Building A Resilient Campus: Fundamentals and Best Practices

Chara Kontaxi

Systems Engineer,
ckontaxi@cisco.com

The Resilient Enterprise Campus

High-Availability Design Requirements

- Campus network design is evolving in response to multiple drivers

  User Expectations: Always ON Access to communications

  Business Requirements: Globalization means true 7x24x365

  Technology Requirements: Unified Communications

  Unexpected Requirements: Worms, Viruses, …

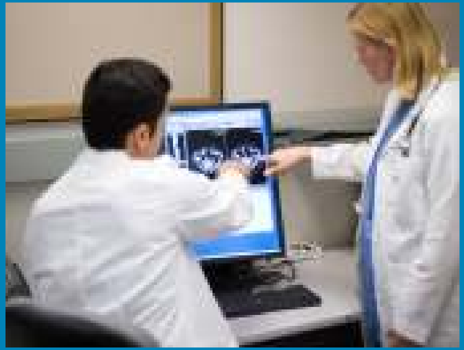
- Designing for availability is no longer just concerned with simple component failures

- Campus design needs to evolve to a ‘resilient’ model

Requires a Structured ‘and’ Resilient Design

Global Enterprise Availability

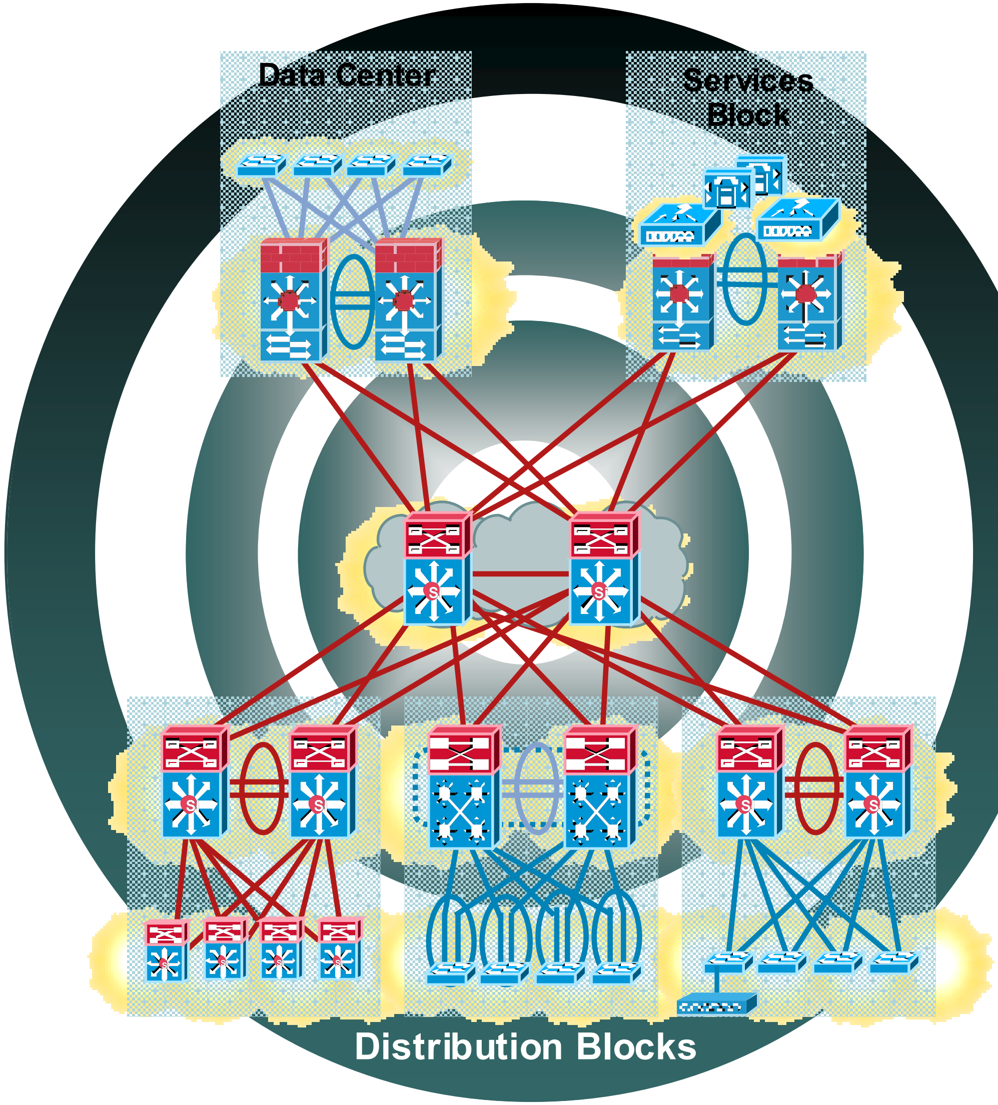
Collaboration and Real-Time Communication



Security

Agenda

- Multilayer Campus Design principles
- Campus Design Best Practices
- Hardening the Campus Network Design
- Summary



High-Availability Campus Design
Structure, Modularity, and Hierarchy

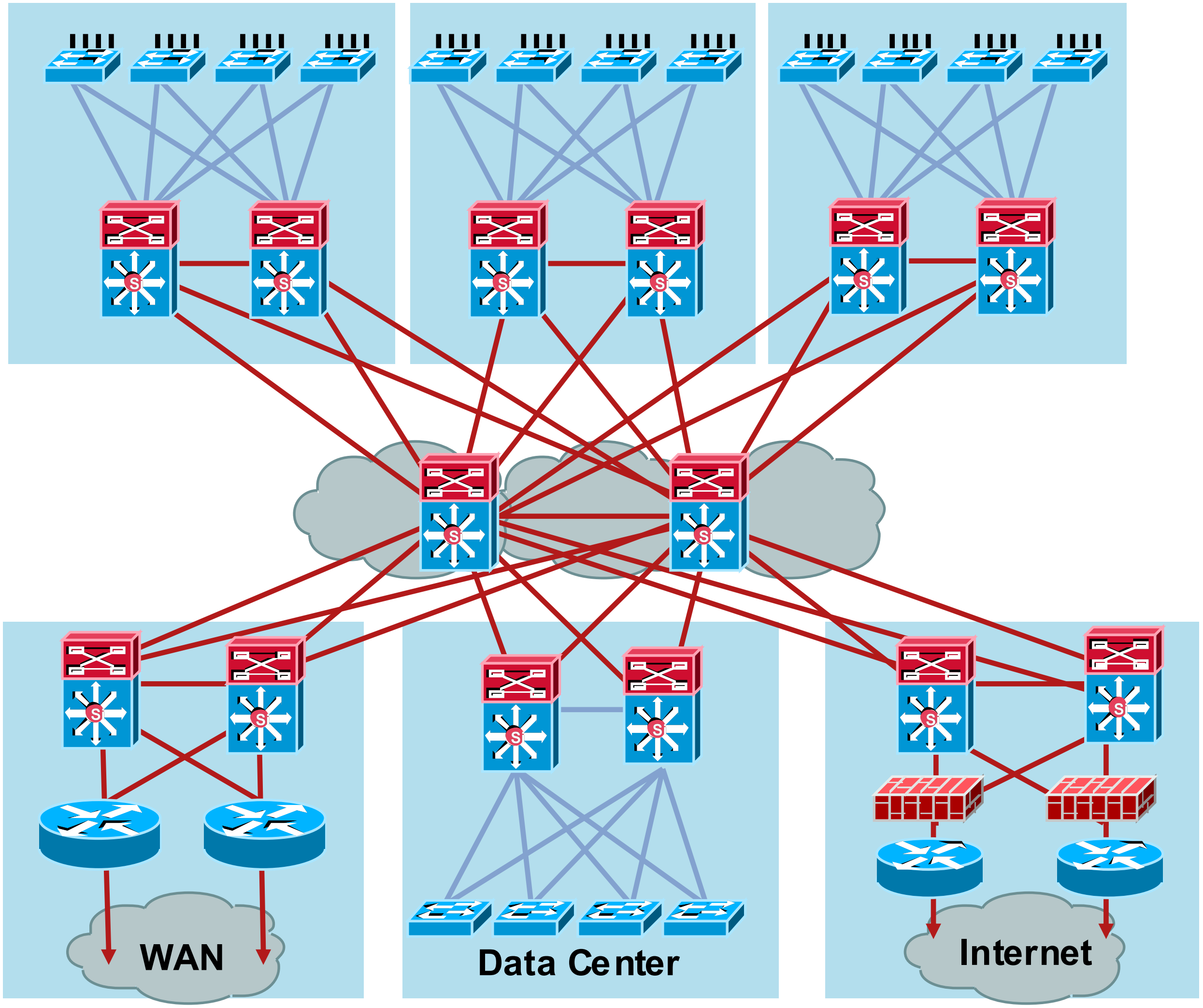
Access

Distribution

Core

Distribution

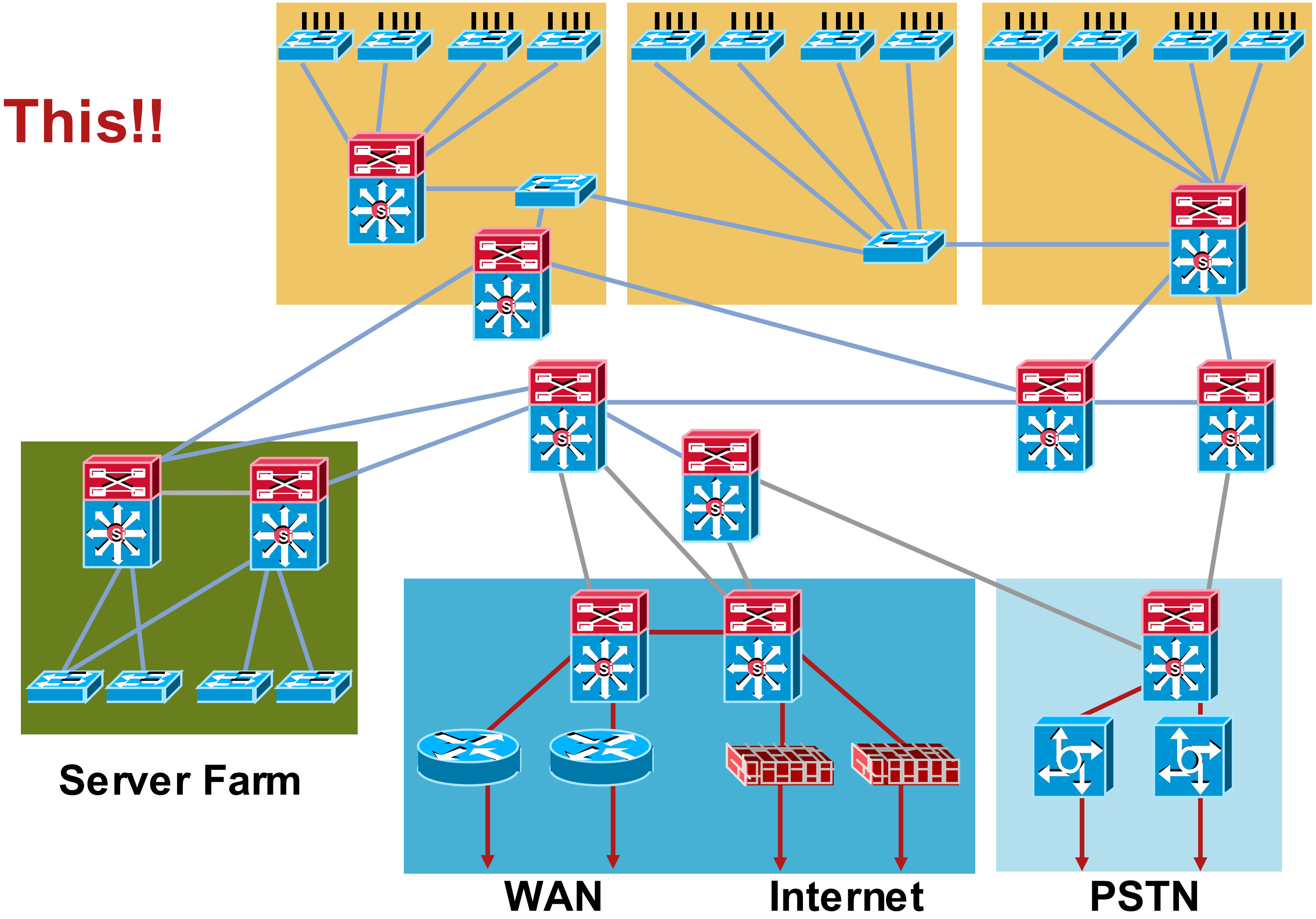
Access



Hierarchical Campus Network

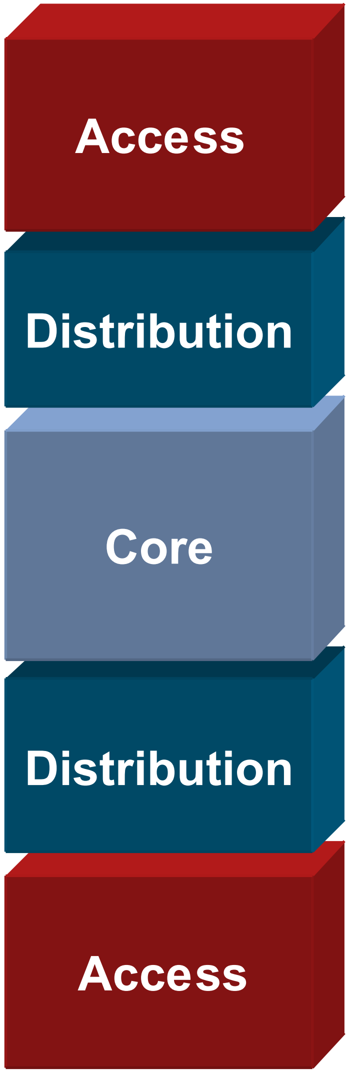
Structure, Modularity and Hierarchy

Not This!!



