CSE/EE 461 Lecture 19
Multicast

Tom Anderson

tom@cs.washington.edu

Peterson, Chapter 4.4

Multicast

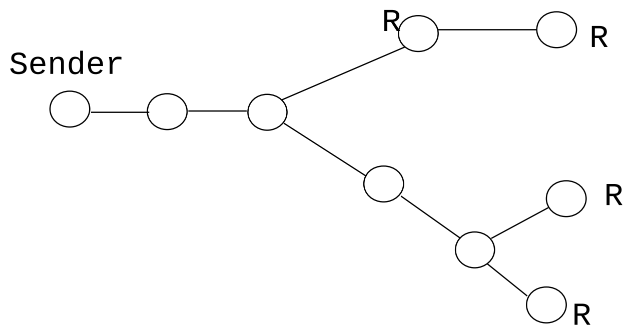
- Challenge: how do we efficiently send messages to a group of machines?
  - Need to revisit all aspects of networking
    - Routing
    - Autonomous systems
    - Address allocation
    - Congestion control
    - Reliable delivery
    - Ordered delivery

Multicast Motivation

- Send data to multiple receivers at once
  - broadcasting, narrowcasting
  - telecollaboration
  - software update
  - group coordination, subcasting
- Send question to unknown receiver
  - resource discovery
  - distributed database
  - anonymous directory services

Multicast Efficiency

- Send data only once down link shared by multiple receivers



Multicast Deployment

- How do we add multicast services to the Internet?
- IP multicast
  - special IP addresses to represent groups of receivers
  - receivers subscribe to specific channels
  - modify routers to support multicast sends
- Overlay network
  - PC routers, forward multicast traffic by tunneling over Internet
  - Works on existing Internet, with no router modifications

IP Multicast Service Model

- Provided by internetwork, with help from LAN
- Best effort delivery (unreliable, unordered, …)
  - Packets addressed to group address (allocated from special range)
- Receivers
  - zero, one or many receivers
  - dynamic -- anyone can join, leave
- Senders
  - Any number of senders -- just send packet to group address

IP Multicast Routing

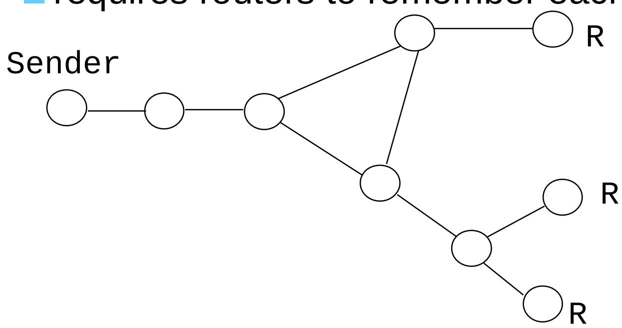
- How do we distribute packets across thousands of LANs?
  - Each router responsible for its attached LAN
- Reduces to:
  - How do we forward packets to all interested routers? (DVMRP, M-OSPF, MBone)
  - How do hosts declare interest to their routers? (IGMP)

Multicast Routing Approaches

- Broadcast data and prune where there are no members
  - based on distance vector routing (DVMRP)
- Broadcast membership and compute forwarding tables
  - based on link state routing (MOSPF)
- Broadcast rendezvous for group addresses
  - independent of unicast routing (PIM)

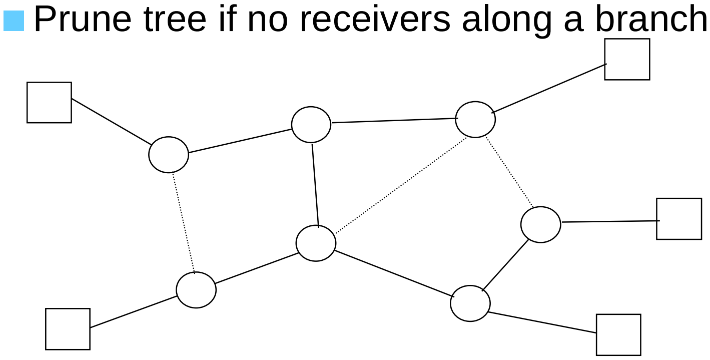
Why not Reliable Flooding?

- If haven't seen a packet before
  - forward it on every link but incoming
  - requires routers to remember each pkt!



