# COMP9332
# Network Routing & Switching

*Multicast Routing*

- www.cse.unsw.edu.au/~cs9332

# This lecture

- Multicast Routing
  - What
  - Why
- Group membership management
- IP Multicast routing

# IP Unicast and Broadcast

- IP Unicast
  - One sender and one receiver
  - Destination address in the IP packet = Unicast IP address of the receiver
- IP Broadcast
  - The message is sent to ALL hosts in a subnet or network
  - Destination identified by a broadcast IP address

# What is Multicast?

- Multicast
  - A sender sends to many receivers
- Multicast groups
  - Only the group members receive packets from that group
- Multicast groups can be dynamic or static
  - Dynamic: Hosts can dynamically join or leave a group

# Multicast Applications

- Webcast of live events on the Internet
  - Sports events/seminars/e-learning
  - TV or radio stations on the Internet
  - One sender, many receivers
- Teleconferencing
  - Example: vic (next slide)
  - Many senders, many receivers
- And many others

# vic—A Video Conferencing Tool

# IP Multicast address

- IP multicast address
  - Class D addresses (starts with 1110)
  - Each address identifies a group of receivers
  - Placed in the IP destination address field
- 32-4=28 bits to define a multicast address
  - 2^28 possible groups!
- Multicast addresses allocation can be
  - Static / Permanent
  - Dynamic

# Some Permanent Multicast Addresses

- Some examples are
  - 224.0.0.2: All routers on this subnet
  - 224.0.0.5: All OSPF routers
  - 224.0.0.9: RIP2 routers

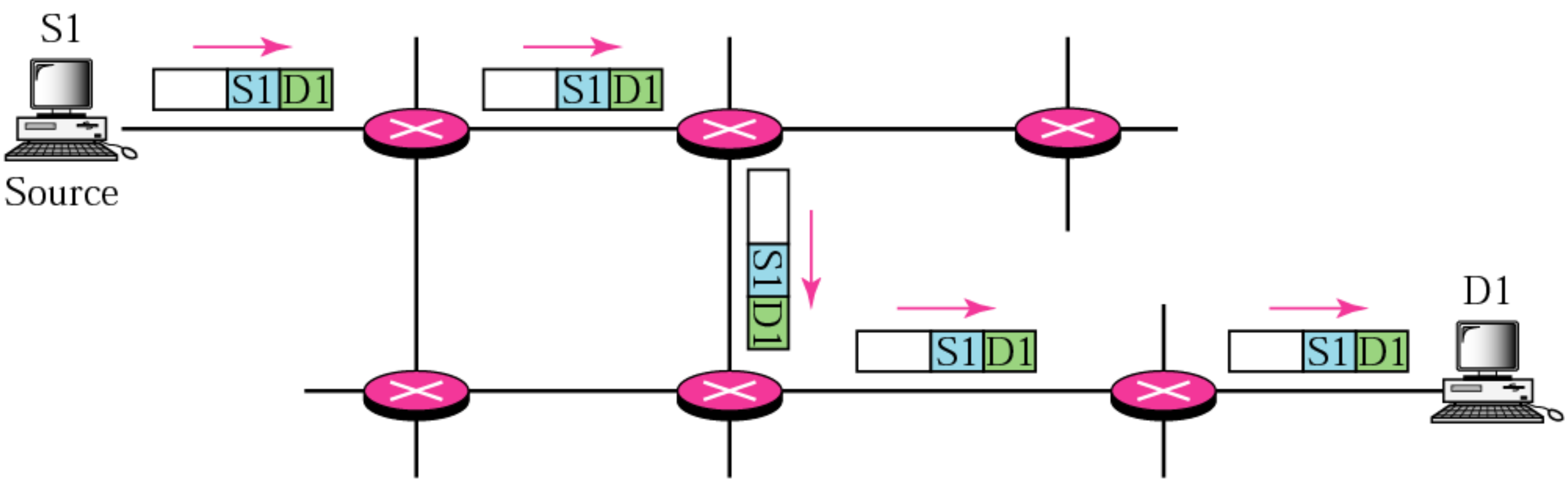
# Dynamic multicast address allocation

- Need to avoid collision
  - A multicast group may have members in multiple networks and autonomous systems
  - A multicast address may be used in another network/domain already
- Intra-domain address allocations
  - Multicast address allocation protocol (AAP)
  - A group of multicast address allocation servers coordinate the address allocation

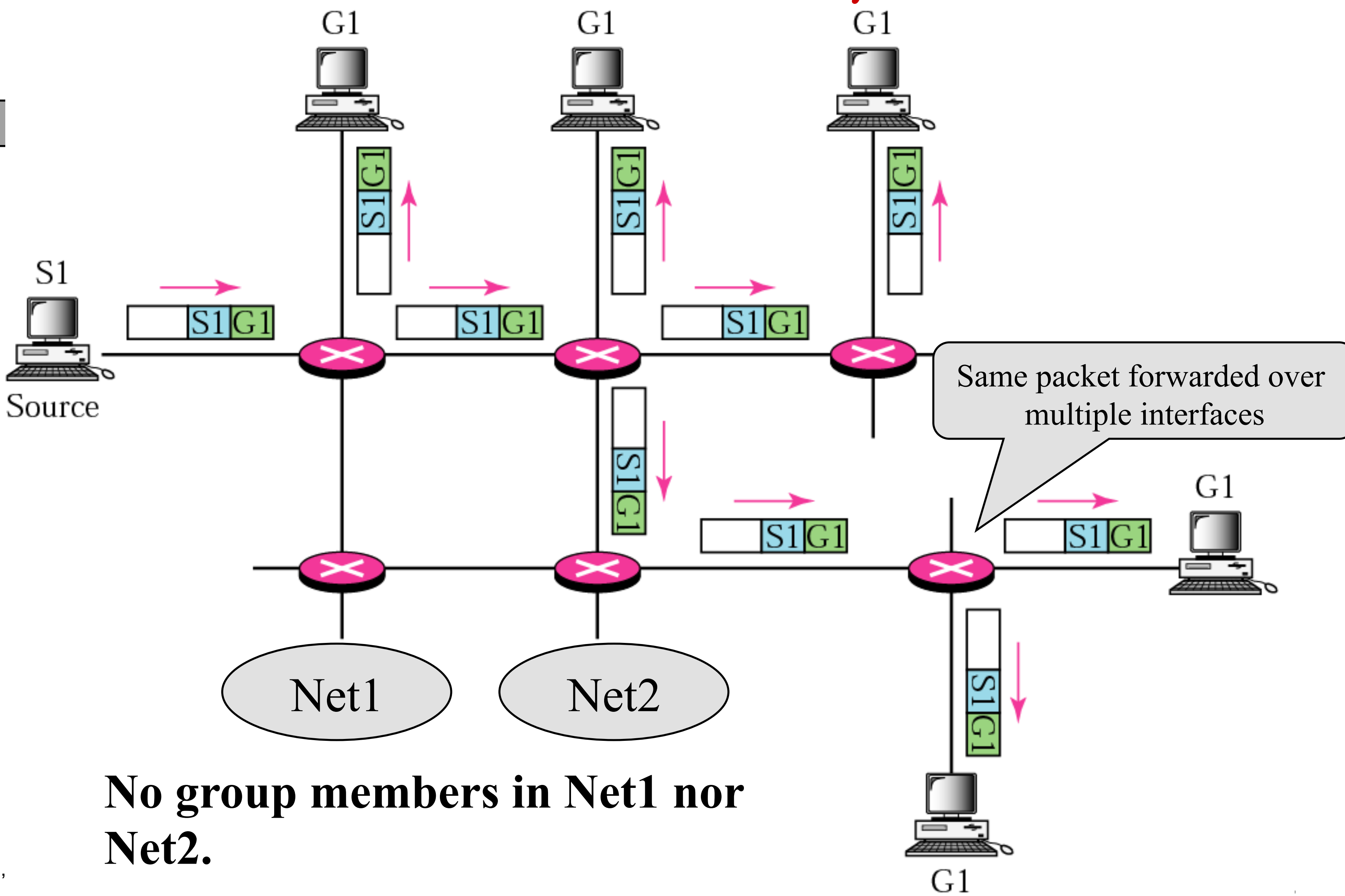
# Unicast versus multicast

- Unicast [Next slide]
  - Destination address is a unicast IP address
  - A path from source to destination
  - Forward packet through only one interface
- Multicast [The slide after]
  - Destination address is a multicast IP address
  - A tree rooted at the source to all receivers
  - May forward packet to multiple interfaces
  - Routers at the branching point make copies of the packets

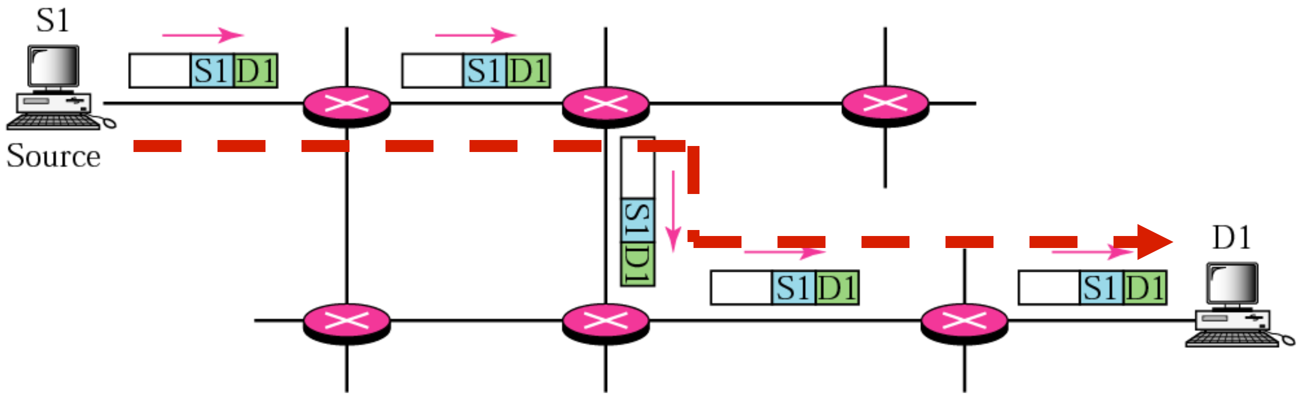
# Unicast delivery mechanism



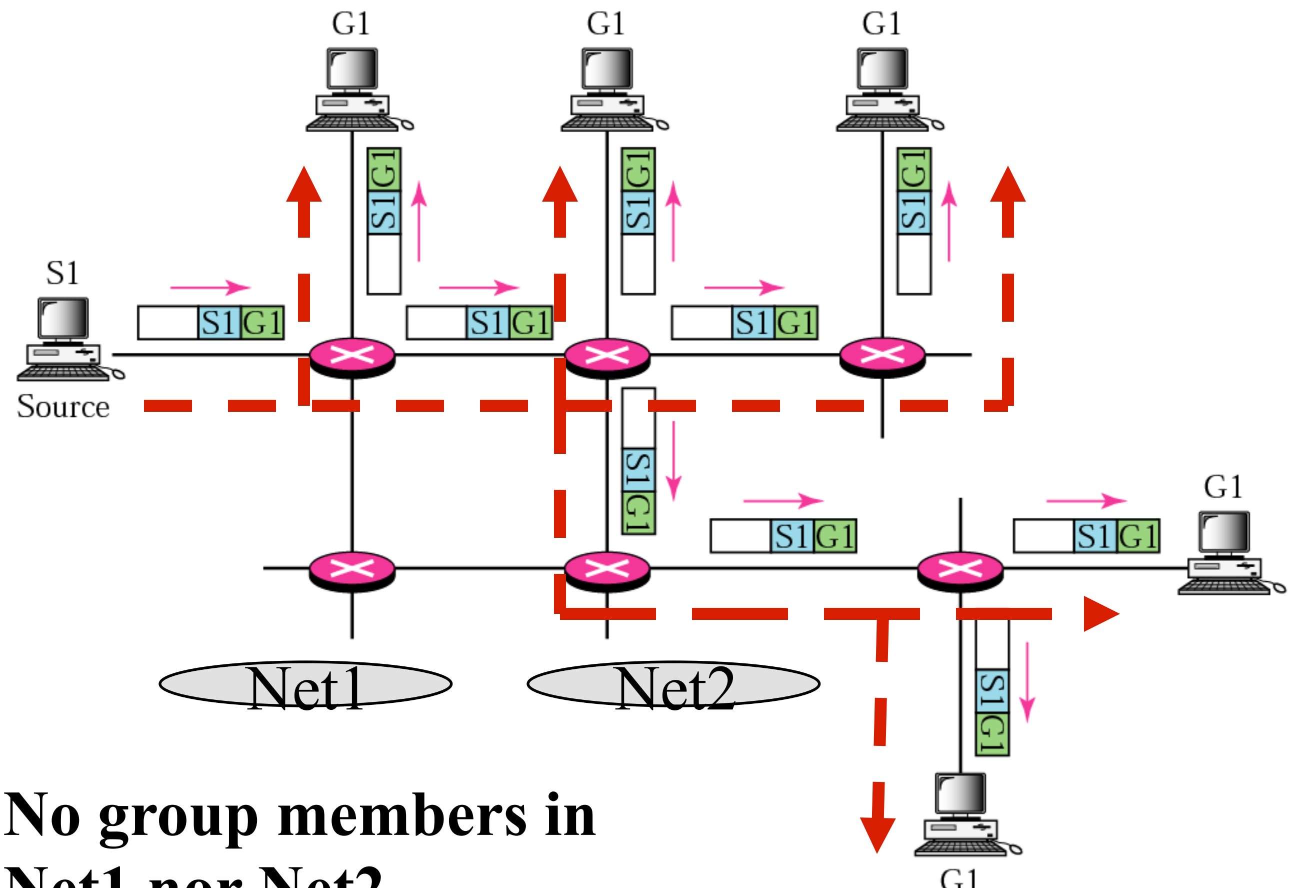
# Multicast delivery mechanism




**No group members in Net1 nor Net2.**




Unicast: A path from source to destination.



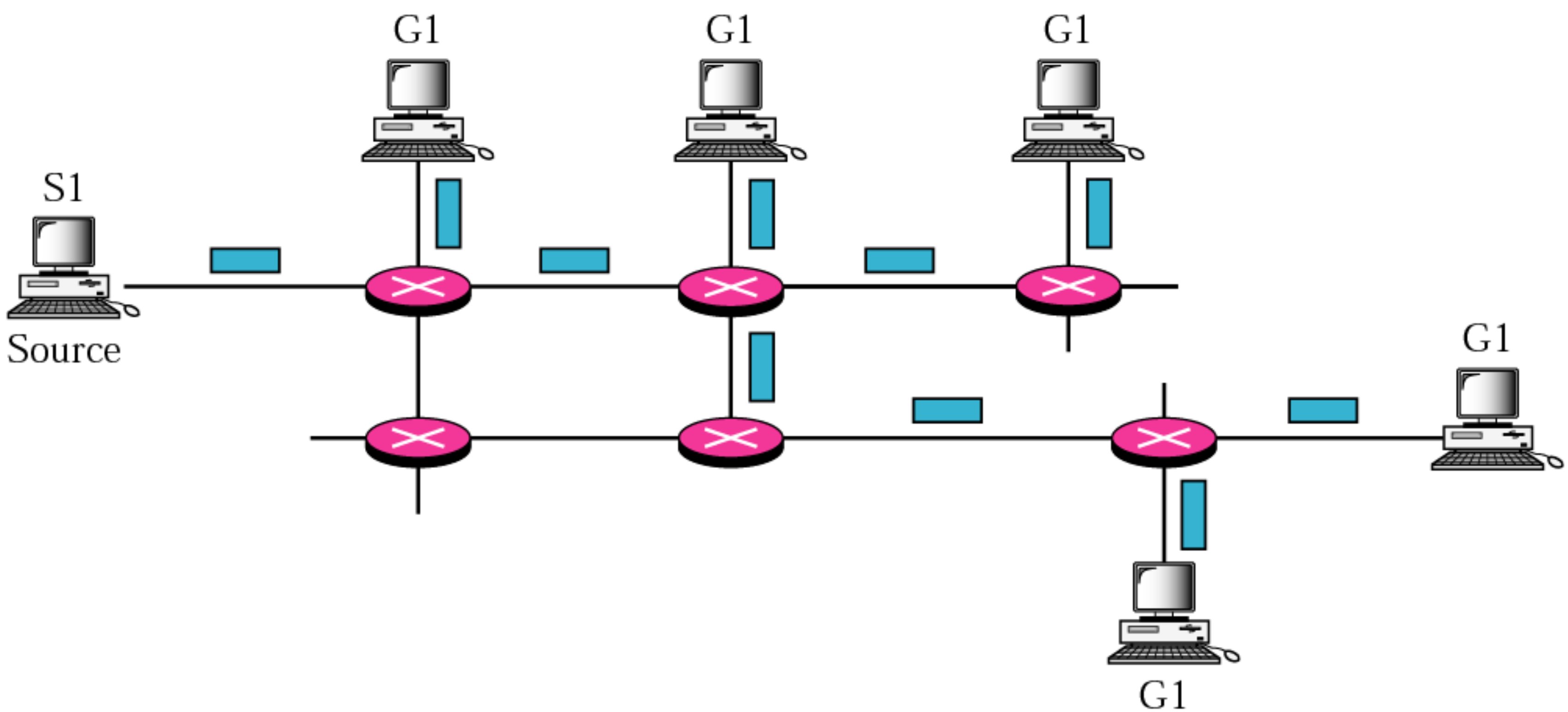

Multicast: A tree rooted at the source. Receivers are nodes of the tree.

