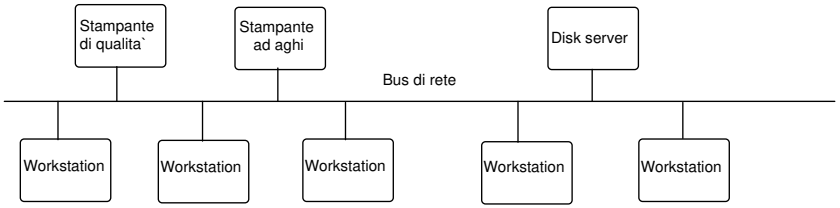
Corso di Laurea in Ingegneria Informatica ed Elettronica
Corso di Fondamenti di Informatica 2

Introduzione a Internet e al World Wide Web

Ing. Franco Zambonelli
Febbraio 2000

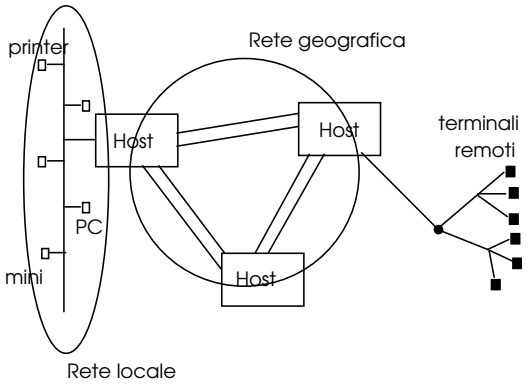
TIPOLOGIE DI RETE

Reti Locali: connettono elaboratori fisicamente vicini (nello stesso ufficio o stabilimento).



***LAN* (Local Area Network)**

Reti geografiche: collegano elaboratori medio-grandi situati anche a grande distanza.



***WAN* (Wide Area Network)**

INTERNET: Nome che designa la rete risultante dall'interconnessione mondiale di tutte le reti

INTERNET

Rete mondiale basata sul protocollo TCP/IP (definito dal DARPA a metá degli anni '70). **Creata** storicamente per esigenze militari e **Un tempo** usata solo per scopi specialistici, fra università e centri di ricerca **OGGI** sempre più **il principale strumento** per

- scambiarsi rapidamente informazioni
- trovare informazioni di ogni tipo *in tempo reale*
- sviluppare applicazioni distribuite
- ....

I computer di Internet sono associati ad:
un nome (o più) simbolico

`deis.unibo.it, linus.apple.com, www.nasa.gov`

il nome è diviso in domini e sottodomini:
(struttura gerarchica)

in USA domini associati al tipo di organizzazione: `gov, edu, com ....`

nel resto del mondo - tipicamente - domini associati alla nazione e sottodomini:

`it, fr, uk....`
`unibo, unimi.....`

INDIRIZZO INTERNET

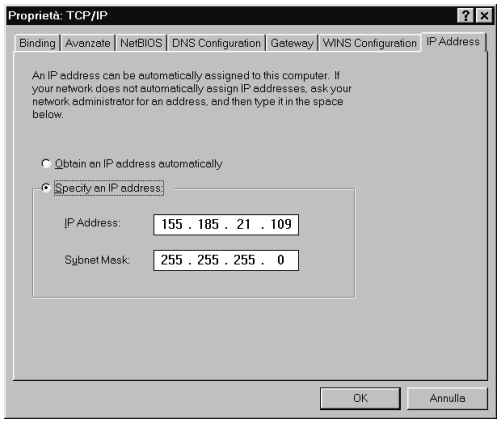
A ogni computer è associato un indirizzo Internet (IP number) unico

`137.204.57.34`
`192.31.254.130`

4 campi di 1 byte
*255*255*255*255 (4 miliardi circa)*
di computer possono essere collegati a Internet....

L'ultimo numero identifica di solito un computer all'interno di una sottorete, gli altri la sottorete....

