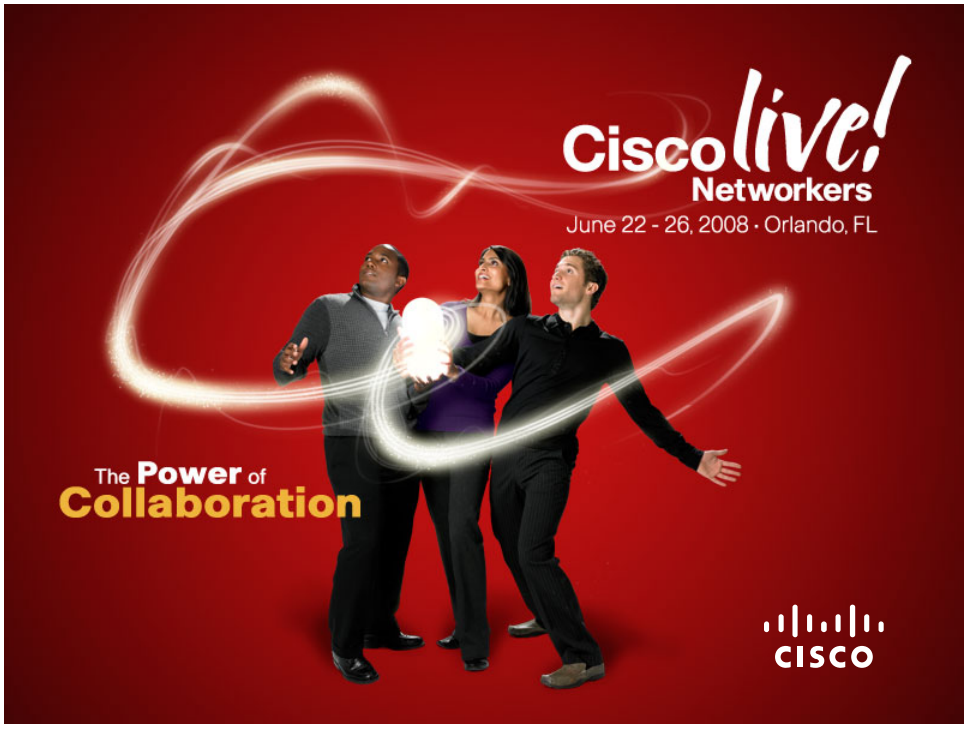
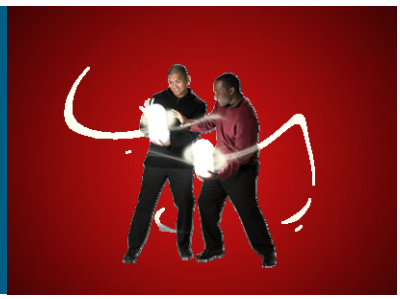


CISCO

Cisco live! Networkers

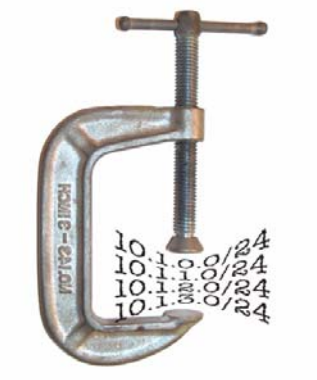
Deploying OSPF in a Large Scale Network



BRKRST-2310

Deploying OSPF in a Large Scale Network

- Market Segments
  - Service Provider Deployments
  - Enterprise Deployments
- Dialup Design techniques
- Design Best Practices
- Fast Convergence
- OSPF as PE-CE Protocol

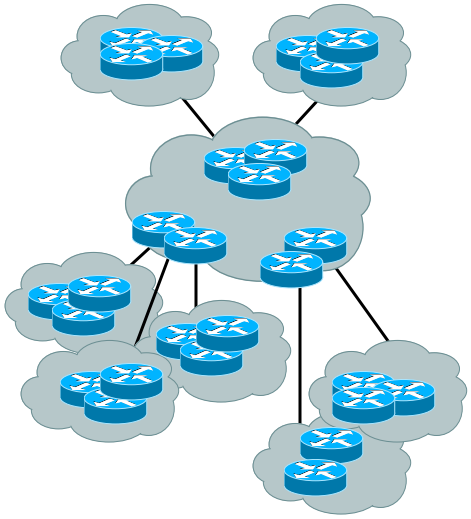


Market Segments

- Market Segments
  - a) Service Providers
  - b) Enterprise
    - Manufacturing
    - Retail

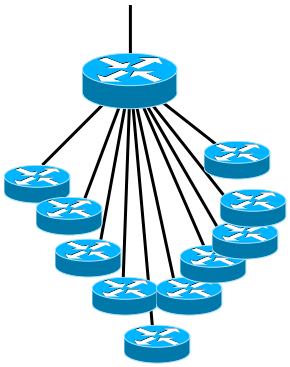
Service Providers

- SP networks are divided into pops
- Transit routing information is carried via BGP
- IGP is used to carry next hop only
- Optimal path to the next hop is critical



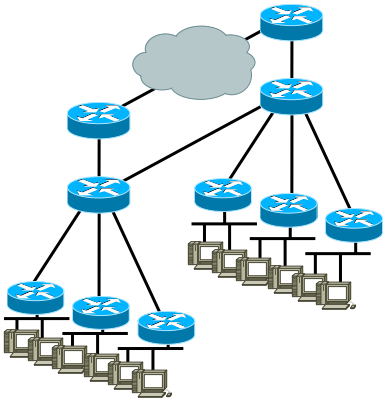
Enterprise Retails

- IGP scalability requirements on retail side is on the Hub and Spoke
- OSPF were designed keeping the Service Provider infrastructure in mind
- Lacks clear vision towards hub and spoke architecture
- Acquisitions brings more topological restrictions
- Distance Vector are better choice for e.g., EIGRP, RIPv2, ODR
- BGP?

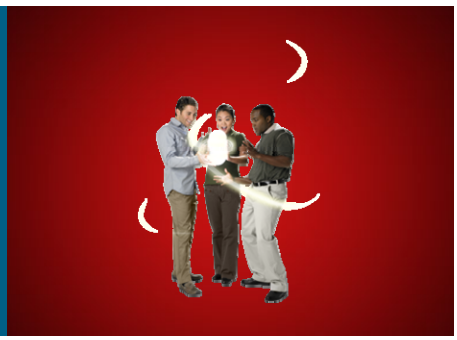


Enterprise Manufacturing

- Closet type infrastructure for host to host communications
- Large number of end devices connections
- Campus networks

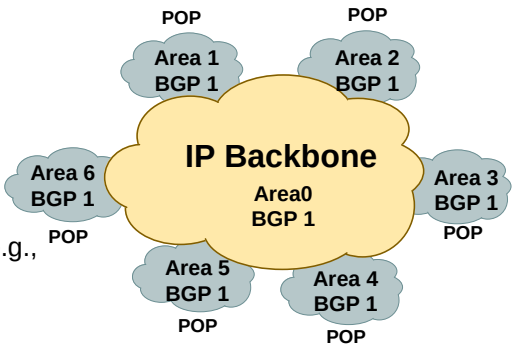


Service Providers Deployments



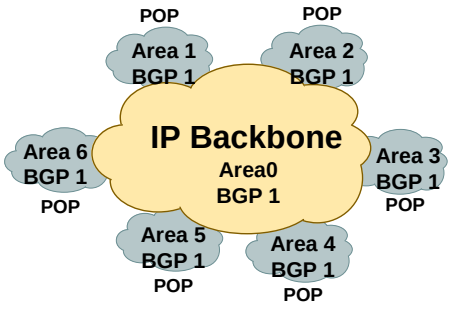
SP Deployment Characteristics

- SPs should have only one instance of IGP running throughout the network (exceptions are there)
- BGP carries external reachability
- IGP carries only next-hop (loopbacks are better for e.g., next-hop-self)



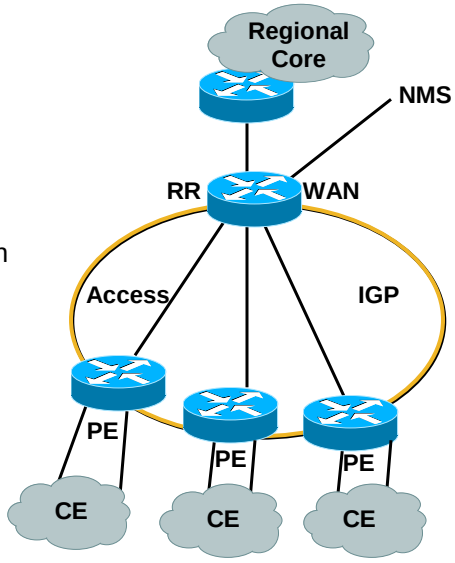
SP Architecture

- Major routing information is ~250K via BGP
- Largest known IGP routing table is ~6–7K
- Total of 267K
- 6K/267K ~ 2% of IGP routes in an ISP network
- A very small factor but has a huge impact on network convergence!



