



IP Multicast

Luca Veltri

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

Corso di Reti di Telecomunicazioni C, a.a. 2009/2010

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

Unicast/Multicast/Broadcast

- Unicast: comunicazione punto-punto
 - una sorgente e una destinazione (relazione one-to-one)
 - in routing unicast, quando un nodo riceve un pacchetto lo rilancia su una sola interfaccia di uscita (link)
 - nel caso di commutazione senza connessione CL (datagram) la destinazione è identificata da un singolo indirizzo unicast (del nodo di destinazione)
- Multicast: comunicazione da uno a molti (potenzialmente)
 - una sorgente e un gruppo di destinazioni (relazione one-to-many)
 - in routing multicast, quando un nodo riceve un pacchetto lo rilancia su diverse interfacce di uscita (link)
 - la replica avviene nel nodo di commutazione (router o switch)
 - nel caso di commutazione senza connessione CL (datagram) la destinazione è identificata da un singolo indirizzo multicast (indirizzo di gruppo) o attraverso una lista di indirizzi
 - may require dynamic management of group memberships
- Broadcast: comunicazione da uno a tutti
 - una sorgente e tutte le possibili destinazioni (relazione one-to-all)
 - viene normalmente usato un indirizzo speciale per identificare tutte le possibili destinazioni (indirizzo broadcast)

2

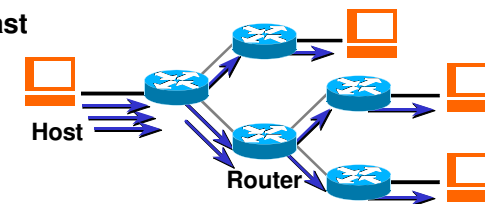
Multi Unicast

- Non tutte le reti e i protocolli supportano nativamente il multicast
- Se in un tratto di rete non è supportato il multicast, il nodo sorgente realizza il collegamento multicast verso un insieme di N nodi di next-hop con N collegamenti unicast
 - **multiple unicast: comunicazione one-to-many attraverso una molteplicità di collegamenti one-to-one**
 - **ogni pacchetto è replicato e rilanciato separatamente verso ogni next-hop**
- Grosso problema con multiple unicast: spreco di banda dovuta alla replicazione dei flussi

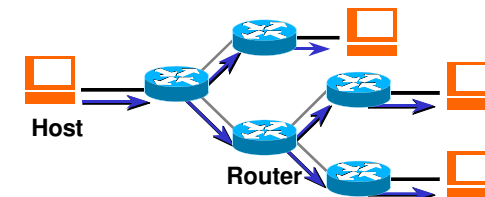
3

Multi Unicast vs Multicast

Multi Unicast



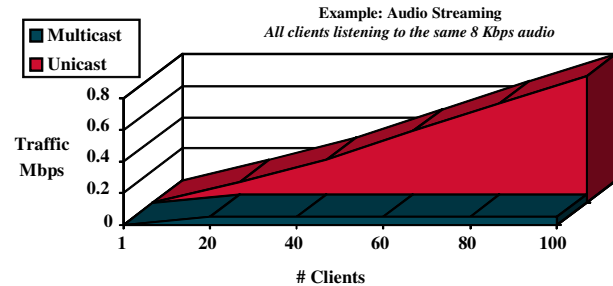
Multicast



4

Perche il Multicast?

- Efficiency and Performance
 - **Eliminates traffic redundancy**
 - Better bandwidth utilization
 - Lesser host/router processing
 - reduces server CPU and output link loads
- Distributed Applications
 - **Makes multipoint applications possible (same data to multiple receivers)**
- Receivers' addresses unknown



