Multicasting

(Z:\Polys\Multicast\Multicasting-intro.fm- 25 November 2014 12:01)

Outline

- Introduction
- Multicast Adressing
- Multicast Trees
- Multicast Tree Computation
- Conclusion

Bibliography

- L. Giuliano. Interdomain Multicast Routing. Addison Wesley, 2002
- R. Wittmann, M. Zitterbart. Multicast Communication : Protocols and Applications. Morgan Kaufmann publishers, 2001
- C. Huitema. Routing in the Internet. Prentice Hall, 1999
- S. Paul. Multicasting on the Internet. Kluwer academic publishers, 1998
- W. Stallings. High Speed Networks. Prentice Hall, 1998
- Stephen E. Deering, Multicast Routing in a Datagram Internetwork, Ph.D. Thesis, Stanford University, 1991.

1. Introduction to Multicasting

1.1. Presentation

Many applications could send the same data packets to several destinations:

- Videoconference, audioconf., web phone, radio, etc.
    - . e.g. : CU-SeeMe, IVS, Netmeeting
- Collaborative tools: whiteboard, shared editor, etc.
- Data broadcasting: updating, etc.

With or without multicast overlay:

- Application layer:
    - . packet duplication is done by hosts: scalability
    - . requires usual unicast forwarding service
    - . link stress could be high,
    - . routing path is suboptimal
        - - ex: CU-SeeMe reflectors, "Peer to Peer" broadcasting.
- Network layer:
    - . packet duplication is done by routers
    - . requires multicast forwarding service
    - . the network knows the best path!
    - . some access networks support broadcast

Some broadcasting methods:

Multi- unicasting

Total broadcasting

host

group member

router

Broadcasting server(s)
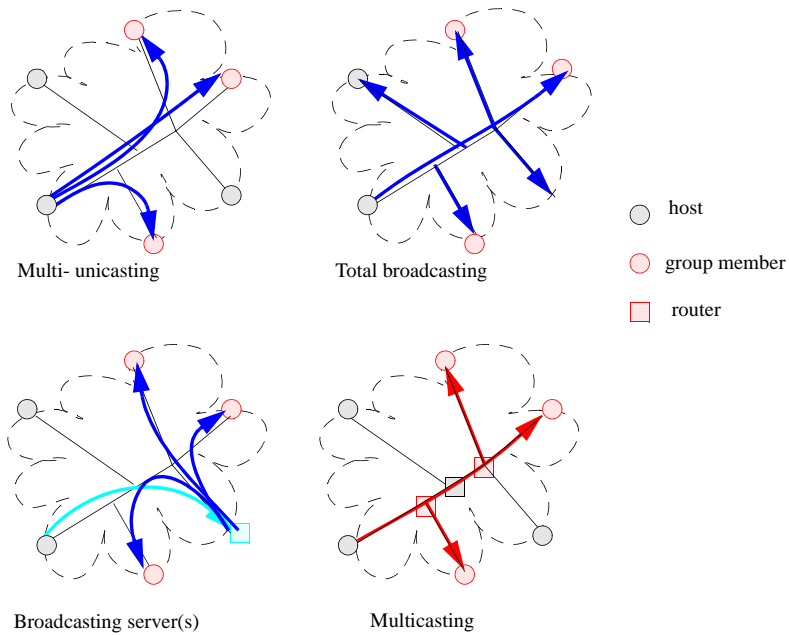
Multicasting

Multicasting in IP: every member of a multicast group receives a copy of the multicast packets sent to the group by any IP source.

1.2. Different Types of Multicast Protocols

Multicast routing protocols

- Network layer protocol
- update the routing tables
- (the routing table) assures the forwarding of the multicast packets
  - e.g.: (IGMP), DVMRP, MOSPF, PIM, MSDP (Multicast Source Distribution Protocol (inter-domain)), etc.

Multicast Transport protocols

- Transport layer protocol
- assured data delivery
  - congestion control and error recovery
- e.g.: UDP!, ST2 (Stream Transport v2), XTP (eXpress Transport Protocol), MTP (Multicast Transport Protocol), RMP (Reliable Multicast Protocol), SRP (Scalable Reliable Protocol), RMTP (Reliable Multicast Transport Protocol), RAMP (Reliable Adaptative Multicast Protocol), etc.

