Инструменты для обеспечения высокой доступности, отказоустойчивости и резервирования в LAN сети. Часть 2

Юрий Дышлевой
Системный инженер, CCIE

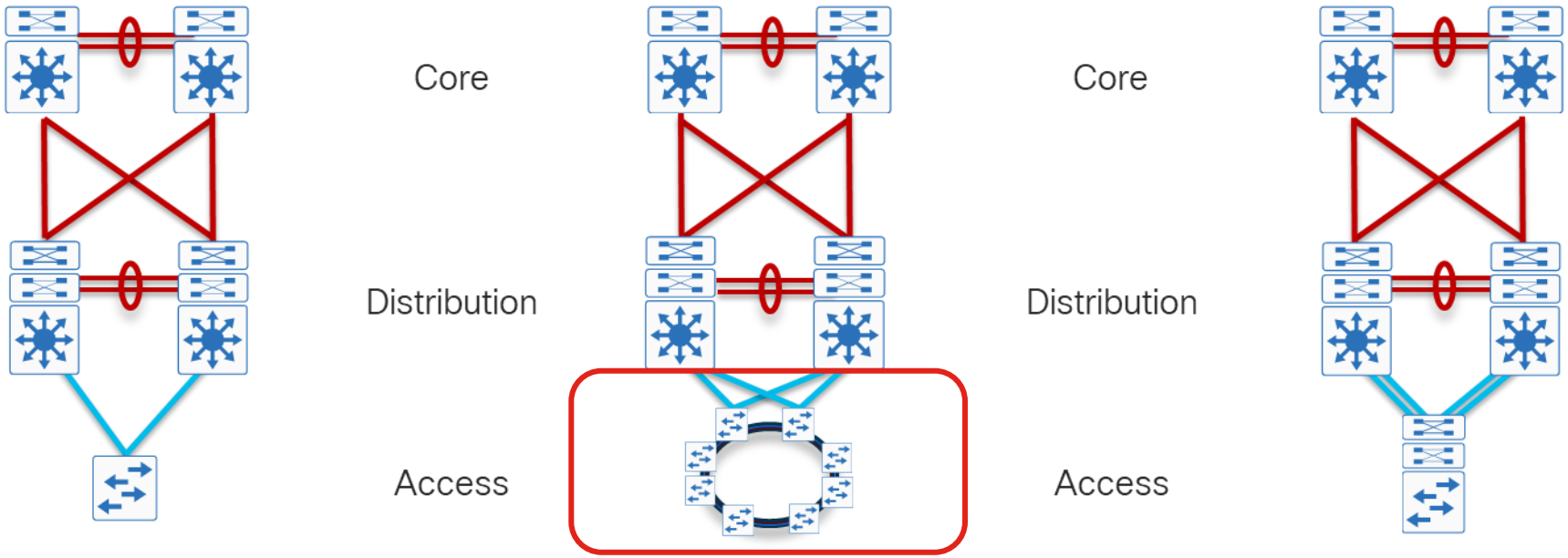
23.03.2021

Agenda

- **Specific Use-Cases**
- Wired campus platform hardware and software features for HA
- Summary and conclusions

Specific Use-Cases

Structured campus network design

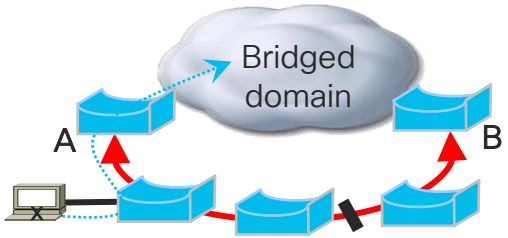


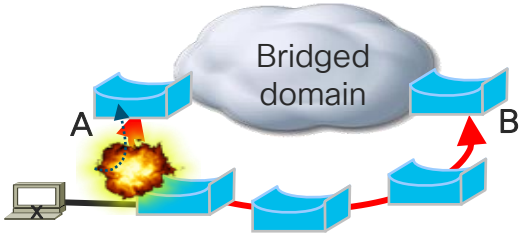
REP (Resilient Ethernet Protocol)

- Overview of REP
  - You can define a "segment" that configures consistent REP settings across multiple switches, and you can specify any protected link or block port (forwarding stop port) in that segment.
  - Supported on a large range of Cisco products.
  - Co-existence with Spanning Tree (TCN from REP to STP)
  - Very easy to configure and troubleshoot
- High-speed switching
  - High-speed switching of up to 50 milliseconds is possible. (About 50 ms to 200 ms)
  - High-speed restoration. Tested on segments with 32 switches.
- Ring independent redundancy
  - It manages path protection and supports complex configurations that connect multiple rings.
  - Failures can be localized and distributed by detecting failures in ring units.
  - Unaffected by other Ring failures.
- Optimal route design
  - Ring network route specification and redundant route design are possible.
  - Flexible bandwidth expansion is possible with EtherChannel.
  - Optimal bandwidth utilization (VLAN Load balancing)

REP configurations

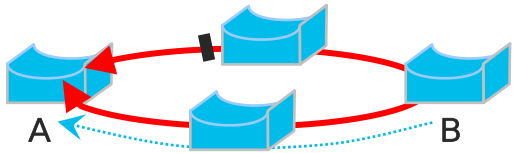
- Segment

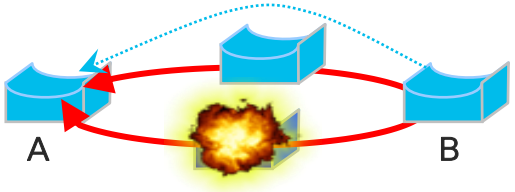




The REP segment provides redundant routes like A or B to other L2 networks.
It solves the block state at the time of failure and provides a redundant route.

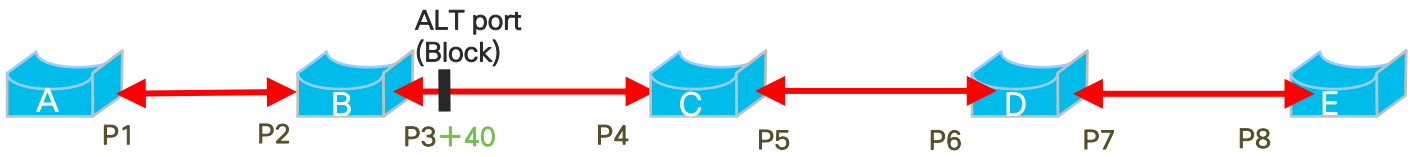
- REP Ring





When configured in a ring, the REP segment provides redundant connectivity between the two switches.
Various networks can be configured by combining rings and segments.

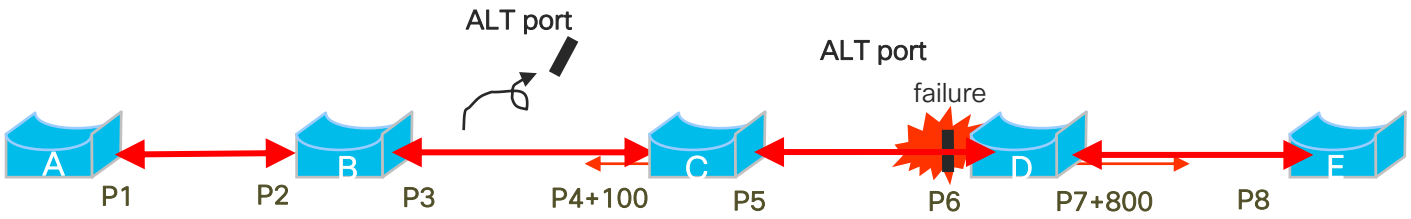
REP segment protocol overview



Each port is configured as part of one segment ID.

Through the Segment REP is Enable on all links in the Segment

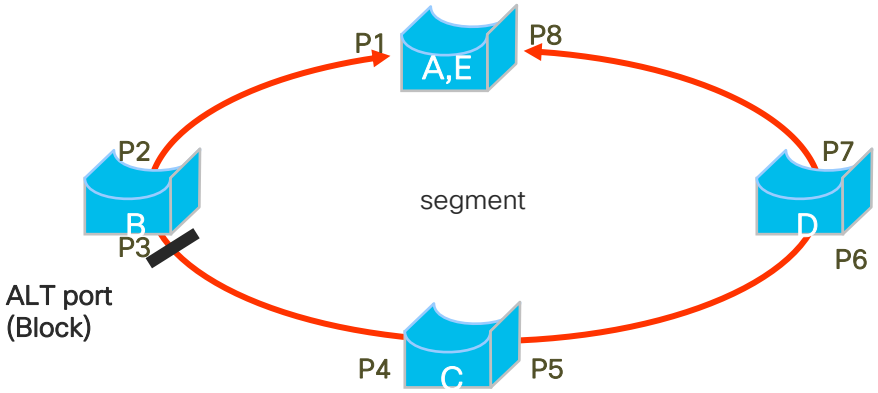
REP Edge Determine the ALT Port (Block) between A” and E” (any part).



If a failure occurs during the REP segment, the blocking port initiates data forwarding.

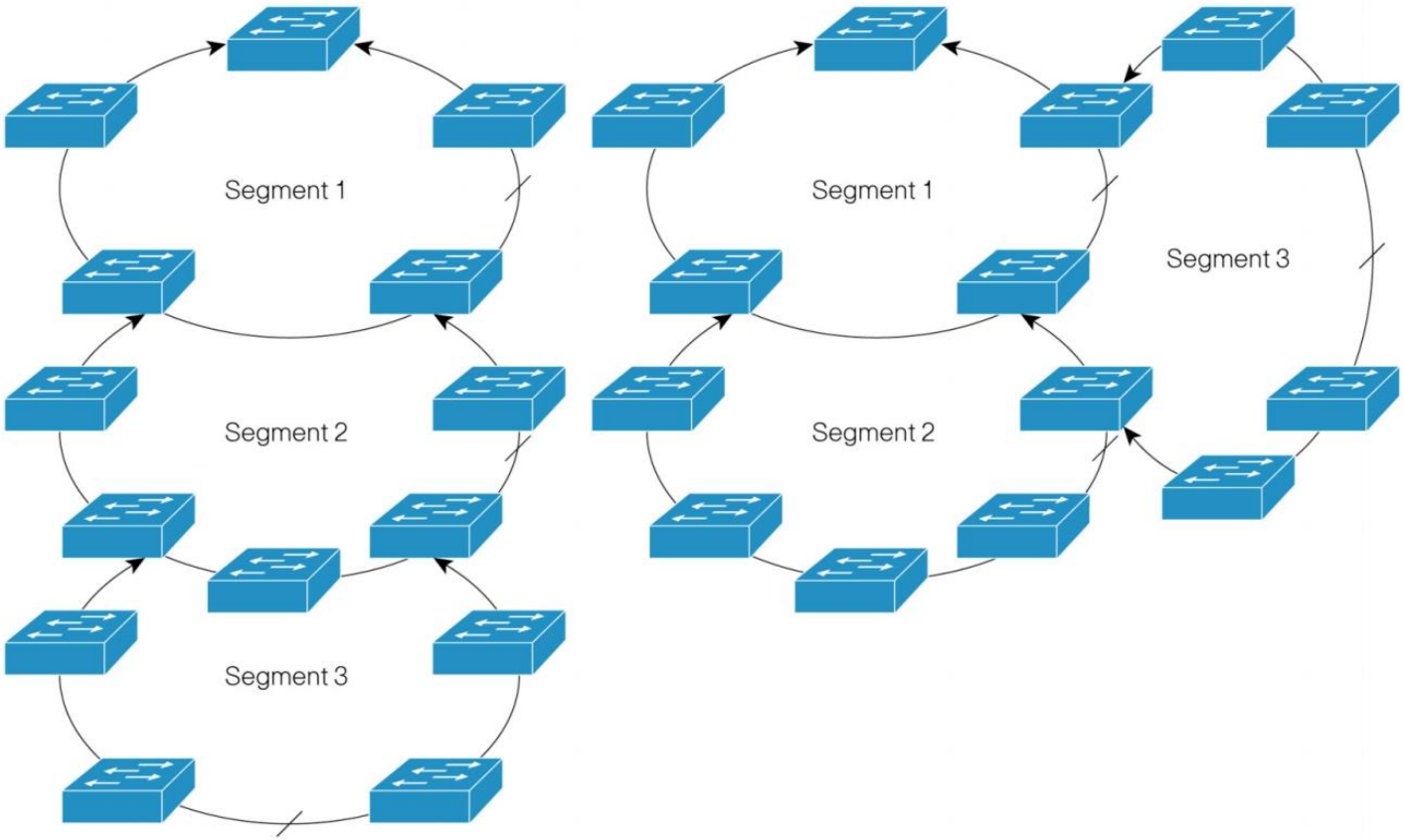
REP segment protocol overview

Ring Topology using REP segment protocol



The REP Edge Port can be expropriated into one Node.

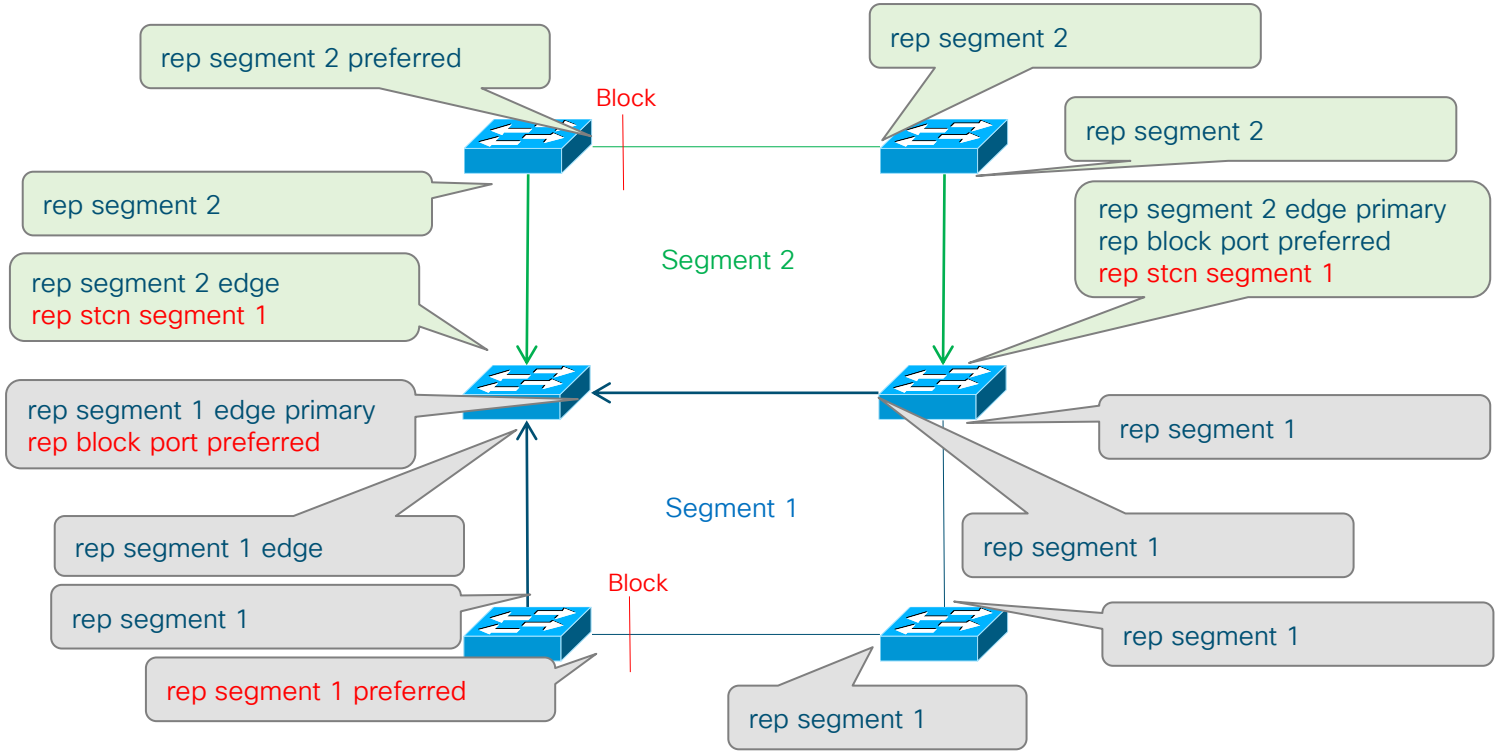
REP segment protocol overview



REP setting example

* One block port (Alt, Failed) per segment

- rep stcn segment 1: notify segment 2 failure to segment 1
- rep block port preferred : setting to specify the block port as Static (set on the primary edge)
- rep segment 1 preferred : designated as a block port for segment 1



REP setting confirmation

```
C9300-1#show rep topology
REP Segment 1
BridgeName                    PortName   Edge Role
------------------------------- ---------- ---- ----
C9300-1                       Gi1/0/3    Pri* Open
C9300-1                       Gi1/0/2         Open
C9300-1                       Gi1/0/2    Sec  Alt

C9300-1#show rep topology detail
REP Segment 1
C9300-1, Gi1/0/3 (Primary Edge No-Neighbor)
  Open Port, all vlans forwarding
  Bridge MAC: 701f.5301.2c80
  Port Number: 003
  Port Priority: 000
  Neighbor Number: 1 / [-3]
C9300-1, Gi1/0/2 (Intermediate)
  Open Port, all vlans forwarding
  Bridge MAC: 701f.5301.2c80
  Port Number: 002
  Port Priority: 000
  Neighbor Number: 2 / [-2]
C9300-1, Gi1/0/2 (Secondary Edge)
  Alternate Port, some vlans blocked
  Bridge MAC: 70b3.17fa.f100
  Port Number: 002
  Port Priority: 000
  Neighbor Number: 3 / [-1]
```

A list of ports belonging to REP segment 1 is displayed.

