Introduction to Link State Protocols

ISP/IXP Workshops

Agenda

- Overview of Link State Protocols
- Concepts in Link State Protocols
- The Dijkstra Algorithm
- Link State Concepts Comparison
- Addressing
- Packet Types
- Areas and Levels
- Adjacencies and Flooding

Overview of Link State Protocols

What are Link State Protocols…?

Why Is It Called a Link State Protocol?

- Traditional Distance Vector Routing Protocols (DVRP) relay information regarding their relative distance to a destination
- Link State Protocols relay specific link characteristics and state information

  Only changes or updates are sent across the network

  Each router uses that information to build a routing table on its own

Link-State vs. Distance Vector Protocols

- Link-State router tells ALL other routers about ONLY its neighbours and links
- Distance Vector router tells ONLY neighbours about ALL routes

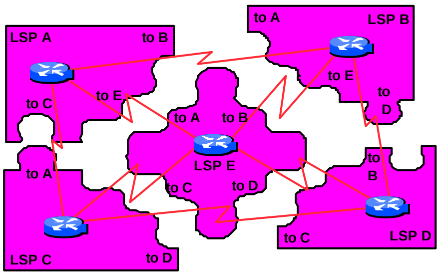
Link-State vs. Distance Vector Protocols

- In Link-State protocols Update and Decision processes are independent
- With Distance Vector protocols a router can't send out new vectors until it has processed them
- Link-State are fast loop-free convergence protocols

What Is a Link State Protocol ?

- The network can be viewed as a “jigsaw puzzle”
- Each piece of the puzzle holds one router
- Each router creates a Link State Packet (LSP) which represents its own jigsaw piece
- LSPs are flooded reliably within the network
- The LSPs are collected by each router to form a Link State Database (LSDB) or complete “picture” of the network
- Use Shortest Path First (SPF) algorithm to put the pieces together

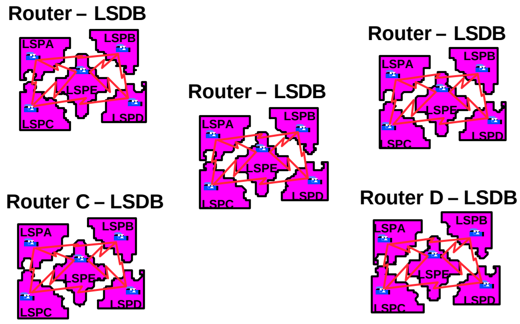
Link State Protocols



All Routers Have the Same View

- All routers exchange all LSPs
  - Via a reliable flooding mechanism
- All routers store all LSPs in a link-state database (LSDB)
  - Separate from the routing table (RIB)
  - All routers should have exactly the same LSDB, but different RIBs

All Routers Have the Same LSDB



Only Keep Track of Own State

- Routers keep track of their own state
- Send “hellos” to detect neighbours
- Keep track of IP addresses
  - interface IP prefixes
  - redistributed IP routes
  - inter-area IP prefixes etc.

When to Send an Update ?

- A router only sends a new version of its LSP when there is a change
  - if no changes, then periodic refreshes
  - a new version of an LSP will overwrite the old version of that LSP in LSDB
  - no partial updates, complete LSP sent

What To Do With LSPs?

- Each router calculates a topology map by executing Dijkstra's Shortest Path First algorithm (SPF)
  - the topology is calculated as a Shortest Path Tree (SPT), with itself as root
  - each router computes a different Shortest Path Tree (SPT)
- From the SPT the RIBs are calculated

Advantages of a Link State Protocol

- Uses metrics (costs) to calculate path
- Typically displays faster convergence than distance vector routing protocols
- Typically more scalable due to hierarchical nature

Disadvantages of a Link State Protocol

- Require more memory to store state information, i.e. databases
- Increased complexity in designing networks

Link State Protocol Concepts

So we know what they are, how do they work…?

