

Reti di Calcolatori e Laboratorio

Federico Matteoni

Indice

1	Introduzione	4	12	HTTP	19
2.1	Tipi di Rete	5	12.1	HTTP URL	19
2.2	Internetwork	5	12.2	Caratteristiche	19
2.3	Switching	6	12.2.1	Modello	19
3	Internet	6	12.2.2	Connessioni	19
3.1	Enti Ufficiali	7	12.3	Esempio HTTP	20
3.2	Reti di accesso	8	12.4	Messaggi HTTP	20
4	Metriche di Riferimento	8	12.5	Header	21
5	Modelli Stratificati	9	12.5.1	Request header	22
5.1	Perché stratificare	9	12.6	HTTP Response	24
5.2	Smistamento Intermedio	10	12.6.1	Response Headers	24
5.3	Elementi fondamentali	10	12.7	Negoziazione del contenuto	24
5.4	Modalità di Servizio	10	12.7.1	Entity Headers	25
5.5	Vantaggi	10			
6	Protocolli	11	13	Web Caching	25
6.1	Incapsulamento	11			
7	OSI RM (Open Systems Interconnection Reference Model)	11	14	Cookies	25
7.0.1	Pila di protocolli	11	15	Telnet	26
8	Flusso dell'Informazione	12	15.1	Introduzione	26
9	Stack protocollare TCP/IP	13	15.2	Protocollo Telnet	26
9.1	I Livelli	13	15.3	NVT	27
10	Livello Applicativo	14	15.4	Architettura	28
10.1	Protocollo a Livello Applicativo	14	15.5	Funzionamento	29
10.2	Paradigmi	14	16	SSH	30
10.3	Componenti di un'Applicazione di Rete	14	17	TCP Port Forwarding	30
10.4	Terminologia	15	18	FTP	31
10.5	Identificazione di un Processo	15	18.1	Modello FTP	31
10.6	Esempio di API: TCP	15	18.1.1	Connessione di Controllo	31
10.7	Uso dei Servizi di Trasporto	16	18.1.2	Connessione Dati	31
11	Applicazioni Web e HTTP	17	18.2	Altri Dettagli	32
11.1	Terminologia	17	18.2.1	Due Modalità	32
11.2	Uniform Resource Identifier	17	18.2.2	Stateful	33
11.2.1	Sintassi	18	18.2.3	Modalità di trasmissione	33
11.2.2	Assolute e Relative	18	18.3	Anonymous FTP	33
			19	DNS	34
			19.1	Motivazioni	34
			19.2	Struttura	34
			19.3	Servizi	35
			19.4	Spazio dei nomi	35
			19.4.1	Indirizzi	35
			19.4.2	Nomi	35
			19.4.3	Top-Level Domains	37
			19.4.4	Struttura di un nome alfanumerico	37
			19.5	Conversione	38
			19.5.1	Name Servers	38
			19.5.2	Root Name Servers	39

19.5.3 Gerarchia dei server	39	23.11.3 Eventi lato destinatario	67
19.6 Risoluzione dei nomi	40	23.11.4 Esempi	68
19.7 Caching e Aggiornamento Record	41	23.12 Calcolo del timeout	72
19.8 Record DNS	41	23.13 Finestra di trasmissione	72
19.9 Messaggi DNS	41	23.14 Controllo della congestione	74
20 SMTP	43	23.14.1 Algoritmo per il controllo della congestione	75
20.1 Agenti Utente	43	23.14.2 cWnd	75
20.2 Mail Server	43	23.14.3 AIMD	75
20.3 Schema di principio	44	23.14.4 Slow Start	76
20.4 Indirizzo	44	23.14.5 Politica Reno per il controllo della congestione	77
20.5 Gestione Alias	45	23.14.6 Politica Tahoe per il controllo della congestione	78
20.6 Modello di riferimento	46	23.15 Throughput	79
20.7 Simple Mail Transfer Protocol	46	23.16 Fairness	79
20.8 SMTP Mail Relaying	46	23.17 Transmission Control Block	79
20.9 Modello SMTP	47		
20.9.1 Fallimenti nella consegna	47		
20.10 Protocollo	47		
20.11 Comandi SMTP	48		
20.12 Esempio	48		
20.13 Formato messaggi mail	49		
20.14 Estensioni multimediali	50		
20.14.1 MIME	50		
20.14.2 Tipi MIME	51		
20.15 Protocolli di accesso alla mail	51		
21 Livello di Trasporto	52	24 Livello di Rete	80
21.1 Obiettivi	52	24.1 Servizi	80
21.2 Caratteristiche	53	24.2 Architettura di rete IP	80
21.3 Servizi offerti	53	24.3 Protocol Data Units	81
22 UDP	53	24.4 Alcuni protocolli della Suite TCP/IP	81
22.1 Proprietà	53		
22.2 Datagramma UDP	54		
22.3 Calcolo del checksum	54		
22.4 TCP vs UDP	55		
23 TCP	56	25 IP	82
23.1 Proprietà	56	25.1 Funzioni	82
23.2 Funzioni del segmento TCP	56	25.1.1 Inoltro	82
23.3 Processi e Socket TPC	57	25.1.2 Instradamento	82
23.4 Trasferimento bufferizzato	57	25.1.3 Indirizzamento	82
23.5 Segmenti TCP	57	25.2 Multiplexing / Demultiplexing	83
23.6 Numeri di sequenza e di riscontro	58	25.3 Datagramma IP	83
23.7 Segmento TCP	58	25.3.1 Formato	83
23.7.1 Campi del segmento TCP	59	25.3.2 Campi	84
23.7.2 Formato del segmento TCP	60	25.4 Frammentazione	84
23.8 Gestione della connessione	60		
23.8.1 Three-Way Handshake	60	26 Indirizzamento IP	86
23.8.2 Perché a tre vie	61	26.1 Indirizzi IPv4	86
23.8.3 Esempio	61	26.2 Strategie di addressing	86
23.8.4 Chiusura della connessione con handshake	62	26.2.1 Classful addressing	87
23.9 Half-Close	63	26.2.2 Classless addressing	87
23.9.1 Scenario Half-Close	64	26.3 Subnet Mask	88
23.9.2 Stato Time-Wait	64	26.4 Indirizzi speciali	88
23.10 Stati del TCP	65	26.5 Assegnazione degli indirizzi	88
23.11 Trasferimento Dati Affidabile	66		
23.11.1 Sequenza e riscontro	66	27 DHCP	89
23.11.2 Eventi lato mittente	66		
28 IP Forwarding	89	28 IP Forwarding	89
28.1 Forwarding Diretto	89	28.2 Forwarding Indiretto	89
29 NAT	90		
30 ICMP	91		
30.1 Formato dei messaggi	91		
31 Ping	91		
32 Traceroute	92		

33 Routing	93	36.4 Router	112
33.1 Router	93		
33.2 Routing Engine	94		
33.3 Forwarding Engine	94		
33.4 Algoritmi di Routing	94		
33.4.1 Tipologie	95		
33.4.2 Algoritmo Link State	96		
33.4.3 Distance Vector Routing	98		
33.4.4 Count-To-Infinity Problem	99		
33.4.5 Link-State vs Distance Vector	99		
34 Struttura di Internet	100		
34.1 Sistemi Autonomi	100		
34.2 Protocolli d'Instradamento	101		
34.3 Routing Gerarchico	102		
34.4 Protocolli	102		
34.4.1 RIP	102		
34.4.2 OSPF	103		
34.4.3 BGP	104		
35 Livello Data Link	105		
35.1 Collegamenti	105		
35.2 Servizi	106		
35.3 Indirizzi	106		
35.3.1 Indirizzi MAC	106		
35.3.2 Indirizzamento	106		
35.3.3 ARP	107		
35.4 Ethernet	110		
35.4.1 Topologia a Stella	110		
35.4.2 Pacchetti	110		
36 Dispositivi di interconnessione	111		
36.1 Repeater	111		
36.2 Hub	111		
36.3 Switch	111		
36.3.1 Switch con auto apprendimento . . .	111		
37 VLAN	112		
37.1 Definire le VLAN	113		
37.1.1 Appartenenza per gruppo di porta .	113		
37.1.2 Appartenenza per indirizzo MAC .	113		
37.1.3 Appartenenza per informazioni sul			
protocollo	113		
37.2 Comunicare l'appartenenza	114		
38 Cenni di applicazioni P2P	114		
38.1 Paradigma Peer-to-Peer	114		
38.2 Directory Centralizzata	115		
38.3 Reti Decentralizzate	115		
38.4 Reti Non Strutturate	115		
38.5 Copertura Gerarchica	116		
38.6 BitTorrent	116		
38.6.1 Strategie principali	116		
38.7 Reti Strutturate	117		
38.7.1 DHT	117		
39 Sicurezza	118		
39.1 Cifratura a Chiave Simmetrica	118		
39.2 Cifratura a Chiave Asimmetrica	119		
39.3 Message Digest	120		
39.4 Message Authentication Code	120		
39.5 Firma Digitale	121		
39.6 IPsec	121		
39.6.1 ESP	122		
40 Esercizi	123		
40.1 SMTP	123		
40.2 DNS	124		
40.3 TCP	125		
40.4 IP	130		
40.5 Routing	132		

1 Introduzione

Appunti del corso di **Reti di Calcolatori** presi a lezione da **Federico Matteoni**.

Prof.: **Federica Paganelli**, federica.paganelli@unipi.it

Riferimenti web:

- elearning.di.unipi.it/enrol/index.php?id=169

Password: **RETI2019**

Esame: scritto (o compitini), discussione orale facoltativa + progetto con discussione (progetto + teoria di laboratorio, progetto da consegnare 7gg prima della discussione)

Libri e materiale didattico:

- Slide su eLearning
- IETF RFC
 - tools.ietf.org/rfc
 - www.ietf.org/rfc.html
- "Computer Networks: A Top-Down Approach" B. A. Forouzan, F. Mosharraf, McGraw Hill

Ricevimento: stanza 355 DO, II piano

Immagini: Copyright 1996–2005 J.F Kurose and K.W. Ross

2 Rete

Definizione di rete Interconnessione di dispositivi in grado di scambiarsi informazioni, come end system, router, switch e modem.

Gli end system possono essere di due tipi:

Host: una macchina, in genere di proprietà degli utenti, **dedicata ad eseguire applicazioni**. Esempi: desktop, portatile, smartphone, tablet...

Server: una macchina, tipicamente con elevate prestazioni, destinata ad eseguire programmi che **forniscono servizi** a diverse applicazioni utente. Esempi: posta elettronica, web, ...

Con il termine host si può anche indicare un server.

2.1 Tipi di Rete

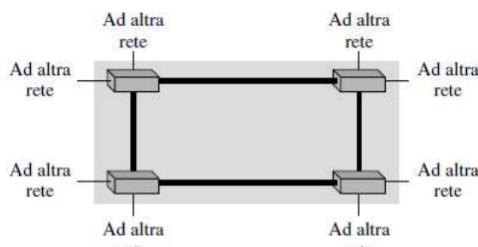
Local Area Network Una LAN è una rete di area geografica limitata: un ufficio, una casa ecc.. I dispositivi comunicano attraverso una determinata tecnologica: switch, BUS, HUB ecc..

In una rete locale tipicamente una serie di host comunicano tra loro attraverso, ad esempio, uno switch centrale.

Wide Area Network Alcuni esempi:



Legenda
■ Dispositivo di comunicazione
— Mezzo trasmissivo

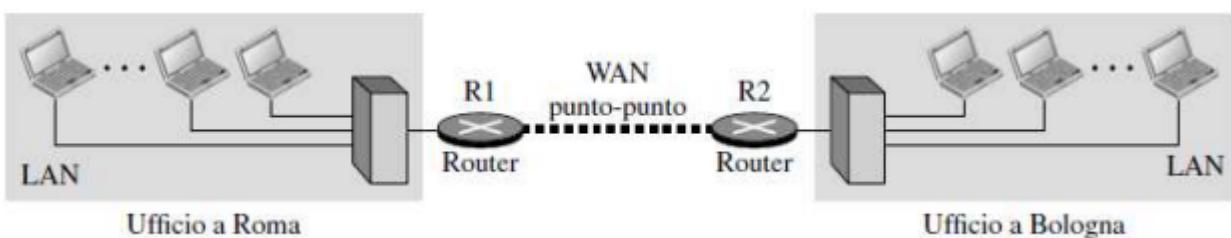


Legenda

■ Switch
— Mezzo trasmissivo

2.2 Internetwork

Una internetwork si crea quando si interconnettono diverse reti. Alcuni esempi:





2.3 Switching

Una rete internet è formata dall'interconnessione di reti composte da link e dispositivi capaci di scambiarsi informazioni. In particolare, i sistemi terminali comunicano tra di loro per mezzo di dispositivi come switch, router ecc. che si trovano nel percorso tra i sistemi sorgente e destinazione.

Switched Network Reti a commutazione di circuito, tipico delle vecchie reti telefoniche

Le risorse sono riservate end-to-end per una connessione. Le risorse di rete (es. bandwidth) vengono suddivise in pezzi, e ciascun pezzo è allocato ai vari collegamenti. Le risorse rimangono inattive se non vengono utilizzate, cioè **non c'è condivisione**. L'allocazione della rete rende necessario un setup della comunicazione.

A tutti gli effetti vi è un circuito dedicato per tutta la durata della connessione. Ciò è rende poco flessibile l'utilizzo delle risorse (**overprovisioning**).

Packet-Switched Network Reti a commutazione di pacchetto, più moderno

Flusso di dati punto-punto suddiviso in pacchetti. I pacchetti degli utenti condividono le risorse di rete. Ciascun pacchetto utilizza completamente il canale.

Store and Forward: il commutatore deve ricevere l'intero pacchetto prima di ritrasmetterlo in uscita.

Le risorse vengono usate **a seconda delle necessità**. Vi è **contesa per le risorse**: la richiesta di risorse può eccedere la disponibilità e si può verificare **congestione** quando i pacchetti vengono accodati in attesa di utilizzare il collegamento. Si possono anche verificare perdite.

3 Internet

L'internet più famosa ed utilizzata è **internet**, ed è composta da migliaia di reti interconnesse. **Ogni rete** connessa ad internet **deve utilizzare il protocollo IP** e rispettare certe convenzioni su nomi ed indirizzi. Si possono aggiungere nuove reti ad internet molto facilmente.

Dispositivi in internet I **dispositivi** connessi ad internet possono essere host, end systems come PC, workstations, servers, pda, smartphones ecc...

I **link di comunicazione** possono essere fibre ottiche, doppini telefonici, cavi coassiali, onde radio... Le **entità software** in internet possono essere:

Applicazioni e processi

Protocolli: regolamentano la trasmissione e la ricezione di messaggi (TCP, IP, HTTP, FTP, PPP...)

Interfacce

Standard di internet e del web: RFC (Request for Comments) e W3C.

Internet è una visione dei servizi. L'infrastruttura di comunicazione permette alle applicazioni distribuite di scambiare informazioni (WWW, e-mail, giochi, e-commerce, controllo remoto...) e fornisce loro **servizi di comunicazione connectionless** (senza garanzia di consegna) o **connection-oriented** (dati garantiti in integrità, completezza ed ordine).

3.1 Enti Ufficiali

L'**Internet Engineering Task Force** (IETF) è l'organismo che studia e sviluppa i protocolli in uso su internet. Si basa su gruppi di lavoro a cui chiunque può accedere. I documenti ufficiali che pubblica, dove descrivono i protocolli usati in internet, sono gli RFC/STD (Request for Comments/STanDards).

L'**Internet Corporation for Assigned Names and Numbers** (ICANN) si occupa di coordinare il sistema dei **nomi di dominio** (DNS) e assegna i gruppi di indirizzi, gli identificativi di protocollo.



3.2 Reti di accesso

Il collegamento tra l'utente ed il primo router di internet è detto **rete di accesso**. Può avvenire in 3 modi:

Tramite rete telefonica: servizio dial-up, ADSL...

Tramite reti wireless

Collegamento diretto, come collegamenti WAN dedicati ad alta velocità (aziende e università)

4 Metriche di Riferimento

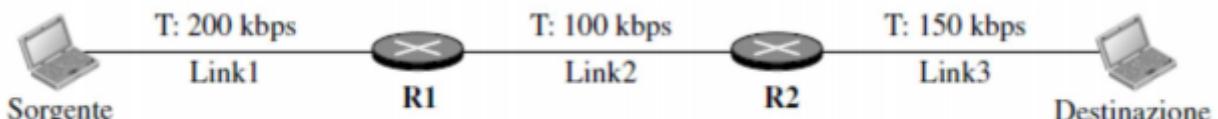
Come misurare le prestazioni della rete? Tramite una serie di metriche:

Bandwidth o ampiezza di banda: è la larghezza dell'intervallo di frequenze utilizzato dal sistema trasmissivo (Hz).

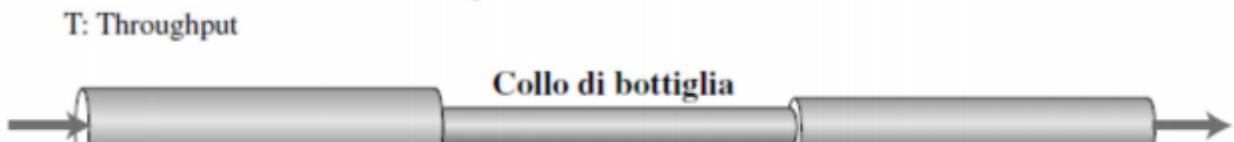
Bitrate o **transmission rate**: quantità di bit che possono essere trasmessi o ricevuti nell'unità di tempo (bit/secondo, bps)

Il bitrate dipende dalla bandwidth e dalla tecnica trasmissiva utilizzata.

Throughput: la quantità di traffico che arriva realmente a destinazione nell'unità di tempo (al netto di perdite sulla rete, funzionamento dei protocolli ecc...).



a. Un percorso attraverso tre link



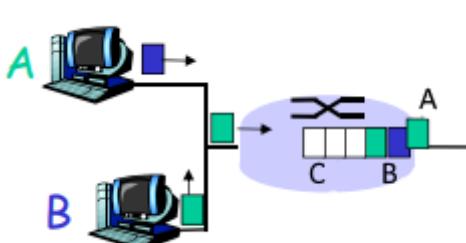
b. Simulazione utilizzando dei tubi

Non è detto che corrisponda alla bandwidth perché ci potrebbe essere un collo di bottiglia.

Latenza o ritardo: il tempo richiesto affinché un messaggio arrivi a destinazione dal momento in cui il primo bit parte dalla sorgente.

latenza = ritardo di propagazione + ritardo di trasmissione + ritardo di accodamento + ritardo di elaborazione

Perdita di pacchetti. Come si può verificare?



A → pacchetti **in attesa** di essere trasmessi (*ritardo*)

B → pacchetti **accodati** (*ritardo*)

C → buffer **liberi** (se non ci sono buffer liberi, i pacchetti in arrivo vengono scartati, *perdita*)

I pacchetti da spedire vengono accodati nei buffer del router. Di solito, il tasso di arrivo dei pacchetti sul router eccede le capacità del router di evaderli, quindi i pacchetti si accodano in attesa del proprio turno.

Il **ritardo di elaborazione** è dato dal controllo sui bit e dalla determinazione del canale di uscita (trascutibile)
Il **ritardo di accodamento** è dato dall'attesa di un pacchetto di essere trasmesso (B)

Il **ritardo di trasmissione** è il tempo impiegato per trasmettere un pacchetto sul link.

$$R_{trasmissione} = R/L$$

R = rate di trasmissione del collegamento, in bps

L = lunghezza del pacchetto in bit

Il **ritardo di propagazione** è il tempo impiegato da 1 bit per essere propagato da un nodo all'altro.

$R_{propagazione} = d/s$

d = lunghezza del collegamento

s = velocità di propagazione del collegamento (si usa la velocità della luce, circa 3×10^8 m/s)

$d_{nodal} = d_{proc} + d_{queue} + d_{trans} + d_{prop}$

d_{proc} = ritardo di elaborazione, pochi microsecondi

d_{queue} = ritardo di accodamento, dipende dalla congestione

d_{trans} = ritardo di trasmissione, L/R e significativo a lunga distanza

d_{prop} = ritardo di propagazione, d/s, da pochi microsecondi a centinaia di millisecondi

5 Modelli Stratificati

https://elearning.di.unipi.it/pluginfile.php/27387/mod_resource/content/1/L02_introduzione_protocolli.pdf

Perché usare un modello a strati Per mandare dei dati da un host ad un altro comunicando su una rete, si devono eseguire una serie di operazioni: trovare il percorso di rete da attraversare, decidere in che modo spedire e codificare i dati, risolvere eventuali problemi di comunicazione e altro ancora.

Programmare ogni volta tutto il procedimento è un lavoro estremamente complesso e ripetitivo. Un modello a strati **astrae su più livelli il problema della trasmissione dati** in modo da fornire di volta in volta strumenti utili al programmatore per poter evitare di "reinventare la ruota".

Definizioni generali Nelle architetture di comunicazione a strati sono importanti una serie di definizioni:

- Stratificazione
- Information hiding
- Separation of concern
- Modello ISO/OSI
- Stack TCP/IP

Tali definizioni verranno viste durante il corso.

Lo Strato Uno strato è un **modulo interamente definito** attraverso i servizi, le interfacce e i protocolli che lo caratterizzano. Si indica anche col nome di livello.

Uno strato **n comunica direttamente** con lo strato **n** di un'altra unità tramite un **protocollo assegnato**. Lo stesso strato **n** può richiedere servizi allo strato **n-1** attraverso la **loro interfaccia**, e fornisce servizi allo strato **n+1** attraverso la **rispettiva interfaccia**.

Es. modello stratificato: sistema postale Vedi slide

https://elearning.di.unipi.it/pluginfile.php/27387/mod_resource/content/1/L02_introduzione_protocolli.pdf, 48

Dal livello più alto al livello più basso per la spedizione, viceversa per la ricezione. Un problema importante che si incontra quando si manda una lettera, ad esempio, dall'Italia al Giappone è la traduzione. In una **serie di passi**, in cui in ognuno viene **eseguito un particolare compito su un messaggio**, che viene poi **trasferito ad un altro livello**. Nell'esempio, la segretaria prepara lettera (traduce in giapponese e imbusta) affinché il postino la possa prendere. Però il "messaggio" della segretaria è "scritto" per essere interpretato dalla segretaria giapponese, il direttore italiano scrive per il direttore giapponese. **Messaggi di un livello del sistema che spedisce sono scritti per essere interpretati dal medesimo livello del sistema ricevente.**

5.1 Perché stratificare

La stratificazione è molto utile per **scomporre il sistema complesso della gestione della comunicazione**. Prendo un sistema estremamente costoso da costruire per una singola coppia di aziende, quindi lo trasformo in strati così che il costo della singola lettera sia irrisorio.

Definisco funzioni di base per effettuare trasferimento e agenti che le svolgono. Principi di base:

Separation of Concern

Far fare ad un determinato strato solo ciò che gli compete, delegando agli altri tutto ciò che è delegabile

Information Hiding

Nascondo ad un determinato strato le informazioni non indispensabili allo svolgere della sua operazione.

Esempio Se traduco il modello postale nel modello a strati ho, ad esempio:



5.2 Smistamento Intermedio

Spedire un pacchetto che è destinato ad determinato livello intermedio. Quindi **arrivo fino al corrispondente livello intermedio per evitare che si possano esporre info sensibili**.

5.3 Elementi fondamentali

Gli **elementi fondamentali** del modello stratificato sono:

Flusso dati

Servizio: una **funzione** che uno strato offre allo strato superiore, attraverso un'interfaccia.

Protocollo

Interfaccia: insieme di regole che governano il formato e il significato dei frame, pacchetti o messaggi che vengono scambiati tra strati adiacenti della stessa entità.

I servizi indicano *cosa* si può fare, le interfaccie regolano *come* si può fare.

5.4 Modalità di Servizio

I **dati** possono essere scambiati in due modalità diverse:

Connection-Oriented: il livello di trasferimento stabilisce una **connessione logica** tra due sistemi. La connessione è quindi **gestita**:

- **Instaurazione** della connessione
- **Trasferimento** dei dati
- **Chiusura** della connessione

Connectionless: i dati vengono **trasferiti senza stabilire una connessione**.

5.5 Vantaggi

Il vantaggio più grosso è che **sviluppare il singolo strato è più semplice ed economico rispetto a sviluppare tutto il sistema complesso**. Questo perché i servizi degli strati inferiori vengono usati da più entità che implementano gli strati superiori.

6 Protocolli

Cos'è un protocollo Un **protocollo** è un **insieme di regole** che dice come comunicare ed esporre dati verso l'esterno. I protocolli **definiscono il formato e l'ordine dei messaggi inviati e ricevuti, con le azioni per trasmettere e ricevere tali messaggi.**

6.1 Incapsulamento

Processo in cui aggiungo strati, "involucri" al messaggio originale che vengono man mano tolti alla destinazione.

7 OSI RM (Open Systems Interconnection Reference Model)

Le prime reti erano chiuse, composte da tecnologie e protocolli proprietari. Alcuni esempi sono ARPANET, SNA (IBM), DNA (Digital). Non potevano intercomunicare tra loro, perché usavano **protocolli diversi**, erano costruite per servizi specifici (TELCO). Insorse quindi un obiettivo: creare un **modello di riferimento per sistemi aperti**, per permettere a qualsiasi terminale di poter comunicare mediante qualsiasi rete. C'era quindi necessità di **accordarsi sulle regole**.

OSI L'OSI è una **collezione di protocolli aperti**: questo significa che i loro **dettagli sono pubblici** e i **cambiamenti vengono gestiti da un'organizzazione con partecipazione aperta al pubblico**. Un sistema che implementa i protocolli aperti è un **sistema aperto**.

7.0.1 Pila di protocolli

L'OSI prevede **sette strati di protocolli**:

7. **Applicativo**: elaborazione dei dati
6. **Presentazione**: unificazione dei dati, preparazione del **pacchetto** da trasmettere/ricevere
5. **Sessione**: controllo del dialogo tra gli host sorgente e destinazione.
4. **Trasporto**: offre il vero e proprio trasferimento dati tra gli host terminali, cioè **astrae la logica** con la quale si scambiano i dati tra host e gestisce gli errori. Realizza il **dialogo end-to-end**.
3. **Rete**: instradamento del traffico (principalmente router, offre il servizio di consegna attraverso il sistema distribuito dei nodi intermedi).
2. **Datalink**: consegna il frame tra le interfacce, interpretato dalla scheda di rete.
1. **Fisico**: modulazione del segnale elettrico per trasmettere correttamente il flusso di bit sul mezzo fisico.

I livelli **7-5** si possono raggruppare in più modi, principalmente sono gli strati di supporto all'elaborazione e all'interazione con l'utente. Sono livelli **software**

I livelli **4-1** sono **software e hardware**. Sono strati di supporto alla rete e all'infrastruttura trasmittiva, cioè gestiscono la vera e propria trasmissione dei dati.

8 Flusso dell'Informazione

Per le reti, l'informazione ha origine al **livello applicativo**, che la genera per mandarla in remoto. Una volta generata, essa discende i vari livelli fino al canale fisico e ogni livello **aggiunge** all'informazione ricevuta dal livello superiore **una – o più – propria sezione informativa** sotto forma di **header**, contenente informazioni esclusive di quel livello. Per l'informazione ricevuta, si segue il **cammino inverso**, quindi dal basso verso il livello applicativo, e ogni livello "spacchetta" l'header del livello corrispondente, ne legge le informazioni esclusive e lo gestisce appositamente. Il **processo di encapsulamento** è quindi **reversibile**: ogni livello esegue una operazione di encapsulamento su dati già encapsulati dal livello precedente, in modo tale da garantire la possibilità di estrarre i dati precedentemente encapsulati.

HEADER —— DATA —— TRAILER

Header: qualificazione del pacchetto per questo livello

Data: payload proveniente dal livello superiore

Trailer: generalmente usato in funzione del trattamento dell'errore



9 Stack protocollare TCP/IP

Il **TCP/IP** è una famiglia di protocolli attualmente utilizzati in internet. Si tratta di una **gerarchia di protocolli** costituita da **moduli interagenti**, ciascuno con funzioni specifiche.

Gerarchia Con il termine **gerarchia** s'intende che ciascun protocollo di livello superiore è **supportato dai servizi forniti dai protocolli di livello inferiore**. Cioè un protocollo a livello n realizza le sue funzionalità grazie ai protocolli a livello $n-1$.

9.1 I Livelli

Lo stack TCP/IP in origine era intesa come **quattro livelli software** sovrastanti **un livello hardware**. Oggi è intesa come semplicemente **composta da 5 livelli**



Livello Applicativo Il livello più alto, con il quale interagisce l'utente

Identificativi risorse: URL, URI, URN

Il web: user agents, http: request, response, connessioni persistenti, GET, POST, PUT, DELETE, status code, proxy server, caching

FTP: connessioni dati e di controllo, rappresentazione TELNET

Posta elettronica: SMTP, POP3, IMAP

DNS e risoluzioni nomi: gerarchia nomi, risoluzione iterativa e ricorsiva, formato pessaggi, nslookup...

Livello Trasporto Livello al quale si definisce la codifica e il protocollo di trasporto

Servizi: mux demux, controllo errore, connectionless

TCP: formato segmenti, gestione connessione, controllo flusso e congestione

UDP: formato segmenti

Livello Rete Dove si gestisce l'indirizzamento dei vari host

Strato di rete e funzioni

Indirizzamento IP: classful IPv4, NAT, sottoreti e maschere, classless, CIDR

Risoluzione IP e MAC, ARP

IPv4: formato datagramma ip, frammentazione

Routing IP e istradamento

Introduzione IPv6

Livello Link Trasferimento dati tra elementi di rete vicini

Ethernet

Livello Fisico Bit sul filo

10 Livello Applicativo

Applicazioni e processi Le applicazioni possono essere composte da **vari processi distribuiti** comunicanti fra loro. Un **processo** sono programmi eseguiti degli host di una rete. Due processi possono anche comunicare all'interno dello stesso host attraverso la **comunicazione inter-processo** (definita dal S.O.).

Nella **comunicazione a livello applicativo fra due host di una rete**, due o più processi vengono eseguiti da ciascuno degli host comunicanti e **si scambiano messaggi**.

I livelli applicazione dei due host **si comportano come se esistesse un collegamento diretto** attraverso cui mandare e ricevere messaggi.

10.1 Protocollo a Livello Applicativo

Un protocollo di livello applicativo **definisce**:

Tipi di messaggi scambiati a quel livello, ad esempio di richiesta o di risposta

Sintassi dei vari tipi di messaggi, cioè i campi

Semantica dei campi

Regole per determinare quando e come un processo invia o risponde ai messaggi

10.2 Paradigmi

I due programmi applicativi devono essere entrambi in grado di richiedere e offrire servizi oppure ognuno deve occuparsi di uno dei due compiti?

Client-Server Un **numero limitato di processi server** che **offrono** un servizio, in esecuzione **in attesa di richieste** dai **processi client**, che **richiedono** servizi.

Client *Parla per primo*, cioè inizia il contatto con il server. Tipicamente richiede un servizio al server, ad esempio: per il web il client è implementato nel browser, per l'e-mail è implementato nel mail reader.

Server Fornisce al client il servizio richiesto e **rimane sempre attivo**. Ad esempio: un web server invia le pagine richieste, un mail server smista ed invia le mail.

Peer-to-Peer Host **peer** che possono **offrire servizi e inviare richieste**.

Misto Un misto tra i due paradigmi sopra.

10.3 Componenti di un'Applicazione di Rete

Due esempi

Web Composto da:

- Web Browser, sul **client**
- Web **Server**
- **Standard per il formato** delle risorse (pagine ecc.)
- **Protocollo HTTP**

Posta Elettronica Composta da:

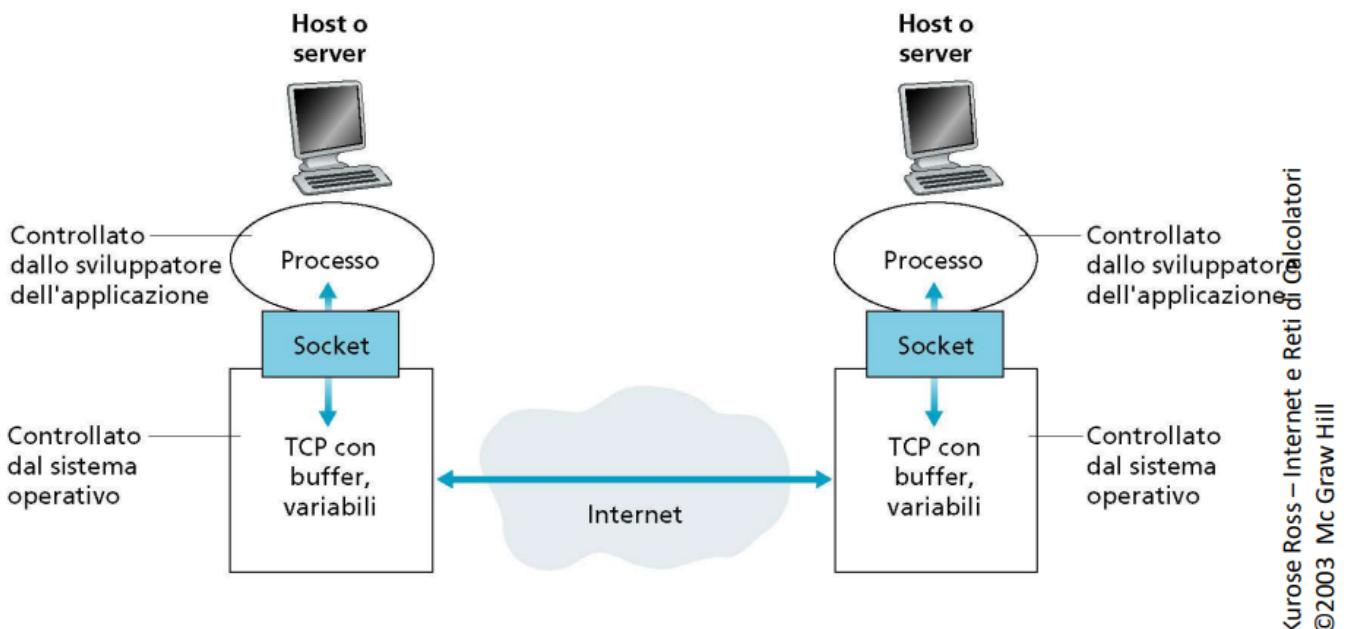
- Programmi di lettura e scrittura sul **client**
- **Server** di posta in internet
- **Standard per il formato** dei messaggi
- **Protocolli SMTP, POP3 ecc.**

10.4 Terminologia

API Application Programming Interface: si tratta di un **insieme di regole** che un programmatore deve rispettare per utilizzare delle risorse.

Socket Una **API** che funge da **interfaccia** tra gli strati di applicazione e di trasporto. A tutti gli effetti è la **API di internet per eccellenza**, due processi comunicano mandando dati sul socket e leggendoli da esso. Forma una **connessione logica**, l'invio e ricezione dei dati sono responsabilità del S.O. e del TCP/IP.

10.5 Identificazione di un Processo



I servizi di trasporto sono offerti al livello applicativo tramite le API. Ogni servizio di transport è **usato simultaneamente** da più processi application.

Come identifico i processi di livello application **di host diversi**? Serve un identificativo che identifichi sia l'host che il processo.

→ Coppie <Indirizzo IP, Numero di porta>



10.6 Esempio di API: TCP

```

Connection TCPopen(IPAddress, int)           //per aprire una connessione
void TCPsend(Connection, Data)                //per spedire dati su una connessione
Data TCPreceive(Connection)                  //per ricevere dati da una connessione
void TCPclose(Connection)                   //per chiudere una connessione
int TCPbind(int)                           //per richiedere l'assegnazione della porta su
                                         //cui attendere le richieste di connessione

void TCPunbind(int)                        //per liberare la porta assegnata
Connection TCPaccept(int)                 //per attendere le richieste di connessione

//Connection: identificata da una quadrupla
//Astraggo dalle possibili eccezioni sollevate e dal loro trattamento

```

10.7 Uso dei Servizi di Trasporto

Una coppia di processi fornisce servizi agli utenti di Internet, siano questi persone o applicazioni. La coppia di processi, tuttavia, **deve utilizzare i servizi offerti dal livello di trasporto** per la comunicazione, poiché non vi è una comunicazione fisica a livello applicativo. Le applicazioni di rete sono quindi **realizzate sopra ai servizi di trasporto dati**.

Nel livello trasporto dello stack protocollare TCP/IP sono previsti **due protocolli di trasporto principali**:

TCP Transfer Control Protocol

Connection-Oriented: è richiesto un setup tra client e server

Trasporto **affidabile** tra mittente e destinatario

Controllo del flusso: il mittente non *inonderà* di dati il destinatario

Controllo di congestione: *limita* il mittente quando la rete è satura

Non offre garanzie di timing né di ampiezza minima di banda

UDP User Datagram Protocol

Connectionless

Trasporto **non affidabile**

NO controllo del flusso

NO controllo di congestione

NO garanzie di timing o di ampiezza minima di banda

Quindi quali applicazioni usano UDP, e perché?

Che tipo di trasporto richiede un'applicazione?

Throughput Anche detta **banda**, è la frequenza alla quale il processo mittente può inviare i bit al processo ricevente. Alcune applicazioni (es. multimedia) richiedono una **banda minima** per essere efficaci, altri (**elastic apps**) usano la banda che trovano a disposizione.

Velocità di trasferimento \neq velocità di propagazione

Perdita di dati Alcune applicazioni (es. audio) possono tollerare alcune perdite, altre (es. telnet, trasferimento file) richiedono un trasferimento dati **affidabile** al 100%

Timing Alcune applicazioni (es. teleconferenze, videogame) richiedono un basso ritardo per essere efficaci

Applicazione	Tolleranza alla perdita dati	Throughput	Sensibilità al tempo
Trasferimento file	No	Variabile	No
Posta Elettronica	No	Variabile	No
Documenti Web	No	Variabile	No
Audio/Video in tempo reale	Si	Audio: 5Kbps – 1 Mbps Video: 10Kbps – 5MKbps	Si, centinaia di millisecondi
Audio/Video memorizzati	Si	Audio: 5Kbps – 1 Mbps Video: 10Kbps – 5MKbps	Si, pochi secondi
Videogame	Si	Fino a pochi Kbps	Si, centinaia di millisecondi
Messaggistica istantanea	No	Variabile	Si e no

Applicazione	Protocollo a Livello Applicativo	Protocollo di Trasporto
Posta Elettronica	SMTP (RFC 2821)	TCP
Accesso a terminali remoti	Telnet (RFC 854)	TCP
Web	HTTP (RFC 2616)	TCP
Trasferimento file	FTP (RFC 959)	TCP
Streaming multimediale	HTTP (es. YouTube), RTP (RFC 1889)	TCP o UDP
Telefonia internet	SIP, RTP, proprietario (es. Skype)	Tipicamente UDP

11 Applicazioni Web e HTTP

11.1 Terminologia

WEB Consiste di *oggetti* indirizzati da un **URL** (Uniform Resource Locator)

Pagine Web Solitamente formate da: *pagine WEB* (HTML, Javascript...) e diversi **oggetti referenziati** (altre pagine, immagini, script...)

Browser Lo user agent per il web, ad esempio: Chrome, Firefox, Netscape, Lynx

Web server Il server per il web, ad esempio: Apache, MS Internet Information Server

11.2 Uniform Resource Identifier

Una **URI** è una **forma generale per identificare una risorsa presente sulla rete** (IETF RFC 2396: *una Uniform Resource Identifier è una stringa compatta di caratteri usata per identificare una risorsa astratta o fisica*).

La sintassi di uno URI è stata progettata ponendo la **trascrivibilità globale** come uno degli obiettivi principali: utilizza caratteri da un **alfabeto molto limitato** (es. le lettere dell'alfabeto latino base, numeri e qualche carattere speciale).

Uno URI può essere **rappresentato in molti modi**, ad esempio: inchiostro su carta, pixel su schermo, sequenza di ottetti... L'**interpretazione di uno URI dipende soltanto dai caratteri utilizzati e non da come essi vengono rappresentati** nel protocollo di rete.

Uniform Uniformità della sintassi dell'identificatore, anche se i meccanismi per accedere alle risorse possono variare.

Resource Qualsiasi cosa abbia un'identità: documento, servizio, immagine, collezione di risorse...

Identifier Oggetto che può agire da riferimento verso qualcosa che ha identità

Esistono due tipi di URI:

URL Uniform Resource Locator: sottotipo di URI che identifica una risorsa attraverso il suo **meccanismo di accesso primario**, ad esempio la *posizione* nella rete.

Esempi:

URL `https://doi.org/10.1109/LCN.1988.10239`

URL `ftp://ftp.is.co.za/rfc/rfc1808.txt`

URL `https://www.apple.com/index.html`

URN Uniform Resource Name: sottotipo di URI che devono essere **globalmente univoci e persistenti**, anche quando la risorsa cessa di esistere o di essere disponibile.

Esempi:

URN `urn:oasis:names:specification:docbook:dtd:xml:4.1.2:`

URN `urn:doi:10.1109/LCN.1988.10239`

11.2.1 Sintassi

La sintassi di un URI è **organizzata gerarchicamente**, con le componenti **elencate in ordine decrescente** di importanza da sinistra a destra.

Una **URI assoluta** può essere formata da **quattro** componenti

```
<scheme>://<authority><path>?<query>
```

<scheme> Obbligatorio, schema per identificare la risorsa.

Lo URI scheme **definisce il namespace** dello URI, quindi potrebbe porre ulteriori vincoli su sintassi e semantica degli identificatori che usano quello schema. Nonostante molti URL scheme prendono il nome da protocolli, **questo non implica che l'unico modo di accedere la risorsa dello URL sia attraverso il protocollo specificato**.

<authority> Elemento gerarchico per richiamare un'authority così che la gestione del namespace definito sia delegato a quella authority. Il **nome di dominio** di un host o il suo **indirizzo IP** in notazione puntata decimale.
authority = [userinfo@]host[:port]

<path> Contiene dati specifici per l'authority (o lo schema) e **identifica la risorsa nel contesto** di quello schema e di quella autorità. Può consistere in una sequenza di segmenti.

<query> L'interrogazione o i dati da passare alla risorsa richiesta

Esempi:

```
foo://example.com:8042/over/there?name=ferret#nose

    scheme = foo
    authority = example.com:8042
    path = /over/there
    query = name=ferret
    fragment = nose

urn:example:animal:ferret:nose

    scheme = urn
    path = example:animal:ferret:nose

http://maps.google.it/maps/place?q=larco+bruno+pontecorvo+pisa&hl=it

    scheme = http
    authority = maps.google.it
    path = /maps/place
    query = q=larco+bruno+pontecorvo+pisa&hl=it
```

11.2.2 Assolute e Relative

Le URI possono essere assolute o relative.

URI Assoluta Identifica una risorsa **indipendentemente dal contesto** in cui è usata.

URI Relativa Informazioni per **identificare una risorsa in relazione ad un'altra URL** (è priva di **scheme e authority**). **Non viaggiano sulla rete**, sono interpretate dal browser in relazione al documento di partenza.