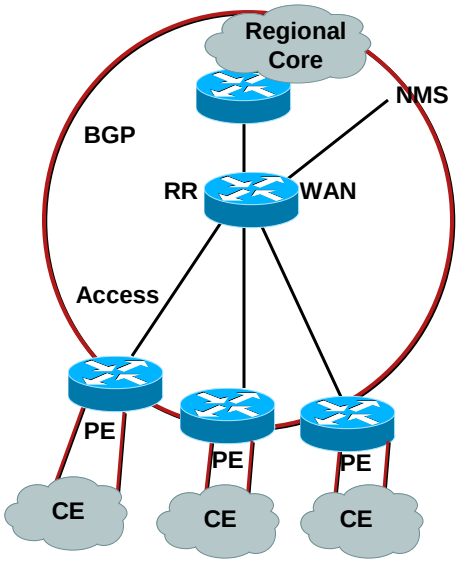
SP Architecture

- You can reduce the IGP size from 6K to approx the number of routers in your network
- This will bring really fast convergence
- Optimized where you must and summarize where you can
- Stops unnecessary flapping



SP Architecture

- The link between PE-CE needs to be known for management purpose
- BGP next-hop-self should be done on all access routers—unless PE-CE are on shared media (rare case)
- This will cut down the size of the IGP
- For PE-CE link do redistributed connected in BGP
- These connected subnets should ONLY be sent through RR to NMS for management purpose; this can be done through BGP communities
- Link Suppression can be done in OSPF via OSPF prefix suppression feature



OSPF Prefix Suppression

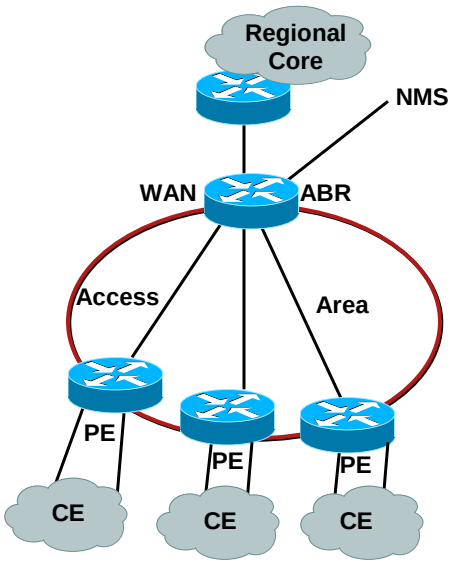
- When OSPF is enabled on the interface, it always advertises directly connected subnet
  - Only way to prevent the connected subnet to be advertised is to make the link unnumbered (not possible on broadcast segments)
  - Users may want to keep the links numbered for management and troubleshooting purposes
- WAN link routes can be suppressed to keep the routing table smaller
- Keep the network more secure
- Makes the network more scalable

OSPF Prefix Suppression (CLI)

- Router Mode
  - [no] prefix-suppression
  - Suppress all prefixes except loopbacks and passive interfaces
- Interface Mode
  - [no] ip ospf prefix-suppression [disable]
  - Suppress all prefixes on interface
  - Loopbacks and passive interfaces are included
  - Takes precedence over router-mode command
  - Disable keyword makes OSPF advertise the interface ip prefix, regardless of router mode configuration
- Available 12.4(15)T

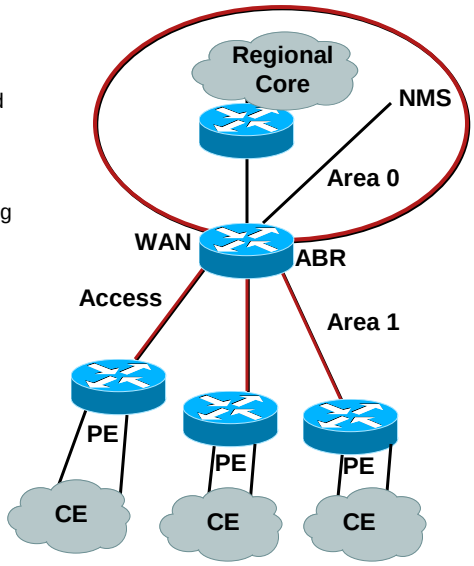
SP Architecture

- Where do we define area boundaries? WAN routers can be the ABRs
- Hide the pop infrastructure from your core
- Traffic engineering if needed can be done in core from WAN routers



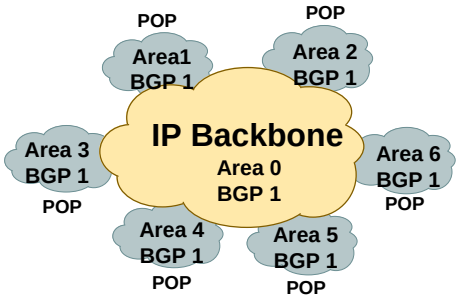
SP Architecture

- Physical address between ABR and PE should be in a contiguous blocks
- These physical links should be filtered via Type 3 filtering from area 0 into other areas or suppressed via prefix suppression feature
- Why? To reduce the size of the routing table within each pop
- Every area will carry only loopback addresses for all routers
- Only NMS station will keep track of those physical links
- These links can be advertised in BGP via redistribute connected
- PE device will not carry other Pop's PE's physical address in the routing table



SP Architecture

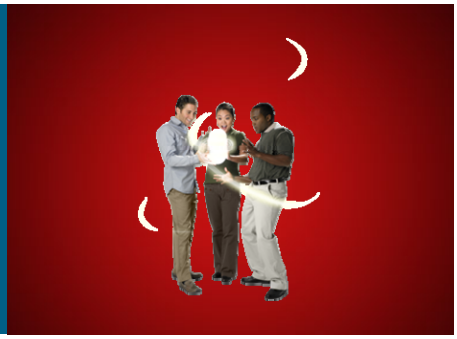
- Area 0 will contain all the routes
- This is the most intelligent form of routing and also there will not be too many routes in IGP
- If there are 500 pops and every pop contains 4 routers; then instead of having 6K routes you will only have 2K
- This is scalable and hack proof network!



Dealing with Redistribution

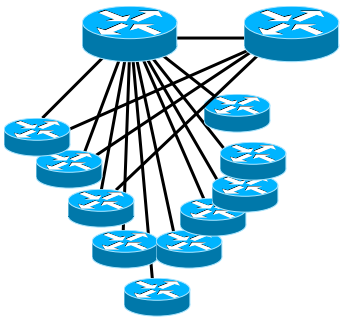
- Don't do it!
- If you're an SP you shouldn't be carrying external information in your IGP
- Let BGP take care of external reachability
- Use OSPF or to carry only next-hop information—i.e., loopbacks

Enterprise Deployments



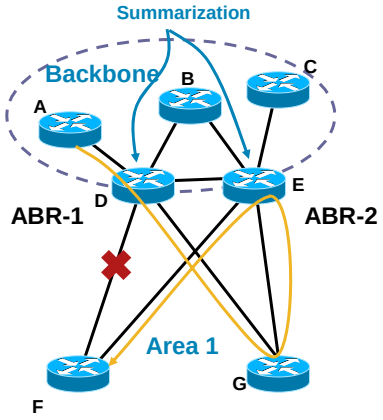
Enterprise Retail

- OSPF is not a very good choice for hub and spokes
- Enhancements are being made in IETF to make OSPF more robust on hub and spoke
- EIGRP, ODR, RIPv2 and BGP are better choices at the moment
- Enterprise BGP is not complicated
- You do not need to play with lot of attributes



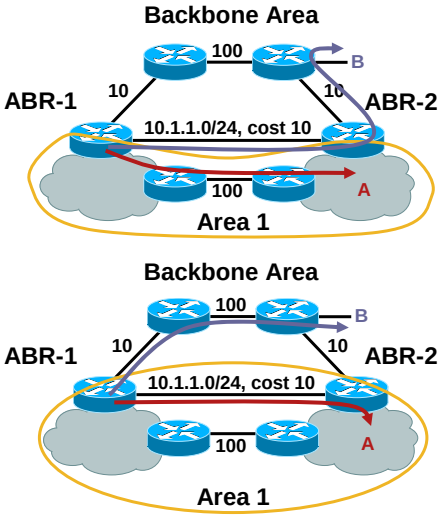
OSPF Border Connections

- Dual homed connections in hub and spoke networks illustrate a design challenge in OSPF: connections parallel to an area border
- Assume the D to E link is in area 0
- If the D to E link fails, traffic from A to F will:
  - Route towards the summary advertised by D
  - Route via the more specific along the path G, E, F



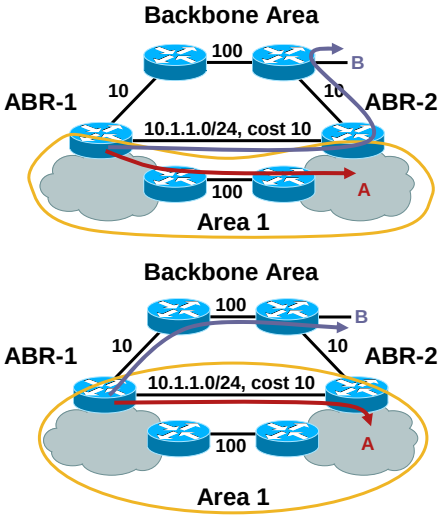
OSPF Border Connections

- Let's take a closer look at the problem
- Traffic prefers to stay within the area no matter what the actual link costs are
  - To reach A, we will take the higher cost link if the border link is in the backbone
  - To reach B, we will take the higher cost link if the border link is in the area
  - This is because we will always use an intra-area path over an inter-area path



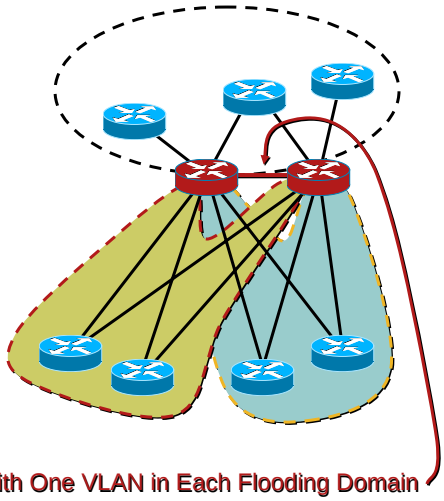
OSPF Border Connections

- Then, either:
  - Decide which traffic you want to route optimally
  - Use either virtual circuits (sub-interfaces) or real links to create one adjacency between the ABRs per routed area
- Configure the link in Area 1 with a virtual link in backbone (won't work for more than 1 area)