Win95: Pannello Controllo –> Rete –> Proprietà TCP/IP

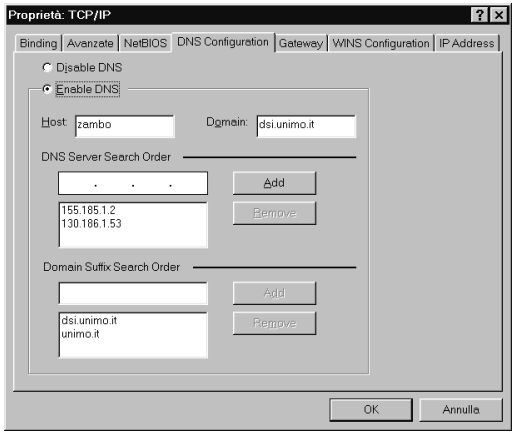


UNIX: diversi file di configurazione: `/etc/netconfig`

DNS (domain name system)

permette di traslare il nome simbolico nell'indirizzo IP. Nodi specializzati (DNS server) risolvono la associazione tra nome simbolico e indirizzo IP. Il tutto è trasparente all'utente

In Win 95:



In UNIX: comando `nslookup`

Se non è settato il domain name system si può creare una associazione tra nomi simbolici e indirizzi IP nel file:

`C:/Windows/hosts (per Windows)`
`/etc/hosts (per UNIX)`

INTERNET… A CASA PROPRIA?

Per essere "in rete" occorre che "il filo di Internet" arrivi sul proprio computer

Ma Internet collega direttamente solo reti geografiche di università, centri di ricerca, etc

E il proprio computer di casa?

IDEA:

- a casa arriva già le rete *telefonica*
- **usare IL FILO DEL TELEFONO** per arrivare da casa a una rete collegato a Internet → *MODEM*

Ma ... dove telefonare??

Perché qualcuno dovrebbe lasciarmi usare la sua rete per collegarmi a Internet??

PROVIDER ("fornitore")

- qualcuno che ha una rete collegata permanentemente a Internet (sempre acceso)
- e decide di **consentire, a pagamento o gratis (per motivi pubblicitari e con qualità di servizio più scarsa) di accedere a Internet tramite il suo collegamento** telefonando presso di lui

→ ABBONAMENTO AL PROVIDER

Il nostro computer, collegandosi a un computer del provider, sito, diventa parte della rete del provider e, come tale, è collegato ad Internet. Indirizzo IP assegnato dinamicamente all'inizio del collegamente dal provider stesso.

PROTOCOLLI DI ACCESSO A INTERNET

Protocollo = schema convenzionale di comunicazione tra entitá

I protocolli di Internet sono realizzati specializzando il protocollo di base TCP/IP

- telnet, rlogin (terminale remoto)
- E-Mail (posta elettronica)
- ftp (trasferimento file)
- http (trasferimento e navigazione in ipertesti)

Protocolli di tipo **CLIENT-SERVER**:

uno strumento di tipo "*cliente*" che richiede (secondo un preciso protocollo) servizi a un programma "*servitore*" su una macchina, tipicamente remota

- *Esempio: HTTP (protocollo), Web Browser (cliente), Web Server (servitore)*

TELNET

terminale remoto: permette di usare il terminale locale come se fosse il terminale di una macchina remota. Server disponibili solo per sistemi UNIX/LINUX. Crea una connessione di rete con la macchina remota basata su terminale a carattere

In Win95: Client Telnet per Win95

In UNIX: comando: `telnet nomemacchina`

RLOGIN è simile a telnet, ma:

- telnet effettua il collegamento a terminale remoto ed è generale (non solo UNIX)
- rlogin tenta il login remoto e si fa solo tra macchine UNIX "fidate"

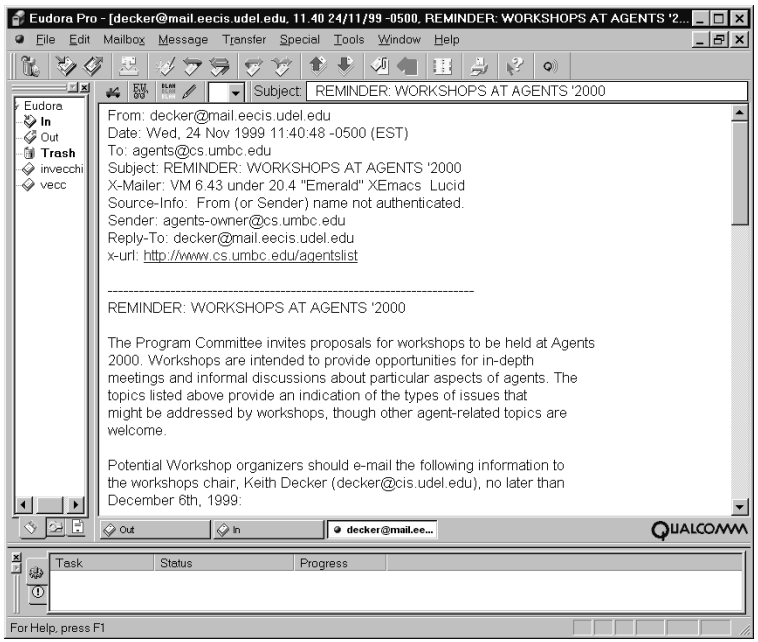
E-MAIL

Un utente su una macchina UNIX ha associato uno username. In generale si identifica l'utente attraverso la sintassi:

```
username@nomemacchina
```

