

Tecnologie Web - MFN0634

Internet ed il WWW

Docenti

Marco **Botta**[marco.botta@unito.it]

Argomenti della lezione

- * Cos'è Internet
 - * Un po' di storia
 - * Persone ed organizzazioni
 - * Tecnologie
- * Il World Wide Web (WWW o Web)
 - * Architetture Client e Server
 - * URL e DNS
 - * Hypertext Transfer Protocol (HTTP)
 - * Linguaggi del Web

Argomenti della lezione

✧ **Cos'è Internet**

- ✧ Un po' di storia
- ✧ Persone ed organizzazioni
- ✧ Tecnologie

✧ **Il World Wide Web (WWW o Web)**

- ✧ Architetture Client e Server
- ✧ URL e DNS
- ✧ Hypertext Transfer Protocol (HTTP)
- ✧ Linguaggi del Web

Internet



- * Wikipedia: <http://en.wikipedia.org/wiki/Internet>
- * Una rete di computer connessi tra di loro che usa l'Internet Protocol (IP)
- * Livelli di protocolli di comunicazione: IP → TCP/UDP → HTTP/FTP/POP/SMTP/SSH...
- * Qual è la differenza tra Internet ed il Web?
 - * Il Web è una collezione di siti e pagine web distribuiti nel mondo; Internet è qualcosa di molto più grande che include altri servizi come le email, la chat, i giochi on line, etc.

Breve storia

- * Partito come evoluzione di un progetto finanziato dal Dipartimento della Difesa degli USA: ARPANET ('60-'70)
- * Primi servizi: posta elettronica, trasferimento file
- * Oggetto di interesse commerciale dai tardi anni '80
- * WWW creato da Tim Berners-Lee nel 1989-91
- * I primi browser web popolari: Netscape 1994, IE 1995
- * Amazon.com on line nel 1995; Google: Gennaio 1996
- * La pagina Hamster Dance, forse il primo “meme” diffusosi in modo importante via Web, è stata creata nel 1999
https://en.wikipedia.org/wiki/Hampster_Dance



Caratteristiche principali di Internet

- * Le sottoreti possono essere rese indipendenti dal resto
- * I computer (host) possono dinamicamente collegarsi alla rete così come possono lasciarla in qualsiasi momento
- * Realizzata su standard aperti. Ognuno può creare un proprio dispositivo da collegare ad Internet senza pagare alcuna “royalty”
- * Per la maggior parte priva di controllo centralizzato
- * Tutti la possono usare in modo semplice, usando principalmente software già disponibile

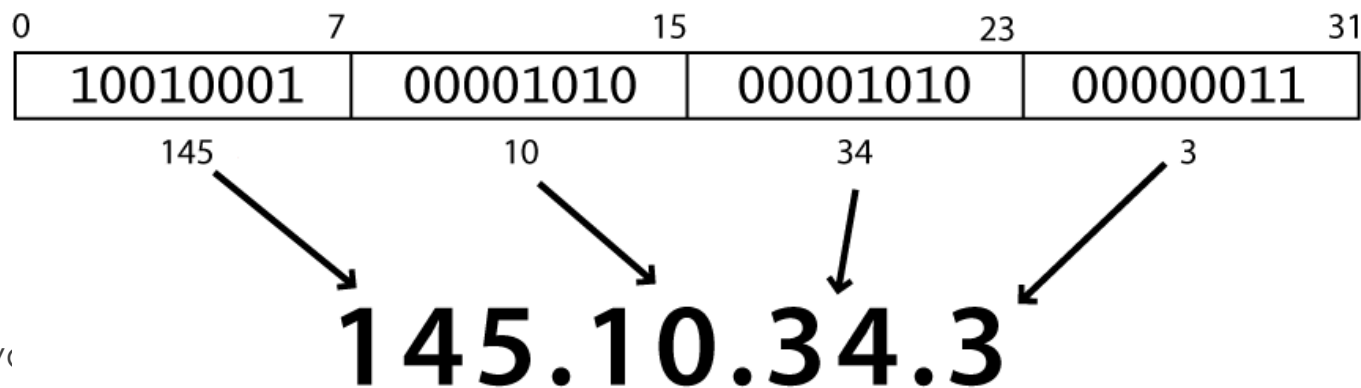
Persone e organizzazioni

- * Internet Engineering Task Force (IETF): definiscono gli standard dei protocolli internet
- * Internet Corporation for Assigned Names and Numbers (ICANN): decide e assegna i nomi dei domini di alto livello
- * World Wide Web Consortium (W3C): definisce gli standard del web



