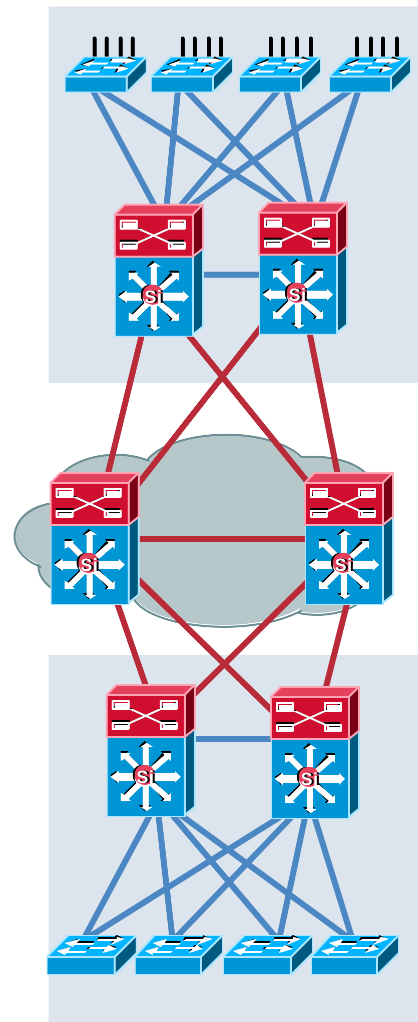
CAMPUS DESIGN: ANALYZING THE IMPACT OF EMERGING TECHNOLOGIES ON CAMPUS DESIGN

SESSION RST-3479

Campus Design
A Multitude of Design Options and Challenges

Cisco.com

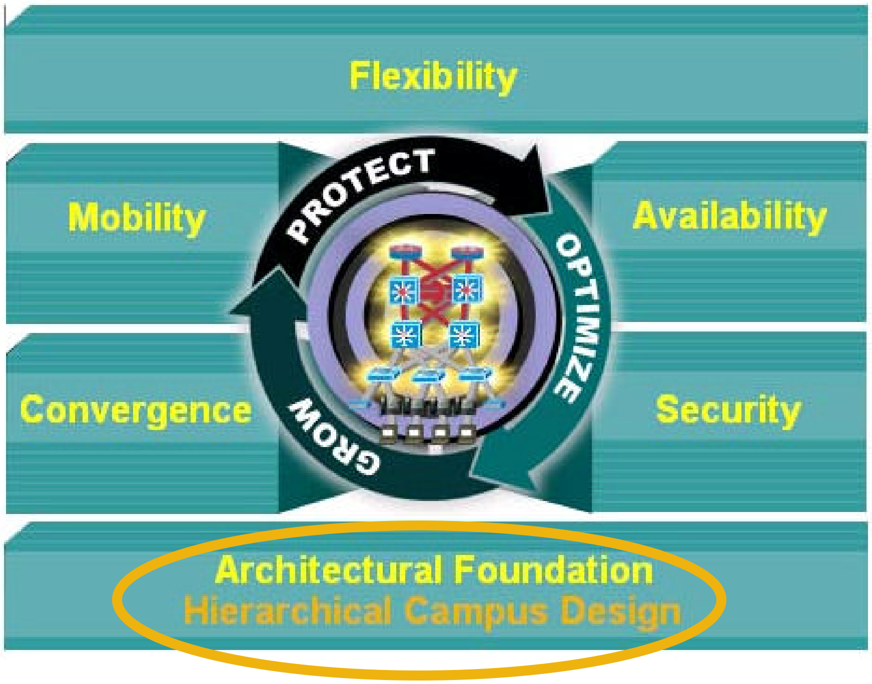
- **Campus network design is evolving in response to multiple drivers**
- **Voice, financial systems driving requirement for 5 nines availability and minimal convergence times**
- **Adoption of Advanced Technologies (voice, segmentation, security, wireless) all introduce specific requirements and changes**
- **The Campus is an integrated system everything impacts everything else**



High Availability Combined with Flexibility and Reduced OPEX

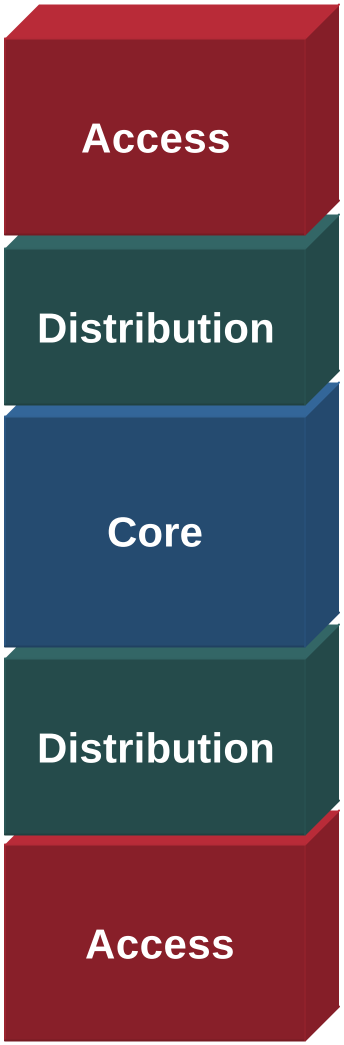
Agenda

- Foundational Design Review
- Convergence—IP Communications
- Wireless LAN and Wireless Mobility
- High Availability
    - Alternatives to STP
    - Device HA (NSF/SSO and Stackwise™)
    - Resilient Network Design
- Segmentation and Virtualization
    - Access Control (IBNS and NAC)
    - Segmentation
- Questions and Answers



Multilayer Campus Design
Hierarchical Building Blocks

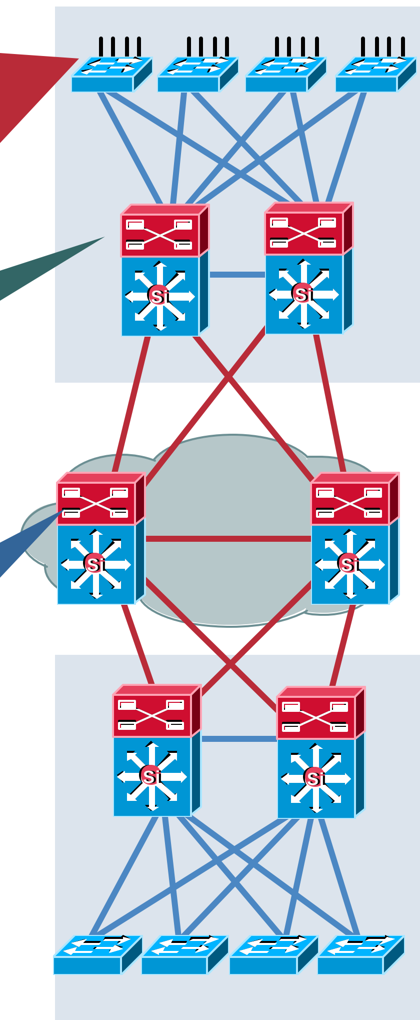
Cisco.com



- Network trust boundary
- Use Rapid PVST+ if you MUST have L2 loops in your topology
- Use UDLD to protect against 1 way up/up connections
- Avoid daisy chaining access switches
- Avoid asymmetric routing and unicast flooding, don't span VLANS across the access layer

- Aggregation and policy enforcement
- Use HSRP or GLBP for default gateway protection
- Use Rapid PVST+ if you MUST have L2 loops in your topology
- Keep your redundancy simple; deterministic behavior = Understanding failure scenarios and why each link is needed

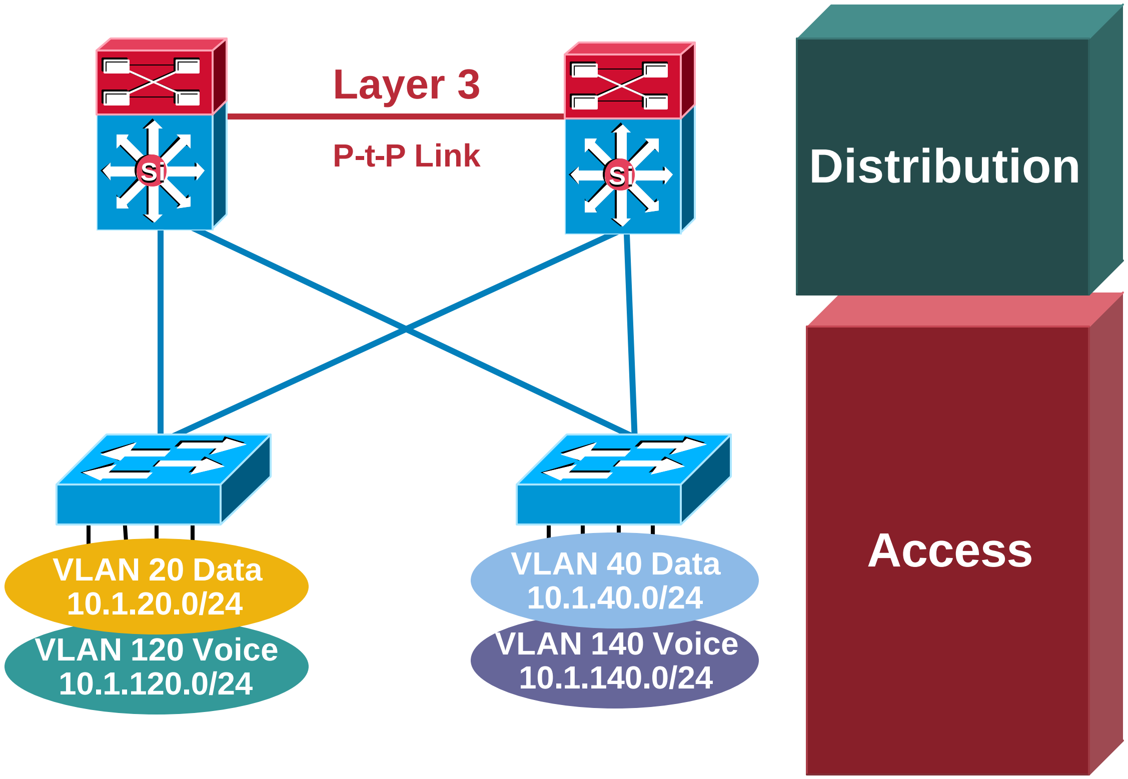
- Highly available and fast—always on
- Deploy QoS end-to-end: Protect the good and Punish the bad
- Equal cost core links provide for best convergence
- Optimize CEF for best utilization of redundant L3 paths



Distribution Building Block

Reference Design—No VLANs Span Access Layer

- Unique Voice and Data VLAN in every access switch
- STP root and HSRP primary tuning or GLBP to load balance on uplinks
- Set Port Host on access layer ports:
  - Disable Trunking
  - Disable Etherchannel
  - Enable PortFast
- Configure Spanning Tree Toolkit
  - Loopguard
  - Rootguard
  - BPDU-Guard
- Use Cisco® Integrated Security Features (CISF) Features



Campus Solution Test Bed

Verified Design Recommendations

Cisco.com

Total of 68 Access Switches, 2950, 2970, 3550, 3560, 3750, 4507 SupII+, 4507SupIV, 6500 Sup2, 6500 Sup32, 6500 Sup720 and 40 APs (1200)

Three Distribution Blocks
6500 with Redundant Sup720

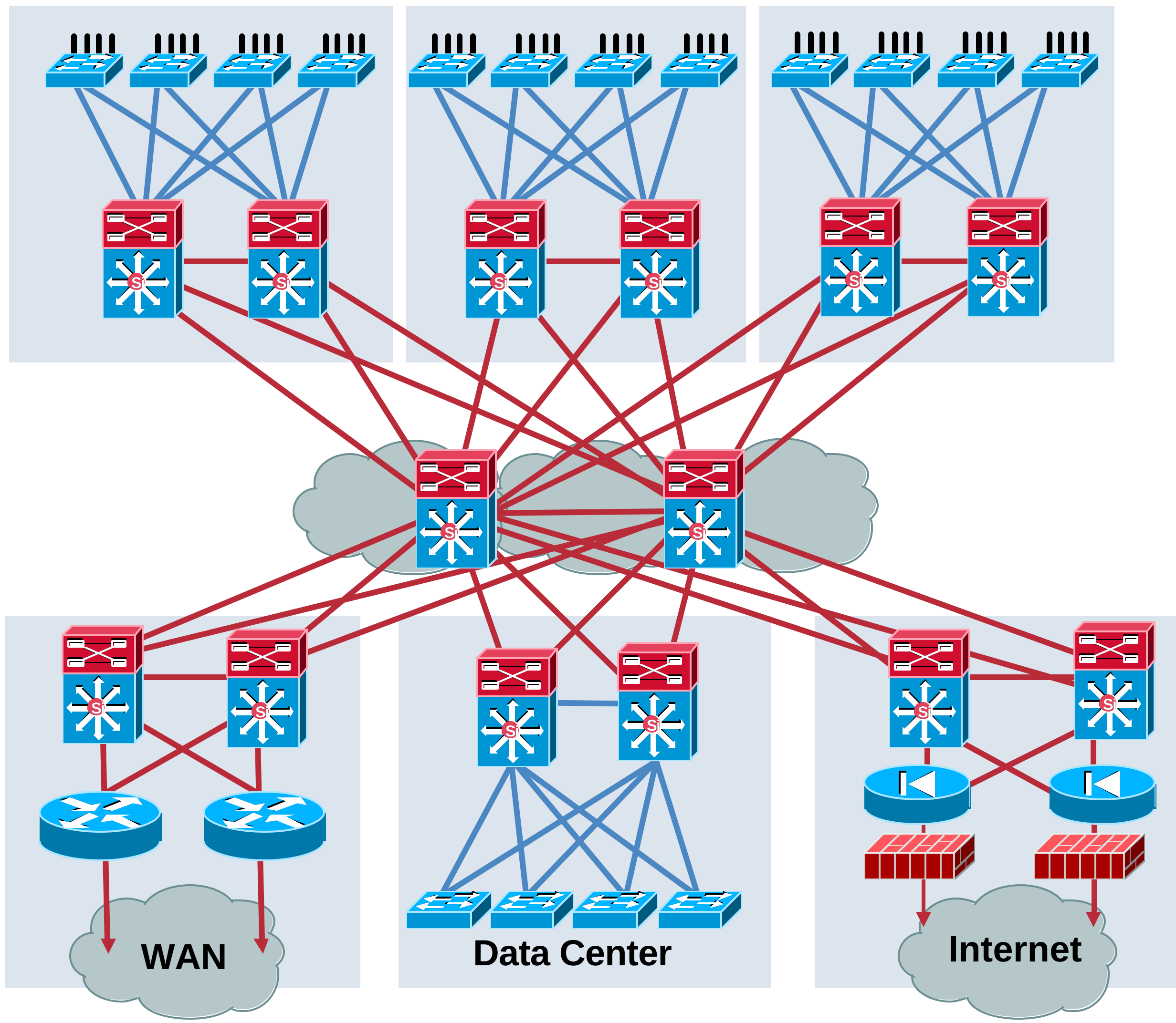
4507 with Redundant SupV

6500 with Redundant Sup720s

Three Distribution Blocks
6500 with Redundant Sup720s

7206VXR NPEG1

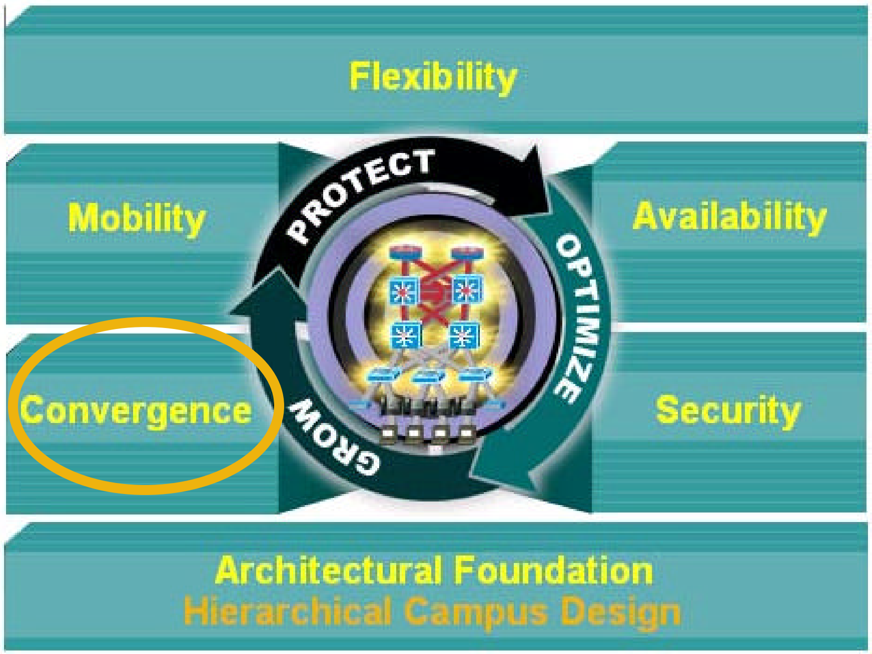
4500 SupII+, 6500 Sup720, FWSM, WLSM, IDSM2, MWAM



Agenda

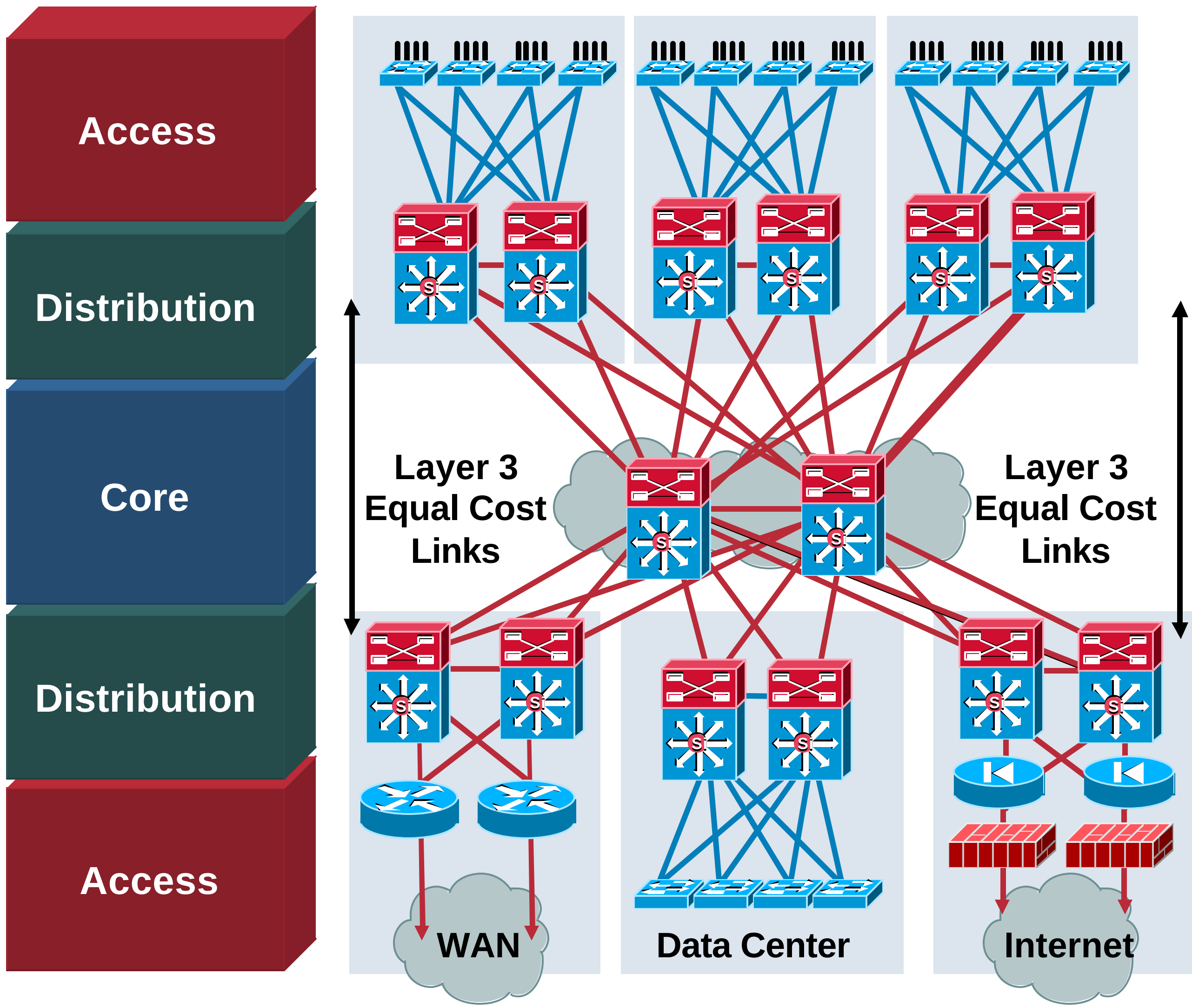
Cisco.com

- **Foundational Design Review**
- **Convergence—IP Communications**
- **Wireless LAN and Wireless Mobility**
- **High Availability**
  - **Alternatives to STP**
  - **Device HA (NSF/SSO and Stackwise)**
  - **Resilient Network Design**
- **Segmentation and Virtualization**
  - **Access Control (IBNS and NAC)**
  - **Segmentation**
- **Questions and Answers**

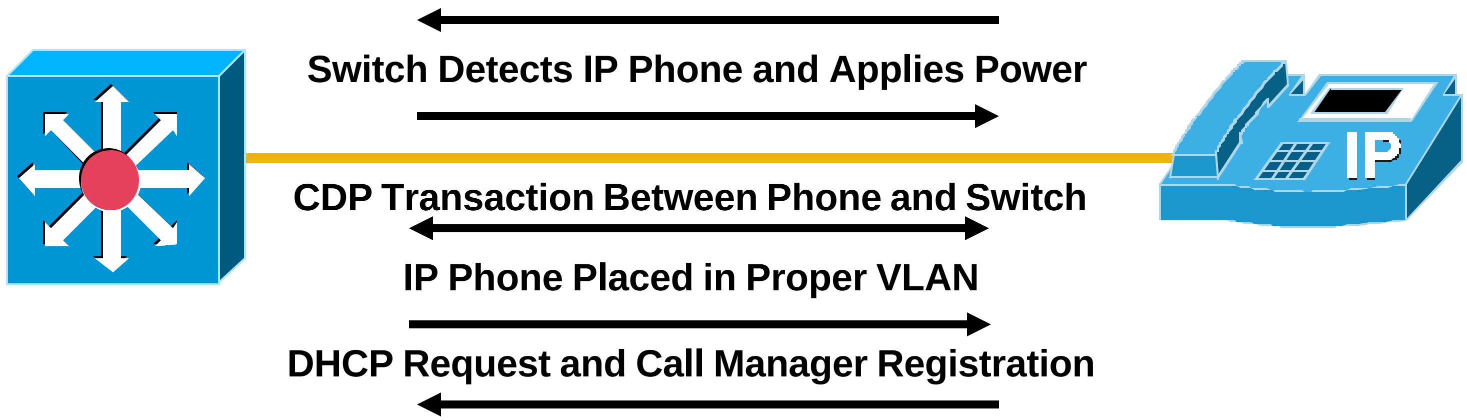


Building a Converged Campus Network
Infrastructure Integration, QoS and Availability

- **Access layer**
  - Auto phone detection
  - Inline power
  - QoS: scheduling, trust boundary and classification
  - Fast convergence
- **Distribution layer**
  - High availability, redundancy, fast convergence
  - Policy enforcement
  - QoS: scheduling, trust boundary and classification
- **Core**
  - High availability, redundancy, fast convergence
  - QoS: scheduling, trust boundary



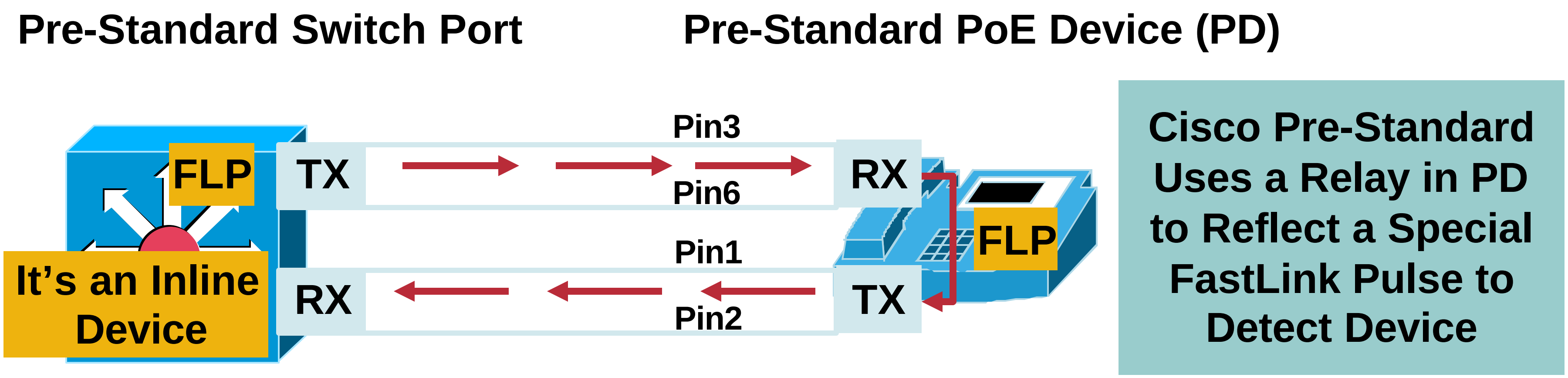
Infrastructure Integration
Extending the Network Edge

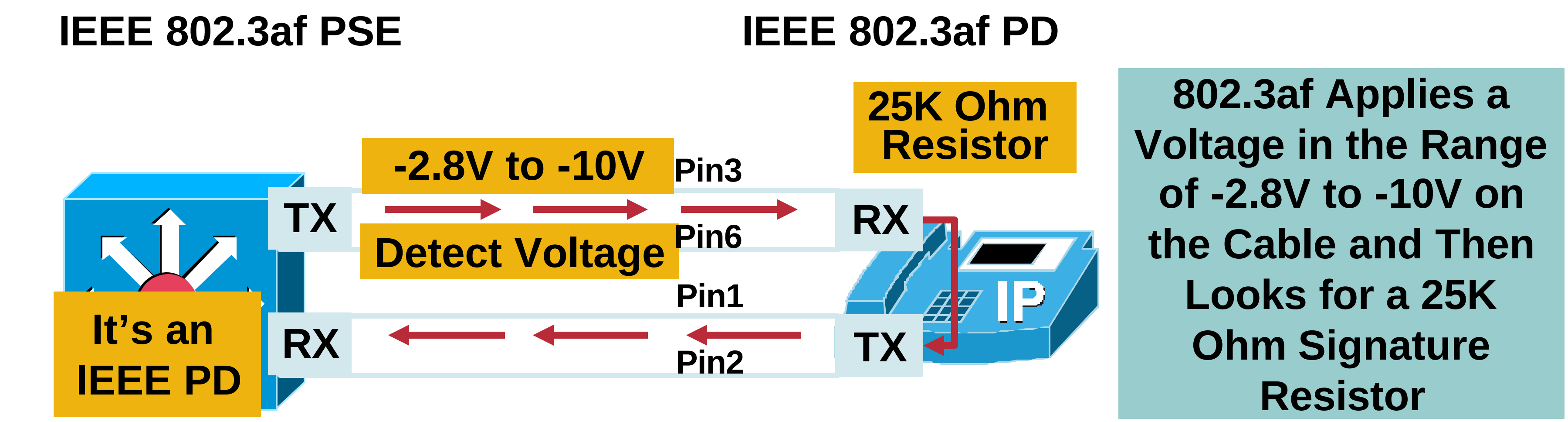


- **Phone contains a 3 port switch that is configured in conjunction with the access switch and CallManager**
  1. **Power negotiation**
  2. **VLAN configuration**
  3. **802.1x interoperation**
  4. **QoS configuration**
  5. **DHCP and CallManager registration**

Infrastructure Integration: First Step

Device Detection



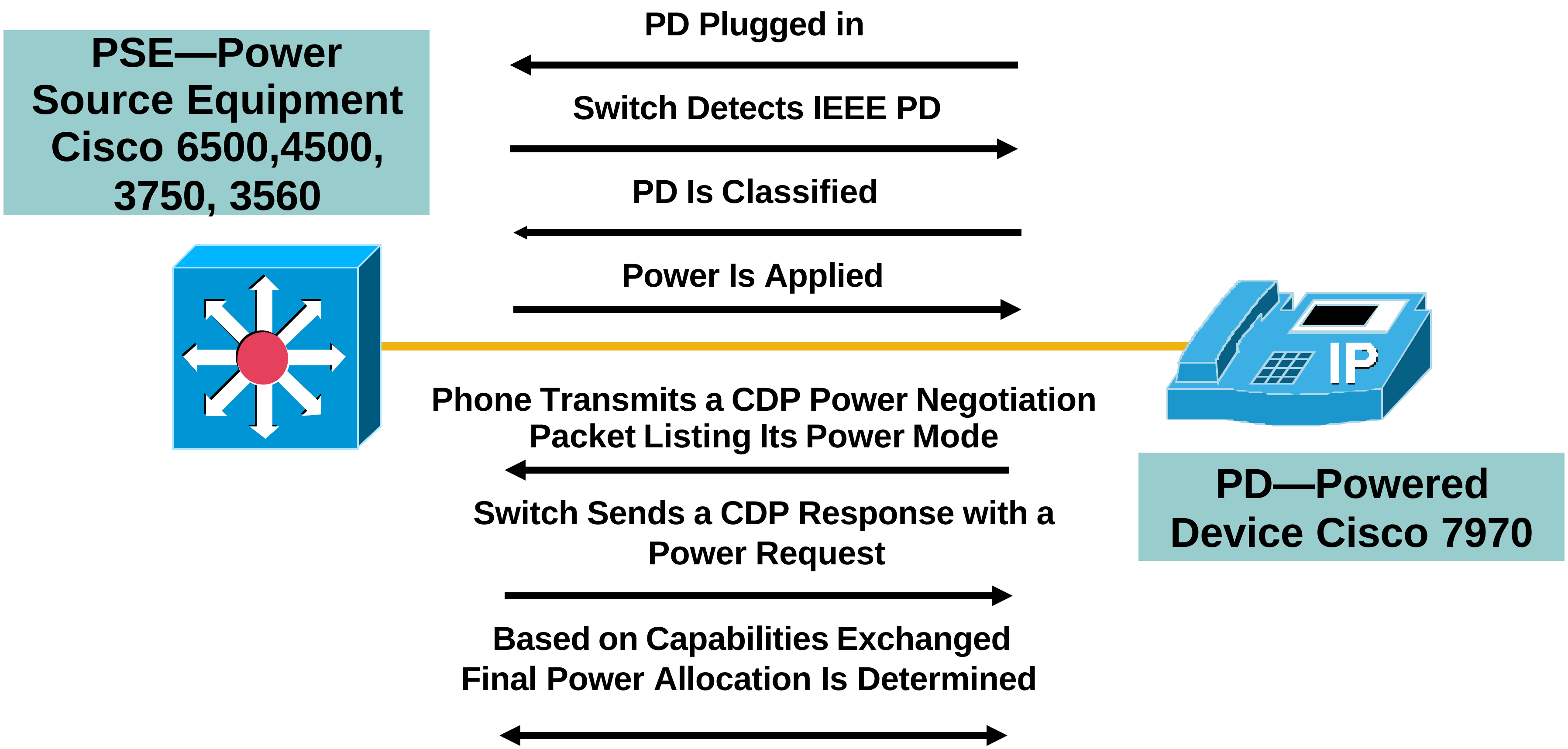


Infrastructure Integration: First Step
Power Requirement Negotiation

- **Cisco pre-standard devices initially receive 6.3 watts and then optionally negotiate via CDP**
- **802.3af devices initially receive 12.95 watts unless PSE able to detect specific PD power classification**

Class	Usage	Minimum Power Levels Output at the PSE	Maximum Power Levels at the Powered Device
0	Default	15.4W	0.44 to 12.95W
1	Optional	4.0W	0.44 to 3.84W
2	Optional	7.0W	3.84 to 6.49W
3	Optional	15.4W	6.49 to 12.95W
4	Reserved for Future Use	Treat as Class 0	Reserved for Future Use: a Class 4 Signature Cannot Be Provided by a Compliant Powered Device

Enhanced Power Negotiation
802.3af Plus Bi-Directional CDP (Cisco 7970)



- **Using bidirectional CDP exchange exact power requirements are negotiated after initial power-on**

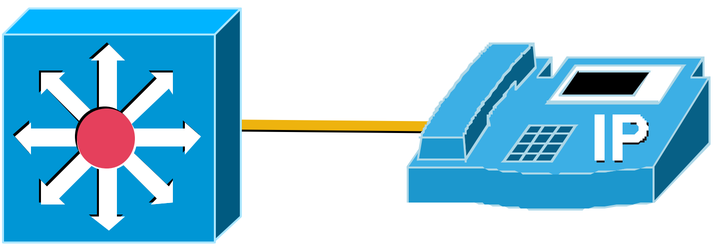
Design Considerations for PoE
Power Management

Cisco.com

- Switch manages power by what is allocated not by what is currently used
- Device power consumption is not constant
- A 7960G requires 7W when the phone is ringing at maximum volume and requires 5W on or off hook
- Understand the power behaviour of your PoE devices
- Utilize static power configuration with *caution*

```
Dynamic allocation:
power inline auto max 7200
```

```
Static allocation:
power inline static max 7200
```

