



IP Multicast

Luca Veltri

(mail.to: luca.veltri@unipr.it)

Corso di Reti di Telecomunicazioni C, a.a. 2004/2005

<http://www.tlc.unipr.it/veltri>

Unicast/Multicast/Broadcast Communications

- Unicast: one to one communication
 - one source and one destination (one-to-one relation)
 - in unicast routing, when a node receive a packet, it forwards the packet through only one output interface (link)
 - in connection-less (datagram) communications single unicast address is used for the destination
- Multicast: (potentially) many to many communication
 - one source and a group of destinations (one-to-many relation)
 - in multicast routing, when a node receive a packet, it may forward the packet through several output interfaces (links)
 - replication happens inside network nodes (routers or switches)
 - in connection-less (datagram) communications a group address or an address list is used for the destinations
 - may require dynamic management of group memberships
- Broadcast: one to all communication
 - one source, but all the others are the destinations
 - a broadcast address is used as destination

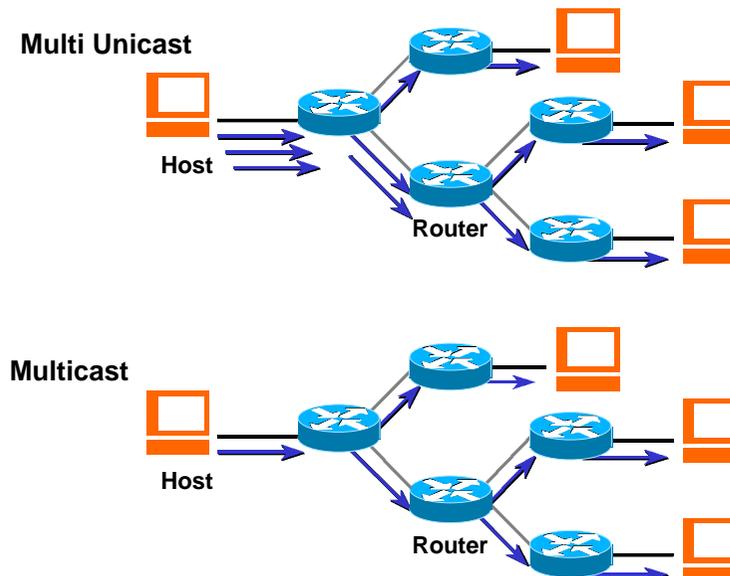
2

Multiple Unicast Communications

- Not all networks natively support multicast or broadcast
- When needed, multicast can be emulated through multiple unicast
- Multiple Unicast: one-to-many communication through multiple one-to-one communications
 - each packet is duplicated and sent separately to every destinations
- Big problem with multi unicast communications: bandwidth waste with multiple data flows
 - with N receivers, sender must replicate the stream N times.
 - Consider good quality audio/video streams are about 1.5Mb/s
 - Each additional receiver requires another 1.5Mb/s of capacity on the sender network

3

Multi Unicast vs Multicast



4

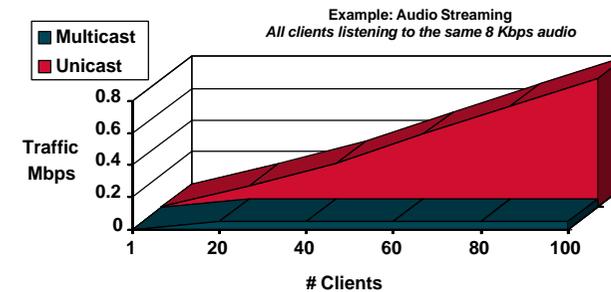
Why Multicast?

