Networking Technologies and Applications

Rolland Vida
BME TMIT

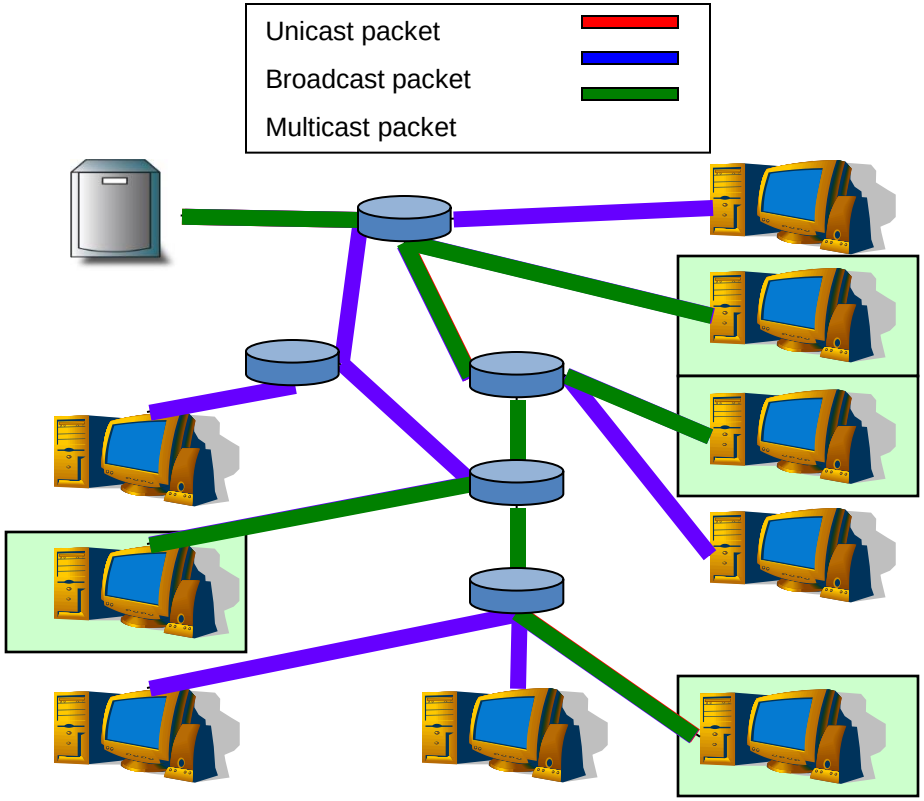
Group communication

- Goal: instead of a single destination node, communicate with a group of nodes
  - „natural” extension of the point-to-point communication (unicast)

- **Multicast**

What is multicast?

- **Unicast**
  - Point – to - point
  - Destination address: the address of one specific receiver
- **Broadcast**
  - Point – to – everyone
  - Destination address: address of the (sub)network
- **Multicast**
  - (Multi)point – to – multipoint
  - Destination address: group address



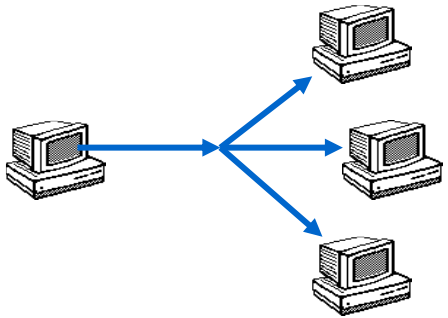
Group communication

- Packets have to be sent to all members of a group, not just a single destination
  - Group membership can be dynamic

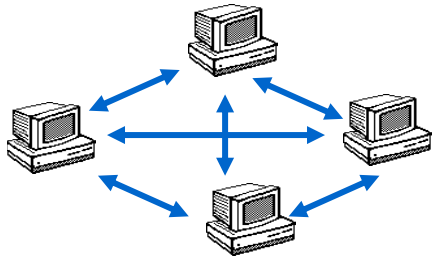
- Basic principle: once a group is created…
  - Interested receivers join the group
  - The network maintains the group and handles data delivery

Multicast applications

- Many applications are not point-to-point
  - Point-multipoint
    - Remote learning
    - Cache update
    - Video on demand



  - Multipoint-to-multipoint
    - Videoconferences, Audio conferences, Chat,
    - Distributed networking games
    - Cooperative applications



Requirements

- No one size fits all solution

- Requirements are different
  - Depending on the application needs
  - Depending on group size
  - Depending on network services / support
  - Depending on member heterogeneity

Participation rules

- Membership control
  - Open group: anybody can join
  - Closed group: limited membership

- Source control
  - Anybody can send a packet to the group
  - Only a group member can be a source
  - Just a selected source can send data

