SE 4C03 Winter 2007

11 Routing Protocols

William M. Farmer

Department of Computing and Software
McMaster University

31 March 2007

Routing Management Goals

- Packets can get from any source to any destination.
- Packets take the shortest route.
- Routing information is automatically distributed.
- Routing tables are automatically initialized and updated.
- Many routers route with partial information about destinations by using default routes.
  - Tables are smaller and routing is more efficient.
  - Management can be performed locally.
  - Routing tables are less likely to be inconsistent with each other.

Core Router Architecture

- The pool of internet routers is divided into core routers and noncore routers.
  - Motivated by the ARPANET which was a single WAN.
  - Works well when the internet has a single backbone.
- Provides optimal routes for all possible destinations.
  - Core routers do not use default routes.
  - Every assigned class network address must be advertised to the core system.
  - A noncore router sends all nonlocal traffic to the core router at its site.
  - Core routers communicate with themselves to preserve consistency.
- Core router architecture is impractical today because:
  - The Internet is no longer built around one backbone.
  - It is not possible to have a core router at each site.
  - Core router architecture does not scale up very well.

Peer Backbone Architecture

- Peer backbone networks have two or more backbone networks with several connections between them.
  - Motivated by connection of the NSFNET backbone to the ARPANET backbone.
- The system is difficult to implement because:
  - Packets cannot simply be routed according to the network portion of their destination.
  - Peers must keep their routes consistent with each other.
  - Defaults routes from one peer to another can create routing loops.

Vector Distance Routing (1)

- Each router keeps a list of route records having the following fields:
  - The destination class network (i.e., the vector).
  - The number of hops to the destination (i.e., the distance).
  - The route (either direct or the name of a router).
- Assumption: Measuring distance as number of hops is a good measure of the time cost of a route.
- When a router boots, the list is initialized to just the routes for the class networks that are directly connected to the router.
- Each router sends a copy of its list to all other routers that are directly connected to it.

Vector Distance Routing (2)

- Suppose that R_1 and R_2 are directly connected routers.
  - If R_1 has a route record (N, D, R) for the class network N but R_2 does not, R_2 will add the record $(N, D+1, R_1)$ to its list.
  - If R_1 and R_2 have route records (N, D_1, R_1') and (N, D_2, R_2'), respectively, with $D_1 + 1 < D_2$, then R_2 will update its record to $(N, D_1 + 1, R_1)$.
  - If R_1 and R_2 have route records (N, D_1, R) and (N, D_2, R_1), respectively, with $D_1 + 1 \neq D_2$, then R_2 will update its record to $(N, D_1 + 1, R_1)$.

Vector Distance Routing (3)

- Advantage: Vector-distance algorithms are easy to implement.
- Disadvantage: Vector-distance algorithms do not scale up well.
  - Each router eventually has a record in its list for each class network.
  - Lots of routing information is transmitted.
  - Routing information is propagated slowly (which can make the system unstable).
  - Any incorrect routing information will be propagated along with correct information.

Gateway-to-Gateway Protocol (GGP)

- Now defunct vector-distance protocol used for sharing routing information among core routers.
- GGP messages were encapsulated in IP datagrams (with protocol field set to 3).
- Kinds of GGP messages:
  - Routing update message.
  - Positive acknowledgment to routing update (update acceptable).
  - Negative acknowledgment to routing update (error detected).
  - Echo request.
  - Echo reply.

Link-State Routing (1)
(Shortest Path First (SPF) Routing)

- Each router keeps a graph of the topology of the internet.
  - A node represents a router.
  - An edge represents a network.
- The routers work to keep the graphs up to date.
  - Each router periodically checks to see if its neighbors are up or down.
  - Each router periodically broadcasts a message that contains the state of each of its links.
  - Each router uses the link-state broadcasts it receives to update its internet topology graph.

Link-State Routing (2)

- Whenever the graph changes, a router uses the Dijkstra shortest path algorithm to compute the shortest route to each destination.
- Advantages over vector-distance routing.
  - Each router independently computes the shortest routes from the same information.
  - It is easy to fix mistakes because link-state information is not modified as it is propagated.
  - The size of link-state messages do not grow as the internet grows.