You can check more detailed contents.

Cisco Catalyst 9000 series REP

	9200 9200L	9300 9300L	9400	9500
Support	○	○	○	○
License	Network Essentials	Network Essentials	Network Essentials	Network Essentials

REP Fast: Enhancing the Resilient Ethernet Protocol With Beacons White Paper

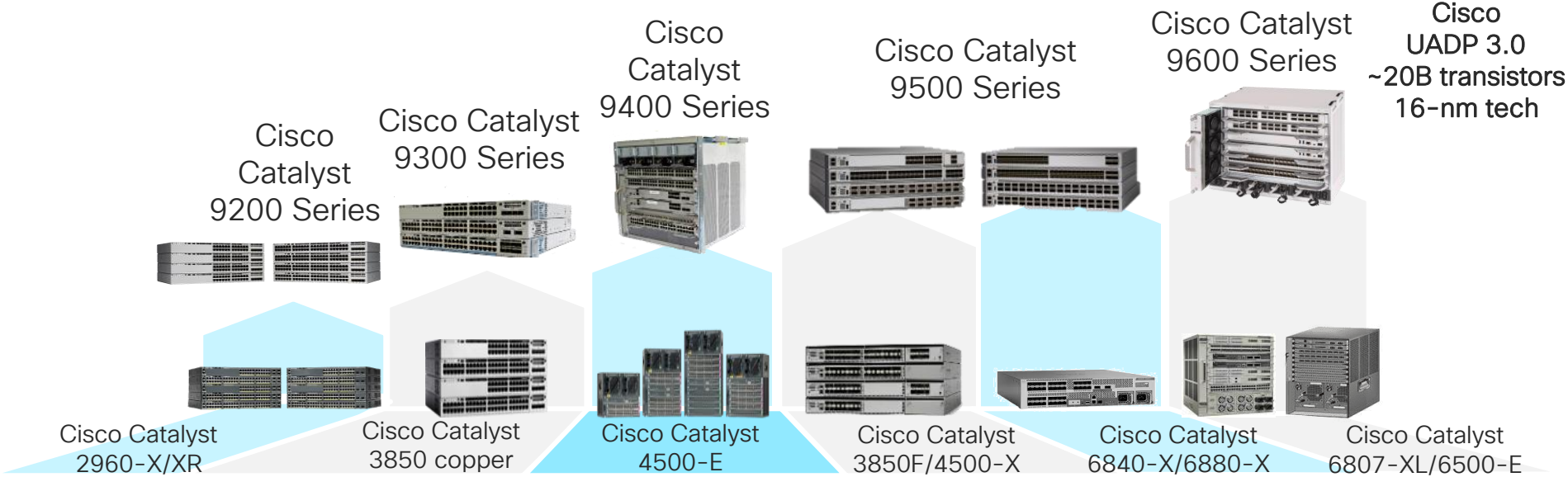
Agenda

- Specific Use-Cases
- **Wired campus platform hardware and software features for HA**
- Summary and conclusions

Cisco Catalyst 9000 Series–switching transitions

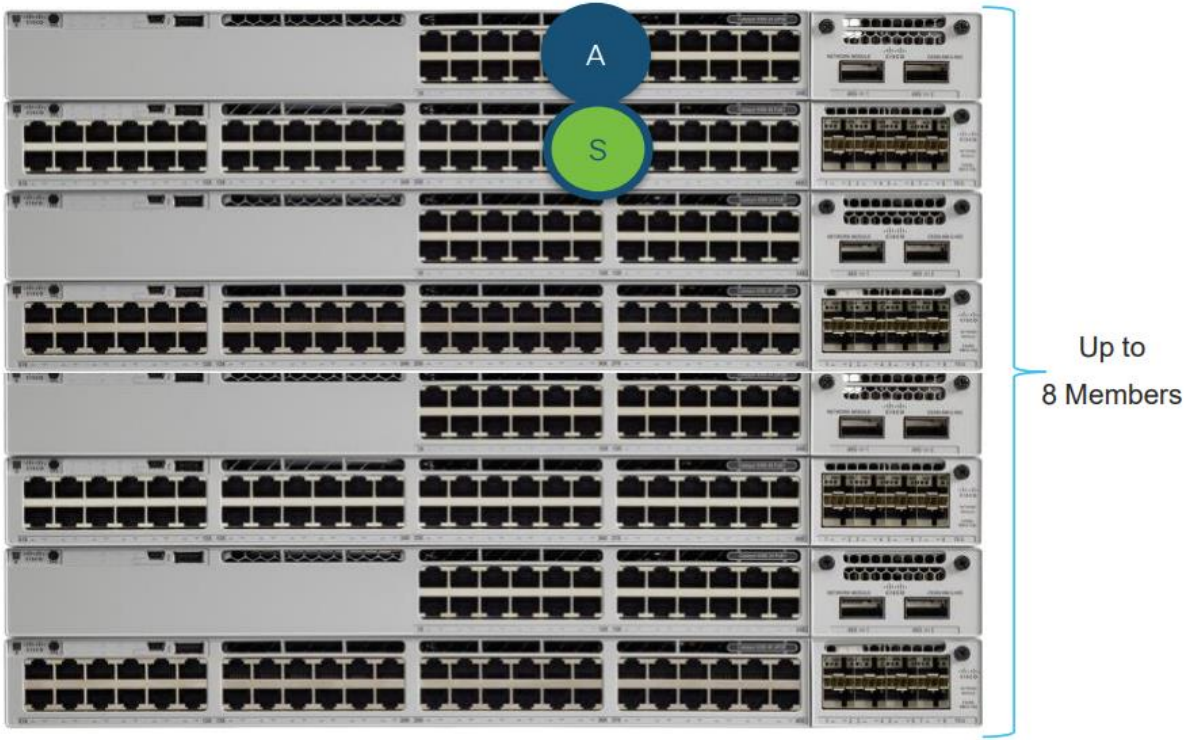
Greater flexibility from small remote site to mission critical campus core.

Cisco UADP 3.0
~20B transistors
16-nm tech



StackWise

High Availability – StackWise



Centralised Control Plane

Distributed Data Plane

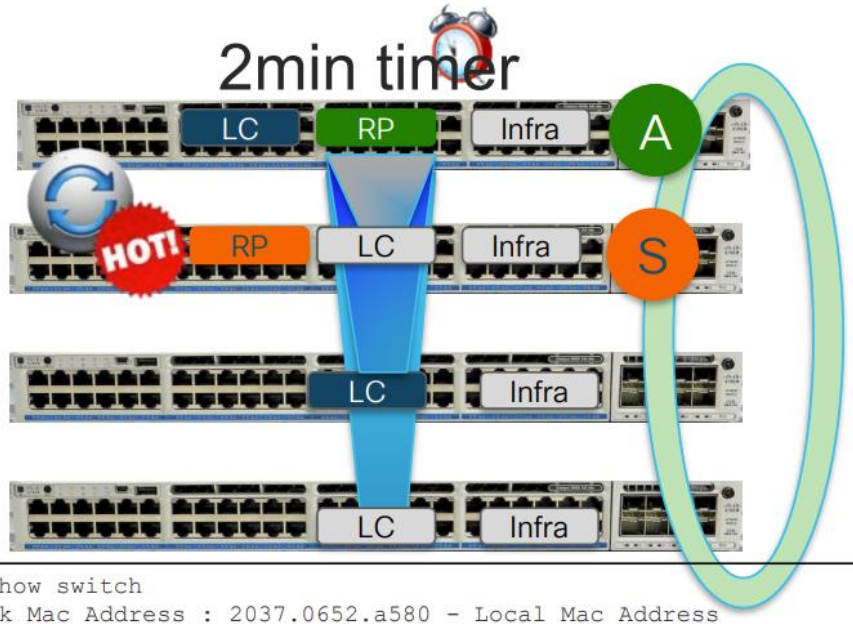
1+1 Stateful Redundancy with Active & Standby

Stateful Switchover SSO/NSF*

* NSF not supported on 9200

Stack initialization

- Active starts RP Domain (IOSd, LC, etc) locally
- Programs hardware on all LC Domains
- Traffic resumes once hardware is programmed
- Starts 2min Timer to elect Standby in parallel
- Active elects Standby
- Standby starts RP Domain locally
- Starts Bulk Sync with Active RP
- Standby reaches “Standby Hot”

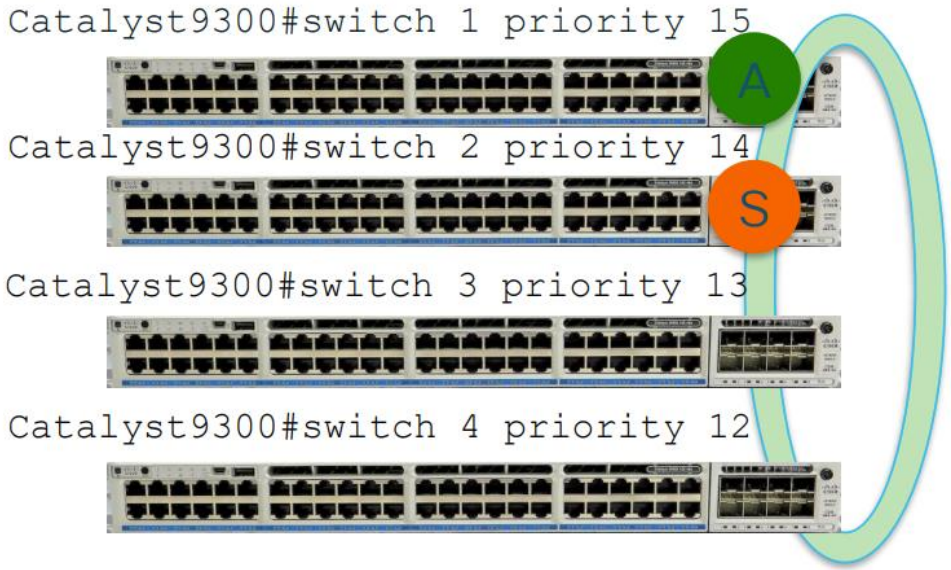



```
%STACKMGR-1-STANDBY_ELECTED: 3 stack-mgr:  Switch 2
has been elected STANDBY.
```

```
GUIDELINE#show switch
Switch/Stack Mac Address : 2037.0652.a580 - Local Mac Address
Mac persistency wait time: Indefinite
                                                    H/W   Current
Switch#   Role     Mac Address      Priority Version  State
-----------------------------------------------------------
 1        Member   2037.0653.ca80     5       P6A     Ready
 2        Standby  2037.0653.db00     10      P6A     HA sync in progress
*3        Active   2037.0652.a580     15      V01     Ready
```

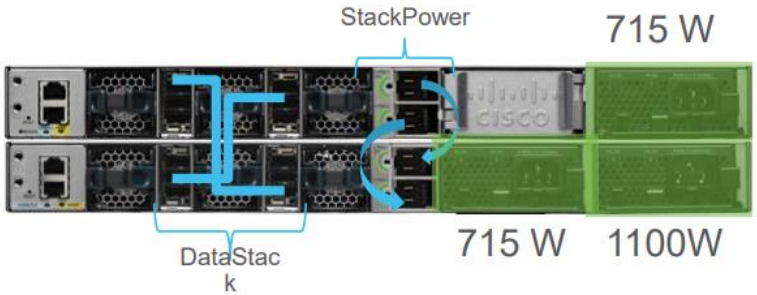
HA Best Practices & Recommendations

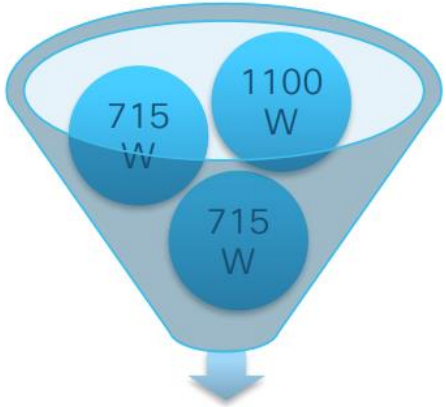
- Power up the first Switch that you want to make it as Active
- Configure Priority of the switch (1-15) – 1 by default – the higher the better
- Power up the second member that you want to make as Standby & then power up rest of the members
- To add a member to an existing stack plug in the stack cable first, then power up the switch
- Avoid stack Merge & Stack split if possible



StackPower

How StackPower Works?



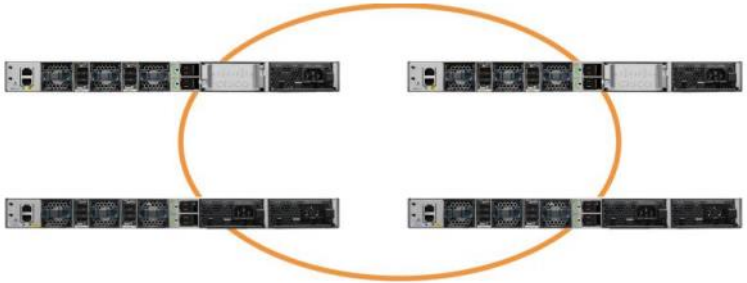


Total Input Power 2530W

- Pools Power from All PS
- All Switches in StackPower share the available Power in Pool
- Each Switch is given their Minimum Power Budget

Power Redundancy Options

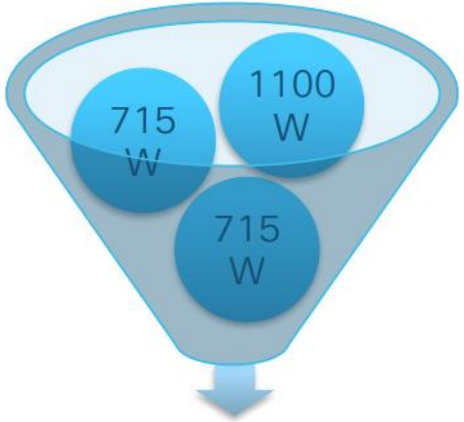
- Zero Footprint RPS OR XPS



StackPower - Zero Footprint RPS

- eXpandable Power System (XPS)

