

Tecnologie di Rete

**Interconnessione
di reti**



Motivazioni

- Le LAN permettono comunicazioni veloci su distanze ridotte
- Le WAN permettono lo scambio di informazioni fra località molto distanti
- Un'azienda dovrà adottare soluzioni diversificate
 - LAN per collegare i calcolatori nello stesso edificio o nella stessa sede
 - ✓ Anche di tipo diverso
 - Servizi WAN (es. Frame Relay o SMDS) per collegare sedi in città diverse

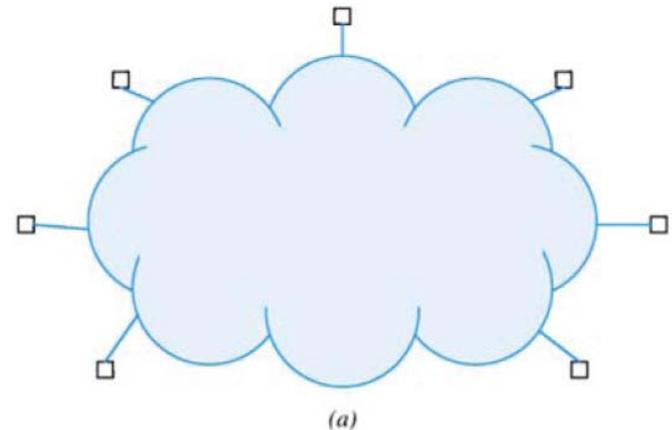
Come far comunicare fra loro calcolatori collegati a reti di tipo diverso?



Servizio di comunicazione universale

Due calcolatori devono poter comunicare indipendentemente dalla rete fisica a cui sono collegati

- Aumento di produttività
- Problemi da risolvere
 - Diversi segnali elettrici,
 - Diverso formato dei frame
 - Diverso schema di indirizzamento



Interconnessione di reti

Per offrire un servizio universale occorre

- Raccordare fisicamente le reti fisiche mediante opportuni dispositivi fisici (router)
- Aggiungere uno strato software sopra l'hardware di rete
 - ✓ Realizza l'astrazione di servizio universale sopra un insieme di reti eterogenee (rete virtuale)

Il sistema che si realizza è detto inter-rete o internet

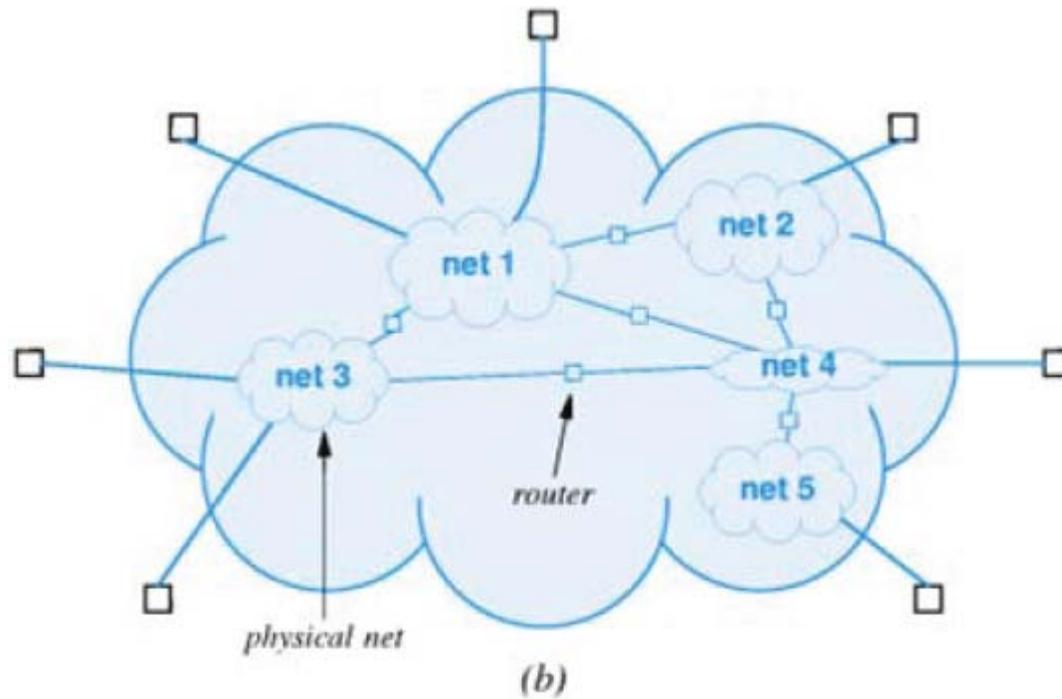
Router

- Calcolatore specializzato dedicato alla interconnessione
 - CPU, memoria, ...
 - una interfaccia per ciascuna rete a cui è collegato
- Si collega come un qualsiasi altro calcolatore
 - Collegato contemporaneamente a (almeno) due reti fisiche
- Può interconnettere reti di tipo qualsiasi
 - LAN + LAN
 - LAN + WAN
 - WAN + WAN



Inter-rete = reti di calcolatori collegate da router

Esempio di inter-rete



Protocolli software

- Simulano una rete virtuale
 - si può collegare un calcolatore come si farebbe con una rete singola
- Nascondono i dettagli delle reti fisiche sottostanti
 - L'utente si può disinteressare del tipo di reti fisiche sottostanti, della presenza o meno di router, ecc.
- Realizzano un servizio universale
 - Ogni calcolatore è individuato tramite un **indirizzo software**
 - Ogni calcolatore può scambiare messaggi con altri calcolatori collegati alla inter-rete

Protocolli TCP/IP

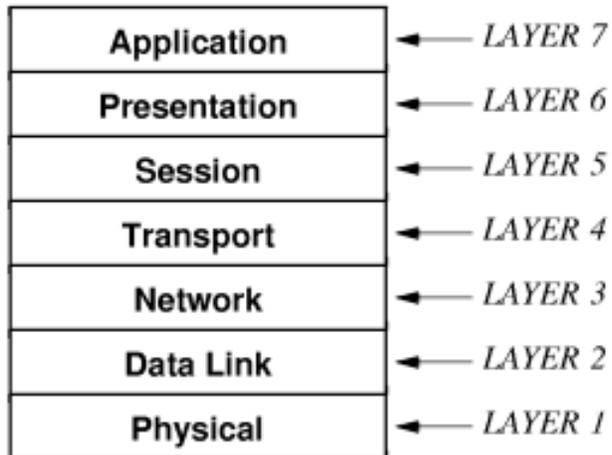
TCP = Transmission Control Protocol

IP = Internet Protocol

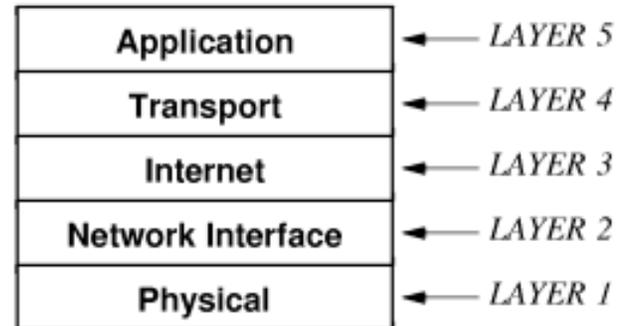
- Famiglia di protocolli usati in **Internet**
- Usati anche per la realizzazione di inter-reti private (Intranet)
- Progettati verso gli inizi degli anni '70 su iniziativa del Pentagono
 - Agenzia ARPA → Arpanet
 - Arpanet → Internet

Architettura a livelli TCP/IP

OSI



TCP/IP



Livello Physical + Network Interface

- Corrispondono complessivamente alla rete fisica cui il calcolatore è collegato
- Il livello fisico di TCP/IP corrisponde al livello fisico di OSI
- Il livello Network Interface corrisponde al livello Data Link

Internet Layer (Protocollo IP)

- Formato dei pacchetti (datagrammi)
- Formato degli indirizzi software (indirizzi IP)
- Instradamento dei datagrammi
- Servizio best-effort
 - Connectionless
 - ✓ Possibili fuori sequenza
 - Non affidabile
 - ✓ Perdite e/o alterazioni dei datagram
 - ✓ Duplicati
 - ✓ Nessuna garanzia di QoS (ritardo, jitter, throughput)

Transport Layer (Protocollo UDP)

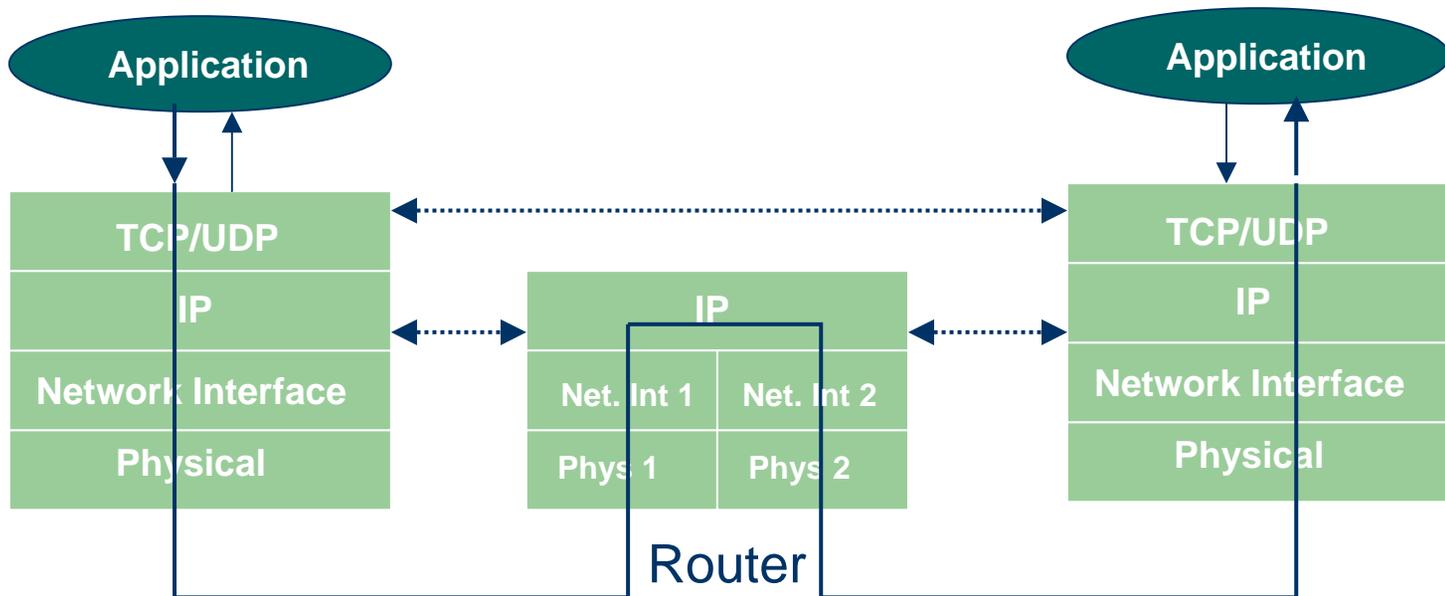
- Demultiplexing dei datagram
 - Riceve un flusso indistinto di datagrammi IP
 - Recapita i datagram ai processi applicativi a cui sono destinati
- Nessun incremento al servizio offerto da IP
 - Servizio connectionless e non affidabile
- Utilizzato per servizi real-time
 - Si richiede solo un servizio di trasporto
 - Ulteriori funzionalità specifiche vengono aggiunte sopra UDP

Transport Layer (Protocollo TCP)

- Flusso di byte
 - ma la comunicazione è sempre a pacchetti (segmenti)
- Trattamento di fuori-sequenza e duplicati
- Rilevazione dei segmenti alterati o persi
- Recupero dei segmenti alterati, persi, ritardati
- Controllo del flusso
- Controllo della congestione
- **Servizio Connection-oriented e affidabile**
 - **Tutti** i segmenti vengono consegnati in sequenza
 - Assenza di duplicati
 - Nessuna garanzia sul ritardo, sul jitter e sul throughput

Host, Router e Protocolli

Host = qualsiasi calcolatore collegato alla inter-rete



Tecnologie di Rete

Protocollo IP

Schema di Indirizzamento

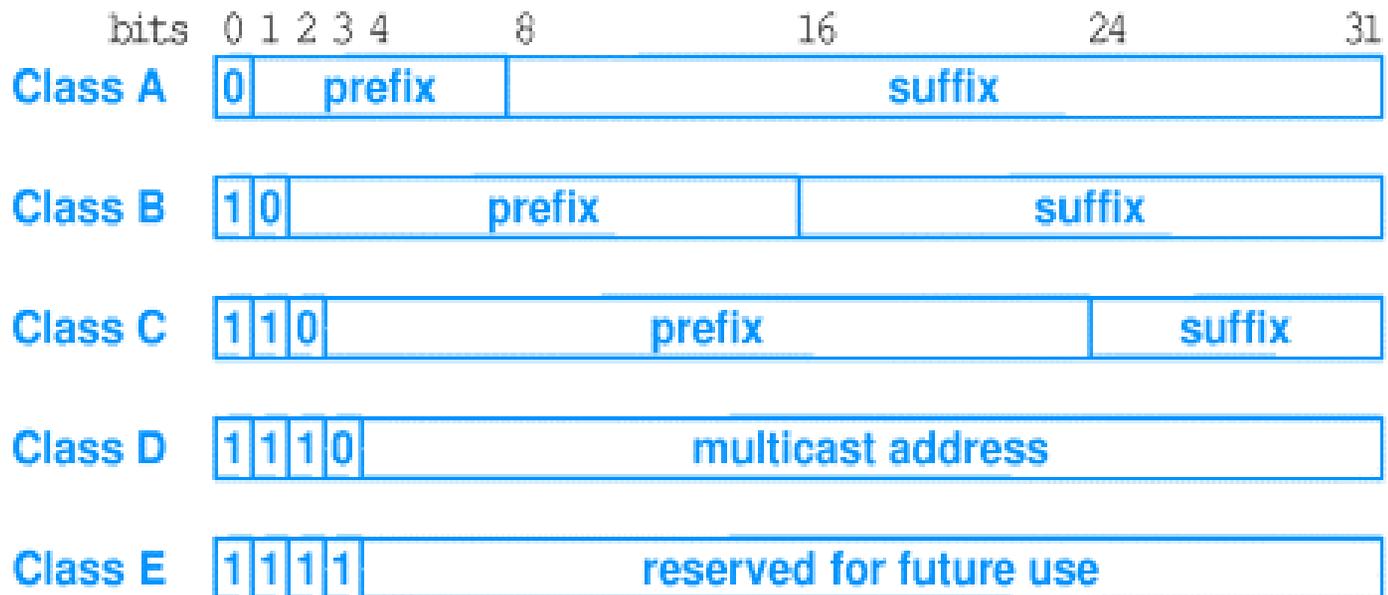
Indirizzi IP

- Indirizzo a 32 bit assegnato a ogni host
- Struttura Gerarchica
 - Indirizzo di rete (prefisso) + Indirizzo di host (suffisso)
- Indirizzo di rete (network number)
 - Identifica una rete fisica
 - Assegnato da una autorità centrale che garantisce l'univocità

