

CISCO *Live!*

GO BEYOND

#CiscoLiveAPJC



Designing High Availability for your Data Center with Resiliency

Mastering High Availability in the Data Center

Moyuresh Biswas

BRKDCN-2947

CISCO *Live!*

#CiscoLiveAPJC

Cisco Webex App

Questions?

Use Cisco Webex App to chat with the speaker after the session

How

- 1 Find this session in the Cisco Live Mobile App
- 2 Click “Join the Discussion”
- 3 Install the Webex App or go directly to the Webex space
- 4 Enter messages/questions in the Webex space

Webex spaces will be moderated by the speaker until November 15, 2024.



<https://ciscolive.ciscoevents.com/ciscolivebot/#BRKDCN-2947>



Hey ChatGPT ... why is the Data Center Important?

From Cat Videos to Critical Infrastructure

Where the Magic Happens

The Nerve Center of the Information Age

From Bits to Business

No Cloud Without a Ground



This has been a year full of engineering excellence that sometimes can go unnoticed. Besides all the visible changes you see on our app, here are some of the most important improvements we have made under the hood.

- Consolidated the tech stacks for For you, Following, Search, Profiles, Lists, Communities and Explore around a singular product framework.
- Completely rebuilt the For you serving and ranking systems from the ground up, resulting in a decrease 90% reduction in lines of code from 700K to 70K, a 50% decrease in our compute footprint, and an 80% increase in the throughput of posts scored per request.
- Unified the For you and video personalization and ranking models, which significantly improved video recommendation quality.
- Refactored the API middleware layer of our tech stack and in doing so simplified the architecture by removing more than 100K lines of code and thousands of unused internal endpoints and eliminating unadopted client services.
- Reduced post metadata sourcing latency by 50%, and global API timeout errors by 90%.

- Blocked bots and content scrapers at a rate +37% greater than 2022. On average, we prevent more than 1M bots signup attacks each day and we've reduced DM spam by 95%.

- Shutdown the Sacramento data center and re-provisioned the 5,200 racks and 148,000 servers, which generated more than \$100M in annual savings. In total, we freed up 48 MW of capacity and tore down 60k lbs. of network ladder rack before re-provisioning it to other data centers.

- Optimized our usage of cloud service providers and began doing much more on-prem. This shift has reduced our monthly cloud costs by 60%. Among the changes we made was a shift of all media/blob artifacts out of the cloud, which reduced our overall cloud data storage size by 60%, and separately, we succeeded in reducing cloud data processing costs by 75%.

- Built on-prem GPU Supercompute clusters and designed, developed, and delivered 43.2Tbps of new network fabric architecture to support the clusters.

- Scaled network backbone capacity and redundancy, which resulted in \$13.9M/year in savings.

- Started automated peak traffic failover tests to validate the scalability and availability of the entire platform continuously.

3:05 PM · Oct 27, 2023 · 37.4M Views

“Optimised our usage of cloud service providers and began doing much more on-prem. This shift has reduced our monthly cloud costs by 60%”

- X (Formerly Twitter) Engineering

Business Drivers

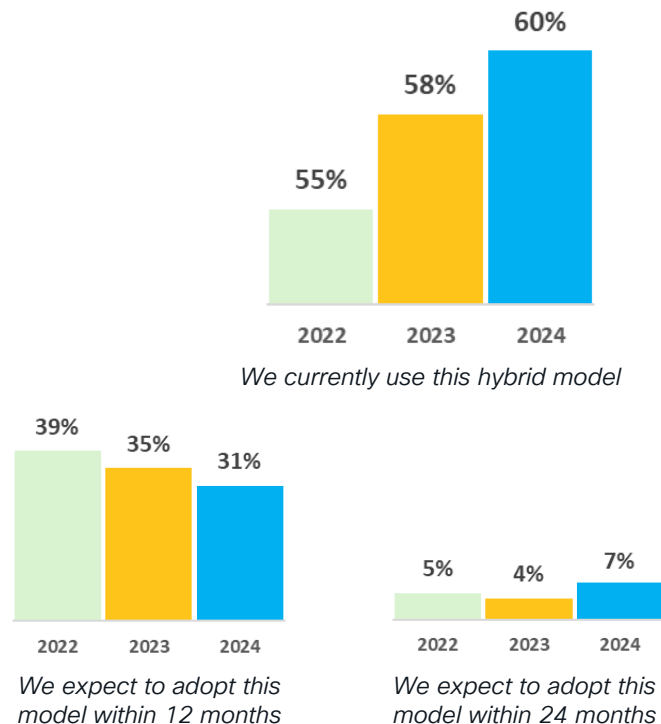
“Optimized our usage of cloud service providers and began doing much more **on-prem**. This shift has **reduced** our monthly cloud **costs** by 60%”

- X (Formerly Twitter) Engineering

Cost Risk Compliance Policy Sovereignty

Hybrid Cloud Adoption

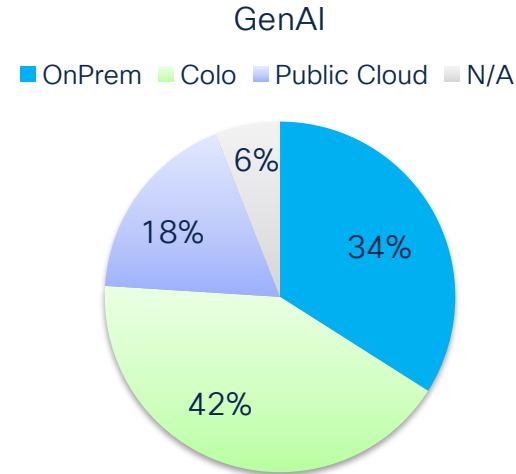
Consistent growth in adoption of Hybrid cloud, encompassing On Premise infrastructure and Colocation facilities



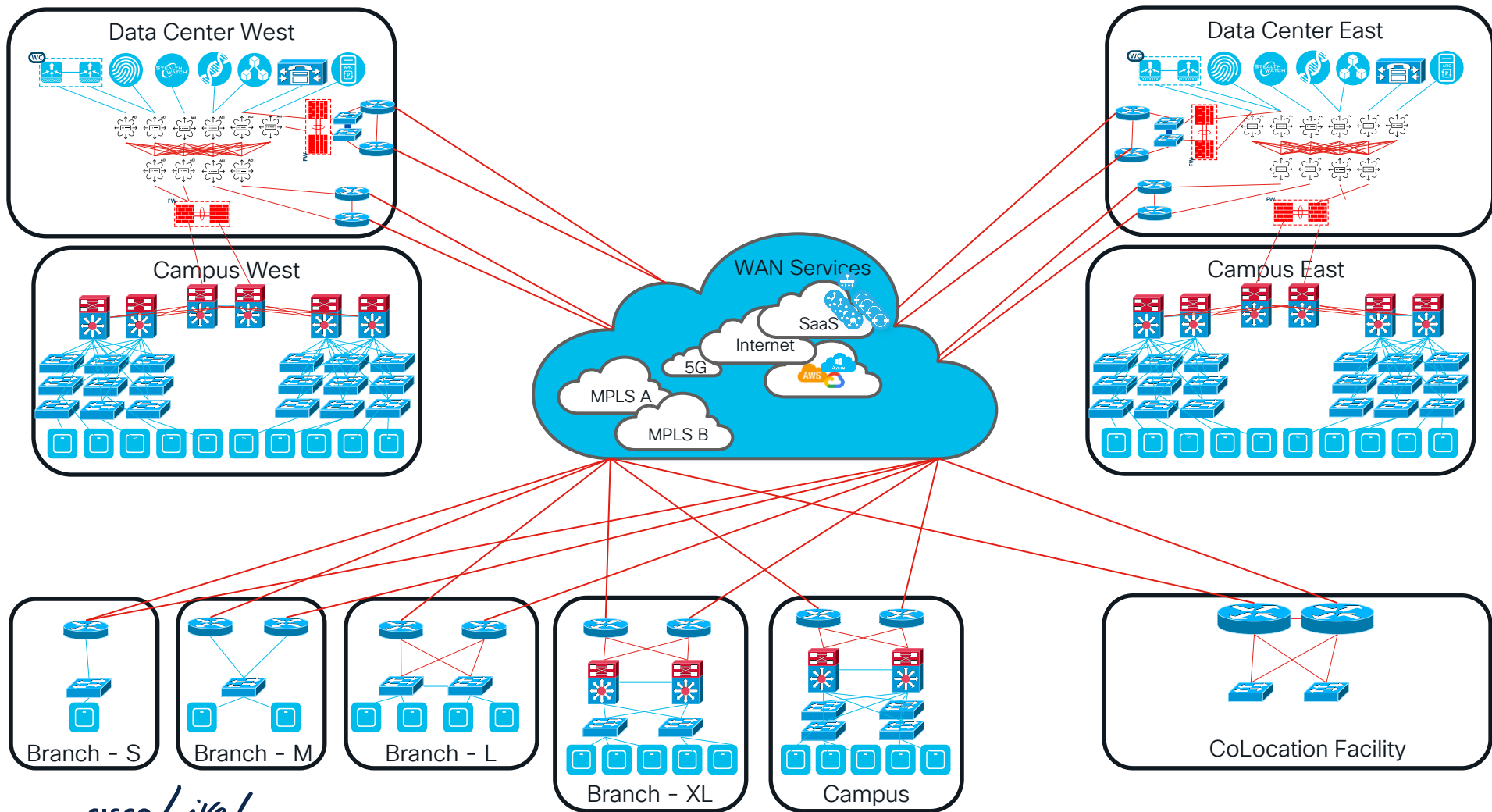
Source: 2024 State of the Data Center, by CoreSite

AI is a key decision driver

The rising expectations surrounding AI, particularly **Generative AI**, are compelling IT leaders to reconsider their strategies for hosting **computationally demanding** workloads

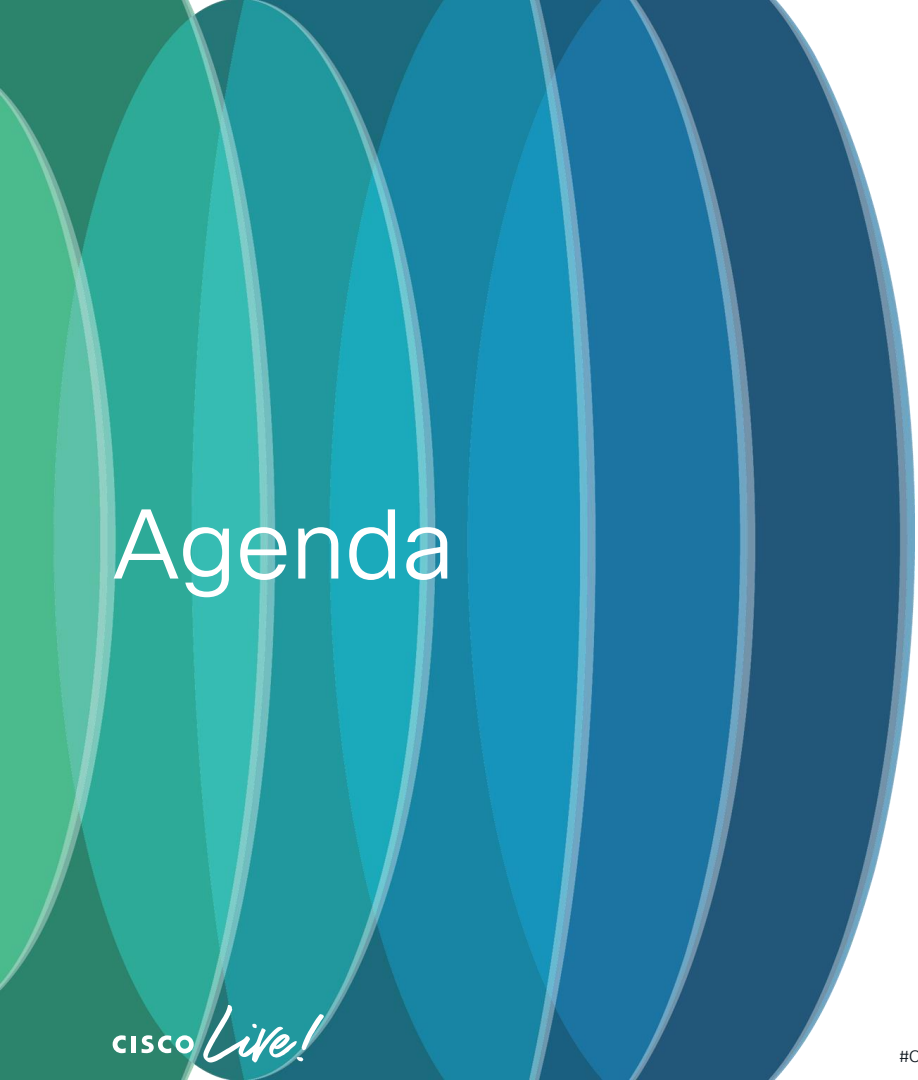


Source: 2024 State of the Data Center, by CoreSite



cisco *Live!*

Mastering High Availability in the Data Center



Agenda

- Fundamentals
- Service Level High Availability
- System Level High Availability
- Network Level High Availability
- Conclusion

