Multicast Communications and Routing Protocols

Redes de Comunicações II

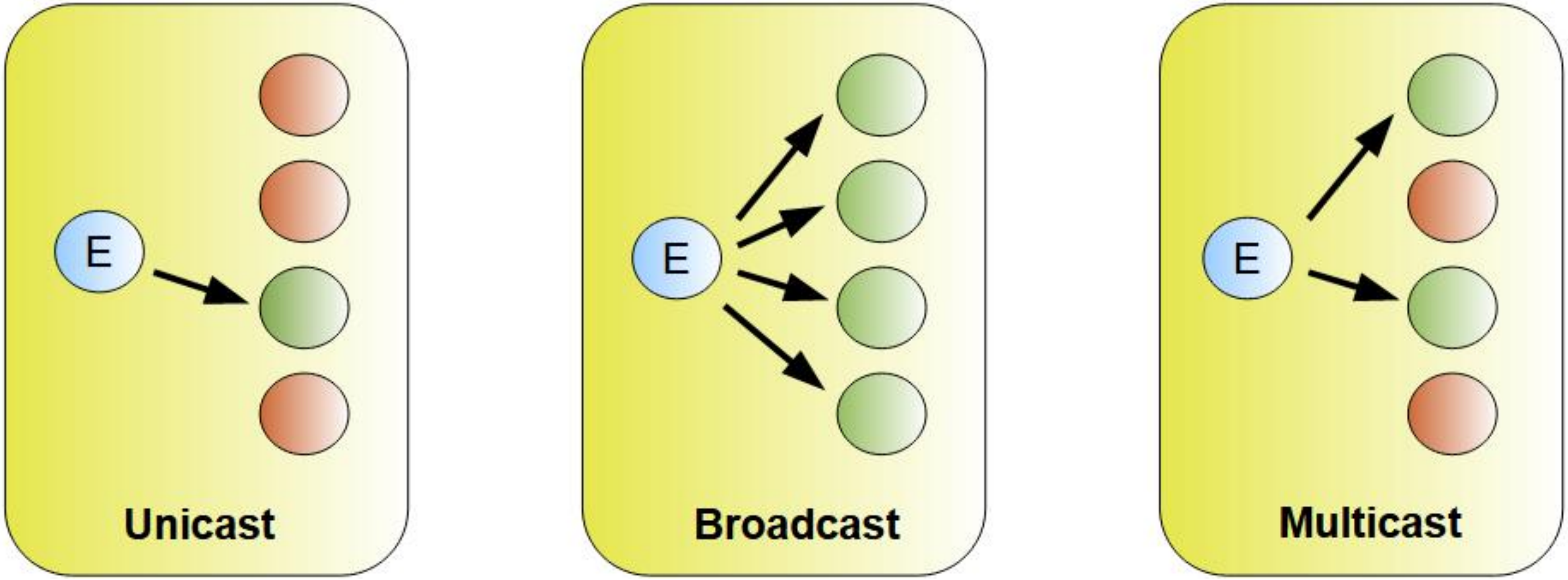
Licenciatura em Engenharia de Computadores e Informática

Prof. Amaro de Sousa (asou@ua.pt)

DETI-UA, 2024/2025

Multicast Communications

- Unicast: one-to-one communications
- Broadcast: one-to-all communications
- Multicast: one-to-many communications
  - multicast at application level.
  - multicast at network level.



Multicast Abstraction

Data stream transmitted by one origin application and destinated to multiple destination applications in different stations.

- **Alternative 1**: The TCP/IP protocolar stack of the sender host establishes a point-to-point connection with each of the destination hosts and sends one copy of the data stream to each destination.
- Advantages:
  - Does not require the network to have multicast capabilities.
  - Allows the use of TCP protocol with all its reliability and congestion control advantages.
- Disadvantages:
  - Requires the sender application to specify a list of destination addresses.
  - There is an inefficient use of network resources (multiple copies of the same data stream are sent from the sender host).

Multicast Abstraction

Data stream transmitted by one origin application and destinated to multiple destination applications in different stations.

- **Alternative 2**: The sender host sends the data once and the network routers copy the data to the multiple destination hosts.
- Advantages:
  - There is a better usage of the network resources.
- Disadvantages:
  - Requires the network to support multicast communications.
  - It can be used only with UDP with its disadvantages (TCP works only in point-to-point communications).
- Issues:
  - How are the multiple destination hosts identified?!
  - How do the routers decide the output interfaces to forward the data?!

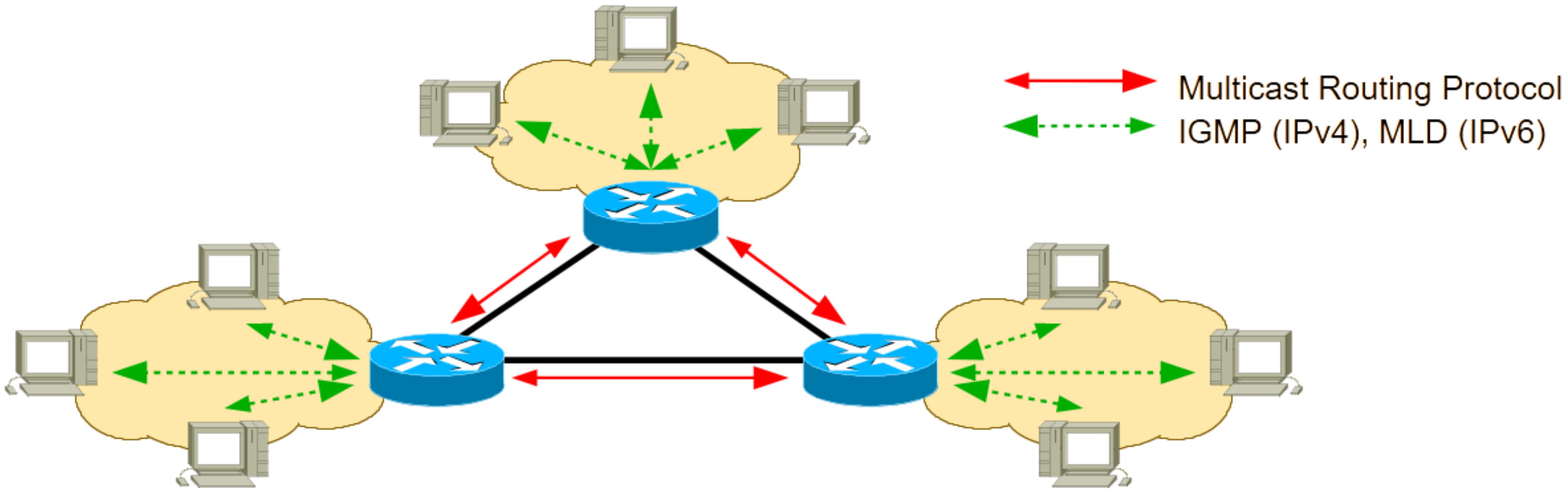
Multicast Abstraction

- In IP multicast networks, a multicast session is identified by a **multicast IP address** agreed among all interested hosts:
  - the sending hosts send IP packets from its unicast IP address to the agreed multicast IP address,
  - the destination hosts accept the incoming IP packets whose destination address is the agreed multicast IP address.
- While IP unicast networks are connectionless networks:
  - unicast IP routing paths are established between routers independently of being used by hosts.
- **IP multicast networks are connection-oriented** networks:
  - a multicast session is created only when there are hosts interested on it,
  - the appropriate multicast IP routing paths are established between routers for the multicast sessions of interest,
  - the multicast IP routing paths are eliminated when the interested destination hosts no longer refresh their requests (soft state).

Types of Multicast Protocols

IP multicast is supported by two types of protocols:

1. a signalling protocol between routers and connected hosts
   - used by destination hosts to declare (to the connected routers) their participation in the session(s) identified by the multicast address(es)
   - IGMP (IPv4), MLD (IPv6).
2. a routing protocol between routers
   - used by routers to establish the multicast routing paths and route to all participating hosts the IP packets sent to the multicast address(es).
   - DVMRP, MOSPF, PIM, etc...



IPv4 Multicast Addresses

	4 bits	28 bits
Class D address	**1110**	**Multicast Group ID**

- Class D addresses range from 224.0.0.0 to 239.255.255.255
- The range 224.0.0.1 to 224.0.0.225 is reserved to routing protocols and other discovery/maintenance protocols. For example:
  - 224.0.0.1 : All Multicast Hosts
  - 224.0.0.2 : All Multicast Routers
  - 224.0.0.5 : All OSPF routers
  - 224.0.0.6 : OSPF designated routers
  - 224.0.0.13 : All PIM routers
- The range 239.0.0.0 to 239.255.255.255 is reserved to private networks

IPv6 Multicast Addresses

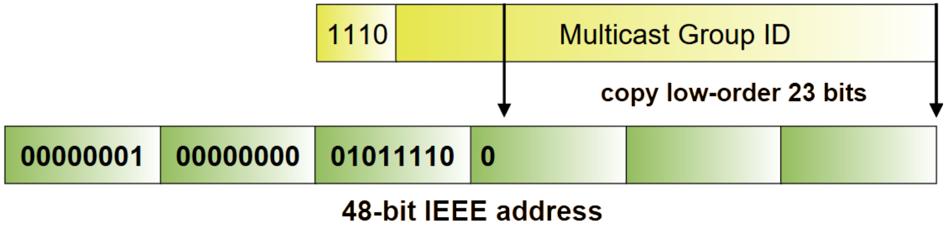
8 bits	4 bits	4 bits	112 bits
1111 1111	**flag**	**scope**	**Multicast Group ID**

- Multicast IPv6 addresses: FF::/8
- The scope field defines the scope of the address:
  - link local, site local, global, etc…
- Like in IPv4, some IPv6 multicast addresses have a standard meaning. For example:
  - FF02::1 : All Multicast Hosts
  - FF02::2 : All Multicast Routers
  - FF02::5 : All OSPFv3 Routers
  - FF02::6 : OSPFv3 designated routers
  - FF02::d : All PIM Routers
  - FF02::16 : MLDv2 reports

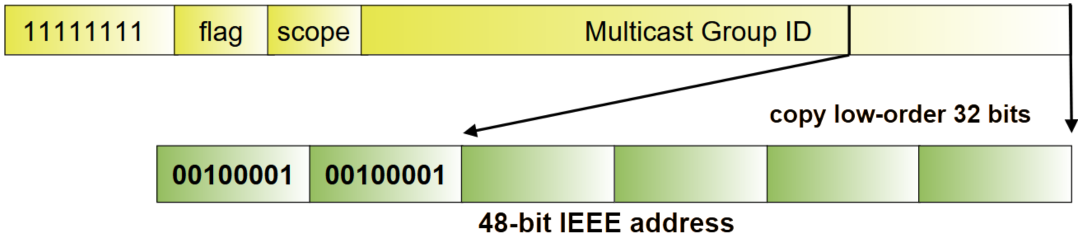
Forwarding IP multicast packets on LANs

- IP multicast addresses are only used as destination addresses in IP packets
- Each IP multicast address has an associated (Layer 2) IEEE address. On each LAN segment:
  - the sender (host or router) of an IP multicast packet uses the associated IEEE address as destination address of the LAN frame,
  - the interested receivers set their NICs to accept LAN frames with the associated IEEE address as destination address.

