

Corso di Laurea in Ingegneria Informatica



Corso di Reti di Calcolatori I

Roberto Canonico (roberto.canonico@unina.it)

Giorgio Ventre (giorgio.ventre@unina.it)

IP multicasting
Multicast routing

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 Canonico, Giorgio Ventre

Reverse path forwarding



- Utilizza le informazioni di routing unicast già note
- Costruisce un nuovo albero ad ogni pacchetto
- Non tiene in considerazione i gruppi,
 - questo lo rende più un algoritmo per la costruzione di un "broadcast tree"

3

Il Reverse Path Forwarding (RPF)



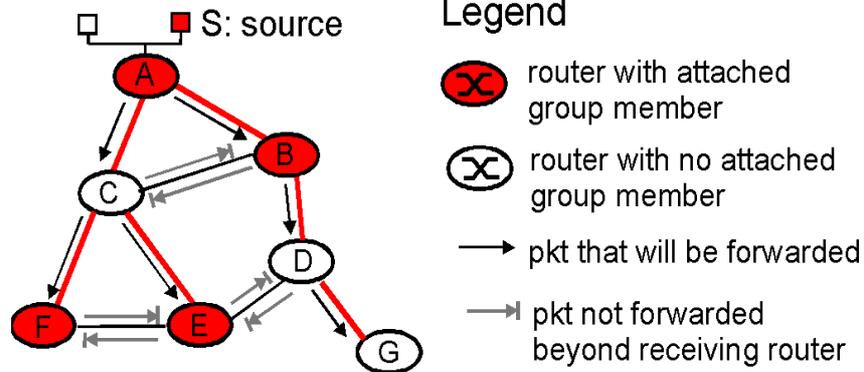
**Reverse Path
Forwarding (RPF)**



Quando un router riceve un pacchetto multicast con un dato indirizzo sorgente, lo trasmette su tutte le interfacce di uscita solo se il pacchetto è giunto da un link appartenente al proprio shortest path verso il sender in questione

4

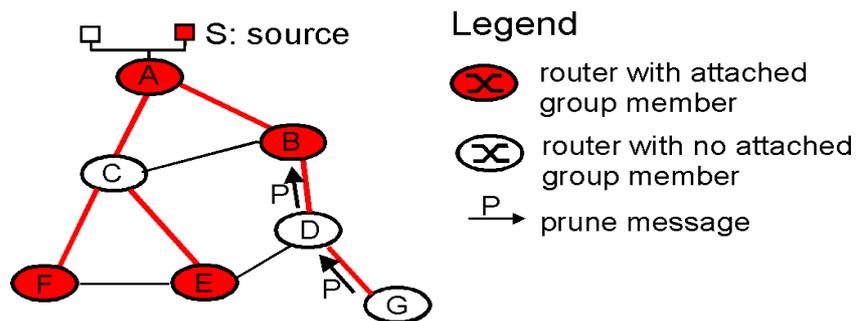
Il Reverse Path Forwarding



N.B.: In tal caso, D inoltra il pacchetto verso G, nonostante G non appartenga al gruppo multicast

5

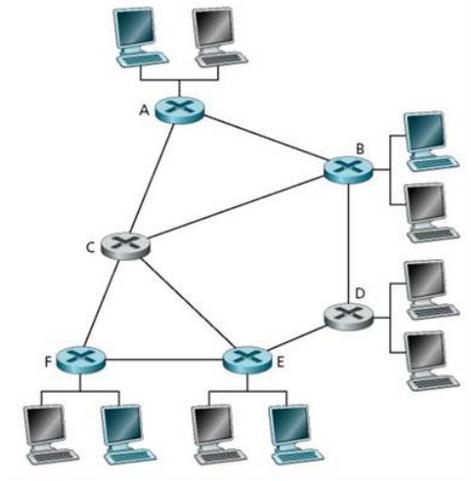
Il Truncated Reverse Path Forwarding



I router che ricevono pacchetti multicast pur non essendo connessi ad host appartenenti al gruppo destinazione, inviano un apposito messaggio di “pruning” verso il router a monte. Un router che riceve tale messaggio da tutti i suoi successori, itera il procedimento

