Resilient Campus Network Design
Next Generation Campus Networks

Evolving Design Requirements

- Campus network design is evolving in response to multiple drivers
  - User Expectations: Always ON Access to communications
  - Business Requirements: Globalization means true 24x365
  - Technology Requirements: Unified Communications, high definition interactive video, new traffic patterns
  - Work Environment Requirements: Anytime anywhere access
  - Unexpected Requirements: Worms, viruses, …

- Campus design needs to evolve to support the evolving business and technology drivers
Resilient Campus Network Design

Agenda

Campus Design Foundation

Distribution Block Design Considerations
  - Fully Routed Campus
  - Cisco Catalyst 6500: Virtual Switching System

Access Layer Design Considerations
  - Architecture Enhancements
  - Availability
  - QoS
  - Security

Managing the Resilient Campus
High Availability Campus Design

Structure, Modularity and Hierarchy

- Optimize the interaction of the physical redundancy with the network protocols
  - Provide the necessary amount of redundancy
  - Pick the right protocol for the requirement
  - Optimize the tuning of the protocol

- The network looks like this so that we can map the protocols onto the physical topology

- We want to build networks that look like this
Hierarchical Campus Network
Structure, Modularity and Hierarchy

Not This!

Server Farm

WAN Internet PSTN
Redundancy and Protocol Interaction

Link Redundancy and Failure Detection

- Direct point-to-point fiber provides for fast failure detection
- IEEE 802.3z and 802.3ae link negotiation define the use of Remote Fault Indicator and Link Fault Signaling mechanisms
- Bit D13 in the Fast Link Pulse (FLP) can be set to indicate a physical fault to the remote side
- Do **not** disable auto-negotiation on GigE and 10GigE interfaces
- Carrier-Delay
  - 3560, 3750 and 4500—0 msec
  - 6500—Leave it at default 50 msec
- The default debounce timer on GigE and 10GigE fiber linecards is 10 msec
- The minimum debounce for copper is 300 msec
Redundancy and Protocol Interaction

Link Neighbor Failure Detection

- Indirect link failures are harder to detect
- With no direct HW notification of link loss or topology change convergence times are dependent on SW notification
- In certain topologies the need for TCN updates or dummy multicast flooding (uplink fast) is necessary for convergence
- In the recommended distribution block design recovery of access to distribution link failures is accomplished based on L2 CAM updates not spanning tree
- No dependence on external events (no need to wait for spanning tree convergence)
- Behavior is deterministic

All Links Forwarding: In an Environment with All Links Active Traffic Is Restored Based on HW Recovery
Redundancy and Protocol Interaction

Layers 2 and 3: Why Use Routed Interfaces?

Configuring L3 routed interfaces provides for faster convergence than an L2 switchport with an associated L3 SVI.

1. Link Down
2. Interface Down
3. Routing Update

8ms

21:38:37.042 UTC: %LINEPROTO-5-UPDOWN: Line protocol on Interface GigabitEthernet3/1, changed state to down
21:38:37.050 UTC: %LINK-3-UPDOWN: Interface GigabitEthernet3/1, changed state to down
21:38:37.050 UTC: IP-EIGRP(Default-IP-Routing-Table:100): Callback: route_adjust GigabitEthernet3/1

1. Link Down
2. Interface Down
3. Autostate
4. SVI Down
5. Routing Update

256ms

21:32:47.813 UTC: %LINEPROTO-5-UPDOWN: Line protocol on Interface GigabitEthernet2/1, changed state to down
21:32:47.821 UTC: %LINK-3-UPDOWN: Interface GigabitEthernet2/1, changed state to down
21:32:48.069 UTC: %LINK-3-UPDOWN: Interface Vlan301, changed state to down
21:32:48.069 UTC: IP-EIGRP(Default-IP-Routing-Table:100): Callback: route_adjust Vlan301
Redundancy and Protocol Interaction

Improving Layer 3 Neighbor Failure Detection

- EIGRP, OSPF, IS-IS, mBGP all have native hello/dead mechanisms
- Bidirectional Forwarding Detection (BFD)* provides a protocol-independent mechanism
  
  Negotiation of timers between peers
  
  BFD control packets are encapsulated in UDP unicast datagrams, destination port 3784
  
  Lightweight process, packets are not sequenced
  
  Must be 1 Hop Layer 3 adjacent

```plaintext
interface Vlan4
dampening
ip address 10.122.0.26 255.255.255.254
bfd interval 100 min_rx 100 multiplier 3
bfd neighbor 10.122.0.27
router eigrp 100
  bfd interface TenGigabitEthernet4/1
```

*Verify Cisco IOS® Release Availability, ESE does not yet have specific configuration guidance
Best Practice: Build Triangles, Not Squares

Deterministic vs. Non-Deterministic

**Triangles:** Link/Box Failure Does Not Require Routing Protocol Convergence

- Layer 3 redundant equal cost links support fast convergence
- Hardware based—fast recovery to remaining path
- Convergence is extremely fast (dual equal-cost paths: no need for OSPF or EIGRP to recalculate a new path)

**Squares:** Link/Box Failure Requires Routing Protocol Convergence

Model A

Model B
Equal Cost Multi-Path Convergence

Time to Recovery CEF Paths

1. Link failure detection
2. Removal of the entries in the routing table
3. Update of the software CEF table to reflect to loss of the next hop adjacencies
4. Update of the hardware tables
5. Routing protocol notification and reconvergence

Remote Link Failure Detection

Software Routing Table (RIB)

