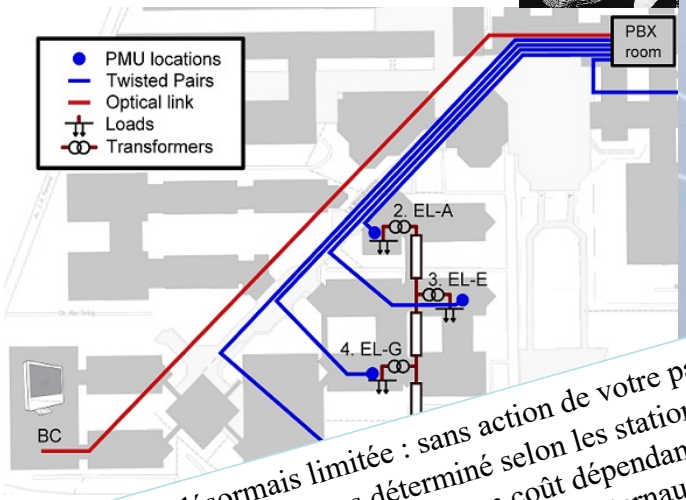
IP Multicast

Jean-Yves Le Boudec

2022

#Mahsa_Amini



EPFL

La durée d'écoute est désormais limitée : sans action de votre part (un simple clic), la diffusion s'arrête au bout d'un temps déterminé selon les stations. En effet, pour nous, diffuseurs, les technologies actuelles imposent un coût dépendant de la durée et du nombre d'auditeurs. Plusieurs éléments nous indiquent que les internautes ayant accès à l'internet illimité ne coupent pas l'écoute, lorsqu'ils quittent leur ordinateur allumé. Radio France ne peut continuer à financer pour celui qui n'écoute pas. C'est pourquoi nous avons mis en place ce système de confirmation, un peu contraignant, mais qui nous permet de mieux contrôler les coûts de diffusion.

http://viphttp.yacast.net/V4/radiofrance/fip_bd.m3u

IP Multicast

Unicast = send to one destination

Multicast = send to a *group* of destinations

IP has multicast addresses:

224.0.0.0/4 (i.e. 224.0.0.0 to 239.255.255.255) and ff00::/8

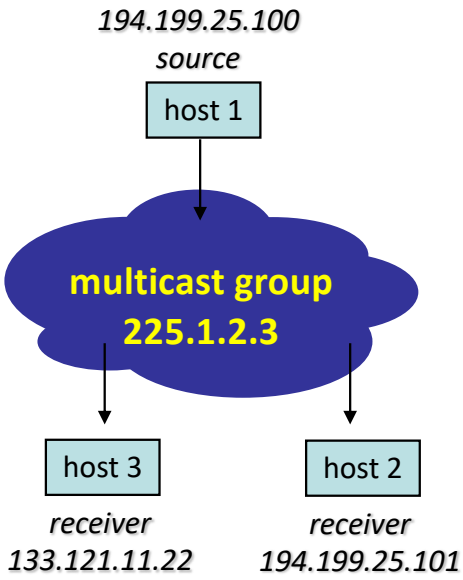
For IPv6, bits 13-16 = scope, e.g. ff02/16 = link local, ff05/16 = site local

An IP multicast address is used to identify a group:

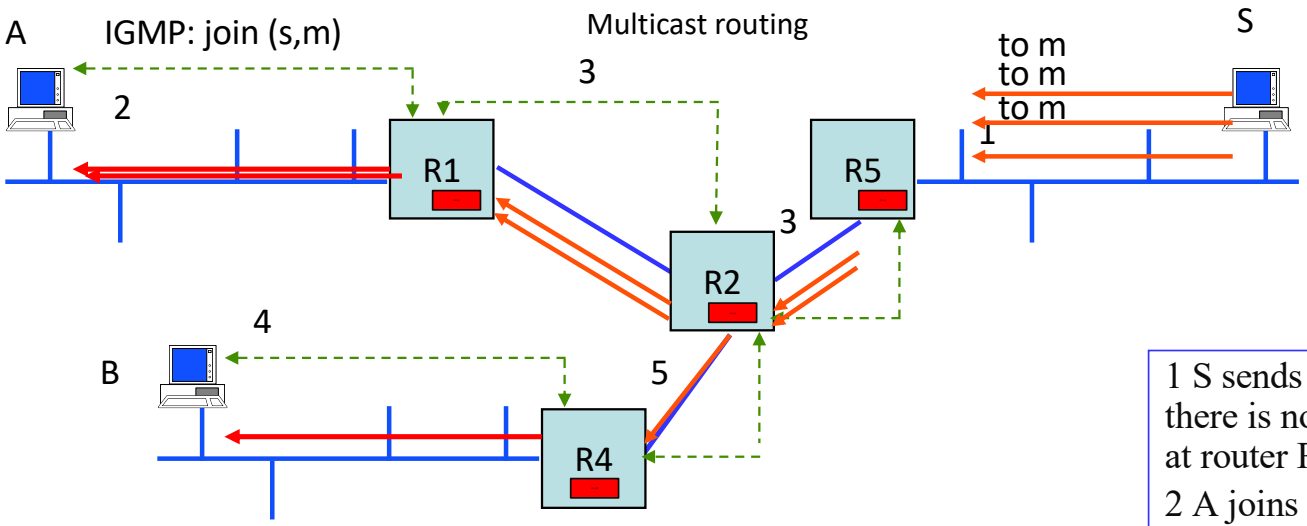
Any Source Multicast (ASM): the group is identified by the multicast address. Any source can send to this group.

