Virtual Switching System (VSS) Best Practices
Session Objectives

Virtual Switching Systems Best Practices

- Understand Multilayer Campus Design and Challenges for Unified Campus
- Key Benefits of Virtual Switching System
- Understanding key VSS components and best practices
- VSS enabled Campus Design
  - Multilayer Design
  - Core and Routed Access Design
- Benefits with VSS
  - High Availability, Convergence, Capacity, Reduced complexity
Age old problem

- Remember old days of flexibility, add/move and mobility promise?
- Spanning VLAN solved that and created some more problems of
  Stability
  Response times
  Inefficient use of resources
  Managing end host behavior
Solution – Oh well compromise!

- Do not span VLANs.
- No Loops no underlying threat to the network
- Solution gave up critical need of not able to span VLANs.
- Indirect consequence were, unless a change imposed by application/technology/business most legacy network remained looped spanning tree based topology
History Repeats – well with some changes

- Solving the same OLD design problem and yet not lose the benefits of stability and mobility
- Virtual Switching allows elimination of loops in the network
- Now its possible to have a cake and eat it too 😊
Catalyst 6500 VSS1440 Key Benefits

- **Maximizes Bandwidth Utilization**
  - Maximize system usage
  - Maximize server usage
  - NIC standardization

- **Simplifies Operational Manageability**
  - Reduce 50% of Managed Nodes
  - Loop-free topology
  - LMS 3.0 integration

- **Boosts Non-Stop Communications**
  - Optimized path selection
  - Increased throughput
  - Deterministic sub-sec network recovery
  - Business continuity with no service disruption

- **Lowers Latency**
  - Loop-free topology

- **Simplifies Operational Manageability**
  - Reduce 50% of Managed Nodes
  - Loop-free topology
  - LMS 3.0 integration
Alternative Campus Design

Agenda

- Multilayer Campus Design
- Virtual Switching Systems
  - VSS Components
    - SSO
    - VSL Link
    - MEC (Multi-Chassis EtherChannel)
  - VSS Recovery
- VSS Enabled Campus Design
  - EtherChannel Optimization
  - VSS Traffic Flow
  - Multilayer Campus Design
  - Routing (core and Routed-access)
  - Convergence
- Summary
Virtual Switch
Virtual Switching System 1440 (VSS)

- Virtual Switching System consists of two Cisco Catalyst 6500 Series defined as members of the same virtual switch domain
- Single control plane with dual active forwarding planes
- Design to increase forwarding capacity while increasing availability by eliminating STP loops
- Reduced operational complexity by simplifying configuration
Virtual Switching System
Single Control Plane

- Uses one supervisor in each chassis with inter-chassis Stateful Switchover (SSO) method in which one supervisor is **ACTIVE** and the other in **HOT_STANDBY** mode.
- Active/standby supervisors run in synchronized mode (boot-env, running-configuration, protocol state, and line cards status gets synchronized).
- **ACTIVE** supervisor manages the control plane functions such as protocols (routing, EtherChannel, SNMP, telnet, etc.) and hardware control (OIR, port management).
- Switchover to STANDBY_HOT supervisor occurs when **ACTIVE** supervisor fails, providing subsecond protocol and data forwarding recovery.
Virtual Switching System
VSS Domain and Switch Priority

