IK2215 Advanced Internetworking

Lecture 4—Multicast
Markus Hidell

Contents

Kurose & Ross: not covered in 7th edition, not much covered in 6th edition

IP Multicast Applications

Unicast is point-to-point

But many applications relay the same information to many receivers

Examples:

- Video conferenceing
- Internet radio, tv distribution
- Distribution of control information
- Distributed games

IP Multicast: Abstraction of HW Multicast

The *Internet abstraction* of hardware multicasting

Prime architect: Steve Deering

Group addresses (class D)

Exploits multicast-capable networking hardware if available

Best-effort delivery semantics

Receiver-based multicast:

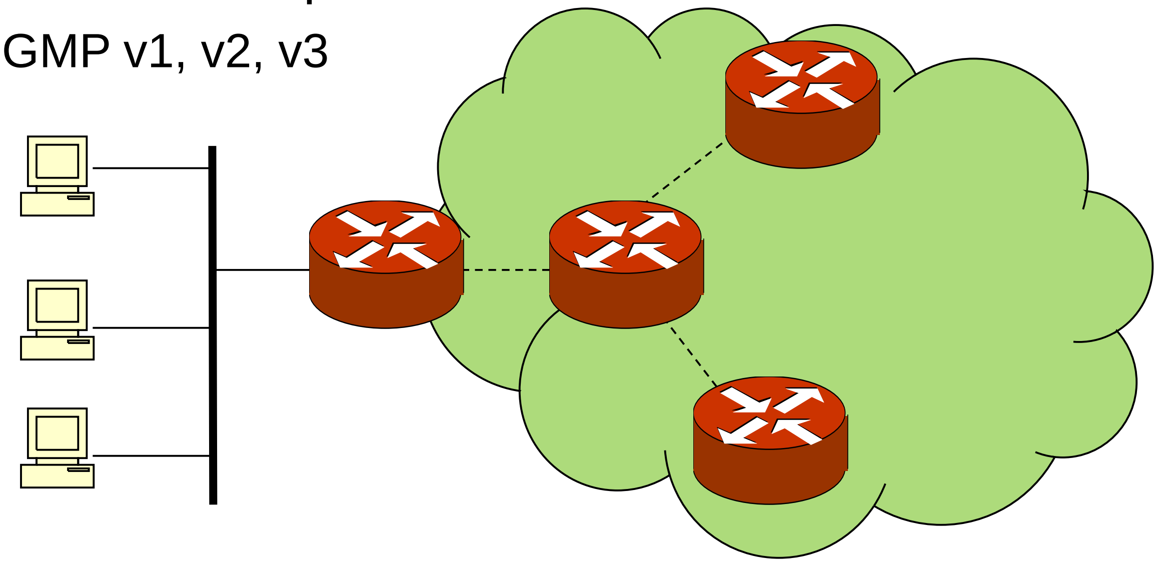
- Senders send to any group
- Receivers join groups

Dynamic group membership

- Hosts leave and join groups dynamically

IP Multicast Service Model

2. Host-to router protocol:
   - IGMP v1, v2, v3



1. Hardware/Link-level Multicast
   - Ethernet

3. Multicast Routing Protocols
   - PIM, CBT, DVMRP, MOSPF, MBGP,...

IP Multicast Addresses

IP-multicast addresses, class D addresses (binary prefix: 1110)

224.0.0.0 - 239.255.255.255

28 bit multicast group id

Selected addresses reserved by IANA for special purposes:

Address	Description
224.0.0.0 – 224.0.0.255	Local Network Control Block (dont forward)
224.0.0.1	All Systems on this subnet
224.0.0.2	All Routers on this subnet
224.0.0.4	DVMRP Routers
224.0.0.9	RIP Routers
224.0.1.0 – 238.255.255.255	Global Scope
239.0.0.0-239.255.255.255	Limited Scope
239.253.0.0/16	Scope restricted to one site
239.192.0.0/16	Scope restricted to one organization

Link-level/Hardware Multicast

Ethernet—good example of hardware multicast

- Most Ethernet NICs support multicast
  - NIC: Network Interface Card

Ethernet multicast addresses:

- The low order bit of the high order byte is 1:

`*1:**:**:**:**:**`

Many NICs on the same network may listen to the same Ethernet multicast address

Other Link-level layers may not support multicast

- E.g., ATM, Frame Relay, X25
- But multicast can still be implemented over these

Mapping IP Multicast to Ethernet

To use HW multicast on a LAN, the *IP multicast address* is translated to an *Ethernet multicast address*.

- How is this done?

The 23 low order bits of the IP multicast address placed in the 23 low order bits of the Ethernet MAC address:

`01:00:5E:00:00:00`

Example, IP multicast address 227.141.54.33 (0xE38D3621):

`0xD3621 into 01:00:5E:00:00:00` ➔ `01:00:5E:0D:36:21`

IP to Ethernet multicast address mapping is *not unique*!

- 32:1 overlap
- IP may receive multicast despite the lack of receiving process
- IP-layer must be able to do filtering (based on IP multicast address)