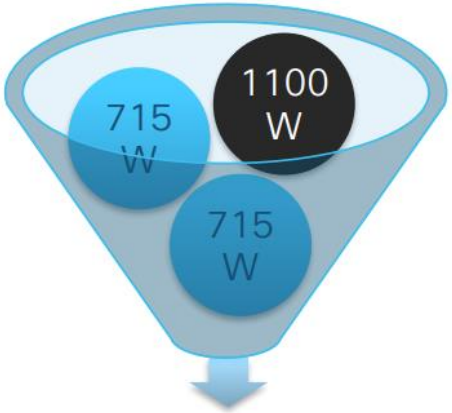
Power Budget Modes



2530W – 30W

Power Sharing Mode

- The Default Mode
- Sum of All PS – 30~60W



1430W – 30W

Redundant Mode

- User Configurable
- Sum of All PS – Largest PS – 30~60W

Global StackPower Reserve = 30W

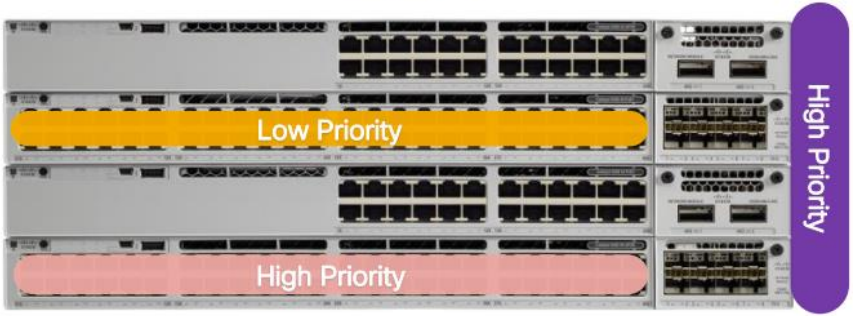
Power Priority

Standalone

Load Shedding Based on configured priority

1. Low Priority Ports
2. High Priority Ports

Stack

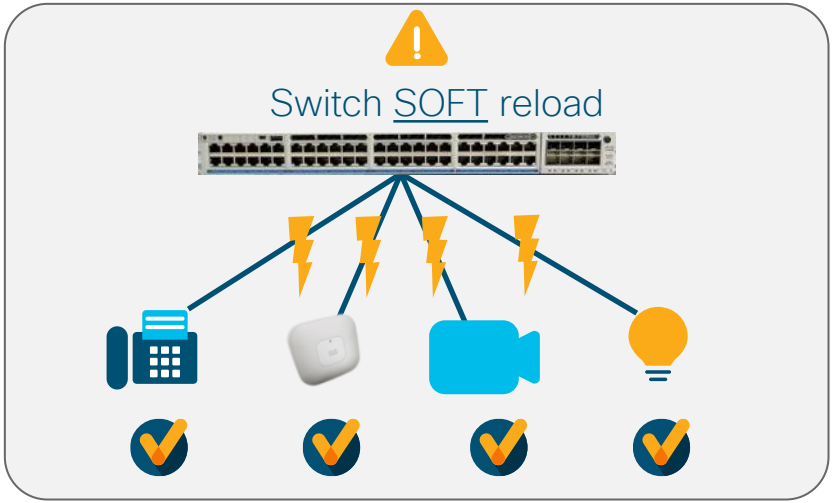


Load Shedding Based on configured priority

1. Low Priority Ports
2. High Priority Ports
3. Switch Priority – **Highest Priority**

Perpetual PoE

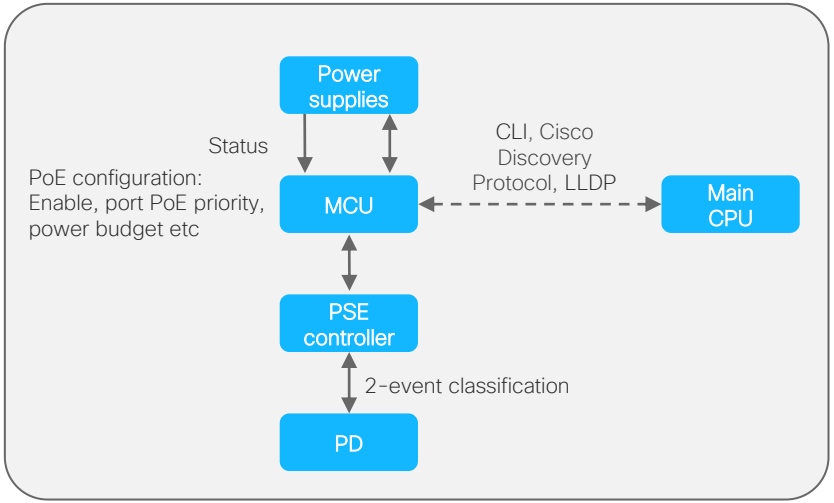
PoE devices connected to switch stay powered even on switch reload




```
Switch> enable
Switch# configure terminal
Switch(config)# interface gigabitethernet2/0/1
Switch(config-if)#
power inline port perpetual-poe-ha
Switch(config-if)# end
```

- PoE devices continue to get last negotiated power
- Applicable to “soft” reload – image upgrade, software crash, manual reboot
- Supported with stacking deployments
- Not applicable during power outage to switch or when front end processor is removed
- Not applicable when switch is in hibernation mode

Fast PoE




```
Switch> enable
Switch# configure terminal
Switch(config)# interface gigabitethernet2/0/1
Switch(config-if)#power inline port
perpetual-poe-ha
Switch(config-if)#power inline port poe-ha
Switch(config-if)# end
```

- Restores power to PD less than 30 seconds after restoration of power
- Works even before Cisco IOS® comes up
- Allocates last power (stored in NVRAM) drawn from PDs
- Works in stacking deployments

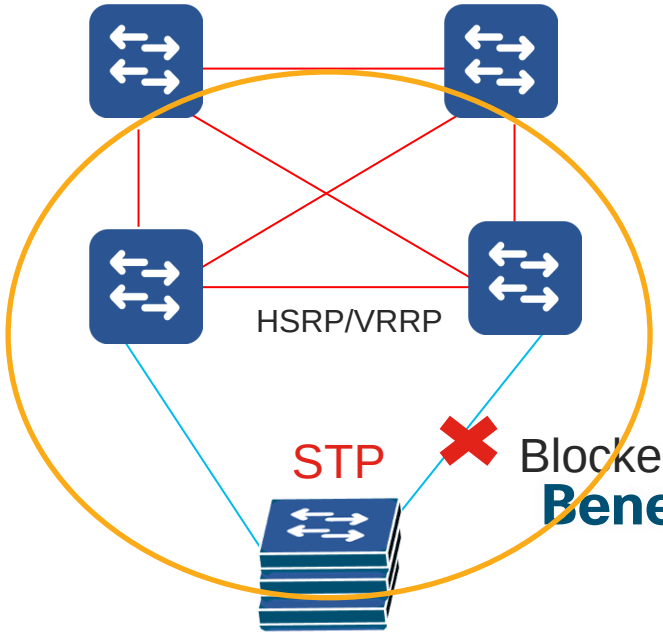
PoE innovations compatibility

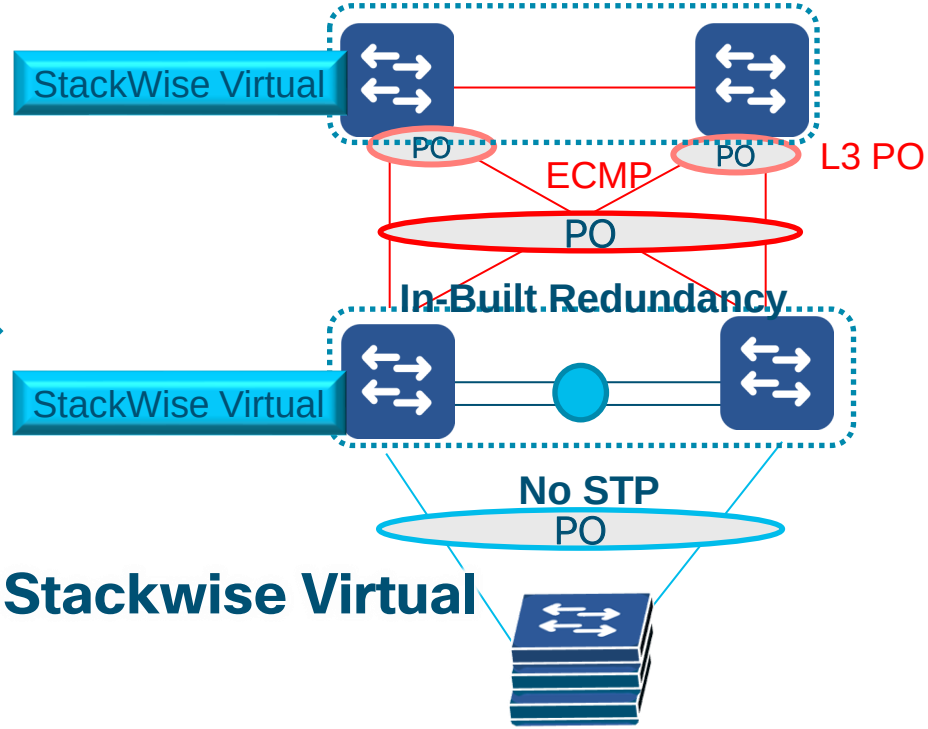
	9200 9200L	9300 9300L	9400
Perpetual PoE	○	○	–
Fast PoE	○	○	–

Stackwise Virtual

Stackwise Virtual

Topology Comparisons





Benefits of Stackwise Virtual

Simplify Operations by Eliminating STP, FHRP and Multiple Touch-Points

Double Bandwidth & Reduce Latency with Active-Active Multi-chassis EtherChannel (MEC)

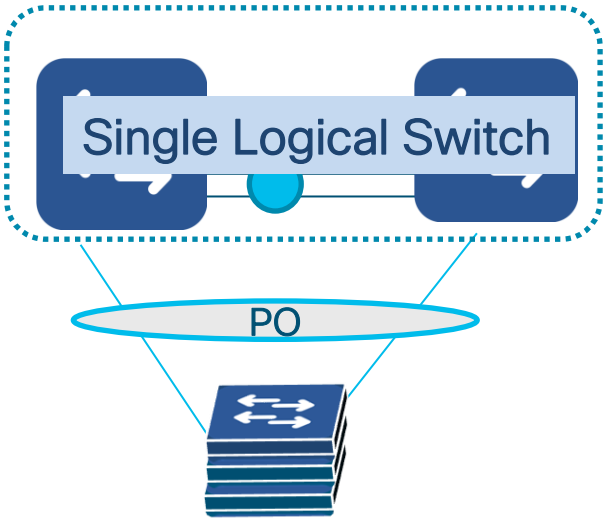
Minimizes Convergence with Sub-second Stateful and Graceful Recovery (SSO/NSF)

Stackwise Virtual Architecture

Control Plane

Unified Control Plane

- Manage, Configure and Troubleshoot single logical switch

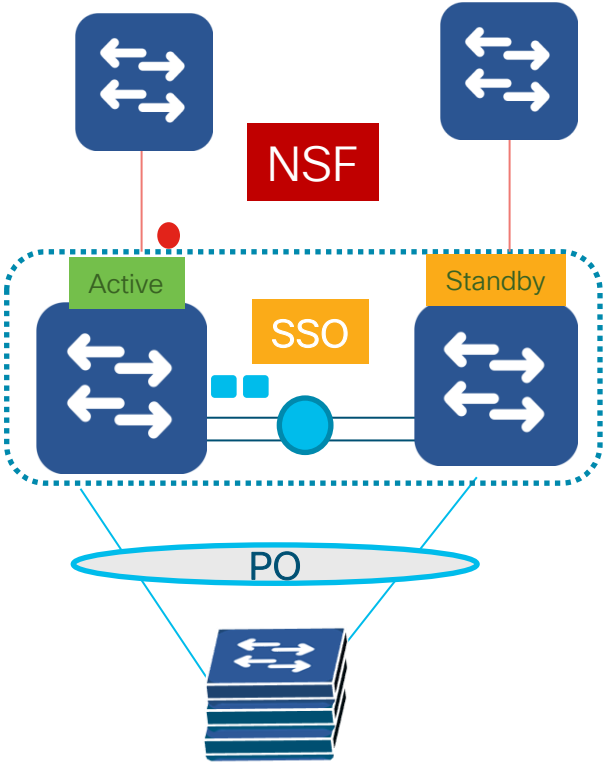


Stackwise Virtual Architecture

Control Plane

Control Plane Synchronization

- Stateful Synchronization of Layer 2 features and Protocols (SSO)
- Non-Stop Forwarding of L3 traffic (NSF)

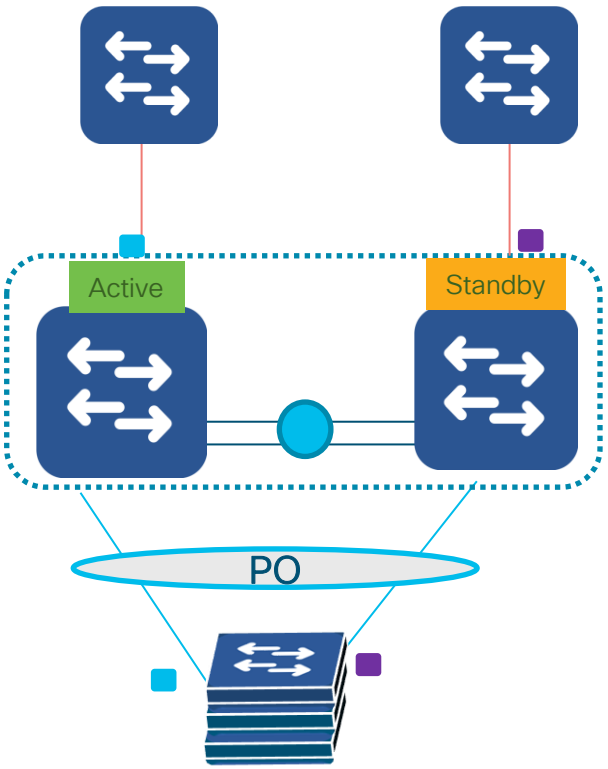


Stackwise Virtual Architecture

Data Plane

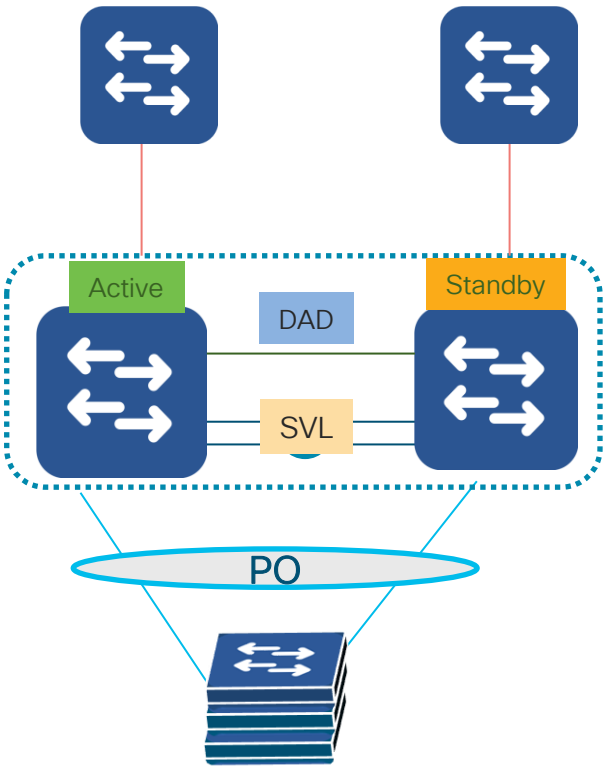
Active/Active Data Plane

- Both the switches are capable of forwarding the traffic locally without sending it over Interconnected-Link



Stackwise Virtual Components

- Stackwise Virtual Link
  - Dedicated Stacking Link facilitating communication between the switches (Internal PO – 8 links)
- Dual Active Detection Link
  - Dedicated Separate Connection to check and avoid dual-active scenario (4 Links)
- Multi-Chassis Ether-channel
  - Port-Channel Spanning across Stackwise virtual switches
  - L2 and L3 Port-channels

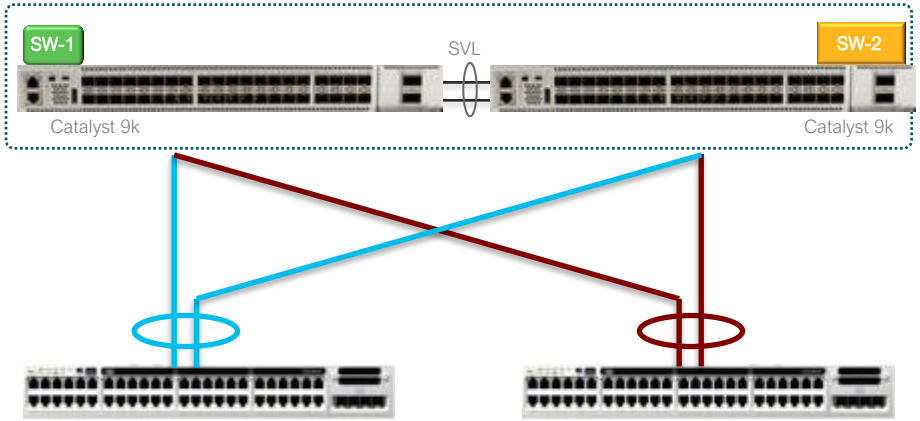


StackWise Virtual – Multi-Chassis EtherChannel

- Multi-Chassis EtherChannel (MEC) in StackWise Virtual enables cross stack-member link bundling into single logical L2/L3 Interface

- MECs can be deployed in three modes –

  Cisco PAgP, LACP and Static (ON)



High Availability
Dual-Active Detection

If the entire SVL bundle fails, the SVL Domain will enter into a "Dual Active" scenario

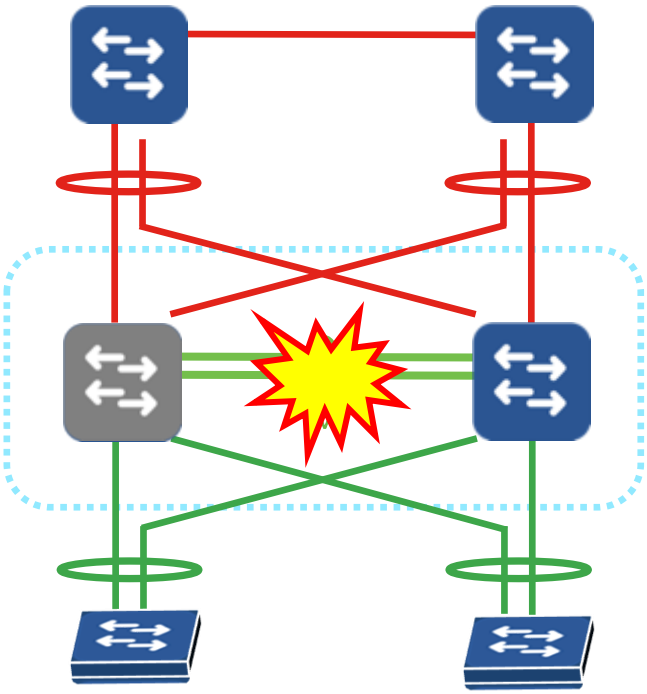
Both switches transition to SSO Active state, and share the same network configuration

- IP addresses, MAC address, Router IDs, etc.

