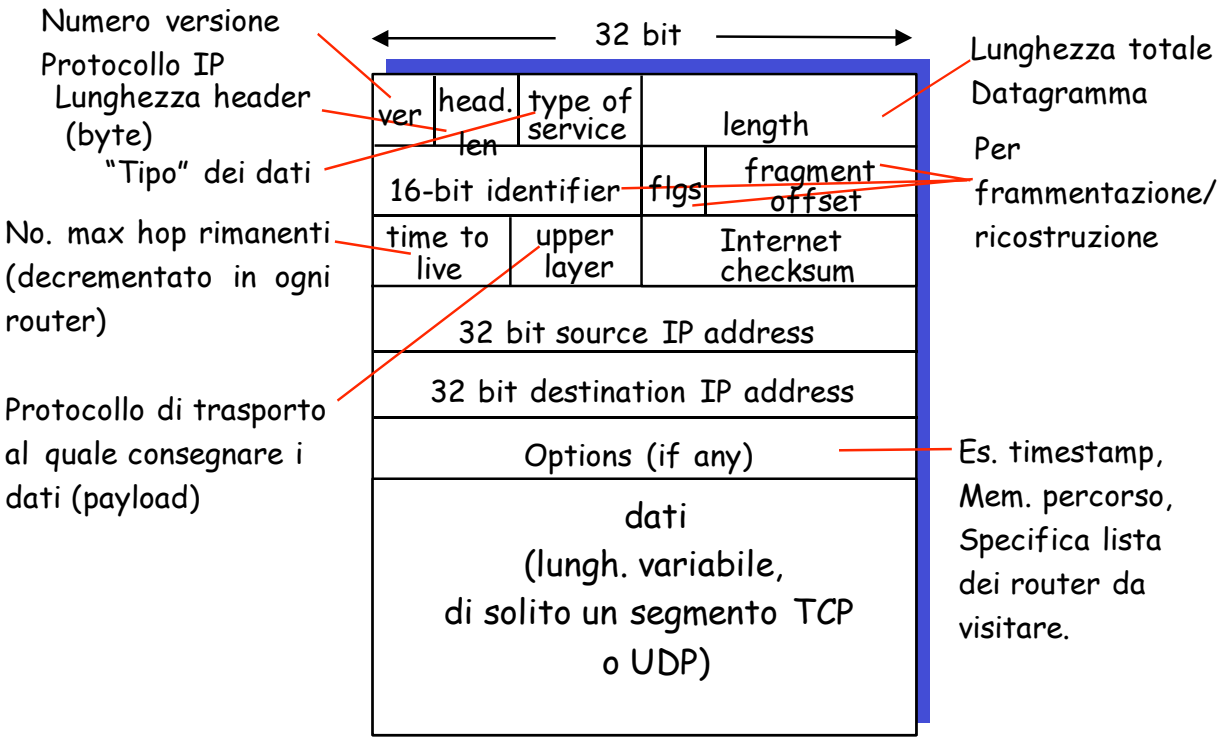
Routing (instradamento) in Internet

- Internet globalmente consist di Sistemi Autonomi (AS) interconnessi:
  - **Stub AS**: istituzione piccola
  - **Multihomed AS**: grande istituzione (nessun transito)
  - **Transit AS**: provider