Esempio di URI relativa Sia `http://a/b/c/d;p?q` il documento di partenza, allora

```
g = http://a/b/c/g
/g = http://a/g
//g = http://g
?y = http://a/b/c/?y
#s = documento corrente#s
g;x?y#s = http://a/b/c/g;x?y#s
.. = http://a/b/
.../..../g = http://a/g
```

12 HTTP

Lo HTTP è usato dal 1990 come **protocollo di trasferimento per il World Wide Web**. Definito nel seguente modo (RFC 2068, RFC 2616): *protocollo di livello applicazione per sistemi di informazione distribuiti, collaborativi ed impermediati*.

Protocollo **generico**, **stateless** e **object-oriented** che può essere usato per molte attività, come name server e sistemi distribuiti di gestione oggetti, attraverso l'estensione dei suoi **request methods** (comandi). Una funzionalità dell'HTTP è la rappresentazione del tipo di dati, consentendo al sistema di essere **costruito indipendentemente dai dati che vengono trasferiti**.

12.1 HTTP URL

Lo schema `http` è usato per accedere alla risorsa attraverso il protocollo HTTP.

Sintassi La sintassi per un URL `http` è:

```
http_URL = http://host[:port][path]
```

Host: un dominio, hostname o indirizzo IP in forma decimale puntata di Internet.

Porta: un numero, se omessa viene usata la porta 80.

La risorsa è localizzata nel server in ascolto per connessioni TCP su quella porta di quell'host. Il path specifica la **Request-URI**.

12.2 Caratteristiche

Il protocollo HTTP è un protocollo **request/response**: la **connessione** viene **iniziatata dal client**, che invia un **messaggio di request** al quale il server risponde con una **response**.

In quanto **generico e stateless** le coppie **richiesta/risposta sono indipendenti**.

12.2.1 Modello

Il modello del protocollo HTTP è **client-server**:

Client: browser che richiede, riceve e visualizza oggetti web.

Stabilisce una connessione con il server e invia una **richiesta sotto forma di request-method, URI e versione di protocollo**, seguito da un messaggio.

Server: web server che invia oggetti in risposta ad una richiesta.

Accetta le connessioni e serve le richieste rispondendo con i dati richiesti.

12.2.2 Connessioni

Una **connessione** è un **circuito logico di livello trasporto** stabilito tra due programmi applicativi per comunicare tra loro.

Non-Persistent Connection http1.0: RFC 1945

Viene stabilita una **connessione TCP separata** per raggiungere ogni URL. **Aumenta il carico** sui server HTTP e **può causare congestioni su Internet** (questo perché, ad esempio, se vengono usate tante immagini allora si creano richieste multiple del solito server in un breve lasso di tempo).

Persistent Connection http1.1: RFC 2616

Se non è indicato altrimenti, il client può **assumere che il server manterrà una connessione persistente**.

Lo standard specifica un **meccanismo con il quale** un client o un server **può segnalare la chiusura di una connessione TCP** (il campo connection nell'header). Una volta che la chiusura viene segnalata, il cliente **non deve più mandare richieste** su quella connessione.

12.3 Esempio HTTP

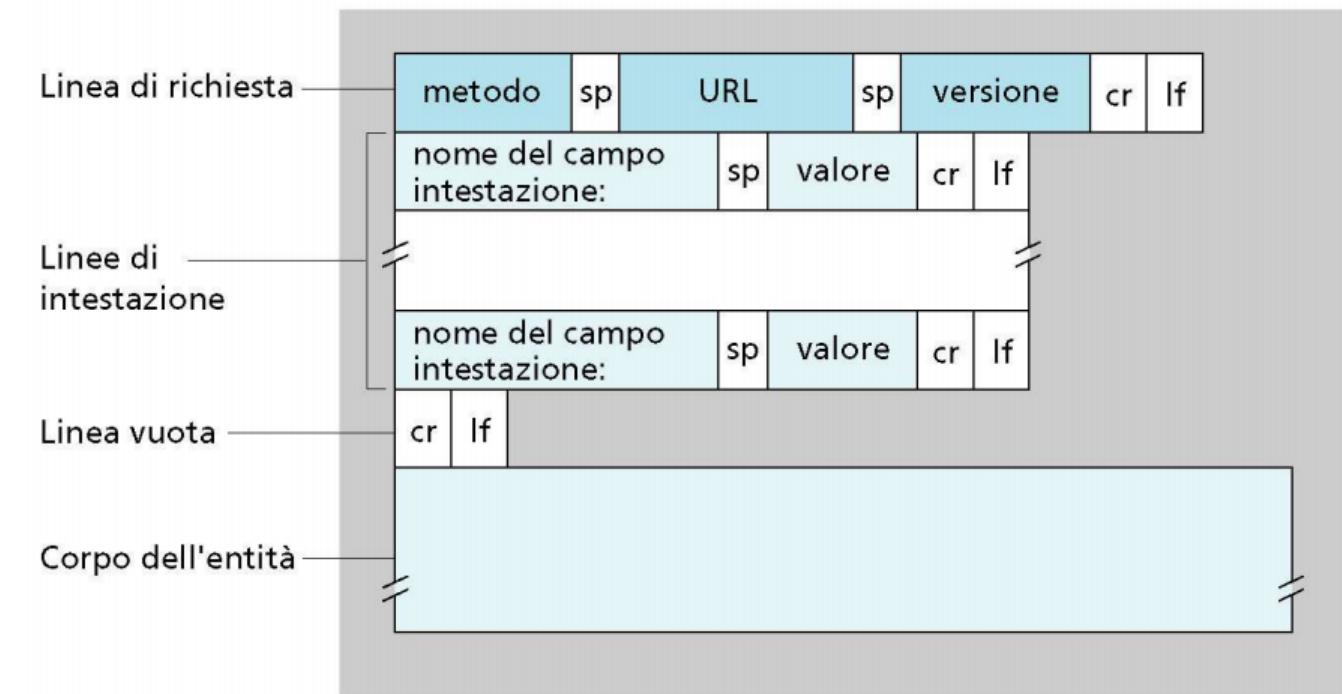
Vedi slide¹

12.4 Messaggi HTTP

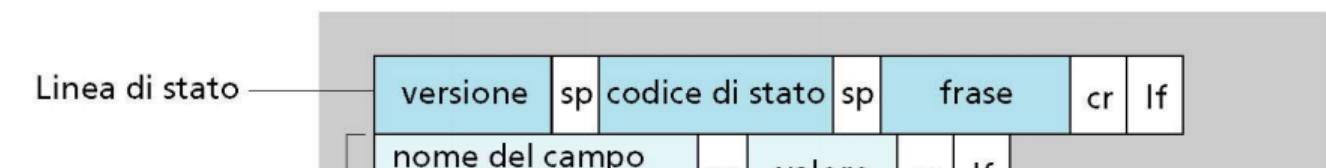
```
generic-message = start-line *message-header CRLF [message-body]
start-line = Request-Line | Status-Line
```

La **start-line** distingue **request** da **response**.

HTTP Request Message



HTTP Response Message



¹https://elearning.di.unipi.it/pluginfile.php/27477/mod_resource/content/2/L03_Applicativo_HTTP.pdf, slide 34

HTTP Request Request-Line *(general-header | request-header | entity-header) CRLF [message-body]

Ad esempio:

```
GET /pub/WWW/TheProject.html HTTP/1.1
Host: www.w3.org
Connection: close
User Agent: Mozilla/4.0
Accept-language: it
```

(Body)

HTTP Request Line Request-Line = Method SP Request-URI SP HTTP-Version CRLF
GET http://www.w3.org/pub/WW/TheProject.html HTTP/1.1

```
Method = GET
Request-URI = http://www.w3.org/pub/WW/TheProject.html
HTTP-Version = HTTP/1.1
Method = "OPTIONS" | "GET" | "HEAD" | "POST" | "PUT" | "DELETE" | "TRACE" | extension-method
```

Method: indica il **metodo** che deve essere eseguito sulla risorsa identificata dal **Request-URI**. Case sensitive.

HTTP-Version: indicare la versione del protocollo è pensato per consentire al mittente di indicare il formato di un messaggio e la sua capacità di capire il resto della comunicazione HTTP.

Le **URI** sono stringhe formattate in modo semplice che identificano una risorsa di rete.

12.5 Header

Gli **header** sono insiemi di coppie <nome:valore> che **specificano alcuni parametri** del messaggio trasmesso o ricevuto.

General header Relativi alla trasmissione

Data: data e ora di generazione del messaggio

Connection: consente al mittente di specificare le opzioni desiderate per quella particolare connessione. L'opzione "close" segnala che la connessione verrà chiusa al completamento della response

Transfer-encoding: indica quale (se presente) tipo di trasformazione è stata applicata al message body per trasferirlo correttamente dal mittente al destinatario (chunked, gzip...)

Cache Control

Public Indica che la response è cachable in qualsiasi cache

Private Indica che tutte o parti della response sono destinate ad un singolo utente e **non devono essere memorizzate** in una cache condivisa (shared cache). Una cache privata (non-shared) potrebbe memorizzare la response

no-cache Indica che tutte o parti della response **non devono essere memorizzate in nessuna cache**

```
general-header = Cache-Control | Connection | Date | Pragma | Transfer-Encoding | Upgrade
| Via
```

Questi header si applicano a **tutto il messaggio**. Esempi:

```
Date: Tue, 15 Nov 1994 08:12:31 GMT
Connection: close
Transfer-Encoding: chunked
```

Entity header Relativi all'entità trasmessa

Content-type, Content-length, data di scadenza...

12.5.1 Request header

Relativi alla richiesta

Chi fa la richiesta, a chi viene fatta, che tipo di caratteristiche è in grado di accettare il client, autorizzazione... consente al client di passare **informazioni aggiuntive** a proposito della richiesta o del client stesso al server. Questi campi agiscono come **modificatori di richiesta**, con **semantica equivalente a quella dei parametri** di un metodo.

```
request-header = Accept | Accept-Charset | Accept-Encoding | Accept-Language | Authorization  
| Proxy-Authorization | From | Host | If-Modified-Since | If-Unmodified-Since | If-Match |  
If-None-Match | If-Range | Max-Forwards | Range | Referer | User-Agent
```

Accept Specifica che tipi di media sono accettabili nella response. Parametro **q** per indicare un fattore di qualità relativo, default a 1.

Accept-Charset Indica il set di caratteri accettato per la risposta

Accept-Encoding Tipi di trasformazioni accettate (es. compressione)

```
Accept: text/plain;q=0.5, text/html, text/x-dvi;q=0.8, text/x-c  
Accept-Charset: iso-8859-5, unicode-1-1;q=0.8  
Accept-Encoding: compress, gzip
```

(Alcuni) metodi request:

OPTIONS Richiede solo le opzioni di comunicazione associate ad un URL o al server stesso (capacità, metodi esposti ecc.). Un esempio:

```
OPTIONS http://192.168.11.66/manual/index.html HTTP/1.1  
host: 192.168.11.66  
Connection: close
```

```
HTTP/1.1 200 OK  
Date: Sun, 14 May 2000 19:52:12 GMT  
Server: Apache/1.3.9 (Unix) (Red Hat/Linux)  
Content-Length: 0  
Allow: GET, HEAD, OPTIONS, TRACE  
Connection: close
```

GET Richiede il trasferimento di una risorsa identificata da un URL o le operazioni associate all'URL stessa. Un esempio:

```
GET http://192.168.11.66 HTTP/1.1  
host: 192.168.11.66  
Connection: close
```

Response

```
HTTP/1.1 200 OK  
Date: Sun, 14 May 2000 19:57:13 GMT  
Server: Apache/1.3.9 (Unix) (Red Hat/Linux)  
Last-Modified: Tue, 21 Sep 1999 14:46:36 GMT  
ETag: "f2fc-799-37e79a4c"  
Accept-Ranges: bytes  
Content-Length: 1945  
Connection: close  
Content-Type: text/html
```

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2  
Final//EN">  
<HTML>...
```

Sono possibili **GET condizionali e parziali**. Esempio di GET condizionale:

```
GET http://192.168.11.66 HTTP/1.1  
Host: 192.168.11.66  
If-Modified-Since: Tue, 21 Sep 1999  
14:46:36 GMT
```

Response:
HTTP/1.1 304 Not Modified
Date: Wed, 22 Sep 1999 15:06:36 GMT
Server: Apache/1.3.9 (Unix) (RedHat/Linux)

HEAD Simile al GET, ma il server **non trasferisce il body** nella response. Utile per controllare lo stato dei documenti (validità, modifiche...). Un esempio:

```
HEAD http://192.168.11.66 HTTP/1.1
host: 192.168.11.66
Connection: close
```

Response (notare la somiglianza con GET, esclusa la mancanza qua del body):

```
HTTP/1.1 200 OK
Date: Sun, 14 May 2000 20:02:41 GMT
Server: Apache/1.3.9 (Unix) (Red Hat/Linux)
Last-Modified: Tue, 21 Sep 1999 14:46:36 GMT
ETag: "f2fc-799-37e79a4c"
Accept-Ranges: bytes
Content-Length: 1945
Connection: close
Content-Type: text/html
```

POST Serve per **inviare dal client al server** informazioni inserite nel body del messaggio.
In teoria lo standard dice che il metodo POST è usato per richiedere che il server **accetti l'entità racchiusa nella richiesta come nuovo subordinato della risorsa identificata** dallo Request-URI nel Request-Line.
Nella pratica, la funzionalità effettiva del metodo POST è determinata dal server e solitamente dipende dalla Request-URI.

DELETE Il client **chiede di cancellare una risorsa identificata** dalla Request-URI.
Solitamente non attivo su server pubblici.

PUT Il client **chiede di creare/modificare una risorsa identificata** dalla Request-URI. Dopo posso usare una GET per recuperarla.
Solitamente non attivo su server pubblici.

Response header Nel messaggio di risposta
Server, autorizzazione richiesta...

Safe Methods Metodi che **non hanno effetti collaterali** (es. non modificano la risorsa): GET, HEAD, OPTIONS, TRACE

Idempotent Methods Metodi che **non hanno effetti ulteriori se vengono fatti N > 0 richieste identiche**: GET, HEAD, PUT, DELETE, OPTIONS, TRACE

12.6 HTTP Response

```
Response = Status-Line *( general-header | response-header | entity-header ) CRLF [message-body]
```

Un esempio:

```
HTTP/1.1 200 OK
Date: Sun, 14 May 2000 23:49:39 GMT
Server: Apache/1.3.9 (Unix) (Red Hat/Linux)
Last-Modified: Tue, 21 Sep 1999 14:46:36 GMT
```

Status-Line La prima linea del messaggio di risposta.

Status-Line = HTTP-Version SP Status-Code SP Reason-Phrase CRLF

Esempio: HTTP/1.1 200 OK

Status-Code Intero a 3 cifre, risultato del tentativo di comprendere e soddisfare la richiesta.

1xx: Informational - Request received, continuing process
2xx: Success - The action was successfully received, understood, and accepted
3xx: Redirection - Further action must be taken in order to complete the request
4xx: Client Error - The request contains bad syntax or cannot be fulfilled
5xx: Server Error - The server failed to fulfill an apparently valid request

<http://www.w3.org/Protocols/rfc2616/rfc2616-sec10.html>

Reason-Phrase Ha l'obiettivo di fornire una breve descrizione testuale dello Status-Code. Lo Status-Code è indirizzato ai computer mentre la Reason-Phrase è per gli umani.

12.6.1 Response Headers

Il campo response-header consente al server di passare ulteriori informazioni sulla response. Questi campi dell'header forniscono informazioni sul server e sull'accesso alla risorsa identificata dallo Request-URI.

response-header = Age | Location | Proxy-Authenticate | Public | Retry-After | Server | Vary | Warning | WWW-Authenticate

Esempio:

```
Age: 150 // età del doc. se tramite Proxy
Location: http://www.w3.org/pub/WWW/People.html
Server: CERN/3.0 libwww/2.17
```

Age Una stima in secondi del tempo passato dalla generazione della risposta dal server di origine

Location Usato per reindirizzare il ricevente verso una destinazione diversa dalla Request-URI per il completamento della richiesta o l'identificazione di una nuova risorsa

Server Informazioni sul software usato dal server d'origine per gestire la richiesta

12.7 Negoziazione del contenuto

Le risorse possono essere **disponibili in multiple rappresentazioni**, ad esempio più lingue, formati, dimensioni e risoluzioni, o variare in altri modi ancora. La **content negotiation** è il meccanismo usato per **selezionare l'appropriata rappresentazione** quando si serve una richiesta. Ogni entità è costituita da un **entity body** e da una serie di **entity headers** che ne definiscono il contenuto e le proprietà. Gli entity header sono **informazioni sulle informazioni**, cioè **metadati**.

12.7.1 Entity Headers

entity-header = Allow | Content-Base | Content-Encoding | Content-Language | Content-Length | Content-Location | Content-MD5 | Content-Range | Content-Type | ETag | Expires | Last-Modified | extension-header

Content-Base URI assoluta da usare per risolvere le URL relative contenute nell'entity-body

Content-Encoding Codifica dell'entity-body (es. gzip)

Content-Language Lingua dell'entity-body (es. en, it)

Content-Type Tipo dell'entity-body (es. text/html)

Expires Valore temporale dell'entity-body (utile nel caching)

Last-Modified Data dell'ultima modifica sul server (utile nel caching)

13 Web Caching

L'obiettivo è **soddisfare una richiesta** del cliente **senza contattare il server**. Si memorizzano copie temporanee delle risorse web (es. pagine HTML e immagini) e si servono al client per ridurre l'uso di risorse (es. banda e workload sul server), diminuendo anche il tempo di risposta.

User Agent Cache lo user agent (il browser) mantiene una **copia delle risorse visitate dall'utente**.

Proxy Cache Il proxy intercetta il traffico e **mette in cache le risposte**. Le successive richieste alla stessa Request-URI **possono essere servite dal proxy senza inoltrare la richiesta** al server.

Proxy Programma intermediario che agisce sia da server che da client, con l'obiettivo di fare richieste per conto di altri client. Le richieste sono servite internamente o passandole oltre, anche traducendole, ad altri server.

14 Cookies

L'HTTP è **stateless**, per cui non mantiene info sui client. Come posso riconoscere il cliente di un'applicazione web (es. Amazon)? Come posso realizzare applicazioni web con stato (es. carrello della spesa)? Ricordiamo che **tipicamente l'utente si connette ogni volta con un indirizzo IP e porta diversi**.

Soluzione: numerare i client e obbligarli a farsi riconoscere ogni volta presentando un cookie.

Funzionamento Il client C invia al server S una normale richiesta HTTP.

Il server invia la normale risposta + una linea **Set-Cookie: 1678453**

Il client memorizza il cookie in un file associato a S, e aggiunge una linea **Cookie: 1678453** a tutte le successive richieste verso quel sito.

Il server confronta il cookie presentato con l'informazione che ha associato a quel cookie.

Utilizzi I cookie vengono utilizzati per:

- Autenticazione
- Ricordare il profilo utente e le scelte precedenti (alla carta-socio)
- Creare sessioni sopra un protocollo stateless (es. carrelli della spesa)

Non accettare dolci dagli sconosciuti: cookiecentral.com

15 Telnet

TERminaL NETwork Protocollo di terminale remoto che permette l'**uso interattivo** di macchine remote: **accesso remoto, accesso multiplo ad un singolo computer**.

Realizza coppie client-server generiche per login remoto, non specializzate per tipo di applicativo. Invece di offrire server specializzati per servizi interattivi, l'approccio consiste nel permettere all'utente di **effettuare una sessione login nella macchina remota** e quindi **inviare i comandi**. Tramite il login remoto gli utenti hanno accesso ai comandi e ai programmi disponibili nella macchina remota.

Puoi mandare comandi attraverso il programma Telnet e verranno eseguiti come se fossi a scriverli direttamente sulla server console. Questo ti consente di controllare il server e comunicare con gli altri server sulla rete
Non è un compito facile, per realizzarlo il Telnet:

Maschera sia la rete che i S.O.

Utilizza un'**interfaccia minima ma veloce**, tipicamente a caratteri

15.1 Introduzione

Telnet permette ad un utente su una macchina di **stabilire una connessione con un login su server remoto**. In seguito, passa la **battute dei tasti** della macchina locale **alla macchina remota**: i comandi vengono **eseguiti come se fossero stati battuti al terminale della macchina remota**. Dopo di che, l'**output** della macchina remota **viene trasportato al terminale utente**.

Questo è un **servizio trasparente**: il terminale dell'utente *sembra* essere **connesso direttamente** alla macchina remota.

Il modello di Telnet include:

Server che **accetta** le richieste

Client che **effettua** le richieste

Il programma Telnet che svolge due funzioni:

Interagisce col terminale utente sull'host locale

Scambia messaggi con il Telnet server

15.2 Protocollo Telnet

RFC 854 Comunicazione generale, bidirezionale e orientata a blocchi di 8bit. L'obiettivo primario è di fornire un metodo standard per **interfacciare dispositivi terminali e processi terminal-oriented tra loro**.

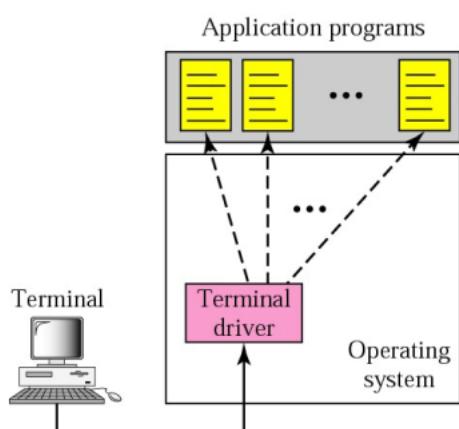
Usa il TCP, con una connessione TCP persistente per tutta la durata della sessione di login, sulla porta 23 del server.

Client **stabilisce una connessione TCP** con il server

Client **accetta le battute di tasti** sul terminale e **le invia al server**.

Accetta i caratteri che il server manda indietro e **li visualizza** sul terminale utente.

Server **accetta la connessione TCP** e trasmette i dati al S.O. locale.



In **Local Login** il S.O. assume che gli input ad un processo vengano forniti dallo standard input (tastiera) e che gli output siano inviati allo standard output (monitor).



In **Remote Login** lo **Pseudo Terminal Driver** è l'entry point del S.O. che consente di trasferire caratteri ad un processo **come se provenissero dal terminale**. Ha il compito di **accettare i caratteri** dal server e **trasmetteri al S.O.** che consegnerà all'applicazione opportuna.

15.3 NVT

Il Telnet deve **poder operare con il numero massimo di sistemi**, quindi gestire **dettagli di S.O. eterogenei** che possono differire per:

Set di codifica dei caratteri

Lunghezza della linea e della pagina

Tasti funzione individuati da diverse sequenze di caratteri (**escape sequence**)

Es. diverse combinazioni per interrompere un processo(CTR+C, ESC), caratteri ASCII diversi per la terminazione di righe di testo.

Per risolvere questo problema si definisce **un ambiente virtuale**. Sulla rete si considera un **unico terminale standard** e in corrispondenza di ogni stazione di lavoro si effettuano le **conversioni da terminale locale a terminale virtuale e viceversa**.

Network Virtual Terminal Telnet assume che sui due host sia in esecuzione un **Network Virtual Terminal**, la connessione TCP è stabilita tra i due terminali NVT.

L'NVT è un dispositivo *immaginario* che **fornisce una rappresentazione astratta di un terminale canonico**. Gli host, sia client che server, traducono le loro caratteristiche locali così da **apparire esternamente come un NVT** e assumo che l'host remoto sia un NVT.

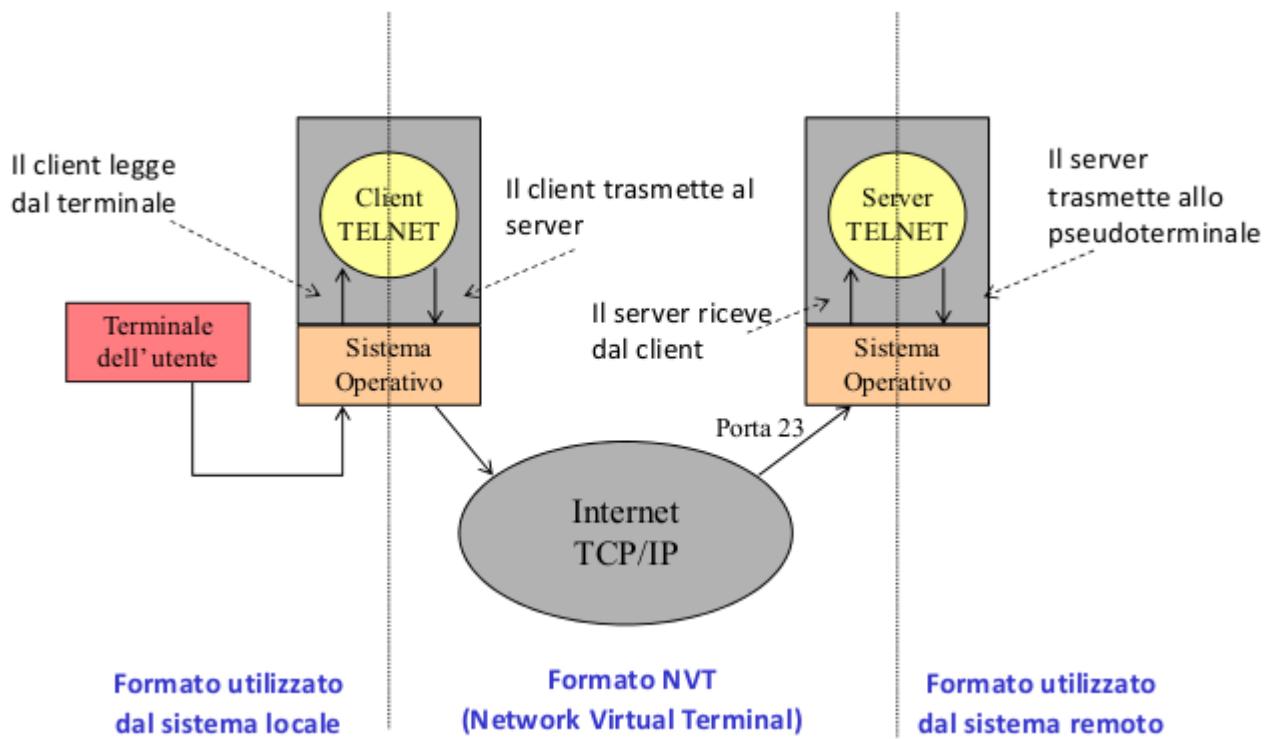
Il Network Virtual Terminal definisce un **set di caratteri e di comandi universale**, che permette di:

Trasformare il set di caratteri in uso localmente **in un set di caratteri universale** (lettere accentate, tasti freccia, backspace...), includendo anche i caratteri di controllo più importanti (break...)

Inviare i caratteri di controllo in maniera **privilegiata** (meccanismo URGENT del TCP)



15.4 Architettura



15.5 Funzionamento

Rispetto alla esecuzione in locale, sono aggiunti una **serie di intermediari**:

Il client Telnet **trasforma in NVT** ed usa il S.O. per inviare sulla rete;
la rete ed il S.O. server portano i dati al server Telnet;
il server Telnet **traduce da NVT a S.O.** remoto;
il S.O. remoto **esegue** il dovuto;
si percorre il cammino inverso con gli stessi soggetti.

Un esempio:

1. L'utente digita INVIO sulla tastiera
2. Il S.O. passa al client il carattere CR
3. Il client converte CR in CR-LF e lo invia sulla connessione TCP
4. Il server riceve CR-LF, lo converte in LF e lo passa allo pseudoterminale
5. Lo pseudoterminale passa all'applicazione LF
6. L'applicazione esegue l'operazione di INVIO

Formato I terminali NVT si scambiano i dati in **formato 7bit US-ASCII**. Ogni carattere è inviato come un ottetto con il primo bit settato a 0. I byte con il bit più significativo a 1 sono usati per le **sequenze di comandi**. I comandi (es. end-of-line trasmesso come la sequenza CR-LF) iniziano con un ottetto speciale (**Interpret as Command** o **IAC**) di 1 → **Inband Signalling** (comandi e dati sulla stessa connessione).

I messaggi di controllo iniziali sono usati per scambiare informazioni sulle caratteristiche degli host (**Telnet Option Negotiation**).

Comandi Qualche esempio

Comando	Codifica Decimale	Significato
IAC	255	Interpreta come comando l'ottetto successivo
EL	248	Erase Line
EC	247	Erase Character
IP	244	Interrupt Process
EOR	239	End of Record

NVT conviene? Suppongo di voler far interoperare N sistemi.

Senza usare NVT, ho bisogno di scrivere N-1 client per ogni sistema, e N server (uno per sistema) = **devo scrivere N(N-1)+N applicazioni**.

Usando NVT invece devo scrivere soltanto N server e N client, cioè **2N applicativi**.
Per N > 2 conviene NVT.

16 SSH

Poiché Telnet passa tutto in chiaro, **anche le password**, con il tempo si è resa necessaria una maggiore sicurezza.

Secure Shell SSH è un'applicazione nata per **sostituire Telnet e risolvere i suoi problemi di sicurezza**. Facilita la comunicazione sicura tra client e server e permette la login remota, resa sicura attraverso **tecniche di cifratura**. Nella realtà SSH offre **funzionalità molto superiori** a quelle di Telnet².



SSH-TRANSPORT Realizza la **connessione sicura** tra due host:

Autenticazione del server

Negoziazione degli algoritmi di cifratura

Scambio di chiavi

SSH-AUTHENTICATION Meccanismi per autenticare l'utente

SSH-CONNECTION Sessioni di login remoto, tunneling

17 TCP Port Forwarding

Port Forwarding L'inoltro delle porte è un meccanismo che permette di creare un **canale di comunicazione sicuro** attraverso il quale veicolare qualsiasi tipo di connessione TCP.

Viene **creato un canale di comunicazione cifrato tra la porta all'indirizzo remoto** al quale ci si vuole collegare **ed una porta locale libera**. In questo modo le applicazioni punteranno il collegamento alla porta locale e la connessione verrà **inoltrata automaticamente** all'host remoto tramite un canale sicuro.



`ssh -L 123:localhost:456 remotehost`

-L specifica che la data porta sull'host locale è da inoltrare alla data porta sull'host remoto.

`<porta locale>:<host>:<porta remota di host>`

²Vedi su Forouzan

18 FTP

Il **File Transfer Protocol** (RFC 959) è usato per il **trasferimento di file da/a un host remoto**. Segue il modello client server:

Client: il lato che **chiede** il trasferimento

Server: l'host remoto

Standard Nelle reti TCP/IP lo FTP è lo **standard per il trasferimento di file**. Questo è un servizio **diverso** dall'accesso condiviso online (che è un **accesso simultaneo** da parte di più programmi **ad un singolo file**). L'FTP fornisce anche **funzionalità aggiuntive** oltre al semplice trasferimento di file:

Accesso interattivo: l'utente può **navigare e cambiare/modificare** l'albero di directory nel file system remoto

Specifiche del formato dei dati da trasferire (es. file di testo o file binari)

Autenticazione: il client può specificare username e password

18.1 Modello FTP

L'FTP ha **due tipi di connessione**:

Control connection: scambio di comandi e risposte tra client e server, segue il protocollo Telnet

Data connection: connessione su cui i dati sono trasferiti con modi e tipi specificati. I dati trasferiti possono essere **parte** di file, **un file** o **un set** di file.

18.1.1 Connessione di Controllo

Il client FTP **contatta il server FTP alla porta 21** usando il TCP come protocollo di trasporto. Il client **ottiene l'autorizzazione** sulla connessione di controllo (es. cambio directory, invio file ecc...).

La **connessione è persistente**.

18.1.2 Connessione Dati

Quando il server riceve un comando per trasferire file (da o verso il client), **apre una connessione TCP** con il client.

Active Mode: una connessione per ciascun trasferimento e dopo il trasferimento di un file il server chiude la connessione.



18.2 Altri Dettagli

Quando un client attiva la connessione di controllo con il server usa un **numero di porta assegnato localmente in modo casuale** e contatta il server ad una porta nota, cioè la 21

FTP usa la connessione di controllo per permettere a client e server di **coordinare l'uso delle porte assegnate dinamicamente** per il trasferimento dati

La connessione di controllo FTP si basa sul protocollo Telnet

18.2.1 Due Modalità

Per creare la connessione TCP per il trasferimento dati sono possibili due modalità:

Active Mode: vedi sopra, una connessione per ciascun trasferimento e dopo il trasferimento di un file il server chiude la connessione. Il server deve conoscere il numero di porta lato client (glielo comunica)

Passive Mode: il **client ottiene un numero di porta dal server** (porta 20, non necessariamente). Il server non deve accettare connessioni da un processo arbitrario.

Da RFC 959 – PASV: questo comando richiede che il server-DTP "ascolti" su una porta dati (che non è quella predefinita) e di aspettare una connessione piuttosto che iniziare una alla ricezione del comando di trasferimento. La risposta a questo comando include l'host e l'indirizzo e porta su cui il server sta ascoltando.

DTP: Data Transfer Process

PI: Protocol Interpreter



I dispositivi dove risiedono client e server FTP sono diversi: S.O., diverse strutture per gestire i file, diversi formati dei file...

Per effettuare il trasferimento di file, il client deve **definire il tipo di file, la struttura dati e la modalità di trasmissione** al fine di risolvere i problemi di eterogeneità tra client e server.

Il trasferimento file viene **preparato attraverso uno scambio di informazioni lungo la connessione di controllo**

La comunicazione sulla connessione di controllo avviene per mezzo di caratteri con una codifica standard NVT ASCII, sia per i comandi che per le risposte

18.2.2 Stateful

FTP è un protocollo stateful. Il server deve **tenere traccia dello stato** dell'utente: tra le altre cose anche della connessione di controllo associata ad un account e della directory attuale.

18.2.3 Modalità di trasmissione

Stream mode: FTP invia i dati a TCP con un **flusso continuo di bit**

Block mode: FTP invia i dati a TCP **suddivisi in blocchi**, ognuno dei quali **preceduto da un header**

Compressed mode: si trasmette il **file compresso**

Comandi di Controllo

USER username

PASS password

LIST, elenca i file della directory corrente

NLST, richiede l'elenco di file e directory (ls)

RETR filename, recupera (get) un file dalla directory corrente

STOR filename, memorizza (put) un file nell'host remoto

ABOR, interrompe l'ultimo comando ed i trasferimenti in corso

PORT, indirizzo e numero di porta del client

SYST, il server restituisce il tipo di sistema

QUIT, chiude la connessione

Codici di Ritorno

I codici di ritorno sono composti da un codice di stato e da un'espressione, come in HTTP:

331 Username OK, password required

425 Can't open data connection

452 Error writing file

200 Comando OK

125 Data connection already open, transfer starting

225 Data connection open

225 Closing data connection. Requested file action succesful (es. trasferimento file o ABOR)

426 Connection closed, transfer aborted

227 Entering passive mode

18.3 Anonymous FTP

Server che supportano connessioni FTP senza autenticazione, spesso consentendo operazioni limitate.
ftp.ed.ac.uk, user: ftp e password la mail.

19 DNS

Identificare il processo Ogni processo di livello applicativo ha necessità di **individuare il processo omologo** con il quale vuole comunicare. Il processo omologo **risiede su una** particolare **macchina remota** anch'essa da **individuare** (sappiamo che usa lo stesso protocollo).

Nome Uno **nome identifica un oggetto**. Consiste in una sequenza di caratteri scelti da un alfabeto finito, scelta per scopo mnemonico.

Nome significativo di alto livello (alfanumerico), vero e proprio identificativo di livello applicativo.

Indirizzo Un **indirizzo identifica dove tale oggetto è situato**.

Host di internet usano indirizzi IP (32bit) per instradare i datagrammi (livello di rete). Il formato è scelto per garantire l'efficienza dell'instradamento.

Disaccoppiamento: ad un nome possono essere associati più indirizzi.

Come associare indirizzo IP e nome?

19.1 Motivazioni

Inizi All'inizio l'associazione fra nomi logici e indirizzi IP era **statica**: tutti i nomi logici ed i relativi indirizzi erano contenuti in un file (**host file**) periodicamente aggiornato da un server ufficiale. Questo approccio è infattibile nella rete internet attuale:

Non è possibile che **ogni host mantenga una copia aggiornata** dell'elenco. Questo per dimensioni dell'elenco, per i volumi di traffico per trasferire i cambiamenti...

Non è possibile **centralizzare un elenco del genere**: si avrebbe così un unico punto di fallimento, un gigantesco volume di traffico sul server...

Questa soluzione non è scalabile.

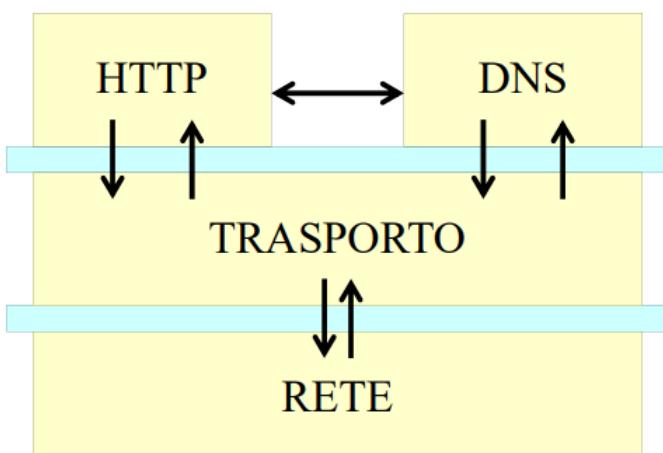
Per questo si utilizza il **Domain Name System**

19.2 Struttura

Il DNS è **posizionato nel livello applicativo**: gira sui terminali a paradigma client-server. Si affida al sottostante protocollo di trasporto punto-punto per trasferire i messaggi tra i terminali.

Non interagisce direttamente con utenti

La complessità è spostata alle estremità della rete



Il DNS è un meccanismo che deve:

Specificare la sintassi dei nomi e le regole per gestirli

Consentire la conversione nomi → indirizzi e viceversa

Il DNS è costituito essenzialmente da:

Schema di assegnazione dei nomi, gerarchico e basato su domini

Database distribuito contenente i **nomi** e le **corrispondenze** con gli indirizzi

Protocollo per la distribuzione delle informazioni sui nomi tra i vari **name server**

19.3 Servizi

I servizi offerti dal DNS sono molteplici:

Risoluzione di nomi di alto livello (**hostname**) in indirizzi IP

Host aliasing: un host può avere più nomi, solitamente il nome canonico + sinonimi.

Esempio: `realy1.west-coast.enterprise` può avere due alias `enterprise.com` e `www.enterprise.com`.

Mail Server aliasing: ci possono essere domain name identici per mail server e web server

Distribuzione del carico tra vari server replicati.

Ad un hostname canonico possono corrispondere diversi indirizzi IP. La lista di indirizzi viene ordinata in modo diverso in ogni risposta alla richiesta di risoluzione del nome, così che ogni server replicato possa essere scelto con uguale probabilità, distribuendo così efficientemente le richieste.

19.4 Spazio dei nomi

Dato che internet è una rete di proporzioni enormi, si presentano i problemi di **identificazione delle macchine** e **intradamento dei pacchetti**. Si adotta un **approccio stratificato** poiché è un sistema complesso.

19.4.1 Indirizzi

In internet esistono più indirizzi

Indirizzi **MAC**, quello della scheda di rete.

Soltanente prefissato

Indirizzo di rete (IP)

Esempio: `150.217.8.21`

Assegnato dal gestore di rete in base al tipo di rete a cui si appartiene (classe di sottorete)

Indirizzo di trasporto

Coppia <indirizzo IP, porta>

La porta è scelta a livello applicativo con regole appropriate

Indirizzo alfanumerico

Esempio: `medialab.det.unifi.it`

Libero, basta che sia mappato in un NameServer

19.4.2 Nomi

Lo **spazio dei nomi** deve permettere di **identificare in modo univoco** un host.

Struttura flat: sequenza di caratteri senza alcuna ulteriore struttura. Poco applicabile.

Struttura gerarchica: un nome è **costituito da diverse parti**.

Requisiti per la partizione dello spazio dei nomi:

Conversione efficiente

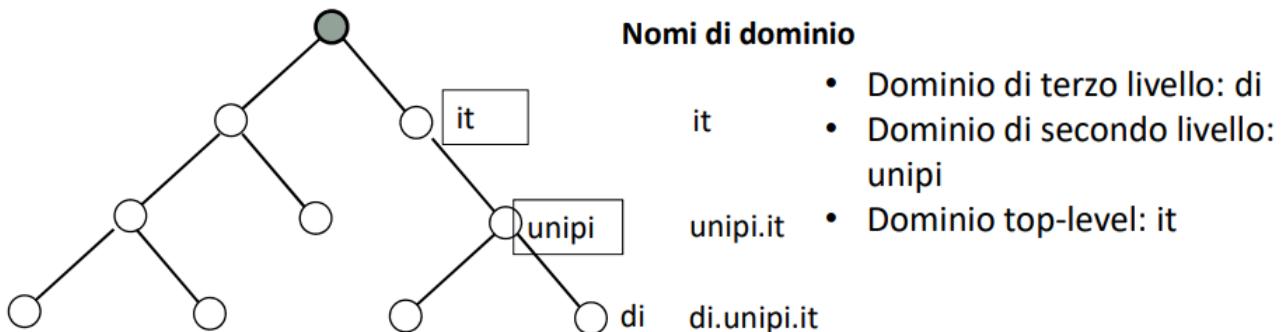
Controllo decentralizzato dell'assegnazione dei nomi.

Delega dell'autorità per le varie parti dello spazio dei nomi e **distribuzione della responsabilità** della conversione tra nomi e indirizzi

Vantaggi e svantaggi

- Minore velocità
- + Maggiore flessibilità
- + Riconfigurazione più veloce
- + Possibilità di aggiornamenti decentrati

Nomi di dominio Spazio dei nomi con **struttura gerarchica**. I nomi hanno una **struttura ad albero** con un numero di livelli variabile. **Ogni nodo definisce un livello gerarchico** ed è individuato da un'etichetta (max 63 caratteri). Alla radice è associata un'etichetta vuota.



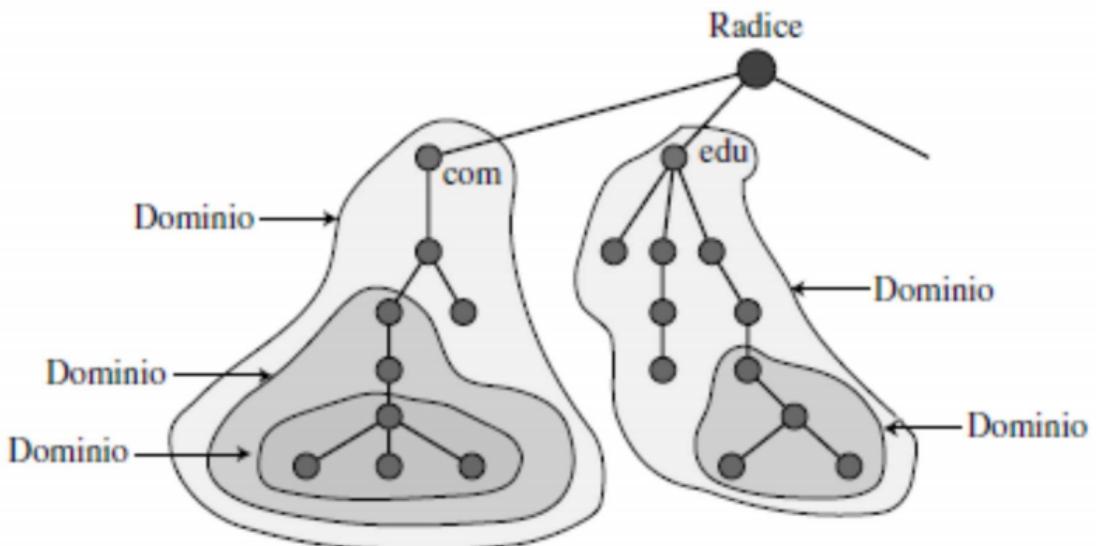
Ogni nodo dell'albero ha un nome di dominio, ovvero una **sequenza di etichette separate da punti** ovvero il cammino foglia → radice.

Dominio Sottoalbero nello spazio di nomi di dominio che viene identificato dal nome di dominio del nodo in cima al sottoalbero.

Può essere suddiviso in ulteriori domini, detti **sottodomini**.

In Internet Su internet i nomi gerarchici delle macchine sono **assegnati in base alla struttura delle organizzazioni** che ottengono l'autorità per porzioni dello spazio dei nomi. La struttura gerarchica permette **autonomia nella scelta dei nomi all'interno di un dominio** perché l'univocità è comunque garantita.
Ad es. server1.di.unipi.it e server1.cs.cornell.edu sono due nomi diversi.

Internet è divisa in diverse centinaia di domini ognuno dei quali **partizionato in sottodomini** e così via.



19.4.3 Top-Level Domains

Internet Assigned Number Authority IANA – iana.org

com Organizzazioni commerciali

edu Istituti d'istruzione (università, scuole...)

mil Gruppi militari

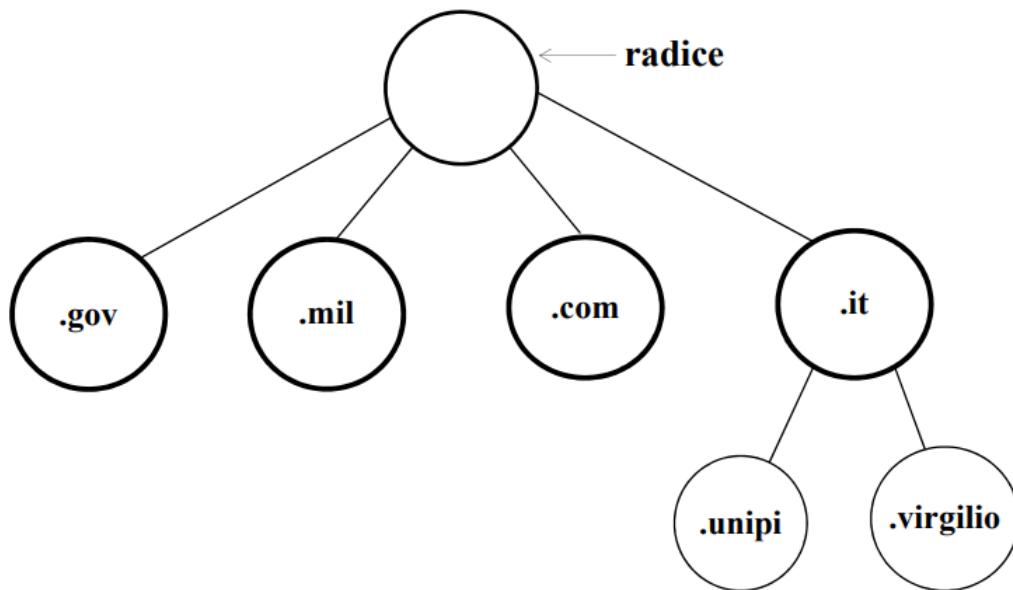
gov Istituzioni governative

net Principali centri di supporto alla rete

org Organizzazioni diverse dalle precedenti

Codici geografici, schema geografico per nazioni

Es. ir, uk, us, fr...



19.4.4 Struttura di un nome alfanumerico

mmedia5.di.unipi.it

mmedia5 nome locale della macchina (etichetta **più specifica**)

di.unipi.it nome del dominio (.it è l'etichetta **meno specifica**)

La parte dominio **può essere ulteriormente suddivisa**, creando così una **struttura logica gerarchica** (di nome locale, unipi.it dominio e così via...)

19.5 Conversione

Indirizzi IP Gli indirizzi IP sono **interi a 32bit**. Vengono rappresentati nella **Decimal Dotted Notation** che divide l'indirizzo IP in 4 **ottetti**, ovvero 4 gruppi di 8bit. Ad es. 150.217.8.21.

Gerarchia L'indirizzo IP è strutturato in una gerarchia a due soli livelli.

Indirizzi alfanumerici Gli indirizzi alfanumerici sono del tipo "medialab.di.unipi.it", sono **logici**, gerarchici e **NON indicano in assoluto la locazione geografica** di un host.

Vengono convertiti da un **Domain Name Server** in un **indirizzo IP**, eventualmente usando un sistema ricorsivo di ricerca.

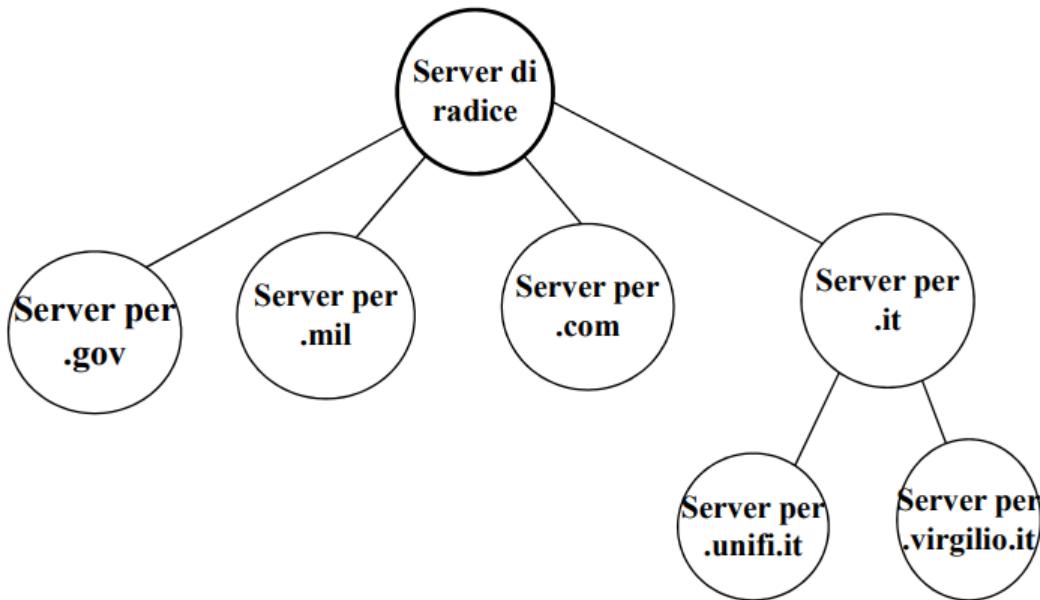
Database DNS Il database DNS è **distribuito** ed implementato in una gerarchia di più name servers.

Protocollo Il protocollo per la **risoluzione dei nomi alfanumerici** in indirizzi IP è dello **strato applicativo**.

NB Una funzione fondamentale di internet è implementata come protocollo dello strato applicazione → una parte della complessità della rete è **gestita alle estremità** della rete stessa.

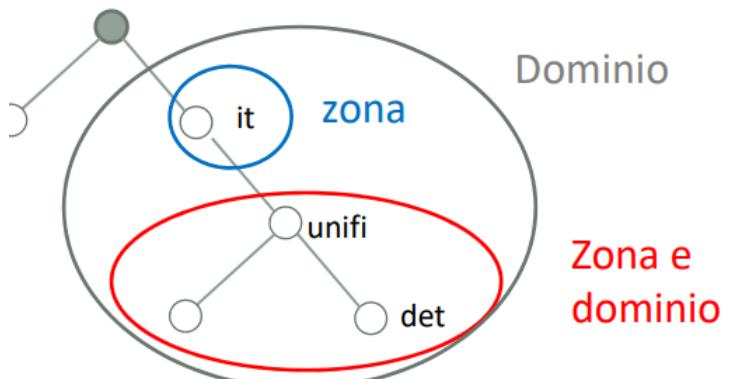
19.5.1 Name Servers

Un **Name Server** è un **programma che gestisce la conversione da nome di dominio ad indirizzo IP**. I Name Server sono strutturati gerarchicamente.



Spesso i DNS sono **accorpati e duplicati** (**sicurezza**). Inoltre per diminuire il traffico di rete ed il carico dei DNS, ogni server usa **una cache** per le ultime richieste espletate.

Informazioni Le informazioni sui domini sono ripartite su più server



Zona Una zona è una **regione di cui è responsabile un name server**, tipicamente una parte contigua dell'albero. Zona e dominio non necessariamente coincidono.

Il server **immagazzina** le **informazioni relative alla propria zona**, inclusi i riferimenti ai server dei domini di livello inferiore.

19.5.2 Root Name Servers

I **root name servers** vengono contattati dai **local name servers** che non possono risolvere un determinato nome.
Il root name server:

Contatta gli **authoritative name server** se quella traduzione non è nota

Recupera la traduzione

Invia la traduzione al local name server

Ce ne sono centinaia in tutto il mondo.



root-servers.org

19.5.3 Gerarchia dei server

Root Name Server, server radice

Top-Level Domain Server: si occupano dei domini .com, .org, .edu...
Ad es. Network Solution gestisce i TLD server per il dominio .com

Authoritative Name Server, server di competenza

Per un host archivia l'indirizzo IP di quello stesso host. Capace di risolvere la traduzione name/address per quell'hostname.

Ogni organizzazione dotata di host internet pubblicamente accessibili (es. web server, server di posta) deve **fornire i record DNS di pubblico dominio** che mappano i nomi di tali host in indirizzi IP (server mantenuti dall'organizzazione o ISP).

Per una certa zona ci possono essere **server primari** e **secondari**.

Server primari mantengono il file di zona

Server secondari ricevono il file di zona e offrono il servizio di traduzione

Local Name Server

Non appartengono strettamente alla gerarchia dei server. Ogni ISP (Università, Società, ISP) ha il **proprio (default) name server locale**. Le query DNS passano prima dal name server locale.

19.6 Risoluzione dei nomi

Esempio: il client vuole l'IP di `www.amazon.com`. Prima approssimazione:

Client **interroga il server radice** per trovare il server DNS `com`

Client **interroga il server DNS `com`** per ottenere il server DNS `amazon.com`

Client **interraga il server DNS `amazon.com`** per ottenere l'indirizzo IP di `www.amazon.com`

Query Ricorsiva

Host `www.tintin.fr` cerca l'IP di `www.topolino.it`.

Contatta il suo DNS locale `dns.tintin.fr`

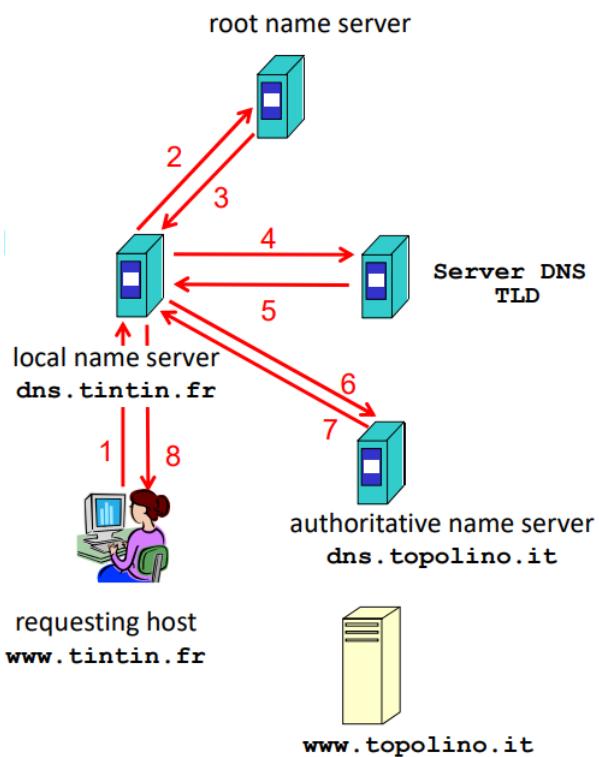
`dns.tintin.fr` contatta il root name server se necessario

Il root name server contatta l'autoritative name server `dns.topolino.it` se è necessario

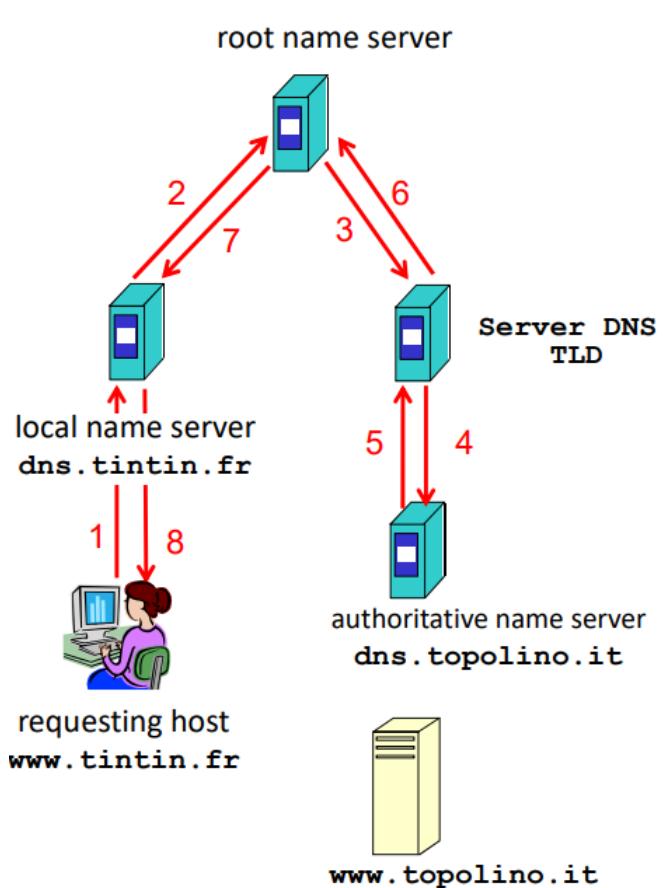
Il local name server ha richiesto una conversione completa

Query Iterativa Le risposte vengono restituite direttamente al client, cioè i name server rispondono con il riferimento al name server da contattare.

Query Iterativa



Query Ricorsiva



19.7 Caching e Aggiornamento Record

Una volta che un name server ha appreso un'associazione, la mette nella **cache**. I record nella cache vengono cancellati dopo un certo tempo (timeout, TTL). L'utilizzo della cache **migliora il ritardo e riduce il numero di messaggi DNS**.

I meccanismi di update/notifica sono descritti nella RFC 2136.

19.8 Record DNS

Il DNS è un database distribuito di "resource records" (RR).

Formato RR: (name, value, type, TTL)

TTL: quando la risorsa dovrà essere rimossa dalla cache

I significati di **name** e **value** dipendono da **type**