This can cause communication problems in the network!

3 Step Process

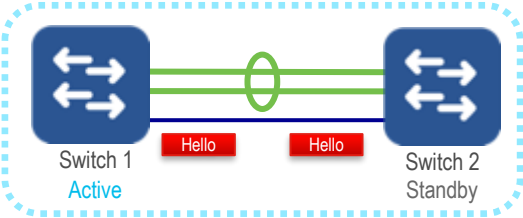
1. **Dual-Active Detection** - using any detection method enabled in the system.
2. Previous SVL Active shuts down ALL interfaces, and enters "Recovery Mode"… preventing further network disruption
3. **Dual-Active Recovery** - when the SVL recovers, the switch in Recovery Mode will reload to boot into a preferred standby state



High Availability

Dual-Active Protocols

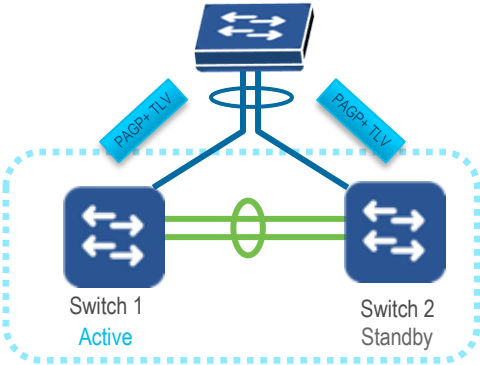
Fast Hello



- **Direct L2 Point-to-Point Connection**

- **Sub-Second Convergence**
  - Typically ~50-100ms

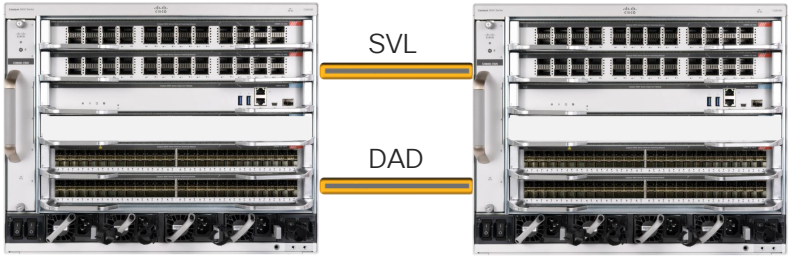
Enhanced PAGP



- **Requires ePAGP capable neighbor:**

- **Sub-Second Convergence**
  - Typically ~200-250ms

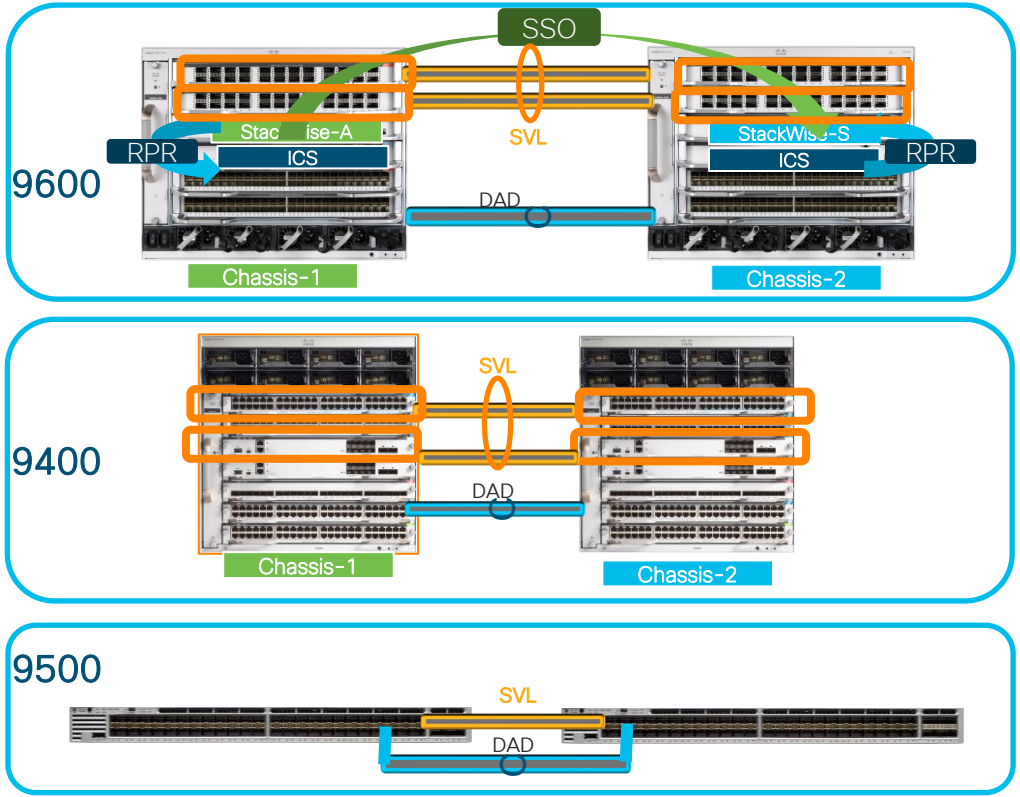
Cisco StackWise Virtual – Catalyst 9600



- SVL: StackWise Virtual Link
  - same speed ports (10G or higher)
  - Up to 8 ports
- DAD: Dual Active Detection:
  - Fast Hello
    - Directly connected
    - Up to 4 links
  - Enhanced PAgP
    - EtherChannel with PAgP
    - Up to 4 port-channels

- Typically, a distribution layer technology, allowing “stacking” of 2 switches
- Supports flexible distances with support of all supported cables and optics
- SVL and DAD are supported on any port with 10G or high speed for SVL and 1G or high speed for DAD.

StackWise Virtual Platforms – Deployment



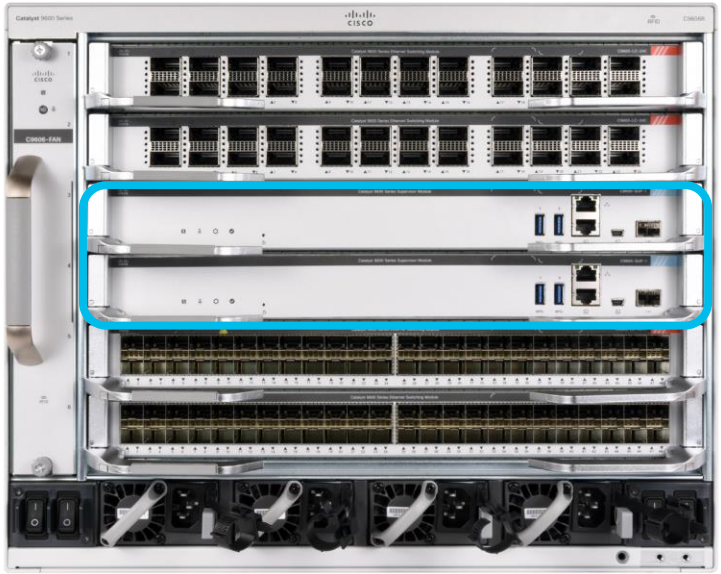
SVL: Linecards – 10/25/40/100G
DAD: Linecards – 1/10/25/40/100G

SVL: Supervisor/Linecards – 10/25/40G
DAD: Linecards – 1/10/25/40G

SVL: Supervisor/Linecards – 10/25/40G
DAD: Linecards – 1/10/25/40G

Quad Sup Support - 9600

Redundant Supervisor high availability mode



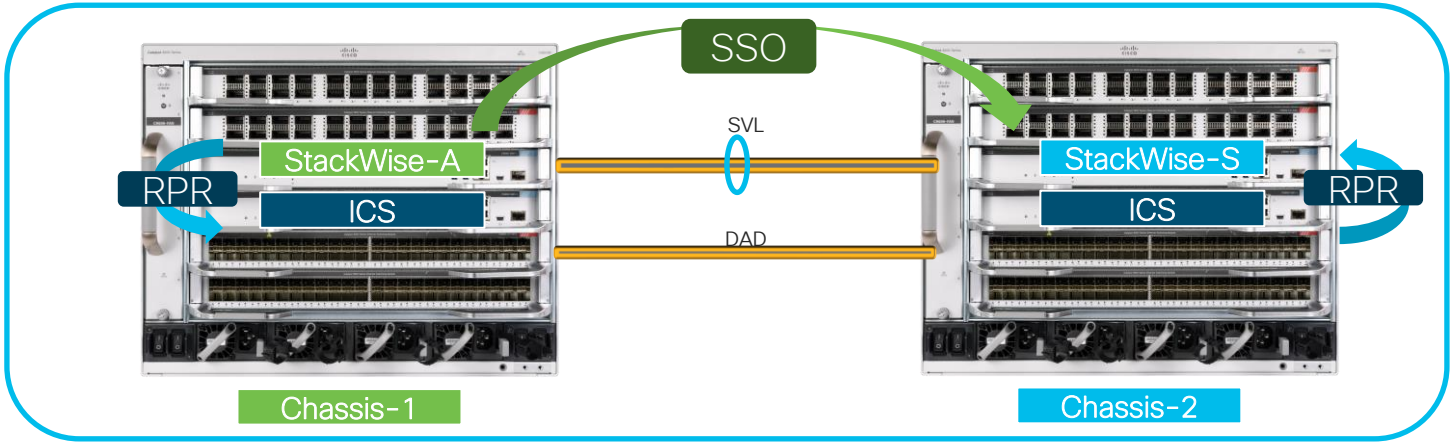
Catalyst 9600

Redundant Supervisors high availability modes:

- RPR: Route Processor Redundancy (Warm Standby)
- SSO: Stateful Switchover (Hot Standby)

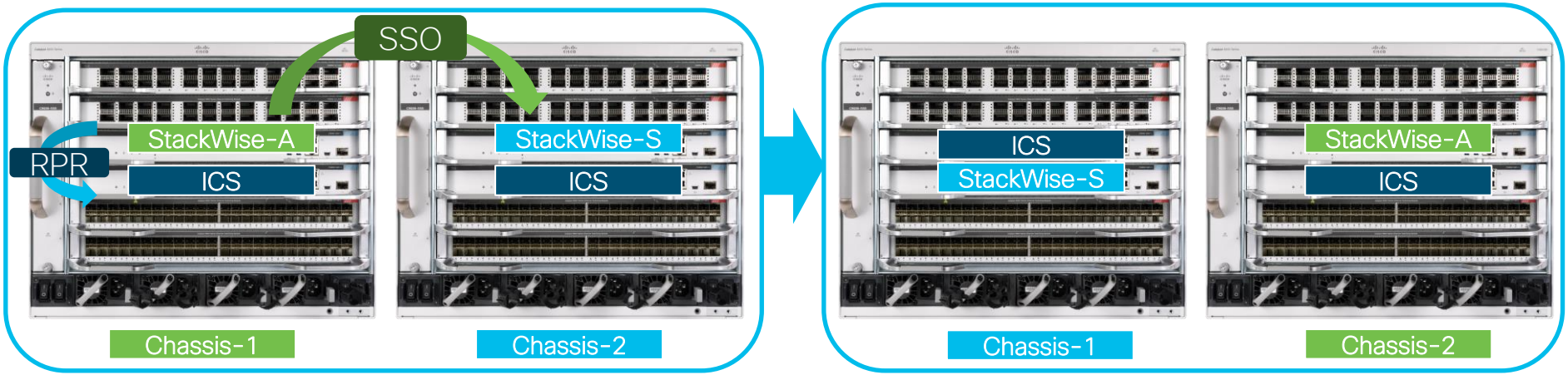
	RPR	SSO
Start-up Configuration Sync between Sups	Yes	Yes
Applications Sync states between Sups	No	Yes
Line Card Reset on Sup switchover	Yes	No

StackWise Virtual Quad Sup RPR



Role	Description	Control Plane	Data Plane
StackWise-A	StackWise Virtual Active In Chassis Active	Active	Active
StackWise-S	StackWise Virtual Standby In Chassis Active	Hot Standby	Active
ICS	In-Chassis Warm Standby	Warm Standby	Warm Standby

Quad Sup StackWise Virtual Switchovers



1. StackWise Active in Chassis-1reload, the StackWise-S in chassis-2 become StackWise Active
2. Cold standby supervisor in Chassis-1 continue the boot process to become StackWise standby while the line cards in chassis-1 get reset
3. ICS in Chassis-2 remain the same. The reloaded Sup in Chassis-1 comes back and become ICS in Chassis-1

Generic OnLine Diagnostics (GOLD)

Generic OnLine Diagnostics (GOLD)

GOLD (Generic Online Diagnostics) checks the health of hardware components to ensure that the system is working properly without any potential failures.

■ Features

The GOLD implementation checks the health of the hardware components to ensure that the system is working properly without any potential failures. Some tests are performed at system startup (bootup diagnostics), while others are performed while the system is running (runtime diagnostics).

■ Note

The test items differ depending on the model of the Cisco Catalyst 9000 Series.

■ Boot-up diagnostic results

```
C9200-01#show diagnostic post
Stored system POST messages:

Switch 1
---------

POST: CRYPTO Tests : Begin
POST: CRYPTO Tests : End, Status Passed

POST: PORT Loopback: loopback Test : Begin
POST: PORT Loopback: loopback Test : End, Status Passed

POST: SIF Tests : Begin
POST: SIF Tests : End, Status Passed

POST: Thermal, Temperature Tests : Begin
POST: Thermal, Temperature Tests End, Status Passed

C9200-01#
```

Generic OnLine Diagnostics (GOLD)

■ Types of runtime diagnostics

```
C9200-01#show diagnostic content switch 1
switch 1:
  Diagnostics test suite attributes:
    M/C/* - Minimal bootup level test / Complete bootup level test / NA
      B/* - Basic ondemand test / NA
    P/V/* - Per port test / Per device test / NA
    D/N/* - Disruptive test / Non-disruptive test / NA
      S/* - Only applicable to standby unit / NA
      X/* - Not a health monitoring test / NA
      F/* - Fixed monitoring interval test / NA
      E/* - Always enabled monitoring test / NA
      A/I - Monitoring is active / Monitoring is inactive
                                                  Test Interval   Thre-
  ID   Test Name                          Attributes      day hh:mm:ss.ms shold
  ==== ================================== ============    =============== =====
    1) DiagGoldPktTest -----------------> *BPN*X**I       not configured  n/a
    2) DiagThermalTest -----------------> *B*N****A       000 00:01:30.00 5
    3) DiagPhyLoopbackTest -------------> *BPD*X**I       not configured  n/a
    4) DiagScratchRegisterTest ---------> *B*N****A       000 00:01:30.00 5
    5) TestUnusedPortLoopback ----------> *BPN****I       not configured  n/a
    6) DiagStackCableTest --------------> ***D*X**I       not configured  n/a
    7) DiagMemoryTest ------------------> *B*D*X**I       not configured  n/a
```

■ Model-specific diagnosis correspondence table

	C9200	C9300	C9500
DiagGoldPktTest (MAC level)	○	○	○
DiagThermalTest (Temperature sensor)	○	○	○
DiagFanTest	—	○	○
DiagPhyLoopbackTest (PHY chip)	○	○	○
DiagScratchRegisterTest (ASIC chip)	○	○	○
TestUnusedPortLoopback (Port and ASIC data path)	○	○	○
TestPortTxMonitoring (Port operation)	—	○	○
DiagStackCableTest	○	○	—
DiagMemoryTest	○	○	○

■ Default value

ID 2) DiagThermalTest
ID 4) DiagScratchRegisterTest
It is automatically executed at predetermined intervals. Other items need to be executed manually or scheduled.

Generic OnLine Diagnostics (GOLD)

- Detailed description of runtime diagnostics

1) DiagGoldPktTest :
A loopback test at the MAC level of each port. It loops back the GOLD packet issued by the ASIC and checks the returned packet against the original packet.
This test does not interrupt the transfer function of the switch (Non-disruptive test).
It cannot be run as a health monitoring test.