IGMP—Internet Group Management Protocol

Group membership communication between hosts and multicast routers

- Not for routing of multicast packets

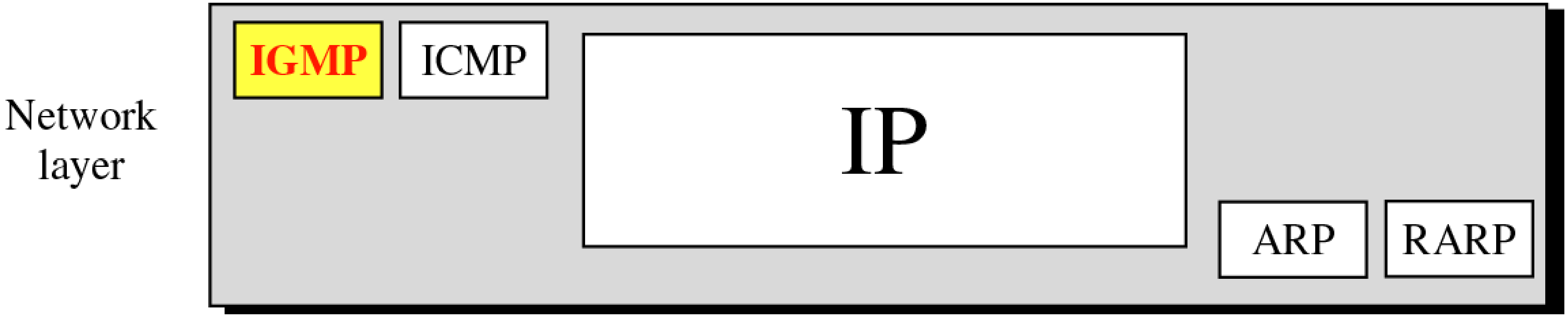
IGMP enables routers to maintain group members to each router interface

- Without IGMP, routers would have to broadcast all multicast packets

Internet Group Management Protocol – RFC 1112

- version 1 – RFC 1112 (Historic)
  - Query from router and response from host
- version 2 – RFC 2236
  - Leave group by host
- version 3 – RFC 3376 (not very common)
  - Source filtering

Position of IGMP in TCP/IP



Part of the network layer

- Encapsulated in IP (like ICMP)

IGMP messages always addressed to a multicast address

- often *all systems (224.0.0.1), all routers (224.0.0.2)*
- or to a specific multicast group

IGMPv2 Messages

General membership query

- Sent regularly by *routers* to query all membership

Specific membership query

- Sent by *routers* to query specific group membership

Membership report

- Sent by *hosts* to report joined groups

Leave group

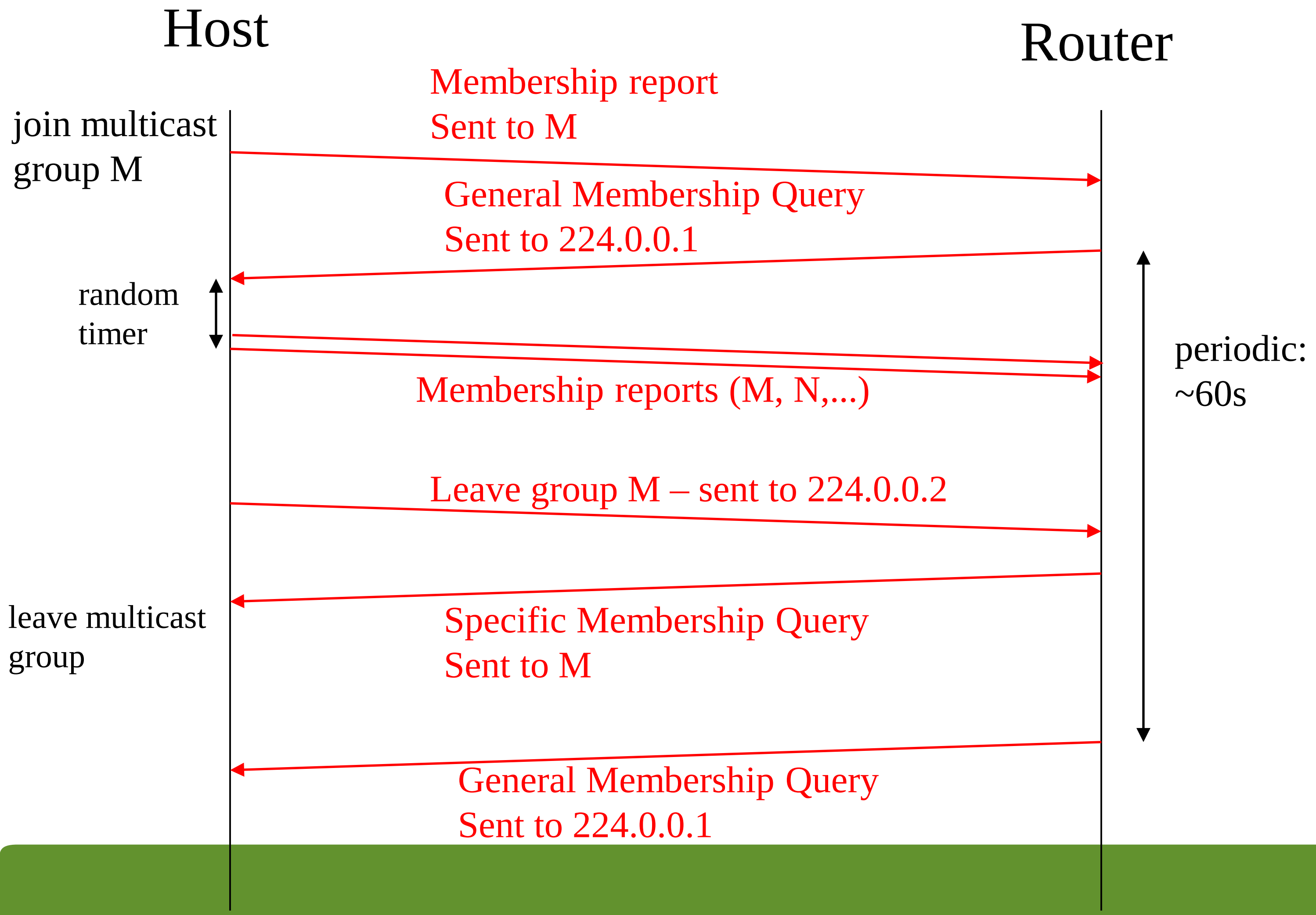
- Sent by *hosts* to leave groups

Host Behaviour

A process joins a multicast group on a given interface

- Host sends *IGMP report* to group address when first process joins a group.
  - Host keeps a table of all groups which have a reference count > 0
- Host sends *IGMP Leave* to 224.0.0.2 when last process leaves group
  - In IGMPv1 hosts did not send explicit leaves
- Router sends *IGMP queries* to 224.0.0.1 at regular intervals.
  - general query: group = 0.0.0.0
  - specific group query: group = multicast address of the group
- Host responds to IGMP query by sending IGMP report to group address
  - Hosts snoop for other hosts' reports
  - Set random timer → Suppress if other host on same segment sends it

Dynamics of IGMP Messages



IGMPv3

Allows selection of senders—not only groups

- Enables: Source specific multicast

A host can join a group and specific sender:

- (S, G) not only (*, G)

This may allow for pruning of certain senders

IGMPv3 is not commonly deployed

Multicast Router

Listens to all multicast traffic and forwards if necessary.