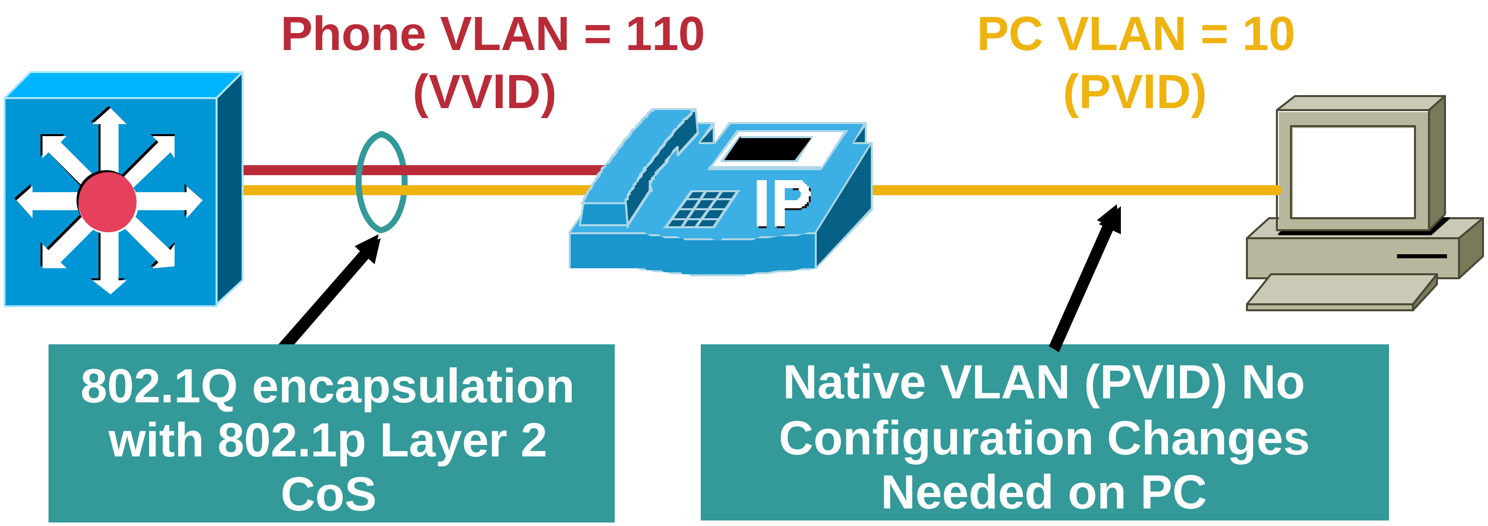


- Use power calculator to determine power requirements

http://www.cisco.com/go/powercalculator

Infrastructure Integration: Next Steps
VLAN, QoS and 802.1x Configuration

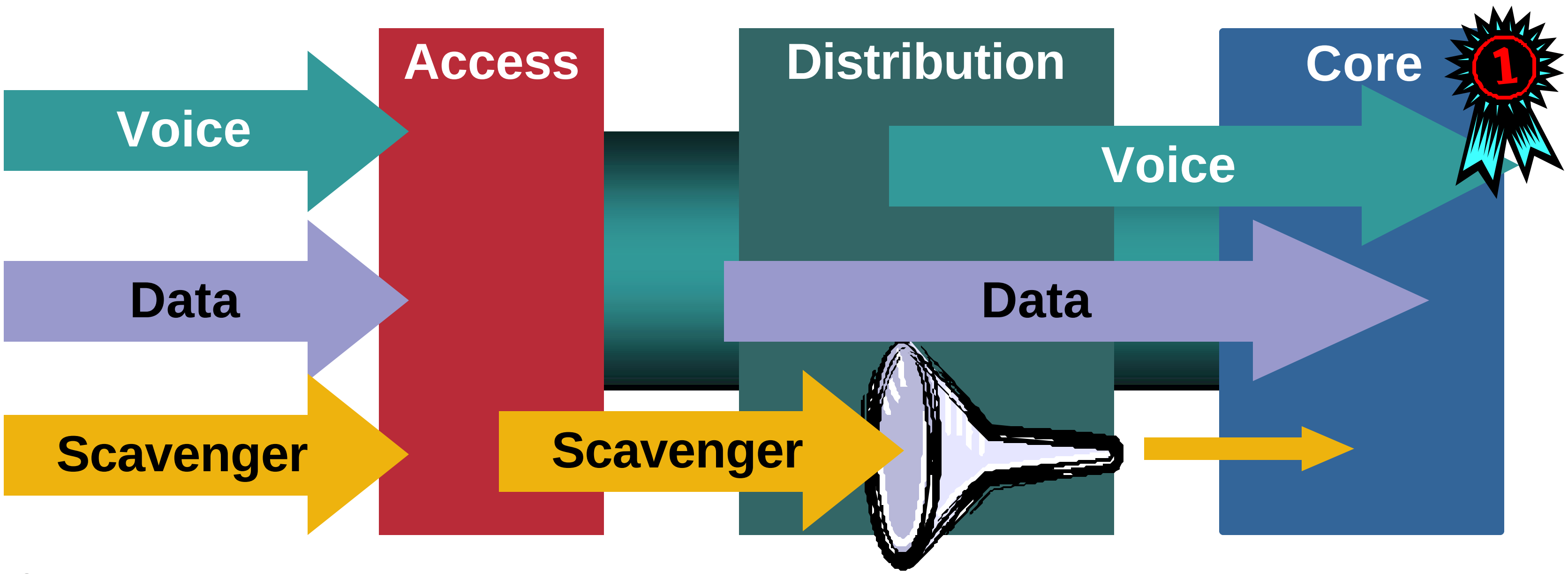
Cisco.com



- **During initial CDP exchange phone is configured with a Voice VLAN ID (VVID)**
- **Phone also supplied with QoS configuration via CDP TLV fields**
- **Additionally switch port currently bypasses 802.1x authentication for VVID if detects Cisco phone**

Why QoS in the Campus
Protect the Good and Punish the Bad

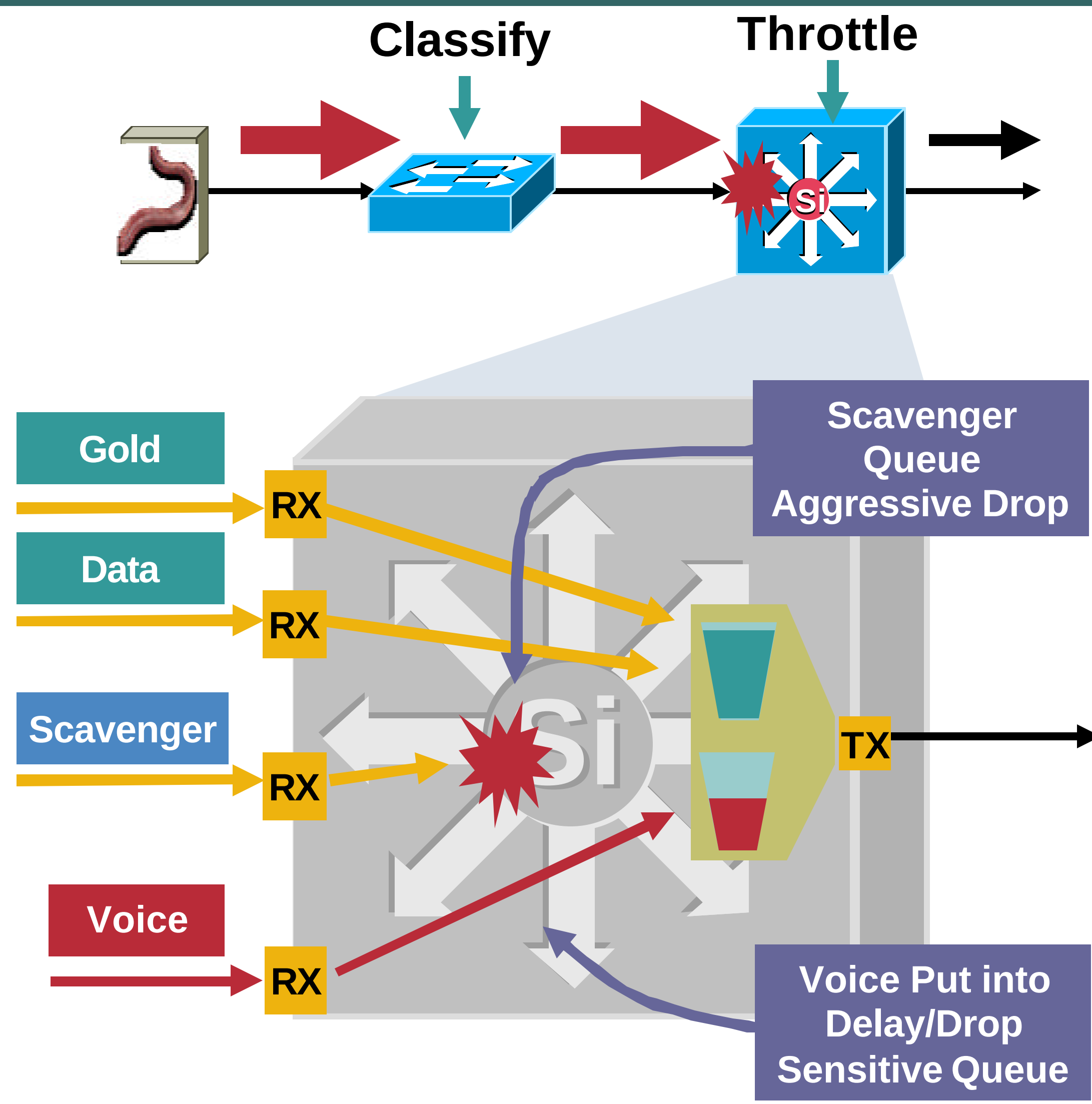
- QoS does more than just protect Voice and Video
- For "best-effort" traffic an implied "good faith" commitment that there are at least some network resources available is assumed
- Need to identify and potentially punish out of profile traffic (potential worms, DDOS, etc.)
- Scavenger class is an Internet-2 Draft Specification => CS1/CoS1



Campus QoS Design Considerations

Classification and Scheduling in the Campus

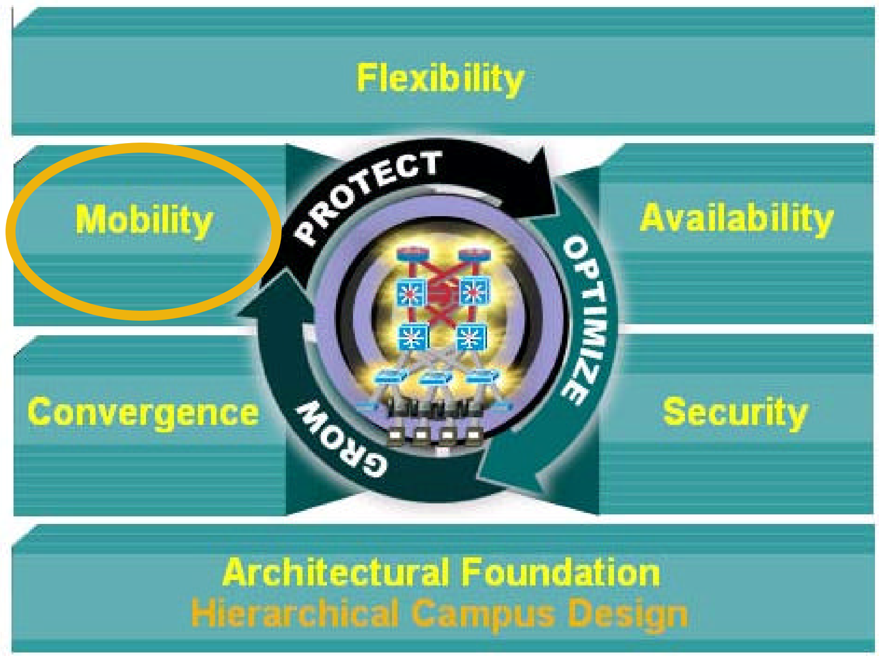
- Edge traffic classification scheme is mapped to upstream queue configuration
- Voice needs to be assigned to the HW priority queue
- Scavenger traffic needs to be assigned its own queue/threshold
- Scavenger configured with low threshold to trigger aggressive drops
- Multiple queues are the only way to "guarantee" voice quality, protect mission critical and throttle abnormal sources



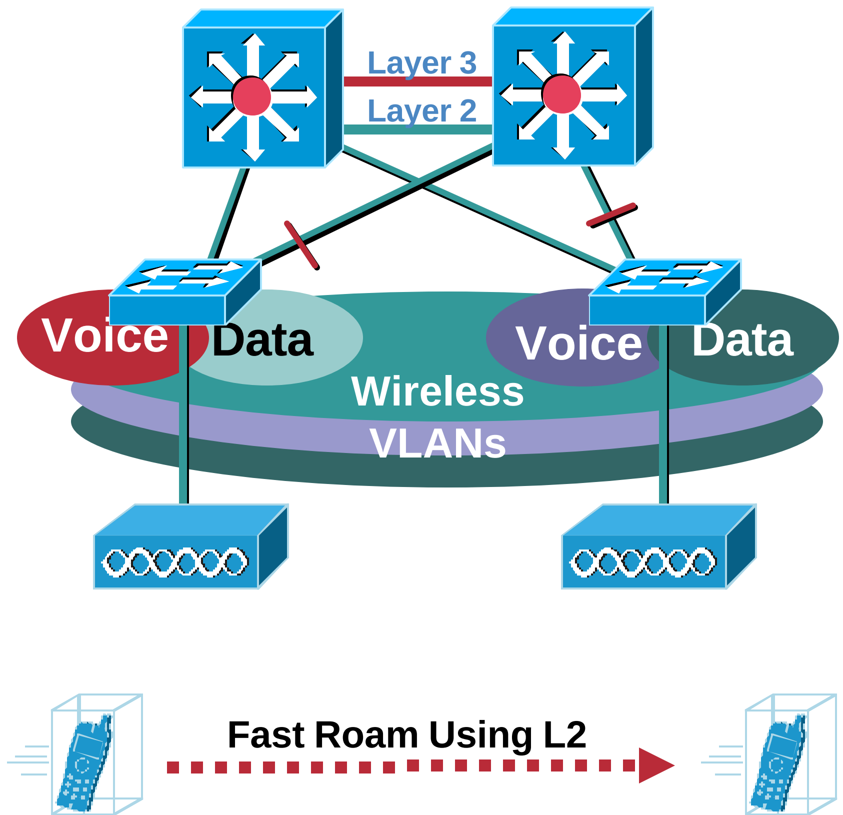
Agenda

Cisco.com

- Foundational Design Review
- Convergence—IP Communications
- Wireless LAN and Wireless Mobility
- High Availability
    - Alternatives to STP
    - Device HA (NSF/SSO and Stackwise)
    - Resilient Network Design
- Segmentation and Virtualization
    - Access Control (IBNS and NAC)
    - Segmentation
- Questions and Answers



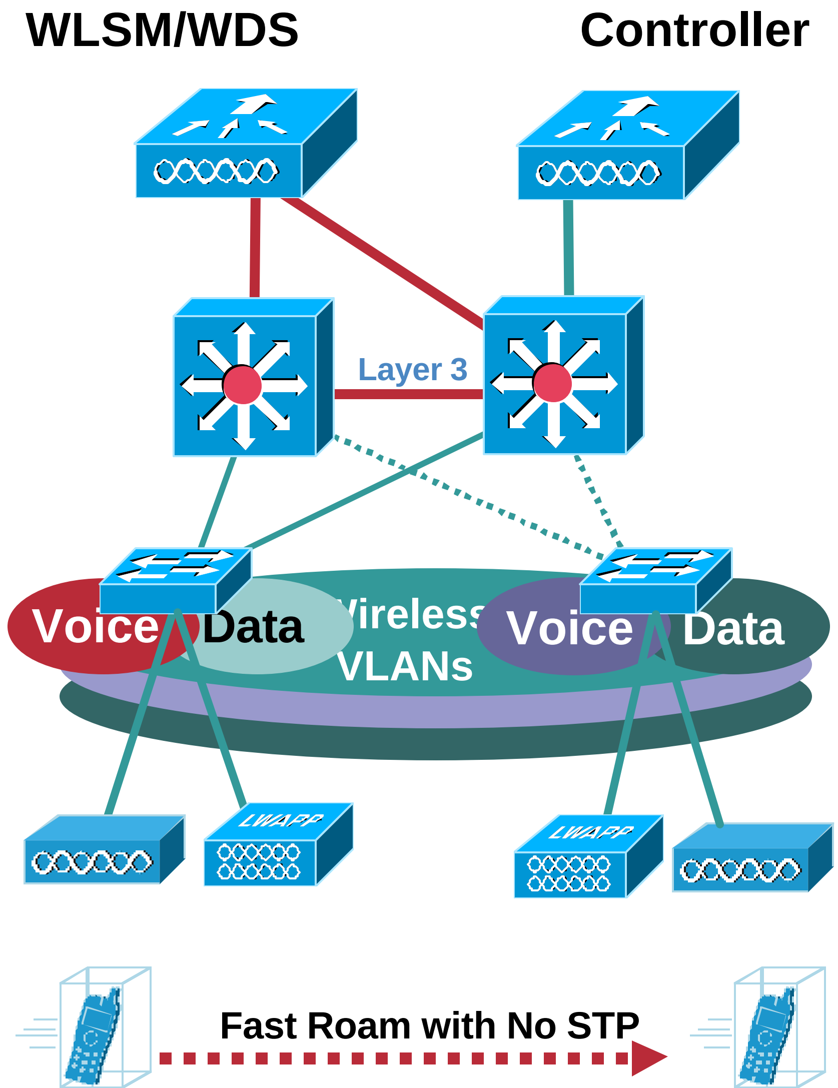
Wireless Integration into the Campus
Non-Controller-Based Wireless



- **Use a 802.1Q trunk for switch to AP connection**
- **Different WLAN authentication/encryption methods require new/distinct VLANs**
- **Layer-2 roaming requires spanning at least 2 VLANs between wiring closet switches**
  1. **Common 'Trunk' or native VLAN for APs to communicate to WDS**
  2. **The Wireless Voice VLAN**

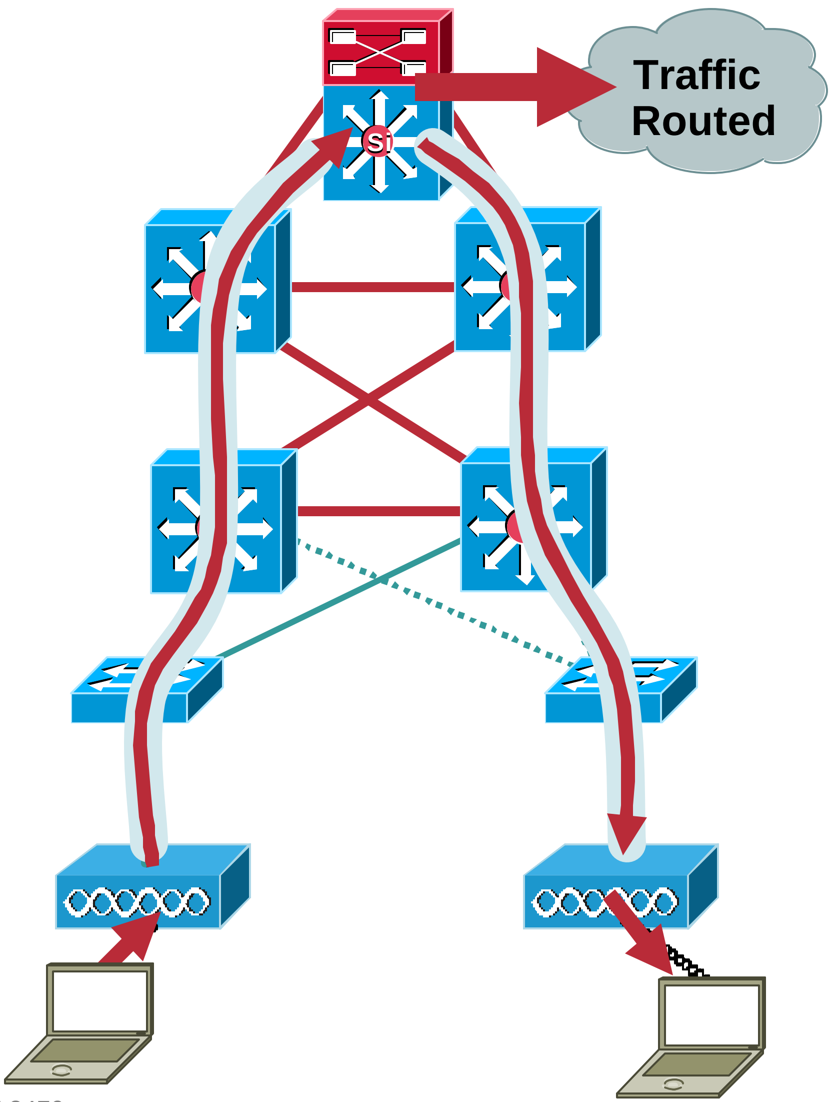
Controller-Based WLAN

The Architectural Shift



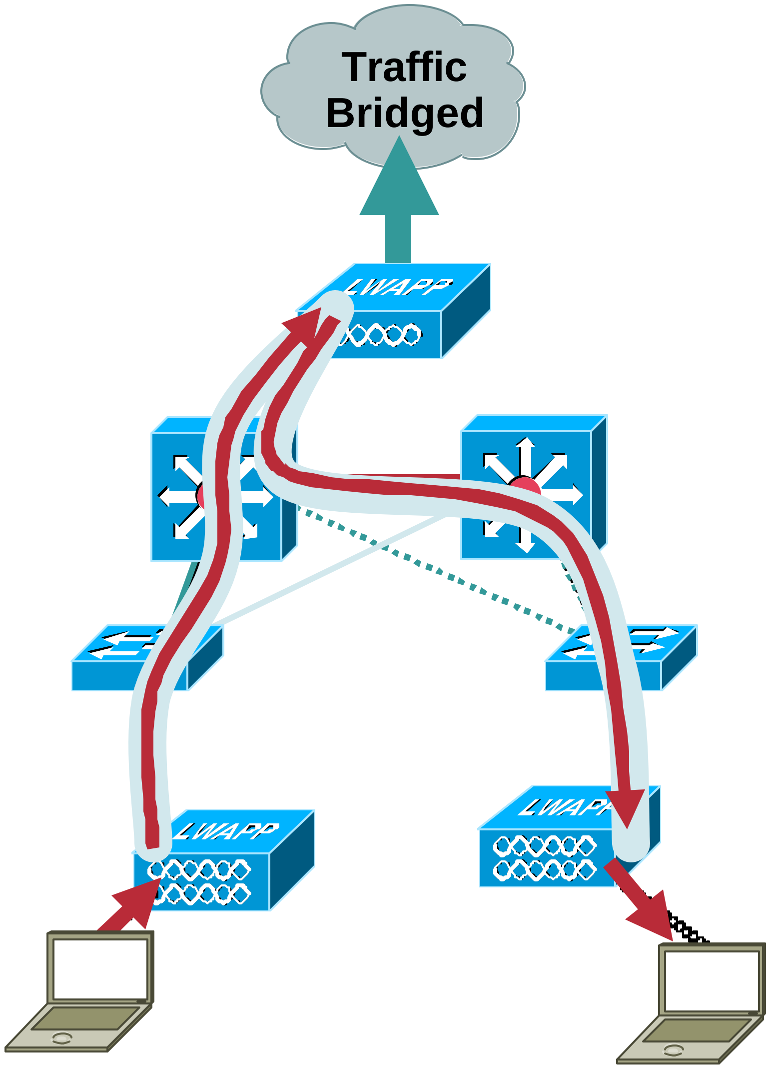
- Wireless LAN Switching Module (WLSM) provides a virtualized centralized Layer 2 domain for each WLAN
- Cisco wireless controller provides for a centralized point to bridge all traffic into the Campus
- AP VLANs are local to the access switch
- No longer a need to span a VLAN between closets
- No spanning tree loops

Wireless LAN Switching Module (WLSM)
Traffic Flows



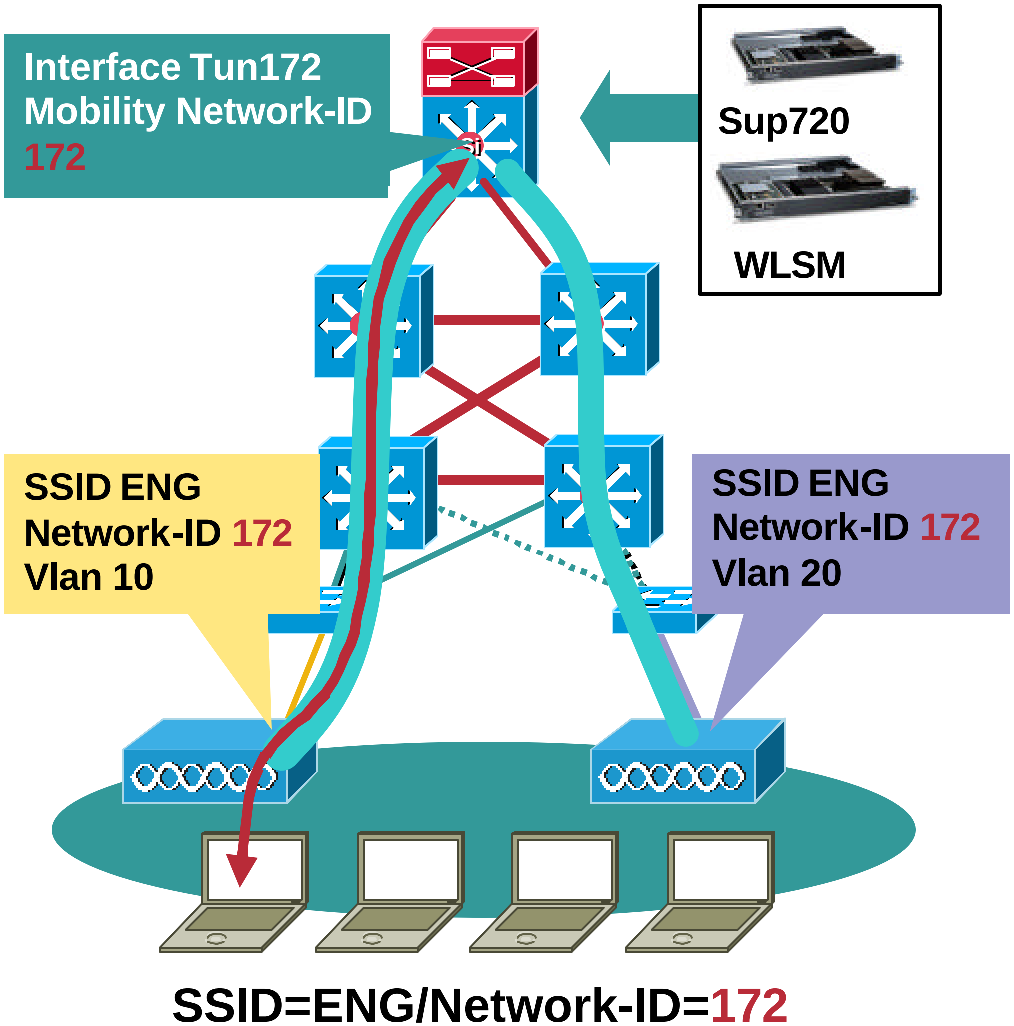
- All traffic from mobile user 1 to mobile user 2 will traverse the GRE tunnel to the Sup720
- Sup720 forwards de-encapsulated packets in HW
- The packet is switched and sent back to the GRE tunnel connected to other AP
- When mobile nodes associate to the same AP traffic still flows via the WSLM/Sup720
- Broadcast traffic either proxied by AP (ARPs) or forwarded to Sup720 (DHCP)
- Traffic to non-APs is routed to the rest of the network

Cisco Wireless Controller
Traffic Flows



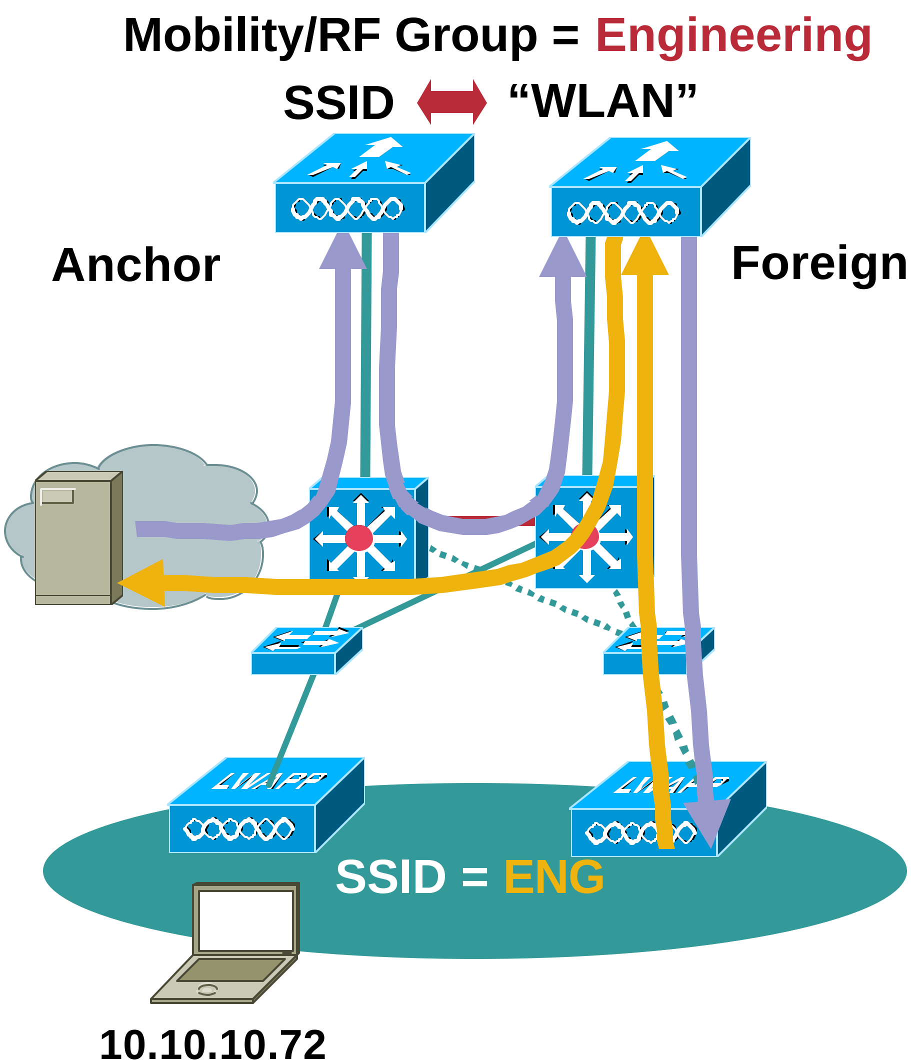
- **Data is tunneled to the Controller in Light Weight Access Point Protocol (LWAPP) transport layer**
- **AP and Controller operate in "Split-MAC" mode dividing the 802.11 functions**
- **The packet bridged onto the wired network uses the MAC address of the original wireless frame**
- **Layer 2 LWAPP is in an Ethernet frame (Ethertype 0xBBBB)**
- **Layer 3 LWAPP is in a UDP / IP frame**
  - **Control traffic uses source port 1024 and destination port 12223**
  - **Data traffic uses source port 1024 and destination 12222**

The Architectural Shift: WLSM
Network-ID Replaces the "VLAN"



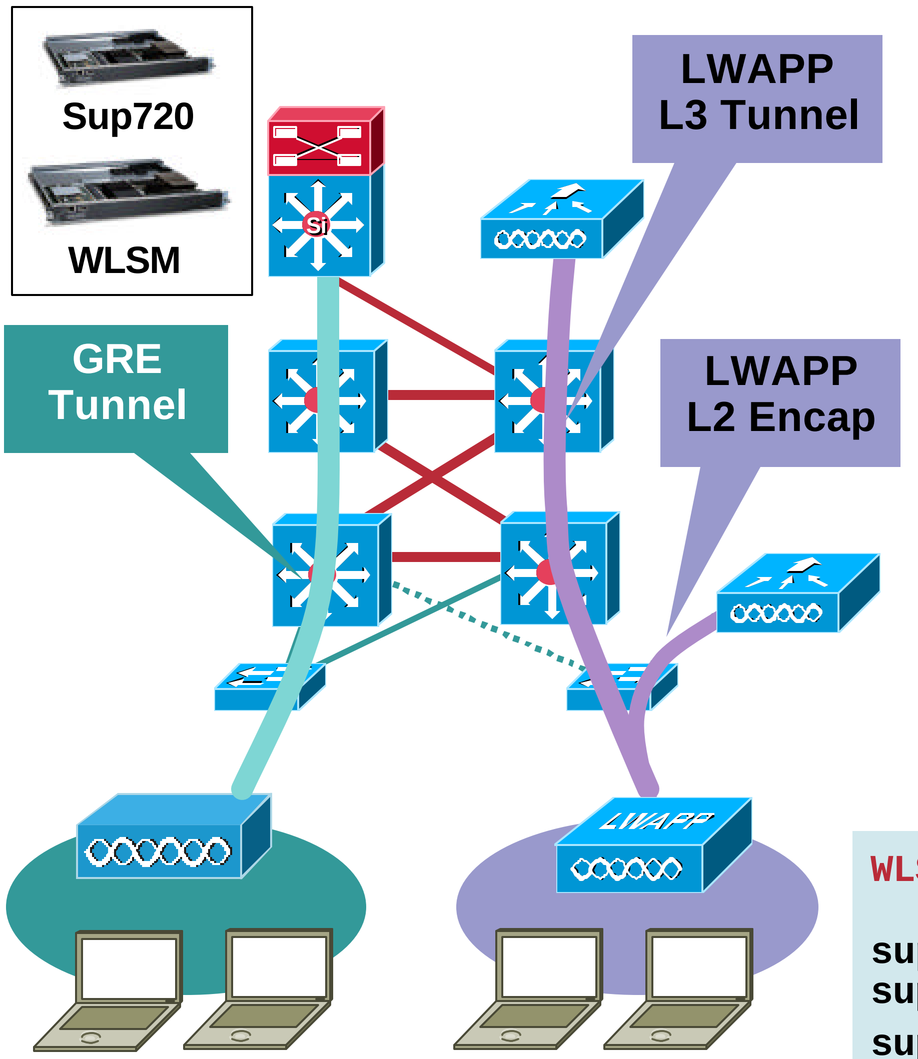
- A Mobility Group is identified by mapping a SSID to a network-ID
- It replaces the mapping of SSID to a wired VLAN
- Define the same SSID Network-ID pair on all APs where mobility is required
- One mGRE tunnel interface is created for each Mobility Group on Sup720
- One SSID/Network-ID = one subnet