Hierarchical Network Design

Without a Rock Solid Foundation the Rest Doesn’t Matter



Access

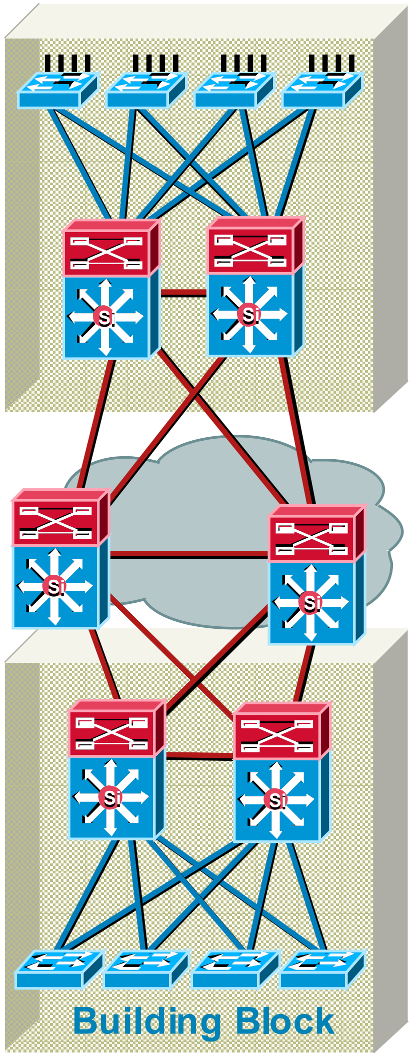
Distribution

Core

Distribution

Access

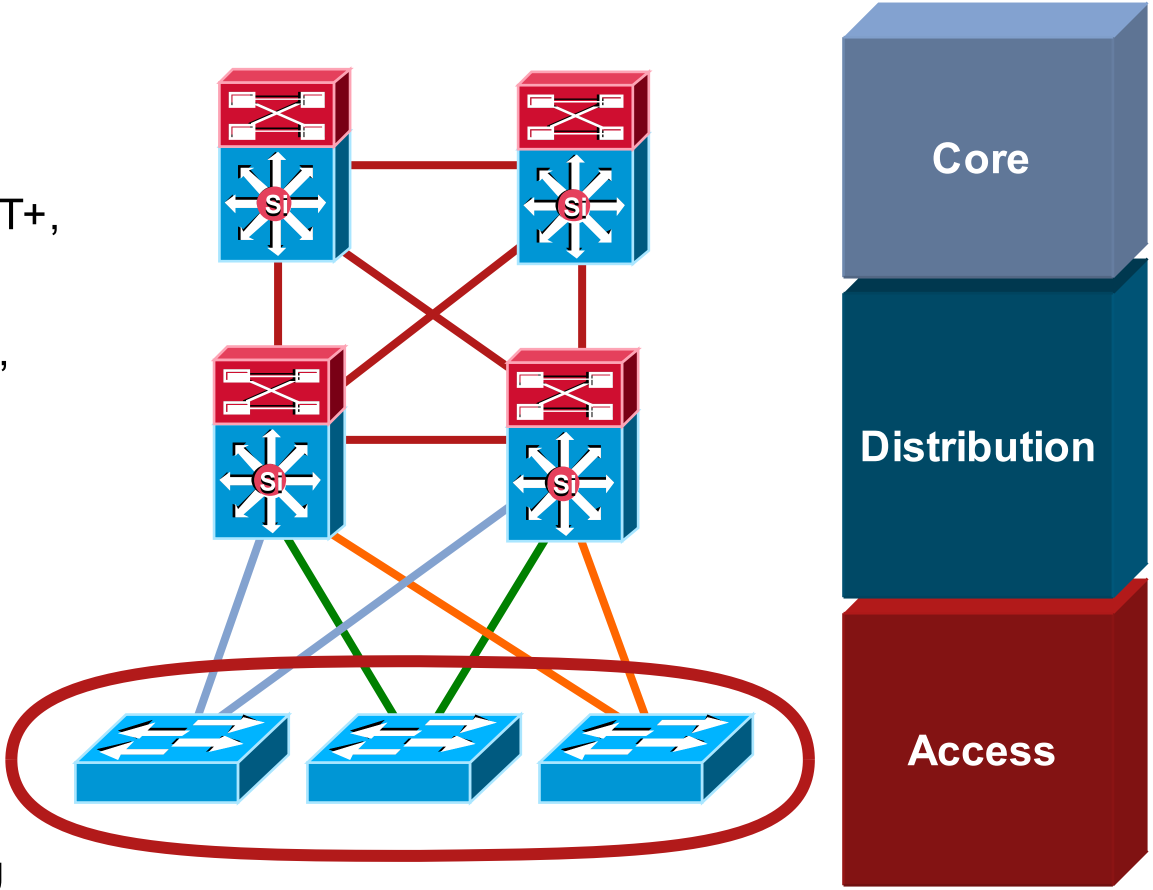
- Offers hierarchy—each layer has specific role
- Modular topology—building blocks
- Easy to grow, understand, and troubleshoot
- Creates small fault domains—clear demarcations and isolation
- Promotes load balancing and redundancy
- Promotes deterministic traffic patterns
- Incorporates balance of both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



Access Layer

Feature Rich Environment – Not JUST connectivity

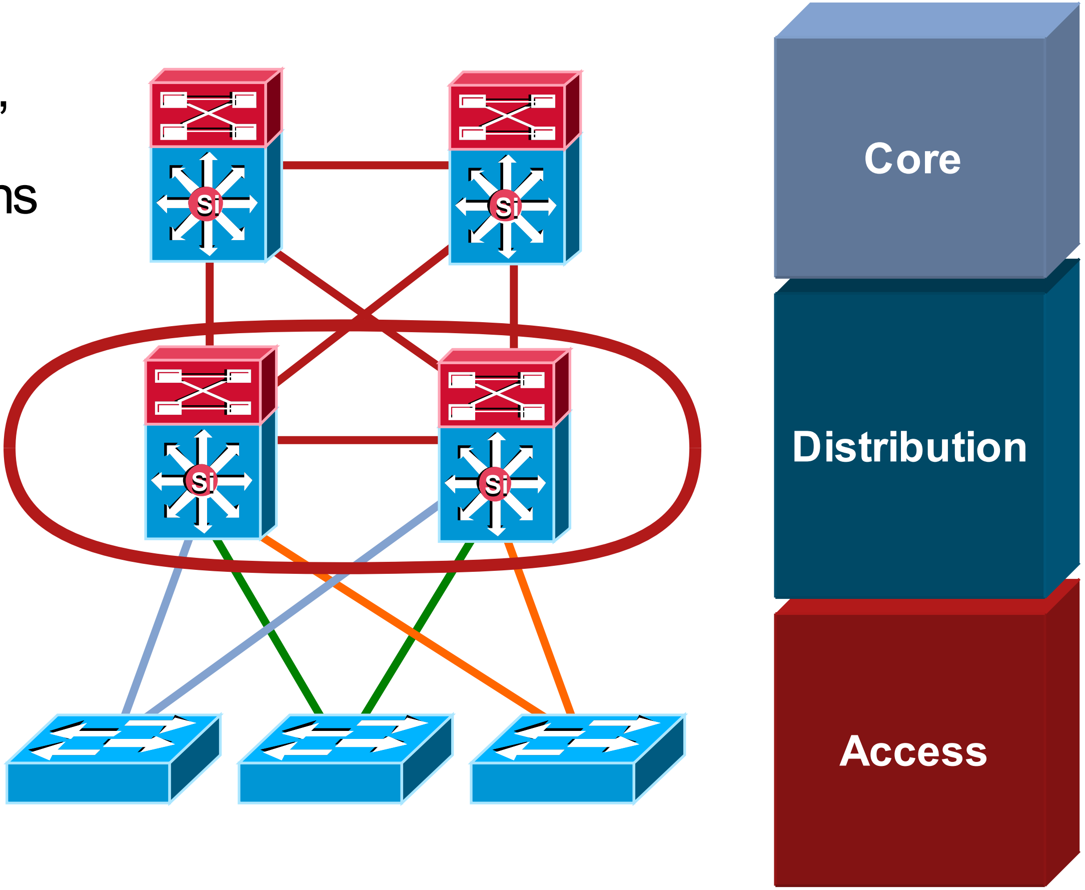
- Layer 2/Layer 3 feature rich environment; convergence, HA, security, QoS, IP multicast, etc.
- Intelligent network services: PVST+, Rapid PVST+, EIGRP, OSPF, PAgP/LACP, UDLD, etc.
- Intelligent network services: QoS, broadcast suppression, IGMP snooping
- Integrated security features IBNS (802.1x), port security, DHCP snooping, DAI, IPSG, etc.
- Automatic phone discovery, 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

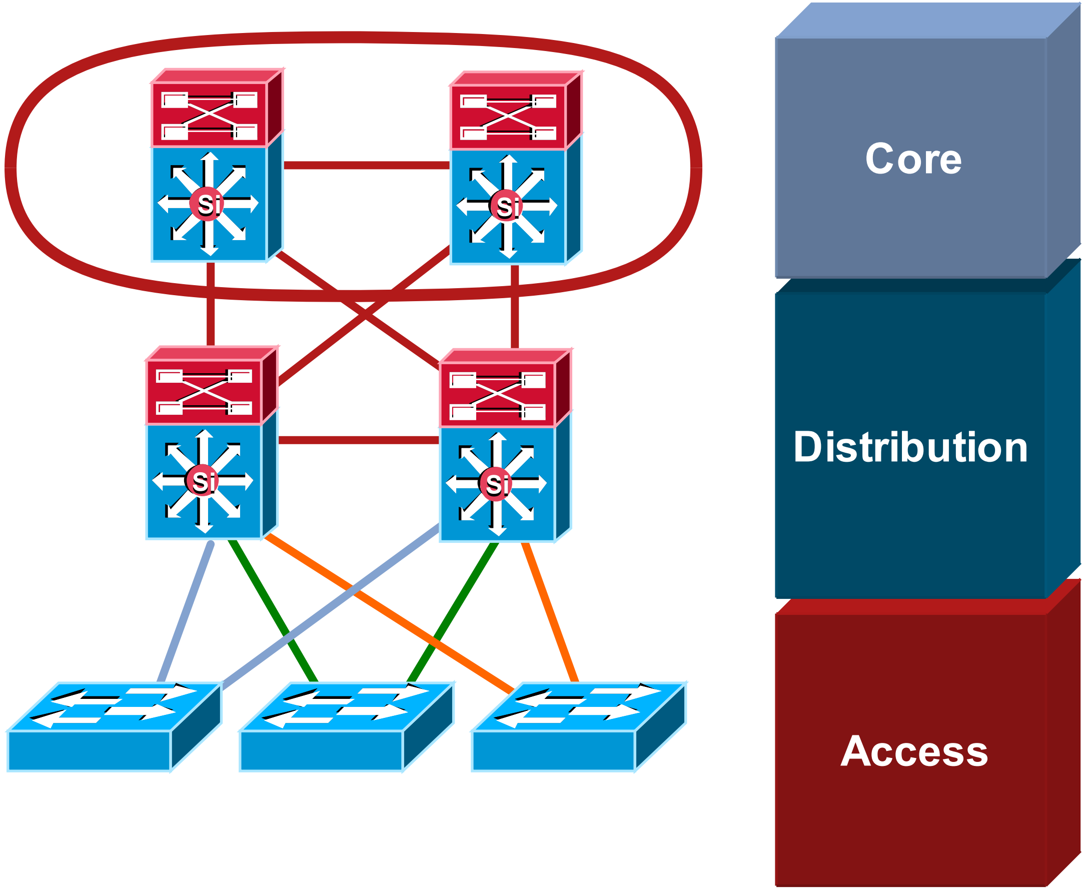
- Availability, load balancing, QoS and provisioning are the important considerations at this layer
- Aggregates wiring closets (access layer) and uplinks to core
- Protects core from high density peering and problems in access layer
- Route summarization, fast convergence, redundant path load sharing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

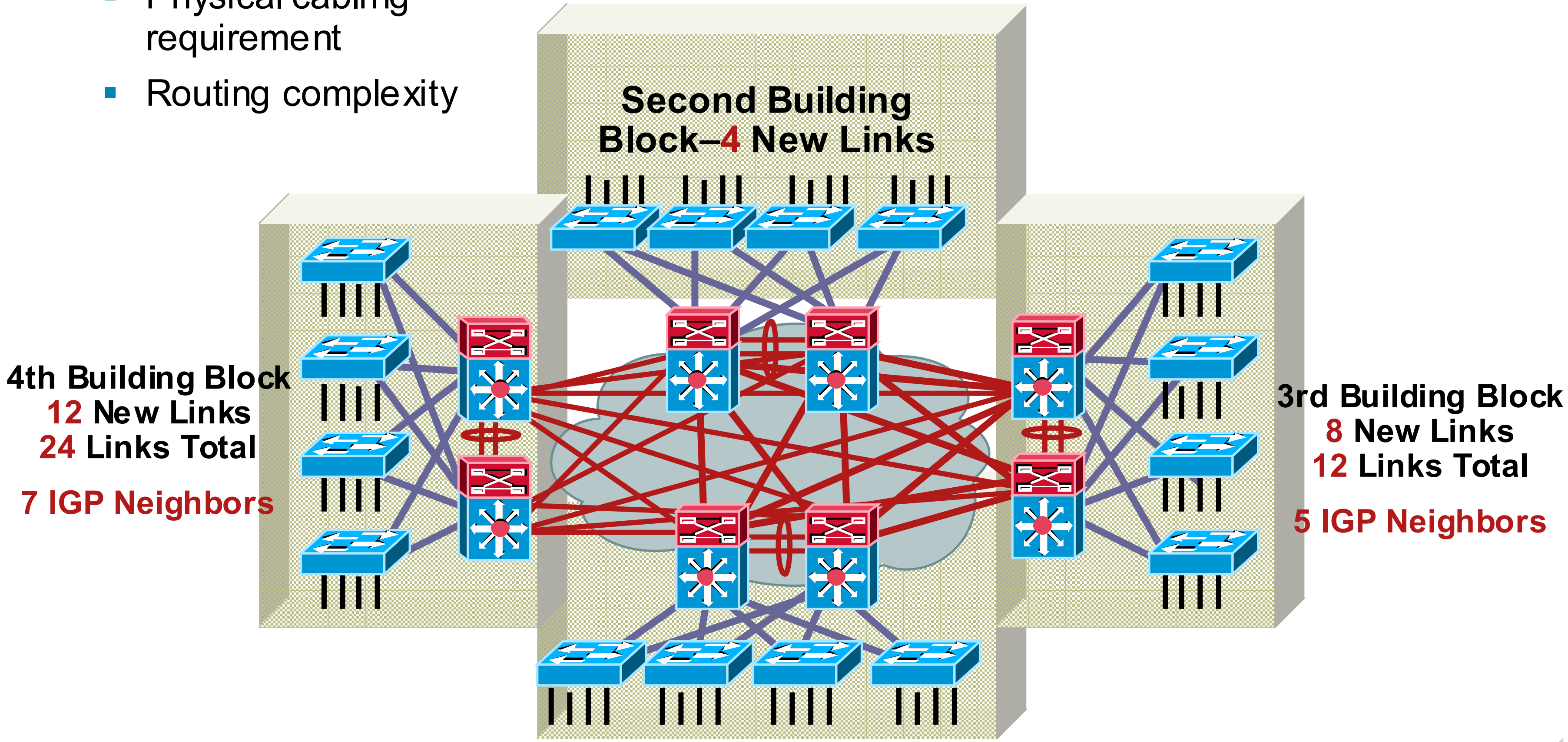


Do I Need a Core Layer?

