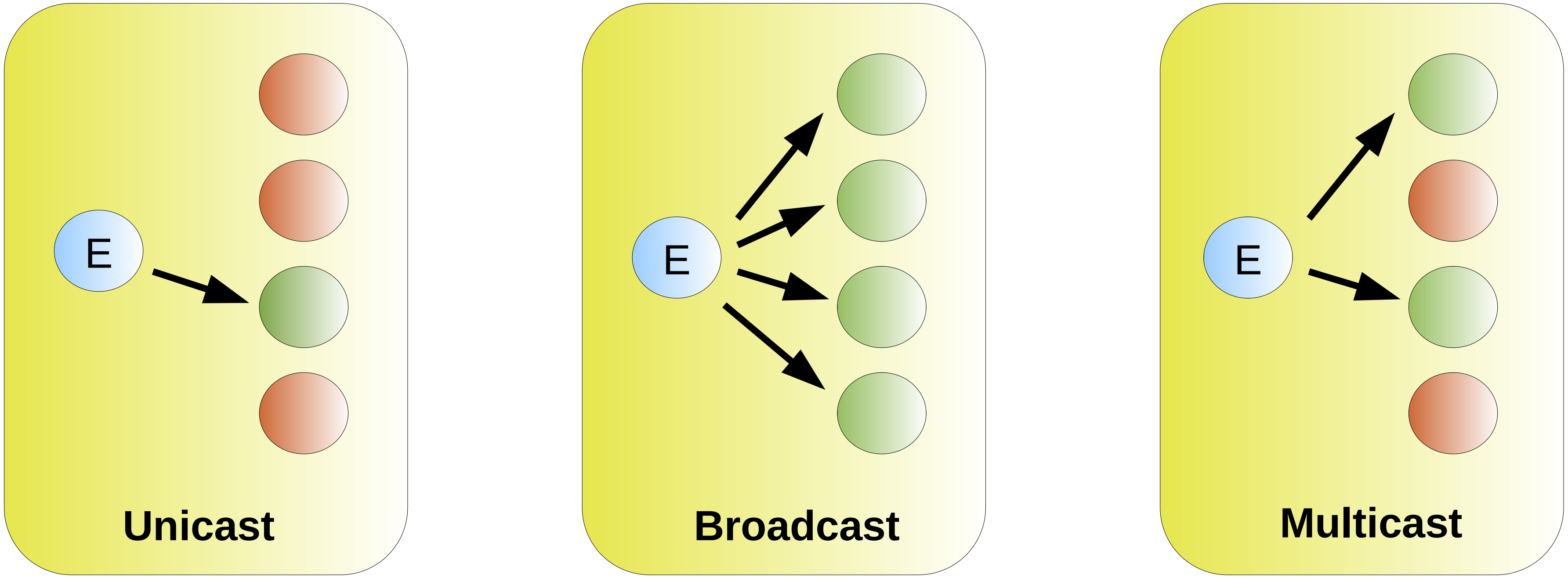
Multicast Routing

universidade de aveiro
deti.ua.pt

Multicast

- Multicast communications refer to one-to-many and many-to-many communications.
  - Multicast at application level.
  - Multicast at network level.



Multicast Abstraction

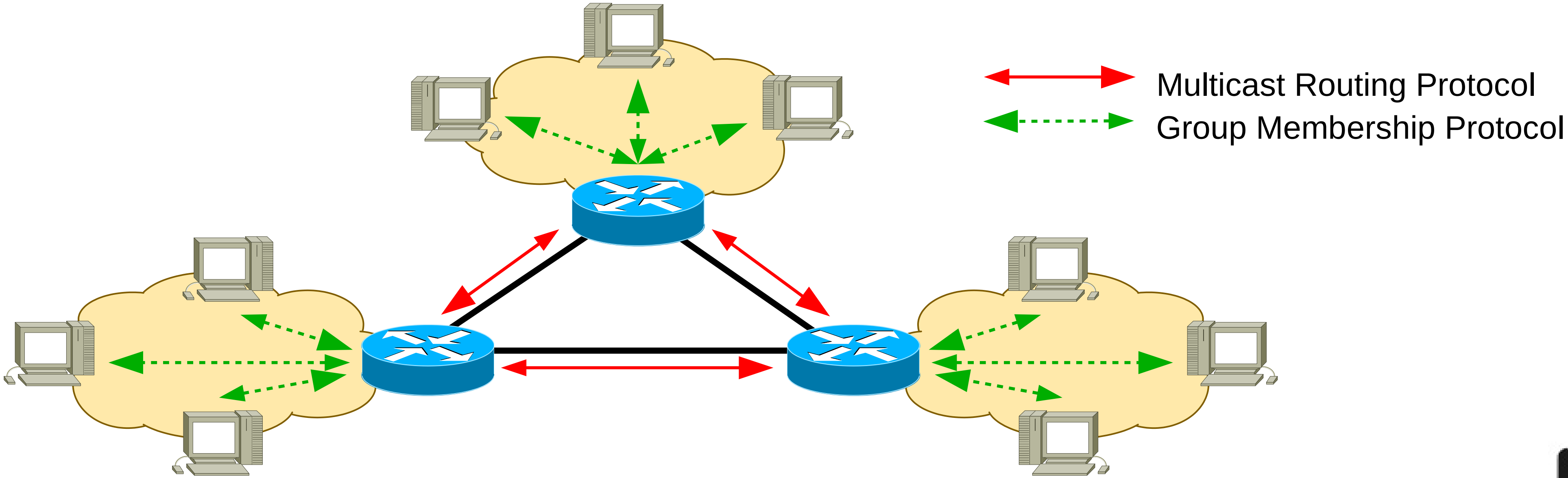
- Information transmited by one origin application is received by multiple destination applications in different stations.
- *Alternative 1*: TCP/IP protocolar stack of the sender station establishes point-to-point connections with all destinations and send multiple copies, one for each destination.
- *Advantages*
    - Allows the use of networks without multicast capabilities.
    - Allows the use of TCP protocol with all its advantages.
- *Disadvantages*
    - Requires that the sender application specifies the destination address list.
    - Results in an inefficient use of network resources.

Multicast Abstraction

- *Alternative 2*: The station sends each IP packet only once and the network is responsible for the copy of the packet to multiple destinations.
- *Disadvantages*
  - Requires that the network has multicast capability.
  - Requires the use of only UDP with all its disadvantages.
- *Advantages*
  - It is possible to have a better usage of network resources.
- *Issues*
  - How do the sender stations specify the destination stations?!
  - How do the routers implement multicast capabilities?!

Multicast Abstraction

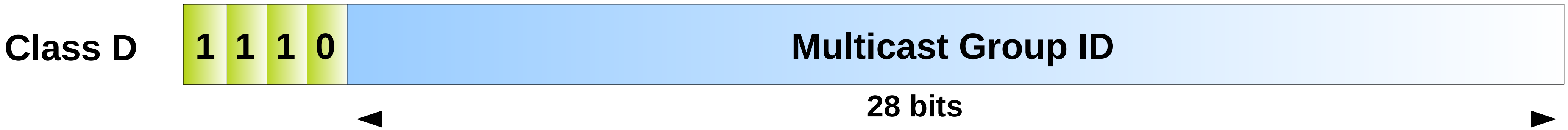
- IP multicast networks are not “*connectionless*” networks as unicast networks.
- It is required to establish multicast paths between the routers for them to know how to route the multicast packets.
- Then, signalling is required (between the stations and the routers) and routing protocols (between the routers) to establish the required multicast paths.



Identification of destination stations

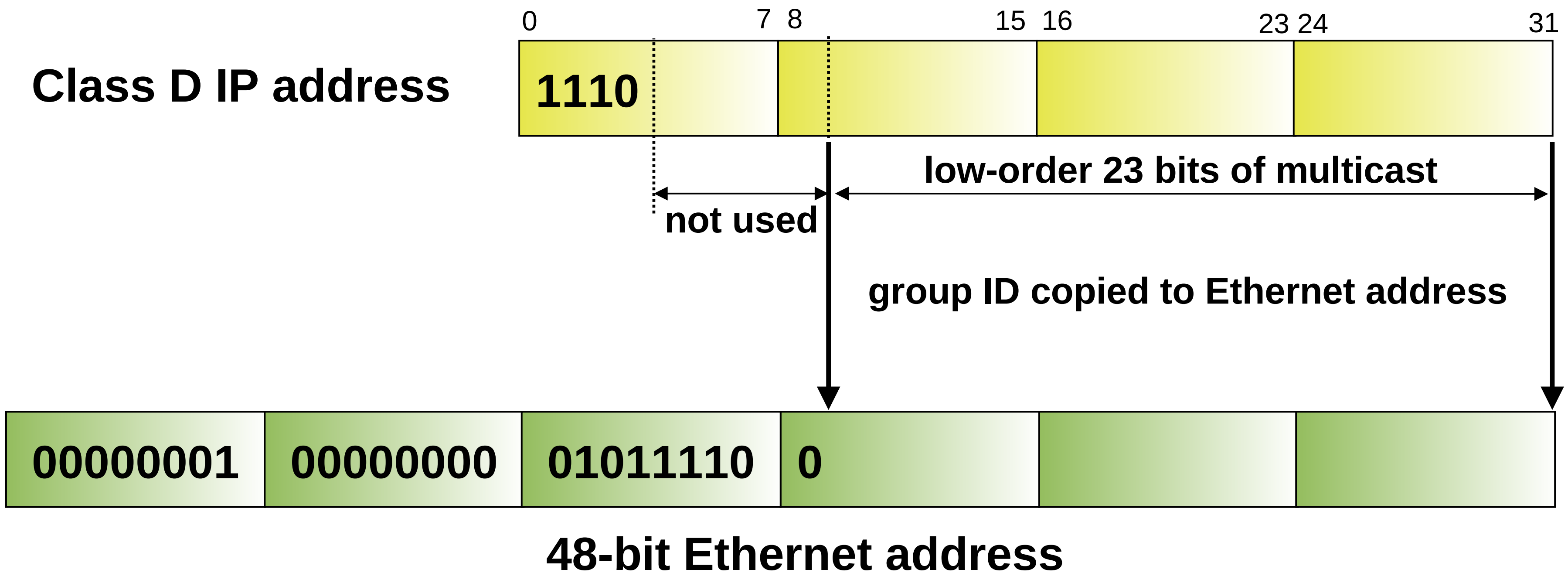
- Explicit identification may not be desirable.
- It makes use of IP addresses of class D (starting by 1110) in IPv4, and addresses of the type FF00::/16 in IPv6.
- The participating stations agree on the use of an address that identifies the session.
- The destination stations announce to the routers their participation in the multicast session identified by the agreed address.
  - IGMP (v4), MLD (v6).
- The routers route the IP packets sent with this agreed address to all networks in which there are participating stations.
  - Routing protocols DVMRP, MOSPF, PIM, etc...

Classe D Addresses

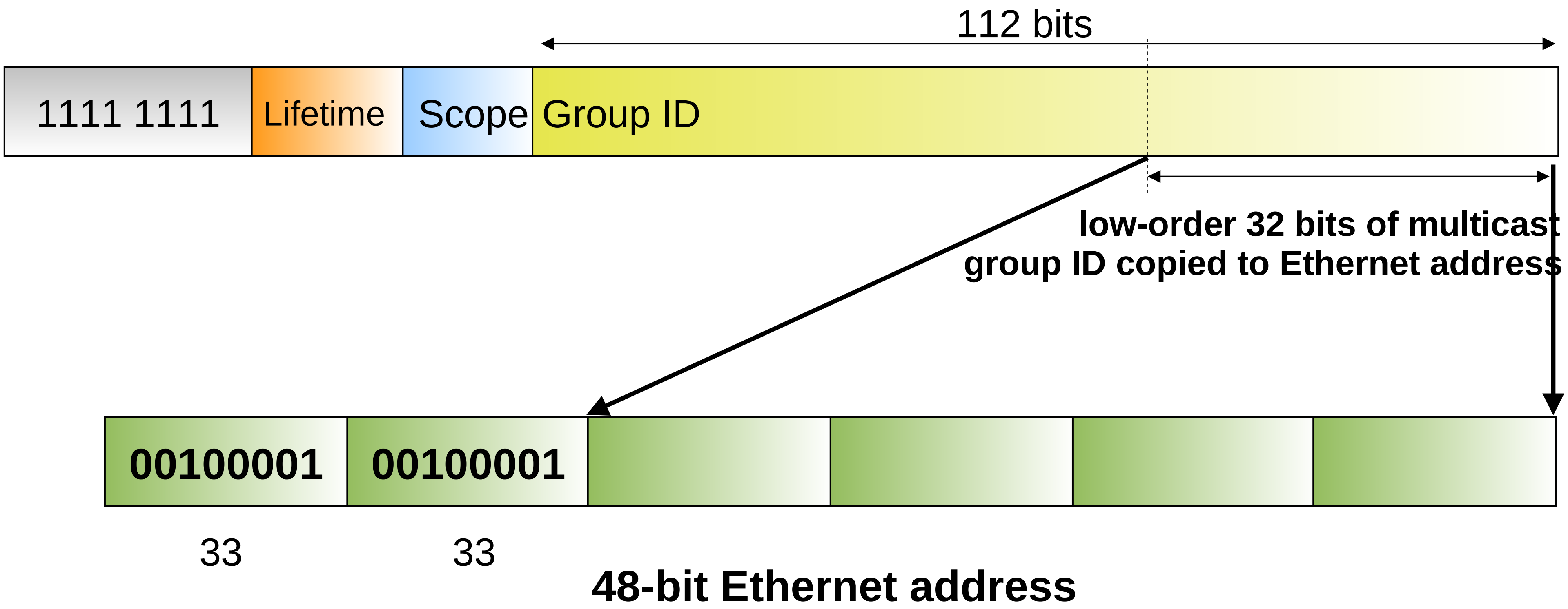


- Some addresses are reserved by the Internet Assigned Numbers Authority (IANA).
- Some addresses have their own meaning. For example, addresses on the form 224.0.0.X:
    - 1 : All Hosts
    - 2 : All Multicast Routers
    - 4 : All DVMRP Routers
    - 5 : All OSPF routers
    - 6 : OSPF designated routers
    - 13 : All PIM routers
- 224.0.0.1 to 224.0.0.225 reserved to routing protocols and other discovery/maintenance protocols..
- The range 239.0.0.0 to 239.255.255.255 is destined to private networks.

