Advanced Networking

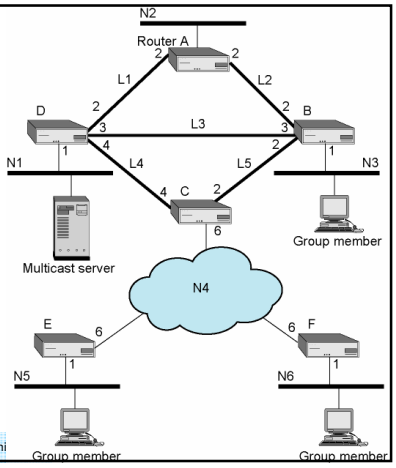
Multicast

Renato Lo Cigno
Renato.LoCigno@dit.unitn.it
Homepage:
disi.unitn.it/locigno/index.php/teaching-duties/advanced-networking

Multicasting

- Addresses that refer to group of hosts on one or more networks
- Applications
  - Multimedia "broadcast" and streaming
  - Teleconferencing
  - Distributed Database
  - Distributed computing (GRID??)
  - Real time workgroups

Example of multicast configuration

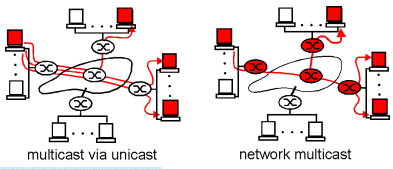


Broadcast and Multiple Unicast

- Broadcast a copy of packet to each network
  - Requires 13 copies of packet
- Multiple Unicast
  - Send packet only to networks that have hosts in group
  - 11 packets

Multicast Routing

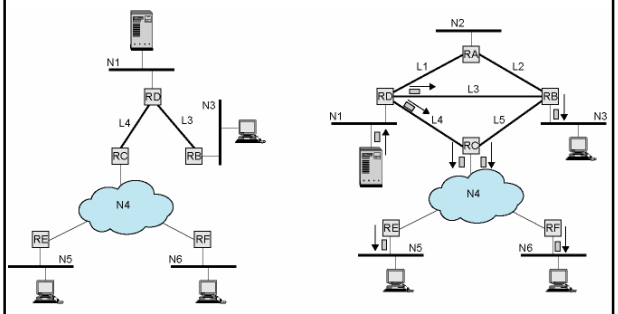
- Multicast: delivery of same packet to a group of receivers
- Multicasting is becoming increasingly popular in the Internet (video on demand; whiteboard; interactive games)
- Multiple unicast vs multicast



True Multicast

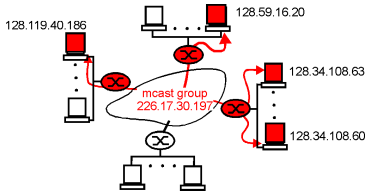
- Determine least cost path to each network that has host in group
  - Gives spanning tree configuration containing networks with group members
- Transmit single packet along spanning tree
- Routers replicate packets at branch points of spanning tree
- 8 packets required

Multicast Transmission Example



Multicast Group Address

- M-cast group address "delivered" to all receivers in the group
- Internet uses Class D for m-cast
- M-cast address distribution etc. managed by IGMP Protocol



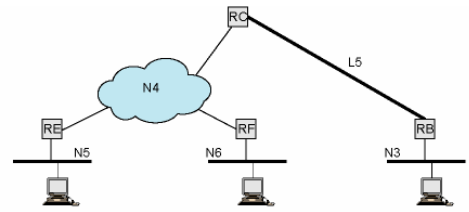
Requirements for Multicasting (1)

- Router may have to forward more than one copy of packet
- Convention needed to identify multicast addresses
  - IPv4 - Class D - start 1110
  - IPv6 - 8 bit prefix, all 1, 4 bit flags field, 4 bit scope field, 112 bit group identifier
- Nodes must translate between IP multicast addresses and list of networks containing group members
- Router must translate between IP multicast address and network multicast address

Requirements for Multicasting (2)

- Mechanism required for hosts to join and leave multicast group
- Routers must exchange info
  - Which networks include members of given group
  - Sufficient info to work out shortest path to each network
  - Routing algorithm to work out shortest path
  - Routers must determine routing paths based on source and destination addresses

Spanning Tree from Router C to Multicast Group



Internet Group Management Protocol (IGMP)

- RFC 3376
- Host and router exchange of multicast group info
- Use broadcast LAN to transfer info among multiple hosts and routers

Principle of Operations

- Hosts send messages to routers to subscribe to and unsubscribe from multicast group
  - Group defined by multicast address
- Routers check which multicast groups of interest to which hosts
- IGMP currently version 3
- IGMPv1
  - Hosts could join group
  - Routers used timer to unsubscribe members