The Architectural Shift: Controller
Controllers Virtualize the "VLAN"



- **An SSID is configured with a "WLAN" identifier**
- **The "WLAN" is configured in in all Controllers that define the "Mobility Group" or roaming region**
- **When a client performs an L3 roam, traffic from the client is bridged directly to the network from the foreign controller**
- **Return path traffic is forwarded to the anchor controller**
- **Anchor forwards traffic to the foreign controller**

Design Considerations
LWAPP and GRE Tunnel Traffic



- There must be 'no' NAT between WLSM/WDS and the APs
- If WLSM behind a Firewall open WLCCP (UDP 2887) and GRE (47)
- GRE adds 24 bytes of header therefore need to tune MTU and MSS adjust on the Wireless subnet
- L3 LWAPP adds 94 bytes of headers
- LWAPP AP and Controller will fragment packets if network not configured to support Jumbo frames

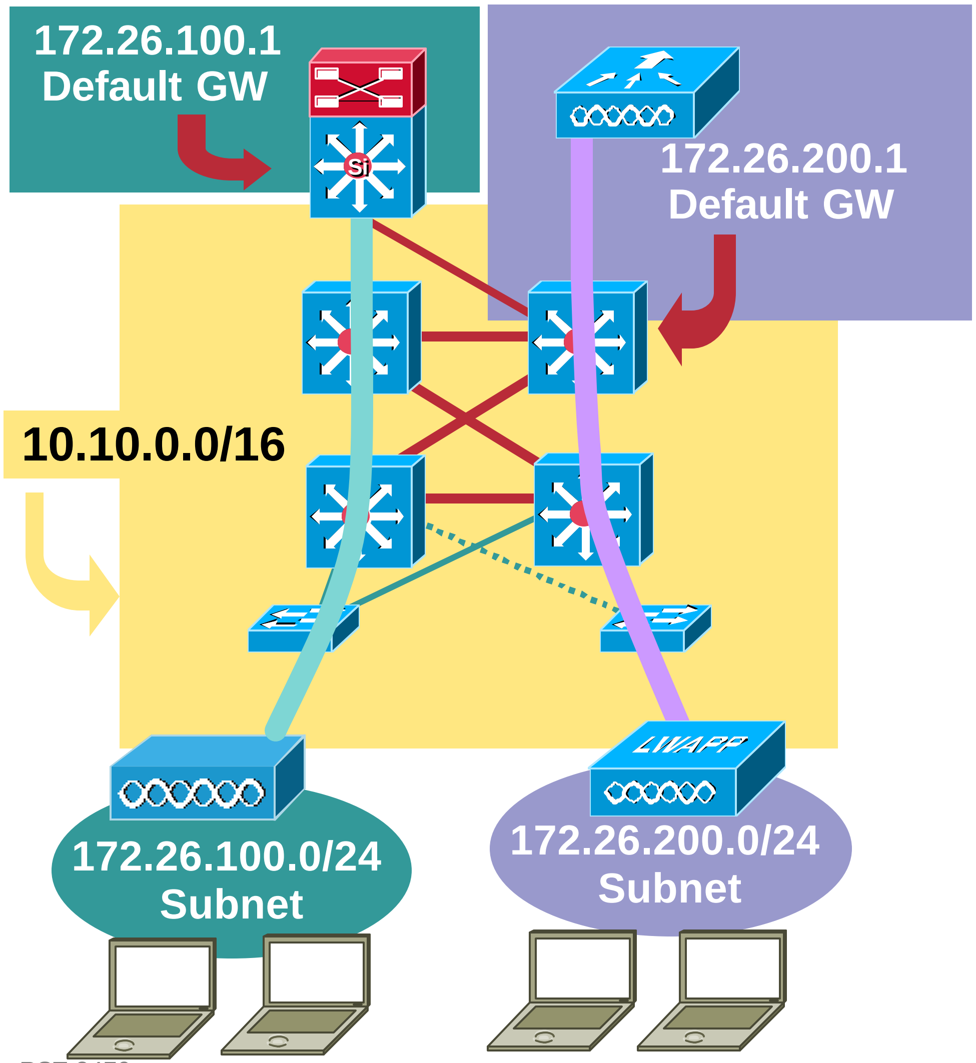
WLSM Switch Config (Cat6k Sup720)

```
sup720(config)#int tunnel 172
sup720(config-if)#ip mtu 1476
sup720(config-if)#mobility tcp adjust-mss
```

Design Considerations

IP Addressing Considerations

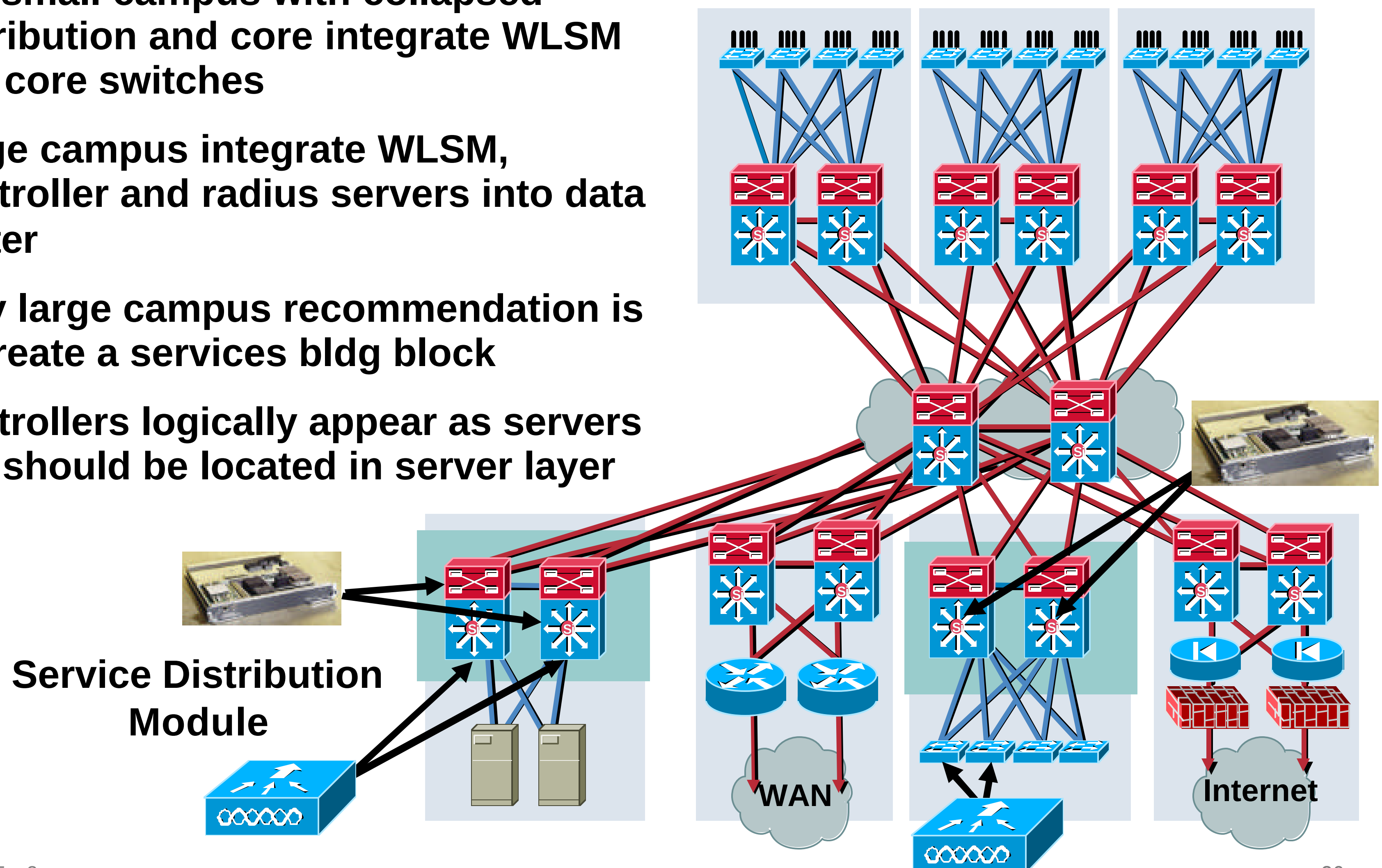
Cisco.com



- The default gateway for all wireless endpoints when using a WLSM Controller is the WLSM Switch
- The default gateway for all wireless endpoints when using a Cisco Controller is the adjacent Catalyst® switch
- The wireless mobile node endpoints are addressed out of the summary range as defined by the location of the controller or the WLSM switch
- Communication between a wired client on an access switch and a wireless client is via the core

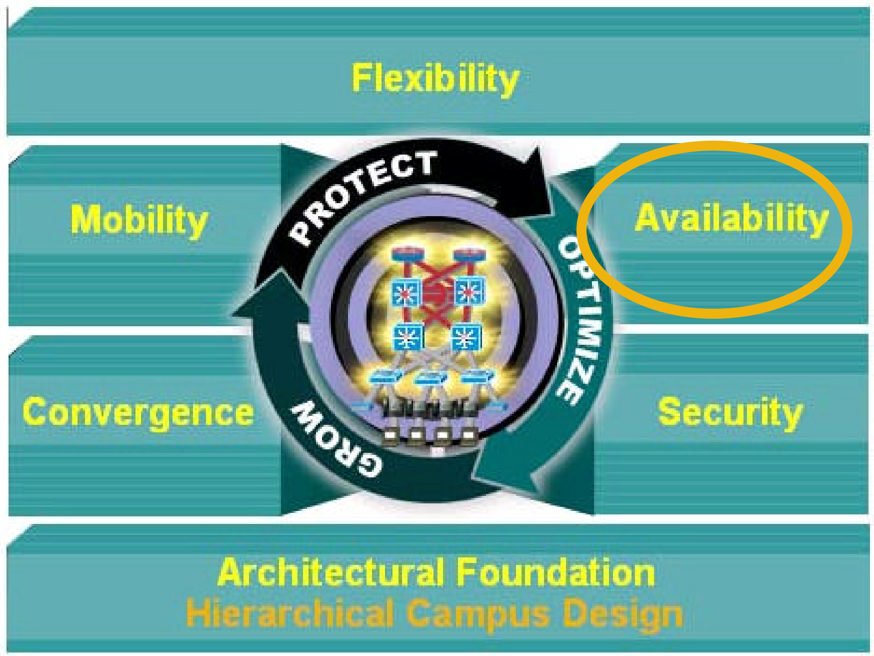
Design Considerations
Location of Controllers

- **In a small campus with collapsed distribution and core integrate WLSM into core switches**
- **Large campus integrate WLSM, Controller and radius servers into data center**
- **Very large campus recommendation is to create a services bldg block**
- **Controllers logically appear as servers and should be located in server layer**



Agenda

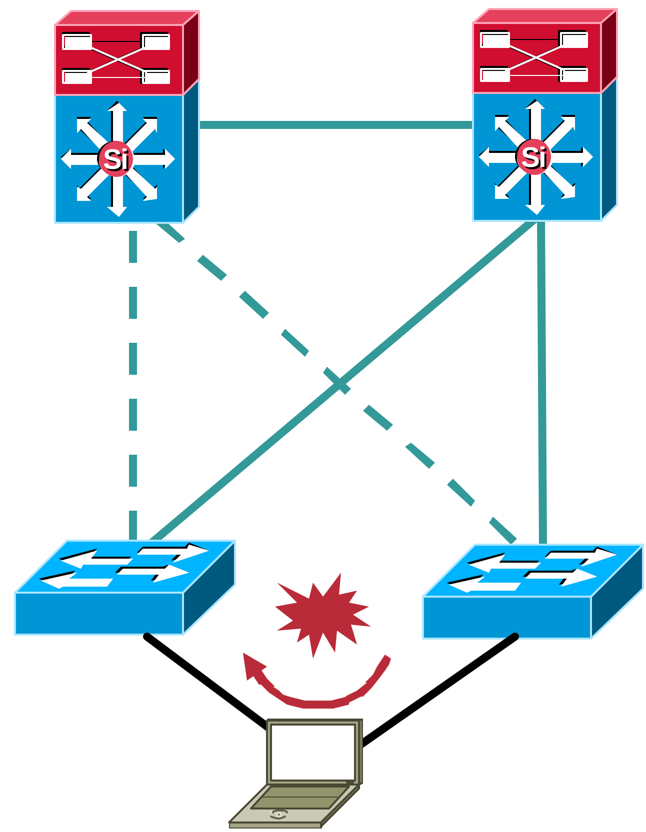
- Foundational Design Review
- Convergence—IP Communications
- Wireless LAN and Wireless Mobility
- High Availability
    - Alternatives to STP
    - Device HA (NSF/SSO and Stackwise)
    - Resilient Network Design
- Segmentation and Virtualization
    - Access Control (IBNS and NAC)
    - Segmentation
- Questions and Answers



Flex-Link

Link Redundancy

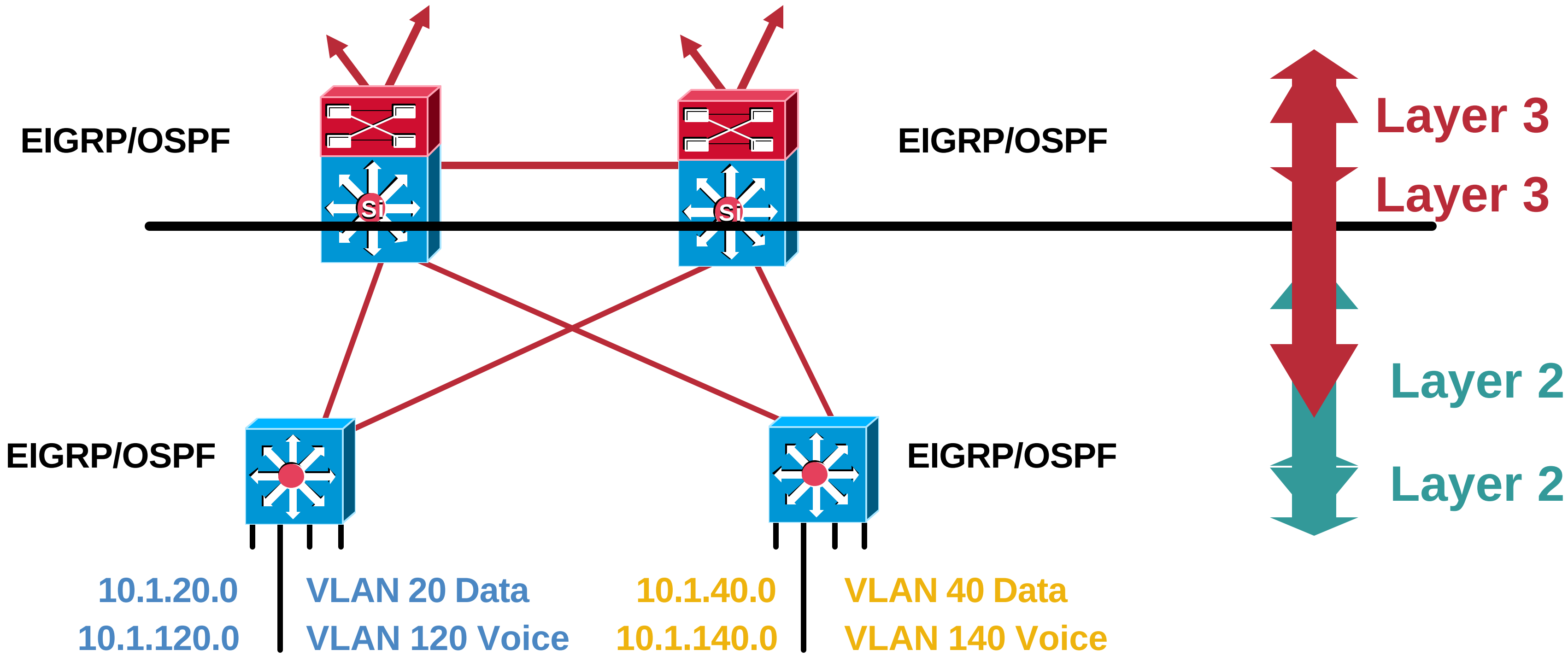
- **Flex-link provides a box local link redundancy mechanism**
- **On failure of the prime link the backup link will start forwarding**
- **Spanning tree is not involved in link recovery however the network is ‘not’ L2 loop free**
- **Spanning Tree should still be configured on access and distribution switches**
- **Flex-link reduces size of the spanning tree topology but does *not* make the network loop free**
- **Supported on 2970, 3550, 3560, 3750 & 6500**



```
interface GigabitEthernet0/1
 switchport backup interface GigabitEthernet0/2
```

Routing to the Edge
Layer 3 Distribution with Layer 3 Access

Cisco.com

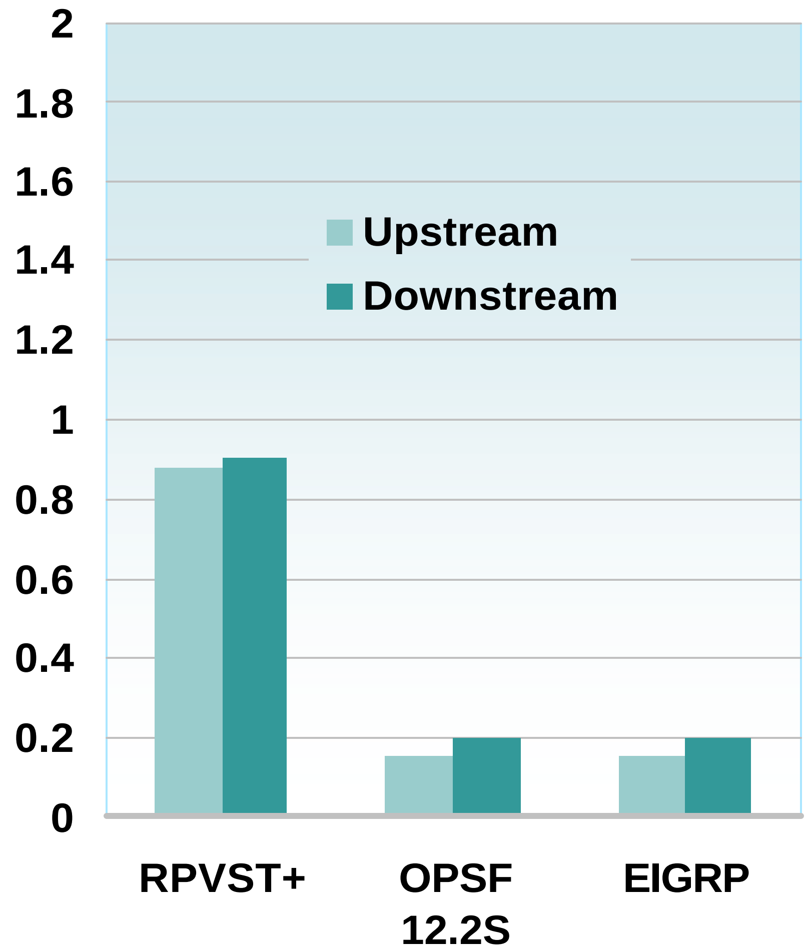


- Move the Layer 2/3 demarcation to the network edge
- Upstream convergence times triggered by hardware detection of light lost from upstream neighbor
- Beneficial for the right environment

Routing to the Edge
Advantages, Yes in the Right Environment

- **Ease of implementation, less to get right**
  - No matching of STP/HSRP/GLBP priority
  - No L2/L3 Multicast topology inconsistencies
- **Single Control Plane and well known tool set**
  - traceroute, show ip route, show ip eigrp neighbor, etc.…
- **Most Catalysts support L3 Switching today**
- **EIGRP converges in <200 msec**
- **OPSF with sub-second tuning converges in <200 msec**
- **RPVST+ convergence times dependent on GLBP/ HSRP tuning**

Both L2 and L3 Can Provide Sub-Second Convergence

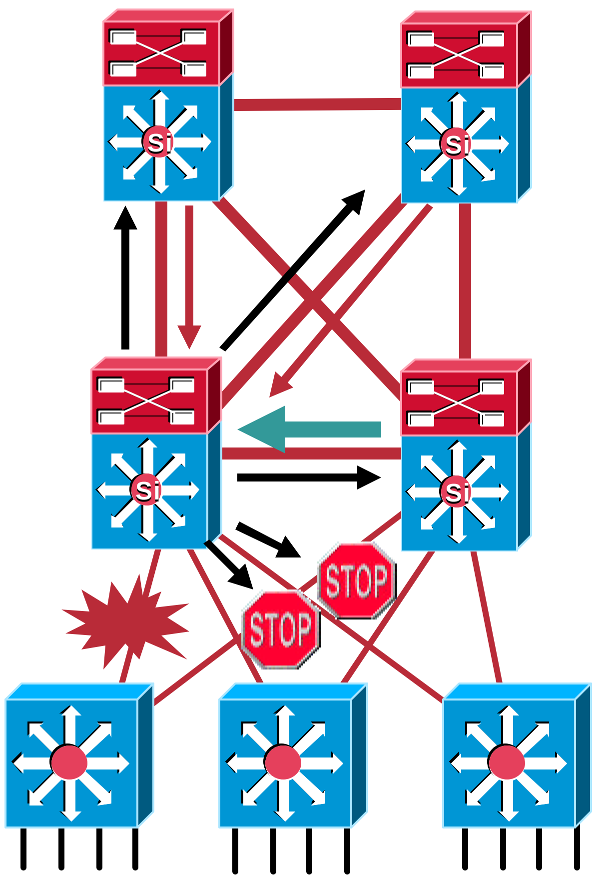


EIGRP Design Rules for HA Campus

High-Speed Campus Convergence

Cisco.com

- EIGRP convergence is largely dependent on query response times
- Minimize the number and time for query response to speed up convergence
- Summarize distribution block routes upstream to the core
- Configure all access switches as EIGRP stub routers
- Filter routes sent down to access switches




```
interface TenGigabitEthernet 4/1
 ip summary-address eigrp 100 10.120.0.0 255.255.0.0 5

router eigrp 100
 network 10.0.0.0
 distribute-list Default out <mod/port>

ip access-list standard Default
 permit 0.0.0.0
```

```
router eigrp 100
 network 10.0.0.0
 eigrp stub connected
```

OSPF Design Rules for HA Campus

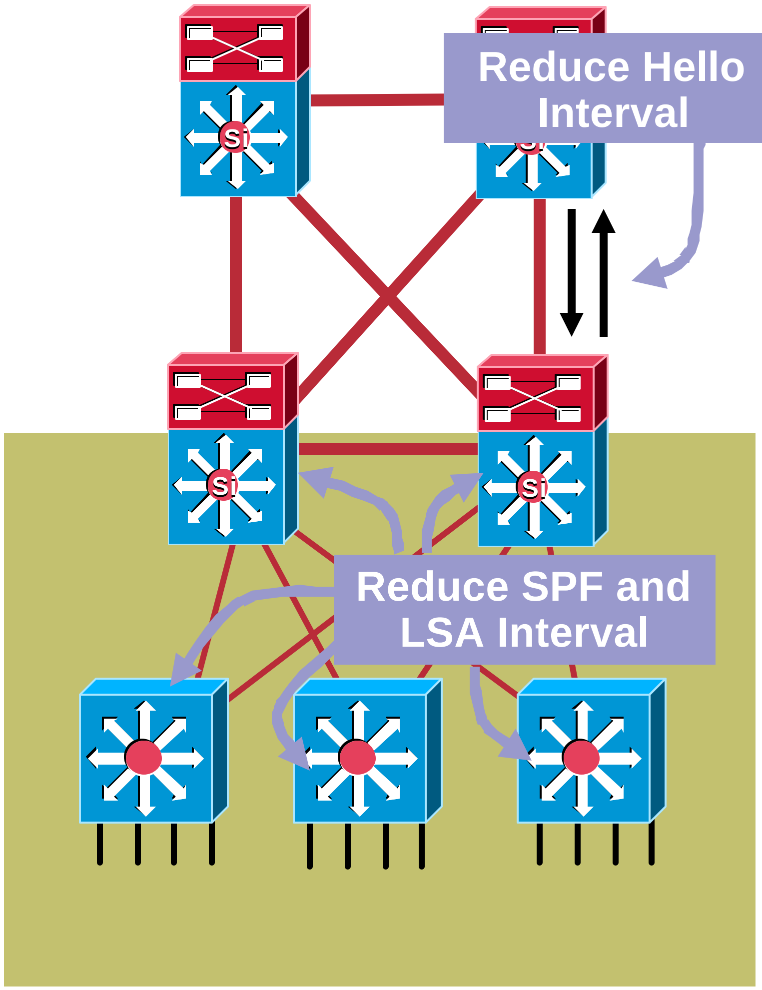
High-Speed Campus Convergence

Cisco.com

- **OSPF convergence is largely dependent on time to compute Dijkstra response times**
- **In a full meshed design key tuning parameters are spf throttle and lsa throttle**
- **Utilize Totally Stubby area design to control number of routes in access switches**
- **Hello and Dead are secondary failure detection mechanism**

```
router ospf 100
 router-id 10.122.102.2
 log-adjacency-changes
 area 120 stub no-summary
 area 120 range 10.120.0.0 255.255.0.0
 timers throttle spf 10 100 5000
 timers throttle lsa all 10 100 5000
 timers lsa arrival 80
 network 10.120.0.0 0.0.255.255 area 120
 network 10.122.0.0 0.0.255.255 area 0
```

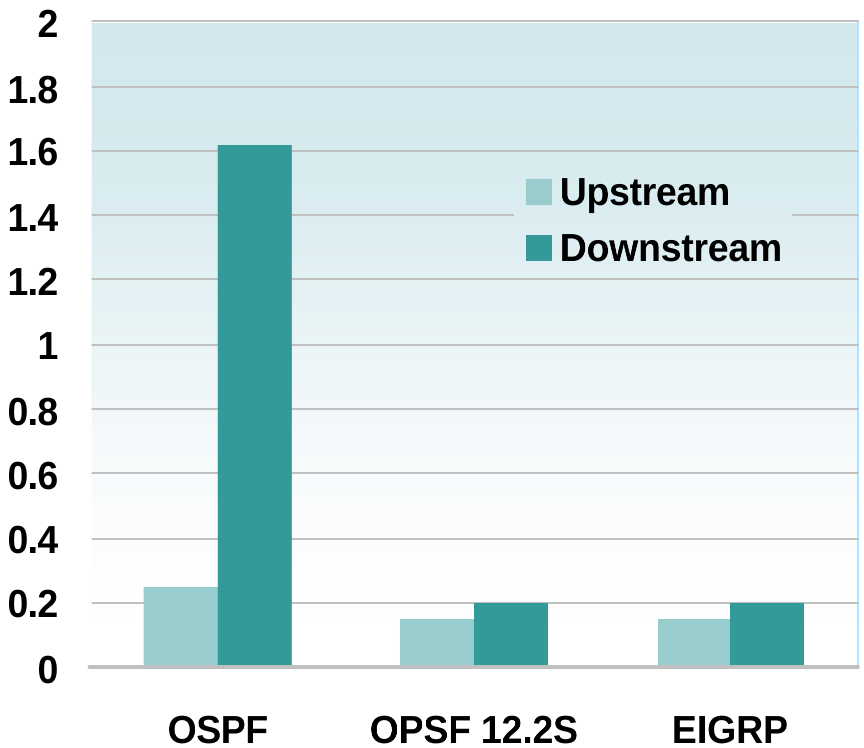
```
interface GigabitEthernet5/2
 ip address 10.120.100.1 255.255.255.254
 ip ospf dead-interval minimal hello-multiplier 4
```



EIGRP vs. OSPF as Your Campus IGP
DUAL vs. Dijkstra

Cisco.com

Both Can Provide Subsecond Convergence

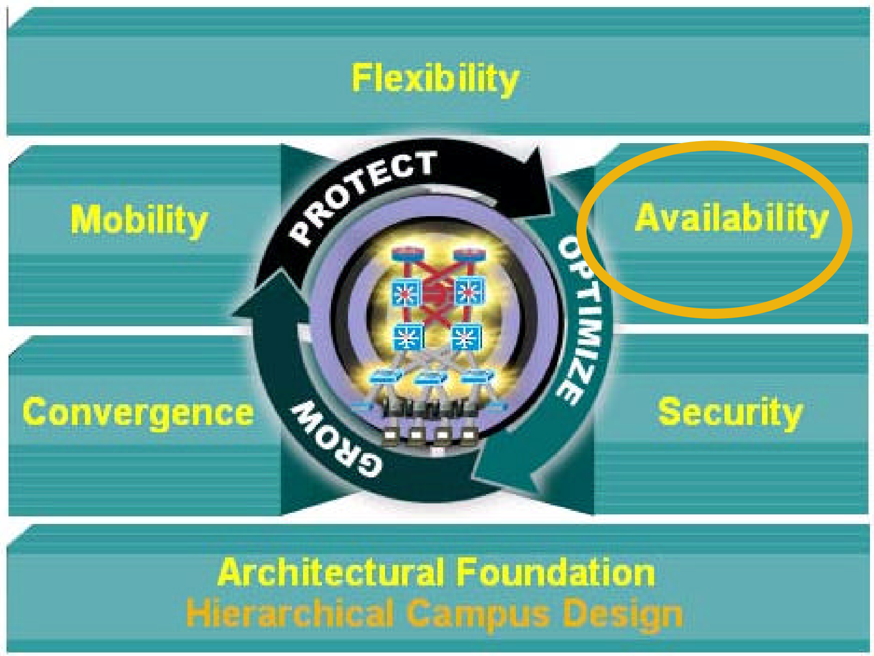


- **Convergence:**
  - Within the campus environment, both EIGRP and OSPF provide extremely fast convergence
  - EIGRP requires summarization and, OSPF requires LSA and SPF timer tuning for fast convergence
- **Flexibility:**
  - EIGRP supports multiple levels of route summarization and route filtering which simplifies migration from the traditional Multilayer L2/L3 Campus design
  - OSPF Area design restrictions need to be considered
- **Scalability:**
  - Both protocols can scale to support very large Enterprise Network topologies

For More Discussion on Routed Access Design Best Practices—RST-2031

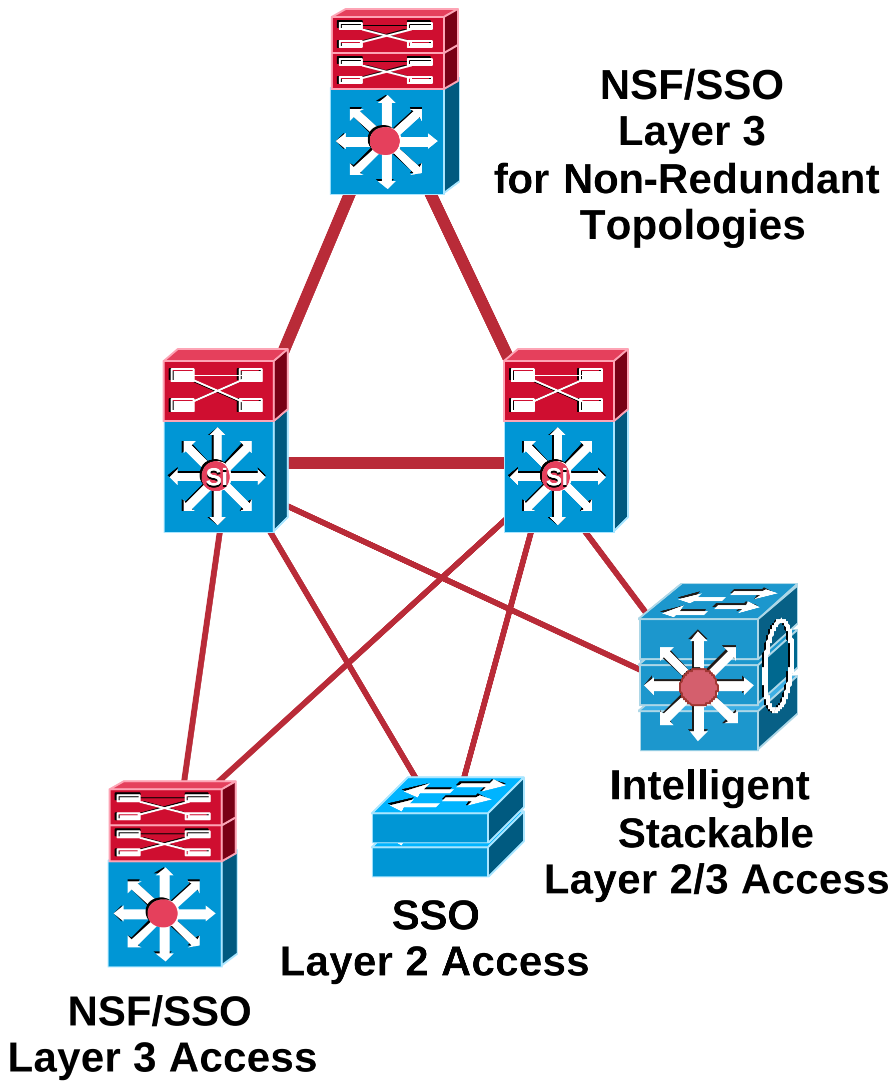
Agenda

- Foundational Design Review
- Convergence—IP Communications
- Wireless LAN and Wireless Mobility
- High Availability
  - Alternatives to STP
  - Device HA (NSF/SSO and Stackwise)
  - Resilient Network Design
- Segmentation and Virtualization
  - Access Control (IBNS and NAC)
  - Segmentation
- Questions and Answers



