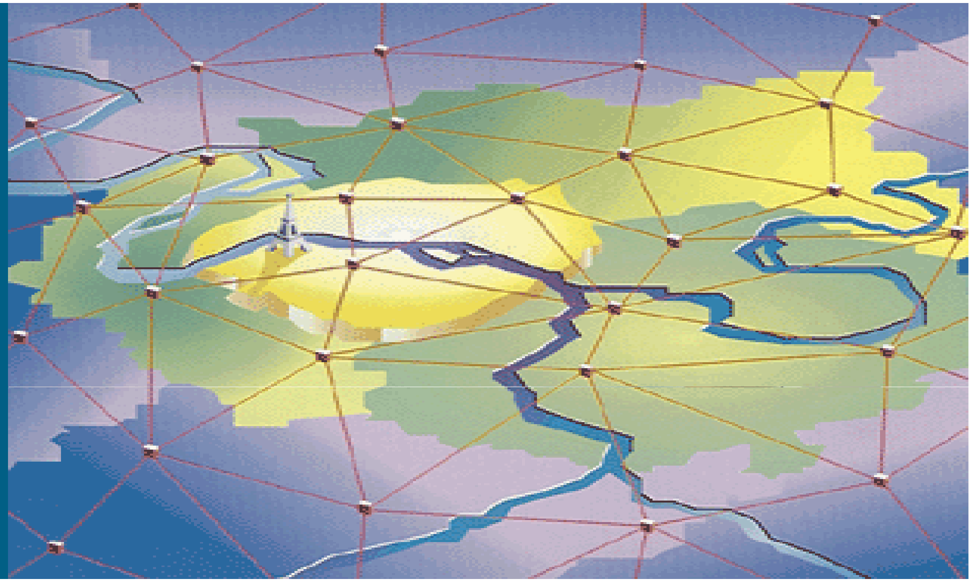
Which Routing Protocol?



Comparison between OSPF & ISIS

Faraz Shamim

Khalid Raza

Issues and Comparison

- OSPF top down view
  - –OSPF is for the most part more "optimized" (and therefore significantly more complex)
  - –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
    - •Accounts for a significant portion of OSPF's implementation complexity
    - •Partially a side effect of granular database (requires many DBD packets)

- ISIS top down view
  - –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
    - – Coarse database granularity makes this easy (just a few CSNPs)
    - –Transient routing issues can be reduced (albeit non deterministically) by judicious use of the "overload" bit

Issues and Comparison

- 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 DoS attacks (use of authentication is strongly advised)

- Encapsulation
  - –IS-IS runs directly over L2 (next to IP)
    - •Sort of makes sense (ISIS was originally designed for CLNS)
    - •Partition repair requires tunneling (rarely implemented)
    - •More difficult to spoof or attack

Terminology

Terminology

OSPF:

- Host
- Router
- Link
- Packet
- Designated router (DR)
- Backup DR (BDR)
- Link-state advertisement (LSA)
- Hello packet
- Database Description (DBD)

ISIS:

- End System (ES)
- Intermediate System (IS)
- Circuit
- Protocol Data Unit (PDU)
- Designated IS (DIS)
- N/A (no BDIS is used)
- Link-state PDU (LSP)
- IIH PDU
- Complete Sequence Number PDU (CSNP)

Terminology (cont.)

OSPF:	ISIS:
LS update	LSP (ISIS runs over layer-2)
LS acknowledgement	Partial Sequence Number PDU (PSNP)
Area	Subdomain (area)
Non-backbone area	Level-1 area
Backbone area	Level-2 area
Area Border Router (ABR)	L1L2 router
Virtual link	Virtual link (not used though)
AS Boundary Router (ASBR)	any IS
Router ID	System ID
Link-state ID	N/A
Advertising router ID	N/A

Packets

Packets

- OSPF basic header is fixed 20 bytes

Version	Type	Packet Length
Router ID		
Area ID		
Checksum	Autype	
Authentication		
Authentication		

