4

Deploying High Availability in Campus

The requirement for network reliability and availability is not a new demand, but one that must be well planned for during the early network design phase. To prevent catastrophic network failures and network outages, it is important to identify network fault domains and define rapid recovery plans to minimize application impact during minor and major network outages.

Because every tier of the LAN network design can be classified as a fault domain, deploying a strong campus network foundation with redundant system components and a resilient network design becomes highly effective for non-stop borderless services operation and business communication. However this introduces a new set of challenges, such as higher cost and the added complexity of managing more systems. Network reliability and availability can be simplified using several Cisco high availability technologies that offer complete failure transparency to end users and applications during planned or unplanned network outages.

Cisco high availability technologies can be deployed based on whether platforms have a critical or non-critical role in the network. Some of the high availability techniques can be achieved with the campus network design inherent within the borderless enterprise

network design, without making major network changes. However the critical network systems that are deployed in the main campus that provide global connectivity may require additional hardware and software components to provide uninterrupted communication. The following three major resiliency requirements encompass most of the common types of failure conditions; depending on the LAN design tier, the resiliency option appropriate to the role and network service type must be deployed:

- *Network resiliency*—Provides redundancy during physical link failures, such as fiber cut, bad transceivers, incorrect cabling, and so on.
- *Device resiliency*—Protects the network during abnormal node failure triggered by hardware or software, such as software crashes, a non-responsive supervisor, and so on.
- *Operational resiliency*—Enables resiliency capabilities to the next level, providing complete network availability even during planned network outages using In Service Software Upgrade (ISSU) features.

1 Borderless Campus High-Availability Framework

Independent of the business function, the goal of the network architect should always be to build a strong, scalable, and resilient next-generation IP network. Networks that are built on these three fundamentals provide the high availability necessary to use the network as a core platform that allows you to overlay advanced and emerging technologies as well as provision non-stop network communications. The Borderless Campus network must be built on the same fundamentals, providing highly available network services for uninterrupted business operation, campus security, and the protection of campus physical assets.

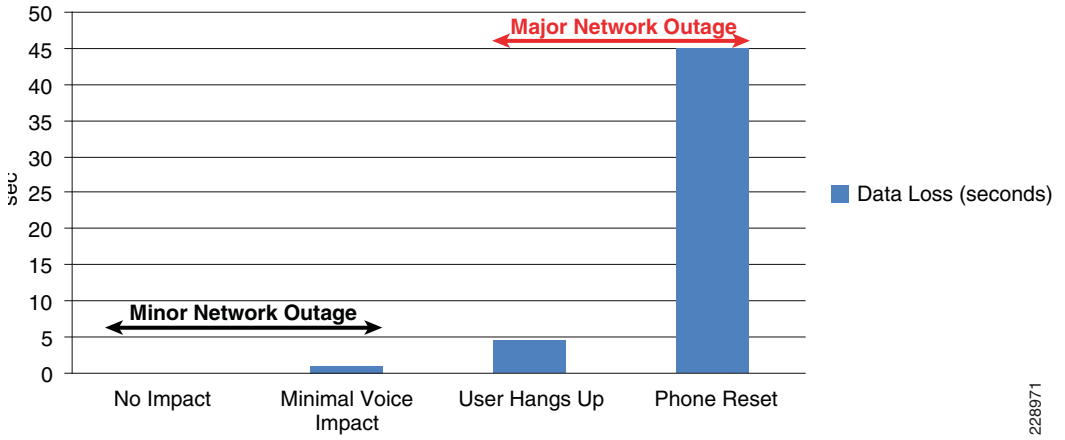
Network fault domains in this reference architecture are identifiable, but the failure conditions within the domains are unpredictable. Improper network design or non-resilient network systems can lead to more faults that not only degrade the user experience, but may severely impact application performance, such as the failure to capture critical physical security video information. The fault levels can range from network interruption to disaster, which can be triggered by the system, humans, or even by nature. Network failures can be classified in two ways:

- *Planned Failure*—A planned network outage occurs when any network system is administratively planned to be “down” for a for scheduled event (software upgrade, etc.).
- *Unplanned Failure*—Any unforeseen failures of network elements can be considered as unplanned failures. Such failures can include internal faults in the network device caused by hardware or software malfunctions, which includes software crashes, linecard or link transceiver failures, etc.

Campus High Availability Baseline

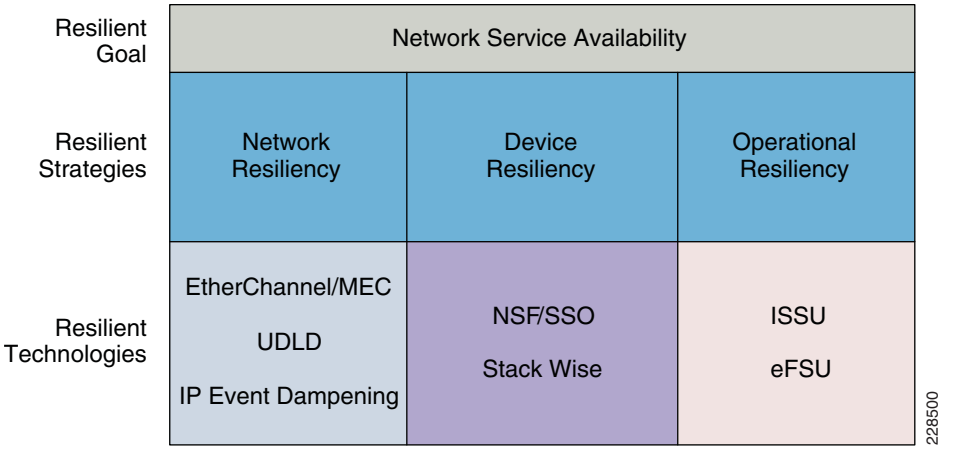
Typical application response time is measured in milliseconds when the campus network is built with high-speed backbone connections and in a fully-operational state. In deterministic network environments, users typically accomplish their work very rapidly. However, during network failures, abnormal traffic loss, congestion, and application retries impact performance and alert the user to a network problem. During major network faults, users determine network connection problem based on routine experience even before an application's protocol mechanism does (e.g,. slow Internet browsing). Protocol-based failure detection and recovery is intentional and is designed to minimize overall productivity impact and allow the network to gracefully adjust and recover during minor failures. While retries for non-critical data traffic may be acceptable, the same level of retries for applications running in real-time may not. Figure 1 illustrates some of the more typical user reactions to varying levels of real-time VoIP application outage, from minor network outages that have no user impact at all to major outages requiring a full device reset.

Figure 1 VoIP Impact During Minor and Major Network Outage



This high availability framework is based on the three major resiliency strategies to effectively mitigate a wide-range of planned and unplanned network outages. Several high availability technologies must be deployed at each layer to provide high network availability and rapid recovery during failure conditions and to prevent communication failure or degraded network-wide application performance (see Figure 2).

Figure 2 High-Availability Goals, Strategy, and Technologies



2 Network Resiliency Overview

The most common network fault occurrence in the LAN network is a link failure between two systems. Link failures can be caused by issues such as a fiber cut, miswiring, linecard module failure, and so on. In the modular platform design, the redundant parallel physical links between distributed modules in two systems reduces fault probabilities and can increase network availability. It is important to remember how multiple parallel paths between two systems also affect how higher layer protocols construct adjacencies and loop-free forwarding topologies.

Deploying redundant parallel paths in the recommended Borderless Campus design by default develops a non-optimal topology that keeps the network underutilized and requires protocol-based network recovery. In the same network design, the routed access model eliminates such limitations and enables full load balancing capabilities to increase bandwidth capacity and minimize application impact during a single path failure. To develop a consistent network resiliency service in the centralized main and remote campus sites, the following basic principles apply:

- Deploying redundant parallel paths is a basic requirement for network resiliency at any tier. It is critical to simplify the control plane and forwarding plane operation by bundling all physical paths into a single logical bundled interface (EtherChannel).
- Implement a defense-in-depth approach to failure detection and recovery. An example of this is configuring the UniDirectional Link Detection (UDLD) protocol, which uses a Layer 2 keep-alive to test that the switch-to-switch links are connected and operating correctly and acts as a backup to the native Layer 1 unidirectional link detection capabilities provided by 802.3z and 802.3ae standards. UDLD is not an EtherChannel function; it operates independently over each individual physical port at Layer 2 and remains transparent to the rest of the port configuration.

- Ensure that the network design is self-stabilizing. Hardware or software errors may cause ports to flap, which creates false alarms and destabilizes the network topology. Implementing route summarization advertises a concise topology view to the network, which prevents core network instability. However, within the summarized boundary, the network may not be protected from flooding. Deploy IP event dampening as an tool to prevent control and forwarding plane impact caused by physical topology instability.

These principles are intended to be a complementary part of the overall structured modular campus design approach and serve primarily to reinforce good resilient design practices.

3 Device Resiliency Overview

Another major component of an overall campus high availability framework is providing device- or node-level protection that can be triggered during any type of abnormal internal hardware or software process within the system. Some of the common internal failures are a software-triggered crash, power outages, line card failures, and so on. LAN network devices can be considered as a single-point-of-failure and are considered to be major failure conditions because recovery may require a network administrator to mitigate the failure and recover the system. The network recovery time can remain undeterministic, causing complete or partial network outage, depending on the network design.

Redundant hardware components for device resiliency vary between fixed configuration and modular Cisco Catalyst switches. To protect against common network faults or resets, all critical Borderless Campus network devices must be deployed with a similar device resiliency configuration. This section provides basic redundant hardware deployment guidelines at the access layer and collapsed core switching platforms in the campus network.

Redundant Power System

Redundant power supplies for network systems protect against power outages, power supply failures, and so on. It is important not only to protect the internal network system, but also the endpoints that rely on power delivery over the Ethernet network. Redundant power systems can be deployed in the following two configuration modes:

- *Modular switch*—Dual power supplies can be deployed in modular switching platforms such as the Cisco Catalyst 6500-E and 4500E Series platforms. Depending on the Cisco Nexus 7000 chassis model, it can be deployed with multiple redundant power supplies, each designed to include two isolated power units. By default, the power supply operates in redundant mode, offering the 1+1 redundant option. In modular Catalyst and Nexus switching systems, the network administrator must perform overall power capacity planning to allow for dynamic network growth with new linecard modules while maintaining power redundancy. Smaller power supplies can be combined to allocate power to all internal and external resources, but may not be able to offer power redundancy.

- *Fixed configuration switch*—Depending on the Catalyst switch, fixed configuration switches offer a wide range of power redundancy options, including the latest innovation, Cisco StackPower, in the Catalyst 3750-X series platform. To prevent network outages on fixed configuration Catalyst switches, they must be deployed with power redundancy:
  - Cisco StackPower technology on 3750-X switches
  - Internal and external redundant power supplies on Catalyst 3560-X switches

  A single Cisco RPS 2300 power supply uses a modular power supply and fan for flexibility and can deliver power to multiple switches. Deploying an internal and external power supply solution protects critical access layer switches during power outages and provides complete fault transparency and constant network availability.

Redundant Control Plane

Device or node resiliency in modular Cisco Catalyst 6500-E, Cisco Nexus 7000, 4500E, and Cisco StackWise Plus platforms provides 1+1 redundancy with enterprise-class high availability and deterministic network recovery time. The following subsections provide high availability design details, as well as graceful network recovery techniques that do not impact the control plane and provide constant forwarding capabilities during failure events.

Stateful Switchover

The stateful switchover (SSO) capability in modular switching platforms such as the Cisco Catalyst 6500-E, Nexus 7000, and 4500E provide complete enterprise-class high availability in the campus network. Cisco recommends the distribution and core layer design model to be the center point of high-availability in the enterprise network. Deploying redundant supervisors in the mission-critical distribution and core system provides non-stop communication throughout the network.

Core/Distribution Layer Redundancy

Increase network- and device-level resiliency by designing the enterprise campus to operate in a deterministic capacity, with network resiliency and the availability of rich, integrated services. The Catalyst 6500-E system running VSS mode must be deployed with a redundant supervisor module in each virtual switch chassis in the aggregation layer and backbone network. In the Cisco best practice campus design, the Cisco 6500-E system provides constant network availability and deterministic recovery with minimal application impact during supervisor switchover.

The system architecture of the Cisco Nexus 7000 system is built to deliver a lossless networking solution in large-scale enterprise campus and data center networks. Decoupling the control plane from the forwarding plane, the supervisor switchover process becomes graceful and hitless in the Cisco

Nexus 7000 system. The resilient hardware and software in the Nexus 7000 architecture is designed to protect campus network capacity and services availability using redundant components—supervisor, I/O, and crossbar fabric modules.

Access Layer Redundancy

Depending on the redundancy capabilities of the access layer system, the campus access layer may become a single-point of failure. To provide 99.999 percent service availability in the access layer, the Catalyst 4500E must be equipped with redundant supervisors to critical endpoints, such as Cisco TelePresence.

Cisco StackWise Plus is a low-cost solution to provide device-level high availability. Cisco StackWise Plus is designed with unique hardware and software capabilities that distribute, synchronize, and protect common forwarding information across all member switches in a stack ring. During master switch failure, the new master switch re-election remains transparent to the network devices and endpoints. Deploying Cisco StackWise Plus according to the recommended guidelines protects against network interruption and recovers the network in less than one second during master switch re-election.

Bundling SSO with NSF capability and the awareness function allows the network to operate without errors during a primary supervisor module failure. Users of realtime applications such as VoIP do not hang up the phone and IP video surveillance cameras do not freeze.

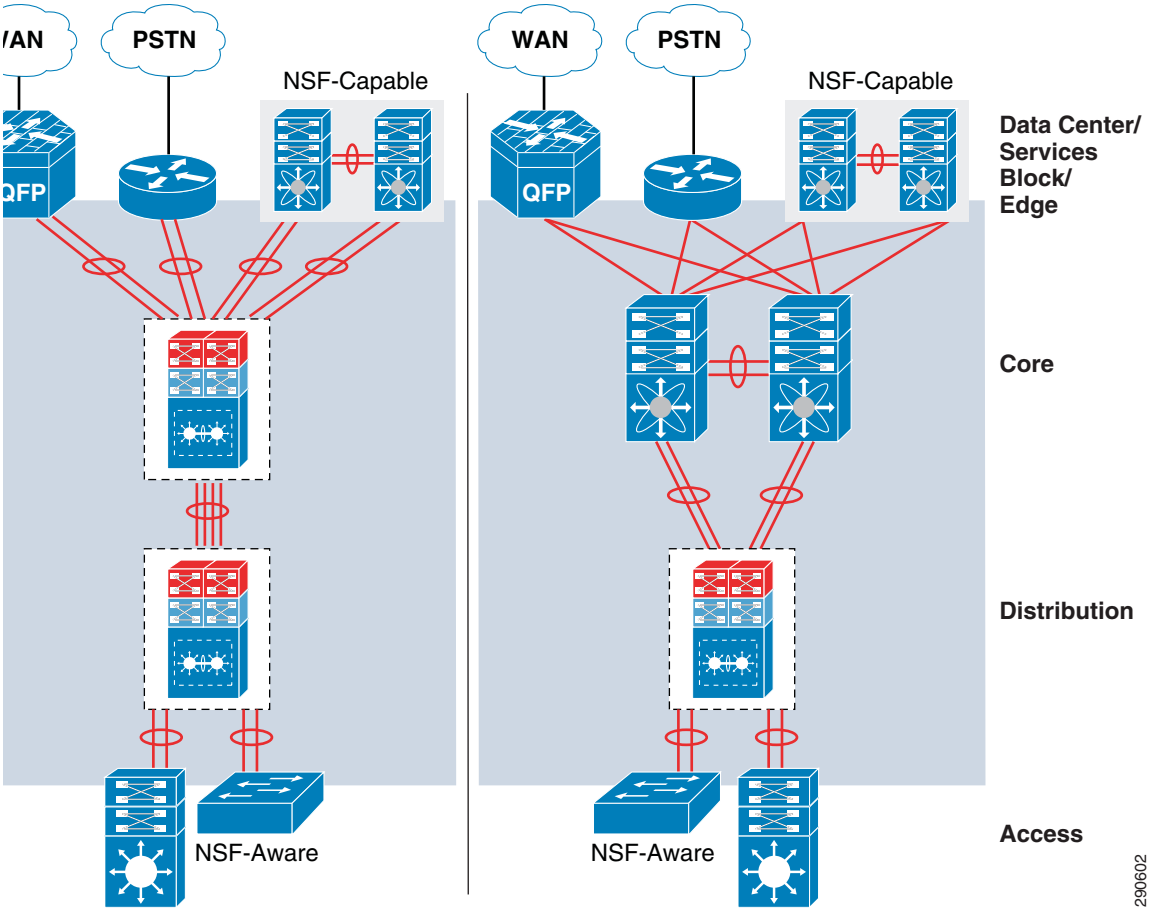
Non-Stop Forwarding

Every borderless campus recommended system deployed in redundant SSO configuration mode provides graceful protocol and network recovery during active supervisor or switch resets. The systems deployed with dual supervisor or route processors are NSF-capable systems that have the capability to initialize graceful protocol recovery with neighbors during the active supervisor or route processor reset. The neighbor system must have the NSF-Aware capability—to support the NSF-capable system to gracefully recover—by protecting routing adjacencies and topology.

It is important to enable the NSF capability for Layer 3 protocols running in a campus network. During the graceful switchover process, the new active supervisor or switch sends graceful recovery signals to neighbors to protect adjacencies and topology reset. Combining SSO with protocol intelligence using NSF technology enables graceful control plane recovery to maintain a bi-directional non-stop forwarding plane for continuous network communication.

As device redundancy is critically important in each campus network tier, the modular Cisco Catalyst and Nexus 7000 systems are designed to support NSF capability for Layer 3 unicast and multicast routing protocols. The non-modular systems, such as the Catalyst 3560-X and Cisco ISR routers, provide network-level redundancy while a SSO-capable neighbor switch is going through the recovery process. (See Figure 1-3.)

Figure 1-3 Borderless Campus NSF/SSO Capable and Aware Systems

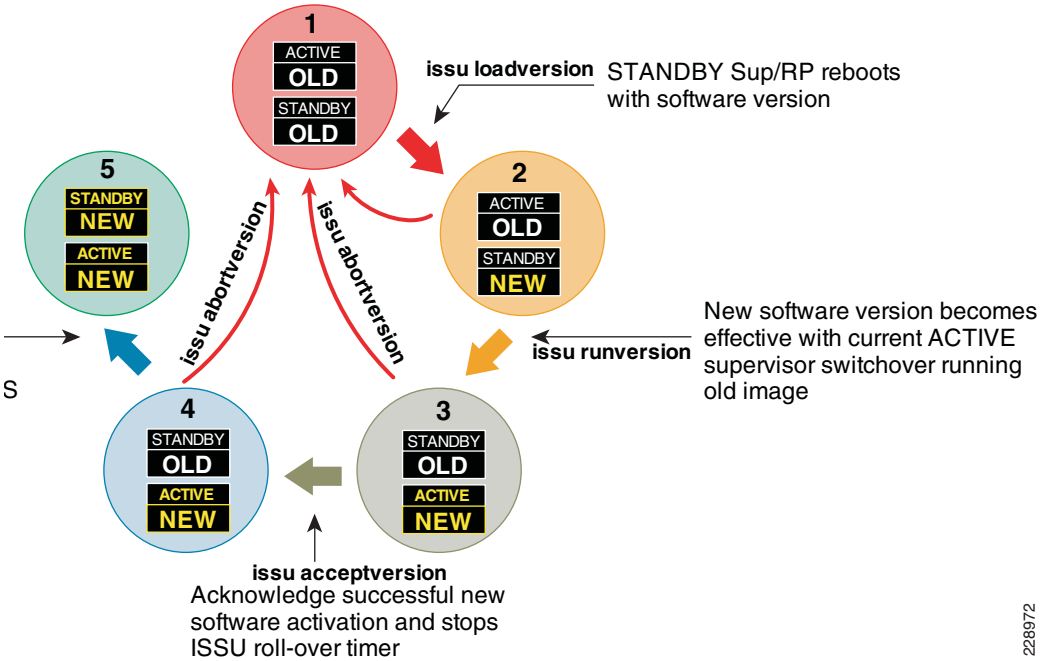


4 Operational Resiliency Overview

Designing the network to recover from failure events is only one aspect of the overall campus non-stop design. Converged network environments are continuing to move toward requiring true 7x24x365 availability. The Borderless Campus network is part of the backbone of the enterprise network and must be designed to enable standard operational processes, configuration changes, and software and hardware upgrades without disrupting network services.

The ability to make changes, upgrade software, and replace or upgrade hardware becomes challenging without a redundant system in the campus core. Upgrading individual devices without taking them out of service is similarly based on having internal component redundancy (such as with power supplies and supervisors) complemented with the system software capabilities. The Cisco Catalyst 6500-E, Nexus 7000, 4507R+E, and ASR 1000 series platforms support real-time software upgrades in the campus without introducing network downtime or impacting network availability. The Cisco In-Service Software Upgrade (ISSU) and Enhanced Fast Software Upgrade (eFSU) leverage NSF/SSO technology to provide continuous network availability while upgrading critical systems. This helps to greatly reduce the need for planned service downtime and maintenance. Figure 1-4 demonstrates the platform-independent Cisco IOS software upgrade flow process using ISSU technology.

Figure 1-4 Cisco IOS ISSU Software Process Cycle



Catalyst 4500E—ISSU

Full-image ISSU on the Cisco Catalyst 4500E leverages dual redundant supervisors to allow for a full, in-service Cisco IOS upgrade, such as moving from IOS Release 12.2(53)SG to 12.2(53)SG1. This leverages the NSF/SSO capabilities and unique uplink port capability to keep ports in an operational

and forwarding state even when supervisor module is reset. This design helps retain bandwidth capacity while upgrading both supervisor (Sup7-E, Sup6-E, or Sup6L-E) modules at the cost of less than a sub-second of traffic loss during a full Cisco IOS upgrade.

Having the ability to operate the campus as a non-stop system depends on the appropriate capabilities being designed into the network from the start. Network and device level redundancy, along with the necessary software control mechanisms, guarantee controlled and fast recovery of all data flows following any network failure, while concurrently providing the ability to proactively manage the infrastructure.

The Catalyst 4500E can perform ISSU with the following two methods:

- Manual—Follow each ISSU process as illustrated in Figure 1-4. The manual IOS upgrade mode is more attentive and requires users to upgrade IOS by manually going through each upgrade cycle. Executing each ISSU upgrade step provides flexibility for users to verify the stability of network operation and services by introducing new IOS software individually, as well as providing an option to abort the upgrade process and roll back to an older IOS version if any abnormal behavior is observed.
- Automatic—Follows the same ISSU upgrade process as illustrated in Figure 1-4. However the automatic upgrade process is the new, single-step automatic IOS-XE upgrade process that automates each ISSU step on the active and standby Sup7-E supervisor modules without user intervention. This simplified upgrade process helps network administrators of large Catalyst 4500E-based campus networks to roll out new Cisco IOS software in the network. The Catalyst 4500E Sup6-E and Sup6L-E supervisor modules currently do not support the automatic upgrade process.

Cisco recommends using both ISSU methods when upgrading the IOS software process on the Cisco Catalyst 4500E Sup7-E module in order to minimize the disruptive impact to network operation and services and upgrade the network rapidly. It is recommended that network administrators first upgrade the Catalyst 4500E Sup7-E system using manual procedures that allow verification of stability at each upgrade step. They should then identify the reliability of the new Cisco IOS version and verify that it is ready to be deployed across the campus network. Later the remainder of the Sup-7E-based systems can be upgraded using the single-step automatic ISSU upgrade procedure.

Catalyst 6500 VSS—eFSU

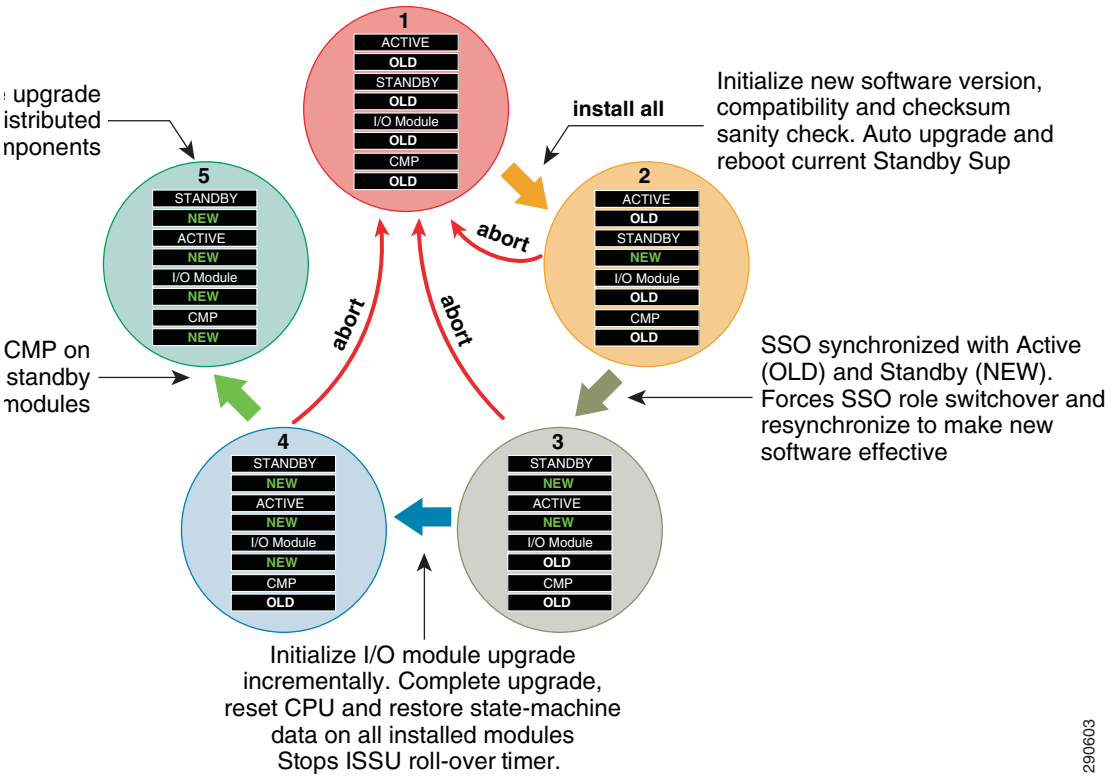
A network upgrade requires planned network and system downtime. VSS offers unmatched network availability to the core. With the Enhanced Fast Software Upgrade (eFSU) feature, VSS can continue to provide network services during an upgrade. With the eFSU feature, the VSS network upgrade remains transparent to applications and end users. Because eFSU works in conjunction with NSF/SSO technology, network devices can gracefully restore control and forwarding information during the upgrade process, while the bandwidth capacity operates at 50 percent and the data plane converges in less than one second.

