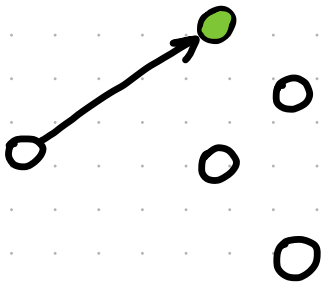
Advanced Topics in Communication Networks

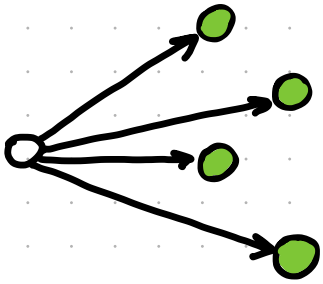
L9: IP MULTICAST | 17.11.2020

Prof. Laurent VANBEVER - nsg.ee.ethz.ch



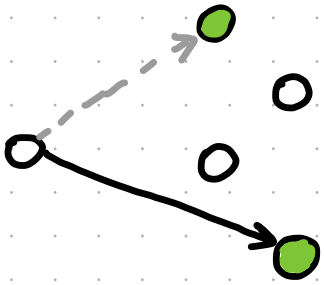
UNICAST

one-to-one



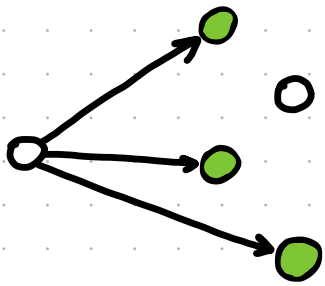
BROADCAST

one-to-all



ANYCAST

one-to-one-of-many



MULTICAST

one-to-many
many-

How do we efficiently transmit data to a set of receivers?

How do we efficiently transmit IP packets to a set of receivers?

Let's consider 2 possible solutions:

1. Source-based solution:
   - Sender simply sends as many copies of each IP packet as there are receivers.
   - Easy to implement / Waste a lot of bandwidth, not efficient!
2. Network-based solution:
   - Sender transmits one copy of each IP packet (as in IP unicast).
   - Network (i.e. the routers) take care of distributing this information to all the receivers.
   - Efficient (each packet crosses each link only once) / Hard to implement (need new protocols and forwarding mechanisms).

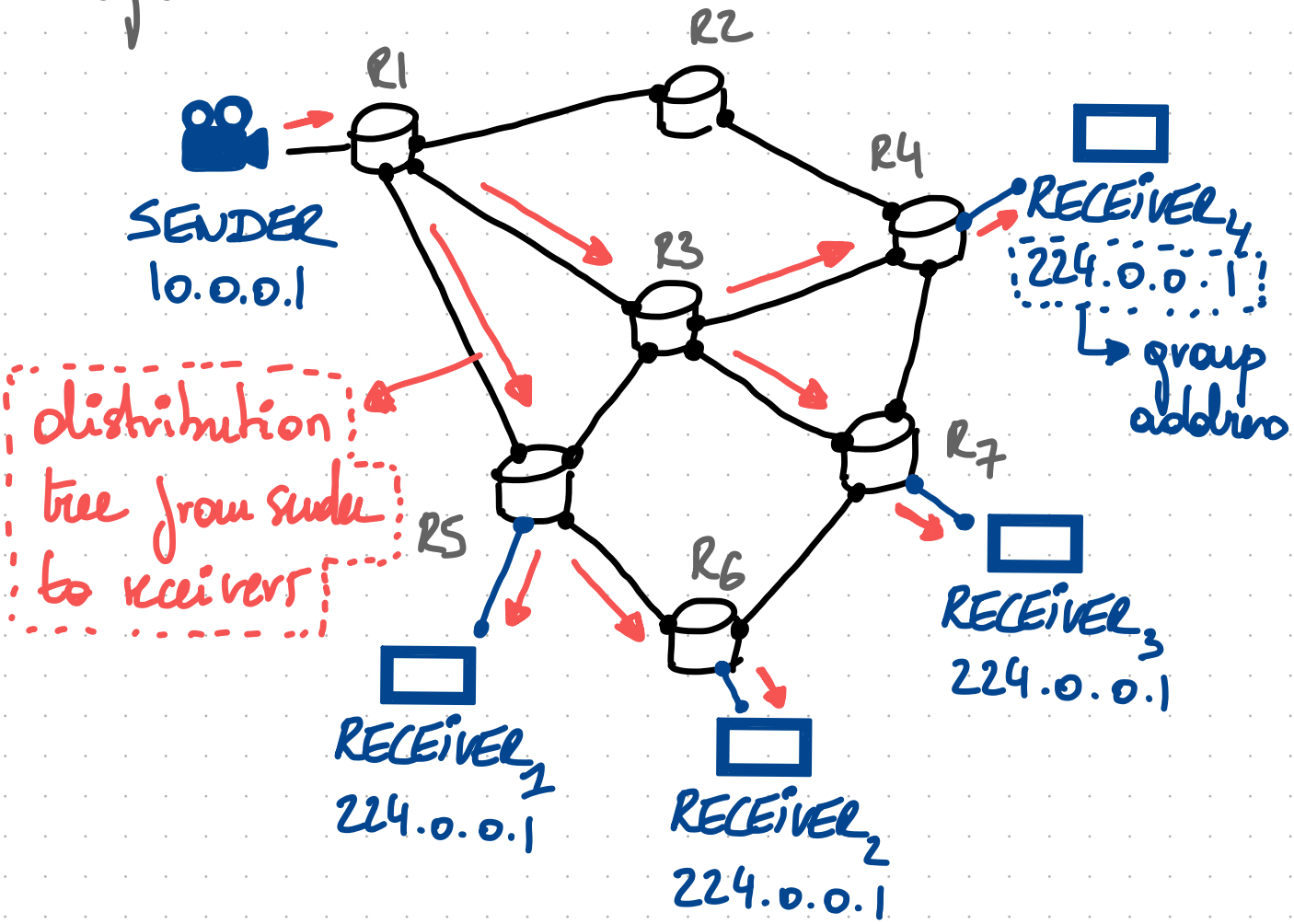
Network-based Multicast:

Principle: Sender sends multicast packets towards group of receivers identified by a group address.

Intermediate routers retransmit each received multicast packet such that:

- packets reach all receivers;
- packets traverse each link only once.

Example:



IP Multicast: Outline

1. How do we address a group of receivers?
2. How does a host receive traffic destined to a multicast address?
3. How do routers figure out which hosts belong to which group?
4. How do routers dynamically construct efficient distribution trees from senders to receivers?
   - 4.1. Pro-active solutions
   - 4.2. Reactive solutions
     - 4.2.1. "Flood and Prune"
     - 4.2.2. Rendez-vous Point
   - 4.3. Protocol Independent Multicast (PIM)

1/ How do we address a group of multicast receiver?

A subset of the IP space is reserved for multicast:

224.0.0.0 - 239.255.255.255
(leading bits "1110")

Some of these addresses have pre-defined allocation:

224.0.0.1 : All hosts
224.0.0.2 : All multicast routers
224.0.0.5 : All OSPF routers
...

Some of these addresses can only be used within an AS (they are not publicly routed)

239.0.0.0 - 239.255.255.255
(typically used for IPTV).

A multicast sender simply sends to a multicast IP address.

2/ How do the hosts receive multicast traffic?

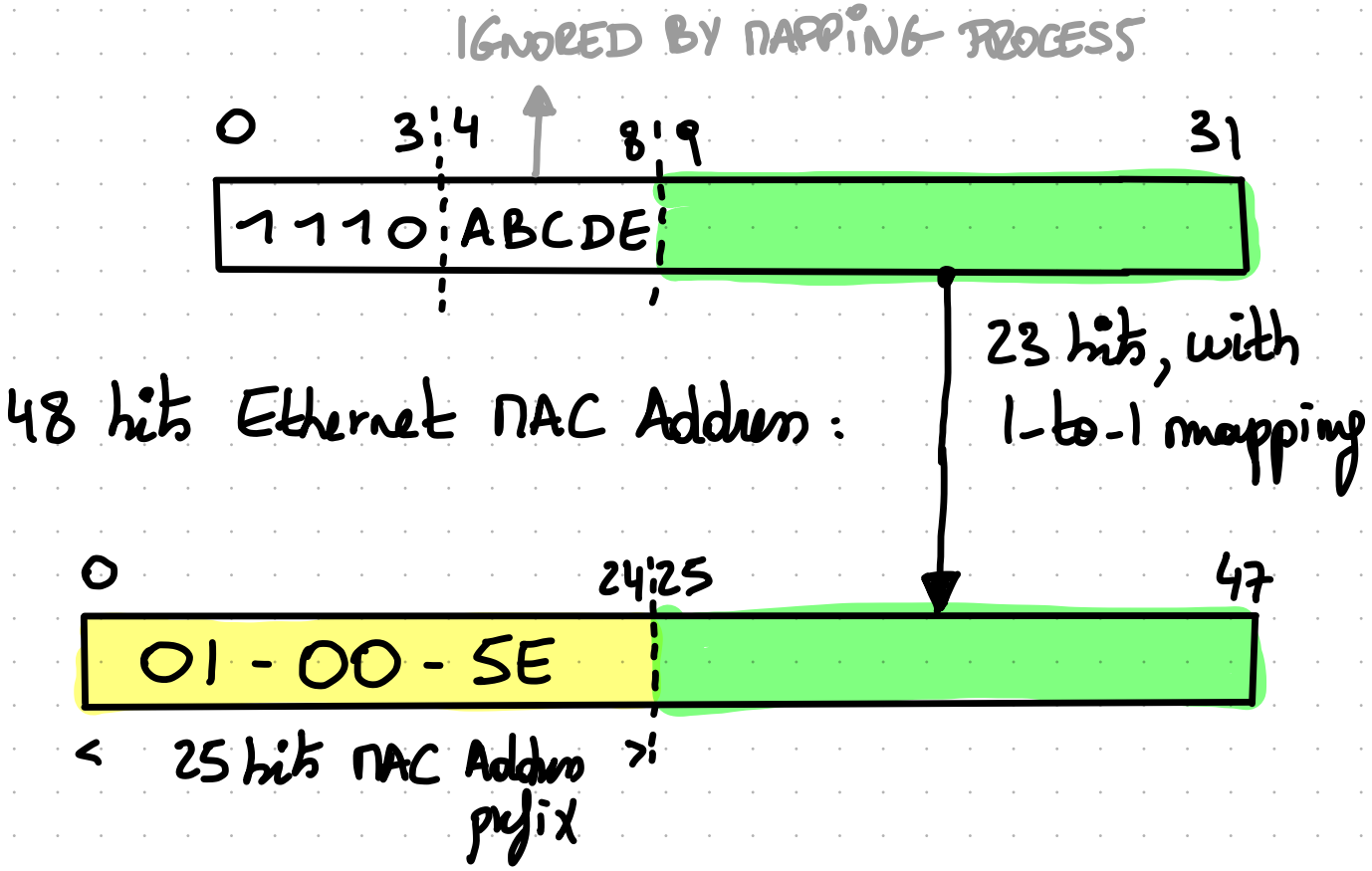
Ethernet addresses are of 2 kinds:

1. Physical addresses which identify one Ethernet adapter.
   -> These addresses start with "0" in the first byte.
2. Logical addresses which identify a group of Ethernet destinations.
   -> These addresses start with "1" in the first byte.

Ethernet adapters can be configured to capture frames whose destination is a set of these logical addresses. (in addition to their unicast address).

Hosts automatically figure out the logical MAC address from the IP multicast directly.

32 bits IPv4 Multicast Address:



Note that the mapping is not lossless: 32 IP Multicast addresses are mapped to the same Ethernet logical address.

→ This can lead to unwanted traffic.

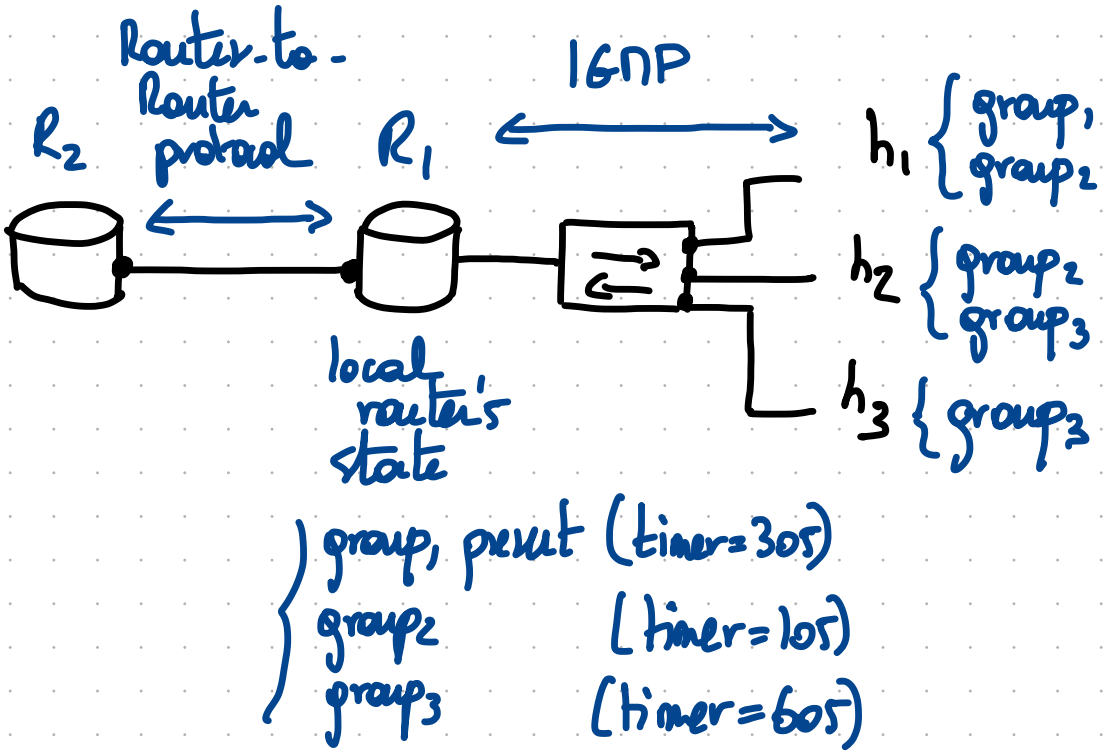
3. How do routers figure out which hosts belong to which group?

