

Multilayer Campus Architectures and Design Principles

– BRKCRS-2031

Mark Montañez
Principal Engineer, CCIE #8798
Enterprise Networking Group

Cisco Plus Canada 2012

Title of Slide Would Go Here

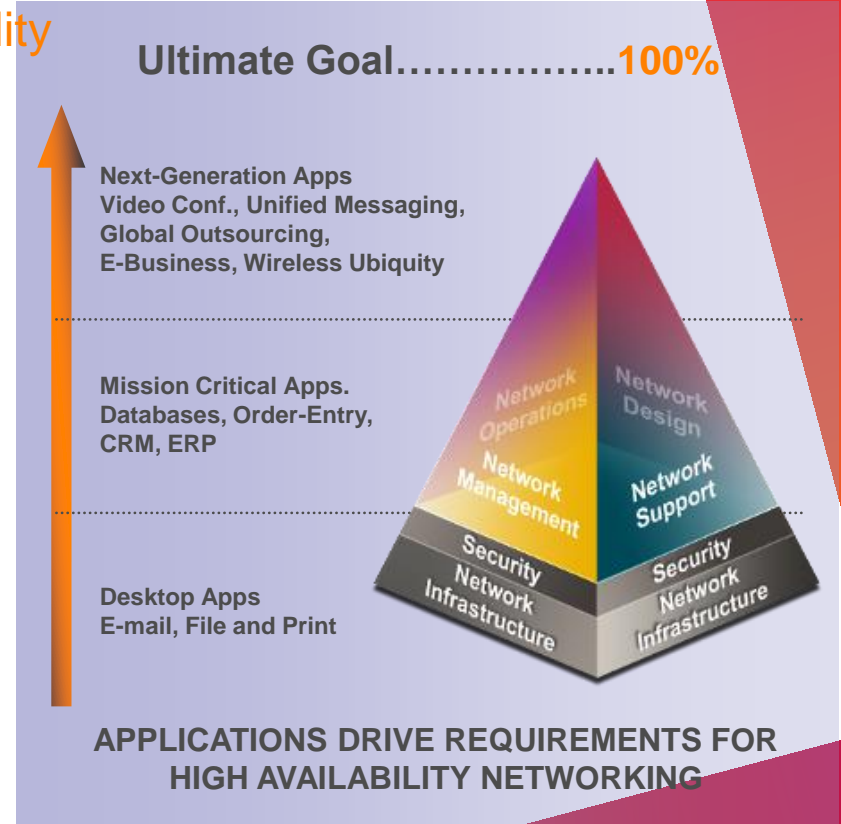
- Bullet Level 1
 - Bullet Leve 2
 - Bullet Level 3
 - Bullet Level 4

Enterprise-Class Availability

Resilient Campus Communication Fabric

Campus Systems Approach to High Availability

- Network-level redundancy
- System-level resiliency
- Enhanced management
- Human ear notices the difference in voice within 150–200 msec
10 consecutive G711 packet loss
- Video loss is even more noticeable
- 200-msec end-to-end campus convergence



Next-Generation Campus Design

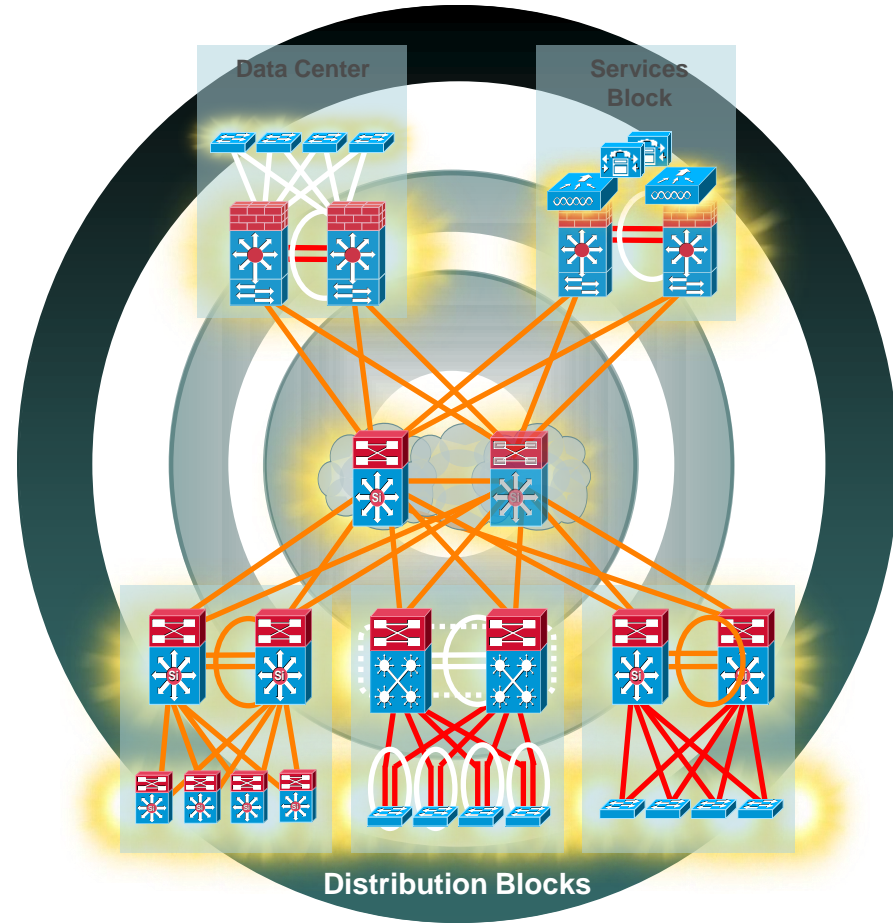
Unified Communications Evolution

- VoIP is now a mainstream technology
- Ongoing evolution to the full spectrum of Unified Communications
- High-definition executive communication application requires stringent Service-Level Agreement (SLA)
 - Reliable service—high availability infrastructure
 - Application service management—QoS

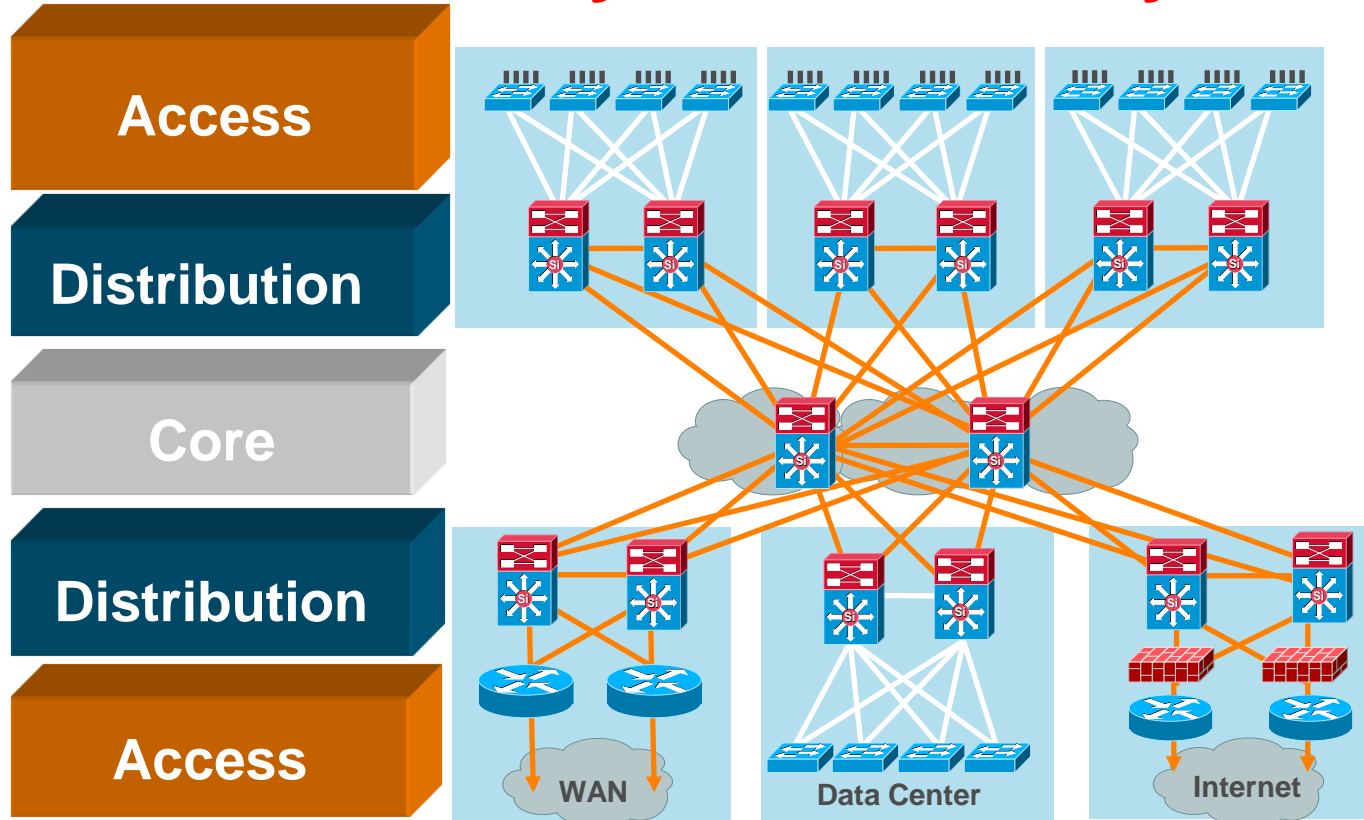


Agenda

- Multilayer Campus Design Principles
- Foundation Services
- Campus Design Best Practices
- IP Telephony Considerations
- QoS Considerations
- Security Considerations
- Putting It All Together
- Summary



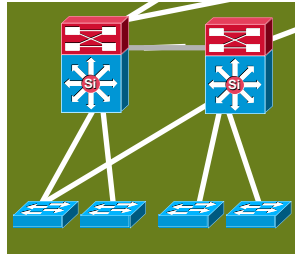
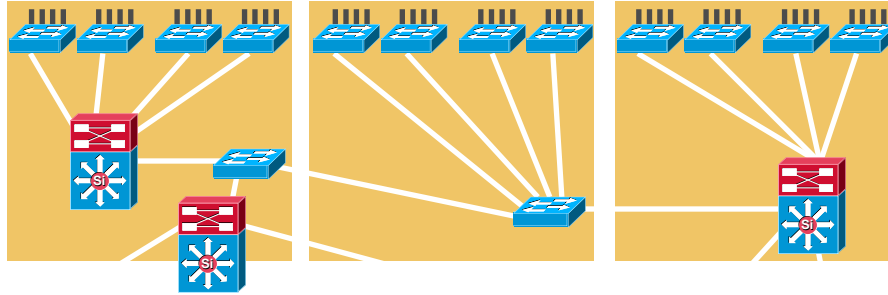
High-Availability Campus Design Structure, Modularity, and Hierarchy



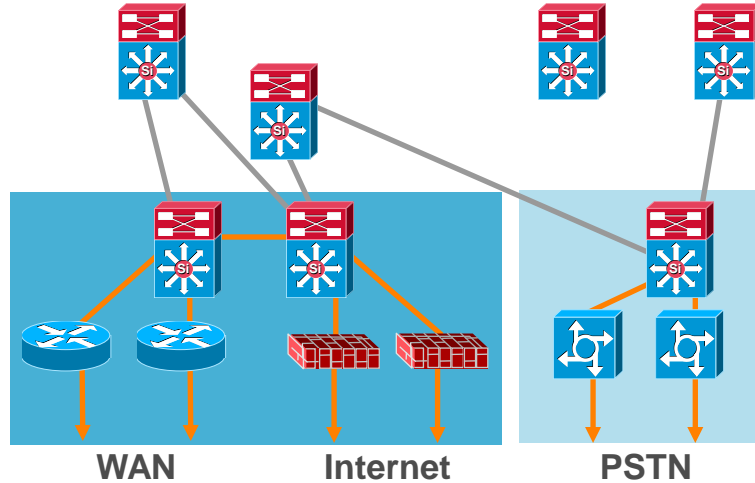
Hierarchical Campus Network

Structure, Modularity and Hierarchy

Not This!!

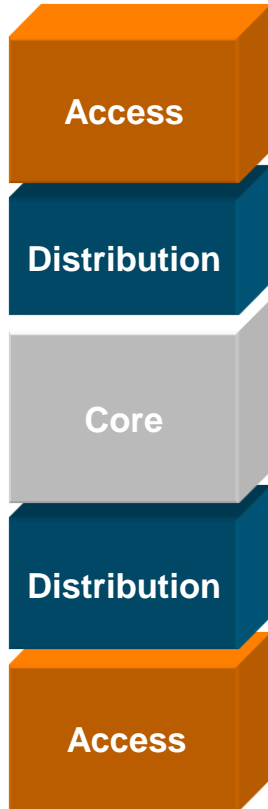


Server Farm

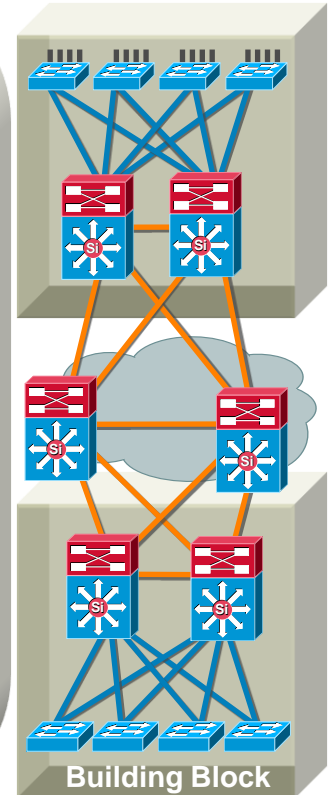


Hierarchical Network Design

Without a Rock Solid Foundation the Rest Doesn't Matter



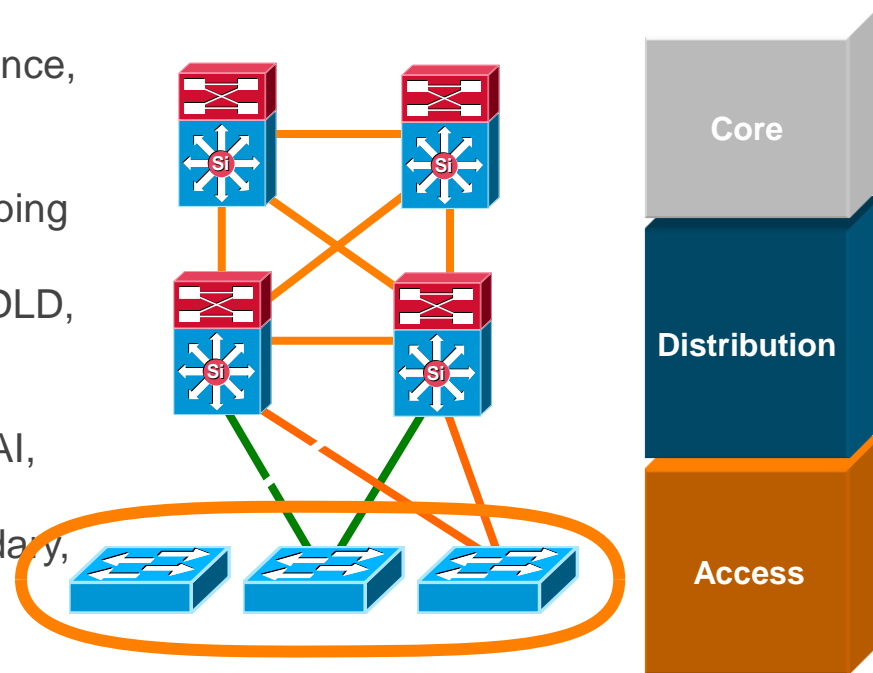
- Offers hierarchy—each layer has specific role
- Modular topology—building blocks
- Easy to grow, understand, and troubleshoot
- Creates small fault domains—clear demarcations and isolation
- Promotes load balancing and redundancy
- Promotes deterministic traffic patterns
- Incorporates balance of both Layer 2 and Layer 3 technology, leveraging the strength of both
- Utilizes Layer 3 routing for load balancing, fast convergence, scalability, and control



Access Layer

Feature Rich Environment

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

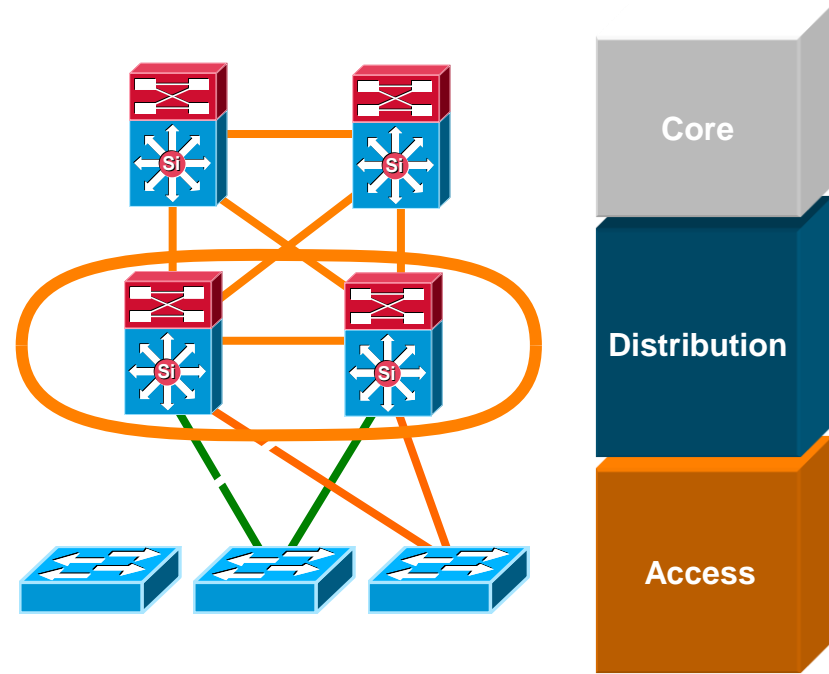


See BRK-CRS 3037—Integrating Intelligent Access

Distribution Layer

Policy, Convergence, QoS, and High Availability

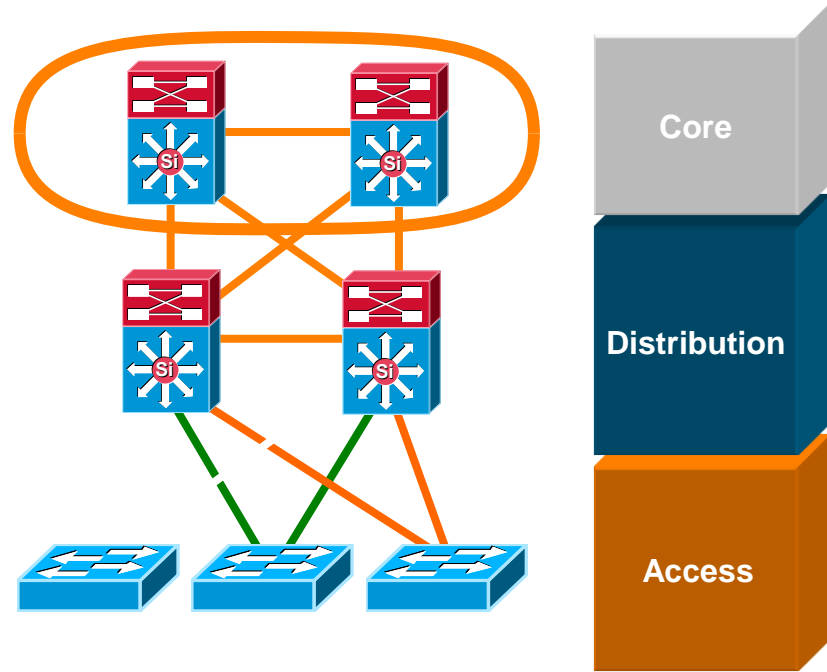
- Availability, load balancing, QoS and provisioning are the important considerations at this layer
- Aggregates wiring closets (access layer) and uplinks to core
- Protects core from high density peering and problems in access layer
- Route summarization, fast convergence, redundant path load sharing
- HSRP or GLBP to provide first hop redundancy



Core Layer

Scalability, High Availability, and Fast Convergence

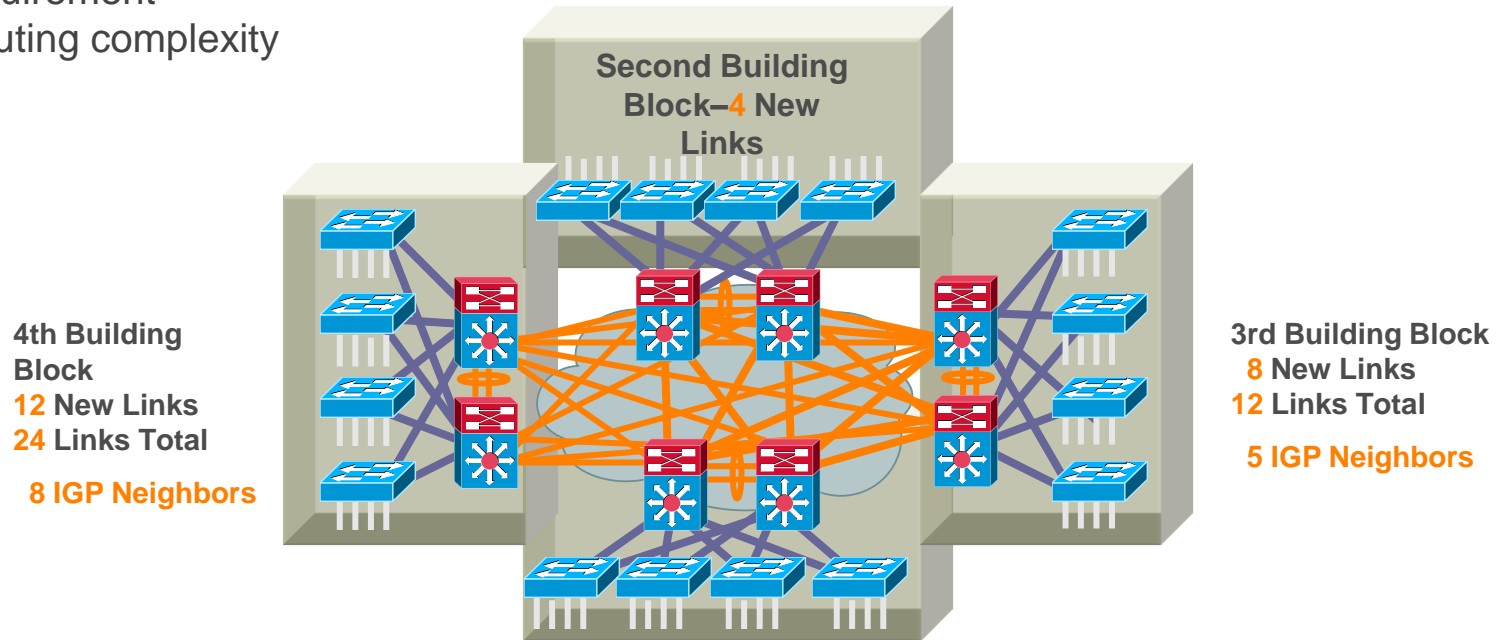
- Backbone for the network— connects network building blocks
- Performance and stability vs. complexity— less is more in the core
- Aggregation point for distribution layer
- Separate core layer helps in scalability during future growth
- Keep the design technology-independent



Do I Need a Core Layer?