For a transparent software update, the ISSU process requires three sequential upgrade events on both virtual switch systems. Each upgrade event causes traffic to be re-routed to a redundant MEC path, causing sub-second traffic loss that does not impact realtime network applications, such as VoIP.

Cisco Nexus 7000—ISSU

To provide non-disruptive network services in the campus core, the Nexus 7000 provides a simplified and resilient upgrade procedure. The distributed hardware components require software upgrades with the latest Cisco NX-OS software and protect against control plane disruption, maintaining campus backbone network availability and capacity. During a graceful software upgrade process on a dual-supervisor module, all I/O module and CMP complexes go through a five-step automatic upgrade procedure initiated by a single user step. Each step performs several non-disruptive checks to ensure the Cisco NX-OS upgrade procedure will not introduce any network instabilities. Combined with the resilient Nexus 7000 system architecture, best practice campus network design, and NSF/SSO capability, the Cisco NX-OS software upgrade process results in zero packet loss. Figure 1-5 illustrates the Cisco NX-OS ISSU-based software upgrade process.

Figure 1-5 Cisco NX-OS ISSU Software Process Cycle



5 Design Strategies for Network Survivability

Each network tier can be classified as a fault domain, with the deployment of redundant components and systems increasing redundancy and load sharing capabilities. However, this introduces a new set of challenges— namely, higher costs and increased complexity in managing a greater number of systems. Network reliability and availability can be simplified using several Cisco high-availability and virtual system technologies such as VSS, which offers complete failure transparency to end users and applications during planned or un-planned network outages. In this sense, minor and major network failures are considered broad terms that includes several types of network faults which must be taken into consideration in order to implement a rapid recovery solution.

Cisco high availability technologies can be deployed based on whether platforms have a critical or non-critical role in the network. Some of the high-availability techniques can be achieved in the campus network design without making major network changes; however, the critical network systems that are deployed in the center of the network to provide global connectivity may require additional hardware and software components to offer non-stop communication.

The network survivability strategy can be categorized using three major resiliency requirements that can encompass most of the common types of failure conditions. Depending on the network system tier, role, and network service types, the appropriate resiliency option must be deployed (see).

Table 1-1 Borderless Campus Network High Availability Strategy

Platform	Role	Network Resiliency	Device Resiliency	Operational Efficienc
Catalyst 3560-X	Access	EtherChannel UDLD Dampening	RPS 2300	None. Standalone Sy
Catalyst 3750-X				
Catalyst 3750-X StackWise Plus			Cisco StackPower NSF-Capable and Aware	StackWise Plus
Catalyst 4500E	Access		Redundant Power Supplies Redundant Linecard Modules Redundant Supervisor Modules SSO/NSF Capable and Aware[1]	ISSU
	Distribution			
	Core			
Catalyst 6500-E	Distribution			VSS eFSU
	Core			
Nexus 7000	Core	EtherChannel UDLD	Redundant Power Supplies Redundant Linecard modules Redundant Crossbar Fabric Module Redundant Supervisor modules SSO/NSF Capable and Aware	ISSU

Table 1-1 Borderless Campus Network High Availability Strategy

ASR 1006	WAN Edge	EtherChannel Dampening	Redundant Power Supplies Redundant ESP modules Redundant Route Processors SSO/NSF Capable and Aware	ISSU
ASR 1004	Internet Edge		Red. Power Supplies SSO/NSF Capable and Aware[2]	ISSU

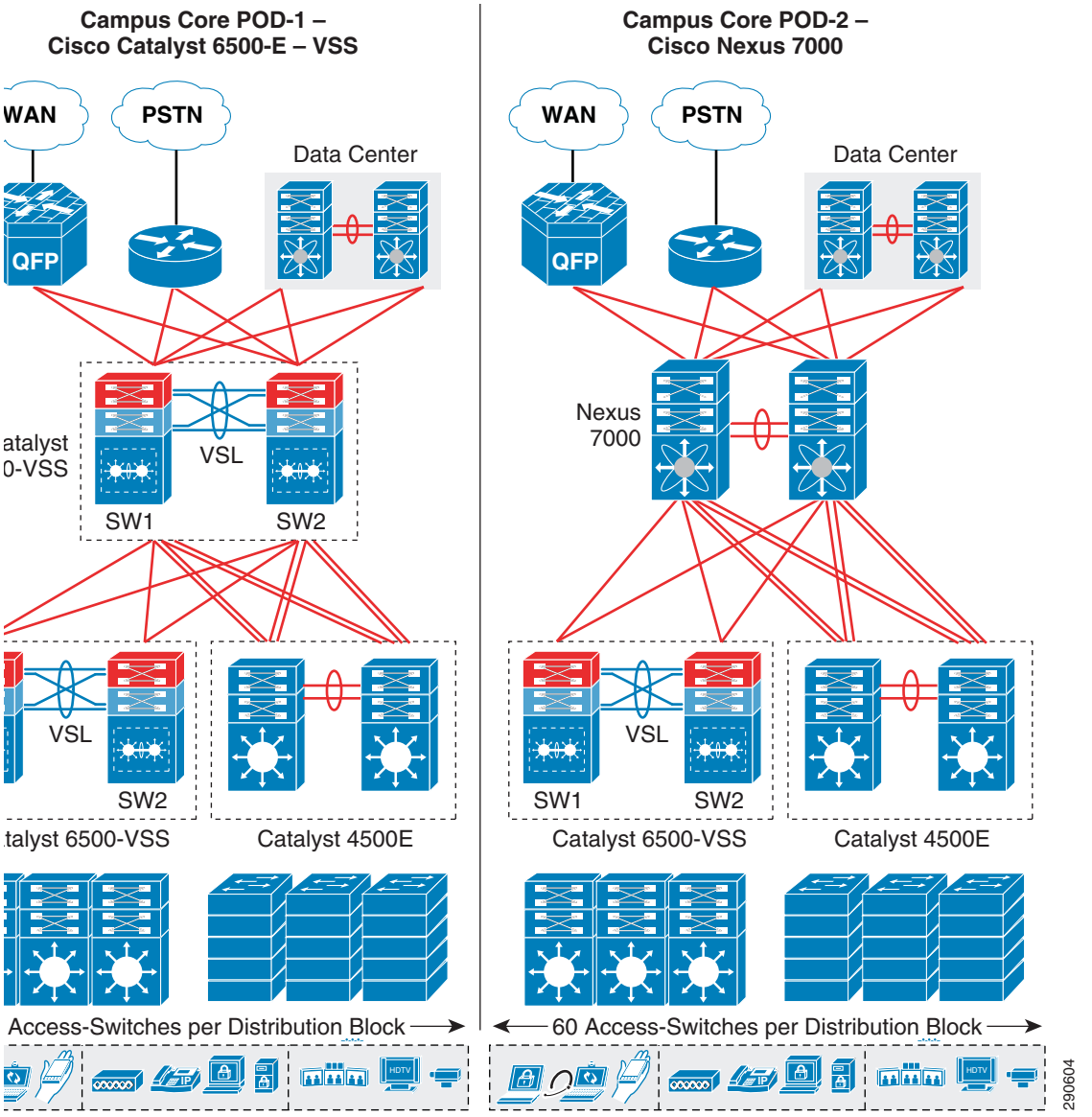
1. Redundant quad supervisor per VSS Domain (two per virtual switch node basis)and dual supervisor module on Catalyst 4500 chassis.
2. Software-based SSO Redundancy.

6 Borderless Campus Design Validation

This design guide validates the operation, integration, and performance of an end-to-end reference borderless campus network architecture. Evaluating network and application performance through thorough and rigorous testing in a Cisco solution lab derives the recommended network design, systems, technologies, and best practices. To align with real-world large enterprise networks and understand the impact to end points and application in an end-to-end, large scale environment, the Cisco solution lab is equipped with a large number of real-time and non-real time devices, such as IP phones, TelePresence units, PCs, laptops, etc.

Previous chapters provided guidance on deploying key foundational technologies and optimizing application performance with QoS techniques. This chapter provides strategies and guidance on building a resilient campus network design. To meet the campus high-availability baseline—enabling real-time applications such as unified and video communication—this document provides validated results obtained by inducing faults on system components and measuring the impact on the network and applications. Figure 1-6 illustrates a sample solution lab network topology for a large campus network design based on a reference architecture.

Figure 1-6 Sample Large Campus Solution Lab Network Topology



To characterize end-to-end application impact during system or network failure, the solution architect collects bi-directional test data to analyze overall application-level impact and the recovery mechanisms in the network. Unicast and multicast data, voice, and video traffic directions are divided into the following categories:

- Unicast upstream—Traffic flows in unique patterns (point-to-point, client-to-one server, client-to-many servers) originated by end points from the access layer and routed towards the data center or a remote medium and small campus over the WAN infrastructure.
- Unicast downstream—Traffic flows originated from data centers by a single or many servers destined to many clients connected at the access layers.
- Multicast downstream—Traffic flows of multicast data and video originated by multicast sources in data centers and sent to many multicast receivers connected at the access layers.

All results described in subsequent sections are validated with the bi-directional traffic patterns described above.

7 Implementing Network Resiliency

The Borderless Campus design guide recommends deploying a mix of hardware and software resiliency designed to address the most common campus LAN network faults and instabilities. It is important to analyze network and application impact using a top-down approach and implement the appropriate high availability solution to create a resilient network. Implementing a resilient hardware and software design maintains the availability of all upper layer network services that are deployed in a Borderless Campus design. This section provides Layer 2 and Layer 3 network design recommendations to build a simplified, flexible, scalable, and resilient multi-tier enterprise campus network. Cisco recommendations and best practices are consistent across different campus network sizes and designs (two-tier versus three-tier).

Campus network stability and reliability are challenged during most common path failures caused by fiber cuts, faulty hardware, or Layer 1 link errors. Such fault conditions de-stabilize the network and result in service disruptions and degraded application performance. Network-level resiliency can be stabilized and service disruptions minimized by suppressing link faults and dampening un-stable network paths by implementing Cisco recommended network resiliency techniques.

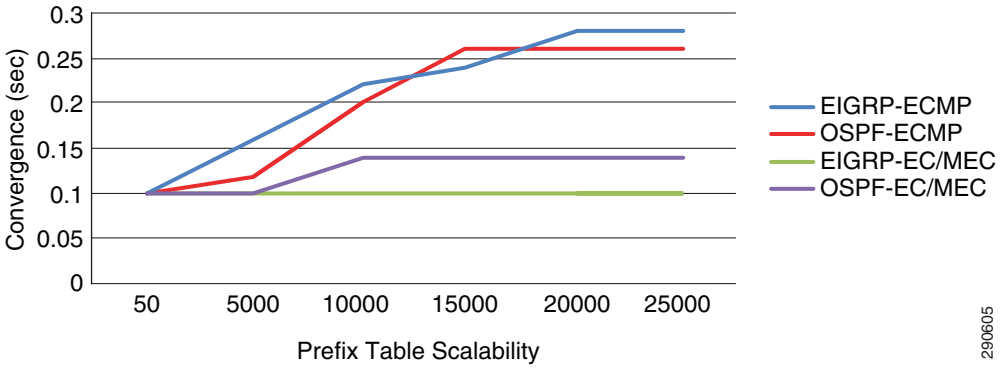
ECMP versus EtherChannel

Chapter 1, "Deploying Network Foundation Services," describes the functional and operational impact of several Layer 2 and Layer 3 network foundation technologies based on ECMP and EtherChannel. The key point to consider in a network design with multiple parallel paths between two system is to simplify the operation with a single logical EtherChannel that builds a concise network routing and switching topology and lets intelligent hardware perform network optimization with parallel forwarding paths that increase network capacity and resiliency.

In the Multilayer campus network design, depending on the aggregation system configuration mode—VSS versus Standalone—the network administrator must deploy EtherChannel/MEC when there are multiple parallel Layer 2 paths between the logical distribution and access layer systems. Bundling Layer 2 paths between two systems offers several architectural and operational benefits (see Chapter 1, "Deploying Network Foundation Services," for more details). If the distribution layer system is deployed in standalone configuration mode, then it may operate in a sub-optimal configuration with two distributed Layer 2 uplinks from the access layer network. Depending on the Layer 2 VLAN design—Flat versus Segmented VLANs in the distribution block—the forwarding path may become asymmetric. Alternatively, the Layer 3 routing boundary can be extended to the wiring closet with a subset routing function to build active/active Layer 3 forwarding paths between the distribution and access layer systems. From a network resiliency perspective, both recommended Layer 2 MEC and Layer 3 routed access designs deliver deterministic sub-second network recovery during link faults.

As next-generation campus systems are evolving with high-performance systems and network virtualization, the redundant and mission-critical enterprise campus distribution and core systems must be simplified to scale, enable borderless network services, improve application quality, and increase user satisfaction. If the distribution and core layer are deployed with the Catalyst 6500-E in VSS mode, then it is highly recommended to build a single unified point-to-point Layer 3 MEC between both campus layer systems. A full-mesh, diversified, and distributed fiber between both virtual switch systems helps increase hardware-driven data load sharing and builds a prefix scale independent campus backbone network. Figure 7 provides evidence of how a well-designed network simplifies and future-proofs network operation and resiliency and delivers consistent, deterministic enterprise-class network recovery independent of prefix scale size in the campus backbone.

Figure 7 6500-E VSS—ECMP versus EC/MEC Link Loss Analysis



EtherChannel technology should be leveraged when the Cisco Nexus 7000 is deployed in the campus core layer. The Nexus 7000 system should be deployed with a single Layer 3 EtherChannel when there are multiple parallel Layer 3 paths with a standalone neighbor device or with distributed Layer 3 paths between logical switches, i.e., VSS or 4500E in redundant mode. Deploying Layer 3 EtherChannel in

the high-scale campus backbone, the Nexus 7000 system is specifically designed to offer the same consistent application performance and user experience as Catalyst 6500-E VSS mode. In a recommended EtherChannel-based campus network design, the Nexus 7000 performs as consistently as the Catalyst 6500-E and delivers network stability and resiliency during path failures.

EtherChannel/Multi-Chassis EtherChannel

In a non-EtherChannel network environment, the network protocol requires fault detection, topology synchronization, and best path recomputation in order to reroute traffic requiring variable timing and to restart the forwarding of traffic. Conversely, EtherChannel or MEC network environments provide significant benefits in such conditions, as the network protocol remains unaware of the topology changes and allows the hardware to self-recover from faults. Re-routing traffic over an alternate member link of EtherChannel or MEC is based on minor internal system EtherChannel hash re-computations instead of an entire network topology re-computation. Hence an EtherChannel and MEC-based network provides deterministic sub-second network recovery of minor to major network faults.

The design and implementation considerations for deploying diverse physical connectivity across redundant standalone systems and virtual systems to create a single point-to-point logical EtherChannel is explained in the Designing the Campus LAN Network in Chapter 1, "Deploying Network Foundation Services."

EtherChannel/MEC Network Recovery Analysis

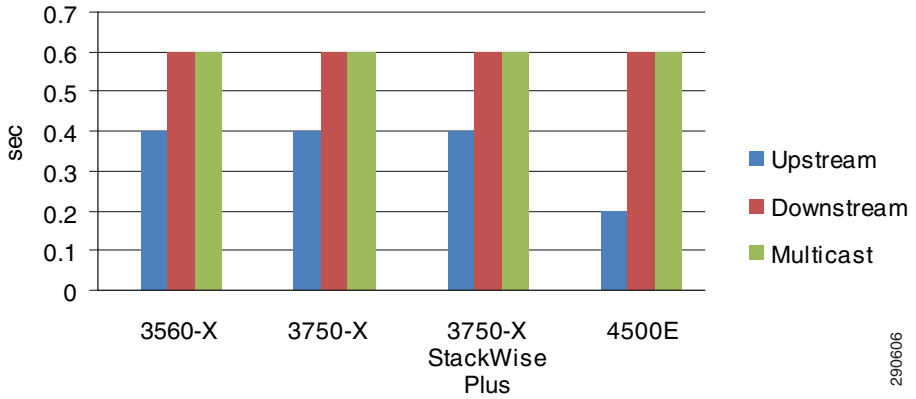
Network recovery with EtherChannel and MEC is platform- and diverse-physical-path-dependent instead of Layer 2 or Layer 3 network protocol dependent. The Borderless Campus design deploys EtherChannel and MEC throughout the network in order to to develop a simplified single point-to-point network topology which does not build any parallel routing paths between any devices at any network tiers.

During individual member link failures, the Layer 2 and Layer 3 protocols dynamically adjust the metrics of the aggregated port channel interfaces. Spanning-Tree updates the port costs and Layer 3 routing protocols like EIGRP update the composite metrics (note that OSPF may change the interface cost). In such events, the metric change will require the generation of minor update messages in the network and will not require end-to-end topology recomputations that impact the overall network recovery process. Since the network topology remains intact during individual link failures, the re-computation to select alternate member links in EtherChannel and MEC becomes locally significant to each impacted EtherChannel neighbor on either end. EtherChannel re-computation requires recreating a new logical hash table and re-programming the hardware to re-route the traffic over the remaining available paths in the bundled interface. The Layer 2 or Layer 3 EtherChannel and MEC re-computation is rapid and independent of network scale.

Catalyst 6500-E VSS MEC Link Recovery Analysis

Several types of network faults can trigger link failures in the network (e.g., fiber pullout, GBIC failure, etc.). Network recovery remains consistent and deterministic in all network fault conditions. In standalone or non-virtual systems using switches such as the Catalyst 3560-X or 4500E, the EtherChannel recomputation is fairly easy as the alternate member link resides within the system. However, with the distributed forwarding architecture in virtual systems like Catalyst 6500-E VSS and Catalyst 3750-X StackWise Plus, extra computation may be required to select alternate member link paths through its inter-chassis backplane interface—VSL or StackRing. Such designs still provide deterministic recovery, but with an additional delay to recompute a new forwarding path through the remote virtual switch node. The link failure analysis chart with inter-chassis reroute in Figure 8 summarizes several types of faults induced in large scale EIGRP and OSPF campus topologies during the development of this Cisco Validated Design guide.

Figure 8 Catalyst 6500-E VSS Inter-Chassis MEC Link Recovery Analysis



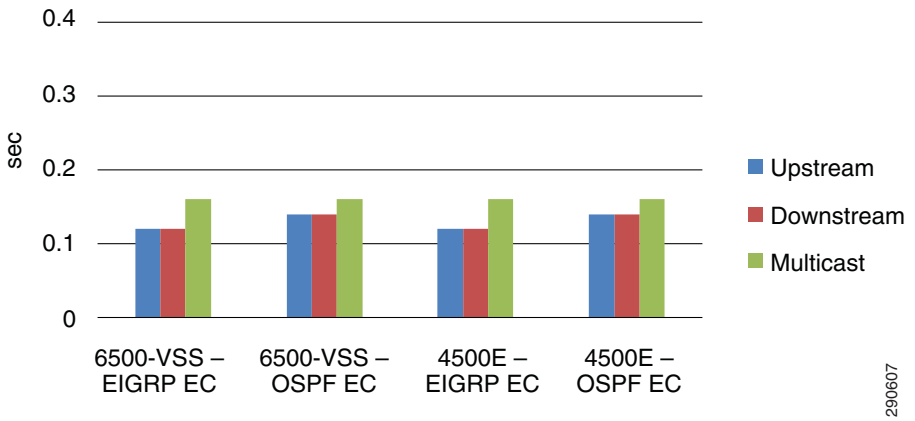
The Borderless Campus network can be designed optimally for deterministic and bidirectional symmetric network recovery for unicast and multicast traffic. For an intra-chassis recovery analysis with the same network faults tested in inter-chassis scenarios, refer to Redundant Linecard Modules.

Nexus 7000 EtherChannel Link Recovery Analysis

As described earlier, designing an EtherChannel-based campus network minimizes routing topology recomputation during individual member link failures. Member link failure in an EtherChannel-based network design suppresses notification to upper layer protocols such as EIGRP, OSPF, and multicast PIM, while the same link fault in an ECMP network design may force a network-wide topology change and could cause forwarding path switchover due to metric adjustments. Based on the best practices in this design guide, the Nexus 7000 maintains next-hop Layer 3 OSPF or EIGRP paths in the URIB/FIB table or the multicast OIF interface in the MRIB/MFIB table during individual member link failures.

With fully-synchronized forwarding information across all system-wide installed I/O modules, the hardware rapidly re-computes the EtherChannel hash and performs data switching based on the new lookup. Even if the new forwarding egress path is within the same I/O module or another I/O module, data plane re-routing within the system across crossbar fabric module remains deterministic and within the campus HA baseline. Figure 9 summarizes several types of campus core layer link faults induced in large-scale EIGRP and OSPF campus core network topologies during the development of this Cisco Validated Design guide.

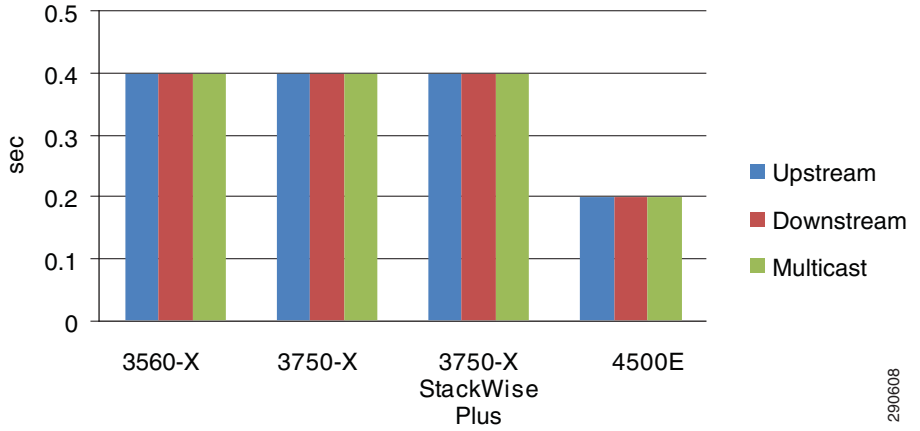
Figure 9 Cisco Nexus 7000 EtherChannel Link Recovery Analysis



Catalyst 4500E EtherChannel Link Recovery Analysis

In the Borderless Campus design, a Catalyst 4500E with redundant hardware components is deployed in the different campus LAN network tiers. A Cisco Catalyst 4500E can only be deployed in standalone mode with in-chassis supervisor and module redundancy. The traffic load balancing and rerouting across different EtherChannel member links occurs within the local chassis. The centralized forwarding architecture in the Catalyst 4500E can rapidly detect link failures and reprogram the hardware with new EtherChannel hash results. The test results in Figure 10 confirm the deterministic and consistent network recovery in large-scale campus topologies running EIGRP and OSPF during individual Layer 2/Layer 3 EtherChannel member link failures.

Figure 10 Catalyst 4500E EtherChannel Link Recovery Analysis



Unidirectional Link Detection (UDLD)

UDLD is a Layer 2 protocol that works with Layer 1 features to determine the physical status of a link. At Layer 1, auto-negotiation takes care of physical signaling and fault detection. UDLD performs tasks that auto-negotiation cannot perform, such as detecting the identity of neighbors and shutting down misconnected ports. When auto-negotiation and UDLD are both enabled, the Layer 1 and Layer 2 detection methods work together to prevent physical and logical unidirectional connections and protocol malfunctions. The UDLD protocol functions transparently on Layer 2 or Layer 3 physical ports. The protocol level, uni-directional communication between two systems should be deployed based on these recommendations:

- Layer 2 Network—In the multilayer standalone or EtherChannel-based network design, the UDLD protocol can be enabled on a per-trunk port level between the access and distribution switches.
- Layer 3 ECMP—In the Layer 3 ECMP-based campus core or in a routed access network design, the uni-directional communication between two systems can be detected by Layer 3 routing protocols as it operates on per-physical interface basis.
- Layer 3 EtherChannel—In a recommended EtherChannel-based network design, the UDLD should be implemented between two Layer 3 systems. Enabling UDLD on each member link of the Layer 3 EtherChannel provides uni-directional path detection at the Layer 2 level.

Copper media ports use Ethernet link pulses as a link monitoring tool and are not susceptible to unidirectional link problems. However, because one-way communication is possible in fiber optic environments, mismatched transmit/receive pairs can cause a link up/up condition even though bidirectional upper layer protocol communication has not been established. When such physical connection errors occur, it can cause loops or traffic black holes. UDLD operates in one of two modes:

- *Normal mode (Recommended)*—If bidirectional UDLD protocol state information times out, it is assumed there is no fault in the network and no further action is taken. The port state for UDLD is marked as undetermined and the port behaves according to its STP state.
- *Aggressive mode*—If bidirectional UDLD protocol state information times out, UDLD attempts to reestablish the state of the port provided it detects that the link on the port is operational. Failure to reestablish communication with UDLD neighbor forces the port into the err-disable state, which either must be manually recovered by the user or the switch if it is be configured for auto-recovery within a specified time interval.