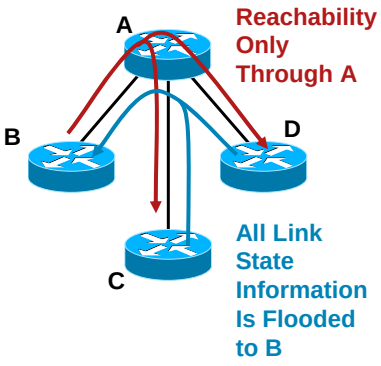
Enterprise Retail

- Summarization of areas will require specific routing information between the ABR's
- This is to avoid suboptimal routing
- The link between 2 hub routers should be equal to the number of areas
- As you grow the number of areas, you will grow the number of VLAN/PVC's–scalability issue
- Possible solution is to have a single link with adjacencies in multiple areas (draft-ietf-ospf-multi-area-adj-07.txt)



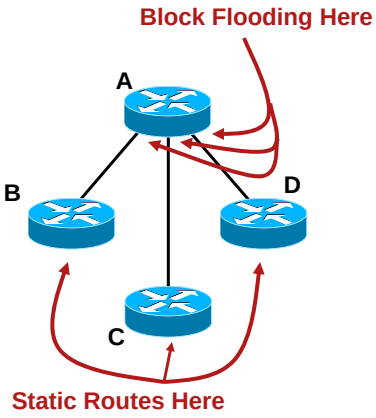
OSPF Hub and Spoke

- Every router within a flooding domain receives the same information
  - Although B can only reach C through A, it still receives all of C's routing information
- Because of this, OSPF requires additional tuning for hub and spoke deployments



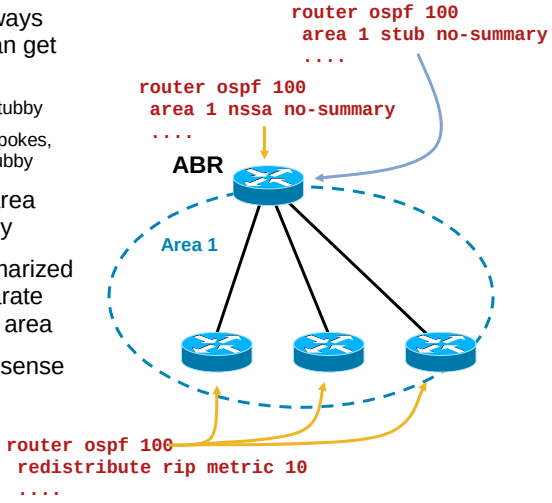
OSPF Hub and Spoke

- One of our primary goals is to control the amount of flooding towards the remotes
- You can reduce flooding by configuring the hub not to flood any information to the remotes at all
  - ip ospf database-filter all out
  - The remote routers must supply their own remote, or the hub router must originate a default locally
- This isn't a common configuration



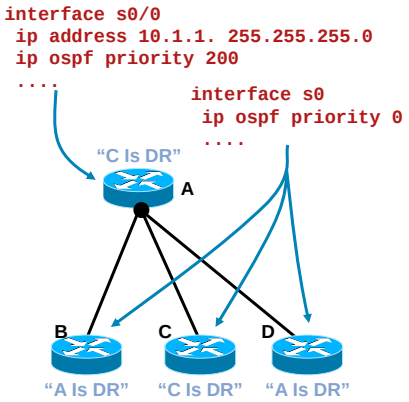
OSPF Hub and Spoke

- The spoke areas should always be the “most stubby” you can get away with
  - If possible, make them totally stubby
  - If there is redistribution at the spokes, make the area totally not-so-stubby
- The fewer spokes in each area the less flooding redundancy
- However, less can be summarized in the backbone and a separate sub-interface is needed per area
- Totally stubby makes more sense in multiple area situation



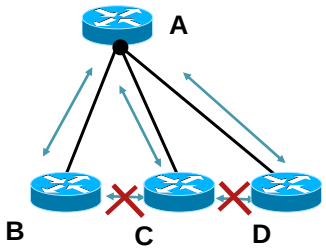
OSPF Hub and Spoke: Network Types

- OSPF could treat a multi-point link as a broadcast network or NBMA, but there are issues with flooding and forwarding
  - B and D don’t receive C’s packets, so they think A has the highest IP address, and elect A as DR
  - C elects itself as DR
  - Flooding will fail miserably in this situation
- We can set the OSPF DR priorities so the hub router is always elected DR
  - Set the spokes to 0 so they don’t participate in DR election



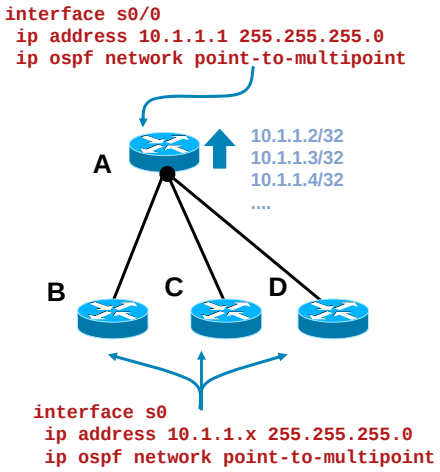
OSPF Hub and Spoke: NBMA/Broadcast

- OSPF will still have forwarding issues since the OSPF broadcast and NBMA assume a full mesh
- There is bi-directional connectivity between A and B, C, and D
- B, C, and D cannot communicate amongst themselves and traffic will be black-holed
- One can get around this using DLCI routing at the Frame Relay layer



OSPF Hub and Spoke: P2MP

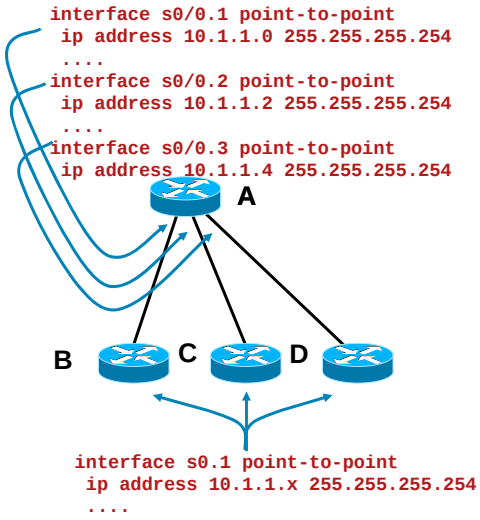
- You can also configure the serial interface at the hub router as a point-to-multi-point type
  - All the remotes are in a single IP subnet
  - OSPF treats each remote as a separate point-to-point link for flooding
- OSPF will advertise a host route to the IP address of each spoke router to provide connectivity
- Smaller DB Size compare to p2p
- Most natural OSPF solution



OSPF Hub and Spoke: P2P Sub-Interfaces

- OSPF can also use point-to-point sub-interfaces, treating each one as a separate point-to-point link
- Increase the DB size
- These use more address space and require more administration on the router

  Use /31 addresses for these point to point links

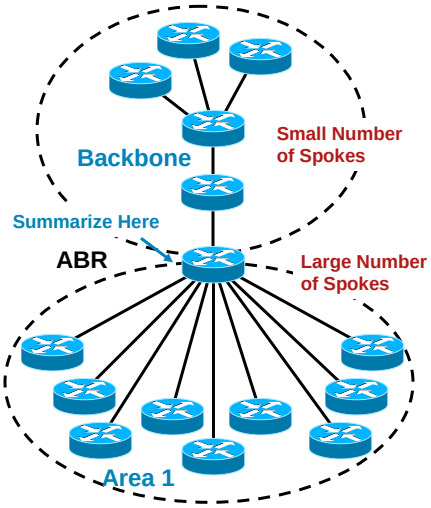


OSPF Hub and Spoke

Network Type	Advantages	Disadvantages
Single Interface at the Hub Treated as an OSPF Broadcast or NBMA Network	Single IP Subnet Fewer Host Routes in Routing Table	Manual Configuration of Each Spoke with the Correct OSPF Priority No Reachability Between Spokes or Labor-Intensive Layer 2 Configuration
Single Interface at the Hub Treated as an OSPF Point-to-Multipoint Network ip ospf Network-Type Point-to-Multipoint	Single IP Subnet No Configuration per Spoke Most Natural Solution Smaller database	Additional Host Routes Inserted in the Routing Table
Individual Point-to-Point Interface at the Hub for Each Spoke ip ospf Network-Type Point-to-Point	Can Take Advantage of End-to-End Signaling for Down State	Lost IP Address Space More Routes in the Routing Table Larger database Overhead of Sub-Interfaces

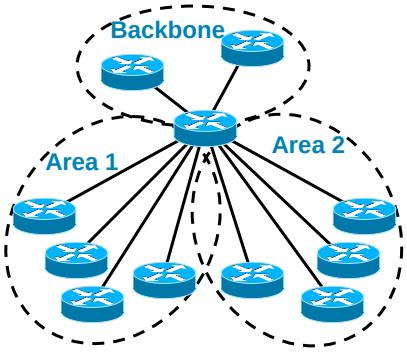
OSPF Hub and Spoke

- The other consideration is determining how many spokes should fall within the area
- If the number of remotes are low, we can place the hub and its spokes within an area
- However, as the count rises, we want to make the hub an ABR, and split off the spokes in a single area
- If you're going to summarize into and out of the remotes, the hub needs to be a border router



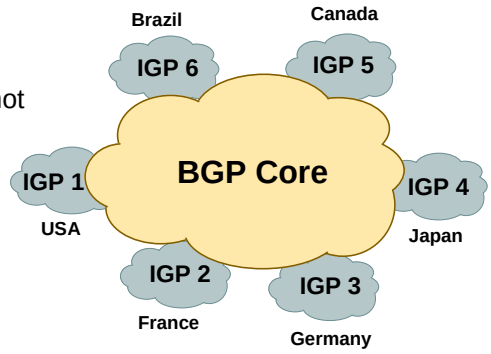
OSPF Hub and Spoke

- Low speed links and large numbers of spoke may require multiple flooding domains
- Balance the number of flooding domains on the hub against the number of spokes in each flooding domain
- The link speeds and the amount of information being passed through the network determines the right balance



Enterprise Manufacturing

- Can have multiple ‘islands’ of IGPs
- Islands tied together by a BGP core
- One Island’s instability will not impact other Islands
- This increases network stability
- May be a requirement for redistribution



Dealing with Redistribution

- The number of redistribution boundaries should be kept to a minimum

  Why?