Some other protocols:

- Group management protocols:
  - e.g.: SAP (Session Announcement Protocol), SDP (Session Description Protocol), SIP (Session Initiation Protocol) uses HTML, etc.
- "QoS" protocols
  - e.g.: RSVP (ReSerVation Protocol), RTP & RTCP (Real Time &Control Protocol), etc
- Application protocols
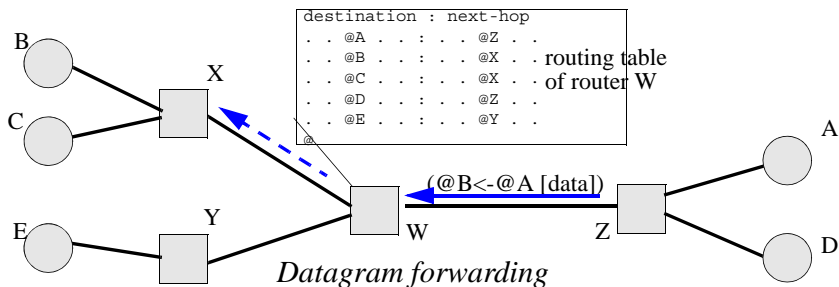 - e.g.: MFTP (Multicast File Transfer Protocol) : MCP + MDP (Control + Data)

Some tools:

- SDR (Session Directory) uses SAP+SDP,
- Mrouted (multicast routing daemon),
- Mrinfo: display the multicast router configuration,
- Mtrace: display the path toward a multicast destination

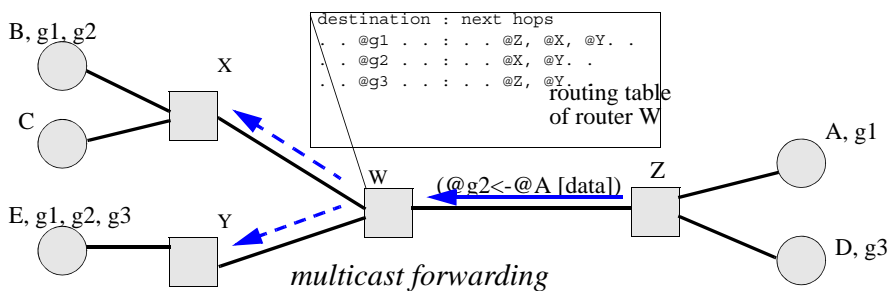=> multicast routing protocols

1.3. Unicast Datagram Forwarding

- IP Forwarding (by routers)

  - based on routing table

    . destination address extraction

    . table look up: destination address => next-hop

```
destination : next-hop
. . @A . . : . . @Z . .
. . @B . . : . . @X . · routing table
. . @C . . : . . @X . · of router W
. . @D . . : . . @Z . .
. . @E . . : . . @Y . .
. . @
```

Datagram forwarding

(@B<-@A [data])

- Routing protocols

    . routing table update

1.4. Multicast Routing

- Multicast forwarding

  - forwarding of the multicast IP datagrams.

    . selection of the next routers (next-hops)

```
destination : next hops
. @g1 . . : . . @Z, @X, @Y. .
. @g2 . . : . . @X, @Y. .
. . @g3 . . : . . @Z, @Y.
```
routing table
of router W

multicast forwarding

(@g2<-@A [data])

- Multicast routing protocols

    . (multicast) routing table update

1.5. Scalability

Routers process thousands of packets every second

- optimization of the forwarding process
- its complexity depends of the size of the routing table
  - usual look up algorithm: *n log(n)*
  - number of network hosts ==> billions!

Solutions:

- Aggregation
- Routing domain structuring

1.6. Aggregation

- IP subnets
  - all the subnet hosts have the same address prefix
   . for instance: 194.10.1.0/24

reste d'Internet

LAN

IP subnet 194.10.1.0/24

194.10.1.7

194.10.1.8   194.10.1.9

194.10.1.6
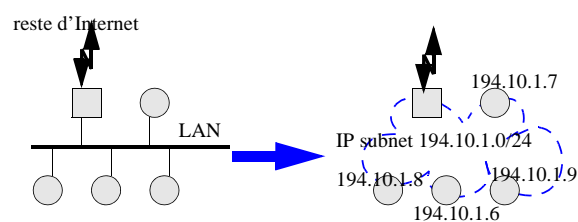
- Best prefix match
  - one table entry for all destinations with binary-successive addresses and same next hop
    . for instance: European prefix 194.0.0.0/7 = [194.0.0.0, 195.255.255.255]
    . (only) 104.000 entries in the routing table of backbone routers (2001)
  - exhaustive table look up to find out the entry with the best prefix match

   => unsuitable for multicast forwarding: no locality!

