Cisco Expo 2008

Catalyst 6500 VSS Design and Implementation

Zaklanovic Ivan

izaklano@cisco.com

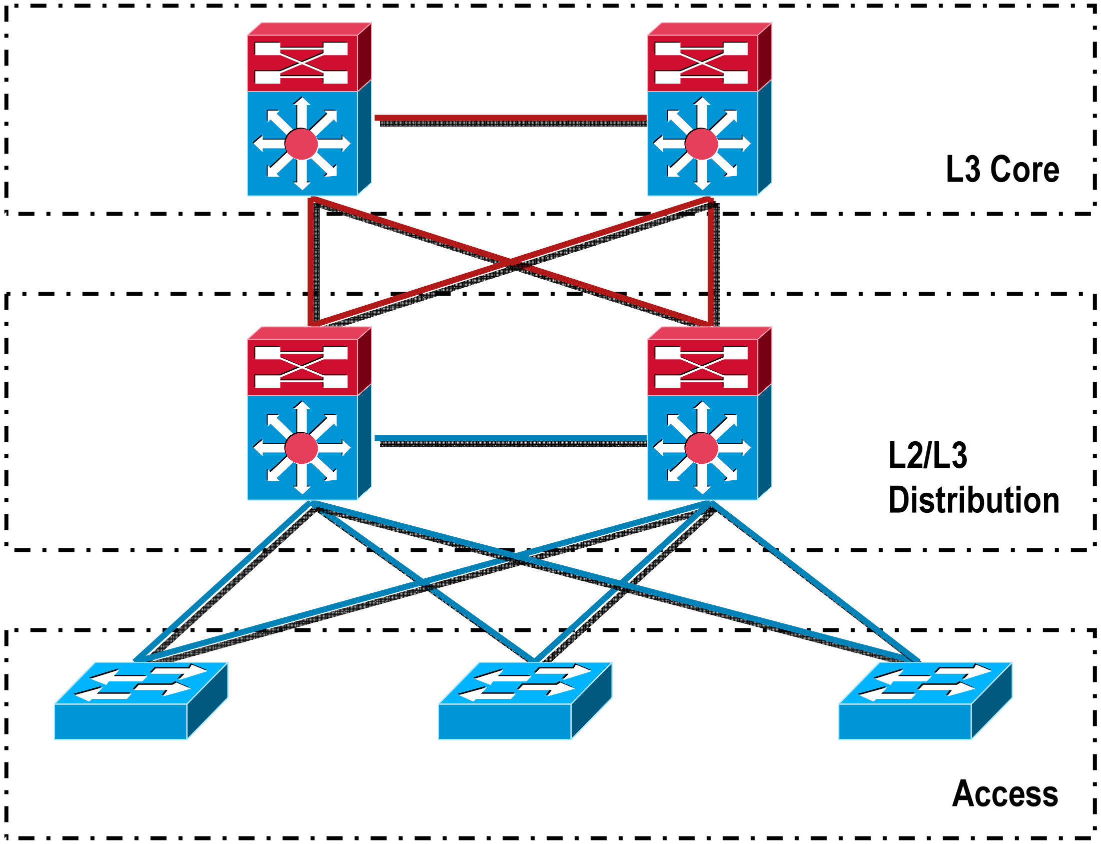
Agenda
Virtual Switching System

- Introduction
- Architecture
- Conversion Process
- High Availability
- Hardware Requirements
- Traffic Flow and Topology Considerations
- Operational Management and ISSU
- Summary

VSS Introduction

Current Network Challenges
Enterprise Campus

Traditional Enterprise Campus deployments have been designed in such a way that allows for scalability, differentiated services and high availability. However they also face many challenges, some of which are listed in the below diagram…



Extensive routing topology, Routing reconvergence

FHRP, STP, Asymmetric routing, Policy Management

Single active uplink per VLAN (PVST), L2 reconvergence

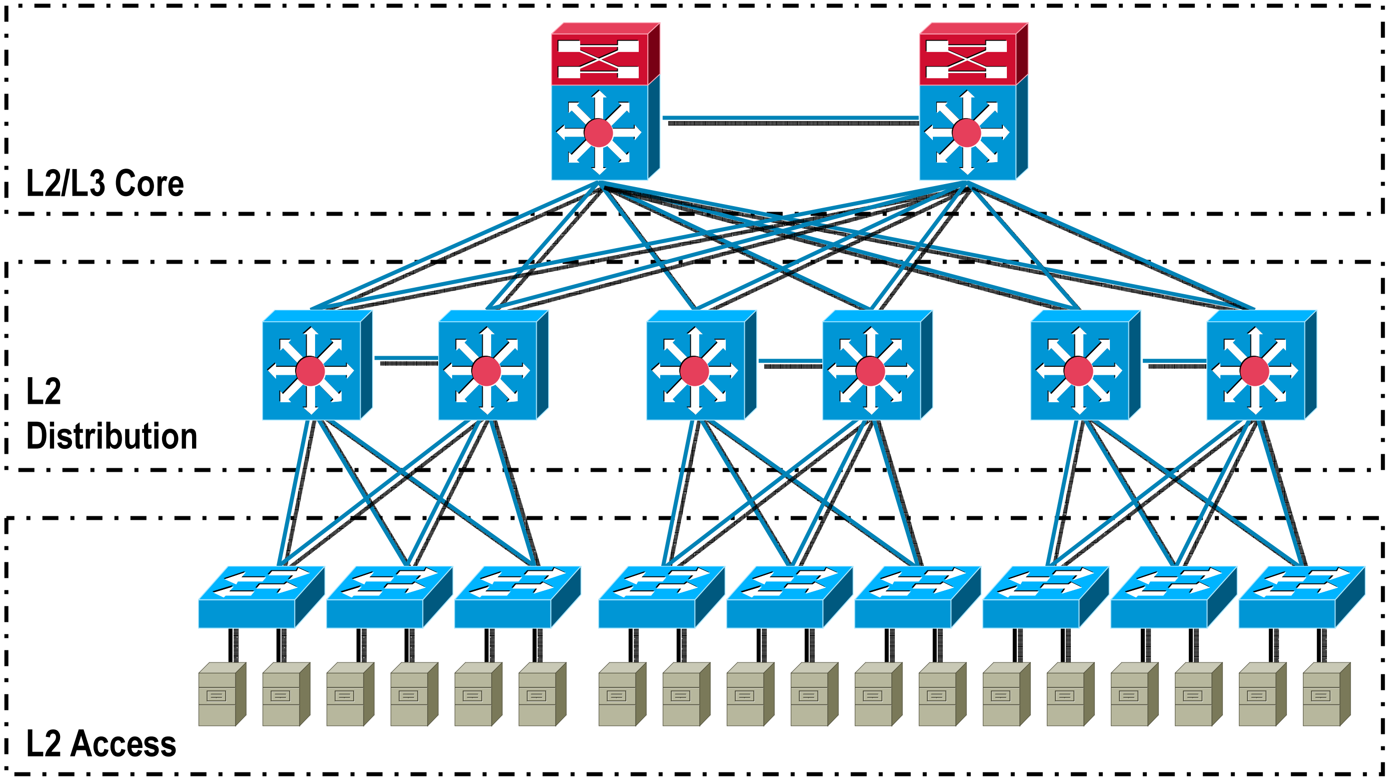
Current Network Challenges
Data Center

Traditional Data Center designs are requiring ever increasing Layer 2 adjacencies between Server nodes due to prevalence of Virtualization technology. However, they are pushing the limits of Layer 2 networks, placing more burden on loop-detection protocols such as Spanning Tree...

FHRP, HSRP, VRRP
Spanning Tree
Policy Management

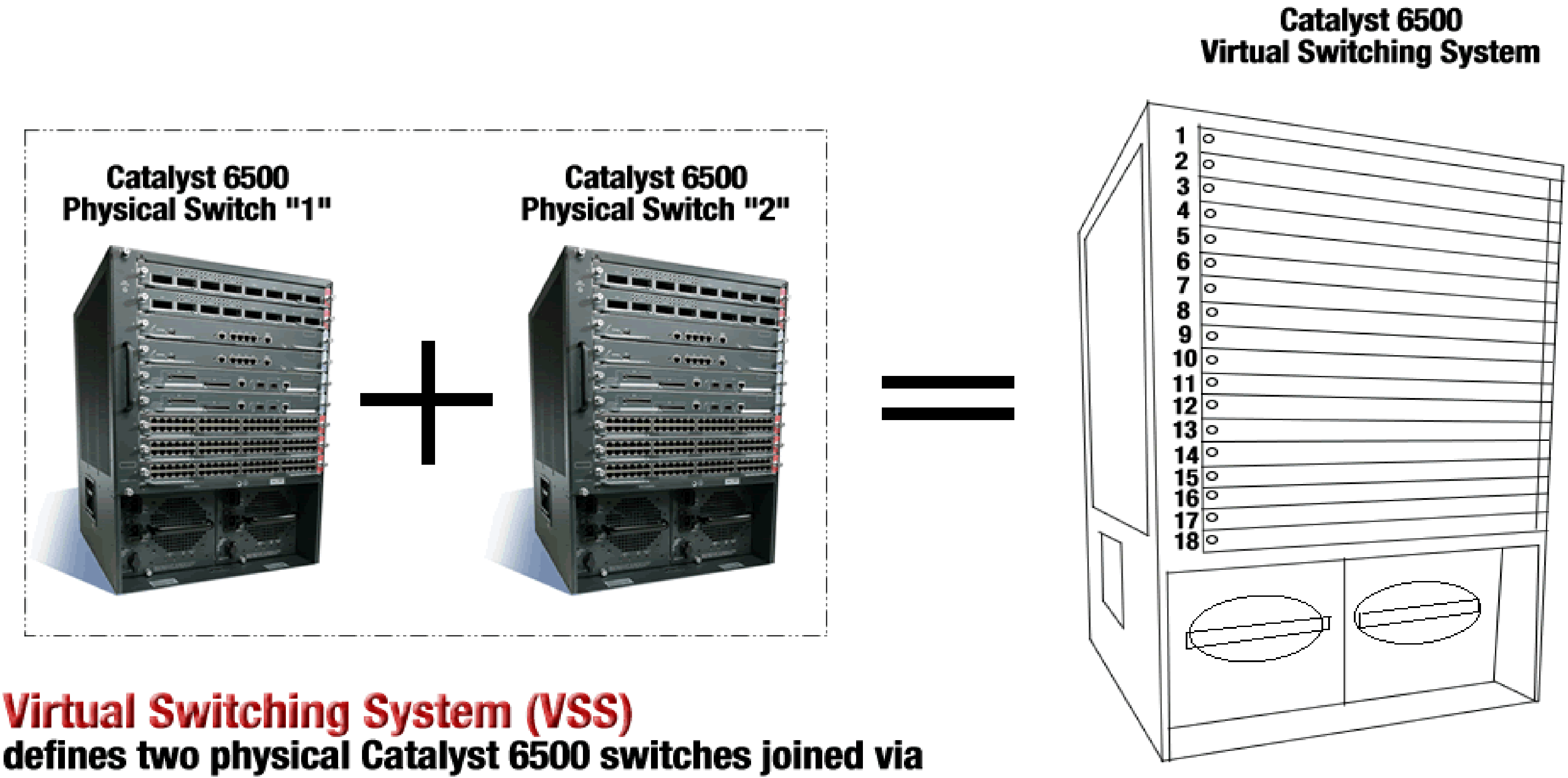
Single active uplink per VLAN (PVST), L2 reconvergence, excessive BPDUs

Dual-Homed Servers to single switch, Single active uplink per VLAN (PVST), L2 reconvergence



Virtual Switching System

Virtual Switching System System is a new technology break through for the Catalyst 6500 family…

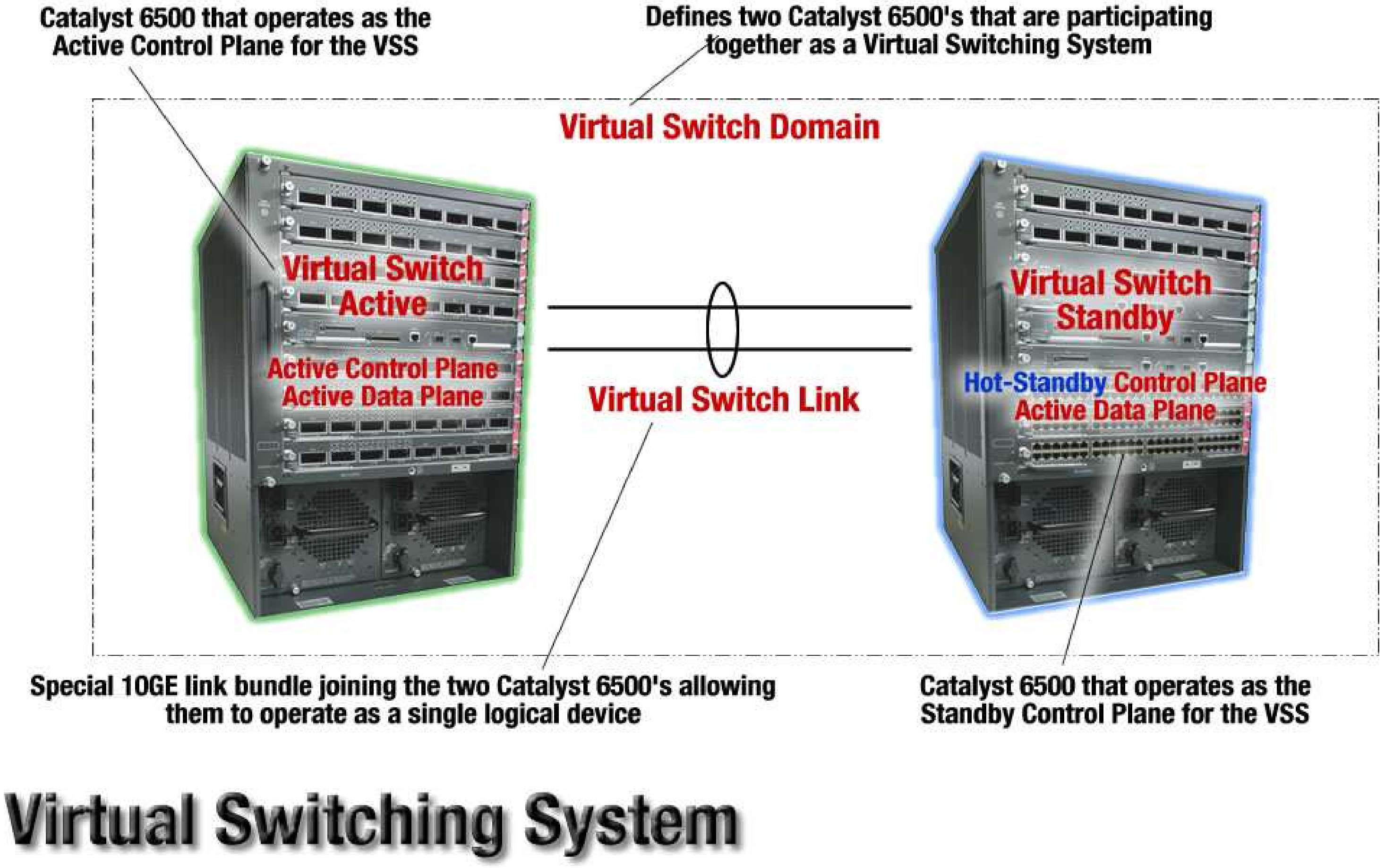


Virtual Switching System (VSS)
defines two physical Catalyst 6500 switches joined via a special link called a Virtual Switch Link (VSL) running special hardware and software that allows them to operate as a single logical switch

Introduction to Virtual Switching System
Concepts

Uses one supervisor in each chassis with inter-chassis Stateful Switchover (SSO) method in with one supervisor is ACTIVE and other in HOT_STANDBY mode

Active/standby supervisors run in synchronized mode (boot-env, running-configuration, protocol state, and line cards status gets synchronized)



Catalyst 6500 that operates as the Active Control Plane for the VSS

Defines two Catalyst 6500's that are participating together as a Virtual Switching System

