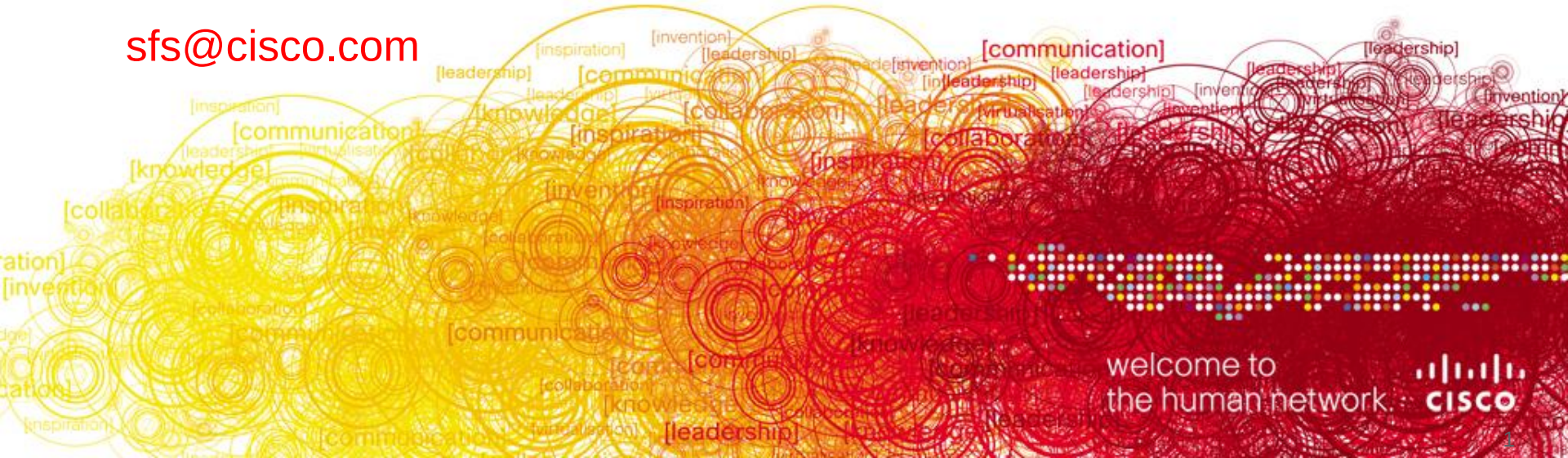
Which Routing Protocol?

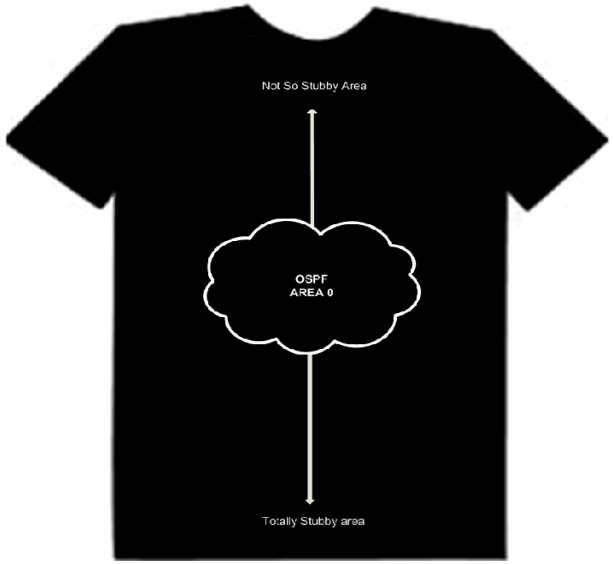
Comparison between OSPF & ISIS

Faraz Shamim

sfs@cisco.com



Is one protocol better than the others? Which routing protocol should I use in my network? Should I switch from the one I'm using? Do the same selection rules apply to IPv4 and IPv6? How will my IPv4 and IPv6 routing protocols coexist?



The Questions

- Is one routing protocol better than any other protocol?
- Define “Better!”

- Converges faster?
- Uses less resources?
- Easier to troubleshoot?
- Easier to configure?
- Scales to a larger number of routers, routes, or neighbors?
- More flexible?
- Degrades more gracefully?
- …

The Questions

- The answer is yes if:

  The network is complex enough to “bring out” a protocol’s specific advantages

  You can define a specific feature (or set of features) that will benefit your network tremendously…

The Questions

- But, then again, the answer is no! ☺
- Every protocol has some features and not others, different scaling properties, etc.
- Let's consider some specific topics for OSPF & ISIS....

Before That … The Twist!

- Most likely the IPv6 IGP will not be deployed in a brand new network and just by itself
- Most likely the existing IPv4 services are more important at first since they are generating most of the revenue
- Redefine “Better!”

- What is the impact on the convergence of IPv4?
- How are the resources shared between the two protocols?
- Are the topologies going to be congruent?
- How easy is it to manage parallel IPv4 / IPv6 environments?
- Opportunity to adapt a new IGP?

Which Routing Protocol

- IPv4 and IPv6 IGPs (OSPF & ISIS)
- Convergence Speed
- Design and Topology Considerations
- Summary

IPv4 and IPv6 IGPs

A comparative overview

“IPv6 is an Evolutionary Not a Revolutionary Step and this is very clear in the case of routing which saw minor changes even though most of the Routing Protocols were completely rebuilt.”

The IPv4 – IPv6 Parallel

OSPF	OSPFv2 for IPv4 OSPFv3 for IPv6 Distinct but similar protocols with OSPFv3 being a cleaner implementation that takes advantage of IPv6 specifics
IS-IS	Extended to support IPv6 Natural fit to some of the IPv6 foundational concepts Supports Single and Multi Topology operation

- For all intents and purposes, the IPv6 IGPs are very similar to their IPv4 counterparts
- IPv6 IGPs have additional features that could lead to new designs

ISIS

High-level overview

High-level perspective	•IS-IS was not designed from the start as an IP routing protocol. •Adjacency is reported once two-way connectivity has been ensured. •IS-IS essentially uses its regular flooding techniques to synchronize neighbors. •Transient routing issues can be reduced (albeit non-deterministically) by judicious use of the “overload” bit.
Encapsulation	IS-IS runs directly over L2 (next to IP) •Sort of makes sense (since it was designed for CLNS) •Does not require a valid interface address to transmit protocol messages. •Agnostic about the type of prefix being transported. •Partition repair requires tunneling (rarely implemented).

ISIS

High-level overview

Database Node	IS-IS database node is an LSPacket •LSPs are clumps of topology information organized by the originating router. •Always flooded intact, unchanged across all flooding hops (so LSP MTU is an architectural constant—it must fit across all links). •Small topology changes always yield entire LSPs (though packet size turns out to be much less of an issue than packet count). •Implementations can attempt clever packing.
Links and Areas	•In IS-IS, if routers do not agree on the area ID, they form L2 adjacency. •Area borders cross links in IS-IS. •In IS-IS, a link can be associated with an L1 and an L2 area simultaneously.

ISIS

Comparative overview

Implementation	Two new TLVs: - IPv6 Reachability TLV (0xEC): Describes network reachability (IPv6 routing prefix, metric information and option bits) - IPv6 Interface Address TLV (0xE8): Contains 128 bit address. Hello PDUs, must contain the link-local address but for LSP, must only contain the non link-local address A new Network Layer Protocol Identifier (NLPID): Allows IS-IS routers to advertise IPv6 prefix payload using 0x8E value
Operational Considerations	**Single Topology** (default for all protocols supported) - potentially beneficial in saving resources (same topology and same SPF) **Multi Topology** (RFC5120) - Independent IPv4 and IPv6 topologies, independent interface metrics **Transition mode available** - both types of TLVs are advertised
Notes	Standardization: draft-ietf-isis-ipv6-07 Evolution: draft-ietf-isis-mi

OSPF

High-level overview

High-level perspective	•OSPF is for the most part more “optimized” (and therefore significantly more complex) •Only LSAs are extensible (not hellos, etc.). •Unrecognized LSA types are not flooded (though opaque LSAs can suffice, if implemented universally). •Uses complex, multistate process to synchronize databases between neighbors. Intended to minimize transient routing problems by ensuring that a newborn router has nearly complete routing information before it begins carrying traffic.
Encapsulation	OSPF runs on top of IP •Traditional IP routing protocol approach •Allows virtual links (if you like them) •Relies on IP fragmentation for large LSAs •Subject to spoofing and DoSattacks (use of authentication is strongly advised).

OSPF

High-level overview

Database Node	OSPF database node is an LS Advertisement •LSAs are mostly numerous and small (one external per LSA, one summary per LSA). •Network and router LSAs can become large. •LSAs are grouped into LS Updates during flooding. •LS Updates are built individually at each hop. •Small changes can yield small packets (but router, network LSAs can be large).
Links and Areas	•An OSPF link can be only in one area, and routers must agree on the area ID. •Area borders cross routers in OSPF.

