Routing

"It's how the bits get there!"

CIS 68C2

UNIX Network Administration

Routing

- *What* is Routing?
  - TCP/IP's directional decision maker for packet flow
    - Routing is ultimately just passing a packet from gateway to gateway until it reaches its destination host
- *Why* are packets routed?
  - TCP/IP is a switching network
    - Packets must be switched across physical network boundaries
    - IP can only deliver packets through known networks and gateways
- *Who* does routing?
  - IP Protocol is responsible for performing packet switching
    - IP decides where to send a packet next

Routing

- *When* are packets routed?
  - Every time a packet is received by IP!
- *Where* are packets routed?
  - Up the TCP/IP stack
  - To an attached physical network
  - To the local loopback network
  - Nowhere – just dropped on the floor
- *How* are packets routed?
  - IP routes packets with the help of a table of instructions

Routing

- What is a route?
  - Simple instructions specifying how to forward a packet
  - No route … no packet delivery!
- Routing consists of two distinct components:
  - Packet Routing
    - Forwards packets based on a routing table
    - Performed in all TCP/IP networking systems
  - Routing Protocols
    - Define the exchange of routing table information
    - Used by systems whose routing tables are updated dynamically
    - Optional

Routing

- Routing Configurations
  - Minimal – Single Route
    - Single route created by **ifconfig**
      - **ifconfig up** typically installs a single route to the local network
  - Static Routes
    - Routing table built manually by system administrator
  - Dynamic Routes
    - Routing table built dynamically by routing table exchange software
    - Protocols help define and determine “best” routes

Routing

- A Packet's Journey
  - Transmitting NIC places packet on wire
    - Packet headers contain:
      - Source and destination Ethernet addresses (MAC addresses)
      - Source and destination IP addresses
  - Receiving NIC picks up packet as it passes on the wire
    - NIC passes packet up the protocol stack to the IP layer
  - IP makes routing decision based on destination IP address
    - If address matches host's, packet is passed up to Transport Layer
    - Otherwise, it is considered for *forwarding* to the next *hop*

Routing

- A Packet's Journey - *continued*
  - Forwarding occurs when…
    - The system is enabled for packet forwarding, *and*
    - The routing table contains either
      - a specific entry for a directly connected network, or
      - a specific entry for a router on a directly connected network, or
      - a default route (a place to go when no specific route exists)
  - When IP forwards a packet…
    - Original **datagram** (IP header plus payload) is intact
    - Packet is re-framed with a MAC header for transmitting interface
      - Source address is interface's MAC address
      - Destination address is next hop's MAC address

Routing

- Building Static Routes
  - Use **route** command to add/delete routes
  - Linux **route** syntax is somewhat unique (and bizarre)
  - Commands to view the route table:
    - **netstat –rn** and **route**
  - Set default route to gateway that accesses maximum number of networks
  - Routes in Linux can be added at boot
    - /etc/sysconfig/static-routes
    - Format: *interface args*
      - Where *args* are the arguments to **route** that follow the **add** keyword

Routing

Reading the Route Table

1) For each row in the route table, apply the network mask in the Genmask column to the packet's destination IP address and if the resulting network address matches the Destination address here …

2) … then IP can forward the packet onto the network or to the gateway listed for that row…

3) … by transmitting the packet through this interface.

```
$ netstat -rn
Destination     Gateway         Genmask         Flags   MSS Window  irtt Iface
10.0.0.0        0.0.0.0         255.255.255.0   U        40 0          0 eth0
11.0.0.0        0.0.0.0         255.255.255.0   U        40 0          0 eth1
127.0.0.0       0.0.0.0         255.0.0.0       U        40 0          0 lo
0.0.0.0         11.0.0.1        0.0.0.0         UG       40 0          0 eth1
```

4) Finally, if there is no previous match, then the 0.0.0.0 entry here *will always match* since applying this row's Genmask of 0.0.0.0 to any IP address results in a 0.0.0.0 network address. This entry is called the Default Route.

