
TOMORROW starts here.
cisco
Cisco live!

# Designing Layer 2 Networks – Avoiding Loops, Drops, and Flooding

BRKCRS-2661

Roland Salinas

Technical Marketing Engineer

Cisco live!

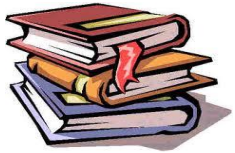
# Abstract

Designing Layer 2 networks is easy.

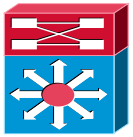
Apparently, In fact there are many traps and dependencies. Three issues of Layer 2 networks - loops, traffic drop and excessive flooding can be demanding. This session is to discuss and present how to avoid them with the standard design techniques or by new mechanisms.

# Presentation Legend

Key Points



Reference Material



Standalone Multilayer Switch

Virtual Switching System



Layer 2 Link



Layer 3 Link

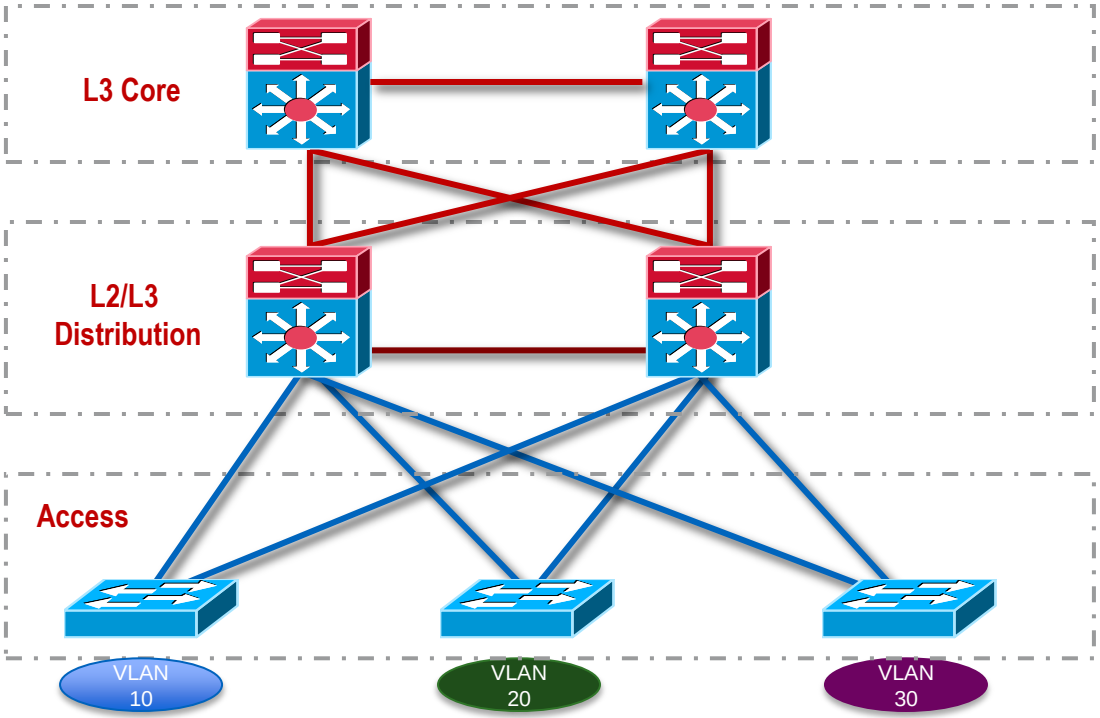
# Agenda

- L2 Network Design Challenges
- Layer 1 and Layer 2 Best Practices
- Spanning Tree Toolkit
- Integrated Security Toolkit
- Control Plane Protection
- Alternative Designs



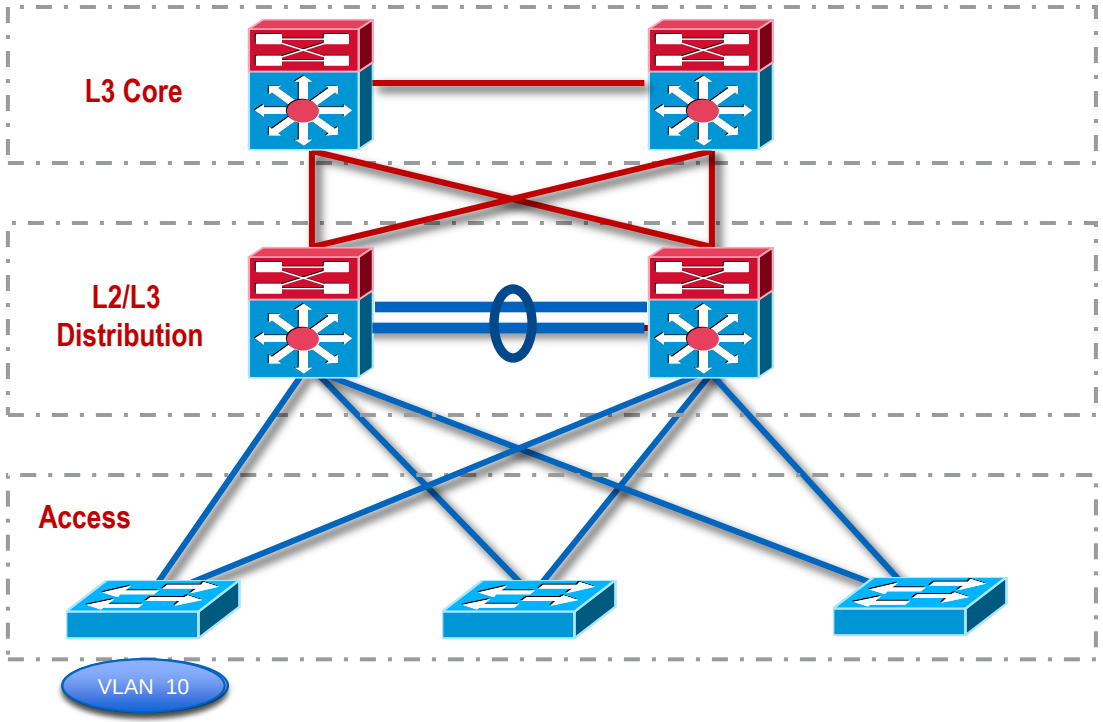

# L2 Network Design Challenges

# Traditional Multi-Layer Design – No L2 Loops



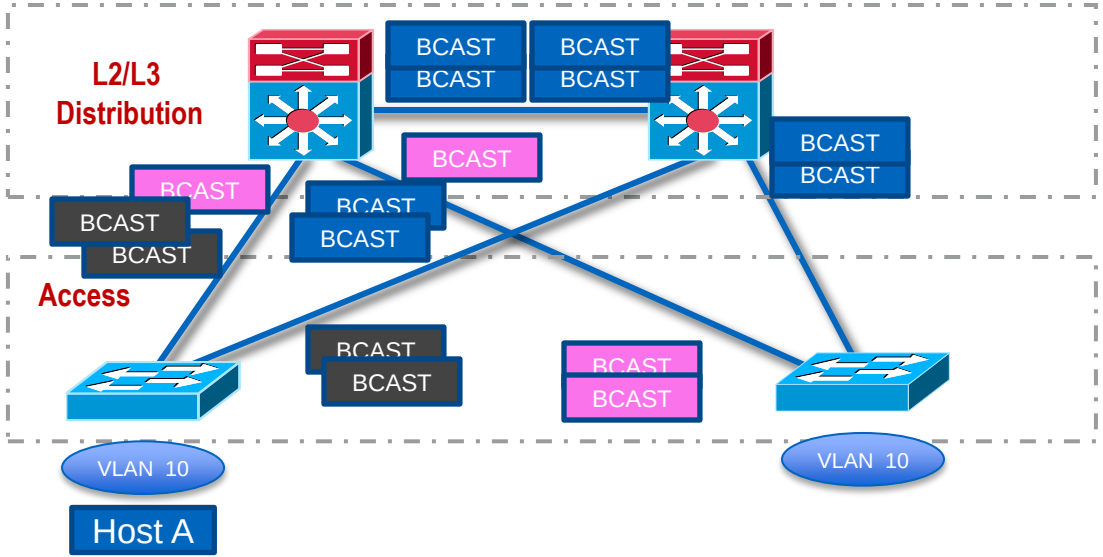

- One switch per subnet per vlan
- Simple design
- Limits L2 domain size to port density to size of the access switch

# Traditional Multi-Layer Design – With L2 Loops




- Extending the L2 domain beyond the single switch
- Best practice says
  - Distribution link must be an L2 link
  - Redundant Links
- Now we have the loop

# L2 Loop – What's the Problem?




- Broadcast and multicast storm
- Source MAC address appear to be moving around as the MAC gets learned on different ports
- Frames are replicated repeatedly

# Effects of a Broadcast Storm

- Bandwidth gets consumed by the frame replication
- CPU utilisation on network attached devices can start to reach high levels due to processing the broadcast traffic
- MAC addresses move from one port another
- Traffic drops
- This can occur with broadcast , multicast and unknown unicast traffic

# Solution: Harden and Mitigate the Design

- Layer 1 Best Practices
- Layer 2 Best Practices
- Spanning Tree Protocol Best Practices

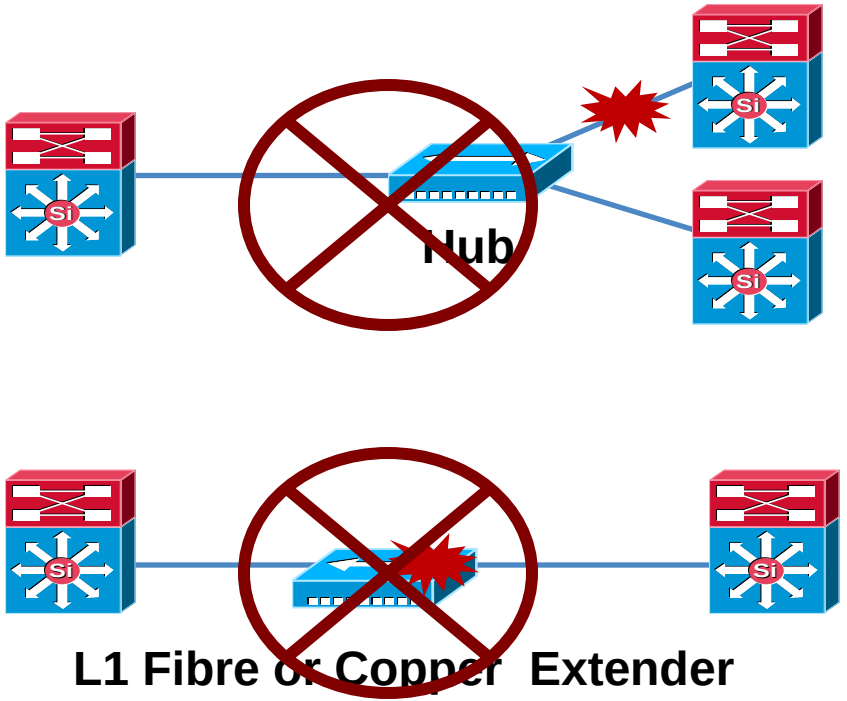
**“Make the network fail closed”**

# Layer 1 and Layer 2 Best Practices

# Layer 1 Best Practice

## Use point-to-point links only

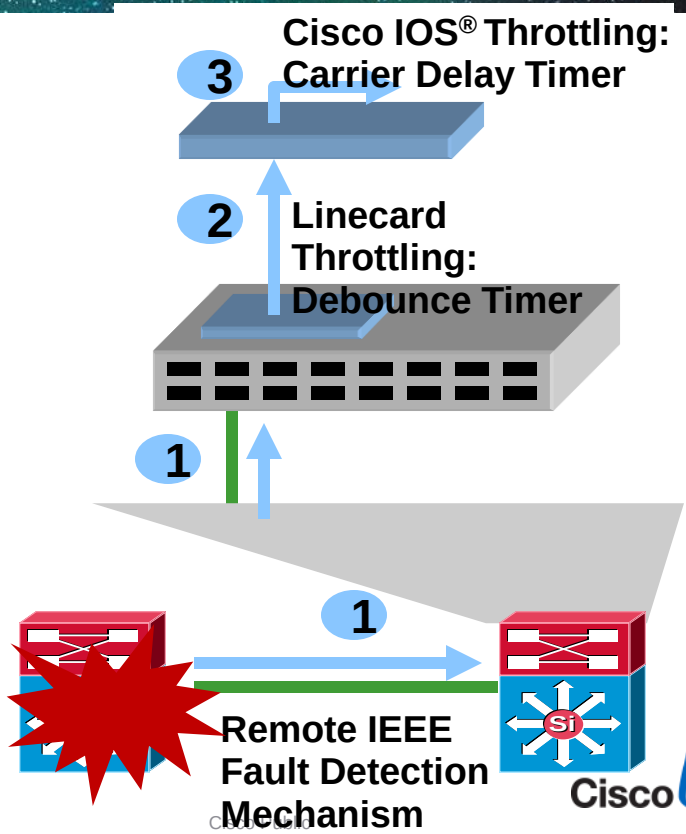
- Eliminate or Avoid at all cost intermediate L1 devices
- Use point-to-point link only



# Redundancy and Protocol Interaction

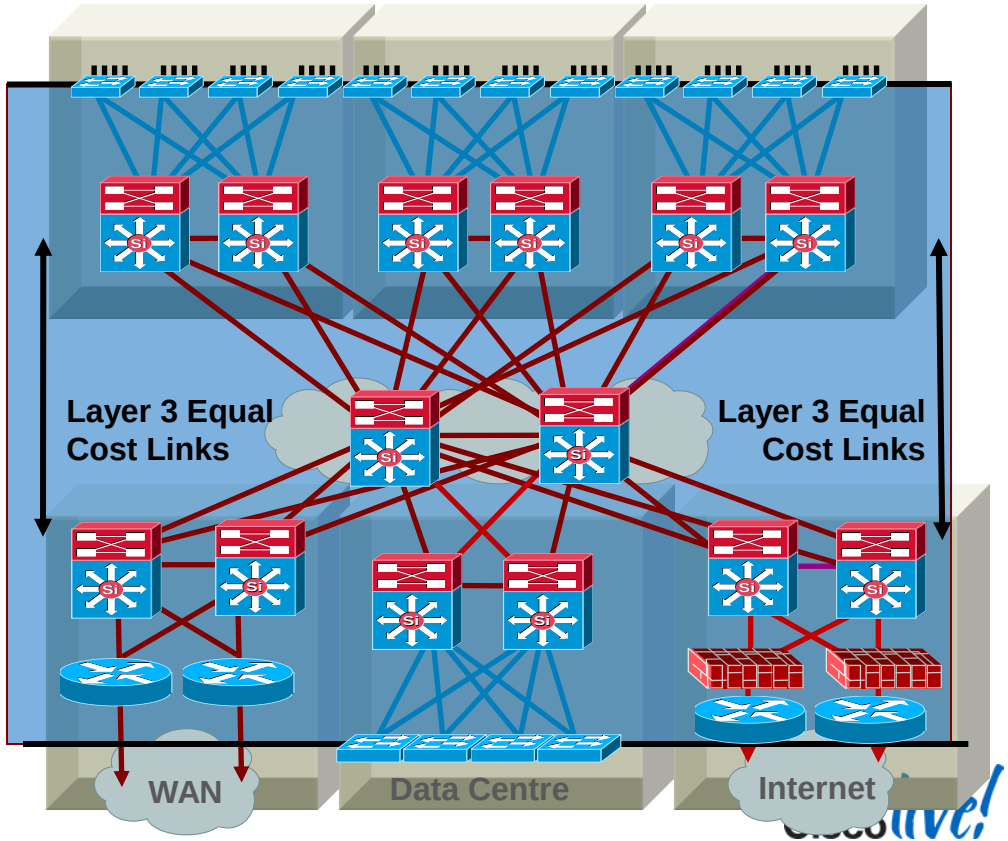
## Fibre Links Versus Copper Links

- Direct point-to-point fibre provides for fast failure detection
- IEEE 802.3z and 802.3ae link negotiation define the use of remote fault indicator and link fault signalling 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
- The default debounce timer on GigE and 10GigE fibre linecards is 10 msec
- The minimum debounce for copper is 300 msec
- Carrier-delay
  - Default 10 microseconds is adequate for most applications

