Lecture 13
More IP

Peter Steenkiste
School of Computer Science
Carnegie Mellon University

15-441 Networking, Spring 2006
http://www.cs.cmu.edu/~prs/15-441

Coming Attractions

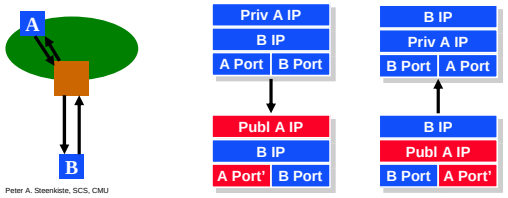
- Project 2 is due Thursday evening …
- Mid-semester grades will be based on the two homeworks, project 1, and the midterm
- Project 3 will be handed out the 2nd week after Spring break and will be due the 6th week after Spring break, i.e. one week before the last day of classes.
- We will also have ~3 more homeworks.
  - At least one will be hands-on lab

Outline

- NAT.
- Tunneling.
- IPv6.
- SNMP.
- IP multicast.

Network Address Translation NAT

- NAT maps (private source IP, source port) onto (public source IP, unique source port)
  - reverse mapping on the way back
  - destination host does not know that this process is happening
- Very simple working solution.
  - NAT functionality fits well with firewalls



NAT Considerations

- NAT translation must be consistent during a session.
  - Set up mapping at the beginning of a session and maintain it during the session
  - Recycle the mapping that the end of the session
- Must determine the end of "sessions" so entries can be retired.
  - Relatively easy for TCP (but be careful about retransmissions)
  - Harder for UDP since NAT does not "understand" protocol
  - Typically uses a timer based mechanism
- NAT has to be consistent with other protocols.
  - ICMP, routing, …
- Many flavors of NAT exist.
  - Basic, network address port translation (NAPT), bi-directional,..

NAT Challenges

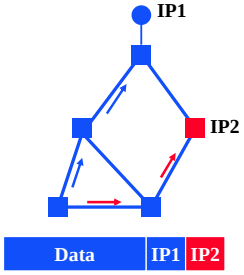
- NAT breaks the basic IP-based connection model used in the Internet.
- NAT creates problems for certain classes of applications.
  - E.g., applications that pass IP information in payload
- Solution is to make NAT aware of these protocols.
  - Unfortunately this only works for standard protocols, e.g. special support for FTP in NATs
- NATs continue to be a problem for some applications, e.g. peer-to-peer applications.
  - Has results in the development of lots of tricks to punch holes through NATs

Outline

- NAT.
- Tunneling.
- IPv6.
- SNMP.
- IP multicast.

Tunneling

- Force a packet to go to a specific point in the network.
  - » Path taken is different from the regular routing
- Achieved by adding an extra IP header to the packet with a new destination address.
  - » Similar to putting a letter in another envelop
  - » preferable to using IP source routing option
- Used increasingly to deal with special routing requirements or new features.
  - » Mobile IP,..
  - » Multicast, IPv6, research, ..

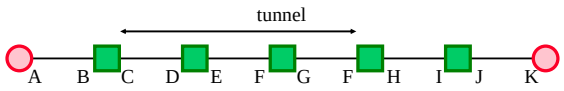


IP-in-IP Tunneling

- Described in RFC 1993.
- IP source and destination address identify tunnel endpoints.
- Protocol id = 4.
  - » IP
- Several fields are copies of the inner-IP header.
  - » TOS, some flags, ..
- Inner header is not modified, except for decrementing TTL.

V/HL	TOS	Length
ID		Flags/Offset
TTL	4	H. Checksum
Tunnel Entry IP		
Tunnel Exit IP		
V/HL	TOS	Length
ID		Flags/Offset
TTL	Prot.	H. Checksum
Source IP address		
Destination IP address		
Payload		

Tunneling Example



Tunneling Considerations

- Tunnels are currently standardized.
  - » Some diversity in initial implementations
  - » Early versions sometimes merged with multicast code
- Performance.
  - » Tunneling adds (of course) processing overhead
  - » Tunneling increases the packet length, which may cause fragmentation
    - – BIG hit in performance in most systems
    - – Tunneling in effect reduces the MTU of the path, but end-points may not know this
- Security issues.
  - » Should verify both inner and outer header
  - » Tunneling often used with "IP sec" – see Security lectures

Tunneling Applications

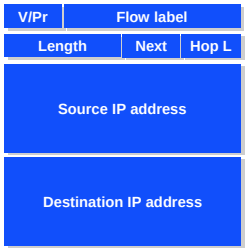
- Virtual private networks.
  - » Connect subnets of a corporation using IP tunnels
  - » Often combined with IP Sec
- Support for new or unusual protocols.
  - » Routers that support the protocols use tunnels to "bypass" routers that do not support it
  - » E.g. multicast
- Force packets to follow non-standard routes.
  - » Routing is based on outer-header
  - » E.g. mobile IP

Outline

- NAT.
- Tunneling.
- IPv6.
- SNMP.
- IP multicast.

