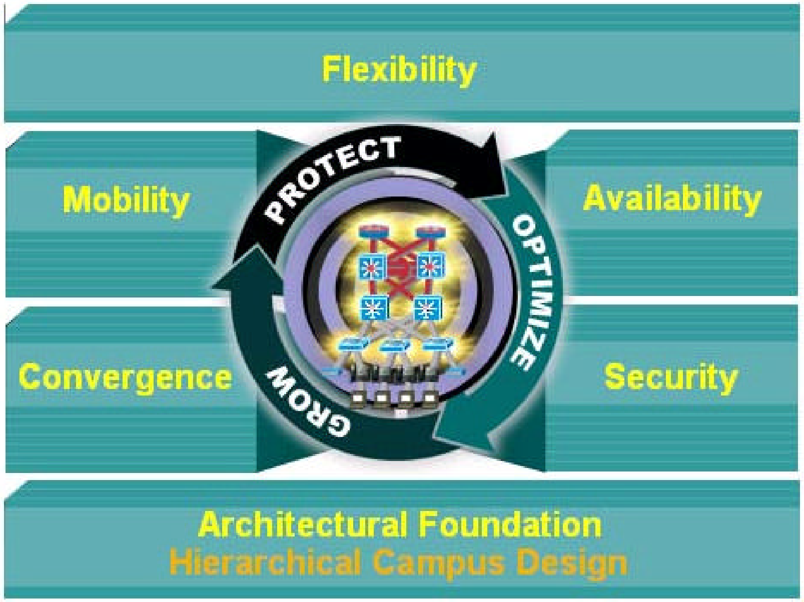
DEPLOYING A FULLY ROUTED ENTERPRISE CAMPUS NETWORK

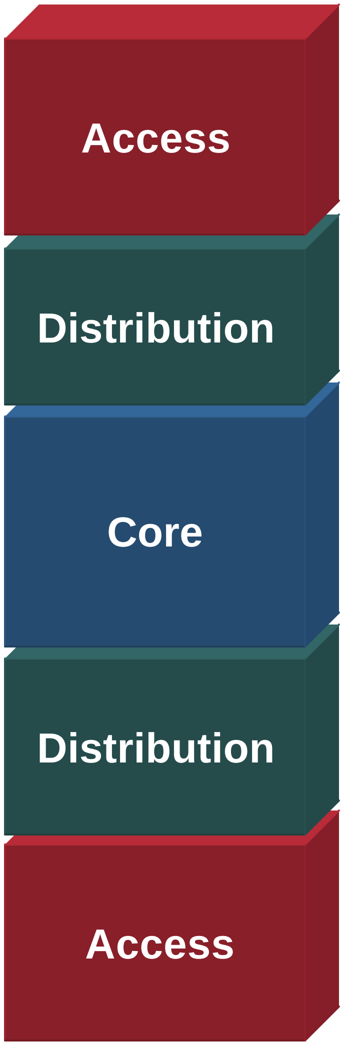
SESSION RST-2031

Agenda

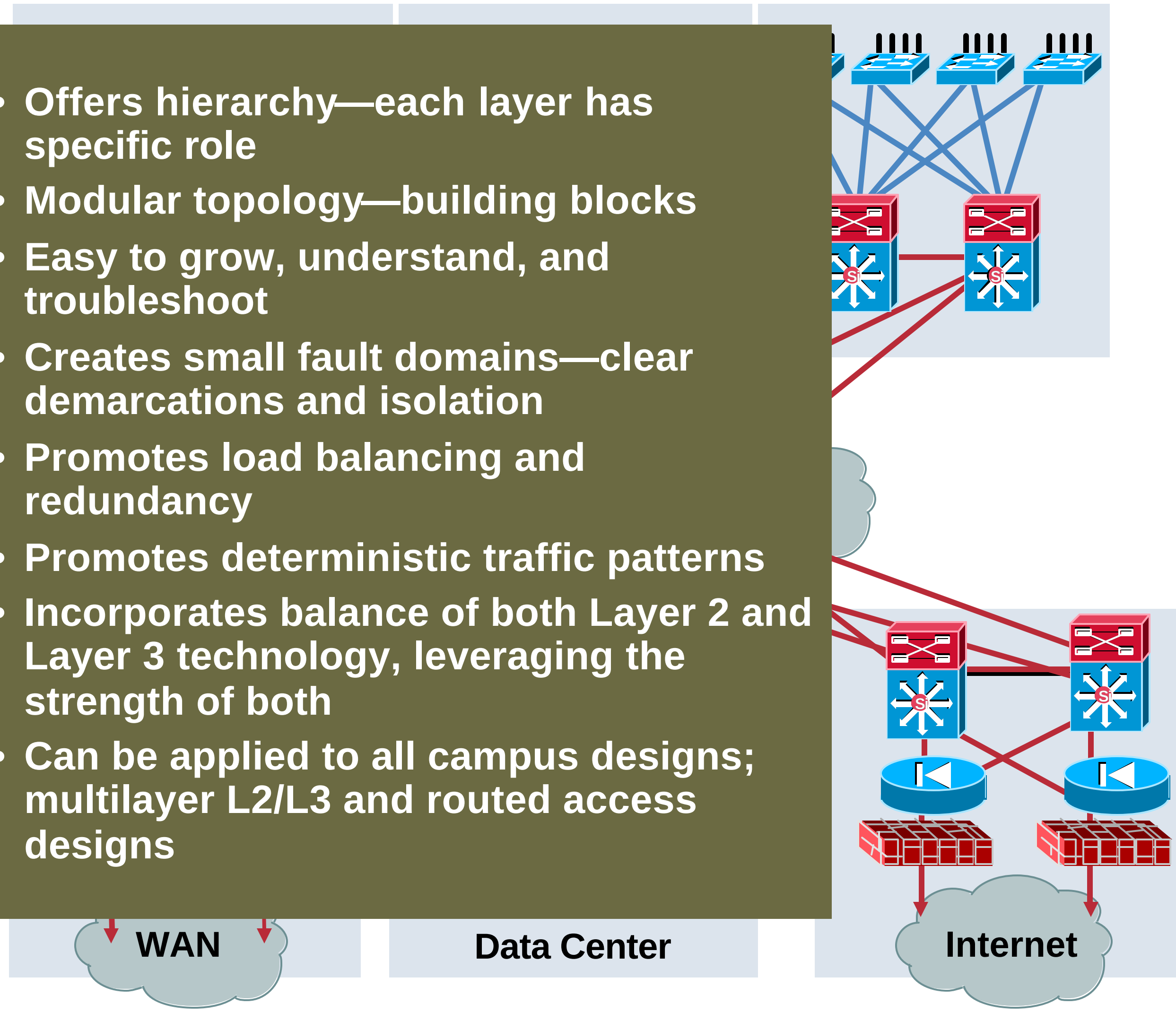
- Campus Network Designs
- Routed Access Design
- EIGRP Design Details
- OSPF Design Details
- PIM Design Details
- Summary



Hierarchical Campus Design
Building Blocks

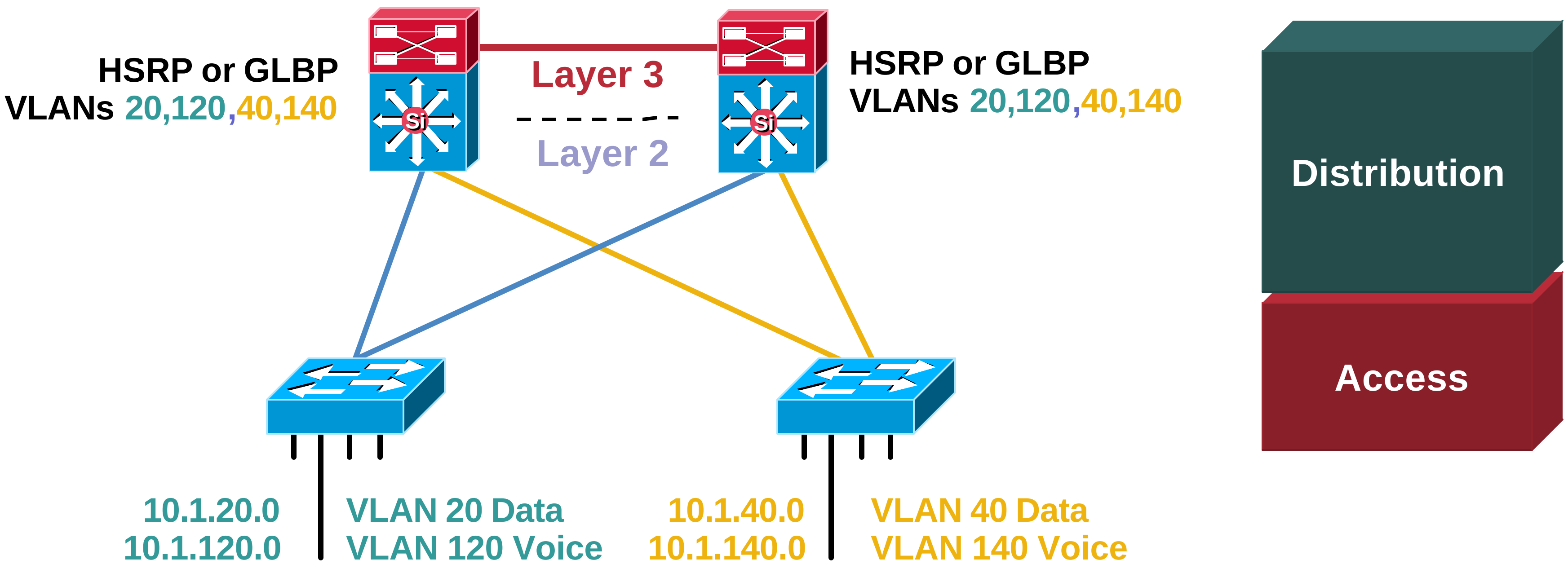


- Offers hierarchy—each layer has specific role
- Modular topology—building blocks
- Easy to grow, understand, and troubleshoot
- Creates small fault domains—clear demarcations and isolation
- Promotes load balancing and redundancy
- Promotes deterministic traffic patterns
- Incorporates balance of both Layer 2 and Layer 3 technology, leveraging the strength of both
- Can be applied to all campus designs; multilayer L2/L3 and routed access designs



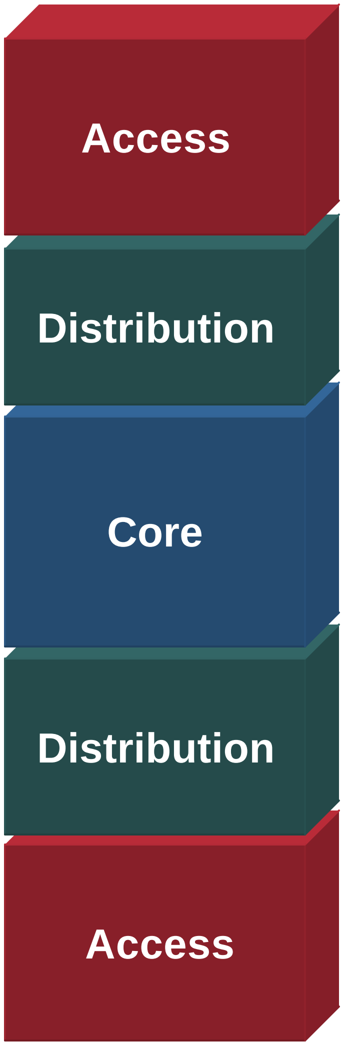
Tried and True: Reference Design

Multilayer L2/L3 Design



- **Consider fully utilizing uplinks via GLBP**
- **Distribution-to-distribution link required for route summarization**
- **No STP convergence required for uplink failure/recovery**
- **Map L2 VLAN number to L3 subnet for ease of use/management**
- **Can easily extend VLANs across access layer switches if required**

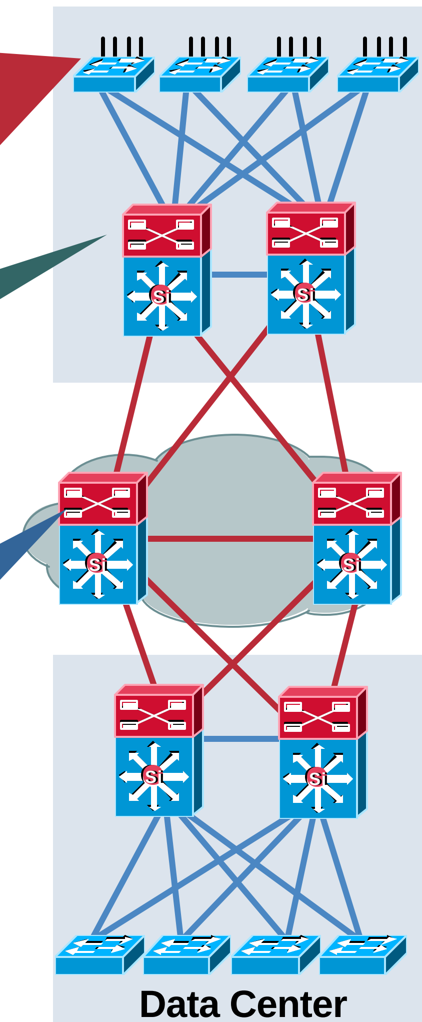
Hierarchical Campus Design
Multilayer L2/L3 Building Blocks



- Network trust boundary
- Use Rapid PVST+ on L2 ports to prevent loops in the topology
- Use UDLD to protect against 1 way interface UP connections
- Avoid daisy chaining access switches
- Avoid asymmetric routing and unicast flooding, don't span VLANS across the access layer

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

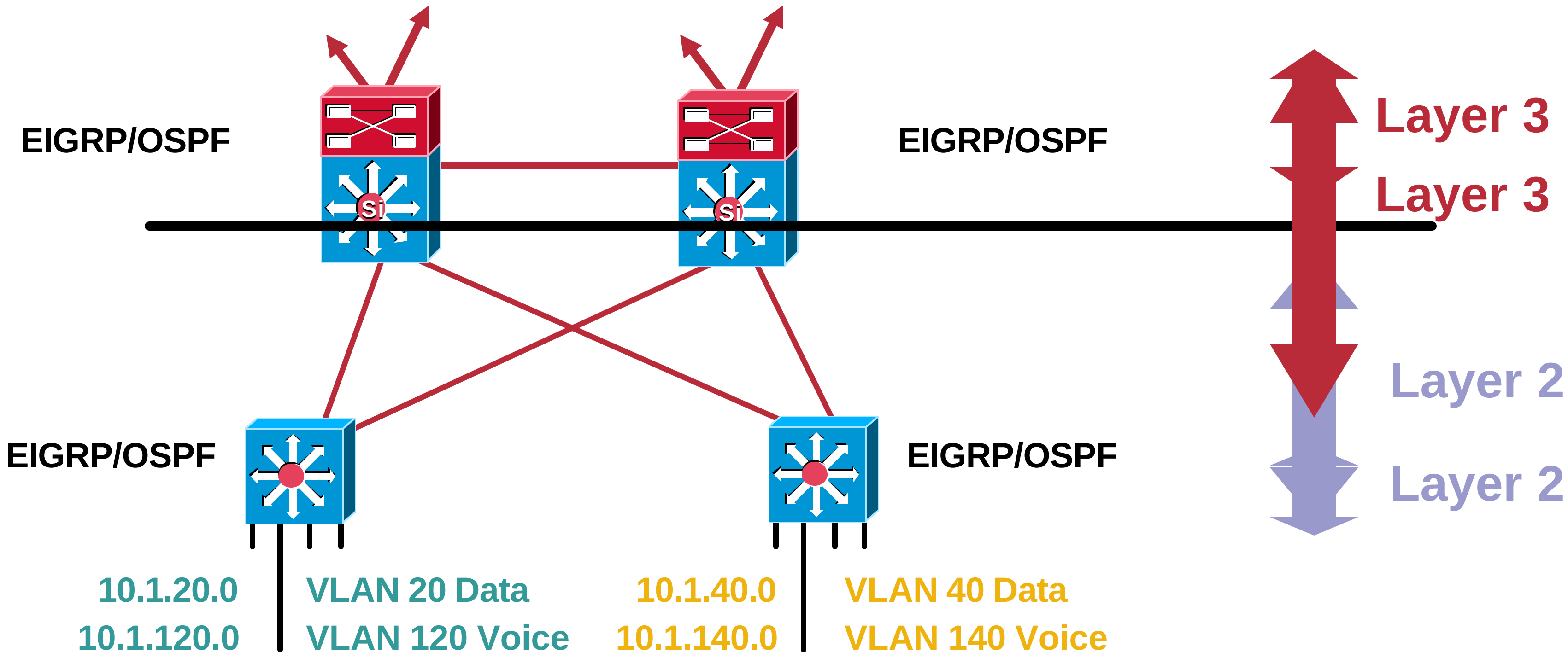
- Highly available and fast—always on
- Deploy QoS end-to-end: protect the good and punish the bad
- Equal cost core links provide for best convergence
- Optimize CEF for best utilization of redundant L3 paths



Routing to the Edge

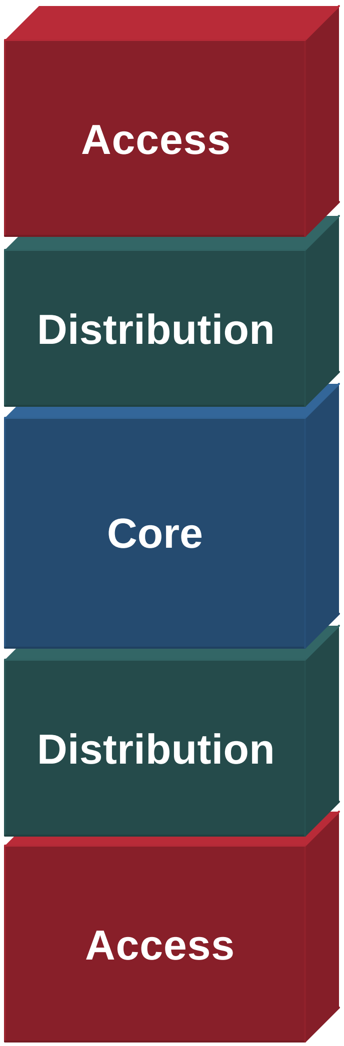
Layer 3 Distribution with Layer 3 Access

Cisco.com



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

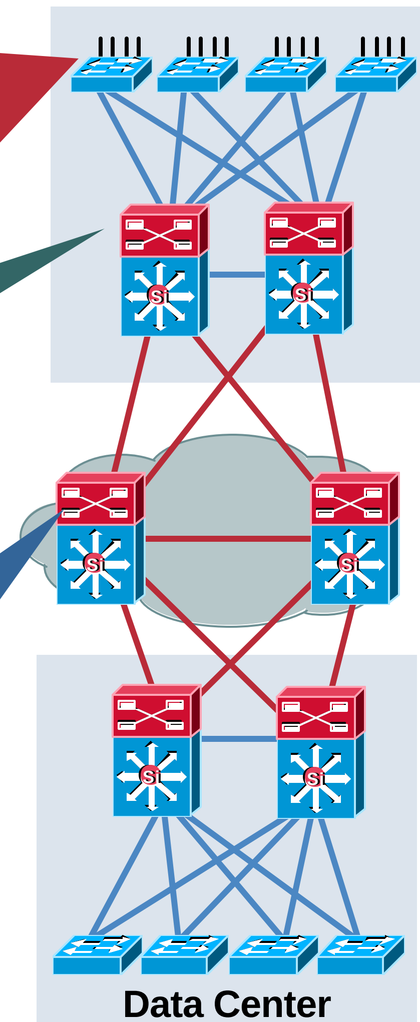
Hierarchical Campus Design
Routed Access Building Blocks



- Network trust boundary
- VLANs are contained to the access switch
- Use EIGRP or OSPF on interfaces to distribution layer
- Use parallel paths for Equal Cost Multi Path (ECMP) routing
- Use EIGRP stub routers or OSPF stub areas to limit scope of convergence events

- Access layer aggregation
- Route summarization to the core to minimize routing events
- Route filtering from the core to minimize routing table size in access
- OSPF stub area border (ABR)
- Keep your redundancy simple; equal cost load balancing between access and core
- Vary CEF algorithm to prevent polarization

- Highly available and fast—always on
- Deploy QoS end-to-end: protect the good and punish the bad
- Equal cost core links provide for best convergence



What Is High Availability?

Availability	DPM	Downtime Per Year (24x365)		
99.000%	10000	3 Days	15 Hours	36 Minutes
99.500%	5000	1 Day	19 Hours	48 Minutes
99.900%	1000		8 Hours	46 Minutes
99.950%	500		4 Hours	23 Minutes
99.990%	100			53 Minutes
99.999%	10			5 Minutes
99.9999%	1			30 Seconds

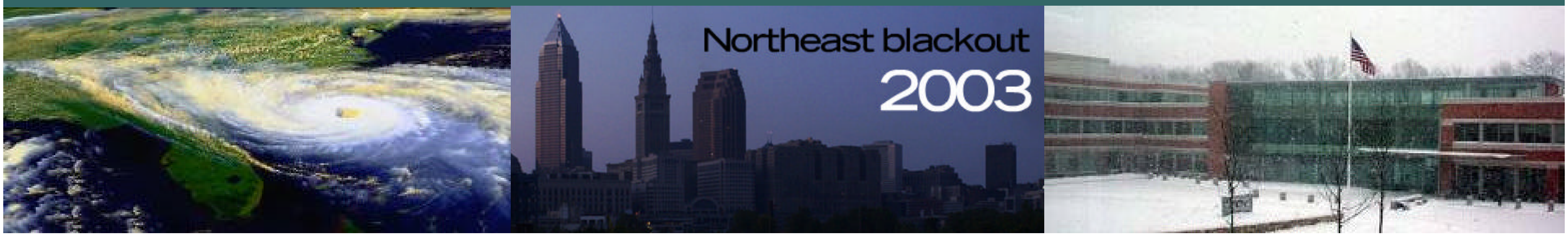
“High Availability” (99.999% and 99.9999%)

DPM—Defects per Million

What If You Could…

Reduce Cost Through Diminished Risk of Downtime

Cisco.com

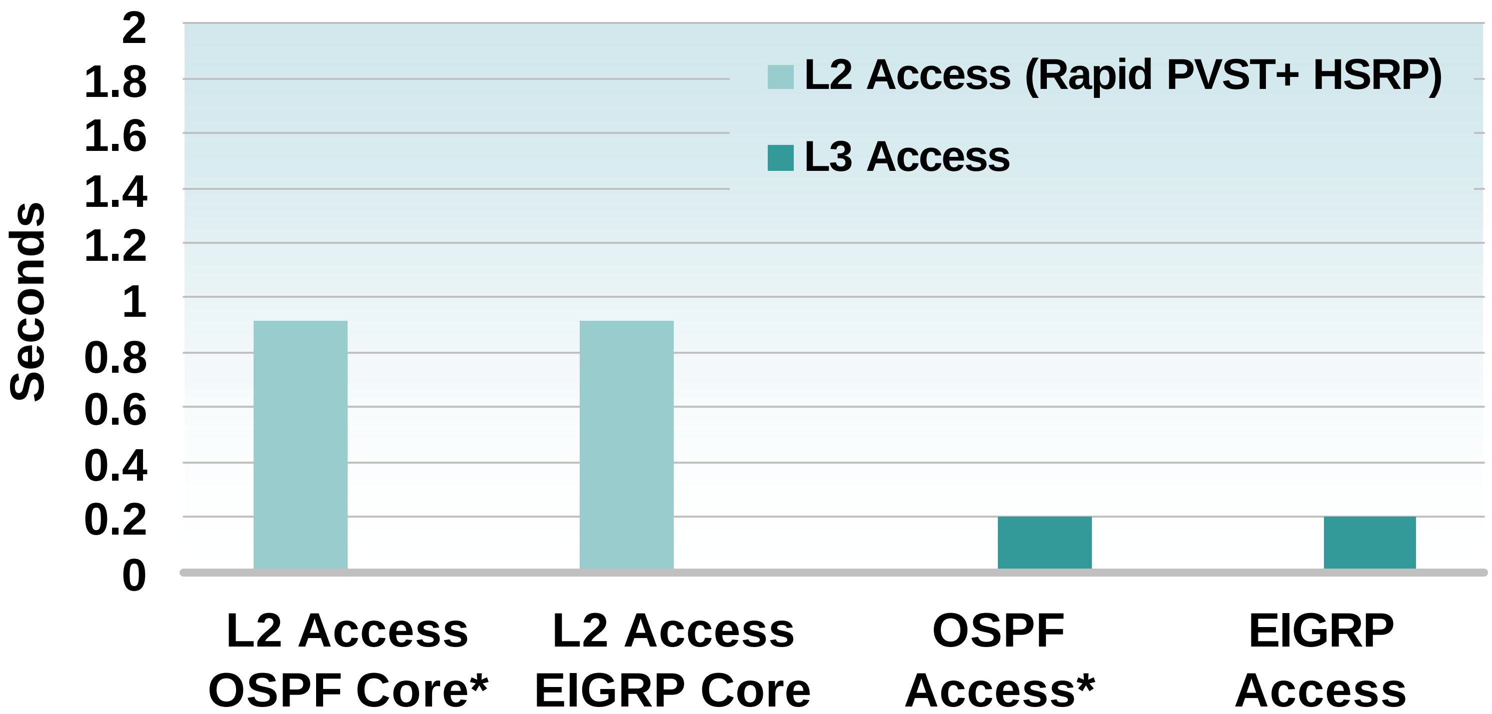


- **Costs for downtime are high**
  - One day cost of lost productivity = $1,644 per employee
    - 100 person office = $164K per day
- **More than just a data network outage**
- **More than just revenue impacted**
  - Revenue loss
  - Productivity loss
  - Impaired financial performance
  - Damaged reputation
  - Recovery expenses

Industry Sector	Revenue/Hour	Revenue/ Employee-Hour
Energy	$2,817,846	$ 569
Telecommunications	$2,066,245	$ 186
Manufacturing	$1,610,654	$ 134
Financial Institution	$1,495,134	$1,079
Insurance	$1,202,444	$ 370
Retail	$1,107,274	$ 244
Transportation	$ 668,586	$ 107
Average	$1,010,536	$ 205