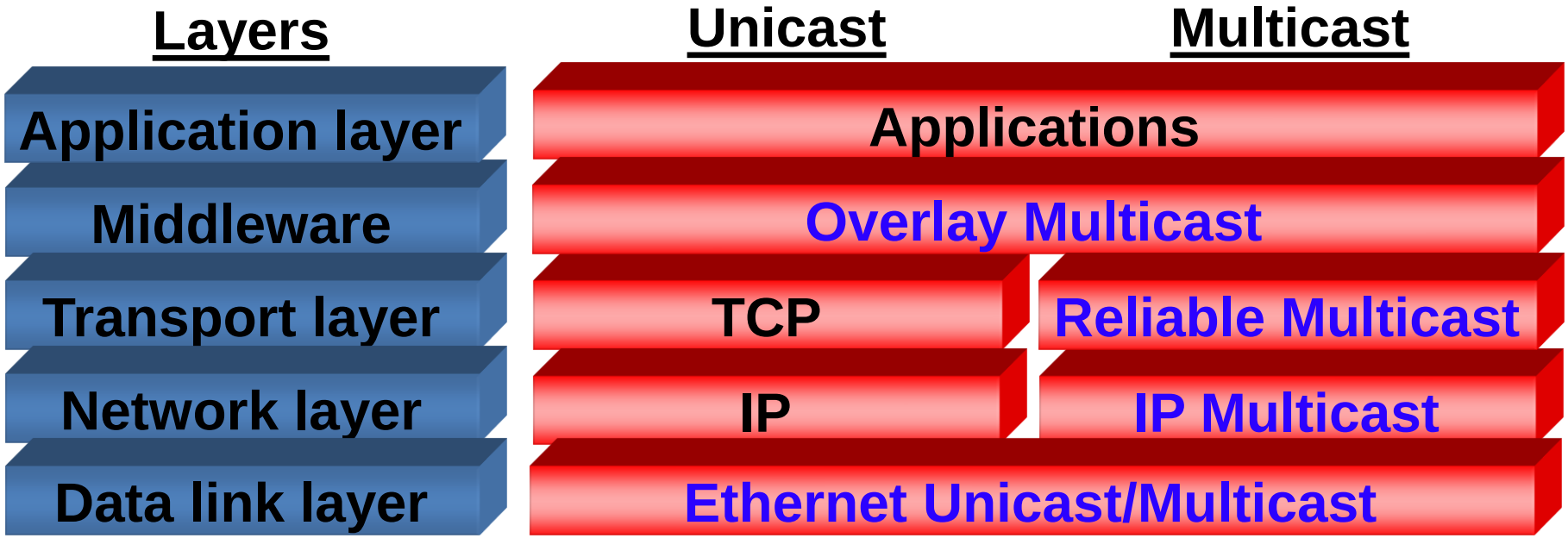
Reliability requirements

- Point-to-point communication
  - Reliable or best-effort (no guarantees)
  - The destination checks the packet: OK, or not
- Point-to-multipoint communication
  - Each receiver perceives the service differently
- Different reliability levels
  - 0-reliability: no receiver is guaranteed reliable transmission
  - 1-reliability: at least 1 receiver will reliably receive the packets
  - k-reliability: at least k receivers will reliably receive the packets
  - Total reliabiliy: all receivers will reliably receive the packets

Multicast at different layers

- The multicast service can be implemented in different layers
  - Data link layer
    - E.g. Ethernet multicast
  - Network layer
    - E.g. IP multicast, Xcast
  - Application layer
    - E.g. Narada, TBCP
- Which solution is the best?
  - It depends, no general solution