Multicast router listens to *all* multicast addresses.

- Ethernet: 2^{23} link layer multicast addresses
- Listens *promiscuously* to all LAN multicast traffic

Communicates with directly connected hosts via *IGMP*

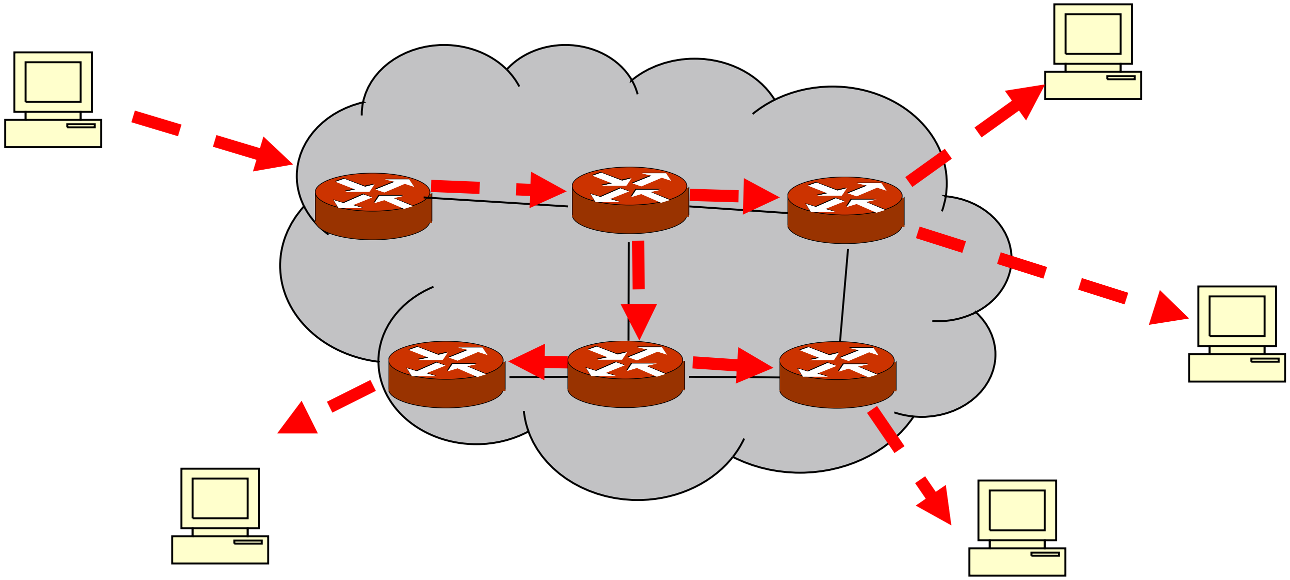
Communicates with other multicast routers with *multicast routing protocols*

Multicast Routing

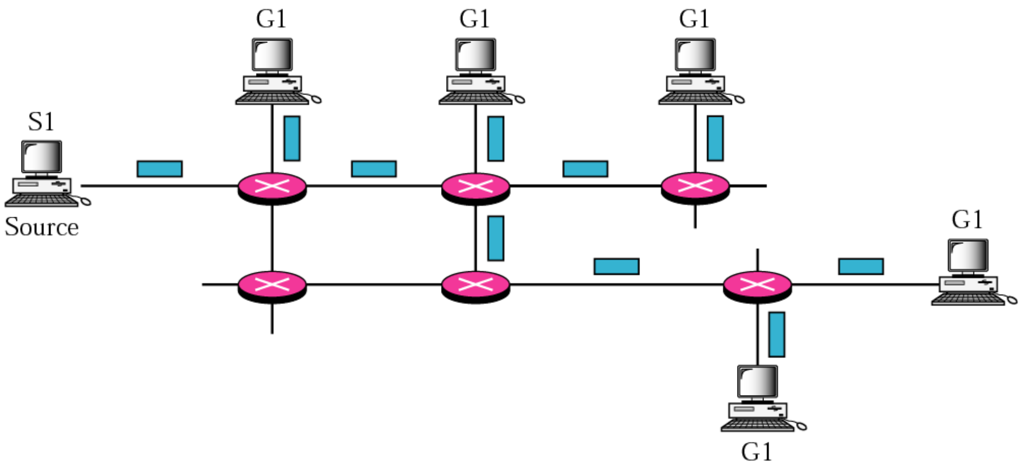
A packet received on a router is forwarded on multiple interfaces

