IS-IS and OSPF: Network Design Comparisons and Considerations

Jeff Doyle

IP FOCUS 2001

Juniper NETWORKS®

Objectives

- Understand protocol similarities and differences
- Understand protocol strengths and weaknesses
- Make more informed design decisions

Agenda

- Overview of link-state protocols
- A parallel history of IS-IS and OSPF
- Comparative analysis of IS-IS and OSPF
- Design considerations
- Less-tangible considerations
- Conclusions

Agenda

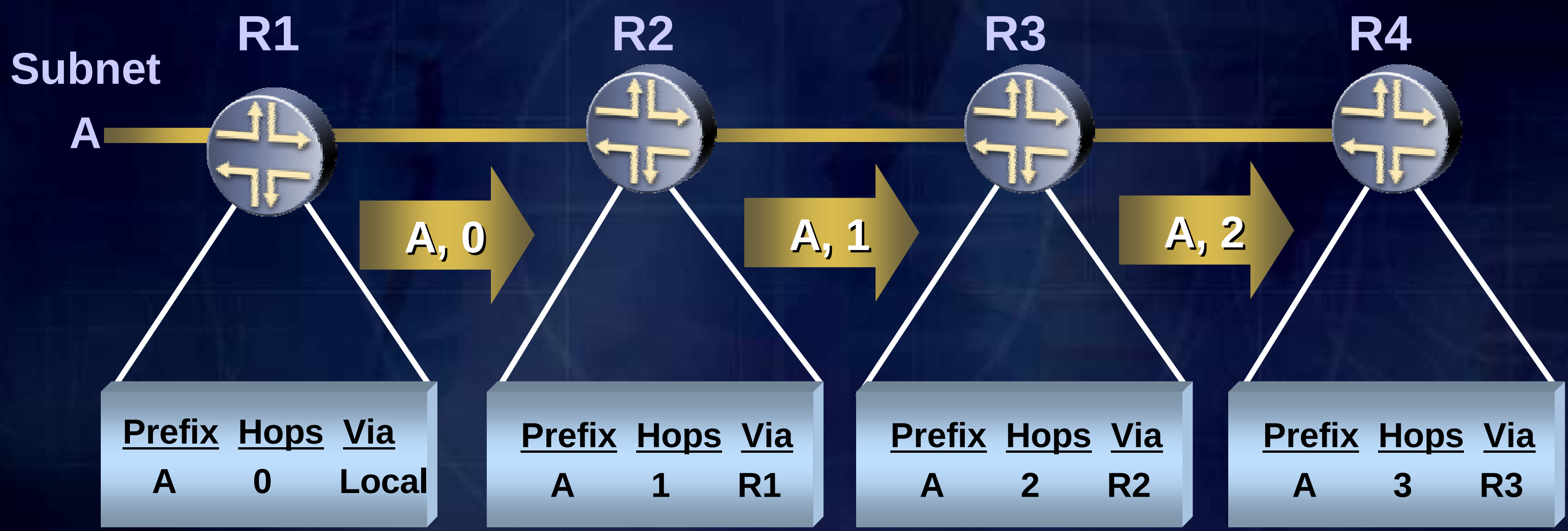
- Overview of link-state protocols
- A parallel history of IS-IS and OSPF
- Comparative analysis of IS-IS and OSPF
- Design considerations
- Less-tangible considerations
- Conclusions

In the Beginning Was Distance Vector...

- Also known as Bellman-Ford, Ford-Fulkerson
- Very simple algorithm
- Distance Vector Protocols include
  - RIP
  - BGP (but usually called Path Vector)
  - Cisco's IGRP
  - Cisco's EIGRP

Routing by Rumor

- Distributed calculation
- Each router knows only what its neighbor tells it



IP FOCUS 2001

Juniper NETWORKS

Problems with Distance Vector

- Slow convergence
  - A direct result of the distributed calculation
  - Triggered updates help
  - Kludges such as hold-down timers reduce transient errors, but increase convergence time
- Single-hop routing loops
  - Solution: split horizon
- Counting to infinity
  - Solution: make infinity finite
- Synchronized periodic updates
  - Solution: update jitter timers