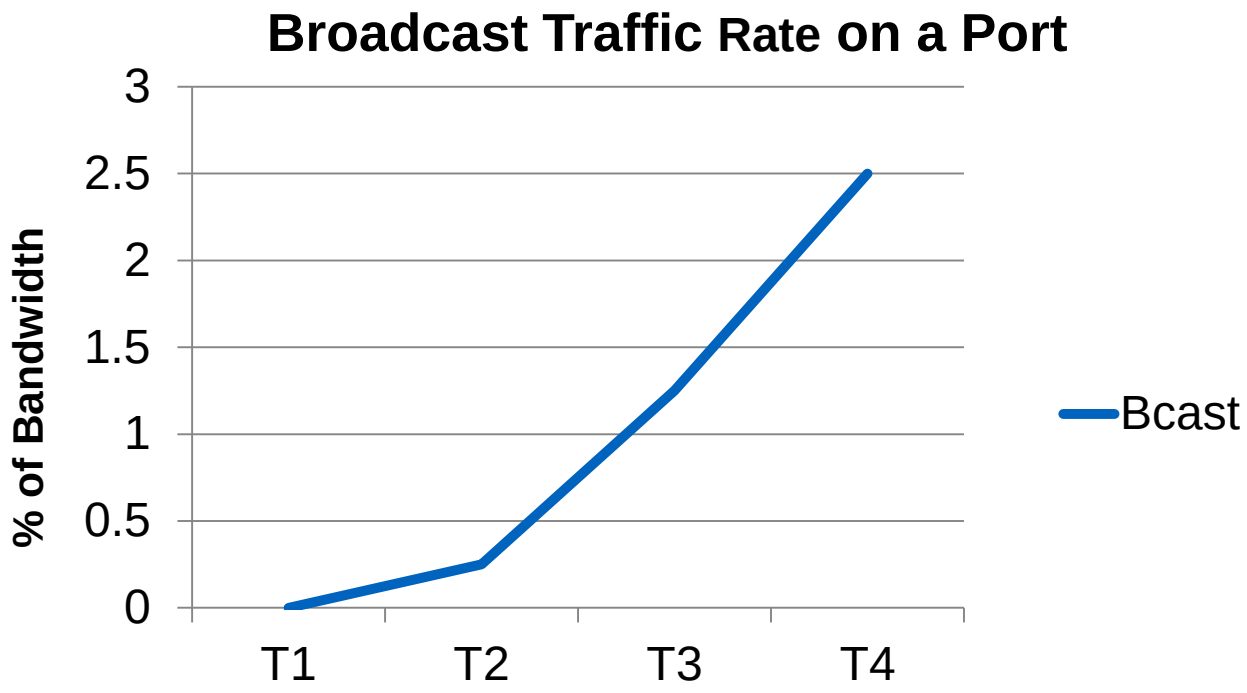


# Best Practices
## Layer 1 Physical Things

- Use point-to-point interconnections—no L2 aggregation points between nodes
- Use fibre for best convergence (debounce timer)

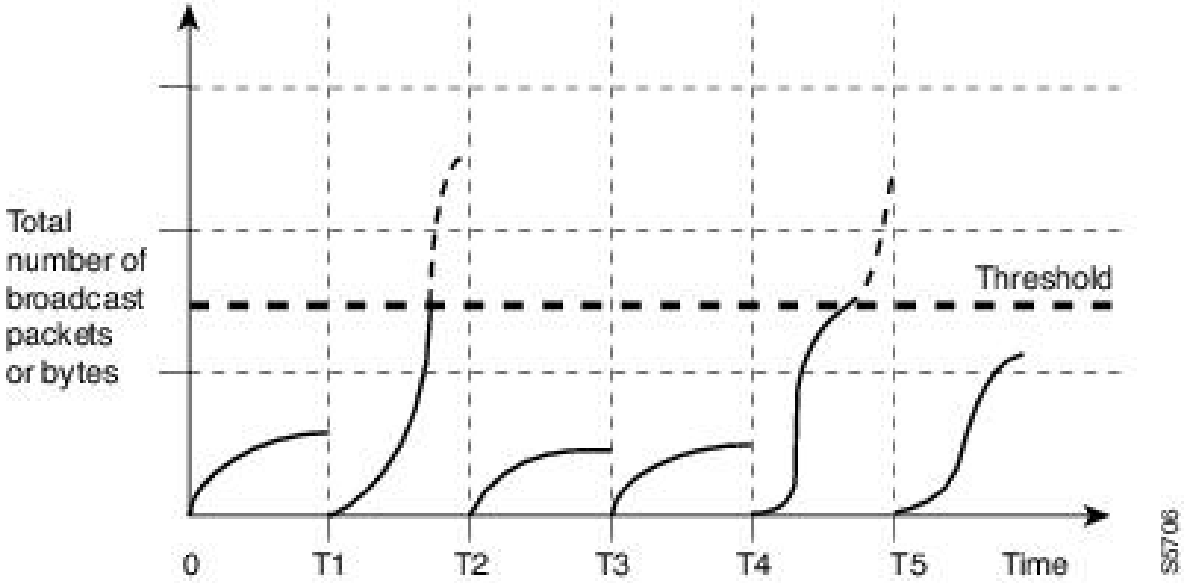


# L2 Loops and the Effect of Frame Replication on Interface Bandwidth



# Storm Control / Traffic Suppression

## Use Hardware-based rate-limiting to Protect and Harden the Broadcast Domain




- Limit broadcast. multicast and unknown unicast to a specific rate
- 1 sec sample interval
- The rate-limiter will drop all profiled traffic above the threshold rate during the sample period
  - All profile traffic including legitimate traffic

# Determining Typical Rates

- Take some time to monitor bcast, mcast rates under normal conditions
  - Use Top N tools (if available on the platform)
  - Netlfow monitoring
  - Interface counters and SNMP tools
  - Wire Shark monitoring

Example using Top N reports on Catalyst 6500

```
VSS01#collect top counters interface ten sort-by broadcast interval 30
TopN collection started.
VSS01#
*Mar  6 16:58:18.735: %TOPN_COUNTERS-SW2-5-STARTED: TopN collection for report 1 started by console
```

# Show Top Reports Example

```
VSS01#collect top counters interface ten sort-by broadcast interval 30
TopN collection started.
VSS01#VSS01#show top counters interface report 1
Started By       : console
Start Time       : 16:58:18 UTC Thu Mar 6 2014
End Time         : 16:58:48 UTC Thu Mar 6 2014
Port Type        : TenGigEthernet
Sort By          : broadcast
Interval         : 30 seconds

Port       Band  Util Bytes            Packets    Broadcast  Multicast  In-  Buf-
           width  (Tx + Rx)           (Tx + Rx)  (Tx + Rx)  (Tx + Rx)  err  ovfl
---------- ----- ---- ------------------ ---------- ---------- ---------- ---- ----

Te1/2/4    10000 0    21559              151        9          142        0    0
Te2/2/4    10000 0    21559              151        9          142        0    0
Te1/2/5    10000 0    18256              140        1          137        0    0
Te2/1/5    10000 0    18160              139        1          137        0    0
Te2/4/1    10000 0    168                2          0          0          0    0
Te1/4/4    10000 9    7072412223         17148485   0          3          0    0
Te2/4/5    10000 0    168                2          0          0          0    0
Te1/4/3    10000 9    7072658842         17148739   0          3          0    0
Te1/4/2    10000 0    6496               82         0          82         0    0
Te1/4/8    10000 0    168                2          0          0          0    0
```

# RMON History Example

```
C3850#show rmon history
Entry 30 is active, and owned by
 Monitors ifIndex.27 every 15 second(s)
 Requested # of time intervals, ie buckets, is 30,
  Sample # 8 began measuring at 3w4d
   Received 128 octets, 0 packets,
   0 broadcast and 0 multicast packets,
   0 undersized and 0 oversized packets,
   0 fragments and 0 jabbers,
   0 CRC alignment errors and 0 collisions.
   # of dropped packet events is 0
   Network utilization is estimated at 0
  Sample # 9 began measuring at 3w4d
   Received 460 octets, 0 packets,
   0 broadcast and 4 multicast packets,
   0 undersized and 0 oversized packets,
   0 fragments and 0 jabbers,
   0 CRC alignment errors and 0 collisions.
   # of dropped packet events is 0
   Network utilization is estimated at 0
  Sample # 10 began measuring at 3w4d
   Received 949 octets, 0 packets,
   0 broadcast and 2 multicast packets,
   0 undersized and 0 oversized packets,
   0 fragments and 0 jabbers,
```

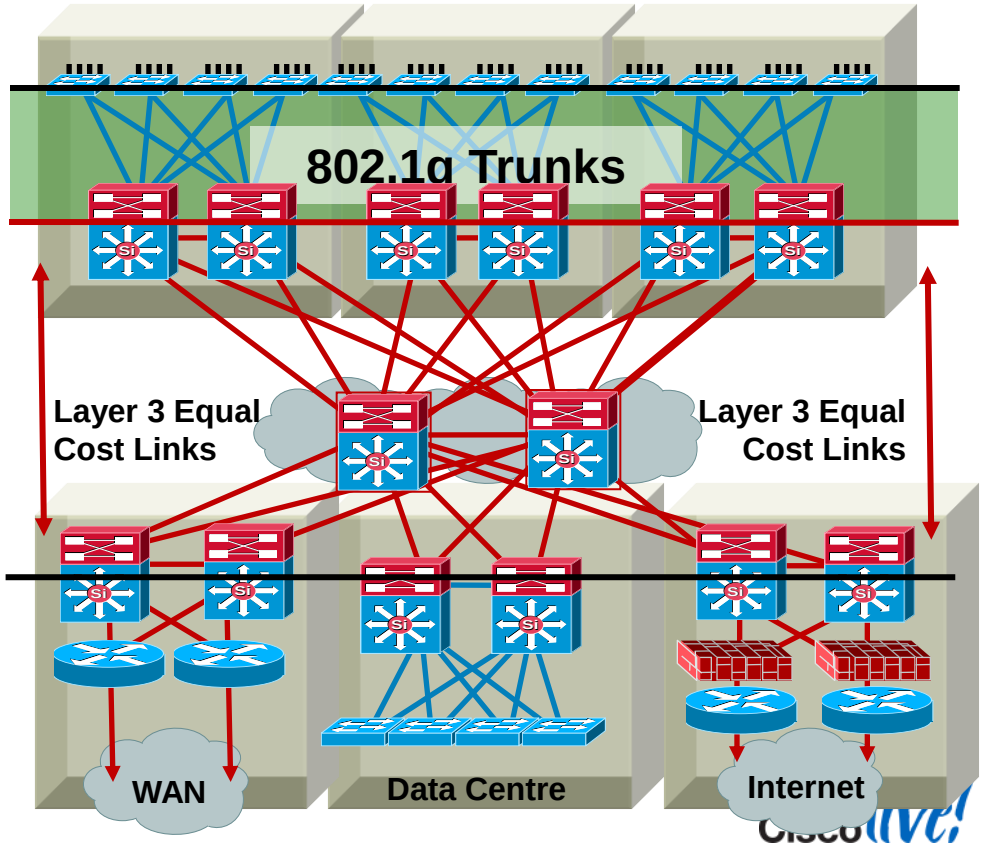
RMON History Example on Catalyst 3850

# Storm Control Recommendations

- Based on your network metrics, choose a value that will allow for peak Broadcast, multicast, unknown unicast  plus 50%
  - 1% bcast rate is common for 1GbE interfaces
  - 0.5% is a common rate for 10GbE interfaces
- The higher the interface speed the less the percentage needs to be
- Be very cautions when configuring storm control for multicast frames, as this can limit BPDUs
- Verify platform specific support and caveats
  - Some platforms treat traffic types individually, some will group bcast and mcast together
  - Some legacy hardware do not support storm-control or implement software-based mechanisums

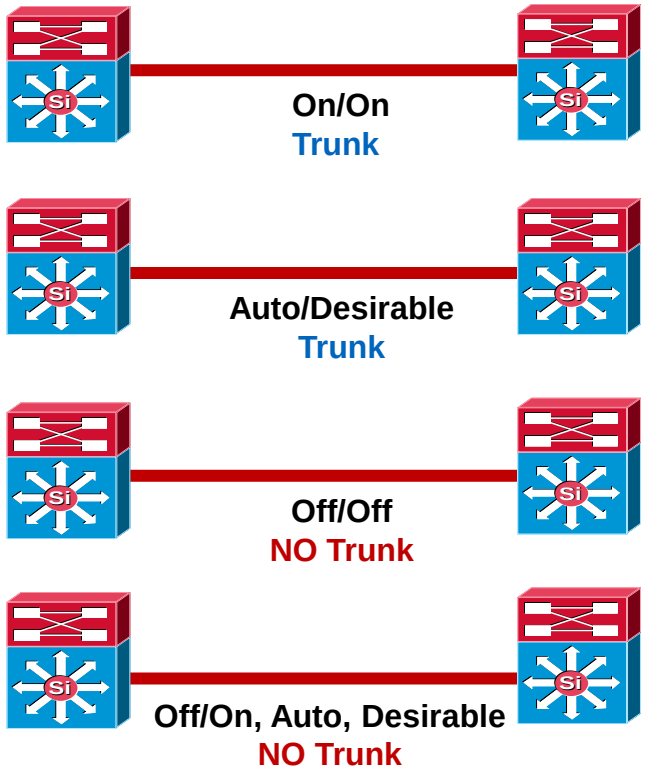
# Best Practices—Trunk Configuration

- Typically deployed on interconnection between access and distribution layers
- Use VTP transparent mode to decrease potential for operational error
- Hard set trunk mode to on and encapsulation negotiate off for optimal convergence
- Change the native VLAN to something unused to avoid VLAN hopping
- Manually prune all VLANS except those needed
- Disable on host ports:
  - Cisco IOS: `switchport host`



