UNIVERSITÀ DI PARMA

DIPARTIMENTO DI SCIENZE MATEMATICHE, FISICHE ED INFORMATICHE
Corso di Laurea in Informatica

Il Livello Network

Parte I : IPv4

RETI DI CALCOLATORI - a.a. 2022/2023

Roberto Alfieri

Il livello Network: sommario

PARTE I

- Scopi del livello Network
- Commutazione di circuito e di pacchetto
- La famiglia dei protocolli TCP/IP
- Il protocollo IPv4: trama indirizzi, instradamento
- Protocolli di servizio per IPv4: ARP, ICMP, DHCP

PARTE II

- IPv6

PARTE III

- Algoritmi e protocolli di routing
- Distance Vector e Link State.

RIFERIMENTI

Reti di Calcolatori, A. Tanenbaum, ed. Pearson
Reti di calcolatori e Internet, Forouzan , Ed. McGraw-Hill

Scopi e servizi del livello Network

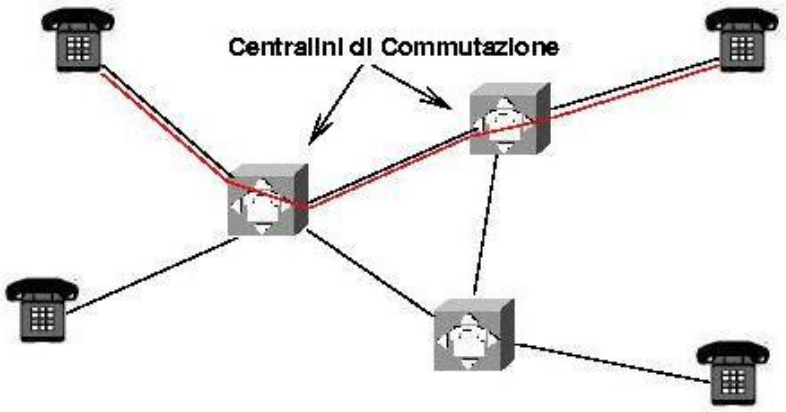
Estendere i servizi che il livello Data-Link offre a macchine connesse anche a macchine che non hanno una connessione diretta.
Compito fondamentale dello strato di rete è trasportare i pacchetti lungo tutto il percorso dal mittente al destinatario, attraversando tutti i nodi di transito dove sono possibili scelte alternative per le linee di uscita.
La funzione di collegamento di una linea di ingresso con una di uscita opportunamente scelta, che viene svolta nei nodi è detta **Commutazione** (Switching)

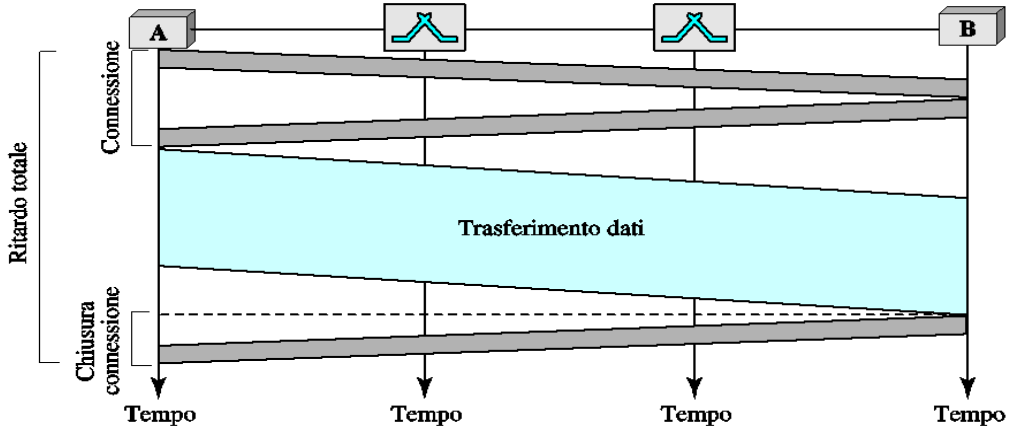
Le due tecniche di commutazione usate tradizionalmente nelle reti sono:

- **Commutazione di circuito** (utilizzata storicamente dai Provider di telefonia)
- **Commutazione di pacchetto** (utilizzata nelle reti di calcolatori)

Commutazione di Circuito

- I nodi di transito sono i Centralini di Commutazione (Manuali, Meccanici o Elettronici)
- L'algoritmo per la commutazione interviene all'apertura del canale fisico
- Nella fase di connessione vengono allocate le risorse necessarie
- Ritardo: è minimo nel trasferimento dati, ma è elevato in fase di apertura e chiusura della connessione
- Efficienza: le risorse allocate sono riservate anche se la connessione è inutilizzata. Questo non avviene per le telefonate, ma può avvenire per il trasferimento dati.

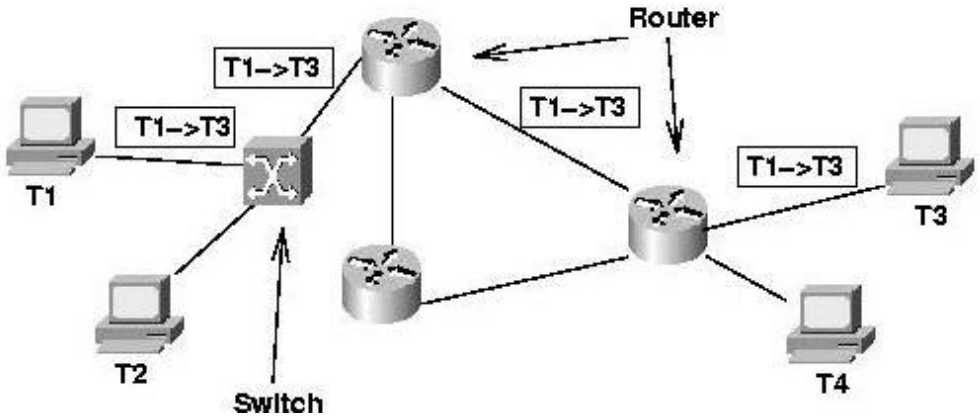




Commutazione di Pacchetto

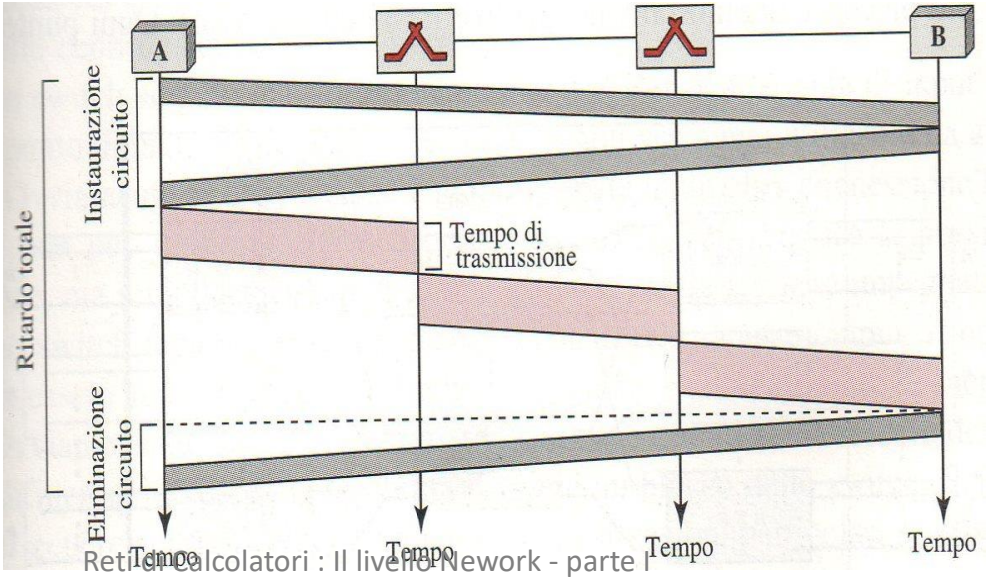
Commutazione di Pacchetto

- Comunicazione frazionata in “pacchetti”
- Algoritmo per la commutazione interviene sui pacchetti
- Esistono diversi tipi di nodi di transito a seconda della loro funzione:
  - Hub, Bridge, Switch, Router o Gateway.
- Esistono 2 tipologie di commutazione a pacchetto:
  - A circuito virtuale
  - A datagramma



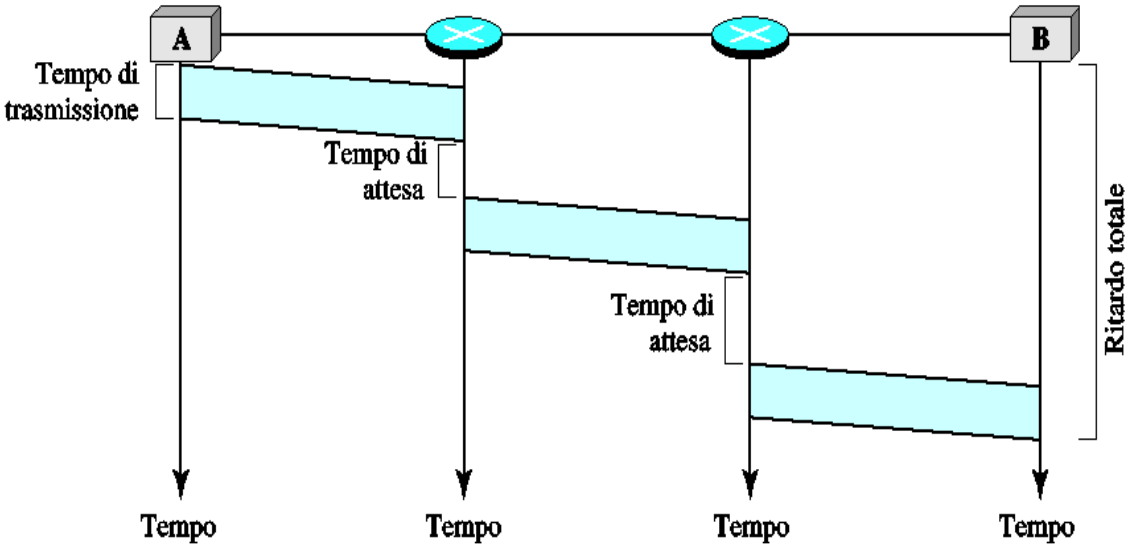
Commutazione di pacchetto a circuito virtuale

- Algoritmo per la commutazione interviene solo all'inizio per l'apertura del Canale Virtuale (VC).
- Ad ogni nuovo VC viene assegnata una etichetta; ogni router viene marcato con l'etichetta del VC e la relativa porta di uscita.
- I pacchetti seguono il percorso individuato
- Implementazioni principali:
  - ATM. E' la rete che utilizza la commutazione di pacchetto a circuito virtuale per la telefonia.
  - In internet è possibile creare isole a circuito virtuale con il protocollo MPLS
  - La versione 6 di IP supporta (anche) reti a circuito virtuale.



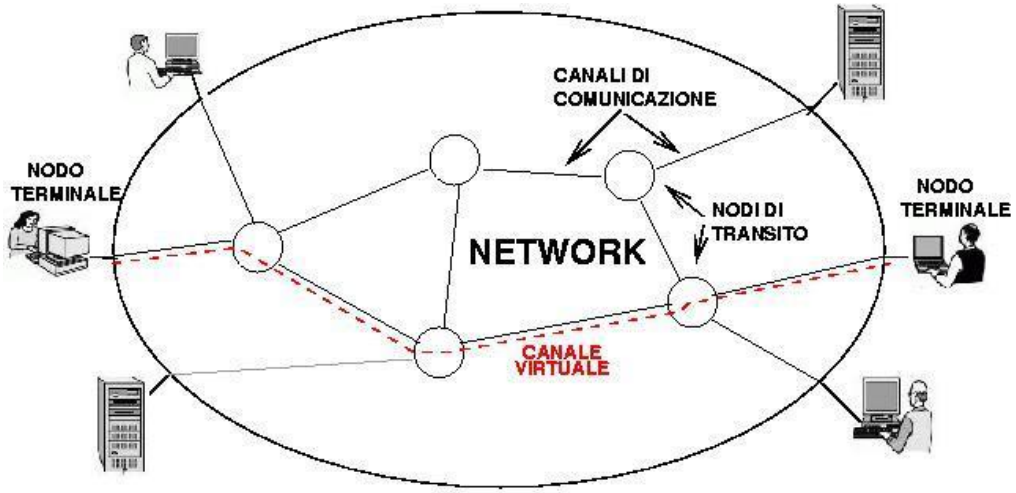
Commutazione di pacchetto a datagramma

- I pacchetti sono instradati in modo indipendente in base all'indirizzo di destinazione
- L'instradamento è determinato dai router attraversati in base "**tabelle di instradamento**" che ogni router costruisce dinamicamente mediante gli "**algoritmi di routing**".
- Pacchetti della stessa connessione possono seguire strade diverse.
- Implementazioni principali: IPv4 e IPv6.



Routing

Il routing è quella parte del software dello strato Network che si preoccupa di dell'instradamento dei pacchetti in transito.



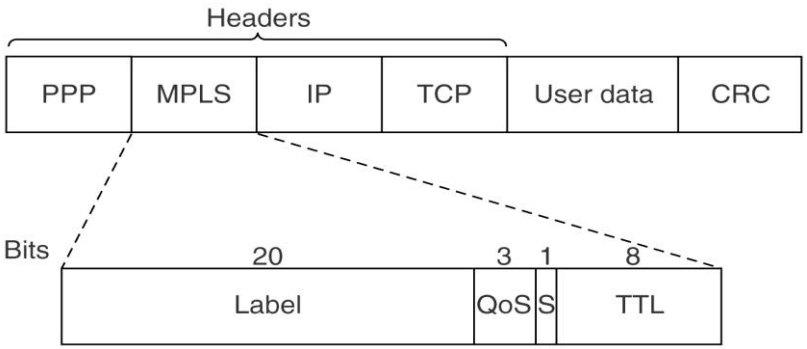
- Se la **Rete è a Datagramma** il routing viene determinato per ogni pacchetto, poiché il percorso migliore può cambiare nel tempo.
- Se la **Rete è a Circuito Virtuale** il routing viene determinato al momento dell'attivazione del circuito. Da quel momento in poi tutti i pacchetti seguono il percorso stabilito.

Confronto tra i metodi di commutazione a pacchetto

Issue	Datagram subnet	Virtual-circuit subnet
Circuit setup	Not needed	Required
Addressing	Each packet contains the full source and destination address	Each packet contains a short VC number
State information	Routers do not hold state information about connections	Each VC requires router table space per connection
Routing	Each packet is routed independently	Route chosen when VC is set up; all packets follow it
Effect of router failures	None, except for packets lost during the crash	All VCs that passed through the failed router are terminated
Quality of service	Difficult	Easy if enough resources can be allocated in advance for each VC
Congestion control	Difficult	Easy if enough resources can be allocated in advance for each VC

Reti a Circuito Virtuale: MPLS

MPLS (MultiProtocol Label Switching) consente di creare in Internet aree a commutazione di Label.
E' uno strato che si pone sotto il livello rete aggiungendo un proprio Header di 4 byte tra l'header di livello rete (IP) e quello di livello dati (ppp o Ethernet). Per questo può essere considerato un protocollo di livello 2.5.
I campi principali sono la Label (20bit), QoS, e TTL.

