CHAPTER 2

Virtual Switching System 1440 Architecture

This chapter addresses the architecture and components of Cisco Catalyst 6500 Series Virtual Switching System (VSS) 1440. Although this design guide focuses on the deployment specifics of the VSS (and not technology itself), sufficient detail is included for all the necessary components of the VSS that can affect campus design. This includes operational characteristics, design tradeoffs, and best-practice configuration recommendations. For further details about VSS technology, refer to the following document:

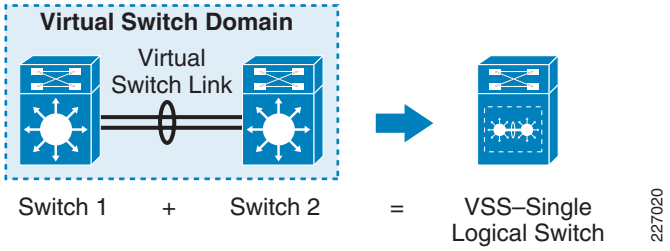
Cisco Catalyst 6500 Series Virtual Switching System (VSS) 1440 White Paper on VSS technology:

http://www.cisco.com/en/US/prod/collateral/switches/ps5718/ps9336/white_paper_c11_429338.pdf

VSS Architecture and Operation

The VSS forms two Cisco Catalyst 6500 switches into a single, logical network entity. Two chassis combine to form a single *virtual switch domain* that interacts with rest of the network as single logical switch and/or router. See Figure 2-1.

Figure 2-1 VSS Conceptual Diagram



The VSS domain consists of two supervisors—one in each member chassis connected via a *Virtual Switch Link* (VSL). A VSL facilitates the communication between two switches. Within the VSS, one chassis supervisor is designated as *active* and the other as *hot-standby.* Both use *Stateful Switch Over (SSO)* technology. The switch containing the active supervisor is called *active switch* and the switch containing hot-standby supervisor is called *hot-standby switch*. VSS operates on a unified control plane with a distributed forwarding architecture in which the active supervisor (or switch) is responsible for actively participating with the rest of the network and for managing and maintaining control plane information. The active switch maintains and updates the hot-standby supervisor with up-to-date information about the states of the system and network protocols via the VSL. If the active supervisor fails, the hot-standby supervisor assumes the active roles in managing the control plane. Both physical chassis do the data forwarding and data forwarding is done in a distributed manner. The devices adjacent

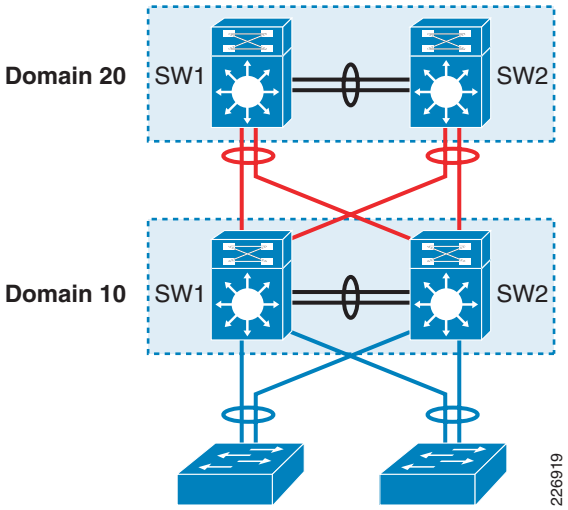
to the VSS are connected via a *Multichassis EtherChannel (MEC)* to form a single logical connection. The single logical switch, in combination with the MEC, forms the foundation of a highly available, loop-free topology. The rest of this section details the operation and best-practice configuration for components of the VSS that influence the deployment of a VSS in the campus distribution block. This design guide does not provide detailed step-by-step instructions about transitioning an existing network to a VSS-based environment, nor does it cover all preparatory work; however, it does cover essential steps that can affect the deployment of a VSS in the campus.

Virtual Switch Domain (VSD) and Switch ID

Virtual Domain

Defining the domain identifier (ID) is the first step in creating a VSS from two physical chassis. A unique domain ID identifies two switches that are intended to be part of the same VSS pair that defines the VSS domain. Assignment of a domain ID allows multiple virtual switch pairs to be connected in a hierarchical manner. Only one VSS pair can participate in a particular domain. The domain ID can have a value ranging from 1 to 255 and must be unique when multiple VSS pairs are connected together. See Figure 2-2.

Figure 2-2 VSS Domain IDs



The domain ID is defined in both physical switches as shown in the following examples.

Standalone Switch 1:

```
VSS-SW1# config t
VSS-SW1(config)# switch virtual domain 10
```

Standalone Switch 2:

```
VSS-SW2# config t
VSS-SW2(config)# switch virtual domain 10
```

The use of a domain ID is important for networks in which VSS is deployed at multiple layers in a manner as shown in Figure 2-2. This unique ID is used in many different protocol and systems configurations—such as virtual Media Access Control (MAC), Port Aggregation Protocol (PAgP), and Link Aggregate Control Protocol (LACP) control packets. If two connected VSS domains contain the same domain ID, the conflict will affect VSS operation.

The following command output example illustrates the domain ID used in the LACP system ID. The last octet—0a (hex)— is derived from the domain ID of 10 (decimal).

```
6500-VSS# sh lacp sys-id
32768,0200.0000.000a <--
```

Tip **The recommendation is to use a unique domain ID as a best practice, even when you are not connecting multiple VSS domains together.**

Switch Identifier

A VSS comprises of pair of physical switches and requires a switch ID to identify each chassis with a unique number. The switch ID can be either 1 or 2 and must be unique on each member chassis. This number is used as part of the interface naming to ensure that the interface name remains the same regardless of the virtual switch role (active or hot-standby switch). Usually the switch ID should be set in configuration mode as shown in the following examples.

Standalone Switch 1:

```
VSS-SW1# config t
VSS-SW1(config-vs-domain)# switch 1
```

Standalone Switch 2:

```
VSS-SW2# config t
VSS-SW2(config-vs-domain)# switch 2
```

However, when a hardware problem on one supervisor prompts the adoption of a new supervisor, you can set the switch ID via the command-line interface (CLI) in enable mode as shown in the following configuration examples.

Standalone Switch 1:

```
6500-VSS# switch set switch_num
```

Standalone Switch 2:

```
6500-VSS# switch set switch_num 2
```

Both methods write the switch identifier in each member chassis ROMMON. The domain ID and switch number via following can be shown via the following CLI example.

```
6500-VSS# show 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
```

Caution Avoid using the command **write erase** to copy a new startup configuration. This command will erase switch numbers stored in ROMMON and any subsequent reboot will cause both switches to come up in standalone mode. Use the **switch set switch_num** *1/2* command only after both switches are rebooted because the CLI to set the switch number is *not* available in VSS mode.

Virtual Switch Link (VSL)

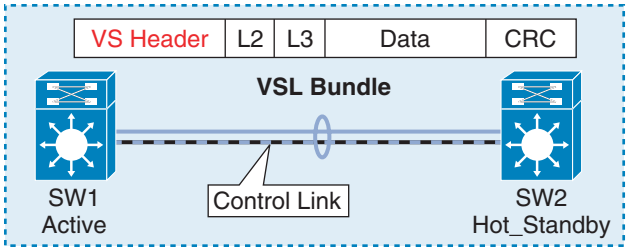
The VSS is made of two physical chassis combined to form a single logical entity. This unification of the control plane is only possible if the system controls, signaling, and backplane exist as a single entity in both chassis. This extension of the system control plane is achieved via a special purpose EtherChannel bundle. The link belonging to EtherChannel is called Virtual Switch Link (VSL). The VSL serves as logical connection that carries critical system control information such as hot-standby supervisor programming, line card status, Distributed Forwarding Card (DFC) card programming, system management, diagnostics, and more. In addition, VSL is also capable of carrying user data traffic when necessary. Thus, the VSL has a dual purpose, supporting system control synchronization and a data link.

The VSL link is treated as a systems control link and encapsulates all traffic into a special system header called the Virtual Switch Header (VSH). This encapsulation is done via dedicated hardware resources and the VSL can only be configured on a 10-Gigabit interface with following hardware ports:

- Sup720-10G, 10-Gbps ports
- WS-X6708
- WS-X6716 (in performance mode only, requires 12.2(33)SXI)

The size of the VSH is the same as that of the internal compact header used by the Cisco Catalyst 6500—it is 32 bytes long. This header is placed after the Ethernet preamble and directly before the Layer-2 header. See Figure 2-3.

Figure 2-3 VSL Header and Control Link



VSL link Initialization and Operational Characteristics

The VSS features a single control plane with a distributed forwarding architecture (see the "Stateful Switch Over Technology" section on page 2-23). Although only one supervisor manages the control plane, both switches participate in learning the control plane information. Any network and system control plane information learned via the hot-standby switch is sent to the active supervisor which in turn updates the hot-standby supervisor. This bidirectional process of learning and updating between switches is carried out over the VSL link.

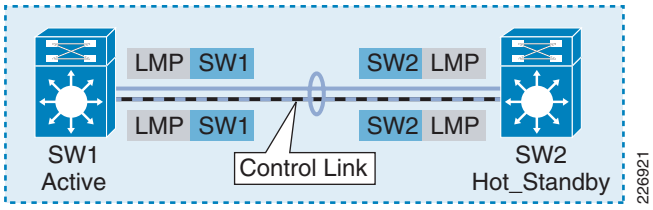
The VSL is an integral component of the VSS because it is the only path through which the two independent systems can become aware of each other during system initialization. This mutual awareness during system initialization is needed to determine the respective role of each physical chassis in becoming either the active or hot-standby virtual switch. For this reason, VSL links are brought up very early in the boot up process—before any other major service or component. To accomplish this task, each switch stores the switch number and other information necessary in ROMMON to activate the VSL link and associated line card booting sequence. During the VSL initialization, the system undergoes a variety of compatibility checks required to form a virtual switch. The VSS requires identical supervisor types and Cisco IOS software versions on both chassis. Please refer to the applicable release note for additional system requirements needed to form a virtual switch. VSL link initialization and maintenance are done through the *VSL Protocol* (VSLP) framework, which consists of two protocols: *Link Management Protocol* (LMP) and *Role Resolution Protocol* (RRP). LMP manages link integrity, while RRP determines the role of each switch member in the virtual switch domain. RRP is covered under "Stateful Switch Over Technology" section on page 2-23.

Link Management Protocol (LMP)

LMP is the first protocol to be initialized, once the VLS line card diagnostics is finished and VSL link comes on line. See Figure 2-4. The LMP is designed to perform following critical functions:

- Establishing and verifying bidirectional communications during startup and normal operation.
- Exchanging switch IDs to detect a duplicate switch ID or member that is connected to another virtual switch. This will be used by the RRP to determine the role of the switch (active or hot-standby).
- Independently transmitting and receiving LMP hello timers to monitor the health of the VSL and peer switch.

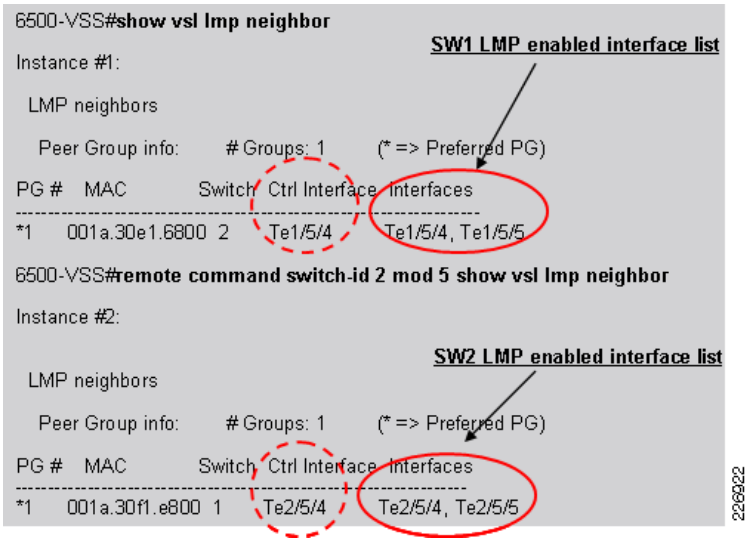
Figure 2-4 Link Management Protocol (LMP)



LMP operates independently on each member switch in the same Virtual Switch Domain (VSD). Unlike other protocols—such as PAgP, LACP and Interior Gateway Protocol (IGP) hello—that use a single control plane over which the active switch originates and terminates the protocol, both switches in the VSD independently originate and terminate LMP control-plane packets on the Switch Processor (SP). See the dotted-red circle highlighting in Figure 2-5. LMP is designed to run on each VSL member link to maintain multiple state machines with the same peer over different ports. In a case in which a unidirectional condition on a port is detected, LMP marks it as *down* and attempts to restart VSLP negotiation—instead of *err-disabling* the port. On each member switch of a VSD, LMP internally forms a single unique peer group (PG) ID with a common set of VSL links. When all VSL interfaces are down, LMP destroys the peer group and notifies RRP to take an appropriate action. The active switch will detach all the interfaces associated with the hot-standby switch. At the same time the hot-standby switch performs switchover, assumes the active role, and detaches all interfaces associated with the previously active switch.

LMP hello packets are Logical Link Control (LLC)/ Sub-network Access Protocol (SNAP)-encapsulated with the destination MAC address matching the Cisco Discovery Protocol (CDP)—01.00.0C.CC.CC.CC. All inter-chassis control plane traffic, including LMP and hello packets, are classified as bridge protocol data unit (BDPU) packets and are automatically placed in transmit priority queue.

Figure 2-5 Output Showing LMP Enabled Interface List



Control Link and Inter-Chassis Control Plane

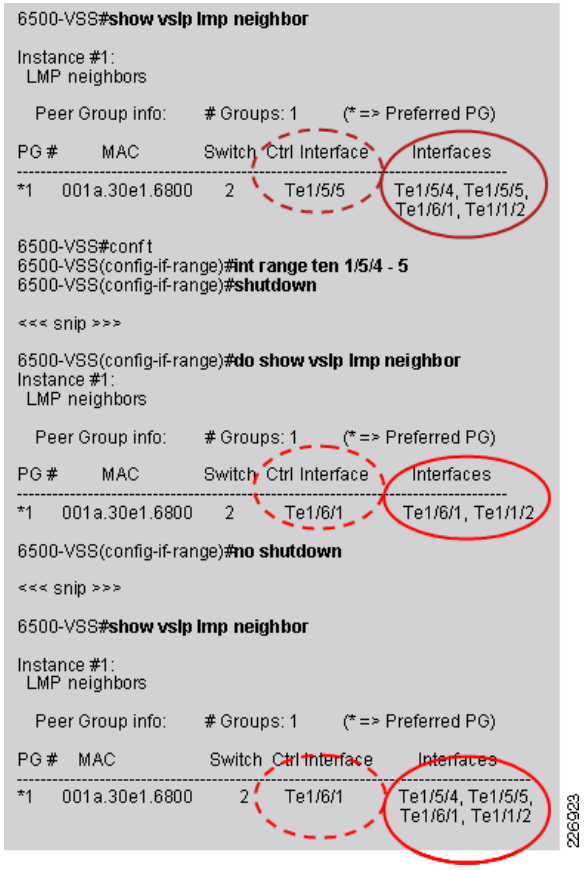
The VSL bundle is special purpose EtherChannel that can have up to eight members. Only one link out of a configured member is selected as the control link and that control link is the only link that can carry the inter-chassis control plane. The control link carries the inter-switch External Out-of-Band Channel (EOBC) control traffic that includes the Switch Control Packet (SCP) for line card communication, Inter-process Communication Packets (IPC), and Inter-Card Communication (ICC) for communicating the protocol database and state—as well as updates to the hot-standby supervisor. In addition, the same link can carry user and other network control traffic—depending how the traffic is hashed (based on source and destination MAC and/or IP addresses). The remaining bundled links carry network control plane and user data traffic, but not the inter-chassis control plane traffic (see the "Traffic Prioritization and Load-sharing with VSL" section on page 2-15). The control link is shown in Figure 2-4

The control-link selection procedure is determined by the VSS system and cannot be managed by the user. During the bootup process, the first VSL link that establishes LMP relationship (state-machine) will be selected as the control link. Based on the Cisco Catalyst 6500 architecture, the supervisor module becomes operational ahead of any other module installed in the switch. If the 10-Gbps port of Sup720-10G module is bundled in the VSL EtherChannel, then it will be selected as control-link interface whenever both switches undergo the boot process.