IPv4 case



IPv6 case

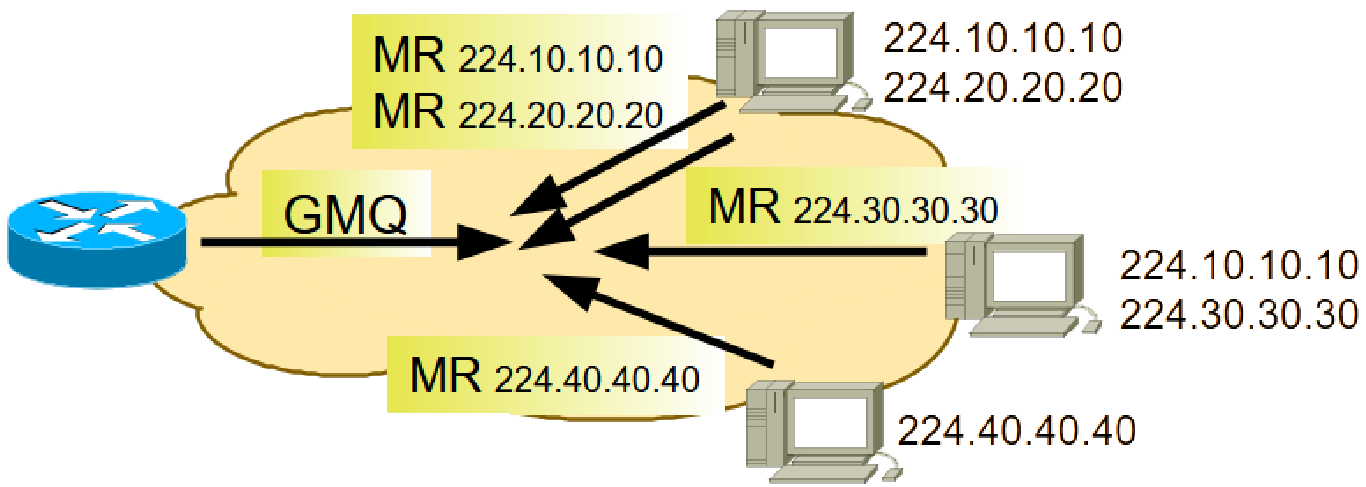


Internet Group Membership Protocol (IGMP)

- IGMP version 2, RFC 2236
  - IGMP runs over IPv4 protocol (protocol type = 0x02).
  - IGMP messages are sent to the destination address 224.0.0.1 ("All Multicast Hosts") with TTL = 1.
- On each network, the Querier Router (QR) is the multicast router with lower IP address on the network.
- There are four types of IGMP messages:
  - GMQ (General Membership Query) – sent by the QR requesting a refresh of the hosts participation in the multicast sessions.
  - MR (Membership Report), sent by the hosts to declare their participation on a multicast session.
  - SMQ (Specific Membership Query), sent by the QR requesting a refresh of the hosts participation in a specific multicast sessions.
  - LGR (Leave Group Report) (optional), sent by the hosts to declare that they are no longer participating in a multicast session.

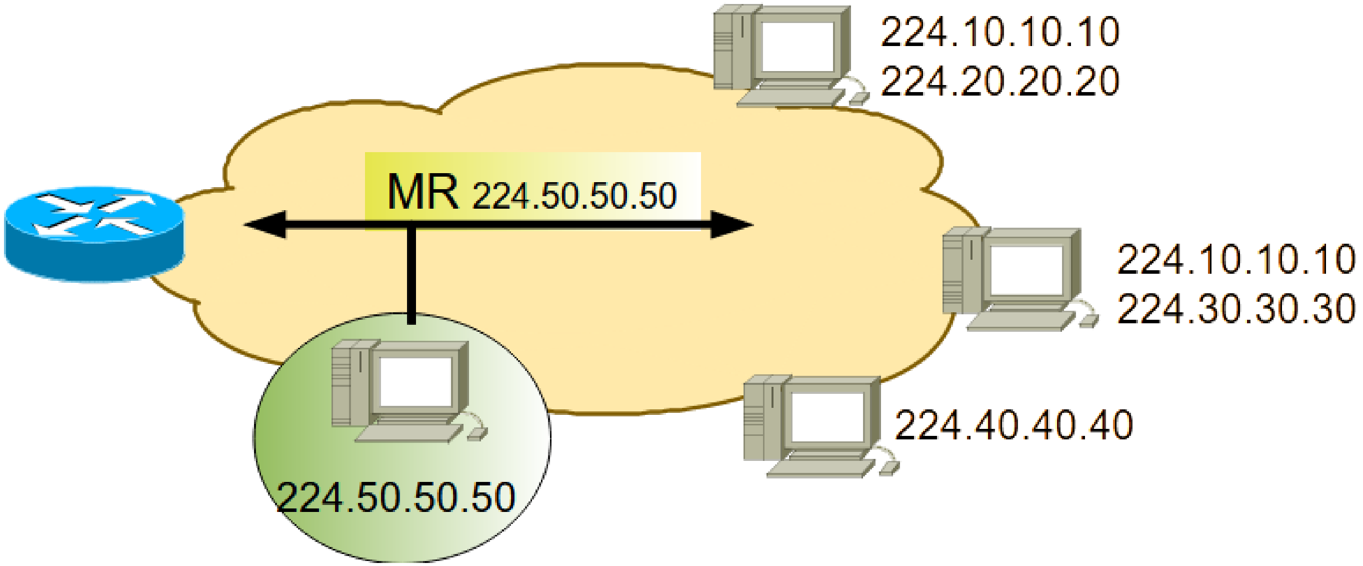
IGMP protocol (I)

- By default, the QR (Querier Router) sends periodically a GMQ specifying a Maximum Response Time (MRT);
  - each host waits a random time between 0 and MRT to reply with a MR specifying its participation on a multicast session,
  - if in the meanwhile the host ‘sees’ a MR for the same multicast session, it aborts the sending of its MR.



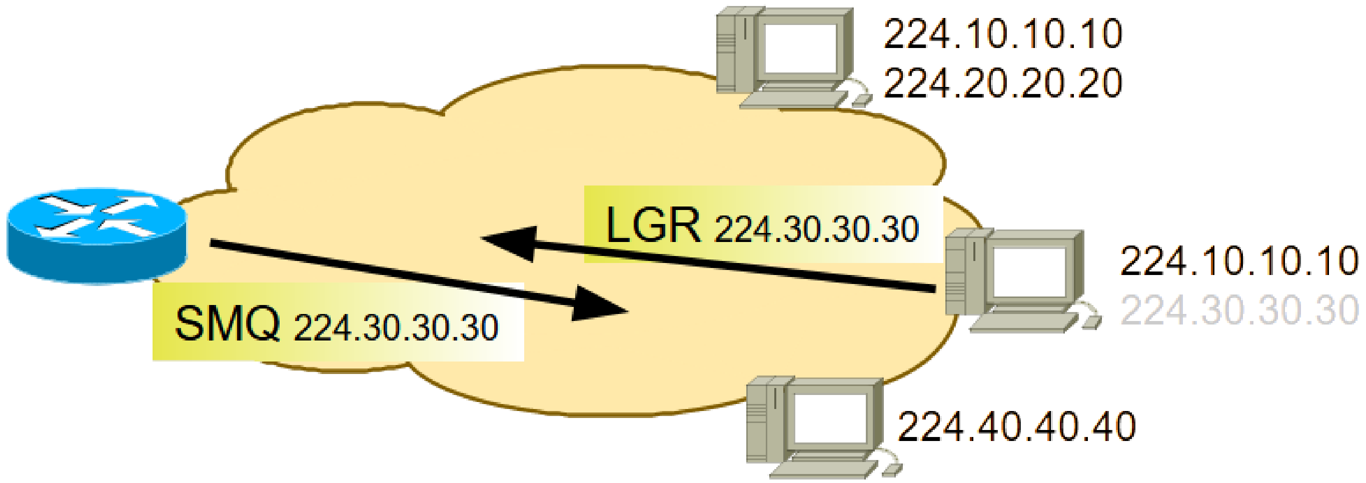
IGMP protocol (II)

- When a host wants to start participating in a multicast session:
  - it immediately sends a MR with the multicast address,
  - i.e., it does not wait for the reception of a GMQ from the QR.



IGMP protocol (III)

- When a host no longer wants to participate in a multicast session:
  - the host sends a LGR (which is optional) or does nothing,
  - when the QR receives a LGR, it sends a SMQ to check if there are still hosts participating in that session.



IGMP version 3

- IGMPv3 adds support to "source filtering“, meaning to define the sources from which the destination host is interested in the specified multicast session.
- Allows a host to declare its participation on a multicast session from :
  - ONLY specific source addresses (INCLUDE Mode).
  - ALL sources EXCEPT specific source addresses (EXCLUDE Mode).
- Has a new “Membership Report” message format
  - that allows source filtering
  - and simultaneous requests to multiple multicast sessions.
- Allows interoperability with the previous versions (IGMPv1 and IGMPv2).

Multicast Listener Discovery (MLD)

- Multicast Listener Discovery (MLD) protocol is the equivalent to IPv6 networks to the IGMP to IPv4 networks.
- It is an adaptation of IGMP to the IPv6 semantics:
  - MLD version 1 is equivalent to IGMP version 2
  - MLD version 2 is equivalent to IGMP version 3
- MLD is embedded in ICMPv6 instead of being a separate protocol (IGMP is a protocol of its own).
- MLD messages are sent with IPv6 link-local addresses as sources.
- MLD messages:
  - Multicast Listener Query (equivalent to GMQ and SMQ in IGMP)
  - Multicast Listener Report (equivalent to MR in IGMP)
  - Multicast Listener Done (equivalent to LGR in IGMP).

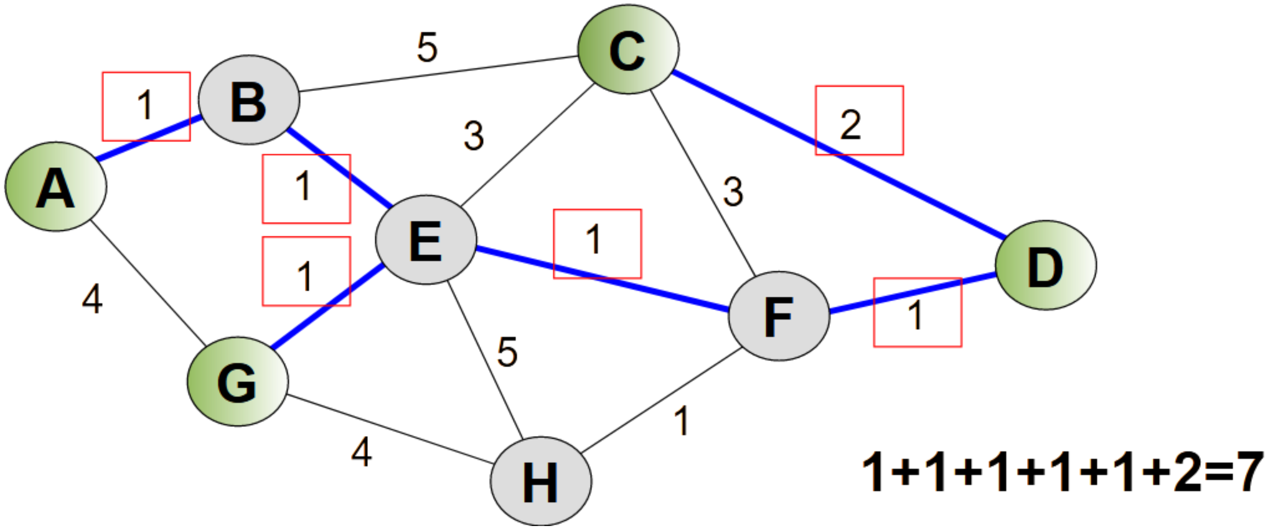
Multicast Routing Principles

- Multicast routing is based on trees:
  - **Multicast Tree (MT)**: a set of links of an IP network connecting a given set of multicast routers and defining a single routing path between any pair of routers (i.e., not containing cycles)
- ***Group-shared tree***: based on determining a routing tree per multicast session that connects all routers with hosts participating in the session.
  - Minimum cost Steiner tree (not used in practice)
  - Minimum cost tree from a central node ("*rendezvous point*").
- ***Source-based tree***: based on determining a routing tree, per multicast session and sender host.
  - Minimum cost tree from the multicast router connected to the sender host.