1.7. Structuring of the Unicast Routing Domain

Two levels of routing domain:

- Interior Gateway Protocols exchange routing data between routers inside a routing domain.
  - routers of a routing domain know a route toward any host in the routing domain
  - the route is made of next-hop routers
  - for instance: RIP, OSPF
- Exterior Gateway Protocols exchange routing data between routers belonging to different routing domains.
  - EGP routers know a route to reach any host with any prefix of any routing domain
  - the route is made of domain border routers
  - for instance: BGP, EGP
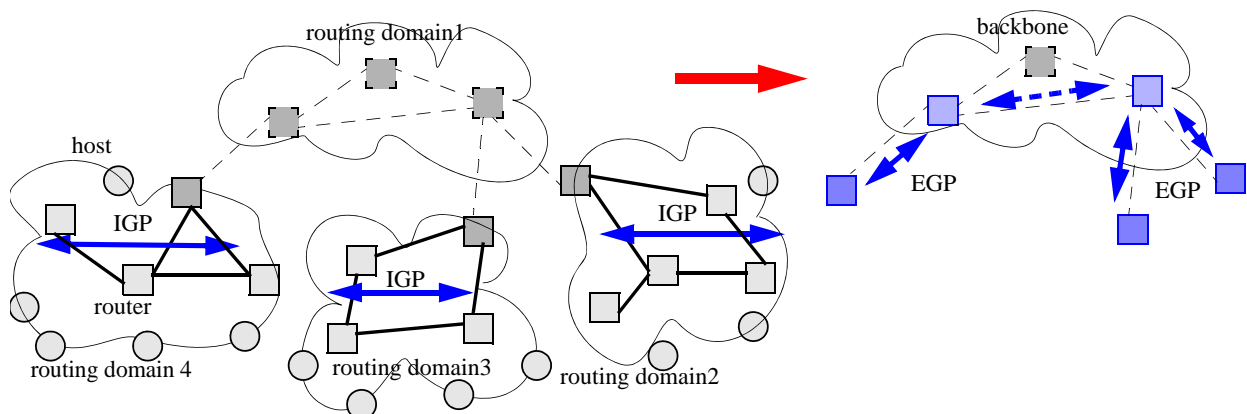
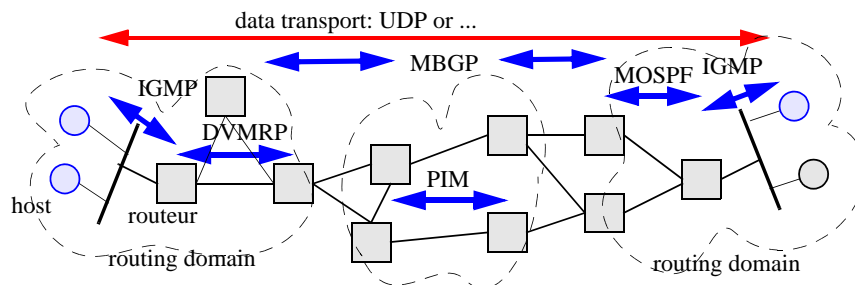=> this domain structuring enhances the network scalability.

Nota: some protocols introduces a extra level
- for instance: OSPF routing areas

1.8. Structuring of Multicast Routing Domains

Three levels of multicast routing domain:
- Between hosts and routers: IGMP (group membership protocol).
- Between routers of the same domain. For instance: DVMRP, MOSPF, PIM, etc.
- Between routers belonging to different domains. MBGP, MSDP.

=> this domain structuring enhances the network scalability.

---

2. Multicast Addressing

2.1. Introduction

A IP multicast address identifies a group (of hosts)
- Class D IP address, with prefix "1110_2",
- In the 224.0.0.0 - 239.255.255.255 interval

Reception:
- Any host with the multicast address receives a copy of every packet sent to this address.

Transmission:
- Any host can send a packet toward any multicast group.

Beware:
- Multicast address can not be used as source address.
- No access control to a multicast group: out of IP scope.
- If several applications share the same multicast address, their entities will receive a copy of any packet sent to the multicast group.

