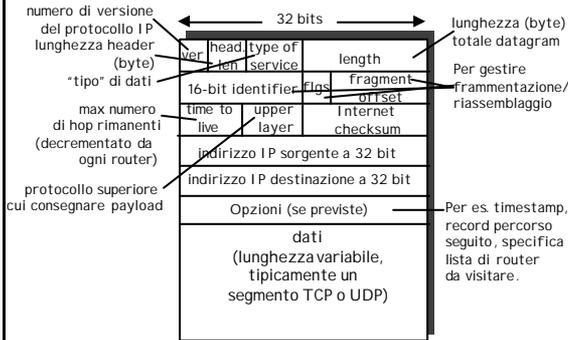
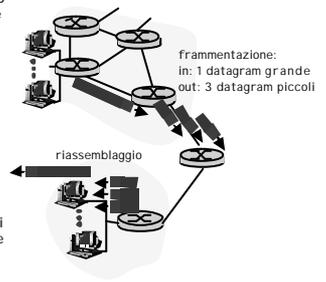


Formato del datagram IP

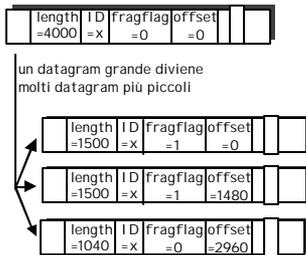


Frammentazione & Riassetaggio

- o I link della rete hanno un MTU (max transfer unit) - la frame di livello link più grande possibile.
 - differenti tipi di link, differenti MTU
- o datagram grandi divisi ("frammentati") dalla rete
 - un datagram diviene molti datagram
 - "riassetati" solo alla destinazione finale
 - I bit di header di IP usati per identificare e ordinare frammenti correlati



Frammentazione & Riassetaggio



ICMP: Internet Control Message Protocol

- o Usato da host, router e gateway per comunicare informazioni di livello network
 - Segnalazione errori: host, network, port, protocollo irraggiungibili
 - Richieste/risposte echo (usato dal ping)
- o È un livello network "sopra" IP:
 - I messaggi ICMP sono portati dai datagram IP
- o messaggi ICMP: tipo e codice più i primi 8 byte del datagram IP che causa errore

Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

II LivelloNetwork 4b-4

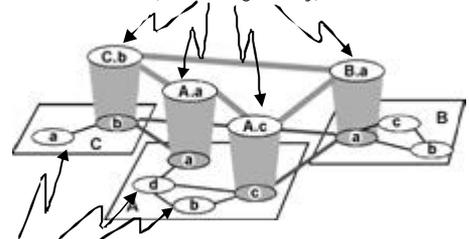
Routing in Internet

- o La Internet Globale consiste di Autonomous Systems (AS) interconnessi l'un l'altro:
 - **Stub AS**: piccole organizzazioni
 - **Multihomed AS**: grandi organizzazioni (di tipo no transit)
 - **Transit AS**: provider
- o Routing a due livelli:
 - **Intra-AS**: l'amministratore è responsabile delle scelte
 - **Inter-AS**: standard unico

II LivelloNetwork 4b-5

Gerarchia degli AS di Internet

Inter-AS border (exterior gateway) routers



Intra-AS Routing

- o Noti come Interior Gateway Protocols (IGP)
- o I più comuni IGP:
 - RIP: Routing Information Protocol
 - OSPF: Open Shortest Path First
 - IGRP: Interior Gateway Routing Protocol (Cisco propr.)

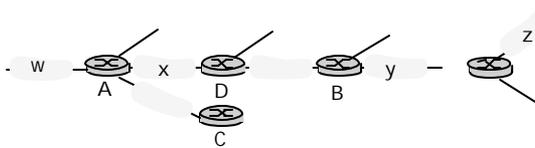
II LivelloNetwork 4b-7

RIP (Routing Information Protocol)

- o Algoritmo Distance Vector
- o Incluso nella versione BSD-UNIX del 1982
- o Metrica della Distanza: # of hop (max = 15 hop)
- o Distance vectors: scambiati ogni 30 sec attraverso Response Message (anche detti **advertisement**)
- o Ogni advertisement: instrada fino a 25 reti di destinazione

II LivelloNetwork 4b-8

RIP (Routing Information Protocol)



Rete Destinazione	Prossimo Router	# di hops per dest.
w	A	2
y	B	2
z	B	7
x	--	1
....