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Next Hop</th>
<th>Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.255.0.0/16</td>
<td>10.10.1.1</td>
<td>gig 1/1</td>
</tr>
<tr>
<td></td>
<td>10.20.1.1</td>
<td>gig 1/2</td>
</tr>
</tbody>
</table>

Cisco IOS Software CEF Tables

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Adjacency Ptr</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.255.0.0/16</td>
<td>Adj1 (gig 1/1)</td>
</tr>
<tr>
<td></td>
<td>Adj2 (gig 1/2)</td>
</tr>
</tbody>
</table>

Adjacency Table
Rewrite Information
AA.AA.AA.AA.AA, VLAN
BB.BB.BB.BB.BB, VLAN

Hardware Tables

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Adjacency Ptr</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.255.0.0/16</td>
<td>Adj1 (gig 1/1)</td>
</tr>
<tr>
<td></td>
<td>Adj2 (gig 1/2)</td>
</tr>
</tbody>
</table>

Adjacency Table
Rewrite Information
AA.AA.AA.AA.AA, VLAN
BB.BB.BB.BB.BB, VLAN

Routing Protocol Process
Resilient Campus Network Design

Agenda

- Campus Design Foundation
- Distribution Block Design Considerations
  - Fully Routed Campus
  - Cisco Catalyst 6500: Virtual Switching System
- Access Layer Design Considerations
  - Architecture Enhancements
  - Availability
  - QoS
  - Security
- Managing the Resilient Campus
Distribution Layer

Policy, Convergence, QoS, and High Availability

- Availability, load balancing, QoS and provisioning are the important considerations at this layer
- Aggregates wiring closets (access layer) and uplinks to core
- Protects core from high density peering and bounds Spanning Tree
- Route summarization, fast convergence, redundant path load sharing
- HSRP or GLBP to provide first hop redundancy
Multilayer Network Design
Layer 2 Access with Layer 3 Distribution

- Each access switch has unique VLANs
- No Layer 2 loops
- Layer 3 link between distribution
- No blocked links

- At least some VLANs span multiple access switches
- Layer 2 loops
- Layers 2 and 3 running over link between distribution
- Blocked links
Multilayer Network Design
Layer 2 Loops and Spanning Tree

- Implement physical L2 loops only when you have to
- Spanning tree protocol is very, very rarely the problem
- L2 has no native mechanism to dampen down a problem
- Utilize Rapid PVST+ for best convergence
- Limit the size of the L2 domain
- Complex L2 topologies take longer to converge
- Clear unnecessary VLANs from trunk configuration
Layer 2 Loops and Spanning Tree

Spanning Tree Should Behave the Way You Expect

- The root bridge should stay where you put it
  - Loopguard and rootguard
  - UDLD
- Only end station traffic should be seen on an edge port
  - BPDU guard
  - Port-Security
- There is a reasonable limit to B-Cast and M-Cast traffic volumes
  - Configure storm control on backup links to aggressively rate limit B-Cast and M-Cast
  - Utilize Sup720 rate limiters or SupIV/V with HW queuing structure
First Hop Redundancy
Capable of Sub-Second Convergence

interface Vlan4
ip address 10.120.4.1 255.255.255.0
ip helper-address 10.121.0.5
no ip redirects
vrrp 1 description Master VRRP
vrrp 1 ip 10.120.4.1
vrrp 1 timers advertise msec 250
vrrp 1 preempt delay minimum 180

Sub-second Hello timer enables < 1 Sec traffic recovery upstream
Preempt delay avoids black holing traffic when ACTIVE gateway recovers and preempt the backup, as upstream routing and link may not be active

interface Vlan4
ip address 10.120.4.2 255.255.255.0
standby 1 ip 10.120.4.1
standby 1 timers msec 250 msec 750
standby 1 priority 150
standby 1 preempt
standby 1 preempt delay minimum 180

interface Vlan4
ip address 10.120.4.2 255.255.255.0
glpb 1 ip 10.120.4.1
glpb 1 timers msec 250 msec 750
glpb 1 priority 150
glpb 1 preempt
glpb 1 preempt delay minimum 180
First Hop Redundancy with Load Balancing

Cisco Gateway Load Balancing Protocol (GLBP)

- Each member of a GLBP redundancy group owns a unique virtual MAC address for a common IP address/default gateway
- When end stations ARP for the common IP address/default gateway they are given a load balanced virtual MAC address
- Host A and host B send traffic to different GLBP peers but have the same default gateway

Note: GLBP supported on Cisco Catalyst® 6500(Sup720) and 4500(Sup6E)
Sub-Second Timer Considerations

HSRP, GLBP, OSPF, PIM

- Evaluate your network before implementing any sub-second timers
- Certain events can impact the ability of the switch to process sub-second timers
  - Application of Large ACLs
  - OIR of line cards in 6500
- The volume of control plane traffic can also impact the ability to process
  - 250/750 msec GLBP and HSRP timers are only valid in designs with less than 150 VLAN instances (Cisco Catalyst 6500 in the distribution)
  - Spanning Tree size discussed above
- Check size of input queue

interface GigabitEthernet3/2
description Downlink to Access
hold-queue 2000 in
hold-queue 2000 out
Multilayer Network Design
Well Understood Best Practices

- Match Cisco CatOS/Cisco IOS EtherChannel® settings and tune load balancing
- Summarize routes towards core
- Limit redundant IGP peering
- STP Root and HSRP primary or GLBP and STP port cost tuning to load balance on uplinks
- Set trunk mode on/negotiate
- Disable EtherChannel unless needed
- RootGuard on downlinks
- LoopGuard on uplinks
- Set port host on access layer ports:
  - Disable trunking
  - Disable EtherChannel
  - Enable PortFast
- RootGuard or BPDU-Guard
Resilient Campus Network Design