- Domain ID is used to identify that two switches are intended to be part of the same VSS pair.
- Switch ID is assigned during initial phase of VSL link configuration. Switch with lowest number (switch ID) always chosen as an ACTIVE switch when both member switches are booting together.
- The order of which switch become ACTIVE can be changed via switch priority command. This is useful in two cases:
  - To change the order of which switch can become ACTIVE after initial configuration.
  - One can configure the priority matching the switch ID so that configuration file showing the priority order helps in change management.
- One can explicitly decide that only designated switch becomes ACTIVE via switch pre-emption configuration. However this will lead to multiple reboot when the ACTIVE switch recovers, forcing the existing ACTIVE switch to reboot and go into Hot_Standby mode.
- ACTIVE switch role do not transition from low priority switch to high priority switch unless switch pre-emption on high priority switch is configured.
- Since both VSS switch member are considered as single logical router (single control plane, dual forwarding plane), it does not matter which switch must become ACTIVE.
- Recommendation is to NOT to configure switch pre-emption.

```
cr2-6500-vss#sh run
...
switch virtual domain 10
switch mode virtual
switch 1 priority 110
switch 2 priority 110
```

```
cr2-6500-vss#show switch virtual
Switch mode : Virtual Switch
Virtual switch domain number : 10
Local switch number : 2
Local switch operational role : Virtual Switch Active
Peer switch number : 1
Peer switch operational role : Virtual Switch Standby
```
Virtual Switching System
VSL—Virtual Switch Link

- VSL (Virtual Switch Link) provides two functions
  - Control plane extension and enables synchronization of protocol states and table
  - Data forwarding when needed
- VSL is treated as system links thus many user level protocol and capabilities are restricted e.g., IP address, flow control, CoS queuing only (no changes) etc.
- VSL can only be defined with 10 Gig port on either Sup720-10G or WS-X6708
- VSL is defined by unique port-channel interface on each switch

```
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

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
```
Virtual Switching System
Resilient VSL Configuration

- Protecting VSL bundle is of the highest priority. VSL bundle is a special purpose EtherChannel however all the best practices of designing and configuring of any general EtherChannel applies to VSL bundle.

- Redundancy of VSL is important to avoid dual ACTIVE condition and instability of VSS thus not to be confused with an idea that if the supervisor fails, then what good is the VSL link redundancy – well in that case the HOT_STANBY is ready to take over and has nothing to do with VSL link resiliency. The VSL resiliency is important when supervisor is NOT failed and some how systems lost VSS links, leading to dual ACTIVE condition.

- Diversify VSL bundle on two separate hardware just like any resilient EtherChannel design. VSL link hardware selection also affect the QOS configuration on the rest of the ports on supervisors.

- Design 1 is most common and intuitive choice, which uses both 10G ports on supervisor which is does not provide hardware diversity as both ports are connected to single internal fabric connection. Diversifying VSL links avoids dual ACTIVE condition. Thus assumption that if one put all VSL links on supervisor card is OK is misleading.

- Design 2 offers a practical solution where the cost is more influential then flexibility. Design 2 uses one port from Sup720-10G uplink and other from WS-X6708 line card. In this configuration the unused ports on Sup720-10G can only be configured with default CoS based queuing. 12.2(33)SXHI will allow diverse QOS configuration on the unused 10G port of Sup720-10G.

- Design 3 is recommended if full flexibility of QOS configuration is required. This option avoids using Sup 720-10G uplink ports and uses two separate WS-X6708 line cards for VSL links but not as cost effective as option 2 or 1.
Virtual Switching System
Link Management Protocol (LMP)

- 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:
  - Verify neighbor is bidirectional
  - Ensure the member is connected to another virtual switch
  - Transmit and receive LMP keepalives (timers) to maintain health of the member and the VSL

- LMP timers, also referred as VSLP timers plays a key role to maintain the integrity of VSS by way to checking peer switch connectivity. Both VSS member takes independent yet deterministic SSO switchover action if they fails to detect LMP hello message within hold-timer settings on last bundled VSL link

- By default LMP hello transmit timer(T4) and receive (min_rx) timer is 500 msec. The hold-timer known as T5 timer derived from a min_rx and default multiplier of 120. Thus by default VSL member link time out is detected in 60 second.

- LMP(VSLP) timer plays secondary role in determining the VSL link integrity and failover. The VSL link hardware is by default enhanced with Fast Link Notification (FLN) feature which gives a failure detection of the order of 50-100 msec., which is triggered much earlier then LMP timer expiry.