Device High Availability

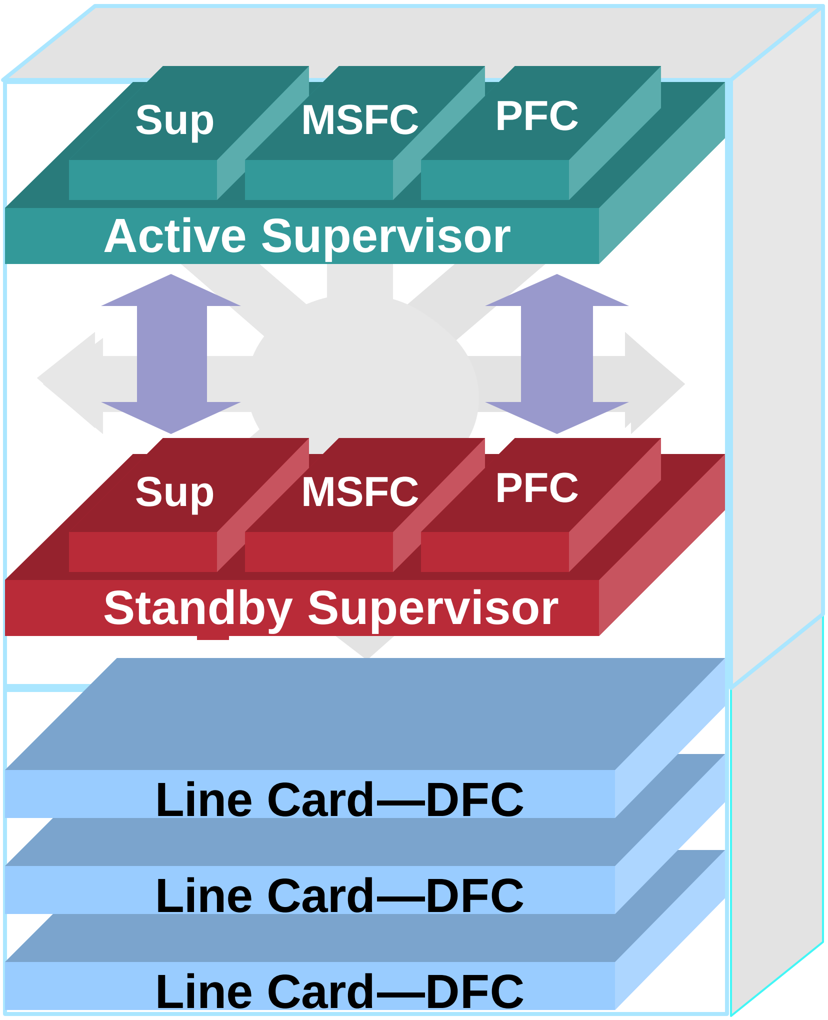
NSF/SSO and 3750 Stackwise



- **Overall availability of the infrastructure is dependent on the weakest link**
- **NSF/SSO provides improved availability for single points of failure**
- **SSO provides enhanced redundancy for traditional Layer 2 edge designs**
- **NSF/SSO provides enhanced L2/L3 redundancy for routed to the edge designs**
- **3750 Stackable provides improved redundancy for L2 and L3 edge designs**

Supervisor Processor Redundancy

Stateful Switch Over (SSO)

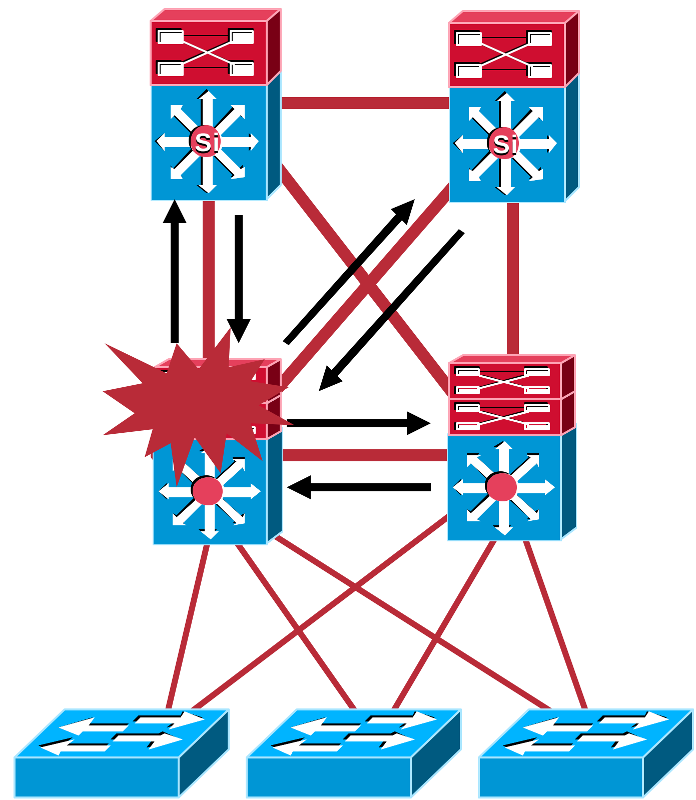


- **Active/standby supervisors run in synchronized mode**
- **Redundant MSFC is in 'hot-standby' mode**
- **Switch processors synchronize L2 port state information, (e.g. STP, 802.1x, 802.1q,…)**
- **PFCs synchronize L2/L3 FIB, Netflow and ACL tables**
- **DFCs are populated with L2/L3 FIB, Netflow and ACL tables**

Non-Stop Forwarding (NSF)

NSF Recovery

1. **DFC enabled line cards continue to forward based on existing FIB entries**
2. **Following SSO recovery and activation of standby Sup synchronized PFC continues to forward traffic based on existing FIB entries**
3. **"Hot-Standby" MSFC RIB is detached from the FIB isolating FIB from RP changes**
4. **"Hot-Standby" MSFC activates routing processes in NSF recovery mode**
5. **MSFC re-establishes adjacency indicating this is an NSF restart**
6. **Peer updates restarting MSFC with it's routing information**
7. **Restarting MSFC sends routing updates to the peer**
8. **RIB reattaches to FIB and PFC and DFCs updated with new FIB entries**

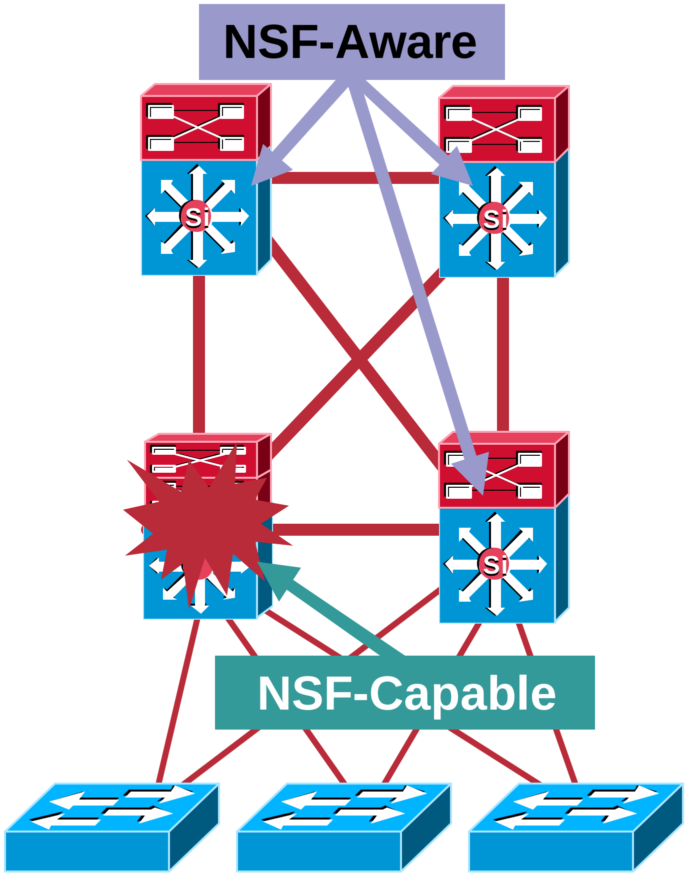


No Route Flaps During Recovery

Non-Stop Forwarding (NSF)
NSF Capable vs. NSF Awareness

Cisco.com

- **Two roles in NSF neighbor graceful restart**
  - **NSF Capable**
  - **NSF Aware**
- **An NSF-Capable router is ‘capable’ of continuous forwarding while undergoing a switchover**
- **An NSF-Aware router is able to assist NSF-Capable routers by:**
  - **Not resetting adjacency**
  - **Supplying routing information for verification after switchover**
- **NSF capable and NSF aware peers cooperate using Graceful Restart extensions to BGP, OSPF, ISIS and EIGRP protocols**



Design Considerations for NSF/SSO

NSF and Hello Timer Tuning?

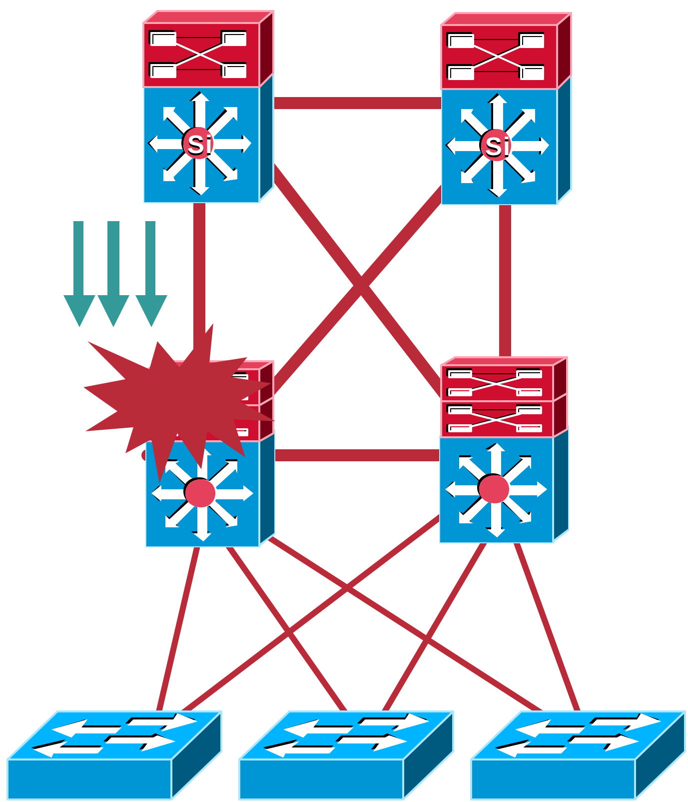
- NSF is intended to provide availability through route convergence avoidance
- Fast IGP timers are intended to provide availability through fast route convergence
- In an NSF environment dead timer must be greater than SSO Recovery + RP restart + time to send first hello

    OPSF 2/8 seconds for hello/dead

    EIGRP 1/4 seconds for hello/hold

- In a Campus environment composed of pt-pt fiber links neighbor loss is detected via loss of light
- RP timers providing a backup recovery role only

Neighbor Loss, No Graceful Restart

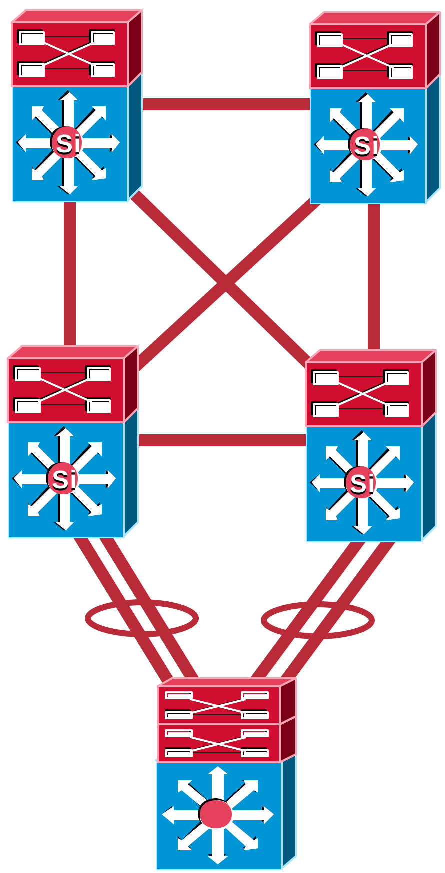


Design Considerations for NSF/SSO

Supervisor Uplinks

Cisco.com

- The use of Supervisor uplinks with NSF/SSO results in a more complex network recovery scenario
- Dual failure scenario
  - Supervisor Failure
  - Port Failure
- During recovery FIB is frozen but uplink port is gone
- PFC tries to forwarded out a non-existent link
- Bundling Supervisor uplinks into Etherchannel links improves convergence
- Optimal NSF/SSO convergence requires the use of DFC enabled line cards



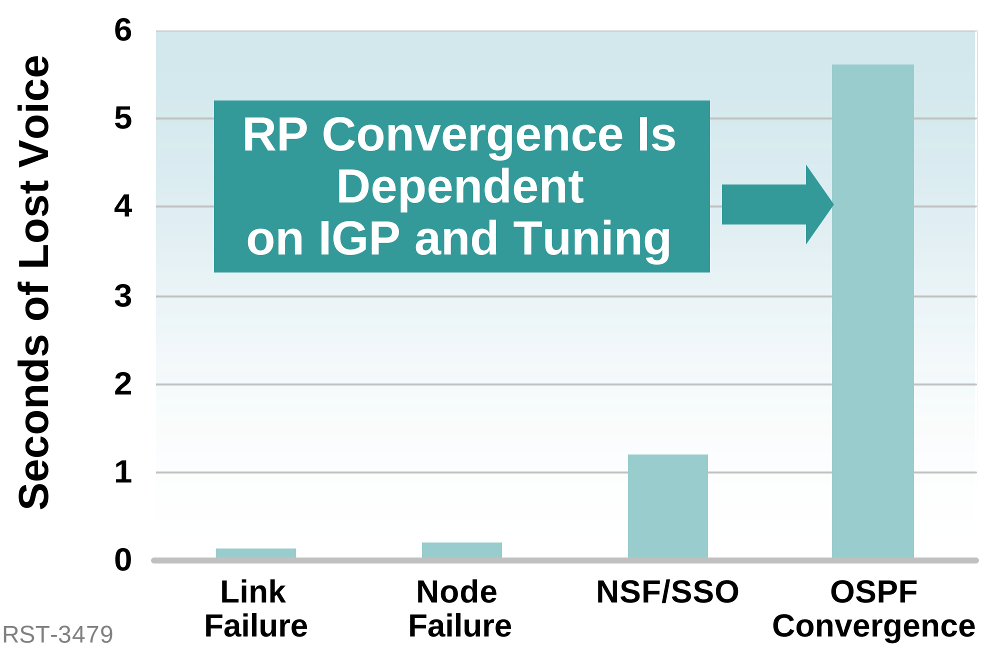
Uplinks on Line Card (msec)	SVI (Etherchannel)	Routed interfaces
920 msec	3100 msec	24 sec

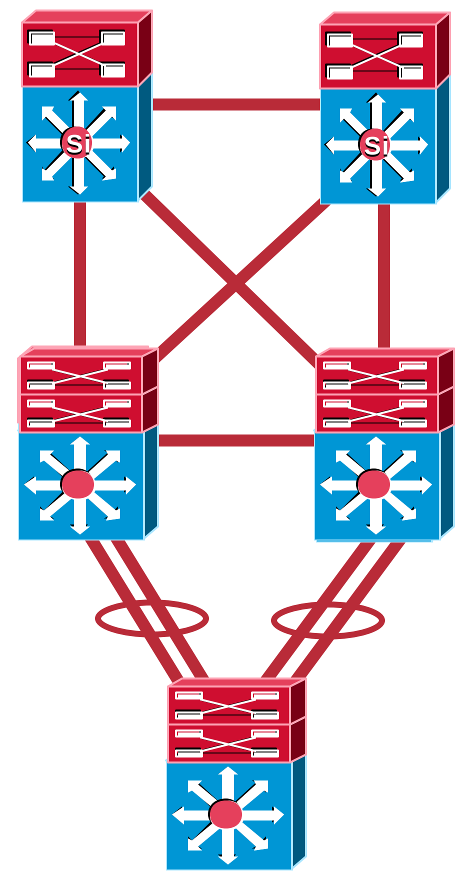
Design Considerations for NSF/SSO

Where Does It Make Sense?

Cisco.com

- **Redundant topologies with equal cost paths provide sub-second convergence**
- **NSF/SSO provides superior availability in environments with non-redundant paths**

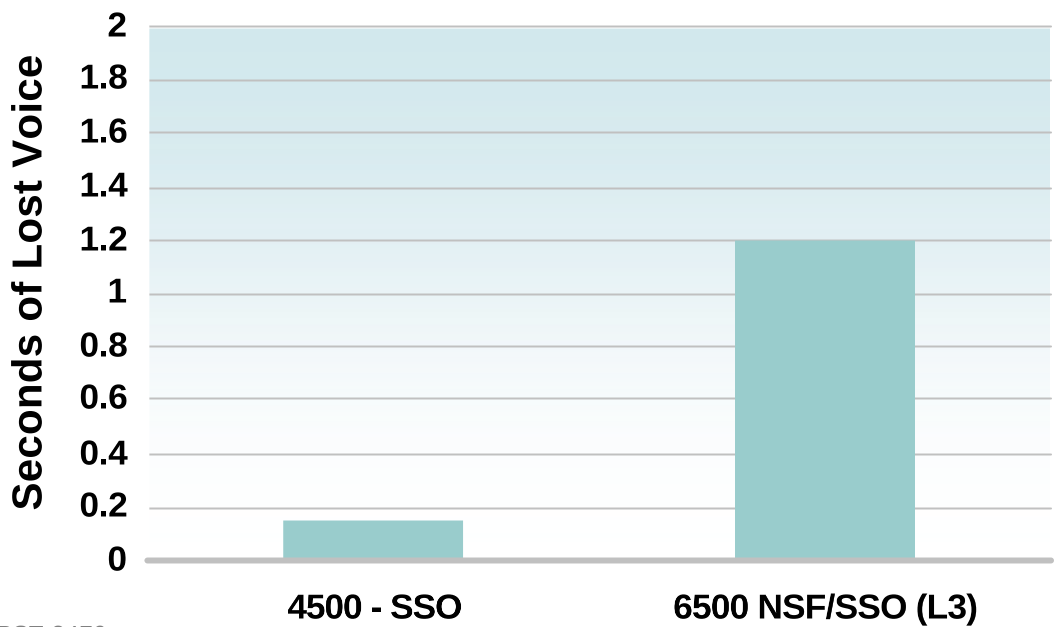


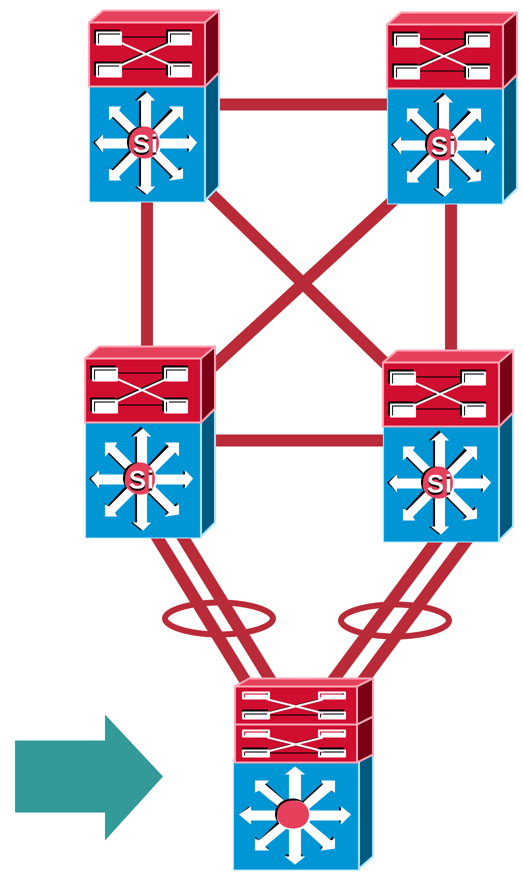


Design Considerations for NSF/SSO

Where Does It Make Sense?

- 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
- SSO provides for sub-second recovery of voice and data traffic
- NSF/SSO provides for sub 1200 msec recovery of voice and data traffic

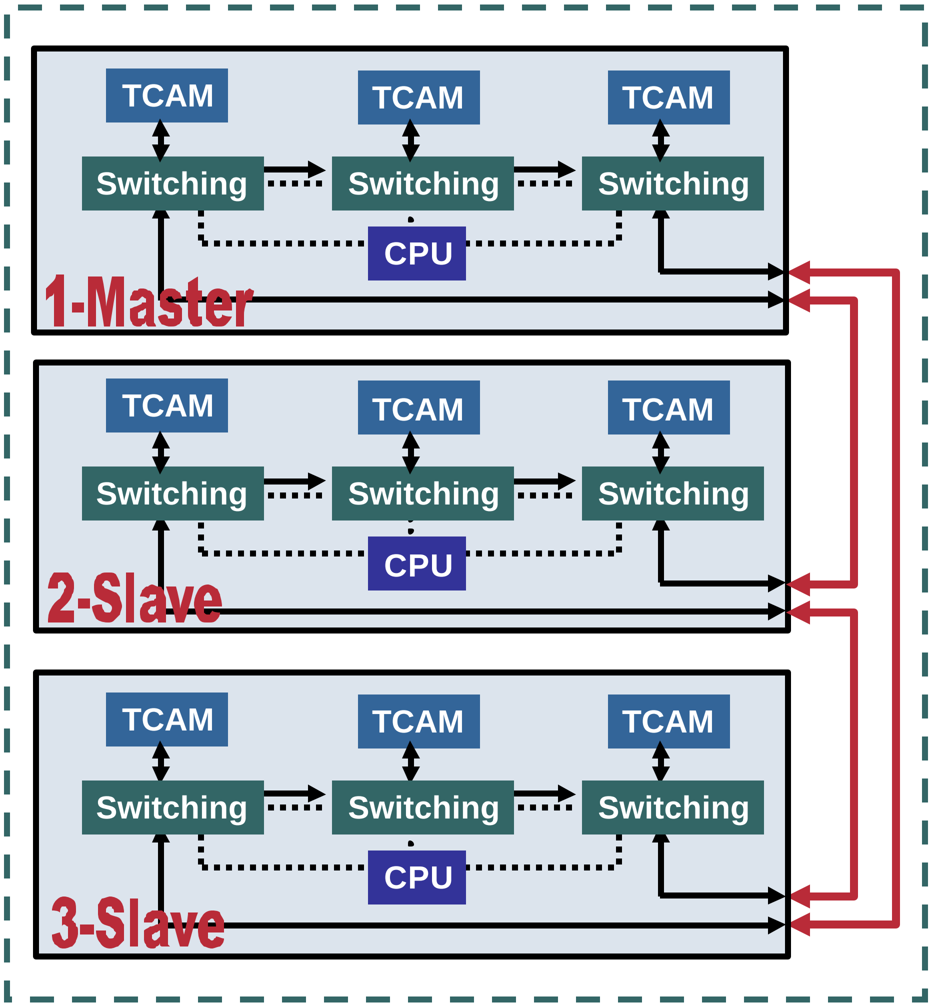




Device High Availability
3750 Stackwise

Cisco.com

- Centralized Configuration and Management
- Switching fabric extended via bidirectional self healing ring
- Each TCAM contains full FIB, ACL and QoS information
- Certain functions are replicated on all switches (e.g. VLAN database, Spanning Tree,…)
- Other functions are managed centrally on the stack master node (e.g. L3 is centrally managed)
- Redundancy is provided via a combination of distributed feature replication and RPR+ like master/slave failover

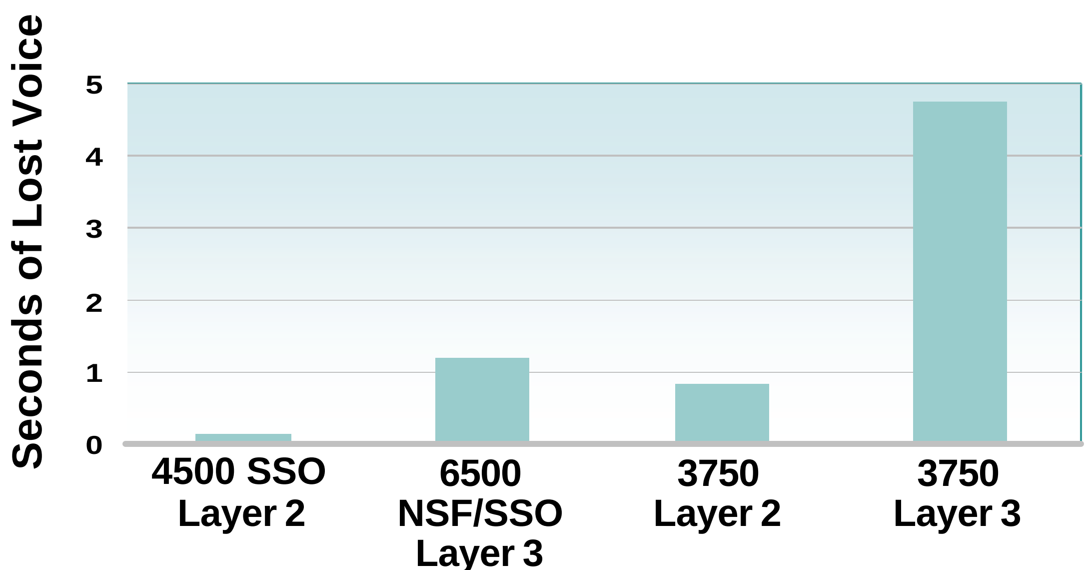


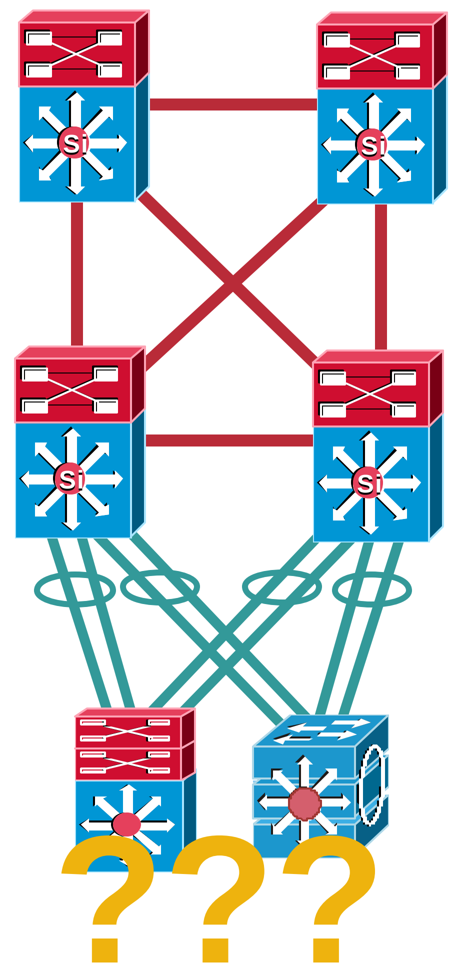
Design Considerations

Chassis vs. Stackable?

Cisco.com

- **Chassis-based systems provide full 1:1 component redundancy**
  - No loss in system switching capacity
  - All edge ports protected
- **NSF/SSO enabled chassis systems provide for both device and network level redundancy**
- **Both provide sub-second L2 convergence**
- **Both support five 9s Campus HA design**

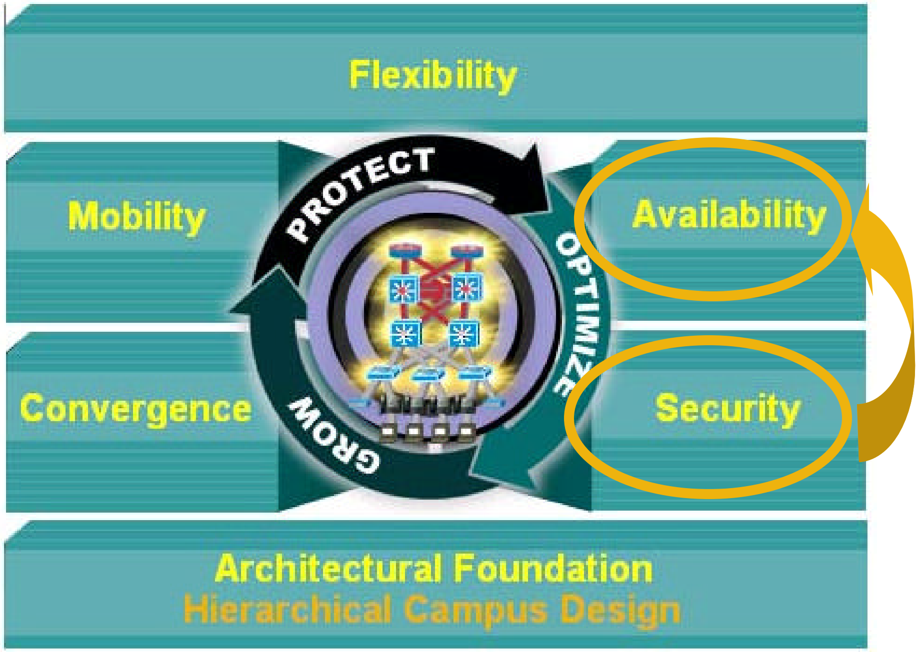




Agenda

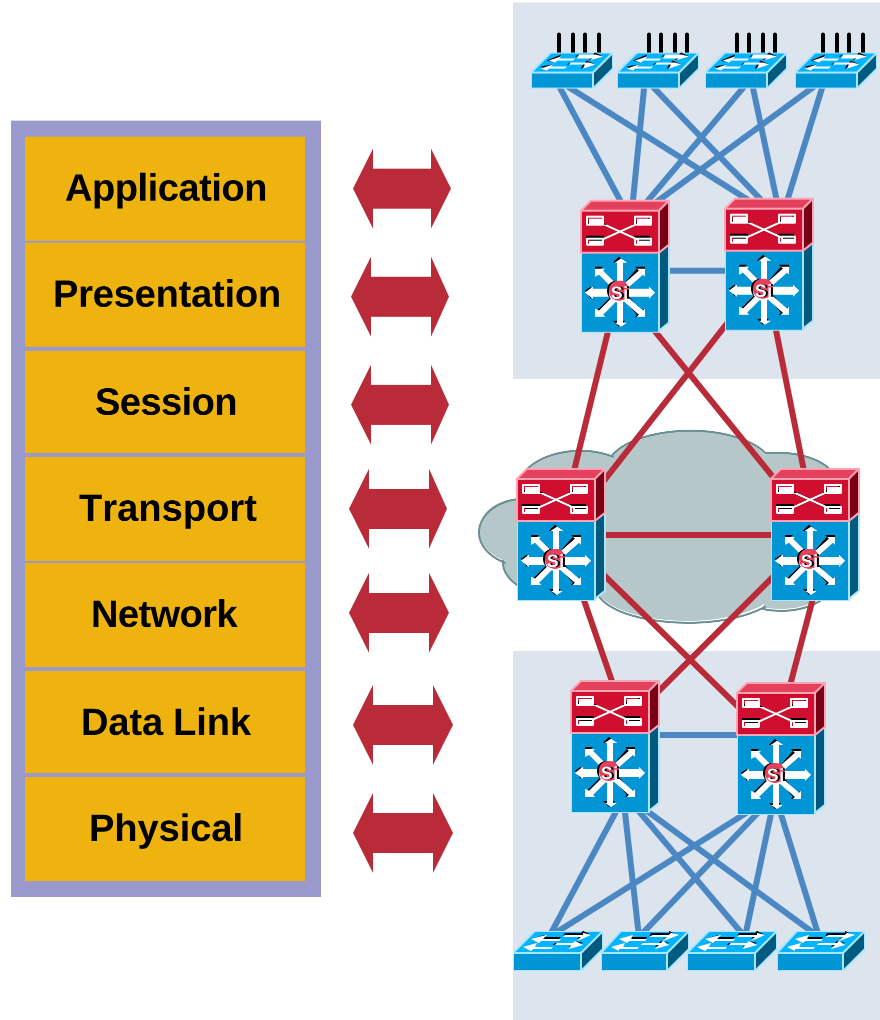
Cisco.com

- Foundational Design Review
- Convergence—IP Communications
- Wireless LAN and Wireless Mobility
- **High Availability**
  - Alternatives to STP
  - Device HA (NSF/SSO and Stackwise)
  - **Resilient Network Design**
- Segmentation and Virtualization
  - Access Control (IBNS and NAC)
  - Segmentation
- Questions and Answers

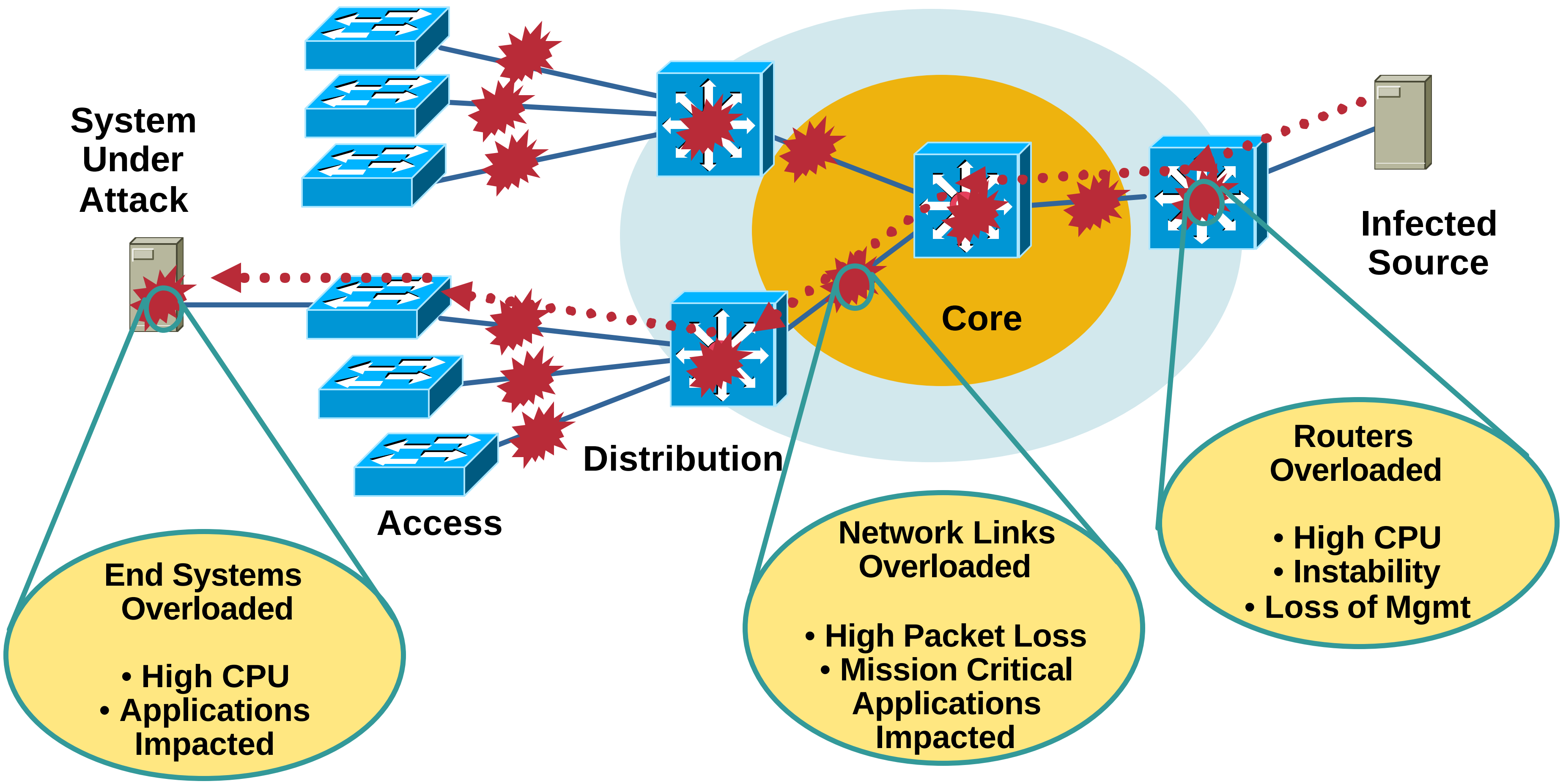


The Resilient Campus Network
Evolution Beyond Structured Design

- **We engineer networks for the expected**
- **We also need to design for the unexpected**
- **Campus design should consider how to prevent or restrict anomalous or bad behaviour**
- **Understand and mitigate the threats at each layer of the network**
- **Protect network resources**



