

5. Instradamento



Complementi di Reti e Sistemi di Telecomunicazioni

5. Instradamento

Complementi di Reti e Sistemi C. Nobile

5.1

Sotto-rete e super-rete

- Anni '80 indirizzamento sotto-rete (subnetting):
 - Suddivisione di una rete di classe in più sottoreti
- Anni '90 indirizzamento senza classi:
 - Supera la suddivisione in classi permettendo una suddivisione arbitraria tra prefisso e suffisso

Complementi di Reti e Sistemi C. Nobile

5.2

Reti fisiche e reti logiche

- Punto-punto
 - le interfacce possono essere “*unnumbered*” (es.: linee dedicate o dial-up)
- Multiaccesso con possibilità di *broadcast*
 - gli host possono comunicare direttamente senza passare per router intermedi (es. : le LAN)
- Multiaccesso senza possibilità di *broadcast*
 - gli host possono comunicare direttamente senza passare per router intermedi (es. : reti a pacchetto commutate)

Complementi di Reti e Sistemi C. Nobile

5.3

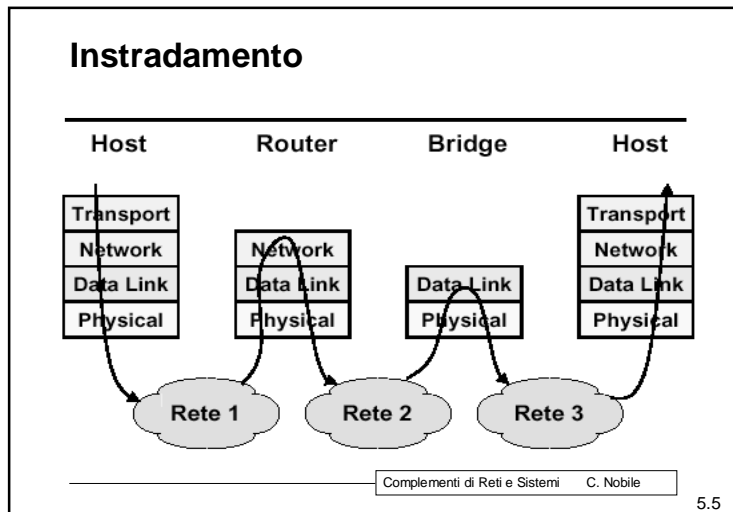
Instradamento

- L'instradamento viene quindi realizzato:
 - all'interno di una rete logica, in modo implicito, direttamente dalla stazione sorgente
 - » ossia deve essere realizzato dal livello 2 e quindi si deve fornire la “traduzione” dell'indirizzo IP di destinazione in indirizzo di livello due (ARP-RARP)
 - Tra reti logiche diverse è gestito esplicitamente dai router, ossia è il router che deve ricevere il pacchetto e che quindi instradarlo verso la sottorete opportuna
 - » La stazione sorgente deve quindi indirizzare il pacchetto al router, questo implica che:
 - Deve essercene almeno uno direttamente connesso alla rete fisica
 - La stazione deve conoscerne l'indirizzo (default router)

Complementi di Reti e Sistemi C. Nobile

5.4

5. Intradamento



Instradamento nei router

- Nei router l'intradamento è realizzato tramite una tabella di instradamento (*Routing Table*, RT) del tipo:

Destinazione	Indirizzo di invio	Interfaccia
200.2.1.0	Invio diretto	interf. 1
205.1.4.0	Invio diretto	interf. 2
200.3.1.0	200.2.1.253	interf. 3
200.4.1.0	205.1.4.252	interf. 2
default	205.1.4.254	interf. 1

Complementi di Reti e Sistemi C. Nobile 5.6

Instradamento nei router

- Due sono quindi gli aspetti di cui si compone l'instradamento:
 - Esecutivo: la scelta della direzione di uscita tramite la tabella
 - Algoritmico: la compilazione/aggiornamento della tabella
- Il secondo aspetto si realizza tramite
 - il calcolo del percorso migliore eseguito secondo un qualche algoritmo
 - Lo scambio di informazioni fra i router per eseguire tale calcolo

Complementi di Reti e Sistemi C. Nobile 5.7

Instradamento nei router

- Per rendere l'instradamento efficiente si deve mantenere le RT di piccole dimensioni.
- Tabelle grandi:
 - Richiedono più tempo per l'individuazione della corretta direzione di uscita (*next hop*)
 - Sono di difficile gestione in fase di calcolo e di aggiornamento.
- La suddivisione net e host, crea una gerarchia che ha l'obiettivo di ridurre la dimensione delle RT.
- Lo stesso vale per la presenza del "default router"

Complementi di Reti e Sistemi C. Nobile 5.8

5. Instradamento

Routing Table

- Tipiche informazioni contenute nelle RT per ciascuna delle reti destinazione sono
 - Indirizzo della rete destinazione
 - Maschera di *subnet*
 - Indirizzo IP del successivo *router* da attraversare (*next hop*) o sul fatto che la destinazione è direttamente raggiungibile
 - Porta di uscita del "*next hop*"
 - Metrica (anche più di una)
 - Identificatore della sorgente dell'instradamento (manuale, locale, ICMP, uno degli algoritmi di instradamento)

Complementi di Reti e Sistemi C. Nobile

5.9

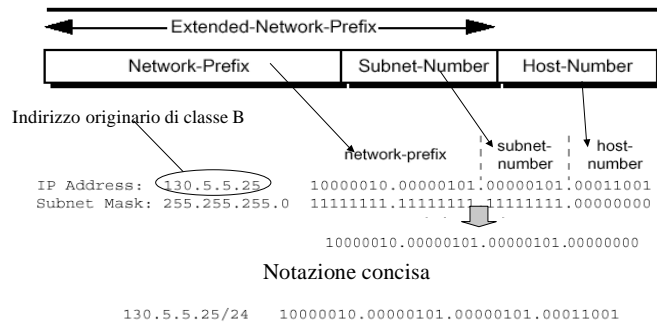
Proxy ARP

- E' una tecnica che permette la corrispondenza di più reti fisiche ad una sola rete o sottorete logica .
- Un *router* risponde agli ARP verso indirizzi IP che sa non appartenere a quella rete fisica con il proprio indirizzo di livello 2.
- L'appartenenza o meno ad una rete viene ricavata attraverso un *subnetting* che solo il *router* è obbligato a conoscere.

