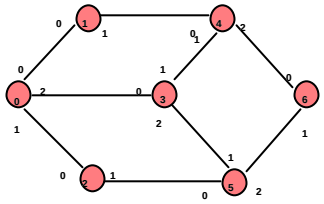
Internet Routing and Multicasting



The author of these slides is Dr. Mark Pullen. Students registered Dr. Pullen's course may make a single machine-readable copy and print a single copy of each slide for their own reference, so long as each slide contains the copyright statement. Permission for any other use, either in machine-readable or printed form, must be obtained from the author in writing.

Lecture Overview

- IP routing
- IP multicasting
- IP multicast routing
- resource reservation

Internet Routing

Routing and Forwarding
What is the difference?

Routing is the process of determining what route a packet shall follow

- requires routers to exchange reachability information using a *routing protocol*

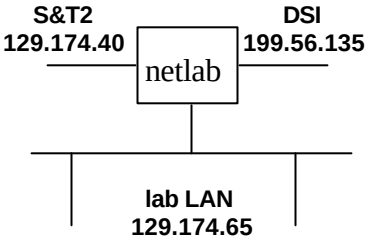
Forwarding is the process of dealing with a newly-arrived packet by sending it to the place indicated by the routing

- this always amounts to deciding which interface to send it on
- in IP the routing is captured in a *routing table* that shows, for every network or group of networks, which interface receives the packet
- in X.25 the routing is captured in a table that shows, for each incoming VC, which VC and interface to use

A Real Routing Table

from netlab.gmu.edu

Destination	Gateway	Interface
127.0.0.1	127.0.0.1	lo0
129.174.120.10	129.174.40.1	
131.182.121.0	199.56.135.1	
131.182.122.0	199.56.135.1	
129.174.142.29	129.174.40.15	
129.174.89.0	129.174.40.83	
129.174.91.0	129.174.40.83	
204.219.6.0	199.56.135.1	
204.218.132.0	199.56.135.1	
192.253.106.0	199.56.135.1	
129.174.65.0	129.174.65.1	le0
129.174.40.0	129.174.40.125	le1
199.56.135.0	199.56.135.2	le2
224.0.0.0	129.174.65.1	le0
default	129.174.40.1	



IP Routing

Some Basics

- A “network” may be a LAN or a WAN composed of packet switches
- IP WANs today are composed of routers acting as packet switches
- Routers exchange pathway information via a “routing protocol”
- Router may also have explicit table of “next hop” addresses and a “default” route
- Some routing protocols are:
  - RIP - distributed with Berkeley Unix
  - OSPF - “open” shortest-path-first
  - BGP - exterior gateways exchange Internet-level routing

Internet Routing Protocols

The major function is distributing the cost and reachability information. How this is done reflects the philosophy of routing and age of protocol

- Routing Information Protocol (RIP)
  - distance vector, uses hop count, uses UDP
  - has messages to exchange routing tables
  - comes with Unix
  - slow convergence
- Open Shortest Path First (OSPF)
  - link state, multiple metrics, uses IP directly
  - type of service routing (uses IP header TOS)
  - load balancing
  - information is exchanged among OSPF routers by Link State Advertisement (LSA) messages

Border Gateway Protocol (BGP)

- Internet gateways exchange Internet routing information among administrative domains
- gateway "advertises" that it can reach certain IP networks and its distance to them
- distance metrics not standardized
- exterior routing through autonomous systems
  - BGP4 allows use of more than one path for backup
  - to do this, advertise higher metric to undesirable paths
- uses TCP to exchange routing information
- important to distinguish between:
  - advertising: telling another administrative domain what nets are reachable through this one
  - routing: setting up the tables to support forwarding

Autonomous Systems (AS)

- An AS is a region of the Internet under the administrative control of a single entity (such as an Internet Service Provider).
  - Each AS has a unique 16 bit identifier (ASN)
  - Allows Internet architecture to scale
  - Allows different routing methods in different domains
  - basic elements: default routes and border routers
- It is common for major ISPs to establish pairwise "peering" gateways to ensure that their customers can communicate well

