Designing Highly Available Networks Using Catalyst 9000 Switches

Minhaj Uddin, Leader Technical Marketing

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
until February 24, 2023.
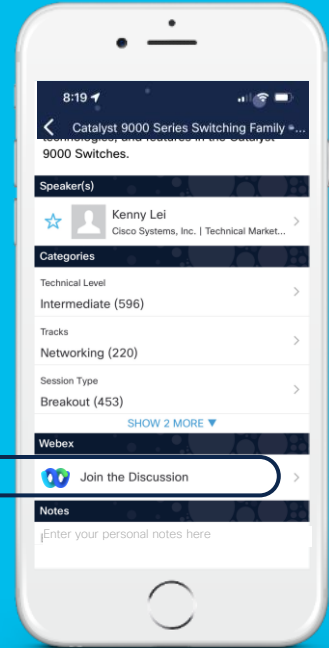
Agenda

- High Availability Overview and Evolution

- High Availability Architecture and Designs

- High Availability Solution on the Campus Access
  - Stackable High Availability Solution
  - Modular High Availability Solution

- High Availability Solution on the Campus Distribution/Core

- Summary/Q&A

Goals

- Efficiently utilize available bandwidth

- Dynamically respond to all types of disruptions

- Leverage most effective design techniques that meet the design requirements
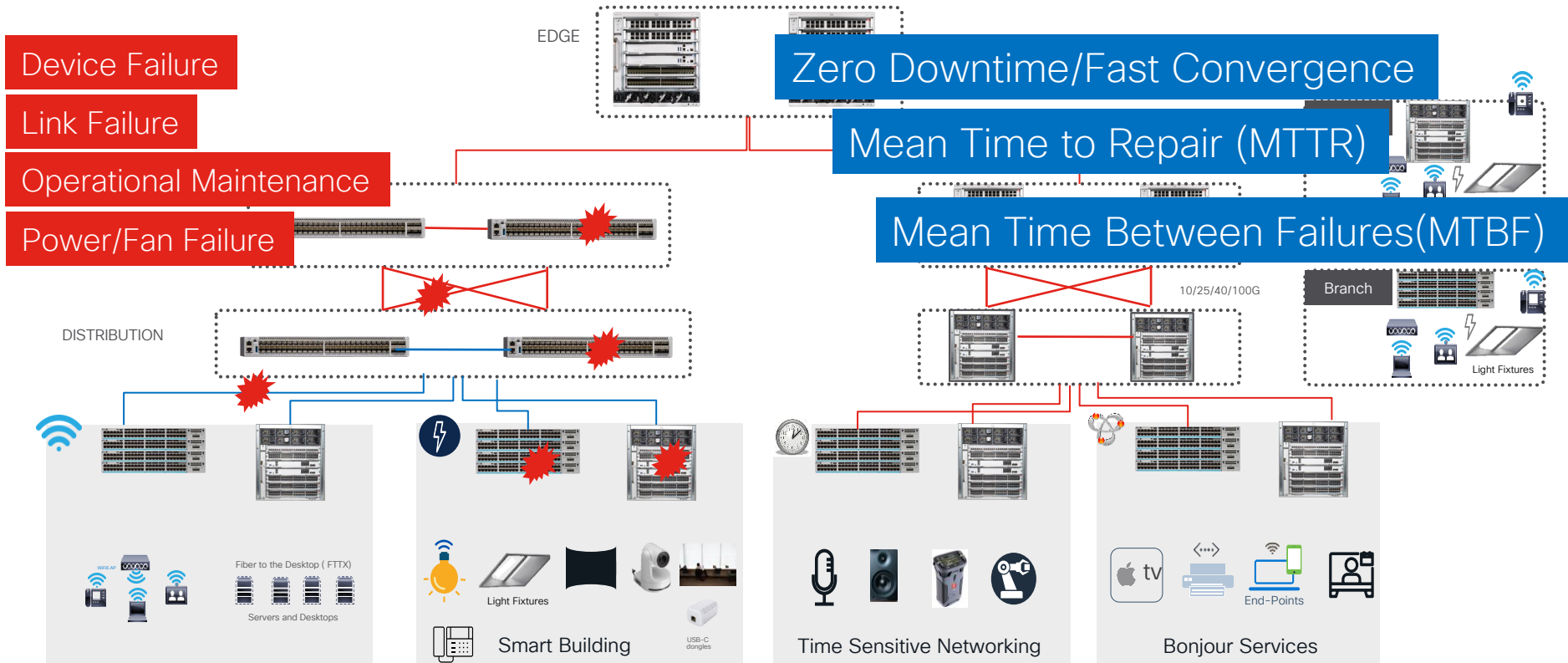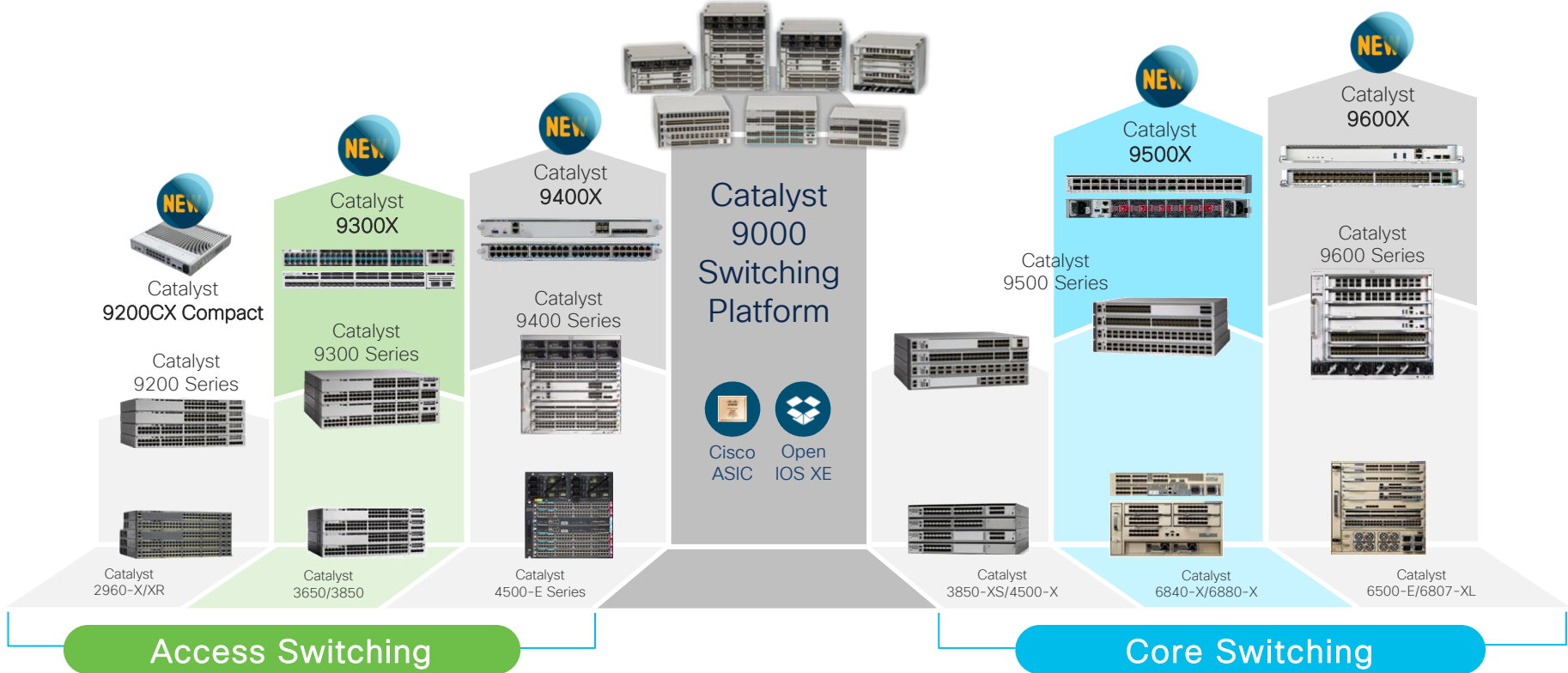
Typical Campus Network Architecture

- Add more links
- Add more devices
- Leverage FHRP like HSRP and VRRP
- Change the timers
- Tune the application performance
- Etc...

Device Failure

Link Failure

Operational Maintenance

Power/Fan Failure

Zero Downtime/Fast Convergence

Mean Time to Repair (MTTR)

Mean Time Between Failures (MTBF)

DISTRIBUTION

10/25/40/100G

Branch

Light Fixtures

Convergence Time?
Failover Detection?

What is the best way ?

Fiber to the Desktop (FTTX)

Servers and Desktops

Light Fixtures

USB-C dongles

Smart Building

Time Sensitive Networking

End-Points

Bonjour Services

Cisco Catalyst 9000 Switching Portfolio

Adding the "X factor" to the industry's leading switching family

Access Switching

- Catalyst 9200CX Compact (NEW)
- Catalyst 9200 Series / Catalyst 2960-X/XR
- Catalyst 9300X (NEW) / Catalyst 9300 Series / Catalyst 3650/3850
- Catalyst 9400X (NEW) / Catalyst 9400 Series / Catalyst 4500-E Series

Catalyst 9000 Switching Platform
- Cisco ASIC
- Open IOS XE

Core Switching

- Catalyst 9500 Series / Catalyst 3850-XS/4500-X
- Catalyst 9500X (NEW) / Catalyst 6840-X/6880-X
- Catalyst 9600X (NEW) / Catalyst 9600 Series / Catalyst 6500-E/6807-XL

Catalyst 9000: Pinnacle of Resiliency

Platform Resiliency

- **Redundant Switches and Supervisors** for Standalone and Modular Switches with Non-Stop Forwarding and Routing (NSF/NSR)

- **Redundant Fan & Power** Supply in case of any hardware failure

Design Resiliency

- **StackWise Virtual:** Redundant System for high availability, simplified design and configuration

- **GIR:** No downtime when device removed for maintenance
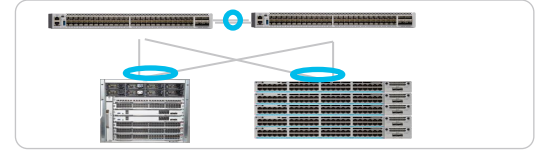
Operational Resiliency

- **ISSU:** Upgrade software with minimal to no traffic loss

- **xFSU :** Upgrade or Reload the Catalyst 9300 with very minimal traffic loss

- **Hot Patching:** No downtime for bug fixes (no reboot)

Eliminate downtime with High Availability designed at every level

Highly Available Architecture with Catalyst 9000 Switches
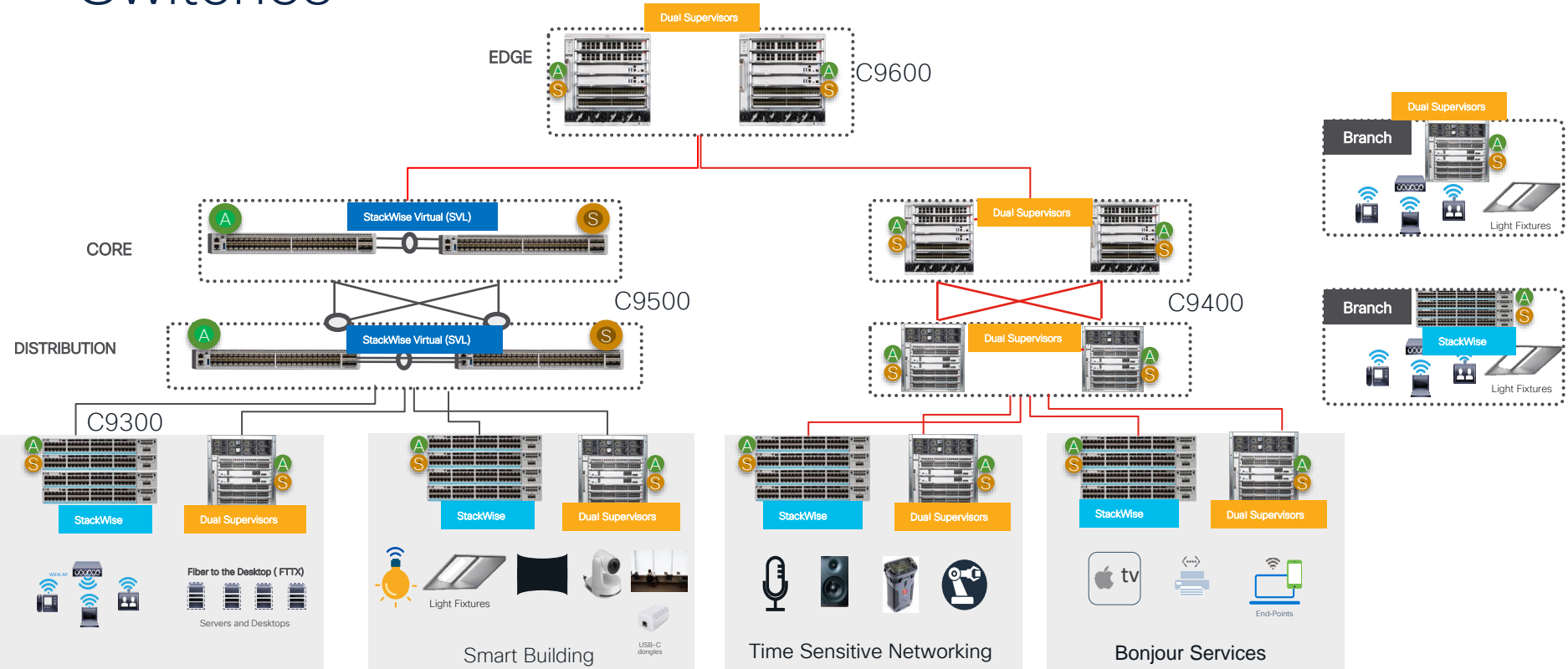
High Availability Components in Catalyst 9000