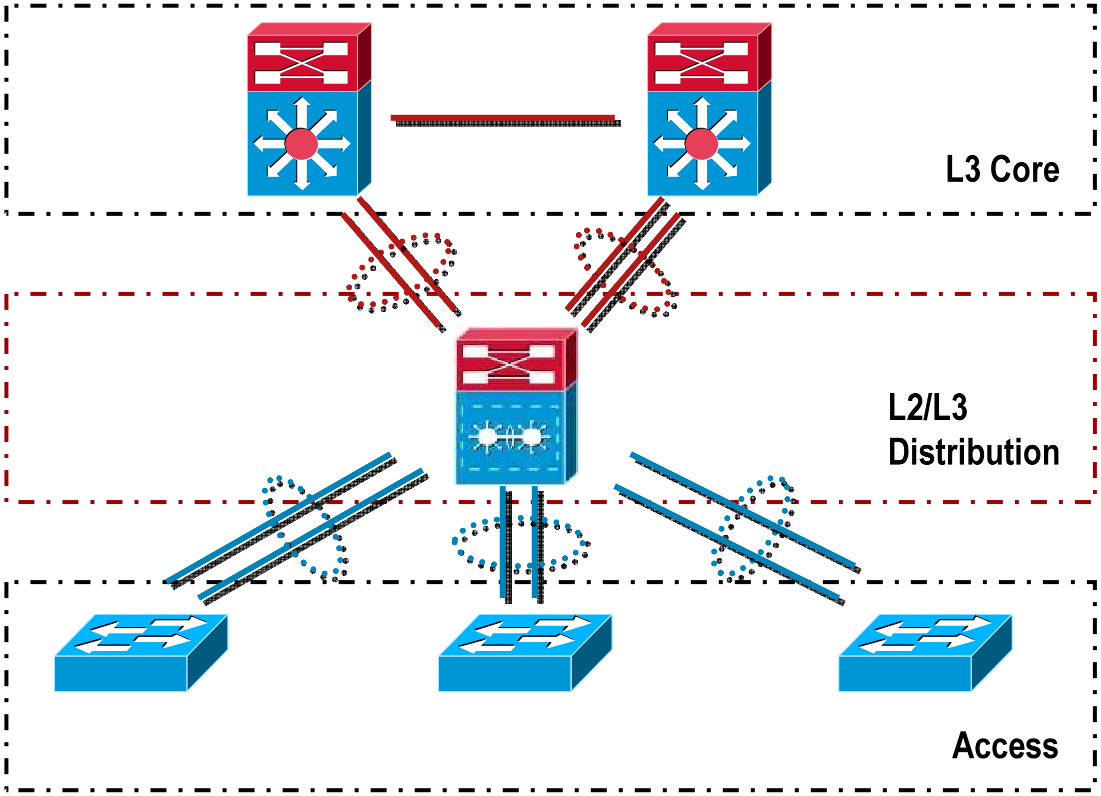
Special 10GE link bundle joining the two Catalyst 6500's allowing them to operate as a single logical device

Catalyst 6500 that operates as the Standby Control Plane for the VSS

Virtual Switching System System

Enterprise Campus

A Virtual Switching System-enabled Enterprise Campus network takes on multiple benefits including simplified management & administration, facilitating greater high availability, while maintaining a flexible and scalable architecture…



Reduced routing neighbors, Minimal L3 reconvergence

No FHRPs
No Looped topology
Policy Management

Multiple active uplinks per VLAN, No STP convergence

Virtual Switching System System

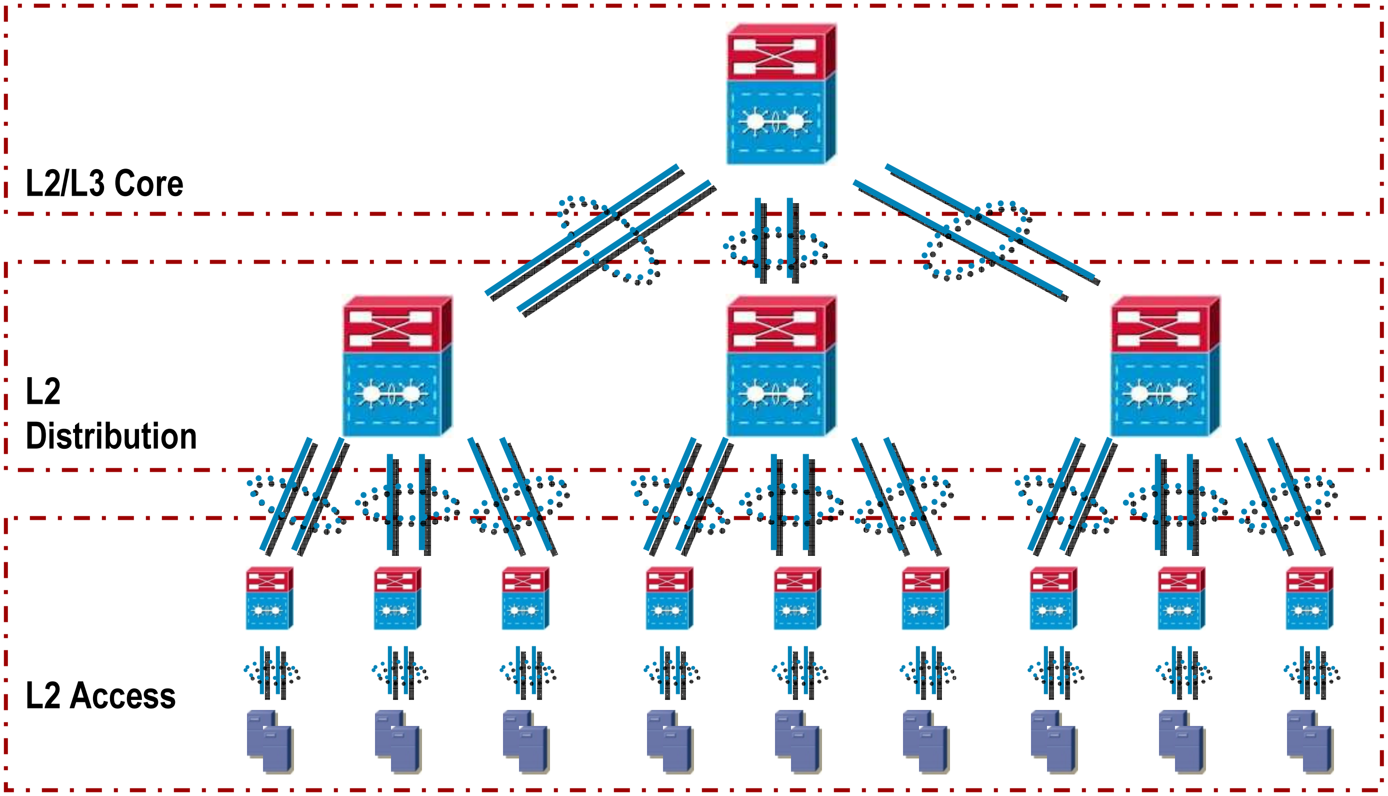
Data Center

A Virtual Switching System-enabled Data Center allows for maximum scalability so bandwidth can be added when required, but still providing a larger Layer 2 hierarchical architecture free of reliance on Spanning Tree…

Single router node, Fast L2 convergence, Scalable architecture

Dual Active Uplinks, Fast L2 convergence, minimized L2 Control Plane, Scalable

Dual-Homed Servers, Single active uplink per VLAN (PVST), Fast L2 convergence



VSS Architecture

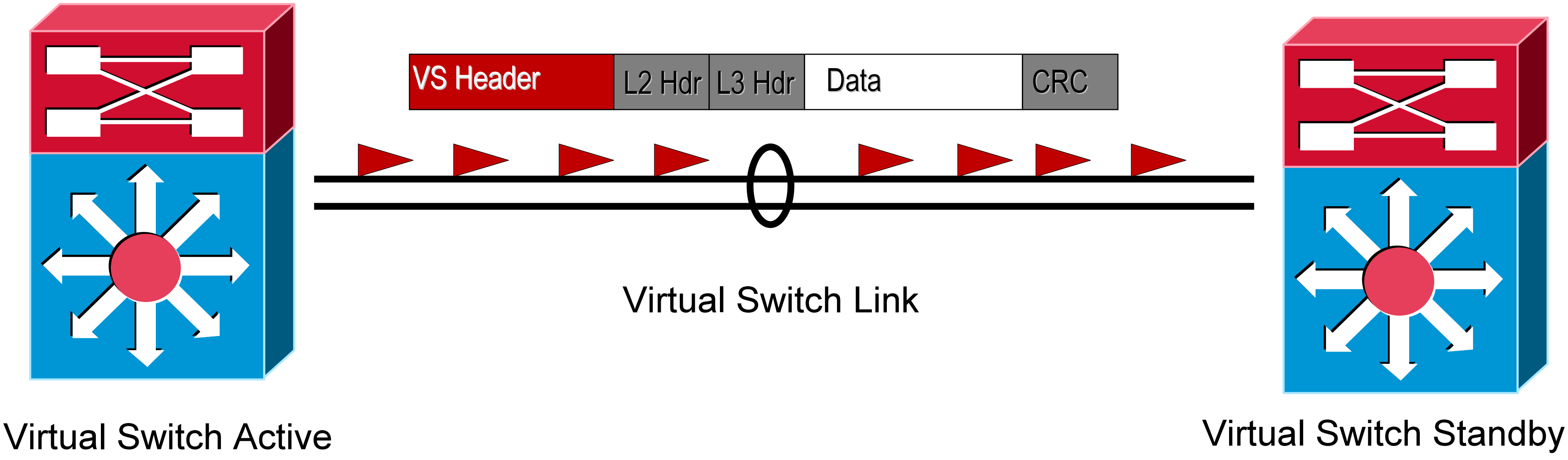
Virtual Switching System Architecture
Virtual Switch Link

The Virtual Switch Link is a special link joining each physical switch together - it extends the out of band channel allowing the active control plane to manage the hardware in the second chassis…

A Virtual Switch Link bundle can consist of up to 8 x 10GE links

All traffic traversing the VSL link is encapsulated with 32 byte “Virtual Switch Header” containing ingress and egress switchport indexes, class of service (COS), VLAN number, other important information from the layer 2 and layer 3 header

Control plane uses VSL CPU to CPU communications while the data plane uses VSL to extend the internal chassis fabric to the remote chassis

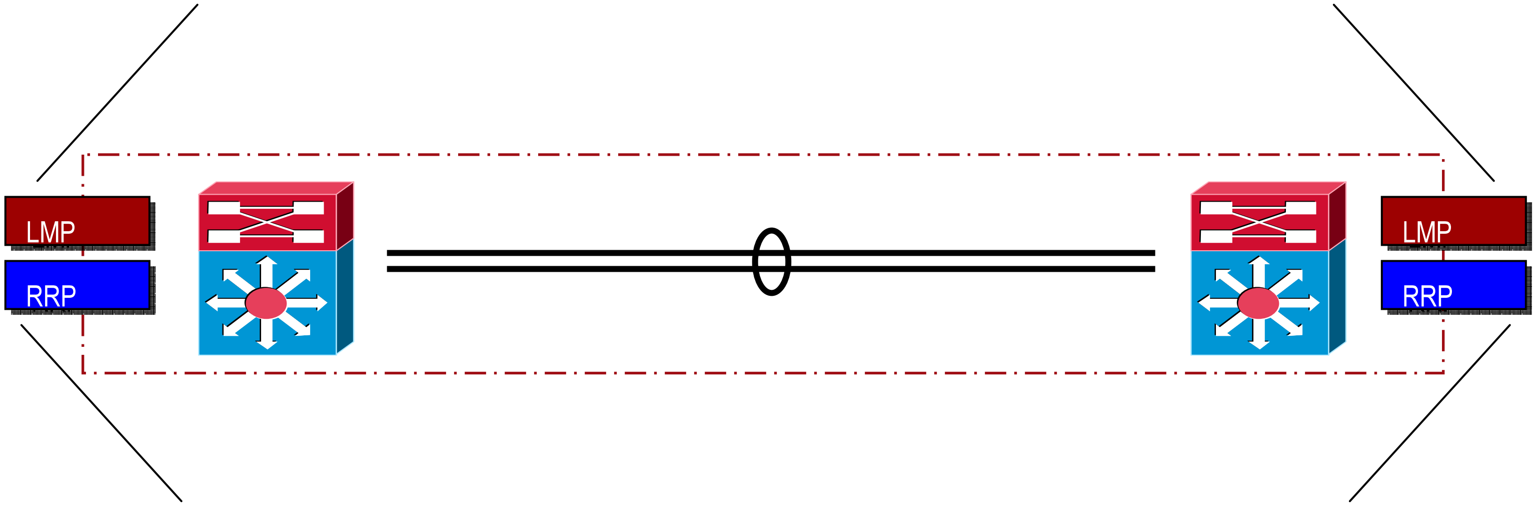


Virtual Switching System Architecture
VSL Initialization

Before the Virtual Switching System domain can become active, the Virtual Switch Link (VSL) must be brought online to determine Active and Standby roles. The initialization process essentially consists of 3 steps:

1. **Link Bringup** to determine which ports form the VSL

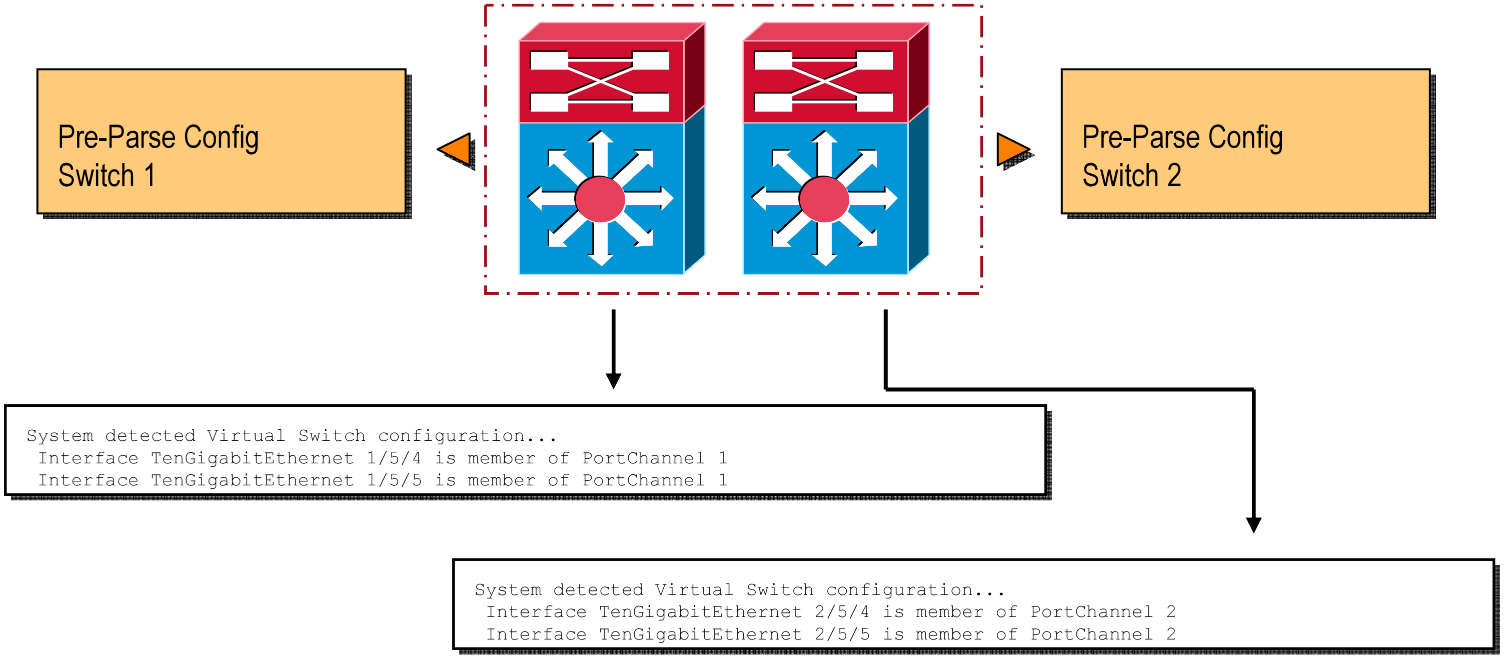
2. **Link Management Protocol (LMP)** used to track and reject Unidirectional Links, Exchange Chassis ID and other information between the 2 switches



3. **Role Resolution Protocol (RRP)** used to determine compatible Hardware and Software versions to form the VSL as well as determine which switch becomes Active and Hot Standby from a control plane perspective

Virtual Switching System Architecture
Link Bringup

Each member of the Virtual Switching System domain must determine which links are candidate for VSL very early on in the bootup cycle. The Switch Processor (SP) pre-parses the configuration to determine which links are configured for VSL…



```
System detected Virtual Switch configuration...
 Interface TenGigabitEthernet 1/5/4 is member of PortChannel 1
 Interface TenGigabitEthernet 1/5/5 is member of PortChannel 1
```

```
System detected Virtual Switch configuration...
 Interface TenGigabitEthernet 2/5/4 is member of PortChannel 2
 Interface TenGigabitEthernet 2/5/5 is member of PortChannel 2
```