Unidirectional fiber cable anomalies can trigger asymmetric communication and may cause network instability, e.g.., STP loops. Normal UDLD operation detects such faults and prevents network instability by disabling the physical port. The default time to detect the unidirectional links and take action in normal or aggressive mode UDLD may still involve a delay of several seconds in a mission critical campus network. To address this, Cisco has introduced fast UDLD technology that can provide sub-second detection of the fault , thus helping to minimize network impact. Currently fast UDLD is supported on Cisco Catalyst 4500 switches running 12.2(54)SG and 6500 12.2(33)SXI4. Cisco Catalyst 4500E Sup7-E running IOS-XE 3.1.0 SG does not support fast UDLD.

While fast UDLD solves the unidirectional link condition with acceptable delay, it introduces the following challenges for large, redundant campus network designs:

- CPU Utilization—Since the fast UDLD hello packets are processed in milliseconds, it requires heavy CPU interruption. Depending on the number of fast UDLD-enabled links and other software processing network applications, fast UDLD may introduce network and system instability challenges for the network administrator.
- SSO Switchover—This design guide recommends deploying Cisco Catalyst modular platforms with dual supervisor modules on each chassis to provide redundancy. Any Layer 2 or Layer 3 protocols implemented with a sub-second timer may trigger a session timeout and create a false positive alarm, which may result in an entire network outage during a supervisor switchover event. The new active supervisor module in the recovery system cannot restart software processing until several seconds have elapsed. Hence, the peer device initiates a session reset due to not receiving the required keepalives within the time period specified by the timeout parameter.

Note It is recommended to avoid implementing UDLD in aggressive mode as well as fast UDLD on the Cisco Catalyst switches deployed with redundant supervisor modules.

The following illustrates a configuration example to implement the UDLD protocol in normal mode:

Cisco IOS

```
cr22-6500-VSS#config t
cr22-6500-VSS(config)#interface range Ten1/1/8 , Ten2/1/8
cr22-6500-VSS(config-if-range)#udld port

cr22-6500-VSS#show udld neighbors
Te1/1/8    TBM14364802     1            Ethernet1/2     Bidirectional
Te2/1/8    TBM14364802     1            Ethernet2/2     Bidirectional
```

Cisco NX-OS

```
cr35-N7K-Core2(config)# feature udld
!Enable UDLD feature set

cr35-N7K-Core2#show udld neighbors
Port             Device Name    Device ID   Port ID         Neighbor State
-------------------------------------------------------------------------
Ethernet1/2      08E3FFFC4      1           Te1/1/8         bidirectional
Ethernet2/2      08E3FFFC4      1           Te2/1/8         bidirectional
```

IP Event Dampening

Unstable physical network connectivity with poor signaling or loose connections may cause continuous port flaps. When the Borderless Campus network is not deployed using best practice guidelines to summarize the network boundaries at the aggregation layer, a single interface flap can severely impact the stability and availability of the entire campus network. Route summarization is one technique used to isolate the fault domain and contain local network faults within the domain.

To ensure local network domain stability during port flaps, all Layer 3 interfaces can be implemented with IP Event Dampening, which uses the same fundamental principles as BGP dampening. Each time the Layer 3 interface flaps, IP dampening tracks and records the flap event. On multiple flaps, a logical penalty is assigned to the port and it suppresses link status notifications to IP routing until the port becomes stable.

IP Event Dampening is a local specific function and does not have any signaling mechanism to communicate with remote systems. It can be implemented on each individual physical or logical Layer 3 interface—physical ports, SVI, or port-channels:

- Layer 3 Port-Channel

```
cr24-4507e-MB(config)#interface Port-Channel 1
cr24-4507e-MB(config-if)#no switchport
cr24-4507e-MB(config-if)#dampening
```

- Layer 2 Port-Channel

```
cr24-4507e-MB(config)#interface Port-Channel 15
```

```
cr24-4507e-MB(config-if)#switchport
cr24-4507e-MB(config-if)#dampening
```

- SVI Interface

```
cr24-4507e-MB(config)#interface range Vlan101 - 120
cr24-4507e-MB(config-if-range)#dampening

cr24-4507e-MB#show interface dampening
Vlan101
  Flaps Penalty    Supp ReuseTm   HalfL  ReuseV   SuppV  MaxSTm    MaxP Restart
      3       0   FALSE       0       5    1000    2000      20   16000       0
…
TenGigabitEthernet3/1 Connected to cr23-VSS-Core
  Flaps Penalty    Supp ReuseTm   HalfL  ReuseV   SuppV  MaxSTm    MaxP Restart
     10       0   FALSE       0       5    1000    2000      20   16000       0
…
Port-channel1 Connected to cr23-VSS-Core
  Flaps Penalty    Supp ReuseTm   HalfL  ReuseV   SuppV  MaxSTm    MaxP Restart
      3       0   FALSE       0       5    1000    2000      20   16000       0
Port-channel15 Connected to cr24-3560X-MB
  Flaps Penalty    Supp ReuseTm   HalfL  ReuseV   SuppV  MaxSTm    MaxP Restart
      3       0   FALSE       0       5    1000    2000      20   16000       0
```

8 Implementing Device Resiliency

Each device in the borderless enterprise LAN and WAN network design is connected to a critical system or end-point to provide network connectivity and services for business operations. Like network resiliency, device resiliency integrates redundant hardware components and software-based solutions into a single standalone or virtual systems. Depending on the platform architecture of the Cisco router or switch deployed in the campus network design, device redundancy is divided into four major categories—Redundant Power Supplies, Redundant Line cards, Redundant Supervisor/RP, and Non-Stop Forwarding (NSF) with Stateful Switchover (SSO).

Redundant Power

To provide non-stop network communication during power outages, critical network devices must be deployed with redundant power supplies. To maintain network services operation and prevent disruption in any campus network tier, the Cisco Catalyst and Nexus 7000 systems are designed to provide power redundancy during power outages or hardware failure. Deploying redundant power supplies offers 1+1 or N+1 power redundancy against power supply unit or power source failure that helps reduce mean-time-to-repair (MTTR) in the mission critical campus system. In the recommended

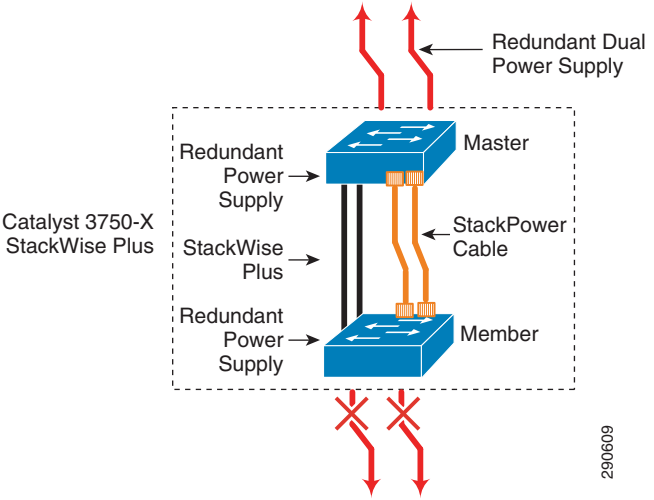
system power redundancy design, campus network communication remains transparent and uninterrupted during power failure, with graceful switchover to redundant power supply units or power input sources.

At the campus access layer, the network administrator must identify the network systems that provide network connectivity and services to mission critical servers. This would also include Layer 1 services such as PoE to boot IP phones and IP video surveillance cameras for campus physical security and communications.

Catalyst 3750-X—Cisco StackPower Redundancy

The next-generation Catalyst 3750-X Series platform introduces innovative Cisco StackPower technology to provide power redundancy solutions for fixed configuration switches. Cisco StackPower unifies the individual power supplies installed in the switches and creates a pool of power, directing that power where it is needed. Up to four switches can be configured in a StackPower stack with the special Cisco proprietary StackPower cable. The StackPower cable is different than the StackWise data cables and is available on all Cisco Catalyst 3750-X models. See Figure 11.

Figure 11 Cisco StackPower Redundancy



A stack member switch experiencing a power fault with its own power supply can derive power from the global power pool so as to provide seamless, continued operation in the network. With the modular power supply design in Catalyst 3750-X Series platform, the defective power supply can be swapped out without disrupting network operation. Cisco StackPower technology can be deployed in two modes:

- *Sharing mode*—All input power is available to be used for power loads. The total aggregated available power in all switches in the power stack (up to four) is treated as a single large power supply. All switches in the stack can provide this shared power to all powered devices connected to PoE ports. In this mode, the total available power is used for power budgeting decisions without any power reserved to accommodate power supply failures. If a power supply fails, powered devices and switches could be shut down. This is the default mode of operation.
- *Redundant mode*—The power from the largest power supply in the system is subtracted from the power budget and held in reserve. This reduces the total power available to PoE devices, but provides backup power in case of a power supply failure. Although there is less available power in the pool for switches and powered devices to draw upon, the possibility of having to shut down switches or powered devices in case of a power failure or extreme power load is reduced. It is recommended to budget the required power and deploy each Catalyst 3750-X switch in the stack with dual power supplies to meet demand. Enabling redundant mode offers power redundancy as a backup should one of the power supply units fail.

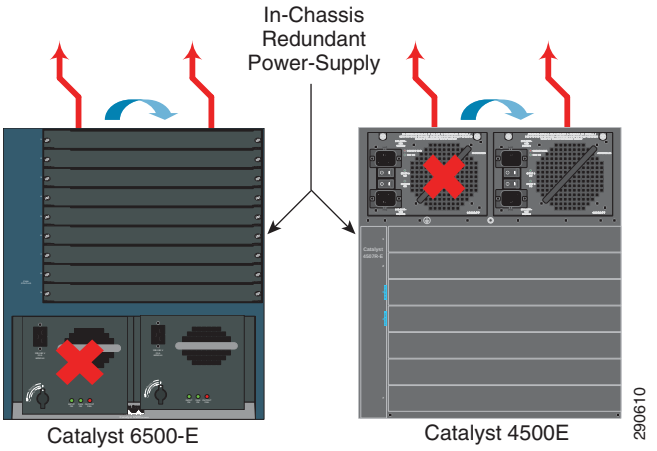
Since Cisco StackWise Plus can group up to nine 3750-X Series switches in the stack ring, Cisco StackPower must be deployed with two power stack groups in order to accommodate up to four switches. The following sample configuration demonstrates deploying Cisco StackPower in redundancy mode and grouping the stack members into power stack groups. To make the new power configuration effective, it is important that network administrator plan for network downtime as all the switches in the stack ring must be reloaded:

```
cr36-3750X-xSB(config)#stack-power stack PowerStack
cr36-3750X-xSB(config-stackpower)#mode redundant

cr36-3750X-xSB(config)#stack-power switch 1
cr36-3750X-xSB(config-switch-stackpower)#stack-id PowerStack
%The change may not take effect until the entire data stack is reloaded

cr36-3750X-xSB(config)#stack-power switch 2
cr36-3750X-xSB(config-switch-stackpower)#stack-id PowerStack
%The change may not take effect until the entire data stack is reloaded
```

Catalyst 4500E and 6500-E (In-Chassis Power Redundancy)

The Cisco Catalyst 4500E and 6500-E Series modular platforms allocate power to several internal hardware components, such as linecards, fans, etc., and externally powered devices, such as IP phones, wireless access points, etc. All of the power is allocated from the internal power supply. With a dual power supply unit hardware design, the Catalyst 6500-E and 4500E systems provide the flexibility to expand the use of power supplies as the network grows. Like linecard module hardware design, power supplies are hot-swappable and implementing 1+1 power redundancy provides network services resiliency while replacing the faulty unit.

Figure 12 Catalyst 4500E and 6500-E Power Redundancy



Dual power supplies in these systems can operate in two different modes:

- *Redundant Mode* (**Recommended**)—By default, power supplies operate in redundant mode offering a 1+1 redundant option. The system determines power capacity and the number of power supplies required based on the allocated power to all internal and external power components. Both power supplies must have sufficient power to provide power to all the installed modules in order to operate in 1+1 redundant mode.

```
cr24-4507e-LB(config)#power redundancy-mode redundant

cr24-4507e-LB#show power supplies
Power supplies needed by system    :1
Power supplies currently available :2

cr22-vss-core(config)#power redundancy-mode redundant switch 1
cr22-vss-core(config)#power redundancy-mode redundant switch 2

cr2-6500-vss#show power switch 1 | inc Switch|mode
Switch Number: 1
system power redundancy mode = redundant

cr2-6500-vss#show power switch 2 | inc Switch|mode
Switch Number: 2
system power redundancy mode = redundant
```

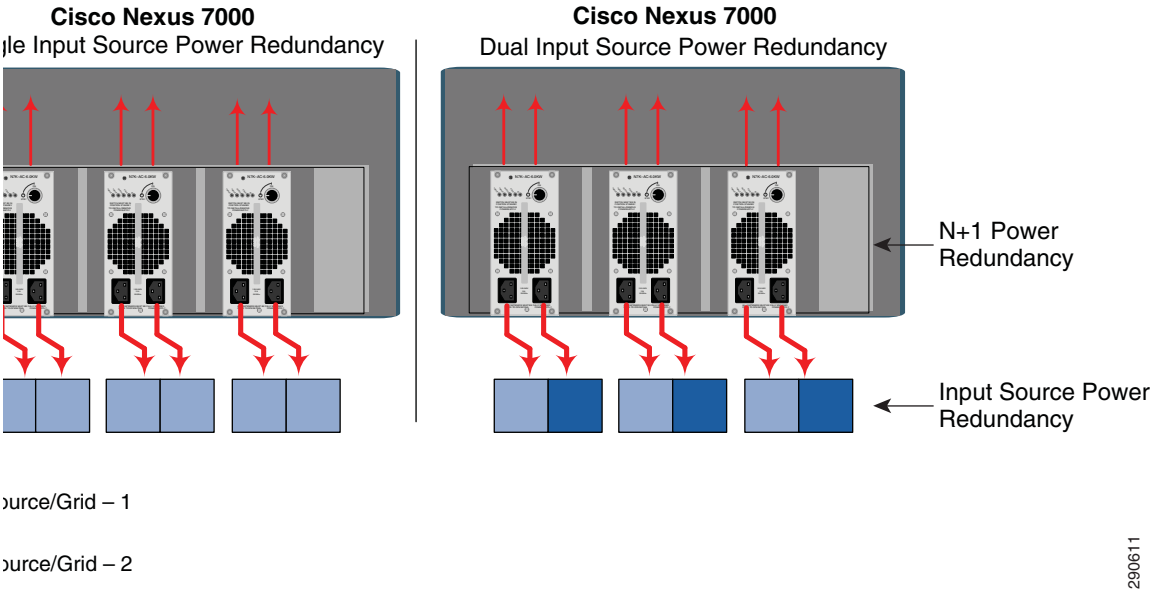
- *Combined mode*—If the system power requirement exceeds the capacity of a single power supply, then the network administrator can utilize both power supplies in combined mode to increase overall capacity. However it may not offer 1+1 power redundancy during a primary power supply failure. The following global configuration enables power redundancy operation in combined mode:

```
cr24-4507e-LB(config)#power redundancy-mode combined

cr24-4507-LB#show power supplies
Power supplies needed by system:2
Power supplies currently available:2
```

Cisco Nexus 7000 (In-Chassis Power Redundancy)

The Cisco Nexus 7000 system can be protected by three internally redundant power supplies with two internal isolated power units that provide up to six active power paths in a fully redundant configuration. Several hardware components, such as supervisor, I/O modules, fan, and crossbar fabric module, consume power from the total aggregated power wattage. All active power supplies use a proportional load sharing method for power distribution to each hardware component that allows efficient use of dissimilar capacity power supplies in the same system. The Cisco Nexus 7000 offers power redundancy to the system in two power source environments—Single Input and Dual Input. The single input source power provides N+1 power unit redundancy, while the dual input source power provides system power protection in multi-failure conditions—power source or grid and power unit failure.

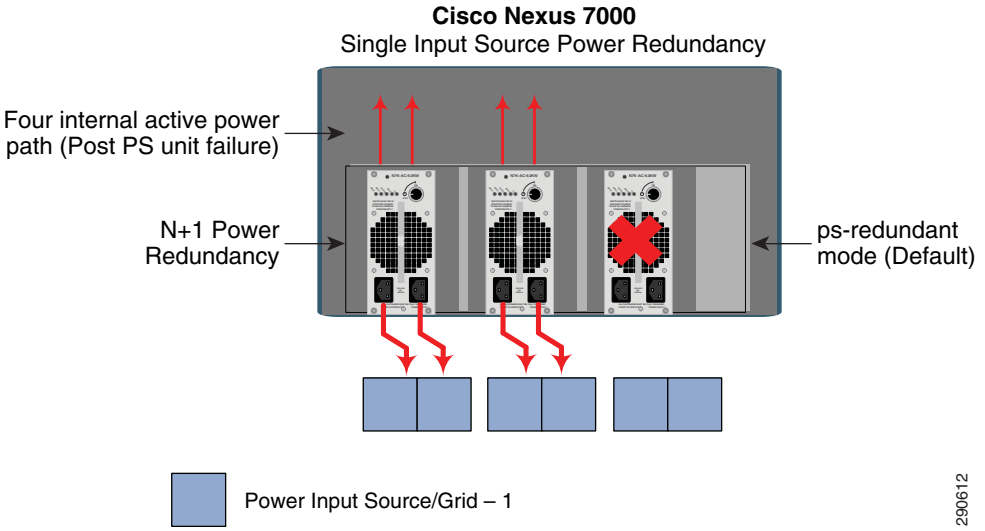
Figure 1-13 Cisco Nexus 7000 Power Redundancy



Implementing redundant power subsystem allows all three units to be configured in the following redundancy modes.

- *PS Redundant mode* (**Recommended**)—By default, deploying redundant power supply units provides N+1 power supply unit redundancy. This redundant mode provides protection against a single power supply unit failure where all power sources are distributed through a single power grid. The cumulative available power to distribute between components is the sum of all installed power supplies minus that of the largest (for redundancy). During single power supply failure, loads are redistributed using the available capacity across the remaining functional power supply units. N+1 power redundancy becomes available with two or three power supplies installed in the system. In a single power circuit/grid environment, the default PS-redundant mode is recommended for N+1 power supply unit redundancy.

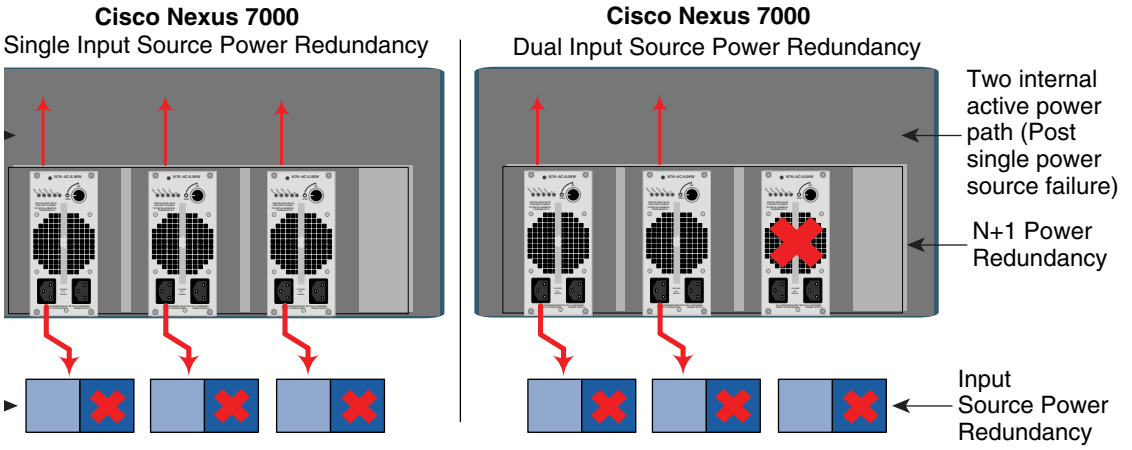
Figure 14 Recommended Single Input Source Power Redundancy




```
cr35-N7K-Core2#show environment power detail | inc Ok|redundancy
1         N7K-AC-6.0KW                515 W       6000 W      Ok
2         N7K-AC-6.0KW                443 W       6000 W      Ok
3         N7K-AC-6.0KW                525 W       6000 W      Ok
Power Supply redundancy mode (configured)                     PS-Redundant
Power Supply redundancy mode (operational)                    PS-Redundant
```

- *Input Source Redundant mode*—In dual power grid network designs, the Nexus 7000 provides the ability to increase power protection against input source failures. To implement input grid power redundancy in the system, each power supply unit must be connected in a distributed model between two independent power sources (grids). During single power grid failure, the cumulative power capacity reduces to half, however with an alternate power source, the remaining half internal power paths remain operational. This mode does not provide power redundancy during individual power unit failure.
- *Redundant mode* (**Recommended**)—The redundant power mode provides multi-failure power protection. Implementing redundant mode provides power protection to the system during power input source (grid) failure and power supply unit failure. This mode provides increased level power redundancy to the Nexus 7000 system by logically combining the N+1 (PS-redundancy) and input grid (*Input Source Redundant*) modes. Each of the power supply redundancy modes imposes different power budgeting and allocation models, which in turn deliver varying usable power yields and capacities. In a dual power input source environment, it is recommended to implement redundancy mode in the Nexus 7000 system.

Figure 1-15 Recommended Dual Input Source Power Redundancy




```
cr35-N7K-Core2(config)# power redundancy-mode redundant

cr35-N7K-Core2# show environment power detail | inc Ok|redundancy
1         N7K-AC-6.0KW                519 W       6000 W      Ok
2         N7K-AC-6.0KW                438 W       6000 W      Ok
3         N7K-AC-6.0KW                521 W       6000 W      Ok
Power Supply redundancy mode (configured)                   Redundant
Power Supply redundancy mode (operational)                  Redundant
```

- *Combined mode*—The cumulative available power watts can be combined with all installed power supplies to provide the sum of all available power to the usable power budget. The combined mode does not provide power redundancy. In this mode the power failure or the unit failure degrades available power to the system. Based on the number of installed hardware components, if power draw is exceeded after failure, it may cause I/O module power down, which may severely impact network services availability and campus backbone capacity. This mode may become an un-reliable power design for power protection during source or unit failure and may introduce network instability or complete outage.

Network Recovery Analysis with Power Redundancy

Each campus LAN router and switch providing critical network services must be powered with either an in-chassis or external redundant power supply system. This best practice is also applicable to the standalone or virtual system devices. Each physical Catalyst 6500-E chassis in VSS mode at the campus distribution and core layer must be deployed with a redundant in-chassis power supply. The Nexus 7000 system at the mission critical core layer must be deployed with three redundant power supply units. Depending on the number of power input sources, the network administrate must implement Cisco recommended power redundancy techniques. The Catalyst 3750-X StackWise Plus must be deployed following the same rule, with the master and member switches in the stack ring deployed using the external redundant power system. Powering virtual systems with redundant power supplies prevents a reduction in network bandwidth capacity, topology changes, and poor application performance in the event of a power failure event.

Several power failures on redundant power systems were conducted during the production of this Cisco Validated Eesign in order to characterize overall network and application impact. Several test cases performed on all redundant power campus systems confirms zero-packet loss during individual power supply failures. Note that the network administrator must analyze the required power capacity that will be drawn by different hardware components (e.g., network modules, PoE+ etc.).

Redundant Linecard Modules

Modular Catalyst platforms support a wide range of linecards for connectivity to the network core and edge. The high-speed core linecards are equipped with special hardware components to build the campus backbone, whereas the network edge linecards are developed with more intelligence and application awareness. Using internal system protocols, each line card communicates with the centralized control plane processing supervisor module through the internal backplane. Any type of internal communication failure or protocol malfunction may disrupt communication between the linecard and the supervisor, which may lead to the linecard and all the physical ports associated with it forcibly resetting to resynchronize with the supervisor.

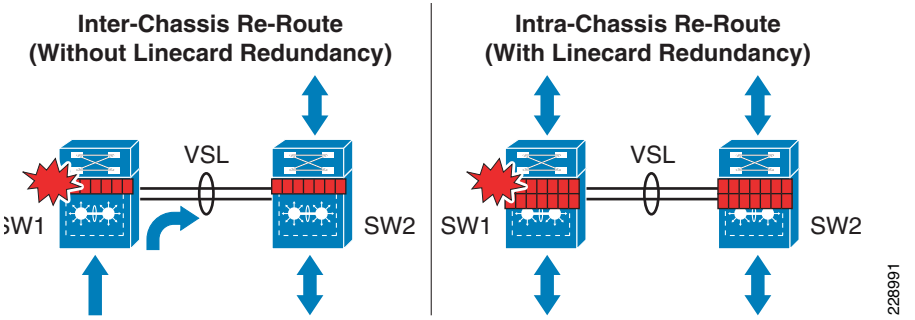
Catalyst 6500-E Linecard Module Recovery Analysis

When the distribution and core layer 6500-E systems are deployed with multiple redundant line cards, the network administrator must design the network by diversifying the physical cables across multiple linecard modules. A full-mesh, diversified fiber design between two virtual switching systems and linecard modules minimizes service disruption and prevents network congestion. The distributed forwarding architecture in hardware is fully synchronized on each DFC-based linecard deployed in the virtual switch. In a steady network state, this software design minimizes data routing across system critical VSL paths. Data traffic traverses the VSL links as a “last-resort” in hardware if either of the virtual switch chassis loses a local member link from the MEC link due to a fiber cut or a major fault

condition like a linecard failure. The impact on traffic could be in the sub-second to seconds range and it may create congestion on the VSL Etherchannel link if the rerouted traffic exceeds overall VSL bandwidth capacity.

