




IP Routing

Silvano Gai
Mario Baldi
Pietro Nicoletti




Nota di Copyright

- Questo insieme di trasparenze (detto nel seguito slides) è protetto dalle leggi sul copyright e dalle disposizioni dei trattati internazionali. Il titolo ed i copyright relativi alle slides (ivi inclusi, ma non limitatamente, ogni immagine, fotografia, animazione, video, audio, musica e testo) sono di proprietà degli autori indicati a pag. 1.
 - Le slides possono essere riprodotte ed utilizzate liberamente dagli istituti di ricerca, scolastici ed universitari afferenti al Ministero della Pubblica Istruzione e al Ministero dell'Università e Ricerca Scientifica e Tecnologica, per scopi istituzionali, non a fine di lucro. In tal caso non è richiesta alcuna autorizzazione.
 - Ogni altra utilizzazione o riproduzione (ivi incluse, ma non limitatamente, le riproduzioni su supporti magnetici, su reti di calcolatori e stampate) in toto o in parte è vietata, se non esplicitamente autorizzata per iscritto, a priori, da parte degli autori.
 - L'informazione contenuta in queste slides è ritenuta essere accurata alla data della pubblicazione. Essa è fornita per scopi meramente didattici e non per essere utilizzata in progetti di impianti, prodotti, reti, ecc. In ogni caso essa è soggetta a cambiamenti senza preavviso. Gli autori non assumono alcuna responsabilità per il contenuto di queste slides (ivi incluse, ma non limitatamente, la correttezza, completezza, applicabilità, aggiornamento dell'informazione).
 - In ogni caso non può essere dichiarata conformità all'informazione contenuta in queste slides.
 - In ogni caso questa nota di copyright non deve mai essere rimossa e deve essere riportata anche in utilizzi parziali.
- 

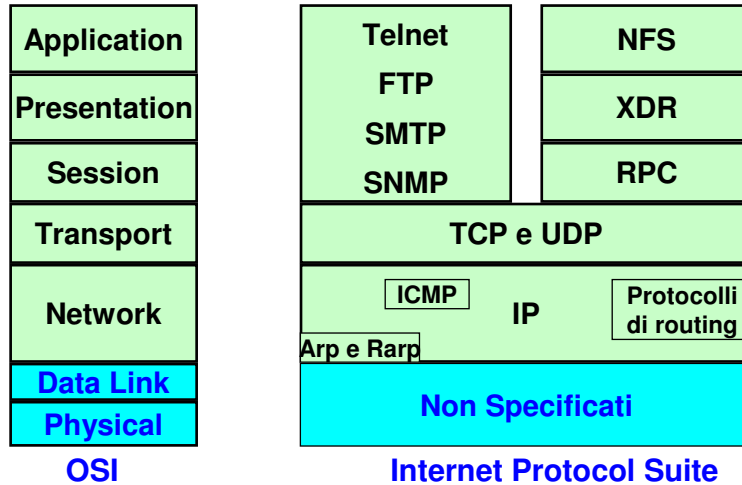
Cenni storici

- Nella prima metà degli anni '70 la DARPA (Defence Advanced Research Project Agency) dimostra interesse per la realizzazione di una rete:
 - a commutazione di pacchetto
 - tra elaboratori eterogenei
 - per le istituzioni di ricerca degli USA
- DARPA finanzia l'Università di Stanford e la BBN (Bolt, Beranek e Newman)

Cenni storici

- Verso la fine degli anni '70 si completa la realizzazione dell'Internet Protocol Suite, di cui i due principali protocolli sono:
 - IP: Internet Protocol
 - TCP: Transmission Control Protocol
- Da cui il nome TCP/IP usato per questa architettura di rete
- Nasce la rete Arpanet, prima rete della ricerca mondiale che evolve e diventa Internet

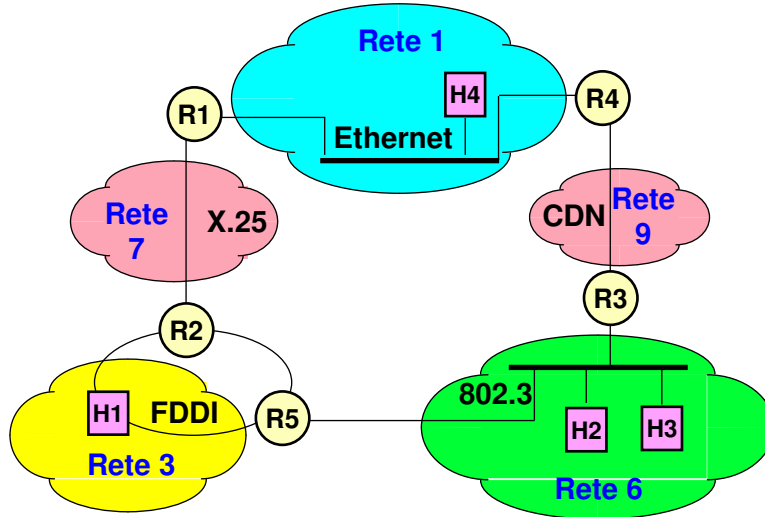
I livelli 1 e 2



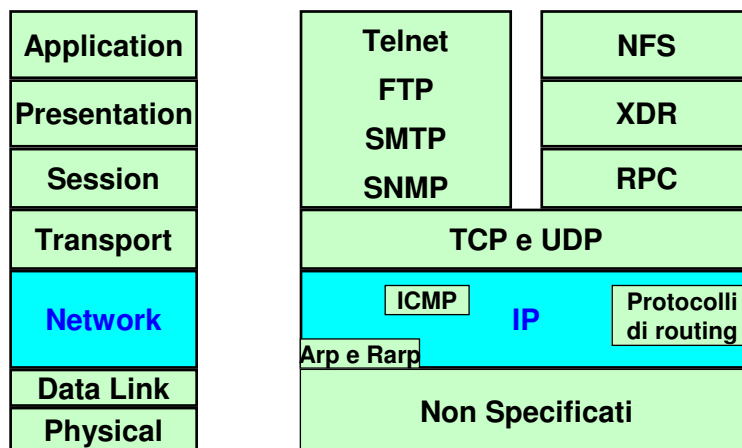
Sotto l'IP - I livelli 1 e 2

- L'architettura TCP/IP è concepita come un mezzo per fare internetworking tra reti (locali o geografiche)
- È in grado di operare su tutte le reti:
 - Ethernet, token-ring, FDDI
 - ATM, SMDS, Frame Relay
 - X.25
 - SLIP, PPP, Dialup
- Esistono realizzazioni di TCP/IP anche per reti non standard

