

# Reti di Calcolatori e Laboratorio

Federico Matteoni

## Indice

<b>1</b>	<b>Introduzione</b>	<b>2</b>	<b>12 HTTP</b>	<b>16</b>
			12.1 HTTP URL . . . . .	16
			12.2 Caratteristiche . . . . .	16
			12.2.1 Modello . . . . .	16
			12.2.2 Connessioni . . . . .	16
			12.3 Esempio HTTP . . . . .	17
			12.4 Messaggi HTTP . . . . .	17
			12.5 Header . . . . .	18
			12.5.1 Request header . . . . .	19
			12.6 HTTP Response . . . . .	21
			12.6.1 Response Headers . . . . .	21
			12.7 Negoziazione del contenuto . . . . .	21
			12.7.1 Entity Headers . . . . .	22
<b>2</b>	<b>Rete</b>	<b>2</b>	<b>13 Web Caching</b>	<b>22</b>
2.1	Tipi di Rete . . . . .	2	<b>14 Cookies</b>	<b>22</b>
2.2	Internetwork . . . . .	3	<b>15 Telnet</b>	<b>23</b>
2.3	Switching . . . . .	3	15.1 Introduzione . . . . .	23
<b>3</b>	<b>Internet</b>	<b>4</b>	15.2 Protocollo Telnet . . . . .	23
3.1	Enti Ufficiali . . . . .	4	15.3 NVT . . . . .	24
3.2	Reti di accesso . . . . .	5	15.4 Architettura . . . . .	25
<b>4</b>	<b>Metriche di Riferimento</b>	<b>5</b>	15.5 Funzionamento . . . . .	26
<b>5</b>	<b>Modelli Stratificati</b>	<b>6</b>	<b>16 SSH</b>	<b>27</b>
5.1	Perché stratificare . . . . .	6	<b>17 TCP Port Forwarding</b>	<b>27</b>
5.2	Smistamento Intermedio . . . . .	7	<b>18 FTP</b>	<b>28</b>
5.3	Elementi fondamentali . . . . .	7	18.1 Modello FTP . . . . .	28
5.4	Modalità di Servizio . . . . .	7	18.1.1 Connessione di Controllo . . . . .	28
5.5	Vantaggi . . . . .	7	18.1.2 Connessione Dati . . . . .	28
<b>6</b>	<b>Protocolli</b>	<b>8</b>	18.2 Altri Dettagli . . . . .	29
6.1	Incapsulamento . . . . .	8	18.2.1 Due Modalità . . . . .	29
<b>7</b>	<b>OSI RM (Open Systems Interconnection Reference Model)</b>	<b>8</b>	18.2.2 Stateful . . . . .	30
7.0.1	Pila di protocolli . . . . .	8	18.2.3 Modalità di trasmissione . . . . .	30
<b>8</b>	<b>Flusso dell'Informazione</b>	<b>9</b>	18.3 Anonymous FTP . . . . .	30
<b>9</b>	<b>Stack protocollare TCP/IP</b>	<b>10</b>	<b>19 DNS</b>	<b>31</b>
9.1	I Livelli . . . . .	10	19.1 Motivazioni . . . . .	31
<b>10</b>	<b>Livello Applicativo</b>	<b>11</b>	19.2 Struttura . . . . .	31
10.1	Protocollo a Livello Applicativo . . . . .	11	19.3 Servizi . . . . .	32
10.2	Paradigmi . . . . .	11	19.4 Spazio dei nomi . . . . .	32
10.3	Componenti di un'Applicazione di Rete . . . . .	11	19.4.1 Indirizzi . . . . .	32
10.4	Terminologia . . . . .	12	19.4.2 Nomi . . . . .	32
10.5	Identificazione di un Processo . . . . .	12	19.4.3 Top-Level Domains . . . . .	34
10.6	Esempio di API: TCP . . . . .	12	19.4.4 Struttura di un nome alfanumerico . . . . .	34
10.7	Uso dei Servizi di Trasporto . . . . .	13	19.5 Conversione . . . . .	35
<b>11</b>	<b>Applicazioni Web e HTTP</b>	<b>14</b>	19.5.1 Name Servers . . . . .	35
11.1	Terminologia . . . . .	14		
11.2	Uniform Resource Identifier . . . . .	14		
11.2.1	Sintassi . . . . .	15		
11.2.2	Absolute e Relative . . . . .	15		

# 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](http://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](http://tools.ietf.org/rfc)  
[www.ietf.org/rfc.html](http://www.ietf.org/rfc.html)
- "Computer Networks: A Top-Down Approach" B. A. Forouzan, F. Mosharraf, McGraw Hill

Ricevimento: stanza 355 DO, II piano

## 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, ...

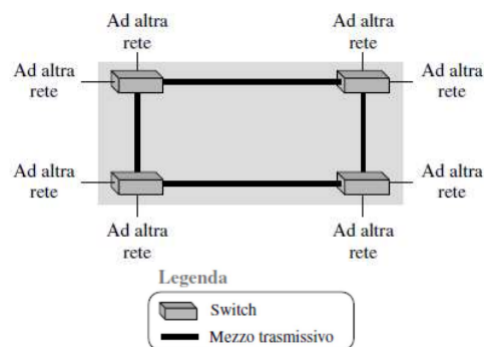
Con il termine **host** si può anche indicare un **server**.