Source: Meta Group

Campus High Availability
Sub-Second Convergence



Worst Case Convergence for Any Campus Failure Even

***OSPF Results Require Sub-Second Timers**

High-Availability Networking in the Campus

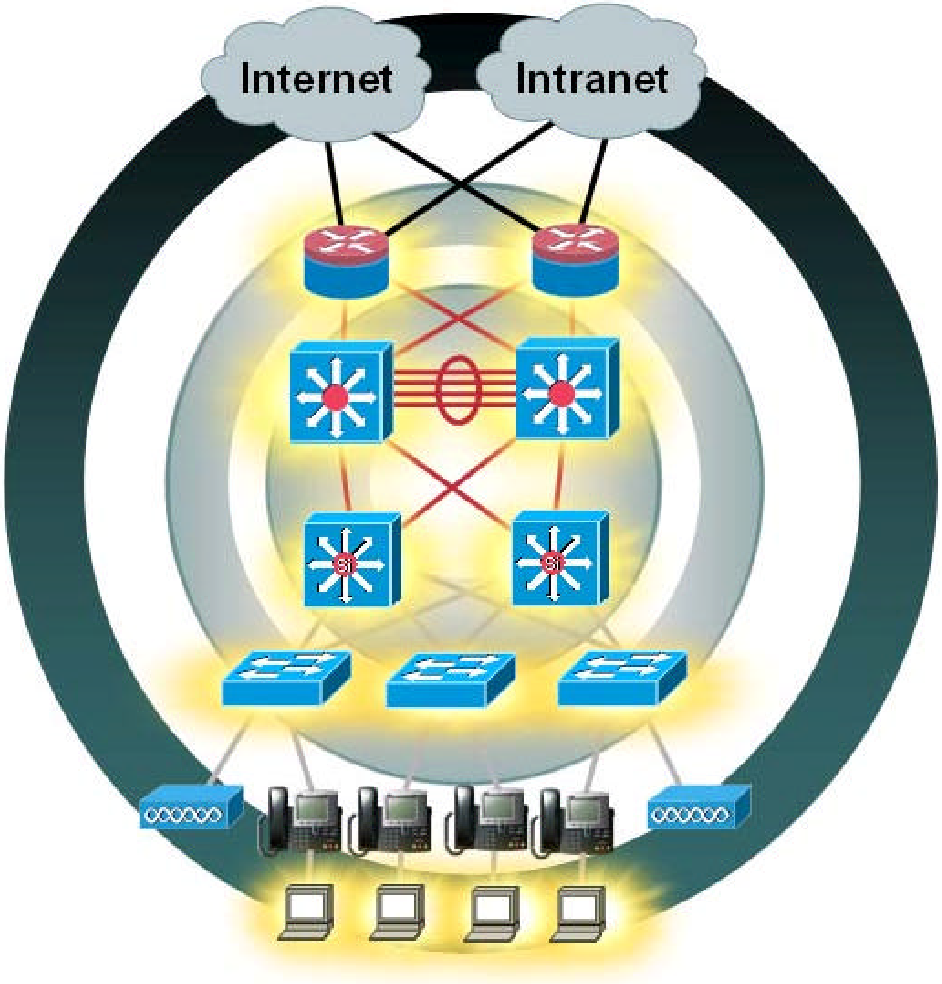
Real World Network Design:
Hierarchical Network Design—
Structured Modular Foundation

Reinforced Network Infrastructure:
Infrastructure Security Hardening
Device-Level and Software Resiliency

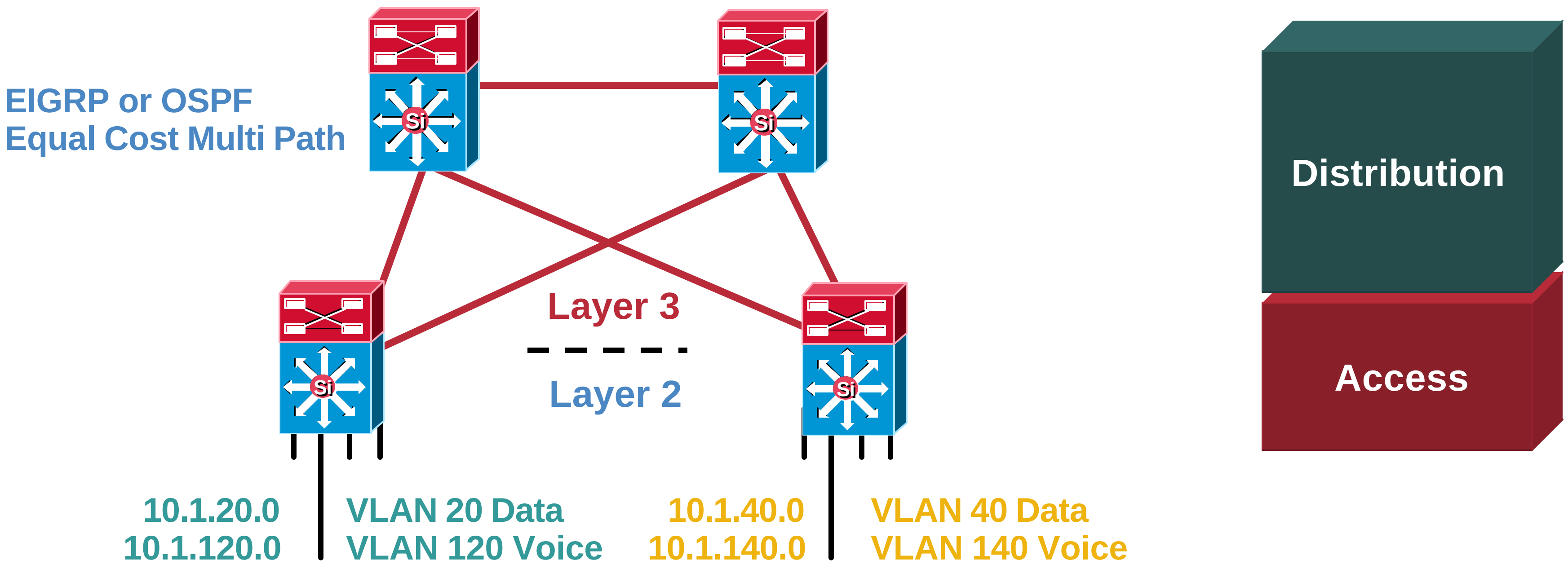
Network Operations:
Best Practices

Real-Time Network Management:
Best Practices

Best-in-Class Support:
TAC, CA, Etc.



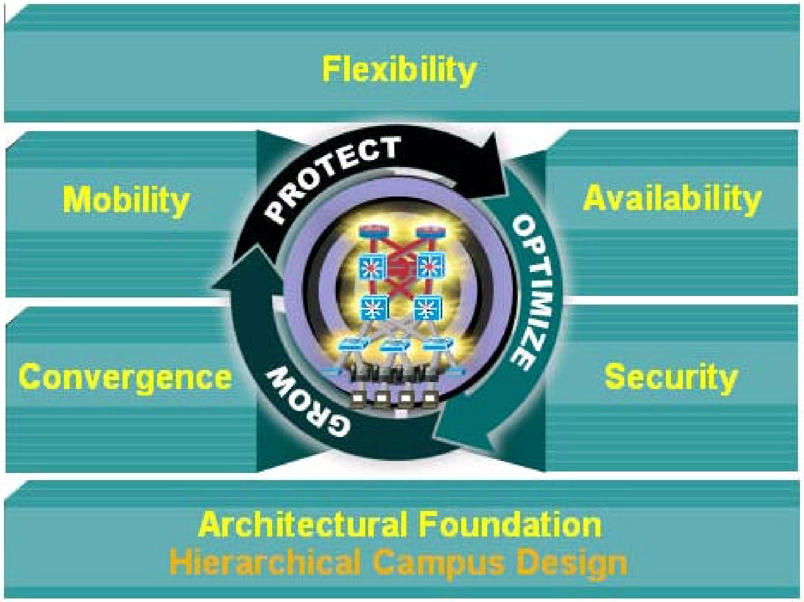
Routed Access Design
Structured Design Foundation



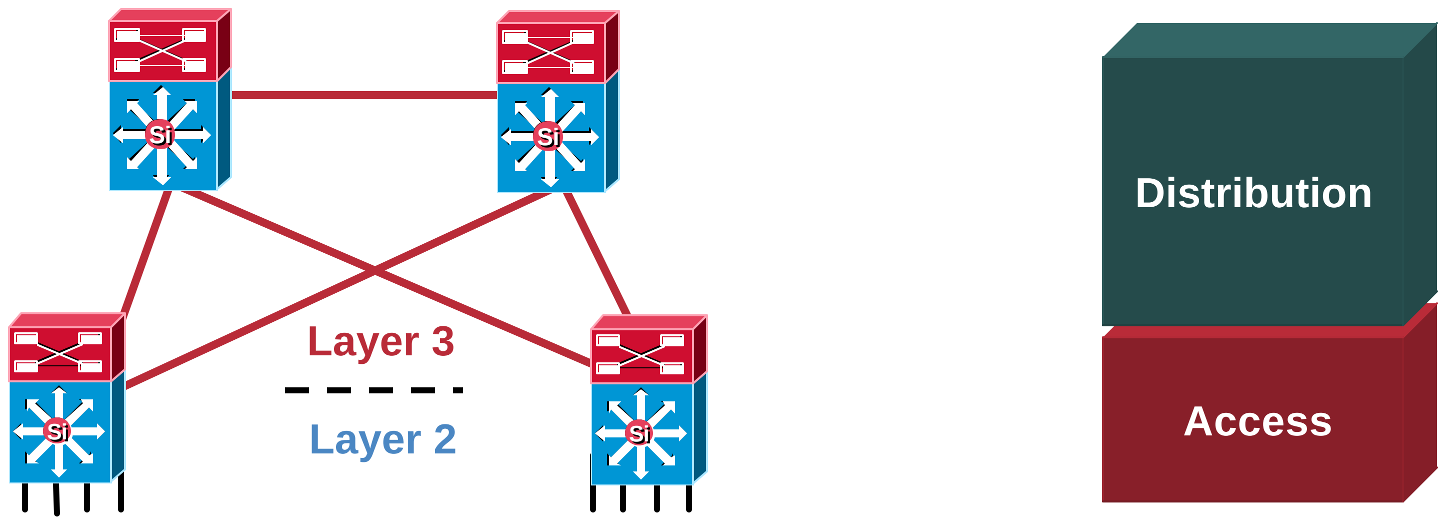
- EIGRP or OSPF routed links between access and distribution
- Routed interfaces, not VLAN trunks, between switches
- Equal cost multi path to load balance traffic across network
- Route summarization at distribution (like L2/L3)
- Single (IGP) control plane to configure/manage (no STP, HSRP,)

Agenda

- Campus Network Designs
- Routed Access Design
- EIGRP Design Details
- OSPF Design Details
- PIM Design Details
- Summary



Why Routed Access Campus Design?



- Most Catalysts® support L3 switching today
- EIGRP/OSPF routing preference over spanning tree
- Single control plane and well known tool set
  - Traceroute, show ip route, sho ip eigrp neighbor, etc…
- IGP enhancements; stub router/area, fast reroute, etc..
- It is another design option available to you

Ease of Implementation

- **Less to get right:**

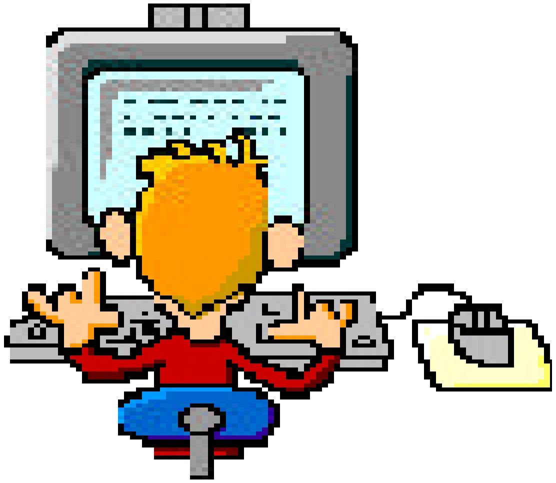
  No STP feature placement core to distribution

    LoopGuard
    RootGuard
    STP Root

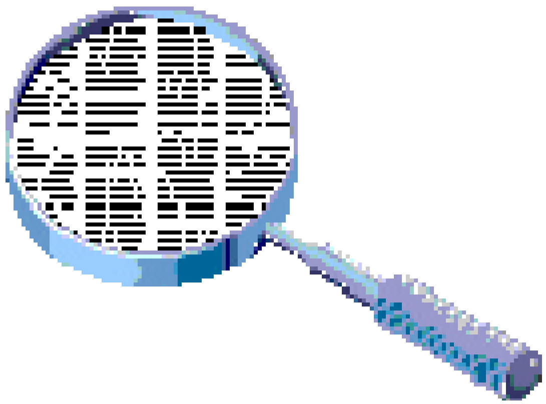
  No default gateway redundancy setup/tuning

  No matching of STP/HSRP/GLBP priority

  No L2/L3 multicast topology inconsistencies



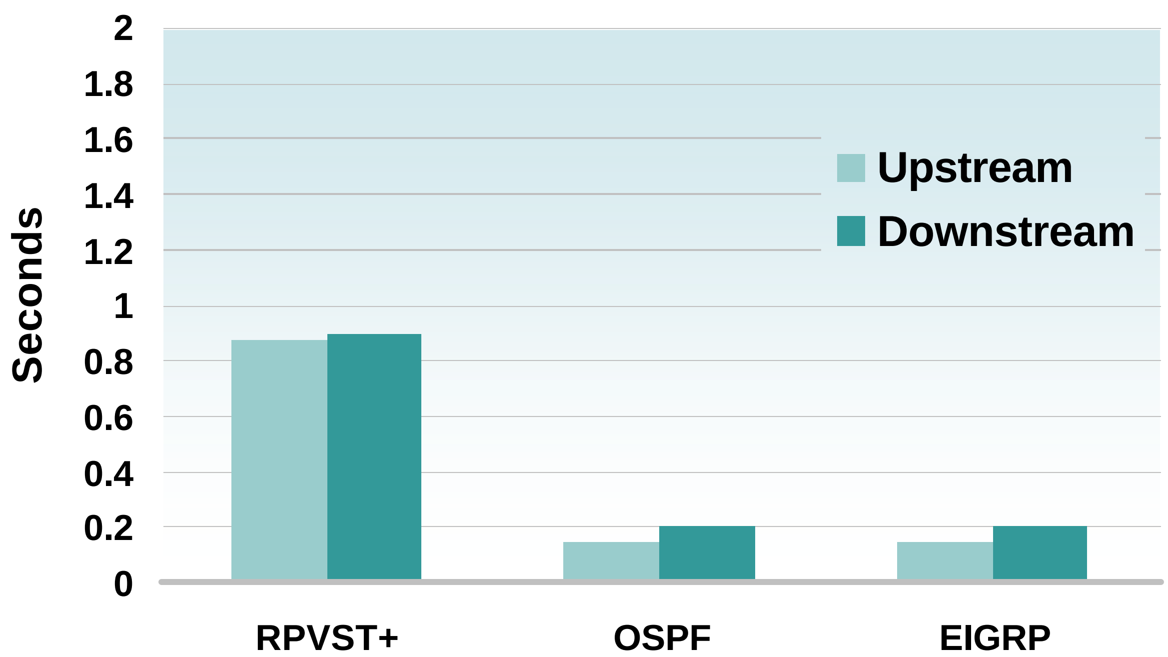
Ease of Troubleshooting

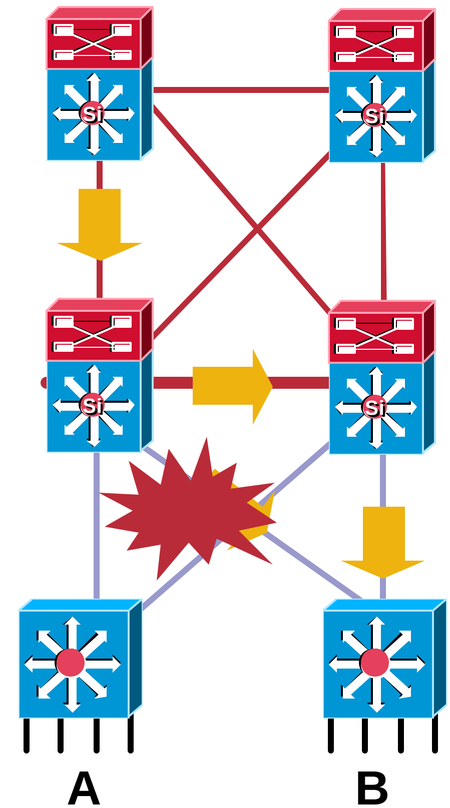


- **Routing troubleshooting tools**
  - Show ip route
  - Traceroute
  - Ping and extended pings
  - Extensive protocol debugs
  - Consistent troubleshooting; access, dist, core
- **Bridging troubleshooting tools**
  - Show ARP
  - Show spanning-tree, standby, etc…
  - Multiple show CAM dynamic's to find a host
- **Failure differences**
  - Routed topologies fail closed—i.e. neighbor loss
  - Layer 2 topologies fail open—i.e. broadcast and unknowns flooded

Routing to the Edge

Advantages? Yes, in the Right Environment





- EIGRP and OSPF converge in <200 msec
- OSPF convergence times dependent on timer tuning
- RPVST+ convergence times dependent on GLBP/HSRP tuning

Routed Access Considerations

Cisco.com

- **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?**
- **Platform requirements;**

  Catalyst 6500 requires an MSFC with hybrid (CatOS and Cisco IOS®) in the access to get all the necessary switchport and routing features

  Catalyst 4500 requires a SUP4 or higher for EIGRP or OSPF

  Catalyst 3500s and 3700s require an enhanced Cisco IOS image for EIGRP and OSPF

Interior Gateway Protocol Options
Static Routing

- **Benefits**
  - Price; in default Cisco IOS feature set for routers and Layer 3 switches
- **Considerations**
  - Configuration intensive and prone to error
  - Potential routing black holes during some failure conditions
- **Design guidance**
  - Default route from the access to the distribution
  - Specific route from the distribution to the access
  - Set next-hop to neighbor's adjacent IP interface address to minimize black holes during failure conditions
  - Redistribute static routes from distribution to core—summarize access subnets when possible

Interior Gateway Protocol Options

RIP Routing

- **Benefits**
  - Widely supported
  - Price; in default Cisco IOS feature set of Catalyst L3 switches
- **Considerations**
  - Slow convergence time
  - Limited network diameter; max hops = 16
  - Redistributing into an advanced IGP?
- **Design guidance**
  - Use RIP version two; VLSM
  - Tune hellos down to one second
  - Summarize routes from distribution to core
  - Use routed interfaces vs. VLAN trunks

Interior Gateway Protocol Options

EIGRP Routing

- **Benefits**
  - Simple to configure
  - Extremely fast convergence without tuning
  - Scales to large topologies
  - Flexible topology options
- **Considerations**
  - Cisco innovation
  - Summarization to limit query range
  - Price; requires enhanced IOS image in some Catalysts
- **Design guidance**
  - Later in the presentation

Interior Gateway Protocol Options

OSPF Routing

- **Benefits**
  - **Fast convergence with tuning**
  - **Widely deployed industry standard**
- **Considerations**
  - **Design and configuration complexity**
  - **Price; requires enhanced IOS image in most Catalysts**
  - **Topology design restrictions**
- **Design guidance**
  - **Later in the presentation**

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

- **Convergence:**

  Within the campus environment, both EIGRP and OSPF provide extremely fast convergence

  EIGRP requires summarization

  OSPF requires summarization and timer tuning for fast convergence

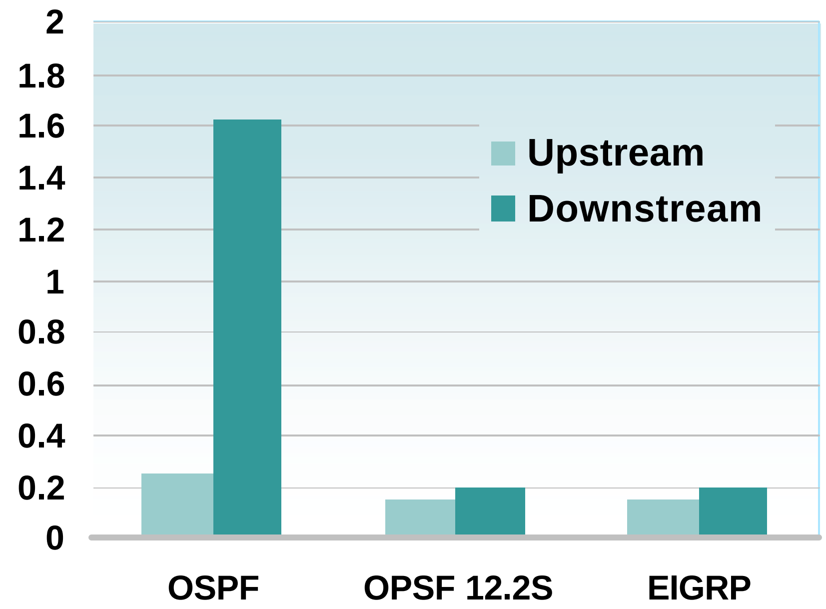
- **Flexibility:**

  EIGRP supports multiple levels of route summarization and route filtering which simplifies migration from the traditional multilayer L2/L3 campus design

  OSPF area design restrictions need to be considered

- **Scalability:**

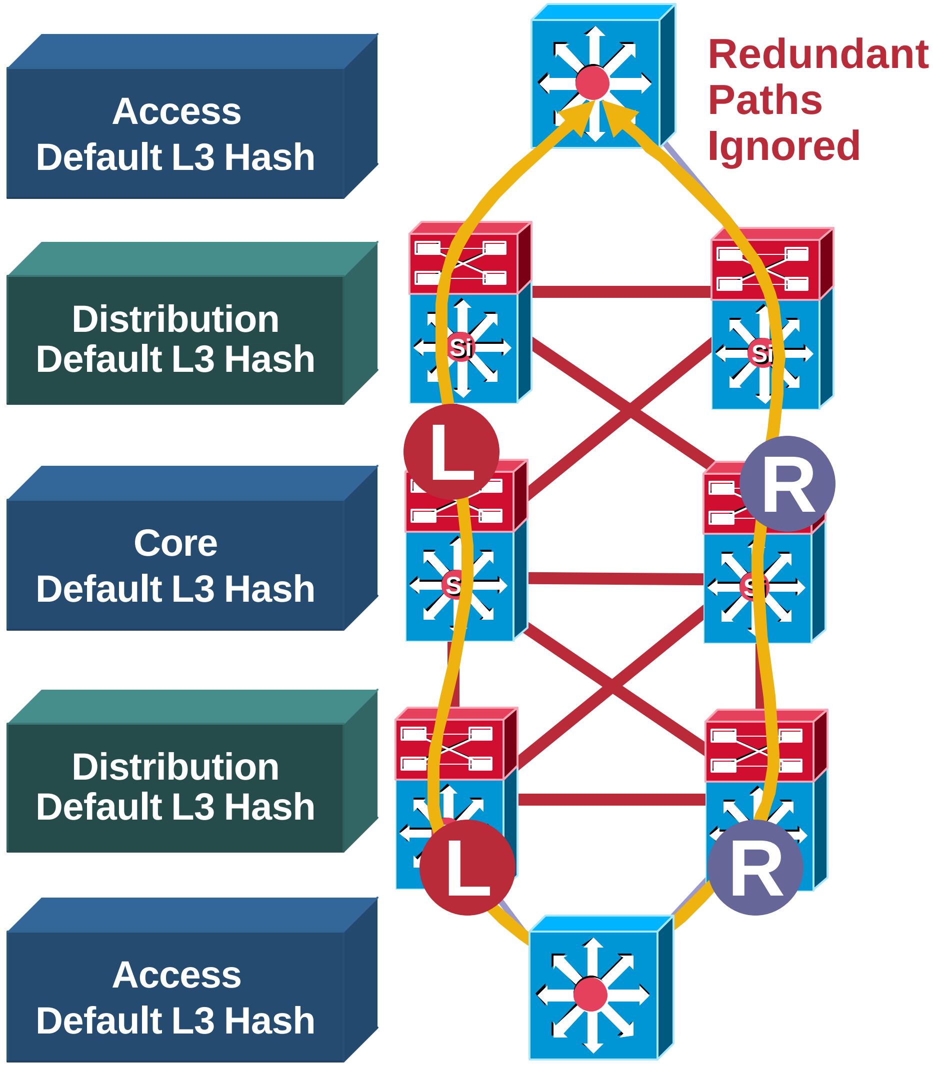
  Both protocols can scale to support very large enterprise network topologies



CEF Load Balancing

Avoid Underutilizing Redundant Layer 3 Paths

- The default CEF hash 'input' is L3
- **CEF polarization:** In a multi-hop design, CEF could select the same left/left or right/right path
- Imbalance/overload could occur
- Redundant paths are ignored/underutilized