How do we measure Availability

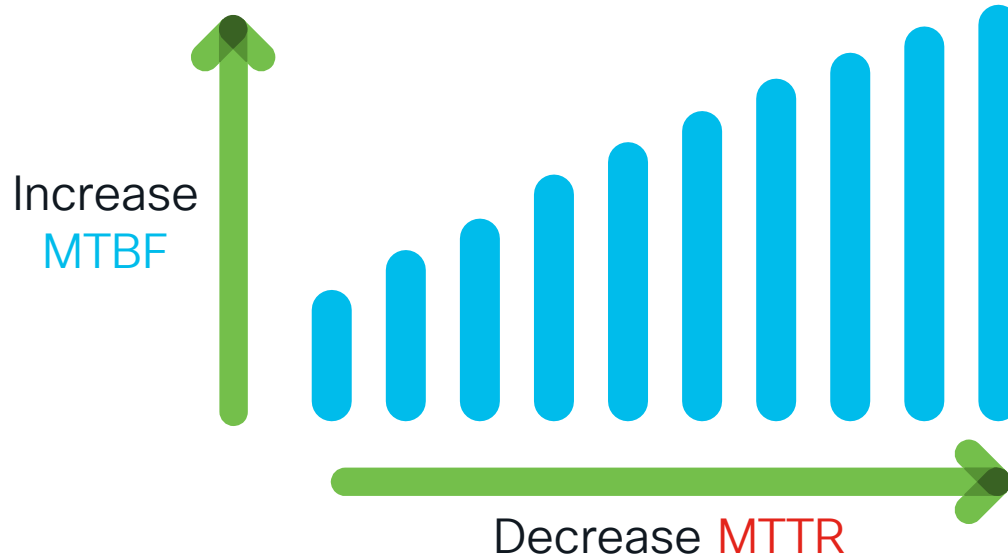
It is a function of

MTBF: Mean Time Between Failures

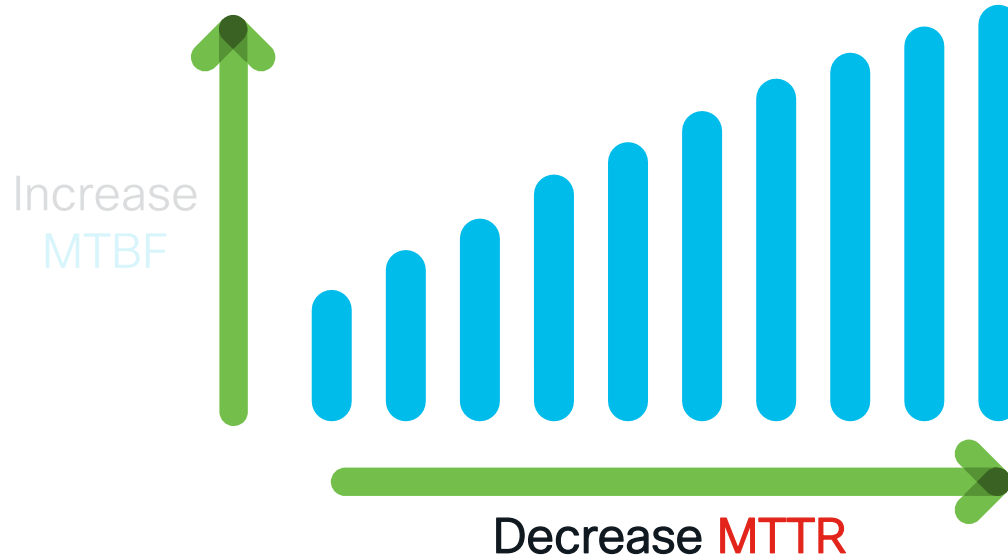
MTTR: Mean Time To Repair

$$\text{Availability} = \frac{\text{MTBF}}{\text{MTBF} + \text{MTTR}}$$

Availability



Availability



Example – Increase availability by reducing MTTR

Component with MTBF=87,600 hours

$$\text{Availability} = \frac{\text{MTBF}}{\text{MTBF} + \text{MTTR}}$$

24 hour depot replacement

2 hour 24 minutes
predicted annual
downtime

$$\text{Availability} = \frac{87,600}{87,600 + 24} = .9997 \text{ (99.9\%)}$$

4 hour depot replacement

24 minutes
predicted annual
downtime

$$\text{Availability} = \frac{87,600}{87,600 + 4} = .99995 \text{ (99.99\%)}$$

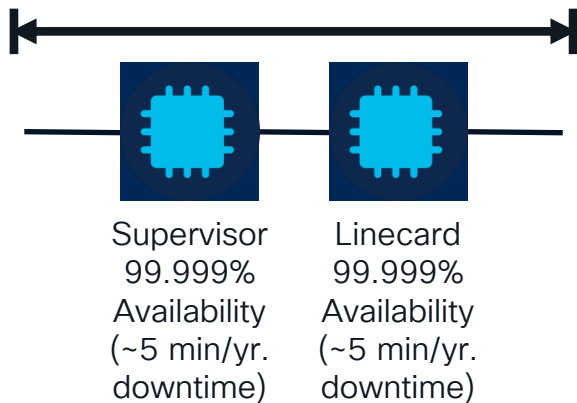
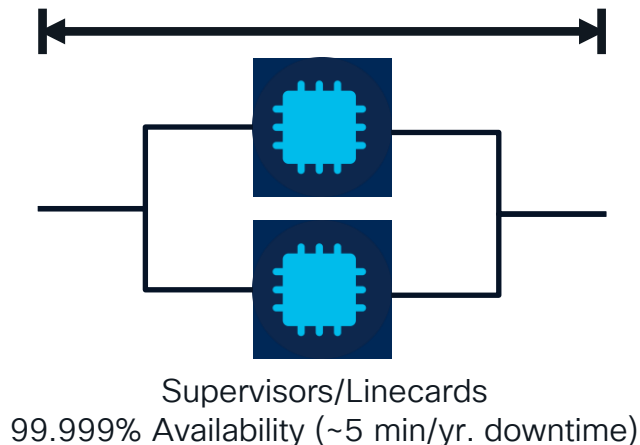
Warm spare (10 minute restore)

1 minute
predicted annual
downtime

$$\text{Availability} = \frac{87,600}{87,600 + .16666} = .999998 \text{ (99.999\%)}$$

The redundancy effect

- Single component functioning in series
- System predicted availability: **99.98%**
(~10 min./year predicted downtime)
- Redundant components functioning in parallel
- System predicted availability: **99.999999%**
(~½ second/year predicted downtime)



Journey to Mastery

Service-Level HA

NX-OS Architecture
Process Restartability
HA Policy

System-Level HA

Physical Redundancy
Upgrades & Patching

Network-Level HA

Layer 2 / 3 Redundancy
DC HA Architecture



Journey to Mastery

Service-Level HA

NX-OS Architecture
Process Restartability
HA Policy

System-Level HA

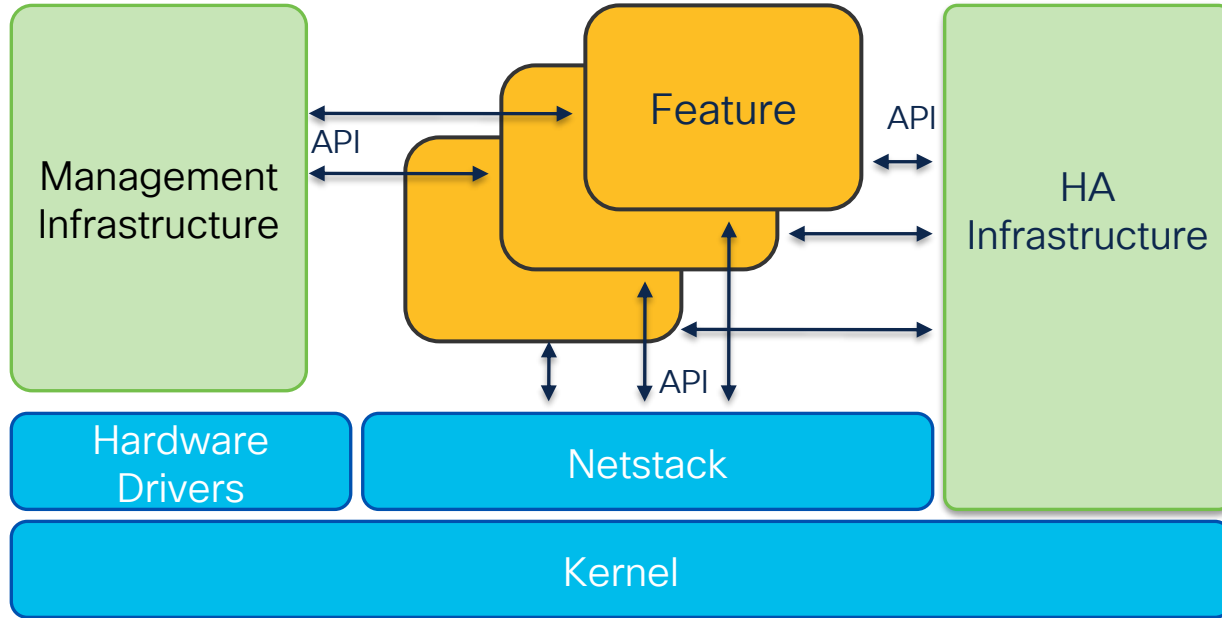
Physical Redundancy
Upgrades & Patching

Network-Level HA

Layer 2 / 3 Redundancy
DC HA Architecture



NX-OS Architecture



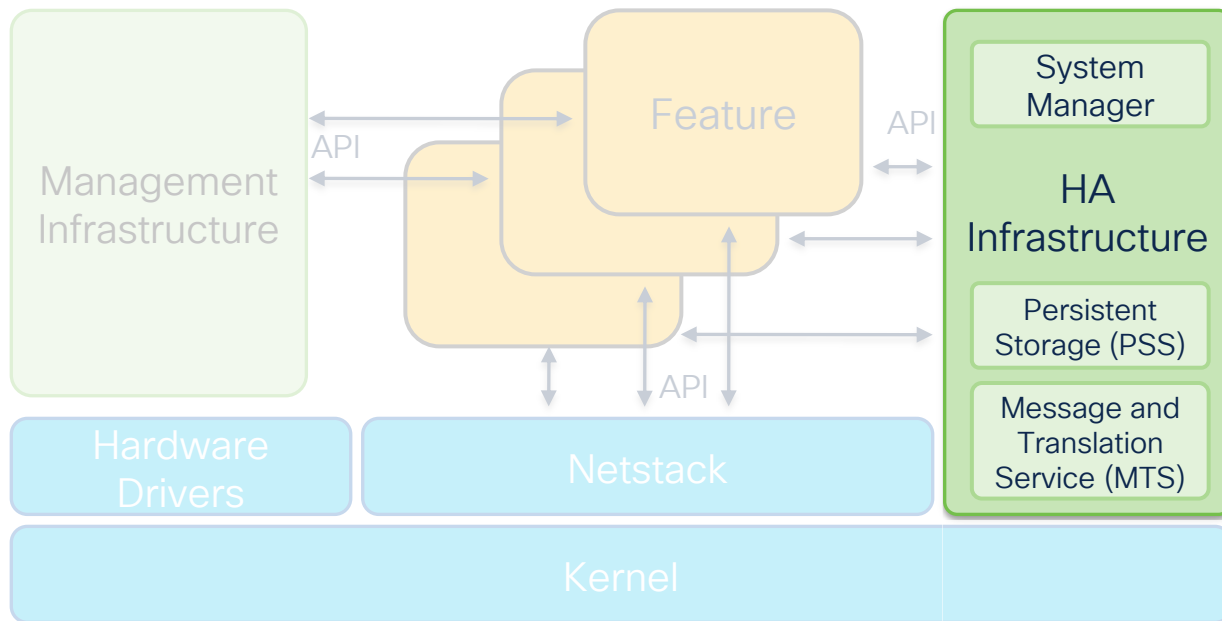
NX-OS Features vs Services vs Processes



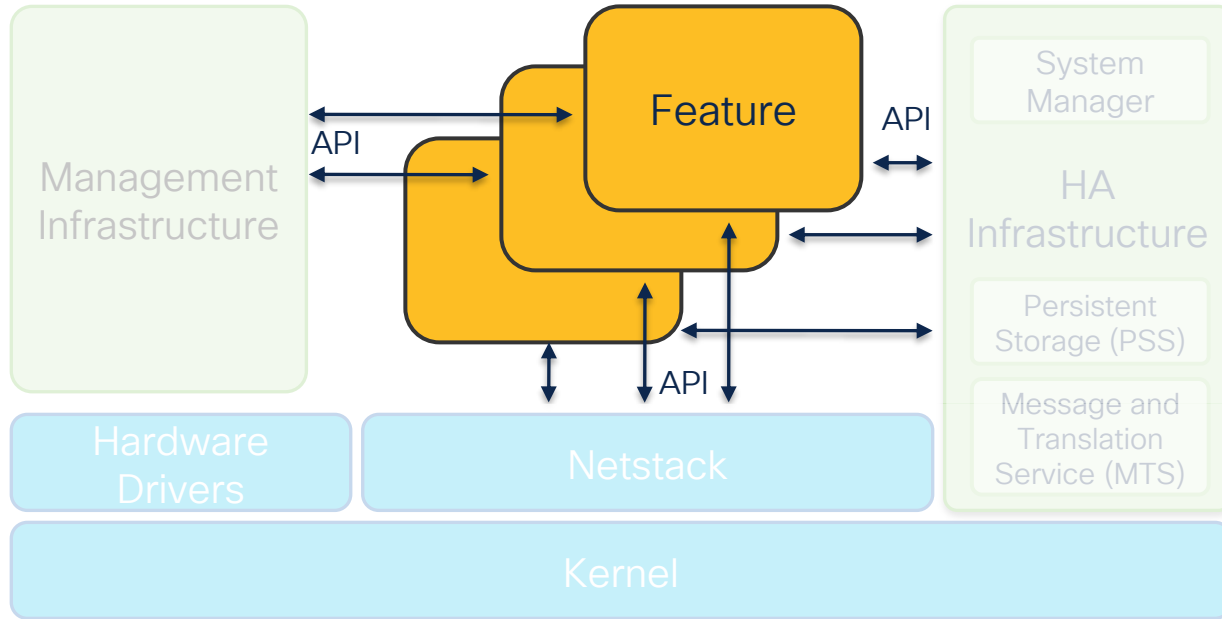
For Your
Reference

- Features are written as user-space services
- A feature can start one or more services
- Each instance of a service runs as an independent process
 - multiple instances of a service possible (e.g. multiple OSPF process)