- Common header is only 8 bytes

Intradomain Routing Protocol Discriminator				1
Length Indicator				1
Version/Protocol ID Extension				1
ID Length				1
R	R	R	PDU Type	1
Version				1
Reserved				1
Maximum Area Address				1
Additional Header Fields				
TLV Fields				

Packets

- Packet Encoding
  - –OSPF is "efficiently" encoded
    - •Positional fields
    - •Holy 32-bit alignment provides tidy packet pictures, but not much else
    - •Only LSAs are extensible (not Hellos, etc.)
    - •Unrecognized LSA types not flooded (though opaque LSAs can suffice, if implemented universally)

- Packet Encoding
  - –IS-IS is mostly Type-Length-Value encoded
    - •No particular alignment
    - •Extensible from the start (unknown types ignored but still flooded)
    - •All packet types are extensible
    - •Nested TLVs provide structure for more granular extension (though base spec does not use them; OSPF is starting to do so)

Packets

OSPF

- 5 type of basic packets

1. Hello
2. DBD
3. LS Request
4. LS Request
5. Link State Ack

ISIS

- 3 types of basic packets granularity within

1. Hello (3 types L1 LAN, L2 LAN, Point-to-point)
2. Link state packet (L1,L2)
3. Sequence number packet (CSNP, PSNP)

Hello

OSPF:

- Fixed format
- Sent every 10 sec by default.
- Intelligent sending on NBMA
- Suppressed for demand circuits

ISIS:

- TLVs (extendable)
- Sent every 10 secs by default
- DIS sends 3 times faster

OSPF LSAs

Type	LSA
1	Router
2	Network
3	Summary Network
4	Summary ASBR
5	External
6	Group Membership
7	NSSA
8	External Attributes
9–11	Opaque

ISIS LSPs

- Up to 256 LSPs per IS
- Each LSP is constructed with TLVs:

TLV	Purpose
2	Neighbor announcement
10	Authentication
22	Extended neighbor info(TE)
128	Internal IP Routing info
129	NLPID announcement (IP)
130	External IP Routing info
132	IP Interface addresses
135	Wide scale metrics

Adjacency Establishment

OSPF:

- LSDB synchronisation is performed before a neighbor is reported in the router-LSA
- On point-to-point links adjacencies are established between every pair of neighbors that can see each other
- On LAN segments adjacencies are established with the DR and BDR
- MTU mismatch is detected

ISIS:

- Adjacency is reported once two-way connectivity has been ensured
- Point-to-point links are treated the same way as in OSPF
- On LAN segments, adjacencies are established with the DIS (no BDIS is elected)
- MTU mismatch is detected

Database Granularity

–OSPF database node is an LSAdvertisement

- •LSAs are mostly numerous and small (one external per LSA, one summary per LSA)
- •Network and Router LSAs can become large
- •LSAs grouped into LSUpdates during flooding
- •LSUpdates are built individually at each hop
- •Small changes can yield small packets (but Router, Network LSAs can be large)

–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

Designated Routers

– Both protocols elect a designated router on multiaccess networks to remove O(N^2) link problem (by creating a pseudonode) and to reduce flooding traffic (DR ensures flooding reliability)

– OSPF elects both a DR and a Backup DR, each of which becomes adjacent with all other routers

- BDR takes over if DR fails
- DRship is sticky, not deterministic
- Complex algorithm

– In IS-IS all routers are adjacent (but adjacency is far less stateful)

- If DR dies, new DR must be elected, with short connectivity loss (synchronization is fast)
- DRship is deterministic (highest priority, highest MAC address always wins)
- DRship can be made sticky by cool priority hack (DR increases its DR priority)

DR Election

OSPF:

- Every LAN interface goes through the Waiting state to listen if the DR and BDR are already elected, if so, the new router does not try to pre-empt
- DR/BDR re-election happens only when current DR/BDR goes down (stability)

ISIS:

- Interfaces also go through a delay (3 seconds), but this is just an attempt to collect as much info for DR election as possible
- New router attached to a segment may cause DR switch-over

