**Motivation:** ISPs charge by bandwidth. They make more money without multicast.

**Overview:**
- Multicast = Send to a group of hosts
- Key ideas:
  - Replace multiple unicasts by a tree (good algorithms for building tree)

**Implementation Issues:**
- 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?
Implementation Issues …

- How are multicast packets addressed?
- How is join implemented?
- How is send implemented?
  - How does multicast traffic get routed?
    - This is easy at the link layer and hardest at the network layer
- How much state is kept and who keeps it?

Internet Group Management Protocol IGMP

- Operates between Router and local Hosts, typically attached via a LAN (e.g., Ethernet)
  - Query response architecture
- Router periodically queries the local Hosts for group membership information
  - Can be specific or general
- Hosts receiving query set a random timer before responding
- First host to respond sends membership reports
- All the other hosts observe the query and suppress their own reports.

- To Join: send a group send an unsolicited Join
  - Start a group by joining it
  - To leave: don’t do anything
  - Soft state

Routing: Approaches

- Kinds of Trees
  - Source Specific Trees
  - Shared Tree
- Tree Computation Methods
  - Intradomain
  - Interdomain: BGMP
    - This is still evolving...

Source Specific Trees

Each source is the route of its own tree.

Can pick “good” trees but lots of state at the routers!
One tree used by all

Can't pick "good" trees but minimal state at the routers

A tree which connects all the group nodes is a Steiner Tree
Finding the min cost Steiner Tree is NP hard

A tree that connects all the nodes in the graph is a spanning tree
Finding a minimum spanning tree is much easier

A tree that connects all the nodes in the graph is a spanning tree
Finding a minimum spanning tree is much easier
Tree Computation

- A tree that connects all of the nodes in the graph is a spanning tree
- Finding a minimum spanning tree is much easier
- Prune back to get a multicast tree

Intradomain Multicast Routing

- Build on unicast Link State: MOSPF
- Build on unicast Distance Vector: DVMRP
- Protocol Independent: PIM
- Problems

Link State: MOSPF

- Use in conjunction with a link state protocol for unicast
- Enhance the LSP updates with group membership
- Compute best tree from source
- Flood Membership in link state advertisements
- Dynamics are a problem

Distance Vector: DVMRP

- An elegant extension to DV routing
- Use shortest path DV routes to determine if link is on the source-rooted spanning tree (tree of shortest paths to source, break ties with lowest router ID)
- Prune back if no group members down the subtree
- Improve scalability: Multicast to a “shared tree” (core) and from there to rest of hosts

PIM

- Popular intradomain method
  - UUNET streaming using this
  - Recognizes that most groups are very sparse
  - Why have all of the routers participate in keeping state?
- Two modes
  - Dense mode: flood and prune
  - Sparse mode: Core-based shared tree approach
Problems with Network Layer Multicast
- 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
  - Difficult to debug problems given the service model

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
- 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

Scalable Reliable Multicast (SRM)
- Randomize NACKs (request repairs)
- All traffic including request repairs and repairs are multicast
- A repair can be sent by any node that heard the request
- A node suppresses its request repair if another node has just sent a request repair for the same data item
- A node suppresses a repair if another node has just sent the repair

Application Layer Multicast
- Provide multicast functionality above the IP unicast
- Gateway nodes could be the hosts or multicast gateways in the network
- Advantages
  - No multicast dial-tone needed
  - Performance can be optimized to application
  - Loss, priorities etc.
  - More control over the topology of the tree
  - Easier to monitor and control groups
- Disadvantages
  - Scale
  - Performance if just implemented on the hosts (not gateways)
Summary

- Large amount of work on multicast routing
- Major problems
  - Preventing flooding
  - Minimizing state in routers
  - 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 are promising