The **show vslp lmp neighbor** command output illustrated in Figure 2-6 depicts the current control link interface (see the dotted red circle) and list of backup VSL interfaces (which can be used if the current control-link path fails). Backup interfaces are member links of the local virtual switch's VSL EtherChannel. For a highly redundant VSL network design, the VSL EtherChannel must be bundled with multiple, VSL-capable 10-Gbps ports between the switch members. In such a case, the first listed interface under the *Interfaces* column of the **show vslp lmp neighbor** command will immediately become control link interface if the current control link interface fails. When the 10-Gbps port of a

Sup720-10G module is restored and rejoins the VSL EtherChannel, it will be placed as the next available control-link backup path without affecting the current control link interface (see second part of the output illustrated Figure 2-6 after control link is brought back up).

Figure 2-6 Control Link Interfaces Selection



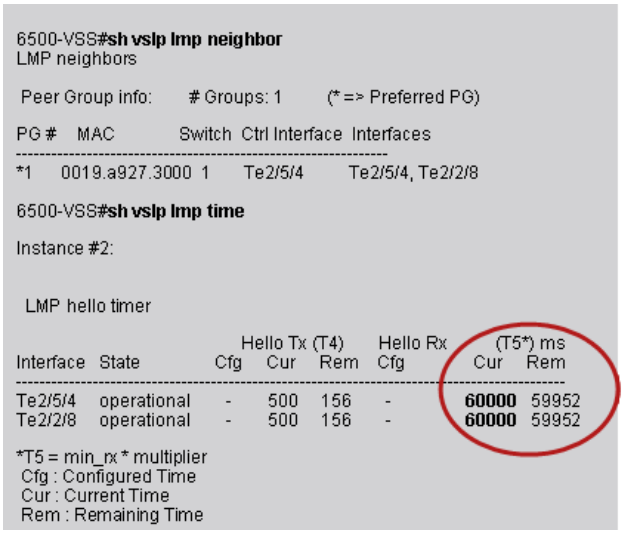
LMP Heart Beat

The LMP heart beat—also referred as the LMP hello timer—plays a key role in maintaining the integrity of VSS by checking peer switch availability and connectivity. Both VSS members execute independent, deterministic SSO switchover actions if they fail to detect the LMP hello message within configured hold-timer settings on the last bundled VSL link. The set of LMP timers are used in combination to determine the interval of the hello transmission applied to maintain the healthy status of the VSL links. The three timers are as follows:

- Hello Transmit Timer (T4)
- Minimum Receive Timer (min_rx)
- T5 Timer (min_rx * multiplier)

Figure 2-7 illustrates an example CLI output depicting the timer values per VSL link member.

Figure 2-7 Timer Values per VLS Link Member



By default, the LMP hello transmit timer (T4) and receive timer (min_rx) are assigned values of 500 msec each. The hold-timer (T5 timer) is derived from a min_rx and default multiplier of 120 (the CLI does not show the default multiplier). By default, a VSL member link time out is detected in 60,000 msec (60 seconds). The expiration of T5 indicates possible instability on the remote peer (active or hot-standby switch). Each switch member will take an independent action if the T5 timer expires. These actions include (but are not limited to) the following:

- If expiration occurs on a VSL port that is the control link, then the switch that detected the problem will force the new control link selection. It is entirely possible that T5 timer would have not yet expired on the remote peer; however, the remote peer switch will respect the request and reprogram internal logic to send control plane traffic to the newly elected VSL port as the control link port.
- If expiration occurs on a non-control link port, the switch that detected the failure selects an available port for user data traffic. Eventually, the remote peer detects a change and removes the link from the bundle.
- If expiration occurs on the last VSL port (a combination control link and user data port) and the timeout is detected on the active switch, then the active switch removes all the peer switch interfaces and announces the change to rest of the network—depending on the configuration on those interface (Layer-2 or Layer-3 protocol). Eventually, the peer switch that is in hot-standby mode will detect the T5 timer expiration and LMP will inform RRP, which forces the hot-standby switch to become the active switch. This triggers a condition known as *dual active* in which both switches declare active roles leading to instability in the network (refer to the "Campus Recovery with VSS Dual-Active Supervisors" section on page 4-18 for information about avoiding such conditions).

Why Timer Should Not be Modified

The LMP timer is used primarily for ensuring the integrity of VSS (during high CPU usage or abnormal software behavior). Normal hardware failure detection or switchover (user or system initiated) is invoked via the hardware mechanism known as *Fast Link Notification* (FLN). When an applicable event occurs, FLN informs the firmware of WS-X6708 or Sup720-10G ports to take any necessary action—typically within 50 to 100 msec. FLN is not designed to detect the failure of the remote switch. In most cases, modifying the default LMP timer to a shorter value does not improve convergence. This is because inter-chassis control plane protocol timer (IPC timer) expires before LMP timers and thus supersede the action taken.

Unintended effects can result when VSLP timers are aggressively modified. For example, modifying the VSLP timer to a lower value will add significant instability. An example scenario description follows:

> When configured with a lower VSLP timer value, the VSL will typically fail to establish neighbor relationships between member switches—leading to continuous rebooting (a *crash* in Cisco IOS parlance) of the switch. During the boot up process, the VSS switch must process multiple high-priority activities. When enabled with an aggressive VSLP timer, each member might be unable to maintain the VSLP session due to the lack of resources. When LMP hello times out on either side, LMP removes the VSL member link from VSL EtherChannel. Eventually, all links between the active and hot-standby switches can fail. For the hot-standby switch, this is considered a catastrophic error and the only way to recover is to immediately send a reset signal (crash) and start over. This can continue until at least one LMP session is established and thus can cause significant network instability.

In addition, a VSS member might also fail to send or receive LMP hello message within a configured T5 timer limit due to VSL link congestion or high CPU utilization. In such situations, the VSS system will become unstable, which can lead to a dual-active condition.

Tip **Cisco strongly recommends that you do *not* modify the default LMP (VSLP) timers.**

Role Resolution Protocol (RRP)

RRP is responsible for determining the operational status of each VSS switch member. Based on the configured parameter, a member switch can assume a role of the active, hot standby, or Route Process Redundancy (RPR). The RRP provides checks for Cisco IOS software compatibility. If the software versions are not compatible, RRP forces one of the switches into RPR mode and all line cards are powered off. The "Virtual Switch Role, Priorities and Switch Preemption" section on page 2-27 provides details of the RRP because its application is more relevant to chassis redundancy and SSO operation.

Configuring VSL Bundle

Configuring VSL is the second step in creating a virtual switch (after defining the domain ID). The VSL bundle is a special-purpose port channel. Each standalone switch is required to be configured with unique port-channel interface numbers; before assigning a port-channel number, you must make sure that the port-channel interface number is not used by an existing standalone configuration on either switch. The following configuration steps are required in each switch to convert the standalone switch to a VSS. The detailed conversion process is beyond the scope of this document. For addressing VSS conversion, refer to the *Migrate Standalone Cisco Catalyst 6500 Switch to Cisco Catalyst 6500 Virtual Switching System* document at the following URL:

http://www.cisco.com/en/US/products/ps9336/products_tech_note09186a0080a7c74c.shtml

Standalone Switch 1:

```
VSS-SW1(config)# interface Port-Channel1
VSS-SW1(config-if) #switch virtual link 1

VSS-SW1(config-if)# interface range Ten5/4 - 5
VSS-SW1(config-if)# channel-group 1 mode on
```

Standalone Switch 2:

```
VSS-SW2(config-if)# interface Port-Channel2
VSS-SW2(config-if)# switch virtual link 2

VSS-SW2(config-if)# interface range Ten5/4 - 5
```

```
VSS-SW2(config-if)# channel-group 2 mode on
```

Since VSL EtherChannel uses LMP per member link, the link-aggregation protocols, such as PAgP and LACP, are not required; each member link must be configured in unconditional EtherChannel mode using the **channel-group group-***number* **mode on** command. Once the VSL configuration is completed, using the **switch convert mode virtual** CLI command at the enable prompt will start the conversion process. The conversion process includes changing the interface naming convention from *slot/interface* to *switch_number/slot/interface*, saving the configuration, and rebooting. During switch rebooting, the systems recognize the VSL configuration and proceeds with their respective VSL ports initialization processes. The two switches communicate with each other and determine which will have active and hot-standby roles. This exchange of information is evident through the following console messages:

Standalone Switch 1 console:

```
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
<snip>
00:00:26: %VSL_BRINGUP-6-MODULE_UP: VSL module in slot 5 switch 1 brought up
Initializing as Virtual Switch active
```

Standalone Switch 2 console:

```
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
<snip>
00:00:26: %VSL_BRINGUP-6-MODULE_UP: VSL module in slot 5 switch 2 brought up
Initializing as Virtual Switch standby
```

A first-time VSS conversion requires that you **must** execute the following command as the final step of accepting the virtual mode operation of the combined switches. If the switch has been converted, or partially converted, you cannot use this command.

```
6500-VSS# switch accept mode virtual
```

The preceding command forces the integration of all VSL-link related configurations from the hot-standby switch and populates the running configuration with those commands. In addition, the startup configurations are updated with the new merged configurations. The following prompt appears:

```
Do you want proceed? [yes/no]: yes
Merging the standby VSL configuration. . .
Building configuration...
[OK]
```

Note Only VSL-related configurations are merged with the conversion step. All other configurations must be managed per your network site implementation requirements. Follow the related Cisco product documentation for further details.

VSL Characteristics

The VSL port channel is treated as an internal systems link. As a result, its configuration, resiliency, mode of operation, quality of service (QoS), and traffic load sharing follow a set of rules that are specific to VSL. This section covers configuration requirements relating to those rules. The logical port channel and its member link both have distinct sets of restrictions.

VSL port-channel logical interface configuration is restricted to VSL-related configuration; all other Cisco IOS features are disabled. The following output illustrates available options:

```
6500-VSS(config)# int po 1
6500-VSS(config-if)# ?
virtual link interface commands (restricted):
  default        Set a command to its defaults
  description    Interface specific description
  exit           Exit from virtual link interface configuration mode
  load-interval  Specify interval for load calculation for an interface
  logging        Configure logging for interface
  mls            mls sub/interface commands
  no             Negate a command or set its defaults
  port-channel   Port Channel interface subcommands
  shutdown       Shutdown the selected interface
  switch         Configure switch link
  vslp           VSLP interface configuration commands
```

VSL member links are in restricted configuration mode once the VSL configuration is applied. All Cisco IOS configuration options are disabled except following:

- EtherChannel
- Netflow configuration
- Default QoS configuration

The following output illustrates available options:

```
6500-VSS(config)# int ten 1/5/4
6500-VSS(config-if)# ?
virtual link interface commands (restricted):
  channel-group   Etherchannel/port bundling configuration
  default         Set a command to its defaults
  description     Interface specific description
  exit            Exit from virtual link interface configuration mode
  load-interval   Specify interval for load calculation for an interface
  logging         Configure logging for interface
  mls             mls sub/interface commands
  no              Negate a command or set its defaults
  priority-queue  Configure priority scheduling
  rcv-queue       Configure receive queue(s)
  shutdown        Shutdown the selected interface
  wrr-queue       Configure weighted round-robin xmt queues
```

When configuring VSL interfaces, only one VSL EtherChannel configuration per virtual-switch is possible. Configuring an additional VSL EtherChannel will result in an error. The following are examples of error messages generated:

```
6500-VSS(config)# interface port-channel 3
6500-VSS(config-if)# switch virtual link 1
% Can not configure this as switch 1 VSL Portchannel since it already had VSL Portchannel
1 configured
6500-VSS(config-if)#
6500-VSS(config-if)# switch virtual link 2
% Can not configure this as switch 2 VSL Portchannel since it already had VSL Portchannel
2 configured
```

In addition, for post VSS conversion, a neighbor switch port cannot be bundled into the local switch's VSL EtherChannel. In the example used here, the Switch 1 VSL EtherChannel cannot bundle a physical port from Switch 2. However, a regular EtherChannel does not have such a restriction. The following example output illustrates the error message generated when attempting to configure this unsupported mode:

```
6500-VSS(config-if)# int te2/1/1
6500-VSS(config-if)# channel-group 1 mode on
VSL bundle across chassis not allowed TenGigabitEthernet2/5/5 is not added to port channel
1
```

The remainder of this section covers design considerations for prioritizing traffic over VSL links, resiliency, and traffic load-sharing options available with VSL.

VSL QoS and Prioritization of Traffic

Today's enterprise application requirements are diverse. Many time-critical services (including voice, video and multicast)—as well as enterprise applications, such as customer relationship management (CRM), SAP, and Oracle—require specific priorities in the campus network. The QoS design guide (available at the URL below) describes an approach for classifying user traffic at the edge and using those priorities intelligently via the Differentiated Services Code Point (DSCP) trust model. This design guide does not cover generic QoS requirements and behavior in VSS; however, this section does address QoS as it applies to a VSL link in terms of default behavior, how the control plane is protected, and implementation options that network designers should consider.

http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/QoS_SRND_40/QoSCampus_40.html

Hardware Configuration Dependency

Due to the ease of configuration, flexibility, and bandwidth requirement, the distribution layer traditionally leverages line cards for connectivity to the core and access layer, thus the use of supervisor ports is limited. The access layer uses the uplink from either a supervisor or a dedicated port on non-modular switches. The Sup720-10G supervisor offers a new option of using the 10-Gbps port available on the supervisor to provide uplink connectivity to the core. This requires an understanding of the current design choices with the Sup720-10G uplink port with respect to QoS when used as the VSL port and/or uplink to the rest of the network. The VSL link can only be configured on 10-Gbps ports. Choosing the VSL link configuration on either a supervisor port or line card affects the QoS capability on the unused ports of the supervisor. The Sup720-10G supervisor has two 10-Gbps ports and three 1-Gbps ports. The Sup720-10G uplink ports can be configured in one of two modes:

- *Default—Non-10 gigabit-only mode*

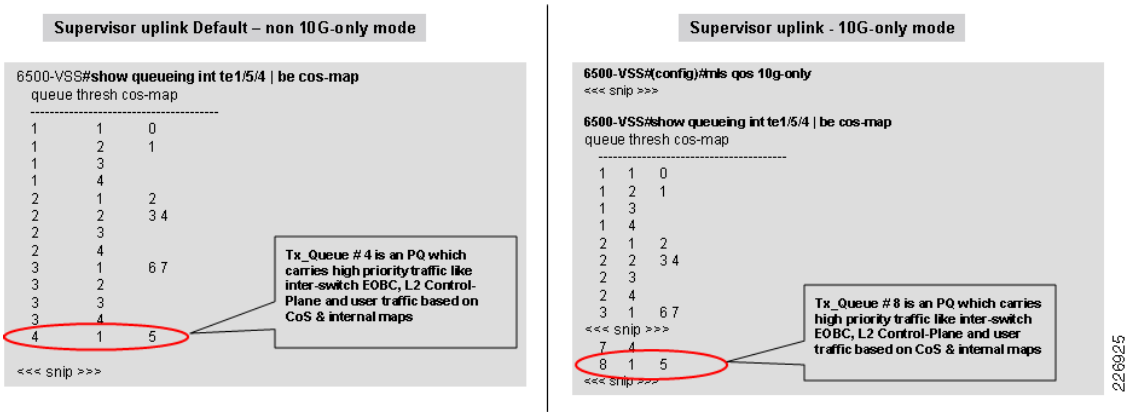
  In this mode, all ports must follow a single queuing mode. If any 10-Gbps port is used for the VSL link, the remaining ports (10 Gbps or 1Gbps) follow the same CoS-mode of queuing for any other non-VSL connectivity because VSL only allows class of service (CoS)-based queuing.