Internet Protocol (IP)

- * http://en.wikipedia.org/wiki/Internet_Protocol
- * Un semplice protocollo per fare scambiare informazioni tra due computer connessi ad Internet (indipendentemente dalla modalità fisica di connessione)
- * Ogni dispositivo ha un indirizzo IP di 32 bit scritto con quattro numeri ad 8 bit (0-255)



- * Trovate il vostro IP locale:
 - * In un finestra terminale, usate il comando: ipconfig (Windows) o ifconfig (Mac/Linux)

Transmission Control Protocol (TCP)

- * http://en.wikipedia.org/wiki/Tcp_protocol
- * Fornisce servizi di multiplexing ed affidabilità ad IP (Internet è inaffidabile e non garantisce la consegna del messaggio)
- * **multiplexing:** più programmi possono usare lo stesso indirizzo IP
 - * **port:** un numero assegnato ad ogni processo o servizio
 - * porta 80: web (port 443 per web sicuro)
 - * porta 25: email
 - * porta 22: ssh
 - * Numeri di porta più comuni:
http://en.wikipedia.org/wiki/List_of_TCP_and_UDP_port_numbers
- * Qualche programma usa il più semplice UDP invece di TCP (ad esempio: giochi, streaming audio/video, etc.)
http://en.wikipedia.org/wiki/User_Datagram_Protocol

Domande riassuntive

- ✧ Quali sono le caratteristiche fondamentali di Internet?
- ✧ Quali organizzazioni internazionali decidono gli standard per le pagine Web e per Internet?
- ✧ Cosa fornisce IP e quali servizi aggiuntivi sono forniti da TCP?
- ✧ Quali di questi indirizzi IP sono legali?
 1. www.google.com
 2. 150.135.1.150
 3. 123.456.789.10
 4. 241.259.17.127
 5. 10.0.0.1

Argomenti della lezione

- * Cos'è Internet
 - * Un po' di storia
 - * Persone ed organizzazioni
 - * Tecnologie
- * **Il World Wide Web (WWW o Web)**
 - * Architetture Client e Server
 - * URL e DNS
 - * Hypertext Transfer Protocol (HTTP)
 - * Linguaggi del Web

Server Web e Browser

- * **Server web:** processo software che rimane in “ascolto” di richieste di pagine web
 - * Apache – www.apache.org
 - * Microsoft Internet Information Server (IIS) (part of Windows) - <http://www.microsoft.com/resources/documentation/windows/xp/all/proddocs/en-us/iiisin2.mspx?mfr=true>
- * **Browser web:** richiede ai server documenti web e poi li visualizza
 - * Mozilla Firefox
 - * Microsoft Internet Explorer (IE)
 - * Apple Safari
 - * Google Chrome
 - * Opera



Domain Name System (DNS)

- * <http://en.wikipedia.org/wiki/Dns>
- * Un sistema distribuito costituito da server con diversi ruoli gerarchici che traducono indirizzi simbolici in indirizzi IP
 - * esempio: `www.cs.washington.edu` → `128.208.3.88`
 - * Esempio: `www.di.unito.it` → `192.168.1.254`
- * Molti SO mantengono una cache locale di indirizzi già risolti in un documento chiamato “hosts file”
 - * Windows: `C:\Windows\system32\drivers\etc\hosts`
 - * Mac: `/private/etc/hosts`
 - * Linux: `/etc/hosts`
- * Per usare direttamente il DNS da finestra terminale usare il comando **nslookup**