2) DiagThermalTest :
System temperature and temperature sensor test. Make sure that the temperature read by the sensor is below the temperature threshold, which is the warning level.
This test does not interrupt the transfer function of the switch (Non-disruptive test).
It can be run as a health monitoring test.

3) DiagFanTest :
This is a test of the cooling fan module. Verify that all cooling fan modules are inserted and working properly.
This test does not interrupt the transfer function of the switch (Non-disruptive test).
It can be run as a health monitoring test.

4) DiagPhyLoopbackTest :
A PHY level loopback test for each port. It loops back the packet at the PHY level and checks the returned packet against the original packet.
The switch's forwarding function is disrupted while this test is running (Disruptive test).
It cannot be run as a health monitoring test.

5) DiagScratchRegisterTest :
Test the state of the ASIC. Write a value to a register on the ASIC, read the value again, and check that the register value is held correctly.
This test does not interrupt the transfer function of the switch (Non-disruptive test).
It can be run as a health monitoring test.

6) TestUnusedPortLoopback :
Perform a loopback test of the data path to the port and ASIC. The CPU sends a packet flooded in the VLAN to an unused port that has been shut down and checks the returned packet against the original packet.
This test does not interrupt the transfer function of the switch (Non-disruptive test).
It can be run as a health monitoring test.

7) TestPortTxMonitoring :
Test that each port is in the correct operating state. Periodically poll the transmit counters on each port to check that forwarded packets are being sent correctly and that there is no stacking.
This test does not interrupt the transfer function of the switch (Non-disruptive test).
It can be run as a health monitoring test.

8) DiagStackCableTest :
Test StackWise's path loopback feature.
The switch's forwarding function is disrupted while this test is running (Disruptive test).
It cannot be run as a health monitoring test.

9) DiagMemoryTest :
Test the memory on the ASIC. Inspect memory using an exhaustive test pattern based on the MBIST standard.
The switch's forwarding function is disrupted while this test is running (Disruptive test).
It cannot be run as a health monitoring test.
The switch will need to be restarted after the test is complete.

<Supplement>
Health monitoring is performed in the background at user-specified intervals. By default, the health monitoring test runs every 30 seconds.

https://www.cisco.com/c/en/us/td/docs/switches/lan/catalyst9300/software/release/17-3/configuration_guide/sys_mgmt/b_173_sys_mgmt_9300_cg/configuring_online_diagnostics.html?bookSearch=true#concept_fqm_vt2_nlb

Graceful Insertion and Removal (GIR)

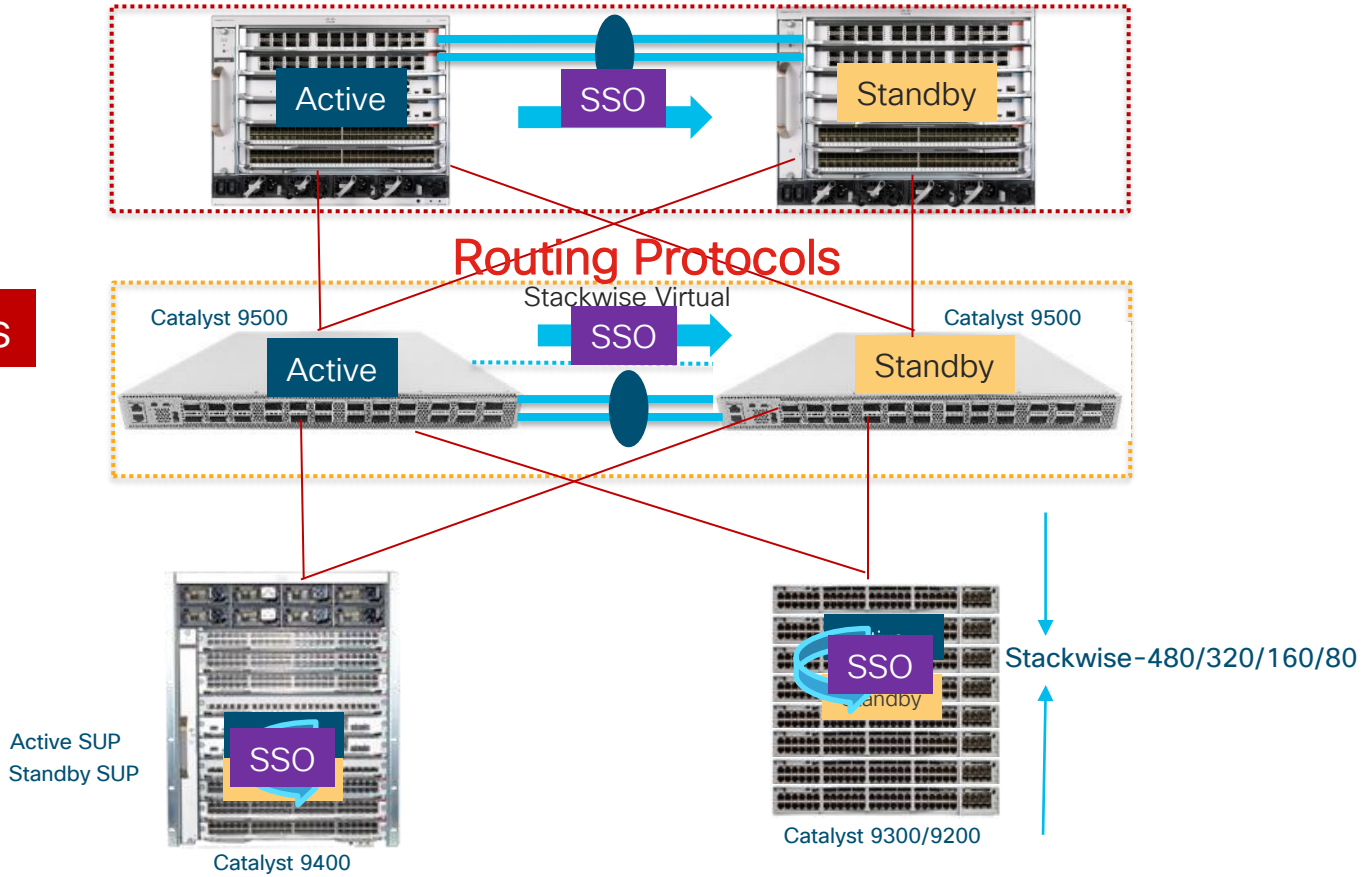
Graceful Insertion and Removal

- Upgrades with no or Minimal Traffic Loss
- Comprehensive Node Isolation Framework
- Easy Execution with a single command
- Highly Customizable workflow

Simple Customizable Non-Traffic Impacting

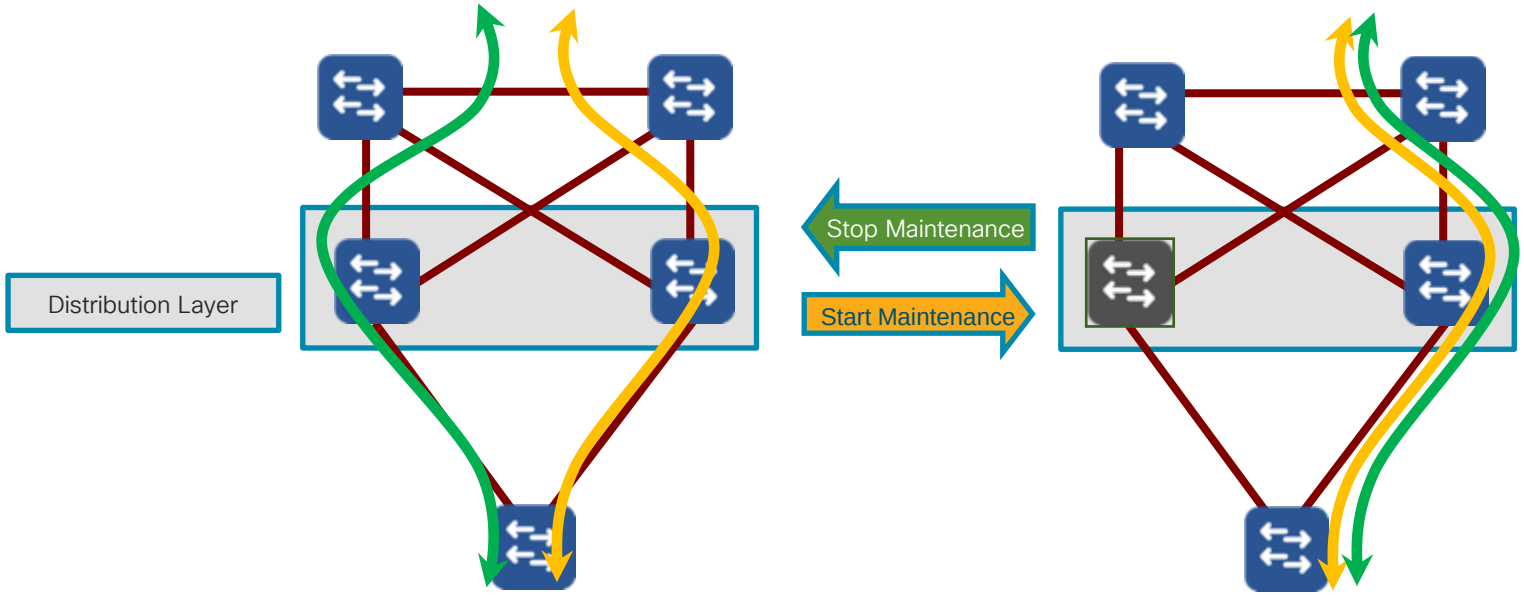
High Availability Architecture in Distro/Core

Routed Access

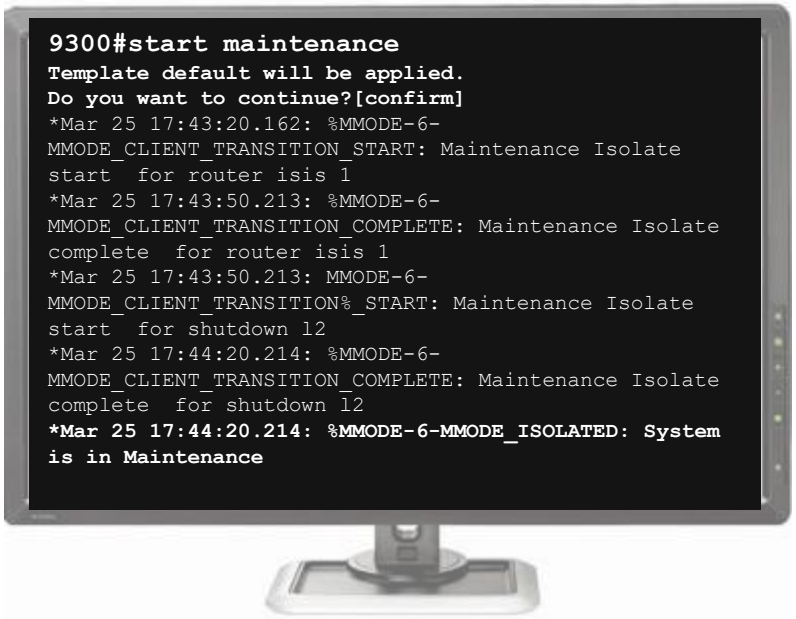


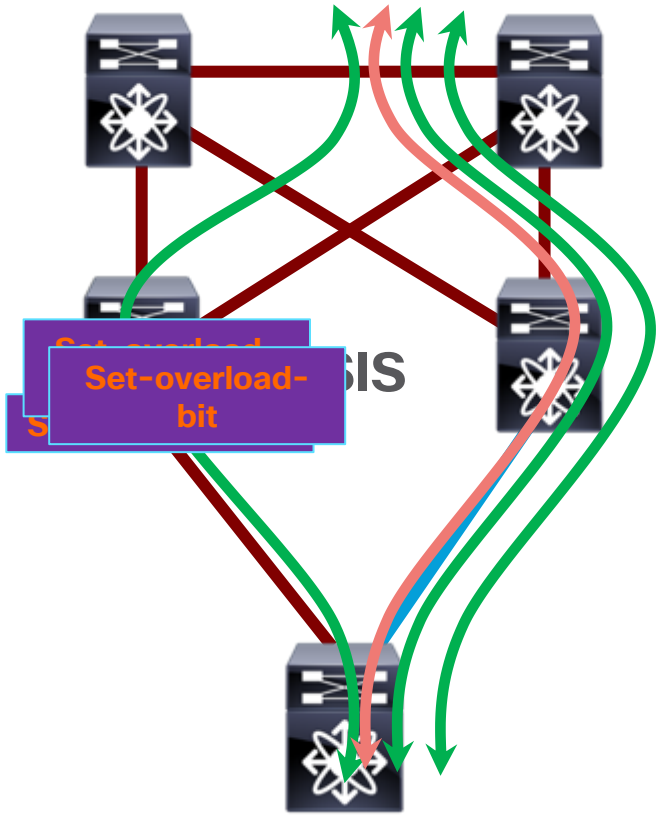
Graceful Insertion and Removal on Catalyst 9000

Isolation of Switch from network Gracefully



L2 and L3 Topology with GIR Isolation



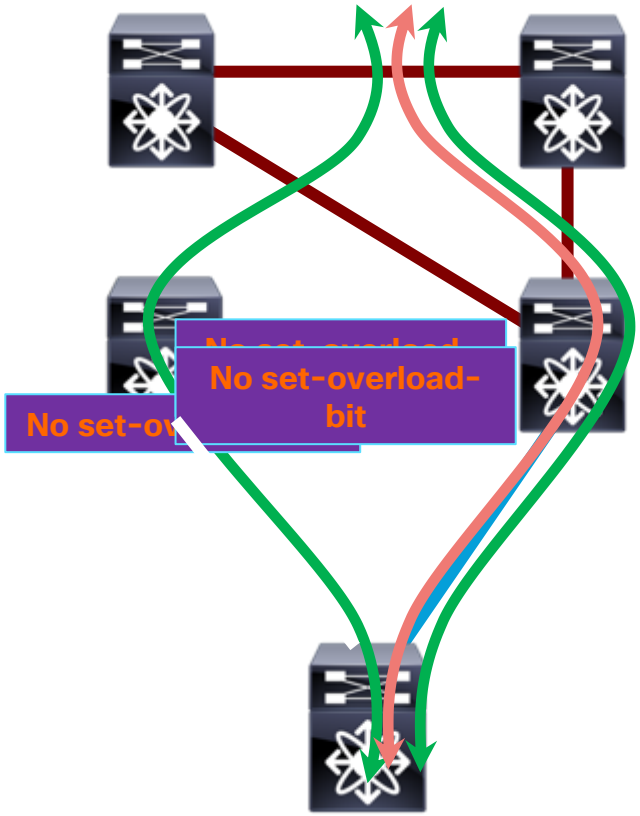


Order for Maintenance:

BGP -> IGPs in parallel (ISIS) -> L2

L2 and L3 Topology with GIR Isolation

```
9300#stop maintenance
*Mar 25 19:15:40.235: %MMODE-6-
MMODE_CLIENT_TRANSITION_START: Maintenance
Insert start  for shutdown l2
*Mar 25 19:16:10.237: %MMODE-6-
MMODE_CLIENT_TRANSITION_COMPLETE: Maintenance
Insert complete  for shutdown l2
*Mar 25 19:16:10.237: %MMODE-6-
MMODE_CLIENT_TRANSITION_START: Maintenance
Insert start  for router isis 1
*Mar 25 19:16:40.288: %MMODE-6-
MMODE_CLIENT_TRANSITION_COMPLETE: Maintenance
Insert complete  for router isis 1
*Mar 25 19:16:40.612: %MMODE-6-MMODE_INSERTED:
System is in Normal Mode
```



Order for Maintenance:

L2 → IGPs in parallel (ISIS) -> BGP

Graceful Insertion and Removal

Default and Customizable Templates

- **Default Template**
  - System Generated Profile based on the switch configuration
- **Customized Template**
  - User Configured Profile based on specific configuration or use case

```
9300L#show system mode maintenance template default
System Mode: Normal
default maintenance-template details:
router isis 1
shutdown l2
9300L#show system mode maintenance template test
System Mode: Normal
Maintenance Template test details:
shutdown l2
```

Configuration Profiles

- Maintenance-mode profile is applied when entering GIR mode,
- Normal-mode profile is applied when GIR mode is exited.