Group-Shared Tree

- **Minimum cost Steiner tree**: determines the set of links with minimum cost that interconnects all multicast routers with hosts participating in a multicast session.
- Not used in practice due to:
  - known computing algorithms have exponential complexity,
  - it requires a link-state protocol (with the topology of the network),
  - it requires being executed every time a multicast router joins/leaves a session.

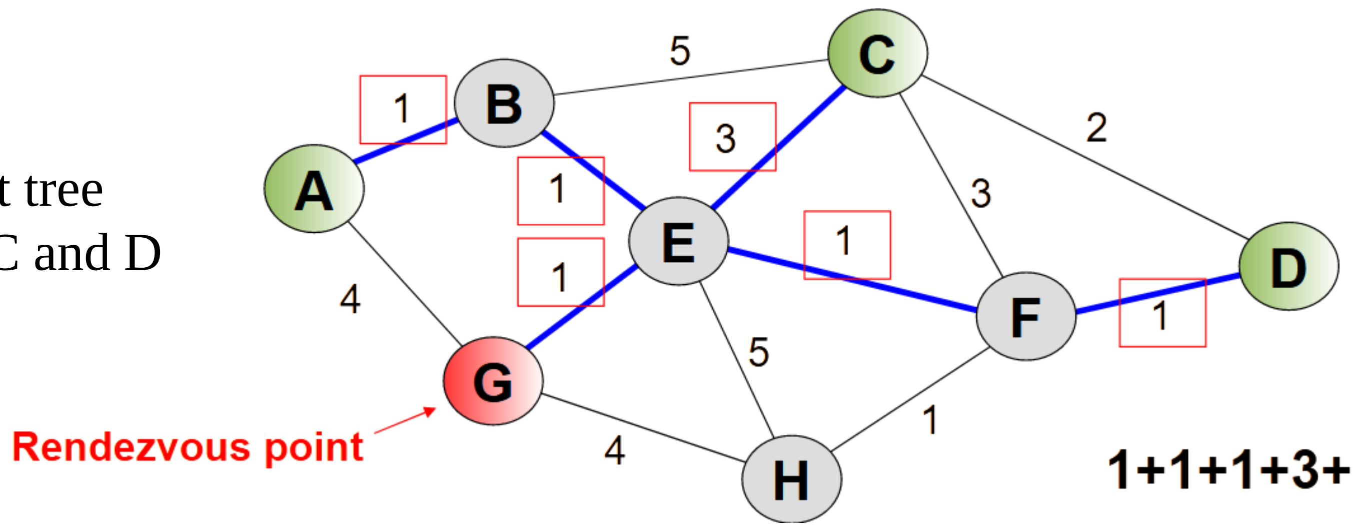
Minimum cost Steiner tree between routers A, C, D and G



Group-Shared Tree

- **Minimum cost tree from a RP** (Rendezvous Point)
  - RP is previously chosen and known by all multicast routers.
    - The RP always belongs to the multicast tree, even if it does not have hosts participating in the session
  - The tree is composed by the links of the shortest paths from the RP to each multicast router with participating hosts.

Minimum cost tree from G to A, C and D

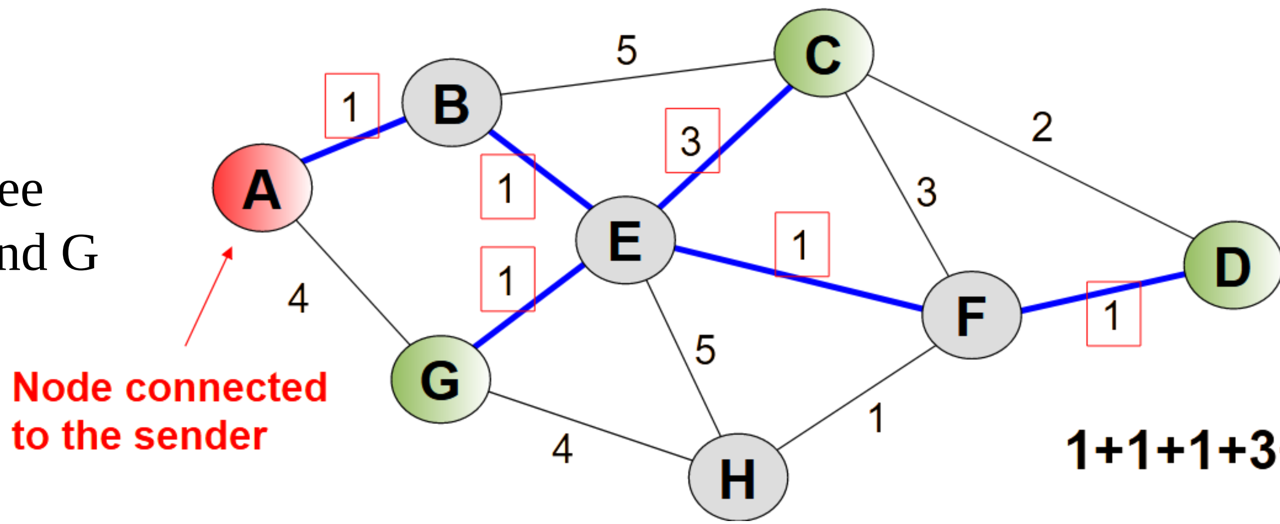


1+1+1+3+1+1=8

Source-Based Tree

- The tree is composed by the links of the shortest paths from the source multicast router (i.e., the multicast router connected to the sender host) to each multicast router with participating hosts.

Minimum cost tree
from A to C, D and G



1+1+1+3+1+1=8

Establishment of a Multicast Tree

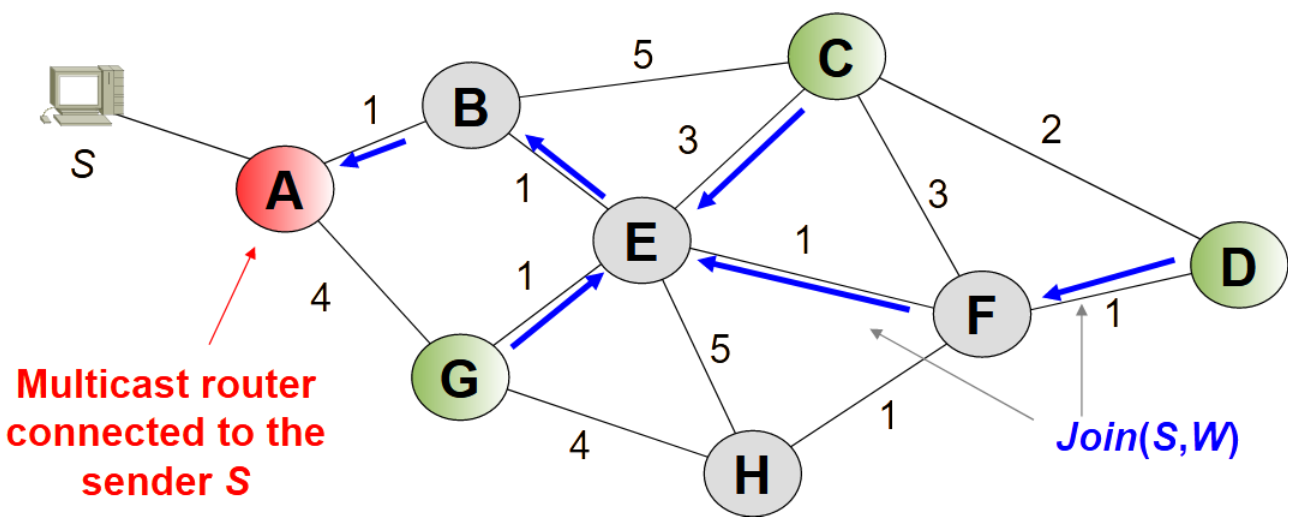
A Multicast Tree (MT) can be established in one of two ways:

1. Receiver-Initiated MT
   - The establishment of the MT is initiated only when receiver hosts declare their participation on a multicast session
   - The routers connected to hosts participating in the multicast session establish the MT and then, the multicast packets start being forwarded
2. Sender-Initiated MT
   - The establishment of the MT is initiated by the sender host(s) of the multicast sessions
   - The multicast packets start being flooded to the whole network and then, routers eliminate themselves from the MT if they do not need to forward the multicast packets

- Group-Shared Trees are always Receiver-Initiated MTs
- Source-Based Trees can be either Receiver-Initiated or Sender-Initiated MTs

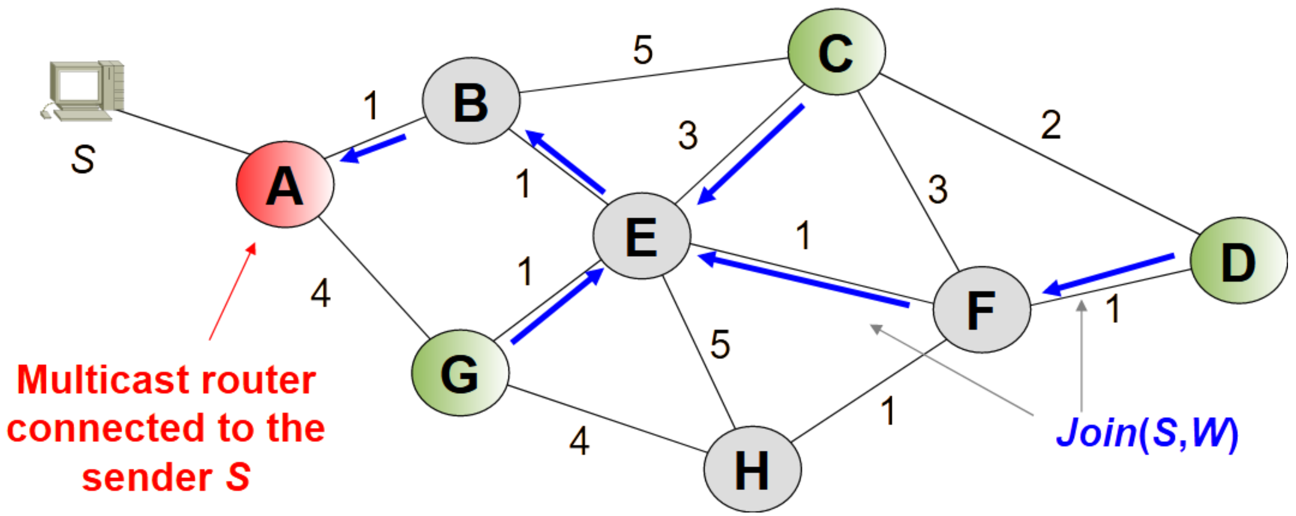
<u>Receiver-Initiated</u> Source-Based Tree (I)

- A multicast router with a first host (in an incoming interface F_{in}) declaring its participation in a multicast session W from a sender with address S:
  - if the entry (S,W) <u>does not exist</u> in its multicast routing table:
    - the router sends a *Join*(S,W) message towards the address S through an outgoing interface F_{out} of the minimum cost path (unicast path),
    - the entry (S,W) is created with the indication that packets to multicast address W received in F_{out} are forwarded to F_{in}.
  - if the entry (S,W) <u>exists</u> in its multicast routing table,
    - the entry (S,W) is updated with the indication that packets to multicast address W (received in F_{out}) are also forwarded to F_{in}.