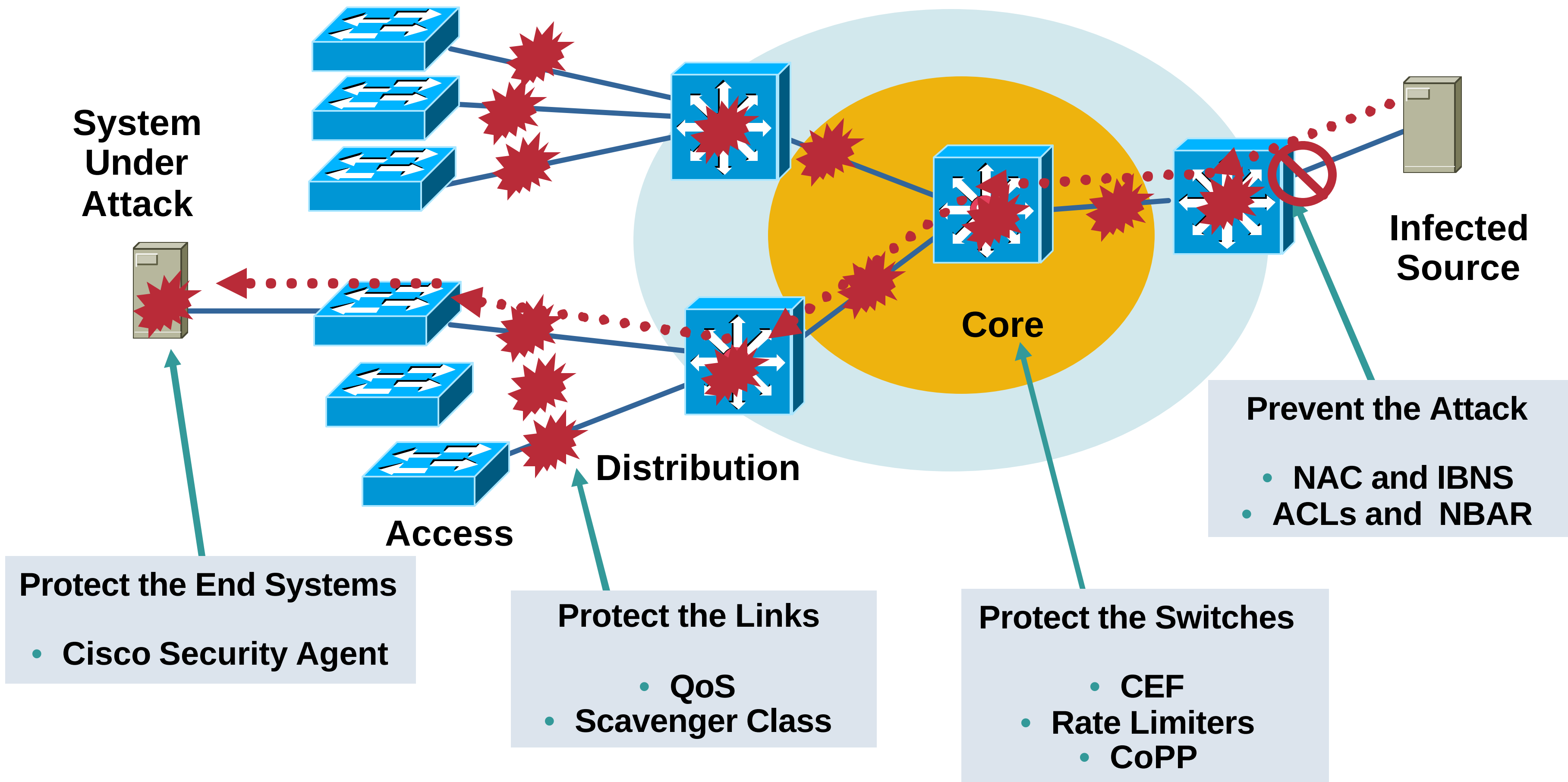
Impact of an Internet Worm
Direct and Collateral Damage



Availability of Networking Resources Impacted by the Propagation of the Worm

Mitigating the Impact

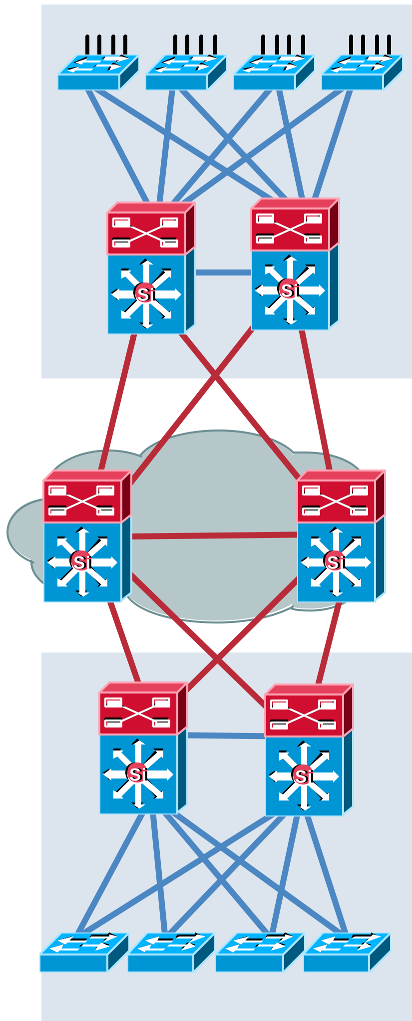
Preventing and Limiting the Pain



Allow the Network to Do What You Designed It to Do but Not What You Didn't

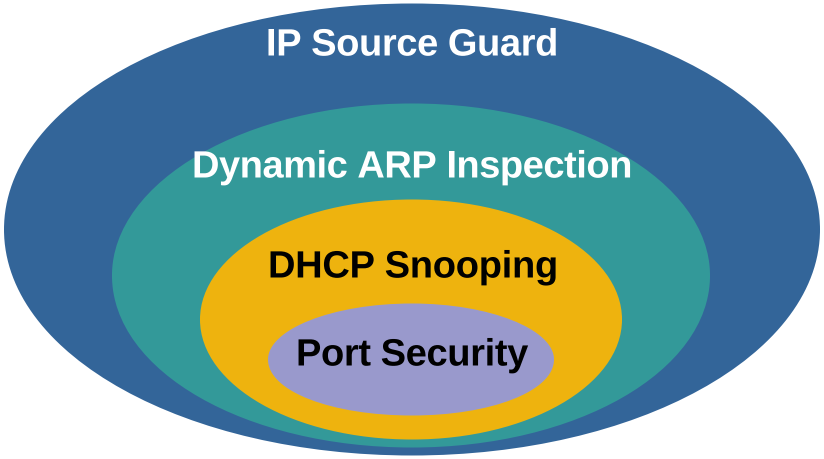
Worms Are Only One Problem
Other Sources of Pain

- **Internet Worms are not the only type of network anomaly**
- **Multiple things can either go wrong or be happening that you want to prevent and/or mitigate**
  - **Spanning Tree Loops**
  - **NICs spewing garbage**
  - **Distributed Denial of Service (DDoS)**
  - **TCP Splicing, ICMP Reset attacks**
  - **Man in the Middle (M-in-M) attacks**
  - **…**



Catalyst Integrated Security Features
Hardening Layer 2/3

Cisco.com

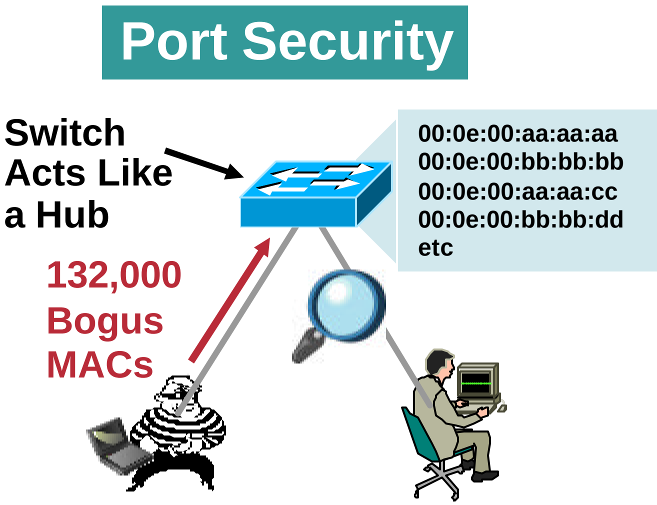


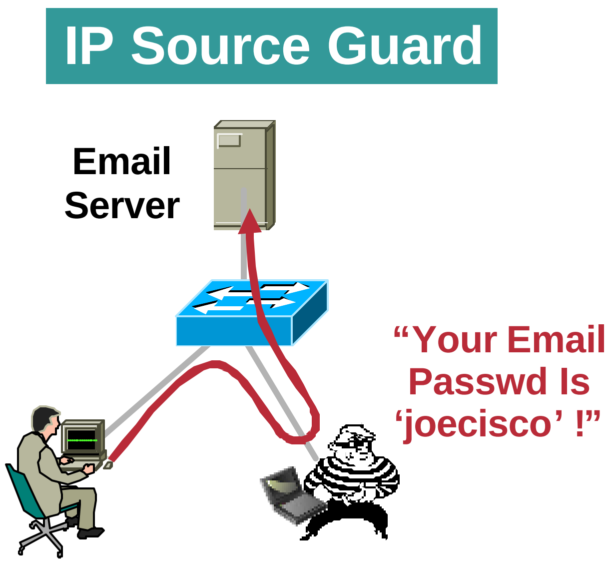
- **Port Security** prevents MAC flooding attacks
- **DHCP Snooping** prevents client attack on the switch and server
- **Dynamic ARP Inspection** adds security to ARP using DHCP snooping table
- **IP Source Guard** adds security to IP source address using DHCP snooping table

```
ip dhcp snooping
ip dhcp snooping vlan 2-10
ip arp inspection vlan 2-10
!
interface fa3/1
switchport port-security
switchport port-security max 3
switchport port-security violation
restrict
switchport port-security aging time 2
switchport port-security aging type
inactivity
ip arp inspection limit rate 100
ip dhcp snooping limit rate 100
!
interface gigabit1/1
ip dhcp snooping trust
ip arp inspection trust
```

Catalyst Integrated Security Features
Hardening Layer 2/3

Cisco.com



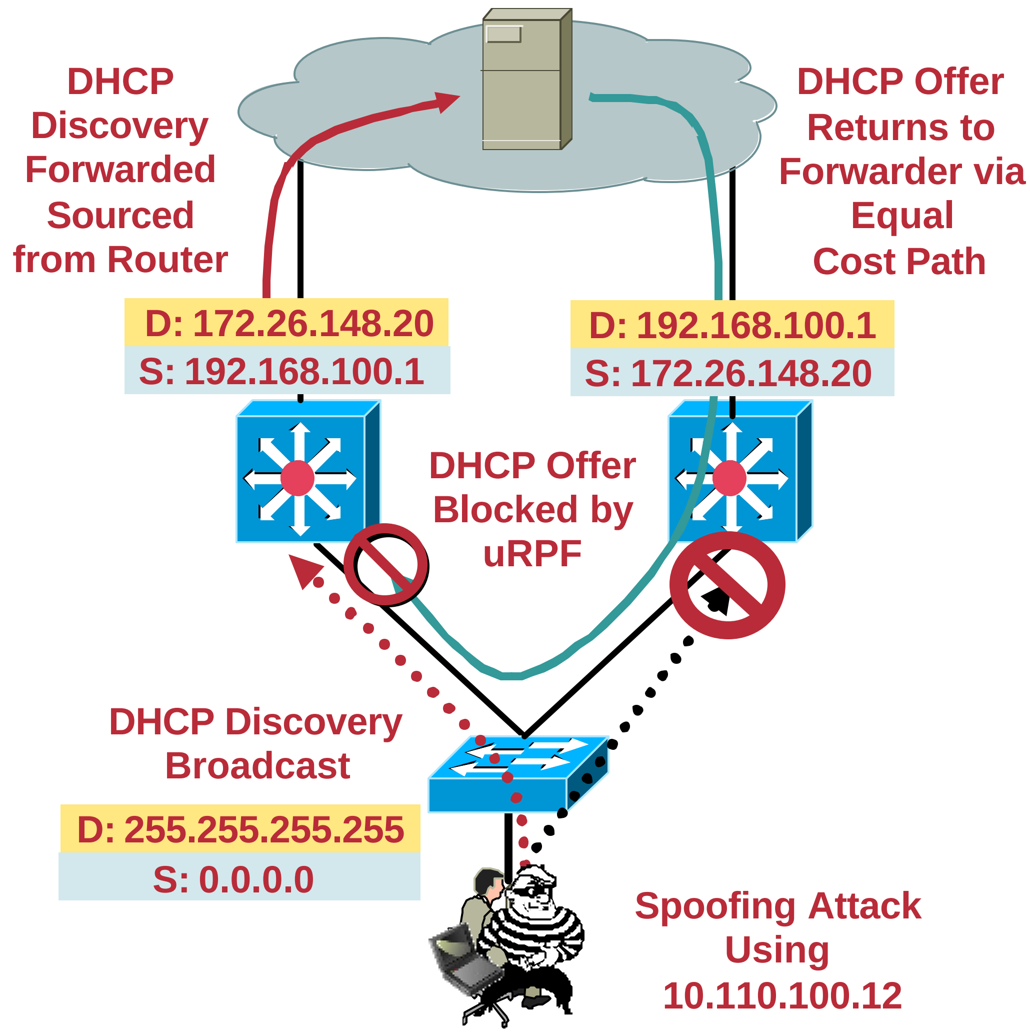


- **Plugging all of the Layer 2 security holes also serves to prevent a whole suite of other attack vectors**
- **Port security mitigates against most Layer 2 based CPU DoS attacks**
- **In addition to preventing M-i-M attacks IP source guard prevents**
  - **DDoS attacks which utilize a spoofed source address, e.g. TCP SYN Floods, Smurf**
  - **TCP splicing and RST attacks**

IP Source Guard vs. uRPF

Preventing Layer 3 Spoofing Attacks

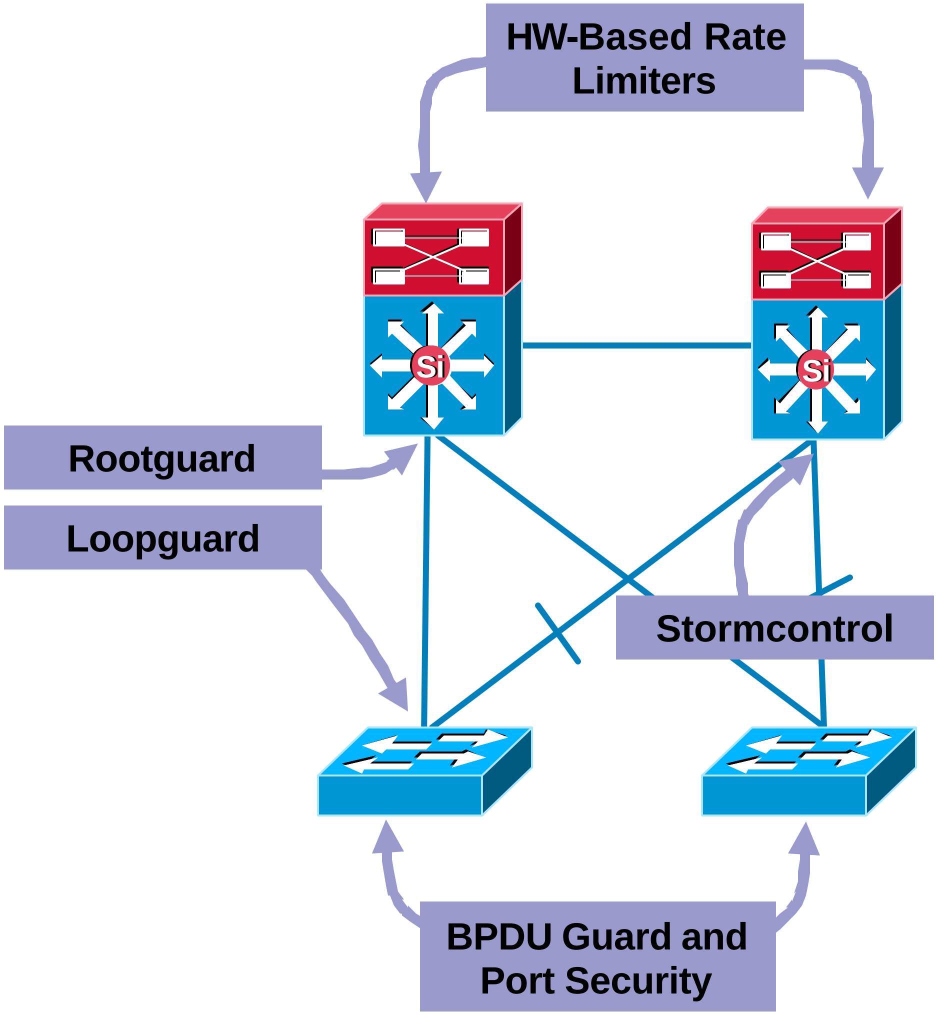
- **Problem:** Infected PC launches a DoS attack using spoofed source address
- Unicast Reverse Path Forwarding (uRPF) checks to see if incoming port is the best route to the source address
- uRPF operates in Strict or Loose mode
- Strict mode complex in a redundant environment
- Loose mode is very valuable for Black Hole Routing
- **IP Source Guard** is the best answer to this problem



Layer 2 Hardening

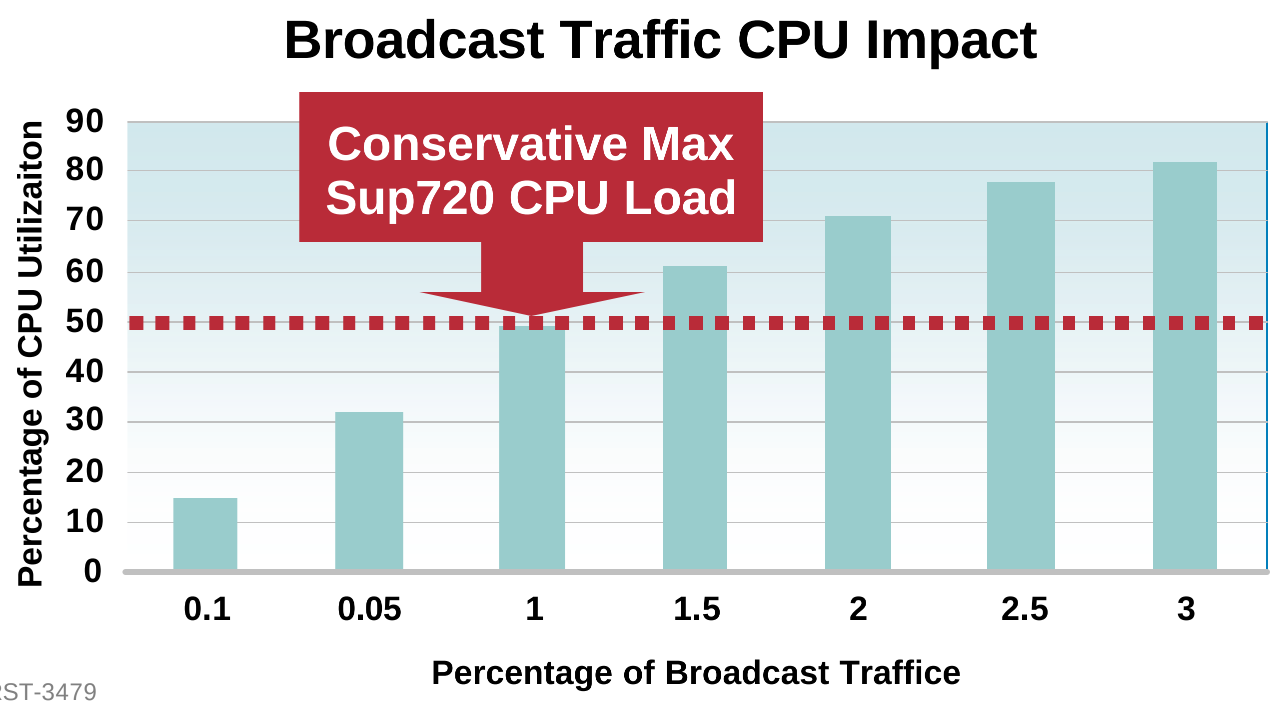
Spanning Tree Should Behave the Way You Expect

- **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

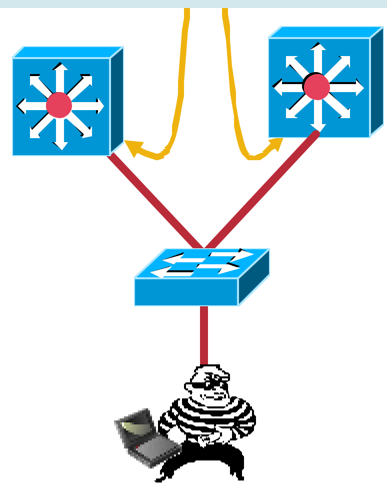


Harden the Network Links
Storm Control

- **Protect the network from intentional and unintentional flood attacks**
  e.g. STP loop
- **Limit the combined rate of broadcast and multicast traffic to normal peak loads**
- **Limit broadcast and when possible multicast to 1.0% of a GigE link to ensure distribution CPU stays in safe zone**

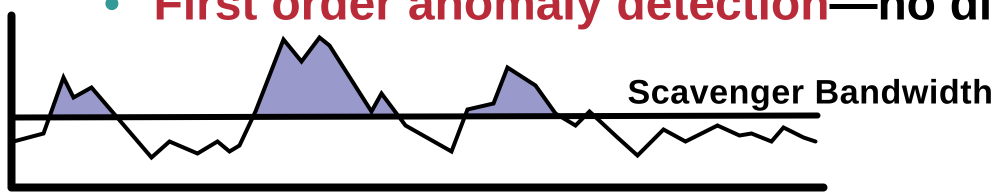



```
! Enable storm control

storm-control broadcast
level 1.0
storm-control multicast
level 1.0
```



Harden the Network Links—QoS
Scavenger-Class QoS

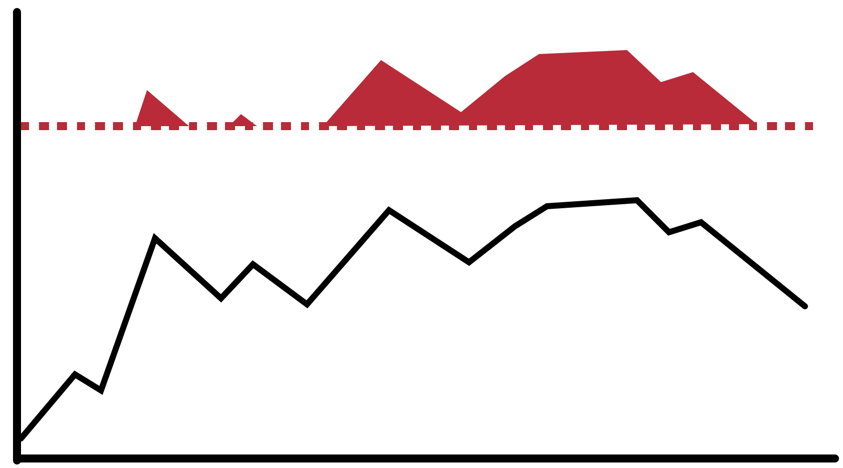
- All end systems generate traffic spikes
- Sustained traffic loads beyond 'normal' from each source device are considered suspect and marked as Scavenger
- First order anomaly detection—no direct action taken



Network Entry Points

- During 'abnormal' worm traffic conditions traffic marked as Scavenger is aggressively dropped —second order detection
- Priority queuing ensuring low latency and jitter for VoIP
- Stations not generating abnormal traffic volumes continue to receive network service

- During 'normal' traffic conditions network is operating within designed capacity

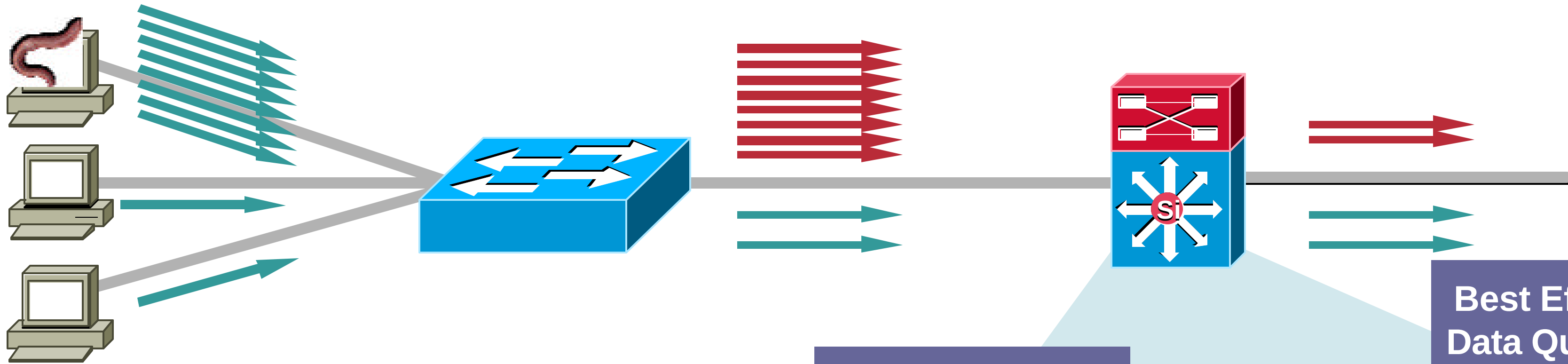


Aggregation Points

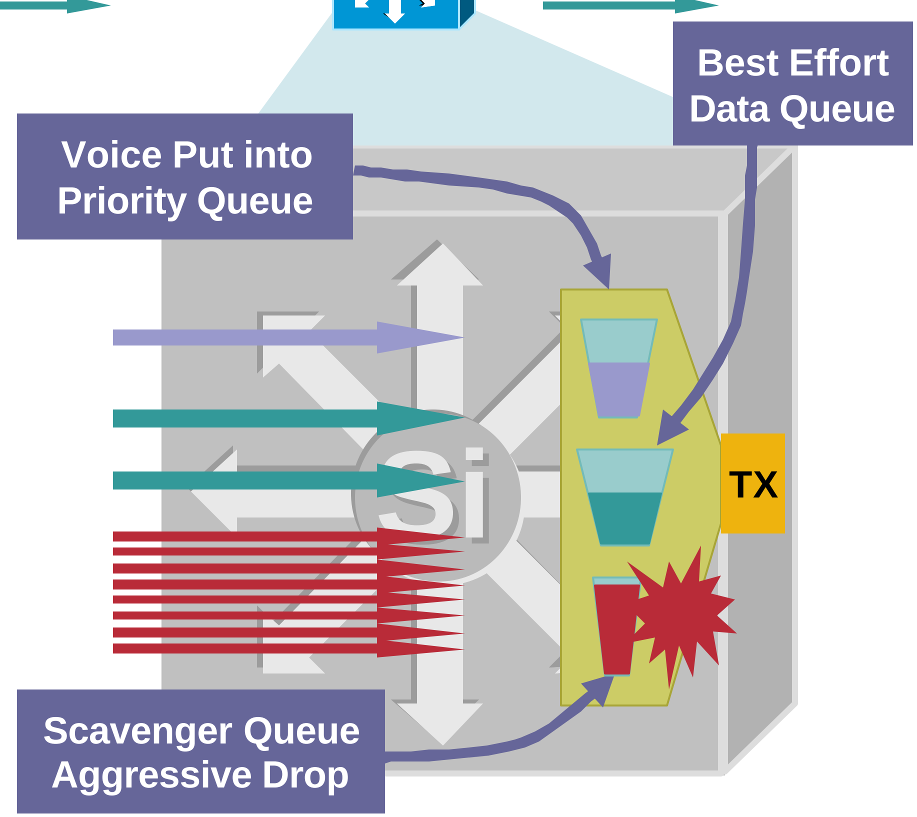
Mitigating the Impact: Scavenger-Class QoS

Scavenger Throttled Back

Cisco.com



- Scavenger traffic is assigned it’s own queue or queue/threshold
- Scavenger queue is configured for aggressive drop



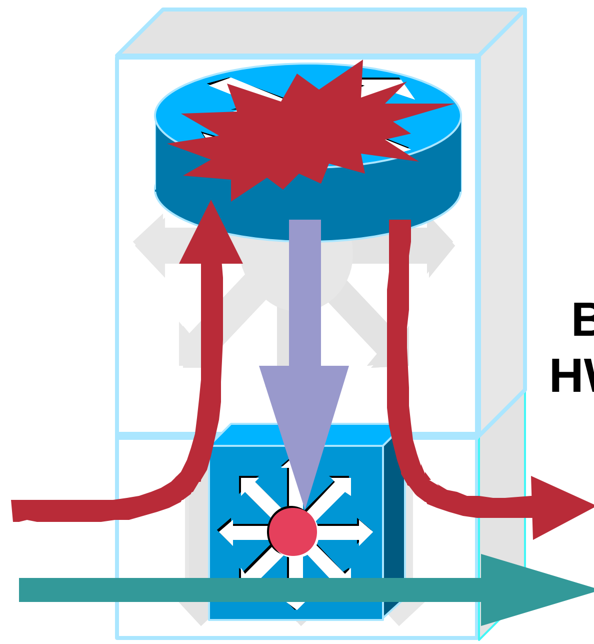
For More Information on Scavenger QoS Please See RST-2501: Campus QoS Design

Catalyst Cisco Express Forwarding

Before CEF…Flow-Based Switching

Cisco.com

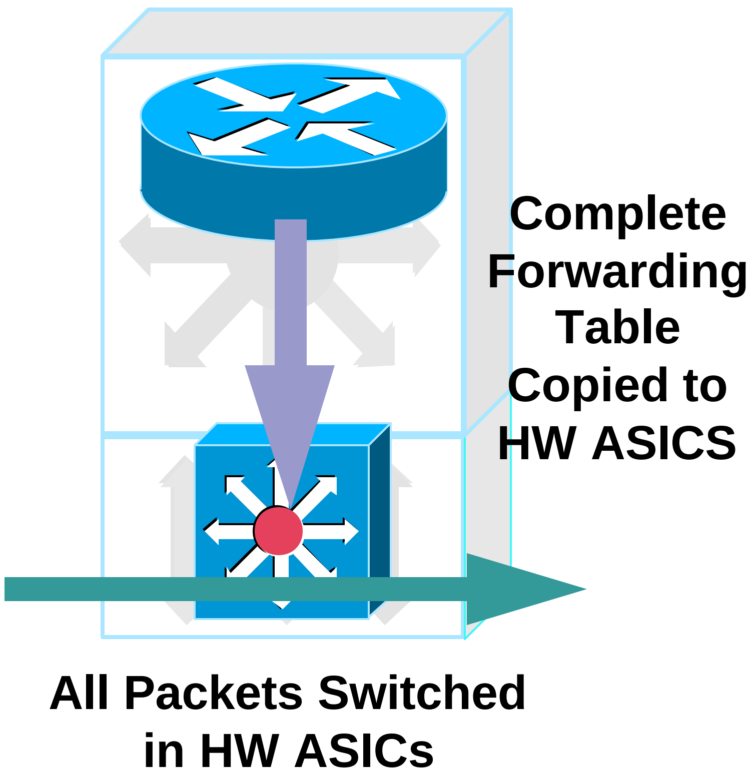
- Nimda, Slammer, Witty and similar worms send packets to a very large number of random addresses looking for vulnerable end systems to attack
- Flow/Prefix-based switching is limited by the ability of the CPU to setup initial flows
- Flow/Prefix-based HW caches may overflow when an abnormally high number of flows established
- Ability of CPU to process control plane traffic (EIGRP, OSPF, BPDU) suffers when flow rate is abnormally high