6

Il routing multicast e l'albero di copertura



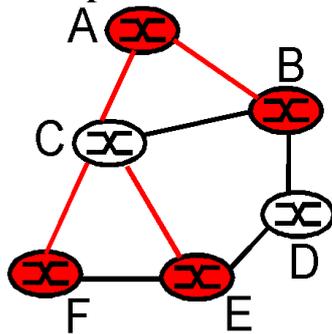
L'obiettivo dell'instradamento multicast è di trovare un albero di link che colleghi tutti i router cui sono attaccati gli host che appartengono al gruppo multicast. I pacchetti multicast saranno allora instradati attraverso questo albero dal sender a tutti gli host appartenenti all'albero multicast.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

7

Il routing multicast: due possibili approcci

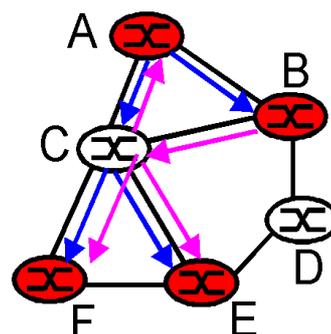


Group-shared Tree



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

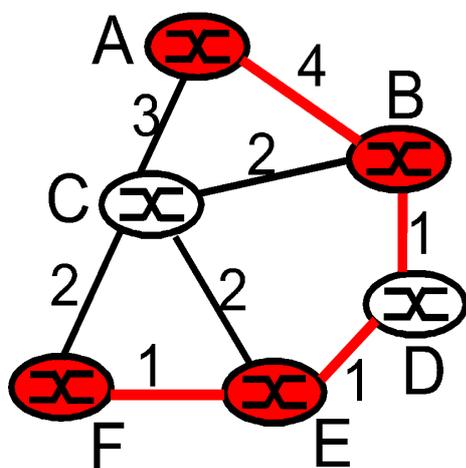
Source-based Tree



Specifico albero d'instradamento costruito per ciascun singolo mittente

8

Group-shared Tree



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



Tale problema è
NP-completo.
Esistono, tuttavia, algoritmi
che approssimano la
soluzione ottimale in
maniera soddisfacente

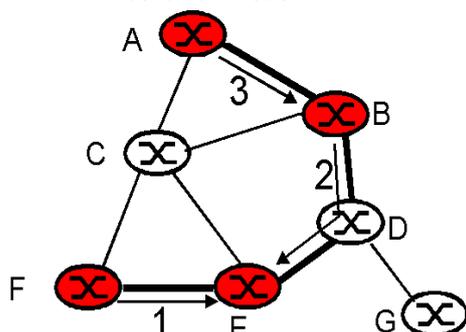
9

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

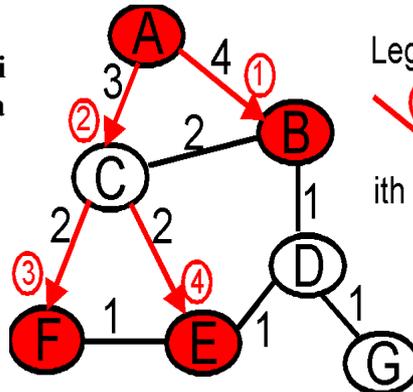
10

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

11

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

12

Protocolli per il multicast in Internet : DVMRP



Distance Vector Multicast Routing Protocol



Utilizza un algoritmo Distance Vector che permette ad ogni router di calcolare il link di uscita sul percorso minimo verso ciascuna possibile sorgente

Implementa un algoritmo source-based tree con RPF, pruning e grafting (innesto)

Pruning: Se non ci sono iscrizioni al gruppo indicato viene mandato un messaggio **prune** ai router limitrofi

Tale messaggio **non** viene mandato sull'interfaccia corrispondente al percorso verso la sorgente, a meno che lo stesso messaggio non venga **ricevuto su tutte le altre interfacce**

13

Protocolli per il multicast in Internet : MOSPF



Multicast Open Shortest Path First



Estende OSPF facendo sì che i router si scambino anche le informazioni relative all'appartenenza ai gruppi

In tal modo, i router possono costruire alberi specifici per ogni sorgente, pre-potati, relativi ad ogni gruppo multicast