Internet Group Management Protocol (IGMP):

IGMP is a protocol used by hosts and adjacent routers to create multicast group membership.

It is quite simple:

- Hosts request membership to a group (i.e. a multicast IP address).
- Adjacent routers listen and keep track of these requests. They also periodically send out subscription queries. (one router is elected to do that)
- Adjacent routers then use Protocol Independent Multicast (PIM) to direct traffic from hosts sending multicast traffic to the hosts that have registered for it.



There are TONS of details about IGMP like how reports/responses are advised, how hosts can leave groups, how timers are organized,...

4/ How do we dynamically construct efficient shortest-path distribution tree from senders to receiver?

Challenge: How do routers learn about where the various receivers are and keep track of them over time.

Two possible solutions:

4.1. Pro active: A routing protocol is used to distribute group membership so that each router knows the exact location of each group member. One can extend link-state protocols for that, e.g. MOSPF.

4.2. Reactive: Assume that group members are everywhere initially → Broadcast the traffic. If a router receives unwanted traffic, it asks the upstream router to stop.

4.1. Pro-actively building distribution trees using link-state routing protocols

Principle:

1. Routers collect membership info using IGMP.
2. Group membership is flooded by link-state protocols using a new type of messages.
3. Each router computes the shortest path tree (S,G) for each source S and each group G.

Note the shortest path tree (S,G) is built on demand, whenever the router receives a packet destined to G.

Pros: Routers have full knowledge.

Cons:

- Possibly important memory overhead on ALL routers, independently on whether or not they see any multicast traffic.
- The flooding of group membership msgs compete with normal link-state msgs (e.g. link down...)
- As the number of sources and groups grows, computing shortest path trees can become problematic.

4.2. Reactively building distribution trees:

2 possible solutions:

4.2.1 "Flood-and-Prune"

Principles:

1. Flood the multicast traffic in the entire network.
2. Prune branches when there is no receivers.

4.2.2. Use rendez-vous point.

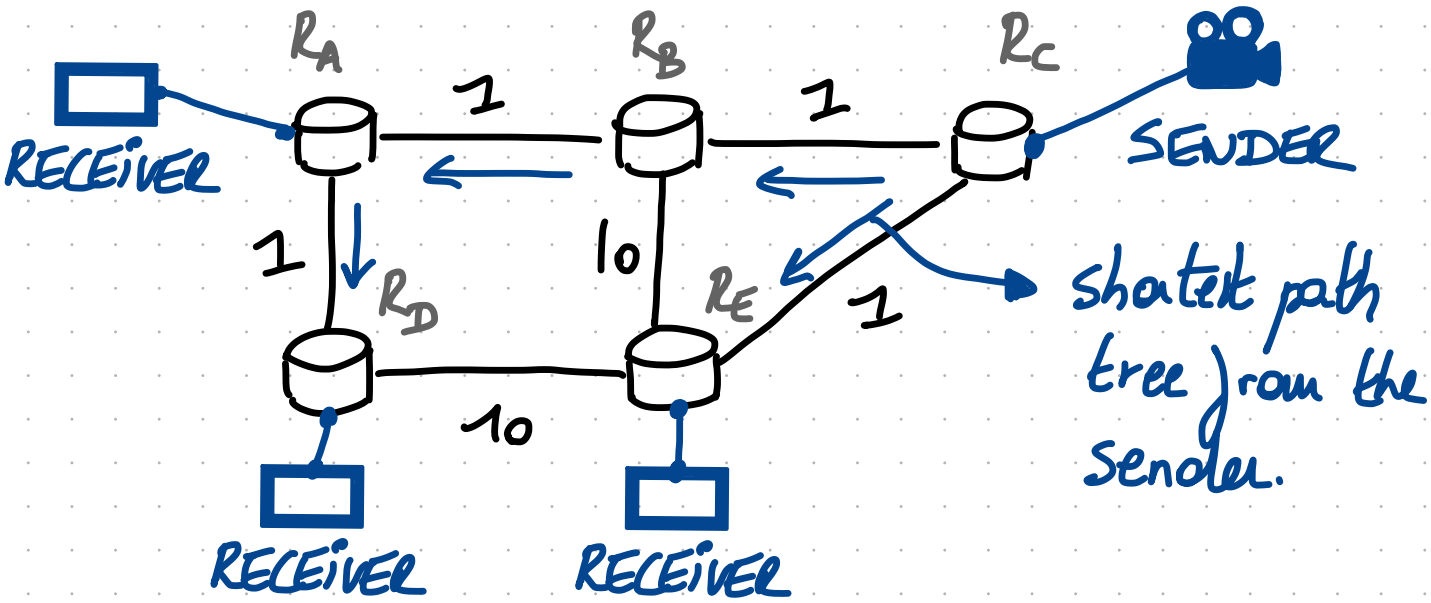
Principles:

1. Have one router acts as root of a shared distribution tree per group.
2. Have the sources encapsulate the traffic to the root.
3. Have the root multicast the encapsulated traffic along side the shared traffic.

42.1. Flood and Prune

How do we broadcast traffic in a large network?

Goal: Broadcast traffic following the shortest path tree.



Insight: Flood traffic only when it arrives from the shortest-path upstream.

This strategy is known as:

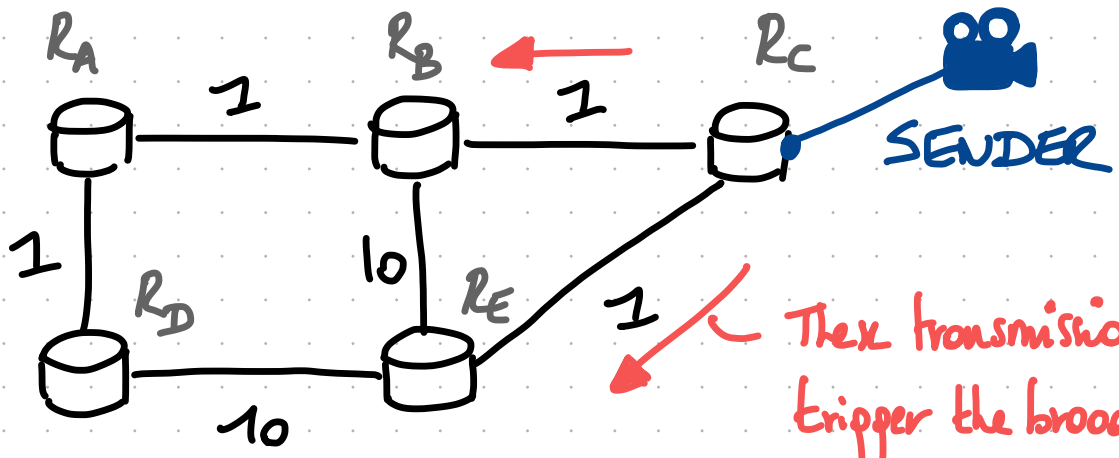
Reverse Path Filtering or RPF

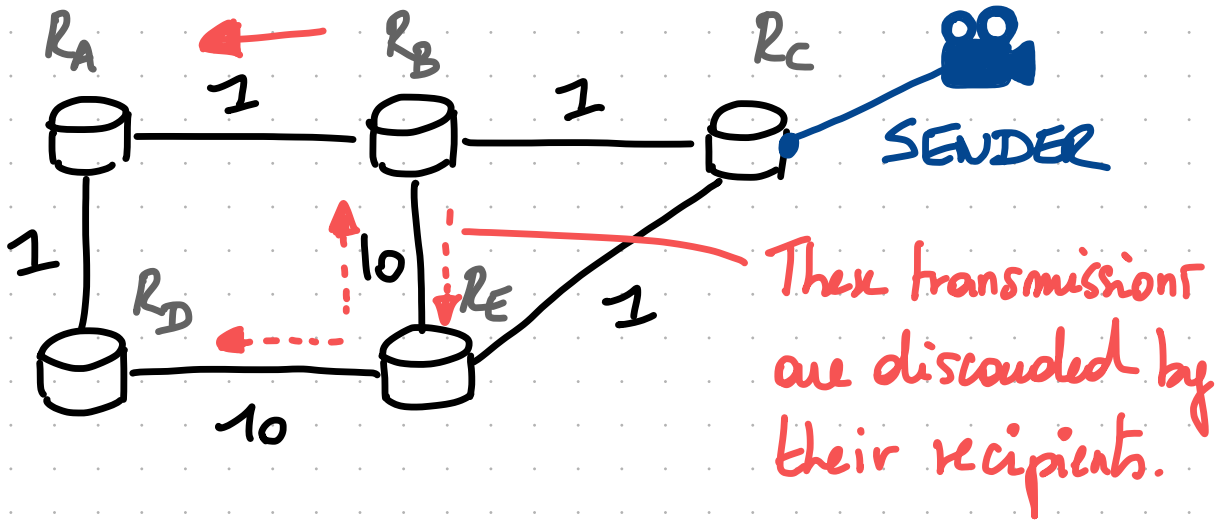
RPF Algorithm:

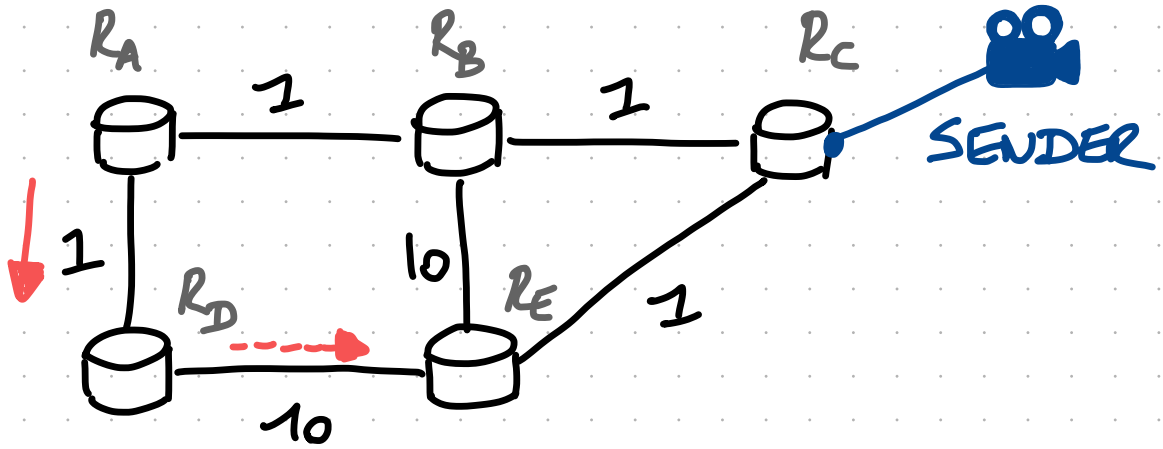
Upon receiving an IP packet from source S on interface in:

- if (in == next_hop_interface (S)):
  - for (interface ij in interfaces):
    - if ($ij \neq in$):
      - send_packet (ij)

An interesting observation is that, unlike unicast routing which depends solely on the destination, multicast routing depends solely on the source address!


R_A
R_B
R_C
SENDER
1
1
1
1
10
10
R_D
R_E
1
These transmissions trigger the broadcast.


R_A
R_B
R_C
SENDER
1
1
1
1
10
10
R_D
R_E
1
These transmissions are discarded by their recipients.


R_A
R_B
R_C
SENDER
1
1
1
1
10
10
R_D
R_E
1

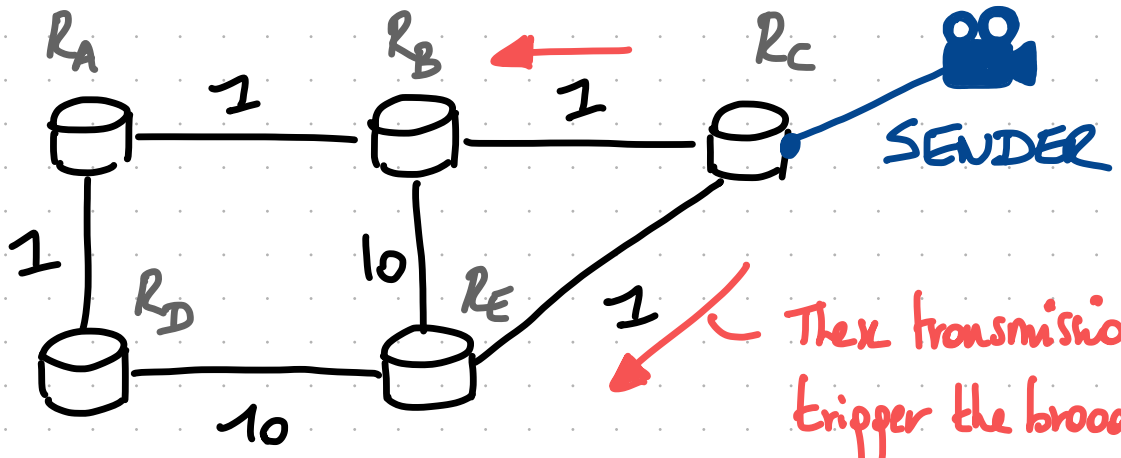
How do we avoid unnecessary transmissions?