Link-state Protocols

- Also known as shortest path, Dijkstra
- Algorithm based on graph theory, providing better loop avoidance
- Local computation means faster convergence
- Link-state protocols include
  - OSPF
  - IS-IS
  - ATM PNNI
  - IBM APPN
  - MPLS CSPF

Fundamental Link-state Concepts

- Adjacency
- Information flooding
- Link-state database
- SPF calculation

SPF Calculation Example

Example Topology

Link-state Database

R1-R2, 1
R1-R5, 2
R2-R1, 2
R2-R3, 1
R2-R4, 2
R3-R2, 3
R3-R4, 2
R4-R2, 4
R4-R3, 4
R4-R5, 4
R4-R6, 2
R5-R1, 3
R5-R4, 5
R5-R6, 3
R6-R4, 1
R6-R5, 2

SPF Calculation Example

Link-state Database

R1-R2, 1
R1-R5, 2
R2-R1, 2
R2-R3, 1
R2-R4, 2
R3-R2, 3
R3-R4, 2
R4-R2, 4
R4-R3, 4
R4-R5, 4
R4-R6, 2
R5-R1, 3
R5-R4, 5
R5-R6, 3
R6-R4, 1
R6-R5, 2

Tentative

LS Entry	Cost to Root
R2-R1, 2	2
~~R2-R3, 1~~	~~1~~
R2-R4, 2	2

Path

R2-R2, 0
R2-R3, 1

IP FOCUS 2001

Juniper NETWORKS

SPF Calculation Example

Link-state Database

R1-R2, 1
R1-R5, 2
R2-R1, 2
R2-R3, 1
R2-R4, 2
R3-R2, 3
R3-R4, 2
R4-R2, 4
R4-R3, 4
R4-R5, 4
R4-R6, 2
R5-R1, 3
R5-R4, 5
R5-R6, 3
R6-R4, 1
R6-R5, 2

Tentative

LS Entry	Cost to Root
~~R2-R1, 2~~	~~2~~
R2-R4, 2	2
R3-R4, 2	3

Path

R2-R2, 0
R2-R3, 1
R2-R1, 2

SPF Calculation Example

Link-state Database

- R1-R2, 1
- R1-R5, 2
- R2-R1, 2
- R2-R3, 1
- R2-R4, 2
- R3-R2, 3
- R3-R4, 2
- R4-R2, 4
- R4-R3, 4
- R4-R5, 4
- R4-R6, 2
- R5-R1, 3
- R5-R4, 5
- R5-R6, 3
- R6-R4, 1
- R6-R5, 2

Tentative

LS Entry	Cost to Root
~~R2-R4, 2~~	~~2~~
~~R3-R4, 2~~	~~3~~
R1-R5, 2	4

Path

- R2-R2, 0
- R2-R3, 1
- R2-R1, 2
- R2-R4, 2

SPF Calculation Example

Link-state Database

R1-R2, 1
R1-R5, 2
R2-R1, 2
R2-R3, 1
R2-R4, 2
R3-R2, 3
R3-R4, 2
R4-R2, 4
R4-R3, 4
R4-R5, 4
R4-R6, 2
R5-R1, 3
R5-R4, 5
R5-R6, 3
R6-R4, 1
R6-R5, 2

Tentative

LS Entry	Cost to Root
~~R1-R5, 2~~	~~4~~
~~R4-R5, 4~~	~~6~~
R4-R6, 2	4

Path

R2-R2, 0
R2-R3, 1
R2-R1, 2
R2-R4, 2
R1-R5, 4

SPF Calculation Example

Link-state Database

- R1-R2, 1
- R1-R5, 2
- R2-R1, 2
- R2-R3, 1
- R2-R4, 2
- R3-R2, 3
- R3-R4, 2
- R4-R2, 4
- R4-R3, 4
- R4-R5, 4
- R4-R6, 2
- R5-R1, 3
- R5-R4, 5
- R5-R6, 3
- R6-R4, 1
- R6-R5, 2