OSPFv3

Comparative overview

Implementation	Similar Concepts as OSPFv2: - Runs directly over IPv6 (port 89) - Uses the same basic packet types - Neighbor discovery and adjacency formation mechanisms are identical (All OSPF Routers FF02::5, All OSPF DRs FF02::6) - LSA flooding and aging mechanisms are identical - Same interface types (P2P, P2MP, Broadcast, NBMA, Virtual) Independent process from OSPFv2
Important Differences	**OSPFv3 Is Running per Link Instead of per Node** **Support of Multiple Instances per Link:** - New field (instance) in OSPF packet header allows running multiple instances per link - Instance ID should match before packet is being accepted - Useful for traffic separation, multiple areas per link **Generalization of Flooding Scope:** - Three flooding scopes for LSAs (link-local scope, area scope, AS scope) and they are coded in the LS type explicitly

OSPFv3

Comparative overview

Important Differences (cont.)	**Address Semantic Changes in LSA:** - Router and Network LSA carry only topology information - Router LSA can be split across multiple LSAs; Link State ID in LSA header is a fragment ID - Intra area prefixes are carried in a new LSA payload called intra-area-prefix-LSAs - Prefixes are carried in the payload of inter-area and external LSA **Explicit Handling of Unknown LSA:** - The handling of unknown LSA is coded via U-bit in LS type - When U bit is set, the LSA is flooded within the corresponding flooding scope, as if it was understood - When U bit is not set, the LSA is flooded within the link local scope **Authentication Is Removed from OSPF:** - Authentication in OSPFv3 has been removed and OSPFv3 relies now on IPv6 authentication header since OSPFv3 runs over IPv6 - Autype and Authentication field in the OSPF packet header therefore have been suppressed

OSPFv3

Comparative overview

Important Differences (cont.)	**OSPF Packet Format has Been Changed:** - The mask field has been removed from Hello packet - IPv6 prefixes are only present in payload of Link State update packet **Two New LSAs Have Been Introduced:** - Link-LSA has a link local flooding scope and has three purposes Carry IPv6 link local address used for NH calculation Advertise IPv6 global address to other routers on the link (used for multi-access link) Convey router options to DR on the link - Intra-area-prefix-LSA to advertise router's IPv6 address within the area
Notes	**Standardization** Main standard: RFC 5340 Obsoletes 2740 Evolution: draft-ietf-ospf-mt-ospfv3 draft-ietf-ospfv3-af-alt

OSPF LSA Types

Comparative overview

	LSA Function Code	LSA Type
Router-LSA	1	0x2001
Network-LSA	2	0x2002
Inter-Area-Prefix-LSA	3	0x2003
Inter-Area-Router-LSA	4	0x2004
AS-External-LSA	5	0x4005
Group-membership-LSA	6	0x2006
Type-7-LSA	7	0x2007
Link-LSA	8	0x0008
Intra-Area-Prefix-LSA New	9	0x2009

The Version Agnostic Perspective

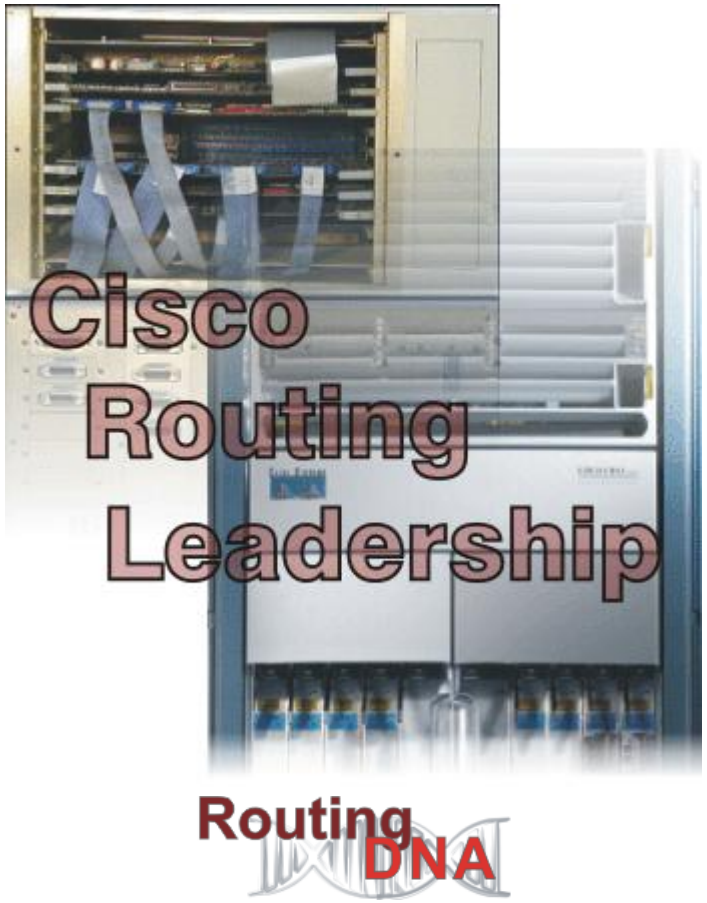
- The similarities between the IPv4 and IPv6 IGP lead to similar network design considerations as far as routing is concerned – **For the rest of the presentation, the analysis is IP version AGNOSTIC!** IPv6 specific considerations are noted where relevant
- The implementation of the IPv6 IGPs achieves parity with the IPv4 counterparts in most aspects but this is an ongoing development and optimization process
- Coexistence of IPv4 and IPv6 IGPs is a very important design consideration.

Convergence Speed

Scenarios and Considerations

Convergence Speed

- Equal Cost Convergence
- Link State Convergence
- Convergence Summary



Convergence Speed

- Which protocol converges faster?
- IS-IS vs OSPF

  IS-IS and OSPF have the same characteristics, from a high level, so we'll consider them both as link state

- Rules of Thumb

  The more routers involved in convergence, the slower convergence will be

  The more routes involved in convergence, the slower convergence will be

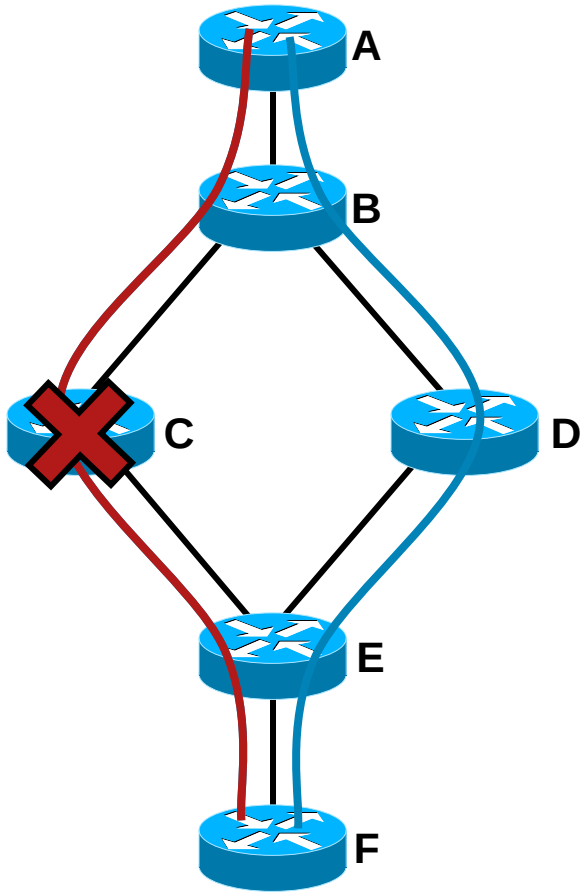
Convergence Speed

- Five steps to convergence
  1. Detect the failure
  2. Flood the failure information
  3. Calculate new routes around the topology change
  4. Add changed routing information to the routing table (RIB)
  5. Update the FIB (possibly distributed)
- Steps 1-4-5 are similar for any routing protocol, so we'll only look at steps 2-3
- *But, it's important to keep in mind steps 1-4-5, since they often impact convergence more than the routing protocol does*

Equal Cost

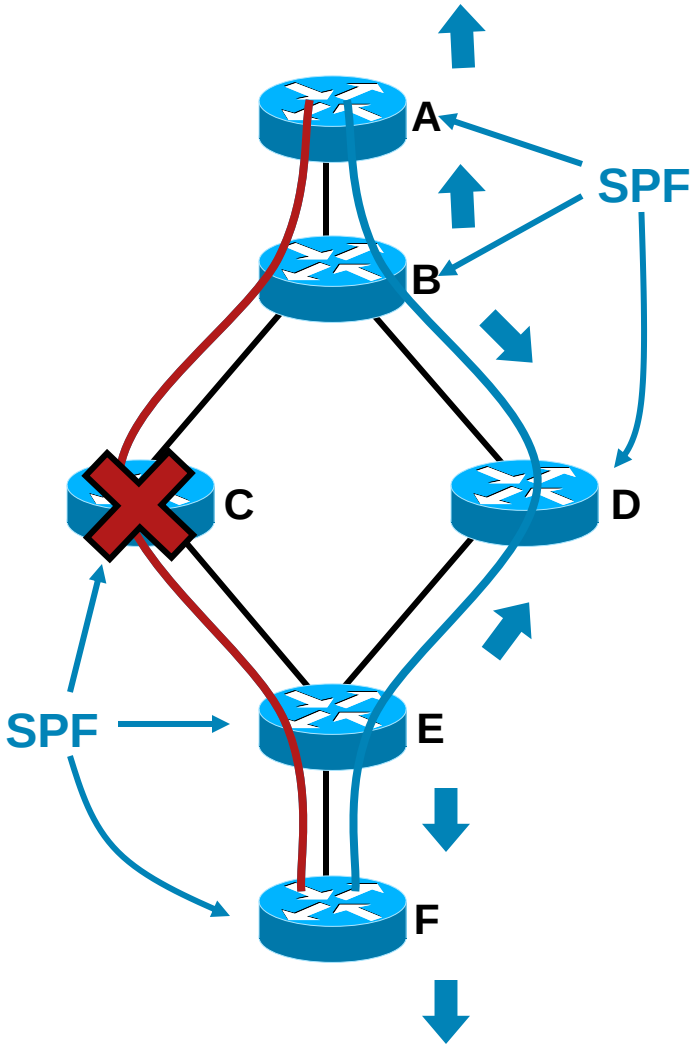
Refresher

- Start with B>C>E and B>D>E being equal cost
- If C fails, B and E can shift from sharing traffic between C and D to sending traffic to D only
- Number of routers involved in convergence: 2 *(B and E)*
- Convergence time is in the milliseconds



Link State

- C fails
- B and E flood new topology information
- All routers run SPF to calculate new shortest paths through the network
- B and E change their routing tables to reflect the changed topology
- Number of routers involved in convergence: 2 *(B and E)*