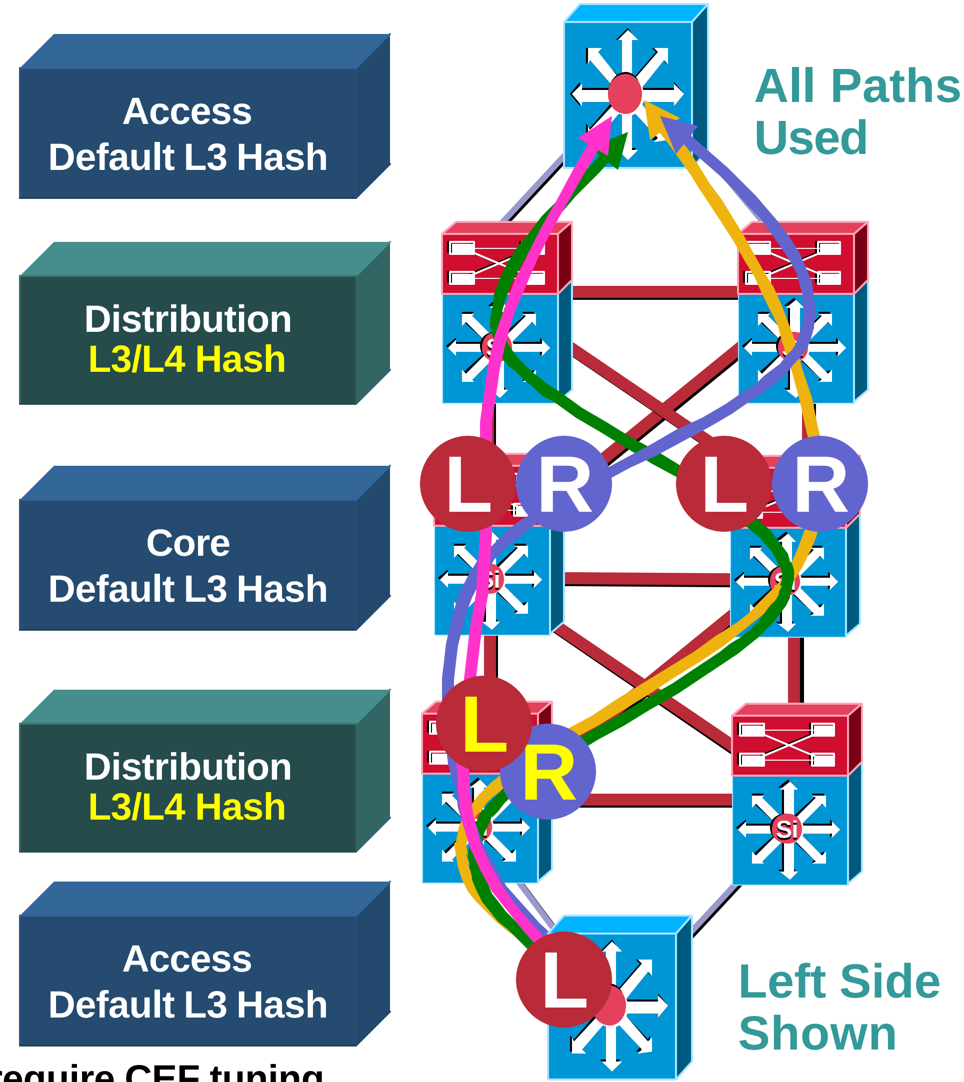
CEF Load Balancing

Avoid Underutilizing Redundant Layer 3 Paths

- With defaults, CEF could select the same left/left or right/right paths and ignore some redundant paths
- Alternating L3/L4 hash and default L3 hash will give us the best load balancing results
- The default is L3 hash—no modification required in core or access
- In the distribution switches use:

```
mls ip cef load-sharing full
```

  to achieve better redundant path utilization

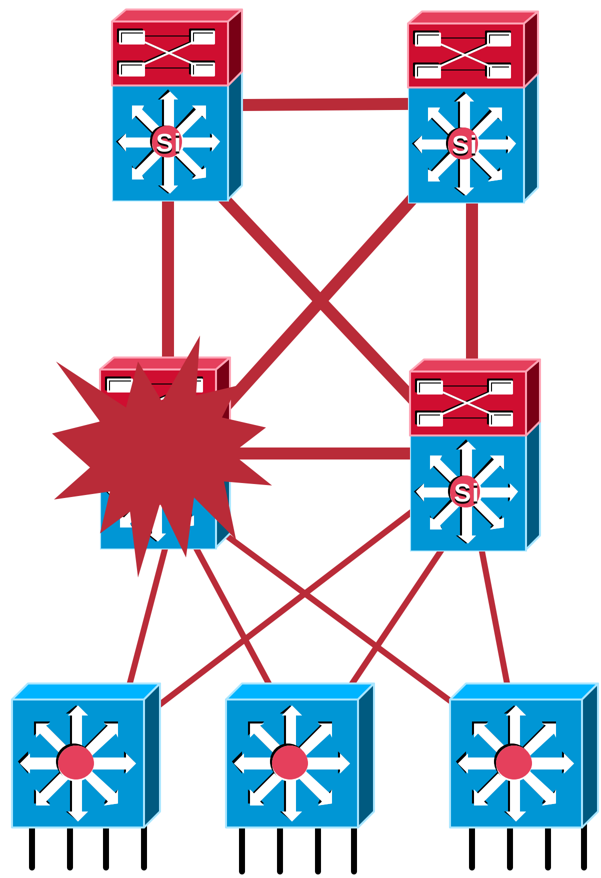


Note: Catalyst 6500 SUP720 does not require CEF tuning

Routed Access Design
High-Speed Campus Convergence

Cisco.com

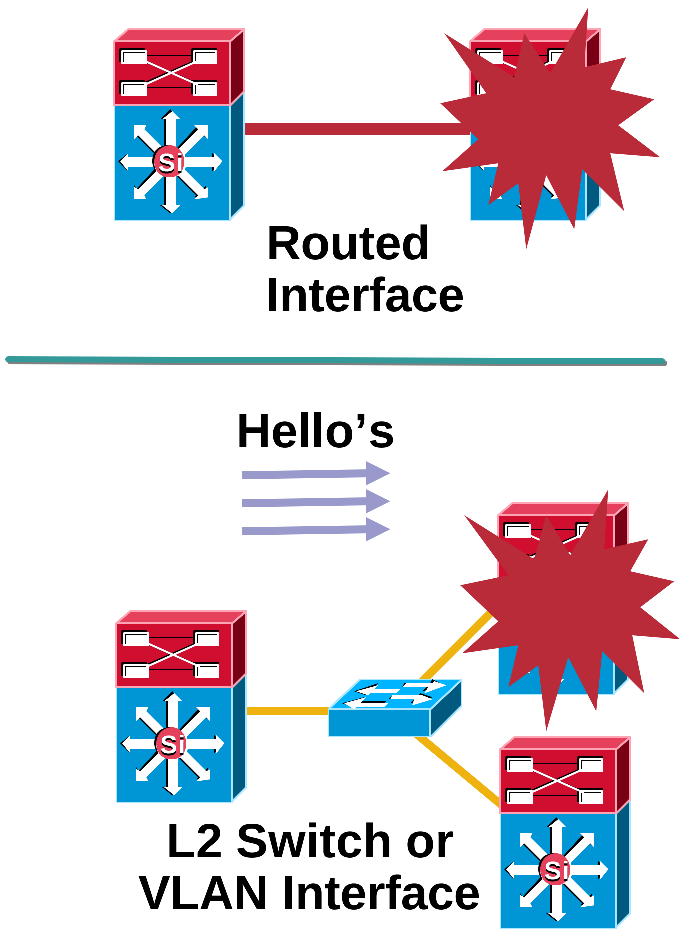
- **Convergence is the time needed for traffic to be rerouted to the alternative path after the network event**
- **Network convergence requires all affected routers to process the event and update the appropriate data structures used for forwarding**
- **Network convergence is the time required to:**
  - **Detect the event**
  - **Propagate the event**
  - **Process the event**
  - **Update the routing table/FIB**



High-Speed Campus Convergence—
Event Detection

Cisco.com

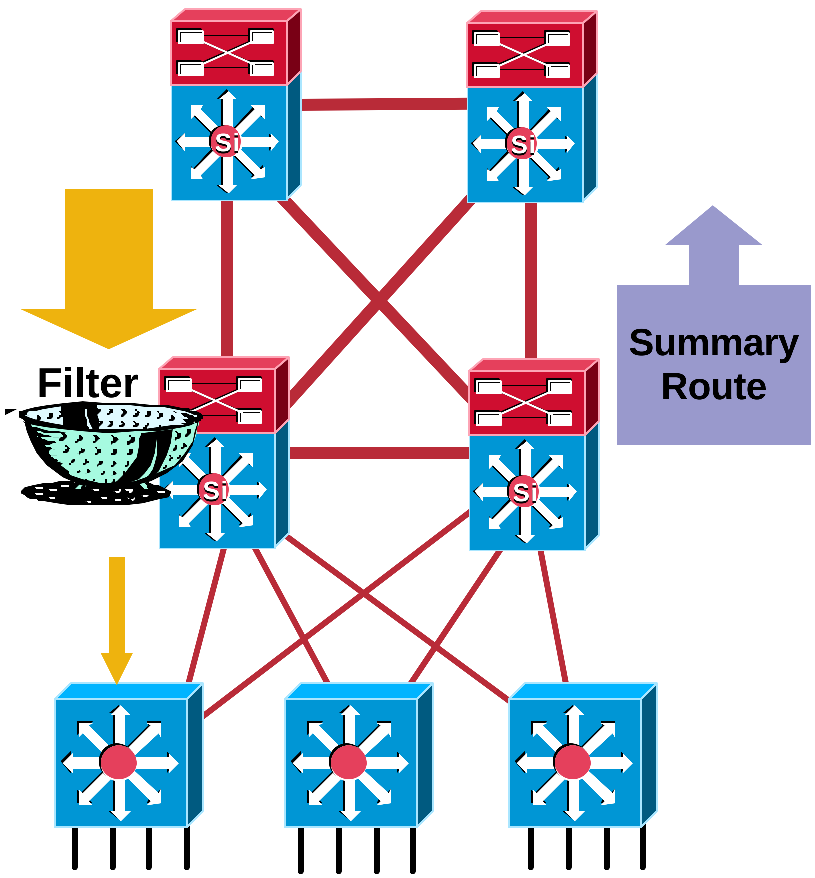
- **When physical interface changes state, the routing process is notified**
  - This should happen in the ms range
- **Some events are detected by the routing protocol**
  - L2 switch between L3 devices is a typical example
    - Neighbor is lost, but interface is UP/UP
  - Hello mechanism has to detect the neighbor loss
- **To improve failure detection**
  - Use routed interfaces between L3 switches
  - Decrease interface carrier-delay to 0s
  - Decrease IGP hello timers
    - EIGRP: Hellos = 1, Hold-down = 3
    - OSPF: Hellos = 250ms



High-Speed Campus Convergence—
Propagate the Event

Cisco.com

- **When an event occurs that changes the topology, all routers that were previously aware of the path need to be notified about the topology change**
- **EIGRP uses the query/reply process to find alternate paths**
- **OSPF propagates LSAs and all affected routers recalculate SPF to find alternate paths**
  - **LSA timer tuning can improve OSPF event propagation performance**
- **Summarization and route filtering can be used to limit the number of routers needing to participate in a network topology change event**



High-Speed Campus Convergence—
Process the Event

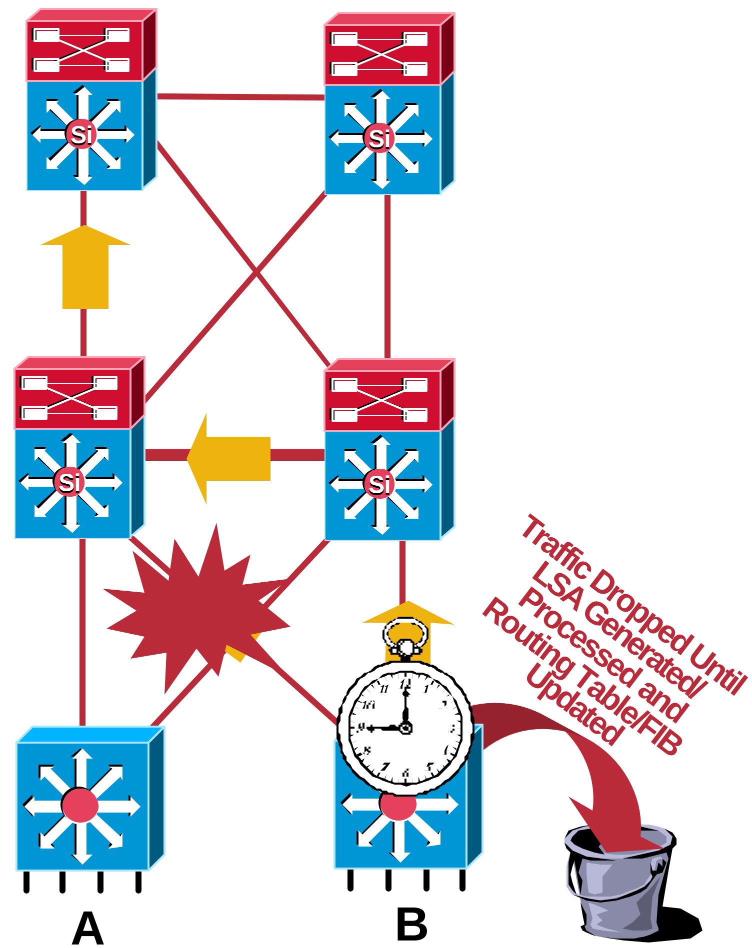
Cisco.com

- **Once a router has been notified that a topology changing event has occurred, it must recalculate a new path or topology for forwarding traffic**
- **EIGRP uses the DUAL algorithm to calculate a next hop successor(s) and possibly feasible successor(s)**
- **OSPF uses the Dijkstra SPF algorithm to calculate a shortest path tree for the new topology**
  - **SPF timer tuning can speed up SPF processing time**

High-Speed Campus Convergence—
Update the Routing Table and FIB

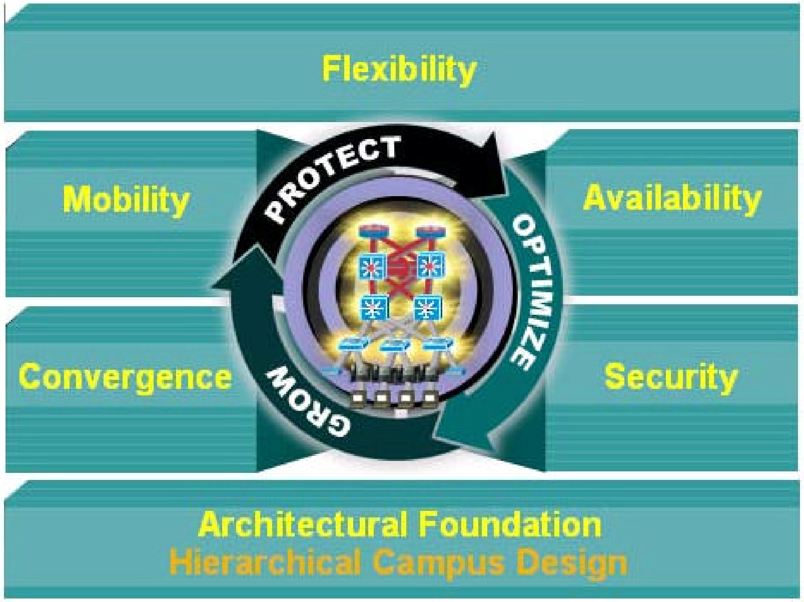
Cisco.com

- **After a new path or topology has been calculated by the protocol algorithm, the routing table must be updated**
  - **Routing Information Base (RIB) is the routing table**
  - **Forwarding Information Base (FIB) is based on the RIB and used by the hardware to forward traffic**
- **Projects are under way to make the RIB faster, more scalable and to improve the FIB info download to the line-cards**
- **Summarization and route filtering can be used to limit the number of routes needed in the RIB and FIB**



Agenda

- Campus Network Designs
- Routed Access Design
- EIGRP Design Details
- OSPF Design Details
- PIM Design Details
- Summary



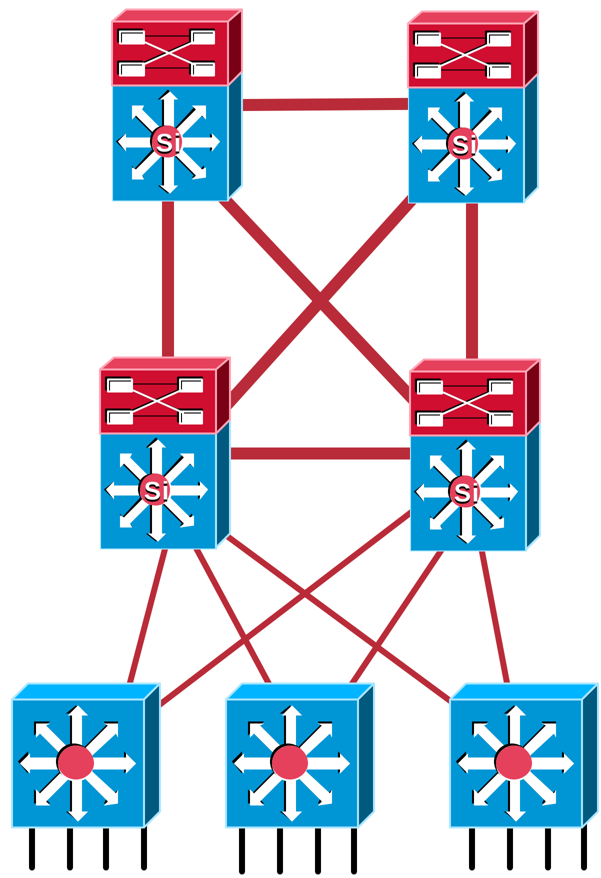
Strengths of EIGRP

- Advanced distance vector
- Maps easily to the traditional multilayer design
- 100% loop free
- Fast convergence
- Easy configuration
- Incremental update
- Supports VLSM and discontiguous network
- Classless routing
- Protocol independent
    - IPv6, IPX and AppleTalk
- Unequal cost paths load balancing
- Flexible topology design options

EIGRP Design Rules for HA Campus

Similar to WAN Design, But…

- **EIGRP design for the campus follows all the same best practices as you use in the WAN with a few differences**
  - **No BW limitations**
  - **Lower neighbor counts**
  - **Direct fiber interconnects**
  - **Lower cost redundancy**
  - **HW switching**
- **WAN → stability and speed**
- **Campus → stability, redundancy, load sharing, and high speed**

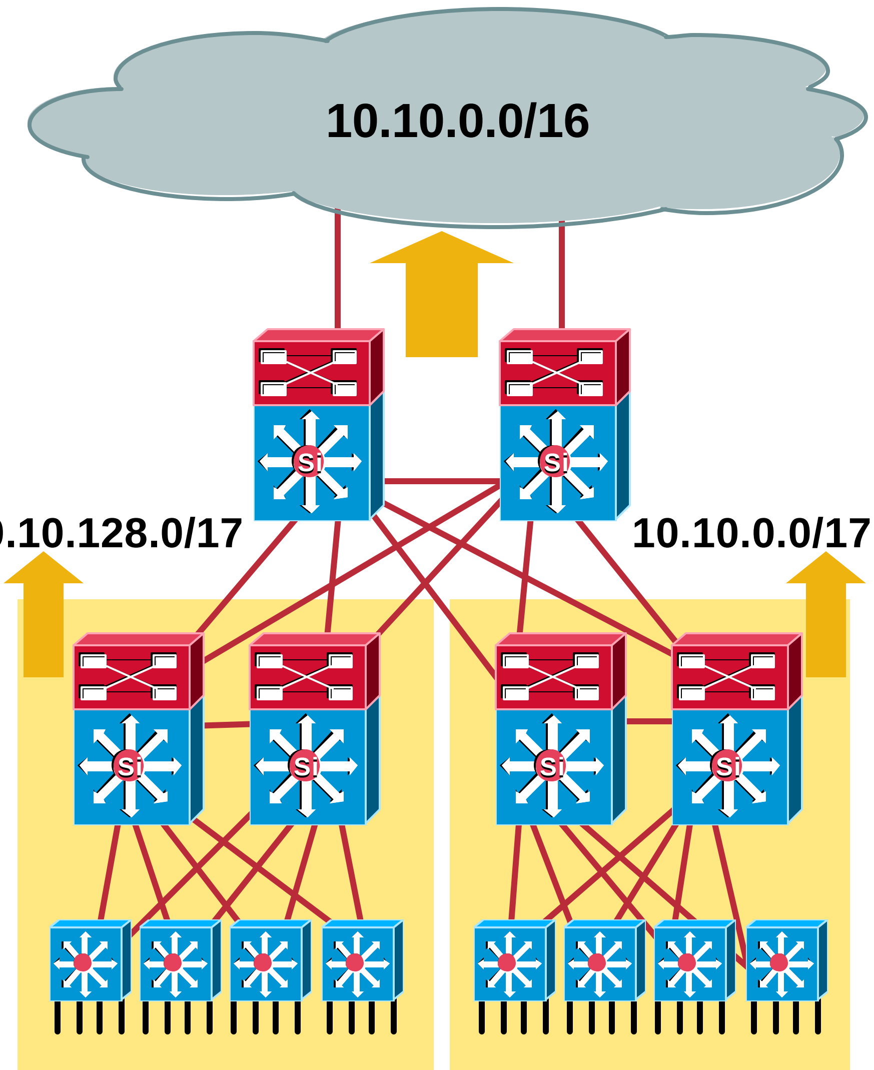


EIGRP in the Campus

Conversion to an EIGRP Routed Edge

Cisco.com

- **The greatest advantages of extending EIGRP to the access are gained when the network has a structured addressing plan that allows for use of summarization and stub routers**
- **EIGRP provides the ability to implement multiple tiers of summarization and route filtering**
- **Relatively painless to migrate to a L3 access with EIGRP if network addressing scheme permits**
- **Able to maintain a deterministic convergence time in very large L3 topology**



EIGRP Protocol Fundamentals

Metric:

- Metric = [K1 x BW + (K2 x BW)/(256 - Load) + K3 x Delay] x [K5/(Reliability + K4)] x 256
  - By Default: K1 = 1, K2 = 0, K3 = 1, K4 = K5 = 0
- Delay is sum of all the delays along the path
  - Delay = Delay/10
- Bandwidth is the lowest bandwidth link along the path
  - Bandwidth = 10000000/Bandwidth

Packets:

- Hello: establish neighbor relationships
- Update: send routing updates
- Query: ask neighbors about routing information
- Reply: response to query about routing information
- Ack: acknowledgement of a reliable packet

EIGRP Protocol Fundamentals (Cont.)

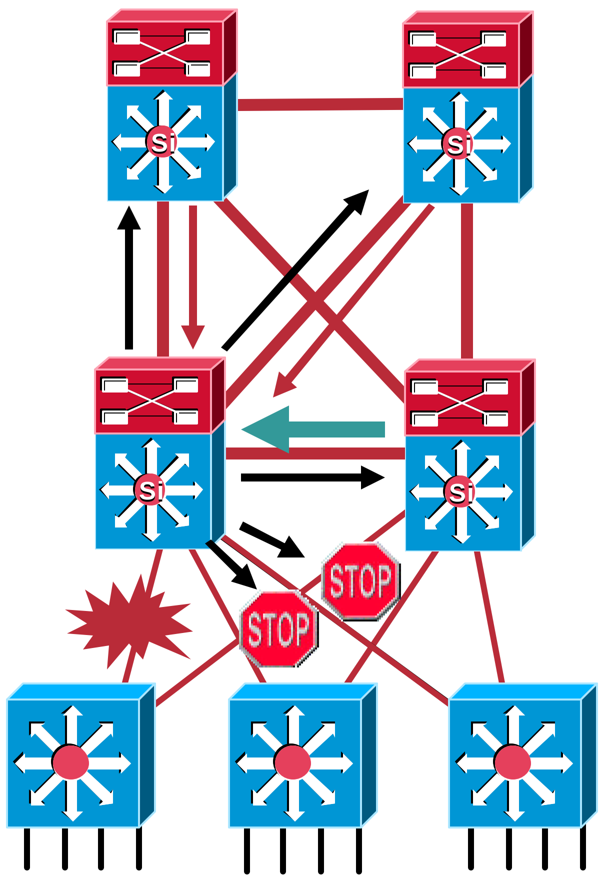
DUAL Algorithm

- Diffusing update algorithm
- Finite-state-machine
  - Track all routes advertised by neighbors
  - Select loop-free path using a successor and remember any feasible successors
  - If successor lost
    - Use feasible successor
  - If no feasible successor
    - Query neighbors and recompute new successor
- A successor is a neighbor that has met the Feasibility Condition (FC) and has the least cost path towards the destination
- Multiple successors are possible (load balancing)
- A feasible successor is the neighbor with the next best loop free next hop towards destination

EIGRP Design Rules for HA Campus

Limit Query Range to Maximize Performance

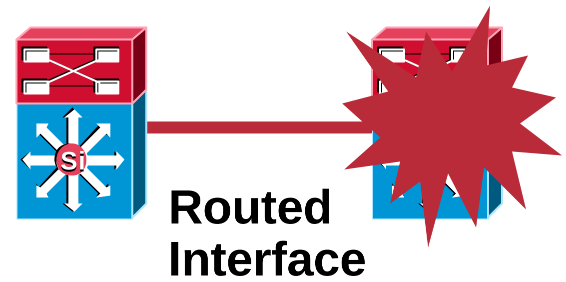
- **EIGRP convergence is largely dependent on query response times**
- **Minimize the number of queries to speed up convergence**
- **Summarize distribution block routes upstream to the core**
  - **Upstream queries are returned immediately with infinite cost**
- **Configure all access switches as EIGRP stub routers**
  - **No downstream queries are ever sent**

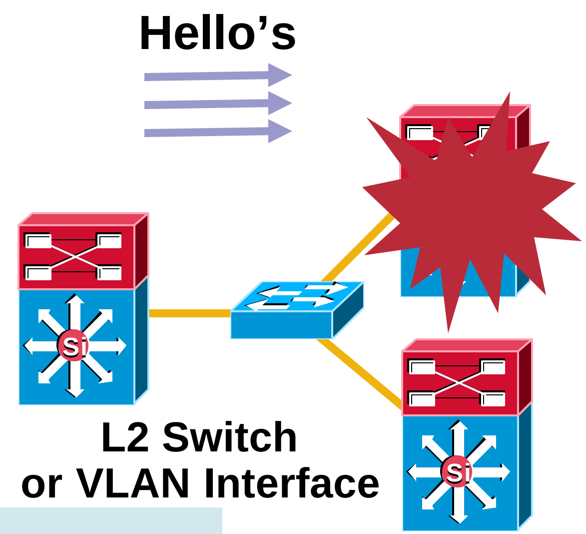


EIGRP Neighbors
Event Detection

Cisco.com

- **EIGRP neighbor relationships are created when a link comes up and routing adjacency is established**
- **When physical interface changes state, the routing process is notified**
  - **Carrier-delay should be set as a rule because it varies based upon the platform**
- **Some events are detected by the routing protocol**
  - **Neighbor is lost, but interface is UP/UP**
- **To improve failure detection**
  - **Use Routed Interfaces and not SVIs**
  - **Decrease interface carrier-delay to 0**
  - **Decrease EIGRP hello and hold-down timers**
    - **Hello = 1**
    - **Hold-down = 3**