Conversion of class D IPv4 address to IEEE 802 address



IPv6 Multicast address conversion to IEEE 802 address



Internet Group Membership Protocol (IGMP)

- IGMP version 2, RFC 2236
  - Operates between the station and the directly connected routers.
  - Is used for the station to announce to the router that it wants to participate in the multicast session (identified by a class D address).

0–8	8–16	16–32
Type	MRT	Checksum
Endereço classe D		

MRT - Maximum Response Time

- IGMP runs through IP protocol (protocol type = 0x02).
- The packets are sent to the destination address 224.0.0.1 (“All Hosts”) with TTL= 1.

IGMP Messages

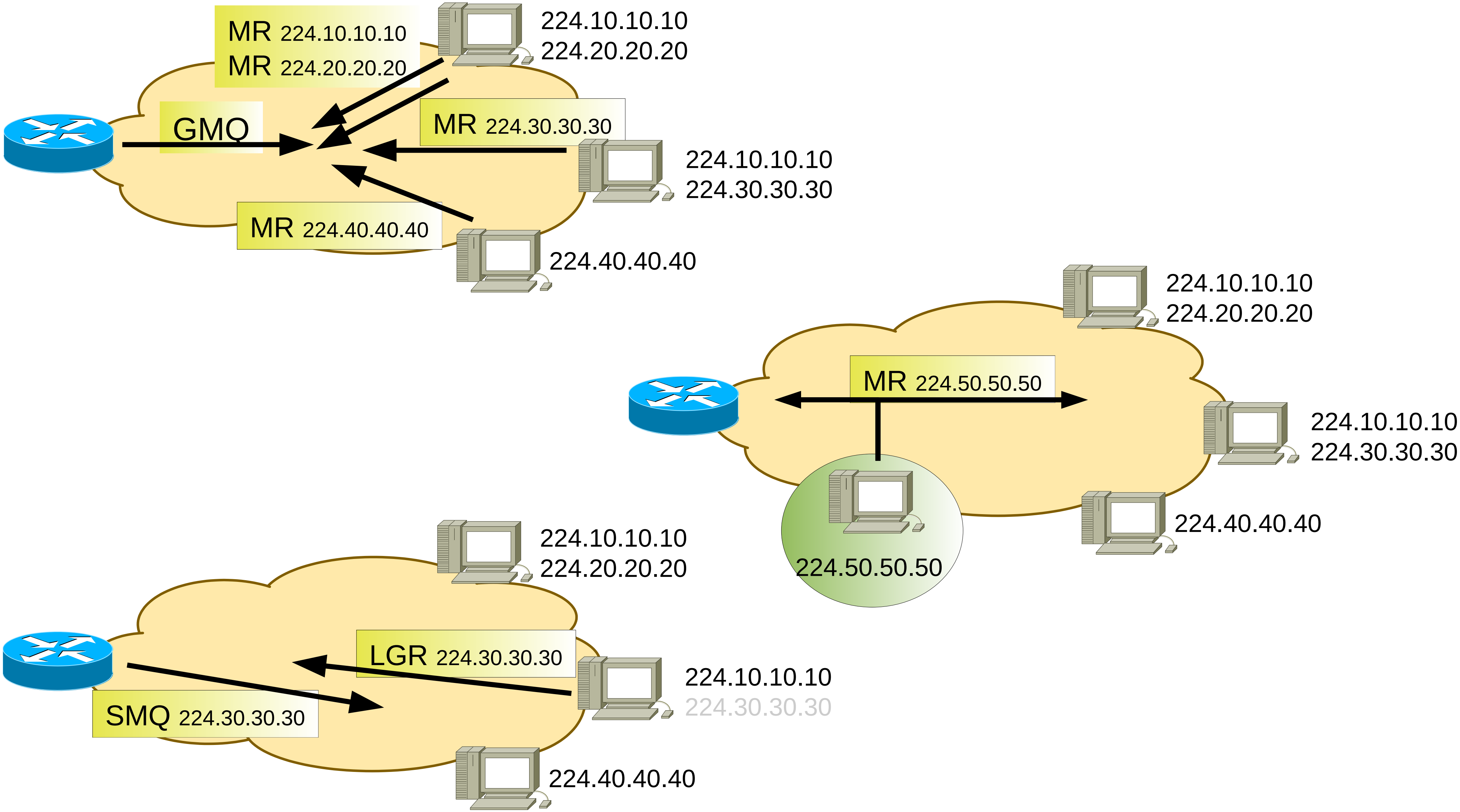
- GMQ - General Membership Query
  - Sent by the routers to ask the stations if they participate in a multicast session.
- SMQ - Specific Membership Query
  - Sent by the routers to ask if there is any station that participates in a specific multicast session.
- MR - Membership Report
  - Sent by the stations to signal that they participate in a multicast session .
- LGR - Leave Group Report (Optional)
  - Sent by the stations to signal that they will leave a multicast session
  - Update can be done through the Membership Report.
- In each network, the Querier Router is the router with lower IP address among all interfaces connected to the network, and is the one that maintains the IGMP messages with the terminals.

IGMP Protocol

- Routers send periodically a GMQ specifying a *Maximum Response Time* (MRT).
- Each station waits a random time between 0 and MRT to answer to a MR specifying a *multicast* address.
- If in the meanwhile the station ‘sees’ a MR for the same session, it aborts the sending of the MR.

- Each station sends a MR when it wants to belong to a multicast session .
- Optionally, a station sends a LGR when it does not belong anymore to a multicast session.
- When a router receives a LGR, it sends a SMQ to verify if there are still any stations belonging to that session.

IGMP Protocol



IGMPv1/v2 – Final Conclusions

- Any station can join a multicast session receiving and sending information.
- The formation of multicast sessions is initiated by the receivers.
    - Senders do not specify nor control the stations that can receive information.
- The network does not provide filtering, ordering or privacy to multicast packets.
- The multicast IP service model follows the same philosophy of the unicast:
    - Simple and reliable protocol layer in which additional functionalities are provided by the upper layers.

IGMPv3

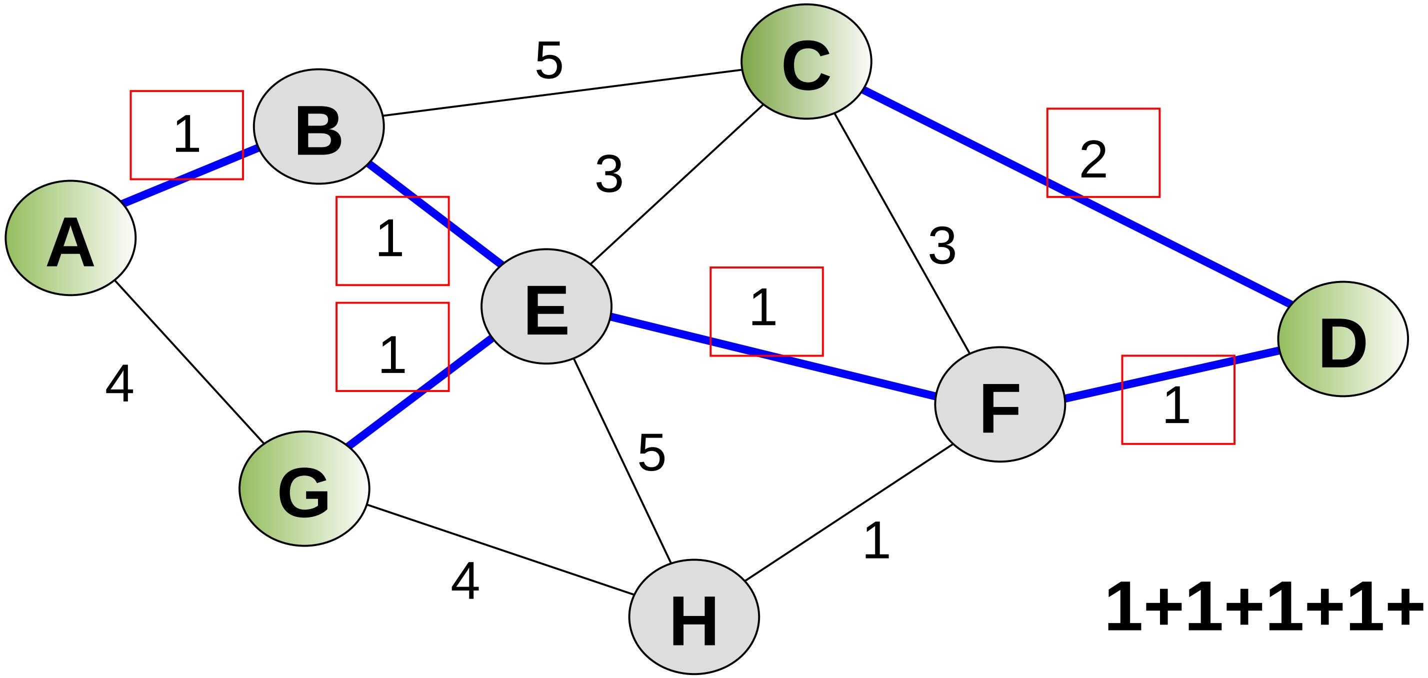
- IGMPv3 adds support to "source filtering".
  - Allows a terminal to report interest in a specific multicast session/group, from
    - ONLY a specific source.
      - INCLUDE Mode.
    - ALL sources EXCEPT specific sources.
      - Exclude Mode.
      - A blank list means interest in all sources.
  - Allows simultaneous requests to multiple multicast sessions.
- Has a new “Report” message format.
    - Version 3 Membership Report.
- Allows IGMPv1 and IGMPv2 interoperability.
  - Supports Version 1 Membership Report, Version 2 Membership Report and Version 2 Leave Group.

Multicast Routing

- Group-shared tree
  - It is based on determining a routing tree per multicast session that connects all routers with stations belonging to the session.
  - Minimal spanning (Steiner tree).
    - It is not used in practice:
      - High computational complexity.
      - Requires information about the overall network.
      - Monolythic: executed every time a router needs to join/leave a session.
  - Minimum cost tree to a central node (“rendezvous point”).
- Source-based tree
  - It is based on determining a routing tree, per multicast session, and per sender.
  - One router is identified as the “central point”.

Group-shared tree

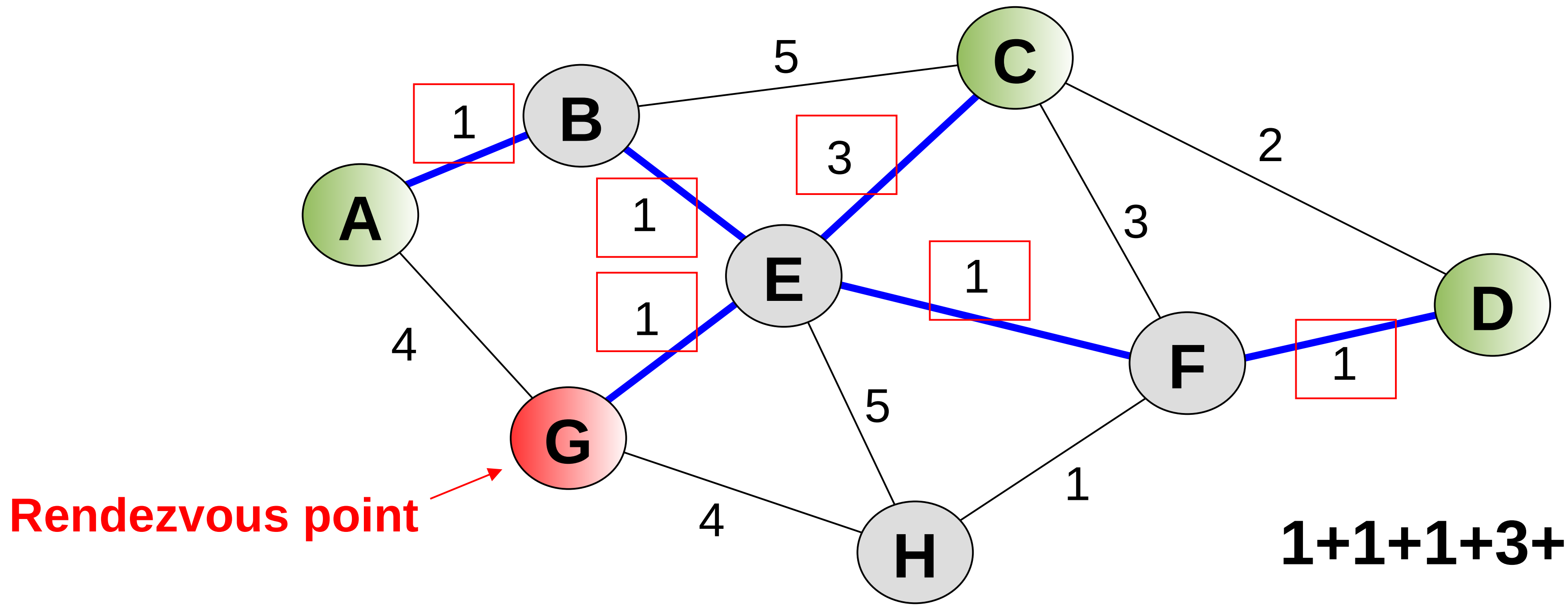
- Steiner Tree: determines the minimum cost tree that interconnects all nodes with stations of a session.
  - Requires a link-state protocol.
  - Algorithm with exponential complexity.