Questo rappresenta anche l'indirizzo di posta elettronica (i mail server sono tipicamente sistemi UNIX). E' possibile associare alias e liste di interesse agli indirizzi E-mail.

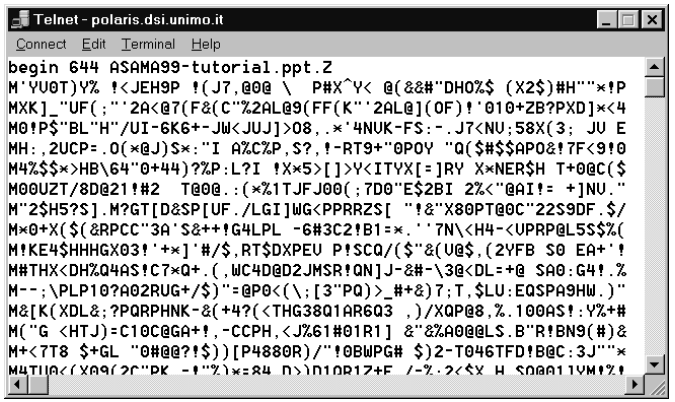
Es: EUDORA, E-mail client per Windows



E-MAIL

Tramite E-mail, utenti su macchine sia locali che remote possono scambiarsi messaggi testuali, cioè solo codice ASCII

Esistono metodi standard di codifica per file di tipo binario (uudecode, uuencode): trasformano codice binario in codice ascii, e sono sfruttati dagli strumenti di “attachment”

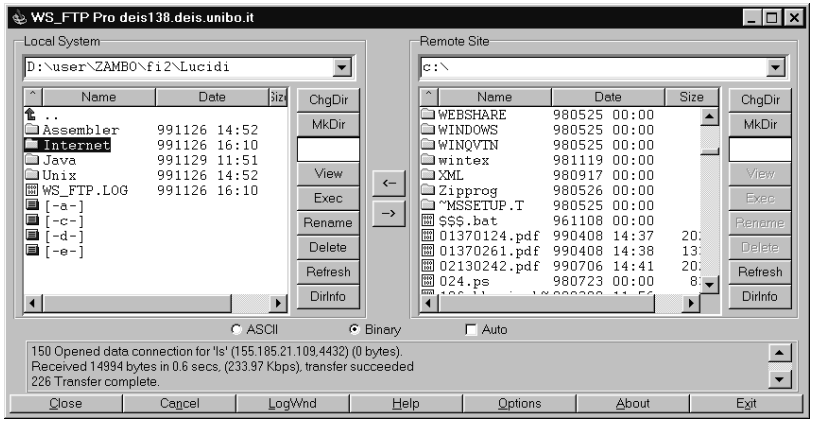


FTP

Protocolli per il trasferimento file

Ci si collega a una macchina remota su cui si deve avere account (e su cui deve eseguire un ftp server!)

In Win95: clienti a interfaccia grafica:



In UNIX: clienti a Linea di Comando

```
prompt> ftp macchina remota
username:
passwd:
ftp>
```

si hanno quindi a disposizione un set limitato di comandi shell per esaminare il file system

COMANDI FTP IN UNIX

`ls, dir, cd` - solito significato

`get` - preleva un file dalla macchine remota

`put` - deposita un file

`mget` - preleva più file con caratteri wild card

`mput` - deposita .....

