
Internet



Internet

- Nata dalla fusione di reti di agenzie governative americane (ARPANET) e reti di università
- E' una **rete di reti**, di scala planetaria, pubblica, a commutazione di pacchetto
 - sistema di comunicazione tra reti e sistemi eterogenei, oltre che geograficamente distribuiti.
- Utilizza protocolli di comunicazione di dominio pubblico derivati dal modello **TCP/IP**
- Al giorno d'oggi è accessibile a tutti tramite **provider** che sono fisicamente collegati a internet

Internet

- Collega fra loro reti locali e metropolitane di tutto il mondo.
- La comunicazione tra le sottoreti avviene sfruttando canali di comunicazione dedicati ad alta tecnologia (ISDN, ATM, fibre ottiche) che consentono velocità di trasmissione dell'ordine di decine o centinaia di Megabit/sec.
- Ovviamente non avrebbe senso usare la rete telefonica, che è troppo lenta.

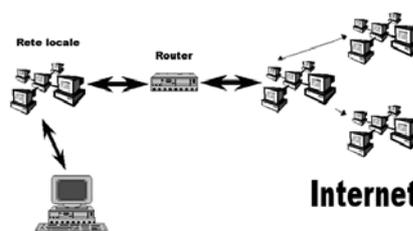
Internet

- Ogni rete locale "si affaccia su Internet" attraverso un dispositivo (un vero e proprio computer) detto **router**. Il router si occupa di smistare il traffico dei pacchetti in uscita ed in entrata nella rete locale rispetto a internet.
- Al router è spesso associato anche un dispositivo detto **firewall**. Il firewall protegge la rete locale da accessi indesiderati dall'esterno (e in alcuni casi, viceversa).

Utilizzo di Internet

- Cosa serve?
 - un computer
 - una connessione alla rete
 - collegamento mediante router
 - *oppure*: linea telefonica + modem + provider
 - programmi client che consentono di usufruire dei servizi della rete

Accesso ad Internet con collegamento diretto



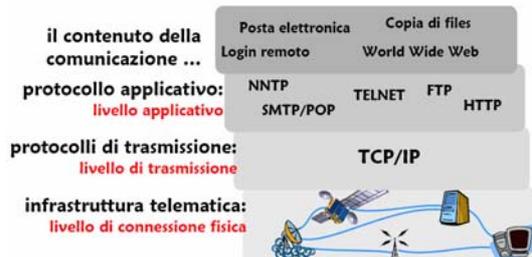
Accesso ad Internet con collegamento commutato



I provider

- **Internet Service Provider (ISP)**
 - organizzazioni, pubbliche o private, che affittano all'utenza privata accessi ad Internet
- Un provider dispone di uno o più host collegati ad Internet con linee dedicate, attive 24 ore su 24
 - gli host possono fornire temporaneamente accesso ai servizi di rete a decine o centinaia di computer mediante modem e linee telefoniche

La struttura di Internet



Indirizzamento in Internet

- Il computer ha anche un **indirizzo fisico**: l'indirizzo con il quale il software che gestisce le comunicazioni in rete localizza e gestisce la comunicazione con quel computer.

Indirizzi numerici vs indirizzi simbolici

- Gli indirizzi IP sono **machine-oriented**, quindi difficili da utilizzare per un utente "umano";
- è stato definito un sistema per passare da indirizzi numerici (gli **indirizzi IP**) a **nomi** facilmente memorizzabili, il **Domain Name System**;
- **Domain Name System (DNS)**
 - associa a ogni indirizzo IP uno o più indirizzi simbolici,
 - gestisce la conversione tra indirizzi simbolici e indirizzi IP
- organizzato in **maniera gerarchica** (domini, sotto-domini, sotto-sotto-domini, ...) per semplificarne l'utilizzo.



Domain Name System

- Gli indirizzi logici Internet hanno la forma:
aaa.bbb.ccc
e sono organizzati in **domini** e **sottodomini**
- I domini possono essere geografici:
xxx.yyy.**it** aaa.bbb.**uk** jjj.kkk.**ca**
- o di altro tipo:
aaa.**com** bbb.ccc.ddd.**net** xxx.**gov**

Domain Name System

- I domini corrispondono a suddivisioni logiche di Internet per facilitare la gestione dei nomi delle risorse
- Internet è suddivisa in una moltitudine di domini radice
 - Domini nazionali: *uk it de*
 - Domini generici: *com edu*
- Un dominio radice include una collezione di host

