Dynamic Routing Protocols II
OSPF
BGP

- Link State Routing and OSPF
- Interdomain Routing: BGP

Distance Vector vs. Link State Routing

- With distance vector routing, each node has information only about the next hop:
  - Node A: to reach F go to B
  - Node B: to reach F go to D
  - Node D: to reach F go to E
  - Node E: go directly to F

- Distance vector routing makes poor routing decisions if directions are not completely correct (e.g., because a node is down).

- If parts of the directions incorrect, the routing may be incorrect until the routing algorithms has re-converged.
Example:
Distance Vector vs Link State in NYC

- Directions from my former residence in New York City (116 W 73rd St) to the National History Museum.
- These are distance vector type directions:

  1. Start out going Northwest on W 73RD ST towards AMSTERDAM AVE by turning left.
  2. Turn RIGHT onto AMSTERDAM AVE.
  3. Turn RIGHT onto W 81ST.
  4. Turn RIGHT onto CENTRAL PARK W.

- The directions are perfect if all parts of the route are operational.
- However, suppose W 81 St is closed!

- Link state routing can be compared to carrying a map. With a map the directions can be recomputed at each turn.

Distance Vector vs. Link State Routing

- In link state routing, each node has a complete map of the topology.

- If a node fails, each node can calculate the new route.

- Difficulty: All nodes need to have a consistent view of the network.
Link State Routing: Properties

- Each node requires complete topology information
- Link state information must be flooded to all nodes
- Guaranteed to converge

Link State Routing: Basic principles

1. Each router establishes a relationship ("adjacency") with its neighbors
2. Each router generates link state advertisements (LSAs) which are distributed to all routers
   
   LSA = (link id, state of the link, cost, neighbors of the link)
3. Each router maintains a database of all received LSAs (topological database or link state database), which describes the network has a graph with weighted edges
4. Each router uses its link state database to run a shortest path algorithm (Dijkstra’s algorithm) to produce the shortest path to each network
Operation of a Link State Routing protocol

Received LSAs → Link State Database → Dijkstra's Algorithm → IP Routing Table

LSAs are flooded to other interfaces

Dijkstra's Shortest Path Algorithm for a Graph

**Input:** Graph \((N, E)\) with
- \(N\) the set of nodes and \(E \subseteq N \times N\) the set of edges
- \(d_{vw}\) link cost (\(d_{vw} = \text{infinity}\) if \((v, w) \notin E\), \(d_{vv} = 0\))
- \(s\) source node.

**Output:** \(D_n\) cost of the least-cost path from node \(s\) to node \(n\)

\[
\begin{align*}
M &= \{s\}; \\
\text{for each } n \notin M & \quad D_n = d_{sn}; \\
\text{while } (M \neq \text{all nodes}) & \quad \text{do} \\
& \quad \text{Find } w \notin M \text{ for which } D_w = \min(D_j; j \in M); \\
& \quad \text{Add } w \text{ to } M; \\
& \quad \text{for each } n \notin M & \quad D_n = \min_w \{ D_n, D_w + d_{wn} \}; \\
& \quad \text{Update route; } \\
\text{enddo}
\end{align*}
\]
OSPF

- OSPF = Open Shortest Path First
- The OSPF routing protocol is the most important link state routing protocol on the Internet
- The complexity of OSPF is significant

History:
- 1989: RFC 1131 OSPF Version 1
- 1991: RFC1247 OSPF Version 2

Features of OSPF

- Provides authentication of routing messages
- Enables load balancing by allowing traffic to be split evenly across routes with equal cost
- Type-of-Service routing allows to setup different routes dependent on the TOS field
- Supports subnetting
- Supports multicasting
- Allows hierarchical routing
Example Network

- Link costs are called Metric
- Metric is in the range $[0, 2^{16}]$
- Metric can be asymmetric

Router IDs are selected independent of interface addresses

Link State Advertisement (LSA)

- The LSA of router 10.10.10.1 is as follows:
  - Link State ID: 10.10.10.1 $\Delta$ Router ID
  - Advertising Router: 10.10.10.1 $\Delta$ Router ID
  - Advertising Router: 10.10.10.1 $\Delta$ Router ID
  - Number of links: 3 $\Delta$ 2 links plus router itself
  - Description of Link 1: Link ID = 10.10.10.2, Metric = 4
  - Description of Link 2: Link ID = 10.10.10.3, Metric = 3
  - Description of Link 3: Link ID = 10.10.10.1, Metric = 0

Each router sends its LSA to all routers in the network (using a method called reliable flooding)
Network and Link State Database

