Introduction to IP Multicast

Colin Perkins <c.perkins@cs.ucl.ac.uk>

Department of Computer Science

University College London

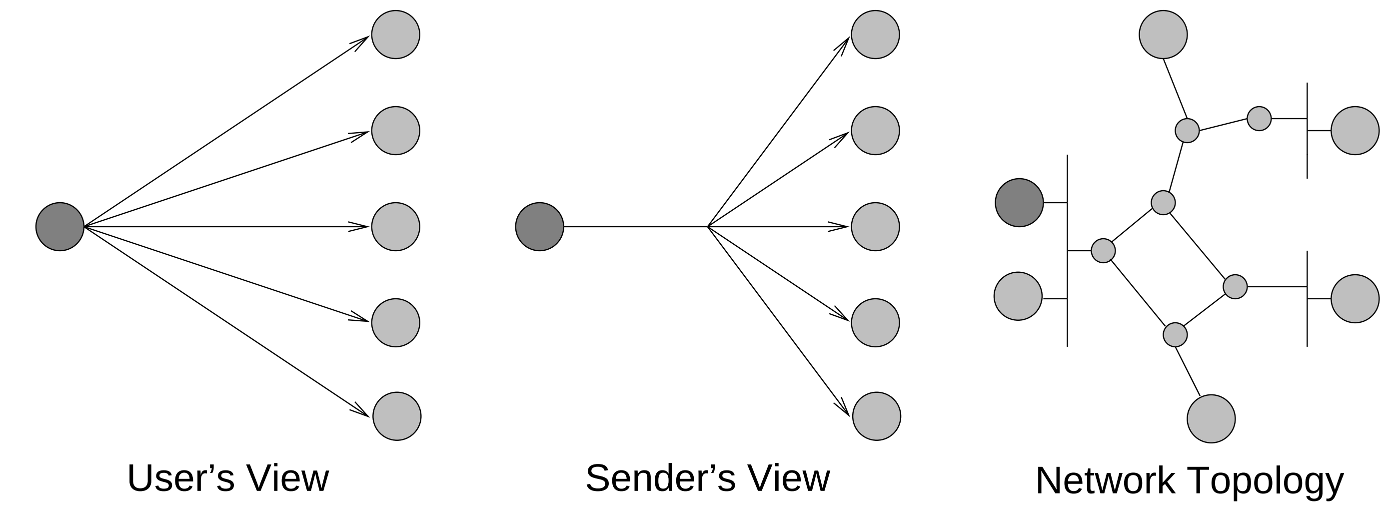
Gower Street

London WC1E 6BT

Contents

Requirements for IP Multicast

- Support many-to-many communication
- Scale to large groups
- Be robust to network failures



IP Multicast Service Model

- Dynamically constructs efficient distribution trees from senders to receivers
- Service model:
  - Receivers announce interest
  - Senders just send
  - Routers conspire to deliver sender's data to all interested receivers
- Precise membership of a multicast group is not known
- No explicit group membership control

IP Multicast Addresses

- Traditionally, IP addresses were divided into three classes:
  - Class A for organizations with millions of hosts
  - Class B for organizations with thousands of hosts
  - Class C for organizations with hundreds of hosts
- Class D addresses have been reserved for IP multicast:
  - Class D addresses are in the range 224.0.0.0 to 239.255.255.255
  - Top 4 bits of a multicast address are 1110
  - The choice of multicast or unicast is based on the address only