Deploying redundant linecards and diversifying paths across the modules prevents inter-chassis re-route, which may cause network congestion if there is not sufficient VSL bandwidth to accommodate the rerouted traffic. Figure 16 demonstrates inter-chassis reroute (without linecard redundancy) and intra-chassis re-route (with linecard redundancy).

Figure 16 Intra-Chassis versus Inter-Chassis Traffic Re-route

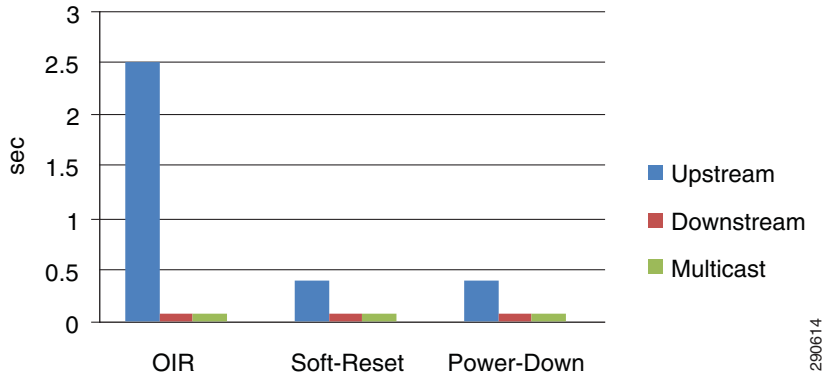


Implementing distributed and diversified fibers between modules mitigates VSL congestion problems. To minimize the service disruption and increase network recovery, the network administrator can follow Cisco recommended best practices to swap or replace module in production network. Removing linecard modules from the modular system while it is in service requires several system internal checks to detect the removal and update distributed forwarding information across all operational modules in the 6500E chassis. This process may take seconds to restore traffic through alternative forwarding paths. To minimize the downtime and restore the service within sub-seconds, Cisco recommends to first disable the linecard from the service and then remove it from the system. The linecard can be put out-of-service in two recommended ways:

- Soft-Reset—Issuing **hw-module switch <1|2> module** <#> from exec mode is a graceful module reset from a software and hardware forwarding perspective, which helps minimize traffic losses bi-directionally. With MEC it also helps minimize control plane changes that trigger topology computation or re-routing. The traffic remains operational through alternate modules and distributed without going through an inter-switch VSL path.
- Power-Down—Disabling power allocation to the network module produces the same impact to the system and network as a soft-reset. The key difference in this procedure is that the module in the specified slot will remain powered down until a new module is installed or power is re-allocated. Power allocation to a module can be disabled using the **no power enable switch <1|2> module** <#> command from global configuration mode.

Both recommended procedures provide graceful network recovery during the linecard removal process. Figure 17 provides an analysis of linecard OIR, soft-reset, and power-down.

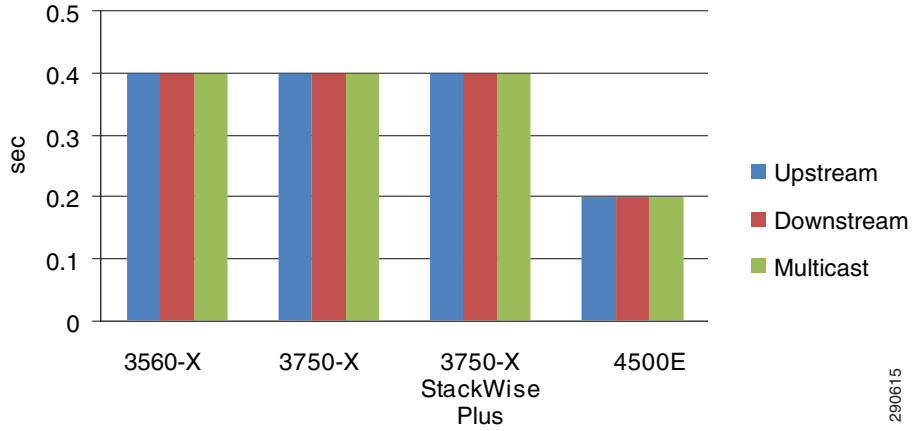
Figure 17 6500E VSS Linecard Recovery Analysis



Catalyst 4500E Linecard Module Recovery Analysis

The centralized forwarding architecture in a Catalyst 4500E programs all the forwarding information on the active and standby supervisor Sup7-E, Sup6-E, or Sup6L-E modules. All the redundant linecards in the chassis are stub and maintain low-level information to handle ingress and egress forwarding information. During a link or linecard module failure, new forwarding information gets rapidly reprogrammed on both supervisors in the chassis. However, deploying EtherChannel utilizing diversified fibers across different linecard modules provides consistent sub-second network recovery during abnormal failure or the removal of a linecard from the Catalyst 4500E chassis. The chart in Figure 18 provides test results associated with removing a linecard from the Catalyst 4500E chassis deployed in various campus network roles.

Figure 18 Catalyst 4500E Distribution Layer Linecard Recovery Analysis

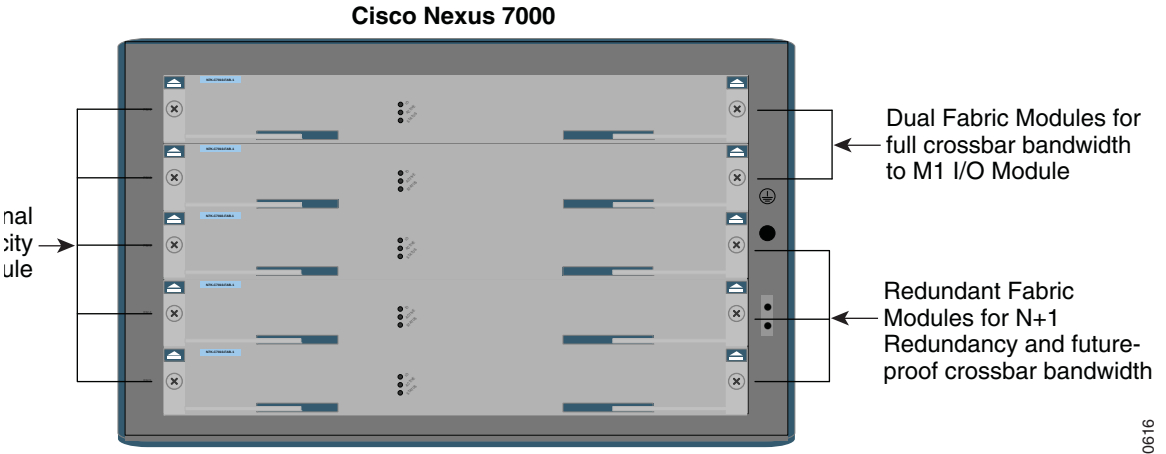


Redundant Nexus 7000 Crossbar Fabric Module

The distributed forwarding architecture in the Nexus 7000 system is built upon an intelligent hardware and software design that decouples the centralized control plane operation from the supervisor module. Each Nexus 7000 I/O module is designed with a distributed forwarding engine that builds and maintains hardware-based forwarding information based on global unicast and multicast RIB. The ingress and egress data switching between ports performs local switching with the I/O module without a centralized lookup procedure or backplane bandwidth involvement to (**??with??**) local switch traffic.

Data traffic switching between different I/O modules is performed through high-speed crossbar modules. The switch fabric capacity per I/O module is determined based on the I/O module's internal throughput capacity and the number of crossbar fabric modules installed in the system. Deploying at least two crossbar fabric modules enables the campus core recommended M108 I/O module to operate at its full 80 Gbps capacity. However during abnormal fabric module failure, the system may introduce backplane congestion due to a lack of sufficient switching capacity from a single crossbar module. It is highly recommended to deploy at least three fabric modules to provide module and backplane capacity redundancy. Cisco recommends deploying additional crossbar fabric modules in the system to future proof the switch fabric bandwidth and increase N+1 fabric module redundancy during abnormal module failure.

Figure 1-19 Cisco Nexus 7000 Crossbar Fabric Bandwidth and Module Redundancy



Crossbar Fabric Bandwidth and Module Redundancy

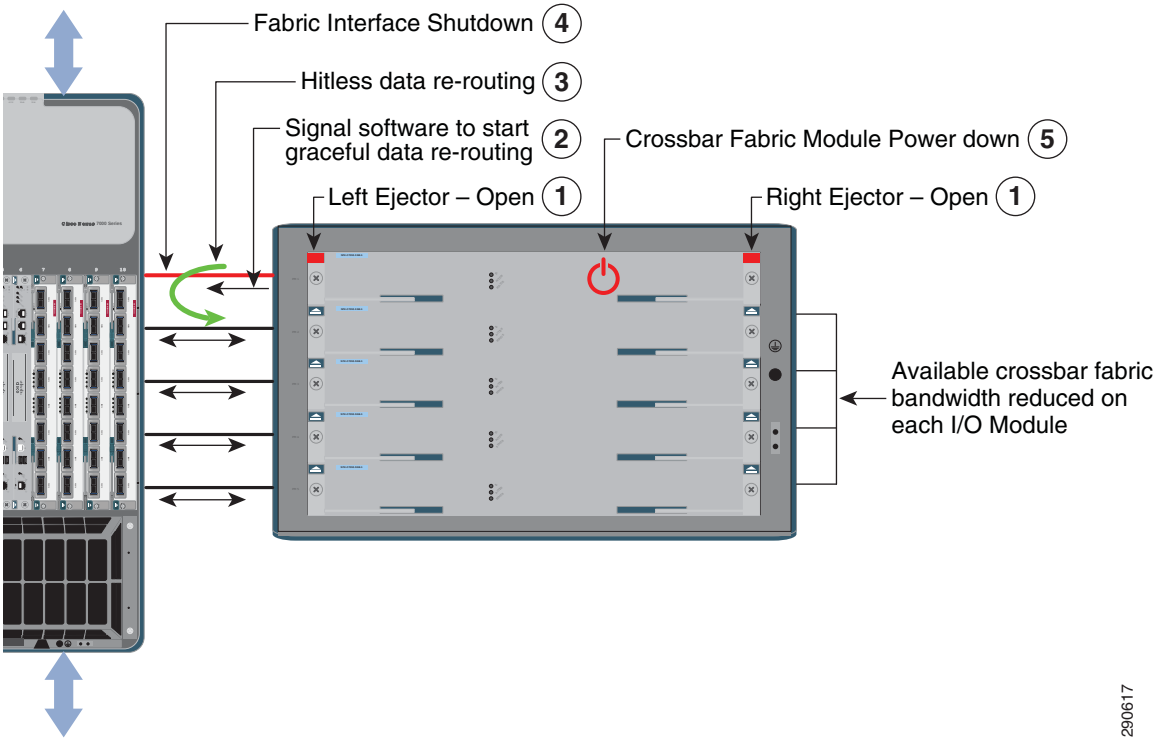
The crossbar fabric modules are hot-swappable. The new fabric module can be inserted in the system without any service disruption or introducing any maintenance window in a production campus network. To swap the fabric module in the Nexus 7000 system, the network administrator must press dual ejector buttons to open and release the internal lock prior to removing the module from operation. The crossbar fabric module get internally shutdown when both ejector buttons are opened; the module remains in an operational state even when one button is open and the other is closed.

```
!Fabric Module remains in operational state with single ejector button in OPEN state
%PLATFORM-3-EJECTOR_STAT_CHANGED: Ejectors' status in slot 13 has changed, Left Ejector is
CLOSE, Right Ejector is OPEN

cr35-N7K-Core2#show module xbar 3 | inc Fabric|Left
3    0      Fabric Module 1                 N7K-C7010-FAB-1    ok
Left ejector OPEN, Right ejector CLOSE, Module HW does support ejector based shutdown.

!Fabric Module status with both ejector button open
%PLATFORM-3-EJECTOR_STAT_CHANGED: Ejectors' status in slot 13 has changed, Left Ejector is
OPEN, Right Ejector is OPEN
%PLATFORM-2-XBAR_REMOVE: Xbar 3 removed (Serial number JAF1442AHKB)

cr35-N7K-Core2# show module | inc Fabric
1    0      Fabric Module 1                 N7K-C7010-FAB-1    ok
2    0      Fabric Module 1                 N7K-C7010-FAB-1    ok
```

To provide hitless fabric module switchover, the hardware sensors are capable of transmitting a signal to the software system for graceful fabric module shutdown when both ejector buttons are opened. This intelligent and highly-available design first gracefully re-routes the data plane to an alternate fabric module prior to internally powering down the fabric module. This graceful OIR proces remains

transparent to the centralized control plane running on a supervisor module and it does not trigger any change in network operation. Figure 1-20 illustrates the step-by-step internal graceful data plane recovery procedure with crossbar fabric module redundancy.

Figure 1-20 Hitless Crossbar Fabric Module Redundancy




```
!System state prior crossbar fabric module failure
cr35-N7K-Core2# show module xbar | inc Fabric
1    0      Fabric Module 1                  N7K-C7010-FAB-1    ok
2    0      Fabric Module 1                  N7K-C7010-FAB-1    ok
3    0      Fabric Module 1                  N7K-C7010-FAB-1    ok
cr35-N7K-Core2# show hardware fabric-utilization
-----------------------------------------------
Slot        Total Fabric        Utilization
            Bandwidth      Ingress % Egress %
-----------------------------------------------
1             138 Gbps           0.0       0.0
2             138 Gbps           0.0       0.0
<snip>
```

```
!System state post crossbar fabric module failure
%PLATFORM-3-EJECTOR_STAT_CHANGED: Ejectors' status in slot 13 has changed, Left Ejector is
OPEN, Right Ejector is OPEN
%PLATFORM-2-XBAR_PWRFAIL_EJECTORS_OPEN: Both ejectors open, Xbar 3 will not be powered up
cr35-N7K-Core2#show module | inc Fabric
1    0      Fabric Module 1                 N7K-C7010-FAB-1    ok
2    0      Fabric Module 1                 N7K-C7010-FAB-1    ok
3    0      Fabric Module                   N/A                powered-dn

!46Gbps Fabric Bandwidth reduced on each I/O module
cr35-N7K-Core2#show hardware fabric-utilization
-----------------------------------------------
Slot        Total Fabric        Utilization
            Bandwidth      Ingress % Egress %
-----------------------------------------------
1              92 Gbps          0.0       0.0
2              92 Gbps          0.0       0.0
<snip>
```

Insufficient Fabric Bandwidth

The high-speed I/O modules may operate under capacity with a non-redundant, single operational crossbar fabric module in a Nexus 7000 system. The 10Gbps M108 I/O module operates at 80 Gbps per slot. With a single operational crossbar fabric, the I/O module remains in an operational state, however backplane switching gets reduced to 46 Gbps. Due to insufficient backplane bandwidth, it may not handle wire-speed campus backbone traffic and may create backplane congestion in the critical campus core.

```
!Steady system state with redundant crossbar fabric modules
cr35-N7K-Core2# show module | inc Fabric
1    0      Fabric Module 1                 N7K-C7010-FAB-1    ok
2    0      Fabric Module 1                 N7K-C7010-FAB-1    ok
3    0      Fabric Module 1                 N7K-C7010-FAB-1    ok
cr35-N7K-Core2# show hardware fabric-utilization
-----------------------------------------------
Slot        Total Fabric        Utilization
            Bandwidth      Ingress % Egress %
-----------------------------------------------
1             138 Gbps          0.0       0.0
2             138 Gbps          0.0       0.0
<snip>

!System state post two crossbar fabric module failure
cr35-N7K-Core2# show module | inc Fabric
1    0      Fabric Module 1                 N7K-C7010-FAB-1    ok
2    0      Fabric Module                   N/A                powered-dn
3    0      Fabric Module                   N/A                powered-dn
```

```
%XBAR-2-XBAR_INSUFFICIENT_XBAR_BANDWIDTH: Module in slot 1 has insufficient
xbar-bandwidth.
%XBAR-2-XBAR_INSUFFICIENT_XBAR_BANDWIDTH: Module in slot 2 has insufficient
xbar-bandwidth.

! Insufficient 46Gbps Fabric Bandwidth for 80Gbps per slot I/O module
cr35-N7K-Core2# show hardware fabric-utilization
-----------------------------------------------
Slot        Total Fabric        Utilization
            Bandwidth      Ingress % Egress %
-----------------------------------------------
1               46 Gbps           0.0       0.0
2               46 Gbps           0.0       0.0
<snip>
```

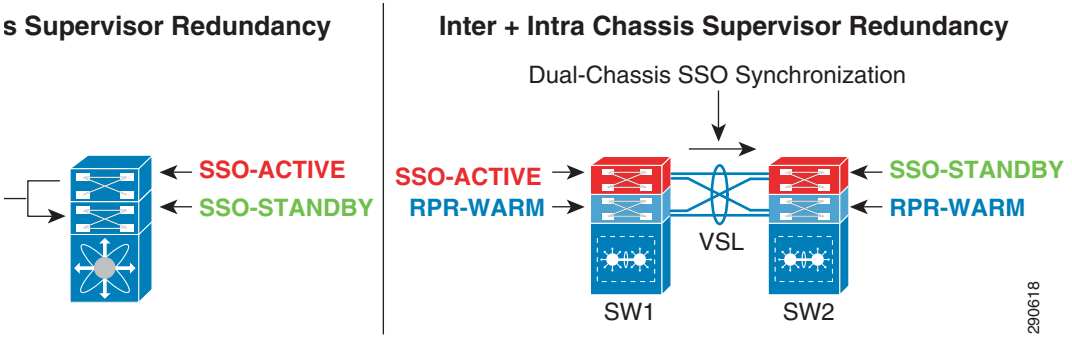
Redundant Supervisor

The enterprise-class modular Cisco Catalyst and Nexus 7000 system support dual-redundant supervisor modules to prevent borderless services disruption due to network control plane and topology resets in the event of supervisor module failure or a forced reset. Deploying redundant supervisor modules in mission critical campus access, distribution, and core layer systems protects network availability and bandwidth capacity during an active supervisor switchover process. Based on the system architecture, the primary supervisor synchronizes all required hardware and software state machines, forwarding information to a secondary supervisor module for seamless operation. The graceful SSO redundancy provides transparent and graceful network recovery that leverages the NSF capability to protect the forwarding plane with completely hitless network recovery. The supervisor redundancy architecture in the recommended modular systems depends on the system hardware design and implemented mode.

- Intra-Chassis Supervisor Redundancy—This mode provides redundancy between two supervisor modules deployed within a single chassis. Depending on the campus system and deployed mode, intra-chassis supervisor redundancy can be in two redundancy modes:
  - SSO—The standalone Nexus 7000 and Catalyst 4500E borderless campus systems provide intra-chassis SSO redundancy. This mode provides single chassis supervisor protection by synchronizing state machines from the active supervisor module to the standby supervisor deployed within the same chassis. The Catalyst 6500-E deployed in standalone mode provides the same intra-chassis SSO redundancy as this system.
  - RPR-WARM—The Catalyst 6500-E deployed in VSS mode is designed to provide inter-chassis redundancy. Deploying each virtual switch with a redundant supervisor module leverages the same set of hardware and supervisor modules to provide quadrupled supervisor redundancy. The intra-chassis supervisor provides virtual switch redundancy if the primary supervisor self-recovery fails.

- Inter-Chassis Supervisor Redundancy—The Cisco VSS innovation with the next-generation Sup720-10GE supervisor module extends the single-chassis SSO supervisor redundancy capability between two separate physical chassis deployed in the same campus network layer. By extending internal backplane communication between two supervisors modules over VSL links, the VSS becomes a single, large, logical, and redundant system to build a unified campus network system. The centralized control plane running on the active supervisor module deployed in the virtual switch, i.e., Switch-1, performs the same SSO synchronization task with the standby supervisor deployed in the remote virtual-switch, i.e., Switch-2.

Figure 1-21 Intra-Chassis versus Inter-Chassis Supervisor Redundancy



Intra-Chassis Supervisor Redundancy

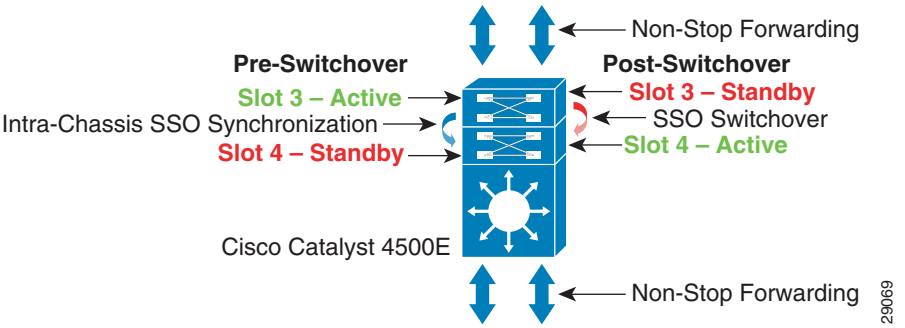
The intra- or single-chassis provides 1+1 supervisor redundancy in the Nexus 7000 system and the Catalyst 4500E switch provides continuous network availability across all the installed modules while the supervisor module is going through the graceful recovery process. Even these systems are modular and provide intra-chassis supervisor redundancy. The hardware and software operation is different when the system is in a steady operational state or going through the switchover process.

Catalyst 4500E

The Catalyst 4500E series platform is modular with a simple hardware and software design. The Catalyst 4500E system is designed for a high-density access layer with end-point-aware intelligence to enable several rich borderless network services at the edge. The active supervisor module holds the ownership of the control and management plane to build the centralized forwarding plane by communicating with end points and upstream network devices. The high-speed linecards are non-distributed and rely on the supervisor for all intelligent forwarding decisions and applying network policies, such as QoS, ACL, etc.

The Catalyst 4500E deployed with a redundant supervisor in SSO configuration dynamically synchronizes the system configuration, network protocol state machines, forwarding information, and more in real-time from the active to the standby supervisor module. During an administrator or software forced supervisor switchover, the Layer 3 network protocols gracefully recover with neighboring systems, however the system maintains its overall network capacity. In the event of supervisor switchover, the uplink ports from both supervisors and linecard modules remains fully operational and in a forwarding state to protect switching capacity and provide continuous non-disruptive borderless services.

Figure 22 Cisco Catalyst 4500E Supervisor Redundancy



Nexus 7000

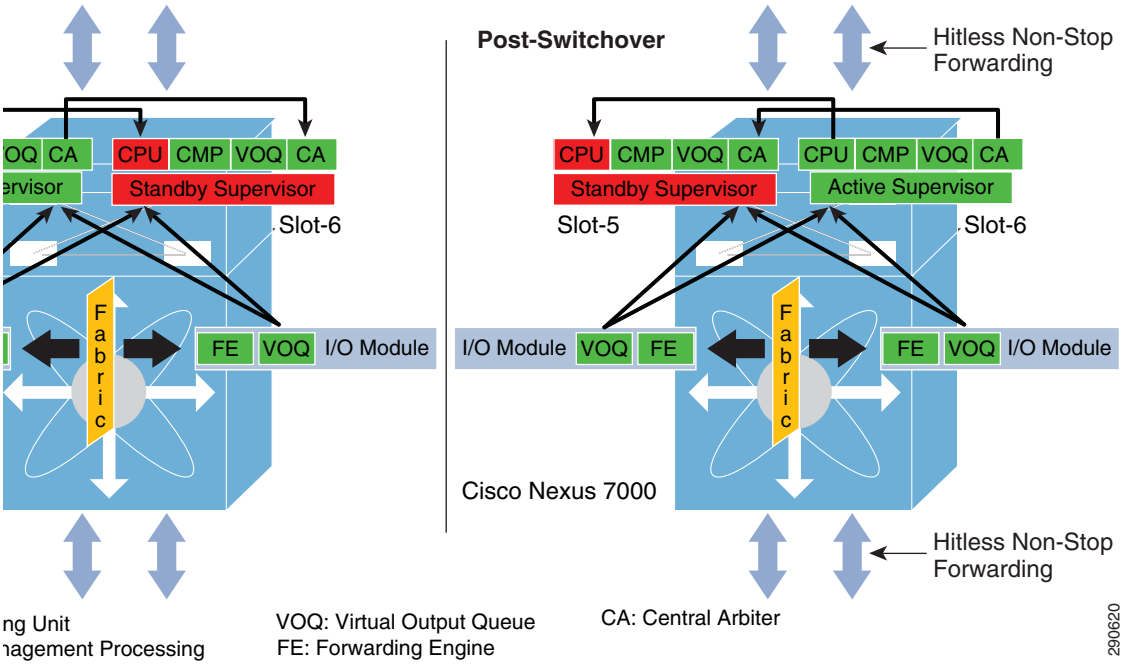
With the increased number of borderless network services and applications running in today's enterprise, the campus network foundation is becoming a core platform to enable a broad ange of digital communication in various emerging forms. As the demands continue to expand at a rapid pace, the backbone network of a next-generation campus network demands a new multi-terabit carrier-class system architecture that scales network capacity several fold. The Nexus 7000 system is a core-class system specifically designed with hardware and software to enable high performance in next-generation campus and data center networks.