Agenda

Campus Design Foundation

Distribution Block Design Considerations
- Fully Routed Campus
- Cisco Catalyst 6500: Virtual Switching System

Access Layer Design Considerations
- Architecture Enhancements
- Availability
- QoS
- Security

Managing the Resilient Campus
Routing to the Edge

Layer 3 Distribution with Layer 3 Access

- Move the Layer 2/3 demarcation to the network edge
- Upstream convergence times triggered by hardware detection of light lost from upstream neighbor

10.1.20.0
VLAN 20 Data
10.1.120.0
VLAN 120 Voice

10.1.40.0
VLAN 40 Data
10.1.140.0
VLAN 140 Voice
Design Implications

Routed Access Simplifies Multicast

- Layer 2 access has two multicast routers on the access subnet, RPF checks and split roles between high and low IP address routers
- Routed Access has a single multicast router which simplifies multicast topology
Resilient Campus Network Design

Agenda

Campus Design Foundation

Distribution Block Design Considerations
  Fully Routed Campus
  Cisco Catalyst 6500: Virtual Switching System

Access Layer Design Considerations
  Architecture Enhancements
  Availability
  QoS
  Security

Managing the Resilient Campus
Virtual Switching System (VSS)

Cisco Catalyst 6500

- Virtual Switching System provides a capability to fundamentally re-architect the Layer 2 distribution block design by collapsing two physical distribution switches into a single ‘logical’ switch.

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

Single Control Plane

- While the Data Planes in both switches are active, only one switch has an active control plane—hence there is only one management point from which to manage the Virtual Switching System.

Note: The Standby Console Is Disabled and Cannot be Used for Normal Operational Purposes.
The Data Planes in both switches are active—hence each has a full copy of the forwarding tables and Security/QoS policies in hardware such that each can make a fully informed local forwarding decision.
Virtual Switch Architecture

Virtual Switch Link

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

A Virtual Switch Link Bundle Can Consist of up to 8 x 10GE Links

All Traffic Traversing the VS Link Is Encapsulated with a 32-Byte “Virtual Switch Header” containing Ingress and Egress Switch Port Indexes, Class of Service (CoS), VLAN Number, and Other Important Information from the Layer 2 and Layer 3 Header

The Control Plane Uses VSL for CPU-to-CPU Communications while the Data Plane Uses VSL to Extend the Internal Chassis Fabric to the Remote Chassis

VS Header  L2 Hdr  L3 Hdr  Data  CRC

Virtual Switch Active

Virtual Switch Standby
Virtual Switch Architecture

Virtual Switch Link

- VSL Encapsulation is only supported on PFC3C and DFC3C forwarding engines
- Only the 10GE ports on the modules shown below are able to support a Virtual Switch Link

VS-S720-10G-3C
VS-S720-10G-3CXL

WS-X6708-10G-3C
WS-X6708-10G-3CXL
Virtual Switch Architecture

Virtual Switch Domain

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

- The configurable values for the domain ID are 1–255. It is always recommended to use a unique VS Domain ID for each VS Domain throughout the network.
Virtual Switching System introduces new connectivity options such as Multi-Chassis EtherChannel.

Multi-Chassis EtherChannel

- Introduces a new deployment option for improving link resiliency
- Allows an EtherChannel link bundle to be terminated across two physical chassis
- Up to eight links can be supported in a Multi-Chassis EtherChannel
- MEC is supported with both standards-based 802.3ad link aggregation and Cisco PAGP
- Attached host sees the other end (Virtual Switch) as a single device
- EtherChannel hash has been modified so that each Virtual Switch will always choose the local link over a link in the other chassis in the same bundle
EtherChannel Concepts

EtherChannel Hash for MEC

- Deciding on which link of a Multi-Chassis EtherChannel to use in a Virtual Switch is skewed in favor toward local links in the bundle—this is done to avoid overloading the Virtual Switch Link (VSL) with unnecessary traffic loads.

Blue Traffic destined for the server will result in Link A1 in the MEC link bundle being chosen as the destination path.

Orange Traffic destined for the server will result in Link B2 in the MEC link bundle being chosen as the destination path.
1. Active Switch incurs a Supervisor outage; NSF/SSO kicks in and Standby Supervisor commences procedures to take over as Virtual Switch Active.

2. VS Standby initiates graceful restart; NSF-aware neighbors exchange updates with new Virtual Switch Active; Non-Stop Forwarding of local packets continues as switch assumes active role.
High Availability
MEC Link Failure Recovery: Layer 2 Uplink
High Availability
MEC Link Failure Recovery: Layer 3 Uplink
High Availability
Actual Virtual Switch Failure Recovery
Virtual Switch System

Benefits

- **Upstream and Downstream**
  Upstream and Downstream neighbors will view the Virtual Switch as a single Layer 2 switching node or as a single Layer 3 routing node thus reducing Layer 2/3 control protocol traffic

- **Single Management Point**
  Administrators will see a single management point from which to configure and administer the VSS which includes a single consolidated configuration file for both physical switches

- **Multi-Chassis EtherChannel**
  Multi-Chassis EtherChannel allows a link bundle to terminate across two physical Cisco Catalyst 6500 chassis

- **As far as the other end is concerned, the link bundle is seen as terminating on the one physical device even though it is actually terminating across two chassis**
Next Generation Campus Design

Evolving the Campus Foundation Architecture

- Traditional Layer 2 designs remain valid
- Evolving architectures provide
  - Simplified Control Plane: Remove dependence on STP
  - Increased Capacity: Provide flow-based load balancing
  - High Availability: 200 msec or better recovery
- Flexibility to provide for the right implementation for each network requirement
Resilient Campus Network Design