### 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 tecnologia: switch, BUS, HUB ecc..

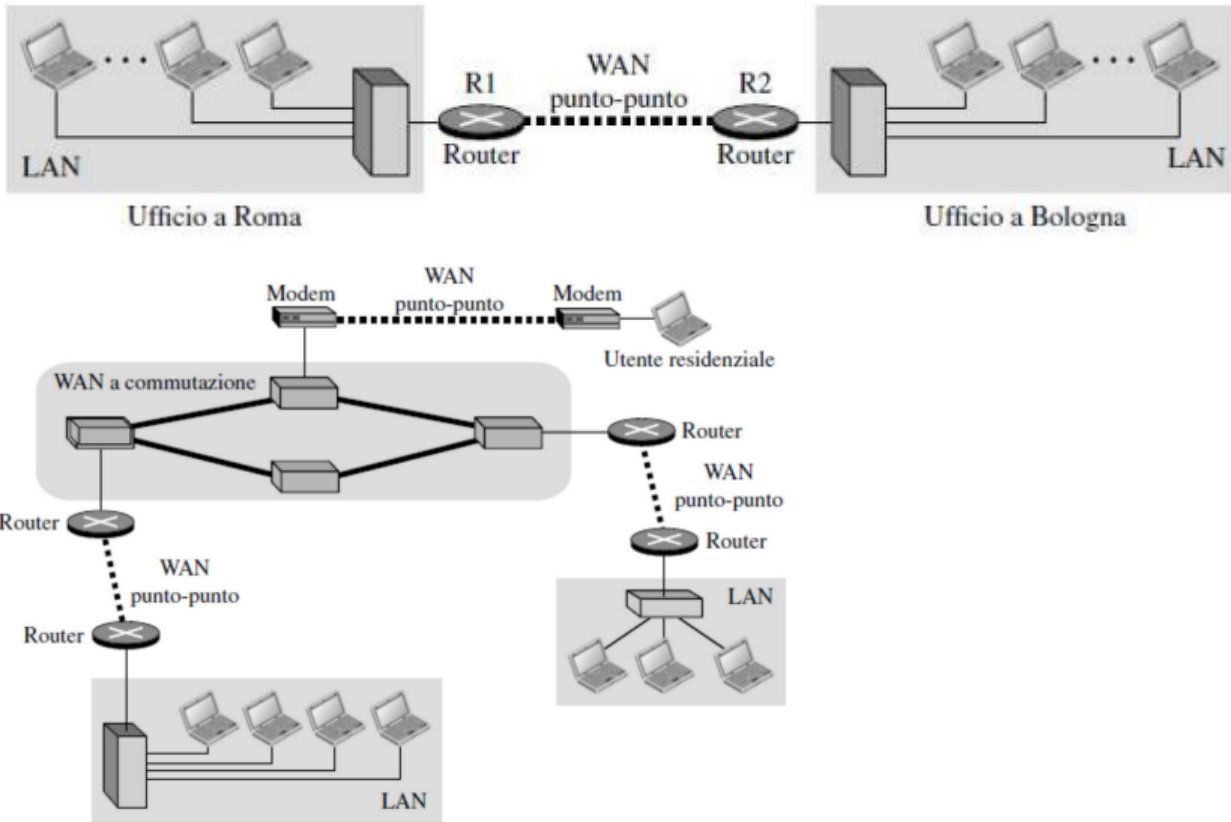
In una rete locale tipicamente una serie di host comunicano tra loro attraverso, ad esempio, uno switch centrale.

**Wide Area Network** Alcuni esempi:



## 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 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, doppiini 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**

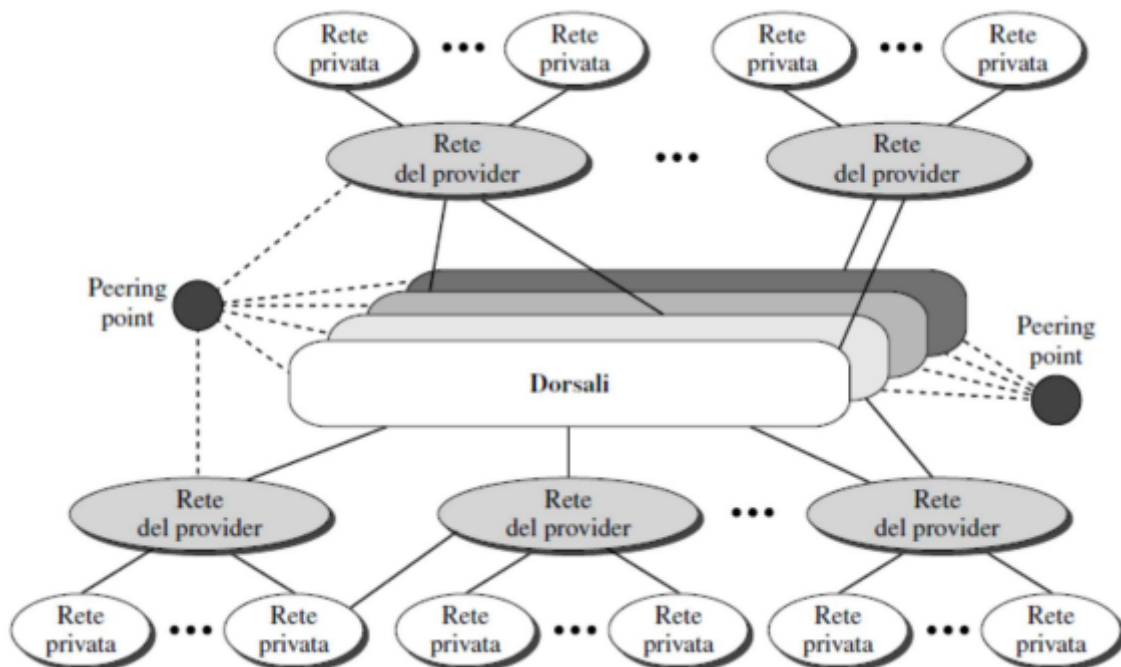
**Standard** di internet e del web: RFC (Request for Comments) e W3C.

**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.



**Peering point:** interconnessione tra due sistemi autonomi

### 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

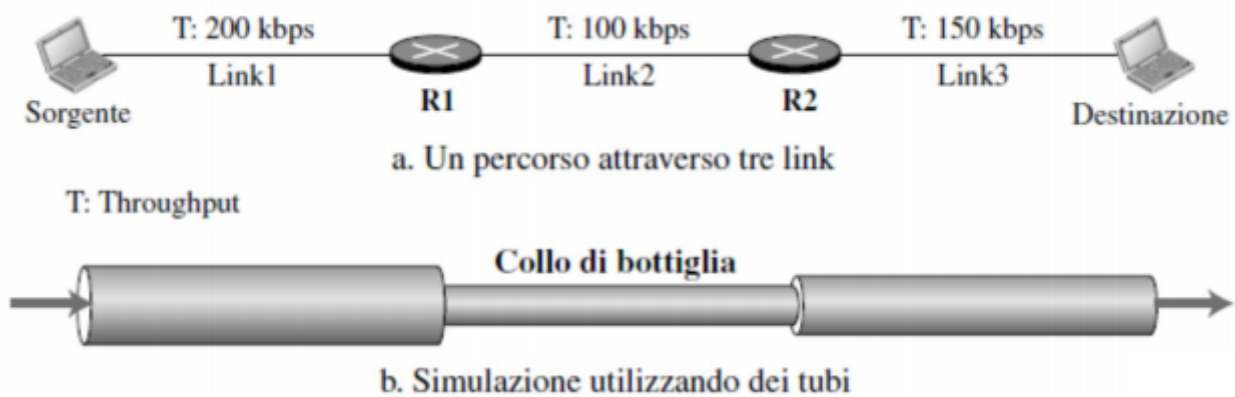
Come misurare le prestazioni della rete? Tramite una serie di metriche:

**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...).

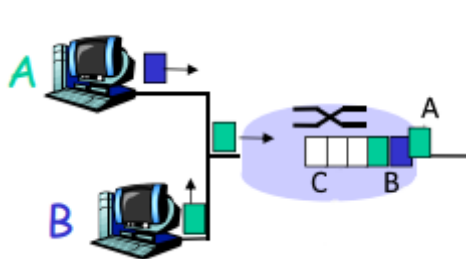


Non è detto che corrisponda alla bandwidth perché ci potrebbe essere un collo di bottiglia.