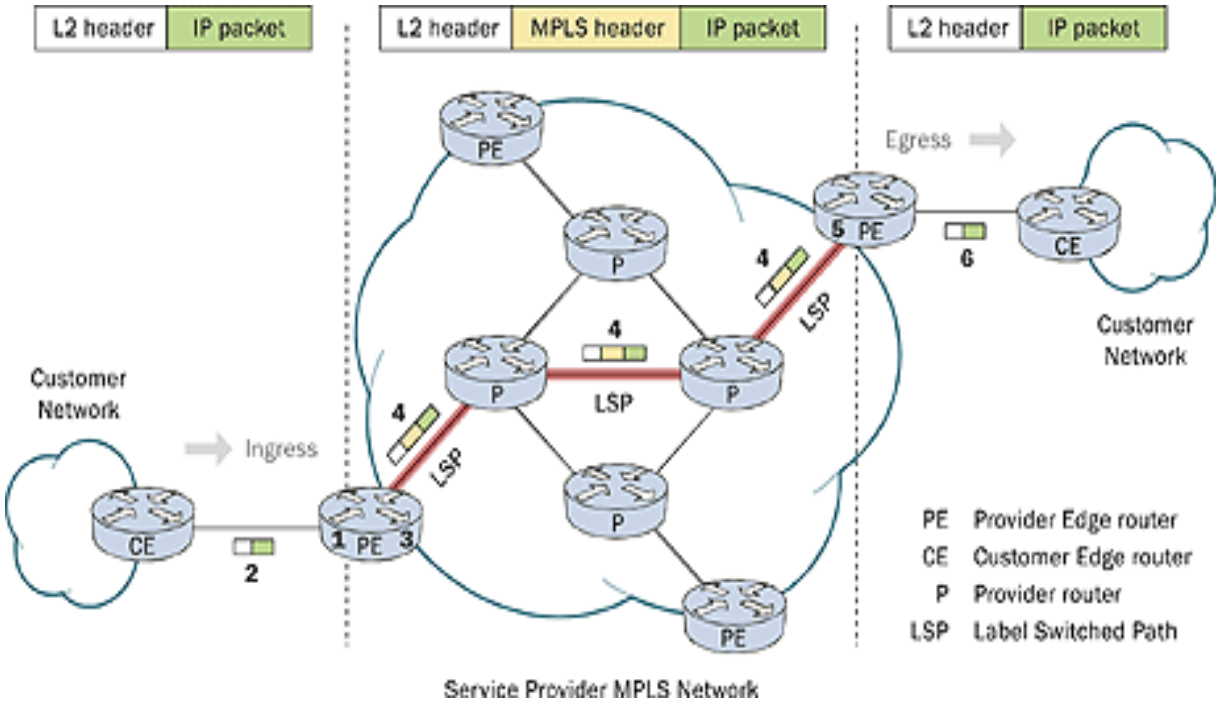


Vantaggi: QoS, Traffic Shaping (e-mail,web, ..), VPN.

Il routing MPLS

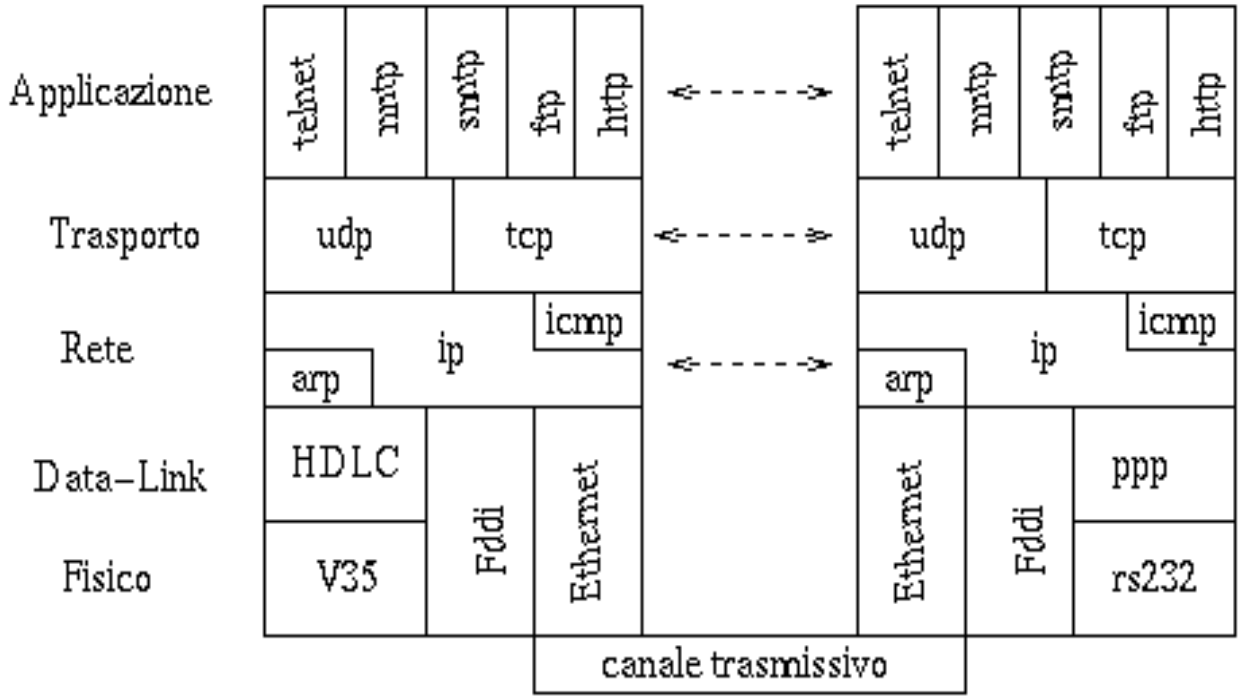
Richiede al proprio interno Router specifici che supportano il protocollo.
Il router di frontiera (Edge) determina il percorso e aggiunge l'header MPLS al pacchetto.

Attraverso le etichette il primo pacchetto definisce un “tunnel” nella rete MPLS.
I pacchetti successivi della stessa connessione seguono il percorso del primo.

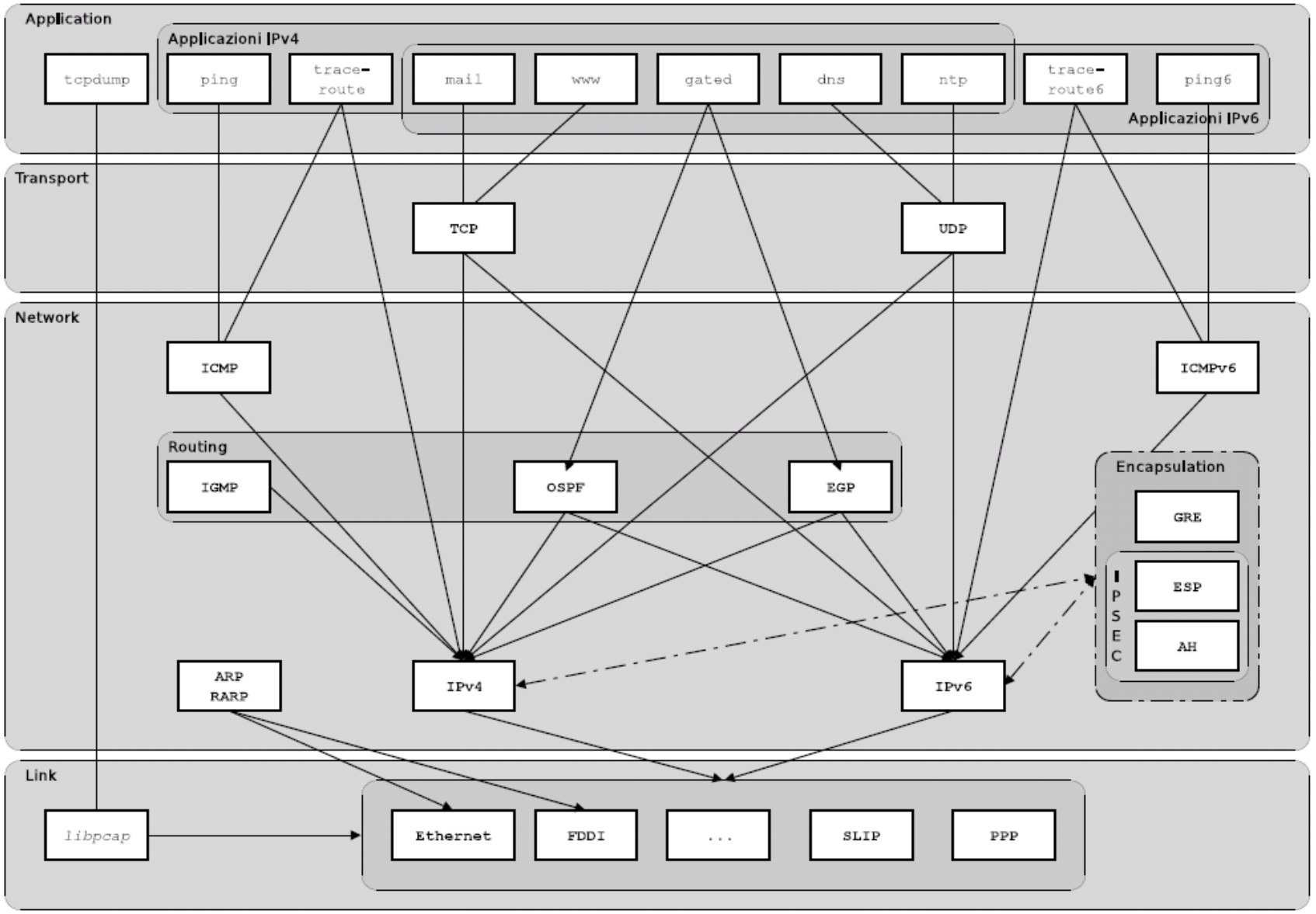


La famiglia dei protocolli TCP/IP

- Nessuna specifica per gli strati sotto IP, in quanto relativi alla singola sottorete.
- IP svolge funzioni di rete e instradamento dei pacchetti
- TCP (o UDP) svolge le funzioni di trasporto e di controllo della connessione end-to-end
- Lo strato di applicazione contiene applicativi utilizzati per fornire servizi all'utente



Quadro generale dei protocolli TCP/IP (da Gapil)



Gli Standard di Internet : Internet Society, RFC

Non esistono veri e propri enti che svolgono la funzione di gestione, ma solo enti di coordinamento delle attività di ricerca e di sviluppo che ora convergono nella **Internet Society**.
Dalla **IS** dipende l'**Internet Advisory Board** (IAB) e si compone di due sottogruppi:

- **Internet Research Task Force (IRTF):** coordina le attività di ricerca
- **Internet Engineering Task Force (IETF):** coordina le attività di ingegnerizzazione ed implementazione
  - IETF pubblica nei **Request For Comment (RFC) -** http://www.ietf.org/rfc.html
  - Tipi di RFC
    - Informational (FYI)
    - Best Current Practice (BCP)
    - Standard (STD) 3 stati : **Proposed Standard, Draft Standard, Standard**

Gli Standard di Internet : ICANN , IANA

ICANN (Internet Corp. for Assigned Names and Numbers - http://www.icann.org/) è l'ente no-profit che assegna gli indirizzi IP e l'identificatore di protocollo e gestisce il DNS di primo livello (Top-Level Domain).
Funzione svolta operativamente da IANA (www.iana.org) che è una sua emanazione.

Data la complessità della gestione, sono state individuate 5 organizzazioni denominate RIR (Regional Internet Registries) che in cooperazione con IANA hanno il compito di gestire le allocazioni a livello continentale.
Le RIR sono: ARIN, APNIC, RIPE, LACNIC e AFRINIC.

Principi architetturali di Internet

Descritti nell'RFC 1958 http://www.ietf.org/rfc/rfc1958.txt, in ordine di importanza:

1) Assicurarsi che funzioni
2) Mantenerlo semplice (nel dubbio, la soluzione più semplice)
3) Fare scelte chiare (se si può fare in diversi modi sceglierne uno)
4) Sfruttare la modularità
5) Aspettarsi l'eterogeneità
6) Evitare opzioni e parametri statici
7) Mirare ad un buon progetto (non necessariamente perfetto)
8) Essere rigorosi nell'invio e tolleranti nella ricezione
9) Pensare alla scalabilità (IPv4 e IPv6..)
10) Considerare le prestazioni e i costi

IP: Lo stato Network in Internet

Interconnette più reti di livello Data-Link (LAN o connessioni punto-punto).
Fornisce uno servizio per il trasporto di datagrammi (pacchetti) tra mittente e destinatario indipendentemente dalle loro reti di appartenenza - http://www.ietf.org/rfc/rfc791.txt
La versione del protocollo IP attualmente in uso, descritta in queste slides, è IPv4.

Operazioni:

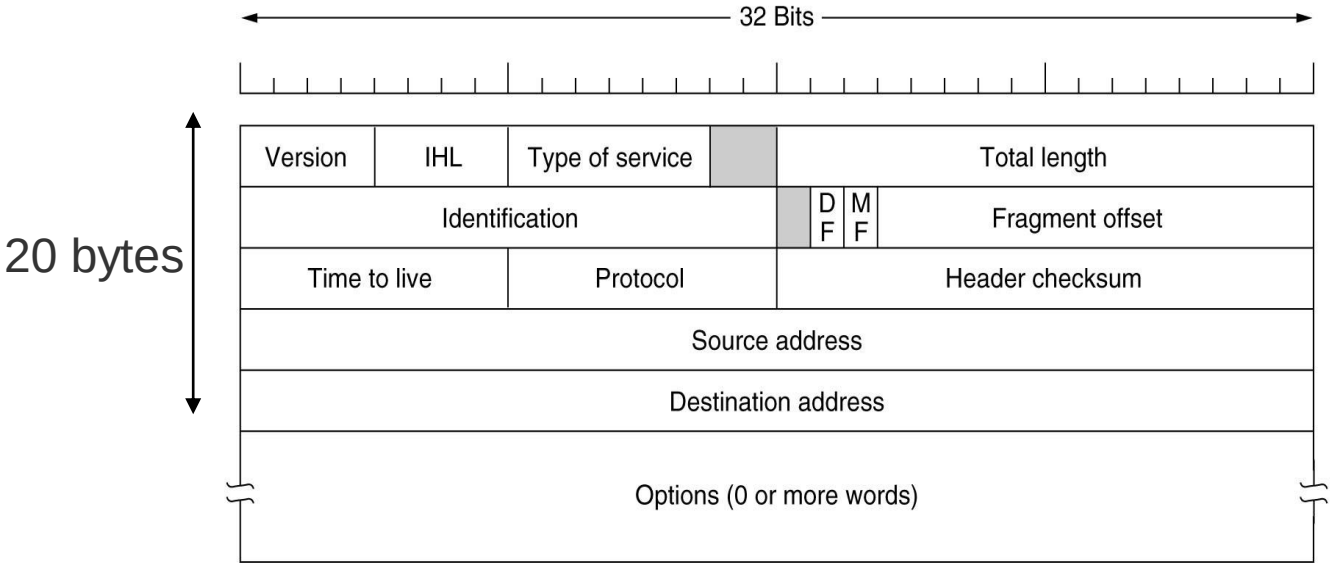
- Lo stato di trasporto prende il flusso di dati e li **divide in datagrammi** che passa allo stato IP. La dimensione massima è di 64KB, ma generalmente vengono scelti datagrammi non superiori a 1500 Byte (per compatibilità con Ethernet).
- Il datagramma di trasporto (detto Segmento) **viene incorporato nella Trama IP** e trasferito da un router all'altro fino a destinazione.
- Il datagramma può subire una **frammentazione** nel caso di passaggio attraverso un livello data-link con dimensione massima (MTU) inferiore. I frammenti vengono riassemblati a destinazione.
  Alcuni MTU (byte): 802.3=1500, 802.11=2312, PPP=576(tipico), FibreChannel=2112
- Il datagramma (eventualmente riassemblato) viene **estratto dalla trama IP** e passato al livello di trasporto, che **ricostruisce il flusso.**
- **QoS**: La consegna è di tipo **"Best Effort"**.