1+1+1+1+1+2=7

Group-shared tree

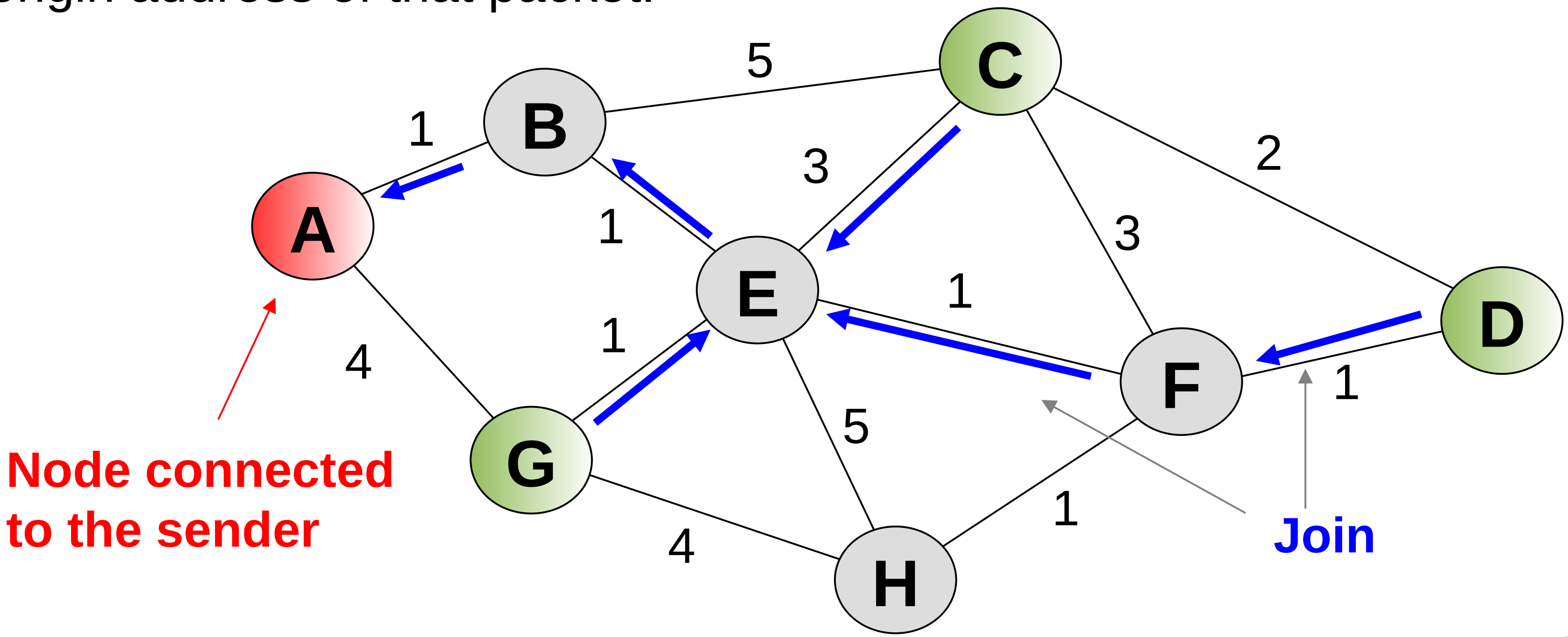
- Minimum cost tree of the paths to a central node (“rendezvous point”).
  - Central node is previously chosen (belonging to the tree even if it does not contain stations that belong to the session) and known by other nodes.
  - To join the tree, the nodes with stations belonging to the session send a “join” message by the path with minimum cost between them and the central node (rendezvous point).



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

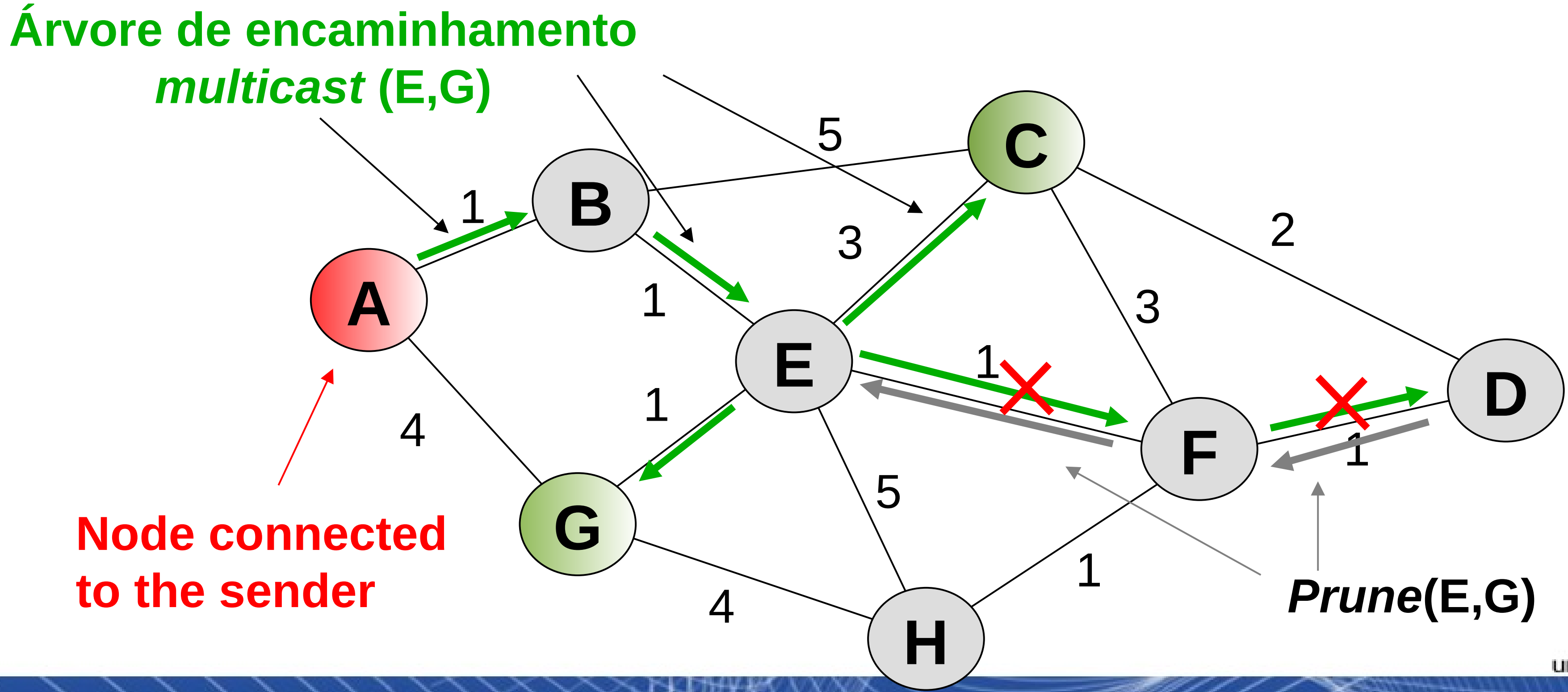
Source-based tree (when the sender is known)

- Each node with receivers interested in a multicast session G of a sender with address E sends a join(E,G) message towards the address of E by the minimum cost path (unicast path).
- In the path of the join message, each node receives the message in interface F0, re-sends it through interface F1 and builds a routing table (E,G) entry stating that multicast packets that enter in F1 can be routed through F0.
- Multicast routing is based not only on the destination address, but also on the origin address of that packet.



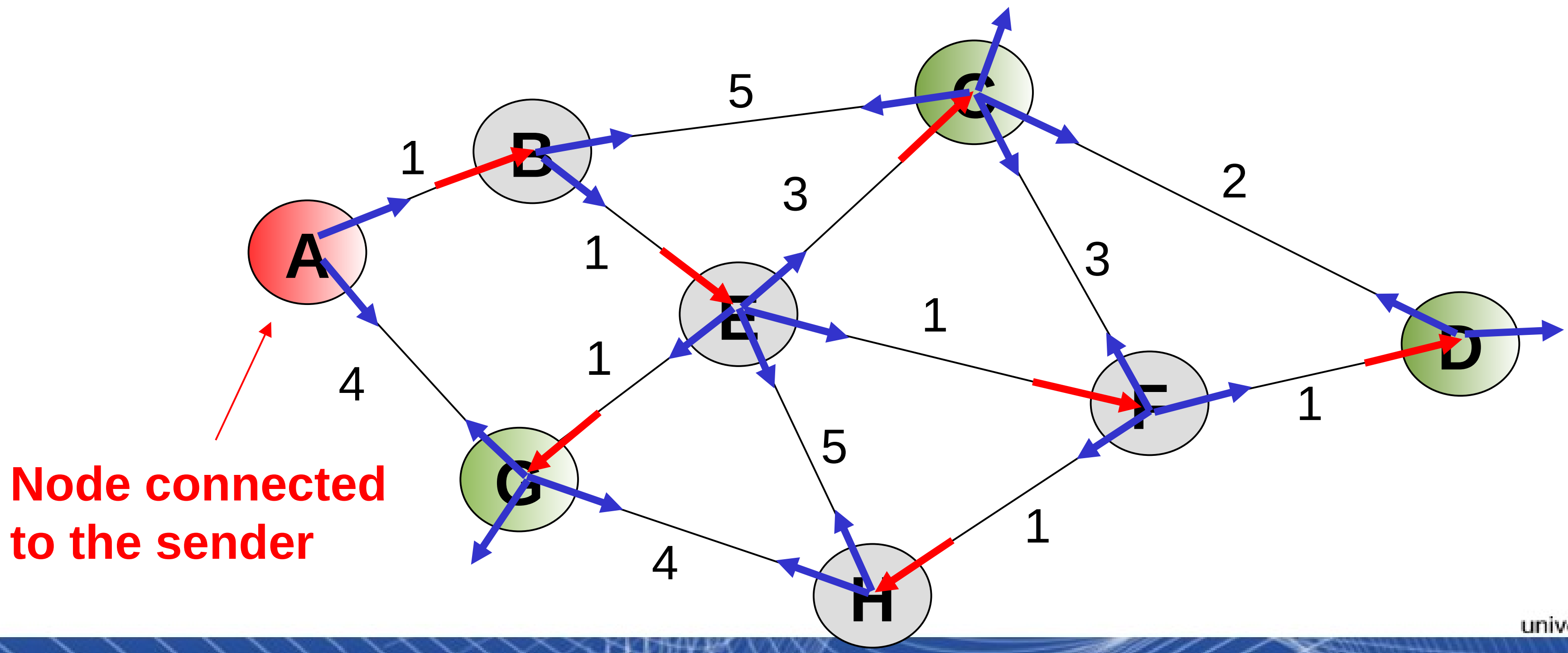
Source-based tree (when the sender is known)

- When there are multiple senders, several multicast routing trees are established, one by each sender.
- When a receiver is not interested in a multicast session, its connected node sends a prune(E,G) message towards the address of sender.
- The prune message is re-sent by the nodes that do not belong anymore to the multicast routing tree.



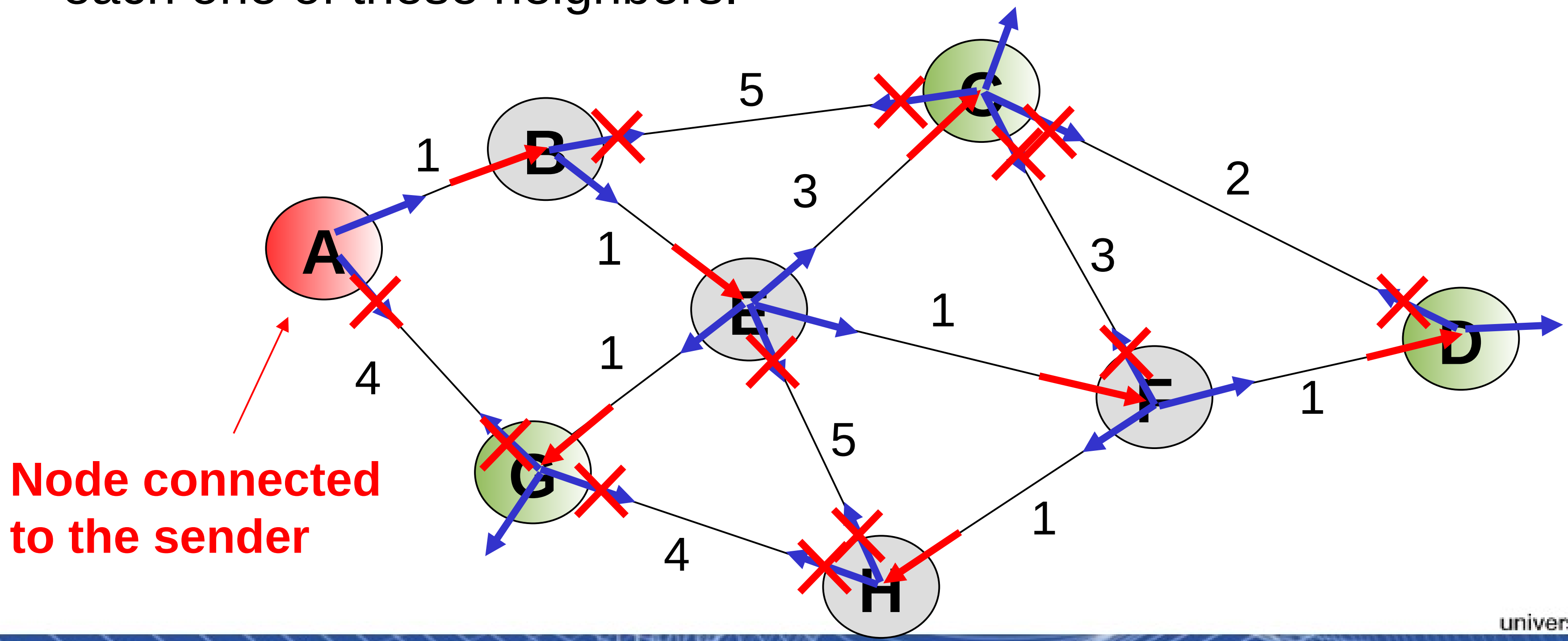
Source-based tree (when the sender is unknown)

- "Reverse Path Forwarding": creation of a virtual spanning tree
  - Each node N routes the packets from the source node O, which are received from neighbor V, to all its neighbors only if V is the last node in the minimum cost path from O to N.
  - Figure: in node E, packets originated from node A are routed to all other ports only if they come from node B because this is the previous node of E in the minimum cost path from A to E.



