第19章 互联网的操作

(2) 组播

南京大学计算机系　黄皓教授

2007年10月19日星期五

Reference

- **TCP/IP Tutorial and Technical Overview, ibm.com**/redbooks
- **Christian Huitma, Routing in the Internet.**

Multicast

Multicasting

- Multicast
  - Act of sending datagram to multiple receivers with single *transmit* operation
- Applications
  - Multimedia conferencing
  - Data distribution
  - Gaming and Simulation
  - Real-time data multicast

Multicast Applications — Multimedia conferencing

- The first multicast applications provided audio conferencing functions.
- These tools have been used on the MBONE for several years.
- They support many-to-many audio-only or audio-video communication.
- When used in conjunction with whiteboard applications, these conferences enhance collaboration while requiring minimal bandwidth.

Multicast Applications

- Data distribution
  - These tools provide the ability to simultaneously deliver data to large numbers of receivers.
  - For example, a central site can efficiently push updated data files to each district office.

Multicast Applications

- Gaming and Simulation
  - These applications have been readily available.
  - However, the integration of multicast services allow the applications to scale to a large number of users.
  - Multicast groups can represent different sections of the game or simulation.
  - As users move from one section to the next, they exit and join different multicast groups.

Multicast Applications

- Real-time data multicast
  - These applications distribute real-time data to large numbers of users.
  - For example, stock ticker information can be provided to sets of workstations.
  - The use of multicast groups can tailor the information received by a specific device.

Multicasting on a single physical network

- The sending process specifies a destination IP multicast address.
- The device driver converts this IP address to the corresponding Ethernet address and sends the packet to the destination.
- The destination process informs its network device drivers that it wishes to receive datagrams destined for a given multicast address.
- The device driver enables reception of packets for that address.

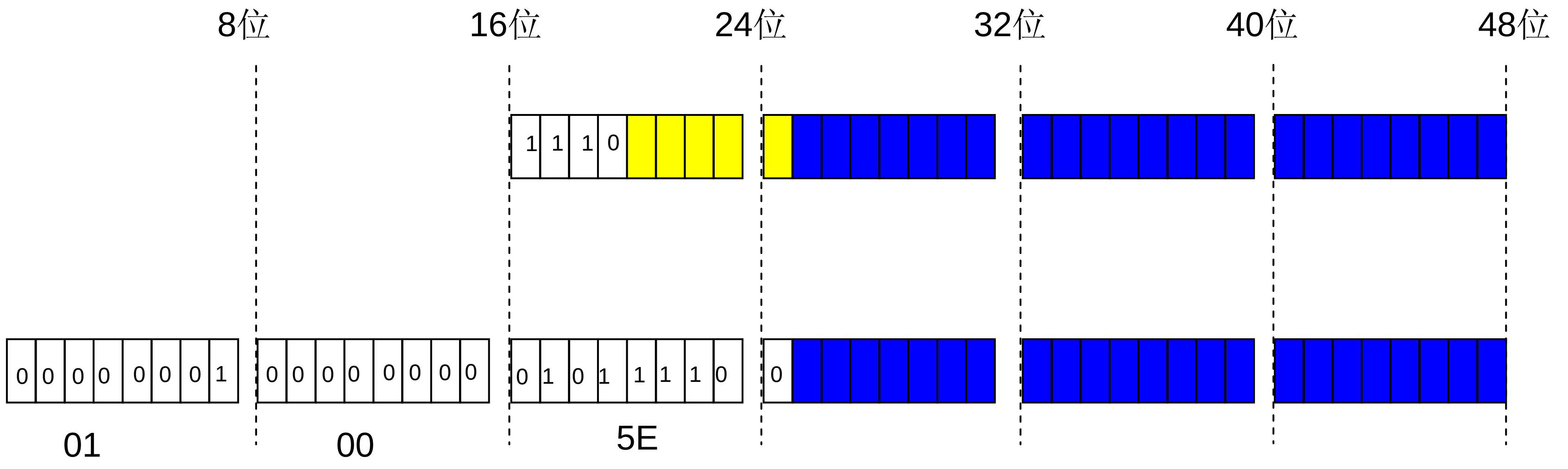
Multicast address

- the mapping between the IP multicast destination address and the data-link address is not done with ARP.
- Instead, a static mapping has been defined.

- Ethernet multicsat address
  - From 0x01-00-5e-00-00-00 to 0x01-00-5e-7f-ff-ff

- IP multicast address
  - 224.0.0.0 to 239.255.255.255

IP组播地址到MAC地址的映射



- MAC地址的第一个字节的最低位置为1；
- 组播MAC地址的范围：0x01-00-5e-00-00-00 到 0x01-00-5e-7f-ff-ff
- 映射：将IP组播地址的后23位映射到MAC组播地址的23位。
- IP组播地址中有5位没有映射，有不同的IP组播地址映射到相同的MAC组播地址，它们被主机链路层拒绝。

reserved multicast addresses

Adddress	meaning
224.0.0.1	All systems on this subnet
224.0.0.2	All routers on this subnet
224.0.0.4	DVMRP Routers
224.0.0.5	All OSPF routers
224.0.0.6	OSPF designated routers
224.0.0.09	All RIP2 Routers
224.0.0.10	All IGRP Routers
224.0.0.12	DHCP Routers
224.0.0.13	ALL PIM Routers
224.0.0.15	ALL CBT Routers
224.0.0.21	MOSPF上的DVMRP

Example of Group

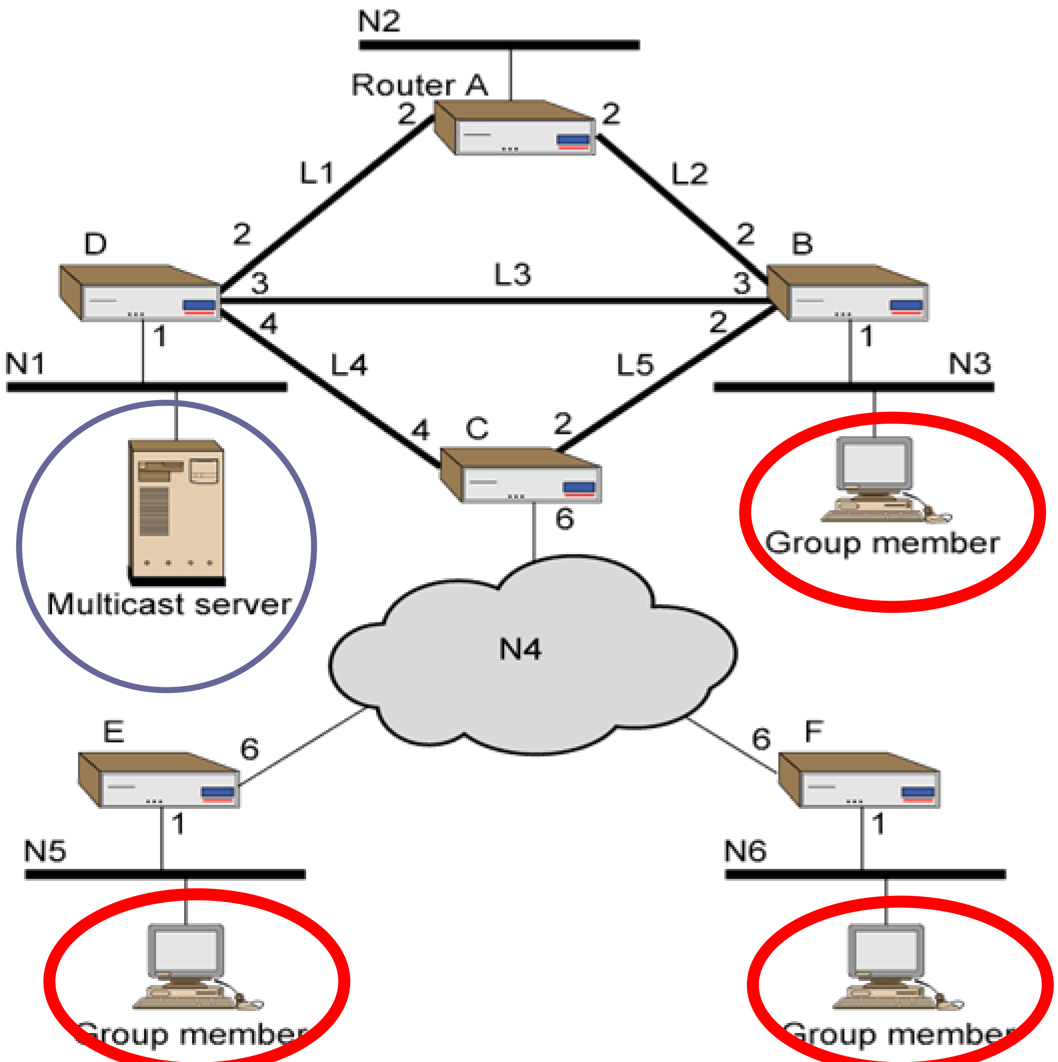
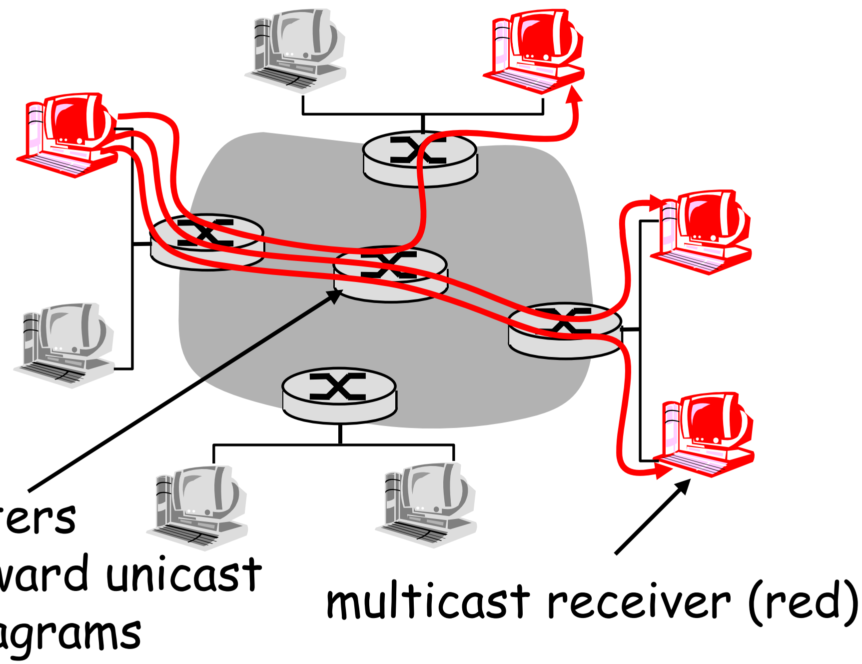