Stateful Switchover (SSO)

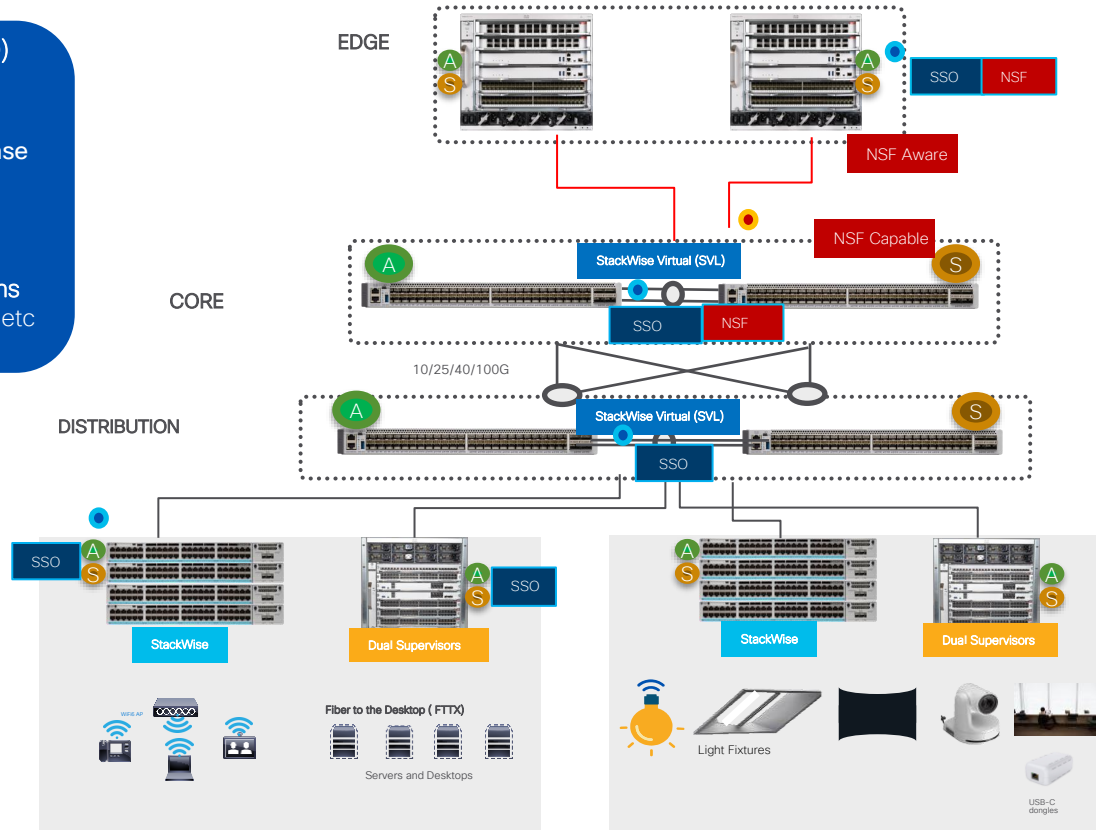SSO Aware Applications

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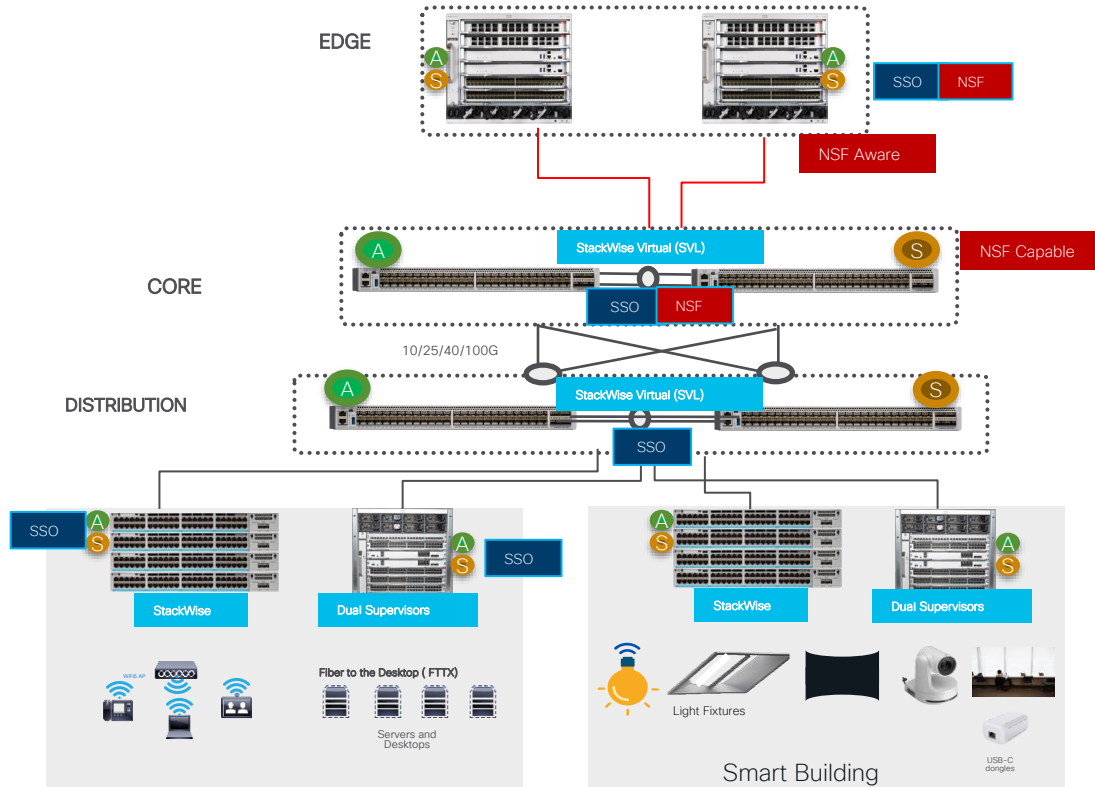PAgP / LACP
...and more

SSO Compliant Applications
Routing Protocols, Netflow, etc

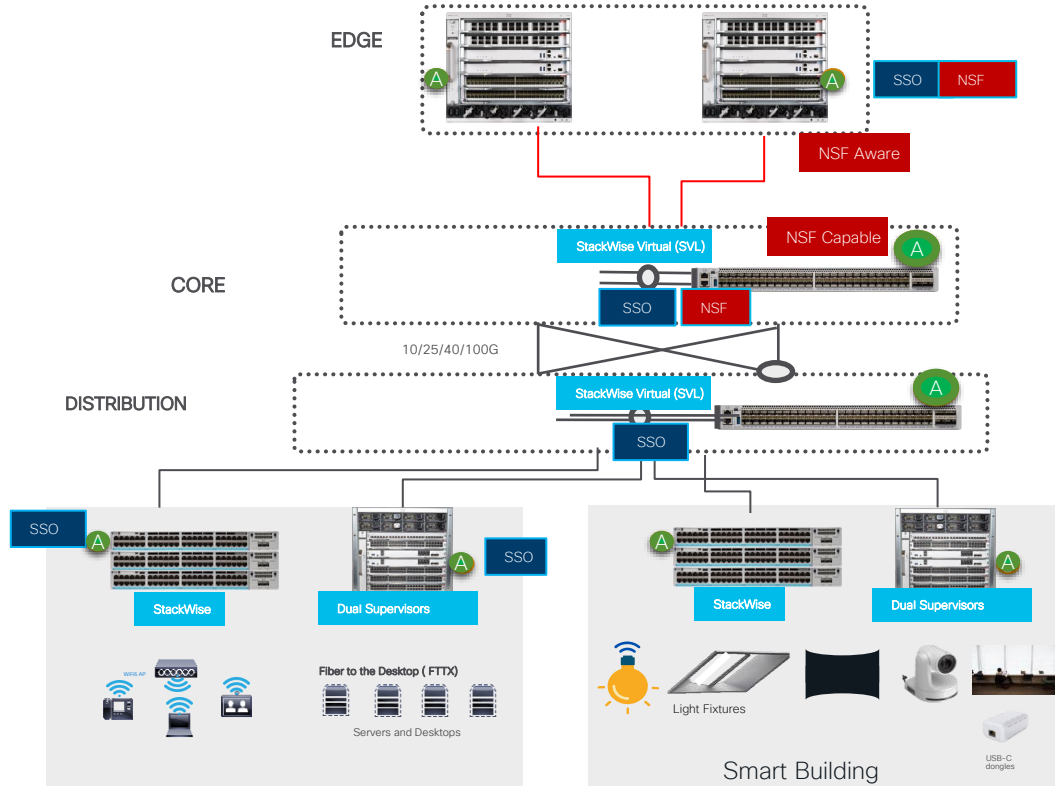**Non Stop Forwarding(NSF) or Graceful Restart**

OSPF,BGP,LDP, etc

EDGE

SSO | NSF

NSF Aware

NSF Capable

StackWise Virtual (SVL)

CORE

SSO | NSF

10/25/40/100G

DISTRIBUTION

StackWise Virtual (SVL)

SSO

SSO

SSO

StackWise

Dual Supervisors

StackWise

Dual Supervisors

WiFi AP

Fiber to the Desktop (FTTX)

Servers and Desktops

Light Fixtures

USB-C dongles

Typical Campus Network Architecture Failures

© 2023 Cisco and/or its affiliates. All rights reserved.   Cisco Public

Typical Campus Network Architecture Recovery

Catalyst 9000: Platform Resiliency

C9300 Fixed Platform

- **StackWise: Redundant System for high availability** with NSF/SSO
- **StackPower**: Redundant Power Supplies providing 1+ N redundancy
- **Redundant Fan & Power** Supply in case of any hardware failure

C9400/C9600 Modular Chassis

- **Redundant Supervisor:** Redundant System for high availability, simplified configuration
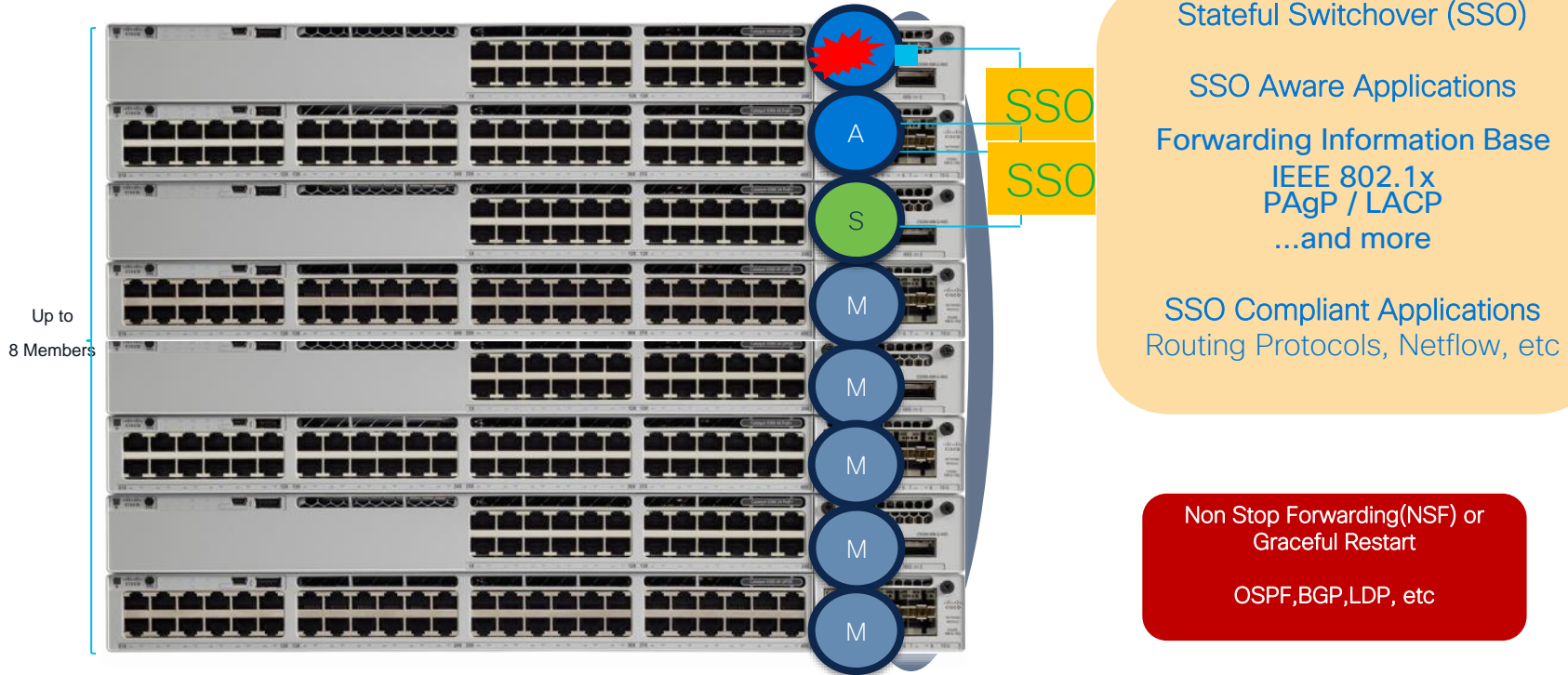- **Redundant Fan & Power** Supply in case of any hardware failure

