Multicasting I
Multicast Groups & Spanning Trees

Kevin Jeffay
Department of Computer Science
University of North Carolina at Chapel Hill
jeffay@cs.unc.edu
October 21, 1999

http://www.cs.unc.edu/~jeffay/courses/comp249f99

Multicast Data Delivery & Multicast Routing
A critical enabling technology for multimedia networking

- **Unicast**
  » 1-to-1
- **“Anycast”**
  » 1-to-anyone
- **Broadcast**
  » 1-to-all
- **Multicast**
  » 1-to-many
Multicast Data Delivery

Outline

- Multicast addressing
- Group management
- Principles of multicast routing
  - Flooding
  - Spanning trees
  - Core-based trees
- Multicast routing algorithms
  - DVMRP
  - MOSPF
  - PIM
- The MBONE multicast backbone

Multicast Addressing

The 4 major classes of IP addresses

- **Class A addresses**
  - 128 networks
  - more than 65,536 hosts
  - 0 1 2 3 4 8 16 24 31
  - net id host id

- **Class B addresses**
  - 16,384 networks
  - 256 to 65,536 hosts
  - 0 1 2 3 4 8 16 24 31
  - 1 0 net id host id

- **Class C addresses**
  - $2^{21}$ networks
  - less than 256 hosts
  - 0 1 2 3 4 8 16 24 31
  - 1 1 0 net id host id

- **Class D addresses**
  - 28-bit multicast addresses
  - no origin or network information is encoded
  - 0 1 2 3 4 8 16 24 31
  - 1 1 1 0 multicast address
  - (IP addresses 224.0.0.0 through 239.255.255.255)
Multicast Addressing
Multicast addressing & multicast groups

- Multicast addresses can only be used as a destination address
- Multicast addresses correspond to a multicast group
- Special group addresses
  - Local/Link-level scope
    - 224.0.0.0 — reserved
    - 224.0.0.1 — all hosts group, all multicast capable hosts & routers
    - 224.0.0.2 — all routers group, multicast capable routers only
    - ...  
    - 224.0.0.255

Multicast Addressing
IPv4 multicast address space

- 224.0.0.0 - 224.0.0.255 — Reserved for routers and group management
- 224.0.1.0 - 238.255.255.255 — Largely unreserved but some address have already been “bought”
  - 224.0.19.0 - 224.0.19.63 owned by Walt Disney
- 239.0.0.0 - 239.255.255.255 — administratively scoped applications
  - 239.192.0.0 - 239.192.63.255 — IPv4 Organizational Scope
  - 239.255.0.0 - 239.255.255.255 — IPv4 Local Scope
Multicast Addressing

Multicast groups

- Groups may be of *any* size
- Group members may be located *anywhere* in the Internet
- Hosts can join and leave groups at will
- There is no "list" of group members
- A sender cannot tell who (or if anyone) received any message
- *Senders need not be members of the group*

Multicast Datagram Delivery

Efficient delivery across LANs

- Virtually all LAN technologies support multicast in hardware
- Simple mapping from IP address (32 bits) to Ethernet addresses (48 bits)
  - Low order 23 bits of IP address mapped to low order 23 bits of Ethernet MAC address (01.00.5E.00.00.00_{16})
Hosts can always send and receive locally generated multicast packets by themselves.

To receive multicast packets from the outside, a multicast router must be present.

Hosts subscribe to a multicast group by issuing IGMP host membership report messages:
- Multicast to the all hosts group with a TTL of 1
- A local multicast router...
  - records the group address
  - begins forwarding packets destined for that group
IGMP Operation

Host membership query messages

- Hosts can join and leave groups dynamically

- Local multicast routers periodically poll hosts for group membership via *host membership query* messages
  - If no response, router stops advertising group membership to other routers

Host Membership Query Messages

Avoiding a *group membership report* explosion

- Multicast routers poll local hosts for membership in *any* multicast group
  - Query message multicast to the *all hosts group*

- Hosts wait a random delay before responding
  - While waiting, if a *host membership report* is heard, a host ignores the original query