**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 dei 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 (trascurabile)

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  $3 \times 10^8$  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](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](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 **interfacce** 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 Interconnction 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 trasmissiva, 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.

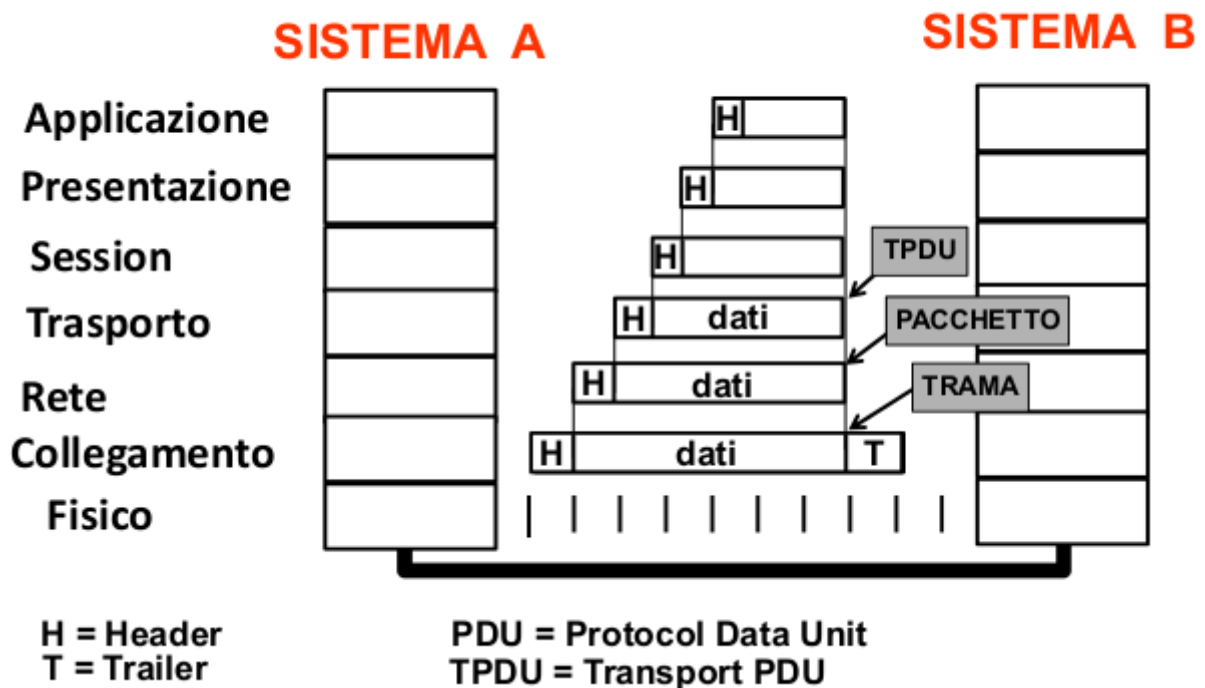
Il **processo di incapsulamento** è quindi **reversibile**: ogni livello esegue una operazione di incapsulamento su dati già incapsulati dal livello precedente, in modo tale da garantire la possibilità di estrarre i dati precedentemente incapsulati.



**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 messaggi, 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

Strutture di rete e funzioni

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

Risoluzione IP e MAC, ARP

IPv4: formato datagramma ip, frammentazione

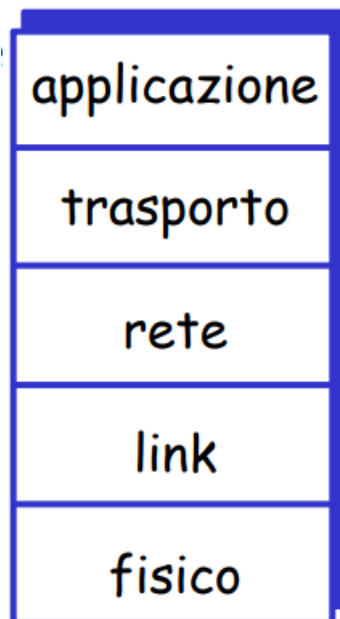
Routing IP e instradamento

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:

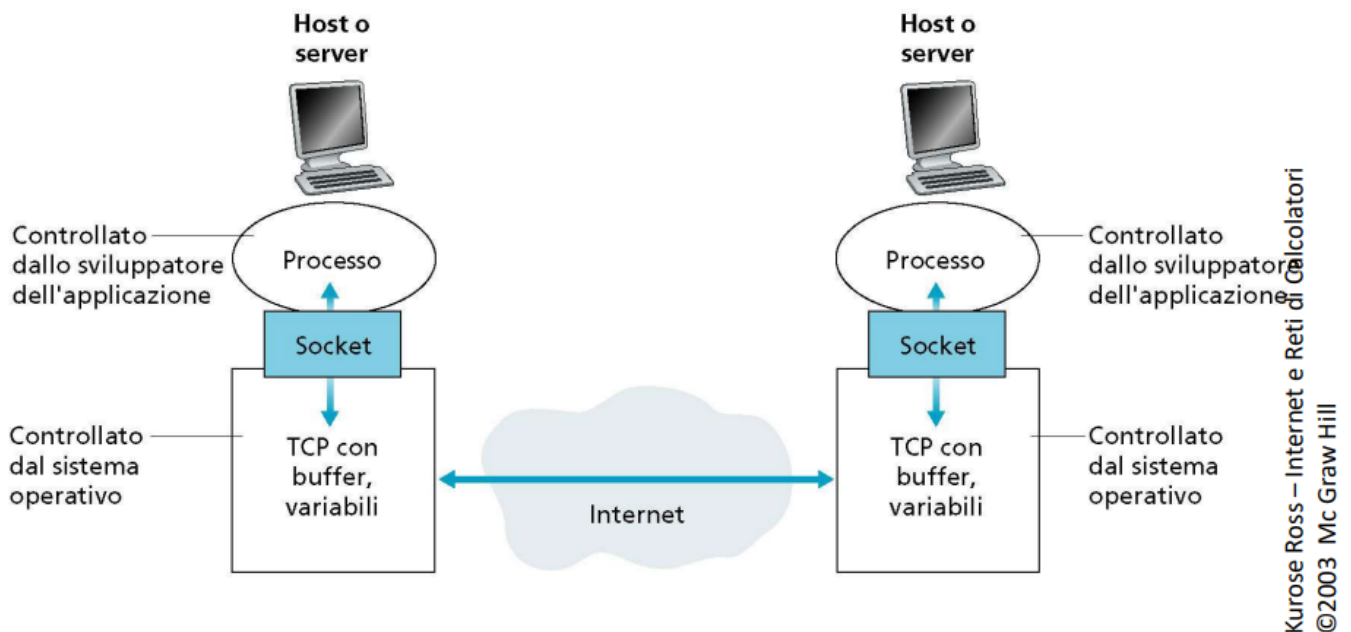
- Programmi di lettura e scrittura sul **client**
- **Server** di posta in internet
- **Standard per il formato** dei messaggi
- **Protocolli** SMTP, POP3 ecc.

## 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 scheme) 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=largo+bruno+pontecorvo+pisa&hl=it
```

```
    scheme = http
```

```
    authority = maps.google.it
```

```
    path = /maps/place
```

```
    query = q=largo+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): *procollo di livello applicazione per sistemi di informazione distribuiti, collaborativi ed impermediali*.

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 **iniziata 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

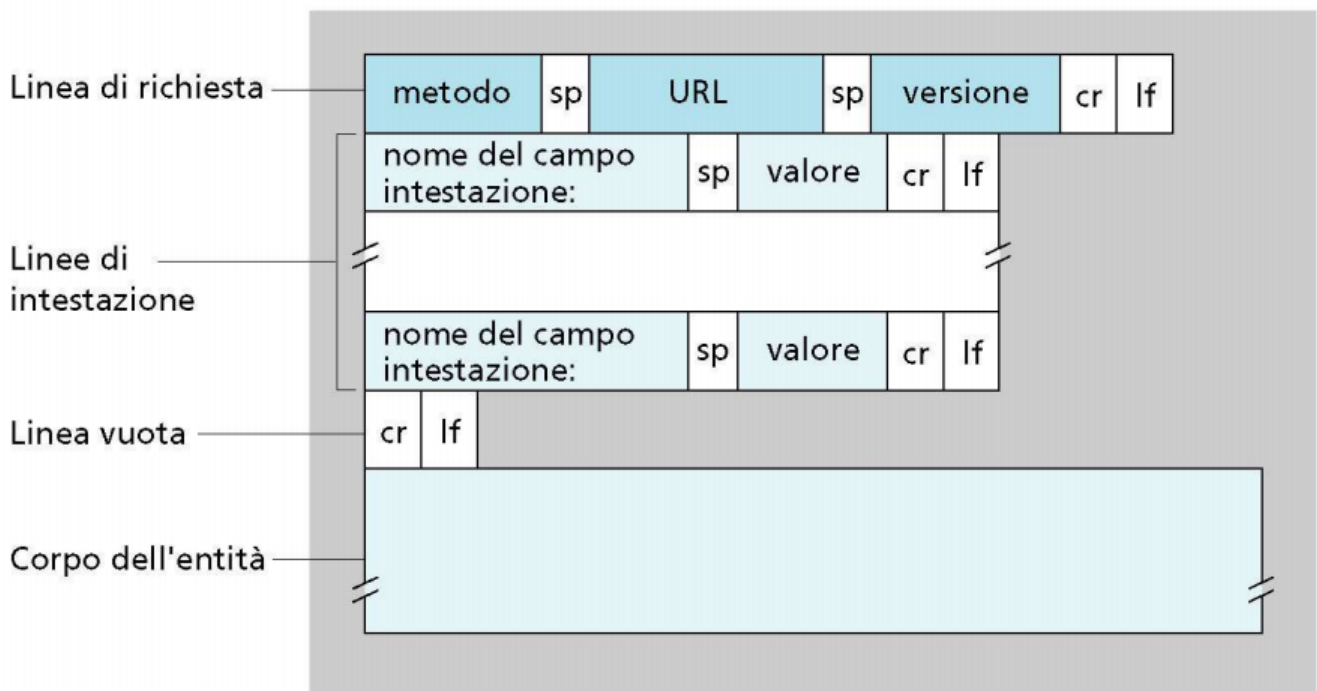
Vedi [slide](#)<sup>1</sup>

## 12.4 Messaggi HTTP

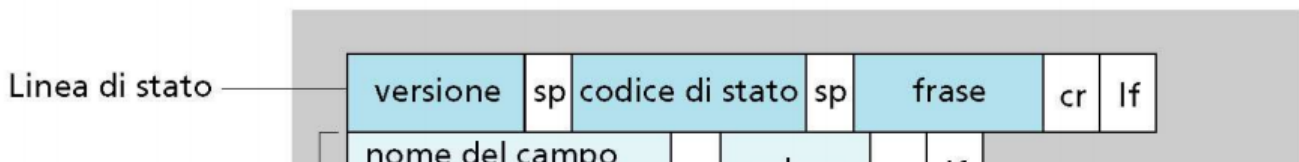
```
generic-message = start-line *message-header CRLF [message-body]
start-line = Request-Line | Status-Line
```

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

### HTTP Request Message



### HTTP Response Message



<sup>1</sup>[https://elearning.di.unipi.it/pluginfile.php/27477/mod\\_resource/content/2/L03\\_Applicativo\\_HTTP.pdf](https://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/WW/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**.

Dopodiché, 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**.

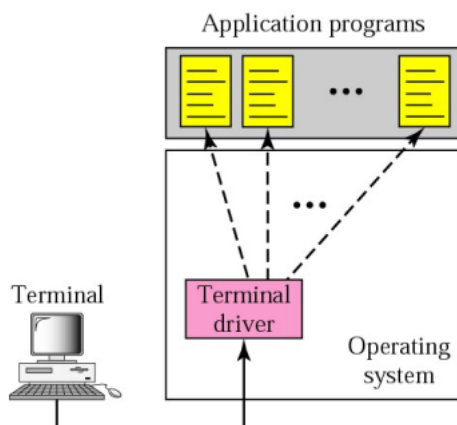
Usa il TCP, con una connessione TCP persistente per tutta la durata della sessione di login, sulla porta 23 del server.