Gerarchia



L'IP: Internet Protocol



OSI

Internet Protocol Suite

IP: Internet Protocol

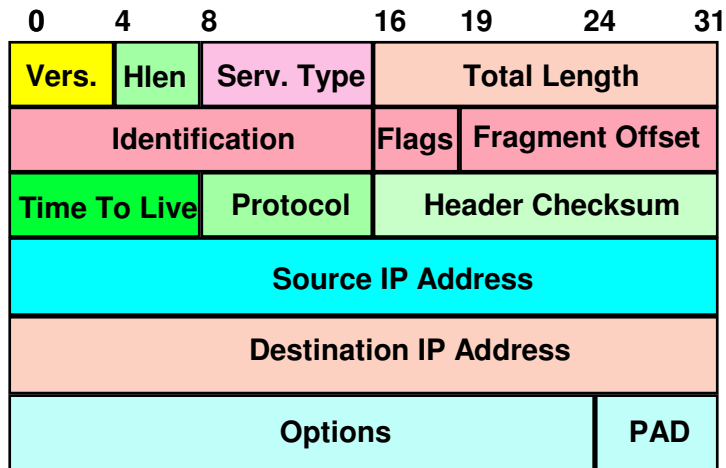
- È il livello Network di TCP/IP
- Offre un servizio non connesso
- Semplice protocollo di tipo Datagram
- Un protocollo datato ...
- ... ma non obsoleto

*Oggi: IP-v.4
Domani: IP-v.6*

IP: funzionalità

- Frammentazione e riassettaggio dei pacchetti
- Gestione indirizzi a 32 bit a livello di rete e di host
- Routing
- Monitoring della connessione
- Configurazione di classi di servizio

Formato del header IPv4



Campi dell'header IPv4

- **VER (4 bit)**
 - versione del protocollo (attualmente 4)
- **HLEN (4 bit)**
 - lunghezza dell'intestazione in multipli di 4 bit
- **TYPE OF SERVICE (ToS = 8 bit)**
 - Tipo di servizio che il pacchetto deve ricevere dalla rete
- **TOTAL LENGTH (16 bit)**
 - lunghezza totale del pacchetto, inclusa intestazione e campo dati
- **IDENTIFIER (16 bit)**
 - Identificatore comune a tutti i frammenti di uno stesso pacchetto

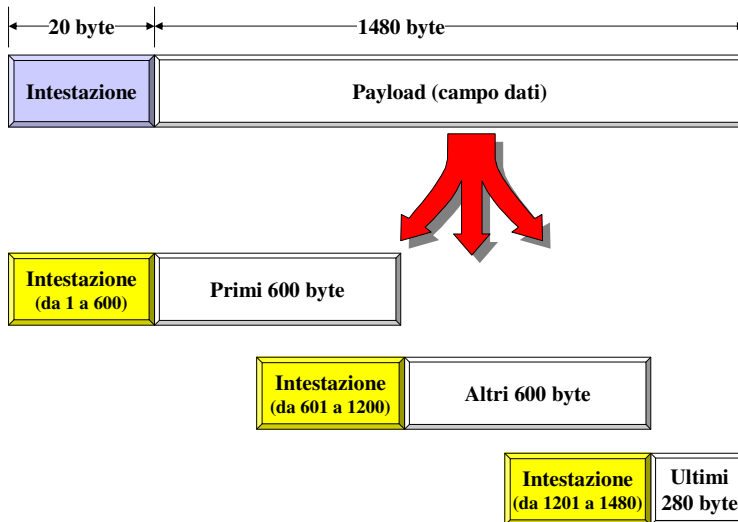
Campi dell'header IPv4

- **FLAGS (3 bit)**
 - Bit 0: reserved, must be zero
 - Bit 1: (DF) 0 = May Fragment, 1 = Don't Fragment.
 - Bit 2: (MF) 0 = Last Fragment, 1 = More Fragments.
- **FRAGMENT OFFSET (13 bit)**
 - posizione (in multipli di 8 byte) del primo byte del frammento nel pacchetto originale
- **TIME TO LIVE (TTL = 8 bit)**
 - tempo di vita rimanente in termini di numero di router che il pacchetto può ancora attraversare prima di essere scartato

Campi dell'header IPv4

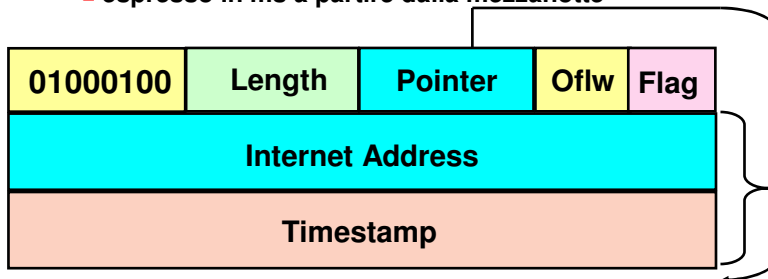
- **PROTOCOL (8 bit)**
 - Identificativo del protocollo trasportato
- **CHECKSUM (16 bit)**
 - Codice per il rilevamento di errori nell'intestazione IP
- **SOURCE ADDRESS (32 bit)**
 - Indirizzo IP del mittente
- **DESTINATION ADDRESS (32 bit)**
 - Indirizzo IP del destinatario
- **OPTIONS (variabile)**
 - campi opzionali
- **PAD (variabile)**
 - Imbottitura (padding) per allineare l'intestazione a 32 bit

Esempio di frammentazione di un pacchetto



Opzione timestamp su header IP

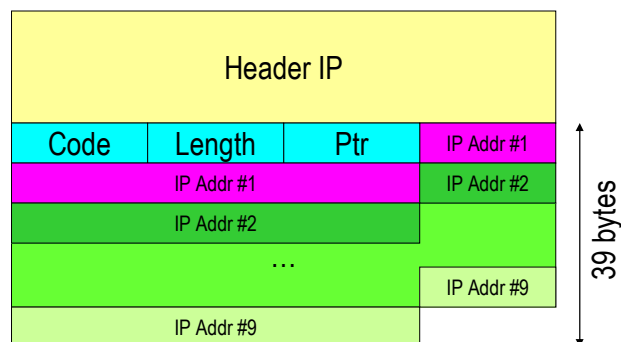
- Raramente adottata
- Internet Address
 - indirizzo IP dell'entità che ha registrato il timestamp
- Timestamp
 - espresso in ms a partire dalla mezzanotte



IP Source Routing

- Opzione prevista dal protocollo IP.
- Permette alla stazione sorgente di inserire nel pacchetto IP una specifica route per raggiungere una data destinazione.
- Esistono due modalità di source routing:
 - Strict (in cui la sorgente determina esattamente i gateway da attraversare),
 - Loose (in cui la sorgente determina una lista di gateway a cui possono aggiungersi altri a seconda dell'esigenze).

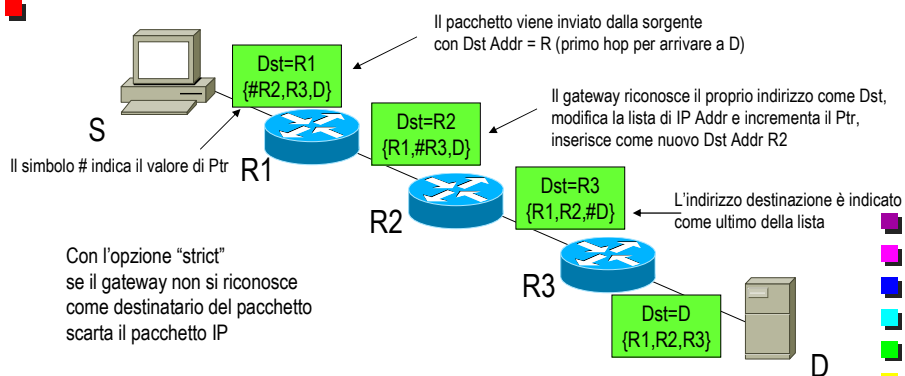
Pacchetto IP con Opzione Source Routing



Campi Opzione Source Routing

- Code = 0x83 per loose, 0x89 per strict.
- Length = numero totale di bytes.
- Ptr = puntatore che indica quale IP Address della lista utilizzare al prossimo hop.
- IP Address = indica l'indirizzo dei gateway da attraversare, l'opzione prevede la possibilità di specificare fino ad un massimo di 9 indirizzi IP.