LAN Flooding

–OSPF uses multicast send, unicast ack from DR

- Reduces flood traffic by 50% (uninteresting)
- Requires per-neighbor state (for retransmissions)
- Interesting (but complex) acknowledgement suppression
- Flood traffic grows as O(N)

–IS-IS uses multicast LSP from all routers, CSNP from DR

- Periodic CSNPs ensure databases are synced
- Flood traffic constant regardless of number of neighbors on LAN

Multiple areas

- OSPF router can sit in many areas
- If backbone is attached, it is an ABR and attracts inter-area traffic
- If no backbone is attached, the router is internal to more than one area and does not attract inter-area traffic
- This is Cisco-specific, OSPF standard says "more than one area, you're an ABR" See RFC 3509 for more details

- Each ISIS router belongs to one area
- In ISIS multi-area has been added - multiple ISIS processes
- One of the processes will be L1L2 to advertise all area addresses from all processes into L2
- Designed to use for CLNS, not for IP

Links and areas

- In OSPF link can be only in one area, and routers must agree on area ID
- Area borders cross routers in OSPF

- In ISIS, if routers do not agree on area ID, they form L2 adjacency
- Area borders cross links in ISIS
- In ISIS, link can be associated with a L1 and a L2 area simultaneously

Area types

- OSPF has ordinary, stub, totally-stub, NSSA (with and without summaries)

- ISIS originally supported areas with no inter-area routes (NSSA, no-summary), now it allows for route leaking (more like NSSA)

Inter-area routing in OSPF

- OSPF has an optimal inter-area routing support---end-to-end metric is calculated
- We can prohibit injection of inter-area routes for stub and NSSA areas by using the “no-summary” keyword on the ABRs
- Inter-area route filtering (CSCdi43518)

Inter-area routing in OSPF (cont.)

- Intra-area routes are announced in type-3 summary-LSAs (possibly aggregated) by ABRs into all attached areas
- If backbone connection is active, ABRs consider only backbone summaries and re-announce them into non-backbone areas
- Standard specifies aggregation to be done only when summaries are created based on intra-area routes
- Inter-area routes can further be aggregated by ABRs when re-announced from the backbone (CSCXXXX)

Inter-area routing in ISIS

- ISIS did not have it, all areas were totally-stub, but allowed external info to be injected at any place
- Route leaking was added to ISIS to solve the problem---good filtering capability

External routing

- Type-5 LSAs are used to announce external routes by ASBRs, one LSA per one external route
- ABRs announce location of ASBRs in type-4 LSAs
- Only one copy of LSA per domain (type-5's are flooded throughout the whole domain except for stub and NSSA areas)
- Administrative tags may be set in OSPF when an external route is injected into the OSPF domain
- External routes are differentiated with internal ones
- May be aggregated by the ASBRs, and by NSSA ABRs.

- TLV 130 is used to announce external routing information, several externals share the same LSP fragment
- Every L1L2 router re-announces it to L2 (and back to L1 if route leaking is configured)
- Remote areas have as many copies of a TLV as many L1L2 routers are leaking it from L2 into these areas
- No administrative tags
- External routes look just like internal in the routing table, only L1 and L2 are differentiated
- May be aggregated by any L1L2 router

Number of neighbors

- Both protocols can maintain hundreds of neighbors (whether it’s a good idea is a different question)
- ISIS has been deployed with more neighbors in the field (people didn’t want areas)

Scalability Issues

Scalability Issues

- Database Size
  - –OSPF topologies limited by Network and Router LSA size (max 64KB) to O(5000) links
    - •External and Interarea routes are essentially unbounded
  - –IS-IS topologies limited by LSP count (256 fragments * 1470 bytes) for all route types
  - –Ultimately a non-issue for even slightly sane topologies