Routing

- ICMP Redirects
  - Router A instructing host to use Router B
  - Causes new static **host route** to be added to route table
    - Only removed via reboot or administrator
    - Use is generally not recommended for security concerns
  - /proc/sys/net/ipv4/conf/*/accept_redirects

Routing Protocols

- Overview
  - Routers use routing protocols to dynamically exchange information used to build route tables
  - Two groups of protocols: **interior** and **exterior**
    - Interior Protocols
      - Used *within* an autonomous system
      - Determine and distribute *best* route information to systems
      - Administrator chooses the protocol
      - Protocols: RIP, RIP-2, OSPF, IS-IS, Hello
    - Exterior Protocols
      - Exchange *reachability information* between autonomous systems
      - ISP or administrator of autonomous system chooses the protocol
      - Protocols: EGP, BGP

Routing Protocols

- Two implementation approaches:
  - **Distance Vector** and **Link State**
- Distance Vector
  - A "gossipy" protocol
  - A router *advertises* its distance to neighboring networks
    - *"How many routers (hops) away is a given network?"*
    - Distance metric is measured in terms of *hops*
      - Each router typically adds one hop (router metric is 1)
  - Other routers listen for *route advertisements*
    - Replace their expensive routes with advertised cheaper routes
    - More expensive routes advertisements are ignored

Routing Protocols

- Distance Vector
  - Limitations
    - Infinite loops are a challenge
      - A receives from B, sends to C, which sends back to B
    - Slow convergence
      - Takes time before routes settle down
    - Requires arbitrary maximum hop count
      - Prevents infinite loops or delays
    - Frequent, bulky updates
      - All information exchanged very frequently
        - RIP: every 30 seconds
        - EGRP/EIGRP: every 90 seconds

Routing Protocols

- Link State
  - Exchange the *state* of adjacent routers
  - *"Router A is connected to router B and link is up"*
  - Requires connectivity map of entire network
    - At most, an entire autonomous system
    - Map is a (possibly very large) *directed graph*
  - Yields an end-to-end overview of the entire network (AS)
    - Cost to reach any end point is known
      - Better than only having cost to next hop
  - Are more complicated than distance vector
  - Requires more CPU and RAM

Interior Routing Protocols

- RIP – Routing Information Protocol
  - Simple Distance Vector protocol
  - Implemented in **routed** daemon
  - Uses hop count as its metric
  - Limitations
    - Bandwidth hog – broadcasts all information
    - Slow convergence after network changes
      - Heuristic solutions: split horizon, poison reverse, triggered updates
    - Beyond 15 hops is considered unreachable
    - No support for CIDR (no support for network masks)

Interior Routing Protocols

- RIP-2
  - RIP w/extensions to overcome RIP's limitations
  - Backwards compatible with RIP
  - Enhancements
    - Supports CIDR (distributes network masks)
    - Minor security enhancements (not widely used)
    - Next hop updates allow advertising routes not your own
    - Route tags to allow propagation of externally discovered routes
    - Uses multicast to reduce network load