It's Really a Question of Scale, Complexity, and Convergence

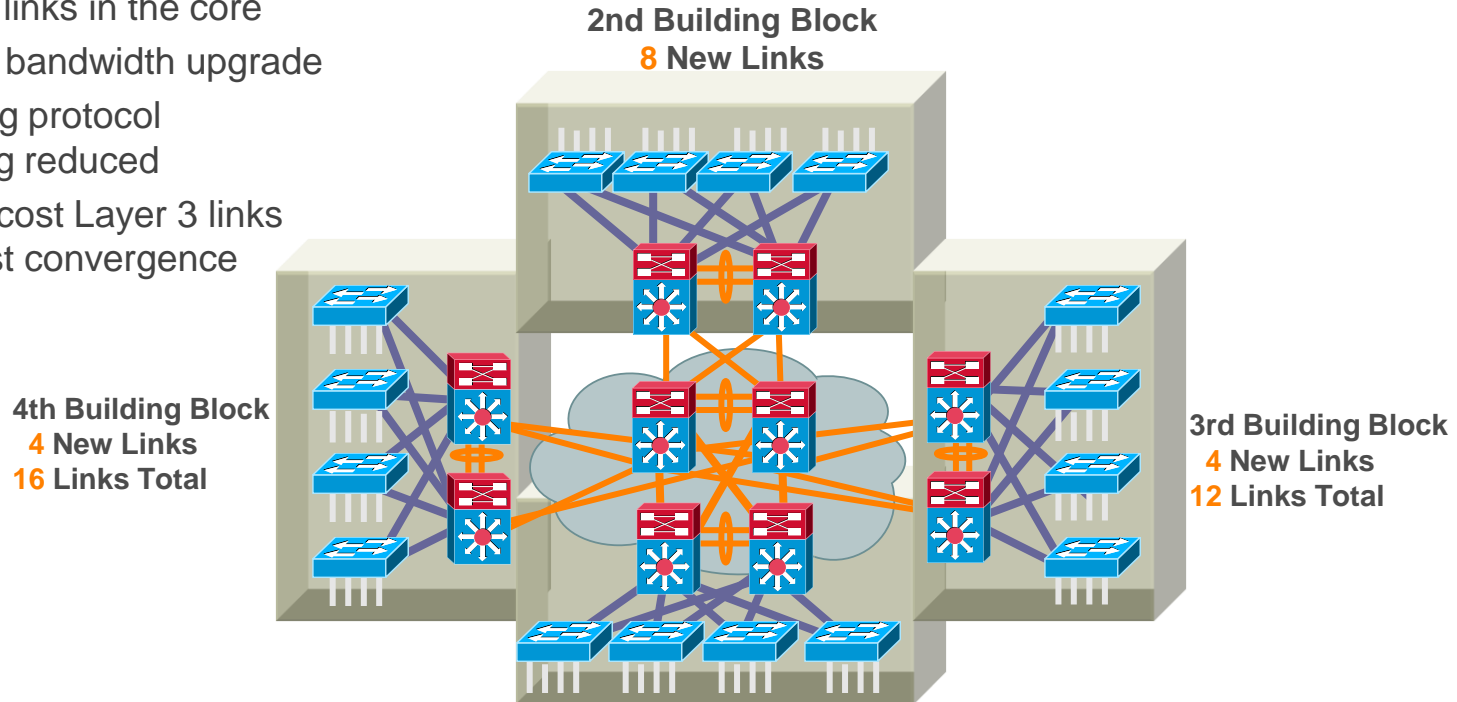
- No Core
- Fully-meshed distribution layers
- Physical cabling requirement
- Routing complexity



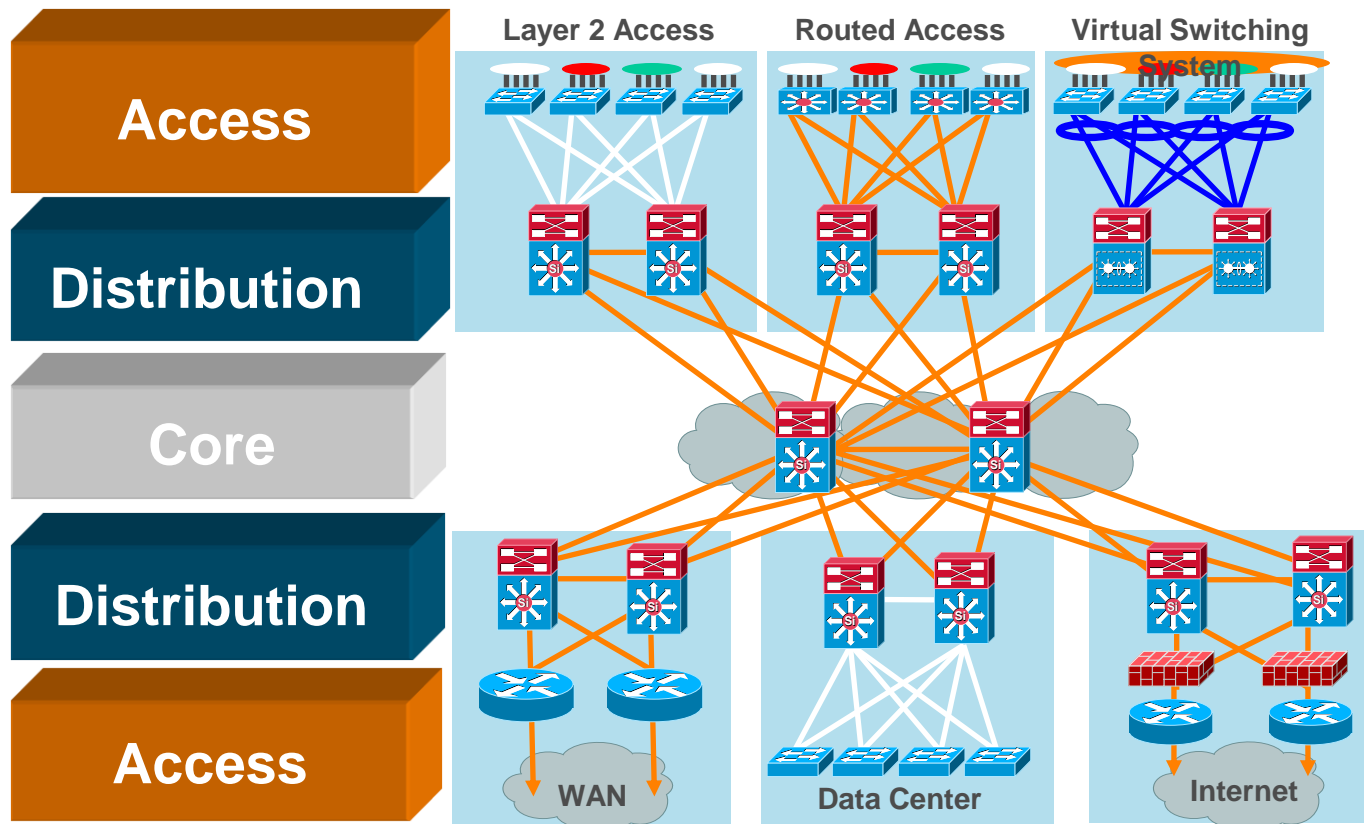
Do I Need a Core Layer?

It's Really a Question of Scale, Complexity, and Convergence

- Dedicated Core Switches
- Easier to add a module
- Fewer links in the core
- Easier bandwidth upgrade
- Routing protocol peering reduced
- Equal cost Layer 3 links for best convergence



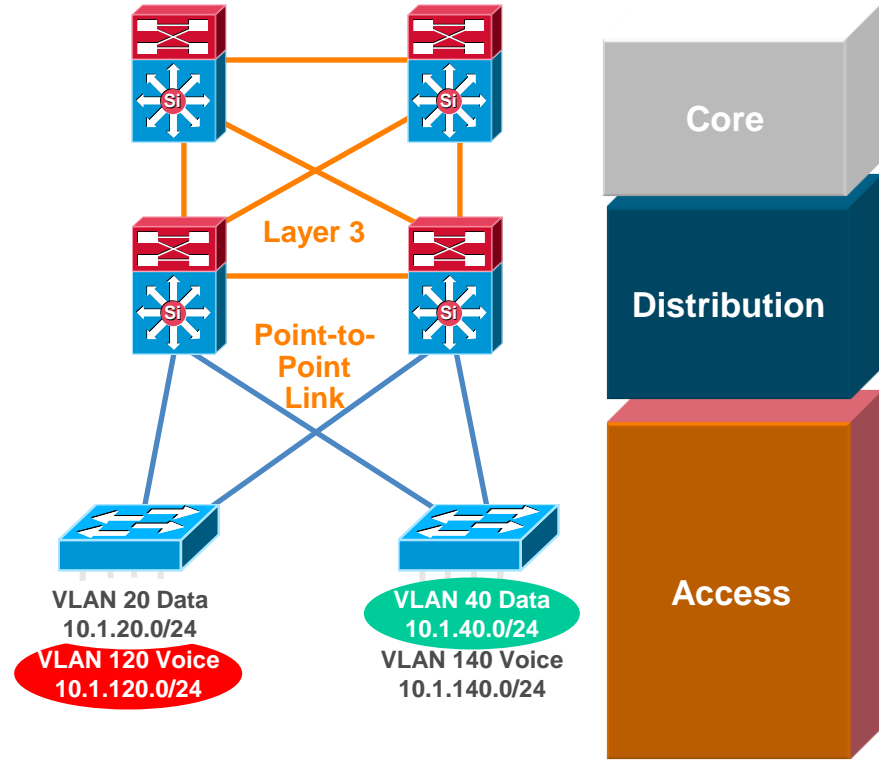
Design Alternatives Come Within a Building (or Distribution) Block



Layer 3 Distribution Interconnection

Layer 2 Access—No VLANs Span Access Layer

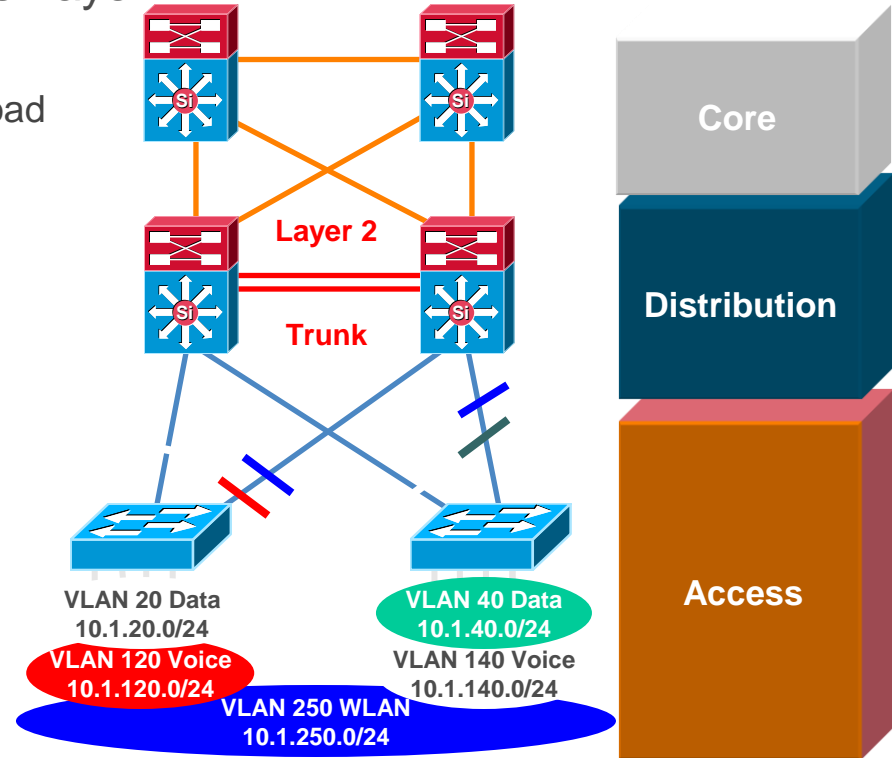
- Tune CEF load balancing
- Match CatOS/IOS EtherChannel settings and tune load balancing
- Summarize routes towards core
- Limit redundant IGP peering
- STP Root and HSRP primary tuning or GLBP to load balance on uplinks
- Set trunk mode on/no-negotiate
- Disable EtherChannel unless needed
- Set port host on access layer ports:
 - Disable trunking
 - Disable EtherChannel
 - Enable PortFast
- RootGuard or BPDU-Guard
- Use security features



Layer 2 Distribution Interconnection

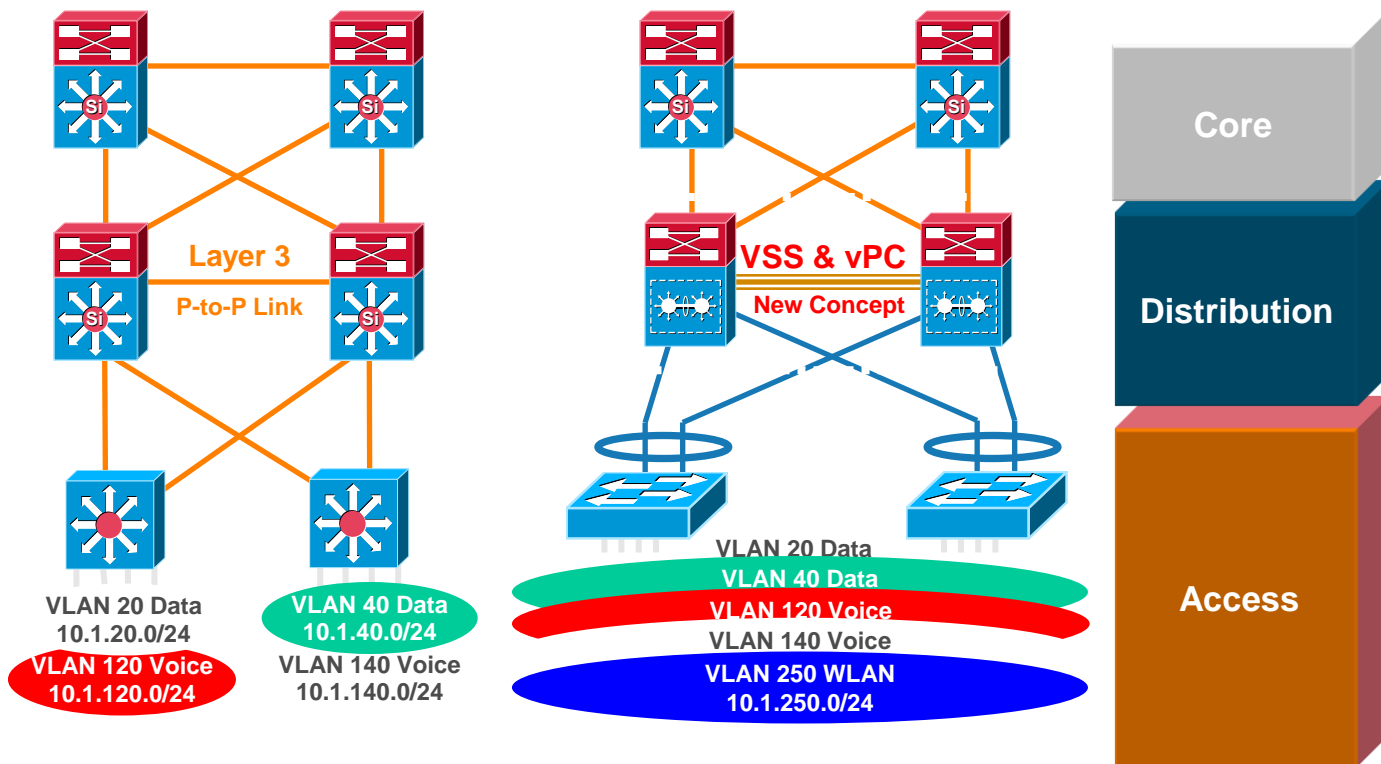
Layer 2 Access—Some VLANs Span Access Layer

- Tune CEF load balancing
- Match CatOS/IOS EtherChannel settings and tune load balancing
- Summarize routes towards core
- Limit redundant IGP peering
- **STP Root and HSRP primary or GLBP and STP port cost tuning to load balance on uplinks**
- Set trunk mode on/no-negotiate
- Disable EtherChannel unless needed
- **RootGuard on downlinks**
- **LoopGuard on uplinks**
- Set port host on access
Layer ports:
 - Disable trunking
 - Disable EtherChannel
 - Enable PortFast
- RootGuard or BPDU-Guard
- Use security features



Routed Access and Virtual Switching System

Evolutions of and Improvements to Existing Designs

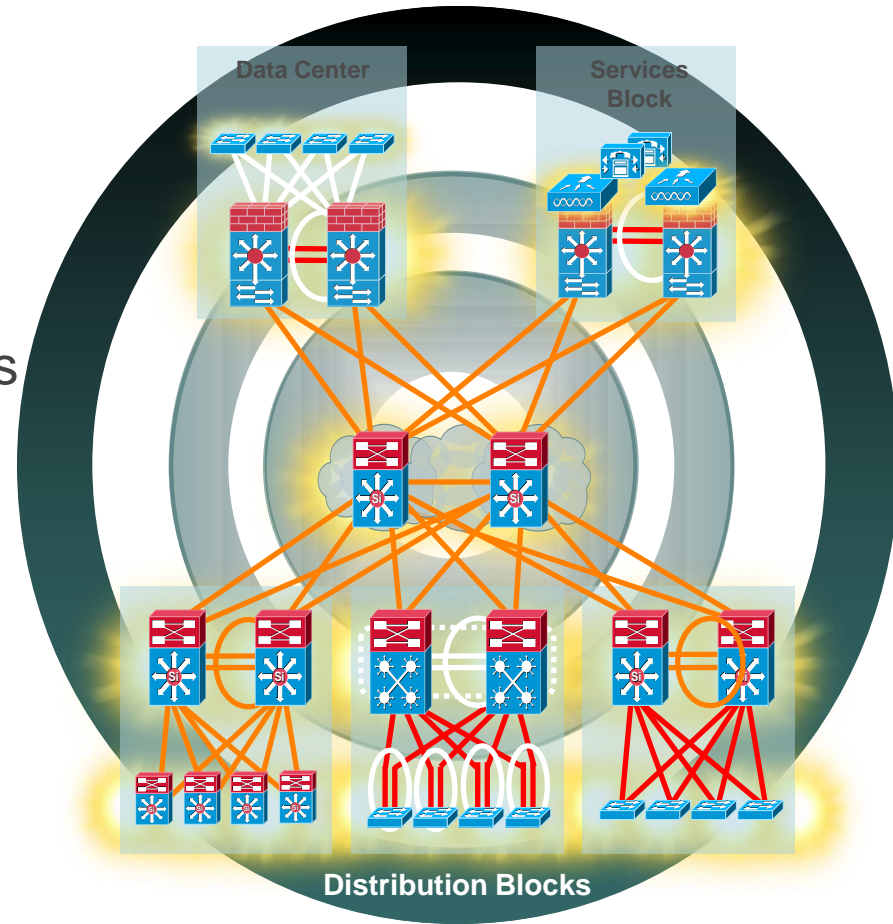


See BRK-CRS3035—Advanced Enterprise Campus Design: VSS

See BRK-CRS3036—Advanced Enterprise Campus Design: Routed Access

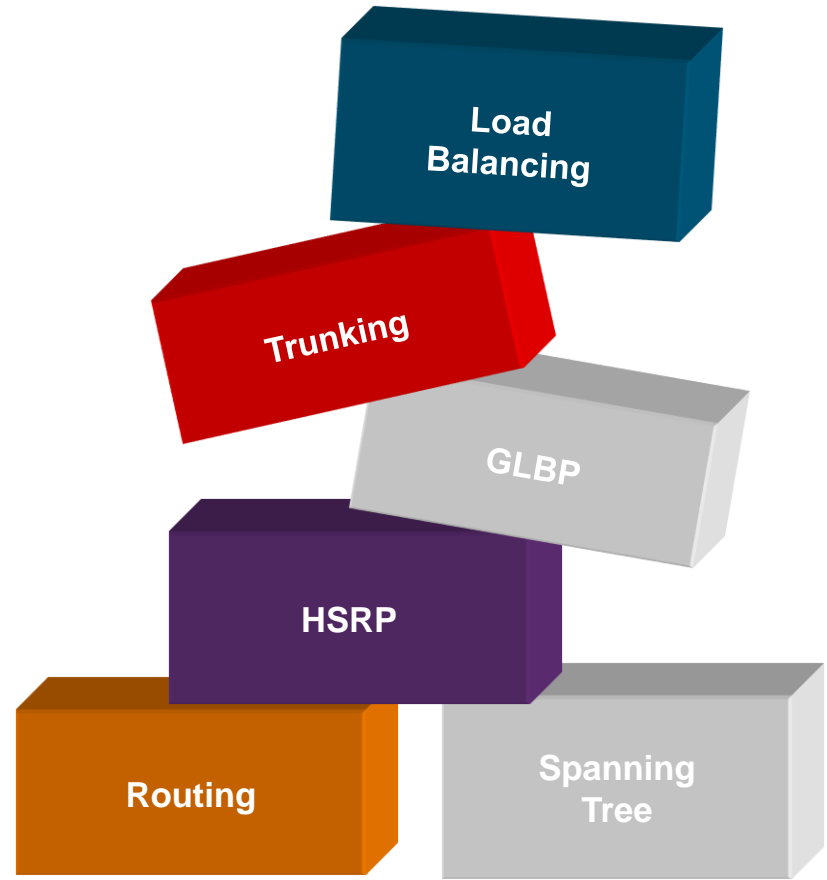
Agenda

- Multilayer Campus Design Principles
- **Foundation Services**
- Campus Design Best Practices
- IP Telephony Considerations
- QoS Considerations
- Security Considerations
- Putting It All Together
- Summary



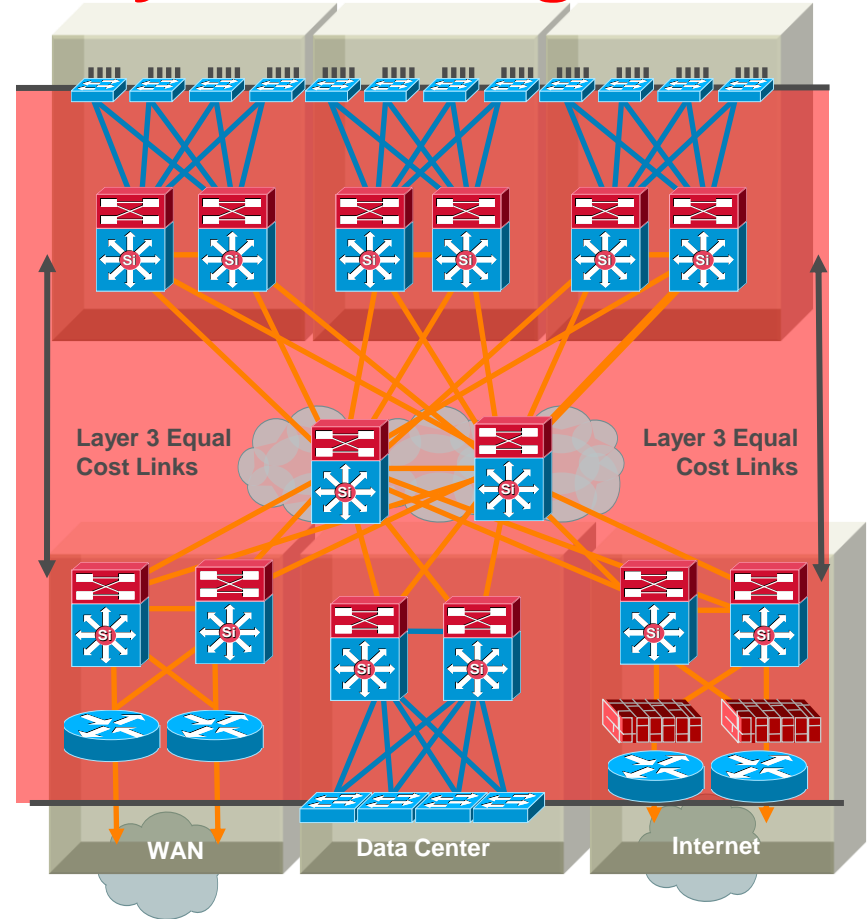
Foundation Services

- Layer 1 physical things
- Layer 2 redundancy—
spanning tree
- Layer 3 routing protocols
- Trunking protocols—(ISL/.1q)
- Unidirectional link detection
- Load balancing
 - EtherChannel link aggregation
 - CEF equal cost load balancing
- First hop redundancy protocols
 - VRRP, HSRP, and GLBP



Best Practices— Layer 1 Physical Things

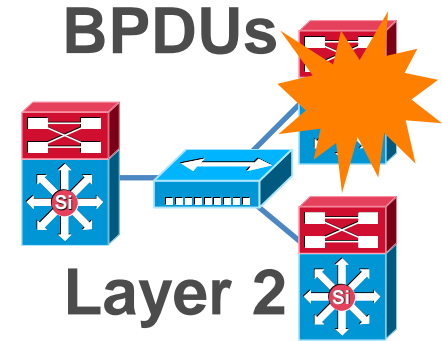
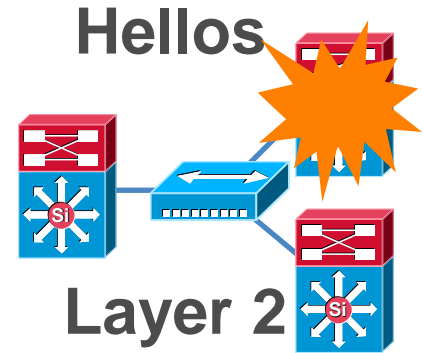
- Use point-to-point interconnections—no L2 aggregation points between nodes
- Use fiber for best convergence (debounce timer)
- Tune carrier delay timer
- Use configuration on the physical interface not VLAN/SVI when possible



Redundancy and Protocol Interaction

Link Neighbour Failure Detection

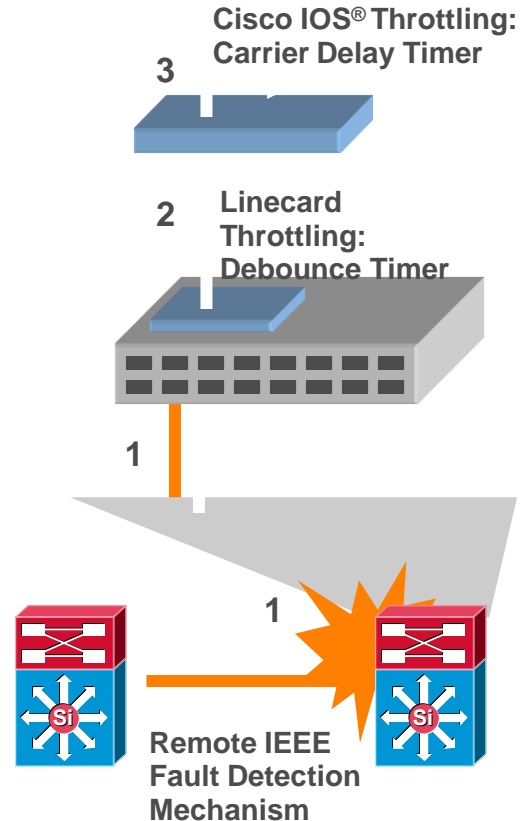
- Indirect link failures are harder to detect
- With no direct HW notification of link loss or topology change convergence times are dependent on SW notification
- Indirect failure events in a bridged environment are detected by spanning tree hellos
- In certain topologies the need for TCN updates or dummy multicast flooding (uplink fast) is necessary for convergence
- You should not be using hubs in a high-availability design