Strict Source Routing



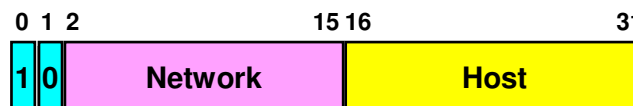
Indirizzi IP

- Sono ampi 32 bit (4 byte)
- Si scrivono come 4 numeri decimali separati dal carattere “.”
- Ogni numero rappresenta il contenuto di un byte ed è quindi compreso tra 0 e 255
- Esempi
 - 131.190.0.2
 - 1.1.2.17
 - 200.70.27.33

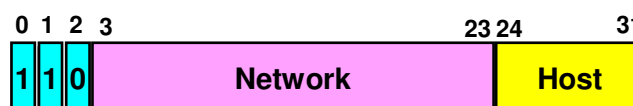
Indirizzamento



Classe A



Classe B



Classe C

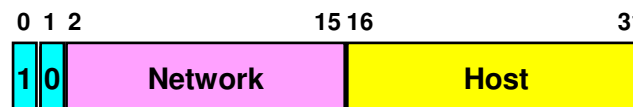
Classe A

- Campo rete
 - 7 bit
 - max 128 reti
 - valori compresi tra 0 e 127 nel primo byte
- Campo host
 - 24 bit
 - max 16M indirizzi di host



Classe B

- Campo rete
 - 14 bit
 - max 16K reti
 - valori compresi tra 128 e 191 nel primo byte
- Campo host
 - 16 bit
 - max 64K indirizzi di host



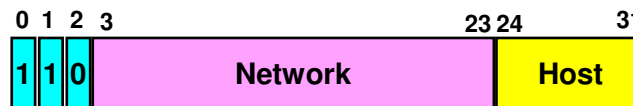
Classe C

■ Campo rete

- 21 bit
- max 2M reti
- valori compresi tra 192 e 223 nel primo byte

■ Campo host

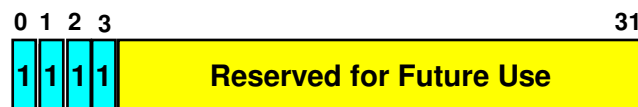
- 8 bit
- max 256 indirizzi di host di 254 utilizzabili per le macchine



Classi D ed E



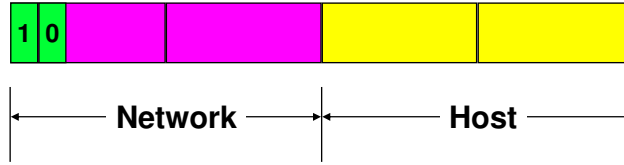
Classe D



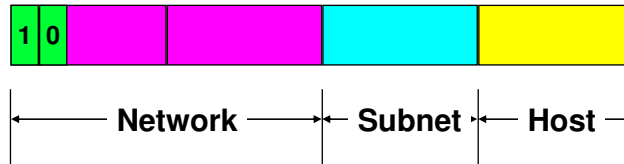
Classe E

Subnetting

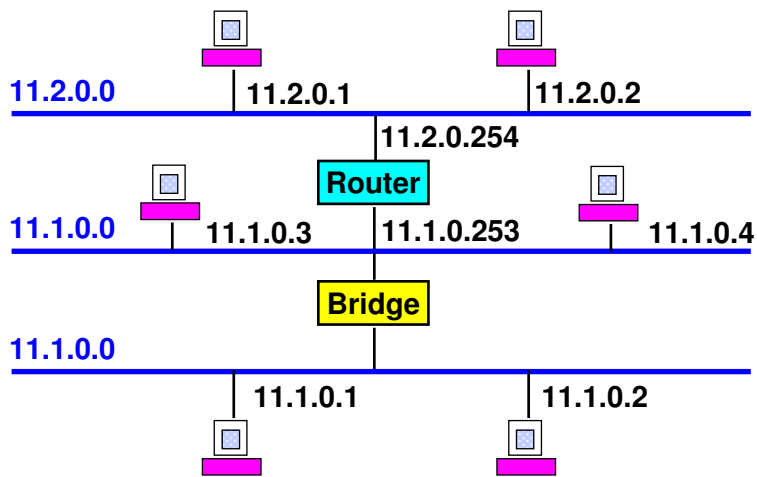
Indirizzo di classe B prima del subnetting



Indirizzo di classe B dopo il subnetting



Reti logiche e fisiche



Il campo Host

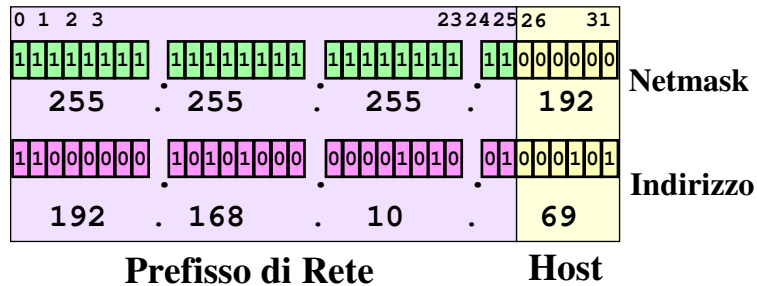
- Non si possono utilizzare tutte le combinazioni binarie del campo host per gli indirizzi delle macchine
 - la combinazione con tutti i bit di valore 1 è riservata per l'indirizzo di broadcast di livello 3
 - la combinazione con tutti i bit di valore 0 è riservata per indirizzare la rete
 - la RFC definisce le specifiche degli host e l'indirizzamento

Netmask

- Parametro che specifica il subnetting
 - bit a 1 in corrispondenza dei campi network e subnetwork
 - bit a 0 in corrispondenza del campo host
- Esempio: si supponga di voler partizionare una rete di classe C in 4 subnet da 62 host
 - Netmask 11111111 11111111 11111111 11000000
 - Netmask esadecimale ff ff ff C0
 - Netmask decimale 255.255.255.192
- Invece che indicare il valore di netmask si può indicare l'indirizzo ed il suo prefisso (lunghezza della netmask espressa in numero di bit a 1)
- La RFC 1878 definisce le modalità di subnetting

Netmask

- Si può dire che all'indirizzo 192.168.10.69 è associata una netmask 255.255.255.192 o un prefisso a 26 bit



AR (Address Range)

- Esempi:
 - 192.106.248.0 255.255.255.0 rappresenta tutti gli indirizzi IP tra 192.106.248.1 e 192.106.248.254
 - 0.0.0.0 0.0.0.0 rappresenta tutti gli indirizzi IP
 - 192.106.248.105 255.255.255.255 rappresenta un unico indirizzo IP
- L'address range associato ad una rete IP è detto "indirizzo della rete"
- L'address range senza subnetting è detto "indirizzo naturale" e la netmask "netmask naturale"

