

C.so Integrato di statistica, informatica e analisi dei dati sperimentali

prof Carlo Meneghini
Dip. di Fisica "E. Amaldi"
via della Vasca Navale 84
meneghini@fis.uniroma3.it
tel.: 06 57337217

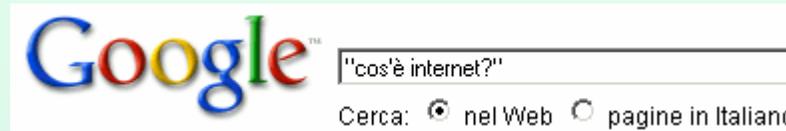
<http://www.fis.uniroma3.it/~meneghini>



La rete internet e il WEB

Bibliografia

Fondamenti di Informatica - Progetto A³ Zanichelli



Introduzione

Internet è un insieme di reti di computer interconnesse tra loro tramite diversi sistemi di collegamento da un cavo coassiale alla fibra ottica, da un cavo telefonico a un ponte radio, ecc.

I computer comunicano tra loro in un linguaggio binario utilizzando un sistema di convenzioni o **protocolli** (TCP/IP, FTP, https, etc..) che si appoggiano sull'Internet Protocol (IP) che rappresenta la base comune per tutti.

Un **protocollo** di comunicazione è un insieme di regole che governano la comunicazione tra due computers.

La comunicazione avviene in modalità client/server: esiste sempre una coppia di programmi, che risiedono rispettivamente presso il computer che vuole utilizzare un certo protocollo e presso il computer che eroga tale servizio. A una determinata richiesta da parte del client si ha una risposta da parte del server.

I computers si identificano tra loro tramite un'indirizzo IP (internet Protocol)

Il **WWW** (World Wide Web) un sistema complesso e intrecciato di server che condividono le proprie informazioni tramite il protocollo HTTP. Il WWW non è internet, ma si appoggia sulla struttura fisica (internet) e utilizza i protocolli di Internet per condividere le informazioni.

I **browsers** sono programmi che forniscono l'interfaccia tra WWW e l'utente.

Le risorse del WWW (informazioni) sono catalogate utilizzando un sistema di indirizzi gestito dall'**URL** (Uniform Resource Locator). Questo sistema permette di identificare in modo univoco un iperoggetto.

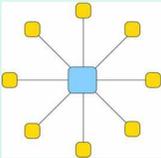
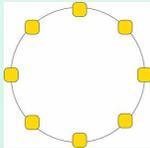
L' **HTML** (HyperText Markup Language) è il linguaggio utilizzato dal browser per visualizzare le informazioni. È un linguaggio ipertestuale ove ogni riferimento a un oggetto esterno è indirizzato tramite un URL.

Una rete di calcolatori è una struttura composta da entità eterogenee:

- diversi Computers: PC, servers, Mainframe, stampanti, etc...;
- diversi sistemi operativi: windows (WXP, Wvista, W2K, ...) Linux, VMS, Linux, MacOS,....;

a loro volta raggruppate in sottoreti diverse per:

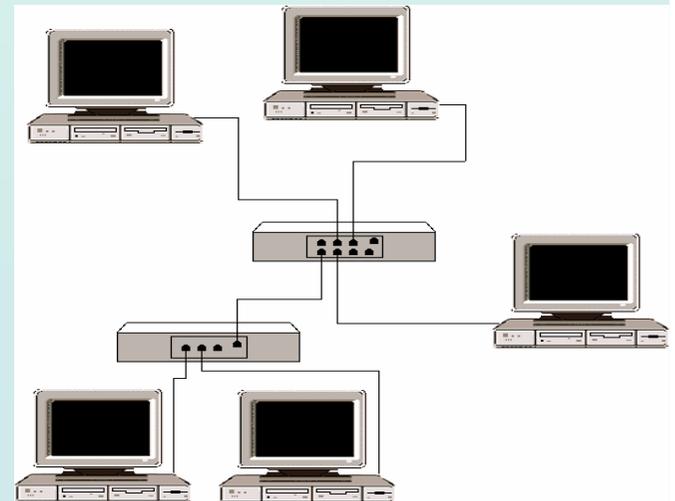
- dimensioni: **LAN**: local area network, **MAN**: Metropolitan area network, **WAN**: wide area network
- supporto di comunicazione: cavo Ethernet, cavo coassiale, doppino telefonico, fibra ottica, connessione senza fili (WiFi, UMTS, Satellitare, etc...);

- topologia della connessione: a stella  ad anello  etc...;

- stabilità della connessione: dedicate (sempre attive), commutate (on demand), mobili (UMTS)

La comunicazione tra questi soggetti estremamente eterogenei è resa possibile dall'esistenza di opportuni **protocolli di comunicazione**, universalmente accettati, riconosciuti e implementati.