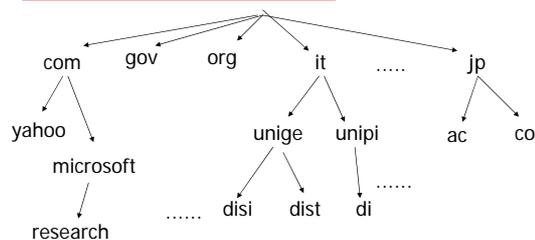
Domain Name System

- All'interno dei domini vi sono i sotto-domini, sotto-sotto-domini, e così via, fino eventualmente a raggiungere lo specifico computer della rete locale relativa:
 - *xxx.unito.it* (sotto-dominio dell'univ. di Torino)
 - *di.unito.it* (sotto-sotto-dominio del Dip. Di Informat., a cui corrisponde fisicamente la rete locale del Dip.)
 - *pianeta.di.unito.it* (uno dei computer della rete locale del Dipartimento di Informatica dell'Univ. Di Torino, Italia)

Domini di primo livello

com - aziende	au - Australia
edu - università americane	ch - Svizzera
gov - istituzioni governative	fr - Francia
mil - istituzioni militari	it - Italia
net - fornitori d'accesso	jp - Giappone
org - organizzazioni non-profit	uk - Inghilterra
....

Spazio dei nomi dei domini



Il Domain Name System

- Ogni dominio è gestito da un computer che contiene l'associazione tra i nomi logici e gli indirizzi fisici di ogni suo sotto-dominio
- Ogni sottodominio è gestito da un computer che contiene l'associazione tra i nomi logici e gli indirizzi fisici di ogni suo sotto-sotto-dominio
- In questo modo è facile localizzare qualsiasi punto della rete usando solo nomi logici.

Il Domain Name System

- Ad esempio, se da un pc negli USA si vuole comunicare con la macchina *pianeta.di.unito.it*.
- Viene contattato il gestore del dominio *.it* (130)
- che contatta il gestore del dominio *.unito* (192)
- che contatta il gestore del dominio *.di* (239)
- che sa che la macchina *pianeta*, all'interno della rete locale *.di* (239) ha numero 1.
- Viene quindi restituito l'indirizzo fisico che sarà usato nella comunicazione: **130.192.239.1**

Domain Name System

- Lo spazio dei nomi è memorizzato quindi sotto forma di database distribuito
- Ogni rete locale ha un proprio server DNS che mappa nomi logici (indirizzi DNS) in indirizzi fisici (indirizzi IP)
- N.B.: Indirizzo IP = codice numerico (binario) utilizzato dal protocollo di invio dati del modello TCP/IP
 - Es. 121.34.16.19

Protocollo di comunicazione TCP/IP

- **TCP/IP (Transmission Control Protocol /Internet Protocol)** e' il protocollo di comunicazione usato in internet e anche nella maggior parte delle altre reti.
- Praticamente tutti i servizi offerti da Internet sono costruiti usando TCP/IP
- Il modello ISO-OSI è in parte ispirato al TCP/IP

Protocollo di comunicazione TCP/IP

- Suddiviso in vari livelli
 - **Livello fisico:** interfaccia fisica tra le stazioni per la trasmissione dei dati e il mezzo di trasmissione
 - **Livello di accesso alla rete:** scambio di dati fra un sistema finale e la rete a cui è collegato, specificando come organizzare i dati in frame e come trasmetterli sulla rete
 - **Livello internet –IP (Internet Protocol):** Scambio di dati tra sistemi che non appartengono alla stessa rete: occorrono delle procedure per attraversare reti multiple interconnesse. Specifica il formato dei pacchetti inviati attraverso la rete ed i meccanismi utilizzati per farli transitare dal calcolatore sorgente attraverso uno o più router verso il destinatario
 - **Livello di trasporto-TCP (Transmission Control Protocol):** trasforma dati in messaggi
 - **Livello di applicazione:** come un'applicazione può usare l'insieme dei protocolli TCP/IP

Protocollo TCP/IP

- Suddivide i dati in uscita in pacchetti recanti le informazioni circa la loro destinazione
 - I pacchetti sono spediti separatamente e poi riassemblati dal ricevente (Internet è una rete a commutazione di pacchetto)
- Provvede all'instradamento dei messaggi
- Controlla che la comunicazione vada a buon fine
 - meccanismo di messaggi di conferma della ricezione avvenuta o di segnalazione di errori, onde la trasmissione venga ripetuta

Protocollo TCP/IP

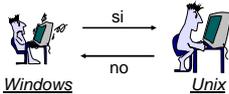
- Il protocollo analizza la parte rete dell'indirizzo:
 - se la rete è quella locale, i dati sono inviati direttamente all'host indicato nell'indirizzo;
 - altrimenti sono inviati al router.
- Ogni router ha in memoria una tabella (tabella di routing) degli indirizzi dei router responsabili di altre sottoreti che conosce direttamente, più uno per gli indirizzi di rete che gli sono sconosciuti.