- *Non-blocking—10 gigabit-only mode*

  In this mode all 1-Gbps ports are disabled, as the entire module operates in a non-blocking mode. Even if only one 10G port used as VSL link, still both 10-Gbps port is restricted to CoS-based trust model. 12.2(33)SXI removes this limitation by allowing the unused (non-VSL configured)10-Gbps port be configured based on user preference, including DSCP-based queuing.

Figure 2-8 shows that a 10-Gbps-only mode increases transmit queue from default *1p3q4t* to the *1p7q4t* setting. It also increases the receive queue size from *2p4t* to *8q4t* similar to the WS-X6708 module. However, the default Tx and Rx queue mapping remains CoS-based with or without 10-Gbps-only mode. As a result, there is no real benefit to using an improved queue structure to map each class of traffic in separate queues because the default COS-to-queue mapping cannot be changed.

Figure 2-8 Comparison of Default and Non-Blocking Modes



If one of the WS-X6708 line card port is used as the VSL link, port queuing for that port is limited to being CoS-based; however, the remaining ports can have independent sets of QoS configuration.

The resilient VSL design uses these facts as a design factor. Table 2-1 summarizes the options and restrictions for Sup720-10G uplink ports. The "Resilient VSL Design Consideration" section on page 2-18 describes incorporating these design factors when developing highly resilient and flexible VSL configurations.

Table 2-1 Options and Restrictions for Sup720-10G Uplink Ports

Sup720-10G Uplink Port	10g-only mode	Non 10g-only mode
Queue Structure	Tx– 1p7q4t Rx – 8q4t	Tx –1p3q4t Rx – 2q4t

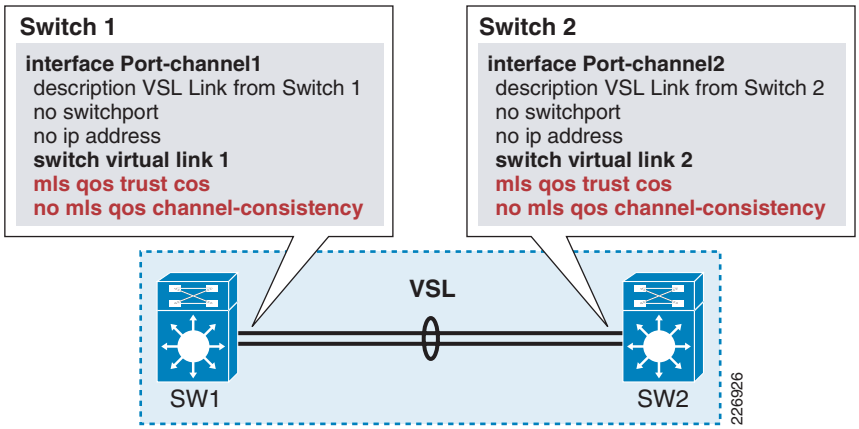
Table 2-1 Options and Restrictions for Sup720-10G Uplink Ports (continued)

Sup720-10G Uplink Port	10g-only mode	Non 10g-only mode
Non-VSS mode (standalone mode) Or VSS mode (no supervisor port used as VSL link)	Only 10-Gbps ports available, all 1-Gbps ports are disabled. Both 10-Gbps ports can have independent sets of QoS configuration including trust model and queuing mode. The DSCP-based queuing allowed.	All ports are available. All uplinks can *only* use a single QoS configuration (maps, thresholds, and queues). Default queuing mode is CoS only. No DSCP based queuing allowed.
10-Gbps ports used for VSL link	If both 10-Gbps ports are used as the VSL link, only CoS-based trust mode supported. Even if only a single 10-Gbps port is used as the VSL link, both 10-Gbps ports are restricted to the CoS-based trust model in Cisco IOS 12.2(33) SXH. Cisco IOS 12.2(33) SXI removes this limitation by allowing the remaining 10-Gbps port to be configured based on user preference, including DSCP-based queuing. The remaining 1-Gbps ports are disabled.	Both 10-Gbps ports are used or a single 10-Gbps port is used as the VSL uplink; you can only define one QoS configuration. Because VSL only allows CoS-based queuing, the remaining non-VSL ports follow the COS-based queuing for any non-VSL connectivity.

Default QOS Setting for VSL

The VSL (by default) uses a CoS-based trust model. This default CoS-trust setting on VSL EtherChannel cannot be modified or removed. Regardless of the QoS configuration in global mode, the VSL EtherChannel is set to CoS-based queuing mode. Figure 2-9 shows the QoS configuration of VSL link on SW1 and SW2. As with a standalone switch, the VSS uses the internal CoS-based queuing with various mapping tables for placing user traffic into the egress queue. Traffic traversing the VSL link uses the same facilities with a CoS-based trust model. The Weighted Round Robin (WRR) queuing scheduler is enabled by default on each VSL member link. The VSL link's QoS configuration does not alter the QoS marking set in the user traffic when that traffic traverses the VSL link.

Figure 2-9 VSL CoS Configuration Example Comparison



The best-practice recommendation for VSL link resiliency is to bundle two 10-Gbps ports from different sources. Doing this might require having one port from the supervisor and other from a Cisco 6708 line card. By default, the Sup720-10G supervisor 10-Gbps port has a non-Distributed Forwarding Card (DFC) 1p3q4t Tx-queue structure in contrast with the WS-X6708 linecard which has a DFC-based 1p7q4t Tx-queue structure. For the conventional configuration of EtherChannel, a bundle between non-congruent queue structures would fail. However, the VSL bundle is by default enabled with the configuration which allows the EtherChannel to be formed. This is enabled via the **no mls qos channel-consistency** command.

The following QoS configuration restrictions apply once the port is bundled in the VSL EtherChannel:

- The CoS-mode setting cannot be changed or removed from VSL port-channel.
- Any QoS configuration, such as DSCP-based trust, receive-queue bandwidth, or threshold limit, cannot be modified on any of the 1-Gbps or 10-Gbps ports of the Sup720-10GE module after bundling the link to the VSL port-channel. Applying such restricted QoS configurations will result in an error.
- User-defined Modular QoS CLI (MQC) service-policies cannot be attached to the VSL EtherChannel.
- Bundling of the 10-Gbps port in the VSL port-channel will fail if unsupported QoS capabilities are preconfigured. All QoS configuration must be removed prior bundling the port into the VSL EtherChannel.

Note The WS-X6708 module supports independent QoS policies on non-VSL configured ports, even if one of its 10-Gbps ports is bundled in the VSL EtherChannel.

Traffic Prioritization and Load-sharing with VSL

This section addresses the following topics:

Prioritizing Specific Traffic Types

The VSL link can potentially carry three types of traffic and uses QoS mapping to distinguish between each. Traffic types carried over a VSL link include the following:

- User data traffic
- Network control plan traffic
- VSS inter-switch communication and Layer-2 per link control traffic

These are described briefly in the sections that follow.

User Data Traffic

In a recommended best-practice configuration implementation, all devices are attached to the VSS via MEC-based connections (see the "MEC Configuration" section on page 2-43). In dual-homed MEC connectivity, pass-through user data traffic does not traverse the VSL link. However, during certain conditions user data must traverse VSL link; applicable conditions include (but are not limited to) the following:

- Failure of uplink from access layer to VSS causing downstream traffic to flow over the VSL link
- Remote Switched Port Analyzer (SPAN) traffic from one VSS switch member to the other
- Service-module traffic flow from FWSM, Wireless Services Module (WiSM), Intrusion Detection System (IDS), and other modules

VSL itself does not alter the QoS marking of user data traffic. It simply categorizes the preceding types of traffic using the CoS-based queuing. As a result, any ingress traffic QoS marking that is not based on CoS must use the internal QoS mapping table to provide the translation to CoS-based queuing. Any end-user application traffic marked with 802.1p CoS 5, DSCP 46, or IPP 5 will be placed in the priority-queue.

Network Control Plane Traffic

The active switch always originates and terminates network control plane traffic from participating adjacent devices. The active switch always uses a locally attached interface (link) to forward the control plane traffic. The network control plane traffic that must traverse the VSL due to either a failure of local links connected to active switch or traffic sourced by a hot-standby switch, including the following:

- *Layer-3 protocol traffic for Layer-3 Equal Cost Multipath (ECMP) links on the hot-standby switch*—Routing protocol control traffic, such as hello, update, database, and so on
- *Traffic intended for the VSS supervisor*—Internet Control Message Protocol (ICMP) responses from other devices, time-to-live (TTL) with value of 1 in increments of hop counts (it must terminate at the active switch), SNMP, Telnet/SSH, and so on.

VSS Inter-Switch Communication and Layer-2 per Link Control Traffic

All communications among VSS member switches are defined as inter-switch traffic. VSS systems automatically classify the following inter-switch control plane protocols as Bridge Protocol Data Unit (BPDU)-type traffic:

- Inter-Switch Communication
  - *Inter-Chassis Ethernet Out Band Channel (EOBC) traffic*— Serial Communication Protocol (SCP), IPC, and ICC
  - *Virtual Switch Link Protocol (VSLP)* —LMP and RRP control-link packets
- *Any Layer-2 per link protocol*—Spanning Tree Protocol (STP) BPDU, Port Aggregation Protocol (PagP)+, LACP, CDP and Unidirectional Link Detection (UDLD), Link Layer Discovery Protocol (LLDP), Root Link Query (RLQ), Ethernet Operations, Administration, and Maintenance (OAM), 802.1x, Dynamic Trunking Protocol (DTP), and so on.

These BPDU packets are automatically placed in transmit priority-queue ahead of any other traffic

Note Network control plane traffic (Layer 2 and Layer 3) is always sent out links connected to the active switch. It will only cross over the VSL either due to a local link being unavailable or the protocol frame needing to be originated from the hot-standby port (e.g. PAgP, LACP or UDLD).

Note Priority-queues are shared by the high-priority user data traffic (marked as expedited forwarding -EF) along with control plane traffic. However, an internal mechanism always ensures that control plane traffic takes precedence over of any other priority-queued traffic. This ensures that user data does not inadvertently affect the operational stability of the entire VSS domain.

Load-Sharing with VSL

As described in the preceding section, the VSL carries multiple types of the traffic over the VSL bundle. From the traffic load-sharing perspective, the VSL bundle is just like any other EtherChannel. It follows the same rules of EtherChannel hashing algorithm available in any given Cisco IOS software. In standalone or virtual-switch mode, a single EtherChannel load-balance hashing is applicable system wide. This means the VSL bundle uses the same configured load-sharing mode that applies to all EtherChannel groups in the VSS. The following output example illustrates load-balancing options:

```
6500-VSS(config)# port-channel load-balance ?
  dst-ipDst 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
```

Note This section only covers the characteristics of EtherChannel related to VSL. For generic EtherChannel design and recommendation refer to the "MEC Configuration" section on page 2-43.

The load-balancing method implemented applies to all network control traffic and user data traffic. The only exceptions to this rule are inter-switch communication traffic (always carried by control link) and LMP hello (sent on every VSL link). The network control plane and user data traffic use source and/or destination MAC addresses and/or the IP address as input to the hash calculation that is used to load-share the traffic. The network control plane traffic that can be load-shared between VSL links includes, but is not limited to, the following (note the difference in QoS categorization of the traffic):

- *Layer-2 protocol traffic*—Broadcast, STP BPDU, PAgP+, LACP, CDP and UDLD , LLDP, RLQ , Ethernet OAM, 802.1x, and DTP
- *Layer-3 protocol traffic for Layer-3 ECMP links on the hot standby switch*—Routing protocol control traffic (such as Hello, Update, and database) and ICMP response
- *Traffic designated to VSS supervisor*—ICMP response from other devices, TTL with 1, and so on

The type of user data traffic crossing VSL links are described in the "Prioritizing Specific Traffic Types" section on page 2-15 section.

Hashing Methods—Fixed versus Adaptive

Traffic across port-channel is distributed based on Result Based Hash (RBH) computation for each port-channel member link. Whenever a port-channel member link is added or removed from a group, RBH must be recomputed for every link in the group. For a short period of time, each flow will be rehashed—causing disruption for the traffic. This hash implementation is called *fixed*.

As of Cisco IOS Release 12.2(33) SXH, Cisco Catalyst 6500 supports the enhanced hash algorithm that pre-computes the hash for each port-channel member link. When the link member fails, dynamic pre-computation of hashing allows new flows to be added to the existing link and also reduces the loss of packets for the flows that were already hashed to the link. This enhanced hash implementation is called *adaptive*. The following example illustrates the hash-distribution options:

```
6500-VSS(config-if)# port-channel port hash-distribution ?
  adaptive  selective distribution of the bndl_hash among port-channel members
  fixed     fixed distribution of the bndl_hash among port-channel members

VSS(config-if)# port-channel port hash-distribution fixed
This command will take effect upon a member link UP/DOWN/ADDITION/DELETION event.
Please do a shut/no shut to take immediate effect
```

By default, the load-sharing hashing method on all non-VSL EtherChannel is fixed. In contrast, the default hash algorithm for the VSL bundle is adaptive because it carries critical inter-switch control plane traffic. The default hashing method can be changed, but the only way to make the hash algorithm effective is to reset the link. Applying this to the VSL will trigger the dual-active condition because both chassis will lose connection with each other when the VSL links bounce (see the "Campus Recovery with VSS Dual-Active Supervisors" section on page 4-18).

Tip **It is recommended to keep the VSL link load-sharing hash method to default (adaptive) as that method is more effective in recovering flows from failed links.**

The current EtherChannel hardware can only load-share with three unique binary buckets, thus any combination of links in the EtherChannel bundle that can fill the all the buckets would optimally use all the links in the bundle. This translates into a number of links in the bundle with a formula of the power of 2 for optimal load sharing.

