EE122: Multicast

Kevin Lai
October 7, 2002
Internet Radio

- **www.digitallyimported.com** (techno station)
  - sends out 128Kb/s MP3 music streams
  - peak usage ~9000 simultaneous streams
    - only 5 unique streams (trance, hard trance, hard house, eurodance, classical)
  - consumes ~1.1Gb/s
    - bandwidth costs are large fraction of their expenditures (maybe 50%?)
  - if 1000 people are getting their groove on in Berkeley, 1000 unicast streams are sent from NYC to Berkeley
Multicast Service Model 1

- receivers join a multicast group which is identified by a multicast address (e.g. G)
- sender(s) send data to address G
- network routes data to each of the receivers
Motivation

- Conserve bandwidth
  - use same bandwidth/link to send to n receivers as 1 receiver
    - internet radio example: reduce bandwidth consumed by 9000/5=1800
  - deals with flash crowds
  - e.g., video/audio conferencing, streaming, news dissemination, file updates

- Separate identifier from address (logical addressing)
  - receiver can change location-dependent addresses without notifying sender
  - sender doesn’t need to know about receivers
  - e.g., service location, mobility, anonymity, naming
Multicast Service Model 2

- Membership access control
  - open group: anyone can join
  - closed group: restrictions on joining

- Sender access control
  - anyone can send to group
  - anyone in group can send to group
  - only one host can send to group

- Packet delivery is best effort
Multicast and Layering

- Multicast can be implemented at different layers
  - data link layer
    - e.g. Ethernet multicast
  - network layer
    - e.g. IP multicast
  - application layer
    - e.g. as an overlay network like Kazaa

- Which layer is best?
Multicast Implementation Issues

- How are multicast packets addressed?
- How is join implemented?
- How is send implemented?
- How much state is kept and who keeps it?
Ethernet Multicast

- Reserve some Ethernet MAC addresses for multicast
- join group G
  - network interface card (NIC) normally only listens for packets sent to unicast address A and broadcast address B
  - to join group G, NIC also listens for packets sent to multicast address G (NIC limits number of groups joined)
    - implemented in hardware, so efficient
- send to group G
  - packet is flooded on all LAN segments, like broadcast
  - can waste bandwidth, but LANs should not be very large
- only host NICs keep state about who has joined → scalable to large number of receivers, groups
Problems with Data Link Layer Multicast

- single data link technology
- single LAN
  - limited to small number of hosts
  - limited to low diameter latency
  - essentially all the limitations of LANs compared to internetworks
- broadcast doesn’t cut it in larger networks
IP Multicast

- Overcomes limitations of data link layer multicast
- Performs inter-network multicast routing
  - relies on data link layer multicast for intra-network routing
- Portion of IP address space defined as multicast addresses
  - $2^{28}$ addresses for entire Internet
- Open group membership
- Anyone can send to group
  - flexible, but leads to problems
IP Multicast Routing

- **Intra-domain**
  - Distance-vector multicast (DVM)
  - Link-state multicast (LSM)

- **Inter-domain**
  - Protocol Independent Multicast (PIM)
  - Single Source Multicast (SSM)
Distance Vector Multicast

- Extension to DV unicast routing
- Routers compute shortest path to each host
  - necessary for unicast delivery
- No join required
  - every link receives a copy, even if no interested hosts
- Packet forwarding
  - iff incoming link is shortest path to source
  - out all links except incoming
  - Reverse Path Flooding (RPF)
  - packets always take shortest path
    - assuming delay is symmetric
  - link may have duplicates

laik@cs.berkeley.edu
Reverse Path Broadcasting (RPB)

- Extend DV to eliminate duplicate packets
- Combine DV and spanning tree
- Choose parent router for each link
  - router with shortest path to source
  - lowest address breaks ties
  - each router can compute independently from already known information
  - each router keeps a bitmap with one bit for each of its links
- Only parent forwards onto link
Truncated Reverse Path Broadcasting (TRPB)

