Lecture 10
OSPF continued

John Ioannidis
AT&T Labs – Research

ji+ir@cs.columbia.edu
Announcements

Lectures 1-10 are available.
Homework 3 is due this thursday 10/10 at 3am.
Submit only plain ASCII or PDF!
Still looking for a TA.
Have you been thinking about your project?
http://www.cs.columbia.edu/~ji/F02/proj/
Hello Protocol

• Sent every $HelloInterval$ (default: 10s).
• Neighbor discovery.
• Parameter announcement/discovery.
  – No negotiation!
• Used as keepalive.
  – Dead after $RouterDeadInterval$ (default: 4*$HelloInterval$).
• Establishes bi-directional communication.
• On broadcast and NBMA networks:
  – Elects DRs and BDRs ([Backup] Designated Routers).
Hello Packet Contents

- **Router ID** of originating router (32 bits):
  - Highest IP address on loopback interfaces.
  - If no lb, highest IP address on regular interfaces.
  - Unchanged even if interfaces go down.

The rest of the fields pertain to the originating interface.

- **Area ID** (32 bits):
  - Area ID 0 is the **backbone** area.

- Checksum (16 bits).
- Authentication type (16 bits) and information (64 bits).
  - None, cleartext (bad!), or keyed hash.
  - The hash is appended to the packet and is not considered part of the packet for checksumming purposes.
Hello Packet Contents (cont’d)

- **HelloInterval** (16 bits).
- **RouterDeadInterval** (32 bits).
- Options (5 of 8 bits).
- Router Priority (8 bits).
- DR and BDR (32 bits each).
- List of neighbors.
Hello Packet Processing

- Receiving routers (on same link) check:
  - AreaID, Authentication, Netmask, HelloInterval, RouterDeadInterval, and Options.
  - If they don’t match its own, packet is dropped.
- If RouterID is known to the receiving interface:
  - RouterDeadInterval timer is reset.
  - RouterID is added to the table of known neighbors.
- If receiving router sees its own ID in the list of neighbors in the hello packet, it knows that it has bi-directional communication with the sender.
- Adjacencies may now be formed, if appropriate.
  - Depends on network type.
Adjacencies

- If hello parameters match, neighbors may become adjacent.
- Adjacent neighbors exchange LSAs.
- Neighbors always become adjacent on:
  - Point-to-point networks.
  - Point-to-multipoint networks.
  - Virtual Links.
- How about Multiaccess networks?
  - A Designated Router is elected.
  - Multicast used on Broadcast networks.
  - Unicast used on NBMA networks.
    - Addresses preconfigured or discovered with Inverse ARP (RFC2390).
Adjacencies on Broadcast Networks

- If \( n \) routers are on a bc link, \( n(n-1)/2 \) adjacencies could be formed.
- \( n^2 \) LSAs would be originating from this network (why?).
Adjacencies, cont’d

• If routers formed pairwise adjacencies:
  – Each would originate \((n-1)+1=n\) LSAs for the link.
  – Out of the network, \(n^2\) LSAs would be emanating.

• Routers would also send received LSAs to their adjacencies.
  – Multiple \((n-1)\) copies of each LSA present on the network.
  – Even with multicast, \((n-1)\) responses would still result.

• To prevent this, a Designated Router is elected.
  – Routers form adjacencies only with DR.
  – Link acts as a (multi-interface) virtual router as far as the rest of the area is concerned.
• One router is selected as the DR.
• Actually, another is selected as the BDR.
  – If the DR fails, we want the BDR to take over within RouterDeadInterval rather than go over a new election.
    • During which no traffic would be forwarded.
• Routers form adjacencies with both DR and BDR.
• DR and BDR also form adjacencies with each other.
Adjacencies, cont’d

Routers connected by data links ⇔ nodes connected by adjacencies.
DR Election

- When router joins in:
  - Listen to hellos; if DR and BDR advertised, accept it.
    - This is the case if all Hello packets agree on who the DR and BDR are.
    - Unlike IS-IS, status quo is not disturbed!
  - If there is no elected BDR, router with highest priority becomes BDR.
  - Ties are broken by highest RouterID.
    - RouterIDs are unique (IP address of lb if).
  - If there is no DR, BDR is promoted to DR.
  - New BDR is elected.
DR Election Details

- Routers who believe they can be BDRs or DRs put their own IDs in their Hello packets.
- Once 2-way communication has been established, all routers know who the candidates are.
- They can now all pick a BDR.
  - Highest priority, then Router ID.
- And then a DR.
- If only one router claims he’s the DR, he becomes the DR.
- First two routers to come up become the DR and BDR.

- Election is identical on NMBA networks, except done with unicast Hellos.
OSPF Interface Data Structure

Relationship of router with its attached network.