Tip **Always bundle the numbers of links in the VSL port-channels in the power of 2 (2, 4, and 8) to optimize the traffic flow for load-sharing.**

Resilient VSL Design Consideration

The VSL can be configured as single member EtherChannel. Configuration of a resilient VSL link follows the same design principles that apply to deploying a resilient EtherChannel-connected device. Resilient EtherChannel design consists of avoiding any single point-of-failure in terms of line cards or ports. Redundancy of VSL is important so as to avoid the dual-active condition and instability of the VSS. VSL redundancy is useful whether or not the supervisor fails. In the case of an active supervisor failure, the hot-standby switch (supervisor) is ready to take over—VSL resiliency notwithstanding. VSL resiliency *is* important when the supervisor has *not* failed and somehow the systems have lost their VSS links—leading to a dual-active condition.

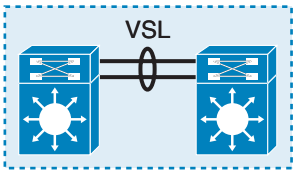
The following key factors should be considered when designing resilient VSL:

- Use port-channels with more than one member in a bundle to reduce the number of potential single points of failure (ports, line cards)

- Use redundant hardware (ports, line cards, and internal resources connecting the port)
- Use diverse fiber paths (separate conduits, fiber terminations, and physical paths) for each VSL links.
- Manage traffic forwarded over the VSL link to avoid single homed devices. This is covered under the "Traffic Flow in the VSS-Enabled Campus" section on page 3-5.
- Since the VSL can only be configured on 10-Gbps port, choices for deploying the VSL bundle are limited to the Sup720-10G, WS-X6708, or WS-X6716 hardware. Besides redundancy, capacity planning also influences number of VSL members per VSL bundle. The capacity planning is explained under the "Capacity Planning for the VSL Bundle" section on page 3-12. There are three design options for avoiding a single point-of-failure:
  - *Use two 10-Gbps ports available with Sup720-10G supervisor*

    Design option 1 (Figure 2-10) is the most common and most intuitive choice. It uses both 10-Gbps ports on the supervisor. However, this option does not provide optimal hardware diversity as both ports are connected to single internal fabric connection. The probability of having both port connections to the fabric backplane having hardware errors is low, but not impossible.

Figure 2-10 VSL over Supervisor Port



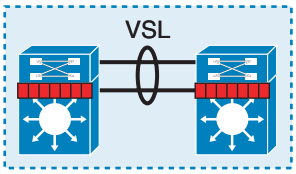
Design 1 – VSL Bundle over Supervisor Port

226927

  - *Use one 10-Gbps port from the Sup720-10G supervisor and another from a VSL capable line card (WS-X6708 or WS-X6716)*

    Design option 2 (Figure 2-11) diversifies the VSL links onto two separate hardware line cards—one port from the Sup720-10G uplink and another from the WS-X6708 line card. This is the best baseline and most practical option for balancing cost and redundancy. This design restricts unused ports on Sup720-10G with CoS-based queuing. Cisco IOS 12.2(33) SXI removes this restriction.

Figure 2-11 Distributed VSL Bundle between Bundle and 67xx Line Card



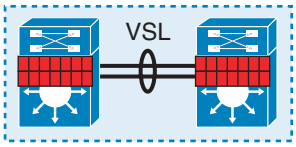
Design 2 – Distributed VSL Bundle between Supervisor and 67xx Line Card

226928

  - Use both 10-Gbps ports from VSL capable line cards (WS-X6708 or WS-X6716)

    Design option 3 (Figure 2-12) uses separate line cards for VSL connectivity. This provides the best flexibility in term of QoS configurations for uplink ports on supervisors, but is not as cost effective as design option 2.

Figure 2-12 Distributed VSL Bundle between Dual 67xx Cards



Design 3 – Distributed VSL Bundle between Dual 67xx Line Cards

226929

Besides avoiding single point-of-failure, the right design selection depends on the availability of the hardware and usage of the uplink ports on the supervisor (see the "VSL QoS and Prioritization of Traffic" section on page 2-12). If one 10-Gbps uplink port is used as the VSL link, then the other 10-Gbps port can only have CoS-based queuing. The Cisco IOS software as of Cisco IOS 12.2(33) SXI removes the CoS-based queuing restriction and allows the other non-VSL 10-Gbps port to be configured for DSCP-based queuing. If any of the Sup720-10G ports are used for connectivity to the core or other network layer in which QoS requirements require flexibility, then design options 2 and 3 are the available choices.

For superior resiliency, having one VSL link on the supervisor and other on a line card is the best option. This design option avoids possible catastrophic disconnection of line cards and allows more ports to be added in the VSL bundle for future growth. Table 2-2 details the design choices for the VSL links. You should be aware that there is no clear choice that emerges from the following options for all implementations; however,Table 2-2 does provide some risk categorization for available choices.

Table 2-2 Design Choices with VSL Links

	Design 1-1 Default - Non 10g-only	Design 1-2 10g-only (mls qos 10g-only)	Design 2-1 Default - Non 10g-only	Design 2-2 10g-only (mls qos 10g-only)	Design – 3
Hardware Configuration	VSL link both on Sup 720-10G uplinks. All uplink ports are available.	VSL link both on Sup 720-10G uplink. Only 10-Gbps ports are available.	One port on Sup 720 and other on 10-Gbps line card. All uplink ports are available.	One port on Sup 720 and other on line card. Only 10-Gbps ports are available.	Both VSL link on different 10-Gbps line cards.
Port Usage and Performance Mode	All uplink ports are available. 10-Gbps and one Gbps are sharing the total 20 Gbps of bandwidth.	Only 10-Gbps ports are available. 10-Gbps ports are non-blocking.	All uplink ports are available. 10-Gbps and one Gbps are sharing the total 20 Gbps of bandwidth.	Only 10-Gbps ports are available. 10-Gbps ports are non-blocking.	All sup uplink ports are available. WS-X6708–2:1 oversubscription.
QoS[1] Configuration Flexibility	Least – all ports can only have CoS based queuing.	CoS-based only since all 10-Gbps ports are used as VSL link. However loses three 1-Gbps ports.	Limited though practical. All uplink ports can only have CoS-based queuing. 10-Gbps line card can have independent QoS.	More Limited then option 2-1. The reminder 10-Gbps port follows the same QoS configuration as VLS port. 12.2(33) SXI removes this restriction. 10-Gbps line card can have independent QoS.	Most Flexible. All sup uplink can be used. All ports on 10 Gbps can have independent QoS in 10g-only mode.

Table 2-2 Design Choices with VSL Links (continued)

VSL Single Point-of-Failure	Possible, due to failure of fabric interconnects. Risk of failure – very low.	Possible, due to failure of fabric interconnects. Risk of failure –very low.	Possibility of losing both ports on distinct hardware is rare.	Possibility of losing both ports on distinct hardware is remote. Risk of failure—very rare.	Possible, due to loss of connectivity to line card in extreme conditions. Risk–Extremely low.
VSS Boot Behavior[2]	Optimal	Optimal	Optimal	Optimal	Delayed
Overall Choice Criteria	Cost effective, efficient, least flexible for QoS configurations on one Gbps uplink. Not recommended due to least diversity of hardware.	Cost effective, performance guarantee. Not recommended due to least diversity of hardware.	Practical. Future enhancement makes it best overall from cost, efficiency and link redundancy. Though reduced QoS flexibility.	Practical. Future enhancement makes it best overall from cost, efficiency and link redundancy.	Most efficient port usage and flexible QoS. Not as cost effective and optimized as design option 1-1 and 1-2.

1. Queue structure and depth is not a factor in VSL since the mapping and ration remains the same. Refer to "VSL QoS and Prioritization of Traffic" section on page 2-12.
2. The VSL link on Sup720-10G ports will come up faster than VSL on line cards since line cards need more time for image download, initialization etc. VSS is optimized for faster booting with VSL link on supervisor port.

VSL Operational Monitoring

This design guide does not cover operational monitoring and troubleshooting of VSS; however, critical information is included to emphasize the need to manage bandwidth utilization—and the health of the VSL port-channel and its member links. Relevant CLI output examples are shown throughout this section.

Troubleshooting of the VSS and VSL port-channel interface might require the port-channel to be spanned to the port to which the network packet decoder is attached. The VSL port-channel can be spanned. However, you can only span a local port-channel to a local destination. Refer to the following CLI command output.

```
6500-VSS# show inteface vsl

VSL Port-channel: Po1
 Port: Te1/5/4
 Port: Te1/5/5
VSL Port-channel: Po2
 Port: Te2/5/4
 Port: Te2/5/5

6500-VSS(config)# monitor session 2 source int po1
6500-VSS(config)# monitor session 2 destination int gi1/4/10

6500-VSS# show monitor session 2
```

```
Session 2
---------
Type                   : Local Session
Source Ports           :
    Both               : Po1
Destination Ports      : Gi1/4/10

Egress SPAN Replication State:
Operational mode       : Centralized
Configured mode        : Centralized (default)
```

As shown in the preceding output, you can monitor the VSL port-channel by spanning it to a switch where that port-channel is local. See the following output example illustrating an attempt to create port monitoring with a destination belonging to the peer (remote) switch. This restriction removes the possibility of looping traffic and avoids over-utilization of VSL links. The port-channel interface numbers are usually created with matching switch number IDs and thus you can easily identify the affinity of a port-channel interface with a switch number.

```
6500-VSS# show interface vsl
VSL Port-channel: Po1
 Port: Te1/5/4
 Port: Te1/5/5

VSL Port-channel: Po2
 Port: Te2/5/4
 Port: Te2/5/5

6500-VSS(config)# monitor sess 1 source int po2
6500-VSS(config)# monitor sess 1 destination int gi1/4/10
% VSL cannot be monitor source with destination on different core
```

Note that the *Giants* counters on the VSL interface (see the output that follows) might lead to an assumption that something is wrong. In fact, this is a normal output. The reason that the interface counters notice the giants is due to fact that VSL inter-switch control frame packets are sent at 1518 bytes + 32 byte of DBUS header between the active and hot-standby switches. Such oversized packets are seen as giants on VSL EtherChannel.

```
6500-VSS# show switch virtual link counters
.
.
! <snip>

Port          Single-Col Multi-Col  Late-Col Excess-Col Carri-Sen     Runts      Giants
Po1                    0         0         0          0         0         0    19788377
Te1/2/8                0         0         0          0         0         0          34
Te1/5/4                0         0         0          0         0         0    19788414
<snip>

Port          Single-Col Multi-Col  Late-Col Excess-Col Carri-Sen     Runts      Giants
Po2                    0         0         0          0         0         0      693910
Te2/2/8                0         0         0          0         0         0          89
Te2/5/4                0         0         0          0         0         0      693821
```

Stateful Switch Over—Unified Control Plane and Distributed Data Forwarding

Stateful Switch Over Technology

The Stateful Switch Over (SSO) technology enables supervisor redundancy in a standalone Cisco Catalyst 6000 Series platform. SSO keeps the necessary control plane and protocol states replicated to the backup supervisor. As a result, if an active supervisor fails, a hot-standby supervisor has enough information about that system and network to continue forwarding packets and to continue in network protocol participation with the rest of the network devices. The dual supervisor-enabled system goes through various states during power-up. During initialization, Cisco IOS determines whether the system has dual supervisors, determines the hardware mode—*simplex* (single supervisor) or *duplex* (dual supervisor), and identifies which supervisor assumes the active or hot-standby role. Cisco IOS software also checks the software version on each supervisor and determines whether the state of the supervisor can enter into SSO or RPR mode. Each supervisor follows the redundancy facility (RF) states described in Table 2-3, depending on the adopted role (active or hot standby). For the supervisor that is elected as *primary,* the supervisor transitions from lowest (disabled) to highest (active) mode after successful SSO startup. The hot-standby supervisor goes through a separate state transition as described in Table 2-3 under the heading *States when becoming Standby-hot.*

Table 2-3 RF States

RF States and Code	RF State Activity
Common States to both supervisor	
RF_UNKNOWN = 0,	Unknown redundancy state; for example, supervisor booting
RF_DISABLED = 1,	Redundancy is disabled; for example, no dual supervisor exists
RF_INITIALIZATION = 2,	First phase of sync between supervisors
RF_NEGOTIATION = 3,	Discovery mode and who becomes active or hot-standby
States when becoming Standby-hot	
RF_STANDBY_COLD = 4,	State on non-active supervisor, peer is active, RPR state
RF_STANDBY_CONFIG = 5,	Sync config from active to hot-standby
RF_STANDBY_FILESYS = 6,	Sync file system from active to hot-standby
RF_STANDBY_BULK = 7,	Protocols (client) state—bulk sync from active to hot-standby
RF_STANDBY_HOT = 8,	Standby ready to be active and getting updates from active
States when becoming ACTIVE	
RF_ACTIVE_FAST = 9,	Immediate notification of hot-standby going active
RF_ACTIVE_DRAIN = 10,	Client clean up—drain queued messages from peer
RF_ACTIVE_PRECONFIG = 11,	Pre-processing configuration, boot environment
RF_ACTIVE_POSTCONFIG = 12	Post-processing the configuration
RF_ACTIVE = 13,	Control and data plane active and participating with network

Among these thirteen states, *13-Active* and *8-Standby-Hot* are critical for determining operational redundancy. These are summarized in the following brief descriptions:

- *State 13-ACTIVE*—In this *active* state, the supervisor is responsible for packet forwarding and managing the control plane. The control plane functions includes handling Layer-3 routing protocol, Layer-2 protocols (STP, BPDU), management—such as Telnet, Simple Network Management Protocol (SNMP), and secure shell (SSH), link and device management (SPAN and CDP), and so on. The active supervisor synchronizes configuration with the secondary supervisor. Finally, the active supervisor synchronizes the state and database of the protocols to the secondary supervisor once the hot-standby supervisor assumes the state of *Standby-HOT* (hot standby).
- *State 8-Standby-Hot*—In this hot-standby state, the supervisor is fully synchronized with the active supervisor and is capable of assuming the active role when needed. This is the final state of the hot-standby supervisor. In this state, each SSO-aware protocol, based on relevant events (such as interface state change, MAC update/change/up/down, and so on), triggers a message from the active supervisor to the hot-standby supervisor. Whenever the primary active supervisor fails for some reason, the protocol state on the hot-standby supervisor goes into the execution (run) state. For example, Cisco Express Forwarding (CEF) is a SSO-aware client. Whenever a change in the CEF's table occurs, the hot-standby supervisor receives an update. This ensures that when the hot-standby unit becomes active, the updated copy of the forwarding information base (FIB) can forward data packet in the hardware, while the control plane undergoes the recovery process. For more information on SSO, refer to following URL: http://www.cisco.com/en/US/docs/switches/lan/catalyst6500/ios/12.2SXF/native/configuration/guide/nsfsso.html

SSO Operation in VSS

The SSO is the core capability that enables VSS high availability. SSO operational and functional support is similar to standalone node operating with a dual supervisor. The two major differences are as follows:

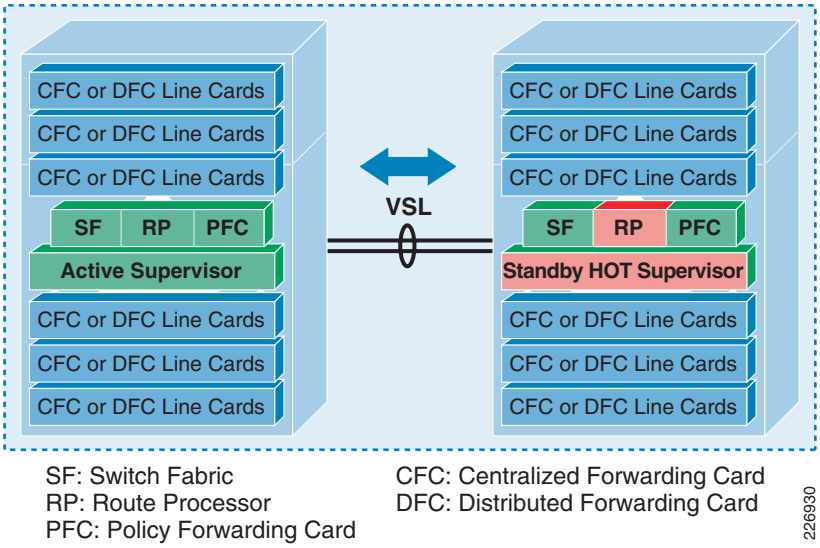
- SSO operation is extended over two chassis, where one supervisor is elected as the active supervisor and the supervisor in the other chassis is designated as the hot standby. This function is defined as *inter-chassis SSO*. See Figure 2-13.
- The packet forwarding occurs on both chassis and supervisors, hence the VSS is a *dual forwarding* solution, although the control plane is managed by only one supervisor.