- When sending same data to multiple receivers
- Better bandwidth utilization
- Lesser host/router processing
- Receivers' addresses unknown

5

Why Multicast? - Advantages

- Efficiency and Performance
 - Eliminates traffic redundancy
 - Controls network traffic
 - reduces server CPU and output link loads
- Distributed Applications
 - Makes multipoint applications possible



6

Multicast Applications

- Typical applications
 - Multimedia conference (video, audio, digital whiteboard)
 - Video Distribution
 - Distance Learning
 - Software/File Distribution
 - Replicated Database Updates
 - Resource discovery (e.g., auto-topology)
 - Commercial apps (e.g., transactions, news distribution)
 - Routing protocols (e.g., both EIGRP and OSPF use multicast to send updates to neighbors)
 - Games (e.g., distributed arcades)
 - etc.
- Examples
 - VIC - Video conferencing
 - VAT/RAT - Audio conferencing
 - WB - Whiteboard

7

Disadvantages

- Possible packet duplication
 - in case of path redundancy between source and destinations
 - latency in convergence of multicast protocols
- Unreliable delivery
 - TCP works only with unicast connections
 - UDP is normally used as transport protocol (best effort)
- Possible network congestion
 - no congestion control with UDP
- Hence, more work is needed at application level, or new multicast transport protocols should be implemented on top of IP multicast
 - however, reliable multicast is still an area open for much research

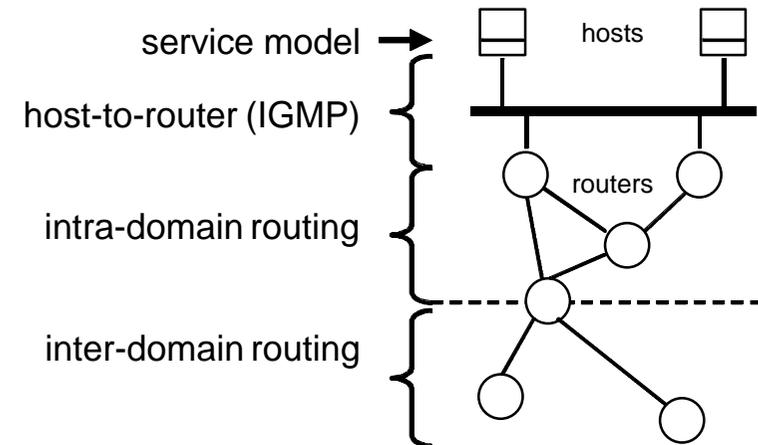
8

Principles of IP Multicast

- Special IP addresses are used to identify multicast groups
- Hosts notify multicast routers about the multicast groups to which they (want to) belong
- Multicast groups are managed by the routers using multicast routing protocols

9

Components of IP Multicast



10

Original IP Multicast Service Model (RFC-1112)

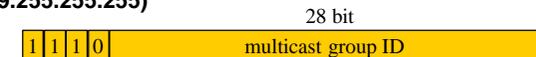
- Each multicast group identified by a class D IP address
- A multicast IP packet reaches a subset (group) of hosts on the network; those hosts have indicated an interest in the multicast group address
- groups may be of any size
- Members of the group could be present anywhere in the Internet
- Members join and leave the group and indicate this to the routers
- Senders and receivers are distinct, i.e., a sender need not be a member
- Routers listen to all multicast addresses and use multicast routing protocols to manage groups

11

Multicast Group Addresses

- IP addresses from 224.0.0.0 to 239.255.255.255 are designated as multicast addresses

Class D (224.0.0.0 - 239.255.255.255)



- Group addresses have inherent scope:
 - **Reserved for link scope: 224.0.0.0 -- 224.0.0.255**
 - These are never forwarded by any router
 - Some special addresses
 - 224.0.0.1: all multicast systems on a subnet
 - 224.0.0.2: all multicast routers on a subnet
 - **Global scope (Internet-wide): 224.0.1.0 -- 238.255.255.255**
 - Can be delivered throughout the Internet
 - **Administrative/Limited scope (local): 239.0.0.0 -- 239.255.255.255**
 - Not forwarded beyond an organization's intranet (like RFC 1918)
 - Reusable

12

Multicast Group Addresses (cont.)