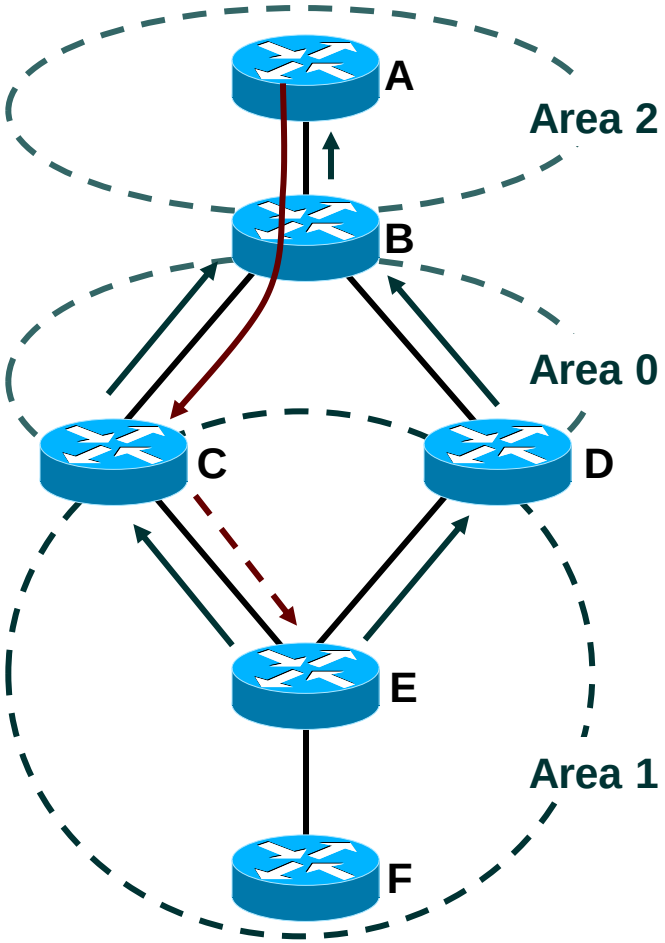


Link State

- Within a single flooding domain
  - A single area in OSPF
  - A single flooding domain in IS-IS
- Convergence time depends on flooding timers, SPF timers, and number of nodes/leafs in the SPF tree
- What happens when we cross a flooding domain boundary?

Link State

- E floods topology changes to C and D
- C and D summarize these topology changes (removing the topology information), and flood it to B
- B builds a summary from the summary flooded to B, and floods it into area 2
- A calculates a route to B, then recurses C onto E



Link State

- Between flooding domains, link state protocols have “distance vector” characteristics
- This can have negative or positive impacts on convergence time in a large network
  - Reduces tree size
  - Allows partial SPFs, rather than full SPFs
  - Introduces translation and processing at the flooding domain boundaries
- The impact is primarily dependant on the network design

Link State Fast Convergence

OSPF

- Carrier Delays
- Hello/dead timers (fast hellos)
- Bidirectional Forwarding Detection(BFD)
- LSA packet pacing
- Interface event dampening
- Exponential throttle timers for LSA & SPF
- MinLSArrivalInterval
- Incremental SPF

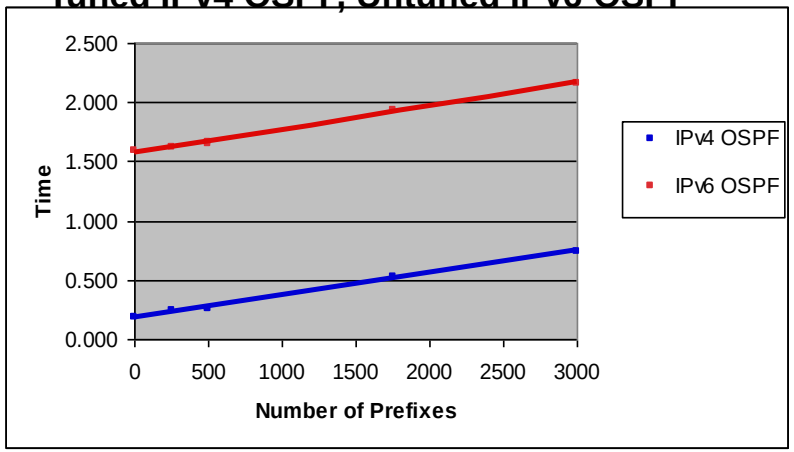
ISIS

- Carrier Delays
- Hello/dead timers (fast hellos)
- Bidirectional Forwarding Detection (BFD)
- LSP pacing
- Interface event dampening
- Exponential throttle timers for LSA & SPF
- PRC interval
- Incremental SPF

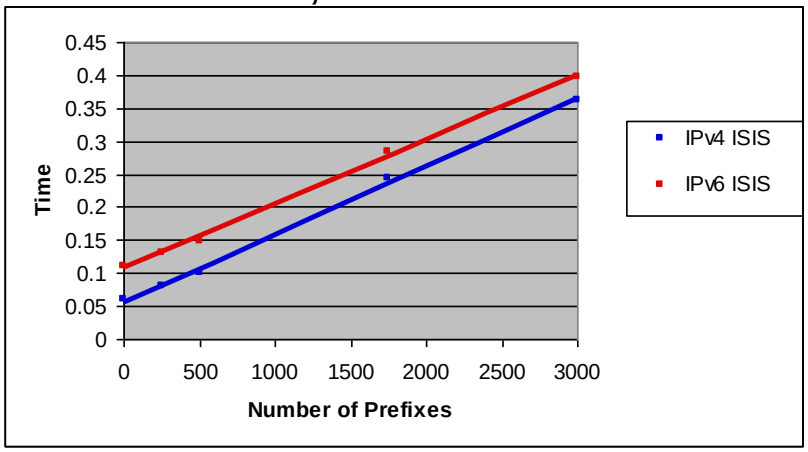
Link State – Convergence Data

- Convergence time with default timers and tuned timers
- IPv4 and IPv6 IGP convergence times are similar
  - The IPv6 IGP implementations might not be fully optimized yet
  - Not all Fast Convergence optimizations might be available

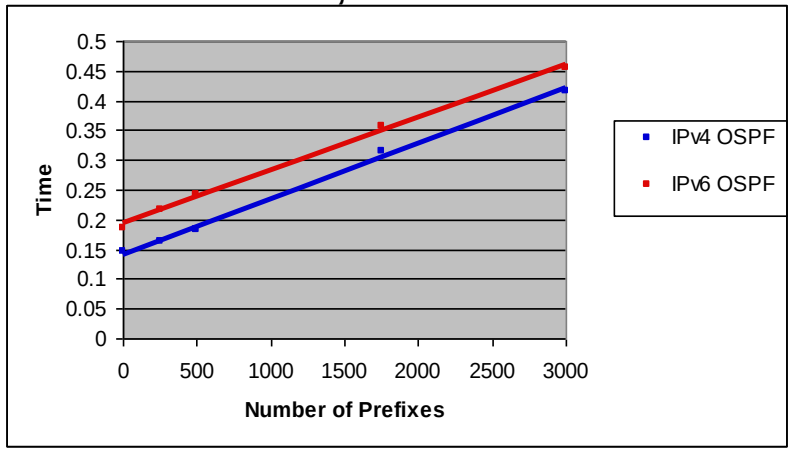
Tuned IPv4 OSPF, Untuned IPv6 OSPF



Tuned IPv4 ISIS, Tuned IPv6 ISIS



Tuned IPv4 OSPF, Tuned IPv6 OSPF



Link State Convergence

Summary

- Within a flooding domain

  The average convergence time, with default timers, is going to be in the order of seconds

  With fast timers, the convergence time can be in the 100s of milliseconds

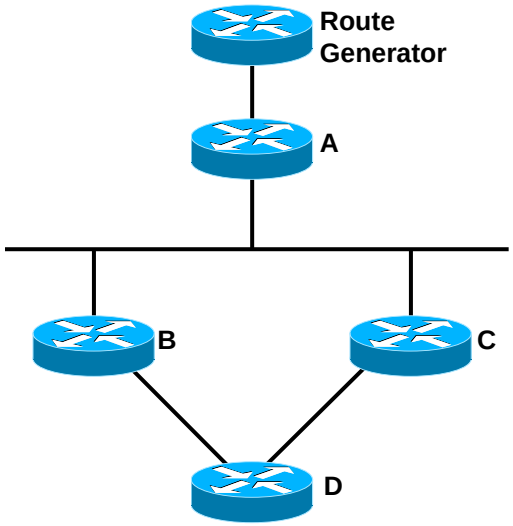
  Note: There are operational 200 node IS-IS and OSPF networks with 500 millisecond convergence times

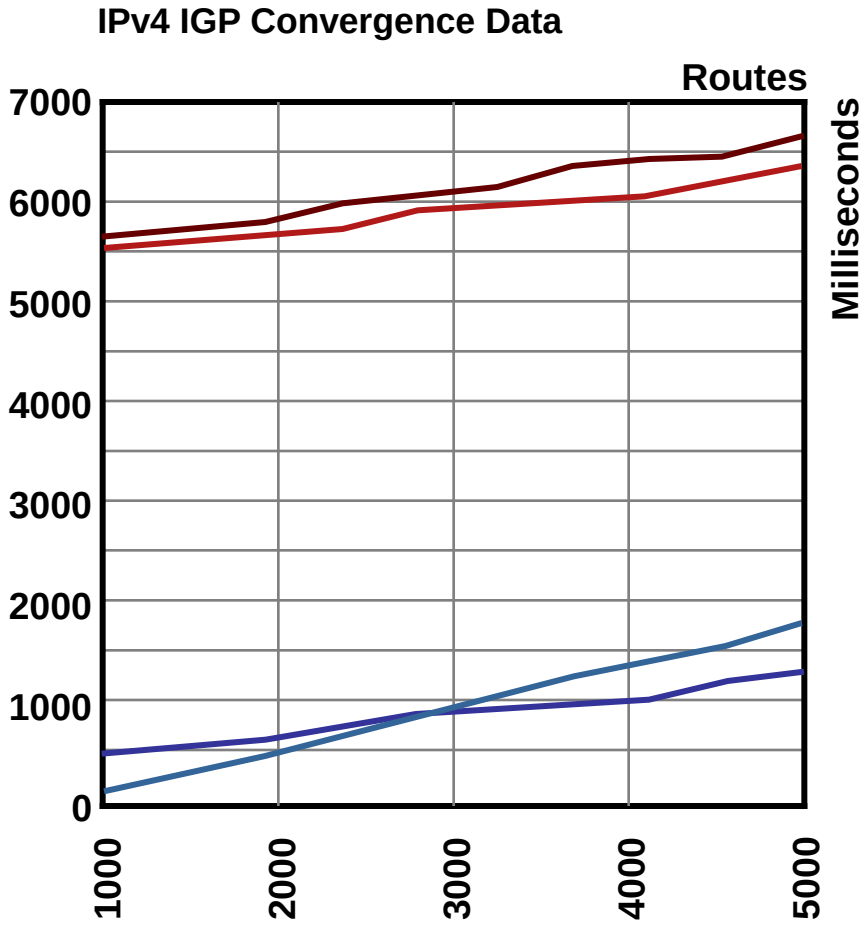
- Outside the flooding domain

  Network design and route aggregation are the primary determining factors of convergence speed