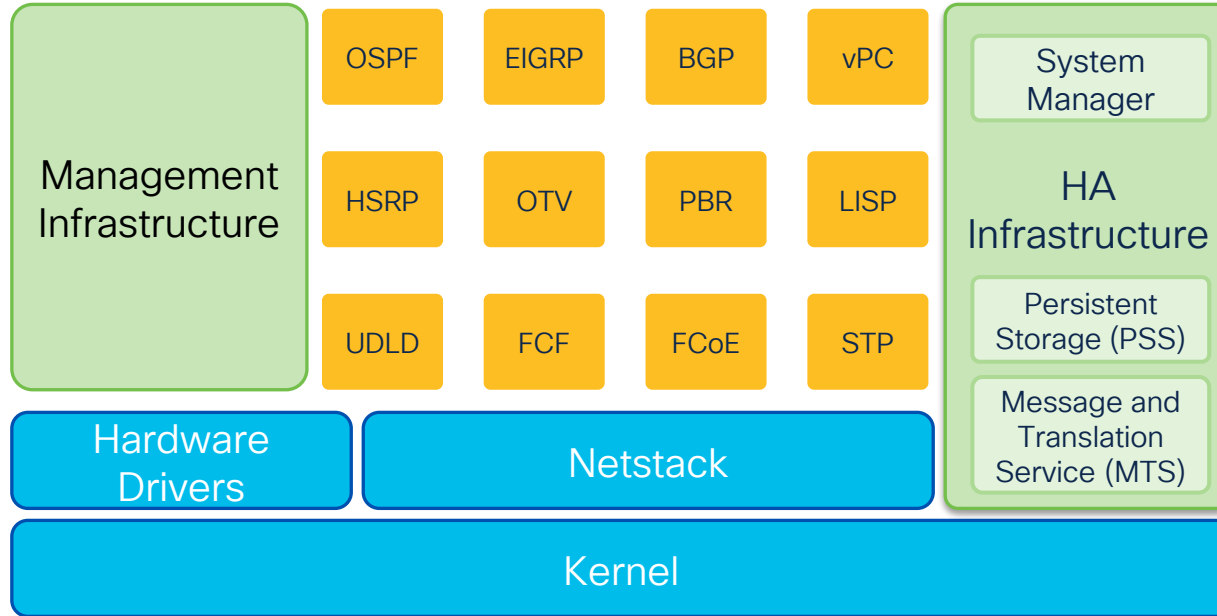
NX-OS HA Infrastructure



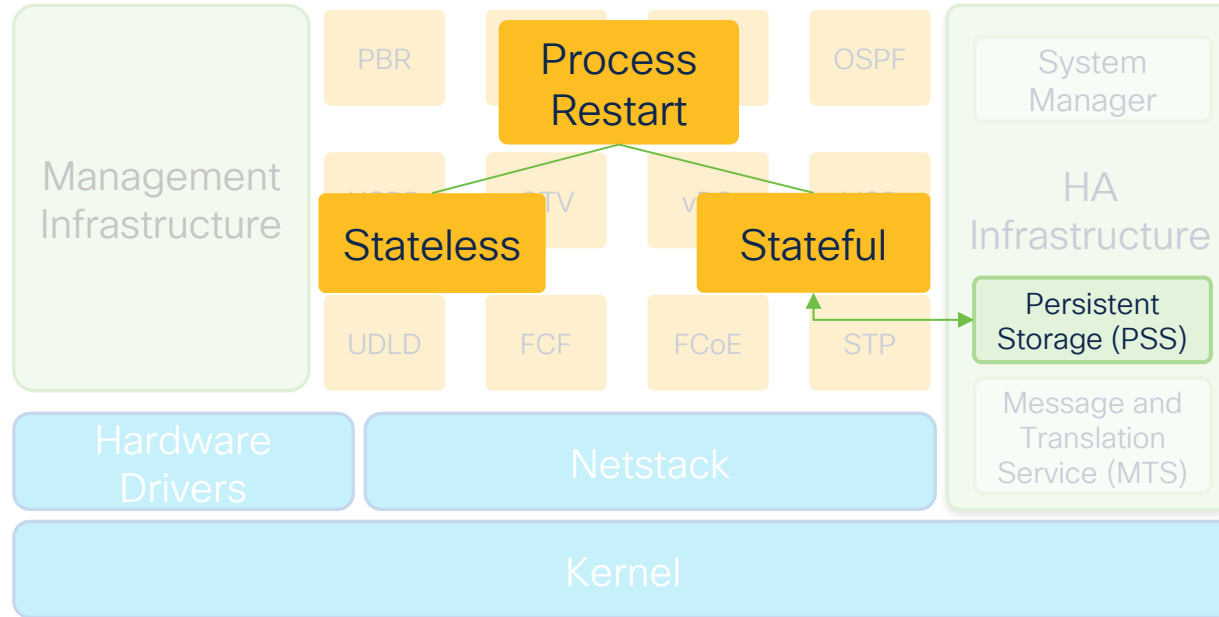
NX-OS HA Infrastructure



NX-OS Process Restartability



NX-OS Process Restartability



High Availability Policies for Services

- Each service can have its own internal HA policy
- Primary and Secondary policy, for single and dual supervisor

Service HA Policies	Action (single SUP)	Action (Dual SUP)
Stateful restart	Stateful restart with running config	Stateful restart with running config
Stateless restart	Stateless restart with startup, default or no config	Stateless restart with startup, default or no config
No restart	Don't restart the service	Don't restart the service
Reset	Reload this SUP	Switchover

```
N9K-S1-Leaf1# show system internal sysmgr service name bgp restart-policies
```

HBTO - HeartBeat TimeOut in seconds

HBML - HeartBeat Max Loss allowed

HAPolicy - (minlife in minutes,maxretry,restart type)

Name	UUID	HBTO	HBML	Single Sup Policy		
				Dual Sup Policy	Primary Policy, Single SUP	Secondary Policy, Single SUP
<hr/>						
bgp	283	30	6	(4,3,stateful)	(4,0,reset)	
				(4,3,stateful)	(4,0,switchover)	

Primary Policy, Dual SUP

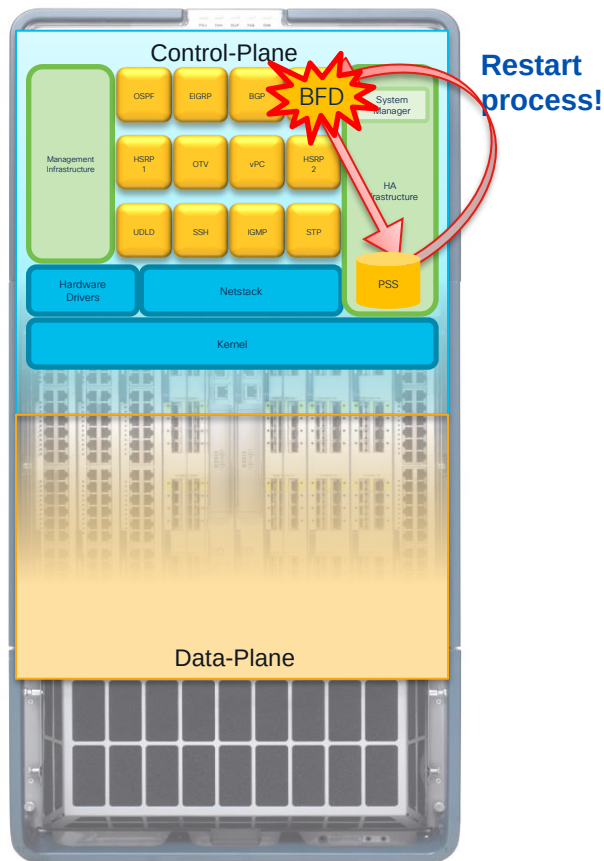
Secondary Policy, Dual SUP

Stateful Process Restart Example

- NX-OS services **Checkpoint** their runtime state to the Persistent Storage Service (PSS) for **Recovery**

A service failure occurs ...

- System manager determines best **Recovery Action** based on HA Policy
- State is recovered, operation resumes
- Total **Recovery Time ~10s ms**
- Process restarts with **No impact on data plane**



Journey to Mastery

Service-Level HA

NX-OS Architecture
Process Restartability
HA Policy

System-Level HA

Physical Redundancy
Upgrades & Patching

Network-Level HA

Layer 2 / 3 Redundancy
DC HA Architecture

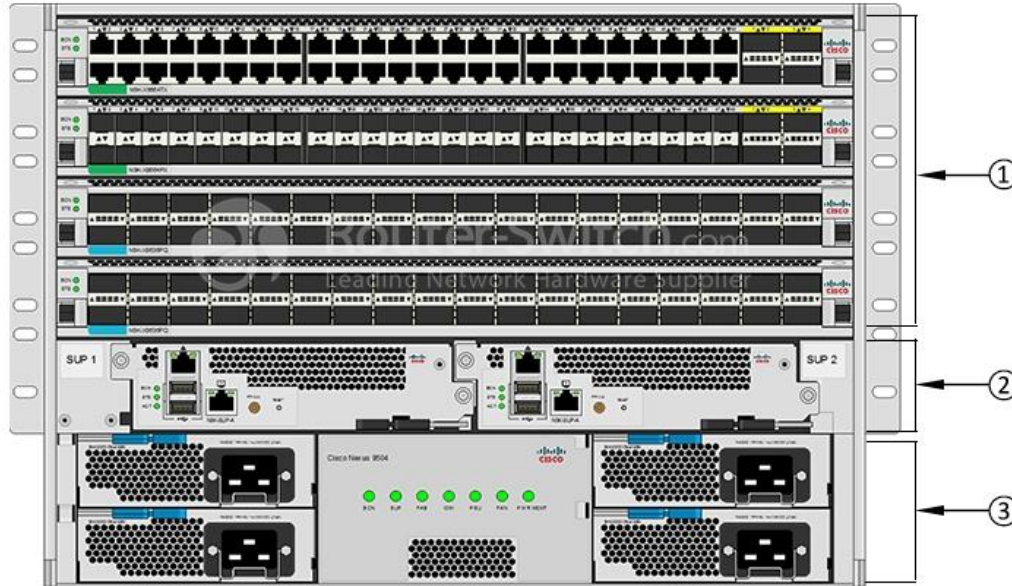


Physical Redundancy

Switch Family	Nexus 9200 Nexus 9300	Nexus 9500 Nexus 9800
Type	Fixed Form Factor Top of Rack (ToR)	Modular Chassis End of Row (EoR)
Redundancy Supported	Power supply Fan tray	Power supply Fan tray Line card Supervisor Switch fabric System controller



Redundant Hardware – Front Panel



Line Cards

Redundant Supervisors

Power Supplies

Nexus 9504

Stateful Switchover (SSO)

with Redundant Supervisors

Nexus 9504



Dual Supervisors

Nexus 9508



Dual Supervisors

Nexus 9516



Dual Supervisors

```
switch# show system redundancy status
```

```
Redundancy mode
```

```
-----  
administrative: HA  
operational: HA
```

```
This supervisor (sup-1)
```

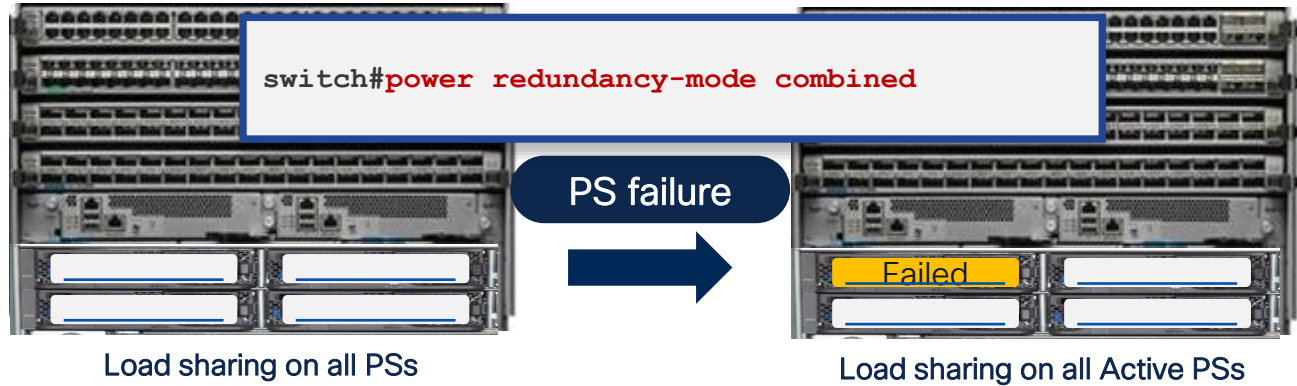
```
-----  
Redundancy state: Active  
Supervisor state: Active  
Internal state: Active with HA standby
```

```
Other supervisor (sup-2)
```

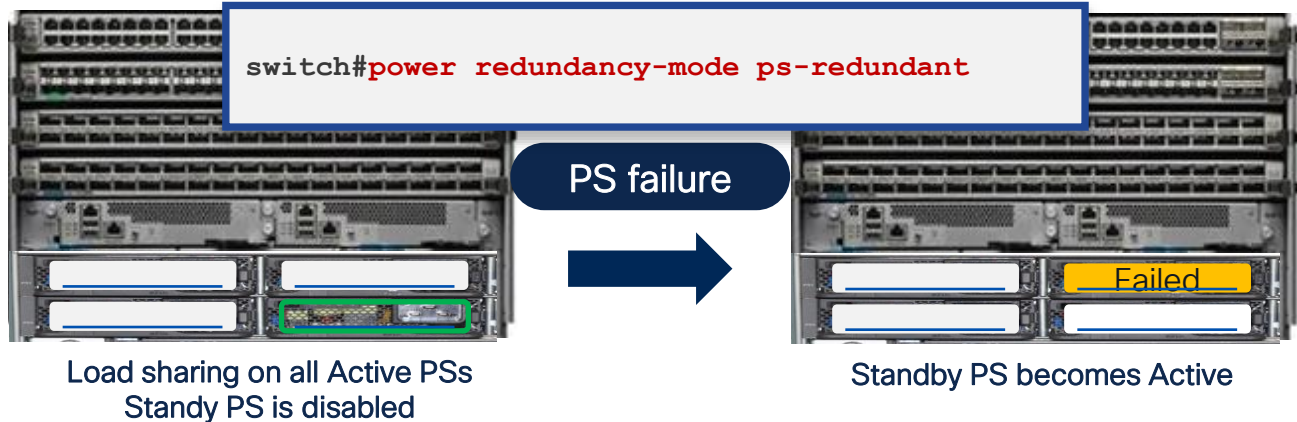
```
-----  
Redundancy state: Standby  
Supervisor state: HA standby  
Internal state: HA standby
```