- Need to map multicast group addresses to data link multicast addresses (e.g. Ethernet)
 - RFC 1112 defines OUI 0x01005e
 - 25th bit is set to 0 (value 1 is reserved for further uses)
 - Low-order 23-bits of IP address map into low-order 23 bits of IEEE address (eg. 224.2.2.2-01005e.020202)
- Ethernet and FDDI use this mapping

13

IP Multicast Service — Sending

- Uses normal IP-Send operation, with an IP multicast address specified as the destination
- Must provide sending application a way to:
 - specify outgoing network interface, if >1 available
 - specify IP time-to-live (TTL) on outgoing packet

14

IP Multicast Service — Receiving

- Two new operations:
 - Join-IP-Multicast-Group (group-address, interface)
 - Leave-IP-Multicast-Group (group-address, interface)
- Receive multicast packets for joined groups via normal IP-Receive operation

15

TTL Scope

- As in unicast IP forwarding, the TTL field of a multicast IP packet is decremented by every router that forwards it
- A source can choose any TTL for multicast packets that it transmits
- TTL controls how far (multicast) packets travel before being dropped (limiting the packet scope)
 - prevents clients that are really far away from source
- Router interfaces can be configured to drop multicast packets with a TTL less than some arbitrary (positive) value, rather than allowing the TTL to count down to zero
 - Cisco IOS interface configuration command:
 - ip multicast ttl-threshold 16
 - These thresholds create TTL scope boundaries
 - E.g. packets must start with a TTL of at least 16 to be forwarded beyond the campus network

16

TTL Scope (cont.)

- General standards for TTL (from IETF)
 - 1 for local net
 - 15 for site
 - 63 for region
 - And 127 for world

17

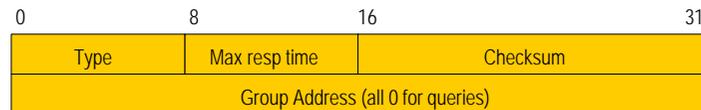
Internet Group Management Protocol (IGMP)

- The protocol by which hosts indicate their multicast group memberships (i.e. interest in receiving packets addressed to a particular multicast group G) to neighboring routers
- Routers solicit group membership from directly connected hosts
- IGMP messages aren't forwarded by routers
- Supported on several end systems: UNIX, PCs, and MACs
- IGMP was originally defined in RFC 1112
- IGMP v2 and IGMP v3 enhancements
 - RFC 1112 specifies version 1, the original standard
 - RFC 2236 specifies version 2, the most widely used, backward-compatible with version 1
 - RFC 3376 specifies version 3, new standard, backward-compatible with version 2 and 1

18

IGMPv2 message format

- IGMP messages are encapsulated in IPv4 datagrams, with an IP protocol number of 2
- Every IGMP message is sent with IP TTL=1



- Type: IGMP message type
Membership Query, Membership Report, Leave Group
- Max Resp Time: max allowed time before sending a responding report in units of 1/10 second (only in Membership Query)
- Checksum
- Group address

19

IGMPv2 messages

- Membership Query (0x11)
 - There are two sub-types of Membership Query messages:
 - General Query, used to learn which groups have members on an attached network
 - Group-Specific Query, used to learn if a particular group has any members on an attached network
 - These two messages are differentiated by the Group Address.
- Version 2 Membership Report (0x16)
- Leave Group (0x17)

20

IGMP Overview

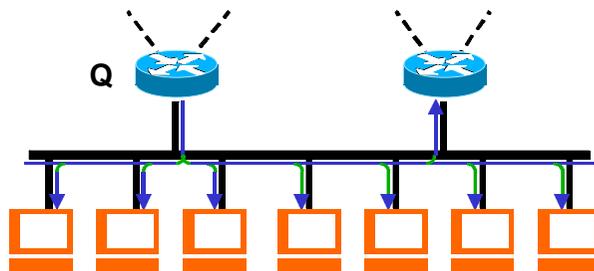
- Periodically, a router attached to a particular subnet sends a General Membership Query to determine what group addresses have members on the subnet
 - These queries are sent to multicast address 224.0.0.1 (the all systems group). All hosts (and routers) listen to this group
 - Query interval is typically 60—90 seconds
- One router on every subnet is designated as the IGMP Querier
 - The querier is responsible for sending membership queries to the subnet to determine group membership
 - All routers on the subnet listen to the membership reports sent by hosts and maintain forwarding states accordingly, regardless of which router is the querier