# DTP Dynamic Trunk Protocol

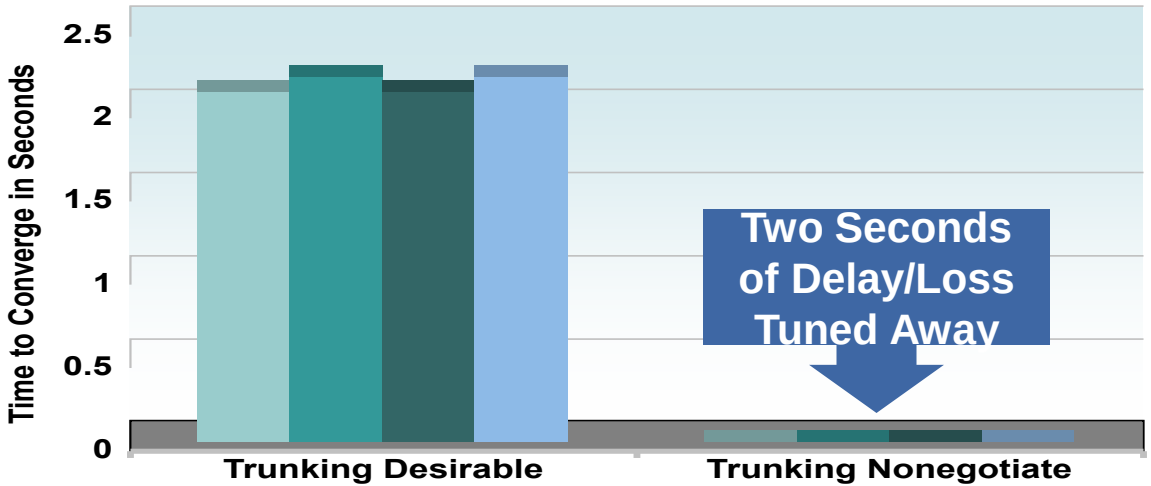
- Automatic formation of trunked switch-to-switch interconnection
  - On: always be a trunk
  - Desirable: ask if the other side can/will
  - Auto: if the other sides asks I will
  - Off: don't become a trunk
- Negotiation of 802.1Q or ISL encapsulation
  - ISL: try to use ISL trunk encapsulation
  - 802.1q: try to use 802.1q encapsulation
  - Negotiate: negotiate ISL or 802.1q encapsulation with peer
  - Non-negotiate: always use encapsulation that is hard set

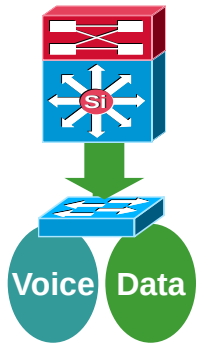


# Optimising Convergence: Trunk Tuning

## Trunk Auto/Desirable Takes Some Time

- DTP negotiation tuning improves link up convergence time
  - `IOS(config-if)# switchport mode trunk`
  - `IOS(config-if)# switchport nonegotiate`



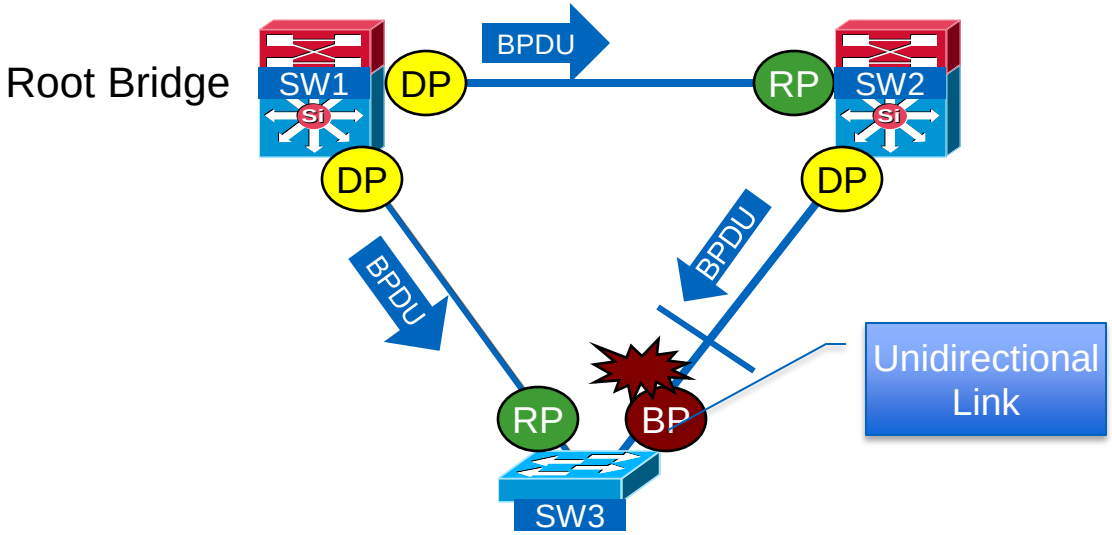



# Trunking/VTP/DTP—Quick Summary

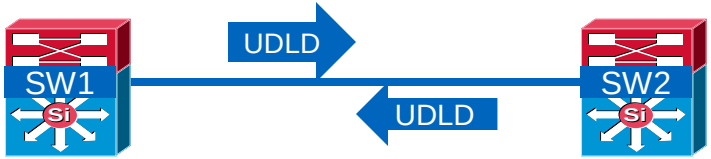
- VTP transparent should be used; there is a trade off between administrative overhead and the temptation to span existing VLANS across multiple access layer switches
- One can consider a configuration that uses DTP **ON/ON** and **NO NEGOTIATE**; there is a trade off between performance/HA impact and maintenance and operations implications
- An **ON/ON** and **NO NEGOTIATE** configuration is faster from a link up (restoration) perspective than a desirable/desirable alternative. However, in this configuration DTP is not actively monitoring the state of the trunk and a misconfigured trunk is not easily identified
- It's really a balance between fast convergence and your ability to manage configuration and change control …

# Unidirectional Link Detection (UDLD)

- Example topology where a uni-directional can cause an L2 loop
- If SW3 stops receiving BPDUs form SW2, SW3 will change its Blocking Port to Forwarding after STP timeout

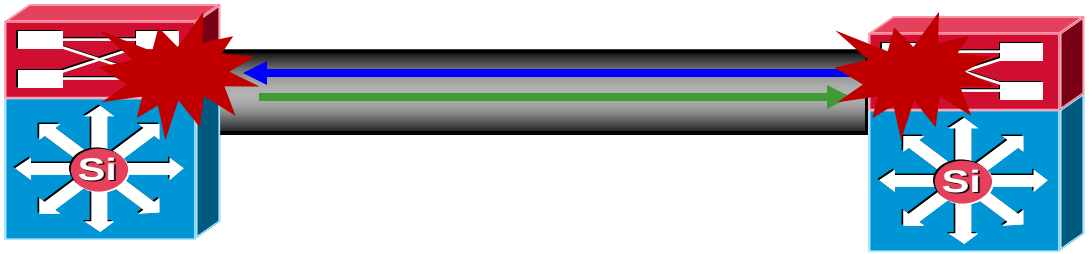


# UDLD Operation




- UDLD works by exchanging protocol packets between the neighbouring devices.
- Both devices on the link must support UDLD and have it enabled on respective ports.
- UDLD protocol packets contain the port's own device/port ID, and the neighbour's device/port IDs seen by UDLD on that port. neighbouring ports should see their own device/port ID (echo) in the packets received from the other side.

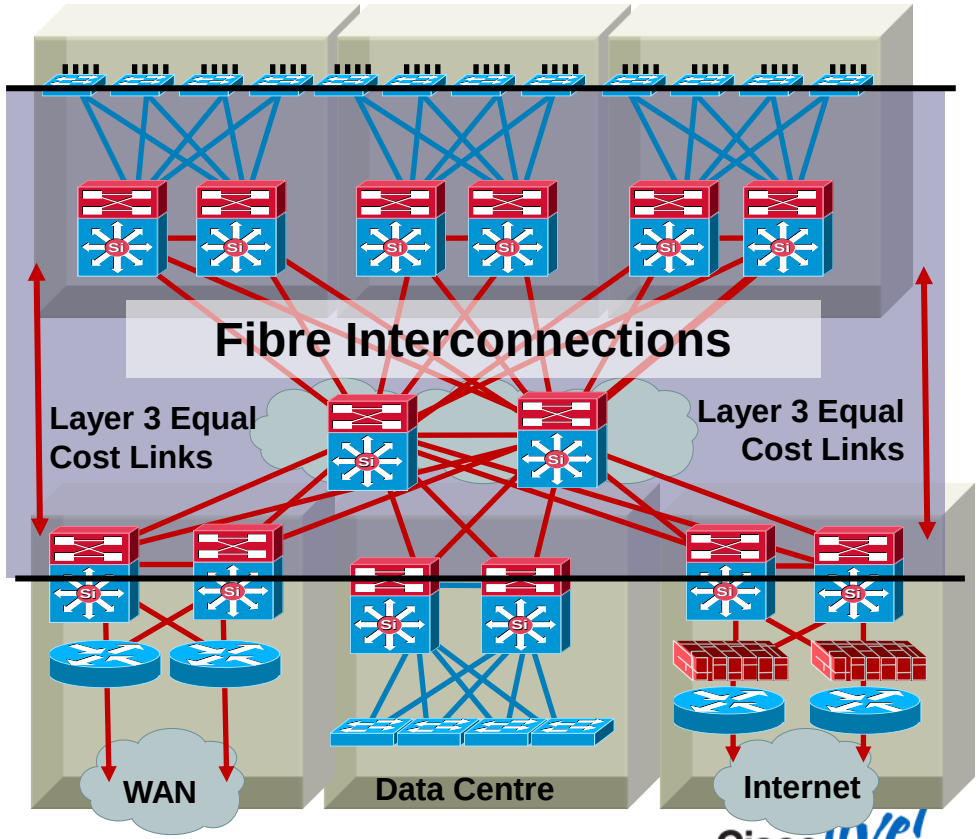
# UDLD Aggressive and UDLD Normal




- Timers are the same—15-second hellos by default
- UDLD—Normal Mode—will transition a port to “undetermined” state
- UDLD—Aggressive—err-disable **both** ends of the connection due to err-disable when aging and re-establishment of UDLD communication fails
  - Aggressive Mode—after aging on a previously bi-directional link—tries eight times (once per second) to reestablish connection then err-disables port

# Best Practices—UDLD Configuration

- Typically deployed on any fibre optic interconnection
  - Use UDLD aggressive mode with caution
  - Ensure an out-of-band console connection to the device in then event that an err-disabled port will cut-off in-band management
- Turn on in global configuration to avoid operational error/misses
- Config example
  - Cisco IOS:
    `udld enable`

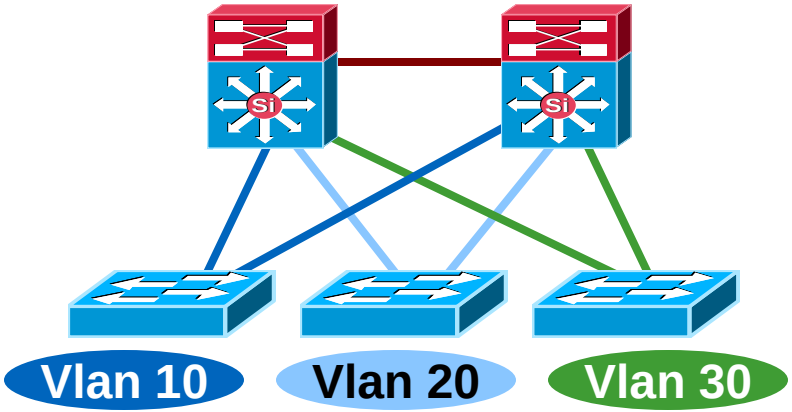




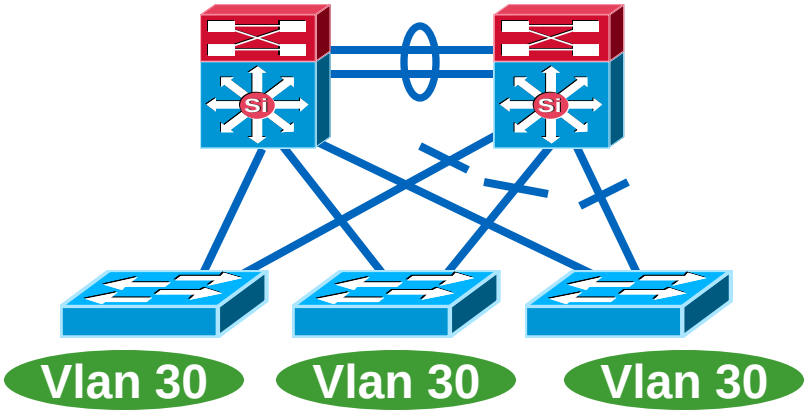

# Spanning Tree Toolkit

# 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

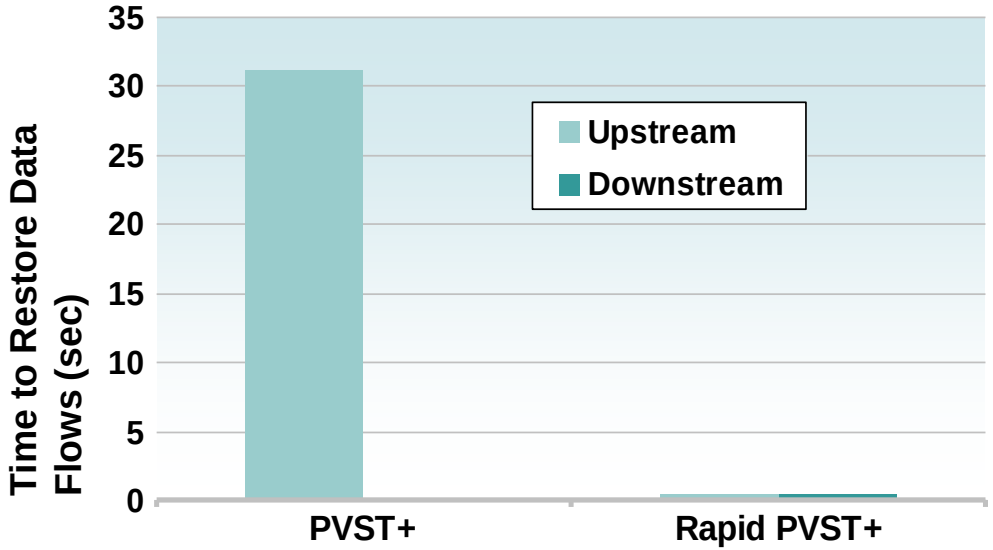



- At least some VLANs span multiple access switches
- Layer 2 loops
- Layer 2 and 3 running over link between distribution
- Blocked links