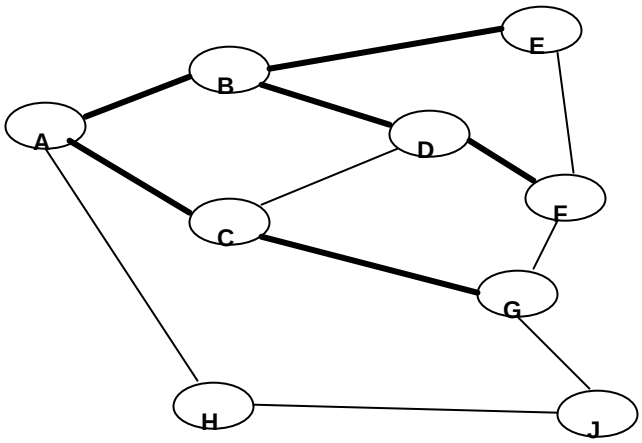
Internet Multicast

IP Multicast

- Multicasting: sending a packet to a group of addresses in a network
  - broadcast: sending to all addresses - impractical in WAN
- IPmc allows packets to be sent to “group” addresses
  - delivered in parallel
  - “class D” address 224.0.0.0 and greater
  - LANs have inherent broadcast/multicast capability
- In WAN, set of all nodes in a multicast group forms a tree
  - grows from the node that starts the group
- Routers at forks in tree must replicate packets
  - commercial routers available that perform this function
- Multicast requires a different routing protocol from unicast

Network with Multicast Tree

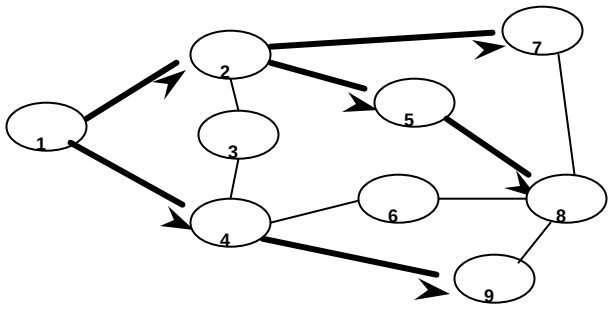
the network duplicates the packets



IP Multicast Directionality

- IPmc uses a many-to-many model of packet delivery
  - equivalent to full duplex for multicast
  - any participating host can send to the group
  - useful for collaborative-style applications
  - a very appealing model but increases complexity of the network layer considerably
- IETF recently developed an approach called Source Specific Multicast (SSM) that uses a one-to-many model
  - equivalent to half duplex for multicast
  - only one host can send to the group
  - useful for presentation-style applications

One-to-Many Multicast



IP Multicast Routing

- Internet Group Management Protocol (IGMP) runs over IP
  - distributes information between router and hosts on LAN
  - unlike ICMP it does not have a router-to-router role
  - like “ARP for multicast” put “pull” rather than “push”
- IPmc routing protocols are required for separate from unicast routing protocols
  - routing protocols support exchange of information among routers
  - different from routing (using the information to optimize flow)
  - also different from forward (implementing the routing by determining what to do with each arriving packet)
- IPmc routers use IP “tunnels” to span sections of the Internet that do not route IPmc
  - IPmc packets encapsulated in IP unicast

Multicast Routing Protocols

1 of 2

- Distance vector: each node maintains distance in multicast hops from itself to each destination (ex. RIP)
  - multicast version: Distance Vector Multicast Routing Protocol (DVMRP), a tunnel counts as one hop
- Link state: each router creates a “link state packet” for each link to a neighbor, containing cost, these are transmitted throughout network whenever cost changes (e.g. OSPF)
  - multicast version: Multicast Open Shortest Path First (MOSPF)
- Border Gateway: exterior routers advertise distance to other nets they can reach (e.g. BGP)
  - Multicast Border Gateway Protocol (MBGP) for inter-domain routing