La trama IP (1/3)

Il datagramma IP è costituito dall'intestazione (header) IP seguita dal segmento del livello di trasporto.

		Segmento liv. Trasporto
HeaderLiv2	Header IP	Payload (max 64KB)

L'header ha una parte fissa e una parte opzionale variabile e viene trasmessa in ordine big endian.



http://www.networksorcery.com/enp/protocol/ip.htm

La trama IP (2/3)

- **Version (4 bit):** i primi 4 bit di ogni pacchetto IP contengono il numero di versione.
- **HLEN (4 bit) :** dimensione dell'header espressa in parole di 4 byte (da 5 a 15)
- **Type of Service (6 bit):**
  - Inizialmente per controllo della rete (priorità e segnalazioni).
  - Con l'RFC 2474 diventa Servizi Differenziati per la codifica delle Classi di Servizio.
  - In realtà Internet è "best effort": questo campo è quasi sempre inutilizzato.
- **Total Lenght (16 bit):** Numero di byte totali header+dati (fino a 64K)
- **Identification (16 bit):** Tutti i frammenti di datagramma hanno lo stesso valore
- **DF (1 bit):** Don't Fragment → ordina ai router di non frammentare
- **MF (1 bit):** More Fragments → 1 per tutti i frammenti tranne l'ultimo
- **Fragment Offset (13 bit):** Indica la posizione del frammento nel datagramma corrente, espressa in blocchi di 8 byte (max. 8192 frammenti).
- **TTL (8 bit):** Numero max di salti; si decrementa ad ogni passaggio. Quando arriva a 0 il pacchetto viene eliminato.

La trama IP (3/3)

- **Protocol (8 bit):** Protocollo di livello superiore (ICMP=1,TCP=6, UDP=17, ..)
- **Header Checksum (16 bit):** Checksum dell'**header**
  - ricalcolato da ogni router , perché il TTL cambia ad ogni salto.
  - Aiuta a rilevare errori generati da locazioni di memoria difettose nei router.
  - Somma tutte le sequenze di 16 bit (con l'aritmetica del complemento a 1) e poi prende il complemento a 1 del risultato.
- **Source e Destination Address (32+32 bit):** Indirizzi di sorgente e destinazione
- **Options**: Pensato per poter aggiungere estensioni non previste.

  Lista completa: http://www.iana.org/assignments/ip-parameters

  Formato: Opt. code (1 byte) - Opt. Length (1byte) - Opt. Data (n Byte)
  - 0 - End of Option List
  - 130 - Security: lista di reti vietate, non usato.
  - 7 - Route record: ogni router aggiunge il proprio indirizzo
  - 68 - Time Stamp: ogni router aggiunge il proprio indirizzo e data/ora.
  - 137 - Source routing: lista dei router da percorrere
- **Padding:** bit aggiunti per rendere il campo Options multiplo di 32 bit.

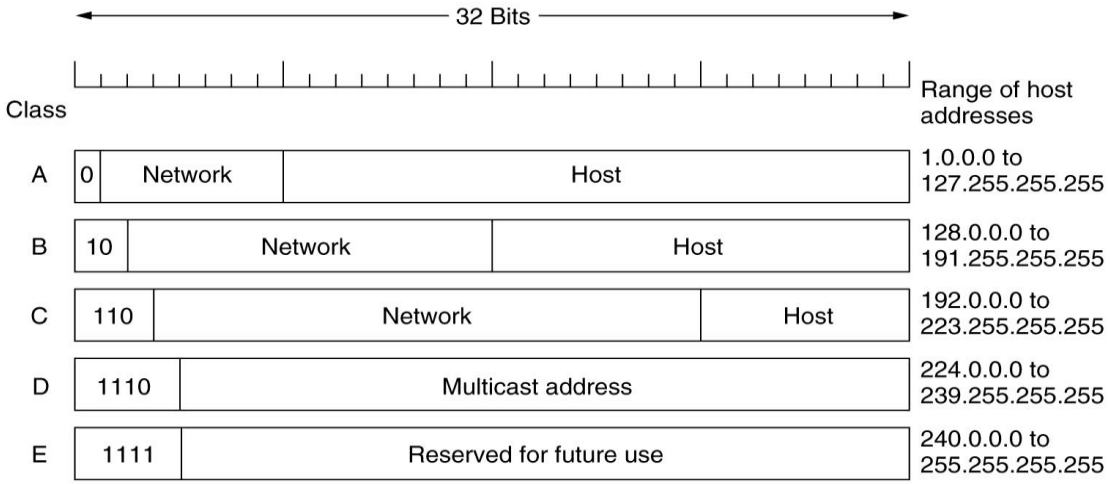
Indirizzi IP

Indirizzi a 32 bit con notazione "**dotted decimal**": 4 decimali (0-255) separati da punto
Esempio: 0x89CCD401 = 10001001.11001100.11010100.00000001 -> **137.204.212.1**
Numero max. di indirizzi 2^{32} = 4.294.967.296
Per motivi di routing la sequenza è suddivisa in due parti:

– **NETid**: Identifica una Rete liv.2 Utilizzata dai router per l'instradamento dei pacchetti

– **HOSTid**: Distingue gli Host della stessa Rete.

Classfull Addressing:



Classi IP

Classe	bit-iniziali	inizio	fine	indirizzi	default-mask	CIDR-equiv	reti	host
A	0	0.x.x.x	127.x.x.x	2G	255.0.0.0	/8	126	16M
B	10	128.x.x.x	191.x.x.x	1G	255.255.0.0	/16	16K	64K
C	110	192.x.x.x	223.x.x.x	0.5G	255.255.255.0	/24	2M	254
D	1110	224.x.x.x	239.x.x.x	0.25G	Multicast			
E	1111	249.x.x.x	255.x.x.x	0.25G	Reserved			

La numerazione è gestita da IANA che delega gerarchicamente alle RIR:

Vedi ad esempio il comando: whois 160.78.0.0

Indirizzi IP di rete e di Broadcast

< network >	00000000000000000000
< network >	11111111111111111111

Address of the network
Broadcast of a specific network

Se la parte Host è di N bit, il numero di indirizzi effettivamente assegnabili agli host è **2^N-2**, poiché il primo indirizzo (tutti zeri nella parte host) identifica la rete, mentre l'ultimo indirizzo (tutti uni nella parte host) è l'indirizzo di broadcast.

Indirizzi IP per uso privato

Le seguenti reti sono riservate da ICANN per uso privato (Intranet) e gli indirizzi non possono essere annunciati dai Router (http://www.ietf.org/rfc/rfc1918.txt):

Classi	inizio	Fine	indirizzi
1 classe A	10.x.x.x		16M
16 classi B	172.16.x.x	172.31.x.x	1M
255 classi C	192.168.0.x	192.168.255.x	64K

LOOPBACK

La rete 127.0.0.0/16 è riservato per il loopback (RFC 3330).
Per convenzione su ogni host viene definita una interfaccia virtuale di loopback con indirizzo IP predefinito 127.0.0.1, con nome **localhost**, che consente la comunicazione TCP/IP tra due processi locali senza il coinvolgimento di interfacce fisiche.

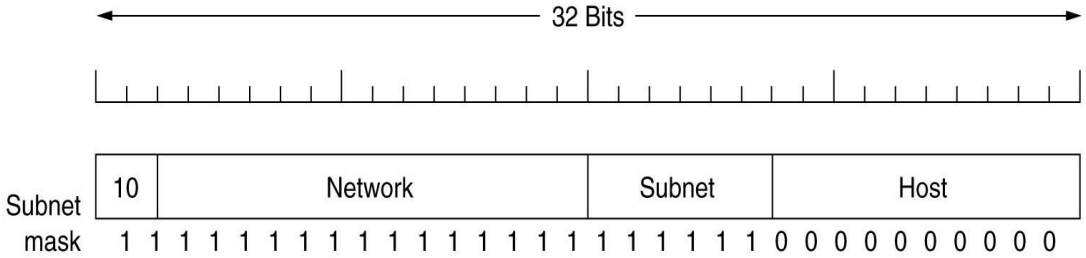
ZEROCONF

La rete 169.254.0.0/16 (IPv4 link-local) è utilizzata dal servizio Zeroconf (http://www.ietf.org/rfc/rfc3927.txt) per assegnare un indirizzo IP agli host di una LAN senza dipendere da una infrastruttura, ovvero quando non è possibile ottenere un indirizzo dinamico da un server DHCP.

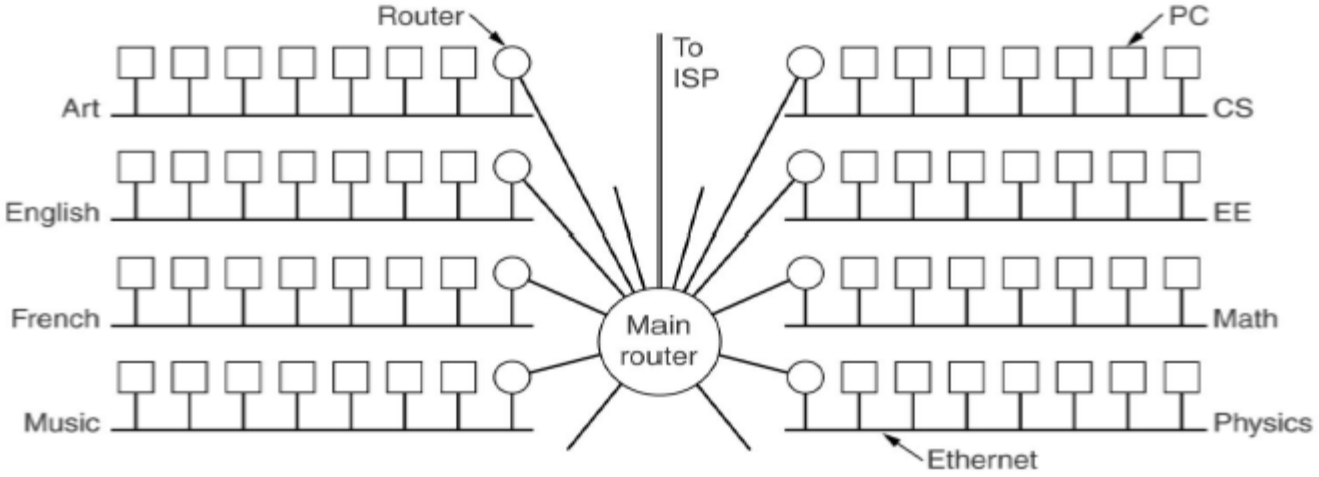
IP subnetting

Consente un ulteriore livello di gerarchia per gli indirizzi IP: NET-SUBNET-HOST
Il **NETMASK** è un parametro di 32 bit che stabilisce la suddivisione:

- Bit a 1 in corrispondenza del campo NET o SUBNET
- Bit a 0 in corrispondenza del campo HOST



Esempio: rete di classe B 160.78.0.0 partizionata in 256 Subnet da 256 indirizzi:
NETMASK 255.255.255.0 -> 11111111 11111111 11111111 00000000



CIDR – Classless Inter-Domain Routing

- Il numero di indirizzi IP (4G) è insufficiente
- Molte reti di classe B usano meno 50 indirizzi

CIDR http://www.ietf.org/rfc/rfc1519.txt
- soluzione temporanea in attesa di IPv6
- Assegna gli indirizzi IPv4 rimanenti in blocchi di dimensione variabile nella forma
  netaddress/NetMaskBit
- routing più complicato (tabelle lunghe)

Il **Supernetting** (route aggregation) consente di accorpare più reti contigue come fossero un'unica rete, per ottimizzare i tempi di routing

- In caso di sovrapposizioni tra 2 reti vince la netmask più lunga

no. of addrs	bits	pref	mask
1	0	/32	255.255.255.255
2	1	/31	255.255.255.254
4	2	/30	255.255.255.252
8	3	/29	255.255.255.248
16	4	/28	255.255.255.240
32	5	/27	255.255.255.224
64	6	/26	255.255.255.192
128	7	/25	255.255.255.128
256	8	/24	255.255.255
512	9	/23	255.255.254
1 K	10	/22	255.255.252
2 K	11	/21	255.255.248
4 K	12	/20	255.255.240
8 K	13	/19	255.255.224
16 K	14	/18	255.255.192
32 K	15	/17	255.255.128
64 K	16	/16	255.255
128 K	17	/15	255.254
256 K	18	/14	255.252
512 K	19	/13	255.248
1 M	20	/12	255.240
2 M	21	/11	255.224
4 M	22	/10	255.192
8 M	23	/9	255.128
16 M	24	/8	255
32 M	25	/7	254
64 M	26	/6	252
128 M	27	/5	248
256 M	28	/4	240
512 M	29	/3	224
1024 M	30	/2	192

Instradamento dei Datagrammi

La rete di appartenenza di un Host è fondamentale per determinare la modalità di consegna, che può essere **diretta** o **indiretta**.

Direct delivery : host sorgente e destinatario condividono la stessa rete.

- trova l'indirizzo fisico del **destinatario** (con ARP) che associa all'IP del destinatario
- inoltra il pacchetto al livello Link indirizzando il destinatario:

```
1) ARPrequest   to:Broadcast from: MACmitt          Who has IPdest?
2) ARPreply     to:MACmitt   from: MACdest
3) Send IP      to:MACdest   from: MACmitt          toIP:dest, fromIP:mitt
```

Indirect delivery : sorgente e destinatario appartengono a reti IP diverse

- individua il router da contattare consultando la propria **Tabella di Routing**
- trova l'indirizzo fisico del **router** (con ARP) che associa all'IP del destinatario
- inoltra il pacchetto al livello Link indirizzando il router.

```
1) ARPrequest     to:Broadcast from: MACmitt        Who has IProuter?
2) ARPreply       to:MACmitt   from: MACrouter
3) Send IP        to:MACrouter from: MACmitt        toIP:dest, fromIP:mitt
```

La scelta del tipo di consegna avviene consultando la tabella di routing locale.

Tabella di routing

E' una tabella che contiene le destinazioni e i percorsi per raggiungerle.

Esempio di tabella di routing (comando “route” di linux) per l'host 160.78.124.1:

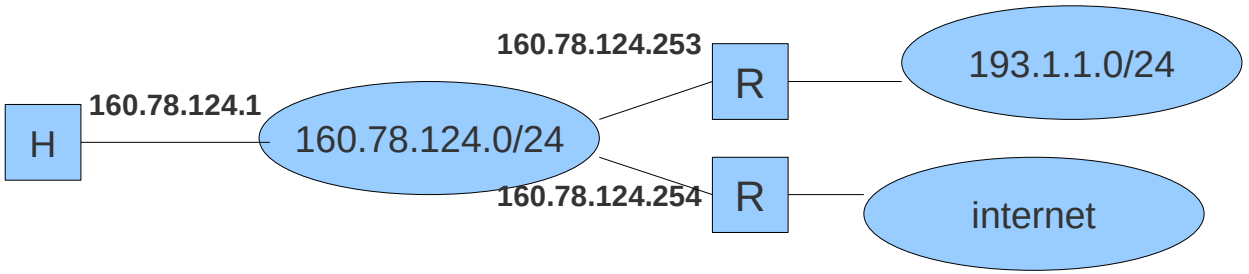
Destination	Router	Mask	Interface
160.78.124.0	*	255.255.255.0	eth0 (consegna diretta)
193.1.1.0	160.78.124.253	255.255.255.0	eth0 (consegna indiretta)
default	160.78.124.254	0.0.0.0	eth0 (consegna indiretta)

La prima riga (Router *) indica che gli host della rete 160.78.124.0/24 vengono raggiunti in consegna diretta.