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

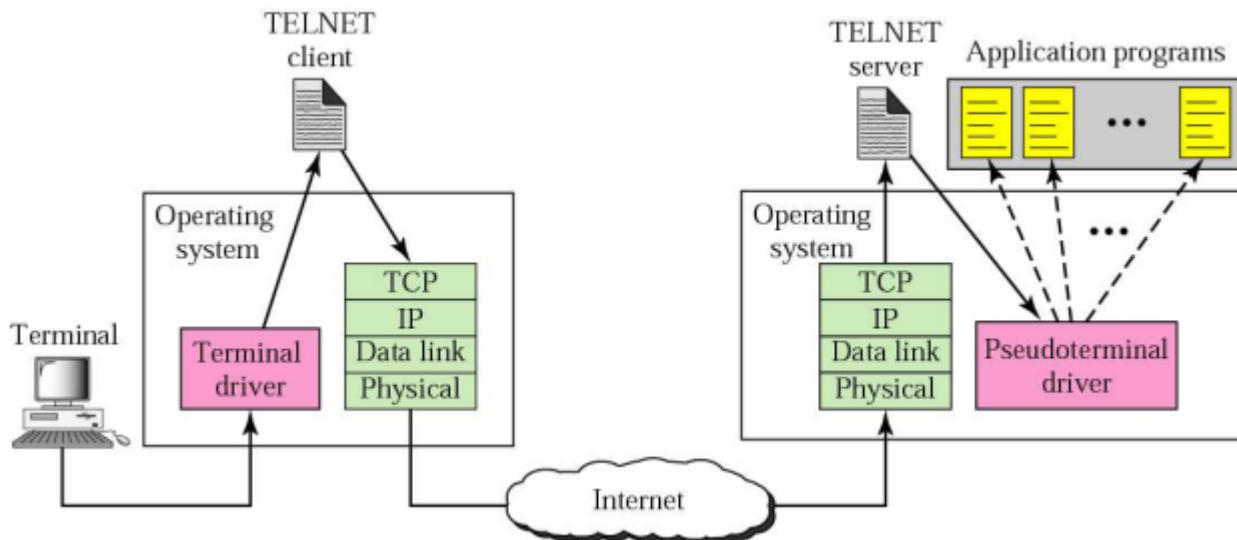
Client **accetta le battute di tasti** sul terminale e **le invia al server**.

**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 trasmetterli al S.O** che li consegnerà all'applicazione opportuna.

### 15.3 NVT

Il Telnet deve **poter operare con il numero massimo di sistem**, 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.

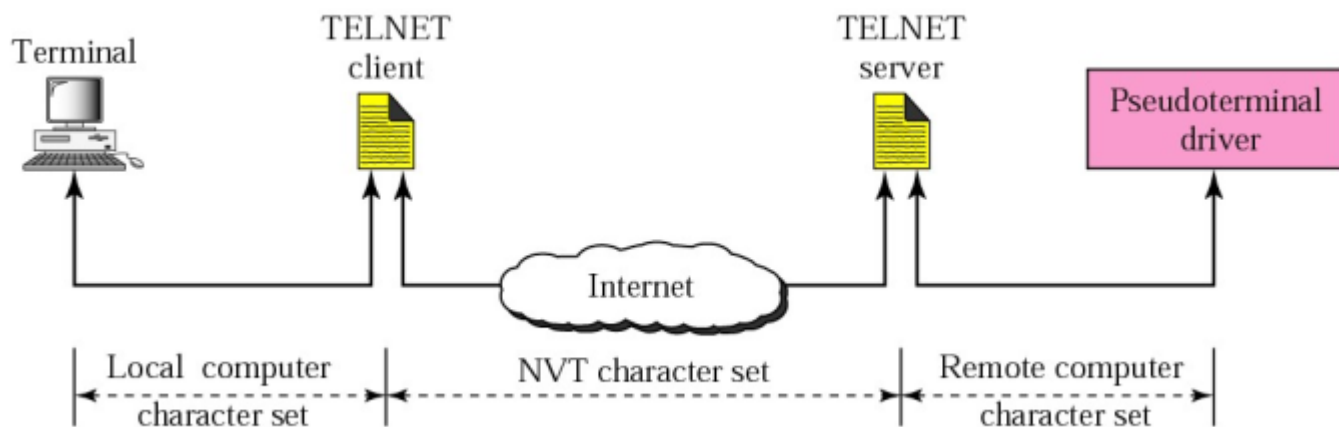
L'NVT è un dispositivo *immaginario* che **fornisce una rappresentazione astratta di un terminale canonico**. Gli host, sia client che server, traducono le loro caratteristiche locali così da **apparire esternamente come un NVT** e assumo che l'host remoto sia un NVT.