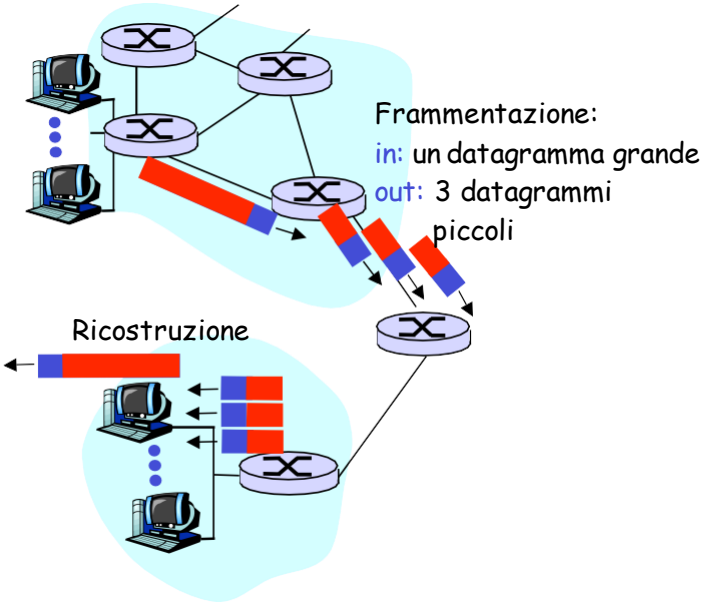
- Routing a due livelli :
  - **Intra-AS:** l'amministratore di rete è respnsabile della scelta
  - **Inter-AS:** standard unico

Formato del datagramma IP

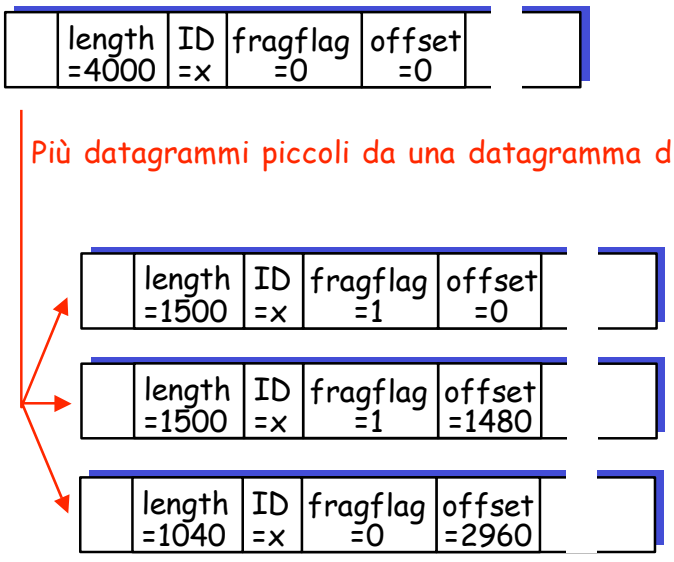


IP: frammentazione e ricostruzione

- I link hanno una MTU (max.transfer unit) - dim. Max frame del livello link.
  - Diversi tipi di link, MTU differenti
- Datagrammi IP grandi divisi ("frammentati") nella rete
  - Più datagrammi piccoli da un singolo datagramma grande
  - "Ricostruzione" solo alla destinazione finale
  - Alcuni bit di intestazione IP usati per identificare e ordinare frammenti dello stesso datagramma di partenza



IP: frammentazione e ricostruzione (cont.)

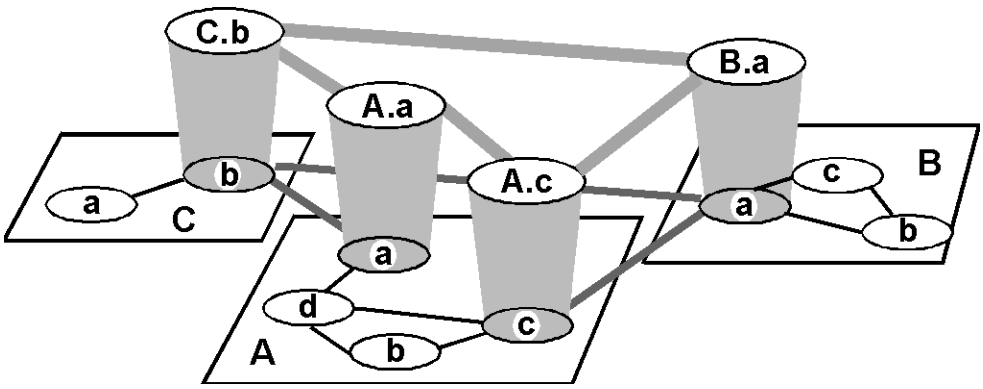


ICMP: Internet Control Message Protocol

- Usato da host, router, gateway per trasmettere info di rete
  - Notifica di errore: host irraggiungibile, rete, porta, protocollo
  - Richiesta/risposta eco (usato da ping)
- Si trova "sopra" IP:
  - Messaggi ICMP trasportati da datagrammi IP
- Messaggio ICMP : tipo, codice, più i primi 8 byte del datagramma IP che ha causato un 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