Agenda

- Campus Design Foundation
- Distribution Block Design Considerations
  - Fully Routed Campus
  - Cisco Catalyst 6500: Virtual Switching System
- Access Layer Design Considerations
  - Architecture Enhancements
  - Availability
  - QoS
  - Security
- Managing the Resilient Campus
Access Layer

Feature Rich Environment

- It’s not just about connectivity
- Layer 2/Layer 3 feature rich environment; convergence, HA, security, QoS, IP multicast, etc.
- Intelligent network services: QoS, trust boundary, broadcast suppression, IGMP snooping
- Intelligent network services: PVST+, Rapid PVST+, EIGRP, OSPF, DTP, PAgP/LACP, UDLD, FlexLink, etc.
- Cisco Catalyst integrated security features IBNS (802.1x), (CISF): port security, DHCP snooping, DAI, IPSG, etc.
- Automatic phone discovery, conditional trust boundary, power over Ethernet, auxiliary VLAN, etc.
- Spanning tree toolkit: Portfast, UplinkFast, BackboneFast, LoopGuard, BPDUGuard, BPDUFilter, RootGuard, etc.
Access Layer Network Services

UC Integrated with Network QoS and Security

- Phone contains a three-port switch that is dynamically configured by the access switch and Cisco Call Manager
  - Power negotiation
  - VLAN configuration
  - 802.1x interoperation
  - QoS configuration
  - DHCP
  - Cisco CallManager® registration

Endpoints Dynamically Participate in the Overall Network QoS and Security
Access Layer Network Services

Evolving Application Traffic Requirements

- Application Traffic Requirements are evolving
- Desktop based Unified Communications
- Collaborative applications
- High Definition Video
- Web portals and front-ends leveraging common HTTP transport for a wide variety of business requirements
- Requirements for more granular application awareness in the network services, QoS, Security, HA, ...

Application Intelligence, Security and Flow Information Required at All Layers
Resilient Campus Network Design

Agenda

Campus Design Foundation
Distribution Block Design Considerations
  Fully Routed Campus
  Cisco Catalyst 6500: Virtual Switching System
Access Layer Design Considerations
  Architecture Enhancements
  Availability
  QoS
  Security
Managing the Resilient Campus
Supervisor 32 PISA

Hardware-Based Feature Acceleration

- PISA is a superset of the MSFC2a daughter card
- It contains both a faster Route Processor (RP) and a dedicated Network Processor
  - Improved Route Processor (750 MHz)
  - 1GB RP DRAM by default
  - PISA CF bootdisk (256 MB)
- Network Processor provides HW accelerated L4-7 IP Services
- WS-S32P-10GE and WS-S32P-GE
Supervisor 32 PISA
Application Forwarding and Visibility

- Traffic is redirected to PISA when NBAR or FPM is configured on an interface—PISA acts as bump in the wire service.

![Diagram showing Traffic Flow through PISA and Traffic Flow bypassing PISA]
Cisco Catalyst 4500 E-Series: Centerflex

Next Generation Centralized Architecture

- Centralized 320 Gbps, 250 Mpps, L2/3/4
- IPv4 (250Mpps) and IPv6 (125Mpps) in HW
- Supports full redundancy with SSO/NSF/ISSU (7R and 10R)
- Four active 10GbE ports or Eight active SFPs in redundant mode
- 512MB DRAM (upgradeable to 1 Gig)
- 256K FIB entries, 128K security/QoS
- 3, 6, 7R, and 10R support
- Supports “classic” (6 Gig/slot) and “E” series line cards (24 gig/slot) with no performance degradation if mixed
## Cisco Catalyst 4500 E-Series: Centerflex
### Next Generation Centralized Architecture

<table>
<thead>
<tr>
<th>Feature</th>
<th>Classic Performance</th>
<th>E-Series Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Throughput</td>
<td>102M(IPv4)/30K(IPv6)</td>
<td>250(IPv4)/125(IPv6) Mpps</td>
</tr>
<tr>
<td>Transmit Queues</td>
<td>4 Queues Fixed</td>
<td>8 Queues per Port (MQC Model)</td>
</tr>
<tr>
<td>Configurable Queues</td>
<td>1P3Q1T</td>
<td>1P7Q2T up to 8184 Queues per Port/100K Pool</td>
</tr>
<tr>
<td>Flexible Policers</td>
<td>8K Ingress/8K Egress</td>
<td>16K in 2K Increments</td>
</tr>
<tr>
<td>Security and QoS ACE</td>
<td>64K</td>
<td>128K Flexible</td>
</tr>
<tr>
<td>MLDv2 Snooping</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>EtherChannel</td>
<td>64 with Global Balancing</td>
<td>512 with Per-Channel Load Balancing Algorithm</td>
</tr>
<tr>
<td>SPAN/RSPAN</td>
<td>6 Sessions</td>
<td>8 Sessions</td>
</tr>
<tr>
<td>URPF</td>
<td>No</td>
<td>Strict Mode</td>
</tr>
<tr>
<td>Output Policy on Input Marketing</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Resilient Campus Network Design

Agenda

- Campus Design Foundation
- Distribution Block Design Considerations
  - Fully Routed Campus
  - Cisco Catalyst 6500: Virtual Switching System
- Access Layer Design Considerations
  - Architecture Enhancements
  - Availability
  - QoS
  - Security
- Managing the Resilient Campus
Access Layer Design: Topology