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

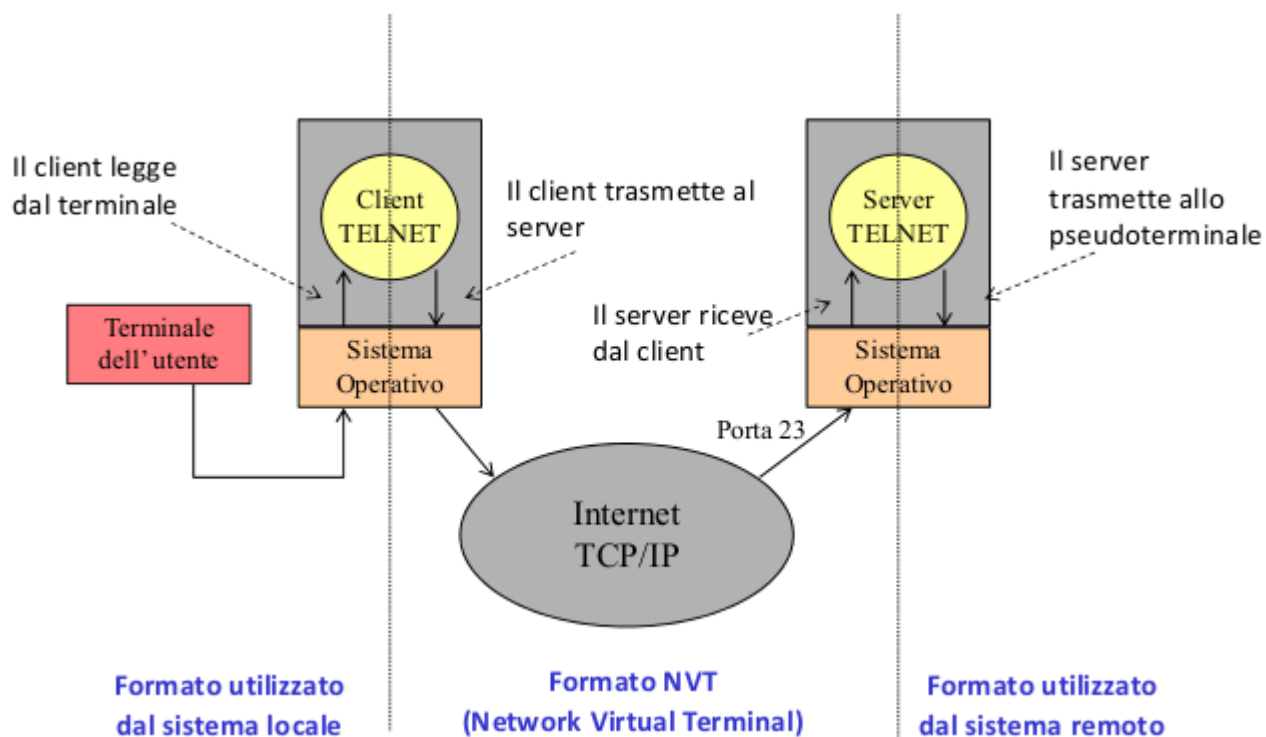
**Trasformare il set di caratteri in uso localmente in un set di caratteri universale** (lettere accentate, tasti freccia, backspace...), includendo anche i caratteri di controllo più importanti (break...)

**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**.

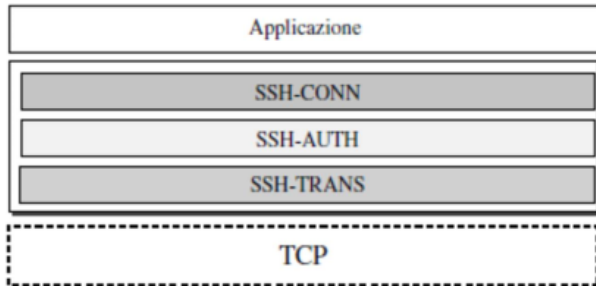
Usando NVT invece devo scrivere soltanto N server e N client, cioè **2N applicativi**.

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 Telnet<sup>2</sup>.



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

- Autenticazione del server
- Negoziazione degli algoritmi di cifratura
- Scambio di chiavi

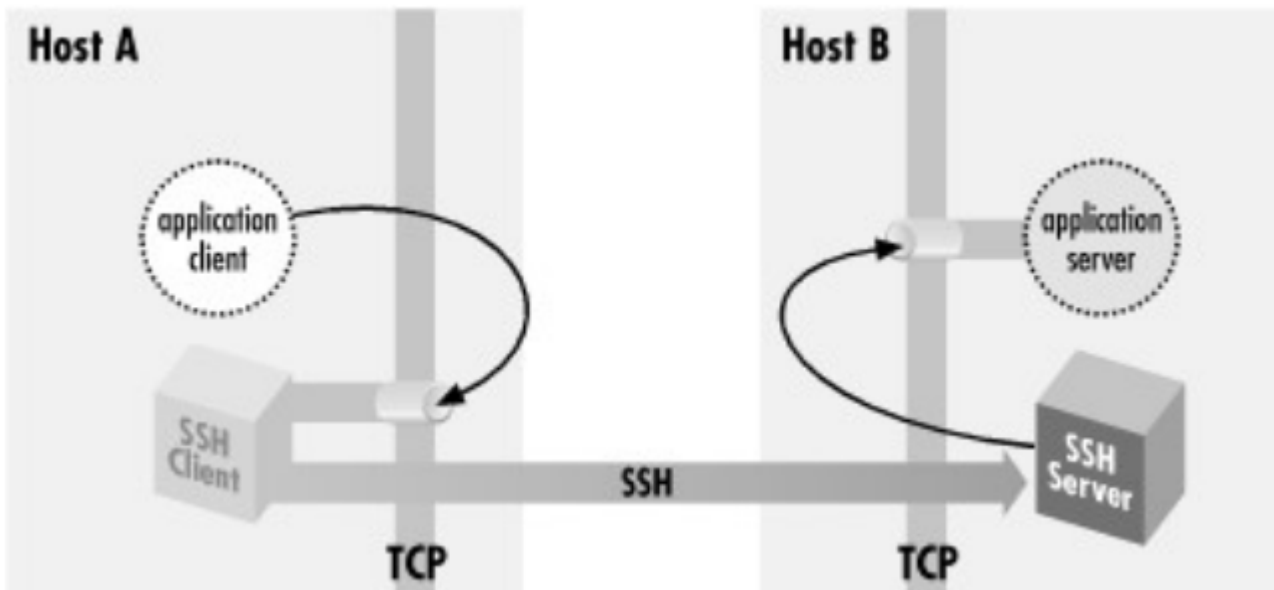
**SSH-AUTHENTICATION** Meccanismi per autenticare l'utente