Source Specific Multicast (SSM): the group is identified by (s, m) where m is a multicast address and s is a (unicast) source address. Only s can send to this group.

By default 232.0.0.0/8 and ff3x::/96 are SSM addresses (x=scope bits. e.g. ff35::/96 = site-local). See RFC7371.



Operation of IP Multicast: destinations need to explicitly join multicast group



source simply sends one single packet for *n* destination

destinations subscribe via IGMP(Internet Group Management Protocol, IPv4) or MLD (Multicast Listener Discovery --IPv6); **join** messages sent to router

routers either build distribution tree via a multicast routing protocol (PIM-SM) or use tunnels (BIER)

packet multiplication is done by routers

1 S sends packets to multicast address m; there is no member, the data is simply lost at router R5.

2 A joins the SSM group (s,m).

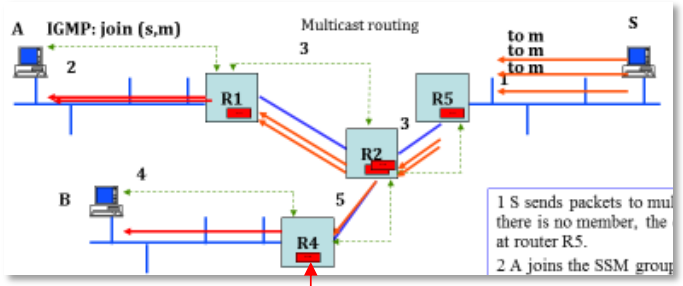
3 R1 informs the rest of the network that (s,m) has a member at R1 using e.g. the multicast routing protocol PIM-SM; this results in a tree being built. Data sent by S now reach A.

4 B joins the multicast address m.

5 R4 informs the rest of the network that m has a member at R4; the multicast routing protocol adds branches to the tree. Data sent by S now reach both A and B.

Traditional Multicast enabled Routers Must Keep Additional State Information

In addition to IP principles #1 and #2,
a multicast IP router does exact match for multicast groups.



Multicast state information is kept in router for every known multicast group:

(s, m) or (*,m)	// id of group
valid incoming interfaces	// for security
outgoing interfaces	// this is the routing info
other information required by multicast routing protocol	

This per-flow state information is dynamic and cannot be aggregated : scalability is an issue.

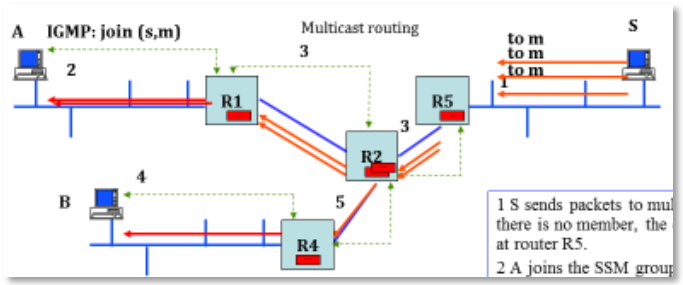
Multicast Routing

There are many multicast routing protocols. Widespread is PIM: Protocol Independent Multicast. It supports ASM and SSM and exists in two versions: sparse and dense.

PIM-DM (Dense Mode) makes heavy use of broadcast and can be used only in small, tightly controlled networks.

PIM-SM (Sparse Mode) is more reasonable and is used e.g. for TV distribution.

When used with SSM, PIM-SM is very simple: it uses Reverse Path Forwarding: when a router (such as R1) needs to add a receiver, it sends a PIM/JOIN towards the source, using unicast routing. This creates the distribution tree on the fly.



PIM-SM for ASM is more complicated; it uses one multicast router as Rendez-vous Point (RP): destination routers create a tree from RP, using RPF; router closest to source sends source packets to RP; if there exists an interested receiver in the domain, RP creates a tree from source (using RPF) otherwise drops; destinations create trees from sources, using RPF.