21

- When a host receives a general membership query, it does not respond immediately
 - The host sets a timer for every group G for which it is a member, using a random value between zero and D seconds.
 - If the timer for a group G expires, the host sends a Membership Report addressed to G
 - While the timer for a group G is non-zero, if the host overhears a report from another member of G, it clears the timer and does not send
- Normally, only one report message per group is sent in response to a query (routers need not know who all the members are)
- The maximum value of the query response timer D depends on the version of IGMP:
 - version 1: D = 10 seconds
 - version 2: the membership query contains a field that specifies the value of D, in tenths of a second

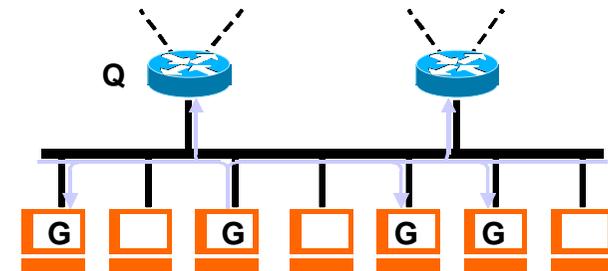
22

Membership Query



23

Membership Report

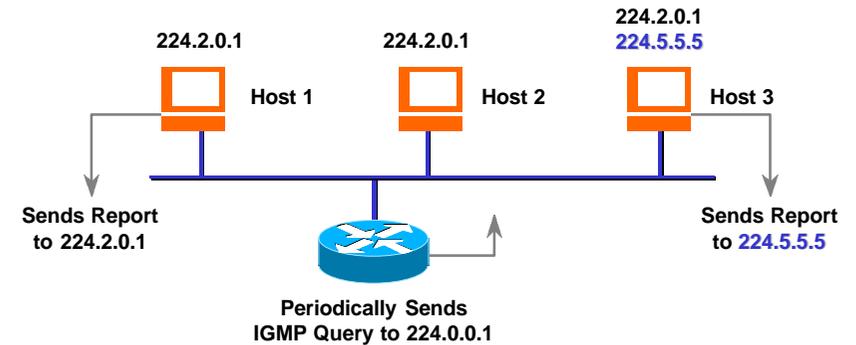


24

- When a host wishes to join a group G it immediately sends an unsolicited Group Membership Report
 - **Speeds up the join process when the host is the first on the subnet to join group G**
- When any router on a subnet receives a membership report for group G
 - if there are no states for group G, the router creates (*,G) state, and sets the oif (output interfaces) to the interface on which the report was received
 - if one or more states exist for group G, the oif list for every state involving G is updated to include the interface on which the report was received. If the interface is already in the oif list, its timer is refreshed

25

IGMP Example — Joining a Group



26

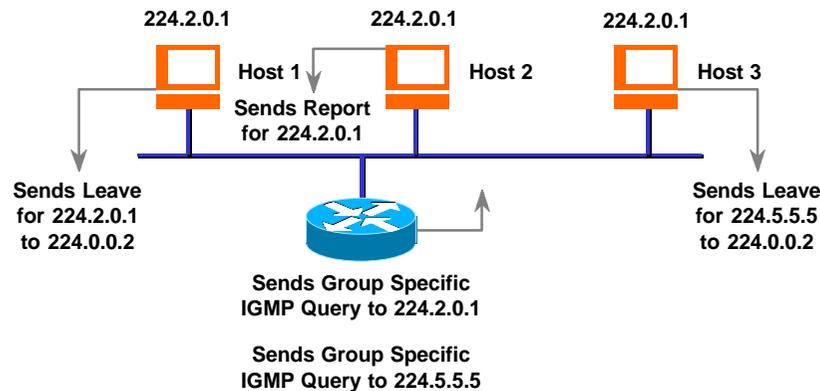
- When an IGMPv2 host wishes to leave a group G it may send a Leave Group message
 - **Sent to 224.0.0.2 (all routers) since other hosts don't care when any particular host leaves**
 - **Not required if the host wasn't the last membership reporter for G**
 - **Can be sent even if the host wasn't the last reporter for G**
- When a host sends a leave group message for a group G, the router sends a Group-Specific Membership Query
 - **Addressed to the group, G**
 - **Hosts follow the same rules as for the general query (i.e. delay before replying, and don't reply if someone else does first)**