**SSH-CONNECTION** Sessione 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>

<sup>2</sup>Vedi 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

**Specifica 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

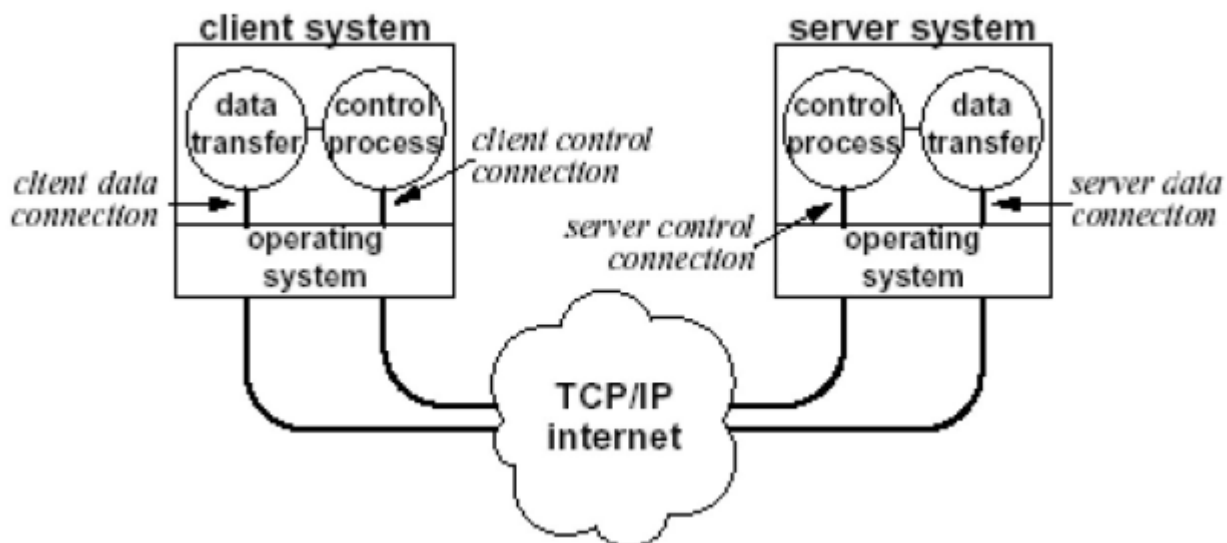
**Data connection:** connessione su cui i dati sono trasferiti con modi e tipi specificati. I dati trasferiti possono essere **parte** di file, **un file** o **un set** di file.

#### 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)

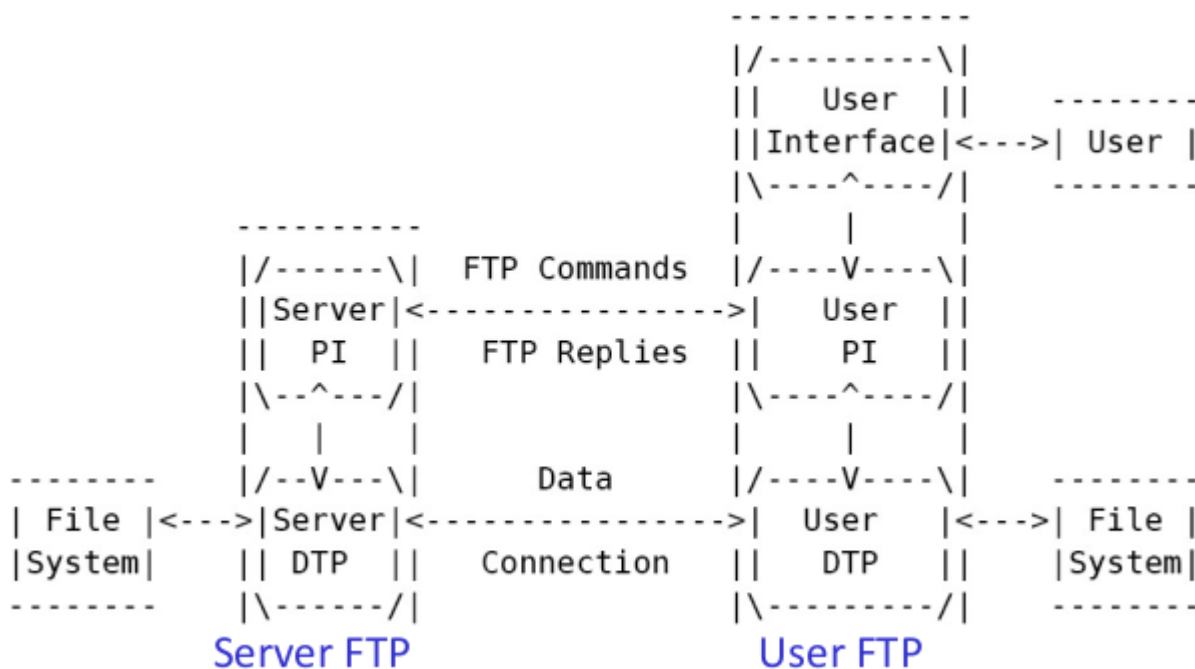
**Passive Mode:** il **client ottiene un numero di porta dal server** (porta 20, non necessariamente).

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 iniziarne 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.

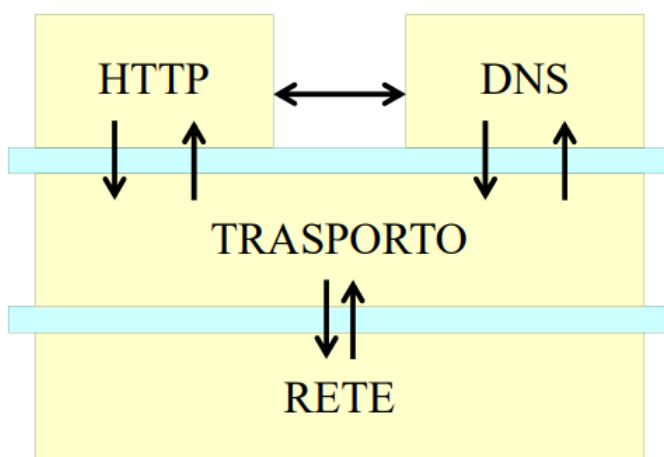
Per questo si utilizza il **Domain Name System**

### 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.

Es. `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 instradamento 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.  
Solitamente prefissato

Indirizzo **di rete (IP)**

Es. `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**

Es `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.

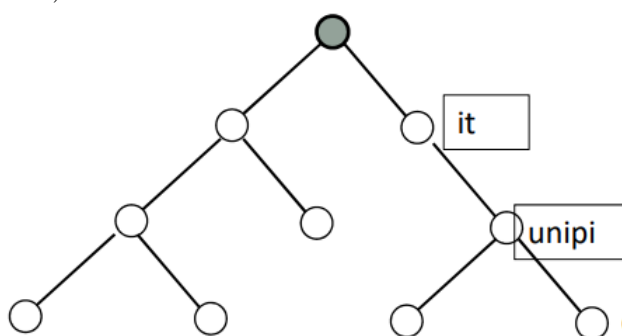
Delega dell'autorità per le varie parti dello spazio dei nomi e **distribuzione della responsabilità** della conversione tra nomi e indirizzi

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.