Daisy Chaining Access Layer Switches

- Daisy chained access switches rely on STP between members of the ‘pile’
- Unicast flooding is common in daisy chained access switch networks
- If you use modular (chassis based) switches these problems are not a concern
  - Loopback links not required
  - No longer forced to have L2 link in distribution
Access Layer Design: Capacity
10/100/1000 Access and Distribution/Core Aggregation

Typical 20:1 Data Over-Subscription

Typical 4:1 Data Over-Subscription

Core

Distribution

Access

10GE and 10GE Channels
Access Layer Design: Availability
NSF/SSO, Cisco IOS Modularity and ISSU

- Access layer availability affects UC/VoIP
  - How to ensure 5x9’s voice availability
  - Non-disruptive to phone (recovery does not trigger a phone reboot)

- Cisco Catalyst 4500 and 6500 Supervisor hardware redundancy (1+1) leverage four key mechanisms to improve network resiliency and provide for enhanced operational change processes
  - SSO—Stateful Switchover
  - NSF—NonStop Forwarding
  - Cisco IOS Modularity
  - ISSU—In Service Software Upgrade
Access Switch Supervisor Redundancy
Routing Protocol Graceful Restart Capabilities (NSF)

- Non-Stop Forwarding provides graceful restart enhancements to EIGRP, OSPF, IS-IS and BGP
- An NSF-capable router continuously forwards packets during an SSO processor recovery
- NSF-aware and NSF-capable routers provide for transparent routing protocol recovery

Graceful restart extensions enable neighbor recovery without resetting adjacencies
Routing database re-synchronization occurs in the background
Access Design Considerations for NSF/SSO

Supervisor Uplinks

- **Cisco Catalyst 4500 E-Series**
  - All uplinks can active on both Supervisors
  - Uplink ports do **not** go down during SSO recovery

- **Cisco Catalyst 6500**
  - All uplinks are active on both Supervisors
  - Uplink ports go down when the supervisor is reset
  - Recommended to run 12.2(18)SXF5 or later when using Supervisor uplinks to carry L3 traffic (routed ports or SVI)
  - Recommended to run 12.2(18)SXF12 when using Supervisor uplinks configured as an EtherChannel to carry L3 traffic
Cisco Catalyst 4500 Series Classic Supervisor

Uplink Redundancy as of 12.2(25)SG

**Supervisor II+, Supervisor IV**
- 2 x GigE Ports Are Active

**Supervisor V**
- 4 x GigE Ports Are Active Concurrently

**Supervisor II+10GE**
- 2 x 10GE and 4 x GigE Ports Are Active

**Supervisor V-10GE**
- 2 x 10GE and 4 x GigE Ports Are Active
Cisco Catalyst Supervisor 6-E
Uplink Redundancy for 4507R-E and 4510R-E

10 GbE Default Uplink Configuration
2 x GigE (Full Line Rate)

10 GbE Optional Uplink Configuration
4 x GigE (2:1 Oversubscribed)

GE SFP (with Twin Gig) Default Configuration
4 x 1GE (Full Line Rate)

GE SFP (with Twin Gig) Optional Configuration
8 x 1GE (Full Line Rate)
Access Layer Design - Availability
In Service Software Upgrade (ISSU)

- ISSU provides a mechanism to perform software upgrades and downgrades without taking the switch out of service.
- Leverages the capabilities of NSF and SSO to allow the switch to forward traffic during supervisor IOS upgrade (or downgrade).
- Network does not re-route and no active links are taken out of service.
Full image ISSU upgrade on the Cisco Catalyst 4500 is a four-step process.

Possible to rollback (abort) up until you complete the fourth step (commit to final state).

Leverages NSF/SSO to implement supervisor transition.

Less than 200 msec of traffic loss during a software upgrade.
Access Layer Design: Availability

Access Switch Supervisor Redundancy

- Access switch is the single point of failure in best practices HA campus design.
- Supervisor failure is most common cause of access switch service outages after L1 failures (power and fiber).
- Recommended design NSF/SSO provides for sub second recovery of voice and data traffic.

![Diagram showing access switch supervisor redundancy](image-url)
Resilient Campus Network Design

Agenda

Campus Design Foundation

Distribution Block Design Considerations
  - Fully Routed Campus
  - Cisco Catalyst 6500: Virtual Switching System

Access Layer Design Considerations
  - Architecture Enhancements
  - Availability
  - QoS
  - Security

Managing the Resilient Campus
The Real-Time Interactive service class should... be configured to provide a high assurance for bandwidth for CS4 marked packets to ensure that they get forwarded...

Note that this service class may be configured as a second EF PHB that uses relaxed performance parameter, a rate scheduler, and CS4 DSCP value.” RFC 4594 Section 4.4
How Many Classes Does the Campus Need?

Expanding the Number of Classes of Service over Time

4/5 Class Model
- Real Time
- Call Signaling
- Critical Data
- Best Effort
- Scavenger

8-Class Model
- Voice
- Video
- Call Signaling
- Critical Data
- Network Control
- Critical Data
- Bulk Data
- Best Effort
- Scavenger

11-Class Model
- Voice
- Interactive-Video
- Streaming Video
- Call Signaling
- IP Routing
- Network Management
- Mission-Critical Data
- Transactional Data
- Bulk Data
- Best Effort
- Scavenger

Time

Presentation_ID © 2007 Cisco Systems, Inc. All rights reserved. Cisco Confidential
Expanding the Number of Classes of Service
Cisco Catalyst 4500E Enhanced Queuing

**Classic Supervisors**
- Four Static Tx Queues per Port

**Flexible Resources**
- Eight User Configurable Tx Queues per Port

**Supervisor II+, II+10GE, IV, V, V-10GE**
- Four queues (1P3Q1T)
- Fixed queue size
- Q3 is priority queue

**Supervisor 6-E**
- Configurable up to eight queues (1P7Q2T) with configurable thresholds
- Queue size configurable up to 8184 packets per queue (100K Pool)
- Class based queuing
- User configurable priority queue
With auto-qos configured default switch behavior is to not trust edge ports and remark all traffic to configured CoS/DSCP.