14



Core-Based Tree



Costruisce un albero "group-shared" bidirezionale, con un unico centro ("core")

- L'aggiunta di rami avviene mediante appositi messaggi di "join"
- La gestione dell'albero è affidata a meccanismi di refresh (soft-state)

15



Protocol Independent Multicast



E' in grado di funzionare utilizzando un qualunque protocollo di routing unicast come supporto
Puo' essere utilizzato per routing sia inter-domain che extra-domain
Se usato come extra-domain allora riesce ad interoperare con qualsiasi altro protocollo usato inter-domain

Prevede due scenari alternativi:

- dense mode
- sparse mode

16

Il protocollo PIM



Dense mode



Dato che la maggior parte dei router è coinvolta nella trasmissione, utilizza un approccio RPF (simile a quello adottato da DVMRP)

DM: progettato ed ottimizzato per reti densamente popolate da utenti di un certo gruppo

Sparse mode



Utilizza un approccio center-based, in cui i router interessati alla trasmissione inviano messaggi espliciti di "join" (simile a CBT)

SM: progettato per reti scarsamente popolate

17

Il routing multicast inter-dominio



Per instradare datagrammi multicast tra differenti Sistemi Autonomi (Autonomous Systems -- AS)



DVMRP:
standard de facto



BGMP:
Border Gateway Multicast Protocol
•approccio group-shared
•in corso di sviluppo

18

La rete MBone: Multicast BackBone



- Un banco di prova semi-permanente per il multicast
- Una rete virtuale che si appoggia su porzioni dell'Internet fisica
- Composta da “isole” capaci di supportare il multicast IP (es: reti locali dotate di meccanismi hardware per il multicasting, quali Ethernet), collegate mediante link virtuali di tipo punto-punto chiamati “tunnel”

19

MBone : i tunnel multicast



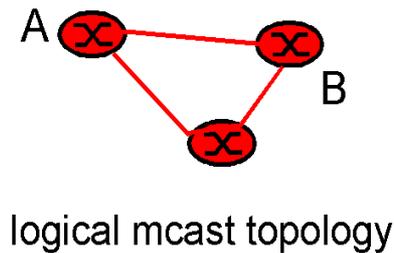
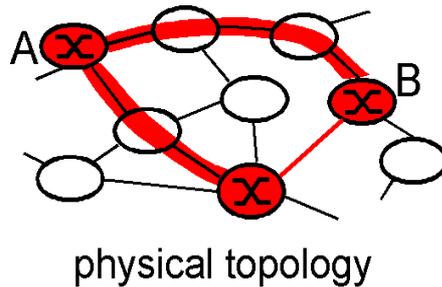
- I pacchetti IP multicast vengono incapsulati prima di essere trasmessi attraverso i tunnel, in modo da apparire, all'esame dei router e delle sottoreti intermedie, come normali datagrammi unicast
- Un multicast router intenzionato a trasmettere un pacchetto all'altro capo di un tunnel deve aggiungere 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

20

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



21

Il programma Mrouterd



- Si occupa del routing multicast sui sistemi UNIX
- Il suo funzionamento è del tutto simile a quello del demone routed nel caso unicast:
 - opera in stretta collaborazione col sistema operativo per installare le informazioni relative all'instradamento dei pacchetti multicast
 - Può essere utilizzato solo con una versione speciale di UNIX, conosciuta come "multicast kernel", contenente:
 - una tabella apposita per il routing dei pacchetti multicast
 - il codice necessario per il loro smistamento