Complementi di Reti e Sistemi C. Nobile

5.10

Subnetting



Complementi di Reti e Sistemi C. Nobile

5.11

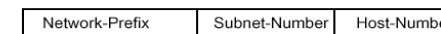
Subnetting

- Il *subnetting* aggiunge un livello di gerarchia all'indirizzamento, contribuendo a ridurre la dimensione delle RT.

Two-Level Classful Hierarchy



Three-Level Subnet Hierarchy



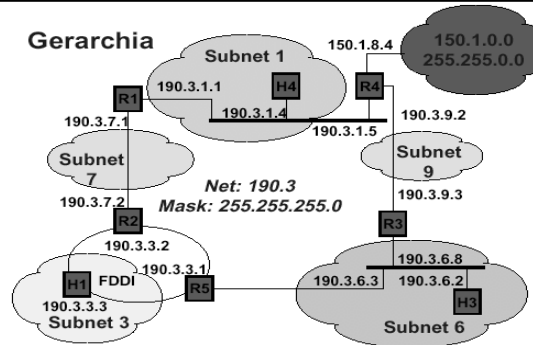
- All'interno di una classe di indirizzi la destinazione è ora individuata dalla coppia (*IP-address, Netmask*)

Complementi di Reti e Sistemi C. Nobile

5.12

5. Intradamento

Subnetting



5.13

Subnetting

- Il router R5 avrà una tabella del tipo

Destinazione	next hop
190.3.6.0/24	255.255.255.0
190.3.3.0/24	255.255.255.0
190.3.1.0/24	255.255.255.0
190.3.7.0/24	255.255.255.0
190.3.9.0/24	255.255.255.0
150.1.0.0/16	255.255.0.0
Default	190.3.6.8

Gli host dovranno conoscere il proprio ind. IP, la Netmask e l'indirizzo del router di default. ad es. per l'host H3 avrà:

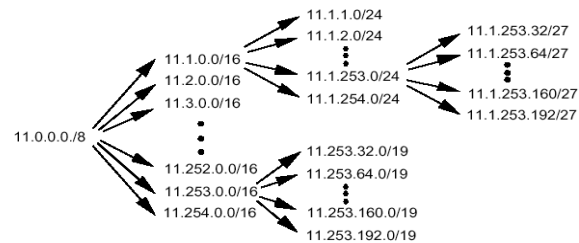
- Ind. IP = 190.3.6.2, NetMask = 255.255.255.0 (ff.ff.ff.0), default router = 190.3.6.8.

Complementi di Reti e Sistemi C. Nobile

5.14

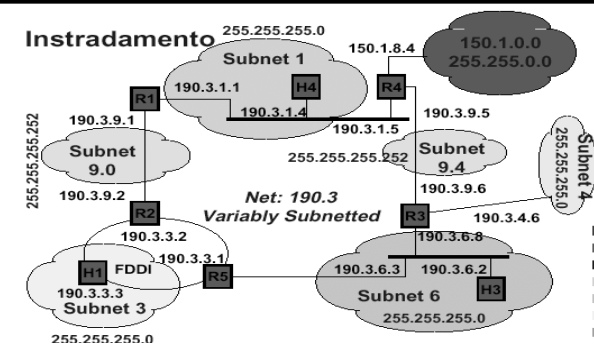
Variable Length Subnetting

- All'interno di una singola classe si possono realizzare gerarchie multiple:



5.15

Esempio



5.16

Variable Length Subnetting

- La maschera variabile permette di sfruttare meglio lo spazio di indirizzamento e ridurre ulteriormente le RT.
- In presenza di più di una scelta per una destinazione si sceglie quella con *subnet mask* più lunga
- Si consideri ad es. la RT di R5

Destinazione	Next hop
190.3.6.0/24 255.255.255.0	diretta
190.3.3.0/24 255.255.255.0	diretta
190.3.9.0/24 255.255.255.0	190.3.3.2
190.3.9.4/30 255.255.255.252	190.3.6.8
190.3.4.0/24 255.255.255.0	190.3.6.8
190.3.0.0/16 255.255.0.0	190.3.3.2
150.1.0.0/16 255.255.0.0	190.3.6.8
Default	190.3.6.8

Complementi di Reti e Sistemi C. Nobile

5.17

Classless InterDomain Routing (CIDR)

- La crescita degli utenti nella rete ha velocemente portato verso l'esaurimento lo spazio di indirizzamento disponibile.
- La ragione principale è legata al fatto che in molte situazioni le reti di classe C sono troppo piccole, quindi viene richiesto un indirizzo di classe B di cui però va sprecato gran parte dello spazio di indirizzamento.
- Per cui si è definito un meccanismo di "supernetting" che consiste nel accorpare indirizzi di classe C contigui in un unico spazio di indirizzamento creando suddivisioni netid-hostid ad hoc.

Complementi di Reti e Sistemi C. Nobile

5.18

Classless InterDomain Routing (CIDR)

- Il CIDR trasforma lo spazio di indirizzamento della classe C in un unico spazio "senza classe" che viene suddiviso usando come "quanti" le reti di classe C con un meccanismo di subnetting.
- I router in grado di gestire tale meccanismo, operano usando la coppia indirizzo IP - netmask per identificare il "next-hop"
- Quindi "annunciano" la coppia ed in presenza della netmask ignorano la definizione di classe C (effettuano un *supernetting*)

Complementi di Reti e Sistemi C. Nobile

5.19

Classless InterDomain Routing (CIDR)

- Si supponga che ad un ISP sia stato assegnato il blocco di indirizzi 206.0.64.0/18.
- Questo blocco rappresenta 16.384 IP *address* che possono essere interpretati come 64 reti da x/24.
- Se un cliente richiede 800 *host addresses*, invece che assegnargli una classe B (e perdere 64.700 indirizzi che non ha) o quattro Classi C (introducendo 4 nuove reti nelle *routing table* di Internet), l'ISP può assegnare al cliente il blocco 206.0.68.0/22, con 1.024 indirizzi IP

```
ISP: 206.0.64.0/18 11001110.00000000.01000000.00000000
Client: 206.0.68.0/22 11001110.00000000.01000100.00000000
Class C #0: 206.0.68.0/24 11001110.00000000.01000100.00000000
Class C #1: 206.0.69.0/24 11001110.00000000.01000101.00000000
Class C #2: 206.0.70.0/24 11001110.00000000.01000110.00000000
Class C #3: 206.0.71.0/24 11001110.00000000.01000111.00000000
```