Indirizzo di host (host number)

- Identifica un particolare host all'interno della rete fisica
- Assegnato localmente dall'amministratore

Classi di indirizzi IP



Notazione decimale puntata

- I 4 byte sono interpretati come numeri decimali
 - ✓ compresi fra 0 e 255
- Indirizzo letto come 4 numeri decimali separati da punti

<u>32-bit Binary Number</u>				<u>Equivalent Dotted Decimal</u>
10000001	00110100	00000110	00000000	129 . 52 . 6 . 0
11000000	00000101	00110000	00000011	192 . 5 . 48 . 3
00001010	00000010	00000000	00100101	10 . 2 . 0 . 37
10000000	00001010	00000010	00000011	128 . 10 . 2 . 3
10000000	10000000	11111111	00000000	128 . 128 . 255 . 0

Classi e notazione puntata

<u>Classe</u>	<u>Range di Valori</u>
A	da 0 a 127
B	da 128 a 191
C	da 192 a 223
D	da 224 a 239
E	da 240 a 255

Spazio di indirizzamento

Address Class	Bits In Prefix	Maximum Number of Networks	Bits In Suffix	Maximum Number Of Hosts Per Network
A	7	128	24	16777216
B	14	16384	16	65536
C	21	2097152	8	256

Come ottenere un indirizzo IP

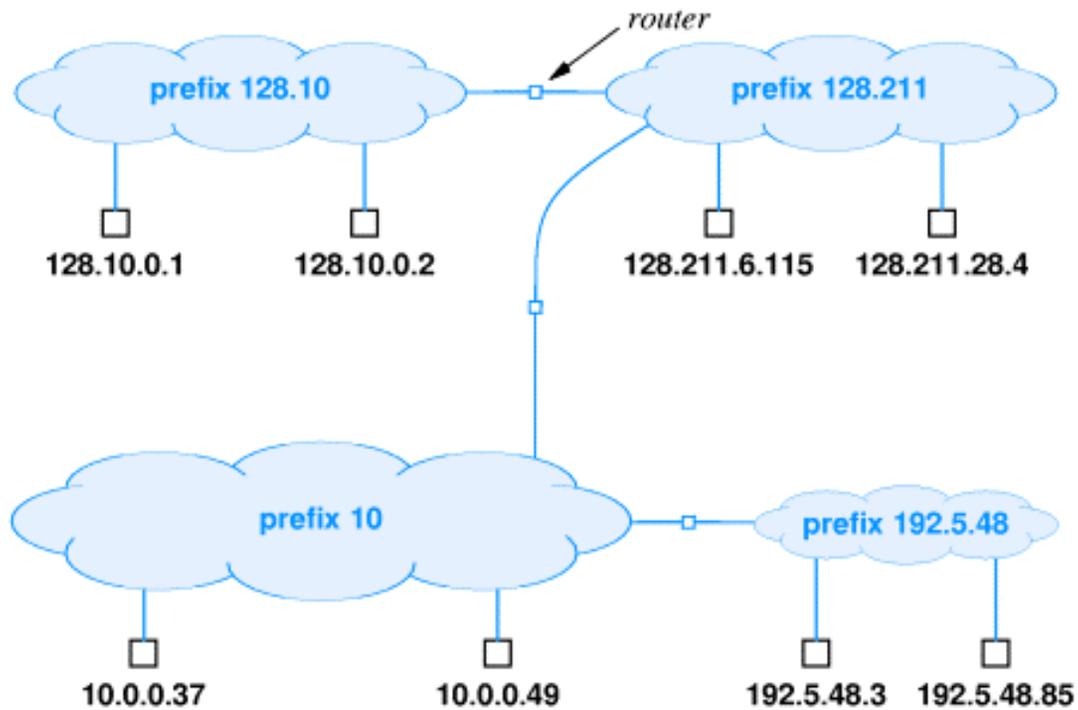
- Internet

- Indirizzo di rete fornito da Internet Service Provider (ISP)
- Indirizzi di host assegnati dall'amministratore locale
- Gli ISP sono coordinati da una autorità centrale (IANA=Internet Assigned Number Authority)

Come ottenere un indirizzo IP?

- Intranet
 - Indirizzo di rete scelto direttamente dal proprietario della inter-rete
 - Il RFC 1597 raccomanda l'uso di indirizzi di classe A, B e C nelle inter-reti private
 - Interconnessione a Internet tramite NAT

Assegnamento di indirizzi



Internet

Indirizzi IP riservati

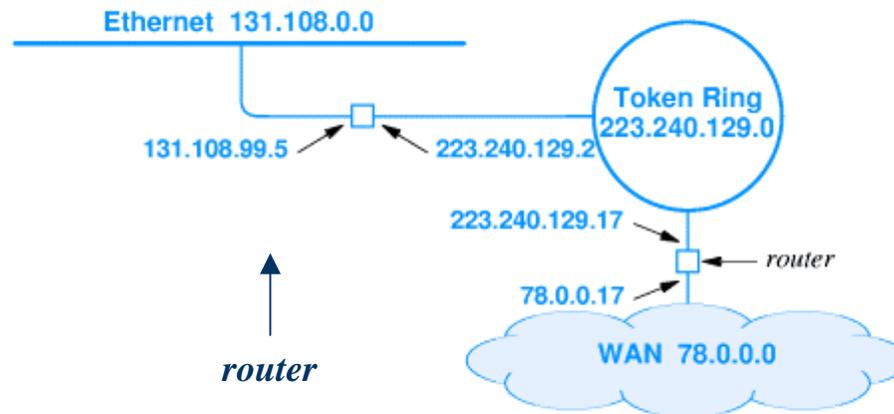
- Indirizzo per l'intera sottorete
 - Suffisso di tutti zeri (131.114.0.0)
- Indirizzo di trasmissione broadcast orientata
 - Suffisso di tutti 1 (131.114.255.255)
- Indirizzo di trasmissione broadcast ristretta
 - Costituito da tutti 1 (255.255.255.255)
- Indirizzo di “questo calcolatore”
 - Indirizzi di tutti 0
 - Usato all'avvio
- Indirizzo loopback
 - 127.0.0.1
 - Usato nello fase di sviluppo di applicazione di rete

Indirizzi IP riservati

Prefisso	Suffisso	Tipo di indirizzo	Note
tutti 0	tutti 0	“questo calcolatore”	Usato in fase di avvio
rete	tutti 0	Indirizzo di rete	Identifica una rete
rete	tutti 1	Broadcast orientato	Broadcast verso una rete specifica
tutti 1	tutti 1	Broadcast ristretto	Broadcast sulla rete locale
127	--	Loopback	Usato in fase di sviluppo

Router e indirizzi IP

- Un indirizzo IP non individua uno specifico calcolatore ma un collegamento fra calcolatore e rete fisica



- Calcolatori multi-homed
 - Hanno più di una interfaccia

Tecnologie di Rete

Datagram IP



Datagram IP

- Ogni rete fisica ha il suo formato di pacchetto (frame)
- Ogni rete fisica ha il suo schema di indirizzamento
- Il router non puo' quindi
 - instradare il frame fa reti fisiche diverse
 - ristrutturare il frame



Viene definito un formato di messaggio virtuale e universale

Formato datagramma IP (1)

HEADER

PAYLOAD

- Intestazione (Header)
 - Indirizzo IP del mittente
 - Indirizzo IP del destinatario
 - Altre informazioni di controllo
- Blocco Dati o Carico (Payload)
 - Lunghezza variabile (stabilito dall'applicazione)

Formato datagramma IP

0	4	8	16	19	24	31
VERS	H. LEN	SERVICE TYPE	TOTAL LENGTH			
IDENTIFICATION			FLAGS	FRAGMENT OFFSET		
TIME TO LIVE		TYPE	HEADER CHECKSUM			
SOURCE IP ADDRESS						
DESTINATION IP ADDRESS						
IP OPTIONS (MAY BE OMITTED)					PADDING	
BEGINNING OF DATA ⋮						

Tecnologie di Rete

Instradamento dei
datagrammi IP



Procedura di instradamento

- Presso host mittente
- Presso un router intermedio

Instradamento presso host mittente

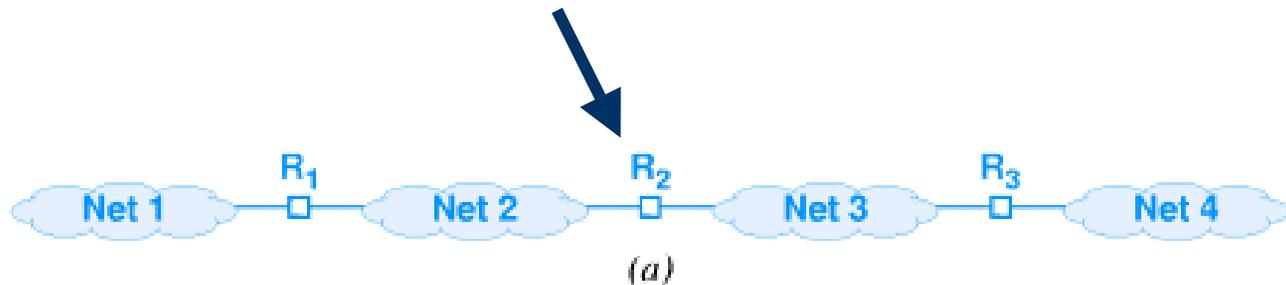
- È noto l'indirizzo IP di un router (gateway) di default
- Confronta la parte di rete dell'indirizzo IP destinatario con l'indirizzo di rete dell'host
 - Se coincide trasmette il datagramma al destinatario finale (tramite la rete fisica comune)
 - Altrimenti invia il datagramma al router di default

```
NetworkNum=MyNetworkMask & Dest_IP_Addr  
If((NetworkNum of Dest_IP_Addr)==MyNetworkNum)  
then deliver datagram to Dest_IP_Addr directly  
else forward datagram to default router
```

Instradamento presso un router

- Il router esamina l'indirizzo IP destinatario
- Inoltra il datagram
 - al destinatario finale
 - al prossimo router
- Tabella di forwarding
 - Contiene coppie (Destination, Next Hop)
 - Costruita all'avvio e modificata dinamicamente

Tabella di instradamento



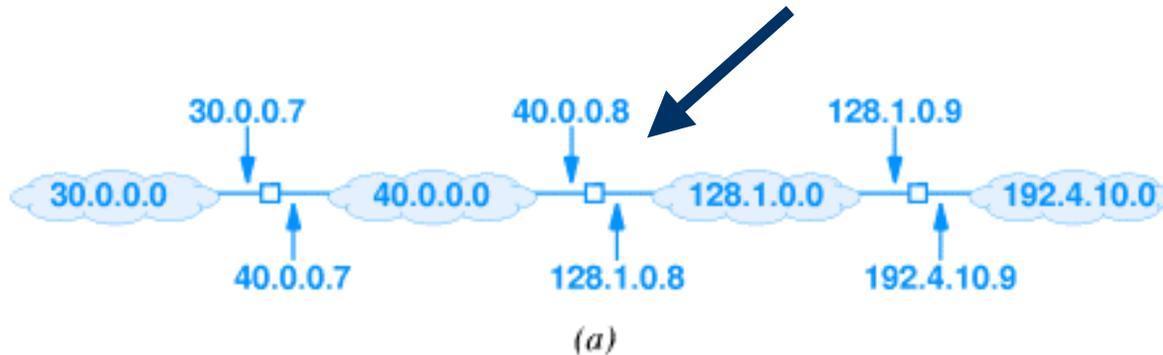
Indirizzi di rete e non di host →

Destination	Next Hop
net 1	R ₁
net 2	deliver direct
net 3	deliver direct
net 4	R ₃

(b)

Tabella di instradamento del router R2

Tabella di instradamento (2)



Destination	Mask	Next Hop
30.0.0.0	255.0.0.0	40.0.0.7
40.0.0.0	255.0.0.0	deliver direct
128.1.0.0	255.255.0.0	deliver direct
192.4.10.0	255.255.255.0	128.1.0.9

(b)

For each entry i

If((Mask[i] & Dest_IP_Addr)==Destination[i]
then forward to NextHop[i]

Costruzione tabella di forwarding

- Routing statico

- Tabella viene generata in fase di avvio
 - ✓ Manuale
 - ✓ Mediante lettura da file di configurazione
- Gestione semplice
- Nessun overhead di comunicazione
- Scarsa flessibilità
 - ✓ Non si adatta a cambiamenti di configurazione o guasti

Costruzione tabella di forwarding

- Routing dinamico
 - Tabella modificata dinamicamente
 - Adattabilità a guasti dei router e interruzioni di link
 - Adattabilità a condizioni del traffico
 - Gestione complessa
 - ✓ Overhead di comunicazione
 - ✓ Ritardi
 - ✓ Protocollo di routing

Protocollo di Routing

- Protocollo per la costruzione della tabella di forwarding
- Routing
 - Processo per la costruzione della tabella di forwarding
- Forwarding
 - Processo di instradamento dei datagram

Protocolli di routing

- Algoritmo di Dijkstra
 - Rete rappresentata come un un grafo
 - A ogni arco (link) è associato un peso
 - ✓ distanza, banda, ecc.
 - Calcolo della rotta più breve fra un nodo (router) e un qualsiasi altro destinatario
 - Costruzione della tabella di forwarding
 - ✓ Come insieme delle rotte calcolate mediante l'algoritmo di Dijkstra

Protocolli di routing

- Calcolo distribuito delle rotte
 - I router si scambiano periodicamente informazioni
 - Ogni router costruisce localmente la sua tabella...
 - ... e informa tutti gli altri dei risultati
 - Permette l'adattamento a situazioni di guasto

Protocolli di routing

- Distance Vector
 - Basato sullo scambio periodico di informazioni
 - I router scambiano vettori di coppie del tipo
(destination, distance)
- Link State
 - Basato sull'algoritmo di Dijkstra
 - I nodi scambiano informazioni sullo stato dei link
 - ✓ Link State Advertisements (LSA)
 - Costruiscono il grafo della rete
 - Derivano la tabella di forwarding

Protocollo RIP (Routing Information Protocol)

- Basato su Distance Vector
- La tabella di forwarding contiene la distanza
- I router scambiano con i vicini coppie del tipo (dest, distance) ogni 30 sec.
- Un router modifica la propria tabella di routing se riceve una rotta con distanza inferiore

Destin.	Next Hop	Distance
net1	R1	1
net2	deliver direct	0
net3	deliver direct	0
net4	R3	1

Internet

Protocollo RIP: esempio

Protocollo OSPF (Open Shortest Path First)

- Basato sull'algoritmo di Dijkstra
- Scambio di Link State Advertisements (LSA)
- LSU trasmessi in broadcast a tutti i router
- Grafo della rete sulla base dei LSA
- Algoritmo di Dijkstra per la determinazione dei percorsi minimi
- Costruzione della tabella di forwarding

