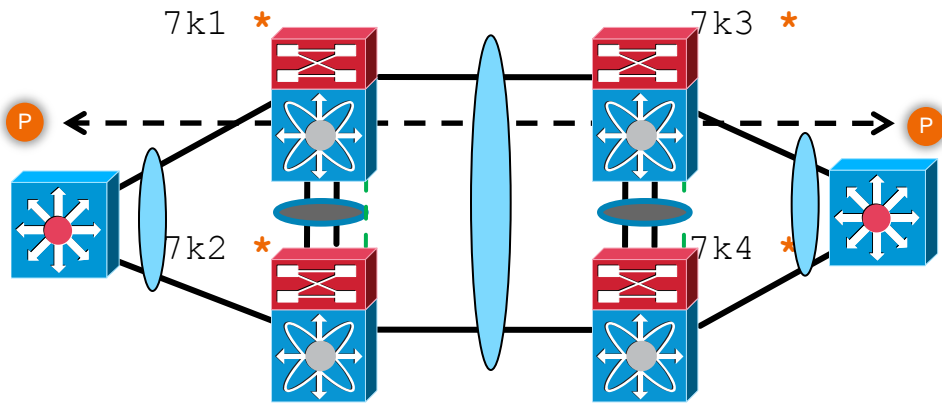
Layer 3 and vPC Designs




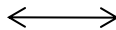
Layer 3 and vPC Interactions: Supported Designs



1. Peering between vPC Device

2. Peering over an STP inter-connection
NOT using a vPC VLAN (Orange VLANs/Links)



-  Switch
-  Router/Switch
-  Routing Protocol Peer
-  Dynamic Peering Relationship

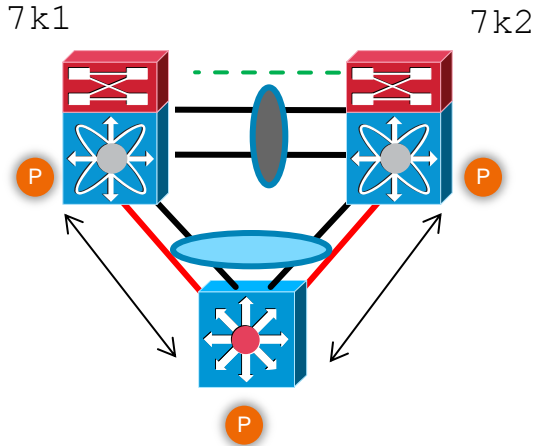
* Nexus 7000 configured for L2 Transport only

3. Peering between 2 routers with vPC devices as transit

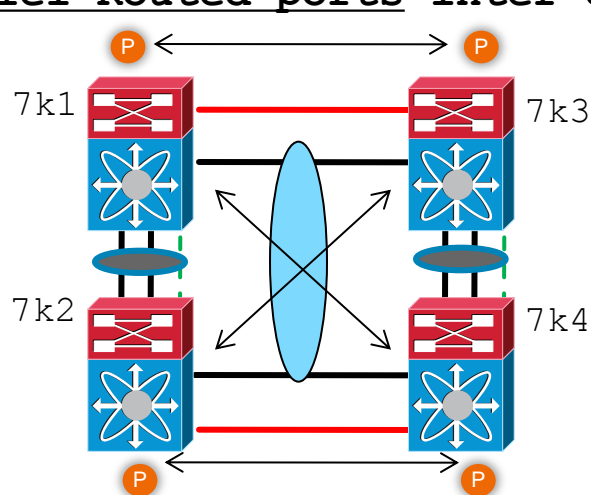
Layer 3 and vPC Designs





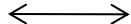
Layer 3 and vPC Interactions: Supported Designs

1. Peering with an external Router on parallel Routed ports inter-connection



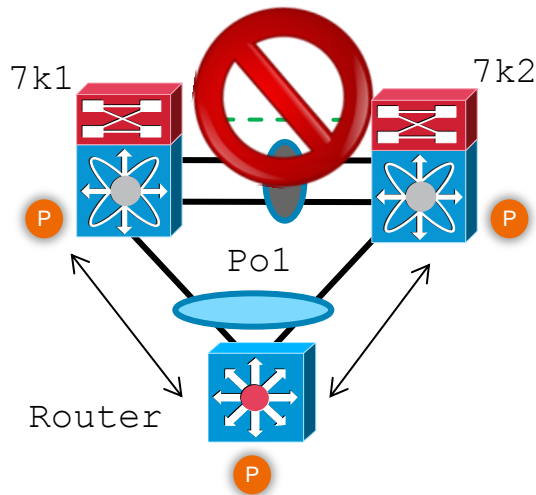
2. Peering over a vPC inter-connection (DCI case) on parallel Routed ports inter-connection