5

Esempi di applicazioni multicast

- 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**

6

Svantaggi

- 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**

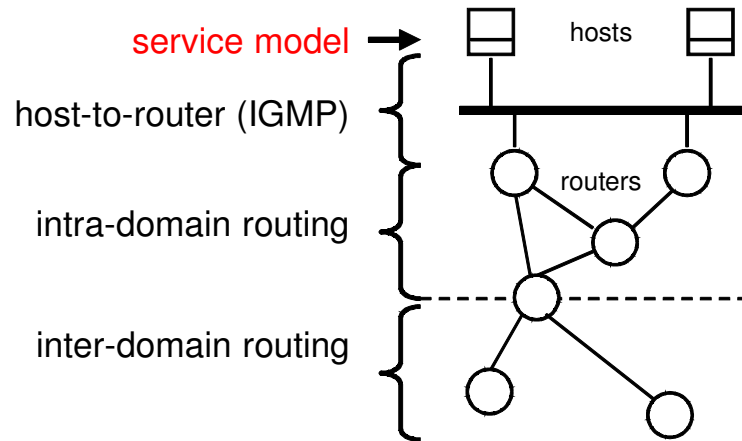
7

IP Multicast: principi base

- Speciali indirizzi IP usati per identificare differenti gruppi multicast
- Gli host notificano la loro volontà a partecipare ad uno o più gruppi multicast (ovvero ricevere i pacchetti a loro destinati) identificati dai corrispondenti indirizzi multicast
- I gruppi multicast vengono gestiti a livello di routing dai router della rete tramite appositi protocolli di routing

8

IP Multicast: Componenti



9

Original IP Multicast Service Model (RFC-1112)

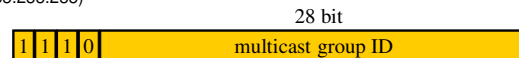
- Each multicast group identified by a class D IP address
- A multicast IP packet reaches a subset (group) of hosts on the network; those hosts have indicated an interest in the multicast group address
- groups may be of any size
- Members of the group could be present anywhere in the Internet
- Members 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

10

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 (224.0.0.22 = all IGMPv3 routers)
 - **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

11

Multicast Group Addresses (cont.)

- Need to map multicast group addresses to data link multicast addresses (e.g. Ethernet)
 - **RFC 1112 defines OUI 0x01005e (24bit)**
 - **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

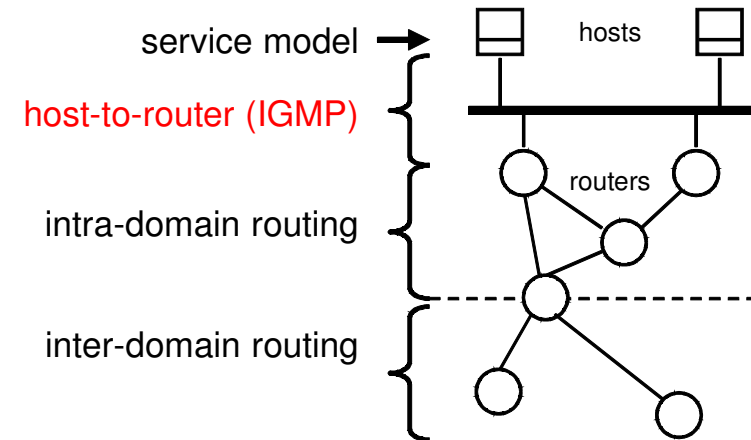
12

IP Multicast Service

- Invio
 - uses normal IP-send operation, with an IP multicast address specified as destination
 - must provide sending application a way to:
 - specify outgoing network interface, if >1 available
 - specify IP time-to-live (TTL) on outgoing packet
- Ricezione
 - 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
- TTL
 - as in unicast IP forwarding, the TTL field 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

13

Components of IP Multicast



14

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
- 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