```
type = A  
name = hostname  
value = indirizzo IP  
  
type = CNAME  
name = hostname (sinonimo)  
value = nome canonico dell'host  
  
type = NS  
name = dominio (es. unifi.it)  
value = hostname dell'autoritative name server per quel dominio  
  
type = MX  
name = nome di dominio  
value = nome canonico del server di posta associato a name  
  
Ci sono altri type...
```

Esempi di record DNS

```
(hostname, indirizzoip, A, ...)  
(www.cnn.com, 157.166.224.25, A, ...)  
(www.cnn.com, 157.166.224.26, A, ...)  
  
(dominio, nomeDiAuthoritativeServerPerDominio, NS, ...)  
(cli.di.unipi.it, nameserver.cli.di.unipi.it*, NS, ...)  
[* ANSWER conterrà anche (nameserver.cli.di.unipi.it, 131.114.120.2,A, ...)]  
  
(alias, hostnameMailServerConTaleAlias, MX, ...)  
(cli.di.unipi.it, mailserver.cli.di.unipi.it*, MX, ...)  
[* ANSWER conterrà anche (mailserver.cli.di.unipi.it, 131.114.11.39, A, ...)]
```

19.9 Messaggi DNS

Query e risposte hanno lo stesso formato.

Le query contengono **QName** e **QType** nella sezione "question".

Le risposte contengono 0 o più RR nella sezione "answer" e 0 o più RR nella sezione "additional".

Per esempio:

```
dns.poly.edu —> TLD server  
Question: QName=gaia.cs.mass.edu QType=A  
  
dns.poly.edu —< TLD server  
Answer: (umass.edu, dns.umass.edu, NS)  
Additional: (dns.umass.edu, 128.115.40.41, A)
```

Header dei messaggi DNS

Identification: 16 bit. Identification usato dalle query, la reply ad una query usa lo stesso identification.

Flags

Query (0) o reply (1)

Recursion desired

Recursion available

Reply is authoritative

Il protocollo DNS usa tipicamente **UDP sulla porta 53 per le query DNS** (dimensioni del messaggio inferiori a 512B). Possibile usare TCP, ad esempio per trasferire i dati tra DNS server – Zone Transfers)



Domande: campi per il nome richiesto ed il tipo di domanda

Risposte: RR nella risposta della domanda

Competenza: record per i server di competenza

Informazioni aggiuntive

Questions

0	31
Nome di dominio dell'interrogazione	
Tipo di interrogazione	Classe di interrogazione
...	

Tipo di Interrogazione: tipo di domanda (nome di una macchina o indirizzo di posta)

Il **client** riempie la sezione di **domanda**

Il **server**, nel proprio messaggio di **risposta**:

ricopia la sezione di **domanda**

riempie le altre sezioni (risposta, autorità, altre info) con RR

Le **RFC di riferimento** sono:

RFC 1034: DOMAIN NAMES – CONCEPTS AND FACILITIES
Definisce i seguenti tipi di record:

A indirizzo dell'host

CNAME nome canonico

HINFO CPU e S.O. usato dall'host

MX mail exchange

NS Authoritative Name Server

PTR name space pointer

SOA Inizio di una zona d'autorità

RFC 1035: DOMAIN NAMES – IMPLEMENTATION AND SPECIFICATION

20 SMTP

Posta Elettronica Uno dei primi servizi nati su internet è la posta elettronica (1982). Essa consiste nel **trasferimento di un messaggio tra** uno user **mittente** e uno user **destinatario**. Il **destinatario potrebbe non essere disponibile** in quel momento, e quindi non poter accettare i messaggi: utente impegnato, **computer spento**.

Il servizio di posta elettronica si **basa su componenti intermediari** per trasferire i messaggi: **disaccoppiamento** lato mittente e destinatario, analogia con posta tradizionale.

Il trasferimento di messaggi di posta elettronica si basa sul **Simple Mail Transfer Protocol** (RFC 821, RFC 2821, RFC 5321).

Componenti principali

Agenti utente

Mail Server

Protocolli

20.1 Agenti Utente

Detti anche **mail reader**, usati per la **composizione**, editing, **lettura** di messaggi di posta. Ad es:

Eudora

Outlook

Thunderbird

I messaggi in entrata e uscita **vengono archiviati sul server**

20.2 Mail Server

Le **mailbox** (casella) **contengono**

i messaggi in ingresso diretti all'utente ancora da leggere

una coda di messaggi in uscita ancora da inviare

il **protocollo SMTP** per il dialogo tra mail server, allo scopo di scambiare i messaggi

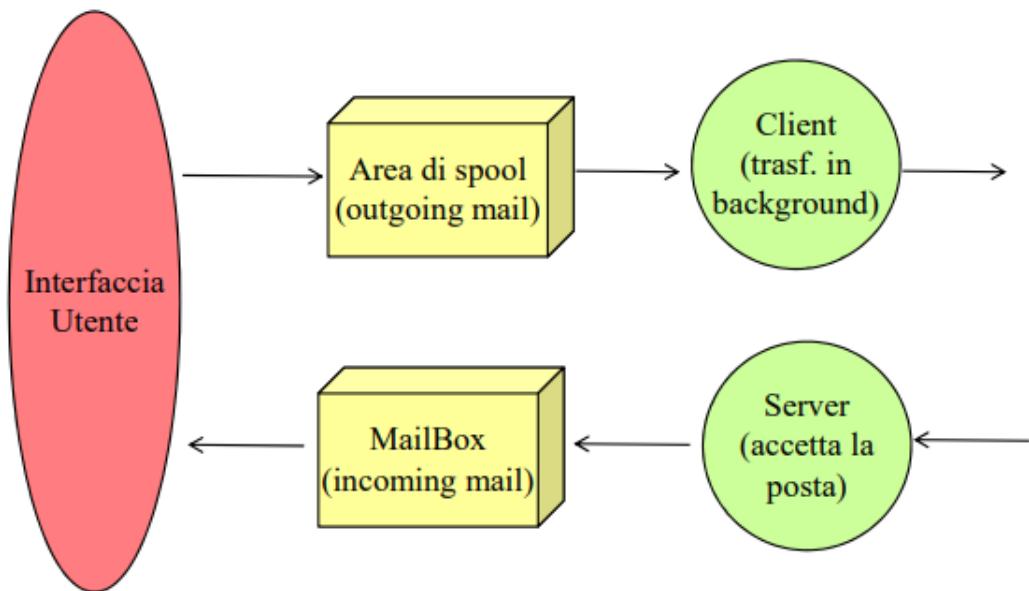
Client: mail server che invia

Server: mail server che riceve

Esempio di scenario: Alice invia un messaggio di posta a Bob.

1. Alice, grazie al suo mail user agent **compone il messaggio**.
Alice **fornisce** allo user agent l'**indirizzo di destinazione**, cioè l'indirizzo di Bob
2. Lo user agent di Alice **spedisce il messaggio** al suo server di posta e il messaggio **viene accodato in attesa di invio**
3. Il **lato client** dell'SMTP sul server di posta di Alice **vede il messaggio** e **apre una connessione** al server SMTP sul server di posta di Bob
4. L'**SMTP client** **invia il messaggio nella connessione**
5. Sull'**host server** di posta di Bob, il **lato server** dell'SMTP **riceve il messaggio** e lo **colloca nella casella di posta di Bob**
6. Bob, quando vuole, chiede al suo user agent di leggere il messaggio

20.3 Schema di principio



La tecnica adottata dai server di posta si chiama **spooling**.

L'utente invia un messaggio, il sistema ne pone una copia in memoria spool – o **area di accomodamento della posta**, insieme all'id mittente, id destinatario, id macchina di destinazione e tempo di deposito.

Il sistema **avvia il trasferimento alla macchina remota**. Il sistema (client) stabilisce una connessione TCP con la macchina destinazione.

Se la connessione viene aperta, **inizia il trasferimento del messaggio**

Se il trasferimento va a buon fine il **client cancella la copia locale del film**

Se il trasferimento non va a buon fine, il **processo di trasferimento scandisce periodicamente l'area di spooling e tenta il trasferimento dei messaggi non consegnati**. Oltre un certo intervallo di tempo (definito dall'amministratore del server), se il messaggio non è stato consegnato **viene inviata una notifica all'utente mittente**.

20.4 Indirizzo

Un ricevente è identificato da un indirizzo email del tipo:

`local-part @ domain-name`

domain-name: specifica un **mail server**. Determina il nome di dominio della destinazione verso la quale la mail deve essere consegnata.

Nome di dominio del server di posta

local-part: specifica una **casella di posta** sul mail server. Spesso identica al login o al nome completo dell'utente.

Indirizzo della casella di posta del destinatario

20.5 Gestione Alias



L'alias è una **casella postale virtuale** che serve a **ridistribuire i messaggi** verso uno o più indirizzi di posta elettronica personali.

Molti – Uno: il sistema di alias permette ad un singolo utente di avere identificatori di mail multipli, assegnando un set di identificatori ad una singola persona.

Un utente → più indirizzi postali

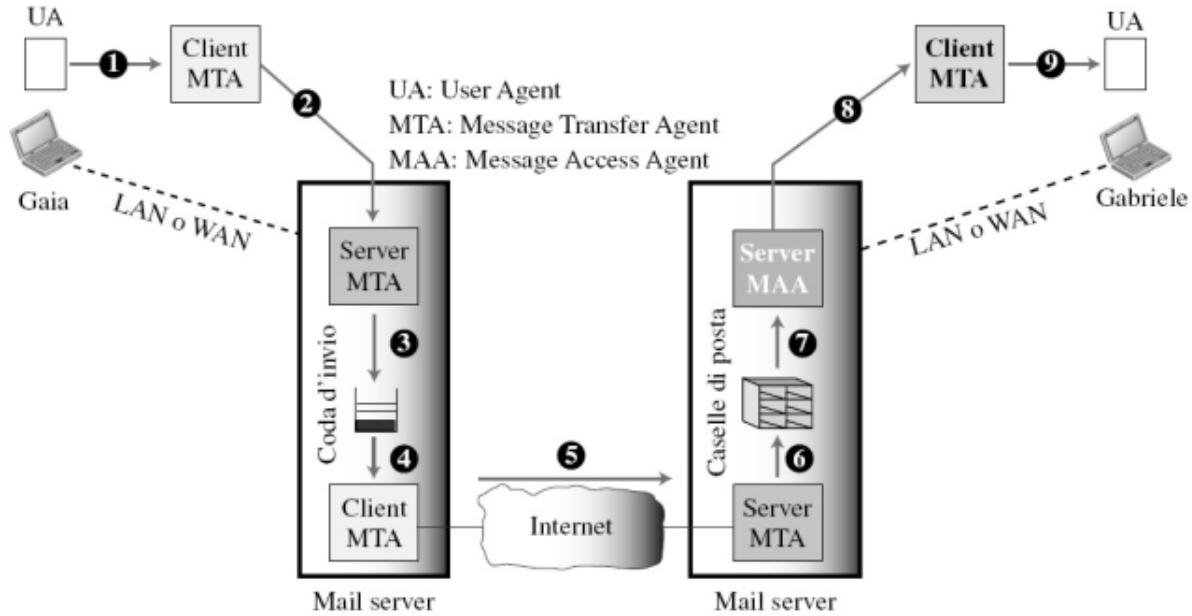
Uno – Molti: il sistema di alias permette di associare un gruppo di destinatari ad un singolo identificatore.

Un indirizzo postale → più utenti destinatari (es. mailing list)

Espansione degli alias postali Conversione di identificatori di indirizzi postali in uno o più indirizzi postali nuovi.

Se l'alias database specifica che all'indirizzo x deve essere assegnato il nuovo indirizzo y, l'espansione dell'alias riscriverà l'indirizzo di destinazione. Si provvederà poi a determinare se y specifichi un indirizzo locale o remoto e lo posizionerà nell'area di spool relativa.

20.6 Modello di riferimento



20.7 Simple Mail Transfer Protocol

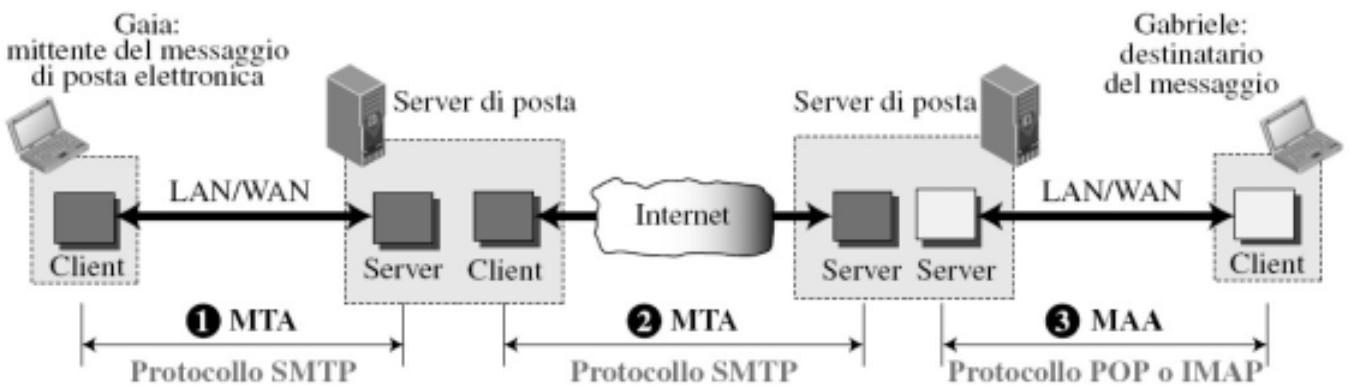
RFC 821, RFC 2821, RFC 5321

L'obiettivo dell'SMTP è trasferire le mail in maniera affidabile ed efficiente. L'SMTP è indipendente dal particolare sottosistema di trasmissione e richiede solo un canale che trasmetta dati ordinati in maniera affidabile.

Nonostante (nelle RFC) si parli specificamente del TCP, altri mezzi di trasporto sono possibili.

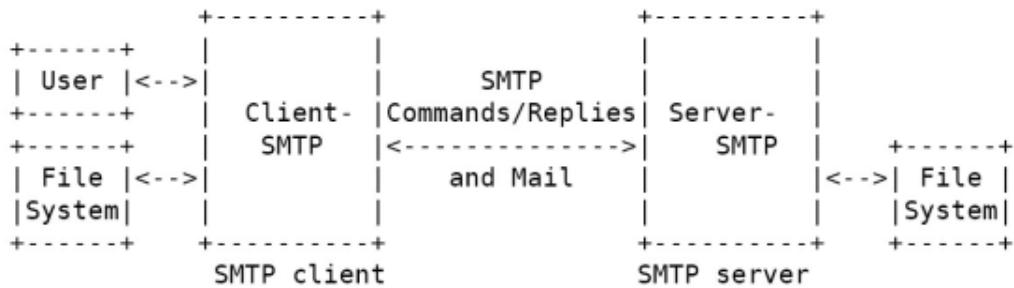
20.8 SMTP Mail Relaying

Una funzionalità importante dell'SMTP è la capacità di **trasportare mail attraverso reti multiple**. In questo modo, un messaggio di posta può passare attraverso molti dispositivi intermedi sul suo percorso dal mittente al destinatario.



20.9 Modello SMTP

Quando un client SMTP ha un messaggio da trasmettere, stabilisce una **canale di trasmissione bidirezionale** verso un server SMTP. La **responsabilità** del client SMTP è di **trasferire messaggi di posta a uno o più server SMTP**, o comunicarne il fallimento.



Un client SMTP determina l'indirizzo di un host appropriato che esegue un server SMTP **risolvendo un nome di dominio di destinazione** ottenendo o un **mail server intermedio** o l'**host finale** (tramite query DNS).

Trasferimento Il trasferimento di un messaggio può avvenire **in una singola connessione** tra mittente e destinatario o tramite **una serie di salti tra sistemi intermedi**.

In entrambi i casi, una volta che il server ha risposto positivamente alla fine della trasmissione dei dati della mail, avviene una **cessione di responsabilità del messaggio**: il **protocollo richiede che un server debba accettare la responsabilità di o consegnare il messaggio o segnalare il fallimento** della consegna.

20.9.1 Fallimenti nella consegna

Possibili problemi

Connessione con il mail server del mittente

Server **inesistente o irraggiungibile**

Connessione con il mail server del destinatario

Server **inesistente o irraggiungibile**

Inserimento nella casella di posta del destinatario

User sconosciuto, casella piena

In tutti i casi **il mittente riceve una notifica!**

L'unico caso in cui il destinatario non riceve il messaggio e il mittente non viene avvisato è se **qualcuno rimuove il messaggio** (intrusi, filtri antispam...)

20.10 Protocollo

Nella pratica viene usato il **protocollo TCP** sulla porta 25 per **consegnare in modo affidabile** i messaggi dal client al server. Questo avviene in tre fasi:

Handshaking

Trasferimento del messaggio

Chiusura della connessione

L'**interazione** avviene tramite **comando/risposta**, con comandi in testo ASCII e risposte con codici di stato e descrizione facoltativa.

I messaggi, header e body, sono in caratteri ASCII a 7bit.

Protocollo semplice, di comandi e risposte I comandi e (a meno di alterazioni dovute ad un'estensione del servizio) i dati dei messaggi SMTP sono trasmessi da mittente a destinario attraverso il canale di trasmissione in "right".

Una **risposta SMTP** è un **riconoscimento** (positivo o negativo) mandato in "righe" da destinatario a mittente attraverso il canale di trasmissione in risposta ad un comando.

La forma generale di risposta è un **codice numerico**, solitamente seguito da una stringa di testo.

```

S: MAIL FROM: <Smith@Alpha.ARPA>
R: 250 OK

S: RCPT TO:<Jones@Beta.ARPA>
R: 250 OK

S: RCPT TO:<Green@Beta.ARPA>
R: 550 No such user here

S: RCPT TO:<Brown@Beta.ARPA>
R: 250 OK

S: DATA
R: 354 Start mail input; end with <CRLF>.<CRLF>
S: Blah blah blah...
S: ...etc. etc. etc.
S: <CRLF>.<CRLF>
R: 250 OK

```

Handshaking Il client stabilisce la connessione e attende che il server invii "220 READY FOR MAIL", cioè la disponibilità a ricevere la posta.

Il client risponde con il comando HELO

Il server risponde identificandosi

A questo punto il client può trasmettere i messaggi

```

S: 220 Beta.GOV Simple Mail Transfer Service ready
C: HELO Alfa.EDU
S: 250 Beta.GOV

```

20.11 Comandi SMTP

Alcuni comandi:

```

HELO <client identifier>
MAIL FROM:<reverse-path> [SP <mail-parameters>
<CRLF>
RCPT TO:<forward-path> [SP <rcpt-parameters>]
<CRLF>
DATA
QUIT

```

Nota bene <CRLF>.<CRLF> per terminare la fine di un messaggio

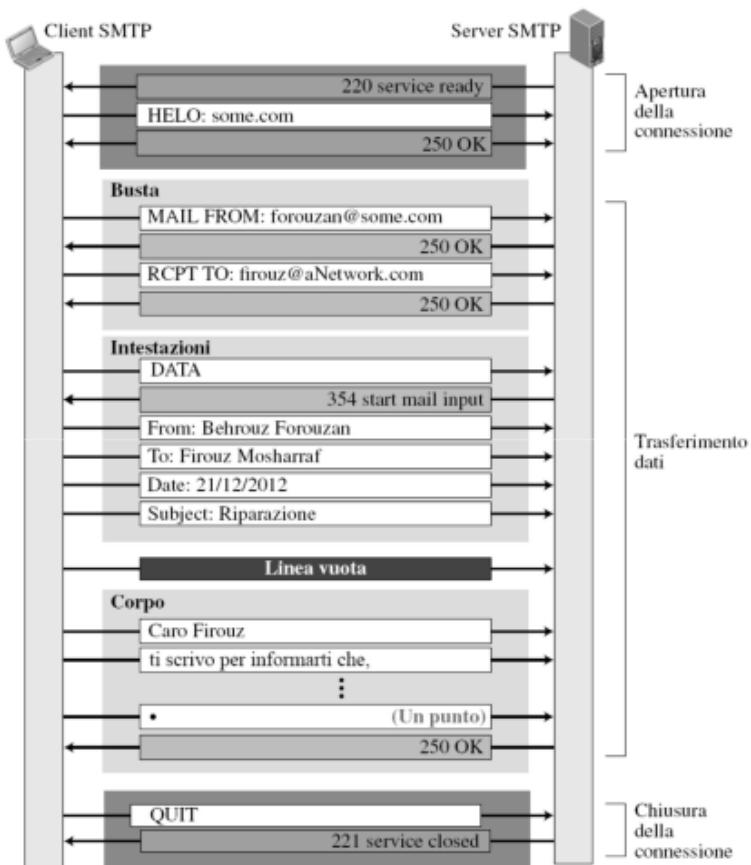
20.12 Esempio

```

S: MAIL FROM:<Smith@Alpha.ARPA>
R: 250 OK
S: RCPT TO:<Jones@Beta.ARPA>
R: 250 OK
S: RCPT TO:<Green@Beta.ARPA>
R: 550 No such user here
S: RCPT TO:<Brown@Beta.ARPA>
R: 250 OK
S: DATA
R: 354 Start mail input; end with <CRLF>.<CRLF>
S: Blah blah blah...
S: ...etc. etc. etc.
S: <CRLF>.<CRLF>
R: 250 OK

```

Esempio d'interazione



20.13 Formato messaggi mail

L' SMTP è il protocollo per lo scambio dei messaggi
L'RFC 2822 è lo standard formato di testo:

Linee d'intestazione ad esempio

To:
From:
Subject:
Diversi dai comandi SMTP!

Linea vuota

Body

Il "messaggio", solitamente in caratteri ASCII 7bit

Esempi di header:

A.3.1, minimo richiesto

```
Date: 26 Aug 76 1429 EDT
From: Jones@Registry.Org
Bcc:
-oppure-
Date: 26 Aug 76 1429 EDT
From: Jones@Registry.Org
To: Smith@Registry.Org
```

Il campo Bcc può essere buoto, il campo **To** deve avere almeno un indirizzo

A.3.2, con qualche campo addizionale

```
Date: 26 Aug 76 1430 EDT
From: George Jones<Group@Host>
Sender: Secy@SHOST
To: "Al Neuman"@Mad-Host, Sam.Irving@Other-Host
Message-ID: <some.string@SHOST>
```

A.3.3, il più complesso possibile