When switch and phone exchange CDP the trust boundary is extended to IP phone.

Phone rewrites CoS from PC port to ‘0’, switch rewrites DSCP.

Sup32 PISA provides an intelligent QoS remarking override for specifically defined applications on the Data VLAN.
Enhanced Access Trust Boundary

PISA NBAR and MQC QoS

- HW NBAR works together with QoS to assign QoS actions based on application classification

- Modular QoS traffic classification:
  - Define match criteria (class-map)
  - Associate actions for a given match criteria in a policy-map
  - Assign policy to an interface

- New match criteria: “match protocol <protocol_name>”

```
Policy Map
Policy Map can contain up to 32 class maps
(config)#policy-map NBAR_policy
(config-pmap)#class-map myApp

Class Map
Refers to a set of classification criteria for the following action criteria—these can be DSCP, ACL, or protocol
(config)#class-map myApp
(config-cmap)#match access-group 101
(config-cmap)#match protocol http
(config-cmap)#match protocol rtp

Policing/Trust Actions
Action settings for trust and policing
(config)#policy-map NBAR_policy
(config-pmap)#class-map myApp
(config-pmap)#set dscp 40
```
NBAR Real-Time Transport Protocol

Classification of UC Applications

- Eases classification of voice and video traffic
- Distinguishes between RTP packets based on payload type and CODECS
- Removes dependencies on UDP Port Range and DSCP markings
- Example

  NBAR to match RTP traffic with the payload-types 0, 1, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 64

<table>
<thead>
<tr>
<th>CODEC</th>
<th>Payload Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>G.711 (Audio)</td>
<td>0 (mu-law)</td>
</tr>
<tr>
<td></td>
<td>8 (a-law)</td>
</tr>
<tr>
<td>G.721 (Audio)</td>
<td>2</td>
</tr>
<tr>
<td>G.722 (Audio)</td>
<td>9</td>
</tr>
<tr>
<td>G.723 (Audio)</td>
<td>4</td>
</tr>
<tr>
<td>G.728 (Audio)</td>
<td>15</td>
</tr>
<tr>
<td>G.729 (Audio)</td>
<td>18</td>
</tr>
<tr>
<td>H.261 (Video)</td>
<td>31</td>
</tr>
<tr>
<td>MPEG-1 (A/V)</td>
<td>14 (Audio)</td>
</tr>
<tr>
<td>MPEG-2 (A/V)</td>
<td>32 (Video)</td>
</tr>
<tr>
<td></td>
<td>33 (A-V)</td>
</tr>
<tr>
<td>Dynamic</td>
<td>96–127</td>
</tr>
</tbody>
</table>

match protocol rtp payload-type "0, 1, 4 - 0x10, 10001b - 10010b, 64"
HTTP Port Overloading
NBAR HTTP Identification and Classification

- Migration to common HTTP interface for multiple applications
- Challenge to distinguish priority based on port numbers
- NBAR deep packet inspection allows marking based on HTTP content

PISA(config-cmap)#match protocol http ?
  content-encoding  Encoding mechanism used to package entity body
  from              E-mail of human controlling the user-agent
  host              Host name of Origin Server containing resource
  location          Exact location of resource from request
  mime              Content-Type of entity body
  referer           Address the resource request was obtained from
  server            Software used by Origin Server handling request
  url               Uniform Resource Locator path
  user-agent        Software used by agent sending the request
  <cr>

match protocol http user-agent *Mozilla/4.0*
match protocol http referer *http://www.cisco.com/go/nbar*
# NBAR Protocol Discovery

## Real-Time Application Visibility

- **NBAR Protocol Discovery**: discover what apps are running on your network and provide real-time statistics.
- Per-interface, per-protocol, bi-directional statistics (bit rate (bps); packet count; byte count)
- SNMP accessible for centralized monitoring
- Supported by Partner products (Concord, CA, InfoVista, Micromuse, IBM) and MRTG

```
Sup32-PISA#show ip nbar protocol-discovery top-n 5
```

### Vlan611

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Packet Count</td>
<td>Packet Count</td>
</tr>
<tr>
<td></td>
<td>Byte Count</td>
<td>Byte Count</td>
</tr>
<tr>
<td></td>
<td>5min Bit Rate (bps)</td>
<td>5min Bit Rate (bps)</td>
</tr>
<tr>
<td></td>
<td>5min Max Bit Rate (bps)</td>
<td>5min Max Bit Rate (bps)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Remote_Desktop</td>
<td>319</td>
<td>252</td>
</tr>
<tr>
<td></td>
<td>157009</td>
<td>47083</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>7000</td>
<td>3000</td>
</tr>
<tr>
<td></td>
<td>1591</td>
<td>1591</td>
</tr>
<tr>
<td></td>
<td>162282</td>
<td>162282</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>1000</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>1000</td>
</tr>
</tbody>
</table>

**NBAR Protocol Discovery**/

Real-Time Application Visibility
Resilient Campus Network Design

Agenda

Campus Design Foundation

Distribution Block Design Considerations
- Fully Routed Campus
- Cisco Catalyst 6500: Virtual Switching System

Access Layer Design Considerations
- Architecture Enhancements
- Availability
- QoS
- Security

Managing the Resilient Campus
Controlling Access to the Network
Who and What Is Connecting to the Network?