Alternative to Multicast Routing: BIER

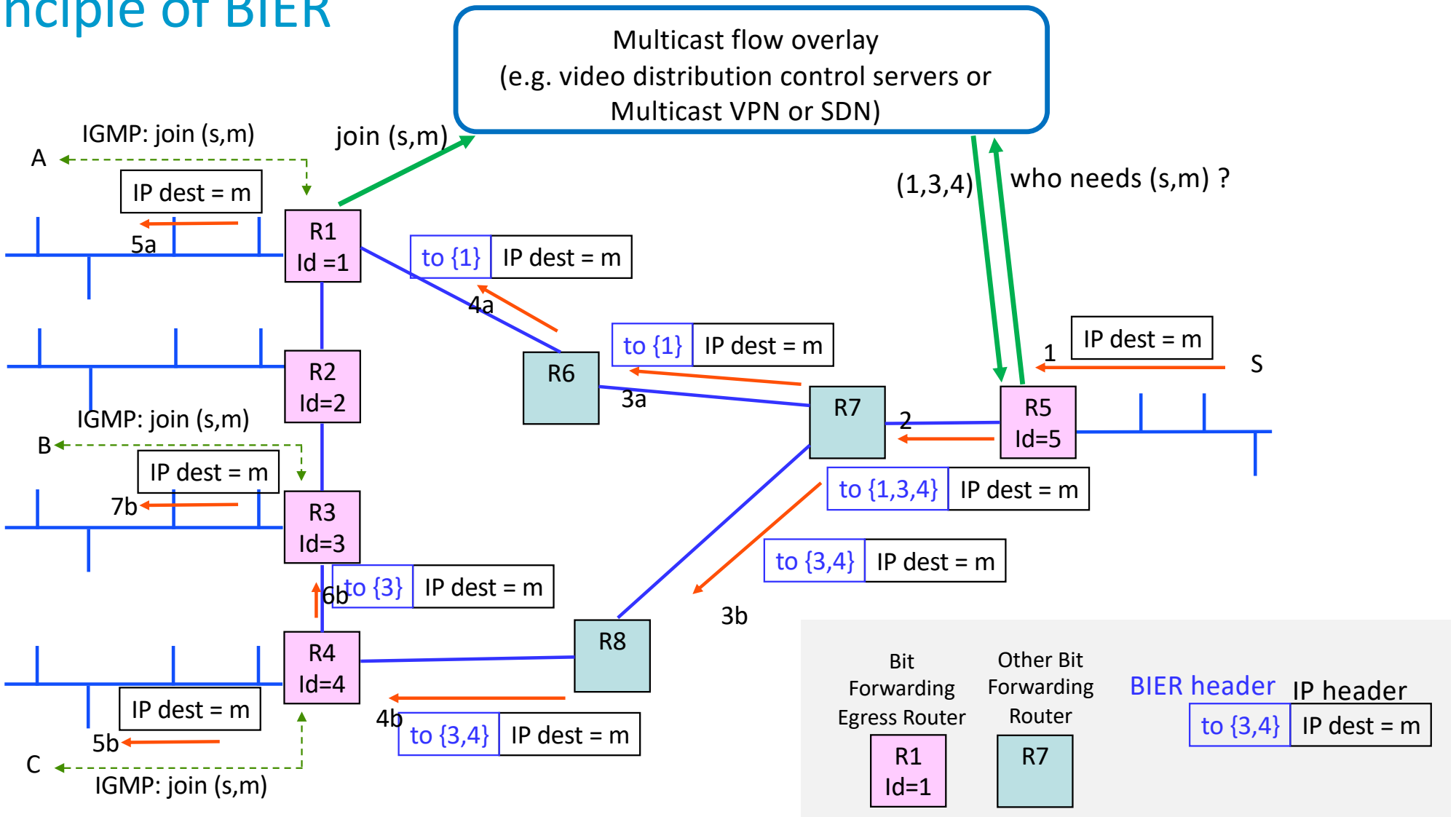
Why ? Multicast routing requires routers to keep per-flow state (dynamic, depends on who listens to the group).

BIER (Bit Index Explicit Replication) is an alternative that avoids this.

How ?

- A multicast packet receives a BIER header that contains (roughly speaking) the list of destination BIER routers (Bit-Forwarding Egress Routers, BFERs)
- BIER routers use unicast routing in order to deliver packets to the set of BFERs indicated in the BIER header
- If several BFERs are reached by the same next hop, only one packet is sent to next hop
- BIER routers multiply packets if the BFERs are reached via different next-hops. In this case the BIER header is modified accordingly.

Principle of BIER



Group membership information is distributed by an external infrastructure called “multicast flow overlay”, for example: a special set of servers used to control the distribution of the video content, or a system to manage multicast virtual private networks using MPLS and BGP (see later in MPLS module) or a central network controller (SDN).

All routers on the figure are BIER routers (Bit Forwarding Routers, BFRs). Routers R1-R5 are egress routers (they need to forward multicast packets to the outside); such routers have a BFR-id, for example R1’s BIER-id is 1.

Router R5 learns from the multicast flow overlay that the group (s,m) has members in routers with BFR-ids 1,3 and 4.