Multicast Routing Models

DVMRP Distance Vector Multicast Routing Protocol

PIM Protocol Independent Multicast

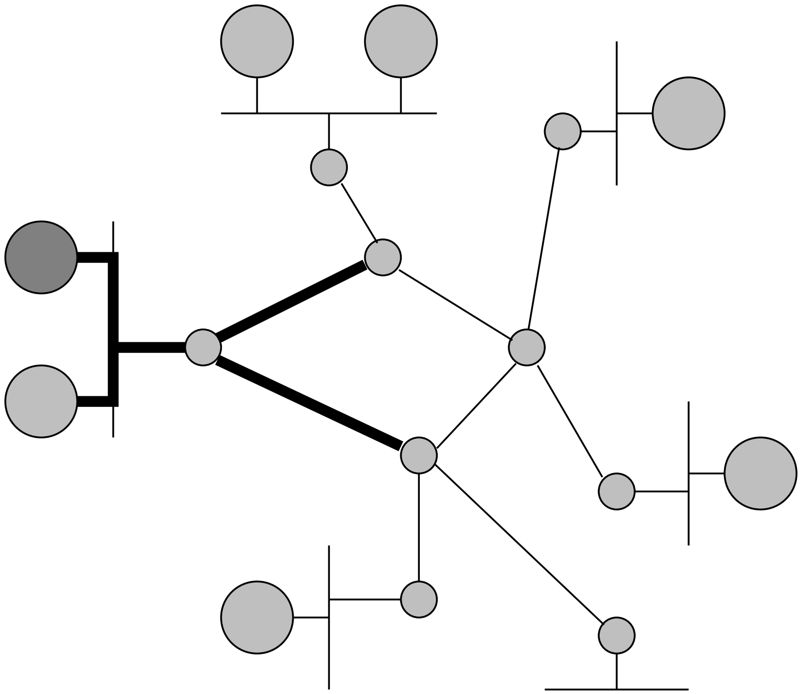
MOSPF Multicast OPSF

CBT Core Based Trees

The current Mbone uses mainly DVMRP with some PIM deployment.

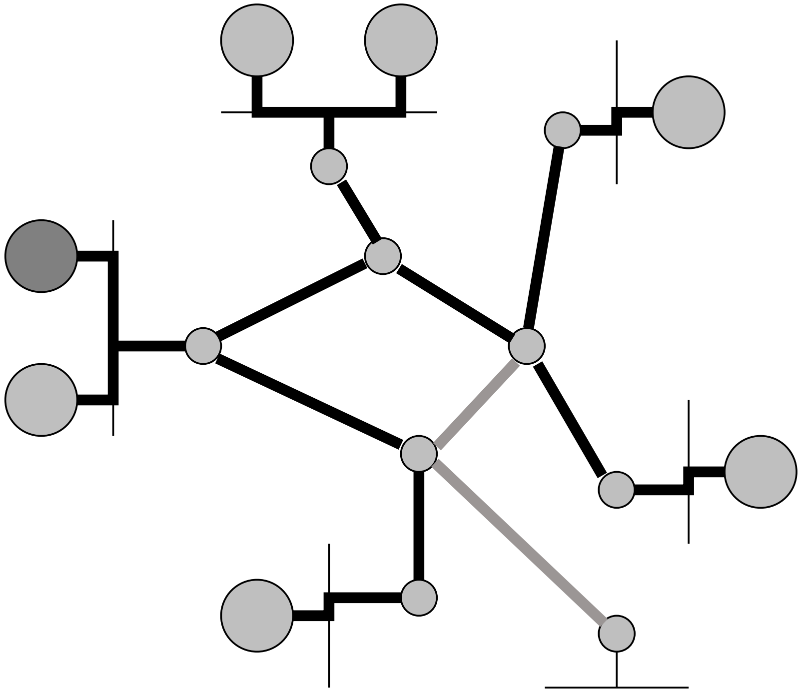
DVMRP: Basic Model

Basic model is "flood and prune"