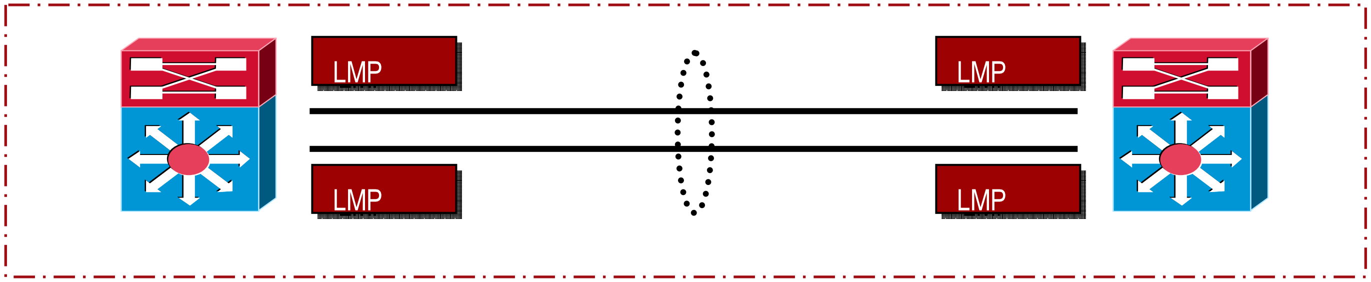
The SP will then bring up the line card/s where the VSL is configured, download the required configuration and initiate Link Management Protocol (LMP)

Virtual Switching System Architecture
Link Management Protocol (LMP)

LMP used to perform basic connectivity check on VSL link

LMP runs on each individual link that is part of the VSL, and is used to program information such as member details, forwarding indices, as well as perform the following checks:

1. Verify neighbor is Bi-Directional
2. Ensure the member is connected to another Virtual Switch
3. Transmit and receive keepalives to maintain health of the member and the VSL



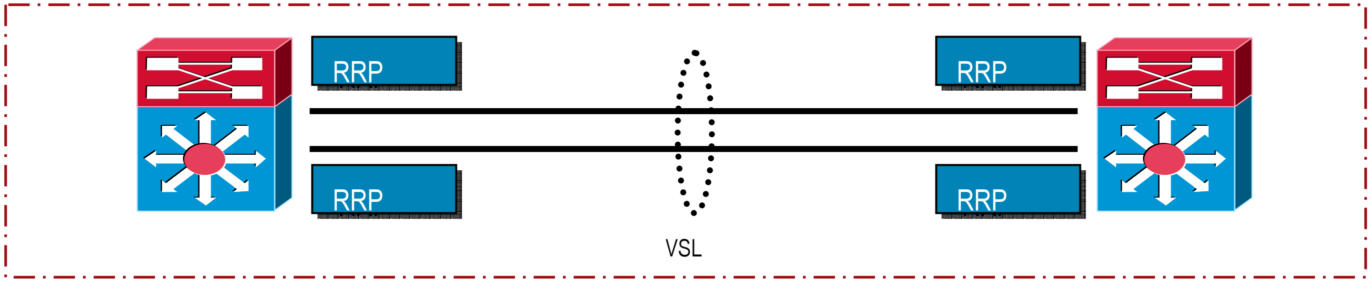
After successful LMP negotiation, a Peer Group (PG) is formed which is a collection of all VSL members that connects to the same VSS. For each PG, a Peer Group Control Link (PGCL) is elected to carry further control information…

Virtual Switching System Architecture
Role Resolution Protocol (RRP)

RRP (Role Resolution Protocol) is used to negotiate the role for each chassis. It is also part of VSLP and performs the following functions:

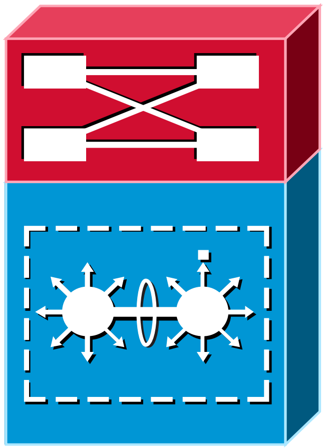
1 Determine whether hardware and software versions allow a Virtual Switching system to form

2 Determine which chassis will become Active and Hot Standby from a control plane perspective



Virtual Switching System Architecture
VSL Configuration Consistency Check

After the roles have been resolved through RRP, a Configuration Consistency Check is performed across the VSL switches to ensure proper VSL operation. The following items are checked for consistency:



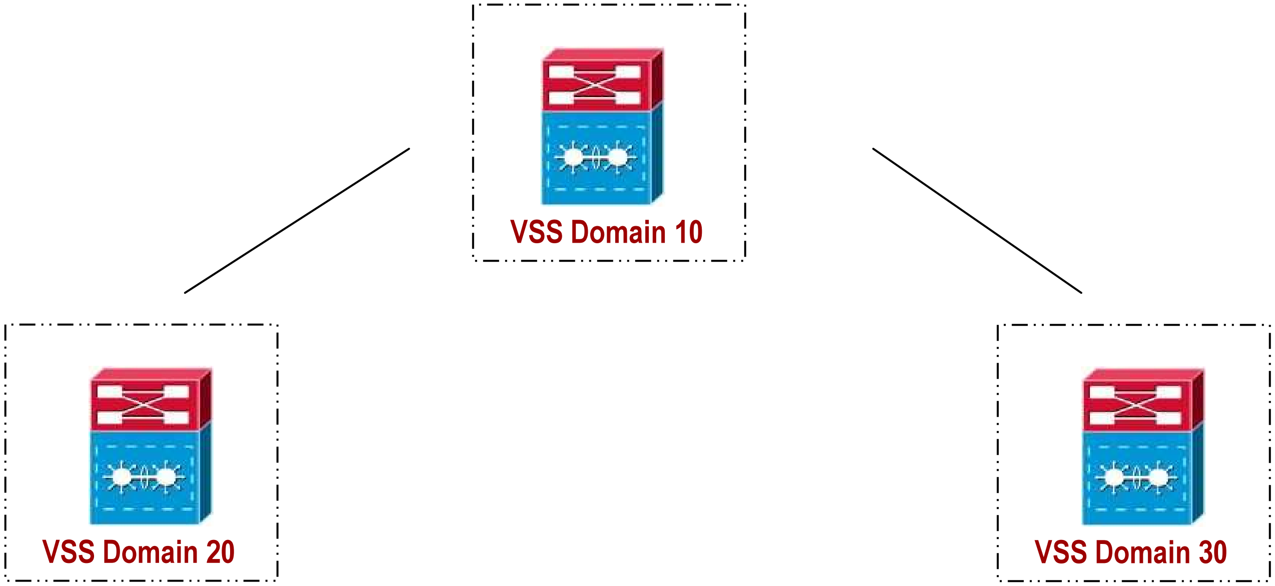
- Switch Virtual Domain ID
- Switch Virtual Node Type
- Switch Priority
- Switch Preempt
- VSL Port Channel Link ID
- VSL Port state, interfaces…
- Power Redundancy mode
- Power Enable on VSL cards

Note that if configurations do not match, the Hot-Standby switch will revert to RPR mode, disabling all non-VSL interfaces…

Virtual Switching System Architecture

Virtual Switch Domain

A Virtual Switch Domain ID is allocated during the conversion process and represents the logical grouping the 2 physical chassis within a VSS. It is possible to have multiple VS Domains throughout the network…



The configurable values for the domain ID are 1-255. It is always recommended to use a unique VSS Domain ID for each VSS Domain throughout the network…

Virtual Switching System Architecture
Router MAC Address Assignment

In a Virtual Switching System, there is also only ONE router MAC address to represent both physical chassis as one logical device.

- Each physical member of VSS pair consist of pool of MAC address stored in backplane EEPROM
- The VSS logical pair MAC address pool will be determined during the role resolution negotiation, thus every interfaces MAC are derived from ACTIVE chassis EEPROM residing on backplane.
- MAC address remains consistent across the switchover keeping ARP table consistent during switchovers
- Default gateway MAC remains the same, which is SVI MAC
- Individual VSS member MAC address are used during dual active condition

```
cr2-6500-vss#sh idprom switch 1 backplane detail | inc mac
  mac base = 0019.A927.3000

cr2-6500-vss#sh idprom switch 2 backplane detail | inc mac
  mac base = 0019.A924.E800
```

Router MAC = burnt-in or virtual mac-address

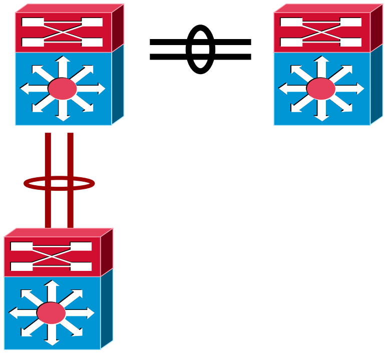
Instead of using default chassis mac-address assignment, from 12.2(33)SXH2 on wards virtual mac-addres can be specified as shown below:

```
VSS(config-vs-domain)# switch virtual domain 100

VSS(config-vs-domain)# mac-address mac-address <0020.1456.2456>
```

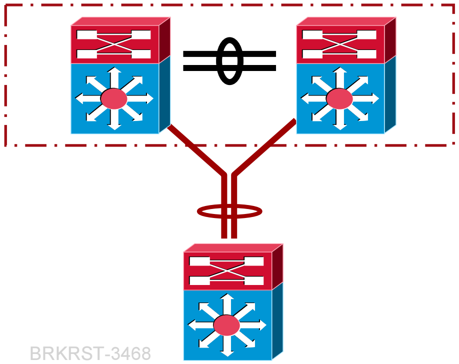
Etherchannel Concepts
Multichassis EtherChannel (MEC)

Prior to Virtual Switching System, Etherchannels were restricted to reside within the same physical switch. In a Virtual Switching environment, the 2 physical switches form a single logical network entity - therefore Etherchannels can now also be extended across the 2 physical chassis…

Standalone: **Regular Etherchannel on single chassis**



Virtual Switch: **Multichassis EtherChannel across 2 VSL-enabled Chassis**

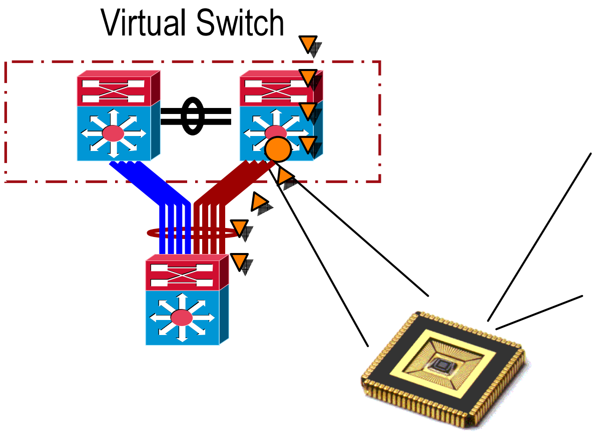


Both LACP and PAGP Etherchannel protocols and Manual ON modes are supported…

- MEC links on both switches are managed by PAgP or LACP running on the ACTIVE switch via internal control messages
- All the rules and properties of EtherChannel applies to MEC such as negotiation, link characteristics (port-type, trunk), QoS, etc.
- Do not use "on" and "off" options with PAgP or LACP protocol negotiation:
  - PAgP—Run Desirable-Desirable with MEC links
  - LACP—Run Active-Active with MEC links
- L2 MEC enables loop free topology and doubles the uplink bandwidth as no links are blocked
- L3 MEC provides reduced neighbor counts, consistent load-sharing (L2 and L3) and reduced VSL link utilization for multicast flows

Etherchannel Concepts
Etherchannel Hash for MEC

Deciding on which link of a Multi-chassis Etherchannel to use in a Virtual Switch is skewed in favor towards local links in the bundle - this is done to avoid overloading the Virtual Switch Link (VSL) with unnecessary traffic loads. Localizing the decision to use a link in the bundle that is resident on the local Switch (thus avoiding forwarding over the VSL) is done as follow…



- The BUNDLE_SELECT register in the port ASIC is programmed to see only the local links of the Etherchannel bundle even though links may exist in the same bundle are resident in the VSS peer chassis… This behavior is fixed and cannot be changed by any configuration option…

NOTE: If all links in the local bundle go down, then the BUNDLE_SELECT register is programmed to point packets to the VSL…

RBH (for MEC) 8 Link Bundle Example	
Bit 7	Link 1
Bit 6	Link 1
Bit 5	Link 2
Bit 4	Link 2
Bit 3	Link 3
Bit 2	Link 3
Bit 1	Link 4
Bit 0	Link 4

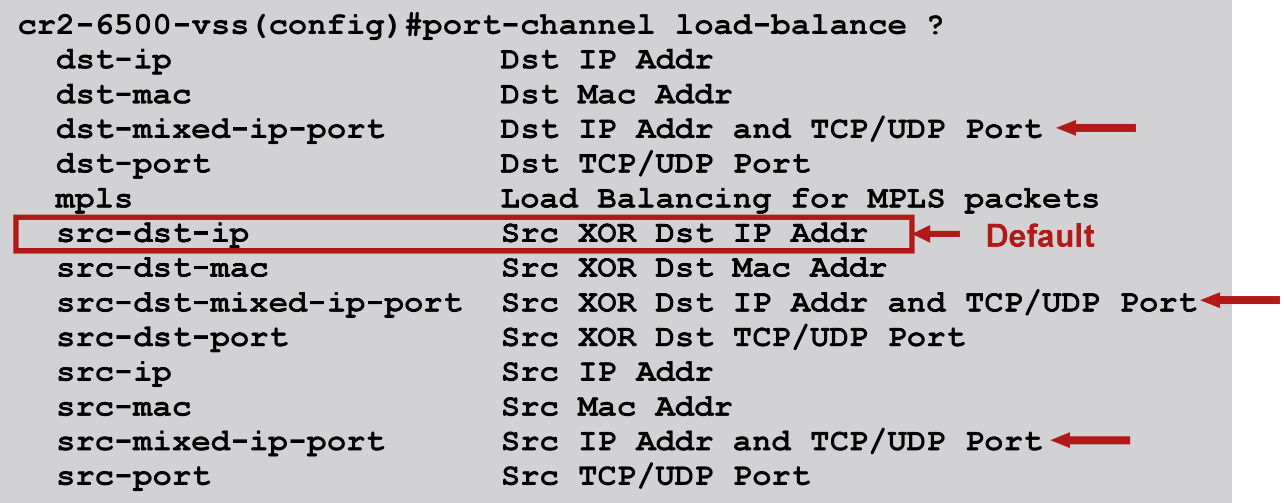
- RBH values are reprogrammed for each core to reflect only the local links that are in the Etherchannel bundle… A new hash distribution algorithm has been introduced with the 12.2(33)SXH release which allows for members of a port channel to be added or removed without the requirement for all traffic on the existing members to be temporarily dropped…

Etherchannel Concepts

Load-balancing

Load balancing is based on a specific hardware hash algorithm that maps to eight possible buckets.