```
Date      : 27 Aug 76 0932 PDT
From      : Ken Davis <KDavis@This-Host.This-net>
Subject   : Re: The Syntax in the RFC
Sender    : KSecy@Other-Host
Reply-To  : Sam.Irving@Reg.Organization
To        : George Jones <Group@Some-Reg.An-Org>, Al.Neuman@MAD.Publisher
cc        : Important folk:
          Tom Softwood <Balsa@Tree.Root>,
          "Sam Irving"@Other-Host;;
Standard Distribution:
  /main/davis/people/standard@Other-Host,
  "<Jones>standard.dist.3"@Tops-20-Host;;
Comment   : Sam is away on business. He asked me to handle his mail for him. He'll be able to provide
           a more accurate explanation when he returns next week.
In-Reply-To: <some.string@DBM.Group>, George's message
X-Special-action: This is a sample of user-defined field-names. There could also be a field-name
                  "Special-action", but its name might later be preempted
Message-ID: <4231.629.XYzi-What@Other-Host>
```

20.14 Estensioni multimediali

L'RFC 822 permette di inviare **solo messaggi di testo** in ASCII. ricevere messaggi

Problema: permettere agli utenti di Internet di inviare/ricevere messaggi

In lingue con accenti (come l'italiano), in alfabeti non latini (come il russo o l'ebraico) o in lingue senza alfabeto (come il cinese o il giapponese)

Contenuti audio o video

20.14.1 MIME

Idea di base Continuare ad usare il formato del messaggio specificato in RFC 822 ma aggiungendo una struttura al message body e definendo regole di encoding per il trasferimento di testo non-ASCII.

Questo ha permesso di inviare messaggi MIME usando protocolli e mail server esistenti, rendendo però necessario cambiare gli user agent.

Multipurpose Internet Mail Extension RFC 2045 e RFC 2046. Questo standard di Internet estende il formato delle mail per supportare:

Testo in character set diversi dall'US-ASCII

Allegati non di testo

Corpo del messaggio diviso in più parti

Informazioni nell'header in character set non ASCII

Si realizza con linee di intestazione aggiuntive per dichiarare il tipo di contenuti MIME. La RFC 2045 definisce una serie di metodi per rappresentare dati binari in formato ASCII.



From: alice@crepes.fr
To: bob@hamburger.edu
Subject: Picture of yummy crepe.
MIME-Version: 1.0 \\Versione MIME
Content-Transfer-Encoding: base64 \\Metodo usato per la codifica dei dati
Content-Type: image/jpeg \\Dati multimediali (tipo, sottotipo, parametri...)

base64 encoded data \\Dati codificati
.....base64 encoded data

Campi nell'header:

MIME-Version: un numero di versione per dichiarare a quale versione aderisce il formato del messaggio

Content-Type: usato per specificare il tipo del media e il sottotipo dei dati nel corpo del messaggio, e per specificare completamente la rappresentazione nativa di quei dati

Content-Transfer-Encoding: specifica la codifica applicata al corpo del messaggio. La codifica è solitamente applicata ai dati per consentire loro di passare attraverso i meccanismi di trasporto delle mail, che potrebbero avere limitazioni sui dati o sui character set.

La maggior parte dei tipi di media che è utile trasportare via mail sono **nativamente rappresentati con caratteri di 8bit** o dati binari. I **client e i server** SMTP si **aspettano messaggi ASCII** (caratteri a 7 bit), i **dati binari invece usato tutti e 8 bit** del byte (es. immagini, eseguibili, set estesi di caratteri). MIME fornisce **cinque schemi di transfer encoding** tra cui:

Codifica ASCII dei dati binari: **codifica base64**

Gruppi di 28bit divisi in 4 unità da 6bit, ciascuna unità inviata come un carattere ASCII

Quoted-printable encoding: per messaggi con pochi caratteri non-ASCII, più efficiente

Oggi giorno i mail server possono negoziare l'invio di dati in codifica binario (8 bit), se la negoziazione non ha successo si usano i caratteri ASCII.

20.14.2 Tipi MIME

`Content-Type: type/subtype; parameters`

Text

Esempi di subtype: `plain, html`

Image

Esempi di subtype: `jpeg, gif`

Audio

Esempi di subtype: `basic` (8bit mu-law encoded), `32kadpcm` (32 kbps)

Video

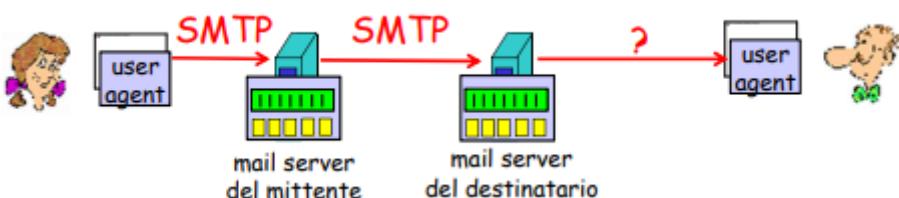
Esempi di subtype: `mpeg`

Application

Altri dati che devono essere processati da un'applicazione prima di essere visualizzabili.

Esempi di subtype: `mword, octet-stream` (dati arbitrari binari)

20.15 Protocolli di accesso alla mail



Per trasferire la posta fino al server del destinatario si utilizza l'SMTP, che è un protocollo di tipo push: la maggior parte dei dati è dal client al server. Per **leggere la posta** invece, è richiesto un protocollo di tipo pull, cioè in cui i dati viaggiano prevalentemente dal server al client.

Messagge Access Agent Attualmente si utilizzano due protocolli di accesso alla posta

POP Post Office Protocol

Semplice ma con funzionalità limitate. Attualmente è in uso la terza versione, POP3.

Fase di autorizzazione

L'accesso avviene con lo **user agent** che **apre una connessione TCP sulla porta 110** e si autentica con i comandi

`user`: specifica lo username

`pass`: specifica la password

ed il server risponderà con `OK` o `ERR`.

Fase di transazione

Il client può utilizzare i comandi:

- list: visualizza la lista dei messaggi
- retr: preleva il messaggio per numero
- dele: elimina il messaggio dal server
- quit: chiude la sessione

Fase di aggiornamento

Dopo aver ricevuto quit il server cancella i messaggi marcati per la rimozione

Il POP3 ha due modalità:

Delete: i messaggi vengono automaticamente eliminati dalla mailbox dopo il prelievo

Keep: i messaggi vengono conservati dopo il prelievo

Inoltre non mantiene le informazioni di stato tra le sessioni, solo le cancellazioni sono permesse.

IMAP Internet Mail Access Protocol

Ha più funzionalità, ed è più complesso, del POP3. Consente tra le altre cose di **manipolare i messaggi memorizzati sul server** con cartelle, e di **estrarre solo alcuni componenti dei messaggi** come solo l'intestazione se si sta usando una connessione lenta.

HTTP

Es. Hotmail, Yahoo! Mail, ...

21 Livello di Trasporto

21.1 Obiettivi

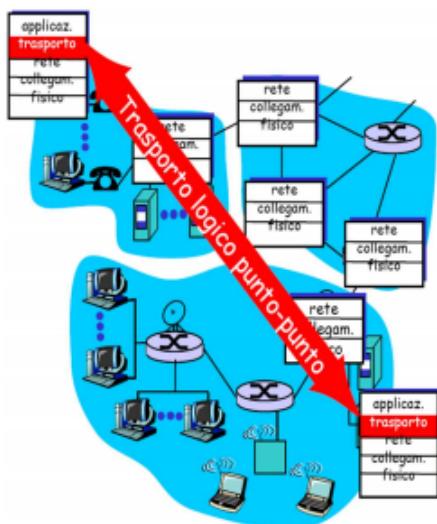
Realizzare una connessione logica fra processi residenti in host diversi.

Logico I processi si comportano come se gli host fossero direttamente collegati, non si preoccupano dei dettagli dell'infrastruttura usata fisica per la comunicazione.

Offrire servizi allo strato applicativo

Un'applicazione interagisce con i protocolli di trasporto per trasmettere e ricevere dati. L'applicazione sceglie lo stile di trasporto: sequenza di messaggi singoli o una sequenza continua di byte. Il programma applicativo passa i dati nella forma richiesta al livello di trasporto per la consegna.

Utilizza i servizi dello strato di rete.



I servizi di trasporto:

Forniscono la **comunicazione logica** fra processi applicativi di host differenti

I **protocolli** di trasporto vengono eseguiti nei sistemi terminali

21.2 Caratteristiche

Servizio privo di connessione In un servizio privo di connessione il **processo mittente consegna i messaggi al livello di trasporto uno per uno**. Il livello di trasporto **tratta ogni messaggio come entità singola** senza mantenere alcuna relazione fra essi. I segmenti possono **non essere consegnati o non arrivare in ordine**.

Servizio orientato alla connessione In un servizio orientato alla connessione client e server **stabiliscono una connessione logica**.

21.3 Servizi offerti

Multiplexing e Demultiplexing Il termine **multiplexing** si riferisce al caso in cui **un'entità riceve informazioni da più di una sorgente**. Lo strato trasporto provvede **allo smistamento dei pacchetti fra rete e applicazioni**.

Demultiplexing invece al caso in cui **un'entità trasmette informazioni a più di un destinatario**. Lo strato trasporto provvede all'**accompagnamento dei flussi dati dai processi verso la rete**, "imbustando" i dati ricevuti dall'alto con un preambolo.

Il livello di trasporto effettua il **multiplexing sul mittente** e il **demultiplexing sul destinatario**.

Si basano sui socket address dei processi e quindi dipendono dal numero di porta su cui i processi sono attivi.

Controllo degli errori

22 UDP

Rispetto al TCP, lo UDP è meno complesso, offre meno servizi, ma è **più indicato in contesti in cui occorre un controllo completo della temporizzazione** cioè in applicazioni time-sensitive come lo streaming.

Nessuna garanzia Lo **User Datagram Protocol** è un **protocollo di consegna a massimo sforzo**. I segmenti UDP possono andare perduti o essere consegnati fuori sequenza. Notare però che l'affidabilità può essere aggiunta a livello applicazione.

Orientato al messaggio Ogni **datagramma UDP** è **indipendente** dall'altro. I processi devono inviare **messaggi di dimensioni limitate, incapsulabili in un datagramma UDP**.

22.1 Proprietà

Nessuna connessione quindi non si introduce ritardo.

Semplice: non viene gestito lo strato di connessione. Le intestazioni sono corte e **non ha controllo di congestione e di flusso**, quindi si possono sparare dati a raffica.

Checksum facoltativo

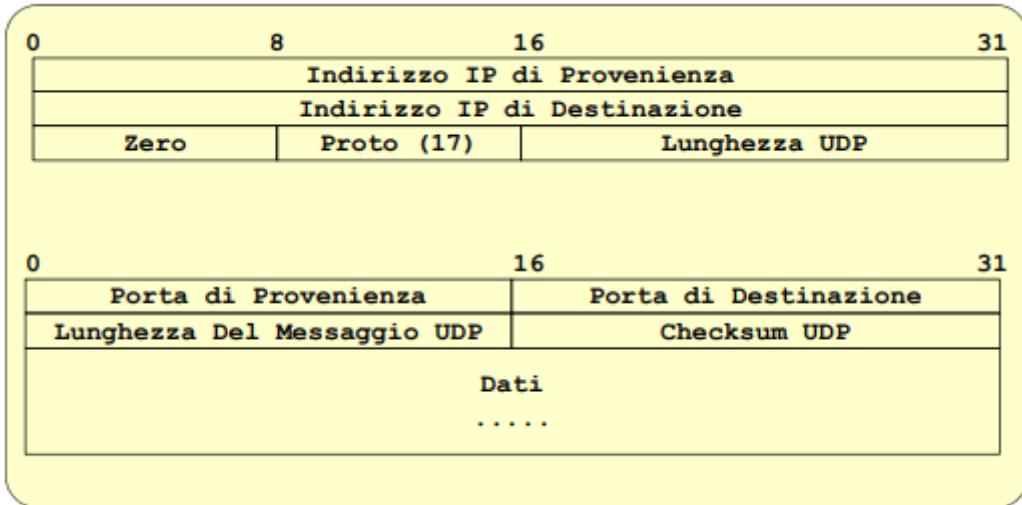
Facile e leggero da gestire, non richiede particolari meccanismi

Concepito solo in funzione del DNS e del TFTP, nel 1980

Utilizzato spesso in **applicazioni multimediali**: tolleranza a piccole perdite e sensibilità alla frequenza

Altri impieghi: DNS, SNMP...

22.2 Datagramma UDP



8 Byte di intestazione

La prima parte è uno pseudo-header che *non* viene trasmesso

Porta: numeri di porta della comunicazione (per il demultiplexing è usato solo quello di destinazione).

La **porta di provenienza** è un campo facoltativo. Quando è significativo, indica la porta del processo mittente, e può essere considerata come la porta alla quale rispondere in assenza di altre informazioni. Se non è usata, viene riempita col valore 0.

Lunghezza del messaggio: lunghezza totale del datagramma UDP, header+dati, è di 65535 Byte.

Checksum checksum dell'intero pacchetto, compreso lo pseudo-header ovvero con il controllo dell'indirizzo host). **Opzionale**, se non è usato è posto a 0xFF (complemento a 1 di 0x00)

Controllo dell'errore end-to-end: il pacchetto corrotto è scartato ma l'utente non viene notificato

L'indirizzo di livello trasporto è una coppia composta dall'indirizzo IP e dalla porta del destinatario (con indicazione dell'indirizzo IP e della porta del mittente)

22.3 Calcolo del checksum

Mittente

Tratta il contenuto del segmento **come una sequenza di interi da 16 bit**

Checksum: somma (complemento a 1) i contenuti del segmento

Il mittente pone il valore della checksum nel campo checksum del segmento UDP

Destinatario

Calcola il checksum del segmento ricevuto

Controlla se il checksum calcolato è **uguale** al valore del campo checksum

No – Errore rilevato

Si – Nessun errore rilevato

22.4 TCP vs UDP

UDP

Trasferimento dati non affidabile

Non fornisce: setup della connessione, affidabilità, controllo del flusso, controllo della congestione, timing o garanzia di banda

Richiede **minor overhead**

TCP

Connection-oriented: richiesto **handshake** tra client e server

Trasporto affidabile tra i processi

Controllo di flusso: il mittente non satura il destinatario

Controllo della congestione: limita il mittente a rete sovraccarica

Non fornisce: **timing, banda minima**

Pro e contro Nella programmazione di rete si deve ricordare che

Il TCP offre un **servizio di trasporto a stream**, quindi si può leggere da un input di rete quanti byte si desiderano

L'UDP invece offre un **servizio a messaggi**, quindi occorre leggere **tutto** il messaggio in arrivo

L'UDP è adeguato per

Processi che richiedono uno scambio di dati con volume limitato **in caso di scarso interesse al controllo del flusso e degli errori**

Processi con **meccanismi interni di controllo del flusso e degli errori**

Trasmissioni **multicast**

Applicazioni interattive in tempo reale che non tollerano ritardi variabili

Considerazioni La RFC 768 che definisce l'UDP è del 1980, non è stata cambiata o integrata da altre RFC ed è di tre pagine.

La RFC 793 che definisce il TCP è del 1981, è stata aggiornata o cambiata da diverse altre RFC, ad esempio

2018 - TCP Selective Acknowledgement Options

1146 - TCP Alternate Checksum Options

2581 - TCP Congestion Control

1323 - TCP Extensions for High Performance

1693 - An Extension to TCP : Partial Order Service

1792 - TCP/IPX Connection Mib Specification

e la lunghezza totale è di 85 pagine, con un glossario di 67 voci.

23 TCP

23.1 Proprietà

Orientato allo stream La lunghezza in byte è **indefinita** a priori. Il servizio di consegna a destinazione passa esattamente la medesima sequenza di byte che il mittente ha passato al servizio di consegna nell'origine. Il TCP vede i **dati come un flusso di byte ordinati**, ma **non strutturati**.



Orientato alla connessione I processi effettuano un **handshake** prima dello scambio dei dati. L'**handshake** è uno **scambio di informazioni preliminari per preparare il trasferimento dei dati**.

Orientato perché lo stato della connessione risiede sui punti terminali, non sugli elementi intermedi della rete (es. router).

La **connessione** è vista dagli applicativi (USERS) come un **circuito fisico dedicato**. Il **TCP** è quindi capace di fornire servizi di tipo **connection-oriented**, mentre il protocollo **IP** su cui si appoggia è in grado di **fornire servizi connection-less**.

Connessione full-duplex Il **flusso dati tra due host può avvenire contemporaneamente nelle due direzioni**. Le due direzioni sono slegate e la connessione è punto-a-punto.

Trasferimento bufferizzato Il software del protocollo TCP è libero di **suddividere lo stream in segmenti indipendenti** (pacchetti), in maniera indipendente dal programma applicativo che li ha generati. Per fare questo è necessario disporre di un **buffer dove immagazzinare la sequenza di byte**. Appena i dati sono sufficienti per riempire un **datagramma** ragionevolmente grande, questo viene trasmesso attraverso la rete.

23.2 Funzioni del segmento TCP

Funzioni base per il trasferimento di dati:

Capacità di trasferire un **flusso continuo** di byte

Trasferimento bidirezionale (**full duplex**)

Multiplexing: consente di assegnare una data **connessione** ad un dato processo.

Permette una comunicazione da processo a processo.

Controllo della connessione: meccanismi di inizio e fine trasmissione (controllo della sessione)

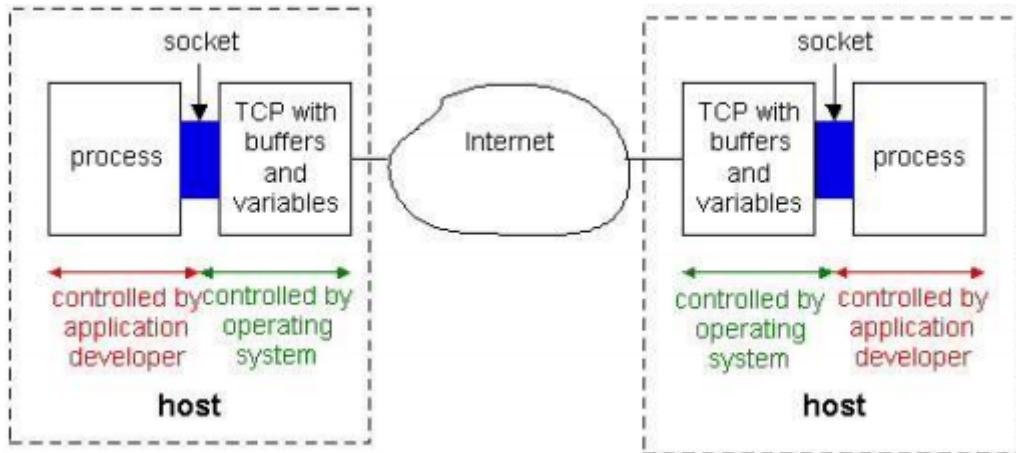
Trasferimento dati ordinato e affidabile: si intende la capacità di correggere tutti i tipi di errore, quali:

- dati corrotti
- segmenti persi
- segmenti duplicati
- segmenti fuori sequenza

Controllo di flusso: evitare di spedire più dati di quanti il ricevitore sia in grado di trattare

Controllo di congestione: ha lo scopo di recuperare situazioni di sovraccarico nella rete

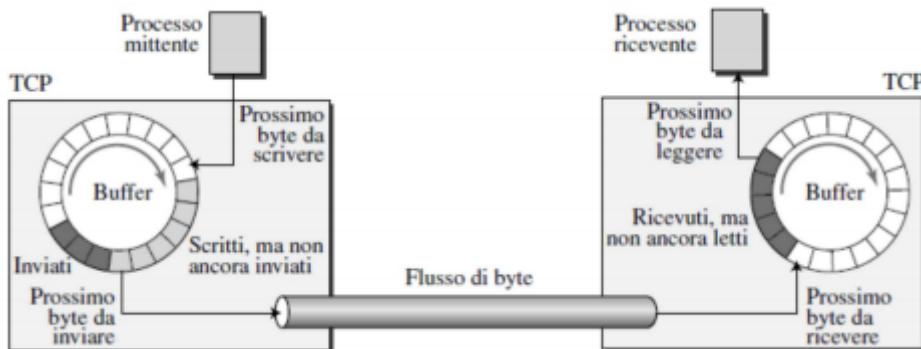
23.3 Processi e Socket TPC



I processi in esecuzione su diverse macchine comunicano tra loro mandando messaggi **attraverso i socket**. Un po' come se ogni processo fosse una casa e ogni socket fosse la porta, **un socket è la porta tra il processo dell'applicazione e il TCP**.

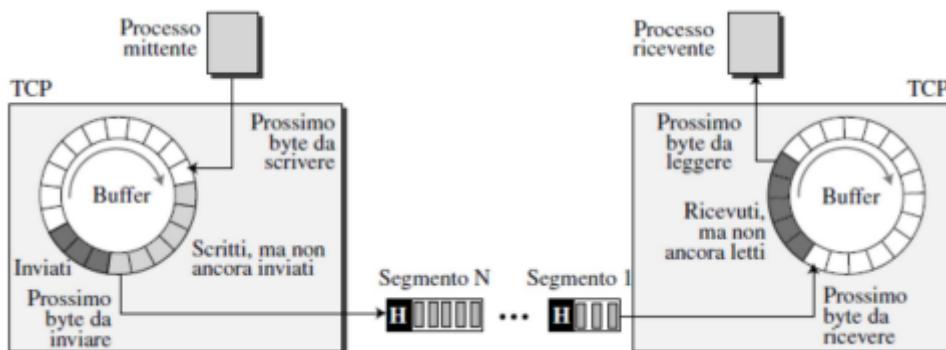
Lo sviluppatore ha il controllo su tutta la parte del **lato applicazione del socket**, ma non può controllare il **lato del livello di trasporto**. Al più, lo sviluppatore può specificare alcuni parametri del TCP, come la dimensione massima del buffer e dei segmenti.³

23.4 Trasferimento bufferizzato



I processi a livello applicativo **scrivono e leggono byte nel/dal buffer**, e questo può avvenire a velocità diverse.

23.5 Segmenti TCP



Il flusso di byte è **partizionato in segmenti**: ogni segmento ha il suo header e viene consegnato al livello IP.

³<http://www3.gdin.edu.cn/jpkc/dzxnw/jsjkj/chapter2/26.htm>

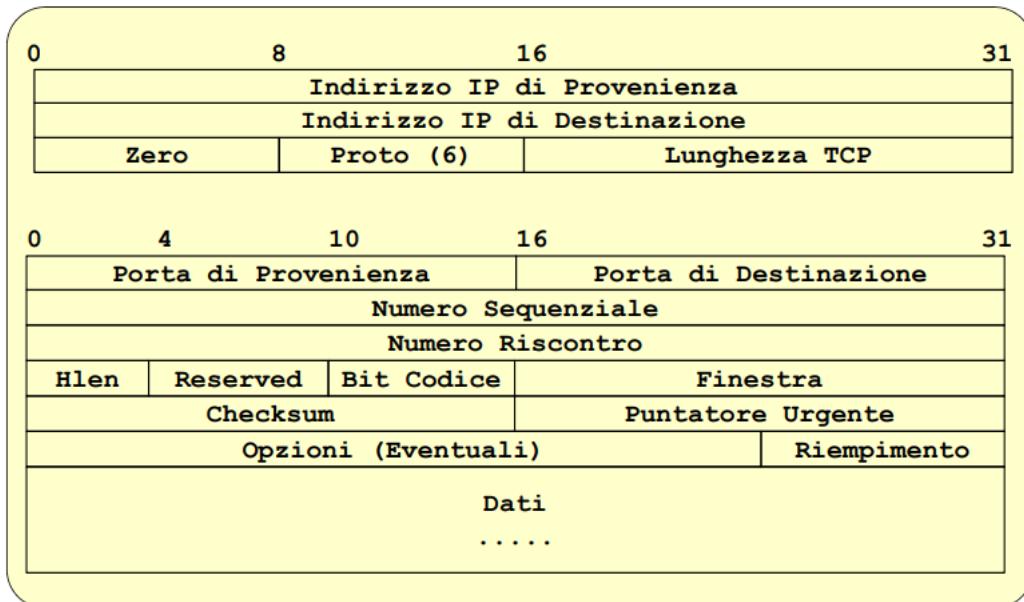
23.6 Numeri di sequenza e di riscontro

Il TCP numera i byte anziché i segmenti

Numero di sequenza Associato al segmento, è il numero (nel flusso) del primo byte (di dati) del segmento. In genere si parte da un **initial sequence number** generato in modo casuale e diverso da 0.

Numero di riscontro $1 + \text{numero dell'ultimo byte correttamente ricevuto}$. I riscontri sono interpretati come **cumulativi**: se ricevo $\text{ACK} = y$ significa che *aspetto il byte y* e quindi ho ricevuto tutti i byte fino a $y - 1$.

23.7 Segmento TCP



Lo **pseudo header** NON viene trasmesso.

Ha valenza logica, contiene **IP sorgente** e **destinazione**. Queste info **vengono inserite** effettivamente e trasmesse dal livello inferiore.

Proto: codice del protocollo.

Lunghezza TCP: lunghezza del segmento TCP escluso lo pseudoheader, server per il calcolo del checksum.

Nel segmento TCP bisogna notare la presenza di:

Numero di **sequenza**

Numero di **riscontro**

Finestra

Essi permettono il **controllo del flusso**, il meccanismo di **ritrasmissione** ed il **riordino** dei pacchetti in ricezione, necessari per la struttura stream-based del TCP

Inoltre è presente un campo **urgent** che permette la trasmissione dei dati "fuori banda", ovvero a **priorità maggiore degli altri** (la loro gestione però è affidata all'applicazione).

23.7.1 Campi del segmento TCP

Porta (16 bit): numeri di porta della comunicazione

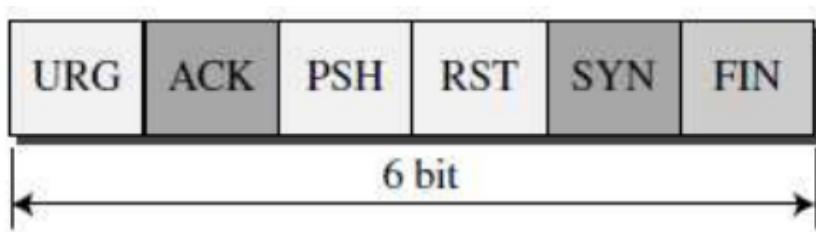
Numero di sequenza (32 bit): è il **numero di sequenza** nello stream **del primo byte di dati di questo segmento**.

Se il flag SYN è settato, allora il numero di sequenza è **ISN** (Initial Sequence Number) e il primo byte di dati è **ISN + 1**.

Numero di riscontro (32 bit): se il flag ACK è settato, allora questo campo **contiene il valore del prossimo numero di sequenza che il mittente del segmento si aspetta di ricevere dall'altro host**. Una volta che la connessione è stabilita è sempre inviato.

Hlen (4bit): **lunghezza dell'header** TCP espressa in parole da 4 byte
La lunghezza dell'header può variare da 20 a 60 byte)

Bit codice: sono **6 flag** e, da sinistra a destra, servono per:



URG: il campo **puntatore urgente** contiene dati significativi da trasferire in via prioritaria

ACK: il campo **numero di riscontro** contiene dati significativi

PSH: **funzione push**, cioè **trasferimento immediato** dei dati in un **segmento dal livello trasporto al livello applicativo**

RST: reset della connessione

SYN: sincronizza il numero di sequenza

FIN: non ci sono altri dati utente – chiusura della connessione

Finestra di ricezione (16 bit): indica il **numero di byte di dati** a partire da quello indicato nel campo **numero di riscontro che il mittente di questo segmento è in grado di accettare**.

Serve per il controllo del flusso.

Checksum (16 bit): checksum dell'intero pacchetto (compreso lo pseudo header).

Serve per il **rilevamento degli errori** in caso di alterazione dei bit del segmento, e si calcola come per l'UDP (ma per il TCP + obbligatorio)

Opzioni (facoltativo, lunghezza variabile di massimo 40 byte): **negoziazione di vari parametri**, ad esempio: dimensione massima del segmento (MSS), selective acknowledgment supportato e blocchi di dati riscontrati selettivamente.

Le opzioni sono sempre multipli di 8 bit e il loro valore è incluso nel checksum.

Puntatore urgente (16 bit): questo campo è un **offset positivo a partire dal numero di sequenza del segmento corrente**. Viene **interpretato solo** se il bit URG è uguale a 1.

Punta al primo byte di dati non urgenti in attesa nella coda di ricezione. Nel segmento contenente dati urgenti **deve essere presente almeno un byte di dati**.

Ad es. se un segmento contiene 400 byte di dati urgenti e 200 byte di dati non urgenti, il puntatore urgente vale 400.

23.7.2 Formato del segmento TCP



23.8 Gestione della connessione

23.8.1 Three-Way Handshake

Handshake a tre vie Dopo l'handshaking a livello di trasporto, non c'è più distinzione tra client e server. Sequenza:

- Il client invia una richiesta di connessione ad un server TCP.
SYN è attivo
Il segmento non contiene dati
Si trasmette anche un numero di sequenza iniziale casuali
Es. SYN = 1, clientISN = 41
- Il server estrae il segmento, alloca i buffer e le variabili TCP per la connessione.
Invia in risposta un segmento di connessione garantita chiamato SYNACK
SYN è attivo
Il numero di sequenza è il valore iniziale (es. 78)
ACK è attivo, il server aspetta clientISN + 1 (es. 42)
Es. SYN = 1, ACK = clientISN + 1, serverISN = 78
- Il client alloca buffer e variabili di connessione, poi manda un riscontro positivo del messaggio del server.
SYN è inattivo, questo segmento può già trasportare dati.
Il prossimo dato data clientISN + 1 (es. 42) ed il client attende serverISN + 1 (es. 79)
SYN = 0, riscontro serverISN + 1

- Inizia lo scambio di dati

I primi segmenti non hanno carico utile. All'arrivo del primo segmento il server inizializza due buffer e le variabili, necessari per il controllo del flusso e delle congestione.

All'arrivo del riscontro del primo segmento il client alloca due buffer e le variabili, per lo stesso motivo.

Alla ricezione del terzo segmento la connessione è instaurata.

23.8.2 Perché a tre vie

Esempio degli scalatori:

Ti tengo la corda? →

← Si, tienimi

Ok →

Le cose possono andare male in tanti modi diversi, alcuni di questi sono **situazioni indistinguibili per A mittente**

Ti tengo la corda? →

Ti tengo la corda? → X

X ← Si, tienimi

Altre invece sono **situazioni indistinguibili per B destinatario**

Ti tengo la corda? →

Ti tengo la corda? → X

← Si, tienimi

Ok → X

23.8.3 Esempio



SYN e SYN + ACK non contengono dati utente ma consumano un numero di sequenza

23.8.4 Chiusura della connessione con handshake

In una connessione TCP normale, dopo l'ACK finale per la chiusura si aspetta un tempo di **timeout molto grande**.

$\text{timeout} = 2 * \text{MSL}$

MSL è il Maximum Segment Lifetime = 120s



Invece con il three-way handshake, dopo l'ACK finale la connessione è immediatamente chiusa.



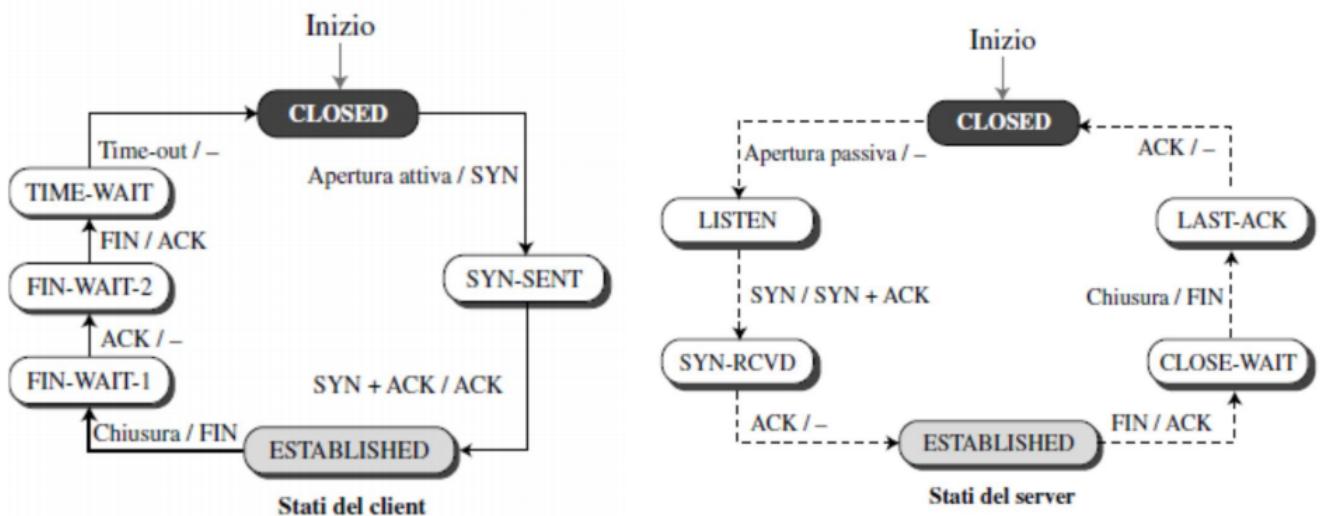
Un FIN che non trasporta dati ma consuma un numero di sequenza. Idem per FIN + ACK.

23.9 Half-Close

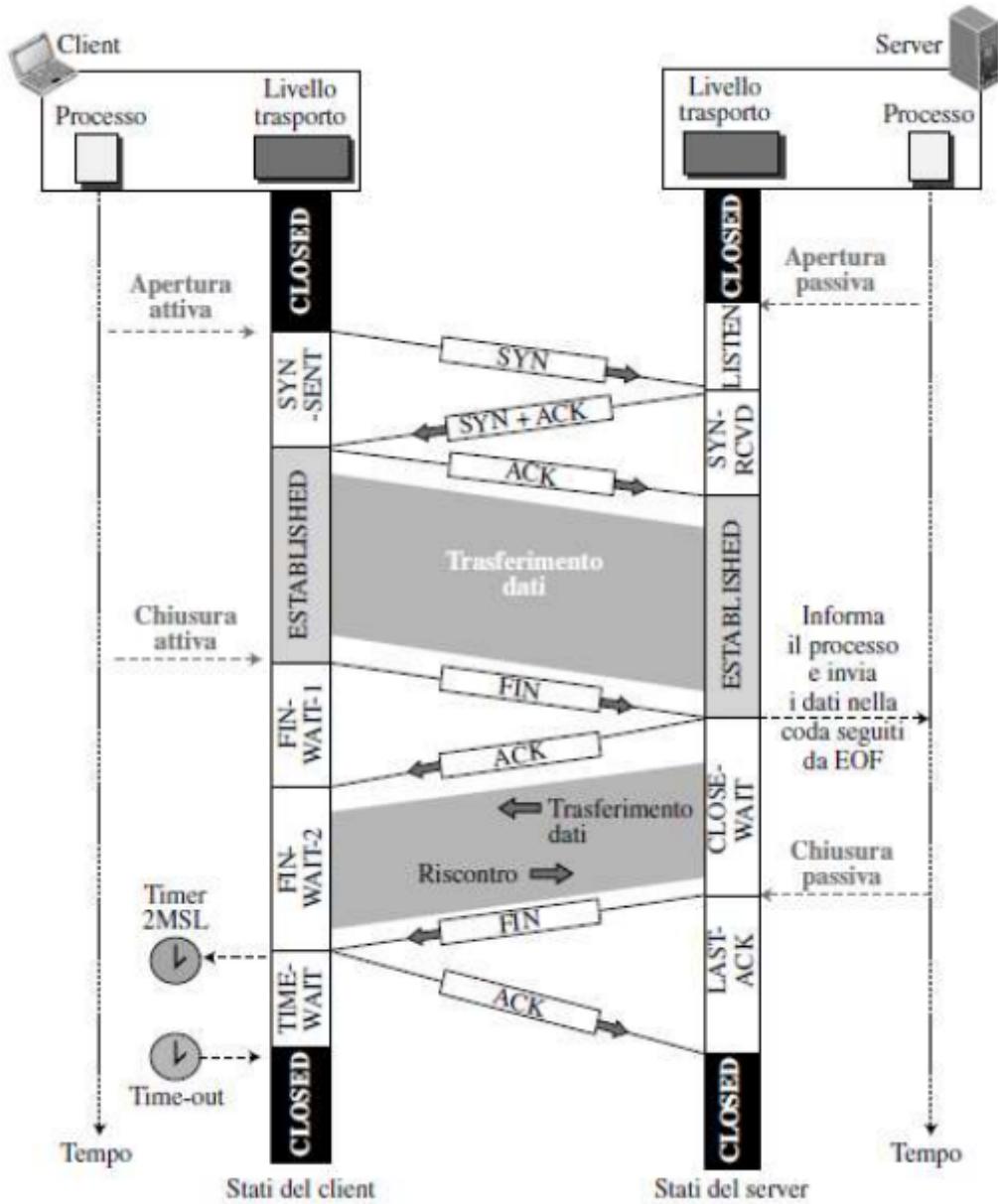
Uno dei due processi può smettere di inviare dati mentre sta ancora ricevendo dati.
`TCP.close(conn) = "I have no more data to send."`



Di seguito l'ASF corrispondente agli stati dell'**half-close**



23.9.1 Scenario Half-Close



23.9.2 Stato Time-Wait

Finale Time-Wait è lo stato finale in cui il capo di una connessione che esegue la chiusura attiva resta prima di passare alla chiusura definitiva della connessione. Dura due volte la Maximum Segment Lifetime.

MSL La Maximum Segment Lifetime è la stima del massimo periodo di tempo che un pacchetto IP può vivere sulla rete.

Questo tempo è limitato perché ogni pacchetto IP può essere ritrasmesso dai router un numero massimo di volte detto **hop limit**.

Ogni implementazione del TCP sceglie il proprio valore per la MSL (RFC 1122 indica 2 minuti, Linux usa 30 secondi).

Perché Time-Wait Lo stato Time-Wait viene usato dal protocollo per due motivi principali:

- Implementare in maniera affidabile la terminazione della connessione in entrambe le direzioni.
Se l'ultimo ACK della sequenza viene perso, chi esegue la chiusura passiva manderà un ulteriore FIN, chi esegue la chiusura attiva deve mantenere lo stato della connessione per essere in grado di rinviare l'ACK.
- Consentire l'eliminazione di segmenti duplicati in rete

23.10 Stati del TCP



Gli stati del TCP:

LISTEN: attesa di una richiesta di connessione remota TCP

SYN-SENT: attesa di ricezione di una connection request corrispondente dopo aver mandato la propria connection request

SYN-RECEIVED: attesa dell'ACK dopo aver entrambi spedito la connection request

ESTABLISHED: connessione aperta, i dati possono essere trasferiti all'utente. Questo è lo stato normale per la fase di trasferimento dati della connessione

FIN-WAIT-1: attesa della richiesta di terminazione della connessione dal TCP remoto, oppure un'ACK sulla richiesta di terminazione della connessione precedentemente spedita

FIN-WAIT-2: attesa della richiesta di terminazione attiva della connessione dal TCP remoto

CLOSE-WAIT: attesa della richiesta di terminazione della connessione dall'utente locale. Si è ricevuta la richiesta di chiusura della connessione e si è già mandato l'ACK, quindi la chiusura passiva è già avvenuta e si attende di avviare la chiusura attiva

CLOSING: attesa dell'ACK sulla chiusura della connessione prima di andare in TIME-WAIT

LAST-ACK: attesa dell'ACK sulla chiusura attiva della connessione. Chiusura passiva effettuata e chiusura attiva iniziata

TIME-WAIT: attesa di abbastanza tempo per assicurarsi che l'host remoto riceva l'ACK per la chiusura della connessione

CLOSED: nessuna connessione

23.11 Trasferimento Dati Affidabile

Un **segmento TCP** può essere **smarrito** o **corrotto**. Il TCP crea un **servizio di trasferimento dati affidabile** sul servizio inaffidabile dell'IP.

Checksum I controlli sono **obbligatori** ed i segmenti corrotti vengono scartati.

Riscontri

Numero di sequenza di un segmento è il **numero del primo byte del segmento nel flusso di byte**. I numeri di sequenza si applicano ai byte, non ai segmenti trasmessi.

Numero di riscontro: numero di sequenza successivo che l'host attende.

Riscontro cumulativo: si effettua il **riscontro dei byte fino al primo byte mancante** del flusso.

Timer

23.11.1 Sequenza e riscontro

Esempio: Telnet su TCP

A SEQ = 42, ACK = 79, DATA = 'C' → B

Viene inviato il carattere "C" da A verso B, *ti mando il 42 aspetto il 79*

Nella direzione opposta si rimanda quanto ricevuto

A ← SEQ = 79, ACK = 43, DATA = 'C' B

ACK dell'host per ricevuta di "C", restituisco "C", *ti mando il 79, aspetto il 43*

A SEQ = 43, ACK = 80 → B

ACK dell'host sulla ricevuta di "C" di ritorno, *ti mando il 43 (o niente) e aspetto l'80*

Pipeline Con il **pipeline** il **mittente può inviare più segmenti senza attendere il riscontro**. Permette di aumentare la produttività

A SEQ = 42 → B

A SEQ = 43 → B

A ← ACK = 43 B

A ← ACK = 44 B

23.11.2 Eventi lato mittente

TCP riceve i dati dall'applicazione:

- assegna un numero di sequenza
- avvia il timer (intervallo di scadenza – timeout)

Timeout

- ritrasmette il segmento non riscontrato
- riavvia il timer

ACK duplicato

Se il mittente **riceve tre ACK duplicati**, allora significa che il **segmento successivo a quello riscontrato è andato perso**.

Ritrasmissione rapida (fast retransmission) **prima** dello scadere del timer.

Ritrasmissione dei segmenti

- se timeout
- se il mittente riceve tre ACK duplicati (fast retransmission)

Segmenti fuori sequenza

I dati possono arrivare fuori sequenza ed essere temporaneamente memorizzati dall'entità TCP destinataria. Il TCP non dice come il destinatario deve gestire i pacchetti fuori sequenza, dipende dall'implementazione.

Nelle versioni più recenti si implementa SACK: i pacchetti fuori sequenza vengono memorizzati, il riscontro dei pacchetti fuori sequenza e duplicati è inviato in OPTIONS.

Ritrasmissione dovuta a riscontro perso



23.11.3 Eventi lato destinatario

Tutti i segmenti inviati per trasmettere dati includono l'ACK. Se il destinatario non ha dati da inviare e riceve un segmento "in ordine", allora ritarda l'invio dell'ACK di 500ms a meno che non riceva un nuovo segmento.

Se il destinatario riceve un segmento atteso ma il precedente non è stato riscontrato allora invia immediatamente l'ACK.

Se il destinatario riceve:

segmento fuori sequenza

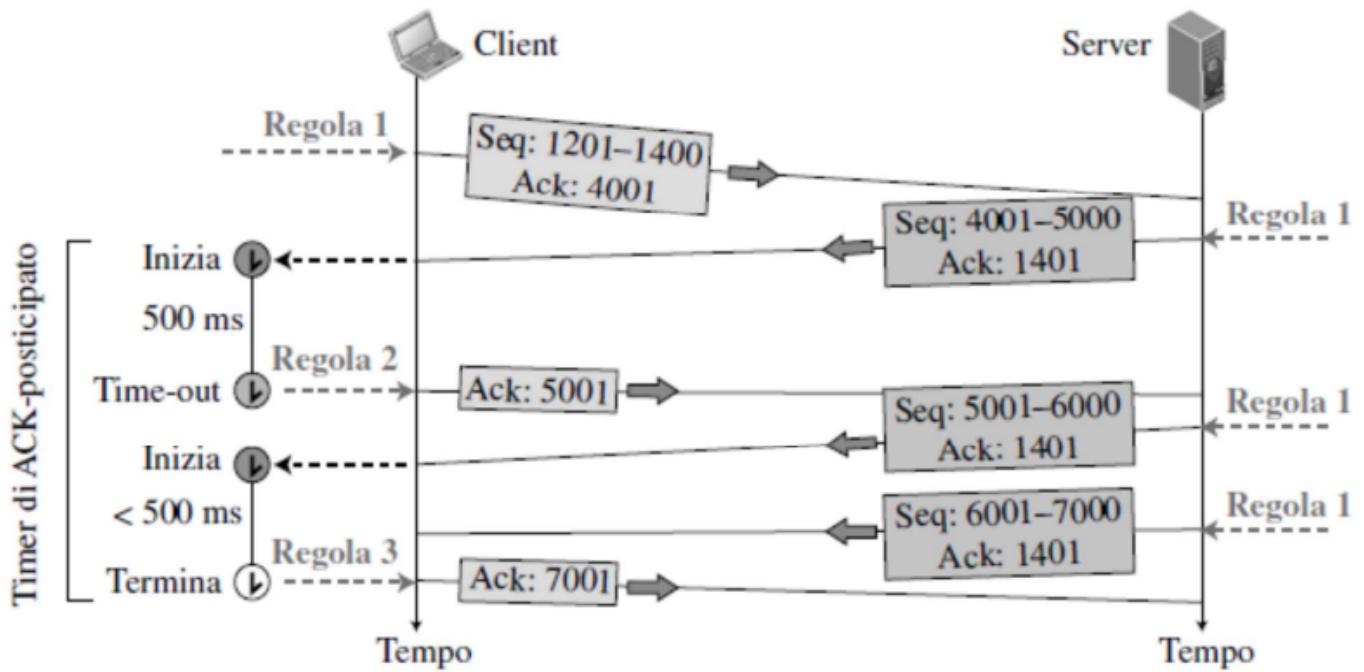
segmento mancante ("buco" nella sequenza")

segmento duplicato

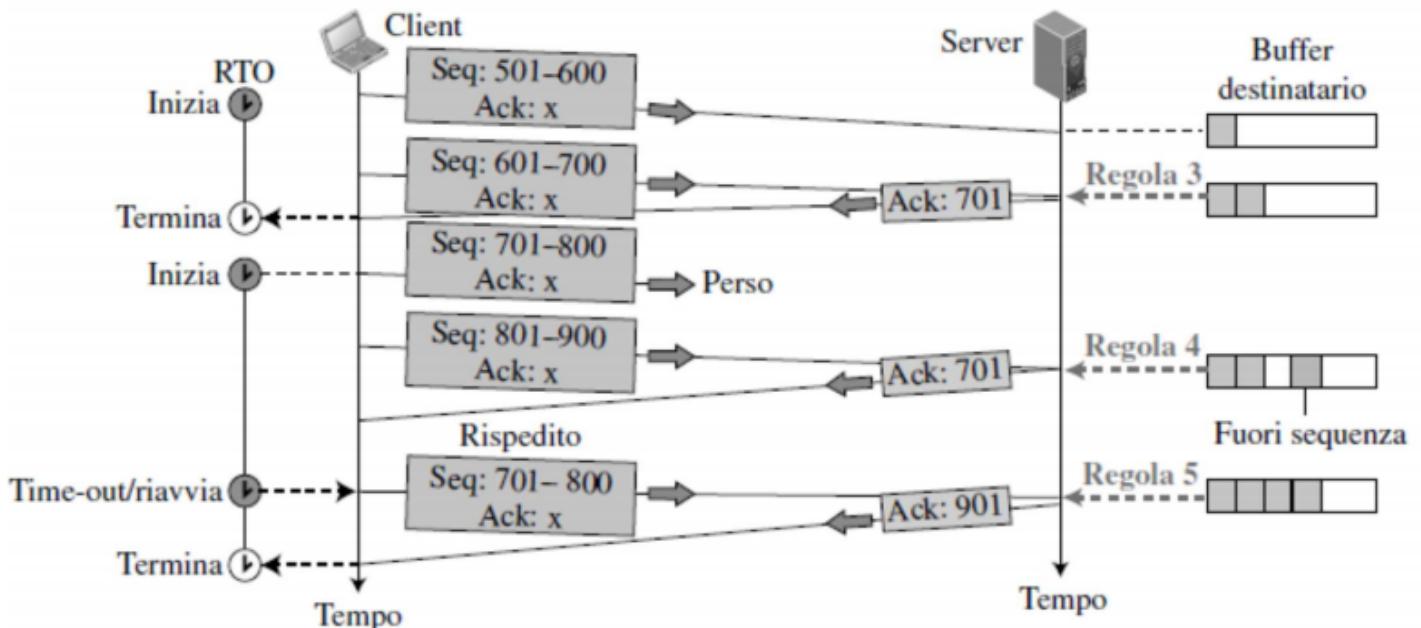
allora invia immediatamente l'ACK indicando il prossimo numero atteso.

23.11.4 Esempi

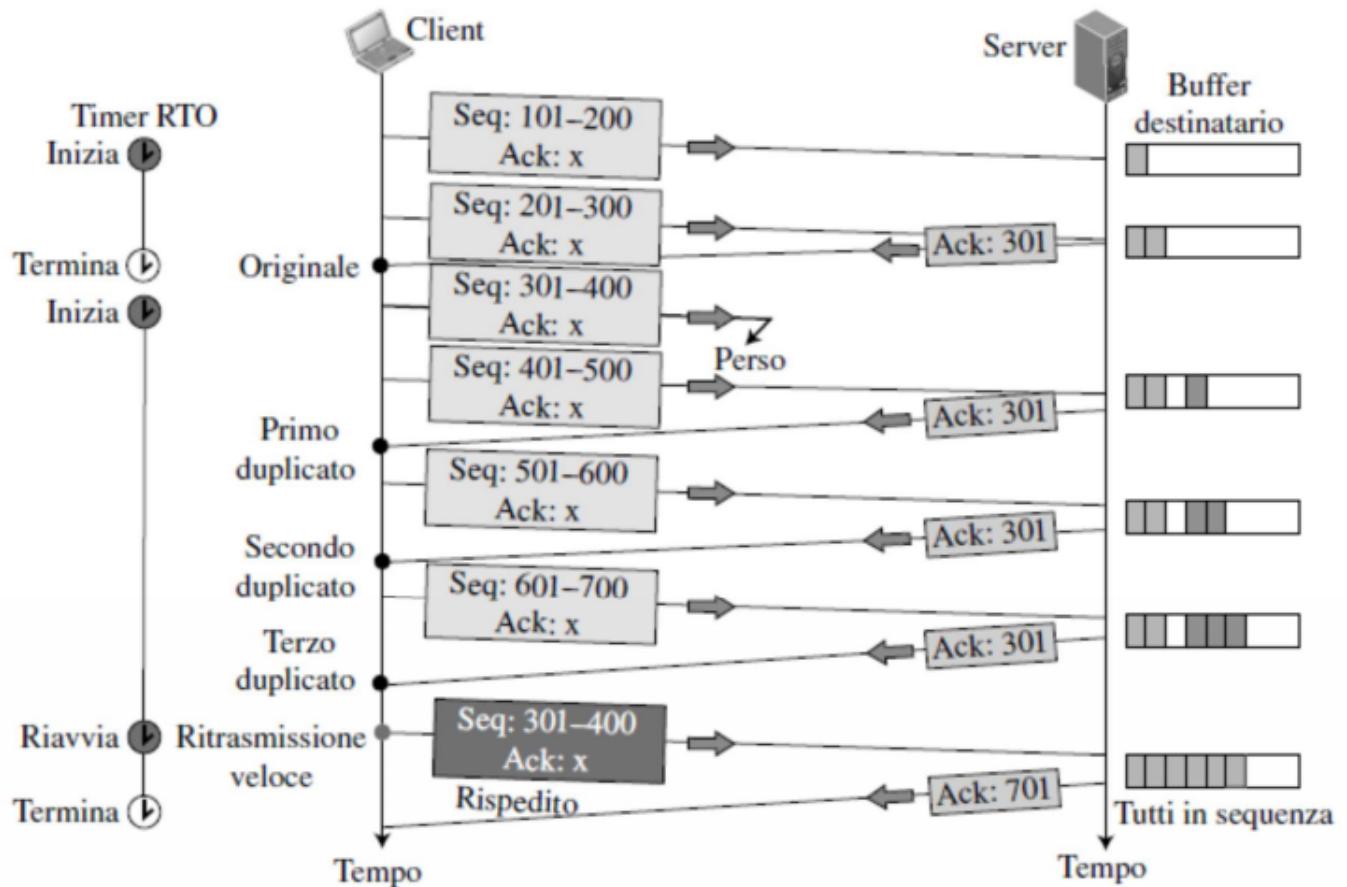
Operatività normale



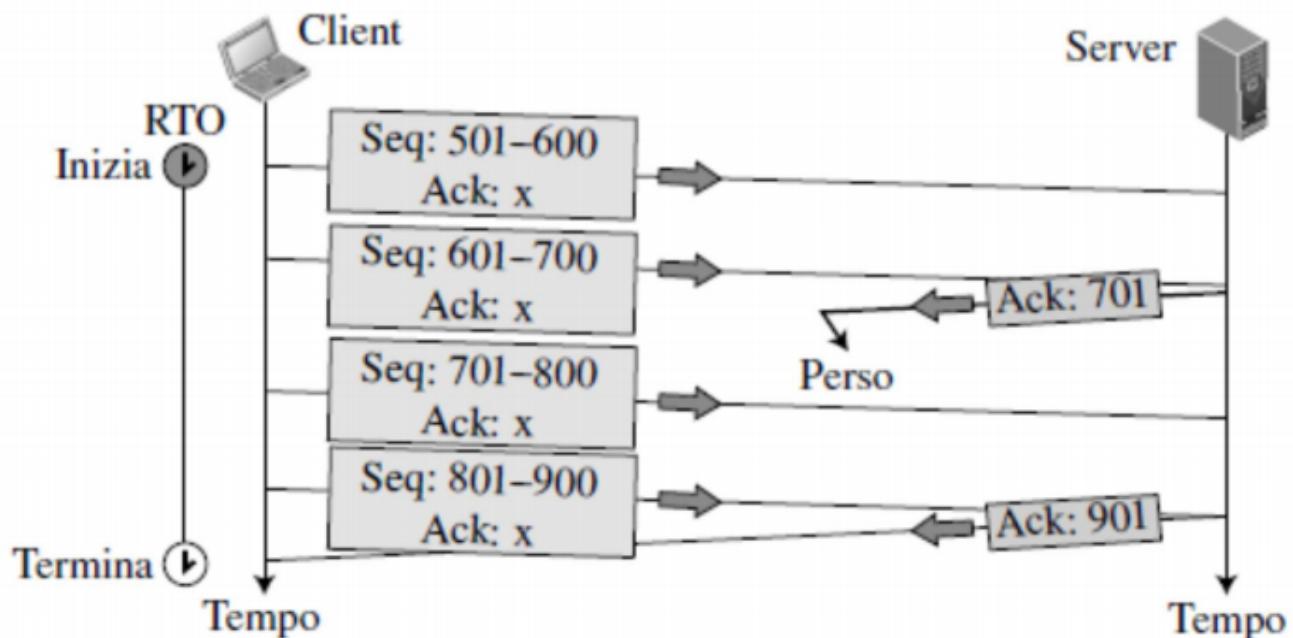
Segmento smarrito



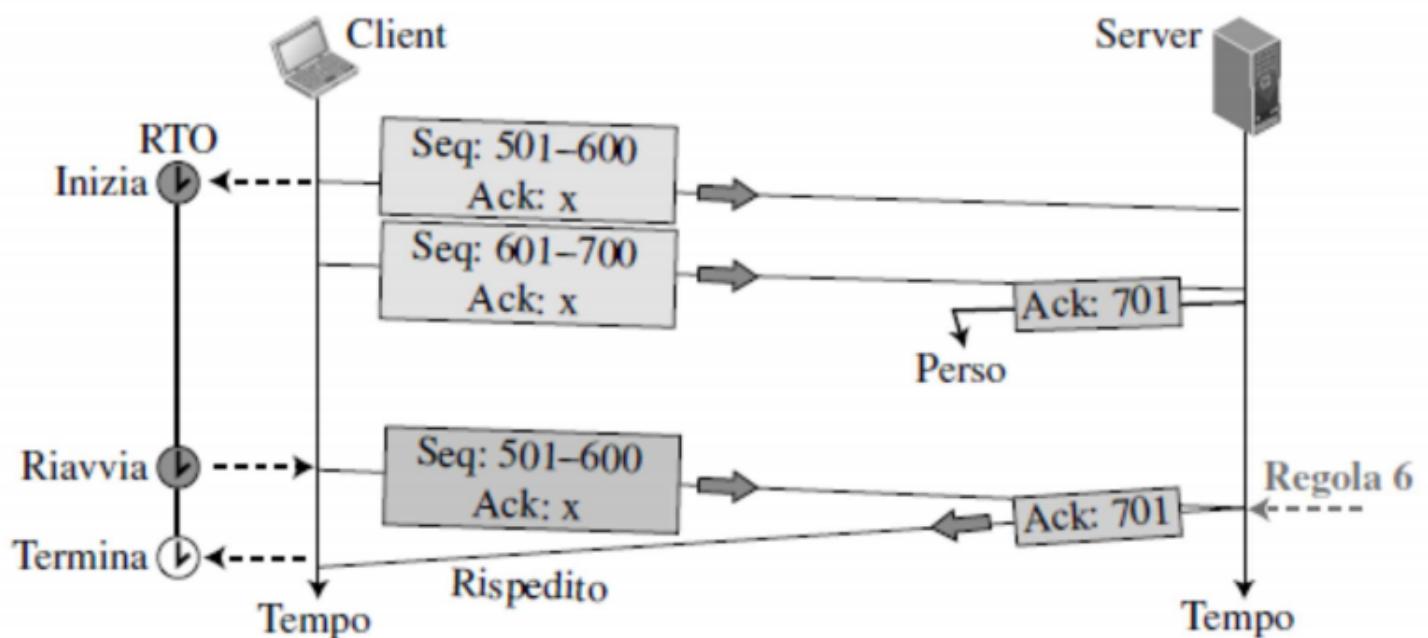
Ritrasmissione veloce



Riscontro smarrito



Riscontro smarrito corretto da ritrasmissione



ASM semplificato per il mittente



ASM semplificato per il destinatario



23.12 Calcolo del timeout

RTO Il Retransmission Timeout, o tempo di timeout, è **fondamentale** per il funzionamento del TCP. L'RTO deve essere maggiore dell'RTT (Round Trip Time)

RTT Il Round Trip Time **tempo trascorso da quando si invia un segmento a quando se ne riceve il riscontro.** Viene calcolato analizzando gli RTT dei segmenti **non ritrasmessi:** Sample RTT, stimato per un segmento trasmesso, non per ogni invio)

$$\text{EstimatedRTT} = (1 - \alpha) * \text{EstimatedRTT} + \alpha * \text{SampleRTT}$$

Poiché SampleRTT può fluttuare, si considera EstimatedRTT, cioè la **combinazione** dei precedenti valori di EstimatedRTT con il nuovo valore SampleRTT.

RFC 2988 – Il valore α viene posto a $\frac{1}{8}$, in modo da rendere via via meno importanti gli RTT dei pacchetti più vecchi.

$$\text{EstimatedRTT} = 0.875 * \text{EstimatedRTT} + 0.125 * \text{SampleRTT}$$

Variabilità Oltre al valore RTT stimato, è necessaria anche una **stima della variabilità** di RTT, data dalla seguente formula che **stima di quanto SampleRTT si discosta da EstimatedRTT:**

$$\text{RTT}_{DEV} = (1 - \beta) \text{RTT}_{DEV} + \beta * |\text{RTT}_{SAMPLE} - \text{RTT}_{ESTIMATED}|$$

RFC2988 – Il valore di β viene posto a $\frac{1}{4}$

Calcolo del timeout Una volta ottenuti questi valori, il timeout viene normalmente calcolato come:

$$\text{RTO} = \text{RTT}_{ESTIMATED} + 4 * \text{RTT}_{DEV}$$