Multicast via Spanning Tree

- Send copies along the spanning tree
  - Ensures every host gets a copy
  - Prune tree if no receivers along a branch

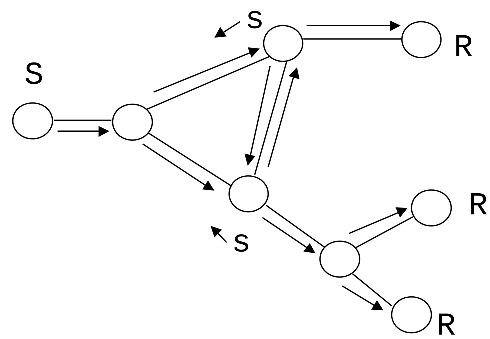


Distance Vector Multicast

- Intuition: unicast routing tables form inverse tree from senders to destination
  - why not use backwards for multicast?
  - Various refinements to eliminate useless transfers
- Implemented in DVMRP (Distance Vector Multicast Routing Protocol)

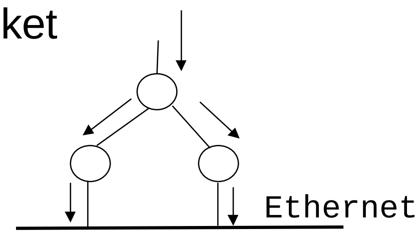
Reverse Path Flooding (RPF)

- Router forwards packet from S iff packet came via shortest path back to S



Redundant Sends

- RPF will forward packet to router, even if it will discard
  - each router gets pkt on all of its input links!
- Each router connected to LAN will broadcast packet

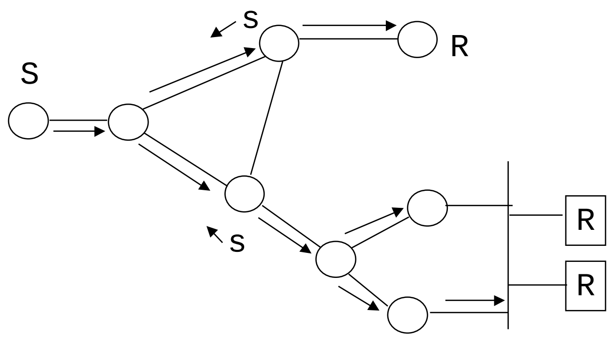


Reverse Path Broadcast (RPB)

- Neighbors exchange distance vector routing tables
  - enables router to know if forwarded multicast packet will be discarded (not on shortest path to S)
- Only send to a neighbor if router is on *neighbor's* shortest path back to source
- Only send on a LAN if router is on shortest path back to source
  - break ties arbitrarily

RPB Example

- Router forwards packet from S iff router is on *neighbor's* shortest path back to S



Truncated RPB

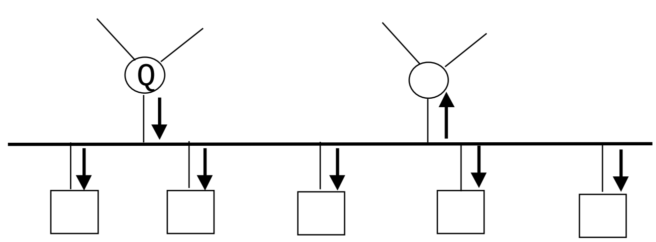
- End hosts tell routers if interested
- Routers forward on LAN iff there are receivers
- Challenges:
  - needs to be robust to router/host failures
  - need to avoid overloading LAN with announcements

Internet Group Management Protocol (IGMP)

- To join, a host multicasts a join request to the LAN, with TTL=1
  - only other hosts, routers on LAN will see request
- Each forwarding router periodically broadcasts a query on the LAN, anyone receiving group?
  - if no one replies (receiver crashed), stop forwarding
- When host gets query, sets random timer
  - When timer expires, multicast join request, TTL=1
  - All other hosts will see the join, stay silent

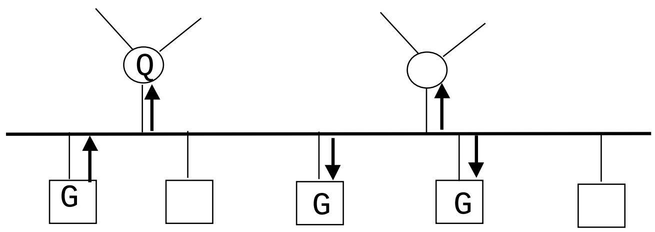
IGMP (step 1)

- Router (on shortest path back to S) periodically broadcasts a query about group (TTL=1)



IGMP (step 2)

- One group member responds to group with TTL=1, everyone else defers
  - All receivers can listen if anyone does



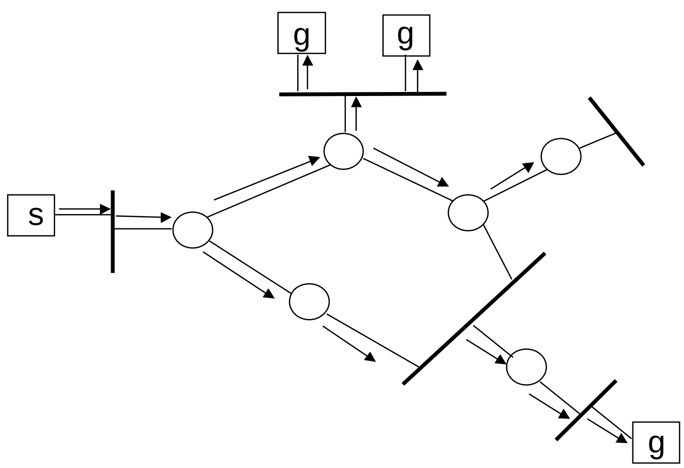
Reverse Path Multicast (RPM)