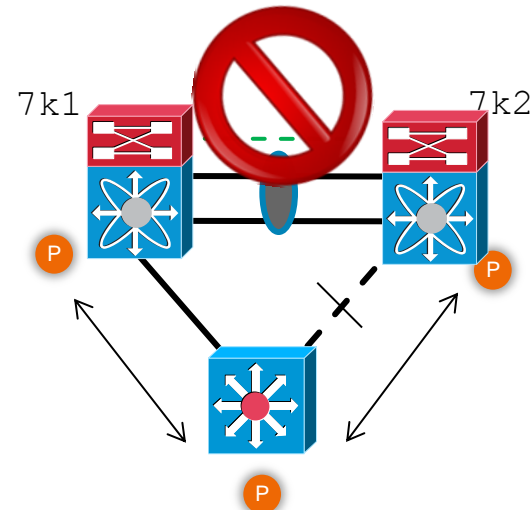
-  Routed Link
-  Switch
-  Router/Switch
-  Routing Protocol Peer
-  Dynamic Peering Relationship

Layer 3 and vPC Designs

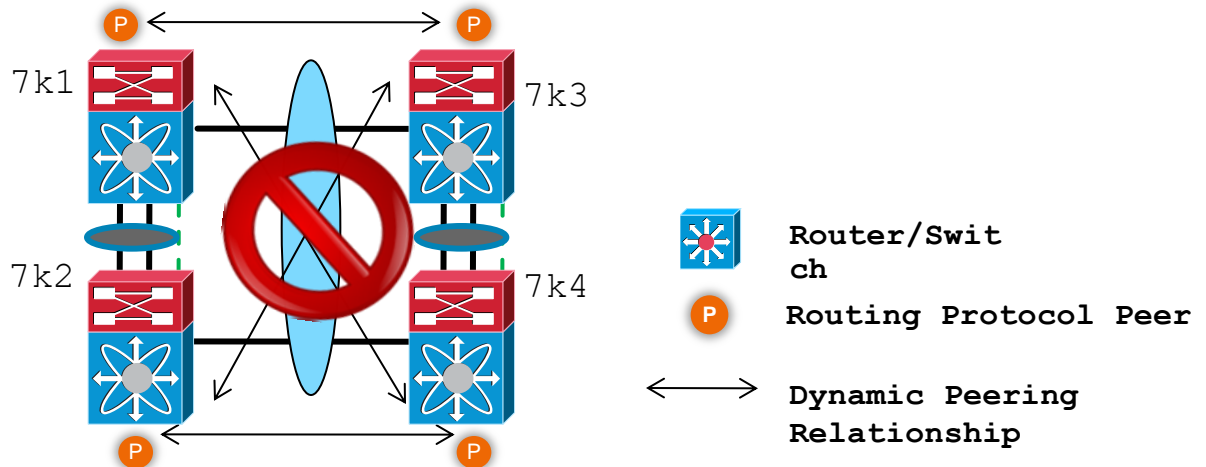
Layer 3 and vPC Interactions: Unsupported Designs



1. Peering over a vPC inter-connection



2. Peering over an STP inter-connection using a vPC VLAN



3. Peering over a vPC inter-connection (DCI case)

Failover Overview/ Type1 and Type 2

Failure Type	Primary Action	Secondary Action
Keep-alive failure	No action	No action
Peer-link failure	No action	Suspend vPCs and vPC SVIs
Peer-link and keep-alive failure	No action, assumes secondary switch went offline	Becomes operational primary and continues forwarding on all vPCs
Keep-alive fails, then after some time peer-link fails	No action, remains vPC primary	Becomes vPC primary (dual active scenario)
Peer-link fails, then after some time keep-alive fails	No action, remains vPC primary	Peer-link failure suspends vPCs. Links remain suspended when keep-alives are lost.

- If vPCs are suspended, we must wait until the peer adjacency is reestablished before bringing back up the vPCs. I.e., both the keep-alive and peer-link must come back online before any vPCs will be brought back up on the operational standby switch

POLLING QUESTION - 3

ABCCorp have configured STP mode as RSTP on N7k primary and MST on N7k secondary, what will be the problem?

- 1) This will create type-1 inconsistencies
- 2) It will shutdown all its vpc-legs in the secondary and wait until STP configs are matched
- 3) It will create just type-2 configs mismatch, traffic will be still forwarded normally.
- 4) One has to check show vpc consistency parameters global command to verify what is happening.
- 5) I don't like VPC, so I don't care.