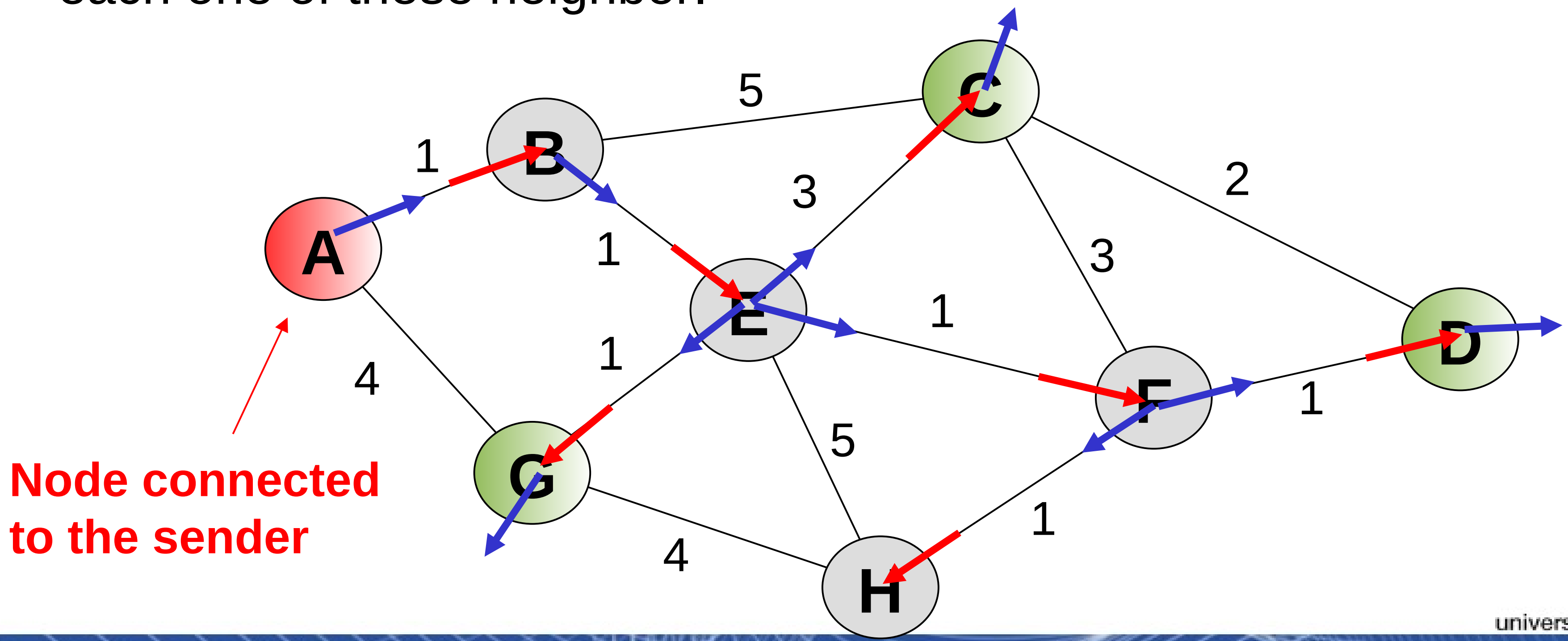
Source-based tree (when the sender is unknown)

- "Reverse Path Forwarding with Pruning" (I)
  - For each origin node O, node N knows to which neighbors V it is the last node in the minimum cost path from O to V.
  - Routing is made to these neighbors.
  - Figure: node E only routes packets from origin node A to neighbors C, F and G because node E is the last node in the minimum cost path from A to each one of these neighbors.



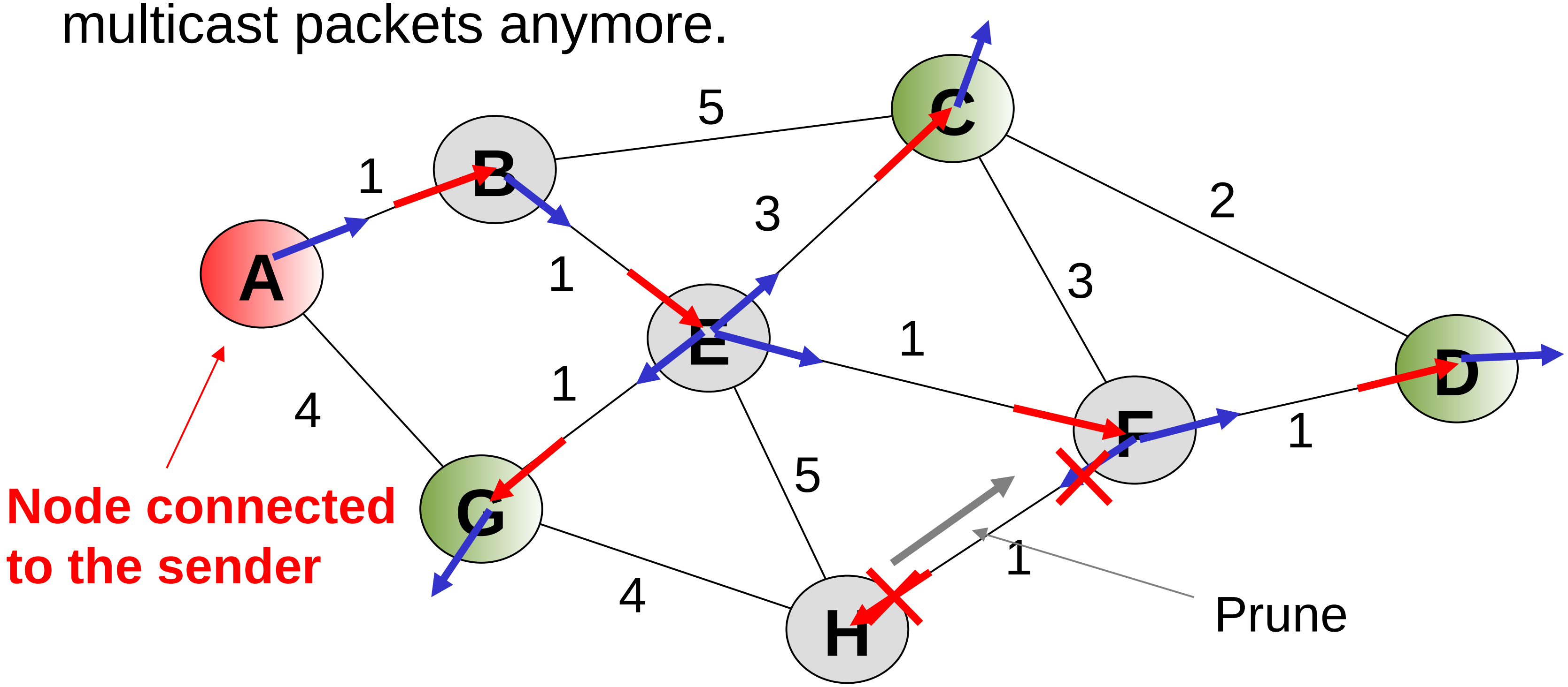
Source-based tree (when the sender is unknown)

- "Reverse Path Forwarding with Pruning" (I)
  - For each origin node O, node N knows to which neighbors V it is the last node in the minimum cost path from O to V.
  - Routing is made to these neighbors.
  - Figure: node E only routes packets from origin node A to neighbors C, F and G because node E is the last node in the minimum cost path from A to each one of these neighbor.



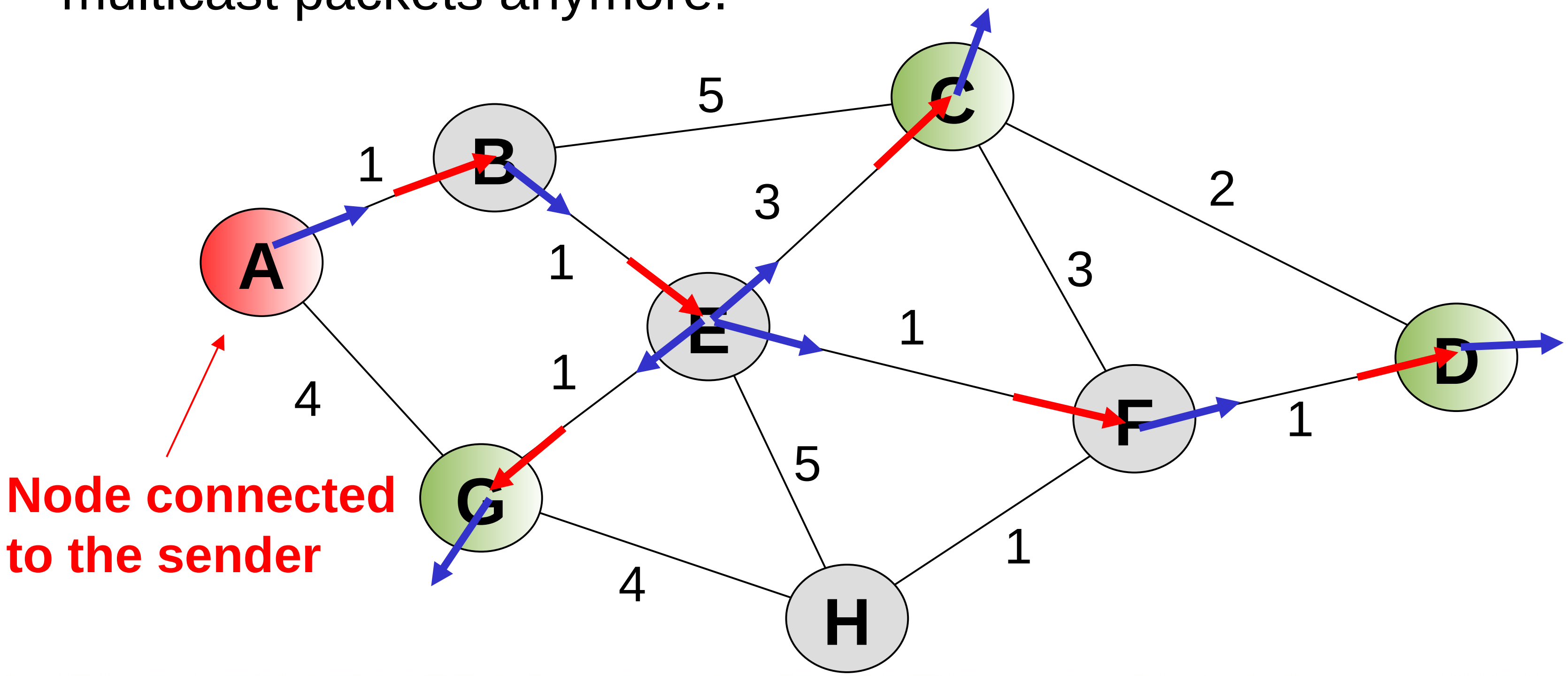
Source-based tree (when the sender is unknown)

- "Reverse Path Forwarding with Pruning" (II)
  - A node without any terminal stations interested in the multicast session, and without any neighbor nodes to forward the multicast packets to, sends a prune message to the neighbor from which it receives the multicast packets.
  - Figure: node H sends a prune message to neighbor F and will not receive multicast packets anymore.



