



Internet Protocol

Luca Veltri

(mail.to: luca.veltri@unipr.it)

Corso di Reti di Telecomunicazioni A, a.a. 2010/2011

<http://www.tlc.unipr.it/veltri>

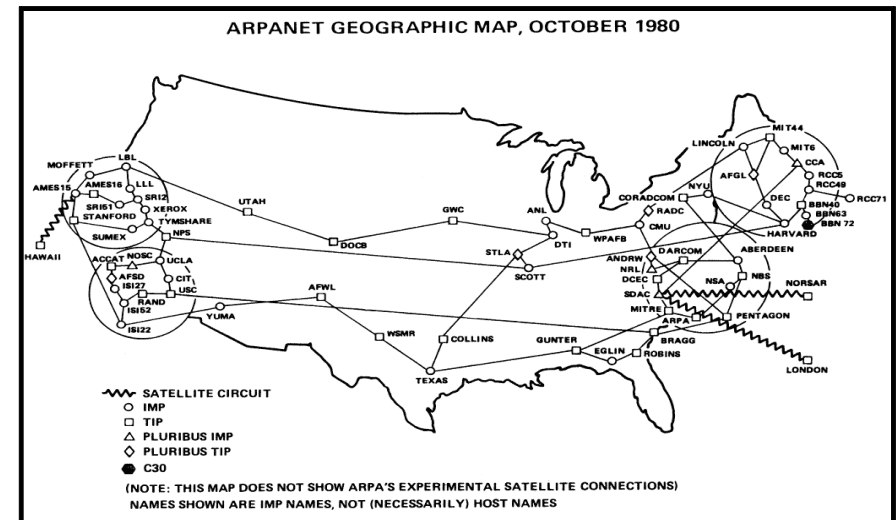
Sommario

- Introduzione ad Internet
- L'architettura TCP/IP
- Il protocollo IP
- Indirizzamento IP
- Routing IP
- ARP
- ICMP
- DHCP
- DNS
- Cenni su IPv6

Internet: storia

- Le origini di Internet si possono far risalire al progetto DARPA (Defense Advanced Research Project Agency) del Department of Defense (DOD) Americano (inizi anni '70)
- Necessità di interconnettere reti dei centri militari, universitari e di ricerca: definizione della rete ARPANET
- 1973 viene commissionato all'Università di Stanford il progetto di una suite di protocolli standard che garantissero connettività a livello di rete
- Verso la fine degli anni '70, tale sforzo portò al completamento dell'Internet Protocol Suite, di cui i due protocolli più noti sono il TCP e l'IP
 - comunemente si fa riferimento ad essa con la sigla TCP/IP
- RFC 791, "INTERNET PROTOCOL", Sept. 1981

ARPANET: 1980



Internet: storia (cont.)

- 1990: ARPANET cessa le sue attività
- 1990: Berners-Lee (CERN) definisce il WWW
- 1993: Andreessen (NCSA, Illinois) sviluppa il primo WWW browser
- Evoluzione:
 - **applicazioni e servizi principali:**
 - e-mail, ftp, telnet, news → www, e-mail, streaming, peer-to-peer
 - **velocità tipica delle portanti:**
 - 64 kbit/s → 2÷1000 Mbit/s
- Nuove problematiche:
 - **Qualità del servizio, multimedia, reti mobili**

5

Motivazioni del successo

- Sovvenzioni statali (\$ 200 milioni nel periodo '86-'95)
- Assenza di vincoli architeturali forti: non richiede che tutti i sistemi siano basati sulla stessa architettura protocollare
 - **interconnessione di reti eterogenee**
- Modalità di trasferimento nello strato di rete senza connessione
 - **flessibilità rispetto a tecnologie di (sotto)rete anche molto diverse tra loro**
- Semplicità dei protocolli usati (inclusi quelli di gestione e controllo)
- Disponibilità immediata e gratuita del software, della documentazione tecnica e degli standard (anche nei primissimi stadi di sviluppo)
 - **protocolli da subito implementati in software**
 - **protocolli integrati nel sistema operativo UNIX**
 - **vasta disponibilità (spesso gratuita) di applicazioni (modello client/server e peer-to-peer)**

6

Organi Regolatori

- ISOC (Internet SOCIety): Organo politico di rappresentanza di Internet
- IAB (Internet Architecture Board): Commissione di supervisione complessiva degli aspetti tecnici di Internet
- IETF (Internet Engineering Task Force): Comunità che specifica i protocolli di Internet ed emette documenti ad essa relativi. È diviso in Aree e Working Groups (WGs)
- IESG (Internet Engineering Steering Group): Organo di supervisione dell'IETF, costituito prevalentemente da IETF Area chairs
- IANA (Internet Assigned Number Authority): Autorità che disciplina l'uso di tutti i numeri, valori, costanti, well-known ports usati in Internet
- InterNIC (Internet Network Information Centre): Ente che assegna gli indirizzi IP ed i nomi di dominio

7

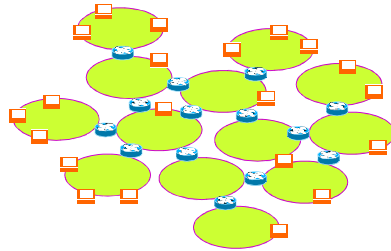
IETF

- L'organo di standardizzazione più importante per Internet dal punto di vista tecnico è l'IETF (Internet Engineering Task Force)
 - **definisce i protocolli riguardanti la rete Internet**
 - **gli standard prodotti vengono denominati RFC (Request For Comments) e numerati in ordine crescente**
 - **http:// www.ietf.org/**
- Processo di standardizzazione dell'IETF:
 - **Processo aperto a tutti i contributi**
 - **Nessuna votazione, bensì consenso**
 - **Prototipi disponibili prima (non dopo) la standardizzazione**
- Tutto ciò riassunto dal motto dell'IETF:
 - **"Rough consensus and running code!"**
 - **David D. Clark, MIT, in speech to IETF Plenary, ~1994.**

8

Struttura di rete

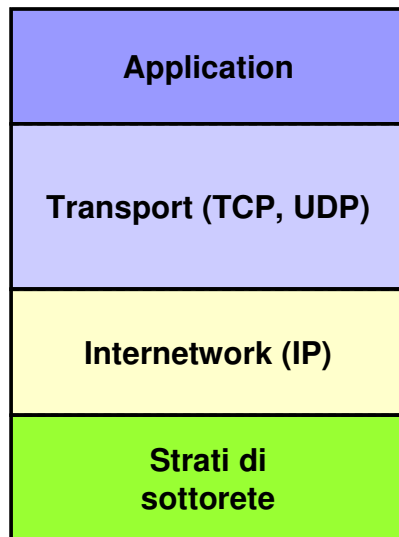
- Una rete IP consiste di un insieme di (sotto)reti interconnesse tramite il protocollo IP
 - queste reti sono collegate da nodi di commutazione IP chiamati router
 - i router instradano i pacchetti IP sulla base dell'indirizzo della rete di destinazione con modalità di trasferimento (commutazione) senza connessione (CL)
 - lo strato IP, di base, non fornisce alcuna garanzia sulla qualità di servizio (integrità informativa, e trasparenza temporale) dei dati
- La struttura di rete è non gerarchica
- Internet è una (particolare) rete IP
 - pubblica con estensione globale



Struttura di rete (cont.)

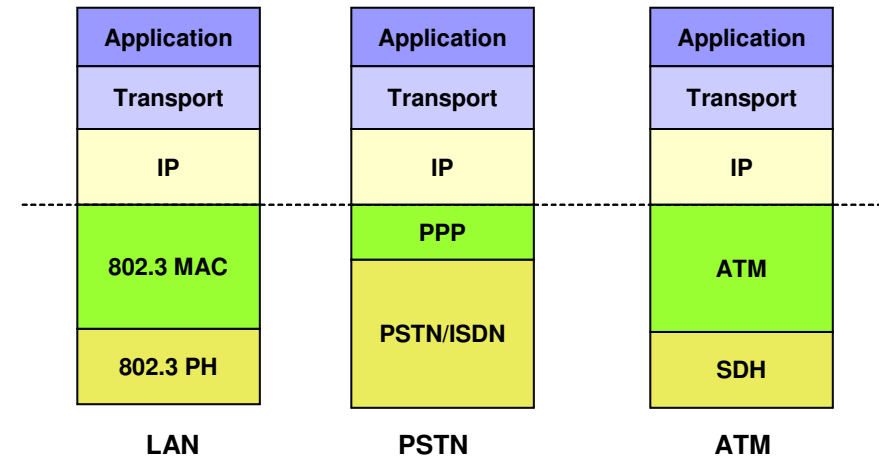
- Il protocollo IP tratta tutte le sotto-reti in modo uguale (link layer)
- Esempi di sotto-reti
 - una rete in area locale (e.g. Ethernet)
 - una rete in area metropolitana (e.g. anello SDH)
 - una rete geografica (e.g. rete PSTN/ISDN, o PLMN)
 - una connessione punto-punto dedicata
 - ...
- I protocolli TCP/IP (inteso come stack protocollare basato su IP)
 - sono posizionati al di sopra dei vari protocolli implementati all'interno delle singole sotto-reti (eventualmente duplicandone le funzioni)
 - sono in grado di operare su diverse piattaforme hardware (PC, PDA, Smartphone, Embedded PC, etc.) utilizzando svariati sistemi operativi (UNIX, Mac OS, MS-DOS, Windows, Linux, Symbian, etc.)

Architettura TCP/IP



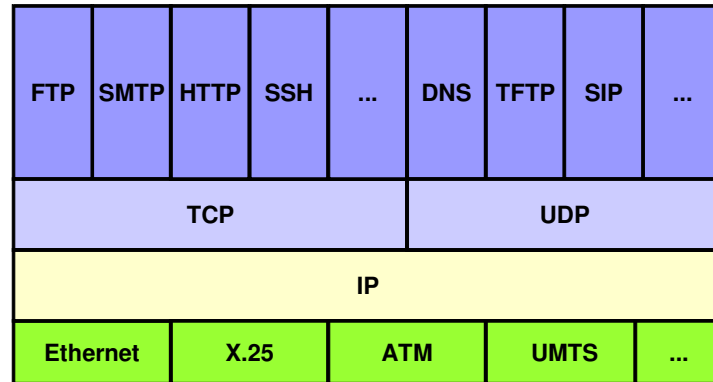
- corrisponde in parte allo strato di sessione e agli strati di presentazione e di applicazione del modello OSI
- corrisponde allo strato di trasporto e a parte dello strato di sessione OSI
- due tipi di servizio/protocolli: servizio affidabile (TCP), servizio senza connessione (UDP)
- consente l'interconnessione delle varie sottoreti con funzionalità che nel modello OSI sono collocate nello strato di rete
- include le funzioni che nel modello OSI sono comprese negli strati fisico, di collegamento e di rete, quest'ultimo per ciò che riguarda le singole sottoreti

Architettura TCP/IP: Esempi



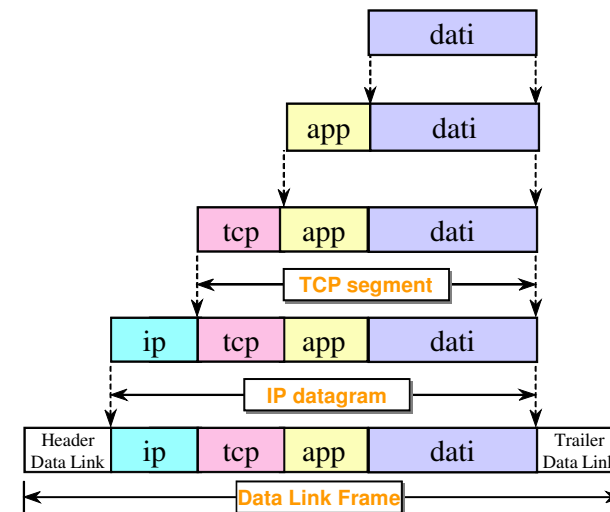
Architettura TCP/IP

- Esempio di pila dei protocolli Internet



13

Architettura TCP/IP: esempio di imbustamento



14

Architettura TCP/IP

- L'insieme di protocolli TCP/IP
 - riesce a inter-connettere tutti i tipi di sotto-rete in quanto assume che le funzionalità degli strati sottostanti costituiscano solo una piattaforma per il trasferimento fisico
 - realizza tutte le funzioni tipiche per il trasporto dell'informazione: controllo di errore, indirizzamento, instradamento, frammentazione e aggregazione delle unità informative, inoltre in rete
- Se alcune o tutte queste funzioni non erano state svolte da una particolare sotto-rete, TCP/IP le realizza; se erano già state svolte le duplica, realizzandole nuovamente
- Ciò conduce ad eventuali duplicazioni di funzioni ma consente di non imporre alcun vincolo sulla tecnologia e sui protocolli delle sotto-reti di trasporto che interconnette (X.25, Frame Relay, ATM, LAN, MAN, ISDN, UMTS, etc.)

15

Architettura TCP/IP (cont.)

- Le prestazioni da estremo a estremo (velocità di trasmissione delle informazioni, grado di trasparenza temporale e di integrità informativa) sono fortemente legate alle caratteristiche delle sottoreti attraversate
- Il trasferimento delle unità informative può richiedere una frammentazione delle stesse laddove le dimensioni delle unità informative gestite dalle sotto-reti non coincidano con le dimensioni massime consentite

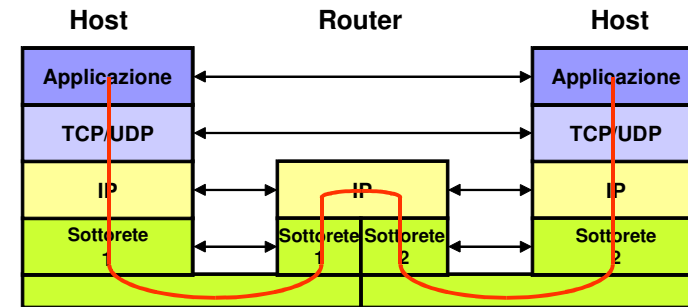
16

Host and Router

- I nodi IP si distinguono in host e i router
 - entrambi possono avere una o più interfacce di rete
 - ognuna delle quali permette di interconnettere il nodo con una rete sottostante
 - per ogni interfaccia di rete il nodo ha un indirizzo IP, coerente con gli altri nodi collegati alla stessa sottorete
- Host = Nodo terminale IP
 - sorgente e/o destinazione dei pacchetti IP (datagrammi)
 - sebbene in molti casi abbiano solo una interfaccia di rete, possono in generale avere più interfacce attive su sottoreti differenti
 - possono essere dei normali PC, workstation, palmari, telefoni cellulari o qualsiasi altro apparecchio connesso ad una rete IP
- Router = Nodo di commutazione IP
 - inoltrano i pacchetti IP da una rete ad un'altra (instradamento)
 - in genere hanno due o più interfacce di rete
 - possono essere delle macchine dedicate (la maggior parte dei router commerciali) o PC con un OS comune (e.g. Linux)

17

Architettura TCP/IP



- I router in quanto nodi di commutazione IP devono implementare il protocollo IP, ICMP e eventualmente protocolli di routing (e.g. RIP, OSPF)

18

Tipi di reti sottostanti lo strato IP

- IP tratta in modo uniforme qualsiasi rete sottostante su cui si appoggia per comunicare con nodi IP adiacenti
 - link level
- E' però possibile distinguere le seguenti tipologie:
 - **Singolo ramo punto-punto**
 - le interfacce dei router possono essere "unnumbered"
 - e.g.: linee dedicate o dial-up (PPP, SLIP)
 - **Rete multipunto**
 - più host possono comunicare direttamente senza passare per router intermedi
 - è necessario un meccanismo di mappaggio di indirizzi (e.g. ARP)
- Nel caso di sottoreti multipunto è possibile distinguere due casi:
 - reti con possibilità di broadcast (e.g. LAN)
 - reti senza possibilità di broadcast (e.g. X.25, Frame Relay, ATM, SDH, PSTN/ISDN, etc.)