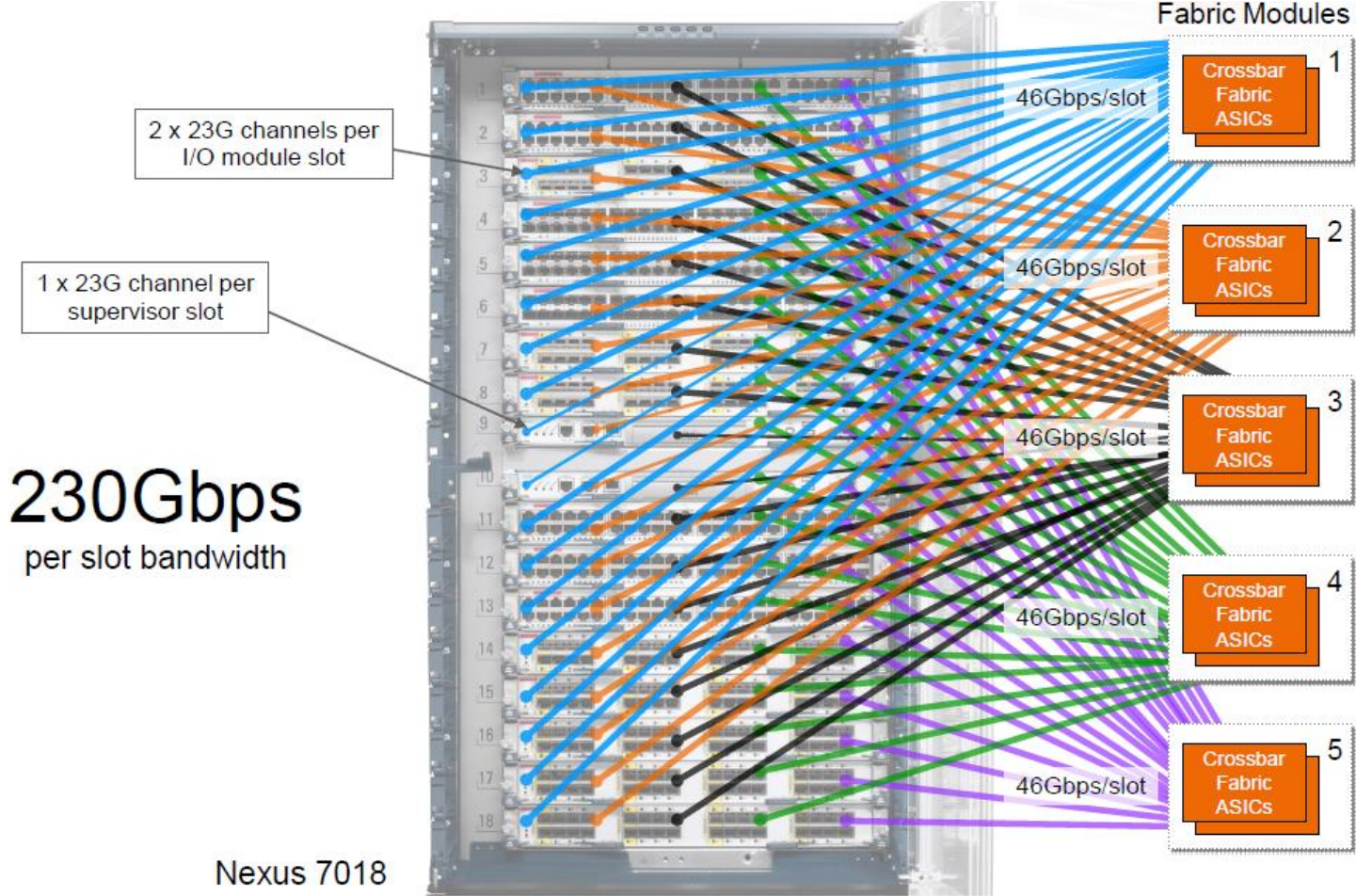
ISSU Failures/Upgrade

- Release notes (EPLD – Always Intrusive)
- Compatibility Check
- Known Defects
- Any Release before 5.x to 6.2 and above – Hard reboot preferred.
- Show install all impact kickstart <> System <>
- Avoid STP loop and Routing loops when ISSU
- Show tech ISSU -> For Contacting TAC
- Show Cores vdc-all

Module Failures

- GOLD Tests
- Test results: (. = Pass, F = Fail, I = Incomplete, U = Untested, A = Abort, E = Error disabled)
- What to Look for?
- Show diagnostics result module x
- Show tech module x
- Show module internal exceptionlog module x
- Show log log
- Show log nvram
- Show cores vdc-all

Fabric Module Capacity (Ex: FAB-1)



FAB-2 is 550Gbps per slot bandwidth

Fabric Troubleshooting

- Show module internal exceptionlog module x
- System Errorcode : 0x40240012 xbar sync failed during module bringup (DevErr is LinkNum)
- Show system error-id 0x40240012
- Show system internal xbar event-history errors
- Show system internal xbar sw(Internal Mapping)
- show hardware internal errors
- show hardware internal xbar-driver event-history errors
- Show tech detail

MTS Buffers/ Memory Leaks

- What is a MTS?
- Message Transactional services – Used for inter-process communications.
- Three fields characterize an MTS message:

OpCode, Message-ID and Payload

Useful Commands:

Show system internal mts buffers summary

Show system internal mts sup sap <number> Desc

MTS Buffers/Memory Leaks

- %KERN-2-SYSTEM_MSG: mts_is_q_space_available_old(): NO SPACE - node=4, sap=221, uuid=410, pid=7451, sap_opt = 0x1, hdr_opt = 0x0, rq=4205(430160), lq=0(0), pq=1(108), nq=0(0), sq=0(0), fast: rq=0, lq=0, pq=0, nq=0,

What do we understand from the above message?