As described previously, the Nexus 7000 system decouples the control, management, and data plane within the system. The active supervisor builds the routing adjacencies and forwarding information that gets dynamically updated on each I/O module designed with a distributed forwarding architecture. The system configuration, network protocol state machines, and active supervisor are constantly synchronized to the standby supervisor module for graceful switchover. The distributed forwarding information from the supervisor is stored in a forwarding engine on an I/O module to maintain a local copy of unicast and multicast forwarding information for rapid egress port or module lookup without supervisor involvement. The forwarding engine also provides distributed services like QoS, ACL, Netflow, etc. to optimize throughput and improve application performance with a rapid lookup and forwarding decision process. The multi-stage crossbar fabric module enables backplane

communication between I/O modules. The I/O modules access to the switch fabric is based on VoQ buffer requests and a granting process that involves a central arbiter operating in an active/active state on both supervisor modules.

With a fully-distributed forwarding information and decoupled crossbar switch fabric module design, the active supervisor module switchover remains completely hitless and transparent to other hardware components in the Nexus 7000 system. To provide hitless forwarding, the crossbar fabric module remains operational and the distributed I/O module maintains local forwarding information to seamlessly switch data traffic while the standby supervisor goes through the recovery process.

Figure 1-23 Cisco Nexus 7000 Supervisor Redundancy



Inter- and Intra-Chassis Supervisor Redundancy

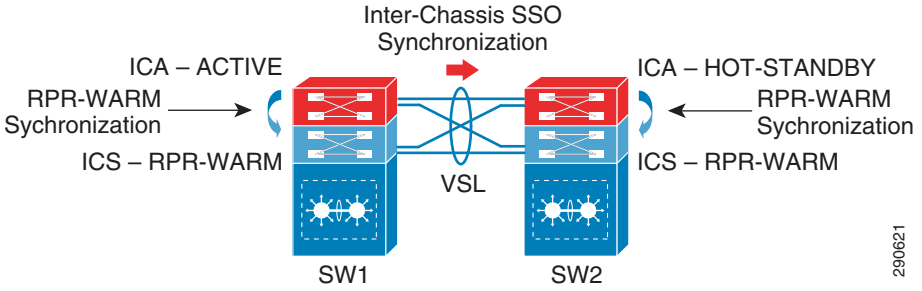
The Cisco VSS solution extends supervisor redundancy by synchronizing SSO and all system internal communication over the special VSL EtherChannel interface between the paired virtual systems. Note that VSS does not currently support stateful intra-chassis supervisor redundancy on each individual virtual node. The virtual switch node running in the active supervisor mode is forced to reset during the switchover. This may disrupt the network topology if it is not deployed with the best practices defined in this design guide. The “triangle”-shaped, distributed, full-mesh fiber paths combined with

single point-to-point EtherChannel or MEC links play a vital role during such network events. During the failure, the new active virtual switch node performs a Layer 3 protocol graceful recovery with its neighbors in order to provide constant network availability over the local interfaces.

6500-E VSS Intra-Chassis RPR-WARM Redundancy

As described earlier the Cisco Catalyst 6500-E introduces innovations aimed at providing intra-chassis stateless supervisor redundancy with the quad-supervisor in the VSS domain. These new innovations allow each redundant supervisor module in each virtual switch chassis to operate in a hybrid role—Supervisor in RPR mode and Distributed Line card. With this hybrid redundancy role, Cisco VSS deployed in quad-sup design operates in a dual redundancy mode—Inter-Chassis SSO with remote virtual-switch chassis and Intra-Chassis RPR within the virtual switch chassis as illustrated in Figure 24.

Figure 24 VSS Quad-Sup Synchronization Process



The ICA supervisor from each virtual-switch chassis synchronizes all critical configurations to the local ICS supervisor module to provide transparent switchover. Even with the stateless intra-chassis redundancy implementation, Cisco VSS offers the ability to maintain full system virtualization, up to date network configuration and protocol state information between both virtual switch chassis. The SSO communication and the synchronization process between ICA supervisors in each virtual switch chassis remains transparent and independent of RPR-WARM. Cisco VSS RPR-WARM provides intra-chassis or local redundancy options, hence it synchronizes the following set of system critical parameters between ICA and ICS supervisor modules:

- Startup-Configuration—Saving the configuration in NVRAM forces the ICA supervisor modules to synchronize their startup configuration with the local in-chassis ICS supervisor module. As part of the SSO synchronization process, the running configuration is synchronized with the remote STANDBY supervisor module in order to maintain an up-to-date configuration.
- BOOT Variables—The boot parameters and registers defined by network administrators are stored as boot parameters in the ROMMON on all four supervisor modules. This synchronization process helps all supervisor modules have consistent bootup information in order to maintain quad-sup redundancy.

- VSS Switch ID—The ICA supervisor module automatically synchronizes the virtual switch ID from its own ROMMON setting to the local ICS ROMMON. Automatically synchronizing the virtual switch ID provides these benefits:
  - Ease of deployment of the in-chassis redundant supervisor module without any additional configuration to synchronize with existing ICA supervisor in the virtual switch.
  - Ability to quickly swap Sup720-10GE module with previous VSS configuration. The ICA supervisor module rewrites old the VSS switch ID to align with its own ID.
- VLAN Database—All VLAN database information is fully synchronized between the ICA and ICS supervisor module.

Deploying Cisco VSS with quad-sup is a plug-n-play operation and no extra configuration is required to enable RPR-WARM. All the intra-chassis ICA and ICS role negotiation and configuration synchronization occurs without any additional settings. The following sample **show** command depicts the SSO synchronized state between the ICA supervisors of SW1 and SW2, which are also running full Cisco IOS software. The in-chassis redundant supervisor modules have been initialized with special Sup720-LC IOS software that enables the hybrid role capability to synchronize RPR-WARM with the local ICA module:

```
cr22-6500-LB#show switch virtual redundancy | inc Switch|Software

My Switch Id = 1
Peer Switch Id = 2

Switch 1 Slot 5 Processor Information :
Current Software state =
Image Version = Cisco IOS Software, s72033_rp Software (s72033_rp-ADVENTERPRISEK9_WAN-M),
Version 12.2(33)SXI4, RELEASE SOFTWARE (fc3)

Switch 1 Slot 6 Processor Information :
Current Software state = RPR-Warm
Image Version = Cisco IOS Software, s72033_lc Software (s72033_lc-SP-M), Version
12.2(33)SXI4, RELEASE SOFTWARE (fc3)

Switch 2 Slot 5 Processor Information :
Current Software state = STANDBY HOT (switchover target)
Image Version = Cisco IOS Software, s72033_rp Software (s72033_rp-ADVENTERPRISEK9_WAN-M),
Version 12.2(33)SXI4, RELEASE SOFTWARE (fc3)

Switch 2 Slot 6 Processor Information :
Current Software state = RPR-Warm
Image Version = Cisco IOS Software, s72033_lc Software (s72033_lc-SP-M), Version
12.2(33)SXI4, RELEASE SOFTWARE (fc3)
```

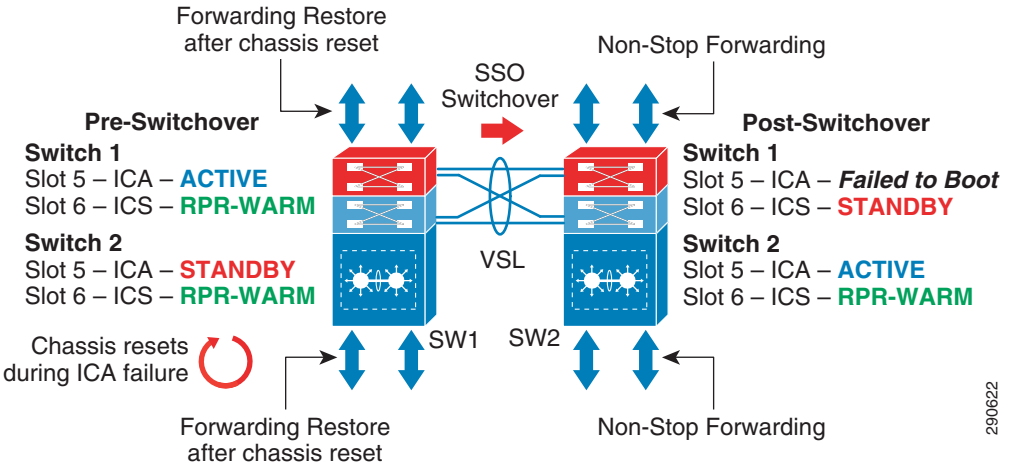
Note Users must use the **show switch virtual redundancy** or **show module** commands to verify the current role of the quad-sup and their status. The **show redundancy** command continues to provide dual-sup role and status information, however it doesnot provide any quad-sup specific information.

6500-E VSS Intra-Chassis Supervisor Switchover

The design using the quad-sup stateless intra-chassis redundancy option is same as the dual-sup VSS design. During an ICA supervisor (ACTIVE or STANDBY) failure, the entire chassis and all modules in the impacted chassis are reset. If the original ICA supervisor fails to reboot, the redundant ICS supervisor module takes over chassis ownership and bootup in the ICA role in order to restore the original network capacity and reliability in the virtual switch system. Administrative reset or failure of a redundant ICS supervisor module does not cause virtual switch chassis reset, since it also acts as a distributed linecard and is not actively handling any of the control plane or switch fabric ownership in the chassis.

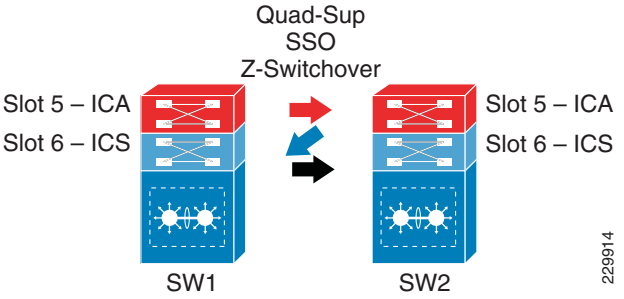
Deploying Catalyst 6500-E in VSS mode with quad-sup capability continues to provide the same level of inter-chassis SSO redundancy as the dual-sup design. The SSO ACTIVE supervisor synchronizes all the run time and stateful information from the SSO HOT-STANDBY supervisor module that resides on the peer virtual switch chassis. Hence during ACTIVE supervisor failure, the operation of the network remains transparent, as the remote virtual switch gracefully takes over software control plane ownership (as illustrated in Figure 25).

Figure 25 VSS Quad-Sup Switchover



Since the VSS domain is now equipped with a quad-sup, the logic utilized to synchronize information and identify the redundancy role of each supervisor module is different than that utilized in the dual-sup deployment. Depending on the reason supervisor reset, Cisco VSS internally sets a bootup parameter that modifies ICA or ICS role preference in the next bootup process. Such software design provides built-in system reliability in order to detect the fault, take over ICA ownership, and stabilize the overall virtual switch and network operation. This integrated quad-sup switchover capability is known as "Quad-Sup SSO Z-switchover" and is transparent to the user. It does not require any manual user intervention for optimization. Figure 26 illustrates the deterministic supervisor roles that occur during multiple switchover events.

Figure 26 VSS Quad-Sup Z-Switchover Process



Network administrators can verify the target supervisor module of the next SSO switchover event using the **show switch virtual redundancy exec** command.

Implementing SSO Redundancy

To deploy SSO supervisor redundancy, it is important to remember that both supervisor modules must be identical in hardware type, software version, and all the internal hardware components—memory and bootflash must be the same to provide complete operational transparency during failure. The default redundancy mode on all modular Catalyst and Nexus 7000 series platforms is SSO. Hence it does not require any additional configuration to enable SSO redundancy. The SSO redundant status can be verified using the following command on each recommended system:

Cisco IOS—Catalyst 6500-E VSS Mode

```
cr23-VSS-Core#show switch virtual redundancy
My Switch Id = 1
Peer Switch Id = 2
Configured Redundancy Mode = sso
Operating Redundancy Mode = sso

Switch 1 Slot 5 Processor Information :
```

```
-------------------------------------------
Current Software state = ACTIVE
<snippet>
Fabric State = ACTIVE
Control Plane State = ACTIVE

Switch 2 Slot 5 Processor Information :
-------------------------------------------
Current Software state = STANDBY HOT (switchover target)
<snippet>
Fabric State = ACTIVE
Control Plane State = STANDBY
```

Cisco IOS—Catalyst 4500-E

```
cr40-4507-1#show redundancy states
       my state = 13 -ACTIVE
     peer state = 8  -STANDBY HOT
       <snip>
Redundancy Mode (Operational) = Stateful Switchover
Redundancy Mode (Configured)  = Stateful Switchover
Redundancy State              = Stateful Switchover
    Manual Swact = enabled
 Communications = Up
<snip>
```

Cisco NX-OS—Cisco Nexus 7000

```
cr35-N7K-Core1# show redundancy status
Redundancy mode
---------------
      administrative:   HA
         operational:   HA
This supervisor (sup-5)
-----------------------
    Redundancy state:   Active
    Supervisor state:   Active
      Internal state:   Active with HA standby
Other supervisor (sup-6)
------------------------
    Redundancy state:   Standby
    Supervisor state:   HA standby
      Internal state:   HA standby
```

Non-Stop Forwarding (NSF)

When implementing NSF technology in systems using SSO redundancy mode, network disruptions are transparent to campus users and applications and high availability is provided even during periods where the control plane processing module (Supervisor/Route-Processor) is reset. During a failure, the underlying Layer 3 NSF-capable protocols perform graceful network topology re-synchronization. The preset forwarding information on the redundant processor or distributed linecard hardware remains intact and continues to switch network packets. This service availability significantly lowers the Mean Time To Repair (MTTR) and increases the Mean Time Between Failure (MTBF) to achieve the highest level of network availability.

NSF is an integral part of a routing protocol and depends on the following fundamental principles of Layer 3 packet forwarding:

- *Cisco Express Forwarding (CEF)*—CEF is the primary mechanism used to program the network path into the hardware for packet forwarding. NSF relies on the separation of the control plane update and the forwarding plane information. The control plane provides the routing protocol with a graceful restart and the forwarding plane switches packets using hardware acceleration where available. CEF enables this separation by programming hardware with FIB entries in all Catalyst switches. This ability plays a critical role in NSF/SSO failover.
- *Routing protocol*—The motivation behind NSF is route convergence avoidance. From a protocol operation perspective, this requires the adjacent routers to support a routing protocol with special intelligence that allows a neighbor to be aware that NSF-capable routers can undergo switchover so that its peer can continue to forward packets. This may bring its adjacency to a hold-down state (NSF recovery mode) for a brief period and request that routing protocol information be resynchronized.

A router that has the capability for continuous forwarding during a switchover is *NSF-capable*. Devices that support the routing protocol extensions such that they continue to forward traffic *to* a restarting router are *NSF-aware*. A Cisco device that is NSF-capable is also NSF-aware. The NSF capability must be manually enabled on each redundant system on a per-routing-protocol basis. The NSF-aware function is enabled by default on all Layer 3 platforms. describes the Layer 3 NSF-capable and NSF-aware platforms deployed in the campus network environment.

Implementing EIGRP NSF Capability

The following sample configuration illustrates how to enable the NSF capability within EIGRP (the same procedure applies to OSPF) on each Layer 3 campus LAN/WAN system deployed with redundant supervisors and route-processors or in virtual-switching modes (i.e., Cisco VSS, Catalyst 4500E, and StackWise Plus). EIGRP NSF capability is enabled by default on the Cisco Nexus 7000 system:

Cisco IOS—Catalyst Platforms

```
cr23-vss-core(config)#router eigrp 100
cr23-vss-core (config-router)#nsf
```

```
cr23-vss-core #show ip protocols | inc NSF
*** IP Routing is NSF aware ***
  EIGRP NSF-aware route hold timer is 240
  EIGRP NSF enabled
     NSF signal timer is 20s
     NSF converge timer is 120s
```

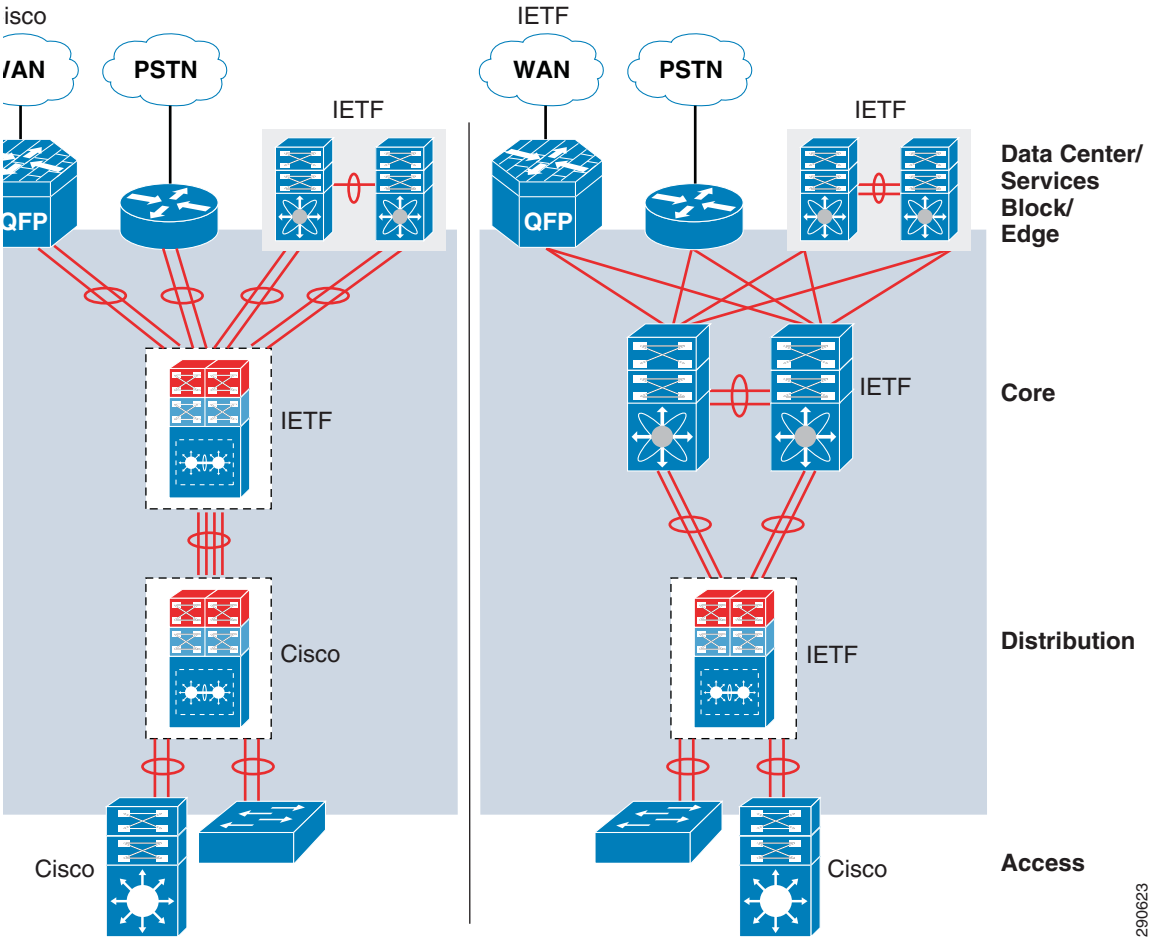
Cisco NX-OX—Nexus 7000

```
cr35-N7K-Core1#show ip eigrp | inc Grace
  Graceful-Restart: Enabled
```

Implementing OSPF NSF Capability

The OSPF NSF capability and helper function in Cisco IOS-based systems is supported in two modes—Cisco proprietary and IETF standard-based. The NX-OS running on the Nexus 7000 system supports OSPF NSF capability and helper function based on the IETF standard. Depending on the campus network design, the network administrator must implement the correct OSPF NSF capability between two adjacent Layer 3 campus systems to recognize and respond to the graceful restart capability in an OSPF TLV packet during supervisor switchover. By default, enabling OSPF NSF capability on Cisco IOS routers and switches enables the Cisco proprietary NSF function, whereas the IETF NSF capability is by default enabled on the Nexus 7000 system. Figure 27 illustrates the recommended OSPF NSF capability in each campus network design.

Figure 27 Recommended OSPF NSF Capability in Campus



Cisco IOS—Cisco NSF Capability

```
cr23-vss-core(config)#router ospf 100
cr23-vss-core (config-router)#nsf
cr23-vss-core# show ip ospf | inc Non-Stop|helper
 Non-Stop Forwarding enabled
 IETF NSF helper support enabled
 Cisco NSF helper support enabled
```

Cisco IOS—IETF NSF Capability

```
cr23-vss-core(config)#router ospf 100
cr23-vss-core(config-router)#nsf ietf
cr23-vss-core#show ip ospf | inc Non-Stop|helper
 IETF Non-Stop Forwarding enabled
 IETF NSF helper support enabled
 Cisco NSF helper support enabled
```

Cisco NX-OS—IETF NSF Capability

```
!IETF OSPF NSF capability is enabled by default
cr35-N7K-Core1#show ip ospf | inc Stateful|Graceful
 Stateful High Availability enabled
 Graceful-restart is configured
```

Graceful Restart Example

The following example demonstrates how the EIGRP protocol gracefully recovers when active supervisor/chassis switchover on a Cisco VSS and Nexus 7000 core system is forced by a reset:

- Cisco IOS

```
cr23-VSS-Core#redundancy force-switchover
This will reload the active unit and force switchover to standby[confirm]y

! VSS active system reset will force all linecards and ports to go down
!the  following logs confirms connectivity loss to core system
%LINK-3-UPDOWN: Interface TenGigabitEthernet2/1/2, changed state to down
%LINK-3-UPDOWN: Interface TenGigabitEthernet2/1/4, changed state to down

! Downed interfaces are automatically removed from EtherChannel/MEC,
! however additional interface to new active chassis retains port-channel in up/up
state
%EC-SW1_SP-5-UNBUNDLE: Interface TenGigabitEthernet2/1/2 left the port-channel
Port-channel100
%EC-SW1_SP-5-UNBUNDLE: Interface TenGigabitEthernet2/1/4 left the port-channel
Port-channel100

! EIGRP protocol completes graceful recovery with new active virtual-switch.
%DUAL-5-NBRCHANGE: EIGRP-IPv4:(613) 100: Neighbor 10.125.0.12 (Port-channel100) is
resync: peer graceful-restart
```

- Cisco NX-OS

```
cr35-N7K-Core1#system switchover

! EIGRP protocol completes graceful recovery with new active supervisor
%DUAL-5-NBRCHANGE: EIGRP-IPv4 100: Neighbor 10.125.10.1 (Port-channel3) is resync:
peer graceful-restart
```

NSF Timers

The OSPF routing information stalls the routes and forwarding information for several seconds to gracefully recover the OSPF adjacencies and re-synchronize the database. By default the OSPF NSF timer on Cisco Catalyst switches is 120 seconds and the Nexus 7000 system can hold routing information for up to 60 seconds. Lowering the timer values may abruptly terminate graceful recovery, which can cause network instability. The default timer setting is tuned for a well-structured and concise campus LAN network topology. It is recommended to retain the default route hold timers in the network unless it is observed that NSF recovery takes more than the default values.

NSF/SSO Recovery Analysis

As described in a previous section, the NSF/SSO implementation and its recovery process differ on the Nexus 7000, Catalyst 4500E (Intra-Chassis), and Catalyst 6500-E VSS (Inter-Chassis) in the Borderless Campus LAN design. In each deployment scenario, the Cisco enterprise solution architecture validated the network recovery and application performance by inducing several types of active supervisor faults that trigger Layer 3 protocol graceful recovery. During each test, the switches continued to provide network accessibility during the recovery stage.

The Nexus 7000 and Catalyst 4500E systems retain the operational and forwarding state of the linecard and fabric modules for non-stop forwarding while the new active supervisor module goes through the graceful recovery process.

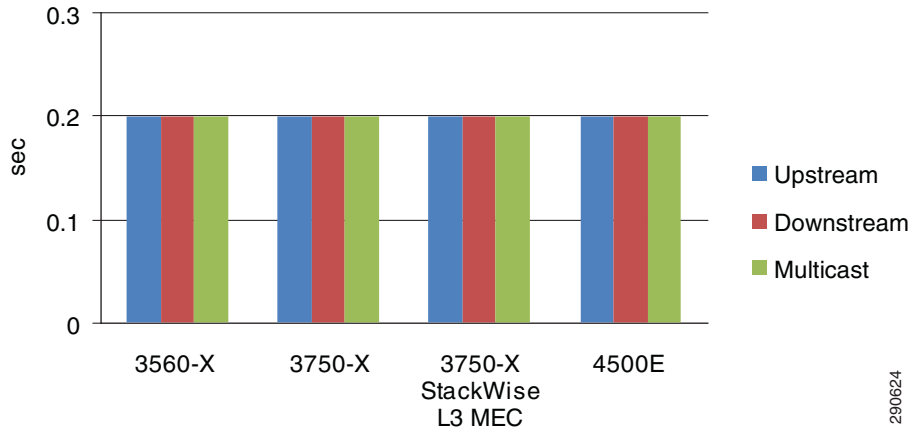
The inter-chassis SSO implementation in Catalyst 6500-E VSS differs from the single-chassis redundant implementation in that during active virtual switch node failure the entire chassis and all the linecards installed reset. However, with Layer 2/3 MEC links, the network protocols and forwarding information remain protected via the remote virtual switch node that can provide seamless network availability.

Catalyst 6500-E VSS NSF/SSO Recovery Analysis

As described earlier, in dual-sup or quad-sup Cisco VSS designs, the entire Catalyst 6500-E chassis and all installed linecard modules are reset during an in-chassis active (SSO ACTIVE or HOT-STANDBY) virtual switch switchover event. With a diverse full-mesh fiber network design, the Layer 2/Layer 3 remote device perceives this event as a loss of a member link since the alternate link to the standby switch is in an operational and forwarding state. The standby virtual switch detects the loss of the VSL Etherchannel and transitions into the active role and initializes Layer 3 protocol graceful recovery with the remote devices. Since there are no major network topology changes and member links are still in an operational state, the NSF/SSO recovery in Catalyst 6500-E VSS system is identical to the scenario where individual links are lost.