  Because you have better things to do in life besides; build the access lists

- When redistributing try to place the DR as close to the ASBR as possible to minimize flooding

Dealing with Redistribution

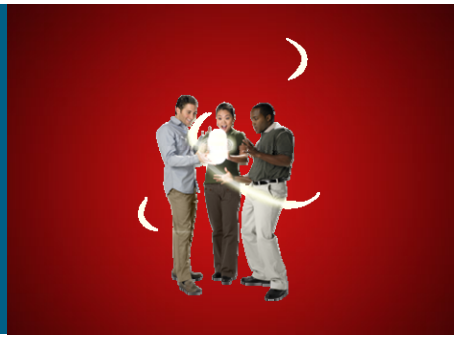
- Be aware of metric requirements going from one protocol to another
  - RIP metric is a value from 1–16
  - OSPF Metric is from 1–65535
- Include a redistribution default metric command as a protection

```
router ospf 1
  network 130.93.0.0 0.0.255.255 area 0.0.0.0
  redistribute rip metric 1 subnets
```


Dealing with Redistribution

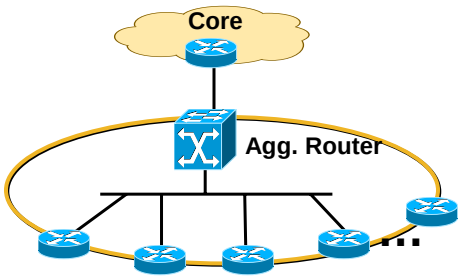
- Redistribute only what is absolutely necessary
  - Don’t redistribute full Internet routes into OSPF
- Default route into BGP core; let the core worry about final destination

Dialup Design Tips



Dialup Design Practices

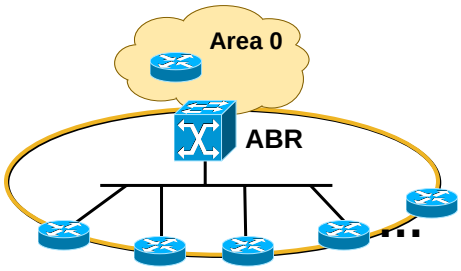
- Two kinds of Pools can be defined on NAS: Static Pools and Distributed Pools
- Static Pool: address range remain within a single NAS—easier to manage from routing perspective
- Dynamic Pool: address range may be distributed over multiple NAS's—hard to manage from a routing perspective



1000+ Routes Injected by Each NAS

Dialup Design with Static Pool Addresses

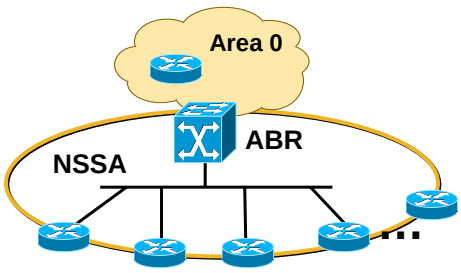
- Three ways to propagate dialup routes from NAS:
  - Either Static route to pool address to null 0 with redistribute static on NAS or
  - Assign the pool add on a loopback on NAS with OSPF p2p network-type including loopback in an OSPF area or
  - Static route on ABR for the pool address pointing towards NAS (ASBRs)—
- Static pool do not require redistribute connected subnets on NAS



1000+ Routes Injected by Each NAS

Dialup Design with Dynamic Pool Addresses

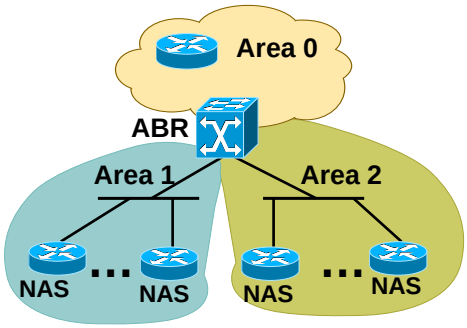
- Distributed pool REQUIRES redistribute connect subnets
- If pool is distributed, you can't summarize the pools at ABR because of redistribute connected subnets on NASs' unless its an NSSA, why?
- NSSA can summarize the Type 7 routes at the ASBR as well as ABR



1000+ Routes Injected by Each NAS

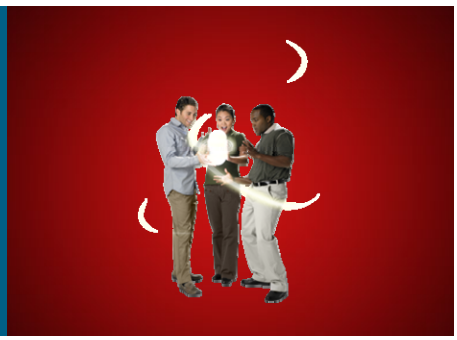
Dialup Design
Practices Scalability Issues

- If an area has too many routes injected by NAS then break it up into more than one area
- Area should be configured as NSSA for controlling type 5 at ABR level
- NSSA ABR can filter type 5 originated by NAS servers
- Configure totally NSSA to filter internal area routes from each other



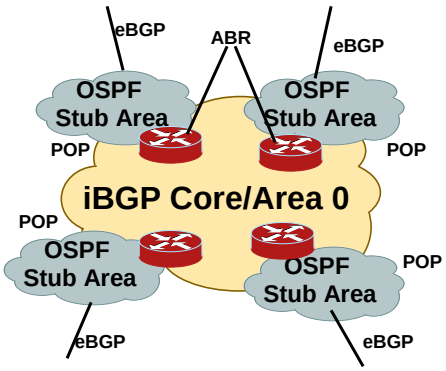
1000+ Routes Injected by Each NAS

Design Best Practices



Accidental Redistribution Prevention (OSPF)

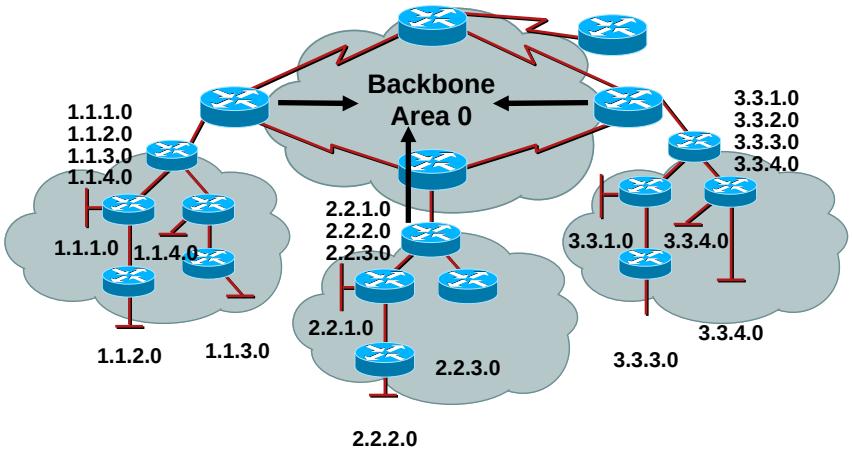
- Areas should be defined as stub to prevent accidental redistribution of eBGP into OSPF
- Type 3 LSA filtering should be used at ABR's and only routers' loopbacks should be allowed to leak into other areas
- Loopback should be in private address space to make LSA type 3 filtering easier; for e.g., 10.0.0.0/8
- iBGP routes can not be redistributed into IGP by default
- NMS resides in area 0 here



Area Size: How Many Router in an Area?

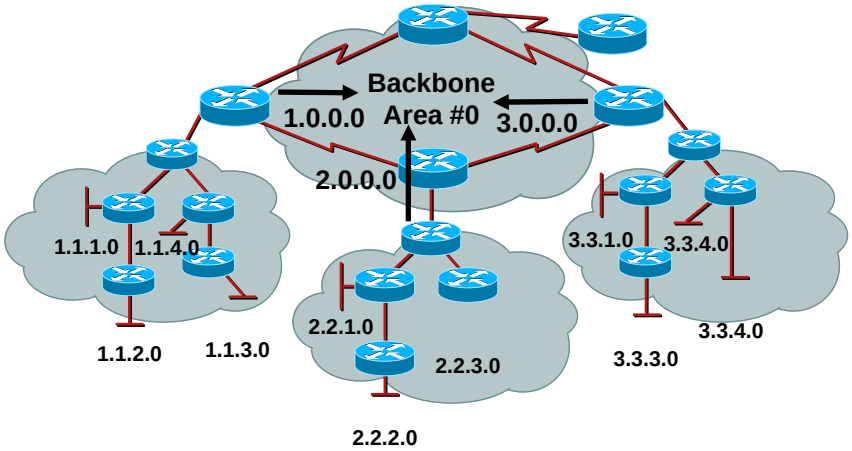
- Number of adjacent neighbors is more a factor
- More important from the standpoint of amount of information flooded in area
- Keep router LSAs under MTU size—exceeding is bad
  - Implies lots of interfaces (and possibly lots of neighbors)
  - Exceeding results in IP fragmentation which should be avoided

OSPF Border Area Filtering: Ranges



- Only summary LSA advertised out
- Link-state changes do not propagate

OSPF Border Area Filtering: Ranges



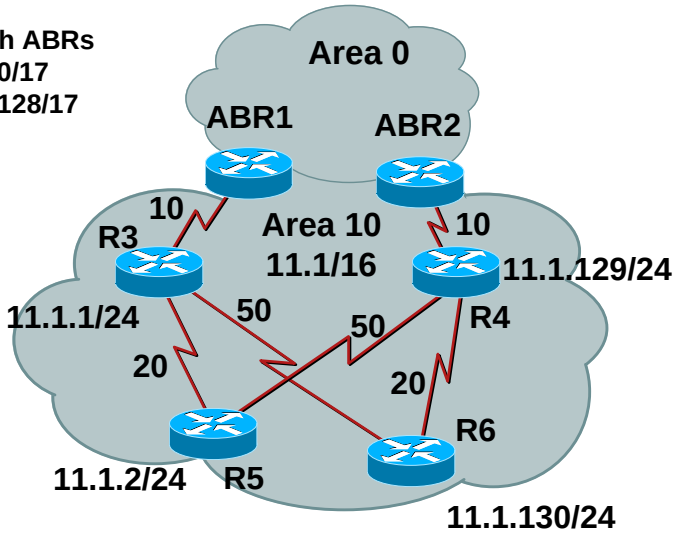
- Only summary LSA advertised out
- Link-state changes do not propagate