- Modifying default timer to a shorter value does not improve convergence of VSL link failure detection in most cases. In addition, it adds significant instability such that in most cases VSL will fails to establish neighbor relation between member switches, leading to continuous crash of hot_standby switch.

- It’s is strongly recommended NOT to modify default LMP(VSLP) timer.
Virtual Switching System
Role Resolution Protocol (RRP)

- RRP runs on control link of the VSL bundle
- Determines whether hardware and software versions allow a virtual switch to form
- Determines which chassis will become Active or Hot Standby from a control plane perspective

<table>
<thead>
<tr>
<th>Switch Number</th>
<th>Switch Status</th>
<th>Preempt</th>
<th>Priority</th>
<th>Role</th>
<th>Session ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOCAL</td>
<td>1</td>
<td>UP</td>
<td>FALSE(N )</td>
<td>ACTIVE</td>
<td>0</td>
</tr>
<tr>
<td>REMOTE</td>
<td>2</td>
<td>UP</td>
<td>FALSE(N )</td>
<td>STANDBY</td>
<td>9924</td>
</tr>
</tbody>
</table>

RRP runs on control link of the VSL bundle. It determines whether hardware and software versions allow a virtual switch to form. It also determines which chassis will become Active or Hot Standby from a control plane perspective.
Virtual Switching System
Multichassis EtherChannel (MEC)

- MEC is an advanced EtherChannel technology extending link aggregation to two separate physical switches.
- MEC enables the VSS appear as single logical device to devices connected to VSS, thus significantly simplifying campus topology.
- Traditionally spanning VLANs over multiple closets would create STP looped topology, MEC with VSS eliminates these loops in the campus topology.
- MEC replaces spanning tree as the means to provide link redundancy and thus doubling bandwidth available from access.
- MEC is supported only with VSS.
Virtual Switching System
Dual Active

- VSL is the heart of the VSS functionality
- Protecting VSL link bundle is the best practice design
  - Use one port from Supervisor and other from line cards to form a VSL bundle
  - Use diverse fiber path for each VSL links
  - Manage traffic forwarded over VSL link by avoiding single homed devices
- In case of loss of all members of the VSL bundle, the standby supervisor will go active, creating dual active condition
- Dual active leads to
  - Two independent routers with same control plane information e.g. IP address, router ID etc.
  - Error disabling of access-layer due to two STP BPDU sent with different source MAC
- Three mechanism to provide dual active state detection and recovery
  - Enhanced PAgP
  - BFD
  - Fast VSLP Hello (supported in 12.2(33)SXHI)
### High Availability Dual-Active Detection

<table>
<thead>
<tr>
<th></th>
<th>ePAgP</th>
<th>Dual-Active IP BFD</th>
<th>Dual-Active Fast Hello</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Communication Medium</strong></td>
<td>EtherChannel L2 or L3</td>
<td>Layer-3</td>
<td>Layer-2</td>
</tr>
<tr>
<td><strong>Detection</strong></td>
<td>Requires ePAgP compatible neighbor to detect ePAgP based detection</td>
<td>Direct-L3-Link, BFD session triggers the detection</td>
<td>Direct-L2-Link</td>
</tr>
<tr>
<td><strong>Software release</strong></td>
<td>65xx - 12.2(33)SXH1 45xx - 12.2(44)SG 2960/3560/3750 – 12.2(46)SE No support for 3750 stack (PAgP not supported on stack)</td>
<td>12.2(33)SXH1</td>
<td>12.2(33)SXI</td>
</tr>
<tr>
<td><strong>Recovery</strong></td>
<td>Sub-second based on routing protocol and topology</td>
<td>Most in seconds. Sub-second only in some cases</td>
<td><strong>Sub-seconds</strong></td>
</tr>
</tbody>
</table>
Virtual Switching System

Dual Active—Enhanced PAgP

- Enhanced PAgP provides a new TLV which uses MAC address of an active switch as an ID to identify dual active detection. 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 generates its own ePAgP message with its own ID via ePAgP supporting neighbor to switch 1.