Tentative

LS Entry	Cost to Root
~~R4-R6, 2~~	~~4~~
~~R5-R6, 3~~	~~7~~

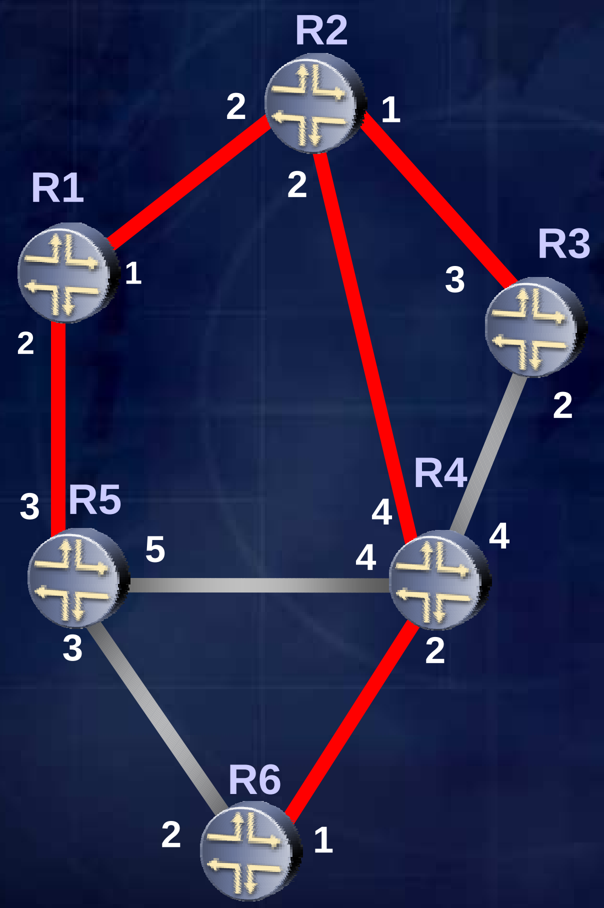
Path

- R2-R2, 0
- R2-R3, 1
- R2-R1, 2
- R2-R4, 2
- R1-R5, 4
- R4-R6, 4

- Tentative is empty
- All nodes in link-state database are in path
- SPF calculation is finished

SPF Calculation Example

- Loop-free, lowest-cost path to every node



Path
R2-R2, 0
R2-R3, 1
R2-R1, 2
R2-R4, 2
R1-R5, 4
R4-R6, 4

Problems with Link State

- Information flooding load
  - Solution: sequence numbers and aging
  - Solution: areas
- Stale LS database entries
  - Solution: periodic database refresh
- .5(n^2-n) adjacencies on multi-access networks
  - Solution: designated routers
- Memory and CPU overload
  - Solution: areas

Agenda

- Overview of link-state protocols
- A parallel history of IS-IS and OSPF
- Comparative analysis of IS-IS and OSPF
- Design considerations
- Less-tangible considerations
- Conclusions

In the Beginning Was DECnet...

- Radia Perlman
- Adopted by ISO for OSI model
- IS-IS extended to support IP
  - Interim solution until OSI makes it extinct (don't hold your breath)
  - RFC 1195
  - Also known as integrated IS-IS, dual IS-IS

IS-IS = 0

- Parallel initiative by IETF to develop an IP routing protocol
- OSPF based on initial work on and experience with IS-IS

Protocol History

- (Disclaimer—biased, foggy memory)
- 1987
  - IS-IS (from DEC) selected by ANSI as OSI intradomain protocol (CLNP only)
- 1988
  - NSFnet deployed, IGP based on early IS-IS draft
  - OSPF work begins, loosely based on IS-IS mechanisms (LS protocols are hard!)
  - IP extensions to IS-IS defined

Protocol History

- 1989
  - OSPF v.1 RFC published
  - Proteon ships OSPF
  - IS-IS becomes ISO proposed standard
  - Public bickering ensues; OSPF and IS-IS are blessed as equals by IETF, with OSPF somewhat more equal
  - Private cooperation improves both protocols
- 1990
  - Dual-mode IS-IS RFC published

Protocol History

- 1991
  - OSPF v.2 RFC published
  - Cisco ships OSPF
  - Cisco ships OSI-only IS-IS
- 1992
  - Cisco ships dual IS-IS (part of DEC Brouter)
  - Lots of OSPF deployed, but very little IS-IS
- 1993
  - Novell publishes NLSP (IPX IS-IS knockoff)