Summarization Technique

Configure on Both ABRs
Area-Range 11.1.0/17
Area-Range 11.1.128/17

Cost to Range 1:
Via ABR1: 30
Via ABR2: 80

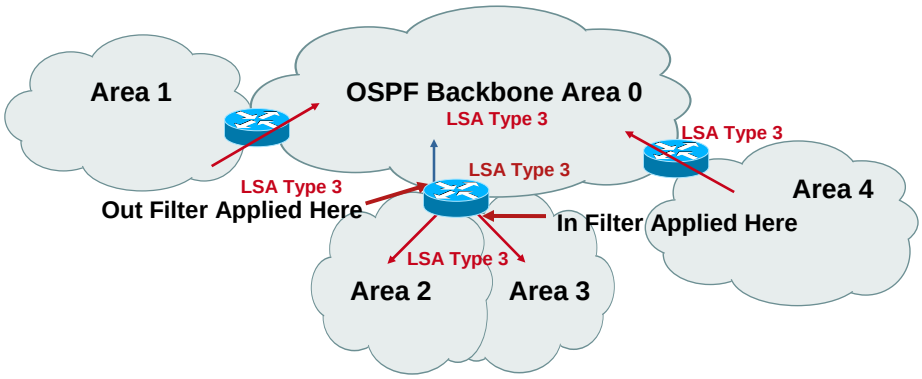
Cost to Range 2:
Via ABR1: 80
Via ABR2: 30



OSPF Border Inter-Area Filtering

- Type-3 LSA filtering on ABR
- Uses prefix-list to filter prefixes from being advertised from/to a specific area
- Works both directions—out of a specific area or into a specific area

OSPF and Filtering: Inter-Area Filtering

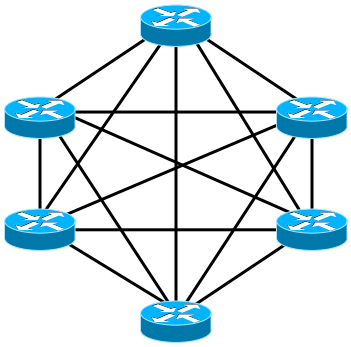


area 0 filter-list prefix AREA_0_OUT out
area 0 filter-list prefix AREA_0_IN in

OSPF Full Mesh

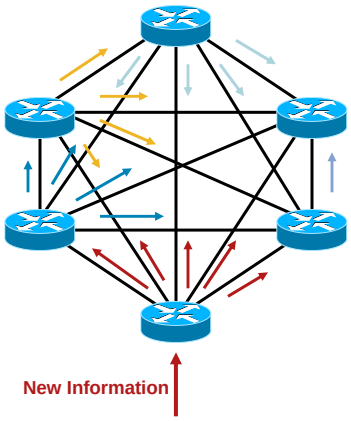
Full Mesh Topologies Are Complex:

- 2 routers == 1 link
- 3 routers == 3 links
- 4 routers == 6 links
- 5 routers == 10 links
- 6 routers == 15 links
- N routers == ((N) (N-1)) / 2
- …



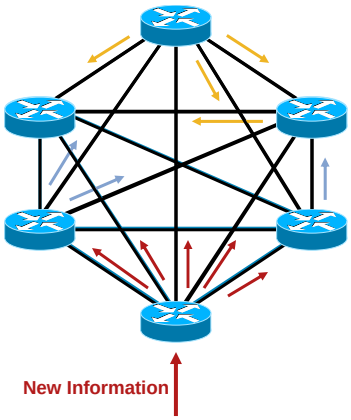
OSPF Full Mesh

- Flooding routing information through a full mesh topology is the MAIN concern
  - Each router will, with optimal timing, receive at least one copy of every new piece of information from each neighbor on the full mesh
- Link down is n^2 and node down is n^3. For a very large router LSA it could become n^4
- There are several techniques you can use to reduce the amount of flooding in a full mesh
  - Mesh groups reduce the flooding in a full mesh network
  - Mesh groups are manually configured "designated routers"



OSPF Full Mesh

- Pick a subset (>=2) of routers to flood into the mesh and block flooding on the remainder; leave flooding to these routers open
- This will reduce the number of times information is flooded over a full mesh topology




```
! Point-to-Point
interface serial 1/0
 ip ospf database-filter all out
 ....

! Point-to-Multipoint
router ospf 1
 neighbor 10.1.1.3 database-filter all out
```

OSPF Flood Reduction

- LSA have an age (1–3600)
- When LSA is originated its age is set to 0
- LSA is flushed from the area/domain when its age reaches MAXAGE (3600 sec)
- Each router must periodically refresh all self-generated LSAs every 1800 seconds (jittered)
- Demand Circuits (RFC 1793) introduces concept of DoNotAge LSAs—high order or DoNotAge bit of LSA header age indicates not to age the LSA
- RFC 4136 extends this concept to all interface types
- Eliminates periodic refresh of unchanged LSAs

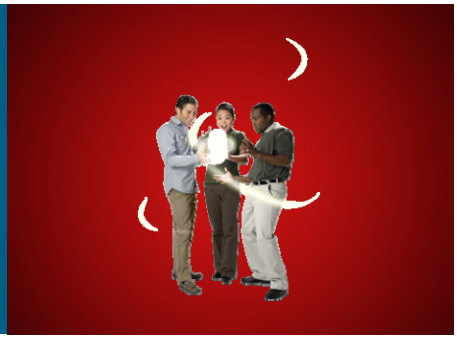
OSPF Flood Reduction (Cont.)

- Configured at interface level

```
 ! All interfaces types
interface ethernet 0/0
ip ospf flood-reduction
....
```

- Very useful in fully meshed topologies
- Possible extension to force refresh at an increased interval, e.g. every four hours; currently not implemented
  - Changes could be missed if the sum of the changes results in an identical checksum
  - Periodic refresh provides recovery from bugs and glitches

Fast Convergence



Detect the Event, Process the Event

Fast Convergence

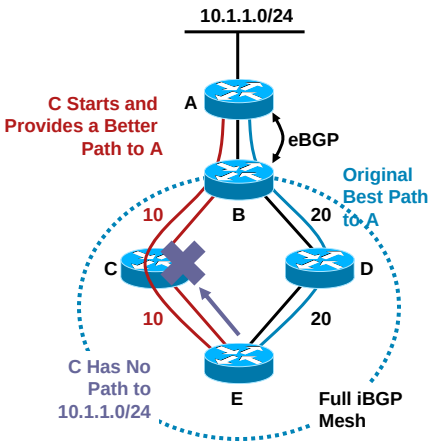
- Network convergence is the time needed for traffic to be rerouted to the alternative or more optimal path after the network event
- Network convergence requires all affected routers to process the event and update the appropriate data structures used for forwarding
- Network convergence is the time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

Network Convergence

- Techniques/tools for fast convergence
  - Carrier Delays **Detect**
  - Hello/dead timers **Detect**
  - Bi-Directional Forwarding Detection—(BFD) **Detect**
  - LSA packet pacing **Propagate**
  - Interface event dampening— **Propagate**
  - Exponential throttle timers for LSA & SPF **Process**
  - MinLSArrival Interval **Process**
  - Incremental SPF **Process**
- Techniques/tools for Resiliency
  - Stub router (e.g., max-metric)
  - Graceful Restart

OSPF Stub Router (RFC 3137)

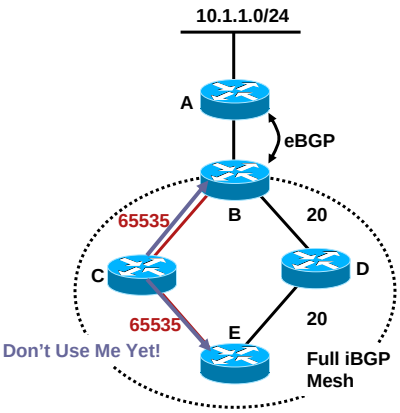
- E is learning 10.1.1.0/24 through iBGP from D with a next hop of A; E's best path to A is through D
- C comes up, and OSPF converges quickly; E now chooses C to reach A
- BGP takes longer to converge; when E forwards packets to C for 10.1.1.1, C hasn't finished building its BGP tables
- C drops the packets



Not to Be Confused with Stub Areas

OSPF Stub Router

- C signals E and B through OSPF that they should not route this direction (through stub router)
- When BGP converges, C changes back to its normal metric, so B and E will now route through it
- OSPF uses max-metric router-lsa on-startup wait-for-bgp to configure this feature



Stub Router Advertisement

- max-metric router-lsa [on-startup {wait-for-bgp | <announce-time>}]
- Syntax Description
- router-lsa Always originate router-LSAs with maximum metric
- on-startup Set max-metric temporarily after reboot
- announce-time Time, in seconds, router-LSAs are originated with max-metric (default is 600s)
- wait-for-bgp Let BGP decide when to originate router-LSA with normal metric (i.e., stop sending router-LSA with max-metric)
- Can be used to take a router out of service since default is indefinite