C9500 Fixed Platform

- **StackWise Virtual:** Redundant System for high availability with NSF/SSO
- **Redundant Fan & Power** Supply in case of any hardware failure

Sub-second Traffic impact during failures across C9k Family

Device Redundancy with StackWise

Stackwise-
160/320480/1T

Stateful Switchover (SSO)

SSO Aware Applications

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

SSO Compliant Applications
Routing Protocols, Netflow, etc

Non Stop Forwarding(NSF) or
Graceful Restart

OSPF,BGP,LDP, etc

Catalyst 9300 Platform Resiliency for Power Supplies

Normal   **PS failure**

Combined (default)

Load sharing on all PSs   Load on Single PS

Redundant

System and PoE Load sharing on active PS
Standby PS shares the load of a system only
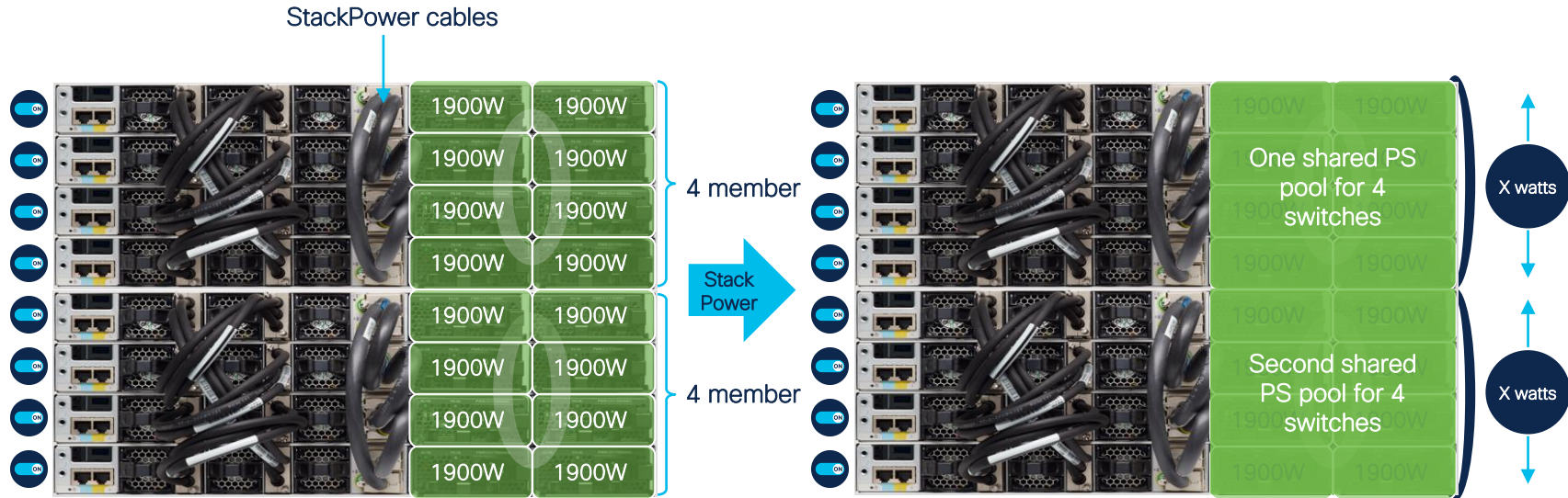
Standby PS becomes active
System enters alarm state

Failed PS   Active   Standby
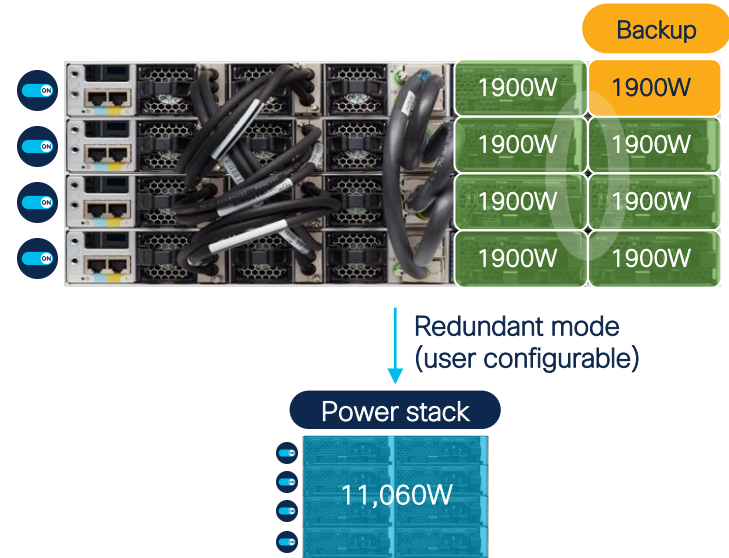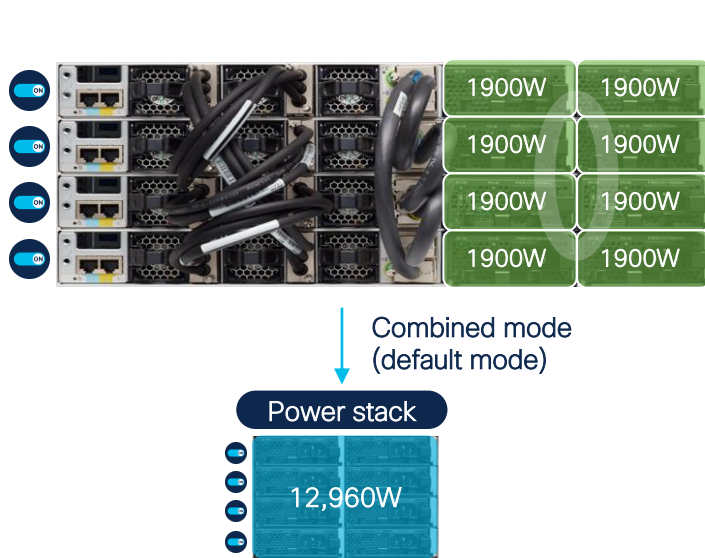
Power Redundancy with Stacking Power Supplies
StackPower

StackPower cables

1900W 1900W
1900W 1900W
1900W 1900W
1900W 1900W
4 member

1900W 1900W
1900W 1900W
1900W 1900W
1900W 1900W
4 member

Stack Power

One shared PS pool for 4 switches — X watts

Second shared PS pool for 4 switches — X watts

- Pools power from all Power Supplies (PS)
- All switches in StackPower share the available power in the pool
- Each switch is given its minimum power budget

- 1+N Redundancy with inline power
- Up to 4 switches in one StackPower Ring
- Multiple Power stacks possible in one data stack

Power stack modes

Combined mode
(default mode)

Power stack

12,960W

Backup

Redundant mode
(user configurable)
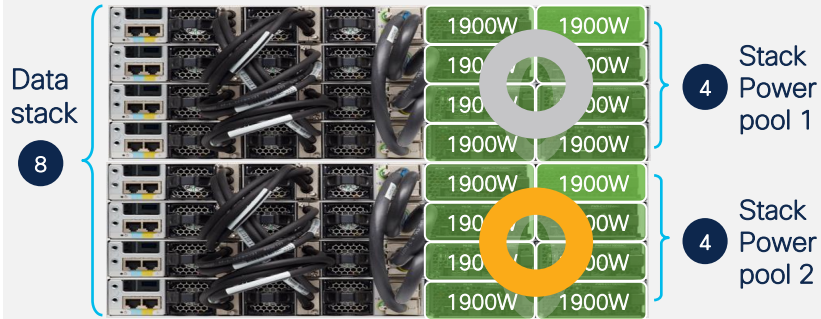
Power stack

11,060W

- All power supplies contribute to common output power budget

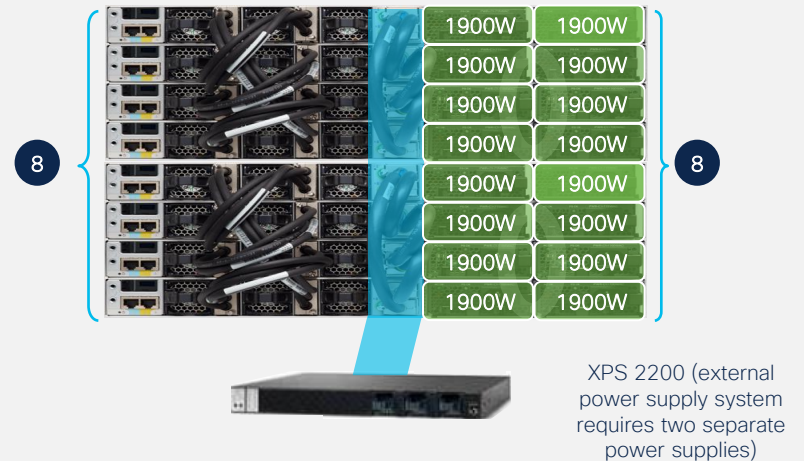- One of the highest valued power supplies is not used in the power budget calculation

Power stack deployment models

Ring model (default)

Data stack

8

1900W 1900W
1900W 1900W
1900W 1900W
1900W 1900W

4 Stack Power pool 1

1900W 1900W
1900W 1900W
1900W 1900W
1900W 1900W

4 Stack Power pool 2

Power sharing and redundancy are available only within individual pools

Star model (requires XPS 2200 device)

8

1900W 1900W
1900W 1900W
1900W 1900W
1900W 1900W
1900W 1900W
1900W 1900W
1900W 1900W
1900W 1900W

8

XPS 2200 (external power supply system requires two separate power supplies)

Power sharing and redundancy extend across all eight members in the stack

PoE Redundancy at Port Level

2-event classification

- Fast power negotiation without Link Layer Discovery Protocol (LLDP)
- Physical layer negotiation < 1 second

Perpetual PoE

- Uninterrupted PoE power during control plane reboot

Fast PoE

- Bypasses Cisco IOS® control plane boot
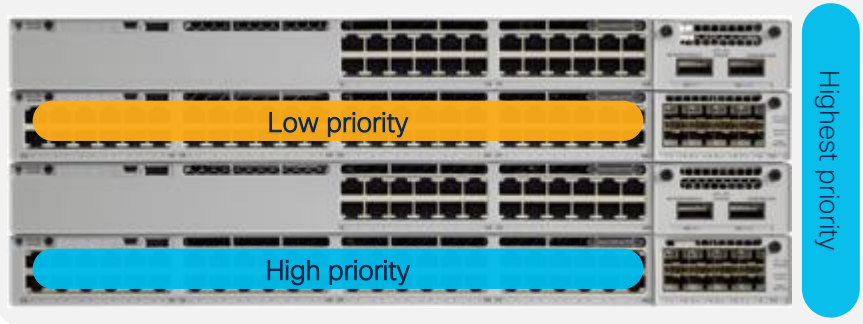- Restores power to Powered Device (PD) within 30 seconds of power resumption
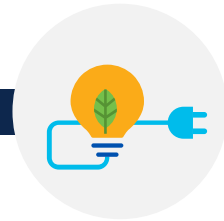
Power priority
Load shedding

Standalone mode

Low priority

High priority

Stack mode

Low priority

High priority

Highest priority

Load shedding based on configured priority

1. Low-priority ports
2. High-priority ports
3. Switch priority – **Highest priority**


```
9300-GIR-Access-144#show stack-power load-shedding switch 1
Power Stack          Stack   Stack    Total  Rsvd   Alloc  Unused Num Num
Name                 Mode    Topolgy  Pwr(W) Pwr(W) Pwr(W) Pwr(W) SW  PS
-------------------  ------  -------  ------ ------ ------ ------ --- ---
Powerstack-1         SP-PS   Stndaln  1100   30     240    830    1   1

      Power Stack            Priority Consumd Consumd Consumd Alloc Alloc
SW    Name                   Sw-Hi-Lo Sw(W)   Hi(W)   Lo(W)   Hi(W) Lo(W)
--    -------------------    -------- ------- ------- ------- ----- -----
1     Powerstack-1           4-13-22  153     0       0       0     0

Switch 1 High Priority Active (powered) Ports:

Switch 1 High Priority Inactive (unused) Ports:

Switch 1 Low Priority Active (powered) Ports:

Switch 1 Low Priority Inactive (unused) Ports:
    Gi1/0/1, Gi1/0/2, Gi1/0/3, Gi1/0/4, Gi1/0/5, Gi1/0/6,
    Gi1/0/7, Gi1/0/8, Gi1/0/9, Gi1/0/10, Gi1/0/11, Gi1/0/12,
    Gi1/0/13, Gi1/0/14, Gi1/0/15, Gi1/0/16, Gi1/0/17, Gi1/0/18,
    Gi1/0/19, Gi1/0/20, Gi1/0/21, Gi1/0/22, Gi1/0/23, Gi1/0/24,
```

Operational Resiliency with Extended Fast Software Upgrade

C9300- 17.3.2

C9300X- 17.7.1

Catalyst® 9300/9300X standalone	Catalyst 9300/9300X stack

#Install add file image activate reloadfast commit

#Install add file image activate reloadfast commit

Control plane

Data plane

Active control plane

Data plane

< 30 seconds of traffic impact

A

S

M

< 30 seconds of traffic impact for all ports in the stack

Catalyst 9000: Platform Resiliency

C9300 Fixed Platform

- **StackWise: Redundant System for high availability** with NSF/SSO
- **StackPower**: Redundant Power Supplies providing 1+ N redundancy
- **Redundant Fan & Power** Supply in case of any hardware failure