Limiti dei protocolli precedenti

- Scalabilità
 - Ogni router deve mantenere informazioni relative a ciascuna rete destinataria
 - ✓ Memoria
 - ✓ Tempo di elaborazione
 - Non applicabili al caso di inter-reti grandi (Internet)
- Soluzioni
 - Routing Gerarchico
 - Subnetting

Dominio di Routing

- Dominio di routing (o Autonomous System, AS)
 - inter-rete in cui router sono tutti sotto lo stesso controllo
 - ✓ campus universitario, azienda
 - Stesso protocollo di routing
 - Router di frontiera per connettersi ad altri domini
- I router esterni instradano verso l'AS
 - Una sola entry nella tabella di forwarding (router di frontiera)
- All'interno del AS si instrada verso l'host dest
 - Solo i router del AS hanno consapevolezza delle varie reti fisiche

Routing Gerarchico

- Routing intra-dominio (intradomain routing)
 - RIP, OSPF, ...
- Routing inter-dominio (interdomain routing)
 - BGP (Border Gateway Protocol)
 - EGP (Exterior Gateway Protocol)

Reti fisiche e indirizzi IP

- Corrispondenza fra rete fisica e indirizzo IP
 - Ogni rete fisica richiede almeno un indirizzo di classe C
 - Ogni rete fisica con più di 255 host richiede un indirizzo di classe B
- Conseguenze
 - Uso inefficiente degli indirizzi
 - Elevata dimensione della tabella di forwarding
 - Processo di instradamento molto lento
- Soluzione
 - Sottoreti (Subnetting)
 - Classless InterDomain Routing (CIDR) o Supernetting

Sottoreti IP

- Uno stesso indirizzo di rete allocato a più reti fisiche
 - Ogni rete fisica costituisce una sottorete
 - Host della stessa rete fisica hanno stesso indirizzo di sottorete
 - Indirizzi di host assegnati in modo contiguo alle sottoreti
- Sottoreti viste dai router lontani come un'unica rete
 - Una sola entrata nella tabella di forwarding
 - Le varie sottoreti devono essere fisicamente vicine
- Vantaggi
 - Migliori prestazioni a livello di routing
 - Migliore utilizzazione dello spazio degli indirizzi

Indirizzo di sottorete

Indirizzo di rete

Indirizzo di host

Indirizzo di classe B

Indirizzo di rete

Ind. di sottorete

Id. di host

Indirizzo di classe B con sottorete

1 0 0 0 0 0 0 0 0

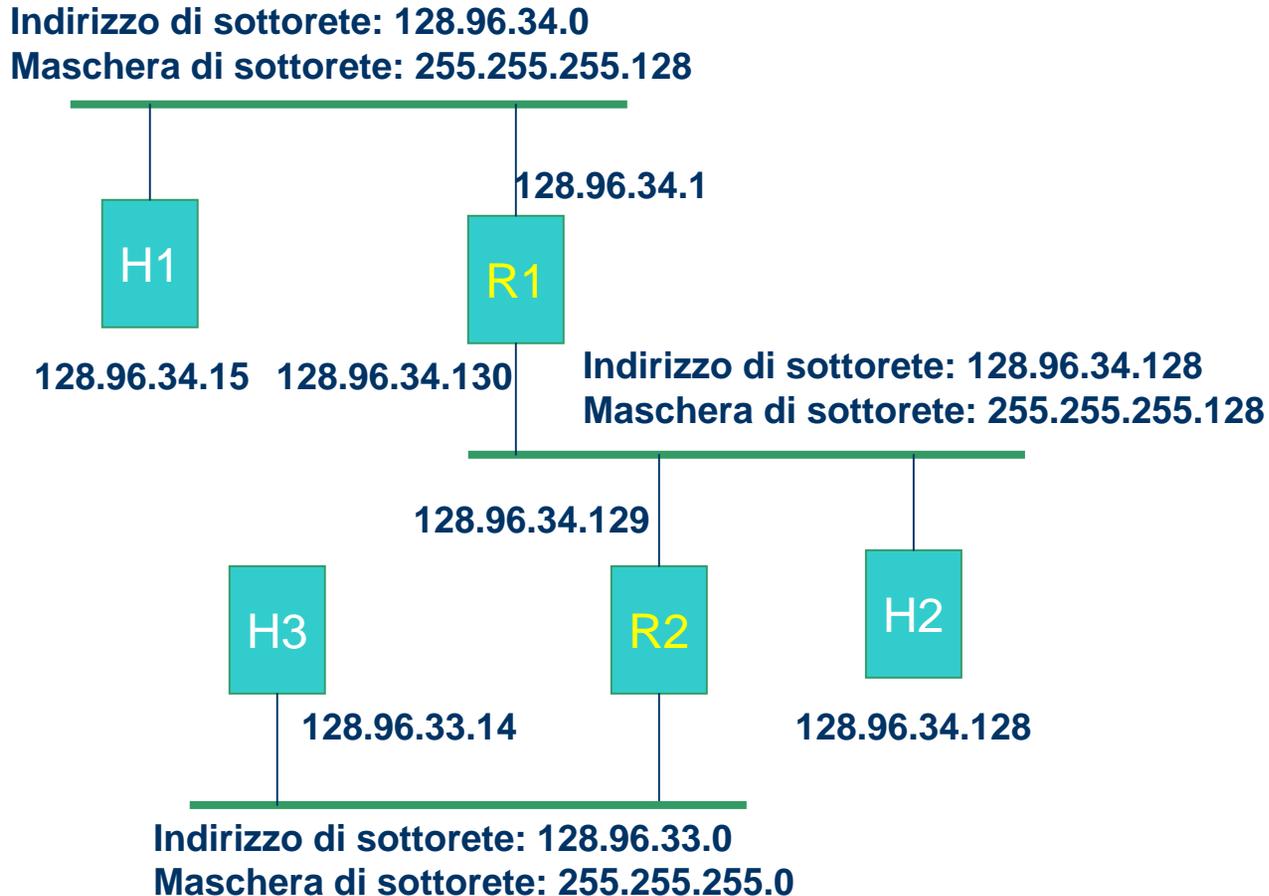
Maschera di sottorete: 255.255.255.0

- Host appartenenti alla stessa sottorete devono avere lo stesso indirizzo di sottorete (ind. di rete + id. di sottorete)

Maschera di sottorete

- Serve a separare indirizzo di sottorete da indirizzo di host
- Non necessariamente allineata al byte:
 - Esempio: 255.255.255.128
 - Si usano 16+9 bit per identificare la sottorete e 7 bit per l'host
- Impostata all'atto della configurazione della rete
 - Indirizzo IP dell'host
 - Maschera di sottorete
 - Router di default (gateway)

Esempio di utilizzo di sottoreti



Instradamento a livello di host

Procedura presso host mittente

```
SubnetNum=MySubnetMask & Dest_IP_Addr  
If((SubnetNum of Dest_IP_Addr)==MySubnetNum)  
then deliver datagram to Dest_IP_Addr directly  
else forward datagram to default router
```

Instradamento a livello di host

Esempio: H1 vuole inviare a H2

- Indirizzo IP Destinatario (H2) 128.96.34.139
- Maschera di sottorete del mittente (H1) 255.255.255.128
- And bit a bit 128.96.34.128
- Indirizzo di sottorete del mittente (H1) 128.96.34.0

Il datagramma viene inoltrato a R1

Tabella di instradamento

<u>SubnetNumber</u>	<u>SubnetMask</u>	<u>NextHop</u>
128.96.34.0	255.255.255.128	interface 0
128.96.34.128	255.255.255.128	interface 1
128.96.33.0	255.255.255.0	R2

Tabella di instradamento presso il router R1 (esempio precedente)

- Indirizzi di sottorete anziché indirizzi di rete
- Maschere di sottorete anziché maschere di rete

Instradamento a livello di router

```
D=Dest_IP_Addr
For each [SubnetNum, SubnetMask, NextHop] in Table
  D1=SubnetMask & D
  If(D1==MyNetworkNum)
  then deliver datagram to Dest_IP_Addr directly
  else forward datagram to default router
```

Router di default per diminuire il numero di entrate

Limiti del Subnetting

- È legato al concetto di classe
 - Il subnetting opera a partire da un indirizzo di una certa classe
- L'uso degli indirizzi rimane non ottimale
 - ✓ Un AS con 256 host richiede indirizzo di classe B (0,39%)
- Soluzione
 - Superare completamente il concetto di classe di indirizzi
 - Classless Interdomain Routing (CIDR)

CIDR (Supernetting)

- Obiettivi

- Utilizzare efficientemente l'uso degli indirizzi
- Non aumentare le entry nelle tabelle di forwarding
 - ✓ Assegnare un blocco di indirizzi di classe C non contigui comporterebbe un numero elevato di entry

- Idea di base

- Assegnare a un AS un blocco contiguo di indirizzi
 - ✓ La parte comune rappresenta l'indirizzo di rete (length)
 - ✓ La parte variabile rappresenta l'indirizzo di host (value)
 - ✓ Indirizzo IP = <length, value>

CIDR (Supernetting)

- L'aggregazioni dei percorsi è generalizzabile
 - Un AS espone un unico indirizzo di rete (length)
 - Un provider espone un unico indirizzo che include tutti gli AS gestiti
- Richiede variazioni nell'instradamento interdominio
 - ✓ Non si specifica più la maschera ma la lunghezza in bit del campo length (es. 231.1.2.0/24)
 - ✓ BGP-4 è in grado di CIDR

Assegnazione dinamica (DHCP)

- Limiti dell'assegnazione statica
 - L'indirizzo viene assegnato in modo permanente
 - Host con collegamenti sporadici occupano un indirizzo che usano raramente
- Assegnazione dinamica
 - Assegnazione all'atto del collegamento
 - ✓ Condivisione statica degli indirizzi
 - Richiede DHCP Server
 - ✓ Dynamic Host Configuration Protocol

NAT (Network Address Translation)

- Permette di usare un unico indirizzo IP per tutti gli host di piccola rete
 - ✓ Reti SOHO (Small Office, Home Office)
 - ✓ Laboratori
 - ✓ ...
- NAT server/router
 - Traduce indirizzi privati nell'unico indirizzo pubblico e viceversa
 - Usa tabella di traduzione NAT per decidere il vero destinatario

Obiezioni all'uso del NAT

- È una soluzione di medio termine
 - La soluzione a lungo termine è IPv6
- Viola il principio di identificazione globale
 - Ogni host dovrebbe avere un indirizzo globale in Internet
- Viola la semantica end-to-end
- Viola la stratificazione dei protocolli
 - Il router manipola informazioni di livello superiore al 3 (numeri di porta)

Tecnologie di Rete

Risoluzione degli
indirizzi IP - ARP



Introduzione

- Determinato il prossimo salto il router (o l'host mittente) trasmette il datagram
 - destinatario finale
 - altro router



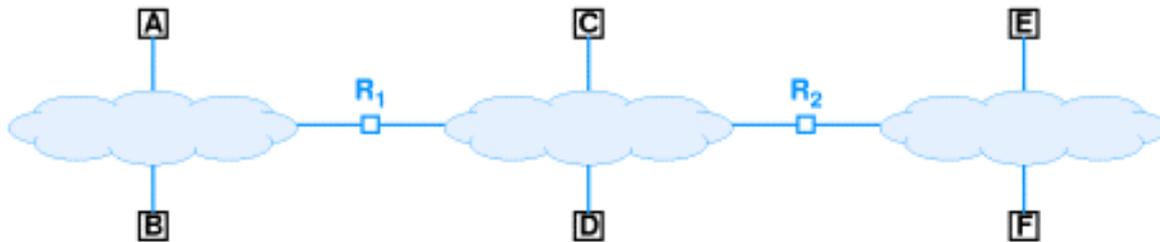
È necessario conoscere l'indirizzo fisico dell'host o del router destinatario

Come si fa a conoscere tale indirizzo?



Risoluzione degli indirizzi

- Traduzione dell'indirizzo software nel corrispondente indirizzo fisico



- La traduzione è necessariamente locale alla rete fisica
 - Ogni nodo risolve solo l'indirizzo relativo al salto successivo

Tecniche di risoluzione

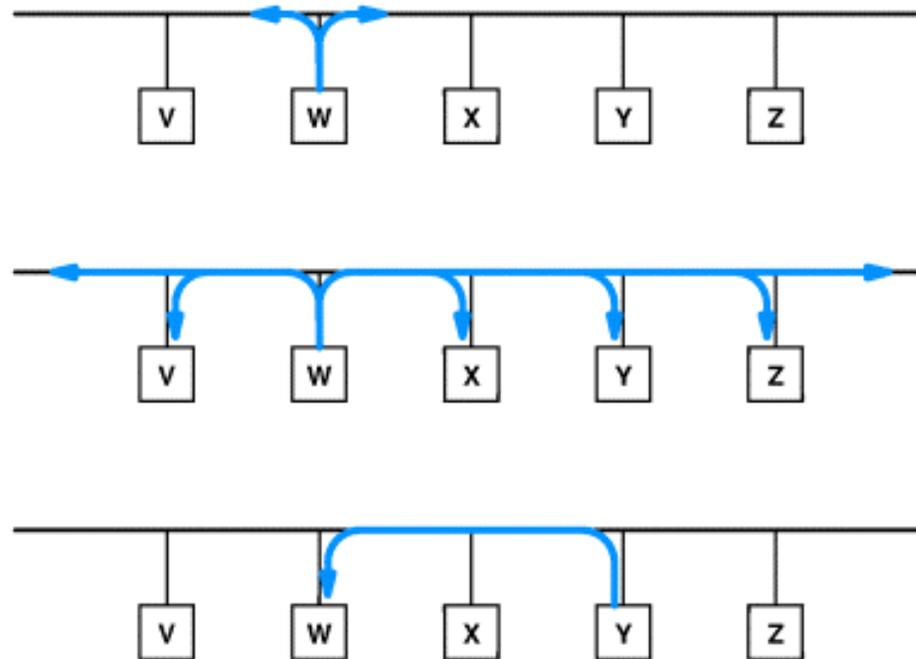
- Ricerca tabellare
 - Tabella in cui sono memorizzate tutte le corrispondenze
- Computazione in forma chiusa
 - Dall'indirizzo fisico si ricava direttamente l'indirizzo software
 - Indirizzi ISP (Novell)
- Risoluzione distribuita
 - Si manda una richiesta broadcast con l'indirizzo sw da risolvere
 - L'host con quell'indirizzo sw invia il proprio indirizzo hw

Risoluzione in Internet

- Protocollo ARP (Address Resolution Protocol)
 - Risoluzione distribuita
- Cache ARP
 - Ricerca Tabellare

Protocollo ARP

W vuole inviare un messaggio a Y ma non conosce il suo indirizzo fisico



Protocollo ARP

- W invia un **ARP request** con indirizzo IP di Y
 - Trasmissione broadcast
- Tutti gli host ricevono la richiesta
- Y risponde con un **ARP response** contenente il proprio indirizzo fisico
 - Gli altri host non considerano la richiesta