Stub Router: show ip ospf max-metric

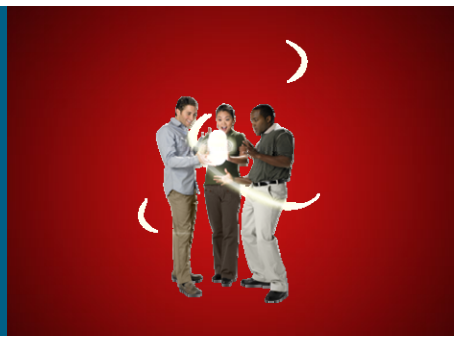
- OSPF allows for a configurable option to wait for BGP at startup
- OSPF announces maximum metric while BGP is coming up
- This enhancement shows real-time status on what is happening internally

```
r201#sh ip ospf max-metric
Routing Process "ospf 1" with ID 1.0.0.201
 Start time: 00:00:01.876, Time elapsed: 00:00:56.668
 Originating router-LSAs with maximum metric, Time remaining: 00:00:09
  Condition: on startup for 66 seconds, State: active

r201#sh ip ospf max-metric
Routing Process "ospf 1" with ID 1.0.0.201
 Start time: 00:00:01.876, Time elapsed: 00:01:12.560
 Originating router-LSAs with maximum metric
  Condition: on startup for 66 seconds, State: inactive
 Unset reason: timer expired, Originated for 66 seconds
 Unset time: 00:01:07.932, Time elapsed: 00:00:06.504

r201#sh ip ospf 2 max-metric
Routing Process "ospf 2" with ID 2.2.2.2
 Start time: 00:07:42.052, Time elapsed: 00:00:20.684
 Router is not originating router-LSAs with maximum metric
```

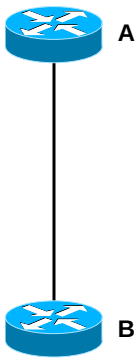

Graceful Restart



Non Stop Forwarding, Grace Restart

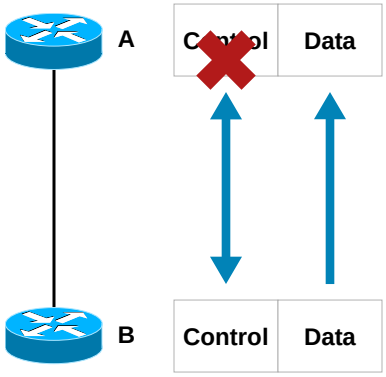
Graceful Restart

- Graceful Restart (GR) allows a router's control plane to reset without impacting global routing
- Consider two routers connected over a single circuit



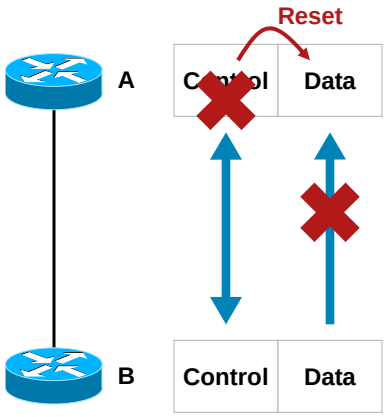
Graceful Restart

- Router A loses its control plane for some period of time
- It will take some time for Router B to recognize this failure and react to it



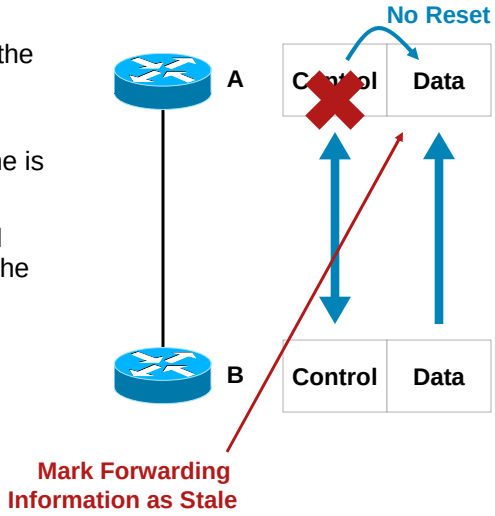
Graceful Restart

- During the time that A has failed and B has not detected the failure, B will continue forwarding traffic through A
- Once the control plane resets, the data plane will reset as well, and this traffic will be dropped
- NSF reduces or eliminates the traffic dropped while A's control plane is down



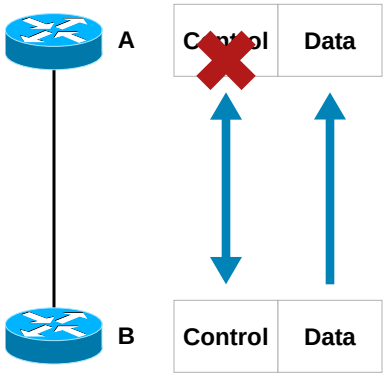
Graceful Restart

- If A is NSF capable, the control plane will not reset the data plane when it restarts
- Instead, the forwarding information in the data plane is marked as stale
- Any traffic B sends to A will still be switched based on the last known forwarding information



Graceful Restart

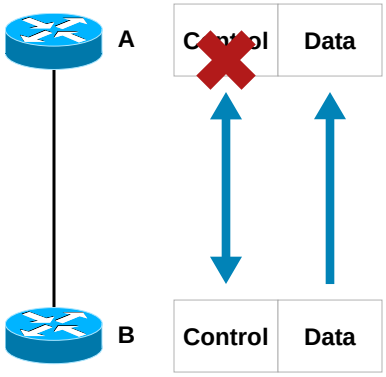
- While A's control plane is down, B's hold timer counts down
- A has to come back up and signal B before B's hold timer expires, or B will route around it
- When A comes back up, it signals B that it is still forwarding traffic and would like to resynchronize
- A and B get resynchronized and this completes the NSF process



Hold Timer: 6

Graceful Restart

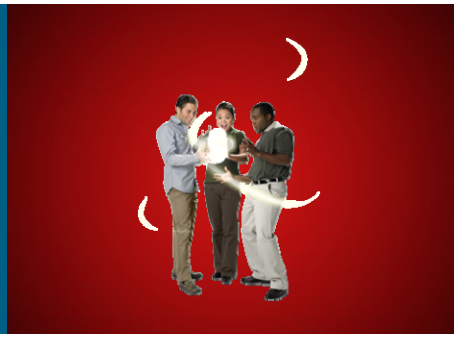
- NSF is a protection scheme
- NSF and fast hellos/BFD do not go well and should be avoided
- NSF makes more sense in a singly homed edge devices
- Multihomed edge devices should go on restoration scheme instead and switch over to the alternate path



Graceful Restart Mechanism Comparison

	GR Signaling	Resynchronization
Cisco NSF (LLS/OOB) RFC 4811, 4812, 4813	Link Local Signaling (LLS), Extending Hellos to Add Optional Information	Out-of-Band Synchronization (OOB), Creates a New Form of Synchronization (Similar to Standard Database Synchronization)
IETF RFC 3623 (Opaque LSAs)	Grace LSA (Opaque Link Local LSA)	Normal Database Exchange (Adjacency Advertised as FULL)

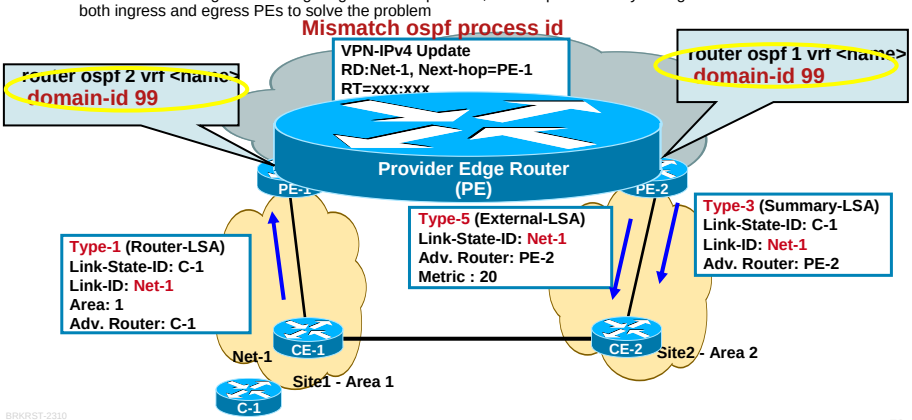
OSPF as PE-CE



MPLS Related Techniques

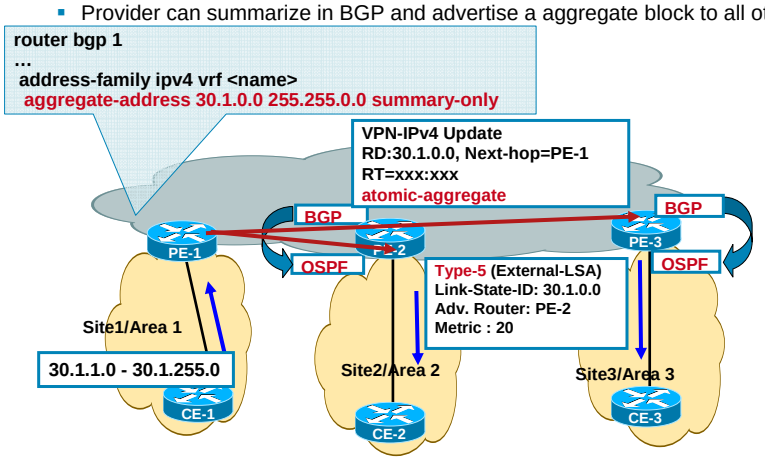
OSPF as PE-CE Routing Protocol
Different OSPF Process ID

- In service provider networks, the OSPF process id should match, otherwise external type-5 routes are generated
- Site 2 expects Type-3 inter-area routes from Site1 via PE2 but instead receives external type-5
- OSPF process-ID is usually locally significant. In MPLS VPN, consider service provider cloud acting as a single router from OSPF perspective
- Instead of removing and reconfiguring the OSPF process, service provider may configure same domain-id on both ingress and egress PEs to solve the problem