- FBI/CSI Risk Assessment*
  
  Many enterprises network ports are open
  Usually any laptop can plug into the network and gain access to the network
  Of companies surveyed total loss was over $130 million
  Average spending per employee $247 per year
  78% said about inside attacks—they did not know if they were under attack

- Secure Access beyond a locked door
  Query for Identity

- Inspect Devices for Posture
  Check OS Patch Levels
  Validate Desktop Security Tools

Hardening the Edge
Cisco Catalyst Integrated Security Features

- **Port Security** prevents CAM attacks, DHCP Starvation attacks and spanning tree loop mitigation
- **DHCP Snooping** prevents Rogue DHCP Server attacks
- **Dynamic ARP Inspection** prevents current ARP attacks
- **IP Source Guard** prevents IP/MAC Spoofing and a wide variety of TCP/UDP splicing and DoS attacks
Edge Security: Intelligent Filtering
Flexible Packet Matching: Sup32 PISA

- ACL can match traffic based on L2-L4 information: legitimate traffic could be blocked.
- Classification is dependant on hardware PFC3 TCAM support:
  - No support for the following match criteria: packet length, TTL.
- Predefined supported classification criteria, only match one stack layer:

- FPM is a stateless solution.
- Matches any characteristics in a packet header and payload:
- Matches L2–L7 information.
- Specify arbitrary bits/bytes at any offset.
- Supports multiple stack layers.
- Can match not only on outer IP header, but also inner header—for instance, the inner header of a GRE encapsulated packet.
- String match and regular expressions.
- Set up custom filters rapidly using XML-based policy language.
Edge Security: Intelligent Filtering

Defining the FPM Rule

- Traffic matching a given protocol stack is subject to an FPM Deep Packet inspection rule
- Supports pattern matching through regular expressions and string matching: up to 48 regular expressions (32 Bytes window in 12.2(18)ZY)

**Slammer**

```
class-map type access-control match-all slammer
description "match on slammer packets"
match field UDP dest-port eq 1434
match field IP length eq 404
match start UDP payload-start offset 196 size 4 eq 0x4011010
```

**Gnutella**

```
class-map type access-control match-all gnutella
match start TCP payload-start offset 0 size 32 regex "^GNUTELLA CONNECT"
```

**Nimda**

```
class-map type access-control match-any cm-nimdal
match start 13-start offset 40 size 32 regex "[\/]csrss\.exe"
match start 13-start offset 40 size 32 regex "[\/]httpodbc\.dll"
match start 13-start offset 40 size 32 regex "[\/]sample\.exe"
match start 13-start offset 40 size 32 regex "[\/]dnsservice\.exe"
match start 13-start offset 40 size 32 regex "[\/]puta\.eml"
match start 13-start offset 40 size 32 regex "[\/]puta\.scr"
match start 13-start offset 40 size 32 regex "[\/]readme\.eml"
```
Resilient Campus Network Design

Agenda

- Campus Design Foundation
- Distribution Block Design Considerations
  - Fully Routed Campus
  - Cisco Catalyst 6500: Virtual Switching System
- Access Layer Design Considerations
  - Architecture Enhancements
  - Availability
  - QoS
  - Security
- Managing the Resilient Campus
Proactive Fault Detection and Reaction

Distributed Management Intelligence

- **Challenge:** In today’s highly available networks improved physical redundancy is not enough, intelligent system failure detection and recovery are key.

- **Detect and Isolate**
- **Enhanced System Stability**
- **Enhanced Network Stability**

- **Generic Online Diagnostics (GOLD) Provides Proactive, Scheduled and Manual System Diagnostics**
- **Enhanced Object Tracking (EOT), Embedded Event Manager (EEM), and Smart Call Home (SCH) Provide Intelligent Response to System Events**

**Enhanced System Stability**
- Memory Corruption
- Software Inconsistency
- System Faults
- Link Faults
Integrated System Operations Toolkit
Detect, React, and Analyze

Proactive Diagnostics
- **Generic Online Diagnostics (GOLD)** can proactively detect software and hardware malfunctions before they adversely impact network traffic

Customizable System Reaction
- **Embedded Event Manager (EEM)** allows customizable reaction to GOLD detected events or Cisco IOS counters and events
- **Smart Call Home** provides automated operations and TAC notification for critical network systems events

Improved Analysis Tools
- **System Event Archive (SEA)** maintains a secure log of system events and critical messages across reboots
- **On Board Failure Logging (OBFL)** is like a “black box” recorder on each line card that records diagnostic and environmental information
Generic Online Diagnostics

An Example: Loopback Test—Line card Data Path Coverage

- Test is disruptive for the tested port (subseconds)
- Verifies the tested port functionality and the datapath
- Newer linecards support non-disruptive loopback tests: ten consecutive failures are treated as fatal and will result in port being error-disabled

"show diagnostic content module X" to see available tests and their impact
Generic Online Diagnostics

**Diagnostic Operation**

**Boot-Up Diagnostics**

```
Switch(config)#diagnostic bootup level complete
```

**Runtime Diagnostics**

**Health-Monitoring**

```
Switch(config)#diagnostic monitor module 5 test 2
Switch(config)#diagnostic monitor interval module 5 test 2
00:00:15
```

**On-Demand**

```
Switch#diagnostic start module 4 test 8
Module 4: Running test(s) 8 may disrupt normal system operation
Do you want to continue? [no]: y
Switch#diagnostic stop module 4
```

**Scheduled**

```
Switch(config)#diagnostic schedule module 4 test 1 port 3 on Jan 3 2005 23:32
Switch(config)#diagnostic schedule module 4 test 2 daily 14:45
```

Run During System Bootup, Line Card OIR or Supervisor Switchover
Makes Sure Faulty Hardware Is Taken out of Service

Non-Disruptive Tests Run in the Background
Serves as HA Trigger

All Diagnostics Tests Can Be Run on Demand, for Troubleshooting Purposes; It Can Also Be Used as a Pre-deployment Tool

Schedule Diagnostics Tests, for Verification and Troubleshooting Purposes
Embedded Event Manager

Proactive Fault Detection and Notification

- EEM is a Cisco IOS technology that runs on the control plane; it is a combination of processes designed to monitor key system parameters such as CPU utilization, interface errors, counters, SNMP and SYSLOG events, and act on specific events or thresholds/counters that are exceeded.