Metrics

- Each link (interface) is assigned a cost
  - Relevant to the outbound interface
- Routers advertise the cost of their links in an LSP(s)
- Numeric value that can be assigned or calculated using link characteristics
- Metric = path cost
- More flexible than hops used in DVRPs

Topology/Link State Database

- The Link State Database (LSDB) contains information regarding all links and routers within a logical area
- A router has a separate LSDB for each area to which it belongs
- All routers belonging to the same area have identical databases
- SPF calculation is performed separately for each area

Neighbour Discovery: Hello Protocol

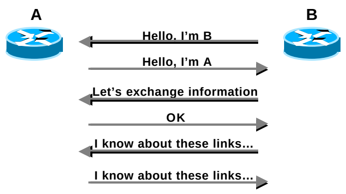
- Hello packets are used to acquire and maintain neighbour adjacencies
- Hello packets ensure that neighbours belong to the same logical subnet, same area, or contain the correct authentication information
- Dead intervals are used to detect when neighbours are no longer present

Building Adjacencies

- Routers participating in a Link State Protocol are uniquely identified throughout the network with a router ID (some form of address)
- Link state protocol routers “discover” their adjacent neighbours with some form of Hello protocol
- Once discovered neighbours form a relationship to exchange/synchronize LSDB information between them

Building the Database

- Hello packets discover neighbors
- Once neighbors are discovered LSDB information is exchanged



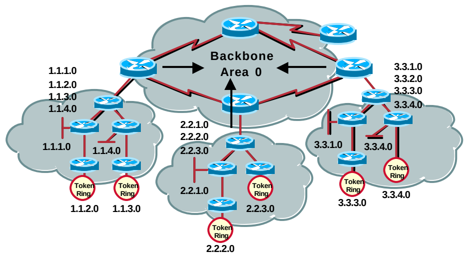
Synchronizing the Database

- Routers exchange lists describing the information they know
- Routers request specific information from their neighbor about things they do not know based on the received list
- Reception of new information is reliably acknowledged

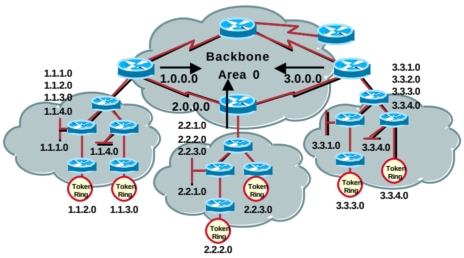
Logical Hierarchy

- Link State Protocols deploy a logical hierarchy in their design
- Usually consists of the concept of a “backbone” level and another sub-level
  - OSPF: Backbone area (area 0), regular areas
  - IS-IS: L2 Backbone areas, L1 areas
- Enables scalability by
  - Allowing summarization
  - Providing a level of abstraction

Not Summarized: Specific Links



Summarized: Summary Links

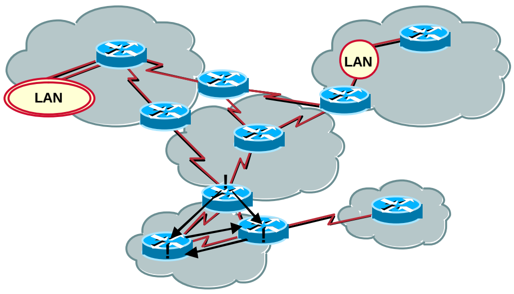


- Only summary LSA advertised out
- Link-state changes do not propagate

Flooding

- Information that changes or is learned from a neighbor is “flooded” across a logical network area
- This is done to maintain consistency of the LSDB across all routers
- Flooding of Link State Packets (LSP) are usually contained within an area

Flooding



Cisco Systems

The Dijkstra Algorithm

How does the SPF Algorithm work?

(This isn’t as complicated as it sounds!)

Shortest Path First algorithm

- We maintain three lists

UNKNOWN list:	all nodes start on this list
TENTative list:	all nodes to which a path has been found which is not yet known to be *minimum* cost.
	Also called candidate list
PATHS list:	all nodes to which we have calculated final paths with definitive cost
	Also called known list

Shortest Path First algorithm

- We execute N steps

  typically N is the number of nodes in the network. During each step we find the path(s) to one node from root.

