“open”: publicly available
uses Link State algorithm
LS packet dissemination
topology map at each node
route computation using Dijkstra’s algorithm

OSPF advertisement carries one entry per neighbor router
advertisements disseminated to entire AS (via flooding)
carried in OSPF messages directly over IP (rather than TCP or UDP)
OSPF “advanced” features (not in RIP)

- **security:** all OSPF messages authenticated (to prevent malicious intrusion)
- **multiple same-cost paths** allowed (only one path in RIP)

For each link, multiple cost metrics for different **TOS** (e.g., satellite link cost set “low” for best effort; high for real time)

integrated uni- and **multicast** support:

Multicast OSPF (MOSPF) uses same topology data base as OSPF

- **hierarchical** OSPF in large domains.
Hierarchical OSPF
Hierarchical OSPF

- **two-level hierarchy**: local area, backbone. Link-state advertisements only in area; each node has detailed area topology; only know direction (shortest path) to nets in other areas.

- **area border routers**: “summarize” distances to nets in own area, advertise to other Area Border routers.

- **backbone routers**: run OSPF routing limited to backbone.

- **boundary routers**: connect to other AS’s.
Hierarchical OSPF

(a) Hierarchical routing.

(b) Full table for 1A

<table>
<thead>
<tr>
<th>Dest.</th>
<th>Line</th>
<th>Hops</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1B</td>
<td>1B</td>
<td>1</td>
</tr>
<tr>
<td>1C</td>
<td>1C</td>
<td>1</td>
</tr>
<tr>
<td>2A</td>
<td>1B</td>
<td>2</td>
</tr>
<tr>
<td>2B</td>
<td>1B</td>
<td>3</td>
</tr>
<tr>
<td>2C</td>
<td>1B</td>
<td>3</td>
</tr>
<tr>
<td>2D</td>
<td>1B</td>
<td>4</td>
</tr>
<tr>
<td>3A</td>
<td>1C</td>
<td>3</td>
</tr>
<tr>
<td>3B</td>
<td>1C</td>
<td>2</td>
</tr>
<tr>
<td>4A</td>
<td>1C</td>
<td>3</td>
</tr>
<tr>
<td>4B</td>
<td>1C</td>
<td>4</td>
</tr>
<tr>
<td>4C</td>
<td>1C</td>
<td>4</td>
</tr>
<tr>
<td>5A</td>
<td>1C</td>
<td>4</td>
</tr>
<tr>
<td>5B</td>
<td>1C</td>
<td>5</td>
</tr>
<tr>
<td>5C</td>
<td>1B</td>
<td>5</td>
</tr>
<tr>
<td>5D</td>
<td>1C</td>
<td>6</td>
</tr>
<tr>
<td>5E</td>
<td>1C</td>
<td>5</td>
</tr>
</tbody>
</table>

(c) Hierarchical table for 1A

<table>
<thead>
<tr>
<th>Dest.</th>
<th>Line</th>
<th>Hops</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1B</td>
<td>1B</td>
<td>1</td>
</tr>
<tr>
<td>1C</td>
<td>1C</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1B</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1C</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>1C</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>1C</td>
<td>4</td>
</tr>
</tbody>
</table>
• BGP (Border Gateway Protocol): *the* de facto standard
• BGP provides each AS a means to:
  – Obtain subnet reachability information from neighboring ASs.
  – Propagate reachability information to all AS-internal routers.
  – Determine “good” routes to subnets based on reachability information and policy.
• allows subnet to advertise its existence to rest of Internet: “*I am here*”
pairs of routers (BGP peers) exchange routing info over semi-permanent TCP connections: **BGP sessions**

BGP sessions need not correspond to physical links.

when AS2 advertises a prefix to AS1: AS2 *promises* it will forward datagrams towards that prefix.

AS2 can aggregate prefixes in its advertisement
using eBGP session between 3a and 1c, AS3 sends prefix reachability info to AS1.
1c can then use iBGP do distribute new prefix info to all routers in AS1
1b can then re-advertise new reachability info to AS2 over 1b-to-2a eBGP session
when router learns of new prefix, it creates entry for prefix in its forwarding table.
Path attributes & BGP routes

advertised prefix includes BGP attributes.

prefix + attributes = “route”

two important attributes:

- **AS-PATH**: contains ASs through which prefix advertisement has passed: e.g., AS 67, AS 17

- **NEXT-HOP**: indicates specific internal-AS router to next-hop AS. (may be multiple links from current AS to next-hop-AS)

when gateway router receives route advertisement, uses **import policy** to accept/decline.
router may learn about more-than-one routes to some prefix. Router must select route.

elimination rules:
- local preference value attribute: policy decision
- shortest AS-PATH
- closest NEXT-HOP router: hot potato routing
- additional criteria
BGP messages exchanged using TCP.

BGP messages:

- **OPEN**: opens TCP connection to peer and authenticates sender
- **UPDATE**: advertises new path (or withdraws old)
- **KEEPALIVE** keeps connection alive in absence of UPDATES; also ACKs OPEN request
- **NOTIFICATION**: reports errors in previous msg; also used to close connection
A, B, C are provider networks.

X, W, Y are customer (of provider networks).

X is dual-homed: attached to two networks.

X does not want to route from B via X to C.

.. so X will not advertise to B a route to C.
A advertises path AW to B
B advertises path BAW to X
Should B advertise path BAW to C?
   No way! B gets no “revenue” for routing CBAW since neither W nor C are B’s customers
   B wants to force C to route to w via A
   B wants to route *only* to/from its customers!
Why different routing schemes?

Policy:
Inter-AS: admin wants control over how its traffic routed, who routes through its net.
Intra-AS: single admin, so no policy decisions needed

Scale:
hierarchical routing saves table size, reduced update traffic

Performance:
Intra-AS: can focus on performance
Inter-AS: policy may dominate over performance