

Corso di Laurea in Ingegneria Informatica



Corso di Reti di Calcolatori

Roberto Canonico (roberto.canonico@unina.it)
Giorgio Ventre (giorgio.ventre@unina.it)

**Tecniche di trasmissione broadcast
IP multicasting: IGMP e introduzione al routing multicast**

I lucidi presentati al corso sono uno strumento didattico
che **NON sostituisce i testi indicati nel programma del corso**

Nota di copyright per le slide COMICS



Nota di Copyright

Questo insieme di trasparenze è stato ideato e realizzato dai ricercatori del Gruppo di Ricerca COMICS del Dipartimento di Informatica e Sistemistica dell'Università di Napoli Federico II. Esse possono essere impiegate liberamente per fini didattici esclusivamente senza fini di lucro, a meno di un esplicito consenso scritto degli Autori. Nell'uso dovranno essere esplicitamente riportati la fonte e gli Autori. Gli Autori non sono responsabili per eventuali imprecisioni contenute in tali trasparenze né per eventuali problemi, danni o malfunzionamenti derivanti dal loro uso o applicazione.

Autori:

Simon Pietro Romano, Antonio Pescapè, Stefano Avallone,
Marcello Esposito, Roberto Canonica, Giorgio Ventre

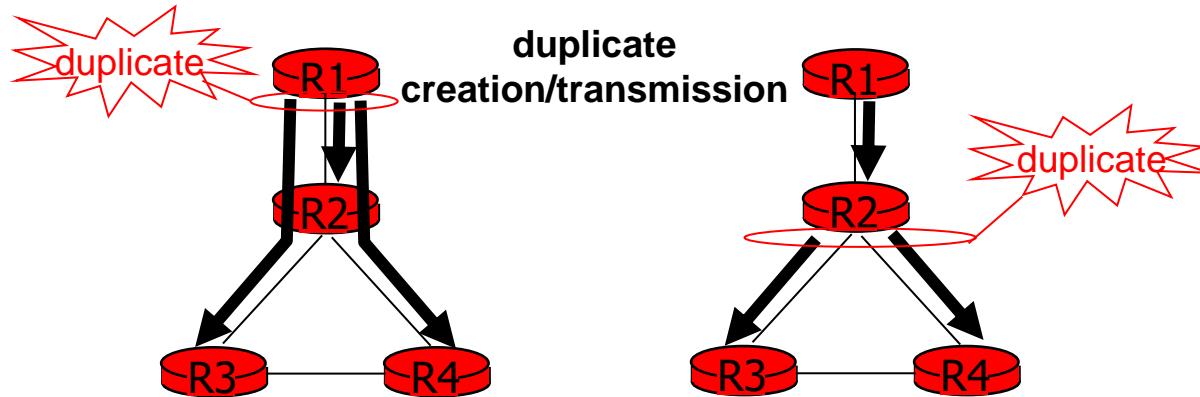


Tecniche di trasmissione broadcast



Broadcast Routing

- deliver packets from source to all other nodes
- source duplication is inefficient:



**source
duplication**

**in-network
duplication**

- source duplication: how does source determine recipient addresses?

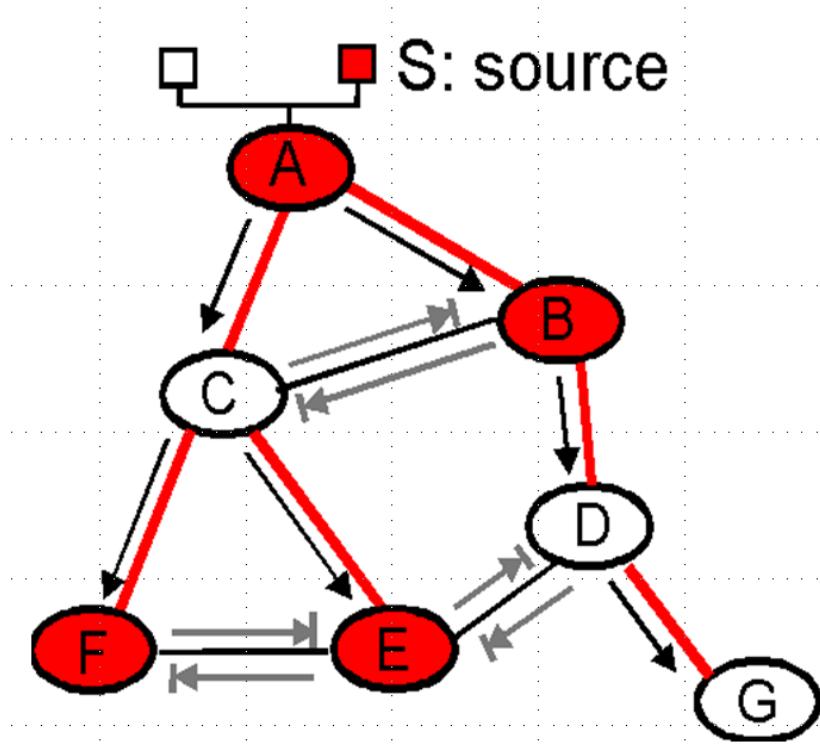


In-network duplication

1. **flooding**: when node receives broadcast packet, sends copy to all neighbors
 - Problems: cycles & broadcast storm
2. **controlled flooding**: node only broadcasts packet if it hasn't broadcasted same packet before
 1. **Sequence number**: node keeps track of pckt ids already broadcasted
 2. **Reverse path forwarding (RPF)**: only forward pckt if it arrived on shortest path between node and source
3. spanning tree
 - No redundant packets received by any node



Reverse packet forwarding (RPF)



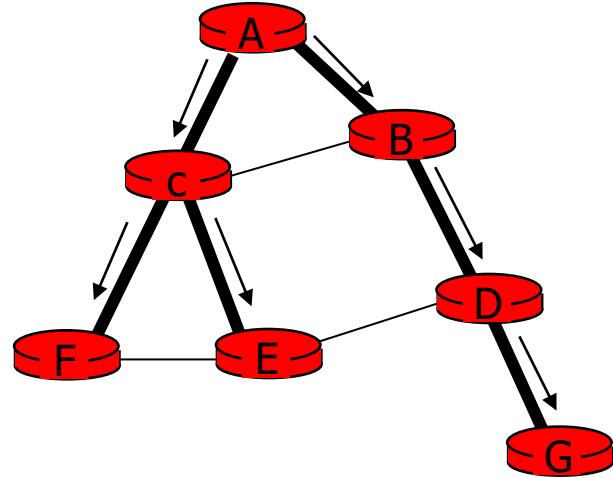
A node forwards a packet only if it arrived from the shortest path between node and source (reverse path)

Still some useless packets are transmitted

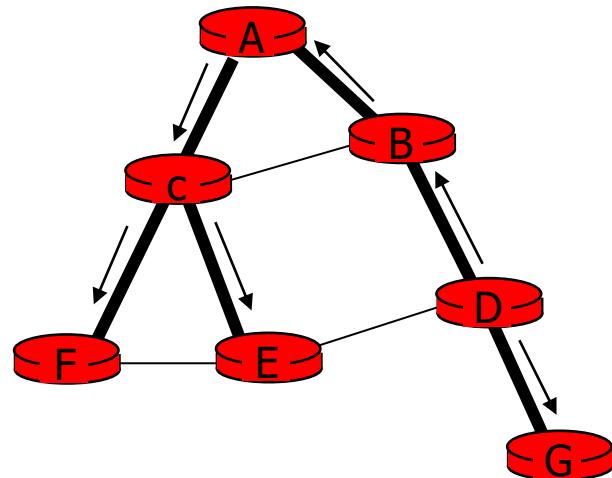


Spanning Tree

- First construct a spanning tree
- Nodes forward copies only along spanning tree