1. Router R5 has an IP multicast packet to send to m, from s. R5 knows that it should send the packet to R1, R3 and R4. From unicast routing, R5 finds that R1, R3 and R4 are reached via the same next-hop (R7) therefore R5 sends one packet with BIER header (1,3,4) to R7.

2. From unicast routing, R7 finds that R3 and R4 are reached via the same next-hop (R8) but R1 requires a different next-hop. Therefore, R7 duplicates the packet and creates 2 packets, one with BIER header (1), sent to R6, and one with BIER header (3,4), sent to R8.

3a. R6 forward the packet to R1.

4a, 5a. R1 belongs to the destination list contained in the BIER header, therefore R1 knows it should forward the packet using native multicast. The BIER header is removed and the packet is forwarded to the west LAN interface where A can receive it.

3b. From its forwarding table, R8 finds that R3 and R4 are reached via the same next-hop (R4). Therefore, R8 sends the packet to R4.

4b, 5b. R4 belongs to the destination list contained in the BIER header, therefore R4 knows it should forward the packet using native multicast. The BIER header becomes (3). 5b. The BIER header is removed and the packet is forwarded to R4’s west LAN interface where C can receive it.

6b. R4 sends one copy of the packet to R3.

7b. The BIER header is removed and the packet is forwarded to R3’s west LAN interface where B can receive it.

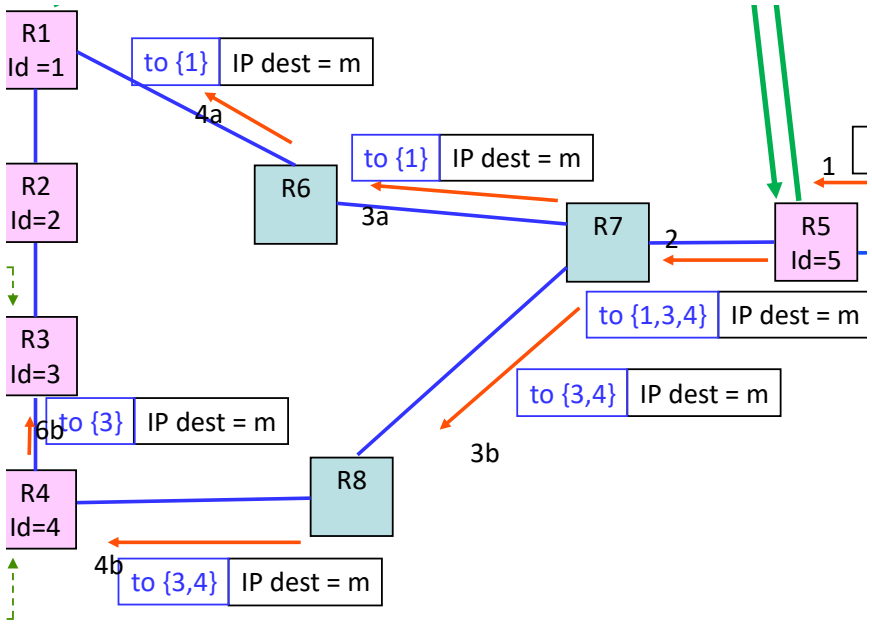
BIER : Packet Forwarding

A BIER Router pre-computes a *"forwarding bit mask"* that indicates, for every destination BFER, the set of destination BFERs that are reached by the same next-hop (as obtained from the IP forwarding table). When BIER router has a packet to forward, with destination set S:

1. Send packet to 1st destination in S; packet has destination set = $S \cap S_1$, where S_1 is the forwarding bit mask of 1st destination.
2. If $S \setminus S_1 \neq \emptyset$, duplicate the packet but with destination $S \setminus S_1$ and goto 1; else leave.

Example of BIER forwarding at router R7:

1. R7 Receive packet with destination set S ={1,3,4}
2. Least significant destination is 1, R7 uses Bier Index Forwarding table and finds that the forwarding bit mask for destination 1 is $S_1 = \{1,2\}$: sends a copy of packet to next-hop (R6), with destination $S' = S \cap S_1 = \{1\}$;
3. $S'' = S \setminus S_1 = \{3,4\}$ is not empty; R7 duplicates the packet with destination set S'' and applies the same procedure → Least significant destination is 3, send packet to next-hop (R8) with destination set $\{3,4\}$. Now $\{3,4\} \setminus \{3,4\} = \emptyset$, so the procedure ends.