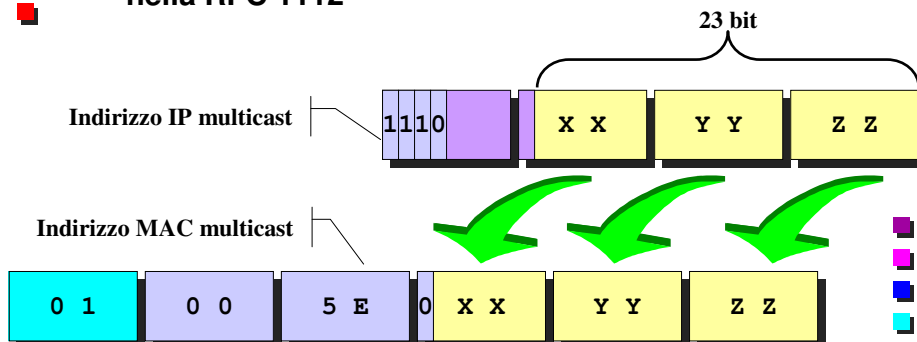
AR e netmask

I valori decimali indicati nei 4 byte che costituiscono la netmask sono quindi:

■ 128	1000 0000
■ 192	1100 0000
■ 224	1110 0000
■ 240	1111 0000
■ 248	1111 1000
■ 252	1111 1100
■ 254	1111 1110
■ 255	1111 1111

Multicast IP e multicast di livello 2

Un indirizzo Multicast IP viene rimappato su un indirizzo di livello 2 secondo le modalità definite nella RFC 1112



Esercizio 1

- Verificare cosa avviene su una rete fisica con 2 subnet diverse

Indirizzo 192.168.10.65	Indirizzo 192.168.10.126
Netmask 255.255.255.192	Netmask 255.255.255.192



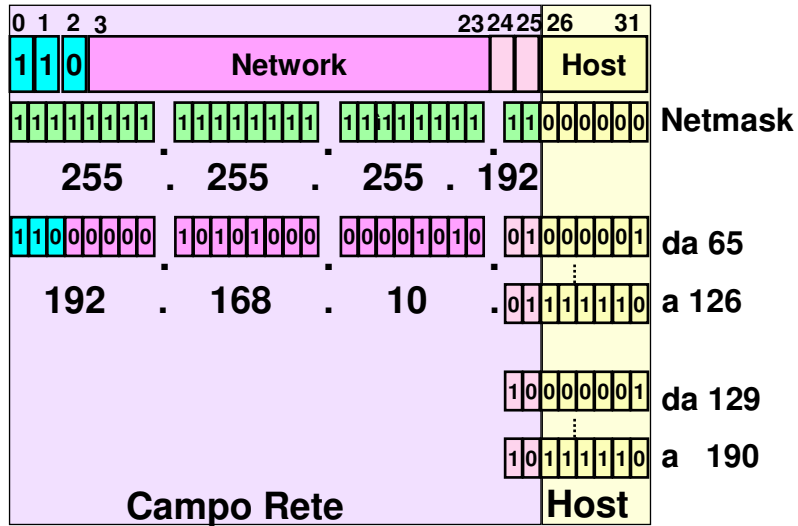
Indirizzo 192.168.10.129
Netmask 255.255.255.192

Indirizzo 192.168.10.190
Netmask 255.255.255.192

Esercizio 1

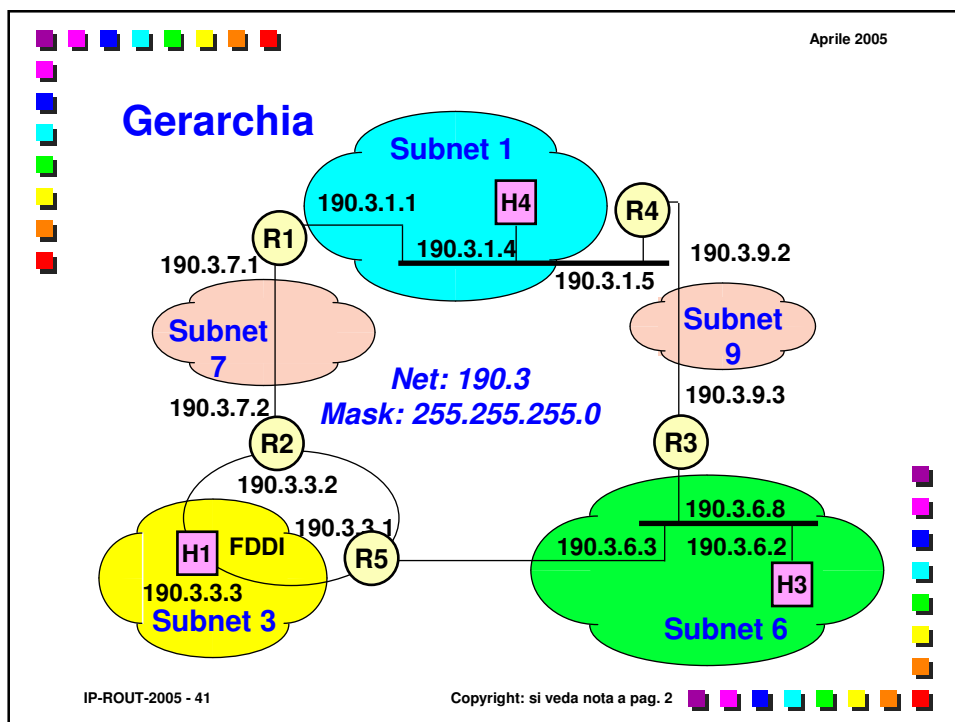
- Ci sono due subnet, con prefisso di Network a 26 bit e parte di host a 6 bit, sulla stessa rete fisica:
 - la subnet che va da 192.168.10.65 a 192.168.10.126 con Netmask 255.255.255.192
 - la subnet che va da 192.168.10.129 a 192.168.10.190 con Netmask 255.255.255.192
- La connessione si stabilisce solo tra PC1-PC2 e tra PC3-PC4

Esercizio 1



Router IP

- I router IP ricevono pacchetti IP da un'interfaccia e li inoltrano su un'altra
- Si distinguono dagli Host perchè:
 - hanno in genere più di un'interfaccia
 - utilizzano "protocolli di routing" più sofisticati
- I router IP hanno (normalmente) un indirizzo IP per ogni interfaccia