C9400/C9600 Modular Chassis

- **Redundant Supervisor:** Redundant System for high availability, simplified configuration

- **Redundant Fan & Power** Supply in case of any hardware failure

C9500 Fixed Platform

- **StackWise Virtual:** Redundant System for high availability with NSF/SSO

- **Redundant Fan & Power** Supply in case of any hardware failure

Sub-second Traffic impact during failures across C9k Family

Catalyst 9400/9600 Platform – Supervisor Redundancy

Non Stop Forwarding(NSF) or Graceful Restart

OSPF,BGP,LDP, etc

NSF Capable
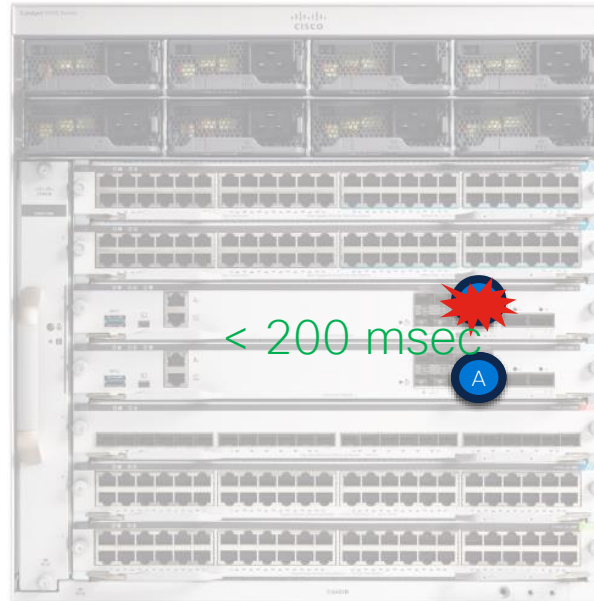
< 200 msec

A

SSO

Stateful Switchover (SSO)

SSO Aware Applications

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

SSO Compliant Applications
Routing Protocols, Netflow, etc

Dual Supervisors Required to provide both Control and Data Plane Redundancy
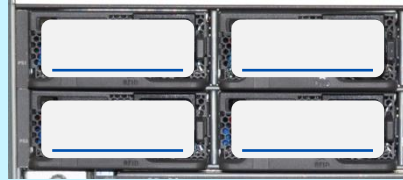
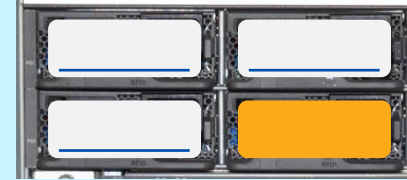# High Availability Architecture with Modular Chassis

C9600 w/Dual Supervisors

Non Stop Forwarding(NSF) or Graceful Restart

OSPF,BGP,LDP, etc

NSF Aware

NSF Aware

GR/NSF Signaling per protocol

Synchronization per protocol

NSF Capable

NSF Capable

C9400 w/Dual Supervisors

C9300 Stack
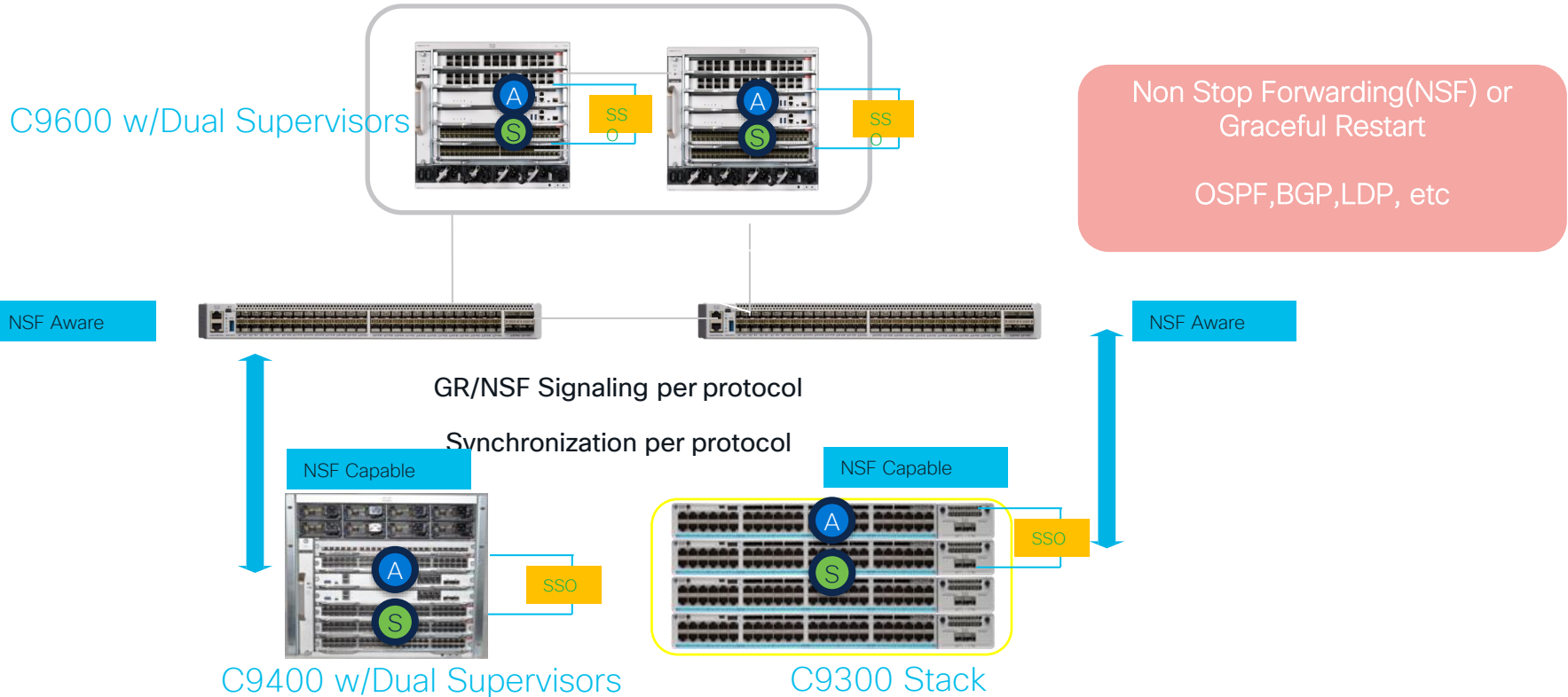
Catalyst 9400 Power Redundancy

	Normal	PS failure
Combined (default)	Load sharing on all PSs	Load sharing on functional PSs
Redundant	Load sharing on active PSs Standby PS in output disabled	Standby PS becomes active System enters alarm state

Failed PS   Active   Standby

Power redundancy: N+1 and N+N

- Default active is PS1 through 4, and standby is PS5 through 8
- Standby power slots are configurable


```
SW(config)#power redundancy-mode redundant ?
  N+N   Redundant N+N (N is active, N is standby)
  N+1   Redundant N+N (N is active, 1 is standby)
SW(config)#power redundancy-mode redundant N+1 ?
  <1-8>  standby slot in N+N mode
SWR(config)#
```

- Default active is PS1 through 7, and standby is PS8
- Standby power slot is configurable


```
SW(config)#power redundancy-mode redundant ?
  N+N   Redundant N+N (N is active, N is standby)
  N+1   Redundant N+N (N is active, 1 is standby)
SW(config)#power redundancy-mode redundant N+1 ?
  <1-8>  standby slot in N+1 mode
SWR(config)#
```

Active   Standby

Operational Resiliency with ISSU

Dual Supervisors

Start ISSU

Uplinks

SSO

ASIC

ASIC

Active Supervisor

Standby Supervisor

Line Cards

Catalyst 9400
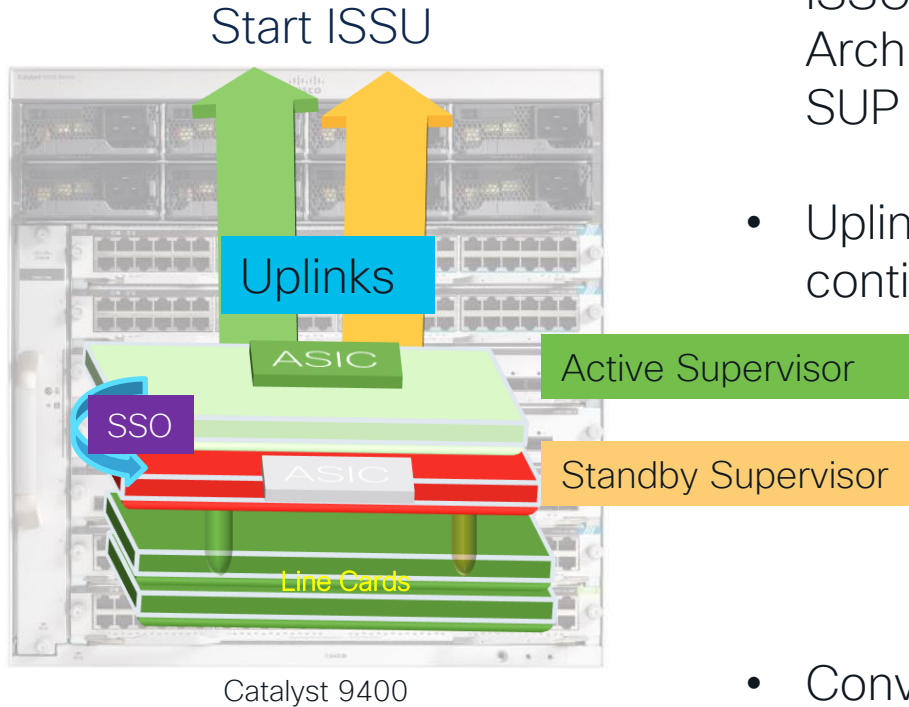
- ISSU Process leverages SSO/NSF Architecture with upgrading the Standby SUP first and then the Active SUP

- Uplinks on both active and standby SUP continue to forward traffic on C9400

- Convergence is less than 200 msec

Campus Network Architecture – Operational Resiliency
Software Upgrades

CORE/EDGE

C9600 w/Dual Supervisors

<200msec

SSO

ISSU with Dual Supervisors

DISTRIBUTION

StackWise Virtual (SVL)

<500msec

<500msec

SSO NSF

C9500

StackWise Virtual ISSU

Extended Fast Software Upgrade(xFSU)
Supported on both Standalone and Stack

<30sec

<30sec

SSO

StackWise

C9300

Dual Supervisors

SSO

C9400

ISSU with Dual Supervisors
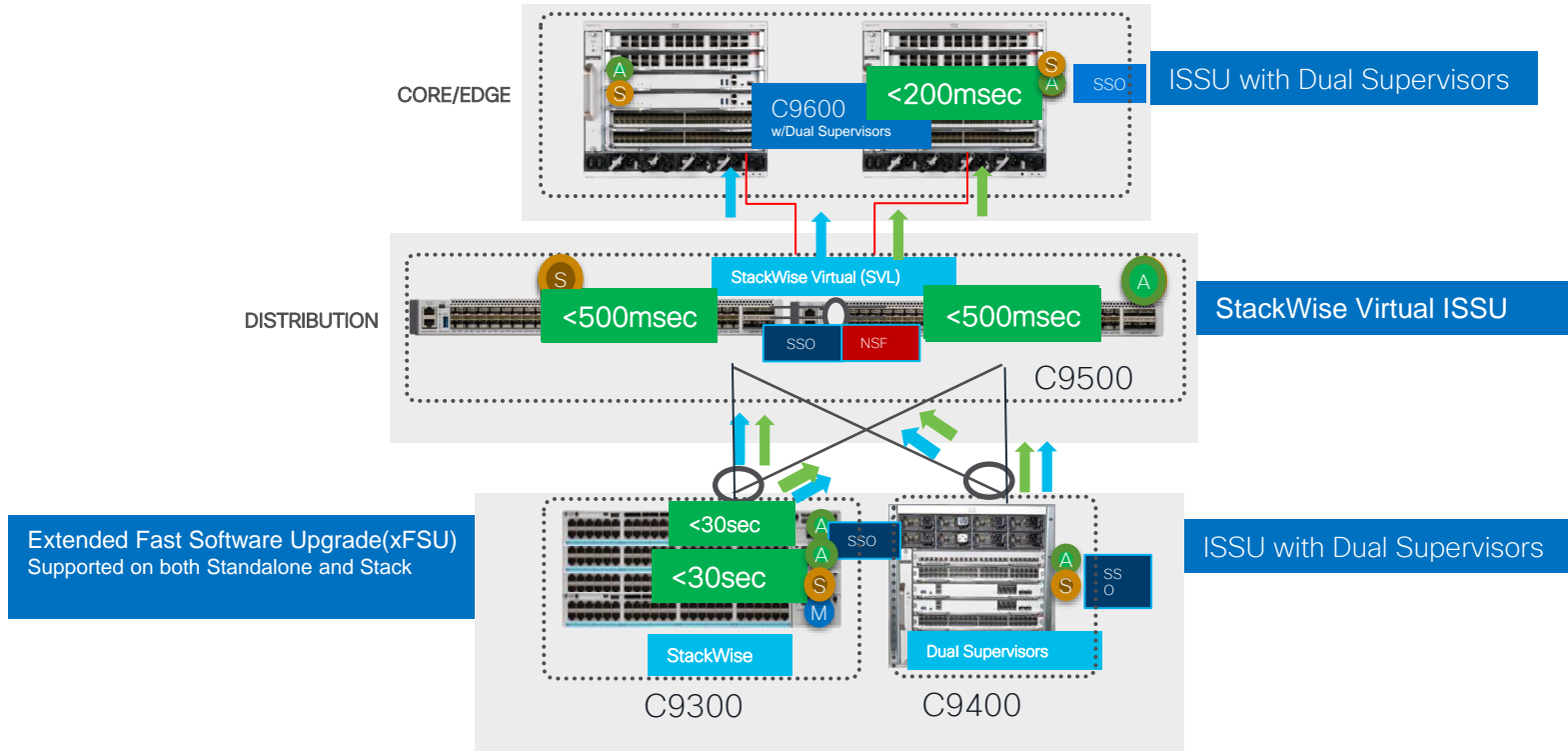
Campus Network Architecture – Operational Resiliency
ISSU Software Upgrades via DNAC