19

Internet Protocol (IP)

Internet Protocol (IP)

- Il protocollo IP (RFC 791, 919, 922, 950, 1349)
 - è il **protocollo portante dell'architettura TCP/IP**
 - è un **protocollo di strato di rete (secondo il modello OSI)**
 - **opera con modalità di trasferimento connectionless (senza connessione), ovvero a datagramma**
 - ciascun pacchetto è trattato indipendentemente da tutti gli altri
 - è un **protocollo NON affidabile e non fornisce alcuna garanzia sulla QoS (servizio "best effort")**
 - non ci sono garanzie che un pacchetto IP giunga a destinazione, né sui tempi di trasferimento
 - i pacchetti possono essere consegnati errati, non consegnati, o consegnati fuori sequenza

21

Internet Protocol (IP)

- Il protocollo IP esegue le seguenti principali funzioni:
 - **indirizzamento**
 - **instradamento**
 - **controllo di errore sull'intestazione IP (IP header, cioè l'IP-PCI)**
 - **se necessario esegue frammentazione e ri-assemblaggio dei pacchetti**
- Le unità informative base di IP sono i pacchetti IP, ovvero le IP-PDU, ovvero i datagrammi IP
 - **lunghezza variabile (max 65536 ottetti/byte)**
 - **intestazione di lunghezza variabile, ma solitamente di 20 byte**

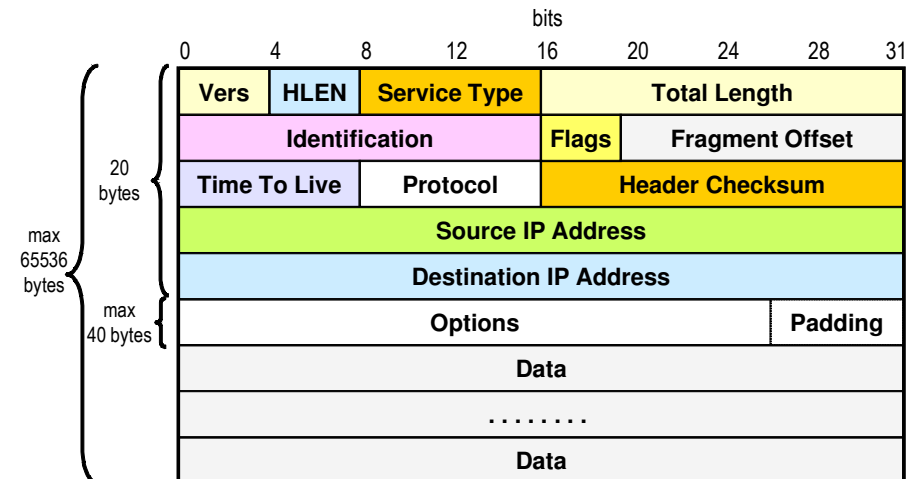
22

Internet Protocol (IP)

- Il protocollo IP in trasmissione
 - **riceve una unità dati dallo strato superiore (in genere lo strato di trasporto)**
 - **costruisce il corrispondente pacchetto IP e predispone l'opportuna intestazione (indirizzi src e dst, etc.)**
 - **sceglie il nodo successivo a cui consegnare il pacchetto**
 - **se necessario esegue frammentazione del pacchetto**
 - **inoltra il pacchetto (o i corrispondenti frammenti) al nodo successivo attraverso l'opportuna interfaccia di rete**
- in ricezione
 - **riceve un pacchetto dallo strato di sottorete**
 - **esamina l'intestazione e ne verifica la validità**
 - **verifica se il pacchetto deve essere rilanciato attraverso altra sottorete; in tal caso esegue le relative operazioni**
 - **al contrario, se corrisponde alla destinazione del pacchetto, estrae il contenuto dati (payload) e lo consegna all'opportuno protocollo di livello superiore**

23

Header IP



24

Header IP

- Vers (4 bits): versione del protocollo usata (attualmente IPv4)
- HLEN (4 bits): lunghezza dell'intestazione specificata in parole di 4 bytes (min 20 - max 60 bytes)
- TOS (Type Of Service): quando supportato dalla rete, permette di specificare un livello di qualità di servizio richiesto dall'utente (ad es. affidabilità o velocità di trasferimento)
 - in passato non usato
 - ora utilizzato da reti a servizi differenziati (Differentiated Services)
- Total length: (16 bits) specifica la lunghezza totale (inclusa l'intestazione) del datagramma, misurata in bytes (valore max: $2^{16}=65536$ bytes)
- Identification: (16 bits) numero identificativo del datagramma; è ereditato dagli eventuali frammenti
- Flags: è un campo di 3 bits usato per la frammentazione: X, DF e MF
 - X: non usato e posto a zero
 - DF: Don't Fragment; se 0 indica che il datagramma può essere frammentato, se 1 no
 - MF: More Fragment; se 0 indica che è l'ultimo frammento, se 1 che ci sono altri frammenti

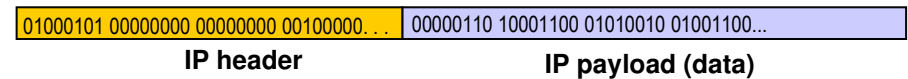
Header IP (cont.)

- Fragment Offset: (13 bits); posizione del frammento all'interno del datagramma, espresso in unità di 8 bytes può numerare 8192 frammenti; se uno o più frammenti non viene ricevuto (a destinazione), verrà scartato l'intero datagramma
- Time to Live: (8 bits); indica il numero massimo di salti residui che il datagramma può effettuare in rete. E' aggiornato da ogni router attraversato
- Protocol: indica a quale protocollo dello stato superiore deve essere trasferito il contenuto informativo del datagramma (es. TCP=6, UDP=17, ICMP=1)
- Header Checksum (16 bits): controllo di errore sull'intestazione; calcolato come somma complemento a 1; se viene rivelato un errore il datagramma viene scartato
- Source Address: (32 bits); indirizzo dell'host sorgente
- Destin. Address: (32 bits); indirizzo dell'host destinazione IP (ovvero dell'host, non dell'utente finale)

Header IP (cont.)

- Options: campo di lunghezza variabile (multipli di 8 bit) che può essere omissso. È composto da tanti ottetti quante sono le opzioni implementate. Ad esempio:
 - Source Route Option (strict o loose): consente al mittente di specificare i nodi attraverso i quali vuole che transiti il datagramma
 - Record Route Option (RRO): consente al mittente di creare una lista vuota di indirizzi IP in modo che ogni nodo attraversato inserisce il suo indirizzo in questa lista
 - Timestamp Option: come RRO con in più l'istante temporale in cui il datagramma attraversa i diversi nodi
- Padding: rende la lunghezza dell'intestazione multiplo intero di 32 bit mediante introduzione di zeri

Esempio di datagramma IP



representazione esadecimale	representazione binaria				
45 00 00 20	01000101	00000000	00000000	00100000	IP header
fe be 00 00	11111110	10111110	00000000	00000000	
80 11 ca b7	10000000	00010001	11001010	10110111	
8d 89 2b 20	10001101	10001001	00101011	00100000	
8d 89 2b 24	10001101	10001001	00101011	00100100	data
06 8c 52 4c	00000110	10001100	01010010	01001100	
00 0c 70 ce	00000000	00001100	01110000	11001110	
63 69 61 6f	01100011	01101001	01100001	01101111	

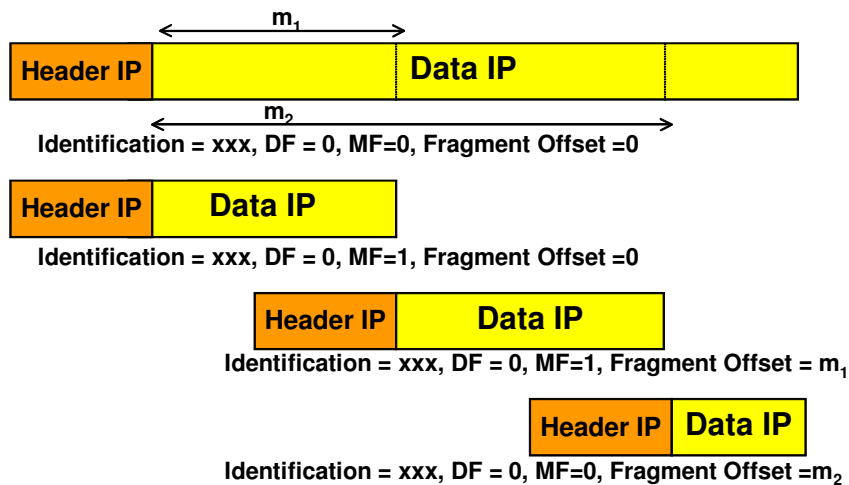
Frammentazione

Frammentazione IP

- Ogni sottorete ha un valore massimo di lunghezza della propria unità informativa
 - **Maximum Transmission Unit (MTU)**
- La frammentazione di un datagramma IP è necessaria se il valore della MTU nella sottorete fisica è inferiore alla lunghezza del datagramma
- La frammentazione è effettuata dal router/host prima del rilancio nella sottorete
- La ricomposizione del datagramma originale è effettuata dall'host di destinazione

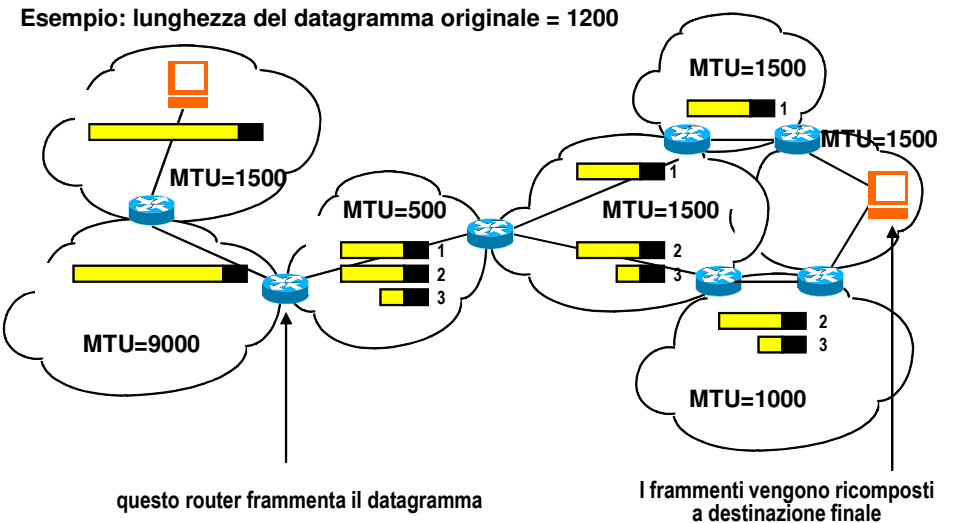
30

Frammentazione (cont.)



31

Frammentazione (cont.)



32

Indirizzamento IP

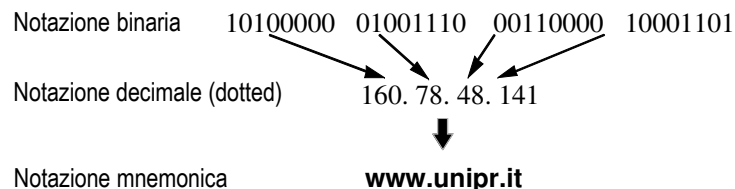
Schema di indirizzamento

- IP consente ad ogni nodo (IP) connesso alla rete di comunicare con ogni altro nodo (IP)
- Al tal fine utilizza un metodo globale di identificazione e indirizzamento di tutti i nodi (host e router) connessi alla stessa rete IP
 - tale schema di indirizzamento è indipendente da quello utilizzato nelle singole sottoreti dai protocolli sottostanti (Ethernet, X.25, PSTN, etc.)
- IPv4 utilizza indirizzi di 32 bit (IP Address)
 - spazio di indirizzamento: $2^{32} = 4.294.967.296$ indirizzi
- Un indirizzo associato ad ogni interfaccia di rete
- Se un nodo è connesso a più di una rete avrà un indirizzo IP per ogni rete (interfaccia di rete); ciò si verifica nei:
 - router
 - multi-homed host (host connessi a più sottoreti)

34

Schema di indirizzamento (cont.)

- Notazione numerica, "dotted" e "mnemonica":



- Un opportuno protocollo applicativo (DNS) provvede a tradurre un indirizzo numerico in mnemonico e viceversa

35

Schema di indirizzamento (cont.)

- Un indirizzo IP identifica un nodo e non uno specifico utente di trasporto o applicativo
 - l'identificazione di un utente all'interno di un host è affidata al campo protocol dell'intestazione IP e agli eventuali protocolli di strato superiore (e.g. TCP, UDP)
- L'identificativo del IP-SAP è dato alla coppia IP_address + protocol_id
- Gli indirizzi devono essere unici in tutta la rete IP
 - è possibili attribuire stessi indirizzi a due differenti host solo se questi sono connessi a due differenti reti IP
 - e.g. host collegati a due intranet diverse
- Lo schema di indirizzamento IP è stato progettato per consentire un instradamento efficiente
 - era però stato pensato per una rete con dimensioni decisamente inferiori alle attuali

36

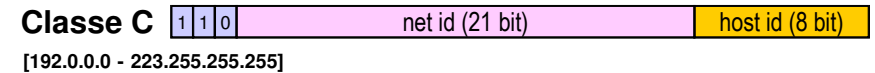
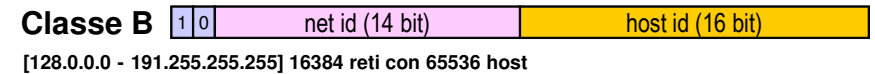
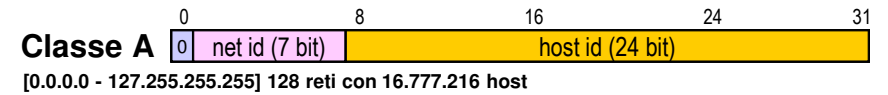
Schema di indirizzamento con classi

- Un indirizzo IP identifica sia il nodo che la rete a cui è connesso
 - l'indirizzo di un nodo IP deve essere coerente con quello della rete in cui si trova
 - se un host si sposta da una rete ad un'altra, il suo indirizzo deve essere cambiato
- In origine (1981, RFC 791) lo schema di indirizzamento era di tipo gerarchico a due soli livelli
 - Indirizzi formati da due parti: $IP_address = net_id + host_id$
 - net_id: identificativo di rete
 - host_id: identificativo del nodo all'interno della rete
 - la divisione tra net_id e host_id non è fissa
 - determinata in modo univoco dai primi bit dell'indirizzo
 - erano state fissate tre classi di indirizzi con net_id di lunghezza variabile da 8, 16, 24, più altre due classi di indirizzi per uso diverso

37

Schema di indirizzamento con classi (cont.)