Each router has a database which contains the LSAs from all other routers

<table>
<thead>
<tr>
<th>LS Type</th>
<th>Link StateID</th>
<th>Adv. Router</th>
<th>Checksum</th>
<th>LS SeqNo</th>
<th>LS Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Router-LSA</td>
<td>10.1.1.1</td>
<td>10.1.1</td>
<td>0x9b47</td>
<td>0x80000006</td>
<td>0</td>
</tr>
<tr>
<td>Router-LSA</td>
<td>10.1.1.2</td>
<td>10.1.2</td>
<td>0x219e</td>
<td>0x80000007</td>
<td>1618</td>
</tr>
<tr>
<td>Router-LSA</td>
<td>10.1.1.3</td>
<td>10.1.3</td>
<td>0x6b53</td>
<td>0x80000003</td>
<td>1712</td>
</tr>
<tr>
<td>Router-LSA</td>
<td>10.1.1.4</td>
<td>10.1.4</td>
<td>0x39a</td>
<td>0x8000003a</td>
<td>20</td>
</tr>
<tr>
<td>Router-LSA</td>
<td>10.1.1.5</td>
<td>10.1.5</td>
<td>0x2a6</td>
<td>0x80000038</td>
<td>18</td>
</tr>
<tr>
<td>Router-LSA</td>
<td>10.1.1.6</td>
<td>10.1.6</td>
<td>0x5c3</td>
<td>0x80000005</td>
<td>1680</td>
</tr>
</tbody>
</table>

Link State Database

- The collection of all LSAs is called the link-state database
- Each router has an identical link-state database
  - Useful for debugging: Each router has a complete description of the network
- If neighboring routers discover each other for the first time, they will exchange their link-state databases
- The link-state databases are synchronized using reliable flooding
OSPF Packet Format

OSPF packets are not carried as UDP payload! OSPF has its own IP protocol number: 89

TTL: set to 1 (in most cases)

Destination IP: neighbor's IP address or 224.0.0.5 (ALLSPFRouters) or 224.0.0.6 (AllDRouters)

OSPF Packet Format

Message types:
1: Hello (tests reachability)
2: Database description
3: Link Status request
4: Link state update
5: Link state acknowledgement

Standard IP checksum taken over entire packet

ID of the Area from which the packet originated
0: no authentication
1: Cleartext password
2: MD5 checksum (added to end packet)

Prevents replay attacks

Authentication passwd = 1: 64 cleartext password
Authentication passwd = 2: 0x0000 (16 bits)
KeyID (8 bits)
Length of MD5 checksum (8 bits)
Nondecreasing sequence number (32 bits)
**OSPF LSA Format**

<table>
<thead>
<tr>
<th>LSA Header</th>
<th>LSA Data</th>
</tr>
</thead>
</table>

- **Link Type**: Link State ID, advertising router, link sequence number, checksum, length
- **Link ID**: Link Data
- **Metric**: #TOS metrics, Metric

**Discovery of Neighbors**

- Routers multicast **OSPF Hello packets** on all OSPF-enabled interfaces.
- If two routers share a link, they can become neighbors, and establish an adjacency

**Scenario:**
Router 10.1.10.2 restarts

- After becoming a neighbor, routers exchange their link state databases
Neighbor discovery and database synchronization

Scenario: Router 10.1.10.2 restarts

Discovery of adjacency

After neighbors are discovered the nodes exchange their databases

Sends database description. (description only contains LSA headers)

Acknowledges receipt of description

Regular LSA exchanges

10.1.10.1 sends requested LSAs

10.1.10.2 has more recent value for 10.0.1.6 and sends it to 10.1.10.1 (with higher sequence number)

10.1.10.2 explicitly requests each LSA from 10.1.10.1

10.1.10.1 sends requested LSAs
Routing Data Distribution

- LSA-Updates are distributed to all other routers via **Reliable Flooding**
- **Example:** Flooding of LSA from 10.10.10.1

Dissemination of LSA-Update

- A router sends and refloods LSA-Updates, whenever the topology or link cost changes. (If a received LSA does not contain new information, the router will not flood the packet)
- Exception: Infrequently (every 30 minutes), a router will flood LSAs even if there are not new changes.
- Acknowledgements of LSA-updates:
  - explicit ACK, or
  - implicit via reception of an LSA-Update

**Question:** If a new node comes up, it could build the database from regular LSA-Updates (rather than exchange of database description). What role do the database description packets play?
Autonomous Systems

- An **autonomous system** is a region of the Internet that is administered by a single entity.
- Examples of autonomous regions are:
  - UVA’s campus network
  - MCI’s backbone network
  - Regional Internet Service Provider
- Routing is done differently within an autonomous system (**intradomain routing**) and between autonomous system (**interdomain routing**).
BGP

- BGP = Border Gateway Protocol
- Currently in version 4
- Note: In the context of BGP, a gateway is nothing else but an IP router that connects autonomous systems.
- Interdomain routing protocol for routing between autonomous systems
- Uses TCP to send routing messages
- BGP is neither a link state, nor a distance vector protocol. Routing messages in BGP contain complete routes.
- Network administrators can specify routing policies

BGP

- BGP’s goal is to find any path (not an optimal one). Since the internals of the AS are never revealed, finding an optimal path is not feasible.
- For each autonomous system (AS), BGP distinguishes:
  - local traffic = traffic with source or destination in AS
  - transit traffic = traffic that passes through the AS
  - Stub AS = has connection to only one AS, only carry local traffic
  - Multihomed AS = has connection to >1 AS, but does not carry transit traffic
  - Transit AS = has connection to >1 AS and carries transit traffic