Agenda

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

High-Availability Campus Design

Structure, Modularity, and Hierarchy

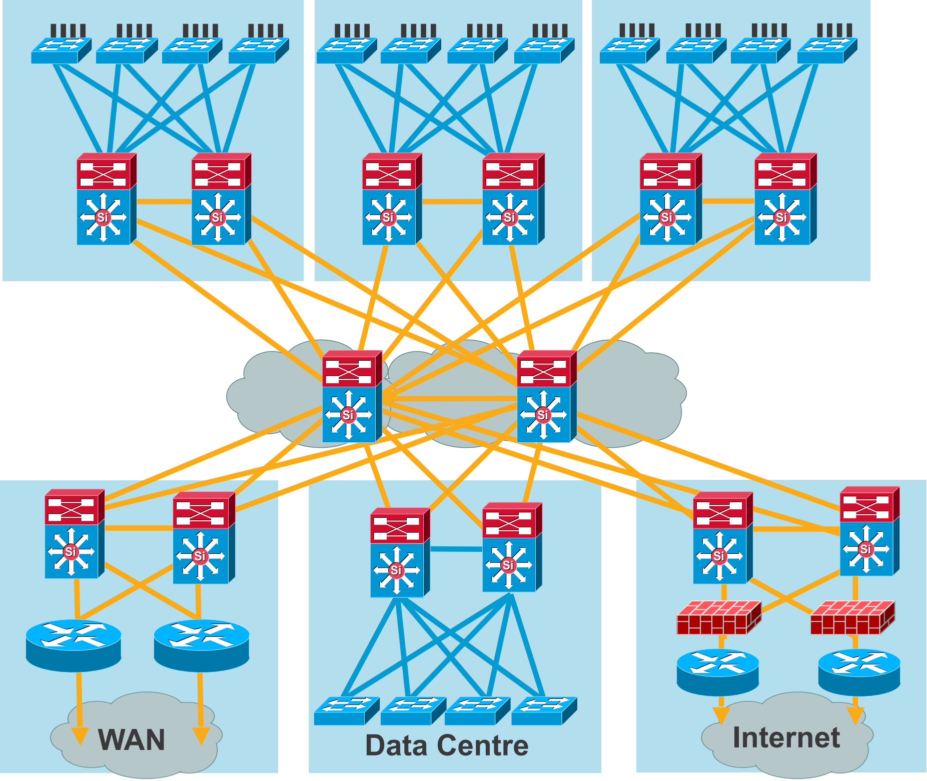
Access

Distribution

Core

Distribution

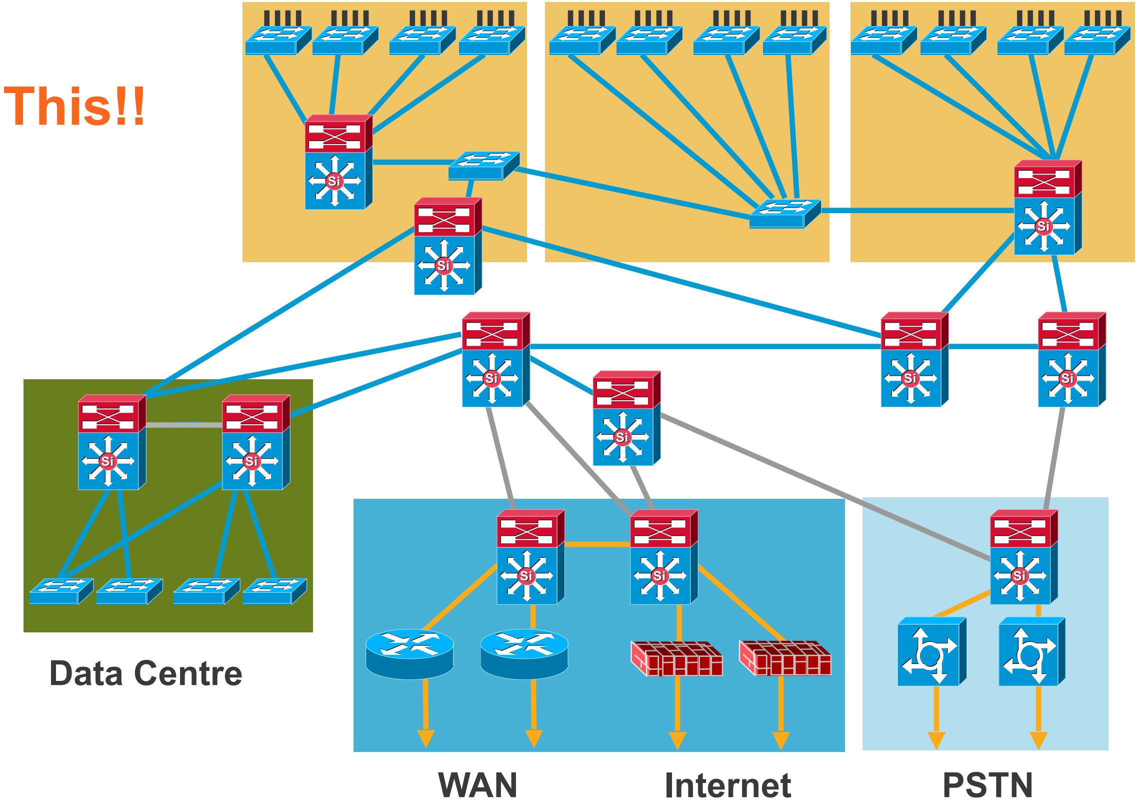
Access



Hierarchical Campus Network

Structure, Modularity and Hierarchy

Not This!!



Hierarchical Network Design

Without a Rock Solid Foundation the Rest Doesn’t Matter

Access

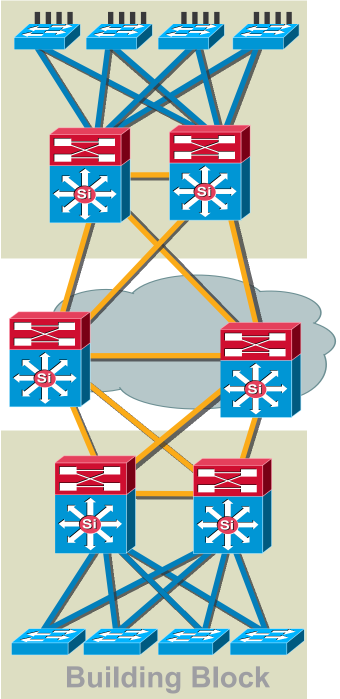
Distribution

Core

Distribution

Access

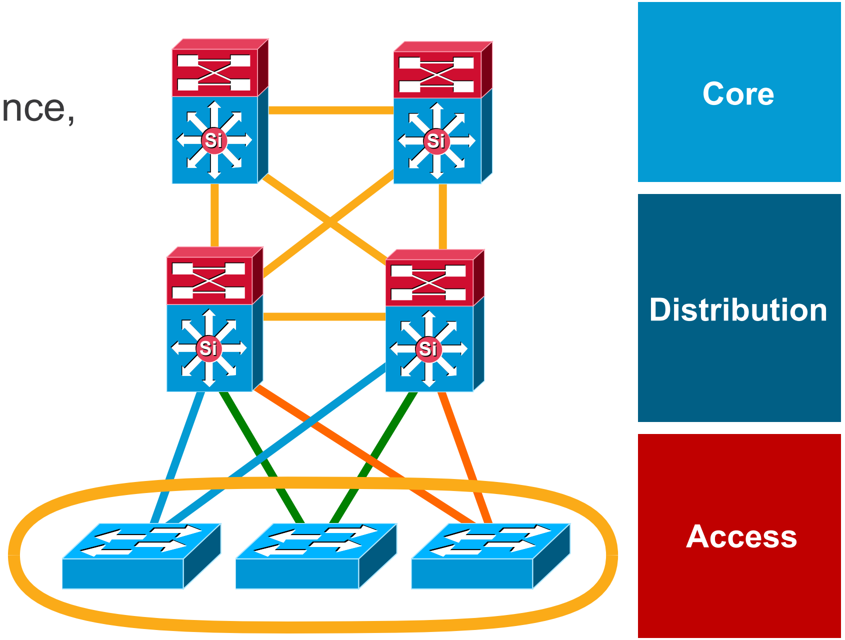
- Offers hierarchy—each layer has specific role
- Modular topology—building blocks
- Easy to grow, understand, and troubleshoot
- Creates small fault domains— clear demarcations and isolation
- Promotes load balancing and redundancy
- Promotes deterministic traffic patterns
- Incorporates balance of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



Access Layer

Feature Rich Environment

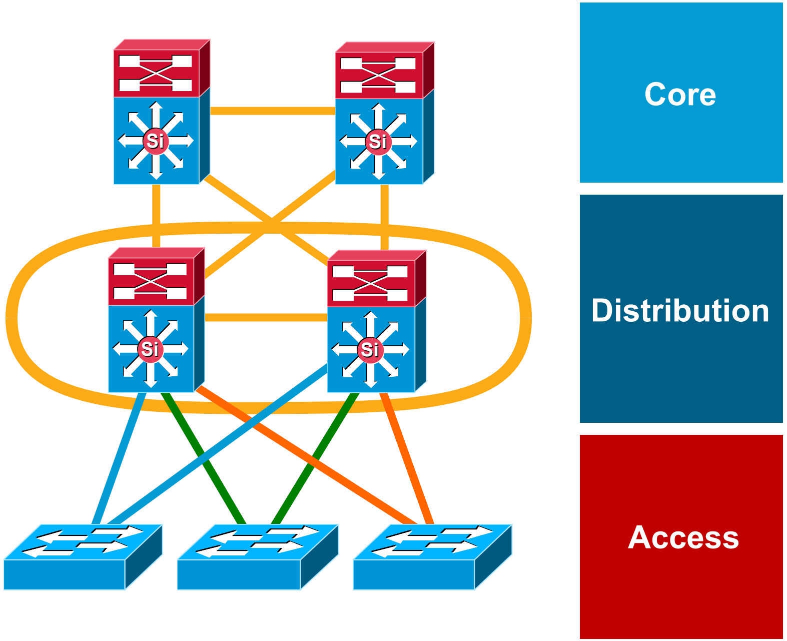
- It’s not just about connectivity
- Layer 2/Layer 3 feature rich environment; convergence, HA, security, QoS, IP multicast, etc.
- Intelligent network services: QoS, trust boundary, broadcast suppression, IGMP snooping
- Intelligent network services: PVST+, Rapid PVST+, EIGRP, OSPF, DTP, PAgP/LACP, 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.



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 summarisation, fast convergence, redundant path load sharing
- HSRP or GLBP to provide first hop redundancy

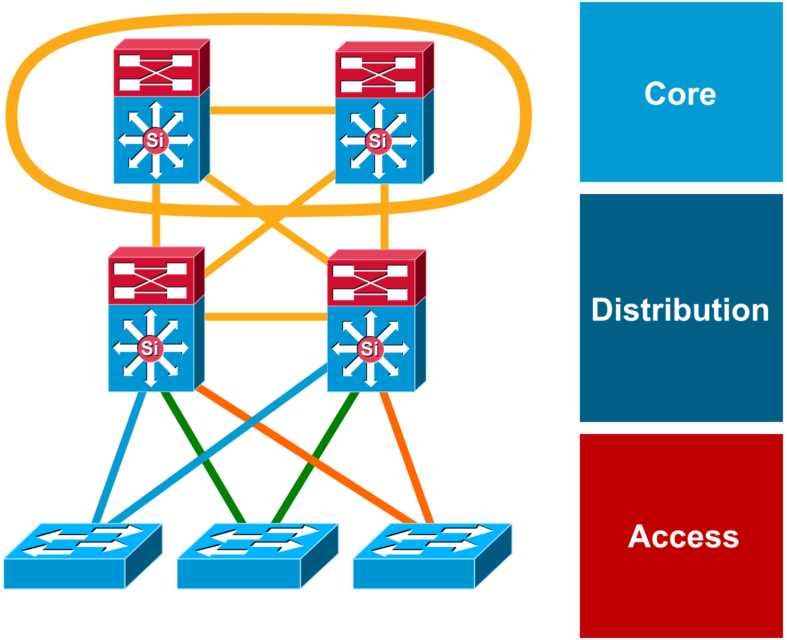


Cisco live!

Core Layer

Scalability, High Availability, and Fast Convergence

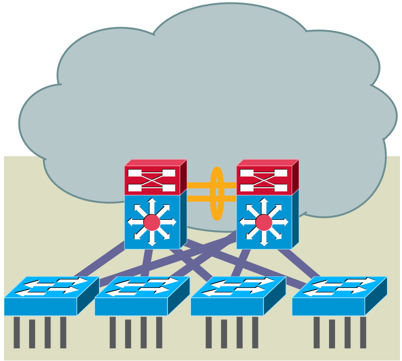
- Backbone for the network—connects network building blocks
- Performance and stability vs. complexity—less is more in the core
- Aggregation point for distribution layer
- Separate core layer helps in scalability during future growth
- Keep the design technology-independent



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

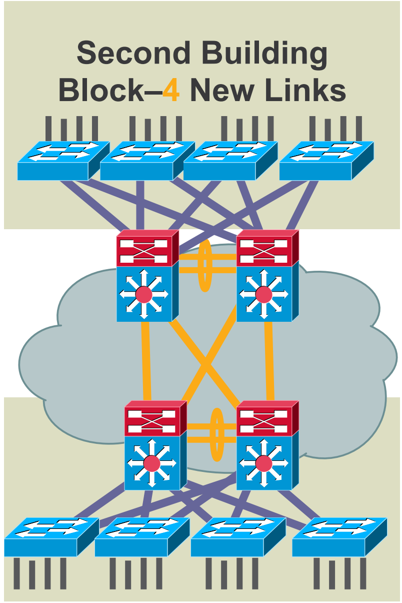


Cisco live!

Do I Need a Core Layer?

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

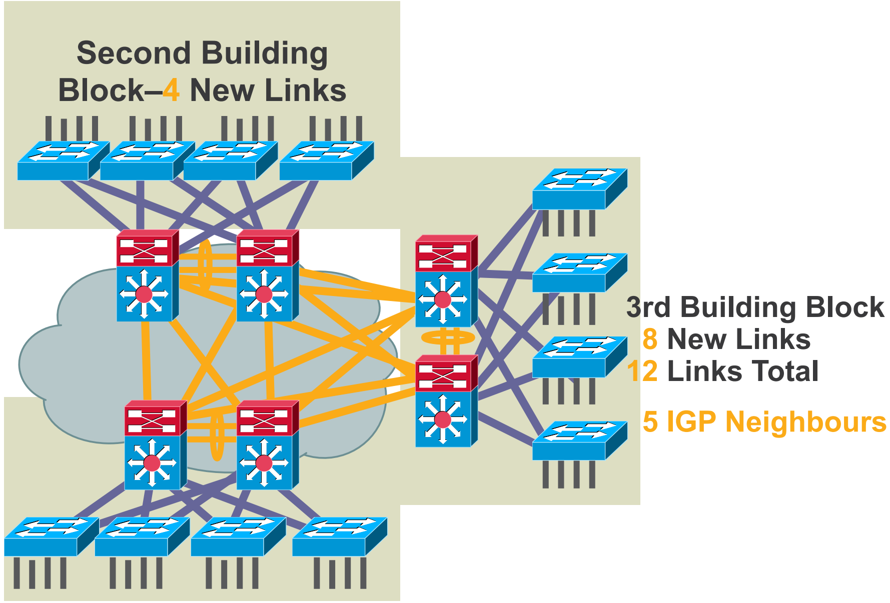
- No Core
- Fully-meshed distribution layers
- Physical cabling requirement
- Routing complexity



Do I Need a Core Layer?

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

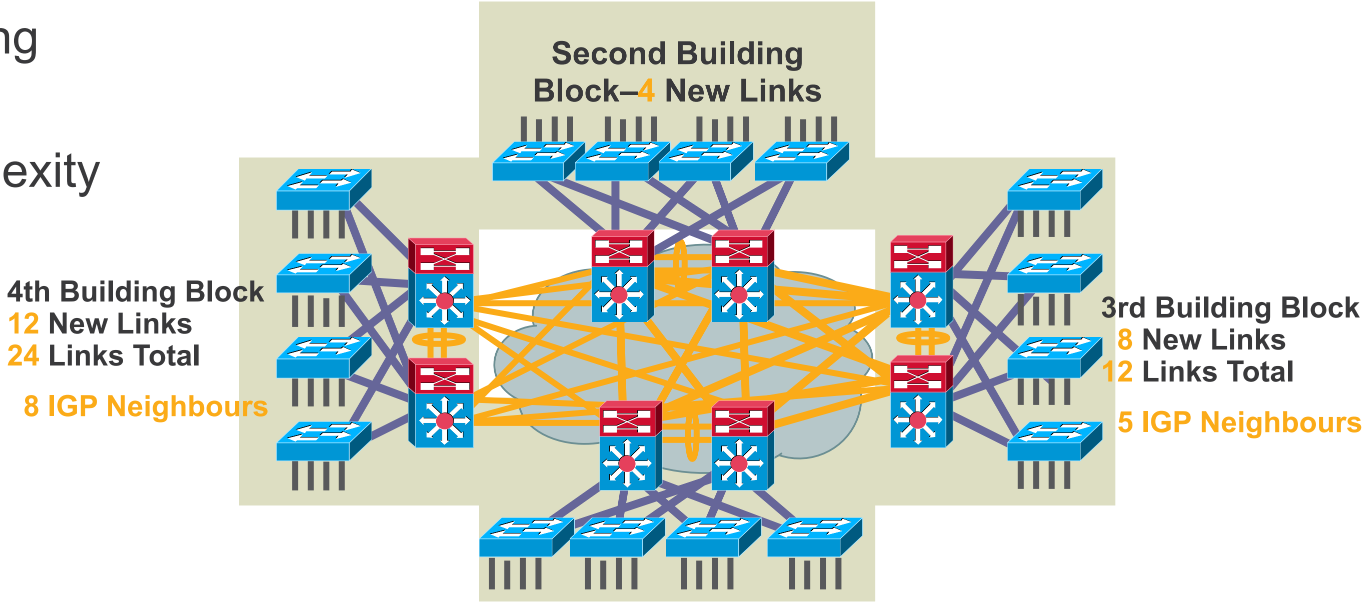
- No Core
- Fully-meshed distribution layers
- Physical cabling requirement
- Routing complexity



Do I Need a Core Layer?

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

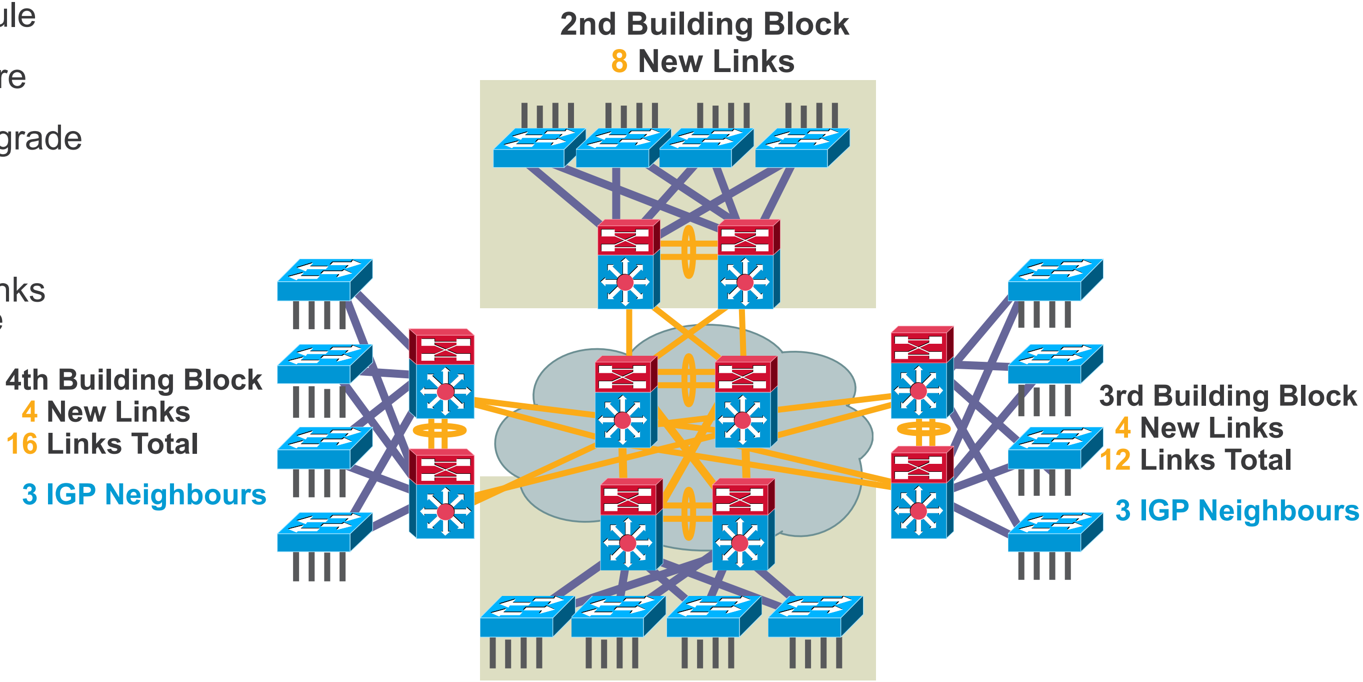
- No Core
- Fully-meshed distribution layers
- Physical cabling requirement
- Routing complexity



Do I Need a Core Layer?