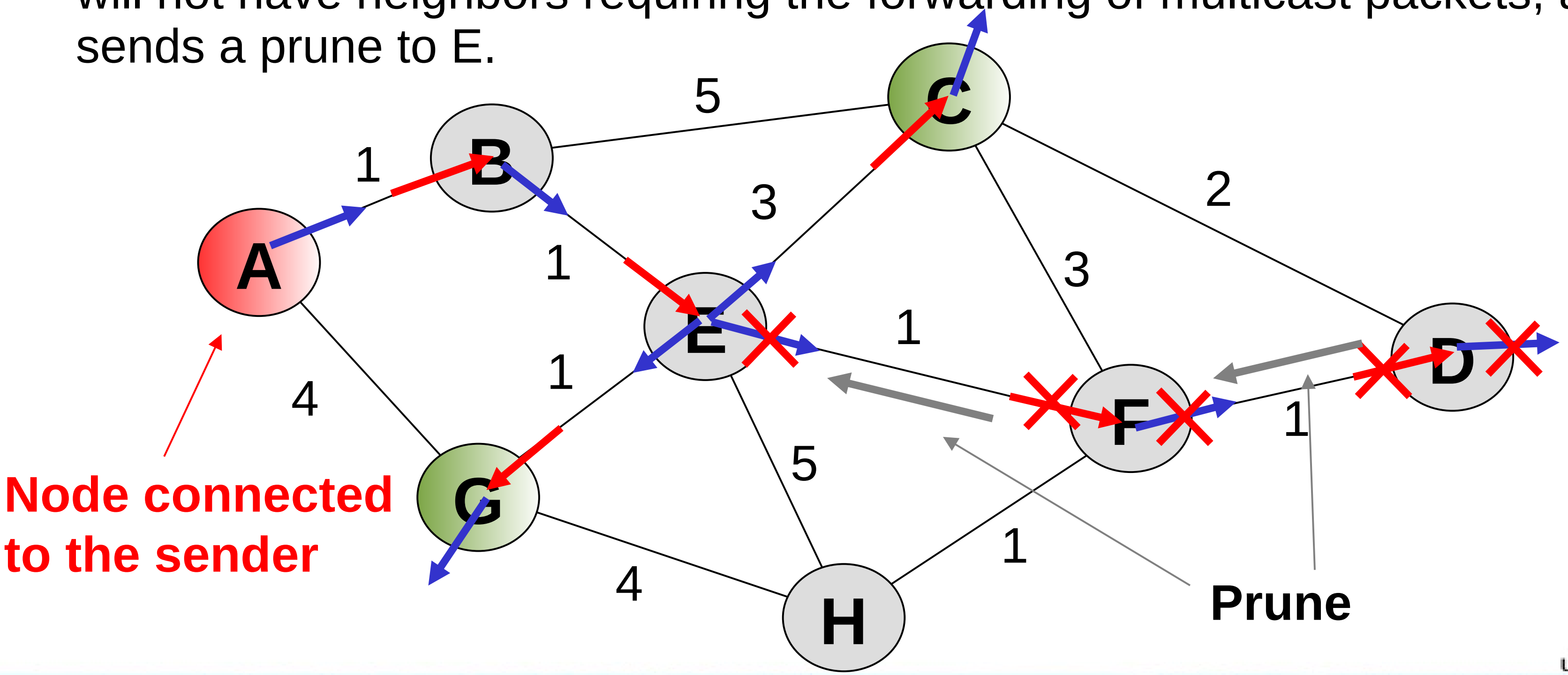
Source-based tree (when the sender is unknown)

- "Reverse Path Forwarding with Pruning" (II)
  - A node without any terminal stations interested in the multicast session, and without any neighbor nodes to forward the multicast packets to, sends a prune message to the neighbor from which it receives the multicast packets.
  - Figure: node H sends a prune message to neighbor F and will not receive multicast packets anymore.



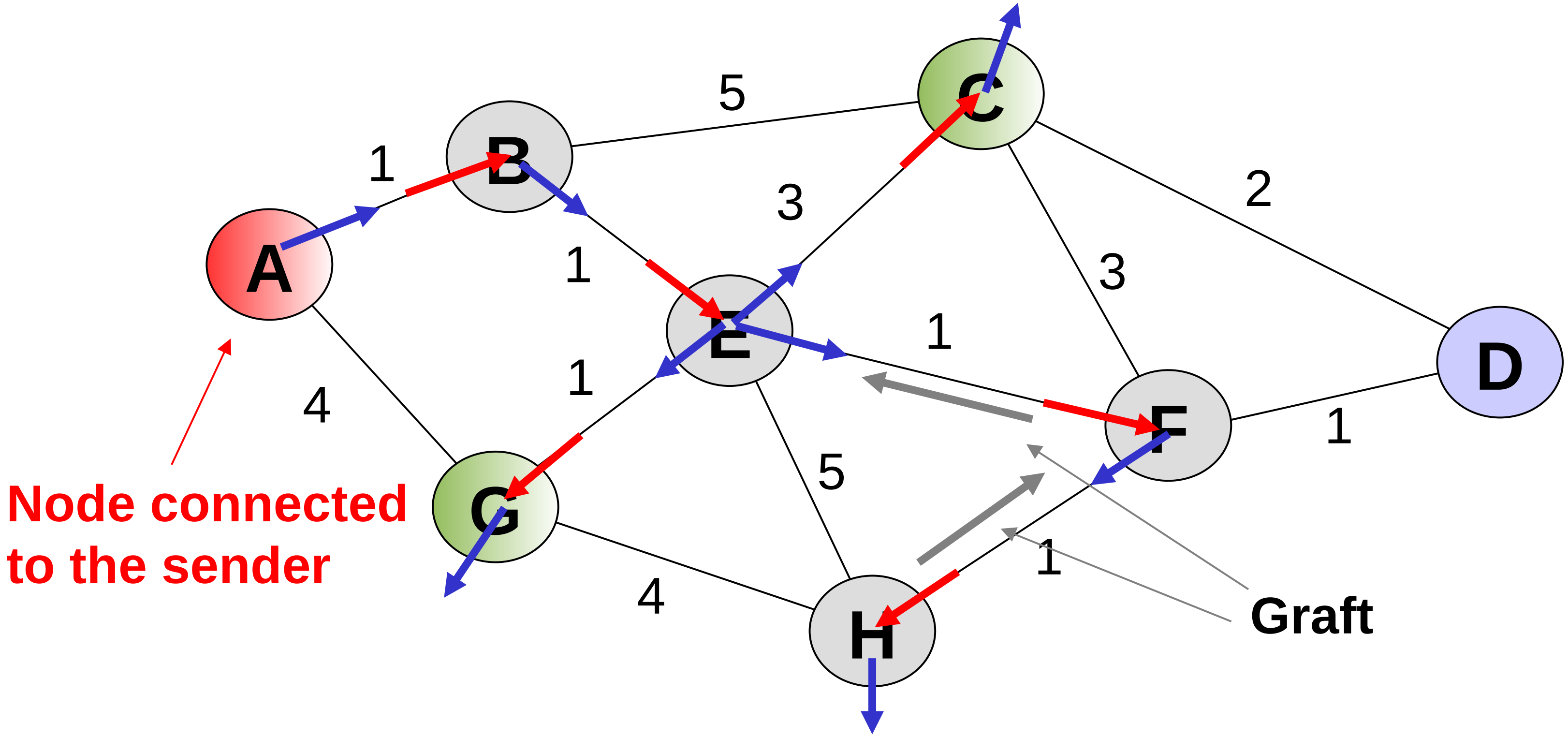
Source-based tree (when the sender is unknown)

- "Reverse Path Forwarding with Pruning" (II)
  - A node without any terminal stations interested in the session, and without any neighbor nodes to forward the multicast packets to, sends a prune message to the neighbor from which it receives the multicast packets.
  - Figure: node H sends a prune message to neighbor F and will not receive multicast packets anymore.
  - Figure: node D has no receivers anymore and sends a prune to F that, in turn, will not have neighbors requiring the forwarding of multicast packets, and also sends a prune to E.



Source-based tree (when the sender is unknown)

- "Reverse Path Forwarding with Pruning" (III): multicast session member nodes appear
  - First strategy: use a graft message to undo the prune.
  - Second strategy: associate a lifetime to the prune; after the lifetime, multicast packets will be forwarded again.



Group-shared tree vs. Source-based tree

- Group-shared tree
  - Minimizes the state information that needs to be maintained in each router.
  - Minimizes the number of connections used to support multicast traffic.
  - Concentrates the traffic congestion problems in a reduced number of nodes.
- Source-based tree
  - Penalizes the state information in each router, since it needs to know information about each origin.
  - Distributes the multicast traffic by a larger number of connections.
  - Less congestion problems.

Distance Vector Multicast Routing Protocol (DVMRP)

- Algorithm of the “source-based tree” type
  - “Distance Vector” (RFC 1075), similar to RIP.
- Uses a strategy of RPF (reverse path forwarding) with “pruning”
  - Such as RIP, the distance is given by the number of hops.
  - Distance vectors represent the distance of each possible origin.
  - For each possible origin, each router announces to its neighbors when they are the last hop in the path from the origin.
- “prune” messages
  - Sent with na associated lifetime.
- “graft” messages
  - To eliminate/recover from a “prune” message.

Distance Vector Multicast Routing Protocol (DVMRP)

- As its name suggests, DVMRP uses a distance-vector routing algorithm.
- Such algorithms require that each router periodically informs its neighbors about its routing table.
- DVMRP routers advertise routes by sending DVMRP report messages.
- For each network path, the receiving router picks the neighbor advertising the lowest cost and adds that entry to its routing table for future advertisement.
- All interfaces are configured with a cost metric and a threshold TTL that limits the scope of the multicast transmission
    - It is possible to change the metric associated to an interface to promote or demote the preference for some routes.
    - A multicast router forwards a multicast datagram through an interface if the TTL in its header is larger than the interface threshold TTL.
- Allows the use of tunnels between multicast routers
    - Tunnels are administratively configured.
    - Border routers act as neighbors.

Multicast Open Shortest Path First (MOSPF)

- Algorithm of the "source-based tree" type
  - It is an extension of OSPF (RFC 1584).
  - Implements a "RPF with pruning" strategy.
- It makes use of the topology knowledge of each router
  - In this way, each router can locally process the minimum cost tree for each multicast session.
- It does not support tunnels
  - In the OSPF messages there is a flag that, when set toNull, indicates that the router does not support multicast.
  - These routers will not belong to the minimum cost tree.
  - This protocol requires that any 2 multicast routers should have at least one path where all intermediate routers support multicast.
- Each MOSPF router has a local database of multicast groups containing a list of the directly connected group members
  - The local router forwards the multicast datagrams to the group members based on this information.
  - The local router will send a group-membership LSA to all other routers in the domain.

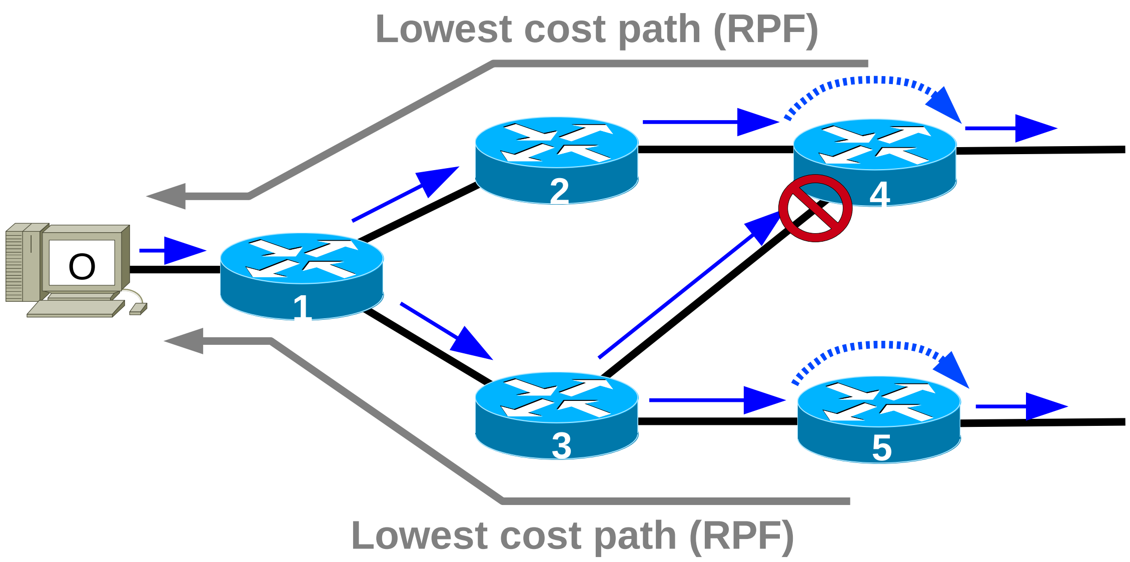
Protocol-Independent Multicast (PIM)

- PIM (RFC 2362) addresses two extreme use cases
- PIM, dense mode
  - When most of the networks contain stations that want to use multicast.
  - Consequently, the majority of the network routers need to route multicast packets.
- PIM, sparse mode
  - Stations that want to use multicast are concentrated on a reduced number of networks.
  - Consequently, the number of routers that need to route multicast packets is small when compared to the total number of routers.