Redundancy and Protocol Interaction

Link Redundancy and Failure Detection

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



Redundancy and Protocol Interaction

Layer 2 and 3—Why Use Routed Interfaces

- Configuring L3 routed interfaces provides for faster convergence than an L2 switch port with an associated L3 SVI



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

~ 8 msec loss

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



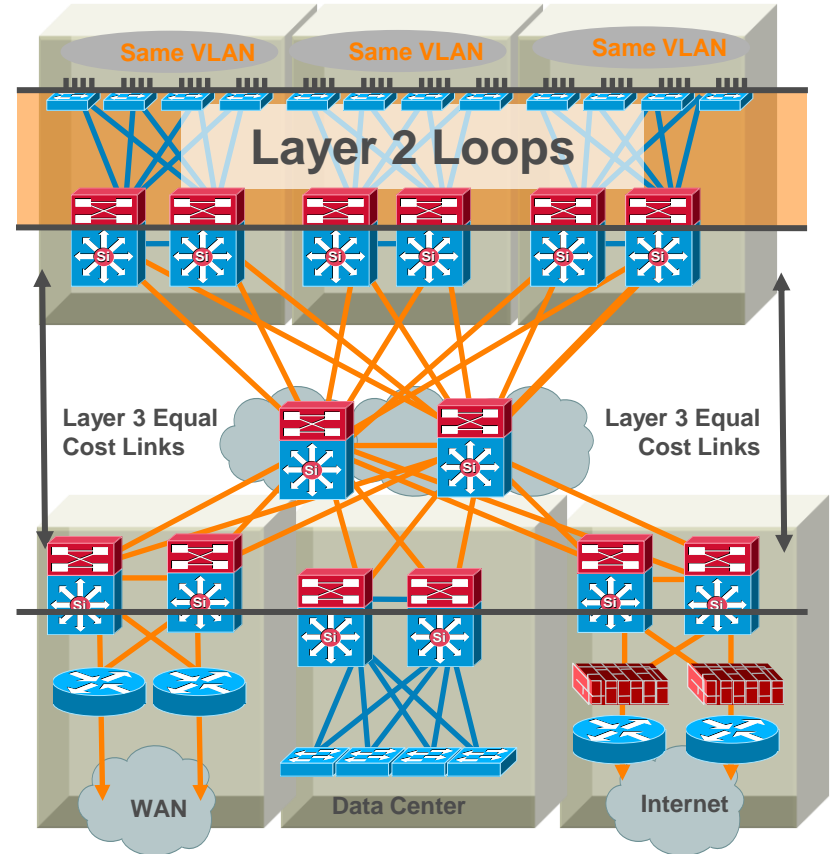
1. Link Down
2. Interface Down
3. Autostate
4. SVI Down
5. Routing Update

~ 150–200 msec loss

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

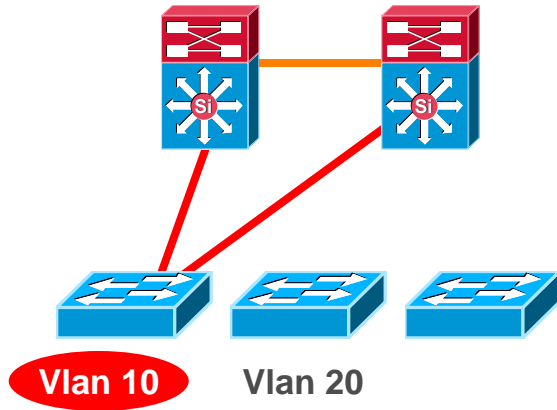
Best Practices— Spanning Tree Configuration

- Only span VLAN across multiple access layer switches when you have to!
- Use rapid PVST+ for best convergence
- More common in the data center
- Required to protect against user side loops
- Required to protect against operational accidents (misconfiguration or hardware failure)
- Take advantage of the spanning tree toolkit

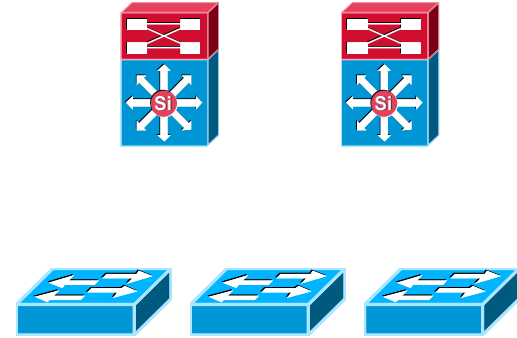


Multilayer Network Design

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

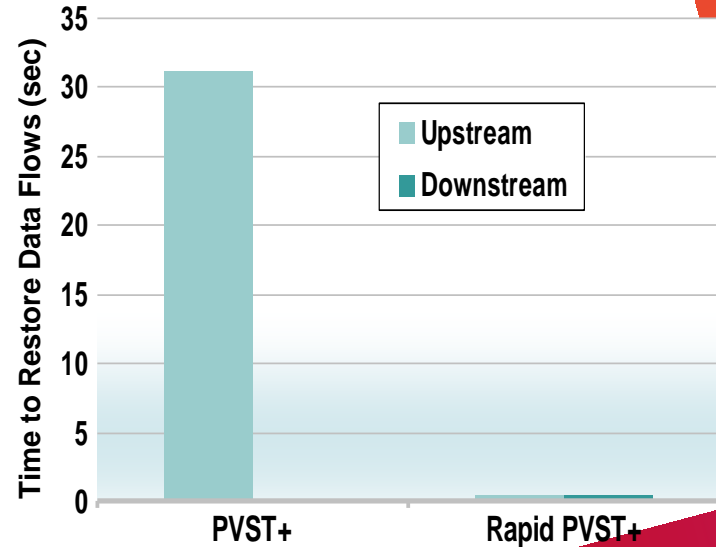


- At least some VLANs span multiple access switches
- Layer 2 loops
- Layer 2 and 3 running over link between distribution
- Blocked links

Optimizing L2 Convergence

PVST+, Rapid PVST+ or MST

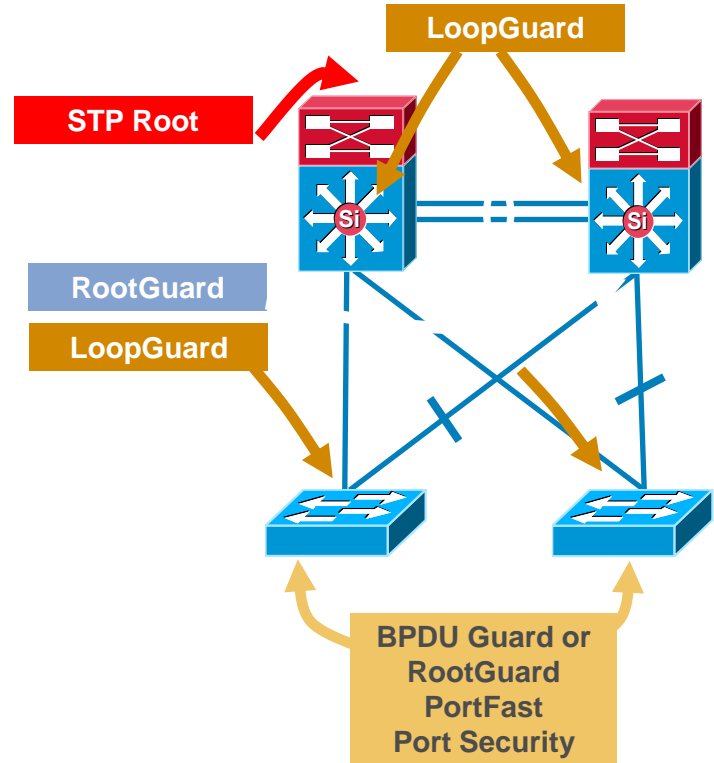
- Rapid-PVST+ greatly improves the restoration times for any VLAN that requires a topology convergence due to link UP
- Rapid-PVST+ also greatly improves convergence time over backbone fast for any indirect link failures
- PVST+ (802.1d)
 - Traditional spanning tree implementation
- Rapid PVST+ (802.1w)
 - 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+



Layer 2 Hardening

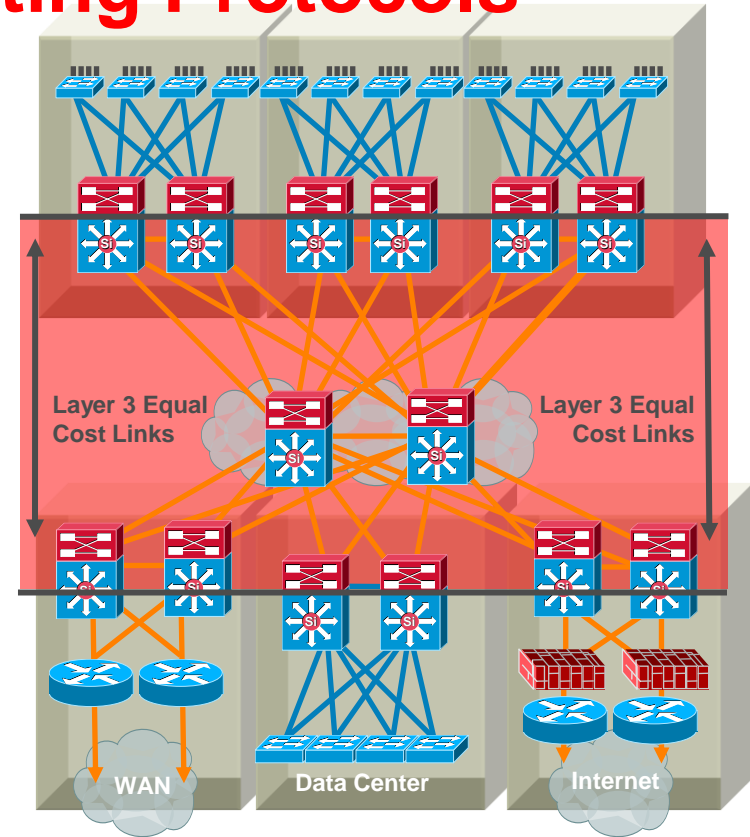
Spanning Tree Should Behave the Way You Expect

- 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



Best Practices - Layer 3 Routing Protocols

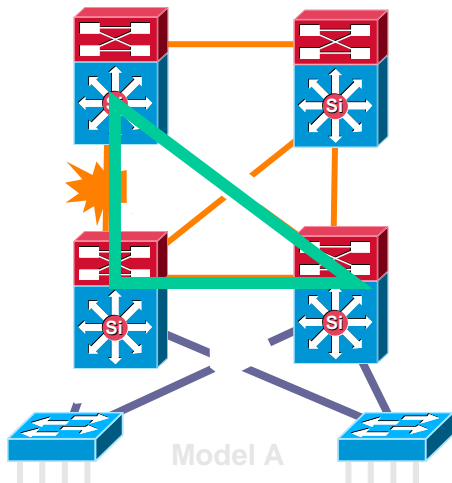
- Typically deployed in distribution to core, and core-to-core interconnections
- Used to quickly reroute around failed node/links while providing load balancing over redundant paths
- Build triangles not squares for deterministic convergence
- Only peer on links that you intend to use as transit
- Insure redundant L3 paths to avoid black holes
- Summarize distribution to core to limit EIGRP query diameter or OSPF LSA propagation
- Tune CEF L3/L4 load balancing hash to achieve maximum utilization of equal cost paths (CEF polarization)



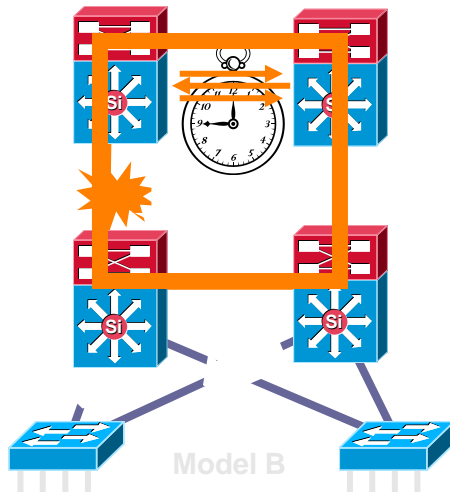
Best Practice—Build Triangles not Squares

Deterministic vs. Non-Deterministic

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



Squares: Link/Box Failure Requires Routing Protocol Convergence

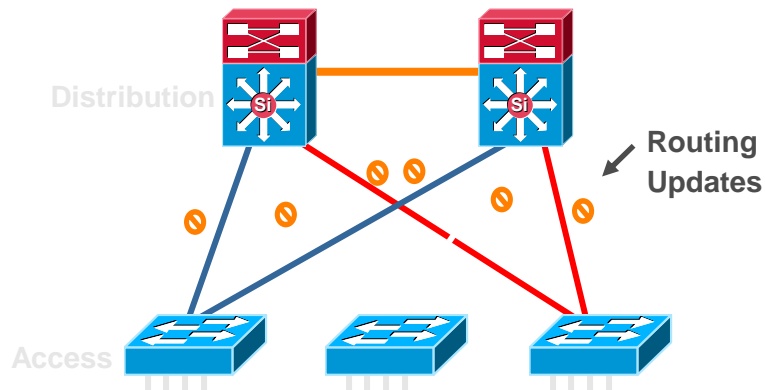


- Layer 3 redundant equal cost links support fast convergence
- Hardware based—fast recovery to remaining path
- Convergence is extremely fast (dual equal-cost paths: no need for OSPF or EIGRP to recalculate a new path)

Best Practice - Passive Interfaces for IGP

Limit OSPF and EIGRP Peering Through the Access Layer

- Limit unnecessary peering using passive interface:
 - Four VLANs per wiring closet
 - 12 adjacencies total
 - Memory and CPU requirements increase with no real benefit
 - Creates overhead for IGP



OSPF Example:

```
Router(config)#routerospf 1
Router(config-router)#passive-
interfaceVlan 99

Router(config)#routerospf 1
Router(config-router)#passive-
interface default

Router(config-router)#no passive-
interface Vlan 99
```

EIGRP Example:

```
Router(config)#routereigrp 1
Router(config-router)#passive-
interfaceVlan 99

Router(config)#routereigrp 1
Router(config-router)#passive-
interface default

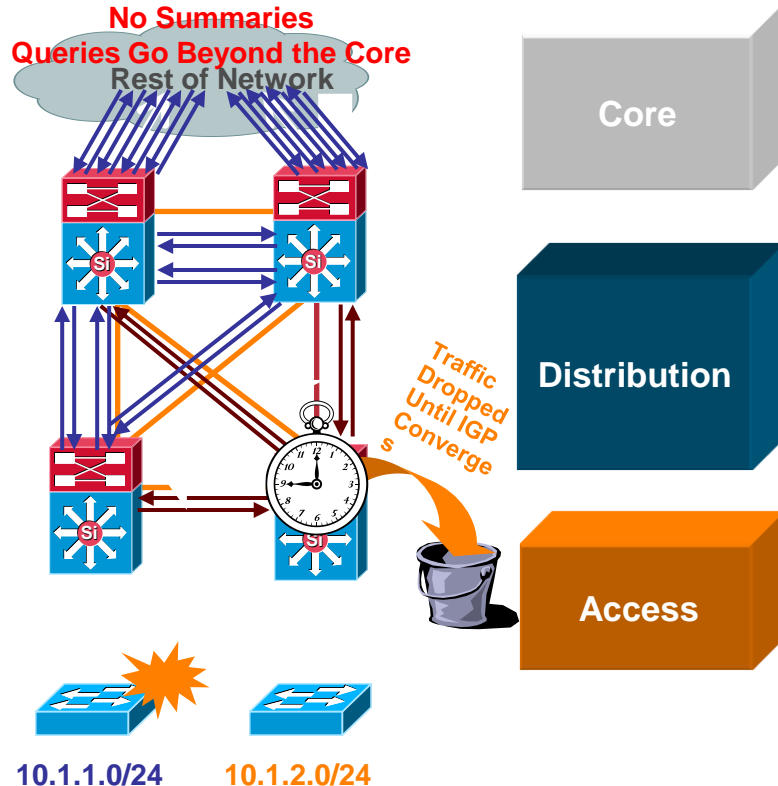
Router(config-router)#no passive-
interface Vlan 99
```

Why You Want to Summarize at the Distribution

Limit EIGRP Queries and OSPF LSA Propagation

- It is important to force summarization at the distribution towards the core
- For return path traffic an OSPF or EIGRP re-route is required
- By limiting the number of peers an EIGRP router must query or the number of LSAs an OSPF peer must process we can optimize this reroute
- EIGRP example:

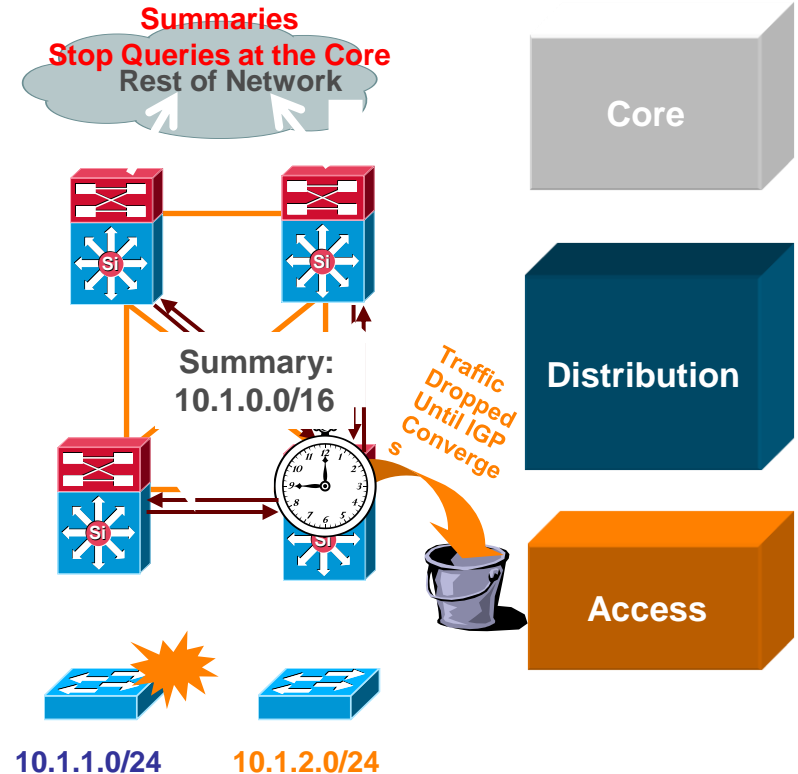
```
interface Port-channell
description to Core#1
ip address 10.122.0.34
255.255.255.252
ip hello-interval eigrp 100
1
ip hold-time eigrp 100 3
ip summary-address eigrp 100 10.1.0.0 255.255.0.0 5
```



Why You Want to Summarize at the Distribution

Reduce the Complexity of IGP Convergence

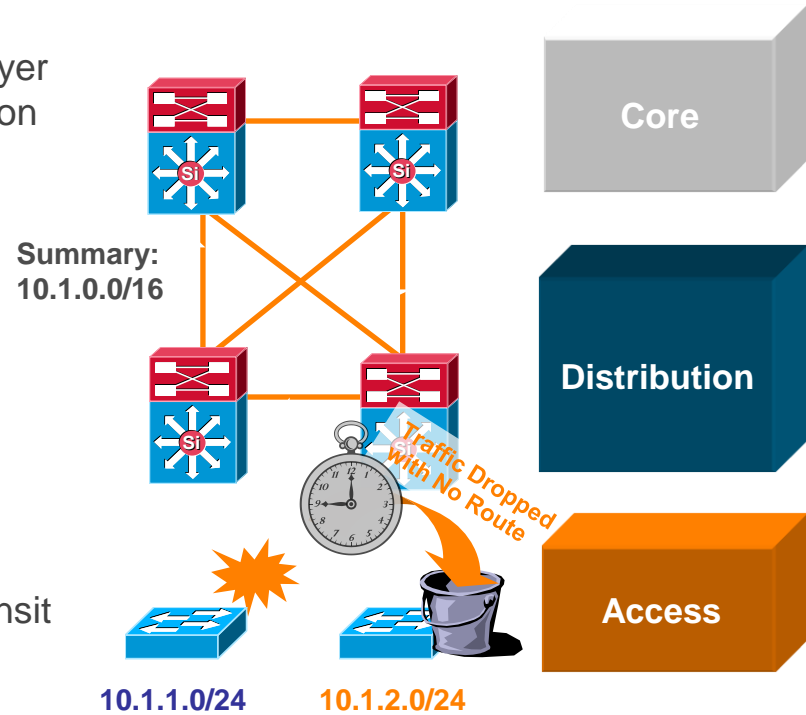
- It is important to force summarization at the distribution towards the core
- For return path traffic an OSPF or EIGRP re-route is required
- By limiting the number of peers an EIGRP router must query or the number of LSAs an OSPF | peer must process we can optimize his reroute
- For EIGRP if we summarize at the distribution we stop queries at the core boxes for an access layer flap
- For OSPF when we summarize at the distribution (area border or L1/L2 border) the flooding of LSAs is limited to the distribution switches; SPF now deals with one LSA not three




Best Practice— Summarize at the Distribution

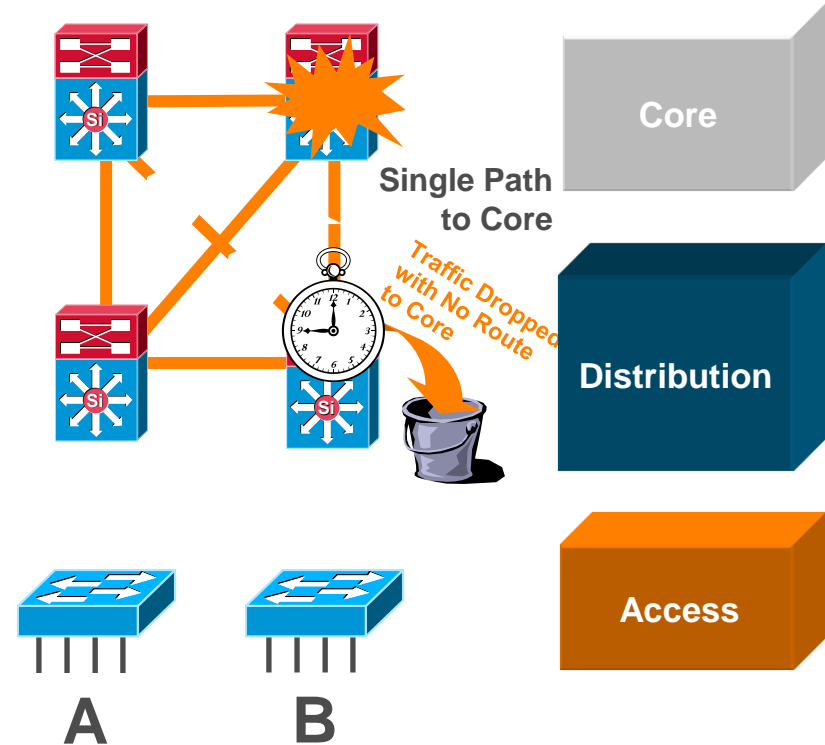
Gotcha—Distribution-to-Distribution Link Required

- Best practice—summarize at the distribution layer to limit EIGRP queries or OSPF LSA propagation
- Gotcha:
 - Upstream: HSRP on left distribution takes over when link fails
 - Return path: old router still advertises summary to core
 - Return traffic is dropped on right distribution switch
- Summarizing requires a link between the distribution switches
- Alternative design: use the access layer for transit



Provide Alternate Paths

- What happens if fails? 
- No route to the core anymore?
- Allow the traffic to go through the access?
 - Do you want to use your access switches as transit nodes?
 - How do you design for scalability if the access used for transit traffic?
- Install a redundant link to the core
- **Best practice: install redundant link to core and utilize L3 link between distribution layer**



Equal-Cost Multipath

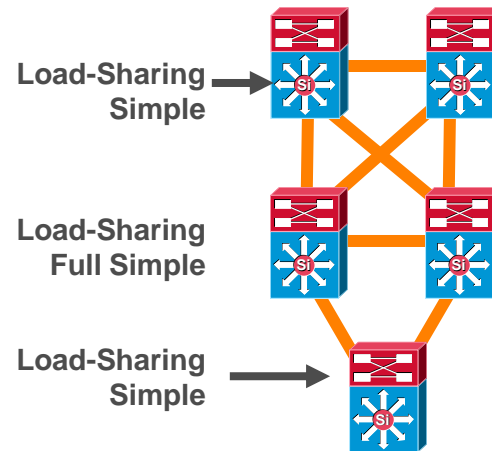
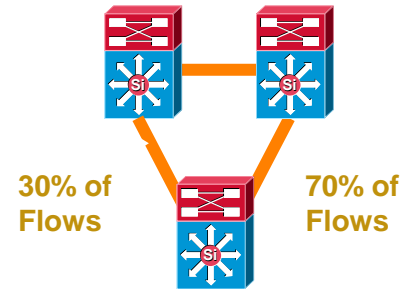
Optimizing CEF Load-Sharing

- Depending on the traffic flow patterns and IP Addressing in use one algorithm may provide better load-sharing results than another
- Be careful not to introduce polarization in a multi-tier design by changing the default to the same thing in all tiers/layers of the network

Original	Src IP + Dst IP
Universal*	Src IP + Dst IP + Unique ID
Include Port	Src IP + Dst IP + (Src or Dst Port) + Unique ID
Default*	Src IP + Dst IP + Unique ID
Full	Src IP + Dst IP + Src Port + Dst Port
Full Exclude Port	Src IP + Dst IP + (Src or Dst Port)
Simple	Src IP + Dst IP
Full Simple	Src IP + Dst IP + Src Port + Dst Port