The *network* replicates the packets—not the hosts

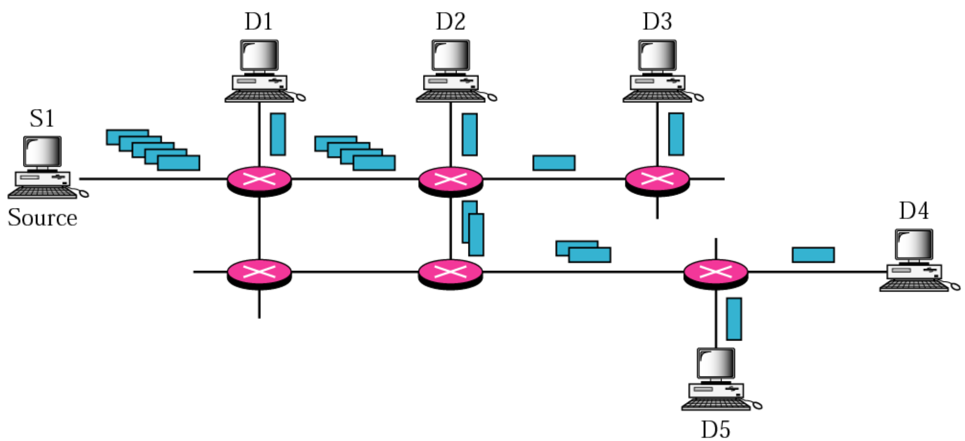
Build a delivery tree through a network



Multicast vs Multiple Unicast



a. Multicasting



b. Multiple unicasting

Delivery Trees

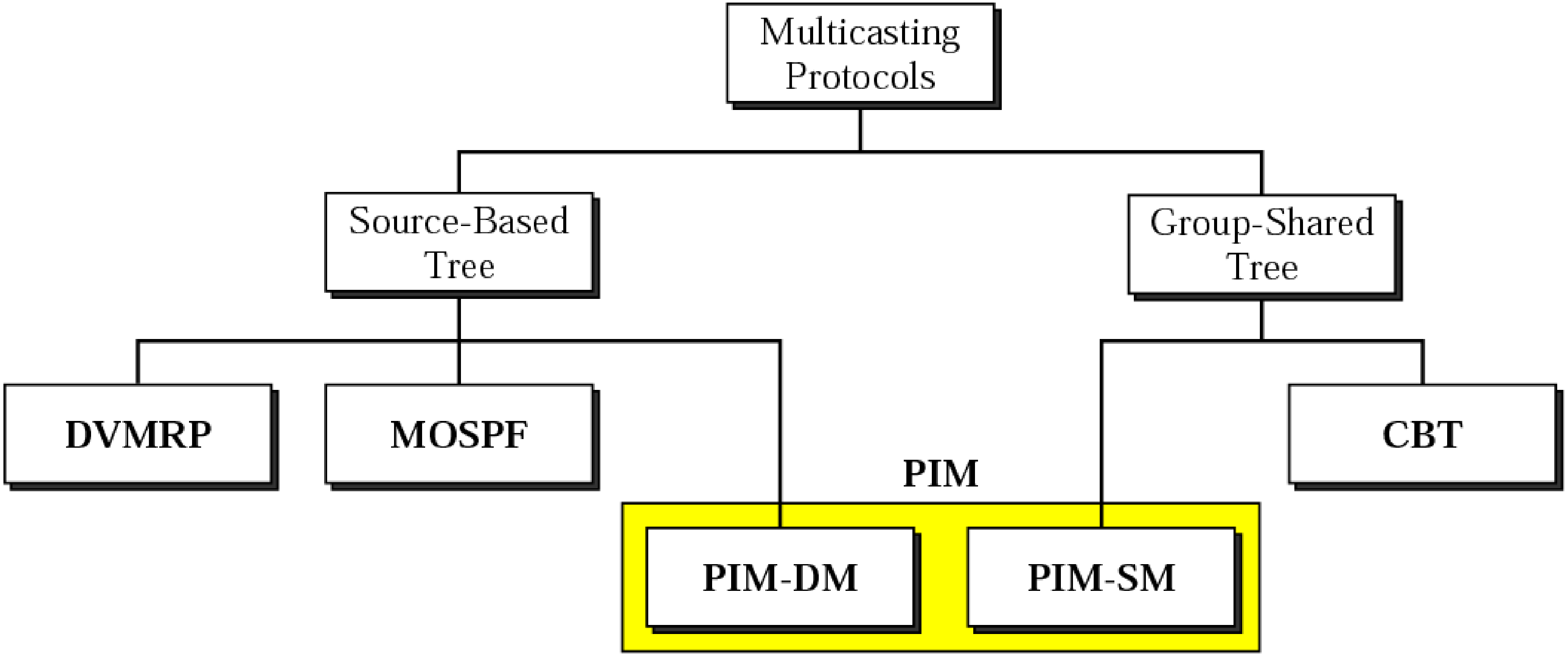
Source Based Trees

- *Each router* needs to have one shortest path tree for *each group*
- Notation: (S, G)
- Uses more memory (O(S*G)), but can give optimal paths and delay

Group Shared Trees

- *One router* (center core or renedevous router) is *responsible* for distributing multicast traffic
- Other routers encapsulates multicast packets in unicast packets and send them to the rendevous point for multicast distribution
- Notation: (*, G)
- Uses less memory (O(G)) but suboptimal paths and delays

Multicast Routing Protocols



DVMRP

Distance-Vector Multicast Routing Protocol

- Based on unicast distance vector (e.g., RIP)
- Routers do not know network topology apart from closest neighbour
- Create multicast routing table by using information from the unicast distance vector tables
- Extend (Destination, Cost, Nexthop) → (Group, Cost, Nexthops)