Campus Network Architecture – Operational Resiliency

Power Supply Resiliency
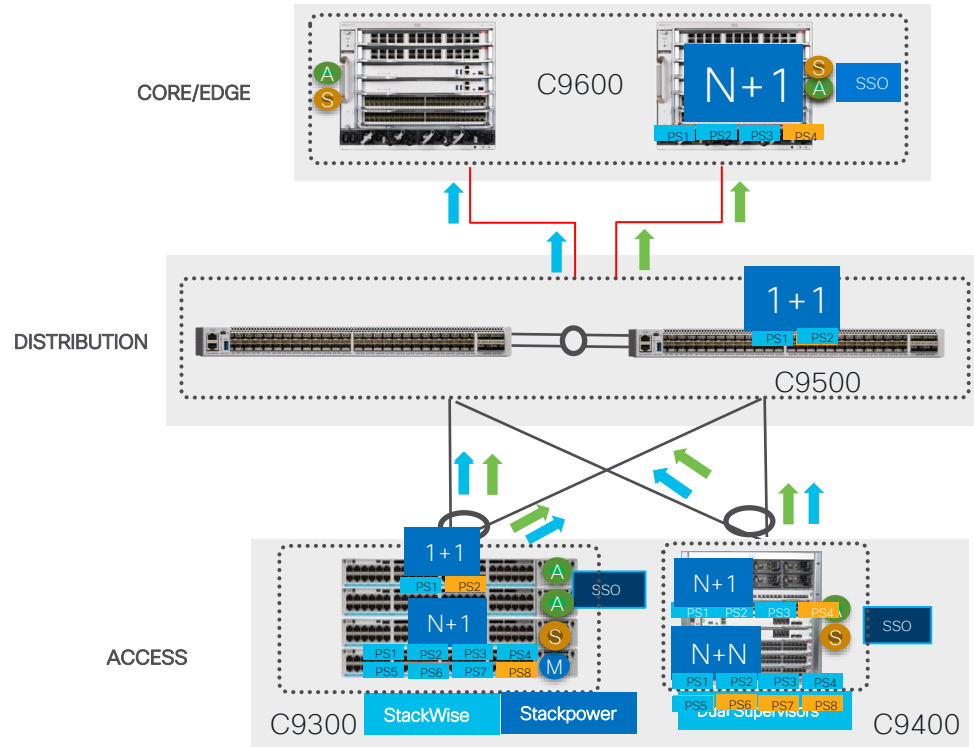
High Availability Architecture for Design Resiliency

FHRP – HSRP/VRRP

L2 + Spanning-Tree

Spanning-Tree

SSO

SSO

SSO

SSO

A

S

High Availability Architecture

FHRP – HSRP/VRRP

L2 + Spanning-Tree

Spanning-Tree

Benefits of Stackwise Virtual
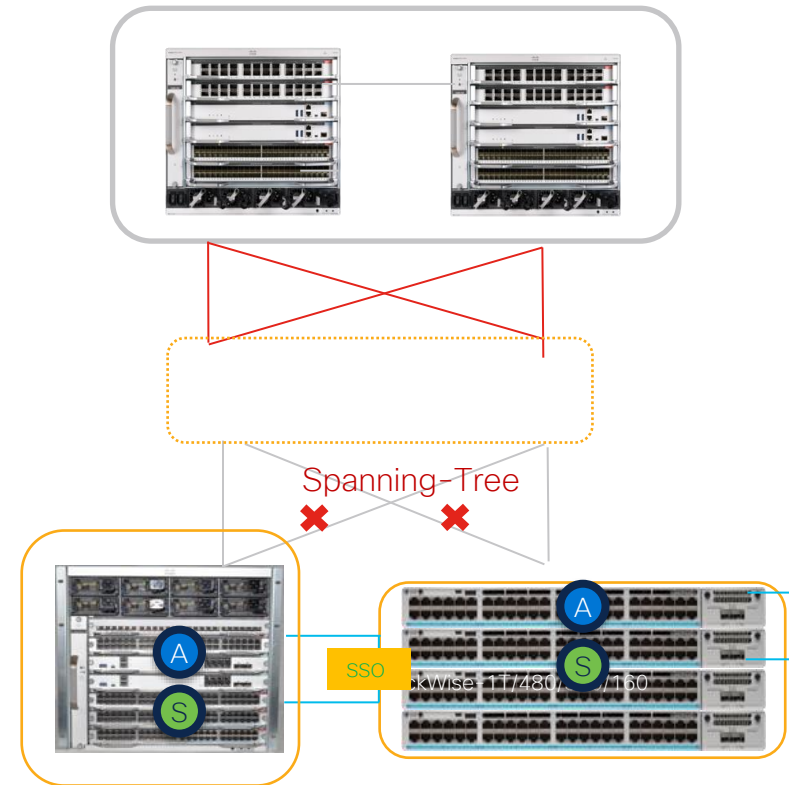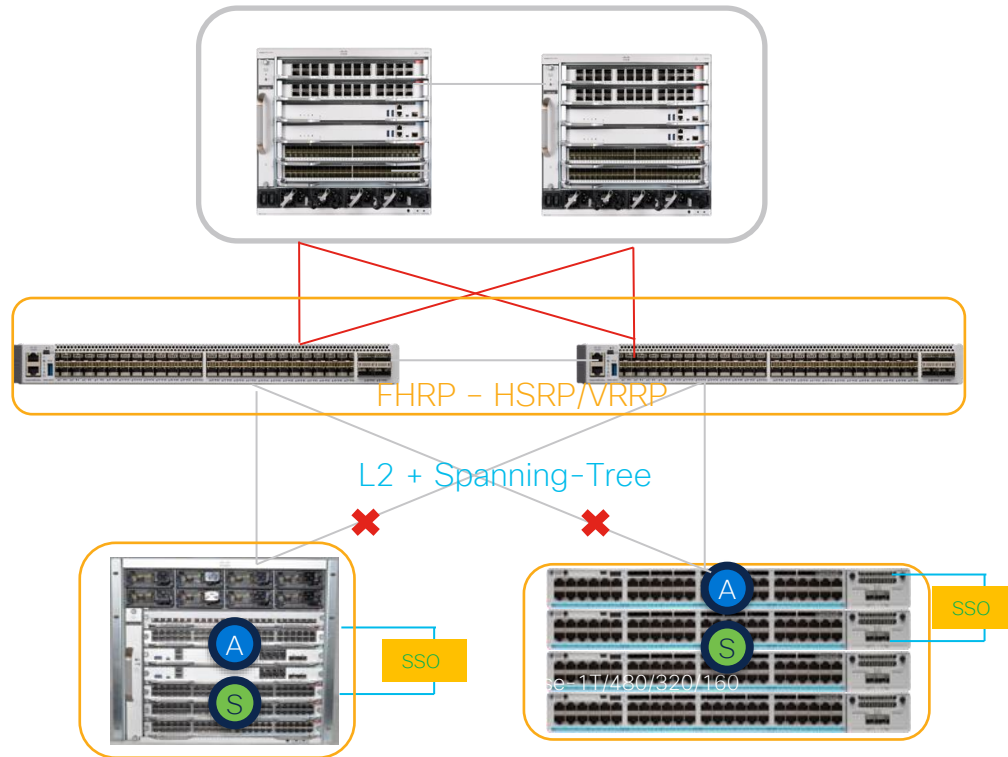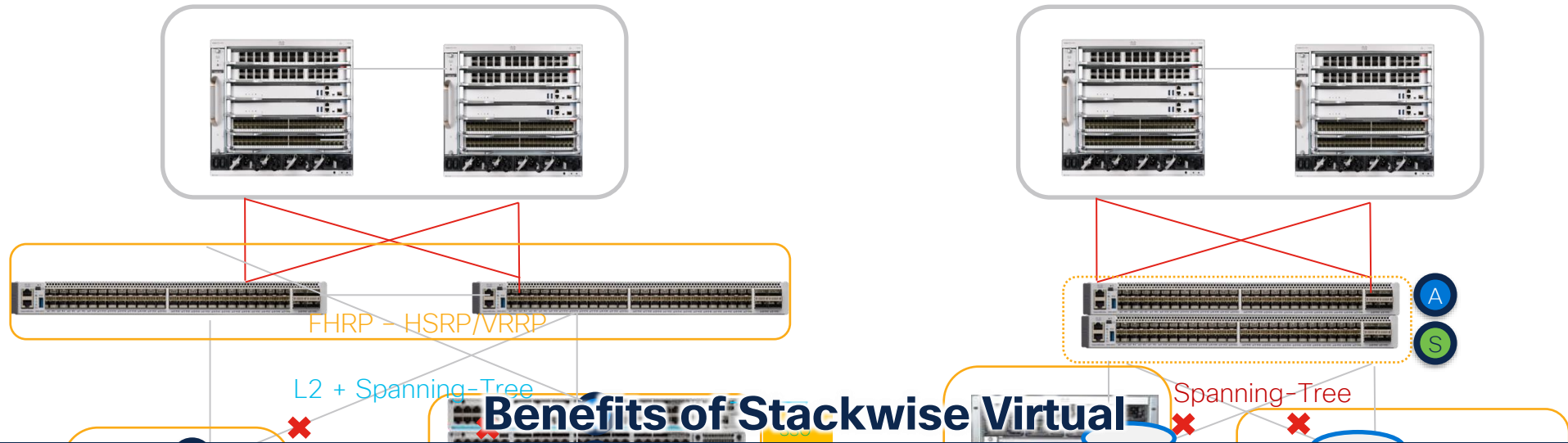
Simplify Operations by Eliminating STP, FHRP and Multiple Touch-Points
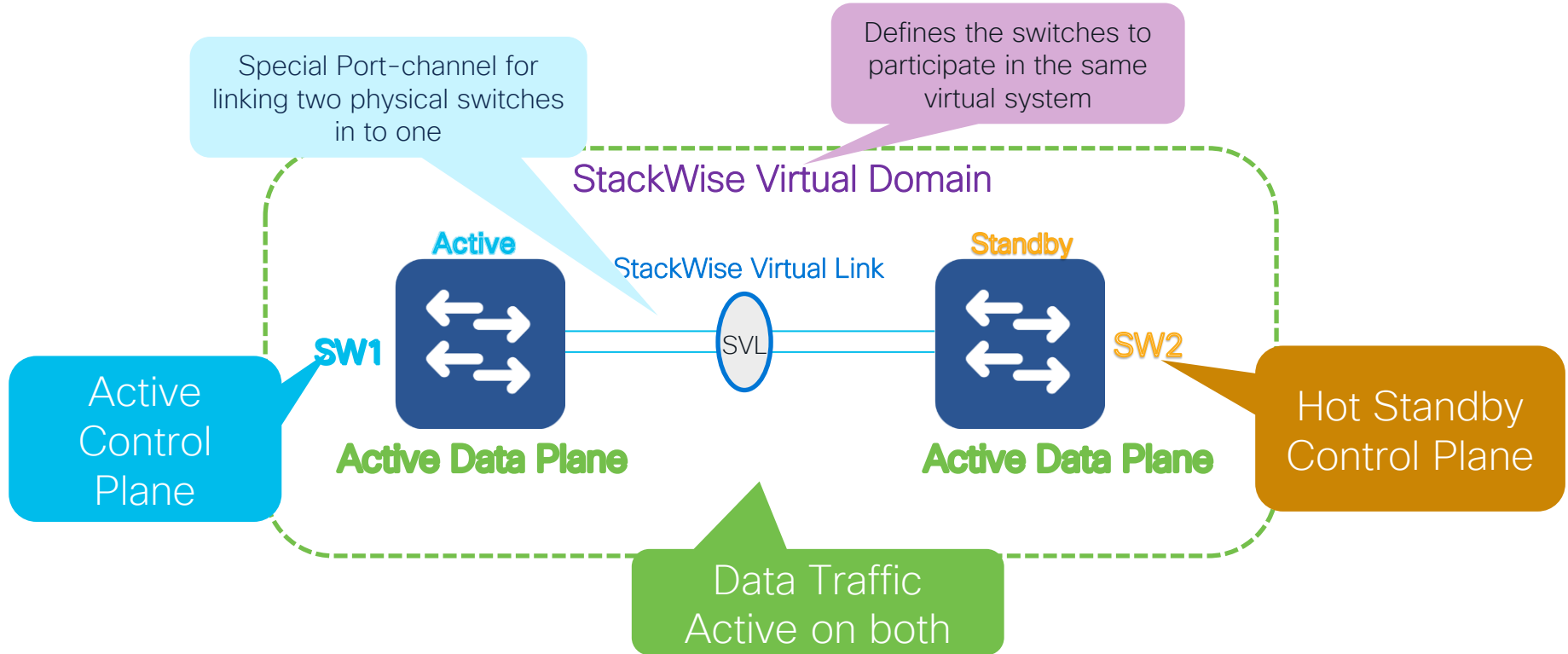
Double Bandwidth & Reduce Latency with Active-Active Multi-chassis EtherChannel (MEC)

