Multicasting in Internets


Definition: Transmission of a packet to a group of hosts.
How about the following approaches?

1. Unicasting to each host in group.

2. Broadcasting to all, with nodes responsible for selectively accepting received packets.
Desirable Properties of Multicasting Protocol

1. Reduce transmission overhead on sender.

2. Reduce network overhead.

3. Sender need not know membership of group.

4. Sender need not be a member of the group.

5. Hosts may join and leave groups at will.

6. Sender can limit scope of multicast packets.
Types of Multicast Groups

**Pervasive groups:** broadcast facility may be useful.

**Sparse groups:** using selective multicast trees may be advisable.

**Local groups:** using selective multicast, or scope control with broadcast.

What is the cost of:

1. Sending packet over all links?
2. Maintaining multicast tree?
3. Host processing on packet arrival?
Single-Spanning-Tree Multicast Routing

- Group members periodically issue *membership-report* packets:
  - Source address: group address.
  - Destination address: all bridges group.

- Link layer bridges learn which incident branches lead to members of a multicast group.

- Bridges maintain following entries: \((address, (outgoing branch, age), (outgoing branch, age) \ldots)\)

- On arrival of packet with multicast destination address:
  - Forward copy of packet along every outgoing branch in the table entry for that multicast address (excluding arrival branch).
  - If no entry for the multicast packet: discard packet.
Reducing Membership-Report Overhead

- Group members issue membership-report with group address as destination.

- Bridges directly attached to reporting group member change destination address to all bridges.

- Other group members on the same LAN hear the report and suppress their own report for period roughly equal to $T_{report}$ interval.

- Result: only one report per group on a LAN every $T_{report}$ period.
Distance-Vector Multicast Routing

Assumptions

1. Symmetric links.
2. Routers know shortest path to each node S, using distance vector protocols.
3. Idea of Reverse path forwarding applicable.

Event: Packet arrives at router R with source address S, destination group G.

- If incoming link is on the shortest path to S: forward packet along all other incident links.
- Otherwise: discard packet.
Distance-Vector Multicast Routing (cont’d)

Consequence: Ensures shortest-path broadcast along the network.

- Hosts discard packets destined for groups to which they do not belong.

Drawback: Multicast packet may be transmitted across a link as many times as the number of router that share that link.

- Multiple routers receive packet along another link that is on shortest path to S.

- Each such router decides to forward packet on this link.
Avoidance of Multiple Transmissions

- Single parent router for each link, relative to source S:
  - S is the parent for the link(s) to which it is attached.
  - On all other links, router with minimum distance to S is parent (w.r.t. S).
  - Router address used to break ties.

- Router forwards a packet only along its child links for S.
Evolution of Distance-Vector-Based Multicast

Reverse Path Flooding (RPF): If a broadcast packet with source address $S$ arrives from the *next-hop-link* for $S$, forward a copy of the packet on all other links.

Reverse Path Broadcasting (RPB): . . . on all child links for $S$.

Truncated Reverse Path Broadcasting (TRPB): . . . except leaf links that have no members of group $G$.

Leaf link: a child link that no other router uses to reach $S$.

Absence of membership-report for $G$ implicitly indicates no members of group $G$ on a link.
Reverse Path Multicasting: Motivation

- Broadcast trees need to be pruned, especially for sparse groups.

- Expensive to prune every broadcast tree (one tree for every source, multicast group pair).

- Only prune the multicast trees actually in use: on-demand pruning.
Reverse Path Multicasting: Protocol

1. Initially deliver multicast packets from source S along shortest-path broadcast tree, as per TRPB.

2. Router, all of whose child links are leaves with no members for the (S, G) pair, sends corresponding non-membership report (NMR) to its next-hop router (predecessor) for S.

3. Router that receives NMRs from all routers along its child links, and has no group members on its child links, sends NMR to its predecessor.

4. Routers use received NMRs to block forwarding of multicast packets.

5. Age associated with NMRs.
Reverse Path Multicasting: Protocol (cont’d)

6. Path truncated by NMR rejoins multicast tree after $T_{maxage}$.

7. If there is still traffic from same source to same group, next packet may trigger new NMR if there is still no member on that path.