Complementi di Reti e Sistemi C. Nobile

5.20

5. Intradamento

Classless InterDomain Routing (CIDR)

CIDR prefix-length	Dotted-Decimal	# Individual Addresses	# of Classful Networks
/13	255.248.0.0	512 K	8 Bs or 2048 Cs
/14	255.252.0.0	256 K	4 Bs or 1024 Cs
/15	255.254.0.0	128 K	2 Bs or 512 Cs
/16	255.255.0.0	64 K	1 B or 256 Cs
/17	255.255.128.0	32 K	128 Cs
/18	255.255.192.0	16 K	64 Cs
/19	255.255.224.0	8 K	32 Cs
/20	255.255.240.0	4 K	16 Cs
/21	255.255.248.0	2 K	8 Cs
/22	255.255.252.0	1 K	4 Cs
/23	255.255.254.0	512	2 Cs
/24	255.255.255.0	256	1 C
/25	255.255.255.128	128	1/2 C
/26	255.255.255.192	64	1/4 C
/27	255.255.255.224	32	1/8 C

Complementi di Reti e Sistemi C. Nobile

5.21

Intradamento

• Requisiti

- Minimizzare lo spazio occupato dalle RT per:
 - » Velocizzare la commutazione
 - » Semplificare i router (meno cari)
 - » Ridurre l'informazione necessaria all'aggiornamento
- Minimizzare il traffico di controllo
- Essere robusto, ossia evitare:
 - » Cicli
 - » Buchi neri
 - » Oscillazioni
- Ottimizzare i percorsi (dal punto di vista della distanza, del ritardo, del costo economico, ...)

Complementi di Reti e Sistemi C. Nobile

5.22

Intradamento: Alternative

- Centralizzato o **distribuito** (o isolato)
- Basato sulla sorgente o "**hop-by-hop**"
- **Deterministico** o stocastico
- **Singolo percorso** o multi-percorso
- Dipendente dallo stato (**dinamico**) o indipendente dallo stato (**statico**)
- **Distance Vector** o **Link State**.

Complementi di Reti e Sistemi C. Nobile

5.23

Distance vector

- L'algoritmo di Bellman Ford può essere realizzato in modo distribuito ed in questo caso viene chiamato **Distance Vector Routing**.
- Ogni nodo (router) conosce l'identità di tutti i nodi della rete e i nodi a lui direttamente connessi (vicini).
- Ogni nodo mantiene un **Distance Vector**, ossia una lista di coppie (destinazione, costo) per tutte le possibili destinazioni.
- Il costo è la somma stimata dei costi sui singoli link sul percorso "più corto" (*shortest path*) verso quella destinazione.
- Ogni nodo inizializza i costi relativi a destinazioni "lontane" ad un valore alto, convenzionalmente indicato infinito.

Complementi di Reti e Sistemi C. Nobile

5.24

5. Instradamento

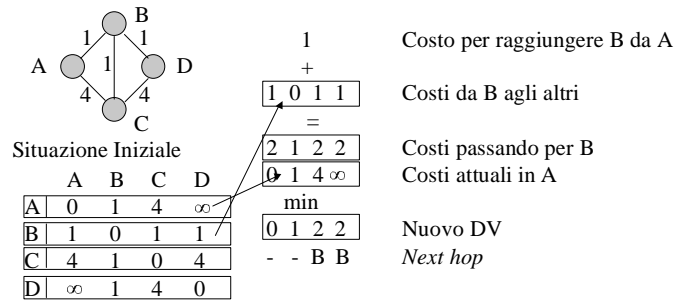
Distance vector

- Periodicamente ogni nodo invia ai propri vicini il proprio DV.
- Quando un *router* A riceve un DV da B (suo vicino), verifica quali sarebbero i costi per le varie destinazioni usando come transitio B; per le destinazioni in cui tali costi risultano minori di quelli attuali, sostituisce il costo vecchio con quello calcolato e lo stesso fa con il next-hop nella RT.

Complementi di Reti e Sistemi C. Nobile

5.25

Distance vector



Complementi di Reti e Sistemi C. Nobile

5.26

Distance vector

- Questo procedimento corrisponde a realizzare in modo distribuito e asincrono l'algoritmo di Bellman-Ford, perché ogni nodo i esegue l'iterazione

$$D_i \leftarrow \min_{j \in N(i)} \{d_{ij} + D_j\}$$
 (dove $N(i)$ è l'insieme dei nodi adiacenti ad i), usando le stime D_j più recenti ricevute dai vicini e trasmettendo D_i ai propri vicini.
- Si dimostra che non è necessaria una inizializzazione con particolari valori di D_j

Complementi di Reti e Sistemi C. Nobile

5.27

Distance vector

- Sia A il numero di archi e N quello dei nodi
- Nel caso peggiore, l'algoritmo di Bellman-Ford centralizzato compie $N-1$ iterazioni, ciascuna su $N-1$ nodi, con al più $N-1$ alternative per nodo, il che porterebbe a complessità $O(N^3)$.
 - Si può mostrare che la complessità è $O(mA)$, con m numero di iterazioni per la convergenza. Questo porta una complessità generalmente compresa fra $O(N^2)$ e $O(N^3)$.
 - Nel caso distribuito, se le iterazioni fossero eseguite in modo sincrono (simultaneamente ad ogni nodo), scambiando ad ogni iterazione i risultati con i vicini, partendo dalle condizioni iniziali $D_i^0 = \infty$ per tutti gli $i \neq 1$ e $D_1^0 = 0$, l'algoritmo convergerebbe in al più $N-1$ passi.