`bin, ascii` - il tipo del file da trasferire: il trasferimento ascii è piú efficiente, sfrutta le ridondanze del codice per limitare i bit da trasferire

`prompt` - chiede conferma o meno

`hash` - visualizza il numero di k trasferiti

FTP ANONIMO

per collegarsi a una macchina su cui non si ha account

banche dati pubbliche

```
login : ftp o anonymous
passwd: E-mail address
```

viene data una visione parziale del file system

Formati dei file:

- *.tar*
- *.Z*
- *.z*
- *.gz*

PROTOCOLLI E STRUMENTI AVANZATI

I precedenti strumenti per l'accesso a Internet non sono dotati del requisito di

trasparenza

si ha coscienza della distribuzione dell'informazione (i.e., nome host)

Strumenti più avanzati permettono di accedere alla rete in maniera trasparente, come se tutto fosse locale. Il collegamento remoto è (o puó essere) nascosto all'utente

- *USENET news*

- *World Wibe Web*

USENET NEWS

"bacheche elettroniche" per lo scambio di:

- informazioni
- opinioni
- domande

o più genericamente:

spazio di discussione

newsgroup:
identifica l'argomento

Esempio:

- `alt.fan.nirvana`
- `comp.unix.networking`
- `soc.culture.italian`

USENET NEWS

macchine che ricevono news su base periodica attraverso protocolli di diffusione punto-punto sulla rete

news server

che si passano le news su base periodica
In questo modo le news diffondono attraverso Internet su scala mondiale

Le news vengono spedite da un utente che vi ha accesso, ovunque esso sia, attraverso strumenti denominati

newsreader (client)

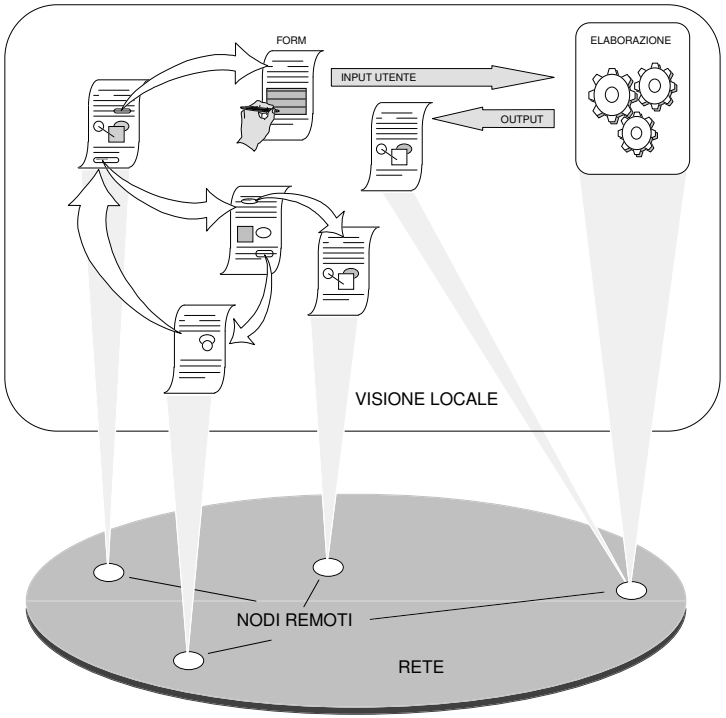
Il newsreader (oggi tipicamente inglobato nel browser o nel mailer) serve ovviamente anche per leggerle

Tutto è trasparente all'allocazione fisica su Internet anche se l'indirizzo di provenienza delle news è presente

NOTA: il formato delle news è quello dei mail! Si può quindi usare un cliente E-mail per leggere le news!

WWW (Mosaic, Netscape, Explorer)

strutturazione ipertestuale delle informazioni (trasparenza della allocazione delle informazioni) e uso di interfacce grafiche (semplicità di utilizzo)



Facendo “click”su una parola/immagine si “espande” una parte del documento che interessa. Non e’ necessario “percepire” il fatto che il documento sia sul file system locale o su una macchina remota.

World Wide Web (WWW)

CERN (1989)
Progetto di integrazione in forma ipertestuale delle risorse esistenti in INTERNET

Scopi

- Trasparenza accesso e allocazione tramite nevigazione su ipertesti
- Presentazione multimediale
- Interfaccia unica per protocolli diversi (integrazione con gli altri protocolli)
- Modificabilità e condivisione delle informazioni

Ampia scelta di interfacce testuali e grafiche
Possibilità di estensioni sperimentali del sistema