# Spanning Tree Protocol Opitons

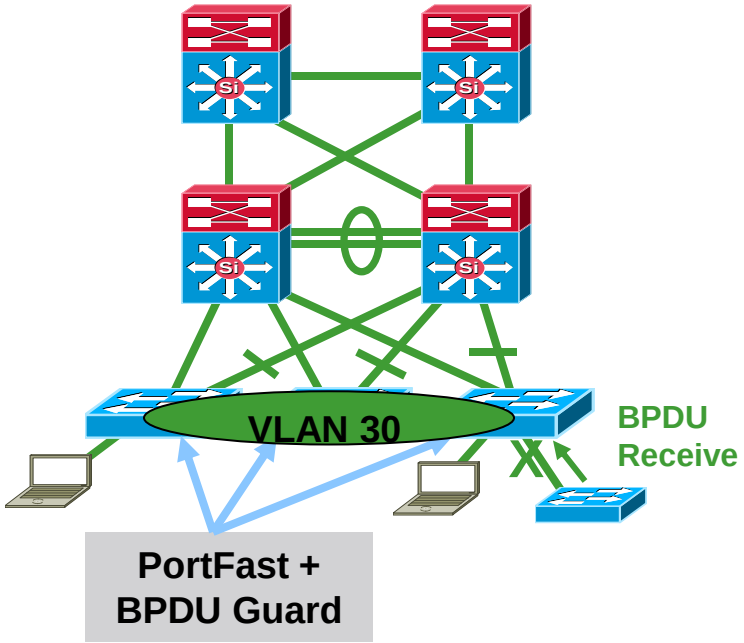
## PVST+, Rapid PVST+ or MST

- Rapid-PVST+ greatly improves the restoration times for any VLAN that requires a topology convergence due to link UP
- Rapid-PVST+ also greatly improves convergence time over backbone fast for any indirect link failures
- PVST+ (802.1d)
  - Traditional spanning tree implementation
- Rapid PVST+ (802.1w)
  - Scales to large size (~10,000 logical ports)
  - **Easy to implement, proven, scales**
- MST (802.1s)
  - Permits very large scale STP implementations (~30,000 logical ports)
  - **Not as flexible as rapid PVST+**



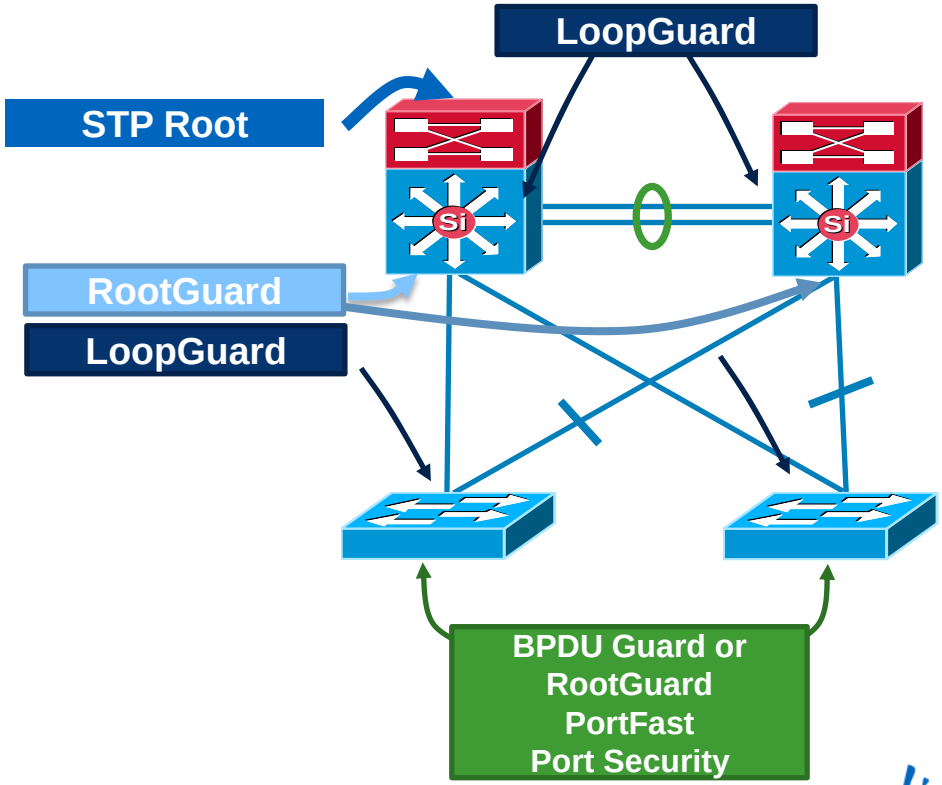
# Optimising the Layer 2 Design
# STP Toolkit – PortFast and BPDU Guard

- PortFast is configured on edge ports to allow them to quickly move to forwarding bypassing listening and learning and avoids TCN (Topology Change Notification) messages
- BPDU Guard can prevent loops by moving PortFast configured interfaces that receive BPDUs to errdisable state
- BPDU Guard prevents ports configured with PortFast from being incorrectly connected to another switch
- When enabled globally, BPDU Guard applies to all interfaces that are in an operational PortFast state



# Force Spanning Tree Perform as Expected

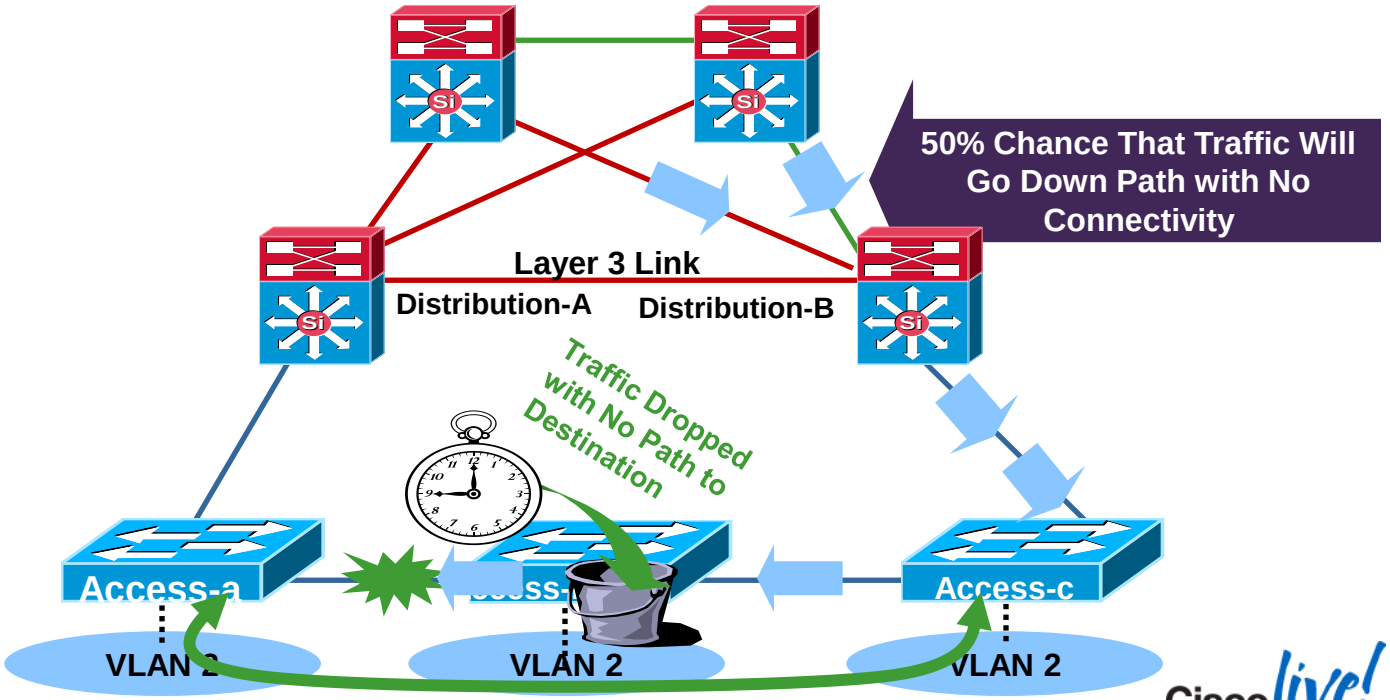
- Place the root where you want it
  - Root primary/secondary macro
- The root bridge should stay where you put it
  - RootGuard
  - LoopGuard
  - UplinkFast
  - UDLD
- Only end-station traffic should be seen on an edge port
  - BPDU Guard
  - RootGuard
  - PortFast
  - Port-security



# Daisy Chaining Access Layer Switches

## Avoid Potential Black Holes

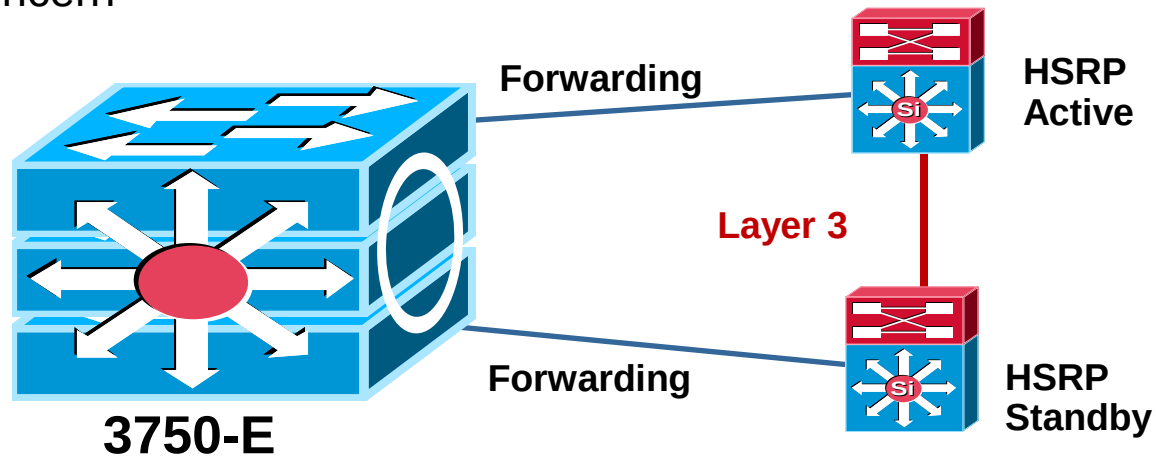

**Return Path Traffic Has a 50/50 Chance of Being 'Black Holed'**



# Daisy Chaining Access Layer Switches

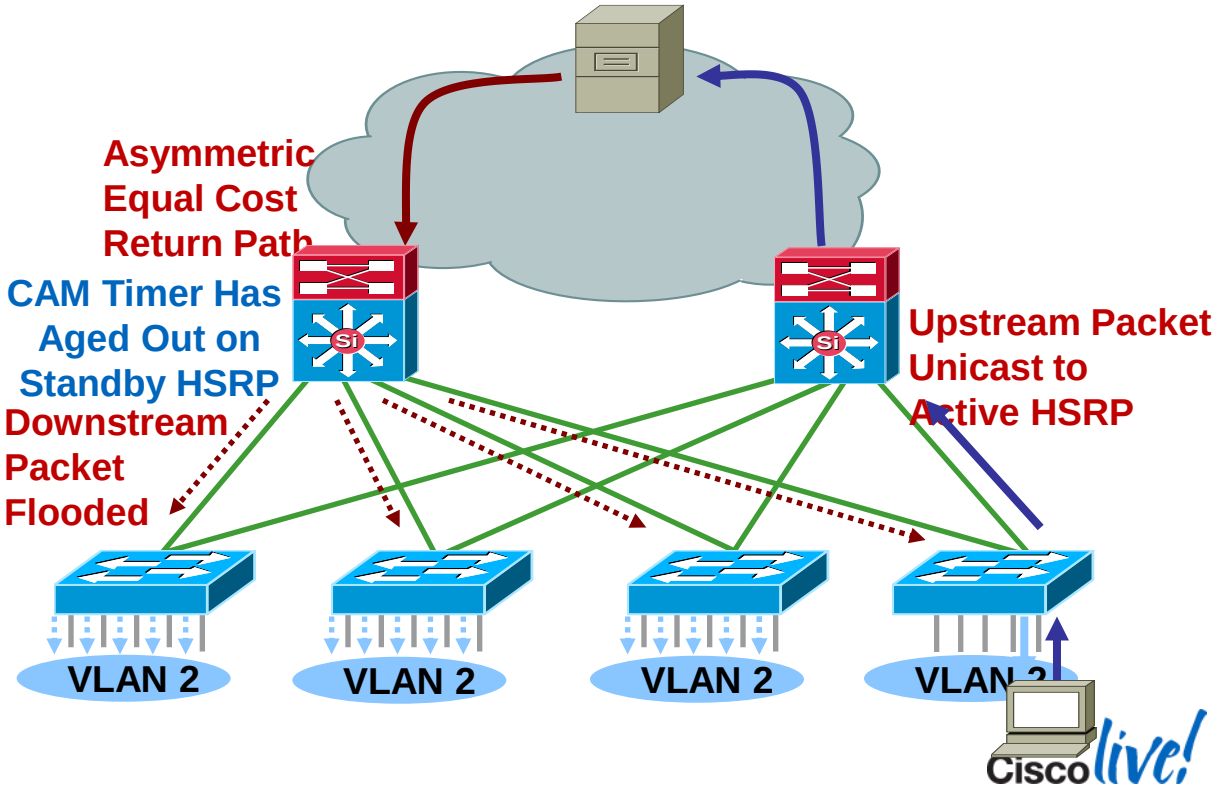
Stacking Technology Addresses Old Problems

- **Stackwise/Stackwise-Plus** technology eliminates the concern
  - Loopback links not required
  - No longer forced to have L2 link in distribution
- If you use modular (chassis-based) switches, these problems are not a concern