- We initialise the computation by moving ourselves to the TENT list

Shortest Path First algorithm

At each step:

- Find the node amongst all nodes on TENT that has the lowest cost, and move it from TENT into PATHS
- Find all neighbours reachable from that node and move them into TENT
- Find all prefixes advertised by the node that was moved to PATHS and install them in the RIB

Shortest Path First algorithm

Two Way Connectivity Check (TWCC)

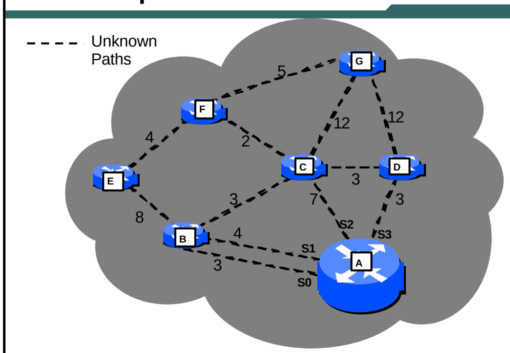
- Before moving a node into TENT, we want to be sure the parent has the same visibility as its child
- The node we want to move to TENT has to report the adjacency back to its parent

Finding the next-hop

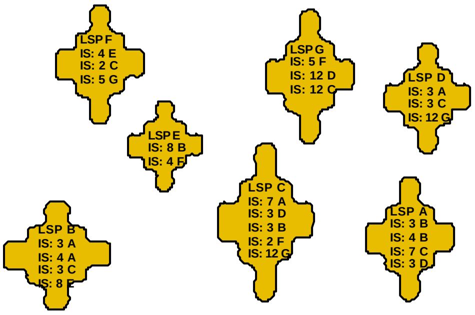
Special actions

- If a node is directly connected to us, search the first-hop info in the adjacency database
- If a node is not directly connected to us, copy the first-hop info from the parent(s)
- For each node on TENT, maintain the cost to get there from the root, and the first-hop info

An Example Network



The Link-State Database



Adjacency Database (Router-A)

Neighbour	Interface	Cost
B	serial0	3
B	serial1	4
C	serial2	7
D	serial3	3

Shortest Path First Example

- **Initial situation**

  TENT: A

  PATHS: empty

  Unknown: B C D E F G

Shortest Path First Example

- **First iteration**
  - ➢**Move A (ourself) to PATHS with cost 0**
  - ➢**Move neighbours of A to TENT: B, C, D**

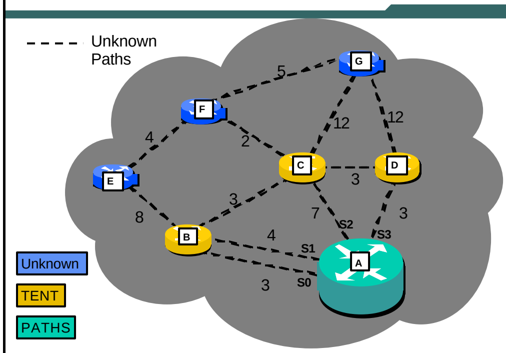
  Note that TWCC succeeds since B, C, D reports an adjacency to A

  - ➢**Find first-hop info in adjacency database**

  TENT: **D cost 3 via S3, [lowest cost]**
  C cost 7 via S2,
  B cost 3 via S0 [lowest cost]

  PATHS: **A**
  Unknown: **E F G**

PATHS and TENT: 1st Iteration



Shortest Path First Example

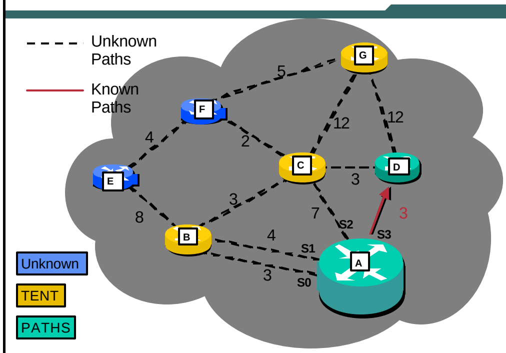
- **Second iteration**
  - ➢**Move D (lowest cost in TENT) to PATHS with cost 3 and insert in RIB its prefixes**
  - ➢**Move neighbours of D to TENT: G, found better path to C, ignore A**