- Load balancing is based on a specific hardware hash algorithm that maps to eight possible buckets. Different hardware implementations on the various platforms (37xx, 4500, and 6500)
- Most hash algorithms are designed to ensure fairness, but will be based on certain assumptions:
  - **Core—Many-to-many flows**
  - **Access—Many-to-few flows**
- There is no ‘One Size Fits All’ configuration. Need to analyze your network and tune based on your requirements. Some rules of thumb:
  - **The more values used in the hash the more likely to be ‘fair’**
  - **Layer 4 hashing tends to be more random than L3**
  - **L2 is not efficient when everyone is talking to a single default gateway**



```
cr2-6500-vss(config)#port-channel load-balance ?
  dst-ip                    Dst IP Addr
  dst-mac                   Dst Mac Addr
  dst-mixed-ip-port         Dst IP Addr and TCP/UDP Port
  dst-port                  Dst TCP/UDP Port
  mpls                      Load Balancing for MPLS packets
  src-dst-ip                Src XOR Dst IP Addr
  src-dst-mac               Src XOR Dst Mac Addr
  src-dst-mixed-ip-port     Src XOR Dst IP Addr and TCP/UDP Port
  src-dst-port              Src XOR Dst TCP/UDP Port
  src-ip                    Src IP Addr
  src-mac                   Src Mac Addr
  src-mixed-ip-port         Src IP Addr and TCP/UDP Port
  src-port                  Src TCP/UDP Port
```

Default

Conversion Process

Conversion Process
Conversion to VSS

Configuration for the conversion takes the following path…

Switch - 1

```
Router(config)#host VSS
VSS(config)#switch virtual domain 100

Domain ID 10 config will take effect only
after the exec command 'switch convert mode
virtual' is issued

VSS(config-vs-domain)#switch 1
VSS(config-vs-domain)#exit

VSS(config)#interface port-channel 1
VSS(config-if)#switch virtual link 1

VSS(config-if)#interface range tenG 5/4 - 5
VSS(config-if-range)#channel-group 1 mode on
```

Switch - 2

```
Router(config)#host VSS
VSS(config)#switch virtual domain 100

Domain ID 10 config will take effect only
after the exec command 'switch convert mode
virtual' is issued

VSS(config-vs-domain)#switch 2
VSS(config-vs-domain)#exit

VSS(config-if)#interface port-channel 2
VSS(config-if)#switch virtual link 2

VSS(config-if)#interface range tenG 5/4 - 5
VSS(config-if-range)#channel-group 2 mode on
```

Conversion Process

Conversion to VSS

Configuration for the conversion takes the following path…

Switch - 1

```
vss#switch convert mode virtual

This command will convert all interface
names to naming convention "interface-type
switch-number/slot/port", save the running
config to startup-config and reload the
switch.
Do you want to proceed? [yes/no]: yes
Converting interface names
Building configuration...
[OK]
Saving converted configuration to bootflash:
...
Destination filename [startup-
config.converted_vs-20071031-150039]?

AT THIS POINT THE SWITCH WILL REBOOT
```

Switch - 2

```
vss#switch convert mode virtual

This command will convert all interface names
to naming convention "interface-type switch-
number/slot/port", save the running config to
startup-config and reload the switch.
Do you want to proceed? [yes/no]: yes
Converting interface names
Building configuration...
[OK]
Saving converted configuration to bootflash:
...
Destination filename [startup-
config.converted_vs-20071031-150018]?

AT THIS POINT THE SWITCH WILL REBOOT
```

Conversion Process

Conversion to VSS

Configuration for the conversion takes the following path…

Switch - 1

```
SWITCH CONSOLE OUTPUT

<…snip…>
vss-demo#switch accept mode virtual
interface Port-channel2
 switch virtual link 2
 no shutdown
interface TenGigabitEthernet2/5/4
 channel-group 2 mode on
 no shutdown
interface TenGigabitEthernet2/5/5
 channel-group 2 mode on
 no shutdown

This command will populate the above VSL
configuration from the standby switch into the
running configuration.
The startup configuration will also be updated
with the new merged configuration if merging is
successful.
Do you want to proceed? [yes/no]: yes
Merging the standby VSL configuration...

Building configuration...

00:11:33: %PFINIT-SW1_SP-5-CONFIG_SYNC:
Sync'ing the startup configuration to the
standby Router. [OK]
```

Switch - 2

```
SWITCH CONSOLE OUTPUT

<…snip…>

Copyright (c) 1986-2007 by Cisco Systems, Inc.
Compiled Wed 10-Oct-07 01:02 by chrisvan
00:02:42: %CRYPTO-6-ISAKMP_ON_OFF: ISAKMP is
OFF
00:02:42: %CRYPTO-6-ISAKMP_ON_OFF: ISAKMP is
OFF
vss-sdby>
Standby console disabled

vss-sdby>
```

Conversion Process

Conversion to VSS

Configuration for the conversion takes the following path…

Switch - 1

```
vss#sh switch virtual
Switch mode                    : Virtual Switch
Virtual switch domain number : 10
Local switch number            : 1
Local switch operational role: Virtual Switch Active
Peer switch number             : 2
Peer switch operational role : Virtual Switch Standby
vss-demo#
```

Switch - 2

```
vss-sdby>enable
Standby console disabled

vss-sdby>
```

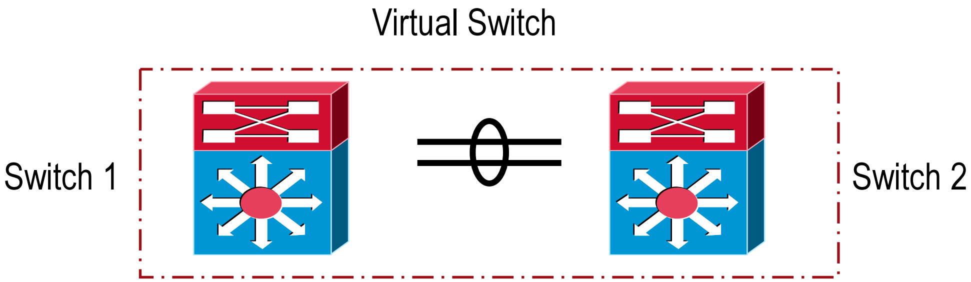
Both switches are now converted with Switch 1 as the VSS Active and Switch 2 as the VSS Hot standby

Switch 2 console is now disabled for normal console activity…

Conversion Process

Boot-up Priority

By default the first switch to boot will assume VSS Active role - this behavior can be changed by configuring switch priority (higher priority uses a higher number). Switch with higher priority will have a higher chance to become active during VSS boot up. This priority configuration only takes effect after a reload of both switches at the same time.

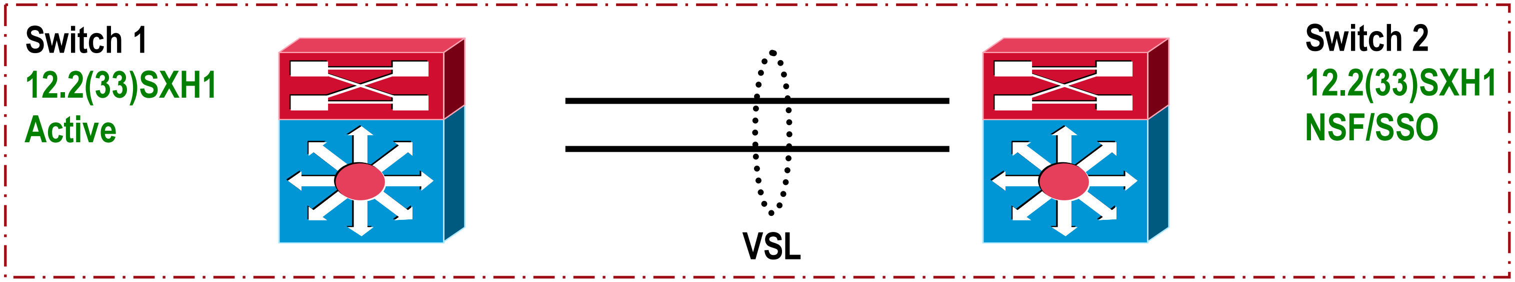


```
VSS#sh switch virtual role

Switch   Switch Status  Preempt     Priority  Role      Session ID
         Number         Oper(Conf)  Oper(Conf)          Local  Remote
------------------------------------------------------------------
LOCAL     1      UP     FALSE(N)    110(110)  ACTIVE    0      0
REMOTE    2      UP     FALSE(N)    100(100)  STANDBY   9114   1391


In dual-active recovery mode: No
```

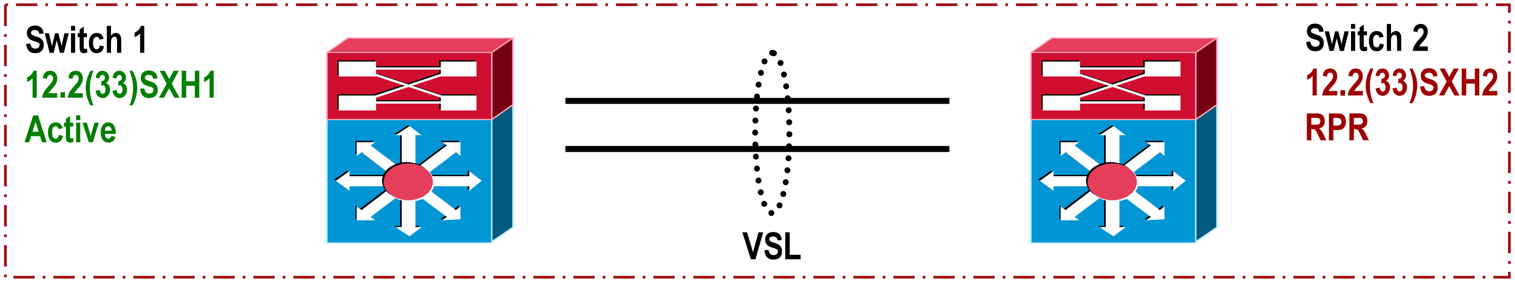
High Availability

High Availability
Redundancy Schemes

The default redundancy mechanism between the 2 VSS chassis and their associated supervisors is NSF/SSO, allowing state information and configuration to be synchronized. Additionally, only in NSF/SSO mode does the Standby supervisor PFC, Switch Fabric, modules and their associated DFCs become active…



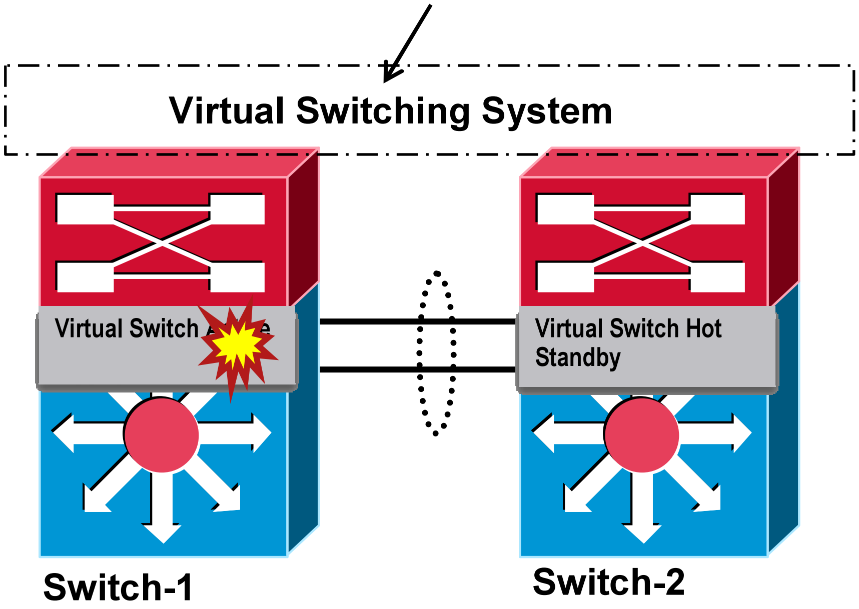
Should a mismatch of information occur between the Active and Standby Chassis, the Standby Chassis will revert to RPR mode, where only configuration is synchronized, but PFC, Switch Fabric and modules will not be brought up



Virtual Switching System

Inter Chassis NSF/SSO

1 Virtual Switch Active incurs a supervisor outage

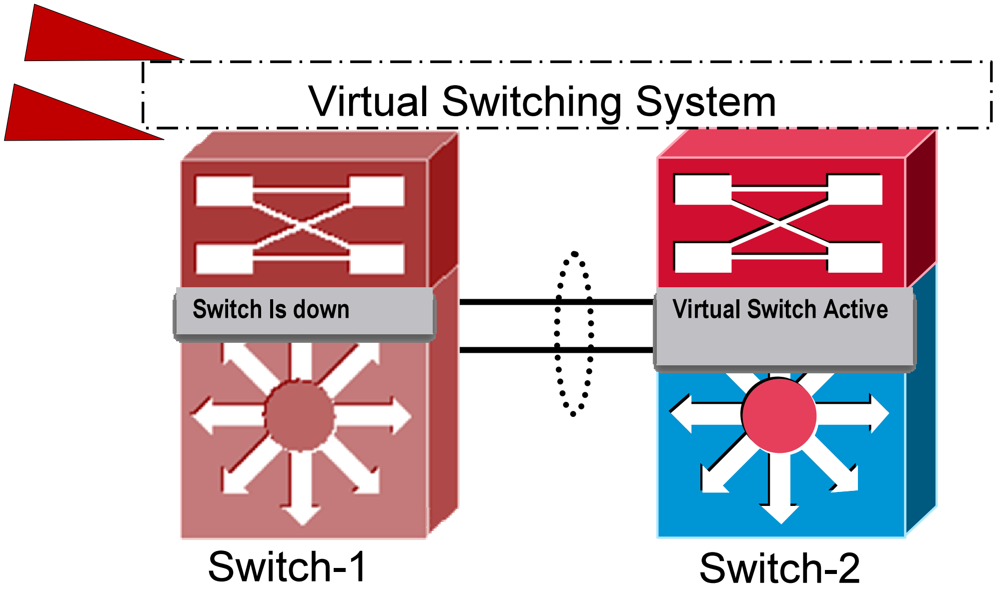


2 Standby Supervisor takes over as Virtual switch Active

Virtual Switch Standby initiates graceful restart

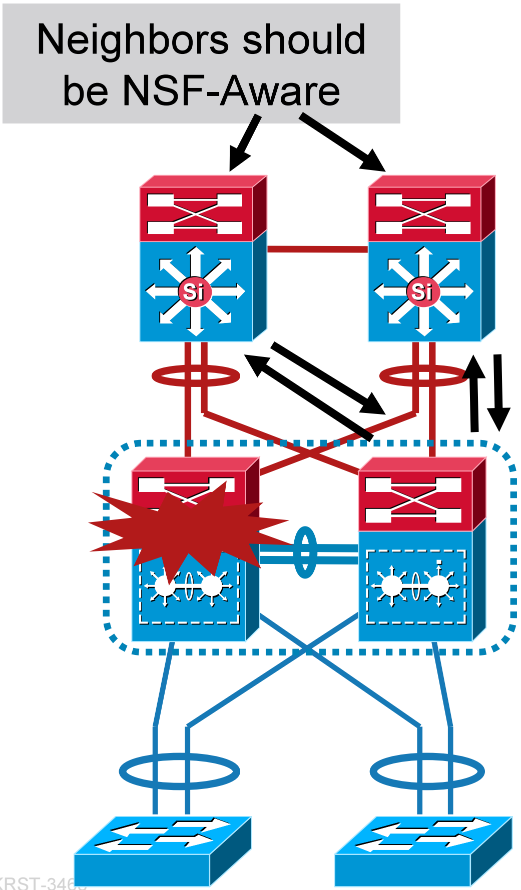
Non Stop forwarding of packets will continue using hardware entries as Switch-2 assumes active role

NSF aware neighbors exchange updates with Virtual Switch Active



High Availability
NSF Aware Layer 3 Neighbors

NSF feature with SSO minimizes the amount of traffic loss following supervisor switchover while continuing to forward traffic using hardware entries. In VSS environment this feature is required to minimize traffic disruption in the event such as supervisor failure that causes supervisor switchover.


Neighbors should be NSF-Aware

- **NSF-aware and NSF-capable routers provide for transparent routing protocol recovery**
- Graceful restart extensions enable neighbor recovery without resetting adjacencies
- Routing database re-synchronization occurs in the background
- **An NSF-capable router continuously forwards packets during an SSO processor recovery**
- EIGRP, OSPF, IS-IS and BGP are NSF capable and aware protocols
- Sup720, Sup32, Sup IV/V and Cat37xx supports NSF functionality

High Availability

Dual-Active Detection

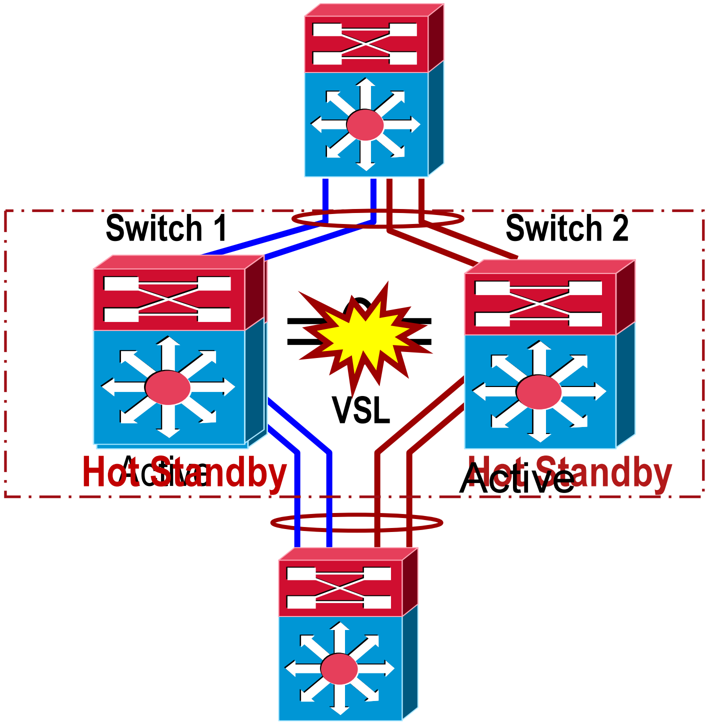
If the entire VSL bundle should happen to go down, the Virtual Switching System Domain will enter a Dual Active scenario where both switches transition to Active state and share the same network configuration (IP addresses, MAC address, Router IDs, etc…) potentially causing communication problems through the network…

Following to Dual Active scenario, if the Virtual Switching System is configured for dual-active detection following steps will take place.

1. Dual-Active detection using the detection method enabled in the system. Dual-Active protocols are Pagp+, Fast Hello and IP BFD

2. Further network disruption is avoided by bringing down VSS active switch interfaces connected to neighboring devices .