15

IGMPv3 Overview

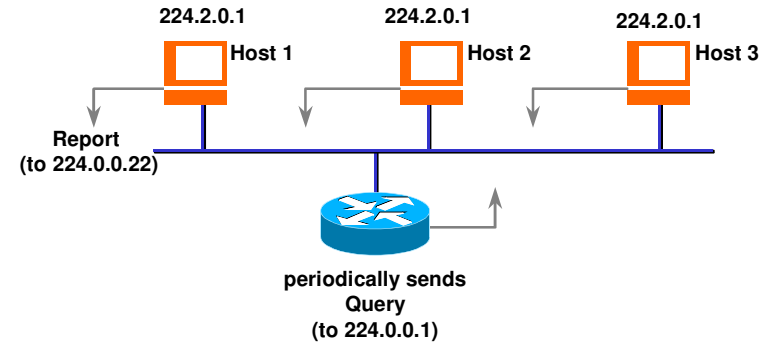
- IGMP is an asymmetric protocol, specifying separate behaviors for group members (hosts or routers that wish to receive multicast packets) and multicast routers
 - a multicast router that is also a group member performs both parts of IGMPv3
 - a system performs the protocol operation over all interfaces on which multicast reception is supported
- The purpose of IGMP is to enable each multicast router to learn, for each of its directly attached networks, which multicast addresses are of interest to the systems attached to those networks
 - IGMP version 3 added the capability for a multicast router to also learn which sources are of interest to neighboring systems, for packets sent to any particular multicast address

16

- Multicast routers send General Queries periodically to request group membership information from an attached network
 - These queries are sent to multicast address 224.0.0.1 (the all systems group). All hosts (and routers) listen to this group
 - Systems respond to these queries by reporting their group membership state (and their desired set of sources) with Current-State Group Records in IGMPv3 Membership Reports
- 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

17

Membership Query/Report



18

IGMPv3 messages

- There are two IGMP message types of concern to the IGMPv3 protocol:
 - **Membership Query (type 0x11)**
 - Membership Queries are sent by IP multicast routers to query the multicast reception state of neighboring interfaces
 - **Version 3 Membership Report (0x22)**
 - Version 3 Membership Reports are sent by IP systems to report (to neighboring routers) the current multicast reception state, or changes in the multicast reception state, of their interfaces
- An implementation of IGMPv3 MUST also support the following three message types, for interoperability with previous versions of IGMP:
 - **Version 1 Membership Report (0x12)**
 - **Version 2 Membership Report (0x16)**
 - **Version 2 Leave Group (0x17)**
- IGMP messages are encapsulated in IPv4 datagrams, with an IP protocol number of 2
- Every IGMP message is sent with IP TTL=1

19

IGMPv3 Membership Query Message

0		8		16		31	
Type		Max resp code		Checksum			
Group Address (all 0 for queries)							
Resv	S	QRV	QQIC	Number of Sources (N)			
Source Address [1]							
Source Address [2]							
...							
Source Address [N]							

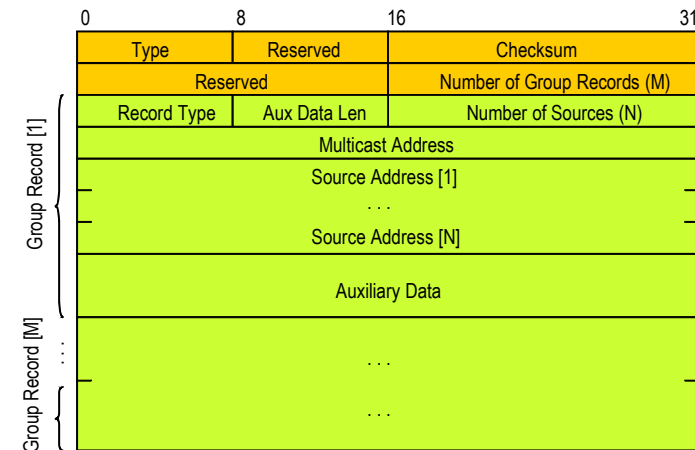
- Type: IGMP message type = 0x11 (Membership Query)
- Max Resp Code: max allowed time before sending a responding report
 - the actual time allowed, called the **Max Resp Time**, is represented in units of 1/10 second
- Checksum: The Checksum is the 16-bit one's complement of the one's complement sum of the whole IGMP message (the entire IP payload)

20

- The Group Address: is set to zero when sending a General Query, and set to the IP multicast address being queried when sending a Group-Specific Query or Group-and-Source-Specific Query
- Resv: reserved field
- S Flag: Suppress Router-Side Processing
 - indicates to any receiving multicast routers that they are to suppress the normal timer updates they perform upon hearing a Query
- QRV: Querier's Robustness Variable
 - If non-zero, it contains the Robustness Variable value used by the querier; Robustness Variable indicates the number of retransmissions
 - this field allow synchronization on non-Queriers
- QQIC: Querier's Query Interval Code
 - specifies the Query Interval used by the querier, that is the interval between General Queries sent by the Querier (Default: 125 seconds)
- Source Addresses: vector of N IP unicast addresses, where n is the value in the Number of Sources (N) field