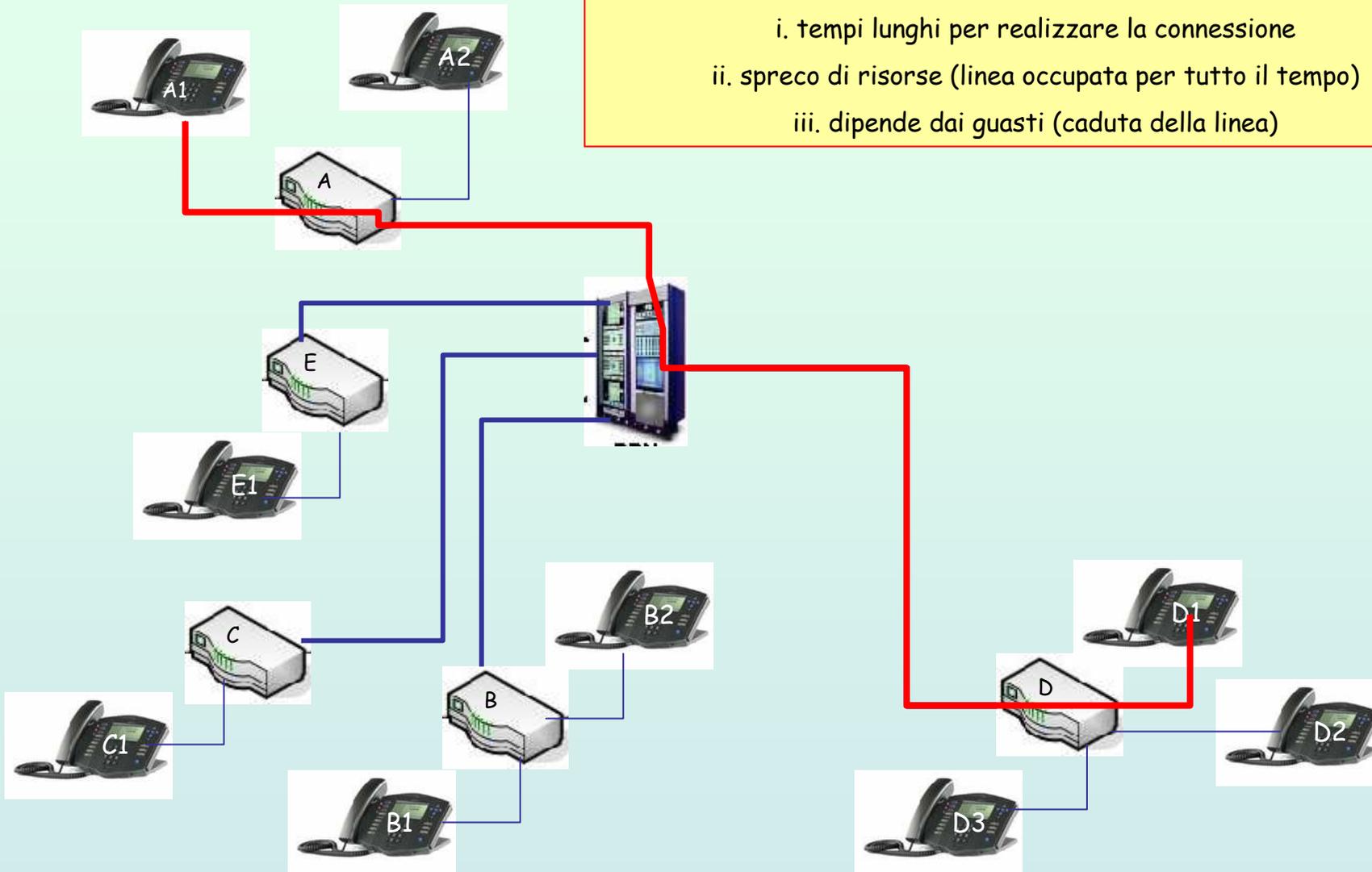
Corso Integrato: I.S.A.D. - modulo
informatica