Messaggi ARP

0	8	16	24	31
HARDWARE ADDRESS TYPE		PROTOCOL ADDRESS TYPE		
HADDR LEN	PADDR LEN	OPERATION		
SENDER HADDR (first 4 octets)				
SENDER HADDR (last 2 octets)		SENDER PADDR (first 2 octets)		
SENDER PADDR (last 2 octets)		TARGET HADDR (first 2 octets)		
TARGET HADDR (last 4 octets)				
TARGET PADDR (all 4 octets)				

HA type

vale 1 se Ethernet

PA type

vale 0x0800 se IP

HADDR len

lungh. Indirizzo fisico

PADDR len

lunghezza indirizzo software

Operation

richiesta (1) – risposta (2)

Sender HADDR/PADDR

indirizzo fisico/software del mittente

Target HADDR/PADDR

indirizzo fisico/software dell'obiettivo

Trasmissione di messaggi ARP

Messaggio ARP trasmesso tramite un frame Ethernet



Dest. Address	Source Address	Frame Type	Data In Frame
		806	complete ARP message

Limiti di ARP

- Richiede scambio ARP request/response
 - Per ogni datagram da consegnare
 - Ad ogni passo del path
- Incrementa il traffico sulla rete
- Incrementa la latenza del datagram

- Soluzione
 - Cache ARP

Cache ARP

- Ogni nodo mantiene una cache ARP
 - contiene risoluzioni più recenti
- Richiesta ARP solo in caso di miss
 - Aumento del throughput
 - Diminuzione della latenza
- Problema: Informazioni scadute
 - Calcolatori che hanno cambiato indirizzo hw
 - Indirizzi assegnati con DHCP
- Soluzione: Tempo di vita
 - Dopo un certo intervallo le entrate vengono eliminate

Tecnologie di Rete

**Frammentazione e
riassembly dei
datagram IP**

Trasmissione di un datagram IP

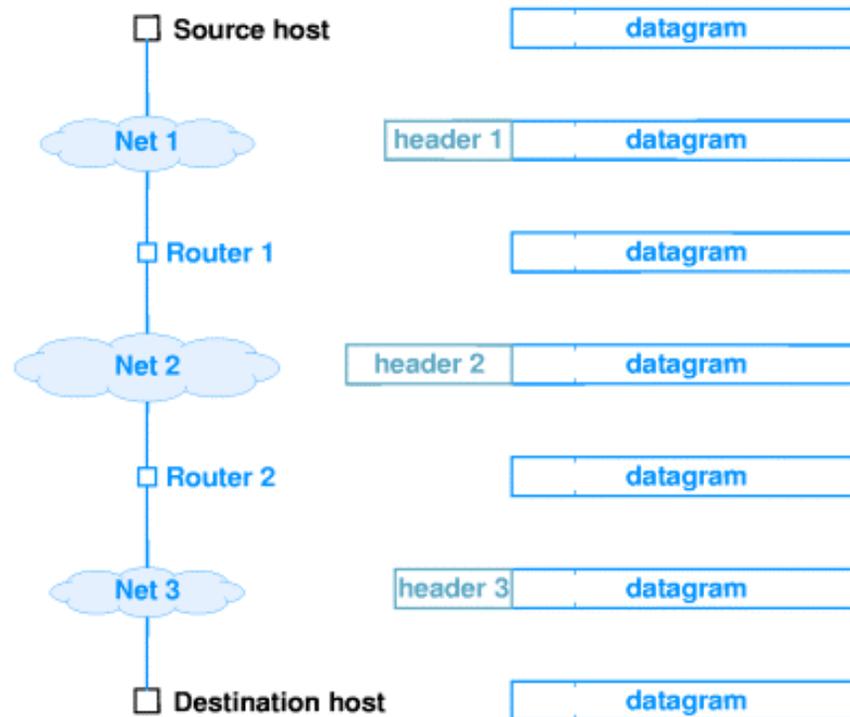
- Incapsulamento

- Il datagram IP viene inserito nel campo dati del frame
- Il campo tipo del frame posto a 0800 (datagram IP)
 - ✓ In ricezione il campo dati verrà consegnato al protocollo IP



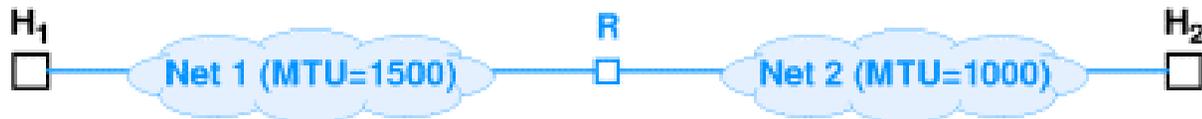
Viaggio di un datagram IP

Prima di attraversare una rete fisica il datagramma viene incapsulato in un frame specifico di quella rete



Massimo carico

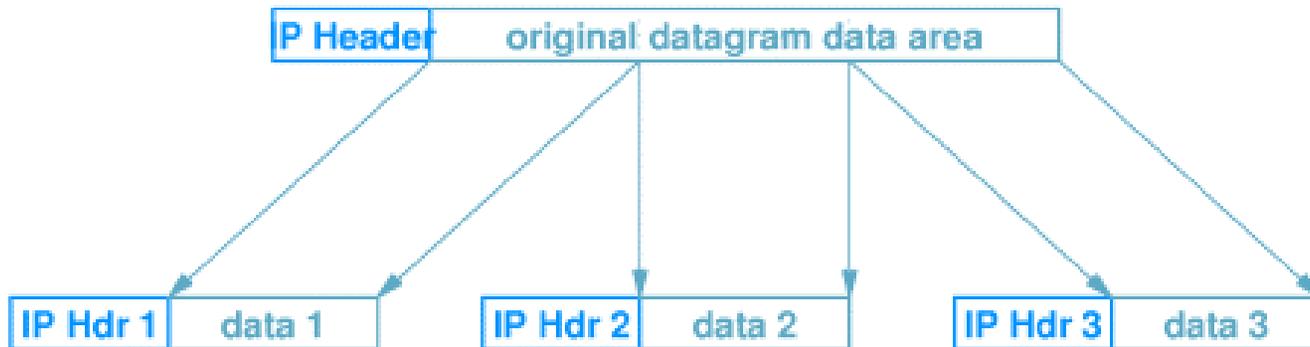
- L'incapsulamento ad ogni passo risolve il problema del diverso formato dei frame
- Ogni rete fisica definisce il proprio MTU
 - Maximum Transmission Unit
- Come conciliare i diversi valori di MTU?



Frammentazione

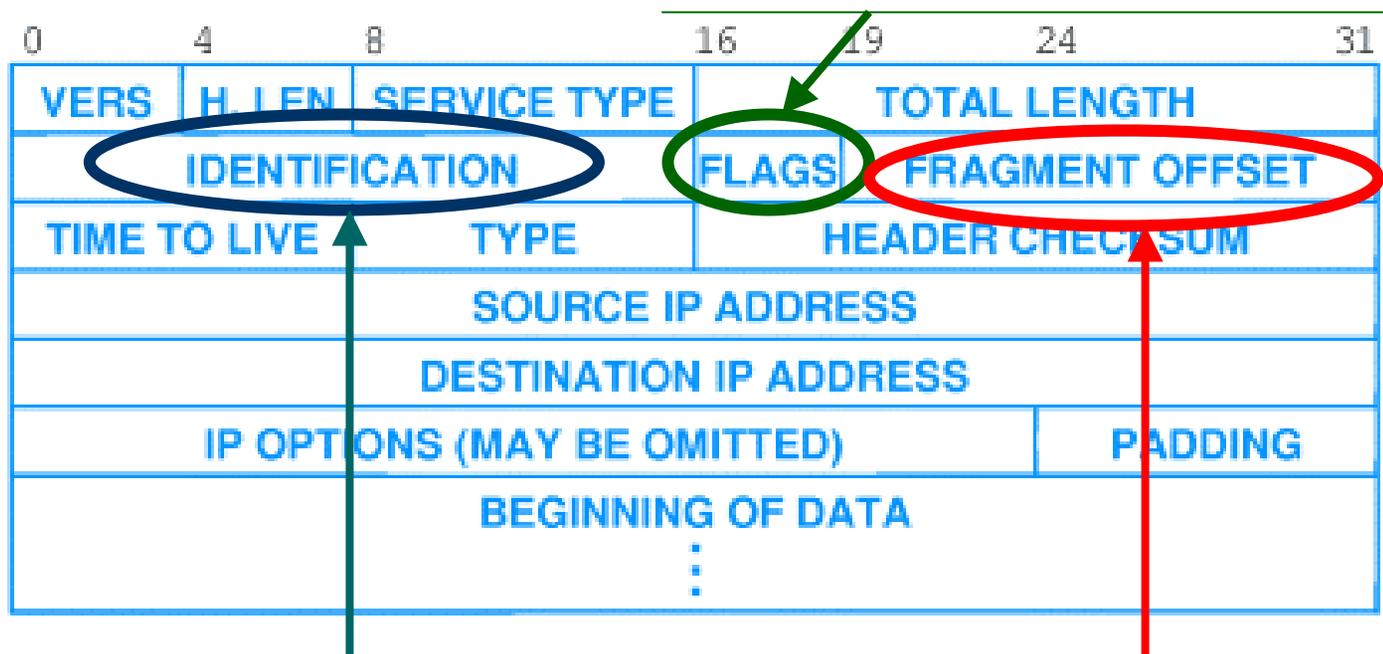
- Se dimensione datagram $>$ MTU → Frammentazione
 - Il router esegue la frammentazione prima della trasmissione
- Ogni frammento ha lo stesso formato di un datagramma
- Tutti i frammenti hanno dimensione pari a MTU
 - Tranne l'ultimo
- I frammenti vengono spediti come datagram indipendenti

Frammentazione



Frammentazione

Un particolare flag indica se si tratta di un frammento. Un altro specifica se si tratta dell'ultimo frammento



Numero di Sequenza. Viene ereditato da tutti i frammenti

Indica la posizione del frammento nel datagramma originale

Ricostruzione

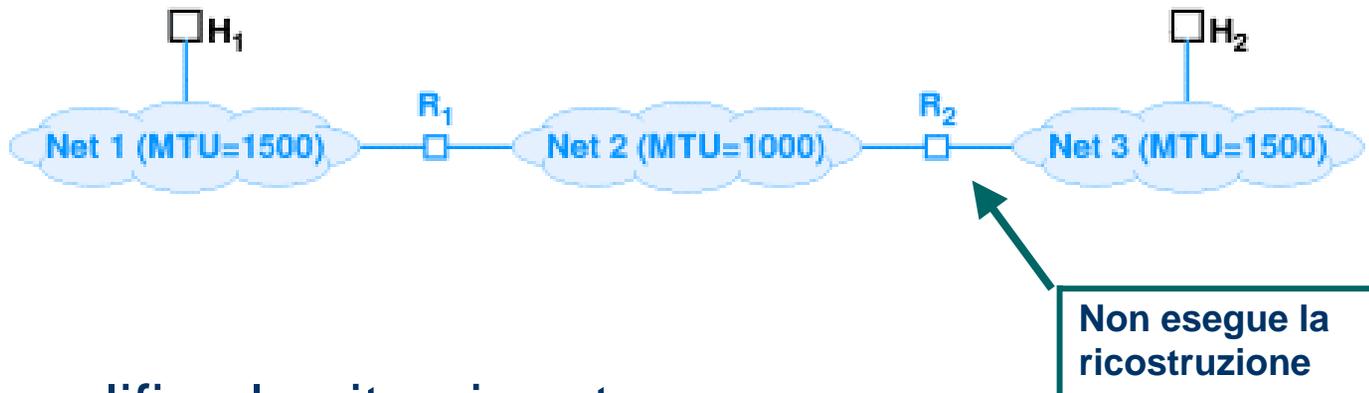
- Tutti i frammenti hanno lo stesso indirizzo destinatario
- Tutti i frammenti hanno lo stesso valore di Identification
- Ogni frammento è caratterizzato da un Offset
 - Posizione nel datagramma originale
- L'ultimo frammento è individuato da un particolare flag



Il datagramma originale può essere ricostruito

Ricostruzione

Ricostruzione presso il destinatario finale



- Semplifica la vita ai router
 - Non distinguono fra datagrammi e frammenti
- Non obbliga i frammenti a seguire tutti lo stesso percorso

Perdita dei frammenti

- La ricostruzione può avvenire solo se sono disponibili tutti i frammenti
- I frammenti possono arrivare fuori sequenza o non arrivare
 - Servizio connectionless e non affidabile
- Il destinatario bufferizza i messaggi mettendoli in ordine
- Passato un certo intervallo di tempo il destinatario butta via tutti i frammenti

Perdita dei frammenti

- Ritrasmissione selettiva dei frammenti inutile
 - Il mittente non è a conoscenza della frammentazione
- Ritrasmissione del datagramma inutile
 - Il datagramma ritrasmesso potrebbe seguire un percorso diverso dal precedente
- La filosofia del “tutto o niente” è coerente con la filosofia best effort

Frammentazione dei frammenti

Cosa succede se un frammento raggiunge una rete con un MTU inferiore alla sua lunghezza?

Viene frammentato!!