Bier Index Forwarding Table at R7

Dest. BFER	Forwarding bit mask	Next-Hop
1	{1,2}	R6
2	{1,2}	R6
3	{3,4}	R8
4	{3,4}	R8
5	{5}	R5

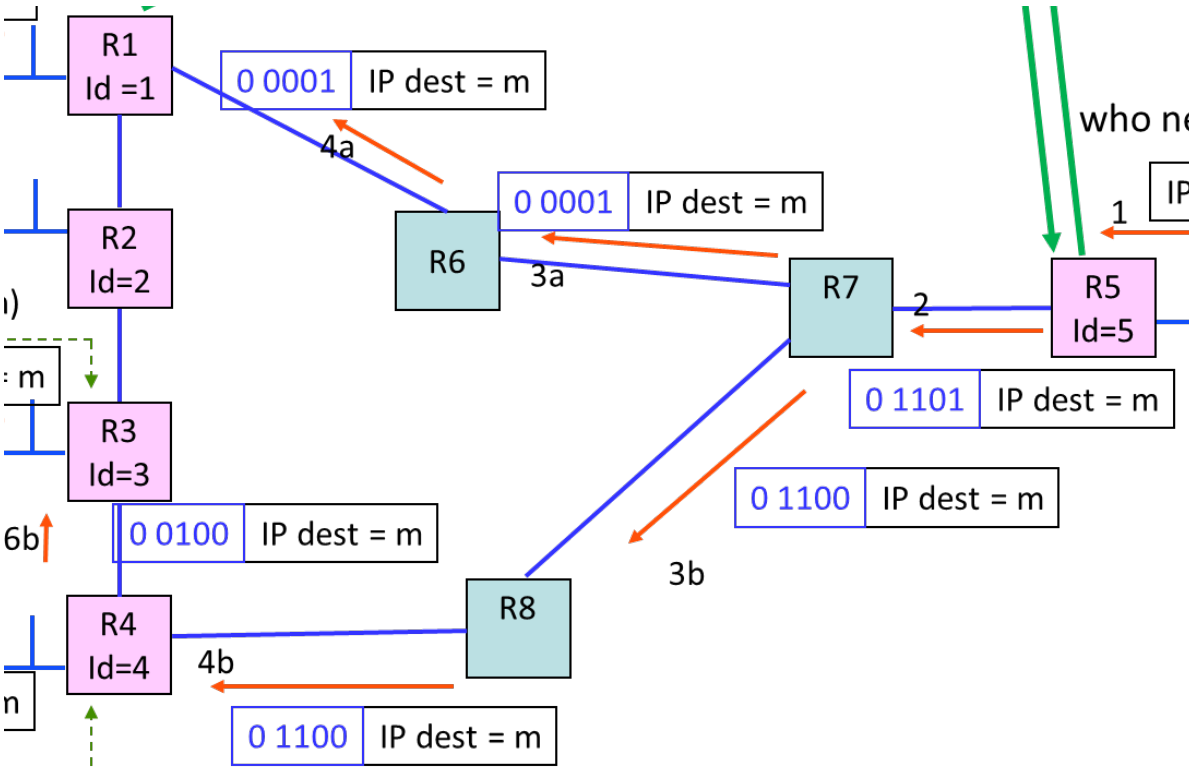
IP Forwarding Table at R7

Dest. IP addr	Next-Hop
R1	R6
R2	R6
R3	R8
R4	R8
R5	R5

How BIER optimizes processing by BIER Routers

- A set of destination (BFERs) is encoded using a *bitstring*. Example with 5 possible BFERs:
  Destination BFERs = {1,3,4}
  → bitstring = 0 1101
  Destination BFERs = {3,4}
  → bitstring = 0 1100
  Destination BFERs = {1}
  → bitstring = 0 0001

- Additional mechanisms exist when the number of BFERs is large, see RFC 8279.

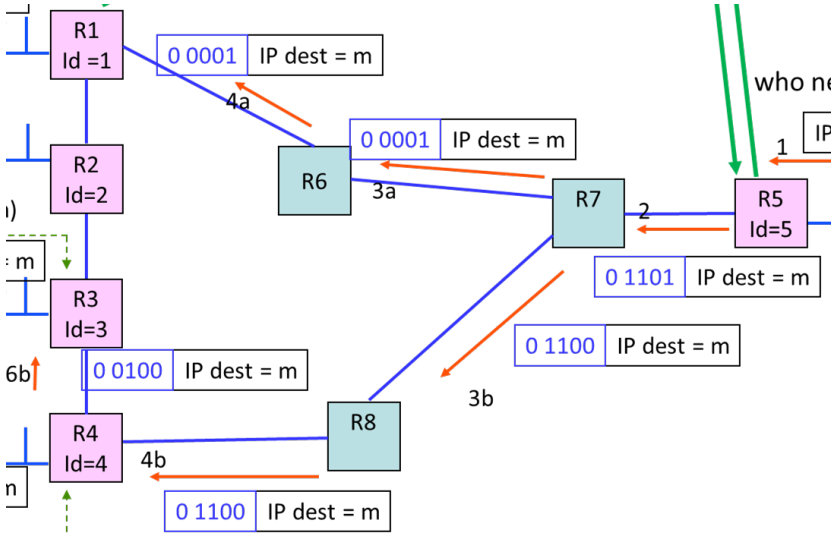


How BIER optimizes processing by BIER Routers (cont'd)

- A BIER Router pre-computes a *forwarding bit mask* that indicates, for every destination BFER, the set of destination BFERs that can be reached with a single packet. Stored in Bier Index Forwarding Table (static). No per-flow state in table.