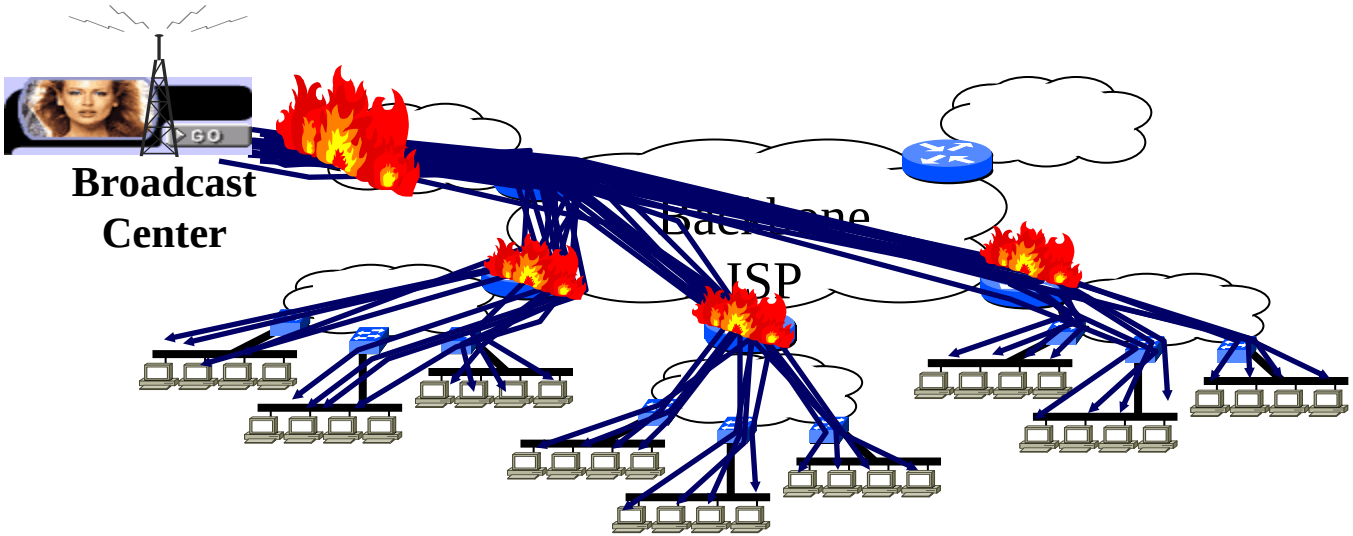
Multicast at different layers



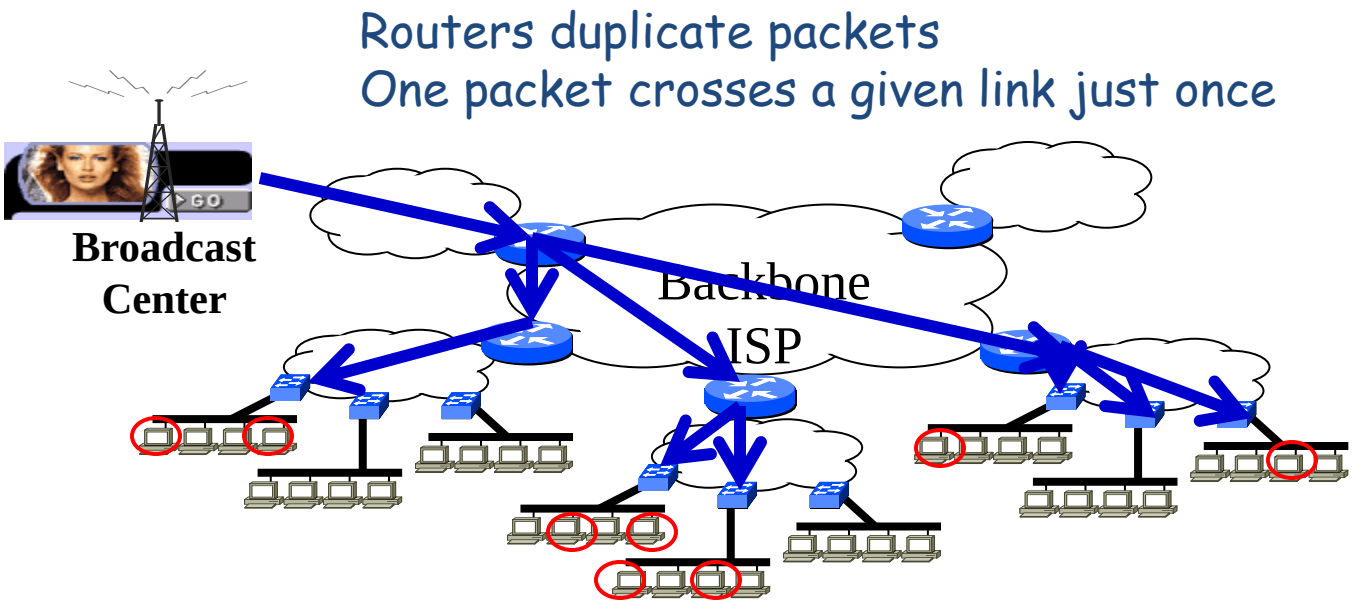
Network layer multicast

- The goal is the optimisation of networ layer resources
  - One packet crosses just once a given link
- Routers build and maintain a multicast tree
  - Traffic forwarding along the tree
  - Routers duplicate packets where needed
    - Branching points on the tree