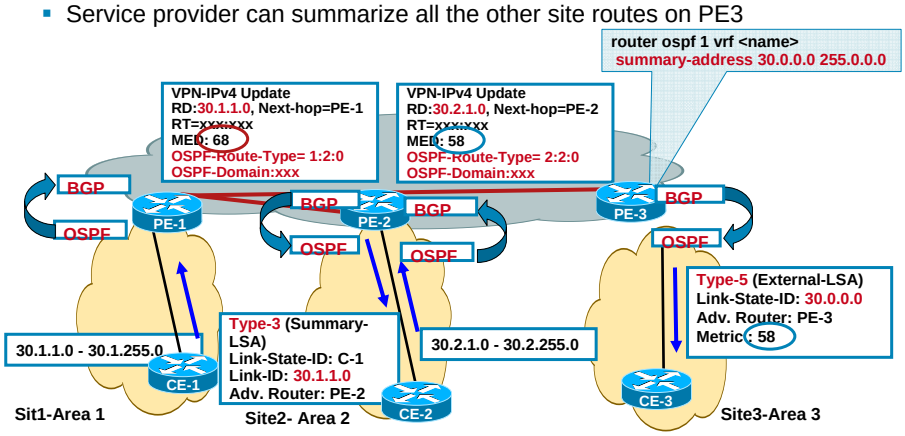
OSPF as PE-CE Routing Protocol
Summarization (Ingress PE)

- Customer wants to send a summary route to all other sites from site1 (area 1)
- Summarization not possible since ABR does not exist in site1
- Provider can summarize in BGP and advertise a aggregate block to all other sites



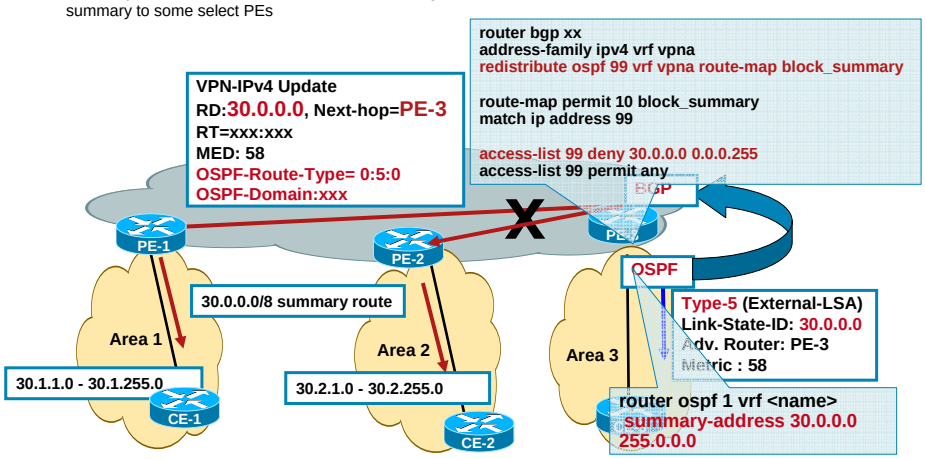
OSPF as PE-CE Routing Protocol
Summarization—Egress PE

- Customer needs to send a summary route to site3–area3
- Customer can't summarize from each individual site. No ABR exists within the site (area 1 or area 2)
- Service provider can summarize all the other site routes on PE3

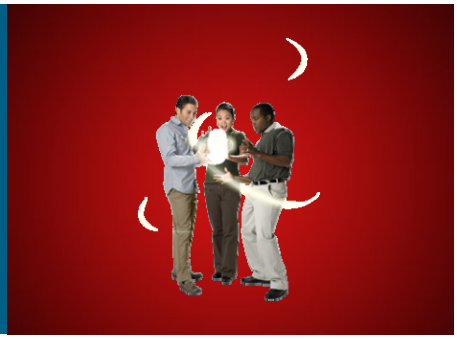


OSPF as PE-CE Routing Protocol
Summarization—Egress PE—Loop Prevention

- Summary is originated in OSPF
- Summary route will be propagated to all customer sites as a result of redistribution from OSPF into BGP
- Summary route should be filtered while redistributing OSPF into BGP on PE3 unless it is desirable to send summary to some select PEs



OSPF as PE-CE Protocol



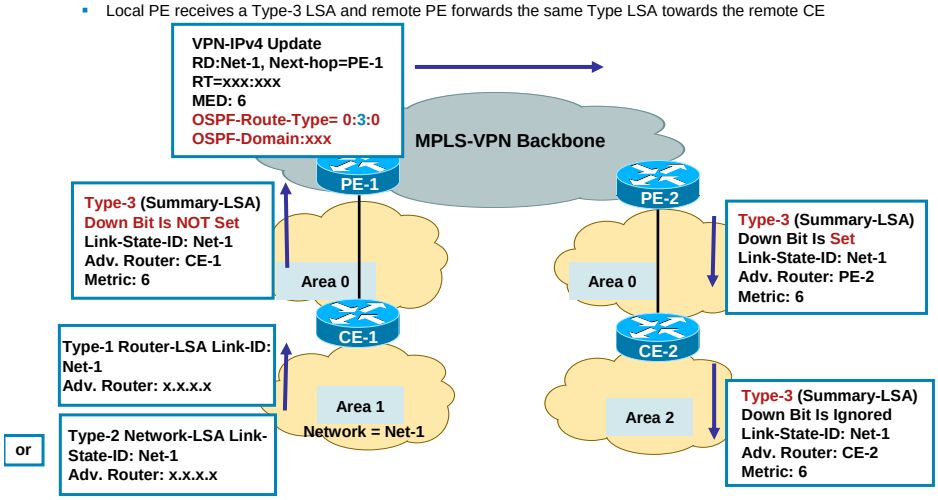
Area Placement Considerations

Common Design Consideration: OSPF Area Placement

- OSPF sites have area 0
- OSPF sites belong to different areas
- All OSPF sites belong to the same area
  - No backdoor exists
  - Backdoor is present

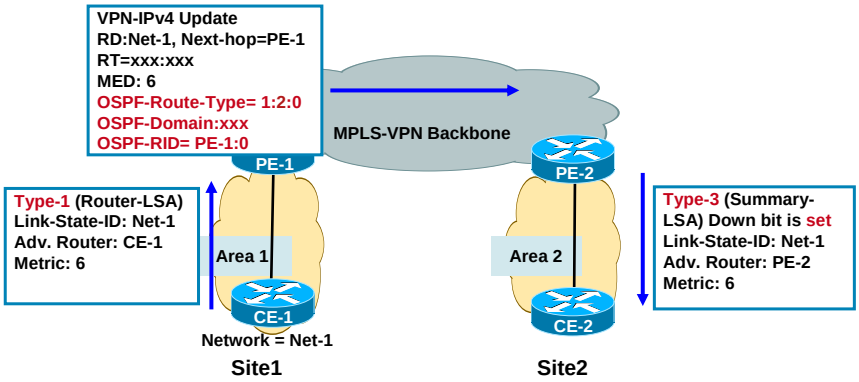
Common Design Consideration–OSPF Area Placement
OSPF Sites Have Area 0

- Type 1 or Type2 LSA converted into summary LSA by the Customer ABR
- Local PE receives a Type-3 LSA and remote PE forwards the same Type LSA towards the remote CE



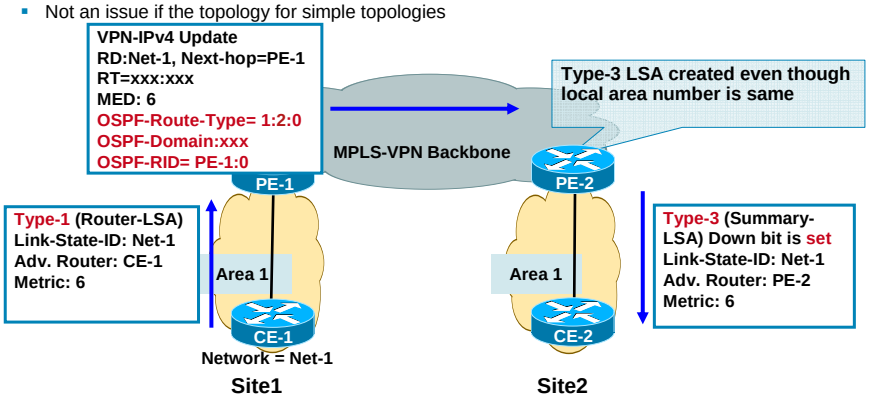
Common Design Consideration—OSPF Area Placement
OSPF Sites Belong to Different Areas

- Area 0 is not mandatory when migrating to MPLS VPN service
- VPN sites may have different Sites configured for different areas
- If Area 0 exists, it must touch MPLS VPN PE routers



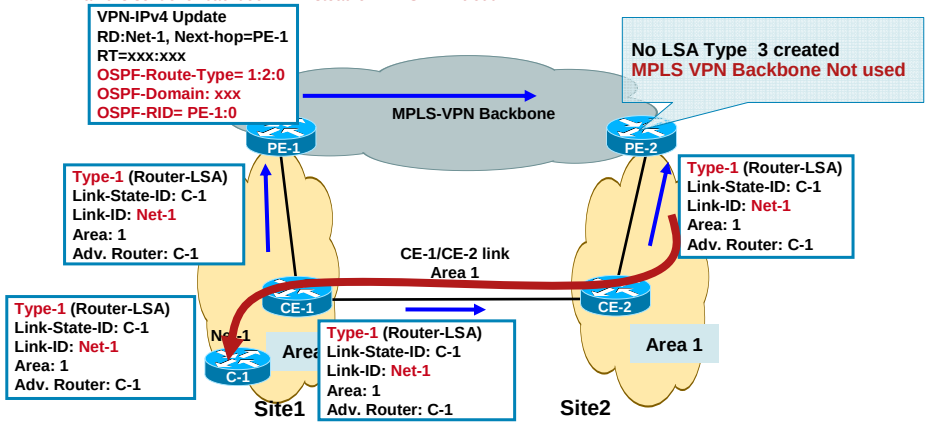
Common Design Consideration—OSPF Area Placement
All Sites Belong to the Same Area—Backdoor Does Not Exist

- All VPN sites belong to same OSPF area (area 1 in the example)
- Service provider PEs acts as OSPF ABR routers
  - Type1 or 2 LSA are always converted into Type 3
- Remote Site receives a summary LSA
- Not an issue if the topology for simple topologies