Protocol History

- 1994
  - Cisco ships NLSP (rewriting IS-IS as side effect)
  - Large service providers need an IGP; IS-IS is recommended due to recent rewrite and OSPF field experience (and to lesser extent, NSF CLNP mandate)
- 1995
  - Service providers begin deployment of IS-IS
    - Cisco implementation firms up
    - Protocol starts to become popular in niche

Protocol History

- 1996-1998
  - IS-IS niche popularity continues to grow (some service providers switch to it from OSPF)
  - IS-IS becomes barrier to entry for router vendors targeting large service providers
  - Juniper Networks and other vendors ship IS-IS capable routers
- 1999-2000
  - Extensions continue for both protocols (for example, traffic engineering)

Agenda

- Overview of link-state protocols
- A parallel history of IS-IS and OSPF
- Comparative analysis of IS-IS and OSPF
- Design considerations
- Less-tangible considerations
- Conclusions

ISOspeak 101

- Intermediate System (IS)
- End System (ES)
- Protocol Data Unit (PDU)
- Subnetwork Point of Attachment (SNPA)
- Link State PDU (LSP)
- Routing Domain
- Level 2 Area
- Level 1 Area

Message Types: OSPF

- Hello packet
- Database description packet
- Link-state request packet
- Link-state acknowledgement packet
- Link-state update packet

Message Types: IS-IS

- Hello PDU
- Link-state PDU
- Sequence number PDUs
  - Complete Sequence Number PDU (CSNP)
  - Partial Sequence Number PDU (PSNP)
- Message types are further divided into Level 1 and Level 2

OSPF LSAs

- Multiple LSA types

Type	LSA
1	Router LSA
2	Network LSA
3	Network Summary LSA
4	ASBR Summary LSA
5	AS External LSA
6	Group Membership LSA
7	Not-so-stubby Area LSA
8	External Attributes LSA
9-11	Opaque LSAs

IS-IS Level 1 LSPs

- Single LSP, multiple TLVs

Type	TLV
1	Area Addresses
2	IS Neighbors
3	ES Neighbors
10	Authentication Information
128	IP Internal Reachability Information
129	Protocols Supported
132	IP Interface Address

IS-IS Level 2 LSPs

- Single LSP, multiple TLVs

Type	TLV
1	Area Addresses
2	IS Neighbors
4	Partition Designated Level 2 IS
5	Prefix Neighbors
10	Authentication Information
128	IP Internal Reachability Information
129	Protocols Supported
130	IP External Reachability Information
131	Inter-domain Routing Protocol Information
132	IP Interface Address
135	Extended IP Reachability (wide metrics)

Message Encoding: OSPF

- Runs over IP (protocol number 89)
- 32-bit alignment
- Only LSAs are extensible
- All OSPF speakers must recognize the extensions

Message Encoding: IS-IS

- Runs directly over data link
- No alignment
- All PDUs are extendable
- Nested TLVs

Media Support

- OSPF
  - Broadcast (LANs)
  - Point-to-point
  - Point-to-multipoint
  - NBMA
- IS-IS
  - Broadcast
  - Point-to-point
  - No NBMA support

Router and Area IDs: OSPF

- Router ID and area ID specified separately
- Each is 32-bit number
- AID associated with interface
- RID
  - Explicitly specified RID
  - Loopback address
  - Highest interface IP address

Router and Area IDs: IS-IS

1-13 bytes	6 bytes	1 byte
Area ID	System ID	SEL

Examples:

01.0000.23a5.7c32.00

49. 0001.0000.23a5.7c32.00

47.0005.80.0000a7.0000.ffdd.0001.0000.23a5.7c32.00

- Area ID and sysID (router ID) specified in Network Entity Title (NET)
- NSAP address format
- In JUNOS™ Internet software, specified on loopback interface

Neighbor Discovery and Maintenance: OSPF

- Hello packets
  - Establish two-way communication
  - Advertise optional capabilities
  - DR/BDR election/discovery
  - Serve as keepalives
  - 10s default hello interval, dead interval 4X
- Most hello fields must match for adjacency
  - Area ID, authentication, network mask, hello interval, router dead interval, options
  - Changing values causes adjacency disruption

Neighbor Discovery and Maintenance: IS-IS

- Hello packets
  - Establish two-way communication
  - L1, L2, L1/L2 neighbor discovery
  - DR election/discovery
  - Serve as keepalives
  - 3s JUNOS default hello interval, dead interval 3X
- Hellos padded to full MTU size (dubious)
- Fewer matches necessary for adjacency
  - Hello and dead intervals can vary
  - Not even IP subnets must match!