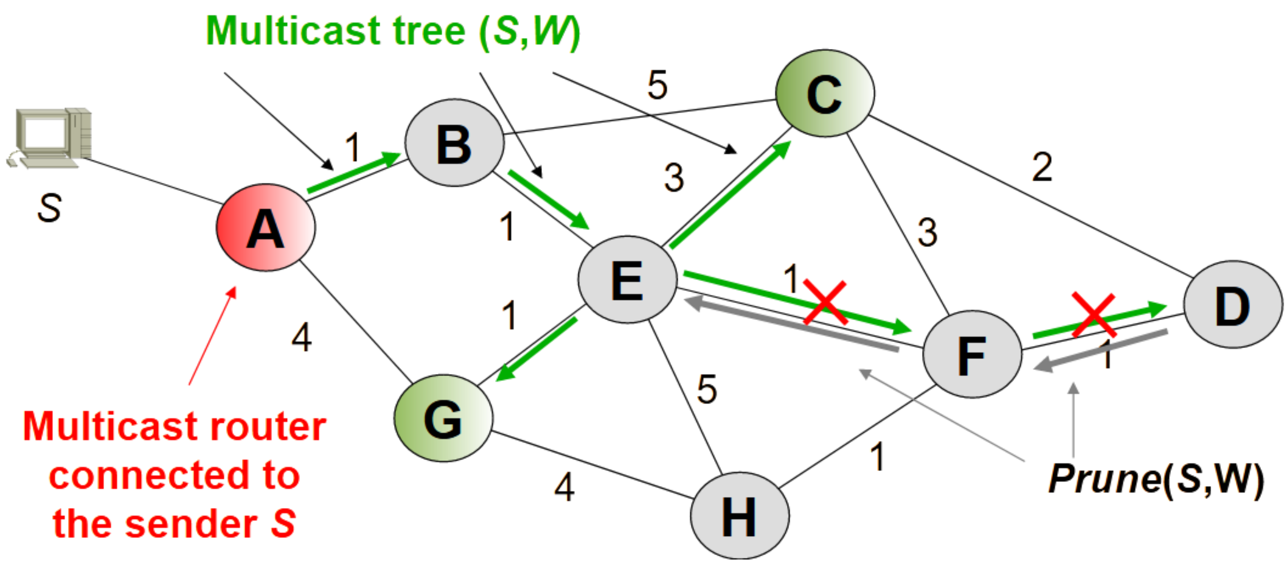
Receiver-Initiated Source-Based Tree (II)

- A multicast router receiving a *Join*(S,W) message (from another multicast router) in an incoming interface F_{in}:
  - if the entry (S,W) <u>does not exist</u> in its multicast routing table:
    - the router sends a *Join*(S,W) message towards the address S through an outgoing interface F_{out} of the minimum cost path (unicast path),
    - the entry (S,W) is created with the indication that packets to multicast address W received in F_{out} are forwarded to F_{in}.
  - if the entry (S,W) <u>exists</u> in its multicast routing table,
    - the entry (S,W) is updated with the indication that packets to multicast address W (received in F_{out}) are also forwarded to F_{in}.



Receiver-Initiated Source Based Tree (III)

- When in a multicast router, all hosts stop participating in an established multicast tree (S,W), it sends a *Prune*(S,W) message towards address S.
- In the path of the *Prune*(S,W) message, each multicast router receives the message in the incoming interface F_{in} and:
  - the router eliminates interface F_{in} of its multicast entry (S,W),
  - the router eliminates the entry (S,W) if F_{in} is the only current forwarding interface of this multicast entry.
- When there are multiple sending hosts, one multicast tree is established per sender.



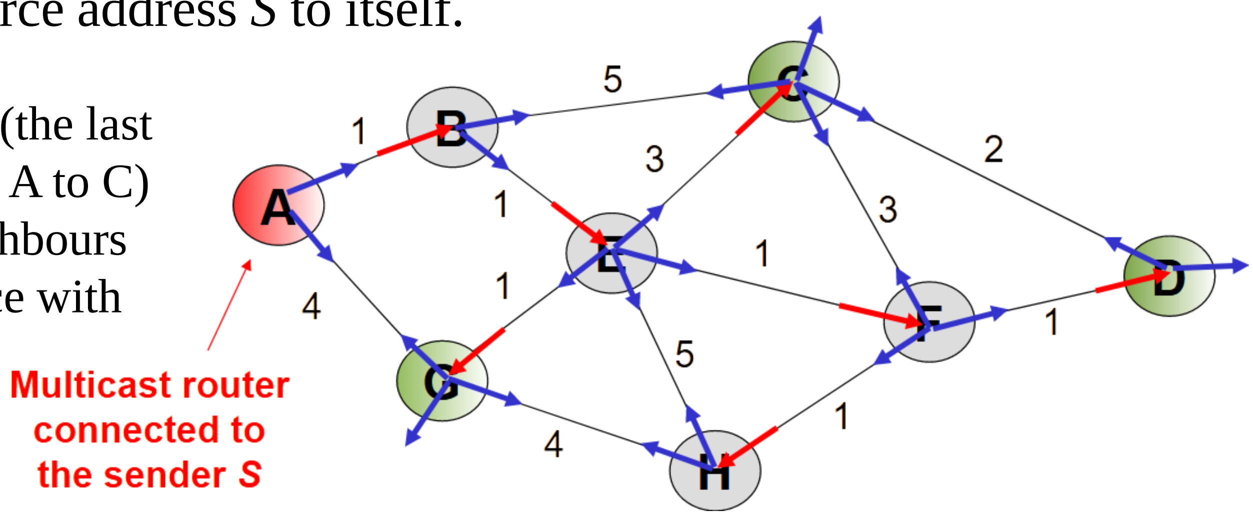
<u>Sender-Initiated</u> Source-Based Tree (I)

"Reverse Path Forwarding with Pruning"

- The multicast router connected to the sender host forwards the packets:
  - to all its neighbour multicast routers and
  - to its interfaces with hosts participating in the multicast session.
- Each of the other multicast routers forward the packets:
  - to all its other neighbour multicast routers and
  - to its interfaces with hosts participating in the multicast session.

  if the packets are received from the neighbour router providing the minimum cost path from the source address *S* to itself.

In C, packets received from E (the last router in the routing path from A to C) are forwarded to all other neighbours (B, D and F) and to its interface with participating hosts



Sender-Initiated Source-Based Tree (II)

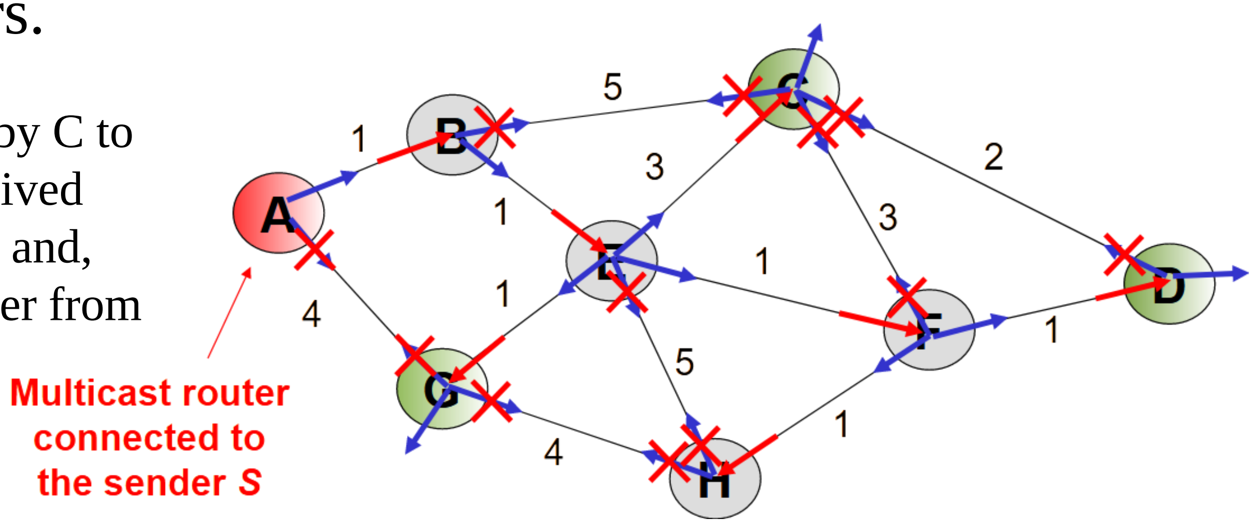
"Reverse Path Forwarding with Pruning"

- When a multicast router receives a packet from a neighbour which is not the last router in the routing path from the source *S* to itself
  - it stops forwarding the packet to him

  because it knows it is not the last router in the path from *S* to the neighbour router.

At this moment, a multicast tree is established from the sender multicast router (i.e., the router connected to the sender host) to all other multicast routers.

Packets stop being forwarded by C to B, D and F because C has received packets from these neighbours and, therefore, C is not the last router from A to each of them



Sender-Initiated Source-Based Tree (III)

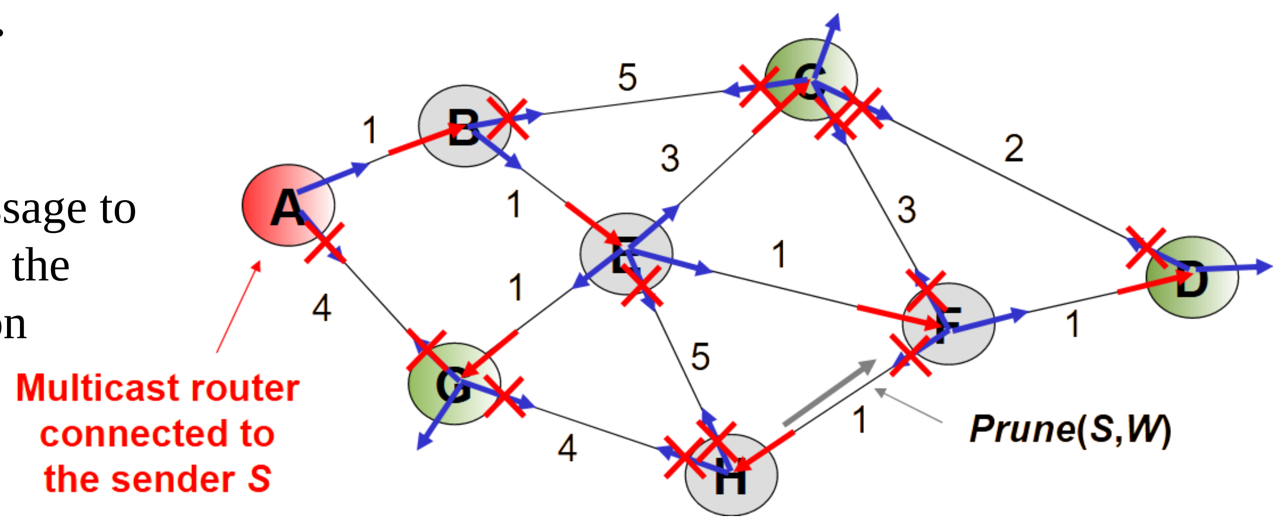
"Reverse Path Forwarding with Pruning"

- A multicast router without:
  - interfaces with hosts participating in the multicast session and
  - neighbour multicast nodes to forward the multicast packets

  sends a *Prune*(*S*,*W*) message to the neighbour multicast router from which it accepts the incoming multicast packets.

