Routing within a network: Internet Helper Protocols

- *Internet Control Message Protocol* (*ICMP*): used to send error messages back to the source when a datagram fails to be processed;
  - e.g.,
    - destination is unreachable
    - TTL expired
    - IP checksum does not match
  - Or advisory messages, such as *ICMP-Redirect, ping,* and *traceroute*
- *Tunneling*: a virtual point-to-point link created by encapsulating IP-in-IP
  - Provides some special capability (e.g., *MBone* (*multicast backbone*))
  - Creates a secure networking
  - Connectivity across other-than-IP networks
  - Forced delivery
    (e.g., mobile networks (MobileIP))

ICMP: Internet Control Message Protocol

- used by hosts, routers, gateways to communication network-level information
  - error reporting: unreachable host, network, port, protocol
  - echo request/reply (used by ping)
- network-layer "above" IP:
  - ICMP msgs carried in IP datagrams
- ICMP message: type, code plus first 8 bytes of IP datagram causing the error

Type	*Code*	*description*
0	*0*	*echo reply (ping)*
3	*0*	*dest. network unreachable*
3	*1*	*dest host unreachable*
3	*2*	*dest protocol unreachable*
3	*3*	*dest port unreachable*
3	*6*	*dest network unknown*
3	*7*	*dest host unknown*
4	*0*	*source quench (congestion control - not used)*
8	*0*	*echo request (ping)*
9	*0*	*route advertisement*
10	*0*	*router discovery*
11	*0*	*TTL expired*
12	*0*	*bad IP header*

Interior Gateway Protocols (IGP):
Routing Information Protocol (RIP)

- The *Routing Information Protocol* (*RIP*) is an example of a use of the Distance-Vector protocol.
- RIP advertises routes to other networks (not other routers).
- RIP exchanges route advertisements every 30 sec and every time that a route is changed due to a received advertisement.
- The *cost* used by RIP is the number of hops, with the value of "16" representing infinity.
- Included in BSD-UNIX Distribution in 1982

Interior Gateway Protocols (IGP):
OSPF (Open Shortest Path First)

- Uses Link State algorithm
  - LS packet (*LSP*) dissemination
  - Topology map at each node
  - Route computation using *Dijkstra's* algorithm
- Each node assumed to know state of links to its neighbors
- Each node broadcasts its state to all other nodes
  - use of reliable flooding mechanism
- Each node locally computes shortest paths to all other nodes from the locally available global state
  - Using *Dijkstra's* shortest path algorithm

OSPF "advanced" features

- Some goals/features of OSPF include:
  - open standard (as opposed to proprietary protocols)
  - type-of-service (TOS) routing (for each link, multiple cost metrics for different TOS (e.g., satellite link cost set "low" for best effort; high for real time))
  - load balancing
  - allows network partition to self-contained subsets termed *areas*
  - variety of authentication schemes (all OSPF messages authenticated (crypto-based authentication))
  - integrated uni- and multi-cast support
  - host- and network-specific routes
  - uses the notion of the designated router to minimize the amount of periodic broadcast link status
  - allows virtual network topology
  - allows the exchange of information from external sites.

Hierarchical OSPF

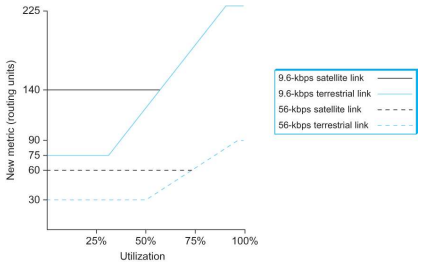
- Two-level hierarchy: local area, backbone.
  - Link-state advertisements only in area
  - Each node has detailed area topology; only knows direction (shortest path) to nets in other areas.
- Area border routers: "summarize" distances to nets in own area, advertise to other Area Border routers.
- Backbone routers: run OSPF routing limited to backbone.
- Boundary routers: connect to other AS's.

Link Metrics

- Simplest metric is *number of hops*
  - Treats all links the same way (e.g., delay, capacity, load)
- Queue-length (lacks the dynamic aspect)
- Average link delay (instability → idle links; dynamic range)

```
Delay = (Depart_Time-Arrival_Time) + Transmission_Time + Latency
Transmission_Time ∝ Bandwidth
Latency ∝ Prop. Delay
```

- Link utilization + "smoothing" (limit the amount of change in one measurement cycle) to eliminate instability + compression of dynamic range



- In practice, metrics are typically static; `∝ 1/link-bandwidth`

Scalability: Improving the Scaling Properties

- Dijkstra's shortest-path algorithm
  - Simplest version: $O(N^2)$, where N is # of nodes
  - Better versions: $O(L*log(N))$, where L is # links
  - Incremental algorithms: great for small changes
- Timers to pace operations
  - Minimum time between LSAs for the same link
  - Minimum time between path computations
- More resources on the routers
  - Routers with more CPU and memory