**No group members in Net1 nor Net2.**

# Can we support multicast with unicasting?

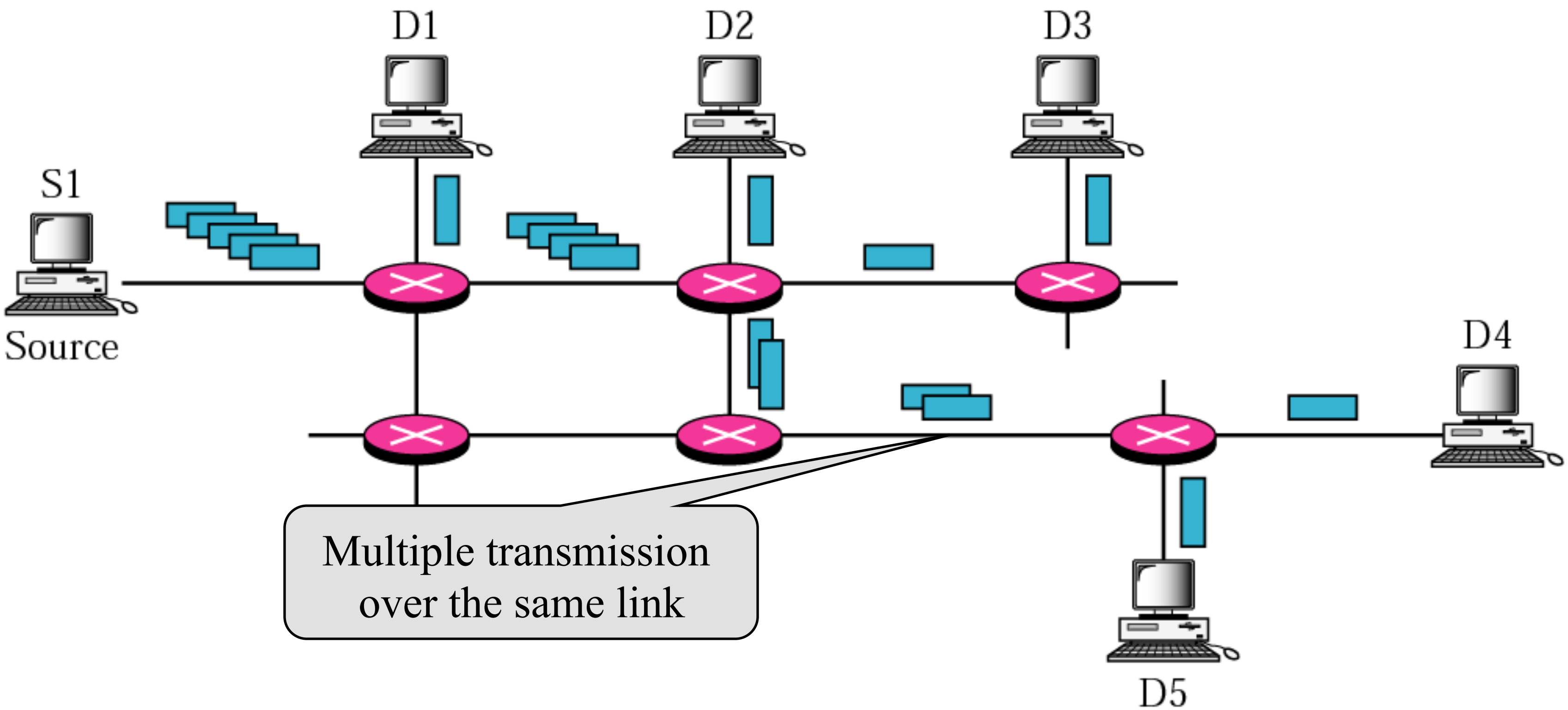
- In theory we can
- We can use separate unicast connections to separate receivers
- The problem is bandwidth overhead
  - does not scale for large number of receivers

# Multicasting versus multiple unicasting




a. Multicasting

# Multicasting versus multiple unicasting



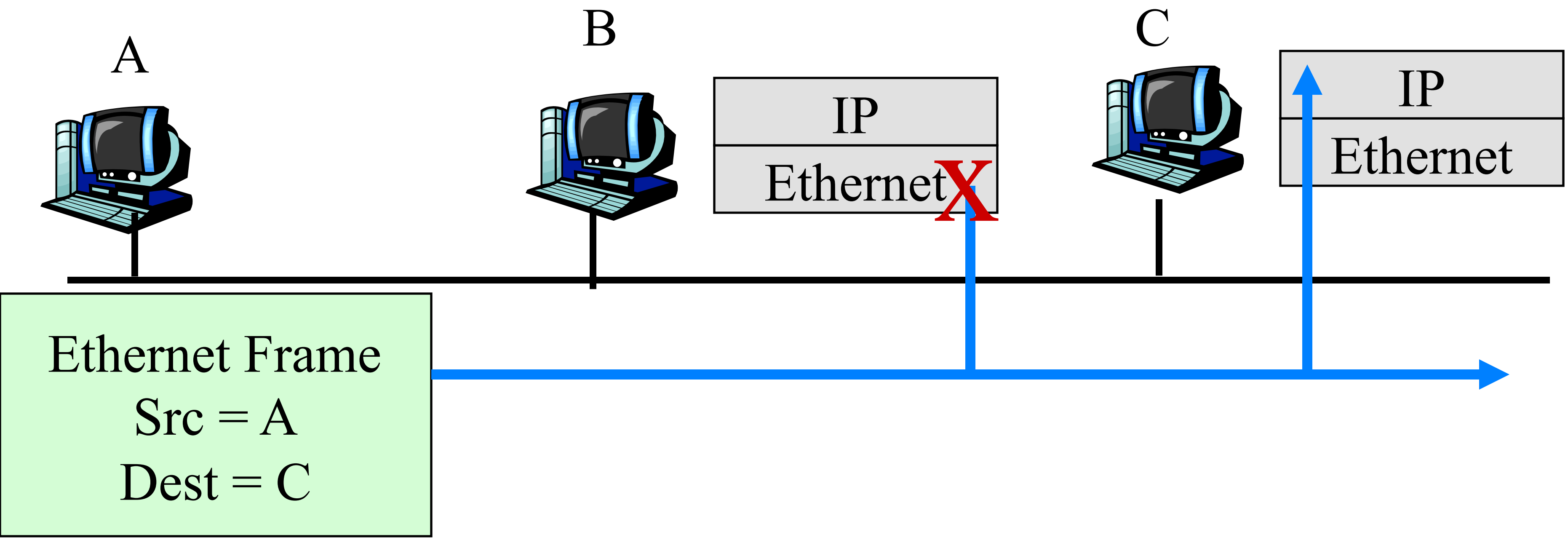

b. Multiple unicasting

# Multicast delivery

- Three cases
  - Within a network
  - Within an autonomous system
  - Within the Internet
- We begin with multicast within a network ( Ethernet specifics)

# Unicast Ethernet frame delivery

- All hosts hear the message
- Ethernet interface checks whether the destination hardware address corresponds to its own hardware address
  - Yes, pass it onto the upper layer
  - No, drop it.



# Broadcast & Multicast in Ethernet

- Broadcast
  - Destination Ethernet address: All 1's
    - » Ethernet address are 48-bit long
  - All hosts accept the frame
- Multicast
  - Low order bit of the high order octet
    - » 0 means unicasting, 00:00:00:00:00:00 (hex)
    - » 1 means multicasting, 01:00:00:00:00:00 (hex)

High-order octet

Low-order octet

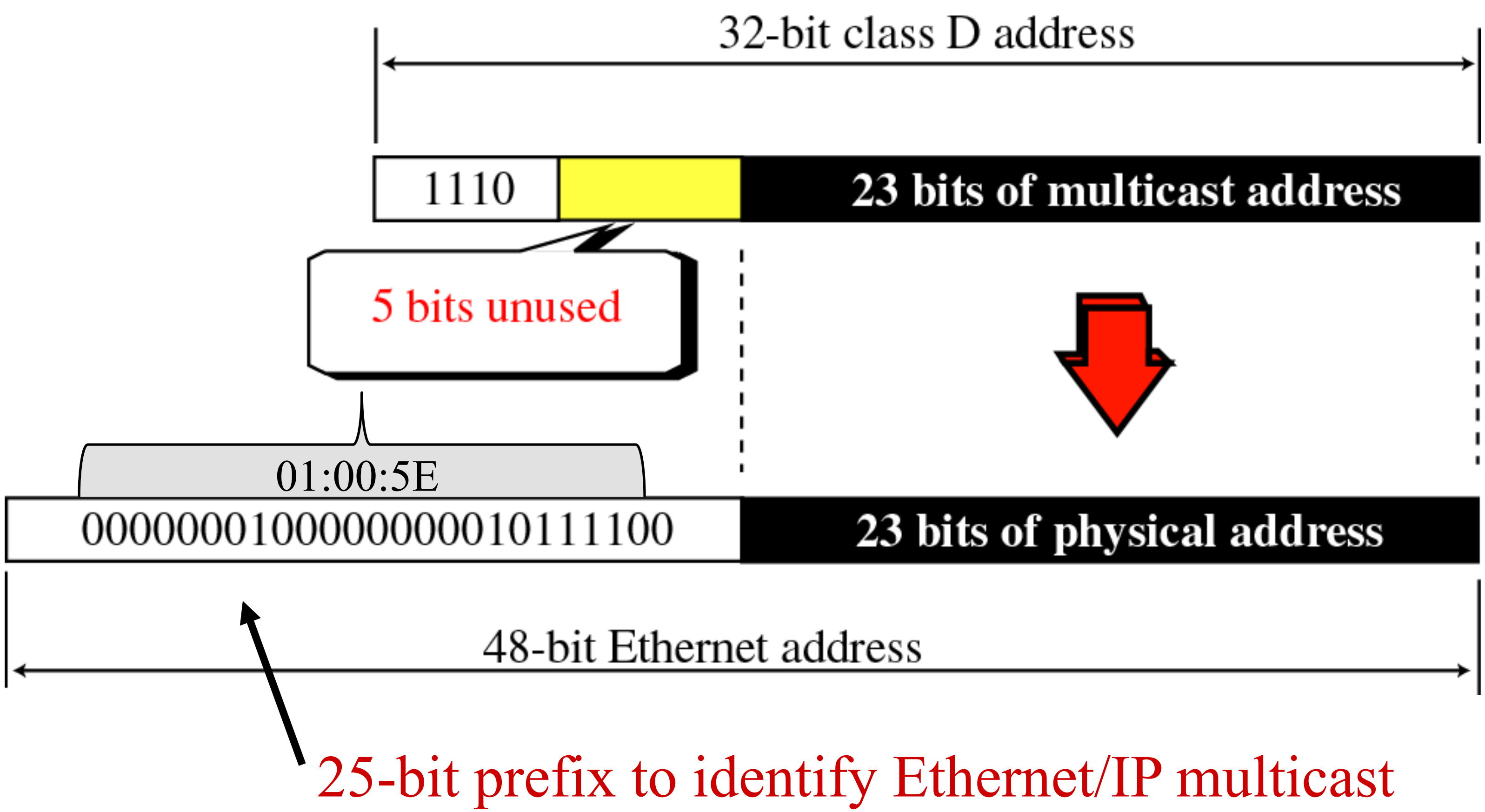
# Multicast in the Ethernet

- By default (no configuration needed), an Ethernet interface accepts
  - Unicast frames with its own Ethernet address
  - All broadcast frames
- The interface software can be configured to recognise one or more Ethernet multicast addresses
  - After this configuration, the interface accepts frames with these Ethernet multicast addresses

# Multicast prefix for MAC

- 48-bit Ethernet is grouped into two 24-bit
  - First 24-bit is organisationally unique identifier (OUI)
- The OUI of “01:00:5E” is used to identify multicasting (at MAC layer)
  - Could be for Ethernet/IP or any other protocols
- 25$^{th}$ bit is set to ZERO for IP multicast over Ethernet (25 bits gone! )
- 48-25 = 23 MAC bits left to carry IP multicast addresses
- But an IP multicast can be defined by 32-4=28 bits!

# Mapping multicast IP address to multicast Ethernet address



# IP multicast in Ethernet

- If a host belongs to an IP multicast group, it will ask its Ethernet interface to pick up the frames with the corresponding Ethernet multicast address

# Exercise

- (a) How many IP multicast groups are mapped to one Ethernet multicast address?
- (b) What is the effect of this many-to-one mapping?

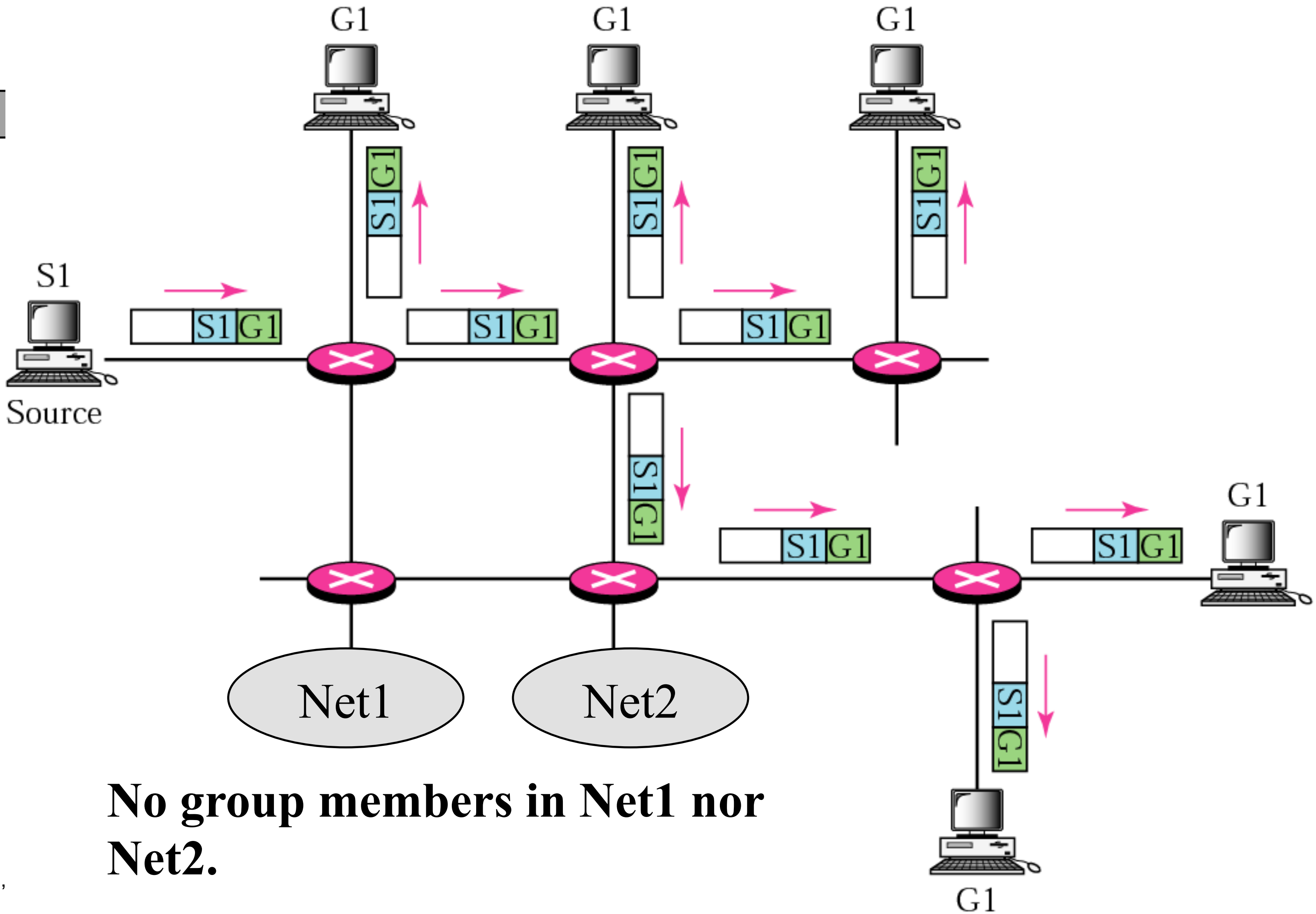
# Solution

- (a) 32
- (b) Some host interfaces may receive multicast packets that are not destined for that host.
  - Isn't this inefficient?
  - Isn't this wrong delivery?