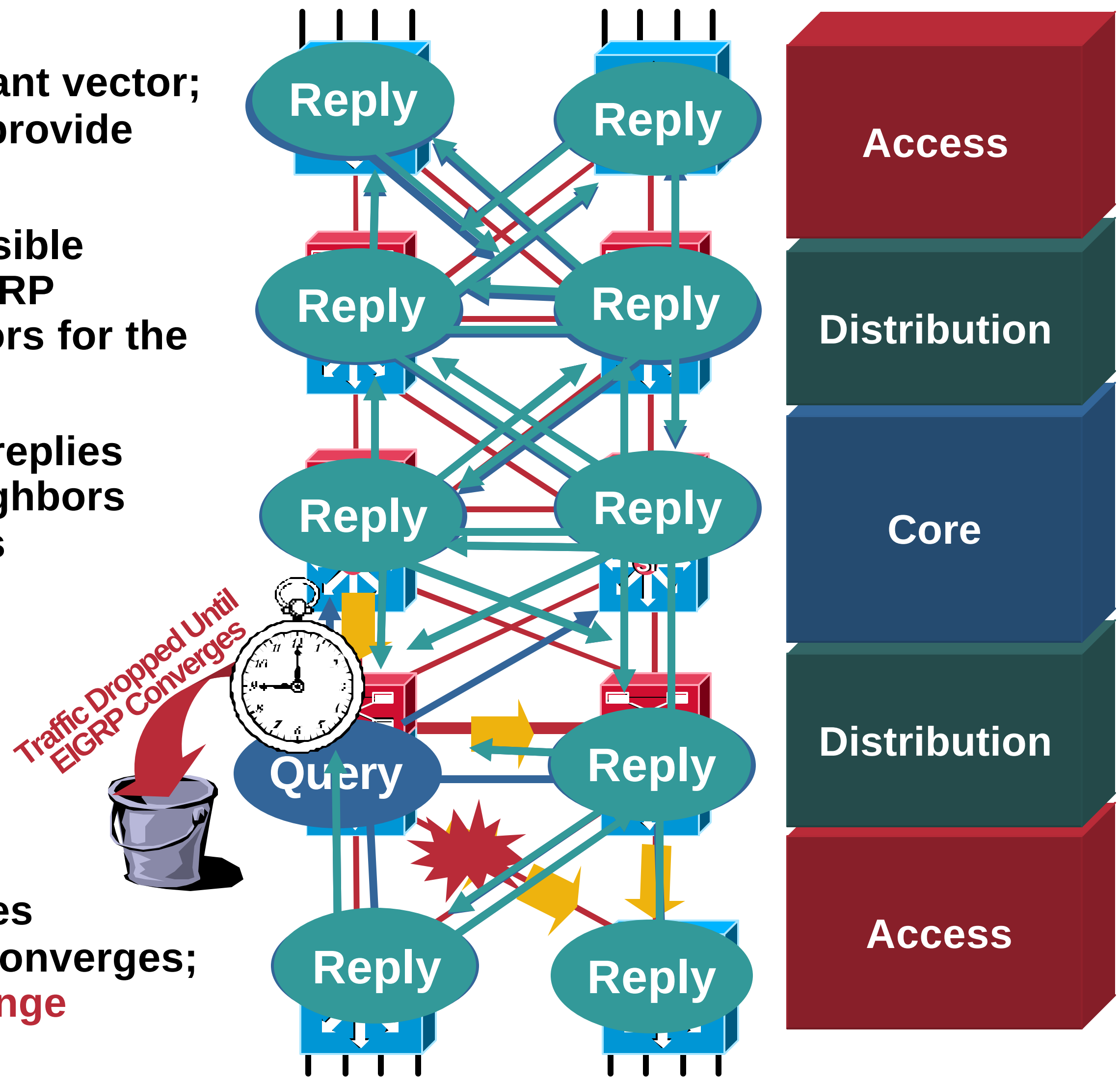



```
interface GigabitEthernet3/2
  ip address 10.120.0.50 255.255.255.252
  ip hello-interval eigrp 100 1
  ip hold-time eigrp 100 3
  carrier-delay msec 0
```

EIGRP Query Process
Queries Propagate the Event

Cisco.com

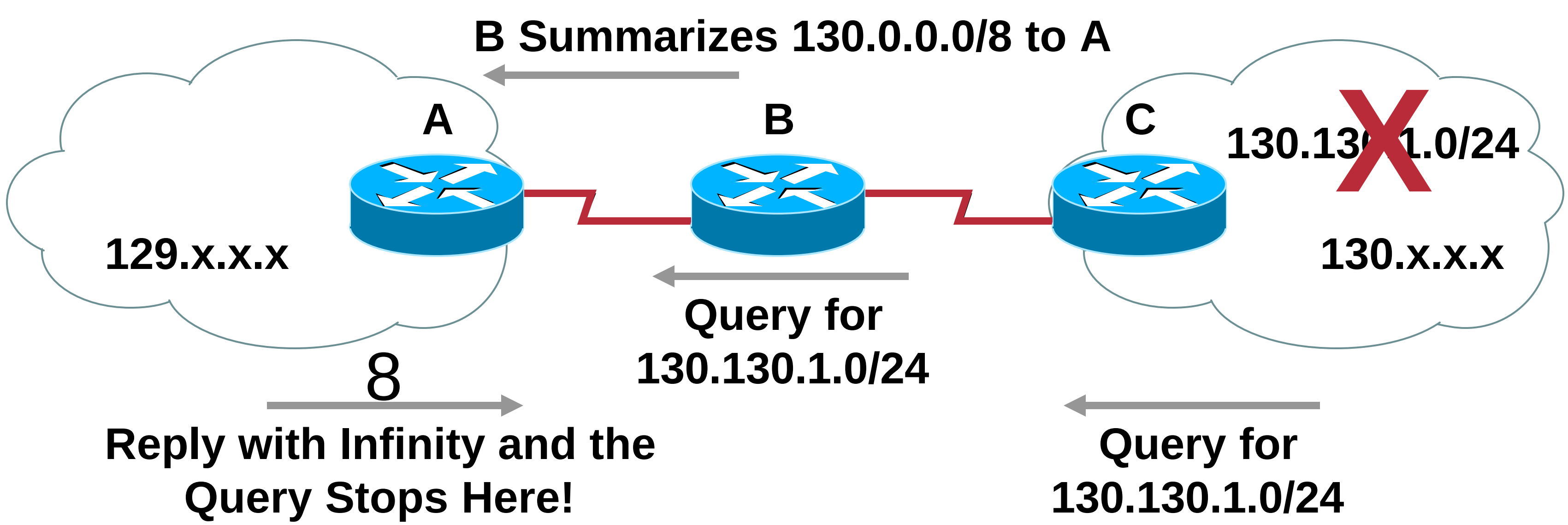
- EIGRP is an advanced distant vector; it relies on its neighbor to provide routing information
- If a route is lost and no feasible successor is available, EIGRP actively queries its neighbors for the lost route(s)
- The router will have to get replies back from ALL queried neighbors before the router calculates successor information
- If any neighbor fails to reply, the queried route is stuck in active and the router resets the neighbor that fails to reply
- The fewer routers and routes queried, the faster EIGRP converges; solution is to limit query range



EIGRP Query Range

Cisco.com

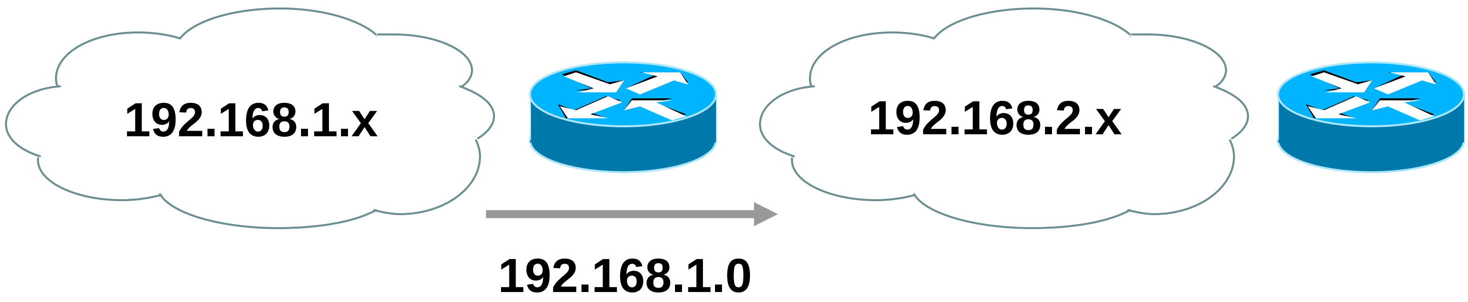
- **Summarization point**
  - **Auto or manual summarization bound queries**
  - **Requires a good address allocation scheme**
- **Stubs also help to reduce the query range**



EIGRP Summarization

Smaller Routing Tables, Smaller Updates, Query Boundary

Cisco.com

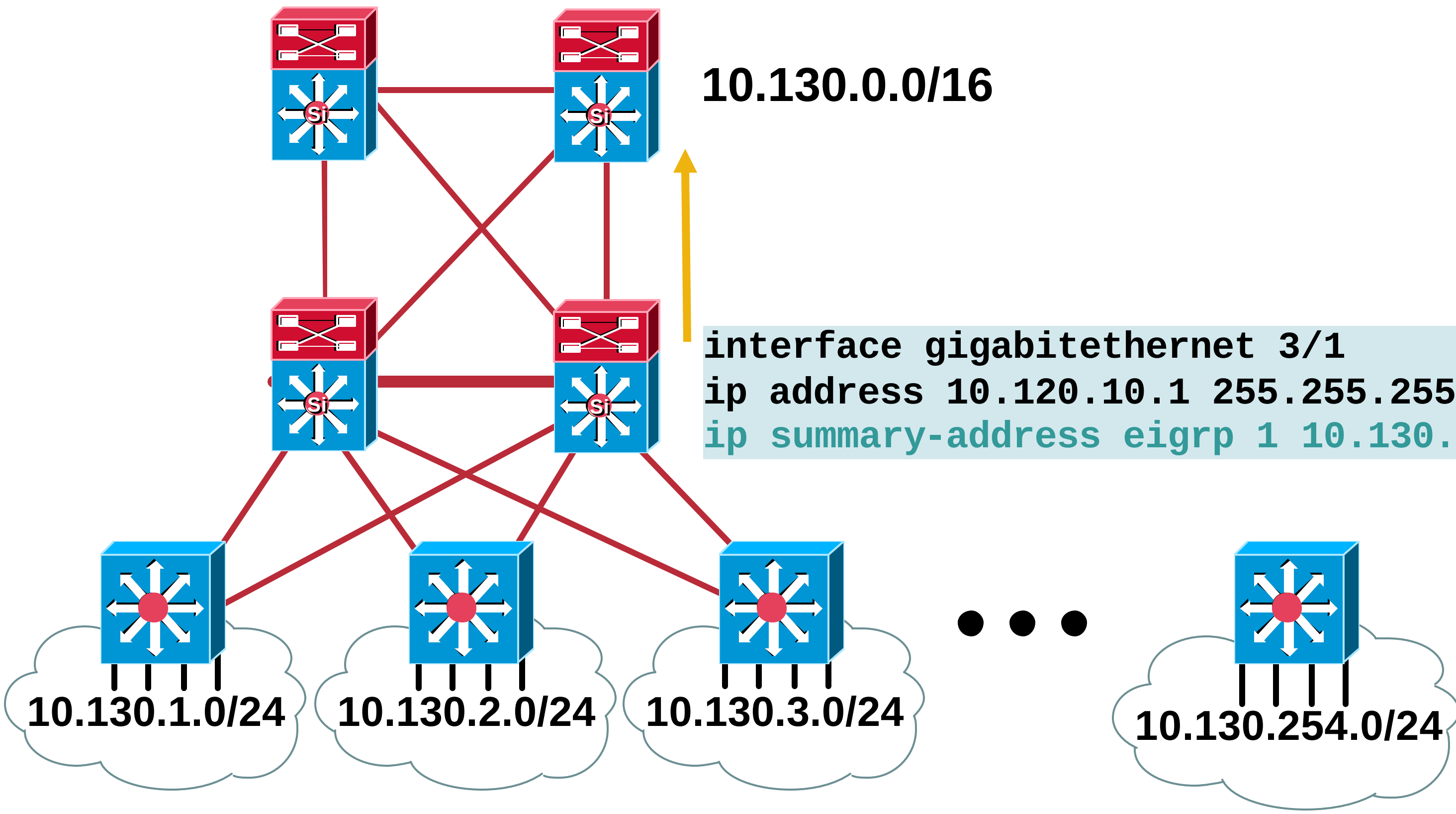


- **Auto summarization:**
  - On major network boundaries, networks are summarized to the major networks
  - Auto summarization is turned on by default
- **Manual summarization**
  - Configurable on per interface basis in any router within network
  - When summarization is configured on an interface, the router immediately creates a route pointing to null zero with administrative distance of five (5)
    - Loop prevention mechanism
  - When the last specific route of the summary goes away, the summary is deleted
  - The minimum metric of the specific routes is used as the metric of the summary route

Manual EIGRP Summarization

Cisco.com

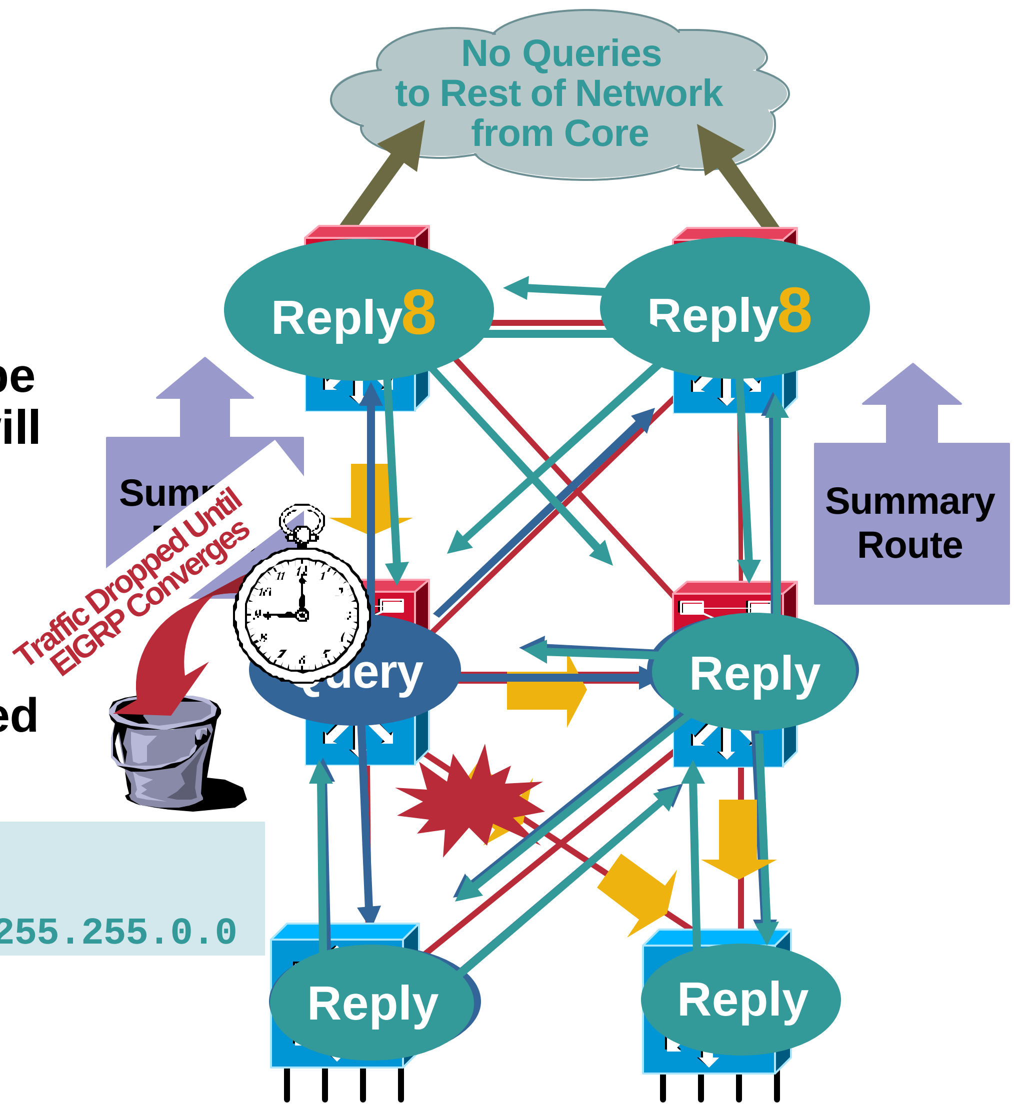
ip summary-address EIGRP <as number> <address> <mask>




```
interface gigabitethernet 3/1
ip address 10.120.10.1 255.255.255.252
ip summary-address eigrp 1 10.130.0.0 255.255.0.0
```

EIGRP Query Process with Summarization

- When we summarize from distribution to core for the subnets in the access we can limit the upstream query/ reply process
- In a large network this could be significant because queries will now stop at the core; no additional distribution blocks will be involved in the convergence event
- The access layer is still queried

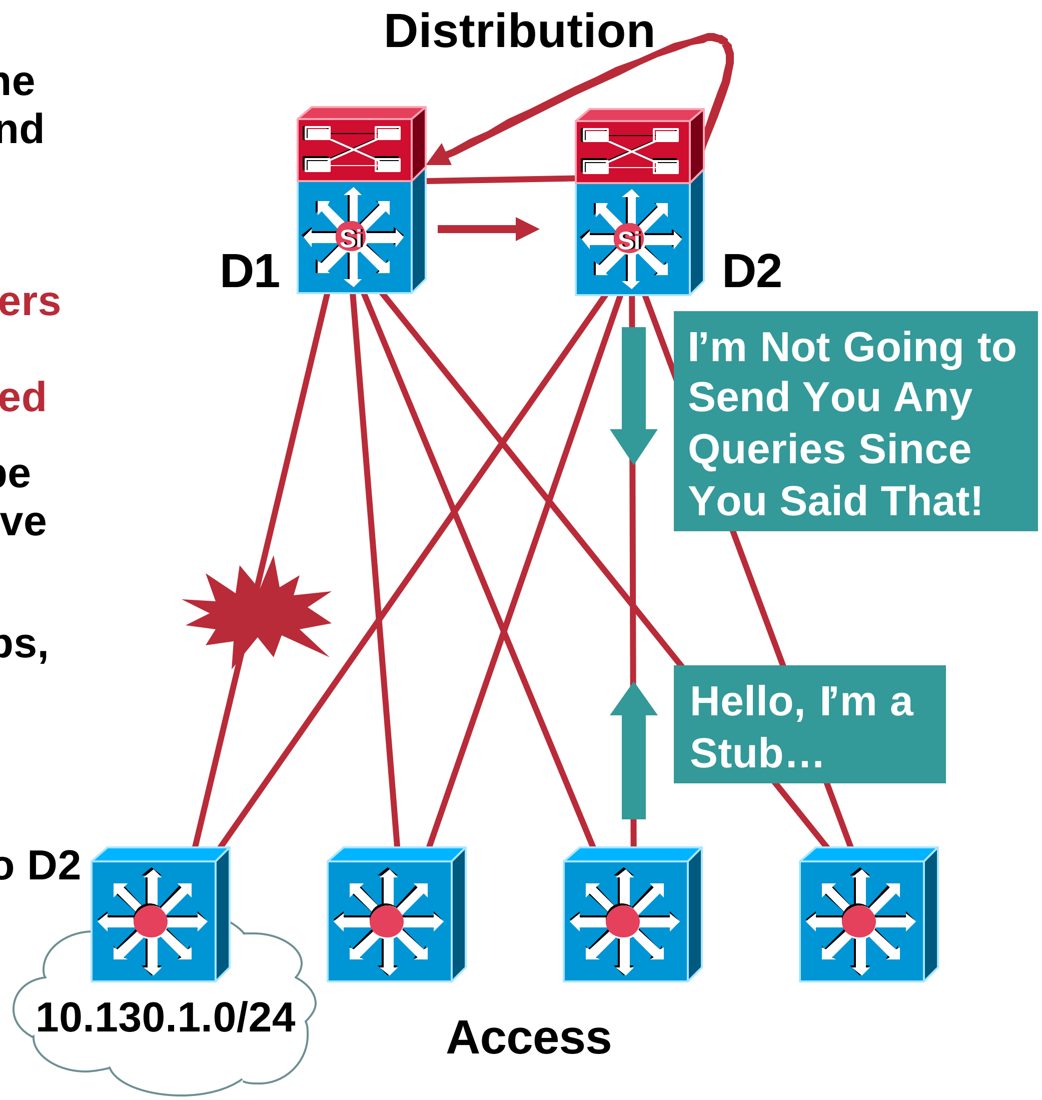



```
interface gigabitethernet 3/1
ip address 10.120.10.1 255.255.255.252
ip summary-address eigrp 1 10.130.0.0 255.255.0.0
```

EIGRP Stubs

Cisco.com

- A stub router signals (through the hello protocol) that it is a stub and should not transit traffic
- Queries that would have been generated towards the stub routers are marked as if a "No path this direction" reply had been received
- D1 will know that stubs cannot be transit paths, so they will not have any path to 10.130.1.0/24
- D1 simply will not query the stubs, reducing the total number of queries in this example to 1
- These stubs will not pass D1's advertisement of 10.130.1.0/24 to D2
- D2 will only have one path to 10.130.1.0/24



EIGRP Stubs

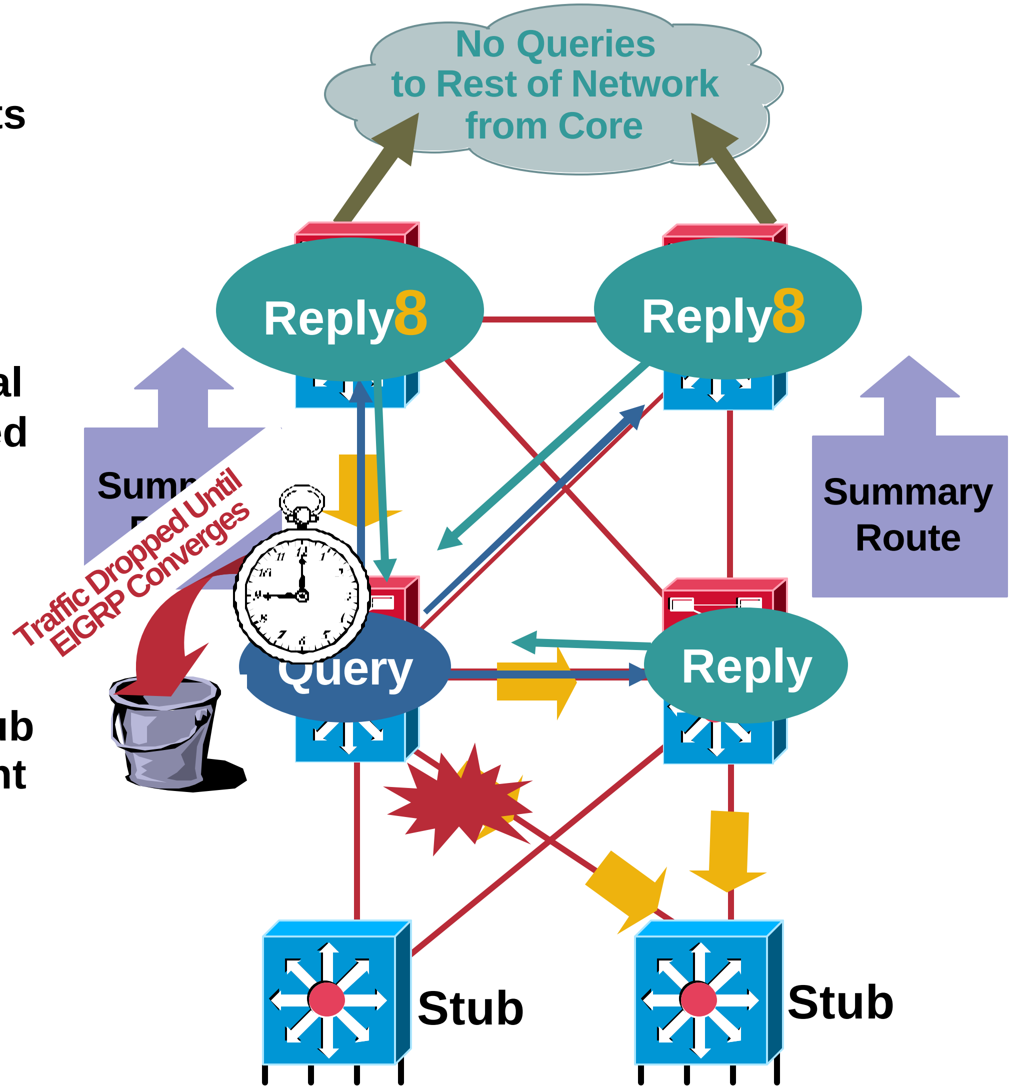

```
router(config-router)#EIGRP stub ?
  connected      Do advertise connected routes
  receive-only   Set IP-EIGRP as receive only neighbor
  static         Do advertise static routes
  summary        Do advertise summary routes
  <cr>
```

- Connected: advertise directly connected networks
- Static: advertise redistributed static routes
- Summary: advertise locally created summaries
- Receive-only: don't advertise anything

EIGRP Query Process

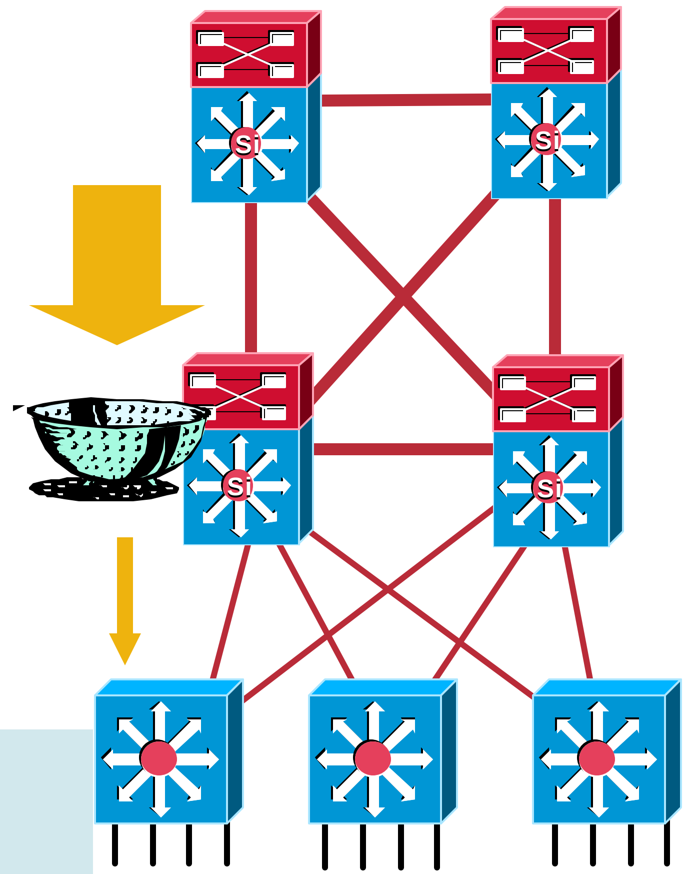
With Summarization and Stub Routers

- When we summarize from distribution to core for the subnets in the access we can limit the upstream query/reply process
- In a large network this could be significant because queries will now stop at the core; no additional distribution blocks will be involved in the convergence event
- When the access switches are EIGRP stub's we can further reduce the query diameter
- Non-stub routers do not query stub routers—so no queries will be sent to the access nodes
- No secondary queries—and only three nodes involved in convergence event



EIGRP Route Filtering in the Campus
Control Route Advertisements

- Bandwidth is not a constraining factor in the campus but it is still advisable to control number of routing updates advertised
- Remove/filter routes from the core to the access and inject a default route with distribute-lists
- Smaller routing table in access is simpler to troubleshoot
- Deterministic topology