NODE

SAP

n7k-1# show system internal mts sup sap 221 desc
L2fm SAP

What is the L2fm? (Layer 2 Feature Manager)

Layer 2 Loop??

High CPU

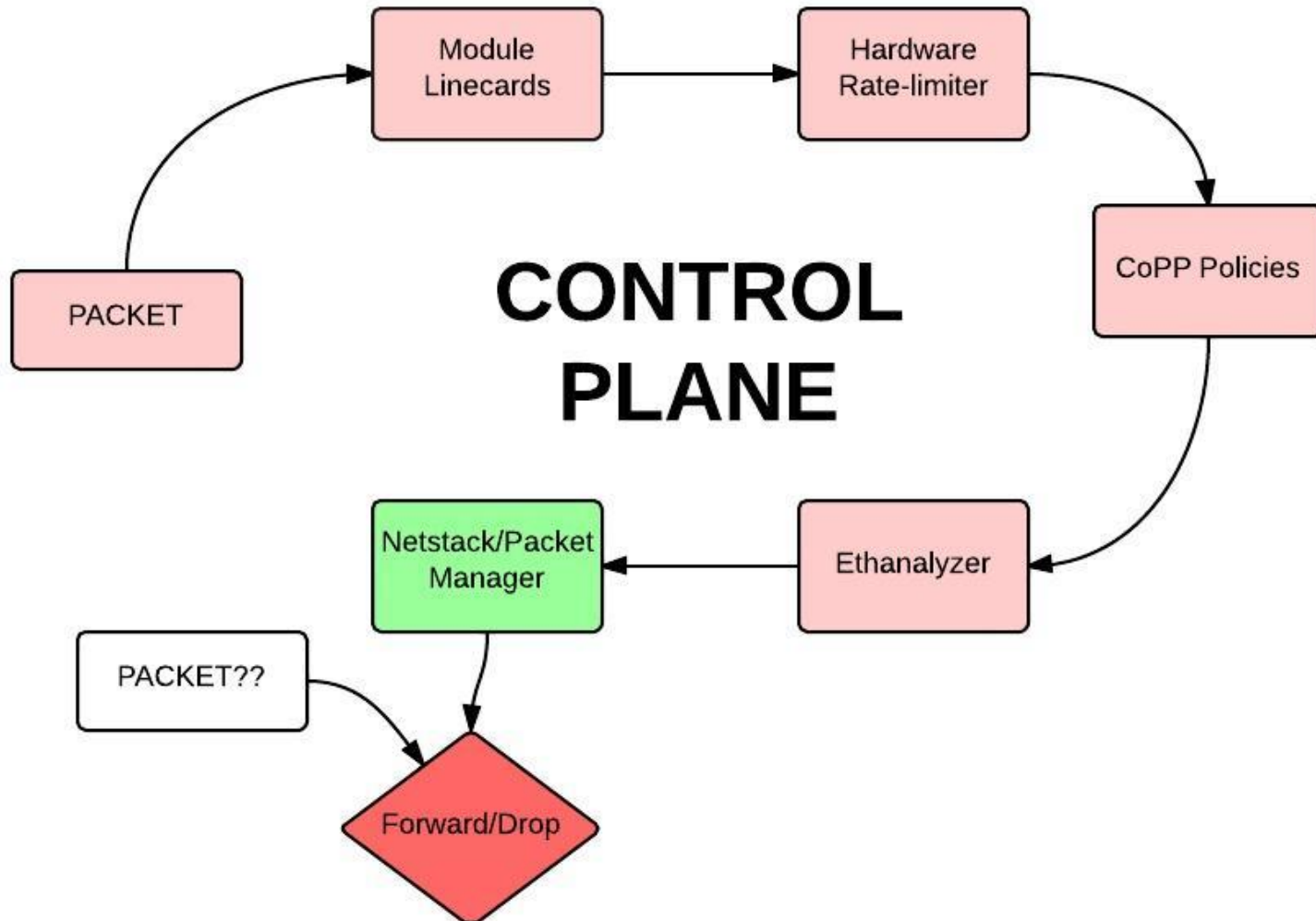
- What is Control-Plane and Data Plane?

BPDU,EIGRP,OSPF, Ping from and to the box is control-plane

Anything that is flow-through between the End hosts.