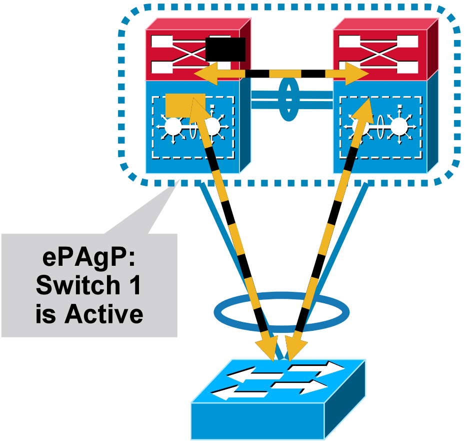
3. Dual-Active recovery, when VSL recovers , the switch that has all it's interfaces brought down in the previous step will reload to boot in a preferred standby state



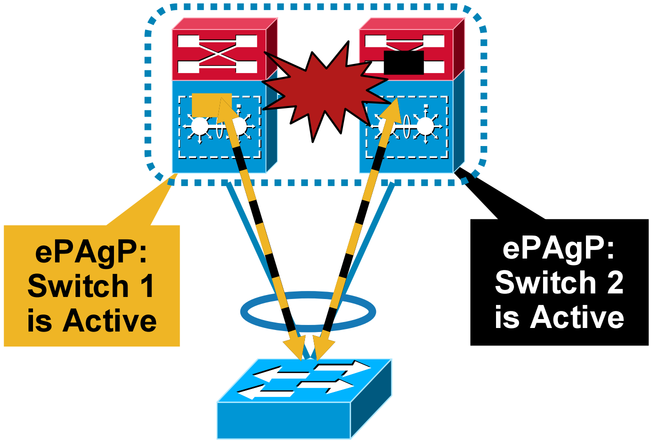
Virtual Switching System

Dual Active—Enhanced PAgP

- Enhanced PAgP provides a new TLV which uses the ID (MAC address) of the active switch. Only the ACTIVE switch originates ePAgP messages in normal mode
- In normal operations all enhanced PAgP neighbors reflects ID of an active switch back upstream on both uplinks
- Once the VSL bundle goes down switch 2 goes active, it generate its own ePAgP message with its own ID via ePAgP supporting neighbor to switch 1



Normal Mode



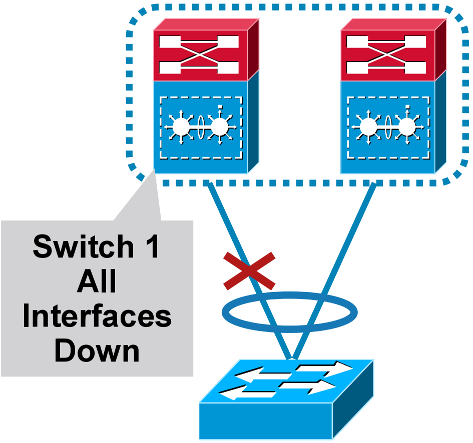
Dual Active Detection

```
cr2-6500-vss#sh switch virtual dual-active summary
Pagp dual-active detection enabled: Yes
Bfd dual-active detection enabled: Yes

No interfaces excluded from shutdown in
 recovery mode

In dual-active recovery mode: Yes
  Triggered by: PAgP detection
  Triggered on interface: Gi2/8/19
  Received id: 0019.a927.3000
  Expected id: 0019.a924.e800
```

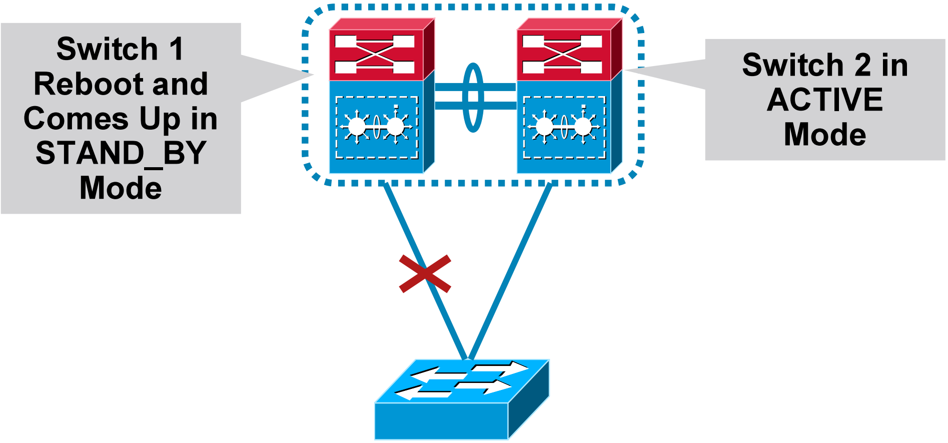
Virtual Switching System

Dual Active Recovery—Enhanced PAgP

- Switch 1 detects that switch 2 is now also active triggering dual active condition thus switch 1 brings down all the local interfaces to avoid network instability
- Until VSL link restoration occurs, switch 1 is isolated from the network; once the VSL link comes up, the role negotiation determines that switch 1 needs to come up in STAND_BY mode hence it reboots itself; finally, all interface on switch 1 are brought on line and switch 1 assumes STAND_BY role
- If any configuration change occurs during the dual active recovery stage, the recovered system will go in RPR+ mode and will require manual intervention



Dual Active Recovery



VSS Restoration

Virtual Switching System

Dual Active Recovery—Enhanced PAgP

- ePAgP dual active detection is enabled by default, however specific MEC group must be trusted via CLI

    Need to explicitly trust enhanced PAgP neighbors and requires MEC in admin down state (both add and remove the trust)

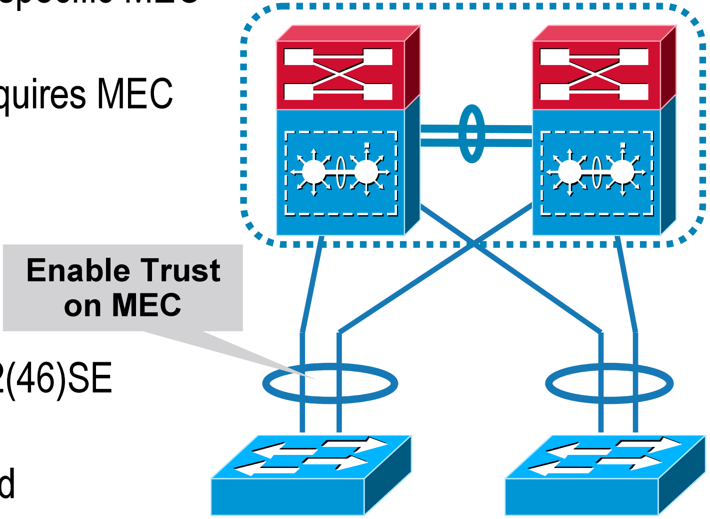
    PAgP protocol must be running on MEC links – L2 or L3

- ePAgP is supported in:

    6500 in 12.2(33)SXH and 4500 in 12.2(44)SG

    3750, will support pagp+ in upcoming release called 12.2(46)SE

- Use “exclude interface” option to keep specified port to remain up during the dual active recovery, e.g., designated management port



Enable Trust on MEC

```
cr2-6500-vss(config)#switch virtual domain 10
cr2-6500-vss(config-vs-domain)#dual-active detection pagp trust channel-group 205
cr2-6500-vss(config-vs-domain)#dual-active exclude interface <port>

cr2-6500-vss#sh switch virtual dual-active pagp
PAgP dual-active detection enabled: Yes
PAgP dual-active version: 1.1

Channel group 205 dual-active detect capability w/nbrs
Dual-Active trusted group: Yes
          Dual-Active     Partner              Partner    Partner
Port      Detect Capable  Name                 Port       Version
Gi1/8/19  Yes             cr7-6500-3           Gi5/1      1.1
Gi1/9/19  Yes             cr7-6500-3           Gi6/1      1.1
```

Virtual Switching System
Configuration

1

```
switch virtual domain 10
 switch mode virtual
 switch 1 priority 110
 switch 2 priority 100
 dual-active detection pagp trust channel-group 202
```

2

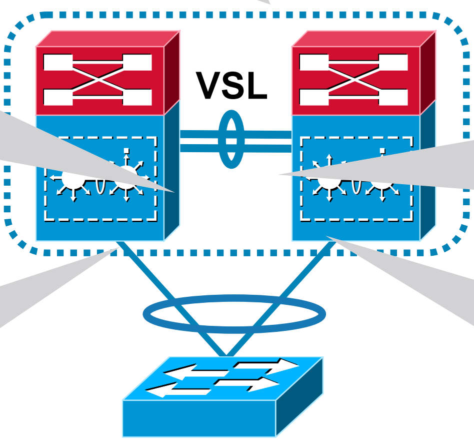
Switch 1

```
interface Port-channel1
 description VSL Link from Switch 1
 no switchport
 no ip address
switch virtual link 1
 mls qos trust cos
 no mls qos channel-consistency
interface range tenGigabitEthernet 1/4 – 5
 channel-group 1 mode on
```

2

Switch 2

```
interface Port-channel2
 description VSL Link from Switch 2
 no switchport
 no ip address
switch virtual link 2
 mls qos trust cos
 no mls qos channel-consistency
interface range tenGigabitEthernet 1/4 – 5
 channel-group 1 mode on
```



MEC

```
interface GigabitEthernet1/8/23
 description Access Switch
 switchport
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 202
 switchport trunk allowed vlan 2,102
 <snip>
channel-protocol pagp
channel-group 202 mode desirable
```

3

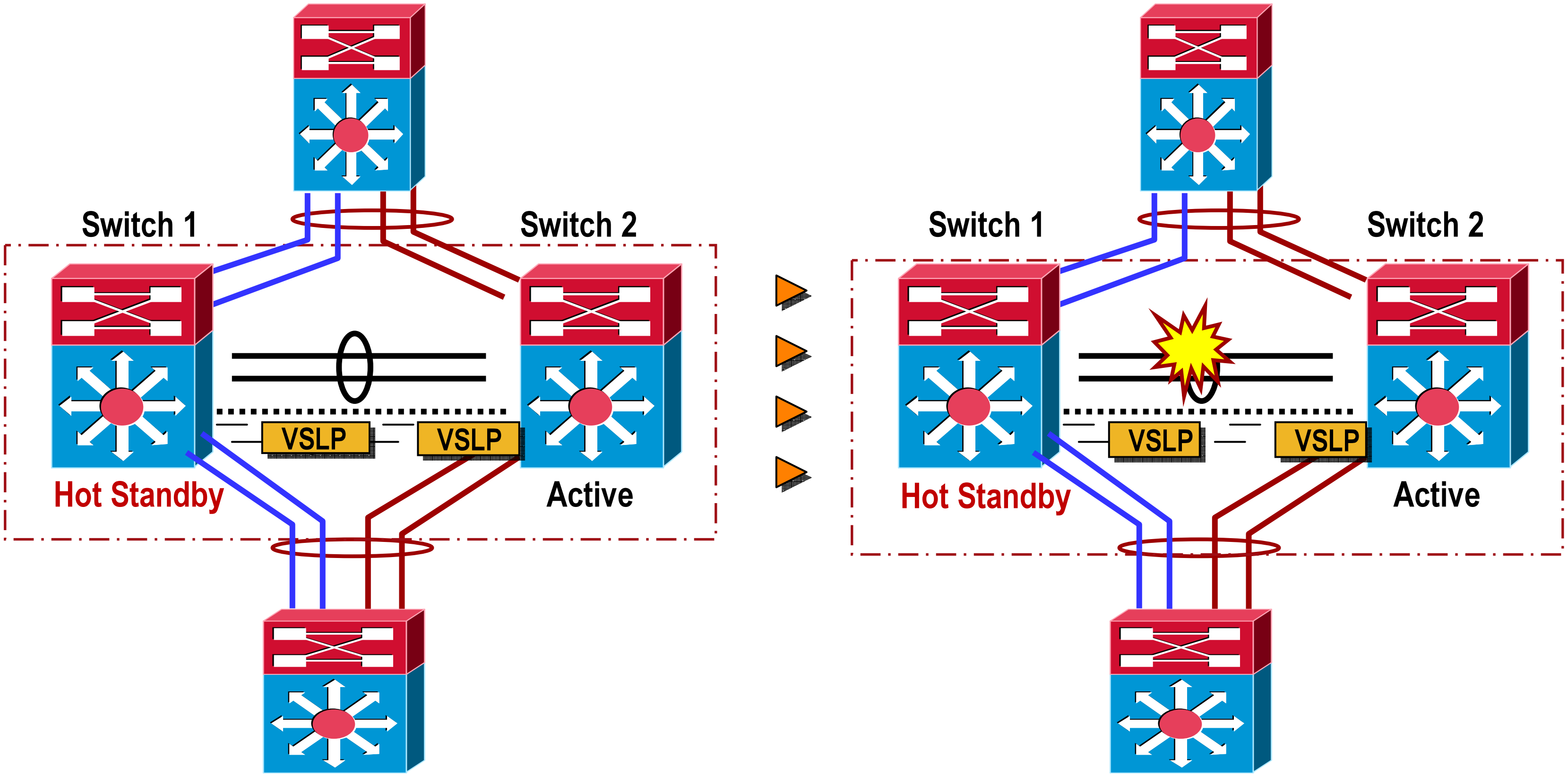
```
interface GigabitEthernet2/8/23
 description Access Switch
 switchport
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 202
 switchport trunk allowed vlan 2,102
 <snip>
channel-protocol pagp
channel-group 202 mode desirable
```

3

```
interface Port-channel202
 description Access Switch
 switchport
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 202
 switchport trunk allowed vlan 2,102
 switchport mode trunk
 switchport nonegotiate
```

High Availability
Dual-Active Detection: VSLP Fast Hello

It is a new feature will be available in 12.2(33)SXI software release. Dual-Active heartbeat messages are exchanged over a heart beat link between switch-1 and switch-2. Information such as Switch-id, Priority and Peer state information exchanged to deterministically decide the switch role during dual-active detection.



Subsequent to Dual Active event, dual active detection takes place and switch-1 will bring down all it’s local interfaces *

High Availability

Dual-Active Detection: VSLP Fast Hello

To Enable VSLP Fast Hello Dual Active detection mechanism. New protocol needs to be enabled on heartbeat link as well globally under Virtual Switching System domain config mode. Configuring more than one heart beat link is allowed using this protocol for redundancy purposes

```
1.   Enable globally
VSS#conf t
Enter configuration commands, one per line.  End with CNTL/Z
VSS(config)#switch virtual domain 1
VSS(config-vs-domain)#dual-active detection fast-hello

2.   Enable at interface level
VSS#conf t
VSS(config)#int range gig 1/9/5 , gig2/9/5
VSS(config-if-range)#dual-active fast-hello
```

Configuration and protocol operational status can be verified using following show command

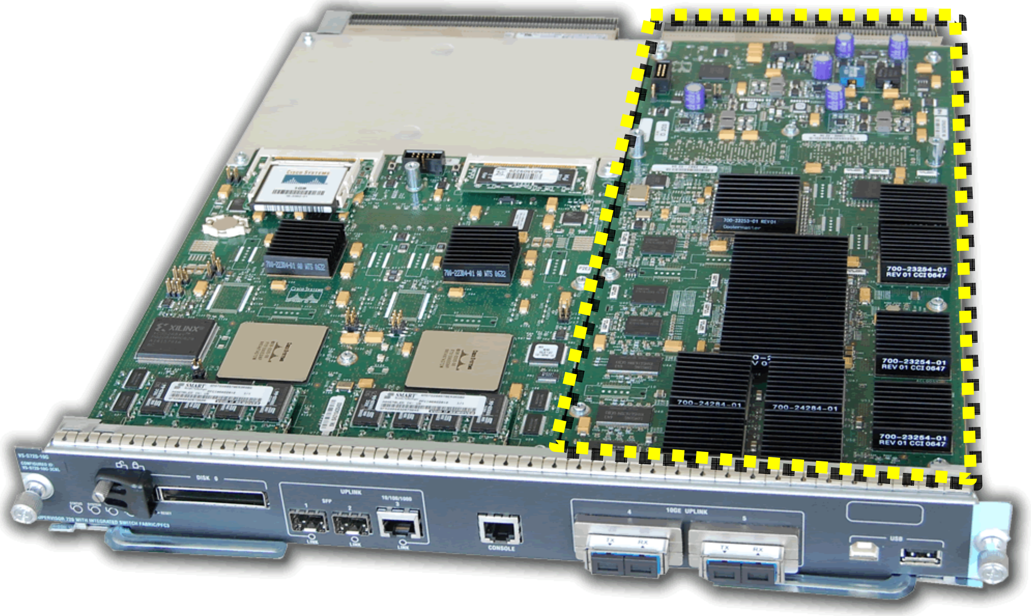
```
VSS#sh switch virtual dual-active fast-hello
Fast-hello dual-active detection enabled: Yes

Fast-hello dual-active interfaces:
Port       State (local only)
-----------------------------
Gi1/9/5    Link up
Gi2/9/5    -
```