Convergence Summary

- IS-IS with default timers
- OSPF with default timers
- OSPF with tuned timers
- IS-IS with tuned timers





Convergence Summary

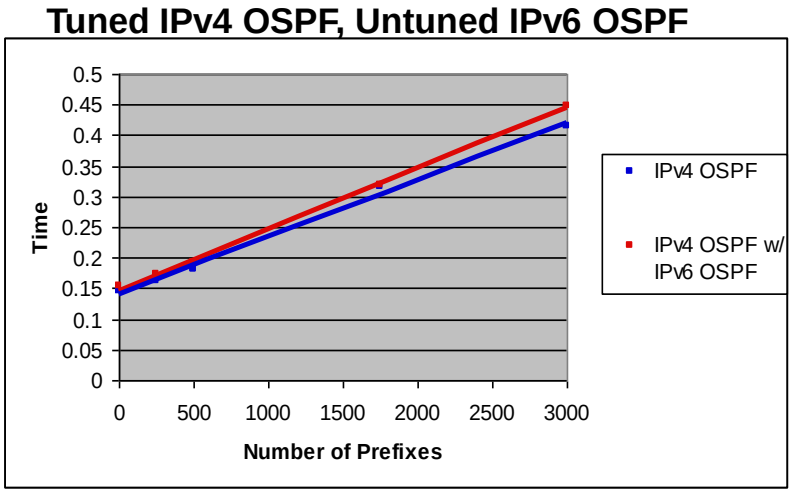
- It's possible to converge in under one second using any protocol, with the right network design
- Rules of Thumb:

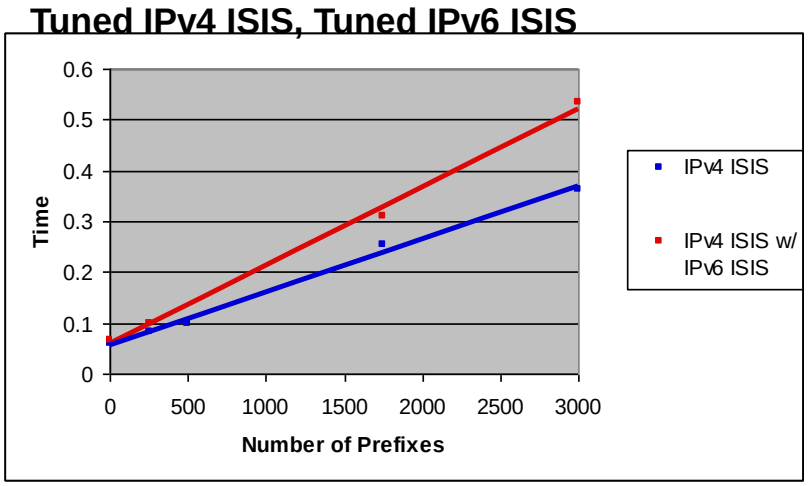
  Less aggregation leads to better performance for link state protocols

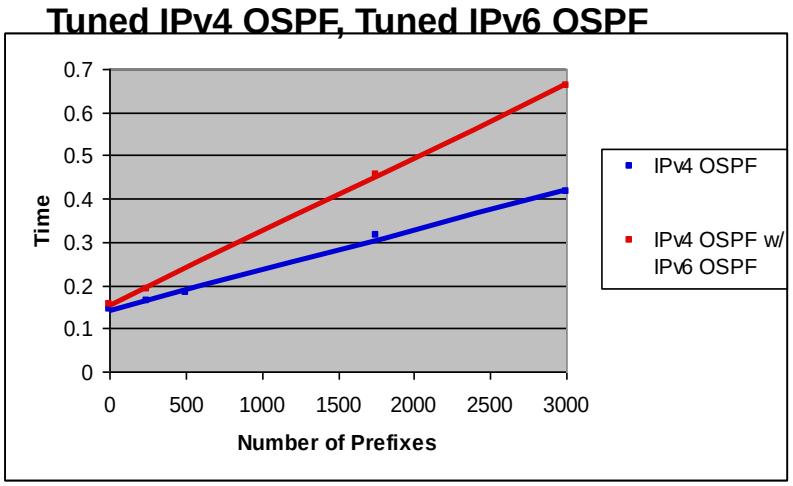
  If you're going to use link state protocols, *tune the timers; but if you tune the timers, be careful with HA features, like GR/NSF*

The Coexistence Twist

- IPv6 IGP impact on the IPv4 IGP convergence
- Aggressive timers on both IGPs will highlight competition for resources
- Is parity necessary from day 1?





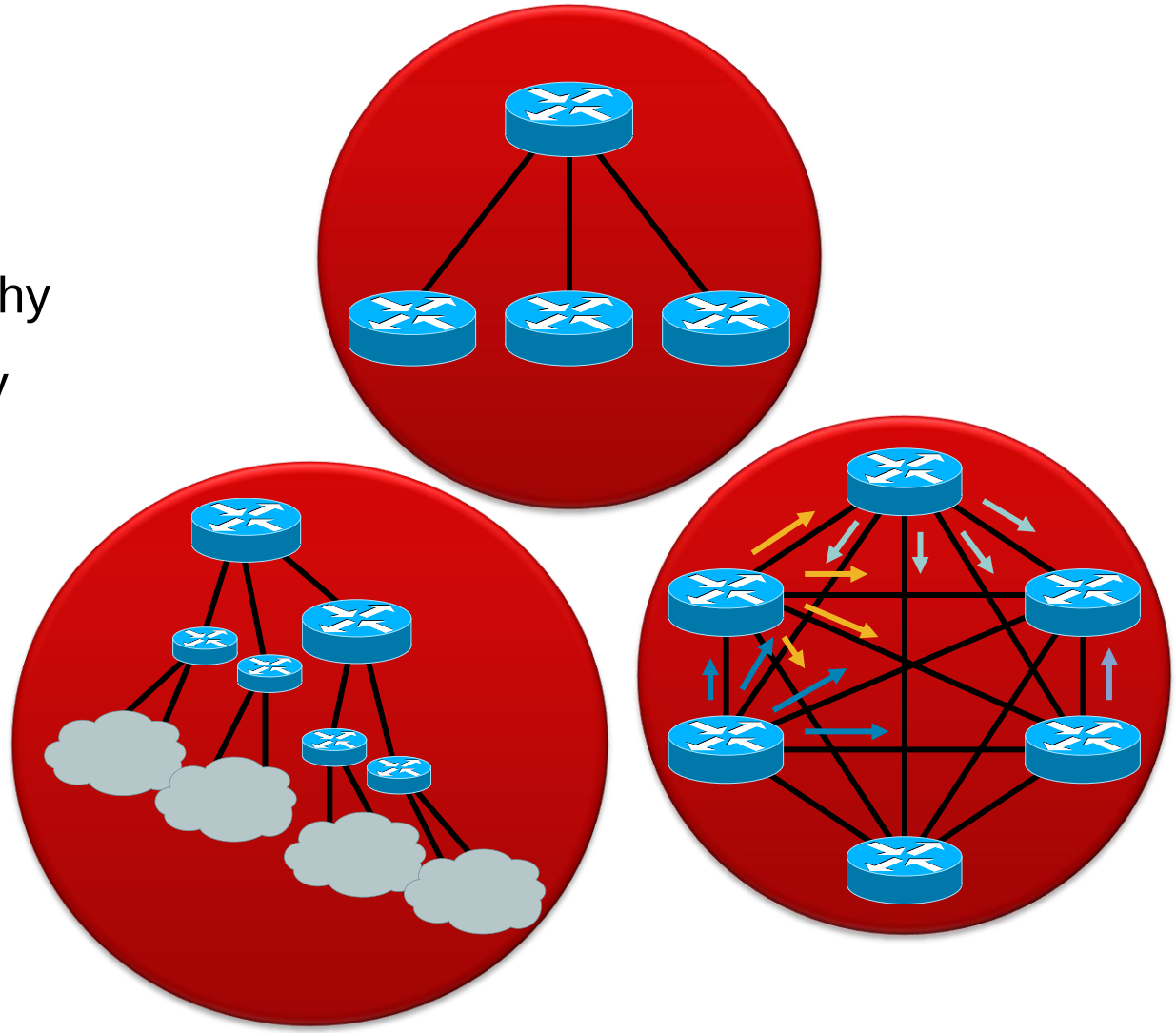


Design and Topology

Considerations

Topology

- Hub and Spoke
- Full Mesh
- Support for Hierarchy
- Topology Summary

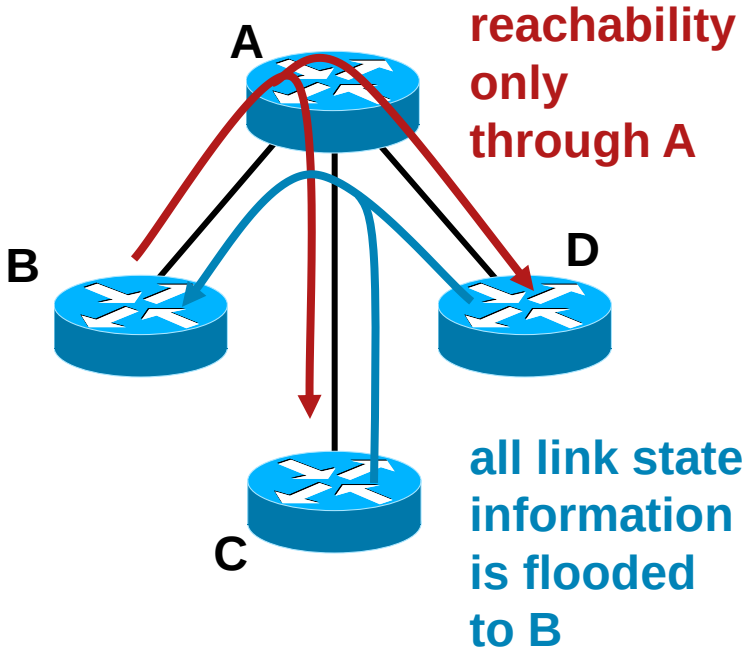


Link State Hub and Spoke

- OSPF and IS-IS are similar when designing for hub and spoke topologies, so we'll look at them together
- Link state protocols rely on every router within a flooding domain having the same view of the network's topology to calculate loop free paths
- Link state flooding rules have implications for scaling and design in hub and spoke networks