DVMRP is data-driven and uses source-based trees

DVMRP uses *Reverse Path Multicasting (RPM)*

RPM is best understood by looking at

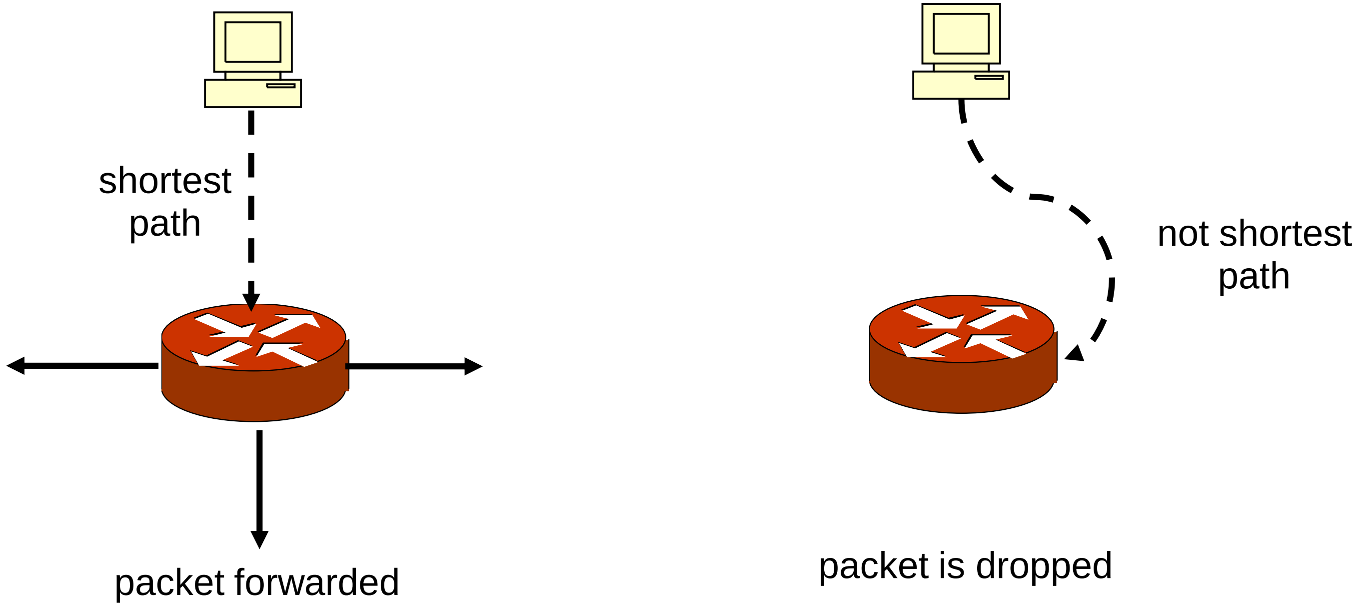
- Reverse Path Forwarding

Reverse Path Forwarding (RPF)

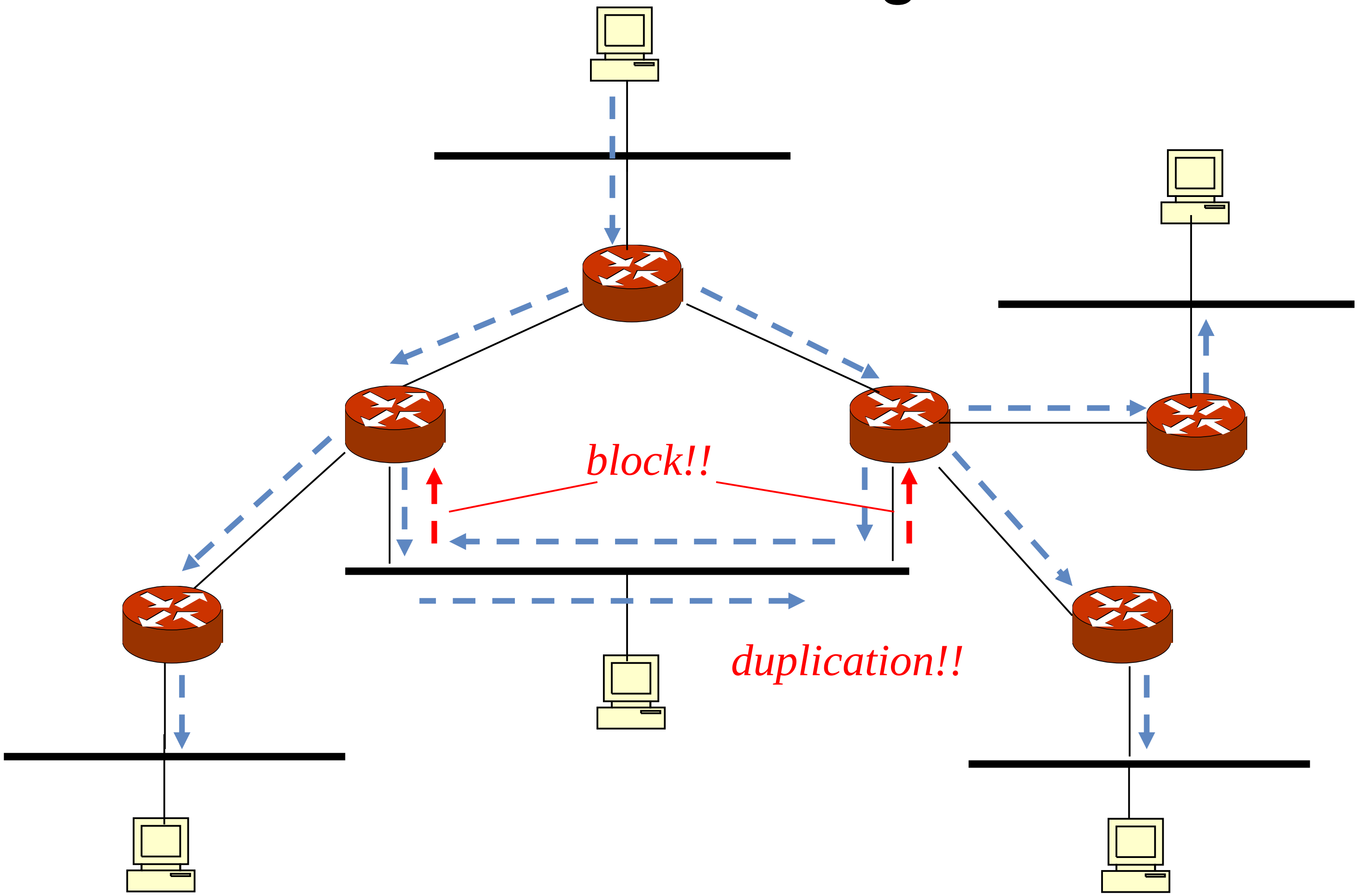
Forward a multicast datagram only if it arrives on the interface that would be used to send unicast to the source

- Send out on all other interfaces
- Flooding!

Make a lookup of the *source* address in the FIB!