Complementi di Reti e Sistemi C. Nobile

5.28

5. Instradamento

Distance vector

“Le buone notizie viaggiano veloci”

via

Y	X	Z
a	X (4) 6	X (1) 6
	X (1) 6	X (1) 6
	X (1) 3	X (1) 3

via

Z	X	Y
a	X 50 (5)	X 50 (5)
	X 50 (5)	X 50 (2)
	X 50 (2)	X 50 (2)

tempo

Complementi di Reti e Sistemi C. Nobile 5.29

Distance vector

“Le cattive notizie viaggiano lente”

via

Y	X	Z
a	X (4) 6	X 60 (6)
	X 60 (6)	X 60 (6)
	X 60 (8)	X 60 (8)

via

Z	X	Y
a	X 50 (5)	X 50 (5)
	X 50 (5)	X 50 (7)
	X 50 (7)	X 50 (7)
	X 50 (9)	X 50 (9)

tempo

Complementi di Reti e Sistemi C. Nobile 5.30

Distance vector

- Questo tipo di algoritmo ha un problema legato all'aggiornamento che è chiamato **Count-to-infinity**:

Costo verso C Prossimo nodo

Iniziale

A	2	B
B	1	C

Si rompe BC

A	2	B
B	∞	-

Dopo il 1° scambio

A	∞	-
B	3	A

Dopo il 2° scambio

A	4	B
B	∞	-

⋮

A	∞	-
B	∞	-

Complementi di Reti e Sistemi C. Nobile 5.31

Distance vector

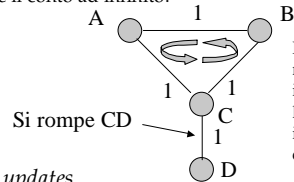
- Ci sono diverse possibili soluzioni al count to infinity**
 - Path vector**
 - oltre al costo si trasmette il percorso (*path-vector*), in questo modo i nodi possono capire quando non esiste più un percorso valido verso una certa destinazione. (BGP)
 - Split horizon**
 - Non viene passato il costo per una certa destinazione ad un vicino se questi è il *next hop* per quella destinazione.
 - Una versione più complessa detta *split horizon with poisonous reverse*, invece di non passare costi passa un costo infinito, questo a volte accelera la convergenza. (RIP)
 - Nel caso precedente, A non invia a B un costo (o lo invia infinito) verso C. Il ciclo quindi non si crea.

Complementi di Reti e Sistemi C. Nobile 5.32

5. Intradamento

Distance vector

- » Se ho ,più nodi coinvolti nella rottura direttamente non si riesce a bloccare il conto ad infinito.



B e A hanno sempre mandato a C un costo infinto verso D, ma non l'uno all'altro; questo innesca un ciclo che coinvolge A, B e C.

– Triggered updates

- » Normalmente, per evitare un numero eccessivo di aggiornamenti delle tabelle e di traffico di controllo, si limita il ritardo minimo fra due aggiornamenti consecutivi (per es. 30 sec.). Nel caso di collegamento caduto, gli aggiornamenti sono fatti immediatamente (riduce il tempo di convergenza). (RIP)

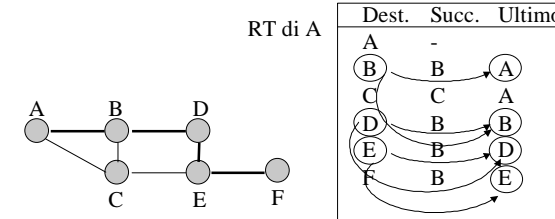
Complementi di Reti e Sistemi C. Nobile

5.33

Distance vector

– Source tracing

- » Insieme al costo i nodi si scambiano anche il nodo da attraversare immediatamente prima della destinazione. Con questa informazione aggiuntiva è possibile ricavare direttamente dalla tabella il percorso complessivo e quindi, mantenendo la RT più piccola si ottiene lo stesso del *path vector*.



Complementi di Reti e Sistemi C. Nobile

5.34

Link-state

- o La filosofia del **Link State (LS) routing** è quella di distribuire a tutti i nodi della rete l'intera sua topologia ed i costi di ogni *link* che la compone.
- o Con questa informazione ogni *router* è in grado di calcolarsi i propri percorsi ottimi verso ogni destinazione.
- o Se tutti vedono gli stessi costi e tutti usano lo stesso algoritmo, i percorsi saranno liberi da cicli.
- o Quindi sono due gli aspetti caratterizzanti questo metodo
 - Il modo in cui la topologia della rete viene diffusa fra i nodi.
 - Il modo in cui ogni nodo calcola i percorsi ottimi.

Complementi di Reti e Sistemi C. Nobile

5.35

Link-state

- o Nel caso LS ogni nodo applica l'algoritmo di **Dijkstra**.
- o A differenza dell'algoritmo di Bellman-Ford che itera sul numero di archi attraversati da un percorso, l'algoritmo di Dijkstra itera sulla lunghezza del percorso.
- o Nel caso peggiore la sua complessità è **$O(N^2)$** , in media si colloca intorno a **$O(A \log A)$** con A numero degli archi .

Complementi di Reti e Sistemi C. Nobile

5.36

5. Instradamento

Link-state

- o La distribuzione dell'informazione topologica (Link state packet) avviene attraverso un *Controlled Flooding*: un LSP "valido" viene propagato solo se proviene da un arco sul percorso a minima distanza verso la sorgente del LSP.
- o Validità di un LSP :Numero di sequenza
 - Per poter decidere se un LSP ricevuto è significativo (contiene una informazione più recente di quella attualmente nel nodo) ogni LSP deve contenere un numero di sequenza progressivo.
 - Il numero di sequenza ha valore locale per ogni tipo di LSP (identificato da coppia ordinata di nodi collegati da una linea)
 - Ogni volta che un nodo riceve un LSP più vecchio di quello in memoria, lo elimina senza propagarlo.

Complementi di Reti e Sistemi C. Nobile

5.37

Link-state

- o Le sequenze realmente utilizzabili sono di lunghezza finita e quindi soggetta ad "avvolgersi" (*wrapping*) bloccando l'aggiornamento.
- o *Wrapped sequence number*
 - Per evitare il problema si può prendere una sequenza molto grande (32 bit => 4.295.967.295) e decidere che quando due numeri distano troppo, il più piccolo sia anche il più giovane. Per esempio supponendo che N sia lunghezza della sequenza, allora a è più vecchio di b se
 - $a < b$ e $|b - a| < N/2$
oppure se
 - $a > b$ e $|b - a| \geq N/2$
- o La presenza di un numero di sequenza pone il problema dell'inizializzazione della sequenza quando un nodo si (ri)attiva.

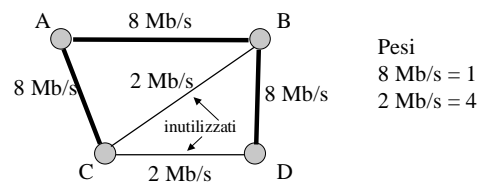
Complementi di Reti e Sistemi C. Nobile

5.38

Metriche e dinamica

• Metriche statiche

- In genere si tratta di valori inversamente proporzionali alla capacità del *link*. La staticità fa sì che le linee a minor velocità tendano ad essere sotto-utilizzate.