2.2. Multicast Addresses

2.2.1 Broadcast addresses

- Full of "1"!
  - Local broadcast : 255.255.255.255
    - . all hosts in the local IP subnet (local to the source)
  - Distant broadcast: netid-A.255.255.255, netid-B.255.255, netid-C.255
    - . all hosts in the distant IP subnet

Local broadcast   Distant broadcast
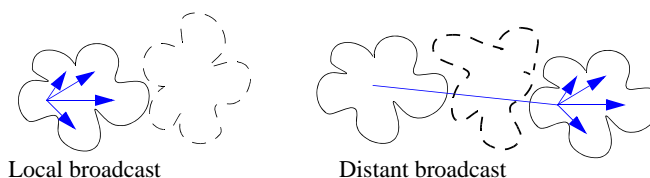
2.2.2 Well known multicast addresses

Identify a group of hosts (routers) for a certain application

- Optimization:
  - no inopportune processor interruption

Some well known multicast addresses:
- 224.0.0.1: all multicast nodes (hosts and routers)
- 224.0.0.2: all routers
- 224.0.0.9: all RIP-2 routers
- etc.

2.3. Multicast IPv6 Address

- IPv6 prefix of the multicast addresses: FF_{16}

IPv6 address : 16 bytes

8 bits	4 bits	4 bits	112 bits
0xFF	flag	scope	group identifier
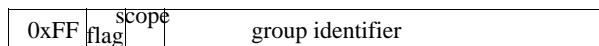
- *Flag* field [4 bits] :
  - T=0: permanent multicast address
  - T=1: temporary multicast address
- *Scope* field [4 bits]
  - 1: node
  - 2: link
  - 5: site
  - 8: organisation
  - E: world

  . 0, F: reserved; 3, 4, 6, 7, 9, A, B, C, D:unused

  . similar to the TTL field used by IPv4.
- *Group identifier* field [14 bytes]

- Some well known multicast addresses:

  . No host belongs to this group:
  - FF0s::0
  . All nodes (stations + routeurs):
  - format: FF0s::1 with s={1, 2, 5, 8, E}
  - e.g.: all nodes of the local organisation = FF08::1

  . All hosts:
  - format: FF0s::2 with s={1, 2, 5, 8, E}
  - e.g.: all link interfaces of the local host = FF01::2

  . All routers:
  - format: FF0s::3 with s={1, 2, 5, 8, E}
  - e.g.: all site routers = FF05::3

  . All NTP servers:
  - format: FF0s::43 with s={1, 2, 5, 8, E}

Goal

  . Optimisation of the discovery process (less processor interruptions)
  . Minimization of the multicast data load

3. Multicast Trees

The multicasting uses broadcast trees (or multicast trees):

- each multicast group is associated with a broadcast tree
- the source is the root of the tree, it sends multicast packets
- the leafs of the tree are group members, they receive a copy of the multicast packets (some group members can be tree nodes)
- branching nodes of the trees are routers

The criteria:

- The height of the tree (the delay)

$$\text{Max}_{p \in \text{Path(Tree(S, G))}}\left(\sum_{l \in p} \text{Cost(l)}\right)$$

- The total volume of the tree (total cost)

$$\sum_{l \in \text{Tree(S, G)}} \text{Cost(l)}$$

3.1. Optimum Cost Tree

Optimum tree which spans all the network nodes:

=> Optimum spanning tree

- Optimum tree which spans the group members

=> Steiner tree

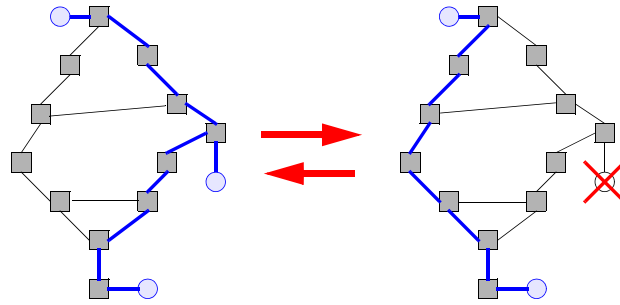Global optimization of the network resources.

The complexity of the Steiner tree computation:

- "NP-complete" problem!