(a) Broadcast initiated at A

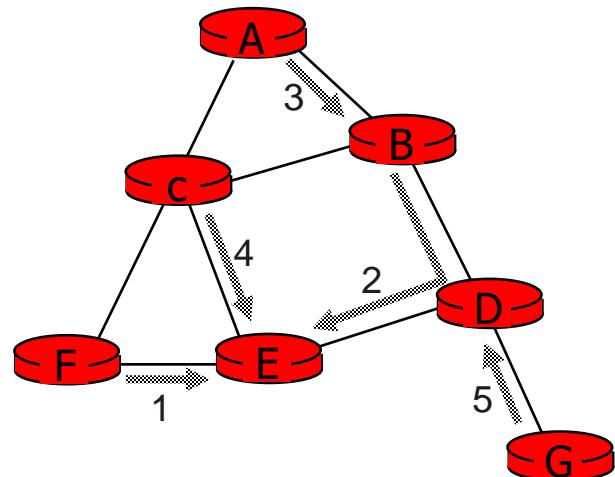


(b) Broadcast initiated at D

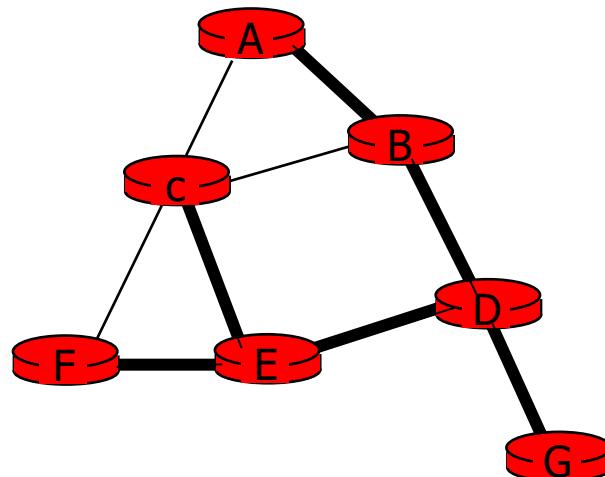
Spanning Tree: Creation



- First select a *Center node* (or *Rendevouz Point*) (eg. Node E)
- Each node sends unicast join message to center node
 - Message forwarded until it arrives at a node already belonging to spanning tree or arrives to the center node



**(a) Stepwise
construction of
spanning tree**



**(b) Constructed
spanning tree**



Trasmissione multicast in reti IP

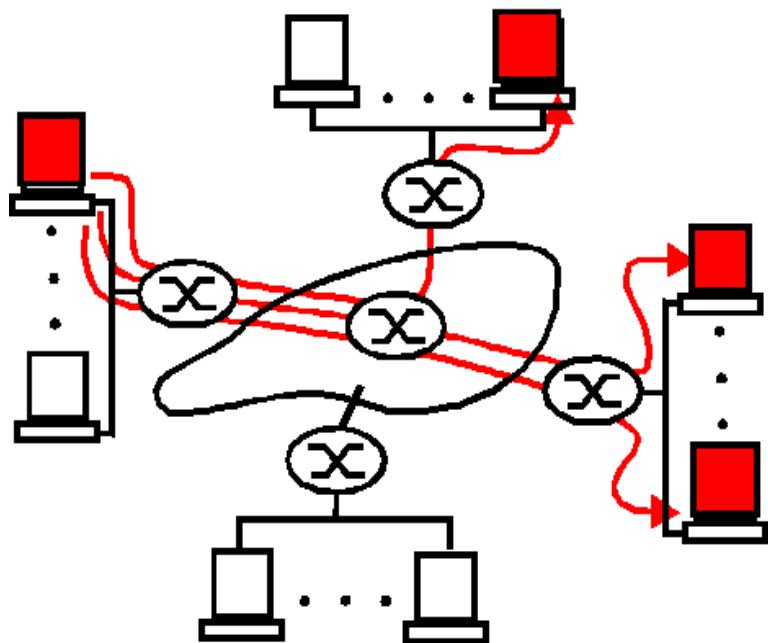


Multicasting: definizione del problema

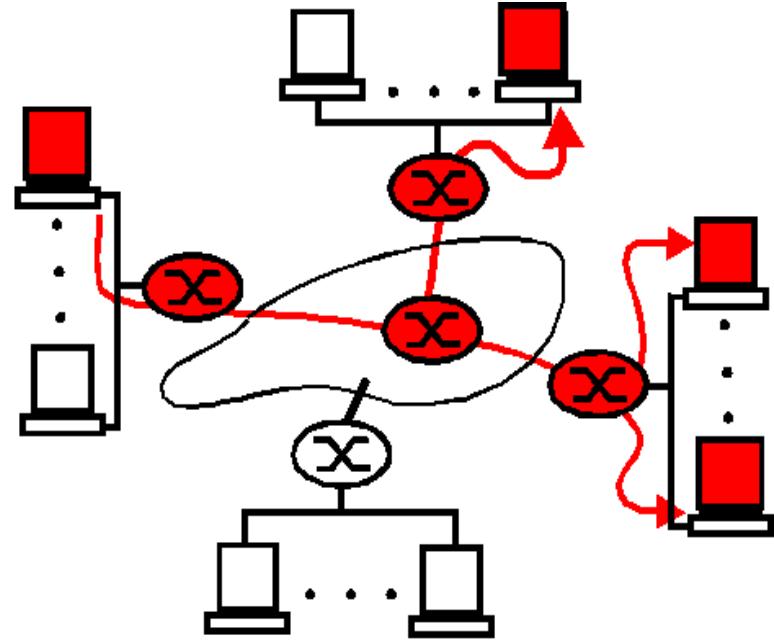
- Un numero sempre maggiore di applicazioni di rete richiedono la spedizione di pacchetti da uno o più sender a un gruppo di receiver
 - la trasmissione dell'aggiornamento di un software dal suo sviluppatore agli utenti che richiedono l'aggiornamento
 - il trasferimento di audio, video e testi per lettura diretta a un gruppo distribuito di partecipanti alla lettura
 - una riunione aziendale o una teleconferenza condivisa tra molti partecipanti distribuiti
 - quotazioni in borsa, distribuzione di listini e cataloghi in tempo reale
 - training, teledidattica
 - sfruttare Internet per trasmissioni di tipo televisivo
- Per ciascuna di queste applicazioni, un'astrazione molto utile è la nozione di **multicast**: l'invio di un pacchetto da un sender a molti receiver con una singola operazione di spedizione.



La trasmissione multicast



multicast via unicast



network multicast



La trasmissione multicast

**Come identificare i ricevitori di un
datagramma multicast?**

**Come inviare un datagramma ai ricevitori, una volta
identificati?**

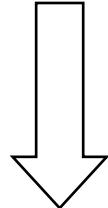




La trasmissione multicast



Address indirection:
da “indirizzo per destinazione” ad “indirizzo per evento”



Si utilizza un identificativo unico per il gruppo di ricevitori e una copia del datagramma è inviata, utilizzando tale identificativo, a tutti i membri del gruppo



IP multicasting: RFC 1112

- RFC 1112 (anno 1989) definisce:
 - Il concetto di gruppo multicast
 - Lo spazio di indirizzamento (classe D) usata per il multicast IP
 - Modalità di trasmissione di datagram IP multicast in reti locali Ethernet
 - Estensioni all'interfaccia di servizio IP verso i livelli superiori dello stack per ricevere datagrammi multicast
 - Definizione di operazioni di adesione esplicita ad un gruppo (*join*) ed abbandono esplicito di un gruppo (*leave*)
 - Definizione di un protocollo di controllo (IGMPv1) per l'interazione locale tra end-system e router multicast finalizzata alla gestione dei gruppi



Gli indirizzi multicast IPv4 (1/2)

Ad ogni gruppo è associato un indirizzo multicast,
cioè un indirizzo IP di classe D



Range da 224.0.0.0 a 239.255.255.255

All'interno di tale classe, un certo numero di indirizzi “well-known”
è riservato a gruppi permanenti