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

No Core

- Fully meshed distribution layers
- Physical cabling requirement
- Routing complexity

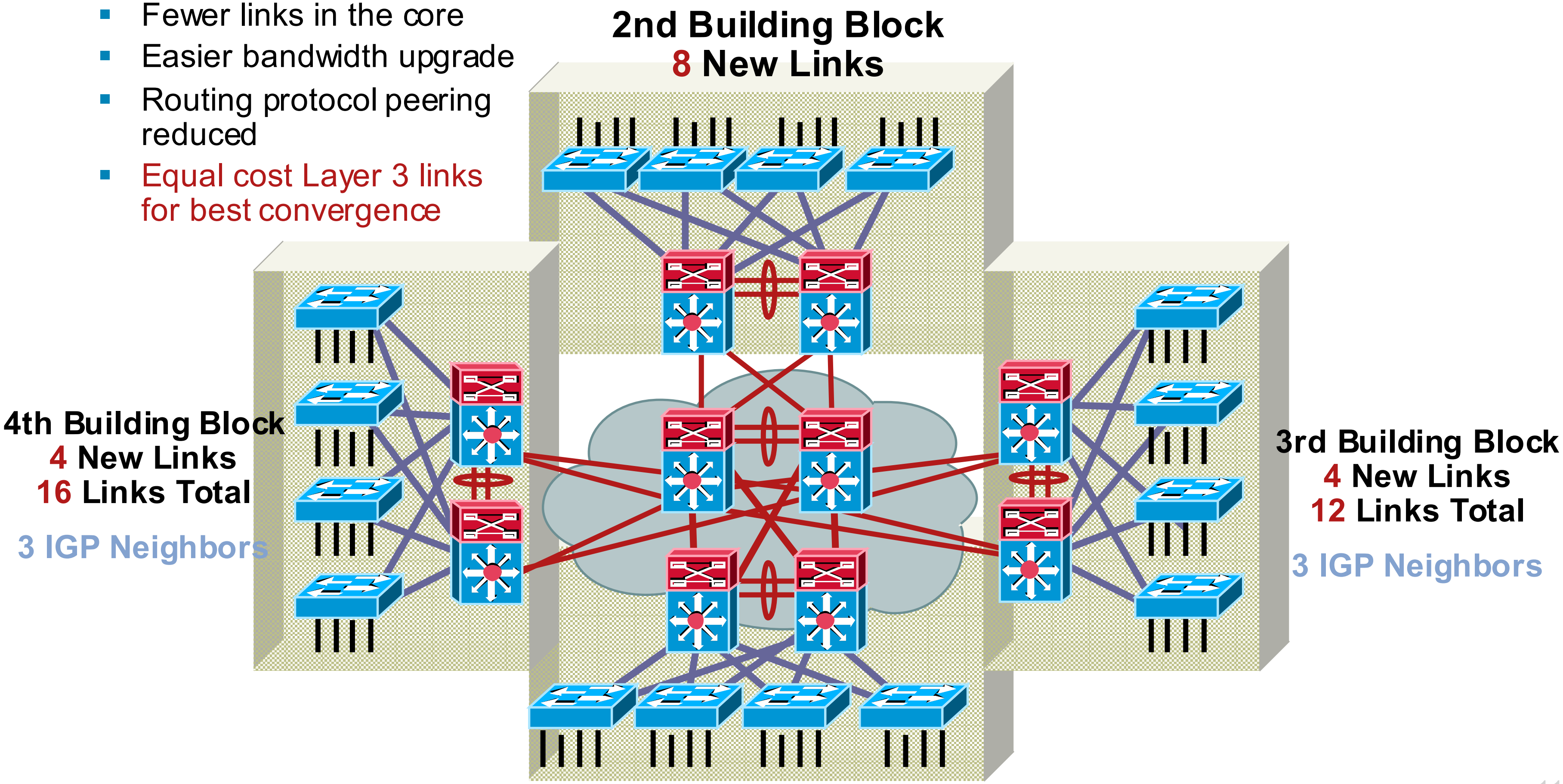


Do I Need a Core Layer?

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

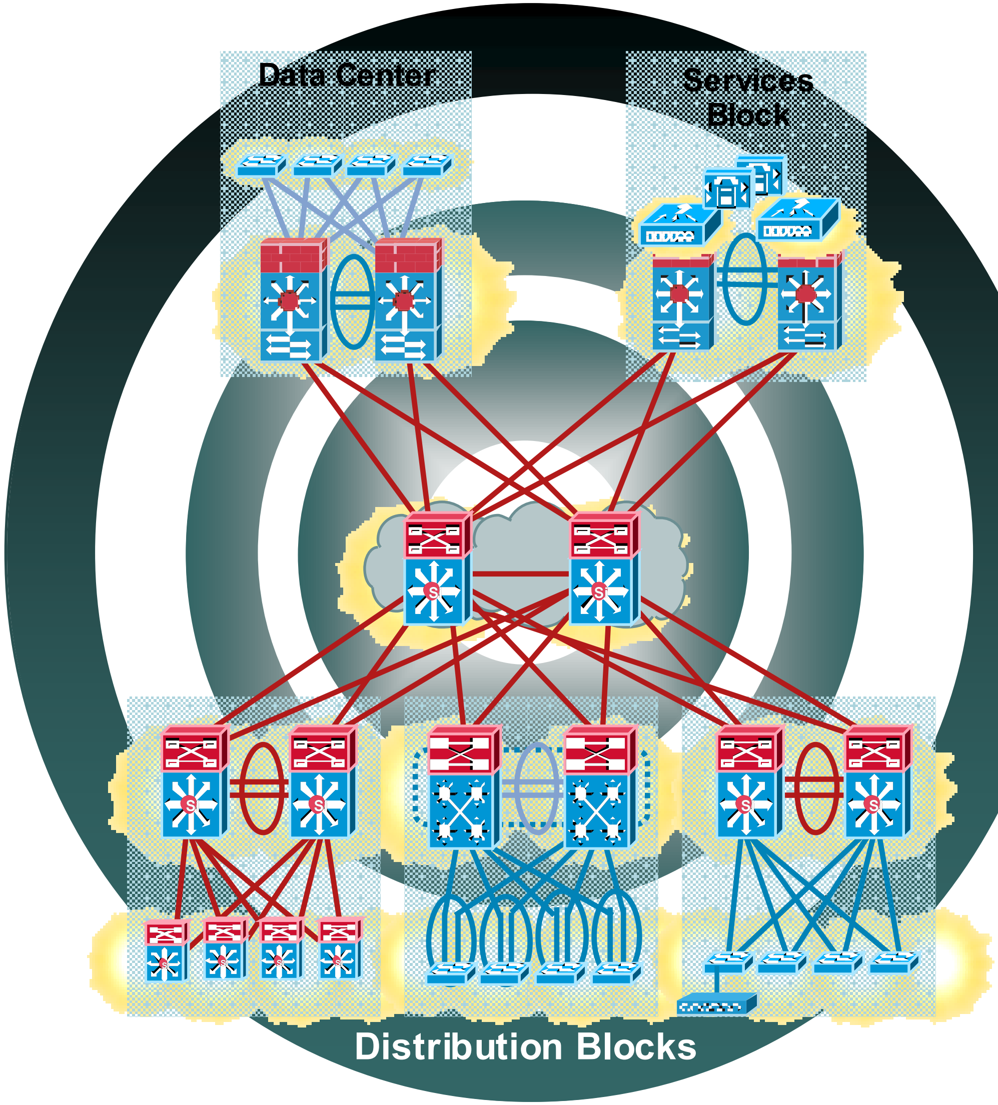
Dedicated Core Switches

- 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



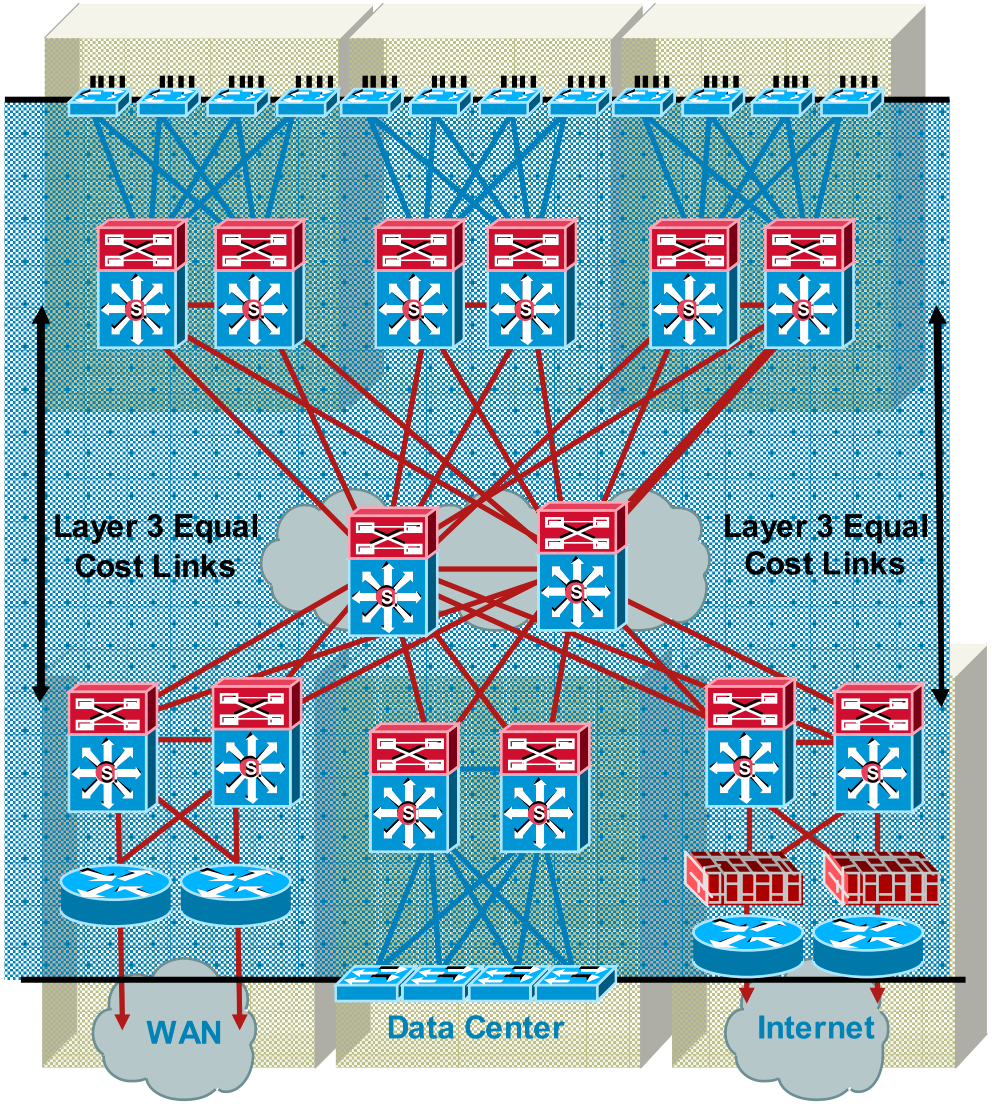
Agenda

- Multilayer Campus Design principles
- Campus Design Best Practices
- Hardening the Campus Network Design
- Summary



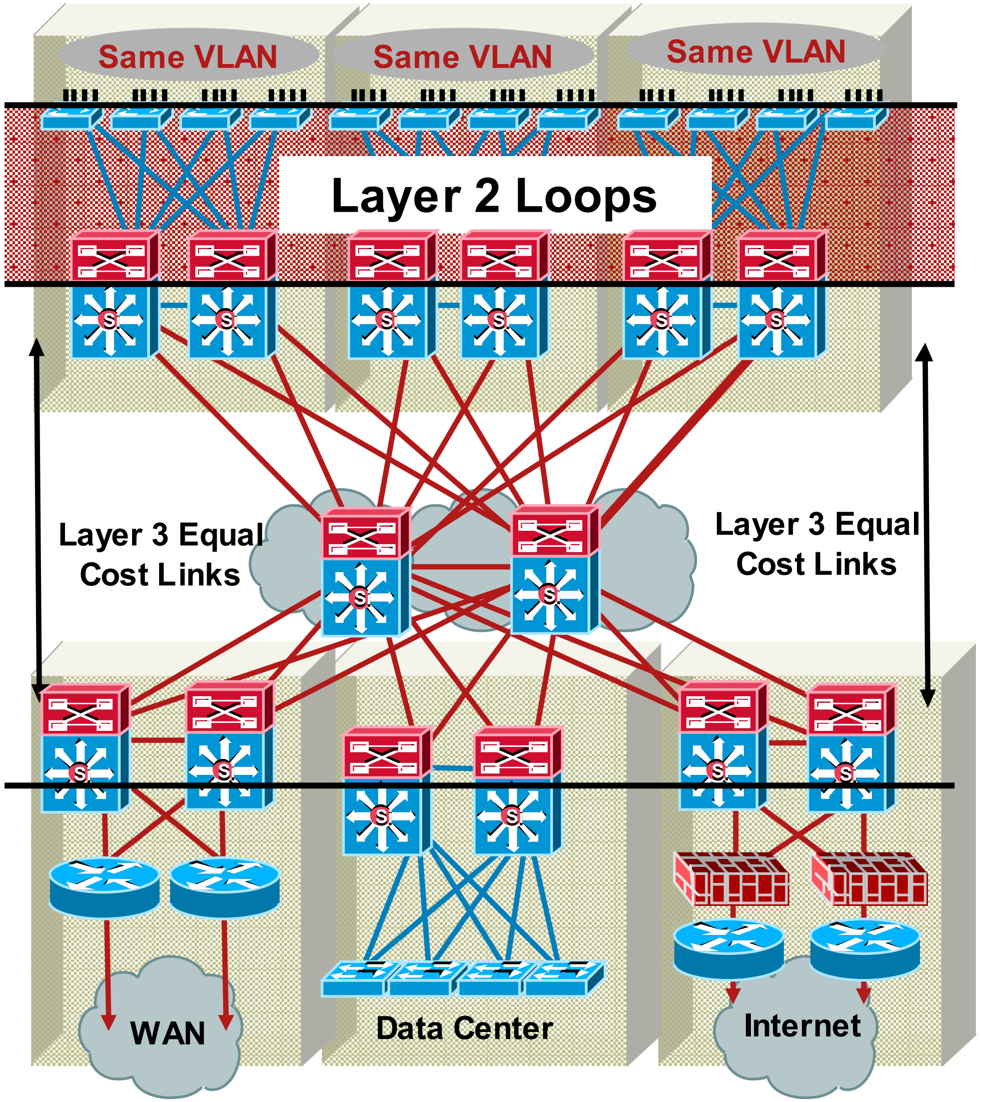
Best Practices— Layer 1 Physical Things

- Use point-to-point interconnections - no L2 aggregation points (Hubs) between nodes as HW detection and recovery is both faster and more deterministic
- Use fiber for best convergence (default debounce timer on GE and 10GE fiber linecards is 10 msec while minimum debounce for copper is 300 msec)
- Use configuration on the physical interface not VLAN/SVI when possible