The SSO operation in the VSS has the same dependency as in the case of a standalone environment. The inter-chassis SSO mode requires identical hardware and Cisco IOS software in both member chassis. Please refer to the URL below for the detailed list of dependencies for the SSO redundancy mode:

http://www.cisco.com/en/US/docs/switches/lan/catalyst6500/ios/12.2SX/configuration/guide/vss.html#wp1059586

Note Cisco IOS 122.(33)SXI removes some of the Cisco IOS software versioning dependencies. Refer to the appropriate release note for more information.

Figure 2-13 Inter-Chassis SSO Operation



Unified Control Plane

As shown in Figure 2-13, one supervisor is actively providing the unified control plane. The chassis carrying the active supervisor is called the *active virtual switch*. In the active switch all three components are active (green)—Switch Fabric (SF), Route Processor (RP), and Policy Forwarding Card (PFC). The active supervisor is also responsible for programming the hardware forwarding information onto all the distributed forwarding cards (DFC) across the entire VSS—as well as programming the policy feature card (PFC) on the hot-standby virtual switch supervisor engine. The unified control plane is responsible for the origination and termination of the traffic types described in the "Network Control Plane Traffic" section on page 2-16—as well as being the sole control-point for maintaining and managing inter-switch communications described in the "VSS Inter-Switch Communication and Layer-2 per Link Control Traffic" section on page 2-16.

Note With the first release of VSS, only single supervisors are supported per physical chassis. Therefore, there is no intra-chassis supervisor engine redundancy. A subsequent software release might offer the ability to add a second supervisor engine into each chassis.

Distributed Data Forwarding

As show in Figure 2-13, both supervisor resources (SF and PFC) are active for user-data forwarding. The Policy Feature Card (PFC) and switching fabric (backplane connectivity for fabric-enabled module) of both supervisors are actively forwarding user data and performing policy functions, such as applying access control lists (ACL) and QoS in hardware. Additionally, all Distributed Forwarding Cards (DFC) can also simultaneously perform packet lookups across the entire VSS. Because the switching fabrics of both switches are also in an active state, the Cisco VSS has the switch fabric capacity of 1440 (720 Mbps x 2) Gbps, or 1.44 Tbps in aggregate.

The active and hot-standby supervisors run in a synchronized mode in which the following system information is synchronized over the VSL link:

- Boot environment
- Synchronization of the running configuration
- Protocol states and the database table—Only protocols capable of supporting SSO redundancy (SSO-aware) are fully capable of supporting SSO-based recovery
- Line card status (interface state table and its capabilities)

During the initialization phase, the hot-standby supervisor undergoes configuration synchronization (RF_STANDBY_CONFIG = 5) with the active supervisor (see Table 2-3). Having an understanding of this configuration synchronization is important when considering switch failures detailed in the "Campus Recovery with VSS Dual-Active Supervisors" section on page 4-18.

Both active and hot-standby switches can learn the address simultaneously; however, the active virtual switch manages the network information from adjacent devices (such as MAC, STP, or CEF). Several protocols are SSO-aware such that active switch synchronizes protocol information (database, protocol state) to the hot-standby supervisor. In addition, the active supervisor manages and updates the information about interfaces and line card status on both chassis.

The state of the supervisor's control plane (active, hot-standby, or any other state) can be checked using the following CLI commands. Notice that the fabric state is active in both the chassis, indicative of the dual forwarding state.

```
6500-VSS# show switch virtual
Switch mode : Virtual Switch
Virtual switch domain number : 200
Local switch number : 1
Local switch operational role: Virtual Switch Active
Peer switch number : 2
Peer switch operational role : Virtual Switch Standby

6500-VSS# show switch virtual redundancy
My Switch Id = 1
Peer Switch Id = 2
Last switchover reason = none
Configured Redundancy Mode = sso
Operating Redundancy Mode = sso
Switch 1 Slot 5 Processor Information :
-----------------------------------------------
Current Software state = ACTIVE
Uptime in current state = 3 weeks, 4 days, 9 minutes
Image Version = Cisco IOS Software, s72033_rp
Software (s72033_rp-ADVENTERPRISEK9_WAN_DBG-M), Version
12.2(SIERRA_INTEG_070502) INTERIM SOFTWARE
Synced to V122_32_8_11, 12.2(32.8.11)SR on rainier, Weekly
12.2(32.8.11)SX76
Technical Support: http://www.cisco.com/techsupport
Copyright (c) 1986-2007 by Cisco Systems, Inc.
Compiled Thu 03-May-07 09:46 by kchristi
BOOT = sup-bootdisk:s72033-
adventerprisek9_wan_dbg-mz.SIERRA_INTEG_070502,1;
CONFIG_FILE =
BOOTLDR =
Configuration register = 0x2102
Fabric State = ACTIVE
Control Plane State = ACTIVE
Switch 2 Slot 5 Processor Information :
-----------------------------------------------
Current Software state = STANDBY HOT (switchover target)
Uptime in current state = 3 weeks, 4 days, 8 minutes
```

```
Image Version = Cisco IOS Software, s72033_rp
Software (s72033_rp-ADVENTERPRISEK9_WAN_DBG-M), Version
12.2(SIERRA_INTEG_070502) INTERIM SOFTWARE
Synced to V122_32_8_11, 12.2(32.8.11)SR on rainier, Weekly
12.2(32.8.11)SX76
Technical Support: http://www.cisco.com/techsupport
Copyright (c) 1986-2007 by Cisco Systems, Inc.
Compiled Thu 03-May-07 09:46 by kchristi
BOOT = sup-bootdisk:s72033-
adventerprisek9_wan_dbg-mz.SIERRA_INTEG_070502,1;
CONFIG_FILE =
BOOTLDR =
Configuration register = 0x2102
Fabric State = ACTIVE
Control Plane State = STANDBY
```

Note Cisco IOS software images on both supervisors must match otherwise the standby supervisor will boot in RPR mode and the line card will not be active on that chassis. Refer to following publication for further information on RPR: http://www.cisco.com/en/US/docs/switches/lan/catalyst6500/ios/12.2SXF/native/configuration/guide/redund.html

Virtual Switch Role, Priorities and Switch Preemption

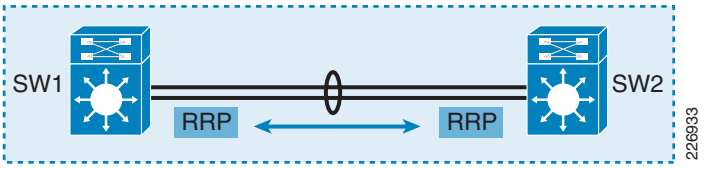
Role Resolution Protocol (RRP)

Having a unified control plane and supporting distributed forwarding with more than one switch requires some form of protocol to determine which switch should (or can) become active, how to change the default setting, and how to deterministically configure which switch member will become active. VSS has a dedicated protocol called Role Resolution Protocol (RRP) that is used to define such behavior. See Figure 2-14.

RRP protocol is used to determine the SSO role (active, hot-standby, or RPR), and to negotiate switch priority and preemption of virtual switch. RRP also checks the software version on each switch which must be the same in order to form a VSS.

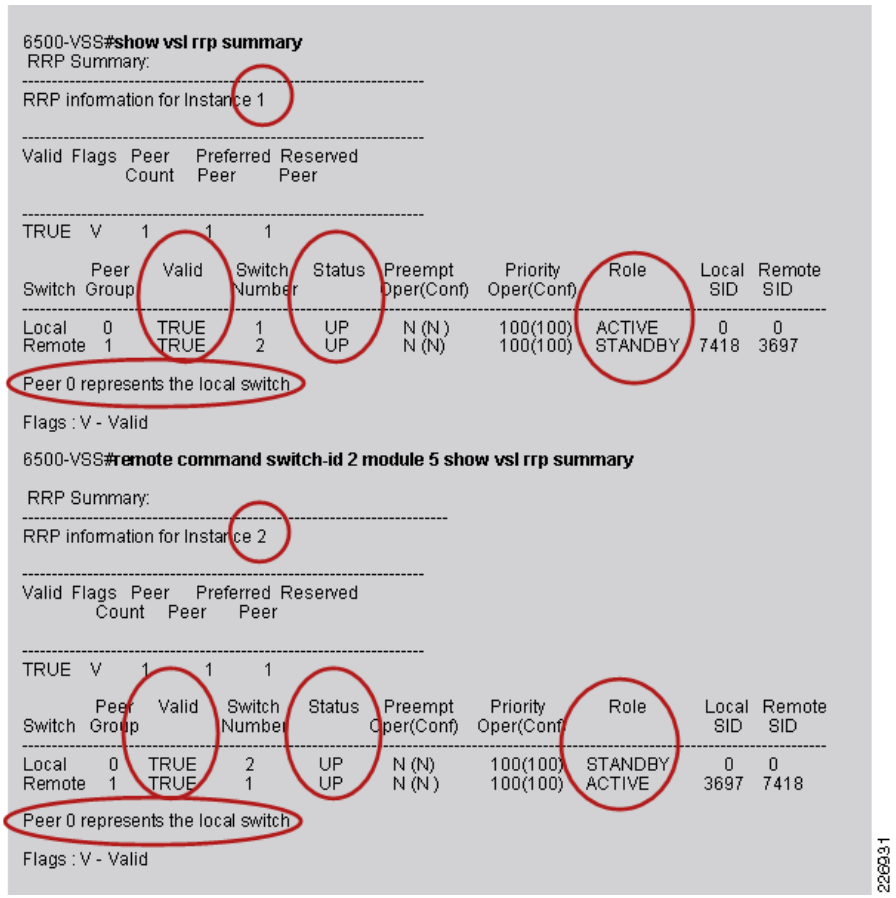
The RRP protocol is initialized once Link Management Protocol (LMP) is fully established on at least one VSL port. The LMP control link is selected by RRP protocol to negotiate the SSO role and switch priority. Each switch member forms local RRP peer group instance and communicate over the control link of the VSL bundle instead running on every VSL link. RRP negotiation packets are encapsulated in the same format as the LMP protocol. Hence, RRP packets are placed into the transmit priority queue in order to prioritize it over data traffic.

Figure 2-14 RRP Negotiation for Role Selection



RRP protocol status can be verified using commands illustrated in Figure 2-15. As shown in the output, the **remote command** *switch-id* command is needed to verify the RRP state of the hot-standby switch. The Peer Group is always 0 on the local switch and 1 for the neighbor switch—regardless of switch ID, priority, preemption, or current SSO role.

Figure 2-15 RRP Protocol Status Information



The RRP session between virtual-switch is negotiated in following conditions:

- When both virtual switches in virtual-switch domain are in bootup mode
- When the hot-standby switch is in the bootup process while the active switch is operational
- When in the dual active recovery phase and the VSL link is restored between virtual switch members

The RRP session is briefly established during the preceding conditions; however, RRP does not maintain a session between two switch members, instead RRP relies on internal notifications from LMP on each switch member. If all VSL member links fail, RRP does not have any control-link path for communication between peers; the LMP protocol on each virtual switch will delete the peer group and notify the RRP process to take SSO-based switchover action. In the absence of VSL link, no RRP action is taken on the active switch if it is still operational; however, the hot-standby switch will transition to the active role because it has no way to determine the remote peer status.

Virtual Switch Priority

The selection of the switch member that is to become the active or hot-standby switch depends on the order and the way in which the switches are initialized. If the switch members boot simultaneously, the switch with the lowest switch ID becomes the active virtual switch. If each switch member boots at different times, then the switch that is initiated first becomes the active virtual switch regardless of switch ID. The VSS can also be configured with switch priority. Usually, the default switch priority (100) need not change; however, you can configure switch priority for two possible requirements:

- You are required to override the default selection of the switch role after initial configuration. This might be due to a need for operational flexibility or requirement for monitoring access. However, the software will not immediately enforce the configured switch priority. Adopting this change requires a reload of both switches in order for the RRP to enable the designated switch to become active. Figure 2-16 illustrates the commands associated with making a change to the switch priority and a syslog message indicating that a reload is required to enforce the new designated priority. If the switch with lower priority comes up first, the software will not force the takeover of the active role when the switch with a higher priority boots up later—unless that switch is configured with preemption.

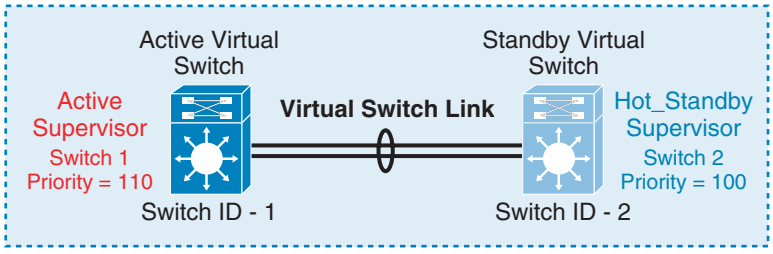
Figure 2-16 Changing Switch Priority

```
6500-VSS#show switch virtual role

Switch  Switch  Status  Preempt      Priority      Role       Session ID
        Number          Oper(Conf)   Oper(Conf)              Local   Remote
-------------------------------------------------------------------------------
LOCAL     1       UP    FALSE(N)     100(100)      ACTIVE      0       0
REMOTE    2       UP    FALSE(N)     100(100)      STANDBY  3071    3108
6500-VSS#conf t
Enter configuration commands, one per line.  End with CNTL/Z.
6500-VSS(config)#switch virtual domain 1
6500-VSS(config-vs-domain)#switch 2 priority 120

Sep 10 11:35:08.945: %VSLP-SW2_SPSTBY-5-RRP_RT_CFG_CHANGE: Configured priority value is different from operational value. Change will take
      effect after config is saved and switch is reloaded.

6500-VSS#show switch virtual role
Switch  Switch  Status  Preempt      Priority      Role       Session ID
        Number          Oper(Conf)   Oper(Conf)              Local   Remote
-------------------------------------------------------------------------------
LOCAL     1       UP    FALSE(N)     100(100)      ACTIVE      0       0
REMOTE    2       UP    FALSE(N)     100(120)      STANDBY  3071    3108
```

226934

- Defining the priority of the switch to match the switch ID will show up in the configuration, which can provide a visual aid in change management. This option is shown in Figure 2-17, where the switch priority is configured to match the switch ID (matching high priority to lowest switch ID).