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

ip access-list standard Default
 permit 0.0.0.0
```

EIGRP Routed Access Campus Design

Overview

- **Detect the event:**

  Set hello-interval = 1 second and hold-time = 3 seconds to detect soft neighbor failures

  Set carrier-delay = 0

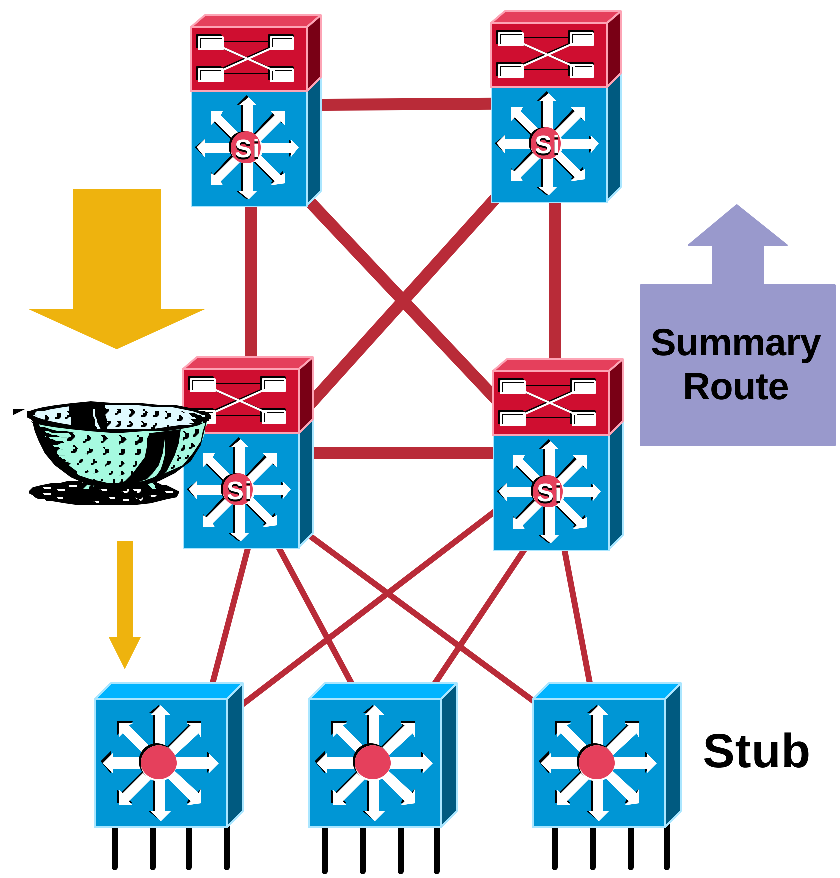
- **Propagate the event:**

  Configure all access layer switches as stub routers to limit queries from the distribution layer

  Summarize the access routes from the distribution to the core to limit queries across the campus

- **Process the event:**

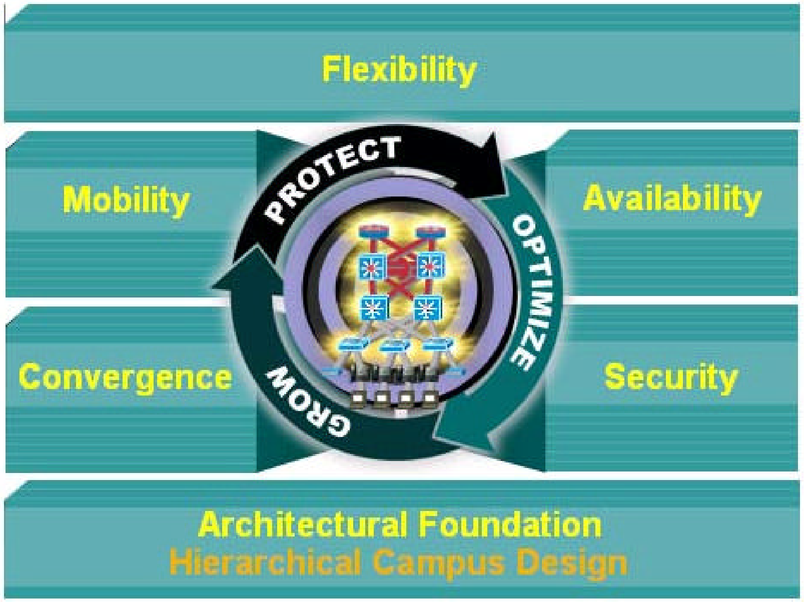
  Summarize and filter routes to minimize calculating new successors for the RIB and FIB



For More Discussion on EIGRP Design Best Practices—RST-3220-3222

Agenda

- Campus Network Designs
- Routed Access Design
- EIGRP Design Details
- OSPF Design Details
- PIM Design Details
- Summary



Open Shortest Path First (OSPF) Overview

- OSPFv2 established in 1991 with RFC 1247
- Goal—a link state protocol more efficient and scaleable than RIP
- Dijkstra Shortest Path First (SPF) algorithm
- Metric—path cost
- Fast convergence
- Support for CIDR, VLSM, authentication, multipath and IP unnumbered
- Low steady state bandwidth requirement
- OSPFv3 for IPv6 support

OSPF Metric
Cost = Metric

- **Cost applied on all router link paths**
- **The lower the more desirable**
- **Route decisions made on total cost of path**
- **Derived from bandwidth**

100000000 ÷ bandwidth	**56-kbps serial link = 1785**
Ethernet = 10	**64-kbps serial link = 1562**
T1 (1.544-Mbps serial link) = 65	**Fast Ethernet = 1**

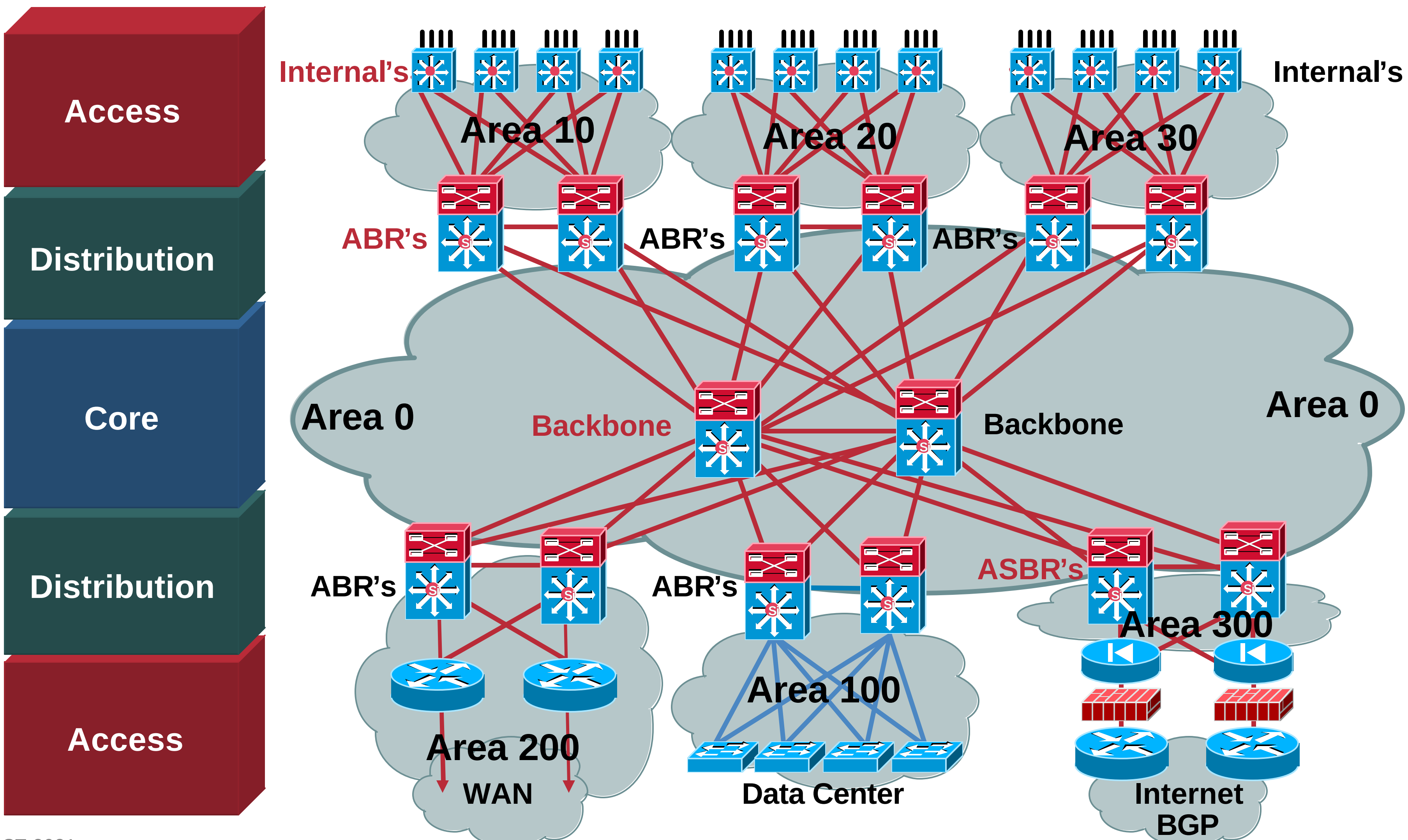
- **Configured via:**

Interface subcommand:	`bandwidth`
Interface subcommand:	`ip ospf cost`
Router subcommand:	`ospf auto-cost reference bandwidth`

Hierarchical Campus Design

OSPF Area's with Router Types

Cisco.com

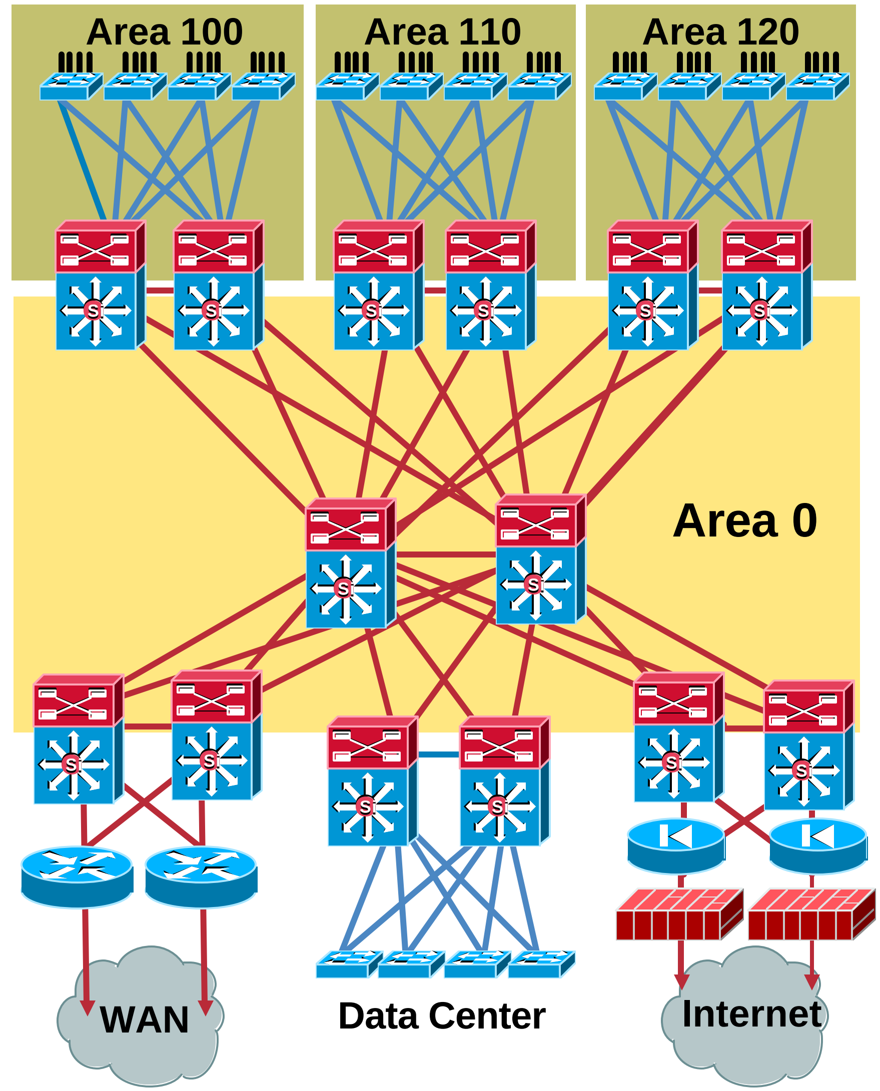


OPSF Design Rules for HA Campus

Where Are the Areas?

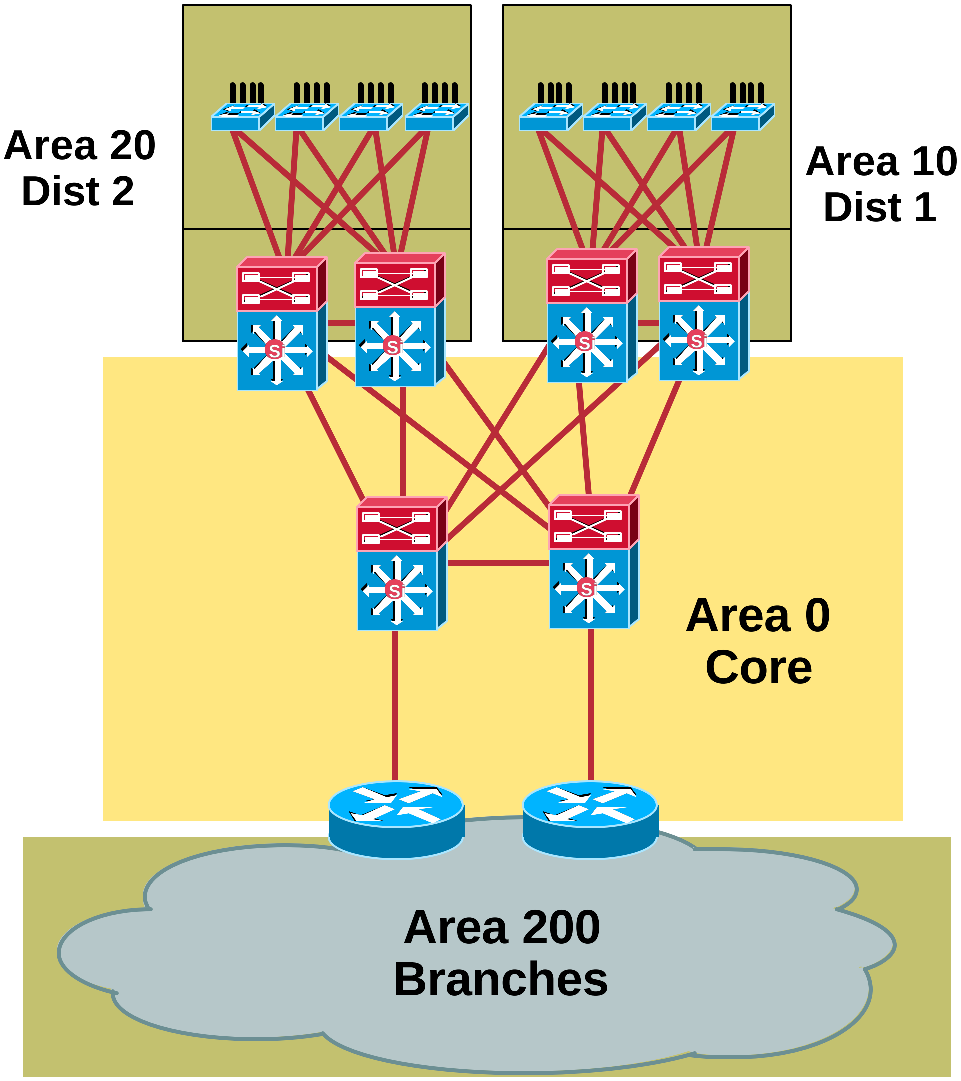
Cisco.com

- **Area size/border is bounded by the same concerns in the campus as the WAN**
- **In campus the lower number of nodes and stability of local links could allow you to build larger areas however…**
- **Area design also based on address summarization**
- **Area boundaries should define buffers between fault domains**
- **Keep area 0 for core infrastructure do not extend to the access routers**



OSPF in the Campus
Conversion to an OSPF Routed Edge

- OSPF designs that utilize an area for each campus distribution building block allow for straight forward migration to Layer 3 access
- Converting L2 switches to L3 within a contiguous area is reasonable to consider as long as new area size is reasonable
- How big can the area be?
- It depends!
  - Switch type(s)
  - Number of links
  - Stability of fiber plant

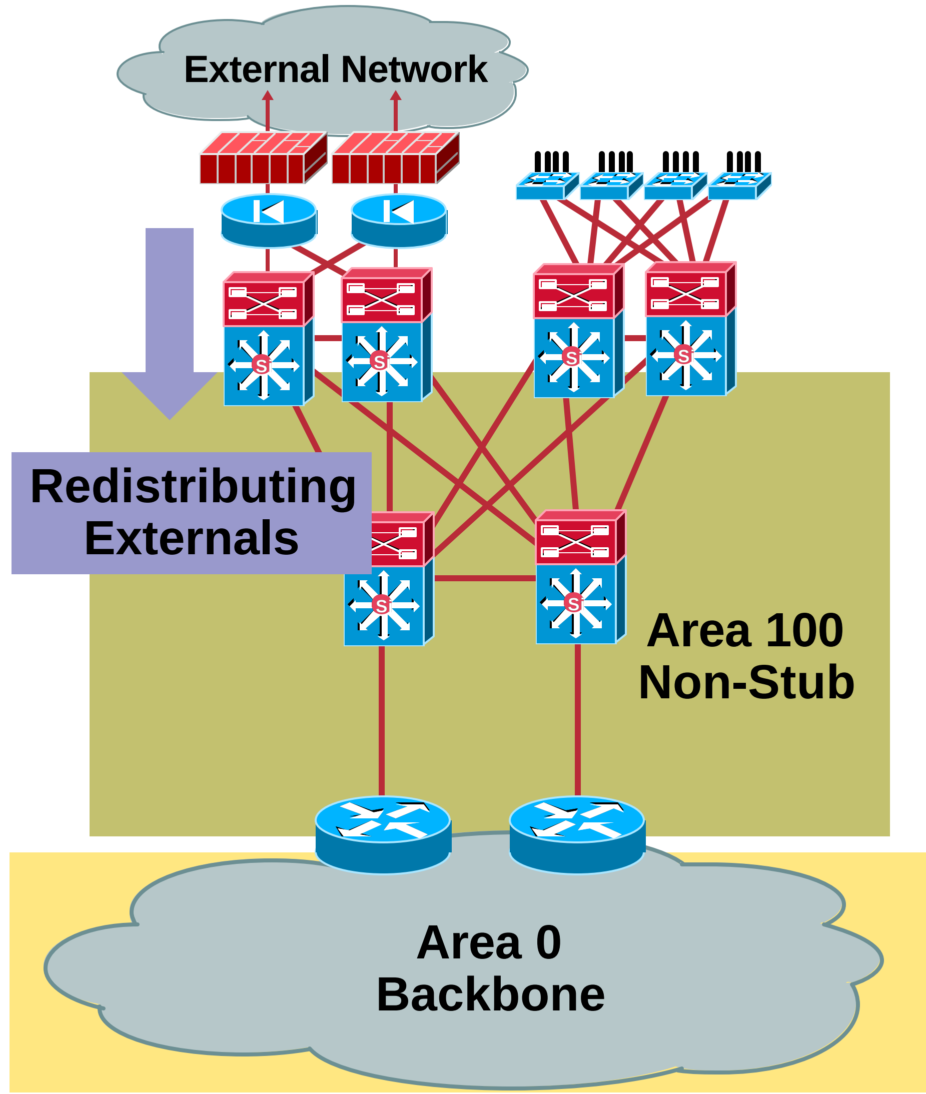


OSPF in the Campus

Conversion to an OSPF Routed Edge

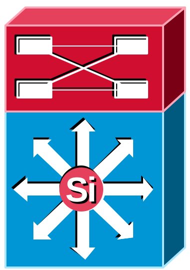
Cisco.com

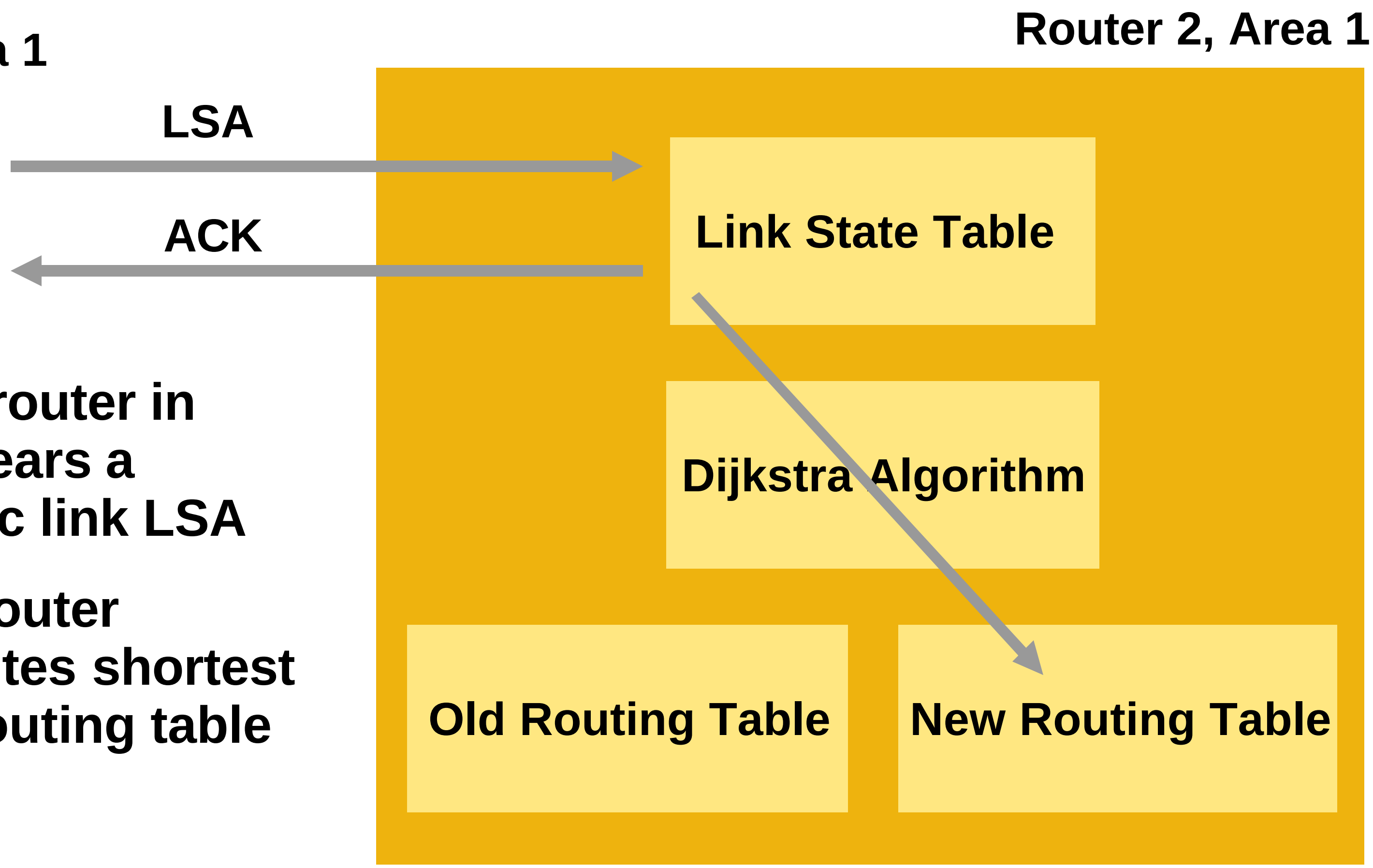
- **Other OSPF area designs may not permit an easy migration to a layer 3 access design**
- **Introduction of another network tier via BGP may be required**
- **Extension of area's beyond good design boundaries will result in loss of overall availability**



When a Link Changes State

Router 1, Area 1





- Every router in area hears a specific link LSA
- Each router computes shortest path routing table

Different Types of LSAs

- Router link (LSA type 1)
- Network link (LSA type 2)
- Network summary (LSA type 3)
- ASBR (LSA type 4)
- External (LSA type 5)
- NSSA external (LSA type 7)

Regular Area

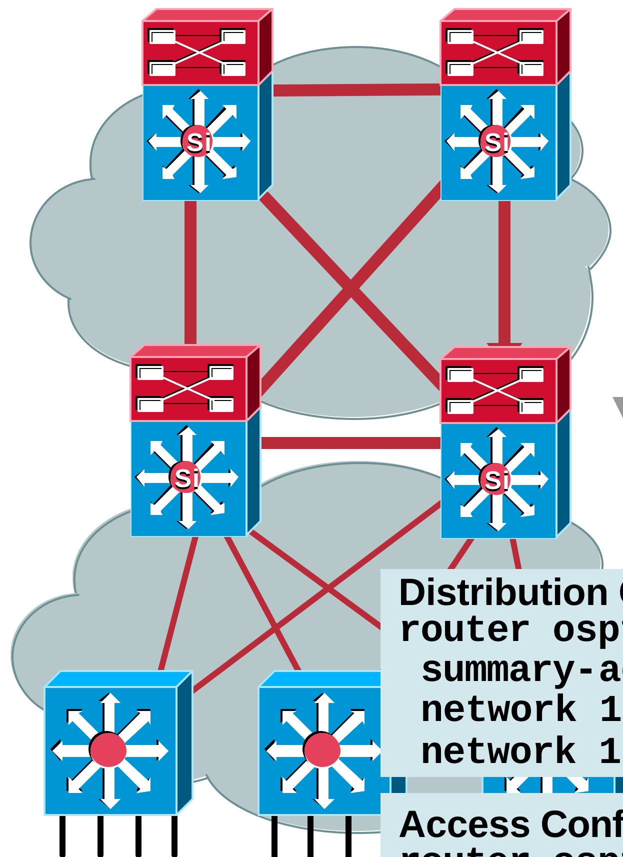
ABRs Forward All LSAs from Backbone

External Routes/LSA Present in Area 120

Backbone Area 0

Area Border Router

Area 120



An ABR Forwards the Following into an Area

- Summary LSAs (Type 3)
- ASBR Summary (Type 4)
- Specific Externals (Type 5)

Distribution Config

```
router ospf 100
 summary-address 10.120.0.0 255.255.0.0
 network 10.120.0.0 0.0.255.255 area 120
 network 10.122.0.0 0.0.255.255 area 0
```

Access Config:

```
router ospf 100
 network 10.120.0.0 0.0.255.255 area 120
```

Stub Area

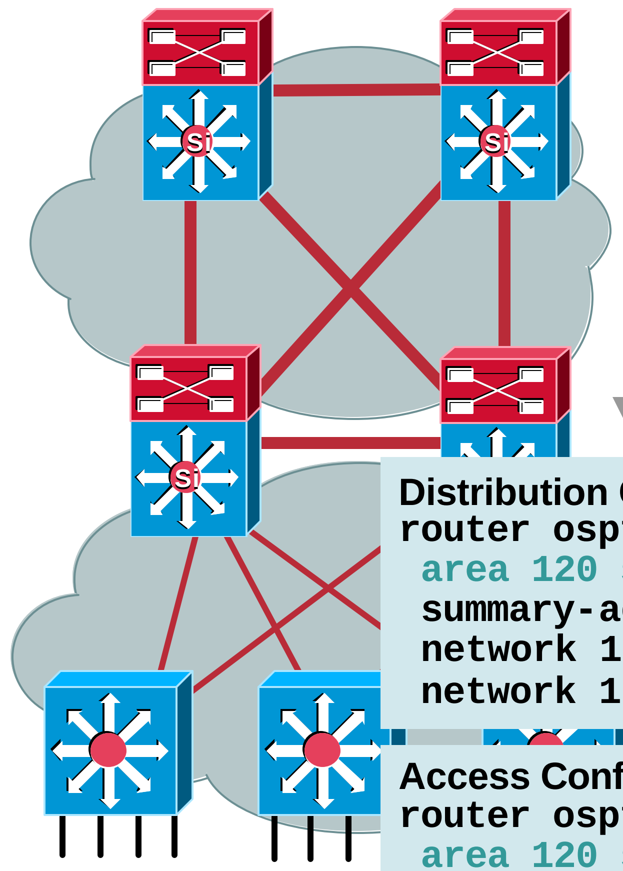
Consolidates Specific External Links—Default 0.0.0.0

Eliminates External Routes/LSA Present in Area (Type 5)

Backbone Area 0

Area Border Router

Area 120



Stub Area ABR Forwards

- Summary LSAs
- Summary Default to ABR

Distribution Config