Normal Mode

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

Dual Active Detection

%PAGP_DUAL_ACTIVE-SP-3-RECOVERY_TRIGGER: PAgP running on Gi6/1 informing virtual switches of dual-active:
new active id 0019.a927.3000, old 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 requires manual intervention of either “reload” or config-sync.
Virtual Switching System
Enhanced PAgP Support and Configuration

- ePAgP requires PAgP protocol to be operational on MEC enabled interfaces
- PAgP is supported in all platforms except 37xx stack configuration where cross-stack EtherChannel (LACP) is required to have MEC connectivity with VSS. Standalone 35xx and 37xx do support PAgP and thus ePAgP.
- Since cross-stack EtherChannel does not support PAgP, it can not use enhanced PAgP for dual ACTIVE detection. Hence the only alternative for cross-stack EtherChannel configuration is to use BFD as a dual ACTIVE detection method
- ePAgP can be enabled either on L2 or L3 PAgP MEC members. This means one can run ePAgP between VSS and core routers (though core routers require proper IOS) in case when access-layer only have 37xx stack configuration
- ePAgP is supported on a specific software version on each platform:

<table>
<thead>
<tr>
<th>Platform</th>
<th>Software</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>6500</td>
<td>12.2(33)SXH</td>
<td>Sup720 and Sup32</td>
</tr>
<tr>
<td>45xx and 49xx</td>
<td>12.2(44)SG</td>
<td></td>
</tr>
<tr>
<td>29xx, 35xx &amp; 37xx</td>
<td>12.2(46)SE</td>
<td>37xx Stack – no support</td>
</tr>
</tbody>
</table>

```
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

<table>
<thead>
<tr>
<th>Port</th>
<th>Dual-Active</th>
<th>Partner</th>
<th>Partner</th>
<th>Partner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gi1/8/19</td>
<td>Yes</td>
<td>cr7-6500-3</td>
<td>Gi5/1</td>
<td>1.1</td>
</tr>
<tr>
<td>Gi1/9/19</td>
<td>Yes</td>
<td>cr7-6500-3</td>
<td>Gi6/1</td>
<td>1.1</td>
</tr>
</tbody>
</table>
```
Virtual Switching System
Dual Active Recovery—BFD

- Utilizes a direct pt-pt link connected to an interface on each switch
- Must have a unique IP subnet on each end of the link; BFD session becomes operational only after VSL link failure
- BFD session establishment triggers the dual active conditions and previously active switch undergoes to recovery mode similar to PAgP detection
- Only use when ePAgP detection method is not available; replaced by better detection in next major release of Cisco IOS software

```plaintext
interface gigabitethernet 1/5/1
no switchport
ip address 192.168.1.1 255.255.255.0
bfd interval 100 min_rx 100 multiplier 3
interface gigabitethernet 2/5/1
no switchport
ip address 192.168.2.1 255.255.255.0
bfd interval 100 min_rx 100 multiplier 3
switch virtual domain 100
dual-active pair interface gig 1/5/1 interface gig 2/5/1 bfd

Console Message:
adding a static route 192.168.1.0 255.255.255.0 Gi2/5/1 for this dual-active pair
adding a static route 192.168.2.0 255.255.255.0 Gi1/5/1 for this dual-active pair
```

Requires Unique IP Subnets on the Two Interfaces
Dual Active

EIGRP Convergence – with Default Timers

- The impact on user traffic is more important than how long each method detects the dual active condition
- ePAgP
  - Both ECMP and MEC based core design give sub-second convergence of user traffic
- BFD
  - The ECMP core design provides same convergence as ePAgP
  - The recovery with MEC based core is more complex. The MEC core design has higher losses due to destabilization of EIGRP adjacency leading to removal routes affecting both upstream and downstream convergence
Dual Active
OSPF Convergence – with Default Timers