Complementi di Reti e Sistemi C. Nobile

5.39

Metriche e dinamica

• Dinamiche

- Le metriche dipendenti dal traffico sono sicuramente più efficaci, ma comportano alcuni problemi.
- Consideriamo la sperimentazione avvenuta su ARPAnet, dove in origine si era usata una metrica proporzionale alla lunghezza delle code di uscita dei *router*, per fare alcune osservazioni:
 - » La lunghezza (metrica) derivava da una media su un orizzonte (10 s). La durata dell'orizzonte è critica:
 - Corta: troppi transienti;
 - Lunga: rete converge lentamente;
 - La durata ottima non è omogenea sulla rete: dipende dalle capacità dei *link*

Complementi di Reti e Sistemi C. Nobile

5.40

5. Instradamento

Metriche e dinamica

- » La dinamica del costo non deve essere alta: altrimenti alcuni percorsi vengono completamente ignorati
- » La lunghezza della coda è usata come "predittore" della situazione futura del *link*: ma linee con code lunghe non verranno scelte nel futuro e quindi si "scaricheranno" (specialmente quelle ad alta capacità) e viceversa.
- » La mancanza di restrizioni fra valori successivi dei costi può generare oscillazioni significative.
- » Il ricalcolo quasi-sincrono delle tabelle tende a raccogliere traffico su alcune linee.

Complementi di Reti e Sistemi C. Nobile

5.41

Metriche e dinamica

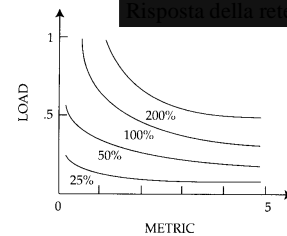
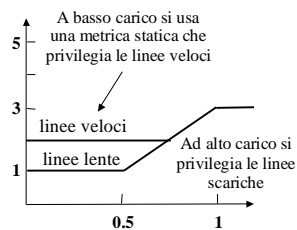
- La soluzione scelta per ARPAnet è stata:
 - Metrica mista capacità-coda (statica dinamica) dove a carico basso prevale la capacità, carico alto la coda.
 - Costi con una dinamica ridotta: valori da 1 a 3.
 - Massima variazione permessa fra due successivi ricalcoli: 1/2.

Complementi di Reti e Sistemi C. Nobile

5.42

Metriche e dinamica

Mappa della metrica



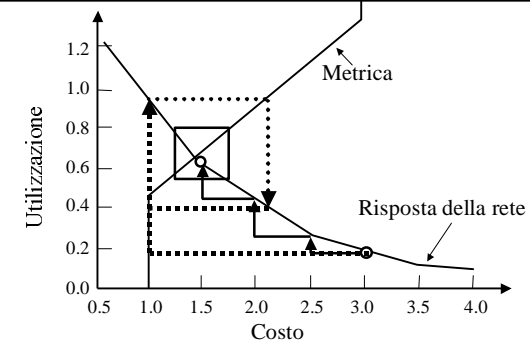
Esprime il valore del costo in funzione del carico sulla linea

Carico medio in funzione del valore del costo su una singola linea (una curva per ogni valore del carico complessivo offerto alla rete)

Complementi di Reti e Sistemi C. Nobile

5.43

Metriche e dinamica



Complementi di Reti e Sistemi C. Nobile

5.44

5. Instradamento

DVR e LSR

- Il confronto DV *Routing* (DVR) e LS *Routing* (LSR) è complesso, proviamo a distinguere diversi aspetti:
 - Velocità di convergenza
 - » In genere si tende a supporre che gli LSR convergano più rapidamente dei DVR, in pratica la velocità di convergenza dipende molto dalla topologia della rete e dalle caratteristiche del traffico.
 - Volume di messaggi di controllo
 - » LS: con N nodi e A archi richiedono lo scambio di $O(NA)$ messaggi per ciascun nodo.
 - » DV: deve solo scambiare i messaggi con i propri vicini.
 - Robustezza; se un nodo comincia a funzionare male:
 - » LS: il nodo propaga un costo sbagliato, ogni nodo si calcola separatamente la propria tabella
 - » DV: il nodo propaga un percorso sbagliato, ogni tabella viene calcolata facendo uso delle altre

Complementi di Reti e Sistemi C. Nobile

5.45

DVR e LSR

- I DVR non escludono la presenza di cicli a priori, ma con le opportune modifiche gli possono evitare efficacemente.
- Gli LSR, per contro, sono più complessi, devono fare uno sforzo significativo per mantenere i DB congruenti (generando anche un traffico di controllo più elevato) ed hanno *Distance Table* più grandi.
- Gli LSR possono usare più metriche diverse contemporaneamente.
- Gli LSR si prestano ad essere estesi per supportare con le stesse tabelle routing *unicast* e *multicast*

Complementi di Reti e Sistemi C. Nobile

5.46

Gerarchia

Ci sono due ragioni importanti per le quali nelle reti di una certa dimensione si tende ad usare meccanismi di instradamento gerarchici:

- **La scalabilità**
 - Per un numero di nodi elevato (WAN), indipendentemente dal tipo di algoritmo, la complessità dell'instradamento e la dimensione delle RT diventano comunque eccessive (oltre al traffico di segnalazione).
 - Per esempio, nel caso LS, con tanti archi quanti nodi, si ha una complessità di circa $O(N \log N)$ ed una RT con dimensione $O(N)$, quindi