Problems with Core Router System

- It is impractical for the group of core routers to include more than a small portion of the internet routers.
- Noncore routers need to know core router routes to avoid "extra hops".
- Core routers need to know about "hidden" networks.

Autonomous System Architecture

- An internet is composed of several autonomous systems each composed of a collection of routers and networks.
- Each autonomous system chooses its own internal routing architecture.
  - An interior gateway protocol (IGP) is used to distribute routing information within an autonomous system.
- One or more routers in an autonomous system advertises local routing information to other autonomous systems.
  - An exterior gateway protocol (EGP) is used to distribute routing information between two autonomous systems.

Border Gateway Protocol (BGP)

- BGP is the EGP currently used in the Internet and most other TCP/IP internets.
  - Each autonomous system has at least one designated border gateway that speaks on behalf of the autonomous system.
  - Border gateways exchange reachability information (but not optimal routes).
- Attributes of BGP:
  - Transport is via TCP and thus is reliable.
  - Local routing policy can be supported.
  - Updates are usually incremental.
  - A receiver can authenticate the sender.
  - Subnet addressing is supported and so, for example, related destinations can be aggregated.

BGP Message Types

1. OPEN: initializes communication.
2. UPDATE: withdraws unreachable destinations and advertises new destinations. Each advertised destination can include:
   - The next hop to the destination.
   - The path of autonomous systems to the destination.
   - The source of the advertisement.

   Each destination is expressed as a compressed address-mask pair.
3. NOTIFICATION: reports an error.
4. KEEPALIVE: tests network connectivity.
   - KEEPALIVE messages have minimum size (they contain only a BGP header).

Routing Information Protocol (RIP) (1)

- An IGP protocol for vector-distance routing.
  - Distance is measured by the number of hops.
- Implemented by the `routed` program designed at the University of California at Berkeley.
  - Became popular because it was distributed with BSD UNIX.
- RIP participants are either active or passive.
  - Active participants are routers that advertise their routes to others.
  - Passive participants are hosts that do not advertise routes but use advertisements to update their routes.

Routing Information Protocol (RIP) (2)

- An active participant broadcasts a message every 30 seconds.
  - RIP messages are encapsulated in UDP datagrams.
  - A RIP server listens at UDP port 520.
- Route restrictions.
  - An old route is retained until a new one with a strictly lower cost is received.
  - Routes timeout after 180 seconds.
  - The maximum possible distance is 16, so RIP only works with relatively small autonomous systems.
- RIP suffers from the problems that are inherent in vector-distance routing.
- RIP does not employ router authentication.

The HELLO Protocol

- A defunct IGP protocol for vector-distance routing.
- Routes are measured by network delay instead of number of hops.

The Open SPF Protocol (OSPF) (1)

- An IGP protocol for link-state (SPF) routing.
- Provides support for:
  - Type of service routing: Routing is done on the basis of both destination address and type of service (precedence plus low delay, high throughput, or high reliability).
  - Load balancing: The same traffic can be distributed over multiple routes.
  - Area organization: Local networks and routers can be organized into independent areas.
  - Authentication: Routers must authenticate each other.
  - Subnet routes: Routes may be directed to subnets.
  - Information forwarding: Routers may forward information received from routers exterior to a site.

The Open SPF Protocol (OSPF) (2)

- Each router keeps a graph of the topology of the internet whose structure is more complex than that of other link-state routing protocols.
  - A node represents a router or an SPN.
  - One router can be directly connected to another router in the graph.
- Kinds of OSPF messages:
  - Hello (to test router reachability).
  - Graph description (to initialize a router's topology graph).
  - Link status (to update a link in a router's topology graph).
- OSPF messages are encapsulated in IP datagrams (with protocol field set to 89).

The gated Program

- Handles multiple routing protocols including both IGPs and EGPs.
- Allows a router to communicate with routers both inside and outside its autonomous system.
- IGPs supported include: RIP, HELLO, OSPF.
- EGPs supported include: BGP.