Inoltre in molte implementazioni, dopo un errore (es. ACK non ricevuto) **si raddoppia il timeout**. Questo è un primo meccanismo di controllo della congestione.

23.13 Finestra di trasmissione

Sliding Window Il controllo del flusso (flow control) del TCP si basa sulla **finestra di trasmissione**. La finestra è sovrapposta alla sequenza da trasmettere, viene **negoziata dinamicamente** e viene fatta avanzare alla ricezione di un ACK.

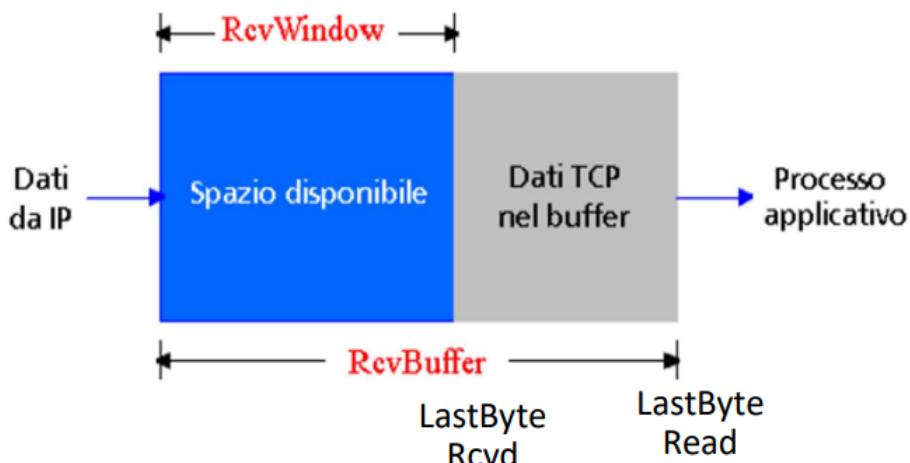




Ogni host impone buffer di invio e ricezione. Il **processo destinatario** legge i dati dal buffer di ricezione (**non necessariamente nell'istante in cui arrivano**)

Con **controllo di flusso** si intende la **capacità del mittente** di evitare la possibilità di saturare il buffer del **ricevitore**, modulando la propria frequenza di invio. Quindi si mette in relazione la **frequenza di invio** del mittente con la **frequenza di lettura** dell'applicazione ricevente.

Receive Window Il TCP implementa questa funzionalità con una **variabile** detta **receive window**, mantenuta dal **mittente**. Questa variabile **fornisce un'idea di quanto spazio è ancora a disposizione** nel buffer del **ricevitore**. Tale valore è comunicato nel **campo window** dell'header TCP.



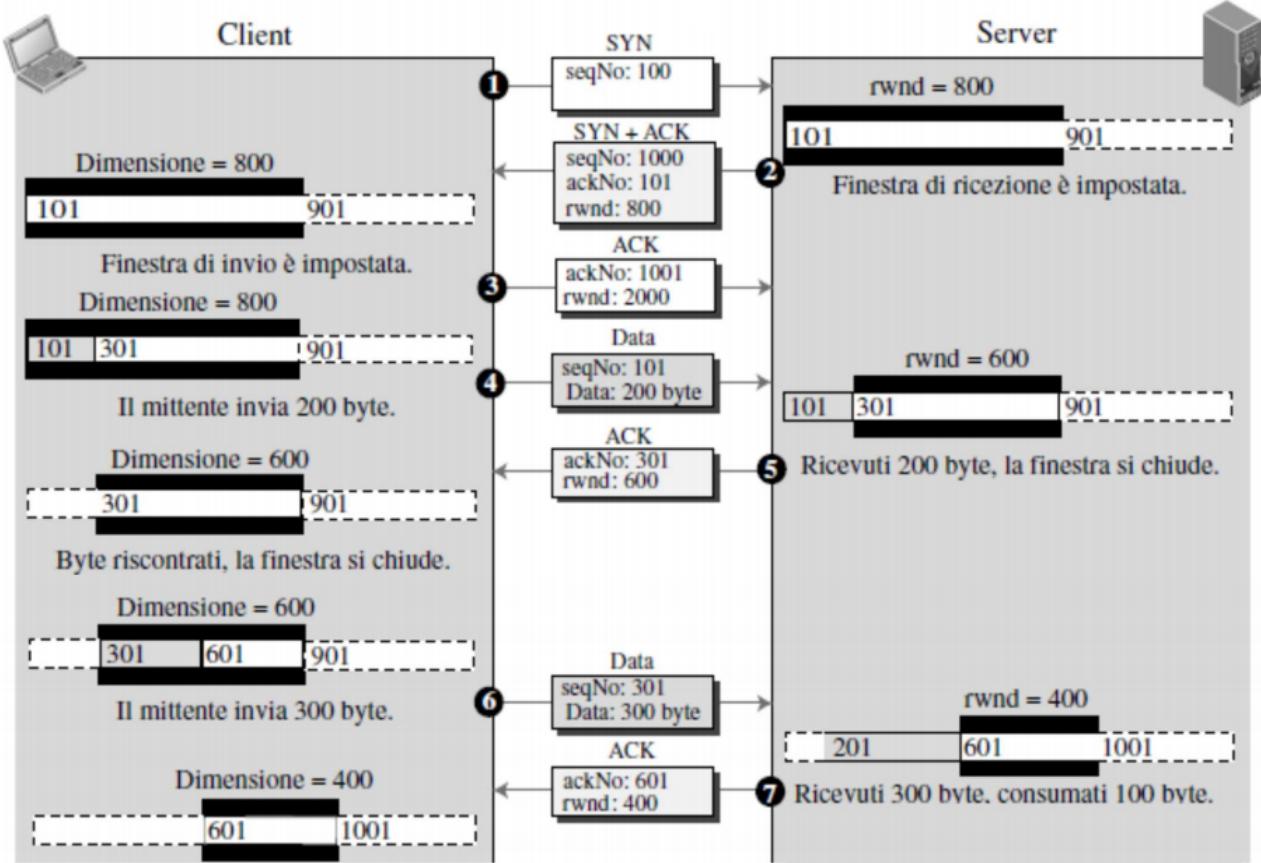
Lo spazio disponibile nel buffer del destinatario è:

$$\text{RcvWindow} = \text{RcvBuffer} - (\text{LastByteReceived} - \text{LastByteRead})$$

RcvWindow è dinamica, il destinatario comunica la dimensione di RcvWindow al mittente. Il mittente si assicura che:

$$\text{LastByteSent} - \text{LastByteAcked} < \text{RcvWindow}$$

Se RcvWindow = 0, il mittente manderà segmenti "sonda" di 1 byte per ricevere l'aggiornamento sulla dimensione di RcvWindow. Di seguito un esempio:



23.14 Controllo della congestione

RFC 2581

Origine Il fenomeno della congestione ha origine quando una o più delle sorgenti **tentano di richiedere più banda di quella disponibile sul percorso**. Troppe sorgenti che trasmettono troppi dati ad una velocità troppo alta, per cui la rete non può gestirli.

Il traffico eccessivo può provocare:

lunghi ritardi, accodamenti nei buffer dei router

perdita di pacchetti, overflow nei buffer dei router

Controlli Il protocollo TCP prevede il **controllo della congestione**, imponendo a ciascun mittente di **limitare la frequenza di invio di pacchetti sulla connessione**, in funzione della **congestione percepita**.

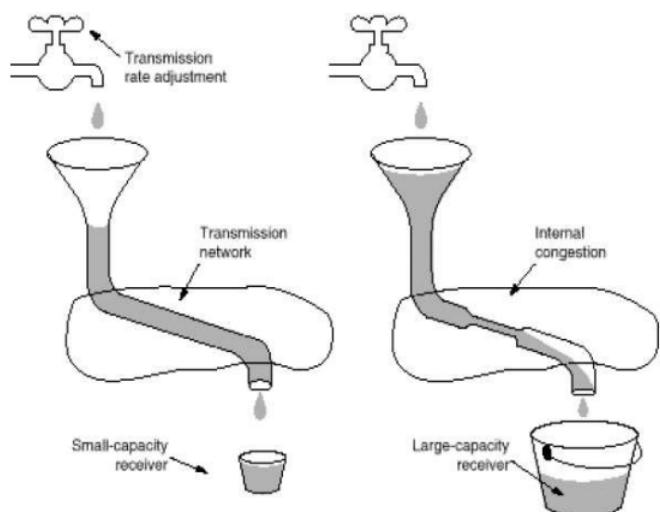
Capacità del TCP di adattarsi alla velocità di rete: se il TCP percepisce scarso traffico, aumenta la frequenza di invio, altrimenti la diminuisce.

Controllo di congestione punto-punto: nessun supporto esplicito della rete, la **congestione è dedotta dai sistemi terminali**

Come gestire entrambi i tipi di congestione?

Receiver window: dipende dalla dimensione del buffer di ricezione

Congestion window: basata su una stima della capacità di rete



I byte trasmessi corrispondono alla **dimensione della finestra più piccola**:

$$\text{dimFinestraInvio} = \min(rWnd, cWnd)$$

<https://www.docenti.unina.it/downloadPub.do?tipoFile=md&id=90675>

23.14.1 Algoritmo per il controllo della congestione

Algoritmo L'algoritmo che il **mittente TCP** utilizza per **regolare la propria frequenza di invio** in funzione della congestione rilevata, è costituito da **tre passi**:

1. **Partenza, slow start**
2. **AIMD: incremento additivo e decremento moltiplicativo**
3. **Ripresa veloce, fast recovery**
4. **Reazione ai time-out**

cWnd impone un **vincolo** alla frequenza di immissione del traffico sulla rete, in base alla congestione percepita.

23.14.2 cWnd

Soltamente è misurata in termini di **Maximum Segment Size, MSS**.

1 MSS è la **quantità massima di dati trasportabili da un segmento**.

Viene **determinato in base alla MTU** (Maximum Transfer Unit), cioè alla lunghezza massima del payload del frame di collegamento inviabile dall'host mittente

MSS scelto in modo tale che **il segmento TCP, incapsulato dentro il pacchetto IP, stia dentro un singolo frame** di collegamento.

RTT (Round Trip Time) è il **tempo impiegato da un segmento per effettuare il percorso di andata e ritorno**.

23.14.3 AIMD

Additive Increase Multiplicative Decrease

Il TCP del mittente **aumenta proporzionalmente la propria finestra di congestione ad ogni ACK ricevuto**. Di quanto aumenta?

Ad ogni ACK la cWnd viene incrementata in modo che si abbia una **crescita pari ad 1 MSS per ogni RTT (congestion avoidance)**

Incremento di $1\text{MSS} * (\text{MSS}/\text{cWnd})$

Ad esempio se $\text{cWnd} = 4\text{MSS}$ allora l'incremento è di $\text{MSS}/4$

Il TCP del mittente **dimezza la propria finestra di congestione ad ogni evento di perdita**, ad es. timeout o ACK duplicati.

Andamento della dimensione della cWnd nel tempo: andamento a "sawtooth"

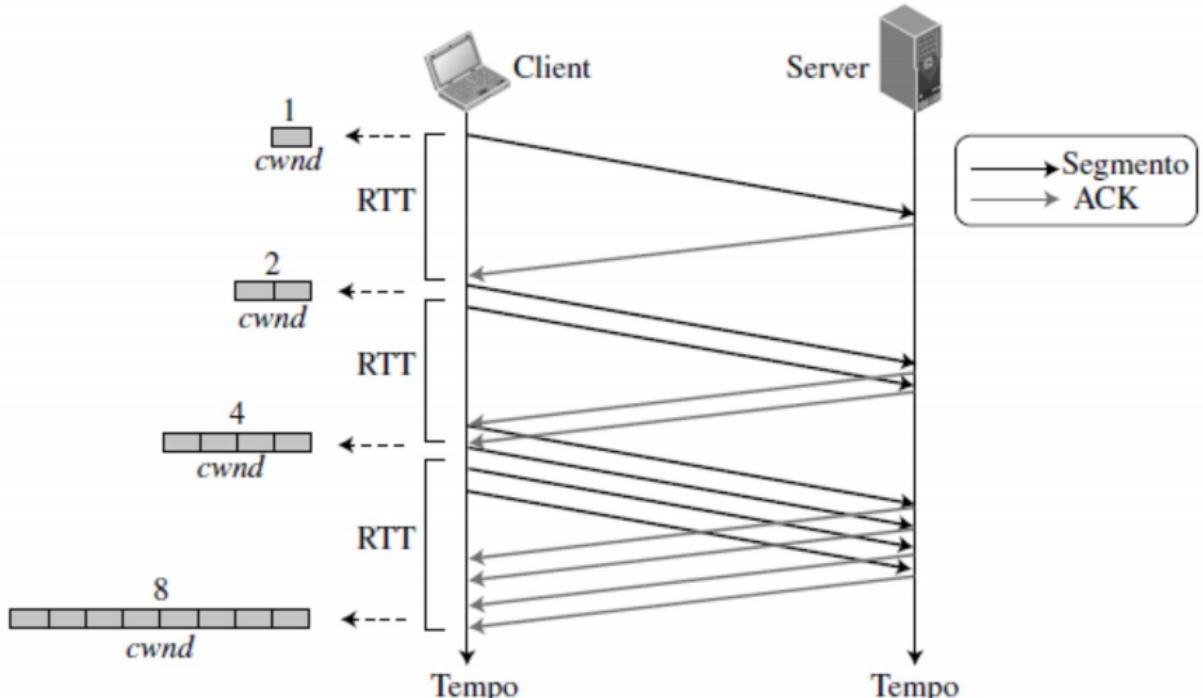


23.14.4 Slow Start

All'inizio, la finestra di congestione cWnd è posta pari a 1 MSS: equivale a dire che la **frequenza di invio è pari a 1 MSS/RTT**.

Esempio Se $MSS = 500\text{Byte}$ e $RTT = 200\text{ms}$ si ha una frequenza di invio di circa 20 Kb/s . Se ho una banda da 1 Mb/s impiegherà molto tempo per sfruttarla con un incremento lineare.

cWnd viene incrementata di 1 MSS ad ogni ACK. L'effetto è che cWnd **raddoppia ad ogni RTT** avendo così una **crescita esponenziale** fino ad un errore, poi cWnd dimezza.



23.14.5 Politica Reno per il controllo della congestione

Soglia Viene definita una variabile "soglia", alla quale è assegnato un valore alto (ad esempio 64 Kb). Questa soglia determina quando termina la slow start ed inizia la congestion avoidance (AIMD).

cWnd < soglia: cWnd aumenta esponenzialmente (*slow start*)

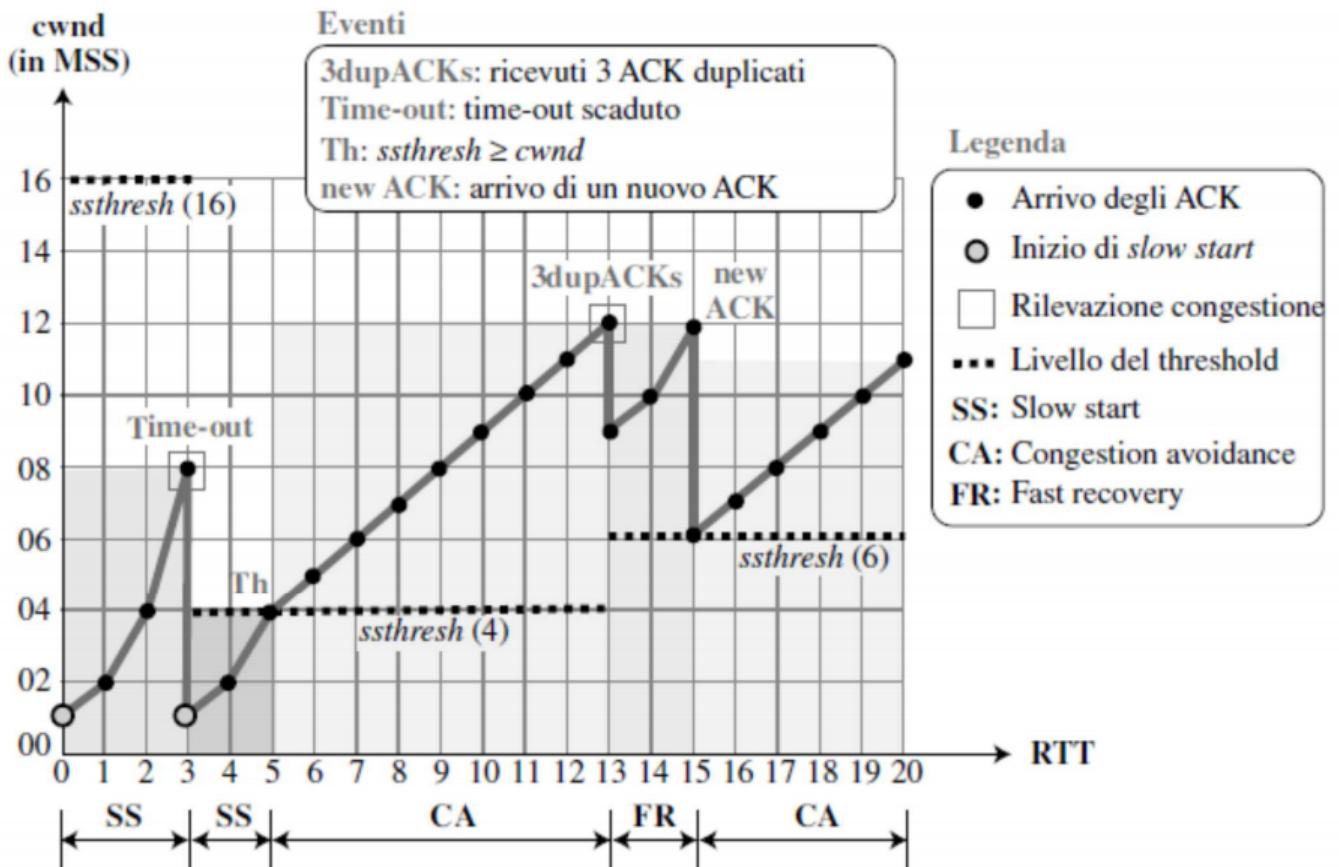
cWnd > soglia: cWnd aumenta linearmente (*AI*)

Evento di perdita Se ho 3 ACK duplicati pongo prima la soglia a metà di cWnd e poi $cWnd = \text{soglia} + 3 \text{ MSS}$ (*fast recovery*)

Se ho un ACK perso per timeout pongo la soglia a metà di cWnd e pongo $cWnd = 1 \text{ MSS}$ (*slow start*)

Nella fast recovery Se avviene un timeout si va in **slow start**, se arriva un ACK non duplicato si va in **congestion avoidance**.

Esempio di controllo della congestione con TCP Reno



TCP Reno



23.14.6 Politica Tahoe per il controllo della congestione

Versione precedente della politica Reno: timeout e 3 ACK duplicati sono trattati allo stesso modo.





23.15 Throughput

Indicando con W il **valore massimo di byte della finestra** (ovvero quando si verifica l'errore), il **TCP in regime stazionario** offre il seguente throughput medio:

$$\text{throughput} = \frac{0.75 * W}{RTT}$$

Calcolo macroscopico che ignora le fasi di slow start. Con W la dimensione della finestra alla perdita:

Quando la finestra è W , il throughput è $\frac{W}{RTT}$

Appena dopo la perdita, la finestra è ridotta a $\frac{W}{2}$ e il throughput a $\frac{W}{2 * RTT}$

Throughput medio: $0.75 * RTT$

23.16 Fairness

Ipotesi

K connessioni TCP esistono su un **unico link di capacità R bit/s**

Non ci sono altri protocolli esistenti sullo stesso link

Risultato Ognuna delle connessioni TCP tende a trasmettere a R/K bit/s.

23.17 Transmission Control Block

Struttura dati Poiché ogni connessione è distinta, bisogna **mantenere delle informazioni** su ogni connessione separatamente.

TCB Il TCP usa una struttura dati speciale per questo scopo, chiamata **Transmission Control Block** o TCB. Il TCB contiene tutte le informazioni a proposito della connessione, come **i due numeri di socket** che la identificano e **i puntatori ai buffer** dove vengono mantenute le informazioni in arrivo e in partenza.

Il TCB è anche usato per **implementare il meccanismo delle sliding windows**. Mantiene delle variabili che tengono traccia dei **numeri di byte ricevuti e riscontrati**, **byte ricevuti e non riscontrati**, **dimensione attuale della finestra** e così via.

Ovviamente, **ogni dispositivo mantiene il proprio TCB per la connessione**.

24 Livello di Rete

Il livello di rete si occupa di **realizzare una connessione logica fra host diversi**: interconnessione di reti eterogenee.

Internet è una **rete logica** costituita da un **insieme di reti fisiche**. Il livello di rete offre un'astrazione che consente ad host e reti eterogenee di funzionare dal punto di vista logico come una singola rete.

Offre **servizi** al livello di trasporto.

Utilizza **servizi** del livello datalink.

24.1 Servizi

Suddivisione in **pacchetti**

Routing (instradamento)

Forwarding (inoltro)

24.2 Architettura di rete IP



Dal mittente, il livello di rete **riceve i segmenti dal livello di trasporto** e li **incapsula nei datagrammi**.

Dal destinatario, **consegna i segmenti al livello di trasporto** (demux TCP o UDP)

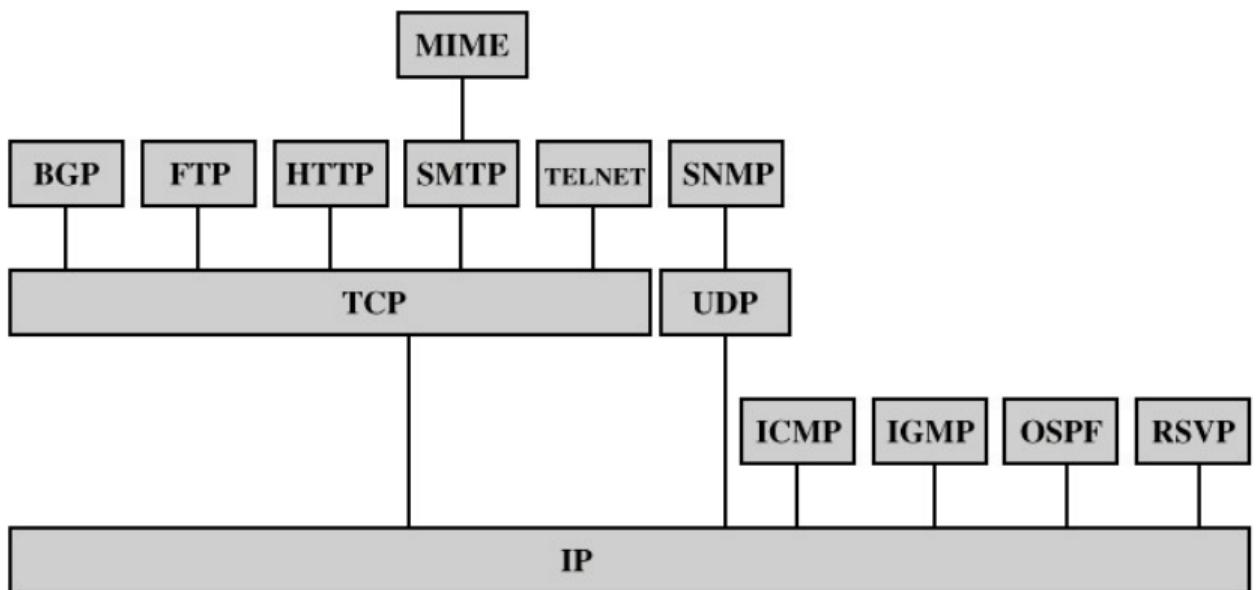
I protocolli del livello di rete sono in ogni host e router. Il router esamina i campi intestazione in tutti i datagrammi IP che lo attraversano, e li inoltra da un collegamento in entrata ad uno in uscita.

24.3 Protocol Data Units



© William Stallings Data and Computer Communications 7th Edition

24.4 Alcuni protocolli della Suite TCP/IP



© William Stallings Data and Computer Communications 7th Edition

BGP = Border Gateway Protocol	OSPF = Open Shortest Path First
FTP = File Transfer Protocol	RSVP = Resource ReSerVation Protocol
HTTP = Hypertext Transfer Protocol	SMTP = Simple Mail Transfer Protocol
ICMP = Internet Control Message Protocol	SNMP = Simple Network Management Protocol
IGMP = Internet Group Management Protocol	TCP = Transmission Control Protocol
IP = Internet Protocol	UDP = User Datagram Protocol
MIME = Multi-Purpose Internet Mail Extension	

BGP = Border Gateway Protocol	OSPF = Open Shortest Path First
FTP = File Transfer Protocol	RSVP = Resource ReSerVation Protocol
HTTP = Hypertext Transfer Protocol	SMTP = Simple Mail Transfer Protocol
ICMP = Internet Control Message Protocol	SNMP = Simple Network Management Protocol
IGMP = Internet Group Management Protocol	TCP = Transmission Control Protocol
IP = Internet Protocol	UDP = User Datagram Protocol
MIME = Multi-Purpose Internet Mail Extension	

25 IP

Internet Protocol RFC 791 Caratteristiche:

Connection less: non c'è circuito virtuale né fisico fra i due host a livello IP, non c'è garanzia di sequenzialità.

Non affidabile: non prevede meccanismi di controllo dell'errore. "Send and pray".

Best effort: non prevede garanzie sulla QoS (Quality of Service), sul tempo di consegna dei datagrammi e sul controllo di flusso.

I modelli di QoS vengono implementati dai router (es: DiffServ)

25.1 Funzioni

25.1.1 Inoltro

Forwarding Trasferimento del pacchetto sull'appropriato collegamento d'uscita



25.1.2 Instradamento

Routing Processo decisionale di scelta del percorso verso una destinazione (**algoritmi di routing** o algoritmi di instradamento)

25.1.3 Indirizzamento

Strumento per identificare gli host nella rete interconnessa

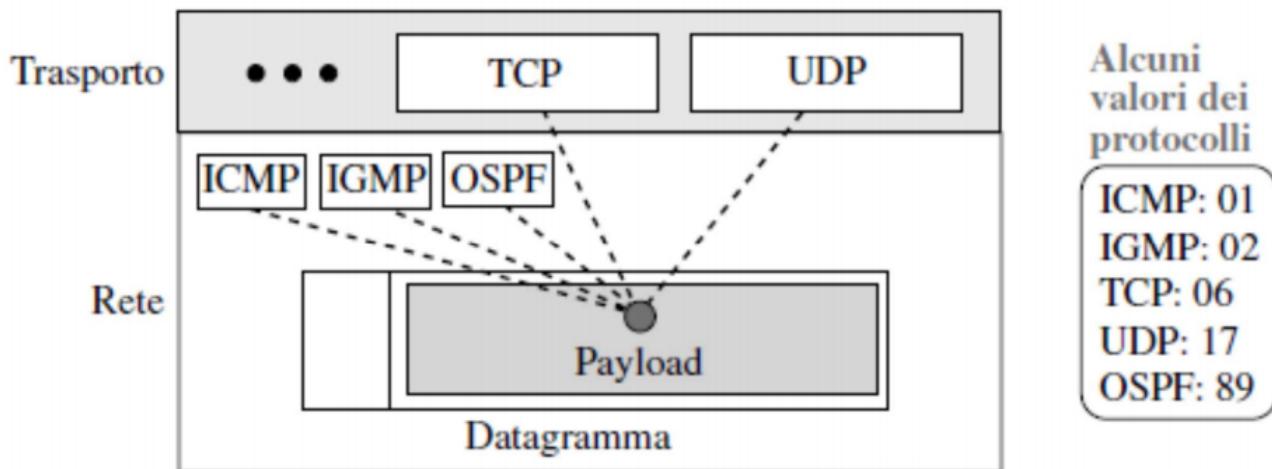
Modello datagram Il modello datagram (senza connessione) per la consegna dei dati:

Segmentazione e riunificazione dei pacchetti (adattamento al data link)

Controllo degli errori solo dell'header

Verifica TTL

25.2 Multiplexing / Demultiplexing



25.3 Datagramma IP

25.3.1 Formato



25.3.2 Campi

Versione, 4 bit: specifica la versione usata. Attualmente IPv4 o IPv6.

Hlen, 4 bit: lunghezza dell'header espressa in parole da 32 bit. Tipicamente vale 0101, cioè 20 Byte.

Tipo di servizio, 8 bit: serve per "colorare" il datagramma IP per classificarlo in base alle necessità (basso ritardo, affidabilità...) e gestirlo appositamente. Nella RFC 791 c'è una prima classificazione, ma si veda anche RFC2474.

Lunghezza del datagramma, 16 bit: lunghezza di tutto il datagramma in byte, header incluso. Max 65535 Byte (nella pratica, max 1500 Byte).

Identificazione, flag, offset: campi per la frammentazione

Tempo di vita, 8 bit: ad ogni passaggio da un router viene decrementato, quando raggiunge lo 0 viene scartato. Assicura che eventuali circuiti ad anello non provochino traffico perpetuo sulla rete.

Il valore dipende dal S.O., esempi: 30, 64, 128, 255.

Protocollo, 8 bit: in ricezione indica quale protocollo dello strato superiore deve ricevere i dati. Inizialmente la tabella era nella RFC 1700, attualmente è un database online su iana.org.

Checsum dell'intestazione, 16 bit: ad ogni router (TTL cambia ad ogni hop) viene calcolato il checksum della sola intestazione ponendo questo campo pari a 0. Se si ottiene un errore si scarta il datagramma.

Indirizzi sorgente e destinazione, 32 bit: indirizzi IP mittente e destinatario

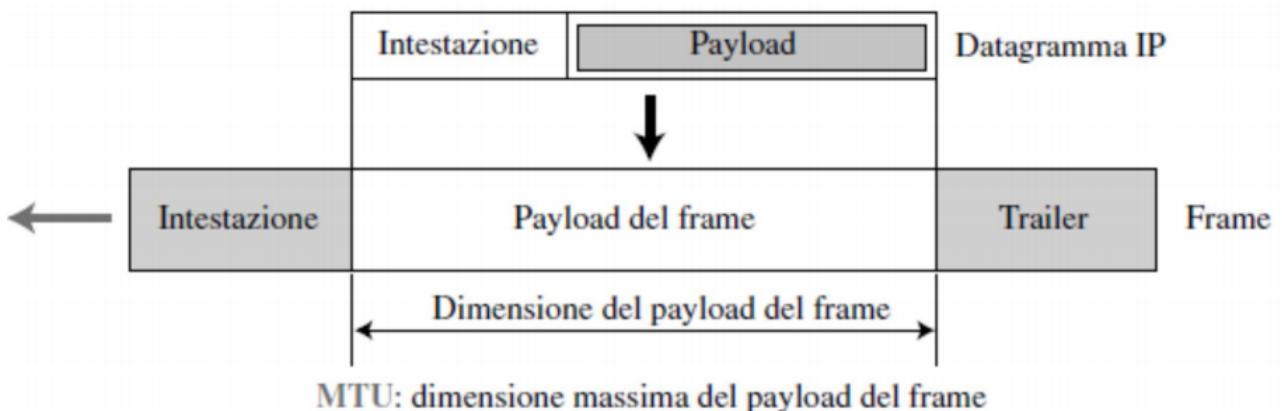
Opzioni, variabile multiplo di 32 bit: principalmente usati a scopo di test o debug

Dati: i dati effettivamente trasportati

25.4 Frammentazione

La MTU pone un limite alla lunghezza dei datagrammi IP e tratte diverse possono porre limiti diversi.

Maximum Transfer Unit è la quantità massima di dati trasportata dal protocollo di collegamento (**in un frame**). Può variare da una tecnologia di collegamento all'altra. Tipicamente è 1500 Byte, da cui bisogna togliere i 20 Byte minimi per l'intestazione IP lasciando **1480 Byte di payload**.



Meccanismo di frammentazione Se un router **riceve un datagramma la cui dimensione supera l'MTU della rete su cui deve spedirlo, lo spezza** in due o più datagrammi più piccoli detti **frammenti**.

Il riassemblaggio è effettuato dall'entità rete nel sistema destinatario. **Se qualche frammento non arriva a destinazione si butta via tutto il datagramma**.

Ciascun frammento è un datagramma IP indipendente e completo, che viene trasmesso attraverso una serie di reti fisiche indipendentemente dagli altri.

Ricomporre i frammenti Per ricomporre si usano i flag **identificazione**, **flag** e **offset**.

Identificazione, 16 bit: identificatore associato al datagramma dall'host sorgente. Associato a IP sorgente e destinazione **identifica quel datagramma in un intervallo di tempo ragionevolmente lungo**. I frammenti mantengono il valore di questo campo e il destinatario riconosce i frammenti che vanno assemblati insieme.

Offset, 13 bit: indica la posizione relativa come multiplo di 8 Byte. Serve ad ordinare i frammenti nell'assemblaggio.

I frammenti **devono essere multipli di 8 Byte**, ed il primo frammento ha offset = 0.

Flag, 3 bit: serve ad identificare l'ultimo frammento. 3 bit

Il bit 0 è **riservato**, per ora vale sempre 0

Il bit 1 è **do not fragment**. Vale

0, se il pacchetto **può** essere frammentato

1, se il pacchetto **non può** essere (ulteriormente) frammentato

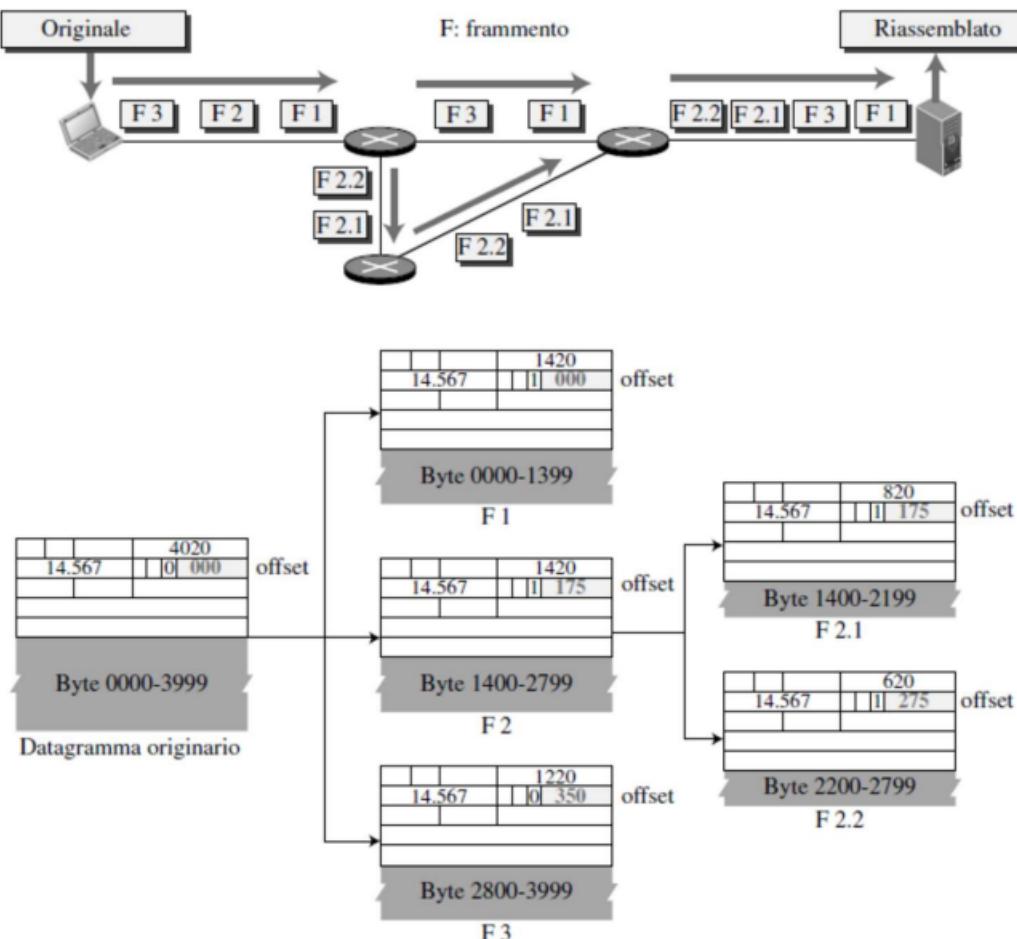
Il bit 2 è **more fragments**. Vale

0, se il pacchetto **è l'ultimo** frammento

1, se il pacchetto **non è l'ultimo** frammento

Ripetuto Il processo di frammentazione **può essere ripetuto**, ad es. nel caso in cui un frammento viene inoltrato su un collegamento con MTU ancora più piccolo.

Critico La ricostruzione del datagramma è un **processo critico**, richiede risorse e se un frammento manca il datagramma è scartato.



26 Indirizzamento IP

Ogni host è connesso ad internet attraverso un'**interfaccia di rete**, confine fra host e il collegamento su cui vengono inviati di datagrammi. Ad **ogni interfaccia è assegnato un indirizzo IP**.

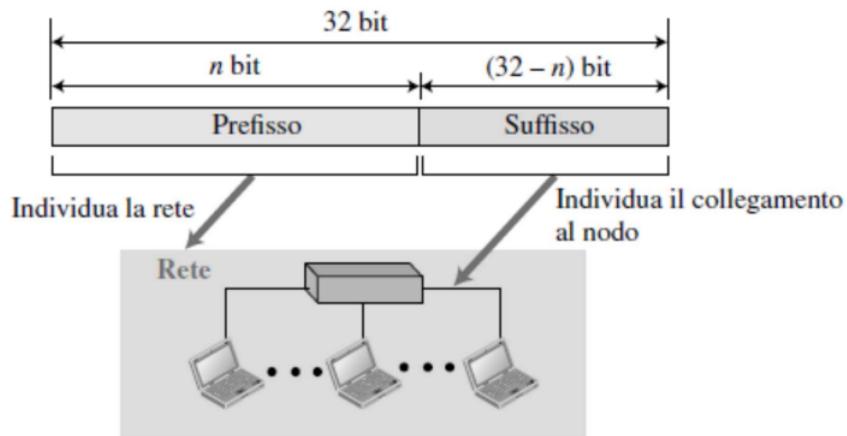
I **router devono necessariamente essere connessi ad almeno due collegamenti**.

26.1 Indirizzi IPv4

CIDR, RFC 1519

Gli indirizzi IP sono costituiti da **32 bit**, cioè 4 Byte, rappresentati in **notazione decimale puntata**.
Es 10000000 00001011 00000011 00011111 → 128.11.3.31

Divisione Ogni host ha un **indirizzo univoco sulla rete**, diviso in due parti, **network ID** e **host ID**, che identificano una rete IP su internet e l'host in quella rete IP.



La rete Net x ha quindi $2^{32-n} - 2$ host (bisogna escludere l'indirizzo di rete .0 e l'indirizzo di broadcast .255)
Una **sottorete IP** o **rete IP**, è una rete i cui terminali sono collegati alle interfacce di router ed host.

26.2 Strategie di addressing

Il **classful addressing** è poco flessibile, si può optare quindi per un **classless addressing**.

Inoltre, gli **indirizzi IP si stanno progressivamente esaurendo**. La specifica IPv6 ha indirizzi di 128 bit invece che di 32 bit, e potrebbe risolvere questo problema.

26.2.1 Classful addressing



5 classi di indirizzi IP:

Classe **A**: (0.0.0.0 → 127.255.255.255) 7 bit per 128 reti, 24 bit per 16M di host

Classe **B**: (128.0.0.0 → 191.255.255.255) 14 bit per 16k reti, 16 bit per 64k host

Classe **C**: (192.0.0.0 → 223.255.255.255) 21 bit per reti e 8 bit per 256 host

Classe **D**: (224.0.0.0 → 239.255.255.255) riservata a **multicasting**

Classe **E**: (240.0.0.0 → 255.255.255.255) riservata per **usi futuri**

15.10.10.90: classe A (primo numero tra 0 e 127), quindi 10.10.90 indica l'host

130.250.42.53: classe B (primo numero tra 128 e 191), quindi 42.53 indica l'host

196.234.12.14: classe C (primo numero tra 192 e 223), quindi 14 indica l'host

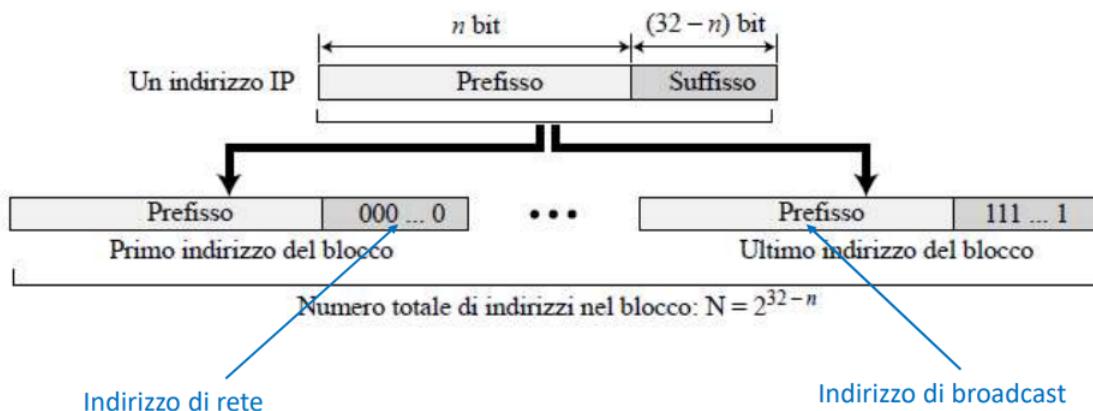
26.2.2 Classless addressing

byte.byte.byte.byte/n, con n lunghezza del prefissi, cioè gli n bit più a sinistra costituiscono il network ID.

12.24.76.8/8

23.14.67.92/12

220.8.24.255/25



26.3 Subnet Mask

La **maschera di sottorete**, o **subnet mask**, distingue **quale parte di un indirizzo IP identifica la rete** e quale l'host.

Ad esempio, all'indirizzo IP 150.217.8.42 è definita la mask 255.255.255.0, quindi la rete o subnet effettiva ha indirizzo 150.217.8.0/24 (i primi 24 bit sono il network ID).

Per determinare l'indirizzo di rete il computer deve effettuare un **AND**.

Indirizzo IP	150.217.8.42	10010110 11011001 00001000 00101010
Subnet mask	255.255.255.0	11111111 11111111 11111111 00000000
AND	150.217.8.0	10010110 11011001 00001000 00000000

26.4 Indirizzi speciali

This Host 0.0.0.0

Usato quando un host ha necessità di inviare un datagramma ma **non conosce il proprio IP**, cioè l'indirizzo sorgente.

Limited Broadcast 255.255.255.255

Usato da router ed host che devono **inviare un datagramma a tutti i dispositivi che si trovano nella rete**.

Il router blocca la propagazione alla **sola rete locale**.

Loopback 127.0.0.1

Il datagram con questo indirizzo di destinazione **non lascia l'host locale (loopback)**.

Usato per test e debug.

Indirizzi Privati

Sono quattro blocchi **riservati per indirizzi privati per reti locali**.

10.0.0.0/8, 172.16.0.0/12, 192.168.0.0/16, 169.254.0.0/16

Indirizzi Multicast 224.0.0.0/4

26.5 Assegnazione degli indirizzi

Partizionare gli indirizzi Se si vuole fare una rete di n host, usare la subnet mask che fornire la potenza di 2 più piccola ma maggiore di n possibile.

Es. 190 host → 256

Per partizionare ulteriormente partire dalla partizione che ha necessità del numero di host più grande e procedere come sopra.

Un IP può essere assegnato all'interfaccia di un host **attraverso due distinte modalità**:

Configurazione manuale

L'amministratore configura direttamente nell'host l'IP ed inserisce ulteriori informazioni di servizio (gateway, netmask...)

DHCP

L'host ottiene il proprio indirizzo e le altre informazioni in modo automatico

27 DHCP

Protocollo Il Dynamic Host Configuration Protocol è un protocollo client-server che assegna indirizzi IP temporanei.

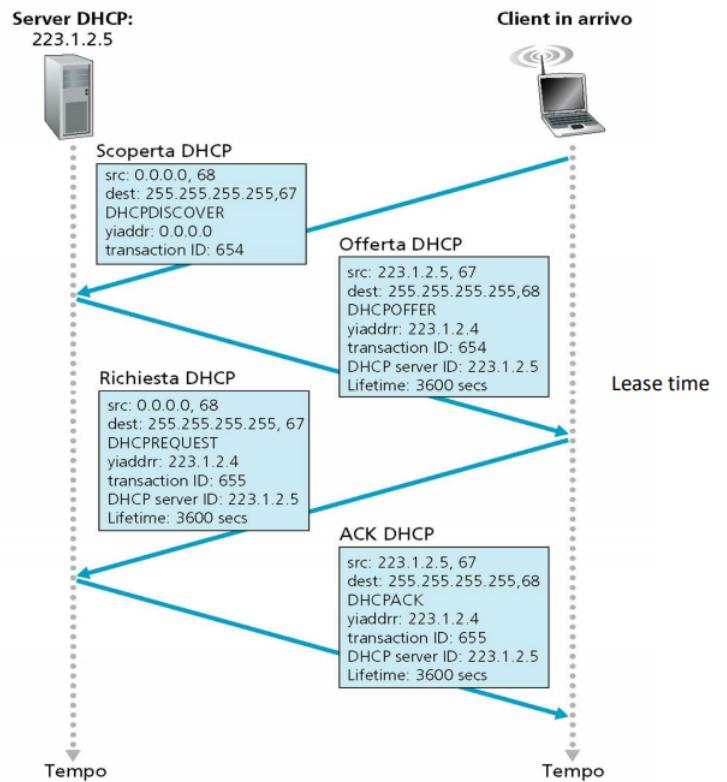
Funzionamento Un host appena connesso alla rete è il **client**, che dà inizio alla procedura di assegnazione verso il server DHCP della sottorete a cui è inserito:

1. Identificazione del server DHCP

2. Offerta DHCP

3. Richiesta DHCP

4. Conferma DHCP



28 IP Forwarding

Ogni datagramma IP è soggetto a **forwarding** da parte dell'host di origine e del router che sta attraversando.

28.1 Forwarding Diretto

Il pacchetto IP ha come **destinazione un host nella propria rete IP**

L'invio è diretto sul destinatario

L'indirizzo di destinazione a livello link è quello del destinatario (MAC address)

Non viene interpellata nessun'altra entità

28.2 Forwarding Indiretto

Il pacchetto IP ha come **destinazione un host di un'altra rete IP**

Viene delegato l'invio ad "un altro", cioè un router

L'indirizzo di destinazione a livello link è quello del router

In entrambi i casi, le **condizioni necessarie** perché tutto funzioni sono che:

Esista un cammino funzionante e diretto, a livello data link, tra tutti gli host che appartengono ad una stessa sottorete

Ogni host coinvolto abbia un indirizzo IP corretto, cioè con **uguale Net ID** (cioè stessa sottorete) e con **Host ID univoco** nella sottorete.

Le due condizioni insieme diventano **condizione necessaria e sufficiente** perché la comunicazione funzioni.

29 NAT

Il Network Address Translation (RFC 2663, RFC 3022) permette di **trasmettere su internet il traffico proveniente da sistemi attestati su sottoreti private**, in cui sono assegnati indirizzi IP privati.

L'accesso di una rete privata ad internet è realizzato attraverso un router abilitato alla NAT, il quale ha un indirizzo IP pubblico unico.

Tutto il traffico in uscita dal router di accesso ha un unico indirizzo IP sorgente pubblico (quello del router).
Tutto il traffico in ingresso alla subnet ha indirizzo IP di destinazione del router di accesso.

Funzionamento Il router di accesso **mantiene in memoria una tabella di traduzione NAT**, le cui righe contengono indirizzi IP privati e numeri di porta.



1. Host 10.0.0.1 manda un datagramma a 128.119.40.186:80
2. Router NAT **traduce** l'indirizzo di origine del datagramma da 10.0.0.1:3345 a 138.76.29.7:5001 e aggiorna la tabella
3. Arriva la risposta, l'indirizzo di destinazione è 138.76.29.7:5001
4. Router NAT **traduce** l'indirizzo di destinazione del datagramma da 138.76.29.7:5001 a 10.0.0.1:3345

30 ICMP

L'Internet Control Message Protocol (RFC 792) è usato da host e router per scambiarsi messaggi di errore o altre situazioni che richiedono intervento. I messaggi ICMP sono incapsulati in datagrammi IP, ma viene comunque considerato parte integrante dello strato di rete. Chi implementa IP deve implementare anche ICMP. Quando un host destinatario deve informare il mittente di errori o eventi avvenuti nell'inoltro di un pacchetto IP, utilizza il protocollo ICMP.

I pacchetti ICMP vengono instradati dai router prima dei pacchetti IP ordinari.

Per i pacchetti IP frammentati, i messaggi ICMP sono relativi solo al frammento 0. Inoltre, i messaggi ICMP non sono mai inviati in risposta a pacchetti IP con indirizzo mittente che non rappresenta un host unico (0.0.0.0, 127.0.0.1, broadcast).

I messaggi ICMP non sono mai inviati in risposta a messaggi di errore ICMP, ma possono essere inviati in risposta a messaggi ICMP di interrogazione.

30.1 Formato dei messaggi



Messaggio di segnalazione errori



Messaggio di richiesta

Valori del tipo e dei codici:

Messaggi di segnalazione errori

- 03: destinazione non raggiungibile (codici da 0 a 15)
- 04: source quench (solo codice 0)
- 05: reindirizzamento (codici da 0 a 3)
- 11: tempo scaduto (codici 0 e 1)
- 12: problema ad un parametro (codici 0 e 1)

Messaggi di richiesta

- 08 and 00: richiesta e risposta eco (solo codice 0)
- 13 and 14: richiesta e risposta timestamp (solo codice 0)

31 Ping

Una delle applicazioni che un host può utilizzare per verificare il funzionamento di un altro host è il **ping**.

Il programma ping si basa sui messaggi di richiesta e risposta eco dell'ICMP. Un host invia una richiesta eco (tipo 8, codice 0) a un altro host che, se attivo, può rispondere con una risposta eco (tipo 0, codice 0). Fornisce anche misure dell'RTT.

In maniera grossolana, ping può anche misurare l'affidabilità e la congestione del router tra due host inviando una sequenza di messaggi richiesta e risposta.

32 Traceroute

Il programma **traceroute** in UNIX (o **tracert** in Windows) può essere utilizzato per **individuare il percorso di un datagramma dalla sorgente alla destinazione** tramite l'identificazione dell'indirizzo IP di tutti i router che vengono visitati lungo il percorso.

Solitamente il programma viene impostato per un massimo di 30 salti, che sono usualmente sufficienti per raggiungere la destinazione.

Funzionamento Traceroute ha un funzionamento molto diverso da ping. Ping è basato su due messaggi query. Traceroute invece è implementato per mezzo di due messaggi di segnalazione degli errori: tempo scaduto e destinazione non raggiungibile.



Traceroute **imposta inoltre un timer per trovare il tempo di round-trip** di ciascun router e della destinazione. La maggior parte dei programmi traceroute **invia tre messaggi** a ogni dispositivo, con lo stesso TTL, per poter effettuare una stima migliore del tempo di round-trip.

Di seguito un esempio di funzionamento con tre messaggi per ogni dispositivo e quindi tre RTT ottenuti:

\$ traceroute printers.com				
traceroute to printers.com (13.1.69.93), 30 hops max, 38 byte packets				
1 route.front.edu	(153.18.31.254)	0.622 ms	0.891 ms	0.875 ms
2 ceneric.net	(137.164.32.140)	3.069 ms	2.875 ms	2.930 ms
3 satire.net	(132.16.132.20)	3.071 ms	2.876 ms	2.929 ms
4 alpha.printers.com	(13.1.69.93)	5.922 ms	5.048 ms	4.922 ms

33 Routing

33.1 Router

Fin'ora abbiamo considerato il **router** come una **scatola che accetta pacchetti in entrata** da una delle interfacce, **usa una tabella d'inoltro per trovare la porta di output** e da quest'ultima **invia il pacchetto**.

Il router può essere suddiviso in **due parti**:

Data Plane o Forwarding Plane

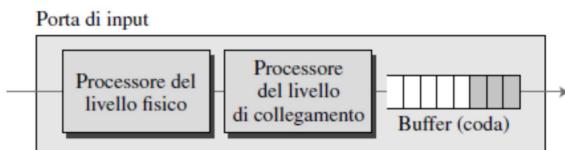
Vengono analizzati i pacchetti in entrata e instradati secondo la propria **tabella di forwarding** e l'**indirizzo di destinazione**, stabilendo il next hop. Il *pacchetto attraversa il router*.

Control Plane

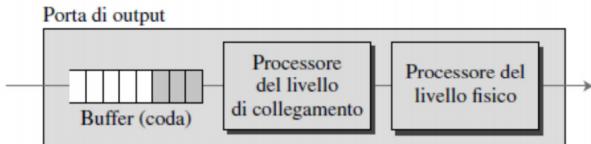
I pacchetti sono **destinati o originati localmente**. Vengono **processati dal router per aggiornare le proprie informazioni di routing**.

Sono individuate quindi una **serie di componenti** di un router:

Porte di input

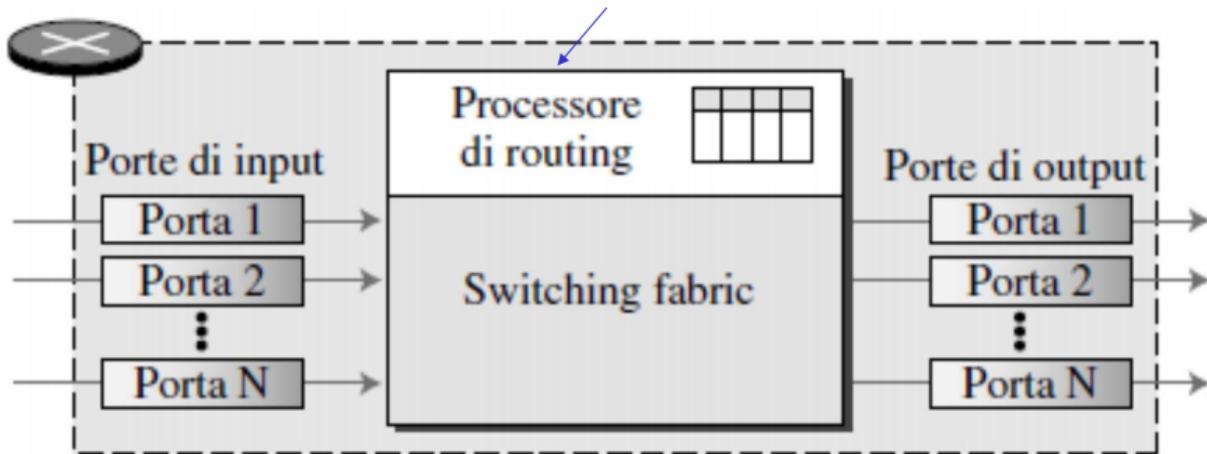
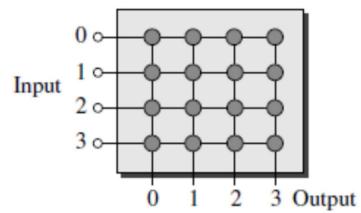


Porte di output



Processore di routing

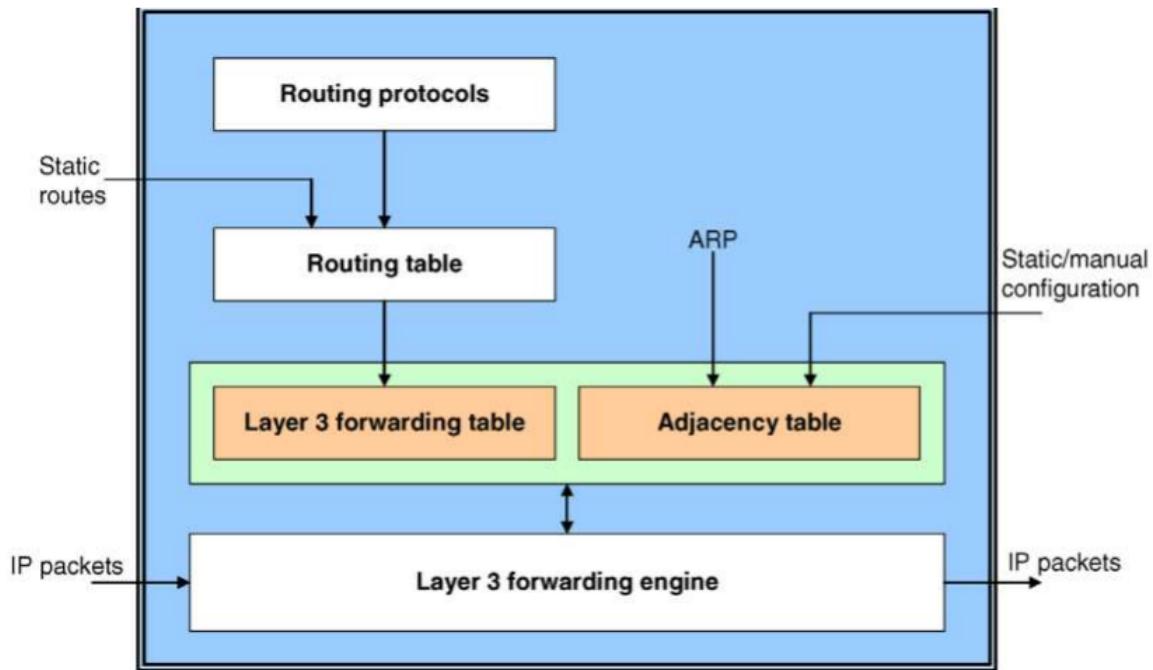
Switching fabric, esempio: crossbar switch



33.2 Routing Engine

Protocolli di Routing (+ static routes) → Routing Table → Tabella d'Inoltro del livello 3

33.3 Forwarding Engine



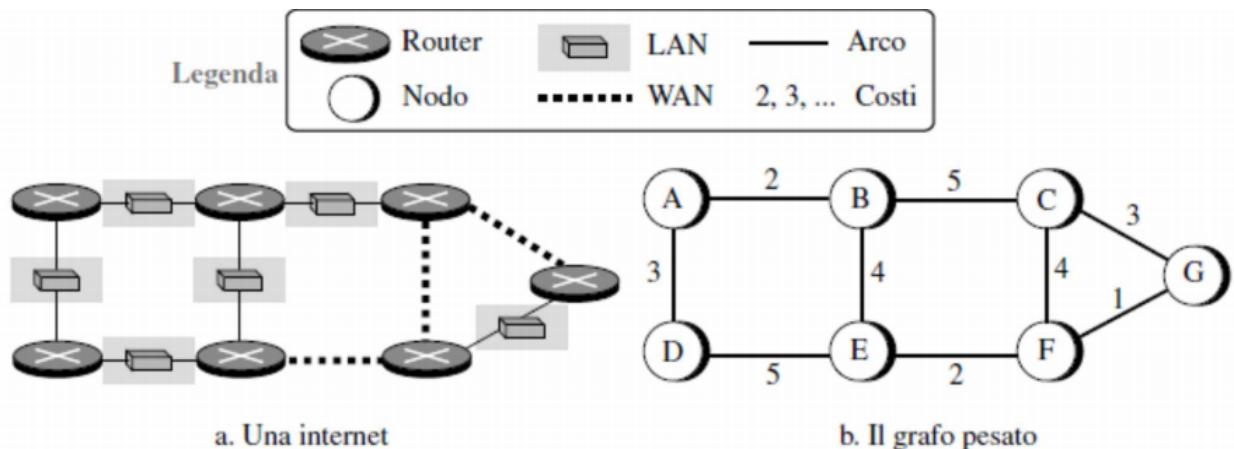
33.4 Algoritmi di Routing

In una rete come Internet lo scopo del livello di rete è quello di **consegnare un datagramma dalla sorgente alla/e destinazione/i**.

Routing Unicast: uno-a-uno, un datagramma è destinato ad una sola destinazione

Routing Multicast: uno-a-molti, un datagramma è destinato ad un gruppo di host

Rappresentiamo una rete di router **come un grafo** con **costo associato agli archi** = la distanza fra i nodi, il numero di hop...



33.4.1 Tipologie

Routing Statico In questo tipo di routing, le **entry** della tabella vengono configurate manualmente dall'operatore. Questo metodo viene usato per reti di piccole dimensioni, la cui topologia è poco variabile e dove è possibile prevedere tutti i possibili percorsi di un pacchetto IP nella rete.

Routing Dinamico Con il routing dinamico si rendono necessari **protocolli specifici** che provvedono automaticamente ad inserire nella tabella di routing le **entry relative ai possibili percorsi**. Viene usato per reti medie-grandi ed a **topologia variabile** come grandi LAN private ed Internet.

Algoritmi Gli algoritmi di routing si possono classificare in:

Globali, se si basano sulla conoscenza della topologia di tutta la rete. Il **calcolo** può essere fatto in un **unico sito (algoritmo centralizzato)** o in più nodi, utilizzando informazioni sulla connettività di tutti i nodi e sui costi di tutti i link. L'algoritmo riceve quindi in ingresso tutti i collegamenti tra i nodi ed i loro costi.

Esempio: **algoritmo Link State**

Decentralizzati, quando **nessun nodo conosce la topologia di tutta la rete** ma ha **informazioni solo sui nodi ed i collegamenti vicini**. Il **calcolo** del percorso è **iterativo distribuito**. Ogni nodo elabora un vettore di stima dei costi (distanza) verso tutti gli altri nodi della rete. Il cammino a costo minimo viene calcolato in modo iterativo distribuito.

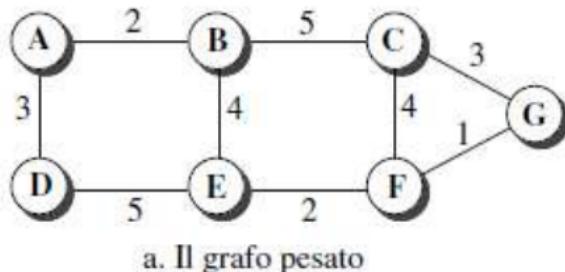
Esempio: **algoritmo Distance Vector**.

33.4.2 Algoritmo Link State

Premesse La topologia di rete e tutti i costi dei collegamenti sono noti a tutti i nodi, attraverso il link-state broadcast. Inoltre, tutti i nodi dispongono delle stesse informazioni.

Obiettivo Calcola il **cammino a costo minimo** da un nodo **origine a tutti i nodi della rete**, e crea una tabella d'inoltro per quel nodo, sfruttando l'**Algoritmo di Dijkstra**: iterativo, alla k-esima iterazione sono noti i cammini a costo minimo fino a k nodi.

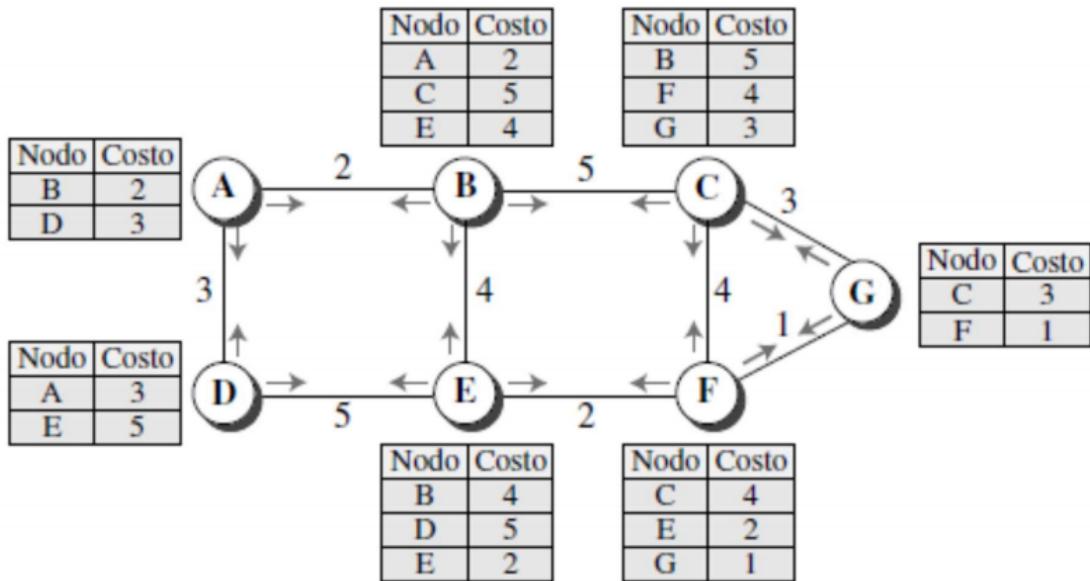
Link-State Database



	A	B	C	D	E	F	G
A	0	2	∞	3	∞	∞	∞
B	2	0	5	∞	4	∞	∞
C	∞	5	0	∞	∞	4	3
D	3	∞	∞	0	5	∞	∞
E	∞	4	∞	5	0	2	∞
F	∞	∞	4	∞	2	0	1
G	∞	∞	3	∞	∞	1	0

b. Il link-state database

Link-State Packet Vengono creati ed inviati da ciascun nodo.



Notazione

$C(x, y)$: costo del collegamento dal nodo x al nodo y
 $= \infty$ se non sono adiacenti.

$D(v)$: costo del cammino dal nodo origine alla destinazione v per quanto riguarda l'iterazione corrente.

$p(v)$: immediato predecessore di v lungo il cammino.

N^* : sottoinsieme di nodi per cui il cammino a costo minimo dall'origine è definitivamente noto.