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

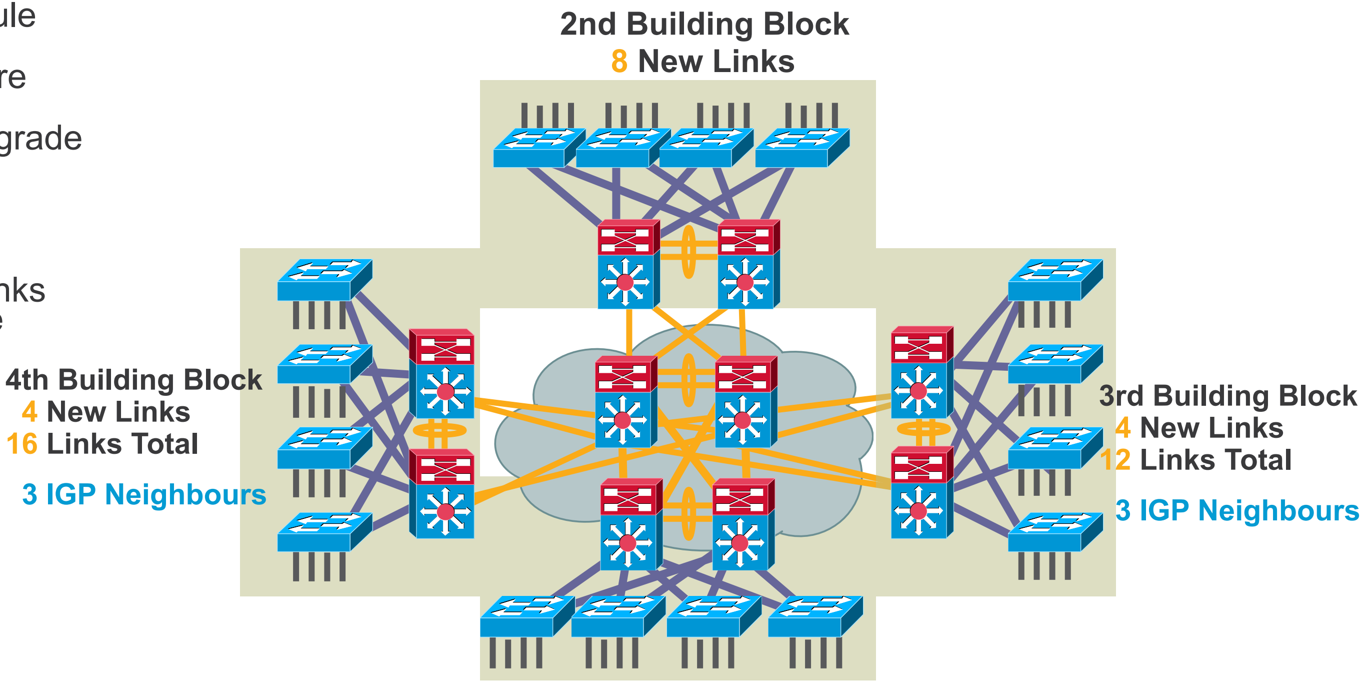
- Dedicated Core Switches
- Easier to add a module
- Fewer links in the core
- Easier bandwidth upgrade
- Routing protocol peering reduced
- Equal cost Layer 3 links for best convergence



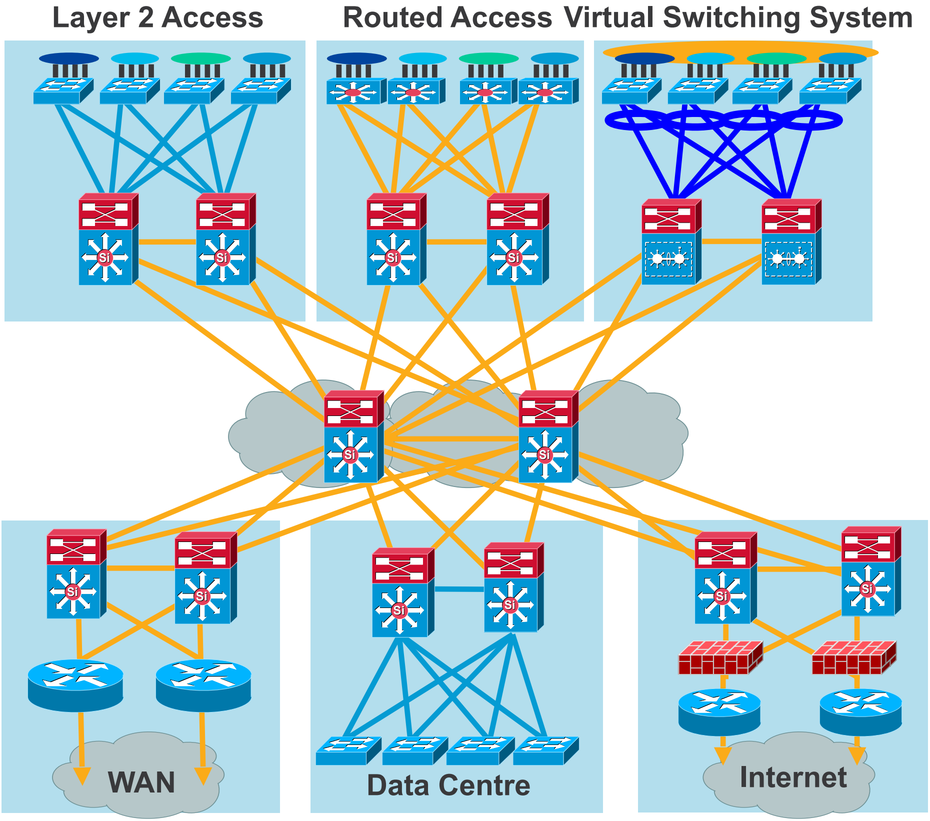
Do I Need a Core Layer?

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

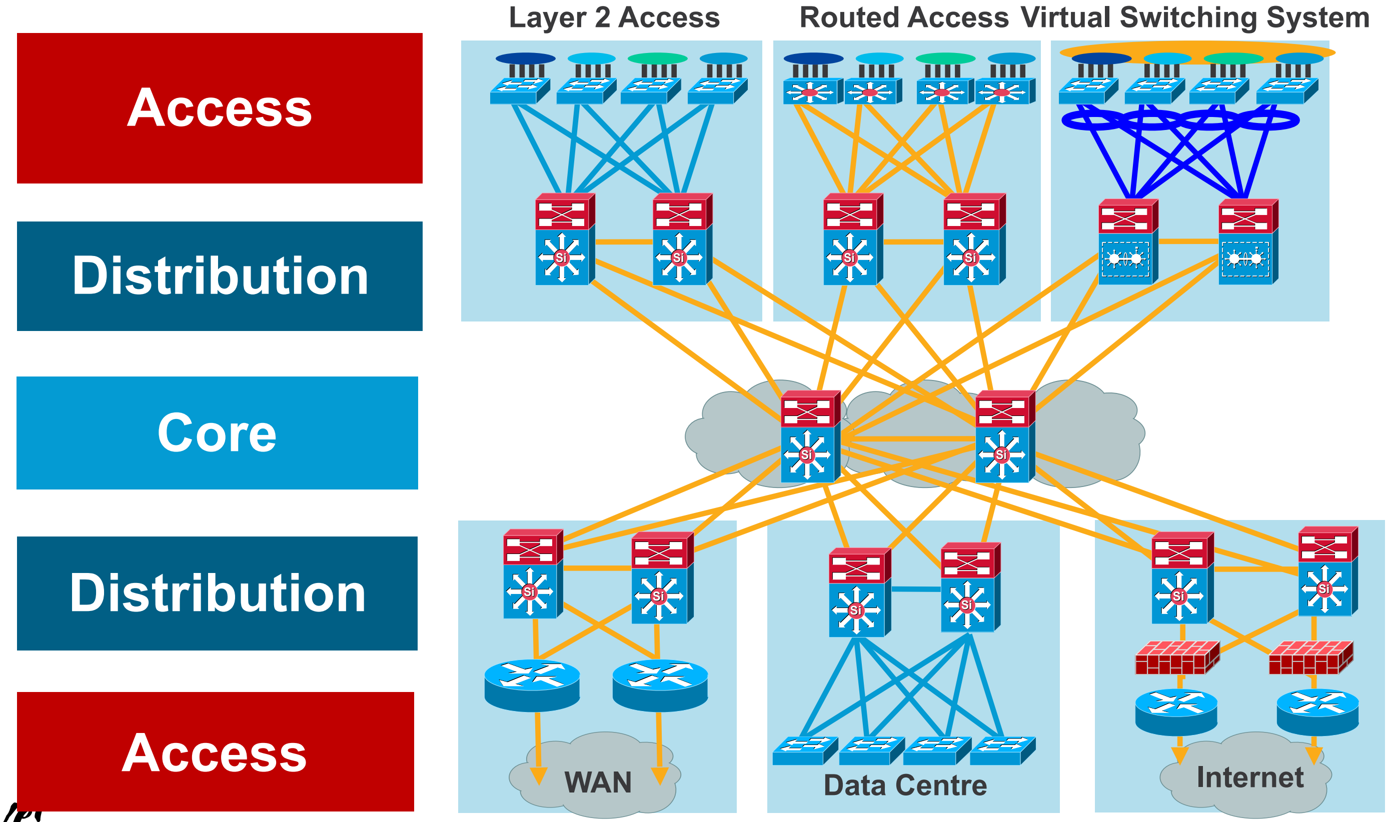
- Dedicated Core Switches
- Easier to add a module
- Fewer links in the core
- Easier bandwidth upgrade
- Routing protocol peering reduced
- Equal cost Layer 3 links for best convergence



Design Alternatives Come Within a Building (or Distribution) Block



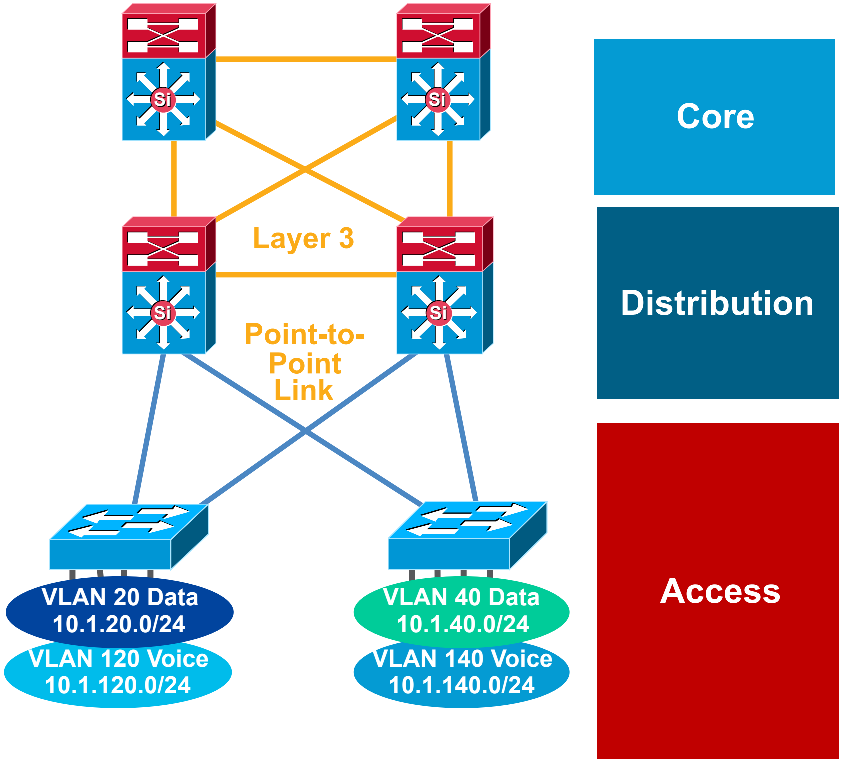
Design Alternatives Come Within a Building (or Distribution) Block



Layer 3 Distribution Interconnection

Layer 2 Access—No VLANs Span Access Layer

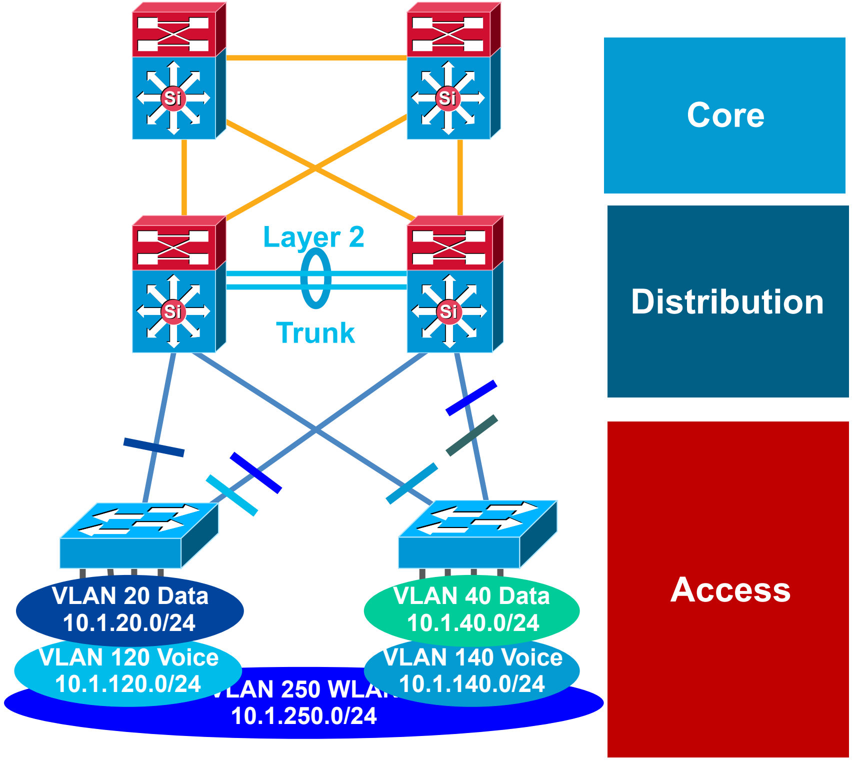
- Tune CEF load balancing
- Summarise routes towards core
- Limit redundant IGP peering
- STP Root and HSRP primary tuning or GLBP to load balance on uplinks
- Set trunk mode on/no-negotiate
- Disable Ether Channel unless needed
- Set port host on access layer ports:
  - Disable trunking
    Disable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



Layer 2 Distribution Interconnection

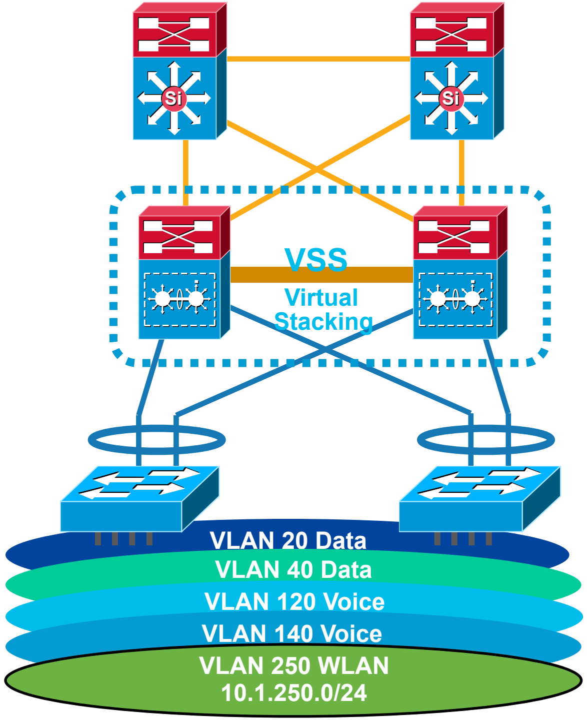
Layer 2 Access—Some VLANs Span Access Layer

- Tune CEF load balancing
- Summarise 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 Ether Channel unless needed
- RootGuard on downlinks
- LoopGuard on uplinks
- Set port host on access Layer ports:
  - Disable trunking
    Disable Ether Channel
    Enable PortFast
- RootGuard or BPDU-Guard
- Use security features



Virtual Switching System & Virtual Stacking

L2 with-out a STP Liability



- Tune CEF load balancing
- Summarise routes towards core
- Limit redundant IGP peering
- Set trunk mode on/no-negotiate
- MUST Ether Channel else blocked ports
- Set port host on access layer ports:
  - Disable trunking
    Disable Ether Channel
    Enable PortFast
- RootGuard or BPDU-Guard
- Use security features

Core

Distribution

Access

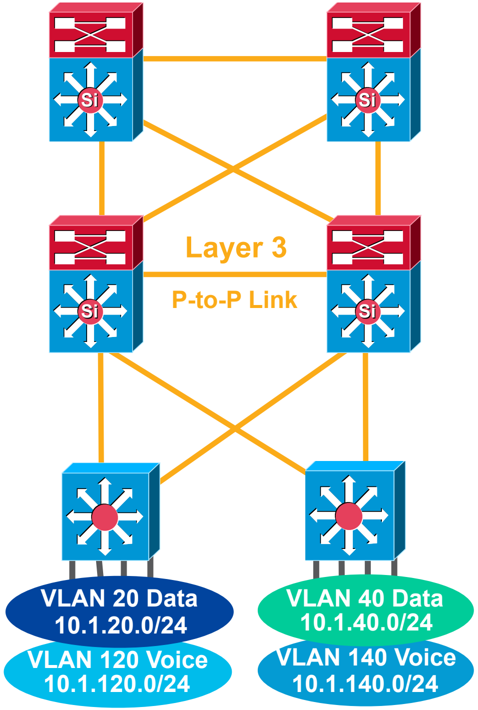
Routing to the Edge
Advantages, Yes in the Right Environment

Advantages:

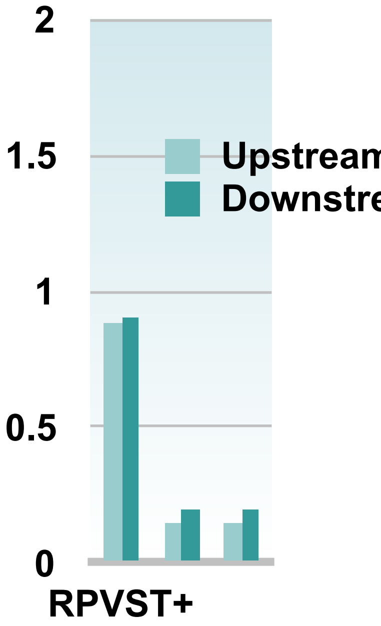
- Ease of implementation, less to get right
  - No matching of STP/HSRP/GLBP priority
  - No L2/L3 Multicast topology inconsistencies
- Single Control Plane and well known tool set
  - traceroute, show ip route, show ip eigrp neighbour, etc.…
- Most Catalysts support L3 Switching today
- EIGRP converges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

Considerations:

- Do you have any Layer 2 VLAN adjacency requirements between access switches?
- IP addressing—Do you have enough address space and the allocation plan to support a routed access design?



Both L2 and L3 Can Provide Sub-Second Convergence

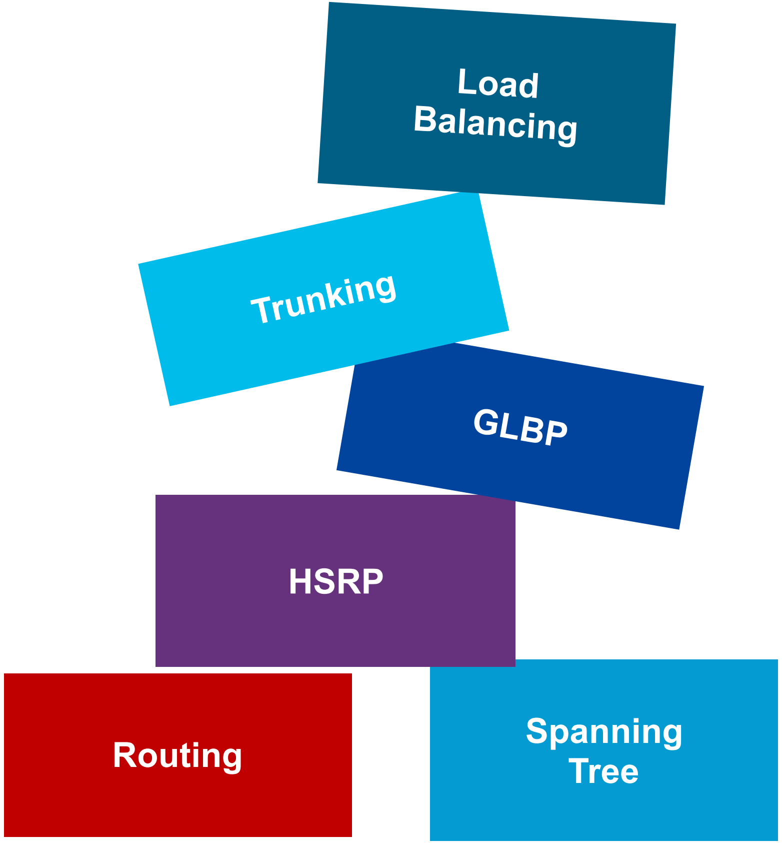


Agenda

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

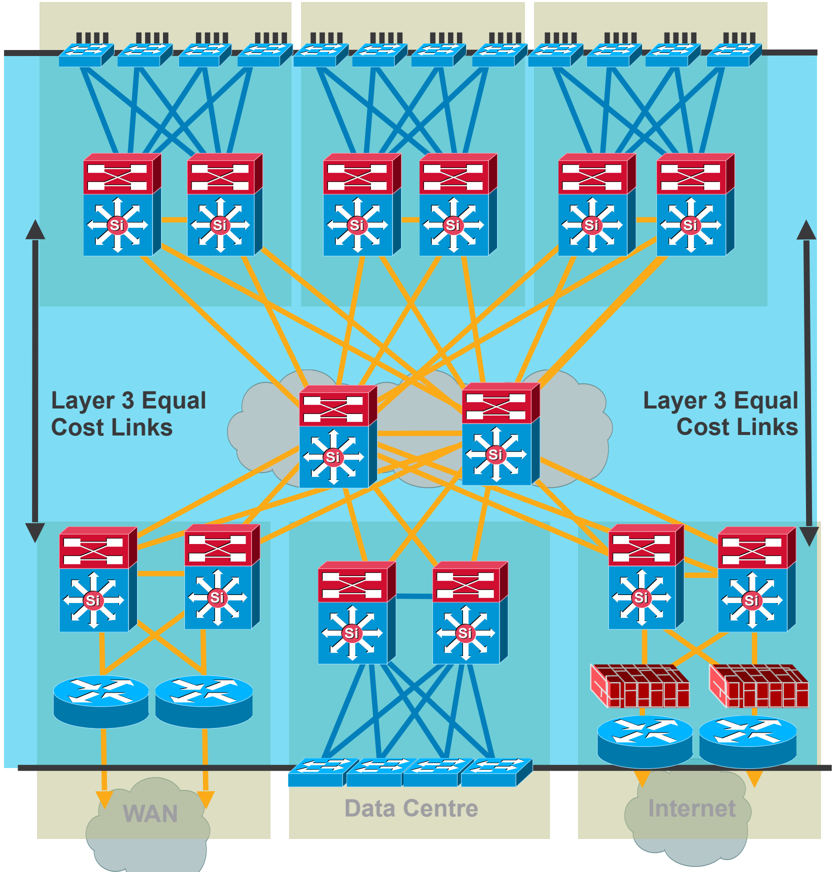
Foundation Services

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



Best Practices - Layer 1 Physical Things

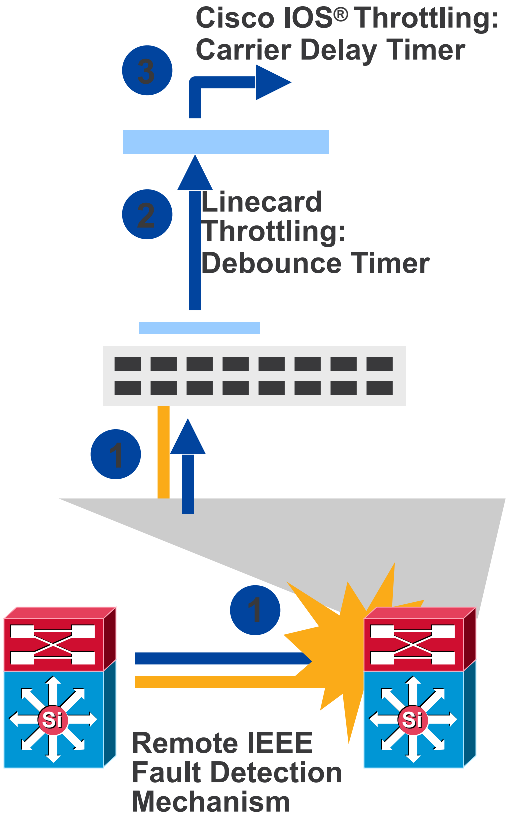
- Use point-to-point interconnections - no L2 aggregation points between nodes
- Use fibre for best convergence (debounce timer)
- Tune carrier delay timer
- Use configuration on the physical interface not VLAN/SVI when possible



Redundancy and Protocol Interaction

Link Redundancy and Failure Detection

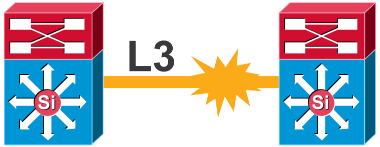
- Direct point-to-point fibre provides for fast failure detection
- IEEE 802.3z and 802.3ae link negotiation define the use of remote fault indicator and link fault signalling mechanisms
- Bit D13 in the Fast Link Pulse (FLP) can be set to indicate a physical fault to the remote side
- Do not disable 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
  - 3560, 3750, and 4500—0 msec
  - 6500—leave it set at default