Additionally, the Cisco Catalyst 6500-E supports Multicast Multilayer Switching (MMLS) NSF with SSO, thereby enabling the system to maintain the multicast forwarding state in PFC3- and DFC3-based hardware during an active virtual switch reset. The new active virtual switch reestablishes PIM adjacency while continuing to switch multicast traffic based on pre-switchover programmed information (see Figure 28).

Figure 28 Catalyst 6500-E Dual-Sup and Quad-Sup VSS NSF/SSO Recovery Analysis



Nexus 7000 NSF/SSO Recovery Analysis

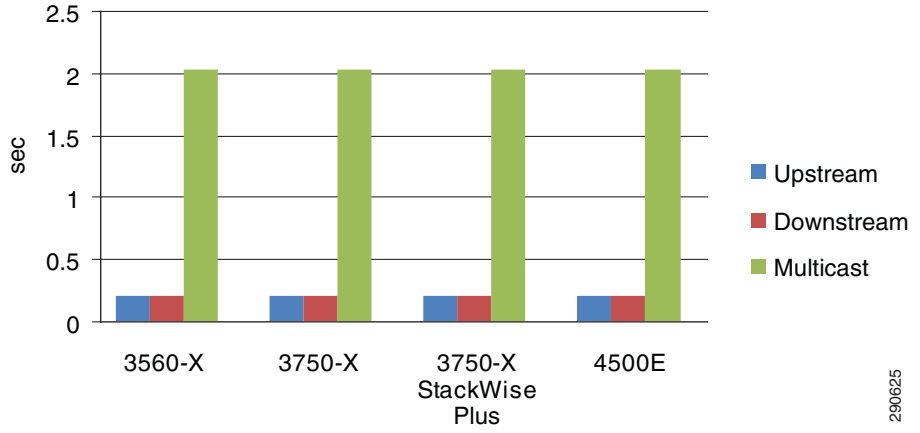
Since the Nexus 7000 system is designed with a distributed architecture to decouple the centralized control plane from the distributed data plane, the supervisor switchover process remains transparent and hitless to the network. During the supervisor switchover process, the distributed I/O and crossbar modules remain intact with synchronized forwarding information across the system. The egress forwarding information lookup and network services, such as QoS and ACL, are perfomed at the I/O module and the campus backbone network remains hitless with zero packet loss during an active or standby supervisor switchover event. The campus core remains hitless when the Cisco Nexus 7000 system is in various supervisor fault conditions, such as administrative forced switchover, manual OIR, or a hardware or software crash.

Catalyst 4500E NSF/SSO Recovery Analysis

Figure 29 illustrates an intra-chassis NSF/SSO recovery analysis for the Catalyst 4500E chassis deployed with Sup7-E, Sup6-E, or Sup6L-E in redundant mode. With EIGRP NSF/SSO capability, the unicast traffic consistently recovers within 200 msec. or less. However, the Catalyst 4500E does not currently support redundancy for Layer 3 multicast routing and forwarding information. Therefore,

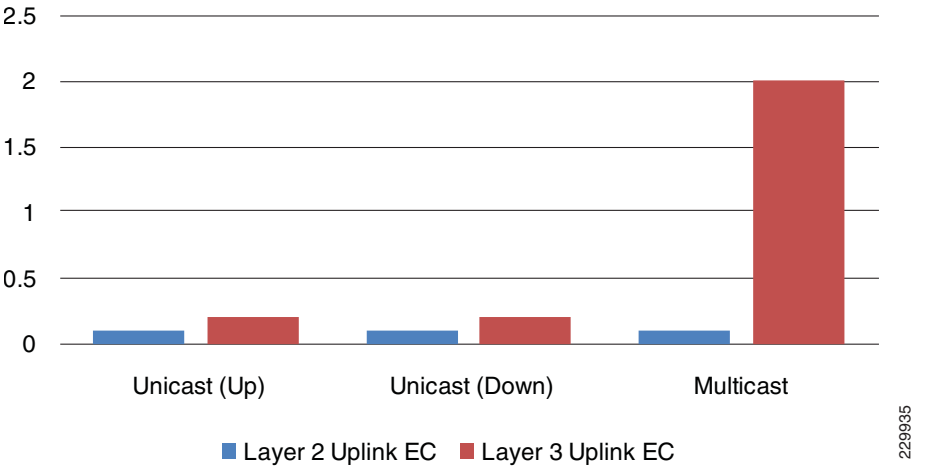
there may be an approximately two second loss of multicast traffic, since the switch has to reestablish all the multicast routing and forwarding information during the switchover event associated with the Sup7-E, Sup6-E, or Sup6L-E.

Figure 29 Catalyst 4500E Distribution Layer NSF/SSO Recovery Analysis



At the campus access layer, the Cisco Catalyst 4500E series platform provides unparalleled high availability to network applications with its unique forwarding architecture. During supervisor switchover, all synchronized Layer 2 forwarding information remains on the standby supervisor module that gracefully takes over control plane ownership; with the uplink port active on the failed supervisor module, the uplink capacity and Layer 2 adjacency are unaffected. Due to the highly-resilient platform design, network recovery is low sub-seconds for unicast and multicast traffic, as illustrated in Figure 30, when the Cisco Catalyst 4500E in the access layer is deployed in Multilayer and Routed-Access mode.

Figure 30 Catalyst 4500E Access Layer NSF/SSO Recovery Analysis



Catalyst 4500E Standby Supervisor Failure and Recovery Analysis

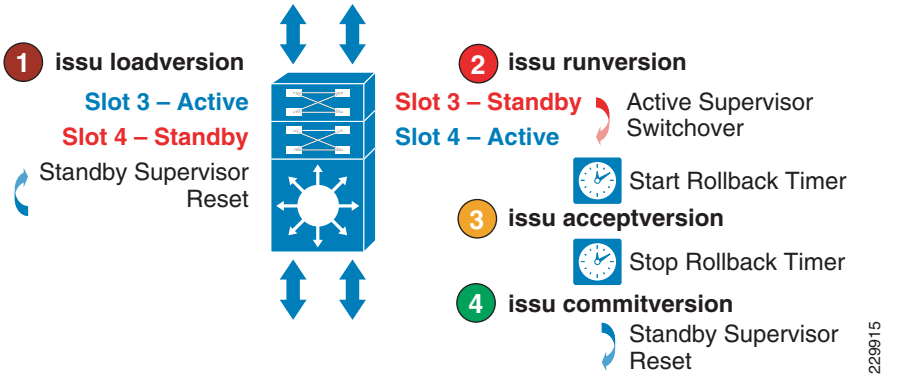
The standby Sup7-E, Sup6-E, or Sup6L-E supervisor remains in redundant mode while the active supervisor is in an operational state. If the standby supervisor gets reset or re-inserted, protocol graceful recovery is not triggered, nor are there any changes in network topology. Hence the standby supervisor remains completely transparent to the system and to rest of the network. The uplink port of the standby supervisor remains in an operational and forwarding state and the network bandwidth capacity remains intact during a standby supervisor soft switchover event.

9 Implementing Operational Resiliency

Path redundancy is often used to facilitate access during maintenance activity. However, single standalone systems are single points of failure and this type of network design simply does not provide user access if a critical node is taken out of service. Leveraging enterprise-class high availability features like NSF/SSO in the distribution and core layer Catalyst 4500E and 6500-E Series platforms enables support for ISSU and real-time network upgrade capability. Using ISSU and eFSU technology, the network administrator can upgrade the Cisco IOS software to implement new features, software bug fixes, or critical security fixes in real time.

Catalyst 4500E ISSU Software Design and Upgrade Process

Figure 31 Catalyst 4500E Manual ISSU Software Upgrade Process



ISSU Software Upgrade Pre-Requisite

ISSU Compatibility Matrix

When a redundant Catalyst 4500E system is brought up with a different Cisco IOS software version, the ISSU stored compatibility matrix information is analyzed internally to determine interoperability between the software running on the active and standby supervisors. ISSU provides SSO compatibility between several versions of software releases shipped during a 18 month period. Prior to upgrading the software, the network administrator must verify ISSU software compatibility with the following **show** command. Incompatible software may cause the standby supervisor to boot in RPR mode, which may result in a network outage:

```
cr24-4507e-MB#show issu comp-matrix stored
Number of Matrices in Table = 1
                My Image ver:  12.2(53)SG
                Peer Version    Compatibility
                -----------------    -------------
                12.2(44)SGBase(2)
                12.2(46)SG              Base(2)
                12.2(44)SG1             Base(2)
    …
```

Software

Prior to starting the software upgrade process, it is recommended to copy the old and new Cisco IOS software on the Catalyst 4500E active and standby supervisor into local file systems—Bootflash or Compact Flash.

```
cr24-4507e-MB#dir slot0:
Directory of slot0:/
 1  -rw- 25442405 Nov 23 2009 17:53:48 -05:00  cat4500e-entservicesk9-mz.122-53.SG1 ← new image
 2  -rw- 25443451 Aug 22 2009 13:26:52 -04:00  cat4500e-entservicesk9-mz.122-53.SG  ← old image

cr24-4507e-MB#dir slaveslot0:
Directory of slaveslot0:/

  1  -rw- 25443451 Aug 22 2009 13:22:00 -04:00  cat4500e-entservicesk9-mz.122-53.SG ← old image
  2  -rw- 25442405 Nov 23 2009 17:56:46 -05:00  cat4500e-entservicesk9-mz.122-53.SG1 ← new image
```

Configuration

It is recommended to save the running configuration to NVRAM and other local or remote locations such as bootflash or TFTP server prior to upgrading IOS software.

Boot Variable and String

The system default boot variable is defined to boot from the local file system. Make sure the default setting is not changed and the configuration register is set to 0x2102.

Modify the boot string to point to the new image to boot from a new IOS software version after the next reset triggered during the ISSU upgrade process. Refer to following URL for additional ISSU pre-requisites:

http://www.cisco.com/en/US/partner/docs/switches/lan/catalyst4500/12.2/53SG/configuration/issu.html#wp1072849

Catalyst 4500E Manual ISSU Software Upgrade Procedure

This subsection provides the manual software upgrade procedure for a Catalyst 4500E deployed in the enterprise campus LAN network design in several different roles—access, distribution, core, collapsed core, and Metro Ethernet WAN edge. The manual ISSU upgrade capability is supported on Catalyst 4500E Sup7-E, Sup6-E, and Sup6L-E supervisors running the Cisco IOS Enterprise feature set. However the automatic ISSU upgrade capability is only supported on the next generation Catalyst 4500E Sup7-E supervisor module.

In the following sample output, the Sup6-E supervisor is installed in Slot3 and Slot4 respectively. The Slot3 supervisor is in the SSO Active role and the Slot4 supervisor is in Standby role. Both supervisors are running identical 12.2(53)SG Cisco IOS software versions and are fully synchronized with SSO.

```
cr24-4507e-MB#show module | inc Chassis|Sup|12.2
Chassis Type : WS-C4507R-E
!Common Supervisor Module Type
 3     6  Sup 6-E 10GE (X2), 1000BaseX (SFP)      WS-X45-SUP6-E      JAE1132SXQ3
 4     6  Sup 6-E 10GE (X2), 1000BaseX (SFP)      WS-X45-SUP6-E      JAE1132SXRQ
!Common operating system version
 3     0021.d8f5.45c0 to 0021.d8f5.45c5 0.4 12.2(33r)SG     ( 12.2(53)SG       Ok
 4     0021.d8f5.45c6 to 0021.d8f5.45cb 0.4 12.2(33r)SG     ( 12.2(53)SG       Ok
!SSO Synchronized
 3     Active Supervisor       SSO Active
 4     Standby Supervisor       SSO Standby hot
```

The following provides the step-by-step procedure to upgrade the Cisco IOS Release 12.2(53)SG to 12.2(53)SG1 Cisco IOS release without causing network topology and forwarding disruption. Prior to issuing the **issu commitversion** command, the ISSU software upgrade can be aborted at any stage by issuing the **issu abortversion** command if any failure is detected.

1. *ISSU loadversion*—This first step will direct the active supervisor to initialize the ISSU software upgrade process.

   ```
   cr24-4507e-MB#issu loadversion 3 slot0:cat4500e-entservicesk9-mz.122-53.SG1 4
   slaveslot0: cat4500e-entservicesk9-mz.122-53.SG1
   ```

   After issuing the above command, the active supervisor ensures the new IOS software is downloaded on both supervisors' file systems and performs several additional checks on the standby supervisor for the graceful software upgrade process. ISSU changes the boot variable with the new IOS software version if no errors are found and resets the standby supervisor module.

   ```
   %RF-5-RF_RELOAD: Peer reload. Reason: ISSU Loadversion
   ```

Note Resetting the standby supervisor will not trigger a network protocol graceful recovery and all standby supervisor uplink ports will remain in operational and forwarding state for the transparent upgrade process.

With the broad range of ISSU version compatibility used in conducting SSO communication, the standby supervisor will then successfully bootup again in its original standby state:

```
cr24-4507e-MB#show module | inc Chassis|Sup|12.2
Chassis Type : WS-C4507R-E
! Common Supervisor Module Type
 3     6  Sup 6-E 10GE (X2), 1000BaseX (SFP)      WS-X45-SUP6-E      JAE1132SXQ3
 4     6  Sup 6-E 10GE (X2), 1000BaseX (SFP)      WS-X45-SUP6-E      JAE1132SXRQ
! Mismatch operating system version
 3     0021.d8f5.45c0 to 0021.d8f5.45c5 0.4 12.2(33r)SG( 12.2(53)SG       Ok
 4     0021.d8f5.45c6 to 0021.d8f5.45cb 0.4 12.2(33r)SG( 12.2(53)SG1      Ok
!SSO Synchronized
 3     Active Supervisor       SSO Active
 4     Standby Supervisor       SSO Standby hot
```

This bootup process will force the active supervisor to re-synchronize all SSO redundancy and checkpoints, VLAN database, and forwarding information with the standby supervisor and will notify the user to proceed with the next ISSU step.

```
%C4K_REDUNDANCY-5-CONFIGSYNC: The config-reg has been successfully synchronized to the
standby supervisor
%C4K_REDUNDANCY-5-CONFIGSYNC: The startup-config has been successfully synchronized to
the standby supervisor
%C4K_REDUNDANCY-5-CONFIGSYNC: The private-config has been successfully synchronized to
the standby supervisor
%C4K_REDUNDANCY-5-CONFIGSYNC_RATELIMIT: The vlan database has been successfully
synchronized to the standby supervisor

%ISSU_PROCESS-7-DEBUG: Peer state is [ STANDBY HOT ]; Please issue the runversion
command
```

2. *ISSU runversion*—After ensuring that the newly-loaded software is stable on the standby supervisor, the network administrator must proceed to the second step:

```
cr24-4507e-MB#issu runversion 4
This command will reload the Active unit.  Proceed ? [confirm]y
%RF-5-RF_RELOAD: Self reload. Reason: Admin ISSU runversion CLI
%SYS-5-RELOAD: Reload requested by console. Reload reason: Admin ISSU runversion
```

This step will force the current active supervisor to reset itself, thereby triggering network protocol graceful recovery with peer devices. However the uplink ports of the active supervisor remain intact and the data plane is not impacted during the switchover process. From an overall network perspective, the active supervisor reset caused by the **issu runversion** command will be no different than in similar switchover procedures (e.g., administrator-forced switchover or supervisor online insertion and removal). During the entire software upgrade procedure, this is the only step that performs SSO-based network graceful recovery. The following syslog on various Layer 3 systems confirm stable and EIGRP graceful recovery with the new supervisor running the new Cisco IOS software version.

- NSF-Aware Core

```
cr23-VSS-Core#
%DUAL-5-NBRCHANGE: EIGRP-IPv4:(415) 100: Neighbor 10.125.0.15 (Port-channel102) is
resync: peer graceful-restart
```

- NSF-Aware Layer 3 Access

```
cr24-3560-MB#
%DUAL-5-NBRCHANGE: EIGRP-IPv4:(100) 100: Neighbor 10.125.0.10 (Port-channel1) is
resync: peer graceful-restart
```

The previously active supervisor module will boot up in the standby role with the older IOS software version instead of the new IOS software version.

```
cr24-4507e-MB#show module | inc Chassis|Sup|12.2
```

```
Chassis Type : WS-C4507R-E
! Common Supervisor Module Type
 3      6   Sup 6-E 10GE (X2), 1000BaseX (SFP)      WS-X45-SUP6-E      JAE1132SXQ3
 4      6   Sup 6-E 10GE (X2), 1000BaseX (SFP)      WS-X45-SUP6-E      JAE1132SXRQ
! Mismatch operating system version
 3      0021.d8f5.45c0 to 0021.d8f5.45c5 0.4 12.2(33r)SG( 12.2(53)SG        Ok
 4      0021.d8f5.45c6 to 0021.d8f5.45cb 0.4 12.2(33r)SG( 12.2(53)SG1       Ok
!SSO Synchronized
 3     Active Supervisor       SSO Standby hot
 4     Standby Supervisor       SSO Active
```

This safeguarded software design provides an opportunity to roll back to the previous IOS software if the system upgrade causes any network abnormalities. At this stage, ISSU automatically starts internal rollback timers to re-install the old IOS image. The default rollback timer is up to 45 minutes, which provides a network administrator with an opportunity to perform several sanity checks. In small to mid-size network designs, the default timer may be sufficient. However, for large networks, network administrators may want to increase the timer up to two hours:

```
cr24-4507e-MB#show issu rollback-timer
       Rollback Process State = In progress
      Configured Rollback Time = 45:00
       Automatic Rollback Time = 19:51
```

The system will notify the network administrator with the following, instructing them to move to the next ISSU upgrade step if no stability issues are observed and all the network services are operating as expected.

```
%ISSU_PROCESS-7-DEBUG: Peer state is [ STANDBY HOT ]; Please issue the acceptversion
command
```

3. *ISSU acceptversion* (Optional)—This step provides confirmation from the network administrator that the system and network is stable after the IOS install and they are ready to accept the new IOS software on the standby supervisor. This step stops the rollback timer and instructs the network administrator to issue the final commit command. The network admnistrator can optionally skip this upgrade step and issue the final commit within the rollback timer window:

```
cr24-4507e-MB#issu acceptversion 4
% Rollback timer stopped. Please issue the commitversion command.

cr24-4507e-MB#show issu rollback-timer
     Rollback Process State = Not in progress
   Configured Rollback Time = 45:00


cr24-4507e-MB#show module | inc Chassis|Sup|12.2
Chassis Type : WS-C4507R-E
! Common Supervisor Module Type
 3      6   Sup 6-E 10GE (X2), 1000BaseX (SFP)      WS-X45-SUP6-E      JAE1132SXQ3
```

```
 4     6  Sup 6-E 10GE (X2), 1000BaseX (SFP)     WS-X45-SUP6-E     JAE1132SXRQ
! Mismatch operating system version
 3     0021.d8f5.45c0 to 0021.d8f5.45c5 0.4 12.2(33r)SG( 12.2(53)SG       Ok
 4     0021.d8f5.45c6 to 0021.d8f5.45cb 0.4 12.2(33r)SG( 12.2(53)SG1      Ok
!SSO Synchronized
 3    Active Supervisor      SSO Standby hot
 4    Standby Supervisor      SSO Active
```

4. *ISSU commitversion*—This final ISSU step forces the active supervisor to synchronize its configuration with the standby supervisor and forces it to reboot with the new IOS software. This stage concludes the ISSU upgrade procedure and the new IOS version is permanently committed on both supervisor modules. If for some reason the network administrator wants to rollback to the older image, it is recommended to perform an ISSU-based downgrade procedure to retain the network operational state without any downtime.

```
cr24-4507e-MB#issu commitversion 3
Building configuration...
Compressed configuration from 24970 bytes to 10848 bytes[OK]
%C4K_REDUNDANCY-5-CONFIGSYNC: The private-config has been successfully synchronized to
the standby supervisor
%RF-5-RF_RELOAD: Peer reload. Reason: ISSU Commitversion

cr24-4507e-MB#show module | inc Chassis|Sup|12.2
Chassis Type : WS-C4507R-E
! Common Supervisor Module Type
 3     6  Sup 6-E 10GE (X2), 1000BaseX (SFP)     WS-X45-SUP6-E     JAE1132SXQ3
 4     6  Sup 6-E 10GE (X2), 1000BaseX (SFP)     WS-X45-SUP6-E     JAE1132SXRQ
! Common new operating system version
 3     0021.d8f5.45c0 to 0021.d8f5.45c5 0.4 12.2(33r)SG( 12.2(53)SG1      Ok
 4     0021.d8f5.45c6 to 0021.d8f5.45cb 0.4 12.2(33r)SG( 12.2(53)SG1      Ok

!SSO Synchronized
 3    Active Supervisor      SSO Standby hot
 4    Standby Supervisor      SSO Active
```

Catalyst 4500E Automatic ISSU Software Upgrade Procedure

The network administrator can use the automatic ISSU upgrade method for large Catalyst 4500E Sup7-E-based campus networks once the manual ISSU upgrade procedure is successfully performed. It is recommended that the status of all network communication, operation, and manageability components on the Catalyst 4500E system now running with the new Cisco IOS-XE software be verified. Once the stability of the new IOS-XE software is confirmed, the network administrator can start a single-step automatic ISSU upgrade procedure on the remaining systems. Cisco IOS-XE also provides the flexibility to program the system for an automatic ISSU upgrade based on a user-defined future time.

The new **issu changeversion** command automates the upgrade of all four ISSU upgrade procedures into a single step that does not require manual intervention from the network administrator. The syntax must include the new targeted Cisco IOS-XE software to be installed on both supervisor modules. This gets set in the BOOT variable and the rest of the upgrade process becomes fully automated. Even with the automatic ISSU upgrade procedure, the standby supervisor module still gets reset when **issu loadversion**, **issu runversion**, and **issu commitversion** are executed by the software.

The following provides the single-step procedure to upgrade the Cisco IOS-XE Release from 3.1.0SG to pre- release Cisco IOS-XE software without causing network topology and forwarding disruption. Like the manual upgrade steps, the automatic ISSU upgrade can also be aborted at any stage by issuing the **issu abortversion** command:

1. ISSU changeversion—The only manual step to initialize automatic ISSU upgrade procedure on Cisco Catalyst 4500E system with Sup7-E supervisor module. The Catalyst 4500E system ensures the correct location and Cisco IOS-XE software information to initialize the automated ISSU software upgrade procedure. Both supervisors perform file system and other checks on the standby supervisor in order to ensure a graceful software upgrade process. The automatic ISSU procedure will modify the boot variable with the new IOS-XE software version if no errors are found. The rest of the ISSU upgrade procedure will automate starting with the standby supervisor module being force reset with a new software version. The network administrator can monitor the automated upgrade status and has flexibility to abort the entire process if any abnormal condition occurs during the new software installation process:

```
cr19-4507-MB#show issu state detail | exclude Pre|Post
    Slot = 3
    RP State = Active
    ISSU State = Init
    Operating Mode = Stateful Switchover
    Current Image = bootflash:cat4500e-universalk9.SPA.03.01.00.SG.150-1.XO.bin

    Slot = 4
    RP State = Standby
    ISSU State = Init
    Operating Mode = Stateful Switchover
    Current Image = bootflash: cat4500e-universalk9.SPA.03.01.00.SG.150-1.XO.bin

cr19-4507-MB#dir bootflash:
59009  -rw-  <truncated>  cat4500e-universalk9.SPA.03.01.00.SG.150-1.XO.bin <- old
image
29513  -rw-  <truncated>  cat4500e-universalk9.SSA.03.01.01.0.74.150.2.SG.bin <- new
image

cr19-4507-MB#dir slavebootflash:
14769  -rw-  <truncated>  cat4500e-universalk9.SPA.03.01.00.SG.150-1.XO.bin <- old
image
14758  -rw-  <truncated>  cat4500e-universalk9.SSA.03.01.01.0.74.150.2.SG.bin <- new
image
```

```
cr19-4507-LB#issu changeversion
bootflash:cat4500e-universalk9.SSA.03.01.01.0.74.150.2.SG.bin

!Automatically triggers issu loadversion that resets current standby supervisor module
% 'issu changeversion' is now executing 'issu loadversion'
% issu loadversion executed successfully, Standby is being reloaded
% changeversion finished executing loadversion, waiting for standby to reload and
reach SSO ...

cr19-4507-MB#show issu state detail | exclu Pre|Post
Slot = 3
RP State = Active
ISSU State = Load Version
Changeversion = TRUE
Operating Mode = not reached
Current Image = bootflash:cat4500e-universalk9.SPA.03.01.00.SG.150-1.XO.bin
Standby information is not available because it is in 'DISABLED' state

!Automatically triggers issu runversion that resets current active supervisor module.
This step will force SSO switchover, the new supervisor module gracefully recovers
protocol with neighbosr. Automatically starts ISSU roll-back timer

%INSTALLER-7-ISSU_OP_SUCC:  issu changeversion is now executing 'issu runversion'
%INSTALLER-7-ISSU_OP_SUCC:  issu changeversion successfully executed 'issu runversion'
Please stand by while rebooting the system...
Restarting system.

%INSTALLER-7-ISSU_OP_SUCC: Rollback timer started with timer value (2700)

cr19-4507-MB#show issu rollback-timer
     Rollback Process State = In progress
   Configured Rollback Time = 00:45:00
    Automatic Rollback Time = 00:43:18
```

Layer 3 neighbors gracefully resynchronize routing information with the new supervisor while maintaining and forwarding traffic across all EtherChannel member links, including on uplink ports of the old active supervisor module.

```
%DUAL-5-NBRCHANGE: EIGRP-IPv4:(100) 100: Neighbor 10.125.0.7 (Port-channel11) is
resync: peer graceful-restart

!Automatically triggers issu commitversion that stops roll-back timer and resets
current standby supervisor module to bootup with new targeted Cisco IOS-XE software

%INSTALLER-7-ISSU_OP_SUCC:  issu changeversion is now executing 'issu commitversion'
%HA_CONFIG_SYNC-6-BULK_CFGSYNC_SUCCEED: Bulk Sync succeeded
%RF-5-RF_TERMINAL_STATE: Terminal state reached for (SSO)

cr19-4507-MB#show issu rollback-timer
```

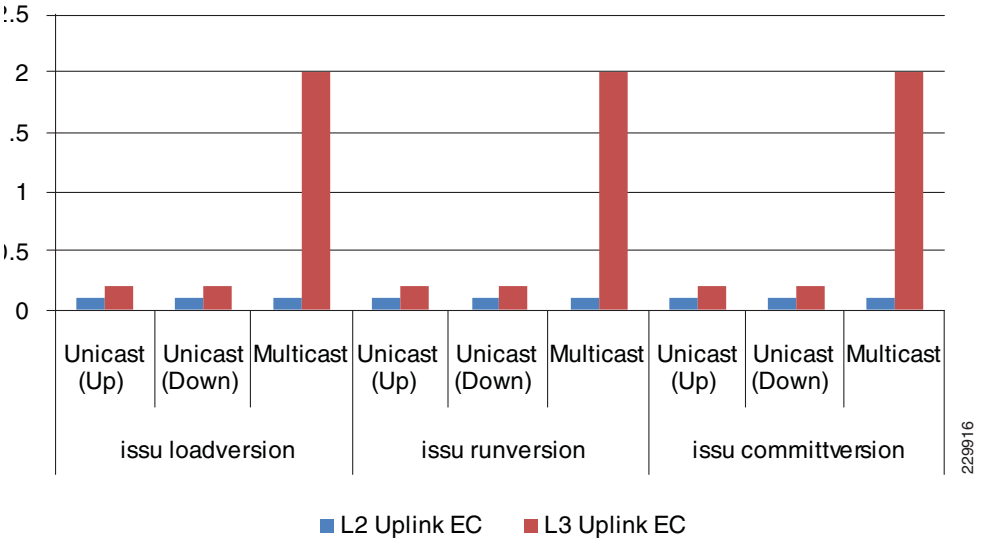
```
      Rollback Process State = Not in progress
    Configured Rollback Time = 00:45:00

cr19-4507-MB#show issu state detail | exc Pre|Post
        Slot = 4
        RP State = Active
        ISSU State = Init
        Operating Mode = Stateful Switchover
        Current Image = bootflash:cat4500e-universalk9.SSA.03.01.01.0.74.150.2.SG.bin

        Slot = 3
        RP State = Standby
        ISSU State = Init
        Operating Mode = Stateful Switchover
        Current Image = bootflash:cat4500e-universalk9.SSA.03.01.01.0.74.150.2.SG.bin
```

Catalyst 4500E Network Recovery with ISSU Software Upgrade

As described in the previous section, the Cisco Catalyst 4500E chassis in redundant supervisor mode gracefully resets the supervisor module without impacting any of its uplink ports. Hence even during the software upgrade procdure, the Cisco Catalyst 4500E chassis maintains its original network capacity and gracefully synchronizes with peer network devices for continuous forwarding of network traffic. This highly resilient architecture provides the network administrator with the flexibility to upgrade the Catalyst 4500E chassis with new Cisco IOS software without downtime or disruption in the operation of the network.