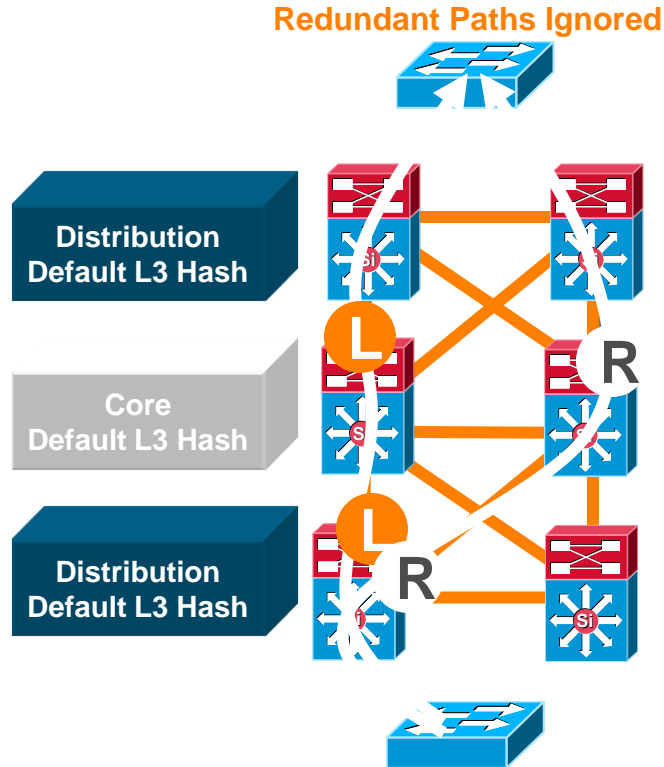
* = Default Load-Sharing Mode

** = PFC3 in Sup720 and Sup32 Supervisors



CEF Load Balancing

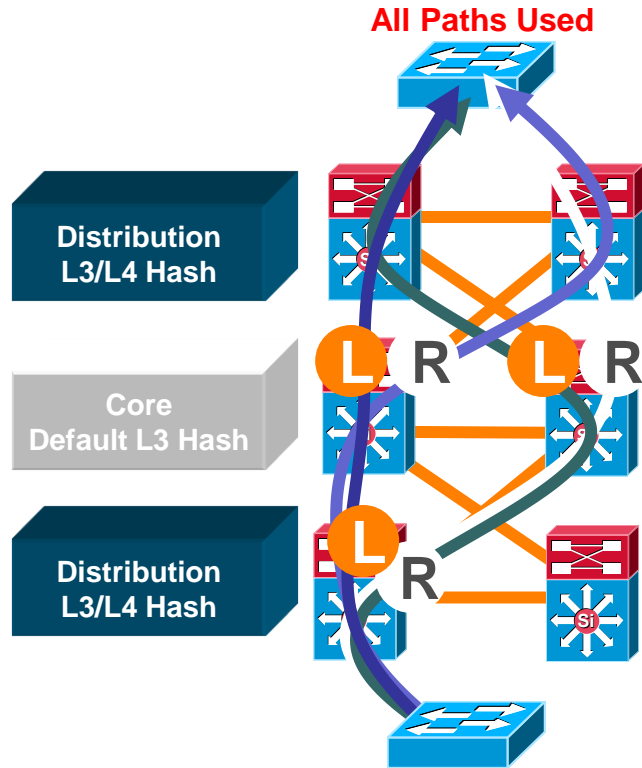
Avoid Underutilizing Redundant Layer 3 Paths



- **CEF polarization**: without some tuning CEF will select the same path left/left or right/right
- Imbalance/overload could occur
- Redundant paths are ignored/ underutilized
- The default CEF hash **input** is L3
- We can change the default to use L3 + L4 information as **input** to the hash derivation

CEF Load Balancing

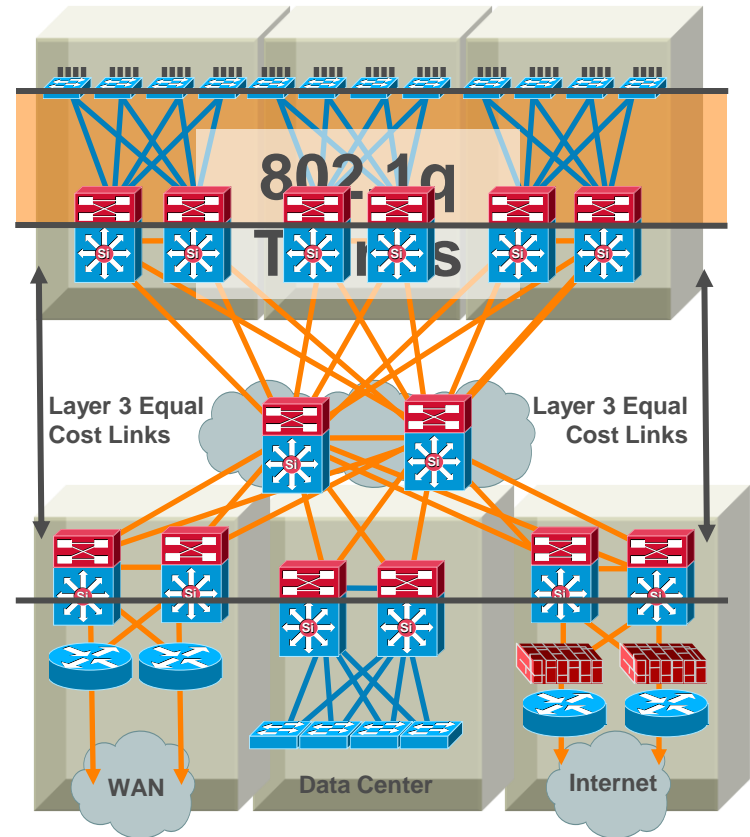
Avoid Underutilizing Redundant Layer 3 Paths



- The default will for Sup720/32 and latest hardware (unique ID added to default). However, depending on IP addressing, and flows imbalance could occur
- Alternating L3/L4 hash and L3 hash will give us the best load balancing results
- Use **simple** in the core and **full simple** in the distribution to add L4 information to the algorithm at the distribution and maintain differentiation tier-to-tier

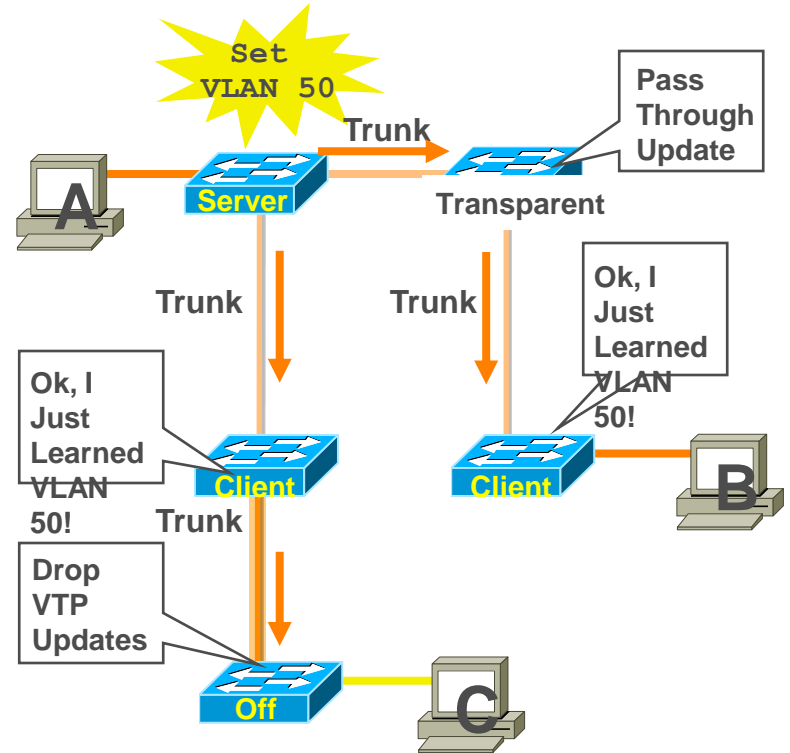
Best Practices—Trunk Configuration

- 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:
 - CatOS: `set port host`
 - Cisco IOS: `switchport host`



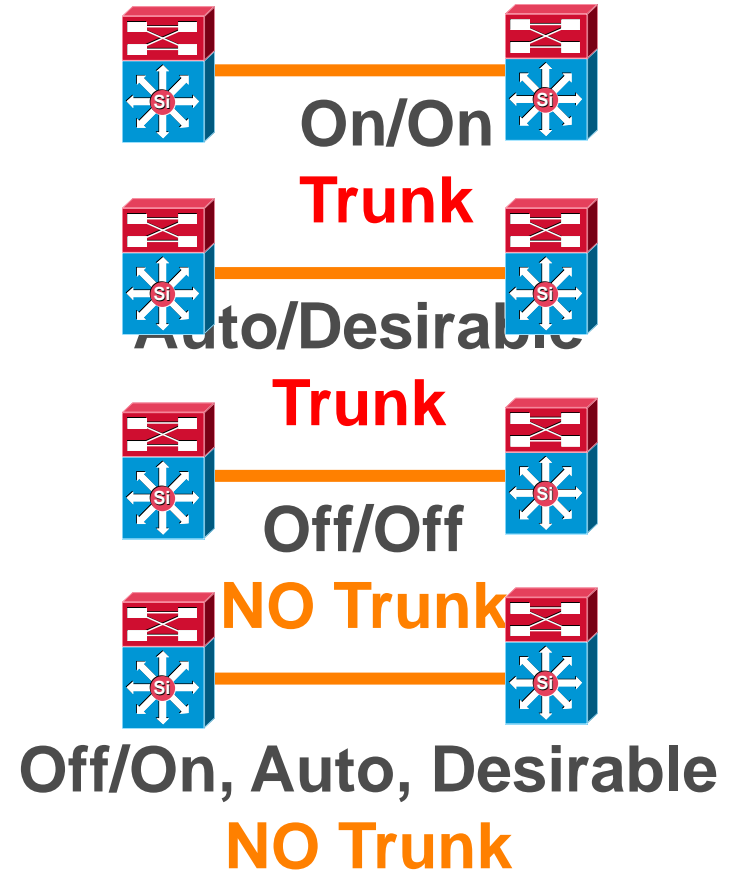
VTP Virtual Trunk Protocol

- Centralized VLAN management
- VTP server switch propagates VLAN database to VTP client switches
- Runs only on trunks
- Four modes:
 - **Server**: updates clients and servers
 - **Client**: receive updates—cannot make changes
 - **Transparent**: let updates pass through
 - **Off**: ignores VTP updates



DTP Dynamic Trunk Protocol

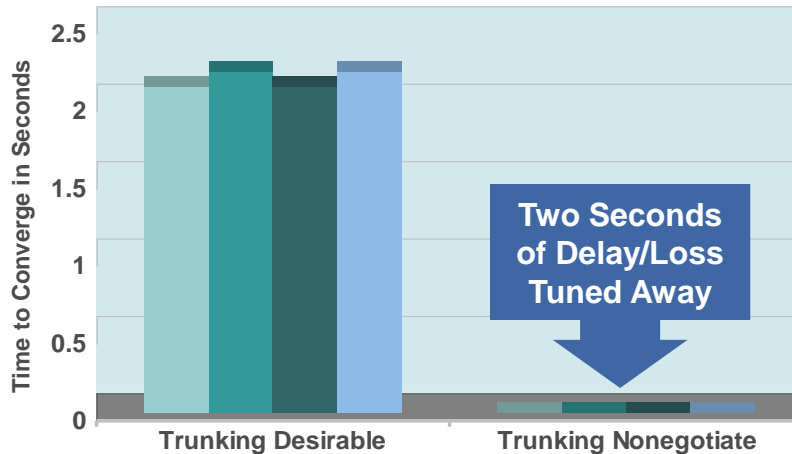
- Automatic formation of trunked switch-to-switch interconnection
 - **On**: always be a trunk
 - **Desirable**: ask if the other side can/will
 - **Auto**: if the other sides asks I will
 - **Off**: don't become a trunk
- Negotiation of 802.1Q or ISL encapsulation
 - **ISL**: try to use ISL trunk encapsulation
 - **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



Optimizing 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`



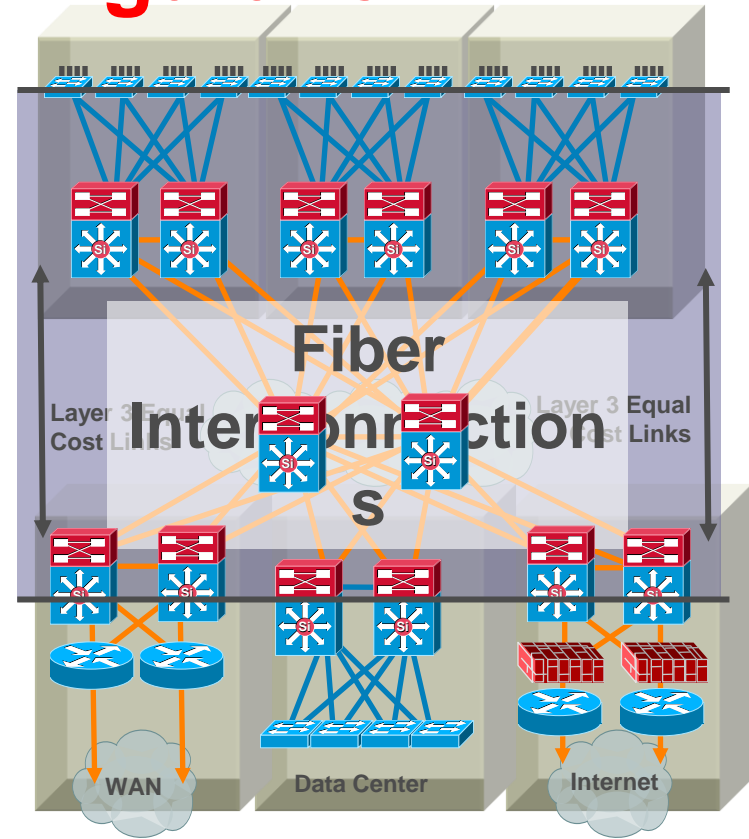
Trunking/VTP/DTP—Quick Summary

- 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
- Emerging technologies that do VLAN assignment by name (IBNS, NAC, etc.) require a unique VLAN database per access layer switch if the rule: A VLAN = A Subnet = AN access layer switch is going to be followed
- 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 ...



Best Practices—UDLD Configuration

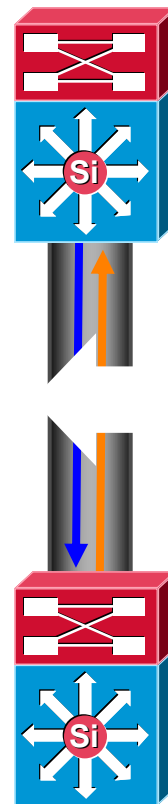
- Typically deployed on any fiber optic interconnection
- Use UDLD aggressive mode for most aggressive protection
- Turn on in global configuration to avoid operational error/**misses**
- Config example
 - Cisco IOS:
`udld aggressive`



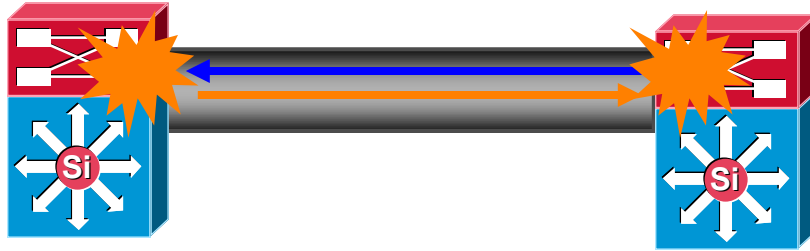
Unidirectional Link Detection

Protecting Against One-Way Communication

- Highly-available networks require UDLD to protect against one-way communication or partially failed links and the effect that they could have on protocols like STP and RSTP
- Primarily used on fiber optic links where patch panel errors could cause link up/up with mismatched transmit/receive pairs
- Each switch port configured for UDLD will send UDLD protocol packets (at L2) containing the port's own device/port ID, and the neighbor's device/port IDs seen by UDLD on that port
- Neighboring ports should see their own device/port ID (echo) in the packets received from the other side
- If the port does not see its own device/port ID in the incoming UDLD packets for a specific duration of time, the link is considered unidirectional and is shutdown



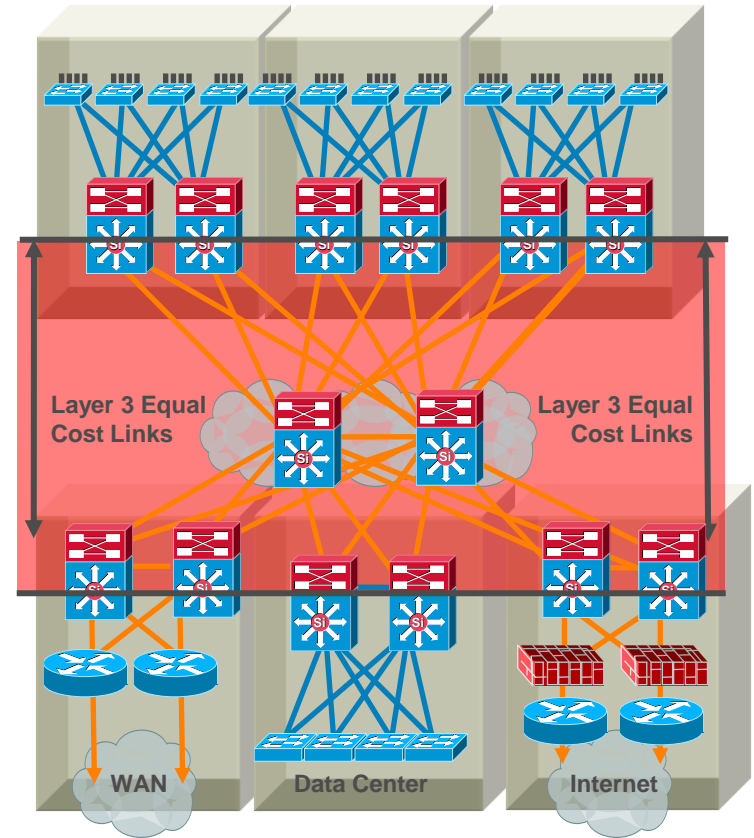
UDLD Aggressive and UDLD Normal



- Timers are the same—15-second hellos by default
- Aggressive Mode—after aging on a previously bi-directional link—tries eight times (once per second) to reestablish connection then err-disables port
- UDLD—Normal Mode—only err-disable the end where UDLD detected other end just sees the link go down
- UDLD—Aggressive—err-disable **both** ends of the connection due to err-disable when aging and re-establishment of UDLD communication fails

Best Practices— EtherChannel Configuration

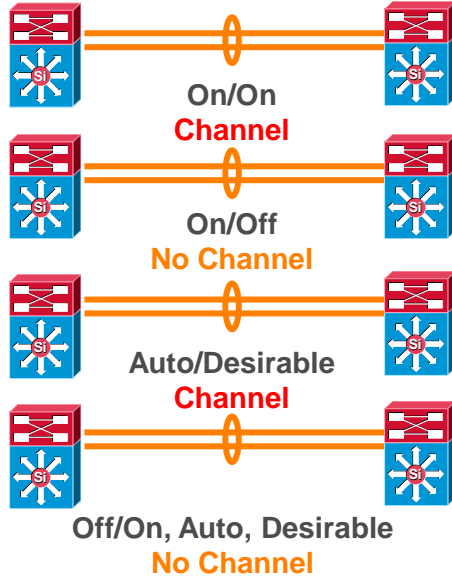
- Typically deployed in distribution to core, and core to core interconnections
- Used to provide link redundancy—while reducing peering complexity
- Tune L3/L4 load balancing hash to achieve maximum utilization of channel members
- Deploy in powers of two (two, four, or eight)
- Match CatOS and Cisco IOS PAgP settings
- 802.3ad LACP for interop if you need it
- Disable unless needed
 - Cisco IOS: `switchport host`



Understanding EtherChannel

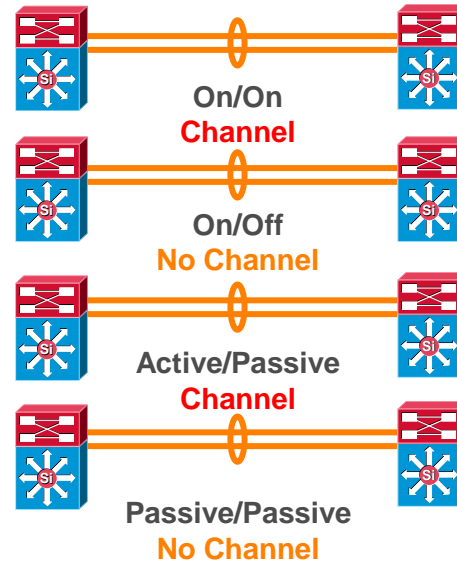
Link Negotiation Options—PAgP and LACP

Port Aggregation Protocol



On: always be a channel/bundle member
Desirable: ask if the other side can/will
Auto: if the other side asks I will
Off: don't become a member of a channel/bundle

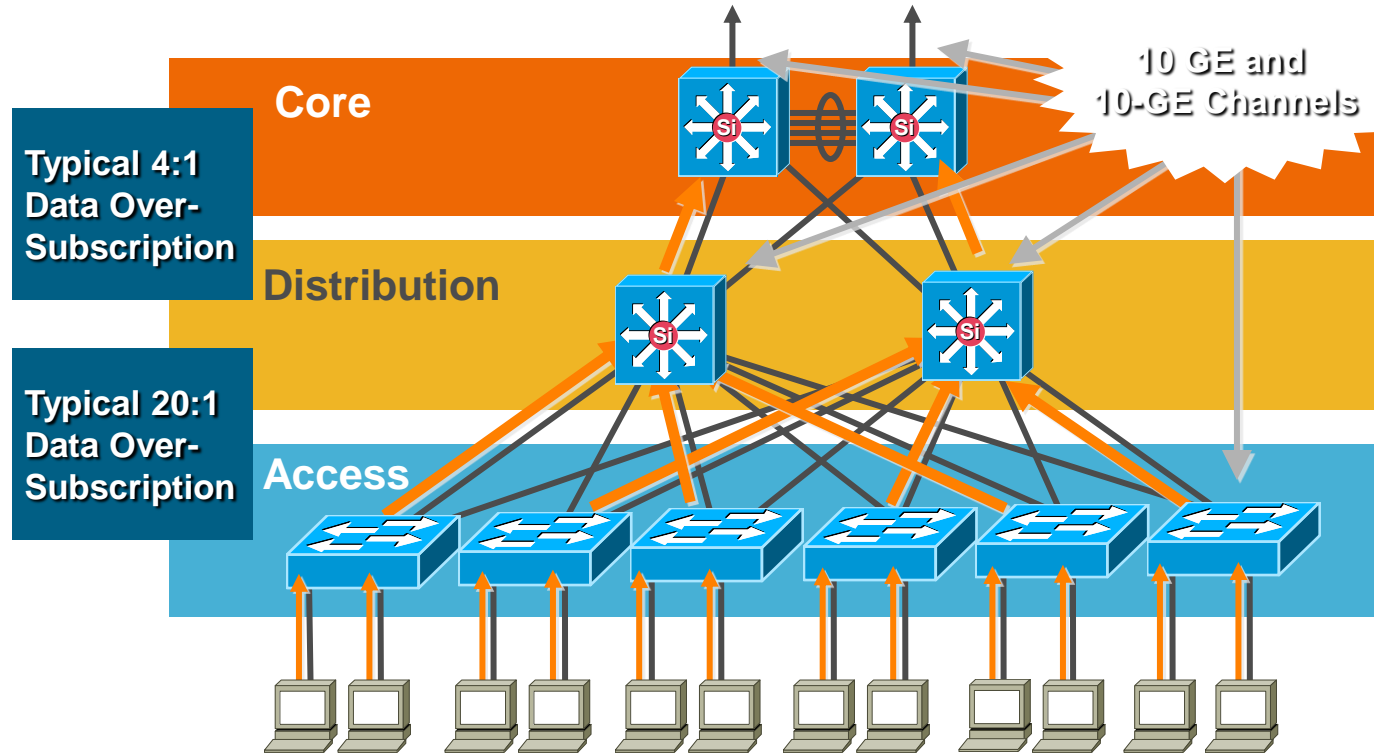
Link Aggregation Protocol



On: always be a channel/bundle member
Active: ask if the other side can/will
Passive: if the other side asks I will
Off: don't become a member of a channel/bundle

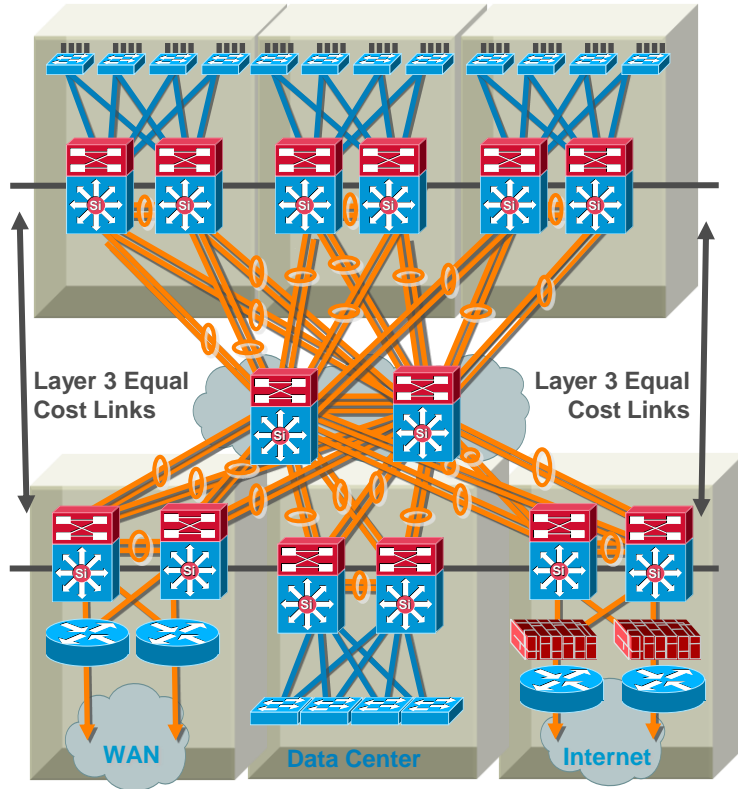
EtherChannels or Equal Cost Multipath

10/100/1000 How Do You Aggregate It?