At this moment, a multicast tree is established from the sender multicast router to all multicast routers with hosts participating in the multicast session.

Router H sends the Prune message to neighbour F to stop receiving the packets of the multicast session



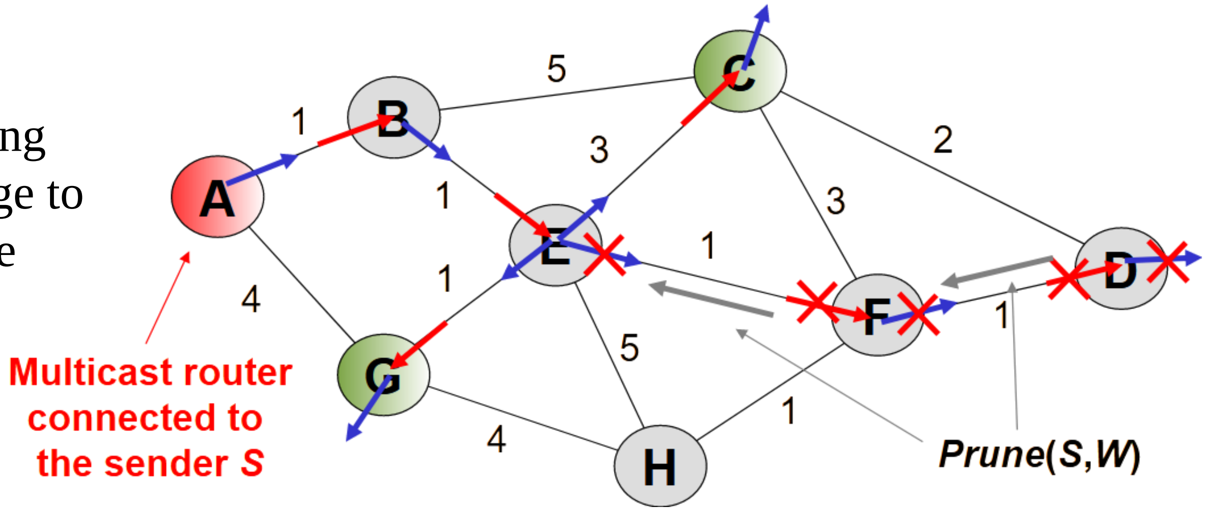
Sender-Initiated Source-Based Tree (IV)

"Reverse Path Forwarding with Pruning"

- When in a multicast router:
  - all connected hosts stop participating in the multicast session and
  - it does not have neighbour multicast nodes to forward the multicast packets

  the router sends a *Prune*(*S*,*W*) message to the neighbour multicast router from which it accepts the incoming multicast packets.
- This behaviour is repeated by the other routers in the path of the *Prune*(*S*,*W*) message.

Router D has no more participating hosts and sends the Prune message to neighbour F, which also sends the Prune message to neighbour E

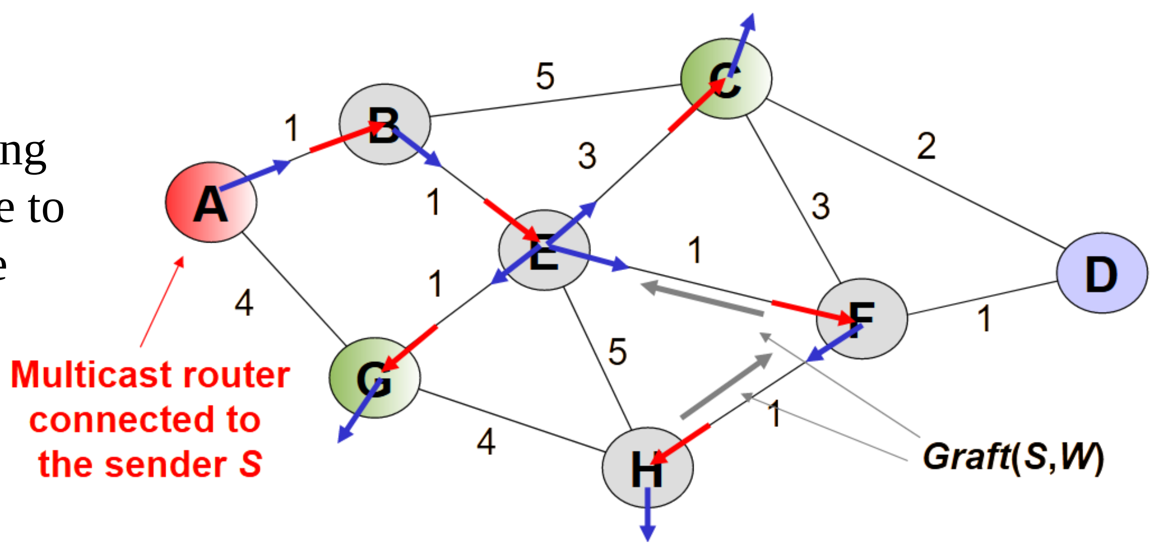


Sender-Initiated Source-Based Tree (V)

"Reverse Path Forwarding with Pruning"

- When a host connected to a multicast router starts participating in a multicast session, if the router is not currently in the multicast tree, it sends a *Graft*(*S*,*W*) message towards the neighbour which is the last router in the path from the source *S* to itself.
- This behaviour is repeated by the other routers in the path of the *Graft*(*S*,*W*) message.

Router H starts having participating hosts and sends the Graft message to neighbour F, which also sends the Graft message to neighbour E



Additional Considerations

- Graft messages in Sender-Initiated Source-Based Trees:
  - Usually, the initial flooding of multicast packets from the sender multicast router (i.e., the router connected to the sender host) to all other routers is repeated periodically.
    - Then, the Prune messages are again used to reach the final MT.
  - In this case, Graft messages (used to reconnect a multicast router to the MT) can be optional. If not used, the interested multicast router waits for the next flood moment and does not remove itself from the MT.
- Establishment of Group-Shared Trees
  - Group-Shared Tree is always of type Receiver-Initiated MT.
  - The establishment process is the same as the one described for the Receiver-Initiated Source-Based Trees
  - The difference is that the address *S*, instead of being the unicast address of the sender host, it is one of the unicast addresses of the RP (Rendezvous Point).

Distance Vector Multicast Routing Protocol (DVMRP)

- DVMRP is a “Distance Vector” type of protocol, similar to RIP
- DVMRP establishes the multicast trees based on Sender-Initiated Source-Based Trees
  - Distance vectors represent the cost from each possible source.
  - For each source, each router announces periodically (through Report messages) to its neighbour routers its cost from the source.
  - For each source, a router receiving Report messages selects the neighbour router announcing the lowest cost and adds the entry to its multicast routing table.
  - Routers use Prune messages (with an associated lifetime) to remove themselves from source-based trees and Graft messages to rejoin themselves to source-based trees.
- Interfaces are configured with a cost metric (default value is 1) and a threshold TTL that limits the scope of the multicast sessions
  - a multicast router forwards a multicast packet through an interface if the TTL in its header is larger than the interface threshold TTL.
- DVMRP allows tunnels between multicast routers

Multicast Open Shortest Path First (MOSPF)

- MOSPF is an extension of OSPF, and therefore, it is a “Link-State” type of protocol
- MOSPF establishes the multicast trees based on Receiver-Initiated Source-Based Trees
  - Through the link-state database, each router is able to compute the source-based tree from whatever source on each active multicast session.
- Each router has a local database of multicast groups containing a list of the directly connected group members (signalled by IGMP)
  - Instead of using Join and Prune messages, the local router sends a group-membership LSA to all other routers.
  - The local router forwards the multicast packets to the group members based on the group-membership LSAs received by all routers.
- MOSPF does not support tunnels
  - In the OSPF messages there is a flag indicating if the router supports multicast.
  - The routers not supporting multicast will not belong to the source-based trees.
  - MOSPF requires any 2 multicast routers to have at least one path where all intermediate routers support multicast.

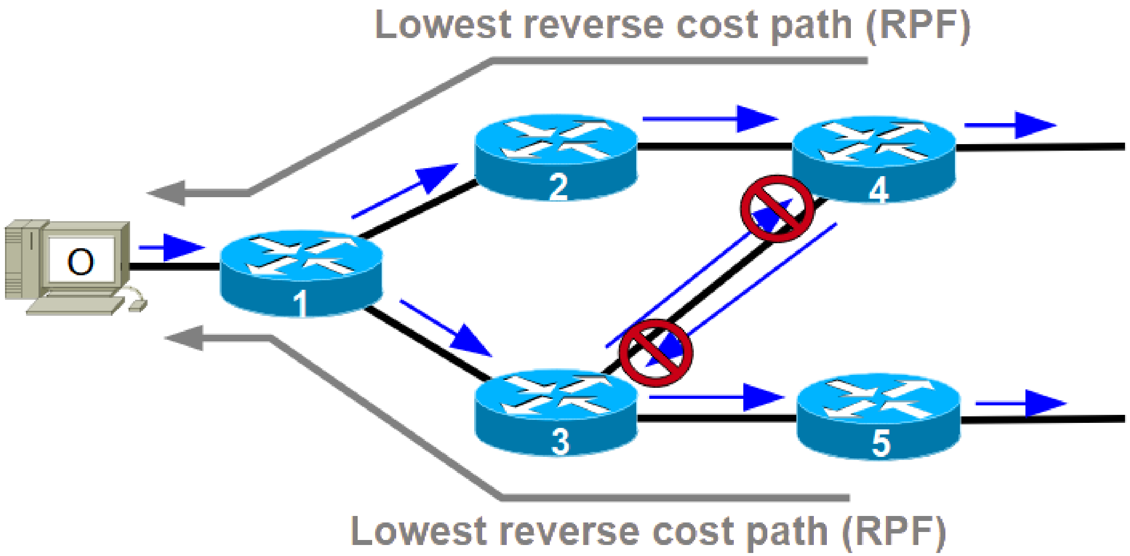
Protocol-Independent Multicast (PIM)

- PIM is a family of multicast routing protocols.
- PIM does not include a topology discovery mechanism but uses instead the routing tables built by a unicast routing protocol.
  - PIM is not dependent on any specific unicast routing protocol;
  - it makes use of any unicast routing protocol in use on the network.
- Initially, PIM defined two protocols used in two extreme cases:
  - PIM Dense Mode (PIM-DM)
    - To be used when most of the networks contain hosts participating in multicast sessions and, consequently, most of the routers need to route multicast packets.
  - PIM Sparse Mode (PIM-SM)
    - To be used when hosts participating in multicast sessions are on a reduced number of networks and, consequently, the number of routers that need to route multicast packets is small when compared to the total number of routers.