Database Synchronization: OSPF

- Database synchronization driven by state machine
- Master/slave election
- Database synchronization
  - Database description packets
  - Link-state request packets
  - Link-state update packets
  - Link-state acknowledgement packets

Database Synchronization: IS-IS

- Simple synchronization based on flooding of sequence number PDUs
- CSNPs
  - Describe all LSPs in the database
  - Analogous to OSPF DD messages
  - Sent by DR every 10 seconds on broadcast networks
  - Sent every hour on point-to-point networks
- PSNPs
  - Request missing or newer LSPs
  - Analogous to OSPF LS Request messages

Database Refresh: OSPF

- LSA refresh every 30 minutes
- MaxAge = 1 hour
- Up-counting timer
- Design flaw: cannot change MaxAge

Database Refresh: IS-IS

- LSP refresh every 15 minutes
  - Minus random jitter timer of up to 25 percent
- LSP Lifetime = 20 minutes (default)
- Down-counting timer
- LSP lifetime configurable up to 18.2 hours
- Major reason IS-IS scales better to large areas

Designated Routers: OSPF

- Highest priority becomes DR
  - 0-255, default 128
  - Highest router ID tie-breaker
- Backup designated router
  - Speeds recovery from failed DR
- DR cannot be pre-empted
  - The DR is usually the first active router
- Adjacencies formed only with DR and BDR

Designated Routers (DIS): IS-IS

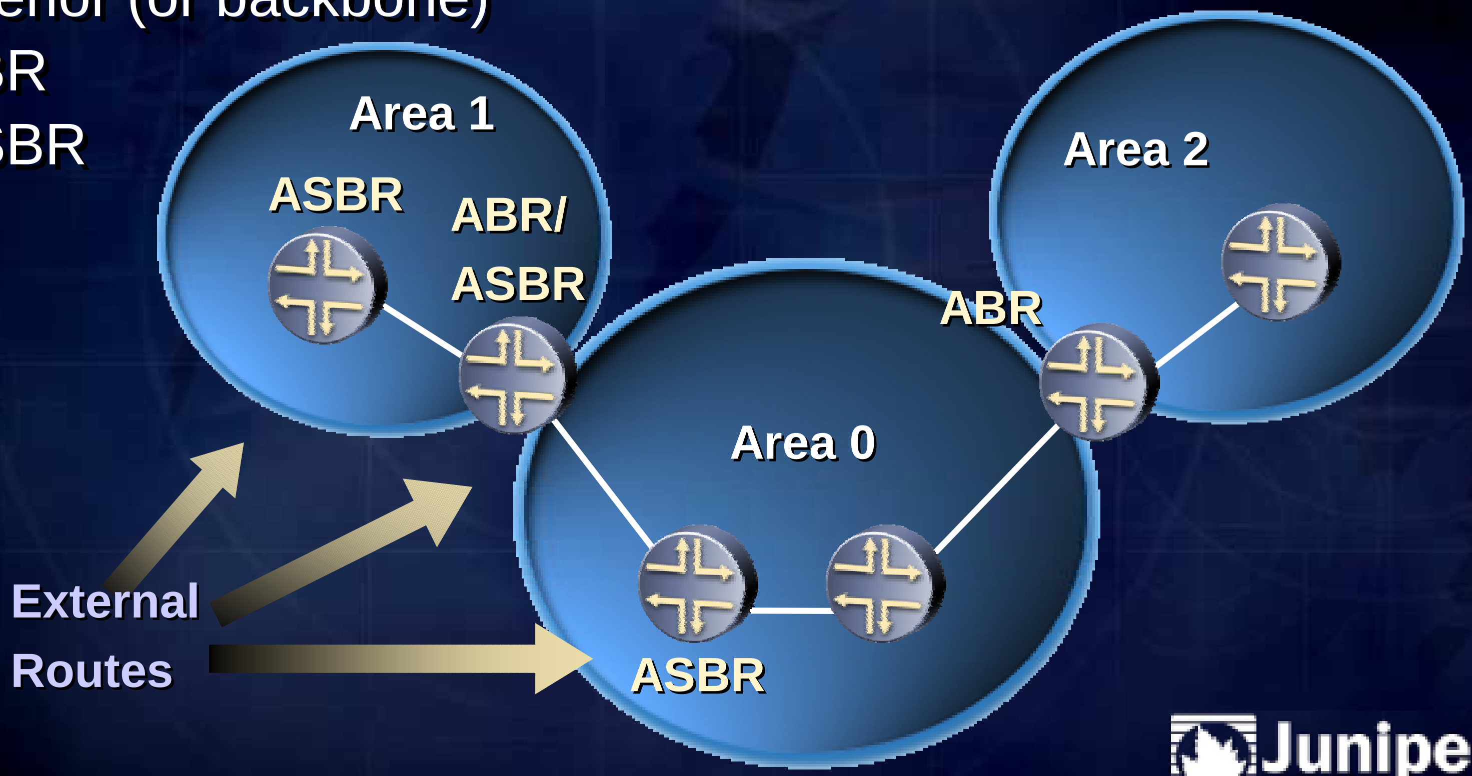
- Highest priority becomes DR
  - 0-127, default 64
  - Highest MAC address tiebreaker
- No backup designated router
- DR can be pre-empted
  - Adding a router to a LAN can cause temporary instability
- Adjacencies formed with all routers on LAN, not just DR
  - Separate L1 and L2 adjacencies on same LAN