Gerarchia degli AS in Internet



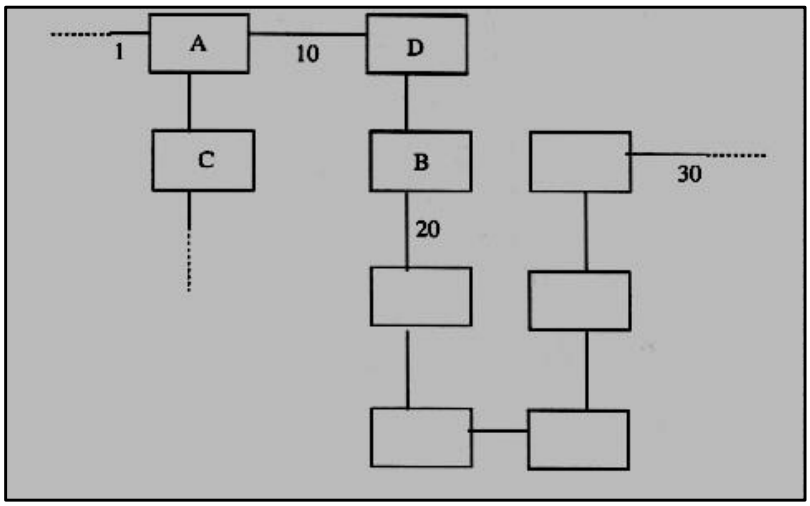
Routing intra-AS

- Noto anche come Interior Gateway Protocols (IGP)
- IGP più comuni :
  - RIP: Routing Information Protocol
  - OSPF: Open Shortest Path First
  - IGRP: Interior Gateway Routing Protocol (proprietario Cisco)

RIP (Routing Information Protocol)

- Schema di tipo distance vector
- Incluso nella distribuzione BSD-UNIX già nel 1982
- Metrica: # di hop (salti, max = 15 hop)
  - *Perché?*
- Vettore di distanza: scambiato ogni 30 secondi mediante un Response Message (detto anche **Advertisement**)
- Ogni Advertisement contiene fino a 25 reti di destinazione

RIP (Routing Information Protocol)



Rete Destinazione	Prox. Router	Num. di hop alla dest.
1	A	2
20	B	2
30	B	7
10	--	1
….	….	….

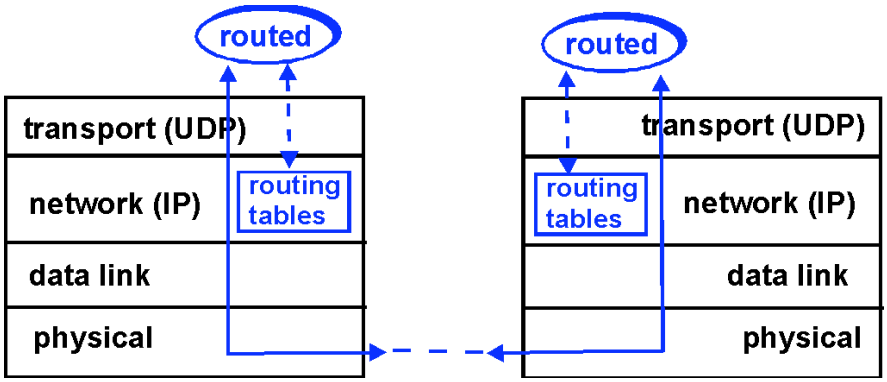
RIP: recupero da guasti

- Se non giunge un advertisement, dopo 180 sec. Si assume che il link sia morto
- Si rimuovono i cammini che usano il link; si mandano messaggi di notifica (advertisement) ai vicini
- I vicini fanno altrettanto se le loro tabelle di routing sono modificate (vedi schema DVR)
- L'informazione si propaga rapidamente all'intera rete