Adapted from Comer p. 296
**Host Membership Report Messages**

**Example**

- Host Membership Report, TTL = 1
  - Group address = <group addr 1>
  - Dest IP address = <group addr 1>
  - Src IP address = <A’s addr>

- Host Membership Report, TTL = 1
  - Group address = <group addr 2>
  - Dest IP address = <group addr 2>
  - Src IP address = <B’s addr>

- Host Membership Report, TTL = 1
  - Group address = <group addr 1>
  - Dest IP address = <group addr 1>
  - Src IP address = <C’s addr>

- Host Membership Report, TTL = 1
  - Group address = <group addr 3>
  - Dest IP address = <group addr 3>
  - Src IP address = <C’s addr>

**IGMP Design**

**Scalability & fault tolerance**

- All communication between hosts and routers uses IP multicast
- Redundant *host membership reports* are suppressed
- *Host membership query* responses are randomly delayed
- Router state is periodically refreshed
Multicast Routing
Recursively apply IGMP into the Internet?

Construct “receiver-based” trees?

Multicast Routing
Basic building blocks

How are multicast packets delivered to all group members everywhere?

*flooding* — Upon receipt of a multicast packet...
  » If a router has never seen the packet before it forwards the packet on all interfaces (except the incoming)
  » Otherwise it is discarded
**Multicast Routing**

**Basic building blocks**

- Avoid the problem of duplicate arrivals by creating a *spanning tree* (an acyclic graph spanning all nodes)

- Build a spanning tree and only transmit packets on branches of the tree

**Flooding-Based Routing**

*reverse path forwarding (RPF)*

- Compute a source-based tree for each multicast source
- When a packet arrives, note source $S$ and input interface $I$
  - If $I$ is on the shortest path to $S$, forward the packet
  - Otherwise, drop
Reverse Path Forwarding
Source-based tree examples

Reverse Path Flooding
Dealing with equal cost paths

- Is RPF guaranteed to build a spanning tree?
  > What to do about equal cost paths?
Using Group Membership Information

Truncated broadcast

- Integrate multicast routing with IGMP
  - A router forwards multicast packets only if it knows one of its downstream hosts in a member of the appropriate group

Using Group Membership Information

RPF with pruning

- First multicast packet from a source $S$ is flooded to all nodes
- Upon receipt at a leaf router...
  - If there are multicast group members behind the router, do nothing
  - Otherwise reply with a “prune ($S$, group)” message
Using Group Membership Information

RPF with pruning

- Prune message are recursively propagated back towards the source

Prune list
(interface = \(I_1\), group = \(yyy.yyy.yyy.yyy\))

Group list
(interface = \(z\), group = \(xxx.xxx.xxx.xxx\))

The Problem With Pruning

Pruning on broadcast networks

- How can A learn that it must forward still forward group \(G\) packets to B?

prune \((S, yyy.yyy.yyy.yyy)\)

prune \((S, ...)?\)
# RPF With Pruning

## Summary

- Multicast group joins are “local”
  - Leaf routers maintain “positive state” — a list of all groups subscribed to by local hosts
  - Joins do not propagate out into the network
- Prunes propagate back to towards the source
  - Interior routers maintain “negative state” — a list of all prune messages
    (Plus RPF routing state)
- Thus, how does someone who has unsubscribed resubscribe?

## Soft-state timeout

- Router state is “soft” and will periodically timeout
  - This is necessary to accommodate dynamic joins and leaves
Spanning Tree-Based Routing

Core-based trees (CBT)

- Fix a point in the network to be the “center” of a multicast group
- Hosts join the group by sending join messages towards the core
  - Intermediate routers record the link on which these messages arrive & forward towards the core
- The resulting spanning tree is used by all sources
  - Per source state is no longer required in routers

Spanning Tree-Based Routing

Core-based trees (CBT)

- To multicast to the group, messages are sent to the core
- If they encounter a router on the CBT along the way, the multicasting begins
Spanning Tree-Based Routing
Core-based trees (CBT)

◆ Advantages
  » Doesn’t depend on local routing tables
  » Eliminates flooding
  » Efficient, low overhead

◆ Disadvantages
  » No load balancing