Catalyst Cisco Express Forwarding
CEF: Topology-Based Switching

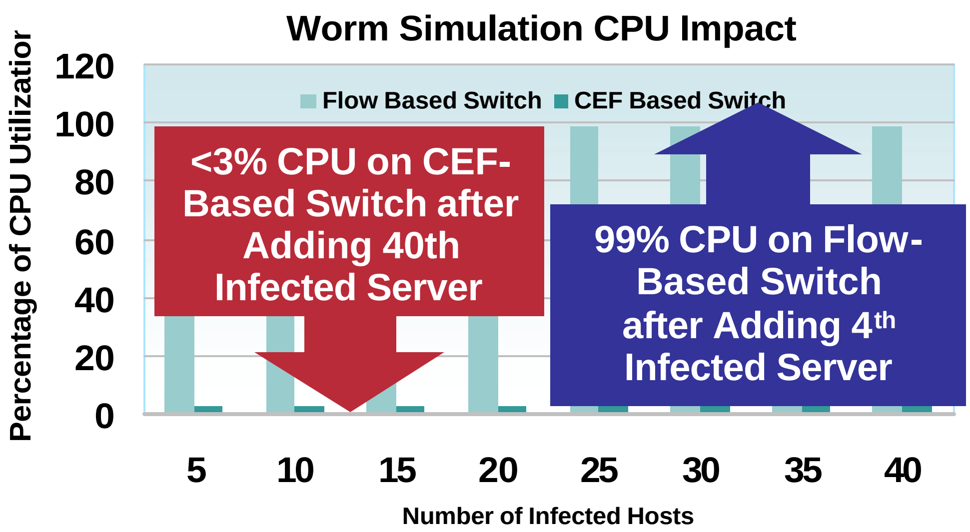
Cisco.com

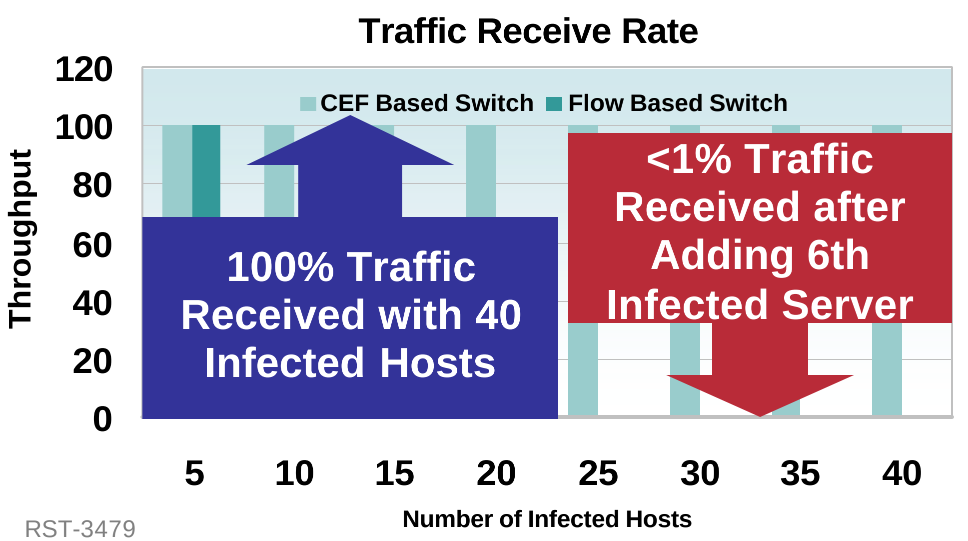
- **Route processor builds a Forwarding Information Base (FIB) calculated based on routing table entries, not traffic flows**
- **Hardware forwarding of first packet in each flow, whether there are one or one million of new flows**
- **Control plane unburdened by traffic forwarding—dedicated to protocol processing**
- **CEF protects campus switches from the abnormal worm traffic behavior**



Mitigating the Impact: CEF

Worm Propagation Impacts Stability



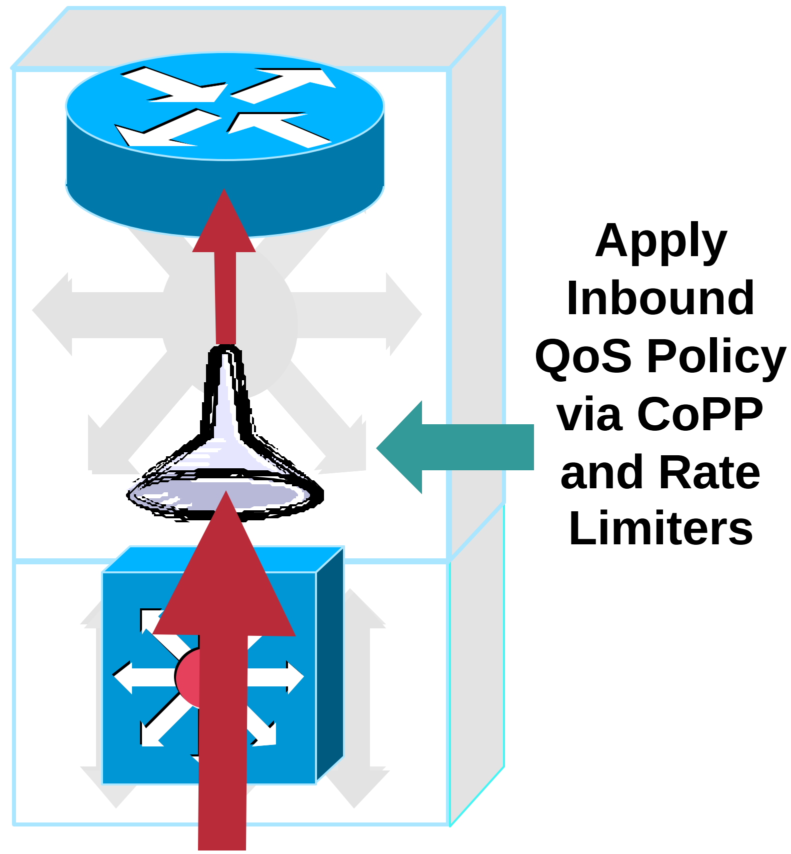


- **Aggressive scanning of network by worm will overload flow-based switching**
- **CPU resources consumed and unable to process BPDU and routing updates**
- **High CPU results in network instability**
- **No traffic loss with CEF**
- **Catalyst 6500 Sup720 and Sup2, Catalyst 4500 Sup IV and SupII+, Catalyst 3x50 all use HW CEF**

DoS Protection: Control Plane Protection

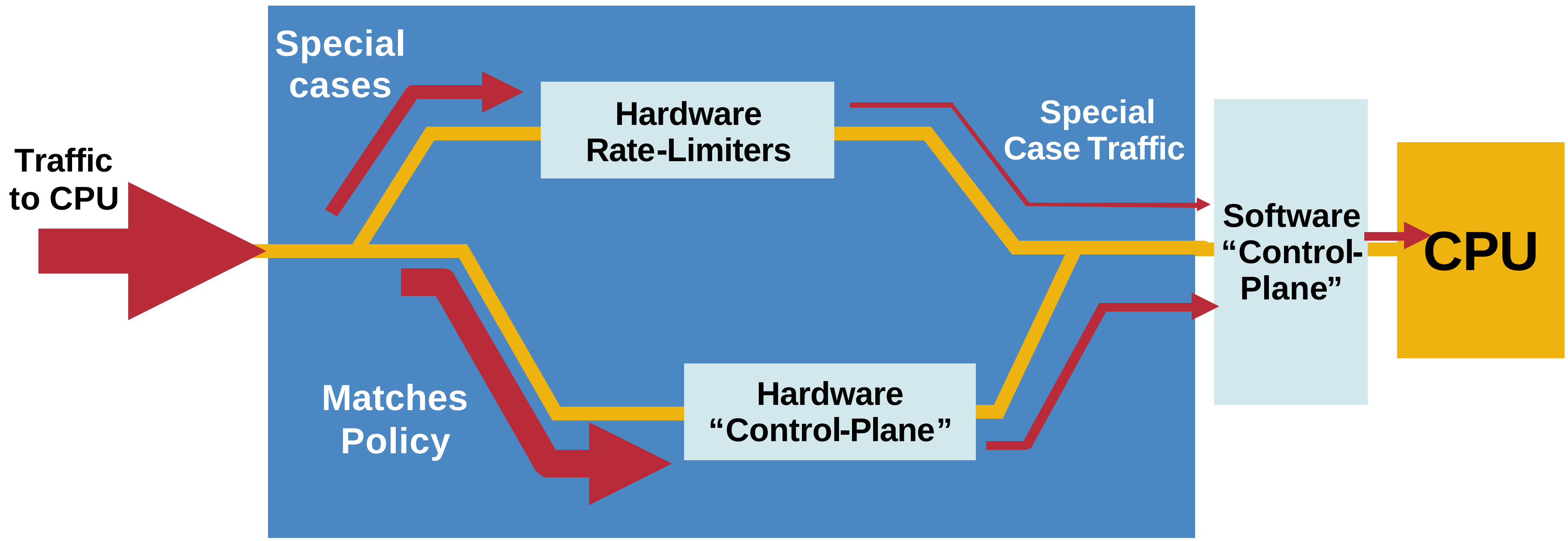
Catalyst 3750, 4500 and 6500

- CEF protects against system overload due to flow flooding
- System CPU still has to be able to process certain traffic
  - BPDUs, CDP, EIGRP, OSPF,…
  - Telnet, SSH, SNMP,…
  - ARP, ICMP, IGMP,…
- System needs to provide throttling on CPU-bound traffic
  - Hardware Rate Limiters and CPU queuing
  - Hardware and Software Control Plane Policing (CoPP)



DoS Protection: Control Plane Protection

Catalyst 6500 Rate Limiting and CoPP



- **Ten Hardware Rate Limiters in 6500 Sup720 (eight are configurable, two reserved)**
  - **Unicast Rate Limiters (CEF Receive, Glean, IP Options,…)**
  - **Multicast Rate Limiters (Multicast FIB Miss, Partial Shortcut,…)**
  - **Layer 2 Rate Limiters (PDU, L2PT)**
  - **General Rate Limiters (MTU Failure, TTL <= 1)**
- **Traffic that matches a configured Rate Limiter bypasses HW CoPP**

DoS Protection: Control Plane Protection
Rate Limiting and CoPP Configuration

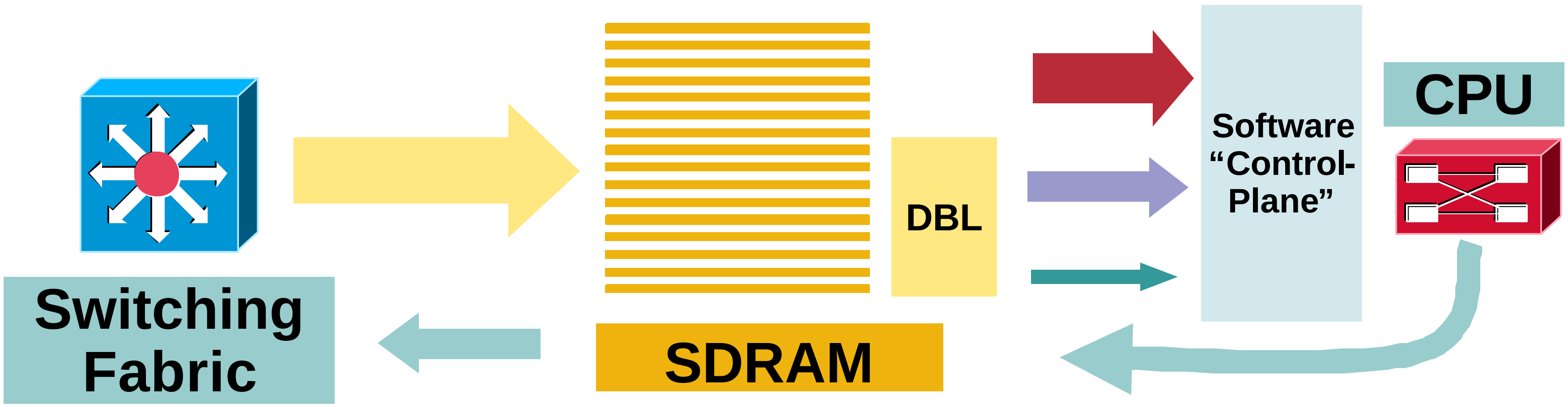
- **Must enable QoS globally, otherwise, CoPP is performed in software only**
- **Define ACLs to match traffic**
  - Permit means traffic will belong to class, deny means fall through
- **Define class-maps**
  - Use "match" statements to identify traffic associated with the class
- **Define policy-map and associate classes and actions to it**
  - Policing is only supported action
- **Tie the policy-map to the control-plane interface**

```
! Partial Sample Config

mls rate-limit multicast ipv4 partial 1000 100
mls rate-limit all ttl-failure 1000 10

mls qos

ip access-list extended CPP-MANAGEMENT
 remark Remote management
 permit tcp any any eq SSH
 permit tcp any eq 23 any
 permit tcp any any eq 23

class-map match-all CPP-MANAGEMENT
  description Important traffic, eg management
  match access-group name important

policy-map copp
 description Control plane policing policy
 class  CPP-MANAGEMENT
  police 500000 12800 12800 conform-action
   transmit exceed-action drop

control-plane
 service-policy input copp
```

Deployment Guide White Paper:
www.cisco.com/en/US/products/sw/iosswrel/ps1838/products_white_paper09186a0080211f39.shtml

DoS Protection: Control Plane Protection

Catalyst 4500 CPU Queue Scheduling and CoPP



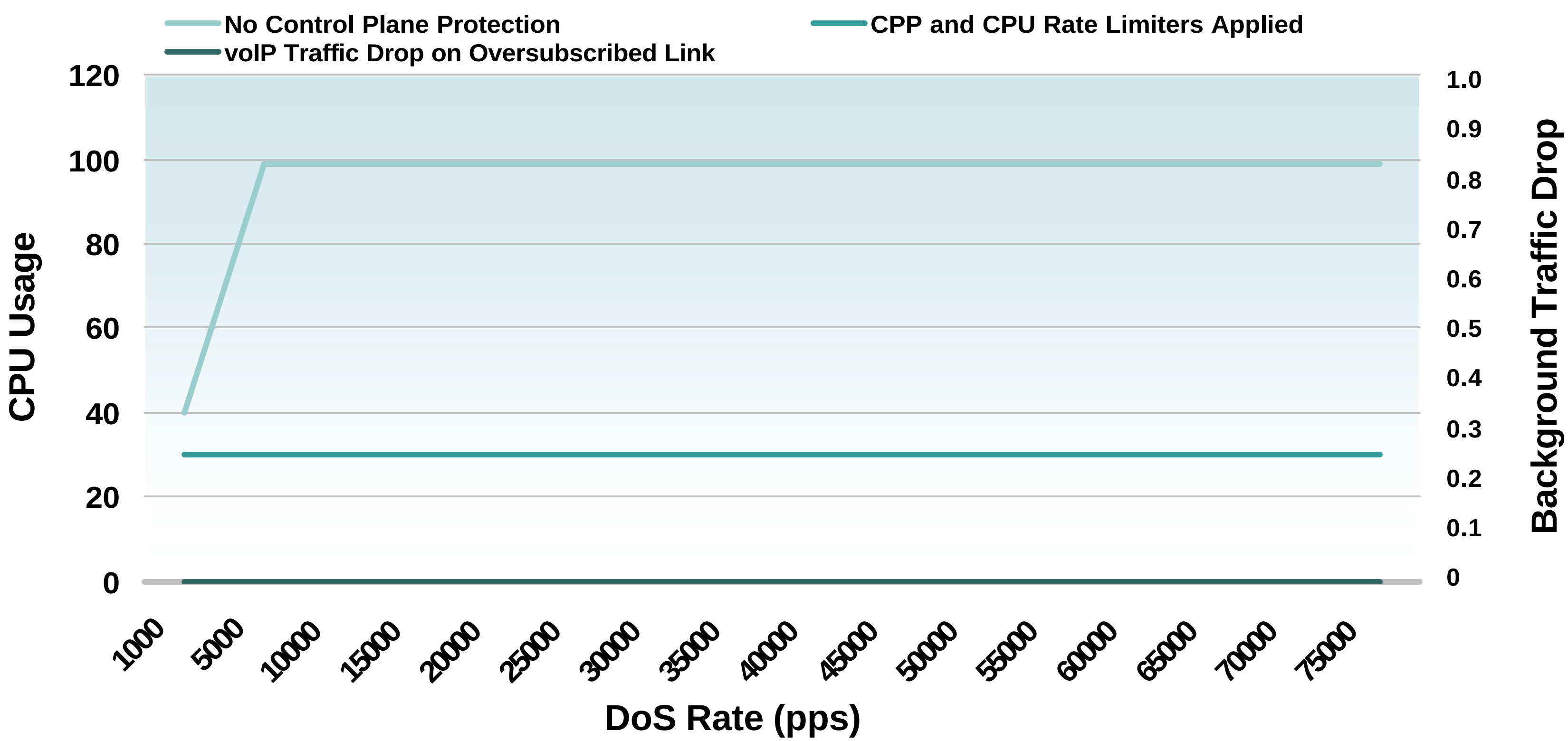
- **16 distinct inbound queues from switching fabric serviced by CPU using a weighted RR scheduling prioritizing control plane packets (e.g., BPDUs)**
- **Dynamic Buffer Limiting (DBL) also performed on CPU queues**

```
4507-SupIV#show platform cpu packet driver
 Queue                  rxTail   received  all guar allJ gurJ   rxDrops  rxDelays
 0 Esmp                62B26C0  522275197   99  100    0    5         0         0
 1 Control             62B2BA0   22814109  595  600    0    5         0         0
…
15 MTU Failure         62B848C          0  102  102    0    5         0         0
```

Mitigating the Impact: CoPP

CoPP and Rate Limiters Compliment CEF

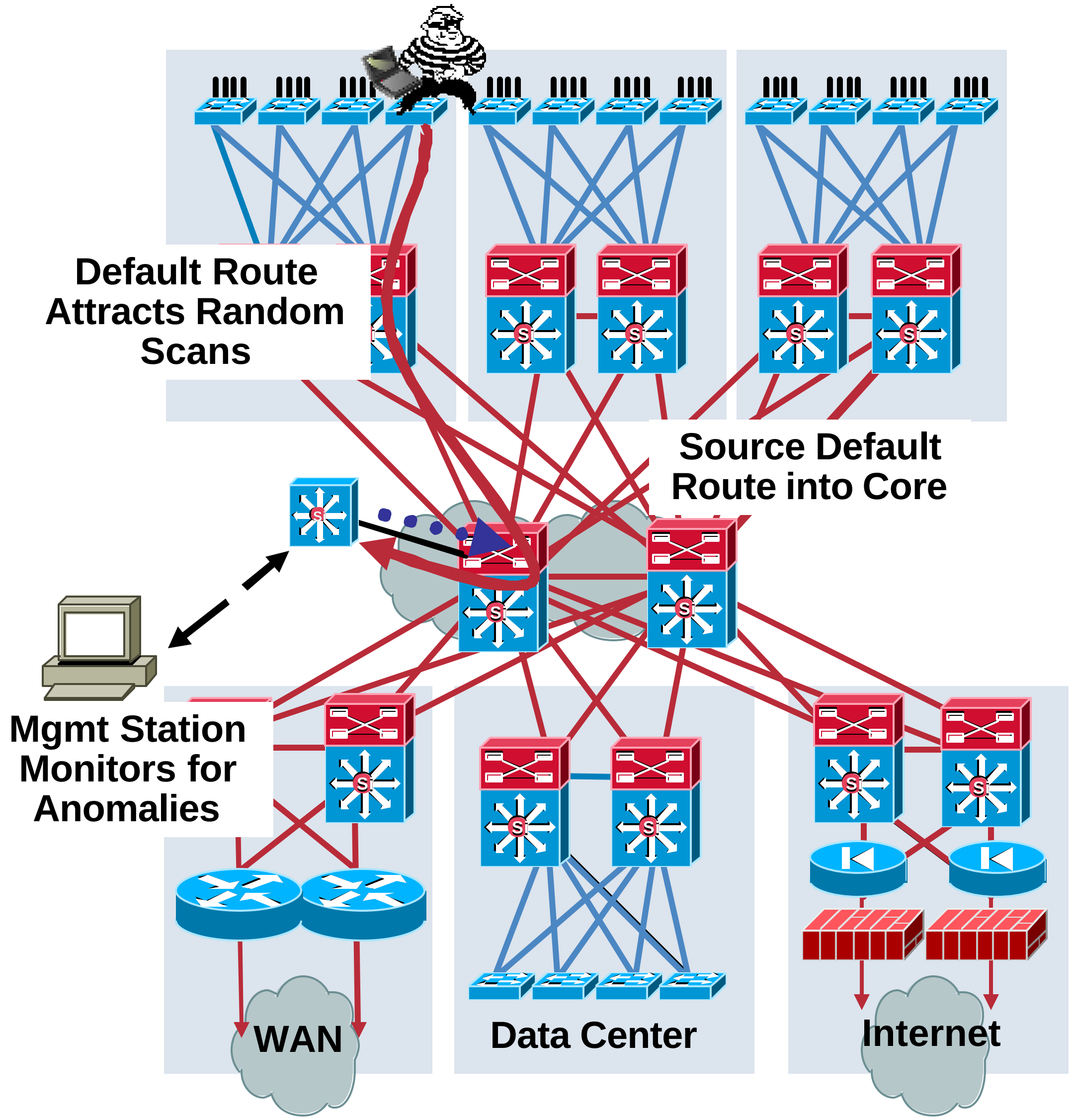
- Multiple concurrent attacks (Multicast TTL=1, Multicast Partial Shortcuts, Unicast IP Options, Unicast Fragments to Receive adjacency, Unicast TCP SYN Flood to Receive Adjacency)
- CPU kept within acceptable bounds with no loss of mission critical traffic



Finding the Worm: Sink Hole Routers
Monitoring for Network Worms

Cisco.com

- **Sink hole router sources a default route (0.0.0.0) into core of network**
- **All traffic with a destination address not in the Enterprise network is sent to the sink hole**
- **Monitor inbound traffic to the sinkhole via ACLs, ip accounting or Netflow**
- **Net mgmt scripts look for common sources sending to random addresses**
- **Does not work when default routing to Internet**

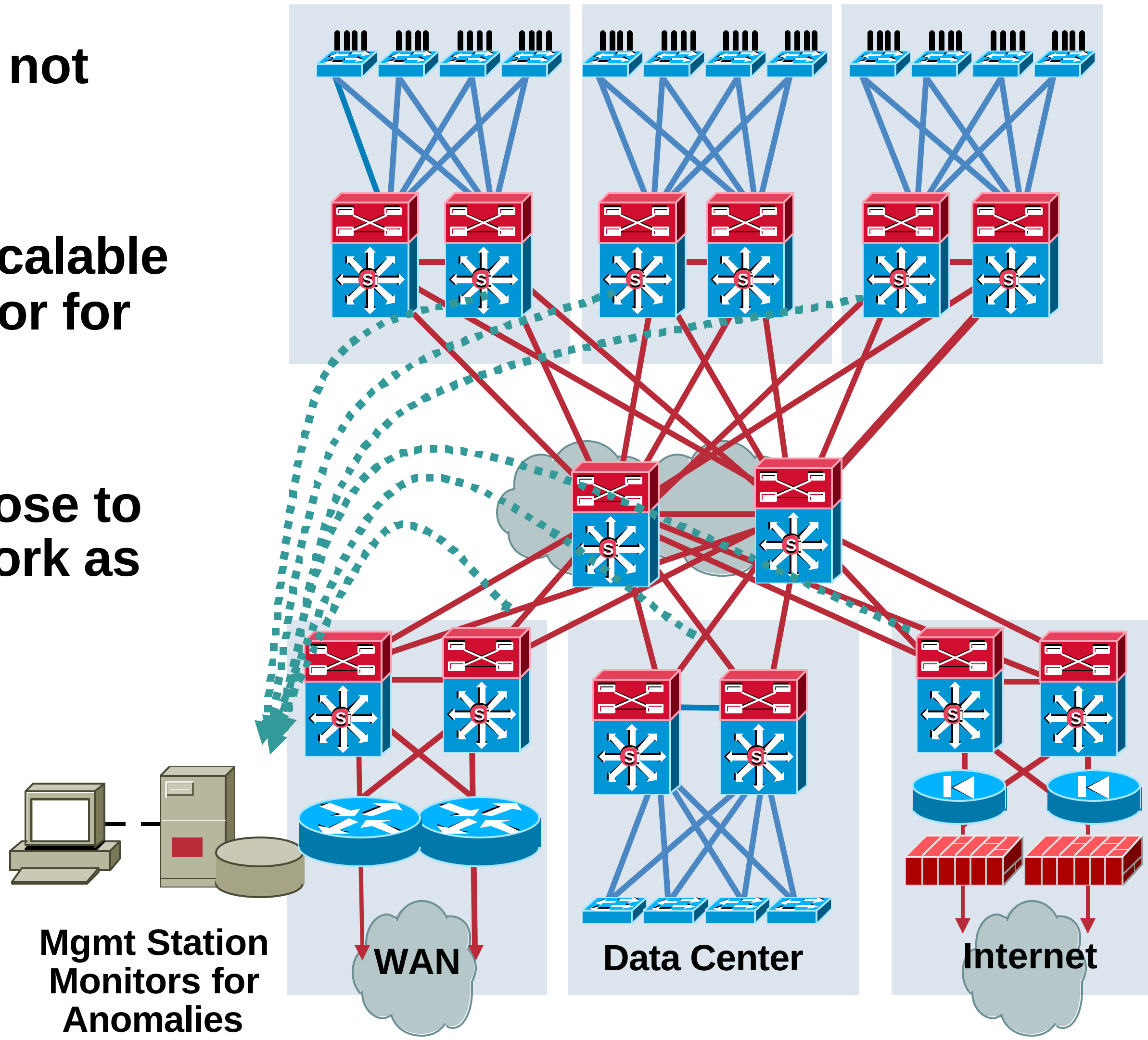


Finding the Worm: NetFlow

Scalable Monitoring for Network Worms

Cisco.com

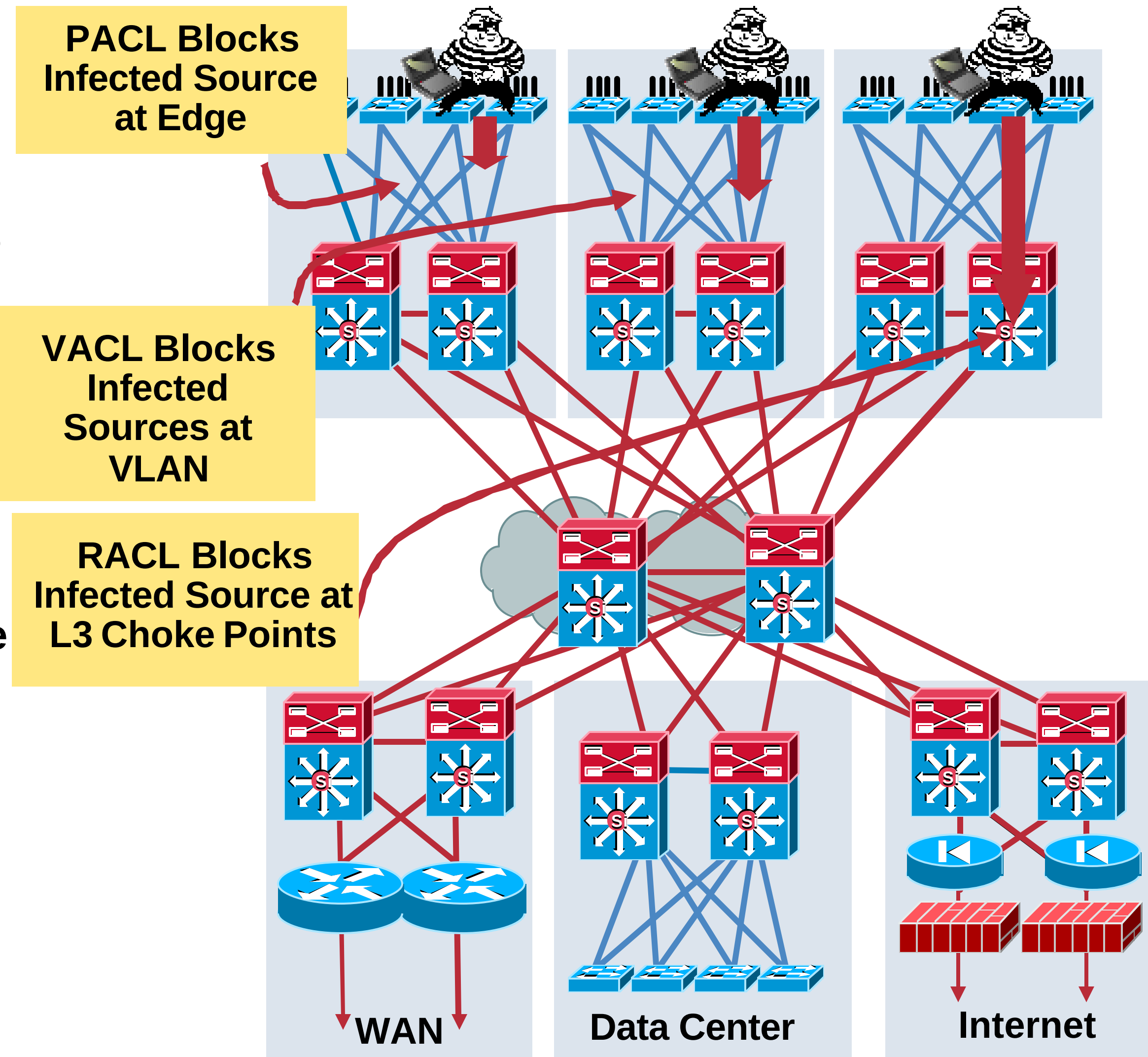
- **Sink hole routers do not detect intelligent scanning worms**
- **Netflow provides a scalable mechanism to monitor for worms throughout the network**
- **Enable Netflow as close to the edge of the network as possible in order to maximize detection accuracy**
  - **Distribution switches**
  - **WAN aggregation**
  - **Internet DMZ**



Worm Containment: Reactive

PACLs, RACLs, VACLs, and CAR

- Hardware access control lists can be utilized at multiple tiers in the network
- Port ACLs allow L3/L4 ACLs to be applied to a L2 port
- Utilize a PACL to block specific worm traffic at network edge
- Utilize VLAN ACLs to block specific worm traffic over a VLAN
- Utilize router ACLs at L3 choke points to block specific worm traffic
- CAR can be used to throttle traffic to destinations under attack (DDoS) on Cisco IOS-based routers

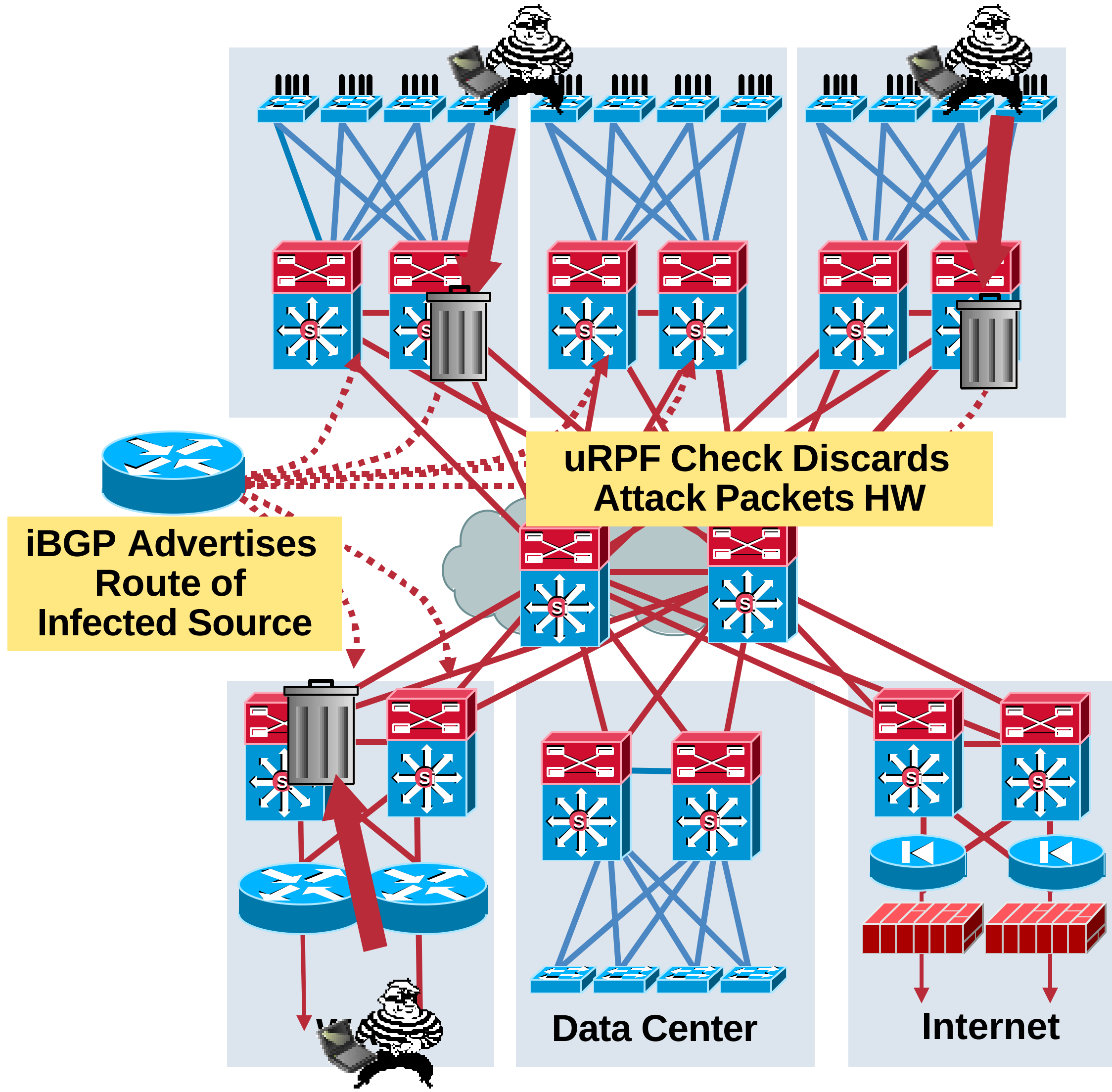


Worm Containment: Block Infected Sources

Triggered uRPF Blackholes

Cisco.com

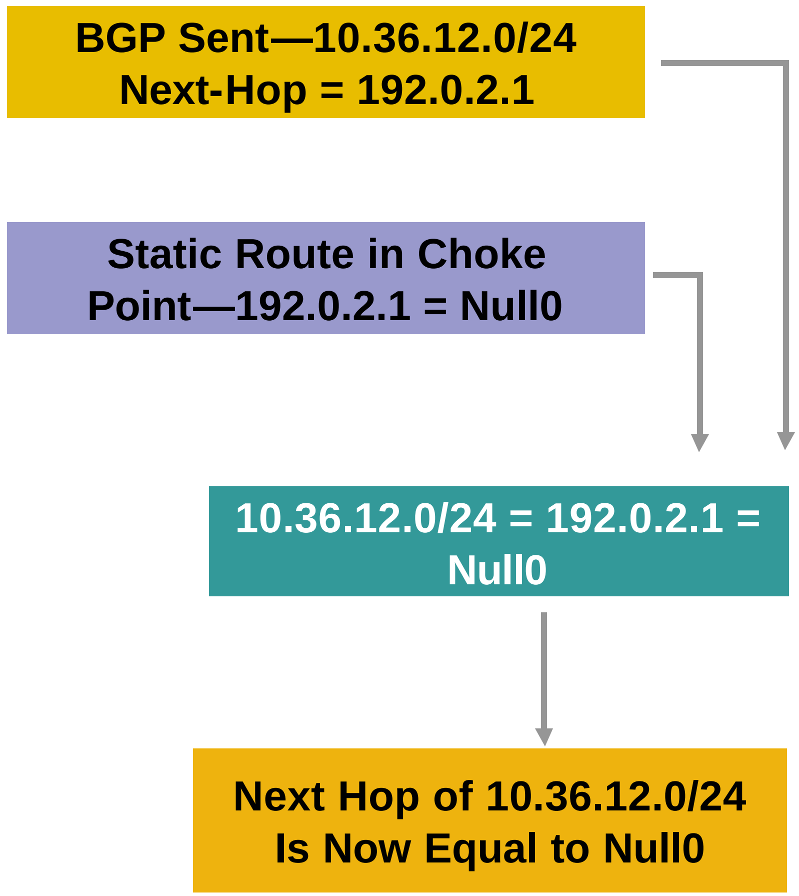
- Need a scalable method to rapidly block traffic **from** infected sources (Worm)
- Need a scalable method to rapidly block traffic **to** destinations under attack (DDoS)
- Triggered uRPF Blackhole Routers will do both
- Using iBGP push message to choke points to discard the attack packets
- Does NOT require to use BGP as your routing protocol
- Recommended **Sup720B** to support; require Sup720