Table 19.1 Traffic Generated by Various Multicasting Strategies

	(a) Broadcast					(b) Multiple Unicast				Multicast
	S → N2	S → N3	S → N5	S → N6	Total	S → N3	S → N5	S → N6	Total	
N1	1	1	1	1	4	1	1	1	3	1
N2										
N3		1			1	1			1	1
N4			1	1	2		1	1	2	2
N5			1		1		1		1	1
N6				1	1			1	1	1
L1	1				1					
L2										
L3		1			1	1			1	1
L4			1	1	2		1	1	2	1
L5										
Total	2	3	4	4	13	3	4	4	11	8

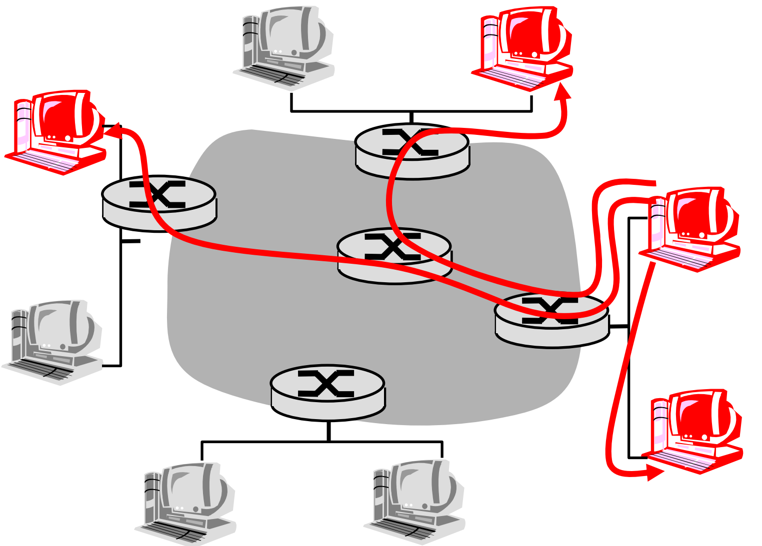
Ways of Achieving Multicast (1)

- Multicast via unicast
  - Source sends N unicast datagrams
  - One addressed to each of N receivers



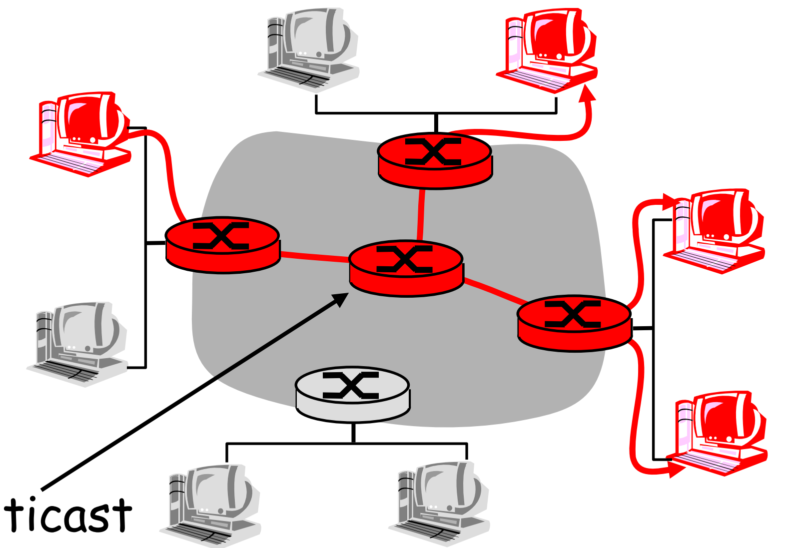
Ways of Achieving Multicast (2)

- Application-layer multicast
- End systems involved in multicast copy & forward unicast datagrams among each other



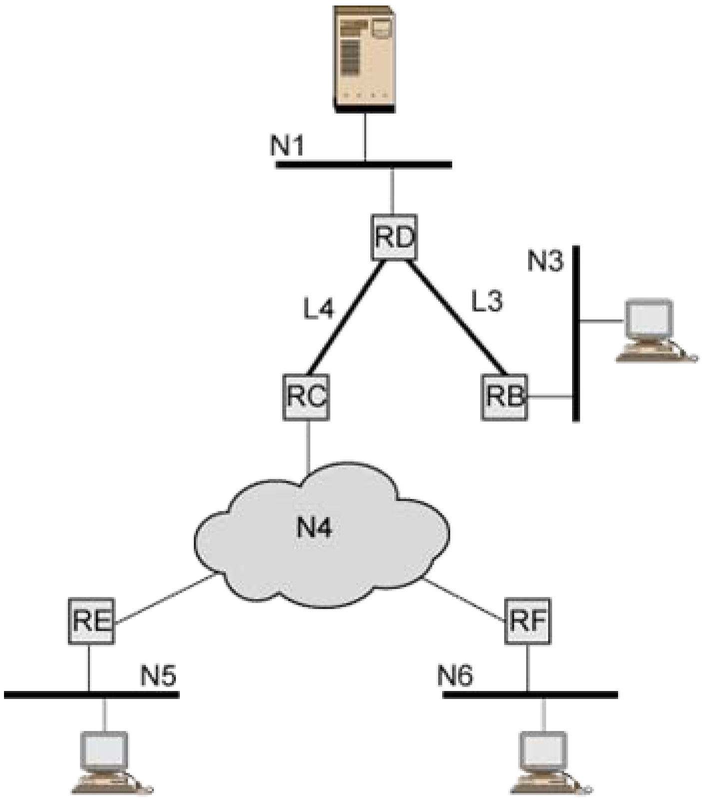
Ways of Achieving Multicast (3)

- Network multicast
  - Routers actively participate in multicast
  - Making copies of packets as needed and forwarding towards multicast receivers
  - Source only sends once

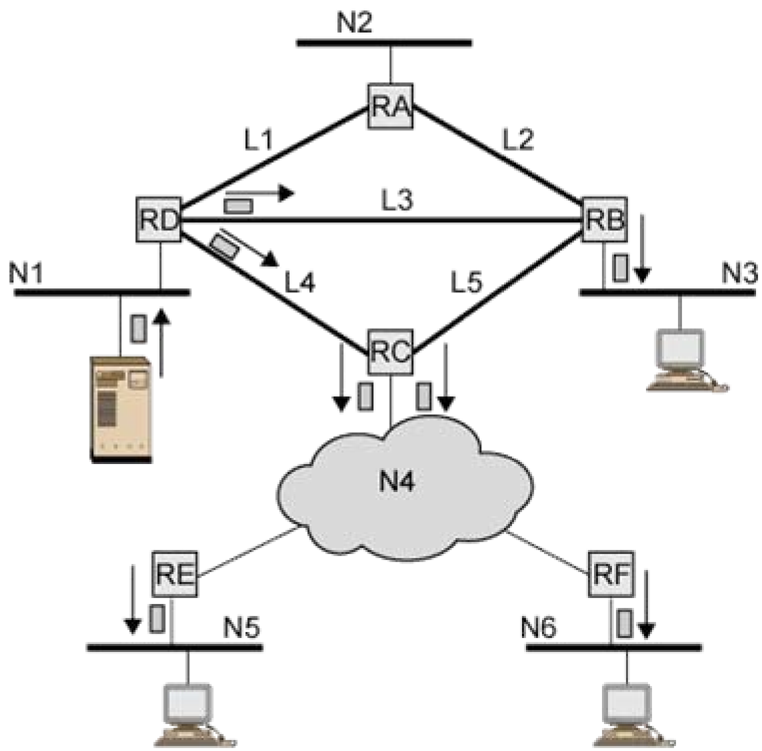


Multicast routers (red) duplicate and forward multicast datagrams

Multicast Example



(a) Spanning tree from source to multicast group



(b) Packets generated for multicast transmission

IP Multicast

- Determine least cost path to each network that has host in group
  - Gives spanning tree configuration containing networks with group members
- Transmit single packet along spanning tree
- Routers replicate packets at branch points of the spanning tree
- 8 packets required, instead of 11 packets in multiple unicast

Multicasting between network segments

- Multicast traffic is not limited to a single physical network.
- If the environment contains multiple routers, specific precautions must be taken to ensure multicast packets do not continuously loop through the network.

Two requirements to multicast data across multiple networks

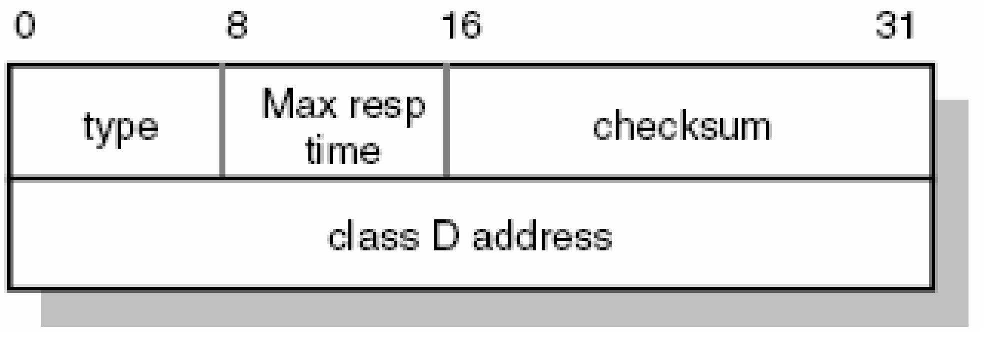
- Determining multicast participants
  - A mechanism for determining if a multicast datagram needs to be forwarded on a specific network.
  - This mechanism is defined in RFC 2236 Internet Group Management Protocol (IGMP).
- Determining multicast scope
  - A mechanism for determining the scope of a transmission. Unlike unicast addresses, multicast addresses can extend through the entire Internet.
  - The TTL field in a multicast datagram can be used to determine the scope of a transmission.