Sender-based solution:

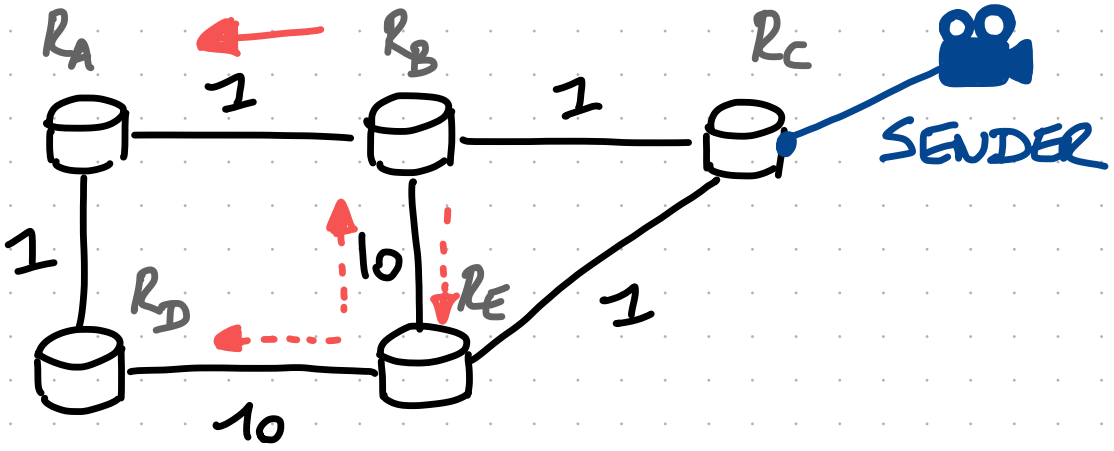
For each possible source, each router computes the list of interfaces on which to broadcast such that it only includes the interfaces on the shortest path tree from the source.
For that routers can rely on the network topology learned through e.g. OSPF or IS-IS. Note that this is different from IGMP. (It has nothing to do with group membership).

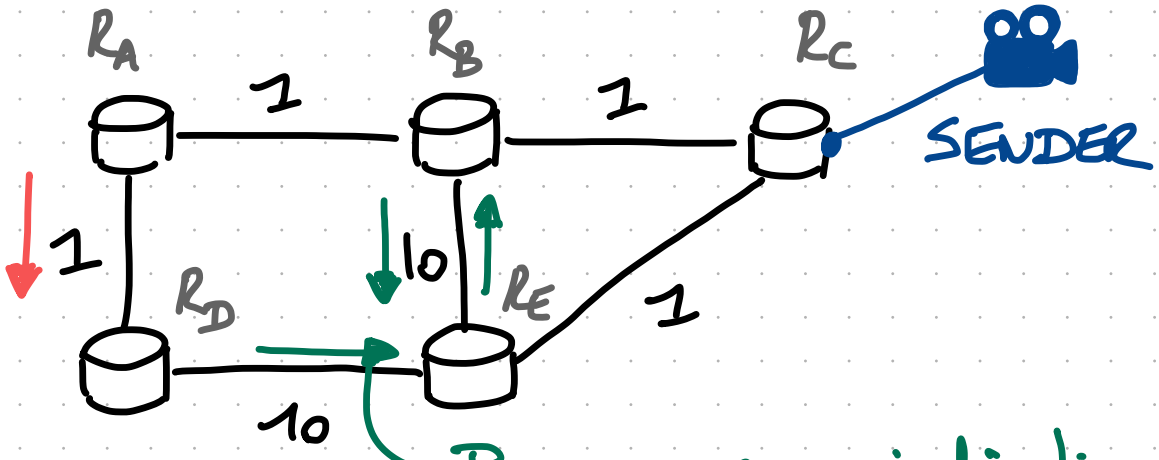
Receiver-based solution:

For each possible source, each router learns the list of outgoing interfaces on which to broadcast by having downstream routers tell them that a given interface is <u>not</u> on the shortest path.
Doing so is less intensive sender-based solution.



These transmissions trigger the broadcast.





Prune message indicating to the upstream router that packets coming from SENDER should not be broadcasted on these interfaces anymore.

In practice, two types of PRUNE messages are sent:

1. Source-specific message if packets are received on non shortest paths.
2. Group-specific message if there is no group member downstream of a router.

The list of outgoing interfaces is periodically refreshed and flooding returns.