- CoPP/ Hardware Rate-limiter
- Show process cpu his -> Will not show accurate stats

High CPU



High CPU

- Show Process CPU
- Show system resources
- Show policy-map interface control-plane | | violated|drops
- Show hardware rate-limiter module x
- Show hardware internal cpu-mac inband stats
- Show hardware internal cpu-mac inband events
- [HIGH CPU](#)

Polling Question - 4

- We have high CPU utilization in the N7k and Memory problems what can be done to prevent such issues?
 - 1) Identify the src that is causing CPU and fix it.
 - 2) Use MTS commands to verify what SAP is causing it.
 - 3) Run NVT tested releases to move from known caveats
 - 4) Does N7k has a CPU?
 - 5) I don't know

LAYER2 CHECK

- Layer 2 : Ethpm/L2fm

Ethernet Port Manager -> Physical Connection

L2fm - > Layer 2 Forwarding Manager

- Spanning-tree Forwarding?

Show spanning-tree vlan x

- VPC links are fine and working?

Show vpc bri

Show vpc consistency parameters global

LAYER2 CHECK

- Is the MAC learnt on Software and hardware Fine?

Show mac-address table address <>

Show hardware mac-address table address <>

- Logging level I2fm 5 -> Enable MAC MOVE
- CBL Forwarding
- Show tech's
- Show tech stp
- Show tech I2fm

2) Verify hardware state on ingress module

```
n7k-2# attach module 4
```

```
Attaching to module 4 ...
```

```
To exit type 'exit', to abort type '$.'
```

```
Expected received MAC from Host
```

```
module-4# show hardware mac address-table address 0000.0c07.ac65 vlan 2500 vdc 1
```

FE	Valid	PI	BD	MAC	Index	Stat	SW	Modi	Age	Tmr	GM	Sec	TR	NT	RM	RMA	Cap
						ic		fied	Byte	Sel		ure	AP	FY			TURE
0	1	1	20	0000.0c07.ac65	0x00409	1	0x000	0	131	1	1	0	0	0	0	0	0
0	0																

```
*** GM (gateway-MAC flag set indicated we should attempt to route this packet)
```

```
Next-hop MAC addresses:
```

```
module-4# show hardware mac address-table address 0000.0000.0156 vlan 55 vdc 1
```

FE	Valid	PI	BD	MAC	Index	Stat	SW	Modi	Age	Tmr	GM	Sec	TR	NT	RM	RMA	Cap
						ic		fied	Byte	Sel		ure	AP	FY			TURE
0	1	0	22	0000.0000.0156	0x009ed	0	0x003	0	159	1	0	0	0	0	0	0	0
0	0																

```
module-4# show hardware mac address-table address 0000.0000.0256 vlan 55 vdc 1
```

FE	Valid	PI	BD	MAC	Index	Stat	SW	Modi	Age	Tmr	GM	Sec	TR	NT	RM	RMA	Cap
						ic		fied	Byte	Sel		ure	AP	FY			TURE
0	1	0	22	0000.0000.0256	0x009ed	0	0x003	0	159	1	0	0	0	0	0	0	0
0	0																

```
Note Index for MAC is 0x9ed. This should map to E3/5. Verification steps in a few slides
```

1) Verify software state

```
n7k-2# show ip route 10.0.3.101 vrf Infrastructure
```

```
IP Route Table for VRF "Infrastructure"
```

```
'*' denotes best ucast next-hop
```

```
'**' denotes best mcast next-hop
```

```
'[x/y]' denotes [preference/metric]
```

```
'%<string>' in via output denotes VRF <string>
```

```
10.0.3.0/24, ubest/mbest: 2/0
```

```
  *via 10.55.0.101, Vlan55, [90/3072], 02:18:54, eigrp-100, internal
```

```
  *via 10.55.0.102, Vlan55, [90/3072], 02:19:13, eigrp-100, internal
```

```
n7k-2# show ip arp 10.55.0.101 vrf Infrastructure
```

```
Flags: * - Adjacencies learnt on non-active FHRP router
```

```
  + - Adjacencies synced via CFSOE
```

```
  # - Adjacencies Throttled for Glean
```

```
  D - Static Adjacencies attached to down interface
```

```
IP ARP Table
```

```
Total number of entries: 1
```

```
Address      Age      MAC Address  Interface
```

```
10.55.0.101  00:13:43  0000.0000.0156  Vlan55
```

```
n7k-2# show ip arp 10.55.0.102 vrf Infrastructure
```

```
Flags: * - Adjacencies learnt on non-active FHRP router
```

```
  + - Adjacencies synced via CFSOE
```

```
  # - Adjacencies Throttled for Glean
```

```
  D - Static Adjacencies attached to down interface
```

```
IP ARP Table
```

```
Total number of entries: 1
```

```
Address      Age      MAC Address  Interface
```

```
10.55.0.102  00:14:56  0000.0000.0256  Vlan55
```

LAYER3 CHECK

2) Verify hardware state on ingress module

```
n7k-2# show forwarding ipv4 route 10.0.3.101 vrf Infrastructure module 4
```

```
IPv4 routes for table Infrastructure/base
```