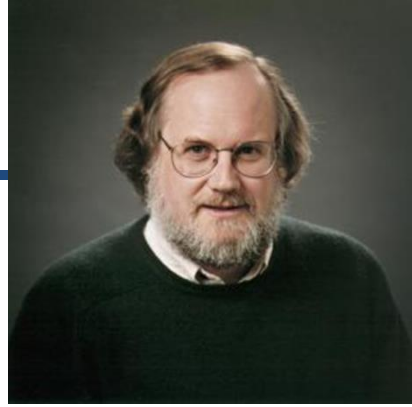
Group level unicast is not scalable



Let’s build trees instead



IP Multicast



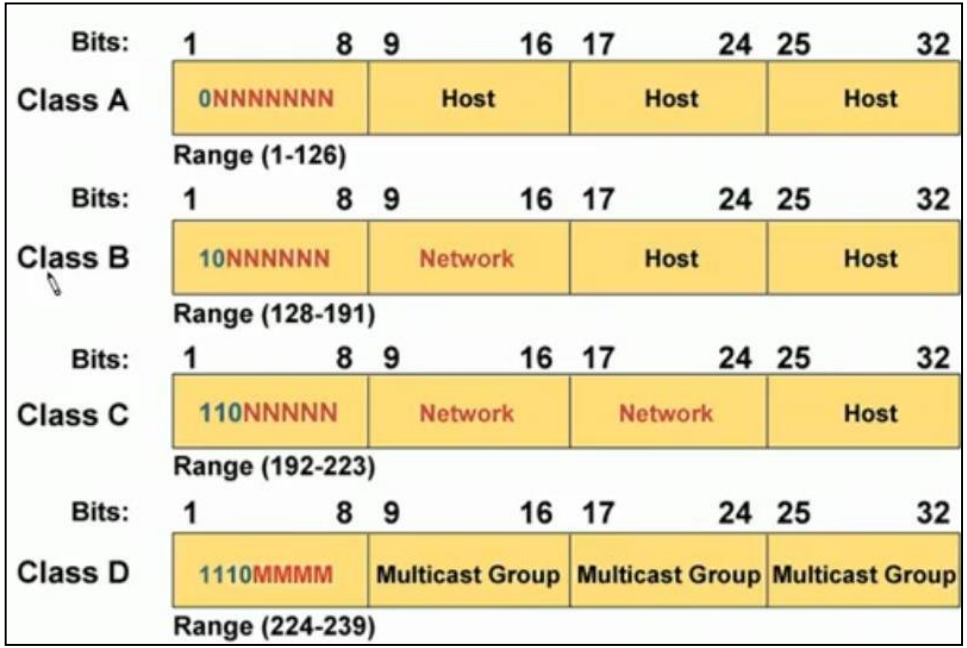
- Steve Deering PhD dissertation (1990)
  - **Any Source Multicast (ASM)**

- Open group communication model
  - Anybody can join the group, no access control
  - One user can be member of several groups in the same time
  - Anybody can send to the group, even non members
  - Group membership is dynamic
  - Nobody knows the size of the group, or its members

IP Multicast

- S. Deering, "Host Extensions for IP Multicasting", RFC 1112, 1989.

- The source sends its packets to a group address
- Anybody who joined the group is „reachable” through this address
  - Receives packets that are sent to this destination address
- A multicast group is identified by a **class D** IP address
  - **224.0.0.0 – 239.255.255.255**
  - 1110 + 28 bit group ID



IP Multicast

- Joining a multicast tree done in two steps
  - On the local area network (LAN)
    - A user announces its local multicast routers about the groups he would like to join
    - IGMP (IPv4), MLD (IPv6)
  - Over the large Internet (WAN)
    - The local router cooperates with the other multicast routers of the network to build the tree and forward the packets along that tree
    - DVMRP, MOSPF, CBT, PIM-DM, PIM-SM, PIM-SSM

IGMP

- **Internet Group Management Protocol**
- An IPv4 protocol, running between the final users and the local multicast routers on the local network
  - Handles multicast group membership
  - Asymmetric protocol
    - User side
    - Router side
- The router learns which groups the end-users on his local network listen to
  - Not interested in how many receivers, important thing is to have at least on receiver
  - Not interested in exactly who are the receivers

IGMPv1

- S. Deering, "Host Extensions for IP Multicasting", RFC 1112, 1989.
- A multicast router sends regular **Query** messages to the multicast address of all the users (224.0.0.1)
- A user answers with a **Report** message, in which specifies the groups he listens to
  - The Report is sent to the multicast addresses of those groups
- To decrease the number of Report messages:
  - Using timers
    - A user does not answer immediately to the Query
  - Host Suppression
    - If someone else answers faster, it deletes its own Report message
- **Unsolicited Report**
  - If a user wants to listen immediately to a new group

IGMPv1 Router

- An IGMPv1 router maintains a multicast membership table
  - Which multicast groups have members on its network
  - When was the last Report message received about those groups
- **Soft-state** protocol
  - If in a given time nobody refreshes its interest in a given groups, the group will be deleted from the multicast membership table
- It forwards to the local network all packets that are sent to a multicast destination address that is contained in its membership table

IGMPv2

- W. Fenner, "Internet Group Management Protocol, Version 2", RFC 2236, November 1997.
  http://www.ietf.org/rfc/rfc2236.txt

- IPv6 version: MLD (Multicast Listener Discovery)
  - S. Deering, W. Fenner, B. Haberman, "Multicast Listener Discovery (MLD) for IPv6", RFC 2710, November 1999.
    http://www.ietf.org/rfc/rfc2710.txt

- Introduces a **Fast Leave** mechanism
  - Do not have to wait until a timer expires to cut off a group

IGMPv2 messages

- Membership Query
  - General Query
  - Group Specific Query
- Membership Report
- **Leave Group Message**

- If a host wants to leave a group, it sends a Leave message to the multicast address of all the multicast routers (224.0.0.2)
- Before cutting off the group, the router has to ask if anybody else is still interested in that group or not
  - Group Specific Query
  - If no answer in a given limited time, the router cuts off the group from its table

- IGMPv3 – later...