22

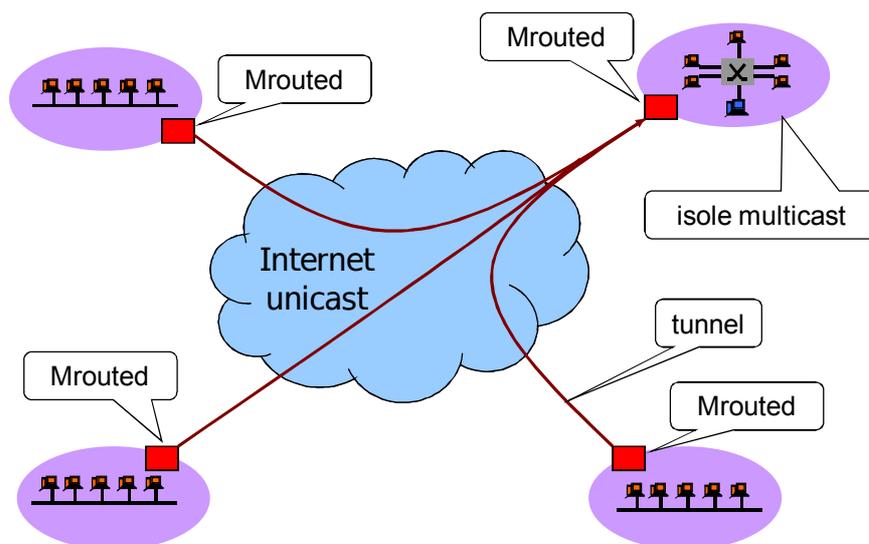
Mrouted : funzioni principali

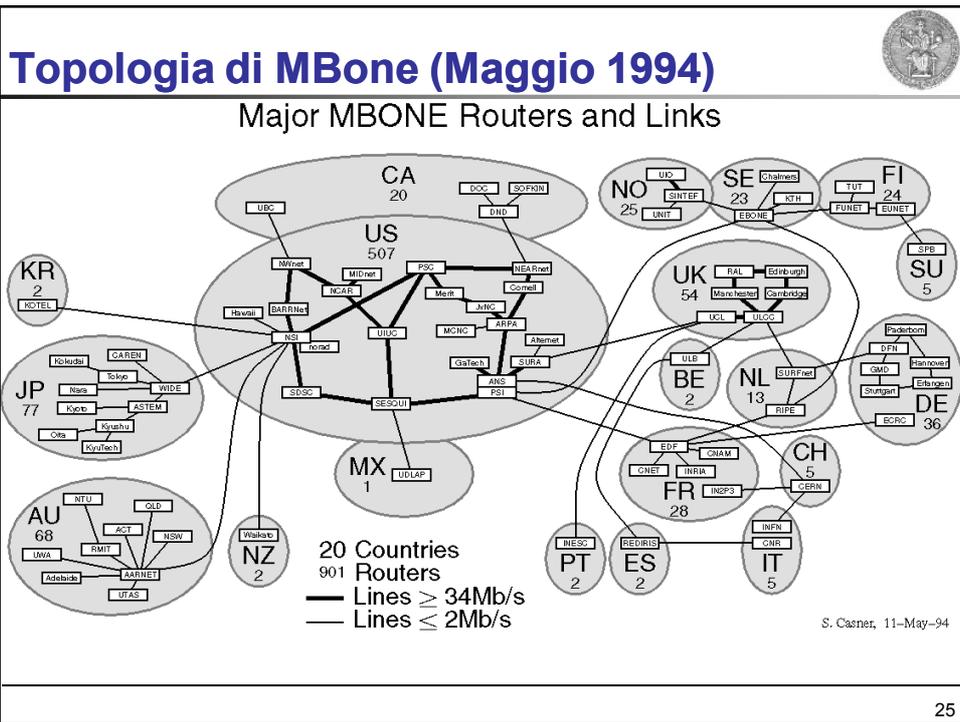


- Propagazione delle informazioni relative al routing :
 - *mrouted* utilizza DVMRP per propagare tali informazioni
 - un calcolatore che supporta *mrouted* è anche in grado di costruire la “multicast routing table”:
 - Impiego di algoritmi quali il Truncated Reverse Path Broadcast (TRPB)
- Creazione dei tunnel multicast:
 - non tutti i router di Internet sono capaci di smistare i datagrammi di tipo multicast
 - *mrouted* si occupa, quindi, della configurazione di un tunnel tra due router, attraverso elementi intermediari che non partecipano al multicasting

23

MBone





25

Mbone: assenza di meccanismi di controllo

Date: Tue, 12 Dec 1995 08:46:36 PST
 From: Van Jacobson <van@EE.LBL.GOV>
 Subject: fh-koeln.de sending 400kb/s of nv screendump
 X-To: mbank@linus.nz.fh-koeln.de, protsche@grisu.nz.fh-koeln.de
 X-Cc: mbone-eu@sics.se
 Content-Length: 543

You may not be aware that the nv screendump video that you are currently sending to Mbone Video session is going to over 20,000 machines all over the world. Your two streams were using 400kb/s of bandwidth. The total bandwidth available to the Mbone is only 500kb/s so you were using all of it. The mbone is pretty saturated this week with (scheduled) traffic from the WWW conference and extra traffic that pushes it above the 500kb/s limit causes packet losses for everyone. Please don't send unnecessary video. Thanks.