Reverse Path Forwarding



Reverse Path Multicasting (RPM)

RPF leads to duplicates and flooding

RPM refines RPF as follows

Only *designated parent router* may forward multicast packets from a source to a link

- Removes duplicates

Flooding (Build the tree)

- First packet broadcast to every network

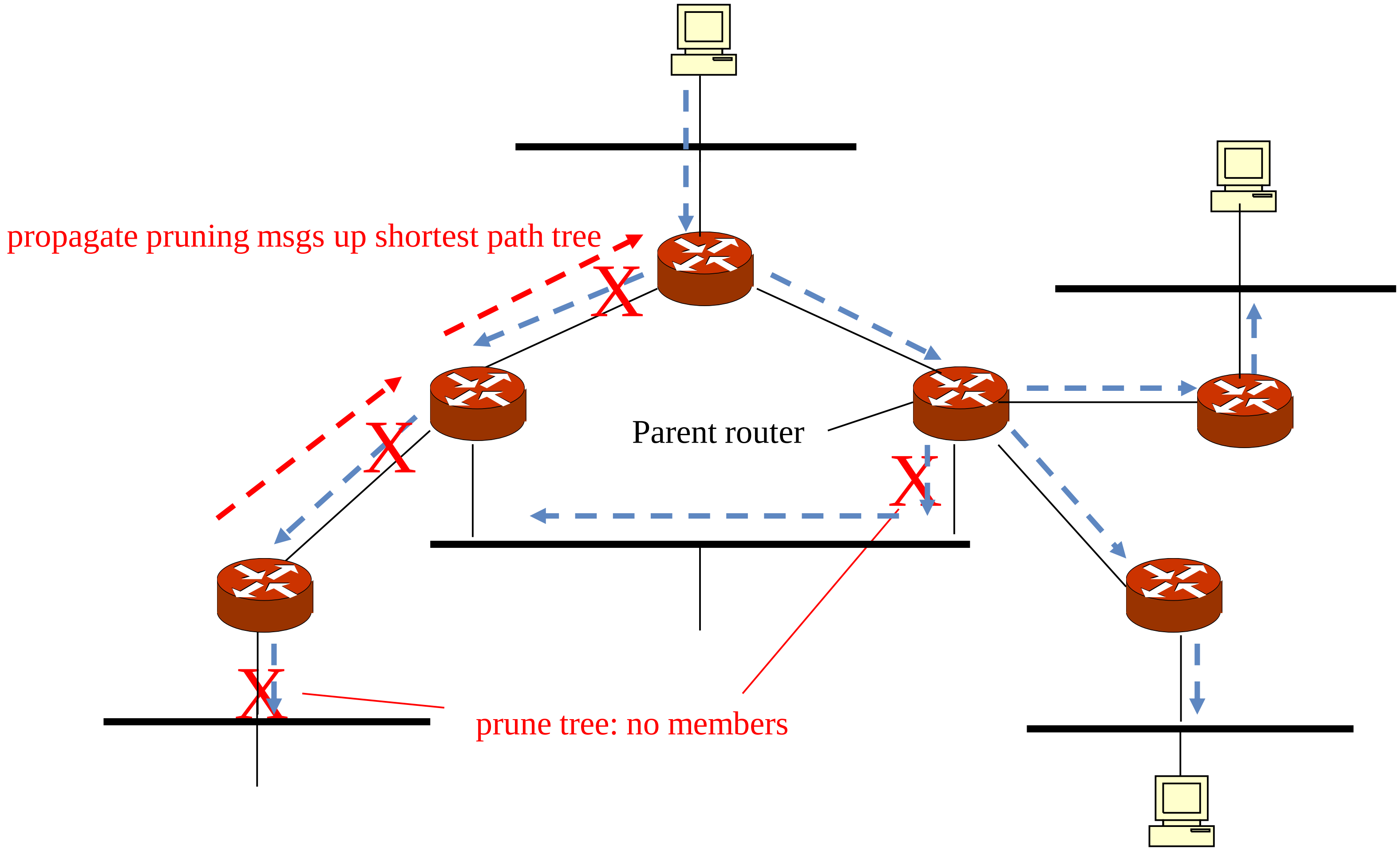
Pruning (Cut the tree)

- Prune networks that do not have members
- IGMP leave (or timeout)
- Propagate prune messages up the shortest path tree.

Grafting (Add a branch to the tree)

- Add a network with a listener
- IGMP join
- Propagate graft messages up the shortest path tree.