# Solution (cont'd)

- The probability that 2 multicast addresses with identical low-order 23 bits in the same network is small.
- The Ethernet accepts the frame but IP will discard the packet if the host is not a group member

# IP multicast - internetwork



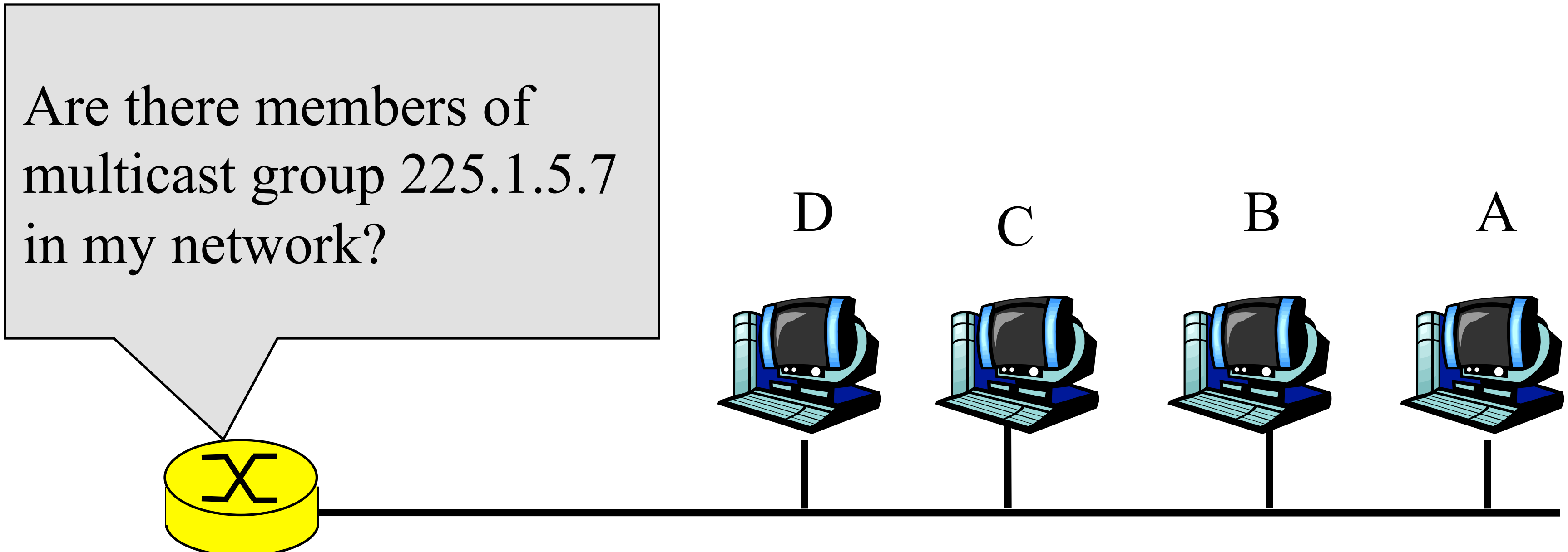

**No group members in Net1 nor Net2.**

# The multicast framework

- Two aspects:
  - A router needs to know whether there are group members behind its interfaces
  - How to efficiently deliver a multicast packet to all its group member?
- First aspect addressed by Internet Group Management Protocol (IGMP)
- Second aspect dealt with *multicast routing*

# IGMP

# Membership management

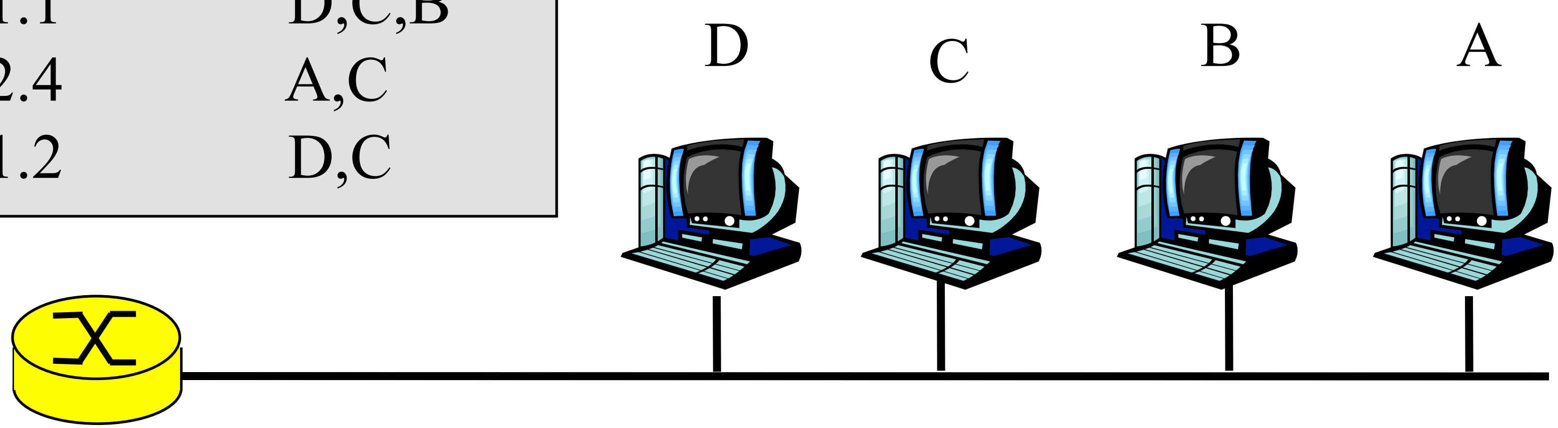



- Host can join and leave a multicast group at any time
  - The membership information has to be up to date
- Overheads
  - Number and frequency of management packets
  - Amount of state information

# Understanding state information: State information - possibility 1

States maintained by the router

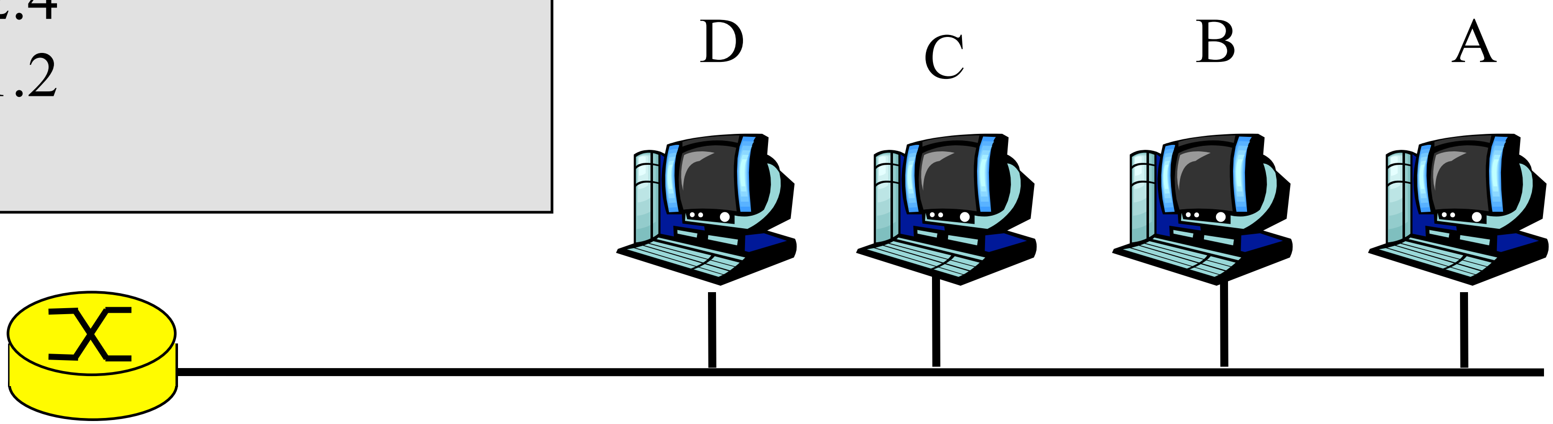
| Active Multicast groups | Active group members |
|---|---|
| 225.1.1.1 | D,C,B |
| 229.1.2.4 | A,C |
| 230.5.1.2 | D,C |



# State information - possibility 2

States maintained by the router

Active multicast groups
225.1.1.1
229.1.2.4
230.5.1.2



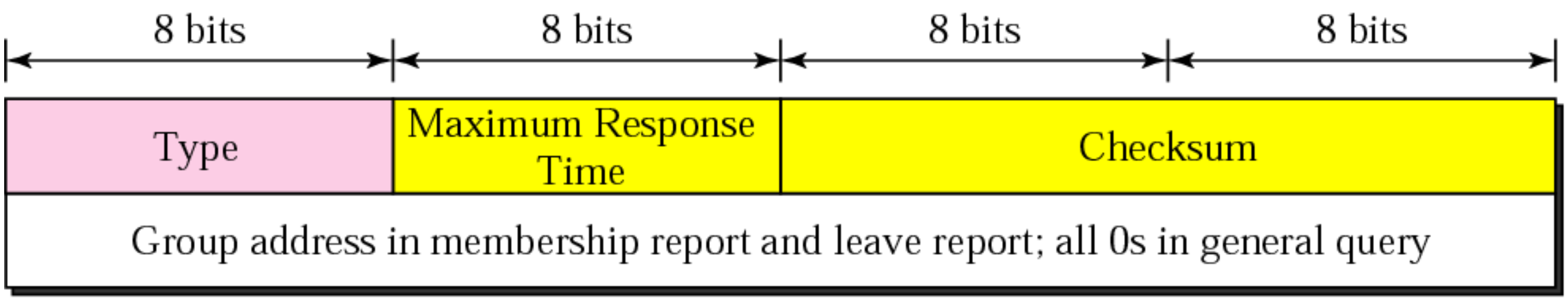
# *Internet group management protocol (IGMP)*

- IETF standard for managing multicast group memberships in a network
- With IGMP, a host can dynamically join and leave a multicast group
- State information
  - Each router interface maintains a list of multicast groups that are currently active
  - These states are *soft.*
    - Soft states expire if not renewed.

# *IGMP message type*

- For joining a group
  - Use *Membership Report*
- For leaving a group
  - Use *Leave Report*
- For a router to maintain group memberships in the network
  - *Query (Specific or General)*

# IGMP message format

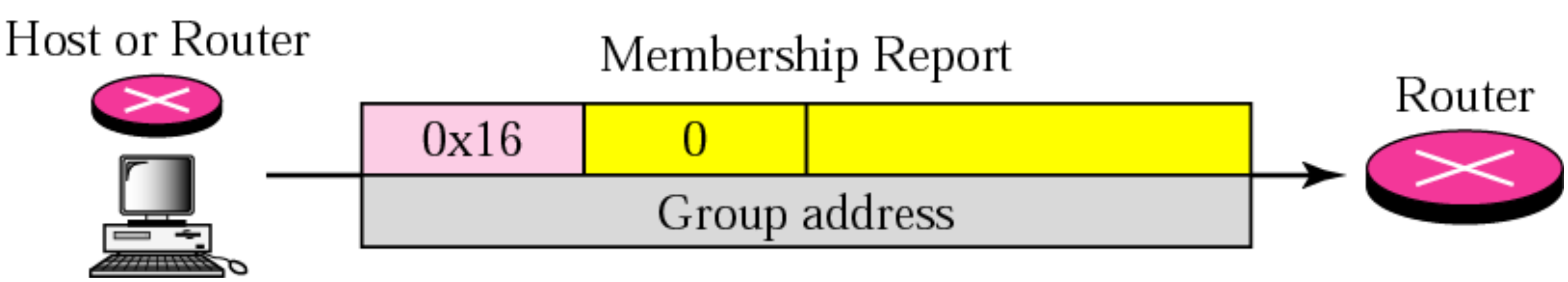



| Type | Value |
|---|---|
| Membership report | 0x16 |
| Leave report | 0x17 |
| General or special query | 0x11 |

Response time is in tenths of a second (0.1s).

# Joining a group

- Routers add the group address to its database of active groups if it is not already there
- Membership report is sent twice, one after another - for reliability

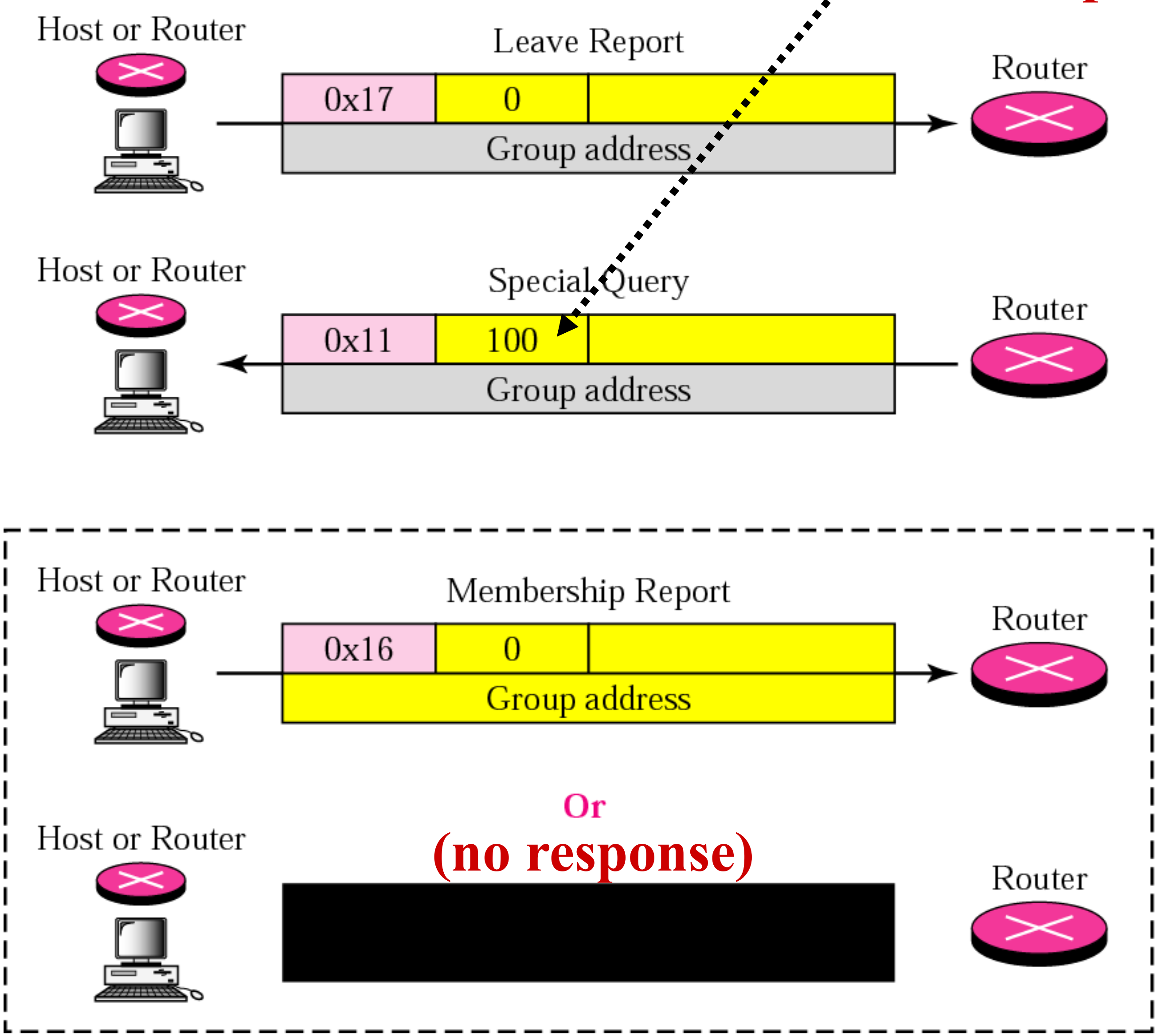


# Leaving a group

- If a host is no longer interested in a group, it sends a leave report message to the router.
- Question: If a router receives a leave report message for group G, should it remove G from its active list of multicast group?

**Answer: No, because other hosts in the network may still be interested.**

**Max. response time**




**Router sends special query to check whether there're still active group members**

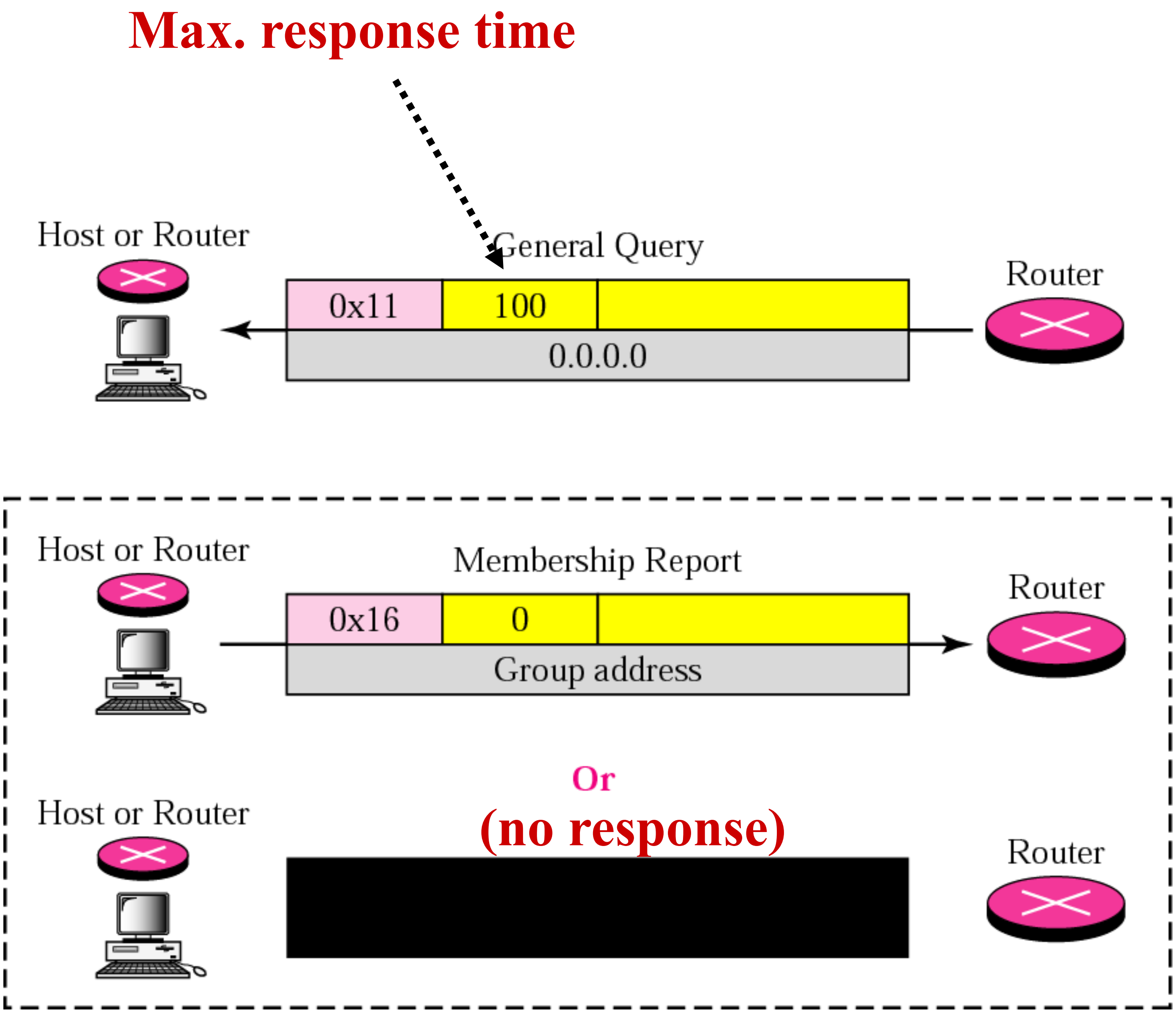
**Case 1: There're active members**

**Case 2: No active member.**

# IGMP: Membership Monitoring

- Imagine only one host in a group, and the host died
  - Router will never receive leave message!
- Router periodically sends general query message
- Router expects response for each group
- Hosts intentionally delay the response (random delay) to avoid unnecessary traffic
  - *A* distributed algorithm!