Example : unitary link cost, total cost of the tree = 14, tree height = 7
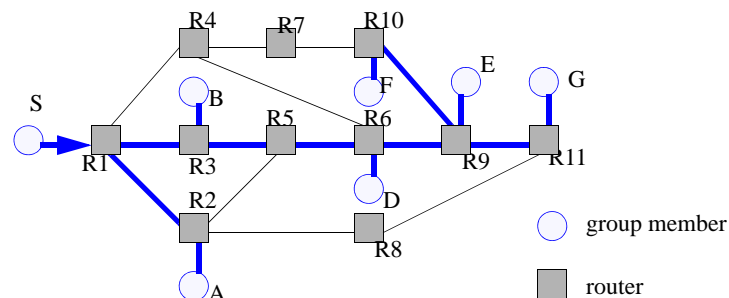
Instability of the Steiner tree:
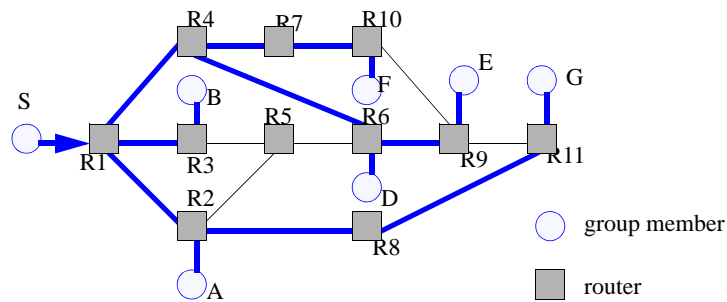

3.2. Shortest Path Tree (SPT)

The multicast packets follows the shortest path between the source and each group member individually.

Minimize the delay (required by some applications)

Complexity of the SPT computation with Shortest Path First algorithm:

- $\mathcal{O}(|V|^2+|E|)$, and if si $|V|^2>>|E|$, $\mathcal{O}((|E|+|V|) \log(|V|))$

Example : unitary link cost, total cost of the tree= 16, tree height= 5

=> SPT is used by most of the internet multicast routing protocols

## 3.3. Shared Trees

The same tree can be shared:

- • one tree for all multicast groups
  - total spanning tree
  - drawbacks: non optimal path, multicast packets delivered to non member hosts
- • a tree for each group and source pair
  - source based multicast trees: <S,G>
  - shortest path but a large number of trees
- • a tree for each group
  - shared multicast tree: <*,G>
  - efficient (bidirectional) tree but non-optimal path
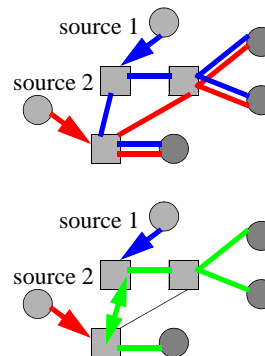
3.4. Local Routing of Multicast Datagrams

When a router receives a multicast packet, it forwards the packet on every interface where a member is active (except on the interface on which the datagram has been received).

A router knows that a group is active on a interface thanks to IGMP.

Lists of group memberships   *Multicast packet reception*   *Multicast packet emission*
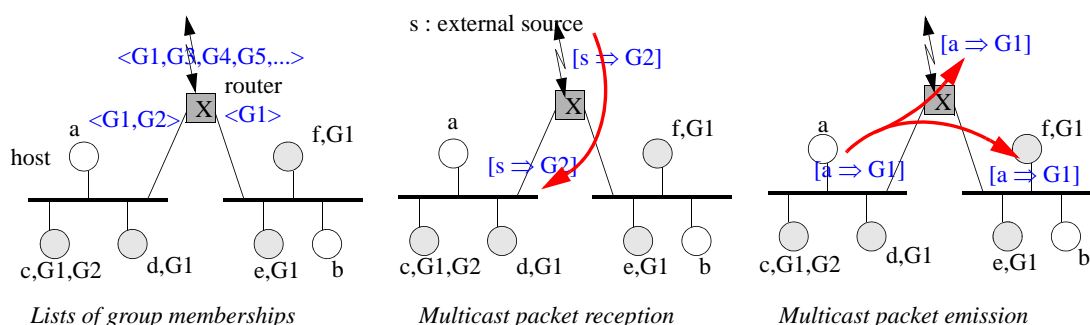
# 4. Multicast Tree Computation

Multicast tree computation is complex.