La seconda riga (Router 160.78.124.253) indica che gli host della rete 193.1.1.0/24 sono raggiunti in modalità indiretta tramite il router 160.78.124.253

Generalmente le reti locali hanno al proprio interno un router di riferimento (indicato come “**Default Router**”) a cui vengono consegnate tutte le destinazioni non note.

Nell'esempio tutte le destinazioni diverse da 160.78.124.0/24 e 192.1.1.0/24 vengono consegnate al router 160.78.124.254.



Ricerca nella tabella

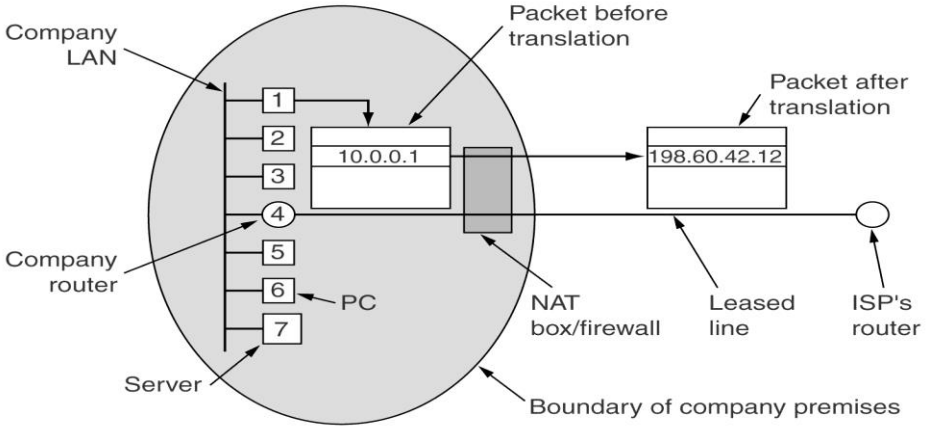
La ricerca avviene utilizzando
- l'IP di destinazione (IPdest)
- La rete di destinazione e Netmask (Mask) di ciascuna riga della tabella

Procedura: IPdest AND Mask

- Se il risultato coincide con la rete presa in esame la riga è quella giusta
- Una volta trovato il risultato il lookup si ferma e il datagramma viene instradato
- Se nessuna riga corrisponde si usa il router di default

NAT (Network Address Translation)

NAT (http://www.ietf.org/rfc/rfc3022.txt): dispositivo che consente agli host di una LAN (con indirizzi privati) di comunicare in Internet utilizzando un solo indirizzo pubblico.
La linea verso Internet possiede un indirizzo IP pubblico e viene visto dagli host della LAN come Default Router.
Le operazioni del NAT sono distinte in base alla direzione:
SNAT (Source-NAT) è la funzionalità che consente di manipolare l'indirizzo sorgente ed è tipicamente utilizzato per consentire ai pacchetti di una LAN privata di uscire in internet. Quando un Host della LAN si rivolge al NAT per uscire in Internet il NAT trasforma l'indirizzo del mittente IP nell'indirizzo IP pubblico del NAT, quindi contatta il destinatario.
DNAT (Destination-NAT) è utilizzata per manipolare l'indirizzo di destinazione.
E' usata tipicamente per dirottare verso una destinazione interna (con indirizzo privato) i pacchetti provenienti da Internet.



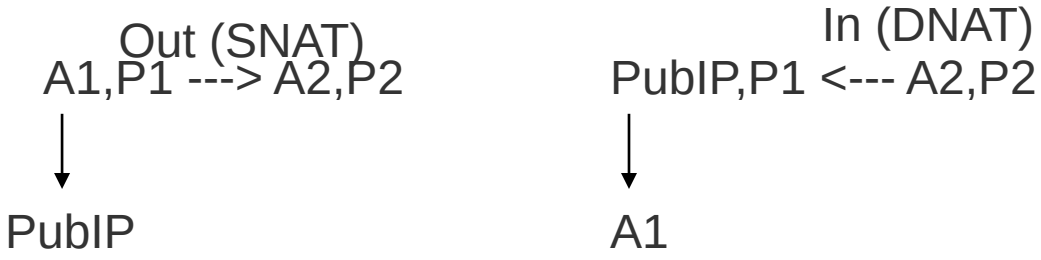
Le tabelle del NAT

Le manipolazioni SNAT e DNAT sono rappresentate in tabelle che vengono consultate per ogni pacchetto che attraversa il NAT. Le entry delle tabelle possono essere statiche o dinamiche.

NAT Dinamico

Cl. interno: A1,P1

Serv esterno: A2,P2

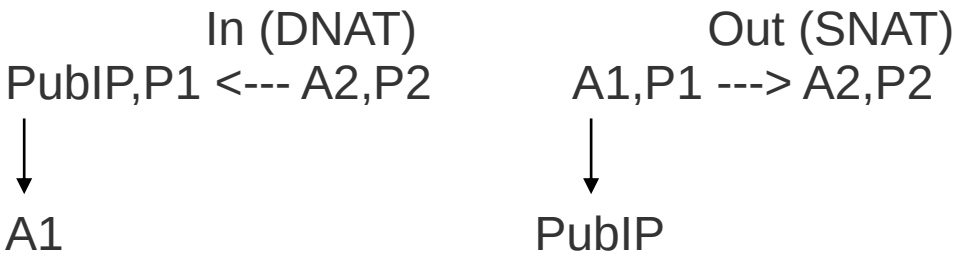


Quando un client della LAN si rivolge al NAT per contattare un server esterno, il NAT genera una entry dinamica associando IP/porta del client con la IP/porta del server quindi applica **SNAT**. L'entry viene utilizzata per il DNAT sulla risposta del server.

NAT Statico

Cl. esterno: A2,P2

Serv interno: A1,P1



Se vogliamo avere un server interno che deve essere contattato da un client esterno dobbiamo istruire il NAT mediante una entry statica che associa una porta del NAT con IP/porta del server interno.

Quando un client esterno contatta il NAT sulla porta viene consultata la entry statica e applicato **DNAT**. Viene inoltre creata una entry dinamica che verrà utilizzata per applicare SNAT sulla risposta.

Problemi dell'architettura NAT

- Violazione dell'univocità degli indirizzi: migliaia di Host usano gli stessi indirizzi privati.
- Sicurezza: è difficile tracciare l'identità dell'indirizzo IP pubblico.
- IP non è più connection-less
- IP non è più stratificato: Il Layer IP non dovrebbe entrare nei layer superiori
- Un guasto al NAT pregiudica tutte le connessioni che lo attraversano.

In realtà NAT ha avuto una grande diffusione e ha ridotto la spinta verso IPv6.

Protocollo ARP

Ogni interfaccia di rete di un nodo (Ethernet, LAN Wireless, Seriale, ecc) possiede un indirizzo fisico e, se utilizzata in internet, almeno un indirizzo IP.
Il protocollo **ARP (Address Resolution Protocol)** ha il compito di determinare l'indirizzo fisico di un nodo IP.
Quando un nodo mittente deve contattare un destinatario in **Direct Delivery** (Terminale o Router) di cui conosce solo l'indirizzo IP utilizzerà il protocollo ARP:

- Il nodo sorgente invia un pacchetto (**ARP Request**) con destinazione Broadcast sulla LAN, contenente l'indirizzo IP del destinatario.
- I terminali con indirizzo IP diverso ignoreranno il Pacchetto, mentre il nodo in oggetto risponderà (**ARP Replay**) con un Unicast inviando il proprio indirizzo fisico.
- Ogni host mantiene una tabella (**ARP Cache**) con le corrispondenze ottenute (comando arp –a). Ogni entry ha un tempo di vita tipicamente di 20 minuti.

Il Frame ARP contiene:

- Un campo codice 1=ARPrequest, 2=ARPreplay
- indirizzo IP e Indirizzo HW di partenza e destinazione

Protocollo RARP

In determinate situazioni alcuni nodi IP al momento dell'attivazione della rete non conoscono il loro indirizzo IP (ad esempio perché non hanno memoria permanente).
Esistono diverse soluzioni, tra cui **RARP** (Reverse ARP) - http://www.ietf.org/rfc/rfc903.txt
E' un protocollo ideato da SUN per risolvere il problema.
Il client invia in modalità Broadcast la richiesta:
"Questo è il mio indirizzo MAC: xx-xx-xx-xx-xx-xx, Qualcuno conosce il mio indirizzo IP?".
Un server RARP, con la tabella MAC-IP , risponderà con l'informazione richiesta.
Svantaggi:

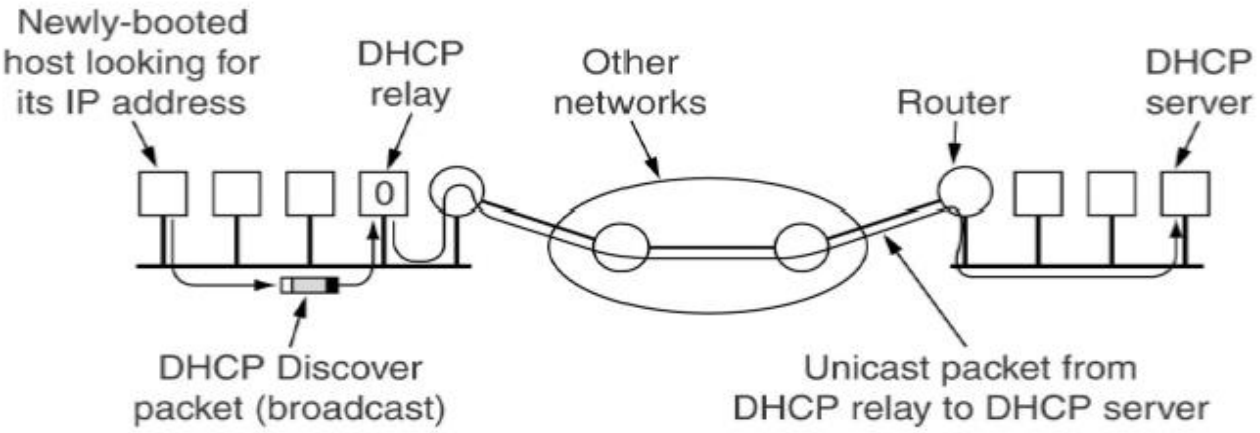
- la richiesta Broadcast non passa i router
- le associazioni MAC-IP sono statiche
- non sono previste altre informazioni

RARP è reso obsoleto dal suo successore DHCP.

Protocollo DHCP

DHCP (Dynamic Host Configuration Protocol, rfc2131.txt e rfc2132.txt) risolve lo stesso problema di RARP aggiungendo nuovi servizi.

- Il server DHCP può fornire più informazioni al client: indirizzo IP, NetMask, Default Router, DNS server, NTP server, ecc.
- L'indirizzo IP fornito può essere statico o dinamico (assegnato al momento della richiesta sulla base di un pool di indirizzi disponibili)
- Il server può risiedere in una LAN diversa dalla LAN del client (tramite relay)



E' un protocollo applicativo: utilizza la porta 67/UDP per il server e la 68/UDP per il client.

Funzionamento del DHCP

1) Il client DHCP invia in modalità broadcast un pacchetto **DHCP Discover**

`0.0.0.0:68 -> 255.255.255.255:67`

2) Il server risponde (tramite l'eventuale Agent) un pacchetto **DHCP Offer** che contiene l'indirizzo richiesto più eventuali altre informazioni.

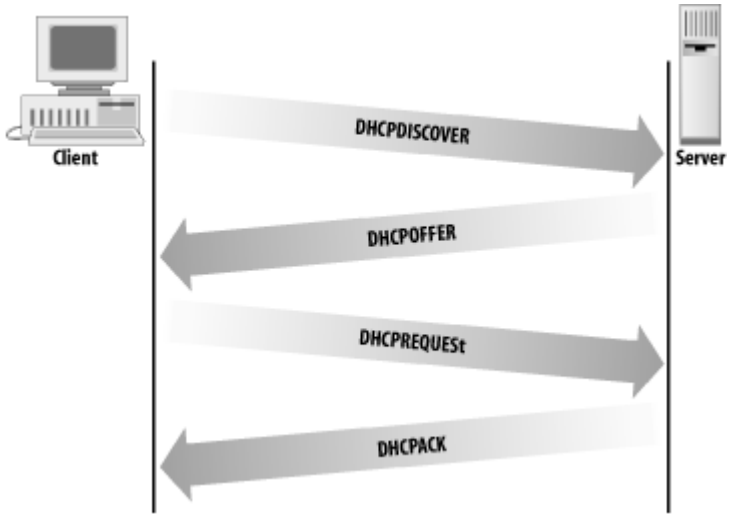
`Ipserver:67 -> Ipclient:68`

3) Il client accetta la prima risposta che ottiene e invia in Broadcast un **DHCP Request** in cui dice da quale server ha ricevuto l'indirizzo.

`0.0.0.0:68 -> 255.255.255.255`

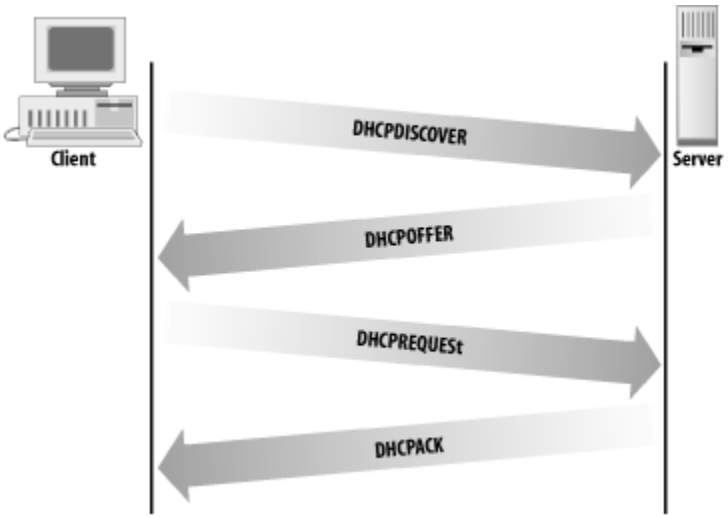
4) Infine il server manda un **DHCP ACK** al client per conferma.

`Ipserver:67 -> Ipclient:68`



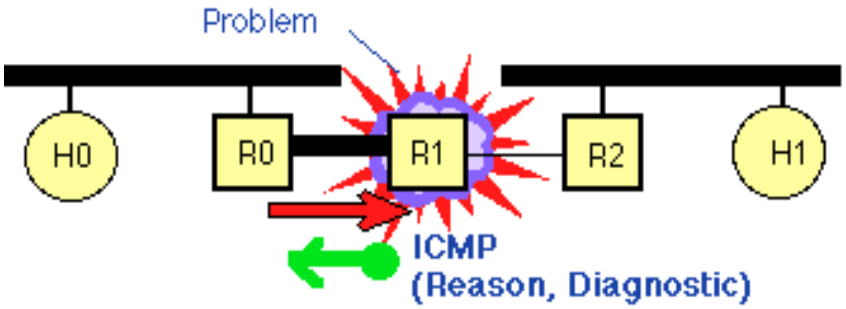
Funzionamento del DHCP: Rinnovo

Il server gestisce una tabella in cui gli indirizzi IP possono essere associati staticamente a indirizzi MAC oppure possono essere “**affittati”** dinamicamente al momento della richiesta. Il “leasing” ha un termine; il client deve chiederne un eventuale rinnovo, altrimenti l’indirizzo viene ritirato ed assegnato ad un altro client.
Le operazioni **DHCPREQUEST/DHCPACK** vengono ripetute per prolungare l'assegnazione dell’indirizzo. La richiesta avviene con 3 tentativi: 2 volte al 50% del tempo utilizzato e un’ultima volta all’ 87,5%



Protocollo ICMP

ICMP - Internet Control Message Protocol (http://www.ietf.org/rfc/rfc792.txt) è un protocollo di servizio di IP per lo scambio di messaggi di errore o di controllo che consentono agli Host e ai Router di accorgersi di eventuali malfunzionamenti della rete. Spedisce i messaggi di notifica dell'errore sempre al mittente del datagramma per il quale si è verificato l'errore.