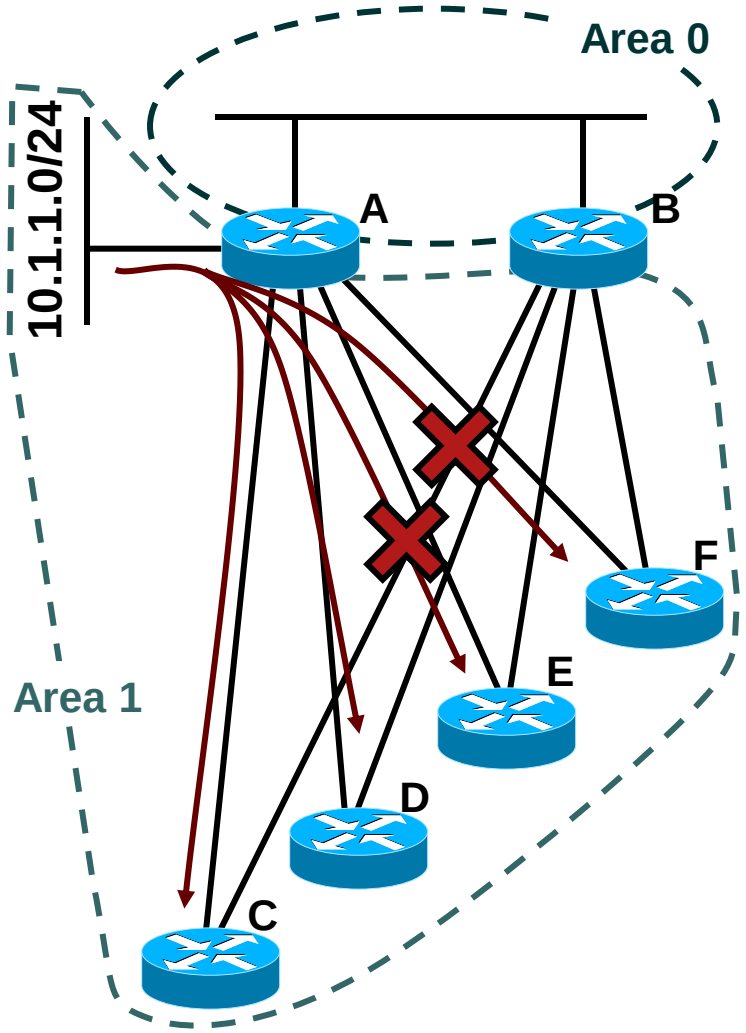
Link State Hub and Spoke

- Although B can only reach C through A, it still receives all of C's routing information
- As the number of remote sites increases, the amount of information each remote site must process and store also increases
- This limits scaling in link state hub and spoke networks



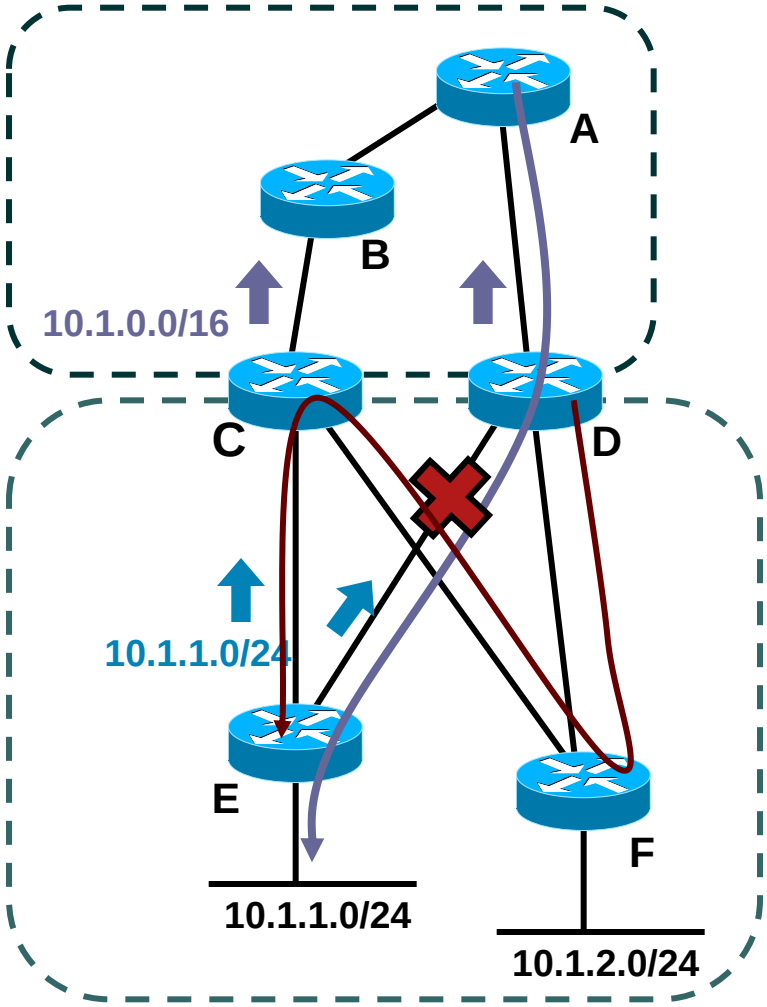
Link State Hub and Spoke

- Controlling route distribution
- There's no way to allow C and D to receive information about 10.1.1.0/24, and not E and F



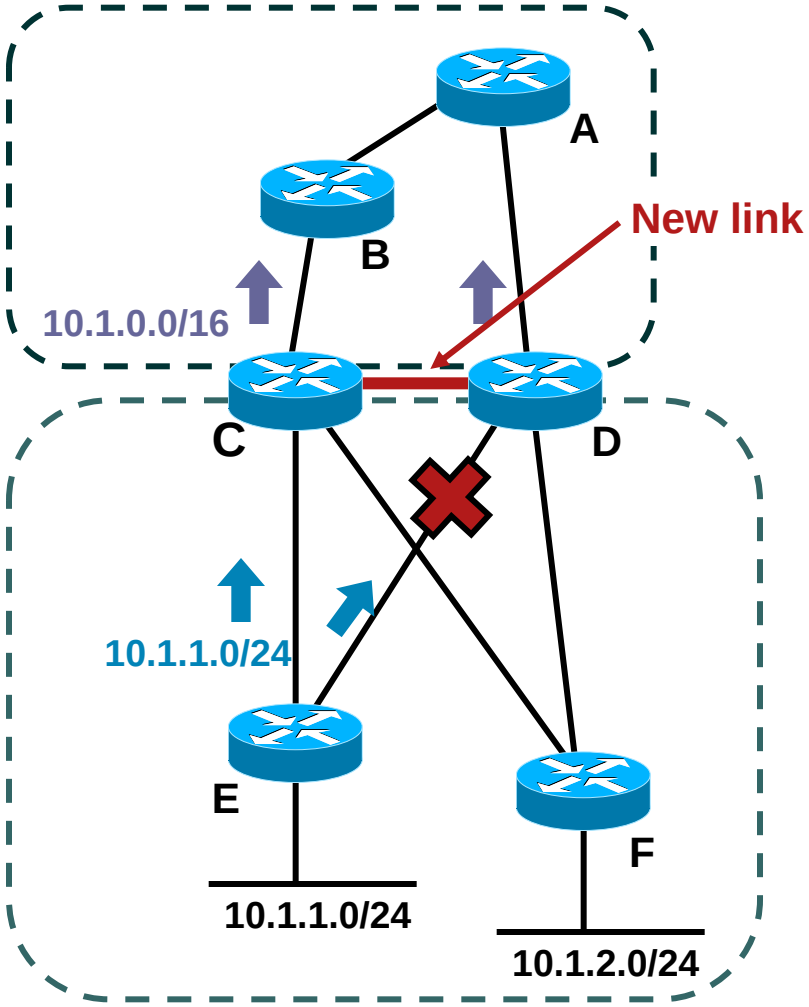
Link State Hub and Spoke

- Transiting remote sites
- C and D issue summaries containing 10.1.1.0/24
- A chooses D as it's best path to the summary
- The D to E link fails
- How can we prevent D from using the link through F to reach 10.1.1.0/24?