# nodi	RT	Calcoli
1000	1000	$O(3000)$
1.000.000	1.000.000	$O(6.000.000)$

- **L'autonomia amministrativa**

Complementi di Reti e Sistemi C. Nobile

5.47

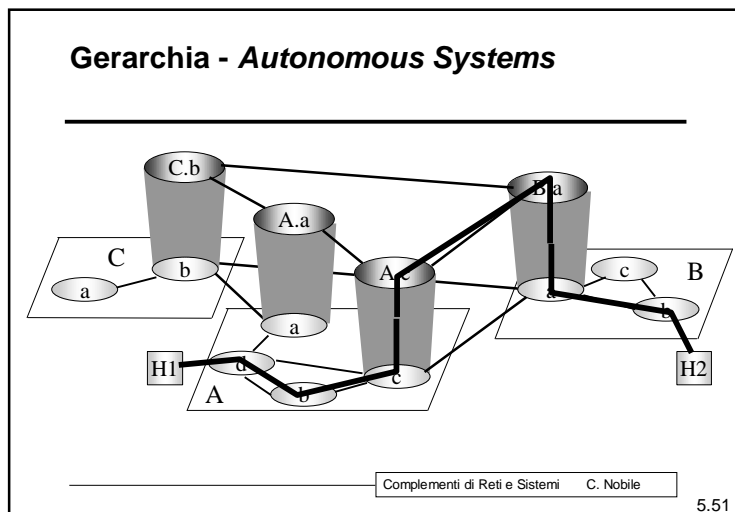
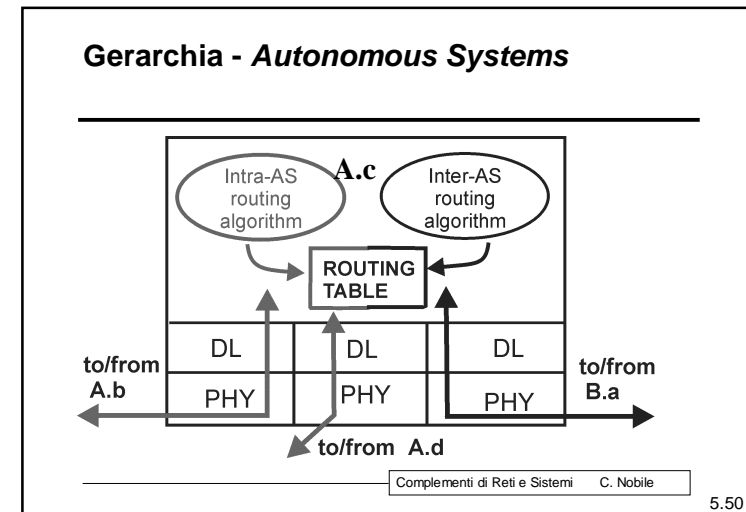
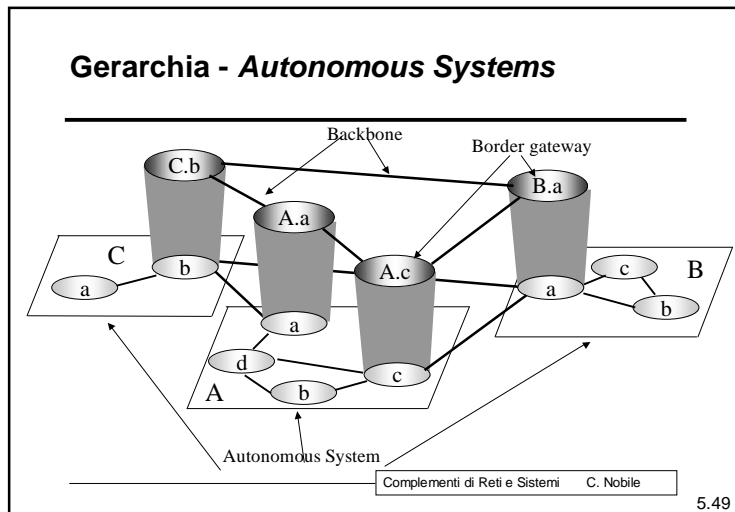
Gerarchia

- Internet distingue tre livelli gerarchici principali:
 - Sottoreti o singoli domini di *broadcast*
 - » All'interno dei quali l'instradamento fa uso dell'ARP.
 - *Autonomous System* (AS)
In cui i protocolli di instradamento prendono il nome di **Interior Gateway Protocol** (IGP) e sono: RIP, IGRP e OSPF.
 - *Backbone*
In cui i protocolli di instradamento prendono il nome di **Exterior Gateway Protocol** (EGP) e sono: EGP e BGP.
- Ulteriori livelli possono essere inseriti tramite alcuni protocolli (OSPF), o sfruttando la "route aggregation".

Complementi di Reti e Sistemi C. Nobile

5.48

5. Intradamento



- ### Gerarchia - Autonomous Systems
- Il partizionamento è realizzato grazie alla gerarchizzazione degli indirizzi.
 - L'obiettivo è avere relativamente "pochi" nodi per ogni livello (e area).
 - Potenzialmente ogni livello può usare algoritmi diversi
 - La gerarchia non è stretta, ossia il collegamento un area di un livello e il livello superiore può avvenire tramite più nodi (*Border Gateway, BG*)
 - Ci sono dei router che partecipano all'instradamento di livelli differenti.
 - Alcuni indirizzi possono non essere omogenei con lo spazio di indirizzamento dell'area/livello (questo diminuisce l'efficacia della gerarchia).
- Complementi di Reti e Sistemi C. Nobile
- 5.52

5. Instradamento

Gerarchia - *Autonomous Systems*

- o I diversi livelli non possono nascondersi reciprocamente tutte le informazioni:
 - Ad es., per poter calcolare l'instradamento più opportuno, un nodo del livello 3 deve conoscere i costi per raggiungere i nodi del livello superiore.
 - Allo stesso modo, un nodo di livello 4 deve conoscere i costi verso i nodi del livello 3.
 - Queste conoscenze sono fornite tramite LSP particolari (detti *external records* e *summary records*) che contengono solo le destinazioni ed i costi per raggiungerle (non la topologia). In pratica le reti dei livelli superiori/inferiori vengono rappresentate come se i loro nodi fossero direttamente collegati ai BG.

Complementi di Reti e Sistemi C. Nobile

5.53

Routing Information Protocol (RIP)

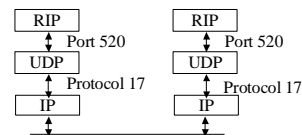
- E' un IGP originariamente progettato dalla Xerox per la propria rete, introdotto dall'Università di Berkley nella propria implementazione di TCP/IP (BSD)
- E' stato standardizzato con RFC 1058 nel 1988, la versione 2 è descritta nel RFC 1723.
- E' un DVR ed usa una metrica statica: il costo di un percorso è il numero di hop (ossia di linee) di cui è composto (ossia ogni linea a costo 1).
- Utilizza lo *split horizon with poisonous reverse*, e i *triggered update*.
- Aggiorna la RT (tramite *RIP response message* o *RIP advertisement*) ogni 30 s. e elimina ogni vettore non aggiornato per 180 s consecutivi (considerando la corrispondente linea non più disponibile).