Scalability Issues

- Database Churn
  - Both protocols have time-limited database entries and therefore require refreshing
  - IS-IS lifetime field is 16 bits, giving 18.7-hour lifetimes (with refresh times close to this)
  - OSPF age (counts up) has an architectural lifetime limit of 1 hour (80,000 LSAs yield a refresh every 23 milliseconds)
  - “Do-not-age” LSAs are not backward compatible
  - Don’t inject zillions of routes into your IGP

Scalability Issues

- Flooding load--the only serious issue
  - –Full-mesh topologies are worst-case for both
  - –N^2 copies of each update (each of which is O(N) in size)
  - –Link failure: information
  - –Router failure: information
  - –IS-IS “mesh group” hack provides backward-compatible way of pruning flooding topology
  - –OSPF has interface blocking

OSPF v3

OSPFv3 addressing v2 issues

–Protocol processing per-link, not per-subnet (next slide..)

–Removal of addressing semantics

–Addition of Flooding scope

–Explicit support for multiple instances per link

–Use of IPv6 link-local addresses

–Authentication method changes

–Packet format, LSA’s header format changes

–Handling of unknown LSA types

OSPFv3 addressing v2 issues

- Protocol processing per-link, not per-subnet
  - –IPv6 uses the term "link" instead of network or subnet to indicate communication
  - –Interfaces connect to links
  - –Multiple IPv6 subnets can be assigned to a single link, and two nodes can talk directly over a single link, even if they do not share a common IPv6 subnet
  - –Change affects the receiving of OSPF protocol packets, and the contents of Hello Packets and Network-LSAs

OSPFv3 and v2 Similarities

packet type	Descrption
1	Hello
2	Database Description
3	Link State Request
4	Link State Update
5	Link State Acknowledgment

- OSPFv3 has the same 5 packet type but some fields have been changed.
- Mechanisms for neighbor discovery and adjacency formation
- Interface types
  - –P2P, P2MP, Broadcast, NBMA, Virtual
- LSA flooding and aging
- Nearly identical LSA types

OSPFv3 Flooding Scope

LS age	Options	LS type

LS age	U	S2	S1	LSA Function Code

- The high-order three bits of LS type {1 bit (U) for handling unrecognized LSA and two bits (S2, S1) for flooding scope} encode generic properties of the LSA, while the remainder, (called LSA function code) indicate the LSA's specific functionality
- OSPFv2 had two flooding scope, AS wide and area wide. OSPFv3 has three flooding scope:
- AS scope, LSA is flooded throughout the AS
- Area scope, LSA is flooded only within an area
- Link-local scope, LSA is flooded only on the local link.

OSPFv3 Flooding Scope

- U (unrecognized) bit is used to indicate a router how to handle an LSA if it is unrecognized

U-bit	LSA Handling
0	Treat this LSA as if it has link-local Scope
1	Store and flood this LSA as if type understood

- S2 / S1 bit indicates the three flooding scopes

S2	S1	Flooding scope
0	0	Link-Local flooding scope
0	1	Area flodding scope
1	0	AS flooding scope
1	1	Reserved

ISIS extension

IPv6 New TLV's

- IPv6 Reachability TLV 236
  - –Defines both IPv6 Internal and External reachability information
  - –Metric is still 32 bits
  - –U: Up/Down
  - –X: External origin bit
  - –S: Sub-TLV present
  - –Prefix length: Length of prefix 8 bits
  - –Prefix: Number of octet is calculated depending on the prefix length

IPv6 New TLV’s

- IPv6 address TLV 232
  - –Modified to carry IPv6 address
  - –For hello PDU interface address must use link local IPv6 address assigned to the interface
  - –For LSP non-link local address must be used

Single SPF rules

- If IS-IS is used for both IPv4 and IPv6 in an area, both protocols must support the same topology within this area.
  - –Could set “no adjacency-check” between L2 routers, but must be used with caution
- All interfaces configured with IS-ISv6 must support IPv6
  - –Can’t be configured on MPLS/TE since IS-ISv6 extensions for TE are not yet defined
- All interfaces configured with IS-IS for both protocols must support both of them
  - –IPv6 configured tunnel won’t work, GRE should be used in this configuration
- Otherwise, consider Multi-Topology IS-IS (separate SPF)