- TTL = 0: A multicast datagram received is restricted to the source host.
- TTL = 1: A multicast datagram reaches all hosts on the subnet that are members of the group.
- TTL = 2 (or more): A multicast datagram with this TTL value reaches all hosts on the subnet that are members of the group.

Internet Group Management Protocol (IGMP)

Internet Group Management Protocol (IGMP)

- IGMP Version 2 is described in RFC 2236.
- IGMP Version 3 is described in RFC 2276.
- IGMP messages are encapsulated in IP datagrams.



IGMP — Basic concept

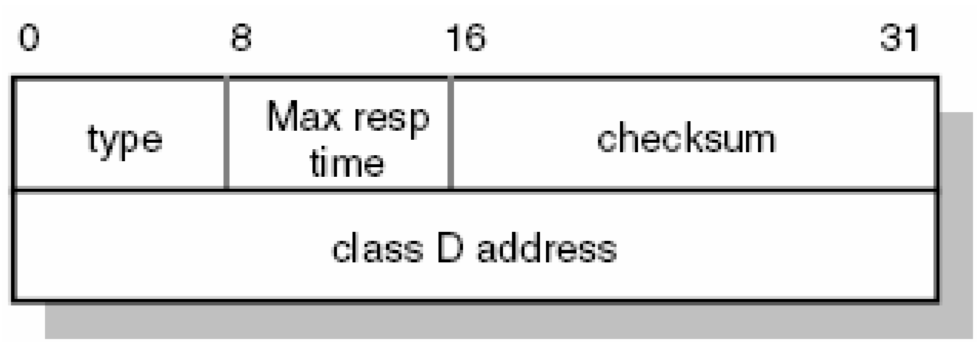
- IGMP is used by IP v4 systems (hosts and routers) to report their IP multicast group memberships to any neighboring multicast routers.
  - Routers
    - collect the membership information needed by its multicast routing protocol .
    - Sufficient info to work out shortest path to each network
    - Determining routing paths based on source and destination addresses
  - Hosts inform itself and other, neighboring multicast routers of its memberships.

- IGMPv1
  - Hosts could join group
  - Routers used timer to unsubscribe members
- Operation Model of IGMPv1
  - Receivers have to subscribe to groups
  - Sources do not have to subscribe to groups
  - Any host can send traffic to any multicast group
- Problems
  - Even if application level filters drop unwanted packets, they consume valuable resources
  - Establishment of distribution trees is problematic
  - Location of sources is not known
  - Finding globally unique multicast addresses difficult

- IGMP v3
  - Allows hosts to specify list from which they want to receive traffic
    - Traffic from other hosts blocked at routers
  - Allows hosts to block packets from sources that send unwanted traffic

IGMP Message Formats

- The IP data field contains the 8-octet IGMP message



- Type: This field specifies the type of IGMP packet:
  - 0x11 Membership Query
  - 0x12 Version 1 Membership Report [RFC-1112]
  - 0x16 Version 2 Membership Report [RFC-2236]
  - 0x22 Version 3 Membership Report [RFC-3376]
  - 0x17 Version 2 Leave Group [RFC-2236]

Membership Query Message

- Membership Queries
  - sent by IP multicast routers
  - to query the multicast reception state of neighboring interfaces.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Type = 0x11  | Max Resp Code |           Checksum            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Group Address                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Resv  |S| QRV |     QQIC      |     Number of Sources (N)     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Source Address [1]                      |
+-                                                             -+
|                       Source Address [2]                      |
+-                              .                              -+
.                               .                               .
.                               .                               .
+-                                                             -+
|                       Source Address [N]                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Membership Query

- General query
  - Which groups have members on attached network
- Group-specific query
  - Does group have members on an attached network
- Group-and-source specific query
  - Do attached device want packets sent to specified multicast address from any of specified list of sources

Membership Query Fields (1)

- Type
- Max Response Time
  - Max time before sending report in units of 1/10 second
- Checksum
  - Same algorithm as IPv4
- Group Address
  - Zero for general query message
  - Multicast group address for group-specific or group-and-source
- S Flag
  - 1 indicates that receiving routers should suppress normal timer updates done on hearing query

Membership Query Fields (2)

- QRV (querier's robustness variable)
  - RV value used by sender of query
  - Routers adopt value from most recently received query
  - Unless RV was zero, when default or statically configured value used
  - RV dictates number of retransmissions to assure report not missed
- QQIC (querier's querier interval code)
  - QI value used by querier
  - Timer for sending multiple queries
  - Routers not current querier adopt most recently received QI
  - Unless QI was zero, when default QI value used

- Number of Sources
  - The Number of Sources (N) field specifies how many source addresses are present in the Query.
  - This number is zero in a General Query or a Group-Specific Query
  - It is non-zero in a Group-and-Source-Specific Query.
- Source addresses
  - One 32 bit unicast address for each source

Version 3 Membership Report Message

- Membership Reports are sent by IP systems to report (to neighboring routers) the current multicast reception state, or changes in the multicast reception state, of their interfaces.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Type = 0x22  |    Reserved   |           Checksum            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Reserved            |  Number of Group Records (M)  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
.                                                               .
.                        Group Record [1]                       .
.                                                               .
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
.                                                               .
.                        Group Record [2]                       .
.                                                               .
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                               .                               |
.                               .                               .
|                               .                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
.                                                               .
.                        Group Record [M]                       .
.                                                               .
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Group Record

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Record Type  |  Aux Data Len |     Number of Sources (N)     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Multicast Address                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Source Address [1]                      |
+-                                                             -+
|                       Source Address [2]                      |
+-                                                             -+
.                               .                               .
.                               .                               .
.                               .                               .
+-                                                             -+
|                       Source Address [N]                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
.                                                               .
.                         Auxiliary Data                        .
.                                                               .
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Group Record

- Aux Data Length
  - In 32-bit words
- Number of Sources
- Multicast Address
  - The Multicast Address field contains the IP multicast address to which this Group Record pertains.
- Source Addresses
  - One 32-bit unicast address per source
- Auxiliary Data
  - Currently, no auxiliary data values defined

Group Record Types

- Current-State Record
  - sent by a system in response to a Query received on an interface.
  - It reports the current reception state of that interface, with respect to a single multicast address.
  - MODE_IS_INCLUDE
    - indicates that the interface has a filter mode of INCLUDE for the specified multicast address.
  - MODE_IS_EXCLUDE
    - indicates that the interface has a filter mode of EXCLUDE for the specified multicast address.

Group Record Types

- Filter-Mode-Change Record
  - CHANGE_TO_INCLUDE_MODE
  - CHANGE_TO_EXCLUDE_MODE
- Source-List-Change Record
  - ALLOW_NEW_SOURCES
  - BLOCK_OLD_SOURCES

IGMP Operation - Joining

- Host using IGMP wants to make itself known as group member to other hosts and routers on LAN
- IGMPv3 can signal group membership with filtering capabilities with respect to sources
  - EXCLUDE mode – all group members except those listed
  - INCLUDE mode – Only from group members listed
- To join group, host sends IGMP membership report message
  - **Address field multicast address of group**
  - **Sent in IP datagram with Group Address field of IGMP message and Destination Address encapsulating IP header same**
  - **Current members of group will receive learn of new member**
  - **Routers listen to all IP multicast addresses to hear all reports**

IGMP Operation – Keeping Lists Valid

- Routers periodically issue IGMP general query message
  - In datagram with all-hosts multicast address(224.0.0.1)
  - Hosts that wish to remain in groups must read datagrams with this all-hosts address
  - Hosts respond with report message for each group to which it claims membership

- Router does not need to know every host in a group
  - Needs to know at least one group member still active
  - Each host in group sets timer with random delay
  - Host that hears another claim membership cancels own report
  - If timer expires, host sends report
  - Only one member of each group reports to router

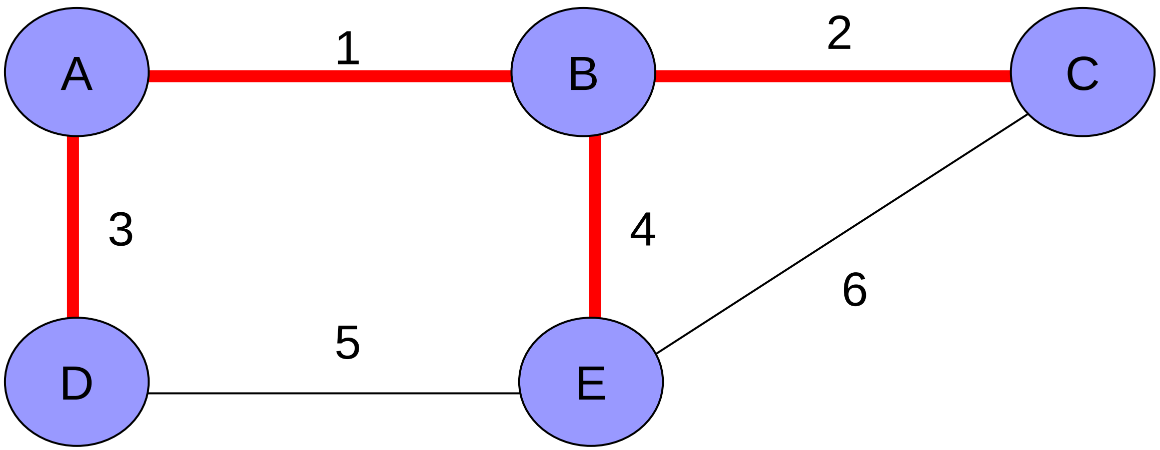
IGMP Operation - Leaving

- Host leaves group, by sending leave group message to all-routers(224.0.0.2) static multicast address
- Send membership report message with EXCLUDE option and null list of source addresses
- Router determine if there are any remaining group members using group-specific query message

Multicast forwarding algorithms

Multicast forwarding algorithms

- Flooding
- Spanning tree



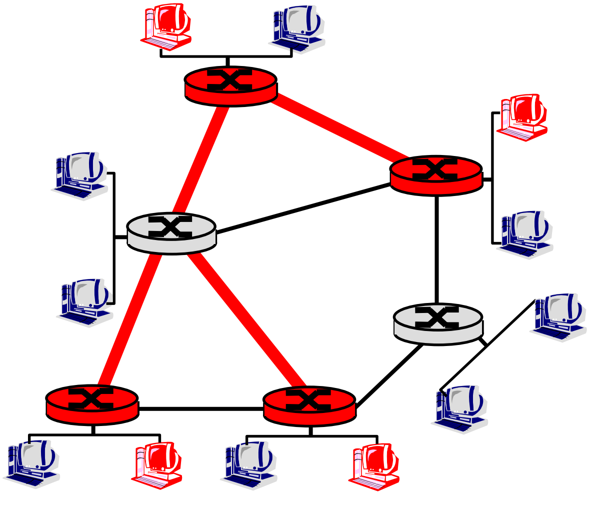
- If the group only consists of A,B and C, it also sends pachet to D and E.
- If the group only consists of C, D and E, it also uses the links 1,2,3,4 to transmit. Obviously, it is better to use the links 5 and 6.