21

IGMPv3 Membership Report Message



22

- Type: IGMP message type = 0x22 (Version 3 Membership Report)
- Multicast Address: contains the IP multicast address to which this Group Record pertains
- Group Record Type:
 - **Current-State Record, in response to a Query**
 - MODE_IS_INCLUDE (1)
 - MODE_IS_EXCLUDE (2)
 - **Filter-Mode-Change Record, unsolicited**
 - CHANGE_TO_INCLUDE_MODE (3)
 - CHANGE_TO_EXCLUDE_MODE (4)
 - **Source-List-Change Record, unsolicited**
 - ALLOW_NEW_SOURCES (5)
 - BLOCK_OLD_SOURCES (6)

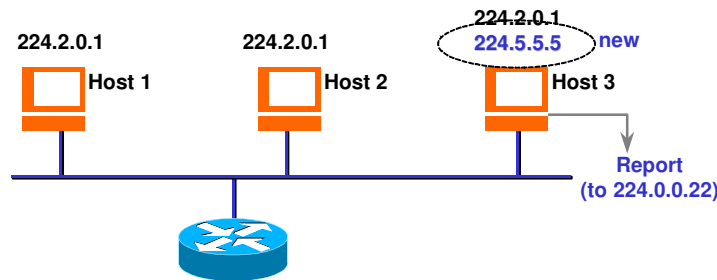
23

Joining a Group

- 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

24

Joining a Group: Example



25

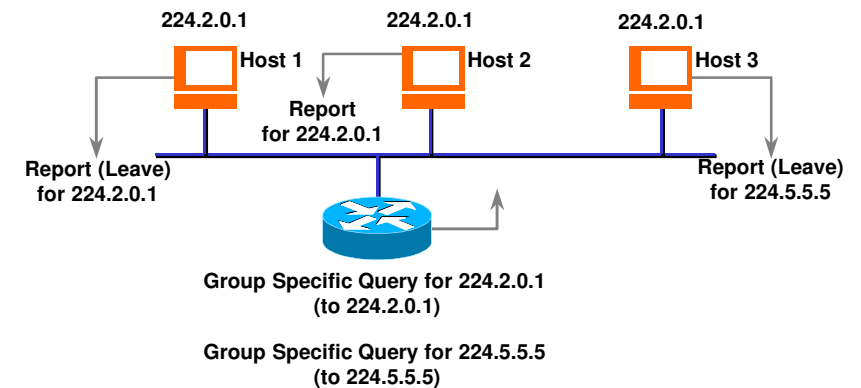
Leaving a Group

- When an host wishes to leave a group G it may send a Membership Report message (or a Leave Group message, in IGMPv2)
 - Sent to 224.0.0.22 (all IGMPv3 routers) since other hosts don't care when any particular host leaves
- 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)