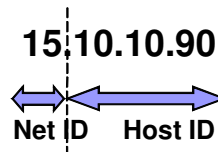
- 5 classi (o gruppi) di indirizzi IP



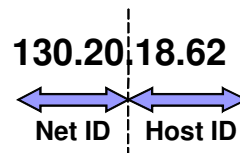
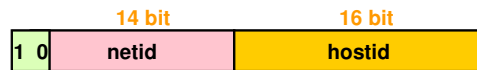
38

Schema di indirizzamento con classi: esempi

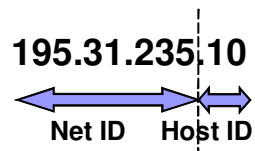
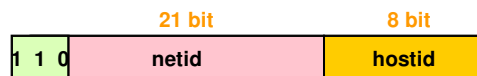
Esempio di indirizzo di classe A:



Esempio di indirizzo di classe B:



Esempio di indirizzo di classe C:



39

Schema di indirizzamento senza classi

- La struttura di indirizzamento a due livelli gerarchici era sufficiente nella fase iniziale di Internet
- Nel 1984 è stato aggiunto un terzo livello gerarchico
 - il livello di sotto-rete (subnet)
 - una rete di classe A, B, C può essere suddivisa in più sottoreti partizionando ulteriormente la parte di host_id
- Successivamente per far fronte all'esaurimento degli indirizzi IP e all'esplosione della dimensione delle tabelle di routing dei router è stato introdotto il CIDR (Class-less Inter-Domain Routing)
 - che a sua volta è basato sul concetto di super-rete (supernet)
 - assegnazione di indirizzi IP in modo aggregato e gerarchico
 - più indirizzi di rete assegnati allo stesso ISP (Internet Service Provider) vengono mantenuti contigui, in modo da poter essere rappresentati tramite un unico indirizzo aggregato
- Il risultato attuale è un indirizzamento senza classi

40

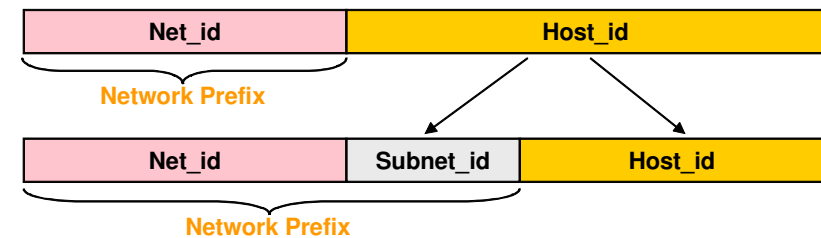
Schema di indirizzamento senza classi (cont.)

- Un indirizzo IP è formato da:
 - **IP_address = network_prefix + host_id**
 - **dove**
 - network_prefix: identificativo/prefisso di rete (lungo i bit)
 - host_id: identificativo dell'host (lungo $n-i$ bit)
 - **Il network_prefix viene utilizzato per verificare l'appartenenza di un nodo ad una sottorete (per l'istadamento)**
- Tale divisione non è più implicita ma quando necessario (nelle tabelle di routing) la divisione viene indicata esplicitamente tramite la *netmask*
 - **la netmask indica la lunghezza del network_prefix**
 - **normalmente vengono impiegate due tipi notazioni**
 - indicando con una sequenza di 1 e 0 la lunghezza del network_prefix, e rappresentandola in notazione decimale (maschera di rete o netmask); e.g. 255.255.255.0
 - indicando semplicemente la lunghezza del network_prefix; e.g. /24

41

Subnetting

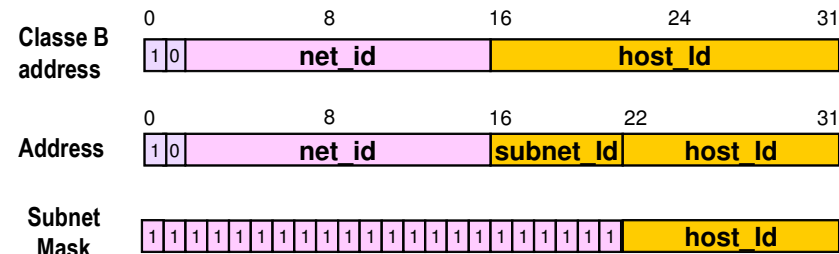
- Si utilizzano alcuni bit dell'host_id originario per codificare il subnet_id
 - **si riduce così la parte di host_id rimanente**



42

Subnetting (cont.)

- Il campo subnet_id e il complessivo network prefix è identificato da una maschera denominata "subnet mask" o brevemente "netmask"
- Una subnet mask è una parola di 32 bit in cui
 - **i bit uguali a "1" identificano i bit del net_id e del subnet_id (network prefix)**
 - **i bit uguali a "0" identificano i bit dell'host_id**



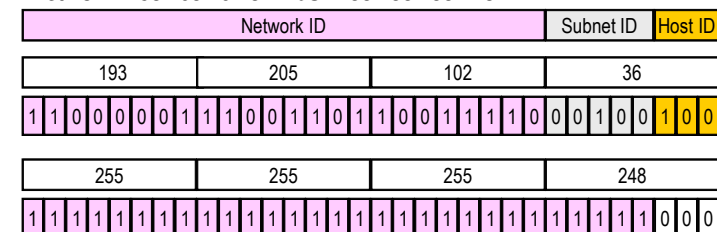
43

Subnetting (cont.)

- Nel caso in cui il network prefix coincide con il net_id (no subnetting) si ha la "netmask naturale" (netmask implicita)
- Si parla di *subnetting* quando la nuova maschera ha un numero di bit a 1 maggiore rispetto alla maschera naturale
- Grazie al *subnetting* aumentano gli indirizzi di rete a disposizione e migliora la gestione degli indirizzi
- Esempio:

➢ **host: 193.205.102.36**

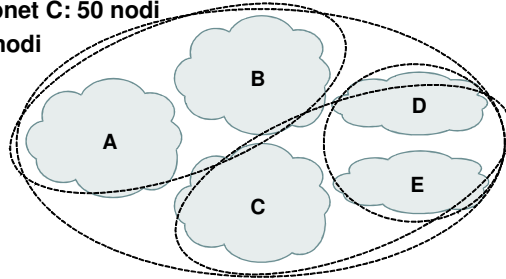
➢ **network: 193.205.102.32 mask 255.255.255.248**



44

Subnetting a lunghezza variabile: Esempio

- A partire da un indirizzo di classe B
 - 165.214.0.0
- Ricavare 5 sottoreti del tipo
 - Subnet A, Subnet B, Subnet C: 50 nodi
 - Subnet D, Subnet E: 20 nodi



- subnetting
 - 4 sottoreti con 64 indirizzi ciascuna (Host_id: 6 bit) (subnet mask 255.255.255.192)
 - 1 sottorete divisa ulteriormente in due sottoreti con 32 indirizzi ciascuna (Host_id: 5 bit) (subnet mask 255.255.255.224)

49

Netmask: valori possibili

- Per come è definita una netmask, i suoi 4 byte possono assumere solo i seguenti valori decimali

decimale	binario	# di host_id (se ultimo byte)
0	0000 0000	(256)
128	1000 0000	(128)
192	1100 0000	(64)
224	1110 0000	(32)
240	1111 0000	(16)
248	1111 1000	(8)
252	1111 1100	(4)
254	1111 1110	(2)
255	1111 1111	(1)

50

Supernetting

- Strategia di assegnazione di indirizzi IP in modo aggregato e gerarchico
- Più indirizzi di rete assegnati ad un ISP (Internet Service Provider) vengono presi contigui, in modo da poter essere rappresentati (nelle tabelle di routing) e pubblicizzati (tramite i protocolli di routing) tramite un unico indirizzo aggregato
 - esempio,
 - le 4 reti: 193.200.16.0/24, 193.200.17.0/24, 193.200.18.0/24, 193.200.19.0/24
 - possono essere viste come una unica rete (aggregata) con indirizzo: 193.200.16.0/22

51

Supernetting: esempio

- Esempio 1
 - Ad un grande Internet Service Provider (ISP) sono assegnati 2048 blocchi di indirizzi di classe C
 - da 198.24.0.0 (11000110.00011000.00000000.0)
 - a 198.31.255.0 (11000110.00011111.11111111.0)
 - CIDR mask per il grande ISP = 198.24.0.0/13
 - Un piccolo ISP locale richiede al grande ISP 16 blocchi di indirizzi di classe C
 - da 198.24.16.0 (11000110.00011000.00010000.0)
 - a 198.24.31.0 (11000110.00011000.00011111.0)
 - CIDR mask per il piccolo ISP locale = 198.24.16.0/20

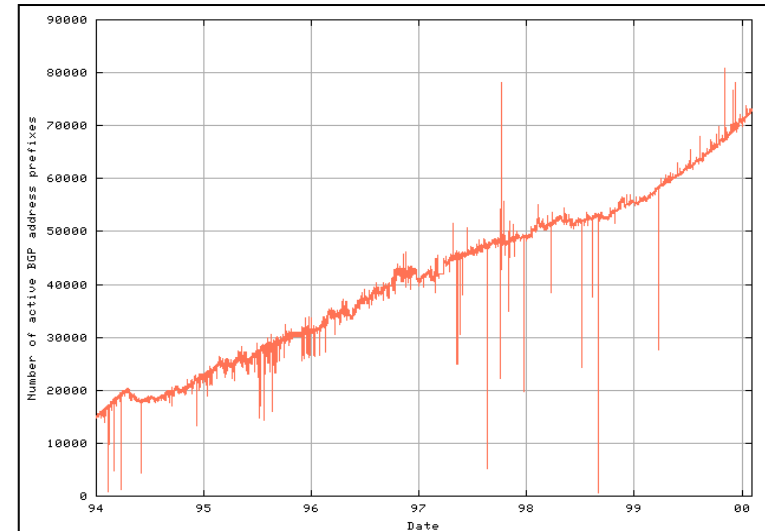
52

Supernetting: esempio

- Esempio 2
 - Ad una organizzazione sono assegnati 2048 indirizzi IP utilizzando 8 blocchi di indirizzi di classe C (8 reti di classe C)
 - da 194.32.136.0 (11000010.00100000.10001000.0)
 - a 194.32.143.0 (11000010.00100000.10001111.0)
 - CIDR mask per il grande ISP = 198.32.136.0/21

53

Esplosione delle tabelle di routing



<http://www.telstra.net/ops/bgptable.html>

54

Esaurimento degli indirizzi IP

- Il progressivo esaurimento degli indirizzi IP unitamente alla rapida crescita delle dimensioni delle tabelle di routing ha spinto l'IETF (Internet Engineering Task Force) ad intraprendere delle azioni preventive
- Tali misure preventive possono essere raggruppate nelle seguenti categorie:
 - Assegnazione razionale degli indirizzi IP
 - Classless InterDomain Routing (CIDR)
 - Indirizzi privati e Network Address Translation (NAT)
 - IP versione 6 (IPv6)

55

Indirizzi IP particolari

All 0s		This host ¹
All 0s	Host	Host on this net ¹
All 1s		Limited broadcast (local net) ²
Net	All 0s	Network address
Net	All 1s	Directed broadcast for net ²
127	Anything (often 1)	Loopback ³

¹ Utilizzabile solo come indirizzo sorgente (usato al bootstrap)

² Può essere usato solo come indirizzo destinazione

³ Non deve essere propagato dai nodi sulla rete

56

Indirizzi privati

- Dallo IANA (Internet Assigned Numbers Authority) sono stati riservati alcuni blocchi di indirizzi IP (RFC 1918) per uso interno ai singoli domini amministrativi (intranet)

10.0.0.0 ÷ 10.255.255.255 (10.0.0.0/8)

172.16.0.0 ÷ 172.31.255.255 (172.16.0.0/12)

192.168.0.0 ÷ 192.168.255.255 (192.168.0.0/16)

- Tali indirizzi NON sono instradabili pubblicamente
 - **IANA-Allocated, Non-Internet Routable, IP Addresses**
- Nella notazione pre-CIDR :
 - **il primo blocco corrisponde ad una rete di classe A**
 - **il secondo blocco a 16 reti contigue di classe B**
 - **il terzo blocco a 256 reti contigue di classe C**

57

Indirizzi privati (cont.)

- Ogni rete può utilizzare al suo interno qualsiasi indirizzo appartenente a questi insiemi
- I router della rete pubblica NON rilanciano pacchetti con indirizzi di destinazione di questo tipo
 - **il router di bordo della rete che usa tali indirizzi non deve propagare all'esterno tramite protocolli di routing informazioni di instradamento relativi alle sottoreti con questi indirizzi (privati)**
- Un host con un indirizzo privato non ha la possibilità di comunicare direttamente a livello IP con nodi pubblici di Internet (esterni alla propria rete)
 - **un host con un indirizzo privato può però comunicare con host esterni tramite**
 - rilancio effettuato a livello applicativo da nodi con indirizzi pubblici
 - Application Level Gateway (Proxy)
 - rilancio effettuato a livello IP da nodi che effettuano traduzione di indirizzi tra domini con indirizzamento differente (e.g. privato-pubblico)
 - NAT router (vedi seguito)
 - rientra nella tipologia di interlavoro classificata come traduzione di protocollo

58

Configurazione di un nodo IP

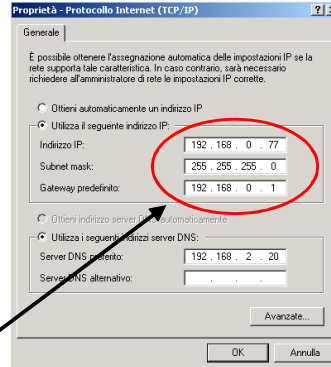
Configurazione di un nodo Unix/Linux

- ifconfig
 - **mostra / configura le interfacce di rete**
- route
 - **mostra / manipola la tabella di routing IP (print, add, delete)**
- arp
 - **manipola la cache ARP del sistema**
- nslookup
 - **effettua interrogazioni al DNS**
- netstat
 - **mostra connessioni di rete, tabelle di routing, statistiche sulle interfacce, connessioni masquerade e messaggi netlink**

60

Configurazione di un nodo windows

- Ipconfig
 - mostra le configurazioni IP di tutte le interfacce di rete
- winipcfg (win98/ME)
 - tool grafico per visualizzare le impostazioni di rete
- netsh (win2000/XP)
 - tool per modificare le impostazioni di rete
- route
- arp
- nslookup
 - (winNT, windows2000/XP)
- netstat
- Impostazioni di rete su WinXP
 - Pannello di controllo → Connessioni di rete → Connessione alla rete locale → Proprietà → Protocollo Internet (TCP/IP)

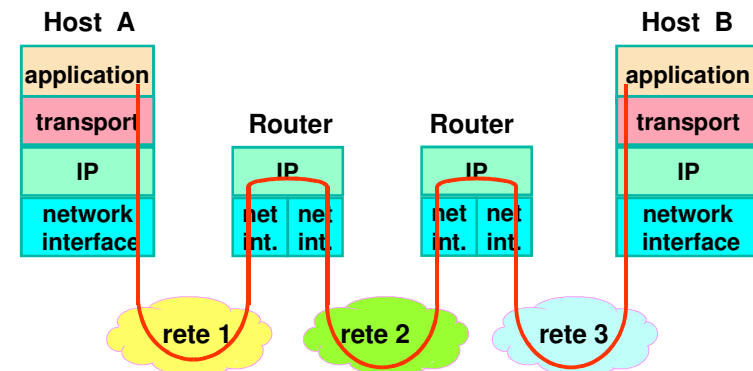


Routing IP

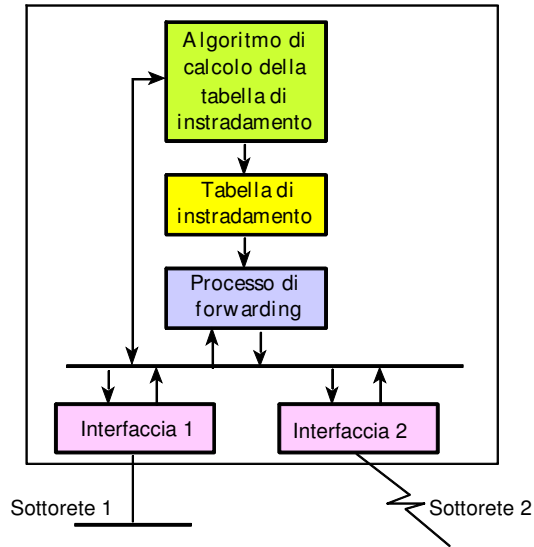
Router

- I *router* sono i nodi di commutazione della rete IP
 - hanno il compito di instradare i pacchetti IP attraverso la rete
 - permettono di far comunicare host che non sono connessi direttamente alla stessa sottorete IP
 - ricevono i datagrammi IP da un'interfaccia di ingresso e li inoltrano attraverso l'opportuna interfaccia di uscita
- Si distinguono dagli *host* perchè:
 - inoltrano i datagrammi IP diretti ad altri nodi
 - possono utilizzare "protocolli di routing" per aggiornare dinamicamente le proprie tabelle di routing

Router (cont.)