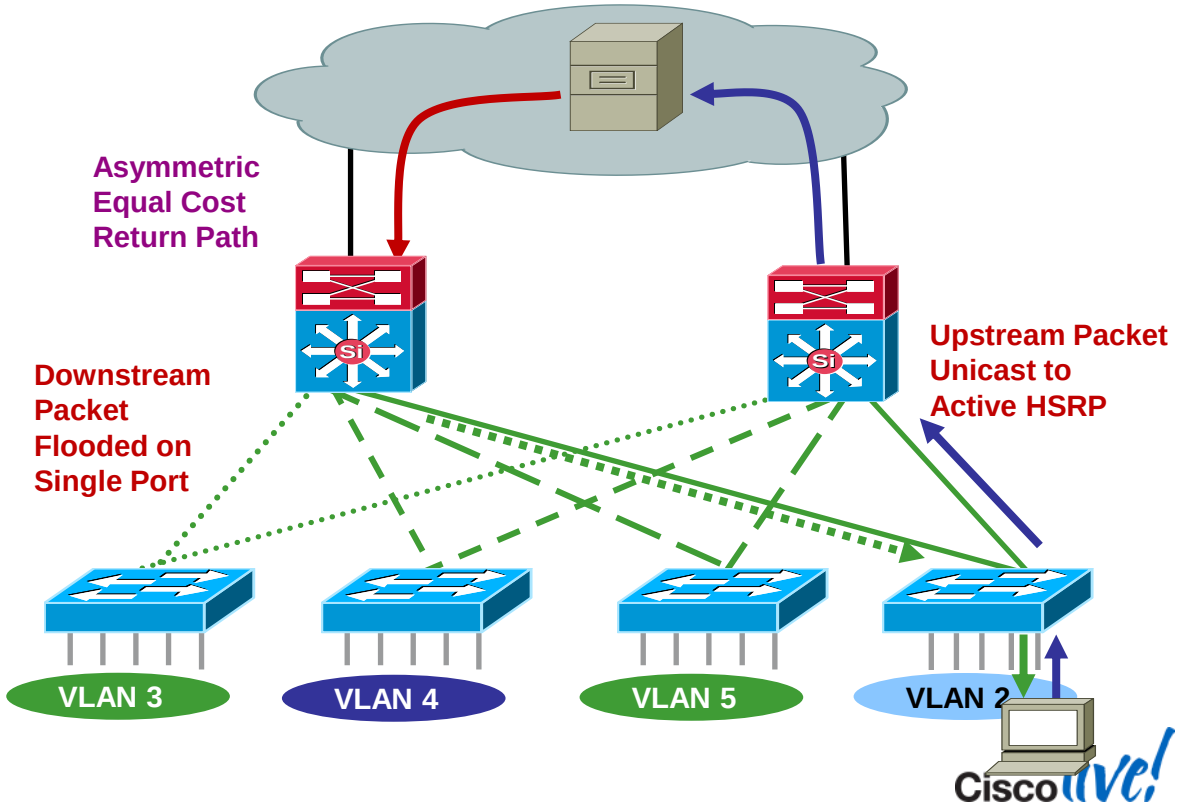
# Asymmetric Routing (Unicast Flooding)

- Affects redundant topologies with shared L2 access
- One path upstream and two paths downstream
- CAM table entry ages out on standby HSRP
- Without a CAM entry packet is flooded to all ports in the VLAN



# Best Practices Prevent Unknown Unicast Flooding

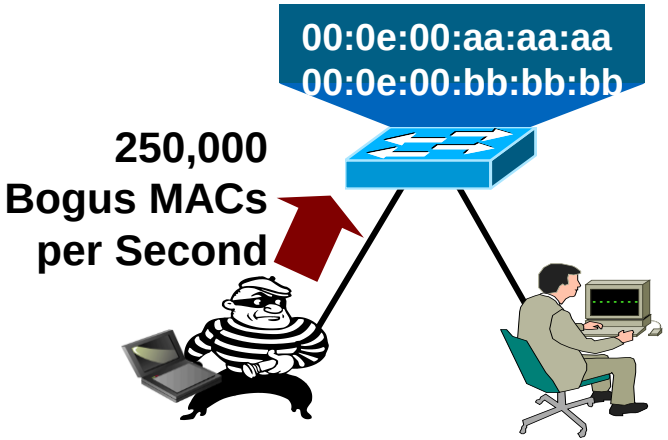
- Assign one unique data and voice VLAN to each access switch
- Traffic is now only flooded down one trunk
- Access switch unicasts correctly; no flooding to all ports
- If you have to:
  - Tune ARP and CAM aging timers; CAM timer exceeds ARP timer
  - Bias routing metrics to remove equal cost routes



# Integrated Security Toolkit

# Securing Layer 2 From Surveillance Attacks

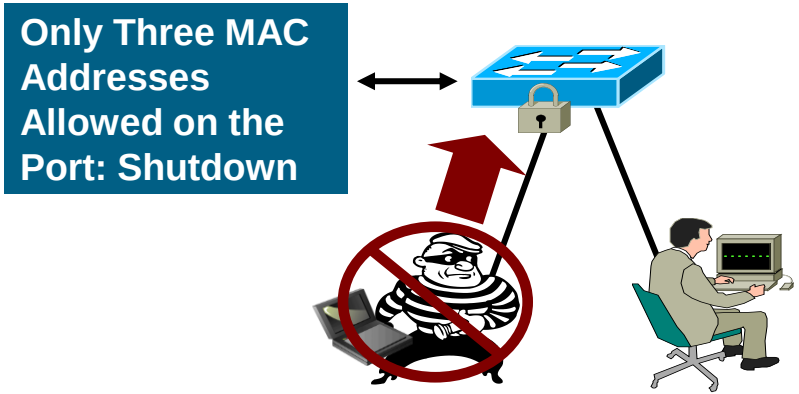
## Cutting Off MAC-Based Attacks




### Problem –

Script Kiddie Hacking Tools Enable Attackers to Flood Switch CAM Tables with Bogus MACs – Turning the VLAN into a Hub and Eliminating Privacy

Switch CAM Table Limit Is Finite Number of MAC Addresses

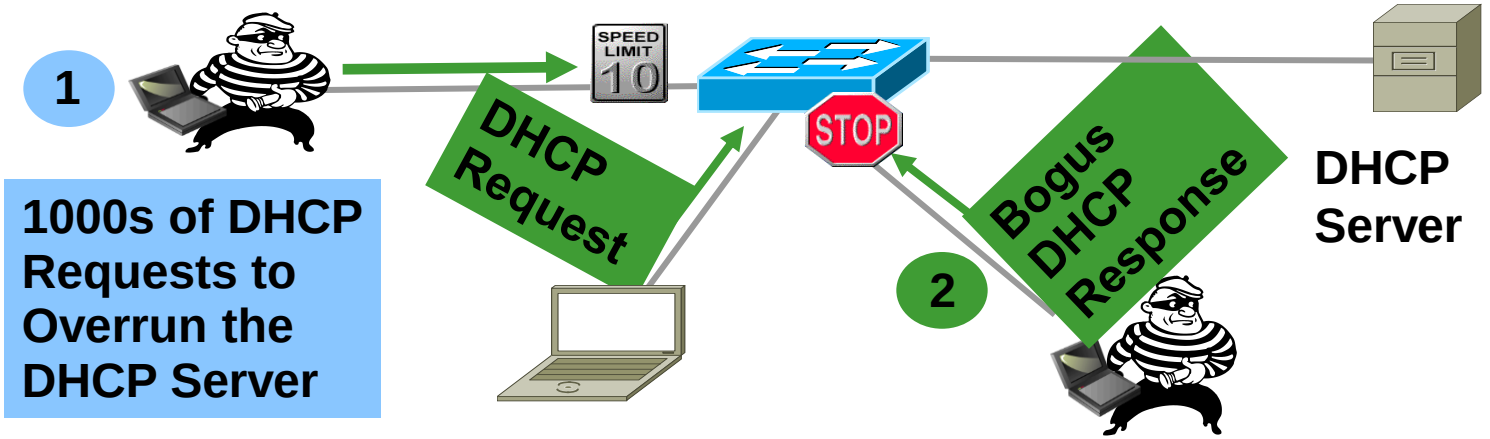



### Solution –

**Port Security** Limits MAC Flooding Attack – Locks Down Port and Sends an SNMP Trap

```
switchport port-security
switchport port-security maximum 10
switchport port-security violation restrict
switchport port-security aging time 2
switchport port-security aging type inactivity
```

# DHCP Snooping
# Protection Against Rogue / Malicious DHCP Server

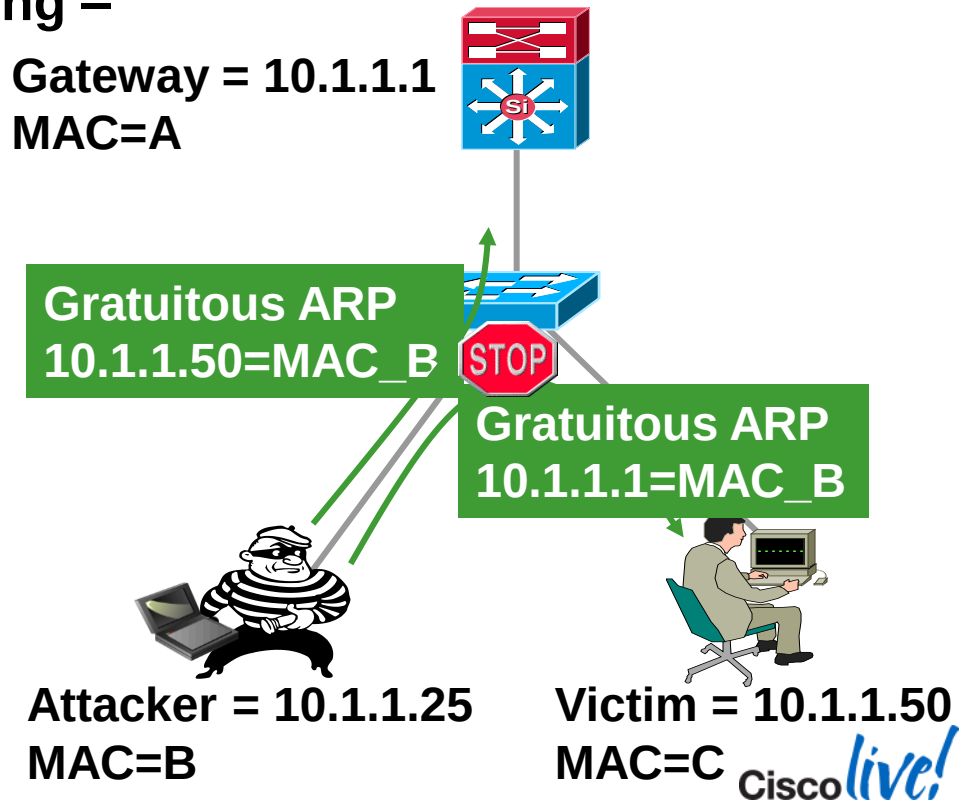



- DHCP requests (discover) and responses (offer) tracked
- Rate-limit requests on trusted interfaces – limits DoS attacks on DHCP server
- Deny responses (offers) on non-trusted interfaces – stop malicious or errant DHCP server

# Securing Layer 2 From Surveillance Attacks

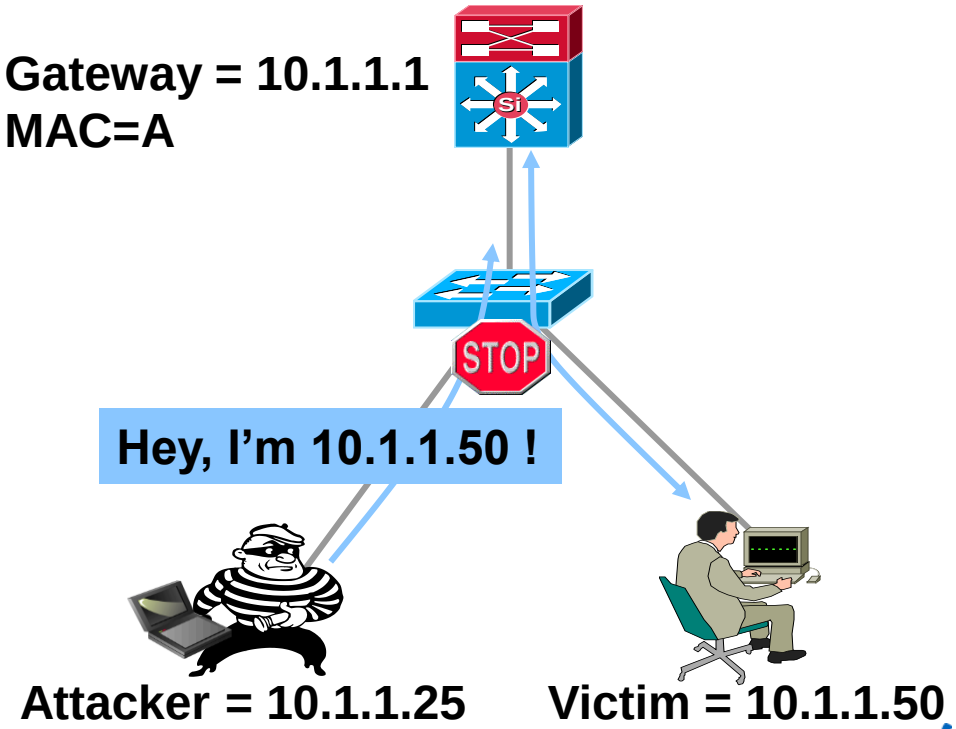
## Protection Against **ARP Poisoning –**

- **Dynamic ARP inspection** protects against ARP poisoning (ettercap, dsnif, arpspoof)
- Uses the DHCP Snooping binding table
- Tracks MAC to IP from DHCP transactions
- Rate-limits ARP requests from client ports
- Drop **bogus** gratuitous ARPs – stops ARP poisoning / MITM attacks



# IP Source Guard
# Protection Against Spoofed IP Addresses

- **IP Source Guard** protects against spoofed IP addresses
- Uses the DHCP Snooping binding table
- Tracks IP address to port associations
- Dynamically programs port ACL to drop traffic not originating from IP address assigned via DHCP



# Catalyst Integrated Security Features Summary

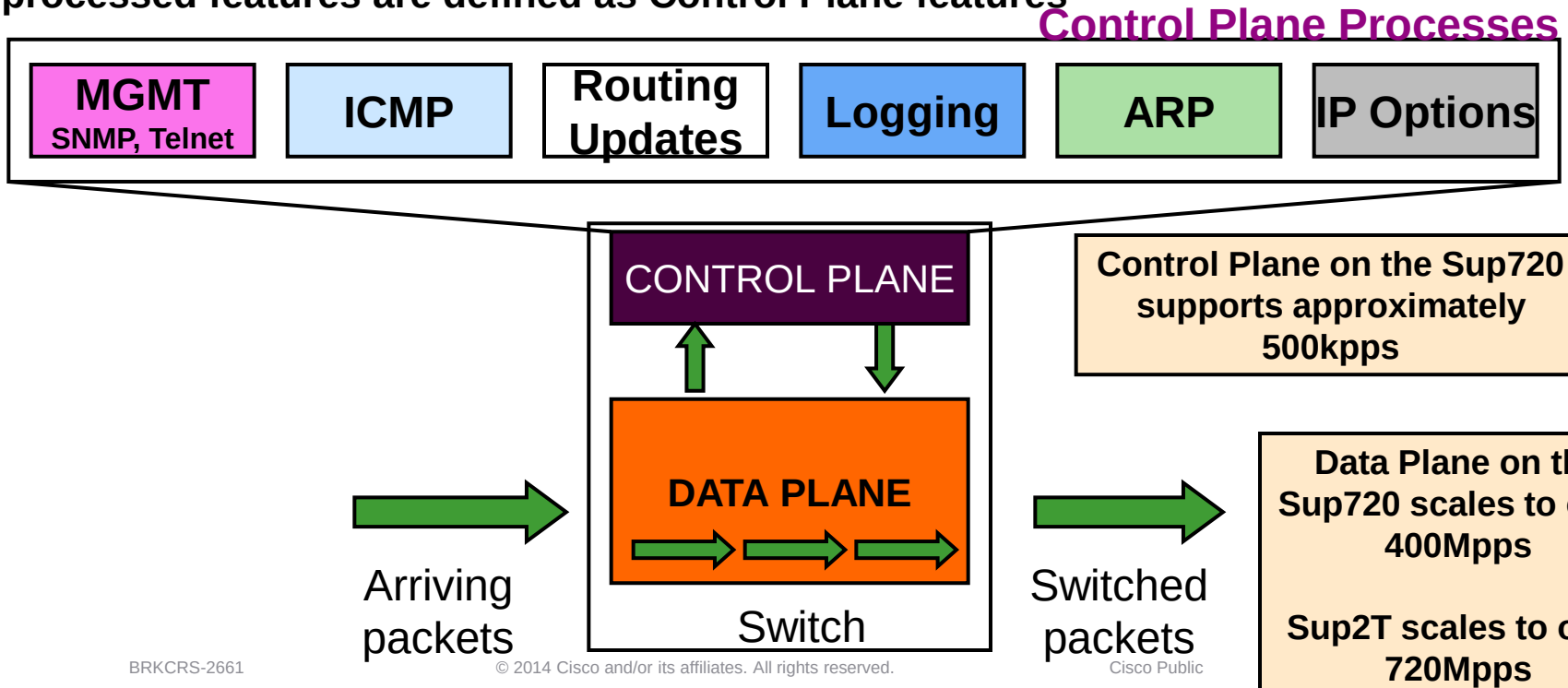

- **Port security** prevents MAC flooding attacks
- **DHCP Snooping** prevents client attack on the switch and server
- **Dynamic ARP Inspection** adds security to ARP using the DHCP snooping table
- **IP Source Guard** adds security to IP source addresses, using the DHCP snooping table

```
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
ip verify source vlandhcp-snooping
!
Interface gigabit1/1
ip dhcp snooping trust
ip arp inspection trust
```