```

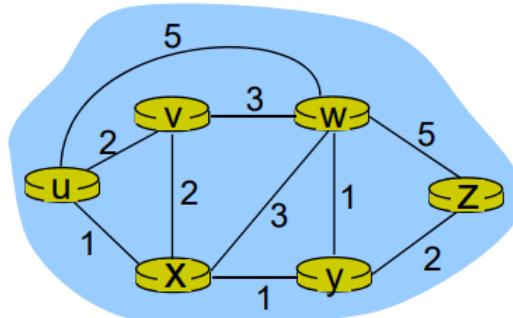
DijkstraAlgorithm() {
    // Inizializzazione
    N* = {u}
    per tutti i nodi v
        if (v adiacente a u) D(v) = C(u, v)
        else D(v) = inf

    // Loop
    while (fino a che tutti i nodi non sono in N*) {
        trova w non in N* tale che D(w) minimo
        aggiungi w a N*
        aggiorna D(v) per tutti i v adiacenti a w e non in N*
        D(v) = min{D(v), D(w) + c(w, v)}
        // Il nuovo costo verso v = il vecchio costo verso v, oppure il costo del
        // cammino minimo noto verso w + il costo da w a v
    }
}

```

Esempio

Step	N'	D(v),p(v)	D(w),p(w)	D(x),p(x)	D(y),p(y)	D(z),p(z)
0	u	2,u	5,u	1,u	∞	∞
1	ux	2,u	4,x		2,x	∞
2	uxy	2,u	3,y			4,y
3	uxyv		3,y			4,y
4	uxyvw					4,y
5	uxyvwz					



33.4.3 Distance Vector Routing

Rappresentazione grafica dell'equazione di Vettori distanza iniziali dei nodi di una rete. **Bellman-Ford.** Aggiornamento di un percorso usando un nodo intermedio.

$$D_{XY} = \min(D_{XY}, (C_{XZ} + D_{ZY}))$$



Aggiornamento dei vettori di distanza.

Nuovo B	Vecchio B	A
A 2	A 2	A 0
B 0	B 0	B 2
C 5	C 5	C ∞
D 5	D ∞	D 3
E 4	E 4	E ∞
F ∞	F ∞	F ∞
G ∞	G ∞	G ∞

a. Primo evento: B riceve una copia del vettore di A.

Nuovo B	Vecchio B	E
A 2	A 2	A ∞
B 0	B 0	B 4
C 5	C 5	C ∞
D 5	D 5	D 5
E 4	E 4	E 0
F 6	F ∞	F 2
G ∞	G ∞	G ∞

b. Secondo evento: B riceve una copia del vettore di E..

```

DistanceVectorRouting() {
    //Inizializzazione: creazione dei vettori distanza iniziali del nodo
    D[this] = 0
    for (y = 1 to N) {
        if (y is vicino) D[y] = c[this][y]
        else D[y] = INF
    }
    Spedisci il vettore {D[1], ..., D[N]} a tutti i vicini

    //Aggiornamento: usare il vettore ricevuto dal vicino w o un qualsiasi cambiamento negli archi
    while (sempre) {
        wait (un vettore Dw da un vicino w o un qualsiasi cambiamento negli archi)
        for (y = 1 to N) {
            D[y] = min(D[y], (c[this][w] + Dw[y])) //Equazione di Bellman-Ford
        }
        if (avviene cambiamento nel vettore)
            spedisci il vettore {D[1], ..., D[N]} a tutti i vicini
    }

    //Fine dell'algorimto DistanceVectorRouting
}

```

33.4.4 Count-To-Infinity Problem



Il problema Se il nodo X inoltra a V i pacchetti destinati a Z, allora:

Split-Horizon

X non invia a V $D_X[Z]$

Se il path per Z è stato dato da V, X non dà indietro a V informazioni su Z

Mai pubblicare un percorso mandandolo attraverso la stessa interfaccia dalla quale è stato appreso

Poisoned Reverse

X invia a V $D_X[Z] = \text{INF}$

I percorsi ricevuti da un'interfaccia devono essere pubblicati attraverso quell'interfaccia con una metrica irraggiungibile

Un percorso appreso tramite un'interfaccia verrà pubblicato come irraggiungibile attraverso la stessa interfaccia

<https://tools.ietf.org/html/rfc7868#page-38>

33.4.5 Link-State vs Distance Vector

Complessità dei messaggi

Link-State: con n nodi, E collegamenti, si inviano $O(nE)$ messaggi.

Distance Vector: richiede scambi tra nodi adiacenti. Il tempo di convergenza può variare

Velocità di convergenza

Link-State: l'algoritmo $O(n^2)$ richiede $O(nE)$ messaggi. Ci possono essere oscillazioni di velocità

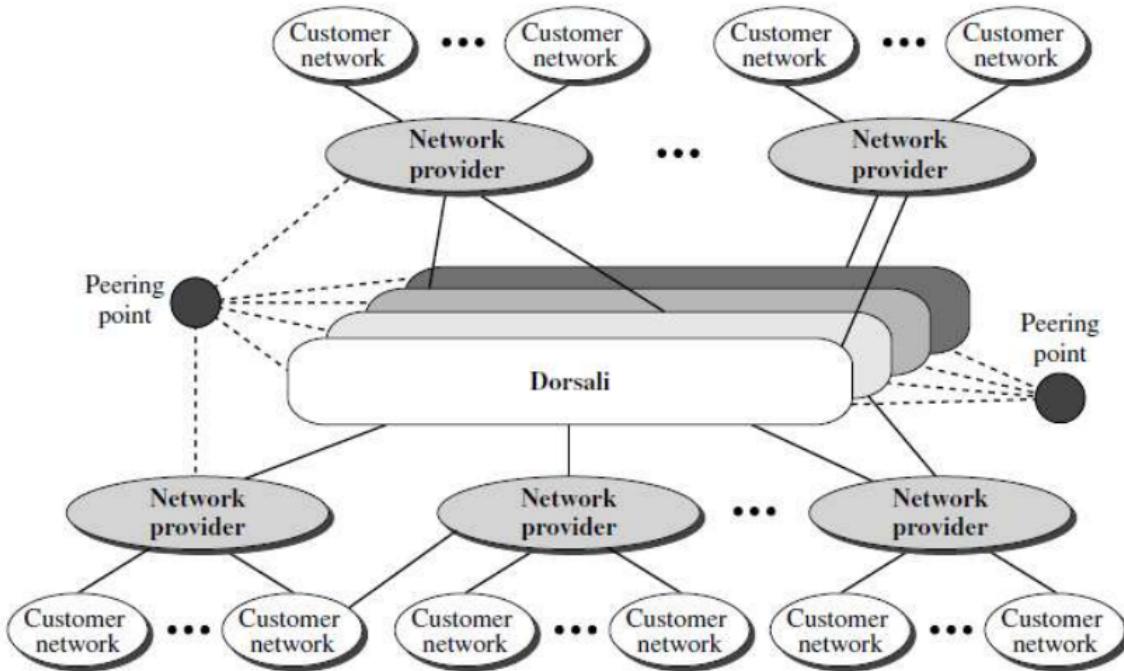
Distance Vector: può convergere lentamente. Può presentare cicli d'instradamento. Può presentare il Count-To-Infinity problem

Robustezza

Link-State: un router può comunicare via broadcast un costo sbagliato per uno dei suoi collegamenti connessi, ma non per altri. I nodi si occupano di calcolare soltanto le proprie tabelle.

Distance Vector: un nodo può comunicare cammini a costo minimo errati a tutte le destinazioni. La tabella di ciascun nodo può essere usata dagli altri. Un **calcolo errato si può diffondere per l'intera rete**.

34 Struttura di Internet



34.1 Sistemi Autonomi

Vedere internet come una rete costituita da un insieme di router omogenei interconnessi è **una visione semplicistica**: problemi di scalabilità e autonomia amministrativa.

Sistemi Autonomi Nella realtà, i router sono organizzati in **sistemi autonomi (AS)**.

Un AS è un **gruppo connesso di una o più reti IP** gestite da uno o più operatori che adottano le stesse politiche di routing verso gli altri AS.

I vari AS decidono autonomamente i protocolli e le politiche di routing che intendono adottare al loro interno. I protocollo di routing all'interno di un AS sono detti **Interior Gateway Protocol (IGP)**.

I protocolli di routing fra le AS sono detti **Exterior Gateway Protocol (EGP)**

All'interno di ciascun AS i router sono sotto uno stesso controllo amministrativo e usano lo stesso **protocollo di routing** (es RIP, OSPF...)



Ciascun **sistema autonomo** sa come inoltrare pacchetti lungo il percorso ottimo verso qualsiasi destinazione interna al gruppo.

I sistemi AS2 e AS3 hanno tre router ciascuno

I **protocolli d'instradamento** dei tre AS **non sono necessariamente gli stessi**

I router 1b, 1c, 2a e 3a sono **gateway**

34.2 Protocolli d'Instradamento

Partizioni Internet è **partizionata in sistemi autonomi**:

AS Stub: collegato solo ad un altro AS

AS Multihomed: collegato a più di un altro AS, ma **trasporta solo traffico di cui è origine o destinazione**
– come stub.

AS di transito

INTRA-AS Un protocollo di routing **INTRA-AS** determina da solo le rotte per le destinazioni interne ad una AS.

Routing Information Protocol (RIP) – Algoritmo di tipo Distance Vector

Open Shortest Path First (OSPF) – Algoritmo di tipo Link-State

INTER-AS Protocolli di routing **INTRA-AS** e **INTER-AS** determinano insieme le rotte per le destinazioni esterne all'AS.

Border Gateway Protocol (BGP), standard *de facto*.

Permette di **conoscere le destinazioni raggiungibili attraverso AS vicini**.

Propaga le informazioni di raggiungibilità ai router interni del proprio AS.

Determina percorsi buoni verso le sottoreti di destinazione.

34.3 Routing Gerarchico

Obiettivi Scalabilità, autonomia amministrativa

Esempio AS1 scopre – tramite un protocollo INTER-AS – che una sottorete X è raggiungibile via AS2, a cui AS1 è collegata tramite un gateway G.

AS1 propaga – con un protocollo INTER-AS – questa informazione al suo interno.

Il router R di AS1 riceve l'informazione "rete X raggiungibile via gateway G":

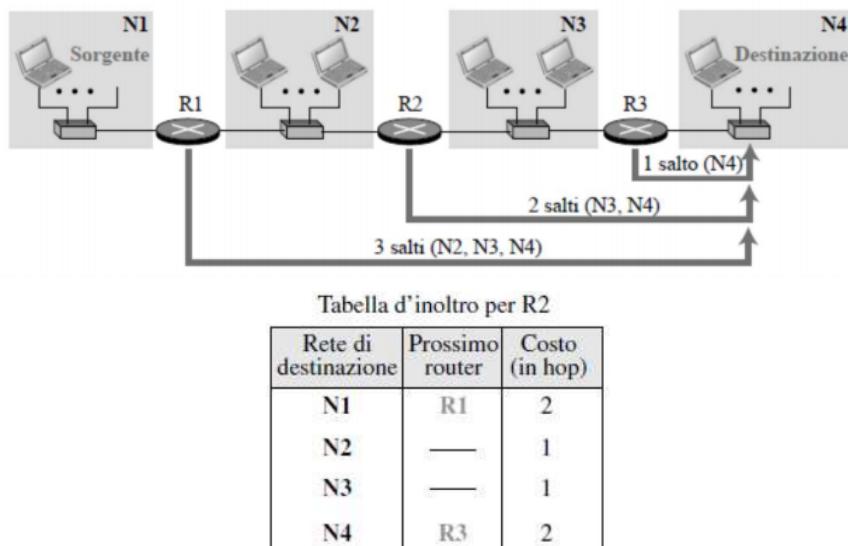
Se R non ha nessuna riga per X, allora R aggiunge (X, I) alla sua tabella d'inoltro, dove I è l'interfaccia su cui R inoltra i pacchetti destinati a G (secondo quanto indicato dal protocollo INTRA-AS).

Altrimenti sceglierà...⁴

34.4 Protocolli

34.4.1 RIP

Routing Information Protocol INTRA-AS. La metrica è il numero di sottoreti attraversate, massimo 15. Usa il Distance Vector ($\infty = 16$, quindi per reti abbastanza piccole)



Caratteristiche Processo demone `routed` su porta 520 in UDP. Gli aggiornamenti vengono inviati ogni 30 secondi, oppure quando cambia la tabella d'inoltro.

Node R riceve advertisement da vicino V con DV[y]: $DR[y] = 1 + DV[y]$