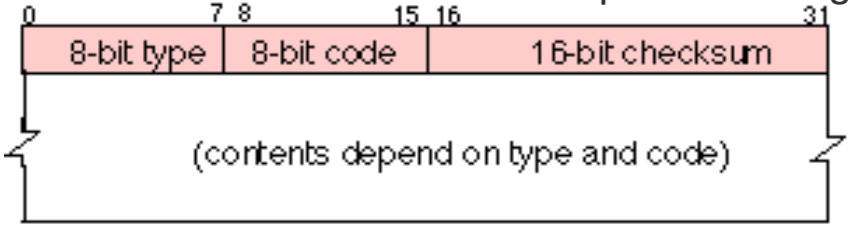


Formato del Frame ICMP

Il formato del frame è costituito da una intestazione e da un'area dati. La prima è composta da tre campi:

- **TIPO** è un numero di 8 bit che identifica il messaggio Tipi principali:
  - 0 = Risposta di ECHO
  - 3 = Destinazione irraggiungibile (esempio datagramma troppo grande, ma DF settato)
  - 4 = Rallentamento della sorgente (Il router informa che il pacchetto è stato eliminato e che la sorgente deve rallentare)
  - 8 = Richiesta di ECHO (comando ping)
  - 11 = TTL scaduto per un datagramma
  - 12 = Problema di parametri (argomento di un opzione scorretto)
  - 13 = Richiesta di contrassegno temporale (per sincronizzare gli orologi)
  - 14 = Risposta di contrassegno temporale
- **CODICE:** Info aggiuntive. Ad esempio se il Tipo è 3 il Codice dice qual'è il tipo di errore
- **CHECKSUM**, di 16 bit, è il CRC del frame ICMP (header+data)

L'area Dati varia in funzione del tipo di messaggio.



Il Frame ICMP è inserito direttamente nel payload di IP:

Header IP	ICMP

UNIVERSITÀ DI PARMA

DIPARTIMENTO DI SCIENZE MATEMATICHE, FISICHE ED INFORMATICHE
Corso di Laurea in Informatica

Il Livello Netowork

Parte II : IPv6

RETI DI CALCOLATORI - a.a. 2022/2023

Roberto Alfieri

Il livello Network: sommario

PARTE I

- Scopi del livello Network
- Commutazione di circuito e di pacchetto
- La famiglia dei protocolli TCP/IP
- Il protocollo IP: trama indirizzi, instradamento
- Protocolli di servizio: ARP, ICMP, DHCP

PARTE II

- IPv6

PARTE III

- Routing: Algoritmi e protocolli. Distance Vector e Link State.

IPv6

Necessità di un nuovo layer IP:

- Supportare molti miliardi di host
- Semplificare il routing per avere backbone veloci
- Offrire meccanismi di sicurezza
- Offrire qualità di servizio (multimedialità)
- Gestire bene multicast e broadcast
- Consentire la mobilità
- Consentire future evoluzioni e garantire compatibilità col passato

Nel 1993 tra varie proposte venne scelta SIPP (Simple Internet Protocol Plus) che prese il nome di **IPv6.**

Indirizzi IPv6

Spazio degli indirizzi grande a sufficienza (16 byte → 128 bit)
Notazione: 8 quaterne di numeri esadecimali separati da ":"
Esempio: 8000:0000:0000:0000:0562:CDAF:2DAF:0001
Notazione compatta: è possibile omettere gli zeri iniziali di ogni quaterna.
Gruppi di 4 zeri possono essere sostituiti con ::
Esempio precedente: 8000::562:CDAF:2DAF:1
Gli indirizzi IPv4 possono essere compresi tra gli indirizzi IPv6 con un prefisso di 96 zeri, mantenendo la notazione dotted decimal. Esempio: ::192.31.20.46

Indirizzi Broadcast: non esistono in IPv6
Indirizzi Speciali: Loopback (127.0.0.1 di IPv4) ::1
Indirizzi Multicast: Indirizzi assegnati a più interfacce (come IPv4)
Indirizzi Anycast (novità): Sono indirizzi assegnati a più interfacce.
Il pacchetto anycast viene consegnato solo all'interfaccia più vicina

Gli indirizzi IPv6 in URL devono essere scritti tra parentesi quadre. Esempio:
http://[2001:1:4F3A:206:AE14]:8888/index.html

Architettura degli indirizzi

Prefix	Hex	Size	Allocation
0x0000 0000 0000 0000 0000 0000		**2**	**Ipv4 compatible**
0000 0000	0000-00FF		Reserved
0000 0001	0100-01FF		Unassigned
0000 001	0200-03FF	2	NSAP
0000 010	0400-05FF		Unassigned
0000 011	0600-07FF		Unassigned
0000 1	0800-0FFF		Unassigned
0001	1000-1FFF		Unassigned
001	**2000-3FFF**	**2**	**IANA to registers**
010,011,100,101,110	4000-CFFF		Unassigned
1110	D000-EFFF		Unassigned
1111 0	F000-F7FF		Unassigned
1111 10	F800-FBFF		Unassigned
1111 110	FC00-FDFF		Unassigned
1111 1110 0	FE00-FE7F		Unassigned
1111 1110 10	**FE80-FEBF**	**2**	**Link-local**

IPv6: Indirizzi Global Unicast

IANA 2000::/3
Gli indirizzi Unicast globali di IPv6 hanno prefisso 001 (2000::/3) e sono gestiti da IANA. IANA ha frammentato questo spazio in diverse reti più piccole che ha poi assegnato in gestione alle RIR continentali (APNIC, ARIN, RIPE, LACNIC e AFRINIC)
http://www.iana.org/assignments/ipv6-unicast-address-assignments/ipv6-unicast-address-assignments.xhtml

RIPE 2001:600::/23
La rete 2001:0600::/23 (2001:06xx: e 2001:07xx:) è stata assegnata a RIPE, che ha suddiviso in reti più piccole (tipicamente /32). http://www.ripe.net/ripe/docs/ripe-510#2e

GARR 2001:760::/32
La rete 2001:760::/32 è stata assegna da RIPE a GARR, che ha suddiviso in reti più piccole (/48).

UNIPR 2001:760:2E04::/48 - INFN-Parma 2001:760:4207::/48
Il GARR ha assegnato la rete 2001:760:2E04::/48 a UNIPR e la rete 2001:760:4207::/48 a INFN-Parma. UNIPR dispone di 64K reti /64 da ripartire alle proprie strutture.

IPv6: possibile ripartizione della rete in UNIPR

Rete di ateneo: 2001:760:2e04::/48

	2001:760:2e04:0000::/52	0000 --> 0fff	4k Reti /64 destinate ad usi futuri
1	2001:760:2e04:1000::/52	1000 --> 1fff	4k Reti /64 destinate alla Sede 1 (Campus)
2	2001:760:2e04:2000::/52	2000 --> 2fff	4k Reti /64 destinate alla Sede 2 (Sede Centrale)
3	2001:760:2e04:3000::/52	3000 --> 3fff	4k Reti /64 destinate alla Sede 3 (Via Gramsci)
4	2001:760:2e04:4000::/52	4000 --> 4fff	4k Reti /64 destinate alla Sede 4 (B.go Carissimi)
5	2001:760:2e04:5000::/52	5000 --> 5fff	4k Reti /64 destinate alla Sede 5 (Via Kennedy/via D'Azeglio)
6	2001:760:2e04:6000::/52	6000 --> 6fff	4k Reti /64 destinate alla Sede 6 (S. Francesco)
7	2001:760:2e04:7000::/52	7000 --> 7fff	4k Reti /64 destinate alla Sede 7 (Momentaneamente dismessa)
8	2001:760:2e04:8000::/52	8000 --> 8fff	4k Reti /64 destinate alla Sede 8 (Via del Taglio)
9	2001:760:2e04:9000::/52	9000 --> 9fff	4k Reti /64 destinate Alla Sede 9 (Via Volturno)

Le sedi in servizio denominate 10 11 12 e 13 sono sedi con un numero di host allocati sensibilmente inferiore a 100 è quindi plausibile supporre che non abbiano grosse esigenze di indirizzamento futuro pertanto si propone di continuare l'allocazione con il seguente schema

2001:760:2e04:a000::/52

10 2001:760:2e04:a100::/60 a100 --> a10f 16 Reti /64 destinate alla Sede 10 (Via S. Michele)
11 2001:760:2e04:a110::/60 a110 --> a11f 16 Reti /64 destinate alla Sede 11 (via Farini)
12 2001:760:2e04:a120::/60 a120 --> a12f 16 Reti /64 destinate alla Sede 12 (Pilotta)
13 2001:760:2e04:a130::/60 a130 --> a13f 16 Reti /64 destinate alla Sede 13 (Paradigna)
14 2001:760:2e04:a140::/60 a140 --> a14f 16 Reti /64 destinate alla Sede 10 (Beni Teatrali)
15 2001:760:2e04:a150::/60 a seguire per le altri sedi attive
16 2001:760:2e04:a160::/60 a seguire per le altri sedi attive
17 2001:760:2e04:a170::/60 a seguire per le altri sedi attive
18 2001:760:2e04:a180::/60 a seguire per le altri sedi attive
19 2001:760:2e04:a190::/60 a seguire per le altri sedi attive

InterfaceID

Le reti assegnate alle strutture per le reti locali sono quindi di 64 bit.

Gli ultimi 64 bit dell'indirizzo IPv6 possono essere assegnati in vari modi:

- Assegnati via DHCPv6
- Configurati manualmente
- Autogenerati con numeri pseudo-random
- Autoconfigurati utilizzando **l'interfaceID**, ovvero una sequnza di 64 bit, univoci di ogni interfaccia di rete, ottenuta partendo dai 48 bit del MAC address

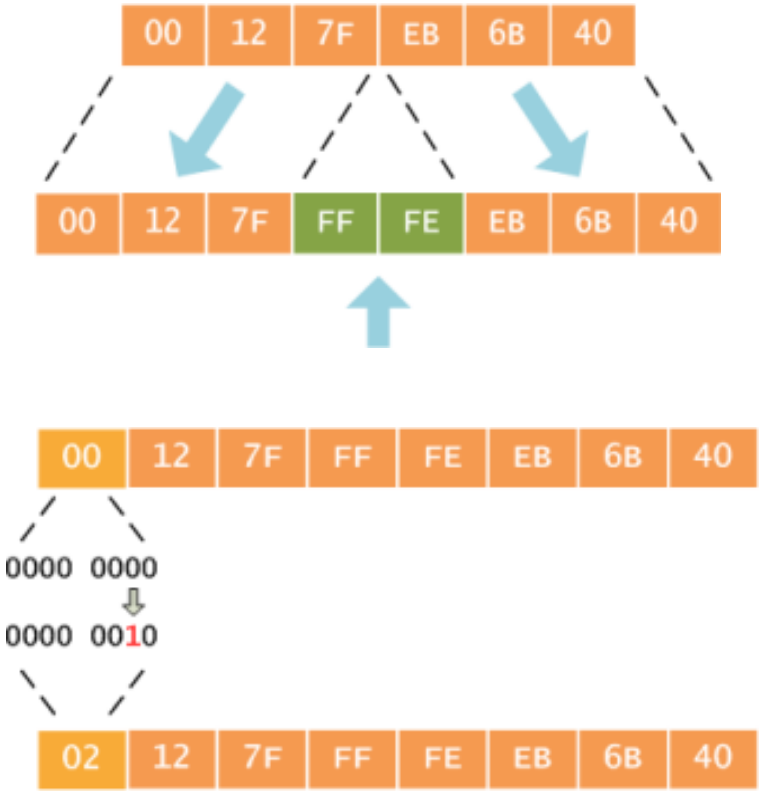
Da MAC48 a InterfaceID

- Gli indirizzi MAC 48 bit utilizzati da Ethernet (MAC48) sono gestiti da IEEE e non si esauriranno prima del 2100.
- IEEE gestisce anche una numerazione a 64 bit, EUI64 (Extended Unique Identifier). La numerazione MAC48 è integrata in EUI64 inserendo 16 bit (FFFE) al centro.

- L' Interface-ID utilizzata per gli indirizzi IPv6

  è una versione modificata di EUI64 (mEUI64) in cui si pone ad 1 il bit 7 di EUI64.

Esempio:
MAC48: 00-12-7F-EB-6B-40
EUI64: 0**0**-12-7F-**FF-FE**-EB-6B-40
mEUI64: 0**2**-12-7F-**FF-FE**-EB-6B-40 (intID)



IPv6: Indirizzi Link-Local

Per Link si intende una rete di livello 2 (LAN o punto-punto). Nodi sullo stesso link sono detti Neighbor (vicini)

Indirizzo Link locale: Destinati ai terminali della stessa rete locale.

Hanno come prefisso 1111 1110 10 (Ad esempio: FE80:)

I pacchetti con questa destinazione non attraverseranno mai un router.

E' un tipo di indirizzo attribuito inizialmente alle interfacce IPv6 con configurazione automatica e viene utilizzato per il processo di Neighbor Discovery.

La configurazione automatica ha il seguente formato:

FE80:0000:0000:0000:xxxx:xxxx:xxxx:xxxx
--interfaceID---

IPv6: Indirizzi Site-Local

Un **Site** è un gruppo di Link gestiti da un'unica autorità (esempio Campus).

Gli indirizzi Site-local sono indirizzi per **uso privato**, analoghi alle reti 10.0.0.0/8, 172.16.0.0/12, e 192.168.0.0/16 di IPv4.
Hanno come prefisso 1111 1110 11 (ad esempio FEC0:)
Rispetto a un indirizzo Link-local cambia il prefisso di formato, aggiungendo la possibilità e la convenienza di suddividere lo spazio di indirizzi in sottoreti. A differenza dagli indirizzi Link-local non sono configurati automaticamente.

IPv6: Indirizzi Multicast e Anycast

Un indirizzo **IPv6 multicast** serve a identificare e a raggiungere un gruppo di nodi simultaneamente. Normalmente il multicast non viene propagato dai router a meno di configurazioni specifiche.
Il prefisso di formato è 1111 1111 (ovvero FF) a cui seguono 4 bit di opzione, 4 bit di ambito e 112 bit per identificare il gruppo.
Vedi http://wwwcdf.pd.infn.it/AppuntiLinux/introduzione_a_ipv6.htm

Gli **indirizzi anycast** sono degli indirizzi con le caratteristiche di quelli unicast che, in base al contesto, sono attribuiti a più interfacce di rete differenti, appartenenti ad altrettanti componenti di rete distinti.
Anycast deve essere supportato dai router, i quali devono gestire il fatto che lo stesso indirizzo IP viene annunciato da luoghi differenti.
L'indirizzo anycast più comune è quello che serve a raggiungere simultaneamente tutti i router nell'ambito link-local.