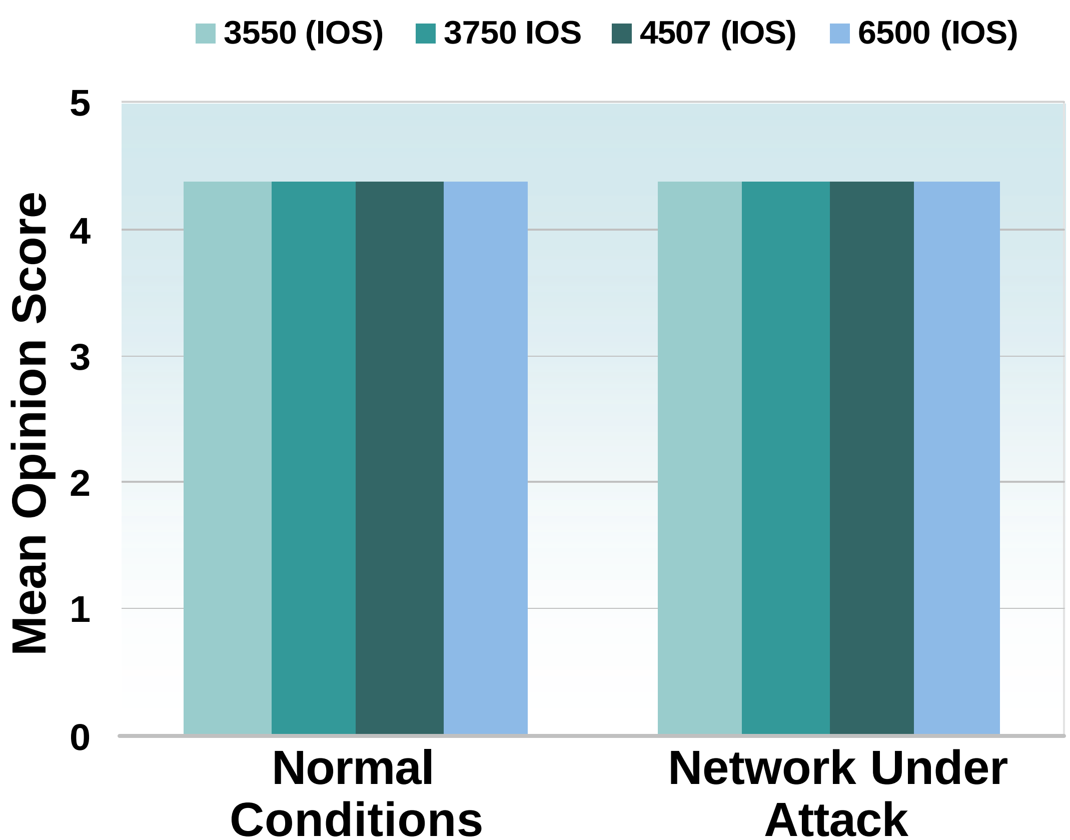
Blackholing Infected Sources
Unicast RPF Loose Check

Cisco.com

- Loose uRPF checks if route is found in the Forwarding Information Base (FIB)
    - If not in FIB, drop the packet
    - If equal to Null0, drop the packet
- Using iBGP insert a route for infected sources that point to Null0
- Activate loose uRPF on downstream stream switch ports
- Choke point switch drops packets with infected source addresses



Does It Work?
Voice Survives the Worm

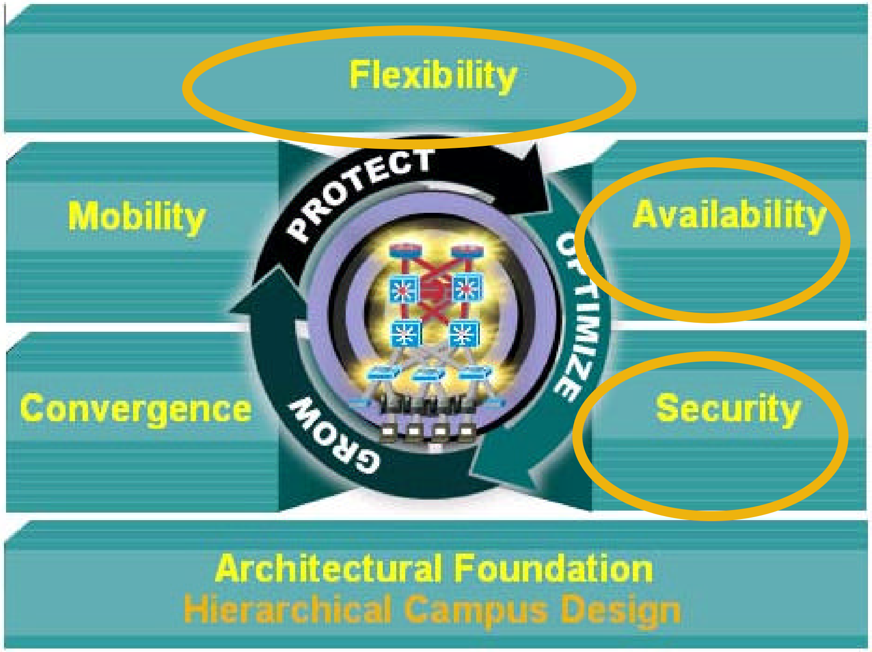


- 90 P4 GigE servers
- Simultaneous attacks
  - Simulated Slammer
  - Macof—L2 DoS
  - Smurf—L3 DoS
- 6500 with Sup720a in the distribution
- Network remained stable
- Mean opinion score for G.711 voice flows unchanged from normal conditions

Agenda

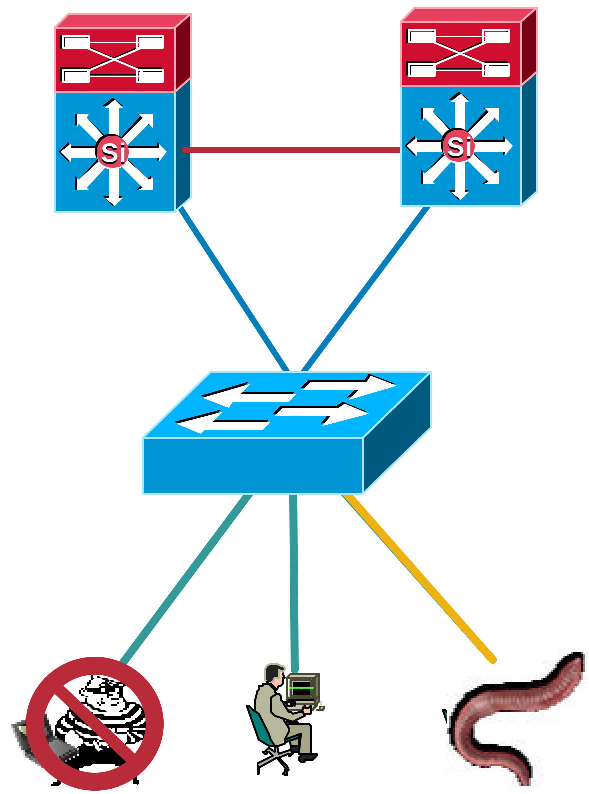
Cisco.com

- **Foundational Design Review**
- **Convergence—IP Communications**
- **Wireless LAN and Wireless Mobility**
- **High Availability**
  - **Alternatives to STP**
  - **Device HA (NSF/SSO and Stackwise)**
  - **Resilient Network Design**
- **Segmentation and Virtualization**
  - **Access Control (IBNS and NAC)**
  - **Segmentation**
- **Questions and Answers**



IBNS (802.1x) and NAC
Access and Policy Control

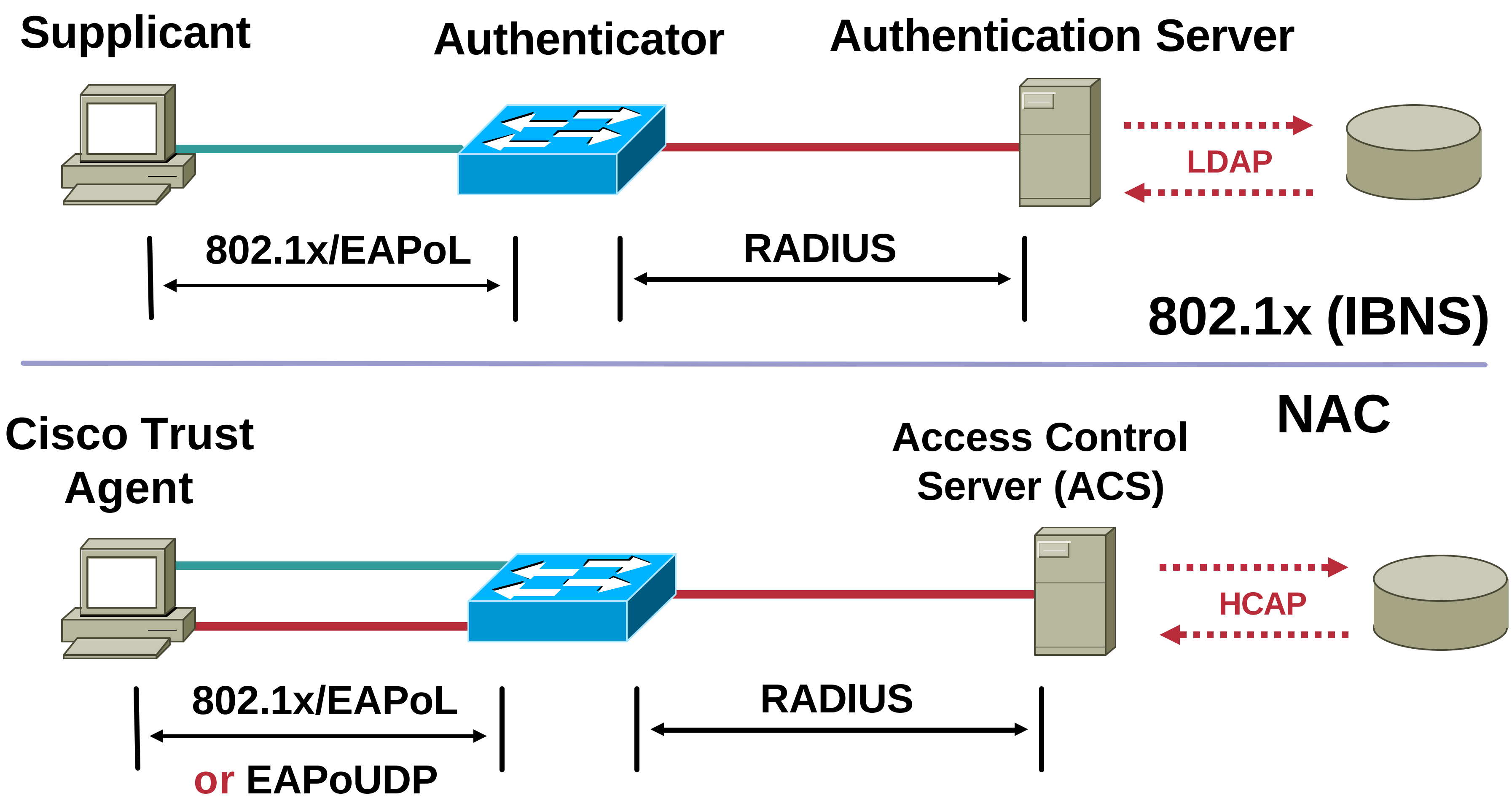
- **Identity-Based Networking Services (IBNS)**
  - **Identifies and authenticates the user or device on the network and ensures access to correct network resources**
- **Network Access Control (NAC)**
  - **Performs posture validation to ensure that machines not compliant with software posture, and therefore vulnerable to infection, can be isolated to a segment of the network where remediation can take place**
- **802.1x provides port-based access control and operates at L2**
- **NAC provides posture assessment and device containment at L3 or L2**
- **Complimentary functions**



Edge Access Control

802.1x and NAC Operation

EAP, EAPoL, RADIUS, and HCAP

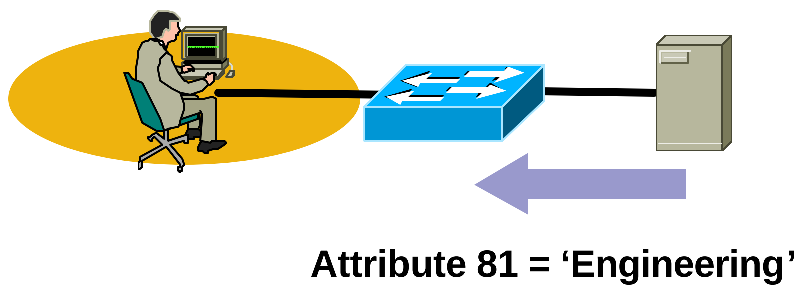


For More Discussion on IBNS and NAC Please See—SEC2005

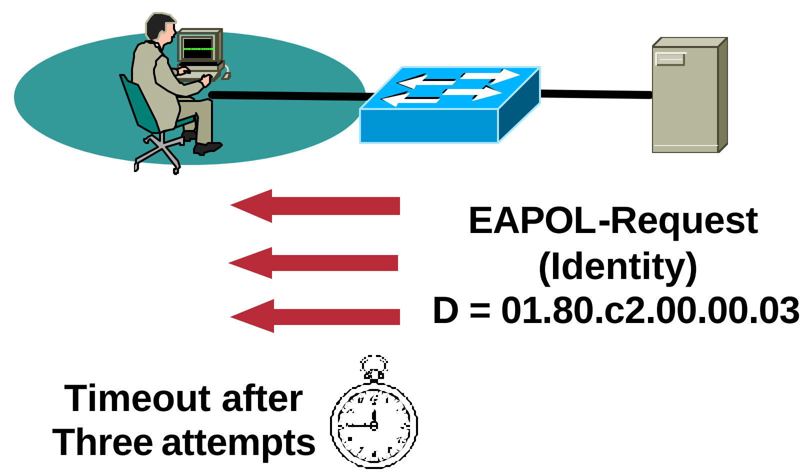
802.1x Use of VLANs
VLAN Assignment and the Guest VLAN

- 802.1x defines an access method for LAN switch ports
- You can permit or deny access based on authorization behavior
- Using RADIUS AV-Pairs we can supply additional policy options for the switch port
- VLAN assignment utilizes AV-Pairs
  - [64] Tunnel-Type—"VLAN" (13)
  - [65] Tunnel-Medium-Type—"802" (6)
  - [81] Tunnel-Private-Group-ID—<VLAN name>
- In the absence of an EAPoL response from the client the switch can assign the port to a locally configured 'Guest' or default VLAN

VLAN Assignment



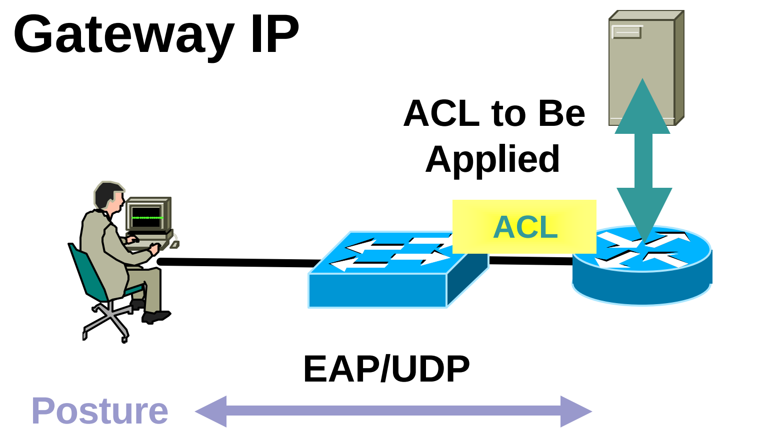
Guest VLAN

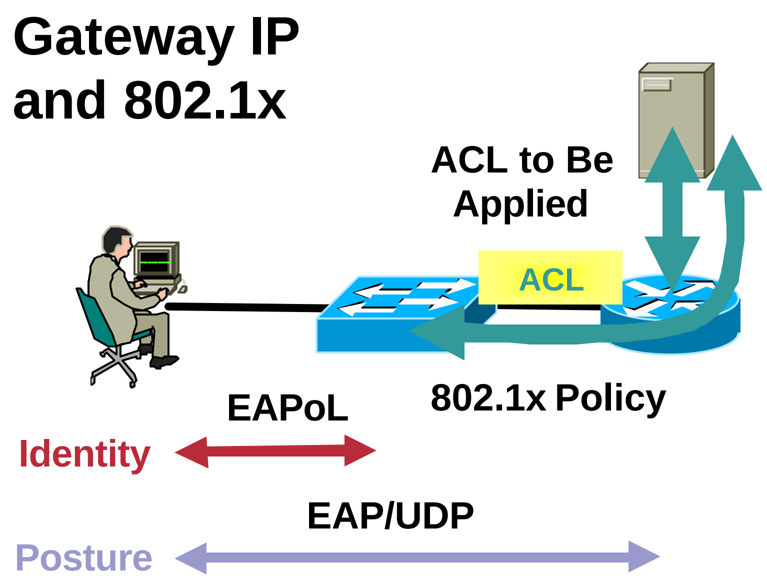


802.1x and NAC
Gateway IP (NAC Version1)

Cisco.com

- NAC posture assessment at the first Layer 3 hop (default GW)
- Cisco IOS 'Intercept ACL' will intercept interesting traffic generated by end station and initiate an EAP/UDP session
- Based on response from RADIUS/Policy server will apply an ACL to permit or control access
- If used 802.1x authentication occurs prior and independently of NAC posture assessment
- Gateway IP is not currently supported on any Catalyst Switch (Cisco IOS Router only)
- Not currently applicable to the Campus today

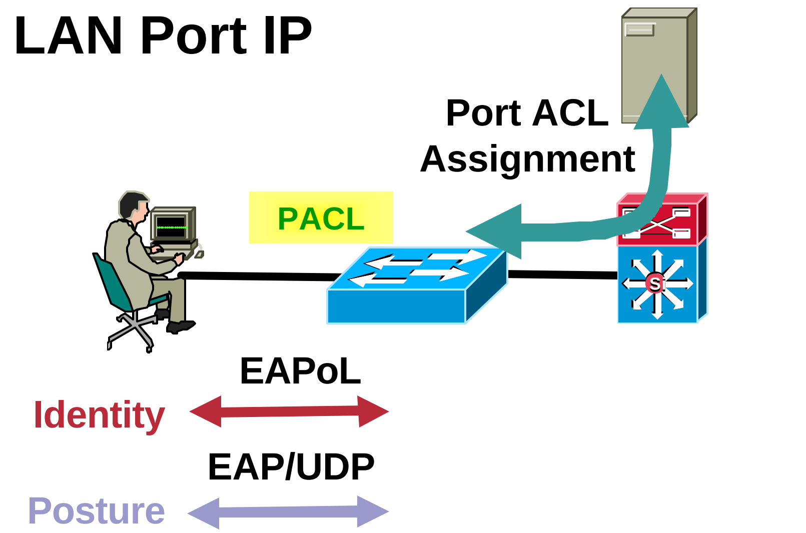


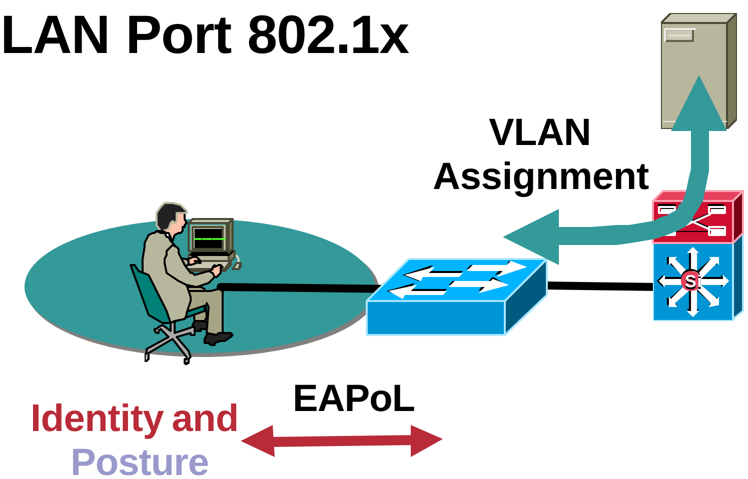


802.1x and NAC in the Campus

LAN Port dot1x and LAN Port IP

- **NAC posture assessment at the first Layer 2 hop (switch port)**
- **LAN Port IP triggers an EAP posture session when first ARP is received on a port or DHCP snooping is triggered**
  - Applies posture policy via a Port ACL
  - Assumes innocent until proven guilty
- **LAN Port 802.1x supplies both identity credentials along with posture data during dot1x login**
  - Applies posture policy via VLAN assignment—Remediation VLAN
  - Assumes guilty until proven innocent





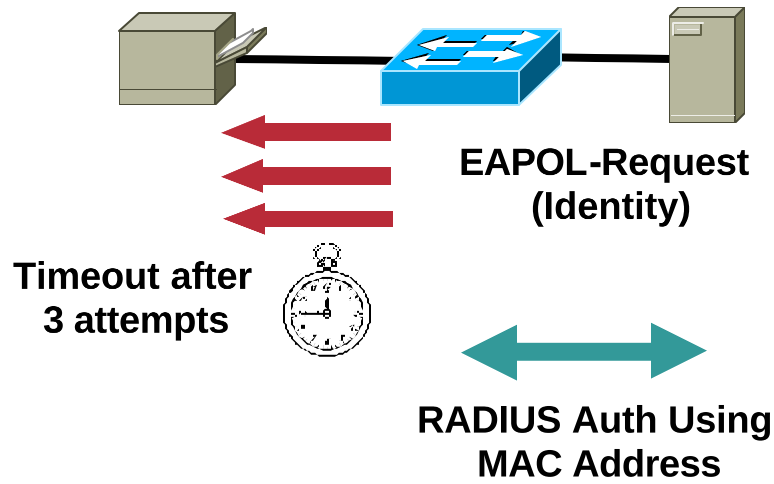
Campus Design Considerations for 802.1x

MAC-Auth and Failed Authentication VLAN

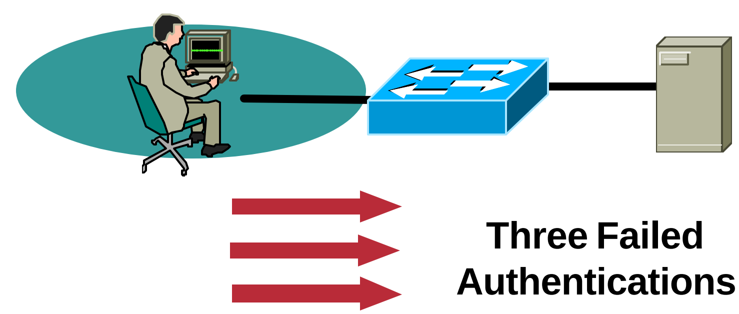
Cisco.com

- **MAC-Auth**
  - Provides a supplementary authentication based on MAC address
  - After timeout of EAPoL MAC address is proxied to ACS to provide credentials for authentication
  - Requires 6500 CatOS 8.5(1)
- **Authentication Fail VLAN**
  - Allows end devices without valid credentials to be assigned to a 'Guest' VLAN
  - Assigns devices to Auth-Failed VLAN after three consecutive login failures
  - Requires 6500 CatOS 8.4(1)

MAC-Auth

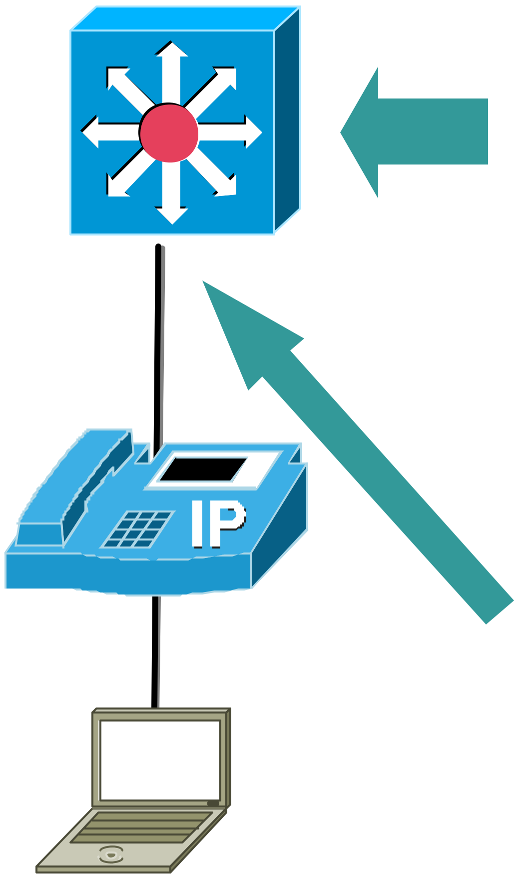


Auth-Failed VLAN



Sample NAC/IBNS switch config

802.1x, NAC, Guest, MAC-Auth and Auth Fail



CatOS Global Configuration

```
set dot1x system-auth-control enable
set radius server 10.1.125.1
set radius key cisco123
```

IOS Global Configuration

```
radius-server host 10.1.125.1
radius-server key cisco123
aaa new-model
aaa authentication dot1x default group radius
aaa authorization default group radius
dot1x system-auth-control
```

CatOS Port Configuration

```
set port dot1x 3/1-48 port-control auto
set port dot1x 3/1-48 guest-vlan 250
set port dot1x 3/1-48 auth-failed-vlan 251
set port mac-auth-bypass 3/1-48 enable
```

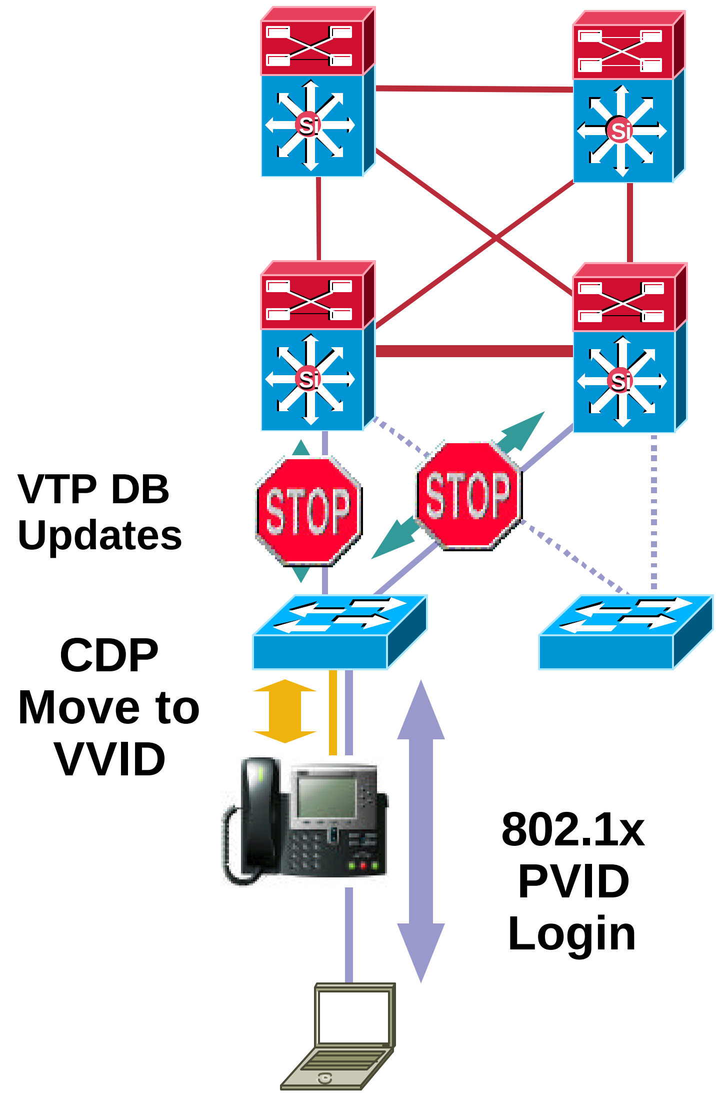
IOS Per-Port configuration

```
int range fa3/1 - 48
dot1x port-control auto
dot1x guest-vlan 250
```

Campus Design Considerations for 802.1x

VTP, CDP and 802.1x Interaction

- VLAN assignment utilizes a string field in the RADIUS attributes to select the VLAN
- This VLAN name should map to a unique VLAN on each access switch
- VTP database will be different on all switches
- Switches need to use either VTP transparent or off
- Switch requires CDP detection of phone to allow phone to connect without 802.1x
- Once identified phone moved to VVID and PC completes 802.1x on the PVID

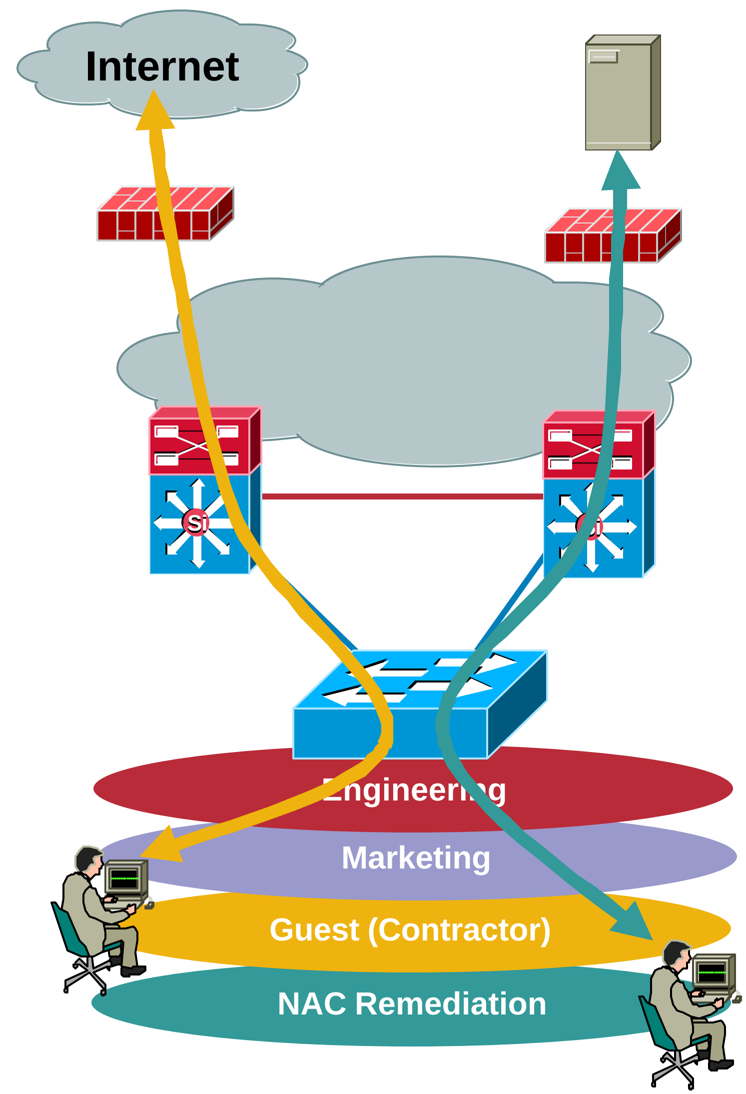


Campus Design for 802.1x and NAC

What Do I Do with Them Once I Have Them in a VLAN?

