Routing to the many

Multicast routing

Multicasting
- Unicast protocols involve communications between one sender and one receiver.
- Multicast protocols involve communications between one sender and many receivers.

N-way-unicast
- Given \(N\) destination nodes, the sender makes \(N\) copies of the data packet, addresses each copy and transmits each to its destination.
- Simple; requires no network layer support.

Simple, but somewhat inefficient
- If the source node is connected to the network via a single link, that router works overtime. Similarly, all down the line.
- More efficient if the network nodes made the copies.
Worse and worser
- How does the broadcaster know the broadcastees and their addresses?
- Link-state routing protocols use broadcast to disseminate the link-state information used to update unicast routes*.

*It would be unwise (at best) to rely on the unicast routing infrastructure achieve broadcast.

Uncontrolled flooding
- When a node receives a broadcast packet, it duplicates the packet and forward it to all of its neighbors (except the one it came from).
- This suffers from a fatal flaw. Name that flaw.

The perfect storm
- But wait, it gets worse, when a node is connected to two or more nodes, it will forward multiple copies of the broadcast packets, each of which will create multiple copies and so forth.
- The result: broadcast storm.

Sequence-number-controlled flooding
- A source nodes put its address and a broadcast sequence number into the packet.
- Each node maintains a list of source address and sequence number of each packet received, duplicated, and forwarded.
- When a node receives a broadcast packet, it first checks whether the packet is in this list. If so, the packet is dropped.
Reverse path forwarding

When a router receives a broadcast packet with a given source address, it transmits the packet on all of its outgoing links (except the one on which it was received)...

...but only if the packet arrived on the link that is on its own shortest unicast path back to the source.

Sequence-number controlled flooding and RPF avoid broadcast storms...

...but, they do not completely avoid transmission of redundant packets.

Spanning-tree broadcast

If broadcast packets were forwarded only along links within a spanning tree, each and every network node would receive exactly one copy of the broadcast packet.

Minimum cost spanning-tree broadcast

If each link has an associated cost and the cost of a tree is the sum of the link cost, then a minimum cost spanning tree would be a good choice through which to broadcast.
So how does one construct a spanning tree?

Center-based spanning tree construction

A tree-join message is forwarded using unicast routing toward the center until it either arrives at a node that already belongs to the spanning tree ...

Nodes then unicast tree-join messages addressed to the center node.

Center-based spanning tree construction

Broadcast algorithms in practice

- Gnutella* used sequence-number-controlled flooding with a 16-bit-identifiers and payload descriptors to detect previously received broadcast queries.
- Gnutella also used a time-to-live field which is decremented at each hop.

*One of the first decentralized peer-to-peer networks.
Multicast

- Broadcast service delivers packets to all nodes in the network.
- Multicast service delivers to only a subset of the network.*
- Two problems: How to identify the receiver and how to address the packets?

*Applications use this to transfer of software upgrades from software developers to users, streaming video, shared data applications, data feeds (stock quotes), Web cache updating, interactive gaming.

The multicast group?

1. How is the group address chosen?
2. How are new hosts added to the group (either as senders or receivers)?
3. Can anyone join a group or is it restricted and, if so, by whom?
4. Do group members know the identities of the other group members as part of the network-layer protocol?
5. How do the network routers interoperate with each other to deliver a multicast datagram to all group members?

Address indirection

- A single identifier is used for a group of receivers, a.k.a. a multicast group.
- Packets are addressed using this single address.

Internet group management protocol

- IGMP operates between a host and its directly attached (first-hop) router.
- Hosts use it to inform their first-hop routers that an application wants to join a multicast group.
- Given that IGMP is limited to communications between a host and its attached router, another protocol is needed to coordinate multicast routers to deliver the message.
IGMP messages
- Like ICMP, IGMP messages are carried (encapsulated within an IP datagram, with an IP protocol number of 2.
- Unlike ICMP, IGMP has only three message type: membership_query; membership_report; and leave_request.