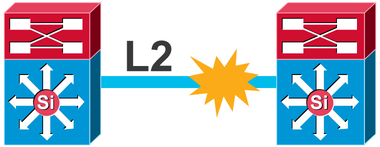
Redundancy and Protocol Interaction

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



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

~ 8 msec loss

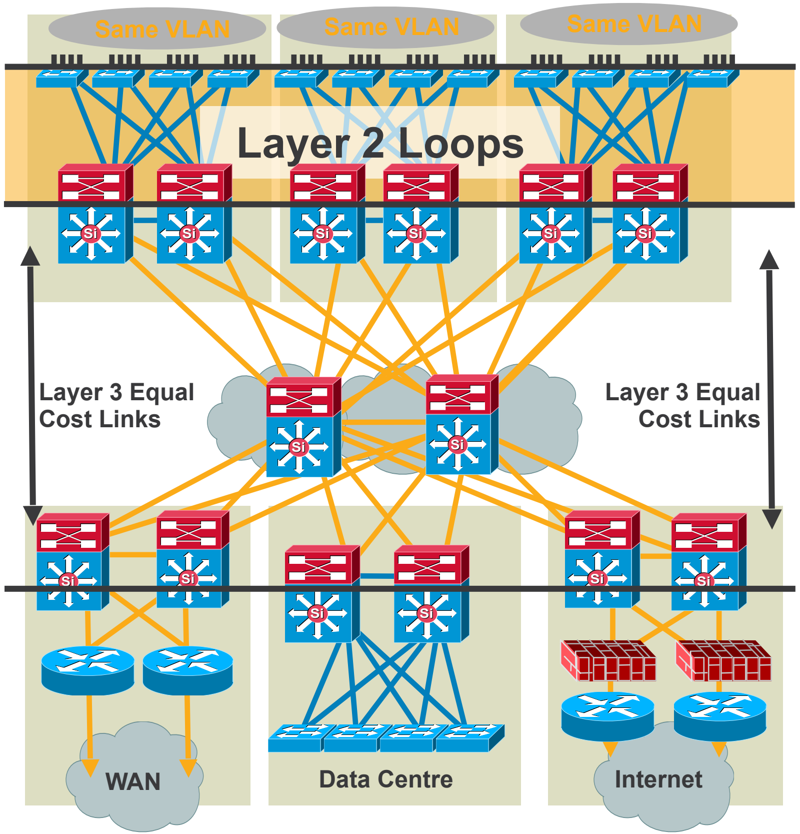
```
21:38:37.042 UTC: %LINEPROTO-5-UPDOWN: Line
protocol on Interface GigabitEthernet3/1, changed state
to down
21:38:37.050 UTC: %LINK-3-UPDOWN: Interface
GigabitEthernet3/1, changed state to down
21:38:37.050 UTC: IP-EIGRP(Default-IP-Routing-Table:
100): Callback: route_adjust GigabitEthernet3/1
```

~ 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
```

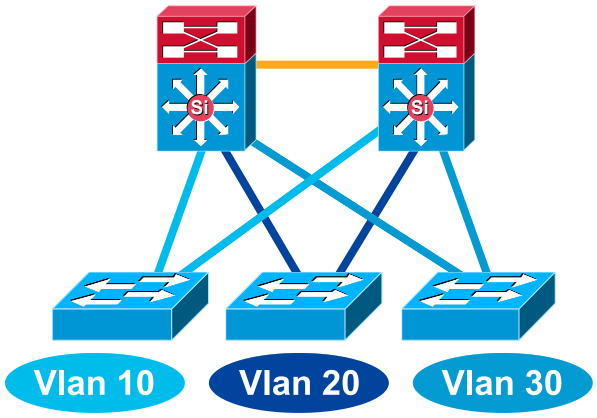
Best Practices - Spanning Tree Configuration

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

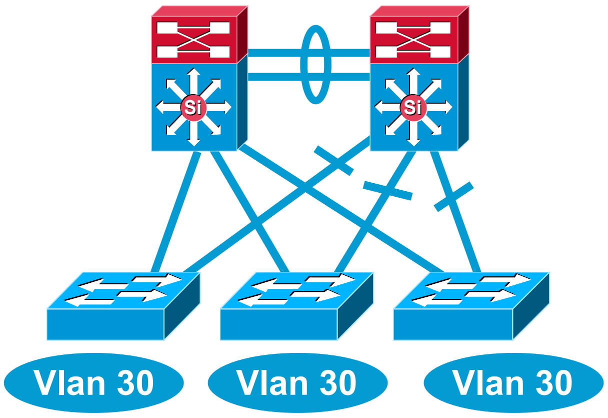


Multilayer Network Design

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

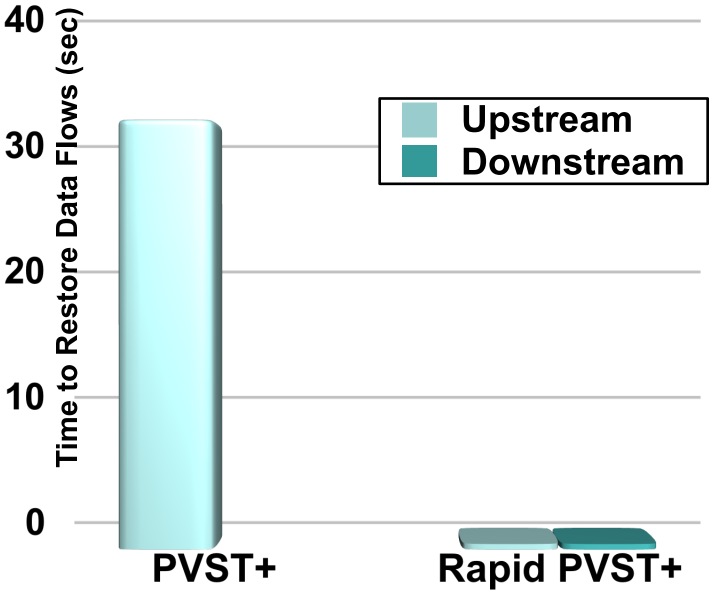


- At least some VLANs span multiple access switches
- Layer 2 loops
- Layer 2 and 3 running over link between distribution
- Blocked links

Optimising L2 Convergence

PVST+, Rapid PVST+ or MST

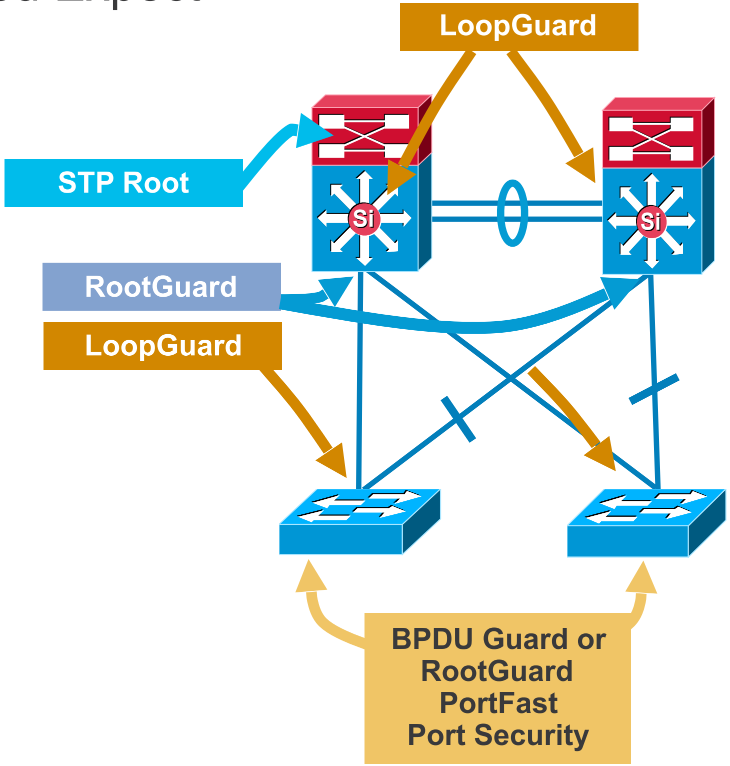
- Rapid-PVST+ greatly improves the restoration times for any VLAN that requires a topology convergence due to link UP
- Rapid-PVST+ also greatly improves convergence time over backbone fast for any indirect link failures
- PVST+ (802.1d)
  - Traditional spanning tree implementation
- Rapid PVST+ (802.1w)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

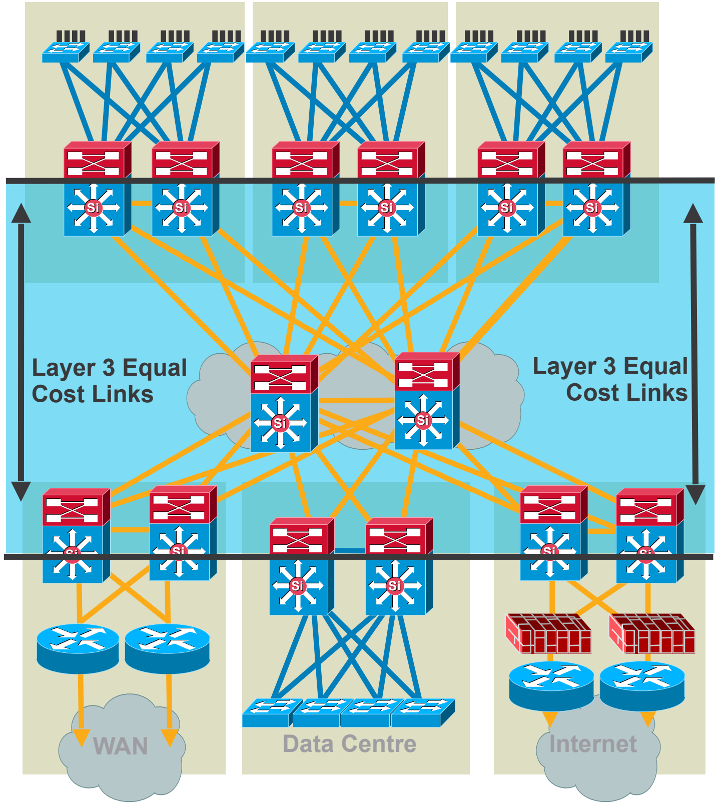
- Place the root where you want it
  - Root primary/secondary macro
- The root bridge should stay where you put it
  - RootGuard
  - LoopGuard
  - UplinkFast
  - UDLD
- Only end-station traffic should be seen on an edge port
  - BPDU Guard
  - RootGuard
  - PortFast
  - Port-security



Best Practices

Layer 3 Routing Protocols

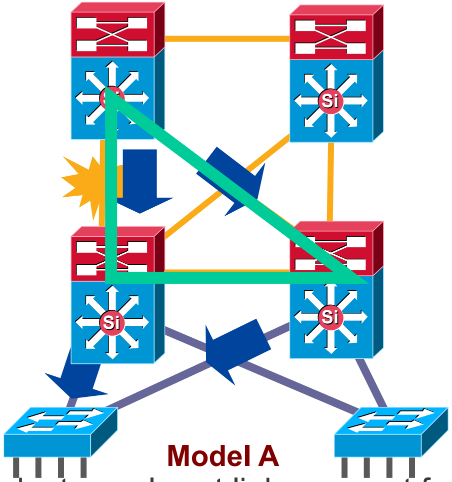
- Typically deployed in distribution to core, and core-to-core interconnections
- Used to quickly reroute around failed node/links while providing load balancing over redundant paths
- Build triangles not squares for deterministic convergence
- Only peer on links that you intend to use as transit
- Insure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



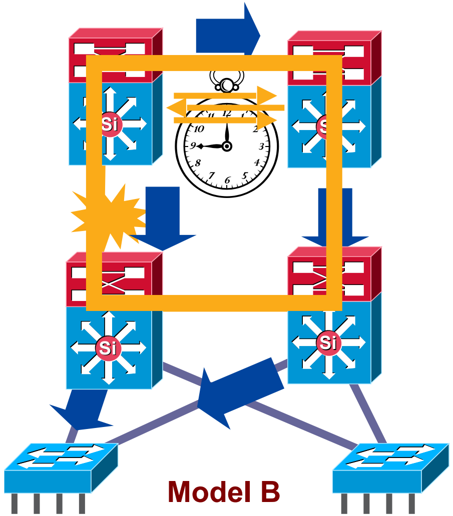
Best Practice - Build Triangles not Squares

Deterministic vs. Non-Deterministic

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



Squares: Link/Box Failure Requires Routing Protocol Convergence

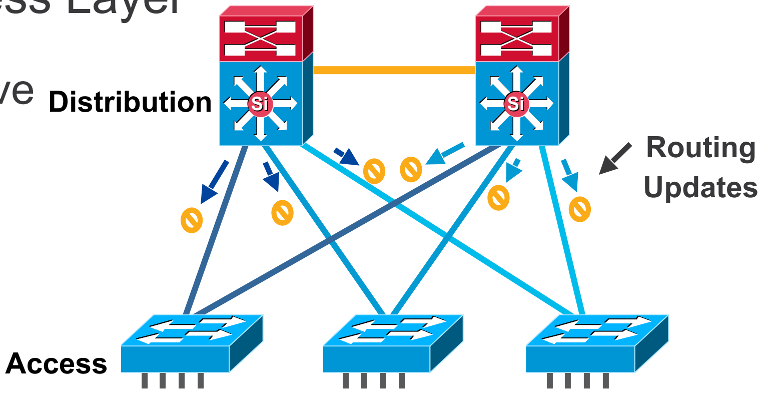


- Layer 3 redundant equal cost links support fast convergence
- Hardware based—fast recovery to remaining path
- Convergence is extremely fast (dual equal-cost paths: no need for OSPF or EIGRP to recalculate a new path)

Best Practice - Passive Interfaces for IGP

Limit IGP 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:

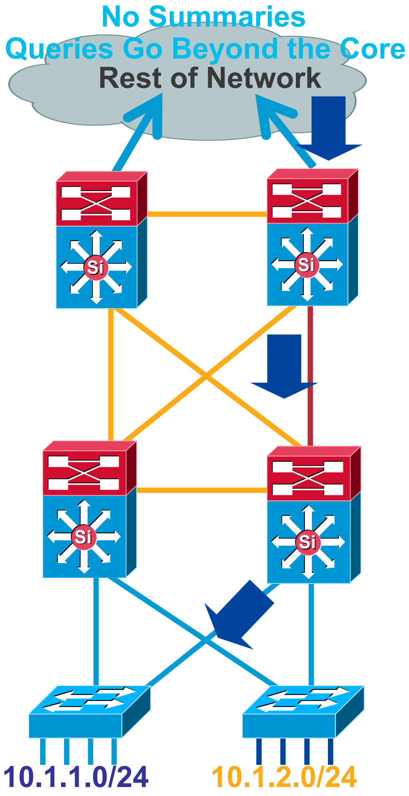
```
Router(config)#routereigrp 1
Router(config-router)#passive-interfaceVlan 99

Router(config)#routereigrp 1
Router(config-router)#passive-interface default
Router(config-router)#no passive-interface Vlan 99
```

Why You Want to Summarise at the Distribution

Limit EIGRP Queries and OSPF LSA Propagation

- It is important to force summarisation 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 optimise this reroute
- EIGRP example:

```
interface Port-channel1
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
```



Core

Distribution

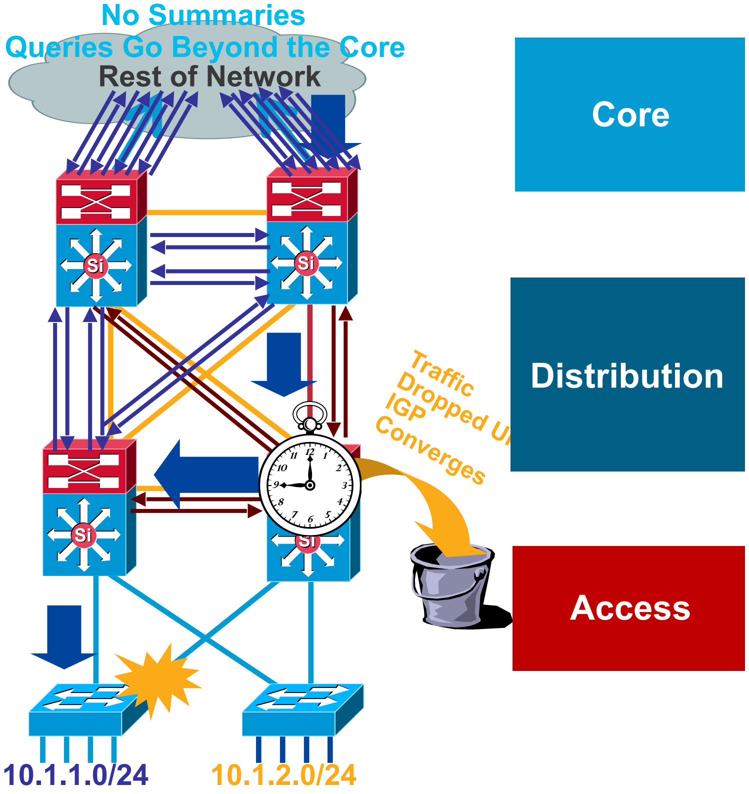
Access

Why You Want to Summarise at the Distribution

Limit EIGRP Queries and OSPF LSA Propagation

- It is important to force summarisation 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 optimise this reroute
- EIGRP example:

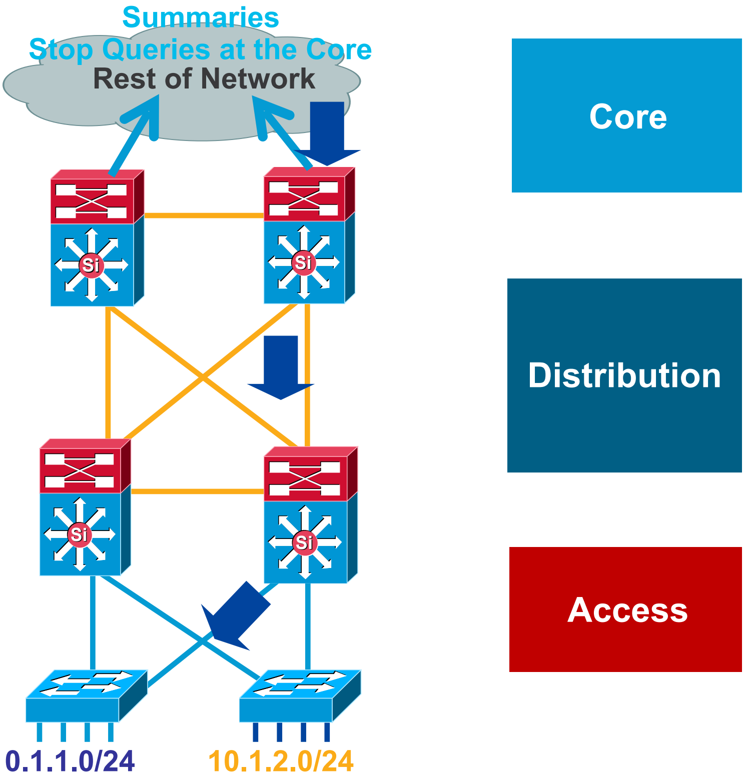
```
interface Port-channel1
description to Core#1
ip address 10.122.0.34
255.255.255.252
ip hello-interval eigrp 100 1
ip hold-time eigrp 100 3
ip summary-address eigrp 100
10.1.0.0 255.255.0.0 5
```



Why You Want to Summarise at the Distribution

Reduce the Complexity of IGP Convergence

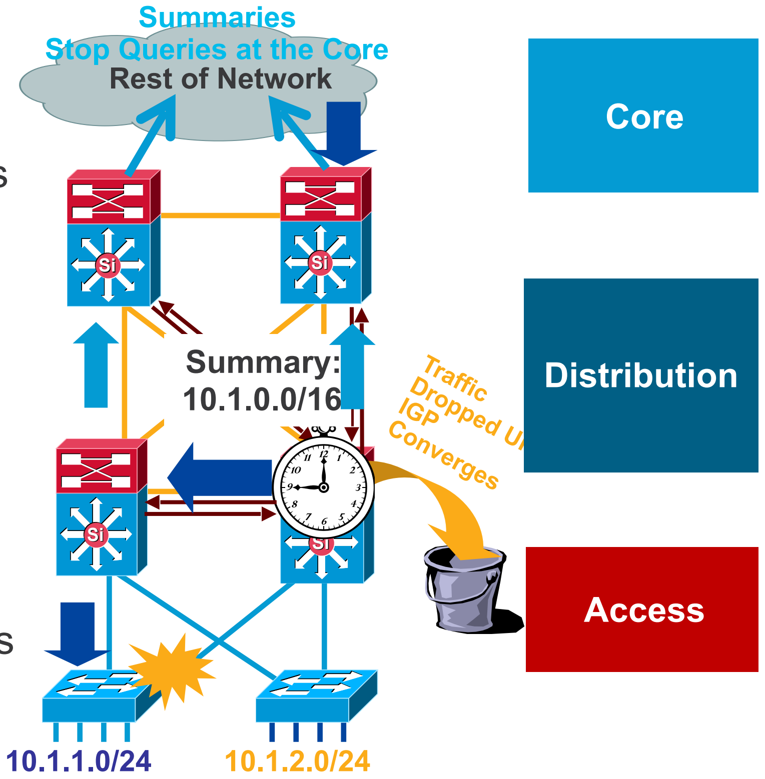
- It is important to force summarisation at the distribution towards the core
- For return path traffic an OSPF or EIGRP re-route is required
- By limiting the number of peers an EIGRP router must query or the number of LSAs an OSPF | peer must process we can optimise his reroute
- For EIGRP if we summaries at the distribution we stop queries at the core boxes for an access layer flap
- For OSPF when we summarise 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



Why You Want to Summarise at the Distribution

Reduce the Complexity of IGP Convergence

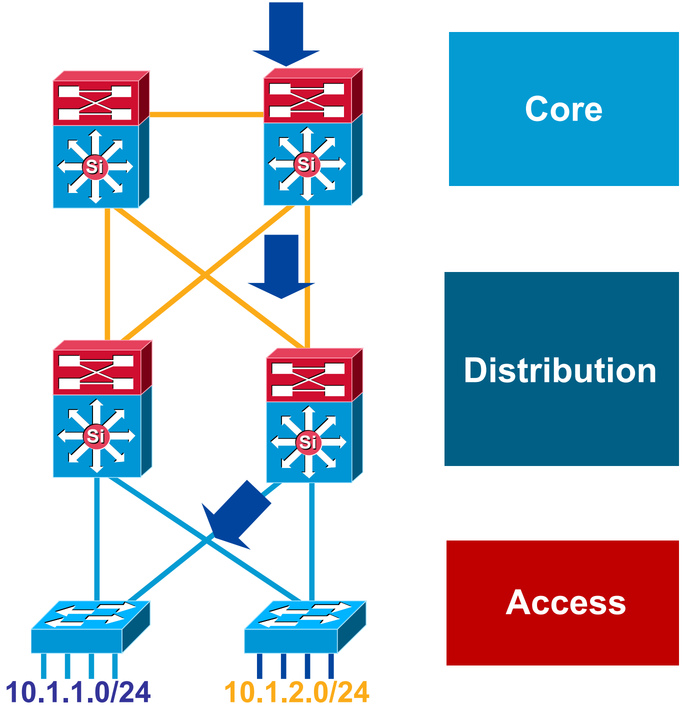
- It is important to force summarisation at the distribution towards the core
- For return path traffic an OSPF or EIGRP re-route is required
- By limiting the number of peers an EIGRP router must query or the number of LSAs an OSPF | peer must process we can optimise his reroute
- For EIGRP if we summaries at the distribution we stop queries at the core boxes for an access layer flap
- For OSPF when we summarise 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 - Summarise at the Distribution

Gotcha—Distribution-to-Distribution Link Required

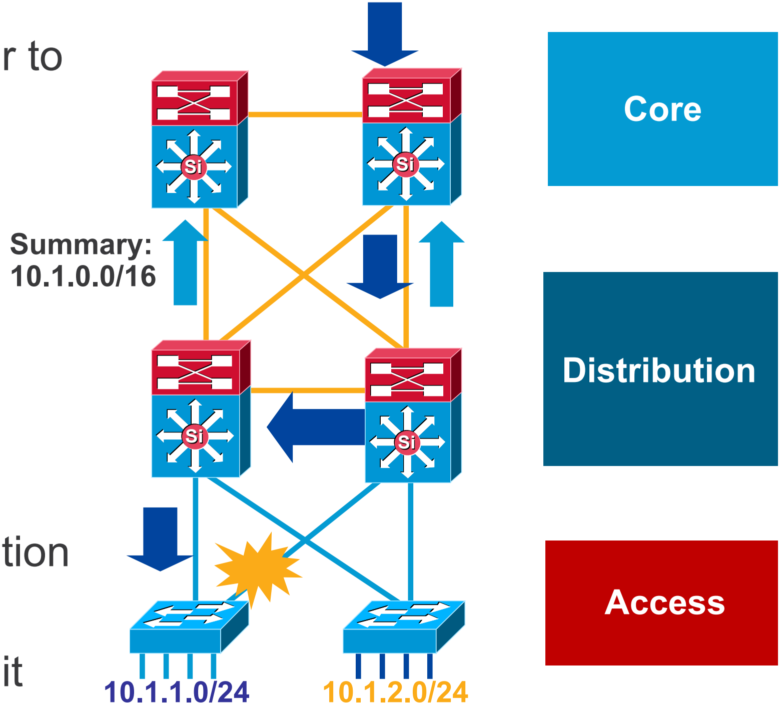
- Best practice - summarise at the distribution layer to limit EIGRP queries or OSPF LSA propagation
- Gotcha:
  - Upstream: HSRP on left distribution takes over when link fails
  - Return path: old router still advertises summary to core
  - Return traffic is dropped on right distribution switch
- Summarising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



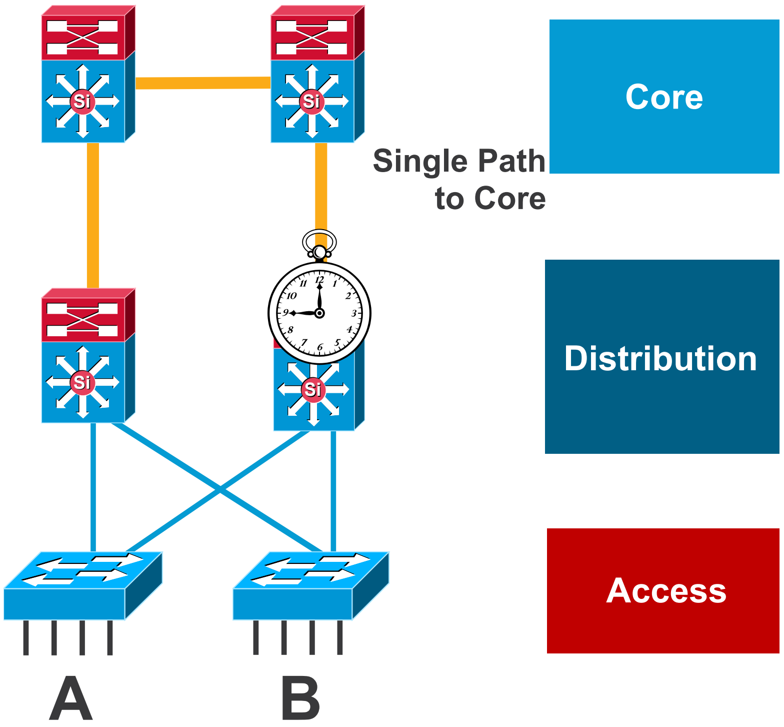
Best Practice - Summarise at the Distribution

Gotcha—Distribution-to-Distribution Link Required

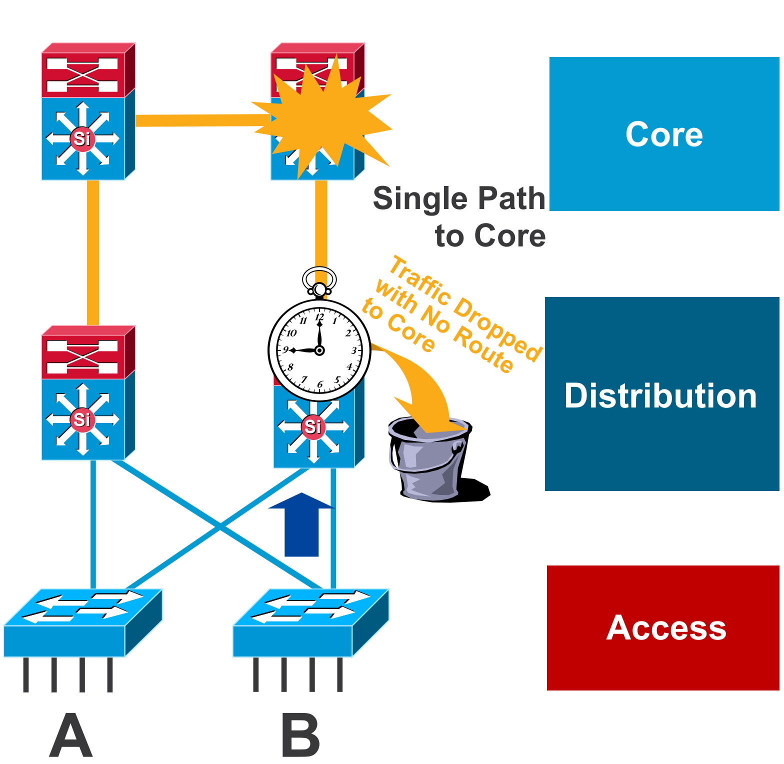
- Best practice - summarise at the distribution layer to limit EIGRP queries or OSPF LSA propagation
- Gotcha:
  - Upstream: HSRP on left distribution takes over when link fails
  - Return path: old router still advertises summary to core
  - Return traffic is dropped on right distribution switch
- Summarising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



Provide Alternate Paths

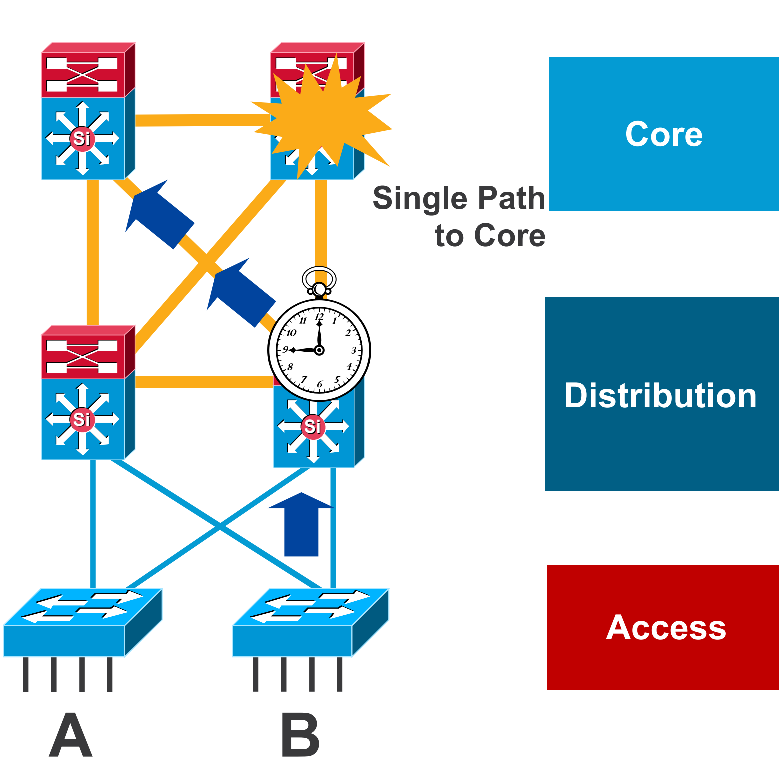


Provide Alternate Paths



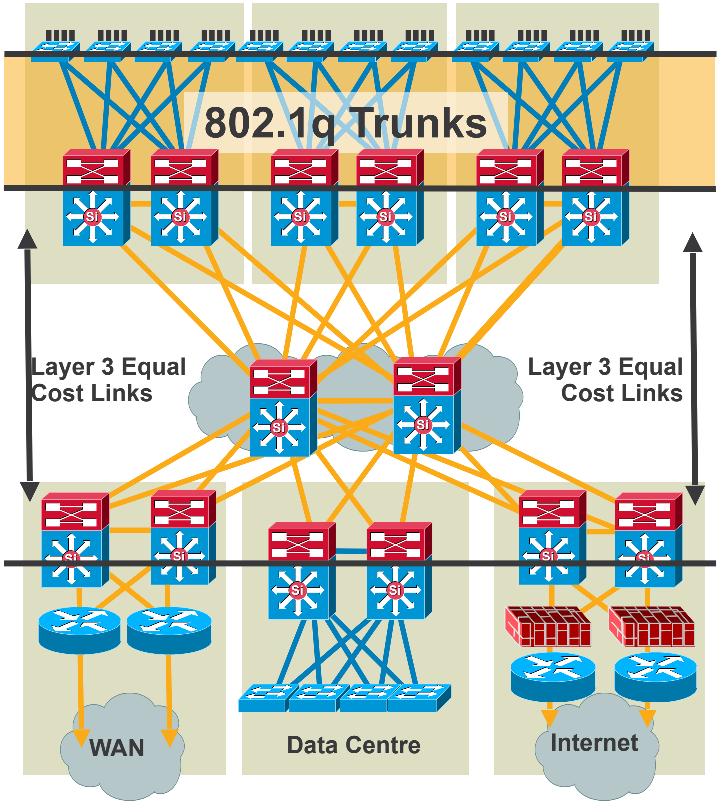
Provide Alternate Paths

- 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 utilise L3 link between distribution layer



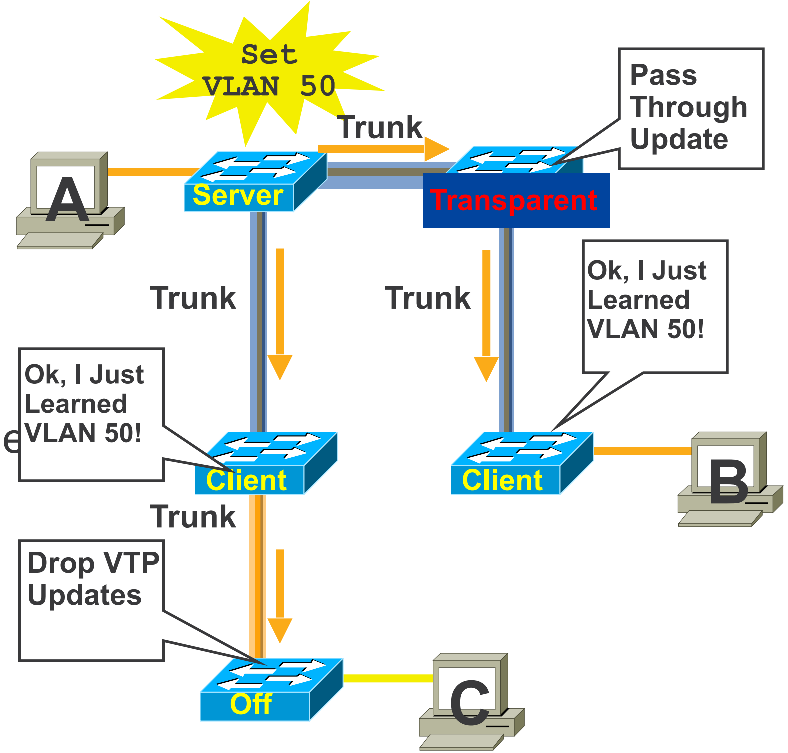
Best Practices - Trunk Configuration

- 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



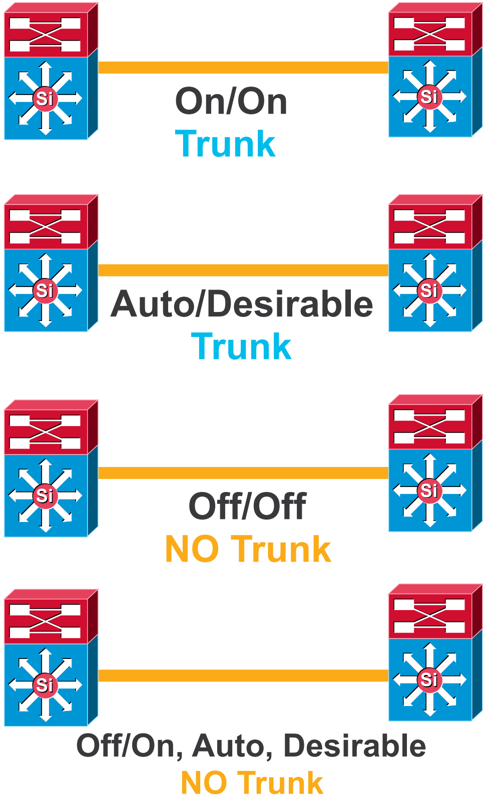
VTP Virtual Trunk Protocol

- Centralised 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 change
  - Transparent: let updates pass through
  - Off: ignores VTP updates



DTP Dynamic Trunk Protocol

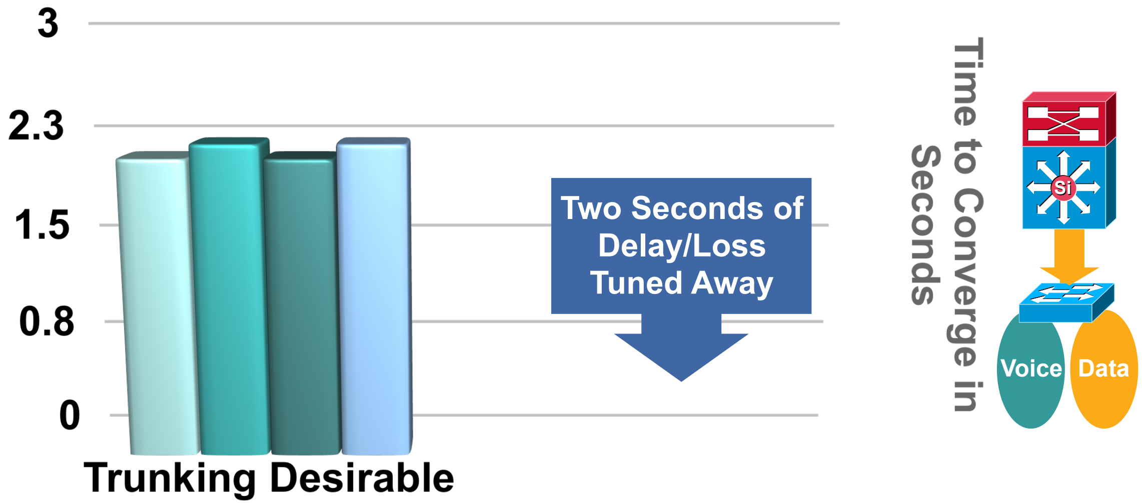
- Automatic formation of trunked switch-to-switch interconnection
  - On: always be a trunk
  - Desirable: ask if the other side can/will
  - Auto: if the other sides asks I will
  - Off: don't become a trunk
- Negotiation of 802.1Q or ISL encapsulation
  - ISL: try to use ISL trunk encapsulation
  -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