- The impact on user traffic is more important than how long each method detects the dual active condition.
- OSPF interaction is more complex and variable due to the nature of OSPF protocol and unique router ID.
- ePAGP
  - ECMP based core design has higher losses due to the fact that OSPF remove access-layer routes in the core during dual ACTIVE.
  - MEC based core design does not suffer from route removal in the core as no OSPF routes withdrawal are sent and ePAGP detection is faster leading to sub-second convergence.
- BFD
  - The ECMP based core design has better convergence compared to ePAGP due to the delayed recovery action by BFD leaving at least one access-layer route in core.
  - The recovery with MEC based core is more complex. The MEC core design has higher losses due to destabilization of adjacency leading to removal routes affecting both upstream and downstream convergence.
  - In addition downstream losses are more due to the fact that core remove the routes for access-layer subnet faster as it detects the adjacency loss, compared to VSS which retains the adjacency and thus the upstream routes.

* The MEC BFD convergence is variable.
Dual Active Recovery
Multiple Mechanisms

- Ensuring the availability of the VSL link is a high priority. Redundant fiber paths recommended to protect against physical fiber failures
- Deploy ePAgP where possible to get the best convergence. BFD will be replaced in next release to have substantially better convergence
- ePAgP can be run either on L2 or L3 MEC
- ePAgP only needs to be run on a single neighbor however, leveraging enhanced PAgP on all interfaces will ensure that in the worst case at least one switch (assuming that not all cable paths are affected in the failure condition) is connected to both members of the same VSS pair then a path will exist for the recovery
- Testing with 30 ePAgP neighbors shows negligible impact on VSS CPU
Summary

- User traffic impact during dual ACTIVE is primarily due to three factors
  - Detection technique deployed
  - The interaction of routing protocol
  - Connectivity from VSS to the core
- ePAgP itself is NOT sufficient to get sub-second convergence
- ePAgP detection with L3 MEC based connectivity to the core gives sub-second convergence for both OSPF and EIGRP
- BFD do provide sub-second convergence however in limited topology
- With BFD detection, EIGRP & OSPF recovery is complex and variable when core connectivity from VSS is L3 MEC based
- For OSPF only one configuration gives sub-second convergence - ePAgP and MEC based connectivity to core. The rest of the combination the convergence is variable
- L2 BFD verification – pending, so final recommendation subject to change
Virtual Switching System
Base Configuration

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

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

Switch 1

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

Switch 2

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

MEC

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
Alternative Campus Design

Agenda

- Multilayer Campus Design
- Virtual Switching Systems
  - VSS Components
    - SSO
    - VSL Link
    - MEC (Multi-Chassis EtherChannel)
    - VSS Recovery
- VSS Enabled Campus Design
  - EtherChannel Optimization
  - VSS Traffic Flow
  - Multilayer Campus Design
  - Routing (core and Routed-access)
  - Convergence
- Summary
VSS Enabled Campus Design
Impact to the Campus Topology

- Physical network topology does not change
  - Still have redundant chassis
  - Still have redundant links

- Logical topology is simplified as we now have a single control plane
  - No unicast flooding
  - Single configuration management

- Allows the design to replace traditional topology control plane with multichassis EtherChannel (MEC)
  - Enables loopfree topology, thus doubles the link capacity
  - Convergence and load balancing are based on EtherChannel
VSS Enabled Campus Design
MEC Configuration

- 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
VSS Enabled Campus Design
Traffic Flow During Failure

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

[Diagram showing traffic flow during different types of failures: VSS Member Failure, Uplink Failure, Access link Failure.]
VSS Enabled Campus Design
Capacity Planning for the Virtual Switch Link