26

- 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

Leaving a Group: Example



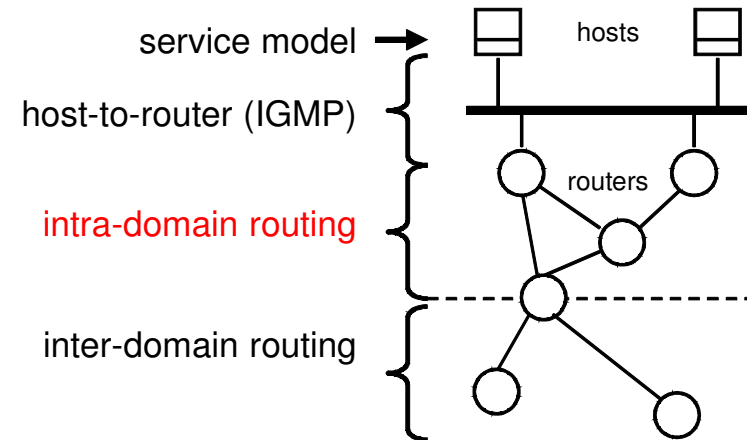
27

28

IP Multicast Routing

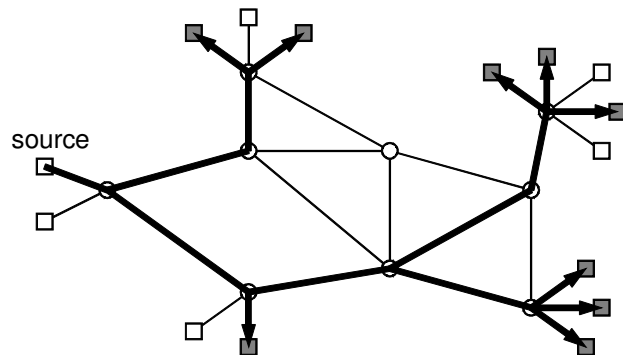
Lezione 35

Components of IP Multicast



30

Multicast Routing and Functions



- Routing (path determination)
- Packet forwarding and possibly replication
- Dynamic membership - path pruning/grafting

31

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 unicast si basava sul solo indirizzo di destinazione - *destination based routing*
 - l'instradamento multicast si basa sia sull'indirizzo della sorgente che sul gruppo multicast (destinazione) - *source-based routing*
- Sono stati individuati differenti algoritmi di instradamento multicast
 - **multicast routing algorithm**
- A partire dai quali sono stati definiti differenti protocolli di instradamento multicast
 - **multicast routing protocol**

32

Algoritmi Multicast: 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 ad esempio nei protocolli di instradamento unicast, (e.g. OSPF) per scambiare le informazioni di routing tra i nodi
- Non distingue il routing per i differenti gruppi multicast

33

Algoritmi Multicast: 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 da cui il pacchetto è arrivato)
- Si evitano così possibili loop
- Traffico solo su una porzione della rete (albero)
- Non distingue il routing per i differenti gruppi multicast

34

Algoritmi Multicast: 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**
- Non distingue il routing per i differenti gruppi multicast

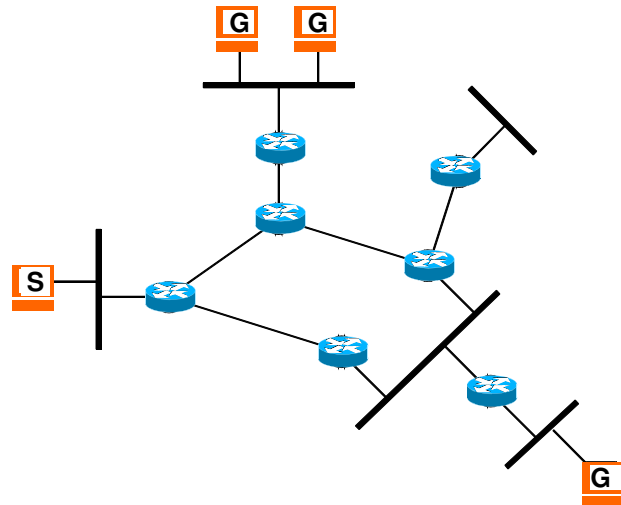
35

Algoritmi Multicast: RPF and prune

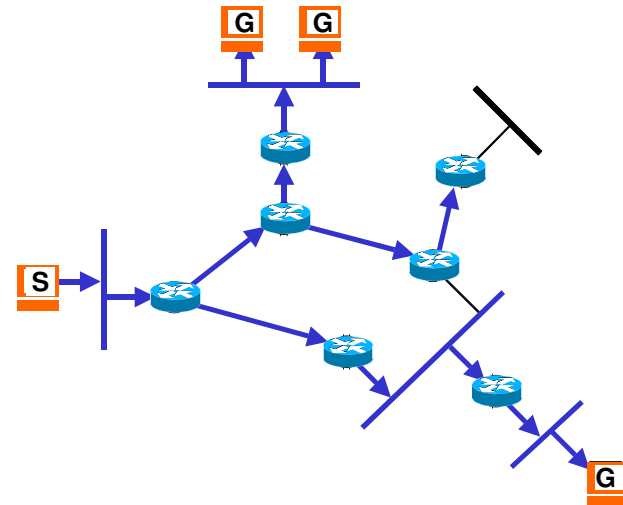
- Variante dell'algoritmo RPF con "potatura"
- L'albero multicast è potato di tutti i rami a cui non è attestato alcun nodo interessato al gruppo multicast
 - **I nodi foglia senza membri del gruppo "G" inviano un messaggio di potatura (prune) al router multicast a monte**
 - **il router a monte è così 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 con potatura 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**