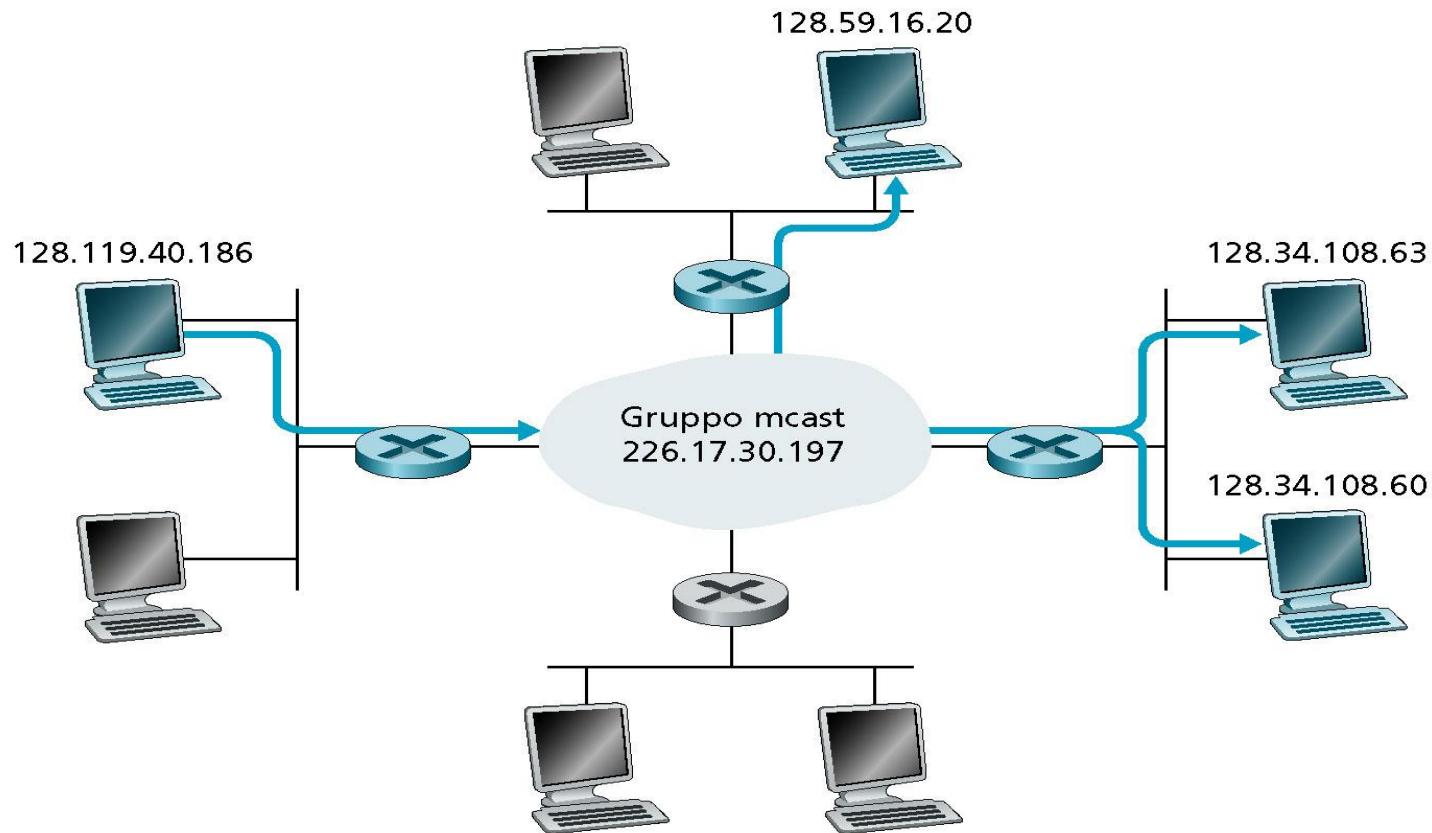


Gli indirizzi multicast IPv4 (2/2)

- Well known addresses stabiliti da IANA
 - Per usi riservati da 224.0.0.0 a 224.0.0.255
 - 224.0.0.1—all systems on subnet
 - 224.0.0.2—all routers on subnet
 - 224.0.0.13 – all PIM routers
 - <http://www.iana.org/assignments/multicast-addresses>
- Indirizzi assegnati dinamicamente:
 - Global scope: 224.0.1.0-238.255.255.255
 - Limited Scope: 239.0.0.0-239.255.255.255
 - Site-local scope: 239.255.0.0/16
 - Organization-local scope: 239.192.0.0/16



Il gruppo multicast



Legenda:

Router cui sono attaccati componenti del gruppo

Router cui non sono attaccati componenti del gruppo



IP multicast : distribuzione su LAN ethernet

- Per la trasmissione di datagram IP multicast su reti LAN ethernet, occorre mappare un indirizzo in classe D su di un indirizzo MAC multicast
- Non è possibile in maniera univoca, dato il range degli indirizzi MAC riservati al multicast

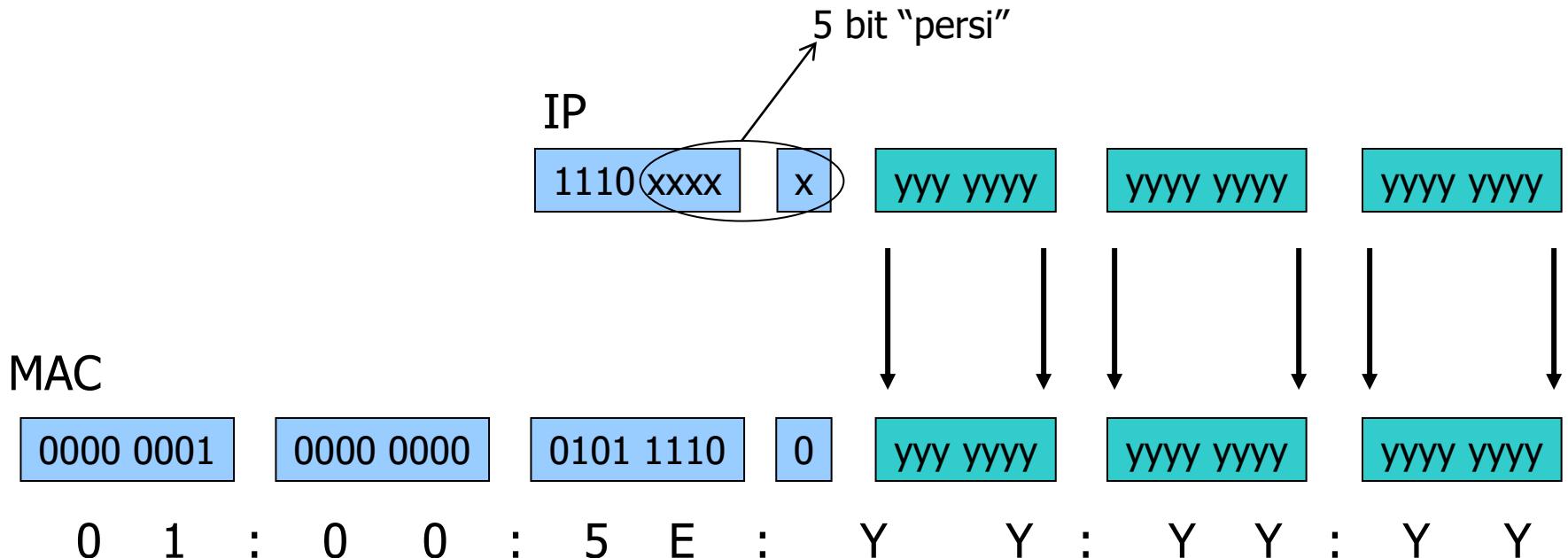
da 01:00:5e:00:00:00 a 01:00:5e:7f:ff:ff

- L'OUI 01:00:5E è riservato al mapping degli indirizzi IP multicast
- Indirizzo IP multicast: 28 bit liberi → Indirizzo MAC ethernet: 23 bit
- Aliasing: $2^5=32$ gruppi multicast IP per ogni MAC multicast
 - Possibili conflitti



IP multicast e Ethernet multicast

Il mapping tra indirizzi IP multicast ed indirizzi MAC multicast avviene come illustrato sotto





Session Announcement Protocol (SAP)

- Per annunciare una sessione multicast e la sua descrizione si utilizza il protocollo SAP. Ci sono diverse possibilità:
 - Sessioni Global Scope → 224.2.127.254 (SAP.MCAST.NET).
 - 224.2.128.0 – 224.2.255.255
 - Sessioni Administrative Scope → indirizzo più alto
 - Es. se gli indirizzi vanno dal 239.16.32.0 al 239.16.33.255, allora l'indirizzo a cui mandare i pacchetti multicast con gli annunci è il 239.16.33.255.
- Protocollo UDP, Porta 9875, TTL 255
- Per cancellare una sessione
 - Explicit Timeout: la durata è parte dell'annuncio
 - Implicit Timeout: se non si riceve nulla per un intervallo prefissato
 - Explicit Deletion



La gestione dei gruppi

- La gestione dei gruppi è di tipo dinamico:
 - Un host può unirsi o abbandonare un gruppo in qualsiasi momento e può appartenere contemporaneamente a più gruppi
 - Non è necessario appartenere ad un gruppo per poter inviare ad esso dei datagrammi
 - I membri del gruppo possono appartenere alla medesima rete o a reti fisiche differenti



Il multicast router

Si occupa dello smistamento dei datagrammi multicast,
in maniera trasparente riguardo agli host interessati
ad un determinato di gruppo



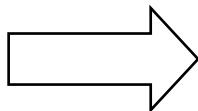
Il multicast router: funzionamento

- Ogni end-system trasmette i datagrammi multicast sfruttando il meccanismo hardware messo a disposizione dalla rete locale su cui si trova
- Se un datagramma giunge al multicast router, quest'ultimo si occupa, se necessario, di instradarlo verso le altre reti