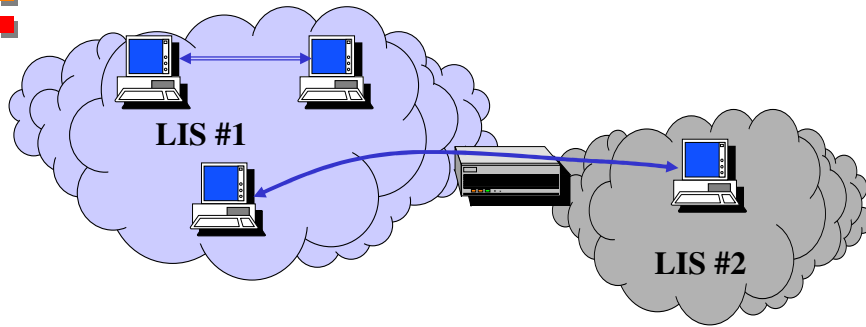
Aprile 2005

LIS (Logical IP Subnet) e reti fisiche

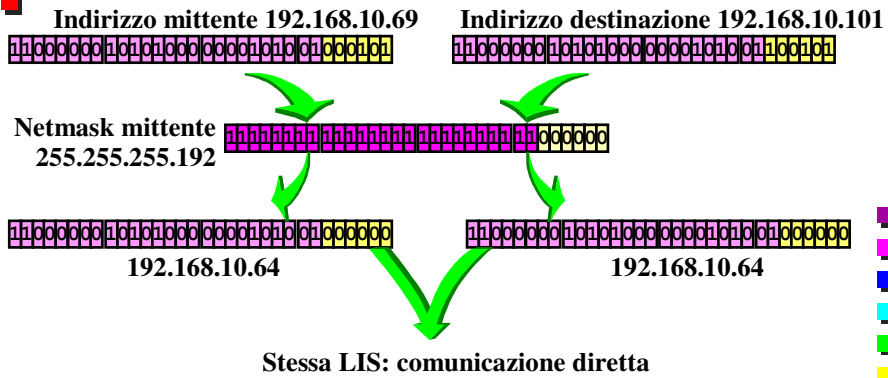
- IP assume una corrispondenza biunivoca tra reti fisiche e subnet:
 - routing implicito all'interno di una LIS
 - realizzazioni più moderne ammettono
 - più LIS sulla stessa rete fisica
 - più reti fisiche sulla stessa LIS (Proxy ARP)
- Il routing tra subnet diverse è esplicito
 - gestito dai router tramite tabelle di routing

IP-ROUT-2005 - 42 Copyright: si veda nota a pag. 2

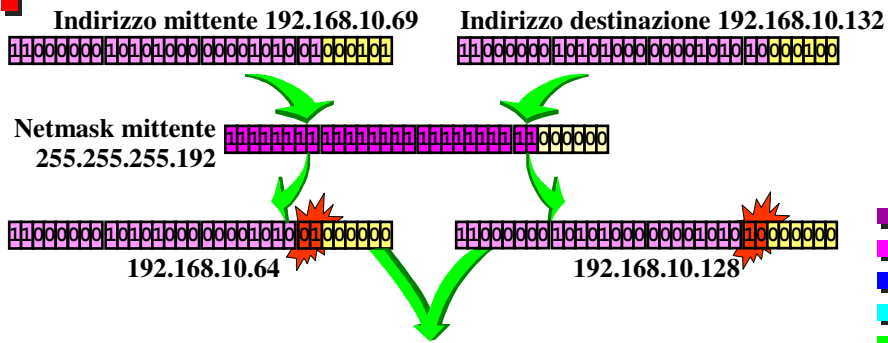
Modello di routing IP



Stessa LIS: comunicazione diretta



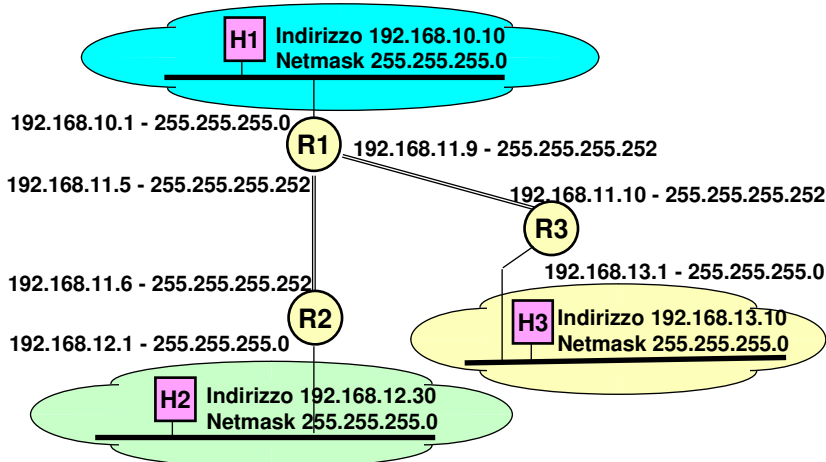
LIS differenti: utilizzo del router



LIS differenti: utilizzo del router

Perchè si usa il subnetting

- Per ottimizzare l'utilizzo di classi e indirizzi



Subnet: instradamento

- All'interno della subnet l'instradamento deve essere fornito dalla rete fisica
- Corrispondenza tra gli indirizzi delle LIS (indirizzi IP) e gli indirizzi di livello 2 gestita da ARP
- Indirizzi di livello 2
 - Indirizzi MAC sulle LAN
 - Indirizzi di DTE in X.25
 - Identificatori di LCI in Frame Relay
 -

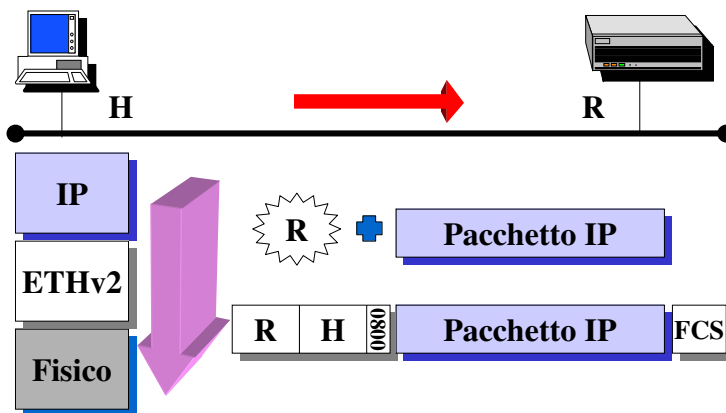
Tipi di reti IP

- punto-punto
 - le interfacce possono essere "unnumbered"
 - es.: linee dedicate o dial-up
- multiaccesso con possibilità di broadcast
 - gli host possono comunicare direttamente senza passare per router intermedi
 - possibilità di broadcast (e di ARP)
 - es. : le LAN
- multiaccesso senza possibilità di broadcast
 - gli host possono comunicare direttamente senza passare per router intermedi
 - non c'è possibilità di broadcast (niente ARP!)
 - es. : reti a pacchetto commutate

Forwarding di pacchetti IP

- Il DA è quello di un'interfaccia dell'host
 - processamento locale
- Il DA appartiene allo stesso AR di una delle interfacce
 - reti punto-punto
 - invio del pacchetto sull'interfaccia
 - reti multipunto con broadcast
 - risoluzione dell'indirizzo con ARP
 - invio diretto verso il destinatario
 - reti multipunto senza broadcast
 - l'indirizzo nativo (data-link) dell'host sulla rete destinataria deve essere configurato staticamente

Imbustamento di un pacchetto IP in una trama Ethernet v2



Forwarding di pacchetti IP

- Il DA del pacchetto non è in nessuno degli address range corrispondenti alle interfacce dell'host
 - consultazione della "routing table"
 - invio del pacchetto al "next hop" (router) indicato sulla routing table