### Nomi di dominio

- Dominio di terzo livello: di
  - Dominio di secondo livello: unipi
  - Dominio top-level: it
- it
- unipi.it
- di di.unipi.it

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

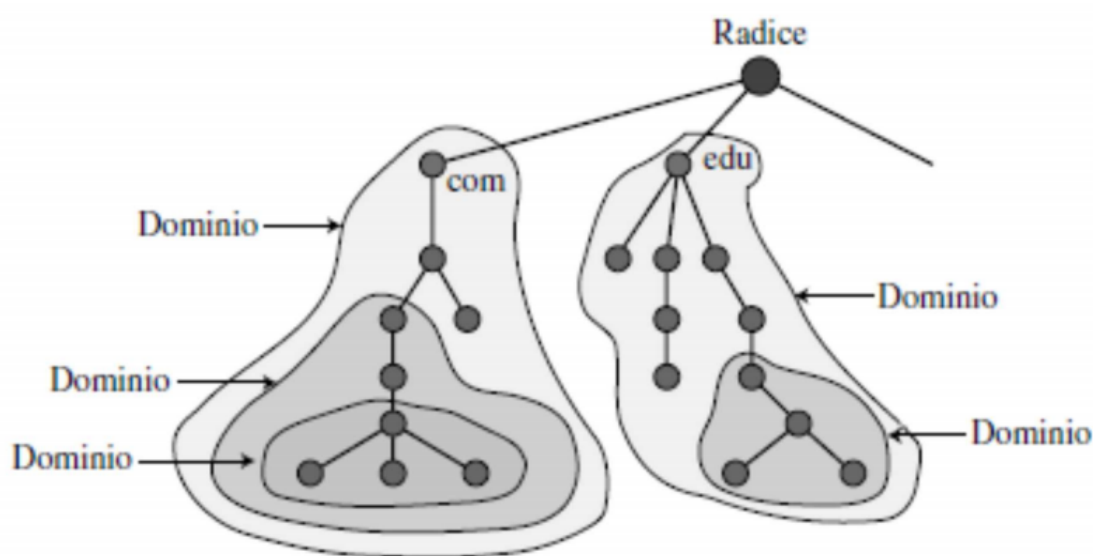
**Dominio** Sottosalbero nello spazio di nomi di dominio che viene identificato dal nome di dominio del nodo in cima al sottoalbero.

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](http://iana.org)

com Organizzazioni commerciali

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

mil Gruppi militari

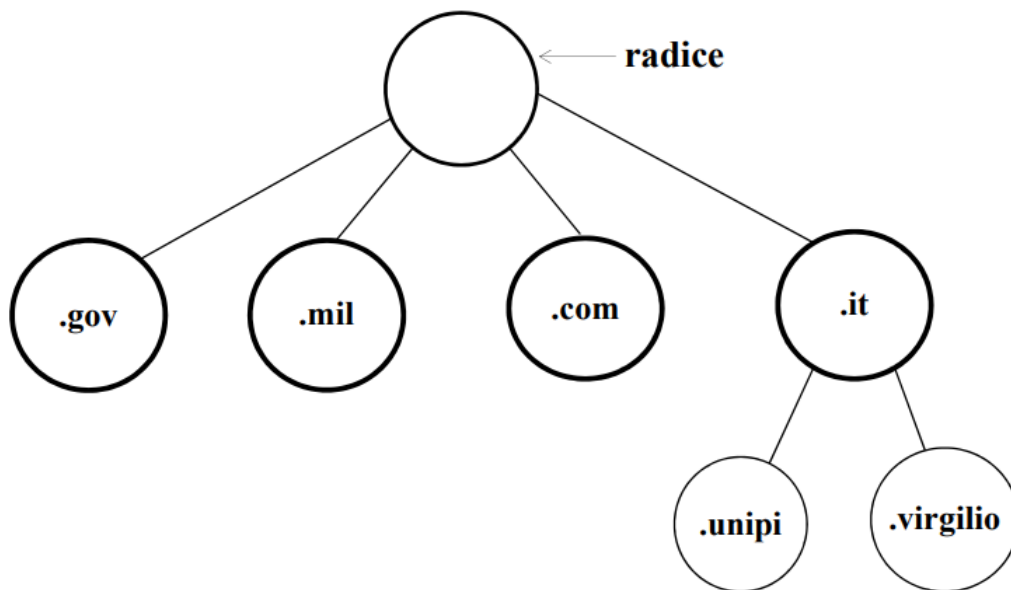
gov Istituzioni governative

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.

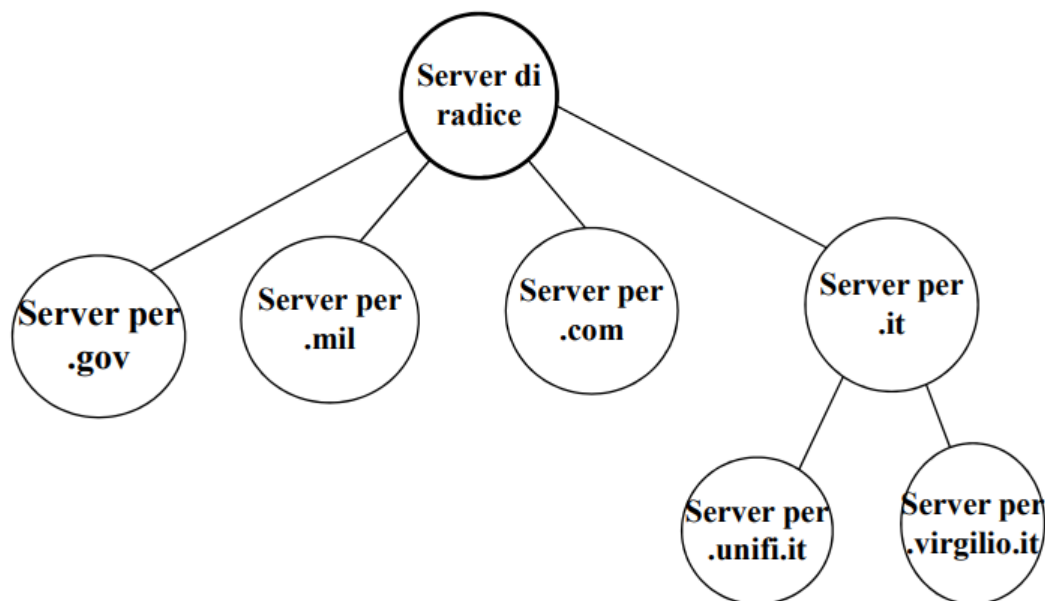
**Database DNS** Il database DNS è **distribuito** ed implementato in una gerarchia di più name servers.

**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.

