15-441 Computer Networking

Lecture 11 – Multicast

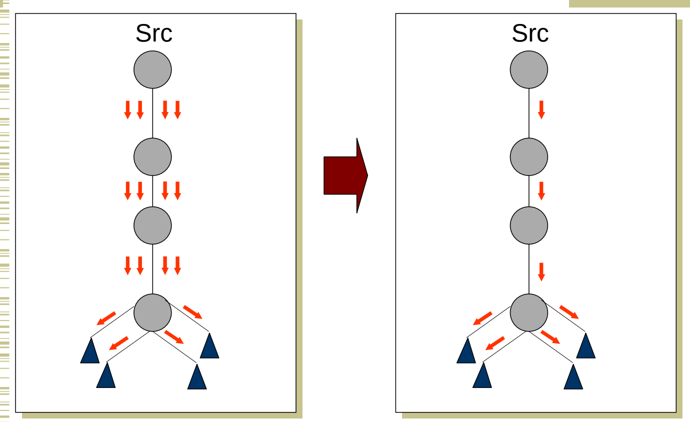
Multicast Routing

- Unicast: one source to one destination
- Multicast: one source to many destinations
- Main goal: efficient data distribution

Overview

- IP Multicast Service Basics
- Host/Router Interaction
- MOSPF/DVMRP
- Overlay Multicast

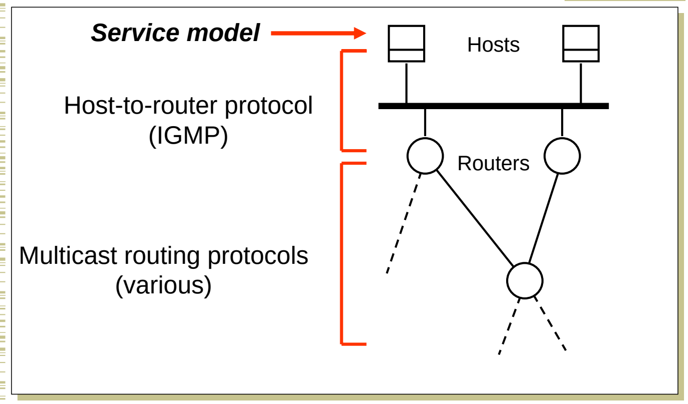
Multicast – Efficient Data Distribution



Example Applications

- Broadcast audio/video
- Push-based systems
- Software distribution
- Web-cache updates
- Teleconferencing (audio, video, shared whiteboard, text editor)
- Multi-player games
- Server/service location
- Other distributed applications

IP Multicast Architecture



Logical Naming

- Single name/address maps to logically related set of destinations
  - Destination set = multicast group
- Key challenge: scalability
  - Single name/address independent of group growth or changes

Multicast Router Responsibilities

- Learn of the existence of multicast groups (through advertisement)
- Identify links with group members
- Establish state to route packets
  - Replicate packets on appropriate interfaces
  - Routing entry:

Src, incoming interface	List of outgoing interfaces

IP Multicast Service Model (rfc1112)

- Each group identified by a single IP address
- Groups may be of any size
- Members of groups may be located anywhere in the Internet
- Members of groups can join and leave at will
- Senders need not be members
- Group membership not known explicitly
- Analogy:
  - Each multicast address is like a radio frequency, on which anyone can transmit, and to which anyone can tune-in.

IP Multicast Addresses

- Class D IP addresses
  - 224.0.0.0 – 239.255.255.255

1 1 1 0	Group ID

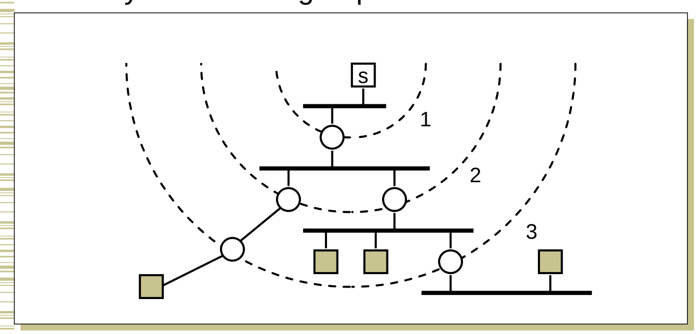
- How to allocated these addresses?
  - Well-known multicast addresses, assigned by IANA
  - Transient multicast addresses, assigned and reclaimed dynamically, e.g., by "sdr" program