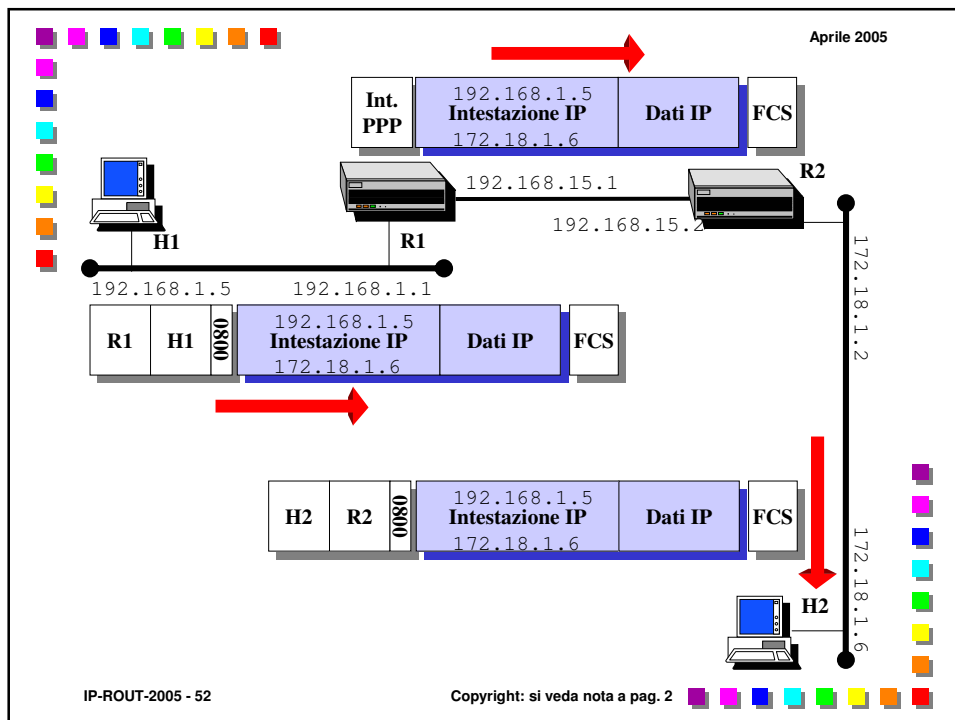


Tabella di routing

- Per ciascun Address Range danno:
 - l'indirizzo IP di un router di "next hop" su una rete IP adiacente
 - il "costo" della destinazione attraverso quel router
- Se un indirizzo IP si mappa su più AR presenti in tabella viene preferito quello con il prefisso di rete più lungo
 - *longest-prefix matching*
 - tra tutte le righe della tabella di routing, il cui prefisso corrisponde con quello dell'indirizzo di destinazione del pacchetto da inoltrare, si sceglie la riga che presenta il prefisso più lungo (maggior numero di bit a 1 nel campo netmask)

Entry sulle tabelle di routing

- Diretti
 - AR corrispondenti alle interfacce del router
- Statici
 - AR configurati staticamente dal gestore
- Dinamici
 - AR appresi attraverso un "protocollo di routing"
- Nel caso uno stesso AR sia appreso da diverse fonti deve essere specificato quale deve essere preferita (costo)

La tabella di routing

```
2503a#
2503a#sho ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF,
       IA - OSPF inter area
```

Gateway of last resort is not set

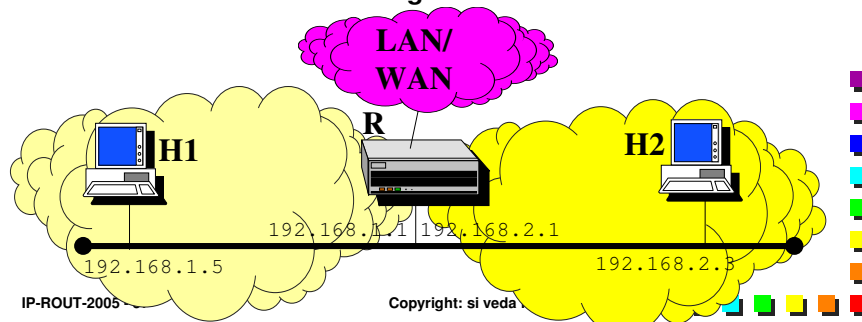
```
R   192.168.64.0/24 [120/2] via 192.168.3.9
R   192.168.65.0/24 [120/3] via 192.168.1.2
S   192.168.66.0/24 [1/0] via 192.168.1.2
R   192.168.67.0/24 [120/4] via 192.168.3.9
C   192.168.1.0/24 is directly connected, Serial0
R   192.168.2.0/24 [120/2] via 192.168.1.2
C   192.168.3.0/24 is directly connected, Serial1
S   192.168.16.0/24 [1/0] via 192.168.3.9
C   192.168.128.0/24 is directly connected, Ethernet0
2503a#
```

Address range multipli

- E' possibile assegnare più indirizzi ad un'interfaccia di un router
 - un indirizzo primario
 - uno o più indirizzi secondari
- L'indirizzo primario è utilizzato come source address di default per i messaggi generati localmente
- L'interfaccia appartiene quindi a più address range contemporaneamente
 - più LIS su una sottorete fisica

Address range multipli

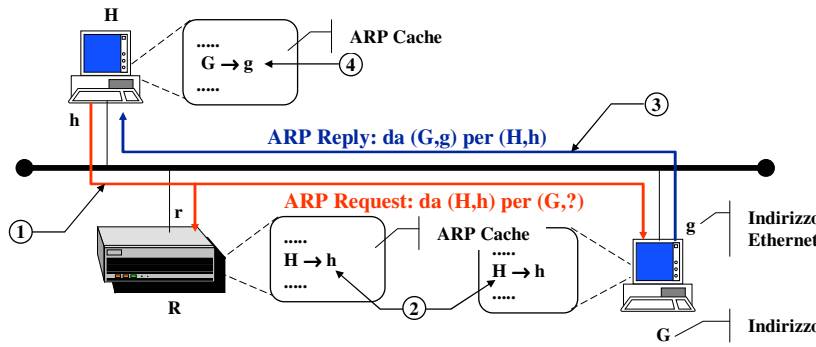
- Gli host che appartengono ad un address range possono inviare i pacchetti destinati agli altri AR:
 - direttamente a destinazione
 - se nelle loro routing table c'è la destinazione
 - al router
 - se nelle loro routing table non c'è la destinazione



Limitazioni nel subnetting

- I protocolli di routing più vecchi non trasportano le maschere
- In tali casi:
 - tutte le sottoreti di una stessa rete naturale devono avere la stessa maschera
 - tutte le sottoreti di una stessa rete naturale devono costituire un insieme connesso
 - un router non connesso ad alcuna delle sottoreti di una stessa rete naturale vede l'insieme delle sottoreti che la costituiscono come un'unica destinazione

Esempio di ARP

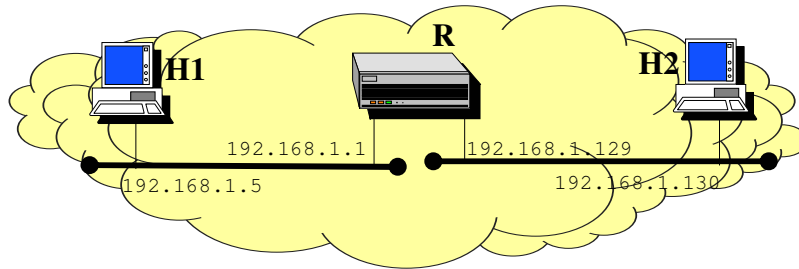


ARP

Hardware Type		Protocol Type
HLEN	PLEN	Operation
Sender Hardware Address		
Sender HW Address		Sender IP Address
Sender IP Address		Target HW Address
Target Hardware Address		
Target IP Address		