  TENT: **C cost 6 via S3,**
  B cost 3 via S0, [lowest cost]
  G cost 15 via S3

  PATHS: **A,**
  D cost 3 via S3

  Unknown: E F

PATHS and TENT: 2nd Iteration



Shortest Path First Example

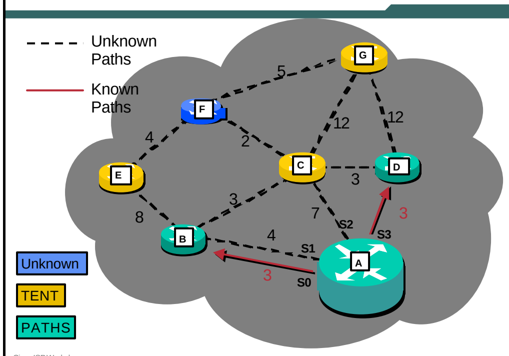
- **Third iteration**
  - ➢**Move B to PATHS with cost 3 and insert in RIB its prefixes**
  - ➢**Move neighbours of B to TENT: E, found equal cost path to C, ignore A**

  TENT: **C cost 6 via S3 & S0 [lowest cost]**
  G cost 15 via S3,
  E cost 11 via S0

  PATHS: **A,**
  D cost 3 via S3,
  B cost 3 via S0

  Unknown: F

PATHS and TENT: 3rd Iteration



Shortest Path First Example

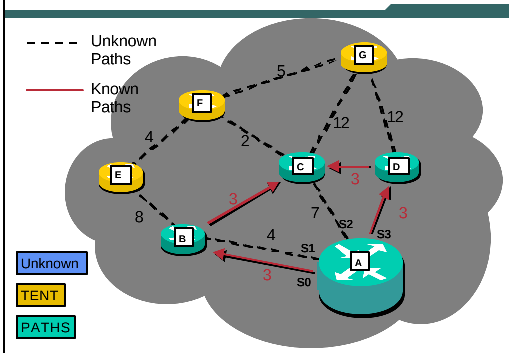
- Fourth iteration
  - ➢Move C to PATHS with cost 6 and insert in RIB its prefixes
  - ➢Move neighbours of C to TENT: F, ignore G, ignore B, ignore D, ignore A

TENT: G cost 15 via S3,
E cost 11 via S0,
F cost 8 via S3 & S0 [lowest cost]

PATHS: A,
D cost 3 via S3,
B cost 3 via S0,
C cost 6 via S3 & S0

Unknown:

PATHS and TENT: 4th Iteration



Shortest Path First Example

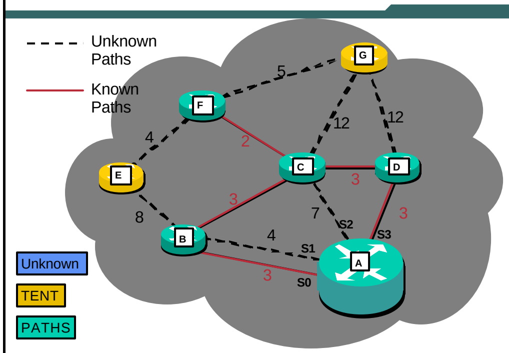
- Fifth iteration
  - ➢Move F to PATHS with cost 8 and insert in RIB its prefixes
  - ➢Move neighbors of F to TENT: Better path to G, ignore C, ignore E

TENT: G cost 13 via S3 & S0,
E cost 11 via S0 [lowest cost]

PATHS: A,
D cost 3 via S3,
B cost 3 via S0,
C cost 6 via S3 & S0,
F cost 8 via S3 & S0

Unknown:

PATHS and TENT: 5th Iteration



Shortest Path First Example

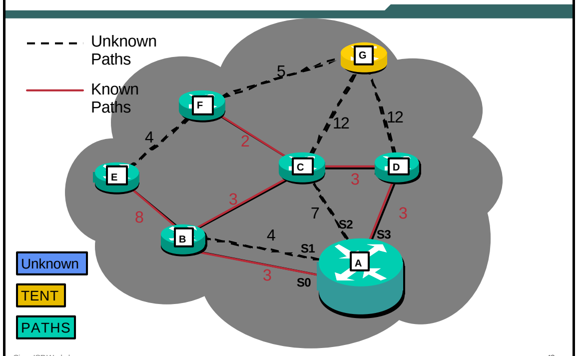
- Sixth iteration
  - ➢Move E to PATHS with cost 11 and insert in RIB its prefixes
  - ➢Move neighbors of E to TENT: Ignore B, ignore F

TENT: G cost 13 via S3 & S0 [last entry]

PATHS: A,
D cost 3 via S3,
B cost 3 via S0,
C cost 6 via S3 & S0,
F cost 8 via S3 & S0,
E cost 11 via S0

Unknown:

PATHS and TENT



Shortest Path First Example

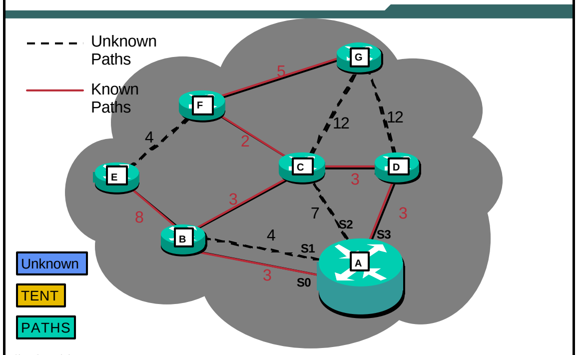
- Seventh iteration
  - Move G to PATHS with cost 13 and insert in RIB its prefixes
  - Move neighbours of G to TENT: Ignore F, ignore C, ignore D

TENT: EMPTY !!!

PATHS: A,
D cost 3 via S3,
B cost 3 via S0,
C cost 6 via S3 & S0,
F cost 8 via S3 & S0,
E cost 11 via S0
G cost 13 via S3 & S0

Unknown:

PATHS and TENT



Shortest Path First Example

- TENT is empty ® iteration stops

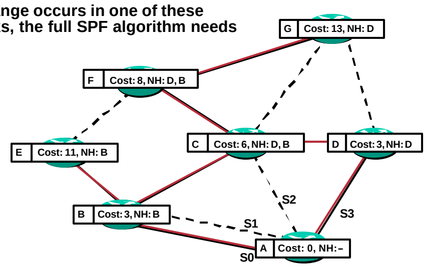
Result: A,
D cost 3 via S3,
B cost 3 via S0,
C cost 6 via S3 & S0,
F cost 8 via S3 & S0,
E cost 11 via S0,
G cost 13 via S3 & S0

Shortest Path First Example

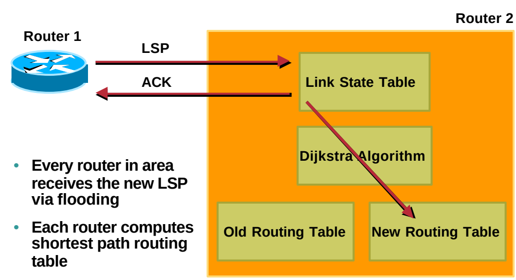
- The Shortest Path Tree

  Not all links are used (dotted links)

  Still if a change occurs in one of these unused links, the full SPF algorithm needs to be re-run



When a Link Changes State



- Every router in area receives the new LSP via flooding
- Each router computes shortest path routing table

Comparing ISIS and OSPF

Both Link State Protocols use the Dijkstra SPF Algorithm

So what's the difference then??

What Is IS-IS ?

- Intermediate System to Intermediate System
- An "IS" is ISO terminology for a router
- IS-IS was originally designed for use as a dynamic routing protocol for ISO CLNP, defined in the ISO 10589 standard
- Later adapted to carry IP prefixes in addition to CLNP (known as Integrated or Dual IS-IS) as described in RFC 1195
- Predominantly used in ISP environment

IS-IS Timeline

- 1978ish "New" Arpanet Algorithm
  - Eric Rosen et al
- 1986 to 90 Decnet Phase V
  - Radia Perlman, Mike Shand
- 1987 ISO 10589 (IS-IS)
  - Dave Oran
- 1990 RFC 1195 (Integrated IS-IS)
  - Ross Callon, Chris Gunner
- 1990 to present All sorts of enhancements
  - Everyone and his dog !

What Is OSPF ?

- Open Shortest Path First
- Link State Protocol using the Shortest Path First algorithm (Dijkstra) to calculate loop-free routes
- Used purely within the TCP/IP environment
- Designed to respond quickly to topology changes but using minimal protocol traffic
- Used in both Enterprise and ISP Environment

OSPF Timeline

- Development began in 1987 by IETF
- OSPFv1 published in 1989 with RFC 1131
- OSPFv2 published in 1991 with RFC 1247
- Further enhancements to OSPFv2 in 1994 with RFC 1583
- Latest revision for OSPFv2 was in 1998 with RFC 2328
- All above OSPF RFCs authored by John Moy