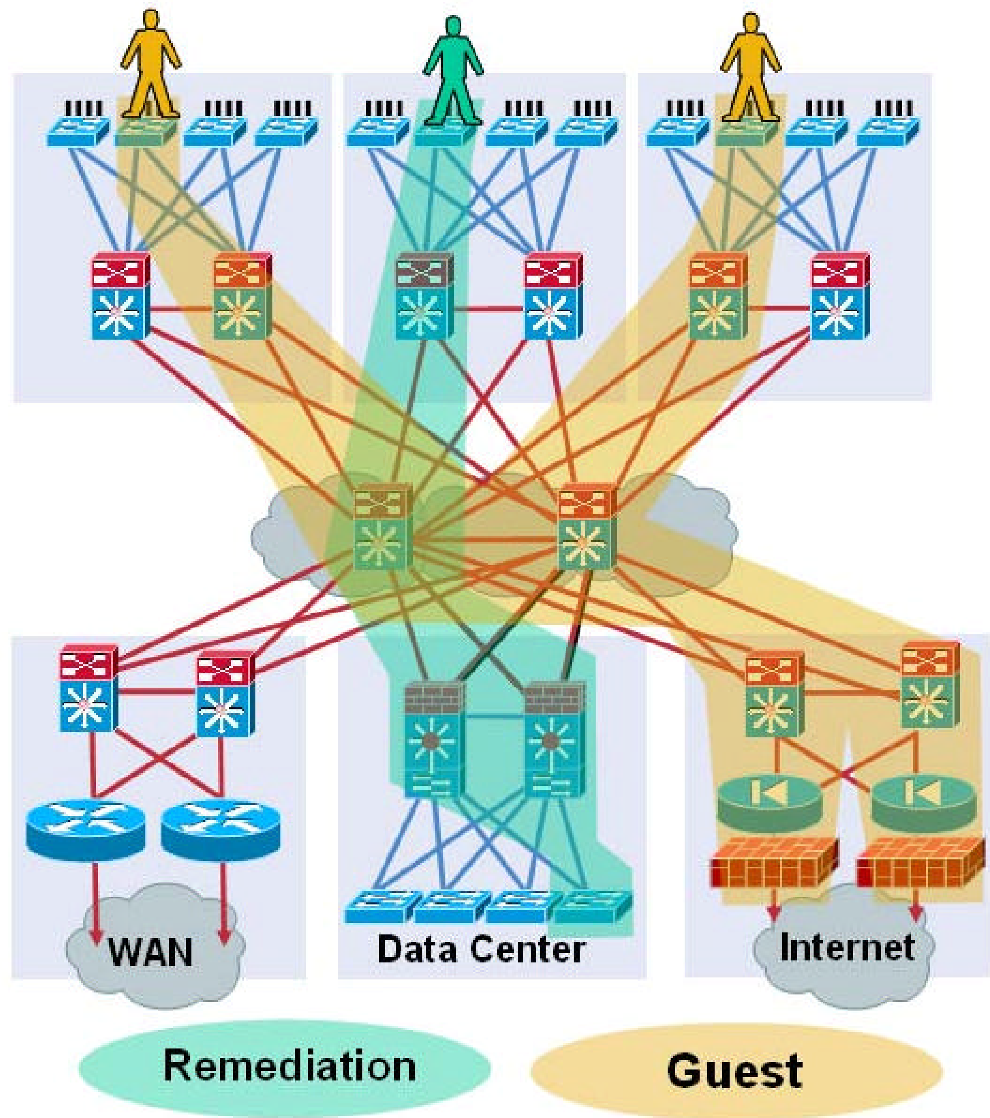
Cisco.com

- **802.1x and NAC LAN Port 802.1x Basic both control network access based on VLAN assignment**
- **Once they are assigned to a specific VLAN the network infrastructure needs to keep the traffic isolated**
- **Potential Solutions**
    - **ACLs**
    - **PBR with GRE**
    - **VRF with GRE**
    - **VRF-Lite**
    - **MPLS**
- **All provide some form of Network Compartmentalization**



Segmentation and Virtualization
Closed User Groups with Centralized Policy

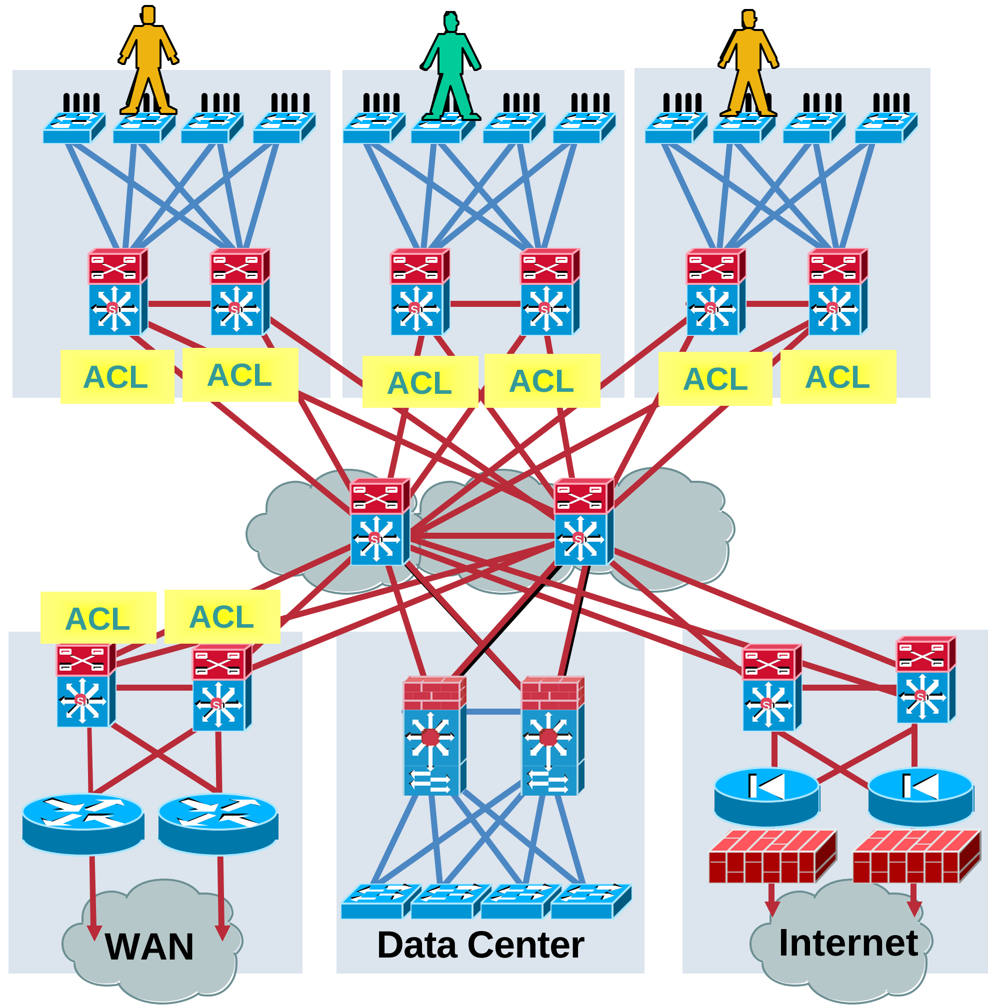
- Guest and Remediation one example of a larger problem
- **Closed User Group** creation
  - Provides secure and independent communication over a shared infrastructure
  - Enable User Mobility
- **Centralization** of policies and services
  - Policies based on groups
  - Enhanced Manageability
- **Sharing** of network intelligence/services
  - Costly resources centrally serve all groups while maintaining privacy



Campus Segmentation
Policy Control—ACLs

Cisco.com

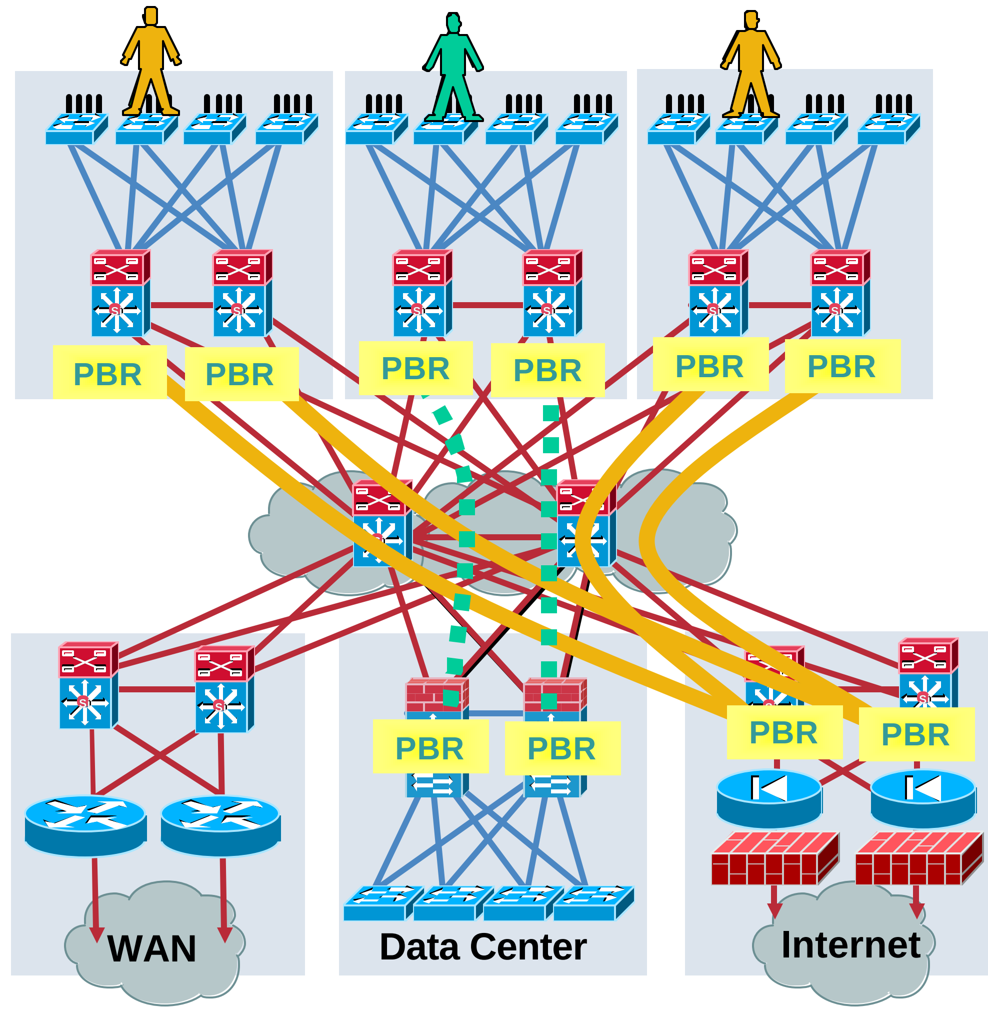
- **Restricting Guest and Remediation traffic via ACLs**
- **Pros:**
  - **HW-Based Forwarding**
  - **Simple Initial Deployment**
- **Cons:**
  - **Distributed static configuration**
  - **ACLs provide for restriction of traffic but not for control of the forwarding path of the traffic**
  - **Restricts user mobility**



Segmentation and Virtualization

GRE and PBR Guest and Remediation

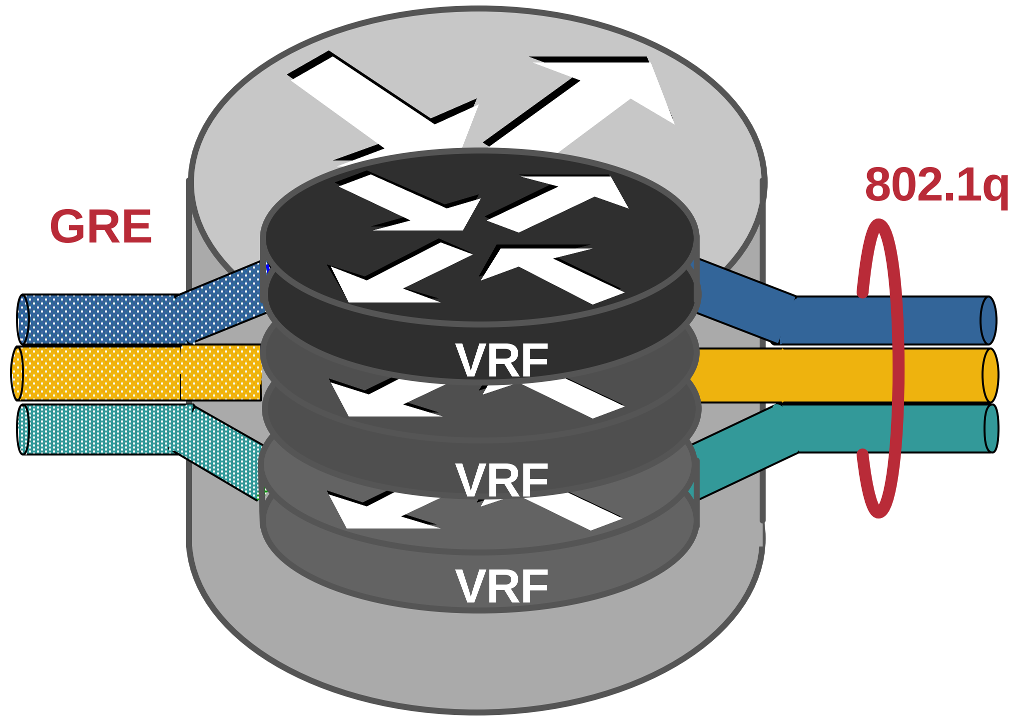
- Configure Policy-Based Routing (PBR) to route all Guest/Remediation traffic via GRE tunnels
- Forces traffic to the DMZ or Remediation Zone




```
interface Vlan250
 description Guest_VLAN
 ip policy route-map Guest-VLAN-to-DMZ

interface Loopback0
 ip address 10.1.250.5 255.255.255.255

interface Tunnel0
 ip address 10.1.250.9 255.255.255.252
 tunnel source Loopback0
 tunnel destination 10.1.250.10

ip access-list extended Guest-VLAN-to-DMZ
 permit ip any any
route-map Guest-VLAN-to-DMZ permit 10
 match ip address Guest-VLAN-to-DMZ
 set interface Tunnel0
```

Virtualized Devices and Data Paths
VRF (Virtual Routing and Forwarding)

- **VRF allows for the creation of multiple logical forwarding tables**
  - Distinct Routing Information Base (RIB)
  - Distinct Forwarding Information Base (FIB)
- **It is possible to associate with each VRF a group of unique logical data paths, e.g.**
  - 802.1q VLANs
  - GRE Tunnels
- **Leverage multipoint GRE (mGRE) and Next Hop Resolution Protocol (NHRP) to ease configuration**



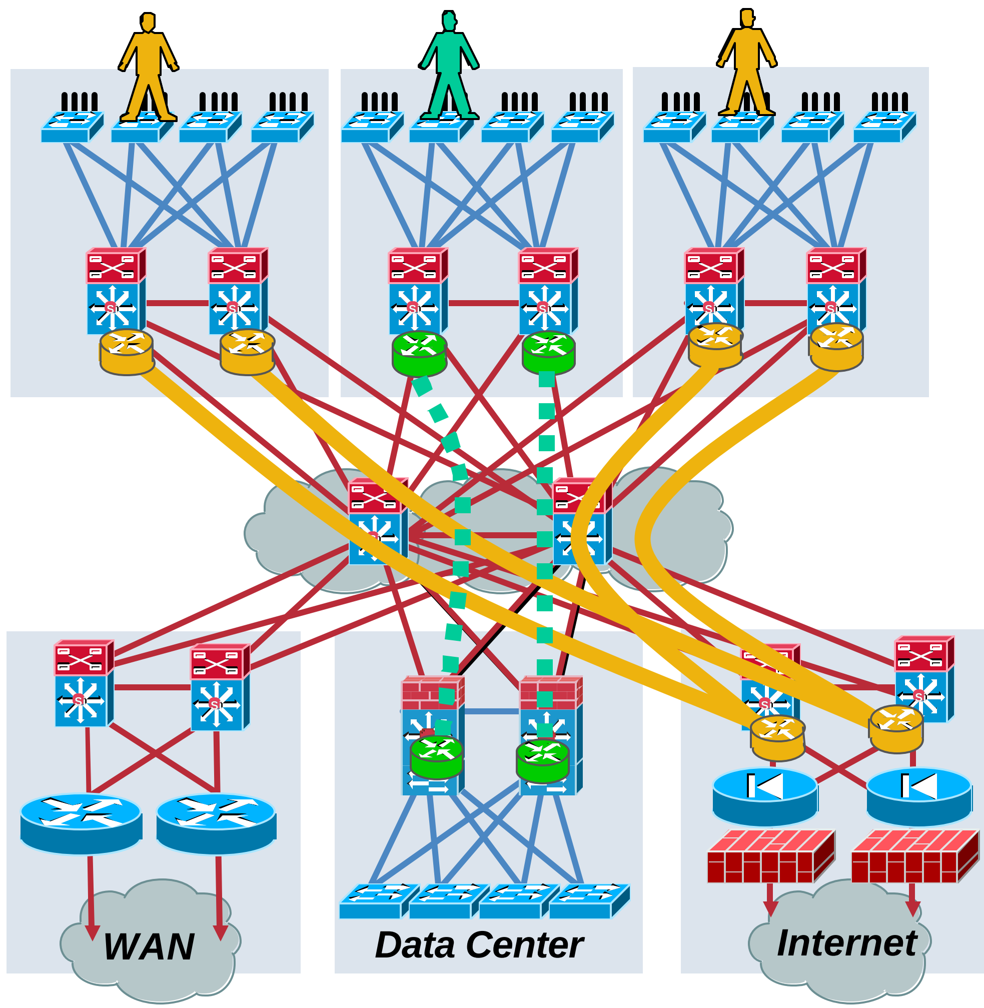
Traffic Is Routed from Each 802.1q VLAN to the Associated GRE Tunnel

Segmentation and Virtualization

GRE and VRF Guest and Remediation

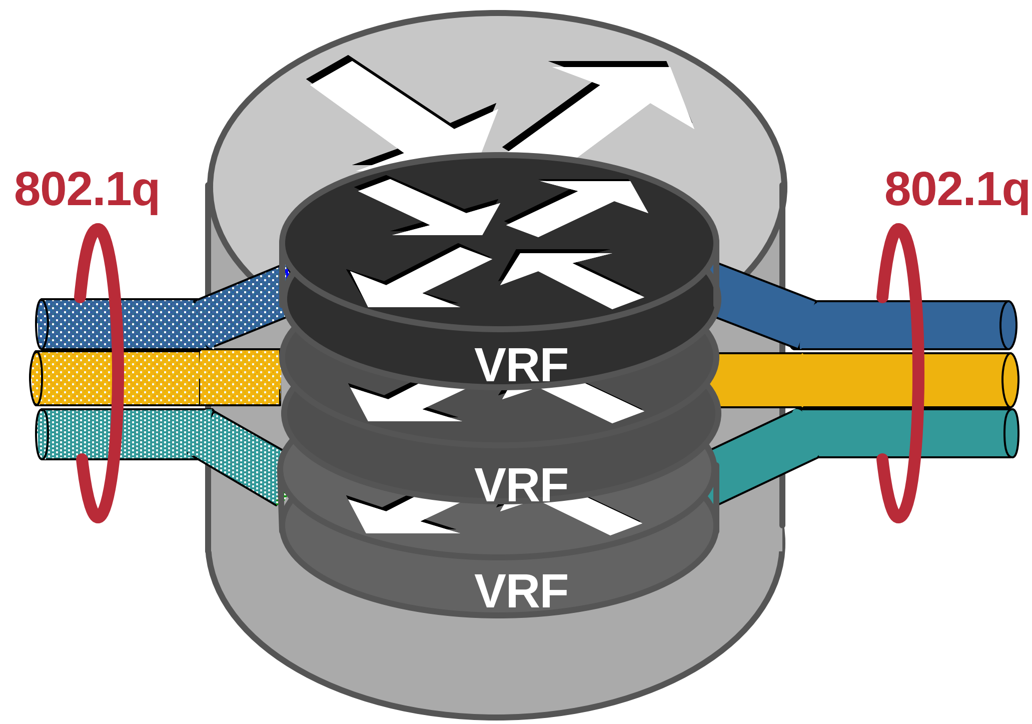
- **Control traffic coming from a specific VLAN to only be able to be forwarded on specific GRE tunnels**
- **Utilize mGRE and NHRP to simplify the configuration of the tunnels**

```
ip vrf GuestAccess
 rd 10:10

interface loopback100
 ip address 10.1.4.3

interface tunnel 0
 ip vrf forwarding GuestAccess
 ip address 192.168.100.2 255.255.255.0
 ip mtu 1416
 ip nhrp map 192.168.100.1 10.126.100.254
 ip nhrp map multicast 10.126.100.254
 ip nhrp network-id 100
 ip nhrp nhs 192.168.100.1
 tunnel source Loopback100
 tunnel destination 10.126.100.254

interface vlan 10
 ip address 192.1.1.4
 ip vrf forwarding GuestAccess
```



Virtualized Devices and Data Paths
End to End VRF-Lite (802.1q Virtual Links)

- **VRF-Lite utilizes hop by hop 802.1q to VRF mapping to build a closed user group**
- **Association of VRF to VLAN is manually configured**
- **Each VRF Instance needs a separate IGP process (OSPF) or address family (EIGRP, RIPv2, MP-BGP)**
- **In this configuration Traffic is routed from each 802.1q VLAN to the associated 802.1q VLAN**



VRF-Lite Supported on 6500, 4500 Sup IV and Sup V, 3560 and 3750

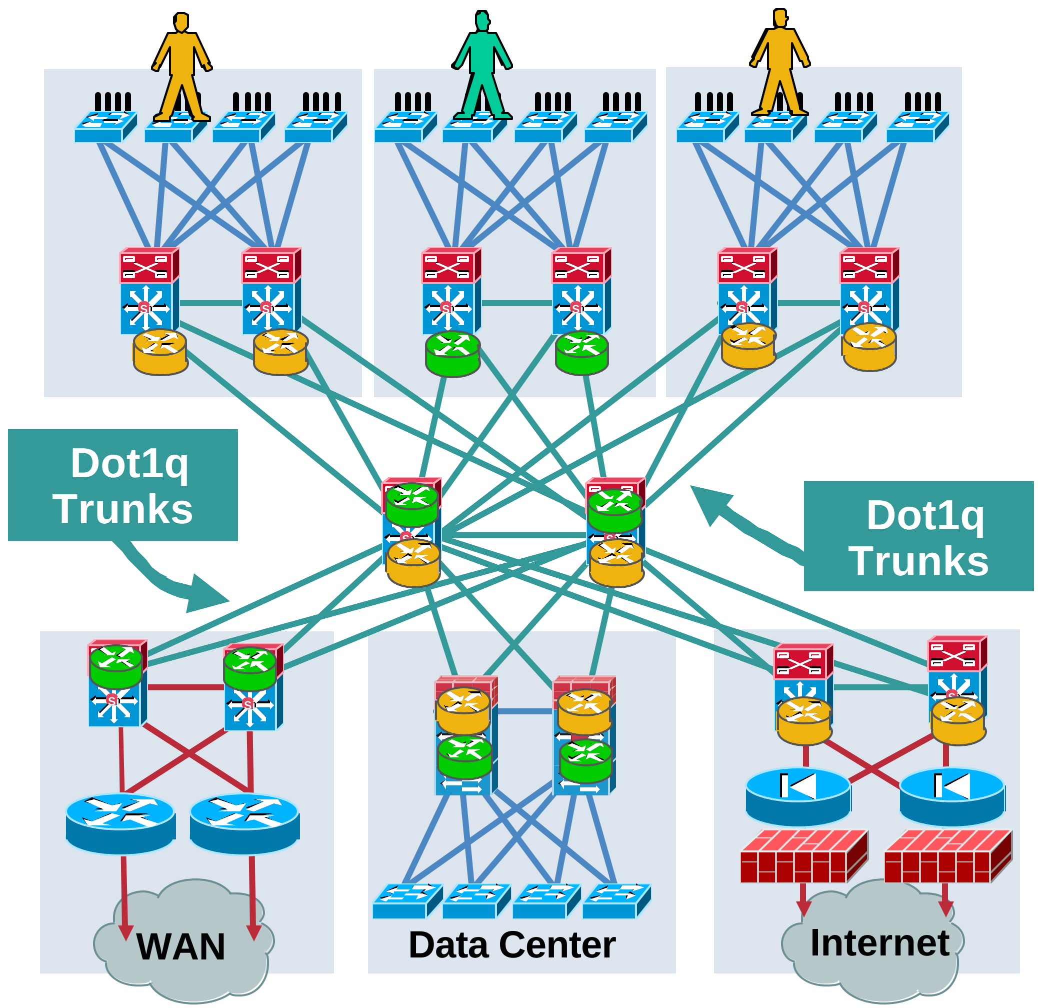
Segmentation and Virtualization

End to End VRF-Lite

Cisco.com

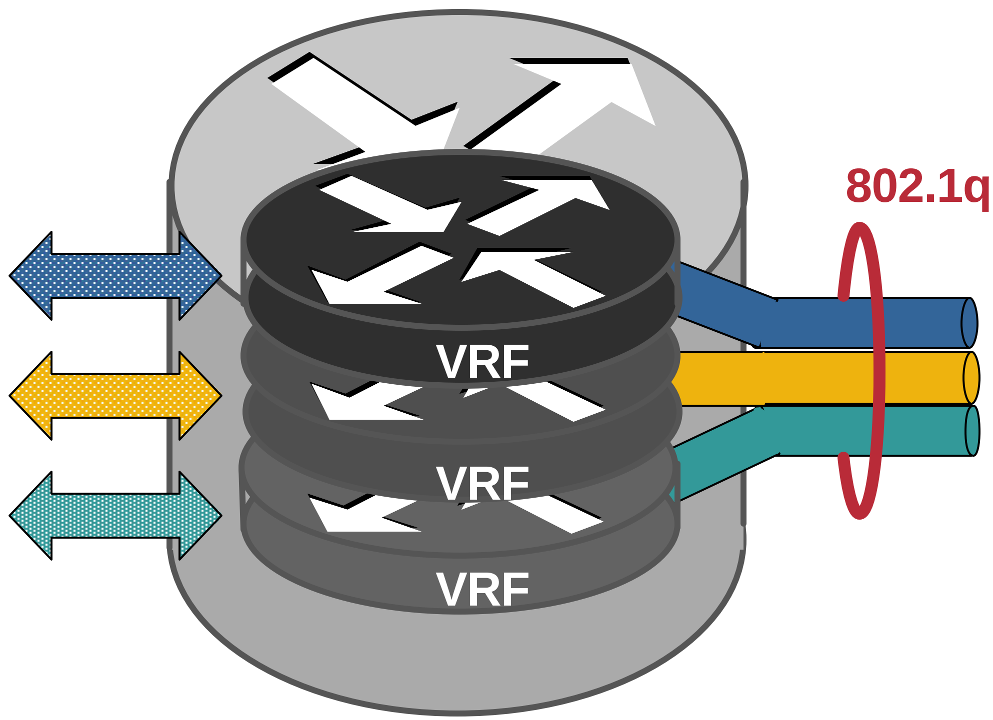
- Configuring distinct Guest and Remediation VRFs allows the network to keep that traffic isolated
- Also can be extended to support other CUGs

```
ip vrf GuestAccess
 rd 10:10

interface vlan 10
 ip address 10.10.2.4
 ip vrf forwarding GuestAccess

interface vlan 110
 ip address 10.100.1.4
 ip vrf forwarding GuestAccess

router eigrp 200
 address-family ipv4 vrf GuestAccess
 network 10.0.0.0
 no auto-summary
 exit-address-family
```



Virtualized Devices and Data Paths
VRF with MPLS Tag Switching

- **When the number of Closed User Group's exceeds 3 or the number of hops > 3 then consider using MPLS tag switching as the virtual data path**
- **No CE, either L2 access or access switch PE**
- **VPN at the first L3 hop (distribution = PE)**
- **MP-iBGP at the distribution only (PE)**
- **MPLS in core and distribution (P and PE)**
- **Overlaid onto existing IGP**

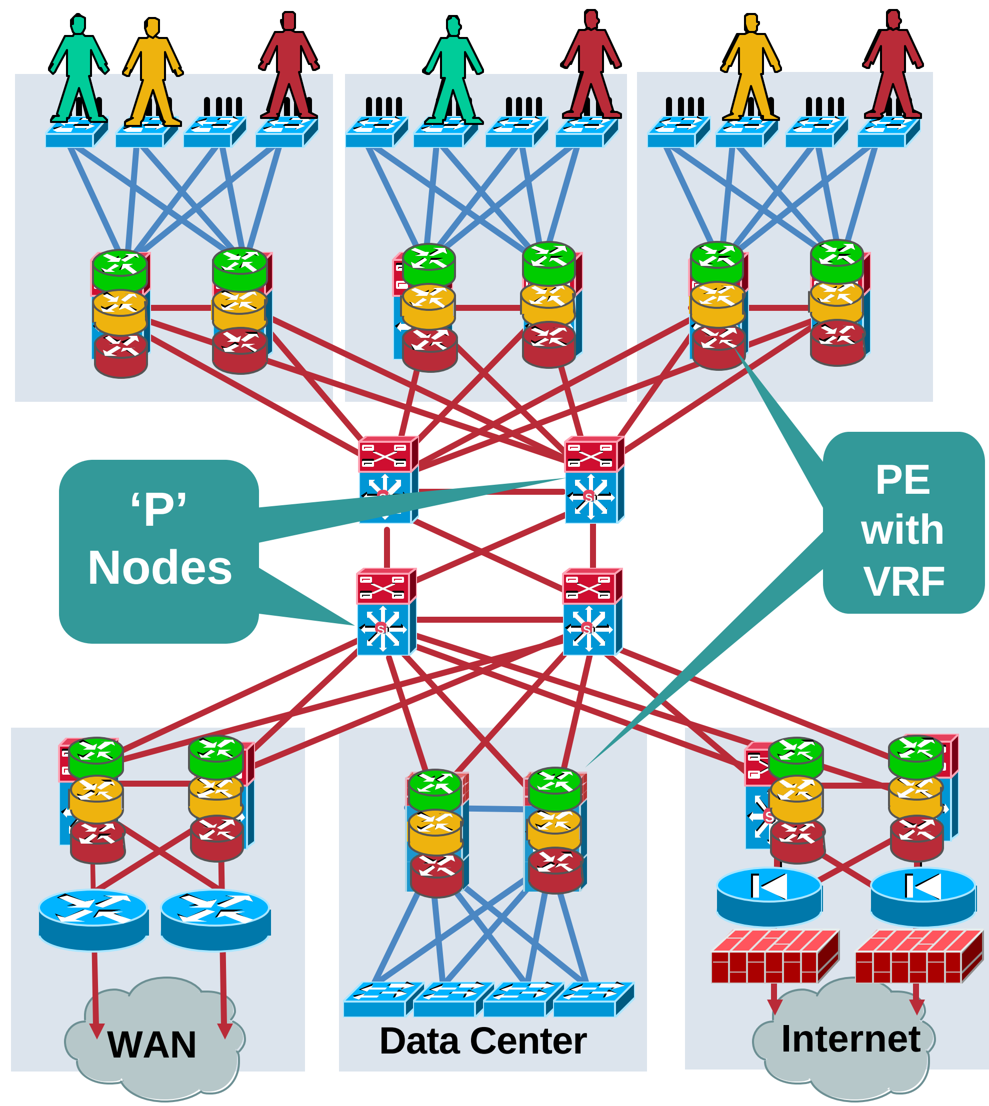


Segmentation and Virtualization

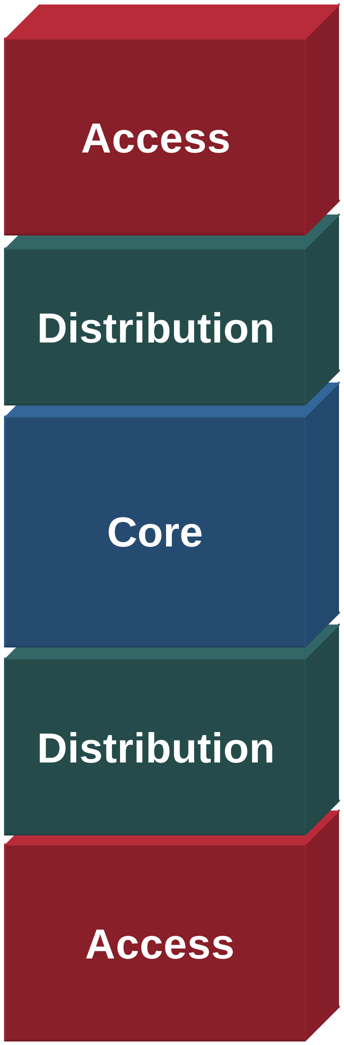
Closed User Group with RFC 2547 VPNs

- **Provides for larger full meshed any-to-any connectivity within each Closed User Group**

```
ip vrf Red
 rd 100:33
 route-target both 100:33
interface FastEthernet0/0
 ip address 10.0.0.11 255.255.255.252
 tag-switching ip
interface Vlan11
 ip vrf forwarding Red
 ip address 10.20.4.1/24

router bgp 100
 no bgp default ipv4-unicast
 neighbor 1.1.1.5 remote-as 100
 neighbor 1.1.1.5 update-source Loopback0
 address-family vpnv4
 neighbor 1.1.1.5 activate
 neighbor 1.1.1.5 send-community extended
address-family ipv4 vrf Red
network 10.20.4.0 mask 255.255.255.0
```



Multilayer Campus Design
Leveraging Advanced Technologies



- Hierarchical design is still the rule
- However there are new ways to implement HA within the distribution block
- QoS is a core feature not an option; it both protects and secures
- Wireless roaming is now possible to do in your structured design
- However wireless and 802.1x are creating the need for additional VLANs
- Data, voice, wireless mgmt, wireless MCast, guest and quarantine VLANs
- Security is not just ACLs and firewalls it is also about integrated anomaly prevention
- High availability depends on being able to survive the unexpected