# The general query message does not define a particular group




**Case 1: There're active members**

**Case 2: No active member.**

# IGMP: Delayed Response (1)

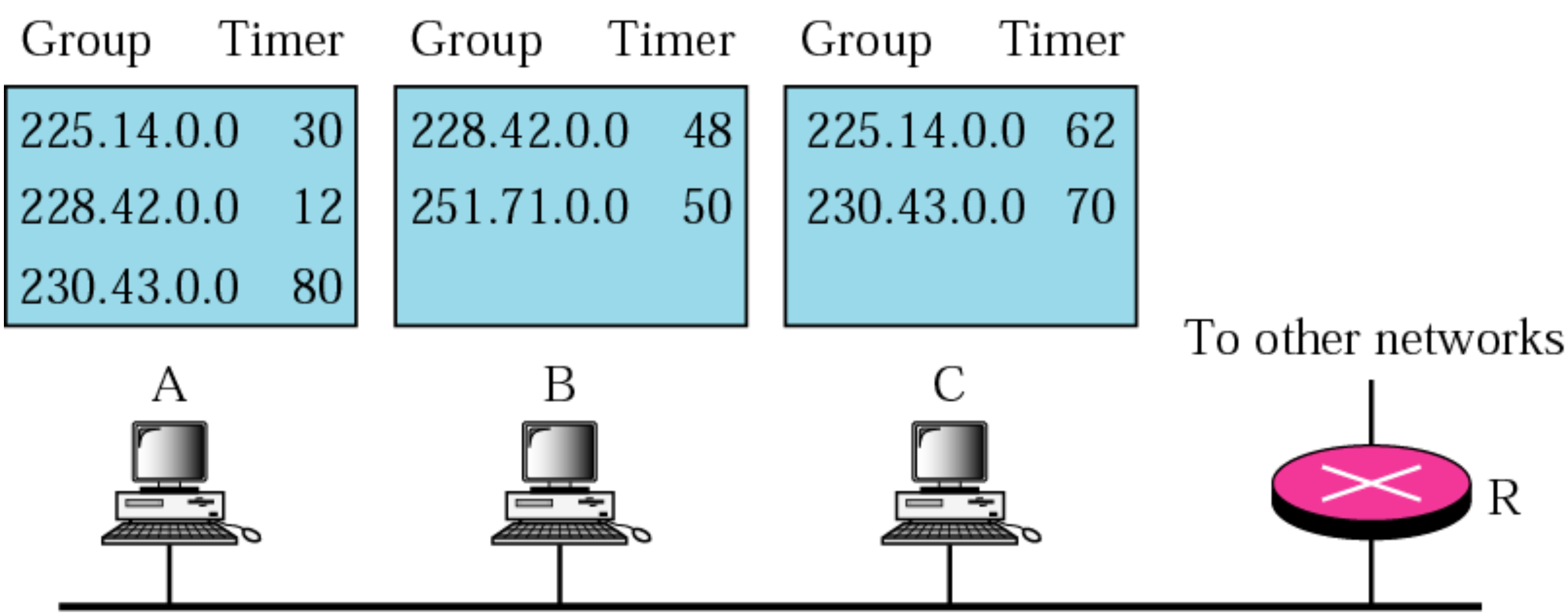
- A network may contain multiple members of a group
  - Too much traffic if all members of a group respond
- Solution: response is randomly delayed
- The host chooses a random response time for each group it belong to
  - The random time is between zero and maximum response time
  - maintains a timer for each group

# IGMP: Delayed Response (2)

- If a host belongs to a group G and sees a membership report for G
  - It no longer has to send a response
  - It cancels its timer
  - Note: The membership report for G is addressed to the group so all members will see it.

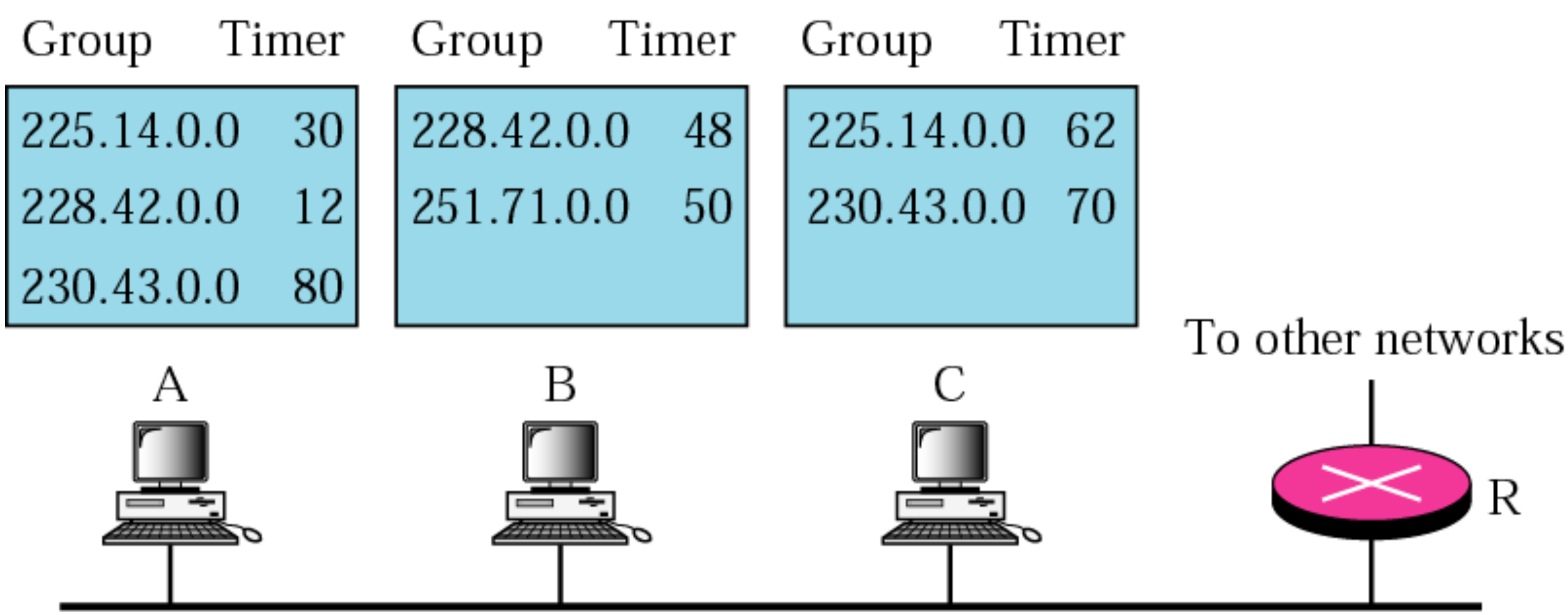
# Exercise

- Imagine 3 hosts as shown. A query message was received at time 0; the random timers (tenths of sec) are shown next to group ids. Show the sequence of report messages on the network.



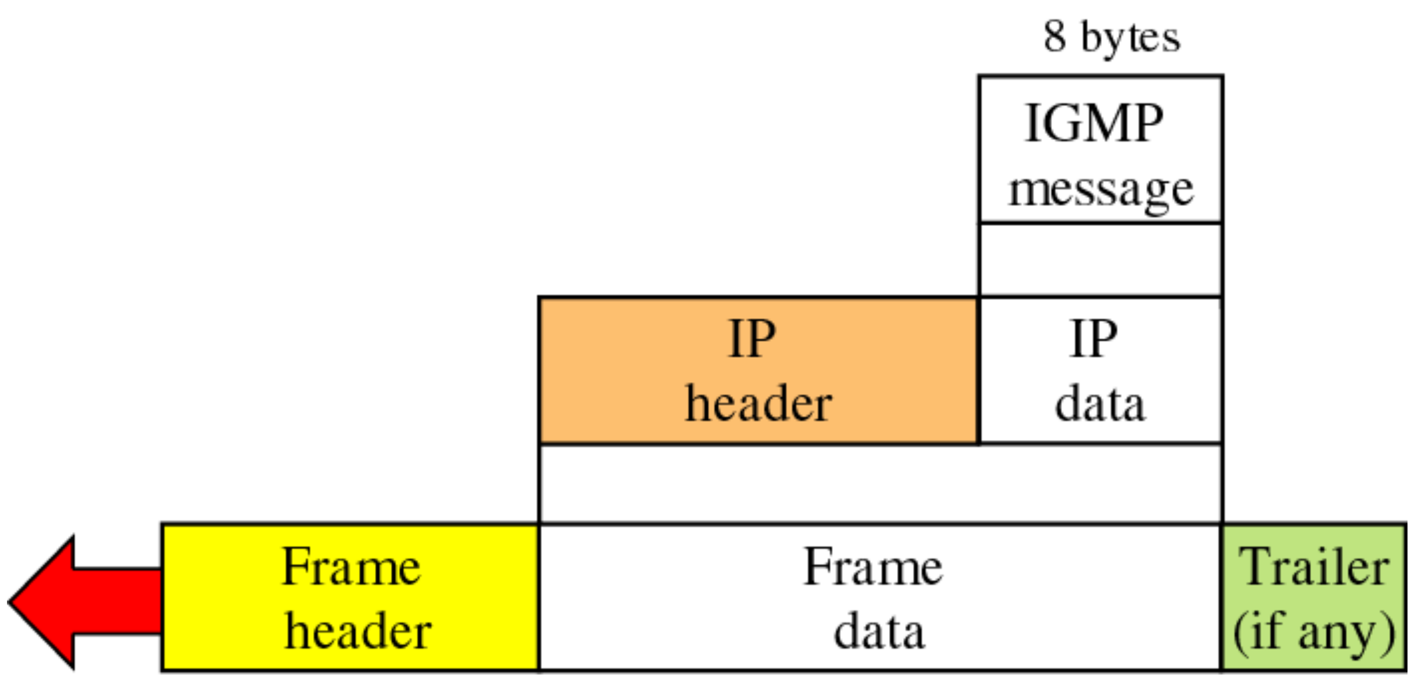
# Solution

- Only 4 reports (instead of 7 reports)
- Time 12 : A sends for 228.42.0.0, B cancels
- Time 30 : A sends for 225.140.0, C cancels
- Time 50 : B sends for 251.71.0.0
- Time 70 : C sends for 230.430.0.0, A cancels



# Encapsulation (1)

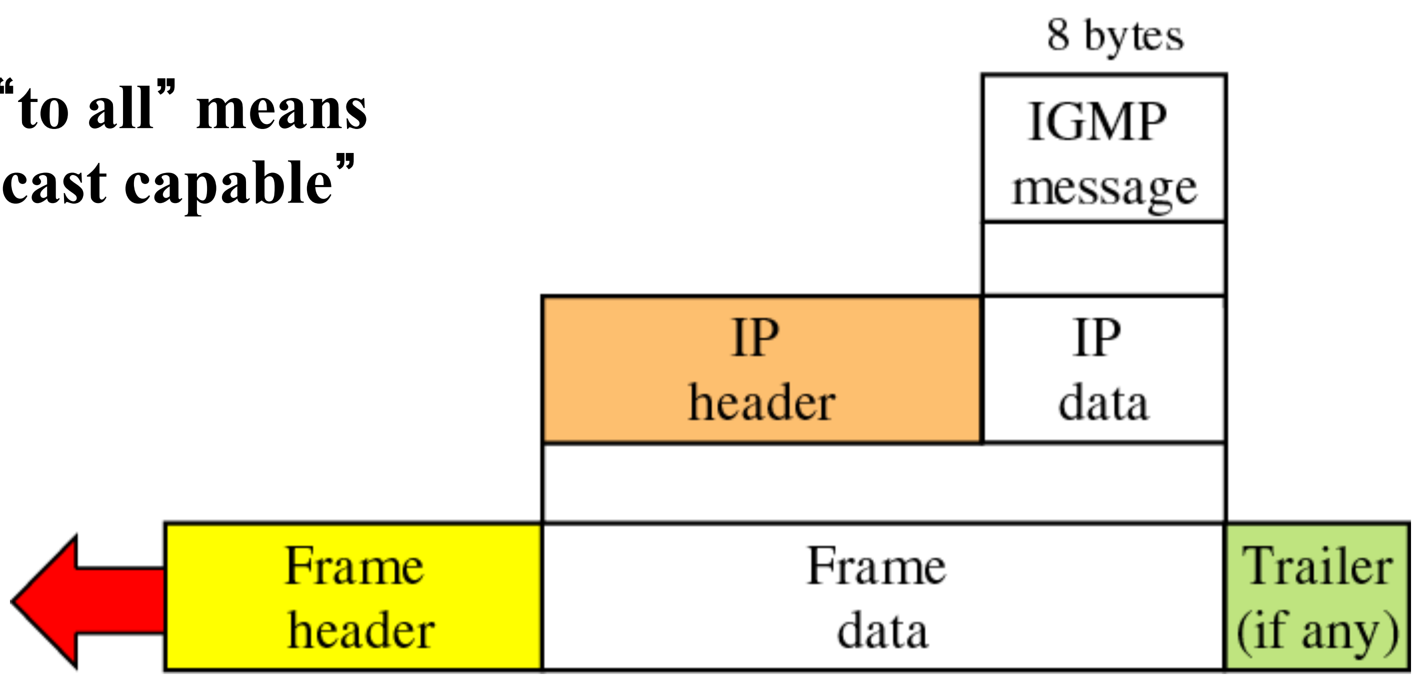
- IGMP messages are directly encapsulated in IP datagrams
- TTL=1 ⇒should not travel beyond LAN



# Encapsulation (2)

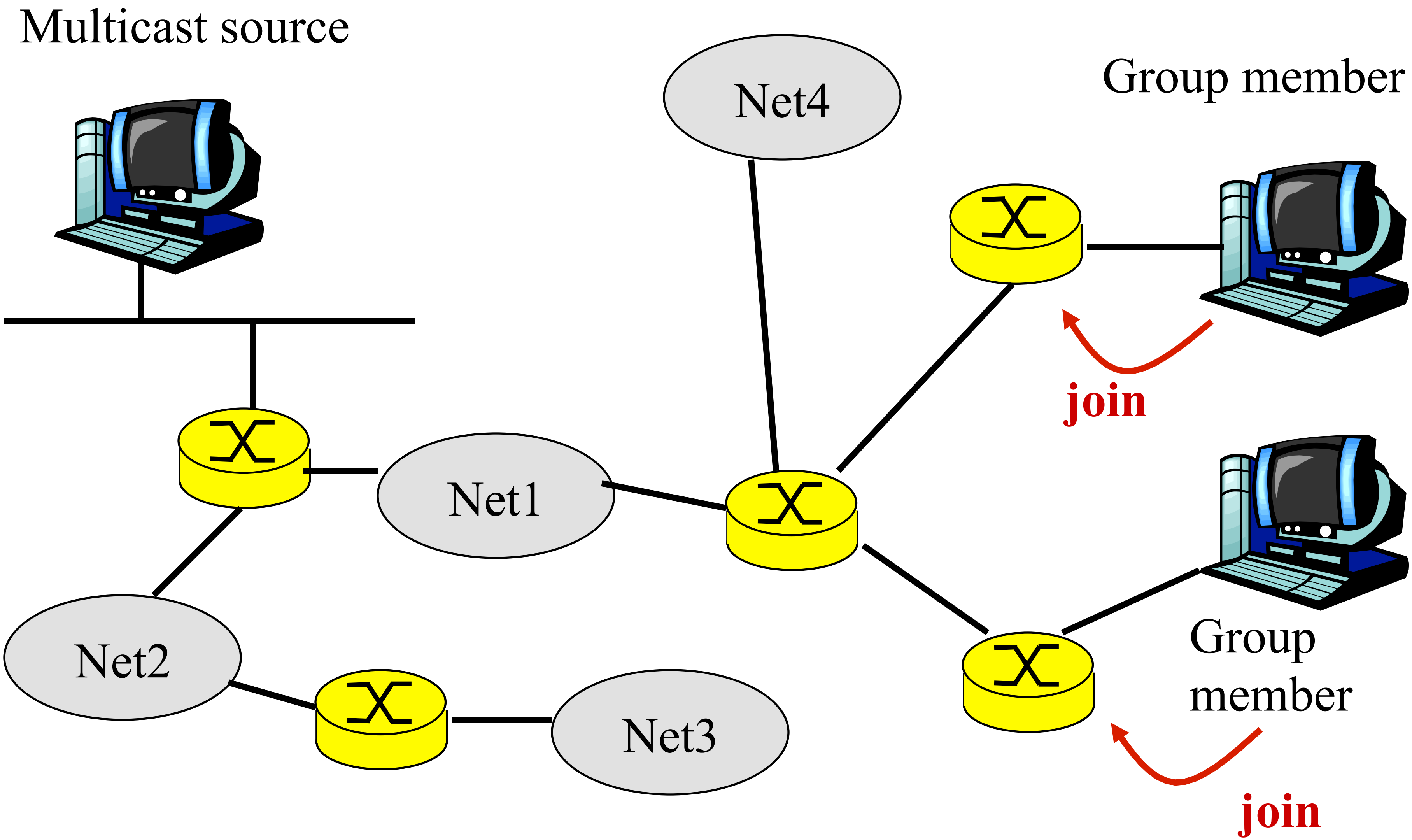
- IP destination address
  - query: to all hosts and routers (224.0.0.1)
  - leave: to all routers (224.0.0.2)
  - membership report: to the multicast group