IP v6

- "Next generation" IP.
- Most urgent issue: increasing address space.
  - » 128 bit addresses
- Simplified header for faster processing.
- Many other changes.
- Support for guaranteed services: priority and flow id
- Options handled as "next header"
  - » reduces overhead of handling options



IPv6 Addressing

- 128 bit addresses with complex structure.
- Examples: format for local configuration, IPv4 backwards compatible, ..
- Provider-based unicast addressing extends the format used in IPv4 with CIDR.
  - » Eventually supposed to be the primary addressing model

Migration from IPv4 to IPv6

- Interoperability with IP v4 is necessary for gradual deployment.
- Two complementary mechanisms:
  - » dual stack operation: IP v6 nodes support both address types
  - » tunneling: tunnel IP v6 packets through IP v4 clouds
- Alternative is to create IPv6 islands, e.g. corporate networks, ...
  - » Use of form of NAT to connect to the outside world
  - » NAT must not only translate addresses but also translate between IPv4 and IPv6 protocols

IPv6 Discussion

- Unfortunately there is little motivation for any one organization to move to IP v6.
  - » the challenge is the existing hosts (using IPv4 addresses)
  - » little benefit unless one can consistently use IPv6
    - – can no longer talk to IPv4 nodes directly
- People have continued to improve the IPv4 infrastructure.
  - » stretching address space through address translation seems to work reasonably well
  - » New standards, e.g. IP Sec, diff serv, ..
- Networking increasingly supports IPsec.

Outline

- NAT.
- Tunneling.
- IPv6.
- SNMP.
- IP multicast.

Network Management - Very Hard Problem

- What to do when there is a problem?
  - Loss of connectivity, complaints of slow throughput, ..
- How do you know how busy your network is?
  - Where are the bottlenecks, is it time for an upgrade, redirect traffic, ..
- How can you spot unusual activity?
  - Somebody attacking a subnet, ..
- How do you plan upgrades and schedule maintenance.
  - Minimize disruption for customers
  - Predict impact of unavailable links
  - Impact of hurricane

Important Component: Monitoring

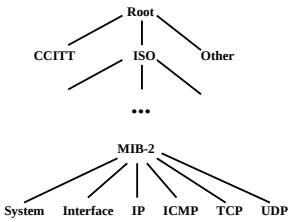
- Need to be able to monitor the health of all the components in the network.
- "Static" information: what is connected to what?
- Dynamic information: what is the load on that link?
- Raise alarms if certain parameters are out of bounds.
- Note that lack of alarms does not necessarily imply that all your customers are happy!

Simple Network Management Protocol (SNMP)

- Protocol that allows clients to read and write management information on network elements.
  - Routers, switches, …
  - Network element is represented by an SNMP agent
- Information is stored in a management information base (MIB).
  - Have to standardize the naming, format, and interpretation of each item of information
  - Ongoing activity: MIB entries have to be defined as new technologies are introduced
- Different methods of interaction supported.
  - Query response interaction: SNMP agent answers questions
  - traps: agent notifies registered clients of events
- Need security: authentication and encryption.

MIB

- Information is represented in an object tree.
  - To identify information you specify a path to a leaf
  - Can extend MIB by adding subtrees
  - Different standard bodies can expand different subtrees
    - E.g. Ethernet and ATM groups are independent
- Uses ASN.1 standard for data representation.
  - Existing standard
  - How is information stored?
  - How is information encoded on the wire (transfer syntax)



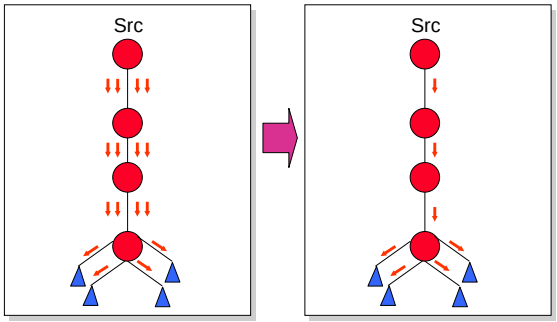
Outline

- NAT.
- Tunneling.
- IPv6.
- SNMP.
- IP multicast.
  - Motivation
  - Routing
  - Transport challenges