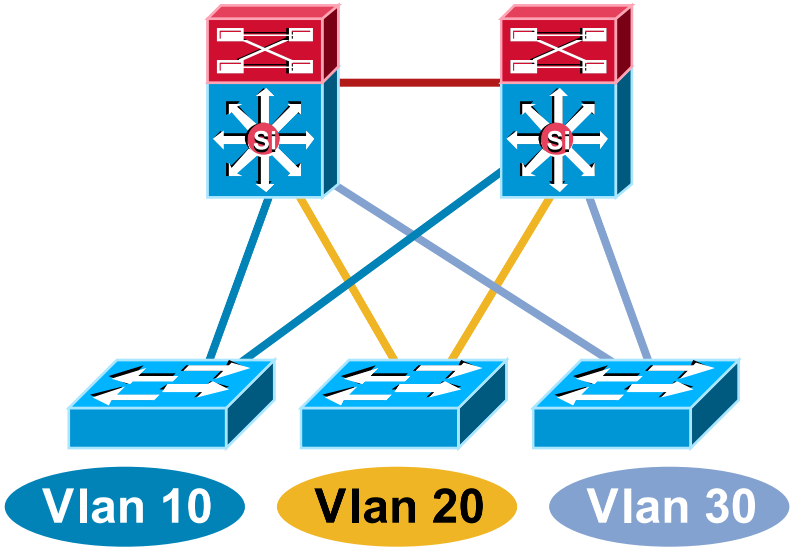
Best Practices— Spanning Tree Configuration

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

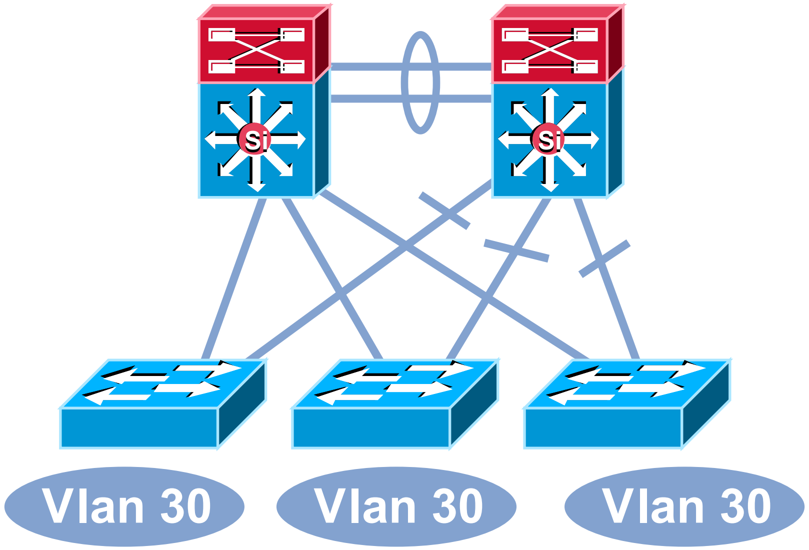


Multilayer Network Design

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
- PVST+ (802.1d)

  Traditional spanning tree implementation

- Rapid PVST+ (802.1w)

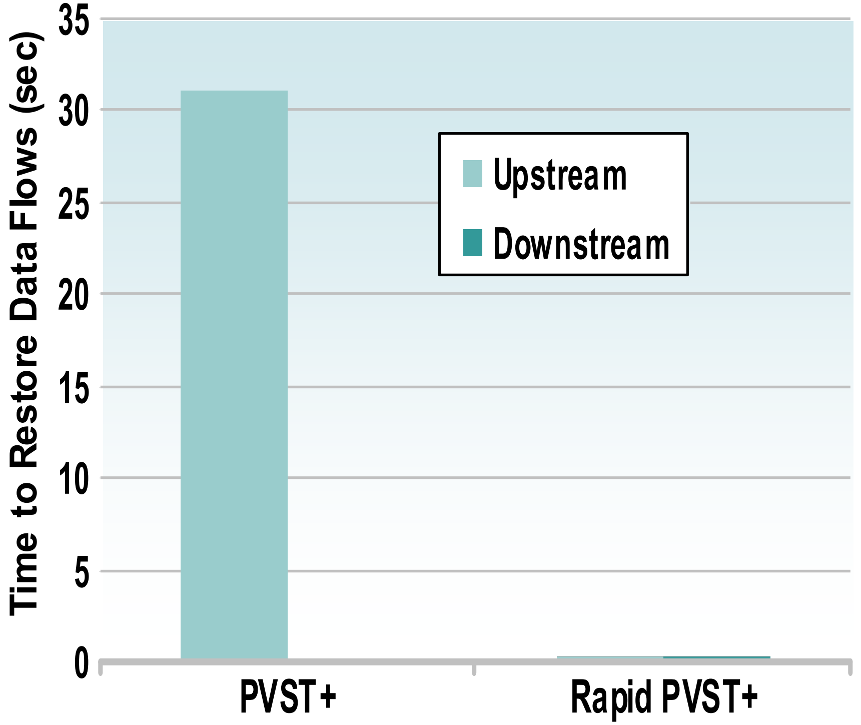
  Scales to large size (~10,000 logical ports)

  Easy to implement, proven, scales

- MST (802.1s)

  Permits very large scale STP implementations (~30,000 logical ports)

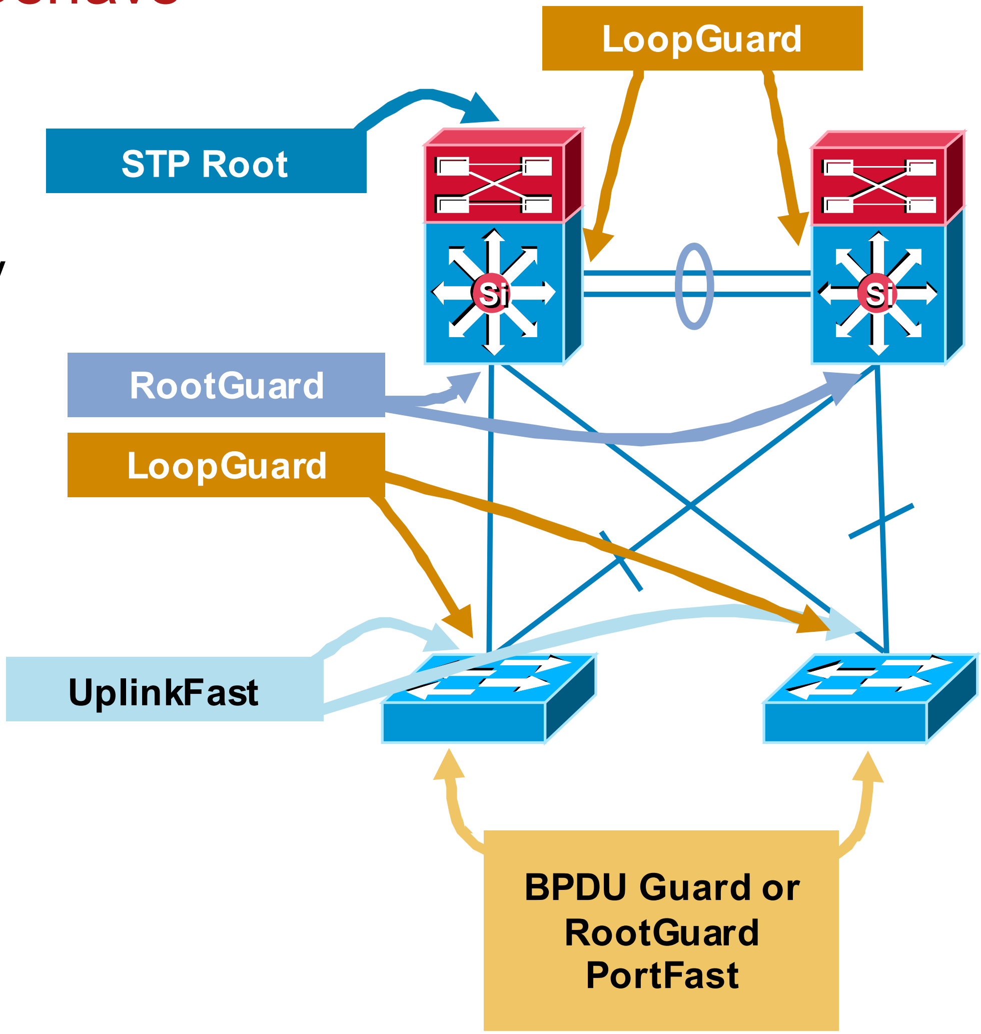
  Not as flexible as Rapid PVST+



Layer 2 Hardening

Spanning Tree Should Behave the Way You Expect

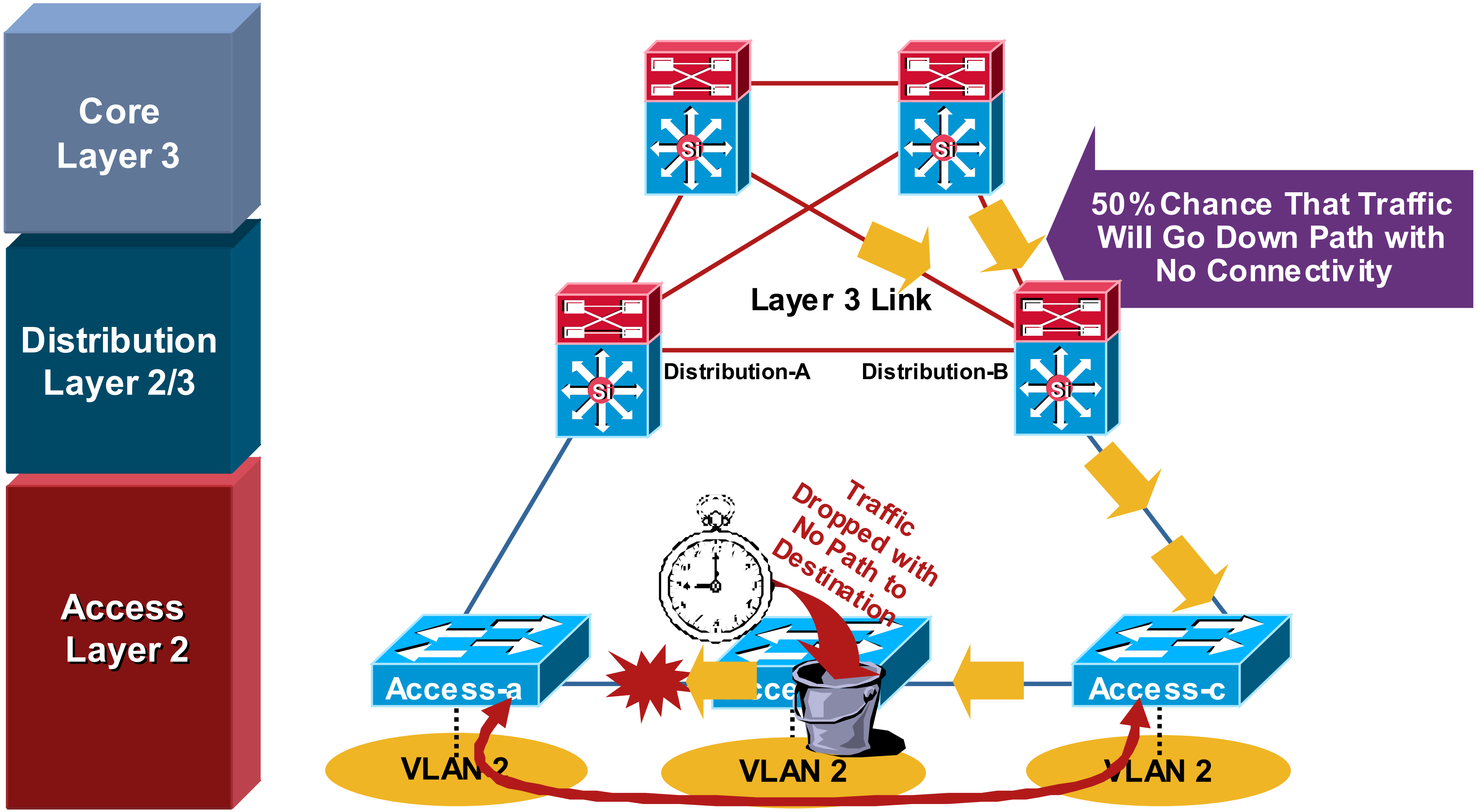
- Place the root where you want it
- The root bridge should stay where you put it
  - RootGuard
  - LoopGuard
  - UplinkFast
- Only end-station traffic should be seen on an edge port
  - BPDU Guard
  - RootGuard
  - PortFast