Multicast forwarding algorithms

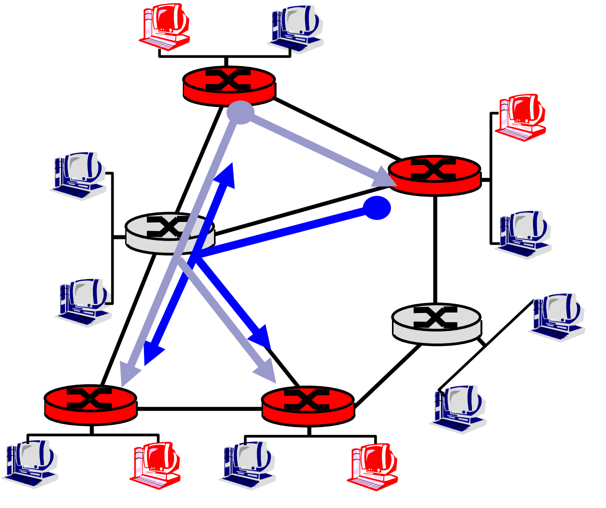
- Multicast algorithms are used to establish paths through the network.
  - **The algorithm must route data only to group members.**
  - **The algorithm must optimize the path from source to destinations.**
  - **The algorithm must maintain loop-free routes.**
  - **The algorithm must provide scalable signaling functions used to create and maintain group membership.**
  - **The algorithm must not concentrate traffic on a subset of links.**

Multicast forwarding algorithms

- The Goal
  - Find a spanning tree (or trees) connecting routers having local mcast group members
- Source-based
  - Different tree from each sender to receivers
- Shared-tree
  - Same spanning tree used by all group members



Shared tree



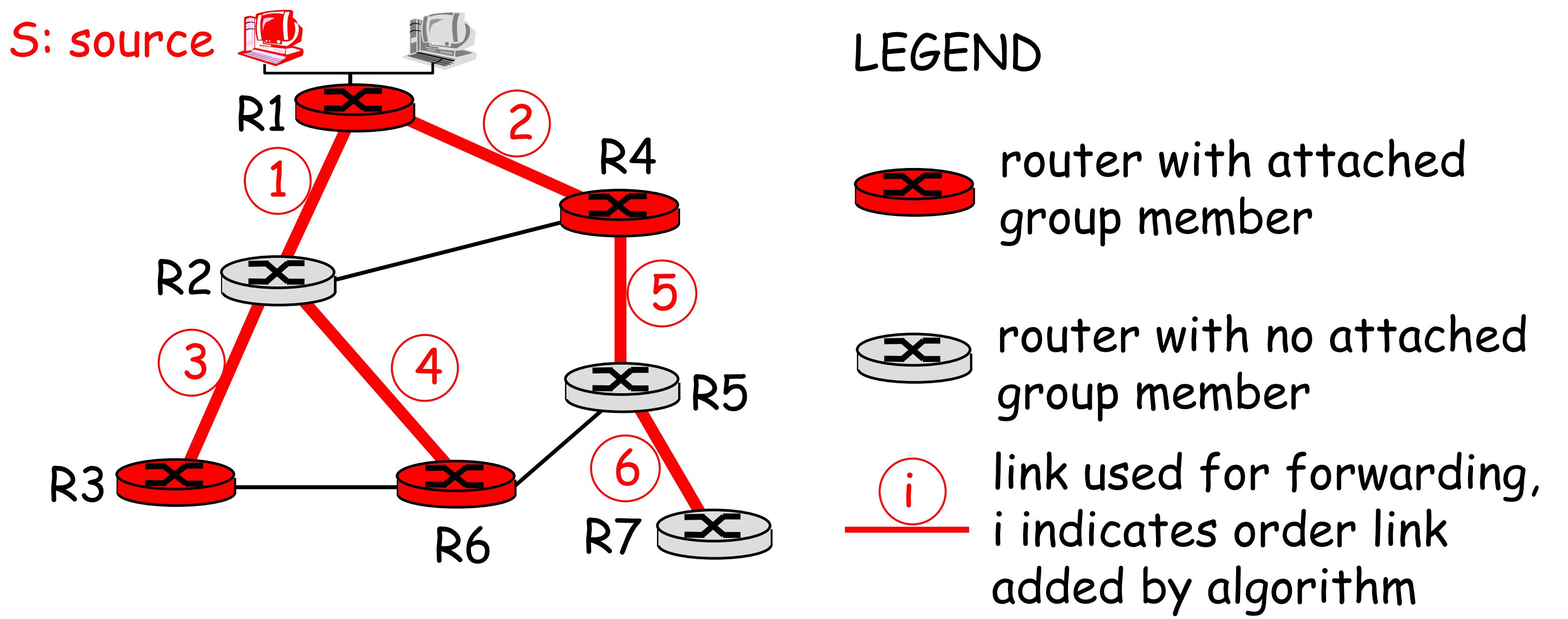
Source-based trees

Approaches for Multicast Trees

- Source-based tree: one tree per source
  - Shortest path trees
  - Reverse path forwarding
- Group-shared tree: group uses one tree
  - Minimal spanning (Steiner)
  - Center-based trees

Shortest Path Trees

- Multicast forwarding tree
  - Tree of shortest path routes from source to all receivers
  - Use Dijkstra's algorithm



Reverse Path Forwarding

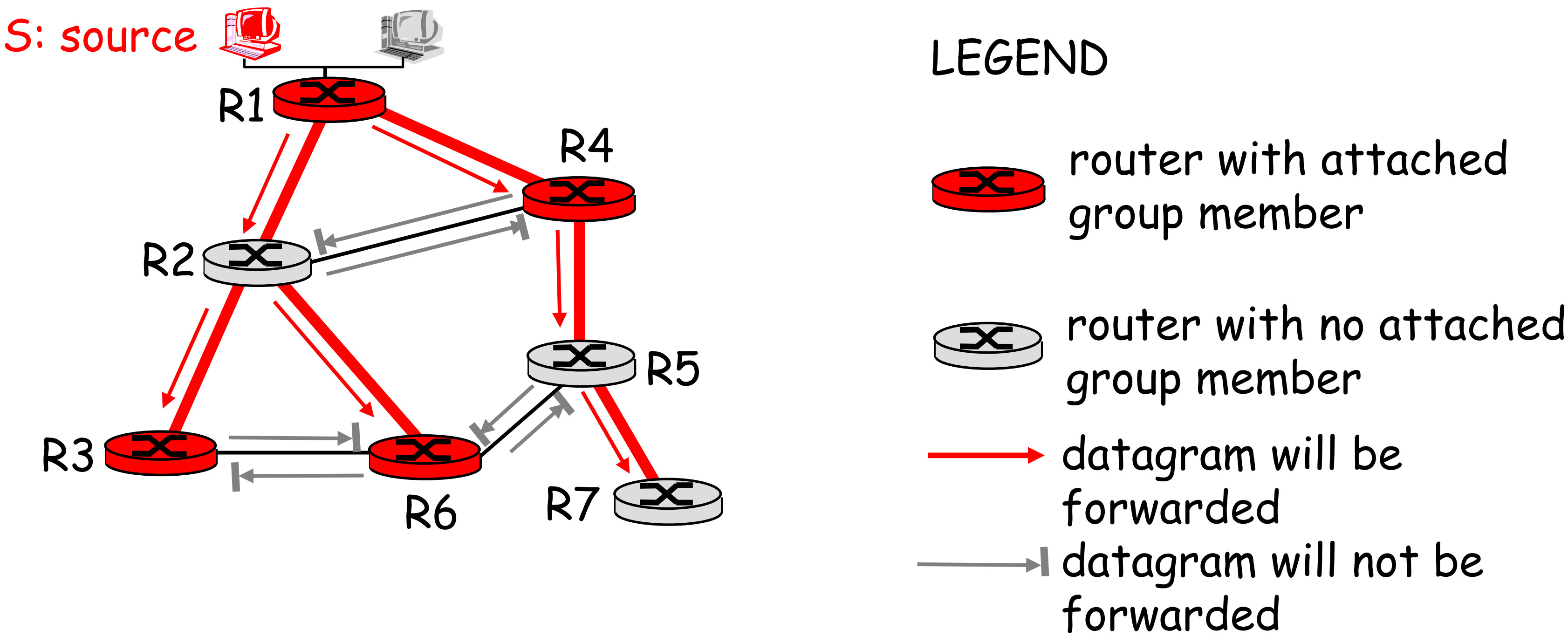
- Rely on router's knowledge of unicast shortest path from it to sender
- Each router has simple forwarding behavior:

if (mcast datagram received on incoming link on shortest path back to sender)