Componenti

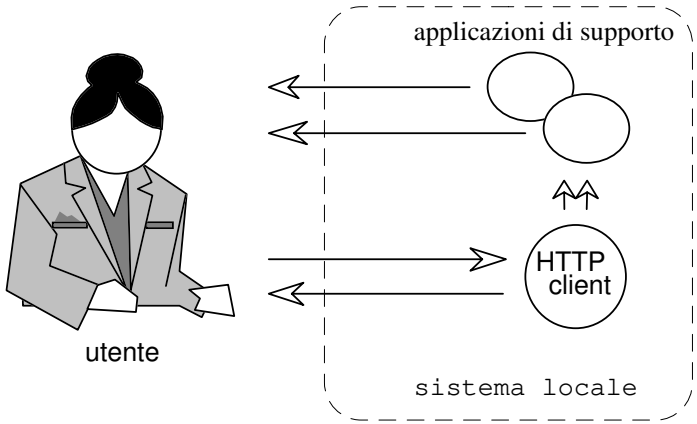
- Browser (presentazione e gestione richieste)
- Server (accesso e invio informazioni)
- Helper applications (particolari presentazioni)
- Applicazioni CGI (esecuzione remota)
- Applet (esecuzione locale)

Specifiche standard

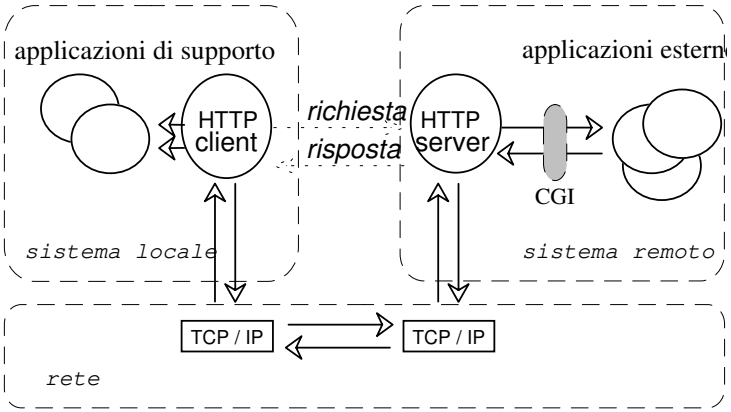
- Sistema di indirizzamento universale URI e URL (Uniform Resource Identifier/Location)
- Protocollo HTTP (HyperText Transfer Protocol)
- Linguaggio HTML (HyperText Markup Language)
- Interfaccia CGI (Common Gateway Interface)
- Linguaggio Java per Applet

SISTEMA **WWW**

Cliente e sua interazione



Il Cliente HTTP usa un modo cliente/servitore nei confronti di un server per volta e può anche interagire con risorse locali



URL **U**niform **R**esource **L**ocators

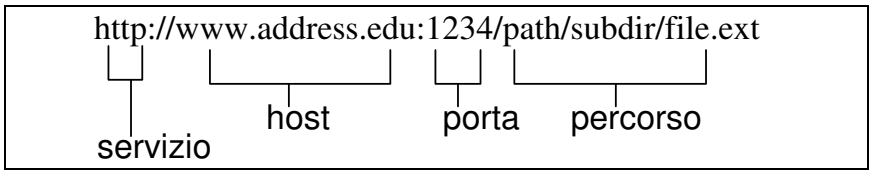
nomi unici per le risorse del sistema
specificati dal cliente per determinare il servitore

o **Uniform Resource Locators (URL):**

- nodo contenente la risorsa (*documento o dati*)
- protocollo di accesso alla risorsa (*e.g. http, gopher*)
- numero di porta TCP (*porta di default del servizio*)
- localizzazione della risorsa nel server.

```
<protocollo>[://<host>][:<porta>][<percorso>]
```

Sono riconosciuti i servizi internet e relativi protocolli =>
http, gopher, ftp, wais, telnet, news, nntp, e mail



Uso di default per localizzare risorse

Un **URL** può anche determinare un insieme di risorse:
ad esempio versioni multilingue tra cui scegliere

HTTP HyperText Transfer Protocol

protocollo di interfaccia tra cliente e servitore
Uso di TCP e di connessione (porta **80** default)

Caratteristiche HTTP:

- request/response
- one-shot connection
- stateless

Request/response: richiesta e ricezione di dati.
One-shot connection: la connessione TCP è mantenuta solo per il tempo necessario a trasmettere i dati
Stateless: non mantiene nessuna informazione tra una richiesta e la successiva

in genere:
- **richiesta** del cliente con **informazioni** per il **server**
- **risposta** con informazioni dal server

il cliente può determinare una forma di scelta (**negoziazione**) sulle informazioni ed i servizi

```
HTTP-message = Simple-Request      ;HTTP/0.9
   / Simple-Response
   / Full-Request                  ;HTTP/1.0
   / Full-Response
```