Automatic Profiles	Custom Profiles
• Generated by default • GIR is applied to all protocols running on the system • GIR state machine uses Registry mechanism to interface with client protocols • Use: Maintenance Windows	• User created profile for maintenance-mode and normal-mode using “templates” • Flexible selection of protocols for isolation • Use: maintenance windows and isolation during troubleshooting using preconfigured templates

Graceful Insertion and Removal

Snapshots

- Automatic Snapshots
  - Snapshots are automatically generated when entering and exiting maintenance mode
  - Captures operational data from the running system like Vlan's, Routes etc.
- User Configured Snapshots
  - Snapshots can be collected manually for comparing and troubleshooting

```
Switch#show system snapshots compare before_maintenance
after_maintenance
================================================================================
Feature             Tag            .before_maintenance .after_maintenance
================================================================================
[interface]

--------------------------------------------------------------------------------
      [Name:Vlan1]
                    packetsinput                 181587         **181589**
      [Name:GigabitEthernet1/0/3]
                    packetsinput                 101531         **101550**
                    broadcasts                   80893          **80910**
                    packetsoutput                211568         **211594**
      [Name:GigabitEthernet1/0/8]
                    output                       00:00:00,      **00:00:04,**
                    packetsinput                 6915           **6918**
                    packetsoutput                57677          **57706**
      [Name:GigabitEthernet1/0/17]
                    packetsinput                 101528         **101550**
                    broadcasts                   80891          **80910**
                    packetsoutput                211570         **211600**
```

In-Service Software Upgrade (ISSU) with Dual (Quad) Supervisors

In-Service Software Upgrade (ISSU) Overview

- ISSU provides a mechanism to perform software upgrades and downgrades without taking the switch out of service
- Leverages the capabilities of NSF and SSO to allow the switch to forward traffic during Supervisor upgrade (or downgrade)
- Key technology is the ISSU infrastructure
- Allows SSO between different extended maintenance versions

<200 ms

Modular Catalyst 9400/9600 with dual Supervisors

C9K ISSU

Dual Supervisor ISSU

3 Step Process

- Install add file <tftp/ftp/flash/disk:*.bin>
- Install activate ISSU
- Install commit

Granular Control on the upgrade process with ability to rollback

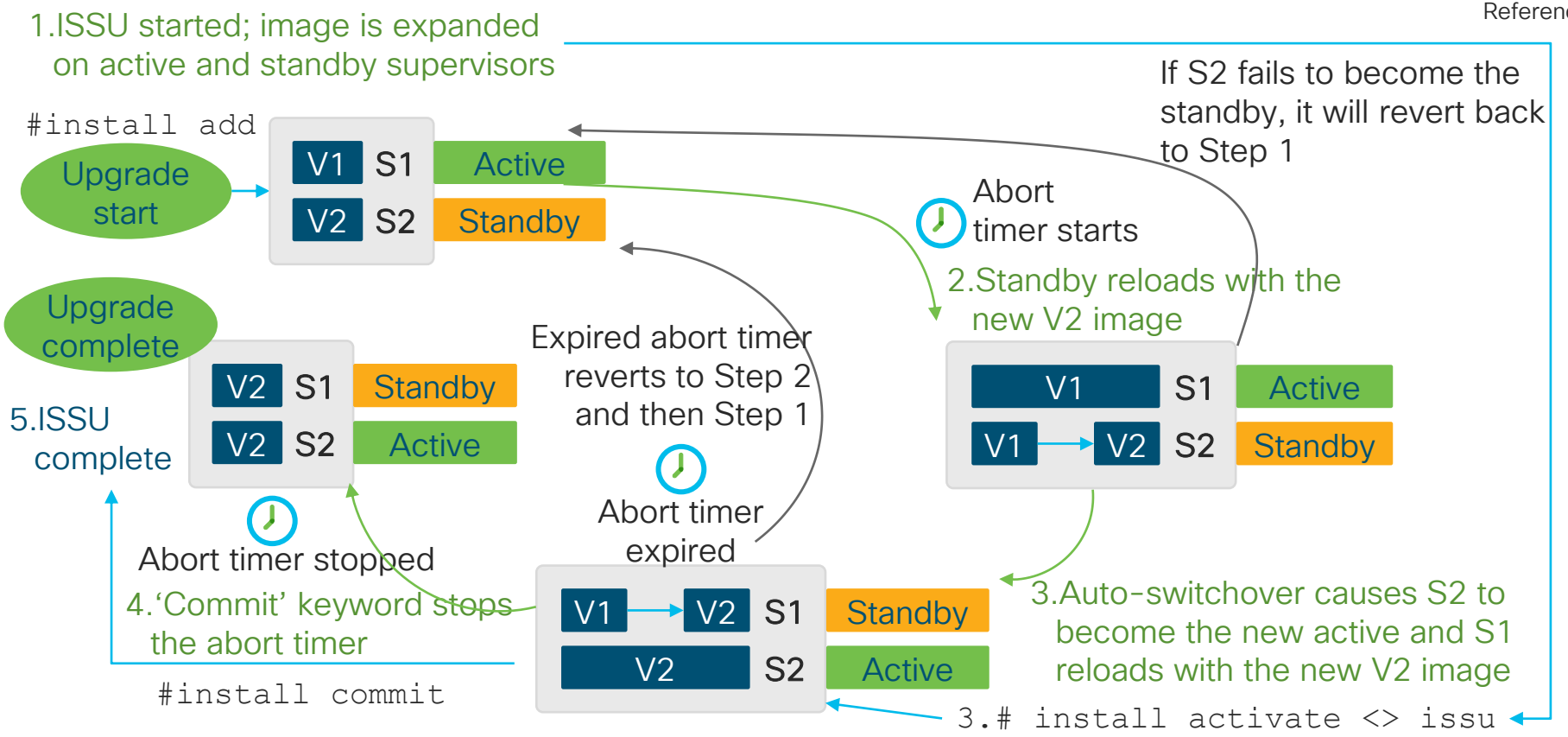
1 Step Process

- Install add file <tftp/ftp/flash/disk:*.bin>activate ISSU commit

Single Command to perform complete ISSU

Cisco Catalyst 9000 Series ISSU workflow

Reference



Install command line interface (CLI) commands

Supported in install mode, extended maintenance releases

Reference

Step-by-step workflow:

```
# install add <tftp://cisco.com/image.bin>
# install activate issu
```

On success

```
# install commit
```

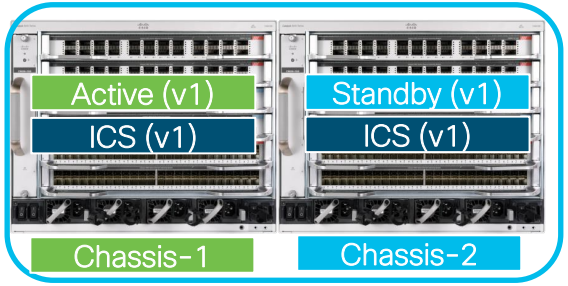
To abort

```
# install abort issu
```

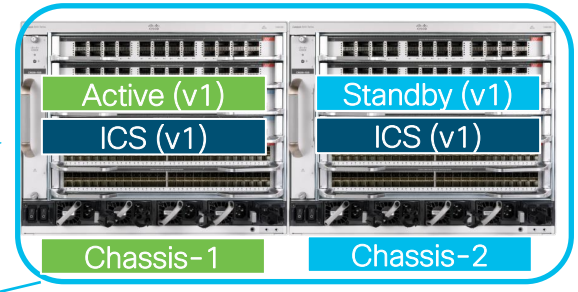
Workflow steps details:

- **`install add`** – performs the image download from the posted location
- **`install activate`** – upgrades the chassis with a new software version
- **`install commit`** – makes the changes permanent and deletes the older version of software from the chassis
- **`install abort issu`** – The operator can issue the abort command to revert the software back to the original state

ISSU



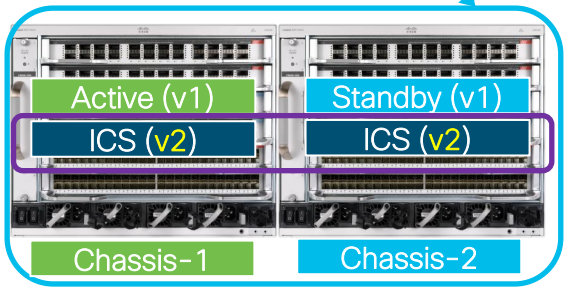
Install add file v2



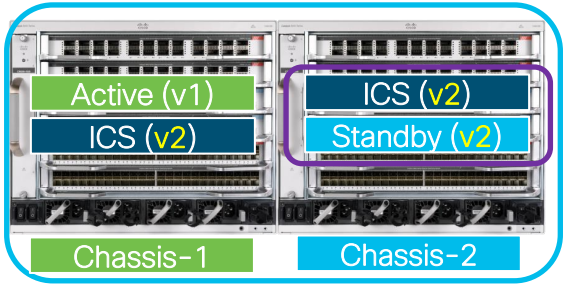
Role change from previous state

Download image(v2) and expand it to all supervisors (No disruption)

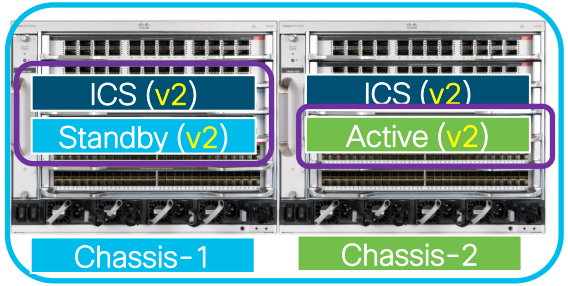
Install activate issu



1. Reload both ICS with v2 image(No disruption)



2. Reload Standby supervisor, ICS in Chassis-2 becomes Standby (Disruption in Chassis-2 with line card reset and supervisor states sync)



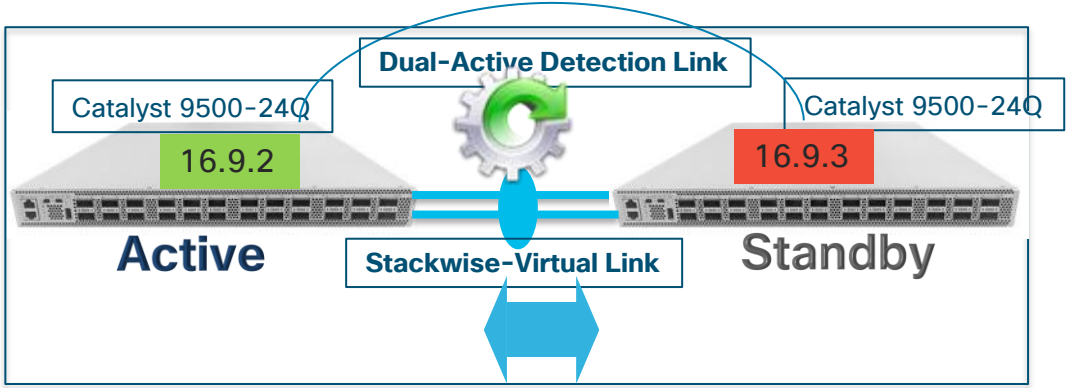
3. Reload Active supervisor to trigger a switchover (Disruption in Chassis-1 with line card reset and supervisor states sync)

Stackwise Virtual ISSU

Stackwise Virtual ISSU

ISSU Overview

- ISSU provides a mechanism to perform software upgrades and downgrades without taking the switch out of service
- Leverages the capabilities of NSF and SSO to allow the switch to forward traffic during Supervisor IOS upgrade (or downgrade)
- Key technology is the ISSU Infrastructure
  - Allows SSO between different versions



C9K ISSU

Stackwise Virtual ISSU

3 Step Process

- Install add file <tftp/ftp/flash/disk:*.bin>
- Install activate ISSU
- Install commit

Granular Control on the upgrade process with ability to rollback

1 Step Process

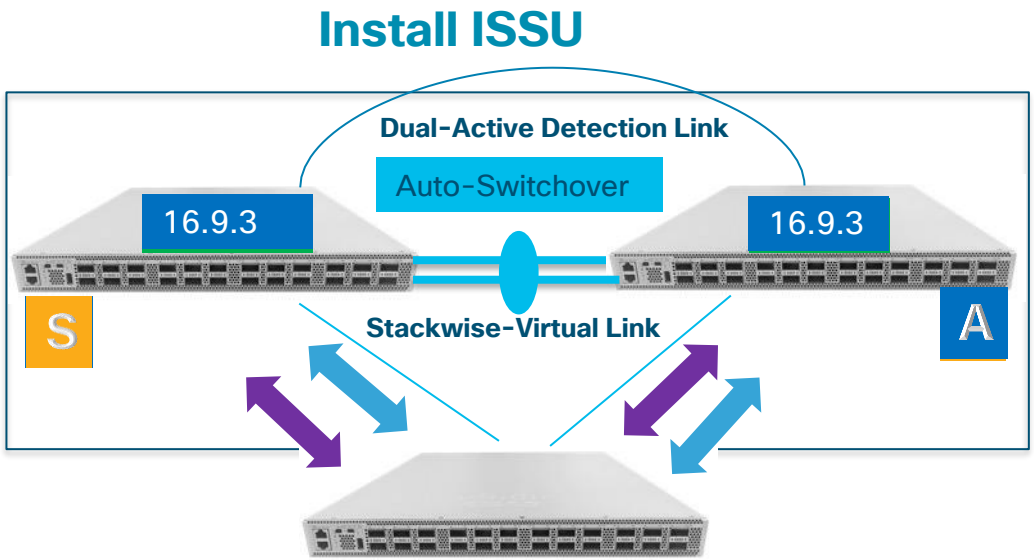
- Install add file <tftp/ftp/flash/disk:*.bin>activate ISSU commit

Single Command to perform complete ISSU

Stackwise Virtual ISSU

ISSU Process

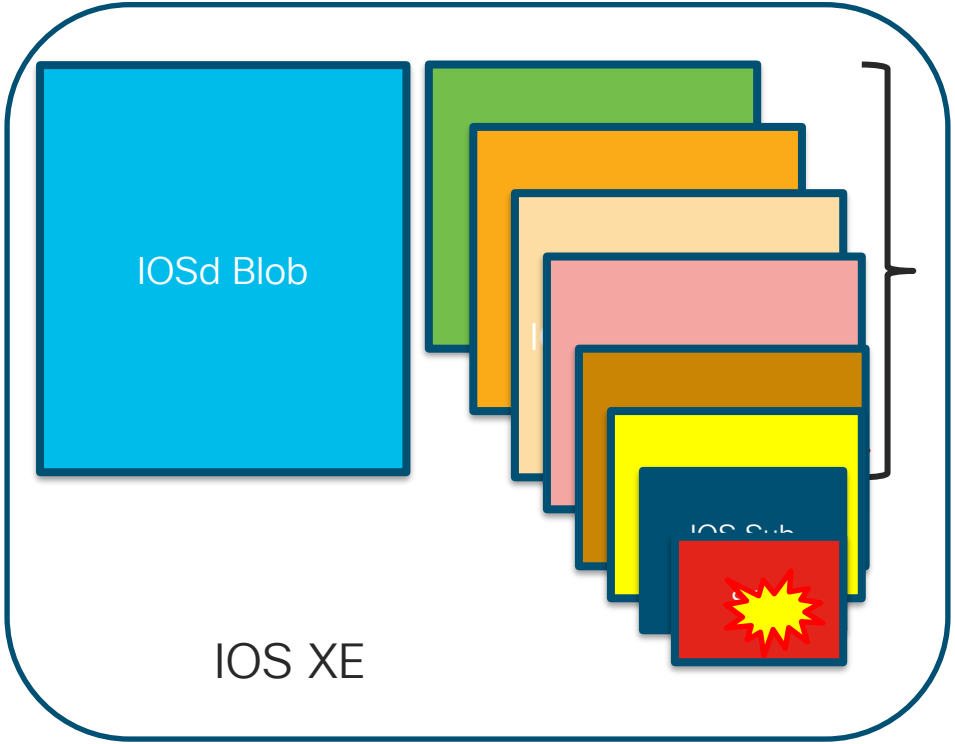
2nd Sub-second traffic convergence



1st Sub-second traffic convergence

Software Maintenance Update (SMU)

Open IOS XE – IOS Sub Systems



BGP

OSPF

MPLS, etc

Ability to patch the failed component without touching the entire operating system

IOS Sub Systems Enhances IOS Resiliency

Ready for Software Patching

SMU is an emergency point fix positioned for expedited delivery to a customer in case of a network down or revenue affecting scenario.

Cold Patching: Install of a SMU will require a system reload in the first release. It is traffic impacting.

Hot Patching: Install of a SMU does not require a reload.