- Forward packets only to those areas of the network with receivers
- “Broadcast and prune”
  - Use IGMP to tell if LAN if no members
  - If no children are members, propagate prune to parent in tree
- Assume membership and prune if wrong vs. assume non-membership and explicit join

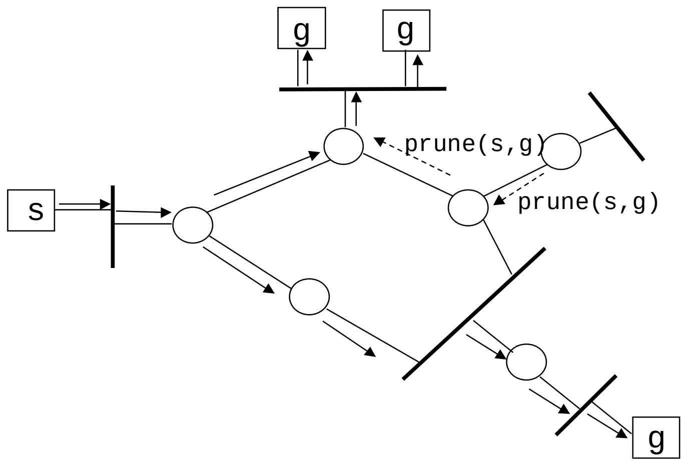
How does a receiver join?

- What if multicast tree has been pruned, and a receiver joins the multicast in progress?
- Graft (prune cancellation)
  - Routers remember where they sent prunes
    - – where multicast traffic came from
  - If child joins, send graft in same direction(s)
- Requires ARQ
  - what if graft is dropped?
  - what if prune is dropped?

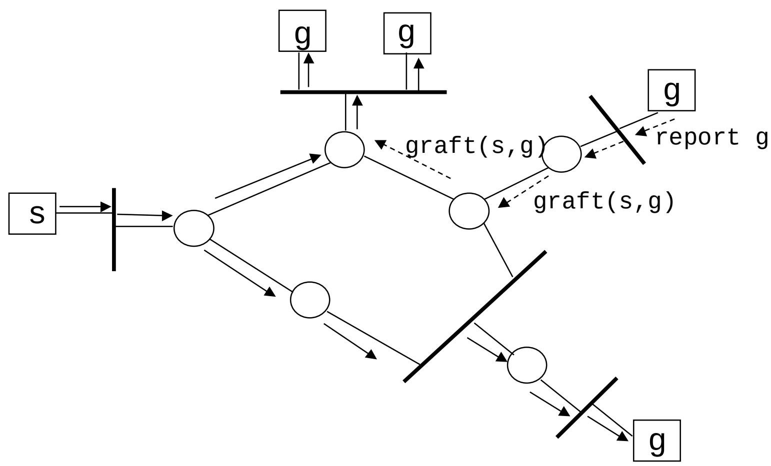
Phase 1: Truncated Broadcast



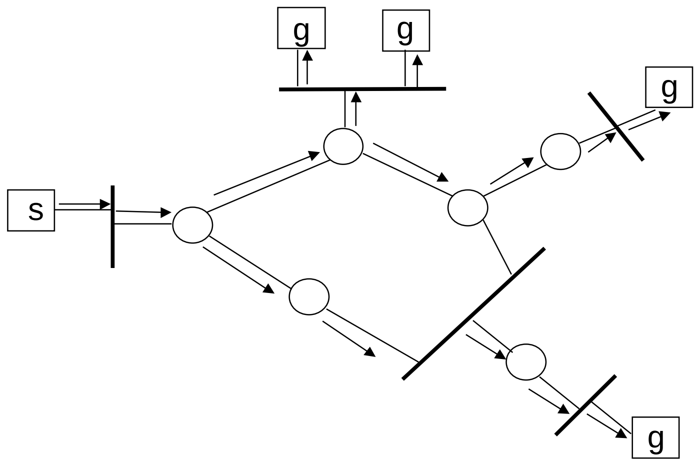
Phase 2: Pruning



Phase 3: Grafting



Phase 4: Steady State



RPM Mechanics

- Data-driven: prune only when data packet arrives
  - Why?
- Periodically time out prunes
  - Why?
- Are loops possible in the multicast routing tree?

Hierarchical Broadcast and Prune

- Reverse Path Flooding
  - Discard incoming packet if not from reverse path
  - Multicast incoming packet to all borders
- Reverse Path Multicast
  - For each neighbor AS, compute if we're on its reverse path to source
  - Multicast incoming packets to all border routers for those AS's
  - Propagate prunes across the AS back towards the source