Multicast Routing

- A source sends its packets to the group's multicast address
- The multicast routers in the network build and maintain a multicast tree
  - Packets are forwarded along that tree
- The local multicast router, based on its IGMP membership table, joins or leaves this tree
- A multicast routing protocol runs among the routers of the network
  - MOSPF, DVMRP, CBT, PIM

MOSPF

- **Multicast Open Shortest Path First**
  - J. Moy, „Multicast Extensions to OSPF”, RFC 1584, March 1994 http://www.ietf.org/rfc/rfc1584.txt
- Link State protocol
- Extends the OSPF unicast routing protocol
  - Multicast group membership information is also distributed among the routers
  - Each MOSPF router learns which multicast groups have listeners on which local network
  - Based on this information they build a shortest path tree for each source and each group
- Large signaling overhead
- Difficult to handle topology changes
  - All the trees have to be recalculated

DVMRP

- **Distance Vector Multicast Routing Protocol**
  - D. Waitzman, C. Partridge, S. Deering, "Distance Vector Multicast Routing Protocol", RFC 1075, November 1988

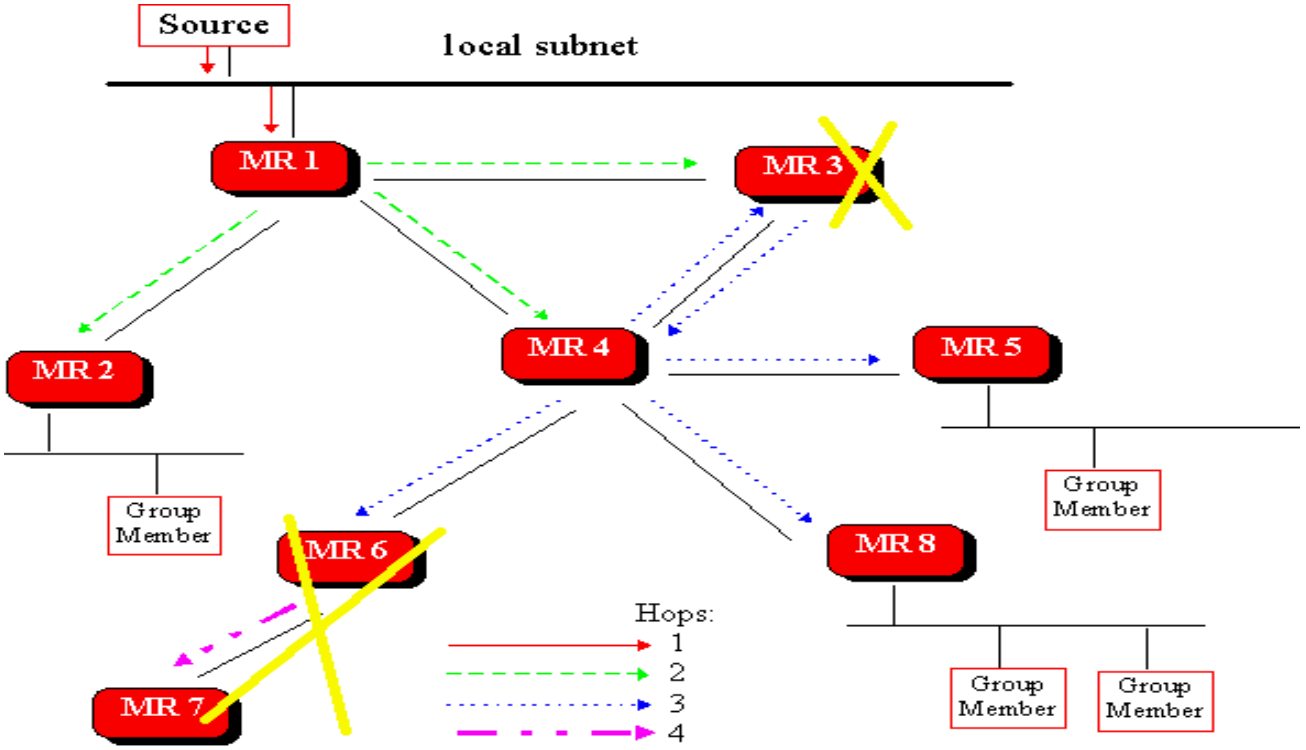
    http://www.ietf.org/rfc/rfc1075.txt

- Distance vector protocol
  - Uses the RIP unicast routing protocol

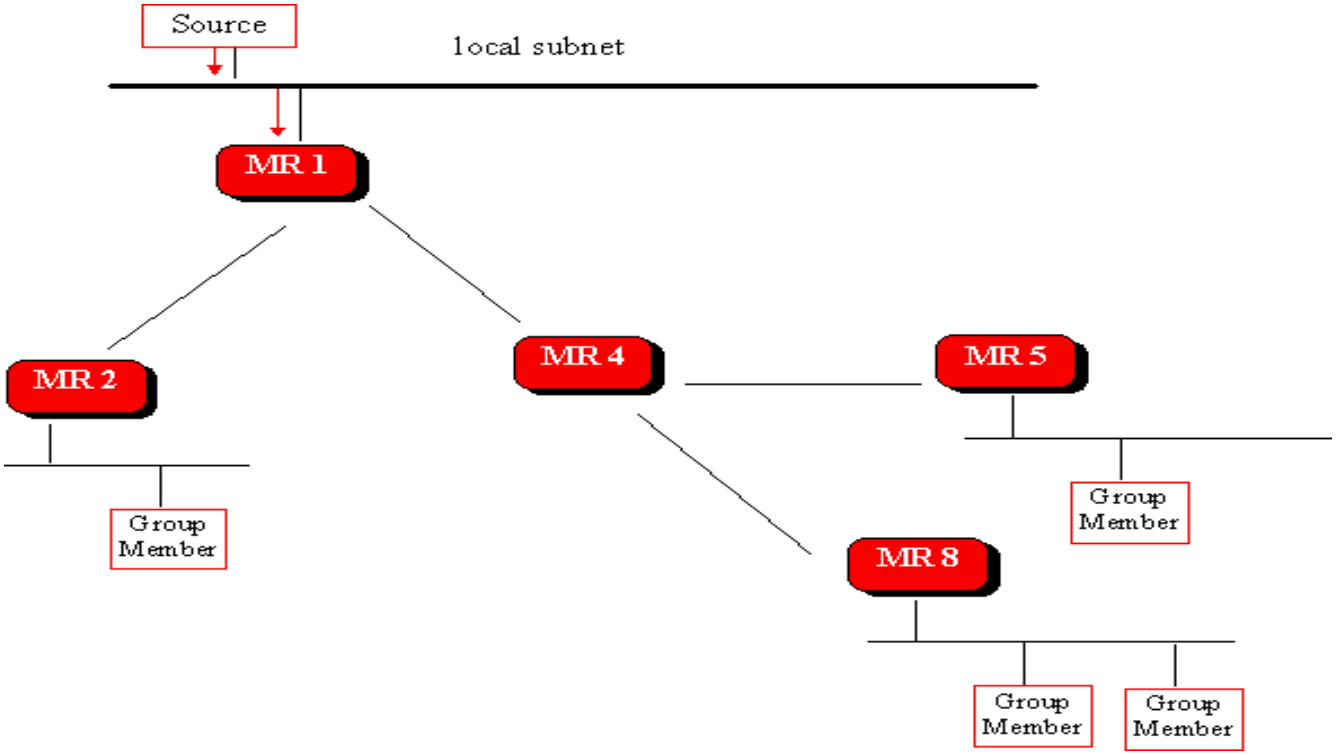
DVMRP

- Flood and prune
  - Flooding
    - Checks the incoming interfaces of the packets
    - If not over the shortest path towards the source, the packets are dropped
    - If yes, packets are flooded over all the interfaces
  - Pruning
    - If no interested receiver on the local network
    - If packet not received over the shortest path
  - An internal router learns its interfaces over which it recieved a Prune message
    - The upcoming packets are not forwarded over those interfaces anymore
    - Prune messages become obsole after a while (one minute by default)

DVMRP flooding



DVMRP prune


Source
local subnet
MR 1
MR 2
MR 4
MR 5
MR 8
Group Member
Group Member
Group Member
Group Member