Prefix	Next-hop	Interface	Labels
10.0.3.0/24	10.55.0.101	Vlan55	
	10.55.0.102	Vlan55	

```
n7k-2# show system internal forwarding vrf Infrastructure ipv4 route 10.0.3.101 det module 4
```

```
RPF Flags legend:
```

```
S - Directly attached route (S_Star)
```

```
V - RPF valid
```

```
M - SMAC IP check enabled
```

```
G - SGT valid
```

```
E - RPF External table valid
```

```
10.0.3.0/24 , Vlan55 , No of paths: 2
```

```
Dev: 1 , Idx: 0x19007 , RPF Flags: V , DGT: 0 , VPN: 9
```

```
RPF_Intf_5: Vlan55 (0x74 )
```

```
AdjIdx: 0x4301d, LIFB: 0 , LIF: Vlan55 (0x74 ), DI: 0x0
```

```
DMAC: 0000.0000.0156 SMAC: 0000.0000.0155
```

```
AdjIdx: 0x4301e, LIFB: 0 , LIF: Vlan55 (0x74 ), DI: 0x0
```

```
DMAC: 0000.0000.0256 SMAC: 0000.0000.0155
```

```
n7k-2# show system internal forwarding vrf Infrastructure adjacency entry 0x4301d module 4 det
```

```
Device: 1 Index: 0x4301d DMAC: 0000.0000.0156 SMAC: 0000.0000.0155  
LIF: 0x74 (Vlan55) DI: 0x0 ccc: 4 L2_FWD: NO RDT: NO  
packets: 0 bytes: 0 zone enforce: 0
```

```
n7k-2# show system internal forwarding vrf Infrastructure adjacency entry 0x4301e module 4 det
```

```
Device: 1 Index: 0x4301e DMAC: 0000.0000.0256 SMAC: 0000.0000.0155  
LIF: 0x74 (Vlan55) DI: 0x0 ccc: 4 L2_FWD: NO RDT: NO  
packets: 0 bytes: 0 zone enforce: 0
```

LAYER3 CHECK

- ACL using LOG keywords Applied to interface
- Netflow Data
- Wireshark
- ELAM
- Ethanalayzer
- Contact TAC

ACL Statistics



- ACLs Statistics are **NOT** enabled by default (fundamental difference with IOS) because they require the ACEs NOT to be merged and this affects the TCAM utilization.
- However Statistics can be enabled on a per-ACL base:

```
DC3(config)# ip access-list my-acl
DC3(config-ip-acl)# deny udp any any
DC3(config-ip-acl)# permit ip any any
DC3(config-ip-acl)# statistics per-entry
```
- Statistics can be viewed with the usual show command:

```
show ip access-list my-acl
```
- Detailed PBACL statistics (showing a separate counter for each group combination) can be obtained with:

```
show ip access-list my-acl expanded
```
- When statistics are not enabled, no counters are shown
- Clearing is done as usual:

```
clear ip access-list my-acl
```

ACL Logging

- ACL logging is enabled by including the **log** keyword in an ACL rule. They are specified as any other matching criterion
- The Sup receives a copy of the packet. The original packet is forwarded/dropped in hardware with **no performance penalty**.
- ACLs are applied in hardware
- The CPU is protected by using one of the available rate limiters. Forwarding engine hardware enforces rate to avoid saturating inband interface CPU. **hardware rate-limit access-list-log** command adjusts rate (100 pps by default)
- ACL Logging can be a useful tool during troubleshooting. Use ACL logging to sample specific packets from data plane. Use onboard ethanalyzer (wireshark) to analyze sampled packets