Optimising Convergence: Trunk Tuning

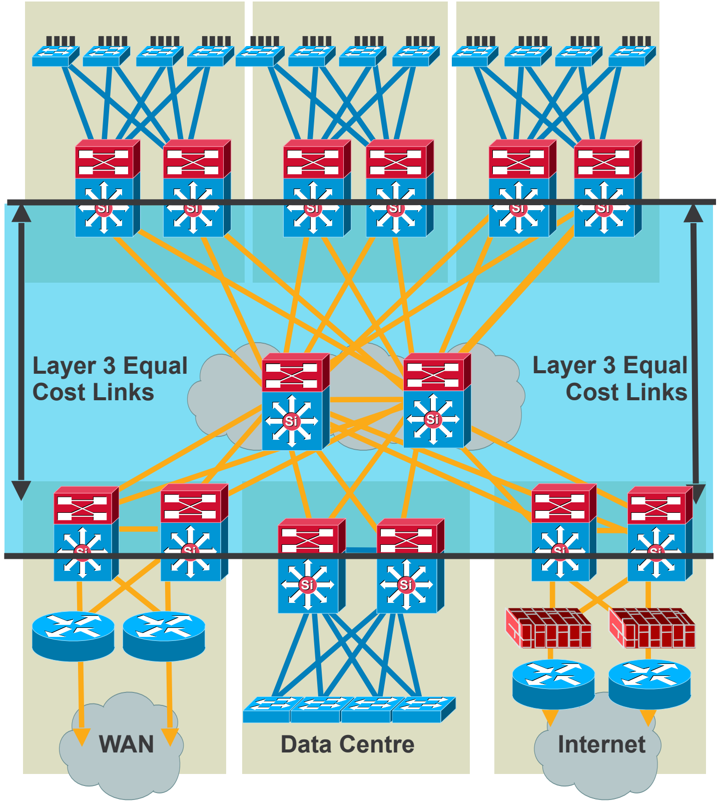
Trunk Auto/Desirable Takes Some Time

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



Best Practices - Ether Channel Configuration

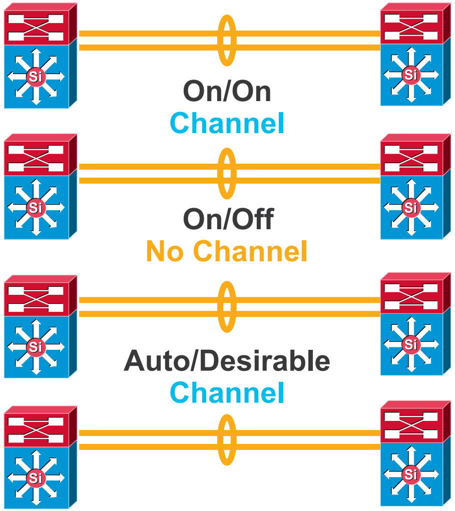
- Typically deployed in distribution to core, and core
  to core interconnections
- Used to provide link redundancy—while reducing peering complexity
- Tune L3/L4 load balancing hash to achieve maximum utilisation of channel members
- Deploy in powers of two (two, four, or eight)
- Match CatOS and Cisco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



Understanding Ether Channel

Link Negotiation Options—PAgP and LACP

Port Aggregation Protocol



On/On
Channel

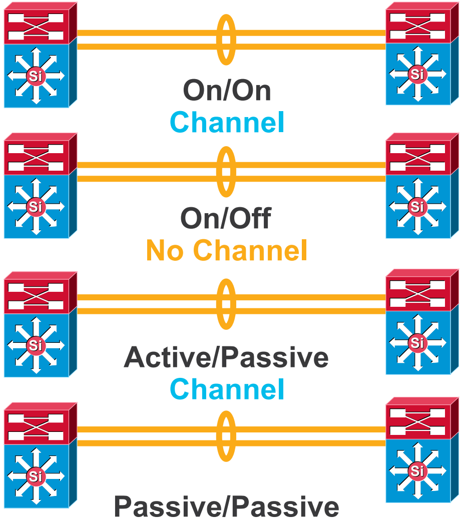
On/Off
No Channel

Auto/Desirable
Channel

Off/On, Auto, Desirable
No Channel

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/On
Channel

On/Off
No Channel

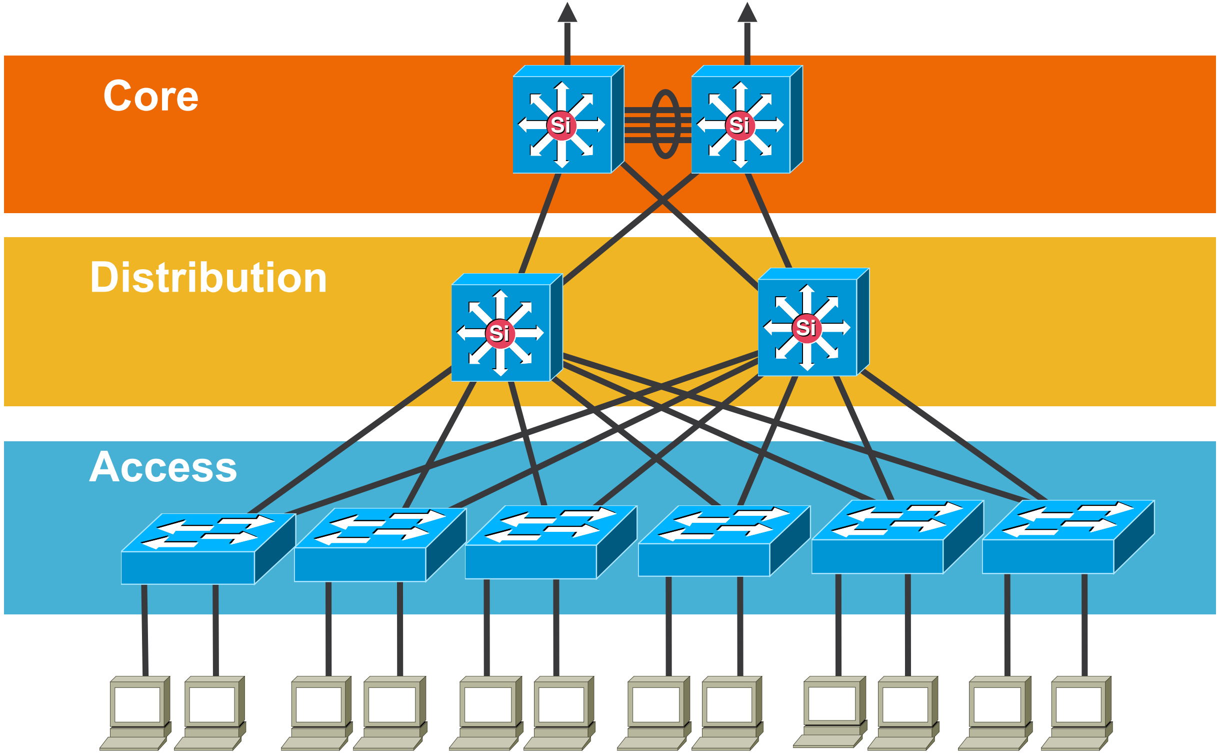
Active/Passive
Channel

Passive/Passive
No Channel

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

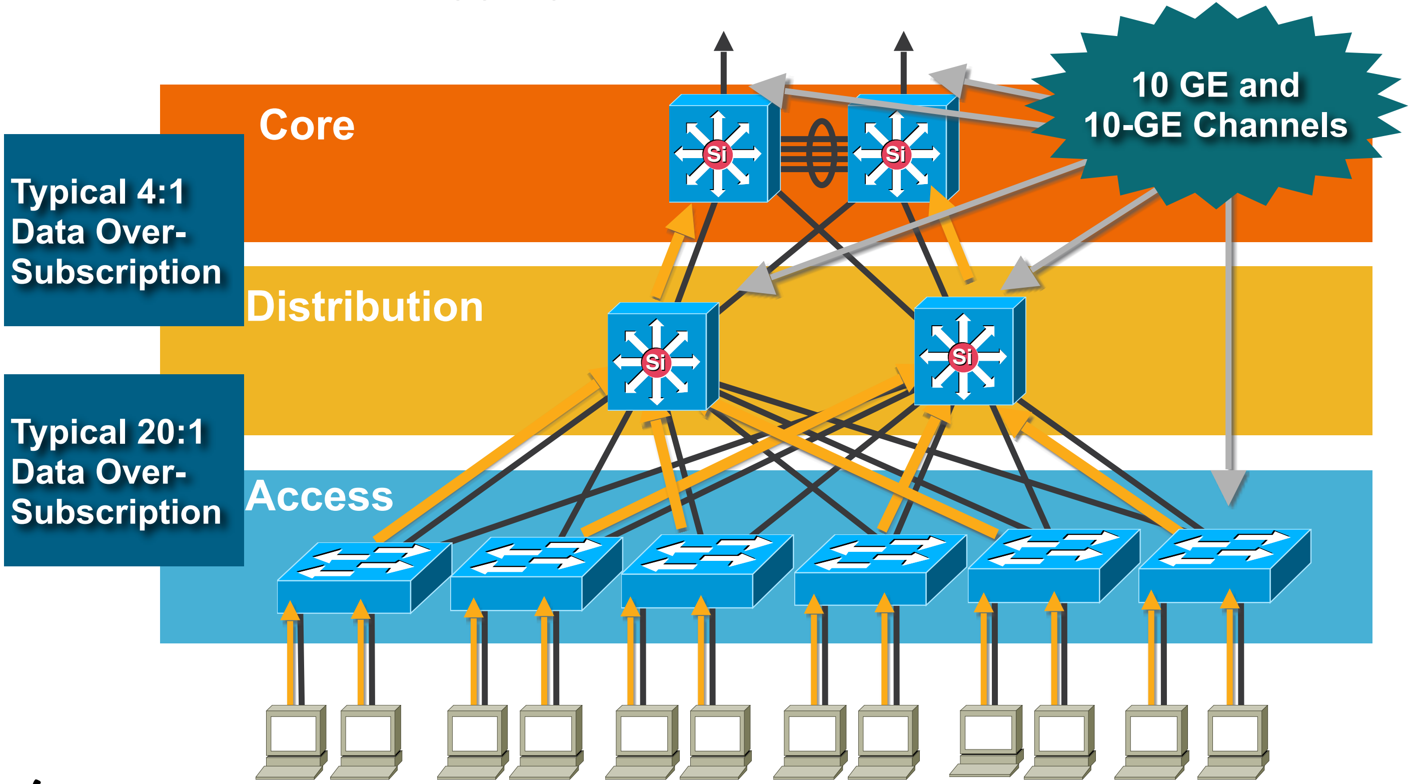
EtherChannels

10/100/1000 How Do You Aggregate It?



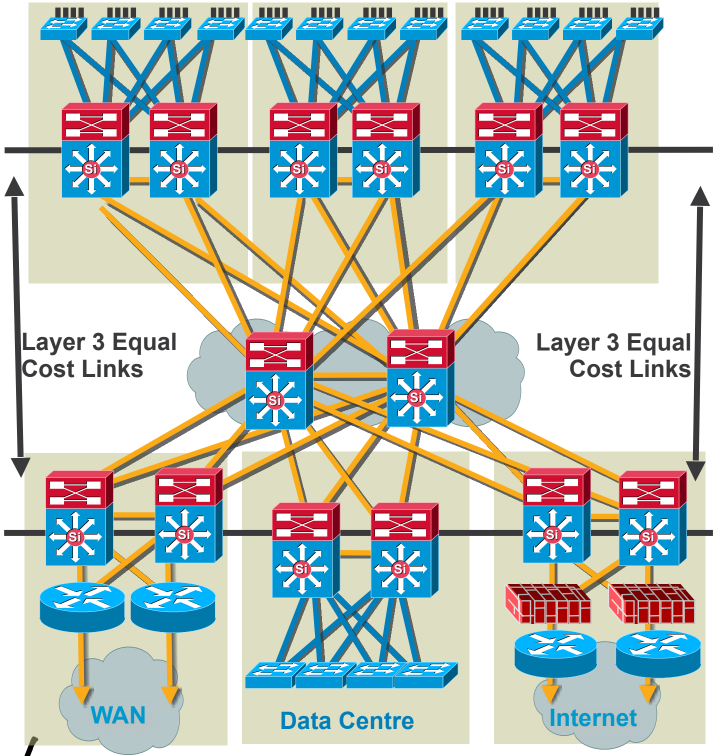
EtherChannels

10/100/1000 How Do You Aggregate It?


Core
Distribution
Access
Typical 4:1 Data Over-Subscription
Typical 20:1 Data Over-Subscription
10 GE and 10-GE Channels
Si
Si
Si
Si

EtherChannels

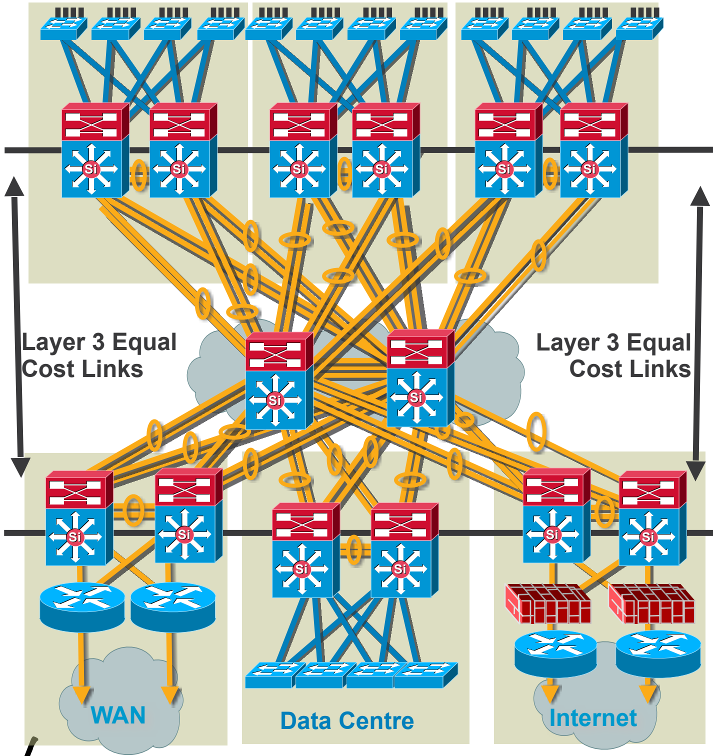
Reduce Complexity/Peer Relationships



- 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

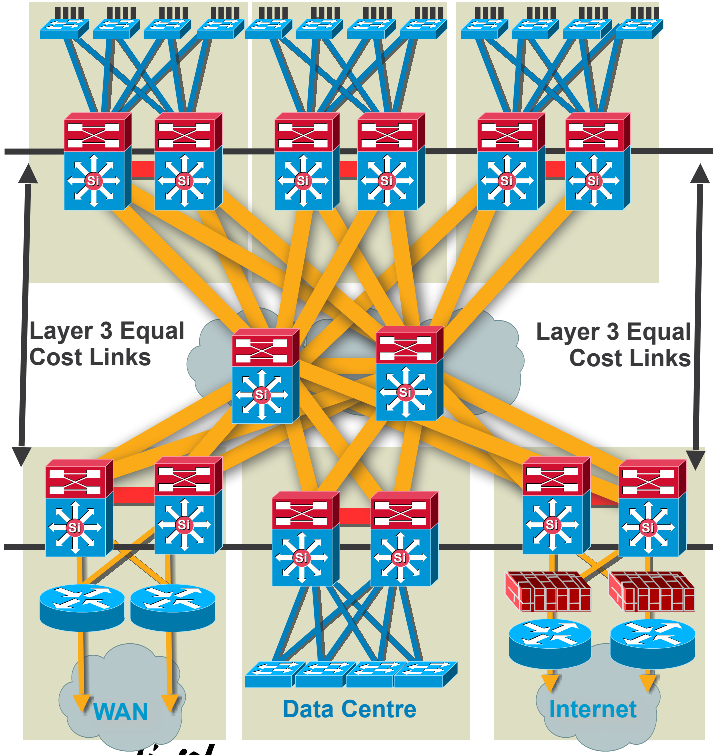
Reduce Complexity/Peer Relationships



- 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

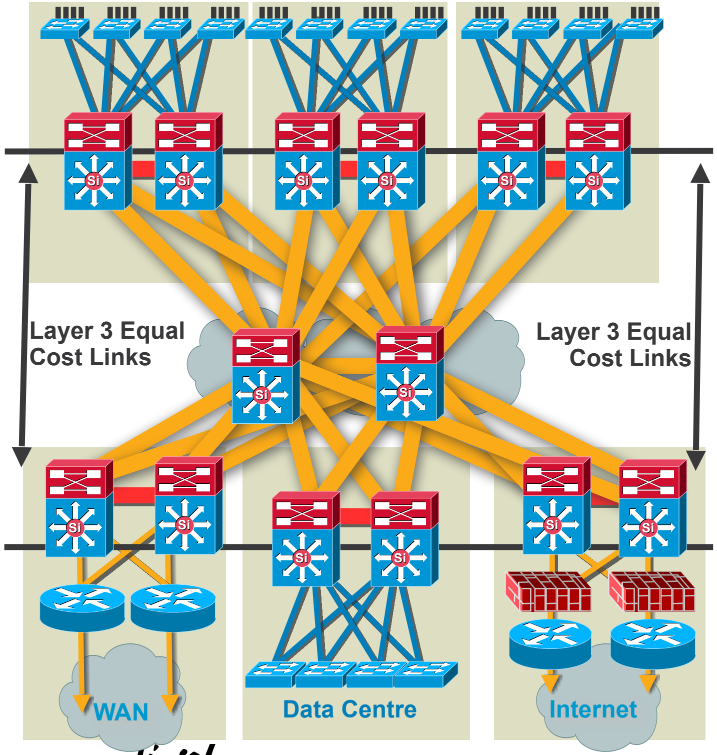
Why 10-Gigabit Interfaces



- 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

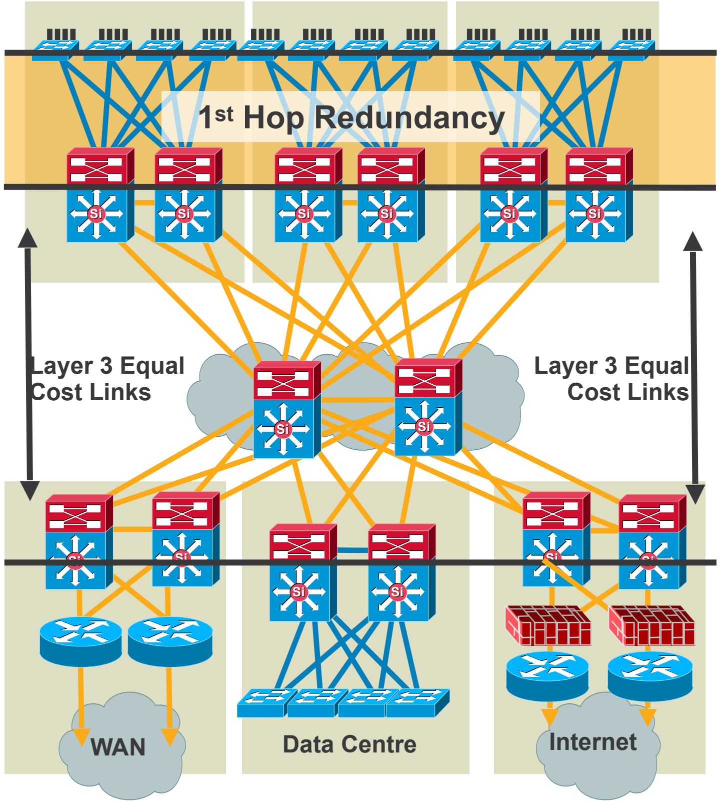
Why 10-Gigabit Interfaces



- 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

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

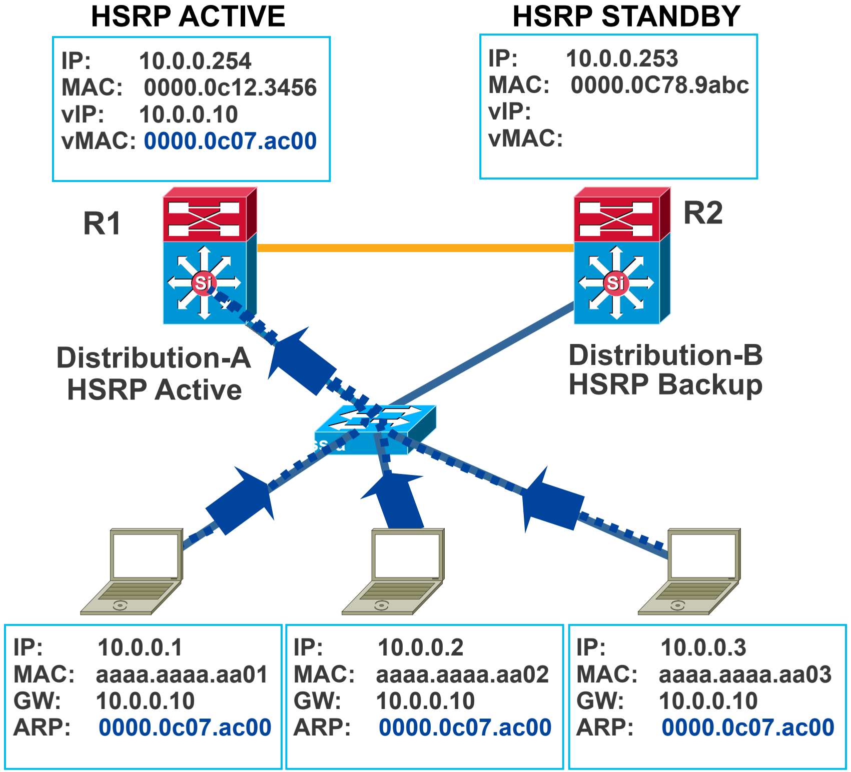


First Hop Redundancy with HSRP

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

R1—Active, Forwarding Traffic;
R2—Hot Standby, Idle



HSRP ACTIVE

IP: 10.0.0.254
MAC: 0000.0c12.3456
vIP: 10.0.0.10
vMAC: 0000.0c07.ac00

HSRP STANDBY

IP: 10.0.0.253
MAC: 0000.0C78.9abc
vIP:
vMAC:

Distribution-A
HSRP Active

Distribution-B
HSRP Backup

IP: 10.0.0.1
MAC: aaaa.aaaa.aa01
GW: 10.0.0.10
ARP: 0000.0c07.ac00

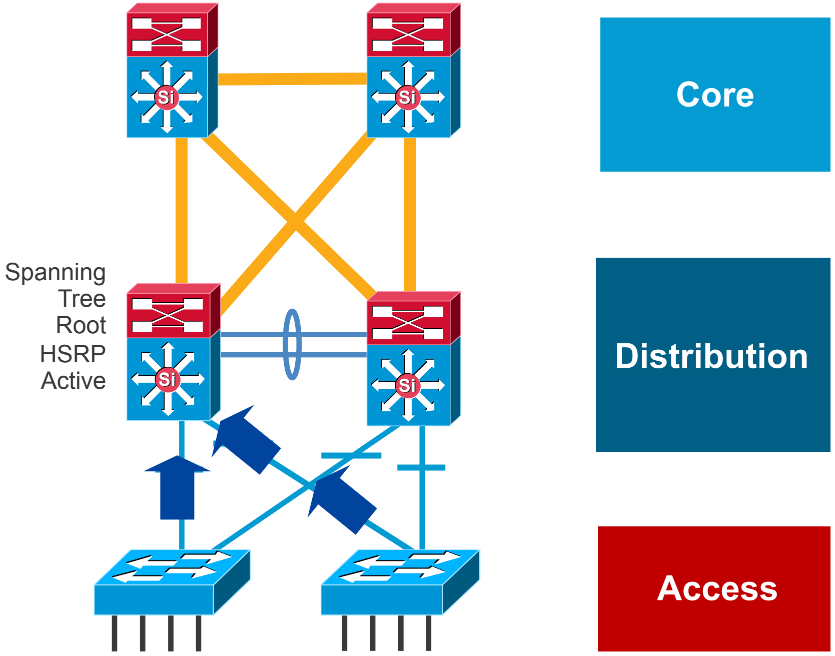
IP: 10.0.0.2
MAC: aaaa.aaaa.aa02
GW: 10.0.0.10
ARP: 0000.0c07.ac00

IP: 10.0.0.3
MAC: aaaa.aaaa.aa03
GW: 10.0.0.10
ARP: 0000.0c07.ac00

Why You Want HSRP Preemption

Avoid 'Black-Hole' during system startup

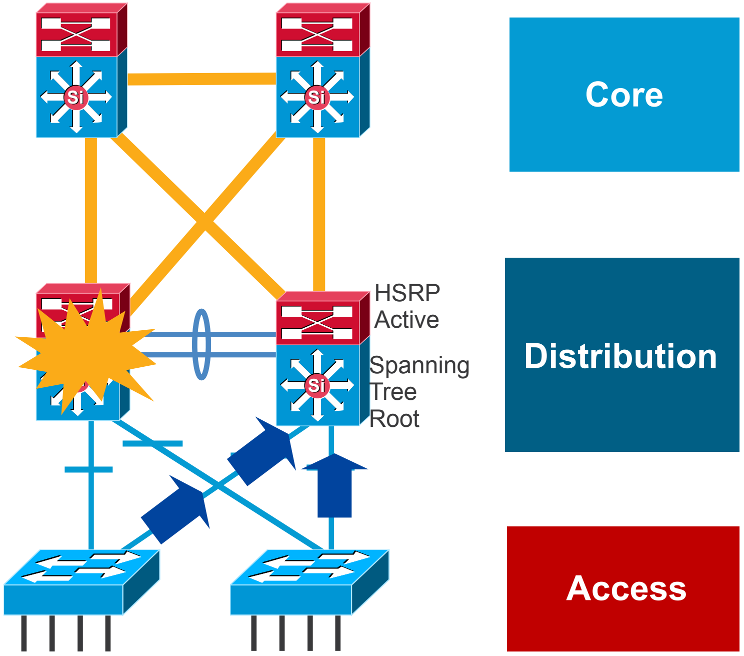
- Spanning tree root and HSRP primary aligned
- When spanning tree root is re-introduced, traffic will take a two-hop path to HSRP active
- HSRP preemption will allow HSRP to follow spanning tree topology



Why You Want HSRP Preemption

Avoid 'Black-Hole' during system startup

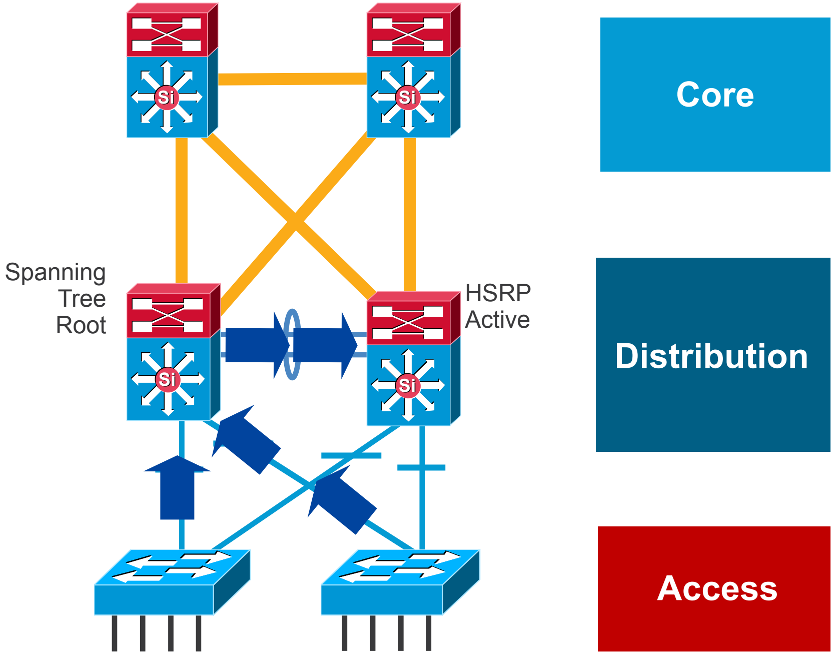
- Spanning tree root and HSRP primary aligned
- When spanning tree root is re-introduced, traffic will take a two-hop path to HSRP active
- HSRP preemption will allow HSRP to follow spanning tree topology



Why You Want HSRP Preemption

Avoid ‘Black-Hole’ during system startup

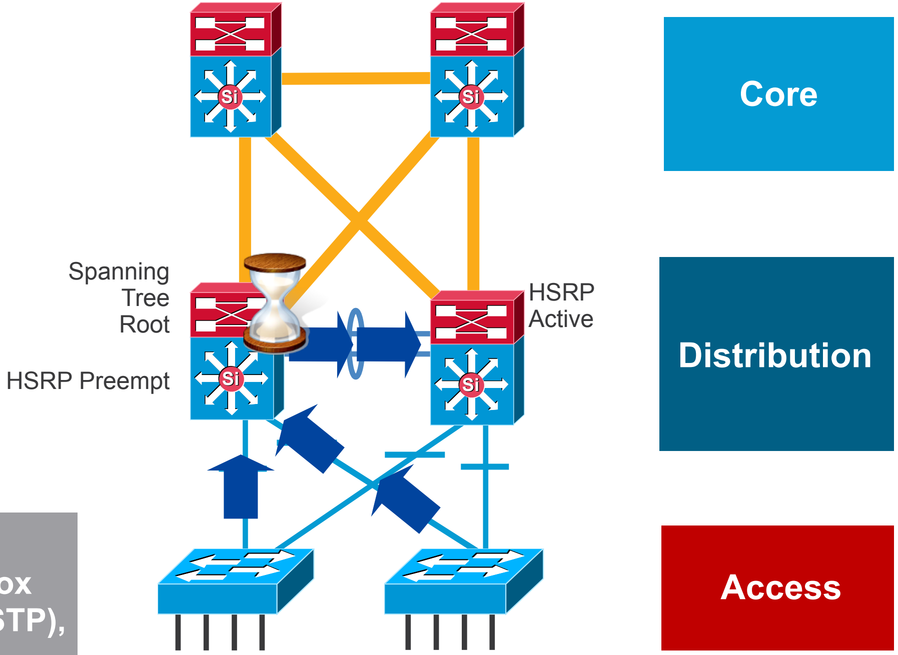
- Spanning tree root and HSRP primary aligned
- When spanning tree root is re-introduced, traffic will take a two-hop path to HSRP active
- HSRP preemption will allow HSRP to follow spanning tree topology



Why You Want HSRP Preemption

Avoid 'Black-Hole' during system startup

- 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



Without Preempt Delay HSRP Can Go Active Before Box Completely Ready to Forward Traffic: L1 (Boards), L2 (STP), L3 (IGP Convergence)