IP Multicast API

- Sending – same as before
- Receiving – two new operations
  - Join-IP-Multicast-Group(group-address, interface)
  - Leave-IP-Multicast-Group(group-address, interface)
  - Receive multicast packets for joined groups via normal IP-Receive operation
  - Implemented using socket options

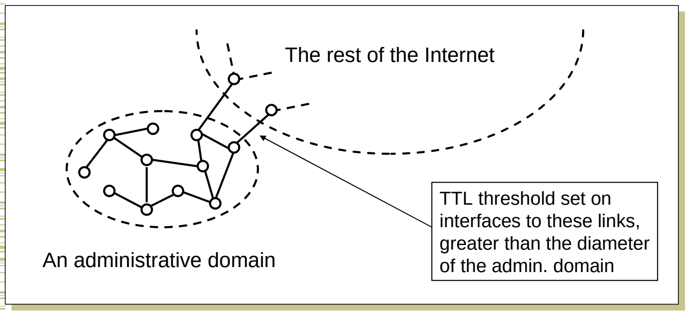
Multicast Scope Control – Small TTLs

- TTL expanding-ring search to reach or find a nearby subset of a group



Multicast Scope Control – Large TTLs

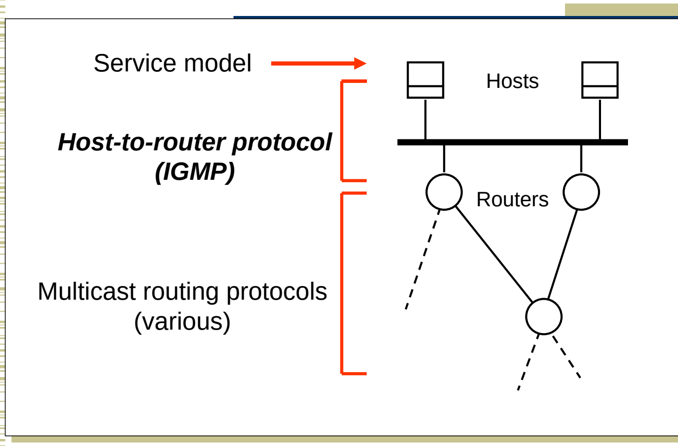
- Administrative TTL Boundaries to keep multicast traffic within an administrative domain, e.g., for privacy or resource reasons



Overview

- IP Multicast Service Basics
- Host/Router Interaction
- MOSPF/DVMRP
- Overlay Multicast

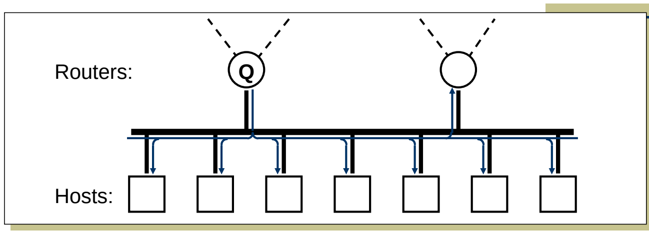
IP Multicast Architecture



Internet Group Management Protocol

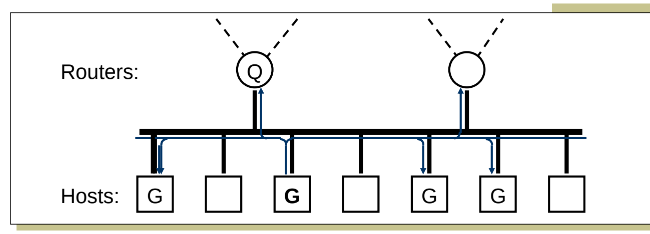
- End system to router protocol is IGMP
- Each host keeps track of which mcast groups are subscribed to
  - Socket API informs IGMP process of all joins
- Objective is to keep router up-to-date with group membership of entire LAN
  - Routers need not know who all the members are, only that members exist

How IGMP Works



- On each link, one router is elected the “querier”
- Querier periodically sends a Membership Query message to the all-systems group (224.0.0.1), with TTL = 1
- On receipt, hosts start random timers (between 0 and 10 seconds) for each multicast group to which they belong

How IGMP Works (cont.)