Power Redundancy Modes

Combined



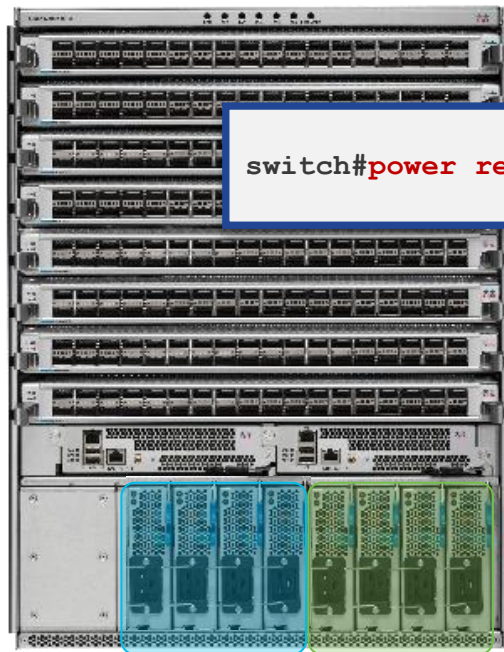
N+1



Power Redundancy Modes

Grid
N+N

```
switch#power redundancy-mode insrc_redundant
```

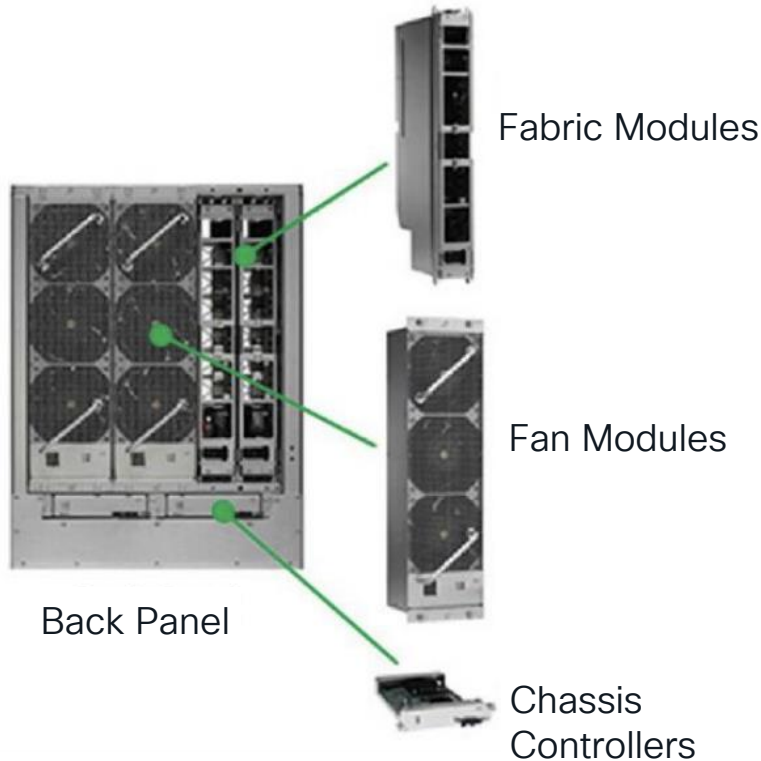


Grid A

Grid B

Power sharing and
redundancy are
available only
within individual Grid

Redundant Hardware – Back Panel



Help in distribution of traffic

Fan Speed Adjust depending on fan failures occurring

Offload the power and fan management
And handle EOBC as well

Show command

```
switch# show module
```

Mod	Ports	Module-Type	Model	Status
1	32	32p 40G Ethernet Module	N9K-X9432PQ	ok
2	52	48x1/10G SFP+ 4x40G Ethernet Module	N9K-X9564PX	ok
5	52	48x1/10G SFP+ 4x40G Ethernet Module	N9K-X9464PX	ok
6	36	36p 40G Ethernet Module	N9K-X9536PQ	ok
7	36	36p 40G Ethernet Module	N9K-X9536PQ	ok
10	32	32p 40G Ethernet Module	N9K-X9432PQ	ok
11	52	48x1/10G-T 4x40G Ethernet Module	N9K-X9564TX	ok
12	52	48x1/10G-T 4x40G Ethernet Module	N9K-X9464TX	ok
15	52	48x1/10G SFP+ 4x40G Ethernet Module	N9K-X9464PX	ok
21	0	Fabric Module	N9K-C9516-FM	ok
22	0	Fabric Module	N9K-C9516-FM	ok
23	0	Fabric Module	N9K-C9516-FM	ok
24	0	Fabric Module	N9K-C9516-FM	ok
25	0	Fabric Module	N9K-C9516-FM	ok
26	0	Fabric Module	N9K-C9516-FM	ok
27	0	Supervisor Module	N9K-SUP-A	ha-standby
28	0	Supervisor Module	N9K-SUP-A	active *
29	0	System Controller	N9K-SC-A	active
30	0	System Controller	N9K-SC-A	standby

Show command

```
switch# show environment power
```

```
Power Supply:
```

```
Voltage: 12 Volts
```

Power Supply	Model	Actual Output (Watts)	Actual Input (Watts)	Total Capacity (Watts)	Status
1	NXA-PAC-1100W-PE2	0 W	0 W	0 W	Shutdown
2	NXA-PAC-1100W-PE2	254 W	262 W	1100 W	Ok

```
Power Usage Summary:
```

```
Power Supply redundancy mode (configured)
```

```
Power Supply redundancy mode (operational)
```

```
PS-Redundant
```

```
Non-Redundant
```

```
Total Power Capacity (based on configured mode)
```

```
1100.00 W
```

```
Total Grid-A (first half of PS slots) Power Capacity
```

```
0.00 W
```

```
Total Grid-B (second half of PS slots) Power Capacity
```

```
1100.00 W
```

```
Total Power of all Inputs (cumulative)
```

```
1100.00 W
```

```
Total Power Output (actual draw)
```

```
254.00 W
```

```
Total Power Input (actual draw)
```

```
262.00 W
```

```
Total Power Allocated (budget)
```

```
N/A
```

```
Total Power Available for additional modules
```

```
N/A
```

```
switch09-n9324#
```

Journey to Mastery

Service-Level HA

NX-OS Architecture
Process Restartability
HA Policy

System-Level HA

Physical Redundancy
Upgrades & Patching

Network-Level HA

Layer 2 / 3 Redundancy
DC HA Architecture



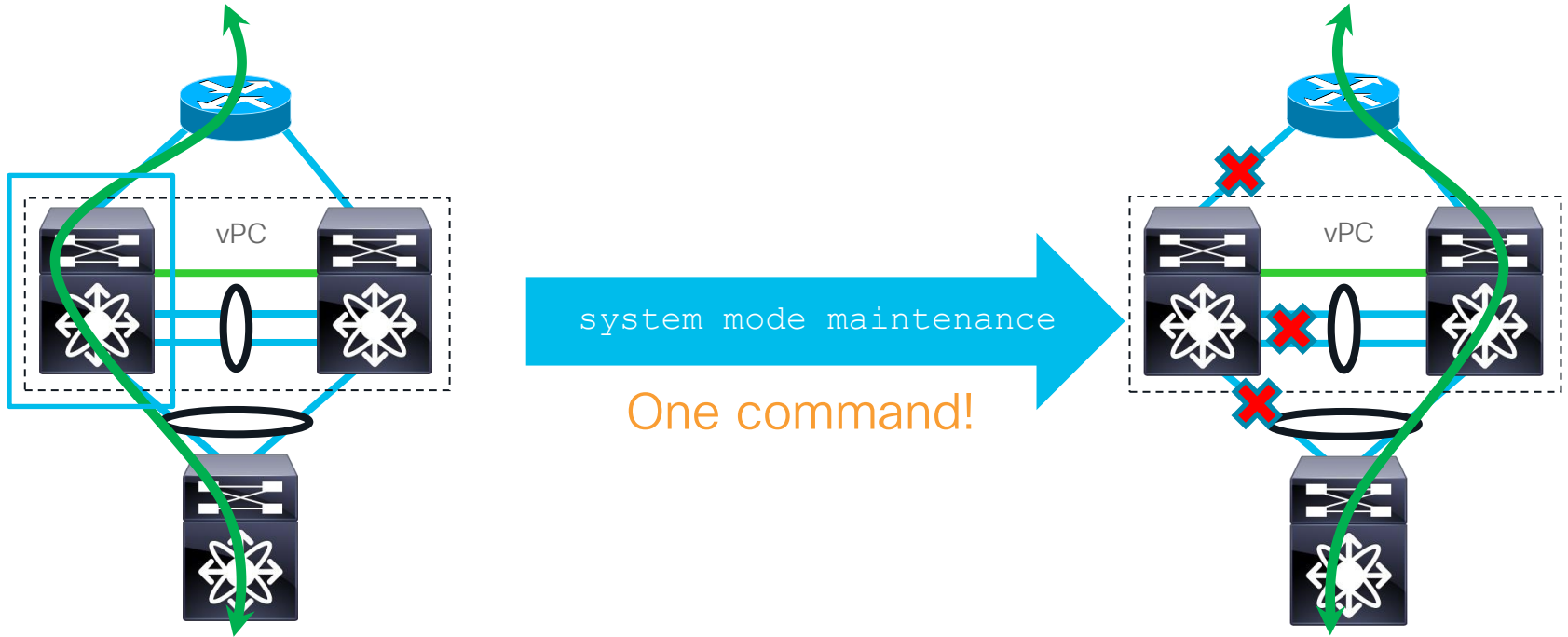
GIR

Graceful Insertion & Removal



Graceful Insertion and Removal

Isolation of Switch from network



Graceful Insertion and Removal

Return of Switch into network



Graceful Insertion and Removal Recommendations

Maintenance Window Timer

- Can be configured to keep the switch in Maintenance Mode for period of time
- Started on entry to maintenance mode
- When configured time elapses, switch automatically returns to normal mode

```
switch# configure terminal
switch(config)# system mode maintenance timeout <Value>
               <5-65535>  Timer value in minute
```

```
switch# show maintenance timeout
Maintenance mode timeout value: 30 minutes
```

SMU

Software
Maintenance
Upgrades



What is SMU Impact

Each SMU comes with a README file

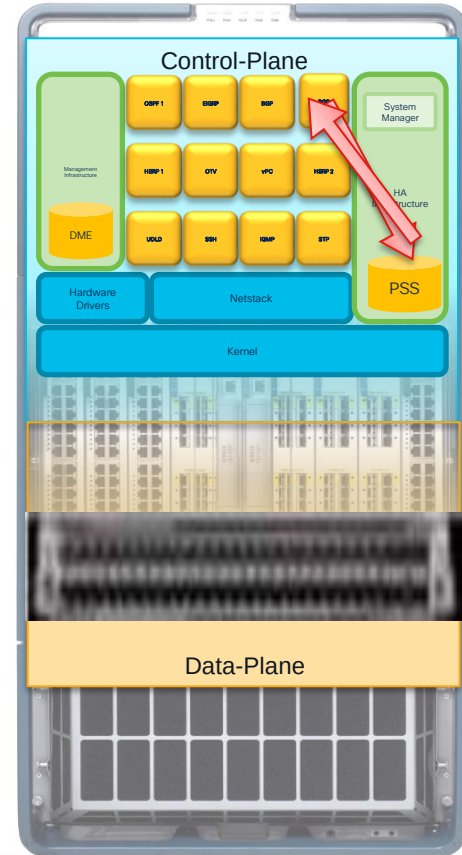
SMU Types	Impact
Hot	Process Restart
Cold	Full Box Reload

```
#####  
#  
# Readme for SMU nxos64.CSCvz98895-n9k_ALL-1.0.0-10.3.1.167.lib32_64_n9000  
#  
# Copyright (c) 2021 by Cisco Systems, Inc.  
# All rights reserved.  
#  
#####  
#####  
Name:      nxos64.CSCvz98895-n9k_ALL  
Version:   10.2.1.167  
Compressed size:89146  
DTS:      CSCvz98895  
Description: Kr2f-122: nxpython3 core while collecting show-tech-detail > icam script is failing  
Type:      non-reload  
MD5 :      a64fa50587d045828421860b840cc366
```

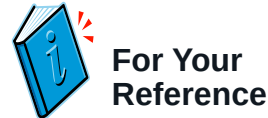
SMU Impact

Non-reload (HOT) SMU update on Nexus 9000

- Services Checkpoints their state in PSS
- Installer asks System Manager to stop the process
- Installer updates file system with files from SMU
- Installer asks System Manager to restart the process
- State is recovered, operation resumes
- Total Recovery Time ~10ms
- **No impact on data plane**



SMU install commands



Adding SMU

```
switch# install add nxos`.CSCxy68793-n9k_ALL-1.0.0-10.3.1.132.lib32_n9000.rpm)
[#####] 100%
Install operation 26 completed successfully at Sat May 4 00:27:02 2024
```

Activating SMU

```
switch# install activate nxos.CSCxy68793-n9k_ALL-1.0.0-10.3.1.132.lib32_n9000.rpm
Activating the patch (/nxos.CSCxy68793-n9k_ALL-1.0.0-10.3.1.132.lib32_n9000.rpm)
[#####] 100%
Install operation 27 completed successfully at Sat May 4 00:27:11 2024
```

Committing SMU

```
switch# install commit
[#####] 100%
Install operation 28 completed successfully at Sat May 4 00:29:15 2024
```