```
standby 1 preempt delay minimum 180
```

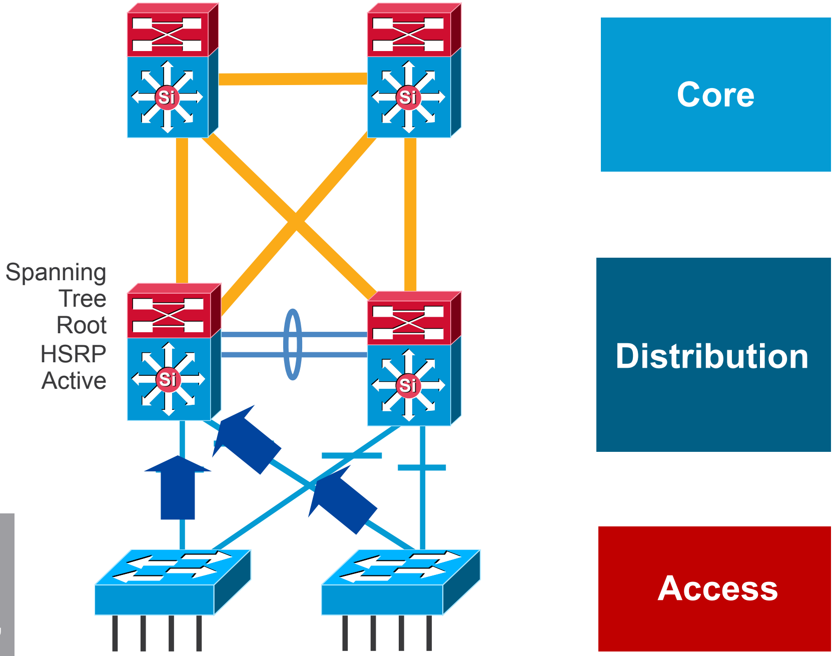
Why You Want HSRP Preemption

Avoid 'Black-Hole' during system startup

- 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

Without Preempt Delay HSRP Can Go Active Before Box Completely Ready to Forward Traffic: L1 (Boards), L2 (STP), L3 (IGP Convergence)

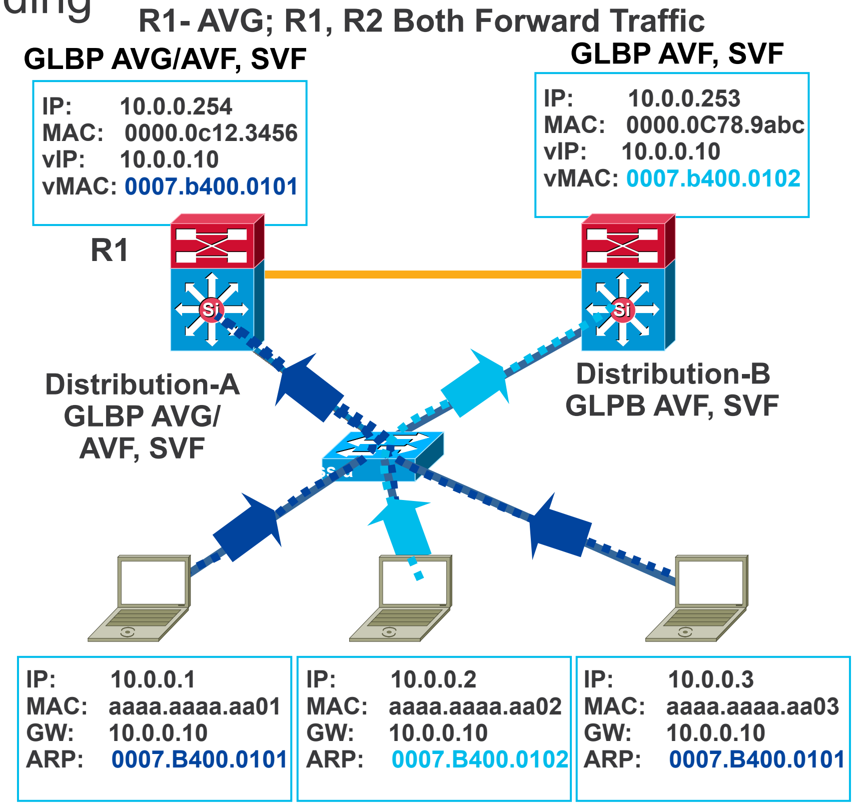
```
standby 1 preempt delay minimum 180
```



First Hop Redundancy with GLBP

Cisco Designed, Load Sharing, Patent Pending

- All the benefits of HSRP plus load balancing of default gateway → utilises 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



R1- AVG; R1, R2 Both Forward Traffic

GLBP AVG/AVF, SVF

IP: 10.0.0.254
MAC: 0000.0c12.3456
vIP: 10.0.0.10
vMAC: 0007.b400.0101

GLBP AVF, SVF

IP: 10.0.0.253
MAC: 0000.0C78.9abc
vIP: 10.0.0.10
vMAC: 0007.b400.0102

IP: 10.0.0.1
MAC: aaaa.aaaa.aa01
GW: 10.0.0.10
ARP: 0007.B400.0101

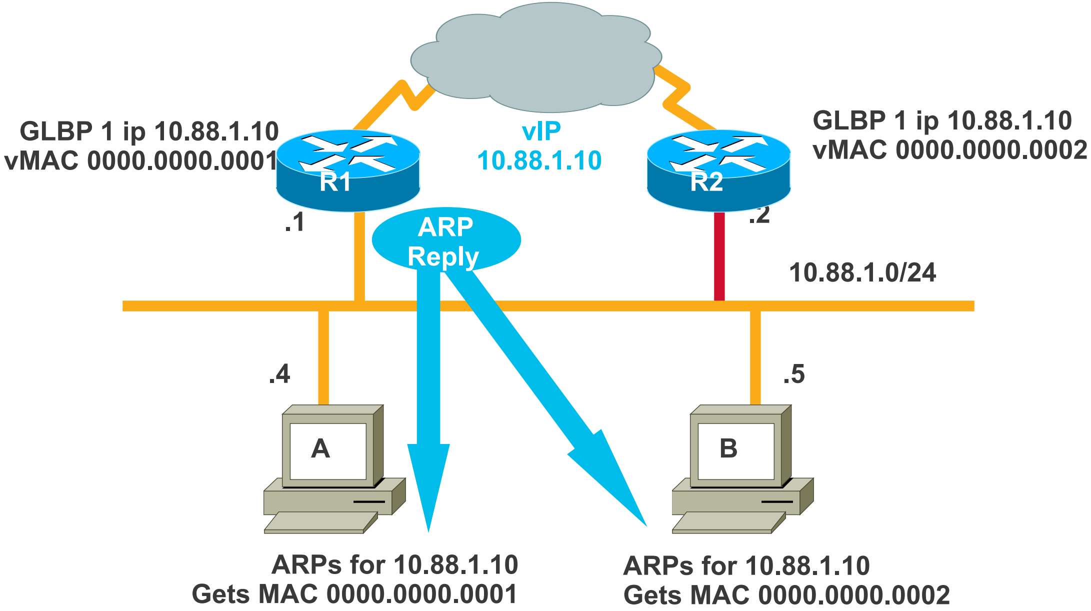
IP: 10.0.0.2
MAC: aaaa.aaaa.aa02
GW: 10.0.0.10
ARP: 0007.B400.0102

IP: 10.0.0.3
MAC: aaaa.aaaa.aa03
GW: 10.0.0.10
ARP: 0007.B400.0101

First Hop Redundancy with Load Balancing

Cisco Gateway Load Balancing Protocol (GLBP)

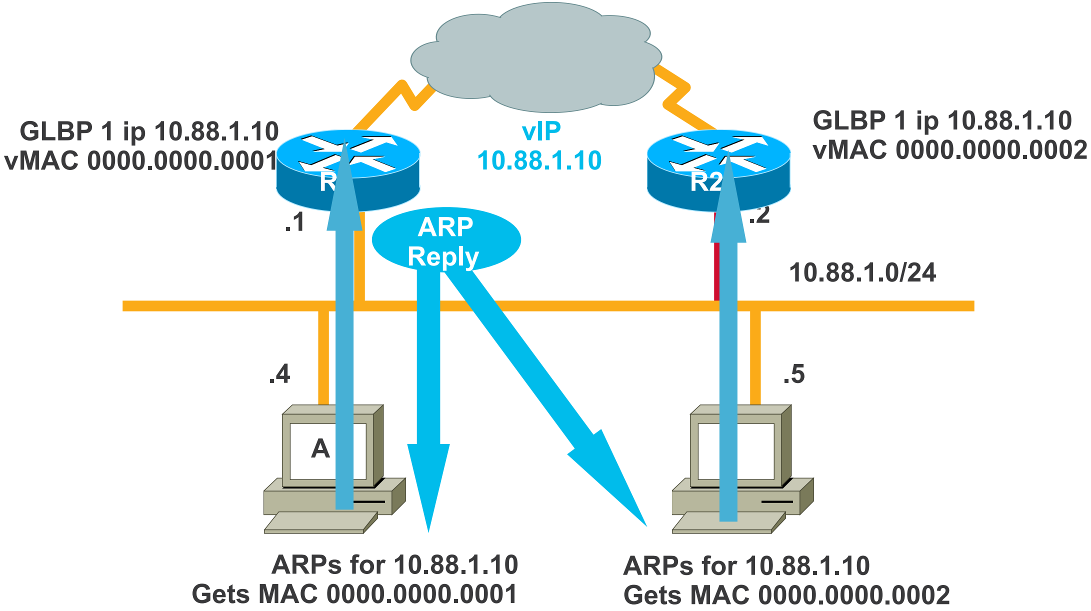
- Each member of a GLBP redundancy group owns a unique virtual MAC address for a common IP address/default gateway
- When end-stations ARP for the common IP address/default gateway they are given a load-balanced virtual MAC address
- Host A and host B send traffic to different GLBP peers but have the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



First Hop Redundancy with Load Balancing

Cisco Gateway Load Balancing Protocol (GLBP)

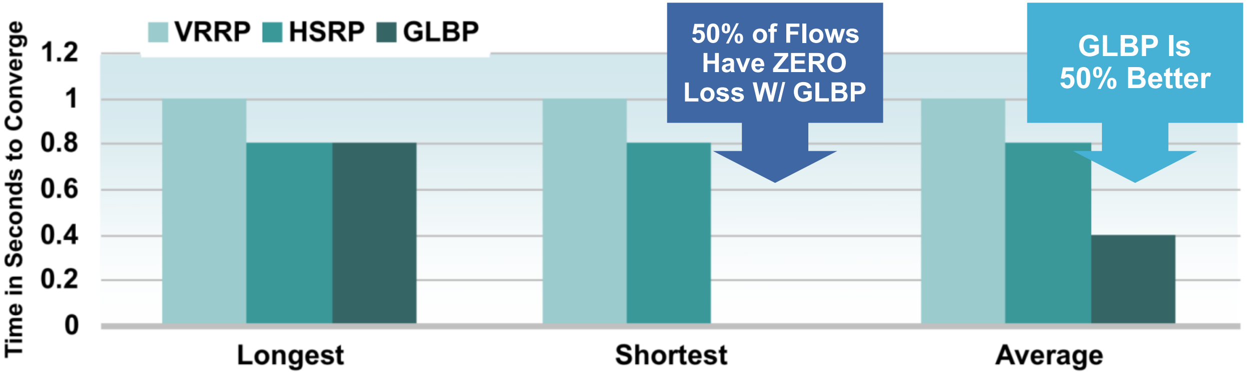
- Each member of a GLBP redundancy group owns a unique virtual MAC address for a common IP address/default gateway
- When end-stations ARP for the common IP address/default gateway they are given a load-balanced virtual MAC address
- Host A and host B send traffic to different GLBP peers but have the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



Optimising 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



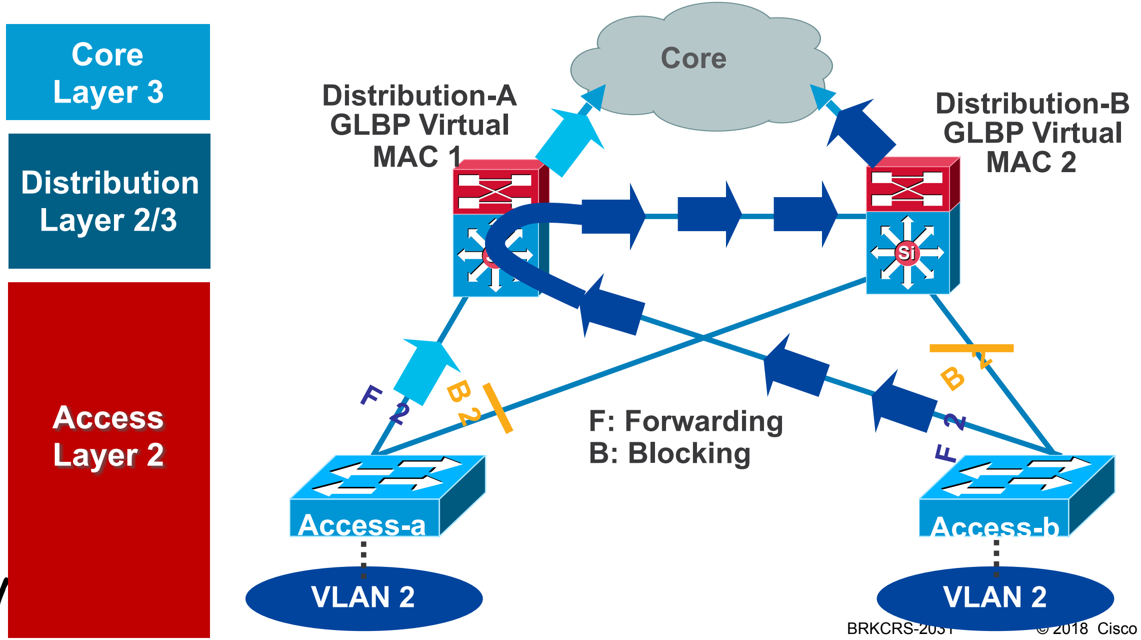
50% of Flows Have ZERO Loss W/ GLBP

GLBP Is 50% Better

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



F: Forwarding
B: Blocking

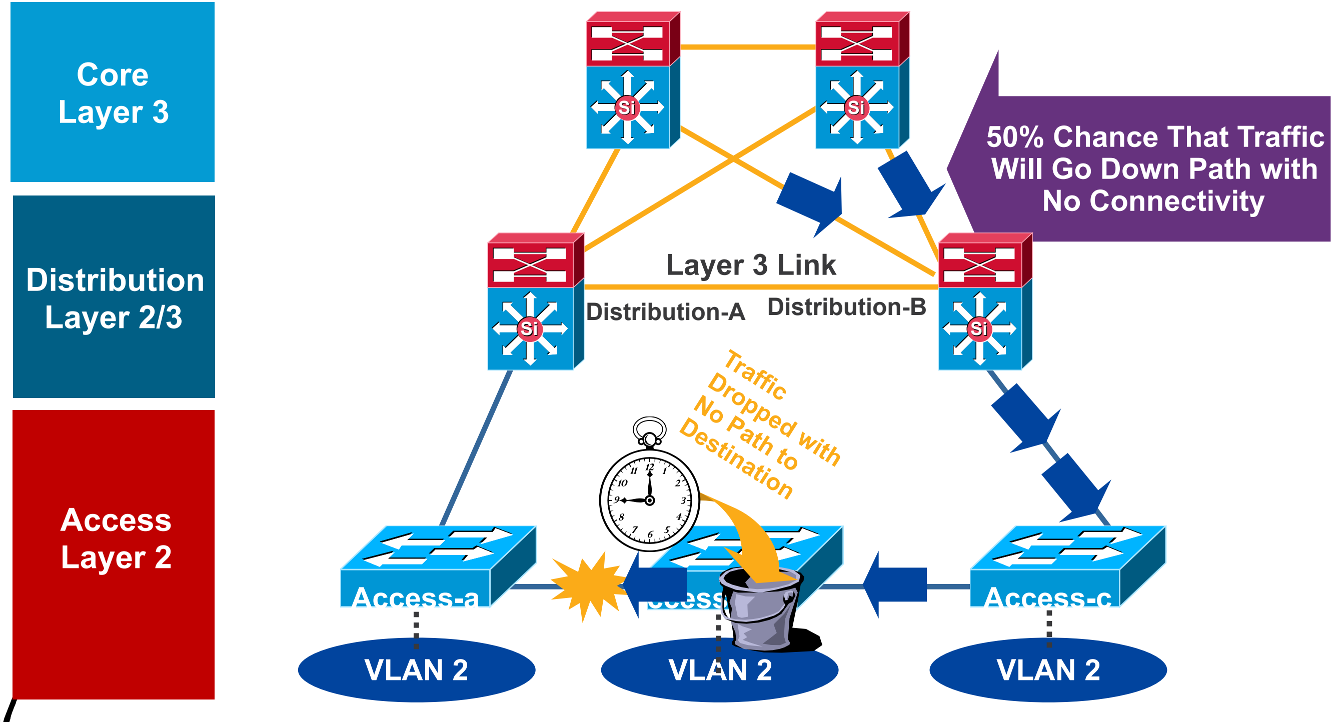
Agenda

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

Daisy Chaining Access Layer Switches

Avoid Potential Black Holes

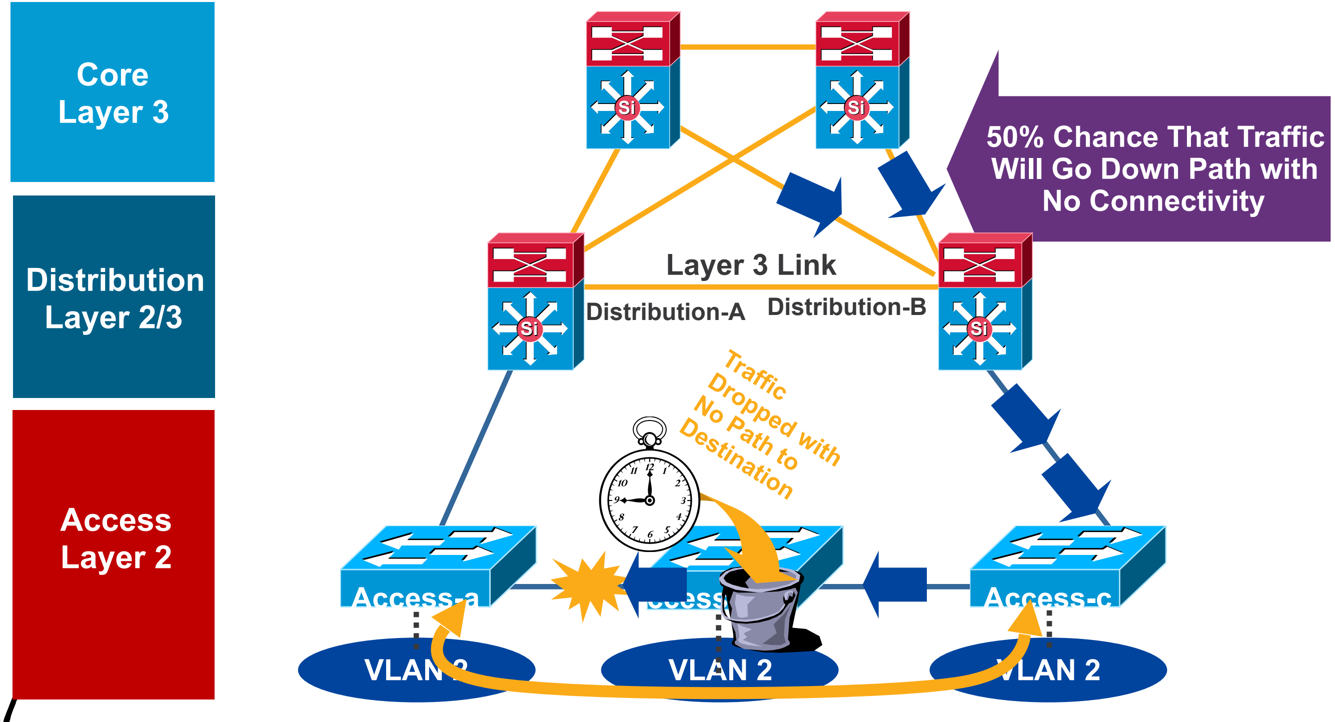
Return Path Traffic Has a 50/50 Chance of Being ‘Black Holed’



Daisy Chaining Access Layer Switches

Avoid Potential Black Holes

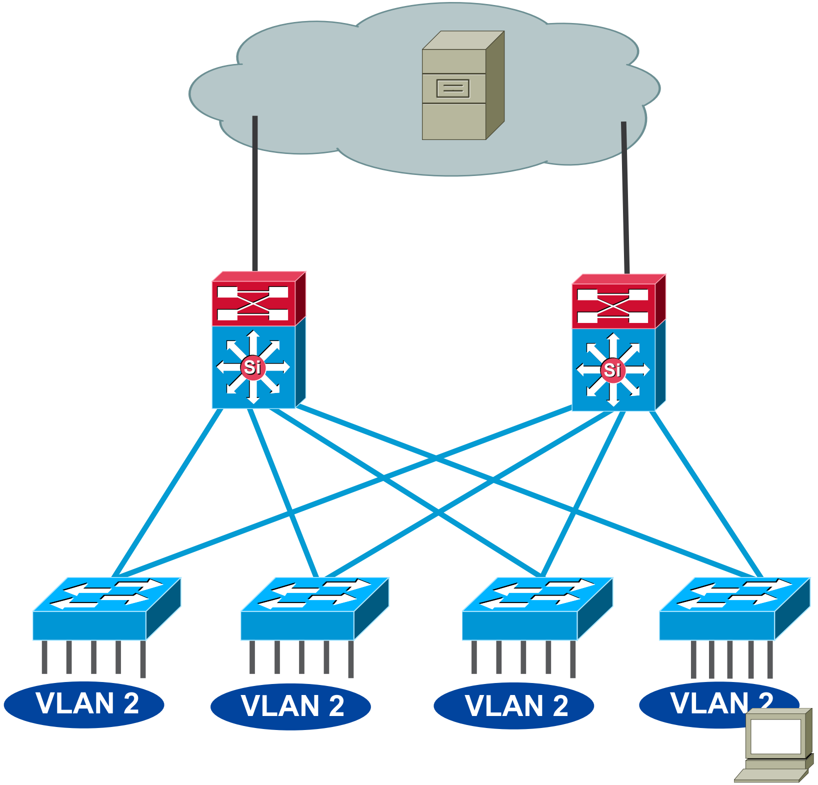
Return Path Traffic Has a 50/50 Chance of Being 'Black Holed'



Asymmetric Routing (Unicast Flooding)

Affects redundant topologies with shared L2 access

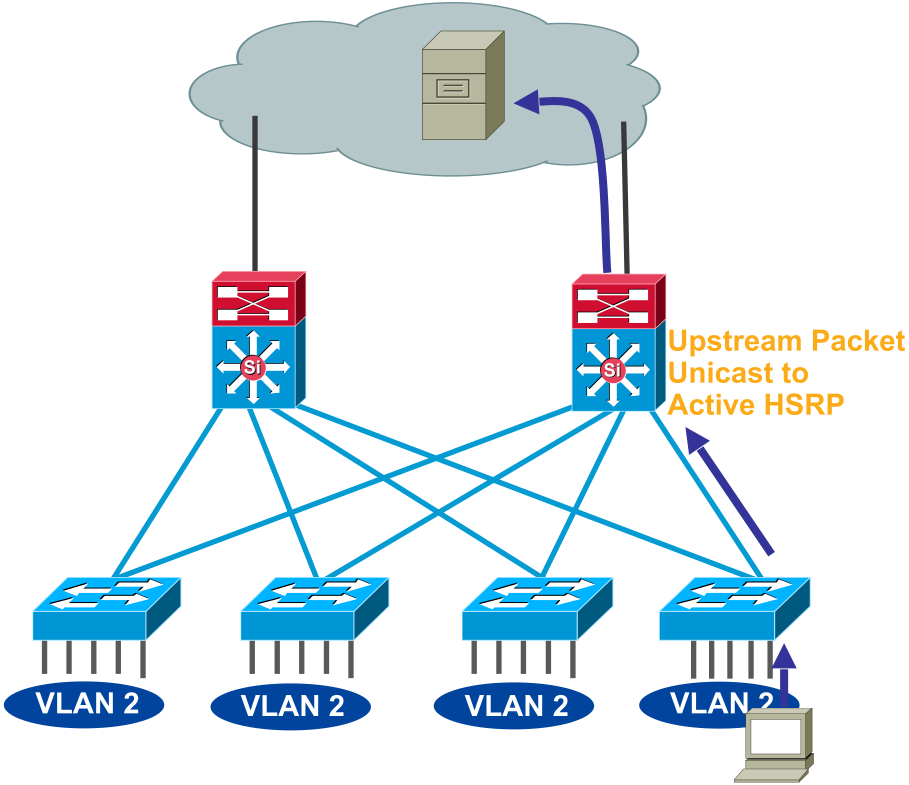
- One path upstream and two paths downstream
- CAM table entry ages out on standby HSRP
- Without a CAM entry packet is flooded to all ports in the VLAN



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

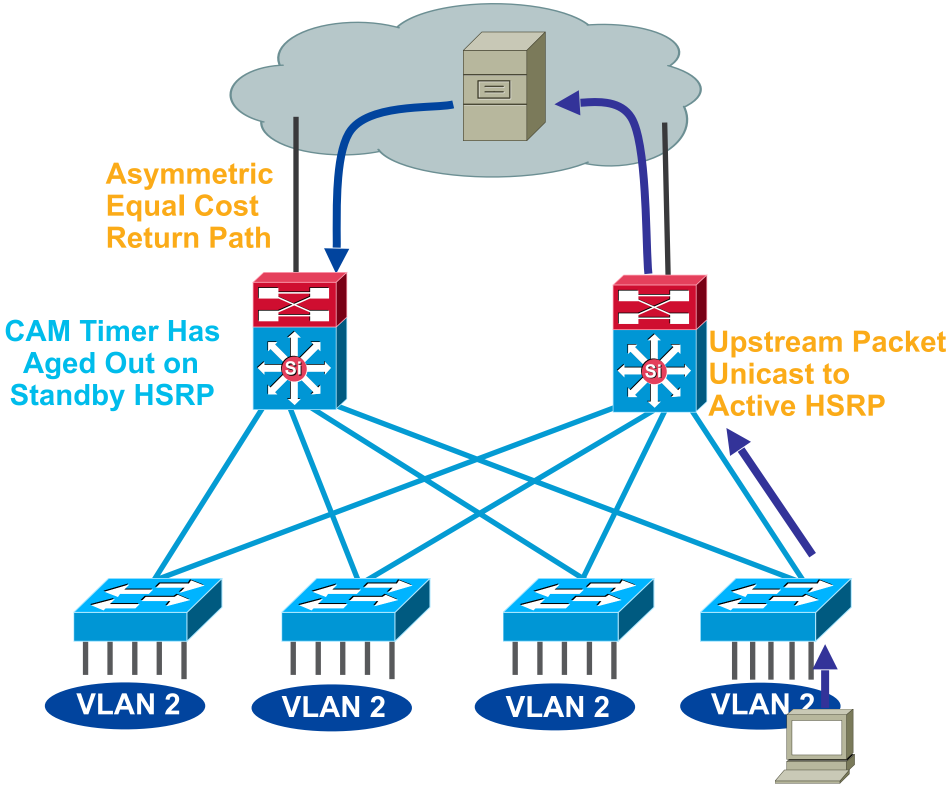


Upstream Packet Unicast to Active HSRP

Asymmetric Routing (Unicast Flooding)

Affects redundant topologies with shared L2 access

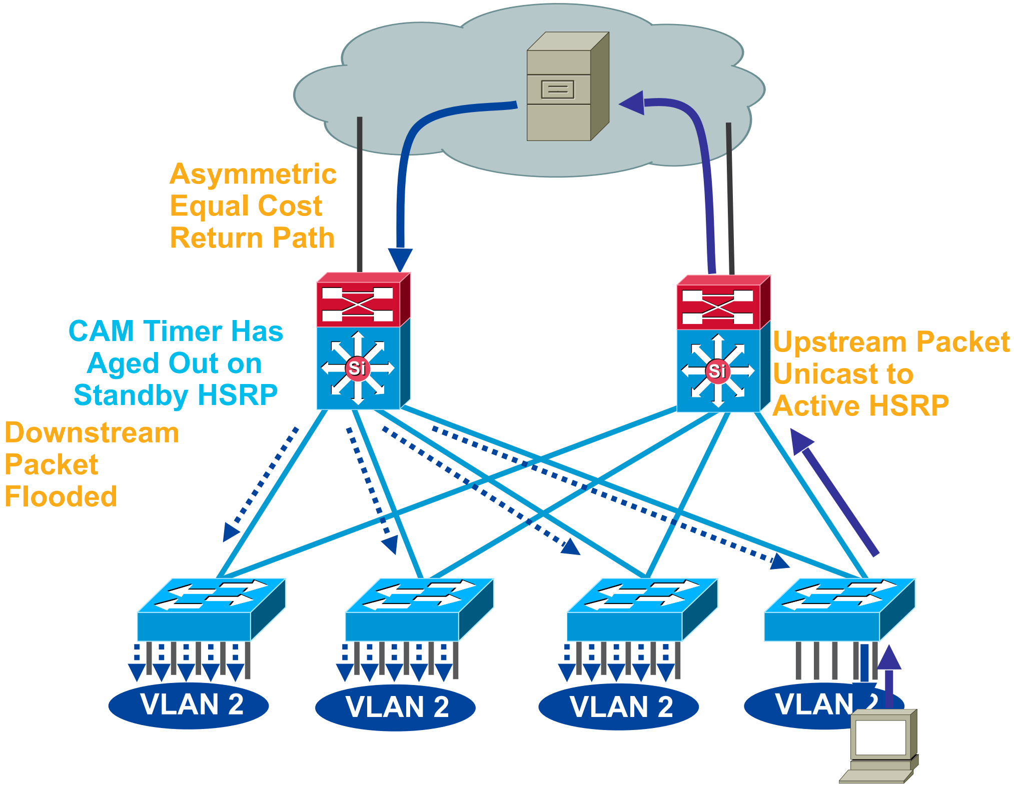
- One path upstream and two paths downstream
- CAM table entry ages out on standby HSRP
- Without a CAM entry packet is flooded to all ports in the VLAN


Asymmetric Equal Cost Return Path
CAM Timer Has Aged Out on Standby HSRP
Upstream Packet Unicast to Active HSRP
Si
Si
VLAN 2
VLAN 2
VLAN 2
VLAN 2

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



Asymmetric Equal Cost Return Path

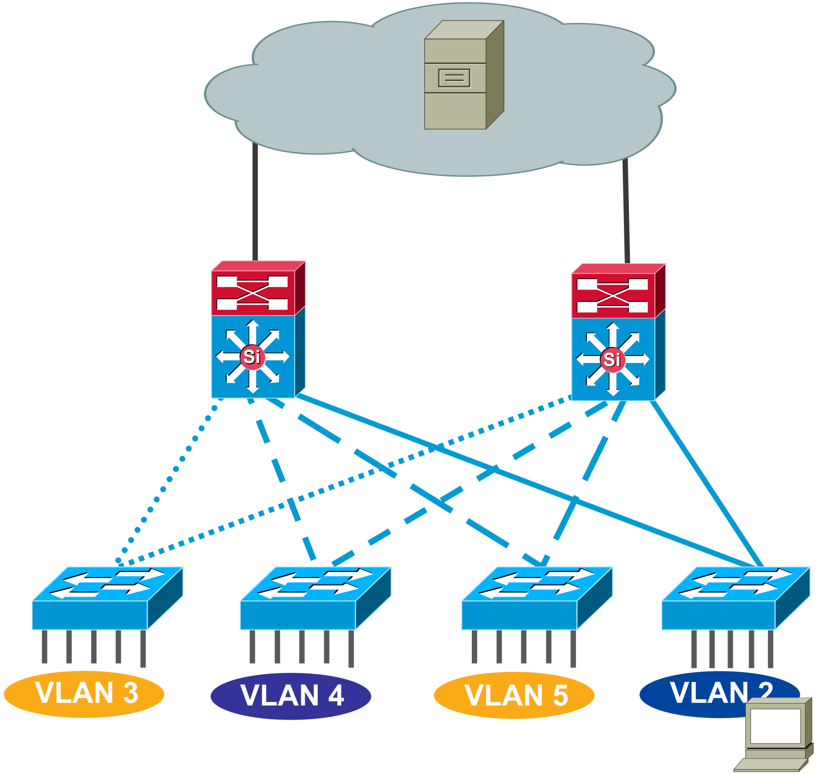
CAM Timer Has Aged Out on Standby HSRP

Downstream Packet Flooded

Upstream Packet Unicast to Active HSRP

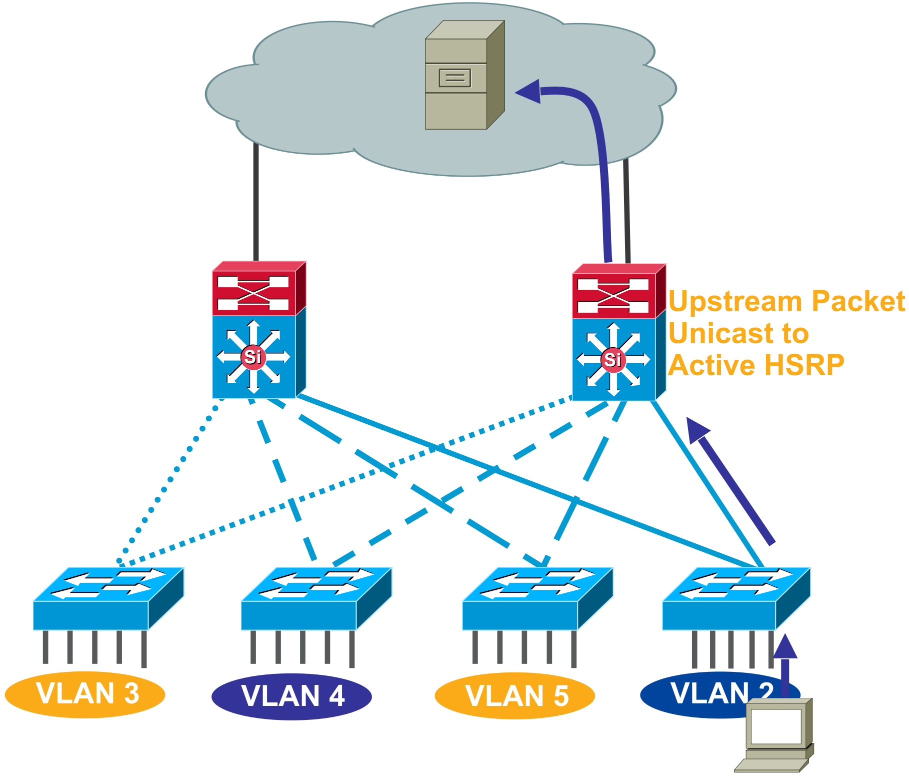
Best Practices Prevent Unicast Flooding

- 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



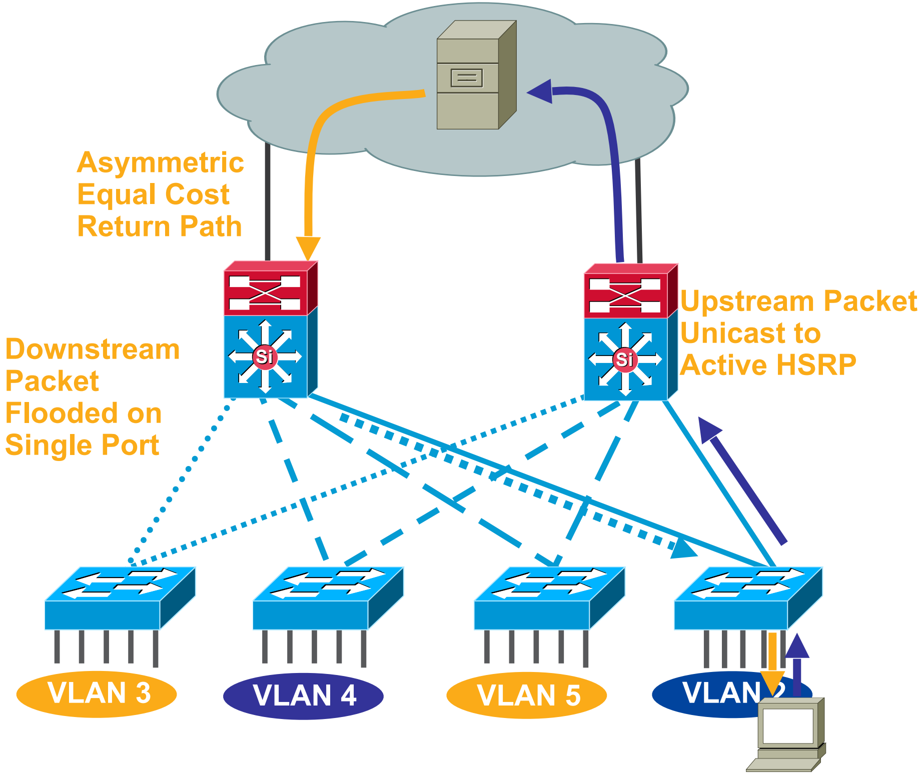
Best Practices Prevent Unicast Flooding

- 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



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

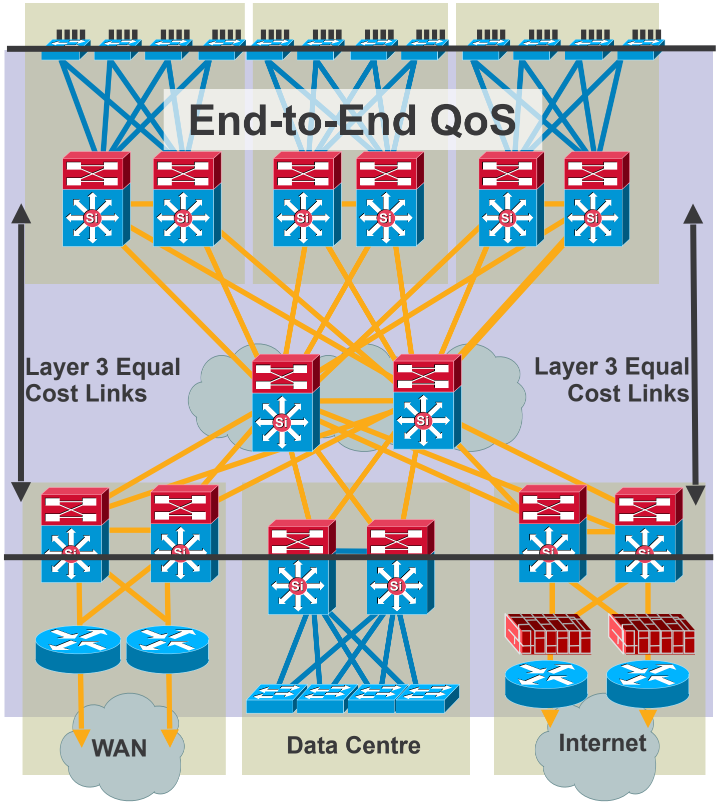

Asymmetric Equal Cost Return Path
Upstream Packet Unicast to Active HSRP
Downstream Packet Flooded on Single Port
Si
Si
VLAN 3
VLAN 4
VLAN 5
VLAN 2

Agenda

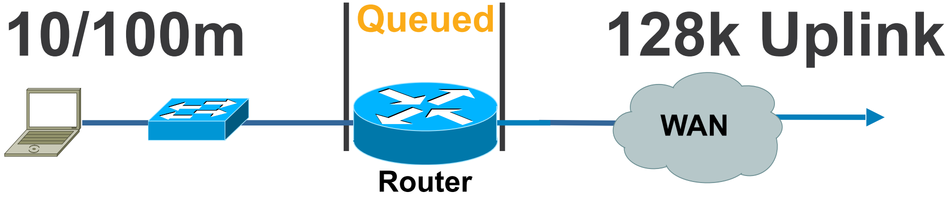
- Multilayer Campus Design Principles
- Foundation Services
- Campus Design Best Practices
- QoS Considerations
- Security Considerations
- Putting It All Together
- Summary

Best Practices - Quality of Service

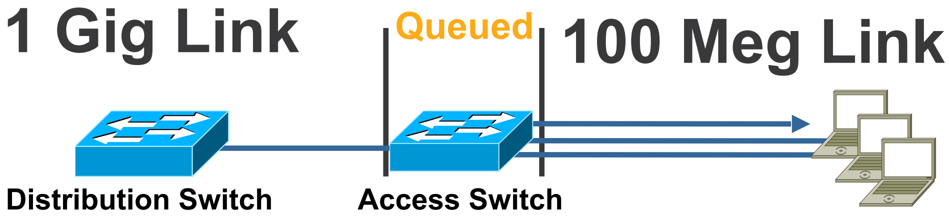
- Must be deployed end-to-end to be effective; all layers play different but equal roles
- Ensure that mission-critical applications are not impacted by link or transmit queue congestion
- Aggregation and rate transition points must enforce QoS policies
- Multiple queues with configurable admission criteria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



Transmit Queue Congestion

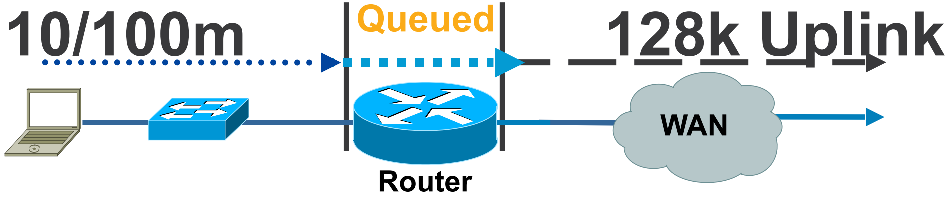


100 Meg in 128 Kb/S out—Packets Serialise in Faster than They Serialise Out
Packets Queued as They Wait to Serialise out Slower Link

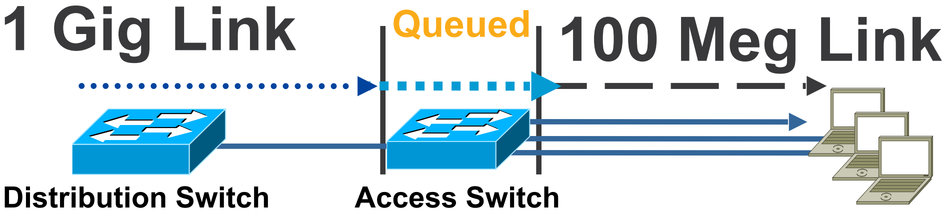


1 Gig In 100 Meg out—Packets Serialise in Faster than They Serialise Out
Packets Queued as They Wait to Serialise out Slower Link

Transmit Queue Congestion



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



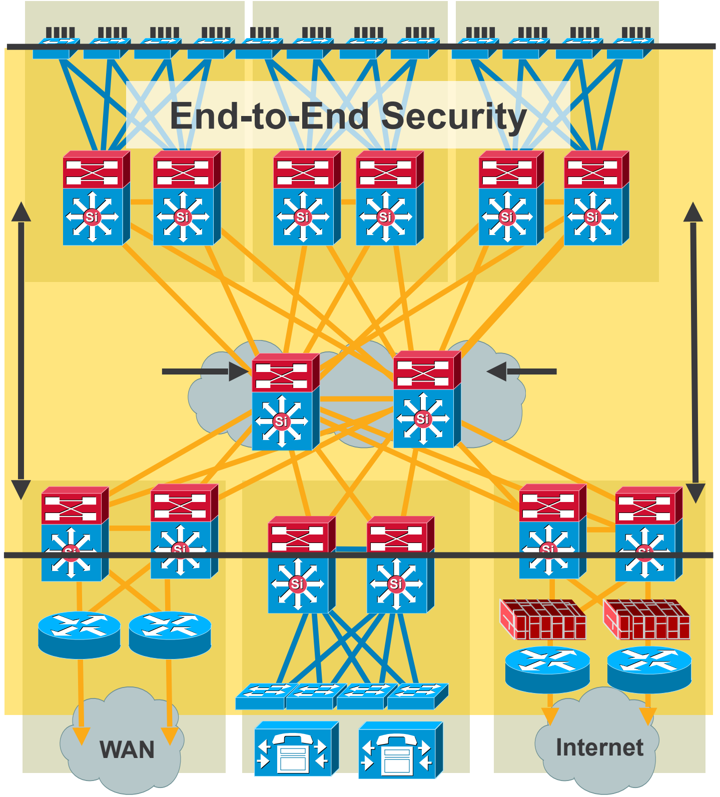
1 Gig In 100 Meg out—Packets Serialise in Faster than They Serialise Out
Packets Queued as They Wait to Serialise out Slower Link

Agenda

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

Best Practices - Campus Security

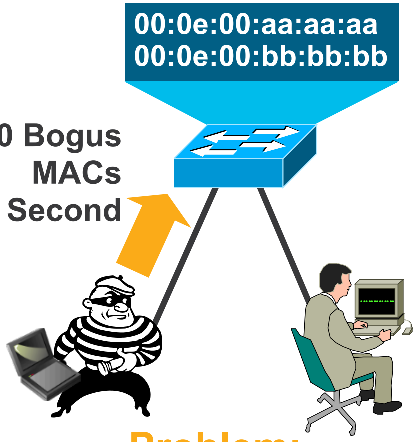
- CISF
  - Dynamic port security
  - DHCP snooping,
  - Dynamic ARP inspection
  - IP source guard



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

Securing Layer 2 from Surveillance Attacks

Cutting Off MAC-Based Attacks



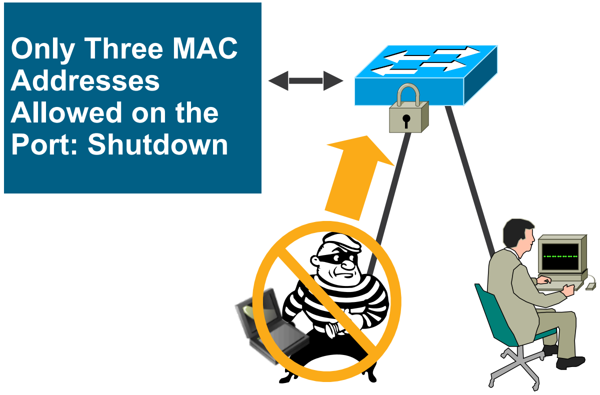
00:0e:00:aa:aa:aa
00:0e:00:bb:bb:bb

250,000 Bogus
MACs
per Second

Problem:

Script Kiddie Hacking Tools Enable Attackers 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



Only Three MAC Addresses Allowed on the Port: Shutdown

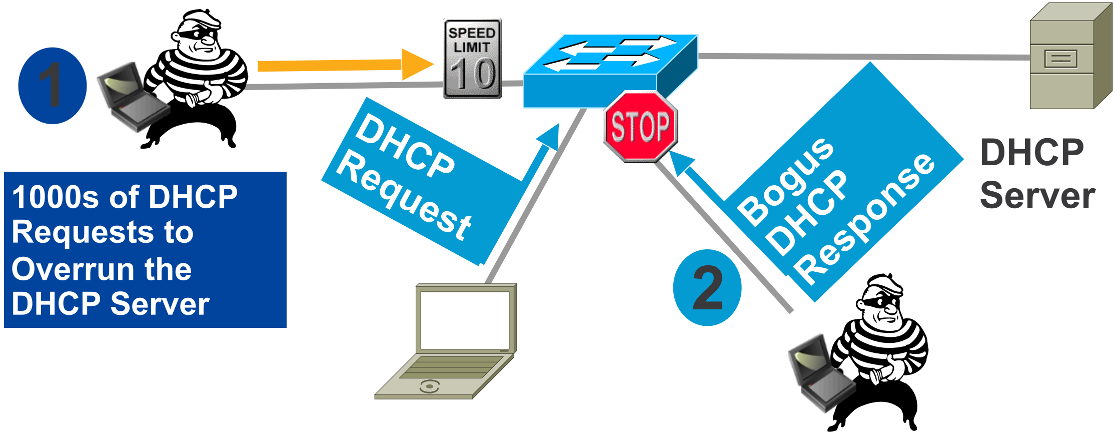
Solution:

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