Strumenti e servizi Internet

- **Telnet** (per accesso a macchine remote)
- **FTP** (per trasferimento file)
- **Talk/Chat** (per conversare in tempo reale)
- **Mailing Lists** (liste di discussione)
- **NewsGroups** (gruppi di discussione)
- **HTTP** (per trasferimento pagine Web)
- **SMTP** (per posta elettronica)

TELNET

- ❑ Protocollo e programma per collegarsi ad un computer remoto
- ❑ Molti servizi per la ricerca di informazioni su internet sono basati su telnet
- ❑ Funziona solo verso macchine multi-user predisposte per telnet (Unix/Linux)

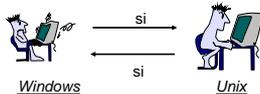


File Transfer Protocol (FTP)

- ❑ Protocollo per trasferire files su internet
- ❑ Programma FTP corrispondente
- ❑ Accesso con parola chiave o "anonimo" (anonymous FTP)
- ❑ versione grafica di FTP (WS_FTP)

Per usare FTP ci vuole

- ❑ Un Indirizzo Internet: **mat.unical.it**
- ❑ Un user-id (login): **utenti**
- ❑ Una Password: **utenti**
- ❑ Un computer predisposto per ftp (di solito una macchina Unix con ftp)



CHAT



- ❑ Evoluzione del servizio TALK
- ❑ Salotti Virtuali
- ❑ Conversazioni "Molti a Molti"
- ❑ Uno dei servizi più usati della rete

Mailing list (liste di discussione)



- ❑ strumenti di discussione attraverso mail elettronica
 - esistono mailing list su svariati argomenti e in tutte le lingue
 - sulle M.L. avvengono discussioni, si fanno annunci, si distribuiscono bollettini elettronici ...
- ❑ una mailing list e' un indirizzo di mail che redistribuisce il messaggio ad una lista di indirizzi (gli iscritti alla mailing list)
 - es: mail mathlist@mat.unical.it

Newsgroups - Usenet (conferenze in rete)



- ❑ A differenza dei messaggi postali, quelli inviati a una conferenza non hanno un vero e proprio destinatario
 - sono semplicemente 'affissi' su una bacheca virtuale.
 - Chi passa di lì, ed è interessato all'argomento, può leggerli, commentarli, rispondere.
- ❑ Possono essere
 - moderate: qualcuno controlla i messaggi prima di pubblicarli nella "bacheca"
 - libere
- ❑ Per accedere occorre un programma "newsreader"

Newsgroups

- Mettono automaticamente in contatto persone che desiderano:
 - cercare/fornire informazioni
 - scambiare opinioni
 - litigare!
- su uno specifico argomento
- Sono una delle fonti principali per recuperare informazioni su Internet: **testi, programmi, immagini, suoni, filmati.**
- NG = **paniere elettronico distribuito**
 - NewsGroups: libertà e anarchia

Newsgroups: funzionamento

- Gestione completamente distribuita
- Molti News-servers sparsi per il mondo:
 - Ogni News-server riceve i messaggi dei suoi utenti e li invia agli altri news-servers
- I messaggi durano un tempo limitato

Newsgroups: gerarchie

- I newsgroups sono organizzati in gerarchie ad albero per argomenti e sotto-argomenti
- La forma generale del nome di un newsgroup e' del tipo:
 - categoria.argomento.sotto-argomento
- alt (gli argomenti più disparati)
 - es: alt.music.pink-floyd
 - alt.astrology
 - alt.fan.simpsons

Newsgroups: gruppi italiani

- gruppi italiani (circa 300)
 - it.cultura.
 - it.sport.
 - it.scienza.
 - it.annunci.
 - it.reti-civiche.
 - it.discussioni.
 - it.discussioni.universita.tesi-di-laurea
- <http://www.mailgate.org>
 - Sito italiano che da accesso a (quasi) tutte i NG italiani