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

StackWise Virtual

StackWise Virtual – Key Concepts

Special Port-channel for linking two physical switches in to one

Defines the switches to participate in the same virtual system

StackWise Virtual Domain

Active

Standby

StackWise Virtual Link

SVL

SW1

SW2

Active Data Plane

Active Data Plane

Active Control Plane

Hot Standby Control Plane

Data Traffic Active on both
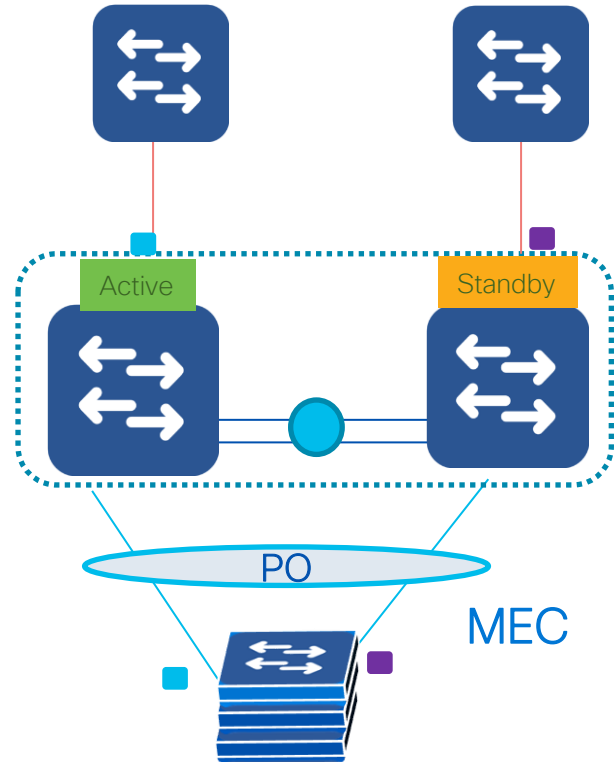
Stackwise Virtual Architecture
Active/Active Data Plane and Multi-Chassis EtherChannel

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

Multi-Chassis EtherChannel
- Port-Channel Spanning across multiple Chassis

Active

Standby

PO

MEC

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)

- StackWise Virtual EtherChannel Support
  - C9400: Support 252 port-channel with IOS-XE 16.12 or later
    - Port channel ID 127 and 128 are reserved for SVL
    - Port channel ID 1–126 and 129–252 are available for L2/L3 network configuration
  - All other C9K platforms: Support 128 port-channel
    - Port channel ID 128 is reserved for SVL
    - Port channel ID 1–127 are available for L2/L3 network configurations

SVL
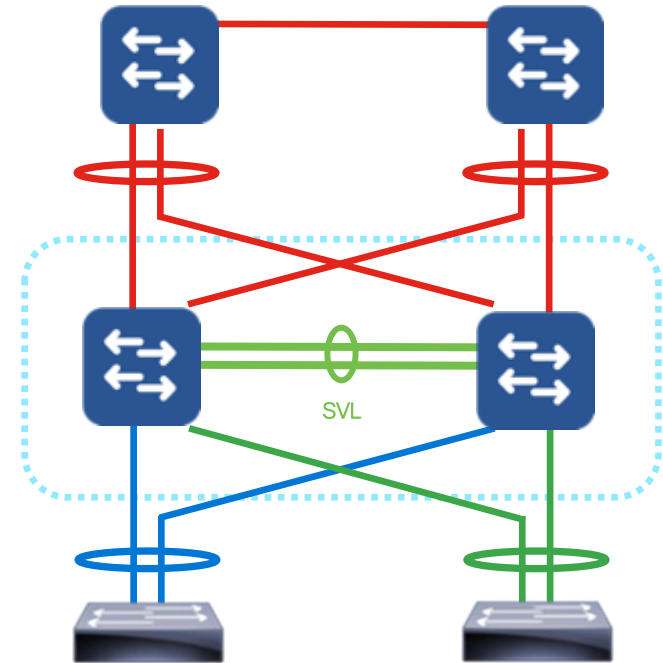
Stackwise Virtual Failure and Recovery
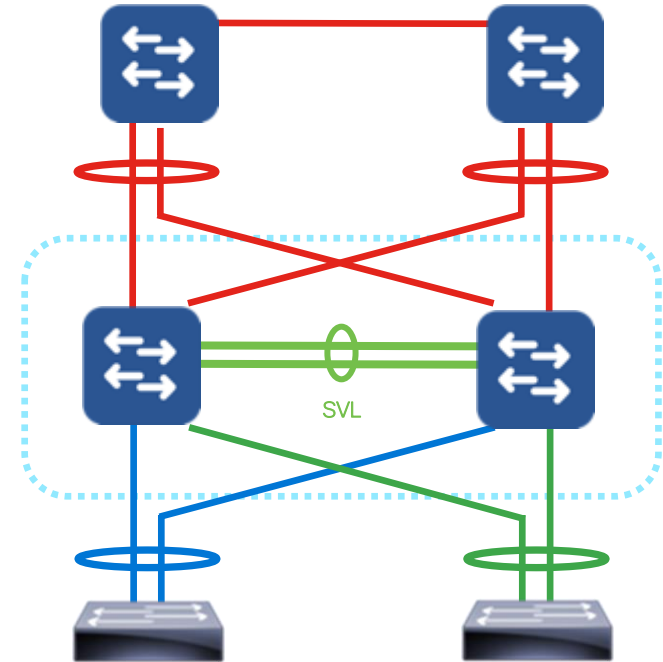
High Availability

Failure Scenario's

In a SVL Domain, one switch is elected as Active and the other as Standby

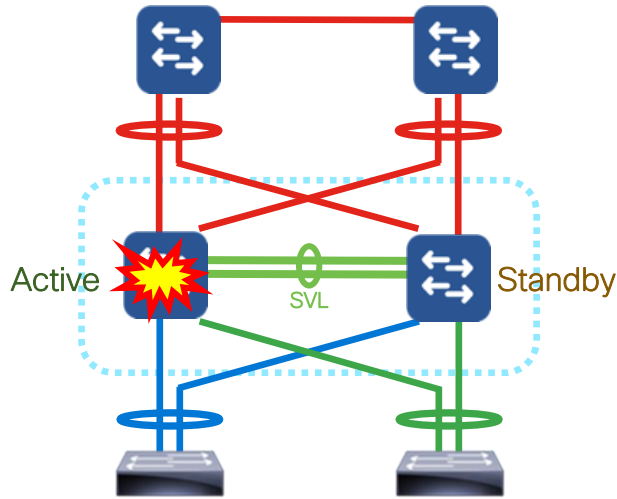All Neighbors view SVL as a single Entity, single MAC, single IP

Since the SVL is always configured as a Port Channel, the chance of the entire SVL going down is remote...
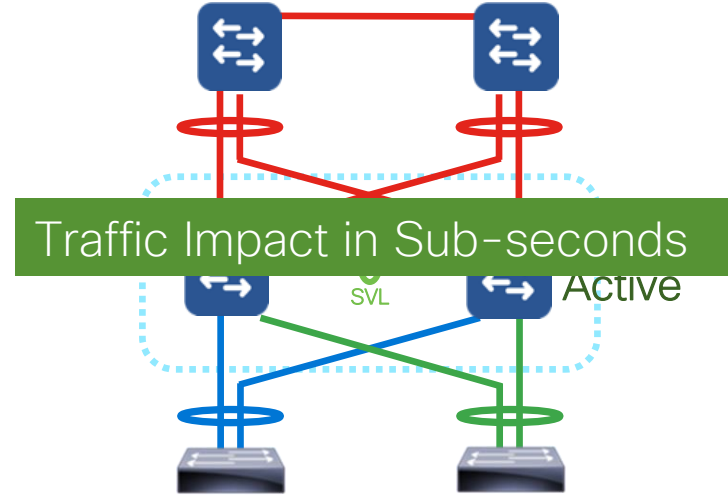
However... IT IS POSSIBLE! ☹

Recommend to deploy the SVL with **2 or more** links, distributed across entities for highest redundancy

StackWise Virtual – Inter Chassis SSO/NSF

Virtual Switch incurs a failure of the (SSO) Active Switch 1

The Standby Switch detects failure by loss of all SVL ports, or no replies to SSO keep-alive packets

Traffic Impact in Sub-seconds

The original Standby Switch now takes over as the new Virtual Switch Active

Virtual Switch initiates Graceful Restart (NSF)

Non Stop Forwarding of packets continues using hardware entries synched to Switch 2

NSF Aware neighbors exchange protocol updates with the new Virtual Switch Active

High Availability
Dual-Active Detection

If the entire SWV bundle fails, the SWV Domain will enter into a "Dual Active" scenario without DAD

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 SWV Active shuts down ALL interfaces, and enters "Recovery Mode"...** preventing further network disruption

3 **Dual-Active Recovery** - when the SWV recovers, the switch in Recovery Mode will reload to boot into a preferred standby state

Recovery Mode   Active
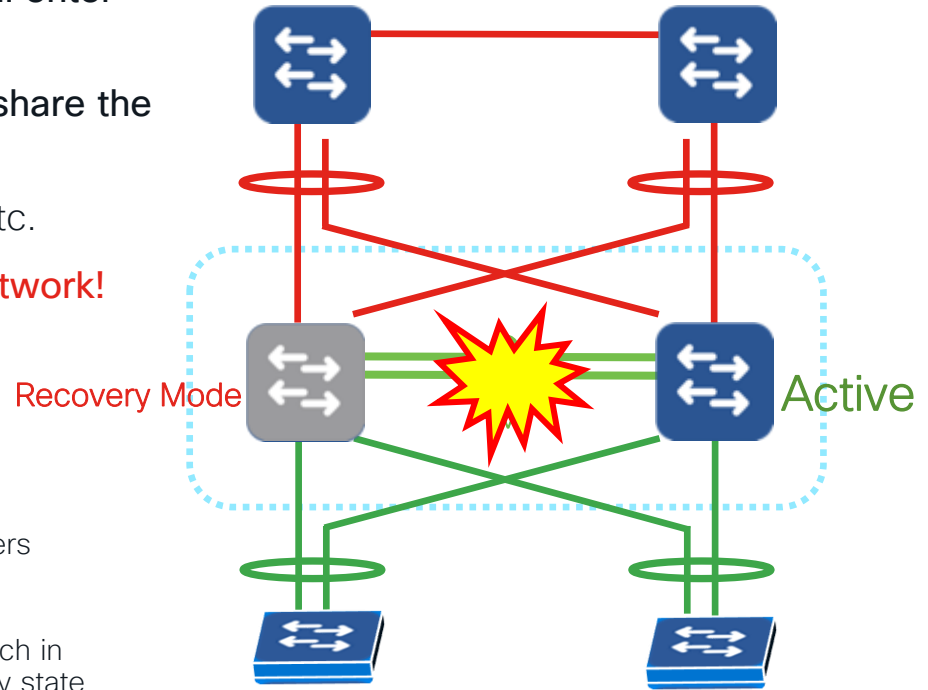
High Availability
Dual-Active Protocols

Fast Hello

Switch 1
Active

Switch 2
Standby
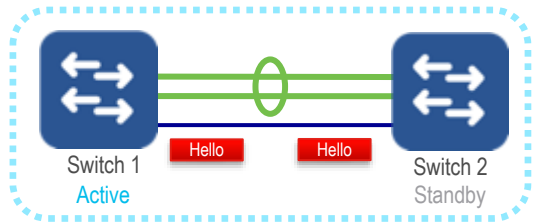
❖ Direct L2 Point-to-Point Connection