then flood datagram onto all outgoing links

else ignore datagram

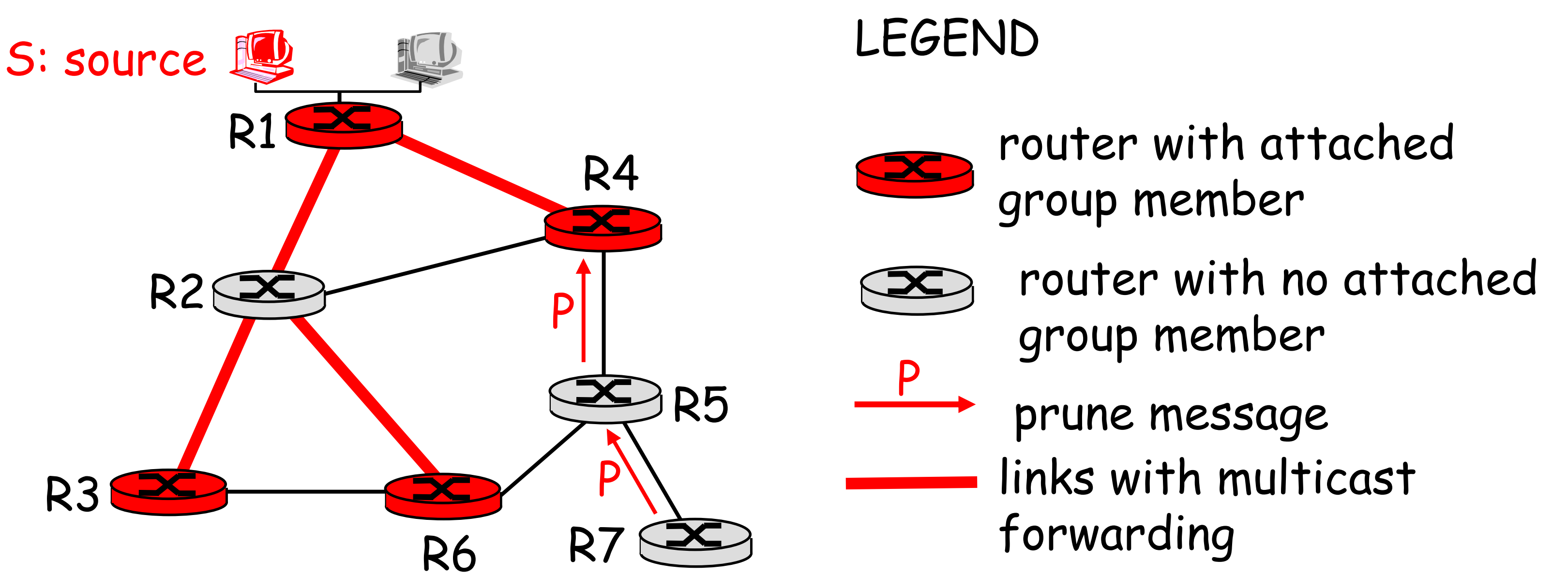
Reverse Path Forwarding: Example



- The result is a source-specific reverse SPT
  - May be a bad choice with asymmetric links
  - Asymmetric link must use forward-path.

Reverse Path Forwarding: Pruning

- Forwarding tree contains subtrees with no mcast group members
  - No need to forward datagrams down subtree
  - "Prune" msgs sent upstream by router with no downstream group members



Shared-Tree: Steiner Tree

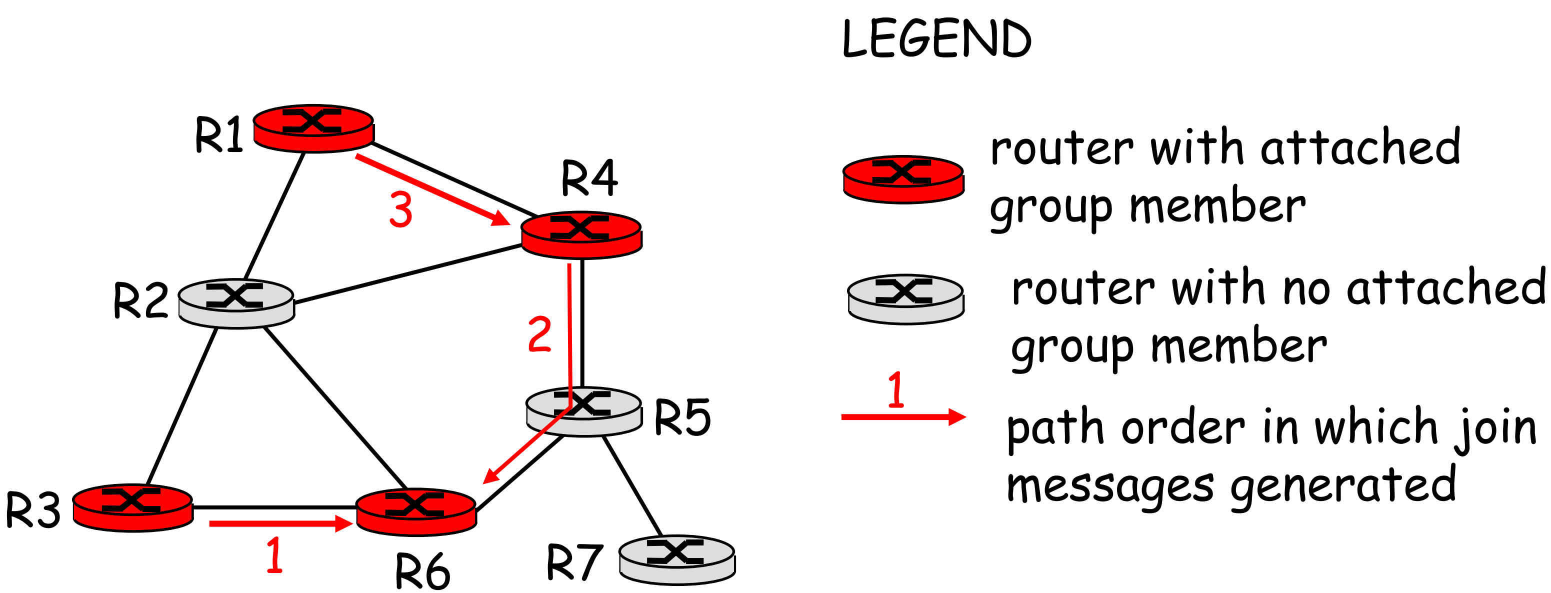
- Steiner Tree
  - Minimum cost tree connecting all routers with attached group members
- Problem is NP-complete
  - It is a Salesman problem
  - But excellent heuristics exists
- Not used in practice
  - Computational complexity
  - Information about entire network needed
  - Monolithic: rerun whenever a router needs to join/leave

Center-based Trees

- Single delivery tree shared by all
- One router identified as center of tree
- Other routers to join:
  - Edge router sends unicast join-msg addressed to center router
  - join-msg processed by intermediate routers and forwarded towards center
  - join-msg either hits existing tree branch for this center, or arrives at center
  - Path taken by join-msg becomes new branch of tree for this router

Center-based Trees: Example

Suppose R6 chosen as center:



Multicast Routing Protocol —— DVMRP

Distance Vector Multicast Routing Protocol (DVMRP)

- DVMRP is an established multicast routing protocol. It was originally defined in **RFC 1075**.
- The standard was first implemented as the ***mrouted*** process available on many UNIX systems.
- It has since been enhanced to support RPF. DVMRP is an interior gateway protocol. It is used to build per-source per-group multicast delivery trees within an AS.
- Any router that processes both multicast and unicast datagrams must be configured with two separate routing processes.

DVMRP — Protocol overview

- DVMRP is described as a ***broadcast and prune*** multicast routing protocol
  - DVMRP builds per-source broadcast trees based upon routing exchanges.
  - DVMRP dynamically prunes the per-source broadcast tree to create a multicast delivery tree. DVMRP uses the RPF algorithm to determine the set of downstream interfaces used to forward multicast traffic.

DVMRP — Protocol overview

- **Neighbor discovery**
  - DVMRP routers dynamically discover each neighbor by periodically sending neighbor probe messages on each local interface.
  - These messages are sent to the *all-DVMRP-routers* multicast address (224.0.0.4).
  - Each message contains a list of neighbor DVMRP routers for which neighbor probe messages have been received.
  - This allows a DVMRP router to verify it has been seen by each neighbor.
  - Once a router has received a probe message that contains its address in the neighbor list, the pair of routers establish a two-way neighbor adjacency.

DVMRP — Protocol overview

- **Routing table creation**
  - DVMRP computes the set of reverse paths used in the RPF algorithm.
  - DVMRP implements its own unicast routing protocol similar to RIP.
  - When a route is received, the interface metric is added to the advertised metric. This adjusted metric is used to determine the best upstream path to the source.
  - DVMRP has one important difference from RIP. RIP manages routing and datagram forwarding to a particular unicast destination. DVMRP manages the return path to the source of a particular multicast datagram.

DVMRP — Protocol overview

- **Dependent downstream routers**
  - DVMRP uses exchanging routing information mechanism to notify upstream routers that a specific downstream router requires them to forward multicast traffic.
  - When the upstream router receives the advertisement, it adds the downstream router to a list of ***dependent downstream*** *routers* for this source.
  - This technique provides the information needed to prune the multicast delivery tree.

DVMRP — Protocol overview

- **Designated forwarder**
  - When two or more multicast routers are connected to a multi-access network, duplicate packets may be forwarded to the network. DVMRP prevents this possibility by electing a designated forwarder for each source.

DVMRP — Building and maintaining multicast delivery trees

- If a datagram was received on the interface representing the best path to the source, the router forwards the datagram out a set of downstream interfaces.
- This set contain each downstream interface included in the multicast delivery tree.

DVMRP — Building and maintaining multicast delivery trees

- **Building the multicast delivery tree**
  - A multicast router forwards datagrams to two types of devices: downstream dependent routers and hosts that are members of a particular multicast group.
  - If a multicast router has no dependent downstream neighbors through a specific interface, the network is a ***leaf network***.
- ***Adding leaf networks***
  - If the downstream interface connects to a leaf network, packets are forwarded only if there are hosts that are members of the specific multicast group.
  - The router obtains this information from the IGMP local group database.
  - If the group address is listed in the database, and the router is the designated forwarder for the source, the interface is included in the multicast delivery tree. If there are no group members, the interface is excluded.

DVMRP — Building and maintaining multicast delivery trees

- ***Adding non-leaf networks***
  - Initially, all non-leaf networks are included in the multicast delivery tree.
  - This allows each downstream router to participate in traffic forwarding for each group.

DVMRP — Pruning the multicast delivery tree

- Routers connected to leaf networks remove an interface
  - when there are no longer any active members participating in the specific multicast group. When this occurs, multicast packets are no longer forwarded through the interface.
- Routers notifies its upstream neighbor that it no longer needs traffic from that particular source and group pair.
  - If a router is able to remove all of its downstream interfaces for a specific group.
  - This continues until all unnecessary branches have been removed from the delivery tree.