Any failures/reloads between activate and commit result in a rollback

ISSU

In-Service Software Upgrades

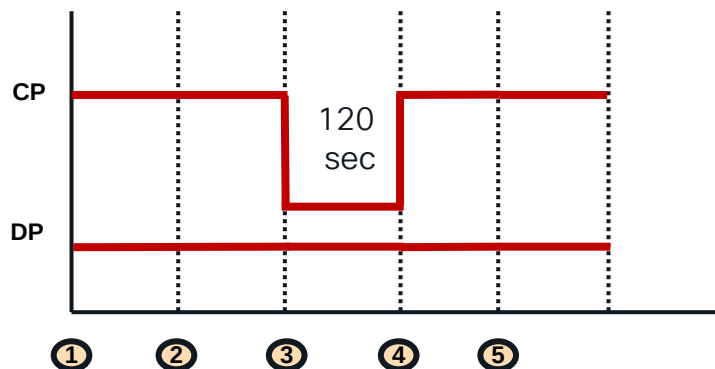


ISSU / Hitless Upgrade

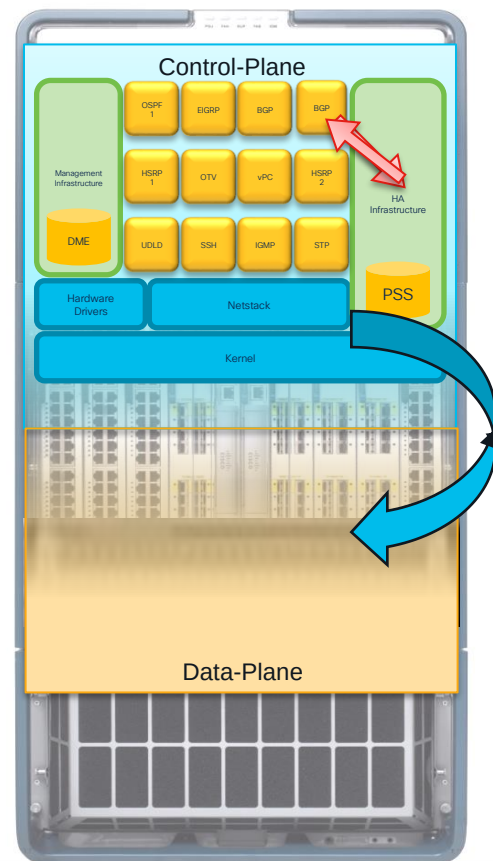
- Upgrade device software while switch continues to forward traffic
- Supported on Nexus ToR fixed switches with single supervisor
- Brings same outcome to single supervisor switches as SSO does to modular switches dual supervisor

ISSU on the Nexus ToR Switch

- Single supervisor
- Hitless Upgrade
 - Control plane is inactive during reload while Data plane is forwarding



1. Pre Upgrade Check
 - Config Locked
 - Stable Network
2. Save State
3. Switch to new Image
4. Restore control plane from saved State
5. Reconcile with Data Plane



Upgrade Impact

New Image Version

```
switch# show install all impact nxos bootflash:nxos64.10.3.1.F.bin
Installer will perform impact only check. Please wait.

Verifying image bootflash:/nxos64.10.3.1.F.bin for boot variable "nxos".
[#####] 100% -- SUCCESS

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

Preparing "nxos" version info using image bootflash:/nxos64.10.3.1.F.bin.
[#####] 100% -- SUCCESS

Preparing "bios" version info using image bootflash:/nxos64.10.3.1.F.bin.
[#####] 100% -- SUCCESS

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

Upgrade Impact

```
switch# show install all impact nxos bootflash:nxos64.10.3.1.F.bin
```

```
Installer will perform impact only check. Please wait.
```

```
Verifying image bootflash:/nxos64.10.3.1.F
```

```
[#####] 100% -- SUCCESS
```

```
Verif Compatibility check is done:
```

```
[#### Module bootable Impact Install-type Reason
```

```
Prepa 1 yes disruptive reset default upgrade is not hitless
```

```
[#### 27 yes disruptive none default upgrade is not hitless
```

```
Prepa
```

```
[####
```

```
Images will be upgraded according to following table:
```

Notif	Module	Image	Running-Version(pri:alt)	New-Version	Upg-Required
[####	1	lcn9k	10.2 (2)	10.2 (2)	no
	27	nxos	10.2 (2)	10.2 (2)	no
	27	bios	v05.45 (07/05/2021) :v05.39 (08/30/2019)	v05.45 (07/05/2021)	no

New Image Version

The impact

Reason

Activate ISSU

- Default upgrade process is disruptive
- ISSU needs to be activated

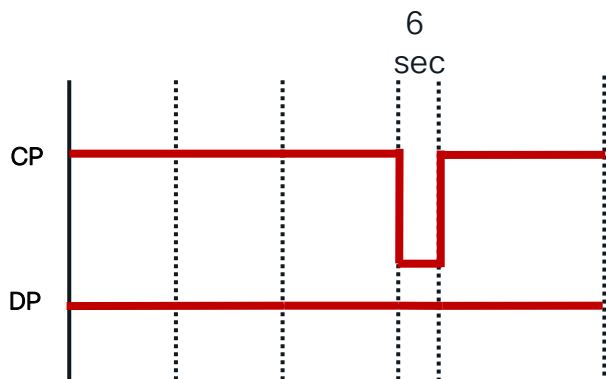
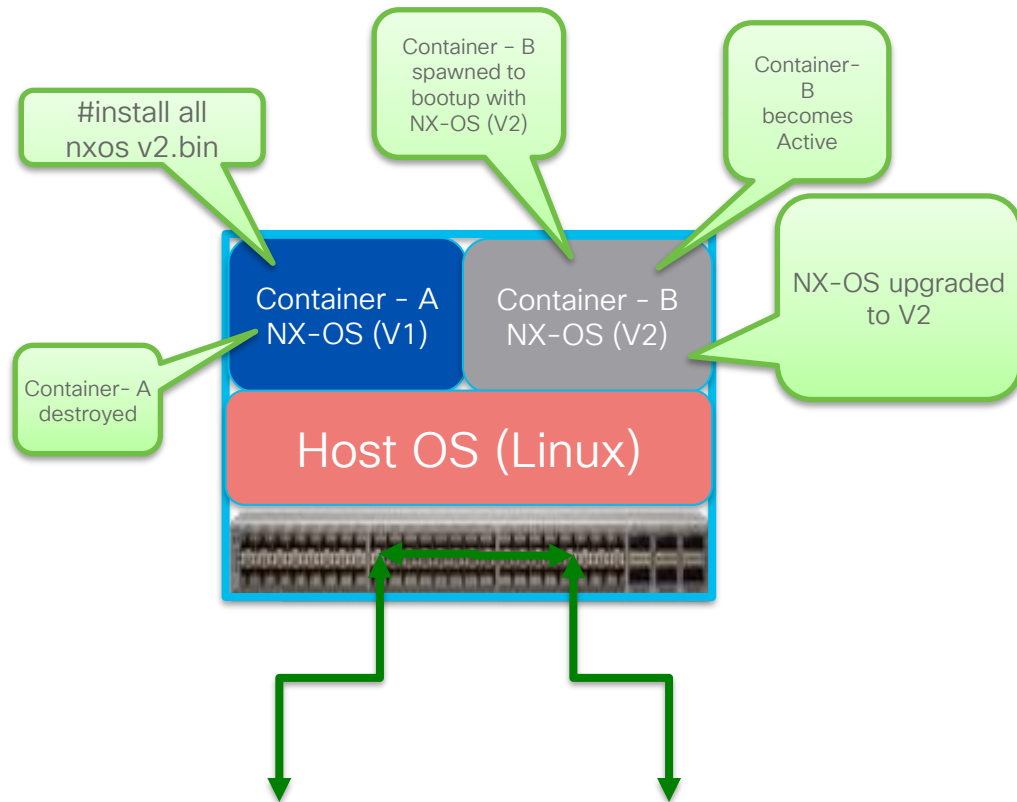
```
install all nxos bootflash:filename non-disruptive
```

Enhanced ISSU (LXC)



Enhanced ISSU on the Nexus ToR Switch

- NX-OS runs on container
- Dual container during upgrade



Activate Enhanced ISSU