```
switchport port-security
switchport port-security maximum 100
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

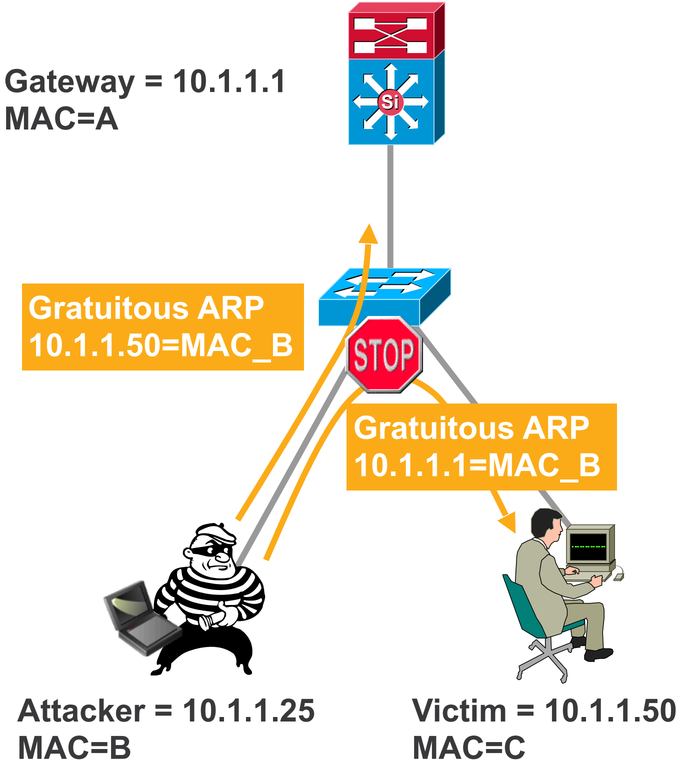


- DHCP requests (discover) and responses (offer) tracked
- Rate-limit requests on trusted interfaces; limits DoS attacks on DHCP server
- Deny responses (offers) on non trusted interfaces; stop malicious or errant DHCP server