Protocolli per il multicast in Internet

IGMP
**Internet Group
Management Protocol**



**Fornisce ad un host i mezzi
per informare il multicast
router ad esso più vicino
che un'applicazione vuole
unirsi ad un determinato
gruppo multicast**

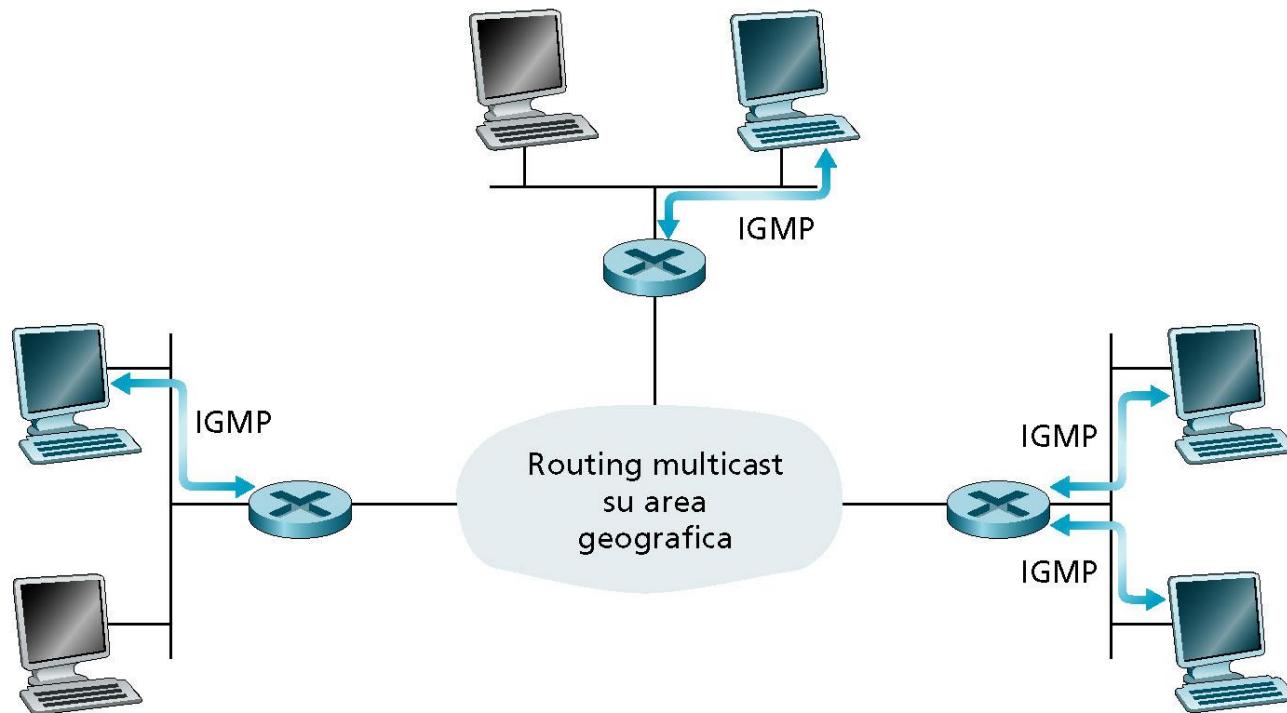
**Algoritmi per il
multicast routing**



**Coordinano i multicast
router all'interno della rete
Internet, per permettere
l'instradamento dei
datagrammi multicast**



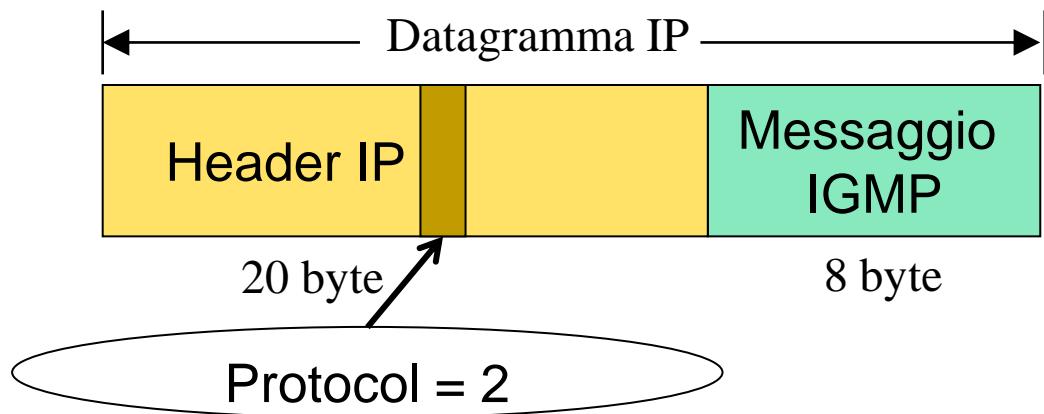
Internet Group Management Protocol opera tra un host ed il router ad esso direttamente collegato





Il protocollo IGMP

- IGMP serve a gestire in ambito locale le informazioni di *membership* ai vari gruppi multicast
 - *Locale* significa che i messaggi IGMP sono scambiati tra end-system e router
- I messaggi IGMP sono trasmessi in datagrammi IP il cui header ha:
il campo *protocol* con valore 2 ed il campo TTL con valore 1
- Tre versioni:
 - IGMP v1 definito in RFC 1112
 - IGMP v2 definito in RFC 2236
 - aggiunge messaggi di *leave group* e messaggi di *query group-specific*
 - IGMP v3 definito in RFC 3376
 - aggiunge messaggi per il *source-specific multicast*