**Note: Here "to all" means "to all multicast capable"**



# *Multicast Routing*

# Multicast routing problem
## IGMP provides only a local solution



# *Multicast routing problem*

- With IGMP, router knows existence of group members in *immediate* subnets
  - Solves local problem
- How can a router in the Internet know through which interfaces it should forward multicast packets so all members in the Internet receive it?
  - This knowledge is gained from multicast routing
  - Solves global problem

# Multicast Routing requirements

- Requirements
  - Every group member should receive only *one* copy of the multicast packet
  - Non-members must not receive a copy
  - No routing loops
  - Either one of the following [next slide]
    - » Least cost path to all members
    - » Least cost multicast tree
      - Tree cost = sum of all link costs in the tree

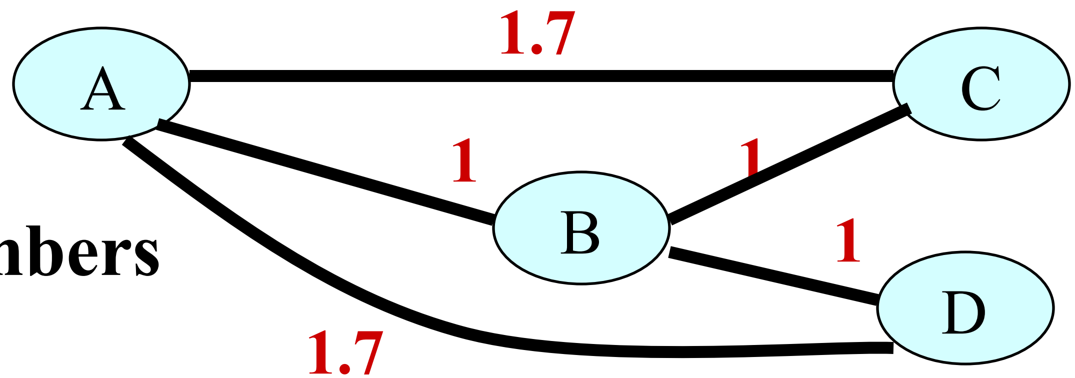
# *Shortest path tree versus Steiner tree*

**Example network:**

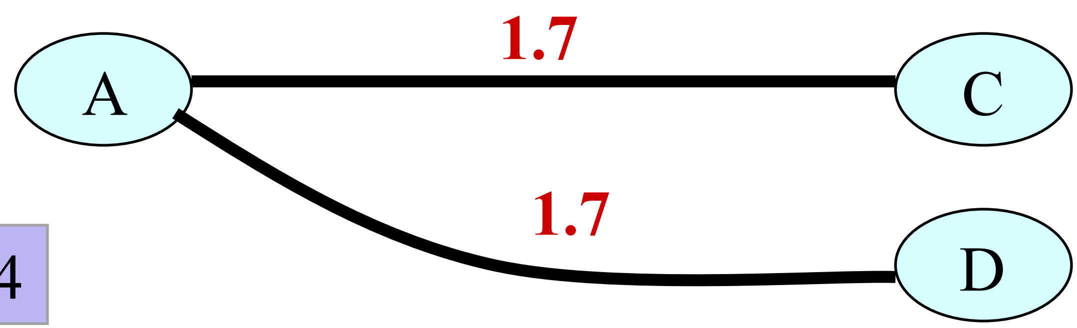
**A = source**

**C, D = multicast members**

**Link cost in red**




**Shortest path tree**

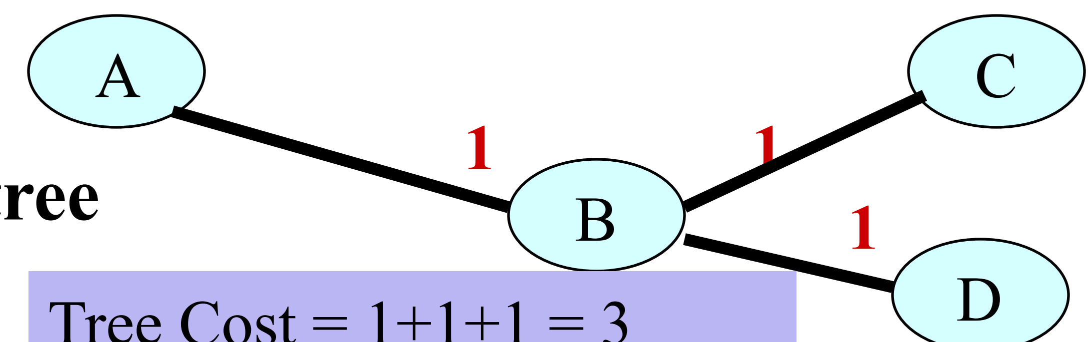



Tree Cost = 1.7+1.7 = 3.4

**Steiner tree =**

**Least cost multicast tree**

**NP-hard to compute**




Tree Cost = 1+1+1 = 3

# Multicast Routing Algorithms

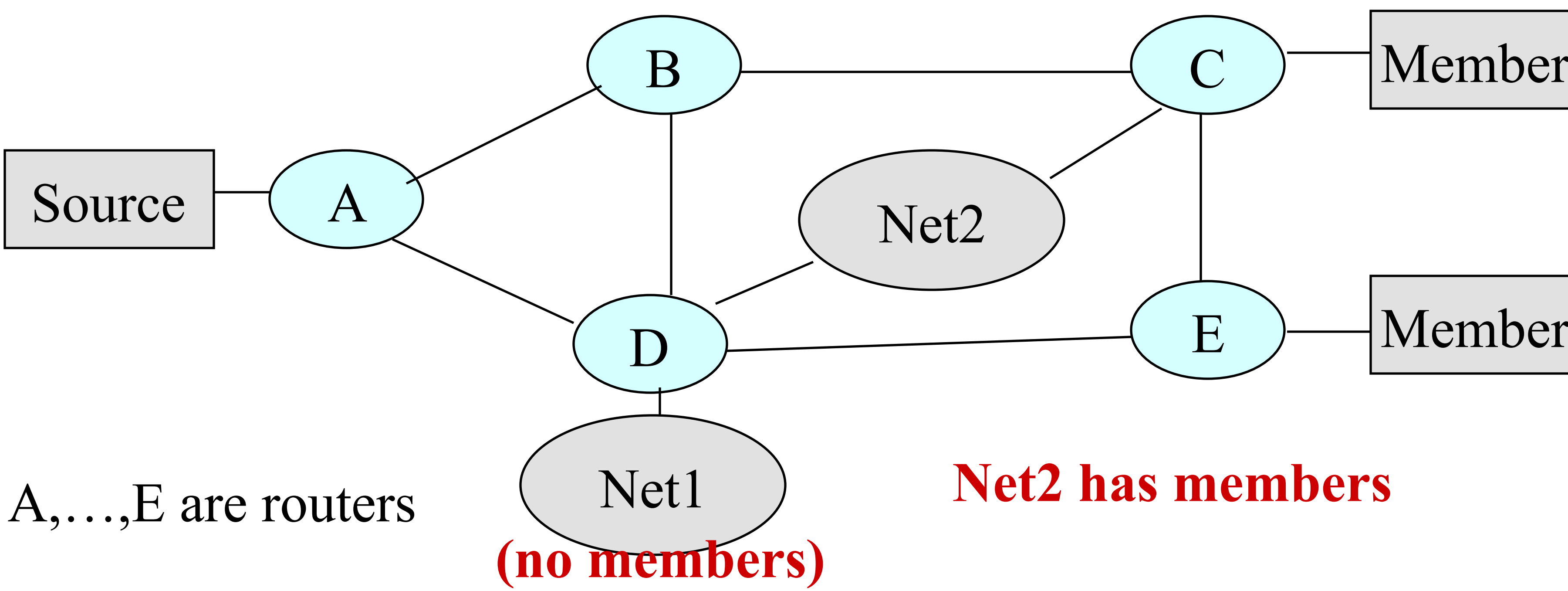
- Multicast routing can be
  - Intra-AS i.e. IGP
  - Inter-AS i.e. EGP
- We will only focus on IGP multicast
- Two classes of algorithms
  - Source based tree
  - Group shared tree
- To motivate source based algorithm, we begin with flooding

# Flooding

- A router forwards a packet out of all of its interfaces except the one from which the packet arrives
  - We will use this definition of flooding in this lecture
  - An improved version checks for duplicates and discard them
    - » Improved version is used in OSPF for LSA updates
    - » Duplicate detection is difficult in general purpose multicast

# Exercise

- Assume flooding is used in the following network, does it meet the following requirements
  - Each member receives only one packet
  - Non-members do not receive any copy
  - No routing loops

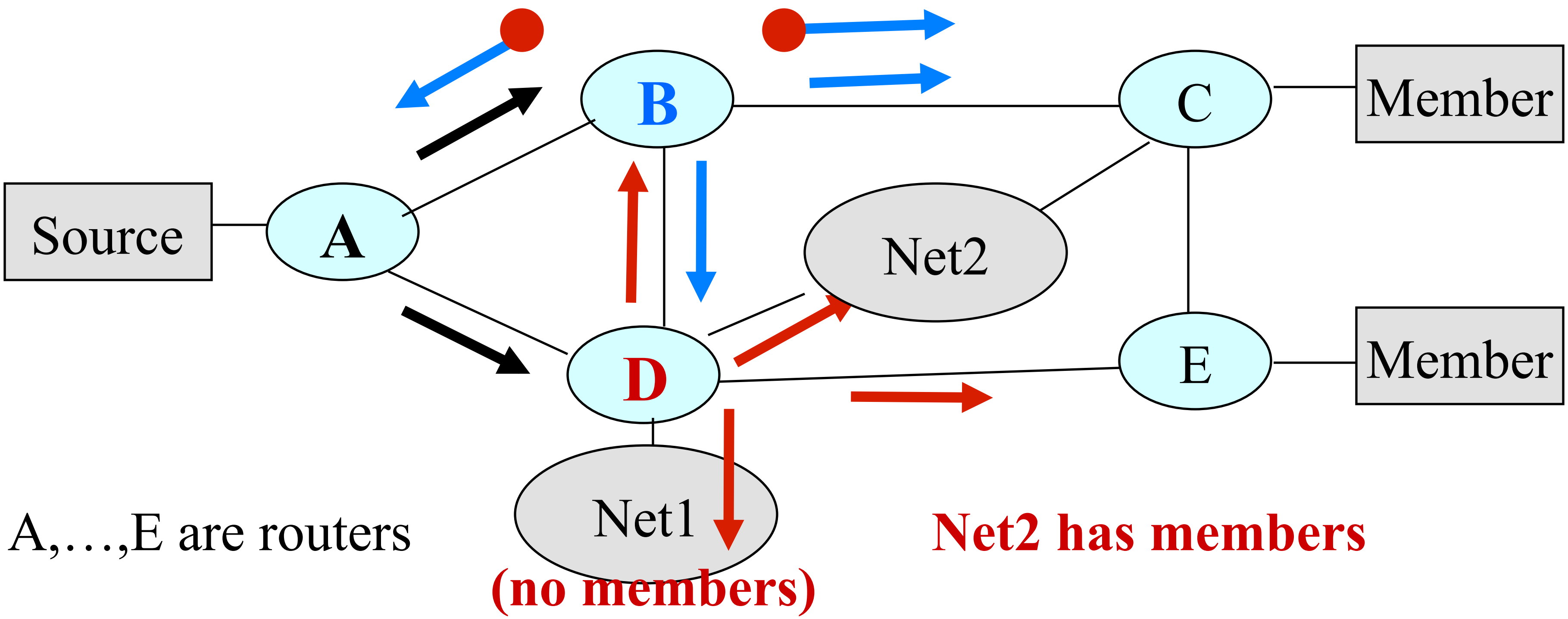


# Solution

Each member receives only one packet? NO

Non-members do not receive any copy? NO

Can you identify any loops? YES



# Flooding as multicast

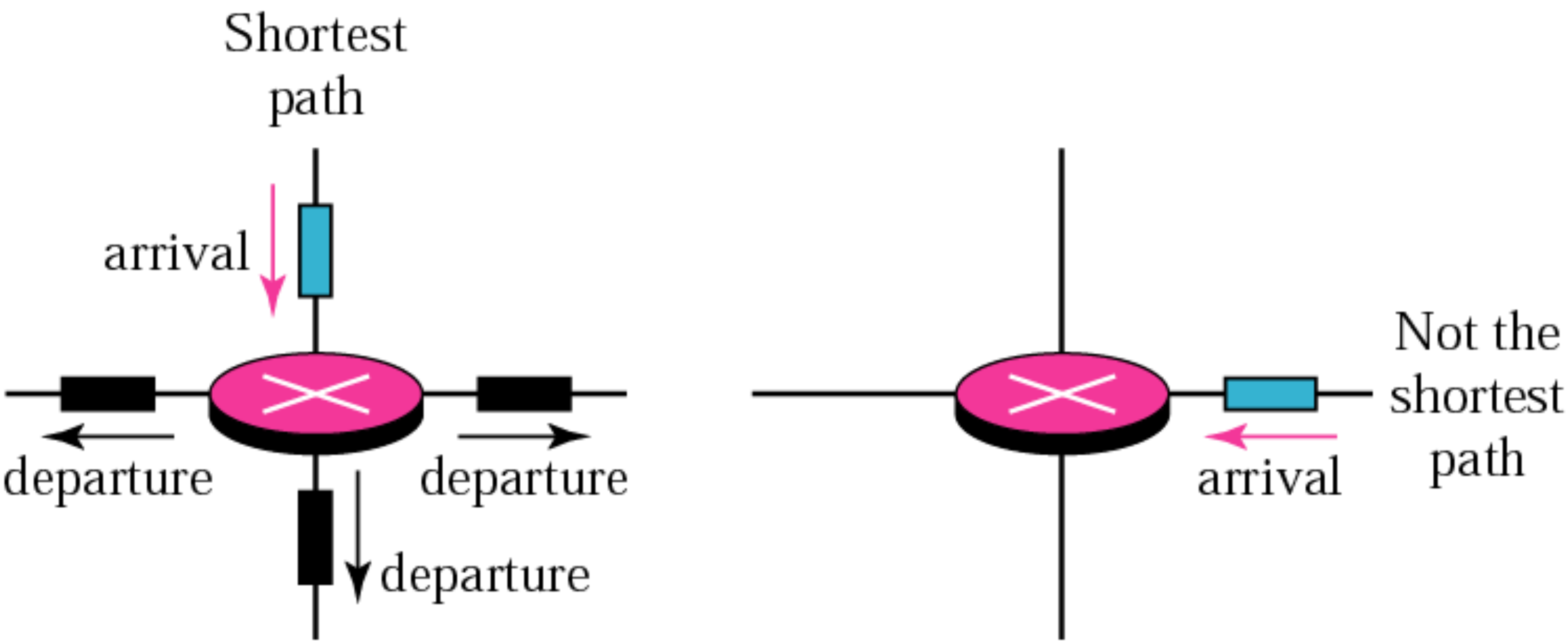
- In flooding, all routers and networks receive at least one copy of the packet
- Pros:
  - Simple to implement
  - Router do not need to maintain group information
- Cons:
  - Duplicate and unnecessary packets
  - Routing loops
- Tradeoff between simplicity and resource usage

# *How to improve flooding?*

- Too many redundant packets
- In flooding with *duplicate detection*, how many times does a node forward a packet?
- We can reduce the number of redundant packets by making sure that each node forwards a packet only once

# Reverse path forwarding (RPF)

- Each node forwards a multicast packet only once
- Forwarding algorithm
  - Given:
    - » Router R has interfaces if-1, if-2, …, if-n
    - » Router R uses interface if-k to reach the node S
      - I.e. Interface if-k is on the shortest path from R to S
  - When a multicast packet whose source address is S reaches router R
    - » If the packet arrives at interface if-k, forward the packets to other interfaces
    - » Otherwise, discard it.