Figure 2-17 Switch Priority Configured to Match Switch ID



227083

Switch Preemption

If the intention is to select one of the switches to be in the active role in all conditions, then simply increasing switch priority will not achieve this goal. For a deterministic role selection, regardless of the boot order, the desired active switch must be configured with a higher switch ID and switch preemption. See Figure 2-18.The CLI allows the preemption feature to be configured on a low-priority switch. However, the switch preemption setting will not be effective.

Figure 2-18 Configuring Preemption

```
6500-VSS#conf t
6500-VSS#(config)#switch virtual domain 1
6500-VSS#(config-vs-domain)#switch 2 priority 120

Sep 15 17:03:24.468: %VSLP-SW2_SPSTBY-5-RRP_RT_CFG_CHANGE: Configured
priority value is different from operational value. Change will take effect after config is
saved and switch is reloaded.

6500-VSS#(config-vs-domain)#switch 2 preempt

Please note that Preempt configuration will make the ACTIVE switch with lower priority to
reload forcefully when preempt timer expires

Sep 15 17:03:30.864: %VSLP-SW2_SPSTBY-5-RRP_RT_CFG_CHANGE: Configured
preempt value is different from operational value(s).
Change will take effect after config is saved and switch is reloaded.

6500-VSS#show vsl rrp summary
 RRP Summary:
-------------------------------------------------------------
RRP information for Instance 1

-------------------------------------------------------------
Valid   Flags  Peer    Preferred   Reserved
               Count   Peer        Peer
-------------------------------------------------------------
TRUE    V      1       1           1

        Peer  Valid   Switch  Status  Preempt      Priority      Role     Local  Remote
Switch  Group         Number          Oper(Conf)   Oper(Conf)             SID    SID
-------------------------------------------------------------------------------------
Local   0     TRUE    1       UP      N (N)        100(100)      ACTIVE   0      0
Remote  1     TRUE    2       UP      N (Y*)       100(120)      STANDBY  3790   7230

Peer 0 represents the local switch

Flags : V - Valid
```

226935

The implication of switch preemption on the network availability is significant and must be evaluated before deployment. The operational impact of switch preemption should not be compared with widely understood HSRP/GLBP protocol behavior in which preemption only allows relegating the role of being the active HSRP/GLBP forwarder to being in standby mode without much impact to the network (no reload or reset of the chassis).

Switch preemption forces multiple reboots of the VSS member—leading to multiple outages, reducing forwarding capacity while the switches decide which supervisor should assume an active role. For example, if the switch that is configured with preemption fails (or otherwise loses connectivity), then the peer switch will assume the role of being active temporarily. When the preemptive switch boots up and finds that its peer switch is active, the preemptive switch will force the newly active peer switch to give up the role of being active. The only way for the peer switch to give up the active role is by resetting and transitioning to the hot-standby role. Similarly, if the non-preemptive (designated hot-standby) switch somehow comes up first (either due to power failure or delayed user action) and assumes an active role, it will be forced to reset when preemptive switch is brought on line.

Tip **Cisco recommends that you do *not* configure switch preemption for the following reasons:**

- It causes multiple switch resets, leading to reduced forwarding capacity and unplanned network outages.
- The VSS is a single logical switch/router. Both switch members are equally capable of assuming the active role because it does not matter which is active—unless required by enterprise policy.

Virtual Switch Member Boot-Up Behavior

The normal VSS boot-up process consists of diagnostics, VSL link initialization, LMP establishment, and switch role negotiation via RRP. RRP determines the role of each switch leading to the SSO state in which one switch is active and the peer is in the hot-standby state. However, the behavior and the end result are different if there are problems or events leading to VSL interface being inactive/disabled/failed after the RRP negotiation phase in which each switch role was assigned. The VSL link typically becomes disabled for one of two primary reasons:

- Due to the VSL interface and/or line card being non-operational.
- The switch that assumed the role of being active has problems leading to reset of the switch and thus the VSL interface is non-operational.

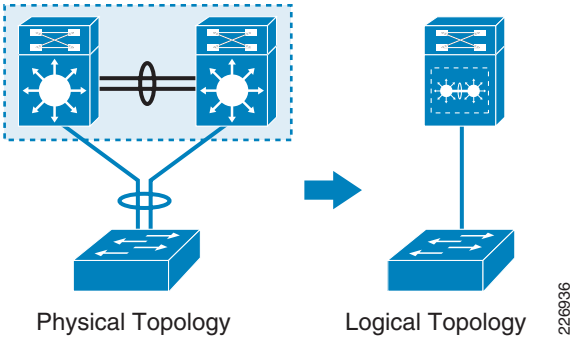
In both cases, the peer switch that has assumed the role of hot standby cannot continue booting, as there is no way to determine (or exchange) the state of the peer switch (due to the VSL being down). The Cisco IOS software will issue a forced reset (described as *crashed* in Cisco IOS parlance) and the peer switch will restart the boot process. If the peer (hot-standby) switch finds that VSL link is still down it will continue booting. In the absence of RRP, it will assume the role of active. This situation can lead to dual-active condition because neither switch has knowledge of the other. For this reason, dual active detection is a mandatory when deploying VSS in campus network. Please refer to the "Campus Recovery with VSS Dual-Active Supervisors" section on page 4-18 for more details.

Multi-chassis EtherChannel (MEC)

Traditional EtherChannel aggregates multiple physical links between two switches. MEC is an advanced EtherChannel technology that extends link aggregation to span over two separate switches. VSS allows for distributed forwarding and a unified control plane so that the MEC appears as single port-channel interface existing on both the active and hot-standby switches. Even though the access-layer is connected to a distinct physical chassis via two physical links, from an access-layer switch perspective, this port-channel connection enables a single logical link connected to a single logical switch (referred to as *VSS with MEC*). Figure 2-19 depicts a physical-to-logical transformation in which the logical topology is simplified and (for the spanning tree) VSS with MEC offers no-loops.

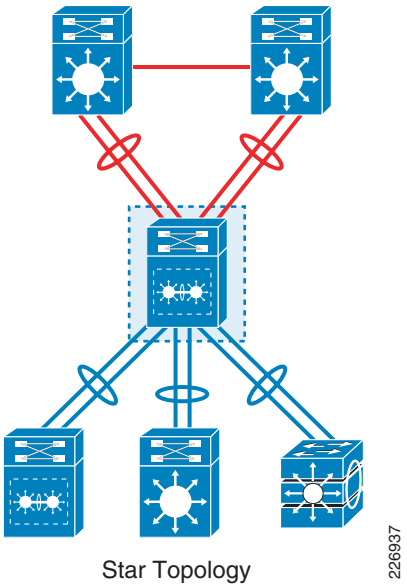
Note MEC configuration is only possible in the VSS; however, access-layer switches requiring connectivity to the VSS are configured with traditional EtherChannel interfaces.

Figure 2-19 MEC—Physical vs. Logical Topology



This capability of spanning EtherChannel over multiple switches as a virtualized, single logical switch creates a topology in which all devices connected via MEC to VSS appear as a star-shaped topology. See Figure 2-20.

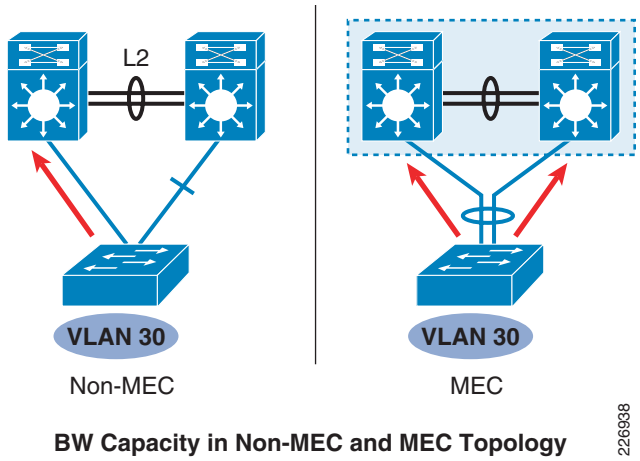
Figure 2-20 Star Topology



The MEC and VSS bring powerful and very effective changes to the campus topology. Two key benefits are as follows:

- *Eliminates loops in multilayer design*—Traditionally, spanning VLANs over multiple closets would create a STP-looped topology because one of the uplinks would be blocked by STP (see Figure 2-21 and Figure 1-4). MEC with VSS together eliminate loops in the campus topology. This is because STP now operates on the EtherChannel logical port and each physical switch appears to be connected via a single logical link to a single logical switch. From the STP viewpoint, this star topology is non-looped. No alternate path exists for STP to be blocked. Figure 2-21 depicts two topologies: The access-layer to VSS topology without MEC in which the uplink is blocked and the access layer to VSS topology with MEC in which both links are forwarding without looping. The advantages to a loop-free network are described in Chapter 3, "VSS-Enabled Campus Design."

Figure 2-21 Bandwidth Capacity in non-MEC and MEC Topologies



BW Capacity in Non-MEC and MEC Topology

- *Doubles the available bandwidth for forwarding*—MEC replaces spanning tree as the means to provide link redundancy. This means that all physical links under the MEC are available for forwarding traffic. The STP can no longer block individual links since its database does not have those links available to calculate a loop-free path. For the network with a looped topology, the total forwarding capacity is half the available bandwidth of physical links. VSS with MEC makes all links available for forwarding and thus doubles the bandwidth available. This offers is an effective change to an existing network in which the lack of a 10-Gbps infrastructure requires choosing a design alternative (EtherChannel on each uplink, routed access, multiple HSRP group, and so on) to efficiently utilize the available links. For any new design, VSS with MEC enables these benefits with simplified topology.

Why MEC is Critical to VSS-Enabled Campus Design

It is important to understand why MEC is critical in a VSS design, before going into detail about MEC functionality, its operation, traffic flow, and implications in campus design. You can have a VSS-enabled network without MEC; however, the resulting topology will consist of either a single point-of-failure (only one link to adjacent network devices) or a looped topology because both links from a given networking device will be connected to a single logical VSS switch in a non-EtherChannel configuration. Either condition reduces the benefits of VSS in any given network. Chapter 3, "VSS-Enabled Campus Design." provides several design proof points to justify the importance of the MEC-enabled design. The following are the major advantages of an MEC-enabled design:

- Enables loop free topology.
- Doubles the available forwarding bandwidth, resulting in reduced application response time, reduced congestion in the network, and reduced operation expenditure.
- Reduces or eliminates control plane activities associated with a single-link failure (either nodal failure or member link failure).
- Allows failure detection in hardware for faster convergence of traffic flows.
- Reduces MAC learning and because the failure of one link does not trigger the Layer-2 control plane convergence.
- Reduces routing protocol announcements—adding efficiency in the Layer-3 network (reduces need for summarization and reduces control plane activity)
- Avoids multicast control plane changes (avoids multicast incoming interface changes), because the failure of one link does not trigger Layer-3 control plane convergence

- Avoids multicast replication over the VSL in an MEC-enabled topology.
- Enables flexibility in deploying dual-active detection techniques (see the "Campus Recovery with VSS Dual-Active Supervisors" section on page 4-18).

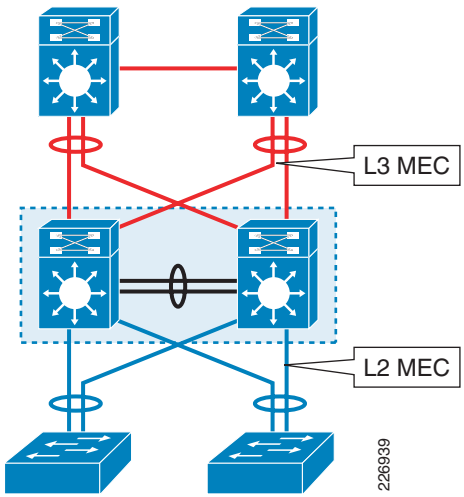
MEC Types and Link Aggregation Protocol

MEC is a distributed EtherChannel environment; however, it inherits all the properties of EtherChannel used in a traditional network. The rest of this section illustrates some, but not all, aspects of EtherChannel technology that apply to MEC. The coverage of EtherChannel here addresses features that might not be obvious or available in other publications—focusing on capabilities that are considered necessary for VSS implementation. Some featured illustrations and protocol configurations are repeated to help readers better understand MEC deployment in the context of a VSS environment.

Types of MEC

Depending upon network connectivity requirements, MEC configuration consists of two modes: Layer 2 and Layer 3. See Figure 2-22.

Figure 2-22 MEC Types



Layer-2 MEC

In a hierarchical, three-layer network, the Layer-2 MEC applies to the connection between the access layer and the distribution layer (where VSS is enabled) as shown in Figure 2-22. In this mode, the Layer-2 MEC is participating in STP, MAC-address learning, and a host of other Layer-2 operations. Layer-2 MEC enabled connectivity can be extended to create large hierarchical Layer-2 topology in which the entire network is loop free. This document does not cover such designs.

Layer-3 MEC

Layer-3 MEC is comprised of a routed port-channel interface. With Layer-3 MEC, the port-channel interface is configured as a routed interface bearing the IP address that participates in Layer-3 functions, such as routing and forwarding using CEF. The natural extension of this type of connectivity is to have multiple, Layer-3 VSS pairs connected together to form the basis for a routed design. The routing application of Layer-3 MEC is discussed in the "Routing with VSS" section on page 3-44.

Link Aggregation Protocol

The EtherChannel is a logical interface. Managing the behavior of a physical member link with (and operational impact on) the rest of the network requires some form of control plane. Two methods available to manage control plane of the underlying link in an EtherChannel group are as follows:

- Port Aggregation Protocol (PAgP)
- Link Aggregation Control Protocol (LACP) or IEEE 802.3ad

Each of these protocols provides the following common functions:

- Ensures link aggregation parameter consistency and compatibility between the VSS and a neighbor switch
- Ensures compliance with aggregation requirements
- Dynamically reacts to runtime changes and failures on local and remote EtherChannel configurations
- Detects and removes unidirectional link connections from the EtherChannel bundle

The EtherChannel is the fundamental building block in a VSS-enabled campus design. The successful deployment of MEC requires operational consistency and interaction with several access-layer switches creating topology that does not introduce unexpected behavior. For these reasons, the MEC interface must be enabled with either PAgP or LACP in order to benefit from functionality described in the preceding discussions. As with traditional EtherChannel, having PAgP or LACP enabled on an MEC provides consistency checks the system configuration and the operational state of the underlying physical links. PAgP and LACP remove a mismatched link from the bundle and thereby provide additional protection against mismatched configurations and operational verification via syslog messages. The following configuration must match in all underlying physical links participating in an MEC bundle.

- Configuration of VLANs on member links
- Trunk type and configuration on member links
- Port status (full- or half-duplex) and QoS support by underlying hardware must be the same in all links

For a complete list of the requirements for forming an EtherChannel, please refer to the individual product release notes and related documentation at www.cisco.com. PAgP and LACP design considerations (as applicable to VSS) that can affect network design are described in "PAgP" and "LACP (IEEE 802.3ad)" sections that follow.

PAgP

PAgP is the mostly widely used link aggregation protocol. PAgP is supported by most Cisco devices and other third-party network interface cards (NIC). In the VSS context, PAgP provides the value-added function of providing assistance in a dual active recovery—in addition to function summarized in the preceding section. PAgP runs on each MEC link member, establishes and manages neighbor adjacency between the VSS and a Layer-2 or Layer-3 partner. PAgP determines multiple attributes on a per member link basis, such as peer-switch state, device name, partner-port, port-aggregation capability, and port-channel bundle group index.