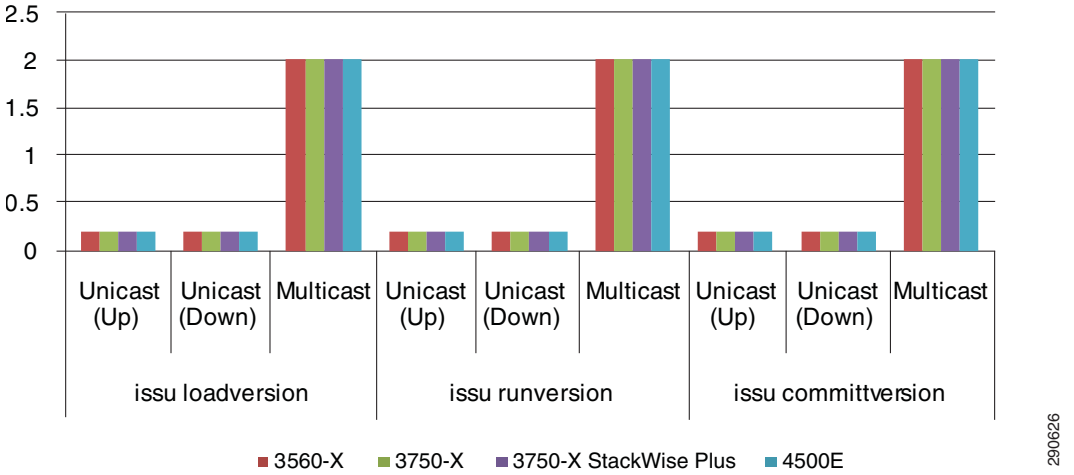
The ISSU software upgrade procedure is even more graceful and transparent with EtherChannel-based network topologies, which offer entire system upgrades with deterministic traffic loss information. The following two charts provide characterized ISSU test results during a supervisor reset that is triggered by **issu loadversion**, **issu runversion**, and i**ssu committversion** via a manual CLI or an automatic ISSU upgrade procedure on the Cisco Catalyst 4500E. These results are collected from a Catalyst 4500E chassis deployed in the campus access layer and at the distribution layer deployed with Layer 2 and Layer 3 EtherChannel:

Figure 32 Catalyst 4500E Access Layer Network Recovery with ISSU Software Upgrade



The Catalyst 4500E can be deployed in Layer 2 or Layer 3 mode at the campus access layer. During each ISSU software upgrade step, the network impact to the unicast or multicast traffic flow is at or below 200 msec range when the Catalyst 4500E system is deployed in Layer 2 mode. However when the routing boundary is extended to the access layer, the current Catalyst 4500E chassis does not fully support Layer 3 multicast high-availability. This means that the multicast traffic loss can be higher than the unicast flows.

Figure 33 Catalyst 4500E Distribution Layer Network Recovery with ISSU Software Upgrade



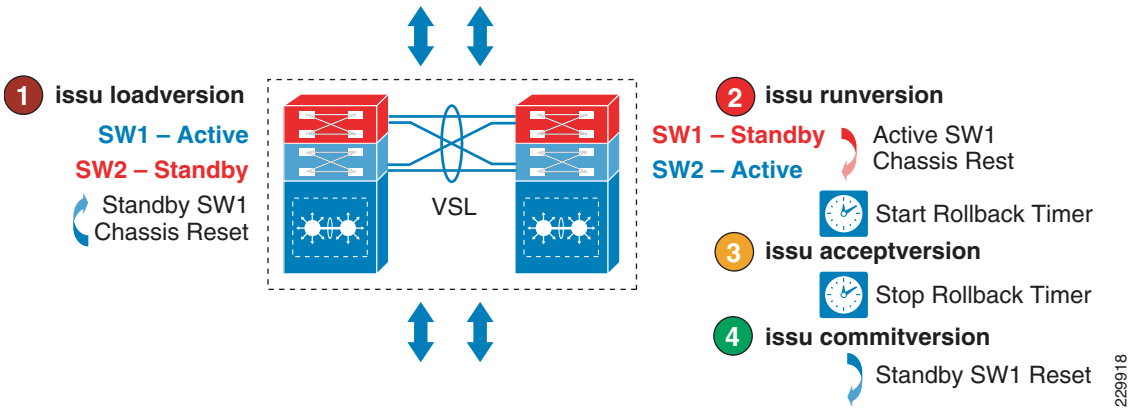
The Catalyst 4507R+E deployed in the distribution layer typically represents a demarcation point between Layer 2 and Layer 3 network boundaries. As described in the previous section, the current software architecture of the Catalyst 4507R+E series platform does not support Layer 3 multicast high-availability. Thus the multicast PIM neighbor adjacency and forwarding information are reset during the supervisor switchover process. This reset causes about a two second multicast traffic loss, but with consistent unicast traffic loss at or below 200 msec baseline range.

Catalyst 6500-E VSS eFSU Software Design and Upgrade Process

Cisco Catalyst VSS was introduced in the initial IOS Release 12.2(33)SXH that supported Fast Software Upgrade (FSU). In the initial introduction, it had limited high-availability capabilities to upgrade the IOS software release. The ISSU mismatched software version compatibility was not supported by the FSU infrastructure, which could cause network down time. This may not be a desirable solution when deploying the Catalyst 6500-E in the critical aggregation or core network tier.

Starting with IOS Release 12.2(33)SXI, the Catalyst 6500-E supports true transparent IOS software upgrade in standalone and virtual switch network designs. Enhanced Fast Software Upgrade (eFSU) made it completely ISSU infrastructure compliant and enhances the software and hardware design to retain its functional state during the graceful upgrade process.

Figure 34 Catalyst 6500-E VSS eFSU Software Upgrade Process

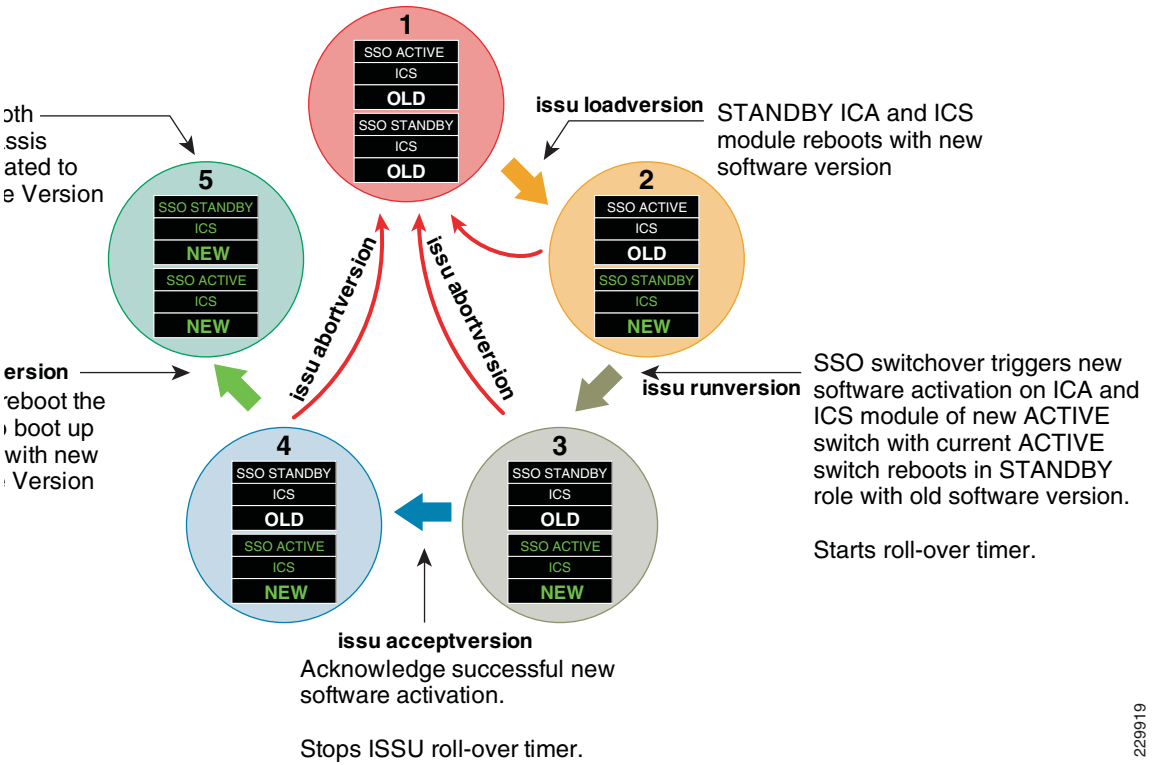


Since eFSU in the Catalyst 6500-E system is built on the ISSU infrastructure, most of the eFSU pre-requisites for Cisco VSS dual-sup design and IOS upgrade procedures remain consistent as explained in a previous sub-section. As described earlier, the Cisco VSS technology enables inter-chassis SSO communication between two virtual switch nodes. However, while the software upgrade procedure for inter-chassis eFSU upgrades is similar, the network operation slightly differs compared to ISSU implemented on intra-chassis based SSO design.

Catalyst 6500-E VSS Quad-Sup eFSU Software Upgrade Process

The eFSU software upgrade process remained simplified even when VSS is deployed with an increased number of supervisor modules in the domain. Upgrading four operational supervisor modules with minimal upgrade time and with reduced complexities, Cisco VSS software is designed to perform parallel ICA and ICS module upgrades by leveraging the existing eFSU cycle. As described earlier, the ICA supervisor module updates the BOOT parameters in ROMMON of the ICS supervisor module in order to to boot up with the new targeted Cisco IOS software. The Cisco VSS quad-sup follows the SSO Z-switchover mechanism as it goes through the entire eFSU upgrade cycle, as illustrated in Figure 1-35.

Figure 1-35 VSS Quad-Sup eFSU Process Cycle



Before going through the eFSU upgrade procedure on Cisco VSS deployed with quad-sup, the network administrator must make sure that the following prerequisites and guidelines for graceful system upgrade are followed:

- To gracefully upgrade all four supervisor modules, Cisco highly recommends that the ICA and ICS supervisor on both virtual switch chassis meet the following requirements:
  - Run common software version and license type.
  - The new IOS software version must be copied to local storage (e.g., disk0, bootdisk) of the ICA and ICS supervisor module.
  - All four supervisor modules are in a fully operational state (SSO ACTIVE/HOT_STANDBY or RPR-WARM mode)
- Do not insert a new ICS supervisor or swap ICS supervisors during any step of the eFSU upgrade procedure.

- During the eFSU software upgrade cycle, intra-chassis role switchover may occur when ISSU triggers a chassis reset. Hence it is strongly recommended to design the network as per the recommendations made in this design document.
- Save the configuration to startup-config, local storage, or at a remote location (e.g.. TFTP/FTP)
- Do not manually modify any BOOT variables and strings.

Catalyst 6500-E eFSU Software Upgrade Procedure

This subsection provides the software upgrade procedure for the Catalyst 6500-Es deployed in VSS mode with quad-sup in the Borderless Campus design. eFSU is supported on the Catalyst 6500-E Sup720-10GE supervisor module running Cisco IOS release with the Enterprise feature set.

In the following sample output, VSS capable Sup720-10G supervisor modules are installed in Slot5 and Slot6 of virtual switch SW1 and SW2, respectively. The virtual switch SW1 supervisor is in the SSO Active role and the SW2 supervisor is in the Standby hot role. In addition, with MEC and the distributed forwarding architecture, the forwarding plane is in an active state on both virtual switch nodes. Both ICA supervisors are running the Cisco IOS Release 12.2(33)SXI4 software version and are fully synchronized with SSO. ICS supervisor modules are running the same Sup720-LC software version and operating in RPR-WARM mode.

```
cr23-VSS-Core#show switch virtual redundancy | inc Mode|Switch|Image|Control
                  My Switch Id = 1
                Peer Switch Id = 2
    Configured Redundancy Mode = sso
     Operating Redundancy Mode = sso
Switch 1 Slot 5 Processor Information :
                  Image Version = Cisco IOS Software, s72033_rp Software
(s72033_rp-ADVENTERPRISEK9_WAN-M), Version 12.2(33)SXI4, RELEASE SOFTWARE (fc3)
           Control Plane State =
Switch 1 Slot 6 Processor Information :
                  Image Version = Cisco IOS Software, s72033_lc Software (s72033_lc-SP-M),
Version 12.2(33)SXI4, RELEASE SOFTWARE (fc3)
           Control Plane State = RPR-Warm
Switch 2 Slot 5 Processor Information :
                  Image Version = Cisco IOS Software, s72033_rp Software
(s72033_rp-ADVENTERPRISEK9_WAN-M), Version 12.2(33)SXI4, RELEASE SOFTWARE (fc3)
           Control Plane State = STANDBY
Switch 2 Slot 6 Processor Information :
                  Image Version = Cisco IOS Software, s72033_lc Software (s72033_lc-SP-M),
Version 12.2(33)SXI4, RELEASE SOFTWARE (fc3)
           Control Plane State = RPR-Warm
```

The following provides a step-by-step procedure to upgrade from Cisco IOS Release 12.2(33)SXI4 to 12.2(33)SXI4a without causing network topology and forwarding disruption. Each upgrade step can be aborted at any stage by issuing the **issu abortversion** command if any failures are detected.

1. *ISSU loadversion*—This first step will direct the active virtual switch node to initialize the ISSU software upgrade process.

```
cr22-6500-LB#show issu state
                           Slot = 1/5
                     RP State =
                    ISSU State = Init
                  Boot Variable =
bootdisk:s72033-adventerprisek9_wan-mz.122-33.SXI4.bin,12;
                           Slot = 2/5
                     RP State = Standby
                    ISSU State = Init
                  Boot Variable =
bootdisk:s72033-adventerprisek9_wan-mz.122-33.SXI4.bin,12;

cr23-VSS-Core#issu loadversion 1/5 disk0:s72033-adventerprisek9_wan-mz.122-33.SXI4a
2/4 slavedisk0: s72033-adventerprisek9_wan-mz.122-33.SXI4a
```

After issuing the above command, the active virtual switch ensures the new IOS software is downloaded on both supervisors' file systems and performs several additional checks on the ICA and ICS supervisor modules on the remote virtual switch for the graceful software upgrade process. ISSU changes the boot variable to the new IOS software version if no error is found and resets the standby virtual switch and installed modules.

```
%RF-SW1_SP-5-RF_RELOAD: Peer reload. Reason: ISSU Loadversion
%SYS-SW2_SPSTBY-5-RELOAD: Reload requested - From Active Switch (Reload peer unit).
%issu loadversion executed successfully, Standby is being reloaded
```

Note Resetting the standby virtual switch node will not trigger the network protocol graceful recovery process and will not reset the ICS supervisor module or the linecards on the active virtual switch. It will remain in an operational and forwarding state for the transparent upgrade process.

With the broad range of ISSU version compatibility for SSO communication, the standby supervisor will successfully bootup again in its original standby state:

```
cr23-VSS-Core#show switch virtual redundancy | inc Mode|Switch|Image|Control
                  My Switch Id = 1
                Peer Switch Id = 2
    Configured Redundancy Mode = sso
     Operating Redundancy Mode = sso
Switch 1 Slot 5 Processor Information :
```

```
                Image Version = Cisco IOS Software, s72033_rp Software
(s72033_rp-ADVENTERPRISEK9_WAN-M), Version 12.2(33)SXI4, RELEASE SOFTWARE (fc3)
           Control Plane State =
Switch 1 Slot 6 Processor Information :
                Image Version = Cisco IOS Software, s72033_lc Software
(s72033_lc-SP-M), Version 12.2(33)SXI4, RELEASE SOFTWARE (fc3)
           Control Plane State = RPR-Warm
Switch 2 Slot 5 Processor Information :
                Image Version = Cisco IOS Software, s72033_rp Software
(s72033_rp-ADVENTERPRISEK9_WAN-M), Version 12.2(33)SXI4a, RELEASE SOFTWARE (fc2)
           Control Plane State = STANDBY
Switch 2 Slot 6 Processor Information :
                Image Version = Cisco IOS Software, s72033_lc Software
(s72033_lc-SP-M), Version 12.2(33)SXI4a, RELEASE SOFTWARE (fc2)
           Control Plane State = RPR-Warm
```

To rejoin the virtual switch domain, both chassis will reestablish distributed VSL EtherChannel communication and force the active supervisor to resynchronize all SSO redundancy and checkpoints, VLAN database, and forwarding information with the standby virtual switch. The network administrator is notified to proceed with the next ISSU step.

```
%HA_CONFIG_SYNC-6-BULK_CFGSYNC_SUCCEED: Bulk Sync succeeded
%PFREDUN-SW2_SPSTBY-6-STANDBY: Ready for SSO mode

%ISSU_PROCESS-SW1_SP-7-DEBUG: Peer state is [ STANDBY HOT ]; Please issue the
runversion command
```

2. *ISSU runversion*—After performing several steps to ensure the new loaded software is stable on the standby virtual switch, the network administrator is now ready to proceed to the runversion step.

```
cr23-VSS-Core#show issu state
                        Slot = 1/5
                   RP State =
                  ISSU State = Load Version
                Boot Variable =
disk0:s72033-adventerprisek9_wan-mz.122-33.SXI4.bin,12

                        Slot = 2/5
                   RP State = Standby
                  ISSU State = Load Version
                Boot Variable =
disk0:s72033-adventerprisek9_wan-mz.122-33.SXI4a,12;disk0:s72033-adventerprisek9_wan-m
z.122-33.SXI4.bin,12

cr23-VSS-Core#issu runversion 2/5
This command will reload the Active unit.  Proceed ? [confirm]y
%issu runversion initiated successfully
```

```
%RF-SW1_SP-5-RF_RELOAD: Self reload. Reason: Admin ISSU runversion CLI
```

This step will force the current active virtual switch (SW1) to reset itself along with its ICS supervisor module and the linecards, which triggers network protocol graceful recovery with peer devices. However the linecard and the ICS supervisor module on the current standby virtual switch (SW2) will remain intact and the data plane traffic will continue to be switched during the switchover process. From a network perspective, the effects of the active supervisor resetting during the ISSU runversion step will be no different than the normal switchover procedure (i.e., administration-forced switchover or supervisor online insertion and removal). In the entire eFSU software upgrade procedure, this is the only time that the systems will perform an SSO-based network graceful recovery. The following sample syslogs confirm stable and EIGRP graceful recovery on the virtual-switch running the new Cisco IOS software version.