PIM Dense Mode (PIM-DM)

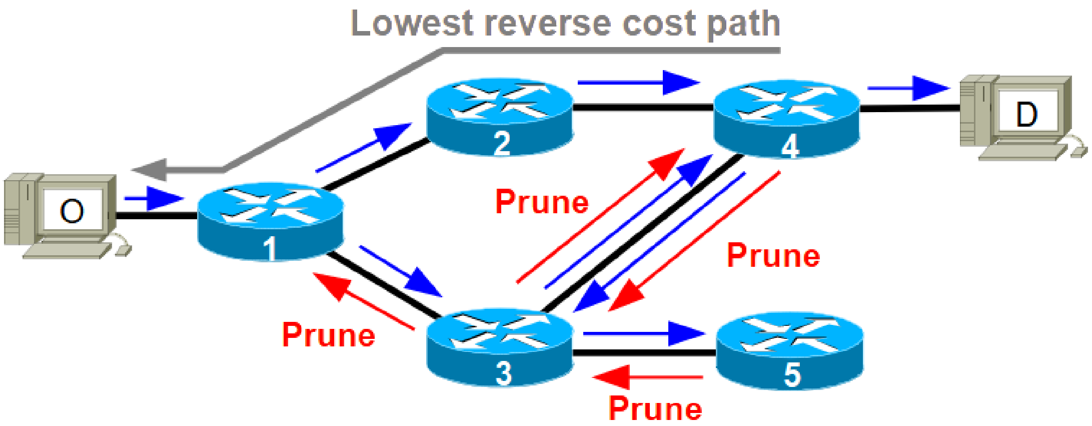
- PIM-DM establishes the multicast trees based on Sender-Initiated Source-Based Trees
  - It requires all routers to have the protocol active.
  - Routers forward packets to all neighbour routers that did not send Prune messages.
  - Routers use Prune messages to signal neighbour routers that they should not forward packets to them.
- Unicast routing tables do not allow to determine to which neighbour router is the last router from the source
  - It assumes that unicast routing paths are symmetric.
  - It forwards multicast packet received from the next-hop router towards the source address (this information is in any unicast routing table).
  - In case of multiple minimum cost paths, it selects the next-hop router with the highest IP address interface (in multicast trees, there is no ECMP).

PIM Dense Mode – initial flooding

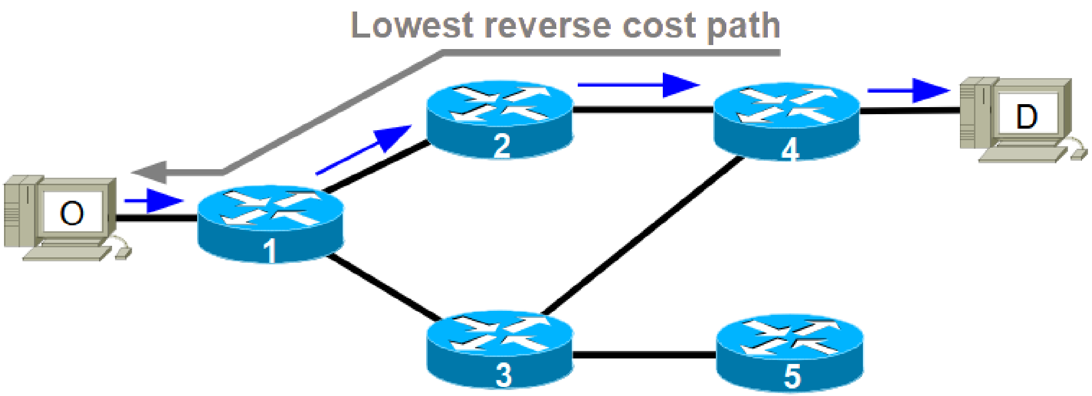


- When a router receives multicast packets in the interface that provides the lowest reverse cost path to the source (RPF interface), it forwards the packets to all other interfaces.
- When a router receives multicast packets in an interface that does not provide the lowest reverse cost path to the source (not the RPF interface), it discards the packets.

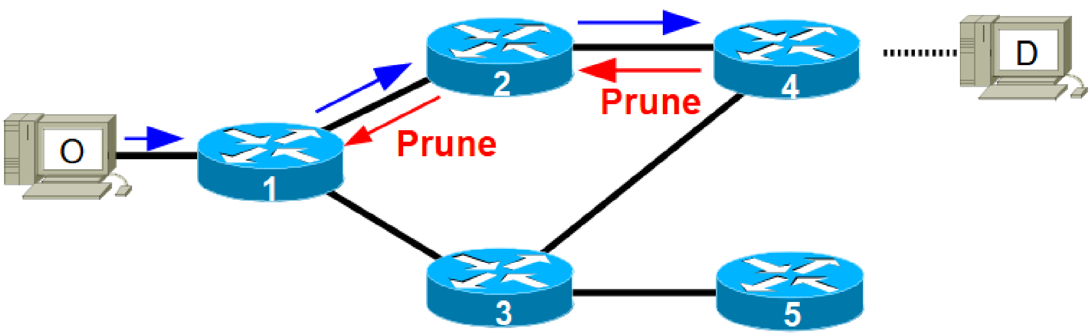
PIM Dense Mode
use of Prune messages



1. Prune messages are sent to avoid multicast packets being forwarded to routers that do not need to receive them.

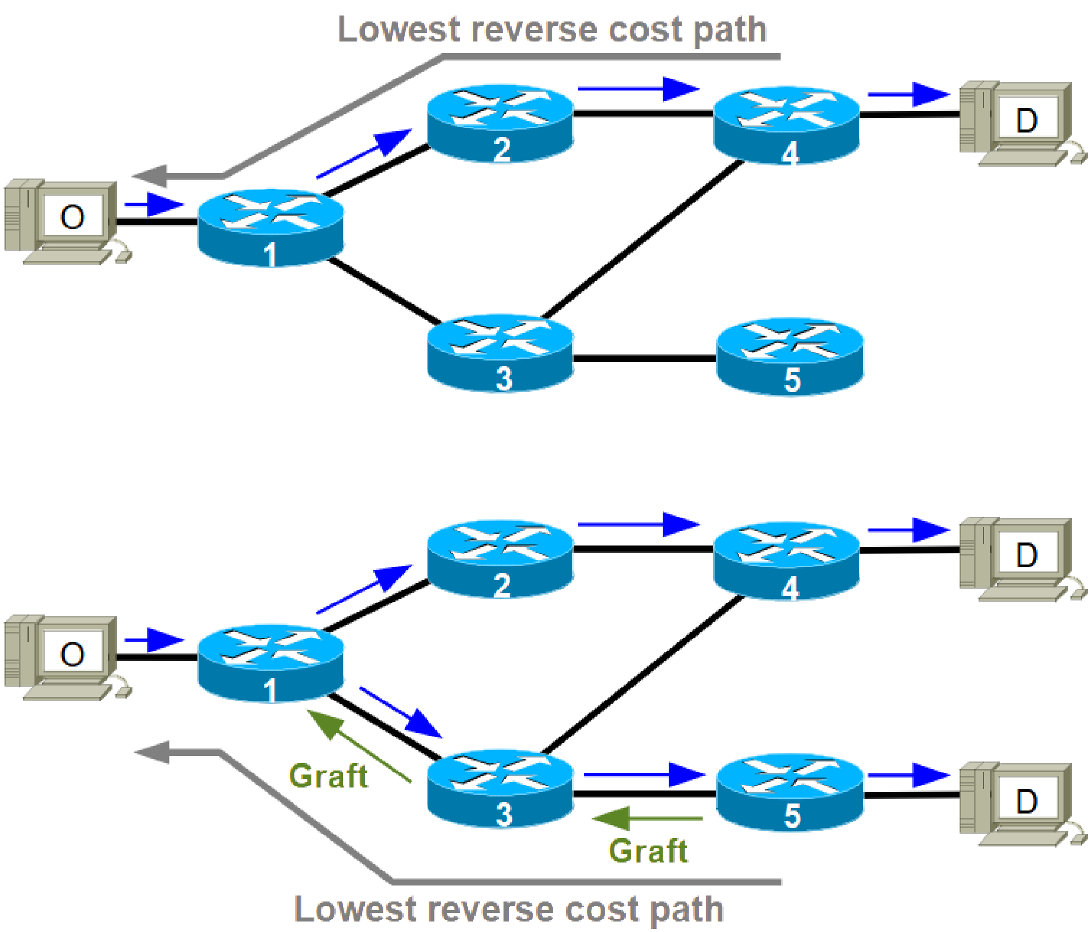


2. In this example, the result is a multicast tree from the active source O to the active receiver D.



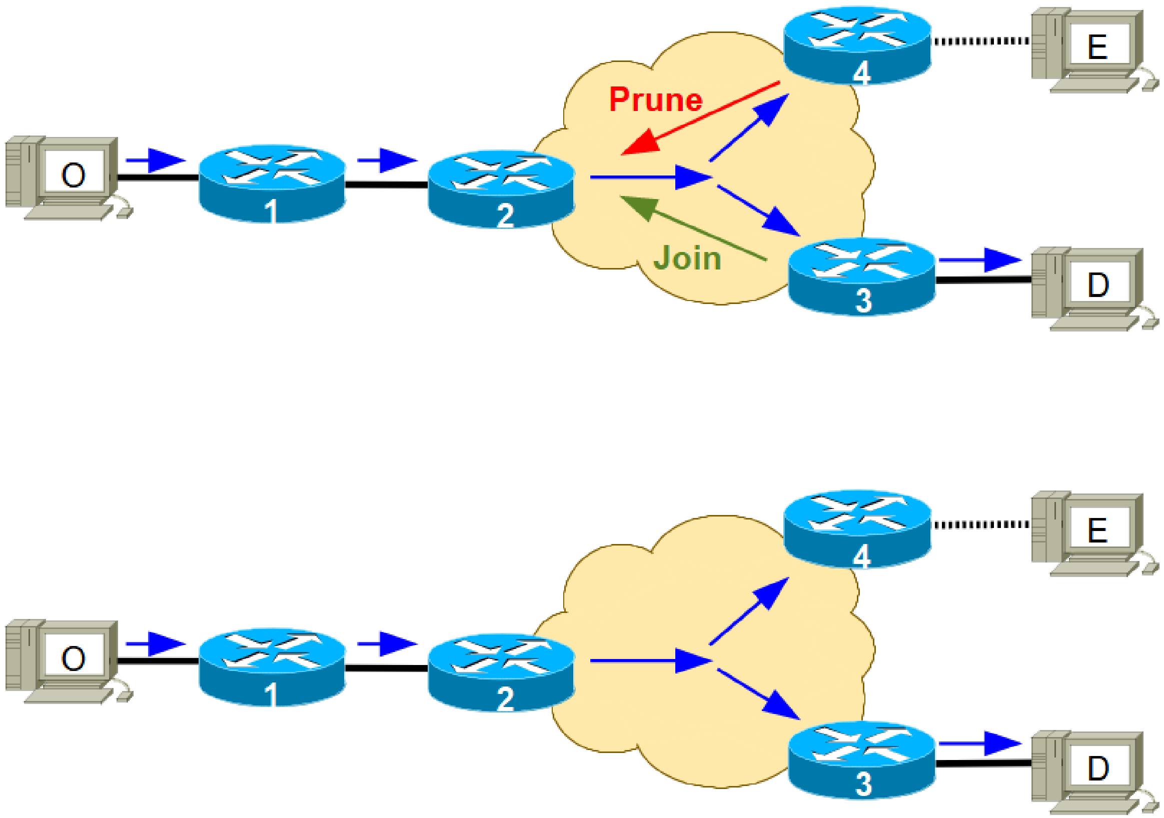
3. If receiver D abandons its participation in the multicast session, again Prune messages are used to update the multicast tree.
4. In this example, the multicast tree becomes empty as there are no participating hosts.