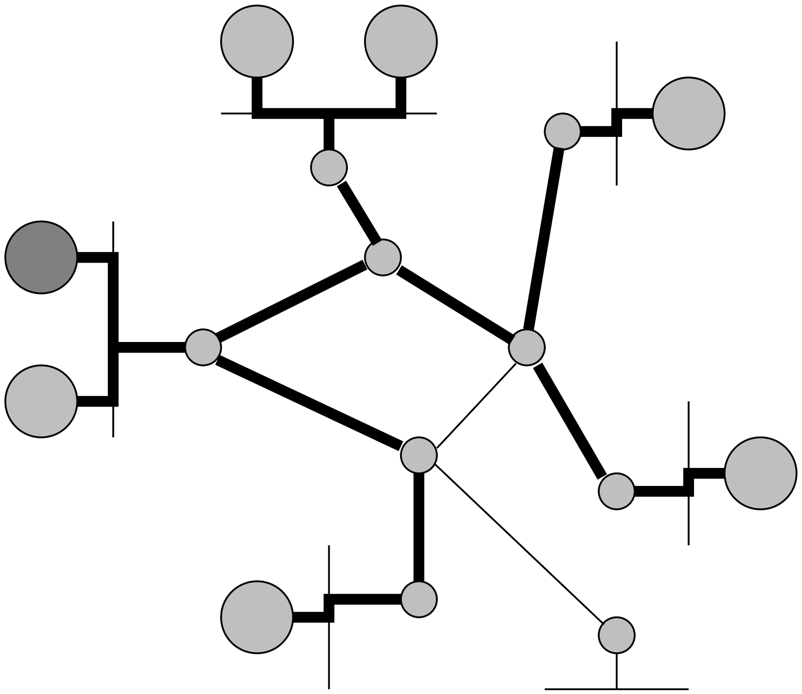
DVMRP: Tree Pruning

Branches not on the shortest path tree from receiver to sender are pruned off



DVMRP: Shortest Path Tree

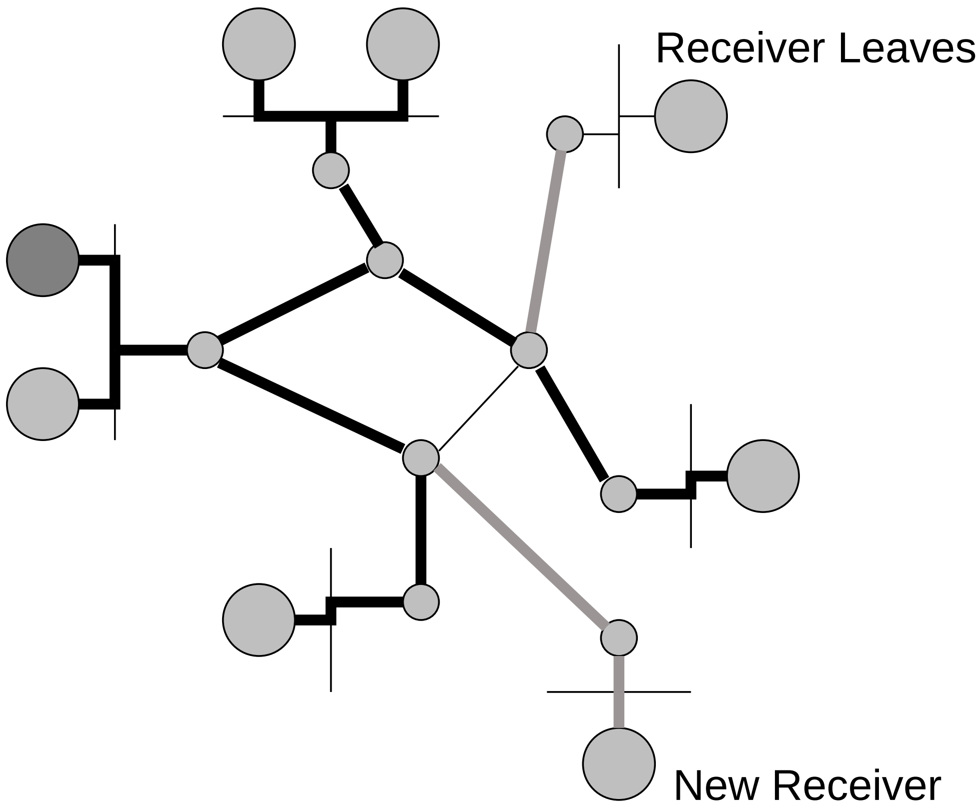
Resulting in a shortest reverse path distribution tree rooted at the sender.



Trees are per {sender,group}, and are generated by data flow.

DVMRP: Membership Changes

Changes in group membership are handled efficiently and locally via grafting and pruning:



DVMRP: Comments

- DVMRP assumes ‘dense-mode’ operation. The default is to transmit data, until an explicit prune is received. Wastes bandwidth for sparsely connected groups.