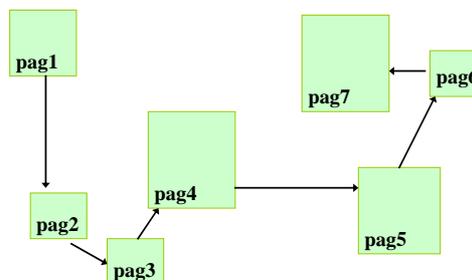
World Wide Web (WWW)



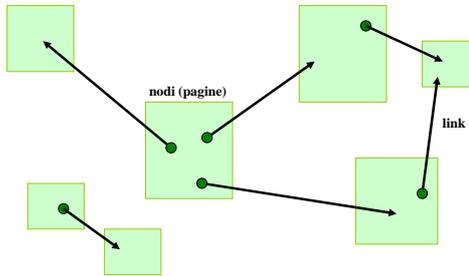
- **Ragnatela mondiale** di documenti dislocati su computer residenti in tutto il mondo
- E' il più grande serbatoio di informazioni esistente
- L'architettura software per la gestione dei dati distribuiti geograficamente è basata sulla nozione di **ipertesto**
- Si appoggia a **TCP/IP** e quindi è compatibile con ogni tipo di macchina collegata ad Internet
 - È un sistema hardware e software che utilizza Internet come ponte di collegamento

Organizzazione di un testo

lettura sequenziale



Organizzazione di un ipertesto



Iperestesi

- Possono contenere testo, immagini, suoni, programmi eseguibili
- Il linguaggio ipertestuale associa indirizzi URL a parti di documento chiamati **link**
- Selezionando un link si scarica una pagina dal server DNS identificato dal corrispondente URL

Terminologia per gli ipertesti sul Web

- **Page (pagina):** singolo "foglio" di un ipertesto
- **Browser:** programma applicativo che gestisce pagine web e permette di navigare in rete
- **Home-page:** "punto di ingresso" di un sito web
- **Hotspot, hotword:** porzione di una pagina che, se selezionata, permette di raggiungere un altro punto dell'ipertesto o una nuova risorsa

URL: indirizzi nel Web

- gli indirizzi per "trovare" i documenti nel Web sono detti **Uniform Resource Locator (URL)** e identificano in modo univoco i nodi (i documenti)
- protocollo:// indirizzo internet del nodo /pathname**
- <http://www.unito.it> è l'indirizzo della home-page dell'Università di Torino
 - <http://www.cisi.unito.it/ateneo/lettere/index.html> è l'indirizzo della home-page della Facoltà di Lettere

URL: indirizzi nel Web

- **URL** (Uniform Resource Location) è uno standard per il formato degli indirizzi delle risorse sul Web
- Specifica:
 - Come si vuole accedere alla risorsa (metodo)
 - Dove si trova la risorsa (**indirizzo server**)
 - Nome della risorsa (**nome**)
- Formato:
 - **metodo://host/nome**

Linguaggio per Pagine Web

- Gli ipertesti del Web sono scritti usando il linguaggio **Hypertext Markup Language (HTML)**
- HTML **NON** è un linguaggio di programmazione!
- HTML descrive la disposizione di tutti gli elementi presenti all'interno di un documento
- E' anche possibile creare dei meccanismi di comunicazione interattiva mediante l'uso del linguaggio HTML opportunamente combinato con altri linguaggi di programmazione (C, Java, Perl)

Linguaggio per Pagine Web

- **HTML (HyperText Markup Language)** utilizza annotazioni per descrivere come verrà visualizzato il documento sul browser di un cliente
 - Es: La prossima parola è in `neretto`
- Il browser interpreta le annotazioni traducendole in effetti grafici
 - Es: La prossima parola è in **neretto**
- Alcuni tool forniscono direttamente l'effetto desiderato senza dover usare HTML

Alcuni esempi

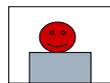
- Pagina web racchiusa nel comando `<HTML>....</HTML>`
- Intestazione: `<HEAD>...</HEAD>`
- Corpo pagina: `<BODY>...</BODY>`
- `<P>`: inizia un paragrafo
- `
`: fine linea
- `...`: testo in bold
- `<I>...</I>`: testo in italico

Sorgente pagina web

```
<HTML>
<BODY>
<b>Marco Rossi</b><br>
PhD Student <br>
Universit&agrave; di Genova<br>
<IMG SRC="marco.gif"><br>
Per scaricare la mia tesi premi qui sotto<br>
<a href="ftp://ftp.disi.unige.it/RossiM/tesi.ps">
<i>TESI</i></a>
</BODY>
</HTML>
```