8. Routers send cancel-NMR messages if hosts join group G downstream from them, or when a new child link or child router (w.r.t. G) joins.
Link-State Multicast Routing

- Link state includes sets of groups that have members on that link.

- Routers attached to the link maintain link state information.

- LAN Designated Router maintains and floods group membership information about the link.

- Routers can compute shortest-path multicast tree from any source to any group using Dijkstra’s shortest path algorithm.

- Multicast trees computed only on demand and information cached.
Link-State Multicast: Operation

Let router receive multicast packet from source S to group G:

- If no cache entry for G, compute shortest-path multicast tree for the (S,G) pair.

- For each incident link leading to members of group G:
  - Determine min-hops, the minimum hops to reach a member of G along that link.
  - If min-hops is less than TTL of packet, forward packet along the link.
Link-State Multicast: Caching

- Old cache records replaced by new records on a least-recently-used basis.

- All cache records discarded on topology change.

- All cache records for a group discarded when link membership of that group changes.

First packet multicast by a source, S, to a group, G, experiences high latency:

- Routers at each hop must compute shortest-path multicast tree for the (S,G) pair.
Core Based Trees (CBT) for Multicasting

Following discussion is based on the paper by Ballardie, Francis and Crowcroft, SIGCOMM’93, pp 85-95.

- Routers need to maintain multicast tree information for every (active source, multicast group) pair: expensive.

- Routers not participating in a multicast tree have to do extra work of generating tree prune/graft messages.

- Distance-vector and Link-state based protocols have scalability problems (?)
Core Based Tree for Multicasting: Protocol

- One router designated as the core for a group.
- Shortest-path multicast tree rooted at the core.
- Router wishing to join the tree unicasts a join packet towards core.
- Branch grafted when the join packet reaches a router already part of the tree.
- Sender unicasts packet towards core.
- Packet sent over the tree once it reaches a router that is part of the CBT.

CBT does not guarantee optimal paths between group members, and from sources.
Protocol Independent Multicast (PIM)

Following discussion is based on the paper by Deering, Estrin, Farinacci, Jacobson, Liu and Wei, SIGCOMM’94, pp 126-135.

- Previous solutions suffer from lack of scalability.
  - Link-state multicast protocols impose high membership broadcast overheads.
  - Distance-vector multicast protocols require non-participating routers to generate and process prune messages: especially undesirable for sparse groups.
  - CBT protocol results in traffic concentration: a subset of links carry most of the traffic.
  - Higher latency for CBT protocol: undesirable for interactive applications.
Shared versus Source-Based Trees

- Shared trees (as in CBT) may be suitable for large number of low data rate sources.

- Source-based trees better suited for high data rate sources.

- PIM is designed to:
  - Avoid broadcast overheads for sparse groups.
  - Support “good-quality” distribution trees for heterogeneous applications.
PIM: Requirements

Efficient sparse-group support

High-quality data distribution: in accordance with different requirements of applications.

Routing protocol independent: multicast protocol uses unicast routing tables, independent of how those tables are computed.

Interoperability: with reverse-path forwarding and link-state multicast protocols.

Robustness: use soft state refresh, and avoid single point of failure.
PIM: Protocol

- PIM uses per-group *rendezvous point(s)* (RPs) for sources and receivers to “meet.”

- When nodes wish to participate in group communication:
  - Receiver’s first-hop PIM speaking router sends a PIM *join* message toward one of the RPs advertised for the group.
  - Similarly, source’s router sends PIM *register* message, piggybacked on first data packet to RP(s) for the group.
  - Intermediate routers use these message to set up the paths to the RP(s).
PIM: Establishing Source-Based Trees

- Receiver’s PIM speaking router sends a join message to source.

- Prune message sent by router to RP after data packets received on new path if the new and old paths diverse at this router.

- Data packets continue to travel from source to RP for delivery to other receivers.

- Receiver continues to receive data packets, generated by other sources, from RP.