Hardware Requirements

Hardware and Software Requirements

In order to enable the Virtual Switching System feature and configure the Virtual Switch Links (VSL) between 2 Catalyst 6500 chassis, the new Catalyst 6500 Virtual Switching Supervisor 720 is required to be used. It is the only Supervisor that will support VSS as it supports both the new PFC3C/XL forwarding engine…

12.2(33)SXH1 or later; current recommendation is 12.2(33)SXH2(a)



VS-S720-10G-3C/XL

The PFC3C/XL contains new hardware to support the extra LTL indices and mappings required to forward traffic across multiple physical chassis, lookup enhancements as well as MAC address table handling enhancements…

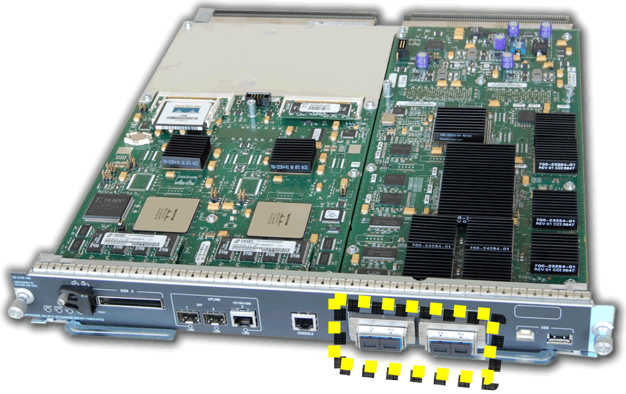
Hardware Requirements

VSL-Capable Interfaces

The VSL is a special link that requires extra headers to be imposed onto the frame. These require new port ASICs that exist only on the 10 GigabitEthernet interfaces on the following modules… WS-X6716-10G-3C/XL module is supported starting from 12.2(33)SXH2* in non VSL config

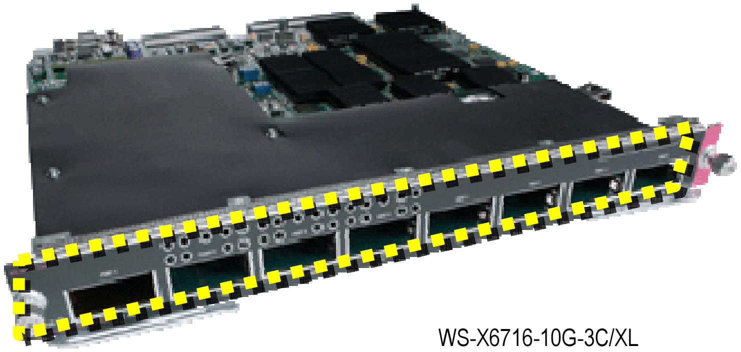
Note that these interfaces may also be used as standard network interfaces

VS-S720-10G-3C/XL

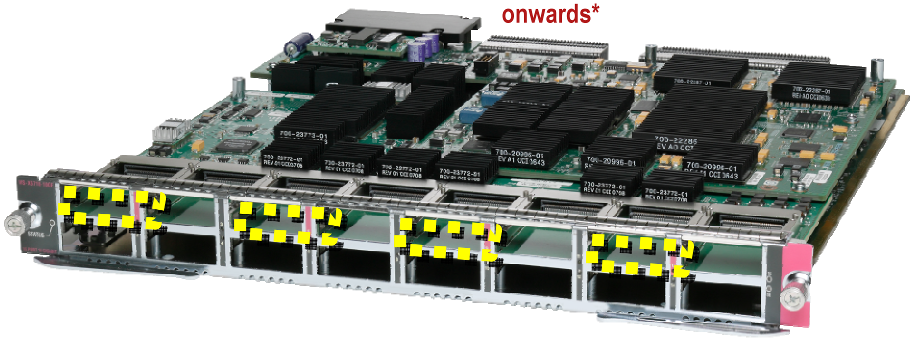


These interfaces are based off the new port ASIC, allowing for frames across the VSL to be encapsulated / de-encapsulated with the VSH…

WS-X6708-10G-3C/XL



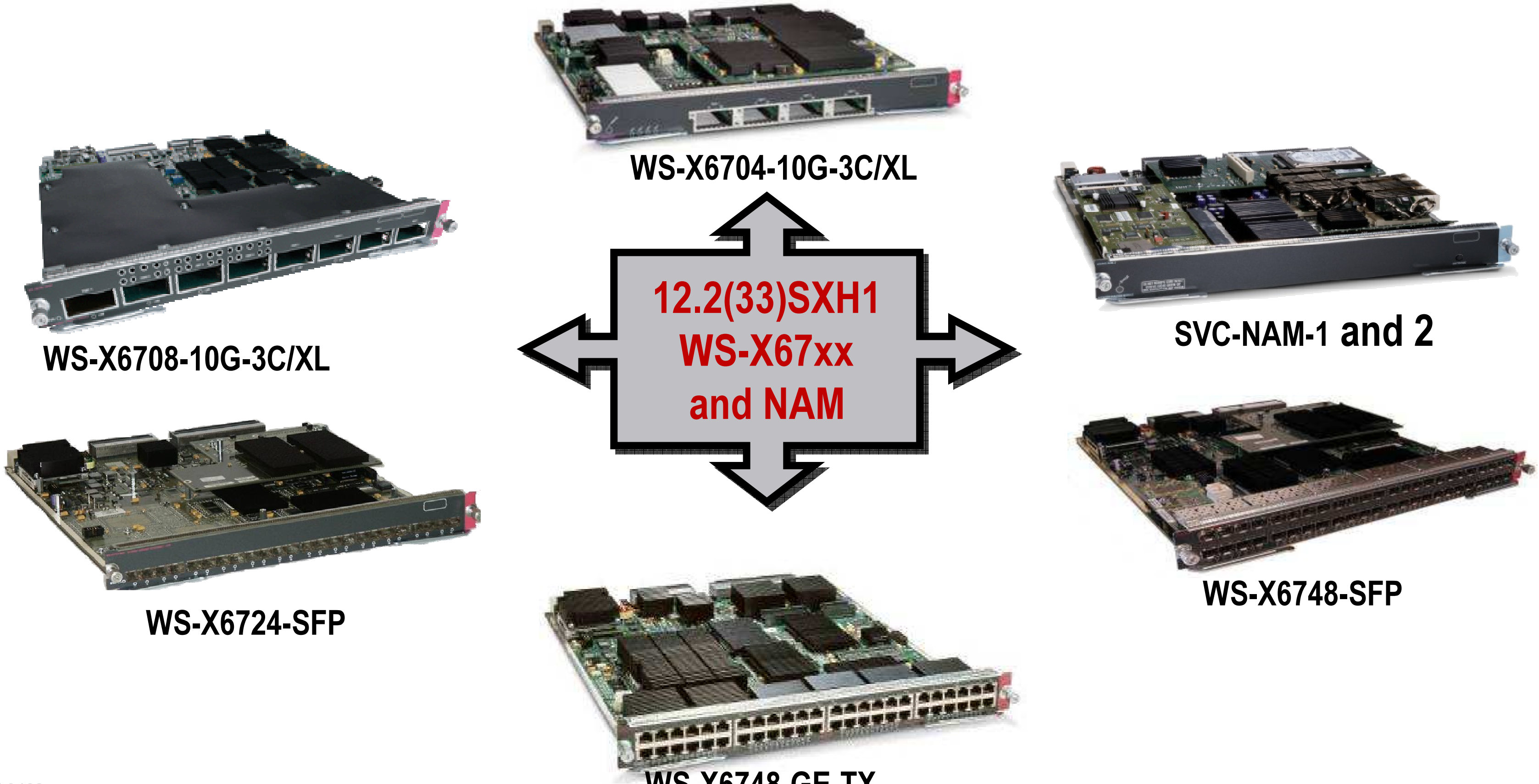
WS-X6716-10G-3C/XL
support for VSL is from 12.2(33)SXI onwards*



Hardware Requirements

Other Supported Modules…

Modules that may exist with current software version in the VSS domain include all WS-X67xx-series, as well as SVC-NAM-1 and SVC-NAM-2.



WS-X6704-10G-3C/XL

WS-X6708-10G-3C/XL

12.2(33)SXH1
WS-X67xx
and NAM

SVC-NAM-1 and 2

WS-X6724-SFP

WS-X6748-SFP

WS-X6748-GE-TX

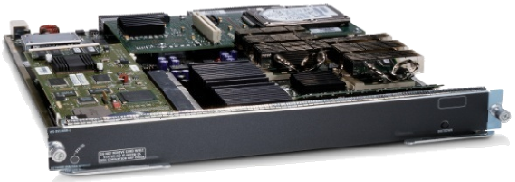
Hardware Requirements
Service Module support…

Other modules that may exist in the VSS domain with upcoming software release 12.2(33)SXI is Service modules FWSM,ACE,IDSM-2 and FWSM.

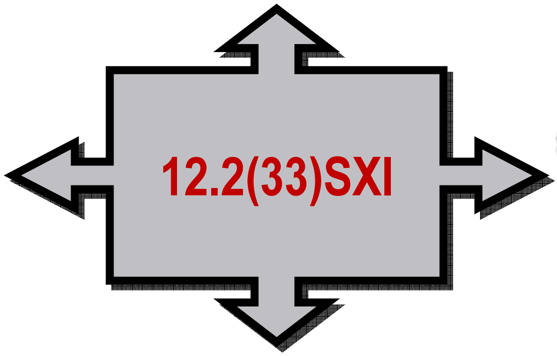
Application Control Engine (ACE)

ACE10/20-6500-K9

Firewall Services Module (FWSM)



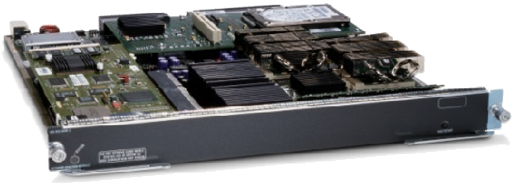
WS-SVC-FWM-1-K9



Wireless Services Module (WiSM)

WS-SVC-WISM-1-K9

Intrusion Detection System Services Module (IDSM-2)



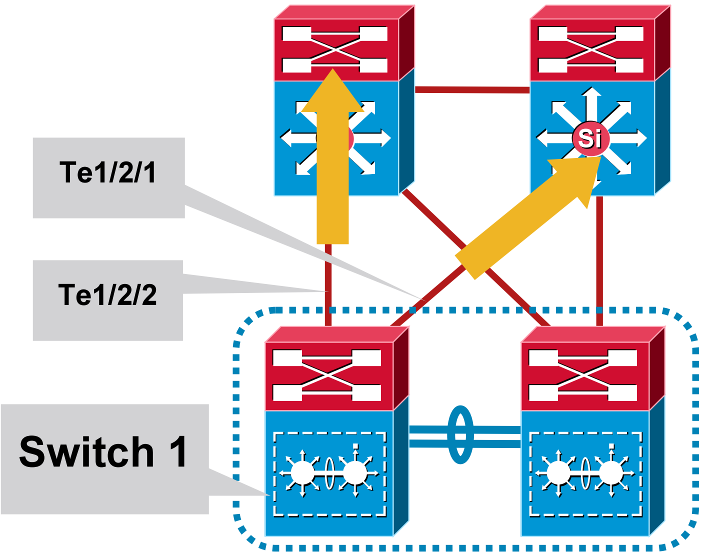
WS-SVC-IDSM2-K9

Traffic Flow and Topology Considerations

VSS Enabled Campus Design

Unicast ECMP Traffic Flows

- ECMP follows a similar behavior, **local** links are preferred and all traffic is forwarded out of a locally attached link
- Hardware FIB inserts entries for ECMP routes using locally attached links
- If all local links fail the FIB is programmed to forward across the VSL link



```
cr2-6500-vss#sh ip route 10.121.0.0 255.255.128.0 longer-prefixes
D       10.121.0.0/17
           [90/3328] via 10.122.0.33, 2d10h, TenGigabitEthernet2/2/1
           [90/3328] via 10.122.0.27, 2d10h, TenGigabitEthernet1/2/1
           [90/3328] via 10.122.0.22, 2d10h, TenGigabitEthernet2/2/2
           [90/3328] via 10.122.0.20, 2d10h, TenGigabitEthernet1/2/2

cr2-6500-vss#sh mls cef 10.121.0.0 17 switch 1

Codes: decap - Decapsulation, + - Push Label
Index  Prefix              Adjacency
102400 10.121.0.0/17       Te1/2/2         , 0012.da67.7e40 (Hash: 0001)
                           Te1/2/1         , 0018.b966.e988 (Hash: 0002)
```

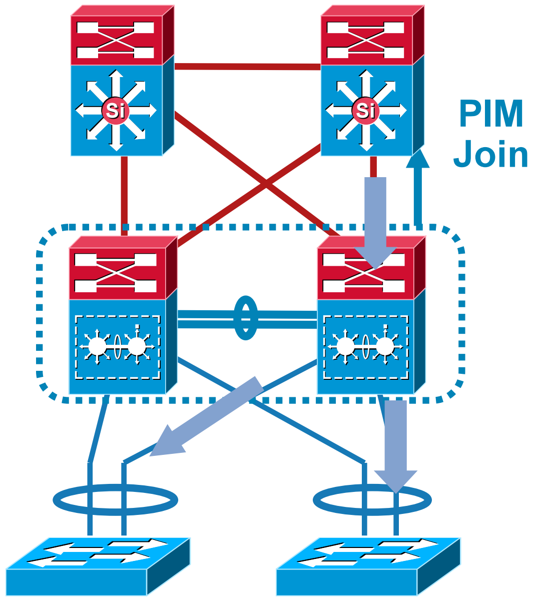
Four ECMP Entries

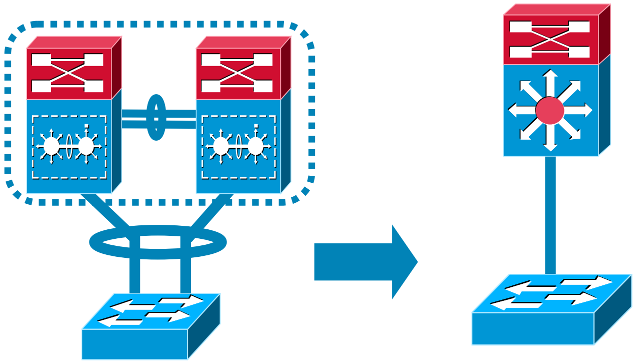
Two FIB Entries

VSS Enabled Campus Design

Multicast Traffic Flows—Layer 2 Access

- VSS represents a single multicast router, therefore a single PIM join is sent upstream, simplified access multicast topology
- With MEC, multicast traffic is forwarded via local line card and does egress replication when DFC line cards available
- Single, logical, multicast router eliminates the non-RPF traffic, efficiently utilizing uplinks



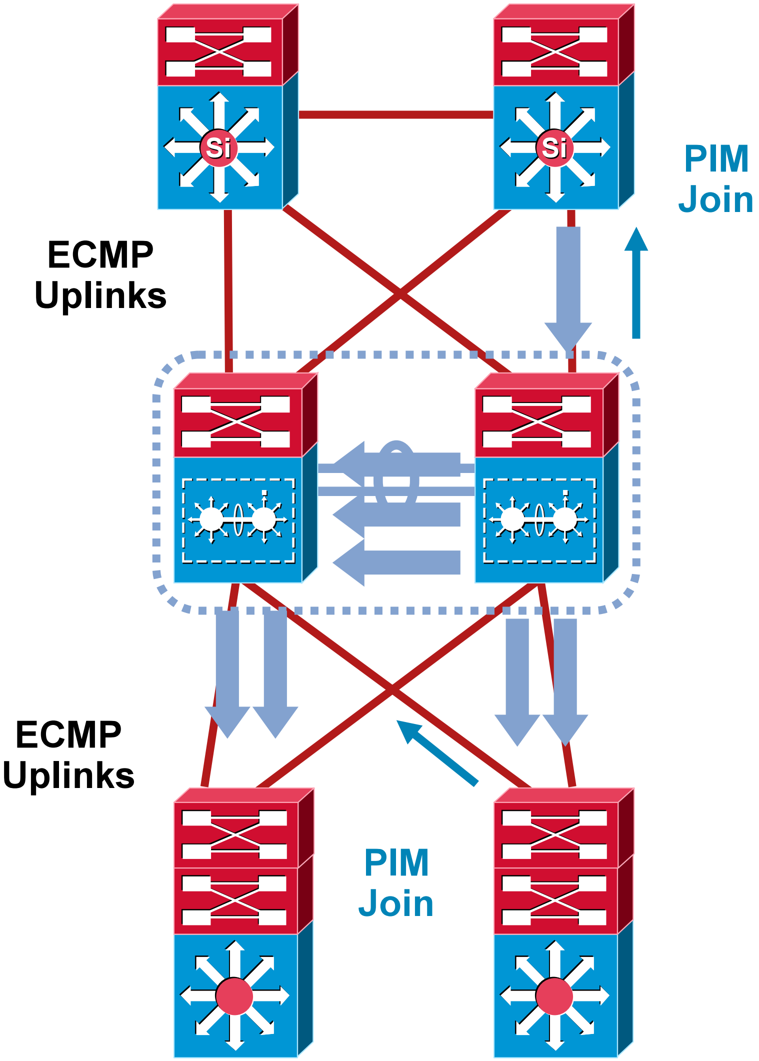


Single Logical Multicast
Designated Router and IGMP Querier

VSS Enabled Campus Design

Multicast Traffic Flows—Layer 3 Access

- In routed access environments use of access to distribution ECMP uplinks can result in multicast traffic forwarded over the VSL links
- VSS represents a single multicast router
- Access PIM joins are sent based on first entry in the routing table out of the two ECMP paths towards the RP
- VSS sends PIM joins upstream on one of it's uplinks
- If the joins are not sent to 'and' from the same physical VSS switch you can get multicast traffic passing across the VSL link