# Control Plane Protection CoPP

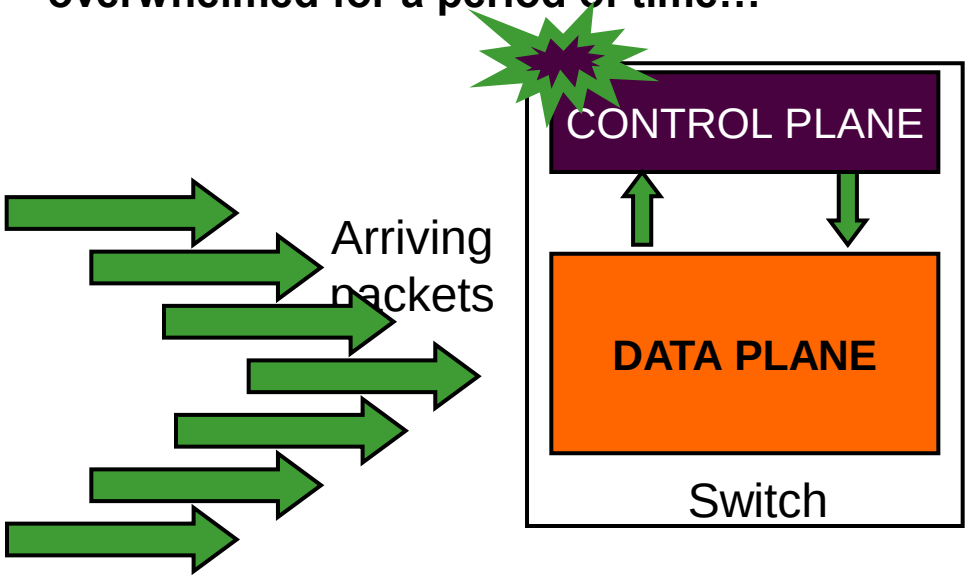
# Catalyst 6500 Control Plane Performance

**All functions within the switch are either performed in hardware or by software. Hardware processed features are defined as Data Plane features, while Software processed features are defined as Control Plane features**



# Catalyst 6500 Control Plane Oversubscription

**Control Plane features are processed by the switch CPU so there is a limited amount of processing power available for these tasks. If that CPU is swamped, all other processes stop. This can have a bad impact on the switch if the CPU is overwhelmed for a period of time…**




## Result of CPU overload

- Dropping Routing Neighbours
- Failure to send Route Updates
- Failure to send STP Updates
- Failure to keep up with Logging requests
- No ARP's processed
- CLI locks up
- Switch locking up
- and more…

# REFERENCE : Example Protocols and Services Processed in Software

| Control Plane Protocols | Control Plane Packet Forwarding |
|---|---|
| UDLD Protocol | IP Options |
| PagP Protocol | Fragmentation |
| LACP Protocol | Select Tunnel Options |
| SNMP Protocol | ICMP Packets |
| Syslog Export | MTU failure |
| Netflow & Netflow Data Export | TTL=1 or TTL=0 |
| Address Resolution Protocol | Packets with Checksum error or error length |
| HSRP, VRRP, GLBP | RPF Check |
| Cisco Discovery Protocol | Packets that require ARP resolution |
| VLAN Trunking Protcol | Non-IP (IPX, Appletalk) |
| Dynamic Trunking Protocol | ACL logging |
| Telnet, IP Sec, SSH | Broadcast traffic denied in RACL |
| BGP, OSPF, EIGRP, RIP, ISIS | Authentication Proxy |
| Web Cache Control Protocol | PBR traffic for certain “match” or “set” arguments |

# Control Plane Protection on Catalyst Switches

| | Catalyst 6500 / 6800 Series | Catalyst 4500 Series | Catalyst 3850 series |
|---|---|---|---|
| Hardware Rate-Limiters | Yes | No | No |
| Modular QoS Class Maps | Yes | Yes | Yes |

# SUP2T Control Plane Protection

## Hardware Rate Limiters Support

| Unicast Rate Limiters | |
|---|---|
| CEF Receive | Traffic Destined to the Router |
| **CEF Receive Secondary** | **Traffic destined to an IP address terminated on the C6500** |
| CEF Glean | Traffic requiring ARP |
| CEF No Route | Packets with Not Route in the FIB |
| ICMP Redirect | Packets that Require ICMP Redirects |
| IP Errors | Packet with IP Checksum or Length Errors |
| ICMP No Route | ICMP Unreachables for Unroutable Packets |
| ICMP ACL Drop | ICMP Uncreachables for Admin Deny Packets |
| RPF Failure | Packets that Fail uRPF Check |
| L3 Security | CBAC, Auth-Proxy, and IPSEC Traffic |
| **ACL Bridged In** | NAT, TCP Int, Reflexive ACLs, Log on ACLs |
| **ACL Bridged Out** | NAT, TCP Int, Reflexive ACLs, Log on ACLs |
| ARP Inspection | Dynamic ARP Inspection Traffic to CPU |
| DHCP Snoop In | DHCP Snooping Traffic to CPU |
| IP Features | Security Features (Auth-Proxy,IP Sec, others) |
| **UCAST UNKNOWN FLOOD** | L2 unknown unicast traffic |
| VACL Logging | CLI Notification of VACL Denied Packets |
| IP Options | Unicast Traffic with IP Options Set |
| Capture | Used with Optimised ACL Logging |

| Layer 2 Rate Limiters | |
|---|---|
| LAYER_2 PT | L2PT Encapsulation/Decapsulation |
| LAYER_2 PDU | Layer 2 PDUs |
| **MAC PBF In** | |
| **IP Admis. on L2 Port** | |
| **LAYER_2 PORTSEC** | |
| **LAYER_2 SPAN PCAP** | |

| General Rate Limiters | |
|---|---|
| MTU Failure | Packets Requiring Fragmentation |
| TTL Failure | Packets with TTL<=1 |
| Capture Pkt | Limits packets punted to the CPU because of Optimised ACL Logging (OAL). |
| **DIAG RESERVED 0** | Reserved |
| **DIAG RESERVED 1** | Reserved |
| **DIAG RESERVED 2** | Reserved |
| **MCAST REPL RESERVED** | Reserved |

# SUP2T Control Plane Protection

## Hardware Rate Limiters Support

| Multicast Rate Limiters | |
|---|---|
| MCAST IPV4 FIB MISS | Packets with No mroute in the FIB |
| MCAST IPv4 IGMP | IGMP Packets |
| **MCAST IPv4 Direct C** | Local Multicast on Connected Interface |
| MCAST IPV4 OPTIONS | Multicast Traffic with IP Options Set |
| **MCAST IPV4 CONTROL PK** | IPv4 Multicast Control Traffic to CPU |
| MCAST IPV6 DIRECTLY C | Packets with No Mroute in the FIB |
| MCAST IPV6 MLD | IPv6 Multicast Control Traffic to CPU |
| **MCAST IPV6 CONTROL PK** | Partial Shortcut Entries |
| **MCAST BRG FLD IP CNTR** | Partial Shortcut Entries |
| **MCAST BRG FLD IP** | Partial Shortcut Entries |
| **MCAST BRG** | Partial Shortcut Entries |
| **MCAST BRG OMF** | Partial Shortcut Entries |
| V6 Route Control | Partial Shortcut Entries |
| V6 Default Route | Multicast Traffic with IP Options Set |
| V6 Second Drop | Mulicast Traffic with IP Options Set |

These HWRL were all covered as a single HWRL in the PFC3

Some new HWRL are added and also more granularity to existing HWRLs is provided in the PFC4 versus PFC3

# Control Plane Protection Comparison Sup720 and Sup2T

| Feature/Capability | Sup720 | Sup2T |
|---|---|---|
| Uses Modular QoS CLI for applying policies to the Control Plane Interface | Yes | Yes |
| Special Case Rate-Limiters for non-IP traffic types (HWRL) | Yes | Yes |
| Max number of HWRLs supported | 8 (L3) 4 (L2) | 31 (L3) 26 (L2) |
| Distributed policing: synchronisation of hardware policers on different linecards | No | Yes |
| HWRL measured in Packets and Bits per second | No, PPS only | Yes, PPS and BPS |
| Traffic Counters | CoPP only | Yes, both HWRL and CoPP |
| Traffic Exceptions configurable in CoPP (allow CoPP based counters and monitoring) | No | Yes |
| Full suite of Multicast HWRL | No | Yes |
| Microflow Policers via CoPP | No | Yes |
| Ability to leak the first packet above the rate (allows packet capture) | No | Yes |
| CoPP Polices to match ARP and RARP traffic | No | Yes |

# Catalyst 3850 Built in Rate Limiters

```
C3850#show platform qos queue stats internal cpu policer

For Asic 0
 Queue                                       Drop
----------------------------------------------
DOT1X Auth                                   0
L2 Control                                   0
Forus traffic                                0
ICMP GEN                                     0
Routing Control                              0
Forus Address resolution                     0
ICMP Redirect                                0
WLESS PRI-5                                  0
WLESS PRI-1                                  0
WLESS PRI-2                                  0
WLESS PRI-3                                  0
WLESS PRI-4                                  0
BROADCAST                                    0
Learning cache ovfl                          0
Sw forwarding                                0
Topology Control                             0
Proto Snooping                               0
BFD Low Latency                              0
Transit Traffic                              0

RPF Failed                                   0
MCAST END STATION                            0
LOGGING                                      0
Health check                                 0
Crypto Control                               0
Exception                                    0
General Punt                                 0
NFL SAMPLED DATA                             0
SGT Cache Full                               0
EGR Exception                                0
Show frwd                                    0
MCAST Data                                   0
Gold Pkt                                     0

C3850#
```

# Control Plane Protection on Catalyst Switches

| | Catalyst 6500 / 6800 Series | Catalyst 4500 Series | Catalyst 3850 series |
|---|---|---|---|
| Hardware Rate-Limiters | Yes | No | No |
| Modular QoS Class Maps | Yes | Yes | Yes |

# Sup2T CoPP Default Class-maps

| Default Class-Maps Supported with Sup2T |
| --- |
| class-map: class-copp-icmp-redirect-unreachable (match-all) |
| class-map: class-copp-ucast-rpf-fail (match-all) |
| class-map: class-copp-vacl-log (match-all) |
| class-map: class-copp-mcast-punt (match-all) |
| class-map: class-copp-mcast-copy (match-all) |
| class-map: class-copp-ip-connected (match-all) |
| class-map: class-copp-ipv6-connected (match-all) |
| class-map: class-copp-match-pim-data (match-any) |
| class-map: class-copp-match-pimv6-data (match-any) |
| class-map: class-copp-match-mld (match-any) |
| class-map: class-copp-match-igmp (match-any) |
| class-map: class-copp-match-ndv6 (match-any) |

# SUP2T Control Plane Protection

## Default Class-maps for Control Plane Protection

Using the class-maps allows for better visibility using show commands

```
Router#show policy-map control-plane input class class-copp-options

 Control Plane Interface

  Service-policy input: policy-default-autocopp

  Hardware Counters:

    class-map: class-copp-options (match-all)
      Match: any
      police :
        100 pps 24 limit 24 extended limit
      Earl in slot 1 :
        0 packets
        5 minute offered rate 0 pps
        aggregate-forwarded 0 packets
                          action: transmit
        exceeded 0 packets action: drop
        aggregate-forward 0 pps exceed 0 pps
      Earl in slot 3 :
        997 packets
        5 minute offered rate 196 pps
        aggregate-forwarded 198 packets
                          action: transmit
        exceeded 798 packets action: dr
        aggregate-forward 99 pps excee
      Earl in slot 5 :
        0 packets
        5 minute offered rate 0 pps
        aggregate-forwarded 0 packets
                          action: transmit
        exceeded 0 packets action: drop
        aggregate-forward 0 pps exceed 0 pps
```

Default Class map "class-copp-options" classifies traffic with IP Options set

Traffic counters matching the class-map on the DFC in slot3

# Catalyst Control Plane Protection
## CoPP Deployment—Step 1

- Step 1: Identify traffic of interest and classify it into multiple traffic classes:
  - BGP
  - IGP (EIGRP, OSPF, ISIS)
  - Management (telnet, TACACS, ssh, SNMP, NTP)
  - Reporting (SAA)
  - Monitoring (ICMP)
  - Critical applications (HSRP, DHCP)
  - Undesirable
  - Default

```
ip access-list extended coppacl-bgp
permit tcp host 192.168.1.1 host 10.1.1.1 eq bgp
permit tcp host 192.168.1.1 eq bgp host 10.1.1.1
!
ip access-list extended coppacl-igp
permit ospf any host 224.0.0.5
permit ospf any host 224.0.0.6
permit ospf any any
!
ip access-list extended coppacl-management
permit tcp host 10.2.1.1 host 10.1.1.1
established
permit tcp 10.2.1.0 0.0.0.255 host 10.1.1.1 eq 22
permit tcp 10.86.183.0 0.0.0.255 any eq telnet
permit udp host 10.2.2.2 host 10.1.1.1 eq snmp
permit udp host 10.2.2.3 host 10.1.1.1 eq ntp
!
```