RPM Pruning Example



Link-State Multicast: MOSPF

Add multicast to a given link-state routing protocol

- MOSPF

Uses the multiprotocol facility in OSPF to carry multicast information

Extend LSAs with group-membership LSA

- Only containing members of a group

Uses the link-state database in OSPF to build delivery trees

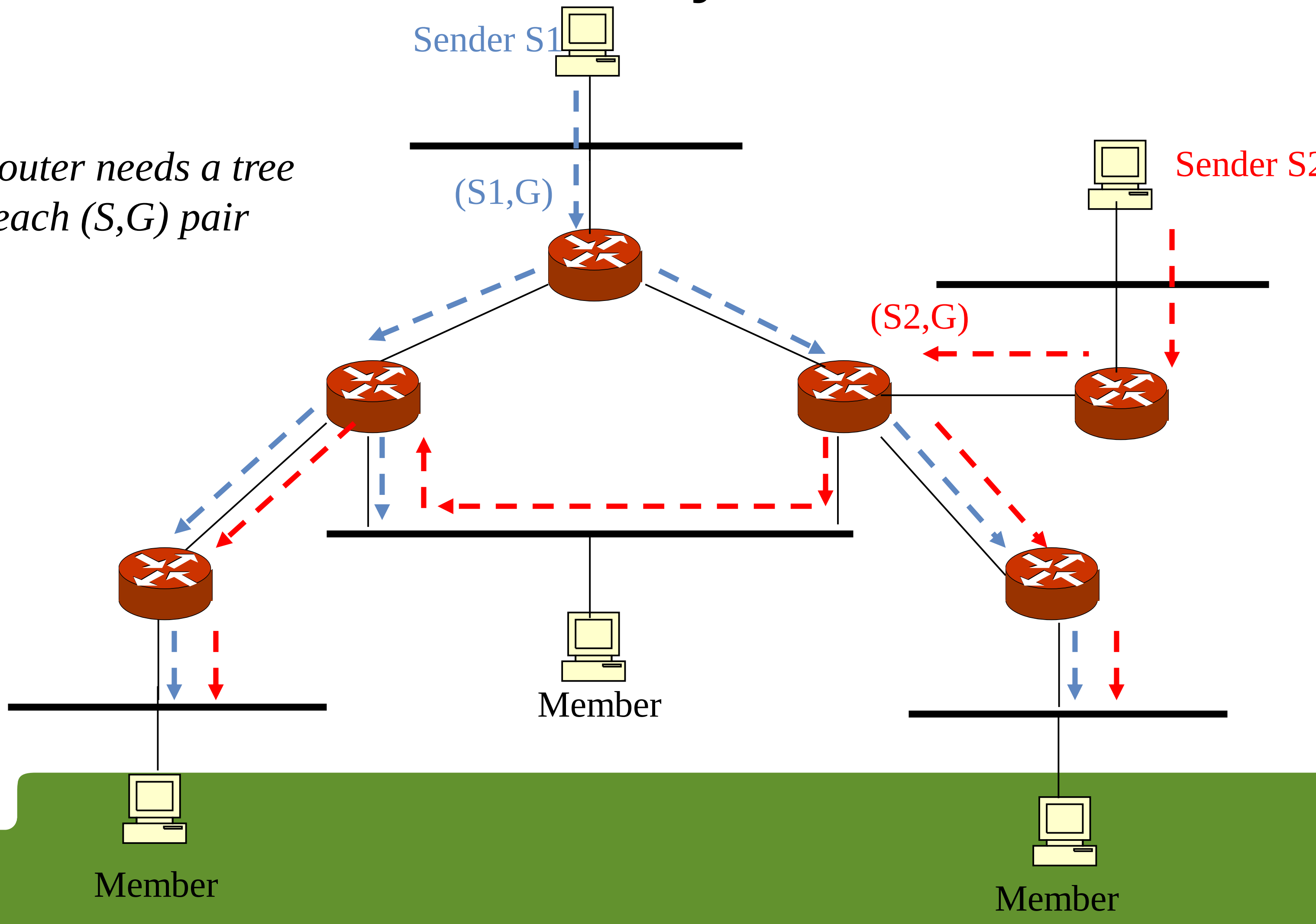
- Every router knows the topology of the complete network
- Least-cost *source-based trees* using metrics
- One tree for all (S,G) pairs with S as source

Expensive to keep all this information

- Cache active (S,G) pairs
- MOSPF is *Data-driven*: computes Dijkstra when datagram arrives

Link-State Multicast Example: Shortest Path Delivery Trees

Each router needs a tree for each (S,G) pair



Core Based Tree—CBT

CBT—Core Based Trees

Group shared multicast trees—(*, G)

Demand-driven

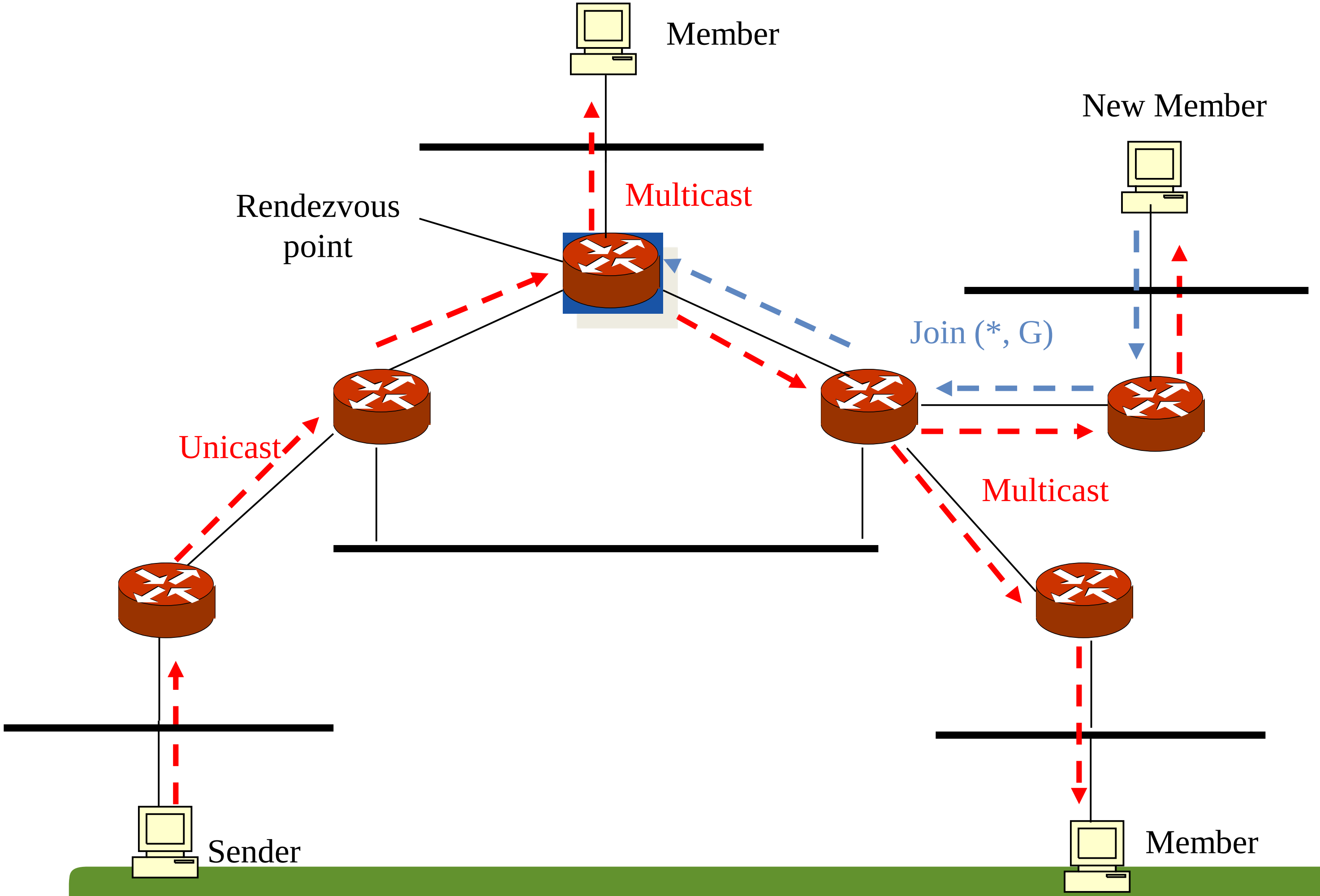
- Routers send join messages when hosts join groups