PIM Dense Mode

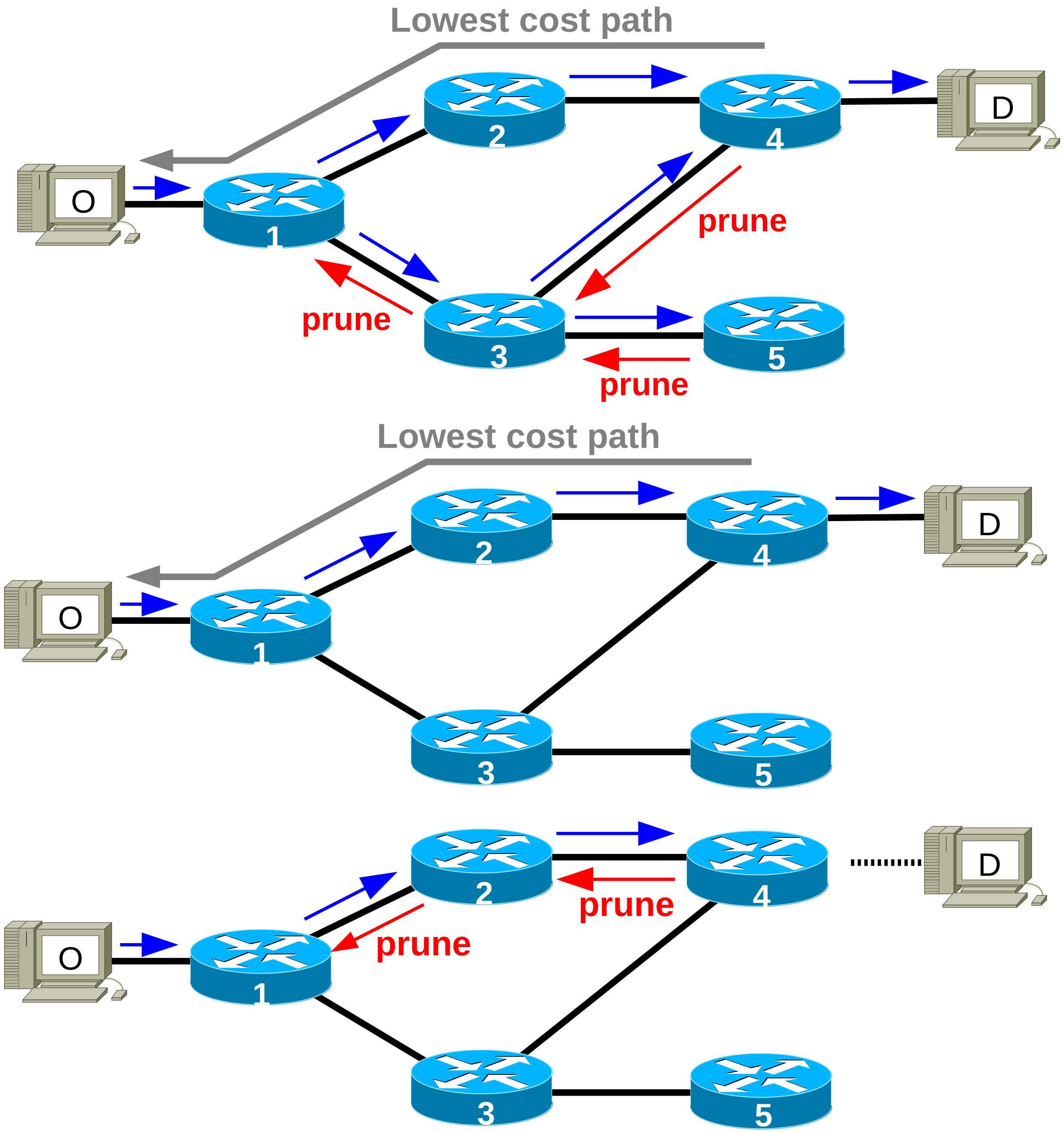
- Implements a “RPF with pruning” strategy.
- Requires all routers to have the protocol active.
- It is simpler because:
  - Does not calculate routing tables.
  - Instead, it uses routing tables built by any other unicast protocol.
    - It is then independent from the routing protocol that is in use.
  - Assumes that all point-to-point routes are symmetric.
- In case of multiple minimum cost paths, it only accepts the ones corresponding to the highest IP address interface.
- Unicast routing tables do not allow to determine to which neighbors the packets should be forwarded to. So:
  - By default, routers forward packets to all neighbors that did not send prune messages.
  - Use prune messages to signal their neighbors that they should not forward packets to them.

PIM Dense Mode – Initial Flooding



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

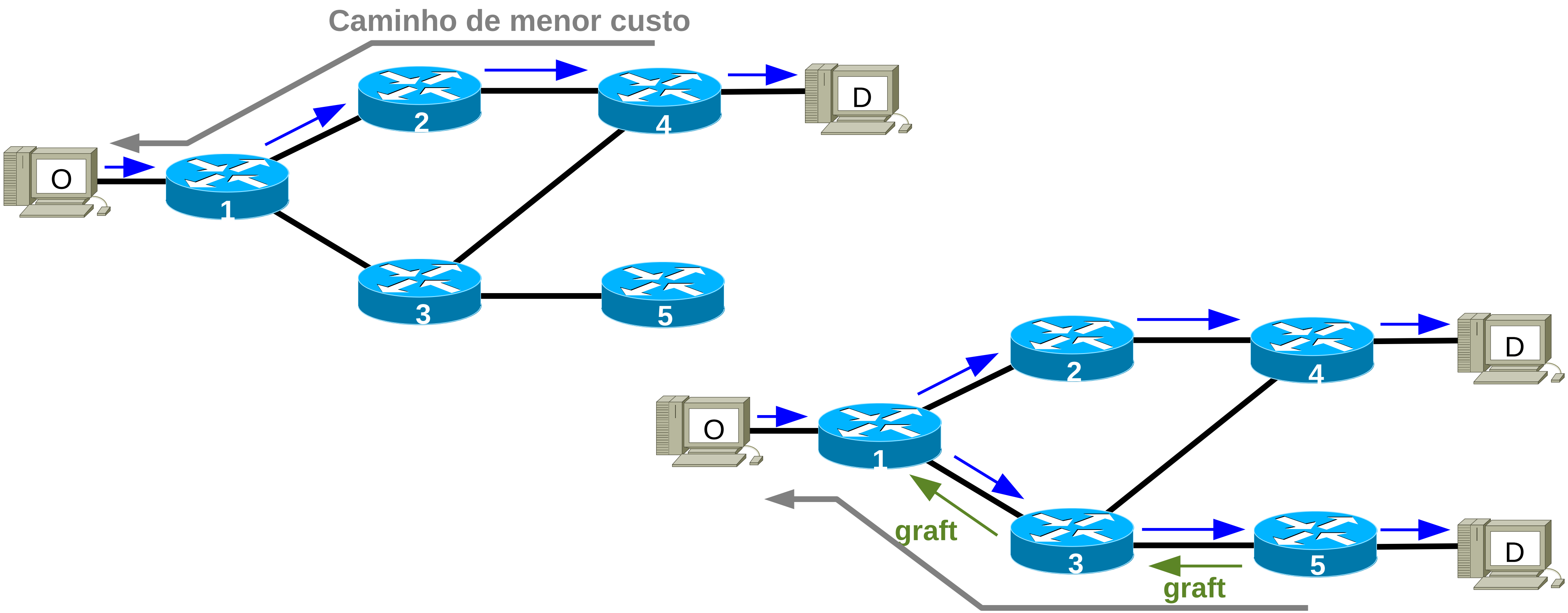
PIM DM – *Prune* Message



- Routers forward multicast trafic received in one interface to to all other interfaces in which they did not received a *Prune* message.
- A *Prune* Message is sent by:
  - Routers with no clients interested in a specific multicast session (e.g., Router 5).
  - Routers that receivbed the same multicast traffic in more than one interface (e.g., Router 4)
    - Send the *Prune* message via all interfaces in which the traffic was received, and do not provide the lowest path cost to source.
  - Routers that received *Prune* messages in all interfaces to which the multicast traffic was forwarded (e.g., Router 3).

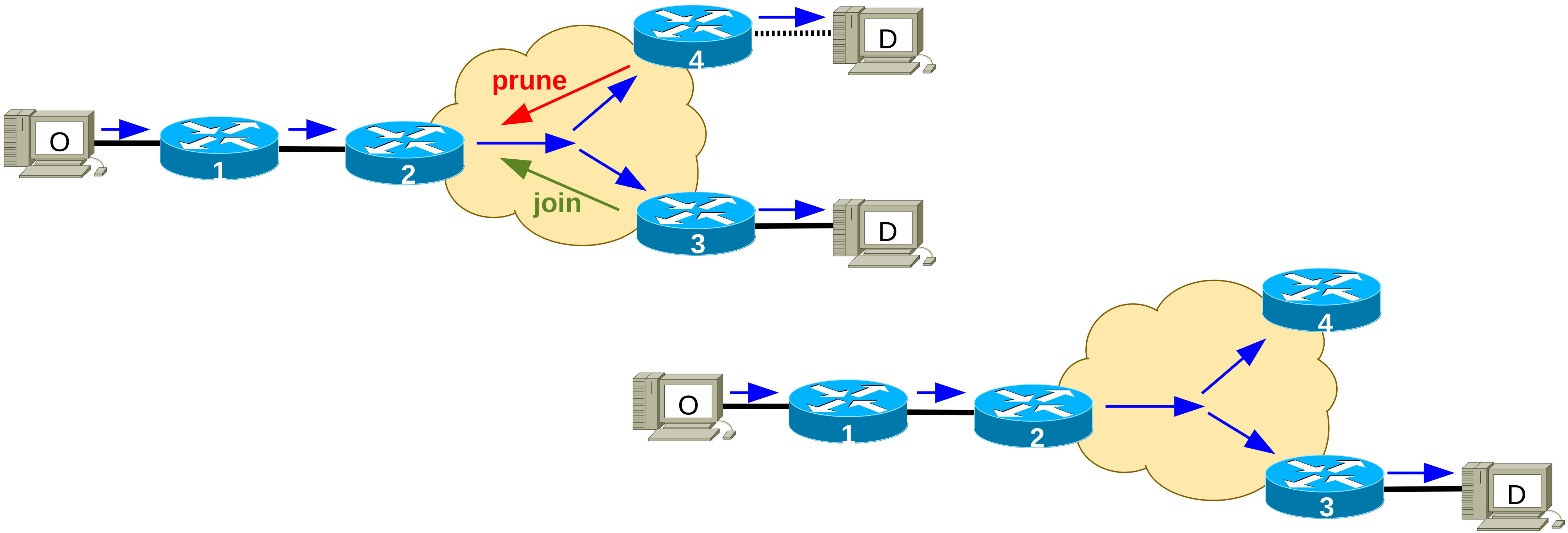
- When a router does not have a more clients interested in a specific multcast session it send a *Prune* message via the interface that receives the multicast traffic.

PIM DM – *Graft* Message



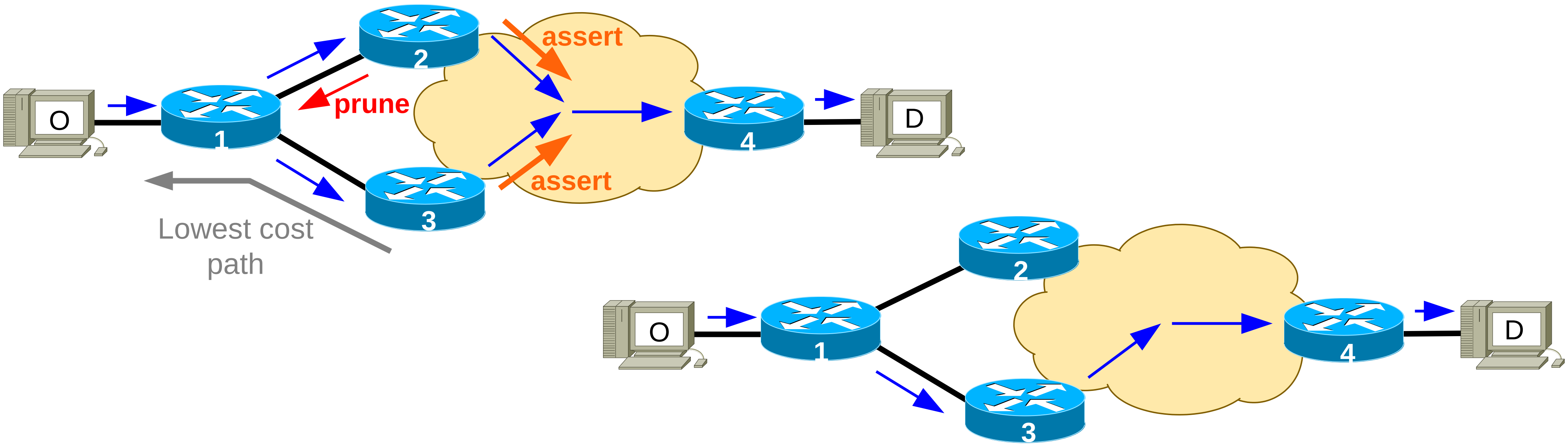
- A router may restart to receive multicast traffic by sending a *Graft* message to cancel a previously sent *Prune* message.

PIM DM – *Join* Message



- When a router has no more clients for a specific multicast session it send a *Prune* message.
- When the *Prune* message is sent through a shared medium (e.g., LAN), and when other routers have clients for the same multicast sessions, these should send a *Join* message to nullify the sent *Prune* message.