- DVMRP is not hierarchical. Routing tables are large, and routing updates consume considerable bandwidth. Will not scale to larger networks.

- DVMRP uses it’s own routing tables. The underlying unicast routing information is not used, requiring a duplication of routing data in many cases.

PIM: Protocol Independent Multicast

There are two forms of PIM:

Dense mode for situations where there are many receivers, and bandwidth is plentiful.

- Similar to DVMRP
- Builds on the underlying unicast routing protocol

Sparse mode for situations where receivers are widely distributed across the network.

- Only forwards data when it is explicitly requested, unlike DVMRP, which forwards until explicit prunes are received.

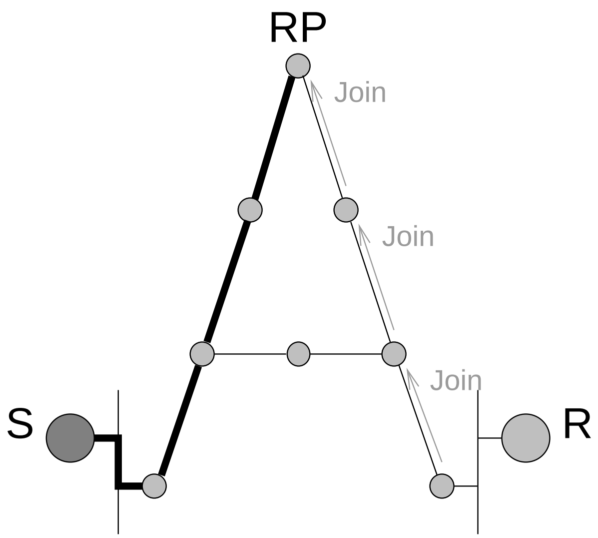
PIM-SM: Basic Model

- Instead of flooding and pruning to directly build a shortest path tree, PIM-SM initially builds shared tree, which is not shortest path.
- The shared tree is initially built by sending join messages towards a well known rendezvous point: traffic is *requested*, rather than being sent by default.
- Once data is flowing, the shared tree can be converted to a shortest path tree if required.

Details included for completeness, else skip to slide 18...

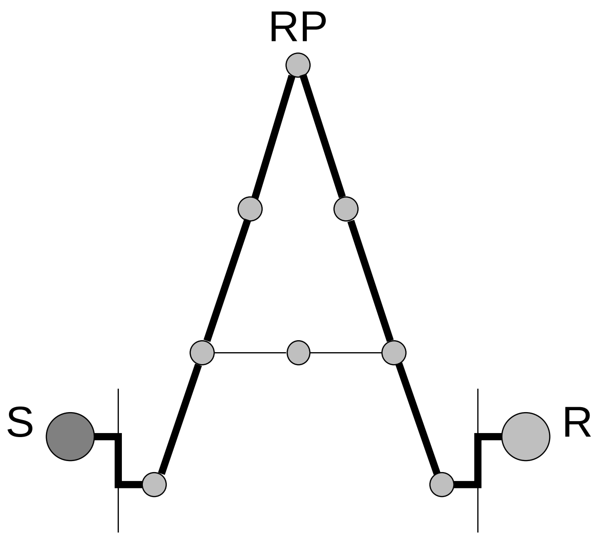
PIM-SM: Build Shared Tree

- Source S sends traffic to group G
- S's local router forwards group G traffic to the rendezvous point, RP, for the group
- R's local router sends a join message for G to the RP