Proxy ARP

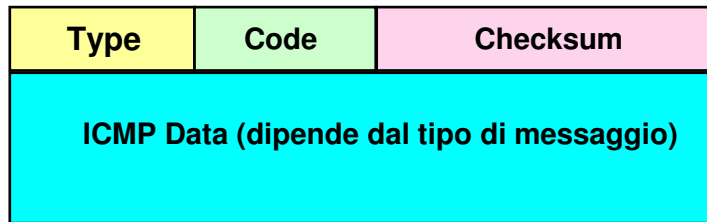
- Permette di usare la stessa subnet su due o più reti fisiche diverse



ICMP

- Internet Control Message Protocol
 - messaggio descritto nella RFC 792
- Verificare lo stato della rete (Ping)
 - Echo request ed Echo reply
- Riportare anomalie
 - Destination Unreachable
 - Time Exceeded for a Datagram
 - Parameter Problem on a Datagram
- Scoprire la netmask
 - Mask Request
 - Address Mask Reply
- Migliorare il routing
 - Redirect

Messaggio ICMP



Contenuto del messaggio ICMP

Type	Code	Descrizione	RFC
0	0	ECHO REPLY	792
3		DESTINATION UNREACHABLE	792
	0	Net Unreachable	792
	1	Host Unreachable	792
	2	Protocol Unreachable	792
	3	Port Unreachable	792
	4	Fragmentation Needed and Don't Fragment was Set	792
	5	Source Route Failed	792
	6	Destination Network Unknown	1812
	7	Destination Host Unknown	1812
	8	Source Host Isolated	
9	Communication with Destination Network is Administratively Prohibited		

Contenuto del messaggio ICMP

Type	Code	Descrizione	RFC
3	10	Communication with Destination Host is Administratively Prohibited	
	11	Destination Network Unreachable for Type of Service	1812
	12	Destination Host Unreachable for Type of Service	1812
	13	Communication Administratively Prohibited	1812
	14	Host Precedence Violation	1812
	15	Precedence cutoff in effect	1812
4	0	SOURCE QUENCH	792
5		REDIRECT	792
	0	Redirect Datagram for the Network (or subnet)	792
	1	Redirect Datagram for the host	792
	2	Redirect Datagram for the Type of Service and Network	792
	3	Redirect Datagram for the Type of Service and Host	792

Contenuto del messaggio ICMP

Type	Code	Descrizione	RFC
6	0	ALTERNATE HOST ADDRESS	
8	0	ECHO	792
9		ROUTER ADVERTISEMENT	1256
	0	Normal router advertisement	
	16	Does not route common traffic	
10	0	ROUTER SELECTION	1256
11		TIME EXCEEDED	792
	0	Time to Live exceeded in Transit	
	1	Fragment Reassembly Time Exceeded	
12		PARAMETER PROBLEM	792
	0	Pointer indicates the error	
	1	Missing a Required Option	1108
	2	Bad Length	

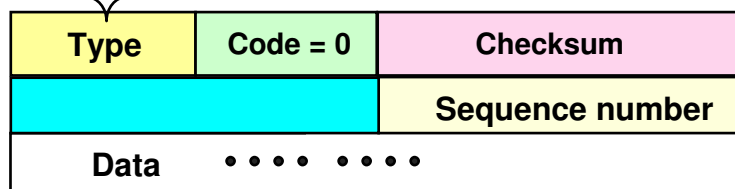
Contenuto del messaggio ICMP

Type	Code	Descrizione	RFC
13	0	TIMESTAMP	792
14	0	TIMESTAMP REPLY	792
15	0	INFORMATION REQUEST	792
16	0	INFORMATION REPLY	792
17	0	ADDRESS MASK REQUEST	950
18	0	ADDRESS MASK REPLY	950
30	0	TRACEROUTE	1393
31	0	DATAGRAM CONVERSION ERROR	1475

Programma diagnostico Ping

- Verifica la raggiungibilità di una destinazione
 - l'entità che effettua il test invia un messaggio ICMP Echo Request
 - l'entità sotto test risponde inviando un messaggio ICMP Echo Reply

0 = Echo request
 8 = Echo Reply



Il programma diagnostico Traceroute

- Traceroute (Tracert in ambito Windows) traccia la strada per raggiungere una determinata destinazione
 - individua i router che vengono attraversati tra la stazione che effettua il test e la stazione terminale di destinazione
- Per il successo del test è necessario che non ci sia Access-List sui router o firewall intermedi che inibiscono il protocollo ICMP

Traceroute: sequenza di messaggi

- Fasi della sequenza di pacchetti del programma Traceroute
 - 1) la stazione che effettua il test invia un messaggio ICMP Echo Request con TTL = 1 al default gateway
 - 2) il primo router scarta il messaggio e risponde inviando un messaggio ICMP TTL Exceeded in Transit
 - 3) la stazione che effettua il test invia apprende l'esistenza del primo router e invia un successivo messaggio ICMP Echo Request con TTL incrementato di 1 al default gateway (TTL = 2)
 - 4) il default gateway inoltra il messaggio al router successivo decrementando il TTL
 - 5) il secondo router scarta il messaggio e risponde inviando un messaggio ICMP TTL Exceeded in Transit