Device ID

The active switch is responsible for origination and termination of PAgP control-plane traffic. PAgP advertises a device-id in order to identify each end of a connection with a unique device status. For VSS, it is a 48-bit number in MAC address format (02.00.00.00.00.*xx*). It consists of a combination of a fixed prefix (02.00.00.00.00) in first five octets and a variable value (*xx*) for last octet. The variable part is the virtual switch domain identifier configured for the system. This PAgP device-id is sent by the two links terminated on two switches that make up a VSS domain. Because the device-id is the same, a remote device assumes that the device-id is coming from the single logical device (the VSS). Even during role switchover in VSS, the device-id remains consistent on each PAgP neighbor to prevent a renegotiation process. This use of the virtual switch domain identifier in PAgP and LACP requires that the identifier to be unique in all VSS domains that are interconnected through the use of MEC. The following command output examples illustrate the device-ID value described in this section, showing the fixed and variable components.

```
6500-VSS# sh run | inc virtual
switch virtual domain 10 ! <-- Device ID
6500-VSS# show pagp neighbor
Flags:S - Device is sending Slow hello.C - Device is in Consistent state.
       A - Device is in Auto mode.        P - Device learns on physical port.

Channel group 10 neighbors
          Partner       Partner          Partner           Partner     Group
Port      Name          Device ID        Port         Age  Flags       Cap.
Gi1/1     6500-VSS      0200.0000.000a   Gi1/4/1      7s   SC          A0001
Gi1/2     6500-VSS      0200.0000.000a   Gi2/4/1      8s   SC          A0001
```

Modes of PAgP Operation

There are many configuration options for PAgP. The best practices for configuring PAgP for EtherChannel are documented in the publication at the following URL:

http://www.cisco.com/en/US/products/hw/switches/ps700/products_white_paper09186a00801b49a4.shtml

This design guide provides selected details from these documents and integrates recommendations that are suited to the VSS-enabled campus environment.

Table 2-4 shows only the best practice-based configuration options. Out of the three choices shown, the recommended mode for PAgP neighbors is *desirable-desirable*. This configuration option enables PAgP on both sides and forces the consistency check mentioned previously in this section. The desirable-desirable option is the best option for ensuring safe and reliable operation of MEC-connected devices. Unlike LACP, PAgP offers strict channel settings and configuration checks prior to bundling the ports. MEC bundle remains in disabled state if PAgP+ detects a configuration mismatch and remains disabled until the error is fixed. This additional measure prevents EtherChannel inconsistency, which would otherwise create operational challenges for managing large-scale EtherChannel deployments.

Table 2-4 Best Practice-based Configuration Options for PAgP

Channel Mode—For both Layer-2 and Layer-3 MEC	VSS	Remote Node	MEC State
	desirable	desirable	operational
	desirable	auto	
	auto	desirable	

The additional configuration options of *silent* or *non-silent* are available based on the network requirements. For most network designs, silent mode is preferred. The silent mode integrates the link into a bundle whether the data is present or not. A non-silent option is used as an indirect measure of link integrity. For the best-practice design, UDLD is the preferred method of link integrity check. As a result, the silent option is more applicable to most network implementations.

Why You Should Keep the PAgP Hello Value Set to Default

By default, PAgP in non-silent mode independently transmits PAgP hello messages at an interval of one per 30 seconds on each member link of an MEC. This is known as *slow-hello* because it takes 105 seconds (3.5 hello intervals) to detect remote partner availability. This timer can be modified so that the PAgP hello is sent every second, which is known as *fast-hello*. Many network designs tend to use this option to accelerate link detection, since UDLD can take longer then fast-hello (3 times 1 second). However, a fast-hello configuration should be avoided in VSS deployment for the following two reasons:

- The VSS control plane might not recover (during the SSO switchover) in 3 seconds, so that the VSS can send a PAgP hello before the remote end declares VSS as being non-responsive. This can lead to false positive
- A fast-hello is sent on a per link basis. For a large-scale deployment, fast-hello transmissions can overrun a switch CPU.

Note Even though one side of the EtherChannel is configured with fast-hello and other side (or device) is configured with slow-hello, operationally they will transmit and receive fast-hellos between devices. This means that even though VSS is configured with slow-hello, improper configuration on remote devices can alter the operational mode of hello messaging.

Tip **The best practice is to keep the PAgP timer settings to default values and to use the normal UDLD to monitor link integrity.**

LACP (IEEE 802.3ad)

LACP is an industry standard port-aggregation protocol that allows for multi-vendor interoperability. LACP is very similar to PAgP in terms of functionality and operation. In VSS, it works for both Layer-2 and Layer-3 MEC interfaces. The following URL provides details of LACP operation and configuration options:

http://www.cisco.com/en/US/products/hw/switches/ps700/products_white_paper09186a00801b49a4.shtml

Device ID

LACP implementation on the VSS uses a pre-computed system-id. The LACP system ID consists of a 48-bit address that is a combination of a fixed prefix in the first five octets (02.00.00.00.00) and variable for last octet (*xx*). The variable part is the virtual switch domain identifier configured for the systems. The following output examples identify the system-id with virtual switch domain number used in last octet of system-id.

```
6500-VSS# sh lacp sys-id
32768,0200.0000.000a ! Device ID info
6500-VSS# sh run | inc virtual
switch virtual domain 10 ! Device ID info
```

LACP Mode of Operation

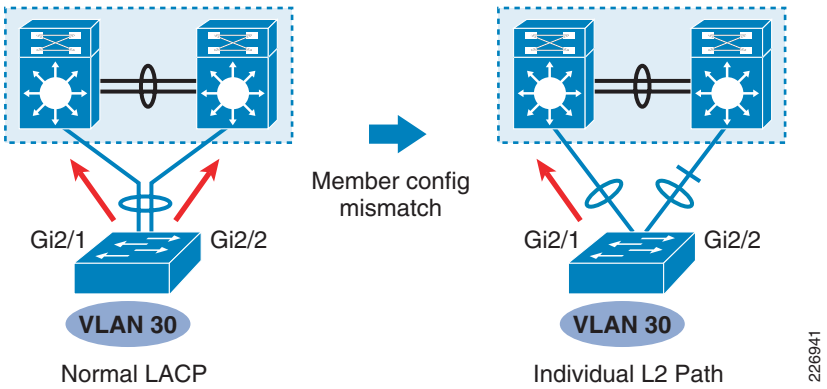
For the same reasons described in preceding PAgP description, Table 2-5 shows only the best practice-based configuration options. Our of the three choices shown, the recommended mode for LACP neighbors is *active-active*

Table 2-5 Best Practice-based Configuration Options for for LACP

Channel Mode—For both Layer-2 and Layer-3 MEC	VSS	Remote Node	MEC State
	active	active	operational
	active	passive	
	passive	passive	

The EtherChannel configured with LACP active-active option allows consistency of configuration on a member link as does PAgP; however, the end result is different. During the EtherChannel bundling process, LACP performs a configuration consistency check on each physical link trying to become a member of the port-channel. If the configuration check fails, a syslog message is generated. In addition, the system generates a special EtherChannel interface and that is assigned with unique alphabetical ID. The system generated LACP MEC will bundle all the physical ports into the MEC that failed the configuration check. See Figure 2-23.

Figure 2-23 LACP Configuration Mismatch Illustration



The following CLI output and configuration process illustrates this behavior. In following output example, the LACP link member is reconfigured with an inconsistent configuration. When a mis-configured interface is attempting to join the port-channel interface, the configuration checks trigger an unique system-generated, port-channel interface.

```
6500-VSS# show etherchannel 20 summary | inc Gi
```

```
Po20(SU)        LACP      Gi2/1(P)       Gi2/2(P)
6500-VSS# show spanning-tree | inc Po20
Po20               Root FWD 3         128.1667 P2p
6500-VSS# config t
6500-VSS(config)# int gi2/2
6500-VSS(config-if)# switchport nonegotiate
6500-VSS(config-if) # shut
6500-VSS(config-if)# no shut
%EC-SPSTBY-5-CANNOT_BUNDLE_LACP: Gi2/2 is not compatible with aggregators in channel 20
and cannot attach to them (trunk mode of Gi2/2 is trunk, Gi2/1 is dynamic)
%EC-SP-5-BUNDLE: Interface Gi2/2 joined port-channel Po20B ! A system generated
port-channel
6500-VSS# show etherchannel 20 summary | inc Gi
Po20(SU)        LACP      Gi2/1(P)
Po20B(SU)       LACP      Gi2/2(P) ! Bundled in separate system-generated port-channel
                                   ! interface

6500-VSS# show spanning-tree | inc Po20
Po20        Root FWD   4         128.1667     P2p
Po20B       Altn BLK   4         128.1668     P2p ! Individual port running STP is blocked
```

This creates two bundles:

- The first bundle is associated with the port that has succeeded in its configuration check.
- The second bundle is system generated and includes the port that did not match configuration.
- As a result, the control plane will be active on both port-channel interfaces (each having one member). The resulting topology consists of two distinct Layer-2 paths created between access-switch and the VSS as shown in Figure 2-23 (this is also true for Layer-3 MEC, but is not shown in the example). The STP topology will consider such network as looped and will block the port with higher STP priority. This is one of the major behavioral considerations for a network topology for LACP compared to PAgP. PAgP offers stricter channel settings and configuration checking prior to bundling the ports. In PAgP, the MEC remains disabled (state is shows as *down*) if PAgP+ detects a configuration mismatch—until the error is fixed.

Why You Should Keep the LACP Hello Value Set to Default

As with PAgP, LACP allows you to configure a hello interval from a default of 30 seconds (slow-hello) to 1 second (fast-hello). Unless the both ends of the LACP neighbor device connection are configured with identical configuration, the LACP hello can be sent at an asymmetric rate from either side. In other words, it is possible that a remote device connected to the VSS can send the LACP with slow-hello and can VSS send the LACP with fast-hello. If the fast-hello method is used for detecting a link failure, the detection time could also be variable based on configuration (30 seconds at one side and three seconds at other end). This is different from PAgP, where the hello transmission rate defaults to fast-hello on both sides (whenever a fast-hello request is made). A fast-hello configuration should be avoided in a VSS deployment for two reasons:

- The VSS control plane might not recover (during the SSO switchover) in 3 seconds (timeout for fast-hello), so that the VSS can send an LACP hello before the remote end declares VSS as being non-responsive. This can lead to false positive
- A fast-hello is sent on a per link basis. For a large-scale deployment, fast-hello transmissions can overrun a switch CPU.
- If the server connected to the VSS is configured based on LACP-MEC with fast-hello, then there is no inherent method to stop the VSS from sending the fast-hello. This can trigger excessive CPU usage when scaled to higher numbers of servers requesting such service.

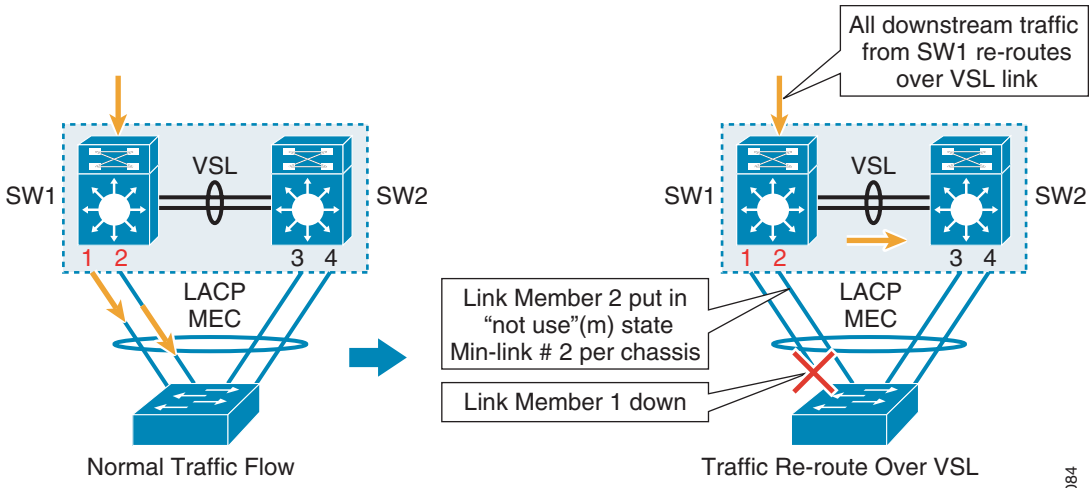
Tip **The best practice is to keep the LACP timer settings to the default values and to use the normal UDLD to monitor link integrity**.

LACP Minimum Link Behavior with VSS

The minimum link feature of LACP is used to maintain a level of bandwidth (in terms of number of interfaces in the *up/up* state) needed for a given application or networking requirement. If the number of links supporting a certain bandwidth falls below a minimum requirement, the bundle is taken out of service. In the standalone design, the minimum link feature is deployed between two separate nodes; the alternate node is available for forwarding traffic as needed. For the VSS, the behavior of the minimum link features is different. In VSS, the LACP EtherChannel minimum link configuration is maintained on a per port-channel basis—although its enforcement is on a per physical chassis basis. The following configuration example and Figure 2-24 illustrate application of this feature.

```
6500-VSS# show etherchannel 10 summary | inc Gi
10     Po10(SU)         LACP      Gi1/4/1(P)     Gi1/4/2(P)     Gi2/4/1(P)     Gi2/4/2(P)

6500-VSS# conf t
6500-VSS(config)# int po10
6500-VSS(config-if)# port-channel min-links 2
6500-VSS(config-if)# int gi1/4/1
6500-VSS(config-if)# shutdown
%LINK-5-CHANGED: Interface GigabitEthernet1/4/1, changed state to administratively down
%LINEPROTO-5-UPDOWN: Line protocol on Interface GigabitEthernet1/4/1, changed state to
down
%LINEPROTO-5-UPDOWN: Line protocol on Interface GigabitEthernet1/4/2, changed state to
down
%EC-SW2_SPSTBY-5-MINLINKS_NOTMET: Port-channel Po10 is down bundled ports (1) doesn't meet
min-links
%EC-SW1_SP-5-MINLINKS_NOTMET: Port-channel Po10 is down bundled ports (1) doesn't meet
min-links
—%LINEPROTO-SW1_SP-5-UPDOWN: Line protocol on Interface GigabitEthernet1/4/2, changed
state to down
%LINK-SW1_SP-5-CHANGED: Interface GigabitEthernet1/4/1, changed state to administratively
down
%LINEPROTO-SW1_SP-5-UPDOWN: Line protocol on Interface GigabitEthernet1/4/1, changed state
to down

6500-VSS# show etherchannel 10 summary
Flags:  D - down        P - bundled in port-channel
        M - not in use, no aggregation due to minimum links not met
        m - not in use, port not aggregated due to minimum links not met

10     Po10(SU)         LACP      Gi1/4/1(D)     Gi1/4/2(m)     Gi2/4/1(P)   Gi2/4/2(P)
```

Figure 2-24 Min-link Configuration Causing Reduced Capacity



Min-link Configuration Causing Reduced Capacity and Undesired Traffic Flow

For a MEC using the LACP control protocol, *minlinks* defines the minimum number of physical links in each chassis for the MEC to be operational. For an example, with the **port-channel min-links 2** configuration on the MEC, each virtual-switch member must match at least two operational local member link ports in order to be associated with the MEC. If one of the member links is down, the other member in the chassis will be put into the *not in use* state. The effect of this enforcement is that one link failure will cause the complete loss of connectivity from one switch member—even though other links are available for forwarding traffic—as shown in the preceding syslog output examples and Figure 2-24. This is will force rerouting of traffic over the VSL link—further congesting the two links that are in the bundle of the other chassis.