27

- When a router does not receive any reply on an interface for a group G for which forwarding states exist:
 - **The outgoing interface timer for every state involving G continues to count down**
 - **When the timer for that interface in a G state expires, the interface is removed from the oif list for that state**

28

IGMP example — Leaving a Group



29

Election of the querier router

- The “election” of a querier is carried out as routers overhear each other’s membership queries
 - When a router first starts up, it assumes it is the querier for each of its interfaces
 - If the router overhears a query on an interface for which it believes itself to be the querier, and the IP address of the source of the query is numerically smaller than its own IP address on that interface, the router stops sending queries

30

Multicast Routing

- Per il trasporto del traffico multicast occorre che nei nodi della rete geografica siano attivate le funzioni necessarie per l'instradamento multicast (multicast routing)
 - si deve creare un albero che ha come radice la sorgente e come foglie i membri del gruppo multicast
 - l'instradamento multicast si basa sia sull'indirizzo della sorgente che sul gruppo multicast (destinazione) - *source-based routing*
 - l'instradamento unicast invece si basava sul solo indirizzo di destinazione (source-based) - *destination based routing*
- Sono stati definiti differenti algoritmi di instradamento multicast
 - **multicast routing algorithms**
- Da cui sono stati specificati differenti protocolli di instradamento multicast
 - **multicast routing protocols**

31

Multicast Algorithms: Flooding (inondazione)

- Estremamente semplice
 - se il pacchetto è ricevuto per la prima volta, allora il router provvede a replicare il pacchetto ricevuto e a ritrasmetterlo attraverso tutte le proprie interfacce, ad eccezione di quella dalla quale il pacchetto è pervenuto
- la difficoltà riscontrata consiste proprio nel determinare se il pacchetto è stato effettivamente ricevuto per la prima volta
 - Una soluzione potrebbe essere quella di tenere traccia di tutti i pacchetti pervenuti al router, ma richiederebbe elevate risorse di memorizzazione e elaborazione
- L'algoritmo non richiede per il funzionamento informazioni sull'instradamento
 - non comporta perciò la predisposizione di alcuna tabella di instradamento multicast
- Utilizzato nei protocolli di instradamento unicast, (e.g. OSPF) per scambiare le informazioni di routing tra i nodi

32

Multicast Algorithms: Spanning Tree

- Soluzione più efficiente del flooding
- Utilizzata ad esempio dai bridge per interconnettere diverse LAN in modo da evitare percorsi chiusi (loop)
- L'algoritmo agisce in una prima fase per individuare i rami che costituiscono lo spanning tree (l'albero ricoprente)
 - si identificano le interfacce dei router agli estremi dei rami dell'albero ricoprente
- Durante l'instradamento ciascun router replica i pacchetti multicast sulle sole interfacce/rami dello spanning tree (ad eccezione dell'interfaccia di arrivato)
- Si evitano così possibili loop
- Traffico solo su una porzione della rete (albero)
- Non distingue il routing per i differenti gruppi multicast

33

Multicast Algorithms: Reverse Path Forwarding

- Reverse Path Forwarding (RPF)
 - alla ricezione di un pacchetto multicast un router analizza l'indirizzo della sorgente "S" e quello dell'interfaccia "I" attraverso la quale è arrivato il singolo pacchetto
 - se "I" si trova sul percorso più breve verso "S" (ovvero se "I" è l'interfaccia usata dal router per instradare i pacchetti verso "S"), allora il pacchetto è replicato ed è inoltrato verso tutte le interfacce ad eccezione di "I"
 - nel caso non si sia verificata la condizione precedente, il pacchetto è scartato
- L'algoritmo richiede una tabella che indichi per ciascuna sorgente l'interfaccia del nodo sul percorso più breve verso la sorgente
 - a questo scopo, potrebbe essere utilizzata la tabella di instradamento unicast
 - siccome il routing non è in generale simmetrico, alcuni protocolli realizzano una tabella ad hoc