IP FOCUS 2001

Juniper NETWORKS

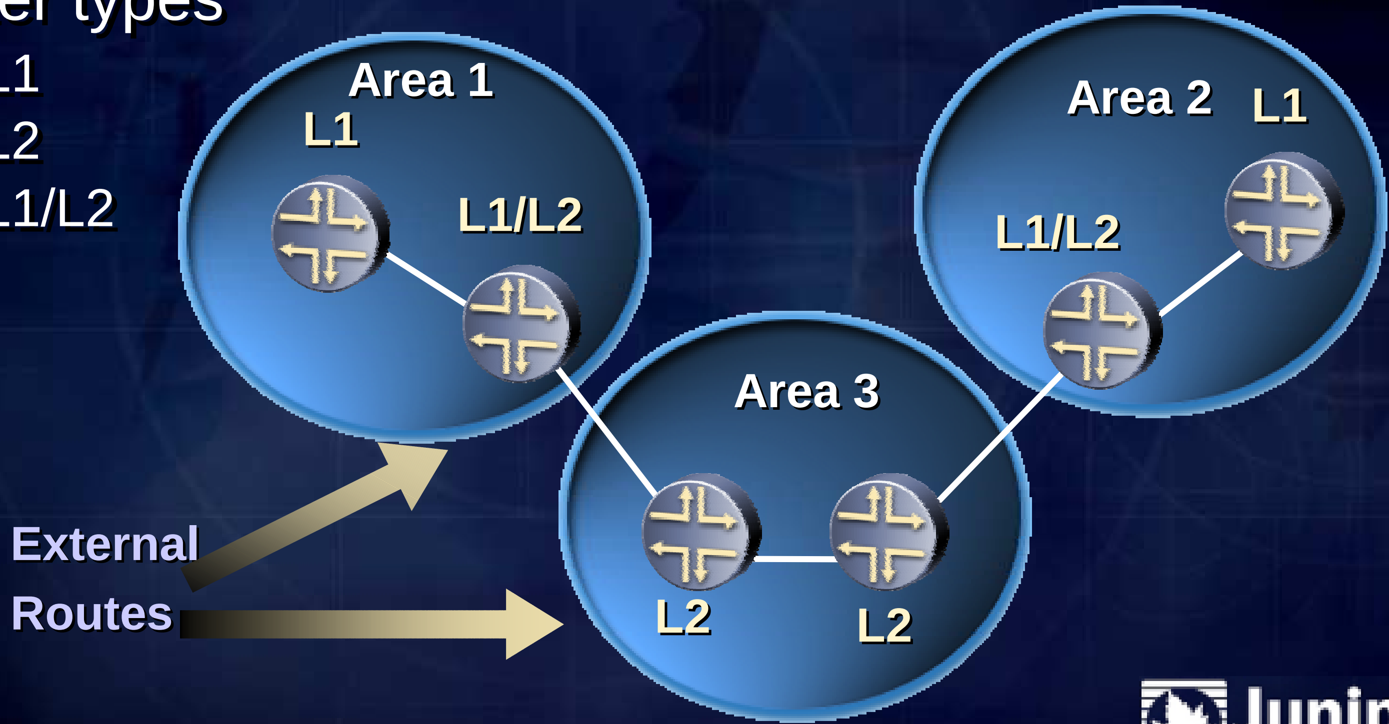
Area Structure: OSPF

- Area boundaries fall on routers
- Router types
  - Interior (or backbone)
  - ABR
  - ASBR



Area Structure: IS-IS

- Area boundaries fall between routers
- External reachability information in L2 LSPs only
- Router types
  - L1
  - L2
  - L1/L2



IS-IS Optimizations for IP Support

- Three-way handshaking
- Dynamic hostname exchange (RFC 2966)
- >256 pseudonode support
- Domain-wide prefix distribution (RFC 2966)
- Wide metrics

Agenda

- Overview of link-state protocols
- A parallel history of IS-IS and OSPF
- Comparative analysis of IS-IS and OSPF
- Design considerations
- Less-tangible considerations
- Conclusions

Metrics: OSPF

- Dimensionless metric
- Large metric field
  - Type 1 LSA = 16 bits
  - Type 3, 4, 5, and 7 LSA = 24 bits
- Cost
  - Cost = Reference BW / Interface BW
  - Default Reference BW = 100 Mbps
  - If (Ref BW / Interface BW) > 1, Cost = 1
  - Cost can also be set arbitrarily
- External Metrics
  - Type 1 (E1) = Assigned cost + cost to ASBR
  - Type 2 (E2) = Assigned cost only

Metrics: IS-IS

- Dimensionless metric
- ISO 10589 defines 4 metric fields
  - Only default used in practice
- Small 6-bit metric field
  - Default = 10 for all interfaces
  - Maximum interface value = 64
  - Maximum route metric = 1,023
  - Possible limited metric granularity in large networks
  - Originally intended to simplify SPF calculation (irrelevant with modern CPUs)
- Wide Metrics
  - Extends metric field to 32 bits