Session Manager for ACL

- Configuration Session mode allows to “dry-run” the configuration against the system resources availability
- After entering the desired configuration, the user can then **verify** it
- The system verifies that the configuration is valid and that there are enough resources available to accommodate such configuration.
- After the configuration has been successfully verified, it can be committed using the “**commit**” CLI command
- Only at this point the system goes ahead and applies the configuration

```
Nexus# configure session My-Acl-Session
Config Session started, Session ID is 1

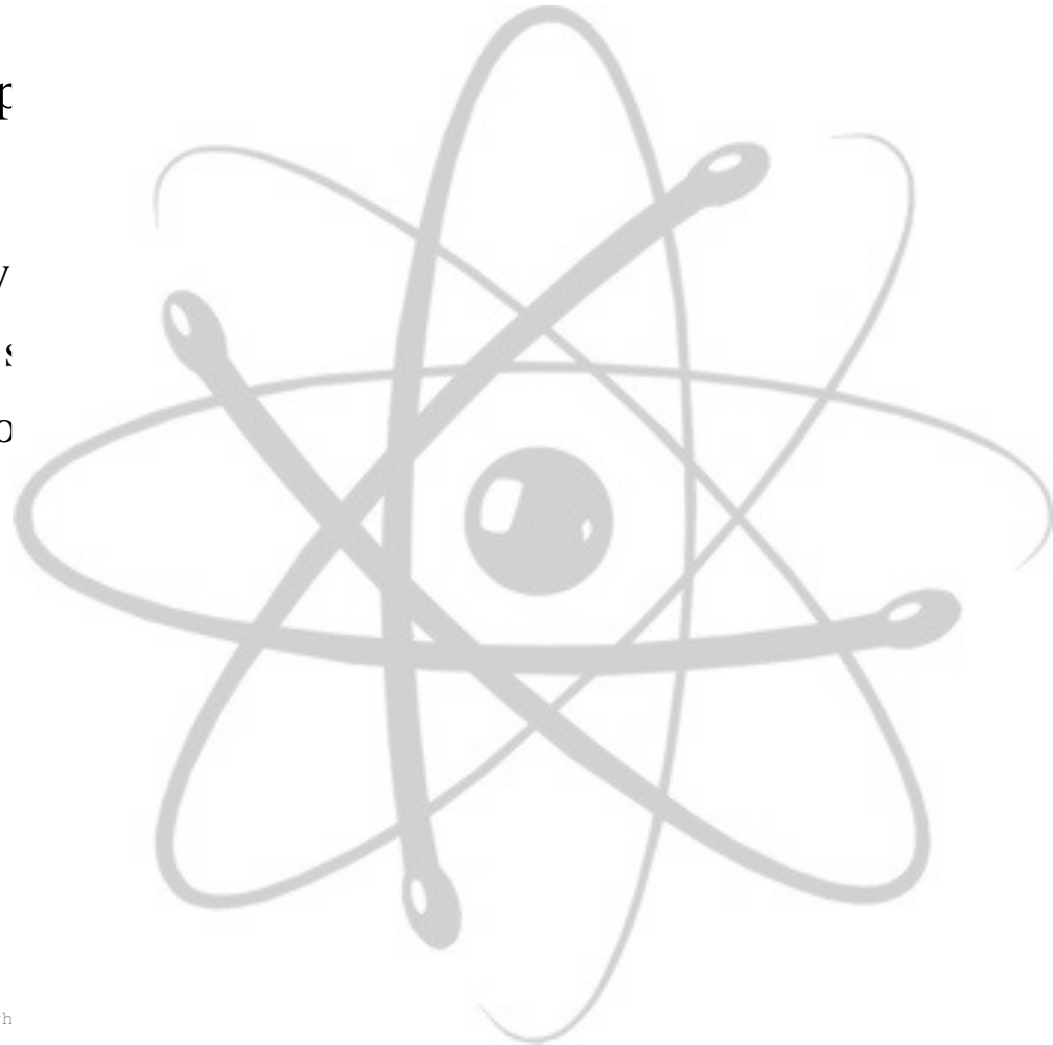
Nexus(config-s)# ip access-list acl1
Nexus(config-s-acl)# permit tcp 11.11.11.11/24 any
Nexus(config-s-acl)# permit tcp 11.11.11.12/24 any
Nexus(config-s-acl)# permit tcp 11.11.11.13/24 any
Nexus(config-s-acl)# exit
Nexus(config-s)#
Nexus(config-s)# interface ethernet 1/1
Nexus(config-s-if)# ip access-group acl1 in
Nexus(config-s-if)#exit
Nexus(config-s)# verify
Verification successful

Nexus(config-s)# commit
Commit successful
```



Atomic Programming

- *Atomic Programming* allows policy updates with *no impact to traffic*
- NX-OS supports *atomic* ACL programming
- A 3 step process:
 - Programming the new policy
 - Enabling the new policy (by setting the ACL to *enable*)
 - Freeing the ACL TCAM resources



Atomic Programming

- **Atomic Programming** requires available hardware resources to succeed: M free entries where M is the number of entries of the new ACL to be programmed.
- Atomic Programming is enabled by default
- If there are insufficient amount of free resources an **error** is returned and no modifications are done to the hardware tables
- In case of insufficient resources the user can disable the atomic programming and just perform the update non-atomically

```
Nexus# conf t
Nexus(config)# no hardware access-list update atomic
```

- Non-atomic programming will affect the traffic briefly
- By default the affected traffic will be **dropped**
- This behavior can be changed by issuing the following command:

```
Nexus# conf t
Nexus(config)#hardware access-list update default-result permit
```

Resource Utilization details

show hardware access-list resource utilization module <mod>

Hardware Modules	Used	Free	Percent Utilization

Tcam 0, Bank 0	1	16383	0.000
Tcam 0, Bank 1	4121	12263	25.000
Tcam 1, Bank 0	4013	12371	24.000
Tcam 1, Bank 1	4078	12306	24.000
LOU	2	102	1.000
Both LOU Operands	0		
Single LOU Operands	2		
TCP Flags	0	16	0.000
Protocol CAM	4	3	57.000
Mac Etype/Proto CAM	0	14	0.000
Non L4op labels, Tcam 0	3	6140	0.000
Non L4op labels, Tcam 1	3	6140	0.000
L4 op labels, Tcam 0	0	2047	0.000
L4 op labels, Tcam 1	1	2046	0.000

CL TCAM banks (ACEs)

LOUs (gt, lt, neq, range operators/operands)

Other resources (TCP flags, static matches)

ACL labels with no L4op pointers
(identify unique policy combinations)

ACL labels with L4op pointers
(identify unique policy combinations)

n7010#

Proxy Routing - Limitations

- M1 – F1 – Who does the routing?
- M2 – F2 – Interoperability
- M2 – F2E – Routing protocols
- M2 – F3 – Independent, will work fine

Where to find the best materials to understand if it works or not?