- In normal condition, control plane load over the VSL link is very small and sent with strict priority over the VSL link.
- Capacity planning and link sizing for VSS is almost identical to traditional multilayer design.
- Two failure points determines bandwidth requirements:
  - Core uplink failure from one switch
  - Downstream link to access-switch failures from one switch.
- In both the cases normal traffic carrying capacity from each switch is determined by links connected to the core from each switch, since each switch can only forward traffic from locally connected interfaces. Thus VSL links should at least carry total bandwidth offered via uplinks from one member switch. This assures the same level of bandwidth as uplink.
- If line card connected to several access-switch fail, one may have over-subscription on downstream but no greater then core uplink capacity from the single member switch.
- The recommendation is to at least have VSL bandwidth equal to uplinks connected to a single physical switch.
- Additional VSL link capacity planning is required for traffic carried by:
  - Single homed devices
  - Remote SPAN from one switch to another and (IDS) Intrusion detection Systems. For both the cases, the traffic hashes only over a single VSL link which can lead to oversubscription of a particular link. The only way to improve the probability of distributing traffic is to have additional VSL link as RBH calculation increases the chance that traffic that hashed along with RSPAN may now choose different link
  - Services traffic – FWSM, WiSM etc

- Always bundle VSL EtherChannel in power of 2 as it will have better hash results for optimized traffic load-sharing.
- Redundancy of the VSL is still critical along with resiliency of VSL links.
VSS Enabled Campus Design
Control Plane Simplification with Multi-layer Design

- VSS makes the network loop-free in normal topology; do not disable spanning tree to safeguard against possible loop introduced at the edge due to user error and daisy chaining.
- Simplifies the topology allowing VLANs to span to increase flexibility in design options supporting various business applications.
- Ease of implementation, less to get it right
  - No need for HSRP, GLBP, or VRRP
  - No reliance on subsecond FHRP timers
  - No asymmetric forwarding
- A Redundant supervisors provide resiliency via SSO-enabled protocols; consistent recovery during the failover of nodes at the distribution.
VSS Enabled Campus Design
Multilayer Topology

- Optimized multilayer topology uses “V” shape design where VLANs do not span closets
- Deploying VSS in such topology without MEC reintroduces STP loops in the networks
- Use of MEC is recommended any time two L2 links from the same devices connected to VSS

Each access switch has unique VLANs
No layer 2 loops
No blocked links

Layer 2 Loop Blocking One Link
MEC Creates Single Logical Link, No Loops, No Blocked Links
VSS Enabled Campus Design
Daisy Chaining Access Layer Switches

- Daisy chained access switch designs challenges
  - Unicast flooding
  - Loop—blocked link
- The use of a virtual switch in the distribution does address the problem of unicast flooding
- You still have a Layer 2 loop in the design with an STP blocked link
- Traffic recovery times are determined by spanning tree recovery in the event of link or node failures
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

#### VSS Switchover Convergence

- **Validated with ESE Campus network environment:**
  - 70 access-layer switches – aka 70 MEC
  - 8 VLANs Spanning multiple switches
  - Total 4K MAC – 720 MAC/VLANs
  - NSF aware adjacent node
  - Default EIGRP & OSPF Timers
  - Native IOS 12.2(33) SXH2a

- **Switchover from ACTIVE to HOT_STANDBY chassis is sub-second, without the complexity of existing design options**

- **L3 to core (MEC or ECMP) does not impact the switchover convergence, though NSF aware core improves downstream convergence**

- **L2 MEC – Access Layer**
  - Average convergence for 37xx and 45xx is 200 msec

#### VSS ACTIVE to HOT_STANDBY switchover convergence (Unicast)

<table>
<thead>
<tr>
<th></th>
<th>Max</th>
<th>Min</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>VSS</td>
<td>0.44</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>HSRP Sub-second Timers</td>
<td>0.9</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>HSRP Default</td>
<td>10</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

![Diagram showing active and hot standby routers connected](image-url)
VSS Enabled Campus Design
NSF – L3 High Availability

- 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 capable and awareness

Recommendation is to not tune IGP hello timers, use default Hello and Dead timers for EIGRP/OSPF in a VSS environment
Modular IOS may require larger OSPF dead timers tuning
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 L3 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 Routed Access Design
Control Plane Simplification & Enhanced Recovery