Introduction

- Mechanism that allows IS-IS, used within a single domain, to maintain a set of independent IP topologies
- Multi-Topologies extension can be used to maintain separate topologies for:
  - –IPv4
  - –IPv6
  - –Multicast
- Topologies need not to be congruent (of course)
- Multiple topologies for same address family is allowed
  - –Think about QBR…
  - –The multicast dimension …
- IETF draft: draft-ietf-isis-wg-multi-topology

The problem

- Current IS-IS spec and implementation forces all protocols carried by IS-IS to agree on a common Shortest Path Tree
  - –Single SPF run for all protocols
- Single SPT means congruent topologies
- Single SPT means all links need to understand all address families present in the domain

IS-IS Multi-Topologies Architecture

- Each router knows on which topologies it will establish adjacencies and build SPTs
  - –Through configuration
- During adjacency establishment, peers need to agree on topologies
  - –Topologies identifiers are exchanged in IIH packets

Two methods

- Multi-Topology
  - –Single ISIS domain with set of independent IP topologies
  - –Common flooding and resource associated with both router and network
  - –Multiple SPF
  - –Large Database
- Multi-instance
  - –Multiple instance of protocol 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

Two methods

- OSPF currently is based on multi-instance
  - Adding multi topology is very easy for OSPFv3
  - Multiple address family support is already there just minor extension for multi-topology needs to be added
- ISIS
  - Multi-topology support has been there for a while
  - Multi-instance draft is there for ISIS now
- Which one is better
  - Depends who you talk to
    - Operation (Multi-instance is better)
    - Development (Multi-Topology is better)

Convergence

Convergence

- Convergence depends on several factors:
  - failure detection
  - change propagation
  - initial wait for SPF computation
  - time to run SPF

Convergence Considerations

The IGPs Will Compete over Processor Cycles Based on Their Relative Tuning

- If you configure the IPv4 and IPv6 IGPs the same way (aggressively tuned for fast convergence), naturally expect a doubling of their stand alone operation convergence time
- If the IPv6 IGP is operating under default settings, the convergence time for the optimally tuned IPv4 IGP is not significantly affected

OSPFv3 Fast Convergence

- Following Techniques/tools are available for fast convergence in OSPFv3
  - Carrier Delays **Detect**
  - Hello/dead timers (Fast Hellos) **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)
  - Cisco NSF (RFC 4811,4812,4813)
  - Graceful Restart (ONLY RFC 3623)

ISIS Fast Convergence

- Following Techniques/tools are available for fast convergence in ISIS
  - – Carrier Delays **Detect**
  - – Hello/dead timers (Fast Hellos) **Detect**
  - – Bi-Directional Forwarding Detection—(BFD) **Detect**
  - – LSP pacing **Propagate**
  - – Interface event dampening - **Propagate**
  - – Exponential throttle timers for LSA & SPF **Process**
  - – PRC-interval **Process**
  - – Incremental SPF **Process**
- Techniques/tools for Resiliency
  - –Cisco NSF
  - –Graceful Restart

Conclusion

Conclusions

- OSPF is much more widely understood
  - –Broadly deployed in enterprise market
  - –Many books of varying quality available
  - –Preserves our investment in terminology
- IS-IS is well understood within a niche
  - –Broadly deployed within the large ISP market
  - –Folks who build very large, very visible networks are comfortable with it

Conclusions

- For all but extreme cases (large full-mesh networks), protocols are pretty much equivalent in scalability and functionality
- Stability and scalability are largely artifacts of implementation, not protocol design
- Familiarity and comfort in both engineering and operations is probably the biggest factor in choosing

Conclusions

- Does the world really need two protocols?
  - –Nearly complete overlap in functionality means (ironically) that few people are motivated to switch
  - –Entrenched constituencies (large ISPs; everyone else) ensure that installed bases will continue to exist
  - –As long as there are two, people will never agree on only one

Questions?