DVMRP — Pruning the multicast delivery tree

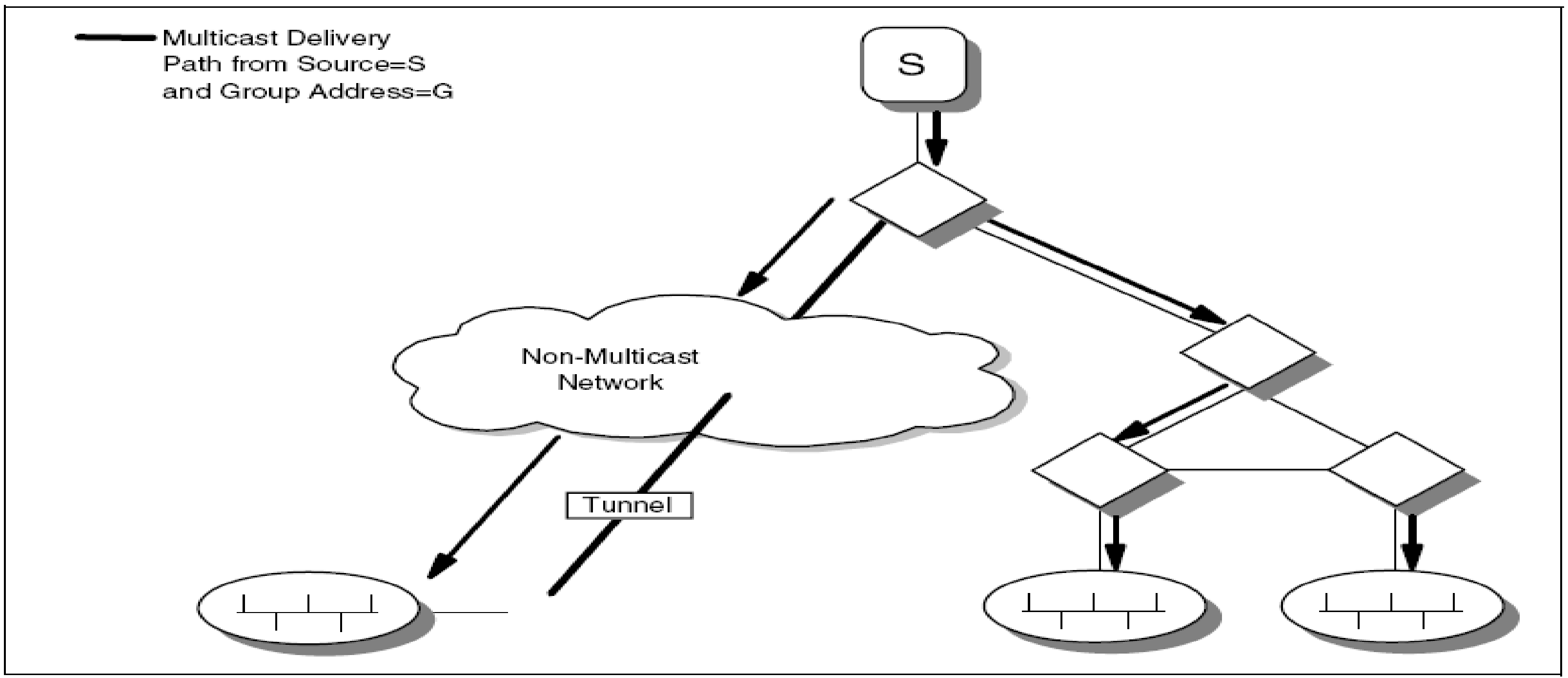
- ***Maintaining prune information***
  - In order to remove outdated prune information, each prune message contains a prune lifetime timer.
  - If the interface is still pruned when the timer expires, the interface is reconnected to the multicast delivery tree.
  - If this causes unwanted multicast datagrams to be delivered to a downstream device, the prune mechanism is reinitiated.

DVMRP — Grafting pruned networks

- IP multicast supports dynamic group membership, hosts may join a multicast group at any time.
- A **graft message** is sent as a result of receiving a IGMP membership report for a group that has previously been pruned.
- Receipt of a graft message is acknowledged with a graft ACK message.

DVMRP — DVMRP tunnels

- Some IP routers may not be configured to support native multicast routing.
- DVMRP provides the ability to tunnel IP multicast datagrams through networks containing non-multicast routers.



Multicast Routing Protocol
—— Multicast OSPF (MOSPF)

Multicast OSPF (MOSPF)

- MOSPF is a multicast extension to OSPF Version 2. It is defined in RFC 1584.
- MOSPF is not a separate multicast routing protocol. It is used in networks that already utilize OSPF for unicast IP routing.
- The multicast extensions leverage the existing OSPF topology database to create a ***source-rooted shortest path delivery tree*.**
- MOSPF forwards multicast datagrams using both source and destination address. This contrasts the standard OSPF algorithm, which relies solely on destination address.

MOSPF — Protocol overview

- **Group-membership LSA**
  - OSPF adds a new type of link state advertisement (thegroup-membership-LSA) to track the location of each group member.
  - These LSAs are stored in the OSPF link state database.
- **Designated routers**
  - On each network segment, one MOSPF router is selected to be the designated router (DR).
  - This router is responsible for generating periodic IGMP host membership queries.
  - It is also responsible for listening to the IGMP membership reports.
  - Every router floods a group-membership-LSA for each multicast group having at least one entry in the router’s local group database.

MOSPF — Protocol overview

- **Shortest-path delivery trees**
  - The path used to forward a multicast datagram is calculated by building a shortest-path delivery tree rooted at the datagram's source
  - Any branch in the shortest-path delivery tree that does not have a corresponding group-membership-LSA is pruned.
  - Initially, shortest-path delivery trees are built when the first datagram is received. The results are cached for use by subsequent datagrams having the same source and destination.
  - In an MOSPF network, all routers calculate an identical shortest-path delivery tree for a specific multicast datagram.
  - The tree is recomputed when a link state change occurs or when the cache information times out.

MOSPF — multiple OSPF areas

- An OSPF ABR can also function as an MOSPF **inter-area multicast forwarder**. It forwards group membership information and multicast datagrams between areas.
- Each inter-area multicast forwarder summarizes the attached areas' group membership requirements and forwards this information into the OSPF backbone.
- Unlike route summarization in a standard OSPF network, summarization for multicast group membership in MOSPF is asymmetric.
  - the non-backbone area is summarized into the backbone.
  - this information is not readvertised into other non-backbone areas.

MOSPF — multiple OSPF areas

- To forward multicast data traffic between areas, a ***wild-card multicast receiveris*** used.
  - This is a router to which all multicast traffic, regardless of destination, is forwarded.
- In non-backbone areas, all **inter-area multicast forwarders are wild-card multicast receivers**.
  - This router sends the multicast datagrams to the backbone area.
  - Since the backbone has complete knowledge of all group membership information, the datagrams are then forwarded to the appropriate group members in other areas.

MOSPF — multiple autonomous systems

- When at least one multicast device resides in another AS, the shortest path tree describing the complete path from source to destination cannot be built.
- In this environment, an **ASBR** in the MOSPF domain is configured as an **inter-AS multicast forwarder**.
- MOSPF designates an inter-AS multicast forwarder as a wild-card multicast receiver.
  - They receive all multicast datagrams, regardless of destination.
  - Since this device has complete knowledge of all group membership outside the AS, datagrams can be forwarded to group members in other autonomous systems.

MOSPF —— interoperability

- Routers configured to support an MOSPF network can be intermixed with non-multicast OSPF routers.
- However, forwarding IP multicast traffic is limited to the MOSPF domain.
- Unlike DVMRP, MOSPF does not provide the ability to tunnel multicast traffic through non-multicast routers.

Multicast Routing Protocol
— Protocol Independent Multicast (PIM)

Protocol Independent Multicast (PIM)

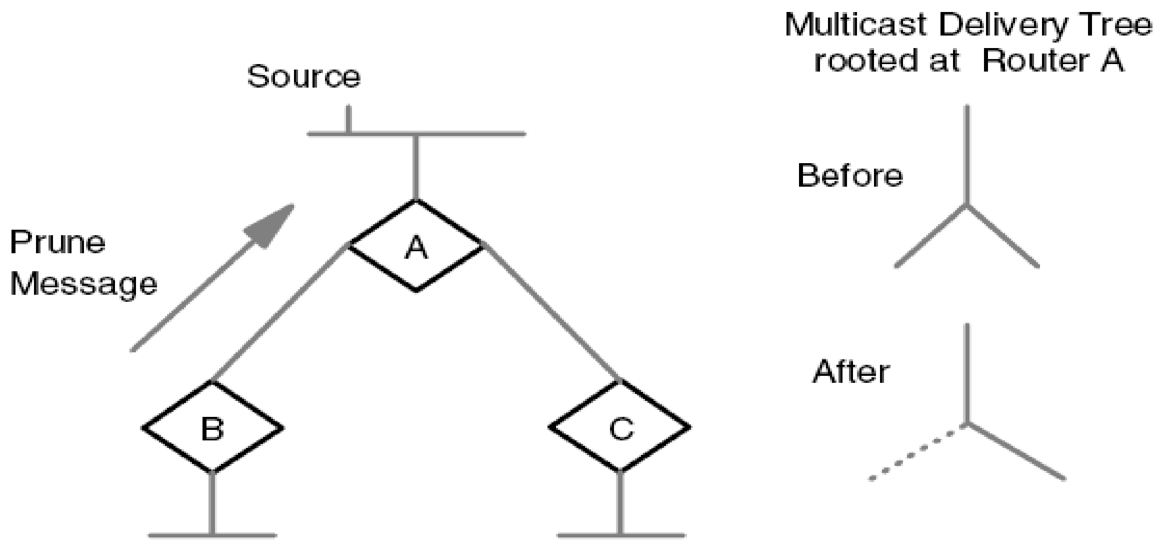
- PIM is independent of any underlying unicast routing protocol.
- It interoperates with all existing unicast routing protocols.
- PIM defines two modes or operation:
  - Dense mode (PIM-DM)
  - Sparse mode (PIM-SM), specified in RFC 2362
- Dense mode and sparse mode refer to the density of group members within an area.