a. Packet is forwarded

b. Packet is discarded

# Exercise

- A multicast router receives a packet with source address 195.34.23.7 and destination address 227.45.9.5 from interface 2. Should the router discard or forward the packet based on the following unicast routing table?
  - *Destination* *Interface*
  - 121.0.0.0 1
  - 185.67.0.0 2
  - 195.34.0.0 3

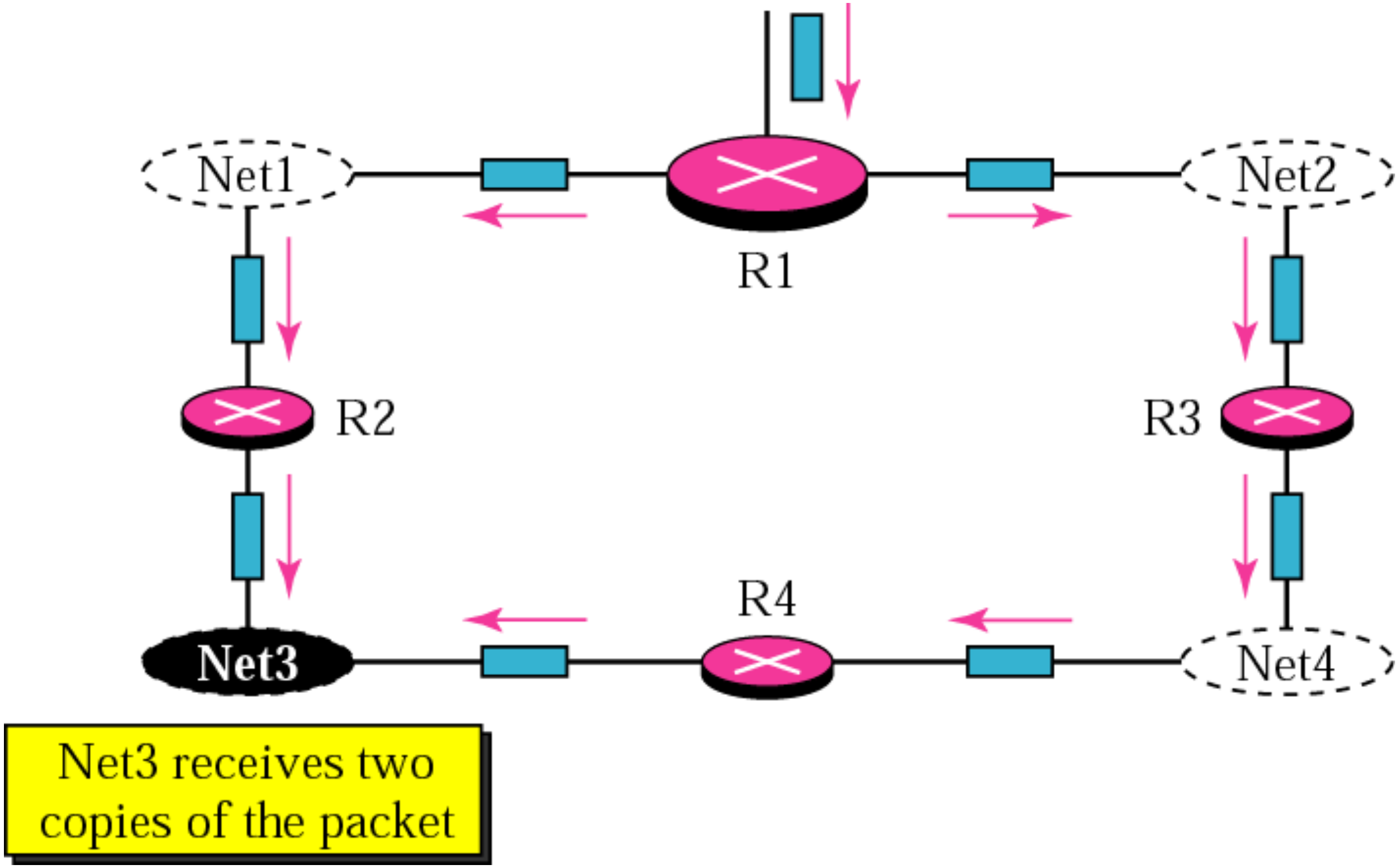
# Solution

- Discard the packet because it has not arrived through the shortest path interface
- It would have been accepted if it has arrived through interface 3
- Note: RPF routing decision depends on both source and destination addresses

# Does RPF meet our multicast routing requirements?

- Requirements
  - Every group member should receive only one copy of the multicast packet NO [next slide]
  - Non-members must not receive a copy NO
  - No routing loops YES

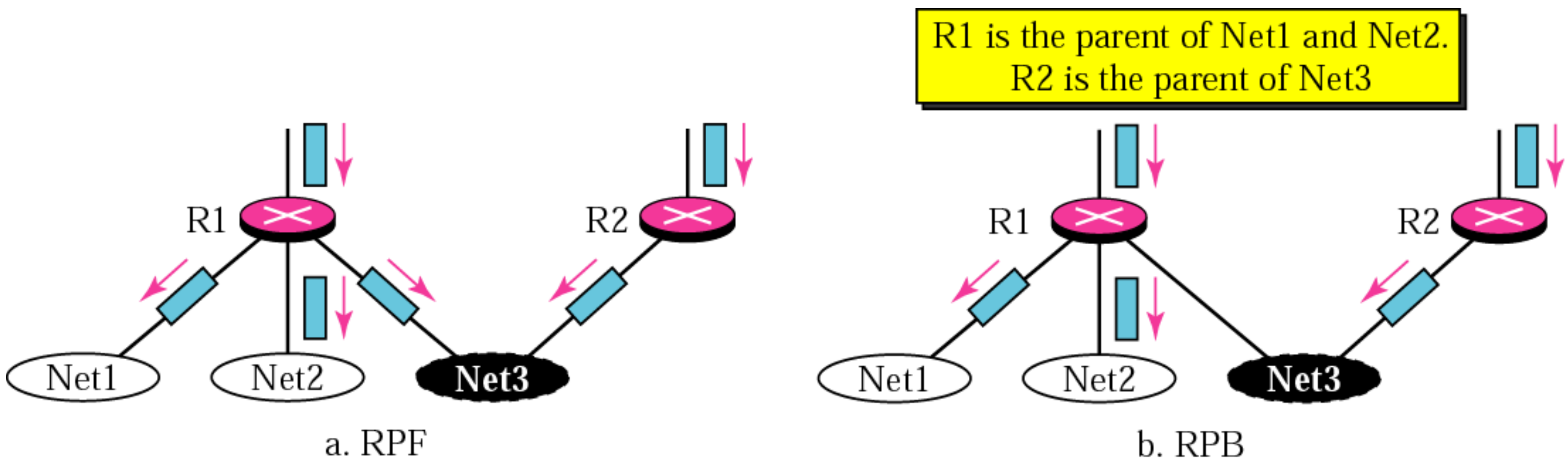
# Problem with RPF: Duplicate packets



# Reverse path broadcasting (RPB)

- An improvement to RPF
- RPB ensures that each network receives only one copy of the multicast message
- If a network has multiple routers
  - A designated router is chosen
  - Only the designated router forwards multicast packets into the network

# RPF versus RPB



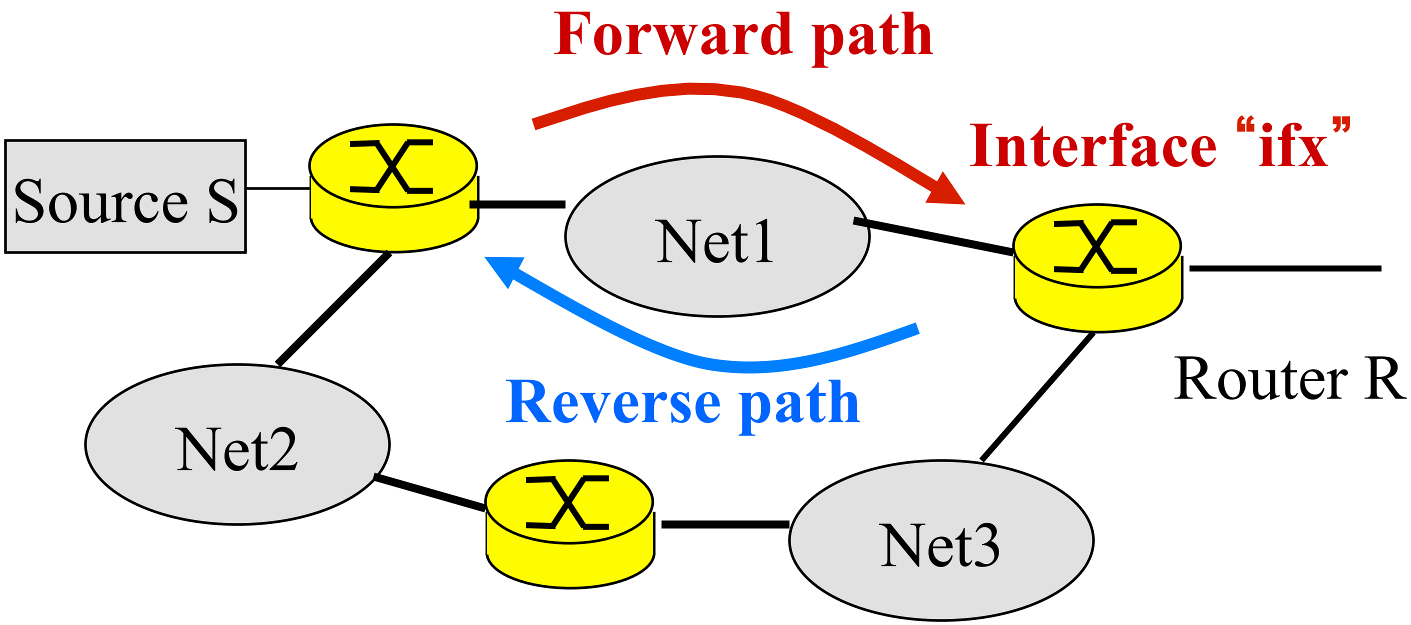

R2 is the designated router for Net3

# Choice of designated router

- Given
  - Multicast source S
  - Routers R1, R2, …, Rn attached to the same network
  - Distance(Ri,S) = routing cost from Ri to S
- Rx is the designated router for multicast packets from S if Distance(Rx,S) is the smallest
- Question: In the example on the previous page, how can R1 know that it shouldn't send to Net3?

# Does RPB meet our multicast routing requirements?

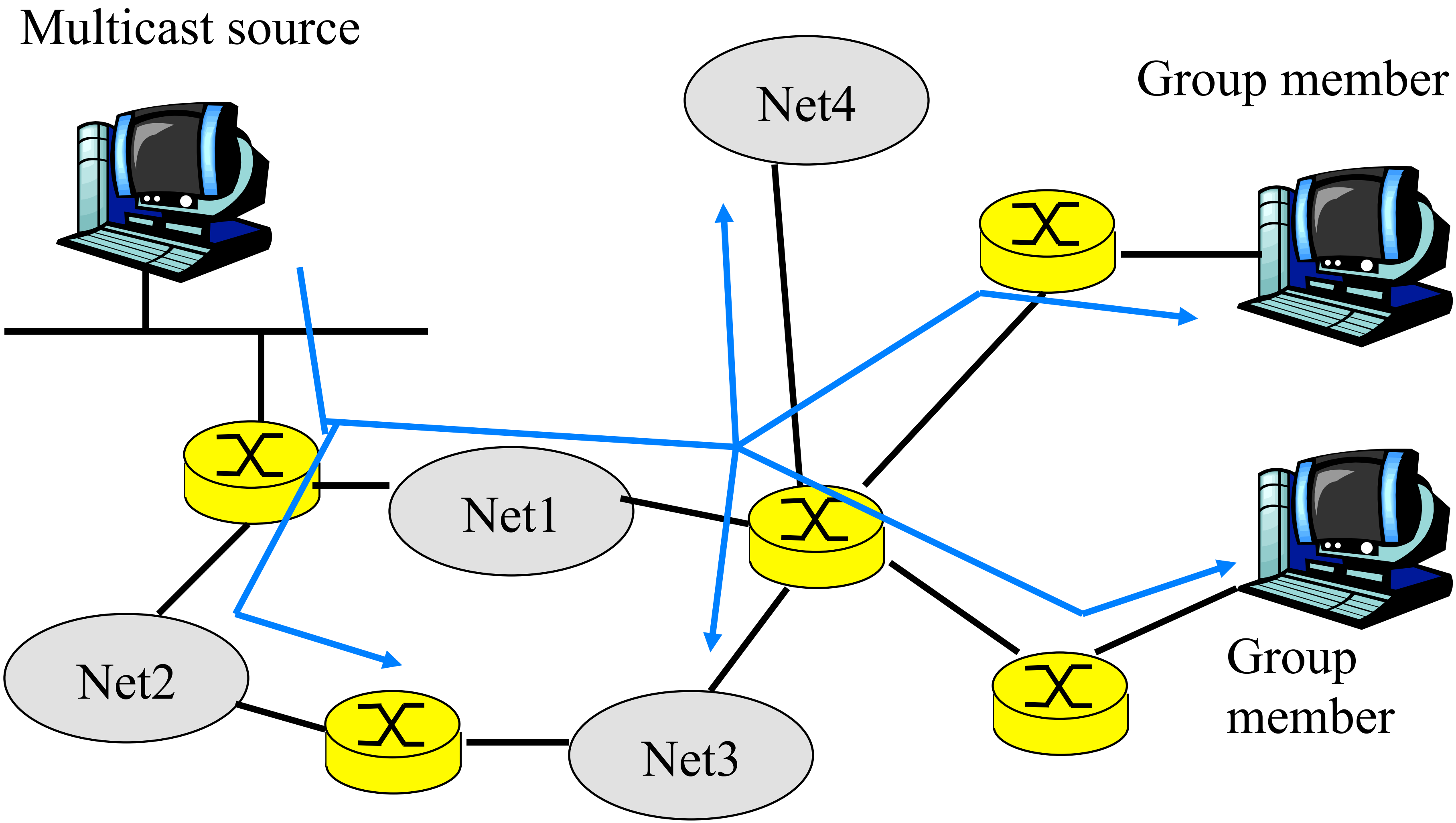
- Requirements
  - Every group member should receive only one copy of the multicast packet YES
    - RPB gives a broadcast tree [The slide after]
  - Non-members must not receive a copy NO
  - No routing loops YES



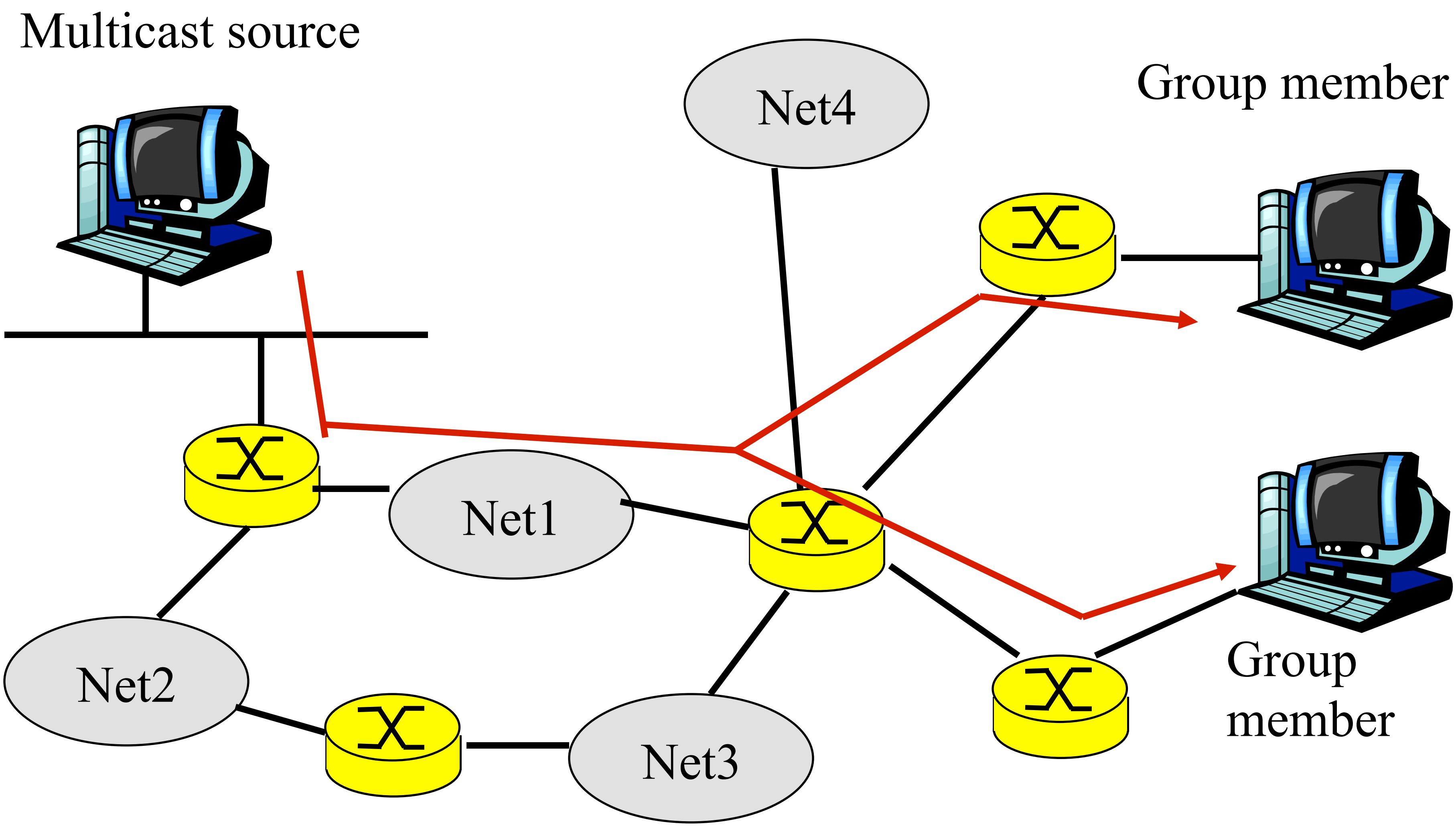

**Given: (1) S = source of the multicast**
**(2) The shortest path from R to S uses interface ifx.**

**If routes are symmetrical,**
**Shortest path from S to R (forward path)**
**= shortest path from R to S (reverse path)**