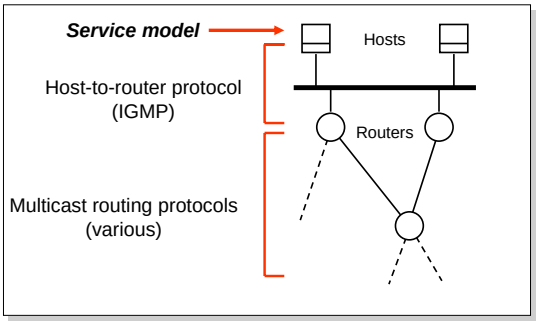
Group Communication Applications

- Broadcast audio/video
- Software distribution
- Web-cache updates
- Teleconferencing (audio, video, shared whiteboard, text editor)
- Multi-player games
- Server/service location
- Other distributed applications

Multicast – Efficient Data Distribution



IP Multicast Architecture

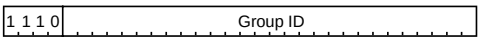


IP Multicast Service Model (RFC1112)

- Each group identified by a single IP address
- Groups may be of any size
- Members of groups may be located anywhere in the Internet
- Members of groups can join and leave at will
- Senders need not be members
- Group membership not known explicitly
- Analogy:
  - Multicast addresses are like radio frequencies, on which anyone can transmit, and to which anyone can tune-in
- What is the common motivation for these design decisions?

IP Multicast Addresses

- Class D IP addresses
  - 224.0.0.0 – 239.255.255.255



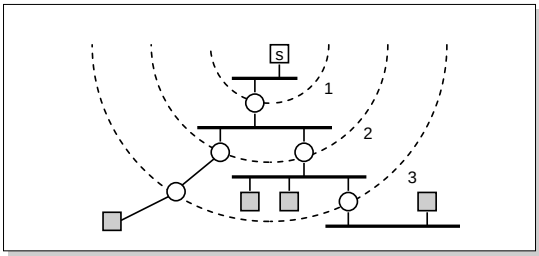
- How to allocated multicast addresses?
  - Well-known multicast addresses, assigned by IANA
  - Transient multicast addresses, assigned and reclaimed dynamically, e.g., by "sdr" program

IP Multicast Service

- Sending – same as before
- Receiving – two new operations
  - Join-IP-Multicast-Group(group-address, interface)
  - Leave-IP-Multicast-Group(group-address, interface)
  - Receive multicast packets for joined groups via normal IP-Receive operation

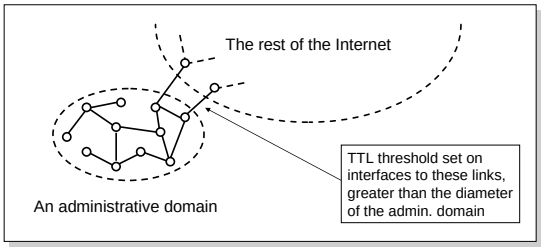
Multicast Scope Control – Small TTLs

- TTL expanding-ring search to reach or find a nearby subset of a group



Multicast Scope Control – Large TTLs

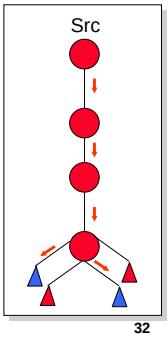
- **Administrative TTL Boundaries to keep multicast traffic within an administrative domain, e.g., for privacy or resource reasons**



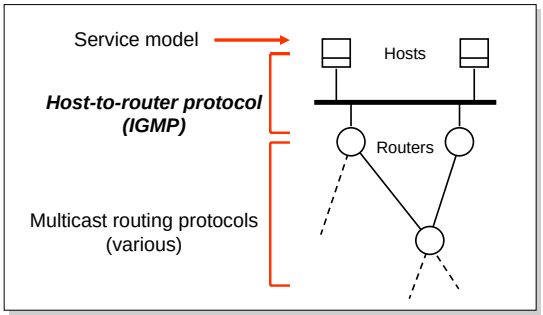
Multicast Router Responsibilities

- **Learn of the existence of multicast groups (through advertisement)**
- **Identify links with group members**
- **Establish state to route packets**
  - » **Replicate packets on appropriate interfaces**
  - » **Routing entry:**

Src, incoming interface	List of outgoing interfaces



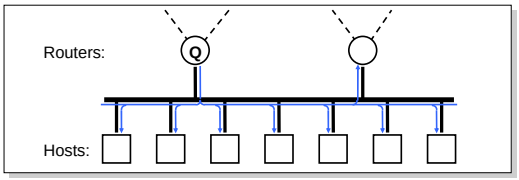
IP Multicast Architecture



Internet Group Management Protocol

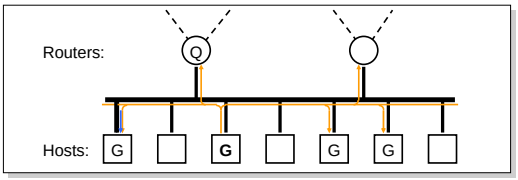
- **End system to router protocol is IGMP**
- **Each host keeps track of which multicast groups it is subscribed to**
  - » **Socket API informs IGMP process of all joins**
- **Objective is to keep router up-to-date with group membership of entire LAN**
  - » **Routers need not know who all the members are, only that members exist**

How IGMP Works



- **On each link, one router is elected the "querier"**
- **Querier periodically sends a Membership Query message to the all-systems group (224.0.0.1), with TTL = 1**
- **On receipt, hosts start random timers (between 0 and 10 seconds) for each multicast group to which they belong**

How IGMP Works (cont.)