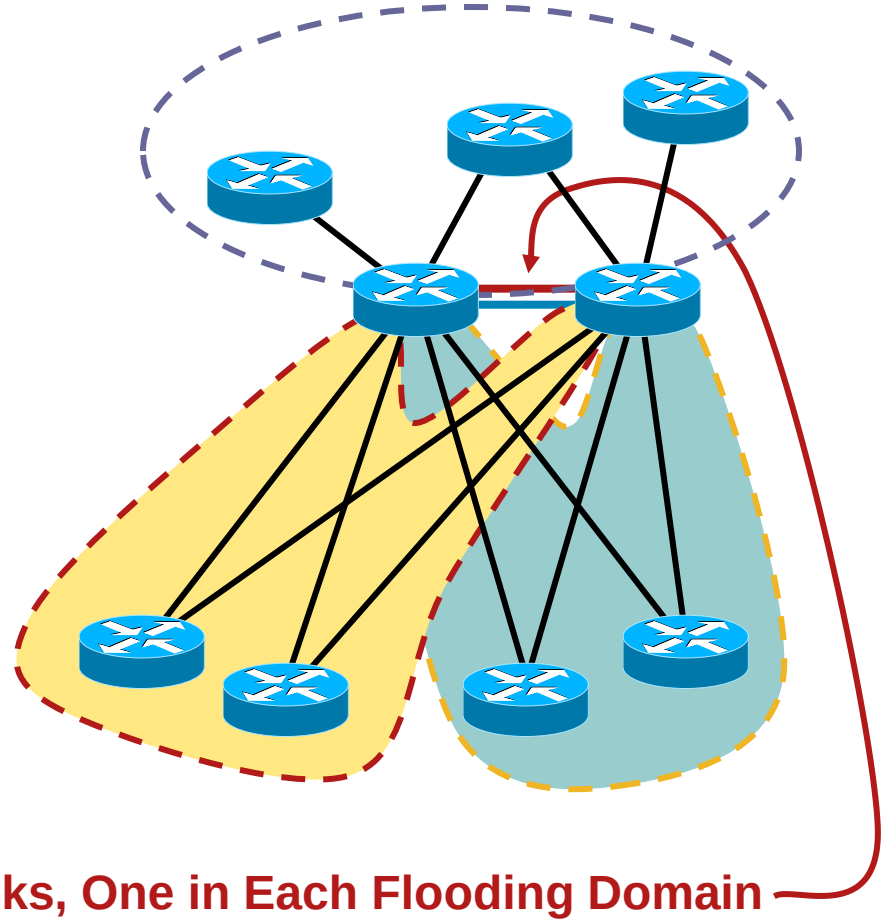
Link State Hub and Spoke

- Place a link between C and D within the same area as the hub and spoke network
- The link cost between C and D should be lower than the link cost through F, causing D to route through this new link



Link State Hub and Spoke

- For each hub and spoke flooding domain you add to the hub routers, you need an additional link between the hub routers in that domain
- You can use virtual links, such as Ethernet VLANS
- This can become difficult to manage in a large scale hub and spoke network

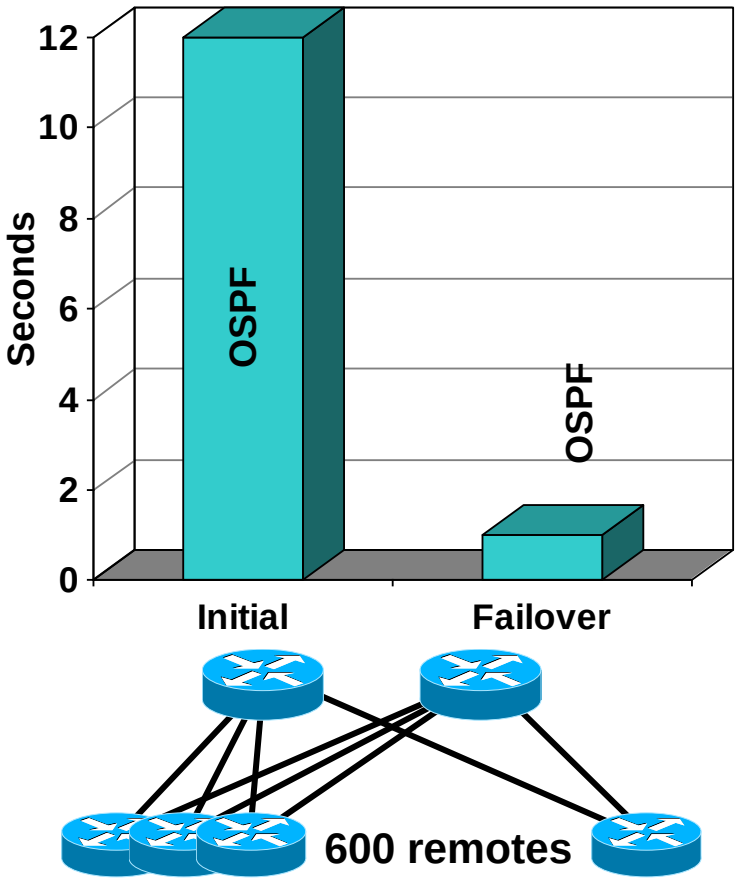


Hub and Spoke Summary

	Scaling	Issues
Link State	All remote sites receive all other remote site link state information; moderate scaling capability	No effective means to control distribution of routing information
		Care must be taken to prevent transiting traffic through remote sites

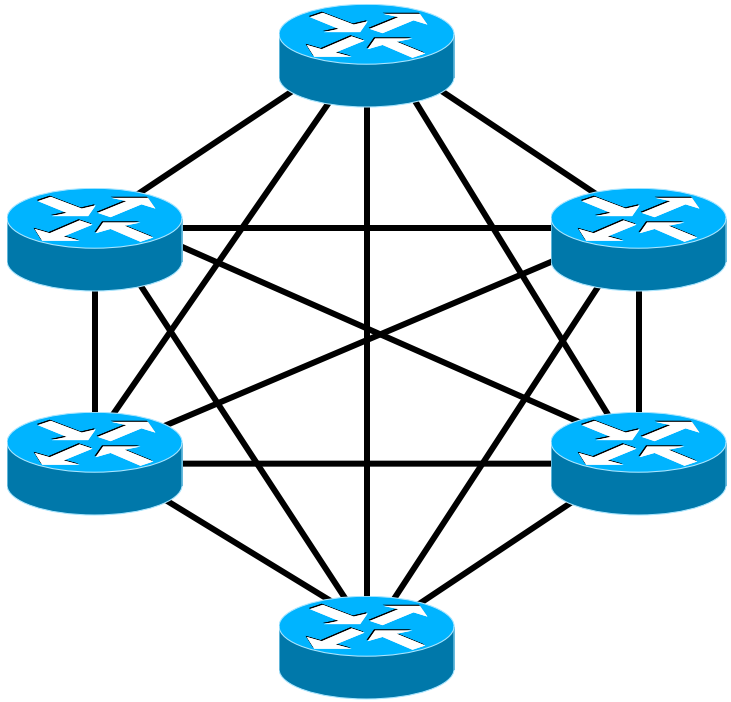
Hub and Spoke

- In the field, we see up to 250 dual homed remotes with OSPF
- Tested initial convergence and hard failover times
  - 600 dual homed remote sites
  - For hard failover, primary hub was powered down
- Testing is still ongoing in this area



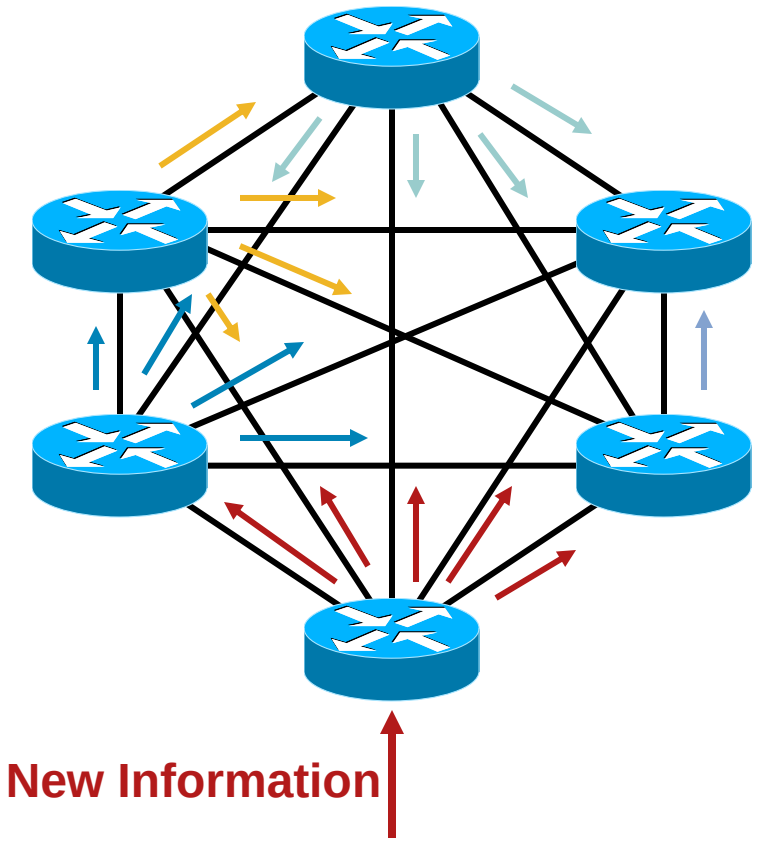
Full Mesh

- 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
  - …
- Adjacencies = links(links-1)/2



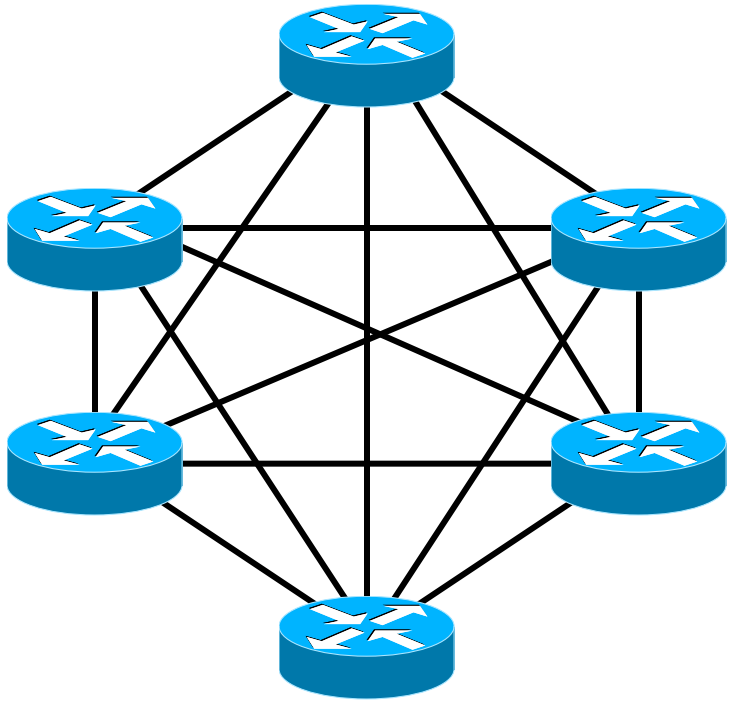
Link State Full Mesh

- Flooding routing information through a full mesh topology is also complicated
- Each router will, with optimal timing, receives at least one copy of every new piece of information from each neighbor on the full mesh
- There are several techniques you can use to reduce the amount of flooding in a full mesh