IS-IS & OSPF: Similarities

- Both are Interior Gateway Protocols (IGP)
  - They distribute routing information between routers belonging to a single Autonomous System (AS)
- With support for:
  - Classless Inter-Domain Routing (CIDR)
  - Variable Subnet Length Masking (VLSM)
  - Authentication
  - Multi-path
  - IP unnumbered links

IS-IS and OSPF Terminology

OSPF	ISIS
Host	End System (ES)
Router	Intermediate System (IS)
Link	Circuit
Packet	Protocol Data Unit (PDU)
Designated router (DR)	Designated IS (DIS)
Backup DR (BDR)	N/A (no BDIS is used)
Link-State Advertisement (LSA)	Link-State PDU (LSP)
Hello packet	IIH PDU
Database Description (DBD)	Complete sequence number PDU (CSNP)

IS-IS and OSPF Terminology (Cont.)

OSPF	ISIS
Area	Sub domain (area)
Non-backbone area	Level-1 area
Backbone area	Level-2 Sub domain (backbone)
Area Border Router (ABR)	L1L2 router
Autonomous System Boundary Router (ASBR)	Any IS

Addressing

How we do addressing within ISIS and OSPF

Transport

- OSPF uses IP Protocol 89 as transport

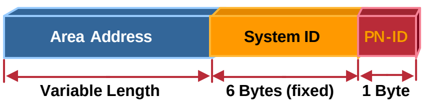
Data Link Header	IP Header	OSPF Header	OSPF Data

- IS-IS is directly encapsulated in Layer 2

Data Link Header	IS-IS Header	IS-IS Data

IS-IS: Basic CLNS NSAP Format

- NSAP: Network Service Access Point
- An NSAP mainly consists of 3 parts:



- Total length is between 8 and 20 bytes
- Example: 49.01.0000.0000.0007.00

Area. System ID. Pseudonode ID

OSPF: Router ID

- Makes use of IPv4 addressing
- Routers are identified by a unique ID
- Router ID can be assigned automatically or manually configured

  router-id <ip address>
- RID: Highest IP address from any loopbacks
- RID: If no loopback exists, highest IP address configured to a physical interface
- Router ID used to build the SPT

Packet Types

OSPF Protocol Packets

- Share a common protocol header
- Routing protocol packets are sent with type of service (TOS) of 0
- Five types of OSPF routing protocol packets
  - Hello – packet type 1
  - Database description – packet type 2
  - Link-state request – packet type 3
  - Link-state update – packet type 4
  - Link-state acknowledgement – packet type 5

Types of LSAs

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

IS-IS PDUs

- Hello PDUs (IIH, ISH, ESH)
- Link State Packet (LSP)
  - Non-pseudonode LSP
  - Pseudonode LSPs
- Complete Sequence Number PDU
- Partial Sequence Number PDU

Node and Pseudonode LSP

- 2 kinds of Link State PDUs
  - Non-Pseudonodes represent routers
  - Pseudonodes represent LANs (created by the DIS)
- Level 1 routers create Level 1 LSP
- Level 2 routers create Level 2 LSP
- Level 1–2 routers create Level 1 LSP and Level 2 LSP

Areas and Levels

Subdividing the network in OSPF and ISIS

Hierarchy Levels

Both IS-IS and OSPF Have a 2 Layer Hierarchy

- IS-IS:
  - The backbone (level-2 domain)
  - The areas (level-1)
- OSPF
  - The backbone (Area 0)
  - The areas

IS-IS: Level-1 Routers

Level-1 Router

- Neighbours only in the same Level-1 area
- Level-1 LSDB with intra-area information
- Level-1 only routers look at the attached-bit to find the closest Level-1-2 router
- Level-1 only routers install a default route to the closest Level-1-2 router in the area

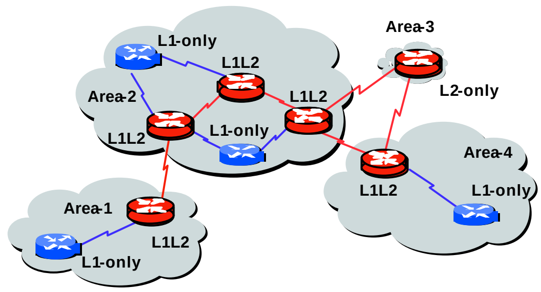
IS-IS: Level-2 Routers

Level-2 Router

- May have neighbours in other areas
- L2 has information about L2 topology
- L2 has information about which L1 areas are reachable and how to reach them via the L2 topology
- L2 routers often do also L1 routing
  - so called L1L2 routers