```
router ospf 100
 area 120 stub
 summary-address 10.120.0.0 255.255.0.0
 network 10.120.0.0 0.0.255.255 area 120
 network 10.122.0.0 0.0.255.255 area 0
```

Access Config:

```
router ospf 100
 area 120 stub
 network 10.120.0.0 0.0.255.255 area 120
```

Totally Stubby Area
Use This for Stable—Scalable Internetworks

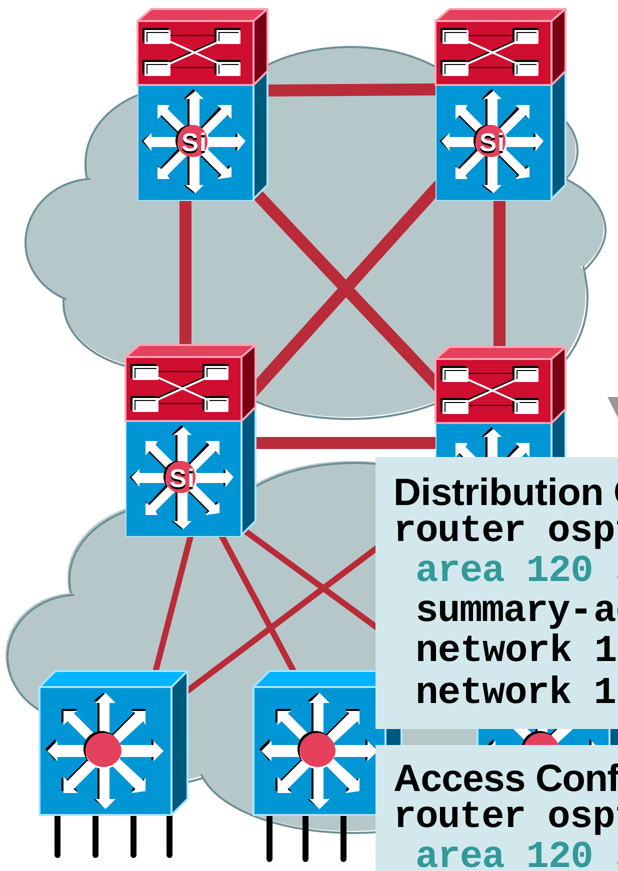
Cisco.com

Minimize the Number of LSA's and the Need for Any External Area SPF Calculations

Backbone Area 0

Area Border Router

Area 120



A Totally Stubby Area ABR Forwards

Summary Default to ABR

Distribution Config

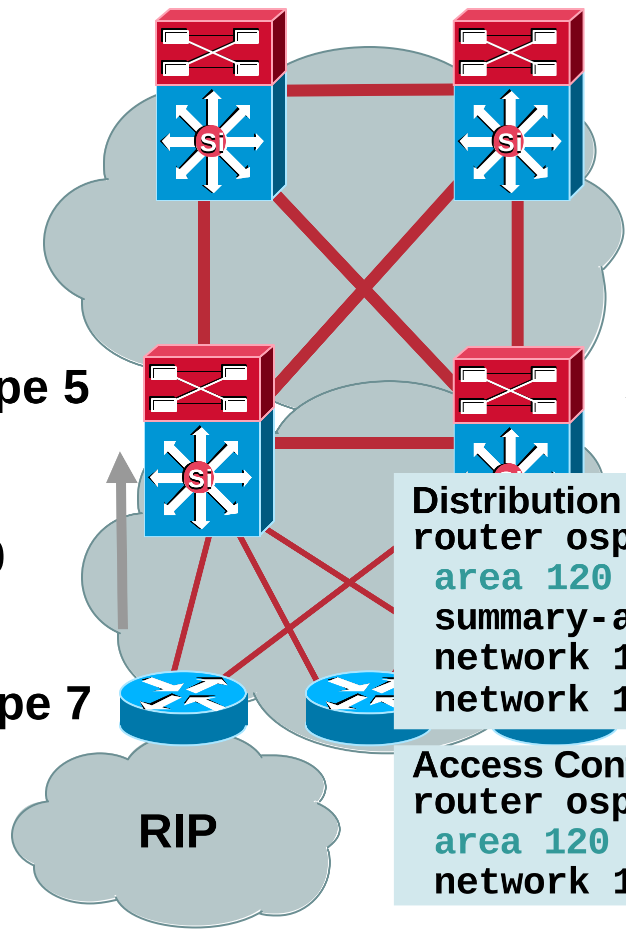
```
router ospf 100
 area 120 stub no-summary
 summary-address 10.120.0.0 255.255.0.0
 network 10.120.0.0 0.0.255.255 area 120
 network 10.122.0.0 0.0.255.255 area 0
```

Access Config:

```
router ospf 100
 area 120 stub no-summary
 network 10.120.0.0 0.0.255.255 area 120
```

Not So Stubby Area (NSSA)

Minimize the Number of LSA’s and the Need for Any External Area While Supporting External Connectivity



NSSA ABR Forwards

- Summary LSAs
- Summary 0.0.0.0 to ABR

Totally Stubby NSSA ABR Forwards

- Summary 0.0.0.0 to ABR

Distribution Config

```
router ospf 100
 area 120 nssa no-summary
 summary-address 10.120.0.0 255.255.0.0
 network 10.120.0.0 0.0.255.255 area 120
 network 10.122.0.0 0.0.255.255 area 0
```

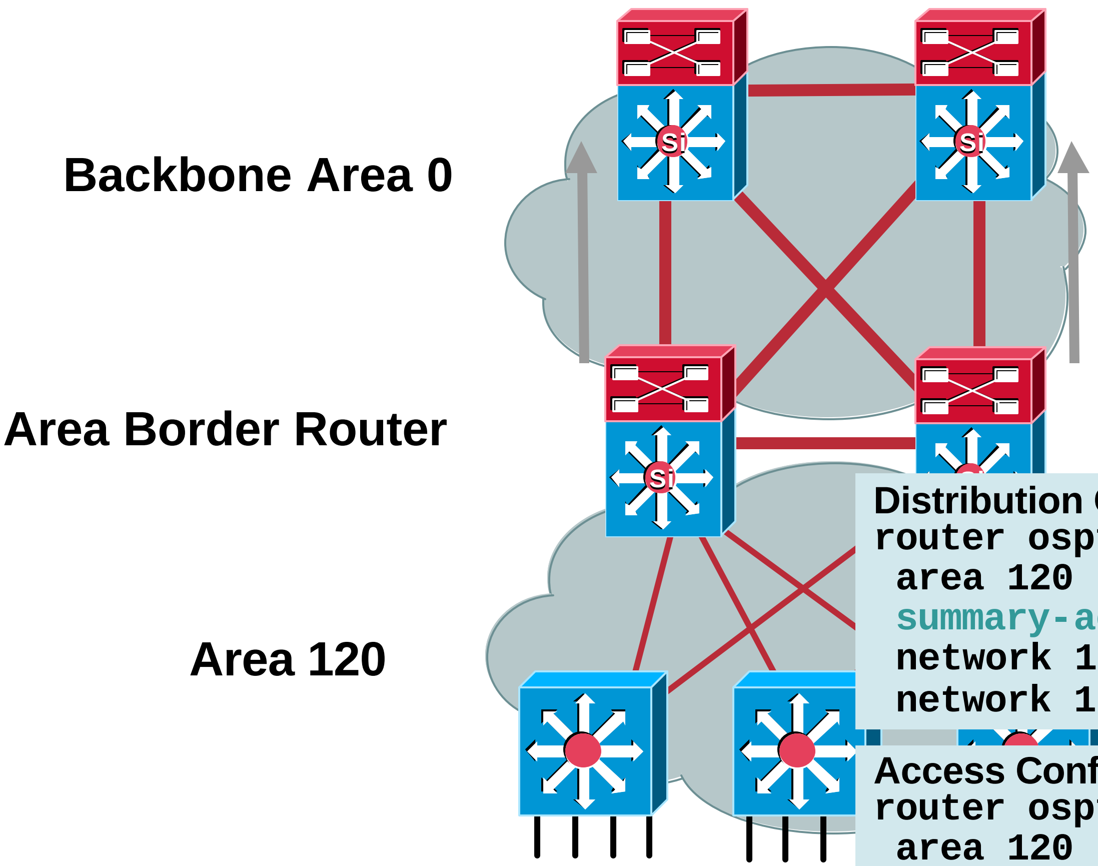
Access Config:

```
router ospf 100
 area 120 nssa no-summary
 network 10.120.0.0 0.0.255.255 area 120
```

Summarization Distribution to Core

Reduce SPF and LSA Load in Area 0

Minimize the Number of LSA's and the Need for Any SPF Recalculations at the Core



Distribution Config

```
router ospf 100
 area 120 stub no-summary
 summary-address 10.120.0.0 255.255.0.0
 network 10.120.0.0 0.0.255.255 area 120
 network 10.122.0.0 0.0.255.255 area 0
```

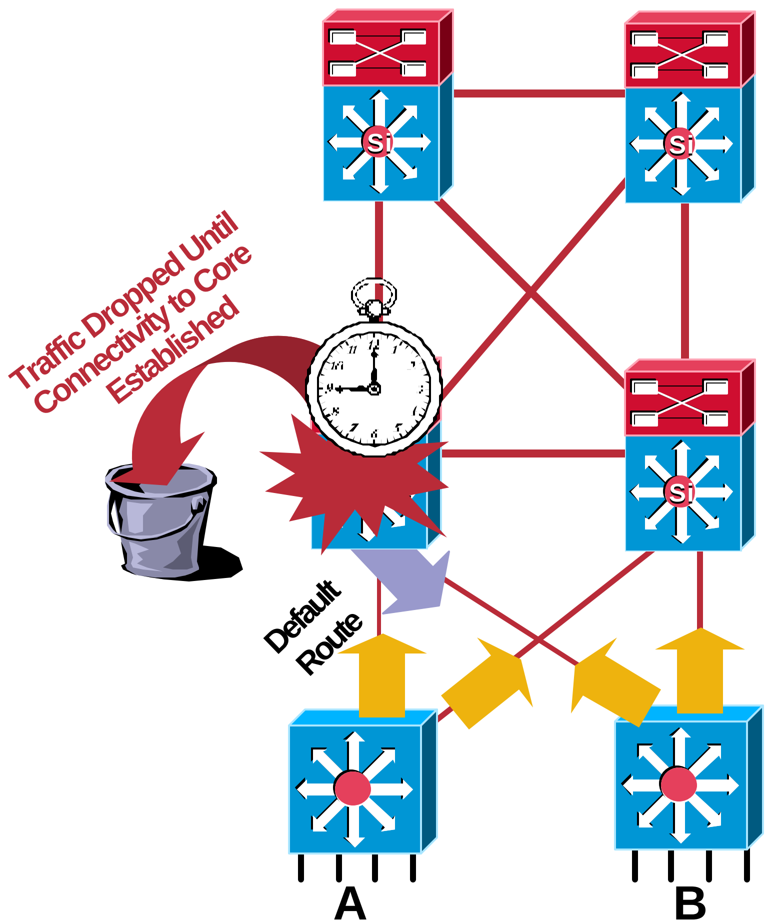
Access Config:

```
router ospf 100
 area 120 stub no-summary
 network 10.120.0.0 0.0.255.255 area 120
```

OSPF Default Route to Totally Stubby Area

Cisco.com

- Totally stubby area's are used to isolate the access layer switches from route calculations due to events in other areas
- This means that the ABR (the distribution switch) will send a default route to the access layer switch when the neighbor relationship is established
- The default route is sent regardless of the distribution switches ability to forward traffic on to the core (area 0)
- Traffic could be black holed until connectivity to the core is established

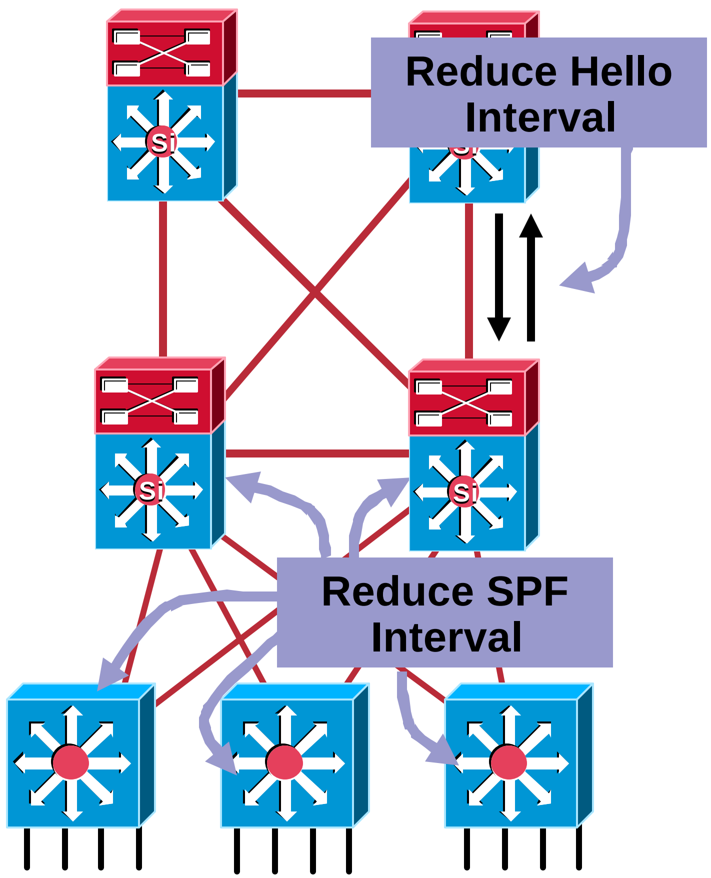


Note: Solution to this anomaly is being investigated.

OSPF Timer Tuning
High-Speed Campus Convergence

Cisco.com

- OSPF by design has a number of throttling mechanisms to prevent the network from thrashing during periods of instability
- Campus environments are candidates to utilize OSPF timer enhancements
    - Sub-second hellos
    - Generic IP (interface) dampening mechanism
    - Back-off algorithm for LSA generation
    - Exponential SPF backoff
    - Configurable packet pacing
    - Incremental SPF

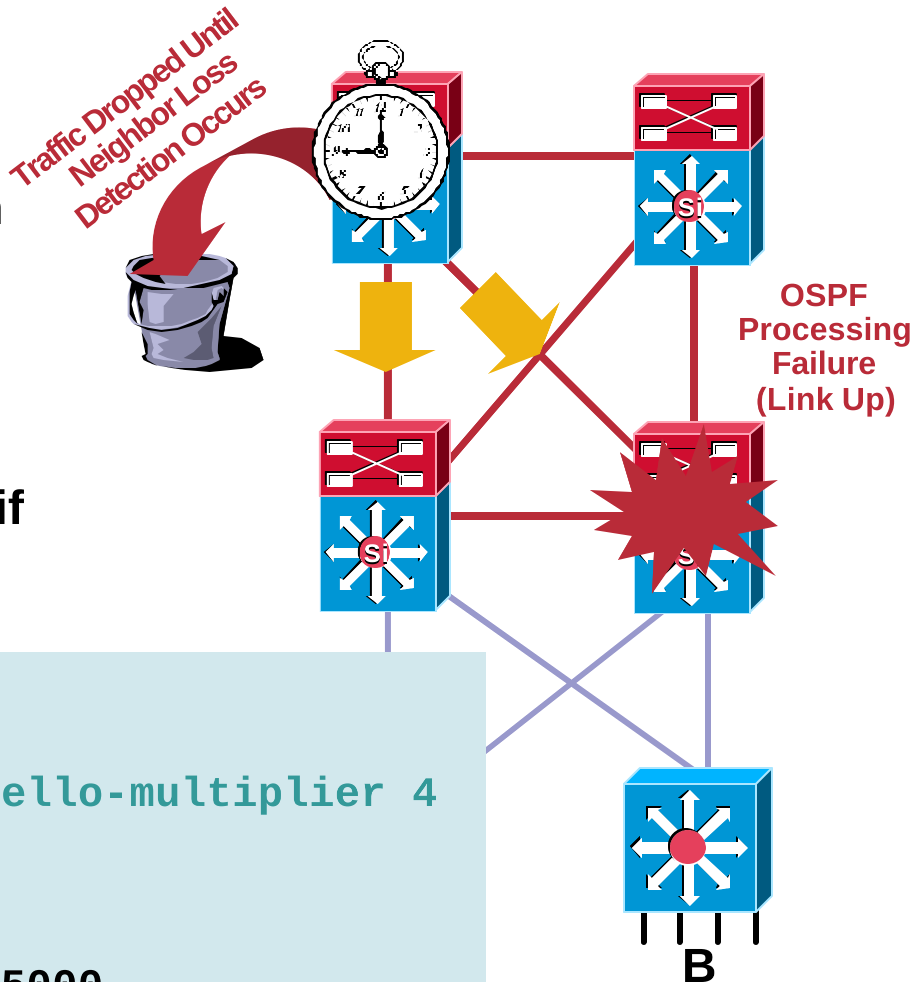


Subsecond Hello's

Neighbor Loss Detection—Physical Link Up

Cisco.com

- OSPF hello/dead timers detect neighbor loss in the absence of physical link loss
- Useful in environments where an L2 device separates L3 devices (Layer 2 core designs)
- Aggressive timers are needed to quickly detect neighbor failure
- Interface dampening is required if sub-second hello timers are implemented



Access Config:

```
interface GigabitEthernet1/1
 dampening
 ip ospf dead-interval minimal hello-multiplier 4

router ospf 100
 area 120 stub no-summary
 timers throttle spf 10 100 5000
 timers throttle lsa all 10 100 5000
 timers lsa arrival 80
```

OSPF LSA Throttling

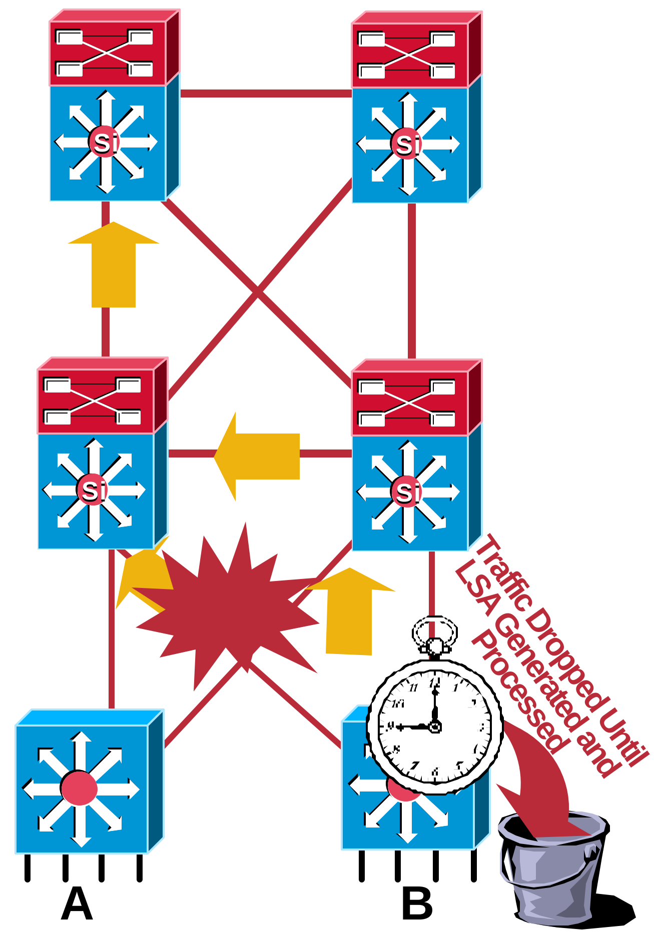
- By default, there is a 500ms delay before generating router and network LSA's; the wait is used to collect changes during a convergence event and minimize the number of LSA's sent
- Propagation of a new instance of the LSA is limited at the originator

```
timers throttle lsa all <start-interval>
      <hold-interval> <max-interval>
```

- Acceptance of a new LSAs is limited by the receiver

```
timers lsa arrival <milliseconds>
```

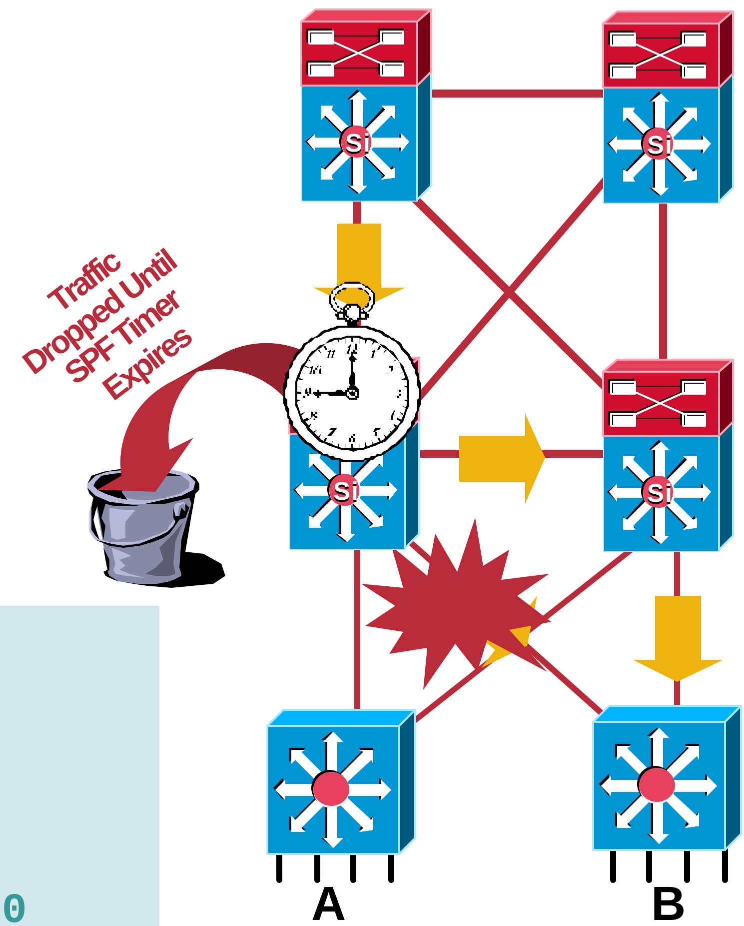
```
Access Config:
interface GigabitEthernet1/1
 ip ospf dead-interval minimal
 hello-multiplier 4

router ospf 100
 area 120 stub no-summary
 timers throttle spf 10 100 5000
 timers throttle lsa all 10 100 5000
 timers lsa arrival 80
```



OSPF SPF Throttling

Cisco.com

- OSPF has an SPF throttling timer designed to dampen route recalculation (preserving CPU resources) when a link bounces
- 12.2S OSPF enhancements let us tune this timer to milliseconds; prior to 12.2S one second was the minimum
- After a failure, the router waits for the SPF timer to expire before recalculating a new route; SPF timer was one second



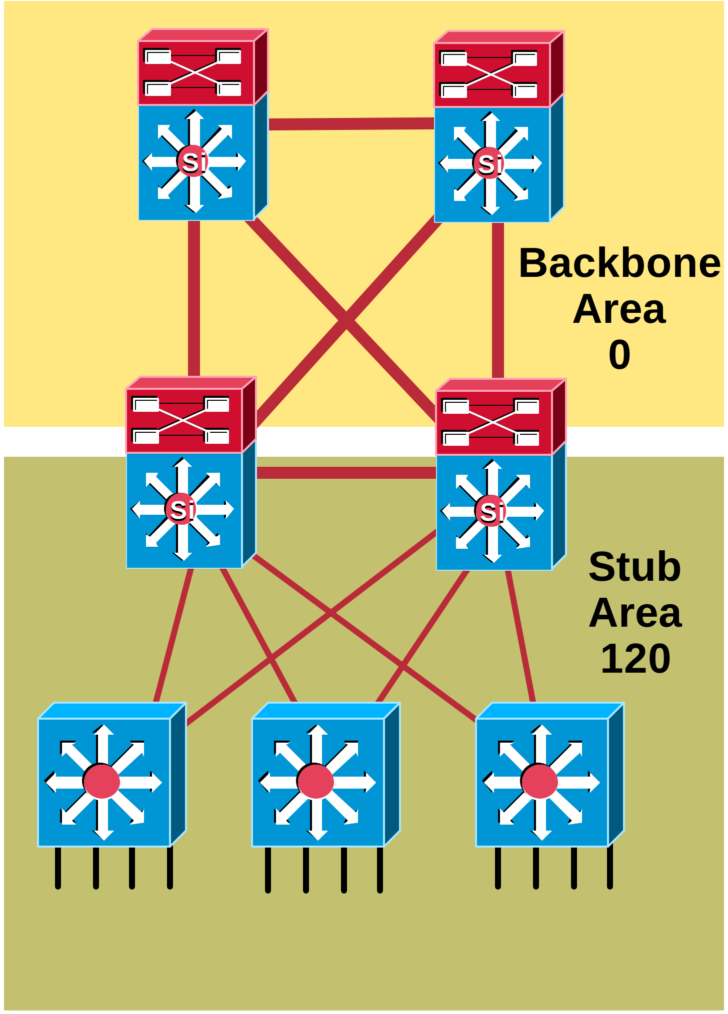
Access Config:

```
interface GigabitEthernet1/1
 ip ospf dead-interval minimal
 hello-multiplier 4

router ospf 100
 area 120 stub no-summary
 timers throttle spf 10 100 5000
 timers throttle lsa all 10 100 5000
 timers lsa arrival 80
```

OSPF Routed Access Campus Design

Overview—Fast Convergence

- **Detect the event:**
  - Decrease the hello-interval and dead-interval to detect soft neighbor failures
  - Enable interface dampening
  - Set carrier-delay = 0
- **Propagate the event:**
  - Summarize routes between areas to limit LSA propagation across the campus
  - Tune LSA timers to minimize LSA propagation delay
- **Process the event:**
  - Tune SPF throttles to decrease calculation delays



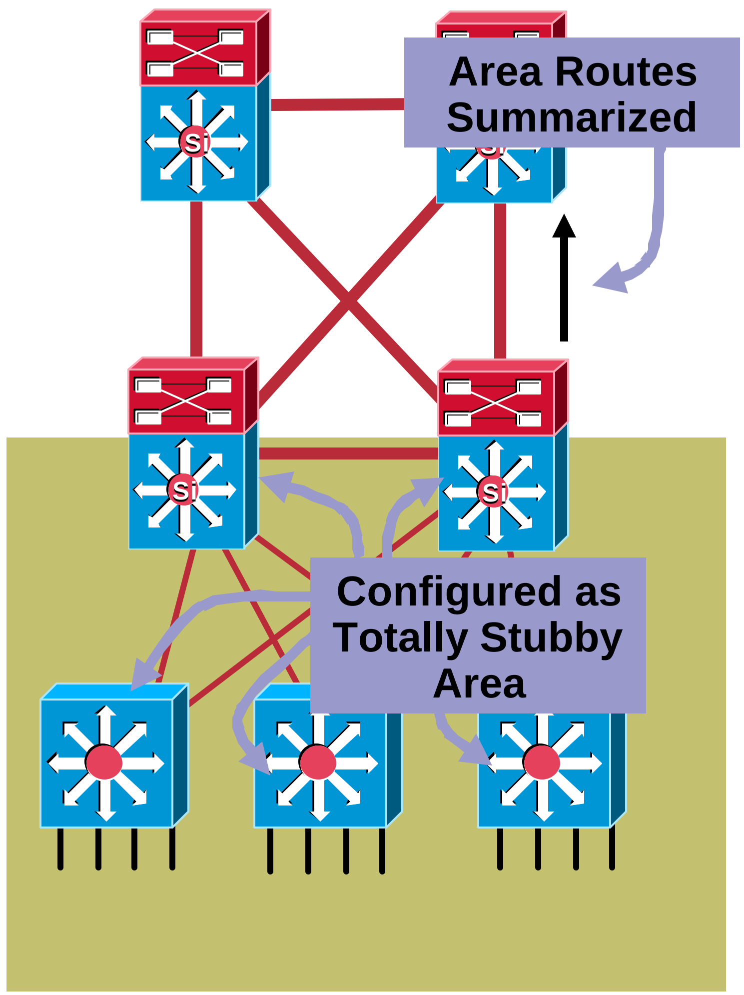
OSPF Routed Access Campus Design

Overview—Area Design

Cisco.com

- Use totally stubby areas to minimize routes in Access switches
- Summarize area routes to backbone Area 0
- These recommendations will reduce number of LSAs and SPF recalculations throughout the network and provide a more robust and scalable network infrastructure

```
router ospf 100
 area 120 stub no-summary
 summary-address 10.120.0.0 255.255.0.0
 network 10.120.0.0 0.0.255.255 area 120
 network 10.122.0.0 0.0.255.255 area 0

router ospf 100
 area 120 stub no-summary
 network 10.120.0.0 0.0.255.255 area 120
```



OSPF Routed Access Campus Design

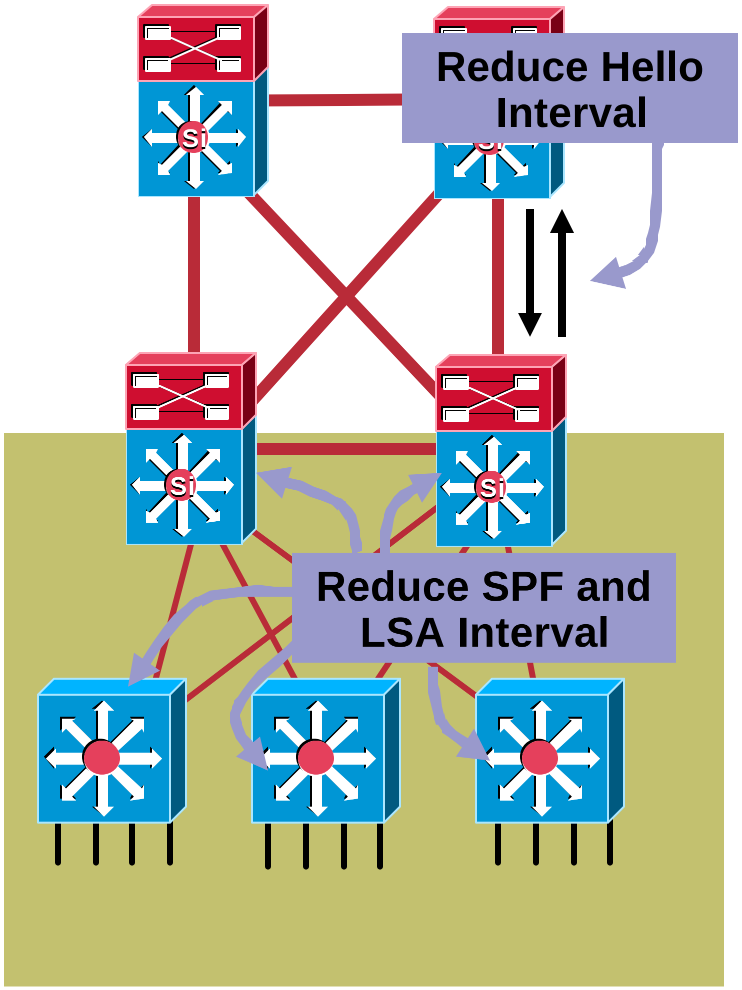
Overview—Timer Tuning

Cisco.com

- **In a hierarchical design, the key tuning parameters are spf throttle and lsa throttle**
- **Need to understand other LSA tuning in the non-optimal design**
- **Hello and dead timers are secondary failure detection mechanism**

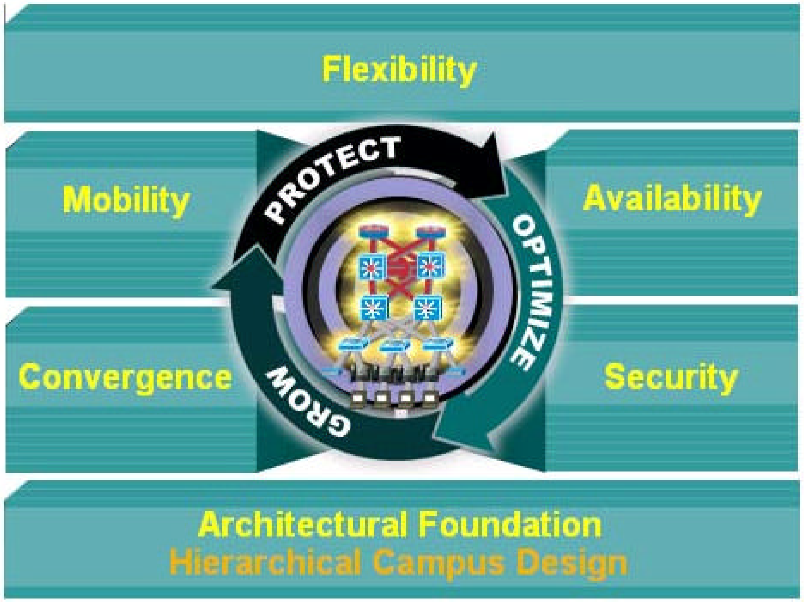
```
router ospf 100
 area 120 stub no-summary
 area 120 range 10.120.0.0 255.255.0.0
 timers throttle spf 10 100 5000
 timers throttle lsa all 10 100 5000
 timers lsa arrival 80
 network 10.120.0.0 0.0.255.255 area 120
 network 10.122.0.0 0.0.255.255 area 0
```

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



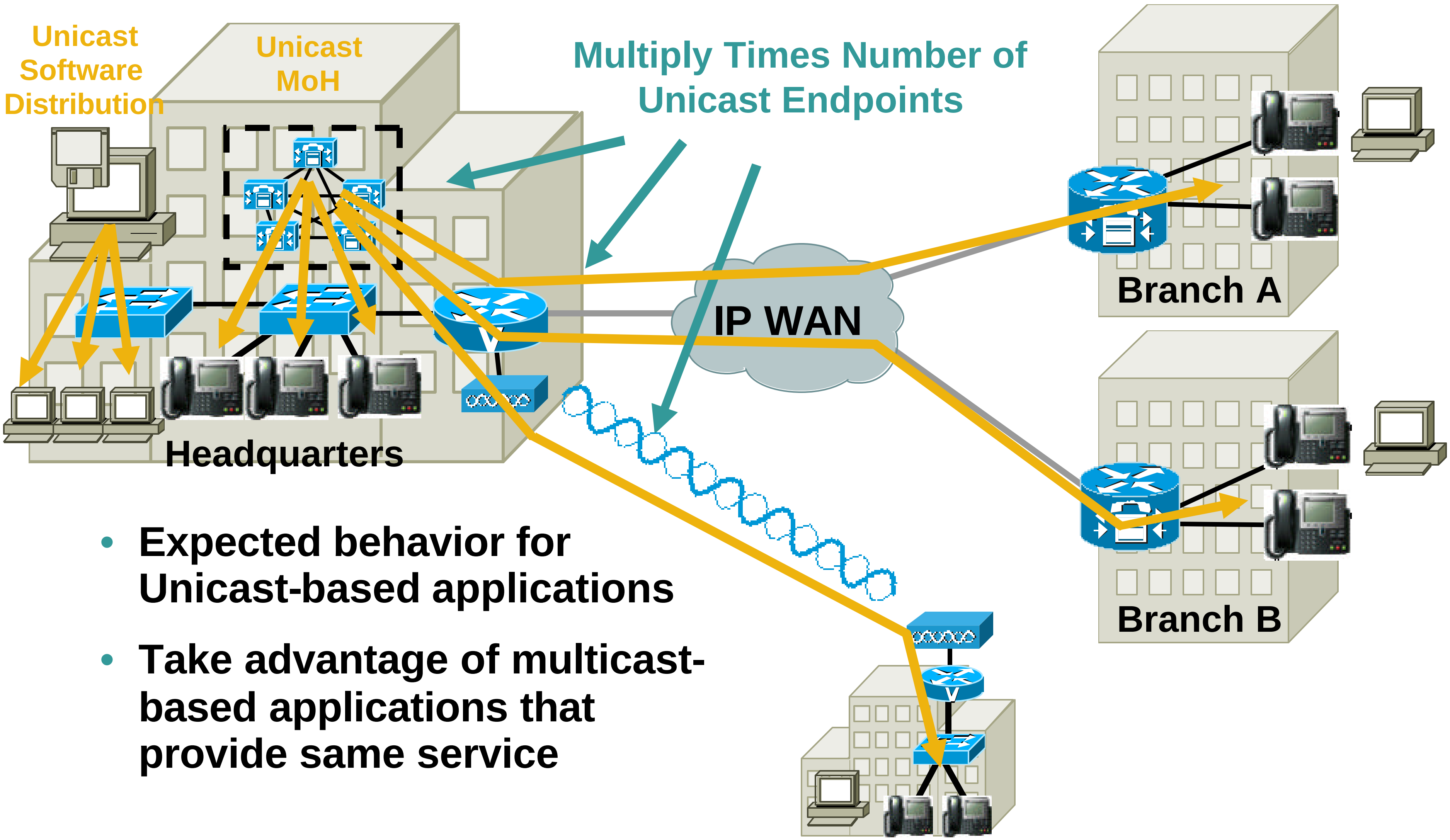
Agenda

- Campus Network Designs
- Routed Access Design
- EIGRP Design Details
- OSPF Design Details
- PIM Design Details
- Summary



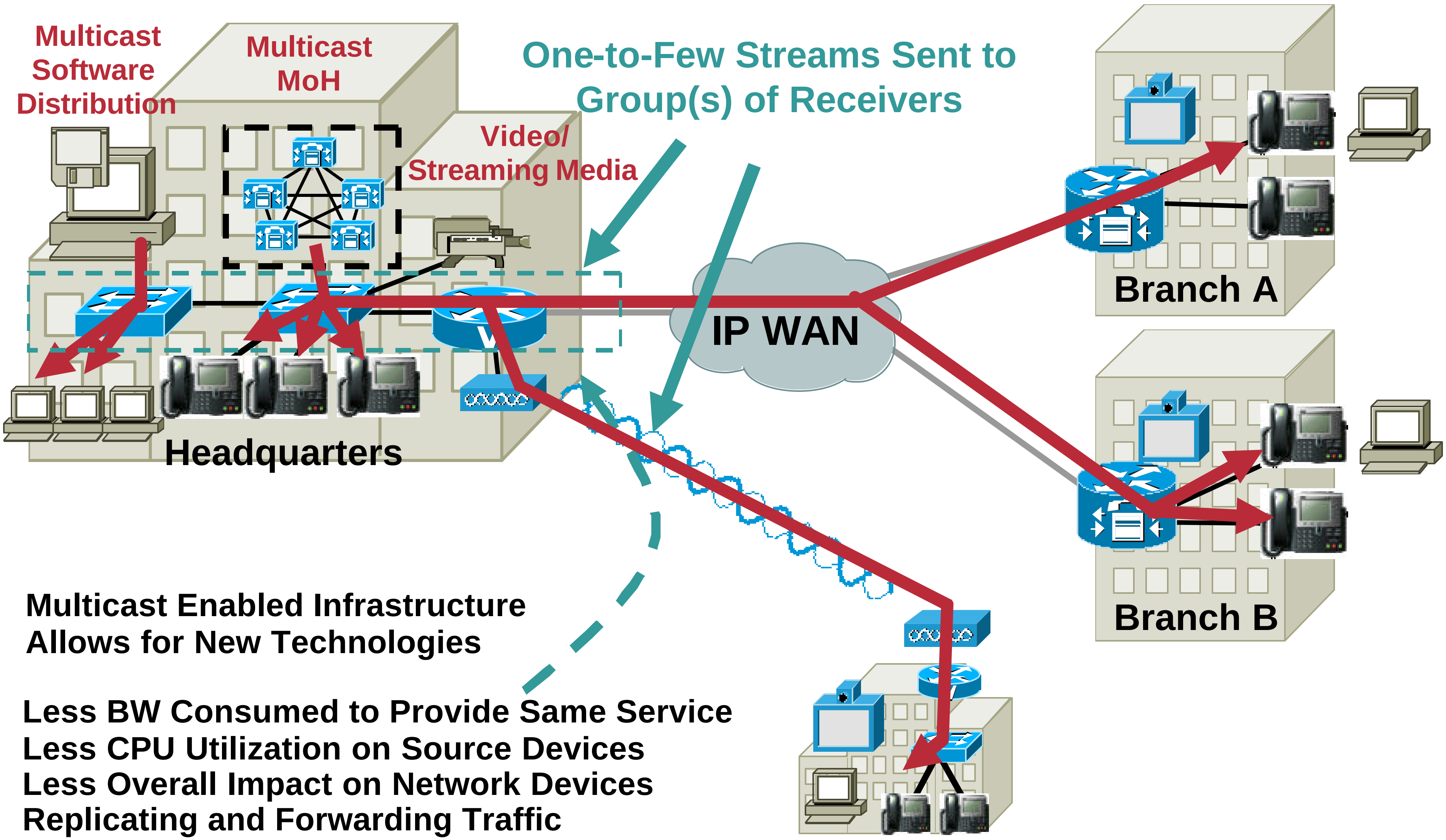
Unicast vs. Multicast

Cisco.com



- Expected behavior for Unicast-based applications
- Take advantage of multicast-based applications that provide same service

Unicast vs. Multicast



Multicast Enabled Infrastructure
Allows for New Technologies

Less BW Consumed to Provide Same Service
Less CPU Utilization on Source Devices
Less Overall Impact on Network Devices
Replicating and Forwarding Traffic

IP Multicast Protocols

- **Dense-mode protocols**
  - Uses "push" model
  - Flood and prune behavior
- **Sparse-mode protocols**
  - Uses "pull" model: traffic sent only to where it is requested
  - Explicit join behavior
- **Enterprise IPmc protocols**
  - PIM, MOSPF, DVMRP,
- **PIM—Protocol independent multicast**
  - Uses underlying Unicast routing protocol to prevent loops
  - Two modes: PIM dense mode and PIM sparse mode

Which PIM Mode—Sparse or Dense

"Sparse mode Good! Dense mode Bad!"

Source: ***"The Caveman's Guide to IP Multicast", ©2000, R. Davis***

PIM Sparse Mode (RFC 2362)

- **Assumes no hosts wants multicast traffic unless they specifically ask for it**
- **Uses a Rendezvous Point (RP)**
  - **Senders and receivers "rendezvous" at this point to learn of each others existence**
    - **Senders are "registered" with RP by their first-hop router**
    - **Receivers are "joined" to the shared tree (rooted at the RP) by their local Designated Router (DR)**
- **Appropriate for…**
  - **Wide scale deployment for both densely and sparsely populated groups in the enterprise**
  - **Optimal choice for all production networks regardless of size and membership density**

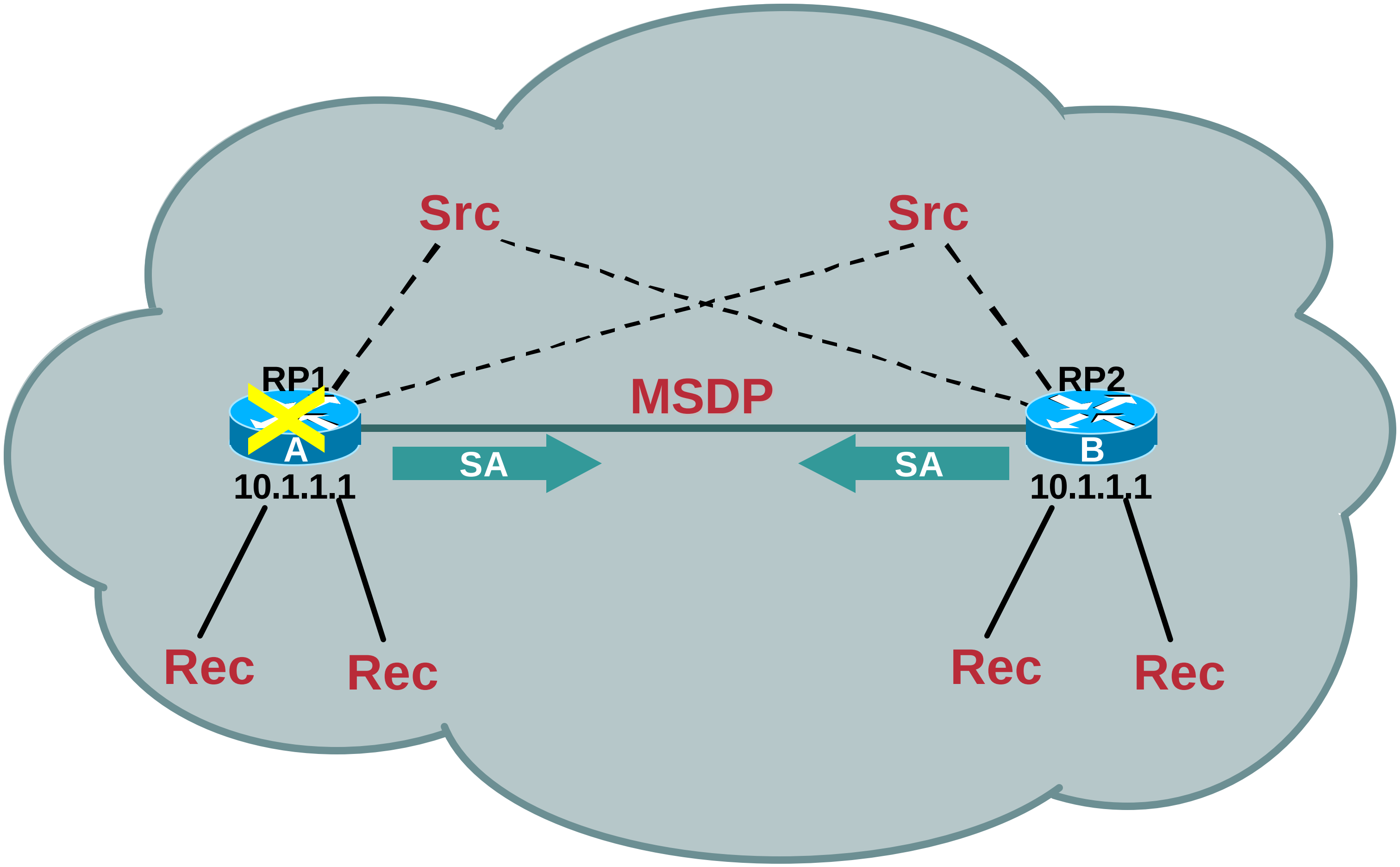
Anycast RP—Overview

Cisco.com

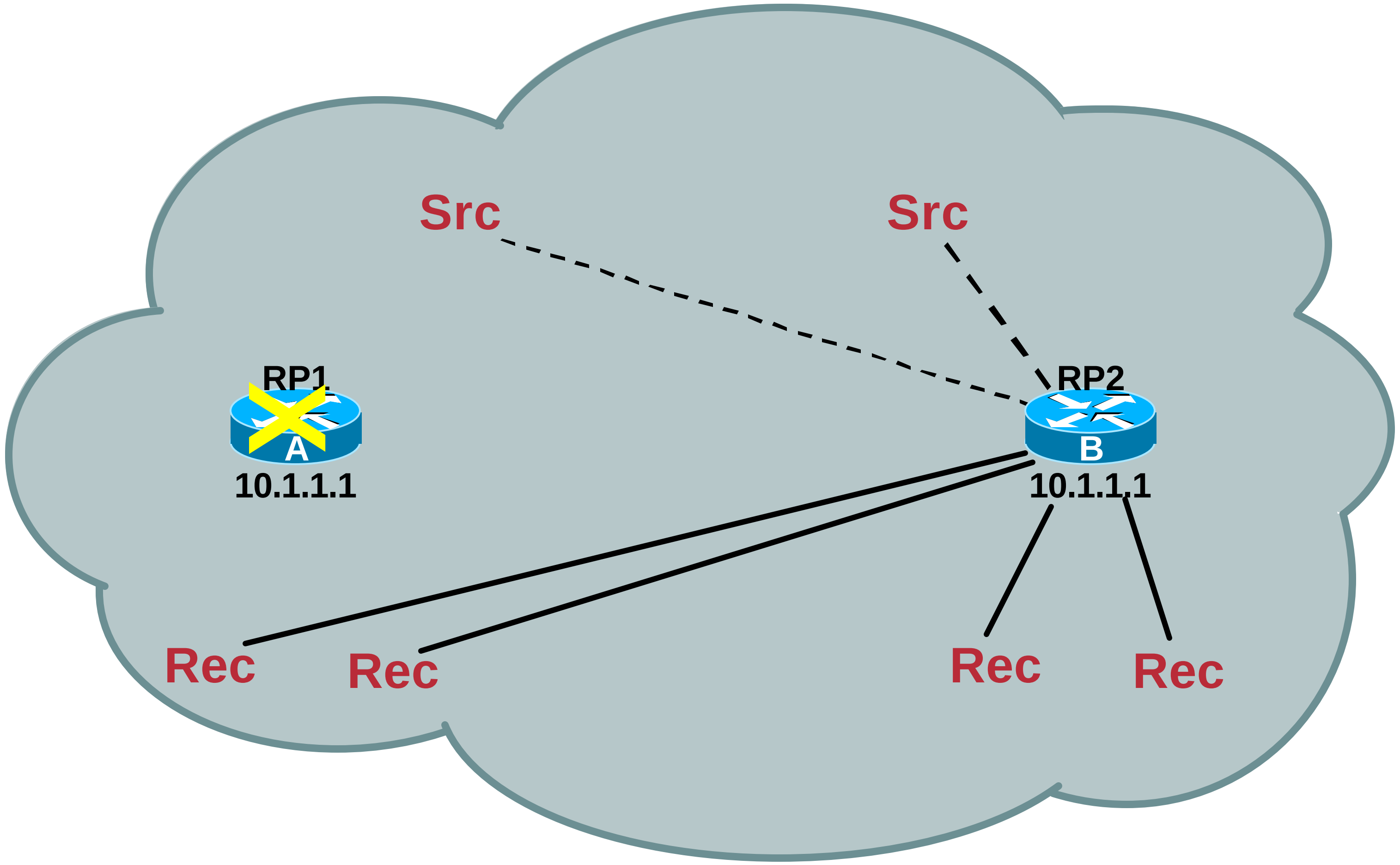
- **PIM RP deployment options**
  - Static, Auto-RP, BSR, and Anycast RP
- **Anycast RP provides fast failover and load-balancing**
  - Multiple RPs use a single IP address
    - Two or more routers have same RP address (anycast)
      - RP address defined as a Loopback Interface
    - Senders and receivers register/join with closest RP
      - Closest RP determined from the Unicast routing table
  - MSDP session(s) run between all RPs
    - Informs RPs of sources in other parts of network
    - Facilitates sharing of source information

Anycast RP—Overview

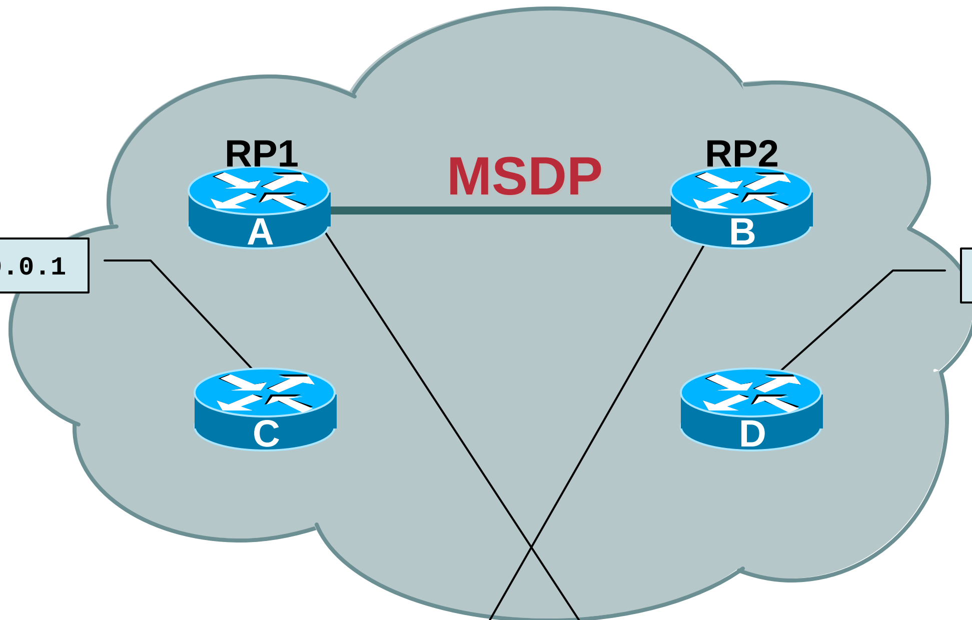
Cisco.com



Anycast RP—Overview



Anycast RP Configuration