VSS Enabled Campus Design

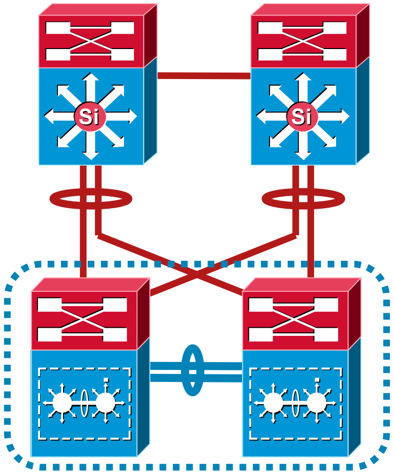
Core Design

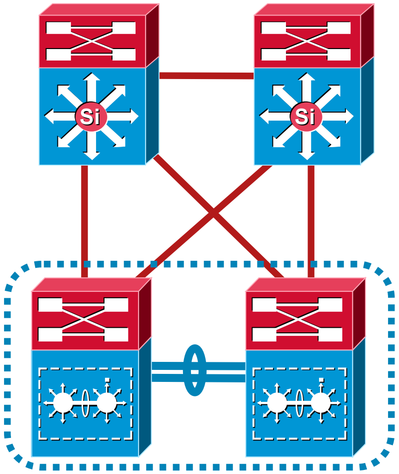
- In a full mesh design two configuration options exist for connecting VSS in the distribution upstream to the core

  4 x ECMP links

  2 MEC links
  (results in 2 x ECMP links)

- Both MEC and HW FIB prefer local links for egress
- Unicast traffic takes the optimal path in both cases (no cross VSL traffic due to the use of one vs. the other)

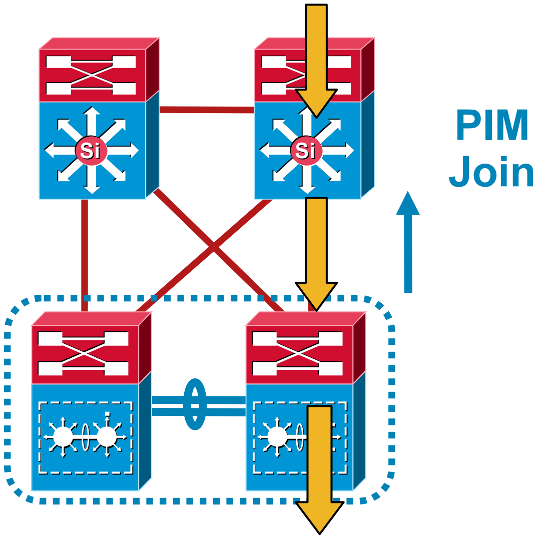


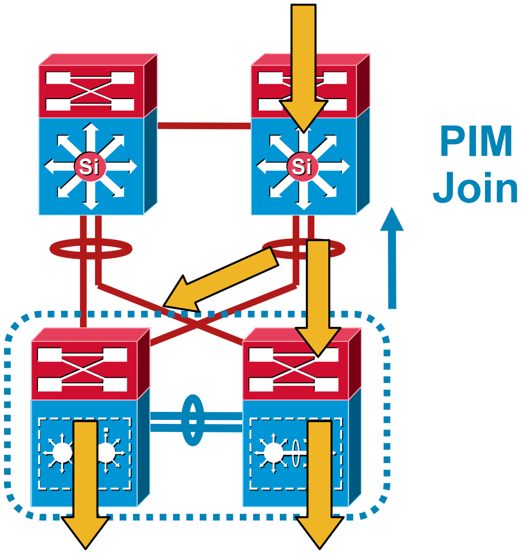


VSS Enabled Campus Design

Core - ECMP & MEC - Multicast Traffic

- PIM joins will be sent on a single L3 path upstream
- In the ECMP configuration, multicast traffic only uses a single link out of four available.
- Traffic takes the optimal path in both cases (no cross VSL traffic due to the use of one configuration vs. the other)
- However, if the PIM join come from core toward the access layer (many to many multicast sources) then MEC to the core is recommended design option
- MEC in the core provides consistent multicast convergence as incoming interface remains the same during the switchover

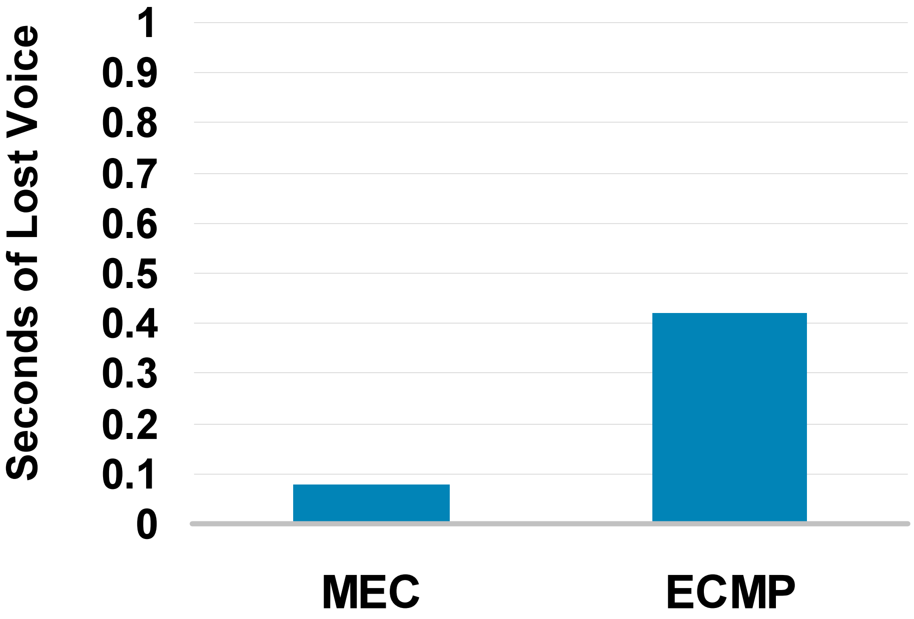


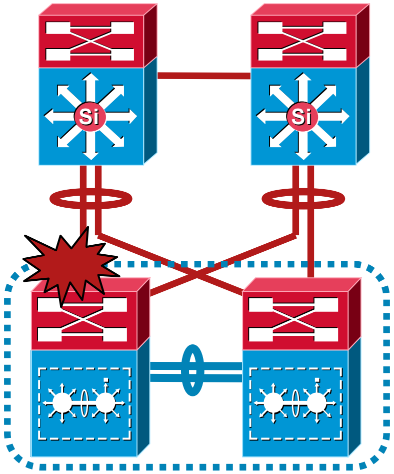


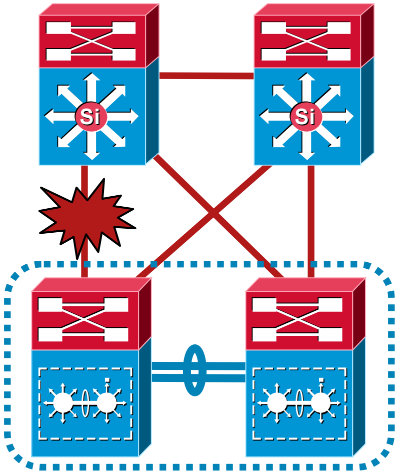
VSS Enabled Campus Design

Core - ECMP & MEC - Link Recovery

- MEC convergence is consistent, independent of the number of routes
- ECMP convergence is dependent on the number of routes



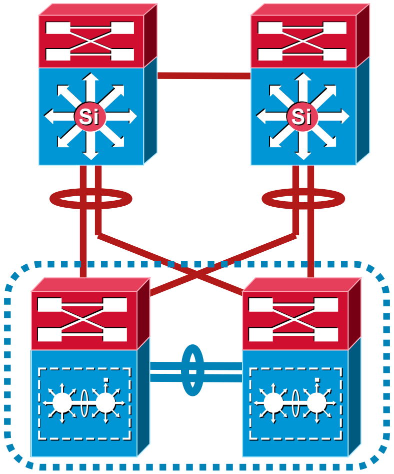


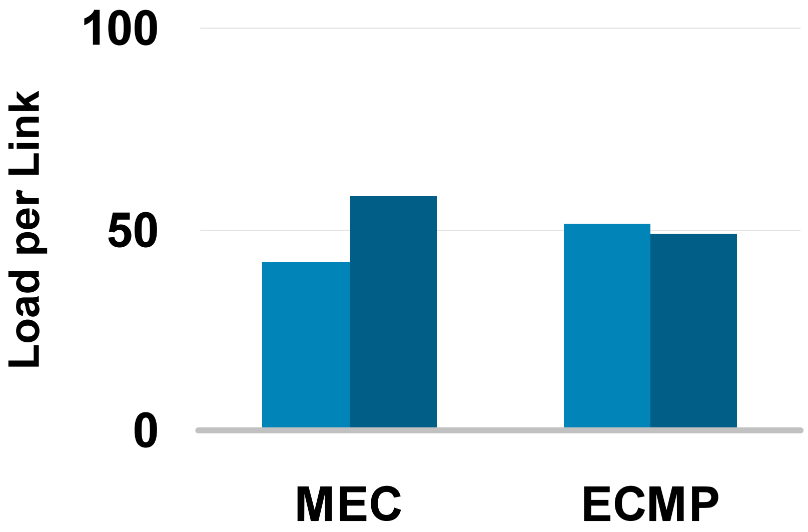


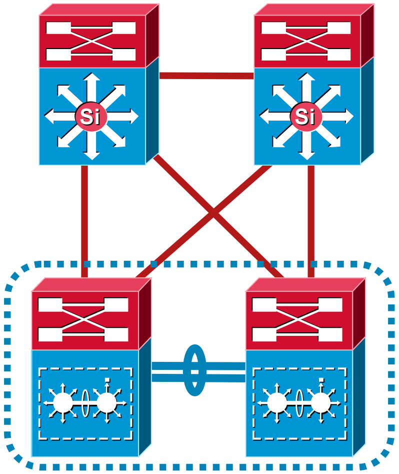
Note: All results are based on pre-FCS code and will be verified and included in upcoming ESE Design Guide

VSS Enabled Campus Design
Core - ECMP & MEC - Load Balancing

- MEC hashes to eight values using a XOR algorithm (Cisco Catalyst 6500 Series)
- ECMP hashes to 16 values and uses a more complex modulo algorithm
- ECMP tends to have better fairness when evaluated against enterprise traffic flows



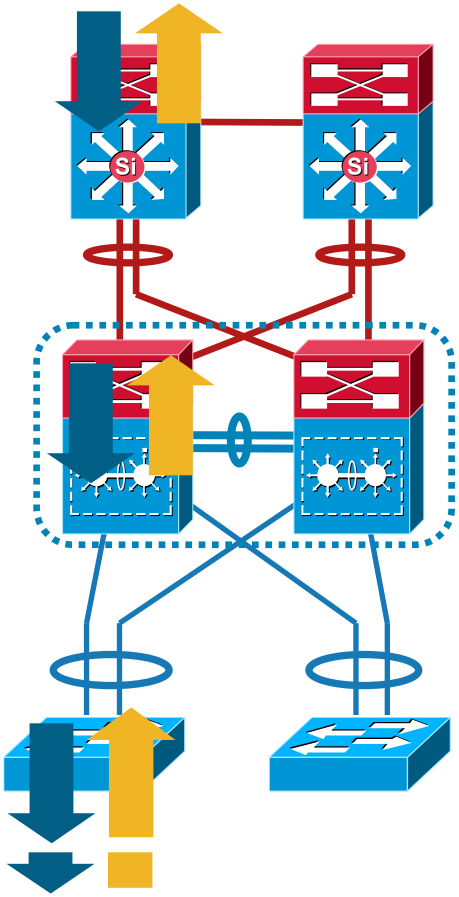




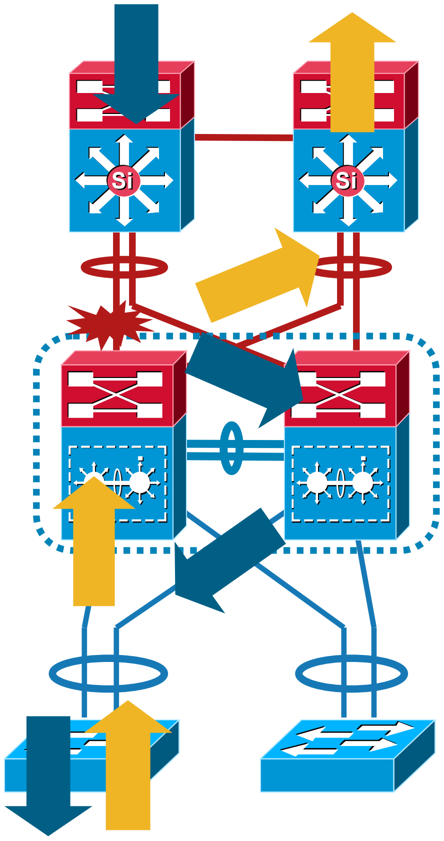
VSS Enabled Campus Design

Failure Recovery

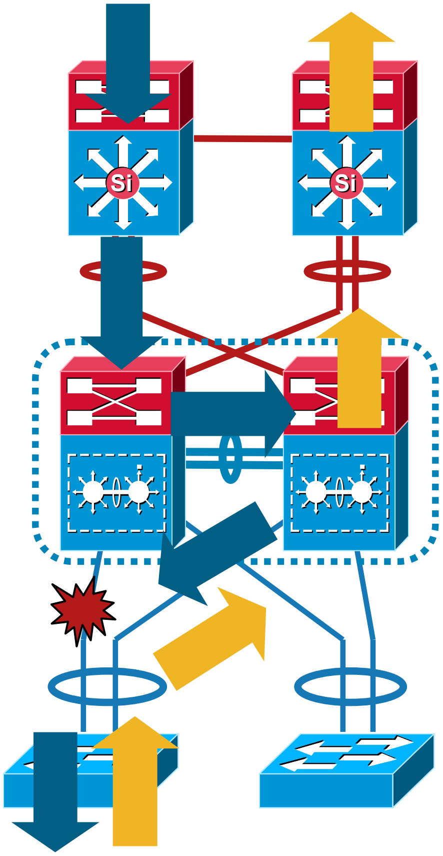
- MEC or ECMP are the primary recovery mechanisms for all link or node failures