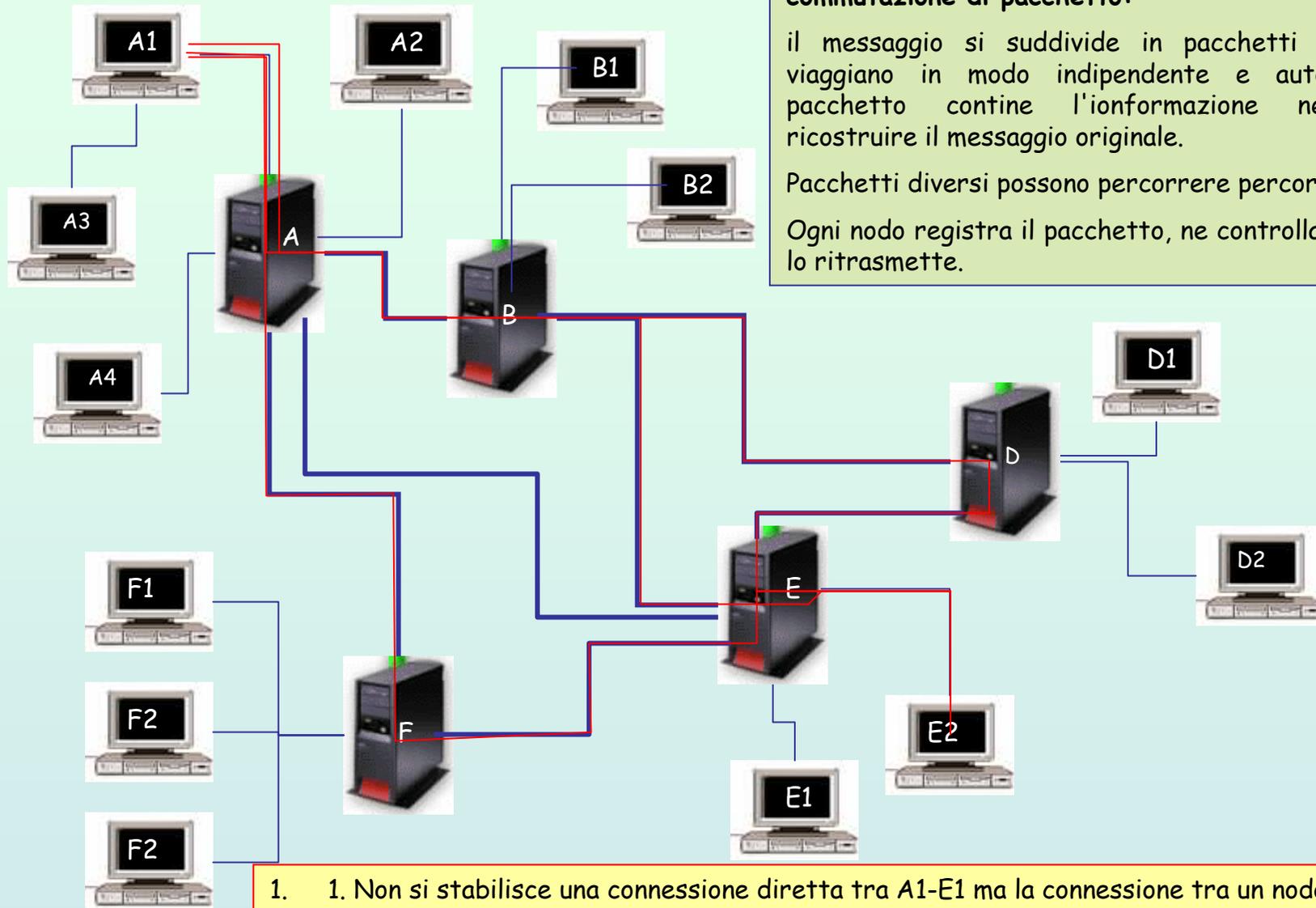


A1

La connessione telefonica (**commutazione di circuito**) non è efficace per Internet:

- i. tempi lunghi per realizzare la connessione
- ii. spreco di risorse (linea occupata per tutto il tempo)
- iii. dipende dai guasti (caduta della linea)





La connessione Internet è basata sulla tecnica a **commutazione di pacchetto**:

il messaggio si suddivide in pacchetti (piccoli) che viaggiano in modo indipendente e autonomo. Ogni pacchetto contiene l'informazione necessaria a ricostruire il messaggio originale.