- When a host’s timer for group G expires, it sends a Membership Report to group G, with TTL = 1
- Other members of G hear the report and stop their timers
- Routers hear all reports, and time out non-responding groups

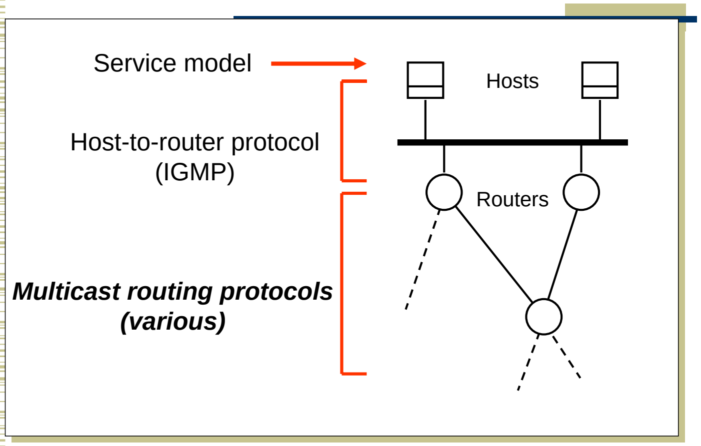
How IGMP Works (cont.)

- Note that, in normal case, only one report message per group present is sent in response to a query
  - Power of randomization + suppression
- Query interval is typically 60-90 seconds
- When a host first joins a group, it sends one or two immediate reports, instead of waiting for a query

Overview

- IP Multicast Service Basics
- Host/Router Interaction
- MOSPF/DVMRP
- Overlay Multicast

IP Multicast Architecture



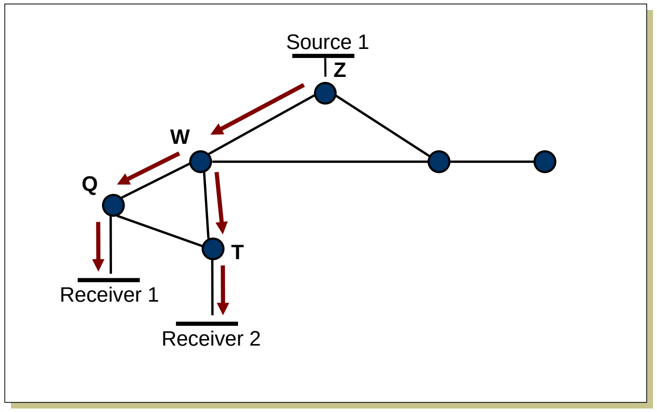
Routing Techniques

- Basic objective – build distribution tree for multicast packets
- Flood and prune
  - Begin by flooding traffic to entire network
  - Prune branches with no receivers
  - Examples: DVMRP, PIM-DM
  - *Unwanted state where there are no receivers*
- Link-state multicast protocols
  - Routers advertise groups for which they have receivers to entire network
  - Compute trees on demand
  - Example: MOSPF
  - *Unwanted state where there are no senders*

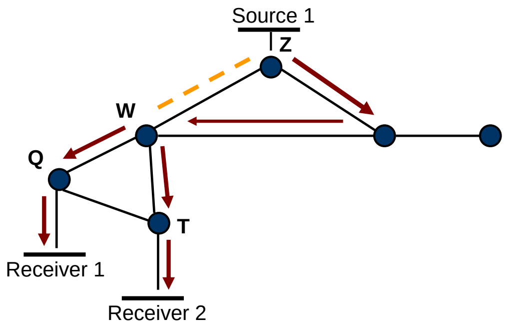
Multicast OSPF (MOSPF)