PIM Dense Mode – use of Graft messages



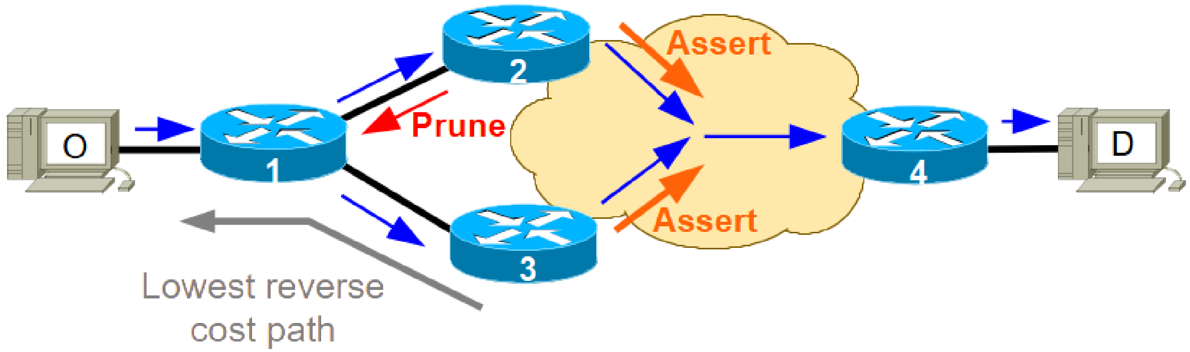
- A router restarts receiving multicast packets by sending a Graft message to join again to the multicast tree.

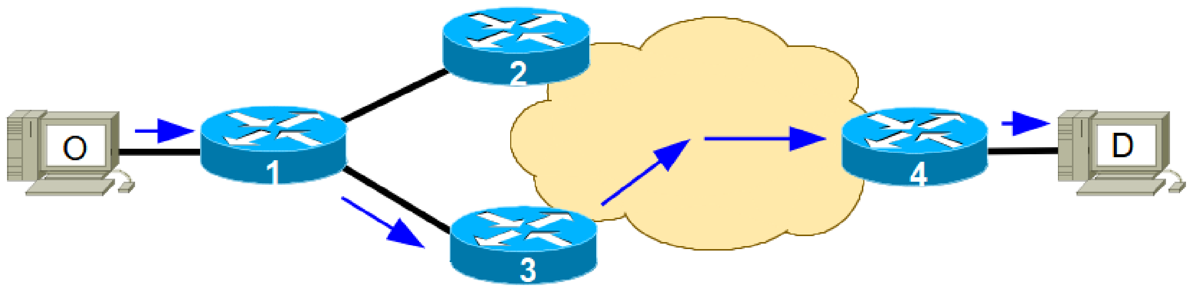
PIM Dense Mode – use of Join messages



- When a Prune message is sent through a shared medium (e.g., LAN), and other routers need to be kept in the multicast tree, they send a Join message to make null the effect of the Prune message.

PIM Dense Mode – use of Assert messages





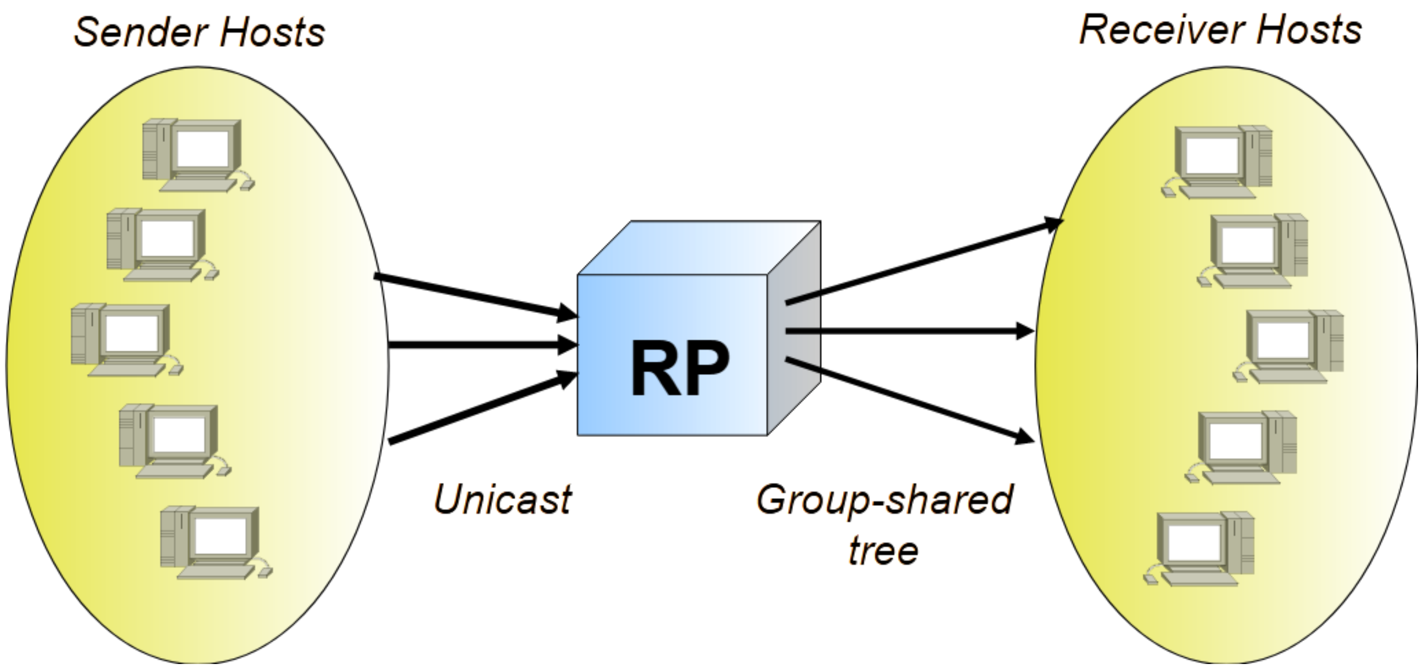
- When there are multiple routers sending multicast packets to a shared medium (e.g., LAN), they must decide which one keeps sending the packets.
- All routers send to the shared medium an Assert message with their unicast cost to the source.
- The chosen router is the one with the lowest cost and, in case of tie, the one with the largest IP address (in the interfaces connected to the shared medium).

PIM Sparse Mode (PIM-SM)

- PIM Dense Mode (PIM-DM) is a data-driven protocol.
  - Routers that do not have multicast clients must periodically send Prune messages to avoid receiving multicast packets.
- PIM Sparse Mode (PIM-SM) is a receiver-driven protocol.
  - Routers with multicast clients must periodically announce (with Join messages) that they want to join to the multicast tree to start receiving the multicast packets.
- PIM-SM starts by establishing a Group-Shared Tree triggered by interested receiver hosts.
  - The RP (Rendezvous Point) address of the Group-Shared Tree can be administratively configured in all routers.
  - There are also automatic mechanisms, such as CISCO RP Discovery Protocol, that uses the multicast address 224.0.1.40.
  - There may be different RPs for different multicast address ranges.
- Then, PIM-SM establishes Source-Based Trees when active sender hosts appear (for routing optimization).

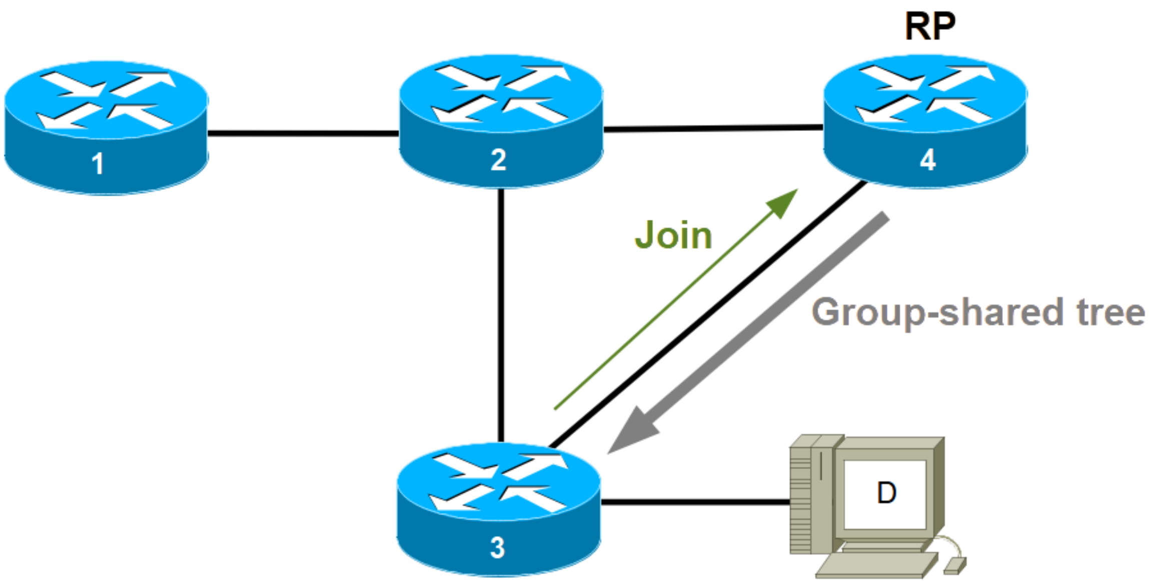
PIM Sparse Mode (PIM-SM)

- At the beginning, multicast packets from sender hosts are forwarded to the RP through unicast routing (using tunnelling).
- At the RP, the multicast packets are forwarded to all interested receiver hosts through the initial established group-shared tree.



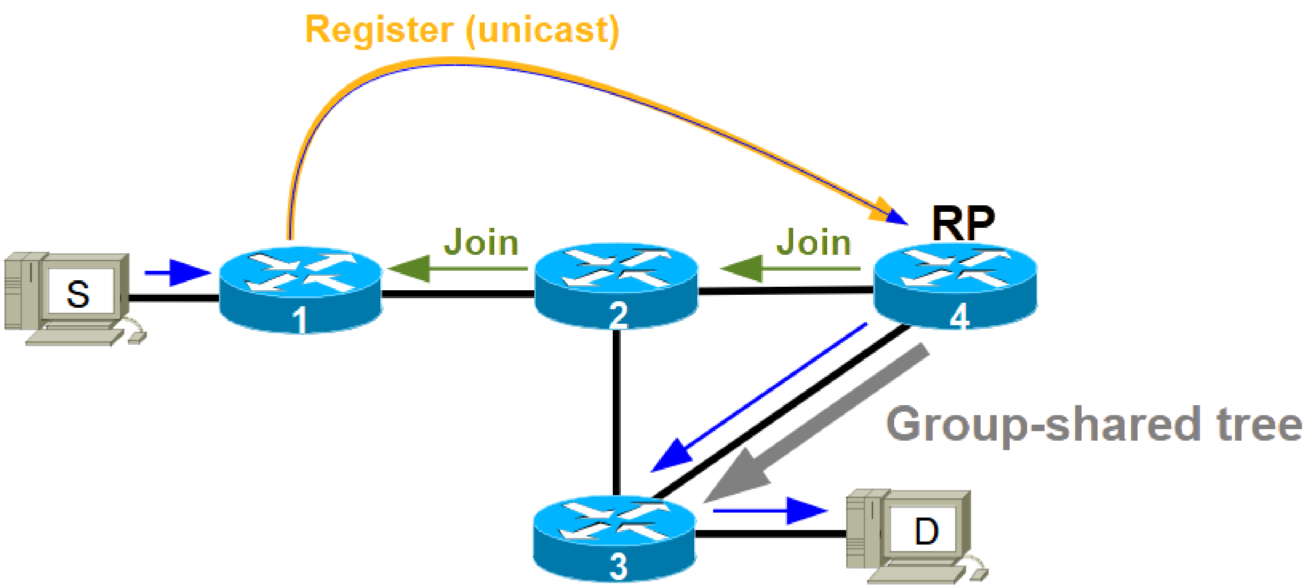
PIM-SM – establishing the initial Group-Shared Tree

- The routers with interested receiver hosts to a specific multicast session send a Join message to the RP in order to join the Group-shared tree (which root is the RP).