PIM dense mode

- The PIM-DM protocol implements the RPF process.
- When a PIM-DM device receives a packet，the router floods the multicast packet provide the incoming interface reflects the best path back to the source.
- The packet is sent out to every interface that has not been pruned from the multicast delivery tree.
- Unlike DVMRP, PIM-DM does not attempt to compute multicast specific routes. Rather, it assumes that the routes in the unicast routing table are symmetric.

PIM dense mode

- Initially， the router floods datagrams to all areas of the network.
- If some areas do not have receivers for the specific multicast group, PIM-DM reactively prunes these branches from the delivery tree.



PIM-DM benefits

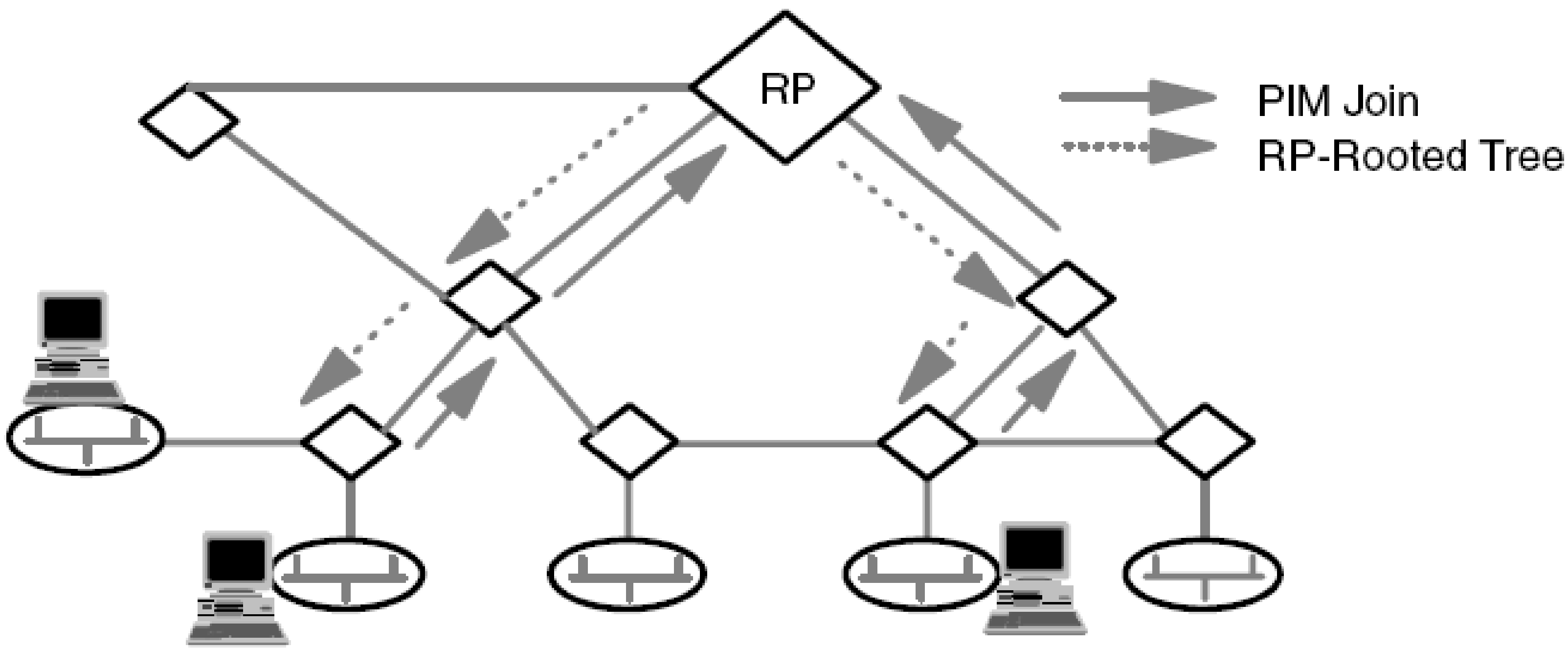
- PIM-DM, this protocol should be used in environments where the **majority of hosts** within a domain need to receive the multicast data.
- The overhead associated with flooding is minimal.
  - Senders and receivers are in close proximity to each other.
  - There are few senders and many receivers.
  - The volume of multicast traffic is high.
  - The stream of multicast traffic is constant.

PIM sparse mode (PIM-SM)

- The PIM-SM protocol uses a variant of the center-based tree algorithm.
- In a PIM-SM network, a ***rendezvous point*** *(RP)* is analogous to the center point described in the algorithm.
- Receivers join a tree rooted at the RP.
- Senders register their existence with the RP.
- Initially, traffic from the sender flows through the RP to reach each receiver.
- Multicast data is blocked from a network segment unless a downstream device specifically asks to receive the data.
  - reduce the amount of traffic traversing the network.
  - No pruning information is maintained for locations with no receivers.
- Because of these benefits, PIM-SM is currently the most popular multicast routing protocol used in the Internet.

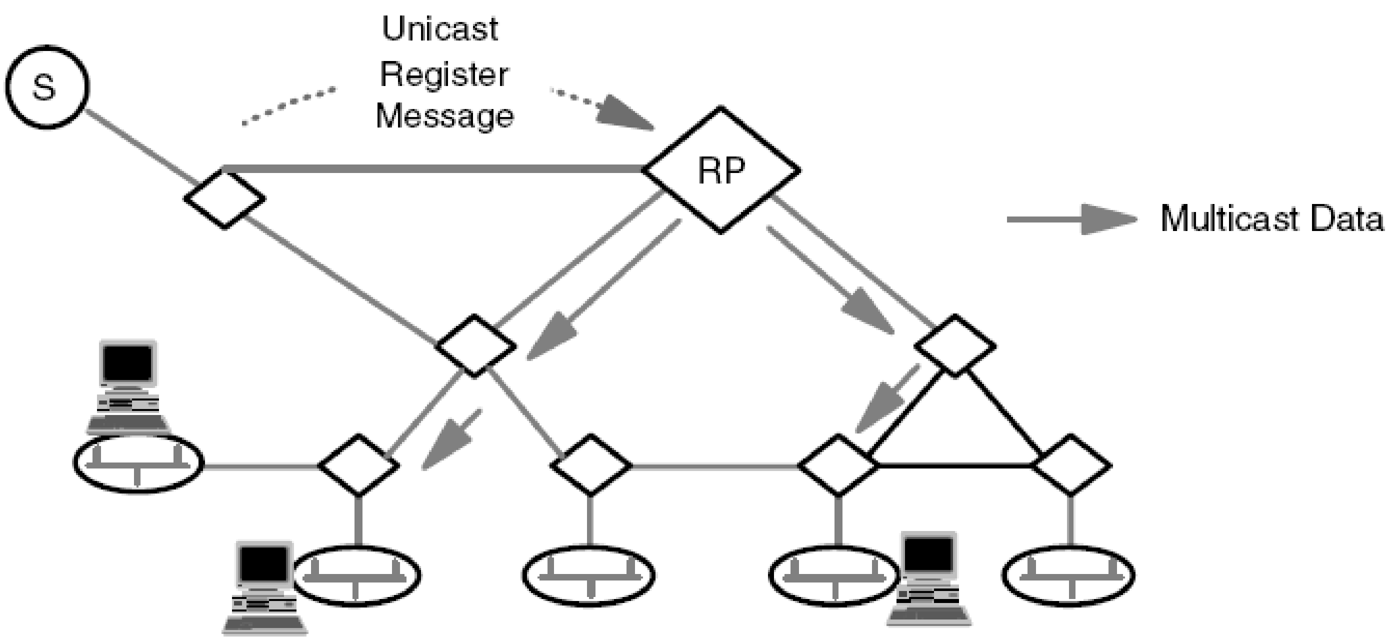
PIM-SM — Building the multicast delivery tree

- A multicast router sends periodic join messages to a group-specific RP.
- Like other multicast protocols, the tree is actually a reverse path tree, because join requests follow a reverse path from the receiver to the RP.



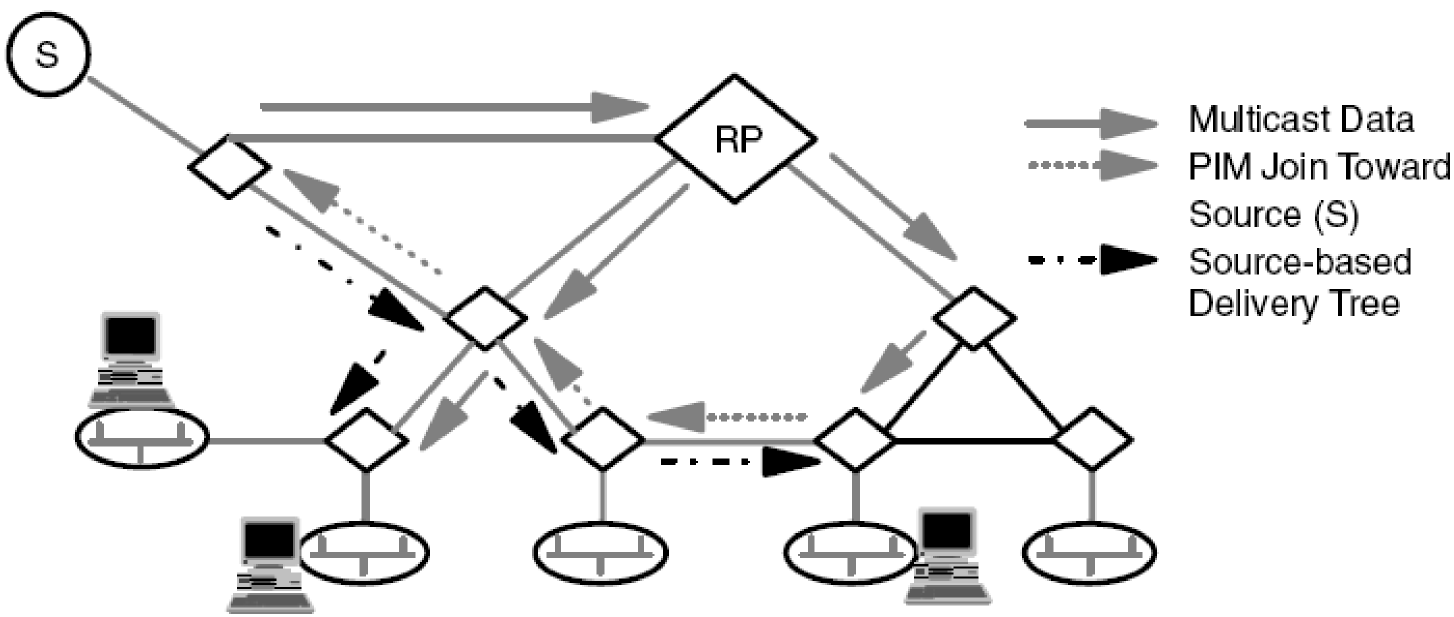
PIM-SM — Building the multicast delivery tree

- The multicast router connecting to the source initially encapsulates each multicast packet in a register message.
- These messages are sent to the RP.
- The RP decapsulates these unicast messages and forwards the data packets to the set of downstream receivers.