NON c'e' stato del server

Il Concetto di URL

L'indirizzo internet di un computer lo specifica univocamente nel mondo

Il nome di un file lo specifica univocamente nel suo file system

QUINDI:

Componendo indirizzo Internet e nome file si identifica univocamente un file a livello mondiale:

NOME COMPUTER/PERCORSO_E_NOME_DEL_FILE

pcfz.unimo.it/CorsoRE/Internet.doc

Ovviamente preceduto dal nome del protocollo che si deve (o si vuole) usare per accedere il file. Di default http.

http://pcfz.unimo.it/CorsoRE/Internet.doc

ftp://pcll.unibo.it/CorsoMO/Windows.doc

Windows98 integra questa visione del file system a livello globale in un ambiente in cui "navigare" nel proprio file system o "navigare" su Internet è la stessa cosa.

NON SOLO: accesso a risorse sul proprio computer (file identificati da un pathname)
MA ANCHE: accesso a risorse remote in Internet (file identificati da una URL)

Come per i file, è possibile collegare al desktop URL.

HTML HyperText Markup Language

HTML è un linguaggio di specifica delle informazioni che deriva da SGML (Standard Generalized Markup Language). E' un **markup language** (TeX, RTF).

I linguaggi markup usano dei **tag** definiti **funzionalmente** per caratterizzare graficamente il testo incluso.

tag HTML

testo di tipo header 1: `<H1>testo</H1>`
testo in grassetto: `<STRONG>testo</STRONG>` oppure `<B>testo</B>`
Visualizzazione dipendente dal browser

link: `<A HREF = "destinazione"> descrizione </A>`

immagini: `<IMG SRC = "myimage.gif">`

applet Java:
```
<APPLET CODE="Hello.class" WIDTH=100 HEIGHT=80>
```

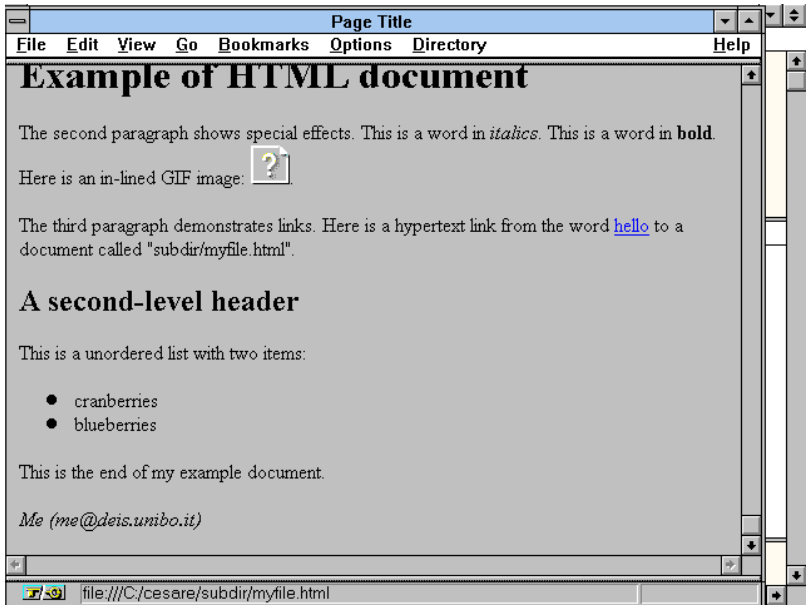
HTML **molto semplice** per non complicare il cliente

versione	browser	proprietà
1.0	storico	header, liste, enfasi
2.0	Mosaic	Inline Image, form
2.1	Netscape/Microsoft	tabelle, allineamento
3.2	Netscape/Microsoft	frame, ...
4.0	Netscape/Microsoft	Stili, JavaScript

Esempio pagina HTML (codice)