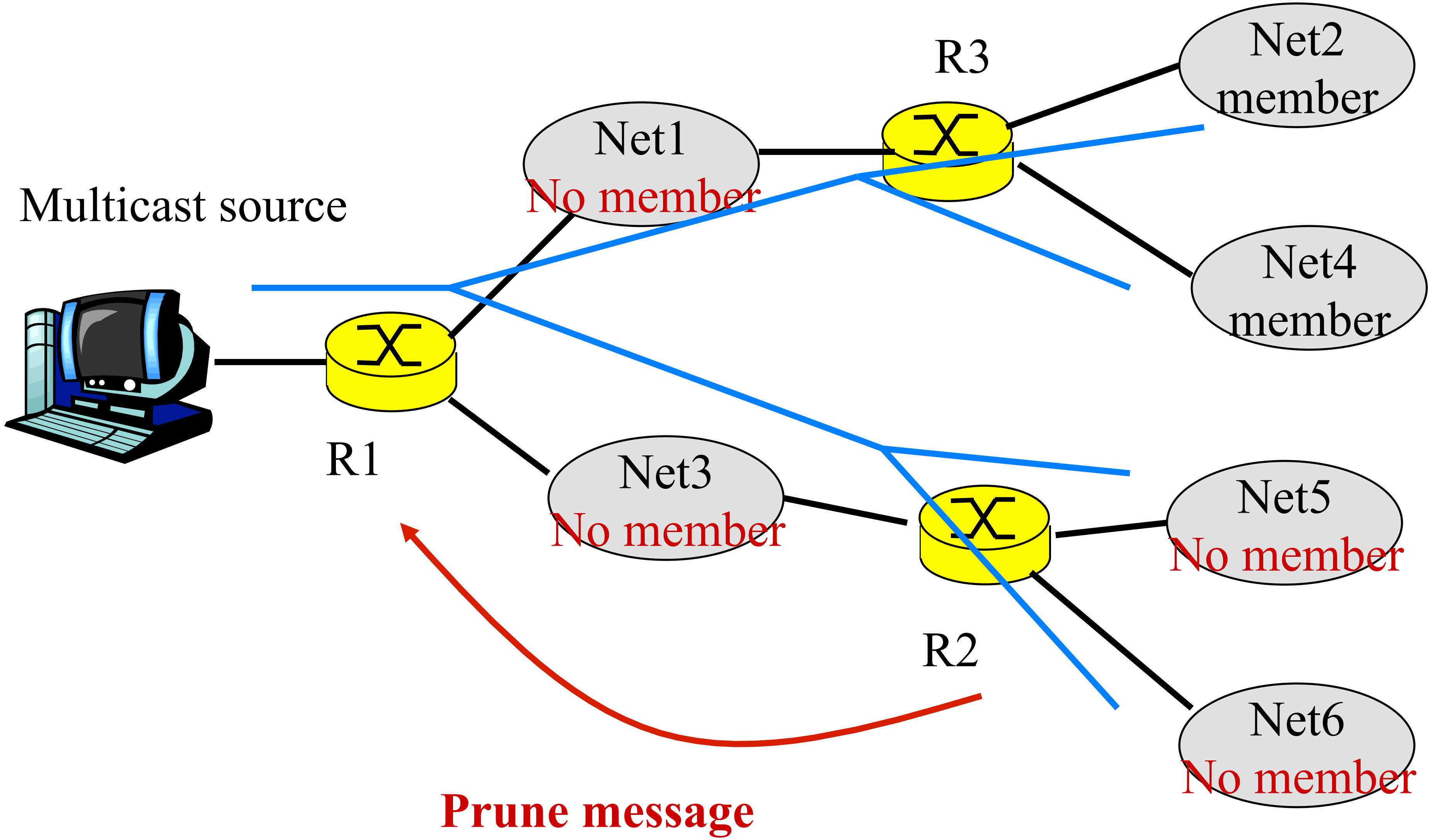
# RPB gives a *broadcast tree* = it delivers to all networks



# but what we need is a *multicast tree*



# Pruning of broadcast tree



# *Pruning and grafting*

- Downstream routers send “prune” messages to upstream routers when no group members exist downstream
- Prune messages help upstream routers to “cross-out” specific interfaces to form a multicast tree
- If new members join on a pruned branch, downstream routers send “graft” message to re-establish a branch ==> *tree is dynamic*

# Reverse path multicasting (RPM)

- RPM (Reverse path multicasting)
  - First builds a per-source broadcast tree (RPB)
    - » The first packet sent by the source creates this tree
    - » Note: The tree is rooted at the source
  - Then *prunes* the broadcast tree to a multicast tree
- Note that for a given multicast group, the multicast trees for 2 different sources can be different.

# Does RPM meet our multicast routing requirements?

- Requirements
  - Every group member should receive only one copy of the multicast packet YES
  - Non-members must not receive a copy YES
  - No routing loops YES

# Distance vector multicast routing protocol (DVMRP)

- IETF standard for IGP multicast routing
- Based on RPM
- DVMRP is a separate routing protocol
  - The routers run either RIP or OSPF for unicast routing
  - DVMRP builds its own routing table for it to make multicast routing decision. It doesn't use the unicast routing table from RIP nor OSPF.

# Multicast forwarding table

- Each row of the multicast routing table has the following entries
  - Source address
  - Multicast group
  - The interfaces over which the node must forward the packets
- Need to match both source and multicast addresses

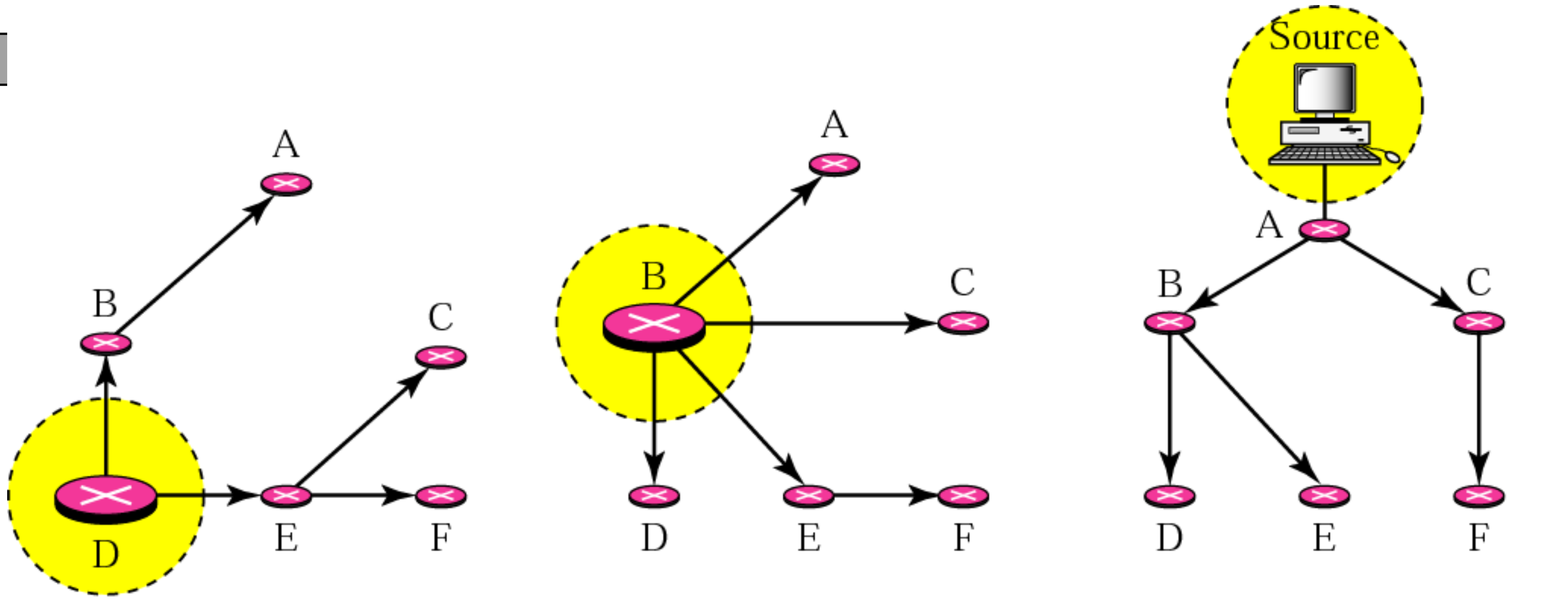
# Correction for Forouzan 3rd Ed. (pages 442-444)

- Figs 15.5, 15.6: Right hand columns in the multicast forwarding tables should have interface numbers instead of next-hops (next hops are used in unicast forwarding tables)
- Top textbox in Page 444: all routers ARE involved in multicast routing in group-shared tree approach

# Multicast OSPF (MOSPF)

- An extension to OSPF, thus an IGP
- Recall in OSPF
  - Each router has the complete network topology
  - Each router finds the shortest path tree rooted at itself

# Unicast tree and multicast tree




a. Unicast tree for D

b. Unicast tree for B

c. Multicast tree for all routers

**Problem:**

**Shortest path tree rooted at D different from Shortest path tree rooted at B**

**Can't use the unicast tree as multicast tree**

# MOSPF (cont'd)

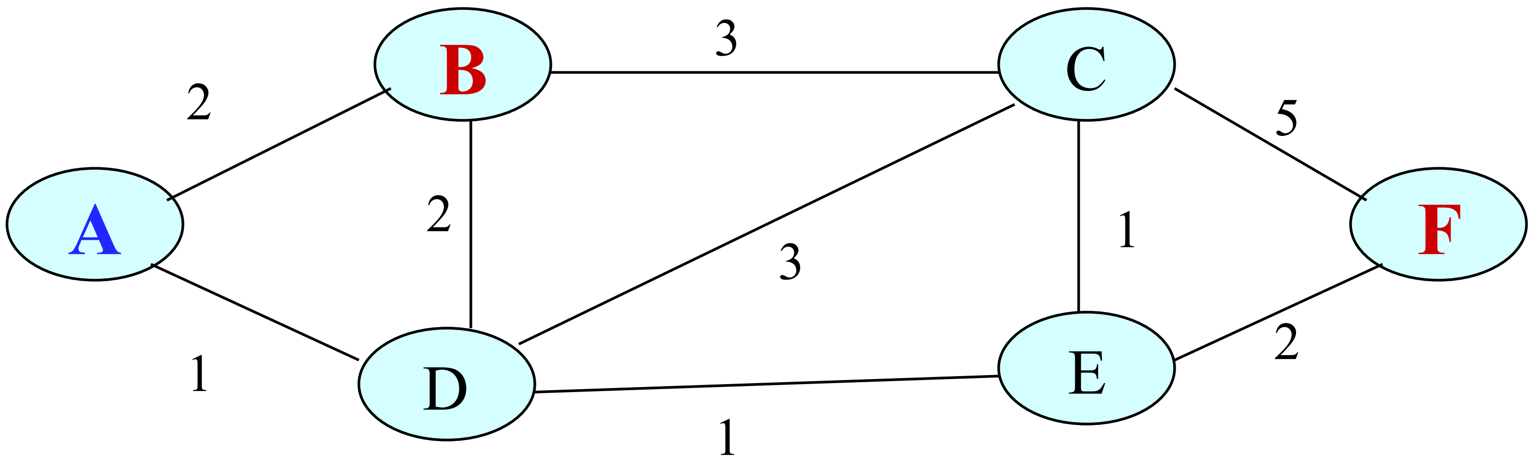
- Since each router has the complete network topology, they can all compute the shortest path tree rooted at the source
  - All routers will have the same tree
  - Each router can compute a tree for each (source, group)
- However, the shortest path tree reaches all network

# MOSPF (cont'd)

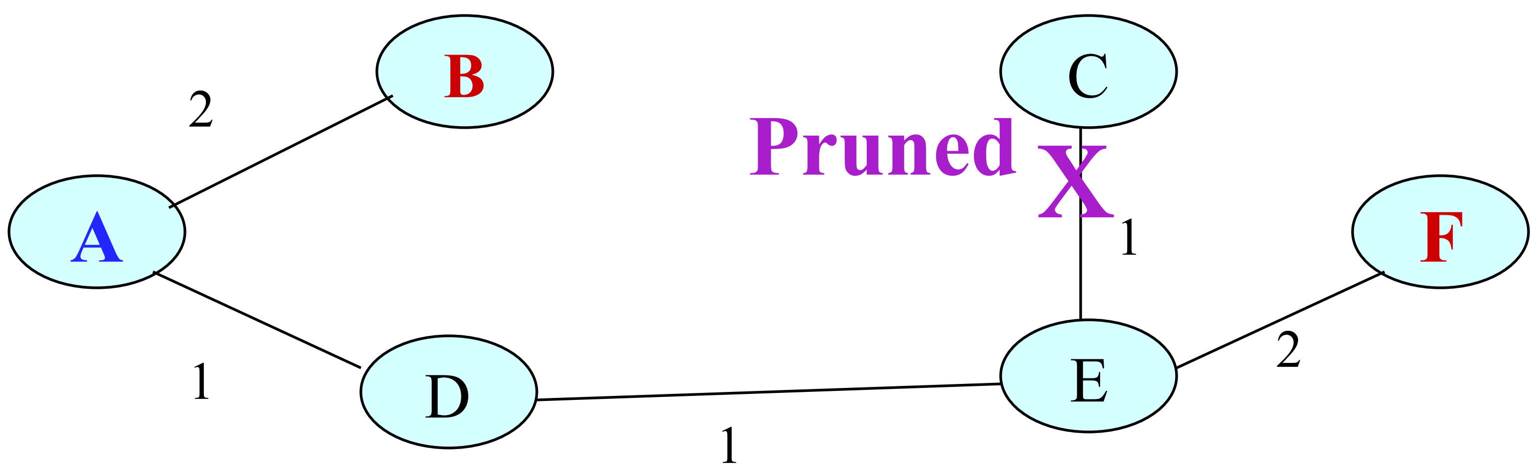
- Group membership LSA
  - A router knows where the group members are
- A router prunes the broadcast tree to obtain the multicast tree
- Both DVMRP and MOSPF form source based trees, each (source,multicast group) has its own tree.