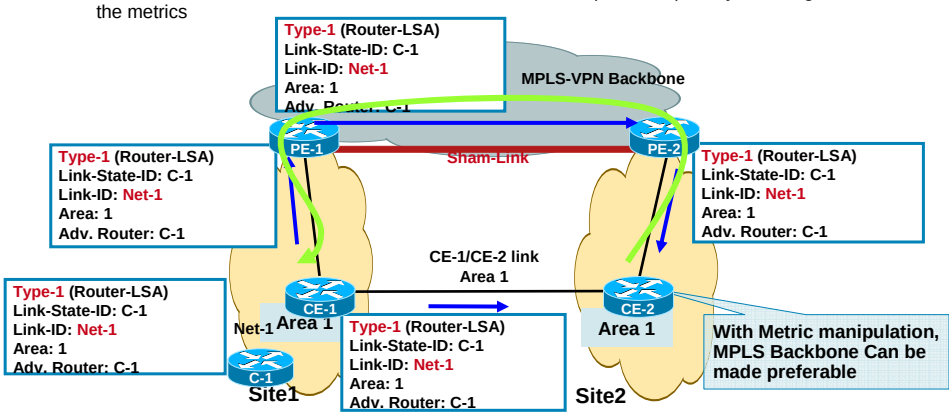
Common Design Consideration—OSPF Area Placement
All Sites Belong to the Same Area—Backdoor Exists

- Customers Sites are in the same area and there is a backdoor link
- Route is advertised to MPLS VPN backbone
- Same prefix is learnt as intra-area route via backdoor link
- PE2 does not generate Type3 LSA once type-1 LSA is received from the site
- Traffic is sent over backdoor link instead of MPLS VPN cloud



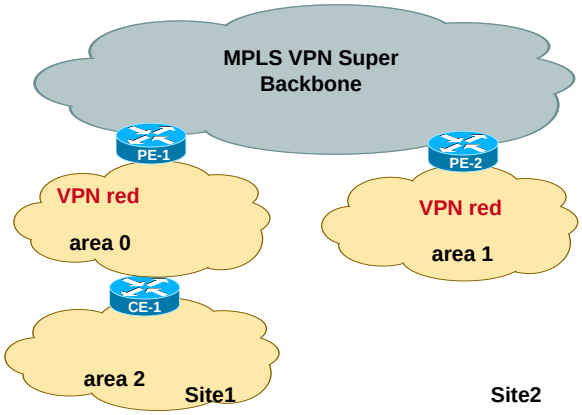
Common Design Consideration—OSPF Area Placement
Sites Belong to the Same Area—Backdoor with Sham Link

- The sham link is treated as a virtual-link : unnumbered, ptp, DC link
- The sham link is reported in the router LSA's type 1 originated by the two routers connecting to the sham link
- The MPLS VPN backbone or the backdoor link can be made preferred path by tweaking the metrics



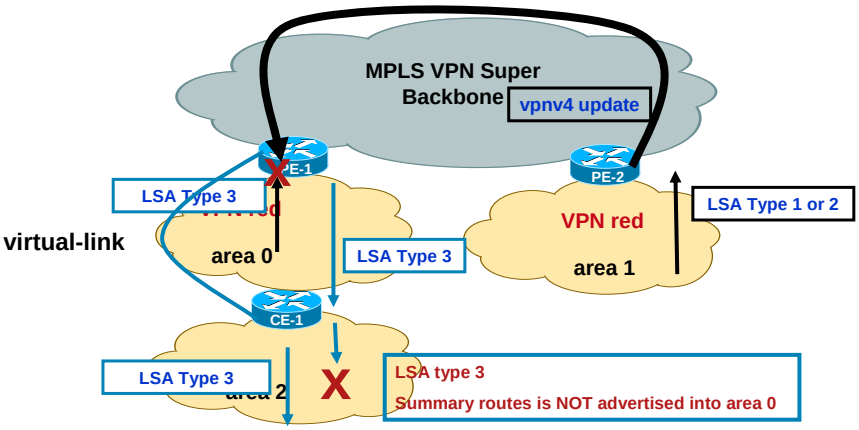
Common Design Consideration—OSPF Area Placement
Other Scenarios

- Some OSPF sites entirely belong to area 0 and some other sites can belong to non area 0
- Some sites may consist of hierarchical OSPF topology consisting of area 0 as well as non-zero areas
- Both scenarios are valid



Common Design Consideration—OSPF Area Placement
Area 0 Placement

- As before some sites may consist of hierarchical OSPF topology consisting of area 0 as well as non-zero areas.
- If site contains area 0, it must touch provider PE router
- OSPF RULE: Summary LSAs from non-zero area's are not injected into backbone area 0
- Inter-area routes will not show up unless a Virtual link is created



Database Overload Protection

- Protects router from a large number of received LSAs
  - Possibly as a result of the misconfiguration on remote router
- Router keeps the number of received (non self-generated) LSAs
- Maximum and threshold values are configured
- When threshold value is reached, error message is logged
- When maximum value is exceeded, no more new LSAs are accepted
- If the counter does not decrease below the max value within one minute we enter 'ignore-state'

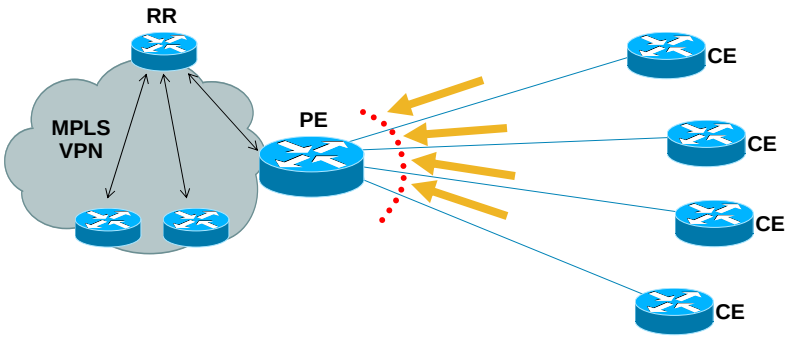
Database Overload Protection (Cont.)

- In ‘ignore-state’ all adjacencies are taken down and are not formed for ‘ignored-interval’
- Once the ‘ignored-interval’ ends we return to normal operations
- We keep the count on how many times we entered ‘ignore-state’—‘ignore-count’
- When ‘ignore-count’ exceeds its configured value, OSPF process stays in the ‘ignore state’ permanently

  Ignore-count is reset to 0, when we do not exceed maximum number of received LSAs for a ‘reset-time’

  The only way how to get from the permanent ignore-state is by manually clearing the OSPF process

Database Overload Protection (Deployment Example)



- PE protected from being overloaded by CE
- One misbehaving VRF affected, other works OK

Database Overload Protection (CLI)

- Router mode

```
max-lsa <max> [<threshold> [warning-only]
              [ignore-time <value>]
              [ignore-count <value>
              [reset-time <value>]]
```

- Available in: 12.3(7)T 12.2(25)S 12.0(27)S 12.2(18)SXE 12.2(27)SBC
- With CSCsd20451 deployable without flapping of all neighbors. Available in 12.2(33)SXH, 12.2(33)SRC, 12.5

OSPF Limit on Number of Redistributed Routes

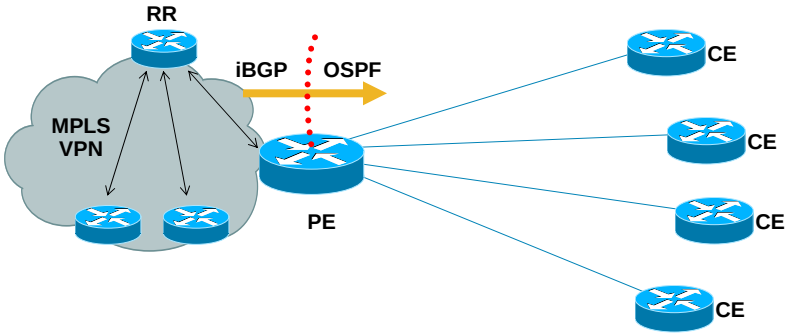
- Maximum number of prefixes (routes) that are allowed to be redistributed into OSPF from other protocols (or other OSPF processes)
- Self-originated LSAs are limited
- Summarized prefixes counted
- Type-7 to Type -5 translated prefixes not counted
- CLI:

```
redistribute maximum-prefix maximum [threshold]
[warning-only]
```

- Availability

  12.0(25)S, 12.3(2)T, 12.2(18)S

OSPF Limit On Number Of Redistributed Routes (Deployment Example)



- PE protected from being overloaded from BGP

Summary

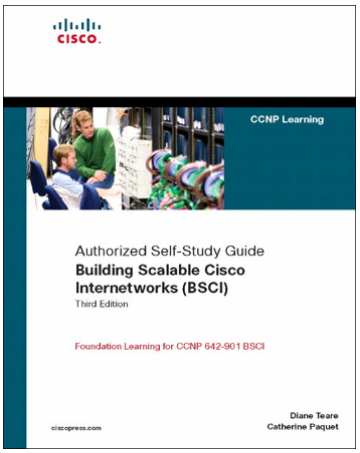
What We Learned?

- Deployment issues in Service Providers and Enterprise Networks
- Dialup deployments techniques
- Design Best Practices
- Understand OSPF fast convergence and resiliency techniques
- OSPF deployments techniques in MPLS environment

Q and A

Recommended Reading

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



Available Onsite at the Cisco Company Store

Complete Your Online Session Evaluation

- Give us your feedback and you could win fabulous prizes. Winners announced daily.
- Receive 20 Passport points for each session evaluation you complete.
- Complete your session evaluation online now (open a browser through our wireless network to access our portal) or visit one of the Internet stations throughout the Convention Center.

Don't forget to activate your Cisco Live virtual account for access to all session material on-demand and return for our live virtual event in October 2008.

Go to the Collaboration Zone in World of Solutions or visit www.cisco-live.com.