36

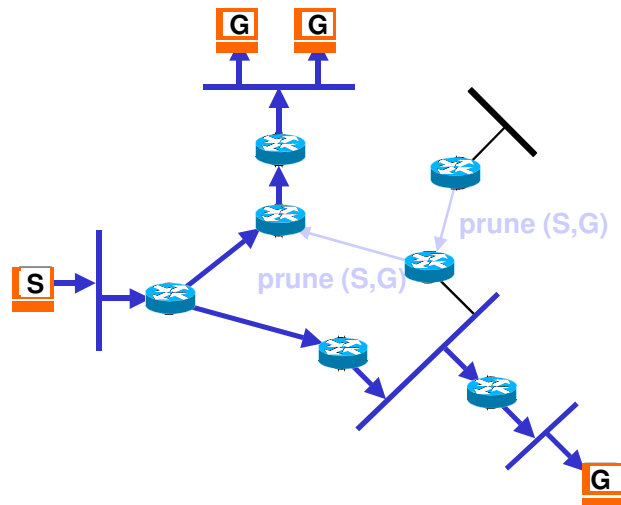
Esempio



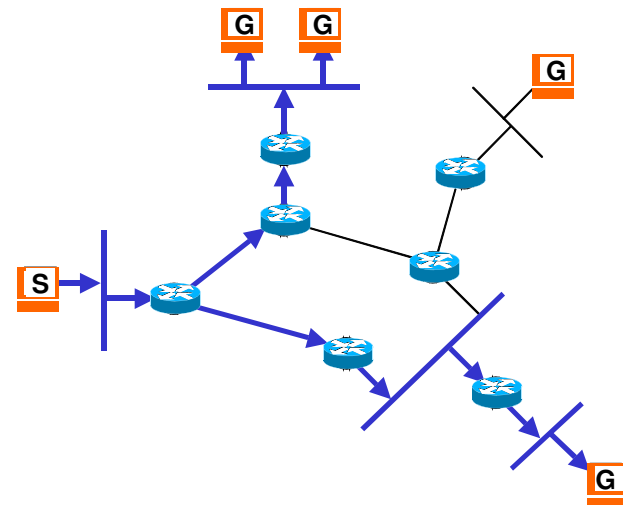
Fase 1: Truncated Broadcast



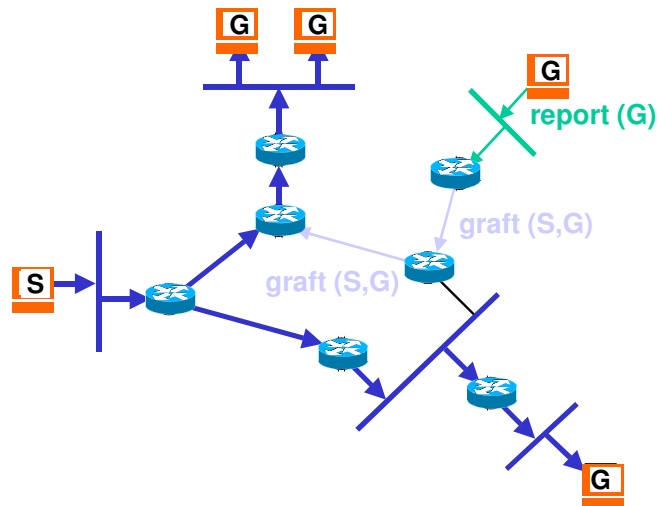
Fase 2: Pruning



Stato stabile

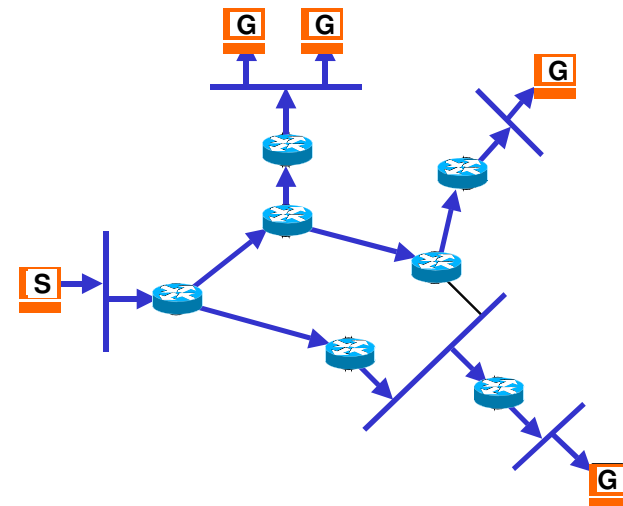


Aggiunta di nuovi ricevitori: Grafting



41

Stato stabile



42

Algoritmi Multicast: Shortest Tree

- Alcuni algoritmi costruiscono l'albero multicas cercando di minimizzare la distanza tra sorgente e destinatari (Shortest Path Tree) o il numero complessivo di rami (Steiner 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

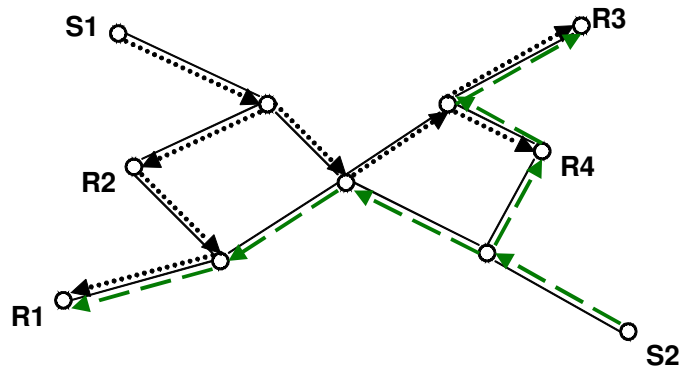
43

Algoritmi Multicast: altre differenze

- 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 ottimi
 - e.g. CBT e PIM-SM

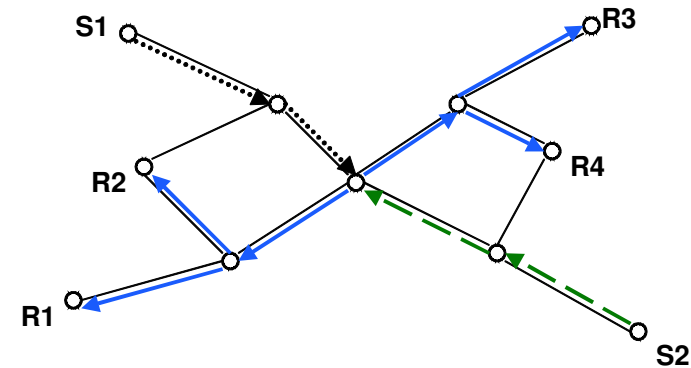
44

Un albero per sorgente



45

Un unico albero per tutte le sorgenti



46

Dense Mode vs Sparse Mode

- Dense Mode
 - impiegano strategie basate su inondazioni e potature periodiche
 - non sono consigliabili per reti di grosse dimensioni, 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, MOSPF & PIM-DM
- Sparse Mode
 - adesione esplicita
 - rendono così minimo il traffico "di segnalazione"
 - indicati per le reti di grosse dimensioni e più in generale per quei contesti in cui la densità di utenza multicast è bassa
 - e.g. CBT & PIM-SM

47

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:
 - **build own routing table (Protocol Dependent)**
 - protocolli di routing multicast che costruiscono le proprie tabelle di instradamento unicast, disaccoppiate da quelle realmente utilizzate per instradare il traffico unicast
 - e.g. DVMR
 - **use unicast routing table (Protocol Independent)**
 - 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