PIM

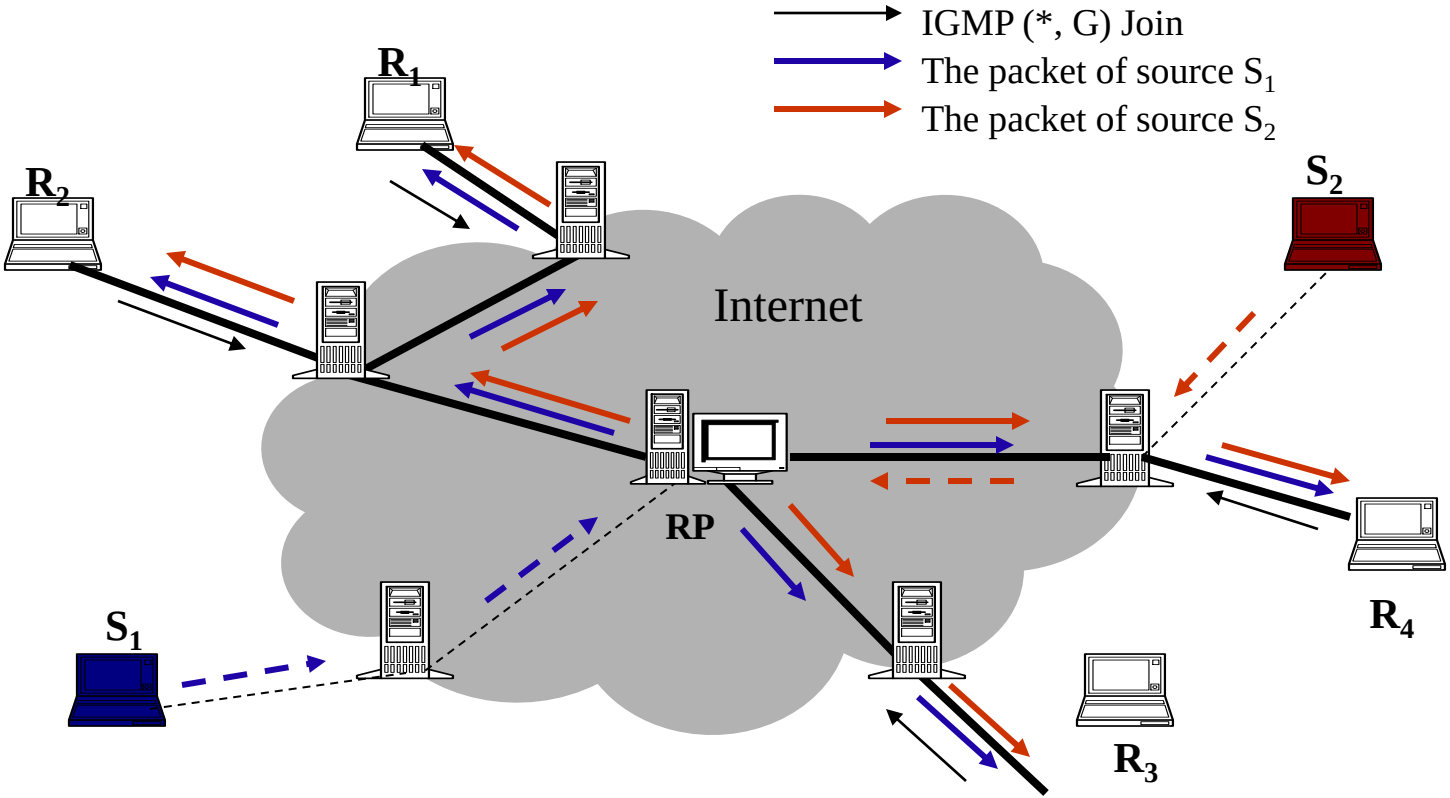
- Protocol Independent Multicast
  - PIM Dense Mode (PIM-DM)
  - PIM Sparse Mode (PIM-SM)

- PIM-SM
  - W. Fenner et al., „Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)” , RFC 4601, August 2006
  - The most used multicast routing protocol today

PIM-SM

- Builds a **shared multicast tree**
- Chooses a rendez-vous point (RP)
  - The RP is the root of the shared tree
    - „Explicit join”
  - Each source sends its message to the RP
    - The RP forwards the messages along the shared tree
  - Optimization to switch after a while from the shared tree to a source-specific tree