SW Patching

Common among OSes

Install Add

Install Activate

Install Commit

Reference

Adding a SMU file

Activating SMU

Committing it

```
9300#install add file flash:cat9k-universalk9.2017-03-
17_21.53_zhangyu.301.CSCuo76464.SSA.smu.bin
install_add: START Sun Mar 26 01:13:29 UTC 2017
SUCCESS: Finished copying package(s) to the selected switch(es)
SUCCESS: install_add /flash/cat9k-universalk9.2017-03-
17_21.53_zhangyu.301.CSCuo76464.SSA.smu.bin Sun Mar 26 01:13:31 UTC 2017

9300#install activate file flash:cat9k-universalk9.2017-03-
17_21.53_zhangyu.301.CSCuo76464.SSA.smu.bin
install_activate: START Sun Mar 26 01:14:12 UTC 2017
2 install_activate: Activating SMU...

This operation requires a reload of the system. Do you want to proceed? [y/n]y
2 install_activate: Reloading the box to complete activation of the SMU...

9300#install commit
install_commit: START Sun Mar 26 01:24:41 UTC 2017
SUCCESS: install_commit  Sun Mar 26 01:24:43 UTC 2017
```

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

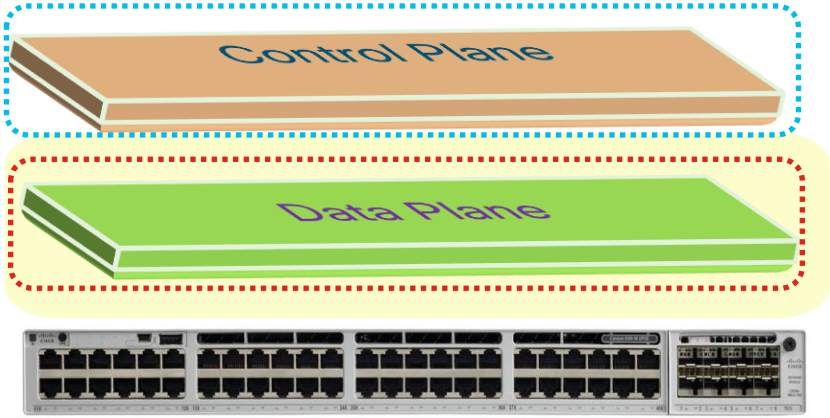
Extended Fast Software Upgrade(xFSU)

Extended Fast Software Upgrade on Catalyst 9300

- **xFSU** provides a mechanism to independently update the control plane and data plane during the upgrade process

- Control plane is upgraded by leveraging Graceful Reload Infrastructure without impacting data plane traffic

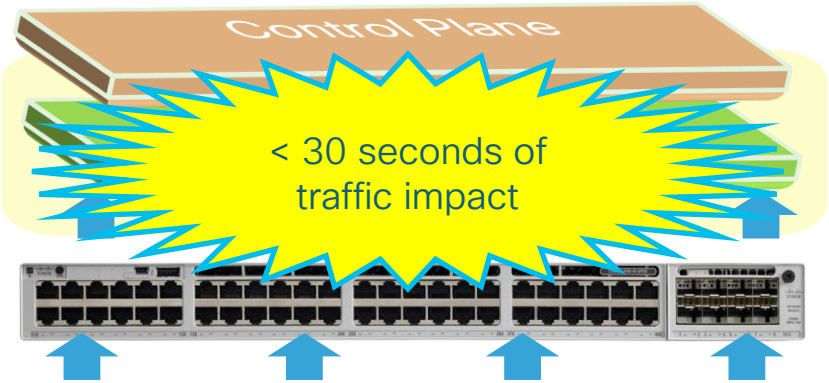
- Data plane(ASIC) is re-programmed in less than 30 seconds by leveraging special cache memory which stores active forwarding entries



Extended Fast Software Upgrade

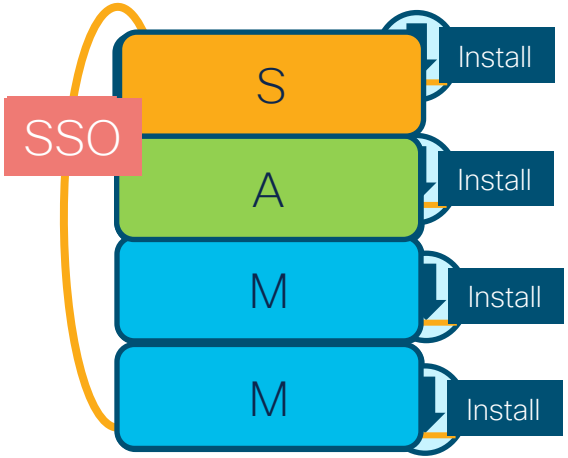
9300 standalone

```
#install add file image activate reloadfast commit
```



Extended Fast Software Upgrade on Stack

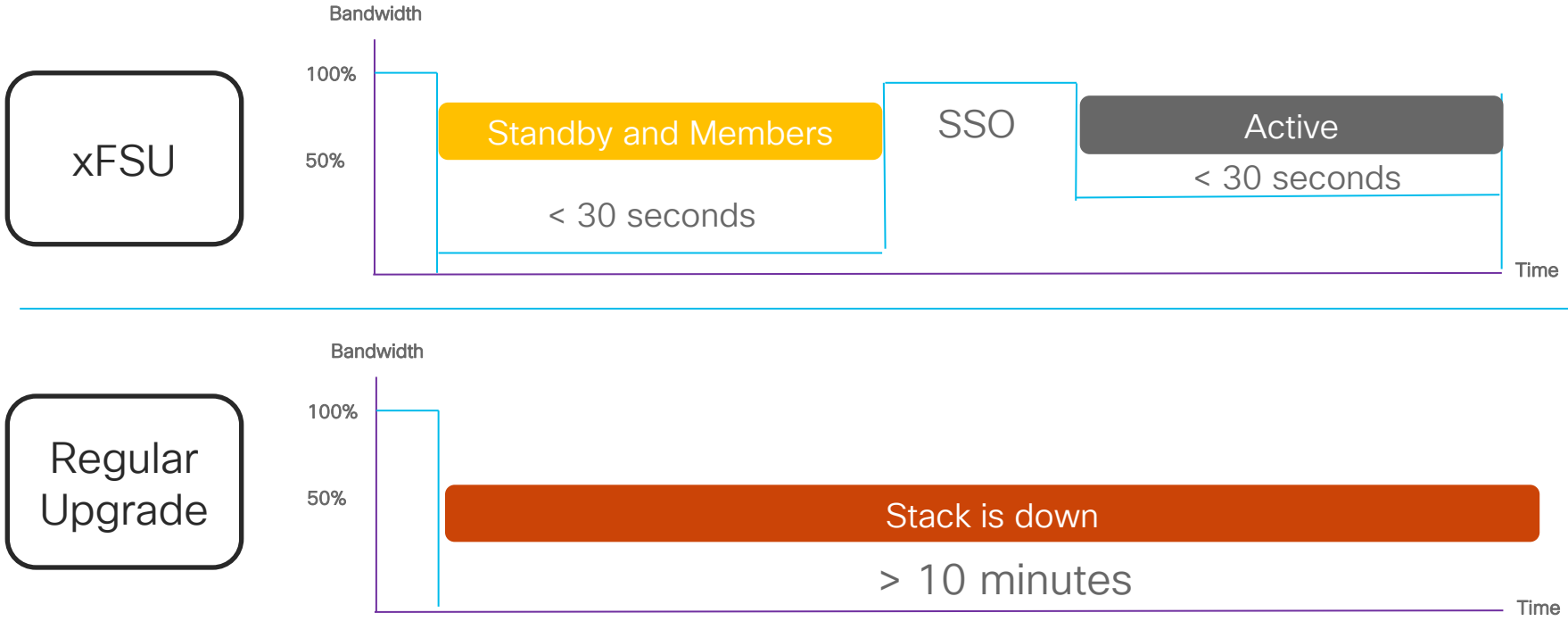
```
#install add file image activate reloadfast commit
```



1. Install the images on all switches
2. Fast reload the standby and member switches
3. Fast reload the active switch only
4. Standby becomes the new active
5. Previous Active switch becomes the new standby

Traffic Impact during the complete upgrade is less than 30 seconds

Convergence



Stateful Switchover (SSO)

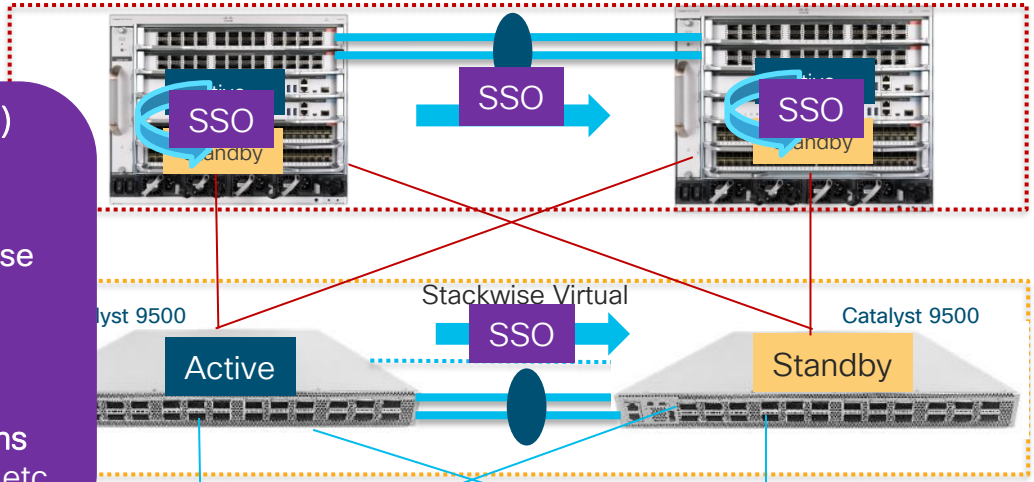
High Availability Architecture in Campus – SSO

Stateful Switchover (SSO)

SSO Aware Applications

Forwarding Information Base
IEEE 802.1x
PAgP / LACP
...and more

SSO Compliant Applications
Routing Protocols, Netflow, etc



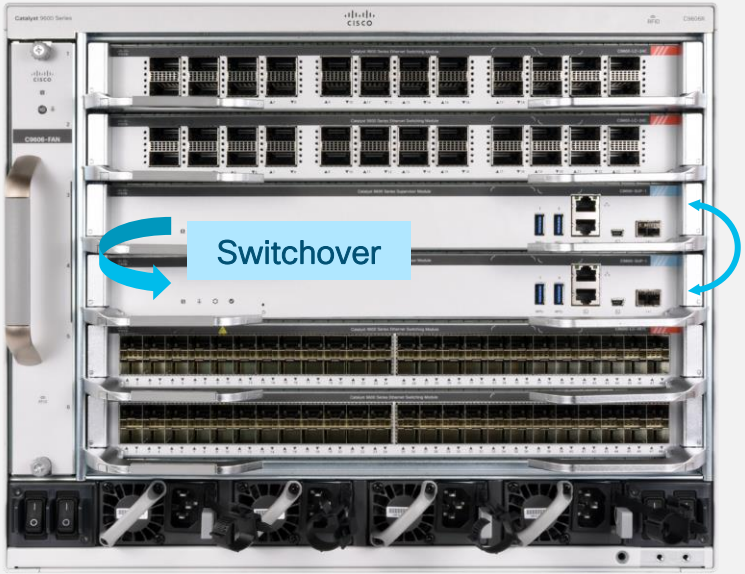
Active SUP
Standby SUP

Catalyst 9400

Stackwise-480/320/160/80

Catalyst 9300/9200

SSO – Catalyst 9000 Series modular chassis



SSO is the default redundancy mode with two supervisors in the system

- The active supervisor is responsible for all control plane processing
- The active supervisor is responsible for hardware programming on both the active and standby supervisors

SSO by itself Does Not Provide Redundancy for the Routing Protocols

Routing Protocol Redundancy With NSF

Active Supervisor/Switch

EIGRP RIB	
Prefix	Next Hop
10.0.0.0	10.1.1.1
10.1.0.0	10.1.1.1
10.20.0.0	10.1.1.1

OSPF RIB	
Prefix	Next Hop
192.168.0	192.168.0.1
192.168.55..0	192.168.55.1
192.168.32.0	192.168.32.1

ARP Table	
IP	MAC
10.1.1.1	aabbcc:ddee32
10.1.1.2	adbb32:d34e43
10.20.1.1	aa25cc:ddeee8

FIB Table	
Prefix	Next HOP
10.1.1.1	aabbcc:ddee32
10.1.1.2	adbb32:d34e43
192.168.0.0	aa25cc:ddeee8

Standby Supervisor/Switch

EIGRP RIB	
Prefix	Next Hop
-	-
-	-
-	-

OSPF RIB	
Prefix	Next Hop
-	-
-	-
-	-

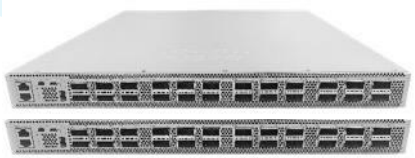
ARP Table	
IP	MAC
-	-
-	-
-	-

FIB Table	
Prefix	Next HOP
10.1.1.1	aabbcc:ddee32
10.1.1.2	adbb32:d34e43
192.168.0.0	aa25cc:ddeee8

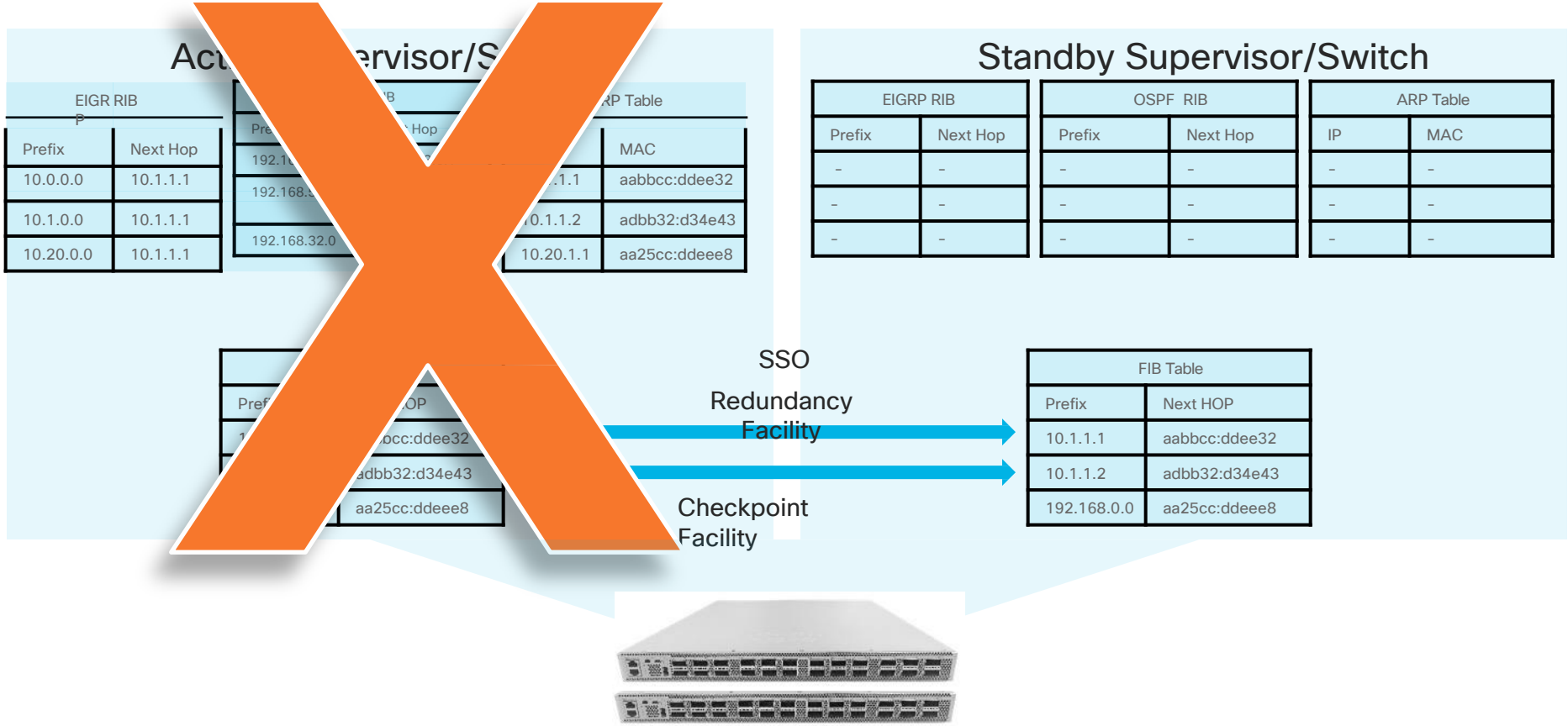
SSO

Redundancy Facility

Checkpoint Facility



Routing Protocol Redundancy With NSF



Routing Protocol Redundancy With NSF

Standby Supervisor/Switch

EIGRP RIB	
Prefix	Next Hop
1-0.0.0.0	-10.1.1.1
-10.1.0.0	-10.1.1.1
-10.20.0.0	1-0.1.1.1

