CSE/EE 461 – Lecture 16

David Wetherall
djw@cs.washington.edu

Last Time

- A whirlwind tour of network security
- Focus
  - How do we secure distributed systems?
- Topics
  - Privacy, integrity, authentication
  - Cryptography and key distribution
  - Firewalls and Denial-of-service
  - TCP/IP vulnerabilities

Application
Presentation
Session
Transport
Network
Data Link
Physical

This Time

- Multicast. See Keshav 11.11.
- Focus
  - How do we communicate efficiently with a group of participants
- Topics
  - Group communication
  - Multicast routing (DVMRP, PIM/CBT)
  - Future: reliable multicast

Application
Presentation
Session
Transport
Network
Data Link
Physical

Group Communication

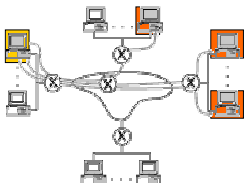
- Many applications involve group communication
  - Quake, conferencing (vic), stock quote distribution,
  - VOD/Internet "TV", software updates, resource discovery …
- Semantics issues
  - Many-to-many or many-to-one communication
  - Consistent delivery order across all members?
- We concentrate on efficient/scalable multicast routing
  - Multicast = send to multiple receivers at once
  - Unicast = send to a single receiver (regular IP)
  - Broadcast = send to all receivers

Why not "Send N Copies"?

- It doesn't scale with group size N
  - For some applications (Internet TV) N can be huge
  - For other applications (Quake) this would be reasonable
- As N grows
  - The source needs to track N members
  - Effective bandwidth near the source is reduced by N
  - Latency to do the multicast grows with N

Example – Repeated Unicast

- Send one copy to each of three receivers
  - Routers do not participate in any special manner

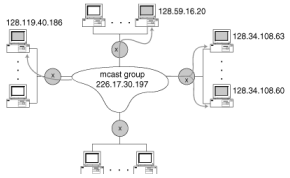


IP Multicast Service Model

- Extend IP "best effort" model for efficient multipoint delivery
  - A single message is sent by any source to reach all receivers
  - Routers take care of the details of delivery!
  - IP multicast address (class D) identifies a group
  - Many-to-many delivery is supported
- Receivers explicitly join/leave a group to receive messages
  - Each receiver contacts the local designated router using IGMP
  - IGMP = Internet Group Management Protocol
  - Receivers learn multicast address via an out-of-band channel
- Senders don't know group membership
  - Multicast address provides a level of indirection
  - Useful for rendezvous / resource discovery
  - Anyone can send to a multicast group w/o explicit setup

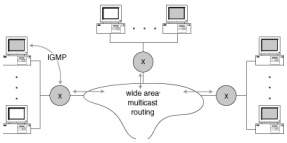
Example – Multicast Group

- Set of receivers associated with a group is dynamic



Example – IGMP

- By convention, hosts don't participate in routing
  - IGMP gives local router sufficient info to act as its agent

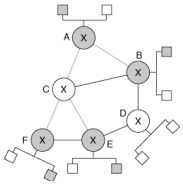


Multicast on Broadcast Media

- The question we really want to answer: how do we route multicast packets in a network so that they reach the right receivers
  - Over a broadcast link this is easy
- Ethernet readily allows all hosts to receive frames
  - Some addresses reserved for multicast
  - Interfaces subscribe to their multicast groups
    - Or receive in promiscuous mode and filter
- Can extend to extended (bridged) LANs
  - Bridges forward all multicast traffic; it will reach all LANs
  - Spanning tree provides loop avoidance

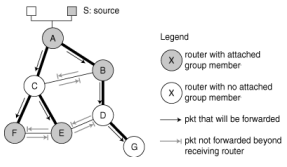
Multicast in an Internet

- Problem: How to set up router forwarding tables to send to different groups (example below)?



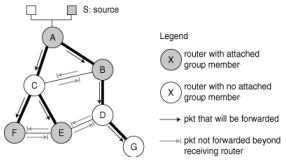
Reverse Path Broadcast (RPB)

- Observation: can broadcast by forwarding along reverse routes!
  - At each node: look up source and check packet came from "output" link using unicast routes
  - If so, send packet in "reverse" on all other interfaces



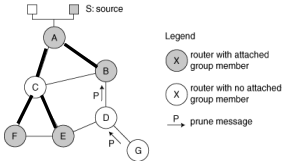
RPB without Duplicates

- Problem: With RPB some nodes get duplicates
- Solution: Use unicast route to determine your children
  - e.g., C knows B doesn't use it to get to A, so C won't send to B
  - Leaves darkened tree only



RPB with Pruning

- Problem: Even w/o duplicates, we still flood network
- Solution: Prune away tree branches with no members
  - E.g, G prunes to D; all children of D are gone, so D prunes too
  - Typically prune on demand and expire prune information

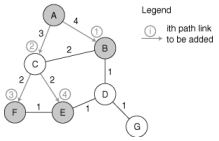


DVMRP

- DVMRP = Distance Vector Multicast Routing Protocol
- Early multicast routing protocol still used in Internet
- Distance vector used to calculate source spanning trees
- Multicast with reverse path forwarding and pruning
  - For each group and source, router maintains next hop routers

Multicast with Link State

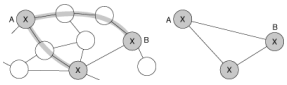
- Can pass around member locations and compute pre-pruned per-source spanning trees at each node



- MOSPF (Multicast OSPF) is an example of this approach

MBONE (Multicast Backbone)

- IP multicast routing requires that routers be upgraded
  - But "native" multicast isn't available everywhere



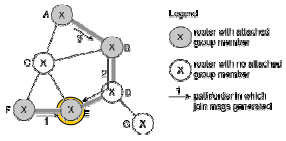
- Multicast in the Internet happens over the MBONE
  - An overlay with tunnels between multicast nodes
  - Within this overlay multicast appears available "everywhere"

Scaling Issues

- DMVRP, MOSPF require each router to maintain state for each group (G) and each source (S): S x G entries
  - This quantity grows quickly if multicast takes off
  - Hierarchical aggregation is difficult for multicast addresses
- Approach to solve for the wide-area:
  - Use a single shared spanning tree per group for all sources
    - Relies on notion of a core or rendezvous point
  - Only routers on the spanning tree keep forwarding state
    - Dense-mode versus Sparse-mode

PIM/CBT

- Protocol Independent Multicast (PIM), Core Based Trees (CBT)
  - Shared tree for each group, explicit joins to tree build routes
  - PIM uses a rendezvous point (RP), CBT a core



- Pros: scalability
- Cons: finding the RP/core, performance, complexity

Multicast Transport Protocols

- We discussed "best effort" (unreliable) IP multicast
- Multicast transport protocols are an open research area

- Heterogeneous receivers
  - How to we send to receivers with different capabilities?
    - "layered" video has been used for bandwidth variation
- Reliable multicast adds more complexities
  - How do receivers acks the source without overwhelming?
  - How to retransmit lost packets to just the receivers who lost?
  - How do we adjust the transmission rate?

Key Concepts

- Many apps can benefit from group communication
- Multicast routing allows efficient multi-point messages to be sent over an internetwork
- Reverse Path Broadcast (RBP) an elegant technique
  - But it's not scalable to wide-area multicast
- Making multicast scale is hard
  - Can be much state in routers (can be S x G)
  - Can't easily aggregate multicast routing info