Some solutions:

- based on the network topology
  - produced naturally by the first phase of link state protocols: routing DB
  - e.g. : OSPF => MOSPF
- based on the unicast routing data
  - in the routing table
  - two main techniques:
    - . Reverse Path Forwarding
    - . Core Base Tree

=> RPM (e.g. DVMRP, PIM-DM)

4.1. Flooding

- flooding broadcast
  - one packet is forwarded to every router interface (except the incoming interface)
  - simple, reliable, optimal (!)
  - independent of the destinations
  - large number of packet copies, and network cycle tend to generate a countless number of copies!

interfaces

multicast source   multicast router

source   router

forwarded packet
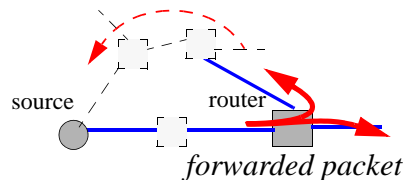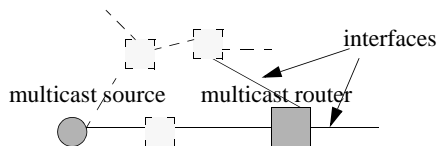
4.2. Reverse Path Forwarding

Suppression of the potential cycles generated by the flooding:

- a packet received on the upstream interface of a router (i.e. the interface which should be used to send a packet toward the source of the multicast packet) is forwarded on all it downstream interfaces.

- a packet received on a downstream interface of a router is discarded.

For each multicast source one and only one interface is designated as the upstream interface.

- the interface designed (by the unicast routing protocol) as the interface toward the multicast source.
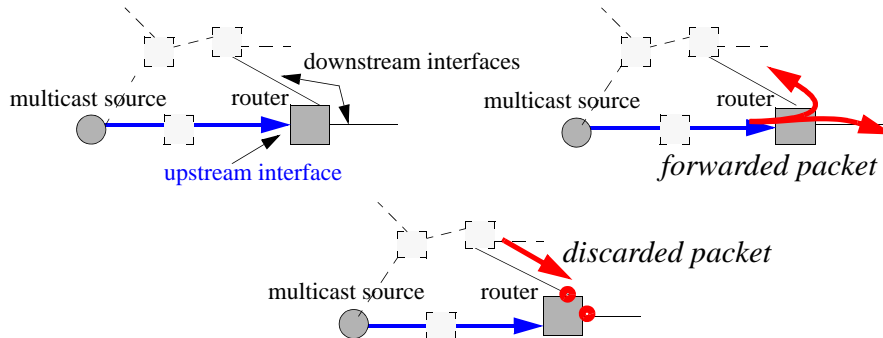
4.3. Reverse Path Broadcasting Tree

If every router applies the RPF algorithm, it produces a total spanning tree:

- the RPB tree
- it doesn't require any additional routing state
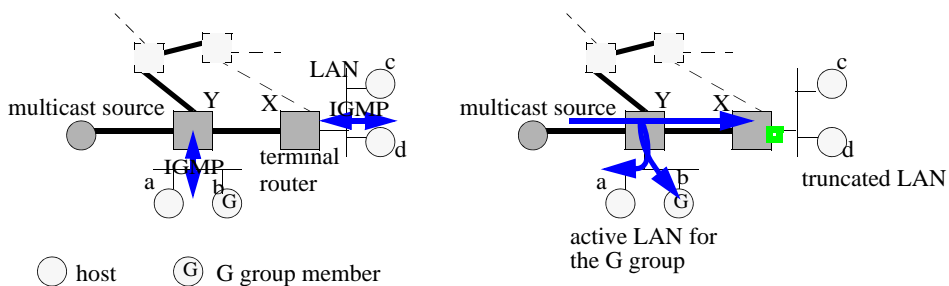- all routers and hosts receive (at least) a copy of every multicast packet sent