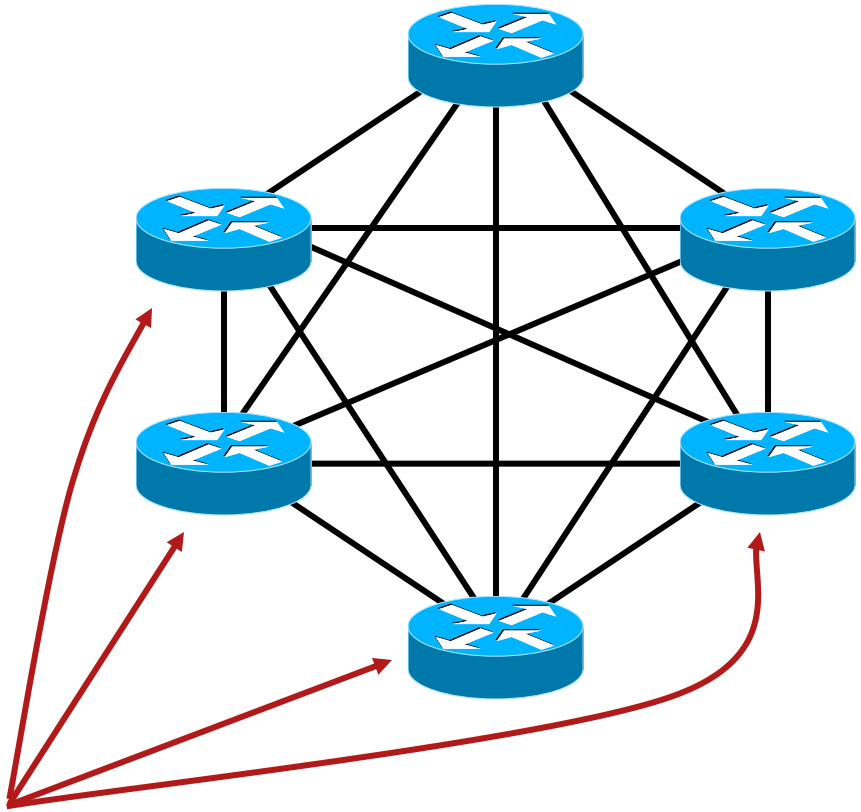
Link State Full Mesh

- OSPF and IS-IS can both use mesh groups to reduce the flooding in a full mesh network
- Mesh groups are manually configured “designated routers” on the full mesh



Link State Full Mesh

- Pick one or two routers to flood into the mesh, and block flooding on the remainder
- This will reduce the number of times information is flooded over a full mesh topology
- This isn't a commonly used configuration

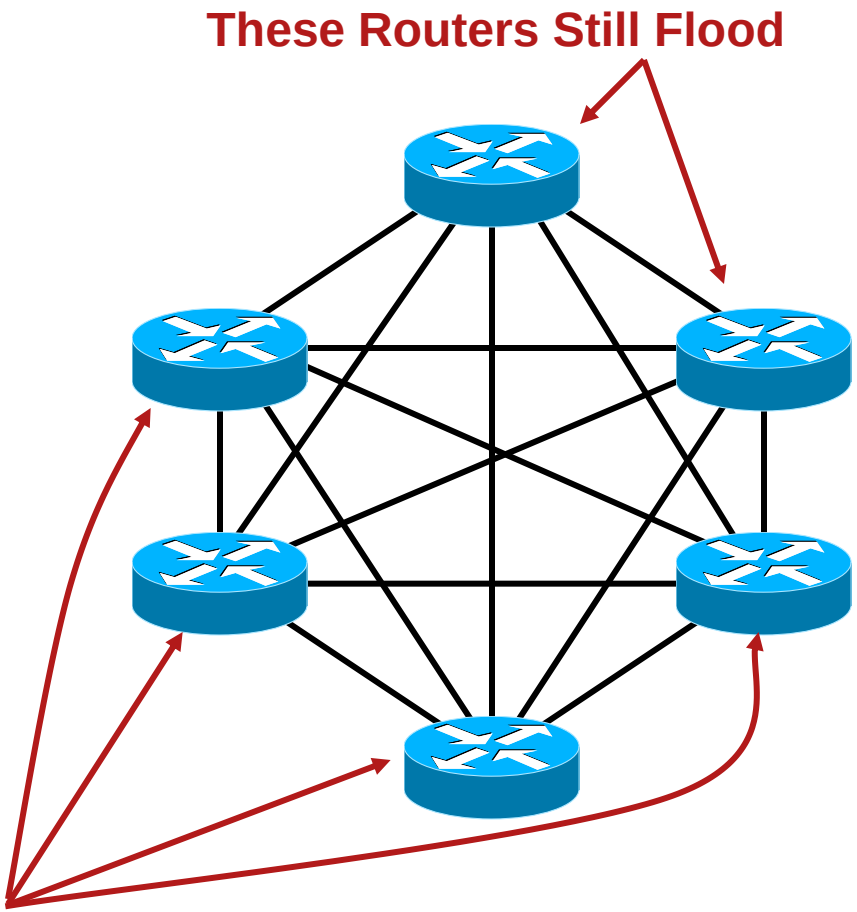


on each serial interface:

```
interface serial x
 ip ospf database-filter all out
 ....
```

Link State Full Mesh

- In IS-IS, each interface is placed in a mesh-group
- Any LSPs received will not be retransmitted back out any other interface on the router in the same mesh-group
- To block all LSP flooding out of an interface, use isis mesh-group blocked
- This isn't a commonly used configuration



On Each Serial Interface:

```
interface serial x
 isis mesh-group 1
 ....
```

Full Mesh Summary

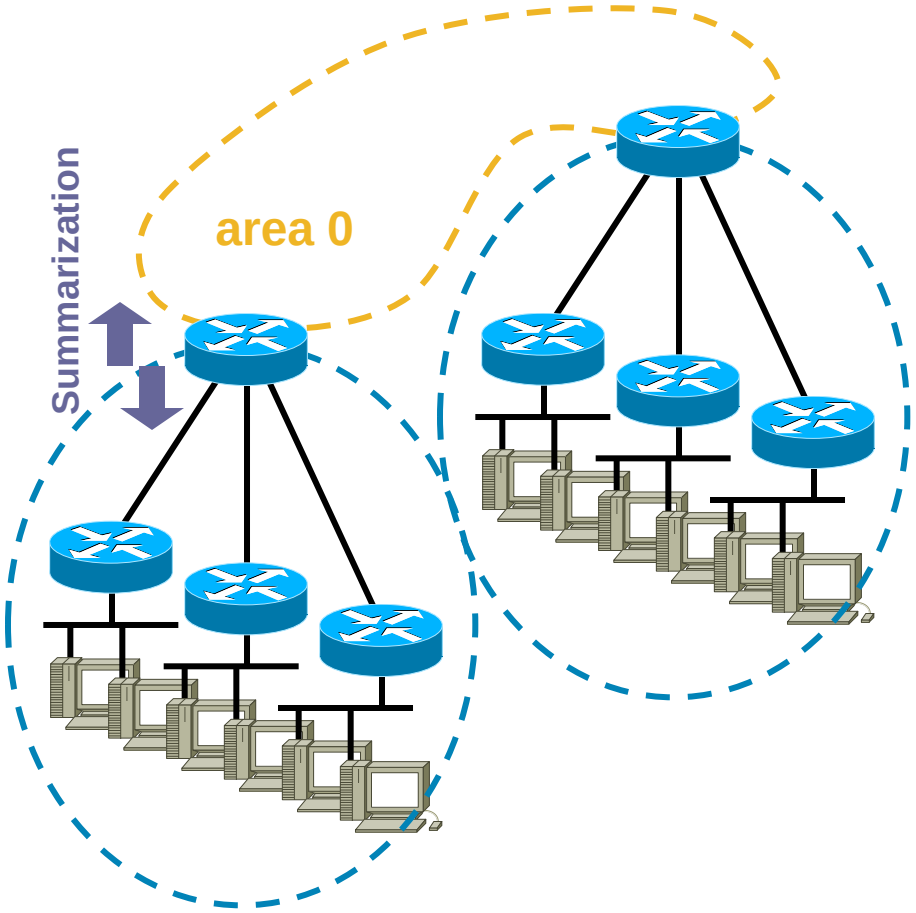
OSPF	Use `ip ospf database-filter all out` to Manually Designate Flooding Points and Increase Scaling Through a Full Mesh
IS-IS	Use `isis mesh-group` or `isis mesh-group blocked` to Manually Designate Flooding Points and Increase Scaling Through a Full Mesh

OSPF Support for Hierarchy

- OSPF has a “hard edge” at flooding domain borders
- Summarization and filtering can occur at this border

  Summarization and filtering can also be configured at routers redistributing routes into OSPF

- In a two layer hierarchy, the flooding domain border naturally lies on the aggregation/core boundary