34

Multicast Algorithms: RPF and prunes

- Variante dell'algoritmo RPF con "potatura"
- L'albero multicast è potato di tutti i rami a cui non è attestato alcun nodo interessato al gruppo multicast
 - I nodi foglia senza membri del gruppo "G" inviano un messaggio di potatura (prune) al router multicast a monte
 - il router a monte è così informato che non deve inoltrare ulteriore traffico multicast destinato a G verso il router a valle
 - in questo modo, partendo dalle foglie e ripercorrendo l'albero verso la radice, sono potati i rami sui quali è inutile inoltrare traffico
- L'algoritmo RPF introduce il concetto di appartenenza ai gruppi e richiede che i router tengano traccia dello stato dell'albero per gruppo e per sorgente
 - lo stato dell'albero viene aggiornato periodicamente

35

Multicast Algorithms: Shortest Tree

- Shortest Path Tree
 - L'algoritmo Shortest Path Tree individua il cammino più breve tra la sorgente (radice dell'albero) e ognuno dei ricevitori (le foglie dell'albero)
 - i due più noti algoritmi utilizzati sono quello di Bellman-Ford e quello di Dijkstra
- Steiner Tree
 - L'albero di Steiner è quello che rende minimo il numero di collegamenti utilizzati per connettere i membri di un gruppo all'interno di un grafo
 - i cammini prescelti, ottimizzano la condivisione delle connessioni e NON la distanza dalla sorgente alle destinazioni

36

Multicast Routing Protocols: Classification

- Source-Based Tree or Shared-Tree
- Intra or Inter-domain
- Dense Mode or Sparse Mode
- Protocol Dependent vs Protocol Independent

37

Source-Based Tree vs Shared-Tree

- SBT (Source-Based Tree)
 - ha come obiettivo la costruzione di alberi di distribuzione da ogni sorgente verso l'insieme completo dei ricevitori di un gruppo multicast
 - tanti alberi quante sono le sorgenti di traffico multicast
 - La costruzione di ogni albero avviene seguendo la strada più breve tra sorgente e destinazioni del traffico
 - scarsa scalabilità
 - e.g. DVMRP e PIM-DM
- Shared-Tree
 - sviluppata con l'obiettivo di superare le limitazioni di SBT
 - prevede la costruzione di un unico albero di distribuzione multicast intorno a un router (o a più di uno) chiamato core o RP (Rendez-vous Point)
 - Per ogni gruppo multicast viene individuato un core router che opera come punto di raccolta dei flussi multicast e come radice dell'albero di distribuzione
 - migliore scalabilità
 - possibilità di costituire percorsi di instradamento non ottimizzati
 - e.g. CBT e PIM-SM

38

Dense Mode vs Sparse Mode

- Dense Mode
 - Impiegano strategie basate su inondazioni e potature periodiche
 - Non sono consigliabili per le reti geografiche, a causa dell'elevato carico del traffico "di segnalazione" introdotto
 - Sono adatti a contesti caratterizzati da un'alta concentrazione di utenti multicast
 - e.g. DVMRP e PIM-DM
- Sparse Mode
 - adesione esplicita
 - rendono così minimo il traffico "di segnalazione"
 - indicati per le reti geografiche e più in generale per quei contesti in cui la densità di utenza multicast è bassa
 - e.g. CBT e PIM-SM

39

Protocol Dependent vs Protocol Independent

- I protocolli di routing multicast utilizzano informazioni di routing unicast per costruire gli alberi di distribuzione del traffico multicast
- Due approcci:
 - protocolli di routing multicast che utilizzano le informazioni di instradamento unicast generate da altri protocolli (OSPF, RIP, routing statico)
 - e.g. PIM-DM e PIM-SM
 - protocolli di routing multicast che costruiscono le proprie tabelle di instradamento unicast, disaccoppiate da quelle realmente utilizzate per instradare il traffico unicast
 - e.g. DVMRP

40