```
ip pim rp-address 10.0.0.1
```

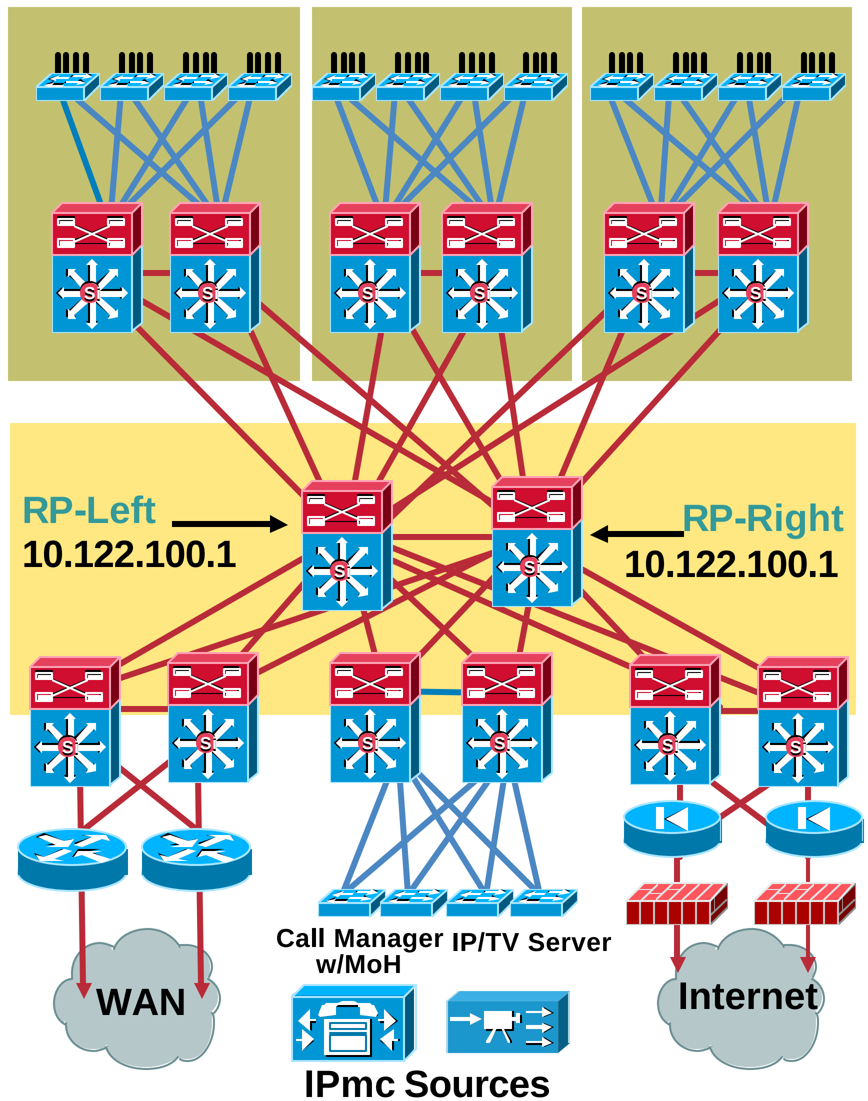
```
ip pim rp-address 10.0.0.1
```

```
Interface loopback 0
   ip address 10.0.0.1 255.255.255.255

Interface loopback 1
 ip address 10.0.0.2 255.255.255.255
!
ip msdp peer 10.0.0.3 connect-source loopback 1
ip msdp originator-id loopback 1
```

```
Interface loopback 0
 ip address 10.0.0.1 255.255.255.255

Interface loopback 1
 ip address 10.0.0.3 255.255.255.255
!
ip msdp peer 10.0.0.2 connect-source loopback 1
ip msdp originator-id loopback 1
```

PIM Design Rules for Routed Campus

- Use PIM sparse mode
- Enable PIM sparse mode on ALL access, distribution and core layer switches
- Enable PIM on ALL interfaces
- Use Anycast RPs in the core for RP redundancy and fast convergence
- IGMP-snooping is enabled when PIM is enabled on a VLAN interface (SVI)
- (Optional) force the multicast traffic to remain on the shared-tree to reduce (S, G) state
- (Optional) use garbage can RP to black-hole unassigned IPmc traffic



Multicast Routed Access Campus Design

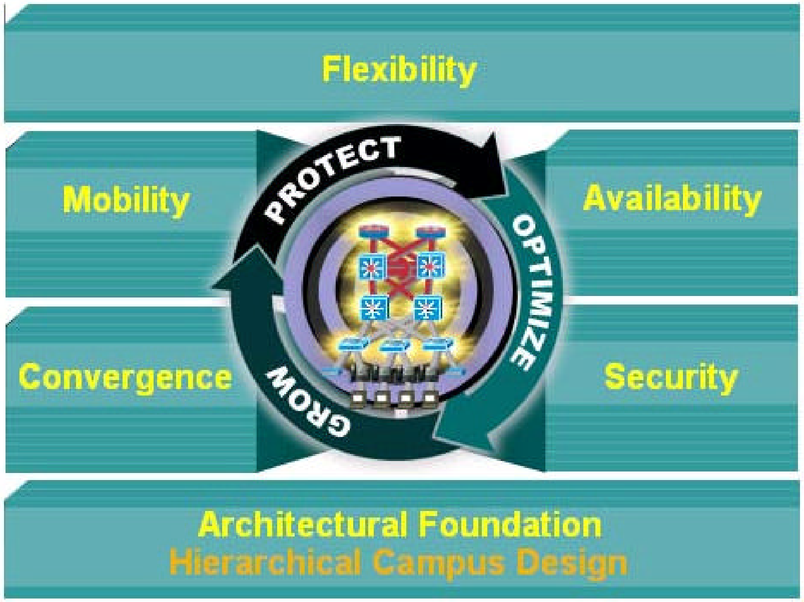
Things You Don't Have to Do…

- Tune PIM query interval for designated router convergence
- Configure designated router to match HSRP primary
- Configure PIM snooping on L2 switches between L3 switches
- Worry about all those L2/L3 flow inconsistency issues

Agenda

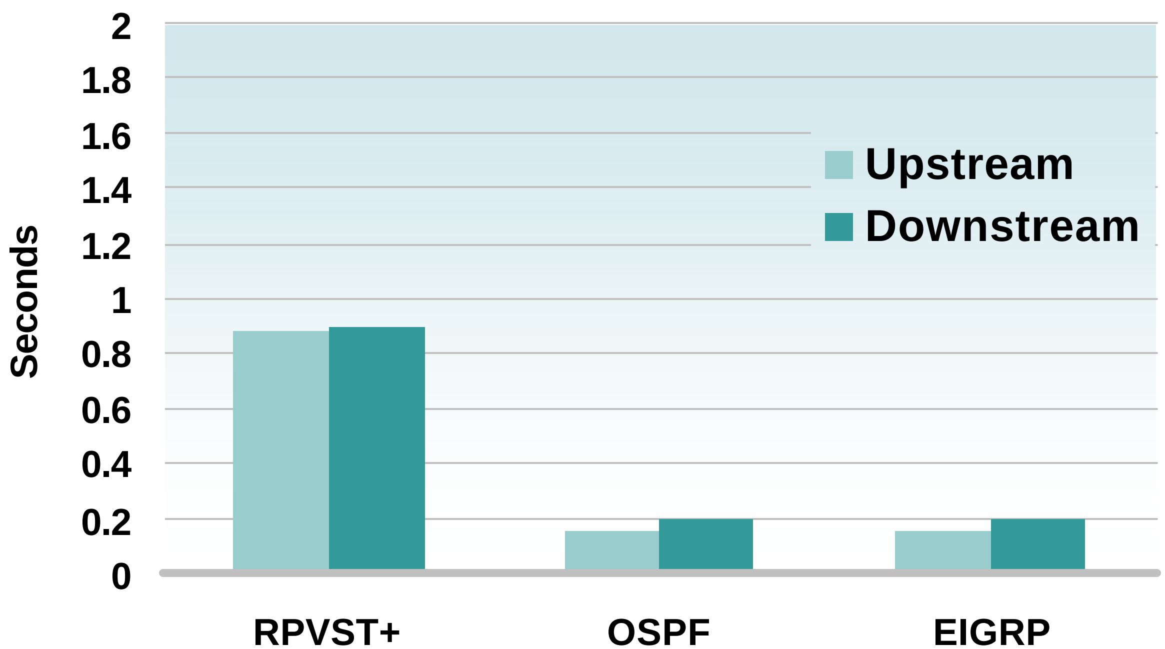
Cisco.com

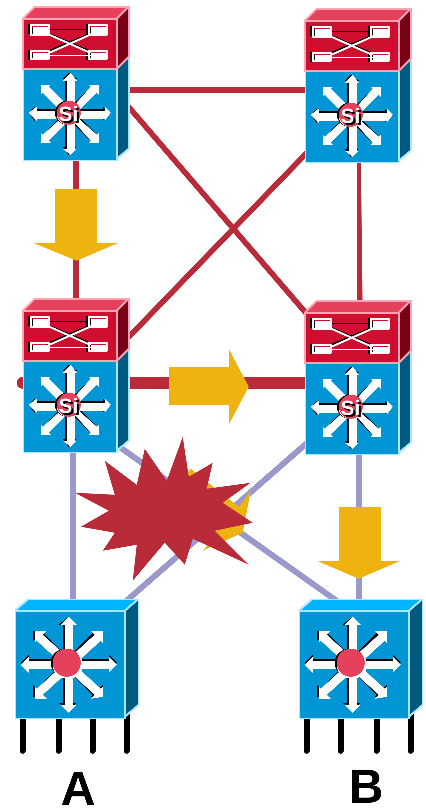
- Campus Network Designs
- Routed Access Design
- EIGRP Design Details
- OSPF Design Details
- PIM Design Details
- Summary



Routing to the Edge

Advantages? Yes, with a Good Design



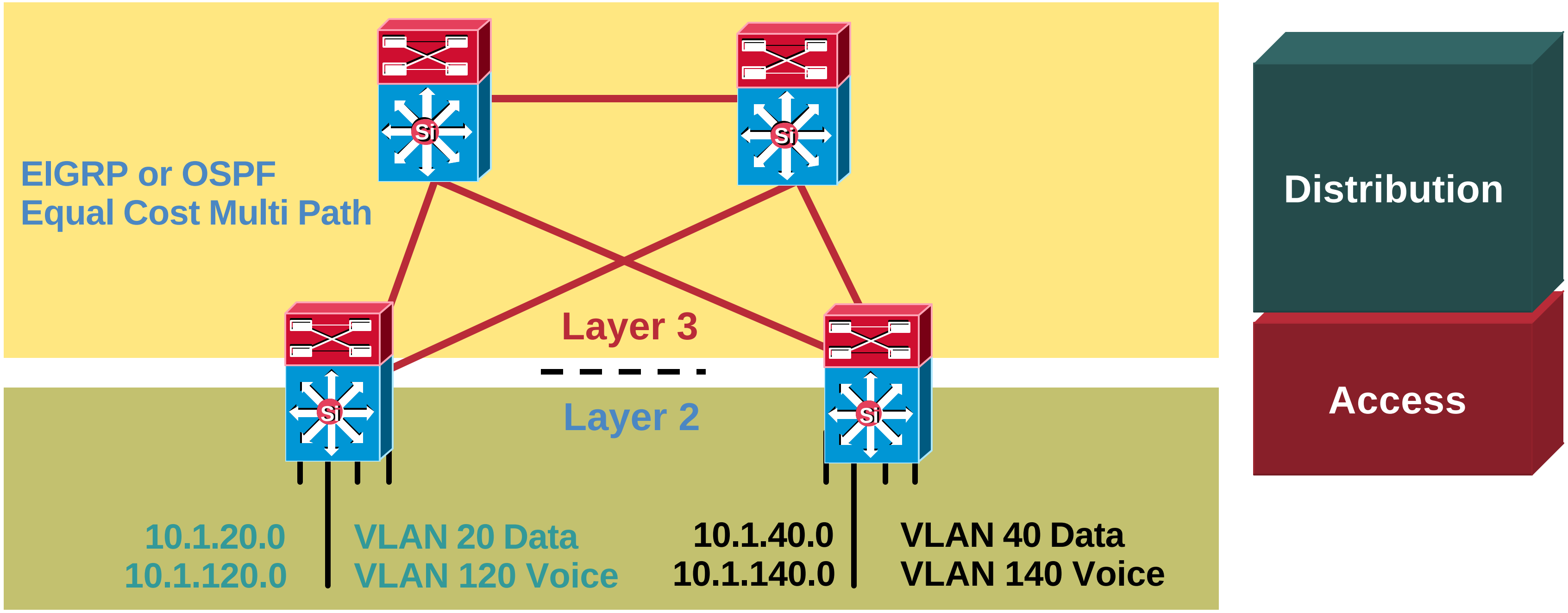


- Sub-200 msec convergence for EIGRP and OSPF
- Ease of implementation; fewer things to get right
- Troubleshooting; well known protocols and tools
- Simplified IP Multicast deployment
- Considerations; spanning VLANs, IP addressing, IGP selection

Routed Access Design

Summary

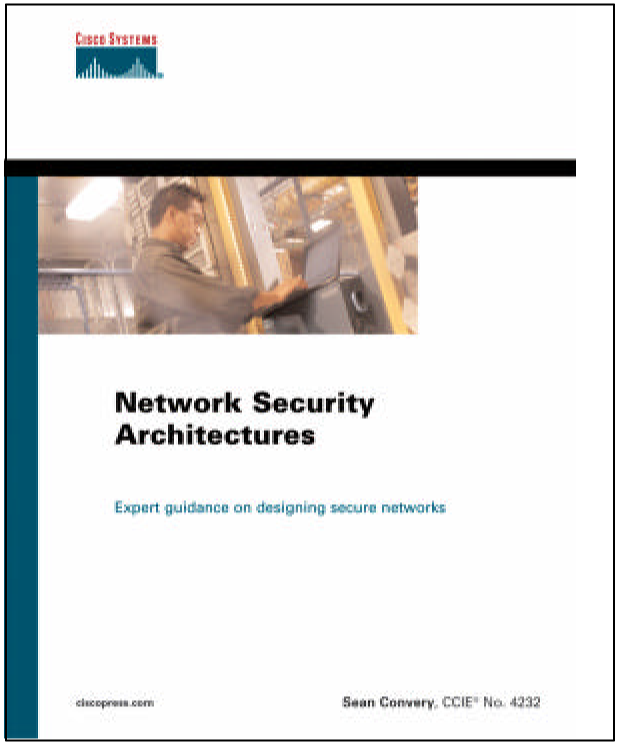
Cisco.com



- EIGRP or OSPF routed links between access and distribution
- Routed interfaces, not VLAN trunks, between switches
- Equal cost multi path to load balance traffic across network
- Route summarization at distribution with stub routers/areas
- Single (IGP) control plan to configure/manage/troubleshoot

Recommended Reading

- Continue your Networkers learning experience with further reading for this session from Cisco Press
- Check the Recommended Reading flyer for suggested books



Available Onsite at the Cisco Company Store

EIGRP
Core Layer Configuration

Cisco.com

6k-core configuration	
interface TenGigabitEthernet3/1 description 10GigE to Distribution 1 ip address 10.122.0.29 255.255.255.252 ip pim sparse-mode ip hello-interval eigrp 100 1 ip hold-time eigrp 100 ip authentication mode eigrp 100 md5 ip authentication key-chain eigrp 100 eigrp carrier-delay msec 0 mls qos trust dscp ! interface TenGigabitEthernet3/2 description 10GigE to Distribution 2 ip address 10.122.0.37 255.255.255.252 ip pim sparse-mode ip authentication mode eigrp 100 md5 ip authentication key-chain eigrp 100 eigrp carrier-delay msec 0 mls qos trust dscp	! router eigrp 100 network 10.0.0.0 no auto-summary

EIGRP
Distribution Layer Configuration

Cisco.com

6k-distribution configuration	

interface GigabitEthernet3/2
  description typical link to Access neighbor
  ip address 10.120.0.50 255.255.255.252
  ip pim sparse-mode
  ip hello-interval eigrp 100 1
  ip hold-time eigrp 100 3
  ip authentication mode eigrp 100 md5
  ip authentication key-chain eigrp 100 eigrp
  carrier-delay msec 0
  mls qos trust dscp
!
interface TenGigabitEthernet4/3
  description 10GigE to Distribution neighbor
  ip address 10.120.0.22 255.255.255.252
  ip pim sparse-mode
  ip hello-interval eigrp 100 1
  ip hold-time eigrp 100 3
  ip authentication mode eigrp 100 md5
  ip authentication key-chain eigrp 100 eigrp
  mls qos trust dscp
	interface TenGigabitEthernet4/2
	  description 10 GigE to Core neighbor
	  ip address 10.122.0.38 255.255.255.252
	  ip pim sparse-mode
	  ip authentication mode eigrp 100 md5
	  ip authentication key-chain eigrp 100 eigrp
	  ip summary-address eigrp 100 10.120.0.0 255.255.0.0 5
	  mls qos trust dscp
	!
	router eigrp 100
	  network 10.0.0.0
	  distribute-list Default out GigabitEthernet3/1
	  distribute-list Default out GigabitEthernet3/2
	  …
	  distribute-list Default out GigabitEthernet9/15
	  no auto-summary
	!
	ip access-list standard Default
	  permit 0.0.0.0
	  permit 10.0.0.0

EIGRP
Access Layer Configuration

Catalyst 4507 configuration

```
interface GigabitEthernet2/1
 description cr3-6500-2 Distribution
 no switchport
 ip address 10.120.0.53 255.255.255.252
 ip hello-interval eigrp 100 1
 ip hold-time eigrp 100 3
 ip authentication mode eigrp 100 md5
 ip authentication key-chain eigrp 100 eigrp
 ip pim sparse-mode
 carrier-delay msec 0
 qos trust dscp
 tx-queue 3
   priority high
!
interface FastEthernet3/5
 description Host port w/ IP Phone
 switchport access vlan 4
 switchport mode access
 switchport voice vlan 104
 qos trust cos
 tx-queue 3
   priority high
 spanning-tree portfast
 spanning-tree bpduguard enable
```

```
interface Vlan4
 ip address 10.120.4.1 255.255.255.0
 ip helper-address 10.121.0.5
 no ip redirects
 ip pim sparse-mode
 ip igmp snooping fast-leave
!
interface Vlan104
 ip address 10.120.104.1 255.255.255.0
 ip helper-address 10.121.0.5
 no ip redirects
 ip pim sparse-mode
 ip igmp snooping fast-leave
!
router eigrp 100
 passive-interface default
 no passive-interface GigabitEthernet1/1
 no passive-interface GigabitEthernet2/1
 network 10.0.0.0
 no auto-summary
 eigrp stub connected
```

OSPF
Core Layer Configuration

Cisco.com

6k-core configuration	

interface Port-channel1
  description Channel to Peer Core node
  dampening
  ip address 10.122.0.19 255.255.255.254
  ip pim sparse-mode
  ip ospf dead-interval minimal hello-multip 4
  load-interval 30
  carrier-delay msec 0
  mls qos trust dscp
!
interface TenGigabitEthernet3/1
  description 10GigE to Distribution 1
  dampening
  ip address 10.122.0.20 255.255.255.254
  ip pim sparse-mode
  ip ospf dead-interval minimal hello-multip 4
  load-interval 30
  carrier-delay msec 0
  mls qos trust dscp
!
	router ospf 100
	  router-id 10.122.10.2
	  log-adjacency-changes
	  timers throttle spf 10 100 5000
	  timers throttle lsa all 10 100 5000
	  timers lsa arrival 80
	  passive-interface Loopback0
	  passive-interface Loopback1
	  passive-interface Loopback2
	  network 10.122.0.0 0.0.255.255 area 0.0.0.0
	!

OSPF
Distribution Layer Configuration

Cisco.com

6k-dist-left configuration	

interface GigabitEthernet3/2
  description 3750 Access Switch
  dampening
  ip address 10.120.0.8 255.255.255.254
  ip pim sparse-mode
  ip ospf dead-interval minimal hello-multip 4
  load-interval 30
  carrier-delay msec 0
  mls qos trust dscp
!
interface TenGigabitEthernet4/1
  description 10 GigE to Core 1
  dampening
  ip address 10.122.0.26 255.255.255.254
  ip pim sparse-mode
  ip ospf dead-interval minimal hello-multip 4
  load-interval 30
  carrier-delay msec 0
  mls qos trust dscp
	router ospf 100
	  router-id 10.122.102.1
	  log-adjacency-changes
	  area 120 stub no-summary
	  area 120 range 10.120.0.0 255.255.0.0
	  timers throttle spf 10 100 5000
	  timers throttle lsa all 10 100 5000
	  timers lsa arrival 80
	  network 10.120.0.0 0.0.255.255 area 120
	  network 10.122.0.0 0.0.255.255 area 0

OSPF
Access Layer Configuration

Cisco.com

3750-Access configuration	

interface GigabitEthernet1/0/1
  description Uplink to Distribution 1
  no switchport
  dampening
  ip address 10.120.0.9 255.255.255.254
  ip pim sparse-mode
  ip ospf dead-interval minimal hello-multip 4
  load-interval 30
  carrier-delay msec 0
  srr-queue bandwidth share 10 10 60 20
  srr-queue bandwidth shape 10 0 0 0
  mls qos trust dscp
  auto qos voip trust
 	
interface FastEthernet2/0/1
  description Host port with IP Phone
  switchport access vlan 2
  switchport voice vlan 102
  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
  spanning-tree portfast
  spanning-tree bpduguard enable
	interface Vlan2
	  description Data VLAN for 3750 Data
	  ip address 10.120.2.1 255.255.255.0
	  ip helper-address 10.121.0.5
	  no ip redirects
	  ip pim sparse-mode
	  ip igmp snooping fast-leave
	!
	interface Vlan102
	  description Voice VLAN for 3750-access
	  ip address 10.120.102.1 255.255.255.0
	  ip helper-address 10.121.0.5
	  no ip redirects
	  ip pim sparse-mode
	  ip igmp snooping fast-leave
	!
	router ospf 100
	  router-id 10.120.250.2
	  log-adjacency-changes
	  area 120 stub no-summary
	  timers throttle spf 10 100 5000
	  timers throttle lsa all 10 100 5000
	  timers lsa arrival 80
	  passive-interface default
	  no passive-interface GigabitEthernet1/0/1
	  no passive-interface GigabitEthernet3/0/1
	  network 10.120.0.0 0.0.255.255 area 120

PIM
Distribution and Access Layer

Cisco.com

6k-dist-left configuration	4507k-access configuration

ip multicast-routing
!
interface Loopback2
  description Garbage-CAN RP
  ip address 2.2.2.2 255.255.255.255
!
interface Y
  description GigE to Access/Core
  ip address 10.122.0.Y 255.255.255.252
  ip pim sparse-mode
!<snip>
!
ip pim rp-address 10.122.100.1 GOOD-IPMC override
ip pim rp-address 2.2.2.2
ip pim spt-threshold infinity
!
ip access-list standard Default
  permit 10.0.0.0
ip access-list standard GOOD-IPMC
  permit 224.0.1.39
  permit 224.0.1.40
  permit 239.192.240.0 0.0.3.255
  permit 239.192.248.0 0.0.3.255
	ip multicast-routing
	ip igmp snooping vlan 4 immediate-leave
	ip igmp snooping vlan 104 immediate-leave
	no ip igmp snooping
	!
	interface VlanX
	  ip address 10.120.X.1 255.255.255.0
	  ip helper-address 10.121.0.5
	  no ip redirects
	  ip pim sparse-mode
	!
	ip pim rp-address 10.122.100.1 GOOD-IPMC override
	ip pim spt-threshold infinity
	!
	ip access-list standard Default
	  permit 10.0.0.0
	ip access-list standard GOOD-IPMC
	  permit 224.0.1.39
	  permit 224.0.1.40
	  permit 239.192.240.0 0.0.3.255
	  permit 239.192.248.0 0.0.3.255

PIM
Core Layer RP Configuration—1

Cisco.com

6k-core Left Anycast-RP configuration

```
ip multicast-routing
!
interface Loopback0
 description MSDP PEER INT
 ip address 10.122.10.1 255.255.255.255
!
interface Loopback1
 description ANYCAST RP ADDRESS
 ip address 10.122.100.1 255.255.255.255
!
interface Loopback2
 description Garbage-CAN RP
 ip address 2.2.2.2 255.255.255.255
!
interface TenGigabitEthernet M/Y
 ip address 10.122.0.X 255.255.255.252
 ip pim sparse-mode
!
ip pim rp-address 2.2.2.2
ip pim rp-address 10.122.100.1 GOOD-IPMC
       override
ip pim accept-register list PERMIT-SOURCES
ip msdp peer 10.122.10.2 connect-source
       Loopback0
ip msdp description 10.122.10.2
       ANYCAST-PEER-6k-core-right
ip msdp originator-id Loopback0
```

6k-core Right Anycast-RP configuration

```
ip multicast-routing
!
interface Loopback0
 description MSDP PEER INT
 ip address 10.122.10.2 255.255.255.255
!
interface Loopback1
 description ANYCAST RP ADDRESS
 ip address 10.122.100.1 255.255.255.255
!
interface Loopback2
 description Garbage-CAN RP
 ip address 2.2.2.2 255.255.255.255
!
interface TenGigabitEthernet M/Z
 ip address 10.122.0.X 255.255.255.252
 ip pim sparse-mode
!
ip pim rp-address 2.2.2.2
ip pim rp-address 10.122.100.1 GOOD-IPMC
       override
ip pim accept-register list PERMIT-SOURCES
ip msdp peer 10.122.10.1 connect-source
       Loopback0
ip msdp description 10.122.10.1
       ANYCAST-PEER-6k-core-left
ip msdp originator-id Loopback0
```

PIM
Core Layer RP Configuration—2

Cisco.com

6k-core Left Anycast-RP configuration	6k-core Right Anycast-RP configuration

! Continued from previous slide
!
ip access-list standard GOOD-IPMC
 permit 224.0.1.39
 permit 224.0.1.40
 permit 239.192.240.0 0.0.3.255
 permit 239.192.248.0 0.0.3.255
!
ip access-list extended PERMIT-SOURCES
 permit ip 10.121.0.0 0.0.255.255
        239.192.240.0 0.0.3.255
 permit ip 10.121.0.0 0.0.255.255
        239.192.248.0 0.0.3.255
	! Continued from previous slide
	!
	ip access-list standard GOOD-IPMC
	 permit 224.0.1.39
	 permit 224.0.1.40
	 permit 239.192.240.0 0.0.3.255
	 permit 239.192.248.0 0.0.3.255
	!
	ip access-list extended PERMIT-SOURCES
	 permit ip 10.121.0.0 0.0.255.255
	        239.192.240.0 0.0.3.255
	 permit ip 10.121.0.0 0.0.255.255
	        239.192.248.0 0.0.3.255