Pacchetti diversi possono percorrere percorsi diversi.

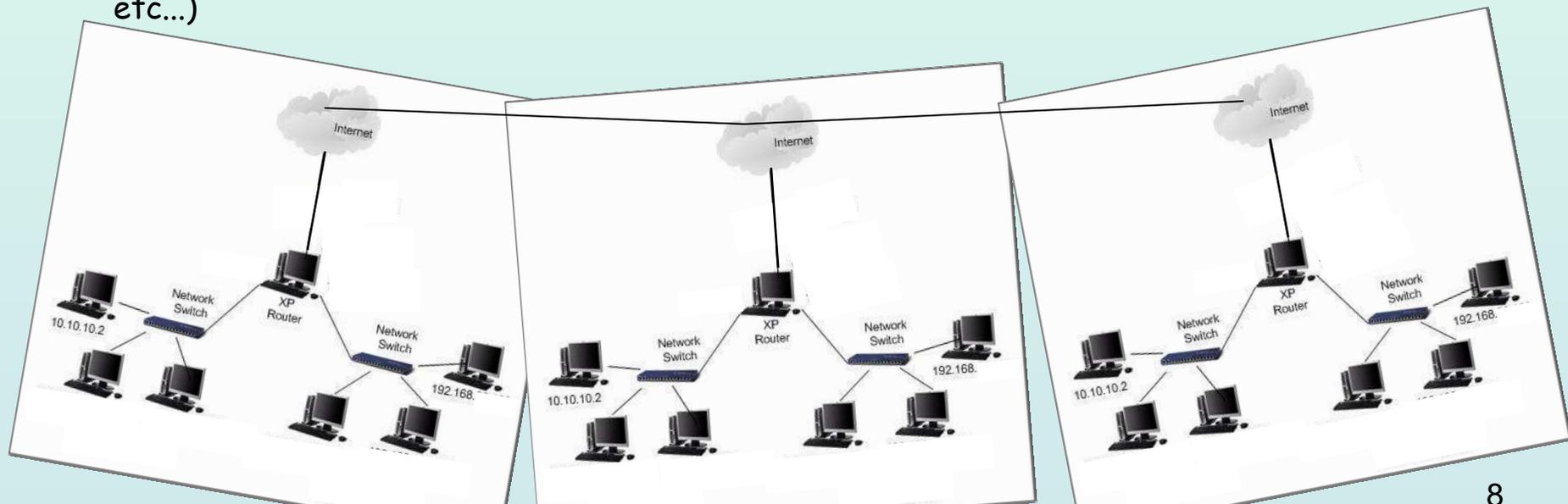
Ogni nodo registra il pacchetto, ne controlla l'integrità e lo ritrasmette.

1. 1. Non si stabilisce una connessione diretta tra A1-E1 ma la connessione tra un nodo e il successivo è attiva solo per il tempo necessario alla trasmissione del pacchetto.
2. 2. Tra un pacchetto e l'altro il nodo è disponibile per ricevere e trasmettere altri pacchetti
3. 3. Se una connessione si interrompe i pacchetti impegnano una strada diversa.
4. 4. Non è necessaria una sincronizzazione: il viene ricostruito quando ricevuto interamente.

La comunicazione avviene in base a **PROTOCOLLI** specifici: un protocollo è un insieme di regole che stabiliscono le modalità della comunicazione.

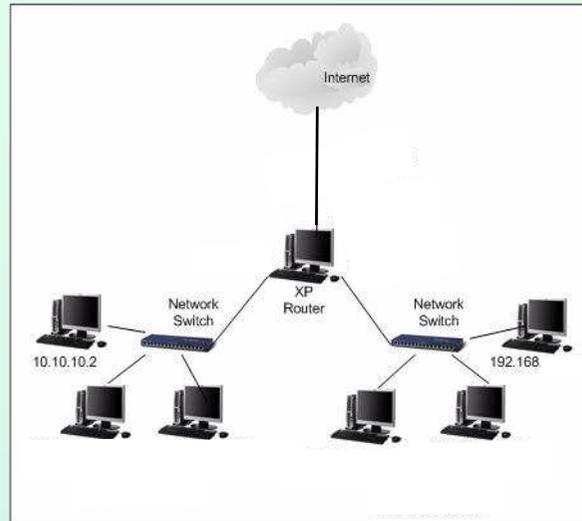
I protocolli operano a livelli diversi:

- Livello **applicazione**: stabiliscono come le applicazioni (programmi) scambiano dati e documenti: HTTP (hyper text transfer protocol), SMTP (simple mail transfer protocol), FTP (file transfer protocol) etc...
- Livello **trasporto** gestiscono il trasferimento dei dati: suddividono il messaggio in pacchetti, controllano l'integrità del messaggio e ne gestiscono la ricostruzione.
- Livello **rete**: gestiscono l'invio e la ricezione dei messaggi sulla rete, in pratica stabiliscono quale percorso deve seguire il pacchetto (IP: internet Protocol)
- Livello **fisico**: stabiliscono le modalità fisiche di trasmissione (frequenza dei segnali, tensione etc...)



La comunicazione tra due reti avviene tramite un Router, esso è il rappresentante della rete

Internet è divisa in modo gerarchico in reti e sottoreti, ogni macchina è individuata da un numero unico: IP number



IP: 192.168.38.121

192: Rete fisica
(I livello)

La IANA (internet assigned number authority) si occupa di assegnare i numeri alle reti

168: Sottorete
(II livello)

I numeri delle sottoreti sono forniti da organismi responsabili per aree geografiche

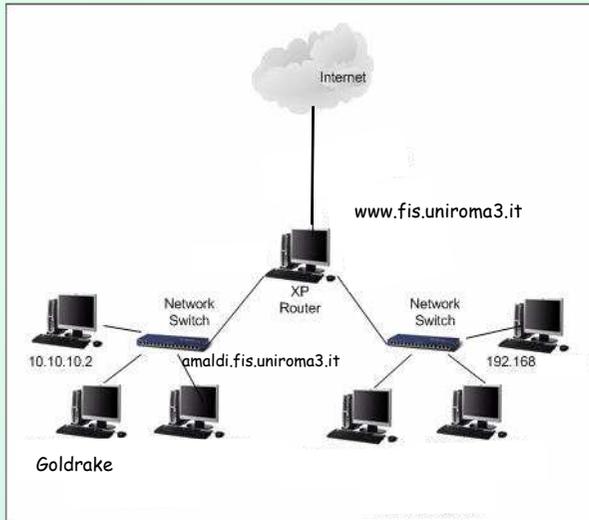
121: macchina
(IV livello)

ad ogni macchina viene assegnato un numero univoco.

38: Nodo
(III livello)

I responsabili delle sottoreti assegnano i numeri ai propri nodi

Indirizzi simbolici (DNS)



Gli indirizzi numerici vanno bene per le macchine, ma sono difficili da ricordare e possono dar luogo a errori di trascrizione. Si usano quindi nomi simbolici:

`www.fis.uniroma3.it` ← Top level domain
`193.204.162.110`

L'insieme di questi nomi simbolici costituisce il Domain Name **System** (DNS)

Il Domain Name Server è una macchina (fisica o virtuale) che fornisce il servizio di Domain Name Resolver, cioè associa ad un indirizzo simbolico il corretto IP. Diversi indirizzi simbolici possono essere associati ad uno stesso indirizzo IP, non è vero il contrario.

I nomi logici (o simbolici) sono assegnati da organismi internazionali in modo gerarchico.

Interazione/comunicazione tra computers

Metodi di comunicazione (architetture)
Protocolli di comunicazione
Servizi
Applicazioni

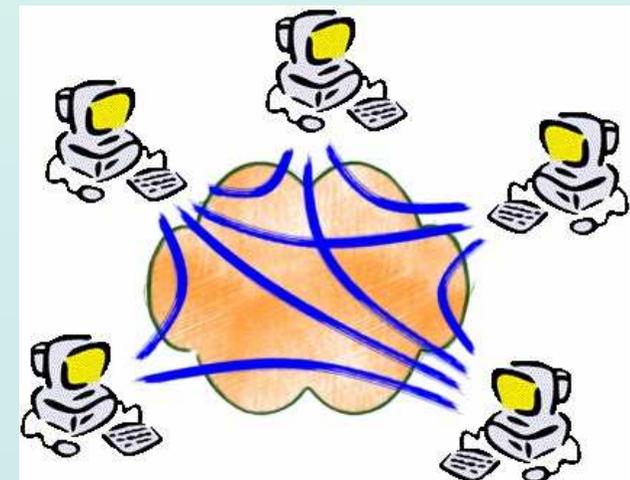
Metodi di comunicazione: Client/server oppure P2P

Client/server:

il client richiede un servizio, il server (un programma o una macchina dedicata) lo fornisce
(DNS, SMTP, etc..)