NSF-Aware Distribution

```
cr24-4507e-MB#
%DUAL-5-NBRCHANGE: EIGRP-IPv4:(100) 100: Neighbor 10.125.0.14 (Port-channel1) is
resync: peer graceful-restart
```

After re-negotiating and establishing the VSL EtherChannel link and going through the VSLP protocol negotiation process, the rebooted virtual switch module boots up in the standby role with the older IOS software version instead of the new IOS software version. The ISSU runversion procedure starts the SSO Z-Switchover process in the VSS quad-sup design where the ICS can take over the ICA ownership during the next boot up process. Hence all the network configuration and the VSL connectivity must be deployed as recommended in this document for transparent network operation:

```
Dist-VSS#show switch virtual redundancy | inc Mode|Switch|Image|Control
                  My Switch Id = 2
                Peer Switch Id = 1
    Configured Redundancy Mode = sso
     Operating Redundancy Mode = sso
Switch 2 Slot 5 Processor Information :
Image Version = Cisco IOS Software, s72033_rp Software
(s72033_rp-ADVENTERPRISEK9_WAN-M), Version 12.2(33)SXI4a, RELEASE SOFTWARE (fc2)
           Control Plane State =
Switch 2 Slot 6 Processor Information :
Image Version = Cisco IOS Software, s72033_lc Software (s72033_lc-SP-M), Version
12.2(33)SXI4a, RELEASE SOFTWARE (fc2)
           Control Plane State = RPR-Warm
Switch 1 Slot 6 Processor Information :
Image Version = Cisco IOS Software, s72033_rp Software
(s72033_rp-ADVENTERPRISEK9_WAN-M), Version 12.2(33)SXI4, RELEASE SOFTWARE (fc3)
           Control Plane State = STANDBY
Switch 1 Slot 5 Processor Information :
Image Version = Cisco IOS Software, s72033_lc Software (s72033_lc-SP-M), Version
12.2(33)SXI4, RELEASE SOFTWARE (fc3)
```

```
            Control Plane State = RPR-Warm
```

Like intra-chassis ISSU implementation, eFSU also provides a safeguarded software design for additional network stability and the opportunity to roll back to the previous IOS software if the system upgrade causes any type of network abnormality. At this stage, ISSU automatically starts a set of internal rollback timers to re-install the old IOS image if there are any problems. The default rollback timer is up to 45 minutes, which provides the network administrator with an opportunity to perform several sanity checks. In small- to mid-size network designs, the default timer may be sufficient. However for large networks, the network administrator may want to adjust the timer up to two hours:

```
%ISSU_PROCESS-SP-7-DEBUG: rollback timer process has been started
cr23-VSS-Core#show issu rollback-timer
        Rollback Process State = In progress
      Configured Rollback Time = 00:45:00
       Automatic Rollback Time = 00:36:08
```

The system will notify the network administrator with the following syslog to continue to the next ISSU upgrade step if no stability issues occur and all the network services are operating as expected.

```
%ISSU_PROCESS-SW2_SP-7-DEBUG: Peer state is [ STANDBY HOT ]; Please issue the
acceptversion command
```

3. *ISSU acceptversion*—This eFSU step provides confirmation from the network administrator regarding the system and network stability after installing the new software. It confirms readiness to accept the new IOS software on the standby supervisor. This step stops the rollback timer and instructs the network administrator to continue to the final commit state. However, it does not perform any additional steps to install the new software on the standby supervisor.

```
cr23-VSS-Core#show issu state
                          Slot = 2/5
                     RP State =
                    ISSU State = Run Version
                  Boot Variable =
disk0:s72033-adventerprisek9_wan-mz.122-33.SXI4a,12;disk0:s72033-adventerprisek9_wan-m
z.122-33.SXI4.bin,12

                          Slot = 1/6
                     RP State = Standby
                    ISSU State = Run Version
                  Boot Variable =
disk0:s72033-adventerprisek9_wan-mz.122-33.SXI4.bin,12


cr23-VSS-Core#issu acceptversion 2/5
% Rollback timer stopped. Please issue the commitversion command.
cr23-VSS-Core#show issu rollback-timer
```

```
Rollback Process State = Not in progress
Configured Rollback Time = 00:45:00
```

4. *ISSU commitversion*—The final eFSU step forces the active virtual switch to synchronize the configuration with the standby supervisor and force it to reboot with the new IOS software. This stage concludes the eFSU upgrade procedure and the new IOS version is permanently committed on the ICA and ICS supervisor modules of both virtual switches. If for some reason the network administrator needs to rollback to the older image, it is recommended to perform the eFSU-based downgrade procedure to maintain the network operational state without any downtime.

```
cr23-VSS-Core#issu commitversion 1/6
Building configuration...
[OK]
%RF-SW2_SP-5-RF_RELOAD: Peer reload. Reason: Proxy request to reload peer
%SYS-SW1_SPSTBY-5-RELOAD: Reload requested - From Active Switch (Reload peer unit).
%issu commitversion executed successfully

cr23-VSS-Core#show switch virtual redundancy | inc Mode|Switch|Image|Control
                  My Switch Id = 2
                Peer Switch Id = 1
    Configured Redundancy Mode = sso
     Operating Redundancy Mode = sso
Switch 2 Slot 5 Processor Information :
                  Image Version = Cisco IOS Software, s72033_rp Software
(s72033_rp-ADVENTERPRISEK9_WAN-M), Version 12.2(33)SXI4a, RELEASE SOFTWARE (fc2)
           Control Plane State =
Switch 2 Slot 6 Processor Information :
                  Image Version = Cisco IOS Software, s72033_lc Software
(s72033_lc-SP-M), Version 12.2(33)SXI4a, RELEASE SOFTWARE (fc2)
           Control Plane State = RPR-Warm
Switch 1 Slot 5 Processor Information :
                  Image Version = Cisco IOS Software, s72033_rp Software
(s72033_rp-ADVENTERPRISEK9_WAN-M), Version 12.2(33)SXI4a, RELEASE SOFTWARE (fc2)
           Control Plane State = STANDBY
Switch 1 Slot 6 Processor Information :
                  Image Version = Cisco IOS Software, s72033_lc Software
(s72033_lc-SP-M), Version 12.2(33)SXI4a, RELEASE SOFTWARE (fc2)
           Control Plane State = RPR-Warm
```

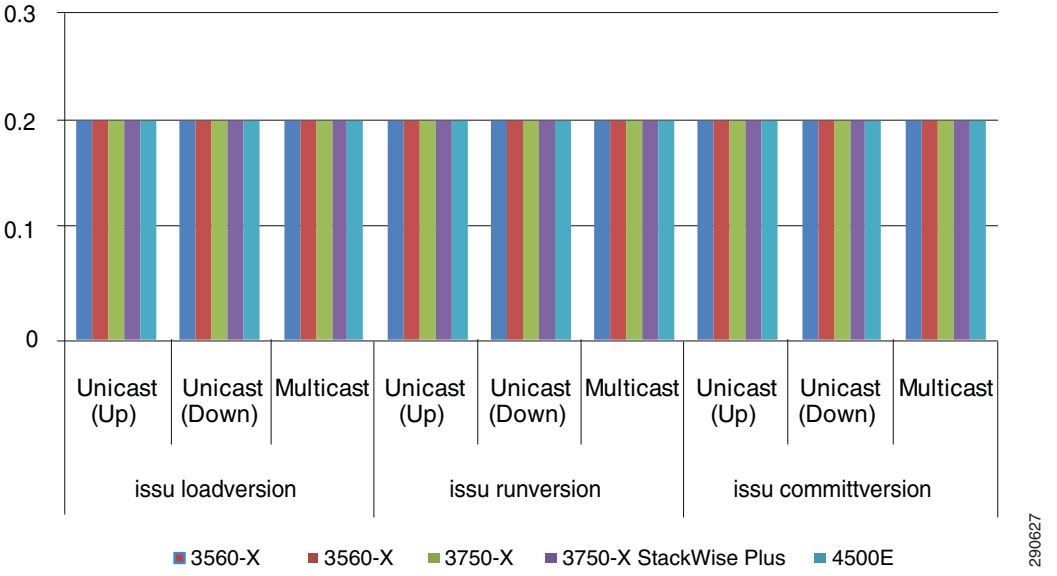
Catalyst 6500-E Network Recovery with eFSU Software Upgrade

As described in a previous section, the Cisco Catalyst 6500-E chassis in virtual switch mode will be reset to gracefully install the new Cisco IOS software version on all supervisor modules. When the campus network is deployed in an environment consistent with the best practices recommended in this design guide, network recovery is transparent to end users and applications, as the other chassis in the virtual switch domain continues to provide constant network availability. Bundling the parallel

physical paths simplifies topology and forwarding paths. Thus, even during virtual switch chassis reset, there is no major topology re-computation or re-programming required on the Cisco VSS or on the peer multihomed network devices.

The eFSU software upgrade procedure becomes more graceful and transparent in Cisco VSS and the MEC-based network topologies, which offer the ability to upgrade the system within a consistent window of time. Figure 36 illustrates the amount of traffic loss during the **issu loadversion**, **issu runversion**, and **issu commitversion** process.

Figure 36 Network Recovery with eFSU Software Upgrade



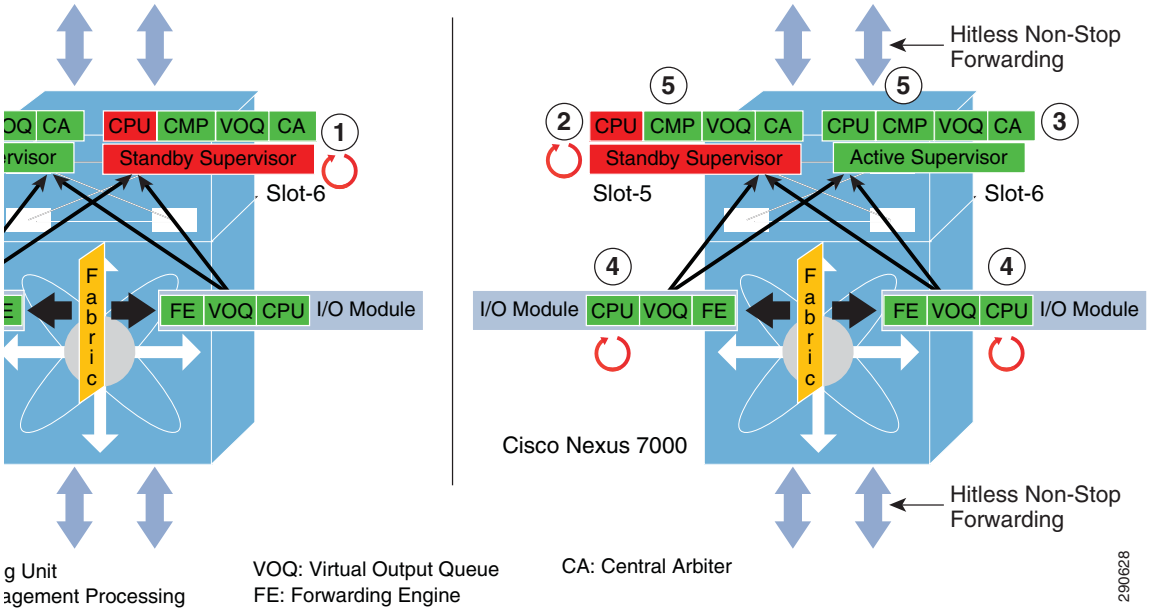
Nexus 7000 ISSU Software Design and Upgrade Process

The ISSU software upgrade process on the Cisco Nexus 7000 running the NX-OS operating system is different than on IOS. Each Nexus 7000 system component is intelligent in design with local hardware and software resources, such as CPU, memory for internal communication, and synchronization. For a common distributed software capability across all distributed hardware, the Nexus 7000 installs a consistent and common software version across the system. The NX-OS software image is bundled and compressed with the system software image, the linecard image to install on each I/O module, and the Connectivity Management Processor (CMP) BIOS and software image. The NX-OS ISSU software provides an option to manually upgrade each component separately or all at once.

The ISSU NX-OS software upgrade process in the Nexus 7000 system offers many benefits:

- Simplified—The network administrator can upgrade redundant supervisor modules, CMP, and all installed I/O modules with a single command or manually upgrade each component incrementally.
- Automated—Once initiated by the user, the Nexus 7000 system goes through five automatic NX-OS upgrade steps without user involvement. Each upgrade step ensures that the preceding step was successful and automatically proceeds through the following steps to upgrade all system components.
- Reliable—Assures the user that upgrading the NX-OS software will be non-disruptive. The Nexus 7000 performs hardware and software compatibility, integrity, and capability checks to prevent an upgrade failure. The system generates an impact report that provides detailed information about the software upgrade process on each distributed module. The Nexus 7000 aborts the upgrade process and automatically reverts to the previous software version if it detects any upgrade failures.
- Hitless software upgrade—Leveraging the distributed forwarding architecture, NSF/SSO technology, and graceful software upgrade design, the entire five-step ISSU upgrade process in the Nexus 7000 is non-disruptive to borderless services and hitless in switching campus core data traffic without dropping a single packet. The software upgrade procedure on each I/O module is also graceful, while data forwarding remains non-disruptive and the active supervisor and I/O CPU protect and restore distributed software applications and run-time data information.

Figure 1-37 Nexus 7000 Hitless ISSU Upgrade Process



The following steps briefly describe the system and network states during a hitless ISSU software upgrade on a Nexus 7000 system as shown in Figure 1-37:

1. The user initiates the NX-OS software process with a single command. The Nexus 7000 system generates a system impact report to verify if the user wants to accept or abort the upgrade procedure. If the user confirms the upgrade, the system synchronizes the kickstart and system image on the standby supervisor module (slot-6), extracts the image from the bundled new software, and upgrades the bios/bootloader/bootrom on each hardware component. On completion, the standby supervisor module is reset to install the new NX-OS software version.
2. The new boot variable gets updated on the current active supervisor module (slot-5) and forces a self-reset to install the new NX-OS software version.
3. This automatic upgrade step performs two key tasks in the system—SSO role switchover and making the new NX-OS software in effect on the new active supervisor module (slot-6). The Layer 3 network protocols perform graceful recovery with neighbor systems and maintain forwarding plane information for continuous data forwarding. This step remains graceful and non-disruptive with complete hitless switchover.
4. The system initializes a graceful software upgrade on the linecard CPU of each module. The I/O module upgrade occurs incrementally, one module at a time. This software design ensures that the new installed linecard software version is non-disruptive, reliable, and successful prior to proceeding to the next I/O module.

5. The final ISSU upgrade step is upgrading CMP on both supervisor modules. The CMP runs a subset of the operating system that operates independently from the NX-OS running on the supervisor module. The CMP software gets installed if the bundled version is the most recent compared to the currently installed version. This step may not proceed if the current CMP software version is the same or older. Since CMP software operates independently of the system, the version may not be the same as NX-OS.

Nexus 7000 ISSU Software Upgrade Procedure

This subsection provides guidelines to gracefully upgrade software and best practice to consider when upgrading NX-OS on the Cisco Nexus 7000 system.

Prior to starting the NX-OS software upgrade procedure on the Cisco Nexus 7000, the network administrator must make sure that the following prerequisites and guidelines for a graceful system upgrade are followed:

- To gracefully upgrade both supervisor modules, Cisco highly recommends that both Sup-1 supervisor modules in the Nexus 7000 chassis meet the following requirements:
  - Install a common new software version on both supervisor modules.
  - The new NX-OS software version must be copied to local storage (e.g., slot0, bootflash) of the supervisor module.
  - Both supervisor modules are in a a full operational and SSO redundant state.
- Inspect the impact report generated as part of the upgrade process to ensure a non-disruptive and hitless software upgrade process.
- Do not insert, swap, or remove the supervisor, crossbar fabric, or I/O module during the ISSU upgrade procedure.
- Use a direct console or CMP connection and login with network-admin privileges on both the active and standby supervisors during the entire ISSU upgrade process.
- During the ISSU software upgrade cycle, active supervisor role switchover will occur. By default NSF capability is enabled for Layer 3 protocol on the Nexus 7000. It is recommended to ensure that the neighbor system is NSF-aware and supports a compatible NSF protocol capability in the network as per the recommendations in this document.
- Save the system configuration to startup-config, local storage, or in a remote location (e.g., TFTP/FTP).
- Do not manually modify any BOOT variables or strings.

The following subsection demonstrate the entire NX-OS software upgrade procedure and sample output that follows the recommended best practices to install a new software version on the Cisco Nexus 7000 system running version 5.0.5 and upgrading to version 5.1.1a.

Preparing for NX-OS Software Upgrade

Prior to initializing the ISSU software upgrade procedure, the network administrator must prepare the Nexus 7000 system with the proper installation and validation to prevent services disruption and upgrade failures. In the sample output below, the Nexus 7000 system is equipped with dual redundant supervisor modules and M108 I/O network modules. The supervisor module is operating in SSO redundant mode and running system image 5.0(5); the I/O network module is running the same version of the linecard image as the supervisor module.

```
cr35-N7K-Core1#show module
Mod  Ports  Module-Type                        Model              Status
---  -----  --------------------------------- ------------------ ------------
1    8      10 Gbps Ethernet XL Module      N7K-M108X2-12L     ok
2    8      10 Gbps Ethernet XL Module      N7K-M108X2-12L     ok
5    0      Supervisor module-1X             N7K-SUP1                 *
6    0      Supervisor module-1X             N7K-SUP1           ha-standby

Mod  Sw              Hw
---  ------------  ------
1    5.0(5)          1.1
2    5.0(5)          1.1
5    5.0(5)          1.8
6    5.0(5)          1.6

cr35-N7K-Core1#show version | inc "CMP version"
  CMP Image version:       5.0(2) [build 5.0(0.66)]
  CMP Image version:       5.0(2) [build 5.0(0.66)]
```

Copy the new NX-OS system and kickstart software images on the local storage of both supervisor modules. If the new software is copied on compact flash, then it is recommended to not remove or swap until the system is successfully upgraded with the new NX-OS software version.

```
cr35-N7K-Core1# dir bootflash://sup-1 | inc 5.1.1a.bin
  145433028    Mar 03 21:52:15 2011  n7000-s1-dk9.5.1.1a.bin
   30484992    Dec 16 20:02:47 2010  n7000-s1-kickstart.5.1.1a.bin
cr35-N7K-Core1# dir bootflash://sup-2 | inc 5.1.1a.bin
  145433028    Mar 05 09:53:31 2011  n7000-s1-dk9.5.1.1a.bin
   30484992    Dec 16 20:08:23 2010  n7000-s1-kickstart.5.1.1a.bin
```

Cisco recommends generating the upgrade impact report to assess if migrating to new targeted NX-OS software will be graceful and non-disruptive or if it will fail due to any particular hardware or software failure. This pre-determination step performs multiple hardware and software integrity checks on each installed system component. The report indicates which system component will be upgraded from the current software version and how gracefully the new software version becomes effective in the system.

The user must enter the following syntax to generate an impact report. This step does **not** initialize the upgrade process; it performs an internal hardware and software verification procedure with the new software version and generates a detailed impact report:

```
cr35-N7K-Core1#show install all impact system bootflash:/n7000-s1-dk9.5.1.1a.bin kickstart
bootflash:/n7000-s1-kickstart.5.1.1a.bin
!Step 1 - Verify the boot variable for kickstart and system image.
Verifying image bootflash:/n7000-s1-kickstart.5.1.1a.bin for boot variable "kickstart".
[####################] 100% -- SUCCESS
<snip>

!Step 2 - Decompress all bundled software from new version (kickstart, system, cmp,
linecard & bios)
Extracting "lc1n7k" version from image bootflash:/n7000-s1-dk9.5.1.1a.bin.
[####################] 100% -- SUCCESS
<snip>

!Step 3 - Verify the new software compatibility with all installed I/O modules
Performing module support checks.
[####################] 100% -- SUCCESS

!Step 4 - Active supervisor sends new software upgrade signal to each distributed system
components.
Notifying services about system upgrade.
[####################] 100% -- SUCCESS

!Step 5 - Generate new NX-OS system impact report with ISSU upgrade. The first following
table briefly describes if new targeted software will be disruptive or non-disruptive. It
also describes the software installation process - supervisor reset (graceful switchover)
and I/O module rolling (incremental I/O upgrade procedure). The second table describes
which hardware components will be upgraded and provides detail information about current
and new software version information.

Compatibility check is done:
Module  bootable     Impact  I          Install-type  Reason
------  --------  --------------      ------------  ------
     1       yes  non-disruptive        rolling
     2       yes  non-disruptive        rolling
     5       yes  non-disruptive        reset
     6       yes  non-disruptive        reset

Module    Image     Running-Version(pri:alt)                New-Version       Upg-Required
------  ----------  ------------------------                ------------      ------------
1       lc1n7k      5.0(5)                                  5.1(1a)               yes
1        bios       v1.10.14(04/02/10):v1.10.14(04/02/10)   v1.10.14(04/02/10)    no
2       lc1n7k      5.0(5)                                  5.1(1a)               yes
2        bios       v1.10.14(04/02/10):v1.10.14(04/02/10)   v1.10.14(04/02/10)    no
5       system      5.0(5)                                  5.1(1a)               yes
5       kickstart   5.0(5)                                  5.1(1a)               yes
5        bios       v3.22.0(02/20/10):v3.22.0(02/20/10)     v3.22.0(02/20/10      no
5       cmp         5.0(2)                                  5.1(1)                yes
5       cmp-bios    02.01.05                                02.01.05              no
6       system      5.0(5)                                  5.1(1a)               yes
```

```
6       kickstart   5.0(5)                                  5.1(1a)             yes
6        bios       v3.22.0(02/20/10):v3.22.0(02/20/10)     v3.22.0(02/20/10    no
6       cmp         5.0(2)                                  5.1(1)              yes
6       cmp-bios    02.01.05                                02.01.05            no
```

Initiating NX-OS Software Upgrade

After confirming a non-disruptive and graceful upgrade based on the new software report, the network administrator can proceed to initiate the simplified NX-OS upgrade procedure on the Cisco Nexus 7000 system. As described earlier, the ISSU upgrade process on the Nexus 7000 is performed with a single user-initiated command. If the system detects any hardware or software failure during the ISSU upgrade process, the Nexus 7000 automatically aborts the upgrade procedure. Even though the upgrade process is completely automated, it is recommended to be vigilent during the process. Any upgrade failures or abnormal errors must be recorded during the process to ensure the new NX-OS software is installed successfully on each hardware component.

Once the user initiates the upgrade procedure, the Nexus 7000 system will go through five installation steps to upgrade each distributed hardware component, as illustrated in Figure 1-37:

1. Initiate the single-step NX-OS software installation process with the following:

```
cr35-N7K-Core1#install all system bootflash:///n7000-s1-dk9.5.1.1a.bin kickstart
bootflash:///n7000-s1-kickstart.5.1.1a.bin
```

This step repeats the same software initialization, extraction, and cross-system verification process as performed in generating the impact table. Once the report is generated, the network administrator must carefully review the upgrade impact and confirm to proceed with installation or abort.

```
<snip>
6    cmp-bios                                  02.01.05             02.01.05
no
Do you want to continue with the installation (y/n)?  [n] Y
```

The first step starts upgrading the bios, bootloader, and bootrom on each I/O and supervisor module. Once the basic software upgrade completes, the boot variable is modified and the standby supervisor module is force reset to boot with the new NX-OS software image.

```
Install is in progress, please wait.

Module 1: Refreshing compact flash and upgrading bios/loader/bootrom.
Warning: please do not remove or power off the module at this time.
[####################] 100% -- SUCCESS

<snip>

%$ VDC-1 %$ %PLATFORM-2-MOD_REMOVE: Module 6 removed (Serial number JAF1432AERD)
```

```
%$ VDC-1 %$ %CMPPROXY-STANDBY-2-LOG_CMP_WENT_DOWN: Connectivity Management processor
(on module 6) went DOWN


Module 6: Waiting for module online.
 -- SUCCESS        <- Standby supervisor online and successfully upgrade

%CMPPROXY-STANDBY-2-LOG_CMP_UP: Connectivity Management processor(on module 6) is now
UP
```

2. The active supervisor gets ready to gracefully reset. It notifies system components about its active role switchover and resets to boot up with the new NX-OS software. After reboot, it changes its role to standby and re-synchronizes with the new active supervisor module.

```
Notifying services about the switchover.
[####################] 100% -- SUCCESS

%MODULE-5-STANDBY_SUP_OK: Supervisor 5 is standby <- The supervisor is successfully
upgraded and reboots in Standby mode
%CMPPROXY-STANDBY-2-LOG_CMP_UP: Connectivity Management processor(on module 5) is now
UP
```

3. This step triggers two important changes in the Nexus 7000 system:
   - The standby supervisor takes over active ownership and performs graceful protocol recovery with neighbors.
   - The newly-installed NX-OS software becomes effective in the Nexus 7000 system.

```
%MODULE-5-STANDBY_SUP_OK: Supervisor 6 is standby   <- Pre-Switchover slot-6
Supervisor role

  %SYSMGR-2-HASWITCHOVER_PRE_START: This supervisor is becoming active (pre-start
phase). <- Slot-5 supervisor got reset and slot-6 supervisor taking over active role.
  %SYSMGR-2-HASWITCHOVER_START: Supervisor 6 is becoming active.
  %SYSMGR-2-SWITCHOVER_OVER: Switchover completed. <- SSO Switchover successfully
complete

  IP-EIGRP(0) 100: Neighbor 10.125.21.1 (port-channel100) is up: new adjacency <-
Gracefully EIGRP adjacency recovered

!Following graceful EIGRP adjacency synch message from EIGRP neighbor system
IP-EIGRP(0) 100: Neighbor 10.125.21.0 (port-channel100) is resync: peer
graceful-restart
```

4. After upgrading and activating the system image on both supervisor modules, in the next step the Nexus 7000 automatically initializes the I/O module upgrade process. Each I/O modules runs a linecard image on its local CPU for various different types of software applications. For consistent internal system communication, each I/O module must be upgraded to the same software version

as the system image running on the supervisor. All internal and external state machines and dynamic forwarding information remains intact on the distributed forwarding engine and other components of the I/O modules to provide non-stop communication. While the linecard CPU is gracefully installing and resetting itself to make the new software version effective, the I/O module remains fully operational and in service to provide a hitless software upgrade.

```
 Module 1: Non-disruptive upgrading.
  -- SUCCESS   <- The CPU on I/O Module in slot-1 successfully upgraded
 Module 2: Non-disruptive upgrading.
  -- SUCCESS   <- The CPU on I/O Module in slot-2 successfully upgraded
```

5. The final automated upgrade step is the CMP software upgrade process. The Nexus 7000 automatically upgrades CMP complexes on each supervisor module. This step may become optional if CMP is running a software version that is more current than the bundle version. The new CMP software version does not become effective after installation. It becomes effective on the next supervisor or system reset. The user may manually reboot the CMP complex to immediately put the new software into effect.

```
Module 6: Upgrading CMP image.
 Warning: please do not reload or power cycle CMP module at this time.
  -- SUCCESS
 Module 5: Upgrading CMP image.
 Warning: please do not reload or power cycle CMP module at this time.
  -- SUCCESS

 Recommended action::
 "Please reload CMP(s) manually to have it run in the newer version." <- Reload CMP
manually to immediately run new software version.

 Install has been successful. <- ISSU Software Upgrade finish

!Reload CMP complex on both supervisor module
cr35-N7K-Core1# reload cmp module 5
This command will reload the CMP on the supervisor in slot 5.  Continue (y/n)?  [no] Y

cr35-N7K-Core1# reload cmp module 6
This command will reload the CMP on the supervisor in slot 6.  Continue (y/n)?  [no] Y

!Verify new upgrades software status in Nexus 7000 system
cr35-N7K-Core1# show version | inc version
  BIOS:      version 3.22.0
  kickstart: version 5.1(1a)
  system:    version 5.1(1a)
  System version: 5.0(5)
  CMP BIOS version:        02.01.05
  CMP Image version:       5.1(1) [build 5.0(0.66)]
  CMP BIOS version:        02.01.05
  CMP Image version:       5.1(1) [build 5.0(0.66)]
```

Nexus 7000 Network Recovery with ISSU Software Upgrade

The distributed forwarding design and resilient software architecture of the Cisco Nexus 7000 system provides a hitless upgrade, thereby reducing the need for a maintenance window to install new software versions. In a hitless software upgrade design, campus backbone availability and switching capacity remain non-disruptive and intact while the Nexus 7000 is rolled out with a new NX-OS software image. The Cisco Nexus 7000 system had zero packet loss in several successful software upgrades in various network designs. Upgrading the Cisco Nexus 7000 system with the recommended procedure and best practices helps ensure a successful software upgrade and minimizes impact to network services.

10 Summary

Cisco Borderless Networks, a Cisco next-generation architecture, delivers a new workspace experience, securely, reliably, and smoothly connecting anyone, anywhere, using any device, to any resource. This borderless experience is only possible with a strong and resilient intelligent network that is designed to meet the needs of a global workspace. The Cisco-enabled network platform is the primary component of this architecture, providing borderless services such as mobility, security, medianet, location, and EnergyWise, to deliver an optimal user experience. Building network designs with intelligence at the edge provides mobility and secure collaboration, as well as the overall infrastructure backbone to provide network-wide differentiated services for a consistent, highly-available, and reliable user experience. Cisco Borderless Networks enable innovative business models, creating new user experiences that lead to increased customer satisfaction and loyalty.