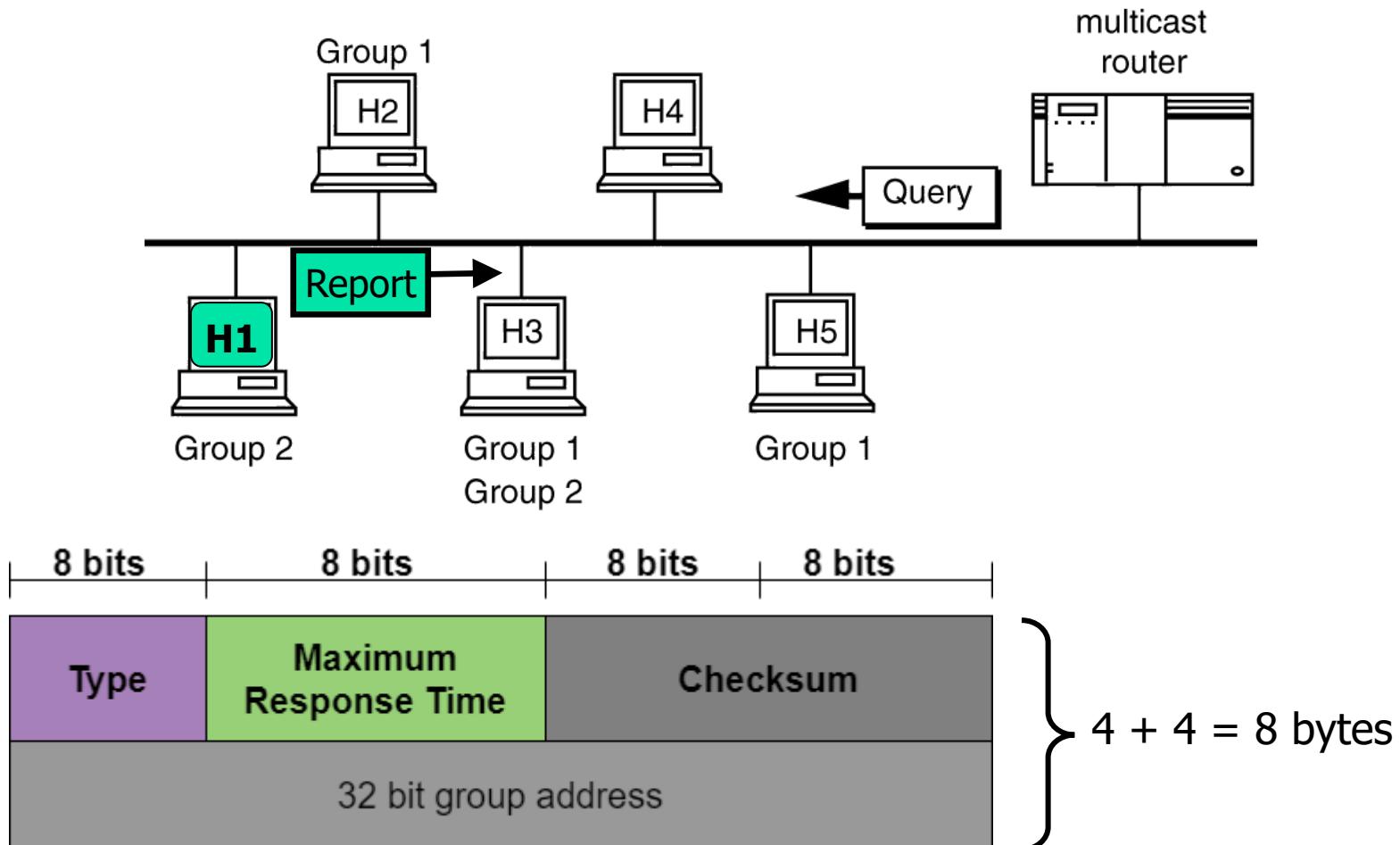


IGMP: funzionamento

- Quando un'applicazione chiede di entrare in un nuovo gruppo (*join*), l'host in cui gira l'applicazione invia un messaggio IGMP *membership report*
 - Il messaggio è detto *unsolicited* perché non è generato come risposta ad un messaggio di *membership query*
 - Gli eventuali multicast router presenti nella rete locale dell'host ricevono il messaggio e stabiliscono i meccanismi di routing atti a ricevere i pacchetti indirizzati al gruppo
- Dovendo gestire i gruppi in maniera dinamica, i multicast router interrogano periodicamente (*polling*) gli host presenti nelle reti locali ad essi collegate, inviando messaggi di *membership query* a cui gli host rispondono inviando messaggi di *membership report*



Il protocollo IGMPv2: formato messaggi





Indirizzi IP destinazione dei messaggi IGMP

IGMP message type	IGMP version	First octet value	Destination Address	
Membership Query (general)	ALL	0x11	224.0.0.1 ALL-SYSTEMS.MCAST.NET	Router asks all systems in the LAN to declare their memberships
Membership Query (group-specific)	v2, v3	0x11	group address G	Router queries for members of a group G in the LAN
Membership Report	v1	0x12	group address G	Join group G
Membership Report	v2	0x16	group address G	Join group G
Membership Report	v3	0x22	224.0.0.22 IGMP.MCAST.NET	Join group G
Source Specific Membership Report	v3	0x22	SSM group address	Join group G with source S
Leave Group	v1	N/A	N/A	N/A
Leave Group	v2	0x17	224.0.0.2 ALL-ROUTERS.MCAST.NET	Leave group G
Leave Group	v3	N/A	N/A	IGMPv3 uses membership reports instead of Leave msg



IGMP Report suppression

- Per evitare di sovraccaricare la rete con un elevato numero di messaggi di *membership report*, alla ricezione di un messaggio *membership query* per un gruppo multicast al quale l'host ha fatto *join*, l'host setta un timer ad un valore D scelto casualmente e si mette in attesa
 - Se prima dello scadere del timer, riceve un *membership report* inviato da un altro host, l'host annulla il timer e si astiene dall'inviare il proprio report
 - Se invece il timer scade senza che siano arrivati altri report, l'host invia il proprio report
 - In questo modo, per ciascun gruppo attivo si genera un solo messaggio report
- Questo meccanismo esiste solo il IGMPv1 ed IGMPv2 ma non in IGMPv3

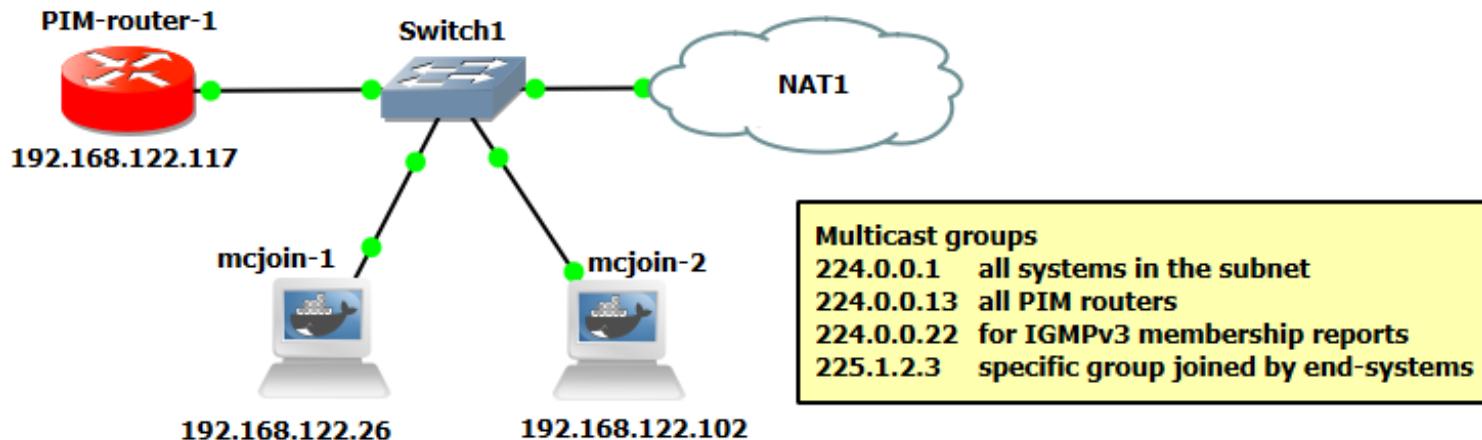


Source-specific multicast (SSM)

- Definito in RFC 4607 (anno 2006)
- Il modello di *Source-Specific Multicast* (SSM) prevede la possibilità per un host di manifestare l'intenzione a ricevere pacchetti inviati ad un gruppo G da una specifico host sorgente S attraverso un'operazione di source-specific join realizzata mediante messaggi IGMPv3 di *source-specific membership report*
- Il modello definito in RFC 1112 è indicato *Any-Source Multicast* (ASM)
- Indirizzi di gruppo IPv4 riservati all'uso SSM:
 - 232/8 (232.0.0.0 to 232.255.255.255)
- Indirizzi di gruppo IPv6 riservati all'uso SSM:
 - Indirizzi definiti dal prefisso FF3x::/32



IGMPv3: esempio



- Un router multicast PIM invia periodicamente messaggi IGMPv3 di membership query generali al gruppo 224.0.0.1
- Due end system fanno il join al gruppo 225.1.2.3 inviando periodicamente messaggi IGMPv3 di membership report al gruppo 224.0.0.22



IGMPv3: Membership query generale

Screenshot of Wireshark showing an IGMPv3 membership query and report.

Table: Captured Packets

Time	Source	Destination	Protocol	Length	Info
30.353496	192.168.122.117	224.0.0.1	IGMPv3	50	Membership Query, general
35.149224	192.168.122.102	224.0.0.22	IGMPv3	54	Membership Report / Join group 225.1.2.3 for any source...

Frame 9: 50 bytes on wire (400 bits), 50 bytes captured (400 bits) on interface -, id 0

Ethernet II, Src: 4a:9a:34:5a:8e:d1 (4a:9a:34:5a:8e:d1), Dst: IPv4mcast_01 (01:00:5e:00:00:01) (highlighted by green oval)

Internet Protocol Version 4, Src: 192.168.122.117, Dst: 224.0.0.1 (highlighted by red oval)

Internet Group Management Protocol

[IGMP Version: 3]

Type: Membership Query (0x11)
Max Resp Time: 10,0 sec (0x64)
Checksum: 0xec8f [correct]
[Checksum Status: Good]
Multicast Address: 0.0.0.0 (highlighted by red oval)

.... 0... = S: Do not suppress router side processing
.... .010 = QRV: 2
QQIC: 12
Num Src: 0

Hex View: 0020 00 01 94 04 00 00 11 64 ec 8f 00 00 00 00 02 0cd

Statistics: Max Response Time (igmp.max_resp), 1 byte | Pacchetti: 24 · visualizzati: 24 (100.0%) | Profilo: Default

- Il messaggio di membership query generale è inviato dal router PIM (192.168.122.117) al gruppo 224.0.0.1
- L'indirizzo multicast della query generale è 0.0.0.0



IGMPv3: Membership report

Screenshot of Wireshark showing an IGMPv3 Membership Report packet.

The packet details:

- Time: 37.450041
- Source: 192.168.122.26
- Destination: 224.0.0.22
- Protocol: IGMPv3
- Length: 54
- Info: Membership Report / Join group 225.1.2.3 for any source...