PIM DM – *Assert* Message



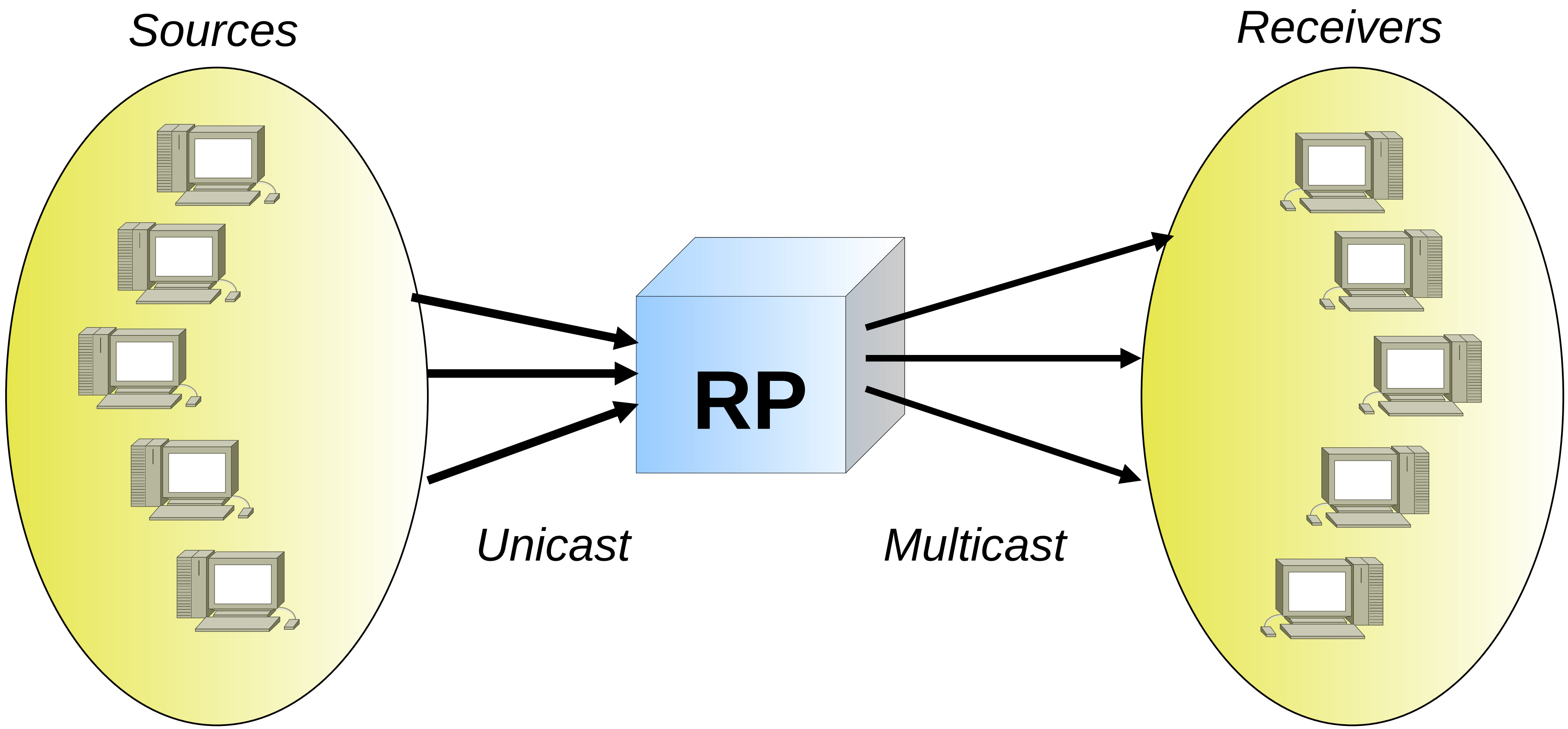
- When there are more than one router sending traffic from a specific multicast session to a shared medium (e.g., LAN), these must decide which one will be responsible for the multicast traffic.
- All routers send an *Assert* message.
  - Message contains the path cost to the multicast source.
- The chosen router is
  - The one that provides the lowest cost path to the source,
  - In case of tie it is chosen the router with the largest IP address (in interfaces connected to the shared medium).

PIM Sparse Mode

- PIM Dense Mode is a data-driven protocol.
  - Requires that routers that do not have multicast clients must periodically send prune messages to avoid receiving multicast packets.
- PIM Sparse Mode is a receiver-driven protocol.
  - Each router announces explicitly (with join messages) that it wants to join specific multicast sessions.
- PIM Sparse Mode initially uses a “Group-Shared tree” strategy based on a Rendezvous Point (RP) router.
  - RP can be administratively configured.
  - 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 sessions.

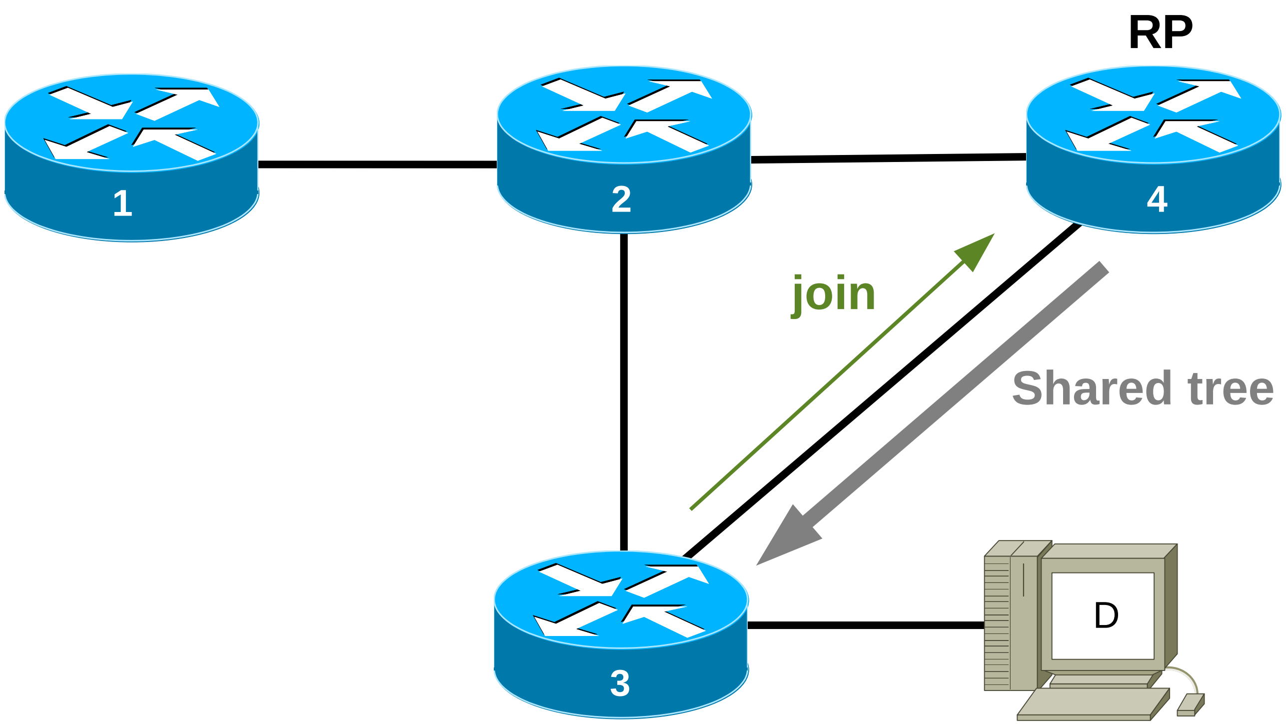
PIM Sparse Mode

- At start, multicast packets from a specific session are sent to the RP using unicast (Multicast over PIM tunnel).
- At the RP, the multicast packets are resent to all interested clients.

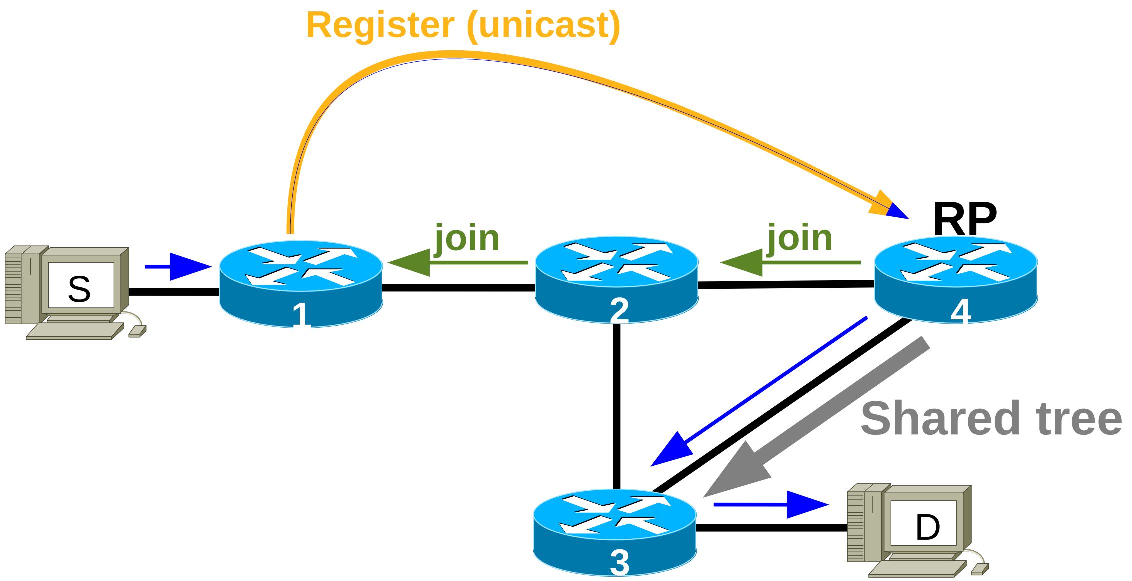


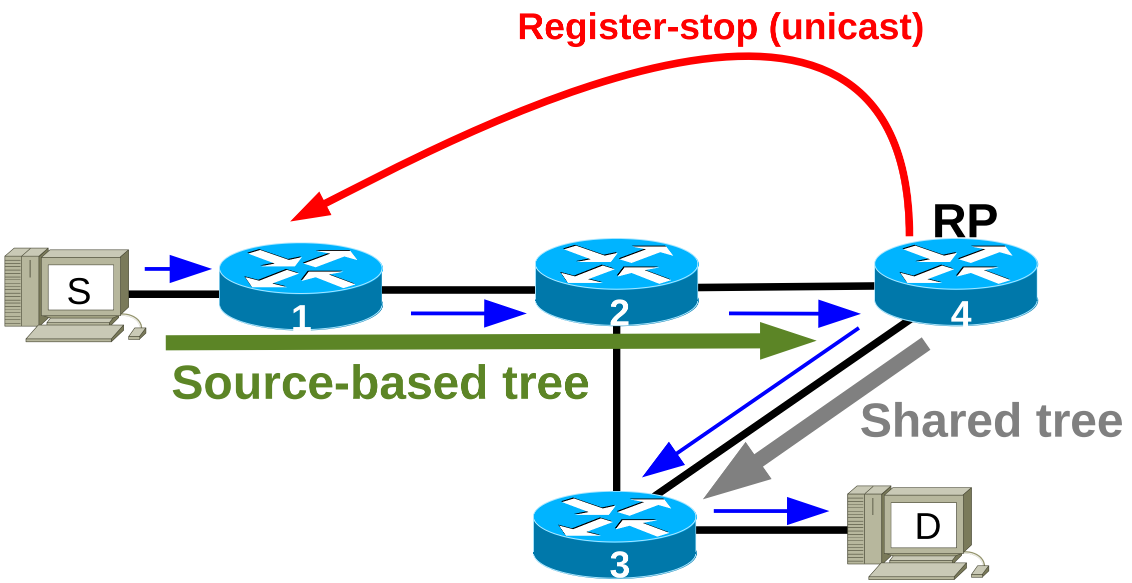
PIM SM – Join to *Group-Shared Tree*

- The routers with clients interested in a specific multicast session, send a *Join* message to the RP in order to join the group-shared tree (which root is the RP).



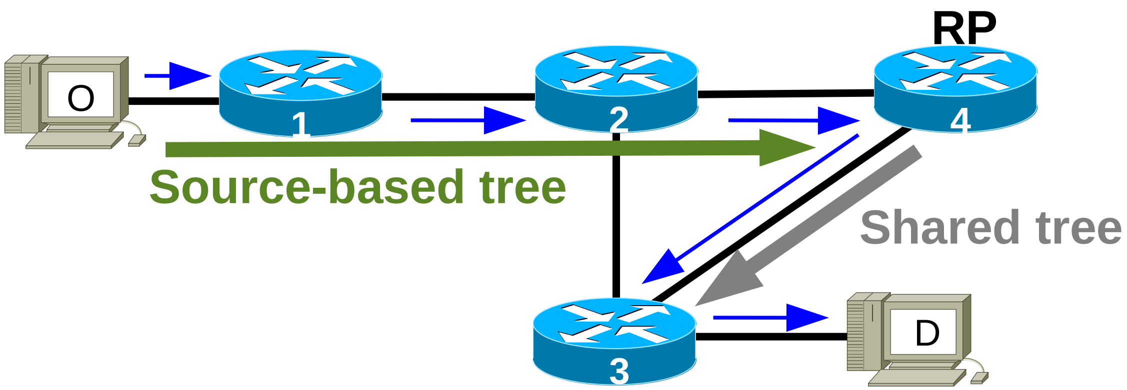
PIM SM – New Source

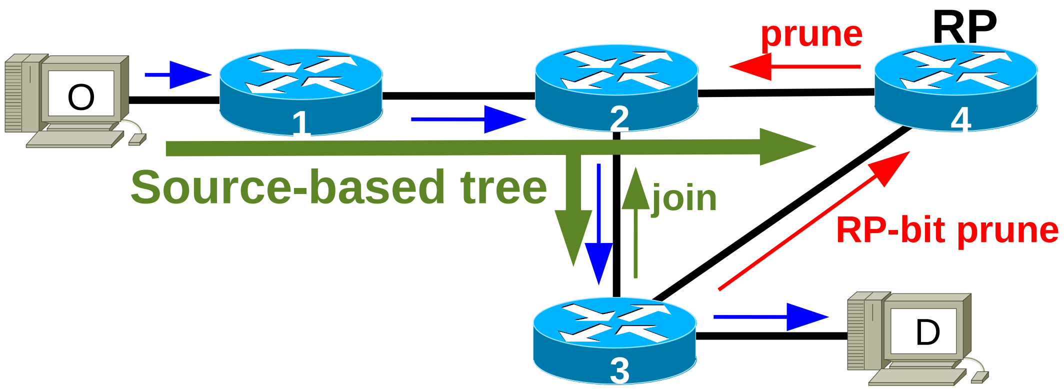




- When a new multicast source appears, the router that provides connection to the source, re-sends the multicast packets encapsulated in PIM packets (PIM *Register* message), and sends them in unicast to the RP.
- The RP desencapsulates the PIM *Register* messages, receives the multicast packets, and simultaneously:
  - Re-sends the multicast packets to the clients, using the Group-shared tree.
  - Just after receiving the first PIM *Register*, sends a *Join* message to the source router to create a source-based tree.
- As soon the source router starts to route the multicast packets over the source-based tree, and the RP receives them, the RP send a PIM *Register-Stop* message to notify the source router to stop sending the encapsulated multicast packets.