EtherChannels or Equal Cost Multipath

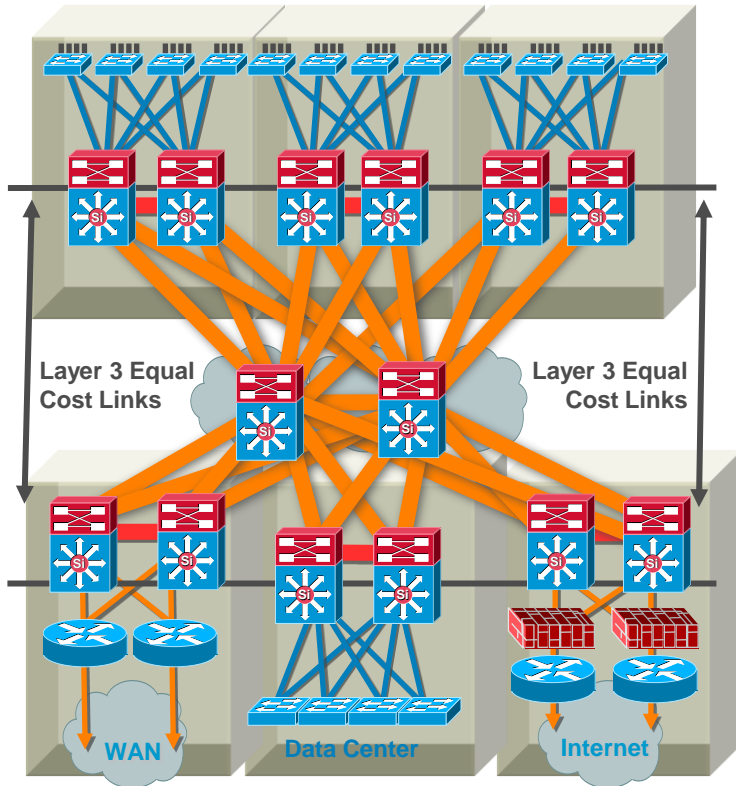
Reduce Complexity/Peer Relationships



- More links = more routing peer relationships and associated overhead
- EtherChannels allow you to reduce peers by creating single logical interface to peer over
- On single link failure in a bundle
 - OSPF running on a Cisco IOS-based switch will reduce link cost and reroute traffic
 - OSPF running on a hybrid switch will **not** change link cost and may overload remaining links
 - EIGRP **may not** change link cost and may overload remaining links

EtherChannels or Equal Cost Multipath

Why 10-Gigabit Interfaces



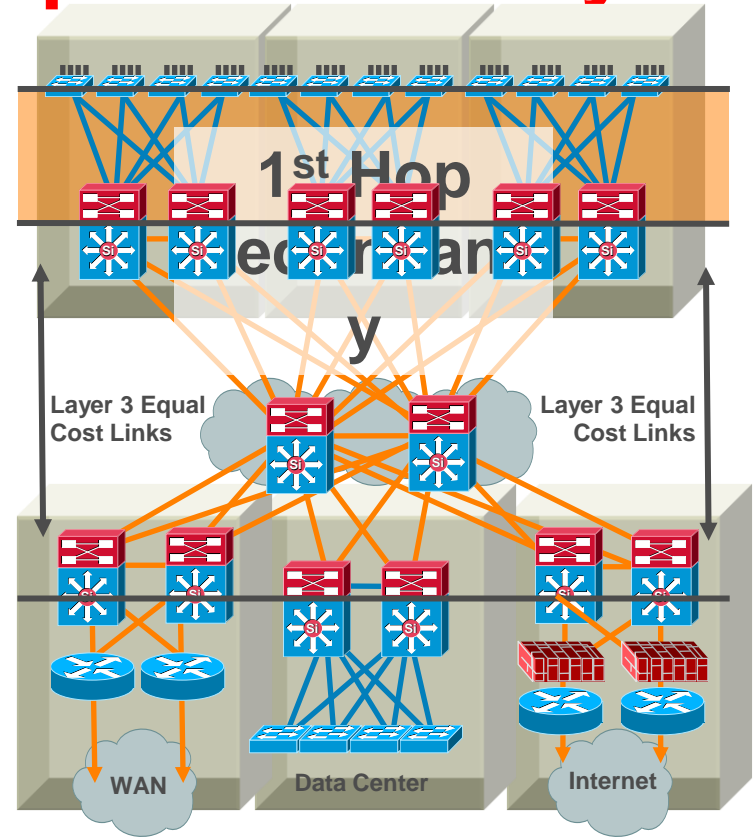
- More links = more routing peer relationships and associated overhead
- EtherChannels allow you to reduce peers by creating single logical interface to peer over
- However, a single link failure is not taken into consideration by routing protocols. Overload possible
- Single 10-gigabit links address both problems. Increased bandwidth without increasing complexity or compromising routing protocols ability to select best path

EtherChannels—Quick Summary

- For Layer 2 EtherChannels: **Desirable/Desirable** is the recommended configuration so that PAgP is running across all members of the bundle **insuring** that an individual link failure will not result in an STP failure
- For Layer 3 EtherChannels: one can consider a configuration that uses **ON/ON**. There is a trade-off between performance/HA impact and maintenance and operations implications
- An ON/ON configuration is faster from a link-up (restoration) perspective than a Desirable/Desirable alternative. However, in this configuration PAgP is not actively monitoring the state of the bundle members and a misconfigured bundle is not easily identified
- Routing protocols may not have visibility into the state of an individual member of a bundle. LACP and the minimum links option can be used to bring the entire bundle down when the capacity is diminished.
 - OSPF has visibility to member loss (best practices pending investigation). EIGRP does not...
- When used to increase bandwidth—no individual flow can go faster than the speed of an individual member of the link
- Best used to eliminate single points of failure (i.e., link or port) dependencies from a topology

Best Practices—First Hop Redundancy

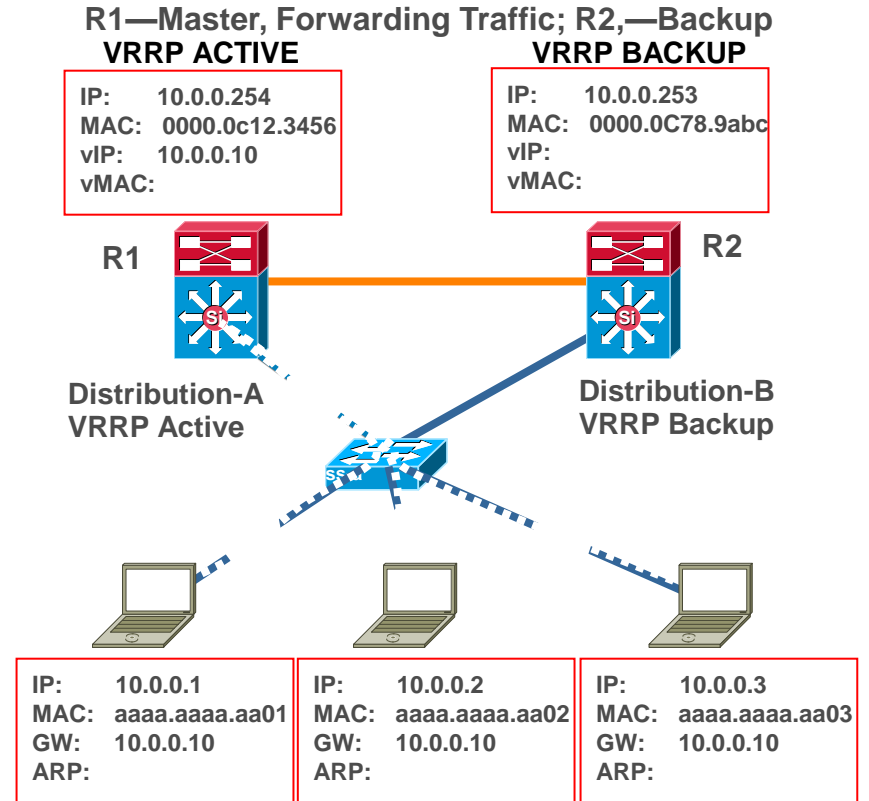
- Used to provide a resilient default gateway/ first hop address to end-stations
- HSRP, VRRP, and GLBP alternatives
- VRRP, HSRP, and GLBP provide millisecond timers and excellent convergence performance
- VRRP if you need multivendor interoperability
- GLBP facilitates uplink load balancing
- Preempt timers need to be tuned to avoid black-holed traffic



First Hop Redundancy with VRRP

IETF Standard RFC 2338 (April 1998)

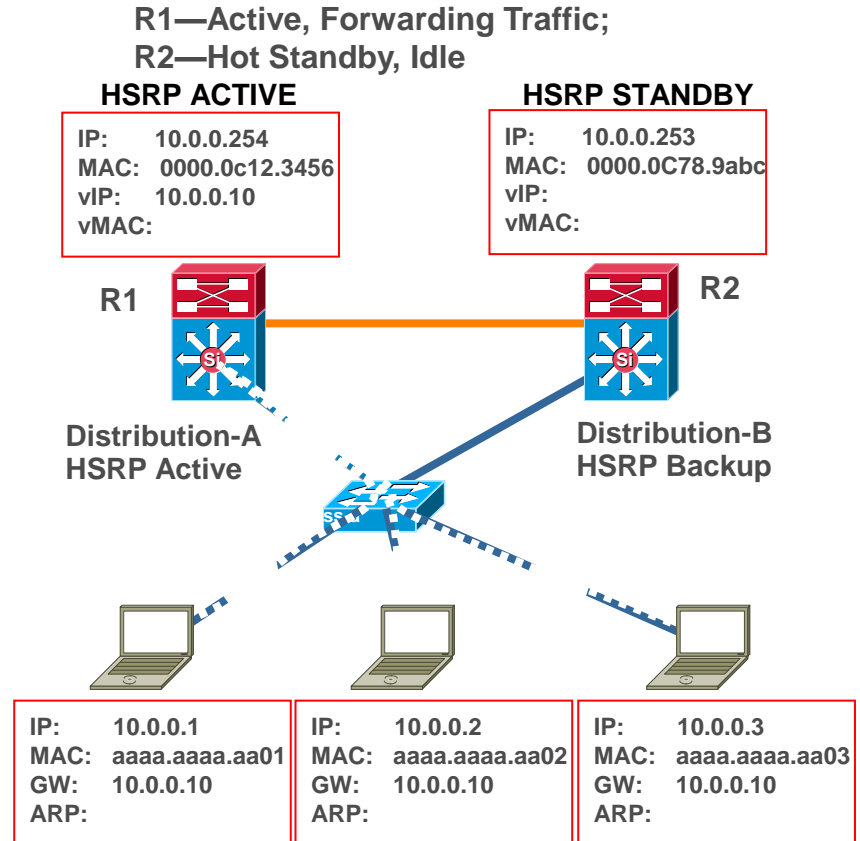
- A group of routers function as one virtual router by sharing **one** virtual IP address and one virtual MAC address
- One (master) router performs packet forwarding for local hosts
- The rest of the routers act as **back up** in case the master router fails
- Backup routers stay idle as far as packet forwarding from the client side is concerned



First Hop Redundancy with HSRP

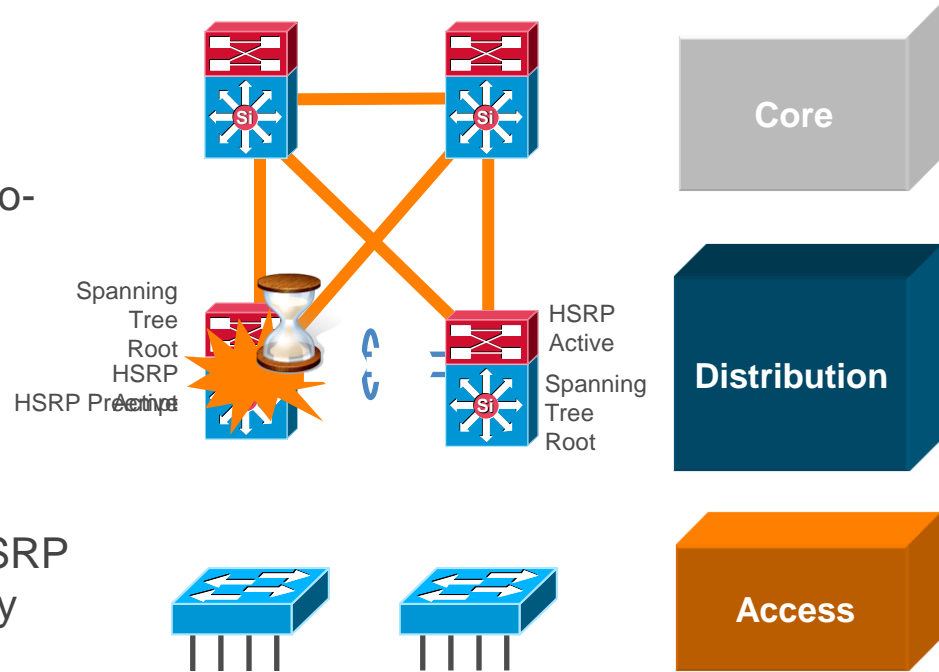
RFC 2281 (March 1998)

- A group of routers function as one virtual router by sharing **one** virtual IP address and one virtual MAC address
- One (active) router performs packet forwarding for local hosts
- The rest of the routers provide **hot standby** in case the active router fails
- Standby routers stay idle as far as packet forwarding from the client side is concerned



Why You Want HSRP Preemption

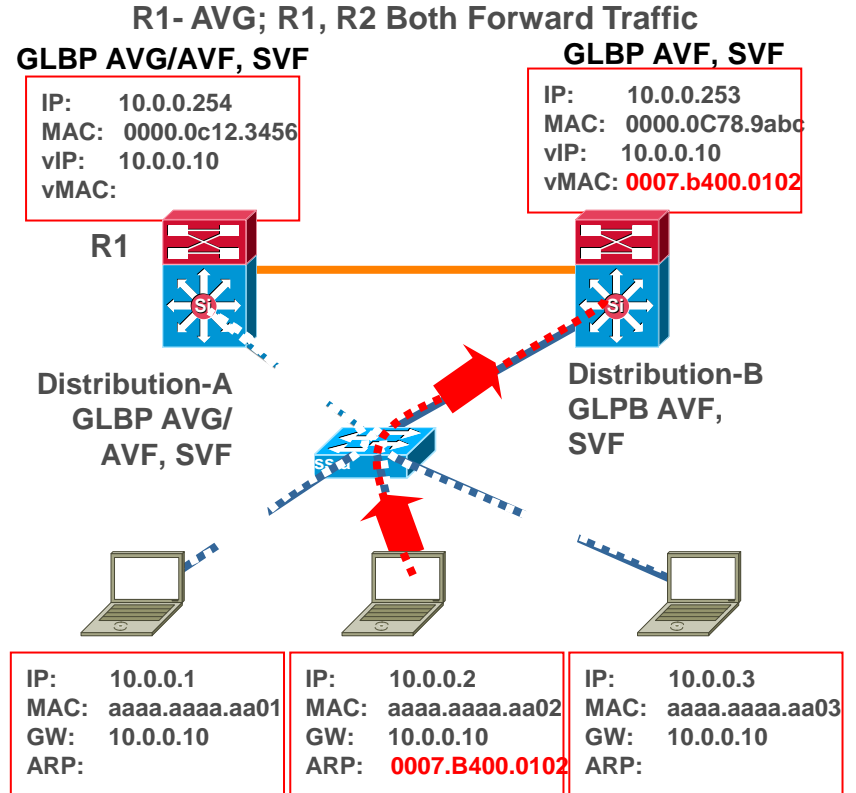
- Spanning tree root and HSRP primary aligned
- When spanning tree root is re-introduced, traffic will take a two-hop path to HSRP active
- HSRP preemption will allow HSRP to follow spanning tree topology



First Hop Redundancy with GLBP

Cisco Designed, Load Sharing, Patent Pending

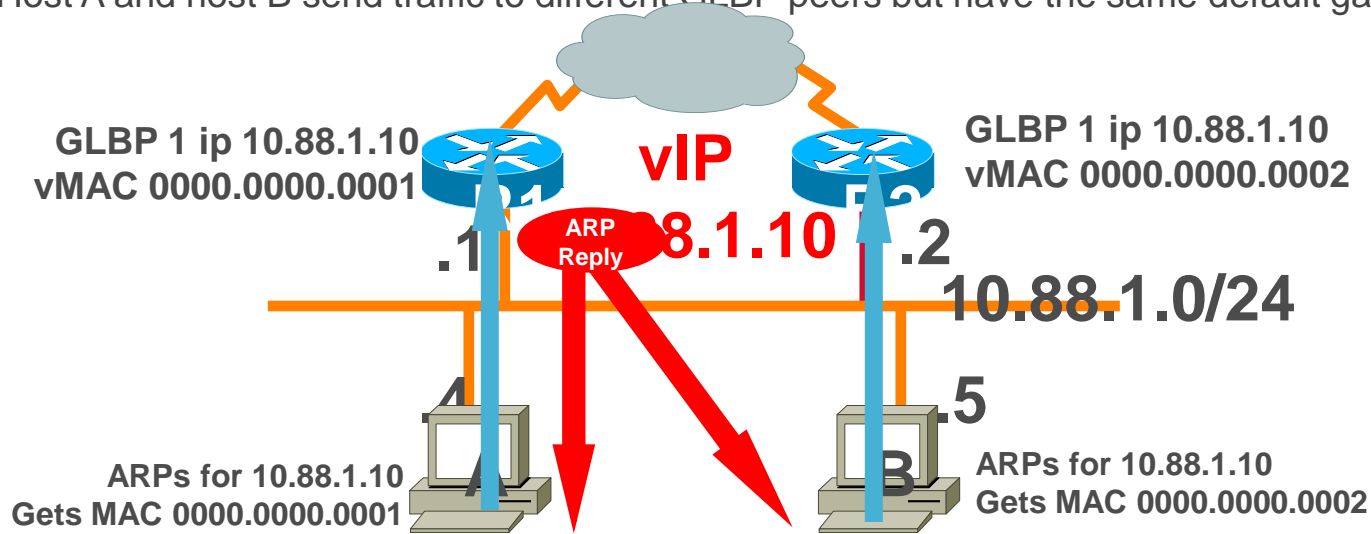
- All the benefits of HSRP plus load balancing of default gateway → utilizes all available bandwidth
- A group of routers function as one virtual router by sharing one virtual IP address but using multiple virtual MAC addresses for traffic forwarding
- Allows traffic from a single common subnet to go through multiple redundant gateways using a single virtual IP address



First Hop Redundancy with Load Balancing

Cisco Gateway Load Balancing Protocol (GLBP)

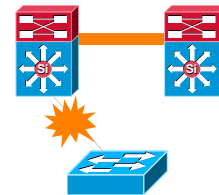
- Each member of a GLBP redundancy group owns a unique virtual MAC address for a common IP address/ default gateway
- When end-stations ARP for the common IP address/default gateway they are given a load-balanced virtual MAC address
- Host A and host B send traffic to different GLBP peers but have the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



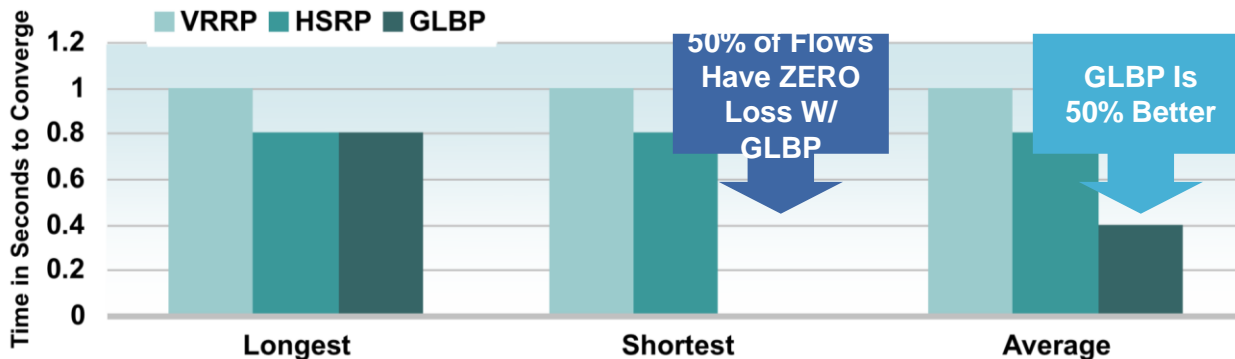
Optimizing Convergence: VRRP, HSRP, GLBP

Mean, Max, and Min—Are There Differences?

- VRRP not tested with sub-second timers and all flows go through a common VRRP peer; mean, max, and min are equal
- HSRP has sub-second timers; however all flows go through same HSRP peer so there is no difference between mean, max, and min
- GLBP has sub-second timers and distributes the load amongst the GLBP peers; so 50% of the clients are not affected by an uplink failure



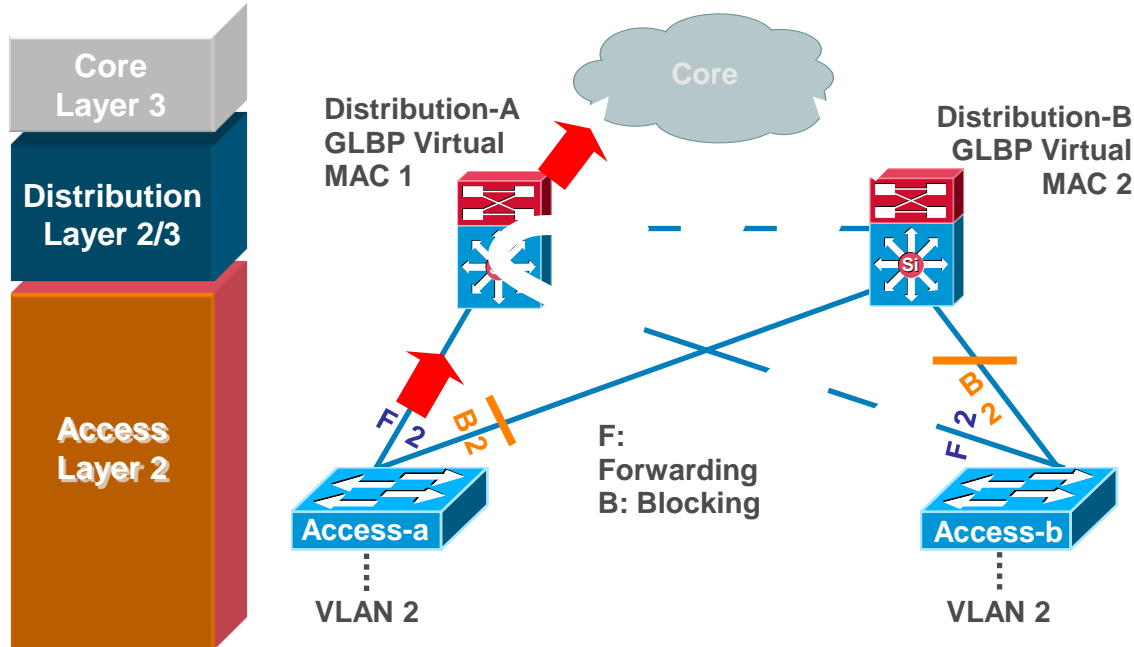
Distribution to Access Link Failure Access to Server Farm