- I router non fanno distinzioni fra datagram e frammenti
- Il destinatario non ha bisogno di distinguere se un frammento proviene da un datagram o da un altro frammento
- Il destinatario ricostruisce il datagramma originale senza prima ricostruire i suoi frammenti

Tecnologie di Rete

**Notifica di Errori
Protocollo ICMP**



Introduzione

- **Gestione degli errori in IP**
 - ✓ I datagram errati vengono cestinati
- **Altre anomalie**
 - ✓ Host/rete destinataria irraggiungibile
 - ✓ Valore di TTL scaduto
 - ✓ Datagram scartato da un router per mancanza di buffer
 - ✓
- **Protocollo ICMP (Internet Control Message Protocol)**
 - ✓ Per gestire le situazioni anomale
 - ✓ E come strumento di notifica di informazioni

Messaggi ICMP

Type	Name
0	Echo Reply
1	Unassigned
2	Unassigned
3	Destination Unreachable
4	Source Quench
5	Redirect
6	Alternate Host Address
7	Unassigned
8	Echo
9	Router Advertisement
10	Router Selection
11	Time Exceeded
12	Parameter Problem
13	Timestamp
14	Timestamp Reply
15	Information Request
16	Information Reply
17	Address Mask Request
18	Address Mask Reply
19	Reserved (for Security)
20-29	Reserved (for Robustness Experiment)
30	Traceroute
31	Datagram Conversion Error
32	Mobile Host Redirect
33	IPv6 Where-Are-You
34	IPv6 I-Am-Here
35	Mobile Registration Request
36	Mobile Registration Reply
37-255	Reserved

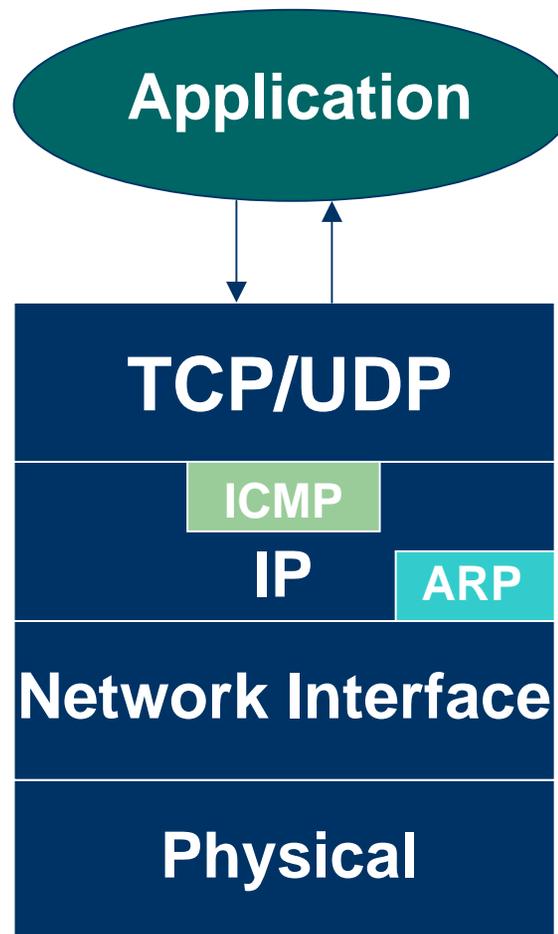
Messaggi di Errore

- Destination Unreachable (3)
 - ✓ Il router notifica al mittente che il destinatario non è raggiungibile o per un problema che riguarda la rete o per un problema dell'host.
- Source Quench (4)
 - ✓ Il router notifica al mittente che ha dovuto scartare un datagram per mancanza di buffer.
 - ✓ Si chiede al mittente di abbassare il rate di trasmissione
- Time exceeded (11)
 - ✓ Un router notifica al mittente che ha cestinato un datagram perché il TTL ha raggiunto il valore 0
 - ✓ Un host notifica al mittente che ha cestinato i frammenti di un datagram perché alcuni di essi non sono arrivati entro il time-out
- Parameter problem (12)
 - ✓ Uno dei parametri del datagram è errato

Messaggi di Notifica

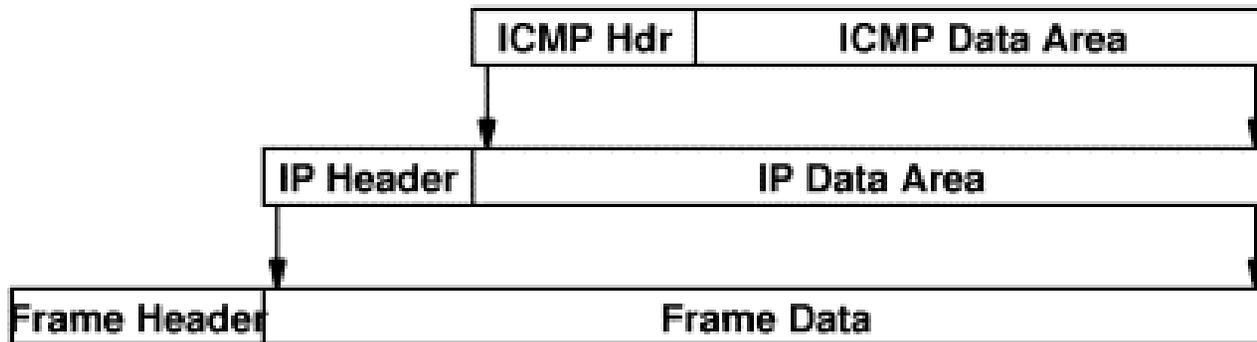
- Echo Request/Reply (8/0)
 - ✓ Se un host riceve una richiesta di eco il suo software ICMP risponde con un messaggio che trasporta gli stessi dati della richiesta
 - ✓ Usato nel programma ping
- Address Mask Request/Reply (17/18)
 - ✓ All'avvio un host genera una richiesta per conoscere la sua maschera
 - ✓ I router rispondono con un messaggio contenente la maschera a 32 bit relativa alla rete (sottorete) dell'host

Posizione del protocollo ICMP



Trasporto dei messaggi ICMP

- ICMP funzionalmente si trova sopra IP
 - Come TCP e UDP



Ping

```
C:\>ping www.unipi.it

Esecuzione di Ping www.unipi.it [131.114.190.24] con 32 byte di dati:

Risposta da 131.114.190.24: byte=32 durata=187ms TTL=245
Risposta da 131.114.190.24: byte=32 durata=125ms TTL=245
Risposta da 131.114.190.24: byte=32 durata=94ms TTL=245
Risposta da 131.114.190.24: byte=32 durata=94ms TTL=245

Statistiche Ping per 131.114.190.24:
    Pacchetti: Trasmessi = 4, Ricevuti = 4, Persi = 0 (0% persi),
Tempo approssimativo percorsi andata/ritorno in millisecondi:
    Minimo = 94ms, Massimo = 187ms, Medio = 125ms

C:\>
```

Ping

- Il programma trasmette una **echo request ICMP** verso l'host destinatario e attende la risposta
- Se la risposta non arriva entro un tempo prefissato viene visualizzata la scritta "risposta scaduta"
- L'host destinatario risponde con un messaggio di **echo reply ICMP**
- Il programma estrae dalla risposta alcune informazioni sul percorso

Traceroute/Tracert

```
Microsoft Windows 2000 [Versione 5.00.2195]
(C) Copyright 1985-1999 Microsoft Corp.

C:\>tracert www.unipi.it

Rilevazione instradamento verso www.unipi.it [131.114.190.24]
su un massimo di 30 punti di passaggio:

 1  <10 ms  <10 ms  <10 ms  rt50.univ.trieste.it [140.105.50.254]
 2  <10 ms  <10 ms  <10 ms  140.105.150.13
 3  <10 ms  <10 ms  <10 ms  utsgw48.univ.trieste.it [140.105.48.231]
 4   31 ms   31 ms   47 ms  rc-units2.ts.garr.net [193.206.132.29]
 5   31 ms   62 ms   47 ms  mi-ts-2.garr.net [193.206.134.53]
 6   47 ms   47 ms   47 ms  bo-mi-2.garr.net [193.206.134.6]
 7  125 ms  125 ms  125 ms  pi-bo-1.garr.net [193.206.134.82]
 8   *      204 ms  281 ms  unipi-rc.pi.garr.net [193.206.136.18]
 9  219 ms  312 ms  250 ms  eth03-gw.unipi.it [131.114.188.61]
10  219 ms  187 ms  204 ms  131.114.186.1
11  250 ms  266 ms  266 ms  solarια.adm.unipi.it [131.114.190.24]

Rilevazione completata.

C:\>
```

Traceroute/Tracert

- Usa datagram IP con TTL crescente
 - Traceroute trasmette un datagramma con TTL=1
 - Il primo router cestina il datagramma e manda al mittente un messaggio di errore ICMP
 - Traceroute viene così a conoscenza del primo router
 - Traceroute trasmette un secondo datagramma con TTL=2
 -

Traceroute/Tracert

- Come scoprire la fine del percorso?
 - I datagram inviati trasportano messaggi UDP destinati ad una applicazione inesistente presso il destinatario finale
 - Il mittente riceve un messaggio di errore ICMP 3 (destinatario irraggiungibile)
 - Capisce quindi che è stato raggiunto il destinatario finale

Traceroute/Tracert

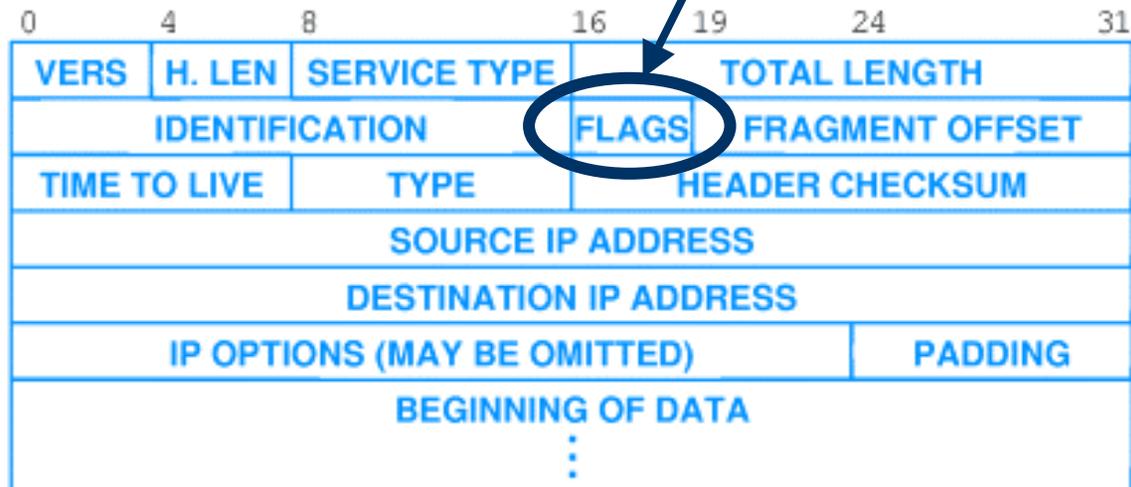
- Alcuni possibili problemi
 - I datagrammi possono perdersi
 - ✓ In questo caso traceroute deve ritrasmettere il datagramma
 - ✓ Dopo quanto tempo? Parametro di utente
 - I datagrammi possono seguire percorsi diversi
 - ✓ Se i datagrammi successivi seguono percorsi diversi la lista dei router si riferisce a tratti di percorsi diversi verso la destinazione
 - ✓ Utilità limitata a inter-reti relativamente stabili

ICMP e Frammentazione

- La frammentazione dei datagram comporta il consumo di risorse
 - ✓ I router devono generare e inviare i frammenti
 - ✓ Il destinatario deve memorizzare i frammenti e ricostruire il datagram
- La frammentazione si può disattivare
 - ✓ Settando un opportuno flag nella intestazione del datagram
 - ✓ Quando un router ha necessità di frammentare un tale datagram lo cestina e manda un messaggio di errore ICMP al mittente
- Si può stimare il valore MTU del percorso
 - ✓ Si inviano datagram di dimensione decrescente
 - ✓ La dimensione del primo datagram che non genera una notifica ICMP è una stima attendibile del MTU del percorso
 - ✓ Se il percorso varia la stima può non essere più significativa

Disabilitazione della frammentazione

Un particolare flag permette di disattivare la frammentazione del datagramma



Tecnologie di Rete

**Evoluzione del
protocollo IP: IPv6**



Perche cambiare?

- Spazio di indirizzi limitato
- Supporto ad applicazioni audio e video
- Supporto a comunicazioni di gruppo
 - Creare e modificare la composizione dei gruppi
 - Consegna efficiente dei msg a tutti i membri del gruppo

Cosa si richiede

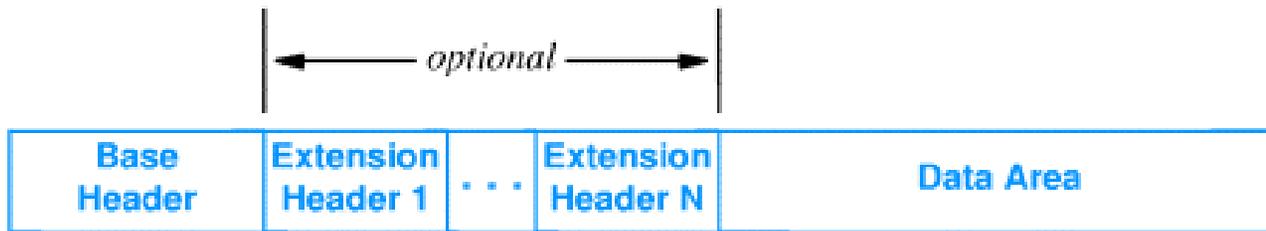
- Deve risolvere i problemi di IPv4
- Deve essere una evoluzione di IPv4
- Deve essere compatibile con IPv4
- Al nuovo protocollo è stato assegnato il numero di versione 6
 - Il numero 5 era già stato assegnato al protocollo ST (protocollo sperimentale)

Caratteristiche di IPv6

- Formato degli indirizzi
 - Indirizzi su 128 bit anziché su 32
- Formato dell'intestazione
 - Quasi tutti i campi sono stati modificati. Alcuni sostituiti.
- Intestazioni aggiuntive
 - Più intestazioni di lunghezza fissa anziché un'unica intestazione di lunghezza variabile
- Supporti per audio e video
 - Si può instaurare un collegamento virtuale e imporre a tutti i datagram di seguire il percorso del collegamento virtuale
 - Il servizio rimane best effort
- Estendibilità
 - Il mittente può allegare informazioni aggiuntive ai datagram

Formato Datagram IPv6

- Intestazione di base
- Intestazioni aggiuntive (opzionali)
- Carico



Indirizzi IPv6

- Analogie con IPv4
 - Un indirizzo per ogni collegamento host-rete
 - Prefisso + Suffisso
- Differenze con IPv4
 - 128 bit invece che 32
 - Non ci sono classi di appartenenza
 - La lunghezza di prefisso e suffisso non sono predeterminate
 - Non esiste un indirizzo riservato per trasmissione broadcast

Tipi di Indirizzi IPv6

- Unicast
 - Un singolo host
- Multicast
 - Gruppo di host anche in località diverse
 - Gruppo dinamico
- Anycast
 - Insieme di calcolatori che condividono lo stesso prefisso (es. Cluster)
 - Consegna a un unico calcolatore del cluster (es. il più vicino)

Notazione esadecimale

Notazione decimale puntata per un indirizzo a 128 bit

105.220.136.100.255.255.255.255.0.0.128.140.10.255.255

Notazione esadecimale con due punti (colon hex)