For a two-link port-channel configuration (typical access switch-to-VSS network topology), the minimum link feature is not applicable because it looks for the minimum two links per physical chassis and, if configured, will not allow the MEC to be operational. The following output example illustrates the behavior of LACP min-link configuration when a two-member EtherChannel is configured to connect to the VSS with each physical chassis having one member link.

```
6500-VSS# show etherchannel 150 sum
Flags:  D - down        P - bundled in port-channel
Group  Port-channel  Protocol    Ports
------+-------------+-----------+-----------------------------------------------
150    Po150(SU)       LACP      Gi1/7/1(P)     Gi2/7/1(P)

6500-VSS# sh spanning-tree int po 150

Vlan                Role Sts Cost      Prio.Nbr Type
------------------- ---- --- --------- -------- --------------------------------
VLAN0050            Desg FWD 3         128.1667 P2p
VLAN0150            Desg FWD 3         128.1667 P2p

6500-VSS# sh int po 150
Port-channel150 is up, line protocol is up (connected)
  <snip>
  input flow-control is off, output flow-control is off
  Members in this channel: Gi1/7/1 Gi2/7/1
  ARP type: ARPA, ARP Timeout 04:00:00
  Last input never, output never, output hang never
  <<snip>>

6500-VSS# conf t
```

```
Enter configuration commands, one per line.  End with CNTL/Z.
6500-VSS(config)# int po 150
6500-VSS(config-if)# port-channel min-links 2
6500-VSS(config-if)#
%LINEPROTO-5-UPDOWN: Line protocol on Interface GigabitEthernet1/7/1, changed state to
down
%LINEPROTO-5-UPDOWN: Line protocol on Interface GigabitEthernet2/7/1, changed state to
down
%LINEPROTO-5-UPDOWN: Line protocol on Interface Port-channel150, changed state to down
%LINK-3-UPDOWN: Interface Port-channel150, changed state to down
%EC-SW1_SP-5-MINLINKS_NOTMET: Port-channel Po150 is down bundled ports (1) doesn't meet
min-links
%SPANTREE-SW1_SP-6-PORT_STATE: Port Po150 instance 50 moving from forwarding to disabled
%SPANTREE-SW1_SP-6-PORT_STATE: Port Po150 instance 150 moving from forwarding to disabled
%EC-SW1_SP-5-MINLINKS_NOTMET: Port-channel Po150 is down bundled ports (1) doesn't meet
min-links
%LINEPROTO-SW1_SP-5-UPDOWN: Line protocol on Interface GigabitEthernet1/7/1, changed state
to down
%LINEPROTO-SW1_SP-5-UPDOWN: Line protocol on Interface GigabitEthernet2/7/1, changed state
to down
%EC-SW2_SPSTBY-5-MINLINKS_NOTMET: Port-channel Po150 is down bundled ports (0) doesn't
meet min-links
%EC-SW2_SPSTBY-5-MINLINKS_NOTMET: Port-channel Po150 is down bundled ports (1) doesn't
meet min-links
```

For an MEC using LACP control protocol, min-links defines the minimum number of physical links in each chassis for a MEC to be operational. With a single member connected to each physical chassis, the configuration violates the minimum link requirements. The following example output illustrates that the port-channel interface is disabled and that the LACP state for each member link is in a wait state. The usage of the min-link feature disables the MEC for the associated connection.

```
6500-VSS# sh int po 150
Port-channel150 is down, line protocol is down (notconnect)
! <<snip>>
  ARP type: ARPA, ARP Timeout 04:00:00
  Last input never, output never, output hang never
  Last clearing of "show interface" counters never
  Input queue: 0/2000/0/0 (size/max/drops/flushes); Total output drops: 0
  Queueing strategy: fifo
  Output queue: 0/2000 (size/max)

6500-VSS# sh etherchannel 150 su
Flags:  D - down        P - bundled in port-channel
        w - waiting to be aggregated

Group  Port-channel  Protocol    Ports
------+-------------+-----------+-----------------------------------------------
150    Po150(SM)       LACP      Gi1/7/1(w)     Gi2/7/1(w)

Last applied Hash Distribution Algorithm: Adaptive

6500-VSS# sh spanning-tree int po 150
no spanning tree info available for Port-channel150
```

Tip **The use of the minimum links feature in the campus environment is not very effective. For all practical deployments of the VSS in the campus (two uplinks from each adjacent network devices), the minimum links feature should not be used.**

MEC Configuration

Configuration requirement for MEC is almost identical with standard EtherChannel interface. This section covers the basic configuration caveats, QoS support, and syslog guidelines for MEC configuration.

The procedure used to configure Layer-2 EtherChannel differs when compared with Layer-3 EtherChannel. Key considerations are as follows:

- Do not create Layer-2 MEC explicitly by defining it via the CLI. Instead, allow the Cisco IOS system to generate the Layer-2 MEC interface by associating to the port-channel group under each member interface.
- Create a Layer-3 MEC interface explicitly via the CLI and associate the port-channel group under each member interface.

Refer to the following URL for more information:

http://cco.cisco.com/en/US/partner/docs/switches/lan/catalyst6500/ios/12.2SX/configuration/guide/channel.html#wp1020478

QoS with MEC

QoS for MEC follows similar procedures as with any standard EtherChannel configuration. For generic QoS support or restrictions related to the VSS, please refer to following white paper QoS chapter:

http://www.cisco.com/en/US/prod/collateral/switches/ps5718/ps9336/white_paper_c11_429338.pdf

Monitoring

Because EtherChannel is ubiquitous in a VSS-enabled campus topology, monitoring EtherChannel is more critical in comparison to non-VSS EtherChannel environments. Generic tools available for monitoring a standard EtherChannel interface are fully applicable to an MEC interface. However, this section provides additional details about newer **show** command and specific **logging** commands that should be enabled to enhance operational understanding of EtherChannel connectivity.

Cisco IOS Release 12.2(33)SXH1 for the Cisco Catalyst 6500 platform now supports previously hidden commands for monitoring traffic flow on a particular link member of a MEC or EtherChannel. The following **remote** command will generate output depicting the interface and the port-channel interface being selected for a given source and destination

Hidden commands:

```
Catalyst6500# remote command switch test EtherChannel load-balance interface po 1 ip
1.1.1.1 2.2.2.2
Would select Gi4/1 of Po1
```

Cisco IOS command available for monitoring generic EtherChannel implementations:

```
6500-VSS# show EtherChannel load-balance hash-result interface port-channel 2 205 ip
10.120.7.65 vlan 5 10.121.100.49
Computed RBH: 0x4
Would select Gi1/9/19 of Po205
```

The following syslog configuration (**logging** command) is recommended in VSS with MEC interfaces. These commands can also be applied to the devices connected to the VSS as long as those devices support the respective syslog functionality.

Port-channel interfaces configurations:

```
interface Port-channel120
 logging event link-status
```

```
 logging event spanning-tree status

logging event link-status
%LINK-5-CHANGED: Interface Port-channel220, changed state to administratively down
%LINK-SW1_SP-5-CHANGED: Interface Port-channel220, changed state to administratively down

logging  event spanning-tree status
%SPANTREE-SW1_SP-6-PORT_STATE: Port Po220 instance 999 moving from learning to forwarding
```

Member link configuration:

```
interface GigabitEthernet1/8/1
 description Link member to port-channel
 logging event link-status
 logging event trunk-status
 logging event bundle-status

logging event link-status
Mar 25 11:43:54.574: %LINK-3-UPDOWN: Interface GigabitEthernet1/8/1, changed state to down
Mar 25 11:43:54.990: %LINK-3-UPDOWN: Interface GigabitEthernet2/8/1, changed state to down

logging event trunk-status
%DTP-SW2_SPSTBY-5-NONTRUNKPORTON: Port Gi2/8/1 has become non-trunk
%DTP-SW1_SP-5-NONTRUNKPORTON: Port Gi2/8/1 has become non-trunk

logging event bundle-status
%EC-SW2_SPSTBY-5-BUNDLE: Interface Gi1/8/1 joined port-channel Po220
%EC-SW2_SPSTBY-5-BUNDLE: Interface Gi2/8/1 joined port-channel Po220
```

MEC Load Sharing, Traffic Flow and Failure

Load sharing, traffic flow behavior, and failure conditions are covered in Chapter 3, "VSS-Enabled Campus Design," because the behavior and impact of MEC load sharing is more related to overall design of the campus network.

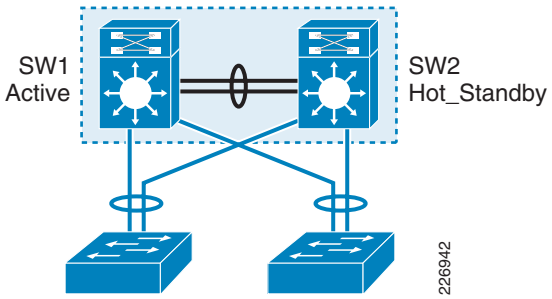
Capacity Planning with MEC

The maximum number of EtherChannels supported by a VSS depends on version of Cisco IOS. A maximum of 128 EtherChannels are supported in Cisco IOS 12.2(33) SXH. This limit is raised to 512 EtherChannels with Cisco IOS 12.2(33) SXI software releases. The scope and scalability as it applies to the VSS in the distribution block is discussed in the "Multilayer Design Best Practices with VSS" section on page 3-14.

MAC Addresses

In a standalone Cisco Catalyst 6500, the MAC addresses used for each interface and control plane is derived from the back plane EEPROM. VSS consists of two chassis (see Figure 2-25). Each physical member of VSS pair consists of pool of MAC addresses stored in backplane EEPROM.

Figure 2-25 VSS Roles



The VSS MAC address pool is determined during the role resolution negotiation. The active chassis pool of MAC addresses is used for the Layer-2 SVI and the Layer-3 routed interface—including the Layer-3 MEC interface. The Layer-2 MEC interface uses one of the link-member MAC addresses. The following CLI output examples reveal the active pool of MAC addresses.

```
6500-VSS# show switch virtual role

Switch  Switch Status  Preempt    Priority  Role
------------------------------------------------------------------
LOCAL    1     UP   FALSE(N )   110(110)  ACTIVE
REMOTE 2     UP   FALSE(N )   100(100)  STANDBY

6500-VSS# show catalyst6000 chassis-mac-addresses
  chassis MAC addresses: 1024 addresses from 0019.a927.3000 to 0019.a927.33ff
```

The MAC address allocation for the interfaces does not change during a switchover event when the hot-standby switch takes over as the active switch. This avoids gratuitous ARP updates (MAC address changed for the same IP address) from devices connected to VSS. However, if both chassis are rebooted together and the order of the active switch changes (the old hot-standby switch comes up first and becomes active), then the entire VSS domain will use that switch's MAC address pool. This means the interface will inherit a new MAC address, which will trigger gratuitous ARP updates to all Layer-2 and Layer-3 interfaces. Any networking device connected one hop away from the VSS (and any networking device that does not support gratuitous ARP), will experience traffic disruption until the MAC address of the default gateway/interface is refreshed or timed out. To avoid such a disruption, Cisco recommends using the configuration option provided with the VSS in which the MAC address for Layer-2 and Layer-3 interfaces is derived from the reserved pool. That takes advantage of the virtual switch domain identifier to form the MAC address. The MAC addresses of the VSS domain remain consistent with the usage of virtual MAC addresses, regardless of the boot order. For the exact formula, see the command and configuration chapter for VSS at www.cisco.com.

Tip **Cisco recommends the configuration of a virtual MAC address for VSS domain using the *switch virtual domain* command.**

```
6500-VSS(config)# switch virtual domain 10
6500-VSS(config-vs-domain)# mac-address use-virtual
```

The individual MAC addresses that reside in each chassis EEPROM are useful for assisting in the dual-active detection process. The following **show** commands illustrate how to find the base address for each chassis.

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

6500-VSS# sh idprom switch 2 backplane detail | inc mac
```

```
mac base = 0019.A924.E800
```

The above base addresses located on each chassis are used by the dual-active detection method described in "Campus Recovery with VSS Dual-Active Supervisors" section on page 4-18.

MAC Addresses and MEC

In VSS, the MAC address of the Layer-2 MEC interface is derived from one of the link-member burn-in (bia) interface MAC addresses. In a normal condition, the MAC address of the first interface added to the port-channel interface is chosen for Layer-2 port-channel (MEC) interface. If the interface whose MAC address is used for the port-channel is disabled, the port-channel interface will start using remaining member interface MAC addresses. However, if the interface that has just been disabled is reactivated, the Layer-2 MEC does not reuse the MAC address of that interface.

The process of inheriting the MAC address of the Layer-2 port-channel is shown below. The burn-in address (bia) of the the port-channel is derived from the first interface that was added to the port-channel.

```
6500-VSS#show etherchannel summary  | inc 220
220    Po220(SU)        LACP      Gi1/8/1(P)     Gi2/8/1(P)

6500-VSS#show interface gig 1/8/1 | inc bia
  Hardware is C6k 1000Mb 802.3, address is 0014.a922.598c (bia 0014.a922.598c)

6500-VSS#show interface gig 2/8/1 | inc bia
  Hardware is C6k 1000Mb 802.3, address is 0014.a92f.14d4 (bia 0014.a92f.14d4)
```

Note that the output where port-channel interface MAC address is derived from Gigabit interface 1/8/1.

```
6500-VSS#show interface port-channel 220 | inc bia
  Hardware is EtherChannel, address is 0014.a922.598c (bia 0014.a922.598c)


6500-VSS# conf t
6500-VSS(config)# interface gi1/8/1
6500-VSS(config-if)# shutdown
```

After disabling the link-member Gigabit-interface 1/8/1 whose MAC address was used by Layer-2 MEC, the port-channel starts using the remaining member (Gigabit-interface 2/8/1) burned-in address.

```
6500-VSS#show interface port-channel 220 | inc bia
  Hardware is EtherChannel, address is 0014.a92f.14d4 (bia 0014.a92f.14d4)

6500-VSS#show interface gig 2/8/1 | inc bia
  Hardware is C6k 1000Mb 802.3, address is 0014.a92f.14d4 (bia 0014.a92f.14d4)
```

If the interface is re-added to port-channel bundle, the MAC address of the port-channel does not change. The below CLI output illustrates that behavior.

```
6500-VSS(config)#interface gig 1/8/1
6500-VSS(config-if)#no shutdown


6500-VSS#show interface port-channel 220 | inc bia
  Hardware is EtherChannel, address is 0014.a92f.14d4 (bia 0014.a92f.14d4)

6500-VSS#show interface g 2/8/1 | inc bia
  Hardware is C6k 1000Mb 802.3, address is 0014.a92f.14d4 (bia 0014.a92f.14d4)
```

```
6500-VSS#show interface g 1/8/1 | inc bia
  Hardware is C6k 1000Mb 802.3, address is 0014.a922.598c (bia 0014.a922.598c)
```

In normal condition, the Layer-2 MAC address is used as sources of BPDU frame (addition of link activates the MEC interface and STP operation on that port). If the interface member (whose MAC address is used by Layer-2 MEC) is disabled, the change of source MAC in BPDU frame is detected by the switches connected to the VSS; however, the root MAC of the STP remains the same. This implies that STP topology did not change. For details, refer to "STP Operation with VSS" section on page 3-36.