# Catalyst 6500 Control Plane Protection

## CoPP Deployment—Step 2

- Step 2: Associate the identified traffic with a class, and permit the traffic in each class
  - Must enable QoS globally(Sup720 only), else CoPP will not be applied in hardware
  - Always apply a policing action for each class since the switch will ignore a class that does not have a corresponding policing action (for example "police 31500000 conform-action transmit exceed-action drop"). Alternatively, both conform-action and exceed-action could be set to transmit, but doing so will allocate a default policer as opposed to a dedicated policer with its own hardware counters.
  - HW CoPP classes are limited to one match

```
class-map match-all copp-bgp
 match access-group name coppacl-bgp
class-map match-all copp-igp
 match access-group name coppacl-igp
class-map match-all copp-management
 match access-group name coppacl-management
class-map match-all copp-reporting
 match access-group name coppacl-reporting
class-map match-all copp-monitoring
 match access-group name coppacl-monitoring
class-map match-all copp-critical-app
 match access-group name coppacl-critical-app
class-map match-all copp-undesirable
 match access-group name coppacl-undesirable
```

```
policy-map copp-policy
 class copp-bgp
  police 30000000 conform-action transmit exceed-action drop
 class copp-igp
  police 30000000 conform-action transmit exceed-action drop
 class copp-management
  police 30000000 conform-action transmit exceed-action drop
 class copp-reporting
  police 30000000 conform-action transmit exceed-action drop
 class copp-monitoring
  police 30000000 conform-action transmit exceed-action drop
 class copp-critical-app
  police 30000000 conform-action transmit exceed-action drop
 class copp-undesirable
  police 30000000 conform-action transmit exceed-action drop
 class class-default
                                                    tion drop
```

```
control-plane
 service-policy input copp-policy
```

# Catalyst 6500 Control Plane Protection

## CoPP Deployment—Step 3

- Step 3: Adjust classification, and apply liberal CoPP policies for each class of traffic
  - show policy-map control-plane displays dynamic information for monitoring control plane policy. Statistics include rate information and number of packets/ bytes confirmed or exceeding each traffic class
  - CoPP rates on Sup720 are bps—pps is not possible. However, HWRL rates are in pps

```
Switch# show policy-map control-plane
 Control Plane Interface
 Service-policy input: copp-policy
<snip>
  Hardware Counters:
  class-map: copp-monitoring (match-all)
   Match: access-group name coppacl-monitoring
   police :
    30000000 bps 937000 limit 937000 extended limit
   Earl in slot 5 :
    0 bytes
    5 minute offered rate 0 bps
    aggregate-forwarded 0 bytes action: transmit
    exceeded 0 bytes action: drop
    aggregate-forward 0 bps exceed 0 bps
   Earl in slot 7 :
    112512 bytes
    5 minute offered rate 3056 bps
    aggregate-forwarded 112512 bytes action: transmit
    exceeded 0 bytes action: drop
    aggregate-forward 90008 bps exceed 0 bps

==========
```

# Catalyst 6500 Control Plane Protection
## CoPP Deployment—Step 3 (Cont.)

- Step 3: Adjust Classification, and Apply liberal CoPP policies for each class of traffic
  - show ip access-lists provides packet count statistics per ACE. Absence of any hits on an entry indicate lack of traffic matching the ACE criteria—the rule might be rewritten
  - Hardware ACL hit counters are available in PFC3B/BXL for security ACL TCAM only (not QoS ACL TCAM)

```
Switch#sh access-list
Extended IP access list coppacl-bgp
  10 permit tcp host 192.168.1.1 host 10.1.1.1 eq bgp
  20 permit tcp host 192.168.1.1 eq bgp host 10.1.1.1
Extended IP access list coppacl-critical-app
  10 permit ip any host 224.0.0.1
  20 permit udp host 0.0.0.0 host 255.255.255.255 eq bootps
  30 permit udp host 10.2.2.8 eq bootps any eq bootps
Extended IP access list coppacl-igp
  10 permit ospf any host 224.0.0.5 (64062 matches)
  20 permit ospf any host 224.0.0.6
  30 permit ospf any any (17239 matches)
Extended IP access list coppacl-management
  10 permit tcp host 10.2.1.1 host 10.1.1.1 established
  20 permit tcp 10.2.1.0 0.0.0.255 host 10.1.1.1 eq 22
  30 permit tcp 10.86.183.0 0.0.0.255 any eq telnet
  40 permit udp host 10.2.2.2 host 10.1.1.1 eq snmp
  50 permit udp host 10.2.2.3 host 10.1.1.1 eq ntp
Extended IP access list coppacl-monitoring
  10 permit icmp any any ttl-exceeded (120 matches)
  20 permit icmp any any port-unreachable
  30 permit icmp any any echo-reply (17273 matches)
  40 permit icmp any any echo (5 matches)
Extended IP access list coppacl-reporting
  10 permit icmp host 10.2.2.4 host 10.1.1.1 echo
Extended IP access list coppacl-undesirable
  10 permit udp any any eq 1434
```

# Catalyst 6500 Control Plane Protection
## CoPP Deployment—Step 4

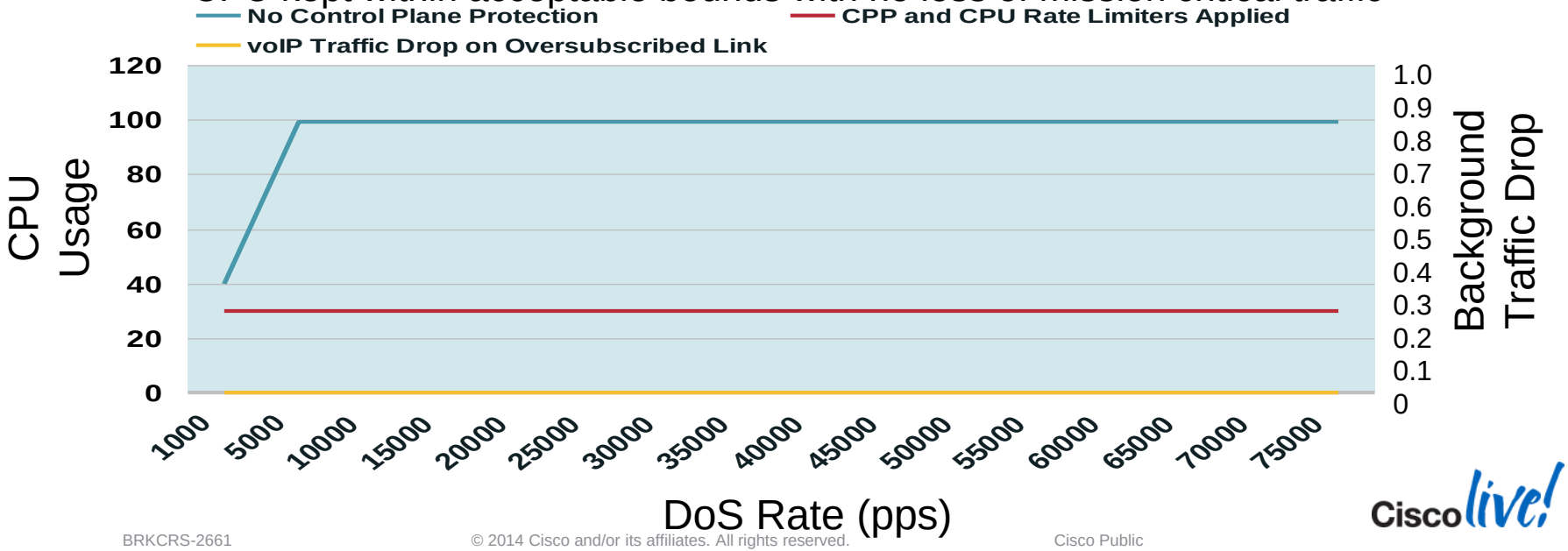
- Step 4: Fine tune the control plane policy
  - Narrow the ACL permit s
    addresses and dependin
    - Routing protocol traffic—n
    - Management traffic—cons
    - Reporting traffic—conserva
    - Monitoring traffic—conserv
    - Critical traffic—conservative rate limit
    - Default traffic—low rate limit
    - Undesirable traffic—drop

```
policy-map copp-policy
 class coppclass-bgp
  police 15000000 conform-action transmit exceed-action drop
 class coppclass-igp
  police 15000000 conform-action transmit exceed-action drop
 class coppclass-management
  police 2560000 conform-action transmit exceed-action drop
 class coppclass-reporting
  police 1000000 conform-action transmit exceed-action drop
 class coppclass-monitoring
  police 1000000 conform-action transmit exceed-action drop
 class coppclass-critical-app
  police 7500000 conform-action transmit exceed-action drop
 class coppclass-undesirable
  police 32000 conform-action transmit exceed-action drop
 class class-default
  police 1000000 conform-action transmit exceed-action drop
```

http://www.cisco.com/en/US/products/hw/switches/ps708/products_white_paper0900aecd802ca5c8.shtml