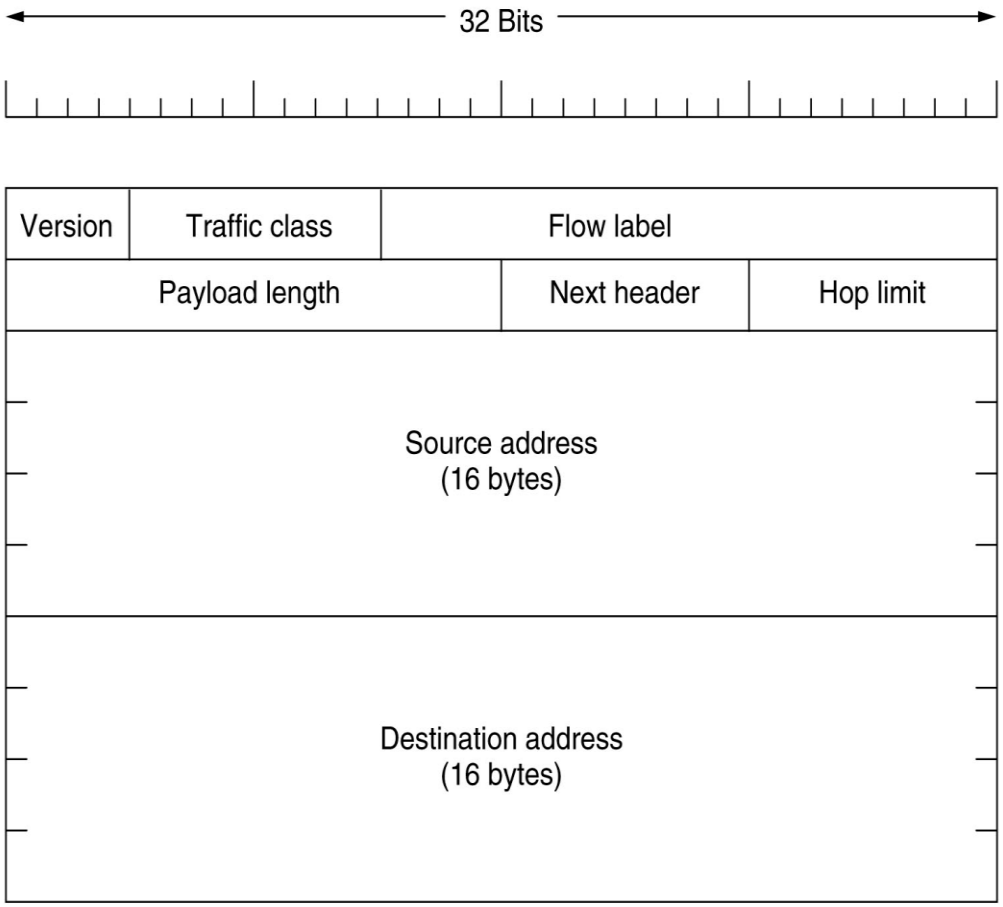
Altri Indirizzi IPv6

Loopback
0:0:0:0:0:0:0:1 (oppure ::1) identifica lo stesso nodo, come 127.0.0.1 in IPv4
Per controllare se lo stack IPv6 funziona: ping6 ::1

IPV4 compatible
Permettono di inserire indirizzi IPv4 in indirizzi IPv6 anteponendo 96 zeri:
Esempio: 10.0.0.1 -> ::A001
vale anche la notazione ::10.0.0.1

Utilizzati per la transizione IPv4-IPv6

La trama IPv6



Cosa è stato eliminato da IPv4

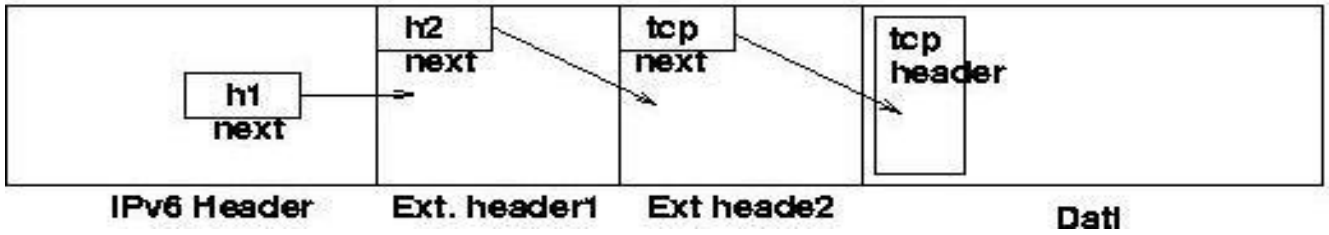
- La frammentazione è stata rimossa perché IPv6 determina dinamicamente la dimensione del datagramma (Path MTU Discovery – rfc 1191)

- Il campo Checksum è stato eliminato perché la sua elaborazione riduce le prestazioni.

- Il campo Protocol è stato rimosso perché questa info è contenuta nel Next Header.

Header Fields

- **Version** (4 bits) -> 0110
- **Traffic Class** (8 bits). E' un nuovo campo utilizzato per supportare la QoS basata sulle Classi. Corrisponde al Type of Service di Ipv4 utilizzato solo sperimentalmente)
- **Flow Label** (20 bits) – Label Switching, per QoS basata sui flussi (nuovo campo).
- **Payload Length** (16 bits) – Lunghezza del payload (esclusa l'intestazione)
- **Next Header** (8 bits) – Per snellire l'intestazione molti campi sono resi opzionali mediante Header numerate che possono essere concatenate



- Hop Limit (8 bits) – Era il TTL che ora assume il nome corretto.
- Source address (128 bits)
- Destination address (128 bits)

Extension Header

Estensioni opzionali nel formato Type-Lunghezza-Valore

L'ultimo NextHeader indica il protocollo del Payload (stessi codici del campo Protocol di IPv4)

Code	Header Type
0	**Hop-By-Hop options –** Informazioni per i router attraversati
43	**Routing Header –** Lista di router da visitare nell'ordine indicato
44	**Fragmentation Header –** In alcuni casi la frammentazione è necessaria
50	**Encapsulating Security Payload (ESP – IPsec) –** Cifratura del datagramma
51	**Authentication Header (AH – IPsec) –** Integrita' del datagramma
60	**Destination Options –** Informazioni per il destinatario
1	**ICMPv4**
58	**ICMPv6**
6	**TDP**
17	**UCP**

ICMPv6

Equivale a ICMP per IPv4, con alcune nuove funzionalità:

- Path MTU discovery
- Neighbor discovery (equivalente in IPv6 di ARP)
- Router Discovery
-

Formato del pacchetto:

Type: indica il tipo - Code: specifica meglio il tipo - Checksum: dell'intero pacchetto

Tipi principali:

1=dest unreacheable (no route to dest, address unreacheable, ...)
2= packet too big (per il discovery automatico dell'MTU ottimale)
3= time exceeded (superato il numero massimo di hop consentiti)
128=echo req (ping)
129=echo replay (risposta al ping)
133=router solicitation (ricerca automatica dei router della LAN)
134=router advertisement
135=neighbor solicitation (sostituisce arp request)
136=neighbor advertisement (sostituisce arp response)

Path MTU discovery

E' un protocollo basato su ICMPv6 che consente di determinare l'MTU ottimale per connessioni TCP.

- Il nodo manda il primo pacchetto con una dimensione pari all'MTU del proprio link
- Se riceve un messaggio ICMPv6 “Packet too big” (tipo 2) manda un nuovo pacchetto con le dimensioni indicate nel messaggio
- Ripete finché non trova più errori

Neighbor discovery

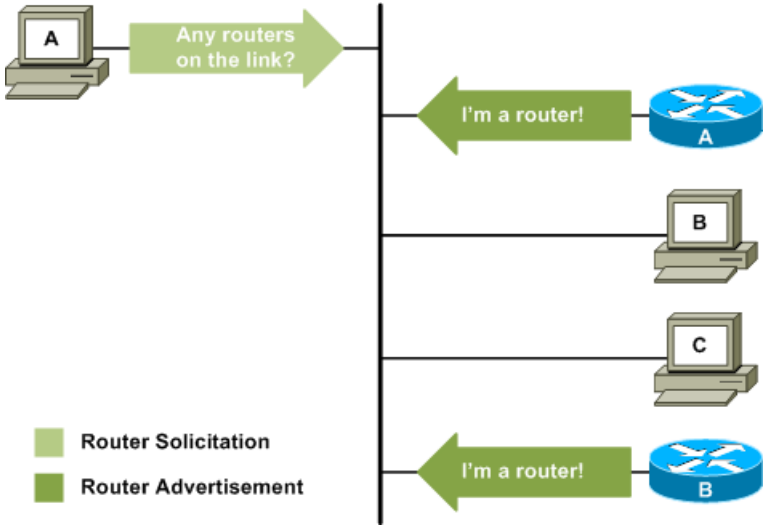
Sostituisce ARP per determinare l'indirizzo di rete LAN.

- Usa pacchetti ICMPv6 anziché ARP, multicast anziché broadcast
- Per ottenere un indirizzo fisico di un altro nodo:
  - Calcola l'indirizzo Solicited-Node (multicast) corrispondente all'indirizzo IPv6 del destinatario, formato aggiungendo gli ultimi 24 bit dell'indirizzo IP (ultime 6 cifre esadecimali del dest) al prefisso ff02::1:ff00:/104
  - Invia all'indirizzo multicast un pacchetto ICMPv6 “Neighbor Solicitation” (125)
  - Il destinatario risponde con un pacchetto ICMPv6 “Neighbor Advertisement” (136)
  - Il nodo memorizza l'indirizzo della Neighbor Cache

Router discovery

In IPv4 il default router deve essere configurato manualmente o via DHCP.
Con IPv6 gli host possono individuare automaticamente i router in un link.
Questo avviene attraverso 2 messaggi ICMPv6:
Router Solicitation (RS, type 133) e **Router Advertisement** (RA, type 124)

Quando un host entra in Link manda un Router Solicitation in multicast all'indirizzo [FF02::2] e ogni router risponde con un Router Advertisement contenente il suo indirizzo e altre informazioni necessarie per il routing.

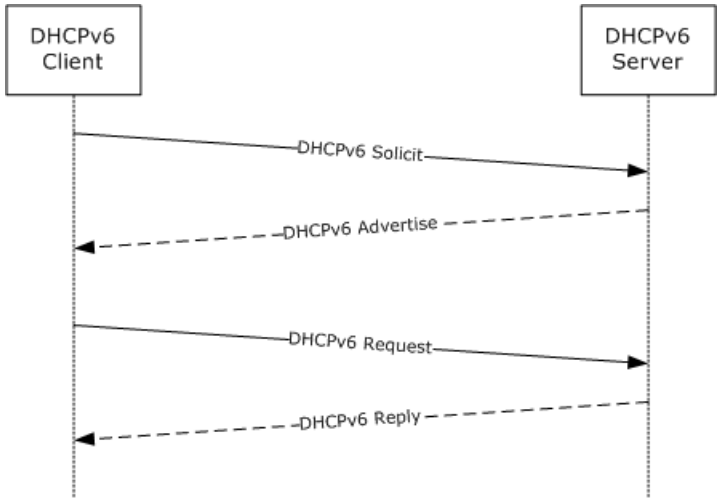


DHCPv6 (statefull autoconfiguration)

E' il protocollo dhcp per IPv6 descritto nell'RFC 3315 e consiste nello scambio del seguenti segmenti UDP:

Il client manda un "**Solicit**" dalla porta 546 a [ff02::1:2]:547 (multicast)
Il server risponde con un "**Advertise**" unicast dalla porta 547 verso la porta 546.
Il client risponde con un "**Request**" dalla porta 546 a [ff02::1:2]:547 (multicast)
Il server completa il protocollo con un "**Reply**" unicast dalla porta 547 verso la 546.

Nota: Per identificare gli host DHCP6 usa il DUID (DHCP UID) che è unico per ogni Host.



Stateless Address AutoConfiguration (SLAAC)

SLAAC è definito nell'RFC 2462
Combinando il protocollo di router discovery con l'autoconfigurazione degli indirizzi Link-local (FE80:0000:0000:0000:mEUI64) è possibile assegnare un indirizzo Global unicast in modalità plug & play, senza la necessità di avere un servizio DHCP.
Al momento del boot l'host ottiene dalla rete il default router ed il prefisso IPv6, quindi genera il Global address combinando LinkPrefix:mEUI64

- Adatto per i client (i server devono essere configurati manualmente)
- Il nome del DNS deve essere ottenuto in altro modo (esempio DHPCv6)
- L'indirizzo non viene automaticamente registrato nel DNS.

Nota: Nei sistemi Linux l'attivazione di SLAAC è controllata dall'opzione IPV6_AUTOCONF
Esempio: IPV6_AUTOCONF=YES

UNIVERSITÀ DI PARMA

DIPARTIMENTO DI SCIENZE MATEMATICHE, FISICHE ED INFORMATICHE
Corso di Laurea in Informatica

Il Livello Network

Parte III : Routing

RETI DI CALCOLATORI - a.a. 2022/2023

Roberto Alfieri

Il livello Network: sommario

PARTE I

- Scopi del livello Network
- Commutazione di circuito e di pacchetto
- La rete ATM
- La famiglia dei protocolli TCP/IP
- Il protocollo IP: trama indirizzi, instradamento
- Protocolli di servizio: ARP, ICMP, DHCP

PARTE II

- IPv6

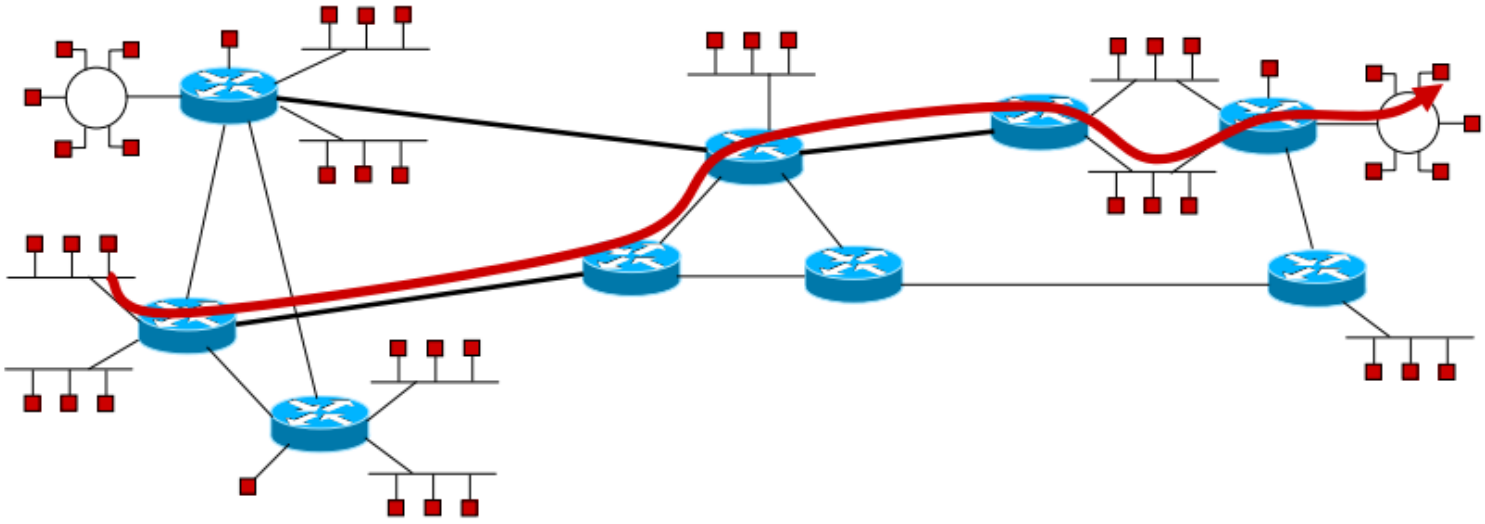
PARTE III

- Routing: Algoritmi e protocolli. Distance Vector e Link State.