69DC:8864:FFFF:FFFF:0:1280:8C0A:FFFF

Compressione delle sequenze di zero

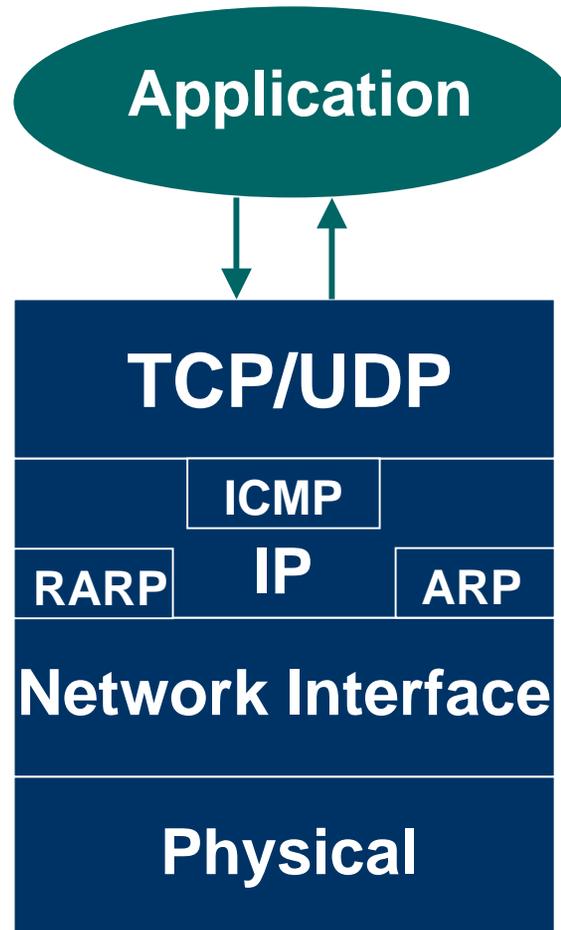
FF0C:0:0:0:0:0:0:B1 → FF0C::B1

Tecnologie di Rete

**Il livello Trasporto
Protocollo UDP**



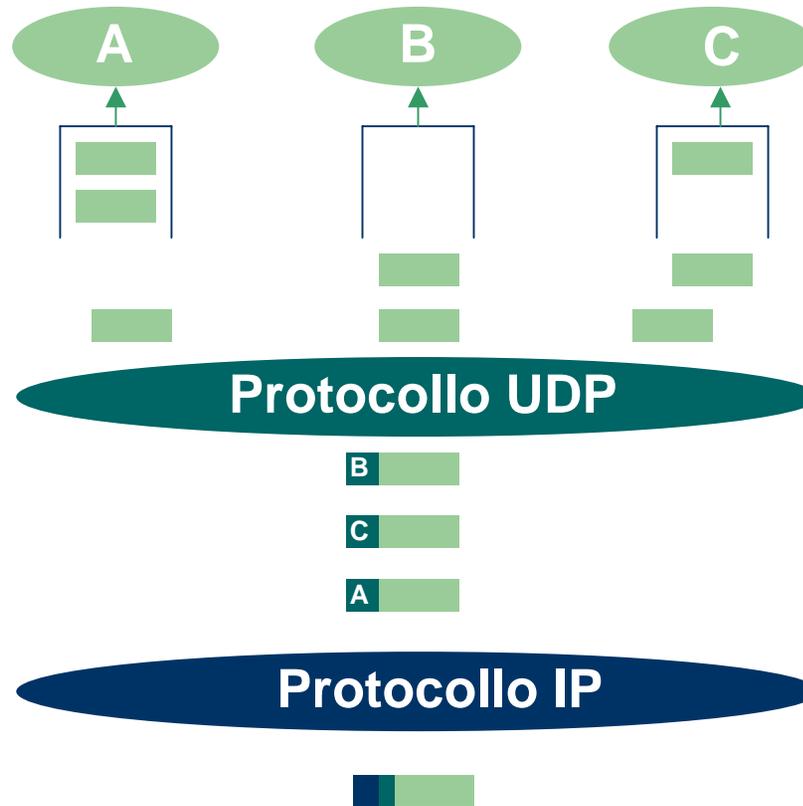
Punto della situazione



Livello di trasporto

- Estende il servizio di trasporto host-to-host in un servizio di comunicazione fra processi
 - Esegue il demultiplexing delle informazioni
 - Basato sul concetto di porta
- I processi vengono individuati mediante la coppia
<Host IP Address, Port Number>
- Il processo mittente deve specificare sia l'indirizzo IP che il numero di porta
- Il SO realizza la porta come coda di messaggi

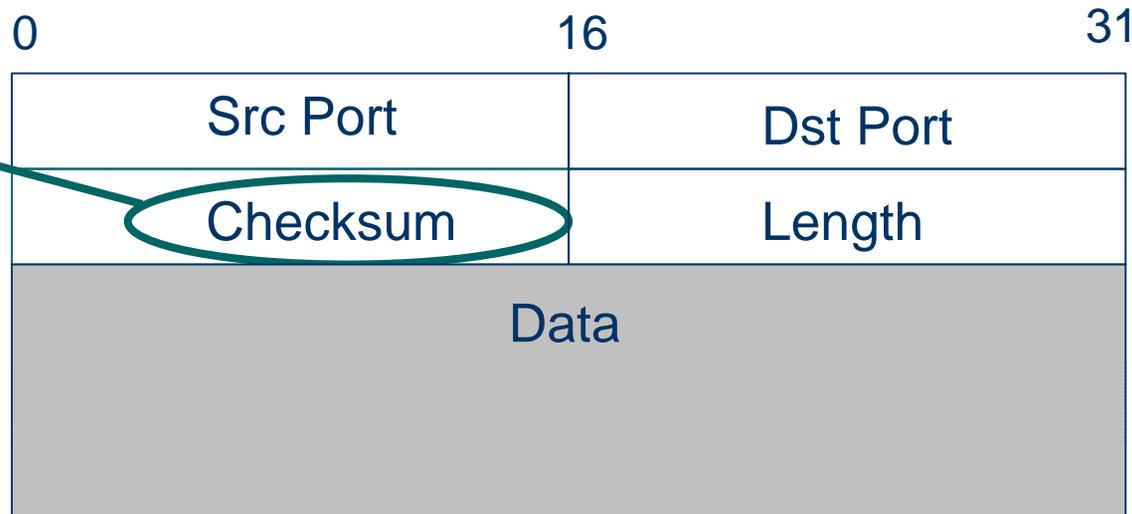
Protocollo UDP



Formato del messaggio UDP

Calcolato su intestazione UDP, corpo del messaggio e pseudo-header IP (Length, Src IP address, Dst IP Addr)

Serve ad assicurare che il messaggio sia consegnato al corretto destinatario finale (facoltativo in IPv4)



Utilizzo di UDP

- Applicazioni con scambio di informazioni limitato
 - Invio di messaggi singoli senza connessione
 - Controllo sull'affidabilità demandato all'applicazione
- Applicazioni real-time
 - Si chiede solo un servizio di trasporto end-to-end
 - Ulteriori funzionalità specifiche realizzate sopra UDP

Tecnologie di Rete

Protocollo TCP



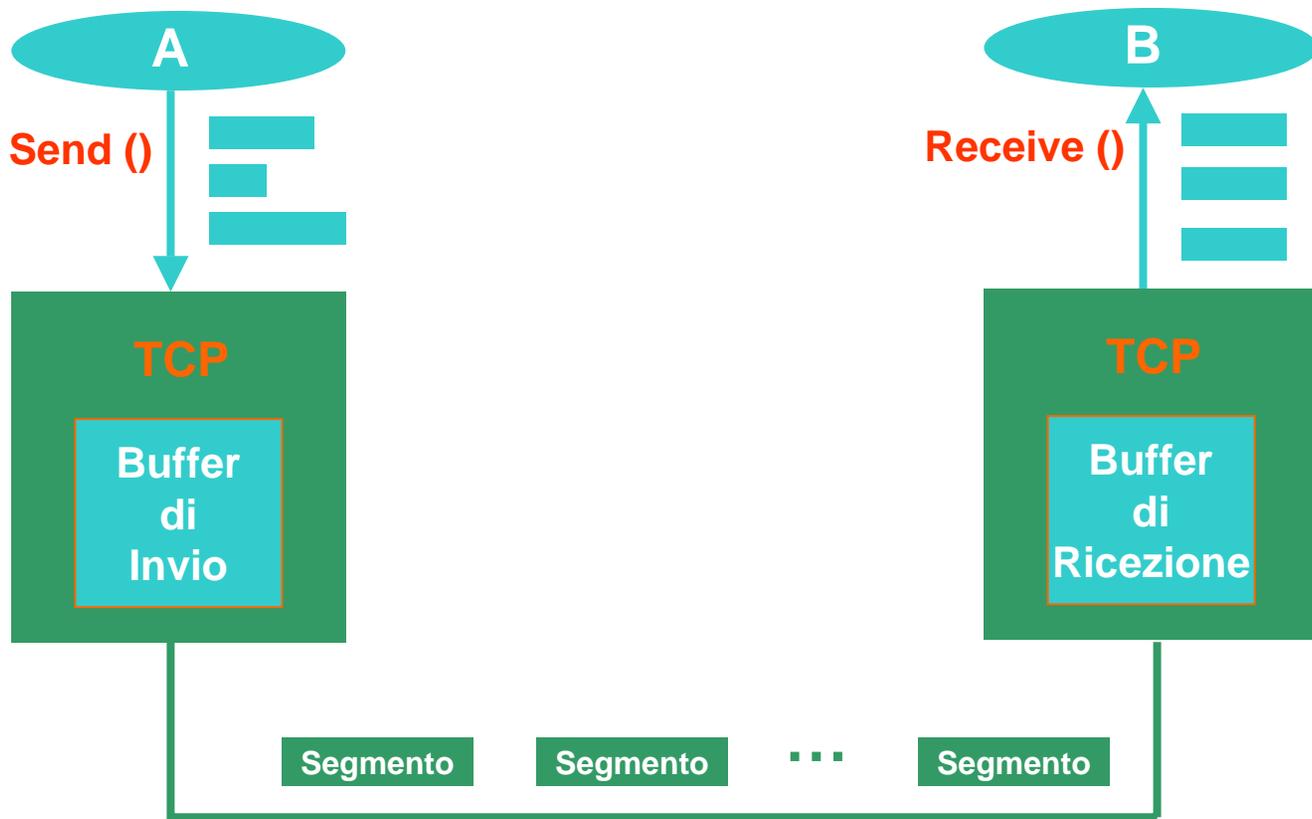
Introduzione al TCP

- Servizio di comunicazione fra processi orientato alla connessione
 - Meccanismo di demultiplexing basato sulle porte
 - Apertura connessione, trasferimento dati, chiusura connessione
- Punto-Punto
 - Ogni connessione TCP collega esattamente due processi
- Affidabile
 - Consegna affidabile, senza duplicati e in sequenza di un flusso di byte
- Full duplex
 - Un flusso di byte in ciascuna direzione

Introduzione al TCP

- Controllo e gestione degli errori
 - Checksum e ACK
 - Timeout e ritrasmissione
- Controllo del flusso
 - Evita che il mittente invii più dati di quanti il ricevitore possa gestire
- Controllo della congestione
 - Evita che il mittente invii più dati di quanti la rete sia in grado di trasportare

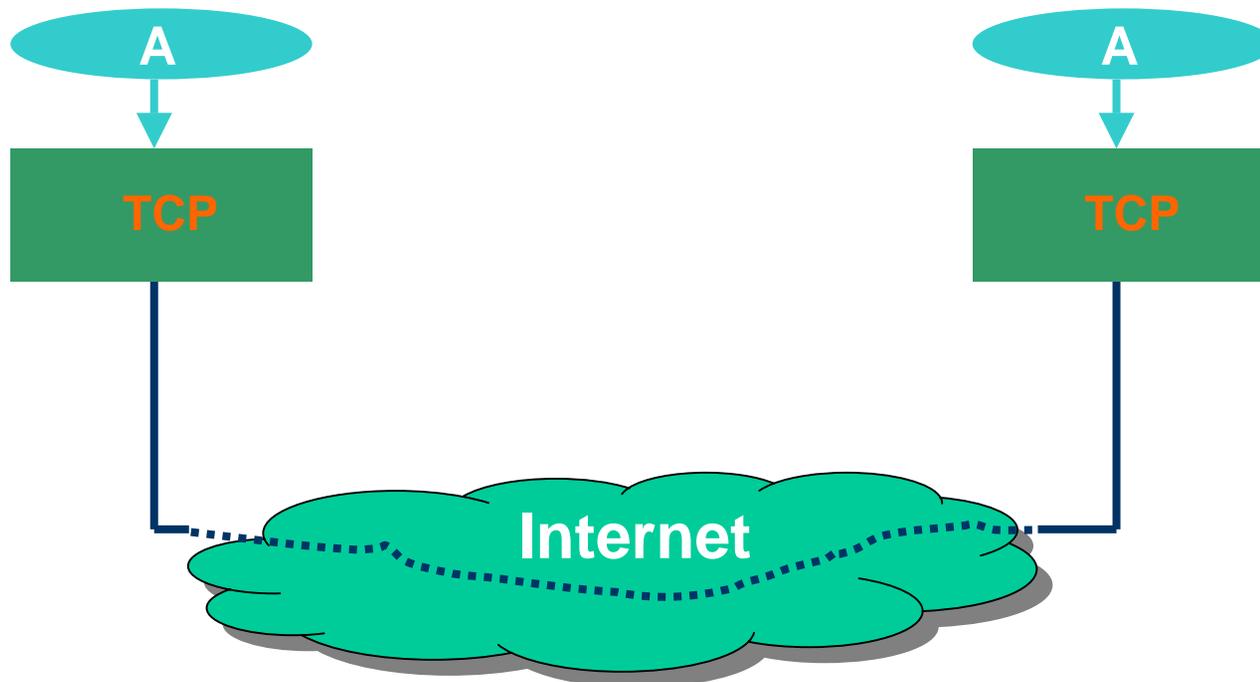
Flusso di byte



Quando inviare un segmento?

- Quando si raggiunge un certo numero di bytes
 - MSS (Maximum Segment Size)
 - $MSS = MSU \text{ locale} - \text{header TCP} - \text{header IP}$
- Quando scatta un timer
 - Per evitare attese infinite
 - Tutti i bytes nel buffer vengono spediti
- Quando il trasmettitore lo richiede esplicitamente
 - In Telnet ogni byte viene inviato in un segmento

Connessione



Fra il TCP mittente e quello destinatario esiste una connessione virtuale

Tutti i datagram IP seguono uno stesso percorso

Internet

Analogie con collegamento punto-punto

- TCP usa un meccanismo a finestra scorrevole
- TCP include un meccanismo di rilevazione degli errori
 - Basato su checksum
- TCP esegue il recupero da errore
 - Meccanismo basato su acknowledgment, time-out e ritrasmissione

Differenze con collegamento punto-punto

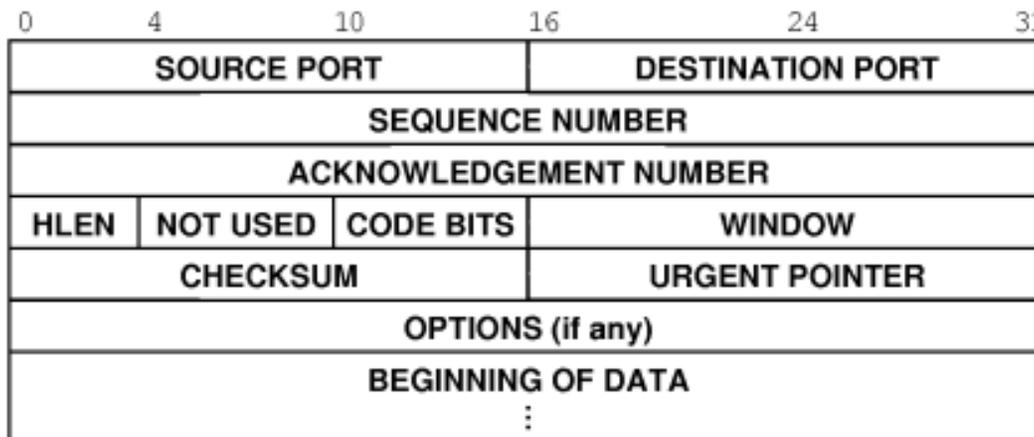
- Connessione non permanente
 - Apertura e chiusura della connessione
 - Nella fase di apertura vengono negoziati i parametri della finestra
- Round Trip Time (RTT) variabile
 - Dipende dalla connessione
 - Variabile nel tempo per una stessa connessione
 - Il time-out per le ritrasmissioni deve essere **adattivo**
- Nessun limite al ritardo dei pacchetti
 - La ricezione di segmenti vecchi potrebbe mettere in crisi il meccanismo a finestra scorrevole