Complementi di Reti e Sistemi C. Nobile

5.54

Routing Information Protocol (RIP)

- Ha come valore massimo del costo 15, 16 corrisponde ad infinito. Quindi non permette reti con percorsi con più di 15 *router* attraversati.
- La limitazione di cui sopra è legata al fatto che per reti più grandi è troppo lento a convergere (non alla dimensione del campo costo).
- Ne esistono due versioni, la seconda (RIPv2) consente l'uso del CIDR, ossia la "*route aggregation*", e la "*default route*".
- Lo scambio di informazioni avviene attraverso un protocollo di livello 4 (UDP)



Complementi di Reti e Sistemi C. Nobile

5.55

Routing Information Protocol (RIP)

- Sollecito per un DV
 - Messaggio di *update* (anche su sollecito) ^{1 o 2}
- Usato solo in RIPv2 per distinguere fra percorsi interni all'AS ed esterni.

Command	Version	Unused
Address Family ID		Route Tag
IP Address		
Subnet Mask		
Next Hop		
Metric		

- Usato per autenticazione
- Nel RIPv2 viene posto a FFFF e viene aggiunto successivamente un campo *password*

Complementi di Reti e Sistemi C. Nobile

5.56

5. Intradamento

Interior Gateway Routing Protocol (IGRP)

- E' nuovamente un DVR, ma di tipo proprietario; infatti è stato sviluppato dalla CISCO verso la metà degli anni '80 ed è disponibile solo sui suoi prodotti.
- Usa una metrica dinamica e sofisticata (considera ritardo, banda, affidabilità, lunghezza del pacchetto ed il carico) in cui il costo della linea viene composto tramite una somma pesato, i cui pesi sono impostabili dal gestore.
- Permette la suddivisione del carico su più linee (multipercorso).
- Usa un meccanismo sofisticato per accelerare la convergenza ed evitare i cicli.

Complementi di Reti e Sistemi C. Nobile

5.57

Open Shortest Path First (OSPF)

- Nasce nel 1990 con l'RFC 1247 per sostituire il RIP
- E' un protocollo di tipo **Link State**
- Quindi ogni nodo costruisce al proprio interno la topologia di tutta la rete e invia in *flooding* i LSP, contenenti i costi dei *link* ad esso connessi.

Complementi di Reti e Sistemi C. Nobile

5.58

Open Shortest Path First (OSPF)

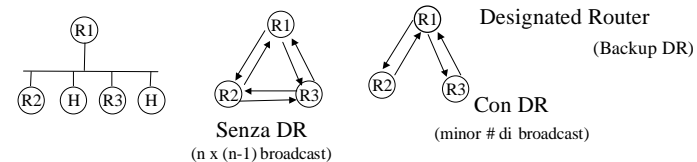
- E' stato realizzato per rispondere a diverse esigenze:
 - **Open**: ossia aperto e non proprietario.
 - **Sicurezza**: gli scambi fra *router* vengono autenticati, per proteggere gli aggiornamenti.
 - **Multi-metrica**: permette l'uso di più metriche anche dinamiche e instradamenti differenziati a seconda del campo TOS.
 - **Multi-percorso**: permette il bilanciamento dei flussi su percorsi a costo uguale.
 - **Gerarchico**: supporta una gerarchia interna.
 - **Multicast**: supporta il multicast (M-OSPF)

Complementi di Reti e Sistemi C. Nobile

5.59

Open Shortest Path First (OSPF)

- Supporta tre tipi di connessione e reti
 - Punto - punto fra router
 - Reti multiaccesso con *broadcast* (LAN)
 - Reti multiaccesso senza *broadcast* (WAN a pacchetto)
- Nel caso di LAN a cui sono connessi più router identifica un *router* di riferimento (*Designated Router*, DR) per ridurre il traffico di LSP in broadcast sulla LAN.

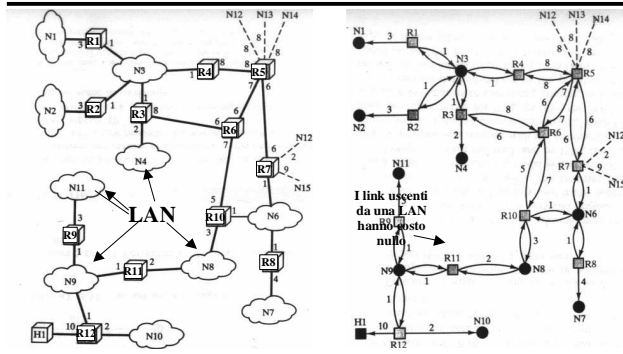


Complementi di Reti e Sistemi C. Nobile

5.60

5. Instradamento

Open Shortest Path First (OSPF)



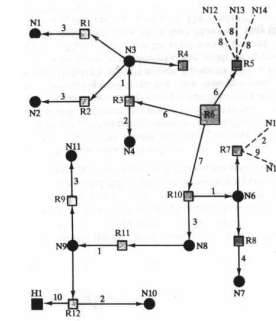
Complementi di Reti e Sistemi C. Nobile

5.61

Open Shortest Path First (OSPF)

TABLE 16.5 Routing Table for RT6.

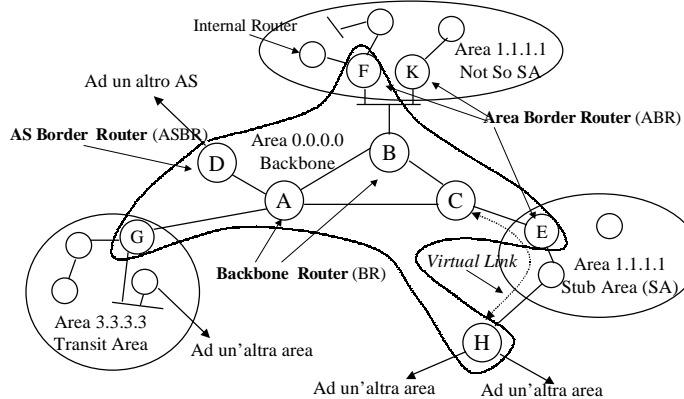
Destination	Next hop	Distance
N1	RT3	10
N2	RT3	10
N3	RT3	7
N4	RT3	8
N6	RT10	8
N7	RT10	12
N8	RT10	10
N9	RT10	11
N10	RT10	13
N11	RT10	14
H1	RT10	21
RT5	RT5	6
RT7	RT10	8
N12	RT10	10
N13	RT5	14
N14	RT5	14
N15	RT10	17