- Add-on to OSPF (Open Shortest-Path First, a link-state, intra-domain routing protocol)
- Multicast-capable routers flag link state routing advertisements
- Link-state packets include multicast group addresses to which local members have joined
- Routing algorithm augmented to compute shortest-path distribution tree from a source to any set of destinations

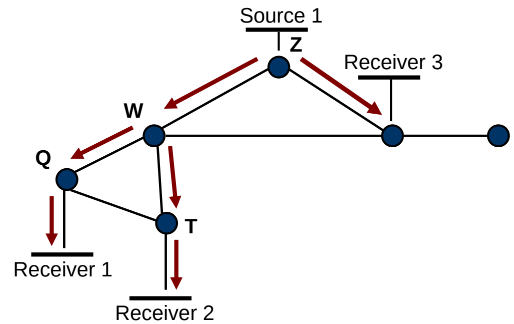
Example



Link Failure/Topology Change



Membership Change



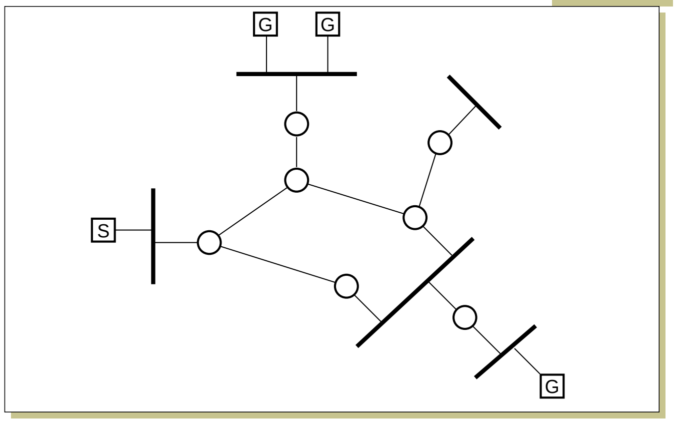
Impact on Route Computation

- Can't pre-compute multicast trees for all possible sources
- Compute on demand when first packet from a source S to a group G arrives
- New link-state advertisement
  - May lead to addition or deletion of outgoing interfaces if it contains different group addresses
  - May lead to re-computation of entire tree if links are changed

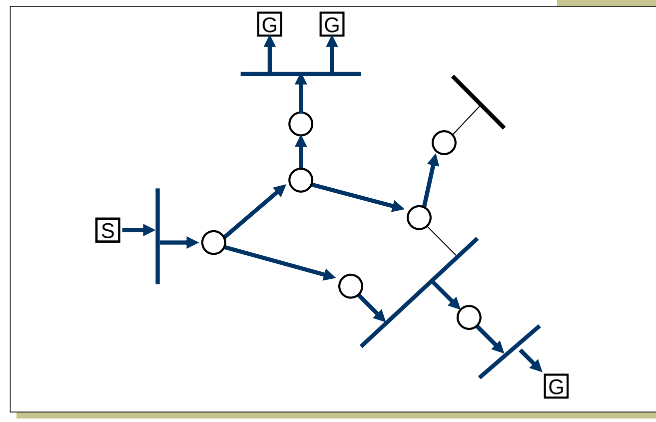
Distance-Vector Multicast Routing

- DVMRP consists of two major components:
  - A conventional distance-vector routing protocol (like RIP)
  - A protocol for determining how to forward multicast packets, based on the routing table
- DVMRP router forwards a packet if
  - The packet arrived from the link used to reach the source of the packet (reverse path forwarding check – RPF)
  - If downstream links have not pruned the tree