Daisy Chaining Access Layer Switches

Avoid Potential Black Holes

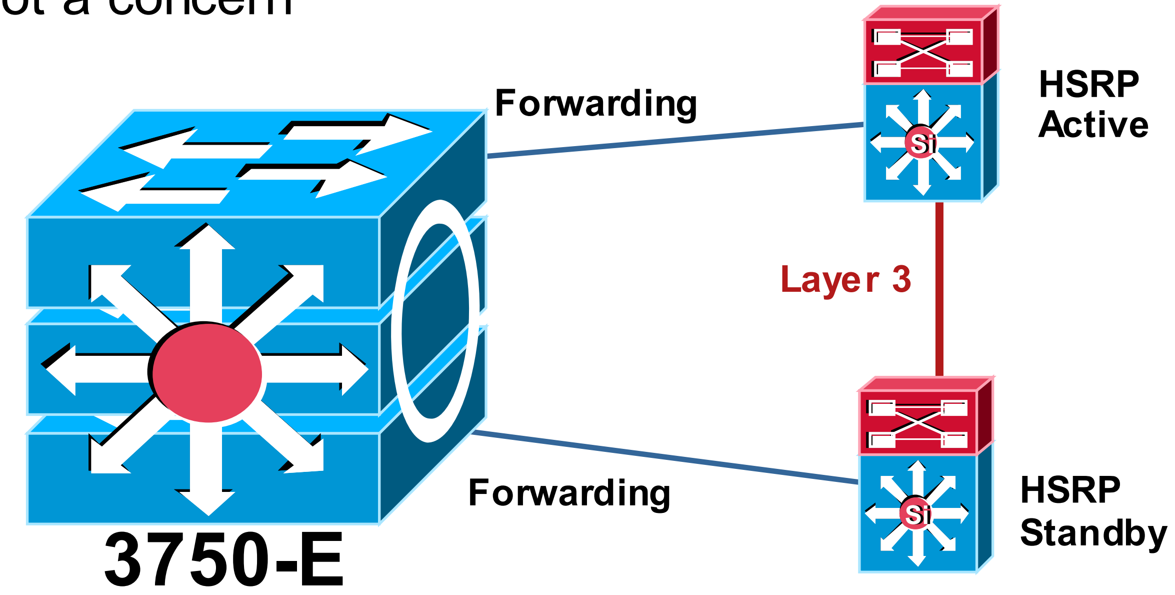
Return Path Traffic Has a 50/50 Chance of Being 'Black Holed'



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

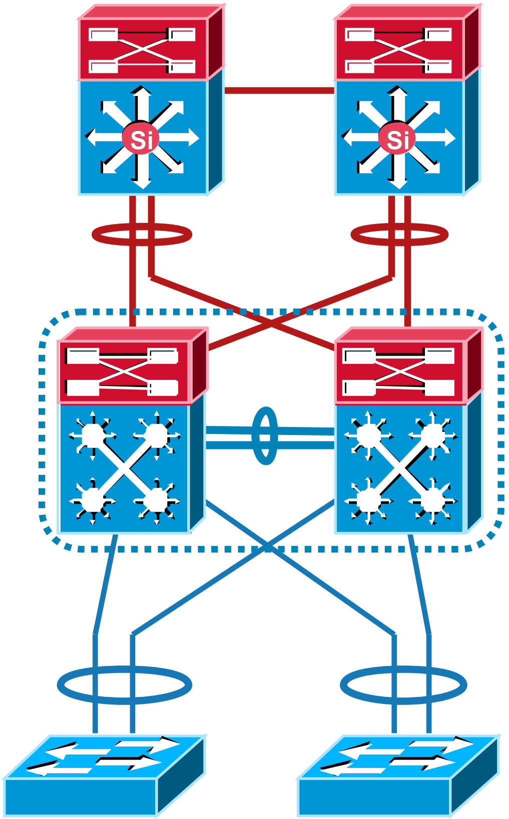


Intelligent Switching

Availability and Resiliency - VSS

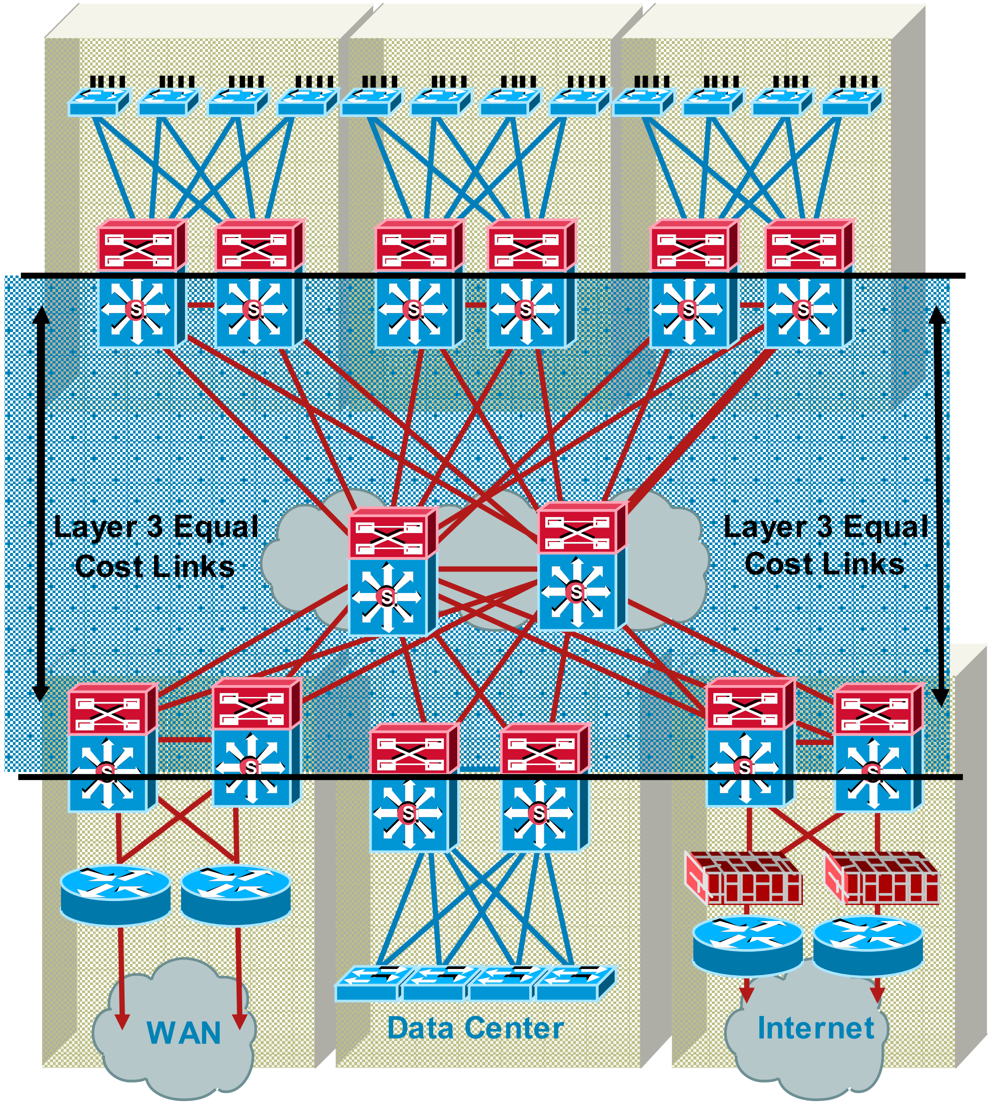
Virtual Switch: Physical redundancy with a single logical control plane

- Catalyst 6500 Virtual Switch System - VSS
  - Extension of control and management planes across chassis
  - Active-Active data plane
  - Stateful switchovers across chassis
  - Single point of configuration and management
- Multichassis Etherchannel (MEC) between Virtual Switch and all neighbors
  - Eliminates L2 loops
  - All links forwarding doubling effective bandwidth
  - Reduce number of L3 routing neighbors



Best Practices—
Layer 3 Routing Protocols

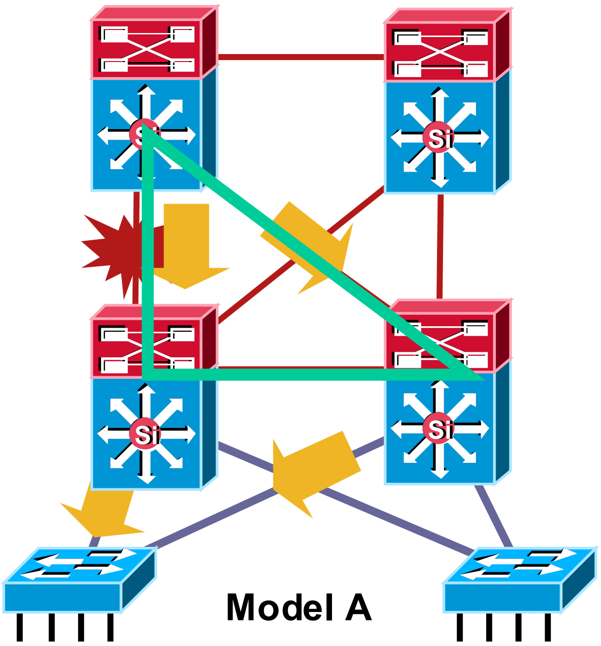
- Typically deployed in distribution to core, and core to core interconnections
- Used to quickly re-route around failed node/links while providing load balancing over redundant paths
- Build triangles not squares for deterministic convergence
- Only peer on links that you intend to use as transit
- Insure redundant L3 paths to avoid black holes
- Summarize distribution to core to limit EIGRP query diameter or OSPF LSA propagation



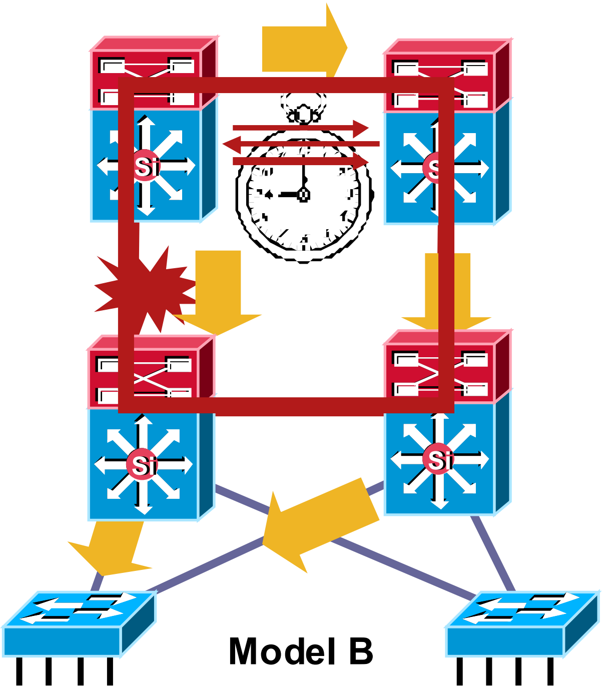
Best Practice— Build Triangles Not Squares

Deterministic vs. Non-Deterministic

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



Squares: Link/Box Failure Requires Routing Protocol Convergence



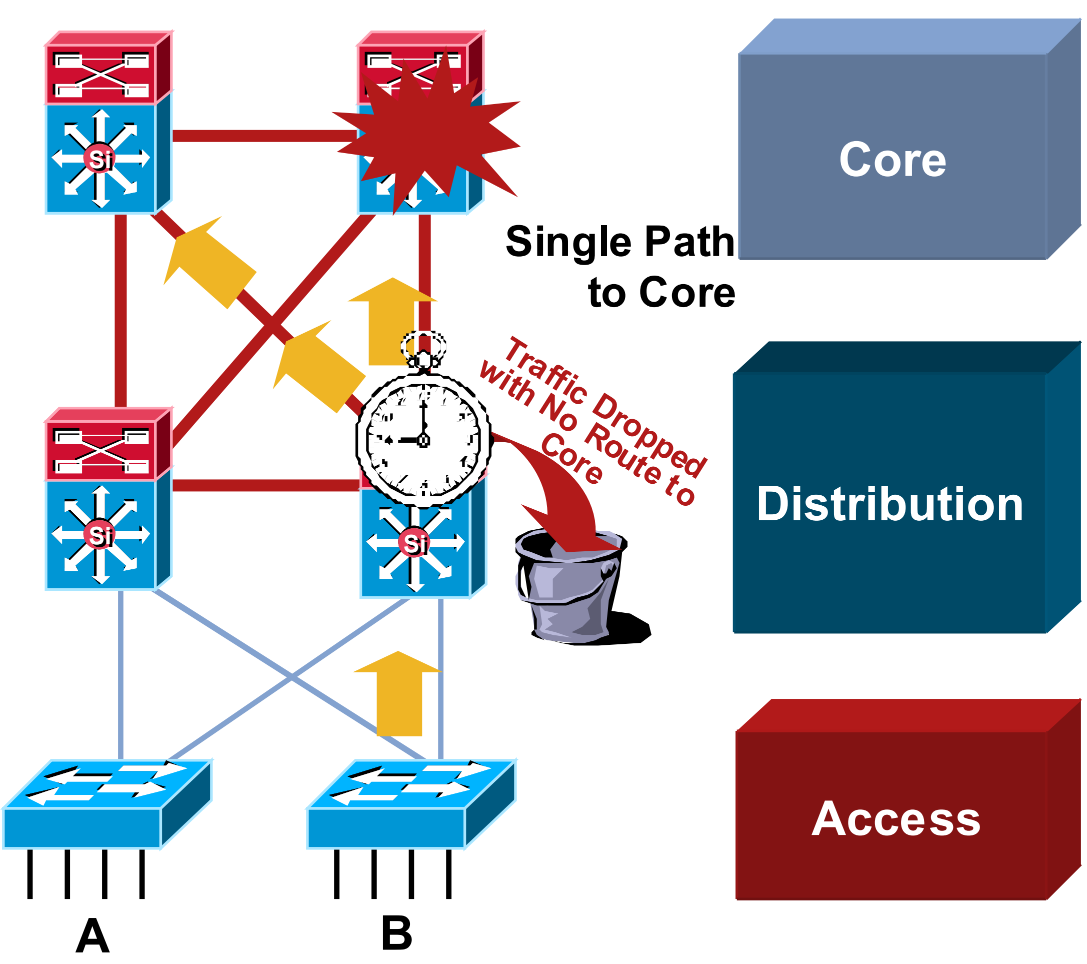
- Layer 3 redundant equal cost links support fast convergence
- Hardware based—fast recovery to remaining path
- Convergence is extremely fast (dual equal-cost paths: no need for OSPF or EIGRP to recalculate a new path)

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?