P2P:

la comunicazione avviene tra "pari"



Protocolli

Servizi

Applicazioni

Protocollo: insieme di regole e convenzioni da rispettare.

Servizio: prestazione fornita da un macchina (server) ad un client

applicazioni: programmi

Uno stesso termine può avere significati diversi:

FTP: è un protocollo(file transfer protocol) per il trasferimento di files tra un computer e l'altro

FTP: è un servizio fornito da server FTP dai quali è possibile prendere o depositare files (usando un protocollo FTP)

FTP: è un'applicazione (programma) che consente di usufruire di servizi FTP messi a disposizione di server FTP mediante il protocollo FTP!!!

Il WEB (World Wide Web) è un ipertesto multimediale distribuito dotato di un'interfaccia.

Ipertesto: tecnologia informatica usata per realizzare testi con collegamenti esterni (iperlink)

Multimediale: le informazioni sono di tipo diverso (scritte, immagini, video, suoni, etc...)

Distribuito: le diverse parti del WEB risiedono su calcolatori diversi

Interfaccia: l'insieme dei programmi che si prendono carico di visualizzare i dati multimediali

L'ipertesto è un documento in cui le diverse parti possono:

- descrivere l'informazione
- rappresentare un collegamento (link) ad altri ipertesti
- contenere istruzioni per rappresentare l'informazione

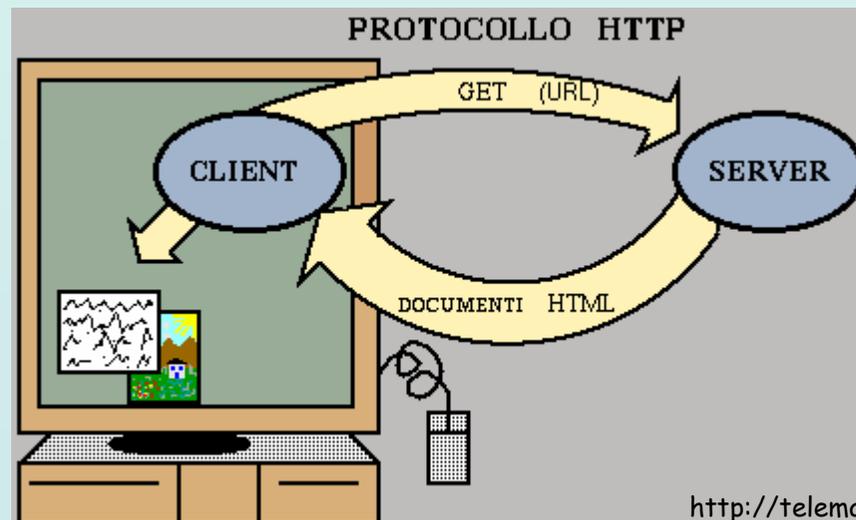
Il Browser consente di visualizzare l'informazione in modo indipendente dalla macchina su cui risiede e opera.

Il WEB è organizzato su un'architettura tipo server/client: i server mettono a disposizione le informazioni (ipertesti) i client accedono alle informazioni (risorse) messe a disposizione.

Il server è un programma, basato sul protocollo HTTP, che interpreta le richieste e fornisce le informazioni.

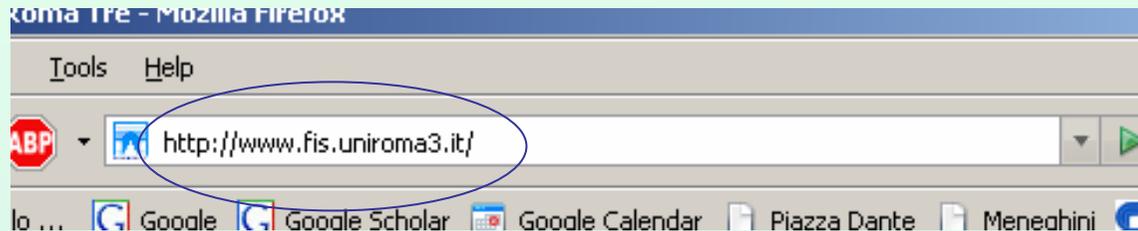
Il client è costituito da un Browser.