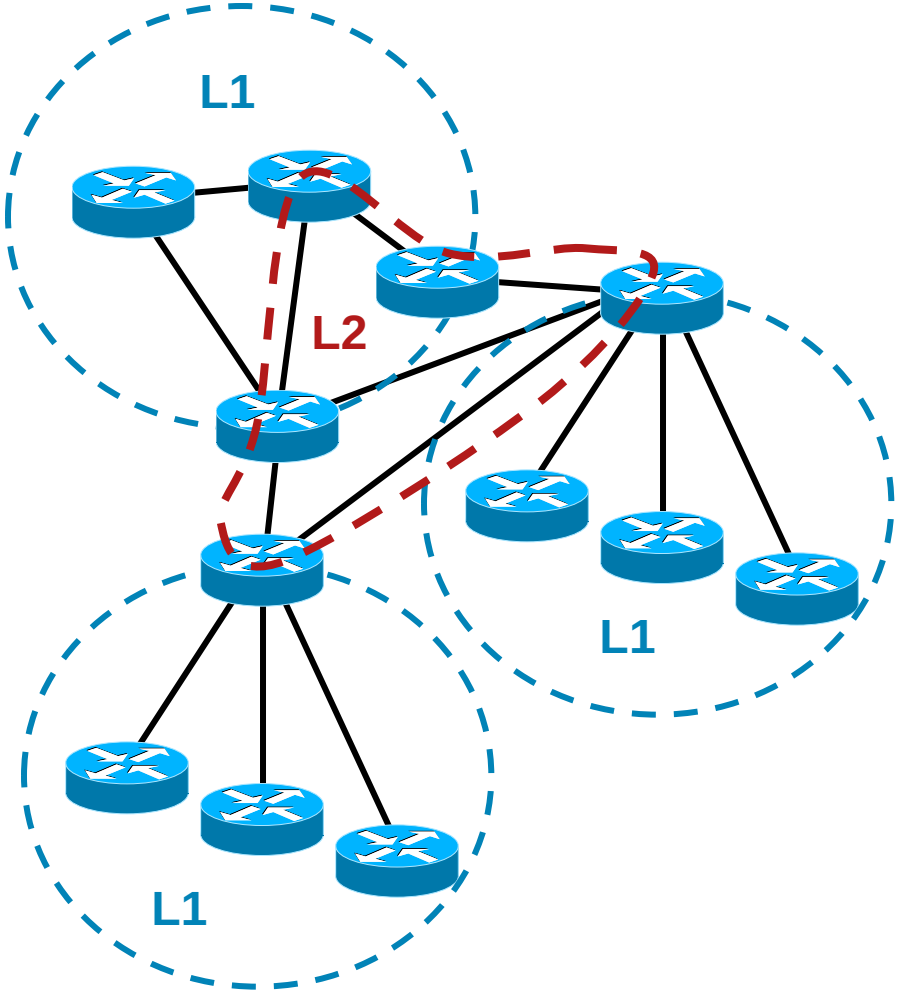


IS-IS Support for Hierarchy

- IS-IS has a “hard edge” at flooding domain borders, as well, but it’s softer than OSPF’s because the L2 routing domain can (and normally does) overlap with the L1 domains
- Summarization and filtering can occur at this border

  Summarization and filtering can also be configured at redistribution points

- In a two layer hierarchy, the flooding domain border naturally lies on the aggregation/core boundary



Hierarchical Division Points Summary

OSPF	“Hard” flooding domain, summarization, and filter border; area borders need to be considered when designing or modifying the network
IS-IS	“Softer” flooding domain, summarization, and filtering border; L2 overlaps L1 domains, providing some flexibility; network design needs to consider flooding domain border

Topology Summary

- Rules of Thumb

  Link state protocols perform better in full mesh environments, *if tuned correctly*

  Link State Protocols tends to perform better in flatter networks

  Note: With IPv6 a great deal of emphasis is placed on hierarchical addressing schemes.

The Coexistence Twist

Multi-Topology vs Multi-Instance

- Multi-Topology
  - Single IS-IS domain with set of independent IP topologies
  - Common flooding and resources associated with both the router and network
  - Multiple SPF
  - Large database.
- Multi-Instance
  - Multiple instances of protocols on a given link
  - Enhances the ability to isolate the resources associated with both router and network
  - Instance-specific prioritization for PDUs and routing calculations.

The Coexistence Twist

Multi-Topology vs Multi-Instance

Multi Process/Topo	Single Process/Topo*
• Clear separation of the two control planes • Non-congruent topologies are very common if not desired in deployments	• Requires less resources • Might provide a more deterministic co-existence of IPv4 and IPv6

*Today most IPv6 IGPs are distinct from their IPv4 counterparts and will run as ships in the night. The only exception is ISIS.

The Coexistence Twist

Multi-Topology vs Multi-Instance

- OSPF currently is based on multi-instance:
  - Adding multi-topology should not be difficult for OSPFv3
  - Multiple-address family support is already (draft) here; just a minor extension for multi-topology needs to be added.
- ISIS
  - Multi-topology support has been available for a while
  - Multi-instance draft is available for IS-IS now.
- Which one is better?

Summary


[collaboration]
[leadership]
[inspiration]
[virtualisation]
[communication]
[knowledge]
[invention]

Is one protocol better than the others? Which routing protocol should I use in my network? Should I switch from the one I'm using? Do the same selection rules apply to IPv4 and IPv6? How will my IPv4 and IPv6 routing protocols coexist?

Did we answer this question???

Summary

- There is no “right” answer!
- Consider:
  - Your business requirements
  - Your network design
  - The coexistence between IPv4 and IPv6
  - Intangibles
- These two advanced IGP’s are generally pretty close in capabilities, development, and other factors

Expertise (Intangible)

- What is your team comfortable with?
- What “escalation resources” and other support avenues are available?
- But remember, this isn’t a popularity contest—you don’t buy your car based on the number of a given model sold, do you?
- An alternate way to look at it: what protocol would you like to learn? ☺

Standardization (Intangible)

- Who's standard?

  OSPF: Standardized by the IETF

  IS-IS: Standardized by the ISO and the IETF

- Standardization is a tradeoff:

  Promises Interoperability

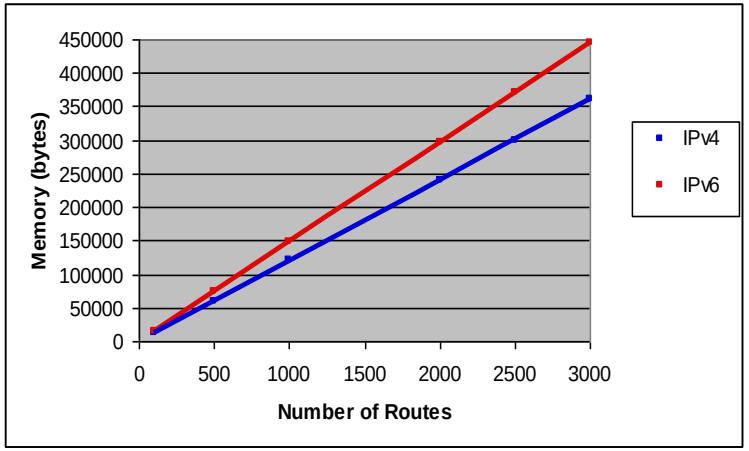
  Larger number of eyes looking at problems and finding new features

  Politics often influence standards and causes compromises

  New features are often difficult to push through standards committees, slowing their release

IPv4/IPv6 Coexistence

- Targeting parity is natural but consider the tradeoffs during the early phases of integration
- IPv4 and IPv6 can be decoupled offering a unique opportunity to try a new design with IPv6. Look at both congruent and non-congruent topology approaches
- Evaluate the additional resources required by IPv6
- Take advantage of the IPv6 addressing resources!



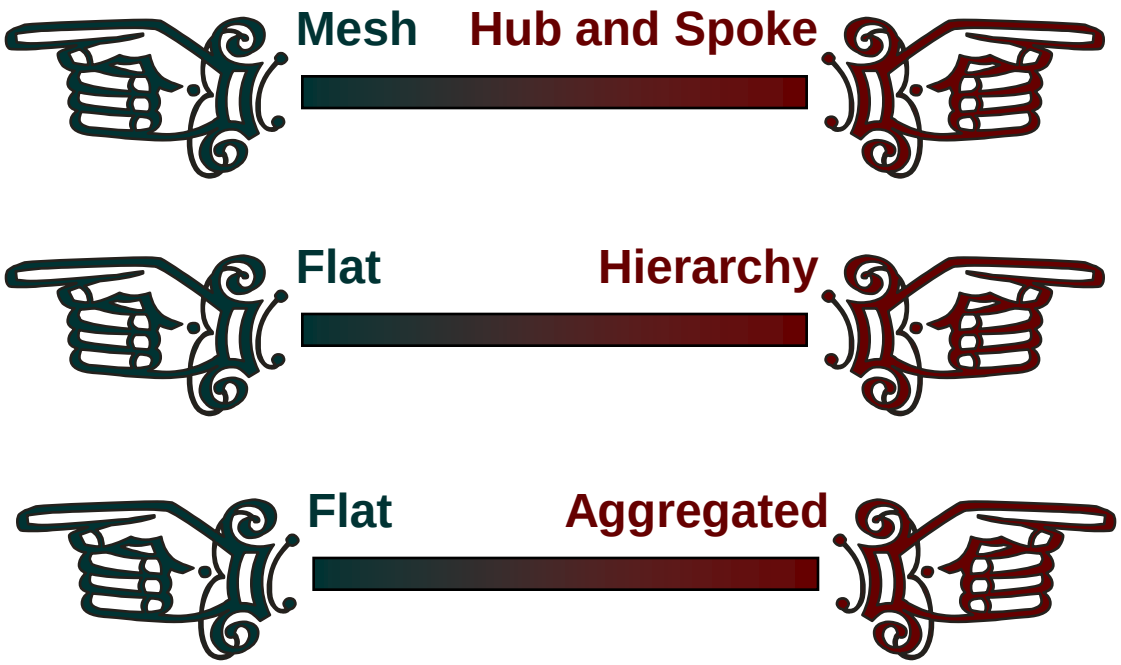

```
show route afi-all summary

IPv4 Unicast:
---------------
Route Source     Routes     Backup     Deleted     Memory (bytes)
connected        5          1          0           720
local            6          0          0           720
local SMIAP      1          0          0           120
static           0          0          0           0
ospf 200         3770       1          0           452520
Total            3782       2          0           454080

IPv6 Unicast:
---------------
Route Source     Routes     Backup     Deleted     Memory (bytes)
connected        3          1          0           592
local            4          0          0           592
ospf 200         3769       1          0           557960
Total            3776       2          0           559144
```

Summary

IP Version Agnostic Rules of Thumb



Q and A


[collaboration]
[leadership]
[inspiration]
[virtualisation]
[communication]
[knowledge]
[invention]