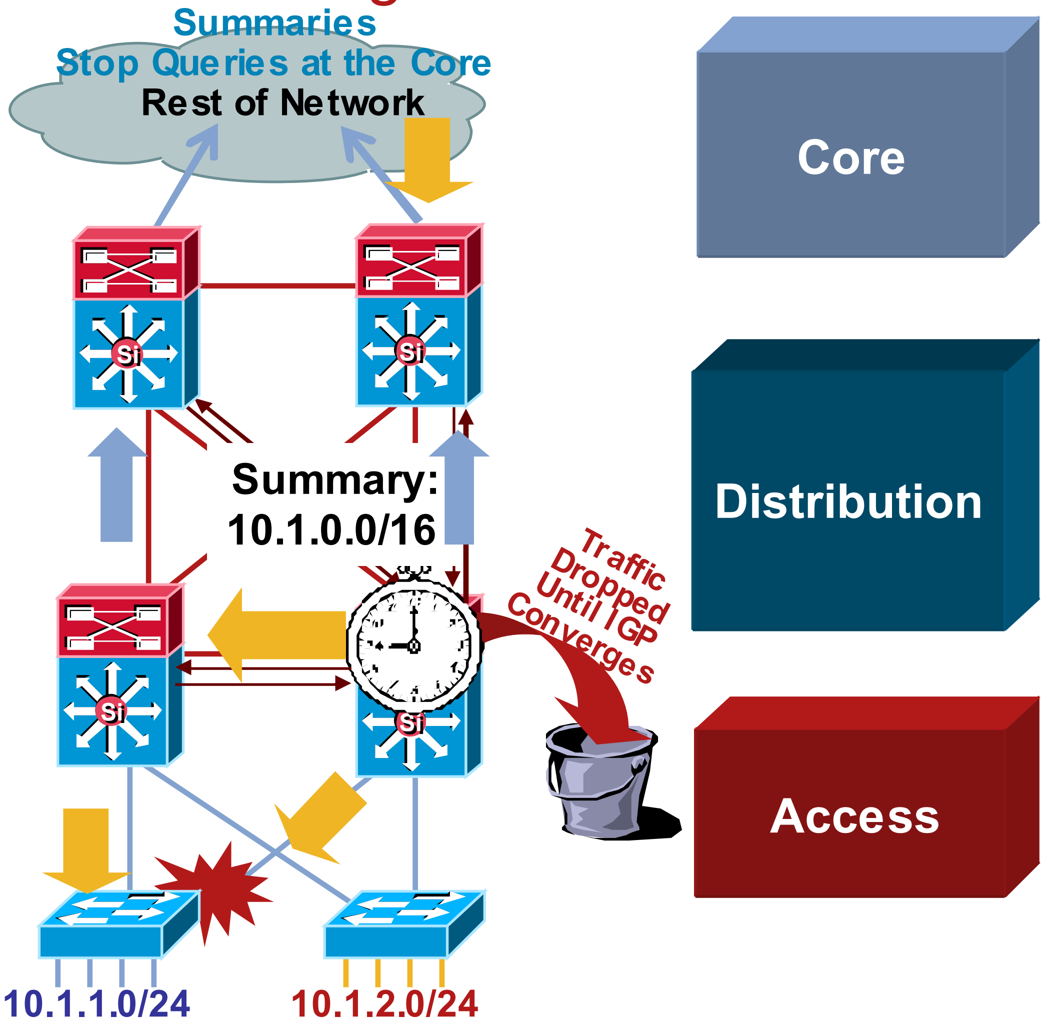
  How do you design for scalability if the access used for transit traffic?
- Install a redundant link to the core
- Best practice: install redundant link to core and utilize L3 link between distribution Layer (summarization—coming)



Why You Want to Summarize at the Distribution

Reduce the Complexity of IGP Convergence

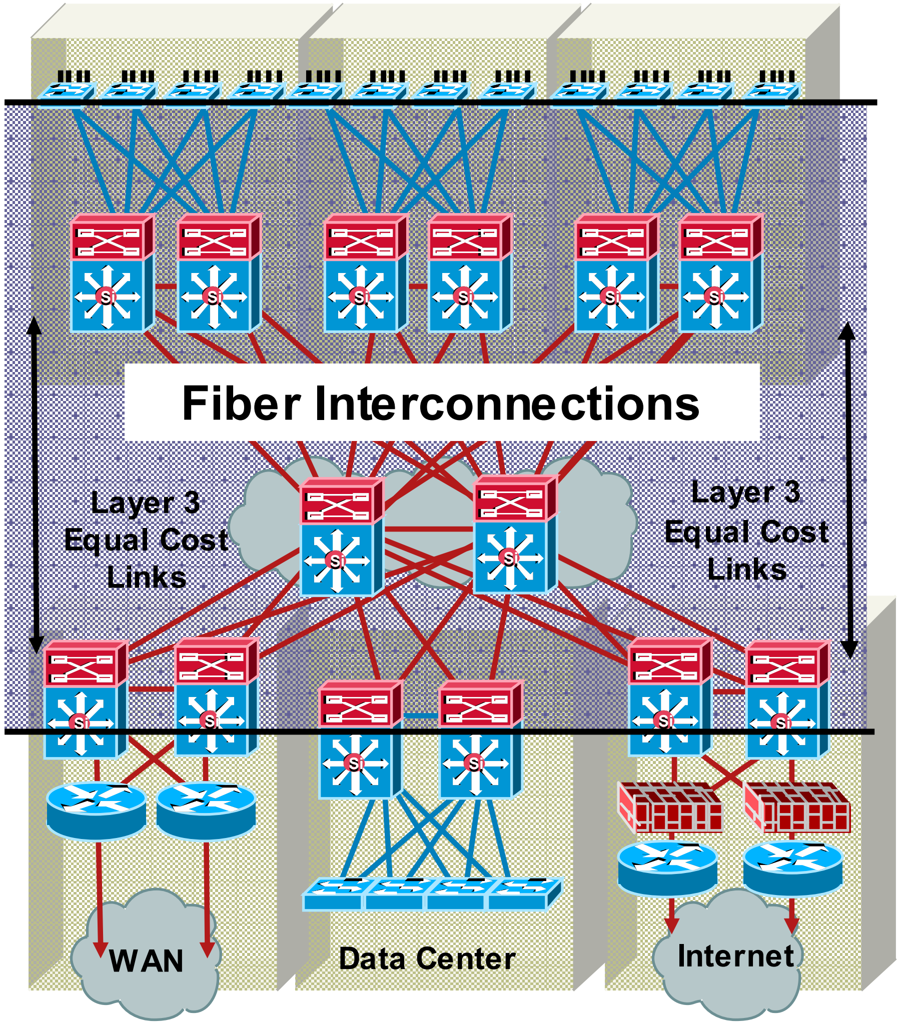
- It is important to force summarization at the distribution towards the core
- For return path traffic an OSPF or EIGRP re-route is required
- By limiting the number of peers an EIGRP router must query or the number of LSAs an OSPF peer must process we can optimize its re-route
- For EIGRP if we summarize 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 – Unidirectional Link Detection

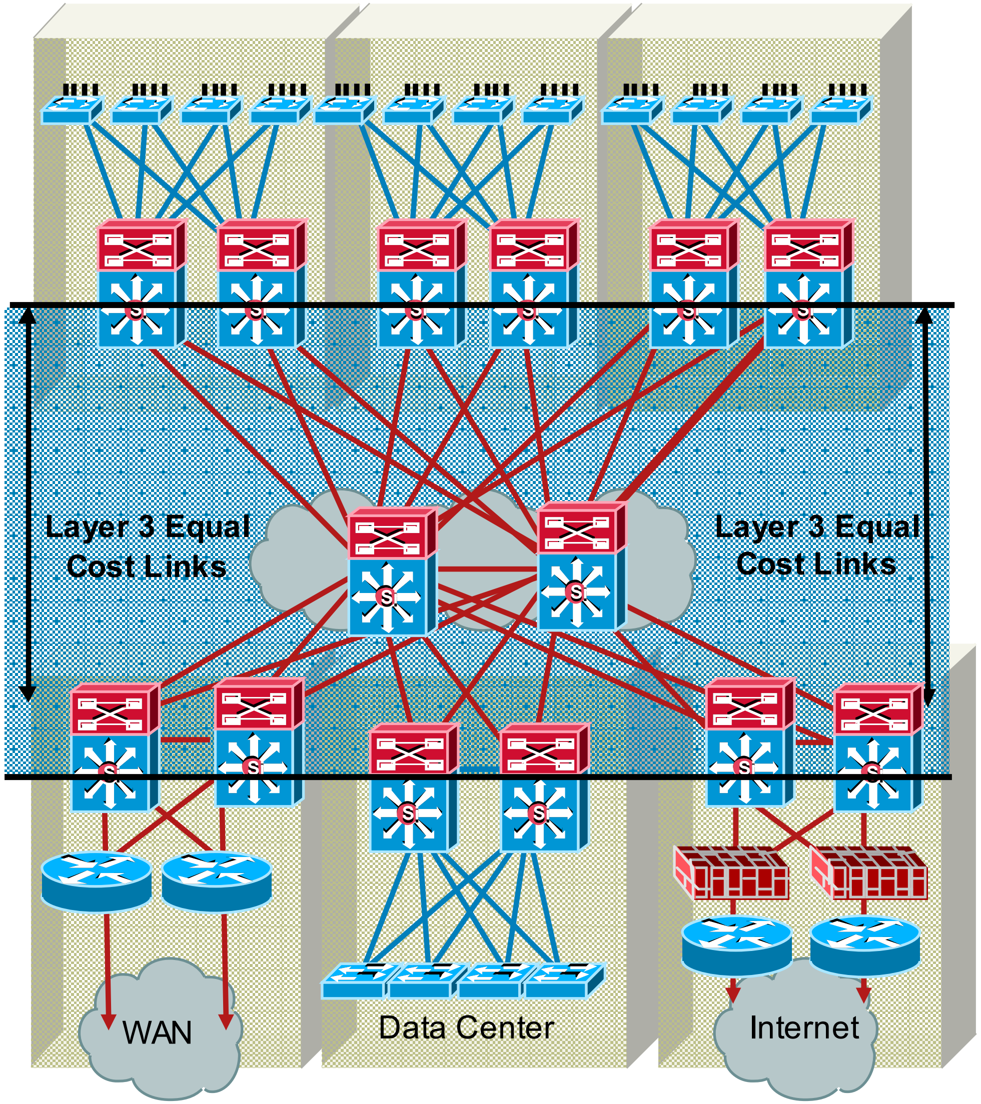
Protecting Against One Way Communication

- Typically deployed on any fiber optic interconnection
- Detects partially failed links and that could impact 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



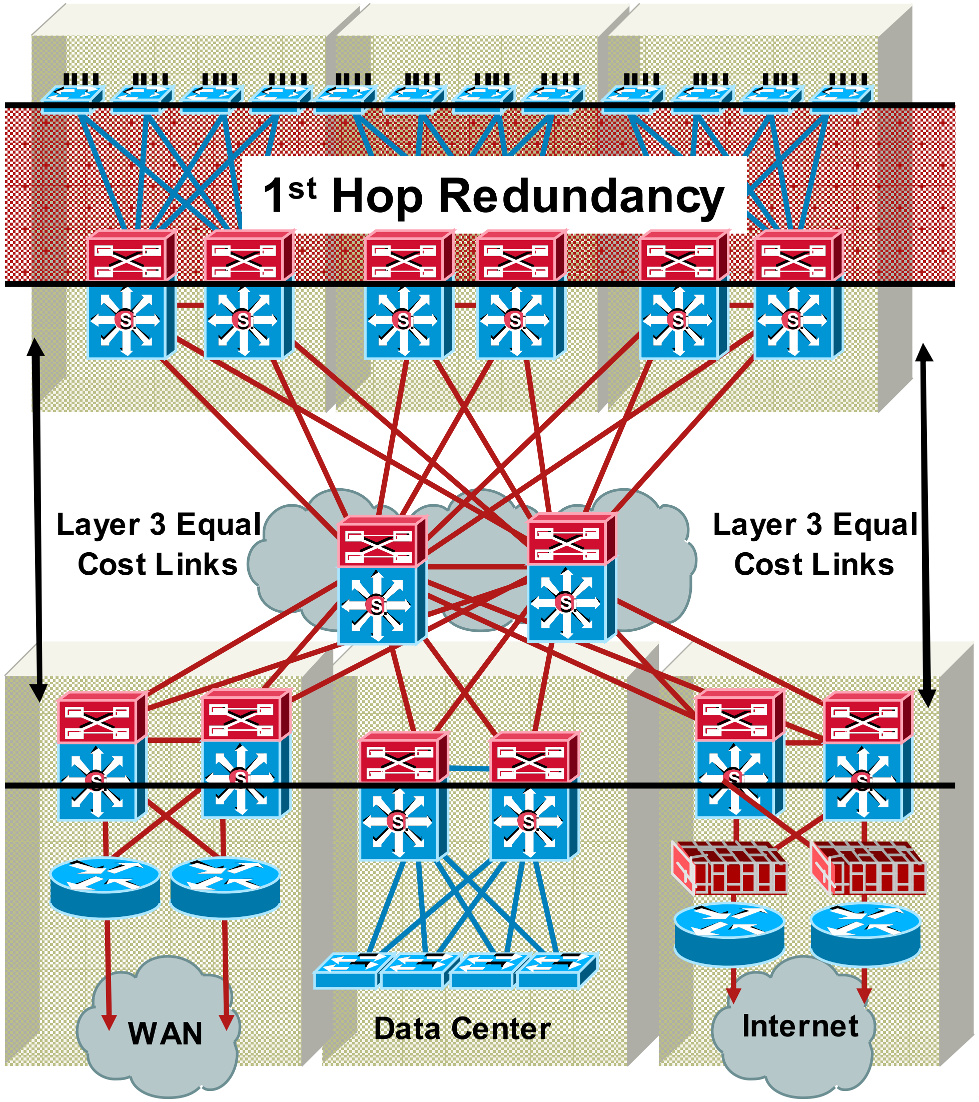
Best Practices—EtherChannel Configuration