- IP Address and Mask
- Area ID
- Router ID
- Network Type
- Cost
- Interface Transit Delay
- State
- Priority
- DR
- BDR
- Hello Interval
- Hello Timer
- Router Dead Interval
- Wait Timer
  - Before DR selection
- Rxmit Interval
  - Ack packets
- Neighbors
- Auth type
- Auth key
OSPF Interface State Machine

Diagram showing the state machine with states:
- point to point
- down
- waiting
- DR
- BDR
- DROther
- loopback
- Election
OSPF Neighbors

• Form adjacencies.
• Pass routing information over them.

• Adjacency establishment:
  – Neighbor discovery.
  – Bidirectional communication.
    • Neighbors listed in each other’s Hello packets.
  – [DR election].
  – Database synchronization.
    • Ensure neighbors have identical LS information.
  – Full adjacency.
• Neighbor State Machine: read about it in RFC2328.
## OSPF Neighbor Data Structure

*Relationship of router with its neighbors.*

<table>
<thead>
<tr>
<th>Feature</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interface</td>
<td></td>
</tr>
<tr>
<td>Area ID</td>
<td></td>
</tr>
<tr>
<td>Neighbor ID</td>
<td></td>
</tr>
<tr>
<td>Neighbor IP Address</td>
<td></td>
</tr>
<tr>
<td>Neighbor Priority</td>
<td></td>
</tr>
<tr>
<td>Neighbor Options</td>
<td></td>
</tr>
<tr>
<td>DR/BDR</td>
<td></td>
</tr>
<tr>
<td>Master/Slave</td>
<td></td>
</tr>
<tr>
<td>State</td>
<td></td>
</tr>
<tr>
<td>Poll Interval (NBMA only)</td>
<td></td>
</tr>
<tr>
<td>Inactivity Timer</td>
<td></td>
</tr>
<tr>
<td>DD sequence number</td>
<td></td>
</tr>
<tr>
<td>Last received DDP</td>
<td></td>
</tr>
<tr>
<td>DB Summary list</td>
<td></td>
</tr>
<tr>
<td>LS Retransmission list</td>
<td></td>
</tr>
<tr>
<td>LS Request list</td>
<td></td>
</tr>
</tbody>
</table>
Database Synchronization

- Last step before full adjacency.
- Neighbors exchange summaries of each LSA they have.
- Master/Slave relationship to determine who starts:
  - Router with highest RouterID.
- Database Description packet:
  - OSPF Header: RouterID, AreaID, Checksum, Auth.
  - Interface MTU. Options.
  - I(nitial), M(ore), M(aster)/S(lave) bits.
  - DD Sequence Number.
  - LSA Header:
    - Age, Options, Type (of LSA).
    - Link State ID (meaning varies by LSA Type).
    - Advertising Router, Sequence Number.
Full Adjacency

- After DDs have been exchanged, routers know what LSAs they are missing.

- LSA Requests.
- LSA Updates.
- LSA Acknowledgements (implicit or explicit).
Areas

- An AS (or Routing Domain) is divided into Areas.
- Group of routers.
- “Close” to each other.
- Reduce the extent of LSA flooding.
- Intra-area traffic.
- Inter-area traffic.
- External traffic.
  - Injected from a different AS.
Areas, cont’d

Area 0

Area 3.141.159.26

Area 51

Areas, cont’d
Router Types

Backbone Routers

Area 0

AS Boundary Routers

Area Border Routers

I J K

G H L

M

F

A B D

C

O N S

P Q R

W

V

BGP EIGRP
Area Partitions

- Link and router failures can cause areas to partition.
- Some partitions are healed automatically.
- Some need manual intervention.
  - Virtual Links.
- Isolated area: link failure results in no path to the rest of the network.
  - Obviously, cannot be healed at all.
  - Redundancy is important!
Partitions Include an ABR
Area 2 gets partitioned, but all its routers can reach an ABR, so traffic is not disrupted.
Isolated area

If AJ fails, A becomes isolated.
Backbone Partition?

Area 0

Area 2

Area 3
If MS fails, Areas 2 and 3 become isolated from each other, as do the two parts of the backbone.
Redundancy is good
Virtual Links

- Link to the backbone through a non-backbone area.
- Unnumbered (unaddressed).
- Connect an area to the BB through a non-BB area.
- Heal a partitioned BB through a non-BB area.
- No physical wires.
  - Exists solely as a result of configuration.
  - An example of a tunnel implemented without encapsulation.
- Configured between two ABRs.
- Transit Area: area through which VL is configured.
- Routers “connected” with VLs become adjacent.
Virtual Link Example 1

Area 3 is a Transit Area
Virtual Link Example 2

Area 3 is a Transit Area