Example of BIER forwarding at router R7:

1. R7 Receive packet with bistring = 0 1101
2. Least significant destination is 1, R7 uses Bier Index Forwarding table: sends one copy of packet to next-hop (R6); ANDs the bitstring with forwarding bit mask; packet sent to R6 has bitstring = 0 0001; R7 computes remaining bitstring (ANDs the original bitstring with inverse of forwarding bit mask) and obtains 0 1100;
3. R7 processes again the packet with the remaining bistring = 0 1100; same procedure → Least significant destination is 3, send packet to next-hop (R8) with bistring 0 1100; compute remaining bitstring = 0 0000. End of procedure.



Dest. BFER	Forwarding bit mask	Next-Hop
1	0 0011	R6
2	0 0011	R6
3	0 1100	R8
4	0 1100	R8
5	1 0000	R5

Bier Index Forwarding Table at R7

BIER Routers

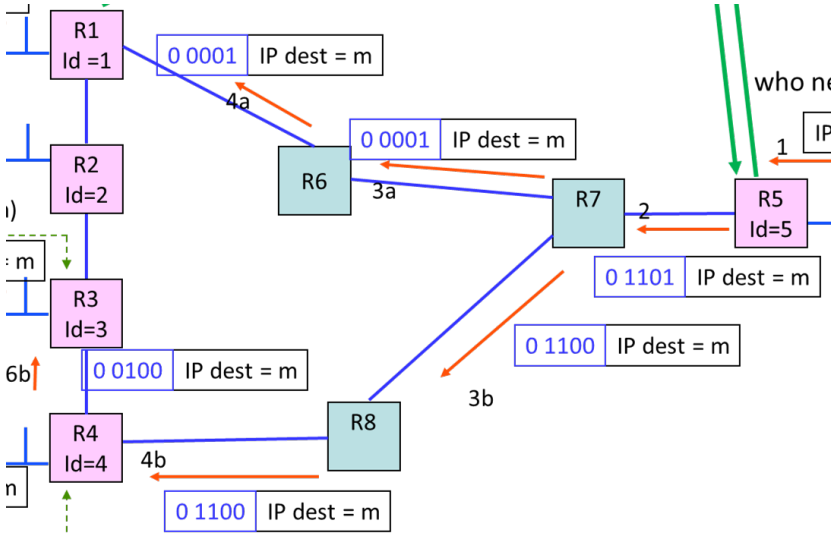
In addition to IP principles #1 and #2,
a BIER router does bitstring processing using its bit forwarding table.

A Bit Forwarding Ingress Router (such as R5) must map destination multicast address to a BIER header (dynamic). Requires out of band mechanism. This is the only per-flow information required by BIER.

Bit forwarding table is automatically derived from this router's unicast IP forwarding table.

Inside a BIER domain, multicast packets have an additional header and the IP destination address is not used (≈ tunnel).

Multiple BIER domains can be interconnected by BFIR-BFER interconnection.



Dest. BFER	Forwarding bit mask	Next-Hop
1	0 0011	R6
2	0 0011	R6
3	0 1100	R8
4	0 1100	R8
5	1 0000	R5

Bier Index Forwarding Table at R7

Is there Multicast ARP ?

Recall ARP = find MAC address that corresponds to an IP address;
here the target MAC address is a multicast MAC address.

There is no ARP for multicast. IP multicast address is algorithmically mapped to a multicast MAC address.

- Last 23 bits of IPv4 multicast address are used in MAC address
- Last 32 bits of IPv6 multicast address are used in MAC address

MAC multicast addr.	*Used for*
01-00-5e-XX-XX-XX	IPv4 multicast
33-33-XX-XX-XX-XX	IPv6 multicast

IP dest address	229.130.54.207
IP dest address (hexa)	e5-82-36-cf
IP dest address (bin)	...-10000010-...
Keep last 23 bits (bin)	...-00000010-...
Keep last 23 bits (hexa)	02-36-cf
MAC address	01-00-5e-02-36-cf