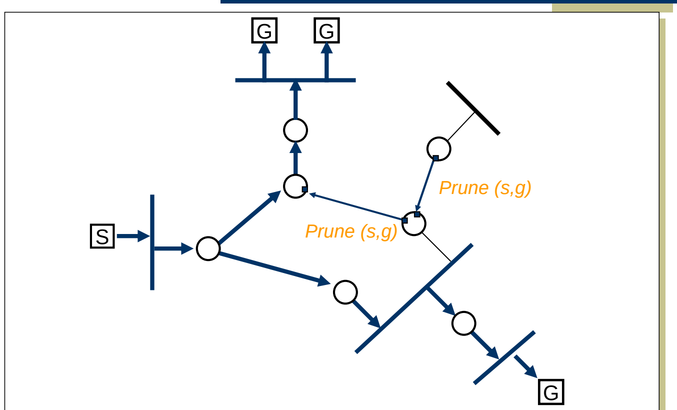
Example Topology



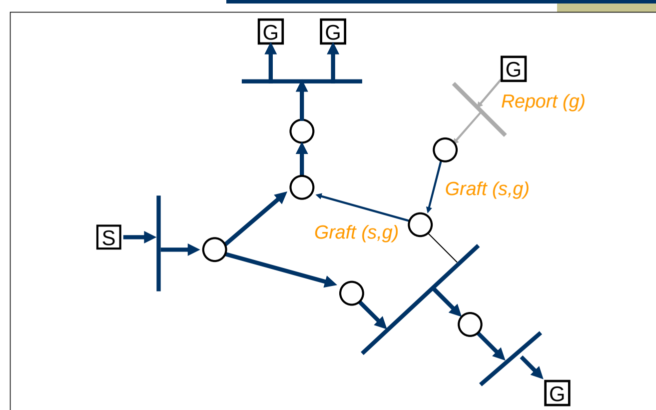
Broadcast with Truncation



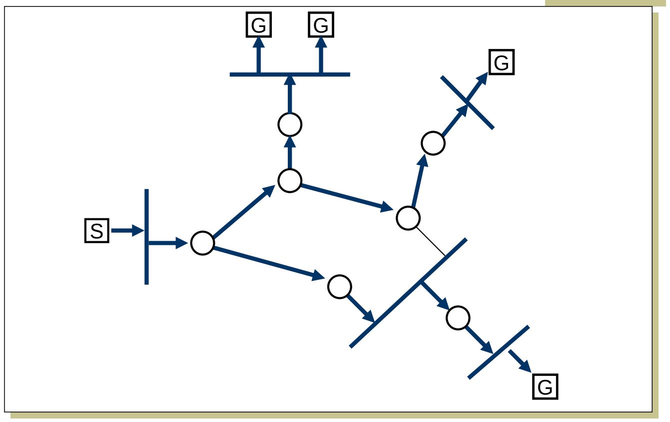
Prune



Graft



Steady State



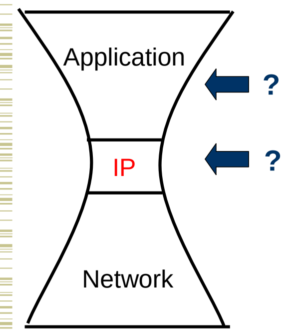
Overview

- IP Multicast Service Basics
- Host/Router Interaction
- MOSPF/DVMRP
- Overlay Multicast

Failure of IP Multicast

- Not widely deployed even after 15 years!
  - Use carefully – e.g., on LAN or campus, rarely over WAN
- Various failings
  - Scalability of routing protocols
  - Hard to manage
  - Hard to implement TCP equivalent
  - Hard to get applications to use IP Multicast without existing wide deployment
  - Hard to get router vendors to support functionality and hard to get ISPs to configure routers to enable