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

Enhanced PAGP

PAGP+ TLV

PAGP+ TLV

Switch 1
Active

Switch 2
Standby
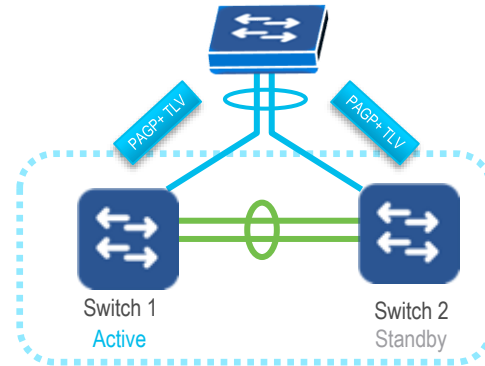
❖ Requires ePAGP capable neighbor:
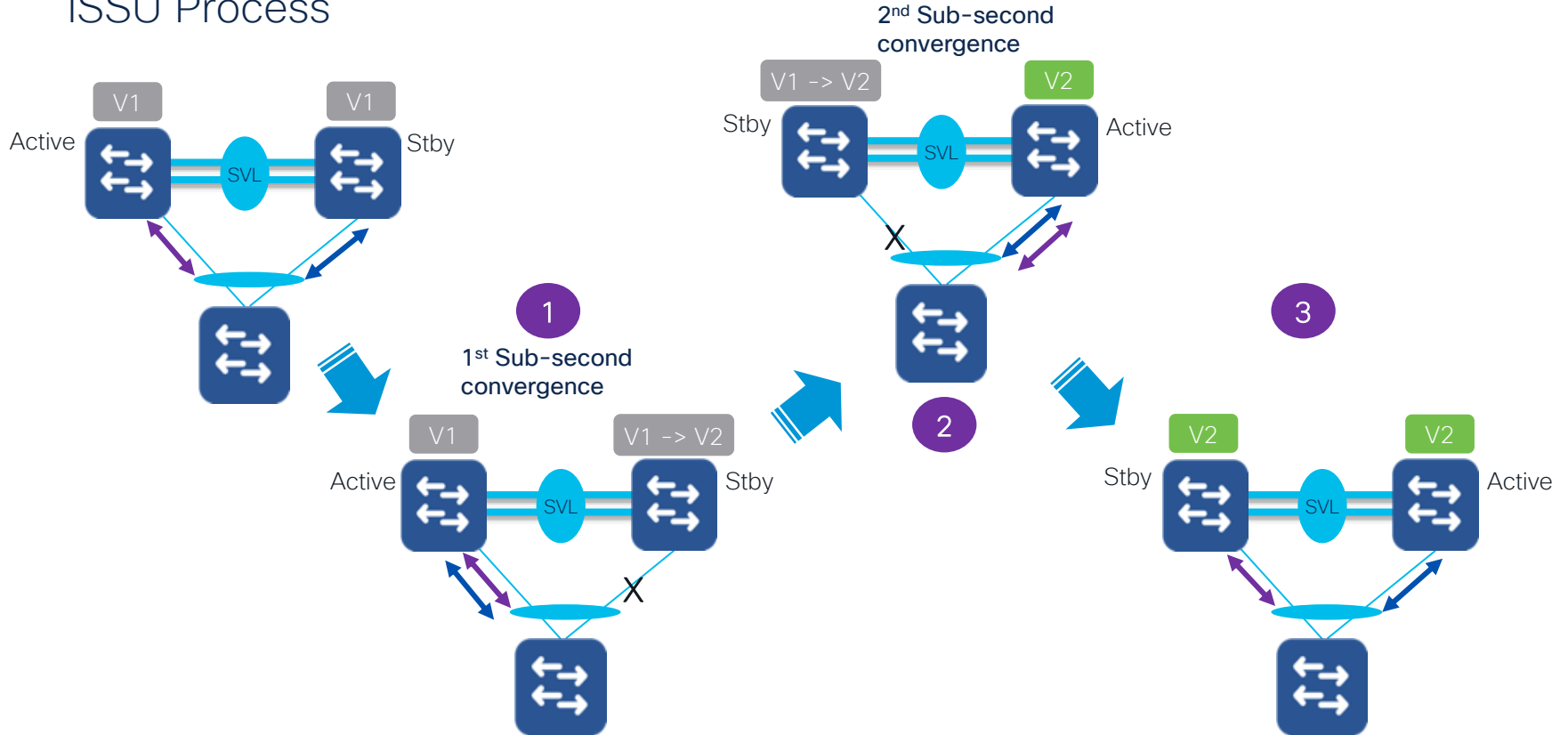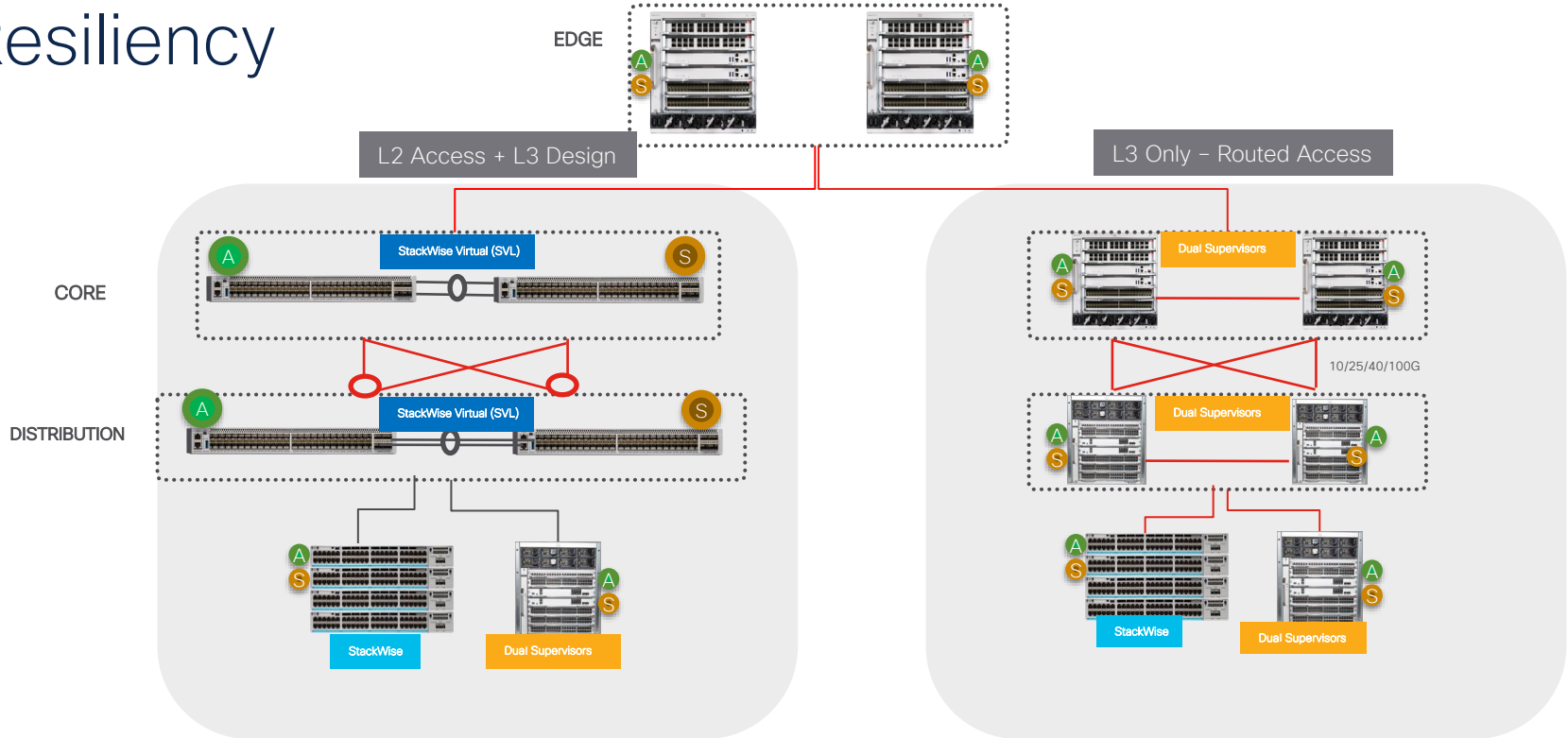
❖ Sub-Second Convergence
  ❖ Typically ~200-250ms

Stackwise Virtual ISSU

ISSU Process

Typical Campus Network Architecture – Design Resiliency

EDGE

L2 Access + L3 Design

L3 Only – Routed Access

CORE

StackWise Virtual (SVL)

Dual Supervisors

DISTRIBUTION

StackWise Virtual (SVL)

Dual Supervisors

10/25/40/100G

StackWise

Dual Supervisors

StackWise

Dual Supervisors

Typical Campus Network Architecture – Design Resiliency
Routed Access Designs

Dual Supervisors

Graceful Insertion and Removal (GIR)

Dual Supervisors

Significant Traffic Impact with routing and Topology Changes

No Traffic Impact

Routing Protocols

Dual Supervisors

StackWise

StackWise

GIR Support: ISIS, OSPF, BGP, HSRP/VRRP
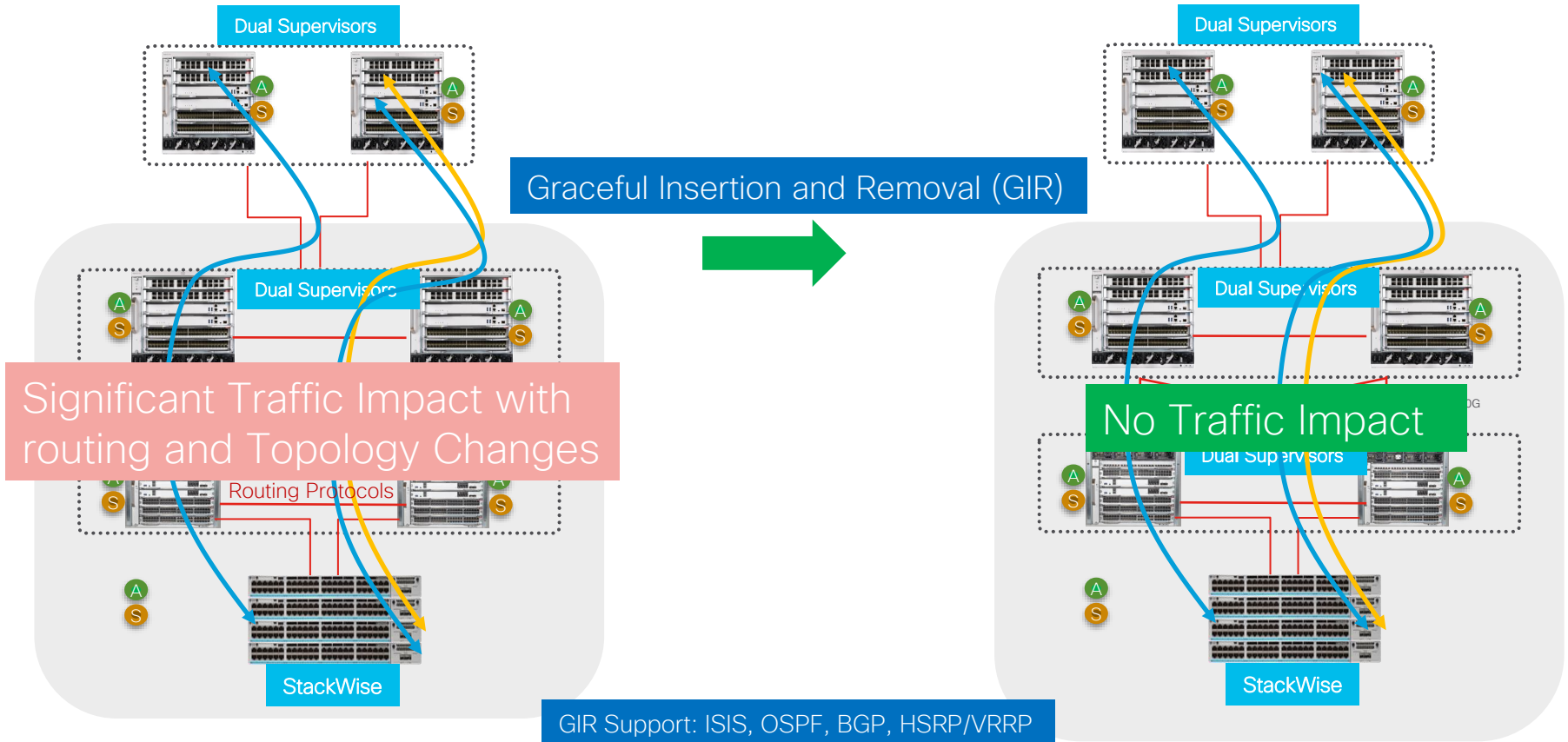
Graceful Insertion and Removal (GIR)