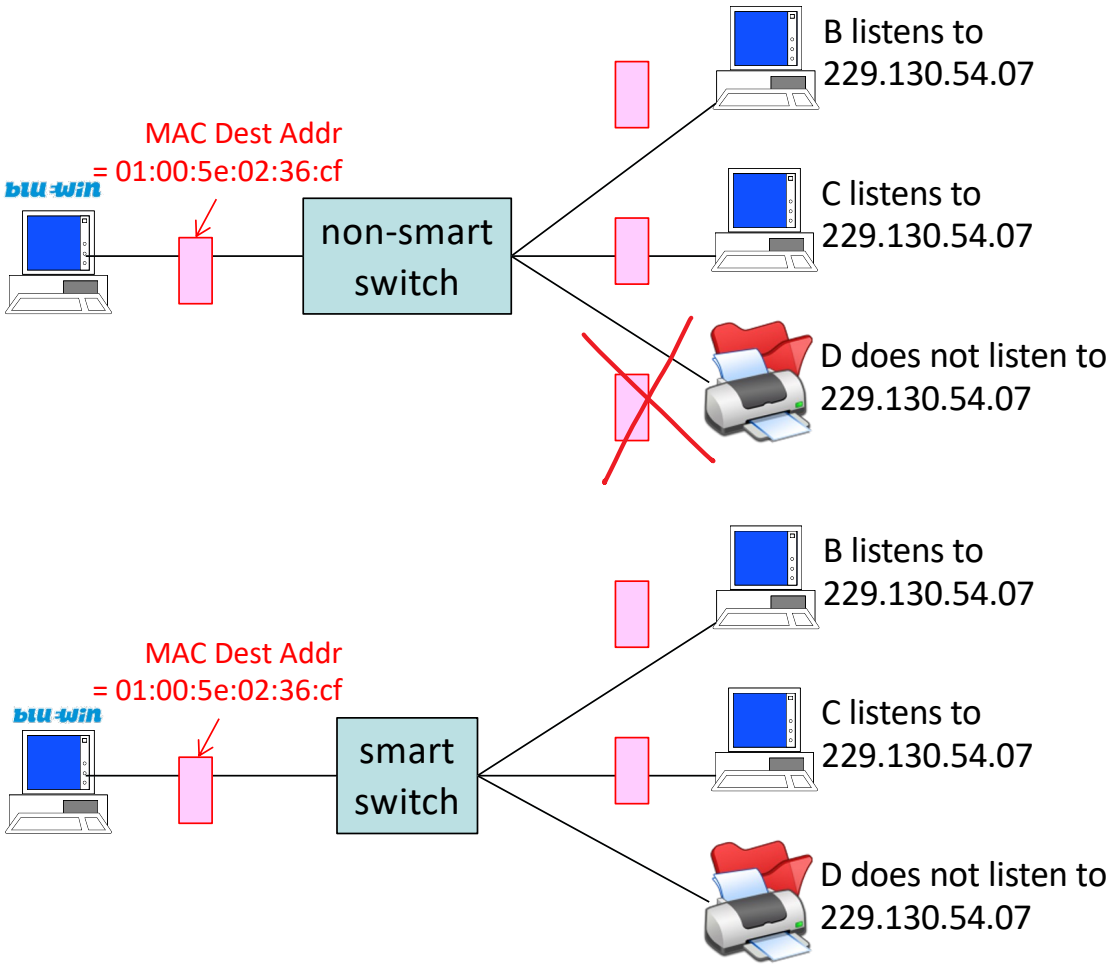
Several multicast addresses may correspond to same MAC address

if needed, operating system removes packets received unnecessarily; it is hoped that this rarely happens

All multicast is handled by MAC layer as ASM

(i.e. MAC multicast address depends only on IP multicast IP address m not on source address s, even if m is an SSM address)

MAC Multicast



Some (non smart) switches simply treat multicast frames as broadcast.

Some smarter switches simply listen to IGMP/MLD and overhear who listens – deliver only to intended recipients (IGMP or MLD snooping) – but do not distinguish SSM from ASM.

Security of IP Multicast

IP multicast with or without BIER makes life easier for attackers (e.g. Denial of Service, witty worm)

- mitigations: limit multicast rate and number of groups; control which multicast group is allowed (access lists)

SSM is safer as routers and destination can reject unwanted sources

IGMP/MLD is not secured and has the same problems as ARP/NDP

- mitigated by same mechanisms: sniffing switches observe all traffic and implement access-lists

Multicast-capable networks must deploy exhaustive filtering and monitoring tools to limit potential damage.

Multicast in Practice

Multicast is good for sources : one packet sent for n destinations -- multiplication is done repeatedly, $O(\log(n))$ times

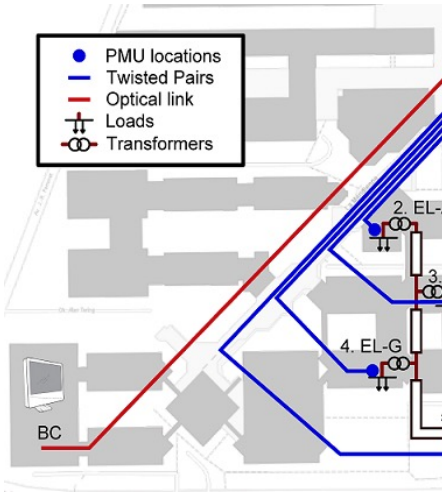
Traditional multicast suffers from per-flow state in routers;

BIER avoids that problem.