Architettura logica di un router



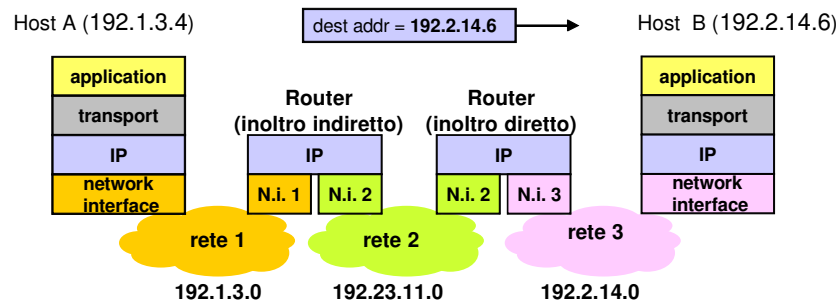
65

Instradamento (1/3)

- Dato un datagramma IP in uscita da un nodo (host o router), si possono distinguere due casi di instradamento:
 - **instradamento diretto:**
 - l'host destinazione è nella stessa sottorete dove si trova il nodo che sta instradando il datagramma IP
 - il pacchetto viene consegnato all'host di destinazione tramite i protocolli della sottorete
 - **instradamento indiretto:**
 - l'host destinazione non si trova in nessuna delle sottoreti a cui è connesso il nodo considerato
 - il datagramma viene consegnato ad un next hop router che avrà il compito di far proseguire il datagramma verso l'host destinazione
 - il pacchetto viene consegnato al router successivo tramite i protocolli della sottorete
- Nel caso di instradamento indiretto
 - **il nodo successivo procederà con instradamento diretto (nel caso sia l'ultimo nodo lungo il percorso) o con instradamento indiretto**

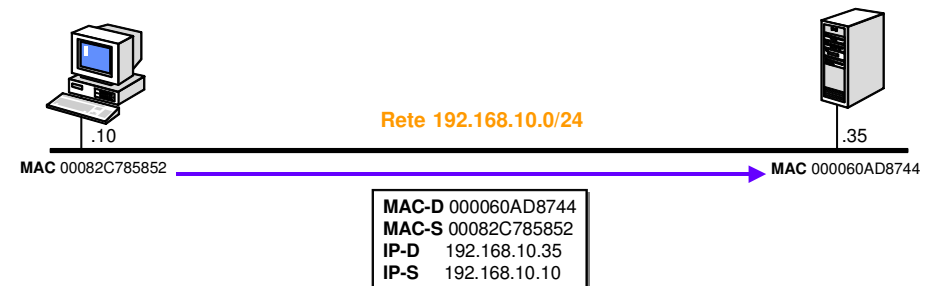
66

Instradamento (2/3)



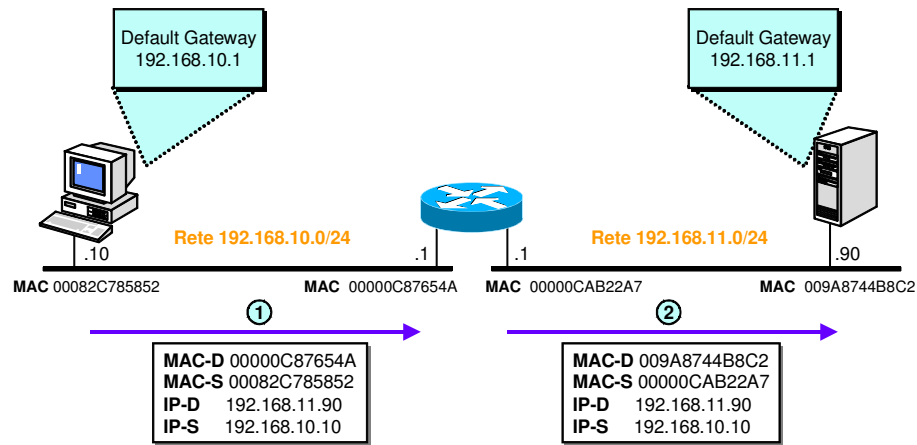
67

Instradamento diretto: esempio di inoltr



68

Forwarding indiretto: esempio di inoltra



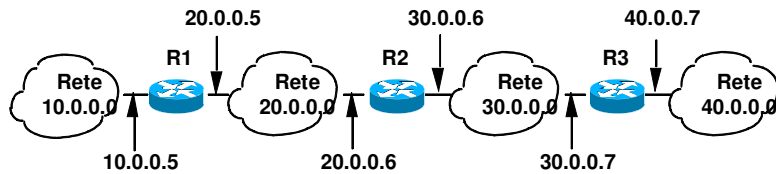
69

Instradamento (3/3)

- La decisione se utilizzare instradamento diretto o indiretto, e nel secondo caso, su quale "next-hop node" utilizzare per instradare il datagramma viene presa in base ad una tabella di instradamento (routing table)
 - la **routing table** di un nodo (host o router) specifica solo un passo lungo il cammino verso l'host di destinazione
 - un host/router non conosce il cammino completo, ma solo il passo successivo verso la destinazione
- La routing table (RT) contiene coppie (dest , next-hop)
 - **dest** : indirizzo della rete di destinazione
 - **next-hop** : indirizzo del prossimo router verso la rete di destinazione
- I nodi devono conoscere l'indirizzo del next-hop router
 - **nota**: nel caso di instradamento diretto l'indirizzo di destinazione apparterrà ad una delle sottoreti a cui il nodo è direttamente connesso

70

Tabelle di instradamento (1/4)



Net_Id	Router_Id
20.0.0.0	Instradamento diretto
30.0.0.0	Instradamento diretto
10.0.0.0	20.0.0.5
40.0.0.0	30.0.0.7

71

Tabelle di instradamento (2/4)

- L'informazione comunemente contenuta in ogni riga di una tabella di instradamento è:
 - **Dest net address**: rete di destinazione
 - **Subnet Mask**: porzione dell'indirizzo IP indicante il prefisso di rete
 - **Next Hop**: IP address del router successivo verso la destinazione
 - **Interface**: identificatore della porta fisica dove trovare il next hop
 - **Metric**: peso assegnato al cammino
- La coppia dest_net_addr + subnet_mask
 - serve per identificare la possibile sottorete/host di destinazione
- La coppia next_hop + interface
 - serve per determinare univocamente dove instradare il datagramma
- Se ci sono più righe che corrispondono, viene scelta quella con network prefix (netmask) più lungo
 - **longest prefix matching**

72

Tabelle di instradamento (3/4)

- Esempio di una tabella di instradamento di un host

windows:

Indirizzo rete	Maschera	Indirizzo gateway	Interfac.	Metric
0.0.0.0	0.0.0.0	150.100.33.1	150.100.33.18	1
127.0.0.0	255.0.0.0	127.0.0.1	127.0.0.1	1
150.100.33.0	255.255.255.0	150.100.33.18	150.100.33.18	1
150.100.33.18	255.255.255.255	127.0.0.1	127.0.0.1	1
150.100.255.255	255.255.255.255	150.100.33.18	150.100.33.18	1
224.0.0.0	224.0.0.0	150.100.33.18	150.100.33.18	1
255.255.255.255	255.255.255.255	150.100.33.18	150.100.33.18	1

linux:

Destination	Gateway	Genmask	Flags	MSS	Window	Iface
150.100.33.0	*	255.255.255.0	U	1500	0	eth0
127.0.0.0	*	255.0.0.0	U	3584	0	lo
default	150.100.33.1	0.0.0.0	UG	1500	0	eth0

73

Tabelle di instradamento (4/4)

- Esempio di una tabella di instradamento di un router (linux)

Destination	Gateway	Genmask	Flags	MSS	Iface
150.100.33.0	*	255.255.255.0	U	1500	eth0
150.100.34.0	*	255.255.255.0	U	1500	eth1
150.100.35.0	*	255.255.255.0	U	1500	eth2
127.0.0.0	*	255.0.0.0	U	3584	lo
default	150.100.35.1	0.0.0.0	UG	1500	eth2

74

Longest Prefix Matching

- Instradamento di 2 datagrammi

- indirizzo dest. 198.15.7.3
- indirizzo dest. 198.15.7.4

- Matching:

- 198.15.7.3
 - R1: matching prefisso 16
 - R7: matching prefisso 24
 - R4: matching prefisso 32
- 198.15.7.4
 - R1: matching prefisso 16
 - R7: matching prefisso 24
 - R4: no matching

- Scelta instradamento:

- 198.15.7.3 ⇒ R4
- 198.15.7.4 ⇒ R7

Tabella di instradamento

Prefix	Next hop
198.15.0.0/16	R1
198.15.7.0/24	R7
198.15.7.3/32	R4

75

Tabelle di instradamento: Esempio di CIDR (1/3)

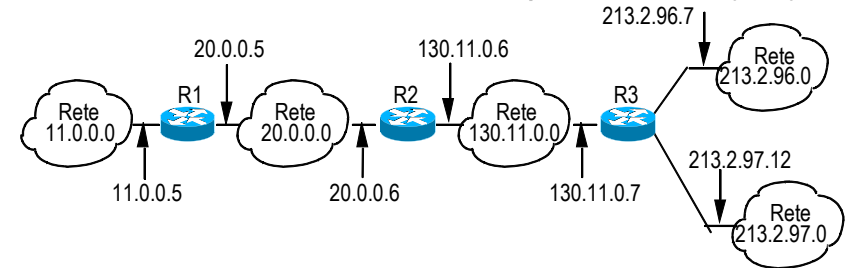
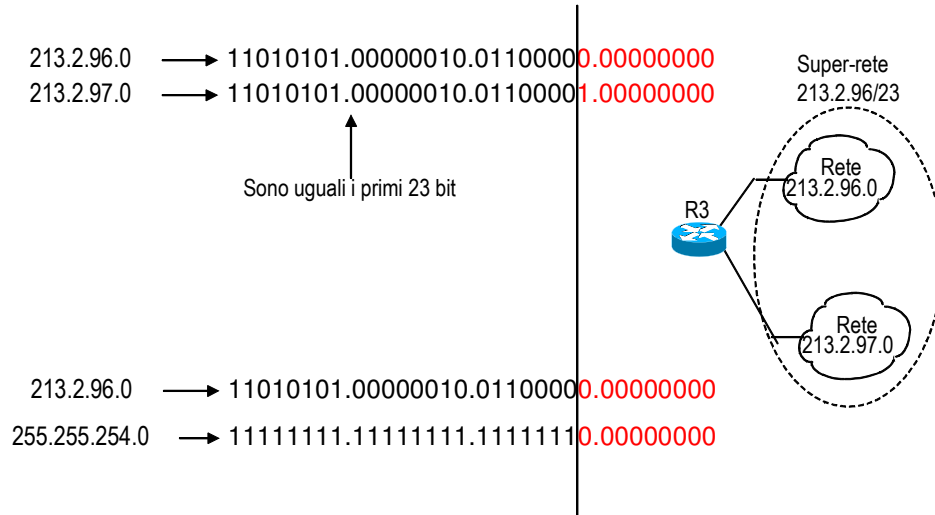


Tabella di instradamento di R2

Dest network	Subnet mask	Next hop
20.0.0.0	255.0.0.0	diretto
130.11.0.0	255.255.0.0	diretto
11.0.0.0	255.0.0.0	20.0.0.5
213.2.96.0	255.255.255.0	130.11.0.7
213.2.97.0	255.255.255.0	130.11.0.7

76

Tabelle di instradamento: Esempio di CIDR (2/3)



77

Tabelle di instradamento: Esempio di CIDR (3/3)

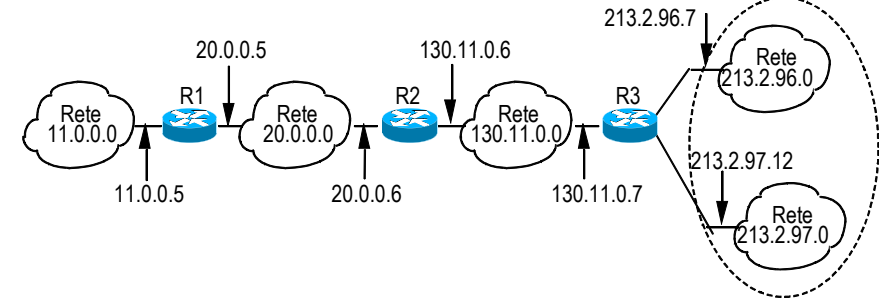


Tabella di instradamento di R2

Dest network	Subnet mask	Next hop
20.0.0.0	255.0.0.0	diretto
130.11.0.0	255.255.0.0	diretto
11.0.0.0	255.0.0.0	20.0.0.5
213.2.96.0	255.255.254.0	130.11.0.7

78

Tabelle di instradamento

- Al fine di
 - nascondere il più possibile i dettagli inerenti la rete,
 - mantenere piccole le tabelle di instradamento, e
 - consentire un instradamento efficiente,
- le tabelle contengono (in genere) solo informazioni sulle reti di destinazione e non sui singoli nodi
- Spesso nelle RT è presente come possibile target di rete di destinazione anche l'indirizzo della massima super-rete 0.0.0.0/0 (ovvero net 0.0.0.0 e mask 0.0.0.0)
 - questo indirizzo di rete include ogni possibile indirizzo di destinazione
 - il router next-hop relativo a questa particolare super-rete viene detto router (o gateway) di default
 - in questo modo si può evitare di includere esplicitamente nella RT tutte le possibili reti di destinazioni
 - tutto ciò vale sia per le RT degli host che dei router
 - In base al Longest Prefix Matching, se per un datagramma non viene trovata nella RT una strada diversa allora viene instradato verso il "router di default" (se presente nella tabella di routing)

79

Classless Inter Domain Routing

- Nel 1996 erano stati assegnati
 - 100 % degli indirizzi di classe A
 - 61.95 % degli indirizzi di classe B (rischio esaurimento)
 - 36.44 % degli indirizzi di classe C
- CIDR è stato ideato per
 - affrontare l'esaurimento dello spazio di indirizzamento di IP (raddoppio degli host ogni anno)
 - diminuire la complessità delle tabelle di instradamento nei router
 - velocizzare le operazioni di instradamento nei router
- CIDR tende ad eliminare la divisione in classi di indirizzi anche nei router più interni alla rete Internet
- CIDR è basato sulla tecnica di *supernetting*

80

Classless Inter Domain Routing (cont.)