Routing

E’ la scelta del percorso su cui inviare i dati quando mittente e destinatario appartengono a 2 reti diverse e quindi la consegna non può avvenire direttamente a livello Link.

In questo caso il mittente affida la consegna ad una struttura interconnessa di **Router** i quali passano i datagrammi dall’uno all’altro finché raggiungono quello che può consegnarli direttamente al destinatario.

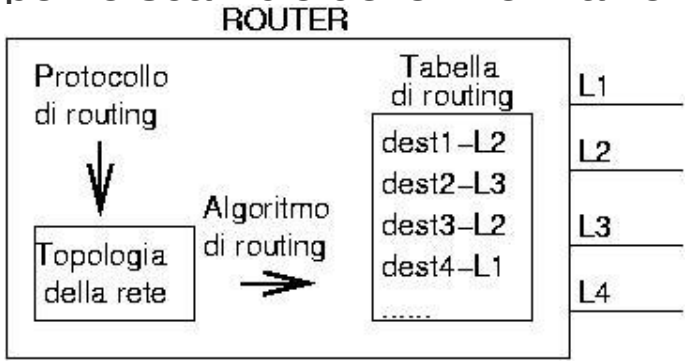


Router

I Router sono dotati di due componenti funzionali: **instradamento e inoltro.**

Instradamento: Creazione di **una tabella di routing (RT),** che contiene le informazioni riguardo la porta di uscita per le destinazioni dei Frame. Componenti:

- **Algoritmi di routing (RA)** si preoccupano di scegliere lungo quale linea di uscita vanno instradati i pacchetti in arrivo. Sono usati per il calcolo della tabella di routing in base alla topologia della rete.
- **Protocolli di routing (RP)** utilizzati per lo scambio delle informazioni necessarie per determinare la topologie della rete.



Inoltro: Applicazione dell'instradamento sui singoli datagrammi ricevuti.

- **Lettura dell'intestazione IP** ed estrazione dell'indirizzo di destinazione.
- **Look-up della tabella di routing** ed identificazione dell'interfaccia di uscita
- **Switching:** trasferimento fisico dei datagrammi da ingresso a uscita.

Algoritmo di Flooding

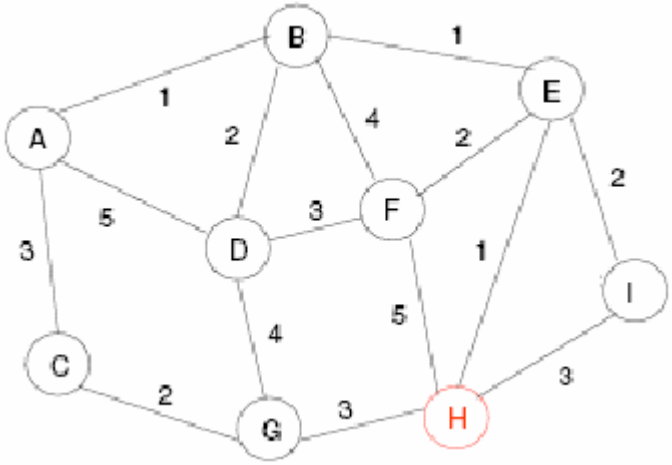
Flooding è un semplice algoritmo di routing in cui ogni pacchetto entrante è spedito verso tutte le linee di uscita eccetto quella di cui è arrivato.

Per gestire la propagazione di pacchetti duplicati le possibili tecniche sono:
- Conteggio del numero di salti, decrementando il valore ad ogni salto.
- Numero di sequenza assegnato dal mittente ad ogni messaggio. Ogni nodo deve gestire una lista con i messaggi già trasmessi, scartando eventuali repliche.

Flooding è utilizzato
- nei bridge per l'invio di broadcast e nella fase di autoapprendimento
- nei protocolli link state.

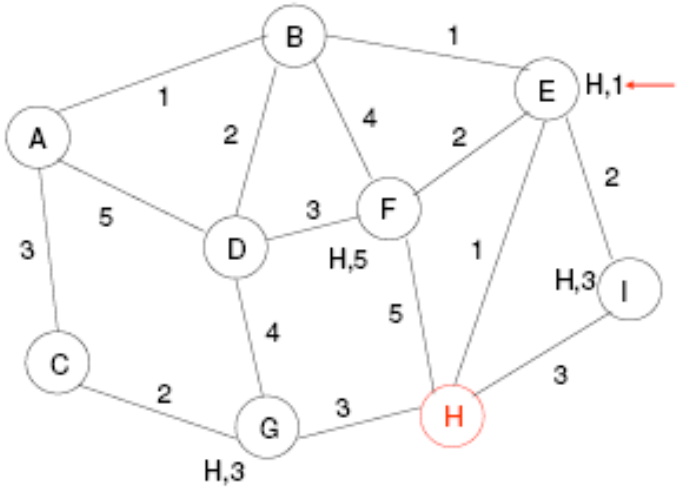
Algoritmo Shortest Path First

- La topologia della rete può essere rappresentata da un grafo pesato non orientato. Il peso attribuito ad ogni arco viene determinato mediante l'attribuzione di una metrica che tiene conto vari parametri di rete (velocità, latenza, ..)
- Per il principio di ottimalità il cammino minimo che un nodo deve percorrere per raggiungere qualsiasi altro nodo del grafo è un albero detto **"Sink Tree"** (a partire dalla destinazione) o **Source Tree** (a partire dall'origine).
- Si conoscono diversi algoritmi per elaborare il percorso più breve tra due nodi. Il più utilizzato è stato ideato da Dijkstra nel 1959 ed è noto con il nome di **Shortest Path First (SPF).**

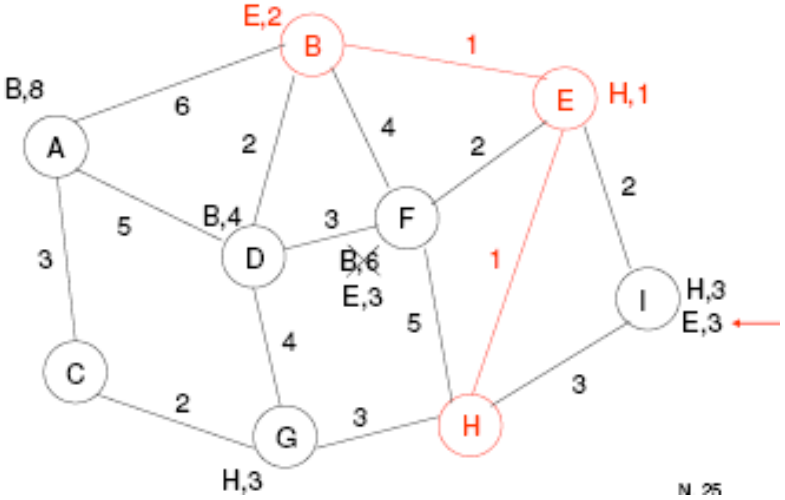


Ricerca del cammino minimo da A verso H. A partire da H (Sink Tree) mettiamo etichette provvisorie sui nodi adiacenti. Scegliamo la più piccola delle distanze.

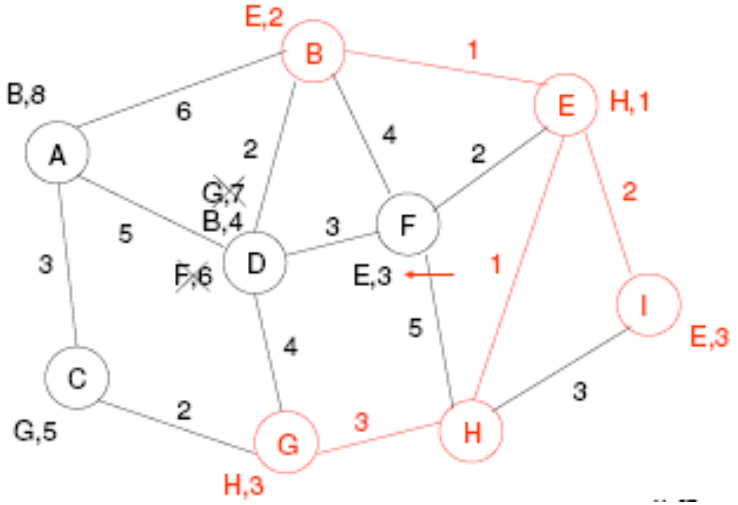
Sink Tree del nodo H con SPF



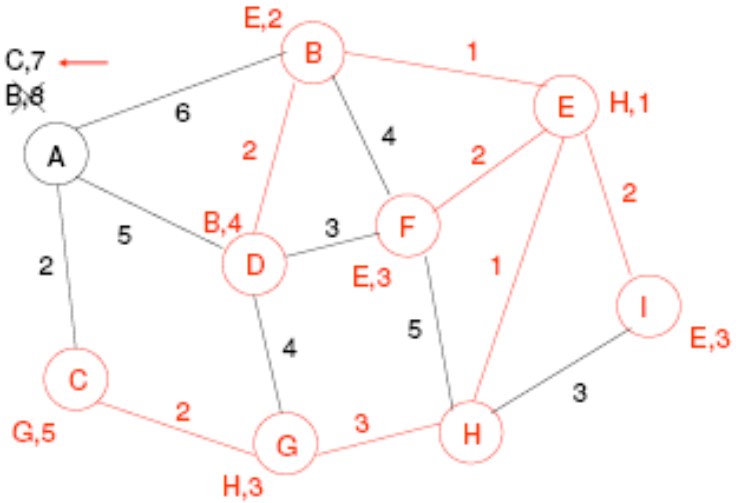
1



2



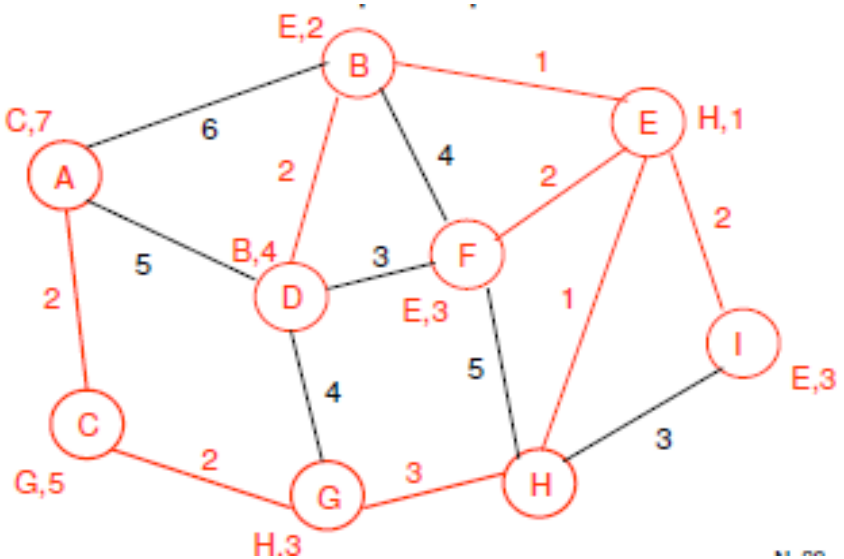
3



4

Risultato finale

- A va ad H attraverso C con distanza 7.



Protocolli di Routing

I RP stabiliscono le modalità di comunicazione tra i router per la costruzione della topologia di rete. Possono essere statici o dinamici (adattivi).

- Routing Statico:
  - La topologia e tabella di routing vengono definite in fase di setup della rete.
  - In caso di variazioni (inserimento o eliminazione di nodi o collegamenti) è necessario l'intervento dell'operatore.
- Routing Dinamico:
  - La topologia della rete è costruita dinamicamente in modo automatico, in base ai cambiamenti della topologia di rete o al traffico.
  - **Routing Dinamico Centralizzato**:
  - - un nodo centrale raccoglie le informazioni sullo stato della rete
  - - calcola (RA) la tabella per ogni nodo e la spedisce.
  - - tabelle consistenti, ma abbiamo un punto di criticità
  - **Routing Dinamico Distribuito**:
  - - i nodi si scambiano informazioni sullo stato della rete
  - - ogni nodo calcola la propria tabella sulla base delle informazioni ricevute.
  - - Tre categorie di protocolli: **Distance Vector, Link State e Gerarchici**

Protocolli Distance Vector

- Nel grafo ogni coppia di nodi ha una **distanza** che dipende dalla “metrica” utilizzata.
- Una metrica ragionevole tra 2 nodi adiacenti potrebbe dipendere dalla velocità, la latenza e il Throughput del canale.
- La distanza di un percorso potrebbe dipendere dalla somma delle singole distanze e/o dal numero di salti.
- **Nel Protocollo Distance Vector (DV) ogni nodo invia ai primi vicini l’elenco delle distanze (a lui note) con tutti gli altri nodi (ovvero il DV), periodicamente e ogni volta che c'è un cambiamento.**
- Le distanze con i primi vicini vengono misurate (ad esempio con un ECHO), mentre le altre distanze sono derivate dalle informazioni ricevute
- Tutte le volte che un Router calcola una nuova tabella di instradamento, la invia agli IS adiacenti (cioè quelli collegati da un cammino fisico diretto) sotto forma di DV
- **La tabella** contiene una entry per ogni nodo presente in rete
- Ogni **entry** è composta da quattro parametri:
  - Indirizzo (del nodo remoto)
  - Hops (numero di salti per raggiungerlo)
  - Costo (determinato in base alla metrica)
  - Linea
- Il DV inviato contiene Indirizzo-Hops-Costo di ogni entry (non la linea).

Protocolli Distance Vector

Il router che riceve il DV prima di tutto verifica se vi sono delle modifiche dal precedente e, in caso affermativo, aggiorna i campi hops e costo, sommando 1 a tutti gli hops e sommando il costo della linea da cui è arrivato il messaggio al campo costo.

Indirizzo	Hops	Costo	Linea
1	3	25	3
2	5	35	2
3	9	50	6
4	1	5	7
5	0	0	0

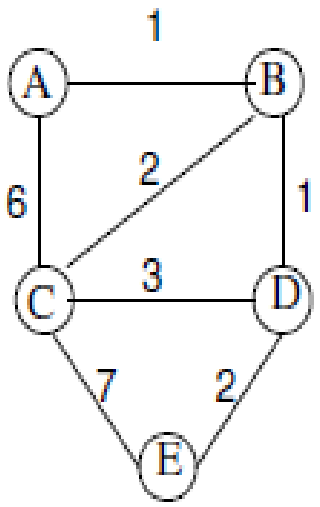
DV

Il passo successivo è l'aggiornamento della propria tabella tramite un processo di **fusione** (merge) di tutti i Distance Vector a lui pervenuti da ogni linea attiva.

Nella fusione vengono esaminate le entry con lo stesso indirizzo di destinazione, scartando quelle con i costi maggiori.
A parità di costo si seleziona quella che ha il **minor numero di hops.**

Il protocollo è semplice ma a lenta convergenza: l'informazione di una modifica della topologia (linea interrotta, router spento, ..) si propaga lentamente.