Interior Routing Protocols

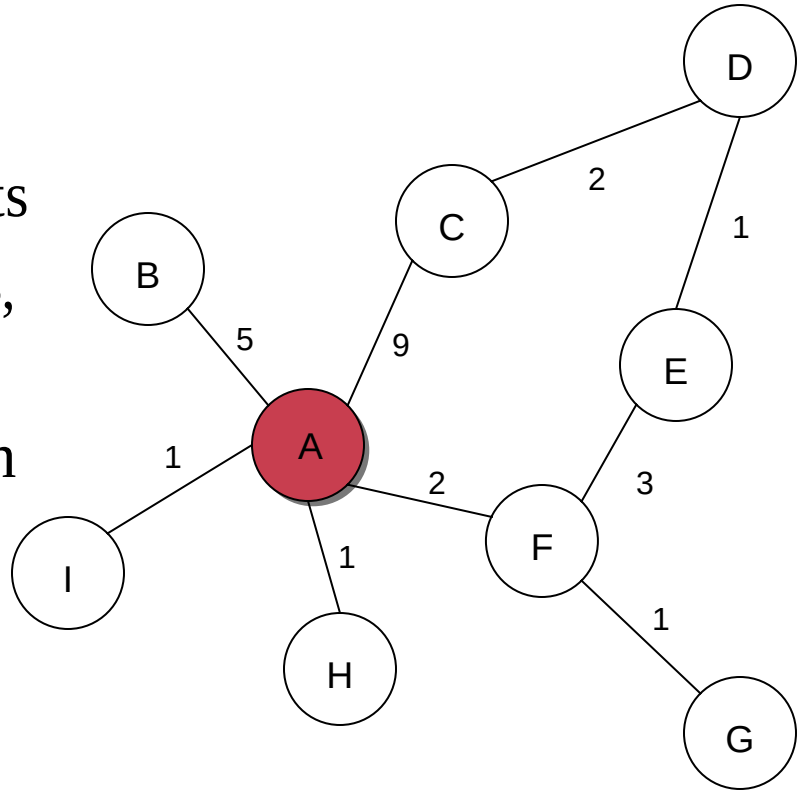
- OSPF – Open Shortest Path First
  - Link-state protocol
  - Implemented in **gated** daemon
    - Widely-used and most popular multi-protocol routing daemon
  - Pros
    - Works well in large, complex environments
    - Much less chatty than RIP
    - Handles multiple paths to a destination
    - Allows arbitrary cost metrics (hops are default w/gated & Cisco)
    - Allows sectioning networks into Areas, Backbone, and Stub Areas
    - Supports authentication (clear text and MD5)

Interior Routing Protocols

- OSPF
  - Cons
    - Very complex (see appendix B, page 449-496)
  - Determines neighbors via HELLO packets
  - Link State Announcements are flooded
    - Neighbors + cost are advertised via each interface
  - Neighbors are added to graph
  - Map is pruned of more expensive routes
  - Use of designated routers improves efficiency
    - Treats all others as neighbors
    - Reduces map size

Interior Routing Protocols

- OSPF Directed Graph
  - **A** sends HELLO packet to neighbors and neighbors respond with HELLO packets
  - **A** floods LSA to neighbors **B**, **C**, **F**, **H**, and **I**
  - Neighbors flood LSAs in turn to their neighbors
  - **A** adds paths to its map
  - **A** prunes map of more expensive routes



Interior Routing Protocols

- Others
  - IGRP / EIGRP
    - Proprietary Cisco routing protocols
  - IS-IS – Intermediate to Intermediate System
    - Rarely used ISO attempt at OSPF-like functionality
  - Hello
    - Was used in 56k NFSNET
  - MOSPF / DVMRP / PIM
    - Multicast IP routing protocols
  - Router Discovery Protocol
    - Not (yet) supported by all routers

Exterior Routing Protocols

- EGP – Exterior Gateway Protocol
  - Only announces its own networks
  - An EGP system *acquires a neighbor*
    - Via *Hello* and *I-Heard-You* messages
  - Does not choose “best” route
    - Distance vector information is not interpreted, just passed
      - Each autonomous system uses its own metric
    - Trusted core gateways evaluate and redistribute best routes
  - No longer popular

Exterior Routing Protocols

- BGP – Border Gateway Protocol
  - Leading exterior routing protocol
  - Path vector protocol
    - Provides entire end-to-end path
  - Supports **Policy-based Routing**
    - Allows non-technical policies to control routing
  - Does not require central routing authority
  - Very efficient

Routing Daemons

- The **routed** Daemon
  - Long time standard in UNIX
  - Supports RIP only
    - Some recent implementations also support RIP-2
  - Options
    - **-s** Run in server mode advertises routes
    - **-q** Run in quiet mode only listens
    - **-g** Advertises a single gateway
  - Configuration file: /etc/gateways
    - For static routes

Routing Daemons

- The **gated** Daemon
  - Generally preferred over **routed**
  - Supported many protocols
    - Interior: RIP, RIP-2, OSPF, IS-IS
    - Exterior: EGP, BGP (v3, v4), HELLO
  - Configuration file: /etc/gated.conf
  - Administration via the **gdc** command

Routing

- Additional Information
  - Linux Networking HOWTO
    - Routing Section
      - http://www.linux.org/docs/ldp/howto/Net-HOWTO/x552.html