PIM SM – Commuting to *Source-based Tree*





- When the aggregated bit rate (of all sources) exceeds a predefined threshold, a router may chose to join the Source-based tree, leaving the Group-shared tree.
  - By sending a *Join* message towards the multicast source, until it finds a router that belongs to that source's 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.
  - Later any router may re-join the Group-shared tree.

Multicast em IPv6

- IGMP is replaced by Multicast Listener Discovery (MLD) protocol.
  - MLD is equivalent to IPV4 IGMP,
  - MLD was adapted to IPv6 semantics:
    - MLDv1 <=> IGMPv2,
    - MLDv2 <=> IGMPv3.
  - MLD messages are transported overICMPv6.
    - Multicast Listener Query, Multicast Listener Report and Multicast Listener Done.
  - MLD uses link local addresses as sources.
- Multicast routing tree may be *Sparse or Source-Specific.*
  - Dense-Mode is not available in IPv6!
  - Source-Specific requires MLDv2.
  - PIM protocols can also be used with IPv6 addresses.
    - PIM-SM and PIM-SSM, only.

MLD Messages

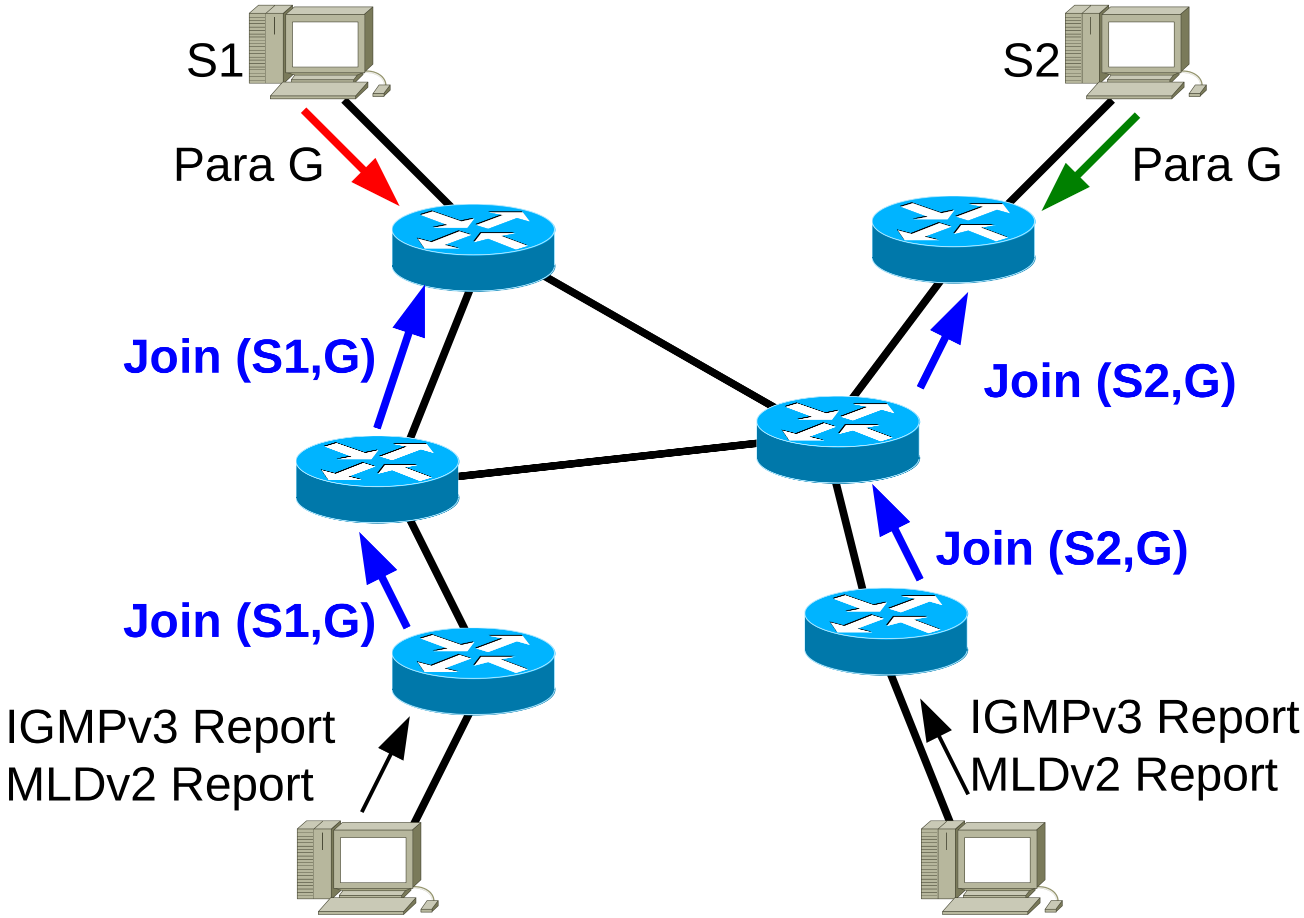
- MLD messages are used to determine group membership on a network segment, also known as a link or subnet.
- Multicast Listener Query
  - Sent by a multicast router to poll a network segment for group members. Queries can be general, requesting group membership for all groups, or can request group membership for a specific group.
- Multicast Listener Report
  - Sent by a host when it joins a multicast group, or in response to an MLD Multicast Listener Query sent by a router.
- Multicast Listener Done
  - Sent by a host when it leaves a host group and is the last member of that group on the network segment.

Source-Specific Multicast (SSM)

- SSM is an extension of the Multicast routing model.
  - Routing is uniquely made using Source-based trees.
  - The deployment of SSM with PIM (PIM-SSM) uses a sub-set of the PIM-SM mechanisms.
- Receivers specify not only the multicast session/group, but also a specific source.
  - The receiver router establishes a Source-based tree by sending a *Join* message to the specific source.
- Reserved SSM addresses:
  - IPv4 range - 232.0.0.0/8.
  - IPv6 range - FF3x::/32.
- Requires IGMPv3 or MLDv2.
  - Because client must specify the desired source.
    - IGMPv3 new INCLUDE mode functionality.

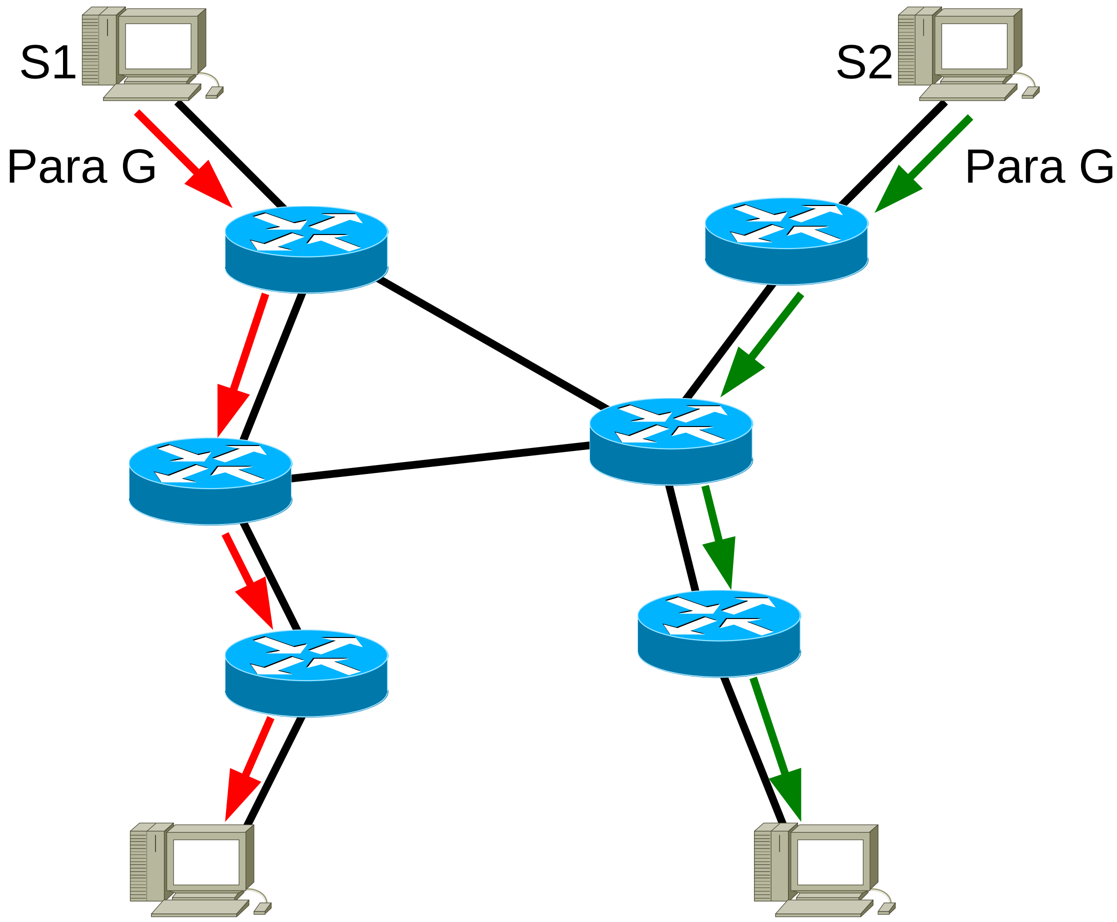
PIM-SSM Example (1)

- Reports are sent to join multicast group G from sources S1 or S2.
- Routers send PIM *Join* messages to construct the two distinct Source-based trees to group G from S1, and group G from S2.



PIM-SSM Example (2)

- After the Source-based trees creations, routers forward the multicast traffic to terminals using the respective source tree.

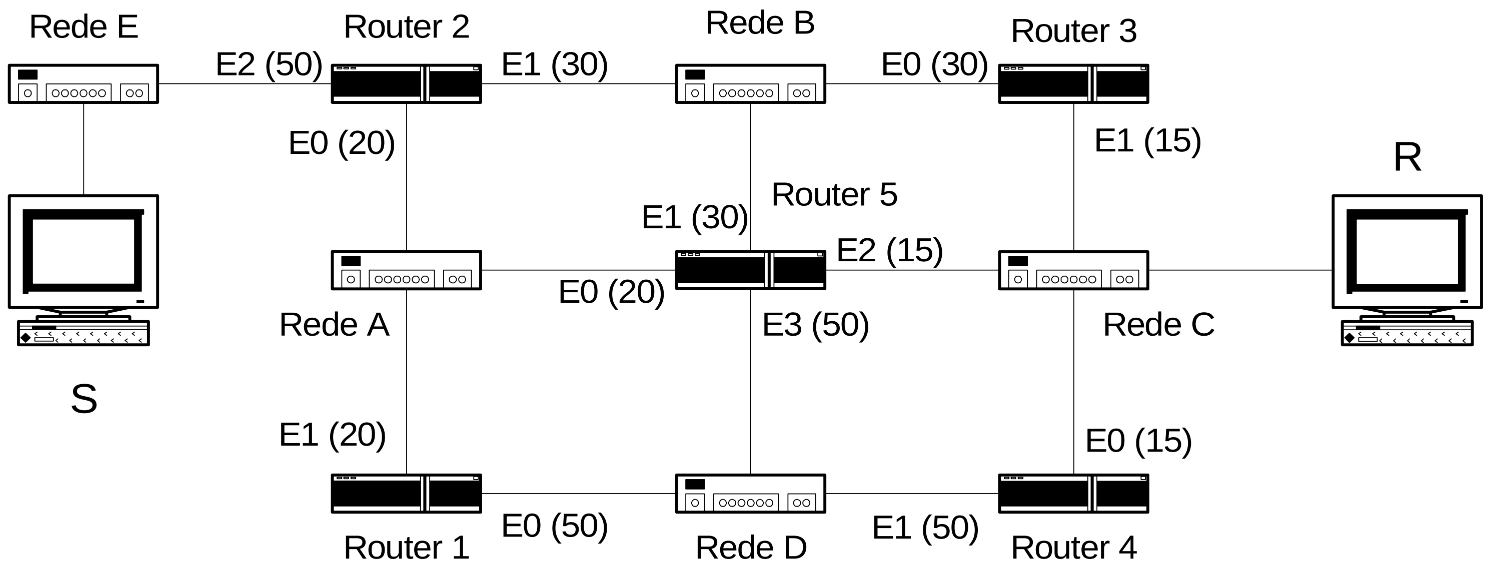


Example 1:

Na rede da figura, considere que os *routers* estão configurados com o OSPF (custos das portas indicados na figura) e com o PIM *dense-mode*. Assuma que o terminal S envia periodicamente pacotes para o endereço 230.20.1.1 e que o terminal R já aderiu à sessão *multicast* 230.20.1.1.

(1) Descreva como é que o primeiro pacote enviado por S se propaga pela rede.

(2) Apresente a tabela de encaminhamento *multicast* de cada um dos *routers; para cada router use a notação: O - G - E -* $S_1,S_2,\ldots$ *(O – endereço origem; G – grupo multicast; E – interface de entrada;* $S_1,S_2,\ldots$ *– lista de interfaces de saída)*



Example 2:

Na rede da figura, considere que os *routers* estão configurados com o OSPF (custos das portas indicados na figura) e com o PIM *sparse-mode* em que o endereço do *Rendezvous Point* é o endereço da interface E1 do *Router* 1. Indique que pacotes são trocados e qual o seu percurso:

(1) no início, quando o terminal R adere à sessão 225.3.3.3;

(2) depois, quando o terminal S envia um pacote para o destino 225.3.3.3.