- Metrics tagged as internal or external (I/E bit)

LSA Scalability: OSPF

- Famous “rules of thumb” carry little real meaning
- 64 KB maximum LSA size
- Only router (type 1) LSAs likely to grow large
  - 24 bytes of fixed fields
  - 12 bytes to represent each link
  - 5,331 links, maximum (but isn’t this enough?)
- Types 3, 4, 5, 7 LSAs
  - One destination prefix per LSA
  - Be careful what you redistribute!

LSP Scalability: IS-IS

- Single LSP per router, per level
- Fragmentation supported, but...
  - Maximum fragment size = 1,470 bytes
  - Maximum number of fragments = 256
  - …Isn’t this enough?
- Be careful what you redistribute!

Stub Areas

- Trade routing precision for improved scalability
- OSPF
  - Stub areas eliminate type 5 LSA load
  - Totally stubby areas extend the concept
  - All area routers must understand stubbiness
- IS-IS
  - L1 routers are “totally stubby” by default
  - Attached (ATT) set by L1/L2 router

IS-IS Inter-Area Route Leaking

- Why leak routes?
  - Improved routing precision
  - More accurate BGP next-hop resolution
  - Using IS-IS metric as BGP MED
- L1-->L2 route leaking happens by default
  - Internal routes only
  - External routes require policy
- L2-->L1 route leaking requires policy
  - Internal or external
  - Up/down bit prevents looping

Not-So-Stubby Areas

- OSPF feature
  - "Trick" to allow advertisement of external routes through stub areas (type 5 LSAs illegal)
  - All routers in area must understand type 7 LSAs
- Similar function with IS-IS
  - Using simple L1-->L2 policy