- Nuova strategia di assegnazione degli indirizzi
 - la metà superiore della classe A (da 64 a 127) è stata riservata per usi futuri
 - un indirizzo di classe B è assegnati solo se la rete ha
 - almeno 32 sotto-reti
 - oltre 4096 host complessivi
 - gli indirizzi della metà inferiore della classe C (da 192.0.0 a 207.255.255) sono divisi in otto blocchi assegnati ciascuno ad una autorità geografica
 - gli indirizzi della metà superiore della classe C (da 208.0.0 a 223.255.255) non sono assegnati
 - ad una rete che non soddisfa i requisiti per la classe B è assegnato un certo numero di blocchi contigui di indirizzi di classe C
 - la rete è caratterizzata da un unico prefisso (insieme dei bit più significativi)
 - la rete sarà individuata nei router solo dal prefisso

81

Classless Inter Domain Routing (cont.)

- Pianificazione geografica degli indirizzi di classe C

Multiregional	192.0.0 ÷ 193.255.255
Europe	194.0.0 ÷ 195.255.255
Others	196.0.0 ÷ 197.255.255
North America	198.0.0 ÷ 199.255.255
Central/South America	200.0.0 ÷ 201.255.255
Pacific Rim	202.0.0 ÷ 203.255.255
Others	204.0.0 ÷ 205.255.255
Others	206.0.0 ÷ 207.255.255

- Tutte le reti appartenenti ad una regione geografica sono identificate dagli stessi 7 bit di prefisso
 - **Esempio: Europa**
 - da 194 = 11000010 0 a 195 = 11000011 1

82

Classless Inter Domain Routing (cont.)

- Da un indirizzo IP a 32 bit e dalla relativa maschera di rete a 32 bit si individua il prefisso con una operazione di AND
- In una routing table un blocco di indirizzi può essere rappresentato da un unico elemento (di lunghezza variabile) corrispondente al prefisso (Supernetting)

```

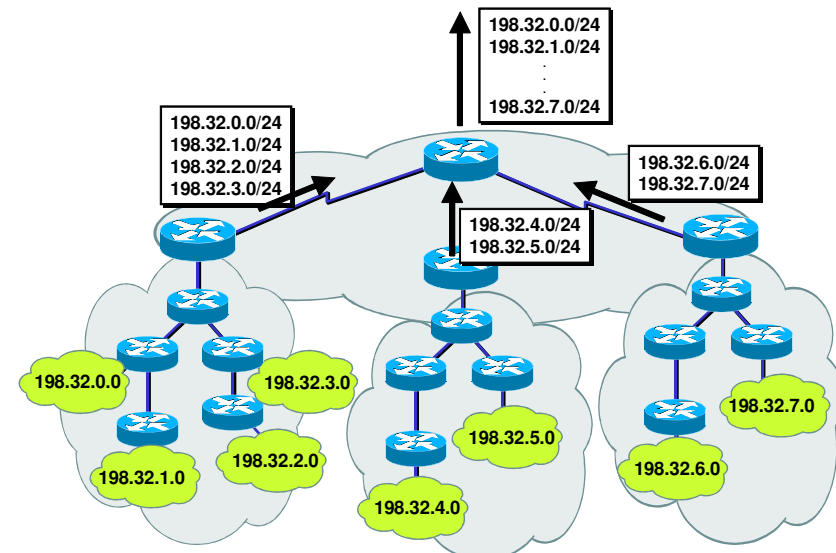
11000000 00100000 10001000 00000000 = 192.32.136.0 (class C address)
11111111 11111111 11111000 00000000 = 255.255.248.0 (network mask)
----- logical_AND
11000000 00100000 10001      = 192.32.136 (IP prefix)

11000000 00100000 10001111 00000000 = 192.32.143.0 (class C address)
11111111 11111111 11111000 00000000 = 255.255.248.0 (network mask)
----- logical_AND
11000000 00100000 10001      = 192.32.136 (same IP prefix)
    
```

- Viene scelto l'instradamento verso la direzione corrispondente al prefisso di lunghezza maggiore (Longest Prefix Matching)

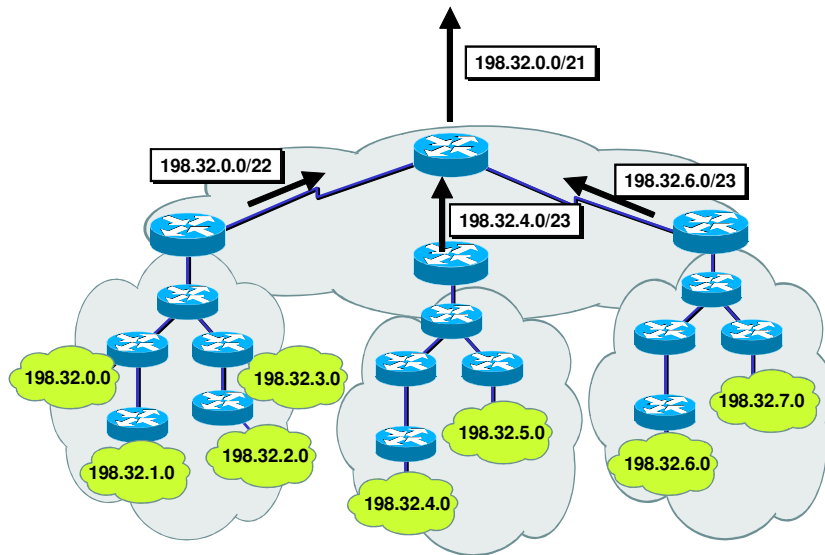
83

Esempio di Routing Tables senza CIDR



84

Esempio di Routing Tables con CIDR



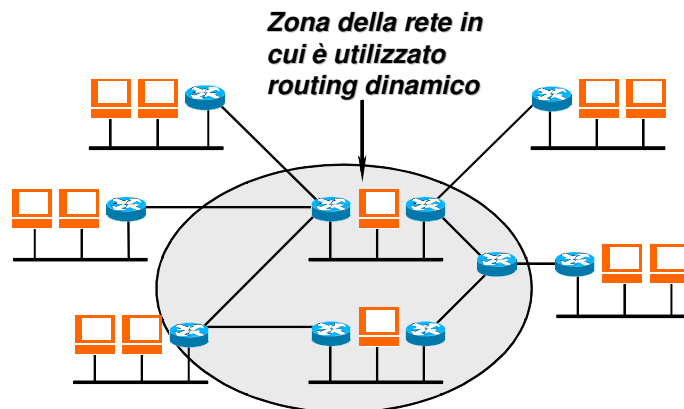
85

Tabelle di instradamento

- In base alle modalità con cui vengono create/aggiornate le tabelle di instradamento nei router si distinguono due tipi di instradamento (routing):
- routing statico
 - le tabelle vengono create/aggiornate staticamente e non sono funzione dello stato della rete
 - le tabelle vengono create/aggiornate dal gestore
 - il gestore ha un totale controllo dei flussi di traffico
 - deve intervenire manualmente per riconfigurare la rete
 - utilizzato ad es.
 - nella parte non magliata di reti IP
 - negli host, che in genere vengono configurati in base alle seguenti informazioni: indirizzo IP, net mask, default "gateway"
- routing dinamico
 - Le tabelle vengono calcolate con appositi algoritmi (di routing) e aggiornate periodicamente al variare dello stato della rete attraverso opportuni **protocolli di routing**

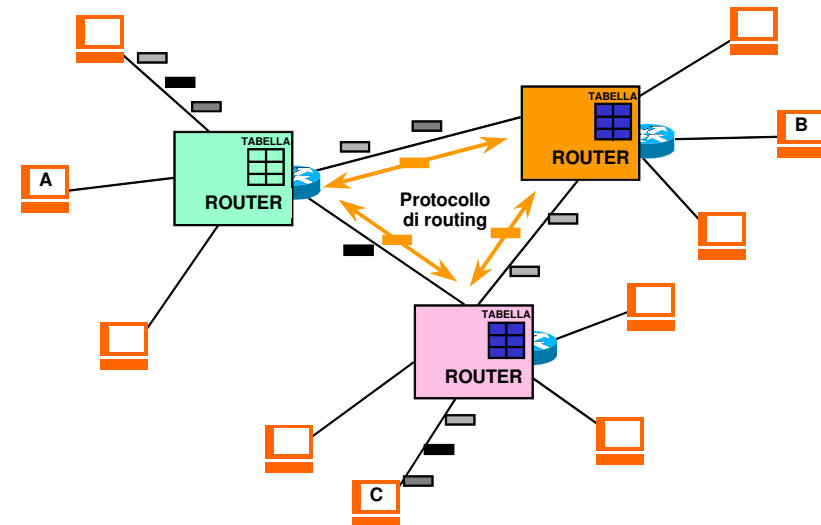
86

Routing Statico e Dinamico



87

Routing Dinamico

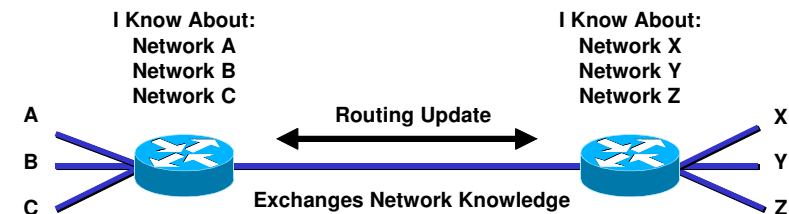


88

Routing protocols

Routing Dinamico

- Ogni router calcola le sue tabelle dialogando con gli altri router
- Tale dialogo avviene tramite dei protocolli (livello 3 o superiore) detti **protocolli di routing**
- I protocolli di routing sono utilizzati dai router per determinare il percorso per raggiungere le reti non direttamente connesse
- Esistono diversi protocolli di routing, ciascuno con caratteristiche più o meno attraenti: **RIP**, **EIGRP (CISCO)**, **OSPF**



90

Routing Dinamico

- Esistono due approcci principali al routing distribuito:
 - **Algoritmi Distance Vector**
 - più semplici
 - impegnano meno risorse sul router
 - meno efficienti
 - adatti a reti piccole
 - **Algoritmi Link State**
 - molto più complessi
 - molto più efficienti
 - impegnano più risorse
 - adatti a reti grandi

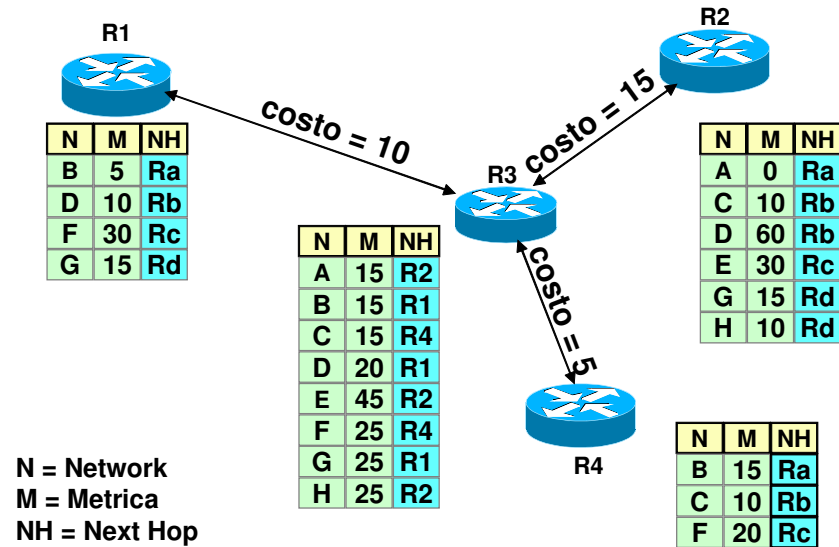
91

Distance Vector

- Noto anche come algoritmo di Bellman-Ford
- Ogni nodo mantiene un database con le distanze minime tra sé stesso e tutte le possibili destinazioni
- Ogni nodo, quando modifica le proprie tabelle di instradamento, invia ai nodi adiacenti un distance vector
- Il distance vector è un insieme di coppie
 - **[indirizzo - distanza]**
- Quando un nodo riceve un distance vector da un nodo adiacente, ricalcola la tabella delle distanze minime; se ci sono modifiche invia il suo nuovo distance vector (aggiornato) ai nodi adiacenti
- La distanza è espressa tramite metriche classiche quali numero di hops e costo

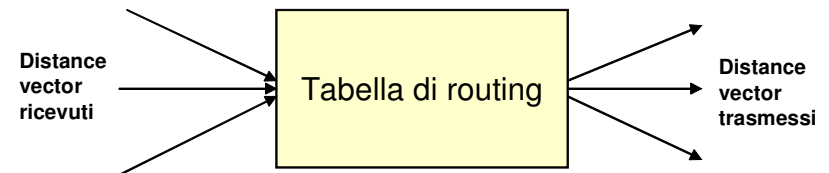
92

Distance Vector



93

Distance Vector



94

Distance Vector: caratteristiche

- Vantaggi:
 - Molto semplice da implementare
- Svantaggi
 - Possono innescarsi dei loop a causa di particolari variazioni della topologia
 - Converge alla velocità del link più lento e del router più lento
 - Difficile capirne e prevederne il comportamento su reti grandi: nessun nodo ha una mappa della rete!
 - L'implementazione di meccanismi migliorativi appesantisce notevolmente il protocollo

95

Link State

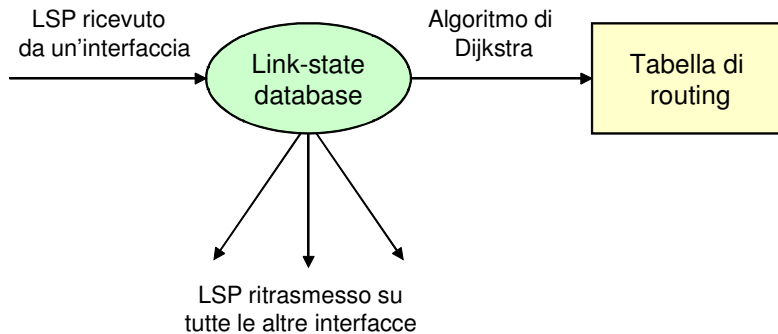
- Ogni router impara il suo ambito locale: linee e nodi adiacenti
- Trasmette queste informazioni a tutti gli altri router della rete tramite un Link State Packet (LSP)
- Tutti i router, memorizzando i LSP trasmessi dagli altri router, si costruiscono una mappa della rete
- Ogni router calcola indipendentemente le sue tabelle di instradamento applicando alla mappa della rete l'algoritmo di Dijkstra o SPF (Shortest Path First)
- La complessità è $E \log N$
 - E è il numero di link, N è il numero di nodi

96

Link State: operazione di un router

- Il LSP è trasmesso in flooding su tutti i link del router
- I LSP memorizzati formano una mappa completa della rete