The packet structure shows:

- Frame 11: 54 bytes on wire (432 bits), 54 bytes captured (432 bits) on interface -, id 0
- Ethernet II, Src: 86:95:0a:9a:da:53 (86:95:0a:9a:da:53), Dst: IPv4mcast_16 (01:00:5e:00:00:16)
- Internet Protocol Version 4, Src: 192.168.122.26 Dst: 224.0.0.22
- Internet Group Management Protocol
 - [IGMP Version: 3]
 - Type: Membership Report (0x22)
 - Reserved: 00
 - Checksum: 0xf8f9 [correct]
 - [Checksum Status: Good]
 - Reserved: 0000
 - Num Group Records: 1
 - Group Record : 225.1.2.3 Mode Is Exclude
 - Record Type: Mode Is Exclude (2)
 - Aux Data Len: 0
 - Num Src: 0
 - Multicast Address: 225.1.2.3

Hex dump:

```
0000  01 00 5e 00 00 16 86 95  0a 9a da 53 08 00 46 c0  ...^..... .S..F.
```

Protocol: Internet Group Management Protocol

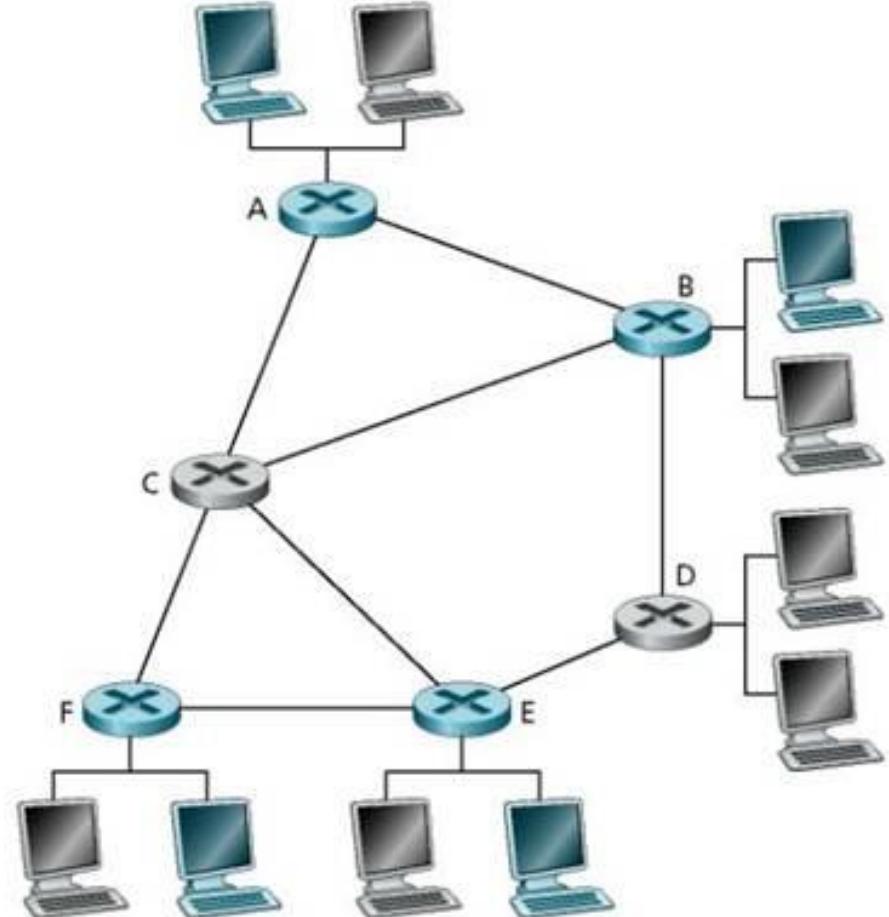
Pacchetti: 24 · visualizzati: 24 (100.0%)

Profilo: Default

- Il messaggio di membership report è inviato da un end-system (192.168.122.25) al gruppo 224.0.0.22
- Il campo Record-Type è "Mode Is Exclude" con numero di sorgenti zero



Il routing multicast e l'albero di copertura

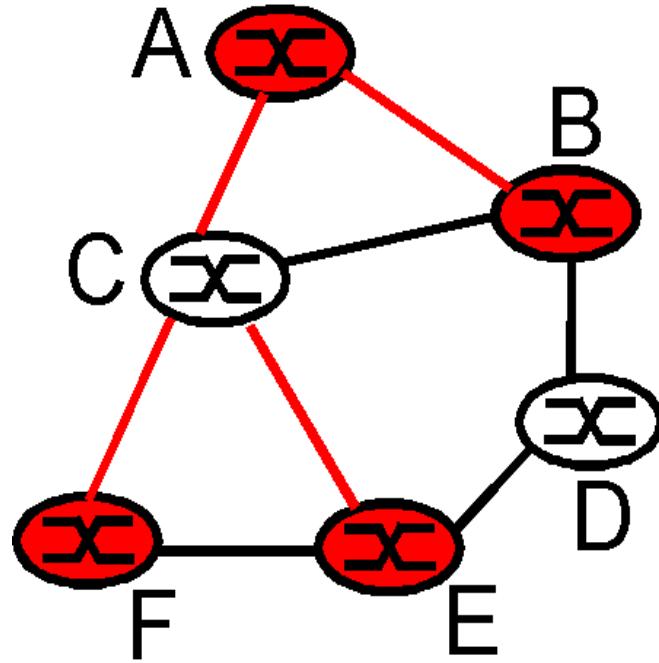


L'obiettivo dell'instradamento multicast è di trovare un albero di link che colleghi tutti i router cui sono attaccati gli host che appartengono al gruppo multicast. I pacchetti multicast saranno allora instradati attraverso questo albero dal sender a tutti gli host appartenenti all'albero multicast. Naturalmente, l'albero può contenere router che non hanno collegati host appartenenti al gruppo multicast

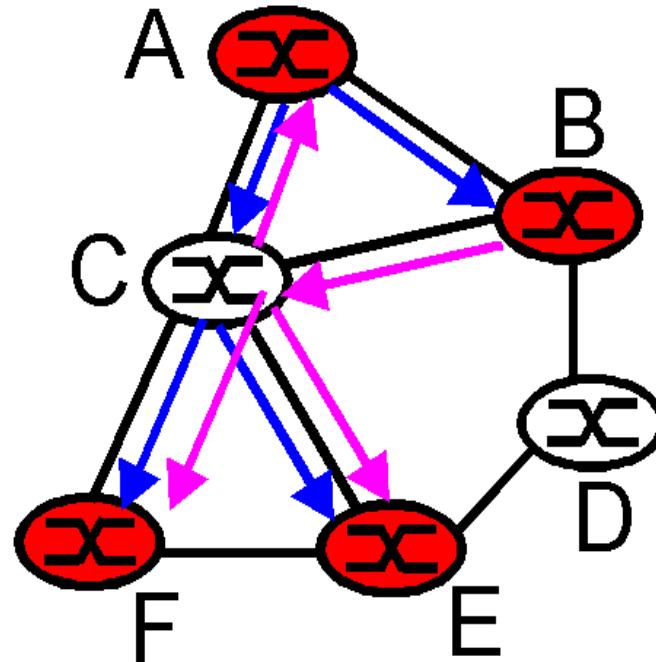


Il routing multicast: due possibili approcci

Group-shared Tree



Source-based Tree

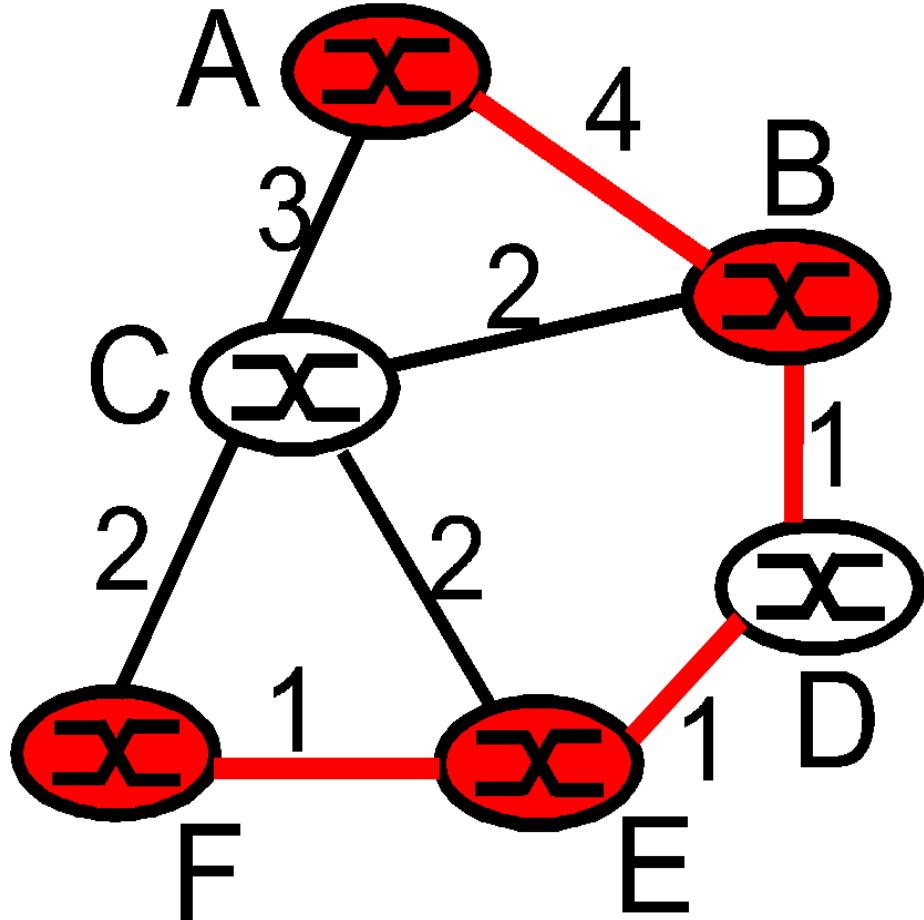


Unico albero condiviso dal gruppo
per distribuire il traffico
di tutti i mittenti