# Catalyst 6500 Control Plane Protection Mitigating Attacks with CoPP and CPU RL (example)

- 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



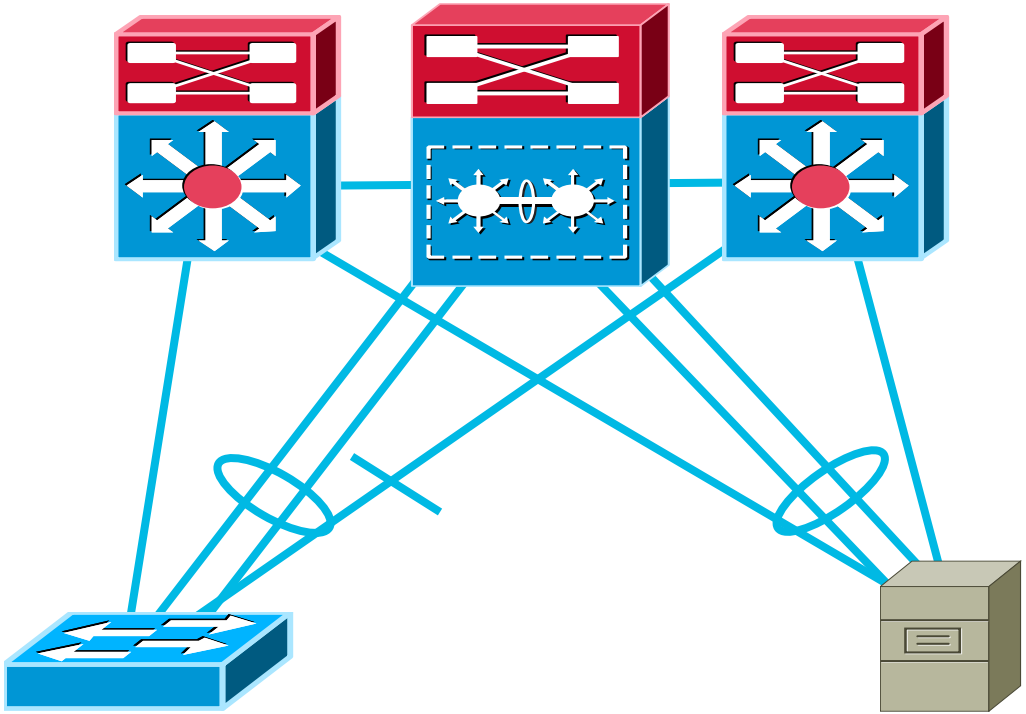



# Alternative Designs

# Virtual Switching System

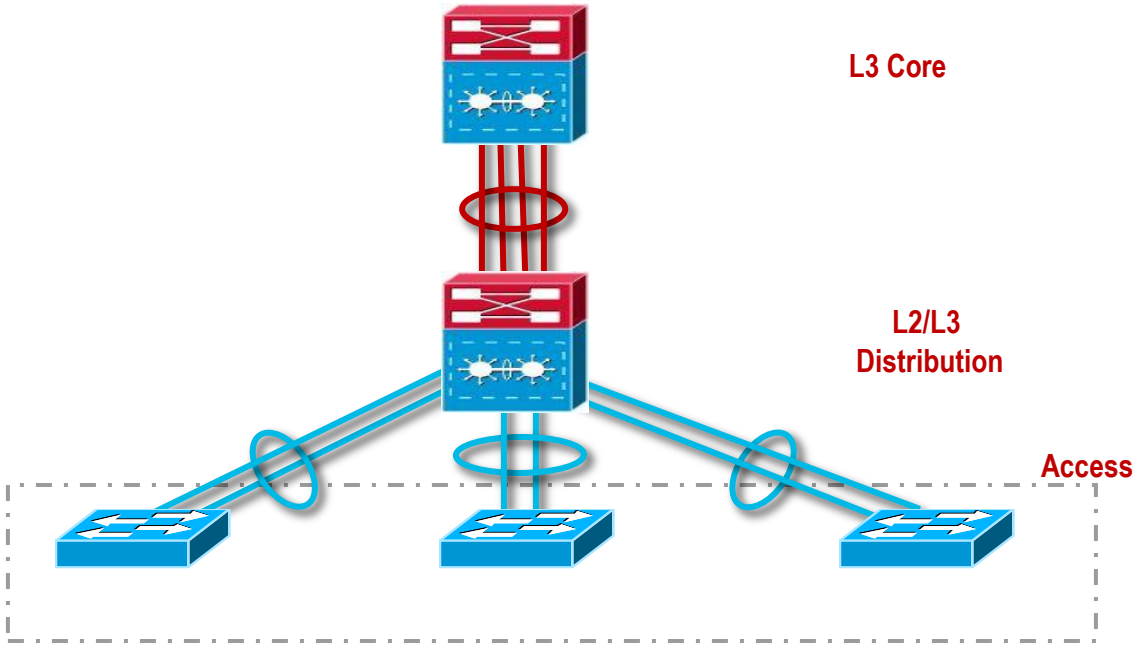
Traditional Design

VSS Design



# Virtual Switching System

## VSS Enterprise Campus



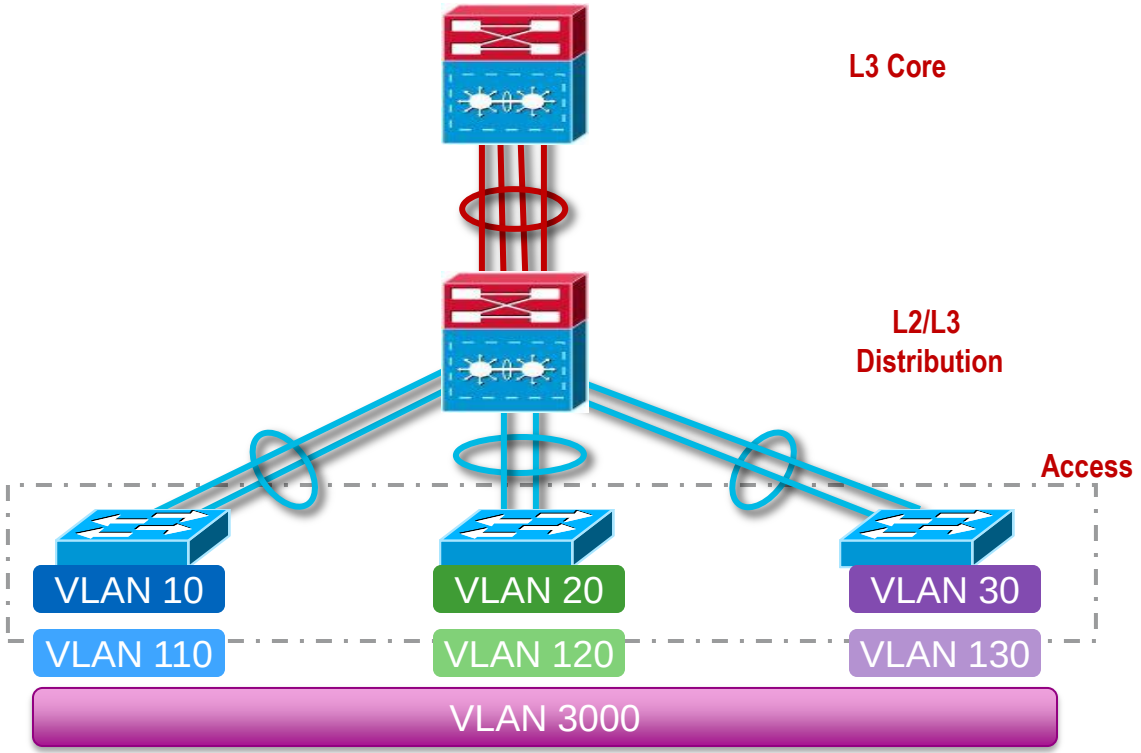

**Reduced routing neighbours, Minimal L3 reconvergence**

**No FHRPs
No Looped topology
Policy Management**

**Multiple active uplinks per VLAN, No STP convergence**

# Virtual Switching System

## VSS Enterprise Campus




**Reduced routing neighbours, Minimal L3 reconvergence**

**No FHRPs**
**No Looped topology**
**Policy Management**

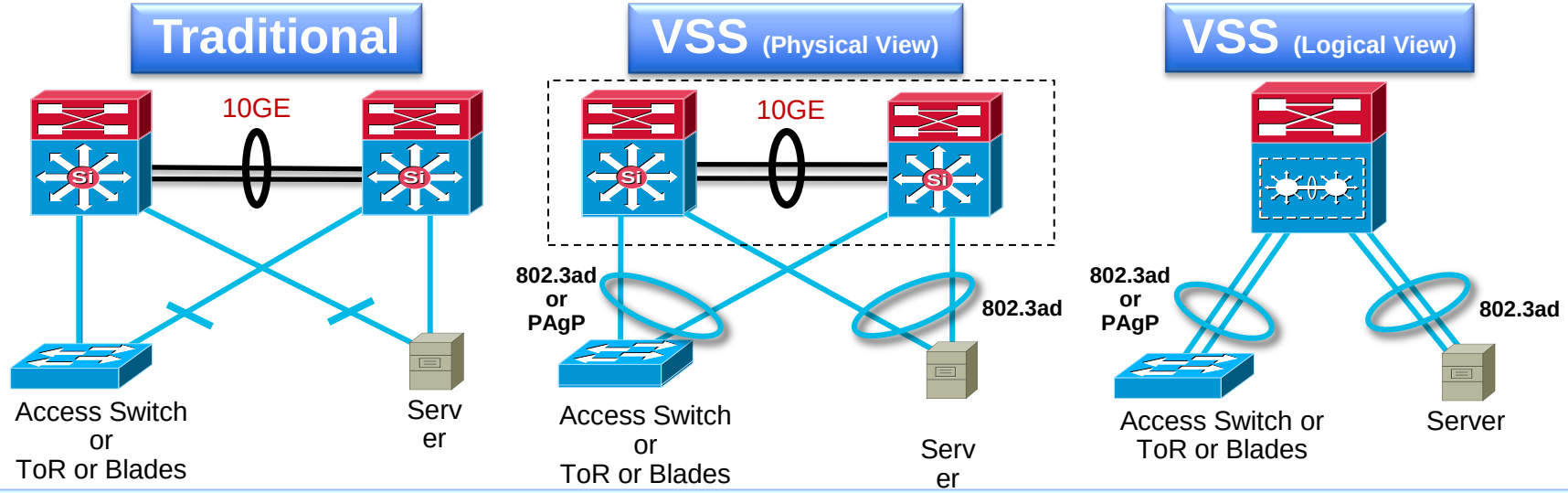
**Multiple active uplinks per VLAN, No STP convergence**

# VSS Simplifies the Configuration

| Standalone Switch 1 (Coordinated Configuration)  | Standalone Switch 2 (Coordinated Configuration)  | VSS (One simplified configuration)  |
|---|---|---|
| **Spanning Tree Configuration** | | |
| ! Enable 802.1d per VLAN spanning tree enhancements.<br>spanning-tree mode pvst<br>spanning-tree loopguard default<br>no spanning-tree optimize bpdu transmission<br>spanning-tree extend system-id<br>spanning-tree uplinkfast<br>spanning-tree backbonefast<br>spanning-tree vlan 2,4,6,8,10 priority 24576! | ! Enable 802.1d per VLAN spanning tree enhancements.<br>spanning-tree mode pvst<br>spanning-tree loopguard default<br>no spanning-tree optimize bpdu transmission<br>spanning-tree extend system-id<br>spanning-tree uplinkfast<br>spanning-tree backbonefast<br>spanning-tree vlan 3,5,7,9,11 priority 24576! | ! Enable 802.1d per VLAN spanning tree enhancements<br>spanning-tree mode rapid-pvst<br>no spanning-tree optimize bpdu transmission<br>spanning-tree extend system-id<br>spanning-tree vlan 2-11 priority 24576 |
| **L3 SVI Configuration (sample for 1 VLAN)** | | |
| ! Define the Layer 3 SVI for each voice and data VLAN<br>interface Vlan4<br>description Data VLAN<br>ip address 10.120.4.3 255.255.255.0<br>no ip redirects<br>no ip unreachables<br>! Reduce PIM query interval to 250 msec<br>ip pim query-interval 250 msec<br>ip pim sparse-mode<br>load-interval 30<br>! Define HSRP default gateway with 250/800 msec hello/hold<br>standby 1 ip 10.120.4.1<br>standby 1 timers msec 250 msec 800<br>! Set preempt delay large enough to allow network to stabilize before HSRP<br>! switches back on power on or link recovery<br>standby 1 preempt delay minimum 180<br>! Enable HSRP authentication<br>standby 1 authentication cisco123 | ! Define the Layer 3 SVI for each voice and data VLAN<br>interface Vlan4<br>description Data VLAN<br>ip address 10.120.4.3 255.255.255.0<br>no ip redirects<br>no ip unreachables<br>! Reduce PIM query interval to 250 msec<br>ip pim query-interval 250 msec<br>ip pim sparse-mode<br>load-interval 30<br>! Define HSRP default gateway with 250/800 msec hello/hold<br>standby 1 ip 10.120.4.1<br>standby 1 timers msec 250 msec 800<br>! Set preempt delay large enough to allow network to stabilize before HSRP<br>! switches back on power on or link recovery<br>standby 1 preempt delay minimum 180<br>! Enable HSRP authentication<br>standby 1 authentication cisco123 | ! Define the Layer 3 SVI for each voice and data VLAN<br>interface Vlan4<br>description Data VLAN<br>ip address 10.120.2.1 255.255.255.0<br>no ip redirects<br>no ip unreachables<br>ip pim sparse-mode<br>load-interval 30 |

# Virtual Switching System

Benefits Summary




**Simplifies** operational Manageability via Single point of Management, Non-loop design, minimise reliance on STP, eliminate FHRP etc

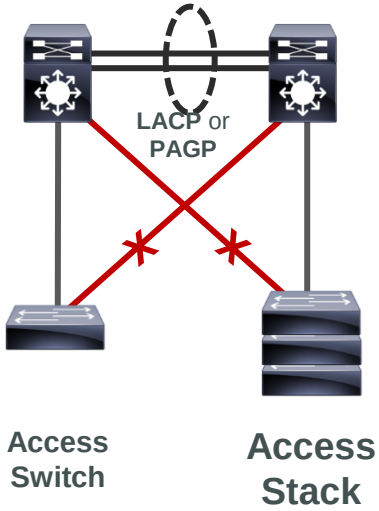
**Scales system capacity** with Active-Active Multi-Chassis Etherchannel (802.3ad/PAgP), no blocking links due to Spanning Tree

**Minimises** traffic disruption from switch or uplink failure with Deterministic subsecond Stateful and Graceful Recovery (SSO/NSF)

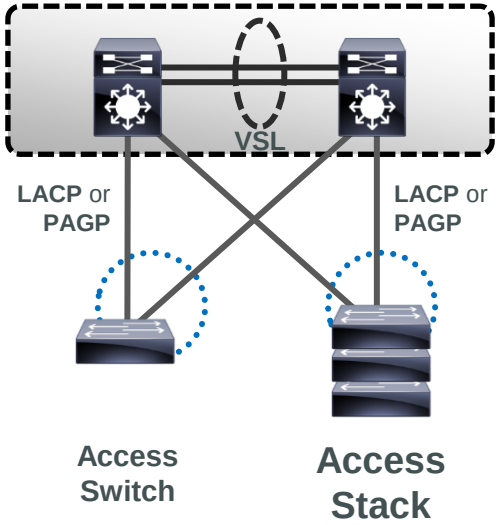
# Catalyst Instant Access

Evolution of the Campus

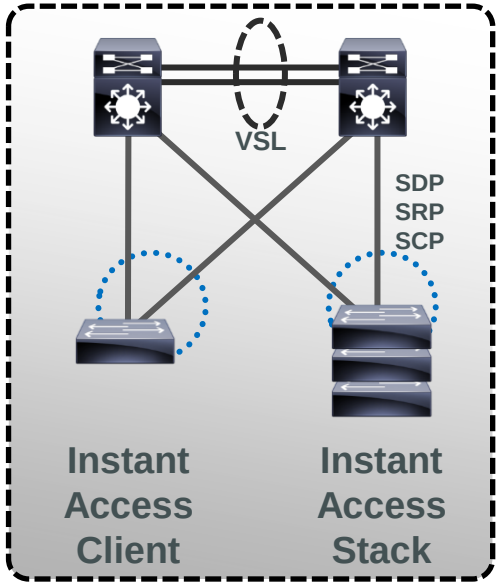
## STANDALONE




## VSS




## INSTANT ACCESS

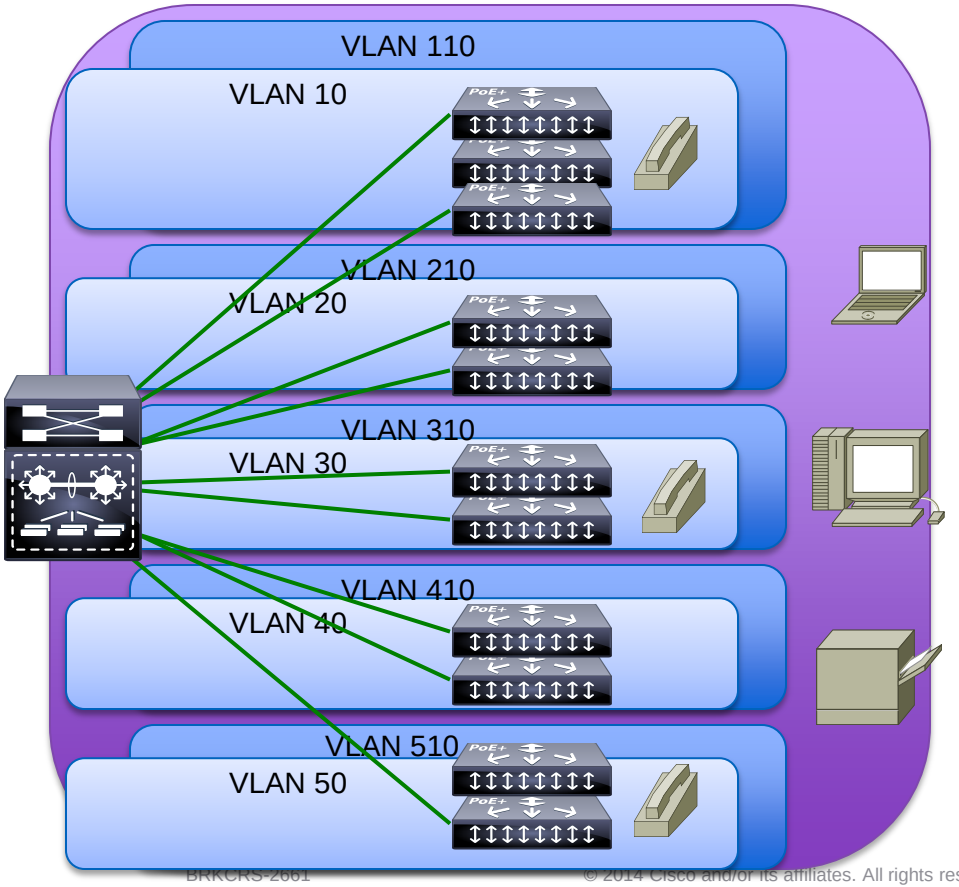


# Case Study #1 – Medium Enterprise – Lower Operations Costs

## Large School District in United States

- Business and Technology Drivers
  - Small operations staff and needs to scale services
  - Spend less time managing the network
  - Many legacy applications requiring L2 connectivity still in use
  - No Cisco certified IT staff onsite
- New building deployment, future  growth planned
  - Instant Access domain size less than 800 ports
- Already using Catalyst 6500 in core, distribution and access in many existing locations

# Instant Access Topology Example




- Five floors with 96 – 144 wired ports per floor
  - 2 X10GbE uplinks per fex
- Instant Access domain size
  - 720 total access ports with PoE
  - RPS2300 for redundant power
- Key applications
  - Third party VoIP
  - Appletalk print services
- Key functionality enabled
  - VLAN bridging for Appletalk
  - Carefully consider L2 domain size whenever extending VLANs across multiple switches




# Summary

# Designing Layer 2 Networks is Easy!

- Limit the size of the L2 domain as much as possible
- Use L1-L2 best practices to harden the network and eliminate common causes of loops
- Implement Spanning Tree tool kit “Make the network fail closed!”
- Use the Integrated Security tool kit to harden the network form malicious or non-malicious network events
- Harden the control plane of the network devices with Control Plane Protection tools
- Consider alternate designs that minimise L2 loops




# Q & A

# Complete Your Online Session Evaluation

## Give us your feedback and receive a Cisco Live 2014 Polo Shirt!

Complete your Overall Event Survey and 5 Session Evaluations.

- Directly from your mobile device on the Cisco Live Mobile App
- By visiting the Cisco Live Mobile Site www.ciscoliveaustralia.com/mobile
- Visit any Cisco Live Internet Station located throughout the venue

Polo Shirts can be collected in the World of Solutions on Friday 21 March 12:00pm - 2:00pm

**Learn online with Cisco Live!**

Visit us online after the conference for full access to session videos and presentations.
www.CiscoLiveAPAC.com


CISCO
TM