```
boot mode lxc
```

Show Module output

LXC enabled

```
tor3-hw-7# show module
```

Mod	Ports	Module-Type	Model	Status
1	54	48x10M/100M/1G + 4x10/25G + 2x40/100G Eth Mod	N9K-C9348GC-FXP	ok
27	0	Virtual Supervisor Module	N9K-C9348GC-FXP	active *

Mod	Sw	Hw	Slot
1	10.3(2) IOV9(0.198)	0.1080	NA
27	10.3(2) IOV9(0.198)	0.1080	VSUP

Mod	MAC-Address(es)	Serial-Num
1	2c-33-11-ec-5f-16 to 2c-33-11-ec-5f-5d	FDO21100DTW
27	2c-33-11-ec-5f-16 to 2c-33-11-ec-5f-5d	FDO21100DTW

Mod	Online Diag Status
1	Pass
27	Pass

Putting it all together

	Critical Bug Fix & PSIRT	Hardware Upgrade	Software Upgrade
SMU	Recommended	Not Recommended	Not Recommended
GIR	Not Recommended	Recommended	Not Recommended
ISSU	Possible	Not Recommended	Recommended

Simplifying Upgrade Process



For Your
Reference

Non-Disruptive NXOS
upgrade

Enhanced ISSU the default boot mode on
FX3 and GX platforms - v10.3(3)
9300-FX, 9300-FX2, 9408 platforms v10.4(1)

Reducing Number of
reloads required during
upgrade

Combining NXOS and EPLD upgrade avoids
additional reload in case of EPLD - v10.1(2)

Combining SMU with NX-OS Upgrade -
v9.3(9)

Prevent downtime for
Kernel Patch

Providing the ability to perform non-disruptive
upgrade without reload due to kernel patch -
v10.2(3)

Journey to Mastery

Service-Level HA

NX-OS Architecture
Process Restartability
HA Policy

System-Level HA

Physical Redundancy
Upgrades & Patching

Network-Level HA

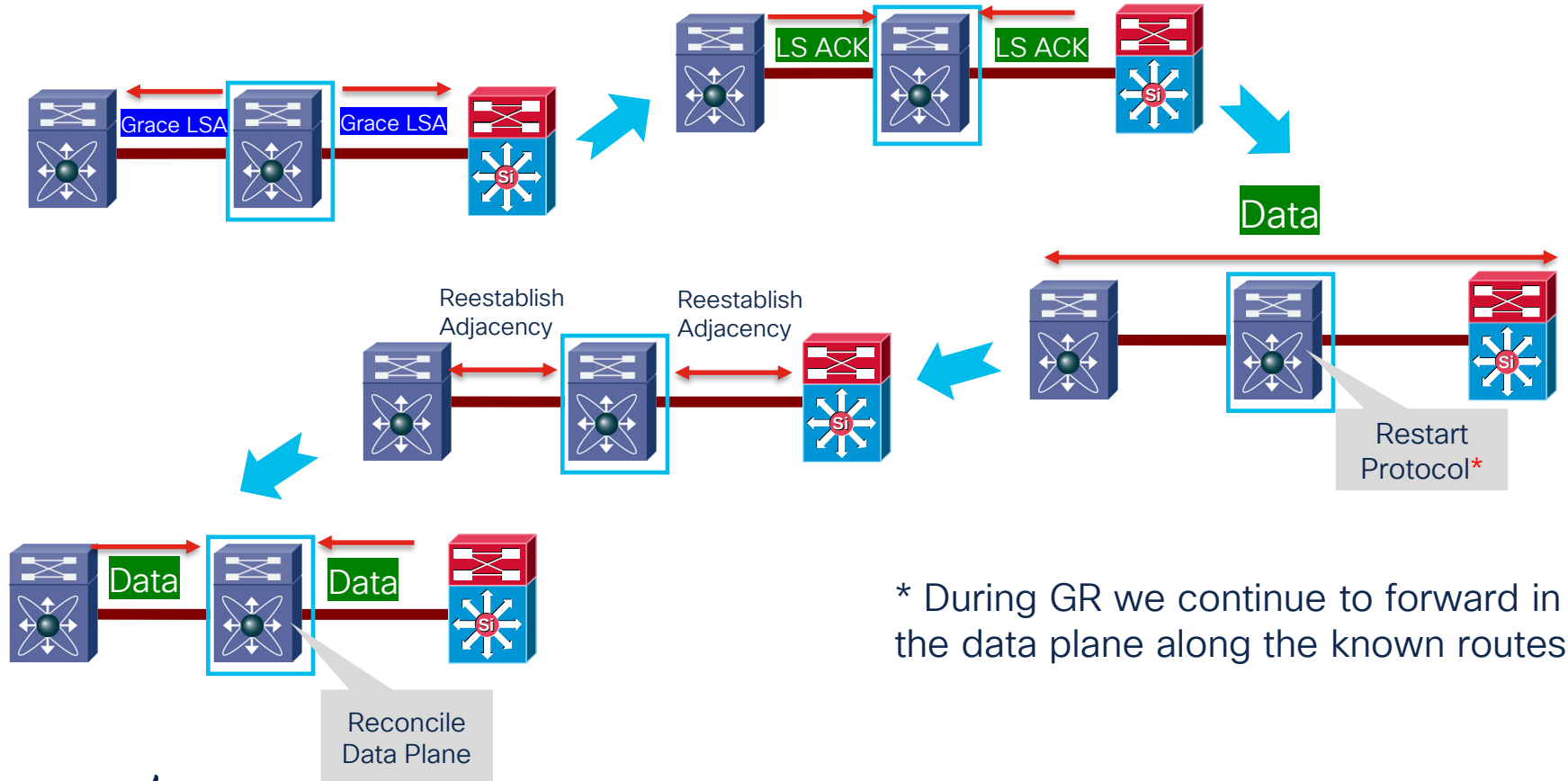
Layer 2 / 3 Redundancy
DC HA Architecture



Graceful Restart Non Stop Forwarding



GR/NSF NXOS Operations



* During GR we continue to forward in the data plane along the known routes

Graceful Restart / Non Stop Forwarding

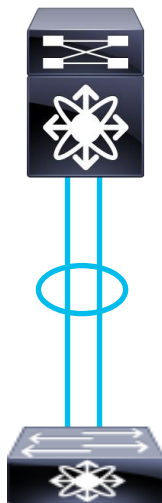
- Enabled by default
- IETF compliant
- Supported routing protocols

Routing Protocol	NSF/GR Definition
OSPFv2	RFC 3623
OSPFv3	RFC 5187
EIGRP	Cisco
BGP	RFC 4724
Intermediate System - Intermediate System	RFC 3847
Label Distribution Protocol	RFC 3478

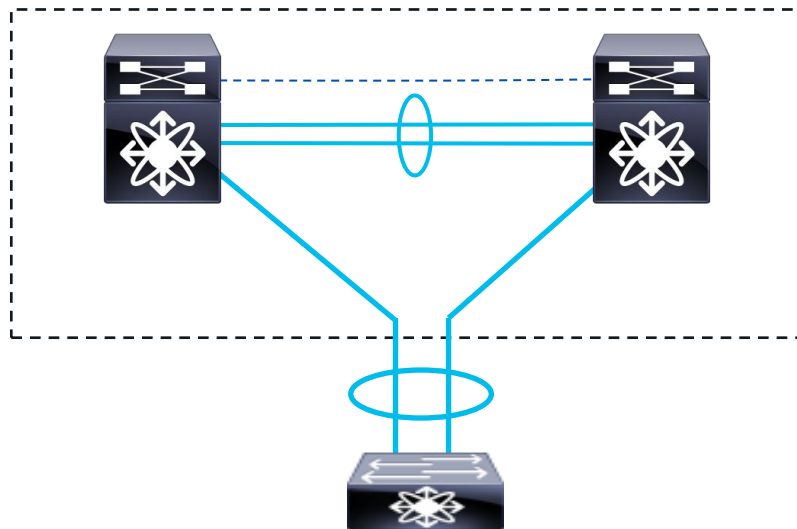
Virtual Port Channel (VPC)



Port Channel



Virtual Port Channel



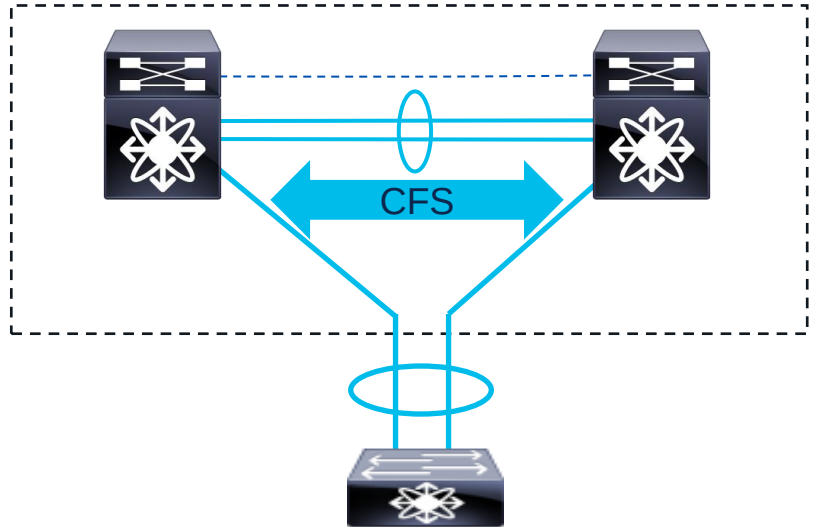
vPC Benefits

- Dual Control Plane
- Active/Active Data Plane

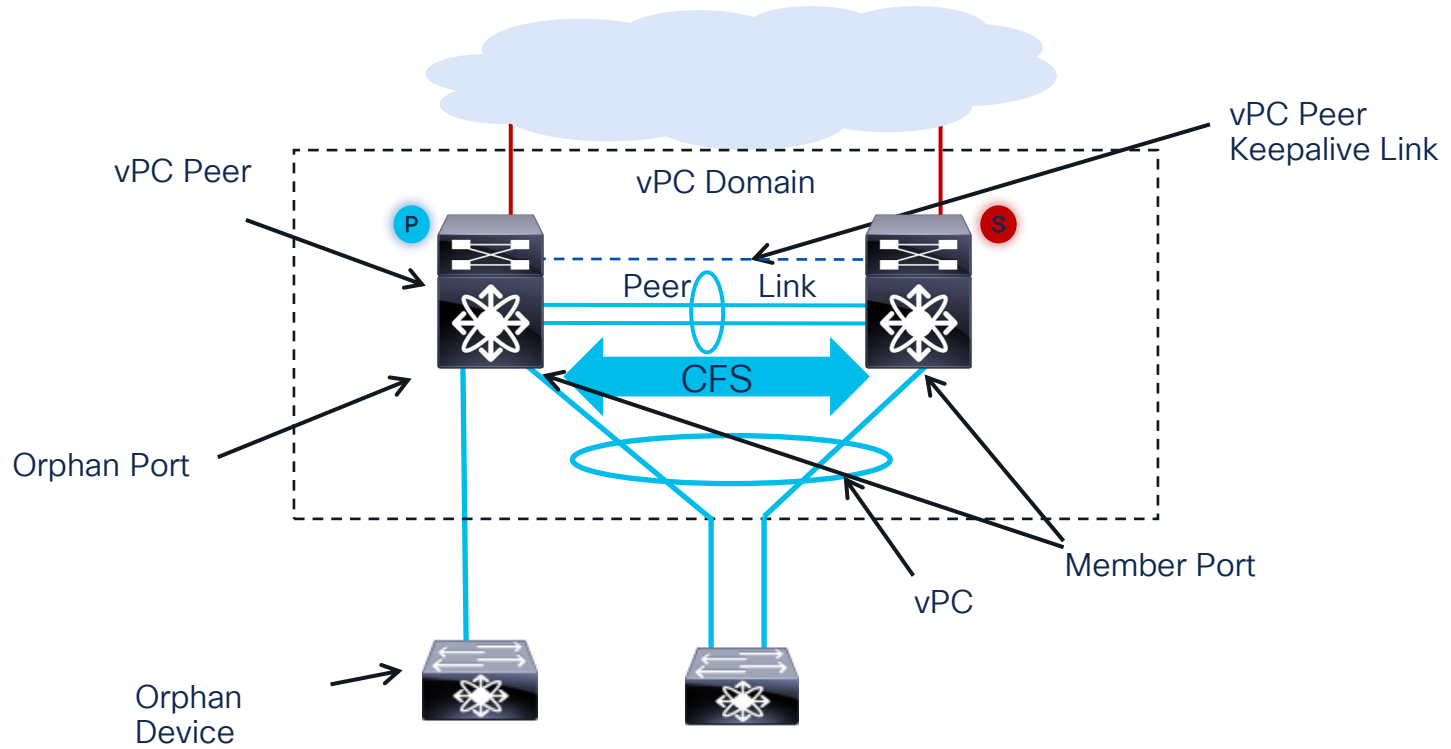
Cisco Fabric Service (CFS)

- Protocol State Synchronization
- System Configuration

Virtual Port Channel



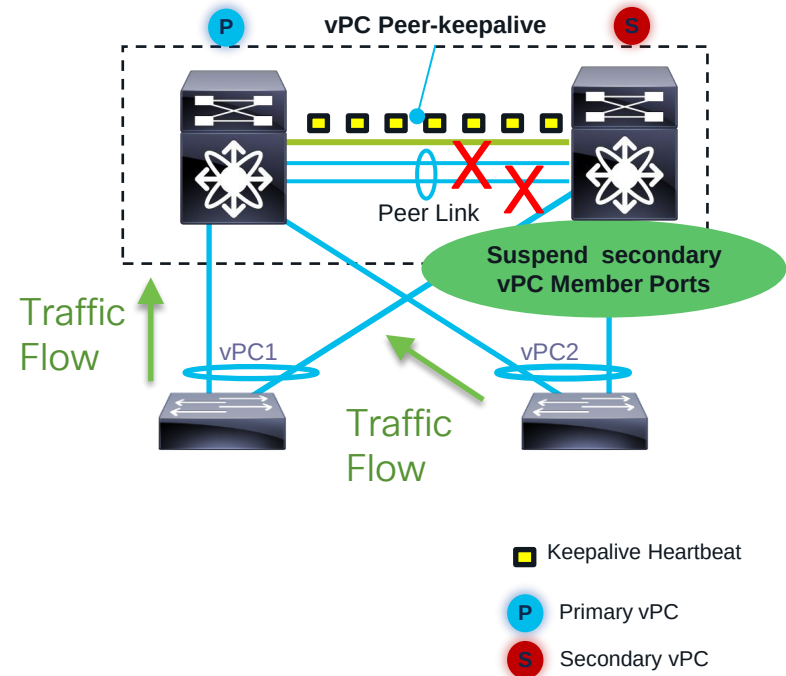
vPC Terminology



vPC Failure Scenario 1

vPC Peer-Link down

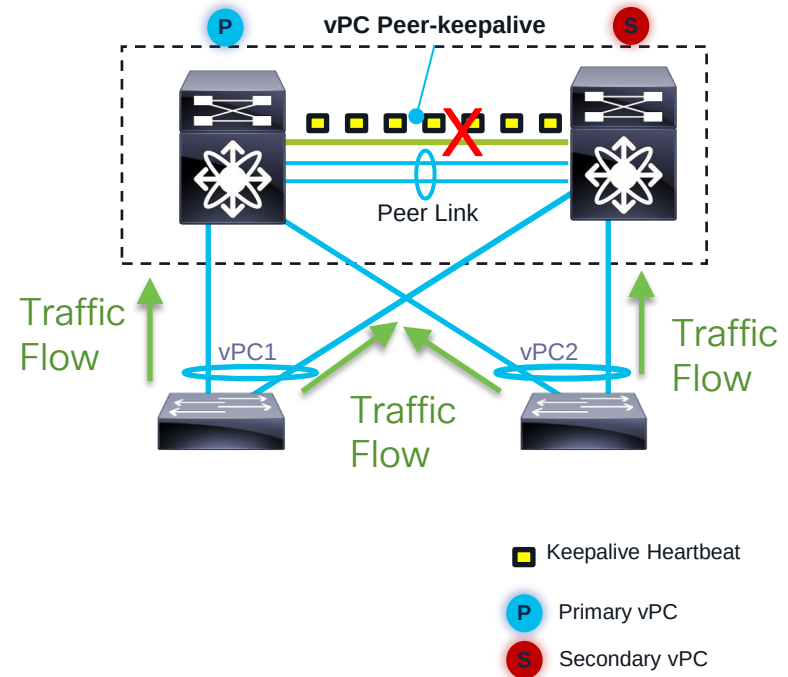
- vPC peer-link failure (link loss)
- vPC peer-keepalive up
- Status of other vPC peer known
- Secondary vPC peer disables all vPC's
- Traffic forwarded to vPC primary



vPC Failure Scenario 2

vPC Peer KeepAlive Link down

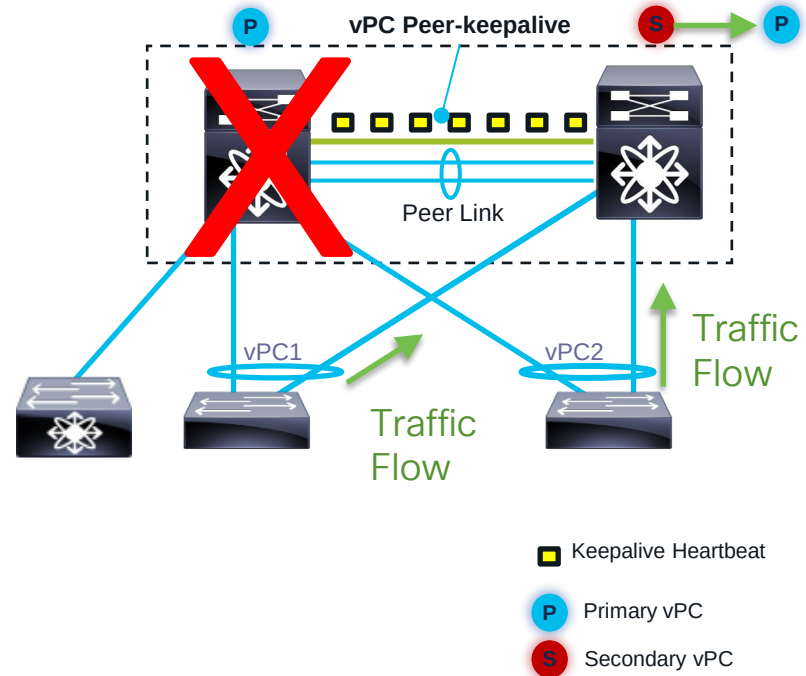
- vPC peer-link is up
- vPC peer-keepalive goes down
- Status of other vPC peer known
- No Impact to the system
- Traffic continues to flow normally



vPC Failure Scenario 3