```
Se  $DR[y] \geq 1 + DV[y]$  e  $nextHopR[y] \neq V$ 
//costo ricevuto inferiore a quello del vecchio percorso
Oppure
Se  $nextHopR[y] = V$ 
//il percorso era noto ed era stato annunciato da V, ma costo cambiato
```

⁴Vedi più avanti

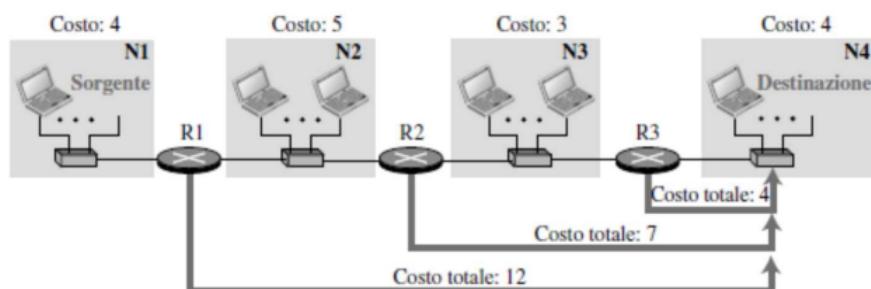
34.4.2 OSPF

Open Shortest-Path First INTRA-AS, Link-State.

Adatto a reti più grandi rispetto al RIP. Come metriche possiamo usare il **numero di sottoreti attraversate**, ma anche altri costi anche definiti dall'amministratore.

Tabella d'inoltro per R2

Rete di destinazione	Prossimo router	Costo
N1	R1	9
N2	—	5
N3	—	3
N4	R3	7



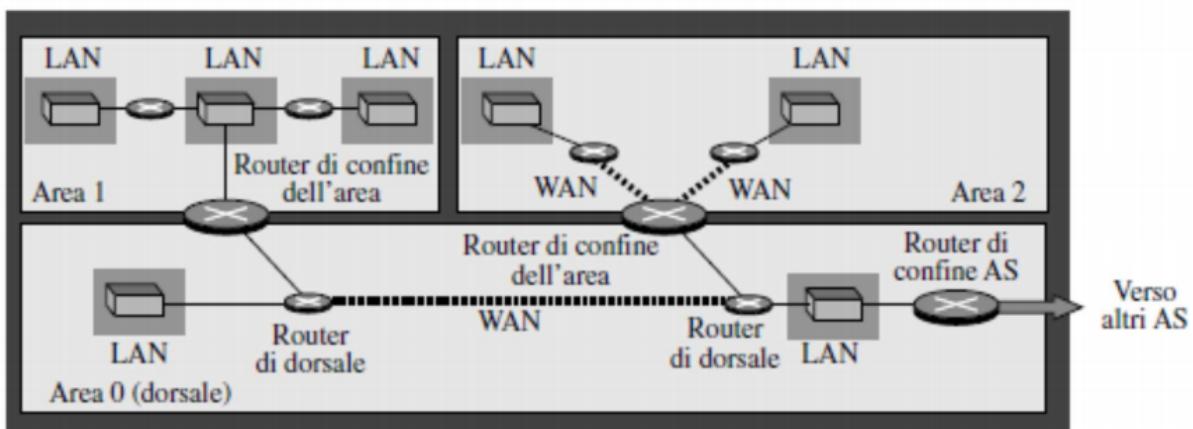
Un nodo deve **ricevere i link-state packet da tutti gli altri nodi** per potersi costruire la topologia di rete. Si ha quindi un **elevato numero di messaggi** quindi un rischio di flooding.

Idea Un AS può essere **partizionato in aree** per ridurre il flooding di pacchetti Link-State, una delle quali fa da **dorsale (backbone area)**, alla quale tutte le altre aree si collegano.

Così vengono ridotti i nodi presenti in un'area, quindi il flooding dei messaggi si può sempre verificare ma solo all'interno della singola area e più raramente.

Usa il protocollo IP.

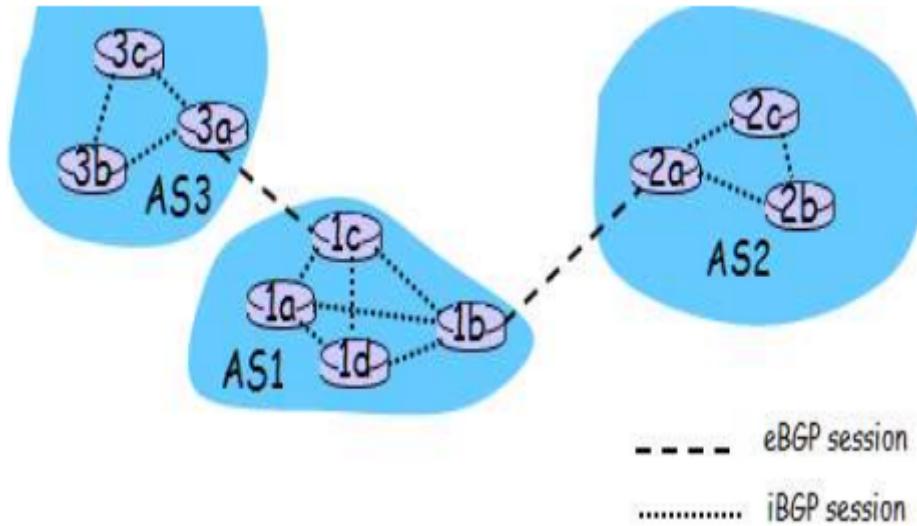
Sistema autonomo (AS)



34.4.3 BGP

Border Gateway Protocol Questo è l'unico protocollo INTER-AS usato su Internet.

Funzionamento Coppie di router si scambiano info sulle connessioni TCP semi-permanenti: sessioni BGP esterne e interne.



Caratteristiche

Aggregazione degli indirizzi, destinazioni rappresentate da prefissi CIDR (Class Inter-Domain Routing). Pubblicare un prefisso significa impegnarsi a instradare pacchetti destinati a reti in quel prefisso.

Distribuzione delle informazioni di raggiungibilità: il gateway riceve info sulla sessione eBGP e distribuisce info sulle sessioni iBGP.

Altro gateway dell'AS può ripubblicizzare info con eBGP...

Politiche d'importazione: quando un gateway riceve un advertisement, usa tali politiche per accettarlo o rifiutarlo.

Scelta delle rotte: un router può ricevere più di una rotta per lo stesso prefisso. Sequenza di regole principale:

1. Attributo di "preferenza locale", scelta dell'amministratore o impostato dai router dell'AS. Vengono selezionati gli advertisement con i valori più alti
2. AS_PATH più breve
3. NEXT_HOP più vicino ("hot potato routing")

Advertisement BGP route = prefix + attributes

I due attributi più importanti sono:

AS_PATH

Sequenza degli AS attraversati dall'advertisement. Usato per scartare gli advertisement già ricevuti e scegliere tra più percorsi per lo stesso prefisso.

NEXT_HOP

Indica l'interfaccia del gateway da usare per il prefisso pubblicato.

RFC 1163: l'attributo NEXT_HOP definisce l'indirizzo IP del router di bordo che dovrebbe essere usato come next hop verso le reti elencate nel messaggio UPDATE. Questo router di bordo deve appartenere alla stessa AS di chi emette l'advertisement.

35 Livello Data Link

Introduzione Introduciamo il **livello link** con alcuni termini utili:

Host e router → **nodi**

Canali di comunicazione, che collegano nodi adiacenti lungo un cammino → **collegamenti (link)**

Collegamenti **cablati**

Collegamenti **wireless**

LAN

Unità di dati scambiate dai protocolli a livello link → **frame**

I **protocolli** a livello link si occupano del trasporto di datagrammi lungo un singolo canale di comunicazione.

Datagrammi Il livello data link quindi muove i datagrammi da un nodo al nodo adiacente su un singolo link di comunicazione. Un **datagramma** quindi può essere gestito da più protocolli su collegamenti diversi.

Esempio Un datagramma può essere gestito da Ethernet sul primo collegamento, WAN sul collegamento intermedio e da PPP sull'ultimo.

Servizi Anche i servizi erogati dai protocolli a livello data link possono essere diversi, ad esempio non tutti i protocolli forniscono un servizio di consegna affidabile.

Analogia con un tour operator Un viaggio da Princeton a Losanna:

Taxi: Princeton → Aeroporto JFK

Aereo: JFK → Ginevra

Treno: Ginevra → Losanna

Turista = datagramma

Ciascuna **tratta = collegamento**

Tipologia del trasporto = protocollo di link

Agente di viaggio = protocollo di routing

35.1 Collegamenti

Tipologie:

Punto-a-punto: collegamento dedicato a due soli dispositivi

Broadcast: collegamento condiviso tra più dispositivi

Quando un nodo trasmette un frame, il canale lo **diffonde a tutti i nodi collegati** che ne ricevono una copia.

Suddividibile in due **sottolivelli**:

Datalink Control

Medium Access Control, accesso al mezzo (in canali broadcast a rischio collisione)

35.2 Servizi

Framing I protocolli encapsulano i datagrammi del livello di rete all'interno di un frame a livello di link. Il framing separa i vari messaggi durante la trasmissione sorgente-destinazione.

Per identificare origine e destinatario sono usati gli indirizzi MAC, diversi rispetto agli IP.

Frame Campo dati — intestazione — eventuale trailer

Consegna Affidabile Considerata non necessaria sui collegamenti con un basso numero di errori sui bit — come fibra ottica, coassiale e doppino intrecciato — viene utilizzata sui collegamenti con alto tasso di errori — come i collegamenti wireless.

Controllo del Flusso Evita che il nodo trasmittente saturi il ricevente

Rilevamento degli Errori Gli errori sono causati dall'attenuazione del segnale e dal rumore elettromagnetico.

Il nodo ricevente individua la presenza degli errori ed è possibile, grazie all'inserimento di bit di controllo dell'errore all'interno del frame da parte del mittente, rilevare l'errore quando avviene.

Correzione degli Errori Il nodo ricevente può anche determinare il punto in cui si è verificato l'errore e correggerlo.

35.3 Indirizzi

Un indirizzo a livello data link viene associato alla scheda di rete e non al nodo. Inoltre è permanente.

Si chiama **indirizzo fisico** o indirizzo LAN o indirizzo MAC (Media Access Control)

Per le LAN ethernet (IEEE 802) è lungo 6 byte (2^{48} possibili indirizzi) ed è espresso in notazione esadecimale: ad esempio 1A-63-F9-BD-06-9B.

La IEEE definisce ed assegna i primi 24 bit (OUI – Organization Unique Identifier), mentre i rimanenti 24 vengono gestiti dalle aziende ed assegnati a livello locale.

→ Quando una società vuole realizzare degli adattatori, compra un blocco di spazio di indirizzi (univocità degli indirizzi)

35.3.1 Indirizzi MAC

Indirizzo MAC Anche noto come indirizzo LAN, indirizzo fisico o indirizzo Ethernet.

Ha una struttura piatta ed è analogo al codice fiscale di una persona: la struttura piatta non varia a seconda del luogo in cui la persona si trasferisce.

Ha una lunghezza di 48 bit.

Per confronto L'indirizzo IP è di 32 bit ed è analogo all'indirizzo postale di una persona: struttura gerarchica e devono essere aggiornati quando una persona cambia residenza.

Broadcast FF-FF-FF-FF-FF-FF

In LAN broadcast il frame viene ricevuto ed elaborato da tutti gli utenti della LAN

Solitamente, invece, ogni scheda di rete controlla se il MAC corrisponde al proprio. In caso positivo estrae il datagramma e lo passa al livello superiore.

In caso contrario scarta il frame.

35.3.2 Indirizzamento

Quando un adattatore spedisce un frame, vi inserisce l'indirizzo MAC della scheda di destinazione. Nel caso di reti broadcast, tutti gli adattatori sulla rete controllano l'indirizzo di destinazione e passano i dati al livello superiore solo se lo riconoscono come proprio.

Se un adattatore vuole che tutte le schede di rete passino i dati agli strati superiori, immette nel campo l'indirizzo fisico broadcast.

35.3.3 ARP

Risoluzione degli indirizzi: problema Normalmente all'accensione una macchina conosce: il proprio MAC, il proprio IP (e la rete locale) ed il suo indirizzo alfanumerico.

Non sa chi ha attorno, le macchine direttamente visibili.

ARP Address Resolution Protocol. Le richieste ARP vengono fatte in broadcast – perché non si conoscono i MAC vicini.

L'ARP fornisce il MAC tramite l'IP.

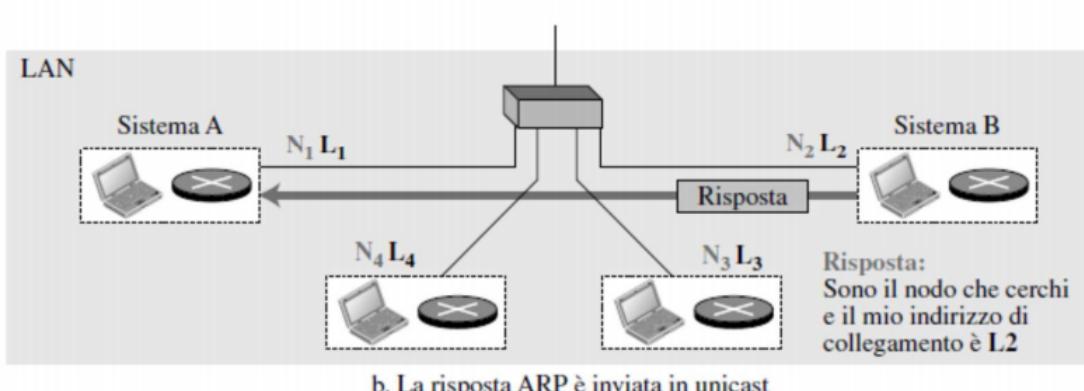
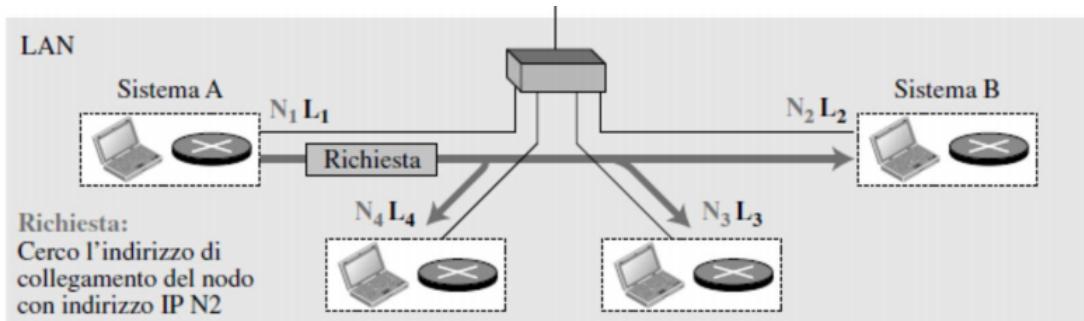
Come determinare MAC se si conosce l'IP Ogni nodo IP (host, router...) nella LAN ha una **tabella ARP** che contiene la corrispondenza IP–MAC.

<IP, MAC, TTL>

Il **TTL** (Time To Live) è un valore che indica quando bisogna eliminare una voce dalla tabella. Tipicamente è di 20 minuti.

Tabella ARP Poiché l'ARP risolve gli indirizzi IP solo per i nodi della stessa LAN, le tabelle ARP contengono la corrispondenza IP–MAC per i nodi della stessa sottorete e non necessariamente la corrispondenza di tutti i nodi.

Funzionamento



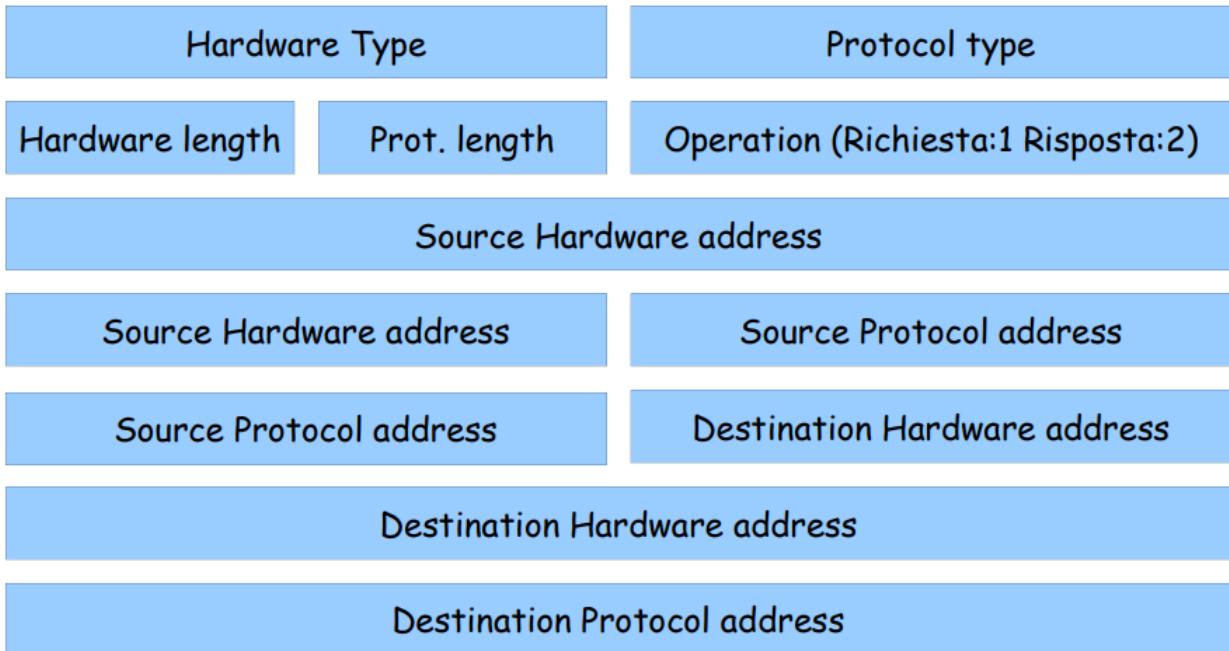
Pacchetto ARP

0

8

16

31



Hardware Type Protocollo di livello data link (es. Ethernet 1)

Protocol Type Protocollo di livello superiore (es. IP)

Source Hardware Address

Source Protocol Address Indirizzi del nodo mittente a livello link e superiore. Lunghezza variabile.

Destination Hardware Address Vuoto nelle richieste

Destination Protocol Address

Caratteristiche

Il protocollo ARP **interagisce direttamente con il livello data link**.

Viene incapsulato in un frame e spedito in broadcast sulla rete.

L'header del frame di livello link **specificava che il frame contiene un pacchetto ARP**

Ogni nodo nella rete locale riceve ed elabora il pacchetto di richiesta ARP.

Il nodo che riconosce in proprio indirizzo IP restituisce un pacchetto di risposta ARP che contiene il proprio IP e MAC. Il **pacchetto di risposta viene inviato in modalità unicast** al nodo originario della richiesta.

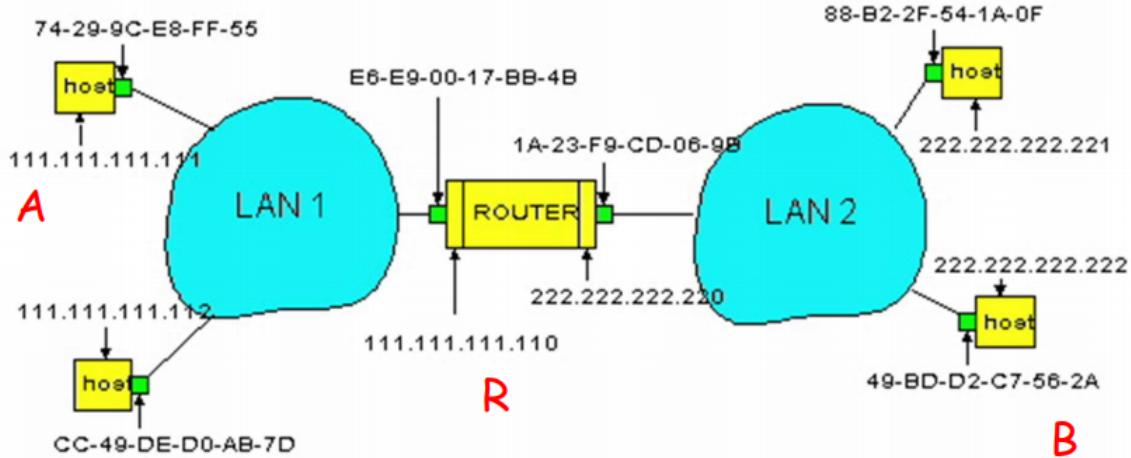
Esempio, con forwarding diretto A vuole inviare un datagramma a B, ma il MAC di B non è nella tabella ARP di A.

A trasmette in un pacchetto **broadcast** il messaggio di richiesta ARP, contenente l'IP di B: in questo pacchetto il "MAC Destinatario" è il broadcast FF-FF-FF-FF-FF-FF.

Tutte le macchine in LAN ricevono la richiesta ARP, ma **solo B risponde** riconoscendo il proprio indirizzo e comunicando ad A il proprio MAC.

Invio di un datagramma da A a B attraverso R, ipotizzando che A conosca l'indirizzo di B.

Due tabelle ARP nel router R, una per ciascuna rete IP (LAN).



A crea una datagramma con origine A e destinazione B

A usa ARP per ottenere l'indirizzo MAC di R

A richiede all'adattatore di inviare un datagramma all'indirizzo MAC di destinazione di R. Il frame contiene il datagramma IP da A a B

L'adattatore di A invia il datagramma

L'adattatore di R riceve il datagramma

R estrae il datagramma IP dal frame ethernet e vede che la sua destinazione è B

R usa ARP per ottenere l'indirizzo MAC di B

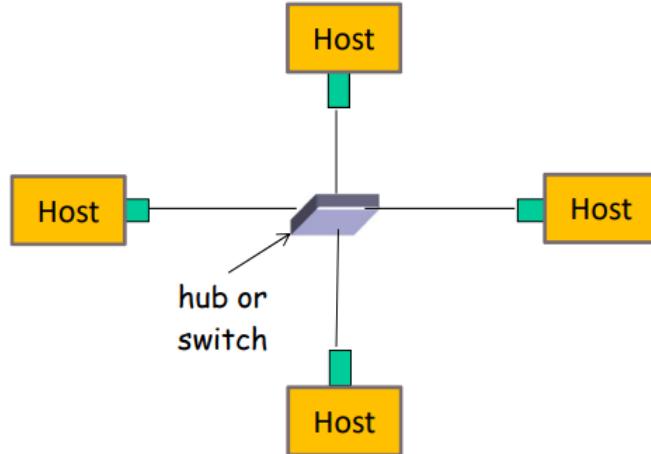
R crea un frame contenente il datagramma IP da A a B e lo invia a B

ARP è "plug-and-play". La tabella ARP di un nodo si costituisce automaticamente e non deve essere configurata dall'amministratore di sistema.

35.4 Ethernet

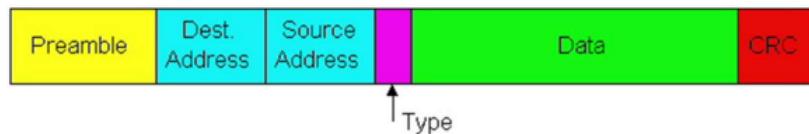
L'Ethernet detiene una **posizione dominante nel mercato** delle LAN cablate, essendo stata la prima LAN ad alta velocità con vasta diffusione. Risulta più semplice e meno costosa di token ring, FDDI e ATM ed è **sempre al passo con i tempi con la velocità di trasmissione** (10 Mbps, 100 Mbps, 1 Gbps, 10 Gbps...)

35.4.1 Topologia a Stella



Quasi tutte le reti ethernet odierne sono progettate con topologia a stella. Al centro vi è collegato un hub o **switch**.

35.4.2 Pacchetti



I datagrammi IP vengono incapsulati, dall'adattatore di rete, in pacchetti ethernet.

Preambolo: i pacchetti ethernet iniziano con un campo di 8 byte.

7 byte hanno i bit 10101010.

L'ultimo byte ha 10101011.

Servono per "attivare" gli adattatori riceventi e sincronizzare gli orologi con il mittente.

Dati: l'MTU varia da 46 byte ad un massimo di 1500 byte.

Se il datagramma è più grande allora deve essere frammentato, mentre se è più piccolo deve essere riempito (**stuffed**)

Indirizzo di destinazione, 6 byte

Quando un adattatore riceve un pacchetto con indirizzo di destinazione **corrispondente al proprio MAC** o broadcast, **trasferisce il contenuto del campo dati del pacchetto a livello di rete**.

I pacchetti con altri indirizzi MAC vengono ignorati.

Tipo, 2 byte

Consente ad ethernet di supportare vari protocolli di rete (IP, ARP, OSP...)

Qua si multiplexano i protocolli.

CRC: consente all'adattatore ricevente di rilevare la presenza di un errore nei bit del pacchetto.

36 Dispositivi di interconnessione

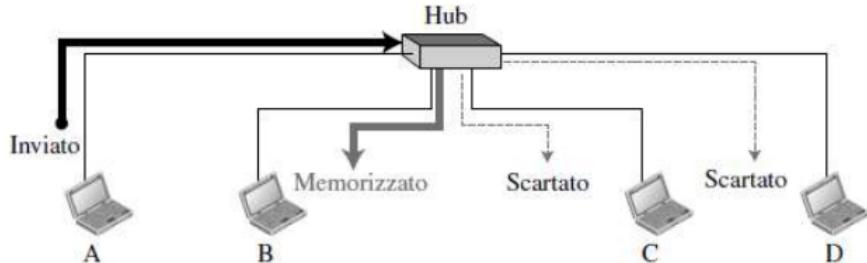
36.1 Repeater

Operano solo a livello fisico, rigenerando il segnale che ricevono. In passato erano usati per collegare segmenti di ethernet con topologia a bus.

36.2 Hub

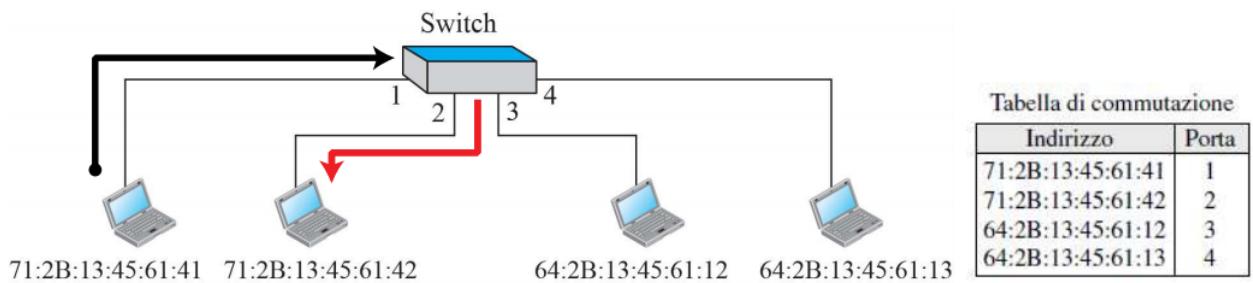
Fondamentalmente dei repeater multiporta.

Entrambi non hanno capacità di filtraggio: ciò che ricevono lo mandano a tutti i dispositivi collegati.



36.3 Switch

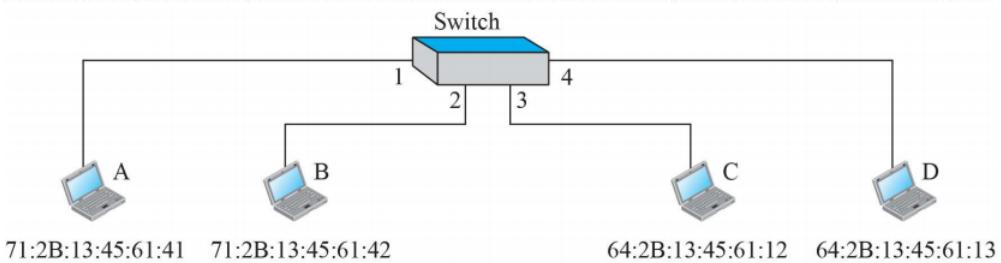
Operano sia a livello fisico, rigenerando il segnale, **sia a livello link**, verificando i MAC contenuti nei frame. Hanno una **tabella di filtraggio** e non modificano i MAC nell'intestazione dei frame



36.3.1 Switch con auto apprendimento

Costruzione graduale della tabella

Indirizzo	Porta	Indirizzo	Porta	Indirizzo	Porta	
a. Originale						
71:2B:13:45:61:41					1	
71:2B:13:45:61:41	1	b. Dopo che A invia un frame a D				
71:2B:13:45:61:41	1	71:2B:13:45:61:41	1	71:2B:13:45:61:41	1	
64:2B:13:45:61:13	4	64:2B:13:45:61:13	4	64:2B:13:45:61:13	4	
c. Dopo che D invia un frame a B						
71:2B:13:45:61:41	1	71:2B:13:45:61:41	1	71:2B:13:45:61:41	1	
64:2B:13:45:61:13	4	64:2B:13:45:61:13	4	64:2B:13:45:61:13	4	
d. Dopo che B invia un frame ad A						
71:2B:13:45:61:42	2	71:2B:13:45:61:42	2	71:2B:13:45:61:42	2	
e. Dopo che C invia un frame a D						
71:2B:13:45:61:42	2	71:2B:13:45:61:42	2	71:2B:13:45:61:42	2	
64:2B:13:45:61:12	3	64:2B:13:45:61:12	3	64:2B:13:45:61:12	3	



36.4 Router

Operano:

Livello fisico: rigenerano il segnale

Livello link: verificando gli indirizzi MAC

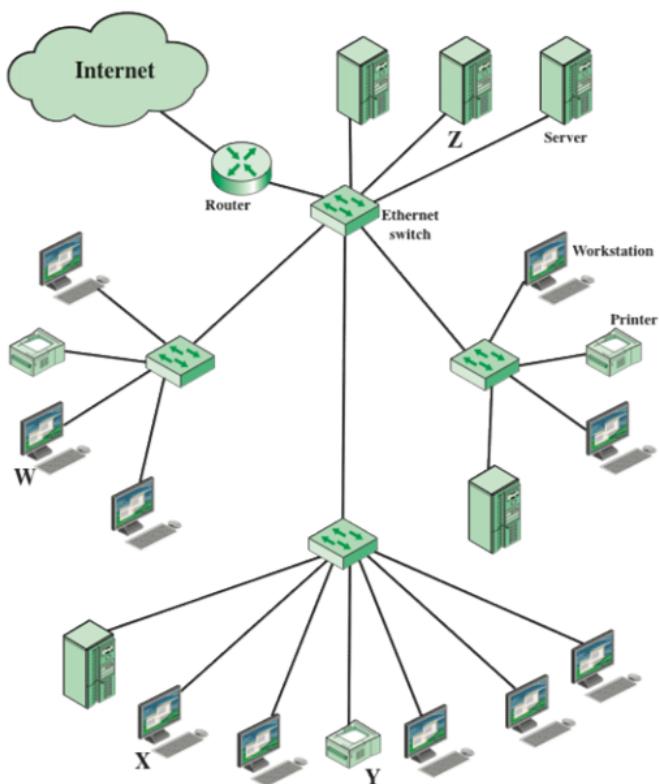
Livello di rete: verificando gli IP

Diversi Un router è fondamentalmente diverso dai ripetitori, switch e hub.

Ogni router ha **1 indirizzo MAC e 1 indirizzo IP per ogni interfaccia**.

I router **operano solo sui frame il cui indirizzo di destinazione (link) è quello dell'interfaccia (link) su cui arrivano**. Inoltre cambiano gli indirizzi link contenuti nei frame che inoltrano.

37 VLAN



Esempio di una LAN gerarchica. I dispositivi sono organizzati in 4 segmenti, ognuno servito da un LAN switch. I tre gruppi inferiori potrebbero corrispondere a diversi dipartimenti, fisicamente separati, e il gruppo superiore può corrispondere ad una server farm centralizzata usata da tutti i dipartimenti.

Subnet IP Partizioni della rete in diverse subnet. Un approccio semplice per migliorare l'efficienza è di **partizionare fisicamente la LAN in diversi broadcast domains**, ad esempio collegando i 4 LAN switch ad un router centrale ottenendo 4 LAN separate e un broadcast da X è trasmesso solo alla sottorete di X.

Esempio di una configurazione LAN Un LAN switch è un dispositivo store-and-forward che si occupa di inoltrare pacchetti ad un numero di sistemi terminali interconnessi che formano un **LAN segment**.

Lo switch può inoltrare un frame MAC da un dispositivo collegato alla sorgente ad un dispositivo collegato alla destinazione.

Vari switch possono essere interconnessi così che più segmenti LAN formino una LAN più grande.

Un LAN switch è anche connesso ad un link di trasmissione o ad un router o ad altri dispositivi di rete che forniscono connettività ad internet o ad altre WAN.

Indirizzamento Unicast Il MAC di destinazione designa un unico destinatario.

Es.: trasmissione di un singolo frame MAC da X a Y. Viene trasmesso da X allo switch locale che lo spedisce a Y.

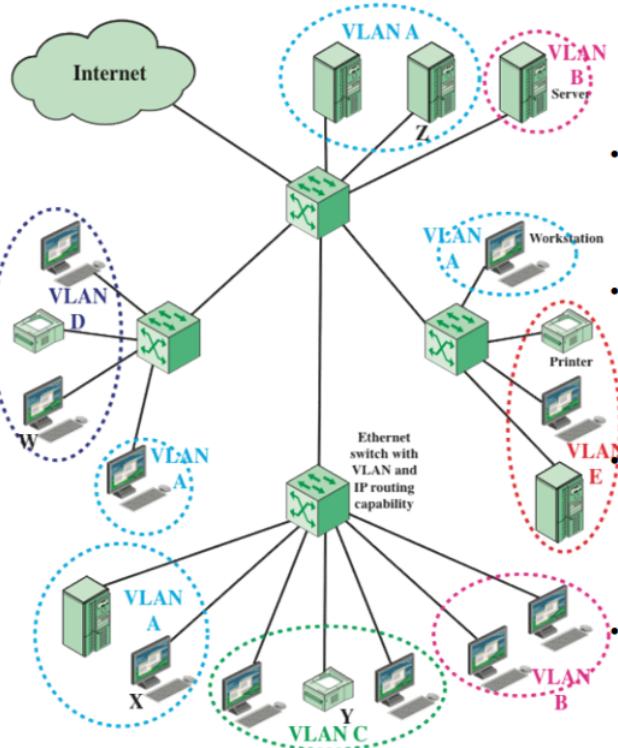
Es.: trasmissione di un singolo frame MAC da X a Z. Viene trasmesso da X allo switch locale, che lo trasmette agli switch appropriati verso la destinazione voluta.

Indirizzamento Broadcast Tutti i dispositivi sulla LAN devono ricevere una copia del frame. Quindi se X trasmette un frame con destinazione broadcast, **tutti i dispositivi collegati a tutti gli switch devono riceverne una copia**.

Dominio Broadcast Il broadcast domain è l'insieme di tutti i dispositivi che ricevono frame broadcast gli uni dagli altri.

In molte situazioni, un frame broadcast è usato per trasportare informazione significativa a livello locale, come per network management o qualche tipo di avviso.

Se un frame broadcast è utile solo per un dipartimento si spreca capacità di trasmissione sulle altre porzioni di rete.



Una configurazione VLAN

37.1 Definire le VLAN

Una VLAN è un dominio broadcast che consiste in un gruppo di dispositivi terminali che non sono vincolati alla loro posizione fisica e possono comunicare come se si trovassero sulla stessa LAN. Una VLAN non è limitata ad uno switch ma può attraversare più switch interconnessi. Bisogna quindi definire chi appartiene ad una VLAN

37.1.1 Appartenenza per gruppo di porta

Ogni switch nella configurazione LAN contiene due tipi di porta: una porta **trunk** che connette due switch ed una **porta end** che connette switch e terminale.

Una VLAN può essere definita assegnando ad ogni porta una specifica VLAN. Ha il vantaggio che è relativamente facile da configurare.

Lo svantaggio principale è che l'amministratore di rete deve riconfigurare le appartenenze alle VLAN quando un sistema terminale cambia porta.

37.1.2 Appartenenza per indirizzo MAC

Le VLAN basate sui MAC consentono agli amministratori di rete di spostare fisicamente un terminale e far sì che quel terminale mantenga automaticamente la configurazione VLAN.

Il problema principale con questo metodo è che **l'appartenenza deve essere assegnata inizialmente**.

37.1.3 Appartenenza per informazioni sul protocollo

Le appartenenze possono essere assegnate anche in base all'indirizzo IP, alle informazioni del protocollo di trasporto o di protocollli a più alto livello.

Questo è un **approccio flessibile**, ma richiede che lo switch esamini porzioni del frame MAC a livelli superiori di quello link, che potrebbe comportare un impatto sulle prestazioni.

VLAN Le Virtual Local Area Network, VLAN, consentono di disperdere fisicamente le varie porzioni di rete a piacere senza perdere l'identità di gruppo.

- Una trasmissione da X a Z avviene **all'interno della solita VLAN A**, quindi anche un broadcast MAC da X è trasmesso a tutti i dispositivi in tutte le porzioni della solita VLAN.

Però una trasmissione da X alla stampante Y attraversa due VLAN, quindi è richiesto dell'instradamento a livello IP per spostare il pacchetto IP da X a Y.

Lo switch determina se il frame MAC in arrivo è destinato alla solita VLAN o meno. In caso non lo sia, lo switch instrada il pacchetto IP contenuto a livello IP.

37.2 Comunicare l'appartenenza

Gli switch devono possedere un modo per comprendere le appartenenze alle VLAN quando il traffico di rete arriva da altri switch.

Frame Tagging Un approccio più comune è il **frame tagging**, dove viene inclusa un'intestazione ad ogni frame che **identifica univocamente a quale VLAN appartiene** un particolare frame MAC. IEEE 802 ha sviluppato uno standard per il frame tagging, l'**IEEE 802.1Q**:

Ad ogni VLAN in una configurazione LAN è **assegnato un VID globalmente univoco**, tra 1 e 4094.

Assegnando il solito VID a sistemi terminali su switch diversi, uno o più domini di broadcast VLAN possono essere estesi attraverso una rete più grande.

38 Cenni di applicazioni P2P

Fin'ora abbiamo trattato di **applicazioni client-server**: il **server offre un servizio** ed il **client chiede di utilizzarlo** (HTTP, FTP, Telnet, SMTP, DNS...)

38.1 Paradigma Peer-to-Peer

P2P Nel **paradigma peer-to-peer** tutti gli host sono peer e agiscono sia da client che da server. Vari esempi includono: Freenet, Napster, Gnutella, FastTrack (KaZaA), eMule, BitTorrent, ...

Peer In linea di principio, tutti i peer (nodi) hanno la stessa importanza. I nodi sono **indipendenti, autonomi e localizzati ai bordi** (edge) di Internet.

Non c'è controllo centralizzato: ogni peer ha funzione di client e server e **condivide delle risorse**.

Sono sistemi altamente distribuiti in cui i nodi possono essere dell'ordine delle centinaia di migliaia. Ma **tutti i nodi sono dinamici e autonomi**: un nodo può entrare e uscire dalla rete P2P in ogni momento. Le opzioni di ingresso e uscita dalla rete sono anche sofisticate con ad esempio ridondanza delle informazioni.

Servent Server + Client

Nella realtà All'atto pratico possono essere presenti **server centralizzati o nodi con funzionalità diverse rispetto agli altri** – i **supernodi**. Esempi di applicazioni:

Distribuzione e memorizzazione di contenuti

Es. file sharing: Gnutella, BitTorrent, eMule...

Es. file storage: Freenet

Condivisione delle risorse di calcolo (elaborazione distribuita)

Es. SETI@home...

Collaborazione e comunicazione

Es. chat, IRC, Jabber...

Telefonia

Es. Skype...

Content Delivery Network

Es. CoralCDN...

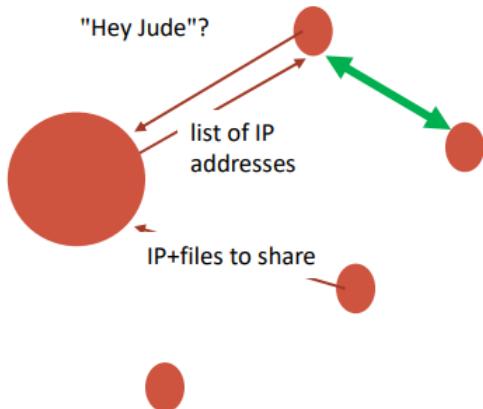
Piattaforme

Es. JXTA (sun.com/software/jxta)

Problema Coppie di peer comunicano direttamente tra loro. Ogni peer però **non è necessariamente sempre attivo** e può cambiare IP ogni volta che si connette. Quindi **come trovare i peer?** Come tenere traccia dei file disponibili nei vari peer?

38.2 Directory Centralizzata

Directory client-server, file transfer P2P (es. Napster)



Problemi Unico punto di fallimento.
Collo di bottiglia per performance
Facile da "oscurare".

38.3 Reti Decentralizzate

Non c'è un servizio di directory centralizzato, i peer si organizzano in una **overlay network** (rete logica). Ci sono due categorie: reti non strutturate e reti strutturate

38.4 Reti Non Strutturate

I nodi sono **organizzati come un grafo casuale**: l'organizzazione della rete segue principi molto semplici. Non ci sono vincoli sul posizionamento delle risorse rispetto alla topologia del grafo: la **localizzazione delle risorse è resa difficoltosa dalla mancanza di organizzazione**. L'aggiunta e rimozione dei nodi è un'operazione semplice e poco costosa. Obiettivo: gestire i nodi con comportamento fortemente transiente (tassi di join/leave elevati).

Esempi: Gnutella, FastTrack, eDonkey, Overnet...

Query Flooding Completamente distribuita (**no server centralizzato**)

Esempio: Gnutella

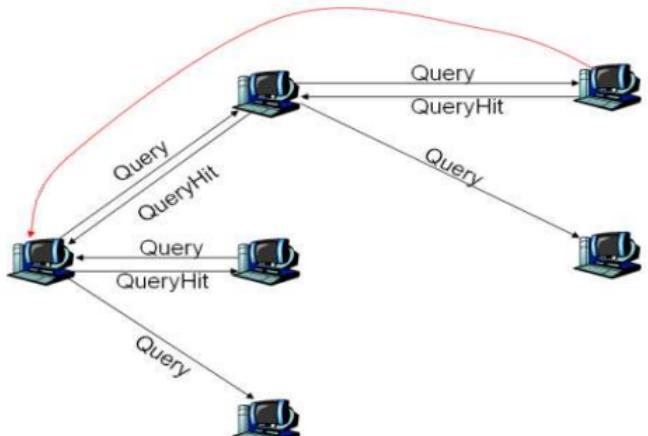
Peer formano una **overlay network**

Archi = connessioni TCP

Ogni nodo tipicamente connesso a 10 vicini

Peer invia query a tutti i suoi vicini

Se un peer riceve una query e ha il file richiesto, allora invia il messaggio **QueryHit** sul reverse path, altrimenti inoltra la query a tutti i suoi vicini.



Problemi:

Flooding

Incompletezza con raggio limitato

38.5 Copertura Gerarchica

Cerca di combinare "il meglio" dei due approcci precedenti.

No server con tutti i contenuti, solo bootstrap servers

I peer non sono tutti uguali, ma esistono group leader

Ogni peer è **group leader** se è "potente" in banda o risorse, oppure viene assegnato ad un **group leader**. Le connessioni TCP sono stabilite tra peer e **group leader** e tra alcune coppie di **group leader**.

Ogni **group leader** tiene traccia del contenuto dei propri figli.

Un **group leader** è un mini-server tipo Napster. Le connessioni leader-leader sono overlay network stile Gnutella.

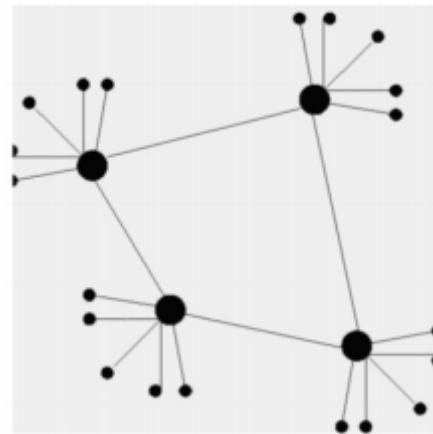
Ad ogni file viene associato un suo hash ed un suo descrittore.

Il client invia una query con una keyword al suo **group leader**.

Il **group leader** risponde con dei "match" del tipo <hash del file, IP>.

Se il leader inoltra la query ad altri leader, questi rispondono con altri match.

Il client sceglie quindi il file da scaricare.



38.6 BitTorrent

Protocollo BitTorrent è un **protocollo** molto diffuso per la **distribuzione di file**. L'idea di base è **dividere un file in pezzi da 256Byte e far distribuire ad ogni peer i dati ricevuti fornendoli a nuovi destinatari**. In questo modo si:

riduce il carico di ogni sorgente

riduce la dipendenza dal distributore originale

fornisce ridondanza

Swarm e Chunk Un insieme di peer (swarm) partecipa alla **distribuzione di un certo file scambiandosi parti** (chunk) del file. Ogni peer allo stesso tempo preleva e trasmette più chunk.

Condivisione Per condividere un file, un peer crea un file .torrent che **contiene informazioni sul file condiviso e sul tracker**.

Tracker: il nodo che coordina la distribuzione del file.

Nuovo Peer Cerca, con un motore di ricerca, "do what u want" e **ottiene il metafile con info sui chunk e l'IP del tracker**.

Contatta il tracker e **riceve gli indirizzi di alcuni peer dello swarm**.

38.6.1 Strategie principali

Tit for Tat Si inviano dati ai peer che inviano dati, scegliendo quelli che stanno trasmettendo a frequenza maggiore.

Ogni peer classifica i suoi vicini in **choked** e **interested**, aggiornando ogni 10s la propria classifica. Ogni 30s sceglie un **choked a caso e lo promuove** cosicché i nuovi arrivati possano entrare.

Rarest Chunks First Così diventano anche meno rari.

38.7 Reti Strutturate

Vincoli Vincoli sul grafo (strutturato) e sul posizionamento delle risorse nei nodi del grafo. L'organizzazione della rete segue rigidi principi, l'aggiunta e rimozione dei nodi è un'operazione costosa.

Obiettivo: migliorare la localizzazione delle risorse.

Es.: Chord, Pastry, CAN...

38.7.1 DHT

Distributed Hash Table Ad ogni peer è assegnato un ID ed ogni peer conosce un certo numero di peer.

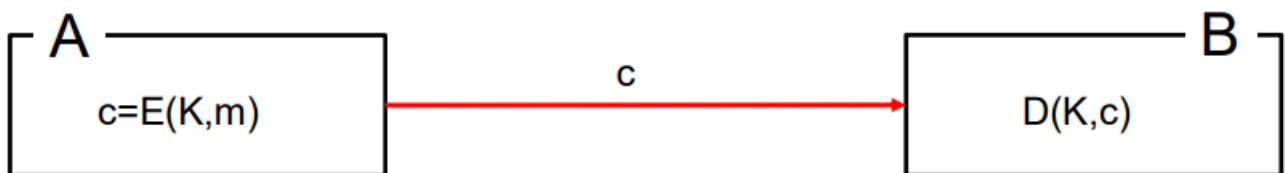
Ad ogni risorsa condivisa (pubblicata) viene assegnato un ID basato su una certa funzione hash applicata al contenuto della risorsa ed al nome.

39 Sicurezza

Obiettivi	Attacchi
Riservatezza/Segretezza Solo mittente e destinatario dovrebbero essere in grado di comprendere il contenuto del messaggio trasmesso	Intercettazione
Integrità Il contenuto della comunicazione non deve subire alterazioni durante la trasmissione	Modifica Spoofing Play-Back Repudiation
Disponibilità/Accessibilità Deve essere garantita disponibilità/operatività di informazioni e servizi	DoS DDoS

39.1 Cifratura a Chiave Simmetrica

Idea Una stessa chiave (segreta) usata per cifrare e decifrare.
La chiave può essere utilizzata per la comunicazione bidirezionale.



Cifrari monoalfabetici
Cifrario di Cesare (k-shift)...

Cifrari polialfabetici

Data Encryption Standard

Cifrario a blocchi con chiave di cifratura di 56 bit

Advanced Encryption Standard

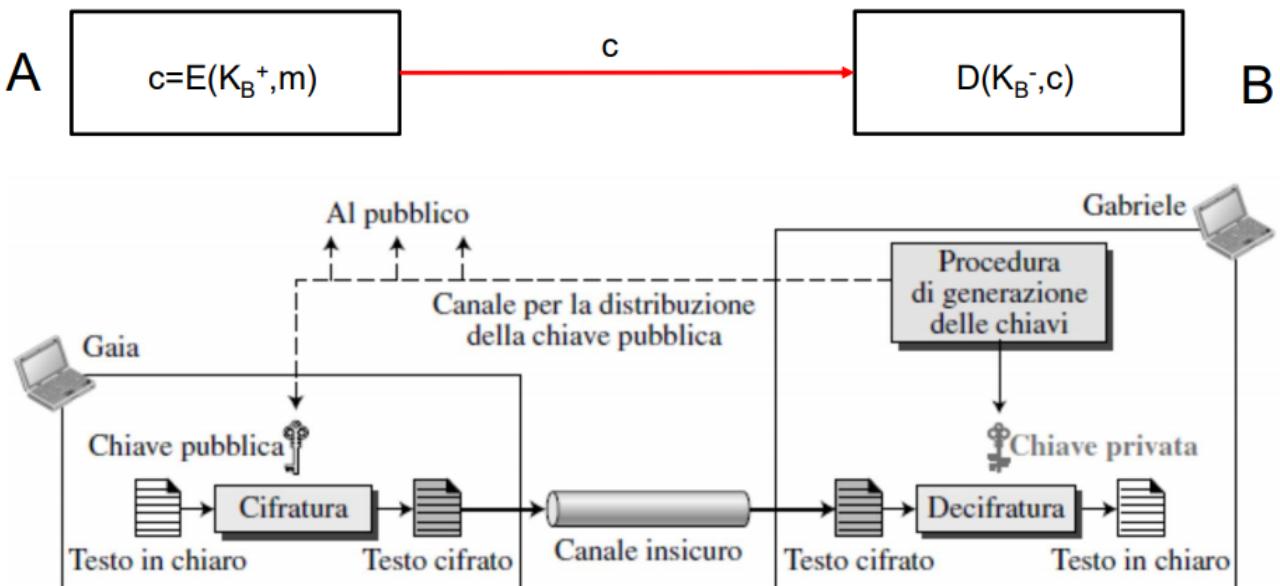
Sostituisce il DES, è del 2001.

Elabora i dati in **blocchi da 128 bit** con chiave di **128/192/256 bit**.

Problemi Come concordare la chiave da usare tra A e B? Soprattutto se non si incontrano mai?
La segretezza della chiave è inversamente proporzionale a quanto viene usata.

39.2 Cifratura a Chiave Asimmetrica

Idea Nessuna chiave segreta condivisa. Ognuno possiede
una propria chiave *pubblica* nota a tutti
una propria chiave *privata* segreta



Requisiti:

$$K_B^+ \text{ e } K_B^- \text{ tali che: } D(K_B^-, E(K_B^+, m)) = m$$

Non deve essere possibile determinare K_B^- a partire da K_B^+

$$\text{Nota bene: } D(K_B^-, E(K_B^+, m)) = D(K_B^+, E(K_B^-, m))$$

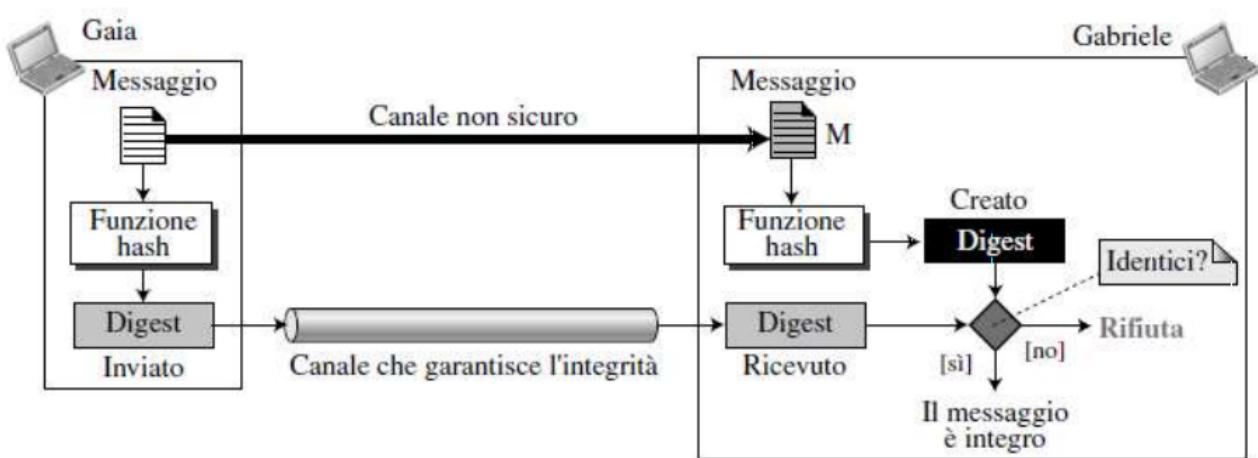
Problema: efficienza

Esempio RSA

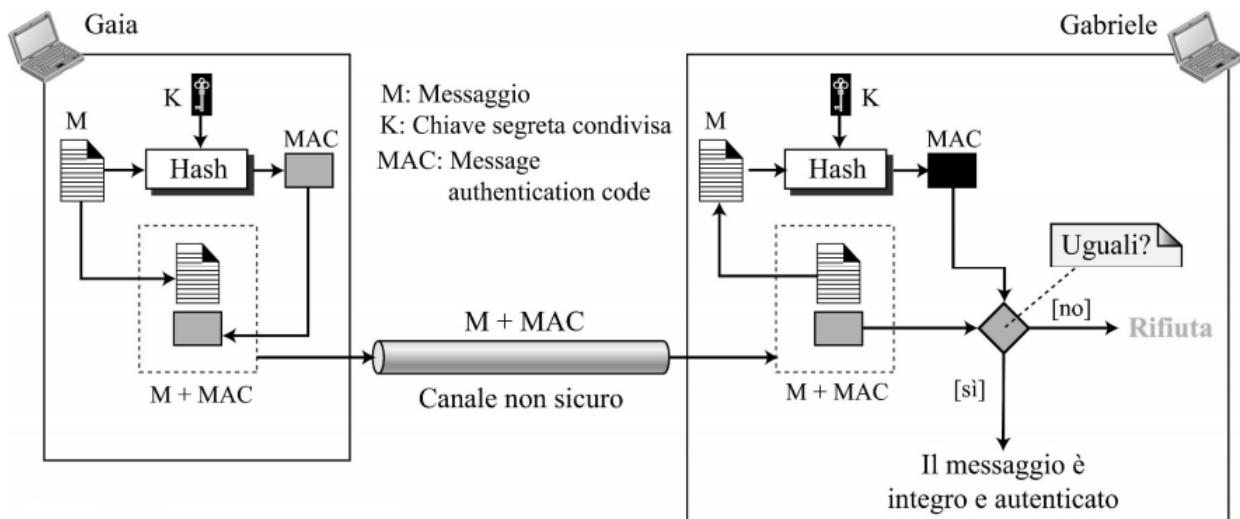
39.3 Message Digest

Vi sono casi in cui, pur non essendo richiesta la riservatezza, risulta di **fondamentale importanza l'integrità**: il messaggio originale non deve poter essere modificato. **Funzione hash**

Problema: confidenzialità



39.4 Message Authentication Code



Integrità e autenticazione con funzione hash e chiave segreta

Gaia e Gabriele **condividono la chiave K**

Mittente:

$$MAC = H(M + K)$$

Invia **M** e **MAC** su canale non sicuro

Destinatario:

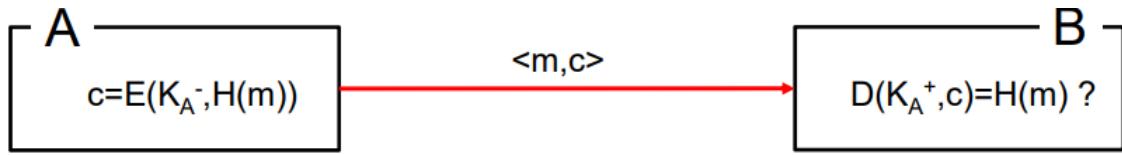
$$\text{Ricalcola } MAC = H(M + K)$$

Confronta **MAC** calcolato con quello ricevuto

Problema: ripudiabilità

39.5 Firma Digitale

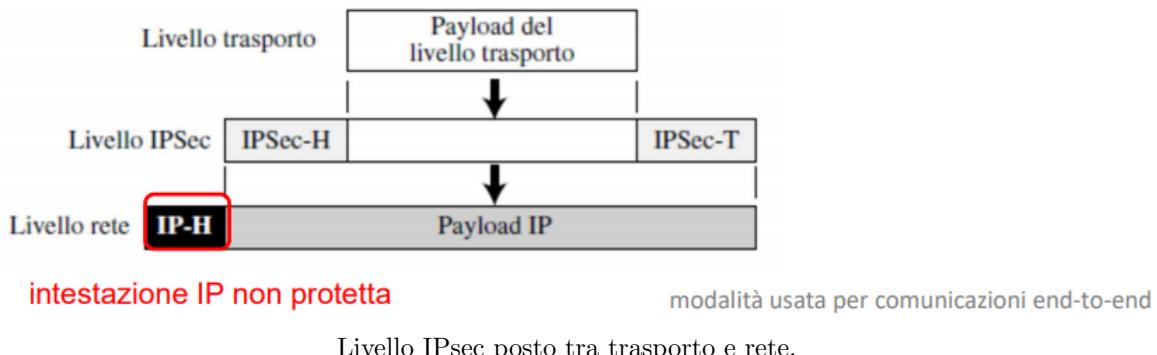
Idea Utilizzare una chiave privata per firmare il digest.



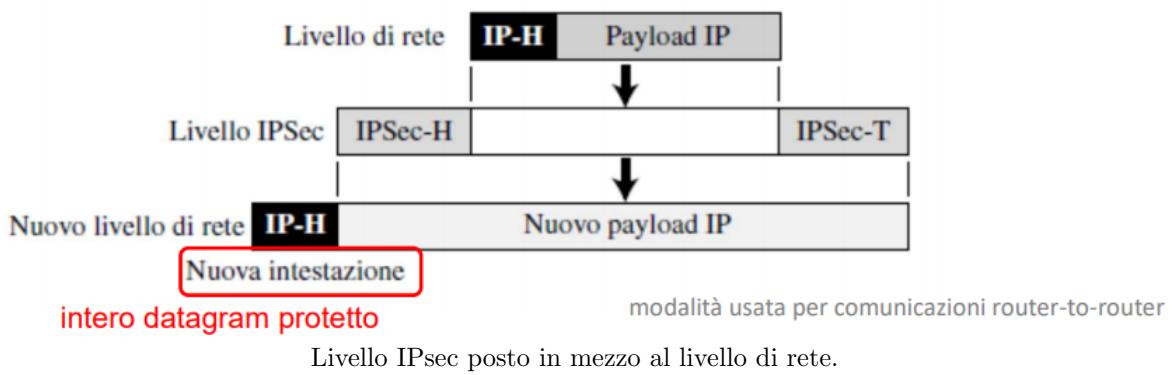
Problema: ripudiabilità (A potrebbe cambiare le chiavi).

39.6 IPsec

Modalità trasporto Protegge i dati passati dal livello trasporto al livello rete.



Modalità tunnel Protegge l'intero pacchetto IP



39.6.1 ESP

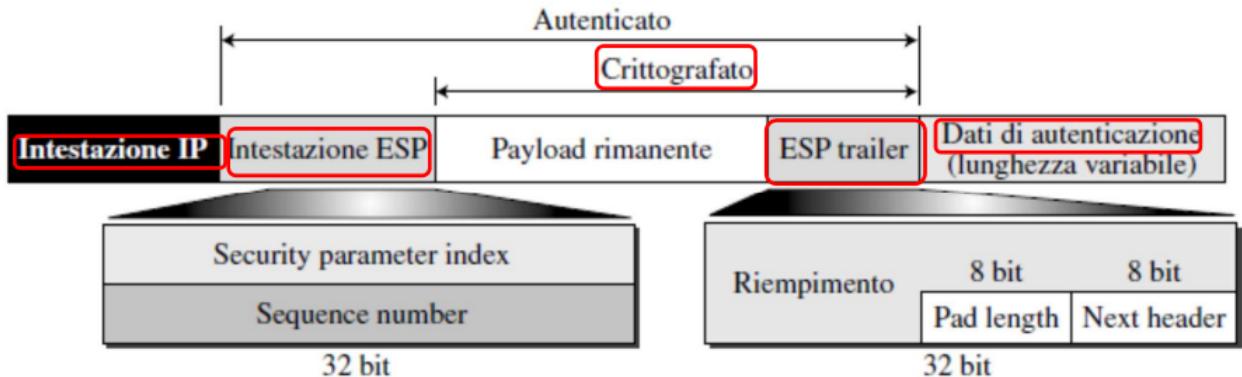
Encapsulating Security Payload Obiettivi: autenticazione della sorgente, integrità e riservatezza

Viene aggiunto un trailer ESP

Payload e ESP vengono **crittografati**

Intestazione ESP, payload e trailer ESP vengono usati per generare i dati di autenticazione, aggiunti dopo il trailer ESP

Viene aggiunta l'intestazione IP, con campo protocollo = 50



40 Esercizi

40.1 SMTP

Testo L'utente mickey@disney.com invia dal suo PC una email a donald@disney.com.
Indicare la sequenza di comandi SMTP inviati e ricevuti dal PC di mickey@disney.com se:

1. Il mailserver disney.com non è raggiungibile
2. Il mailserver disney.com è raggiungibile

Soluzione

1. Il PC del mittente non può stabilire una connessione TCP con il mailserver, di conseguenza nessun messaggio SMTP può essere inviato o ricevuto
2. Il mittente stabilisce una connessione TCP con il mailserver, e avviene lo scambio dei seguenti comandi:

```
R: 220 service ready
I: HELO . .
R: 250 OK
I: MAIL FROM: mickey@disney.com
R: 250 OK
I: RCPT TO: donald@disney.com
R: 250 OK
I: DATA
I: . .
I: . .
I: .
R: 250 OK
I: QUIT
R: 221 service closed
```

Notare che il corpo del messaggio termina sempre con <CRLF>.<CRLF>.

40.2 DNS

Testo Un host deve risolvere il nome simbolico `host.engineering.vanderbilt.edu`, il cui indirizzo IP non è noto al suo resolver (i.e. servizio DNS dell'host).

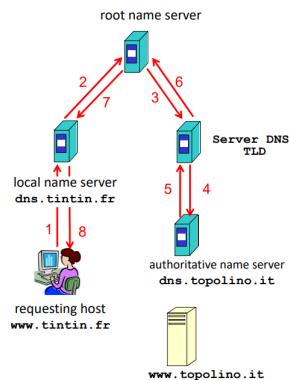
Supponendo che la gerarchia dei name server abbia 4 livelli, indicare – giustificando la risposta - il numero di messaggi DNS che, nel caso peggiore, circoleranno in Internet per risolvere tale nome simbolico, se:

1. Si utilizza ad ogni livello una risoluzione ricorsiva
2. Si utilizza ad ogni livello una risoluzione iterativa

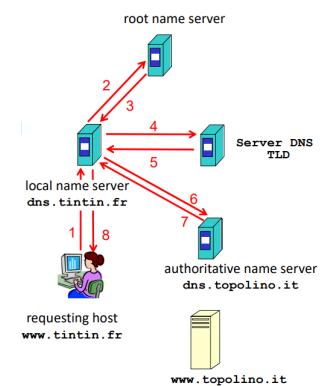
Soluzione

1. Bisogna percorrere tutto l'albero dei nameserver, dal locale fino ad un root server, e poi da questo all'autoritative name server che, nel caso peggiore, è il nameserver locale di `engineering.vanderbilt.edu`, per poi tornare indietro.

In totale sono $4 + 4$ messaggi + 2 dal resolver al nameserver locale del client e viceversa.



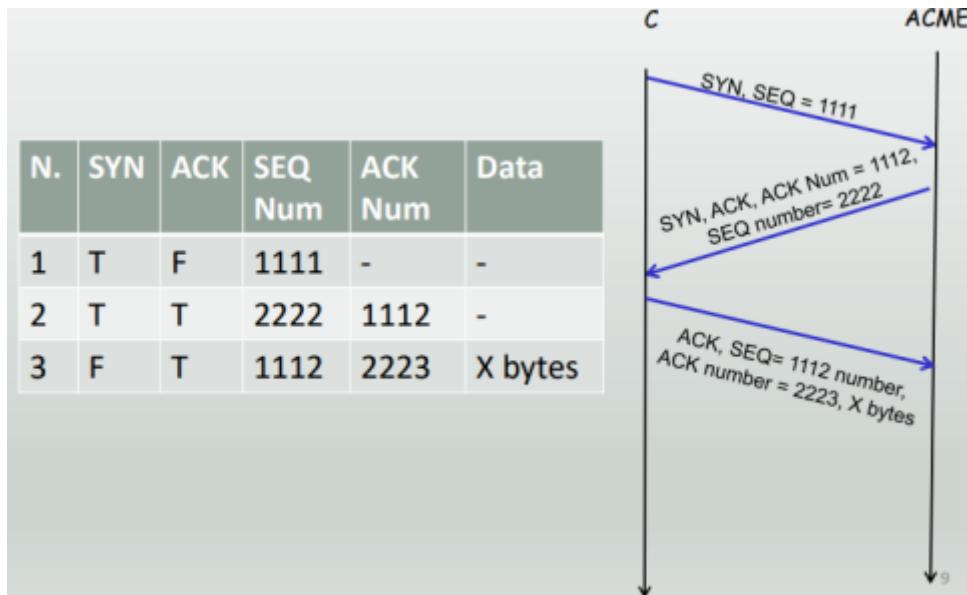
2. Analogamente i messaggi saranno sempre 10, perché i nameserver coinvolti sono 4, nel caso pessimo, più i 2 messaggi tra il nameserver locale del client e il resolver.



40.3 TCP

Testo Un client C chiede la pagina web <http://www.acme.com/home/products.html> al server B di www.acme.com con una GET che è contenuta in un segmento TCP il cui payload (campo Dati) è lungo X byte. Indicare, giustificando la risposta, i valori dei campi sequence number, ACK number, e dei flags ACK e SYN e lunghezza del campo DATA, in ciascuno dei segmenti che C e B si scambiano per aprire la connessione nell'ipotesi che l'ACK finale dell'apertura della connessione sia inviato in piggybacking assieme alla GET. Si supponga che in B ed in C rWnd sia molto grande, che non scada alcun timeout, che non ci siano errori di trasmissione, che nessun segmento vada perduto, e che il numero di sequenza iniziale di C sia 1111 e quello di B sia 2222

Soluzione



Piggybacking L'invio dei primi dati e l'invio dell'ultimo ACK dell'handshake sono uniti.

Testo Un client C ha stabilito una connessione TCP con un server web S per scaricare una pagina web che consiste di tre oggetti. Al tempo t, subito dopo avere inviato la richiesta per il terzo oggetto, l'host di C invia a S un segmento con il flag FIN a true.

Indicare, giustificando la risposta, il tempo minimo necessario al TCP di C per chiudere definitivamente la connessione supponendo che:

1. La dimensione del terzo oggetto sia 1.5 MSS
2. RTT sia costantemente 700ms e che il Maximum Segment Lifetime sia 1100ms
3. Tutti i segmenti vengano ricevuti corretti ed in ordine, e che il valore di cwnd del TCP di S sia 1MSS quando esso riceve il FIN inviato dall'host di C. Trascurare i tempi di preparazione e di trasmissione dei segmenti e assumere che S abbia già a disposizione tutti i dati da inviare.

Soluzione $3RTT + 2MSL = 4300\text{ms}$

Il TCP di C riceve al tempo $t + RTT$ il riscontro S1 del segmento FIN da lui inviato. Se S1 trasporta in piggybacking il primo MSS dei dati del terzo oggetto e il TCP di C invierà un riscontro C2 per tali dati, quindi S potrà inviare l'ultima porzione di dati e attenderne il riscontro, infine il TCP di S invierà un segmento con il flag FIN a true. Il TCP di C invierà un riscontro del FIN e considererà chiusa la connessione dopo avere atteso 2MSL, ovvero al tempo $t + 3RTT + 2MSL = t + 4300 \text{ msec}$.

Testo Si descriva il meccanismo di controllo di flusso del TCP

Soluzione possibile Si intende con controllo di flusso la capacità del mittente di evitare la possibilità di saturare il buffer del ricevitore. Infatti, a livello TCP ogni host imposta un buffer di invio e uno di ricezione. Il processo applicativo destinatario legge i dati dal buffer di ricezione (non necessariamente nell'istante in cui arrivano). Il controllo di flusso ha lo scopo di regolare la frequenza di invio del mittente in base alla frequenza di lettura dell'applicazione ricevente allo scopo di non saturare il buffer del ricevente. TCP implementa questa funzione tramite una variabile detta receive window mantenuta nel mittente: questa variabile fornisce un'idea di quanto spazio è ancora a disposizione nel buffer del ricevitore. Tale valore è comunicato nel campo window dell'header TCP dall'host destinatario.

Il valore di receive window $RcvWindow = RcvBuffer - (LastByteReceived - LastByteRead)$.

L'host destinatario comunica la dimensione di $RcvWindow$ al mittente.

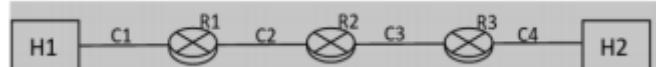
Il mittente si assicura che $LastByteSent - LastByteAcked < RcvWindow$ pari ovvero alla quantità di dati trasmessi e non ancora riscontrati.

Nelle situazioni in cui il buffer risulta pieno ($RcvWindow=0$), il mittente continua a mandare segmenti sonda di 1 byte per ricevere l'aggiornamento sulla dimensione di $RcvWindow$.

Testo Due host H1 e H2 comunicano tramite un canale che attraversa tre router R1, R2 e R3 e 4 link di capacità C1, C2, C3 e C4, rispettivamente come mostrato in figura. La comunicazione avviene tramite commutazione di pacchetto con trasmissione di tipo store and forward.

Assumendo che il ritardo di propagazione sia trascurabile, che i ritardi di accodamento nei router R1, R2 e R3 sia rispettivamente a1, a2 e a3, e che il ritardo di elaborazione nei tre router sia uguale a 1 ms, dire quanto tempo è necessario per la trasmissione da H1 a H2 di un pacchetto di dimensione L nel seguente caso:

- $L = 10 \text{ KBytes}$
- $C1 = C2 = C3 = C4 = 2 \text{ Mbps}$
- $a1 = a3 = a4 = 0.01 \text{ s}$



Soluzione Il ritardo introdotto da ogni router è dato dalla somma tra: ritardo di trasmissione, ritardo di propagazione, ritardo di accodamento e tempo di elaborazione. In particolare:

- il tempo di trasmissione di H1 è pari a $L/C1 = 40 \text{ ms}$.
- il tempo di trasmissione di R1 è pari a $L/C2 = 40 \text{ ms}$.
- il tempo di trasmissione di R2 è pari a $L/C3 = 40 \text{ ms}$
- il tempo di trasmissione di R3 è pari a $L/C4 = 40 \text{ ms}$

Quindi il ritardo introdotto dai router è: $40 + 10 + 1 = 51 \text{ ms}$

Il ritardo dovuto alla trasmissione da H1 è pari a $\frac{L}{R} = 40 \text{ ms}$.

Quindi il tempo complessivo per la trasmissione del pacchetto da H1 a H2 è

$$51 * 3 + 40 = 193 \text{ ms}$$

Testo Tizio manda dal suo account mail tizio@libero.it un messaggio di posta elettronica con testo di 256 caratteri a Caio, di indirizzo caio@occupato.it.

1. Specificare il contenuto dei primi due messaggi TCP inviati dal mailserver di libero.it, ms.libero.it, per ricevere tale messaggio.
2. Specificare, per ciascun segmento, payload, numero di sequenza, ACK number, flags posti ad 1, porta origine e porta destinazione.

Soluzione

1. Il primo segmento avrà il payload vuoto e il payload del secondo conterrà 220 `service ready`
2. Il primo segmento sarà **SYNACK**, e quindi: come già detto, payload vuoto, flags **SYN** e **ACK** ad 1, numero sequenza Z, ACK number Y, porta mittente 25, porta destinazione effimera. Il secondo segmento conterrà come payload 220 `service ready`, nessun flag ad 1, numero sequenza Z+1, ACK number Y, porta mittente 25, porta destinazione effimera.

Testo Dire in quali delle seguenti circostanze il TCP cambia la dimensione della finestra di congestione cWnd, e, nel caso, come viene ricalcolata.

Stato	Evento	Cambia la finestra?	Nuova dimensione della finestra
Slow Start	Ricezione di un ACK duplicato		
Slow Start	Scatta un timeout		
Congestion Avoidance	Scatta un timeout		

Soluzione

Stato	Evento	Cambia la finestra?	Nuova dimensione della finestra
Slow Start	Ricezione di un ACK duplicato	NO	
Slow Start	Scatta un timeout	SI	CongWin = 1 MSS
Congestion Avoidance	Scatta un timeout	SI	CongWin = 1 MSS

Testo Si consideri il seguente scenario TCP in cui, per semplicità, non sono indicati i segmenti inviati o re-inviati dal sender TCP al receiver TCP, che si suppone essere quelli necessari per avere i riscontri descritti di seguito.

- Al tempo t_0 il TCP di un host A ha una connessione già stabilita, per la quale ha 4 segmenti full sized in volo (inviai ma non riscontrati) e nessun nuovo dato da spedire, il primo byte dei segmenti in volo è il byte Y, $ssthresh = 6.5 \text{ MSS}$, $cwnd = 5 \text{ MSS}$. Inoltre, non ha ricevuto nessun riscontro duplicato.
- Tra il tempo t_0 e il tempo t_1 riceve 7 riscontri: i primi due non duplicati e con ACK number uguale a $Y + 1 \text{ MSS}$ per il primo, e $Y + 3 \text{ MSS}$ per il secondo. I seguenti 4 riscontri sono tutti duplicati, ed infine, al tempo t_1 , riceve un settimo riscontro con ACK number uguale a $Y + 4 \text{ MSS}$.

Si supponga che non scatti alcun timeout tra t_0 e t_1 . Indicare, per ciascun riscontro ricevuto, lo stato del TCP e i valori di $ssthresh$ e $cWnd$, giustificando la risposta.

Soluzione

n.1	Riscontro	Cong Window	Threshold	Stato
1	$Y+1 \text{ MSS}$	$5+1 \text{ MSS}$	$6,5 \text{ MSS}$	SS
2	$Y+3 \text{ MSS}$	$(5+1)+1 \text{ MSS}^*$	$6,5 \text{ MSS}$	CA
3-4	$Y+3$	7 MSS	$6,5 \text{ MSS}$	CA
5	$Y+3$	$3,5+3 \text{ MSS}$	$3,5 \text{ MSS}$	FR
6	$Y+3$	$7,5 \text{ MSS}$	$3,5 \text{ MSS}$	FR
7	$Y+4$	$3,5 \text{ MSS}$	$3,5 \text{ MSS}$	CA

Testo Descrivere in modo dettagliato e mediante uno pseudocodice le azioni svolte da un destinatario TCP per realizzare il controllo di flusso. Si assuma che ogni segmento ricevuto dal destinatario contenga anche lo pseudoheader (lunghezza segmento TCP). Inoltre, si supponga che la chiusura della connessione venga fatta dal mittente e che i segmenti di chiusura non contengano dati in piggybacking. Non occorre realizzare la parte iniziale delle azioni svolte (quindi le inizializzazioni delle variabili e l'apertura della connessione). Infine, si supponga che si invii un riscontro appropriato per ogni segmento ricevuto, e che il destinatario non debba inviare dati al mittente. Inoltre, per semplicità, si specifichino solamente i campi dell'header del riscontro relativi al controllo del flusso e al riscontro. Non occorre realizzare le interazioni con il livello applicativo (processo che legge dal buffer), ma solamente quelle con il mittente TCP. Si hanno a disposizione le seguenti procedure:

`receive(segm)`, per ricevere il segmento `segm` dal livello di rete

`OK(segm)`, restituisce true solo se `segm` è corretto

`nuovo(segm.x)`, vettore di booleani che vale true se dal campo `x` di `segm` si deduce che i dati contenuti in `segm` non sono doppioni (specificare `x` nella soluzione)

`insert(segm.y, finestra)`, per inserire `segm.y` nel buffer di ricezione nella posizione corretta

`calcolacknum(segm)`, restituisce il numero di riscontro per il riscontro associato a `segm`

`send(risp)`, per inviare `risp` al mittente

Descrivere il contenuto o la funzionalità delle variabili e delle altre funzioni/procedure eventualmente utilizzate.

Soluzione

```
finito = false;
while(!finito) {
    receive(segm);
    finito = segm.FIN;
    if (!finito) { //controllare l'inizio della chiusura della connessione
        if (OK(segm)) { //se il segmento non è corrotto
            if (nuovo(segm.seqnum)) { //se il segmento contiene nuovi dati
                insert(segm.dati, finestra); //inserisco i dati nel buffer
                rWnd = rWnd - (segm.lTotTCP - segm.HLEN); //nuovo valore di rWnd
                risposta.ACK = true; //flag settato a true
                risposta.ACKnum = calcolachecksum(segm);
                risposta.rWnd = rWnd;
            }
            send(risposta); //Invia o il nuovo ACK o l'ultimo ACK inviato
        }
    }
}
```

Testo Discutere l'affermazione "poichè FTP e HTTP sono protocolli adatti a trasferire file, possono essere usati indifferentemente".

Soluzione possibile Un esempio di risposta potrebbe seguire questo schema:

- Descrivere brevemente l'obiettivo di HTTP e FTP
- Entrambi usano TCP, elencare le differenze ai fini del trasferimento file

FTP

Connessione controllo, persistente, inizializzata dal client, porta 21. Comandi in formato ASCII a 7 bit

Connessione dati, inizializzata da server, porta 20, non persistente

Stateful

Offre funzionalità aggiuntive rispetto a HTTP per la gestione di file e directory (list, retr, put).

HTTP

Unica connessione per dati + comandi

Interazione stateless

40.4 IP

Testo Il sistema autonomo di una grande azienda è logicamente diviso in due sottosistemi: un sottosistema che usa indirizzi IP pubblici nel blocco 113.141.0.0/21, e uno che usa indirizzi IP privati nel blocco 192.168.0.0/16. Quali sono le maschere di rete (in notazione decimale puntata) usate nel sistema autonomo?

Soluzione La prima subnet mask è 255.255.248.0 e la seconda è 255.255.0.0

Testo Una rete aziendale che contiene 8000 hosts e quattro server (web, email, ftp e DNS), adotta al suo interno indirizzi privati del gruppo 192.168.0.0/16, ed usa per i suoi router NAT 6 degli indirizzi pubblici del gruppo 114.113.87.24/29.

Qual è il numero massimo di connessioni TCP uscenti da AS e relative agli host che possono essere contemporaneamente attive nell'AS? Giustificare la risposta.

Soluzione Ho 65536 indirizzi IP privati possibili (netmask di 16 bit, host id di 16 bit e $2^{16} = 65536$), ma ho 8 indirizzi IP pubblici disponibili ($32-29 = 3$ bit, $2^3 = 8$) quindi 8 host possono comunicare verso l'esterno in contemporanea.

Testo Un router IP connette tre sottoreti: 100.100.100.0/24, 111.99.99.0/24 e 200.200.200.0/24,, con MTU (Maximum Transfer Unit) pari a 900, 1000, e 1400, rispettivamente.

I datagrammi gestiti dal router rispettano il formato IPv4, in particolare contengono i campi IP-S (indirizzo sorgente), IP-D (indirizzo destinazione), L (lunghezza del datagramma, inclusa l'intestazione), Id (identificatore del datagramma), F (flag di frammentazione) e O (offset).

Si considerino i seguenti datagrammi ricevuti dal router:

	IP-S	IP-D	L	Id	F	O
D1	100.100.100.130	200.200.200.127	890	66	0	0
D2	111.99.99.12	100.100.100.150	1000	56	0	0
D3	200.200.200.28	100.100.100.65	1390	123	0	0
D4	200.200.200.28	111.99.99.12	1398	98	0	0

Per ognuno di questi datagrammi dire da quale sottorete è stato ricevuto, e quali datagrammi il router invia in uscita specificandone anche la sottorete di destinazione.

Soluzione

Datagramma D1

- Ricevuto dalla sottorete: 100.100.100.0/24
- Inviato alla sottorete: 200.200.200.0/24
- Frammentato in uscita: NO
- Pacchetti in uscita: uguale a D1

Datagramma D2

- Ricevuto dalla sottorete: 111.99.99.0/24
- Inviato alla sottorete: 100.100.100.0/24
- Frammentato in uscita: SI ($L = 1000 > MTU$ della destinazione che è 900)
- Pacchetti in uscita:
Sorgente Destinazione L ID F O
 111.99.99.12 100.100.100.150 900 56 1 0
 111.99.99.12 100.100.100.150 120 56 0 110

Datagramma D3

- Ricevuto dalla sottorete: 200.200.200.0/24
- Inviato alla sottorete: 100.100.100.0/24
- Frammentato in uscita: SI ($L = 1390 > MTU$ della destinazione che è 900)
- Pacchetti in uscita:
Sorgente Destinazione L ID F O
 200.200.200.28 100.100.100.65 900 123 1 0
 200.200.200.28 100.100.100.65 120 123 0 110

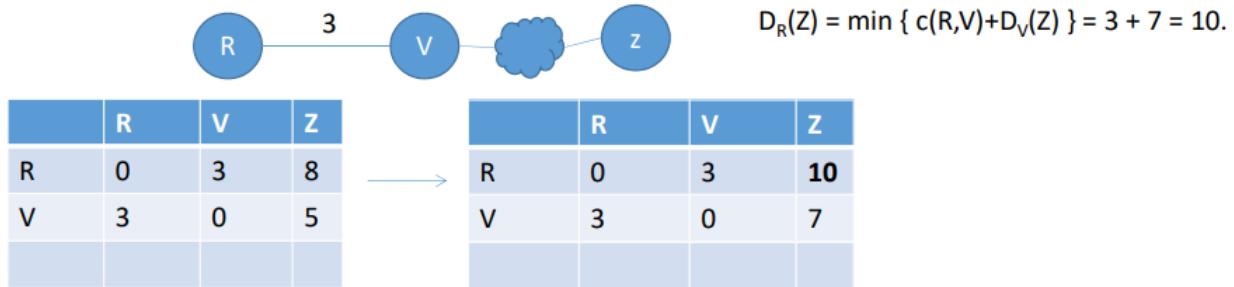
Datagramma D3

- Ricevuto dalla sottorete: 200.200.200.0/24
- Inviato alla sottorete: 111.99.99.0/24
- Frammentato in uscita: SI ($L = 1398 > MTU$ della destinazione che è 900)
- Pacchetti in uscita:
Sorgente Destinazione L ID F O
 200.200.200.28 100.99.99.12 996 98 1 0
 200.200.200.28 100.99.99.12 422 98 0 122

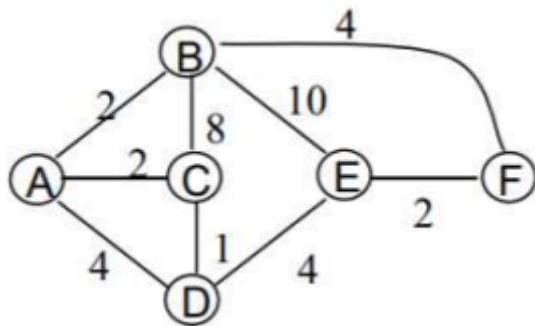
40.5 Routing

Testo In una rete i cui router utilizzano il protocollo distance vector con poisoned reverse, il router R è collegato solo a un router V, il costo del collegamento R-V è 3 e R ha calcolato che la sua distanza minima per Z è 8. Indicare – giustificando la risposta – in che modo R aggiorna il suo vettore delle distanze se V gli comunica che la sua nuova distanza da Z è 7.

Soluzione



Testo Sia data la rete con topologia raffigurata e si ipotizzi che sia in uso l'algoritmo di distance vector. Tutti i distance vector sono stati calcolati in tutti i nodi e ad un certo istante il link da E a B va giù (guasto). Approssimativamente quanti messaggi distance vector il node E manderà a seguito del guasto del link?



Soluzione E, B non viene usato in nessun cammino a costo minimo da E verso altre destinazioni, quindi non manderà nessun messaggio.