Albero d'instradamento costruito
specificamente
per ciascun singolo mittente



Group-shared Tree



Steiner Tree Problem:
il problema di trovare un
albero a costo minimo



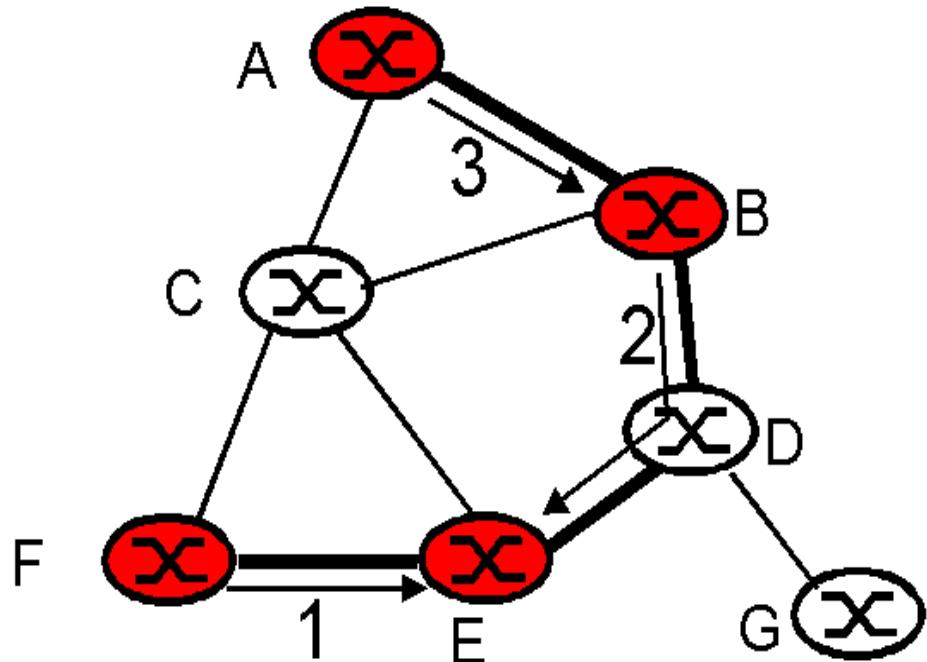
- Problema NP-completo
- Esistono algoritmi che approssimano la soluzione ottima



Instradamento multicast con albero condiviso dal gruppo

Approccio core-based:

N.B.: Il “core” è il nodo E



Legend

- router with attached group member
- router with no attached group member
- path/order in which join msgs generated

- Si usa un approccio centralizzato definendo un nodo centrale come **punto di rendezvous** o **nucleo**
 - Tutti i router cui sono collegati host multicast aderiscono al nucleo con messaggi unicast
 - Il percorso seguito definisce il ramo dell’albero

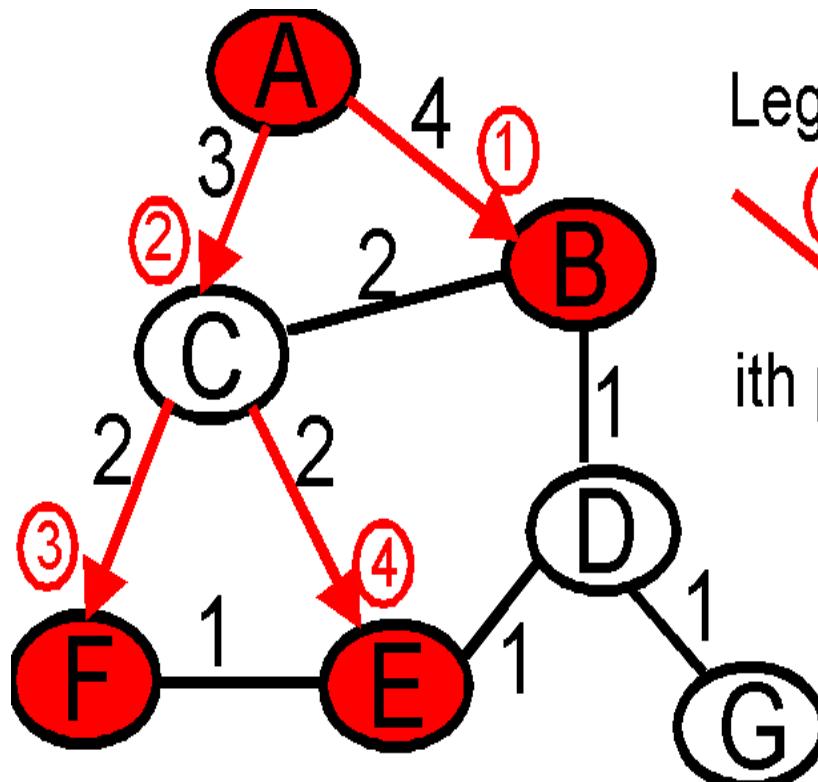


Source-based Tree

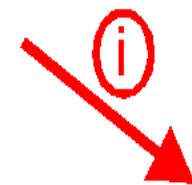
Least unicast-cost

path tree:

l'unione dei percorsi minimi dalla sorgente a tutte le destinazioni



Legend:



ith path link to be added



Differenze tra gli approcci al routing multicast

**Least unicast-cost
path tree**



**Minimizza il costo
dalla sorgente ad
ognuna delle
destinazioni**

Steiner tree

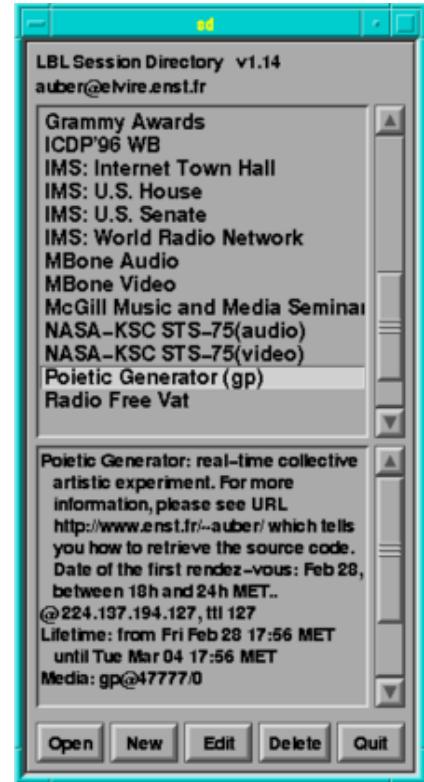


**Minimizza la
somma dei costi
dei link dell'albero
multicast**



La rete MBone: Multicast BackBone

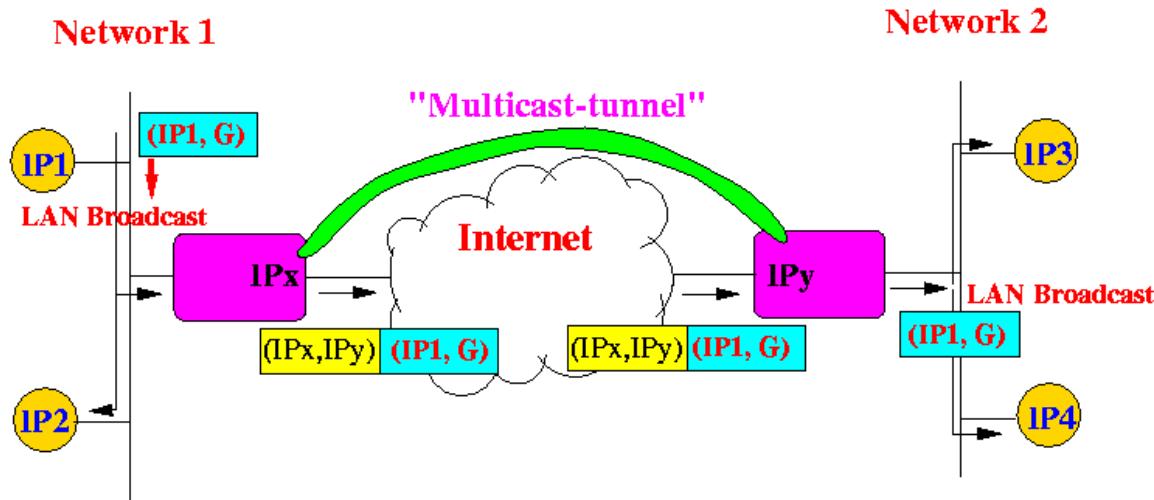
- Un banco di prova semi-permanente utilizzata in passato per il testing delle applicazioni e dei protocolli di routing multicast
- Una rete virtuale (*overlay*) composta da “isole” capaci di supportare il forwarding di pacchetti multicast IP collegate mediante link virtuali di tipo punto-punto chiamati “tunnel”
- Usata per la trasmissione multicasti di flussi multimediali descritti tramite il servizio “SDR” (session directory) basato sulla trasmissione multicast di annunci di sessione in formato SDP





MBone : i tunnel multicast

- I pacchetti IP multicast vengono incapsulati prima di essere trasmessi attraverso i tunnel, in modo da apparire, all'esame dei router e delle sottoreti intermediarie, come normali datagrammi unicast

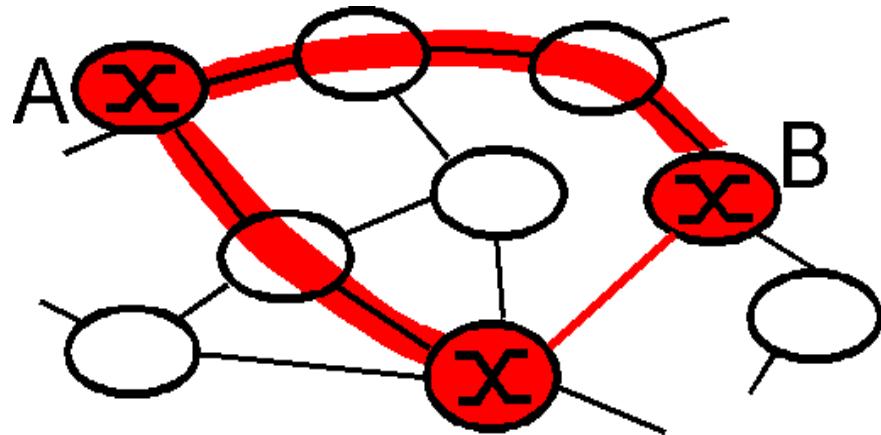


- Un router multicast intenzionato a trasmettere un pacchetto all'altro capo di un tunnel deve aggiungere ad esso un ulteriore header IP in cui sia presente, come indirizzo destinazione, l'indirizzo unicast del router che si trova al capo opposto del tunnel (IP multicast in IP unicast)

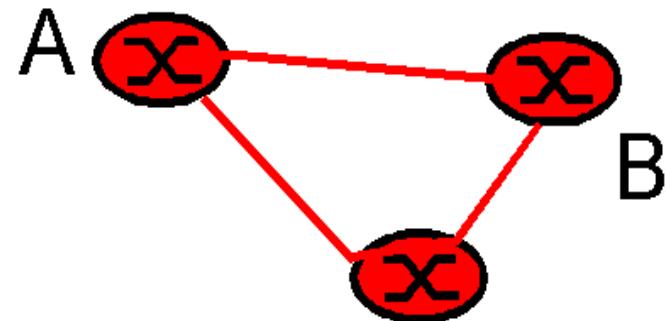


MBone : i tunnel multicast (segue)

- Il router situato all'altro estremo del tunnel deve, alla ricezione del pacchetto, eliminare l'header unicast che fungeva da capsula e smistare il pacchetto multicast nel modo appropriato



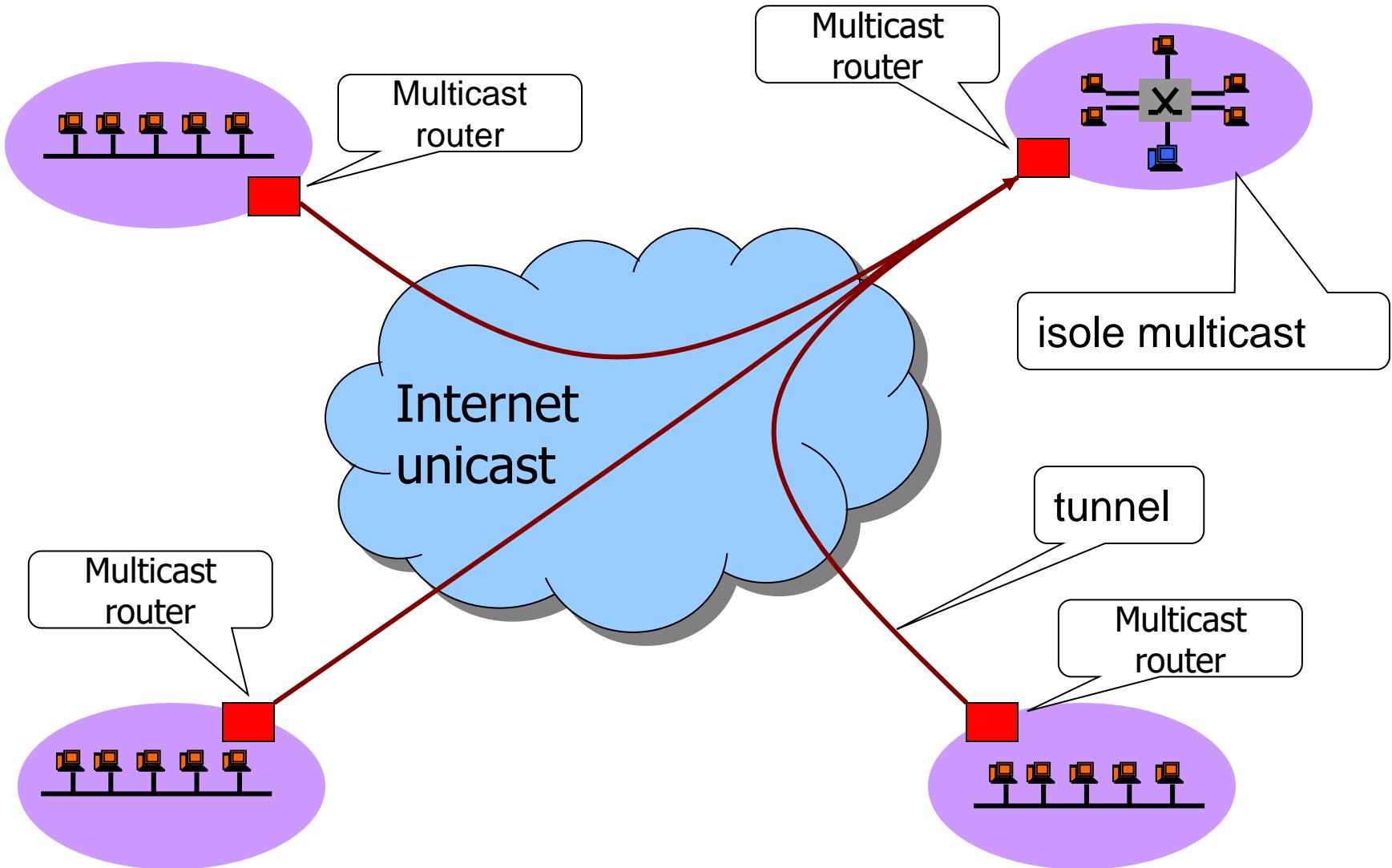
physical topology



logical mcast topology



MBone





IGMP snooping

- Gli switch di rete (operando a livello 2) di default trattano i pacchetti multicast alla stregua di pacchetti broadcast e li inoltrano in *flooding* su tutte le porte con sovraccarico della rete
- La funzionalità di **IGMP snooping** consente agli switch di "imparare" (ascoltando i messaggi IGMP provenienti dagli host) quali porte devono essere interessate alla trasmissione multicast per ciascun gruppo
- Questa funzionalità richiede che lo switch sia in grado di elaborare il payload della PDU di livello 2 dal momento che IGMP è un protocollo di livello 3