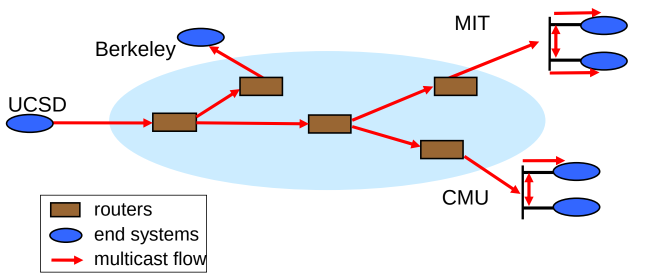
Supporting Multicast on the Internet



Internet architecture

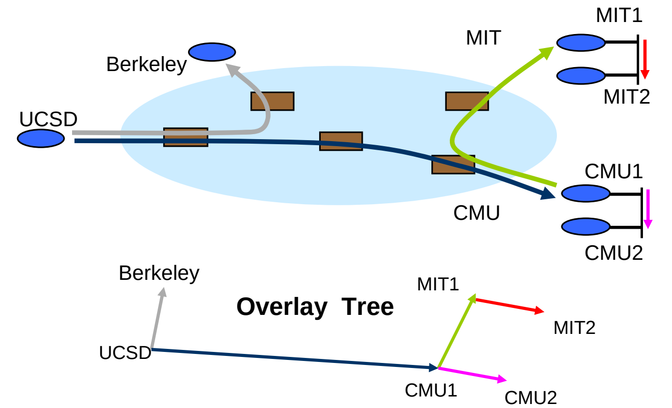
At which layer should multicast be implemented?

IP Multicast



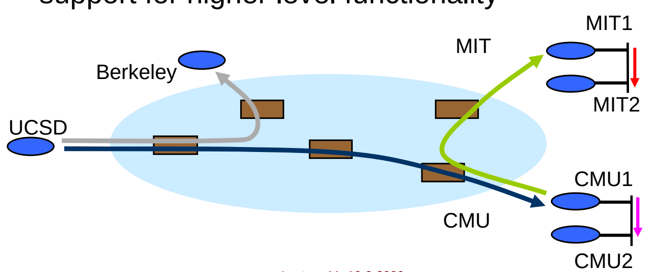
- Highly efficient
- Good delay

End System Multicast



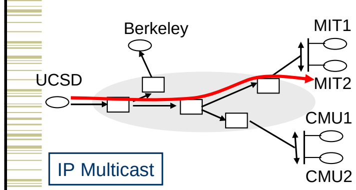
Potential Benefits Over IP Multicast

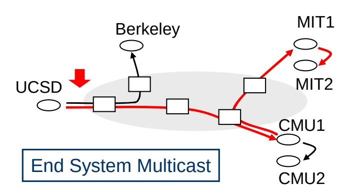
- Quick deployment
- All multicast state in end systems
- Computation at forwarding points simplifies support for higher level functionality



Concerns with End System Multicast

- Self-organize recipients into multicast delivery overlay tree
  - Must be closely matched to real network topology to be efficient
- Performance concerns compared to IP Multicast
  - Increase in delay
  - Bandwidth waste (packet duplication)
  - Penalty can be kept small in practice





Important Concepts

- Multicast provides support for efficient data delivery to multiple recipients
- Requirements for IP Multicast routing
  - Keeping track of interested parties
  - Building distribution tree
  - Broadcast/suppression technique
- Difficult to deploy new IP-layer functionality
- End system-based techniques can provide similar efficiency
  - Easier to deploy

Next Lecture: Wide Area Routing

- How to make routing scale to the size of the Internet
- How to accommodate business relationships in routing

EXTRA SLIDES

The rest of the slides are FYI

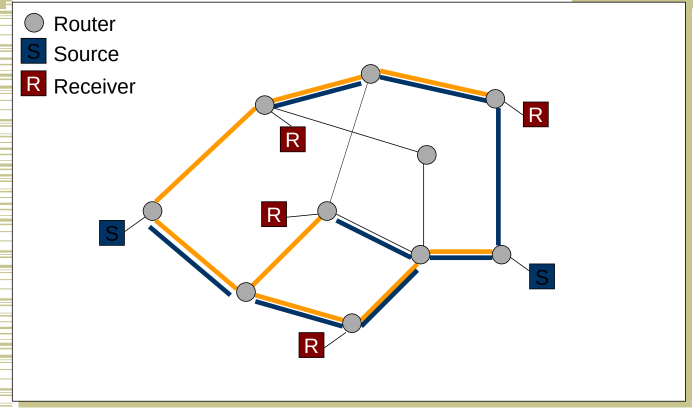
Routing Techniques

- Core based protocols
  - Specify “meeting place” aka core
  - Sources send initial packets to core
  - Receivers join group at core
  - Requires mapping between multicast group address and “meeting place”
  - Examples: CBT, PIM-SM