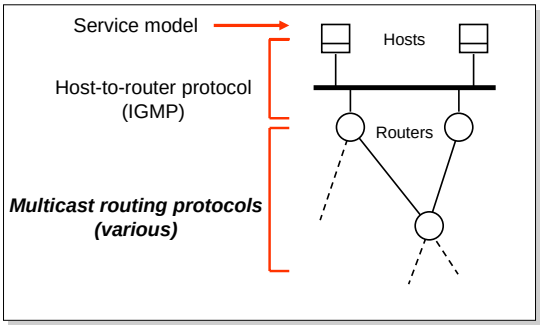


- **When a host's timer for group G expires, it sends a Membership Report to group G, with TTL = 1**
- **Other members of G hear the report and stop their timers**
- **Routers hear all reports, and time out non-responding groups**

How IGMP Works (cont.)

- Note that, in normal case, only one report message per group present is sent in response to a query
- Query interval is typically 60-90 seconds
- When a host first joins a group, it sends one or two immediate reports, instead of waiting for a query

IP Multicast Architecture



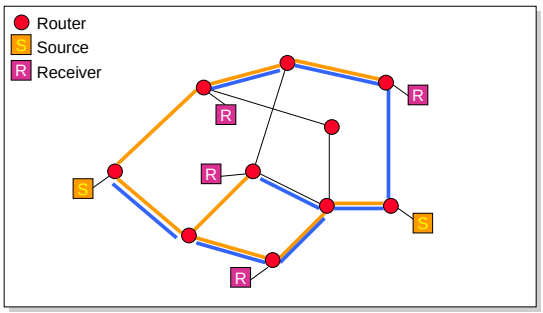
Multicast Routing

- Basic objective – build distribution tree for multicast packets
- Multicast service model makes it hard
  - Anonymity
  - Dynamic join/leave
  - Scalability

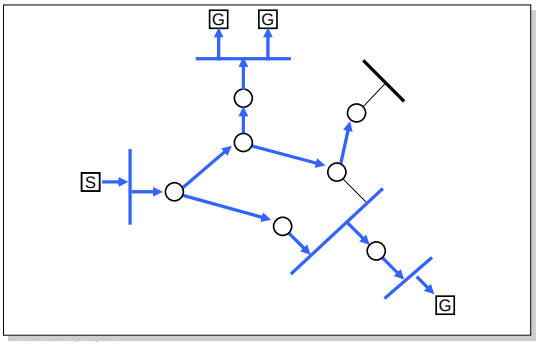
Routing Techniques

- Flood and prune
  - Begin by flooding traffic to entire network
  - Prune branches with no receivers
  - Examples: DVMRP, PIM-DM
  - *Unwanted state where there are no receivers*
- Link-state multicast protocols
  - Routers advertise groups for which they have receivers to entire network
  - Compute trees on demand
  - Example: MOSPF
  - *Unwanted state where there are no senders*
- Rendez-vous based protocols
  - Interested receivers send packet to an IP address
  - When packet reaches a router that is part of the tree, a branch is added to support the new receiver

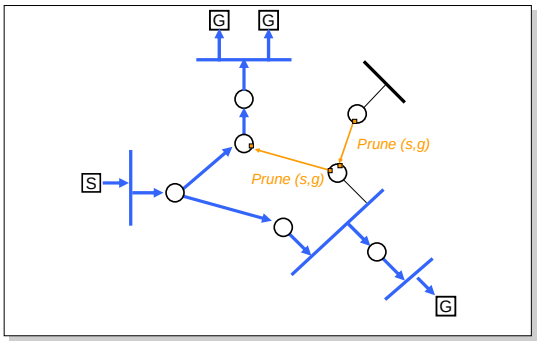
Source-based Trees



Broadcast with Truncation



Prune



Multicast Transport Challenges

- **Recovery from packet loss**
  - Implosion of many receiver ask sender to retransmit
  - How to retransmit: unicast or multicast?
- **Flow control with heterogeneous receivers**
  - Send at the rate supported by the slowest receiver?
  - Penalizes the "fast" receivers
  - How does the sender even know receiver properties
- **Congestion control**
  - Different parts of the network can support different rates
  - How can the sender discover these rates?
- **These are fundamentally hard problems for large multicast groups**
  - Virtually impossible to handle correctly at the IP layer
  - Overlay multicast is a more attractive option (later in the course)