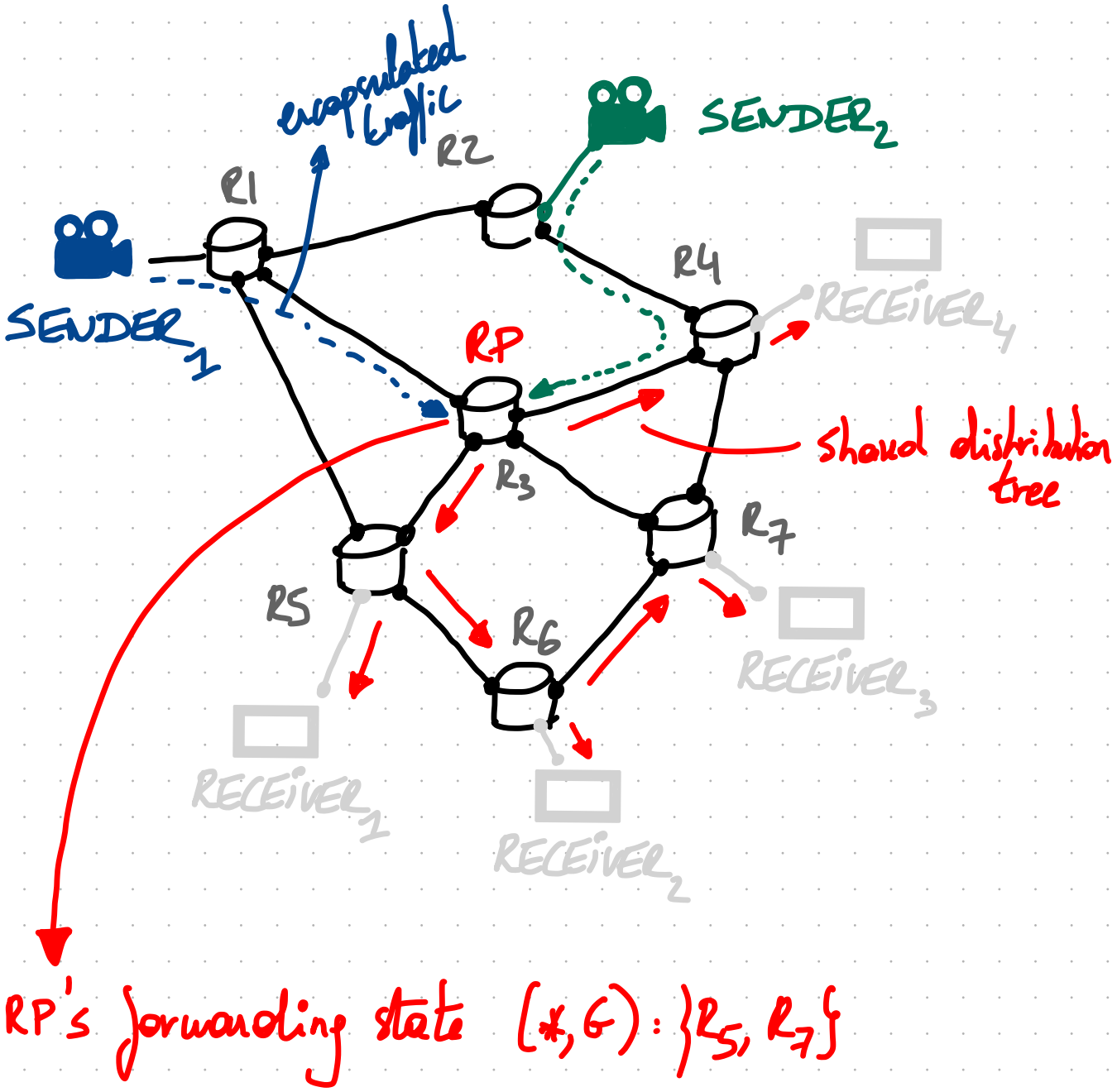
Pros: Approach is simple, "plug and play"

Cons: Approach is relatively costly:
- Routers need to maintain per (S,G) state
- Flooding is frequently reactivated.

4.2.2. Use shared trees and rendez-vous points

Principle :

- One router is configured as a Rendez-vous point (RP).
  - All routers know the RP address.
  - RP acts as the root of a shared tree for the group. This tree is denoted as (*, G).
  - Multicast routers encapsulate packets sent by hosts to G and send them to the RP.
  - RP redistributes these packets over the shared tree (*, G).
  - Receivers dynamically join the shared tree.



Note that RPs need to be configured in advance a a mechanism has to be put in place to discover them.

4.3. Protocol Independent Multicast (PIM):

PIM is the most widely-deployed multicast routing protocol.

PIM does not rely on a specific unicast routing protocol: it leverages the existing unicast routing table (which is populated by whatever protocol) to perform receiver-based optimization for RPF.

PIM has two "modes": - DENSE
- SPARSE

PIM DENSE works by periodically flooding and pruning so as to build source-specific distribution trees.

PIM SPARSE initially builds a shared distribution tree. The shared tree is rooted at a (typically pre-configured) Rendez-vous point.

If required, PIM SPARSE allows for the shared tree to be converted to a shortest-path tree.