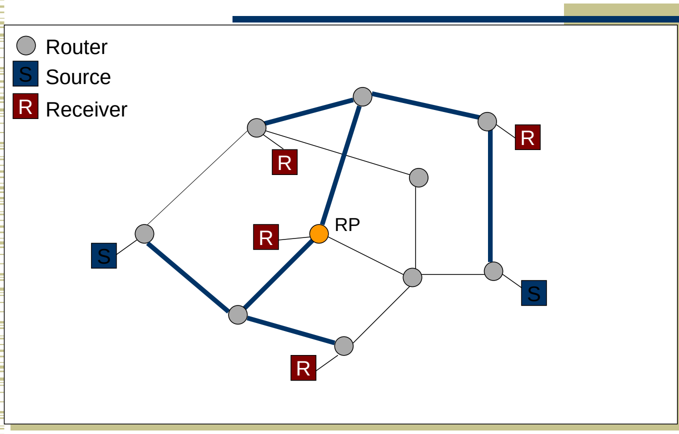
Shared vs. Source-based Trees

- Source-based trees
  - Separate shortest path tree for each sender
  - DVMRP, MOSPF, PIM-DM, PIM-SM
- Shared trees
  - Single tree shared by all members
  - Data flows on same tree regardless of sender
  - CBT, PIM-SM

Source-based Trees



Shared Tree

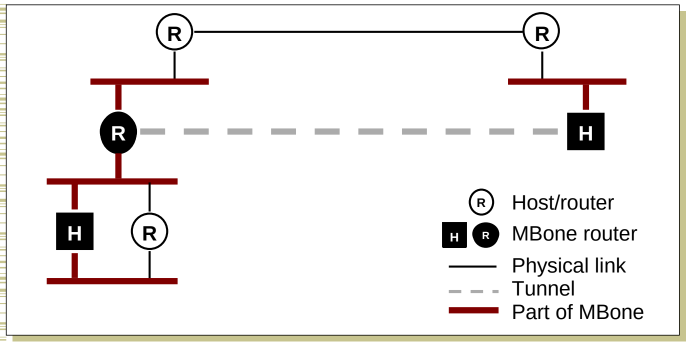


Shared vs. Source-Based Trees

- Source-based trees
  - Shortest path trees – low delay, better load distribution
  - More state at routers (per-source state)
  - Efficient for in dense-area multicast
- Shared trees
  - Higher delay (bounded by factor of 2), traffic concentration
  - Choice of core affects efficiency
  - Per-group state at routers
  - Efficient for sparse-area multicast
- Which is better? → extra state in routers is bad!

Multicast Backbone (MBone)

- An overlay network of IP multicast-capable routers



MBone Tunnels

- A method for sending multicast packets through multicast-ignorant routers
- IP multicast packet is encapsulated in a unicast packet addressed to far end of tunnel:

IP header, dest = unicast	IP header, dest = multicast	Transport header and data…

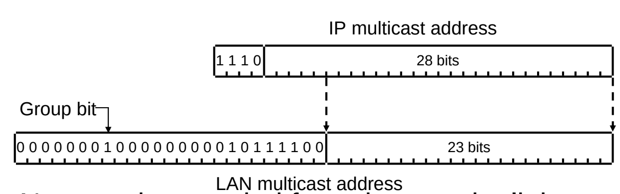
- Tunnel acts like a virtual point-to-point link
- Each end of tunnel is manually configured with unicast address of the other end

Link-Layer Transmission/Reception

- Transmission
  - IP multicast packet is transmitted as a link-layer multicast, on those links that support multicast
  - Link-layer destination address is determined by an algorithm specific to the type of link
- Reception
  - Necessary steps are taken to receive desired multicasts on a particular link, such as modifying address reception filters on LAN interfaces
  - Multicast routers must be able to receive <u>all</u> IP multicasts on a link, without knowing in advance which groups will be used

Using Link-Layer Multicast Addresses

- Ethernet and other LANs using 802 addresses:



- No mapping needed for point-to-point links