Operation of IGMP v1 & v2

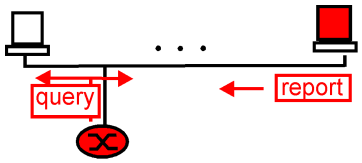
- Receivers have to subscribe to groups
- Sources do not have to subscribe to groups
- Any host can send traffic to any multicast group
- Problems:
  - Spamming of multicast groups
  - Even if application level filters drop unwanted packets, they consume valuable resources
  - Establishment of distribution trees is problematic
  - Location of sources is not known
  - Finding globally unique multicast addresses difficult

IGMP v3

- Allows hosts to specify list from which they want to receive traffic
  - Traffic from other hosts blocked at routers
- Allows hosts to block packets from sources that send unwanted traffic

IGMP dialogues

- IGMP (Internet Group Management Protocol) operates between Router and local Hosts, typically attached via a LAN (e.g., Ethernet)
- Router queries the local Hosts for m-cast group membership info

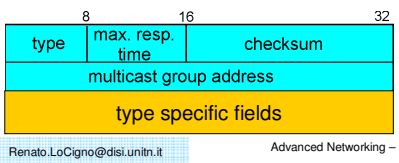


IGMP Protocol

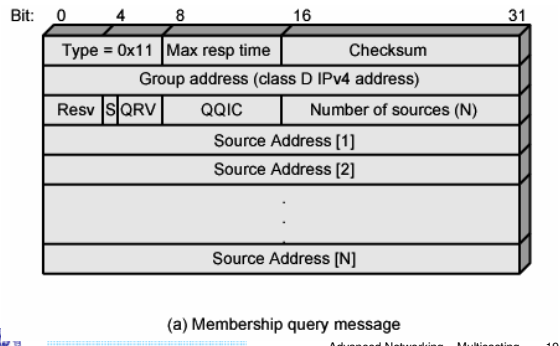
- Router "connects" active Hosts to m-cast tree via m-cast protocol
- Hosts respond with membership reports: actually, the first Host which responds (at random) speaks for all
- Host issues "leave-group" mssg to leave; this is optional since router periodically polls anyway (soft state concept)

IGMP message types

IGMP Message type	Sent by	Purpose
membership query: general	router	query for current active multicast groups
membership query: specific	router	query for specific m-cast group
membership report	host	host wants to join goup
leave group	host	host leaves the group



IGMP Message Formats: Membership Query



(a) Membership query message

Membership Query

- Sent by multicast router
- General query
  - Which groups have members on attached network
- Group-specific query
  - Does group have members on an attached network
- Group-and-source specific query
  - Do attached device want packets sent to specified multicast address
  - From any of specified list of sources

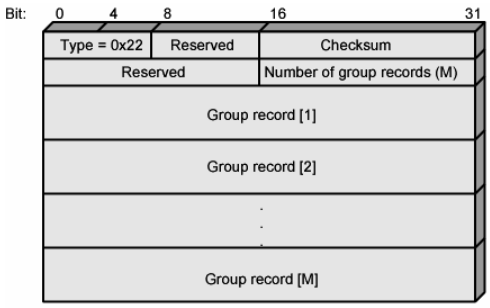
Membership Query Fields (1)

- Type
- Max Response Time
  - Max time before sending report in units of 1/10 second
- Checksum
  - Same algorithm as IPv4
- Group Address
  - Zero for general query message
  - Multicast group address for group-specific or group-and-source
- S Flag
  - 1 indicates that receiving routers should suppress normal timer updates done on hearing query

Membership Query Fields (2)

- QRV (querier's robustness variable)
  - RV value used by sender of query
  - Routers adopt value from most recently received query
  - Unless RV was zero, when default or statically configured value used
  - RV dictates number of retransmissions to assure report not missed
- QQIC (querier's querier interval code)
  - QI value used by querier
  - Timer for sending multiple queries
  - Routers not current querier adopt most recently received QI
  - Unless QI was zero, when default QI value used
- Number of Sources
- Source addresses
  - One 32 bit unicast address for each source

IGMP Message Formats: Membership Report

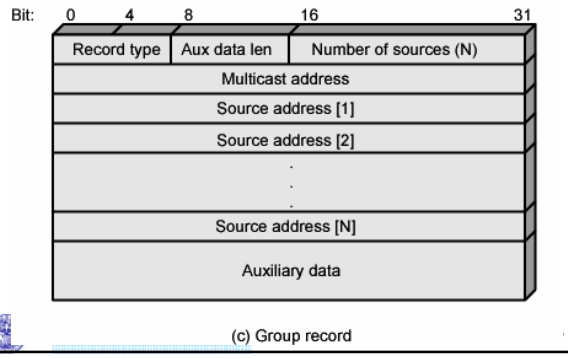


(b) Membership report message

Membership Reports

- Type
- Checksum
- Number of Group Records
- Group Records
  - One 32-bit unicast address per source