Actions
- Execute a Cisco IOS CLI command
- Send a CNS event
- Increment/decrement an EEM counter
- Force an SSO switchover
- Request system information
- Send an email
- Run another EEM policy
- Reload the switch
- Generate an SNMP trap
- Generate a SYSLOG message
Embedded Event Manager

Architecture

Events
- SNMP Agent
- IOS IDB’s
- IOS CLI
- RIB
- OIR
- SYSLOG
- HA
- Routing
- Counters
- Memory

Event Detectors
- SYSLOG Event Detector
- SNMP Event Detector
- Timer Event Detector
- Counter Event Detector
- I/F Counter Event Detector
- CLE Event Detector
- OIR Event Detector
- None Event Detector
- RF Event Detector
- IOS Watchdog Event Detector
- GOLD Event Detector
- Application Event Detector
- SYS Manager Event Detector
- SYS Monitor Event Detector

Embedded Event Manager Server

EEM Event Detector API
EEM Event Client API

EEM Policy Director
Subscribes to Events and Implements Actions

EEM Policy
Embedded Event Manager

EEM Application Example

Upon matching the provided SYSLOG message ‘LINK-3 UPDOWN’, the switch performs the following actions:

- Display error statistics for the link that has gone down
- Start a Time Domain Reflectometry (TDR) test
- Start a GOLD Loopback test
- Send the results using a provided template to a user-configurable address
Embedded Event Manager

Configuration Example

EEM Applet Example

```cisco
event manager applet TEST
  event syslog pattern "%LINK-3-UPDOWN: Interface GigabitEthernet7/1" maxrun 20
  action 1.0 cli command "en"
  action 2.0 cli command "test cable-diagnostics tdr interface G7/1"
  action 3.0 cli command "diagnostic start module 7 test 2 port 1"
  action 4.0 mail server "x.x.x.x" to "email_id@x.com" from "Switch-1" subject "Urgent! Interface went down" body "G7/1 went down"
```

EEM TCL Script Example

```cisco
event manager environment _email_server <IP_address>
event manager environment _email_to email_id@x.com
event manager environment _syslog_pattern .*UPDOWN.*state to down.*
event manager environment _email_from Switch1@mylab.com
event manager environment intchk_template disk1:/interfacecheck.template
event manager directory user policy disk1:/
event manager policy interfacecheck.tcl
::cisco::eem::event_register_syslog occurs 1 pattern $_syslog_pattern maxrun 90
# EEM policy to monitor for a specified syslog message.
# check if all the env variables we need exist
<snip>
namespace import ::cisco::eem::*
namespace import ::cisco::lib::*
# The Body of the code goes here
<snip>
```
Embedded Event Manager
Cisco Beyond: Product Extension Community

- New on Cisco.com
- `/go/ciscobeyond`
- Open source scripts, share, upload, download, learn by example
- Categories include: Ntwk mgmt, routing, QoS, High availability, User interface, etc.
- Comments, ratings, community-managed forum

http://cisco.com/go/ciscobeyond
Smart Call Home
Proactive Problem Identification

TestErrorCounterMonitor ---------> F
- Error code ------------------> 1 (DIAG_FAILURE)
- Total run count --------------> 2484
- Last test execution time ----> Feb 01 2007 10:55:52
- First test failure time -------> Jan 31 2007 11:55:17
- Last test failure time -------> Feb 01 2007 10:55:52
- Last test pass time ---------> Jan 31 2007 11:54:45
- Total failure count ---------> 2474
- Consecutive failure count ---> 2474

Error Records as following:
- ID -- Asic Identification
- IN -- Asic Instance
- PO -- Asic Port Number
- RE -- Register Identification
- RM -- Register Identification More
- EG -- Error Group
- DV -- Delta Value
- CF -- Consecutive Failure
- TF -- Total Failure

<table>
<thead>
<tr>
<th>ID</th>
<th>IN</th>
<th>PO</th>
<th>RE</th>
<th>RM</th>
<th>DV</th>
<th>EG</th>
<th>CF</th>
<th>TF</th>
</tr>
</thead>
<tbody>
<tr>
<td>49</td>
<td>0</td>
<td>255</td>
<td>240</td>
<td>255</td>
<td>8</td>
<td>2</td>
<td>2483</td>
<td>2483</td>
</tr>
</tbody>
</table>

Customer

Indicated developing unrecoverable failure. This is usually a problem related to improper grounding or excessive radiation emitted into the device. Make sure that the device is properly grounded and that neighboring devices are not emitting excessive radiation levels.
Proactive Fault Detection and Reaction

Distributed and Embedded Operations

- Improved fault detection and faster reaction to events
  - Generic Online Diagnostics (GOLD)
  - Embedded Event Manager (EEM) and Enhanced Object Tracking (EOT)
  - Smart Call Home (SCH)

- More visibility into actual state of applications and the network
  - NetFlow
  - IP SLA
  - NBAR
  - ERSPAN