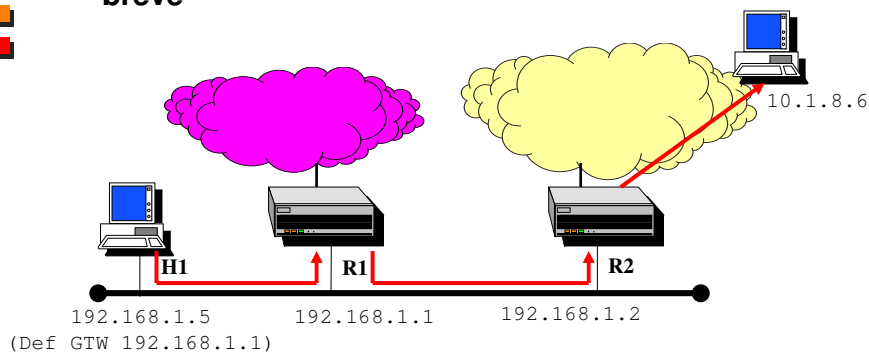
Traceroute: sequenza di messaggi

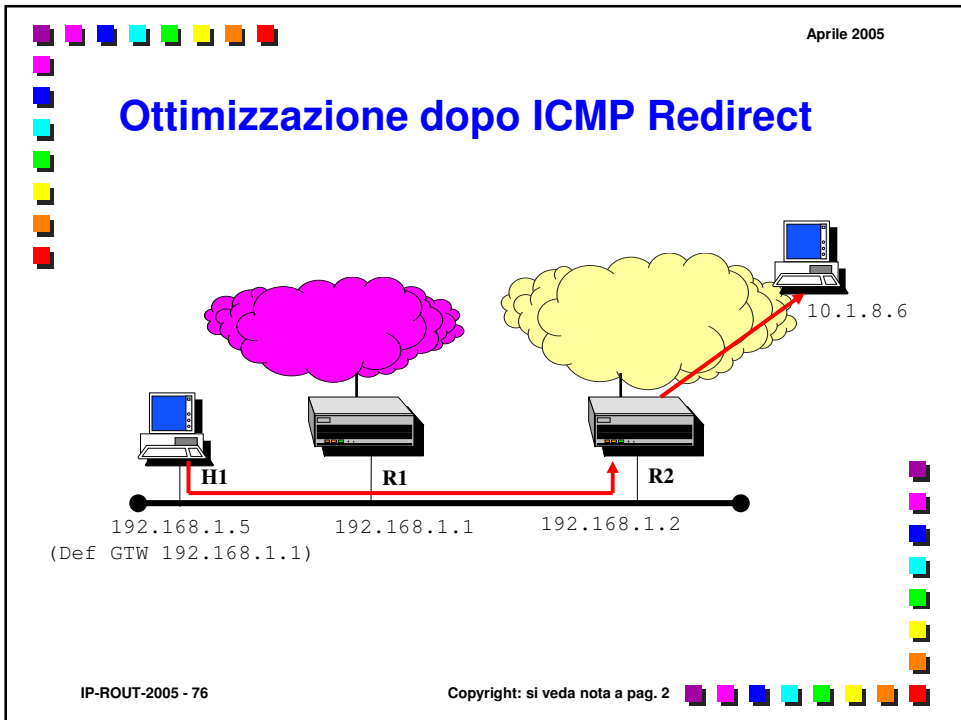
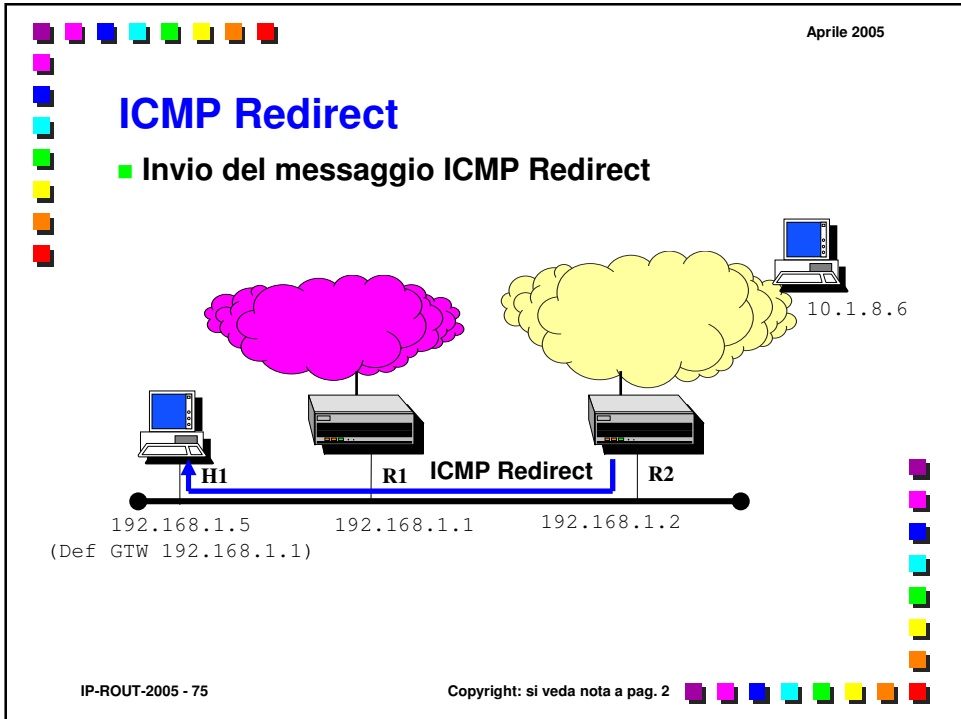
■ Fasi della sequenza di pacchetti del programma Traceroute

- 6) il primo router inoltra il messaggio **ICMP TTL Exceeded in Transit** alla stazione che ha effettuato il test
- 7) la stazione che effettua il test apprende l'esistenza del secondo router e invia un successivo messaggio **ICMP Echo Request** con TTL incrementato di 1 al default gateway (TTL = 3)
- 8) infine la stazione terminale riceve un messaggio **ICMP Echo Request** e risponde inviando un messaggio **ICMP Echo Reply** alla stazione che ha effettuato il test

ICMP Redirect

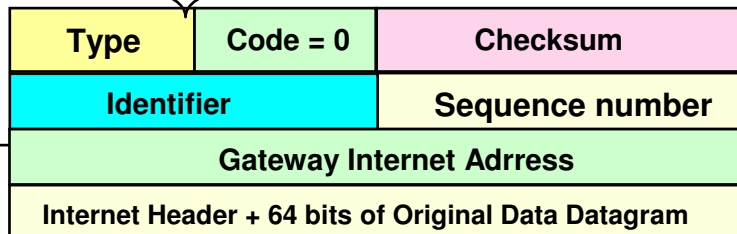
- Ottimizza il routing adottando il percorso più breve





Messaggio ICMP Redirect

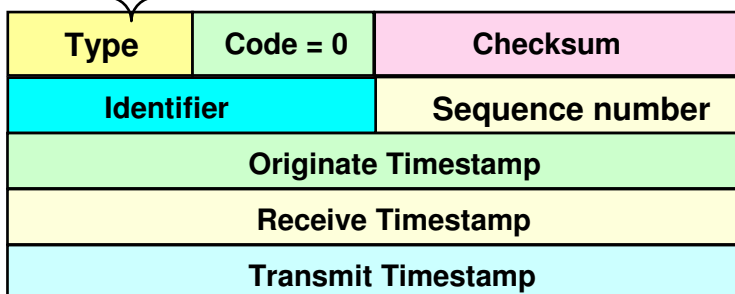
- 0 = Redirect datagrams for the Network.
- 1 = Redirect datagrams for the Host.
- 2 = Redirect datagrams for the Type of Service and Network.
- 3 = Redirect datagrams for the Type of Service and Host.



Indirizzo del router a cui deve essere indirizzato il traffico

Messaggio ICMP Timestamp

- Timestamp espresso in ms a partire dalla mezzanotte
- 13 = Timestamp
- 14 = Timestamp Reply



CIDR

- Classless Inter Domain Routing
- RFC 1517, 1518, 1519 e 1520
- Consiste nel propagare insieme all'indirizzo anche la netmask
- Serve per propagare informazioni di raggiungibilità di cluster di reti
- Esempio:
 - 199.9.4.0 con netmask 255.255.252.0
 - annuncia 199.9.4.0, 199.9.5.0, 199.9.6.0 e 199.9.7.0

Indirizzi privati

- Per Intranet si usano gli Indirizzi IP previsti in:
 - Y. Rekhter, B. Moskowitz, D. Karrenberg, G. J. de Groot & E. Lear, "RFC-1918: Address Allocation for Private Internets," February 1996.
 - nuova versione dello RFC-1597
- Classe A (una sola rete)
 - rete: 10.x.x.x
- Classe B (16 reti adiacenti)
 - reti: 172.16.x.x ... 172.31.x.x
- Classe C (256 reti adiacenti)
 - reti: 192.168.0.x ... 192.168.255.x
- ... e se si utilizzano altri indirizzi non ufficiali e non compresi in questi standard
 - fenomeno dell'accecamento