IGMP Message Formats: Group Record



(c) Group record

Group Record

- Record Type
  - **"Current-State Record"** (in response to a Query)
    - MODE_IS_INCLUDE INCLUDE()
    - MODE_IS_EXCLUDE EXCLUDE()
  - **"Filter-Mode-Change Record"** (when the filter mode change)
    - CHANGE_TO_INCLUDE_MODE TO_IN()
    - CHANGE_TO_EXCLUDE_MODE TO_EX()
  - **"Source-List-Change Record"** (when the source list change)
    - ALLOW_NEW_SOURCES ALLOW()
    - BLOCK_OLD_SOURCES BLOCK()
- Aux Data Length
  - In 32-bit words
- Number of Sources
- Multicast Address
- Source Addresses
  - One 32-bit unicast address per source
- Auxiliary Data
  - Currently, no auxiliary data values defined

IGMP Operation - Joining

- Host using IGMP wants to make itself known as group member to other hosts and routers on LAN
- IGMPv3 can signal group membership with filtering capabilities with respect to sources
  - EXCLUDE mode – all group members except those listed
  - INCLUDE mode – Only from group members listed
- To join group, host sends IGMP membership report message
  - Address field multicast address of group
  - Sent in IP datagram with Group Address field of IGMP message and Destination Address encapsulating IP header same
  - Current members of group will receive learn of new member
  - Routers listen to all IP multicast addresses to hear all reports

IGMP Operation – Keeping Lists Valid

- Routers periodically issue IGMP general query message
  - In datagram with all-hosts multicast address
  - Hosts that wish to remain in groups must read datagrams with this all-hosts address
  - Hosts respond with report message for each group to which it claims membership
- Router does not need to know every host in a group
  - Needs to know at least one group member still active
  - Each host in group sets timer with random delay
  - Host that hears another claim membership cancels own report
  - If timer expires, host sends report
  - Only one member of each group reports to router

IGMP Operation - Leaving

- Host leaves group, by sending leave group message to all-routers static multicast address
- Send membership report message with EXCLUDE option and null list of source addresses
- Router determine if there are any remaining group members using group-specific query message

Multicast Extension to OSPF (MOSPF)

- Enables routing of IP multicast datagrams within single AS
- Each router uses MOSPF to maintain local group membership information
- Each router periodically floods this to all routers in area
- Routers build shortest path spanning tree from a source network to all networks containing members of group (Dijkstra)
  - Takes time, so on demand only

Forwarding Multicast Packets

- If multicast address not recognised, discard
- If router attaches to a network containing a member of group, transmit copy to that network
- Consult spanning tree for this source-destination pair and forward to other routers if required

Equal Cost Multipath Ambiguities

- Dijkstra' algorithm will include one of multiple equal cost paths
  - Which depends on order of processing nodes
- For multicast, all routers must have same spanning tree for given source node
- MOSPF has tiebreaker rule

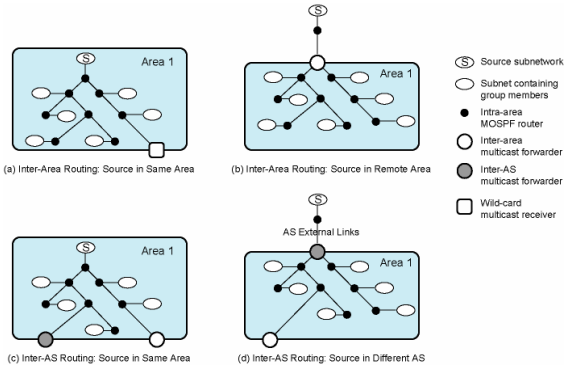
Interarea Multicasting

- Multicast groups may contain members from more than one area
- Routers only know about multicast groups with members in its area
- Subset of area's border routers forward group membership information and multicast datagrams between areas
  - Interarea multicast forwarders

Inter-AS Multicasting

- Certain boundary routers act as inter-AS multicast forwarders
  - Run and inter-AS multicast routing protocol as well as MOSPF and OSPF
  - MOSPF makes sure they receive all multicast datagrams from within AS
  - Each such router forwards if required
  - Use reverse path routing to determine source
    - Assume datagram from X enters AS at point advertising shortest route back to X
    - Use this to determine path of datagram through MOSPF AS

Illustrations of MOSPF Routing

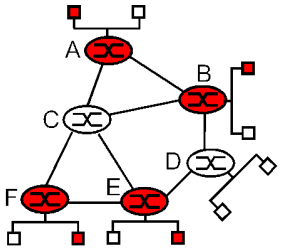


Multicast Routing Protocol Characteristics

- Extension to existing protocol
  - MOSPF v OSPF
- Designed to be efficient for high concentration of group members
- Appropriate with single AS
- Not for large internet