RIP: gestione delle tabelle di routing

- Le tabelle di routing RIP sono gestite da un processo applicativo chiamato route-d (daemon o demone, terminologia Unix)
- I messaggi di advertisement sono incapsulati in pacchetti UDP (non serve un trasferimento affidabile perché i messaggi di notifica sono inviati periodicamente)

RIP: gestione delle tabelle di routing (cont.)



RIP: gestione delle tabelle di routing, esempio

Tabella di routing RIP

(nel router *giroflee.eurocom.fr*):

- Tre reti di classe C (LAN) attaccate
- il router conosce solamente i percorsi verso le LAN direttamente collegate
- il router di default usato per “uscire”
- indirizzo multicast: 224.0.0.0
- Interfaccia di loopback (ritorno) per debugging

Tabella RIP – esempio

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	qaa0
224.0.0.0	193.55.114.6	U	3	0	le0
default	193.55.114.129	UG	0	143454	

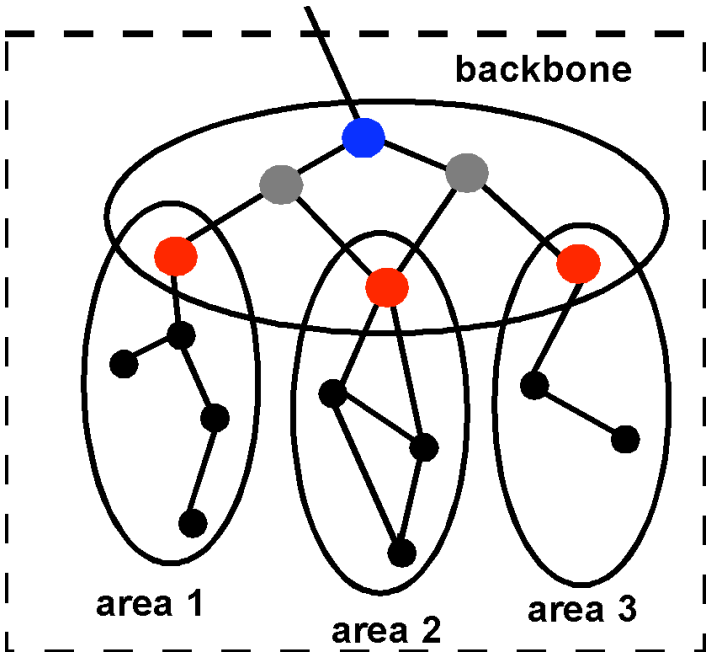
OSPF (Open Shortest Path First)

- "Aperto": disponibile al pubblico (gratis)
- Usa l'algoritmo Link State
  - Broadcast di pacchetti LS
  - Topologia in ogni nodo
  - Calcolo del cammino minimo usando alg. di Dijkstra
- Il messaggio di notifica OSPF contiene un campo per ogni router adiacente
- Il messaggio di notifica è inviato all intero Sistema Autonomo (mediante flooding)

Caratteristiche avanzate di OSPF (non presenti in RIP)

- Sicurezza: tutti i messaggi OSPF sono autenticati; si usano connessioni TCP
- Sono permessi più cammini minimi (solo uno in RIP)
- Per ogni link, metriche diverse per valori diversi del campo TOS (es, costo di un link satellitare "basso" per traffico best effort; alto per traffico real time)
- Supporto per uni- e multicast:
  - Multicast OSPF (MOSPF) usa lo stesso database (tabelle di routing) di OSPF
- OSPF gerarchico in sistemi autonomi di grandi dimensioni.