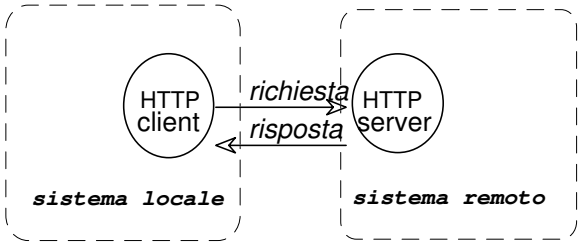
```
<HEAD>
<TITLE>Page Title</TITLE>
</HEAD>

<BODY>
<H1>Example of HTML document</H1>

The second paragraph shows special effects.
This is a word in <I>italics</I>.  This is a
word in <B>bold</B>. Here is an in-lined GIF
image: <IMG SRC = "myimage.gif">. <P>

The third paragraph demonstrates links.  Here
is a hypertext link from the word <A HREF =
"subdir/myfile.html">hello</A>  to  a  document
called "subdir/myfile.html". <P>

<H2>A second-level header</H2>

This is a unordered list with two items: <P>
<UL>
<LI> cranberries
<LI> blueberries
</UL>
This is the end of my example document. <P>
<ADDRESS> Me (me@deis.unibo.it) </ADDRESS>
</BODY>
```

Esempio pagina HTML (visualizzazione)



Per approfondire la conoscenza di HTML consultare il sito Web del corso.

Programmazione client/server in WWW

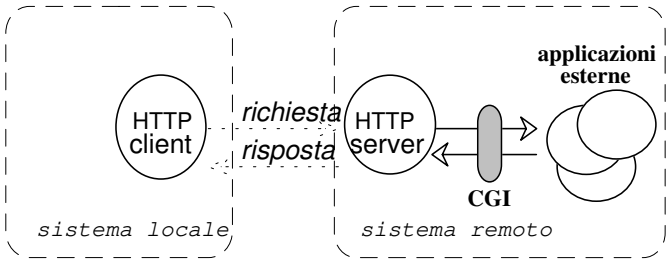


Possibilità di avere ***risposta*** con informazioni dinamiche

Che tipo di elaborazione delle informazioni e **dove** viene eseguita

richiesta	Risposta	tipo di elaborazione
Documento HTML	Statica (la pagina è un file, non modificabile)	semplice trasferimento file dal server
CGI	Dinamica	Codice sul server (qualunque tipo di elaborazione) ed eseguito sul server
Java applet	Statica	codice fonito dal server ed eseguito sul client

Common Gateway Interface (CGI)



CGI è uno **standard** per interfacciare un server WWW con applicazioni esterne (residenti sulla macchina server)

CGI fornisce all'utente la capacità di eseguire una applicazione sulla macchina server remota

richiesta	*risposta*	*tipo di elaborazione*
CGI	dinamica	qualunque, **sul nodo server**

La risposta ottenuta dal server è "dinamica", in quanto risultante dalla esecuzione di un programma sul server.

Programmazione CGI

Una applicazione CGI permette agli utenti di eseguire una applicazione sul nodo dove risiede il server www.

Applicazioni CGI possono essere scritte in: C/C++, PERL, TCL, UNIX shell, Visual Basic, etc…

Normale attivazione di una CGI:

- Si invia al server un messaggio (ciò avviene tipicamente riempiendo moduli, i cui dati serviranno come input al programma)
- Il messaggio scatena l'esecuzione del programma CGI
- Il programma CGI genera come output una pagina HTML in cui inserisce i risultati della sua esecuzione

Interfaccia tra **server www** e applicazione **CGI**:

- **variabili di ambiente** (non le vediamo in dettaglio)
- linea di comando
- **standard input**: il server ridirige sull'ingresso della applicazione CGI i dati ricevuti dal client (browser). Il numero di byte è nella variabile d'ambiente CONTENT_LENGTH, il tipo dei dati MIME nella CONTENT_TYPE.
- **standard output**: l'applicazione CGI manda il risultato dell'elaborazione sullo standard output verso il server, che a sua volta prepara i dati e li spedisce al client.