Uniform Resource Locator

- * Un identificatore della locazione di un documento pubblicato su un sito web
- * Un URL base:
 - * `http://www.aw-bc.com/info/regesstepp/index.html`

protocol host path
- * Dopo aver digitato questo URL nell'apposita barra del browser, questo dovrebbe:
 - * Contattare il server DNS per chiedere l'indirizzo IP `www.aw-bc.com`
 - * Connettersi a quell'indirizzo IP alla porta 80
 - * Richiedere al server il documento con il comando GET `/info/regesstepp/index.html`
 - * Visualizzare la pagina finale all'interno della propria finestra

URL avanzate

- * **anchor:** consentono di *saltare* direttamente ad una determinata sezione della pagina web
 - * `http://www.textpad.com/download/index.html#downloads`
 - * Prima recupera `index.html` quindi salta alla parte della pagina etichettata `downloads`
- * **port:** da specificare esplicitamente per i server web che usano una porta diversa da quella di default 80
 - * `http://www.cs.washington.edu:8080/secret/money.txt`
- * **query string:** un insieme di parametri che vengono passati ad un eseguibile
 - * `http://www.google.com/search?q=miserable+failure&start=10`
 - * Al parametro `q` è passata la stringa `"miserable+failure"`
 - * Al parametro `start` è passato il valore 10

Hypertext Transfer Protocol (HTTP)

- * http://en.wikipedia.org/wiki/Http_protocol
- * Include l'insieme di comandi compresi da un server web ed inviati da un browser
- * Alcuni comandi HTTP (Il vostro browser li invia in modo trasparente):
 - * GET filename : download
 - * POST filename : invia il contenuto di un modulo (form) web, in genere inserito manualmente da un utente
 - * PUT filename : upload
- * Simulazione dell'interazione lato browser da una finestra di comando

```
$ telnet www.cs.washington.edu 80
Trying 128.208.3.88...
Connected to 128.208.3.88 (128.208.3.88).
Escape character is '^]'.
GET /index.html
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 ...">
<html>
...
```


Codici di errore di HTTP

- * Il web server restituisce un codice di risposta che può indicare la corretta esecuzione del comando oppure il tipo di errore che è stato riscontrato. In genere questo codice è seguito da un documento HTML (anche quando si è verificato un errore!)
- * Codici di risposta comuni:

Numero	Significato
200	OK
301-303	Page has moved (permanently or temporary)
403	you are forbidden to access this page
404	page not found
500	internal server error

Lista completa: http://en.wikipedia.org/wiki/Http_error_codes



Internet Media (“Mime”) types

- * http://en.wikipedia.org/wiki/Mime_type
- * A volte, al momento di includere delle risorse in una pagina (fogli di stile, icone, oggetti multimediali), specifichiamo il tipo del dato

MIME type	file extension
text/html	.html
text/plain	.txt
image/gif	.gif
Image/jpg	.jpg
video/quicktime	.mov
application/octet-stream	.exe

- * Liste di tipi Mime:
 - * By type: http://www.w3schools.com/media/media_mimeref.asp
 - * By extension: <http://www.webmaster-toolkit.com/mime-types.shtml>

Linguaggi e tecnologie Web

- * Hypertext Markup Language (HTML): usato per creare le pagine web
- * Cascading Style Sheets (CSS): informazioni stilistiche per le pagine web
- * PHP Hypertext Processor (PHP): crea dinamicamente le pagine su un server web
- * JavaScript: permette la creazione di pagine interattive e programmabili
- * Asynchronous JavaScript and XML (Ajax): consente lo scambio di dati tra applicazioni web
- * eXtensible Markup Language (XML): meta-linguaggio per l'organizzazione dei dati
- * Structured Query Language (SQL): consente l'interazione con il db

Domande riassuntive

- * I nomi di almeno tre browser web? Perché per uno sviluppatore è importante preoccuparsi di quale browser l'utente che accede al sito sta usando?
- * Qual è la URL corretta che consente di scaricare il file `coralize.js` contenuta nella cartella `/bin/bookmark/` sul server `www.coralcdn.org.nyud.net` alla porta 8080?
- * Qual è la differenza tra le richieste POST e GET di HTTP?
- * Cosa significa l'errore HTTP 403? Perché potreste riscontrare un errore HTTP 500?