Securing Layer 2 from Surveillance Attacks

Protection Against ARP Poisoning

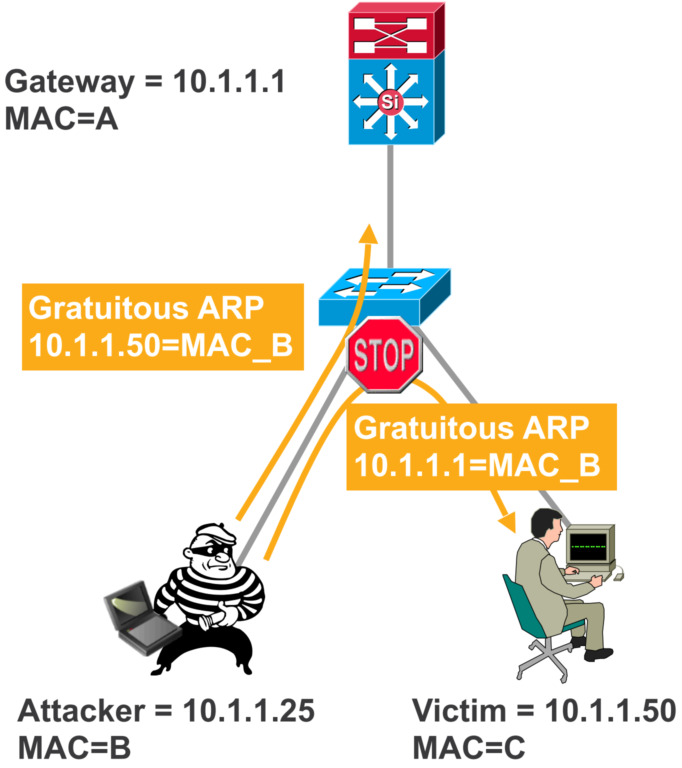
- Dynamic ARP inspection protects against ARP poisoning (ettercap, dsnif, arpspoof)
- Uses the DHCP snooping binding table
- Tracks MAC to IP from DHCP transactions
- Rate-limits ARP requests from client ports; stop port scanning
- Drop bogus gratuitous ARPs; stop ARP poisoning/MIM attacks



Securing Layer 2 from Surveillance Attacks

Protection Against ARP Poisoning

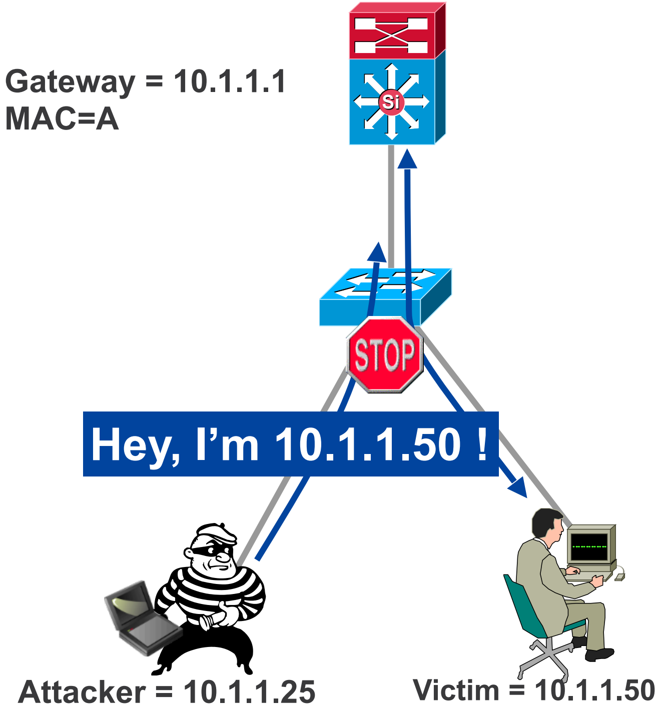
- Dynamic ARP inspection protects against ARP poisoning (ettercap, dsnif, arpspoof)
- Uses the DHCP snooping binding table
- Tracks MAC to IP from DHCP transactions
- Rate-limits ARP requests from client ports; stop port scanning
- Drop bogus gratuitous ARPs; stop ARP poisoning/MIM attacks



IP Source Guard

Protection Against Spoofed IP Addresses

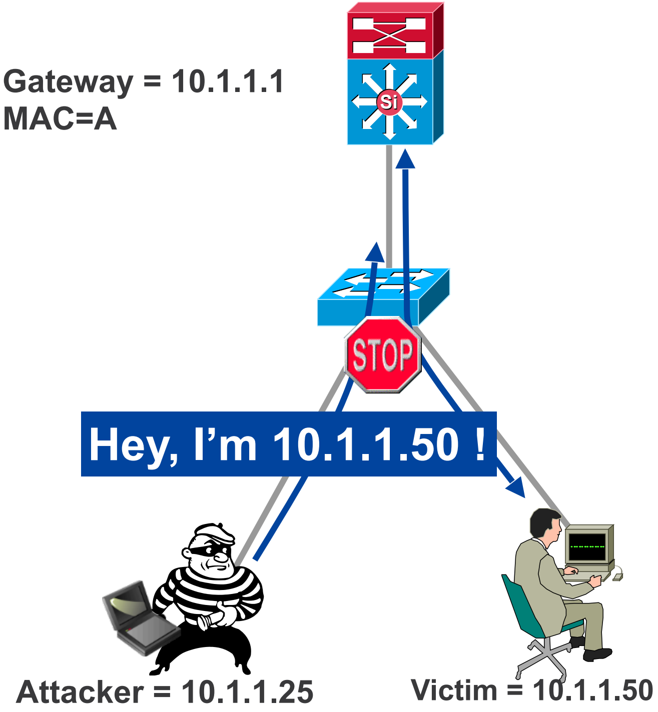
- IP source guard protects against spoofed IP addresses
- Uses the DHCP snooping binding table
- Tracks IP address to port associations
- Dynamically programs port ACL to drop traffic not originating from IP address assigned via DHCP



IP Source Guard

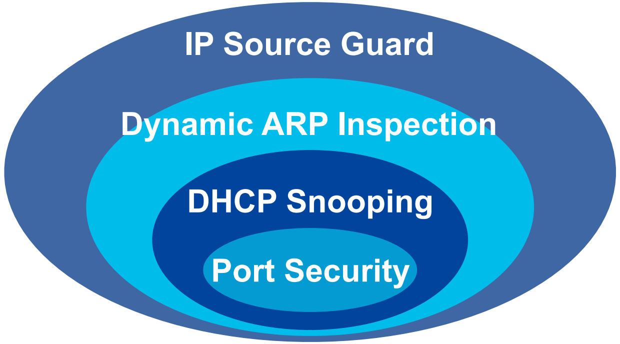
Protection Against Spoofed IP Addresses

- 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



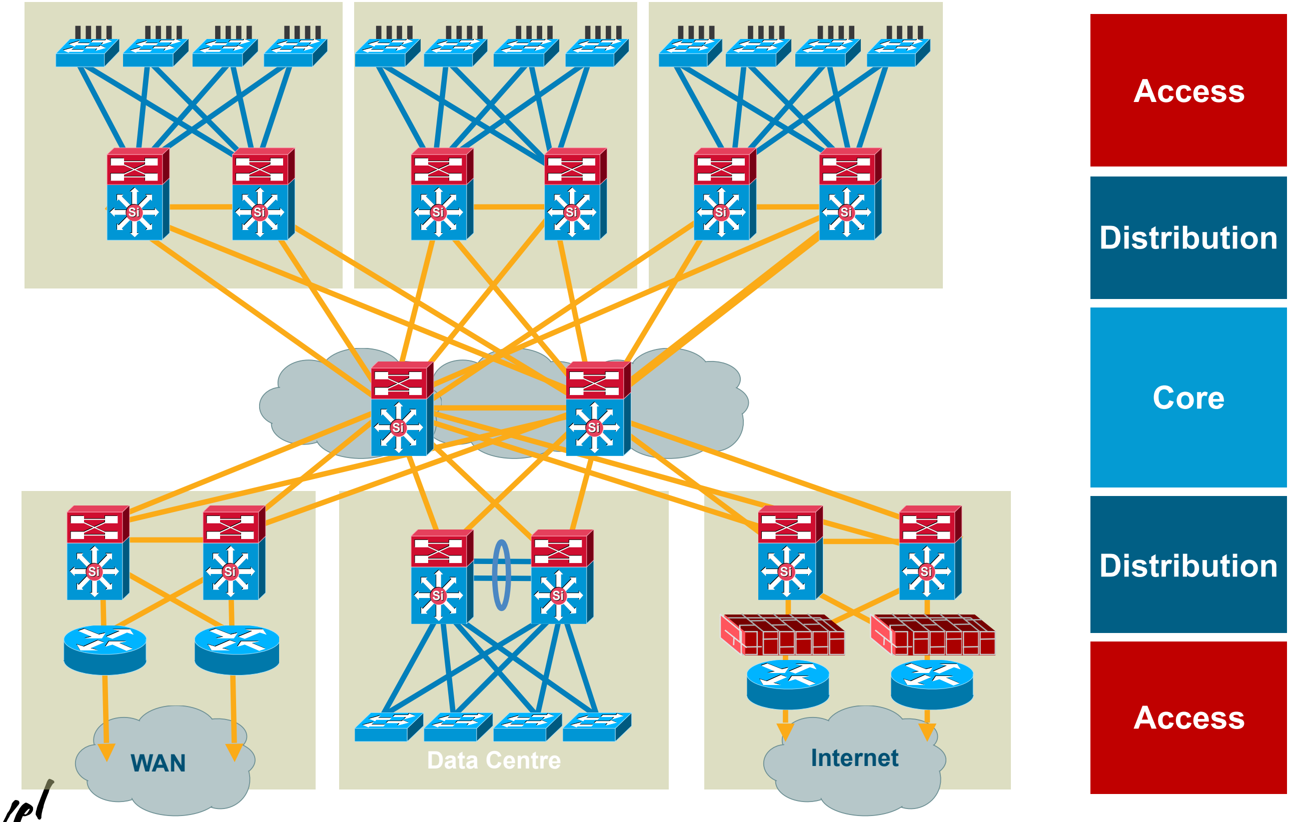
- Port security prevents MAC flooding attacks
- DHCP snooping prevents client attack on the switch and server
- Dynamic ARP Inspection adds security to ARP using DHCP snooping table
- IP source guard adds security to IP source address using DHCP snooping table

```
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
- QoS Considerations
- Security Considerations
- Putting It All Together
- Summary

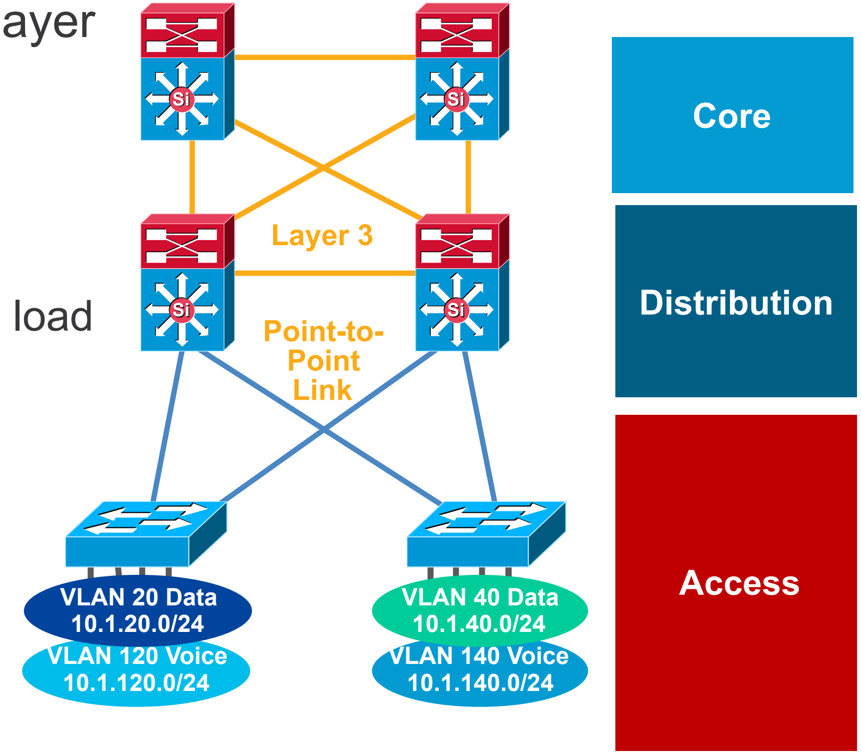
Hierarchical Campus



Layer 3 Distribution Interconnection

Layer 2 Access—No VLANs Span Access Layer

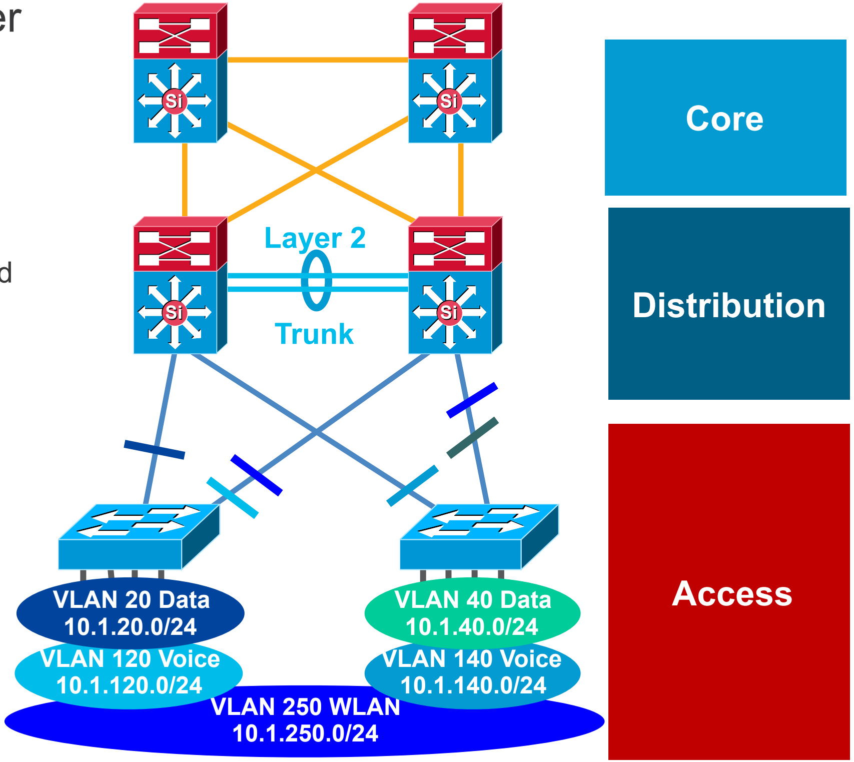
- Tune CEF load balancing
- Summarise routes towards core
- Limit redundant IGP peering
- STP Root and HSRP primary tuning or GLBP to load balance on uplinks
- Set trunk mode on/no-negotiate
- Disable Ether Channel unless needed
- Set port host on access layer ports:
  - Disable trunking
    Disable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


Layer 3
Point-to-Point Link
VLAN 20 Data 10.1.20.0/24
VLAN 120 Voice 10.1.120.0/24
VLAN 40 Data 10.1.40.0/24
VLAN 140 Voice 10.1.140.0/24
Core
Distribution
Access
Si

Layer 2 Distribution Interconnection

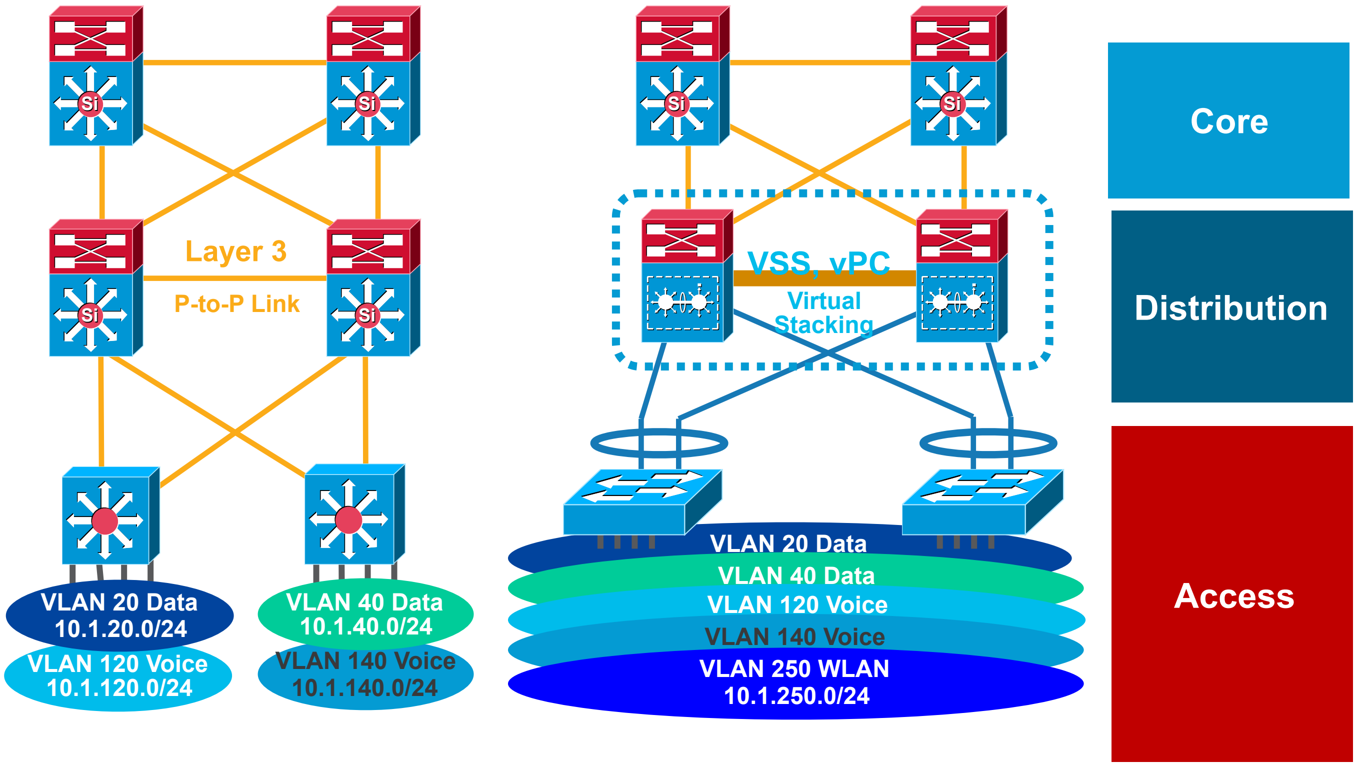
Layer 2 Access - Some VLANs Span Access Layer

- Tune CEF load balancing
- Summarise 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 Ether Channel unless needed
- RootGuard on downlinks
- LoopGuard on uplinks
- Set port host on access Layer ports:
  - Disable trunking
    Disable Ether Channel
    Enable PortFast
- RootGuard or BPDU-Guard
- Use security features



Routed Access and Virtual Switching System

Evolutions of and Improvements to Existing Designs

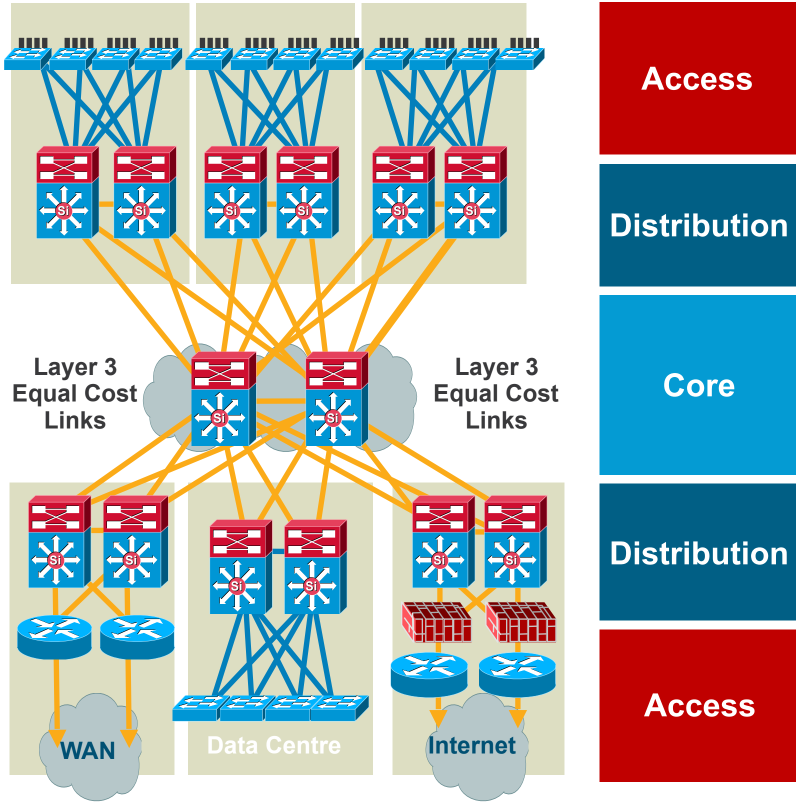


Agenda

- Multilayer Campus Design Principles
- Foundation Services
- Campus Design Best Practices
- 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
- Utilises Layer 3 routing for load balancing, fast convergence, scalability, and control



Hierarchical Network Design

Without a Rock Solid Foundation the Rest Doesn’t Matter

Access

Distribution

Core

Distribution

Access

Load Balancing

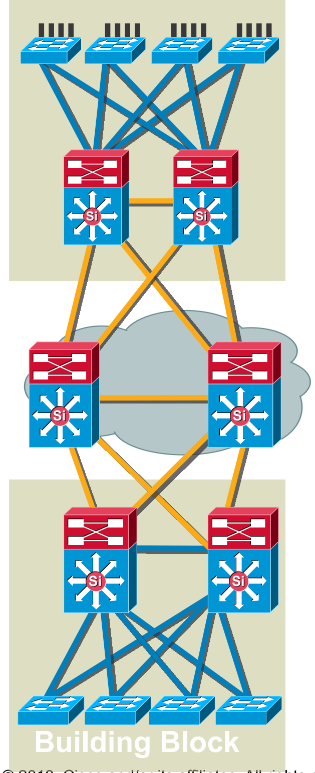
Trunking

GLBP

HSRP

Routing

Spanning Tree


Building Block