RPB tree for source S
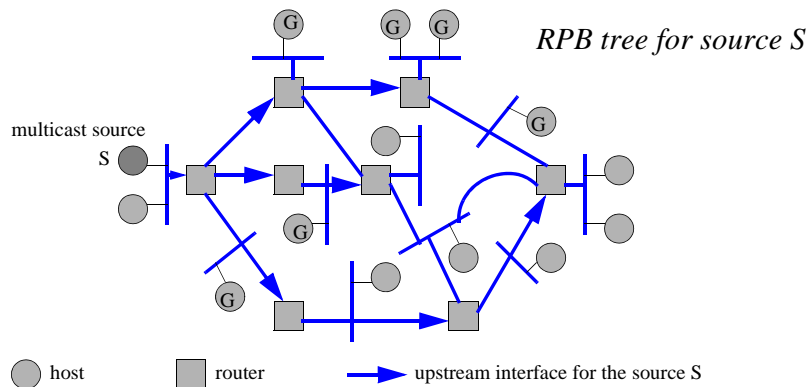
- The spanning tree is dependant of the source
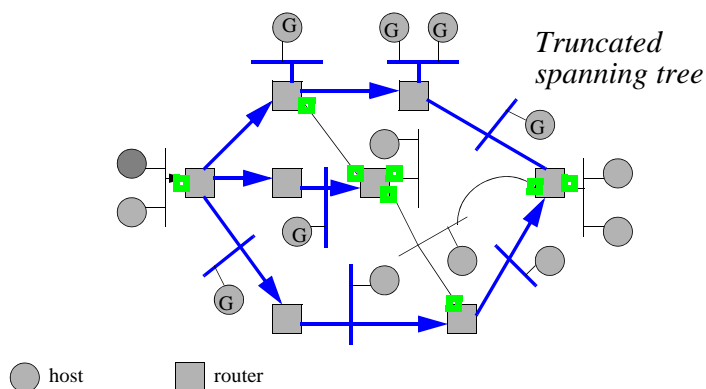
4.4. Truncation of the Spanning Tree

Truncation of the LANs where no group member is active.

• IGMP protocol manages the group memberships.

• the tree becomes dependant of the localization of the group members (and of the source).
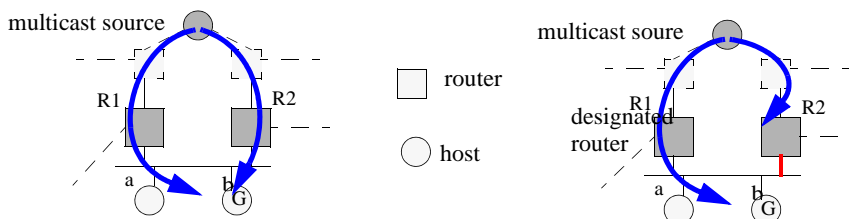

4.5. Truncated Spanning Tree

Truncated spanning tree

4.6. LAN with multiple routers

Some LANs are connected to several routers :

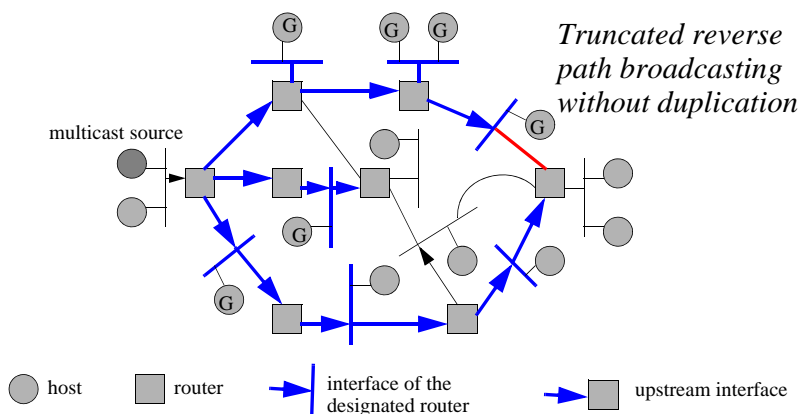• Hosts on these LANs may received several copies of the same multicast packet.

multicast source

R1   R2

□ router

○ host

multicast soure

R1   R2
designated
router

a   b
G

a   b
G

For each LAN, one router is designated as the multicast router. The router with:

• the shortest path toward the multicast source

• the smallest address (tie breaker)

A multicast router may detect the presence of another multicast router on the same LAN, with the help of:

- a routing protocol:

. heterogeneous routing protocols may lead to incoherence

- the IGMP protocol (version 2 or 3)