If You Span VLANS, Tuning Required

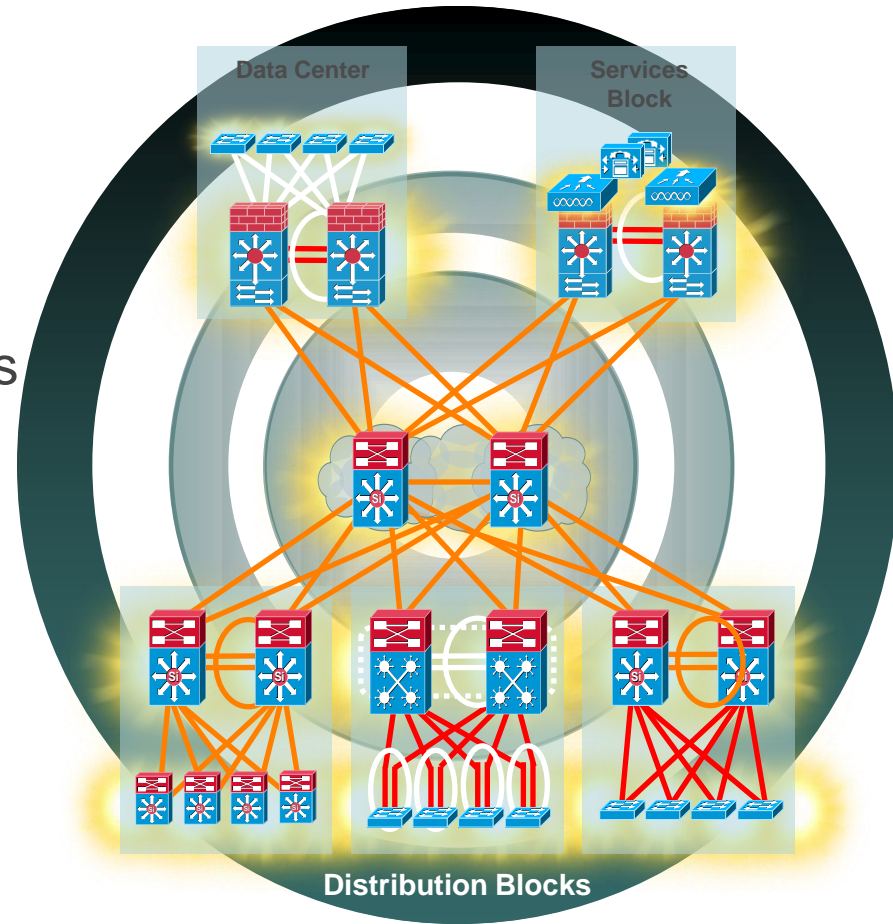
By Default, Half the Traffic Will Take a Two-Hop L2 Path

- Both distribution switches act as default gateway
- Blocked uplink caused traffic to take less than optimal path



Agenda

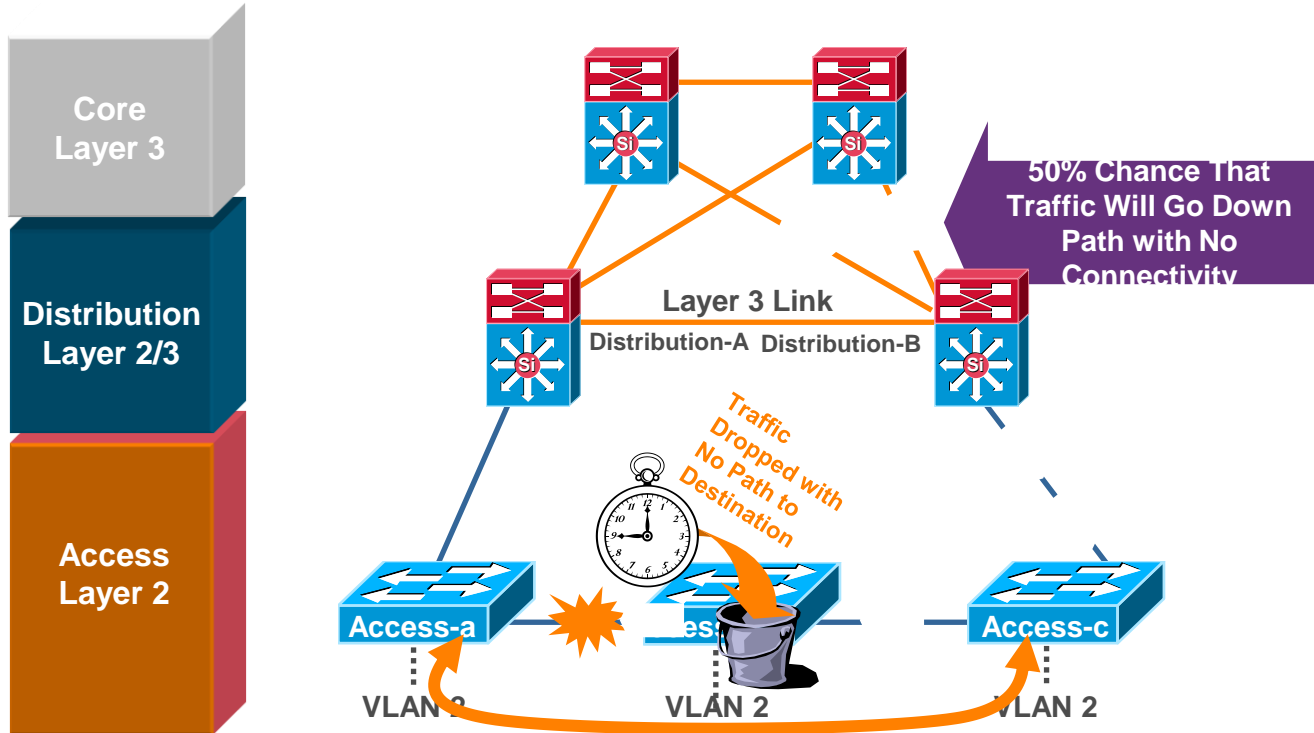
- Multilayer Campus Design Principles
- **Foundation Services**
- Campus Design Best Practices
- IP Telephony Considerations
- QoS Considerations
- Security Considerations
- Putting It All Together
- Summary



Daisy Chaining Access Layer Switches

Avoid Potential Black Holes

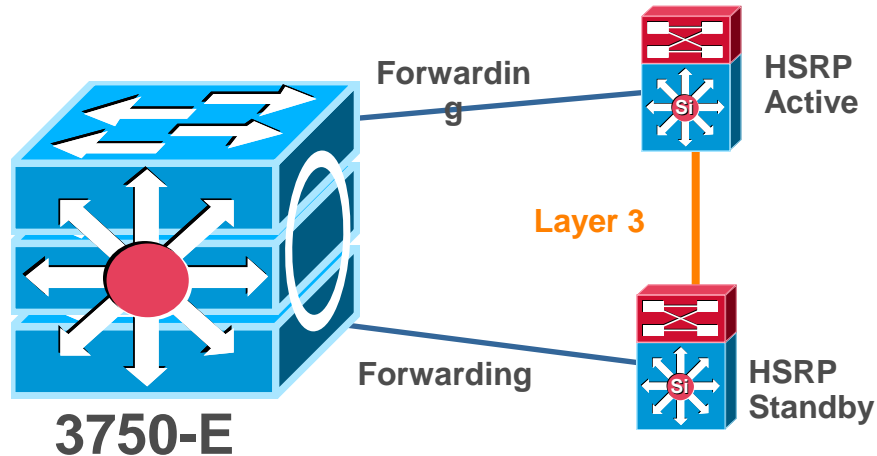
Return Path Traffic Has a 50/50 Chance of Being 'Black Holed'



Daisy Chaining Access Layer Switches

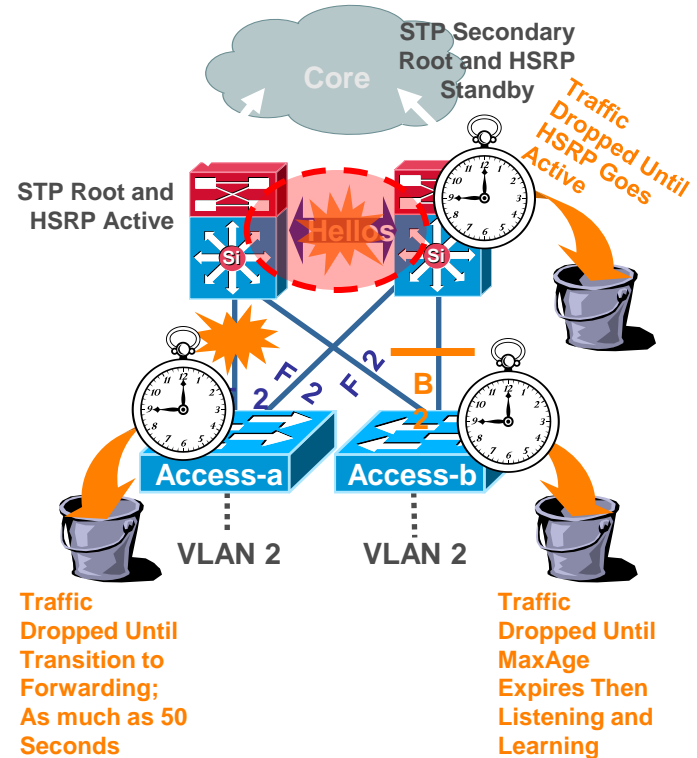
New Technology Addresses Old Problems

- **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



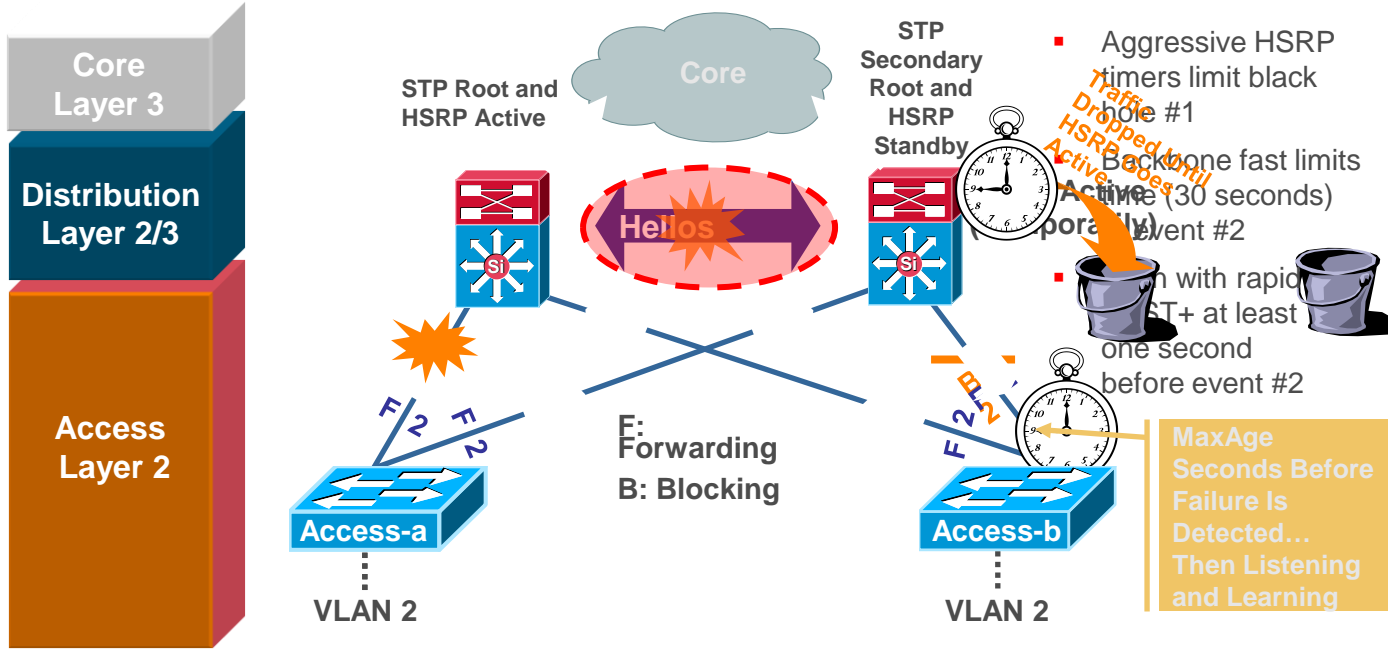
What Happens if You Don't Link the Distributions?

- STPs slow convergence can cause considerable periods of traffic loss
- STP could cause non-deterministic traffic flows/link load engineering
- STP convergence will cause Layer 3 convergence
- STP and Layer 3 timers are independent
- Unexpected Layer 3 convergence and convergence could occur
- Even if you do link the distribution switches dependence on STP and link state/connectivity can cause HSRP irregularities and unexpected state transitions



What if You Don't?

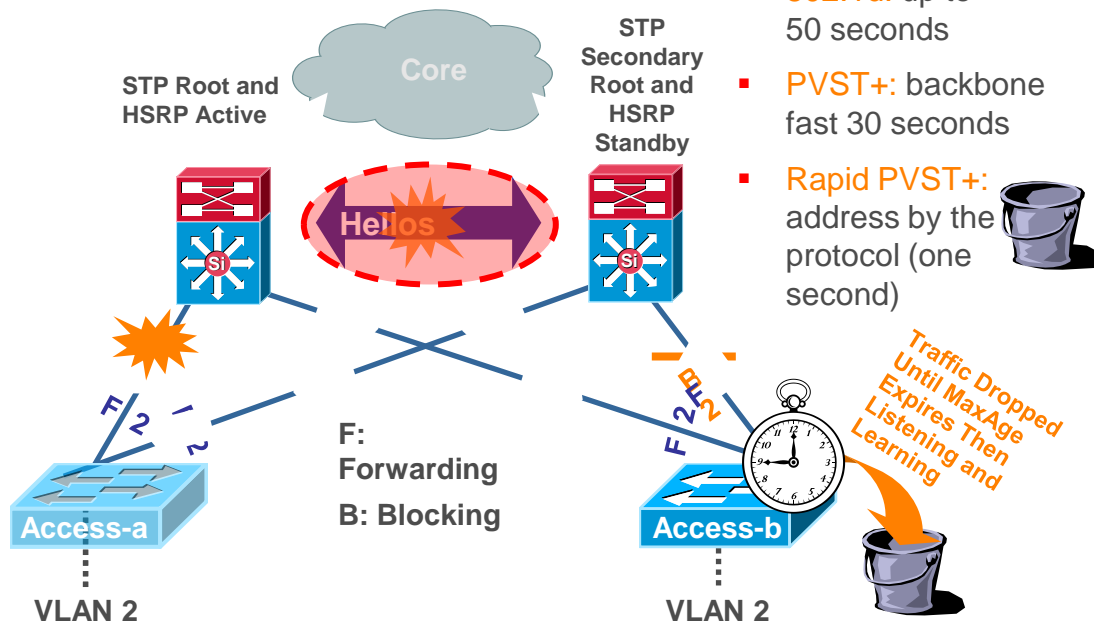
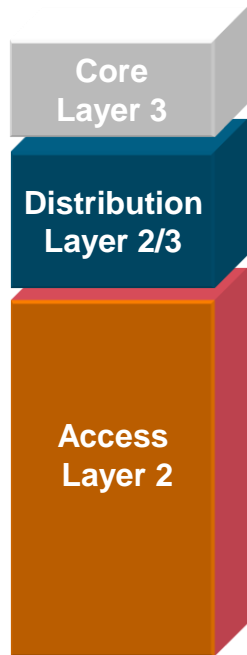
Black Holes and Multiple Transitions ...



- Blocking link on access-b will take 50 seconds to move to forwarding → traffic black hole until HSRP goes active on standby HSRP peer
- After MaxAge expires (or backbone fast or Rapid PVST+) converges HSRP preempt causes another transition
- Access-b used as transit for access-a's traffic

What If You Don't?

Return Path Traffic Black Holed ...

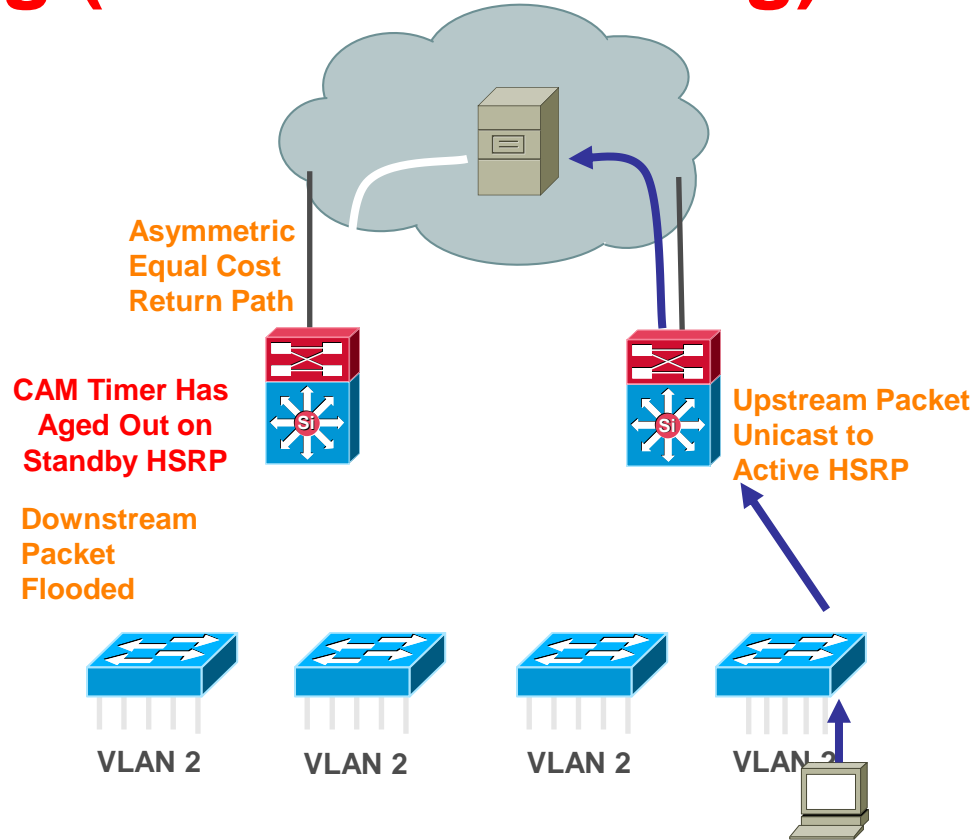


- 802.1d: up to 50 seconds
- PVST+: backbone fast 30 seconds
- Rapid PVST+: address by the protocol (one second)

- Blocking link on access-b will take 50 seconds to move to forwarding → return traffic black hole until then

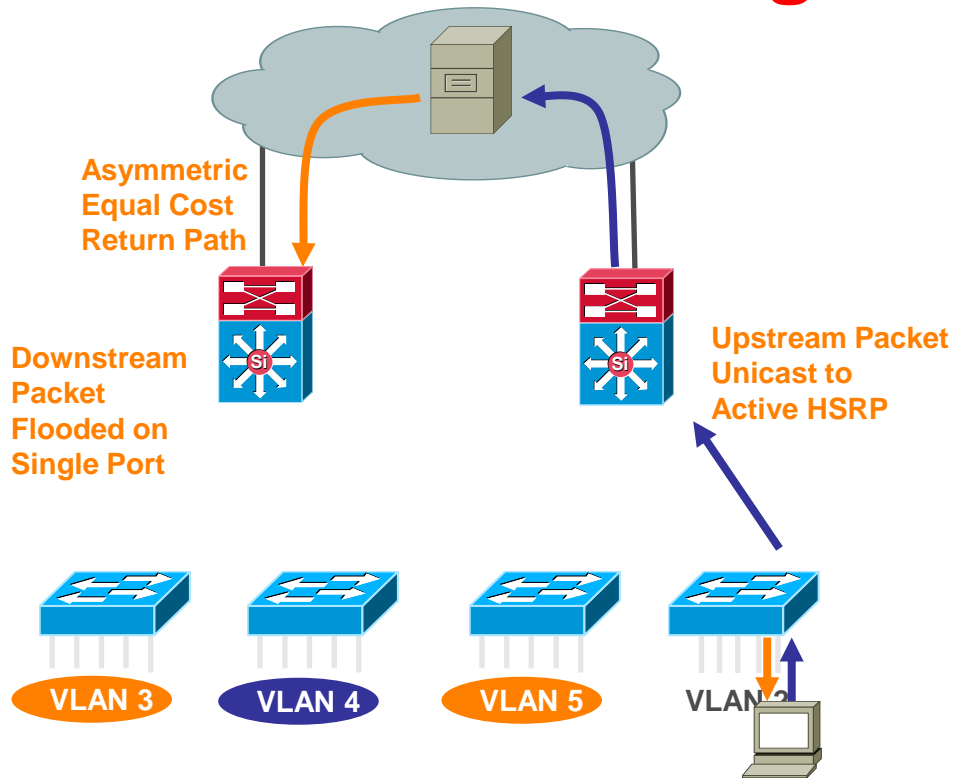
Asymmetric Routing (Unicast Flooding)

- 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



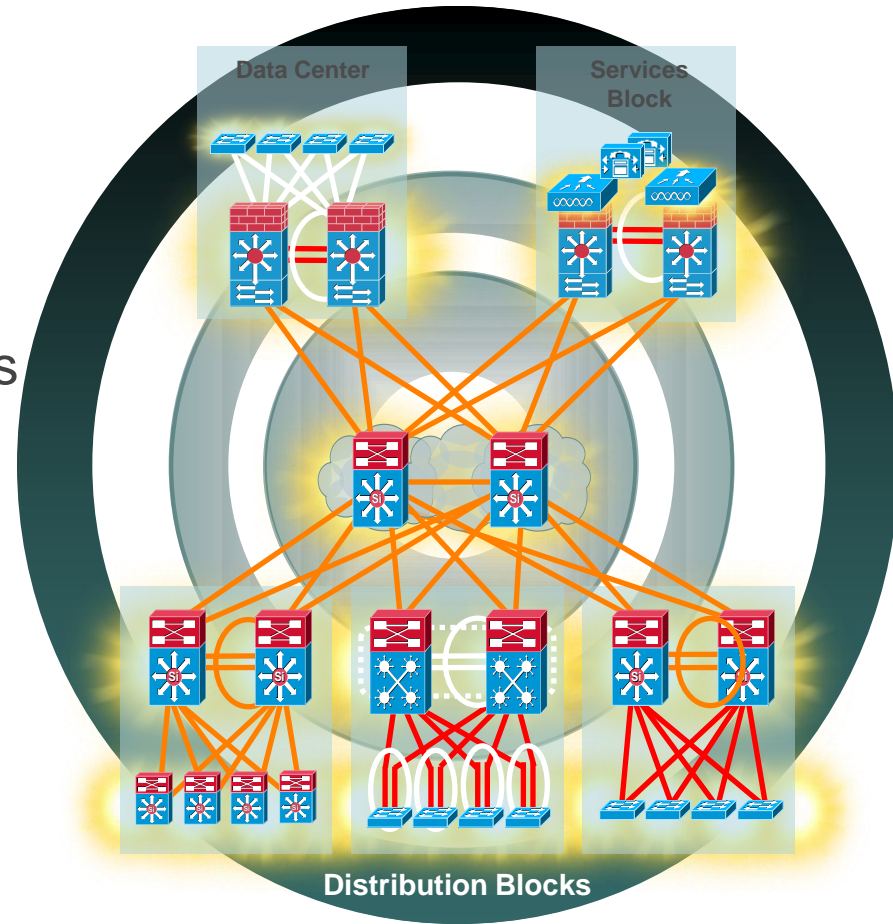
Best Practices Prevent Unicast Flooding

- 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



Agenda

- Multilayer Campus Design Principles
- Foundation Services
- Campus Design Best Practices
- **IP Telephony Considerations**
- QoS Considerations
- Security Considerations
- Putting It All Together
- Summary



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

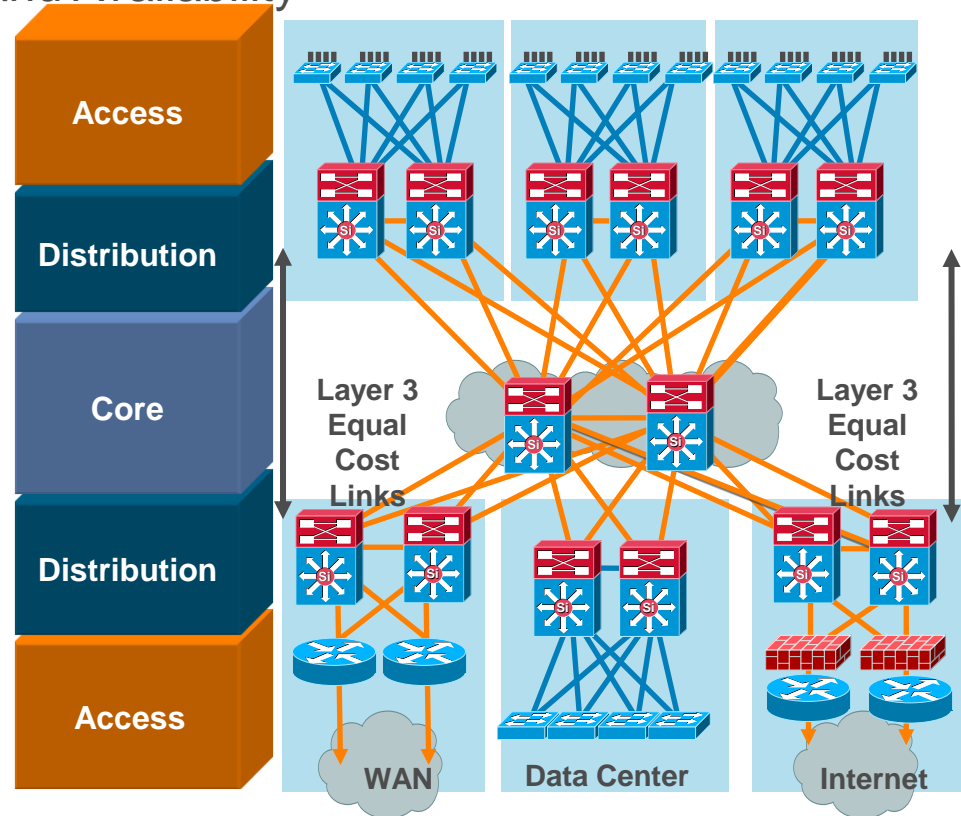
Policy enforcement

QoS: scheduling, trust boundary and classification

Core

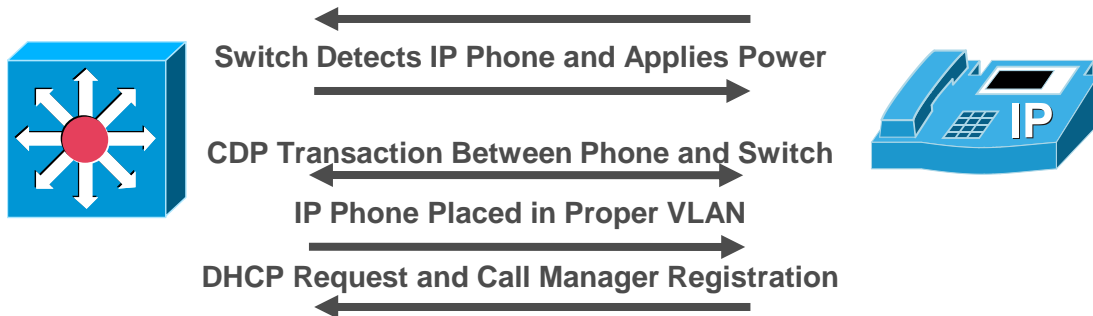
High availability, redundancy, fast convergence