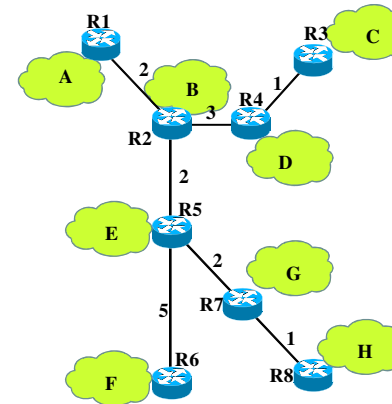
➤ Link State Database



97

Link State: tabella di routing

- Ogni router calcola indipendentemente le sue tabelle di routing applicando alla mappa della rete l'algoritmo di Dijkstra o SPF (Shortest Path First)



Nodo	LS Packet
R1	B(2)
R2	A(2) D(3) E(2)
R3	D(1)
R4	B(3) C(1)
R5	B(2) F(5) G(2)
R6	E(5)
R7	E(2) H(1)
R8	G(1)

Tabella di routing di R2

A	R1
C	R4
D	R4
E	R5
F	R5
G	R5
H	R5

98

LSP Flooding

- I LSP vengono trasmessi in flooding su tutti i link del router che li ha originati
- Un router che riceve un LSP lo ritrasmette in flooding solo se esso ha modificato il LSP database del router stesso (selective flooding)
- All'atto del ricevimento di un LSP un router compie le seguenti azioni:
 - se non ha mai ricevuto LSP da quel mittente o se il num di sequenza del LSP è maggiore di quello del LSP memorizzato nel database, allora memorizza il pacchetto nel database e lo ritrasmette in flooding su tutte le linee eccetto quella da cui l'ha ricevuto;
 - se il LSP ricevuto ha lo stesso numero di sequenza di quello posseduto, allora non viene fatto nulla;
 - se il LSP è più vecchio di quello posseduto, cioè è obsoleto, allora il router ricevente trasmette il LSP aggiornato al router mittente
- Questo meccanismo serve a fare in modo che i LSP database di tutti i router si mantengano perfettamente allineati e coerenti, condizione indispensabile per un corretto instradamento

99

Link State: caratteristiche

- Vantaggi:
 - Può gestire reti di grandi dimensioni
 - Ha una convergenza rapida
 - Difficilmente genera loop, e comunque è in grado di identificarli e interromperli facilmente
 - Facile da capire: ogni nodo ha la mappa della rete
- Svantaggi:
 - Più complesso da realizzare
- È utilizzato nel protocollo OSPF

100

Distance Vector vs. Link State

- Nel LS i router cooperano per mantenere aggiornata la mappa della rete, poi ogni router calcola il proprio spanning tree autonomamente; nel DV i router cooperano per calcolare direttamente le tabelle di instradamento
- L'algoritmo LS può gestire reti di grandi dimensioni (10000 nodi), il DV generalmente non supera i 1000
- LS ha convergenza rapida, difficilmente genera loop, e comunque è in grado di identificarli e interromperli facilmente; ed è facile da capire e prevedere poiché ogni nodo contiene l'intera mappa della rete

101

Protocolli di routing

	Algoritmo	Protocollo
Link State	Dijkstra SPF	OSPF
Distance Vector	Bellman-Ford	RIP IGRP (Cisco)

102

RIP

- Sviluppato dalla Xerox per XNS
- Nel 1982 il RIP è stato adattato per il TCP/IP con lo UNIX BSD
- Si tratta di un protocollo di routing intradominio basato su un algoritmo di tipo **distance vector**
- Definito dall'IETF nello RFC 1058 (1988) e nello RFC 1388 (1993)
- Metrica di costo: basata su hop count
 - Il RIP permette un massimo di 15 hop, superati i quali il percorso viene ritenuto irrealizzabile
- Messaggi di update: inviati ogni 30 s
 - In caso di link failure o modifica di topologia l'update avviene immediatamente
- Memorizzazione in tabella del solo percorso migliore verso la destinazione
- Usato dal demone "routed" in UNIX

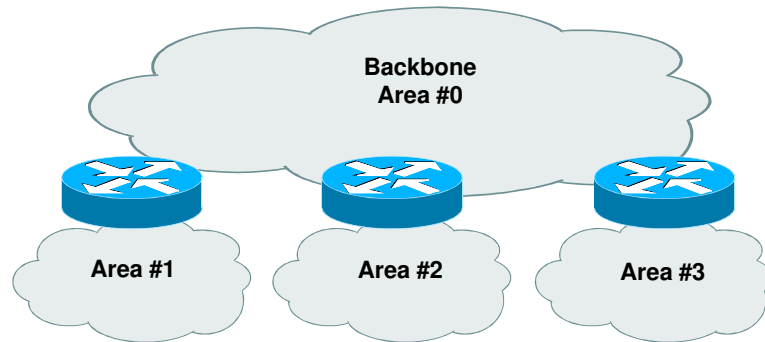
103

OSPF

- Protocollo di tipo link state
- Definito dall'IETF:
 - RFC 1247 (1991)
 - RFC 1583 (1994) - OSPFv2
- OSPF ha il concetto di gerarchia
 - un AS (dominio OSPF) è suddiviso in **aree**
 - le aree contengono un gruppo di reti contigue
 - le aree sono indicate da un area-id su 32 bit
 - deve essere specificato per ogni interfaccia
 - quando un AS ha più di un'area deve esistere una **backbone area** con area-id = 0

104

OSPF: aree



- La topologia di un'area è invisibile all'esterno dell'area
- Riduzione del traffico di routing

105

OSPF: metrica

- Il costo (o metrica) di un'interfaccia può essere legato alla larghezza di banda ad essa associata
 - **costo inversamente proporzionale alla banda**
 - $\text{costo} = 10^8 / (\text{bandwidth in bps})$

106

Protocolli di routing: confronto

	OSPF	RIP
Scalabilità	Buona	Bassa
Banda	Bassa	Alta
Memoria	Alta	Bassa
CPU	Alta	Bassa
Convergenza	Veloce	Lenta
Configurazione	Moderata	Facile

107

Sistemi autonomi

- Si definisce come sistema autonomo (*Autonomous System - AS*) un insieme di hosts, routers e reti fisiche controllate da una singola autorità amministrativa; ogni AS è identificato da un numero assegnato dal NIC
- Ogni AS è libero di scegliere i criteri di determinazione delle strade al suo interno
- Ogni AS deve però affidare in modo specifico ad uno o più routers il compito di comunicare al mondo esterno le informazioni di routing al suo interno
- Le informazioni di instradamento riguardanti le strade all'interno di un sistema autonomo sono gestite tra i router del AS per mezzo degli Interior Gateway Protocols (IGP)
- Le informazioni di instradamento riguardanti strade che coinvolgono più di un sistema autonomo sono scambiate mediante gli Exterior Gateway Protocols (EGP) tra i *core routers*

108

Interior and Exterior Gateway Protocols

- I protocolli di instradamento all'interno di un AS sono detti Interior Gateway Protocol (IGP)
 - **Routing Information Protocol (RIP)**
 - **Open Shortest Path First (OSPF)**
- Le informazioni di instradamento che coinvolgono più di un sistema autonomo sono gestite mediante gli Exterior Gateway Protocols (EGP)
 - **Border Gateway Protocol (BGP)**

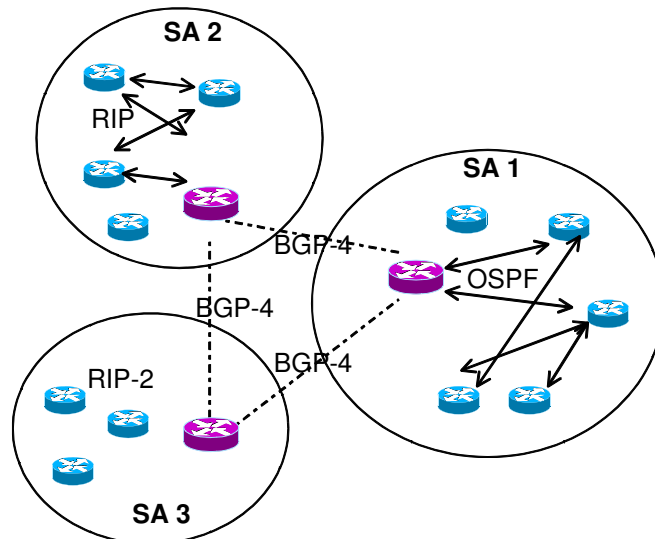
109

Interior and Exterior Gateway Protocols

- Un IGP ha il compito di
 - **individuare i router adiacenti nello stesso AS**
 - **raccogliere e distribuire a tutti i router i dati sulla topologia di un AS e sul costo di attraversamento dei rami**
 - **comunicare tempestivamente eventuali variazioni del costo di attraversamento dei rami di un AS**
- Un EGP ha il compito di
 - **individuazione dei router adiacenti di altri AS con cui scambiare le informazioni di instradamento**
 - **verifica continua della funzionalità dei router interlocutori**
 - **scambio periodico di informazioni di raggiungibilità delle reti**

110

IGP & EGP



Il protocollo ICMP

111

ICMP (Internet Control Message Protocol)

- ICMP (Internet Control Message Protocol) (RFC 792, 950)
- Utilizzato per la trasmissione dei messaggi di errore e di controllo relativi al protocollo IP
 - **errori di instradamento, TTL scaduto, congestione, etc**
- I messaggi vengono manipolati dal software IP, non dagli applicativi utente
- ICMP può quindi essere considerato un sub-strato di IP (visto che serve a trasportare messaggi tra due entità IP) ma è funzionalmente al di sopra di IP (visto che i suoi messaggi governano il funzionamento di IP)
- ICMP è una parte integrante di IP e deve essere incluso in ogni implementazione di IP
- Un messaggio ICMP è incapsulato nella parte dati di un datagramma IP

113

ICMP

- ICMP ha lo scopo esclusivo di notificare errori all'host di origine
 - **ICMP non specifica le azioni che devono essere prese per rimediare ai malfunzionamenti**
 - **spetta all'host di origine decidere le azioni da intraprendere per correggere il problema**
- I messaggi di ICMP viaggiano come comuni datagrammi, anch'essi possono essere soggetti ad errore e contribuire alla congestione di rete
- La procedura di gestione dei datagrammi prevede un'unica differenza tra i datagrammi che trasportano i messaggi ICMP e gli altri:
 - **non vengono generati messaggi ICMP in seguito ad errori causati da datagrammi che trasportano messaggi ICMP**
 - ciò serve ad evitare messaggi di errore relativi a messaggi di errore.
- Ogni messaggio ICMP è in relazione ad uno specifico datagramma
- Un messaggio di errore ICMP contiene quindi anche una parte del datagramma che ha generato l'errore (Intest. IP + primi 8 ottetti dei dati IP, i quali contengono le porte TCP o UDP di sorgente e destinazione)

114

ICMP

- Esempi:
 - **Source Quench**: inviato dal destinatario, interrompe l'emissione di datagrammi del mittente;
 - **Redirect**: il destinatario segnala al mittente di re-instradare il datagramma verso un altro host;
 - **Echo**: controlla se un possibile destinatario è attivo,
 - **Destination Unreachable**: notifica il mittente della non-raggiungibilità di un host

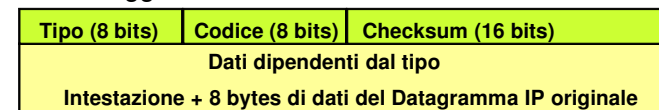
115

ICMP

- Un messaggio ICMP si riferisce ad uno specifico datagramma
- Un messaggio ICMP contiene l'indicazione del particolare datagramma IP che ha generato l'errore
 - **nel caso di frammentazione, un messaggio ICMP viene emesso solo per il frammento 0**
- Incapsulamento di un messaggio ICMP



- Formato messaggio ICMP



116

ICMP: tipi di messaggio

Tipo	Descrizione
0	Echo Reply
3	Destination Unreachable
4	Source Quench
5	Redirect (change a route)
8	Echo Request
11	Time Exceeded for a Datagram
12	Parameter Problem for a Datagram
13	Timestamp Request
14	Timestamp Reply
15	Information Request
16	Information Reply
17	Address Mask Request
18	Address Mask Reply

117

ICMP: codici

- Codici per messaggio 'destinazione irraggiungibile' (tipo 3)

0	Rete irraggiungibile
1	Host irraggiungibile
2	Protocollo irraggiungibile
3	Porta irraggiungibile
4	Frammentazione necessaria e DF settato
5	Fallimento routing sorgente
6	Rete destinazione sconosciuta
7	Host destinazione sconosciuto
8	Host sorgente isolato
9	Comunicazione con rete destinazione proibita amministrativamente
10	Comunicazione con host destinazione proibita amministrativamente
11	Rete irraggiungibile per il tipo di servizio
12	Host irraggiungibile per il tipo di servizio

118

ICMP

- Redirect message
 - se è emesso da un router significa che i successivi datagrammi emessi dall'host verso la rete dovranno essere indirizzati verso il router indicato nel messaggio ICMP
 - causa una modifica della tabella di instradamento dell'host sorgente
- Source quench
 - se è emesso da un router intermedio indica che il router non ha buffer sufficiente per memorizzare il datagramma
 - se è emesso dall'host di destinazione indica che il datagramma non è stato processato dall'host
 - il messaggio è utilizzato dal TCP
- Time exceeded
 - indica che il TTL si è esaurito

119

ICMP

- Echo e Echo replay
 - sono utilizzati per stabilire l'attività di un elemento di un host
- Destination unreachable
 - indica che l'instradamento di un datagramma non è stato completato
- Time Stamp Request e Time Stamp Reply
 - sono utilizzati per effettuare misure di prestazioni (es. ritardi di transito)
- Address mask request e Address mask replay
 - sono usati per determinare la maschera della sotto-rete a cui è connesso un host
 - sono usati da host molto semplici (diskless) dopo aver individuato il proprio indirizzo con il protocollo RARP

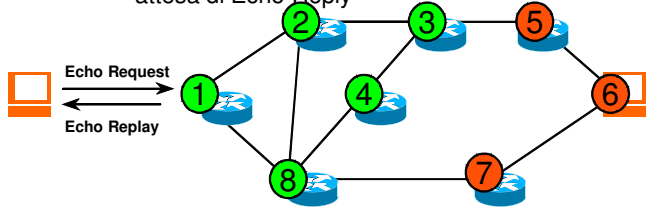
120

Applicazioni dell'ICMP

- Ping
 - è utilizzata per verificare
 - l'installazione della pila TCP/IP
 - l'attività di un host
 - il tempo di transito tra host sorgente e host destinazione
 - utilizza i messaggi ICMP Echo e Echo Replay
- Traceroute
 - determina la sequenza di router attraversati da un datagramma tra l'host sorgente e l'host destinazione
 - utilizza in successione datagrammi con TTL=1, 2, 3, ...
 - Il TTL viene decrementato da ogni router lungo il percorso
 - il datagramma con TTL=1 raggiunge solo il 1° router, TTL=2 raggiunge il 2° router, e così via
 - la sequenza di router viene individuata dai router che in successione scartano i pacchetti per TTL=0 e inviano di conseguenza al host sorgente un messaggio ICMP Time Exceeded

121

Ping

- PING
 - diagnosi di raggiungibilità
 - generazione di pacchetti di Echo Request verso "Echo Server"
 - attesa di Echo Reply
- 
- Problema
 - scarsa capacità diagnostica
 - ★ cosa significa se 5, 6 e 7 non rispondono al PING ?
 - ★ ci sono decine di possibili cause
 - ★ si può migliorare facendo PING da sorgenti diverse

122

Ping

```
[user]$ ping pinco.pallino.net
```

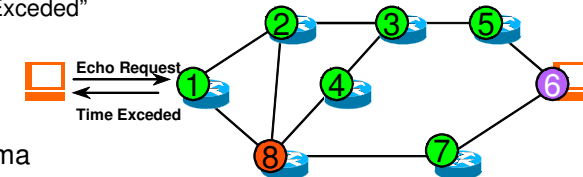
```
PING pinco.pallino.net (193.200.242.5): 56 data bytes
64 bytes from 193.200.242.5: icmp_seq=0 ttl=248 time=111.4 ms
64 bytes from 193.200.242.5: icmp_seq=1 ttl=248 time=90.2 ms
64 bytes from 193.200.242.5: icmp_seq=2 ttl=248 time=116.2 ms
64 bytes from 193.200.242.5: icmp_seq=3 ttl=248 time=80.6 ms
64 bytes from 193.200.242.5: icmp_seq=4 ttl=248 time=80.1 ms
64 bytes from 193.200.242.5: icmp_seq=5 ttl=248 time=537.4 ms
```

```
--- pinco.pallino.net ping statistics ---
6 packets transmitted, 6 packets received, 0% packet loss
round-trip min/avg/max = 80.1/169.3/537.4 ms
```

123

Traceroute

- TraceRoute
 - identificazione dei percorsi sulla rete
 - generazione di più pacchetti successivi di Echo Request
 - TTL inizia da 1 e viene incrementato di 1 ad ogni successivo Echo Request
 - ogni pacchetto percorre un passo in più rispetto al precedente
 - osservazione dell'indirizzo sorgente dei pacchetti di "Time Exceeded"



- Problema
 - come per il ping scarsa capacità diagnostica
 - ★ Esempio: cosa succede se il percorso dei pacchetti nelle due direzioni è diverso (Es. 1,2,3,5,6,5,3,4,8,1) ed il nodo 8 è guasto ?