Distance Vector: Esempio di calcolo



Distance Vector A = {(A,0), (B,1), (C,6), (D,∞), (E, ∞)}
DV B = {(A,1), (B,0), (C,2), (D,1), (E,∞)}
DV C = {(A,6), (B,2), (C,0), (D,3), (E,7)}
DV D = {(A,∞), (B,1), (C,3), (D,0), (E,2)}
DV E = {(A,∞), (B,∞), (C,7), (D,2), (E,0)}

Tabella di routing alla fine del periodo di convergenza (al termine dello scambio dei DV)

1. A riceve DV di B

dest	Costo, next hop
A	0
B	1, A
C	3, B
D	2, B
E	∞

2. A riceve DV di C

dest	Costo, next hop
A	0
B	1, A
C	3, B
D	2, B
E	13, C

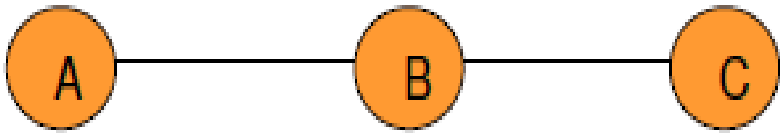
3. B riceve DV di D

dest	Costo, next hop
A	1, B
B	0
C	2, B
D	1, B
E	3, D

4. A riceve DV di B

dest	Costo, next hop
A	0
B	1, A
C	3, B
D	2, B
E	4, B

Distance Vector: il problema “count to infinity”



- Situazione iniziale: $D_{AC} = 2$ e $D_{BC} = 1$
  - Link BC va fuori servizio
  - B riceve il DV di A che contiene l’informazione $D_{AC} = 2$, per cui esso computa una nuova $D_{BC} = D_{BA} + D_{AC} = 3$ e la comunica ad A
  - A calcola la nuova distanza $D_{AC} = D_{AB} + D'_{BC} = 4$
  - Il processo può continuare all’infinito
- Vari rimedi sono stati proposti, nessuno risolutivo

RIP protocol

RIP (Routing Information Protocol) è la prima implementazione di un protocollo DV.
Ne esistono 3 versioni:

- RIPv1 (RFC 1058) usa il routing "classful" (reti senza NetMask)
- RIPng (RFC 2080) estensione del protocollo RIPv1 per supportare IPv6.
- RIPv2 (RFC 2453) usa il routing “classless” (CIDR)

Caratteristiche:

- Metrica: basata solo sulla minimizzazione degli hops (max 15)
- Nodi RIP Attivi (tipicamente Router): annunciano il loro percorsi
  - **ogni 30 secondi e quando si verificano cambiamenti di topologia**
- Nodi RIP Passivi (tipicamente Host): aggiornano senza annunciare

Protocolli Link State

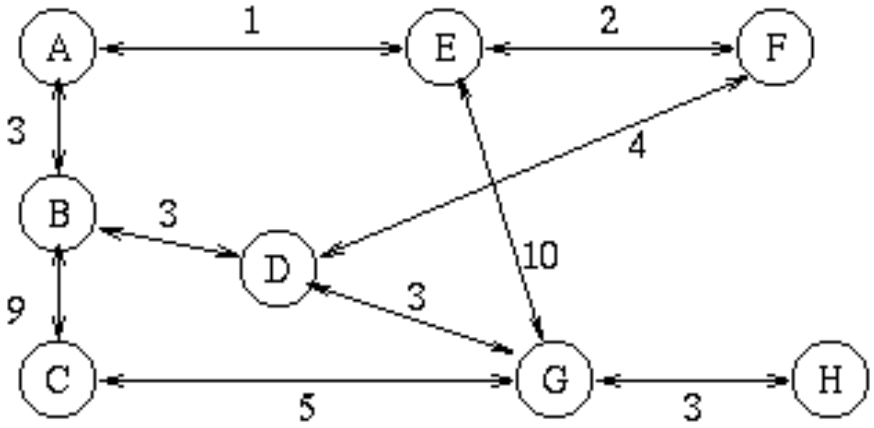
E' un RP con cui ogni nodo determina e mantiene aggiornata la topologia della rete da cui calcola la Tabella di Routing applicando un RA.

Il protocollo si sviluppa nelle seguenti fasi:

1. **Scoperta dei vicini** (neighbor greetings) : invio di un pacchetto HELLO su tutte le linee.
2. **Misurazione costo linea**: invio di un ECHO ai router che hanno risposto all'HELLO.
3. **Costruzione di un pacchetto** (Link State Packet - LSP) con tutte le informazioni ricavate nella fase 2: l'identità del trasmittente, numero di sequenza, dall'età e lista di vicini con il relativo ritardo misurato.
4. **Distribuzione periodica del LSP a TUTTI i nodi della rete**, con un numero di sequenza, utilizzando il Flooding. Se arriva un pacchetto con un numero di sequenza inferiore al numero più alto visto fino a quel momento, il pacchetto viene scartato (ritenuto obsoleto).
5. **Ogni nodo**, dopo aver ricevuto gli LSP da tutti gli altri, **costruisce la topologia della rete** e applica un RA (Shortest Path First di Dijkstra) per il **calcolo della tabella.**

Protocolli Link State: esempio 1/2

- Raccolta Info (HELLO - ECHO)
- Propagazione info (LSP) in flooding multicast



- Per il nodo D il pacchetto LSP sarebbe:

Adiacente	Costo
B	3
F	4
G	3

- Ogni nodo ricostruisce la mappa della rete fondendo i LSP ricevuti in una tabella come quella a fianco.

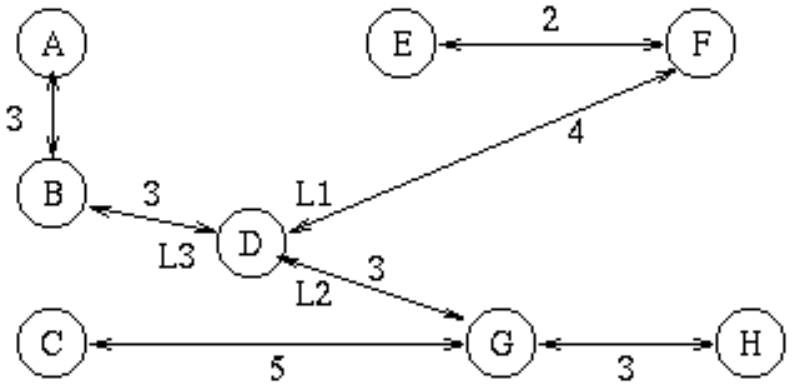
A	B/3	E/1	
B	A/3	C/9	D/3
C	B/9	G/5	
D	B/3	F/4	G/3
E	A/1	F/2	G/10
F	D/4	E/2	
G	C/5	D/3	E/10
H	G/3		

H/3

Protocolli Link State: esempio 2/2

- Il calcolo della tabella di instradamento si riduce ora al calcolo dello spanning tree di tipo SPF (Shortest Path First) e lo si effettua tramite il noto algoritmo di Dijkstra

- Lo spanning tree ad esempio del nodo D risulterà come nella figura a lato e, a seguire la relativa tabella di instradamento



- L'algoritmo può gestire reti di grandi dimensioni grazie alla sua rapida convergenza ed il suo comportamento è prevedibile, poiché ogni nodo ha in memoria la mappa intera della rete.

- Difficilmente si generano loop e, comunque, risulta facile identificarli ed eliminarli.

A	L3
B	L3
C	L2
E	L1
F	L1
G	L2
H	L2

OSPF

OSPF (Open Shortest Path First, RFC2328)
È un protocollo IGP di tipo **Link State Packet** ed è raccomandato da IETF per Internet.

- Ciascun router emette periodicamente (default 10 s) dei **pacchetti Hello** multicast (244.0.0.5) , per valutare possibili modifiche topologiche
- Ogni Router costruisce un pacchetto con l'elenco delle linee attive e dei loro costi (1/larghezza di banda della linea)
- Invia in **flooding** pacchetti **Link State Update (LS-Update)** multicast 244.0.0.5
- Questi vengono riscontrati con un **Link State Ack** (LS-Ack)
- Se, in base ai pacchetti di update ricevuti, si sono verificate modifiche della topologia ricalcola la tabella di routing con l'algoritmo **Shostest Path First (SPF)**

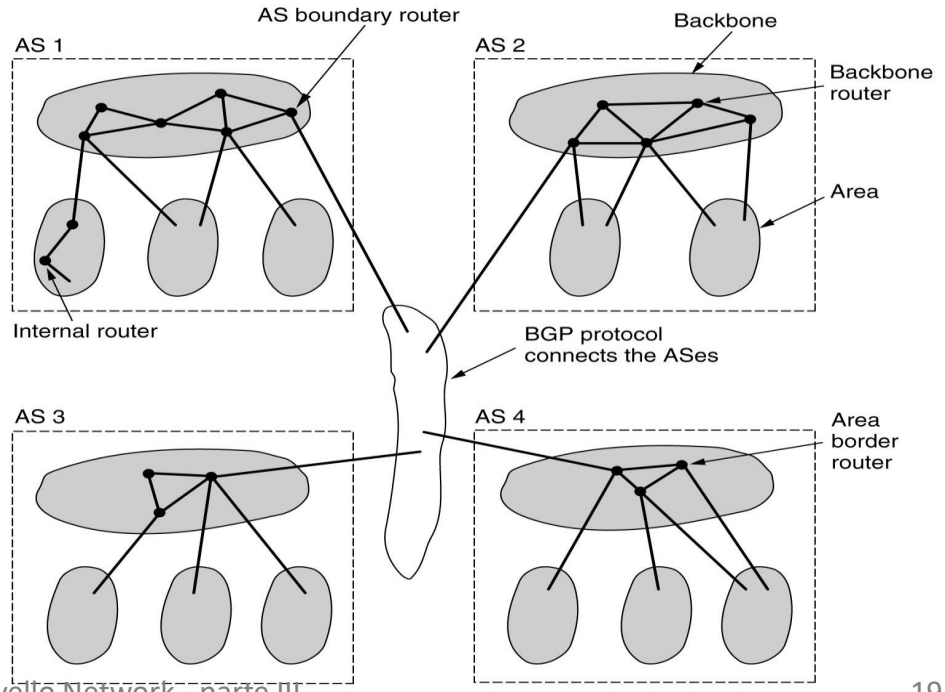
- Scalabilità:
  - Il calcolo del SPF ha complessità *O*(N logN), dove N è il numero di router e reti.
  - Il problema viene risolto suddividendo un AS in aree:
    - Ogni AS-OSPF contiene almeno un'area: l'area di Backbone (area 0)
    - Le eventuali altre aree sono connesse al Backbone
    - Ogni router mantiene informazioni solo riguardo la topologia della propria area.

Protocolli Gerarchici

Nel caso di reti di grandi dimensioni non è possibile gestire le tabelle di routing per l'intera rete in tutti i router, in questo caso il routing deve essere gerarchico:

- la rete viene ripartita in aree, chiamate Autonomous-System
- i router all'interno di un area sono in grado di effettuare l'instradamento relativamente alla sola area
- per destinazioni al di fuori dell'area si limitano ad inviare i pacchetti a dei router "di bordo" che sono a conoscenza della topologia esterna dell'area
- i router "di bordo" si occupano solamente dell'instradamento dei pacchetti fra aree

In linea di principio la ripartizione può essere effettuata tante volte quante si vuole creando più livelli nella gerarchia di routing



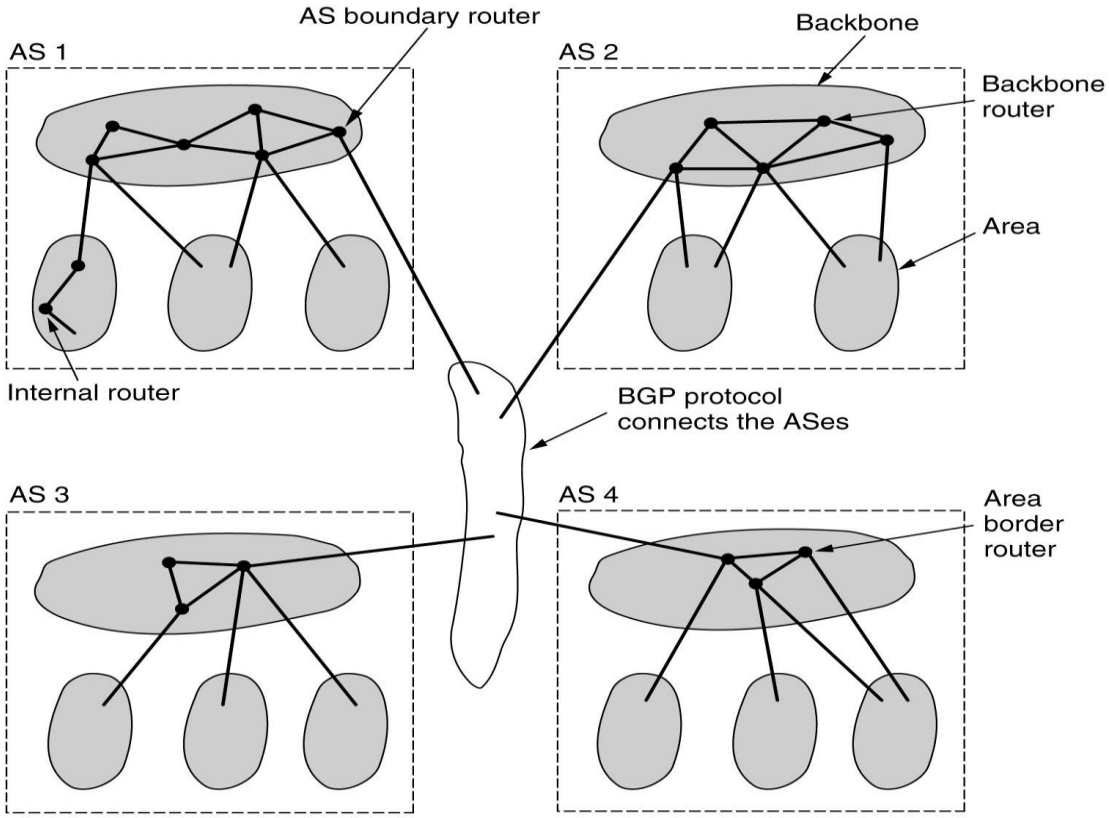
Protocolli di Routing in Internet

In TCP/IP i router sono suddivisi in due classi, **Exterior Router** ed **Interior Router**. I primi interconnettono due insiemi di reti distinti. Ogni insieme di reti, gestito da una singola autorità amministrativa, è un Autonomous System ed i router interni ad essi sono proprio gli Interior.
Un sistema è detto autonomo, poiché è libero di scegliere un'architettura di instradamento interna, ma deve raccogliere informazioni su tutte le sue reti e progettare uno o più gateway, gli Exterior Router, che passino le informazioni di raggiungibilità ad altri sistemi autonomi.
Gli Exterior Router utilizzano protocolli denominati EGP (Exterior Gateway Protocol), mentre gli Interior Router scambiano informazioni di instradamento tramite gli IGP (Interior Gateway Protocol).

I protocolli IGP più utilizzati sono **RIP** (Distance Vector) e **OSPF** (Link State)
BGP è ip protocollo raccomandato in Internet per l'interconnessione di Autonomous System.

BGP

BGP (Border Gateway Protocol, RFC 1771).
Protocollo Path Vector : invece di propagare i costi propaga la sequenza di AS da attraversare per arrivare a destinazione



Routing Anycast

IP Anycast è una tecnica che consente a diverse macchine (server) di condividere lo stesso indirizzo IP, in modo che un client raggiunga il server più vicino (a minore costo) per ridurre la latenza e aumentare la ridondanza.

Gli algoritmi di routing basati sui protocolli Distance Vector o Link State gestiscono automaticamente percorsi multipli per raggiungere una destinazione, selezionando il percorso a minor costo e quindi la destinazione più conveniente.

Ovviamente il routing non è stabilito sugli indirizzi IP, ma sulle reti, quindi sarà necessario definire una rete Anycast, (anche piccola) replicata in diversi punti della rete.