Multicast is not supported everywhere, but is

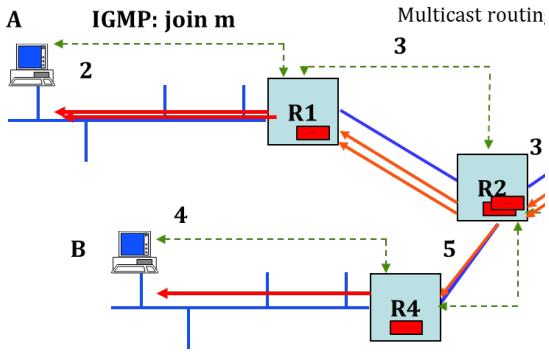
- At EPFL and other academic networks (PIM-SM)
- Data Center Virtualization Services (BIER)
- Internet TV distribution (PIM-SM / BIER)
- In some corporate networks for news, sensor streaming, time synchronization, large videoconferences etc…
- In industrial networks (smart grids, factory automation)

Works only with UDP, not with TCP



Say what is true

A. A
B. B
C. C
D. A and B
E. A and C
F. B and C
G. All
H. None
I. I don’t know



A. In order to send to a multicast group a system must first join the group with IGMP or MLD
B. In order to receive from a multicast group a system must first join the group with IGMP or MLD
C. A system can know whether a packet is multicast by analyzing the IP destination address.

R is an *ingress* edge router used for multicast distribution. In which case must R keep per-flow information ?

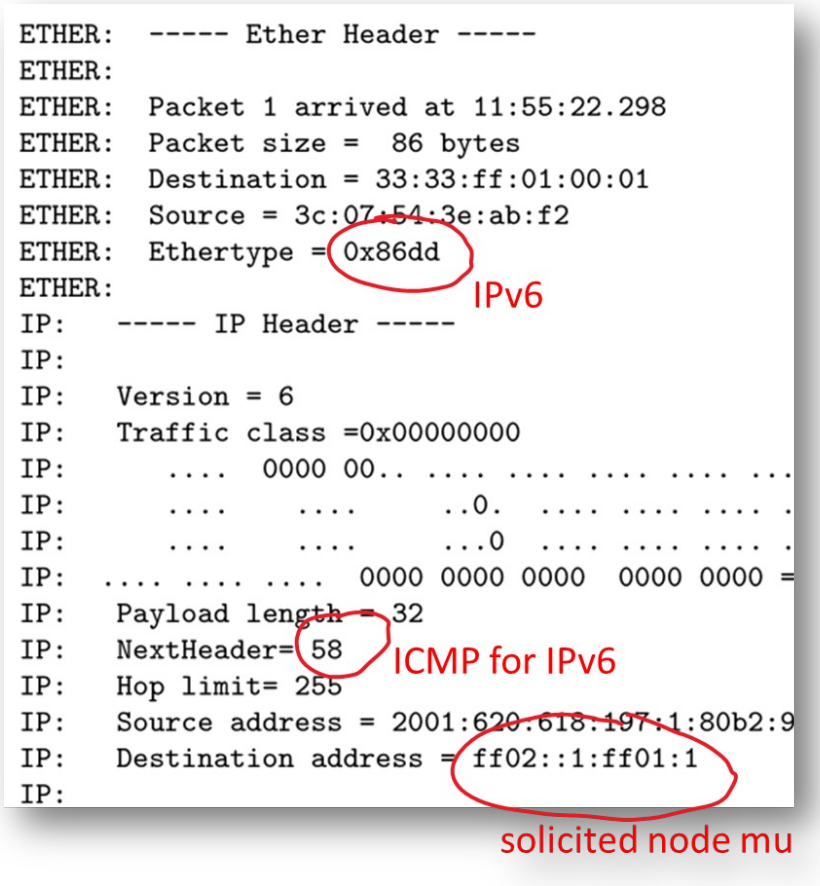
A. If R uses PIM-SM as multicast routing protocol
B. If R uses BIER as multicast routing protocol
C. In both cases
D. In neither case
E. I don’t know

R is a backbone router used for multicast distribution. In which case must R keep per-flow information ?

A. If R uses PIM-SM as multicast routing protocol
B. If R uses BIER as multicast routing protocol
C. In both cases
D. In neither case
E. I don’t know

The destination MAC address is…

A. A group address derived from the last 23 bits of the IPv6 target address
B. A group address derived from the last 24 bits of the IPv6 target address
C. A group address derived from the last 32 bits of the IPv6 target address
D. A broadcast address
E. The MAC address of an ARP server
F. I don't know



Conclusion

IP multicast came as an after-thought and uses a different principle than IP unicast (exact match versus longest prefix match) – deployed only in specific networks

IP multicast addresses cannot be aggregated

Multicast routing requires per-flow state and is non scalable. Tends to be replaced by BIER. BIER uses a different forwarding principle, based on bistrings (which represent sets of destinations).