Multicast Routing Protocols

2 of 2

- Newest: Protocol Independent Multicast (PIM) routing adds “relay points” for multicast-sparse portions of the Internet
  - PIM-Sparse Mode (PIM-SM) is intended to avoid redundant, hand-provisioned tunnels
  - for multicast-dense portions, PIM-Dense Mode (PIM-DM) is like DVMRP
  - for SSM, experimental PIM-SSM
  - primarily available in Cisco routers
- Proposed: Quality-of-Service Path First (QOSPF) would include ability to consider resource availability along the multicast path (see RSVP, below)
  - see http://netlab.gmu.edu/qosip

QoS Issues

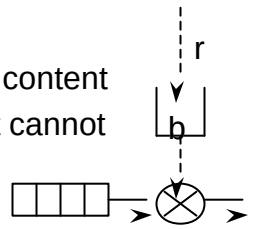
- Service parameters:
  - data rate (bps, pps)
  - end-to-end delay
    possibly also “jitter”, defined as variation in delay
  - packet loss rate
- Quality of Service:
  - best-effort (take what you get)
  - predictive (statistics known) with delay bound
  - guaranteed performance
  - except best-effort, QoS can only be defined in context of capacity requirement (pps or bps)
- The Internet can’t do anything but best-effort at present
- Integrated Services and Differentiated Services aim at changing this
  - thus far, successful only in private networks

Internet Integrated Services

- A group of Proposed Standard RFCs defines QoS categories for the Internet, on an end-to-end connection or "flow" basis, defined to have peak rate in bytes per second, min and max transmission units, and token bucket
- Basic properties: delay, packet loss
- Controlled Load (RFC2211): network expected to offer service as if in lightly-loaded best-effort situation
- Guaranteed QoS (RFC2212): network guarantees to provide service with specified delay and jitter, given definition of peak rate and assurance that this rate will not be exceeded

Token Bucket

- Based on "leaky bucket" flow control*
  - bucket starts out with *b* permits
  - permits arrive at rate *r* per second
  - every packet sent decrements bucket content
  - if bucket is empty, and arriving packet cannot be transmitted
- IntServ Token Bucket (RFC2215) implements this scheme
  - results in "rate-controlled" transmission



to pass, packet needs a token

* see Bertsekas & Gallager *Data Networks 2nd ed* pp 511-512

Bursty Traffic

- Applications may require periodic transmission of large amounts of data
- If data requires multiple packets to send, they become a "burst"
  - correlation destroys assumption of statistical independence
  - for NVE, the bursts also are correlated- bad news!
- A major advantage of packet networks is their ability to send bursts at high data rates while still serving many clients
  - but an overloaded packet network can choke on large numbers of bursts
  - this is why TCP congestion control helps so much

RSVP Messages

- RSVP is a setup or "signaling" protocol to implement IntServ concept
- Downstream: PATH
  - forwarded along data paths of multicast tree
  - path state is stored at each node
  - "soft state" decays if not refreshed periodically
- Upstream: RESV
  - follows multicast tree back to other nodes
  - requests reservation of system resources
- For more information:
  http://www.ietf.org/rfc/rfc2205.txt

Differentiated Services

- RSVP entails a lot of overhead
  - current routers cannot support at scale of today's Internet
- A simpler approach called "differentiated services" works by designating two classes of service
  - limited amounts of QoS-sensitive traffic
  - all other traffic best-effort
- QoS traffic is always processed first
- "Edge" routers are allocated set amounts of QoS traffic
  - edge routers police QoS packet entry

ATM: Networking With Built-In QoS

- Low transmission error rates allow streamlined protocol
  - connection-oriented; does scale as well as IP
  - but can handle heterogeneous traffic mixes with committed cell loss rate
- At high speeds, digital switching is easier than traditional multiplexing
- A solution for a large ISP: install ATM backbone
  - use it to derive circuits among routers as needed

MPLS

- How can ATM and IP work together to do a better job of meeting network requirements?
  - Connectionless IP is very flexible and adaptable
  - Connection-oriented ATM offers high performance and guaranteed QoS
- New approach Multi-Protocol Label Switching (MPLS) addresses this:
  - routers use ATM-derived “pipes” for predefined IP packet flows
  - the pipes are distinguished by “labels” that can be used directly for ATM switching
  - the approach is identified with “traffic engineering”
    - high-level setup by human network engineers