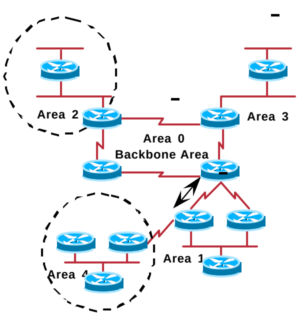
Level-1, Level-2 and Level-1-2 Routers

Backbone *MUST BE* L2 contiguous

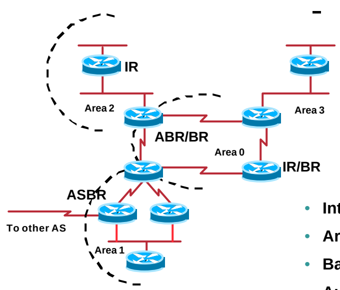


OSPF Areas

- Group of contiguous hosts and networks
- Per area topological database
  - Invisible outside the area
  - Reduction in routing traffic
- Backbone area contiguous
  - All other areas must be connected to the backbone
- Virtual Links

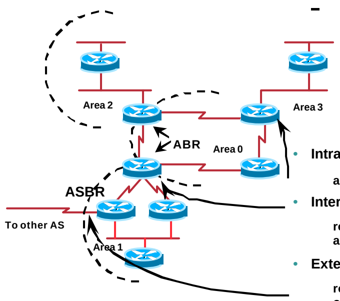


Classification of Routers



- Internal Router (IR)
- Area Border Router (ABR)
- Backbone Router (BR)
- Autonomous System Border Router (ASBR)

OSPF Route Types



- Intra-area Route

  all routes inside an area

- Inter-area Route

  routes advertised from one area to another by an Area Border Router

- External Route

  routes imported into OSPF from other protocol or static routes

Adjacencies and Flooding

How Link State Protocols find and talk to their neighbours

Discovering Neighbours: Hello Protocol

- Both IS-IS and OSPF use a “Hello” mechanism for discovering neighbours and maintaining adjacencies
  - IS-IS uses the following defaults:

    Hello Interval: 10 seconds

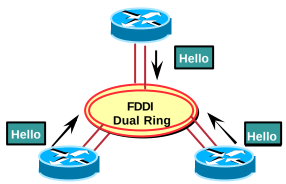
    Hold Time: 30 seconds
  - OSPF uses the following defaults:

    Hello Interval 10 seconds

    Dead Interval 40 seconds

The Hello Protocol

- Responsible for establishing and maintaining neighbour relationships
- Elects designated router on multi-access networks



OSPF: Hello Packets

- The Hello Packet contains:

  Router priority

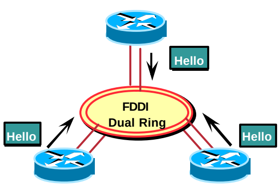
  Hello interval

  Router dead interval

  Network mask

  Options: TOS-bit

  List of neighbours
- Multicast to 224.0.0.5



IS-IS: Hello Packets

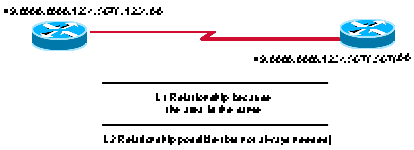
- IIHs are between routers
- Exchanged by IS to form adjacencies

  Point-to-point IIH

  Level 1 LAN IIH

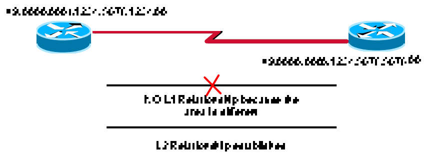
  Level 2 LAN IIH

IS-IS L1 Relationship



- When Area-ID is the same between neighbours a L1 or L2 relationship is possible

IS-IS L2 Relationship



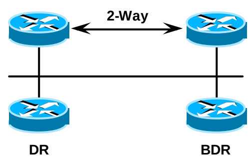
- When Area-ID are different only an L2 relationship is possible

IS-IS Adjacency

- Only two active states: INIT and UP
- Adjacency makes synchronisation of LSDB possible:
  - On p2p through the 'ack'ing of the LSP's
  - On broadcast through frequently sending CSNP packets