. IGMP requests may be periodically sent by multicast routers

4.7. Truncated Reverse Path Broadcasting

Truncated reverse path broadcasting without duplication

○ host   □ router   interface of the designated router   upstream interface

4.8. Pruning

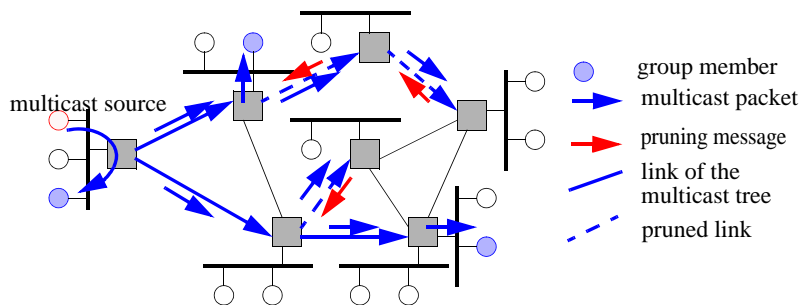Some branches of the Spanning Tree are useless!

- Some specific messages are used to prune these useless branches.

If all the router interfaces are inactive for a multicast group:
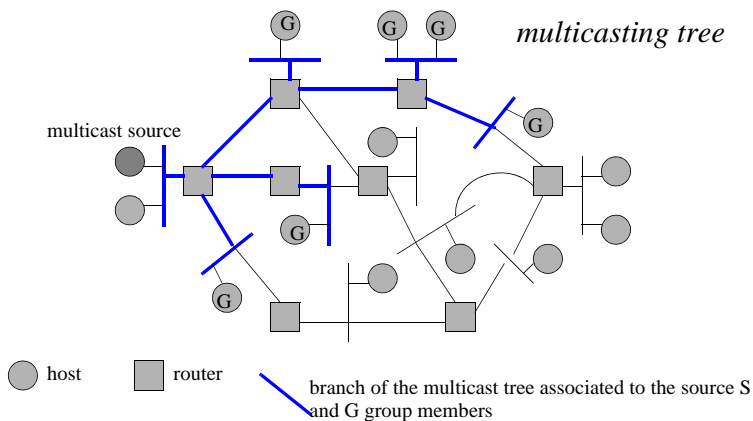
- the router sends a pruning message for this group on the upstream interface.

A router interface is inactive for a group, when:

- there is no active group member (through IGMP), and
- all downstream routers on this interface have sent a prune message for the group.

multicast source

group member
multicast packet
pruning message
link of the
multicast tree
pruned link

4.9. Reverse Path Multicasting Tree

multicasting tree

multicast source

host   router   branch of the multicast tree associated to the source S
and G group members

5. Conclusion

Multicast transmission services are useful, but can require much network resource and, when uncontrolled, can block the network.

The multicast forwarding and multicast IP addressing.

Main types of broadcasting trees:
- Steiner trees
- shortest path trees

One tree distributed computation has been described: reverse path multicasting.

Multicast protocols:
- multicast routing protocol: DVMRP, PIM-DM, MOSPF, PIM-SM, MBGP; MSDP
- explicit multicast protocol: XCAST (for small groups)
- multicast transport protocol: MTP, RMP, SRP RMTP, RAMP, etc.
- multicast group management, etc.

6. Some Additional Information

6.1. LAN Broadcasting

Broadcast service is natural on LAN.
- The medium is shared
- All network cards on a LAN received a copy of all frames sent on the LAN.
  - frames are discarded into the network card
- Same cost: individual frame or group frame
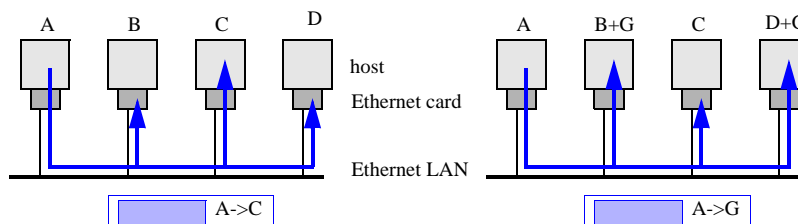
On LAN, a group frame is a frame with a destination address with the G bit set.

6.2. Multicast Address Resolution

- multicast IP adress => group MAC address

    - multicast IP adress:

                          1110 xxxx xabc defg hijk lmno pqrs tuvw

                                  23 lowest bits

    - group IEEE 802 address:
    0000 0001 0000 0000 0101 1110 0abc defg hijk lmno pqrs tuvw

Nota:
    . 16 multicast IP addresses are associated with the same MAC group address
        E.g.: 224.1.2.16 or 225.129.2.16 $\Rightarrow$ 01 00 5E 01 02 10