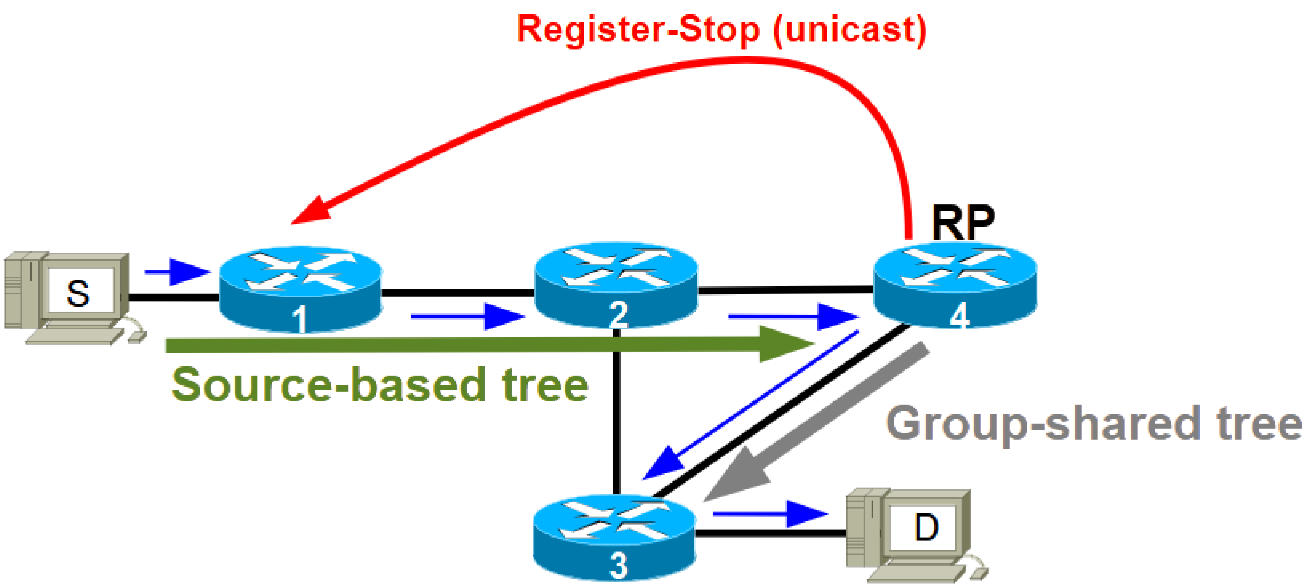
PIM-SM – an active sender host appears (I)

- When a new multicast sender host appears, the source router encapsulates each multicast packet in a PIM Register message and sends the Register message in unicast to the RP.
- The RP extracts the multicast message from the PIM Register message and:
  – forwards the multicast packet through the Group-shared tree,
  – upon receiving the first PIM Register message, sends a Join message towards the source address to establish a Source-based tree.

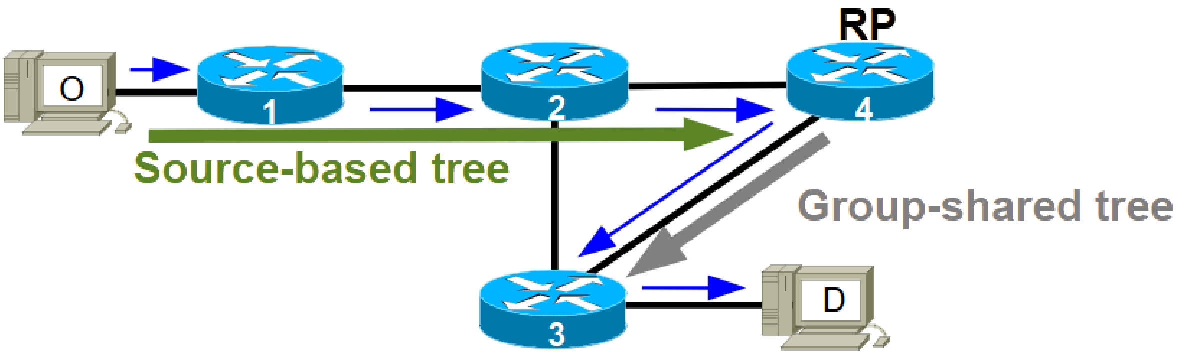


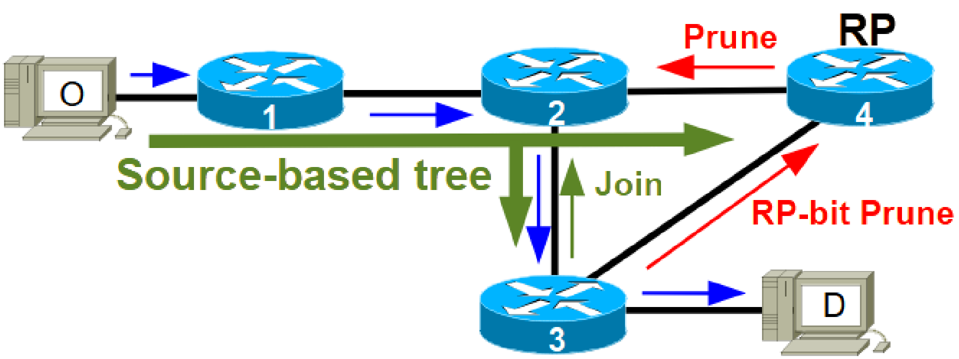
PIM-SM – an active sender host appears (II)

- As soon the source router starts to route the multicast packets through the Source-based tree, and the RP starts receiving them, the RP:
  - forwards the "native" multicast packet through the Group-shared tree,
  - sends a PIM Register-Stop message (in unicast) to notify the source router to stop sending the multicast packets encapsulated in Register messages.



PIM-SM – an active sender host appears (III)



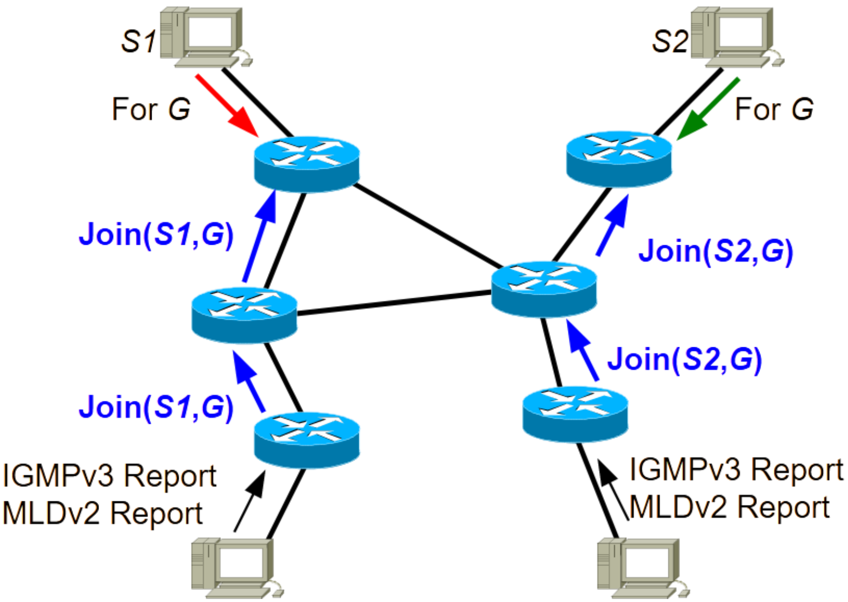


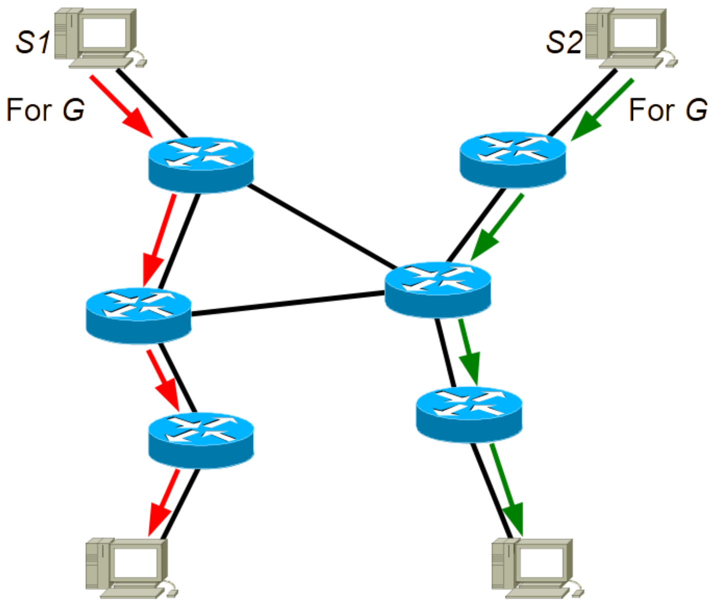
- When the aggregated bit rate of a sender host exceeds a predefined threshold, each router in the Group-shared tree choses to switch to the Source-based tree:
  - The router sends a Join message towards the source address to join the Source-based tree.
  - When a multicast packet arrives through the Source-based tree, the router send a RP-bit Prune message towards the RP to leave the Group-shared tree.
- If all routers join the Source-based tree, the RP leaves the Source-based tree.

PIM Source-Specific Multicast (PIM-SSM)

- More recently, the PIM family of protocols added a new protocol, named PIM Source-Specific Multicast (PIM-SSM).
  - PIM-SSM is uniquely based on Source-Based Trees
    - It can be seen as a sub-set of PIM-SM
  - Besides the address of the multicast session, interested receiver hosts must also specify the interested source addresses
    - PIM-SSM requires IGMPv3
  - The routers connected to interested receiver hosts send a Join message to establish (or join to) a source-based tree towards each specified source address
- There are multicast addresses reserved to SSM routing:
  - IPv4 range: 232.0.0.0 to 232.255.255.255
  - IPv6 range: FF3*x*::/32 (*x* represents the scope of the address)

Illustration of PIM-SSM





- Interested receiver hosts send (IGMVv2 or MLDv2) Report messages to participate in multicast session *G* from different sender host addresses (one from *S1* and the other from *S2*).
- Source routers send PIM Join messages to establish a source-based tree from the sender host requested by the receiver hosts.
- After the source-based trees being established, routers forward the multicast packets to each receiver host only from the requested sender hosts.