Primary Peer Fails

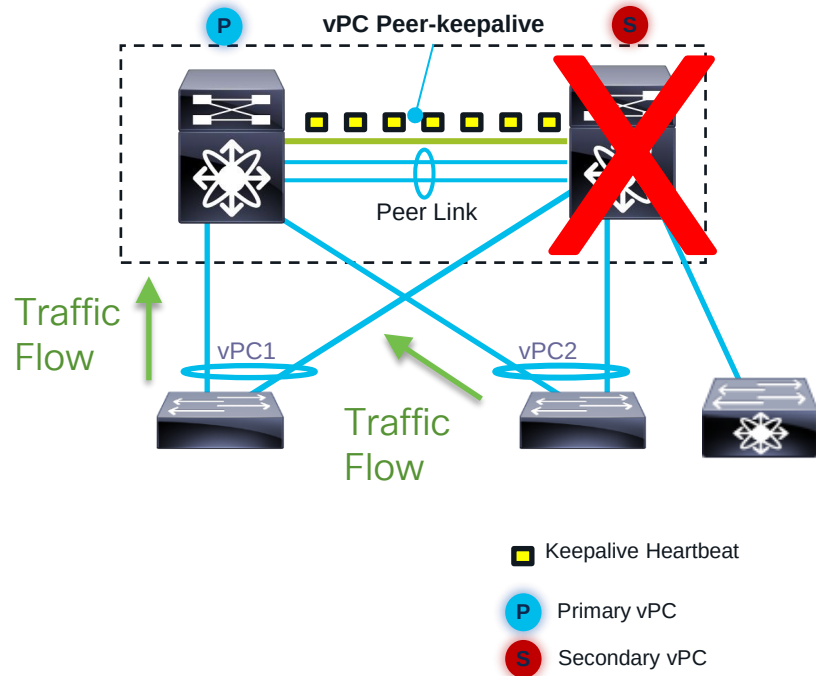
- Primary Peer goes down
- Secondary Detects via Keep alive link
- Secondary Takes on Primary Role
- vPC Traffic will be forwarded to new vPC primary
- Orphan device (to Primary) traffic will be impacted



vPC Failure Scenario 4

Secondary Peer Fails

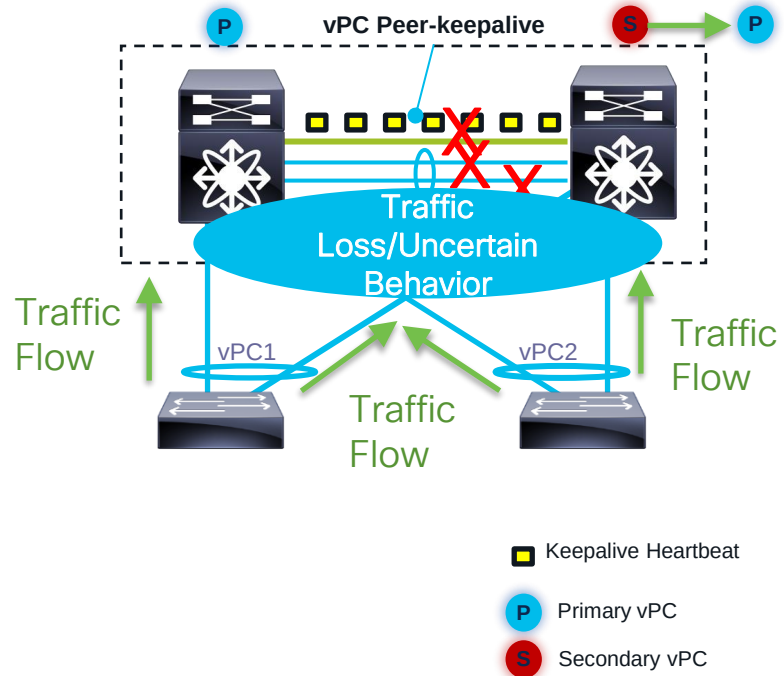
- Secondary Peer goes down
- Primary Detects via Keep alive link
- Traffic will be forwarded to vPC primary
- Orphan device (to Secondary) traffic will be impacted



vPC Failure Scenario 5

Dual Active or Split Brain

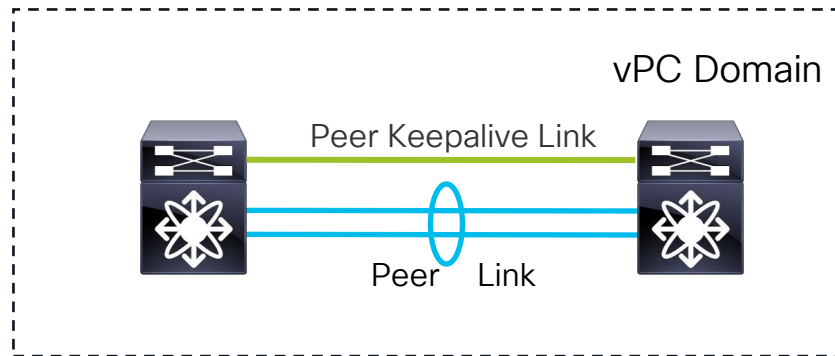
- vPC peer-link goes down
- vPC peer-keepalive goes down
- Status of other vPC peer is **UNKNOWN**
- Secondary also becomes Primary
- Traffic Pattern is Uncertain



vPC Control Plane Recommendations

vPC Peer Link & Peer Keepalive

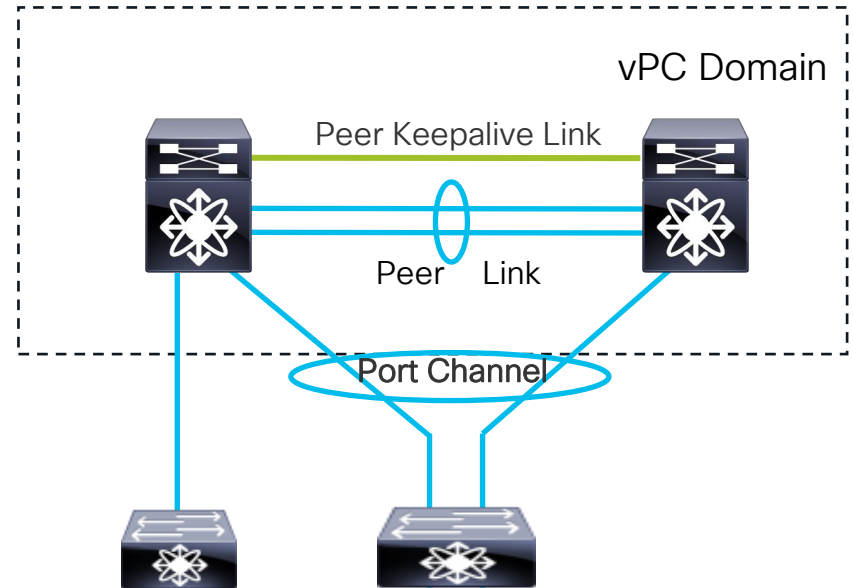
- Peer links should be **point-to-point** connection
- For the Peer Link, recommendation is to use two 10GbE ports for resilient connectivity
- The Peer Keepalive traffic should be over a **separate** link and not over the Peer Link



vPC Member connectivity Recommendations

Dual Home Devices

- Make sure to dual attach all devices into the vPC Domain
- It is best to use LACP for the Port Channels of vPC Member ports
- The configuration must match on both switches
- Use the same value for the port channel ID and the vPC ID



Journey to Mastery

Service-Level HA

NX-OS Architecture
Process Restartability
HA Policy

System-Level HA

Physical Redundancy
Upgrades & Patching

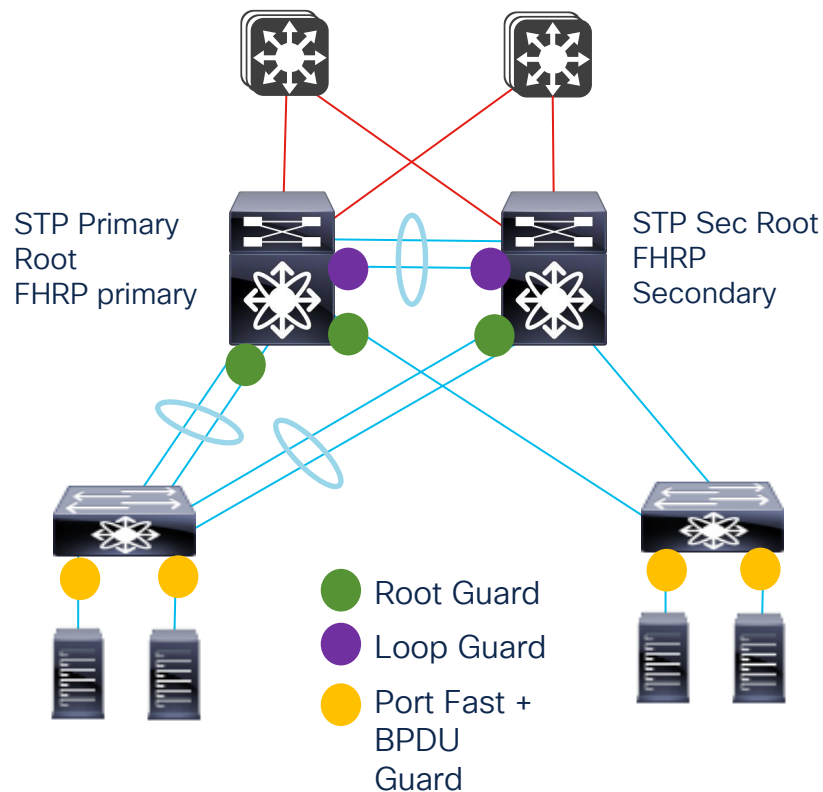
Network-Level HA

Layer 2 / 3 Redundancy
DC HA Architecture



Traditional Data Center Design and Tradeoff

- Well-known placement of STP root and secondary root switches
- Well-known placement of (FHRP) primary and secondary devices
- Forwarding and blocking links from the access layer to the aggregation layer
- Various hardening features to guarantee deterministic spanning-tree behavior



The Spine-Leaf Architecture

Enhanced Performance

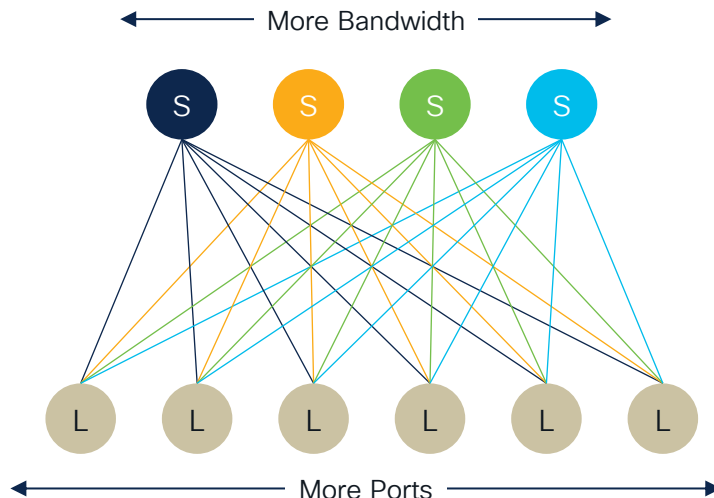
- Equidistant endpoints
- Non-blocking, low-latency core

Scalability

- New leaf switch adds ports
- New spine switch adds bandwidth

Redundancy

- Built-in redundancy and fault tolerance
- ECMP (Equal Cost Multi Path) routing



No STP



No Root Guard



No BPDU

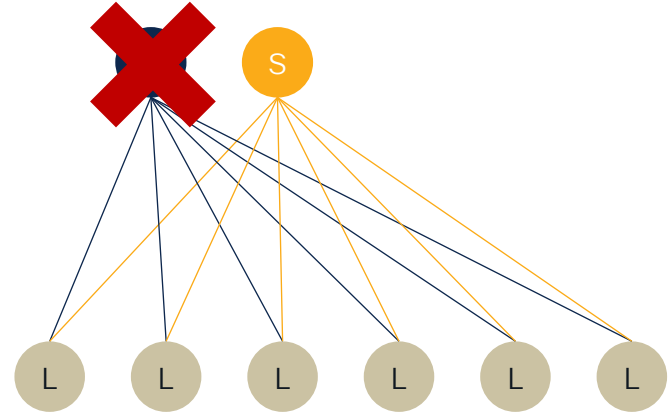


No Loop Guard

The Spine-Leaf Architecture

Availability

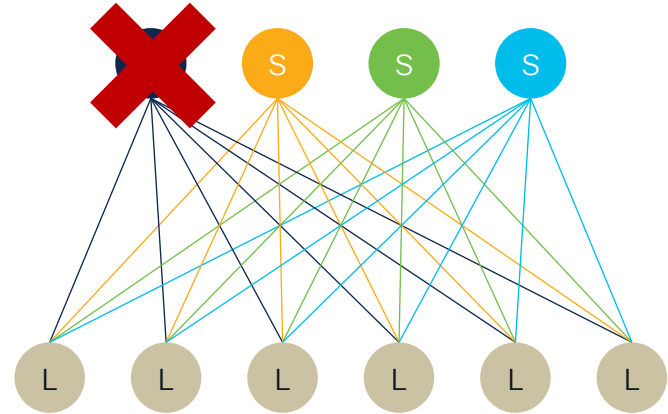
- N+1 Redundancy
- On Spine failure
 - 2 Spine = 50% impact



The Spine-Leaf Architecture

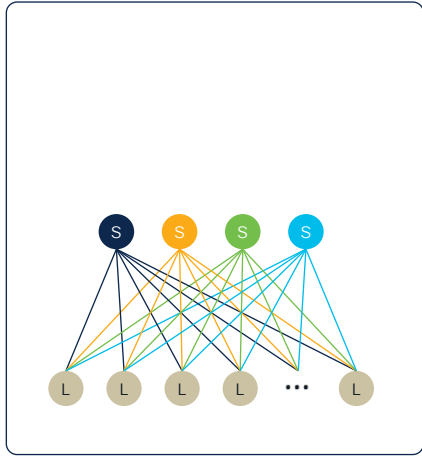
Enhanced Availability

- N+1 Redundancy
- On Spine failure
 - 2 Spine = 50% impact
 - 4 Spine = 25% impact
 - 8 Spine = 12.5% impact



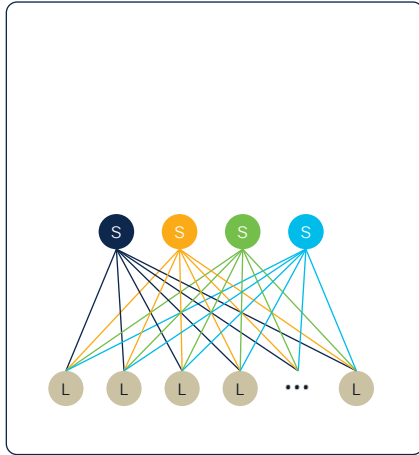
Redundancy increases by building out the topology

Evolution of DC Network Architecture