48

Attuali protocolli di routing multicast

- DVMRP — Distance-Vector Multicast Routing Protocol
 - dense mode protocol, Broadcast and prune
 - unidirectional per-source trees (source trees created on demand) based on RPF rule
 - builds own routing table with a protocol similar to RIP
 - many implementations

- MOSPF — Multicast Extensions to Open Shortest-Path First Protocol
 - broadcast membership
 - unidirectional per-source trees
 - uses unicast routing table

49

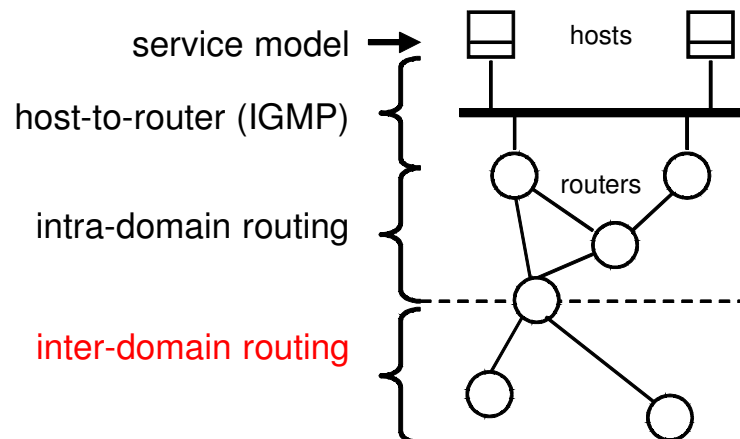
Attuali protocolli di routing multicast (cont.)

- PIM-DM (Protocol-Independent Multicast - Dense Mode)
 - broadcast-and-prune
 - unidirectional per-source trees
 - Protocol Independent, does not perform its own routing information exchange
 - uses unicast routing table made by any unicast routing protocols
 - similar to DVMRP, but:
 - without the routing information exchange part
 - differs in some minor details

- PIM-SM (Protocol-Independent Multicast - Sparse Mode)
 - uses meeting places (“rendezvous points”)
 - initially builds a single (unidirectional) tree per group, shared by all senders to that group
 - once data is flowing, the shared tree can be converted to a per-source, shortest-path tree if needed
 - uses unicast routing table

50

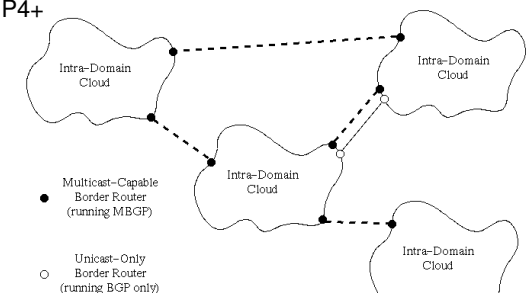
Components of IP Multicast



51

Inter-Domain Route Exchange

- Exchange multicast reachability between Autonomous Systems (AS)
 - Just like unicast routes are exchanged with BGP
 - Protocol is “Multiprotocol extensions to BGP” (RFC 2283)
 - Also known as “Multicast” BGP (MBGP)
 - Also known as BGP4+



52

Problemi con IP multicast

- Difficoltà nel gestire le informazioni di stato all'interno dei router multicast all'aumentare dei membri dei gruppi multicast
- Difficoltà nel gestire alberi multicast attraverso differenti ISP
 - **ad esempio nella scelta dei Rendez-vous Point**
- Topologie unicast e multicast potrebbero non essere congruenti attraverso diversi ISP
 - **a causa di vincoli topologici**
 - **a causa di vincolo amministrativi**
- Maggiore difficoltà ad individuare un modello di tariffazione (sorgente o destinatari?)
- Aumento della possibilità di attacchi di tipo Denial of Service
- Facilità di attacchi di tipo eavesdropping poiché la sorgente invia i dati ad un indirizzo multicast (di gruppo) e non ad una specifica destinazione