- Typically deployed in distribution to core, and core to core interconnections
- Used to provide link redundancy—while reducing peering complexity
- Deploy in powers of 2 (2, 4, or 8)
- 802.3ad LACP for interoperability if you need it



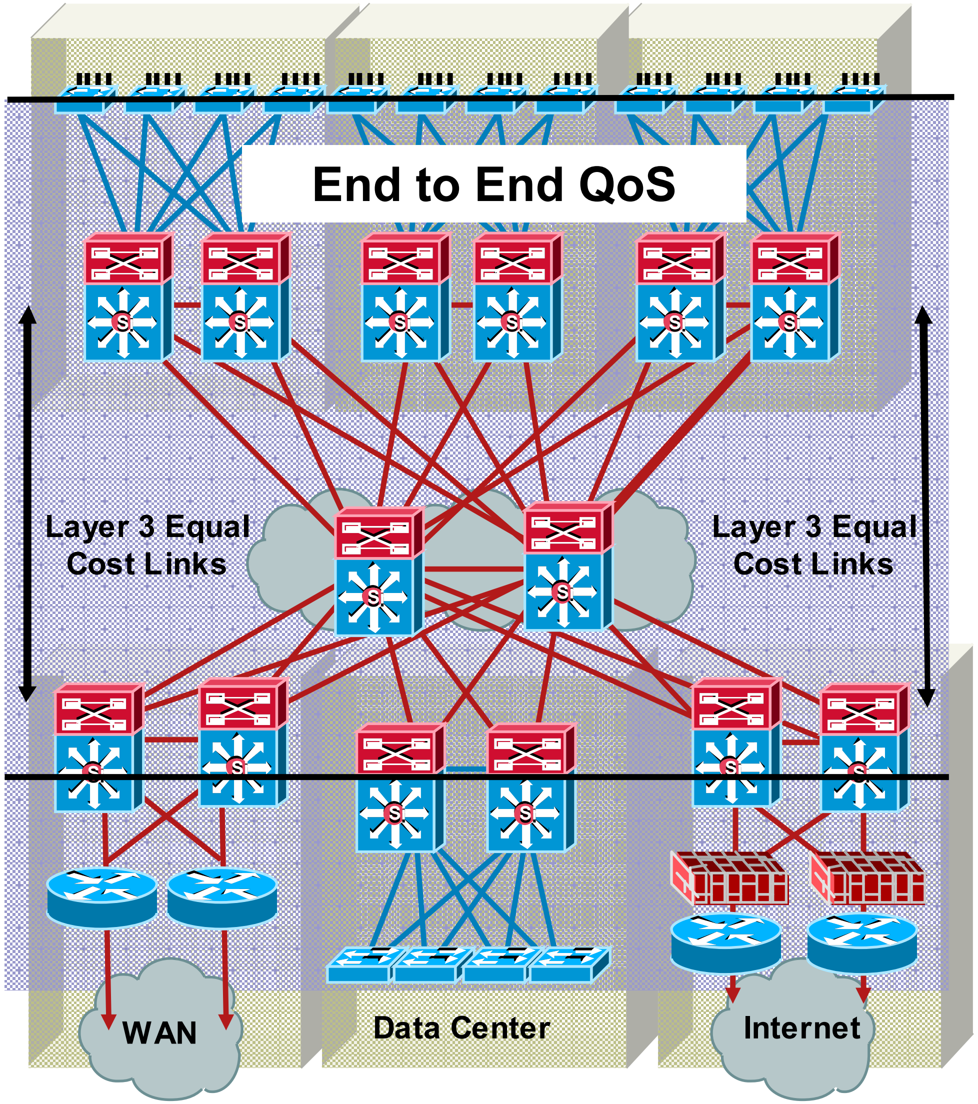
Best Practices—First Hop Redundancy

- 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



Best Practices - Quality of Service

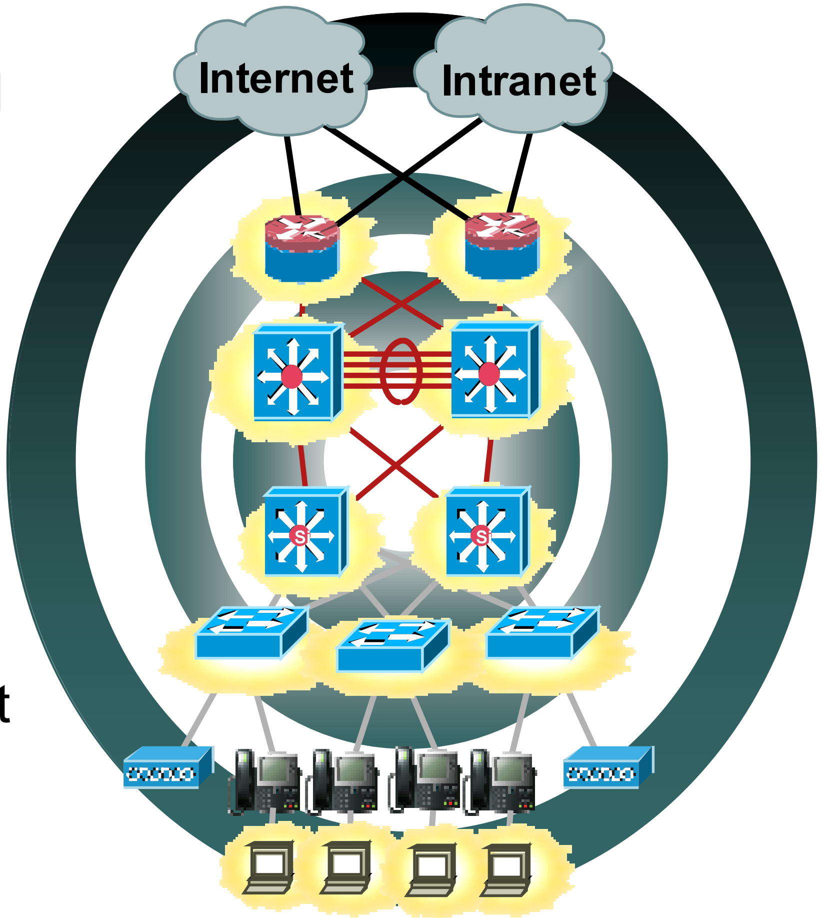
- Must be deployed end-to-end to be effective; all layers play different but equal roles
- Ensure that mission critical applications are not impacted by link or transmit queue congestion
- Aggregation and rate transition points must enforce QoS policies
- Multiple queues with configurable admission criteria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



Intelligent Switching

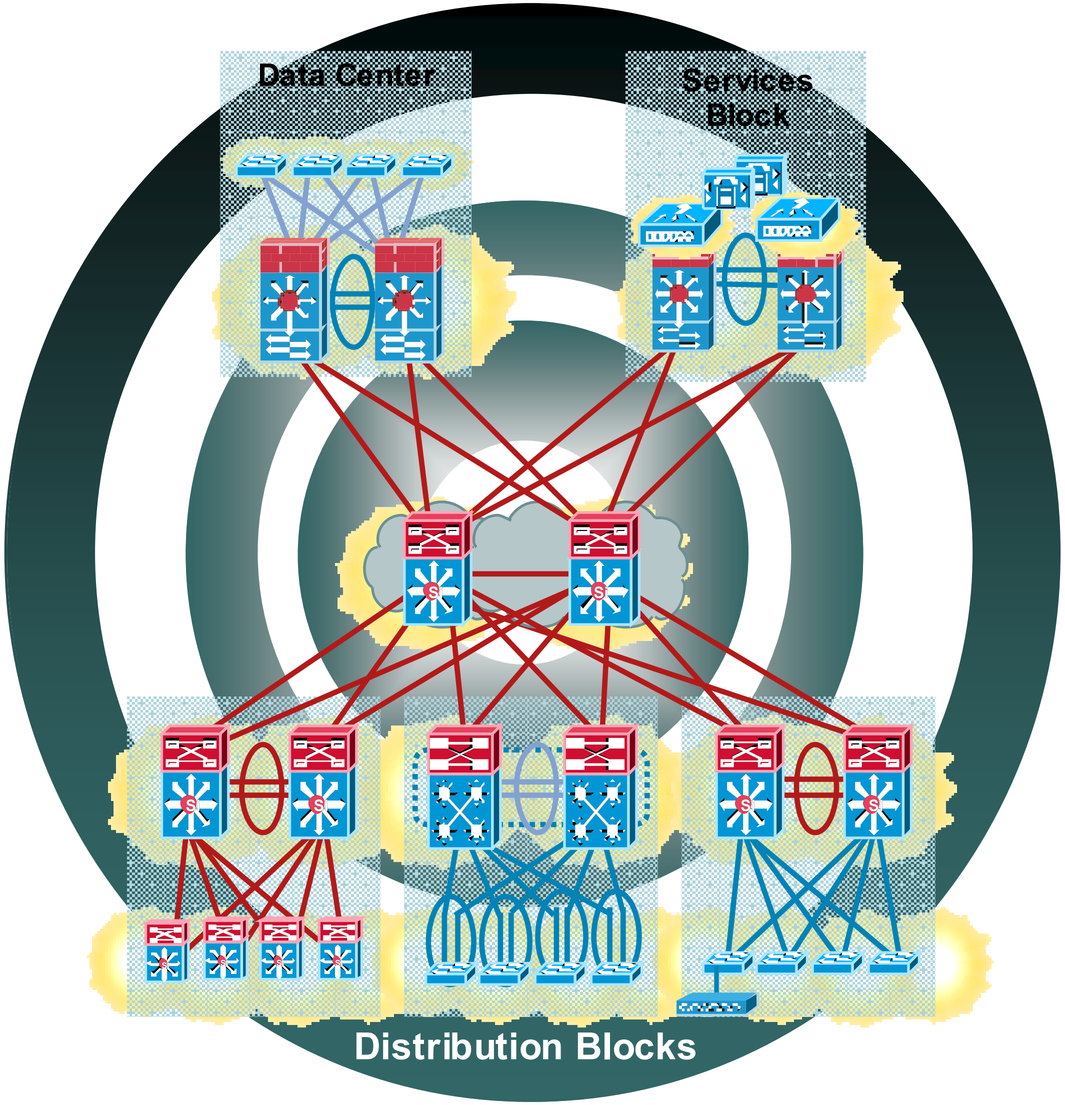
Optimized Delivery

- Congestion Avoidance mechanisms preserve forwarding continuance and bandwidth availability
- Policing and rate-limiting provide bandwidth limit enforcement for variable traffic types, and action due for violating assigned thresholds.
- Traffic Classification allows to distinguish (and therefore aids in prioritizing) one kind of traffic from another by examining variable packet fields
- Traffic Shaping provides control of outgoing traffic rate to ensure it conforms to allowed thresholds



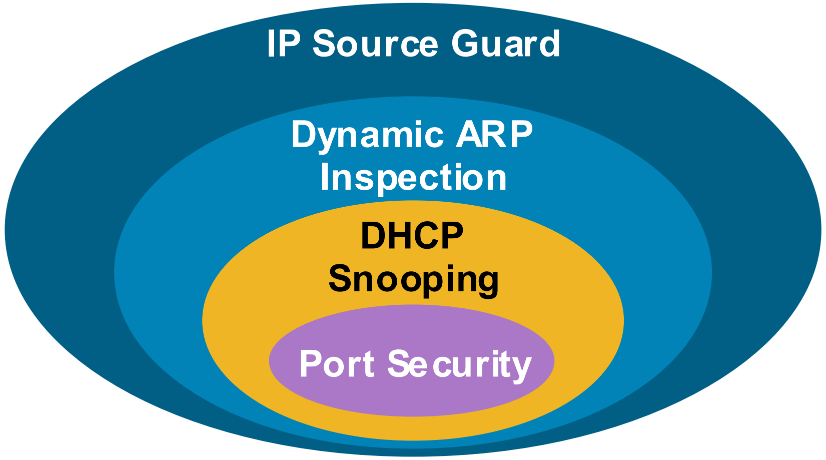
Agenda

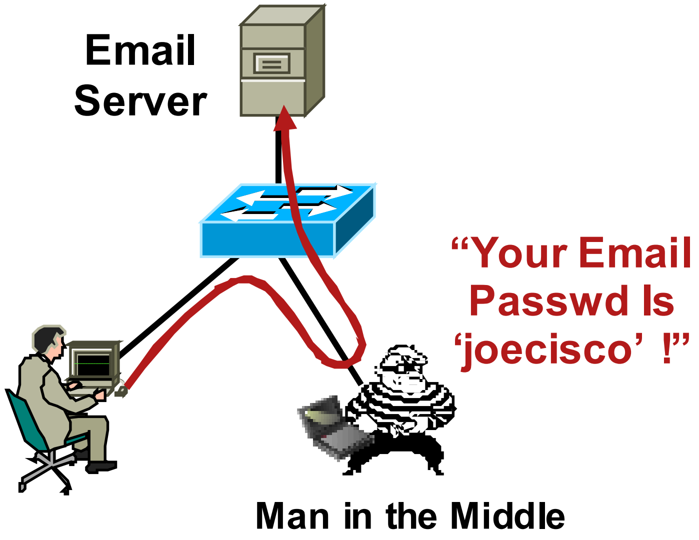
- Multilayer Campus Design principles
- Campus Design Best Practices
- Hardening the Campus Network Design
- Summary