NBMA Networks

- OSPF
  - Point-to-point
  - Point-to-multipoint mode
  - NBMA mode (but why?)
  - P-T-MP and NBMA require manual specification of neighbor addresses
- IS-IS
  - No multipoint support
  - Must configure interfaces as logical P-T-Ps

Virtual Links

- Useful for
  - Patching partitioned backbone areas
  - Area migrations
- Should be a temporary solution!
- Full OSPF support
- No IS-IS support
  - Specified in ISO 10589, but not implemented by major router vendors

Overload Bit

- IS-IS feature
  - Enables router to signal memory overload
  - No transit traffic sent to overloaded router
  - Set separately for Level 1 and Level 2
  - Can be manually set, useful for graceful router turn-up
- No comparable OSPF feature

Mesh Groups

- IS-IS feature (RFC 2973)
  - Can sharply curtail LSP flooding in full-mesh topologies
  - Each router in mesh group receives only one copy of each LSP (one-hop flooding)
  - Risk of lost LSPs—ensure design is robust enough!
  - Interfaces can be manually configured to block LSPs (increased scalability, but increased risk)
- OSPF has no comparable feature

Security

- Both protocols support authentication
  - Plain-text passwords (sniffable!)
  - MD5 cryptographic hash
- Authentication especially important with OSPF
  - Runs over IP, so subject to spoofing and other attacks
- Non-IP nature makes IS-IS inherently more secure
  - But authentication still a good idea

Traffic Engineering Support

- Both protocols extended to disseminate traffic engineering parameters
- OSPF
  - Type 10 Opaque (area scope) LSAs
- IS-IS
  - Extended IS reachability (type 22) TLV
    - Traffic engineering parameters in sub-TLVs
  - Extended IP reachability (type 135) TLV
    - Wide metrics and up/down bit
  - Limited to area by L1 or L2 PDU type

Agenda

- Overview of link-state protocols
- A parallel history of IS-IS and OSPF
- Comparative analysis of IS-IS and OSPF
- Design considerations
- Less-tangible considerations
- Conclusions

Extensibility

- OSPF
  - New extensions require new LSAs (usually)
  - All routers must understand new LSAs (usually)
  - IPv6 support will require new OSPF version
- IS-IS
  - New extensions require new TLVs (usually)
  - Fewer compatibility issues than OSPF
  - Small community of interest (big ISPs) with big vendor clout means faster rollout of extensions
  - Extendable for IPv6

Optimality

- Optimality
  - OSPF was optimized for things that don't matter any more (link bandwidth, CPU alignment)
  - IS-IS was optimized for things that don't matter any more (large LANs, SPF cost)
  - Optimizations turn out to add complexity, but not much value
  - A lot has changed in 10 years

Guru Availability

- OSPF
  - Broad experience base
  - Many books, RFCs, training classes available
- IS-IS
  - Significantly smaller experience base
  - Scarcity of documentation makes it mysterious
  - Simpler than OSPF, easy to learn (with a few shifts in thinking)

Writing OSPF Code

- OSPF spec is an excellent implementation guide
  - If followed to the letter, a working, if naïve, implementation will likely result
  - Spec is complex, but has almost no “why” information; hence, other (potentially more scalable) implementation approaches are at the implementer’s own risk
  - Barrier to entry in high-end router market (you need to know the protocol intuitively)

Writing IS-IS Code

- IS-IS spec uses arcane ISOspeak and has very few implementation hints
  - Spec is inherently simple (once you get the lingo), with fewer implementation issues
  - Boilerplate at front and back of spec means you can lose pages without affecting content
  - Barrier to entry in high-end router market (you need to know the protocol intuitively)

Agenda

- Overview of link-state protocols
- A parallel history of IS-IS and OSPF
- Comparative analysis of IS-IS and OSPF
- Design considerations
- Less-tangible considerations
- Conclusions

Conclusion

- Both protocols are mature and stable (with the right vendor)
- Both protocols continue to be extended
- Enterprise networks
  - IGP requirements can be complex
  - OSPF is a nobrainer
- Service provider networks
  - IGP requirements usually simpler
  - Scalability and stability are paramount
  - Consider your requirements carefully; pick the protocol that fits

Thank You

http://www.juniper.net