Differenze con collegamento punto-punto

- I due end-point non sono sempre gli stessi
 - I due calcolatori potrebbero avere disponibilità di risorse molto diverse
 - Necessità di un meccanismo di controllo del flusso
- Il percorso verso il destinatario non è noto
 - Il mittente non conosce i link (e le relative capacità) attraversati dai segmenti
 - La connessione in esame non è l'unica attiva sulla inter-rete
 - Le condizioni di traffico possono variare da un istante all'altro
 - Il mittente non sa qual è la velocità ottimale per inviare le informazioni
 - Necessità di un meccanismo di controllo della congestione

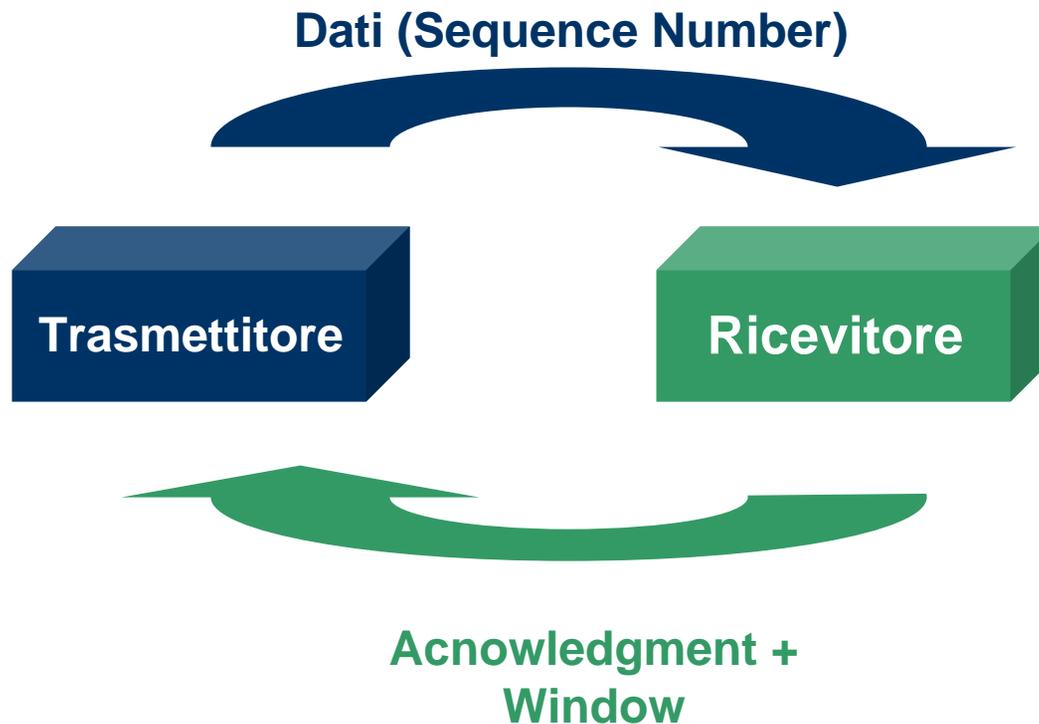
Formato del segmento TCP

Source/Destination Port	Porta associata al processo mittente/destinatario
Sequence Number	Numero di Sequenza (in termini di bytes)
Acknowledg. Number	Numero di byte correttamente ricevuti
Hlen	Lunghezza dell'intestazione
Code Bits	SYN – FIN – RESET – PUSH – URG - ACK
Window	Quantità di buffer ancora disponibile (in bytes)
Checksum	Somma di controllo relativa all'intero segmento
Urgent Pointer	Indica dove iniziano i dati non urgenti (se URG=1)



Flusso delle informazioni

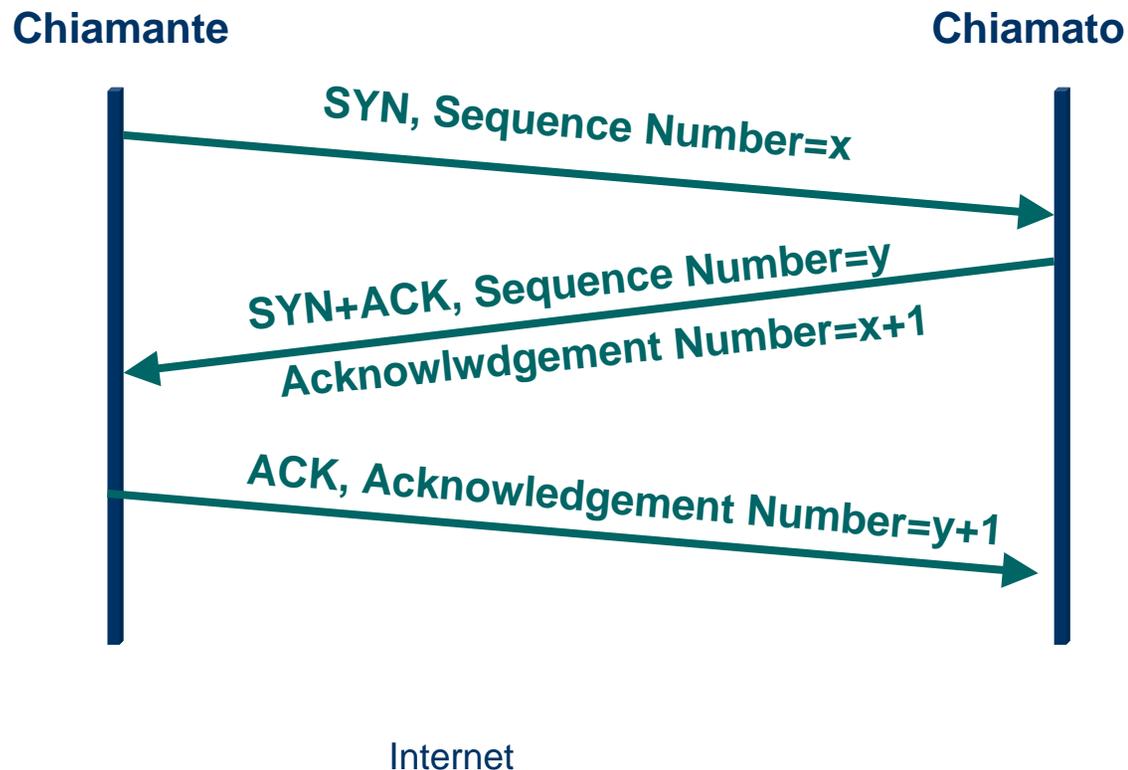
La comunicazione è full duplex



Internet

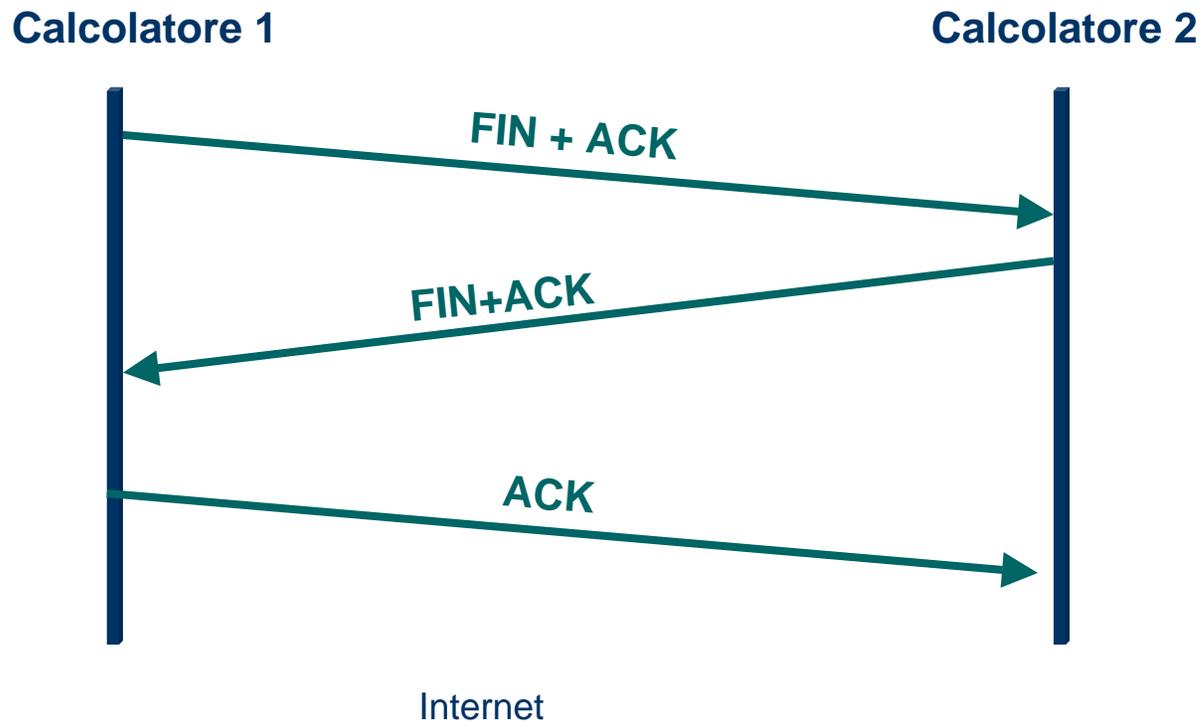
Attivazione della connessione

- Three-Way Handshake



Chiusura della connessione

- Three-way handshake anche per la chiusura della connessione

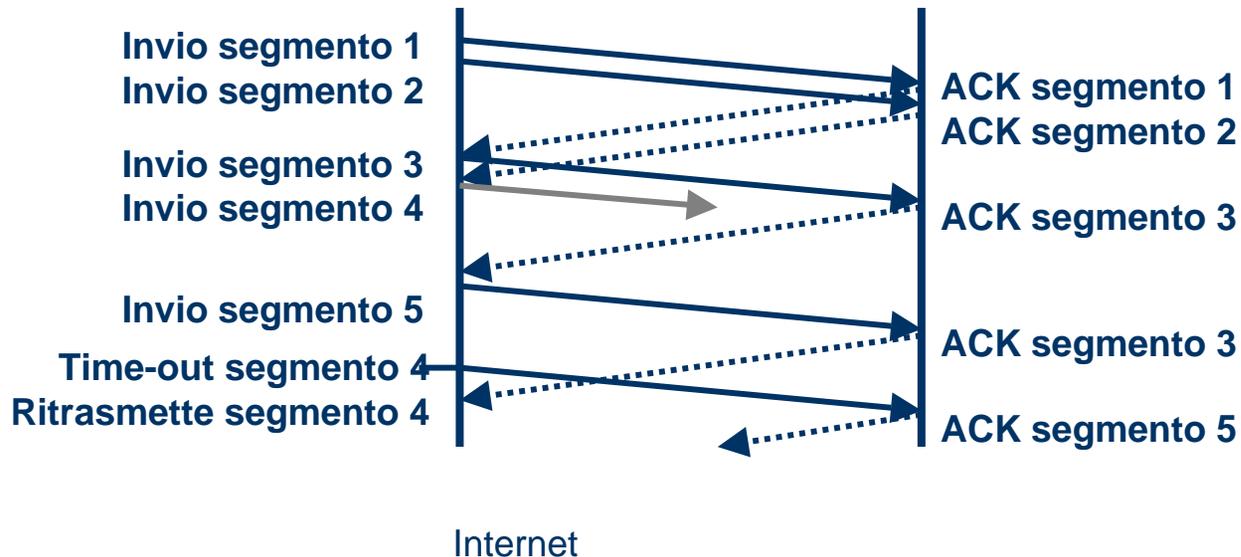


Meccanismo a finestra scorrevole

- Meccanismo a finestra scorrevole per
 - Garantire l'affidabilità della comunicazione
 - Garantire la consegna informazioni secondo l'ordine di invio
 - Esercitare il controllo del flusso
- Dimensione della finestra variabile
 - Comunicata dinamicamente dal ricevitore in base allo spazio di buffer disponibile
 - ✓ Campo window del segmento
 - Evita di inviare più informazioni di quante il ricevitore possa gestire

Gestione delle perdite

- Il trasmettitore invia un segmento e carica un timer
- Il ricevitore manda un ACK per riscontrare i segmenti ricevuti in sequenza
- Se l'ACK non arriva entro un certo time-out il trasmettitore ritrasmette il segmento



Ritrasmissione adattiva

- Qual è il valore ottimale del time-out?
 - Il tempo di andata e ritorno (RTT) dipende dalla distanza
 - e dalle condizioni di traffico sulla inter-rete
- Basato sulla stima dinamica del RTT

Calcolo del Time-out

- Algoritmo Originale

- Si misura il tempo intercorso fra l'invio di un segmento e la ricezione dell'ACK relativo (Sample_RTT)
- Si stima il RTT come media pesata fra la stima precedente di RTT e la misura corrente
- Si imposta il Time-out al doppio del valore stimato di RTT

$$\text{Estimated_RTT} = \alpha \cdot \text{Estimated_RTT} + (1 - \alpha) \cdot \text{Sample_RTT}$$

$$\text{Time-out} = 2 \cdot \text{Estimated_RTT}$$

$$0 < \alpha < 1$$

Calcolo del Time-out

- Limiti dell'algoritmo originale
 - Ambiguità nel caso di ritrasmissioni
 - Non si tiene conto della varianza del RTT
- Algoritmo di Karn-Partridge
 - Per la stima di RTT considera solo i segmenti non ritrasmessi

Calcolo del Time-out

- Algoritmo di Jacobson-Karel
 - Tiene conto della varianza
 - Legato al meccanismo di controllo della congestione

$$\text{Difference} = \text{Sample_RTT} - \text{Estimated_RTT}$$

$$\text{Estimated_RTT} = \text{Estimated_RTT} + (\delta \cdot \text{Difference}) \quad 0 \leq \delta \leq 1$$

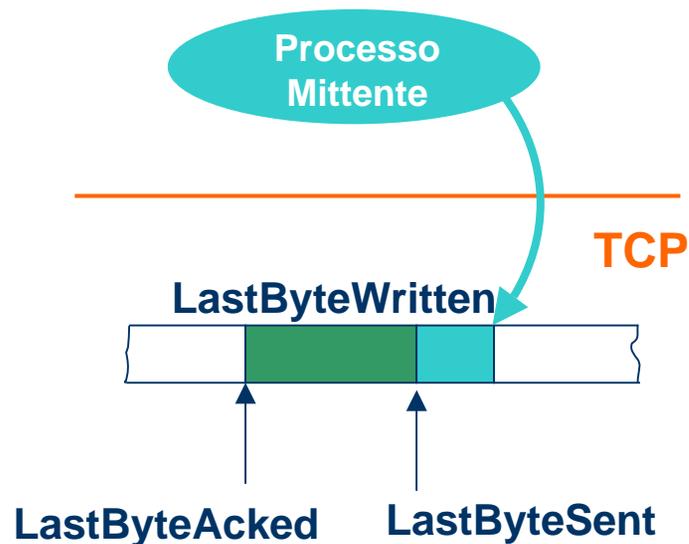
$$\text{Deviation} = \text{Deviation} + \delta (|\text{Difference}| - \text{Deviation})$$

$$\text{Time-out} = \mu \cdot \text{Estimated_RTT} + \phi \cdot \text{Deviation} \quad \mu = 1, \phi = 4$$