Divide the Internet into regions where each region has a *core router*

When a host joins a multicast group the nearest multicast router attaches to the forwarding tree by sending a join request towards its core router

Multicast datagrams to the core router are *encapsulated in unicast datagrams*

Shared Multicast Tree



Protocol Independent Multicasting—PIM

PIM-DM (dense mode)

- For dense multicast environment, like a LAN
- Uses RPF and pruning/grafting strategies—similar to DVMRP
  - Source-based tree
- Does not depend on a specific unicast protocol
- Relies on (any) correct unicast routing tables

PIM-SM (sparse mode)

- For non-broadcast environment (routers involved)
- Demand driven similar to CBT
  - uses *rendezvous points* (RPs) instead of core routers
- Extends CBT in that a router may know of more than one rendez-vous point
- Can build both shared and source distribution trees

MSDP

Multicast Source Discovery Protocol

Interconnects multiple PIM-SM domains

- Enables rendevous-point (RP) redundancy
- Enables inter-domain multicasting

Tunnels can be configured between RPs in various domains

- RPs speak MSDP to each other
- Enough tunnels so that we have connectivity even when an RP fails

Drawbacks:

- Scaling problem—many (S,G) pairs can be active in the Internet
  - Info must be passed about all these pairs
- Configuration-intensive (many tunnels needed)

MBGP

Solves part of the inter-domain problem

Standard BGP configuration facilities

- Extends the BGP multiprotocol attributes
- Exchange multicast routing information
- Policies, capabilities,

Must still use, for example, PIM to build distribution trees and forward multicast traffic

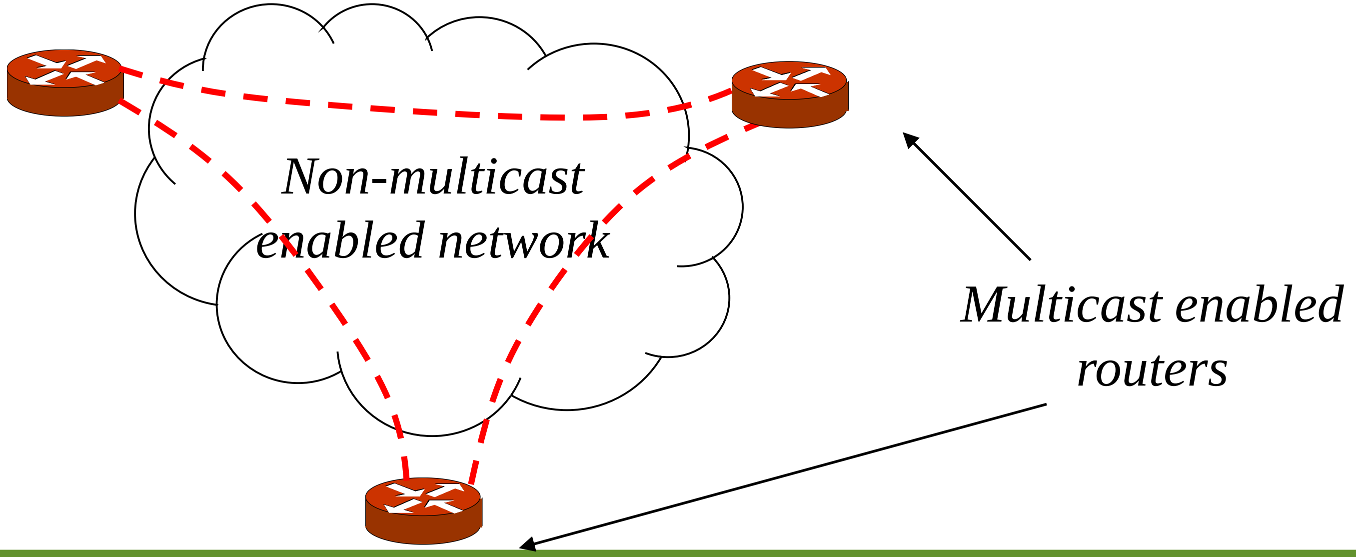
Tunneling

All routers need not be multicast enabled - does this mean that we cannot reach all hosts that want to join a multicast group on the Internet?

- No, because we can use tunneling over non-multicast enabled sub-nets
- cf VPN, IPv6

This is the way the MBONE – the Multicast BackBONE is constructed

- Islands of multicast enabled routers interconnected by tunnels



Deployment

Multicast routing is in general *not* deployed in current networks

Some sites (e.g., metroplitan area networks) have deployed local multicast delivery

- Cable TV distribution

IP multicast is slowly gaining acceptance

Summary: IP Multicast

Multicast routing uses network resources more efficiently than unicast emulation

IP multicast

- Receiver-based
- Best effort delivery

Multicast routing protocols

- DVMRP, MOSPF. CBT, PIM, MBGP

Source-based trees vs shared group trees

Demand-driven vs Data-driven trees

Reverse Path Forwarding (RPF)

Reverse Path Multicasting (RPM)