124

Traceroute

traceroute to www.stanford.edu (171.64.14.203), 30 hops max, 40 byte packets

```

1 151.100.238.1 (151.100.238.1) 2678.773 ms
2 rc-uniromaI.rm.garr.net (193.206.131.49) 1859.746 ms
3 rt-rc-2.rm.garr.net (193.206.134.165) 788.237 ms
4 mi-rm-1.garr.net (193.206.134.17) 766.614 ms
5 ny-mi.garr.net (212.1.200.17) 894.860 ms
6 Abilene-DANTE.abilene.ucaid.edu (212.1.200.222) 1118.096 ms
7 clev-nycm.abilene.ucaid.edu (198.32.8.29) 970.481 ms
8 ipls-clev.abilene.ucaid.edu (198.32.8.25) 1161.797 ms
9 kscy-ipls.abilene.ucaid.edu (198.32.8.5) 967.958 ms
10 denv-kscy.abilene.ucaid.edu (198.32.8.13) 1200.059 ms
11 scrm-denv.abilene.ucaid.edu (198.32.8.1) 985.121 ms
12 BERK--abilene.POS.calren2.net (198.32.249.41) 1166.336 ms
13 SUNV--BERK.POS.calren2.net (198.32.249.14) 1087.366 ms
14 STAN--SUNV.POS.calren2.net (198.32.249.74) 962.810 ms
15 i2-gateway.Stanford.EDU (171.64.1.214) 566.572 ms
16 Core3-gateway.Stanford.EDU (171.64.1.222) 215.399 ms
17 sweet-gateway.Stanford.EDU (171.64.3.110) 215.441 ms
18 www1.Stanford.EDU (171.64.14.203) 215.697 ms
    
```

125

Address Resolution Protocol (ARP)

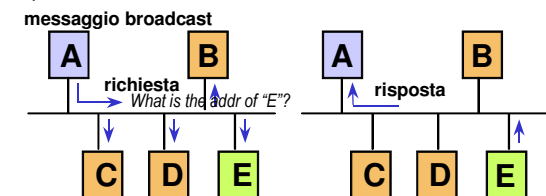
Risoluzione degli indirizzi

- Ogni qual volta si deve rilanciare un pacchetto da un nodo ad un altro bisogna conoscere l'indirizzo di sottorete (relativi al protocollo sotto IP) del nodo next hop identificato dal suo indirizzo IP
 - e' necessaria quindi una funzione di mappaggio da indirizzi IP ad indirizzi di sottorete (e.g. indirizzo MAC di una stazione LAN)
 - questa funzione può essere svolta sulla base di una tabella di mapping
- Mapping statico
 - la tabella di associazione viene predisposta staticamente (ad esempio rete X.25, ISDN, etc.)
- Mapping dinamico
 - la tabella viene costruita dinamicamente attraverso un protocollo ARP (Address Resolution Protocol) RFC826
 - broadcast (sulle LAN)
 - ARP-Server (su reti Non Broadcast)

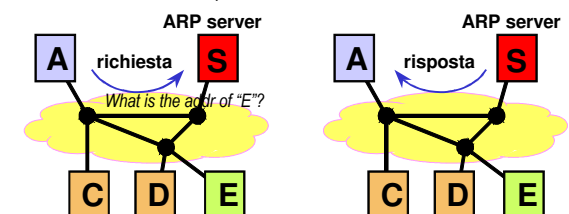
127

Risoluzione degli indirizzi (cont.)

- broadcast (sulle LAN)



- ARP-Server (su reti Non Broadcast)



128

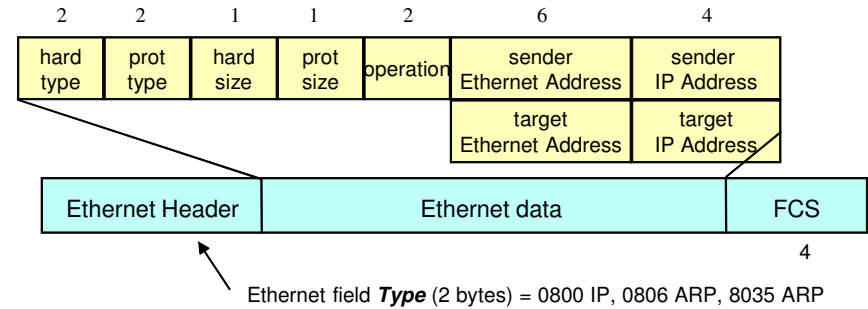
ARP

- Il protocollo ARP (Address Resolution Protocol) fornisce un meccanismo dinamico di associazione fra indirizzi MAC ed indirizzi IP
- Viene utilizzato ogni qual volta un nodo di una LAN debba inviare un pacchetto ad un altro nodo della stessa LAN di cui però conosca solo l'indirizzo IP

129

ARP

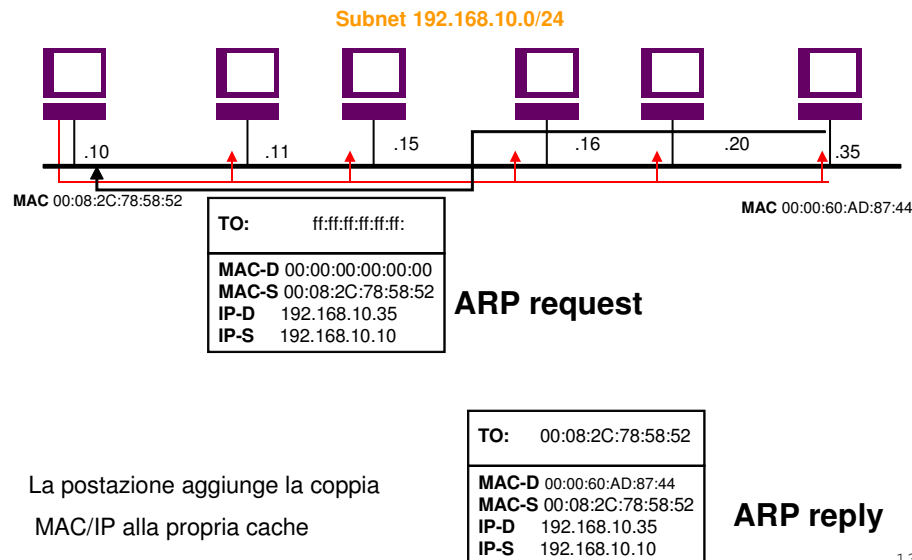
- Il formato dei pacchetti è identico per ARP e RARP
- ARP/RARP si appoggiano direttamente sullo strato MAC (non su IP). Ciò significa che un opportuno campo nell'intestazione della MAC-PDU indica se il contenuto deve essere consegnato a ARP, RARP o IP



- ARP prevede caching delle informazioni
il comando `arp -a` permette di visualizzare il contenuto della cache

130

ARP



131

Configurazione automatica di un nodo IP: da RARP a DHCP e PPP

RARP

- RARP (Reverse Address Resolution Protocol) (RFC 903):
 - dato un indirizzo di scheda di rete, permette di ricavare il corrispondente indirizzo IP
 - si utilizza(va) per le workstation diskless per effettuare il bootstrap tramite immagine su server remoto
 - utilizza un RARP server
 - l'host emette un pacchetto RARP con indirizzo di sottorete (Ethernet) broadcast con il quale richiede il proprio indirizzo IP in base al proprio indirizzo di scheda di rete (Ethernet)
 - il server RARP riceve la richiesta, consulta un file di configurazione e risponde con l'indirizzo IP corrispondente all'indirizzo della scheda di rete del richiedente
 - svantaggi:
 - l'applicazione server deve operare a livello di protocollo di sotto-rete
 - permette di acquisire solo l'indirizzo IP
 - il server deve essere nella stessa sotto-rete

133

BOOTP

- BOOTP (BOOTstrap Protocol) (RFC 951, 1048, 1084)
 - permette di acquisire all'avvio (bootstrap) informazioni di configurazione quali: (proprio IP address, indirizzo di un router, indirizzo di un server,...)
 - è adatto per host senza disco rigido
 - vengono scambiati messaggi (UDP) tra l'host e il BOOTP server utilizzando il limited broadcast address (255.255.255.255) come destinazione
 - il server risponde ad una richiesta e, in base al client identifier specificato dall'host richiedente, fornisce: l'indirizzo IP dell'host, l'indirizzo IP di un router, l'indirizzo IP di un server ecc.
 - lo svantaggio principale è la staticità; ovvero:
 - richiede l'inserimento delle informazioni nel file di configurazione
 - non può essere usato per assegnare dinamicamente gli indirizzi IP a domanda (mappaggio statico)

134

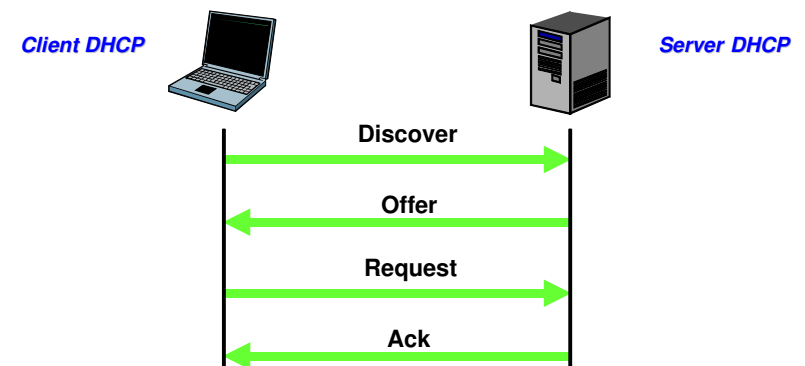
Dynamic Host Configuration Protocol

- DHCP (RFC 2131) è un'estensione di BOOTP che permette di:
 - assegnare dinamicamente gli indirizzi IP
 - ricevere altre informazioni di configurazione (esempio: subnet mask)
- DHCP è utile quando:
 - i computer si connettono e disconnettono frequentemente
 - il numero di computer supera il numero di indirizzi disponibili nella sottorete
 - si vuole rendere automatica l'assegnazione degli indirizzi IP
- DHCP permette tre tipi di configurazione: manuale, automatica e dinamica
- in configurazione dinamica un DHCP server assegna un IP address (tra un insieme di indirizzi disponibili) ad un host che ne effettua richiesta per un tempo limitato

135

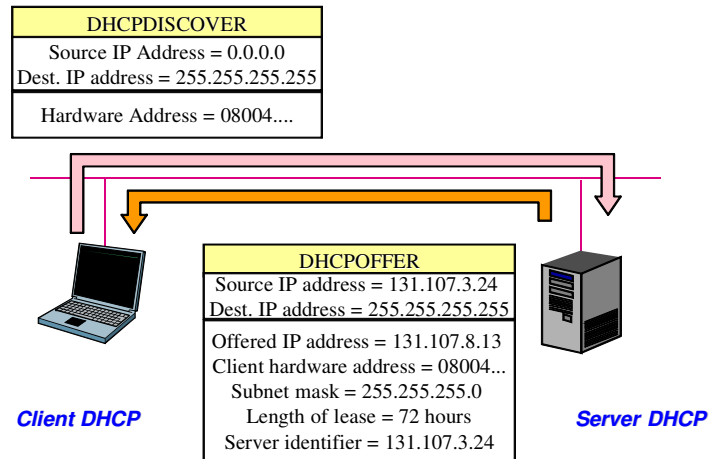
Dynamic Host Configuration Protocol

- DHCP utilizza un processo in quattro fasi per configurare un client
 - discover
 - offer
 - request
 - ack



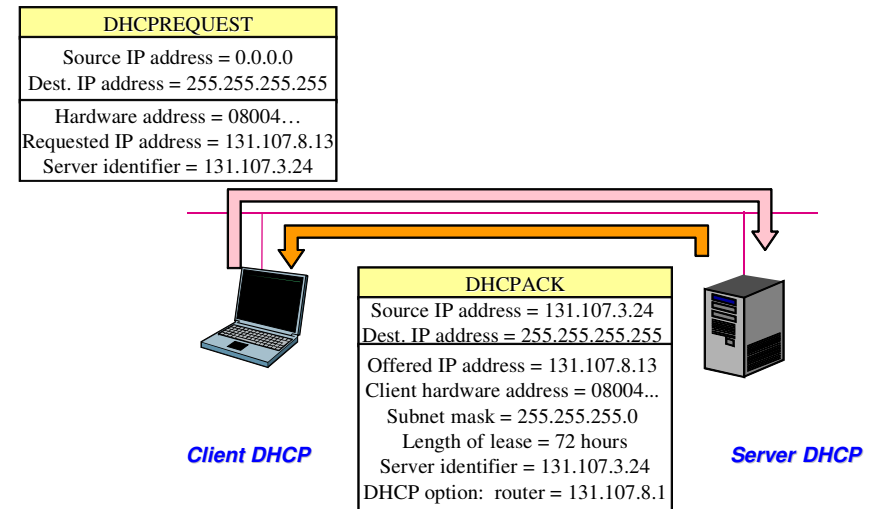
136

DHCP: Discover e Offer



137

DHCP: Request e Ack



138

Point to Point Protocol

- PPP (Point to Point Protocol) è definito nell'RFC1661
- è un protocollo di livello 2 (Data Link) per collegamenti punto-punto
 - attualmente il protocollo di livello 2 più usato in Internet per collegamenti punto-punto
- oltre alle normali funzioni di Data Link (delimitazione di trama, controllo e recupero di errore, ecc) permette di:
 - supportare differenti protocolli di livello 3 (tra cui IP)
 - negoziare informazioni di configurazione di livello 3 (nel caso di IP: host_address, default router/gateway, DNS)