OSPF gerarchico



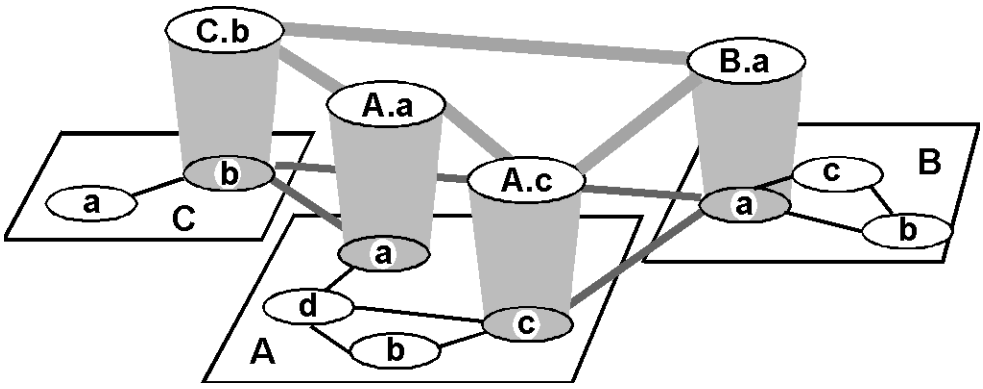
OSPF gerarchico

- Gerarchia a due livelli: area locale e backbone (rete principale).
- I messaggi di notifica link-state non lasciano le aree locali rispettive.
- I nodi di ciascuna area hanno una topologia dettagliata dell'area; conoscono solo la direzione (cammino minimo) verso reti di aree diverse.
- Gli **Area Border router** (router di confine dell'area) "riassumono" le distanze verso altre reti e inviano messaggi di advertisement ad altri Area Border router.
- **Backbone router** (router principali): eseguono un algoritmo di routing OSPF limitato al backbone.
- I **Boundary router** (router di confine) connettono verso altri AS.

IGRP (Interior Gateway Routing Protocol)

- Protocollo proprietario CISCO; successore di RIP (metà degli anni '80)
- Distance Vector, come RIP
- Molte metriche di costo (ritardo, banda, affidabilità, carico ecc.)
- Usa TCP per scambiare messaggi di aggiornamento del routing
- Il contenuto delle tabelle di routing è trasmesso ai vicini solo quando cambia
- Privo di cicli

Routing inter-AS



Routing inter-AS (cont.)

- BGP (Border Gateway Protocol): lo standard di fatto
- Protocollo **Path Vector**: estensione del Distance Vector
- Ogni Gateway invia in broadcast ai vicini (peer) l'intero percorso (la sequenza degli AS attraversati) fino a destinazione
- Per esempio, il Gateway X può memorizzare il seguente cammino verso la destinazione Z:

Path (X,Z) = X,Y1,Y2,Y3,…,Z

Routing Inter-AS (cont.)

- Si supponga che X invii il cammino al Gwy peer W
- W può accettare o rifiutare il cammino proposto da X, per ragioni di costo, strategia ($$$$) o per evitare ciclo.
- Se W accetta il cammino proposto da X, allora:

Path (W,Z) = w, Path (X,Z)

Importante: la selezione del cammino minimo non è basata di solito sul costo (es.,# di AS attraversati), ma piuttosto su considerazioni "politiche"

(es., non instradare pacchetti verso AS di concorrenti)

Routing inter-AS (cont.)

- I gateway peer (pari) scambiano messaggi BGP usando TCP.

- OPEN: apre una connessione TCP verso un vicino e autentica il mittente
- UPDATE: notifica un nuovo cammino o propone l'eliminazione di uno vecchio
- KEEPALIVE: tiene in vita la connessione in assenza di messaggi UPDATE; serve anche come ACK per un messaggio OPEN
- NOTIFICATION: usato per notificare errori nel messaggio precedente e per chiudere la connessione

Perché routing intra- e inter-AS diversi?

- **Poltiche di amministrazione**: routing inter-AS implementa politiche di gestione
- **Scala**: routing inter-AS permette di gestire la complessità della rete
- **Prestazioni**: routing intra-AS ha l'obiettivo di mantenere i costi bassi.

Servono **entrambi**!