Complementi di Reti e Sistemi C. Nobile

5.62

Open Shortest Path First (OSPF)



Complementi di Reti e Sistemi C. Nobile

5.64

Open Shortest Path First (OSPF)

- Le *Stub Area* non propagano informazioni interne o esterne ed accedono al *backbone* tramite un *router* di *default*.
- L'instradamento fra due aree viene realizzato in tre parti:
 - Il percorso nell'area sorgente fra la sorgente stessa ed un *Area Border Router*.
 - Il percorso fra i due ABR delle due aree tramite il *backbone*
 - Il percorso nell'area destinazione fra l'ABR che riceve il pacchetto dal *backbone* e la destinazione.
- In pratica si forza un instradamento a stella in cui il *backbone* rappresenta il centro stella.

5. Instradamento

EGP- "EGP"

- Al più vecchio dei protocolli EGP è stato assegnato lo stesso nome che distingue la categoria: EGP.
- E' un protocollo di stile DV che però non propaga costi ma solo informazioni di raggiungibilità.
- Non è in grado di evitare cicli e quindi non può essere usato in topologie magliate ma solo ad albero.
- La sua struttura di riferimento è composta da "Core Router" (CR) collegati fra loro ad albero.
- Ogni AS può essere collegato ad un unico CR e quindi ogni CR fa da centro stella per un gruppo di AS

Complementi di Reti e Sistemi C. Nobile

5.65

EGP - Border Gateway Protocol (BGP)

- E' il protocollo EGP relativamente recente, definito dal RFC 1654.
- La versione in uso attualmente è la 4 (BGP4).
- Permette la cooperazione fra *router* di AS diversi (chiamati *gateway*) per la realizzazione dell'instradamento fra AS.
- Per lo scambio di informazioni fra i nodi usa il TCP (porta 179).
- Non propaga vere e proprie metriche, ma lascia che la scelta dei percorsi venga determinata tramite "politiche" impostate dai singoli gestori.

Complementi di Reti e Sistemi C. Nobile

5.66

EGP - Border Gateway Protocol (BGP)

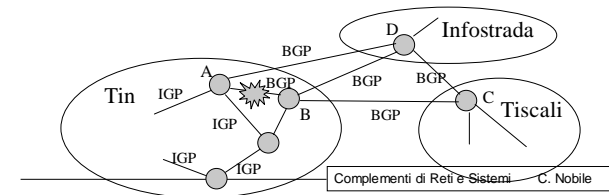
- Opera in tre passi:
 - Identificazione dei nodi adiacenti (*neighbor*)
 - Raggiungibilità dei nodi adiacenti
 - Raggiungibilità delle reti
- Utilizza un algoritmo DV, ed in particolare usa un *Path Vector*.
- Distingue tre tipi di reti
 - *Stub*: che hanno un'unica connessione con il *backbone* e non possono venir usate come transito
 - *Multiconnected*: che potenzialmente possono essere usate per transito (se lo permettono)
 - *Transit*: costruite per realizzare il transito.

Complementi di Reti e Sistemi C. Nobile

5.67

EGP - Border Gateway Protocol (BGP)

- Anche se tutti i *router* all'interno di un AS sono fra loro cooperativi, non è detto che i nodi che interconnettono due AS si "fidino" l'uno dell'altro.
- Questo accade in quanto gli AS sono in genere controllati o posseduti da organizzazioni diverse e quindi l'instradamento deve dipendere anche dagli accordi fra i diversi AS (transito).



Complementi di Reti e Sistemi C. Nobile

5.68

5. Instradamento

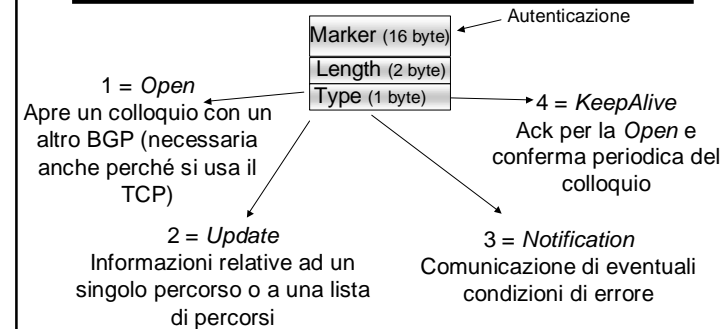
EGP - Border Gateway Protocol (BGP)

- In quanto protocollo di tipo EGP, il BGP differisce dai protocolli IGP in diversi aspetti:
 - Politica: fra AS nella scelta del percorso la “politica” domina, ossia le scelte dipendono da più fattori che coinvolgono considerazioni strategiche, economiche e di sicurezza più che tecniche. Queste considerazioni sono specifiche di ogni AS, quindi la scelta è principalmente sotto il controllo amministrativo.
 - Scala: la scalabilità è molto importante perché le “reti” coinvolte sono generalmente grandi. Invece se un AS cresce eccessivamente lo si può sempre dividere in due.
 - Prestazioni: le prestazioni tecniche contano relativamente poco in un EGP.

Complementi di Reti e Sistemi C. Nobile

5.69

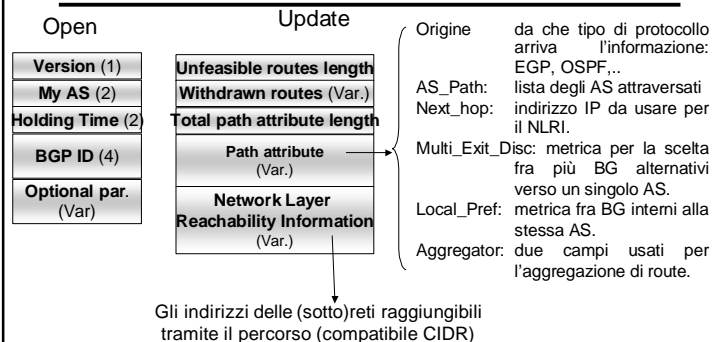
EGP - Border Gateway Protocol (BGP)



Complementi di Reti e Sistemi C. Nobile

5.70

EGP - Border Gateway Protocol (BGP)



Complementi di Reti e Sistemi C. Nobile

5.71