- Extend DV/RPB to eliminate unneeded forwarding
- Identify leaves
  - routers announce that a link is their next link to source S
  - parent router can determine that it is not a leaf
- Explicit group joining
  - members periodically (with random offset) multicast report locally
  - hear an report, then suppress own
- Packet forwarding
  - iff not a leaf router or have members
  - out all links except incoming
Problems with IP Multicast Model

- Few groups have many senders
  - difficult to construct optimal tree for many senders

- Hard to implement sender control → any node can send to the group
  - open group membership

- Multicast address scarcity
  - $2^{28}$ addresses may not be enough for entire Internet
  - how prevent collisions?
Internet Radio using IP Multicast Model

- One sender
  - does not use multiple sender capability of model
- Someone other than Digitally Imported can send to group
  - clog 9000 clients’ links with useless data (Denial-of-Service attack)
- How can Digitally Imported get and keep a multicast address?
  - central organization to manage addresses adds overhead
Single Source Multicast (SSM)

- Network layer multicast
- SSM service model
  - only one sender can send to a group
  - any number of receivers
- Addressing
  - SSM address = (S, G) S: IP address of source, G: 24-bit group address
  - each sender has its own G-space
SSM Join

- receiver sends join to source
- routers on the path read the join packet
- they note a receiver on the incoming link
SSM Send

- router checks that packet is coming from direction of S
- if so, forward it down links that have receivers
SSM v.s. IP Multicast

- Restricted to one sender per group
  - for multiple senders, make multiple groups
- Can prevent denial-of-service attacks on group
- Senders can independently allocate multicast addresses
- Much simpler than other network layer multicast routing schemes
Problems with Network Layer Multicast (NLM)

- Scales poorly with number of groups
  - A router must maintain state for every group that traverses it
  - many groups traverse core routers
- Supporting higher level functionality is difficult
  - NLM: best-effort multi-point delivery service
  - Reliability and congestion control for NLM complicated
- Deployment is difficult and slow
  - ISP’s reluctant to turn on NLM
- Massive reduction in bandwidth not necessary
  - Internet Radio is using 1.1Gb/s
  - network layer multicast would reduce to 5 x 128Kb/s
NLM Reliability

- Assume reliability through retransmission
- Sender can not keep state about each receiver
  - e.g., what receivers have received
  - number of receivers unknown and possibly very large
- Sender can not retransmit every lost packet
  - even if only one receiver misses packet, sender must retransmit, lowering throughput
- N(ACK) implosion
  - described next
(N)ACK Implosion

- (Positive) acknowledgements
  - ack every n received packets
  - what happens for multicast?
- Negative acknowledgements
  - only ack when data is lost
  - assume packet 2 is lost
NACK Implosion

- When a packet is lost all receivers in the sub-tree originated at the link where the packet is lost send NACKs.
Application Layer Multicast

- Provide multicast functionality above the IP layer
- overlay/application layer/end system multicast
- sender sends out multiple copies of packet
- receivers forward them to other receivers
- challenge
  - minimize delay
  - minimize bandwidth consumed
- Narada [Yang-hua et al, 2000]
  - Small group sizes <= hundreds of nodes
  - Typical application: chat
- Peercast
Narada: End System Multicast
Performance Concerns

- **Stretch**
  - ratio of latency in the overlay to latency in the underlying network

- **Stress**
  - number of duplicate packets sent over the same physical link
Performance Concerns

Delay from CMU to Berk1 increases

Duplicate Packets: Bandwidth Wastage

laik@cs.berkeley.edu
Properties

- easier to deploy than NLM
  - don’t have to modify every router on path
- easier to implement reliability than NLM
  - use hop-by-hop retransmissions
- can consume more bandwidth than NLM
- can have higher latency than NLM
- not clear which scales better
  - neither has been used for a group with 1M receivers or 1M groups
- can use network layer multicast where available to optimize performance
Summary

- Large amount of work on multicast routing
- Major problems
  - preventing flooding
  - minimizing state in routers
  - denial-of-service attacks
  - deployment
- Multicast can be implemented at different layers
  - lower layers optimize performance
  - higher layers provide more functionality
- IP Multicast still not widely deployed
  - Ethernet multicast is deployed
  - application layer multicast systems just released