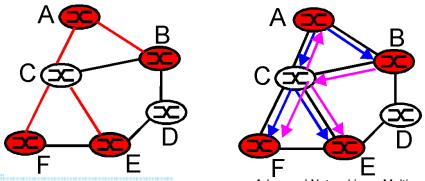
The Multicast Tree problem

- Problem: find the best (e.g., min cost) tree which interconnects all the members



Multicast Tree options

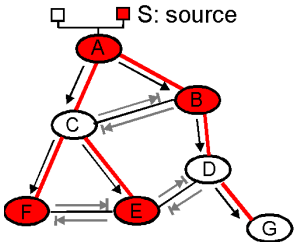
- **GROUP SHARED TREE**: single tree; the root is the "**CORE**" or the "**Rendez Vous**" point; all messages go through the **CORE**
- **SOURCE BASED TREE**: each source is the root of its own tree connecting to all the members; thus N separate trees



Group Shared Tree

- Predefined **CORE** for given m-cast group (eg, posted on web page)
- New members "join" and "leave" the tree with explicit join and leave control messages
- Tree grows as new branches are "grafted" onto the tree
- CBT (Core Based Tree) and PIM Sparse-Mode are Internet m-cast protocols based on GSTree
- All packets go through the **CORE**

Group Shared Tree



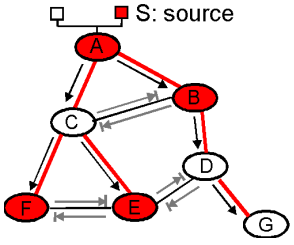
Legend

- router with attached group member
- router with no attached group member
- pkt that will be forwarded
- pkt not forwarded beyond receiving router

Source Based Tree

- Each source is the root of its own tree: the tree of shortest paths
- Packets delivered on the tree using "reverse path forwarding" (RPF); i.e., a router accepts a packet originated by source S only if such packet is forwarded by the neighbor on the shortest path to S
- In other words, m-cast packets are "forwarded" on paths which are the "reverse" of "shortest paths" to S

Source Based Tree



Legend

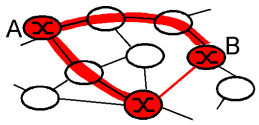
- router with attached group member
- router with no attached group member
- pkt that will be forwarded
- pkt not forwarded beyond receiving router

Source-Based tree: DVMRP

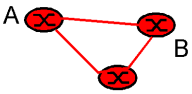
- DVMRP was the first m-cast protocol deployed on the Internet; used in Mbone (Multicast Backbone)
- Initially, the source broadcasts the packet to ALL routers (using RPF)
- Routers with no active Hosts (in this m-cast group) "prune" the tree; i.e., they disconnect themselves from the tree

Source-Based tree: DVMRP

- Recursively, interior routers with no active descendents self-prune After timeout (2 hours in Internet) pruned branches "grow back"
- Problems: only few routers are mcast-able; solution: tunnels



physical topology



logical mcast topology

PIM (Protocol Independent Multicast)

- Is becoming the de facto intra AS m-cast protocol standard
- "Protocol Independent" because it can operate on different routing infrastructures
  - Extract required routing information from any unicast routing protocol
  - Work across multiple AS with different unicast routing protocols
- PIM can operate in two modes:
  - PIM Sparse
  - PIM dense Mode

PIM (Protocol Independent Multicast)

- Initially, members join the "Shared Tree" centered around a Randez Vous Point
- Later, once the "connection" to the shared tree has been established, opportunities to connet **DIRECTLY** to the source are explored (thus establishing a partial Source Based tree)

PIM Strategy

- Flooding is inefficient over large sparse internet
- Little opportunity for shared spanning trees
- Focus on providing multiple shortest path unicast routes
- Dense mode
  - For intra-AS
  - Alternative to MOSPF
- Sparse mode
  - Inter-AS multicast routing

Spars Mode PIM

- A spars group:
  - Number of networks/domains with group members present significantly small than number of networks/domains in internet
  - Internet spanned by group not sufficiently resource rich to ignore overhead of current multicast schemes

Group Destination Router Group Source Router

- Group Destination Router
  - Has local group members
  - Router becomes destination router for given group when at least one host joins group
    - Using IGMP or similar
- Group source router
  - Attaches to network with at least one host transmitting on multicast address via that router

PIM Approach

- For a group, one router designated rendezvous point (RP)
- Group destination router sends join message towards RP requesting its members be added to group
  - Use unicast shortest path route to send
  - Reverse path becomes part of distribution tree for this RP to listeners in this group
- Node sending to group sends towards RP using shortest path unicast route
- Destination router may replace group-shared tree with shortest path tree to any source
  - By sending a join back to source router along unicast shortest path
- Selection of RP dynamic
  - Not critical

Example of PIM Operation