High Availability Architecture
Routed Access Multi-Tier Architecture

Stop Maintenance

Start Maintenance

Standalone

Standalone

Standalone

Standalone

L3

L3

A

S

SSO

A

S

SSO

A

S

SSO

A

S

SSO
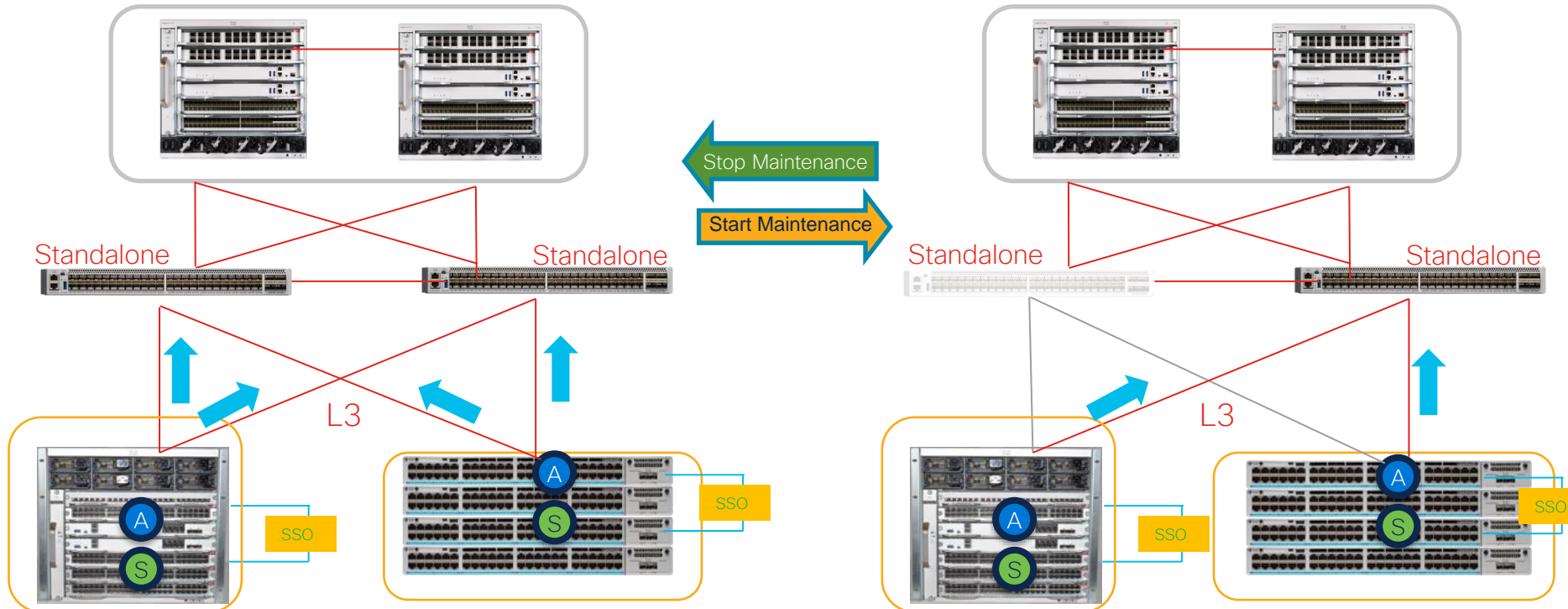
L2 and L3 Topology with GIR Isolation


```
9300#start maintenance
Template default will be applied.
Do you want to continue?[confirm]
*Mar 25 17:43:20.162: %MMODE-6-
MMODE_CLIENT_TRANSITION_START: Maintenance Isolate
start  for router isis 1
*Mar 25 17:43:50.213: %MMODE-6-
MMODE_CLIENT_TRANSITION_COMPLETE: Maintenance Isolate
complete  for router isis 1
*Mar 25 17:43:50.213: MMODE-6-
MMODE_CLIENT_TRANSITION%_START: Maintenance Isolate
start  for shutdown l2
*Mar 25 17:44:20.214: %MMODE-6-
MMODE_CLIENT_TRANSITION_COMPLETE: Maintenance Isolate
complete  for shutdown l2
*Mar 25 17:44:20.214: %MMODE-6-MMODE_ISOLATED: System
is in Maintenance
```

Set-overload-bit

Set-overload-bit

Set-overload-bit

Order for Maintenance:

BGP -> IGPs in parallel (ISIS) -> L2
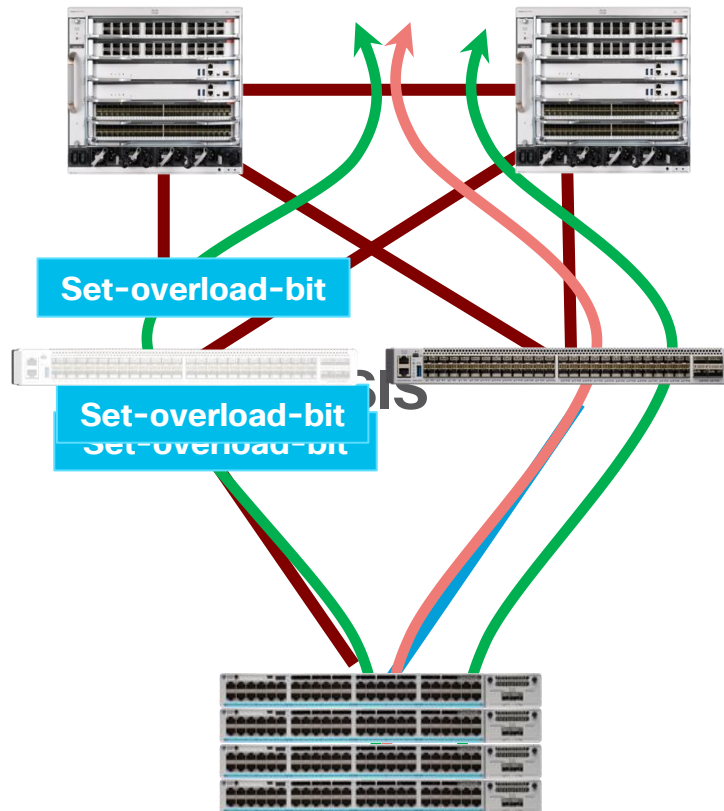
L2 and L3 Topology with GIR Isolation

```
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
```


No set-overload-bit
No set-overload-bit

ISIS

Order for Maintenance:

L2 → IGPs in parallel (ISIS) -> BGP
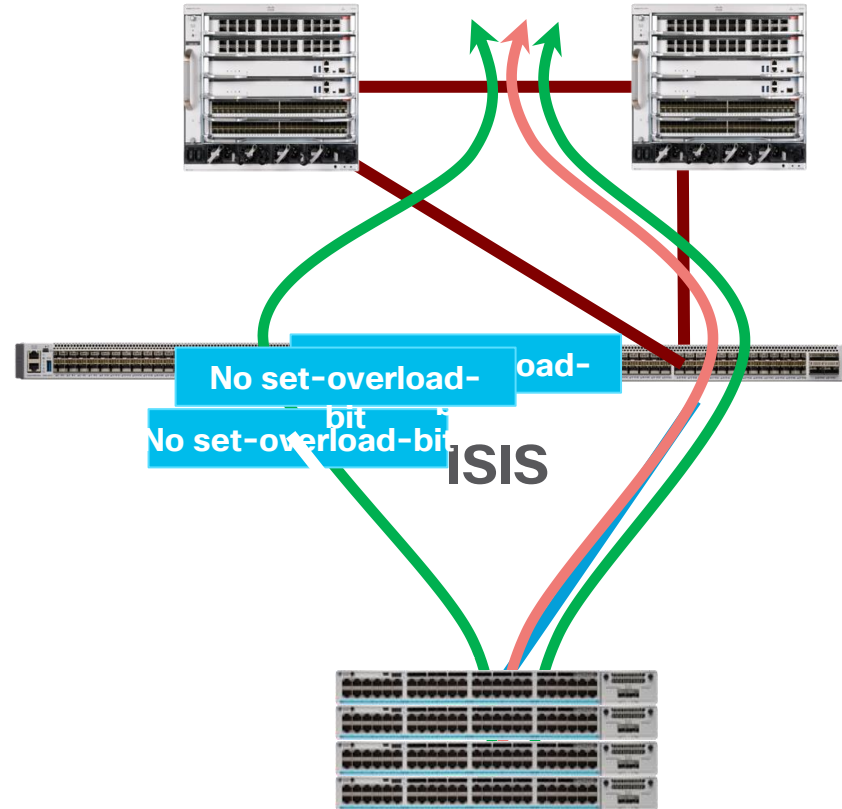
Ready for Software Patching

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.

Install Add → Install Activate → Install Commit

Catalyst 9000: Platform Resiliency

C9300 Fixed Platform

- **StackWise: Redundant System for high availability** with NSF/SSO
- **StackPower**: Redundant Power Supplies providing 1+ N redundancy
- **Redundant Fan & Power** Supply in case of any hardware failure

C9400/C9600 Modular Chassis

- **Redundant Supervisor:** Redundant System for high availability, simplified configuration

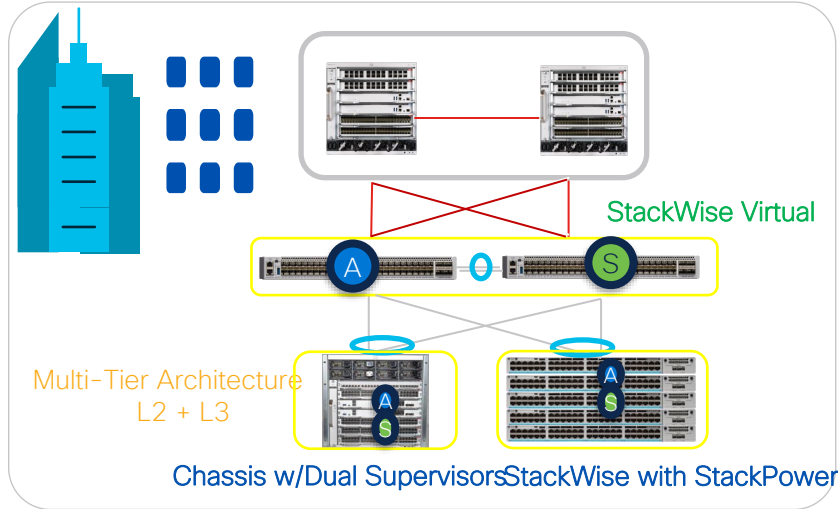- **Redundant Fan & Power** Supply in case of any hardware failure

C9500 Fixed Platform

- **StackWise Virtual:** Redundant System for high availability with NSF/SSO

- **Redundant Fan & Power** Supply in case of any hardware failure

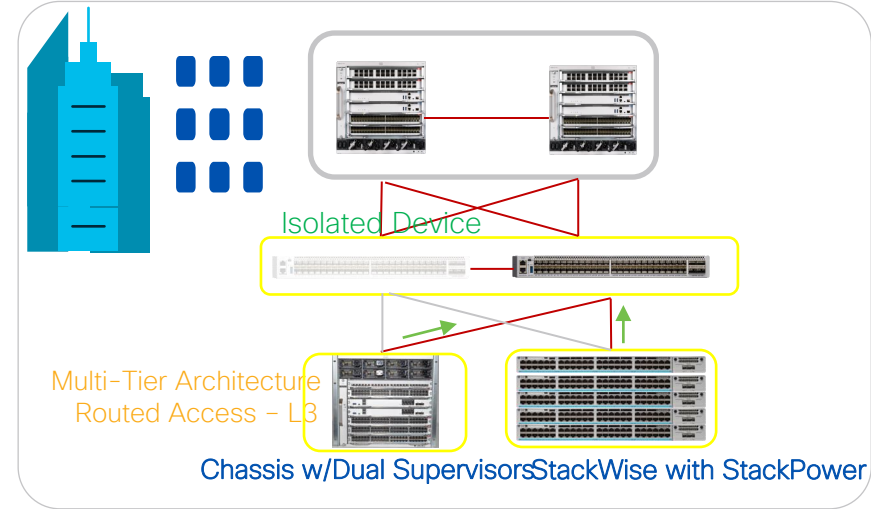## Sub-second Traffic impact during failures across C9k Family

Catalyst 9000: Features for Design Resiliency

StackWise Virtual with Backward StackWise

StackWise Virtual

Multi-Tier Architecture
L2 + L3

Chassis w/Dual Supervisors StackWise with StackPower

Device Resiliency with Simplified Design

Graceful Insertion and Removal (GIR)

Isolated Device

Multi-Tier Architecture
Routed Access – L3

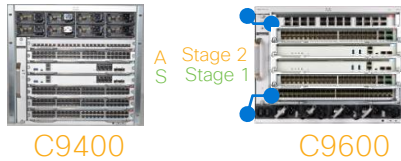Chassis w/Dual Supervisors StackWise with StackPower

Device Isolation without Traffic Impact

Device Resiliency and Sub-second failover for Every Design

Catalyst 9000: Operational Resiliency

C9400/C9600 Modular Chassis

C9400/C9500/C9600 Platform

C9300/C9300L/C9300X

Standalone Mode

StackWise Virtual Mode

Standalone and StackWise Mode

C9400

C9600

A
S

Stage 2
Stage 1

A

S

A Stage 2

In-Chassis Dual Supervisor ISSU

StackWise Virtual ISSU

Extended Fast Software Upgrade (xFS

< 200 msec Traffic Impact

< 1 sec Traffic Impact

< 30 sec Traffic Impact

Requires Dual Supervisor

Hot and Cold Patching for IOS Resiliency

Sub-second Traffic impact during software upgrades across C9k Family

Platform/Solution Support

Features	C9200	C9300/X	C9400/X	C9500	C9500X	C9600	C9600X
StackWise	✅	✅	NA	NA	NA	NA	NA
StackWise Virtual (SVL)	NA	NA	✅	✅	✅	✅	✅
SVL ISSU	NA	NA	✅	✅	✅	✅	Roadmap
Dual Supervisor ISSU	NA	NA (xFSU with stacking)	✅	NA	NA	✅	
Cold/Hot Patching	Cold Patching	✅	✅	✅	✅	✅	✅
GIR	NA	✅	✅	✅	✅	✅	✅
Power Redundancy	✅	✅	✅	✅	✅	✅	✅

Complete your Session Survey

- Please complete your session survey after each session. Your feedback is important.

- Complete a minimum of 4 session surveys and the Overall Conference survey (open from Thursday) to receive your Cisco Live t-shirt.

- All surveys can be taken in the Cisco Events Mobile App or by logging in to the Session Catalog and clicking the "Attendee Dashboard" at
https://www.ciscolive.com/emea/learn/sessions/session-catalog.html

Continue Your Education

Visit the Cisco Showcase for related demos.

Book your one-on-one Meet the Engineer meeting.

Attend any of the related sessions at the DevNet, Capture the Flag, and Walk-in Labs zones.

Visit the On-Demand Library for more sessions at ciscolive.com/on-demand.

Thank you