- Simplified configuration – less to get it wrong
- Simplifies the topology – single logical router end to end - core, distribution and access. Reduced neighbor counts and links state topology as single router interaction at each layer
- No sub-second timer configuration required in core and distribution since access-layer link failure does not require OSPF SPF re-calculation. This simplifies configurations and removes the topological dependencies of extending the timers to non-campus devices
- In best practices design, summarization is still recommended to reduce the control plane instability. However due to single logical router with MEC at the distribution eliminates control plane activity beyond VSS. This leads to the reduced need of summarization. For the network that does not have proper IP addressing plan can benefits the advantage of deploying VSS in routed access design
- Redundant supervisors provide resiliency via SSO-enabled protocols; consistent recovery during the failover of nodes at the distribution. e.g. OSPF or EIGRP NSF/SSO eliminates the dependency of convergence on number of routes in the core and distribution
- A single logical multicast router in core and distribution simplifies the multicast topology resulting in sub-second convergence in core and distribution for a nodal failure
Routed Access

Summary

- EIGRP converges in < 200 msec
- OSPF with subsecond tuning converges in < 200 msec
- Ease of implementation, less to get right
  - No matching of STP/HSRP/GLBP priority
  - No L2/L3 multicast topology inconsistencies
- Single control plane and well known tool set
  - traceroute, show ip route, show ip eigrp neighbor, etc.
- Convergence times dependent on GLBP/HSRP tuning

Both L2 and L3 Can Provide Subsecond Convergence
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
VSS Enabled Campus Design
Operational Considerations

- Avoid preempt configuration between VSS switches
- Use default VSLP(LMP) timers for VSL link configuration
- 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
- Network management – develop baseline what is acceptable polling parameters, since total number of ports in a single chassis has doubled, which can lead to higher CPU
## Convergence Benefits
### VSS vs Multilayer – EIGRP & RPVST+

<table>
<thead>
<tr>
<th>BP Multilayer</th>
<th>Upstream Recovery</th>
<th>Downstream Recovery</th>
<th>Recovery Mechanisms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Active HSRP Distribution Switch Failure</strong></td>
<td>800 msec</td>
<td>200 msec</td>
<td>Upstream—802.1w*, Downstream—802.1w, L3 Equal Cost Path and ARP</td>
</tr>
<tr>
<td><strong>Active or Standby HSRP Distribution Switch Restoration</strong></td>
<td>180 msec</td>
<td>180 msec – 6 sec</td>
<td>Upstream—802.1w, Downstream—802.1w, L3 Equal Cost Path and ARP</td>
</tr>
<tr>
<td><strong>Uplink Fiber Fail to Active HSRP</strong></td>
<td>900 msec</td>
<td>700–1100 msec</td>
<td>Upstream—HSRP, Downstream—EIGRP</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VSS*</th>
<th>Upstream Recovery</th>
<th>Downstream Recovery</th>
<th>Recovery Mechanisms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Active VSS Switch Failure</strong></td>
<td>150- 400 msec</td>
<td>&lt; 200 msec</td>
<td>Upstream—SSO &amp; EtherChannel, Downstream—SSO &amp; ECMP</td>
</tr>
<tr>
<td><strong>Standby VSS Switch Restoration</strong></td>
<td>150- 600 msec**</td>
<td>&lt; 200 msec</td>
<td>Upstream—SSO &amp; EtherChannel, Downstream—ECMP</td>
</tr>
<tr>
<td><strong>VSL Link Failure – Dual Active</strong></td>
<td>200- 400 msec</td>
<td>200- 400 msec</td>
<td>Upstream— EtherChannel, Downstream—ECMP</td>
</tr>
</tbody>
</table>

* For looped topology, 802.1w related losses are higher.
** Linecard failure might have higher losses
VSS Enabled Campus Design
End-to-End VSS Design Option

STP-Based Redundant Topology

Fully Redundant Virtual Switch Topology

\( \text{B} = \text{STP Blocked Link} \)