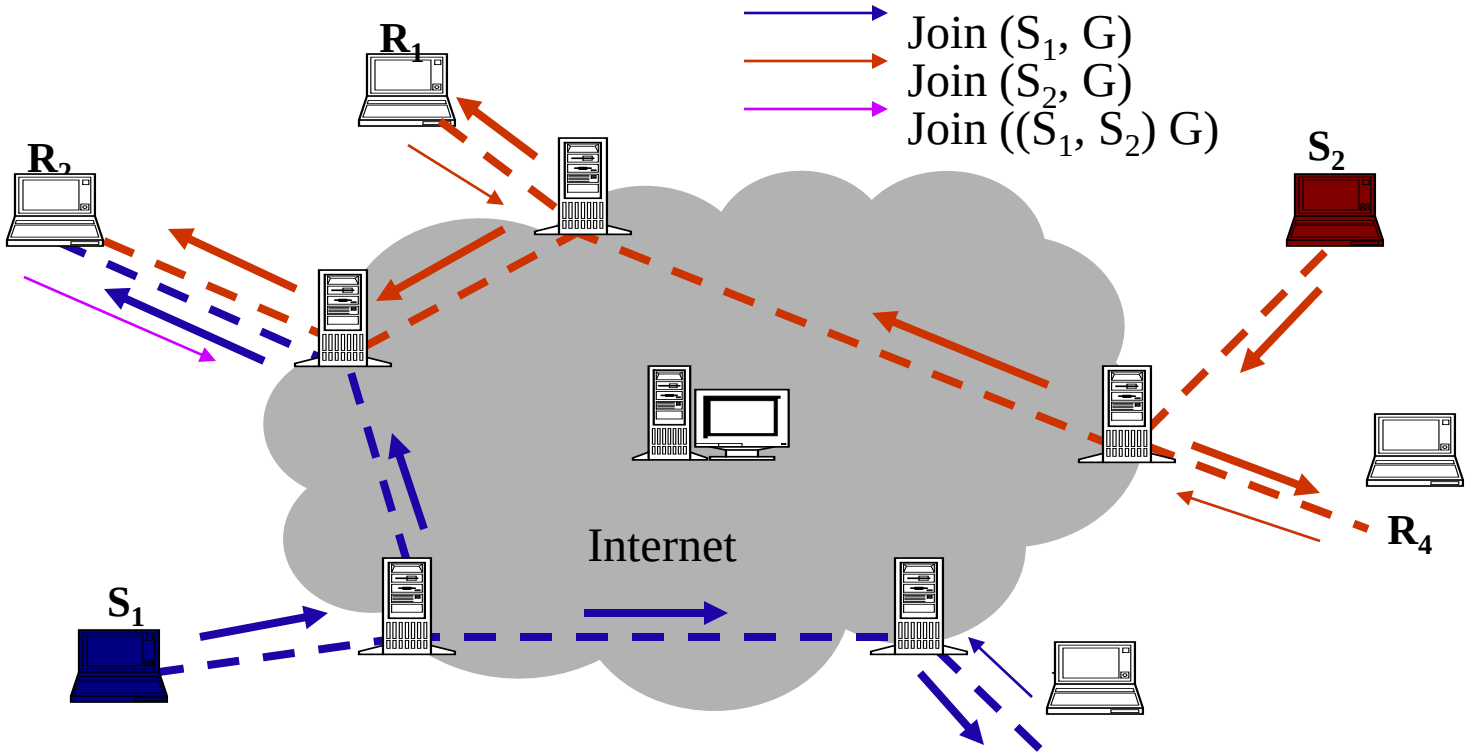
PIM-SM operation



The SSM model

- **SSM - Source Specific Multicast**
  - Based on the Express model
  - H. Holbrook, D. Cheriton, "IP Multicast Channels: Express Support for Large-Scale Single-Source Application", in *Proceedings of ACM SIGCOMM'99*, Cambridge, MA, USA, Sept. 1999.
- The (*,G) multicast group is replaced by the (S,G) multicast channel
  - S the unicast address of the source
  - G the multicast address of the group
  - Only source S can send packets to the receivers of channel (S,G)
  - Traffic is forwarded along a source-specific tree

SSM model



Source filtering

- The SSM model needs source filtering
  - The host specifies not only which group it wants to listen to, but also which source that sends to that group
- IPv4 – IGMPv3
  - B. Cain, et. Al, "Internet Group Management Protocol, Version 3", RFC 3376, October 2002.
    http://www.ietf.org/rfc/rfc3376.txt
- IPv6 – MLDv2
  - R. Vida, L. Costa, „Multicast Listener Discovery Version 2 (MLDv2) for IPv6", RFC 3810, June 2004.
    http://www.ietf.org/rfc/rfc3810.txt

Message types

- IGMP/MLD Query
  - General Query
    - Who listens what?
  - Group Specific Query
    - Does anybody listen this specific group?
  - Group and Source Specific Query
    - Does anyone listen to this specific source that sends to this specific group?
- IGMP/MLD Report
  - Current State Record
    - What do I listen to – e.g. Include (A) or Exclude (B)
      - A and B are source address sets
  - Filter Mode Change Record
    - Changing the filter mode (Include or Exclude)
  - Source List Change Record
    - Allow (A) or Block (B)

IP Multicast

- Considered for several years the „revolutionary technology of the future”
- Advantages
  - Efficient data transfer
    - Usually over the shortest path (DVMRP, MOSPF, PIM-SSM)
    - Taking into account the physical topology
  - Efficient use of resources
    - One packet is sent just once over a specific link
  - Scalable for handling the communication of large groups
    - The group is identified by a virtual group address
      - One routing table entry for a very large group
    - Nobody tracks who is part of the group, and how large is the group

IP Multicast

- Still not deployed at large scale
  - Technical and economic reasons
- Technical reasons
  - Complicated addressing
  - No scalable inter-domain multicast routing
  - Does not scale to a large number of groups
    - The router has to keep one entry per multicast group
    - Multicast addresses are hard to aggregate
  - Lack of support for higher layer services
    - IP multicast is a *best-effort (multi)point-to-multipoint* data transfer service
    - End users are responsible for handling higher layer services
    - Difficult congestion control and reliablility handling

IP Multicast

- Economic reasons
  - Slow and difficult deployment in the network
    - Even though all the routers „speak” today the most important multicast protocols, the ISPs sometimes do not activate them on their networks
    - Really efficient only if used in the entire network
    - Otherwise tunneling is needed
  - „Chicken-egg” problem
    - ISPs do not support it, not enough multicast applications, no need for it
    - Software companies do not develop multicast applications, as there is no network support for them
  - No proper business model behind it
    - Hard to control the usage of network resources for the ISP
    - Hard to control who uses the service, for the content provider
    - Hard to provide a proper charging scheme for it