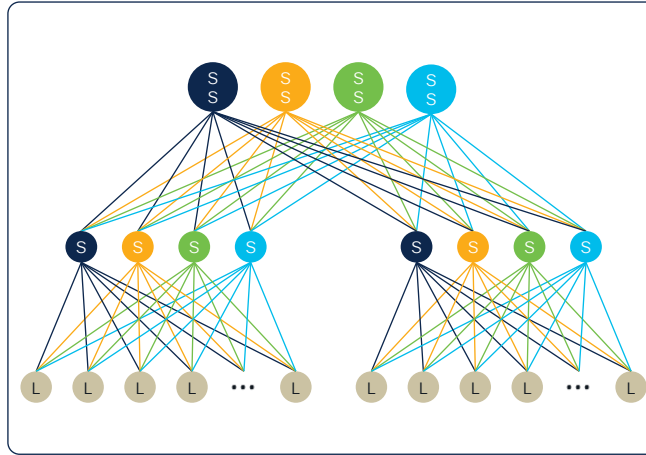


2 Tier Leaf-Spine

Evolution of DC Network Architecture

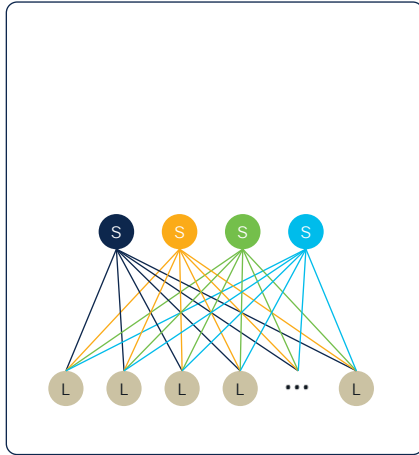


2 Tier Leaf-Spine

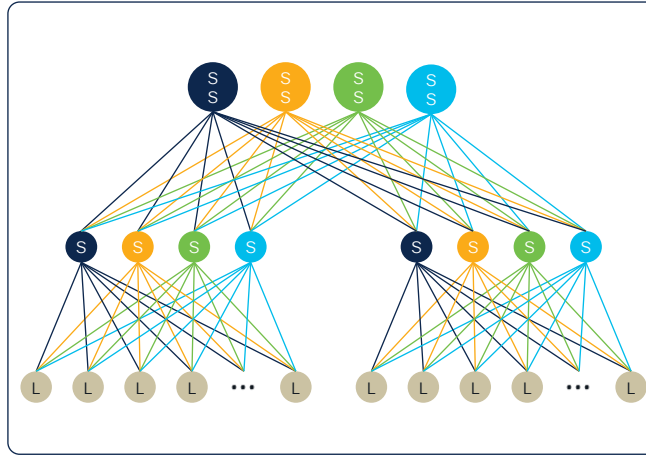


3 Tier Leaf-Spine-SuperSpine

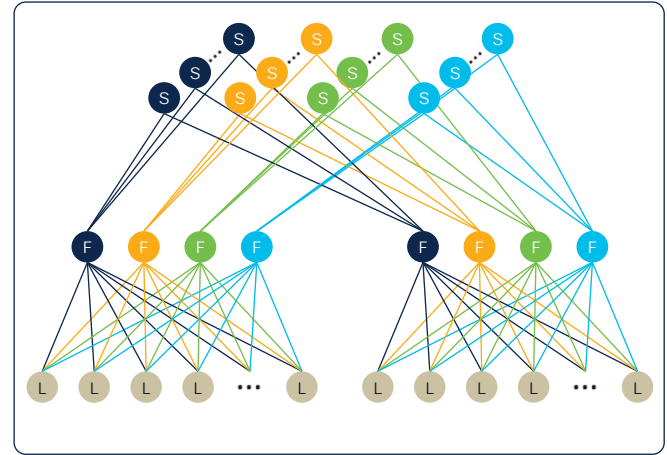
Evolution of DC Network Architecture



2 Tier Leaf-Spine



3 Tier Leaf-Spine-SuperSpine



3 Tier Leaf-Fabric-Spine

What we learned from the Cloud Titans

Building Scalable and Resilient Data Center Networks

#1

Simplicity is Key
Simple Design Principals

#2

Scale as you Go
Scale is Never Finite

#3

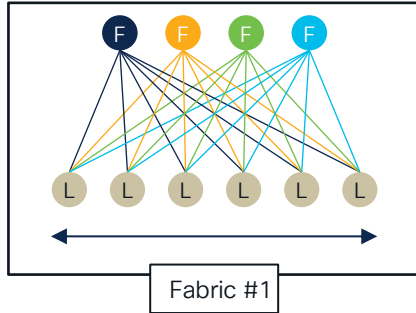
Fail but Fail Fast
Reduce Brown-Out Exposure

#4

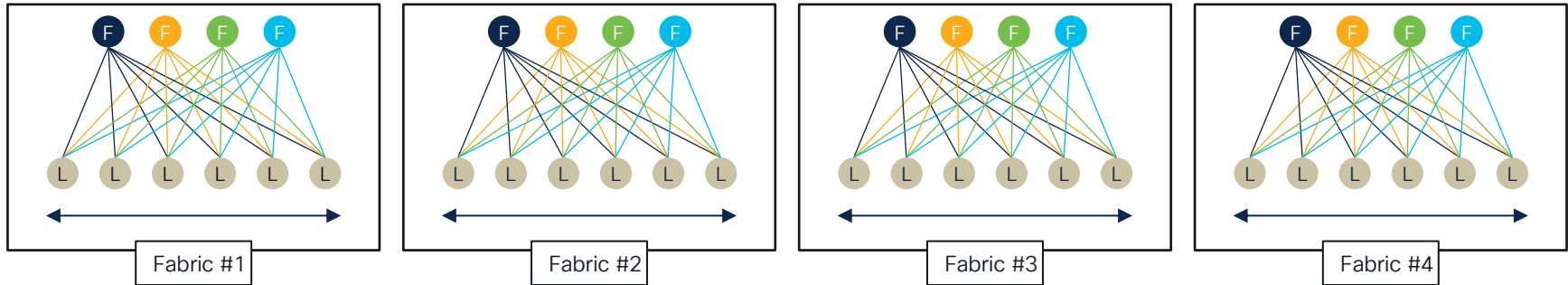
Redundant and Repeatable
Risk is Never an Option

Don't Build Fabric for Maximum Leaf

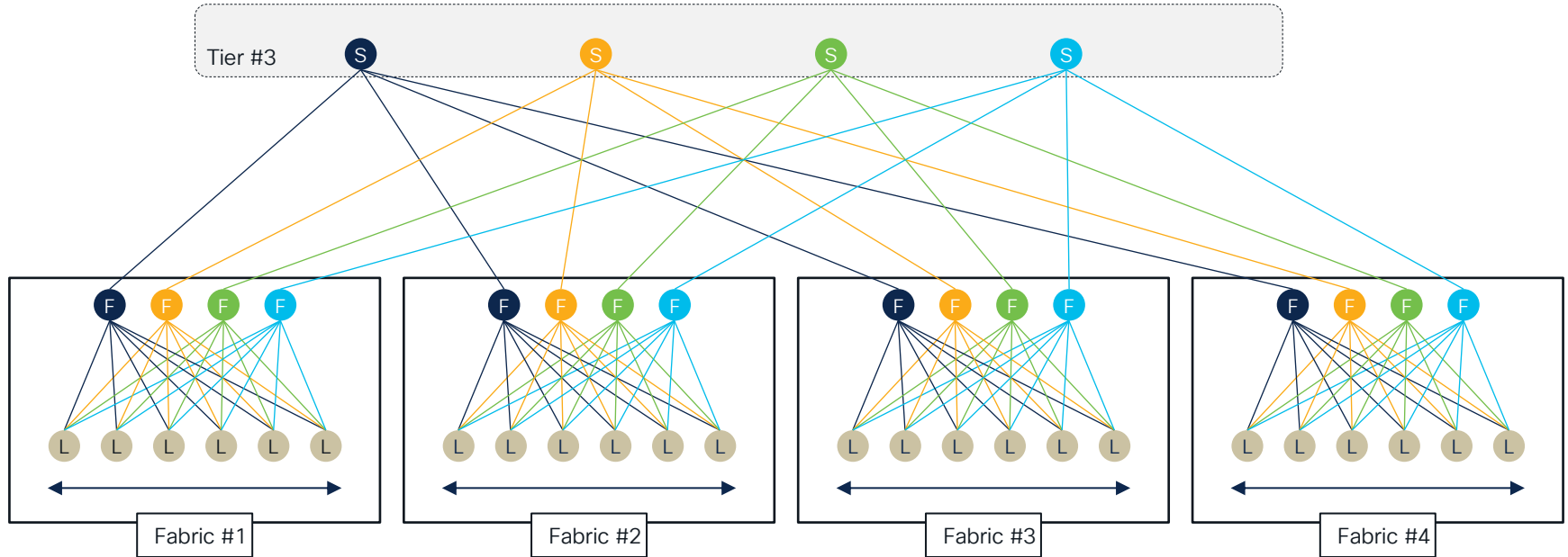
Step back in our Design



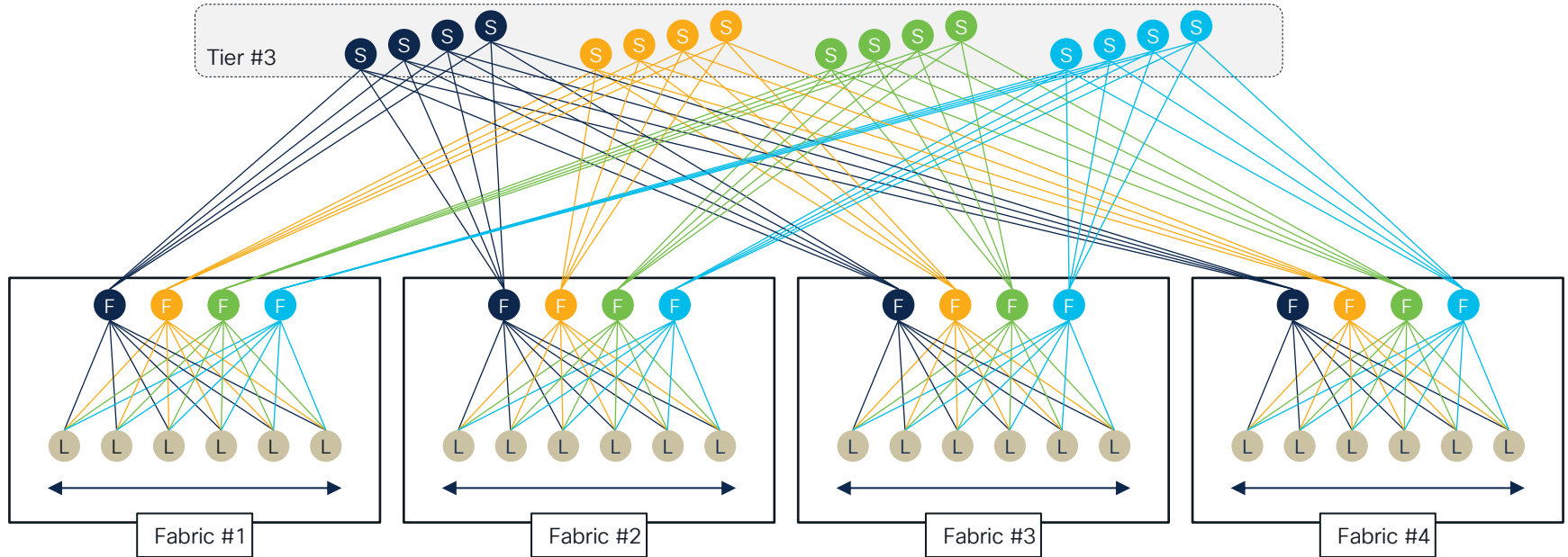
Repeat for Host Port Scale (Scale Out)



Designing Tier #3



Scaling Tier #3



Journey to Mastery

Service-Level HA

NX-OS Architecture
Process Restartability
HA Infrastructure

System-Level HA

Physical Redundancy
Upgrades & Patching

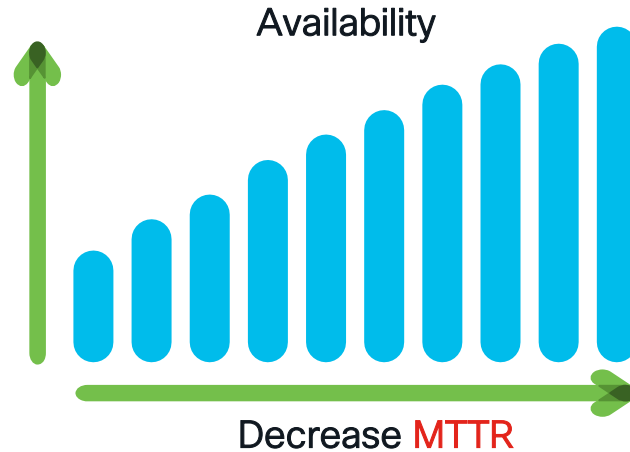
Network-Level HA

Layer 2 / 3 Redundancy
DC HA Architecture



Key Takeaways

- Leverage NX-OS Architecture Feature
 - HA Infrastructure
 - Process Restartability
 - Service HA Policy
- Reduce single point of failure
 - Physical Redundancy
 - System-wide



- Best practice setup for Layer 2/3 HA features e.g. vPC, NSF
- Utilize Enhanced ISSU, GIR for upgrades, maintenance
- Utilize inherent redundancy in spine-leaf architecture
- Start Small and grow as needed

Complete Your Session Evaluations



Complete a minimum of 4 session surveys and the Overall Event Survey to claim a **Cisco Live T-Shirt**.



Complete your surveys in the **Cisco Live mobile app**.



Continue your education



- Visit the On-Demand Library for more sessions at www.CiscoLive.com/on-demand
- Cisco NX-OS High Availability and Redundancy Guide
<https://www.cisco.com/c/en/us/td/docs/dcn/nx-os/nexus9000/105x/configuration/high-availability/cisco-nexus-9000-series-nx-os-high-availability-and-redundancy-guide-105x.html>
- Cisco NX-OS System Management Configuration Guide
<https://www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus9000/sw/105x/config-guides/cisco-nexus-9000-series-nx-os-system-management-configuration-guide-release-105x.html>
- Cisco Nexus Switches Install and Upgrade Guide
<https://www.cisco.com/c/en/us/support/switches/nexus-9000-series-switches/products-installation-guides-list.html>
- NX-OS Modular Package Management
https://www.cisco.com/c/en/us/td/docs/dcn/nx-os/nexus9000/105x/upgrade/cisco-nexus-9000-series-nx-os-software-upgrade-and-downgrade-guide-105x/m_nx-os_optionality_9000_101x.html

Contact me at: moybiswa@cisco.com



Thank you

CISCO *Live!*

#CiscoLiveAPJC

CISCO *Live!*

GO BEYOND

#CiscoLiveAPJC