Client HTTP → server HTTP → CGI

Tipicamente, uso di **form**

```
<TITLE>Esempio di Form </TITLE>
<H1>Esempio di Form </H1>

<FORM METHOD="POST" ACTION="http://www-
lia.deis.unibo.it/cgi-bin/post-query">
<--! Indica di inviare come input al programma
indicato da quell'URL i dati di input inseriti
dall'utente -->

Inserisci del testo: <INPUT NAME="entry">

e premi per invio: <INPUT TYPE="submit"
                    VALUE="Invio">
</FORM>
```

Visualizzazione **form**

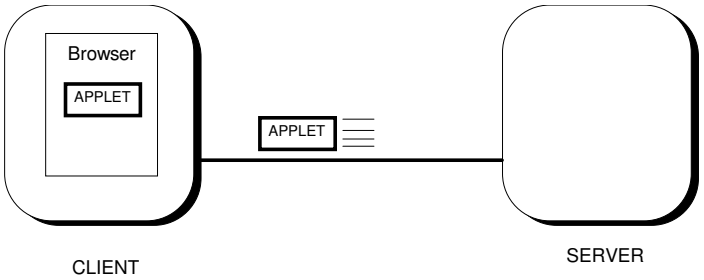
Applicazione CGI

Esempio: codice di CGI con generazione della pagina di risposta (applicazione C)

```
#include <stdio.h>
..............

main(int argc, char *argv[]) {
  int cl;

/* generazione di un documento HTML come output
*/

  printf("Content-type: text/html");

  cl = atoi(getenv("CONTENT_LENGTH"));

/* scopre quanto sono lunghi i dati di input */

  for(x=0;cl && (!feof(stdin));x++) {
/*   ...
     elaborazione dell'input (stdin)
     ...   */

  }

  printf("<H1>Query Results</H1>");
  printf("You submitted ...");

  for(x=0; x <= m; x++)
    printf("...........", ... , ....);
}
```

Applet

- una applet (piccola applicazione) viene eseguita all'interno di un browser WWW come Explorer o Netscape Navigator
- tramite il tag **<applet>** è possibile dire al browser di scaricare il codice (il file .class) e di eseguirlo
- al contrario del CGI, la computazione viene eseguita sul client



Applet Java

Stessa semantica del recupero di una immagine:
- vedere una applet ⇨ eseguirla

Classe

`java.applet.Applet`

Metodi

- per interagire con l'ambiente in cui viene eseguita
- per caricare e gestire immagini e suoni
- per ottenere parametri contenuti nel file HTML
- per ottenere informazioni sul proprio stato (attiva/disattiva, dimensioni, ...)

Sicurezza

- l'applet viene eseguita in una "sandbox" per cui può accedere solo alle risorse dell'host da cui viene prelevata, in particolare:
  - ⇒ connessioni di rete: solo con host di provenienza
  - ⇒ non può scrivere sul file system locale al browser dove viene eseguita
- nel jkd 1.2 maggiore flessibilità
  - ⇒ possibilità di specificare quali operazioni può effettuare un'applet a seconda della provenienza e/o firma digitale

Esempio di applet

File di codice HelloWorld.java

```
import java.applet.Applet;
import java.awt.Graphics;
import java.awt.Font;
public class HelloWorld extends Applet
{  public void paint(Graphics g)
  {   g.setFont(new Font("TimesRoman", Font.PLAIN, 30));
      g.drawString("Hello", 50, 30);
      g.setFont(new Font("TimesRoman", Font.PLAIN, 40));
      g.drawString("World!", 30, 70);
} }
```

File HelloWorld.html per il browser

```
<HTML> <HEAD>
<TITLE> Un semplice programma </TITLE>
</HEAD> <BODY>
Output del programma:<BR>
<APPLET CODE="HelloWorld.class" WIDTH=250 HEIGHT=100></APPLET>
</BODY> </HTML>
```

Senza conoscere la programmazione grafica e ad eventi si fa molto poco….

Java Script

In alternativa alle Applet, si possono rendere le pagine HTML dinamiche grazie all'uso di JavaScript.

Uno speciale TAG HTML può contenere dei piccoli pezzi di programma scritti in un linguaggio di programmazione che è una versione semplificata di Java.

I programmi specificano quando devono essere eseguiti. In particolare, specificano gli "eventi" che ne causano l'esecuzione: il click del mouse in un certo punto, il passaggio del mouse su una certa parola.

L'esecuzione può causare azioni diverse, quali l'apparizione di finestre o il cambiamento dinamico dell'aspetto della pagine.

Esempio:

```
<a HREF="provejs.html"
onClick="
var n = 0;
alert('apro 2 finestre');
n = n + 1;
alert('Apro finestra numero ' + n);
window.open('http://www.dsi.unimo.it');
n = n + 1;
alert('Apro finestra numero ' + n);
window.open('http://www.dsi.unimo.it')
">
DSI</a><br>
```

Guardare il sito Web del corso per approfondire le conoscenze su JavaScript