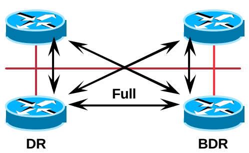
OSPF: Adjacency States

- 2-way
  - Router sees itself in other Hello packets
  - DR selected from neighbours in state 2-way or greater



OSPF: Adjacency States

- Full
  - Routers are fully adjacent
  - Databases synchronised
  - Relationship to DR and BDR



When to Become Adjacent

- Underlying network is point to point
- Underlying network type is virtual link
- The router itself is the designated router
- The router itself is the backup designated router
- The neighbouring router is the designated router
- The neighbouring router is the backup designated router

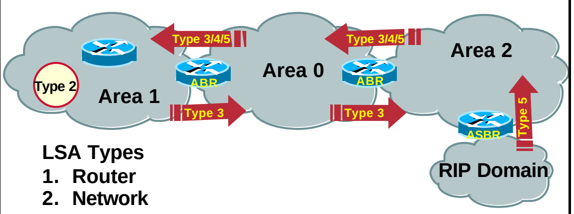
Flooding Scope

- OSPF
  - Intra-area, inter-area, external
- IS-IS
  - Area, backbone

OSPF: Types of Flooding

- Flooding on p2p links
  - Each LSA is acknowleged
- Flooding on LANs
  - DR is responsible (Network LSA)
- General background flooding
  - Refreshing LSAs

OSPF: LSA Flooding



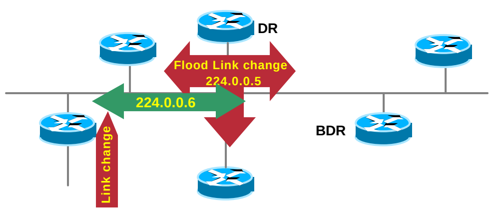
LSA Types

1. Router
2. Network
3. Summary
4. ASBR Summary
5. External

OSPF: Designated Routers

- Reduce OSPF traffic on multiaccess links
- Store and distribute neighbors LSDBs
- Backup DR for redundancy
- OSPF priority used in DR selection
  - ip ospf priority <value>
- Range 1–255, default is 1, 0 for non-candidate

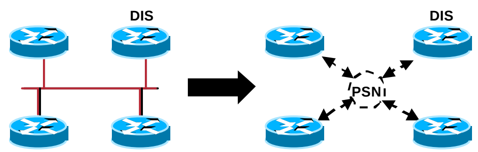
OSPF: Function of DR/BDR



IS-IS: Types of Flooding

- Flooding on p2p links with positive acks
  - each LSP is acknowleged with a PSNP
- Flooding on LANs with negative acks
  - DIS multicasts a full list of LSP descriptions in a CSNP packet every 10 seconds
  - re-transmission requests are done via PSNP
- General background flooding
  - Refreshing LSPs

Pseudonode LSP Generation



- Broadcast link represented as virtual node, referred to as Pseudonode (PSN)
- PSN role played by the Designated Router (DIS)
- DIS election is preemptive, based on interface priority with highest MAC address being tie breaker
- IS-IS has only one DIS; DIS helps routers on broadcast link to synchronize their IS-IS databases

Further Reading

I could go on and on, but here is some homework…

Further Reading: CCO

- OSPF

  http://www.cisco.com/cgibin/Support/browse/psp_view.pl?p=Internetworking:OSPF
- IS-IS

  http://www.cisco.com/cgibin/Support/browse/psp_view.pl?p=Internetworking:ISIS

Recommended Reading

Routing TCP/IP Volume I
ISBN: 1578700418

IS-IS Network Design Solutions
ISBN: 1578702208

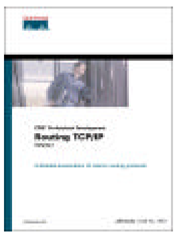
OSPF Network Design Solutions, Second Ed.
ISBN: 1587050323

Available from CiscoPress at www.ciscopress.com and all other good bookstores!

Recommended Reading

Internetworking Technologies Handbook, Fourth Edition
ISBN: 1587051192

Internetworking Technologies, Third Edition
ISBN: 1587050013



Available from CiscoPress at www.ciscopress.com and all other good bookstores!

Introduction to Link State Protocols