- Van Jacobson

26

Mbone tools: SDR – Session Directory



The screenshot displays the SDR software interface. On the left, a window titled 'sdr:elkner@anubis' shows a list of public sessions, including 'CBC Radio On-Line Test', 'Computer Stamp: First Day of Issue', 'FreeBSD Lounge', 'IRB Mbone Team', 'MBone RTP Audio', 'Radio Free Vat', 'Test of Win32 SD/SDR', 'The Stanford Channel', 'UMich IRL (private)', 'USC- CS dgroup VR conference room (private)', and 'UCL Session Directory v2.2a22'. The 'UCL Session Directory v2.2a22' session is selected.

In the center, a 'Sdr: Create New Session' dialog box is open, showing a list of media types to be configured:

<input checked="" type="checkbox"/>	audio	PCM	1	<input type="checkbox"/>
<input checked="" type="checkbox"/>	video	H.261	1	<input type="checkbox"/>
<input checked="" type="checkbox"/>	whiteboard	wb	1	<input type="checkbox"/>
<input checked="" type="checkbox"/>	text	UCL NTE	1	<input type="checkbox"/>
<input checked="" type="checkbox"/>	directory			<input type="checkbox"/>

 Buttons for '<< Back', 'Next >>', 'Accept', and 'Cancel' are at the bottom.

At the bottom, the 'Sdr: Session Information' window is visible, showing session details for 'GMU Netlab Test'. It includes encryption and authentication settings, session timing ('from 15:00 to 17:00 Eastern this afternoon'), contact details, and a table of media configurations:

audio	Format: PCM	Proto: RTP	Addr: 224.2.219.3	Port: 30752	TTL: 63	Key:	Vars: jptime:40
video	Format: H.261	Proto: RTP	Addr: 224.2.237.87	Port: 63806	TTL: 63	Key:	
whiteboard	Format: wb	Proto: judp	Addr: 224.2.224.79	Port: 46017	TTL: 63	Key:	
text	Format: UCL	Proto: judp	Addr: 224.2.181.91	Port: 64440	TTL: 63	Key:	

 Buttons for 'Join', 'Invite', 'Record', 'Edit', 'Delete', and 'Dismiss' are at the bottom.

27

Mbone tools: VIC (video) e RAT (audio)



The screenshot shows the VIC and RAT software interfaces. On the left, the 'vic:vic menu' window is open, displaying transmission and encoder settings. The 'Transmission' section includes 'Transmit' (checked), 'Rate Control' (14 f/s, 417 kb/s), and 'Release' (604 kbps, 15 fps). The 'Encoder' section shows device, port, signal, and options settings. The 'Display' section includes 'Options...', 'Title...', and 'External' settings. The 'Session' section shows destination, name, and key information.

In the center, the 'GMU Netlab Test' window is open, showing a video feed of a person and audio controls (mute, color, info...). The window title is 'VIC v2.8ucl-1.0'.

On the right, the 'CBC Radio On-Line Test' window is open, showing a list of participants:

- Jens Elkner <elkner@irb.cs.uni-magdeburg.de>
- Elizabeth Gentile
- K.Takagishi(Miyazaki Univ.)
- Qin Li (University of Leeds)
- Michael Baldizzi (NASA LeRC)
- Jim Thornton (Xerox PARC)
- CBC Radio On-Line
- lethal
- Mitsuru Maruyama
- Mitch Collinsworth (Cornell PCG)
- Henning Schulzrinne
- CT Chen
- pearl.dct.nih.gov

 The window also features volume controls and a 'RAT v2.6a11' logo.

28