QoS: scheduling, trust boundary



Infrastructure Integration

Extending the Network Edge

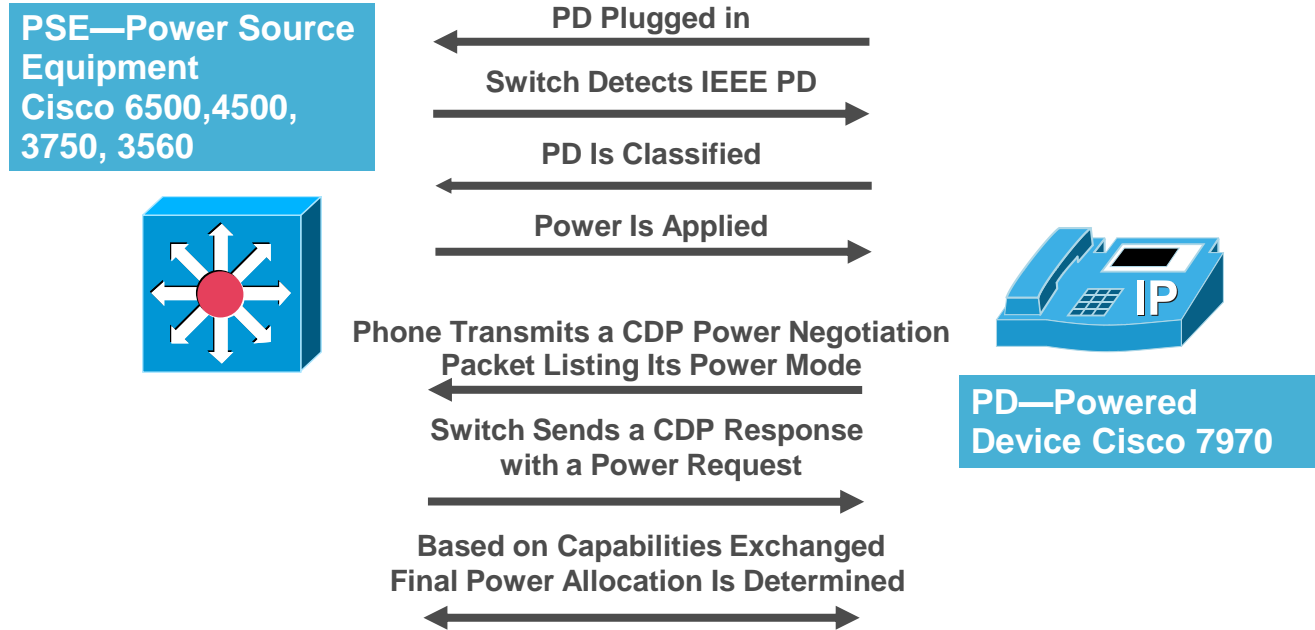


Phone Contains a Three-Port Switch that Is Configured in Conjunction with the Access Switch and CallManager

1. Power negotiation
2. VLAN configuration
3. 802.1x interoperation
4. QoS configuration
5. DHCP and CallManager registration

Enhanced Power Negotiation

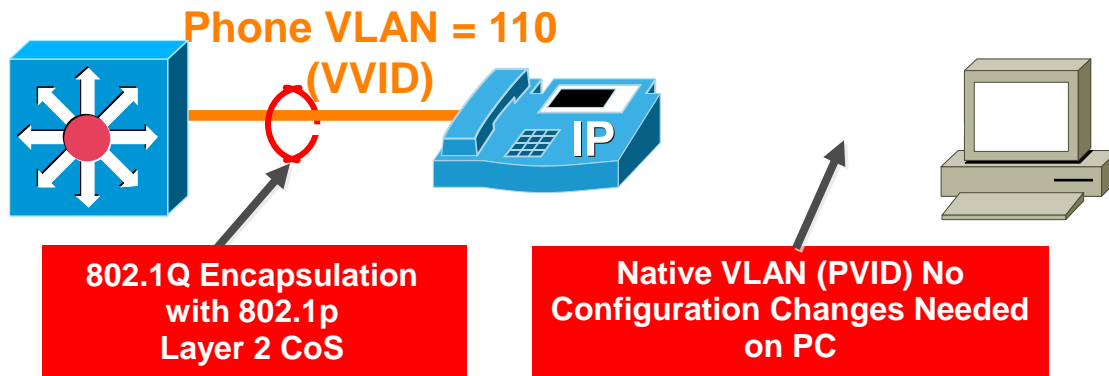
802.3af Plus Bidirectional CDP (Cisco 7970)



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

Infrastructure Integration: Next Steps

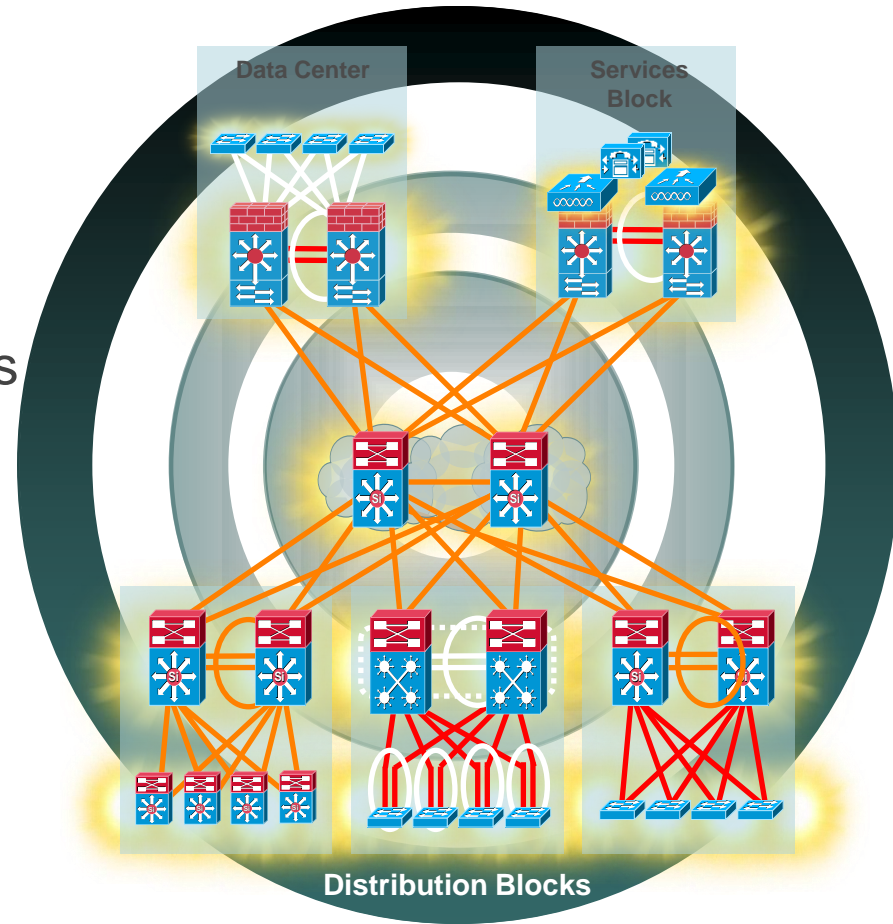
VLAN, QoS, and 802.1x Configuration



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

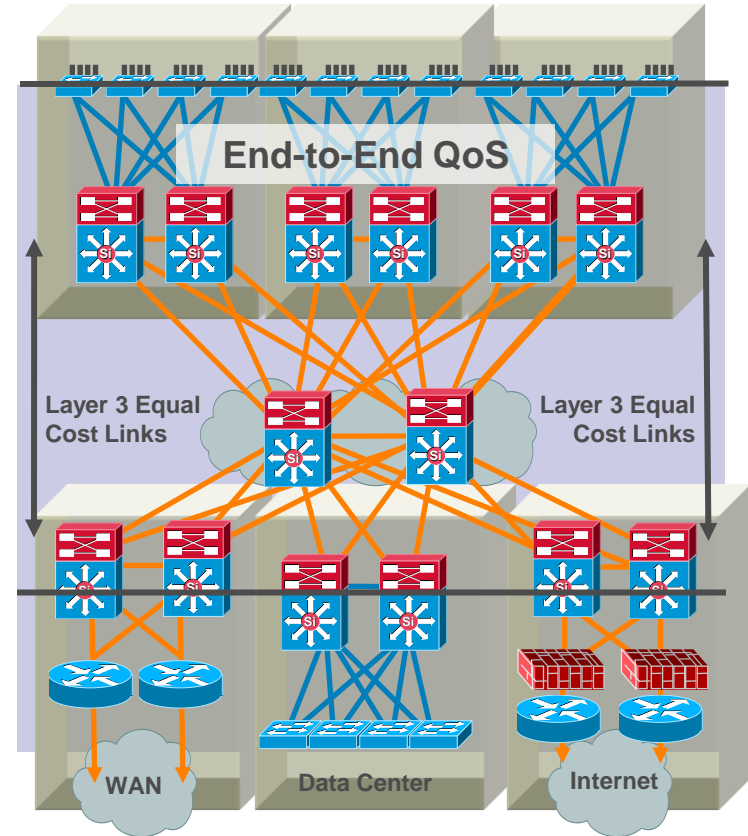
Agenda

- Multilayer Campus Design Principles
- Foundation Services
- Campus Design Best Practices
- IP Telephony Considerations
- **QoS Considerations**
- Security Considerations
- Putting It All Together
- Summary

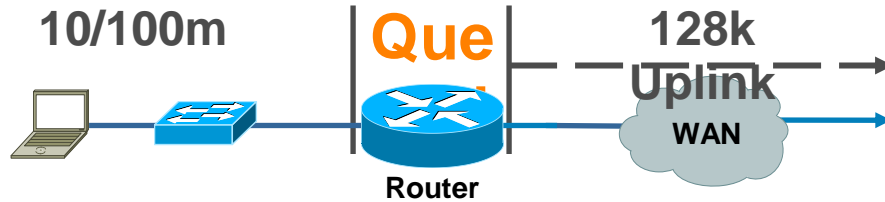


Best Practices—Quality of Service

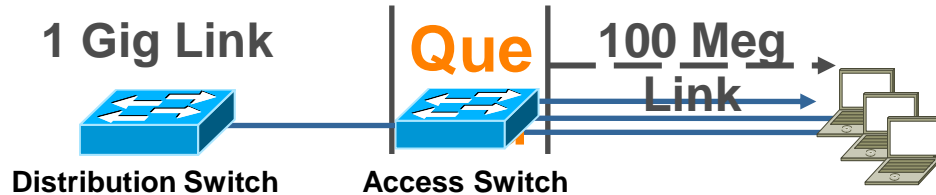
- Must be deployed end-to-end to be effective; all layers play different but equal roles
- Ensure that mission-critical applications are not impacted by link or transmit queue congestion
- Aggregation and rate transition points must enforce QoS policies
- Multiple queues with configurable admission criteria and scheduling are required



Transmit Queue Congestion



100 Meg in 128 Kb/S out—Packets Serialize in Faster than They Serialize Out
Packets **Queued** as They Wait to Serialize out Slower Link



1 Gig In 100 Meg out—Packets Serialize in Faster than They Serialize Out
Packets **Queued** as They Wait to Serialize out Slower Link


Auto QoS VoIP - Making It Easy ...

Configures QoS for VoIP on Campus Switches

```
Access-Switch(config-if)#auto qos voip ?  
cisco-phone      Trust the QoS marking of Cisco IP Phone  
cisco-softphone  Trust the QoS marking of Cisco IP SoftPhone  
trust            Trust the DSCP/CoS marking
```

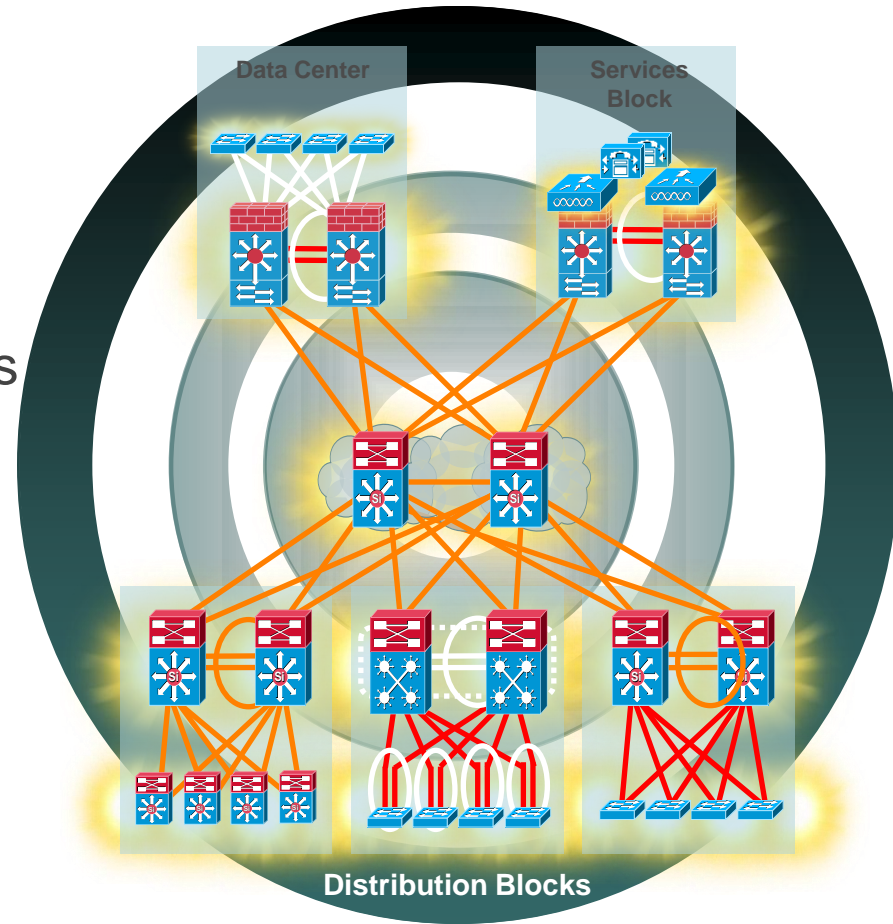
```
Access-Switch(config-if)#auto qos voip cisco-phone  
Access-Switch(config-if)#exit
```

```
!  
interface FastEthernet1/0/21  
srr-queue bandwidth share 10 10 60 20  
srr-queue bandwidth shape 10 0 0 0  
mls qos trust device cisco-phone  
mls qos trust cos  
auto qos voip cisco-phone  
end
```



Agenda

- Multilayer Campus Design Principles
- Foundation Services
- Campus Design Best Practices
- IP Telephony Considerations
- QoS Considerations
- **Security Considerations**
- Putting It All Together
- Summary



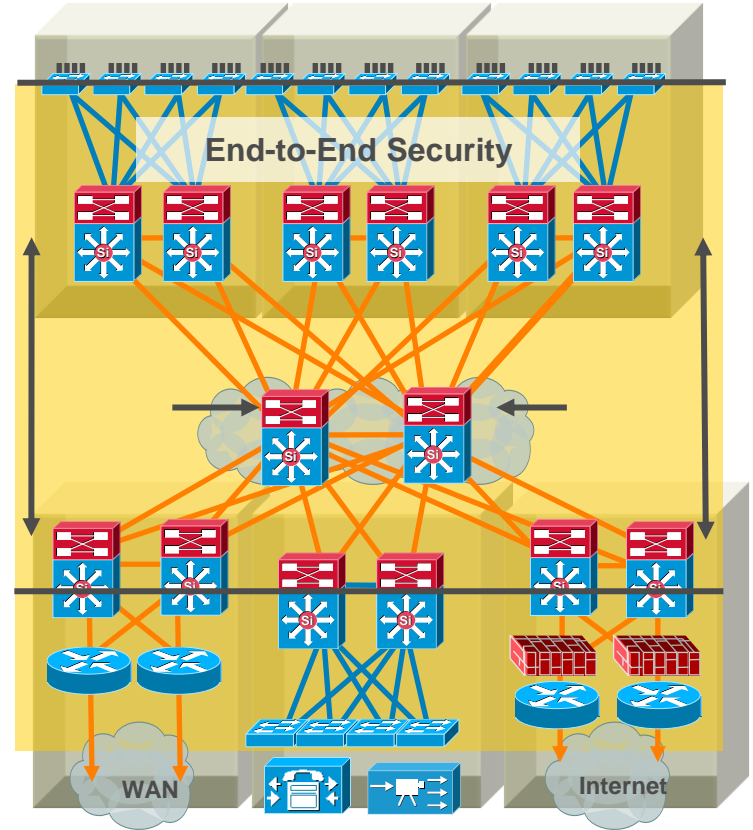
Best Practices—Campus Security

- **New stuff that we will cover!**

- Catalyst integrated security feature set!
- Dynamic port security, DHCP snooping, Dynamic ARP inspection, IP source guard

- **Things you already know—we won't cover...**

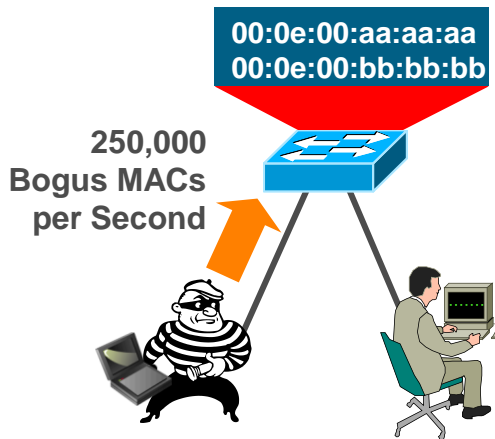
- Use SSH to access devices instead of Telnet
- Enable AAA and roles-based access control (RADIUS/TACACS+) for the CLI on all devices
- Enable SYSLOG to a server. Collect and archive logs
- When using SNMP use SNMPv3
- Disable unused services:
 - No service tcp-small-servers
 - No service udp-small-servers
- Use FTP or SFTP (SSH FTP) to move images and configurations around—avoid TFTP when possible
- Install VTY access-lists to limit which addresses can access management and CLI services
- Enable control plane protocol authentication where it is available (EIGRP, OSPF, BGP, HSRP, VTP, etc.)
- Apply basic protections offered by implementing RFC2827 filtering on external edge inbound interfaces



For More Details, See BRKSEC-2002 Session, Understanding and Preventing Layer 2 Attacks

Securing Layer 2 from Surveillance Attacks

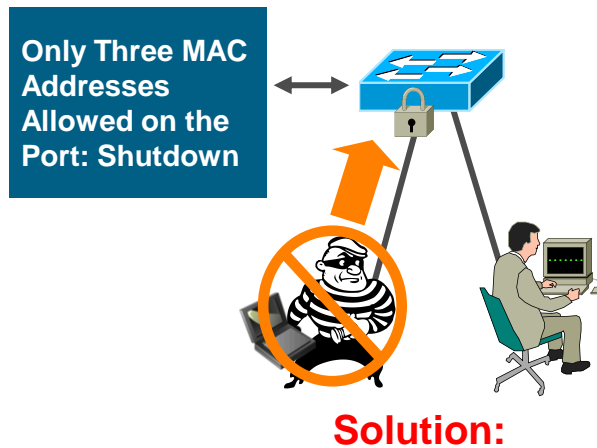
Cutting Off MAC-Based Attacks



Problem:

Script Kiddie Hacking Tools Enable Attackers Flood Switch CAM Tables with Bogus Macs; Turning the VLAN into a **Hu3** and Eliminating Privacy