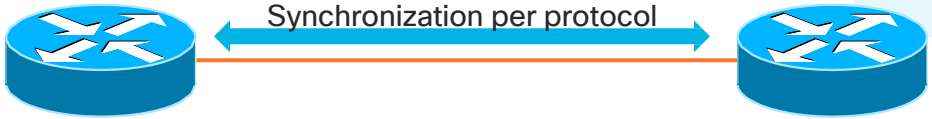
OSPF RIB	
Prefix	Next Hop
192.168.0	192.168.0.1
192.168.55..0	192.168.55.1
192.168.32.0	192.168.32.1

ARP Table	
IP	MAC
-10.1.1.1	a-abbcc:ddee32
-10.1.1.2	-adbb32:d34e43
-10.20.1.1	-aa25cc:ddeee8

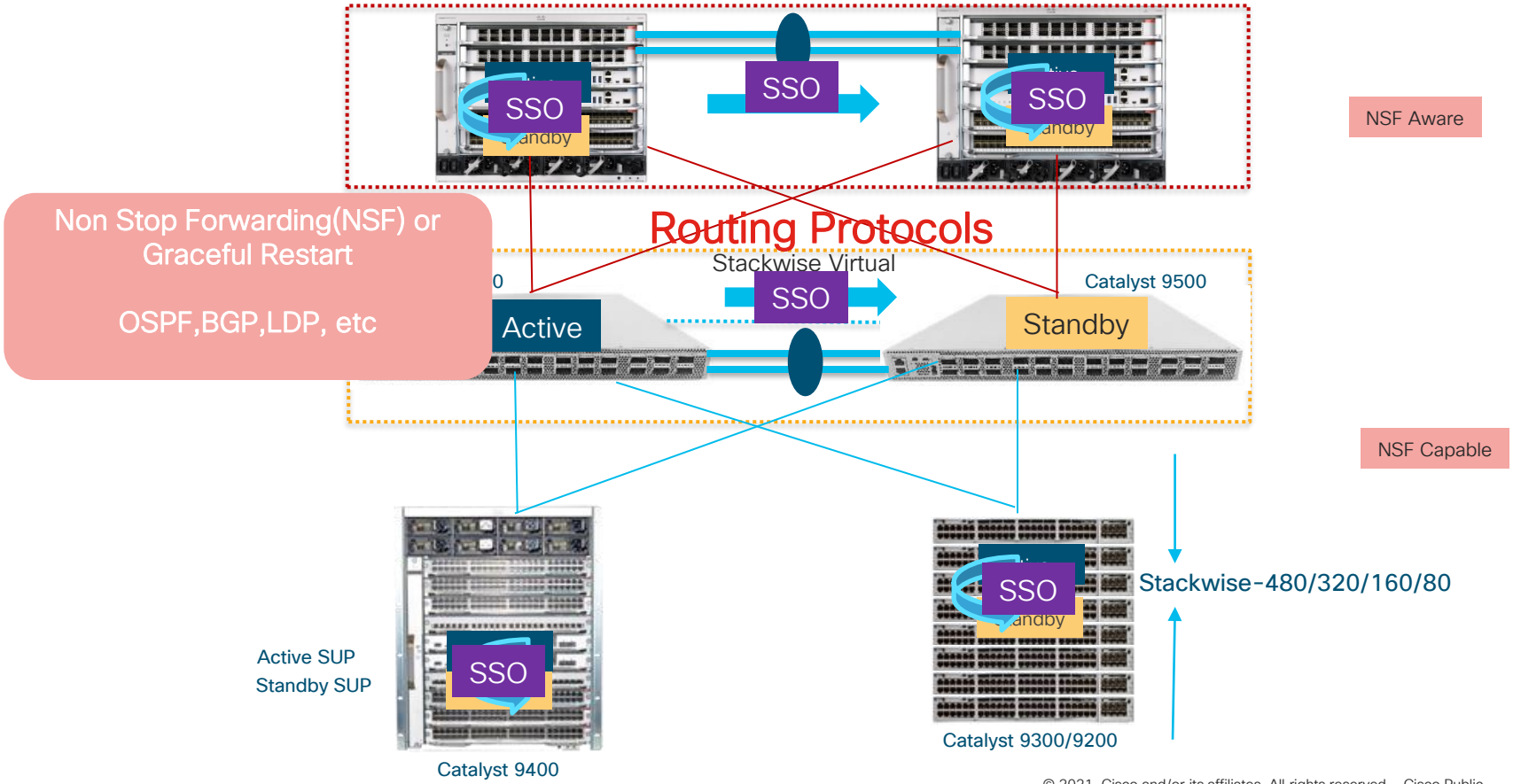
FIB Table	
Prefix	Next HOP
10.1.1.1	aabbcc:ddee32
10.1.1.2	adbb32:d34e43
192.168.0.0	aa25cc:ddeee8

GR/NSF Signaling per protocol

Synchronization per protocol

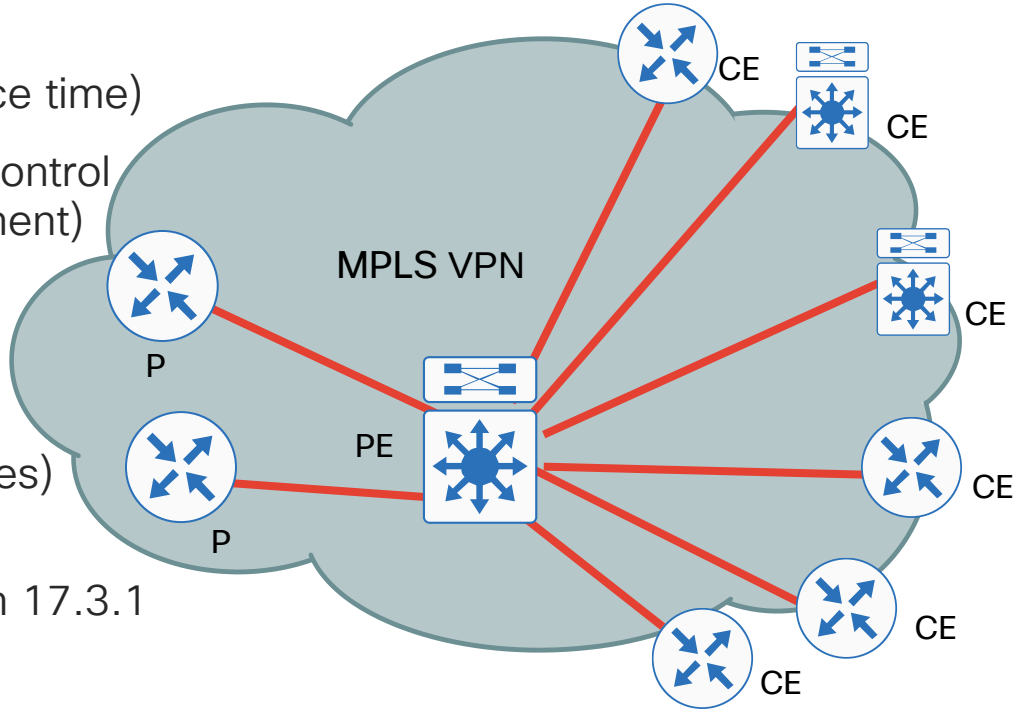


High Availability Architecture in Campus – SSO



Non-Stop Routing (NSR)

- Cisco IOS-XE Non-Stop Routing preserves the full state information (prefixes and related data) in the Routing Information Base across Supervisor Engine (Route Processor) switchover events.
- Avoids reconvergence with peer (versus NSF, which delays during grace time)
- Good for peer config not under your control (Example: CE attached to PE environment)
- Consumes more resources than NSF (memory, CPU)
- Device can also use NSR selectively (peering with P/PE/RR/other CE devices) to reduce resource consumption
- Available on Catalyst 9400, 9600 from 17.3.1



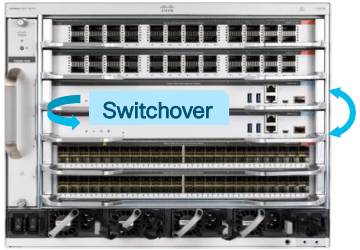
NSR configuring on Catalyst 9600

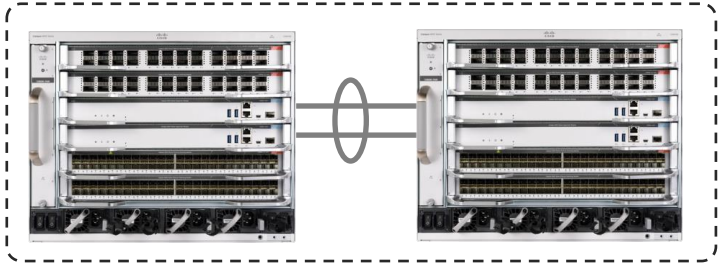
Agenda

- Specific Use-Cases
- Wired campus platform hardware and software features for HA
- Summary and conclusions

Summary and conclusions

Summary: Campus high availability using the Catalyst 9000 Series modular chassis

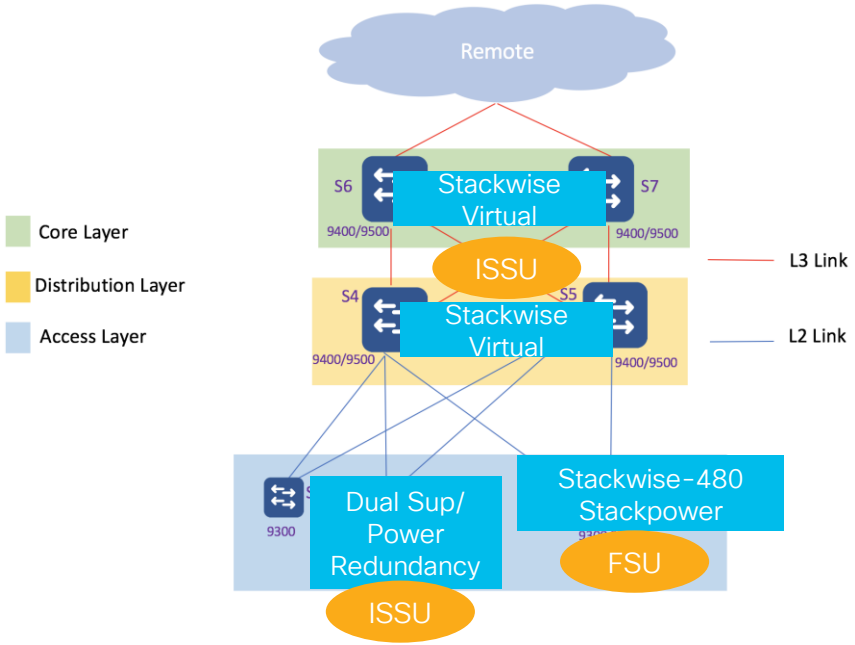




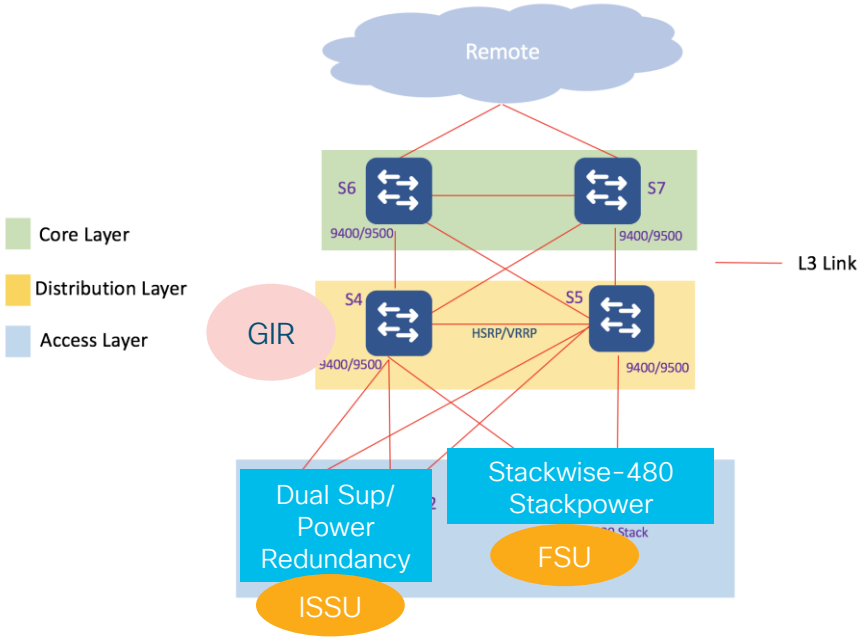
Physical redundancy	Stateful Switchover (SSO)	Non-Stop Forwarding (NSF)	In-Service Software Upgrade (ISSU)	Cisco StackWise Virtual
Redundant hardware • Power supplies • Fans (in tray) • Supervisors • Line cards	**Sub-second failover** • In chassis between Sups • Between chassis: Cisco StackWise-Virtual	**Resilient L3 topologies** • NSF support for OSPF, EIGRP, ISIS, BGP	**Minimize upgrade downtime** • SMU • ISSU • GIR (9600 future)	**Infrastructure resilience** • Multi-chassis EtherChannel (MEC) provides hardware-based failover

Enterprise Campus Network Designs

Multi-Tier **Layer2/3** Topology



Multi-Tier **Layer3** Topology



High Availability on Catalyst 9000

Catalyst 9200	Catalyst 9300	Catalyst 9400	Catalyst 9500/9600
SMU • Cold Patching	**Graceful Insertion & Removal(GIR)** Supported Protocols: ISIS, **OSPF,BGP, HSRP,VRRP** **Software Maintenance Upgrade** • Cold Patching • Hot Patching		
• StackWise	• StackWise • Stack Power • **Extended Fast Software Upgrade**	• StackWise Virtual • ISSU(StackWise Virtual) • ISSU (Dual Supervisor)	• StackWise Virtual • ISSU with StackWise Virtual • ISSU (Dual Supervisor)

Summary: Using the platform features

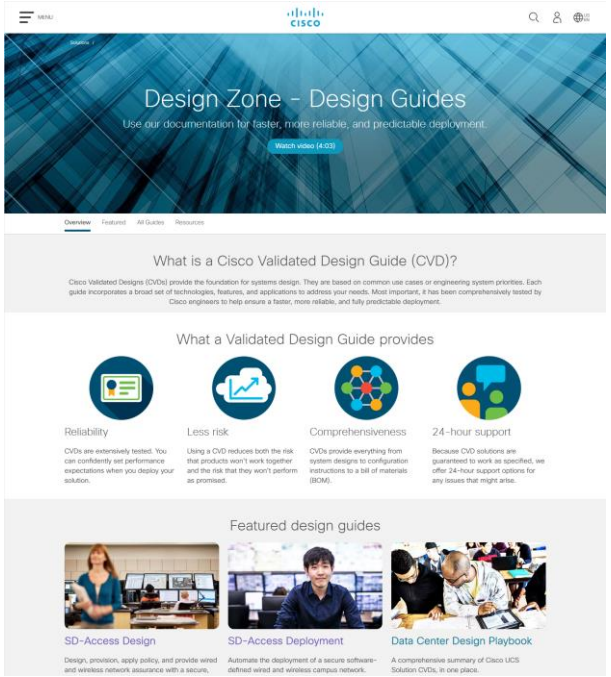
What is the recommendation?

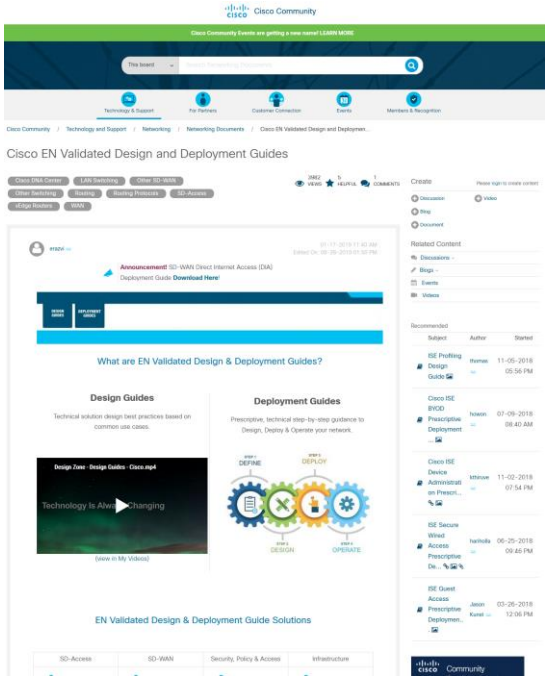
Option \ Situation	Critical Bug Fix & PSIRT	Hardware Upgrade	New Image Version
SMU Patching	✸	X	X
ISSU	✓	X	✸
GIR	X	✸	X
Box reload (Cold Boot)	✓	X	✓

Recommended	✸
Possible	✓
Not recommended	X

Design and deployment guidance available

https://cisco.com/go/cvd and https://cs.co/en-cvds





Thank you