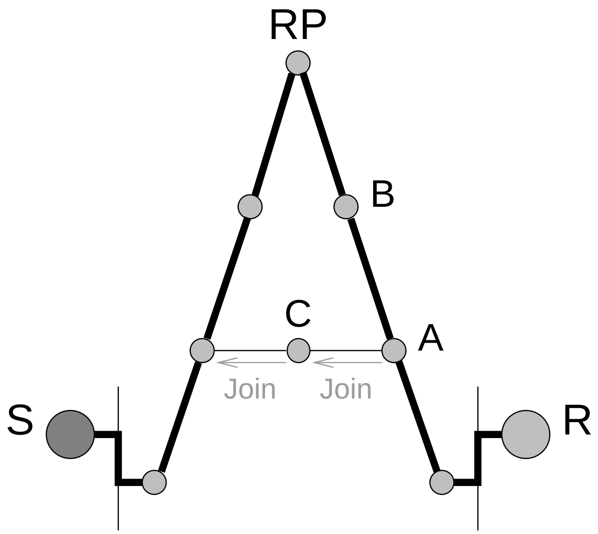
PIM-SM: Shared Tree

- Initially, data flows over the shared RP tree



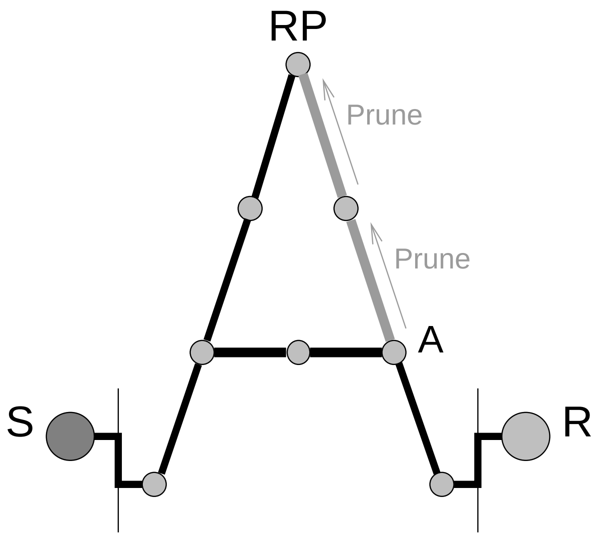
PIM-SM: Transfer to Shortest Path Tree

- Router A sees traffic from S arriving over its B interface, when the unicast route to S is over the C interface.
- If traffic warrants it, A sends a join towards S.



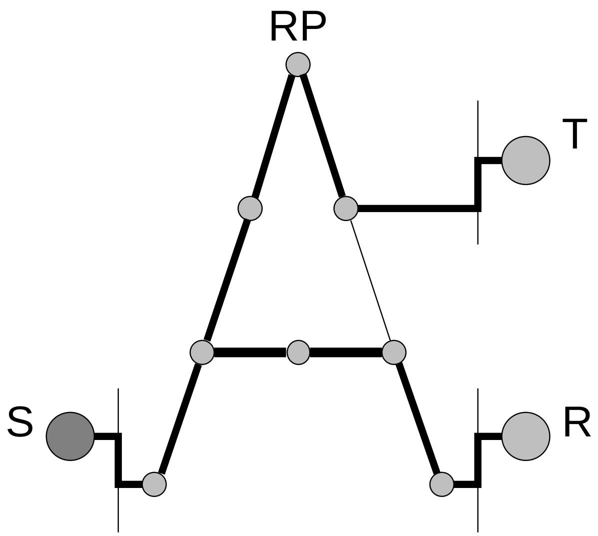
PIM-SM: Transfer to Shortest Path Tree

- When the join is successful, group G traffic from S starts to flow along the shortest path tree. A now receives the traffic twice.
- A then sends a prune back towards the RP



PIM-SM: Membership Changes

- Receiver R receives traffic on the shortest path tree.
- Receiver T joins, and receives traffic through the RP tree. It too can switch to the shortest path tree if required.



PIM-SM: Comments

- Location of the RP is a difficult problem, subject to active research.
- Backup RPs are required, and selection of a backup must be consistent across the entire network.
- Decision of when to switch to the shortest path tree from the RP tree is difficult.
- Routers must maintain significant state to describe sources, groups and RPs. By contrast DVMRP is stateless, driven only by data flow.