Buffer di invio e ricezione

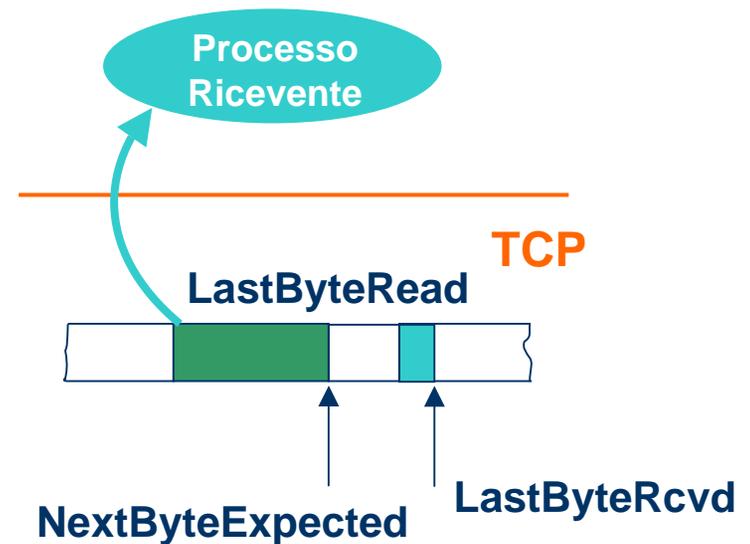
- Buffer di invio

- Messaggi trasmessi ma non ancora riscontrati
- Messaggi scritti dall'applicazione e non ancora trasmessi

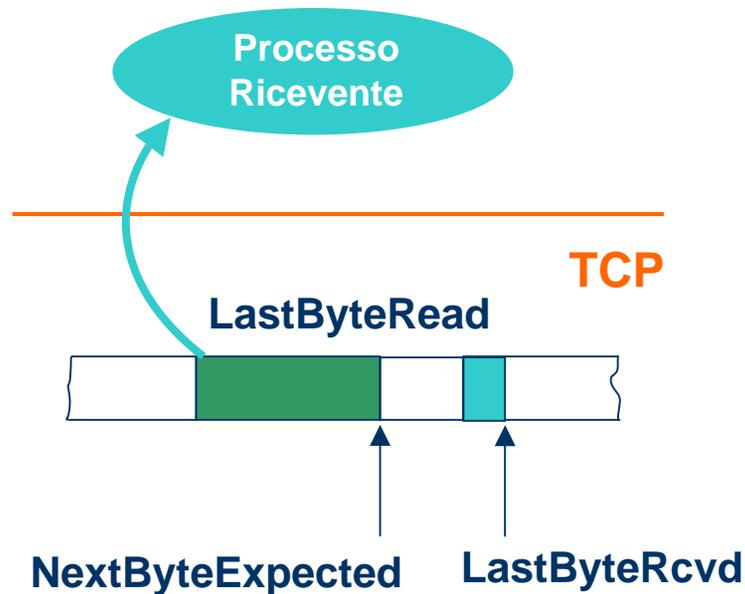


- Buffer di ricezione

- Dati arrivati fuori ordine
- Dati arrivati in ordine ma non ancora letti dal processo ricevente



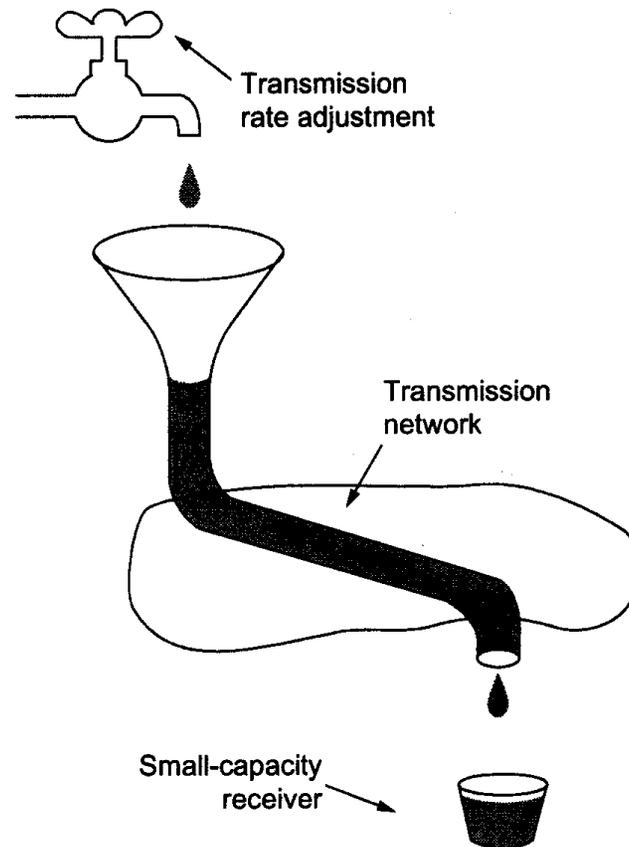
Calcolo della finestra (receiver)



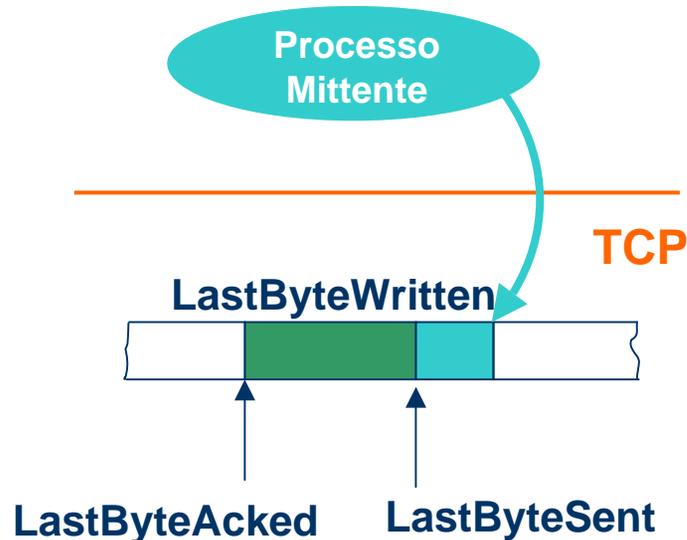
$$\text{LastByteRcvd} - \text{LastByteRead} \leq \text{MaxRcvBuffer}$$

$$\text{AdvertisedWindow} = \text{MaxRcvBuffer} - (\text{LastByteRcvd} - \text{LastByteRead})$$

Controllo del flusso



Calcolo della finestra (sender)

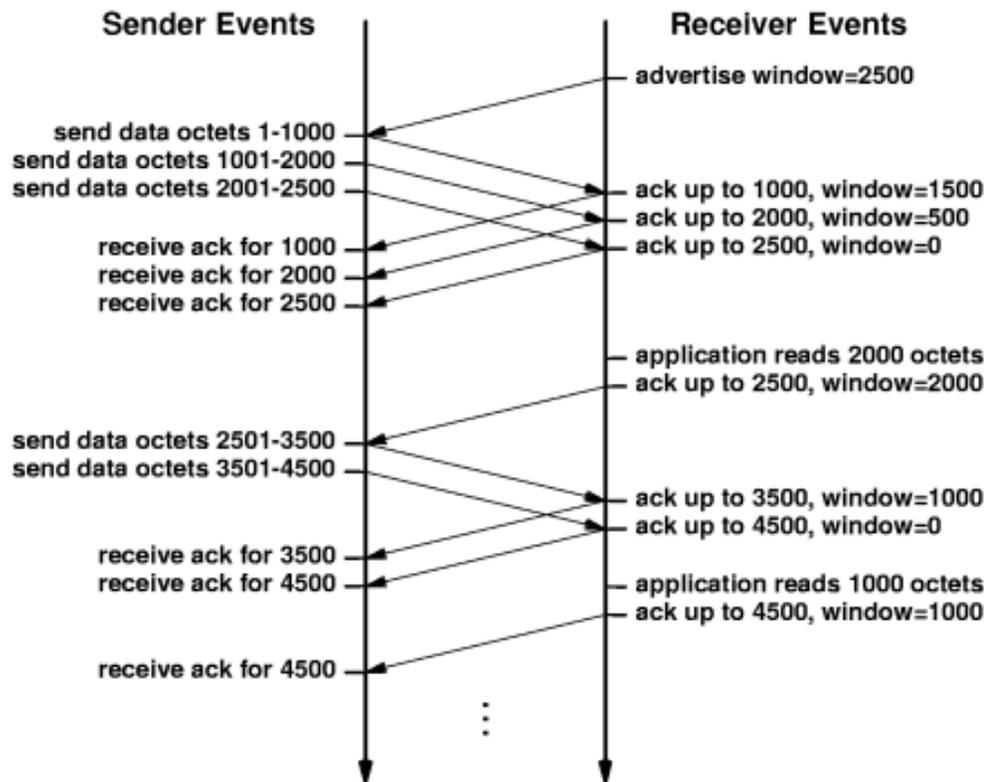


$$\text{LastByteSent} - \text{LastByteAked} \leq \text{AdvertisedWindow}$$

$$\text{EffectiveWindow} = \text{AdvertisedWindow} - (\text{LastByteSent} - \text{LastByteAked})$$

Il TCP mittente deve bloccare il processo mittente quando lo spazio nel buffer di invio è esaurito.

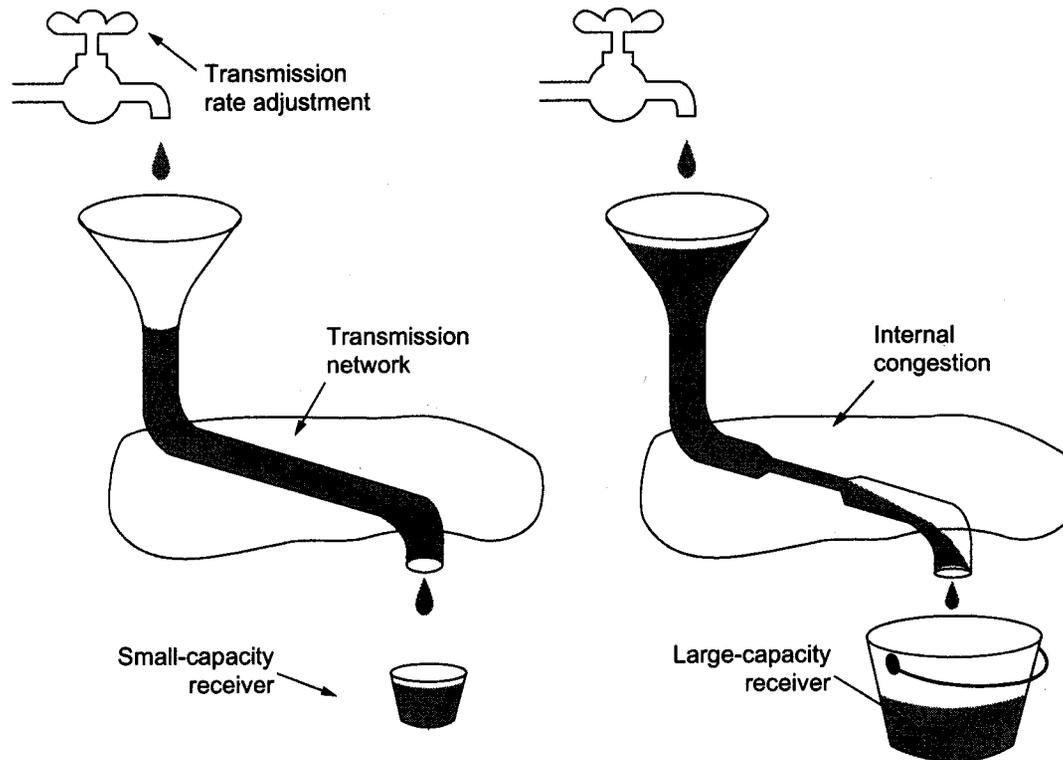
Esempio



Considerazioni finali

- Se `AdvertisedWindow=0` il TCP mittente si blocca
- Quando riprendere la trasmissione di nuovi dati?
 - Il TCP ricevitore non invia segmenti vuoti per annunciare la nuova dimensione della finestra
- Il TCP mittente invia periodicamente segmenti di prova
 - contengono 1 solo byte di dati
 - Il ricevente viene così sollecitato a inviare una risposta

Controllo di flusso e di congestione



Internet

Controllo della congestione

- Idee di base

- Si cerca di prevenirla
- Quando sembra esserci, si riduce drasticamente il tasso di invio
 - ✓ In modo da ridurre la congestione

- Implementazione

- Meccanismo a finestra
- Prevenzione, Rilevazione e Gestione della congestione

Controllo della congestione

- Basata su meccanismo a finestra
 - ✓ come controllo del flusso
- Controllo di flusso e congestione sovrapposti
 - ✓ ma indipendenti (due finestre)
- Valore effettivo della finestra
 - $\text{Min}(\text{cong_win}, \text{flow_win})$

Controllo della congestione

- Prevenzione della congestione
 - Si aumenta gradualmente il valore della finestra
 - ✓ Slow start: crescita esponenziale
 - ✓ Congestion avoidance: crescita lineare
- Rilevazione della congestione
 - Mancata ricezione di ACK assunta come indice di congestione
 - Assunzioni implicite
 - ✓ perdite dovute sempre a congestione
 - ✓ nodi statici

Controllo della congestione

- Gestione della congestione
 - Riduzione drastica della dimensione della finestra di congestione
 - ✓ Riduzione del tasso di invio

Algoritmo di controllo della congestione

Fase Slow Start

- Alla creazione della connessione:
finestra di congestione = MSS
- Si spedisce un numero di segmenti pari alla dimensione (in bytes) della finestra
- Se si riceve l'ACK la finestra di congestione viene raddoppiata
- La finestra di congestione cresce esponenzialmente fino a quando:
 - Si raggiunge la finestra del ricevente (controllo di flusso)
 - Si raggiunge il valore della soglia di congestione

Algoritmo di controllo della congestione

Fase Congestion Avoidance

- Dopo il raggiungimento della soglia la dimensione della finestra cresce linearmente
- Quando si verifica un time-out
 - Soglia = $\frac{1}{2}$ del valore corrente della finestra di congestione
 - Finestra di congestione = 1
 - Si riparte con l'algoritmo slow start

Valore iniziale della soglia = 64 KB

Esempio