Pagina visualizzata su browser

Marco Rossi
PhD Student
Università di Genova



Per scaricare la mia tesi premi qui sotto

[TESI](#)

Motori di Ricerca



- Orientarsi nel Web è difficile e non esiste un'unica sorgente di informazioni
- I **motori di ricerca** consentono di trovare informazioni sul Web
- Basati su immensi **indici** contenenti "descrizioni" delle pagine esistenti
- Aggiornamento degli indici automatico e grazie all'invio di nuove pagine dagli autori
- Permettono all'utente di immettere parole chiavi tramite FORM per cercare informazioni, siti, persone, ecc
- Forniscono un **catalogo** di indirizzi URL di migliaia di siti che l'utente può consultare

Motori di ricerca "classici"

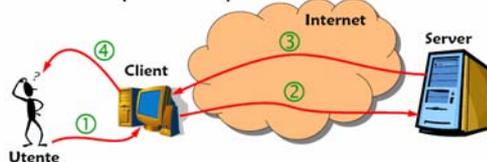
- www.google.com (www.google.it)
- www.yahoo.com (www.yahoo.it)
- www.lycos.com (www.lycos.it)
- www.webcrawler.com
- www.hotbot.com
- www.altavista.com (www.altavista.it)

Ricerche su Internet

- ❑ REGOLA 1: **siate creativi!!**
- ❑ REGOLA 2: **restringete la ricerca!!**
- ❑ usare motori di ricerca specializzati (persona, mail, home page, software, idioma, aree geografiche, argomenti ...)
- ❑ usare le regole di restrizione per le chiavi di ricerca offerte dal motore di ricerca

Il paradigma client-server

1. L'utente usa il client per esprimere le sue richieste
2. Il client si collega al server e trasmette la richiesta
3. Il server risponde al client
4. Il client presenta la risposta all'utente

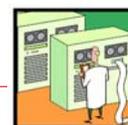


Il Client



- Si preoccupa di dialogare con l'utente
- Sfrutta tutte le possibilità fornite dal calcolatore su cui viene eseguito (audio, video, ...)
- Fornisce all'utente un'interfaccia intuitiva
- Elaborate le richieste dell'utente e le risposte dei server
 - la comunicazione avviene secondo un formato standard (protocollo)

Il Server



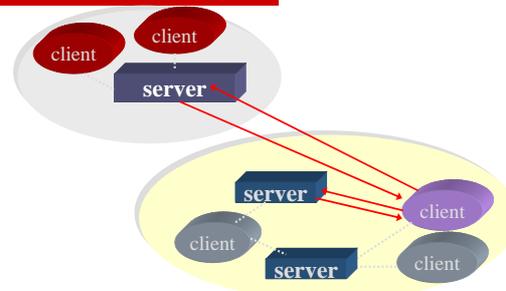
- Rende disponibili delle risorse
- Accetta richieste e risponde automaticamente
 - non bada alla provenienza della richiesta
 - il processo client può trovarsi in qualsiasi punto della rete
- Si può organizzare un insieme di server in modo che siano collegati tra loro
- Potrebbe essere eseguito dallo stesso calcolatore che esegue il processo client!

Il Server



- ❑ interpreta la richiesta ed esegue le operazioni opportune (inoltra all'elaboratore client un file contenente le informazioni desiderate, oppure interroga un database e restituisce il risultato dell'operazione)

Architettura client-server



Architettura client-server

- la comunicazione tra client e server è regolata dai protocolli:
 - **HTTP (Hyper-Text Transfer Protocol)**
per trasmettere documenti ipertestuali / ipermediali sulla rete
 - **FTP (File Transfer Protocol)**
utilizzato per trasmettere testi o file binari
 - **SMTP (Simple Mail Transfer Protocol)**
utilizzato per la posta
 - **Telnet**
utilizzato per collegarsi a elaboratori remoti

Posta elettronica



- Servizio per trasmettere dati (messaggi) tra vari utenti
- Sistemi di posta elettronica supportano:
 - Composizione di messaggi
 - Trasferimento al destinatario
 - Visualizzazione e cancellazione
 - Caselle di posta (Mailbox) per memorizzare i messaggi in arrivo
 - Mailing list (lista di indirizzi)



Invio di messaggi



- Occorre fornire un messaggio, e l'indirizzo del destinatario
- L'indirizzo ha la seguente forma
[Mailbox@locazione](#)
e viene gestito dal DNS server del nostro dominio
- Es. giorgio@disi.unige.it