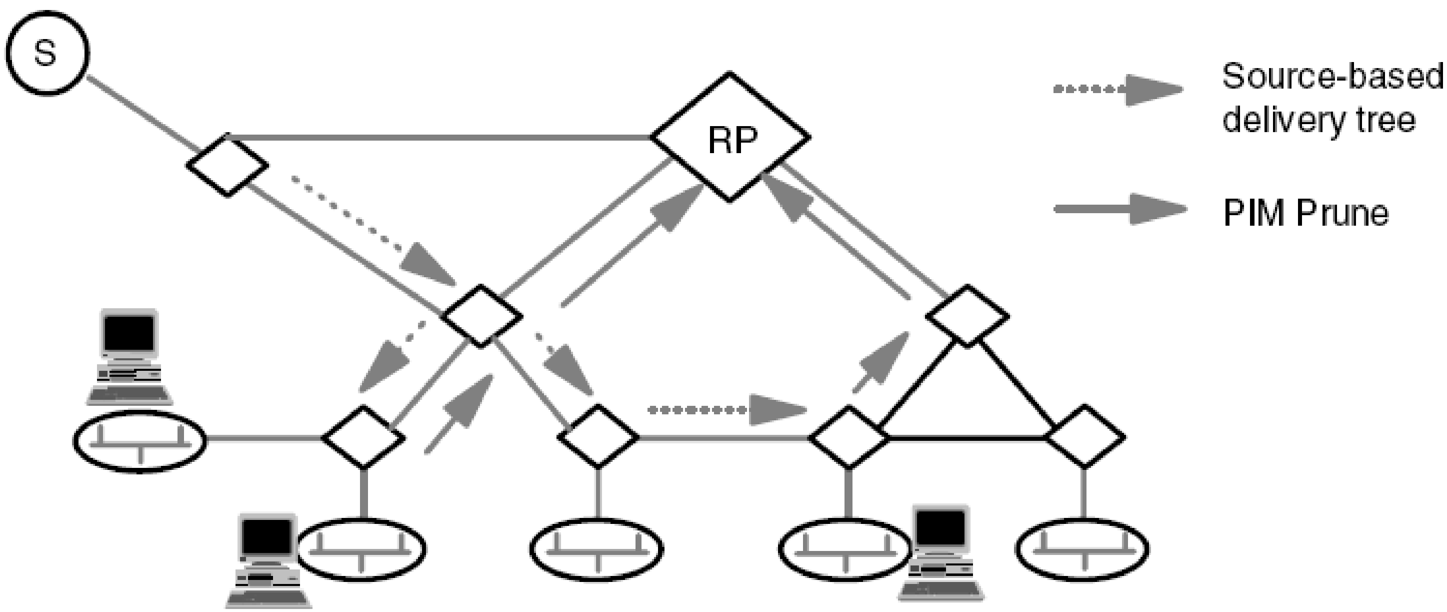
PIM-SM — Building the multicast delivery tree

- The RP-based delivery tree may reflect suboptimal routes to some receivers.
- To optimize these connections, the router may create a *source-based* multicast delivery tree.



PIM-SM — Building the multicast delivery tree

- After the router receives multicast packets through both the source-based delivery tree and the RP-based delivery tree, PIM prune messages are sent towards the RP to prune this branch of the tree.
- When complete, multicast data from the source flows only through the source-based delivery tree.



PIM-SM — Building the multicast delivery tree

- PIM-SM is optimized for
  - environments containing a large number of multicast data streams.
  - Each stream should flow to a relatively small number of the LAN segments.
  - For these groups, the flooding and pruning associated with PIM-DM and DVMRP would be an inefficient use of network bandwidth.

- **PIM-SM benefits**
  - There are few receivers in a group.
  - Senders and receivers are separated by WAN links.
  - The stream of multicast traffic is intermittent.

The Internet multicast backbone (MBONE)

The Internet multicast backbone (MBONE)

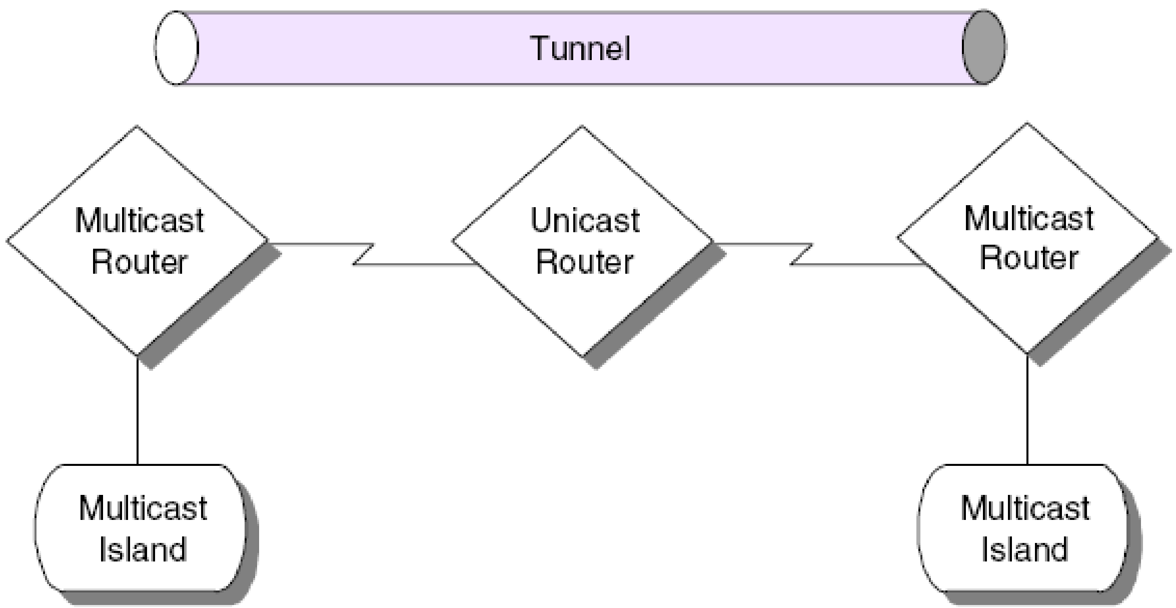
- The Internet multicast backbone (MBONE) was established in March 1992.
- It was initially deployed to provide hands-on experience with multicast protocols.
- The first uses provided audio multicasting of IETF meetings.
- At that time, 20 sites were connected to the backbone.

The Internet multicast backbone (MBONE)

- Two years later, simultaneous audio and video transmissions were distributed to more than 500 participants located in 15 countries.
- Since then, the MBONE has been used to broadcast NASA Space Shuttle missions, rock concerts and numerous technical conferences.
- Commercial and private use of the MBONE continues to increase.

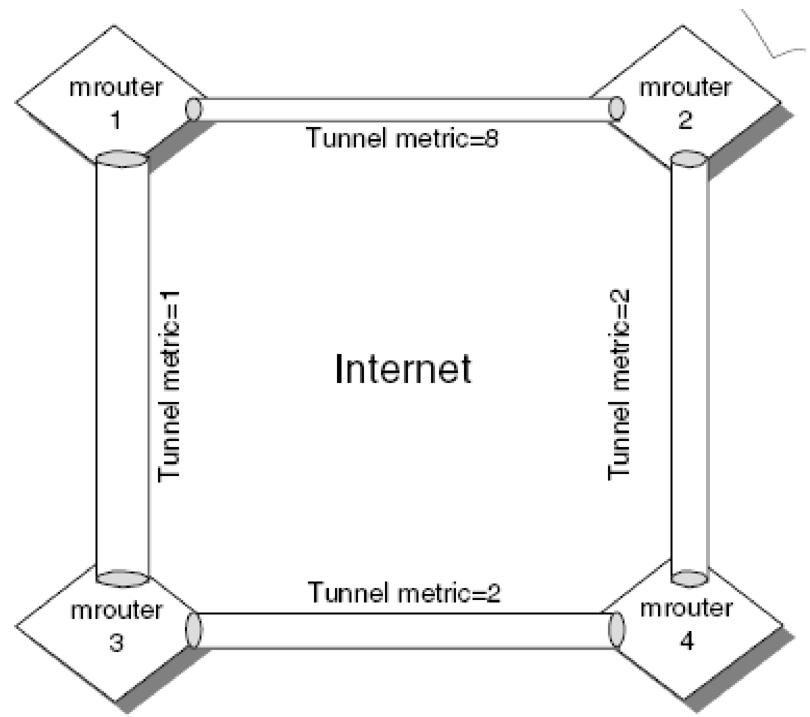
MBONE routing

- Multicast traffic does not flow to every Internet location.
- Until that occurs, the MBONE will consist of a set of multicast network islands.
- These islands are interconnected through virtual tunnels.
- The tunnels bridge through areas that do not support multicast traffic.



MBONE routing

- MBONE tunnels have associated metric and threshold parameters.
- A multicast packet sent from router 1 to router 2 should not use the tunnel directly connecting router 1 and router 2.
- The threshold parameter limits the distribution of multicast packets. It specifies a minimum TTL for a multicast packet forwarded into an established tunnel.



MBONE routing

- In the future, most Internet routers will provide direct support for IP multicast.
- This will eliminate the need for multicast tunnels.
- The current MBONE implementation is only a temporary solution.
- It will become obsolete when multicasting is fully supported in every Internet router.