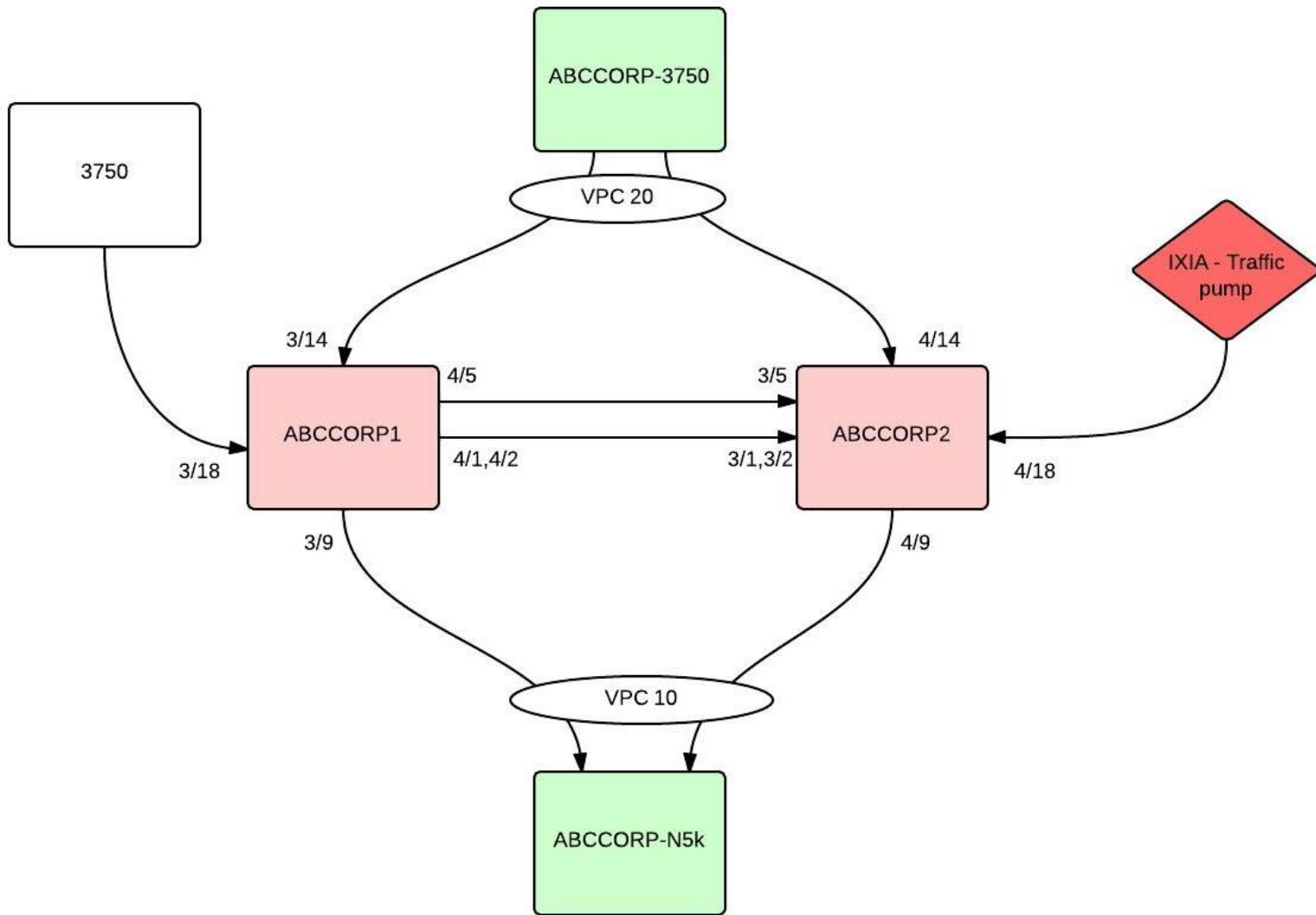
Release Notes

Datasheet

ELAM/ Ethalyzer

- <https://supportforums.cisco.com/document/9878381/innovative-troubleshooting-tools-cisco-switches>
- <https://supportforums.cisco.com/discussion/11975436/packet-capture-capabilities-cisco-routers-and-switches>
- Ethalyzer is only for CPU punt traffic and not for flow-through
- ELAM is for Flow-through traffic as well

DEMO





CISCO