Risoluzione degli indirizzi (DNS)

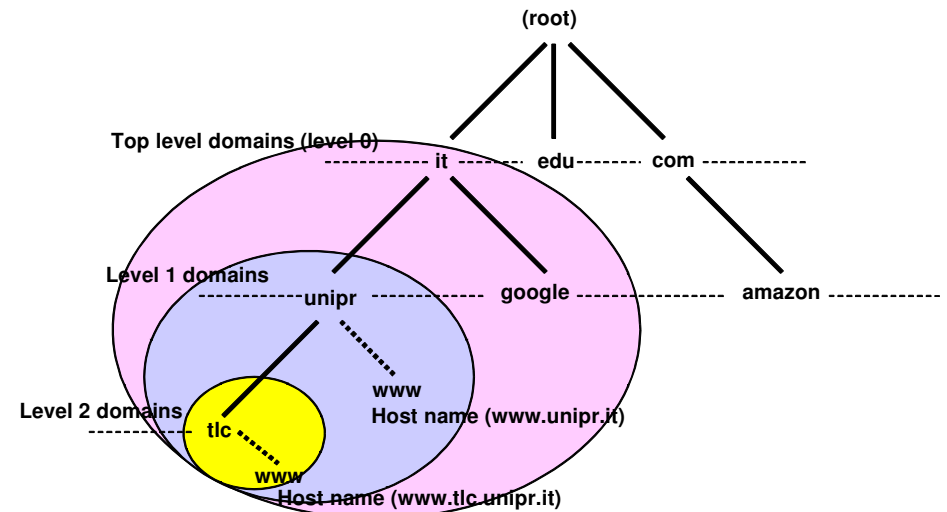
139

Domain Name System (DNS)

- Oltre alla notazione dotted viene spesso utilizzata anche un'altra forma di notazione (mnemonica):
 - e.g. "160.78.48.141" = "www.unipr.it"
- E' necessaria la funzionalità di traduzione di nomi mnemonici in indirizzi e viceversa
 - questa traduzione è attuata da un protocollo di alto livello implementato in un meccanismo noto come Domain Name System (DNS)
- I nomi sono organizzati gerarchicamente in domini
 - i nomi sono costituiti da stringhe separate da "."
 - la parte più significativa è a destra
 - e.g. www.tlc.unipr.it

141

Domain Name System (DNS)



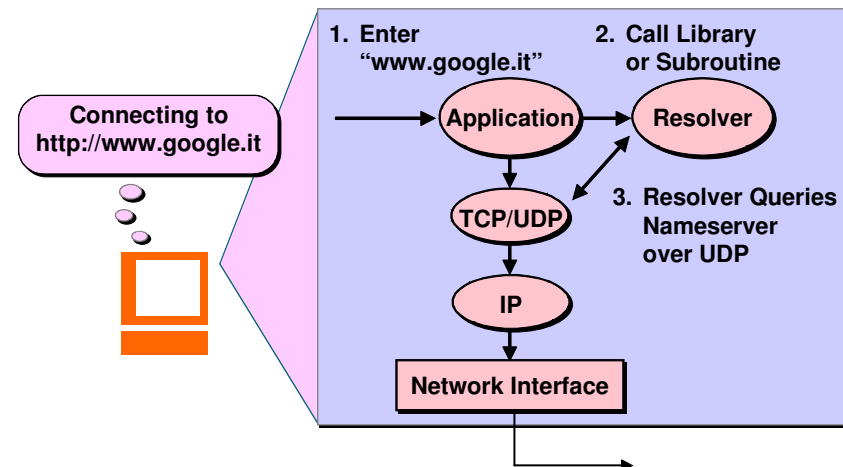
142

DNS (Domain Name System)

- Generic top level domains
 - aereo** compagnie aeree
 - biz** commerciale (simile a .com)
 - com** organizzazioni commerciali (hp.com, sun.com ...)
 - coop** cooperative
 - edu** organizzazioni educative (berkeley.edu, purdue.edu ...)
 - gov** organizzazioni governative (nasa.gov, nsf.gov ...)
 - info** punti di informazione
 - int** organizzazioni internazionali (nato.int ...)
 - mil** organizzazioni militari (army.mil, navy.mil ...)
 - museum** musei
 - name** siti personali
 - net** organizzazione di gestione reti (nsf.net ...)
 - org** organizzazioni non commerciali (eff.org ...)
 - pro** organizzazioni professionali
- National top level domains
 - country-code** codice di due caratteri per indicare una nazione

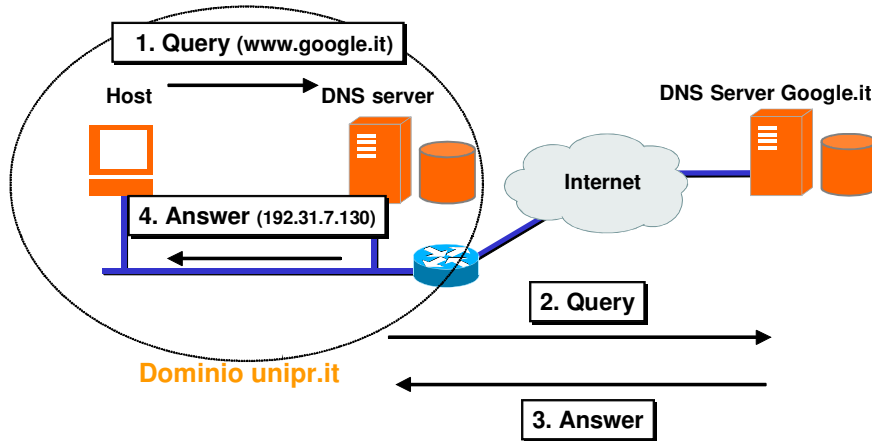
143

Risoluzione dei nomi



144

Risoluzione dei nomi (cont.)



145

Risoluzione dei nomi (cont.)

- Quando un server DNS riceve una query per un nome appartenente ad un dominio di cui non ha autorità effettua le seguenti operazioni:
 - **verifica nella cache se è presente il nome da risolvere**
 - la cache contiene i record dei nomi risolti più di recente
 - **invia la query ad uno dei root name server specificati in un file denominato cache file**
- Un server DNS può ricevere due tipi di richieste di risoluzione
 - **risoluzione ricorsiva**
 - se il server conosce la risposta, questa viene restituita immediatamente, altrimenti il server inoltra la richiesta ad un altro server DNS e attende la risposta
 - **risoluzione iterativa**
 - se il server conosce la risposta, questa viene restituita immediatamente, altrimenti il server restituisce l'indirizzo IP di un DNS a cui rivolgere la richiesta

146

Cenni a IPv6

Perchè IPv6

- Esaurimento dello spazio di indirizzamento IPv4
- Esplosione delle tabelle di instradamento sui router
- Servizi nuovi e/o più efficienti
 - **Soluzione di networking "plug&play"**
 - **Qualità del Servizio**
 - **Sicurezza**
 - **Mobilità**
 - **Multicast**

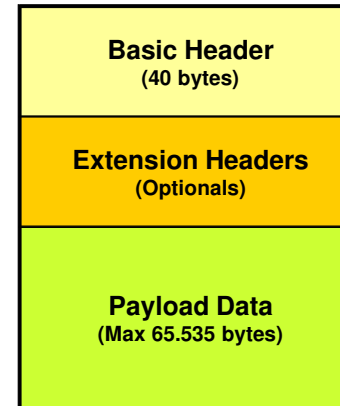
148

Header IPv6/IPv4

- Semplificazioni:
 - l'header IPv6 ha lunghezza fissa (40 byte)
 - le opzioni non sono più trasportate all'interno dell'header IP
 - questa funzione viene svolta dagli extension header di IPv6
 - è stato rimosso il campo IP Header Length (IHL)
 - non più necessario in quanto l'header IPv6 ha lunghezza fissa
 - è stato rimosso il campo Header Checksum
 - quasi tutti i protocolli di livello data link comprendono già il calcolo e la verifica di un checksum
 - non esiste più la procedura di segmentazione hop-by-hop
 - di conseguenza sono stati rimossi i campi Identification, Flags e Fragment Offset
 - è stato rimosso il campo Type Of Service (TOS)
- Nuovi campi:
 - Flow Label
 - Class

149

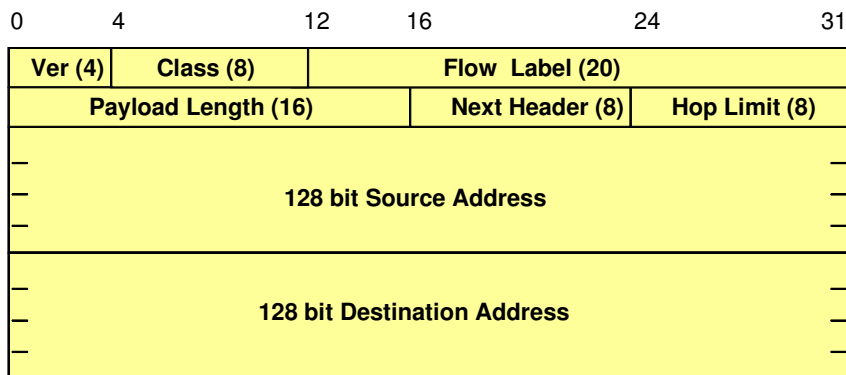
Formato generale datagramma IPv6



- Basic header
 - contiene le informazioni comuni a tutti i datagrammi
- Extension Headers
 - contengono le opzioni utilizzate dai router intermedi e/o dall'host di destinazione
- Payload Data
 - sono i bit informativi elaborati dall'host di destinazione

150

IPv6 Basic Header



40 ottetti
8 campi

151

Spazio di indirizzamento IPv6

- Campi di indirizzo di 128 bit (16 byte)
 - n. di indirizzi IPv6 totale $\approx 3.40 \cdot 10^{38}$
 - paria a $\approx 6,65 \cdot 10^{23}$ indirizzi/m² di superficie della Terra
 - o anche un indirizzo ogni $\approx 5 \text{ cm}^3$ di spazio nella sfera di raggio pari alla distanza Terra-Sole
- In realtà, poiché 64 bit sono riservati per l'identificatore dell'interfaccia (indirizzi unicast), si avranno:
 - n. totale di reti IP $\approx 1.8 \cdot 10^{19}$
 - paria a ≈ 36.000 reti IPv6 per m² di superficie terrestre

152

IPv6 Basic Header

- Version (4 bit)
 - **versione del protocollo (6), è possibile la coesistenza di più versioni di IP**

- Traffic Class (8 bit)
 - **Originariamente definito come campo Priority di 4 bit (RFC 1883)**
 - due classi di priorità
 - Congestion Controlled Traffic (livelli 0 - 7)
 - Noncongestion Controlled Traffic (livelli 8 - 15)
 - La relazione di priorità ha valore solo all'interno di una classe
 - Non è definita nessuna relazione di priorità tra datagrammi appartenenti a classi diverse
 - **Ridefinito come DS Field (Diffserv Field) da RFC 2474**
 - **Stabilisce la classe di traffico o più precisamente il PHB (per-hop behavior) con cui deve essere trattato il pacchetto**

153

IPv6 Basic Header

- Flow label (20 bit)
 - **Originariamente di 24bit (RFC 1883)**
 - **Ha lo scopo di identificare, insieme ai campi source e destination address, un particolare flusso di datagrammi**
 - **E' un numero scelto casualmente dall'host mittente nell'intervallo 1-0xFFFFFFFF (0 identifica traffico non associato a nessun flusso)**
 - **Dovrebbe aiutare a ridurre i tempi di elaborazione dei datagrammi nei router di rete**
 - **Potrebbe facilitare l'istadamento dei datagrammi in hardware mediante consultazione di tabelle di cache evitando l'applicazione della normale procedura di IP forwarding**
 - **Il concetto di flusso si adatta anche a procedure di riservazione di risorse per traffico con qualità di servizio garantita (protocollo RSVP)**

154

IPv6 Basic Header

- Payload Length (16 bit)
 - **indica la lunghezza in byte del datagramma IP (escluso il basic header)**
 - **normalmente la lunghezza massima del payload è 65.535 byte; è possibile l'uso dell'opzione "jumbo payload" (hop-by-hop options header)**

- Next Header (8 bit)
 - **identifica quali header seguono il basic header nel datagramma**
 - **alcuni valori:**

0	Hop-by-hop options header	46	Resource Reservation Protocol
4	Internet Protocol	50	Encapsulating Security Payload
6	Transmission Control Protocol	51	Authentication Header
17	User Datagram Protocol	58	Internet Control Message Protocol
43	Routing Header	59	No Next Header
44	Fragment Header	60	Destination Options Header

155

IPv6 Basic Header

- Hop Limit (8 bit)
 - **l'host sorgente indica il numero massimo di tratte di rete che il datagramma può attraversare**
 - **ogni router decrementa di una unità tale campo**
 - **se il contatore si azzerava prima che la destinazione sia raggiunta, il datagramma è scartato**
 - **evita gli effetti di eventuali loop in rete e può essere utilizzato per effettuare delle ricerche di host in rete a distanza prefissata**

- Source Addresses (128 bit)
 - **indica l'indirizzo IP del host sorgente**

- Destination Addresses (128 bit)
 - **indica l'indirizzo IP del/degli host di destinazione**
 - **potrebbe non essere il destinatario finale, se è presente un Routing header**

156

Sintassi degli indirizzi IPv6

- Per rappresentare formalmente gli indirizzi IPv6 si è scelto di suddividerli in 8 blocchi di 16 bit ciascuno
- I blocchi sono separati mediante il carattere ":" e vengono rappresentati in notazione esadecimale
- Un esempio di indirizzo IPv6 è:
 - **3FFE:1001:7654:3220:FEDC:BA98:789A:32AC**
- Esistono delle semplificazioni:
 - **si possono omettere gli zeri iniziali di ogni blocco**
 - 3ffe:1001:1:100:a00:20ff:fe83:5531
 - **si possono sostituire gruppi di zeri con "::"**
 - 3ffe:1001:1::1

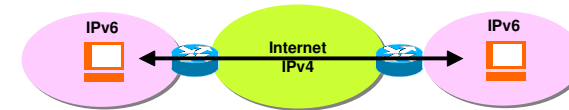
157

La transizione ad IPv6

- Elemento chiave per una transizione con successo ad IPv6 è la compatibilità con la base installata IPv4
- Durante la fase di transizione sono necessari meccanismi per:
 - **Permettere il colloquio tra nuovi host IPv6 e la Internet preesistente IPv4**



- **Realizzare connettività tra isole IPv6 utilizzando l'infrastruttura IPv4**



- Inoltre, reti IPv6 verranno utilizzate per interconnettere reti IPv4 (e.g. backbone UMTS)

158

La transizione ad IPv6

- Lo sviluppo della rete IPv6 avverrà in modalità sovrapposta
 - **Dual stacked nodes**
 - i nodi implementano entrambe le pile protocollari IPv4 e IPv6
 - **Tunneling**
 - il traffico IPv6 viene trasportato da IPv4 mediante tunnel
- E tramite meccanismi di traduzione
 - **protocol translation**
 - il traffico IPv6 viene tradotto in IPv4 e viceversa

159