Membership queries and reports
- A general membership_query message is broadcast by a router to all hosts on an attached interface.
- Hosts respond to membership_query with an IGMP membership_report message.
- A membership_report may also be initiated by the host in which case it is a request to join.

Joining the club
- Joining a multicast group is thus receiver-driven.
- A sender need not be concerned with explicitly adding receivers to the multicast group.
- But neither can the sender control who joins the group and therefore who receives datagrams sent to that group.
- Similarly, there is no control over who sends to the multicast group.

There are issues ...
- Datagrams sent by different host can be arbitrarily interleaved at the various receivers.
- A malicious sender can inject datagrams into the multicast flow.
Even for the well intentioned ...

- Since there is no network-layer coordination of the use of multicast addresses, it is possible that two different multicast groups will choose to use the same multicast address.
- From a multicast application viewpoint, this will result in interleaved extraneous multicast traffic.

Serious problems, but all is not lost

- Although the network layer does not provide for filtering, ordering, or privacy of multicast datagrams, these mechanisms can all be implemented at the application layer.
- There is also ongoing work aimed at adding some of this functionality into the network layer.

Soft State

- The leave message is optional, so how does a router know when a host leaves the multicast group?
- The answer is that the router infers that a host has left when it no longer responds to a membership_query.

Multicast routing

- Multicast hosts and their attached routers still have to communicate over an indifferent network.
- Multicast routing algorithms seek a tree of links that connect all routers that have attached hosts belonging to a particular group.
- We discuss two approaches to finding such a tree.
Group-shared tree

- All members of the group could use a single group-shared tree to distribute messages.
- Ideally, this tree would have minimal cost among all such trees. Finding such a tree is known as the Steiner tree problem, which is known to be NP-complete.
- The good news is that there are good approximation algorithms that work well in practice and are provably within optimal value.

The bad news

- Even though good heuristics exist for the Steiner tree problem they are never used. Why?
- The information needed for them is about all links in the network.
- Another reason is that in order for a minimum-cast tree to be maintained, the algorithm needs to be rerun whenever link costs change.
- Performance is just one of many concerns.

Center-based approach (again)

- A center node, again known as a rendezvous point, is chosen. Routers with attached hosts then unicast join messages to the center node.
- The path that the join message followed defines a branch of the routing tree.

Least unicast-cost path tree

- An alternative approach computes for each source the least unicast path to each destination.
- The union of these paths at each node is that node’s desired tree.
- Isn’t this the same as minimum-cost multicast tree?

*Shades of Dijkstra?
Reverse path forwarding, multicast case

A router transmits the packet on all of its outgoing links (except the one received) only if the packet arrived on the link that is on its own shortest path back to the sender.

Reverse path forwarding does not require that a router know the complete shortest path from itself to the source; it need only know the network hop on its unicast shortest path to the sender.*

Distance vector multicast routing protocol

- DVMRP, the first multicast routing algorithm, implements source-based trees with reverse path forwarding and pruning.
- Unfortunately, only a small fraction of Internet routers are multicast-capable. What does a lone A do in a sea of plain routers?

Pruning

- Router D forwards to G even though G has no attached hosts in the group.
- A multicast router that receives group messages that none of its attached hosts (or downstream routers) are interested in sends a prune message the upstream router.

Answer: Use multicast tunneling to construct a logical multicast network

- To wit, A takes the multicast datagram and encapsulates it inside a standard unicast datagram, addresses and forwards to B.
- B extracts the multicast dataframe and forwards it on to its attached hosts, and/or attached neighboring routers that are multicast-capable, or each via another tunnel to other logical multicast neighbors.
PIM*, an alternative to DVMRP

- **Dense mode:**
  - Used when group members are tightly packed;
  - flood and prune reverse path forwarding technique similar in spirit to DVMRP.

- **Spare mode:**
  - Used when members are far between;
  - is a center-based approach similar to the core-based tree (CBT) multicast protocol.

*Protocol independent multicast, perhaps the most widely used multicast protocol.