# *MOSPF - example*

**A = source. B & F are in the multicast group**



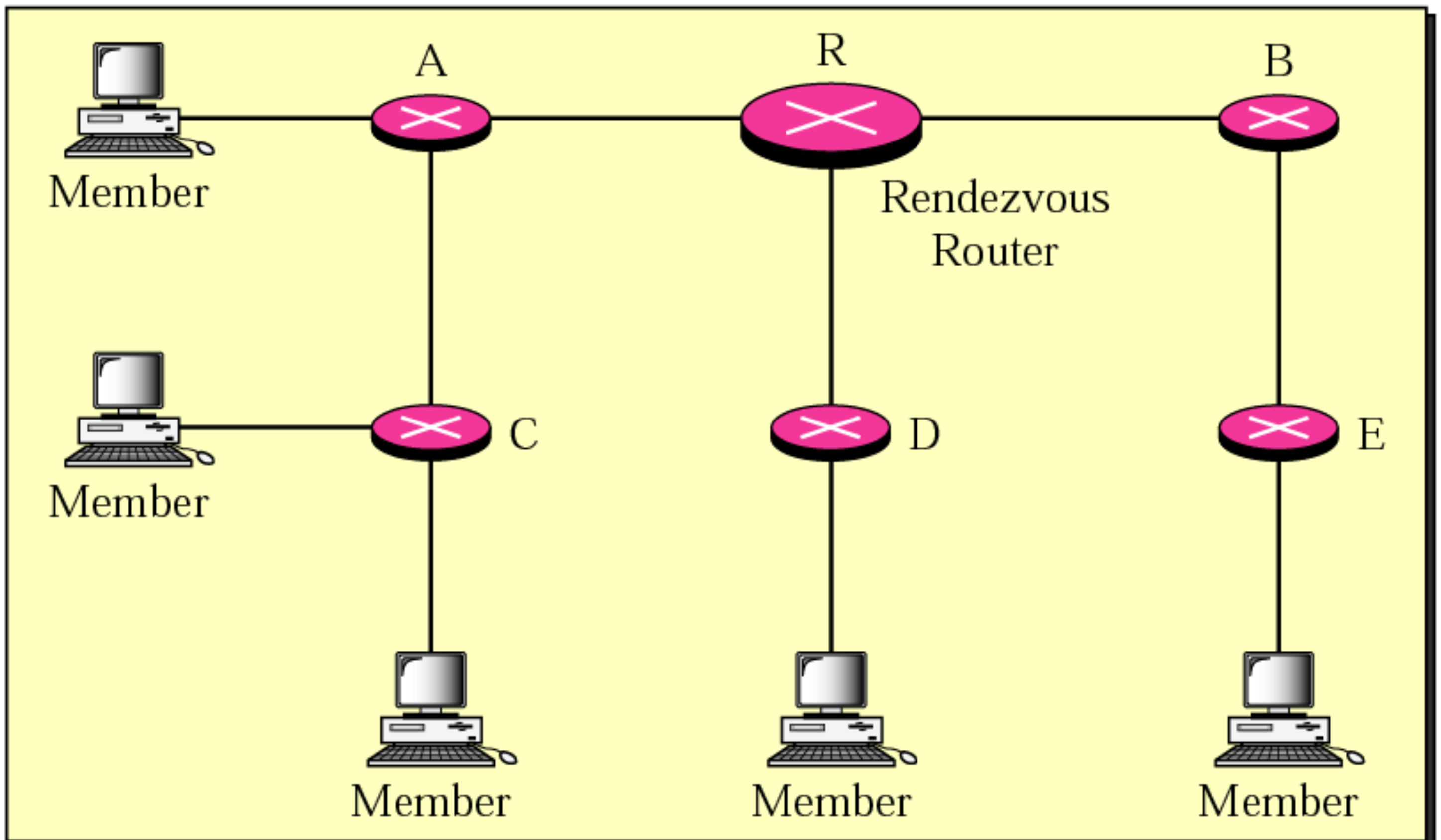

All routers compute the shortest path tree rooted at source node A:



# Core Based Tree (CBT)

- One multicast tree per group
- The root of the multicast tree is called rendezvous router (RR)

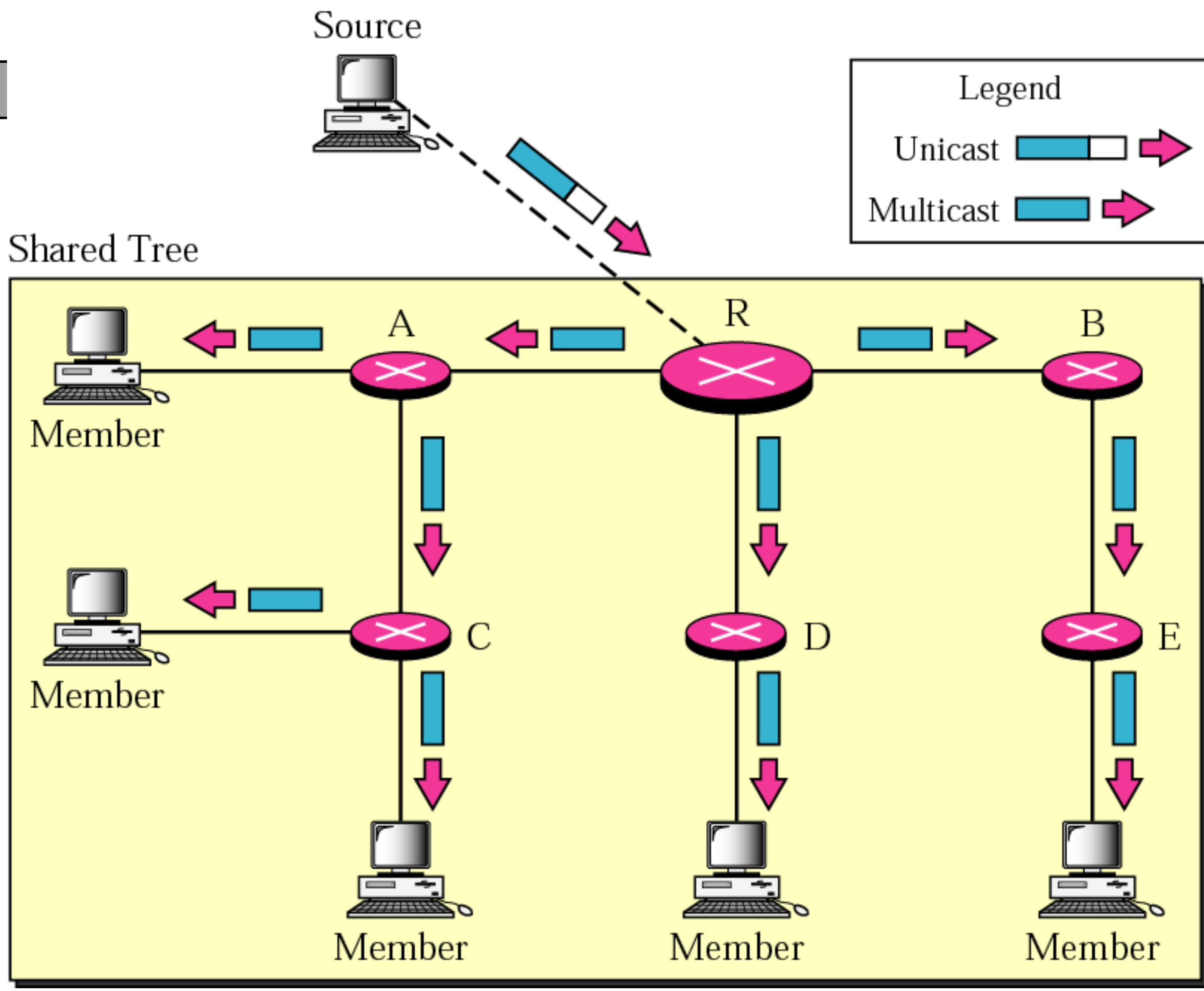
Shared Tree



# Sending a Multicast Packet in CBT

- Source encapsulates multicast packet into unicast packet and sends it to RR
- RR decapsulates and sends the multicast packet to relevant interfaces
- Downstream routers do the same, i.e.,forward the multicast packet to relevant interfaces

# Sending a multicast packet to the rendezvous router



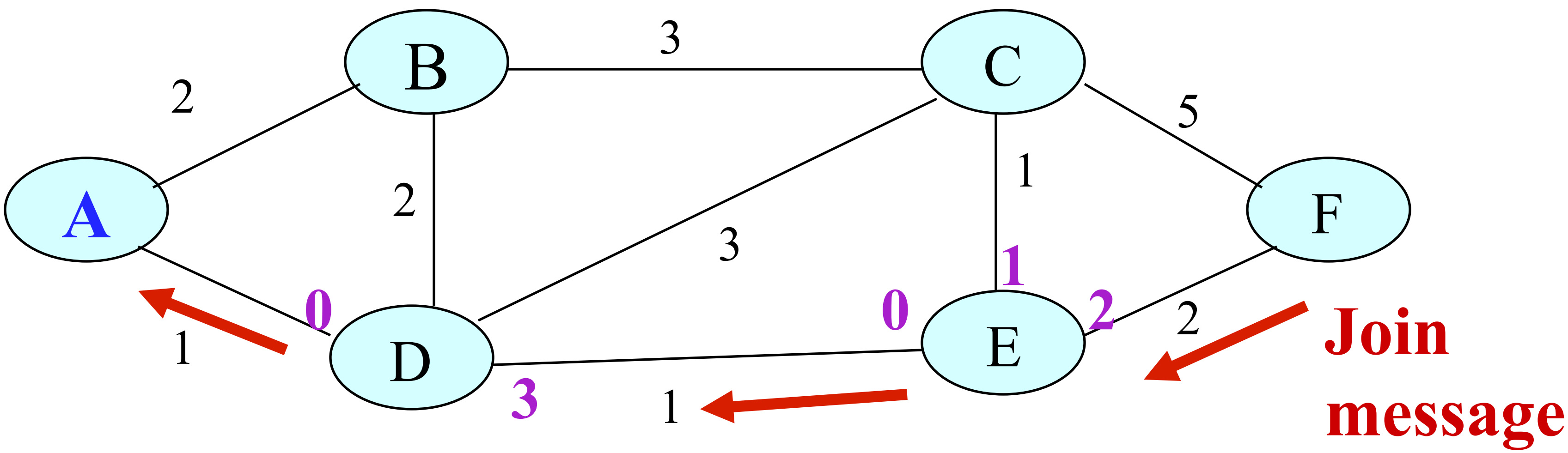
# Tree Formation in CBT

- First RR is selected
- Every router is notified of the unicast address of the RR
- A router wishing to join the group, sends a unicast join message to the RR
- Each intermediate router learns upstream and downstream router from the join message
- Tree is now formed

# CBT - example

Note: Link cost in black.
Router interface number in magenta

A is the Rendezvous router.
F is the first member to join the multicast group G.




**Multicast forwarding table for D**

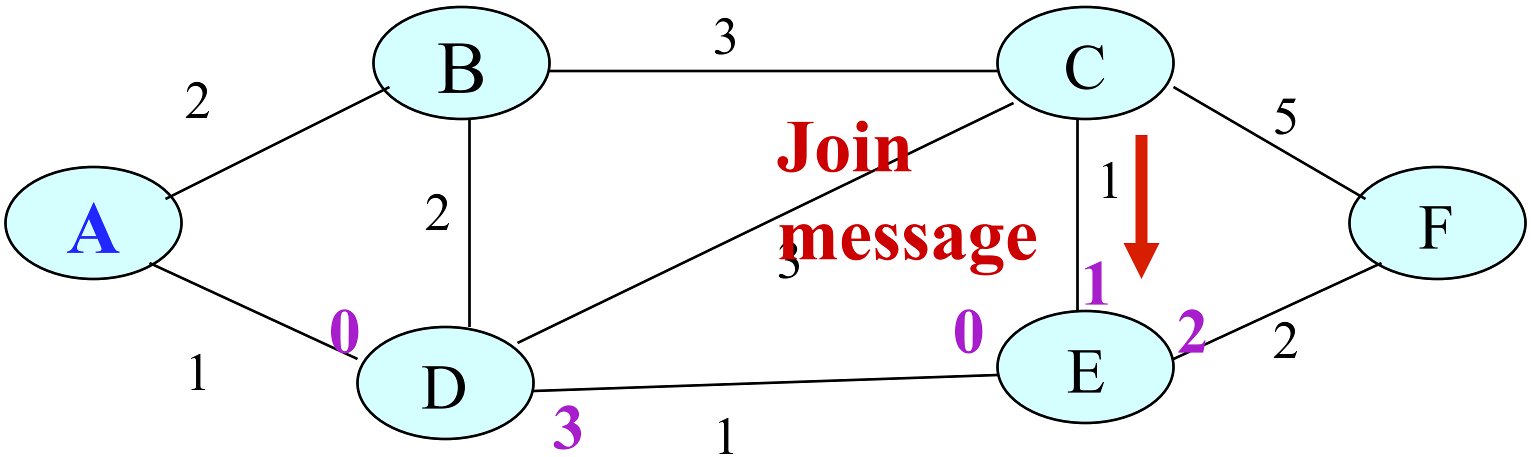
| Group | Interfaces |
| --- | --- |
| G | 3 |

**Multicast forwarding table for E**

| Group | Interfaces |
| --- | --- |
| G | 2 |

# CBT - example (cont'd)

**C is the second member to join the multicast group G.**




**Multicast forwarding table for D**

| Group | Interfaces |
|---|---|
| G | 3 |

**Multicast forwarding table for E**

| Group | Interfaces |
|---|---|
| G | 2,1 |

# Leaving a group in CBT

- A router sends a leave message to its upstream router
- The link is pruned at the upstream router if there are no other downstream routers
- The states in the multicast forwarding table are soft. They will expire if not refreshed.

# DVMRP versus CBT

- DVMRP
  - one tree per (source, multicast group) pair
  - Tree built from root
  - Build a broadcast tree which covers all networks and then prune broadcast to multicast
  - Any node can release a multicast packet to the network
  - An example of "source based tree" method
- CBT
  - one tree per group
  - Tree built from leaves
  - Add a branch only when it is required
  - Only a designated router can generate a multicast packet to the network
  - An example of the "group shared tree" method

# Protocol independent multicast (PIM)

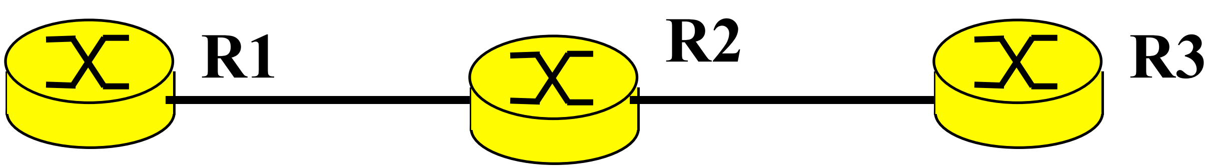
- Two protocols
  - Dense mode (PIM-DM)
  - Sparse mode (PIM-SM)
- PIM-DM
  - Source based tree similar to DVMRP
  - Works with the unicast routing protocol found in the AS
  - Works best when there are many receivers

# PIM (cont'd)

- PIM-SM
  - Similar to core based tree
  - Works best when there are few receivers
  - If there is a lot of traffic, the root of the tree can move from the rendezvous point to the source

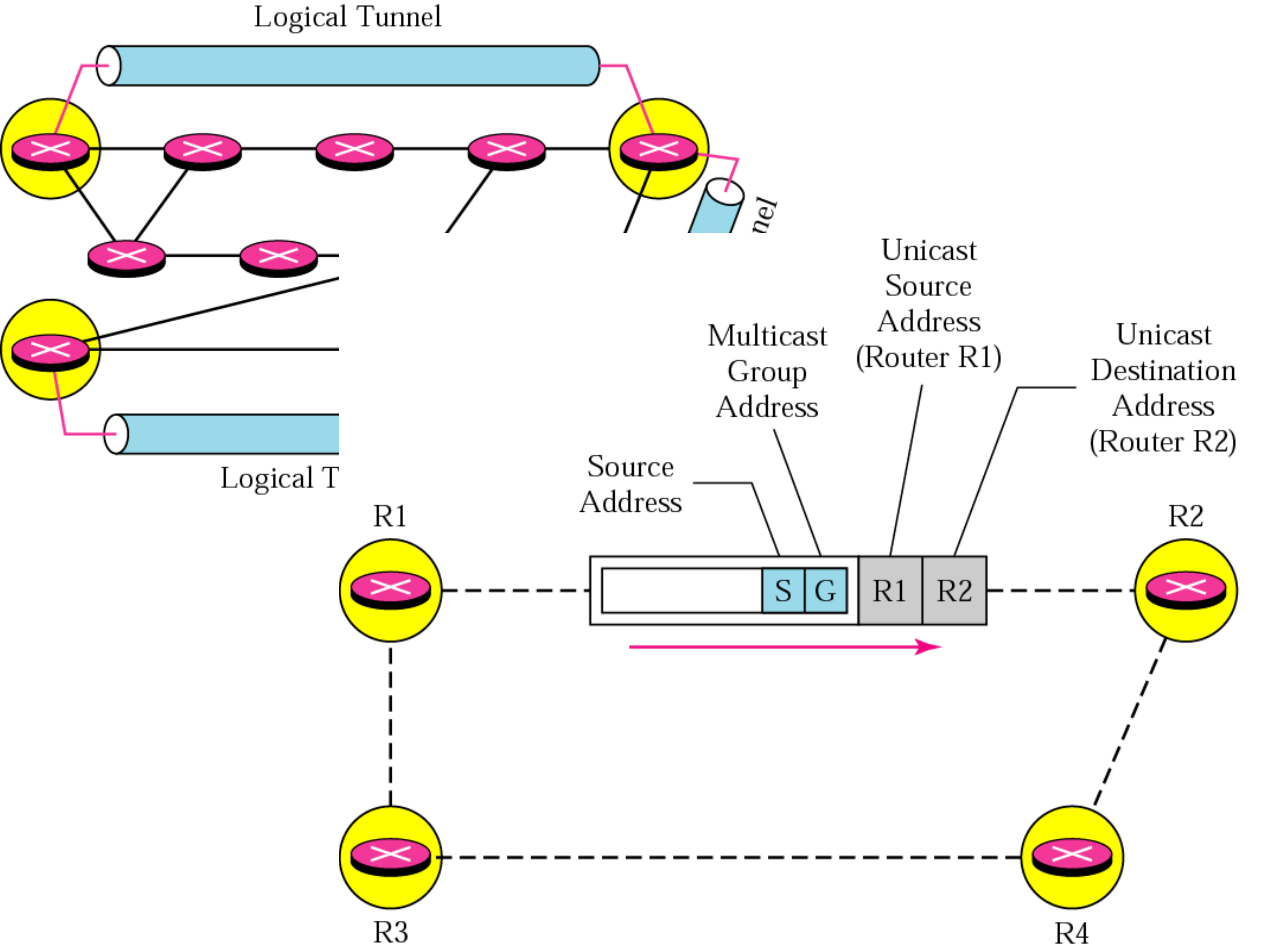
# Multicast in the Internet

- Many routers in the Internet do not run multicast routing protocols
  - they drop IP datagrams with Class D address in the destination field
- How can two hosts in two distant networks participate in multicasting if one or more intermediate routers do not support multicasting?




**Routers R1 & R3 run multicast but R2 doesn't. R2 drops multicast packets**

**Have you seen similar problem before?**


Logical Tunnel
Logical T
Multicast Group Address
Unicast Source Address (Router R1)
Unicast Destination Address (Router R2)
Source Address
R1
R2
S
G
R1
R2
R3
R4

# References

- Chapter 6 (IBM Redbook)
- Chapters 10 and 15 (Forouzan, 3rd Ed.)