Tebella di Routing in D

II LivelloNetwork 4b-9

RIP: Gestione malfunzionamenti dei Link

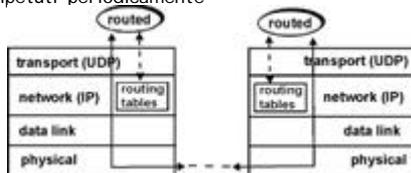
Se non arrivano advertisement entro 180 sec --> il vicino/link viene dichiarato morto

- I percorsi tramite quel vicino sono invalidati
- Viene inviato un advertisement ai vicini
- I vicini a loro volta inviano advertisement (se le tabelle sono cambiate)
- Le info di malfun. dei link si propagano rapidamente all'intera rete
- Si usa il poison reverse per evitare la creazione di ping-pong loop (distanza infinita = 16 hops)

II LivelloNetwork 4b-10

RIP: Elaborazione delle Tabelle

- o Le tabelle di routing del RIP sono gestite da un processo di livello **applicazione** detto route-d (daemon)
- o Gli advertisement sono inviati in pacchetti UDP, ripetuti periodicamente



II LivelloNetwork 4b-11

RIP: Esempio di Tabelle

Router: *giroflie.eurocom.fr*

Destination	Gateway	Flags	Ref	Use	Interface
127.0.0.1	127.0.0.1	UH	0	26492	lo0
192.168.2.	192.168.2.5	U	2	13	fa0
193.55.114.	193.55.114.6	U	3	58503	le0
192.168.3.	192.168.3.5	U	2	25	qa0
224.0.0.0	193.55.114.6	U	3	0	le0
default	193.55.114.129	UG	0	143454	

- o tre reti di classe C (LAN) connesse
- o il Router conosce solo i percorsi verso le LAN
- o Si usa un Default router per tutto il resto
- o Viene usato un indirizzo multicast: 224.0.0.0
- o Interfaccia di Loopback per il debugging

II LivelloNetwork 4b-12

OSPF (Open Shortest Path First)

- o "open": disponibile pubblicamente
- o usa l'algoritmo Link State
 - > dissemina pacchetti LS
 - > ogni nodo possiede la mappa della topologia della rete
 - > il calcolo dei percorsi avviene usando l'algoritmo di Dijkstra
- o gli advertisement OSPF hanno una entry per ogni router vicino
- o gli advertisement vengono disseminati all'intero AS (via flooding)

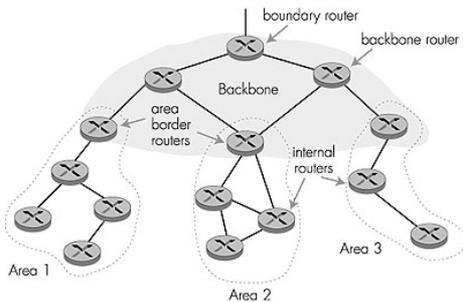
II LivelloNetwork4b-13

OSPF: caratteristiche "avanzate" (assenti in RIP)

- o Security: tutti i messaggi OSPF sono autenticati (per prevenire intrusioni); uso di connessioni TCP
- o Sono consentiti percorsi multipli dello stesso costo (Multipath)
- o Per ciascun link, vengono gestite metriche di costo multiple per diversi TOS (per es., i link satellitari costano "low" per il best effort; "high" per real time)
- o Supporto integrato uni- e multicast:
 - > II Multicast OSPF (MOSPF) usa lo stesso database di topologia dell'OSPF
- o OSPF Gerarchico per domini grandi

II LivelloNetwork4b-14

OSPF Gerarchico



II LivelloNetwork4b-15

OSPF Gerarchico

- o Gerarchia a due livelli: local area, backbone.
 - > Advertisement Link-state solo in area
 - > Ogni nodo possiede la topologia dettagliata dell'area; conosce solo la direzione (il percorso minimo) verso reti di altre aree
- o **Area border routers**: "riassume" le distanze verso le reti nella sua area, annuncia agli altri Area Border router.
- o **Backbone routers**: eseguono routing OSPF limitatamente al backbone.
- o **Boundary routers**: connettono ad altri AS.

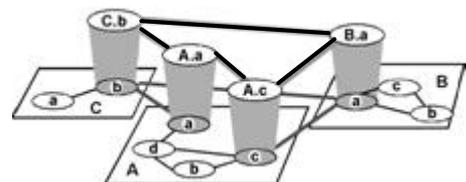
II LivelloNetwork4b-16

IGRP (Interior Gateway Routing Protocol)

- o Protocollo proprietario CISCO; successore di RIP (metà anni 80)
- o Distance Vector, come RIP
- o Svariate metriche di costo (ritardo, larghezza di banda, affidabilità, carico, ecc.)
- o Usa il TCP per scambiare aggiornamenti di routing
- o Routing Loop-free attraverso l'uso del Distributed Updating Algorithm (DUAL) basato sul concetto di *diffused computation*

II LivelloNetwork4b-17

Inter-AS routing



II LivelloNetwork4b-18

Internet inter-AS routing: BGP

- o BGP (Border Gateway Protocol): lo standard de facto
- o protocollo **Path Vector**:
 - simile al protocollo Distance Vector
 - Ogni Border Gateway invia in broadcast ai vicini (peers) *l'intero percorso* (cioè, la sequenza di AS) verso la destinazione
 - Per es., Gateway X può inviare il suo percorso verso la destinazione Z:

Path (X,Z) = X,Y1,Y2,Y3,...,Z

II LivelloNetwork 4b-19

Internet inter-AS routing: BGP

- Poniamo che:* il gateway X invia il suo path al peer gateway W
- o W può oppure no selezionare il path offerto da X
 - Sulla base di motivi di costo, policy (non attraversare AS rivali), loop prevention
 - o Se W seleziona il path annunciato da X, allora:
 - Path (W,Z) = w, Path (X,Z)
 - o Nota: X può controllare il traffico entrante evitando di annunciare i suoi percorsi ai peers:
 - Per es., non voglio instradare traffico a Z -> non invio alcun annuncio a Z

II LivelloNetwork 4b-20

Internet inter-AS routing: BGP

- o i messaggi BGP vengono scambiati usando il TCP.
- o messaggi BGP:
 - OPEN: apre una connessione TCP con il peer e autentica il mittente
 - UPDATE: annuncia un nuovo (o cancella uno vecchio)
 - KEEPALIVE: tieni la connessione in assenza di UPDATES; serve anche per gli ACK alle richieste di OPEN
 - NOTIFICATION: rileva errori nei messaggi precedenti; usato anche per chiudere la connessione

II LivelloNetwork 4b-21

Perchè usare tecniche differenti per il routing Intra- e Inter-AS?

Politica:

- o Inter-AS: l'amministratore vuole il controllo su come il suo traffico viene instradato e su chi instrada attraverso la sua rete
- o Intra-AS: singolo amministratore, non vi sono decisioni di policy

Dimensione:

- o Il routing gerarchico riduce la dimensione delle tabelle ed il traffico di update

Prestazione:

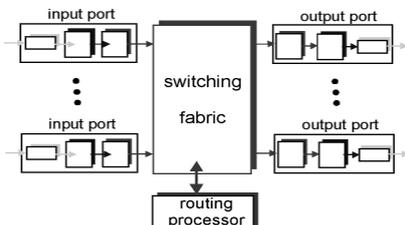
- o Intra-AS: può focalizzarsi sulla prestazione
- o Inter-AS: la politica prevale sulla prestazione

II LivelloNetwork 4b-22

Architettura dei Router: Panoramica

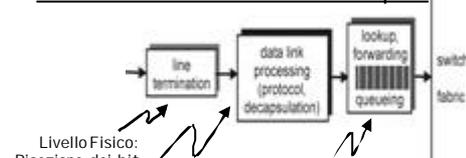
Due funzionalità chiave:

- o Eseguire gli algoritmi/protocolli (RIP, OSPF, BGP)
- o *Commutare (switching)* i datagram dai link di ingresso a quelli di uscita



II LivelloNetwork 4b-23

Funzioni della Porta di Input



Livello Fisico:
Ricezione dei bit
Livello Data link:
Per es., Ethernet

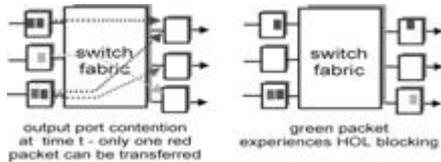
Switching decentralizzato:

- o Data la dest del datagram, cerca la porta di output usando la tabella di routing nella memoria della porta di input
- o Obiettivo: elaborazione completa alla porta di input alla 'velocità di linea'
- o accodamento: se i datagram arrivano più velocemente del tasso di inoltr nella struttura dello switch

II LivelloNetwork 4b-24

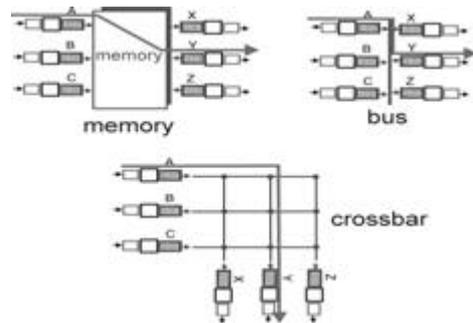
Accodamento in Input

- o Se lo switch è più lento della combinazione delle porte di input-> accodamento in ingresso
- o Head-of-the-Line (HOL) blocking: i datagram accodati in testa impediscono agli altri di avanzare
- o Ritardi di accodamento e perdite per overflow dei buffer di input!