VSS Member Failure



Uplink Failure

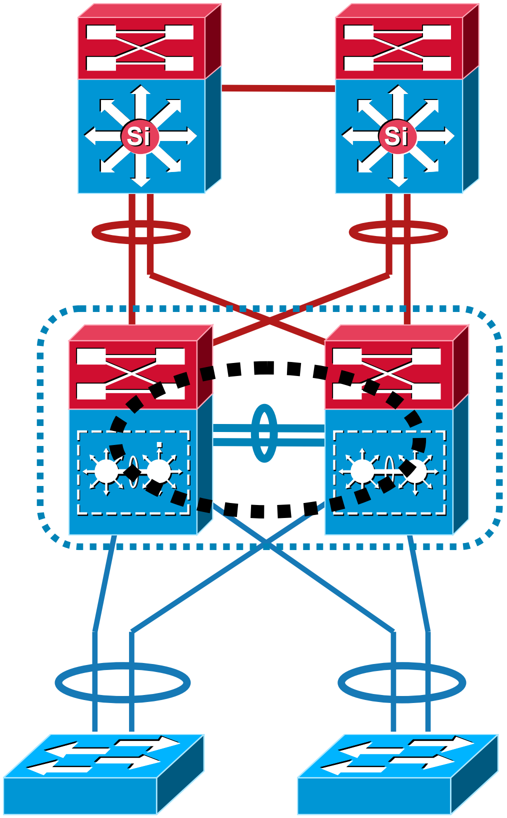


Access link Failure

VSS Enabled Campus Design

Capacity Planning for the Virtual Switch Link

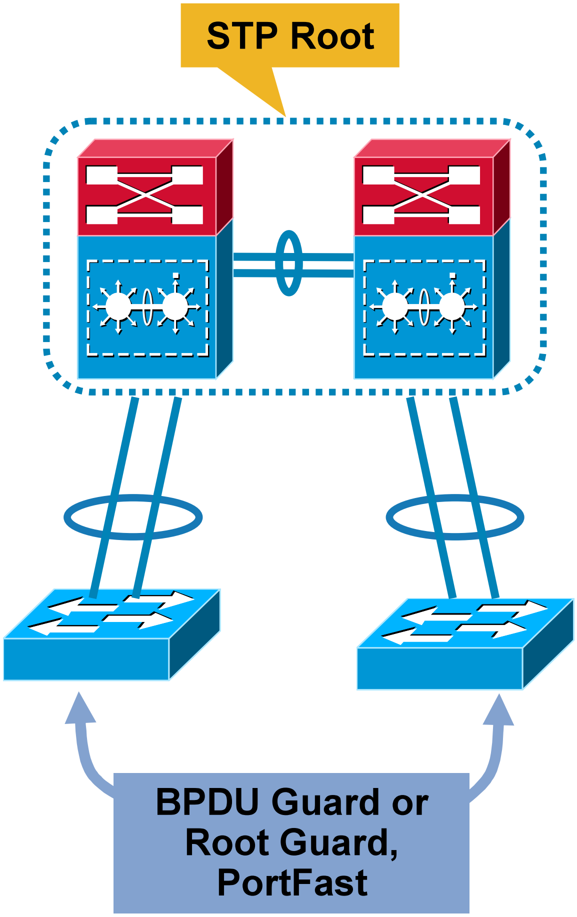
- Capacity planning and link sizing for VSS is almost identical to traditional multilayer design
- The only traffic that should flow across the VSL under normal conditions is control plane traffic
- In an access switch uplink failure half of the downstream traffic will be forwarded across the VSL link
- Control plane load is very small and sent with strict priority over the VSL link
- Redundancy of the VSL is critical and should take priority over capacity planning



VSS Enabled Campus Design

STP Optimization

- Make sure VSS remains root of all VLANs
- Do not use loop guard as it will disable the entire MEC channel on fault detection
- Use root guard at the edge port to protect external switch introducing superior BPDUs, e.g., temporary connectivity
- BPDU guard and root guard are mutually exclusive
- PortFast and BPDU guard is still necessary at the edge switch to prevent accidental loop introduce either due to user error or topology change

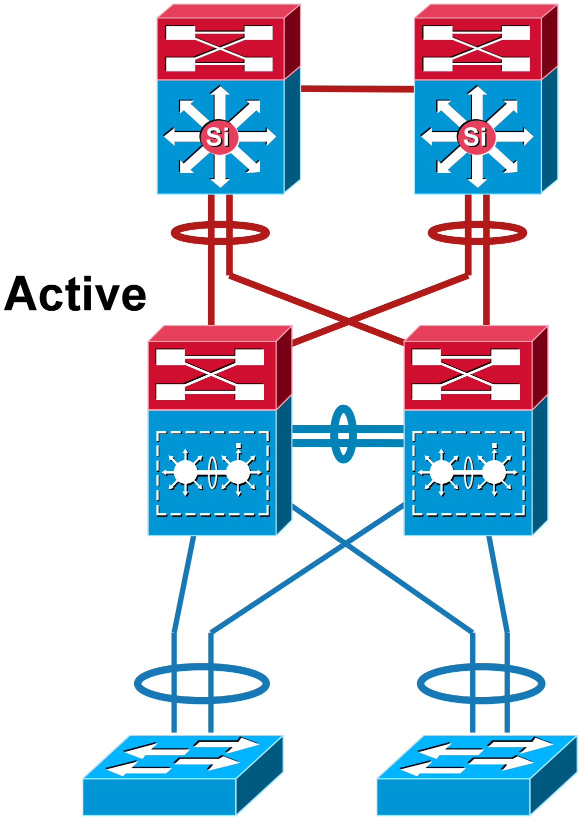


VSS Enabled Campus Design

Summary

- VSS enables highly available campus with sub-second convergence without the complexity of managing dual node at distribution layer
- Eliminates FHRP configuration
- Must use L2 MEC to create loop free topology, STP should remained enabled
- Use of L3 MEC significantly improves convergence for multicast traffic
- Enabled NSF in adjacent routed devices for better convergence

  Use default Hello and Hold timers for EIGRP & OSPF
- Use STP tool kits guidance applicable to loop free “V” shape design



Operational Management and ISSU

Operational Management

Single point of management: Slot/Port Numbering

After conversion, port definitions for switches within the Virtual Switch Domain inherit the Chassis ID as part of their naming convention…

PORT NUMBERING: <CHASSIS-ID><SLOT-NUMBER><PORT-NUMBER>

Chassis-ID WILL ALWAYS be either a “1” or a “2”

```
Router#show ip interface brief
Interface              IP-Address      OK? Method Status                Protocol
Vlan1                  unassigned      YES NVRAM  up                    up
Port-channel1          unassigned      YES NVRAM  up                    up
Te1/1/1                10.1.1.1        YES unset  up                    up
Te1/1/2                192.168.1.2     YES unset  up                    up
Te1/1/3                unassigned      YES unset  up                    up
Te1/1/4                unassigned      YES unset  up                    up
GigabitEthernet1/2/1   10.10.10.1      YES unset  up                    up
GigabitEthernet1/2/2   10.10.11.1      YES unset  up                    up
GigabitEthernet2/1/1   unassigned      YES unset  up                    up
GigabitEthernet2/1/2   unassigned      YES TFTP   up                    up
GigabitEthernet2/1/3   unassigned      YES TFTP   up                    up
Te2/1/4                unassigned      YES TFTP   up                    up
Te2/1/5                unassigned      YES TFTP   up                    up
<snip>
```

Operational Management

Operational Considerations

- Avoid preempt configuration between VSS switches
- Avoid making changes to the configuration during VSS dual active recovery. This will lead to manual syncing of the configuration and reboot
- Beware of SPAN usage

  Avoid replication between chassis which can lead to higher VSL link utilization

  Distributed SPAN requires latest IOS 12.2(33)SXH2(a)
- Reload vs “redundancy force failover”

  Reload causes both VSS chassis to reboot

  Use “redundancy force failover” option to manage both single chassis or dual chassis reboot
- Avoid “write erase” to copy new startup configuration. This will erase switch numbers stored in NVRAM and subsequent reboot will cause both switches to come up in standalone mode. One has to use “switch set switch_num <1/2> command only after both switches are rebooted as the CLI to set switch number is NOT available in VSS mode
- CISCO-VIRTUAL-SWITCH-MIB has been defined to support SNMP access to the Virtual Switching Systeming Configuration - the following MIB variables are accessible to an SNMP manager
- In a Virtual Switching System, with both Data Planes active, Netflow data collection is performed on each Supervisor’s PFC and on each Line cards DFCs - while Netflow export is only performed by VSS Active Supervisor and DFCs that are capable of performing direct export

VSS Software Upgrade, pre 12.2(33)SXI

Preparation Steps

1. Preparation Steps

   a) Ensure the old image and new image files are installed to the local file systems on both Supervisor modules
   b) Configure the boot register to auto-load the specified software image file
   c) Configure the boot string to load the new software image

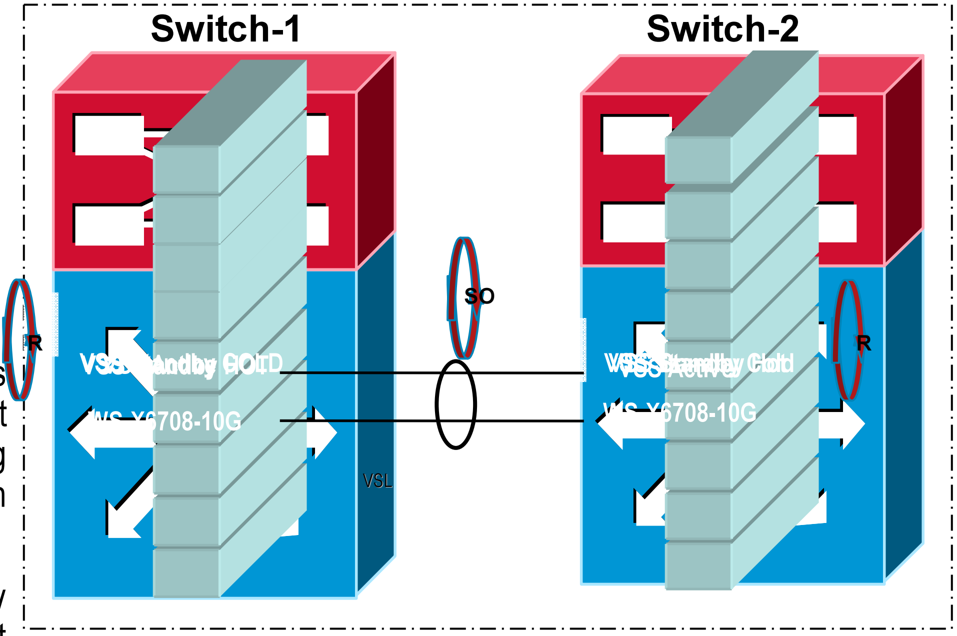
Execute Upgrade

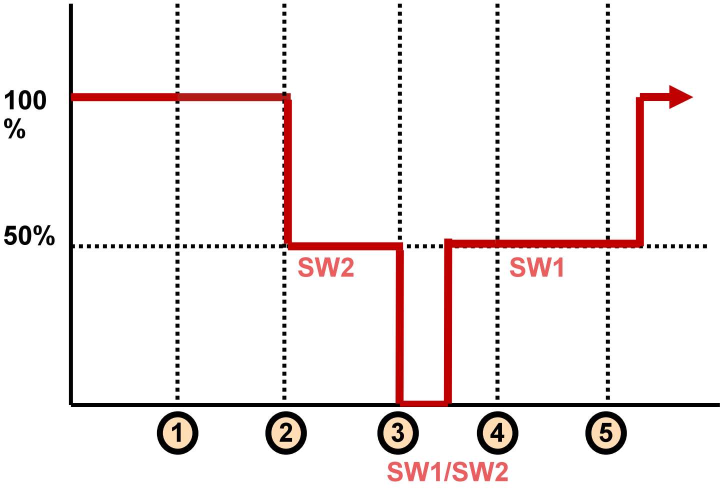
2. Reset the standby Supervisor and ensure it boots successfully to RPR mode (STANDBY COLD). Hot Standby modules are power down and not forwarding traffic at this point , forwarding capacity will be down to 50%

3. Force a Supervisor switchover, forwarding capacity drops to 0%. Standby Supervisor continue to boot and become the new ACTIVE. Old active Supervisor will reset and load the old image and boot to STANDBY COLD (RPR) state

4. Trial Phase

5. Modify boot variable on Switch-1 and reload switch-1 such that it boots up with new software image. Forwarding capacity will resume back to 100%





R = Reset

SO = Switchover

▨ = Old Version

■ = New Version

In Service Software Upgrade process Phase-1

Full Image upgrade

12.2(33)SXI will be The first release of software that will have support for ISSU

Preparation Steps

1. Before ISSU software upgrade, VSS Switch-1 and 2 will be running Old software image. Make sure Active and Standby running a software version that is compatible to perform ISSU

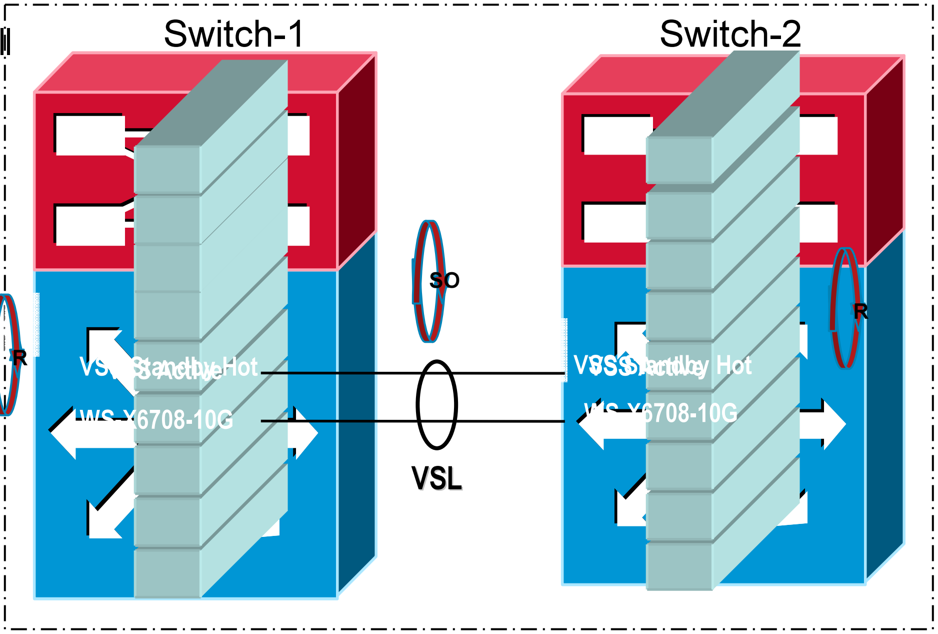
Execute Upgrade

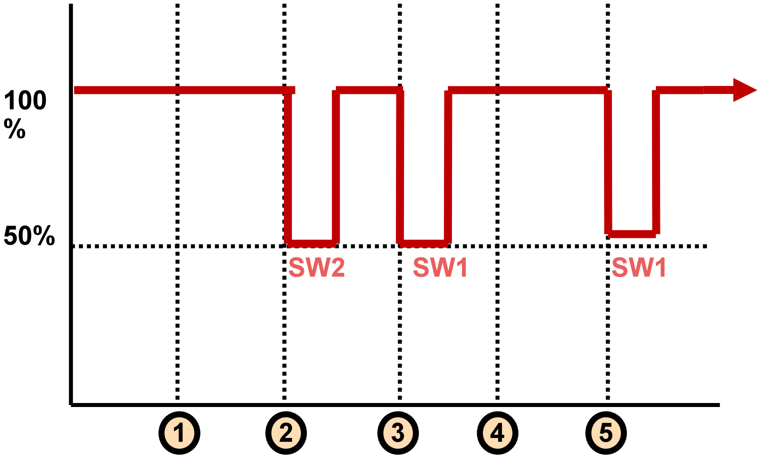
2. ISSU loadversion standby is reloaded to boot new software image and it will be initialized on SSO mode. ISSU infrastructure allows co-existence of different software version now

3. ISSU runversion switchover to the Standby supervisor which is already booted with the new software version ... Old active reloads with old software image and becomes SSO Hot Standby ..

4. ISSU Acceptversion If network is stable issue " ISSU acceptversion " which stops the rollback timer, otherwise ISSU process will aborted intermediately.

5. ISSU Commitversion Once the image is tested and ready to be rolled out .. ISSU commitversion will reload the standby to boot up with new software version





▨ = Old Version

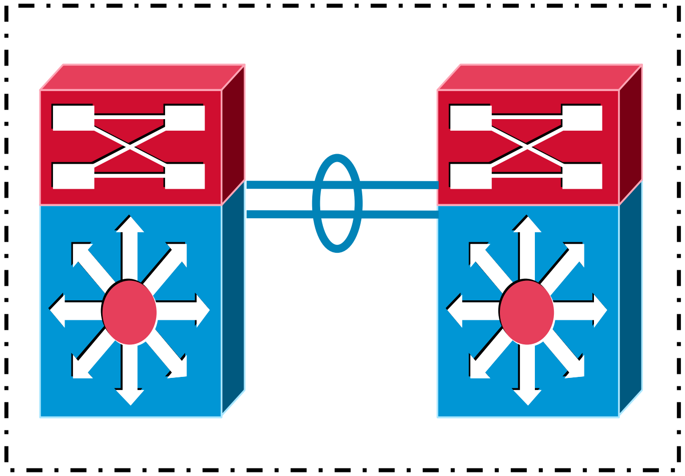
■ = New Version

R = Reset

SO = Switchover

Summary

Virtual Switching System
Summary



Virtual Switching System

Summary

Allows Two Physical Catalyst 6500's to operate as a single logical cat6500 switch.

Provides Single Management point for both switches

Requires supervisor VS-S720-10G-3C or VS-S720-10G-3CXL

Requires 12.2(33)SXH1 or later

Supports WS-X67XX line cards and NAM service module at 12.2(33)SXH1 and FWSM/WISM/ACE/IDSM will be supported at future release called 12.2(33)SXI

Inter-Chassis NSF/SSO used for supervisor failover

Multi-chassis Ether channel provides new benefits for STP elimination and improved resiliency

Dual Active Recovery Mechanisms for VSL Failure

Q and A