Switch CAM Table Limit Is Finite Number of Mac Addresses

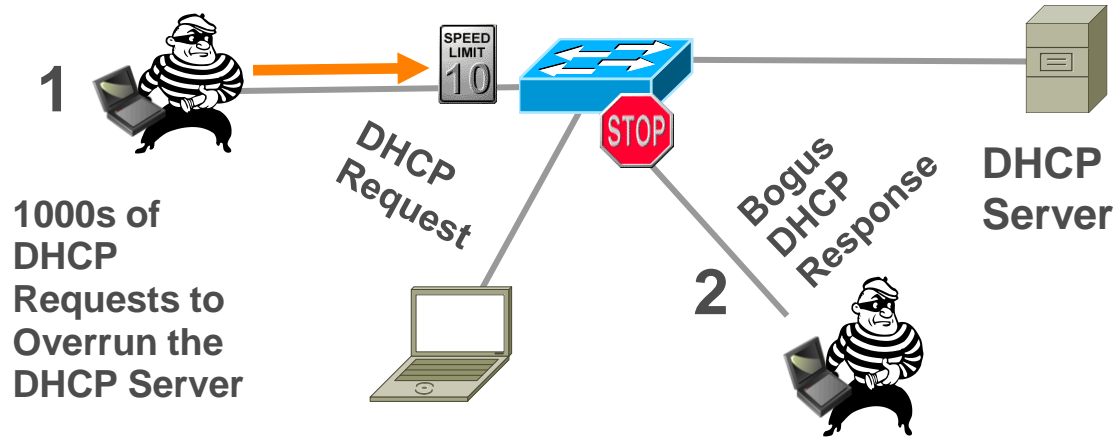


Port Security Limits MAC Flooding Attack and Locks Down Port and Sends an SNMP Trap

```
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

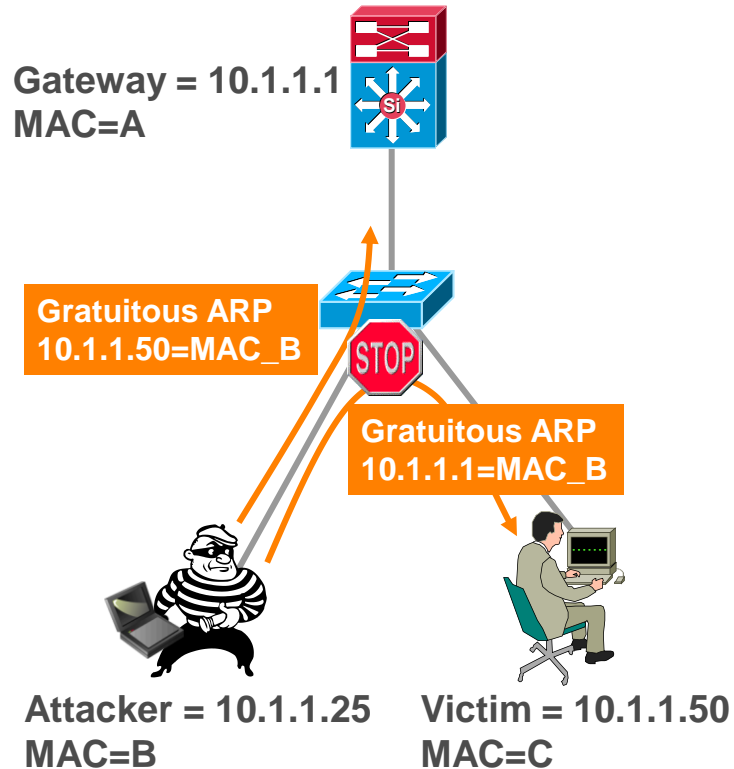


- DHCP requests (discover) and responses (offer) tracked
- Rate-limit requests on trusted interfaces; limits DoS attacks on DHCP server
- Deny responses (offers) on non trusted interfaces; stop malicious or errant DHCP server

Securing Layer 2 from Surveillance Attacks

Protection Against ARP Poisoning

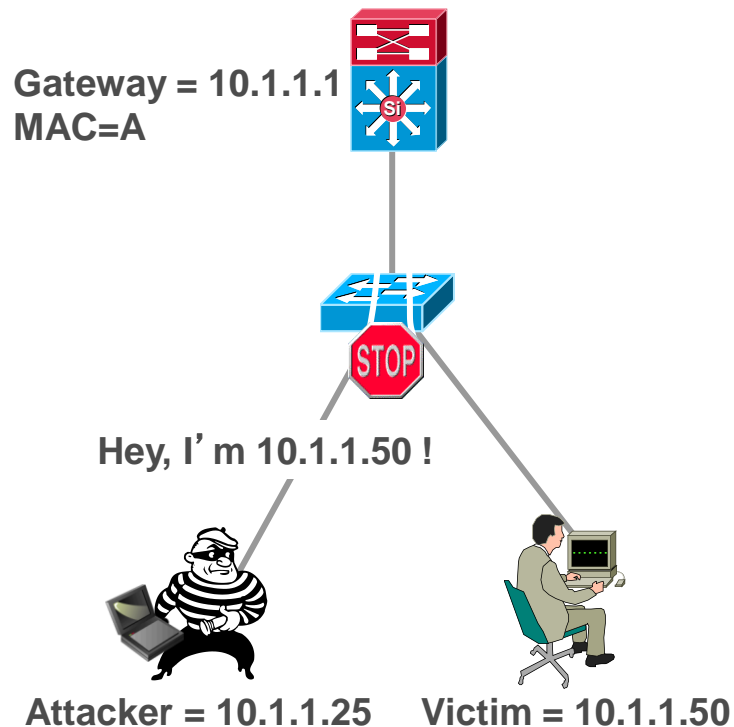
- Dynamic ARP inspection **protects against ARP poisoning** (ettercap, dsnif, arpspoof)
- Uses the DHCP snooping binding table
- Tracks MAC to IP from DHCP transactions
- Rate-limits ARP requests from client ports; stop port scanning
- Drop **bogus** gratuitous ARPs; stop ARP poisoning/MIM attacks



IP Source Guard

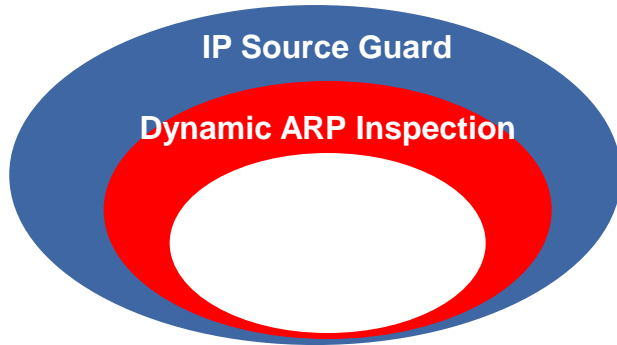
Protection Against Spoofed IP Addresses

- 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 Cisco IOS

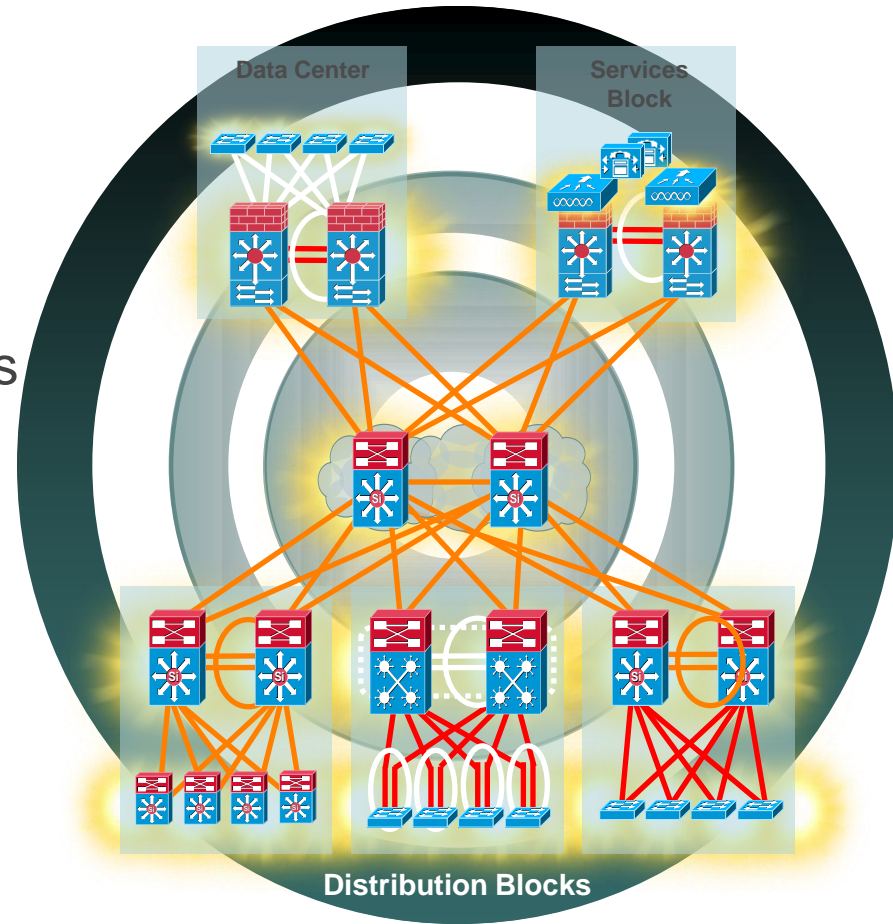


- Port security prevents MAC flooding attacks
- DHCP snooping prevents client attack on the switch and server
- Dynamic ARP Inspection adds security to ARP using DHCP snooping table
- IP source guard adds security to IP source address using DHCP snooping table

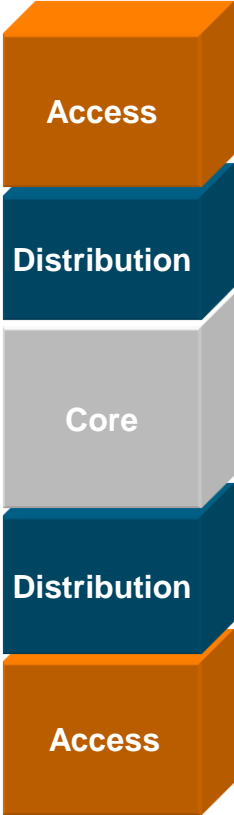
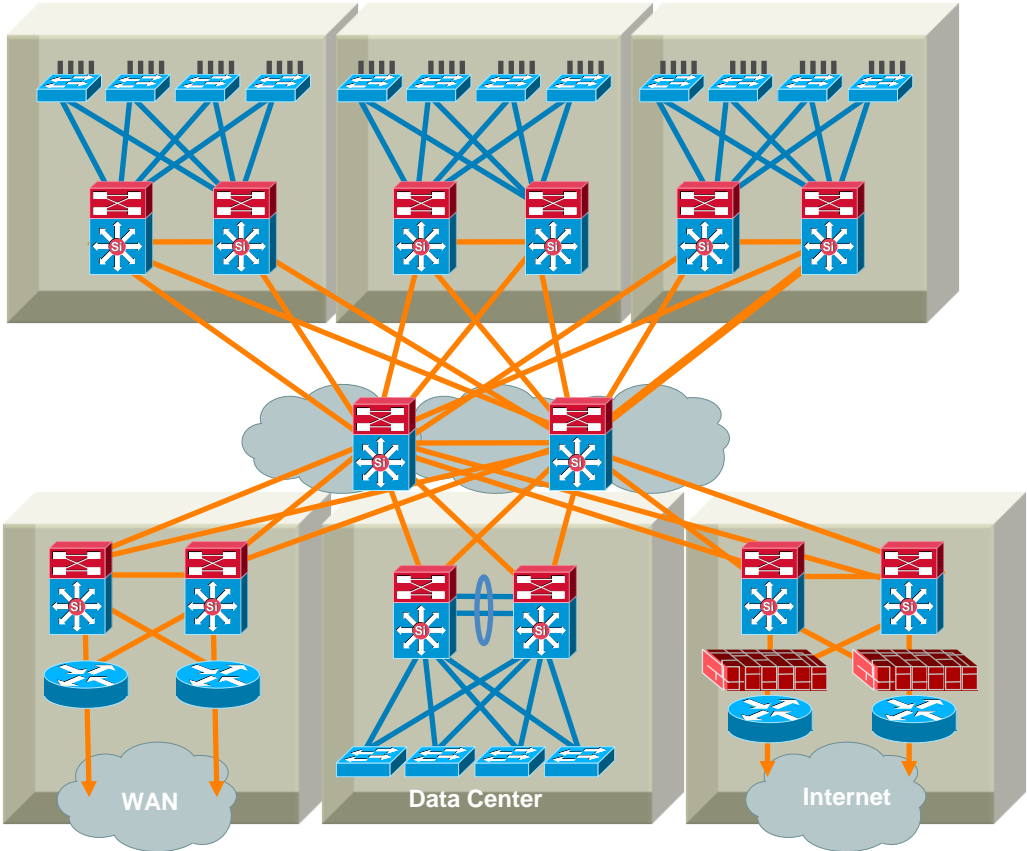
```
ip dhcp snooping
ip dhcp snooping vlan 2-10
ip arp inspection vlan 2-10
!
interface fa3/1
switchport port-security
switchport port-security max 3
switchport port-security violation
restrict
switchport port-security aging time 2
switchport port-security aging type
inactivity
ip arp inspection limit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

Agenda

- Multilayer Campus Design Principles
- Foundation Services
- Campus Design Best Practices
- IP Telephony Considerations
- QoS Considerations
- Security Considerations
- Putting It All Together
- **Summary**



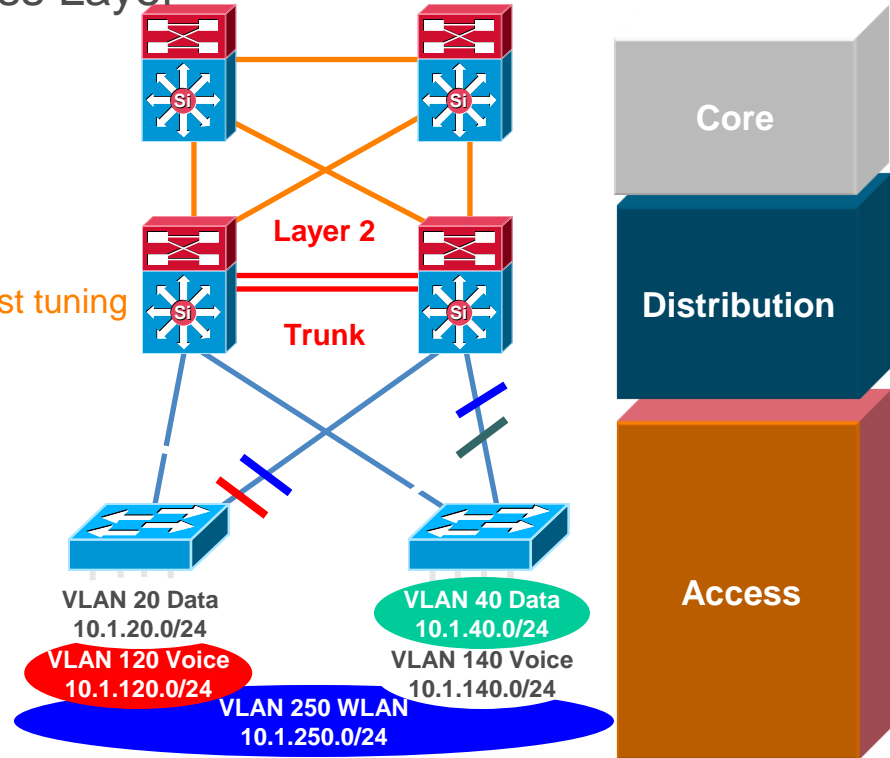
Hierarchical Campus



Layer 2 Distribution Interconnection

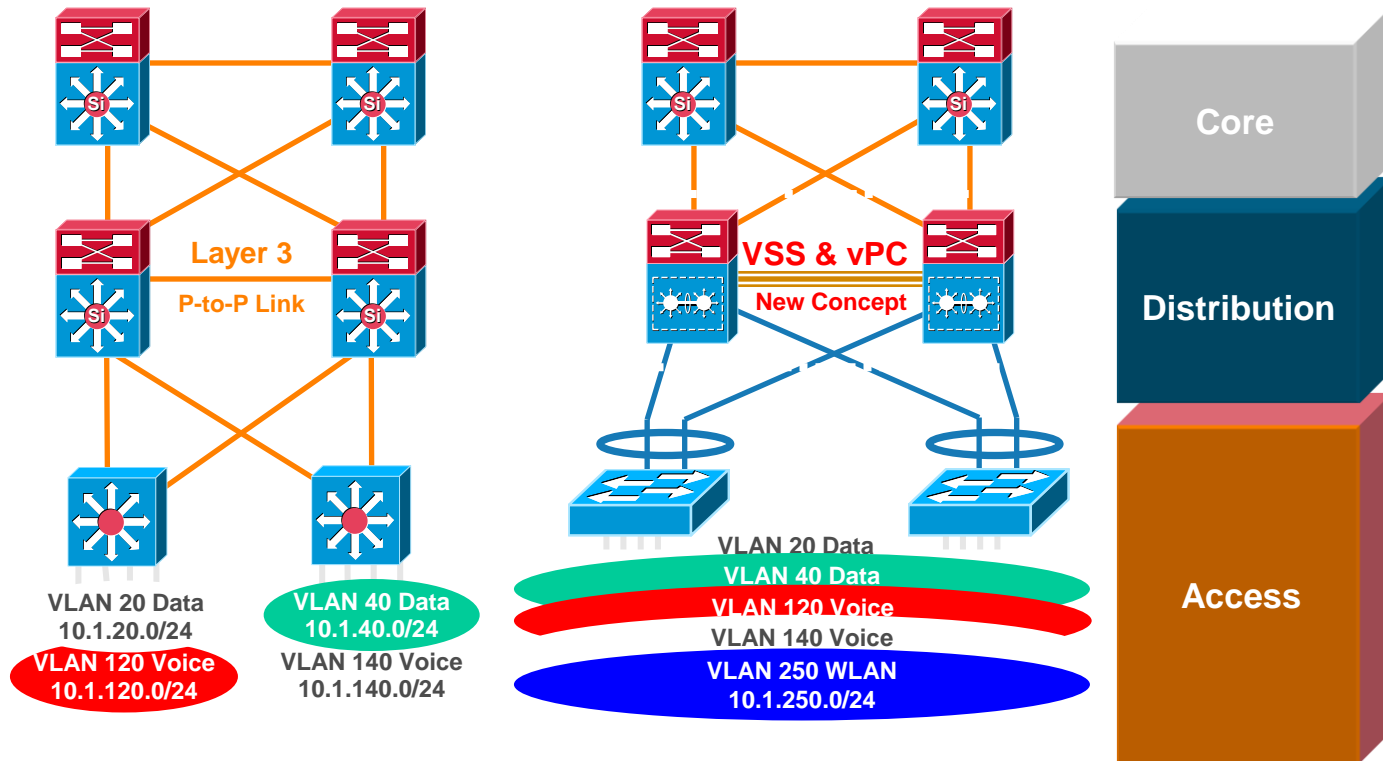
Layer 2 Access—Some VLANs Span Access Layer

- Tune CEF load balancing
- Match CatOS/IOS EtherChannel settings and tune load balancing
- Summarize routes towards core
- Limit redundant IGP peering
- **STP Root and HSRP primary or GLBP and STP port cost tuning to load balance on uplinks**
- Set trunk mode on/no-negotiate
- Disable EtherChannel unless needed
- **RootGuard on downlinks**
- **LoopGuard on uplinks**
- Set port host on access Layer ports:
 - Disable trunking
 - Disable EtherChannel
 - Enable PortFast
- RootGuard or BPDU-Guard
- Use security features



Routed Access and Virtual Switching System

Evolutions of and Improvements to Existing Designs



See BRK-CRS3035—Advanced Enterprise Campus Design: VSS

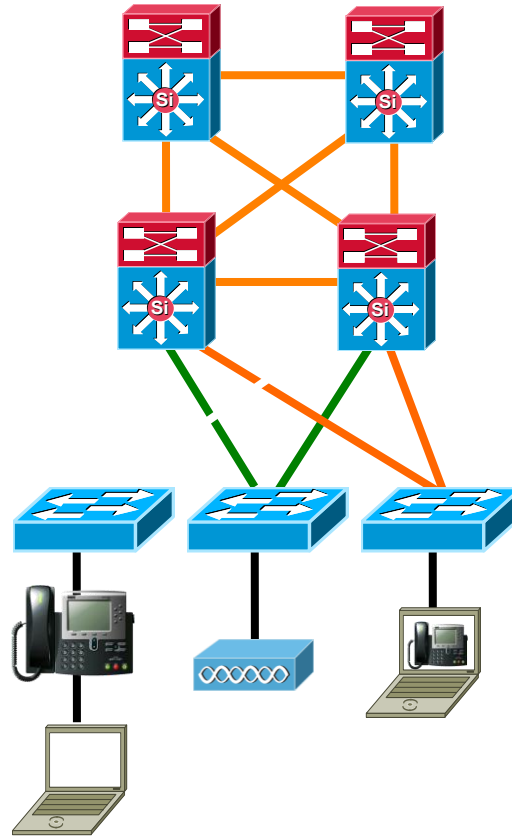
See BRK-CRS3036—Advanced Enterprise Campus Design: Routed Access

SmartPorts—Predefined Configurations

```
Access-Switch#show parser macro brief
  default global      : cisco-global
  default interface: cisco-desktop
  default interface: cisco-phone
  default interface: cisco-switch
  default interface: cisco-router
  default interface: cisco-wireless
```

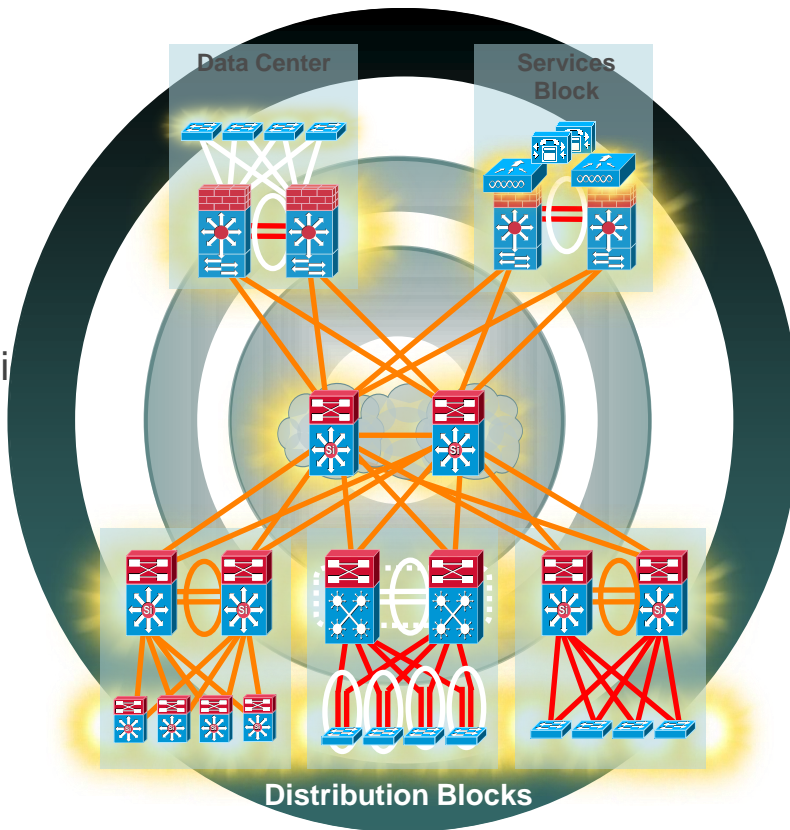
```
Access-Switch(config-if)# macro apply cisco-phone
$access_vlan 100 $voice_vlan 10
```

```
Access-Switch#show run int fa1/0/19
!
interface FastEthernet1/0/19
switchport access vlan 100
switchport mode access
switchport voice vlan 10
switchport port-security maximum 2
switchport port-security
switchport port-security aging time 2
switchport port-security violation restrict
switchport port-security aging type inactivity
srr-queue bandwidth share 10 10 60 20
srr-queue bandwidth shape 10 0 0 0
mls qos trust device cisco-phone
mls qos trust cos
  macro description cisco-phone
  auto qosvoipcisco-phone
  spanning-tree portfast
  spanning-tree bpduguard enable
end
```



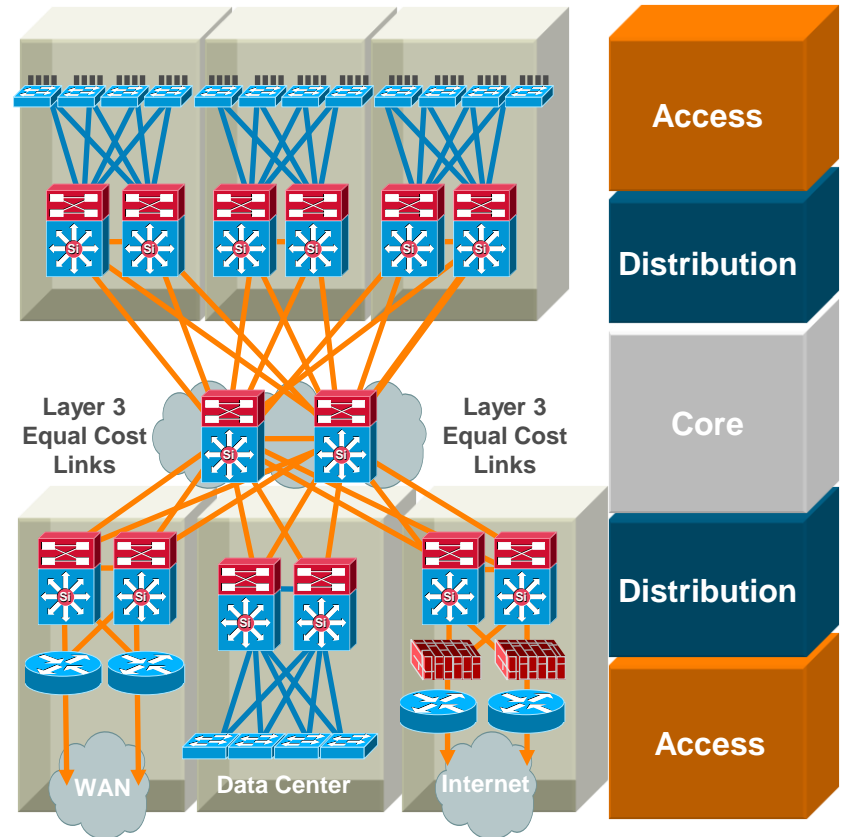
Agenda

- Multilayer Campus Design Principles
- Foundation Services
- Campus Design Best Practices
- IP Telephony Considerations
- QoS Considerations
- Security Considerations
- Putting It All Together
- Summary



Summary

- Offers hierarchy—each layer has specific role
- Modular topology—building blocks
- Easy to grow, understand, and troubleshoot
- Creates small fault domains—clear demarcations and isolation
- Promotes load balancing and redundancy
- Promotes deterministic traffic patterns
- Incorporates balance of both Layer 2 and Layer 3 technology, leveraging the strength of both
- Utilizes Layer 3 routing for load balancing, fast convergence, scalability, and control



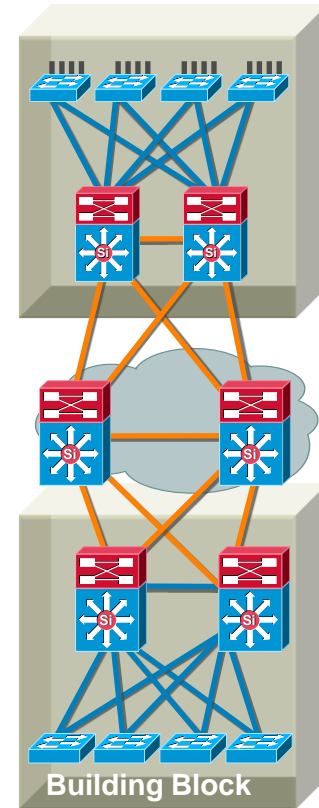
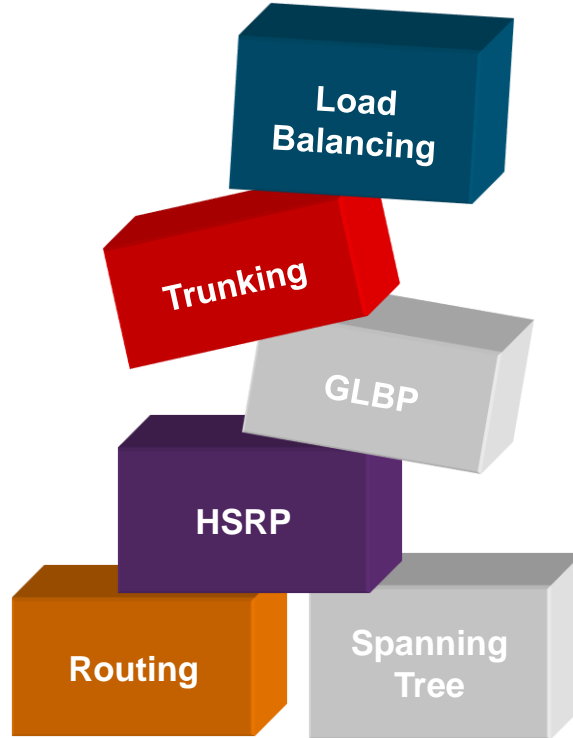
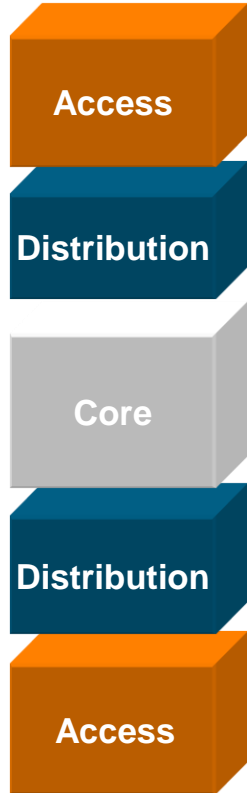
What Questions do you have?

BUILT FOR
THE HUMAN
NETWORK



Hierarchical Network Design

Without a Rock Solid Foundation the Rest Doesn't Matter



BUILT FOR
THE HUMAN
NETWORK



Q&A



We value your feedback.

Please be sure to complete the Evaluation Form for this session.



Access today's presentations at cisco.com/ca/plus



Follow [@CiscoCanada](https://twitter.com/CiscoCanada) and join the [#CiscoPlusCA](https://twitter.com/hashtag/CiscoPlusCA) conversation