II LivelloNetwork 4b-25

Tre tipi di strutture di switching

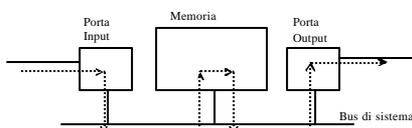


II LivelloNetwork 4b-26

Switching Via Memory

router di prima generazione :

- o pacchetti copiati dalla (singola) CPU di sistema
- o velocità limitata dalla banda di memoria (2 attraversamenti del bus per datagram)

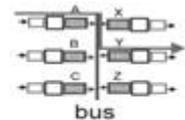


I router moderni:

- o il processore della porta di input port effettua ricerca e copia in memoria
- o Cisco Catalyst 8500

II LivelloNetwork 4b-27

Switching Via Bus



- o Il datagram viene instradato dalla memoria della porta di input a quella di output via bus condiviso
- o Contesa sul bus : la velocità di switching è limitata dalla banda del bus
- o 1 Gbps bus, Cisco 1900: velocità sufficiente per router medi (non regionali o di backbone)

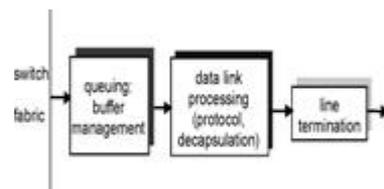
II LivelloNetwork 4b-28

Switching Via Interconnection Network

- o Supera le limitazioni della banda del bus
- o Banyan networks o altre reti di interconnessione originariamente sviluppate per connettere processori nei sistemi multiprocessori
- o Architetture avanzate: i datagram vengono frammentati in celle di lunghezza fissa che vengono instradate nella struttura dello switch
- o Cisco 12000: commuta i Gbps

II LivelloNetwork 4b-29

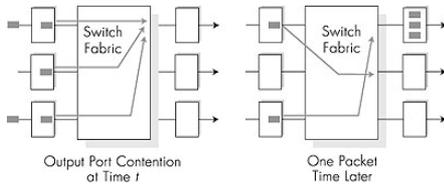
Porte di Output



- o Buffering necessario quando i datagram arrivano più velocemente del tasso di trasmissione
- o Una Scheduling discipline sceglie tra i datagram accodati per la trasmissione

II LivelloNetwork 4b-30

Output port queuing



- o buffering quando la velocità di arrivo attraverso lo switch supera la velocità delle linee di output
- o Ritardi di accodamento e perdite per overflow dei buffer di input!

II LivelloNetwork 4b-31

IPv6

- o Motivazione: spazio di indirizzamento a 32-bit completamente allocato entro il 2008.
- o Motivazioni aggiuntive:
 - > Formato header per velocizzare l'inoltro
 - > Cambiamenti header per facilitare QoS
 - > Nuovi indirizzi "anycast": instrada al "migliore" di un gruppo di server duplicati
- o formato datagram IPv6 :
 - > header a lunghezza fissa di 40 byte
 - > Frammentazione non consentita

II LivelloNetwork 4b-32

Header IPv6 (Cont)

- Priorità*: identifica una priorità tra i datagram
- Flow Label*: identifica datagram in un "flusso" (concetto di "flusso" non ben definito).
- Next header*: identifica il protocollo upper per i dati



II LivelloNetwork 4b-33

Altri cambiamenti da IPv4

- o *Checksum*: rimosso per ridurre il tempo di processo ad ogni hop
- o *Option*: consentite ma fuori dall'header, indicate dal campo "Next Header"
- o *ICMPv6*: nuova versione di ICMP
 - > tipi messaggio aggiuntivi, es., "Packet Too Big"
 - > Funzioni di management dei gruppi multicast

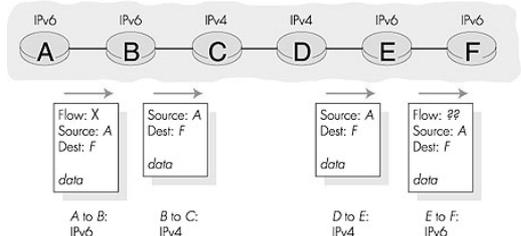
II LivelloNetwork 4b-34

Transizione da IPv4 a IPv6

- o Non tutti i router possono essere aggiornati simultaneamente
 - > Non c'è un "giorno stabilito"
 - > Come opererà la rete con router misti IPv4 e IPv6?
- o Proposti due approcci:
 - > *Dual Stack*: alcuni router con dual stack (v6, v4) "traducono" tra i formati
 - > *Tunneling*: IPv6 porta i datagram IPv4 come payload tra i router IPv4

II LivelloNetwork 4b-35

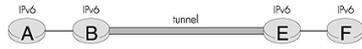
Approccio Dual Stack



II LivelloNetwork 4b-36

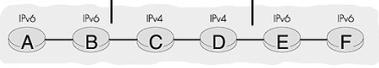
Tunneling

Logical view



Physical view

IPv6 in IPv4 dove necessario



A to B:
IPv6



B to C:
IPv4
(encapsulating
IPv6)



B to C:
IPv4
(encapsulating
IPv6)



E to F:
IPv6