La localizzazione delle risorse avviene attraverso il sistema URL (uniform resource locator).



La barra degli indirizzi: URL (uniform resource locator)

L'URL permette di definire la localizzazione e il tipo di accesso alle risorse di internet



Un server è una macchina dove sono attivi programmi che forniscono servizi (server). Questi programmi, solitamente dormienti, hanno attivi piccoli sottoprogrammi (demoni) che "ascoltano" le comunicazioni in arrivo e "svegliano" il programma server corrispondente. Esempio di server e demoni sono:

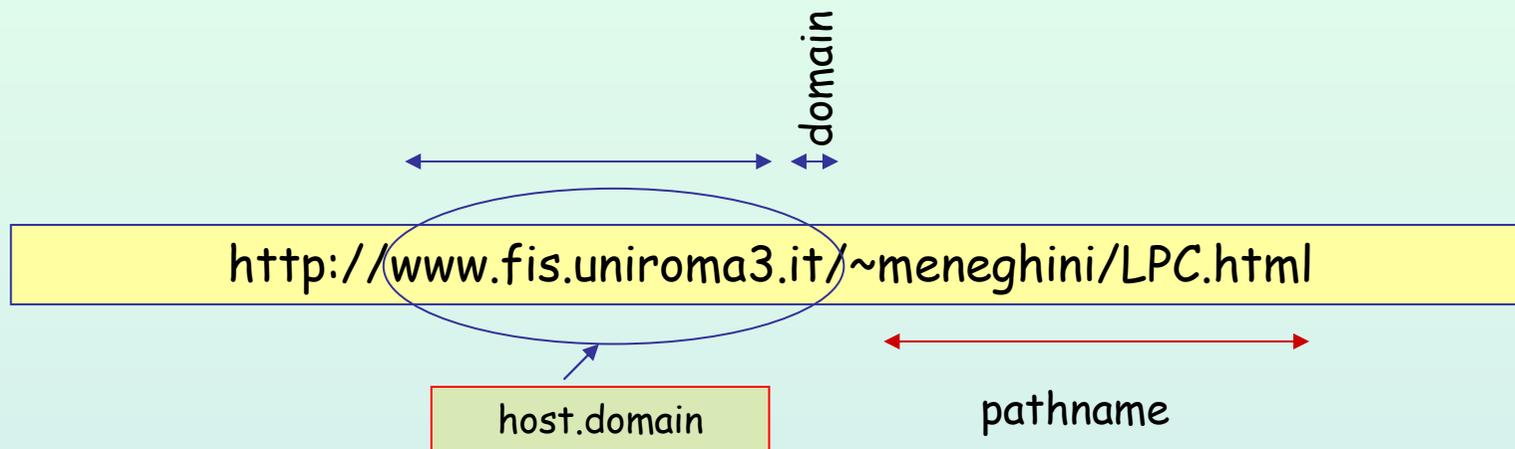
mailer-demon, http-demon, ftp-demon etc...

Un server HTTP è un programma che gira su una macchina ospite (server). Quando il demon HTTP legge una richiesta che inizia con **http://**, il demon indirizza la comunicazione al server HTTP che interpreta il resto del messaggio e fornisce la risorsa richiesta.

http://www.fis.uniroma3.it/~meneghini

protocollo
(server)

indica il modo in cui accedere alla risorsa: quale programma usare e con quale protocollo avviene la comunicazione **http, ftp, mailto, news, file, telnet, https, ...**



Domini

2 caratteri: it, uk, br, ... nazionali

3 caratteri: tematici: com, org, net, edu ...

pathname: indica la pagina da visualizzare. Il pathname è CASE sensitive:

.../~meneghini/**lpc**.html

è diverso da

.../~meneghini/**LPC**.html

Il browser fornisce l'interfaccia tra l'utente e il server HTTP

1. Si collega al server
2. interpreta e trasmette le richieste del client
3. visualizza i documenti e le informazioni
4. consente il collegamento ipertestuale
5. fornisce l'interfaccia ad applicazioni aggiuntive (mediante plug-in)