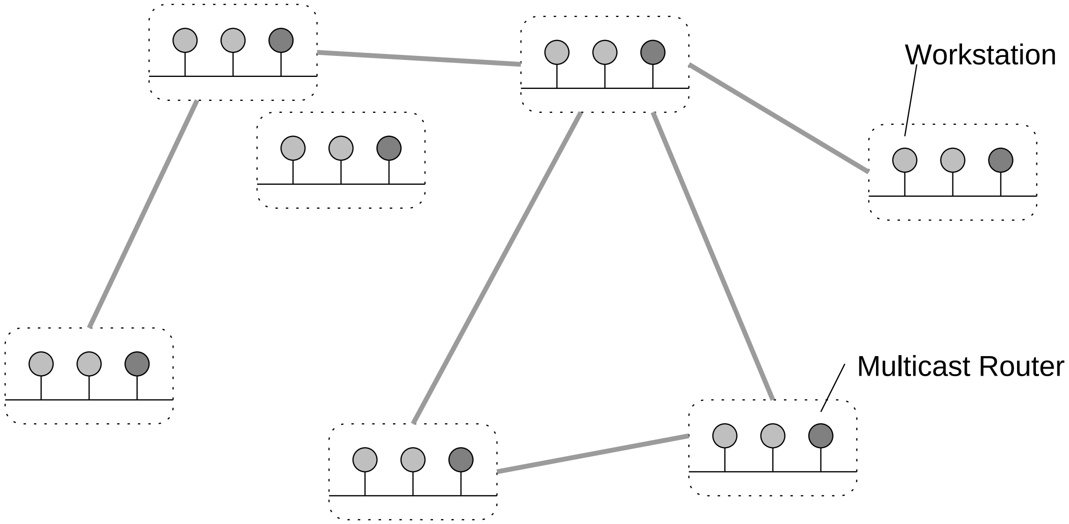
Multicast Routing

- Current multicast routing protocols are not hierarchical. Bandwidth and processing requirements of the routing protocols will grow to be excessive as the Mbone expands.
- The current Mbone is mainly DVMRP, and the scaling problems of this are becoming apparent.
- PIM is better in some ways, but complex and less widely used.

At present, we will be using DVMRP, as implemented in the `mrouted` package.

Islands and Tunnels: The Mbone

Much of the current Internet is *not* currently multicast capable. The *Mbone* consists of a number of islands of multicast capable hosts, connected by tunnels which encapsulate the multicast data for transmission through the rest of the network.



Islands and Tunnels: Tunnel Configuration

- Tunneling multicast packets through unicast links is achieved by encapsulation inside unicast packets.
- Tunneling is a function of the multicast routers.
- Tunnels are configured manually, and have a specific weight, allowing for *TTL scoping*.
- Tunneling is a stop-gap measure, and will not be needed when the entire networks supports multicast.

Islands and Tunnels: TTL Scoping

- All IP packets have a TTL field, which starts at some non-zero value, and is decreased by one for each link crossed.
- When the TTL reaches zero, the packet is discarded.
- Multicast tunnels have a threshold assigned to them.
- If a packet's TTL is less than the threshold, it is not forwarded through a tunnel.
- This lets us crudely scope packets, by adjusting the initial TTL and tunnel weights such that packets are limited to a specific region.

Islands and Tunnels: Administrative Scoping

- In addition to TTL scoping, current multicast routers can be configured so that interfaces/tunnels are administrative boundaries.
- This is done by specifying ranges of multicast addresses which the router will not forward to that interface.
- Addresses in the range 239.0.0.0 to 239.255.255.255 are specified for administrative scoping.

Islands and Tunnels: Summary

- The Mbone is a temporary overlay on the Internet, until all routers support IP multicast.
- Tunnel configuration between multicast capable islands is done manually.
- TTL and administrative scoping allow us to limit the range of a multicast transmission.
- This is not real security though: anyone within the scope can still receive the data.

Security Considerations

- Multicast routing provides a simple, open, service model; with no explicit membership control.
- This means that *anyone* within the scope of a multicast can receive that traffic, unless that traffic is encrypted.
- *Most* multicast conferencing tools support encryption...

Reliability Considerations

- IP multicast is an *unreliable* datagram service.
- Datagrams may be lost, delayed or reordered by the network.
- There are no reliability guarantees.
- Provided sufficient bandwidth is available, problems will be rare.
- As bandwidth constraints are applied, problems will increase in proportion.
- Multicast communication exhibits graceful degradation in quality of service.

Summary

- What is IP Multicast/Mbone
- Multicast routing
- Security: Application level encryption
- Reliability