Hardening the Edge

Catalyst Integrated Security Features



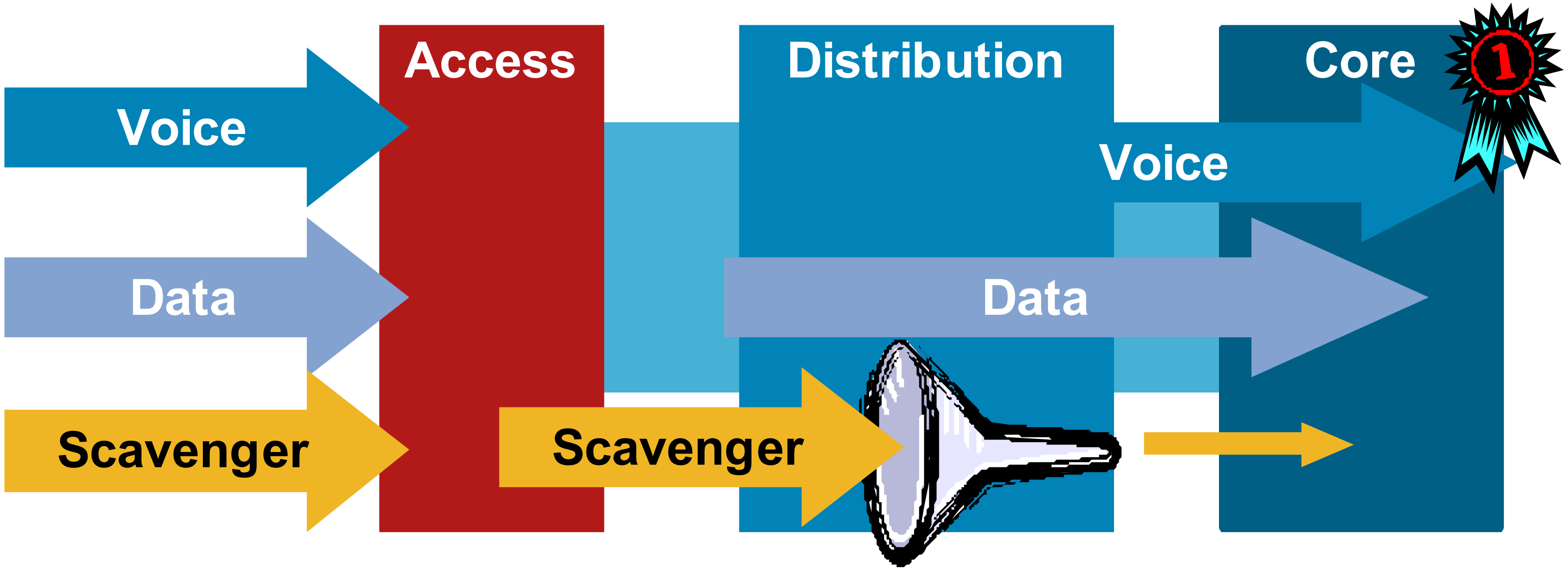


- Port security prevents CAM attacks and DHCP Starvation attacks
- DHCP Snooping prevents Rogue DHCP Server attacks
- Dynamic ARP Inspection prevents current ARP attacks
- IP Source Guard prevents IP/MAC Spoofing

Harden the Network Links

Protect the Good and Punish the Bad

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



Hardening the Switches

Control Plane Protection

- CEF protects against system overload due to flow flooding
- System CPU still has to be able to process certain traffic

  BPDUs, CDP, EIGRP, OSPF

  Telnet, SSH, SNMP, ARP, ICMP, IGMP
- System needs to provide throttling on CPU-bound traffic

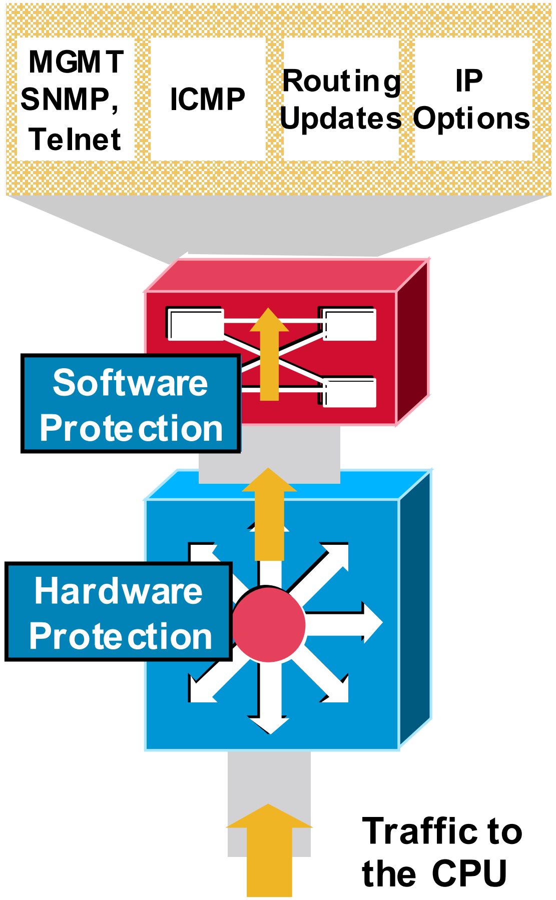
  IOS Based SW Rate Limiters

  Multiple CPU queues on 4500 & 3750

  Hardware Rate Limiters on 6500

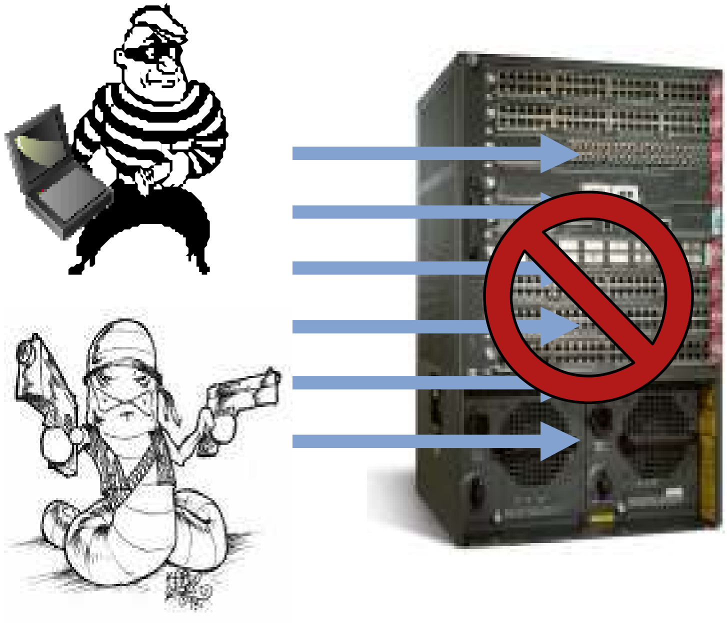
  Hardware Control Plane Policing (CoPP) on 6500 & 4500

  Second tier software Control Plane Policing on 6500



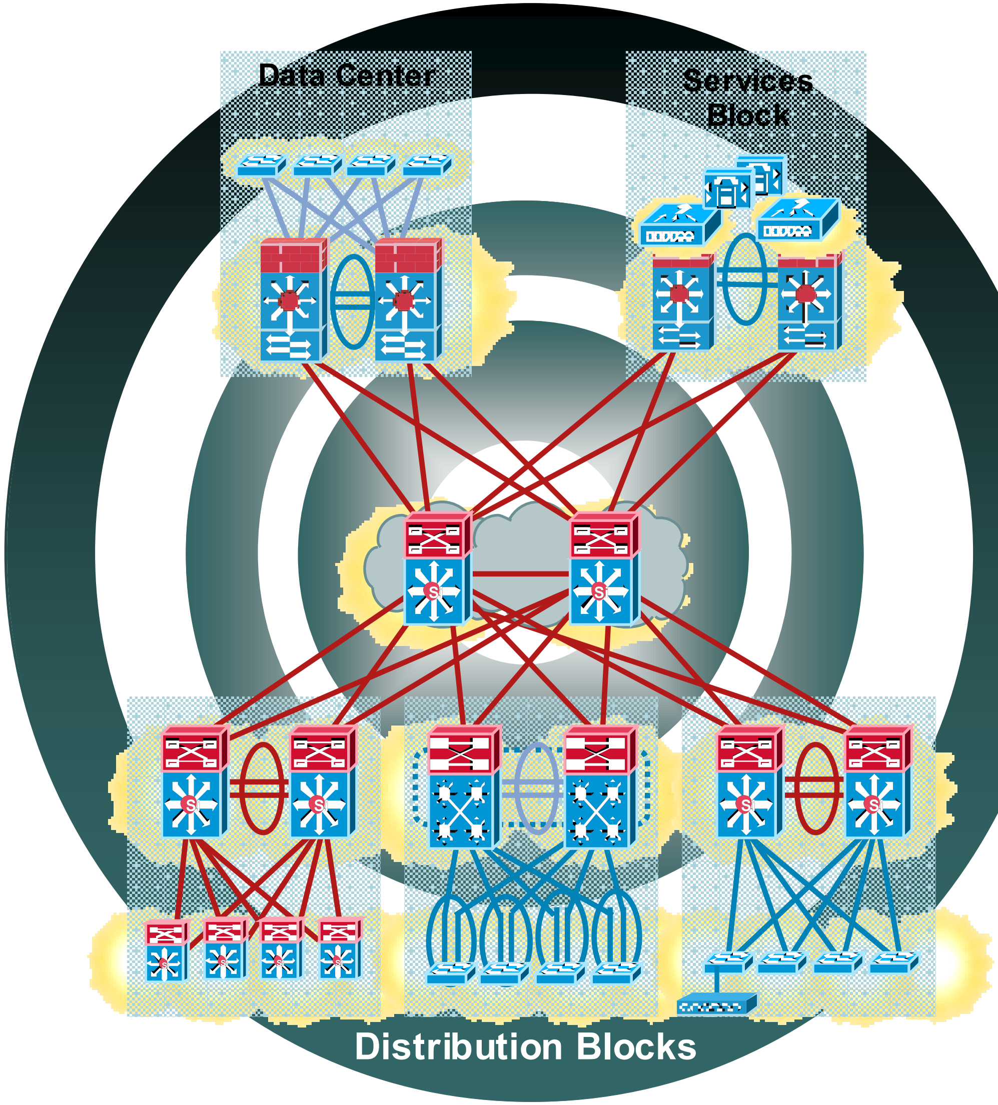
Intelligent Switching
Integrated Security

- Control Plane Policing applies Hardware QoS policies to traffic punted to the CPU to preserve CPU survivability.
- Network Admission Control integration provides Policy Compliance and Remediation
- Broadcast and multicast Storm control prevents service disruption caused by errors in protocol-stack implementation or network configuration that result in flooding, creating excessive traffic and degrading network performance.



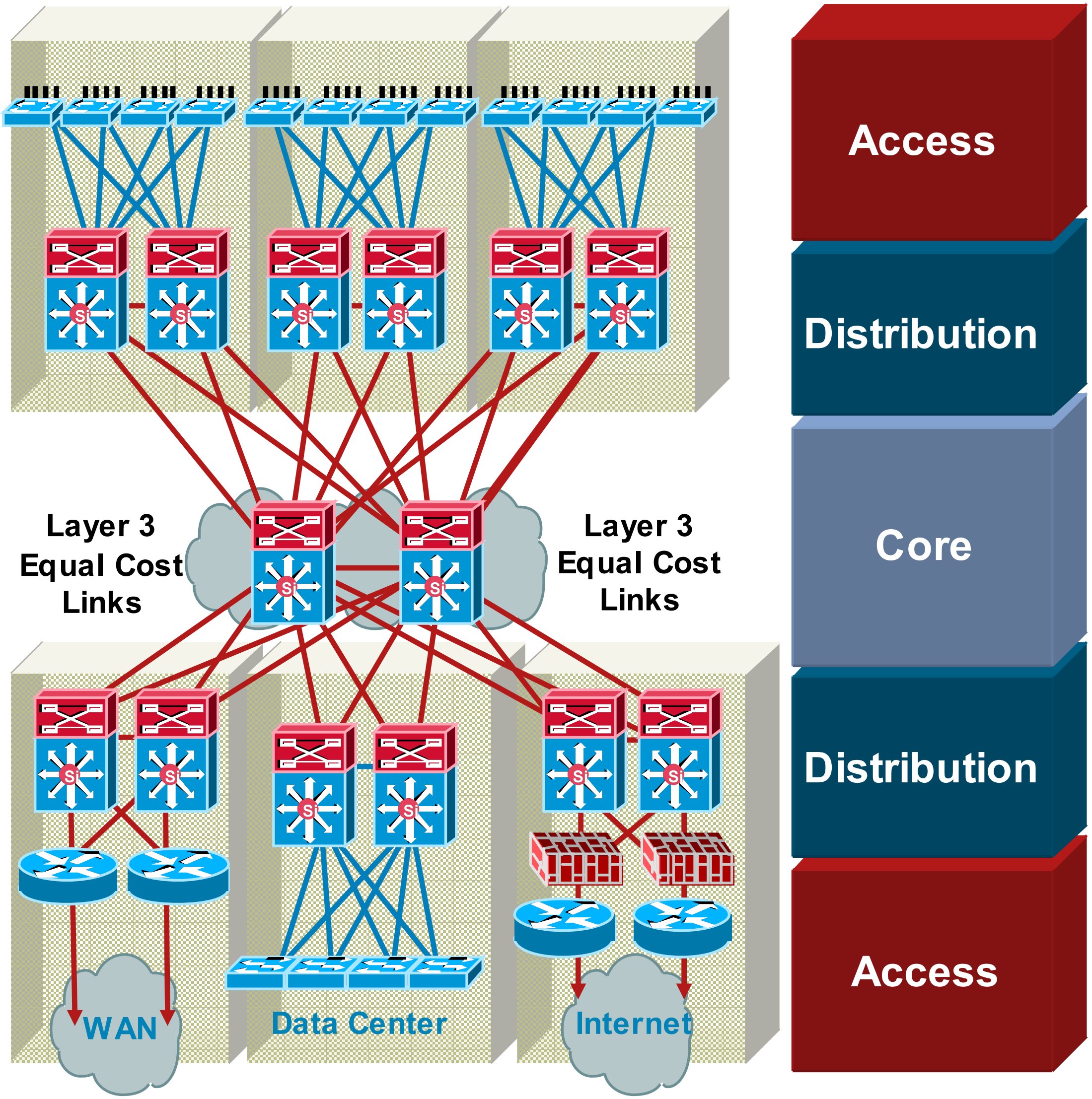
Agenda

- Multilayer Campus Design principles
- Campus Design Best Practices
- Hardening the Campus Network Design
- 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


Layer 3
Equal Cost
Links
Layer 3
Equal Cost
Links
WAN
Data Center
Internet
Access
Distribution
Core
Distribution
Access