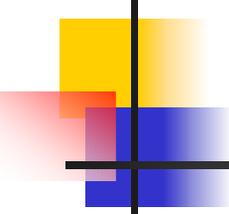


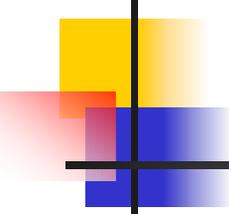
L'informazione nel World Wide Web

- L'ormai capillare diffusione di Internet (ed in particolare del World Wide Web) sta rivoluzionando molte attività.
- L'innovazione chiave della "rete delle reti" è che consente virtualmente a chiunque di pubblicare delle informazioni rendendole immediatamente e facilmente accessibili ad una vasta comunità di utenti.
- La ricchezza delle informazioni reperibili in rete è data anche dal fatto che non sono limitate a contenuti testuali, ma possono essere multimediali (integrando testo, immagini, suoni, animazioni/filmati).
- Tuttavia la libertà di comunicazione suddetta comporta anche lo svantaggio di rendere difficile il reperimento delle informazioni cercate, data la mole di dati presente in Internet e, soprattutto, la mancanza di una struttura uniforme.
- L'espressione "autostrada dell'informazione" (information highway) viene spesso usata come sinonimo di Internet e dei servizi offerti dalla rete. Negli Stati Uniti è in corso un dibattito per stabilire se sia il caso di imporre un controllo sul flusso di informazioni da parte di enti governativi (come avviene per la gestione del traffico su una comune autostrada).



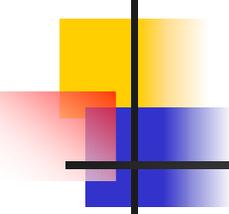
Le origini di Internet

- 1957: gli Stati Uniti istituiscono la Advanced Research Projects Agency (ARPA - Agenzia per i Progetti di Ricerca Avanzata). Si tratta di un dipartimento del Ministero della Difesa con l'incarico di preservare il primato americano nella scienza e nella tecnologia militare.
- 1969: l'ARPA sviluppa ARPANET (il precursore di Internet): si tratta di una rete che collega i quattro computer di quattro università del sud-ovest degli Stati Uniti (UCLA, SRI, UCSB e l'Università dello Utah).
- 1970: alla rete si aggiungono il MIT, Harvard, BBN, e la Systems Development Corp. (SDC) di Santa Monica, California.
- 1971: si aggiungono anche i Lincoln Labs del MIT, Carnegie-Mellon e molti altri centri di ricerca ed università. La rete inizia ad espandersi velocemente.



Gli anni '70 – '80

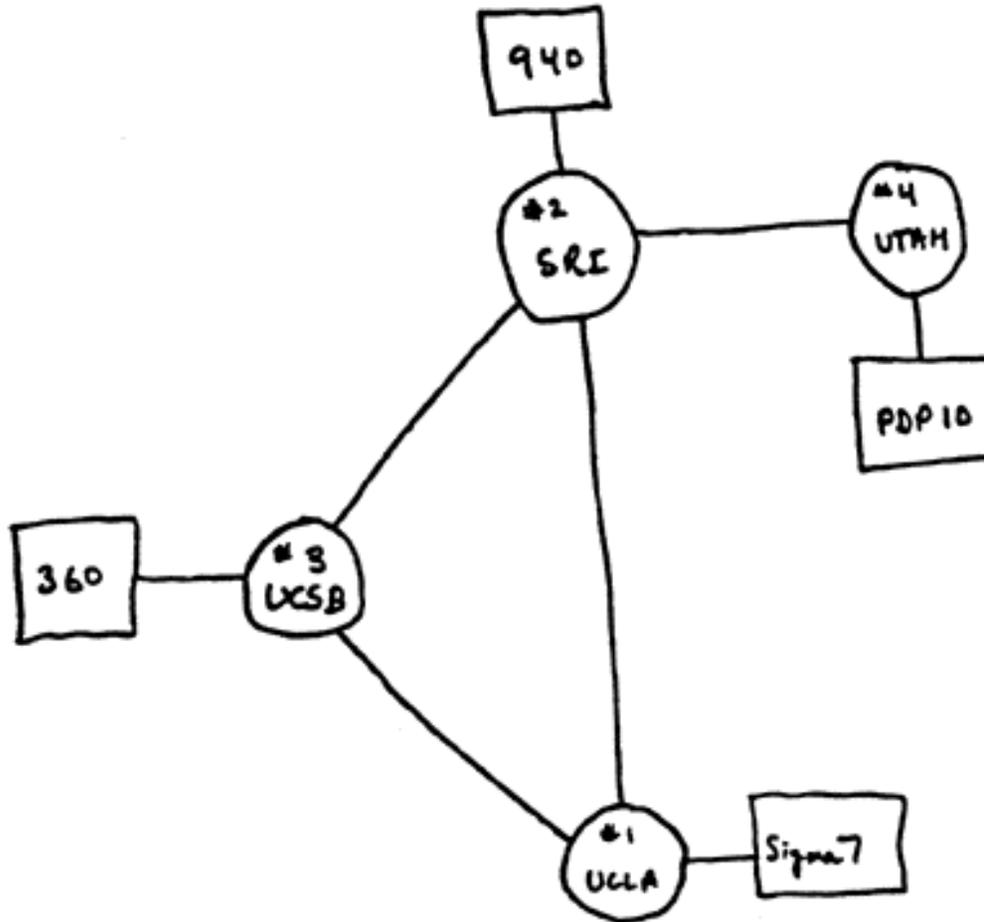
- Grazie all'introduzione del protocollo (UUCP) inizia a diffondersi l'uso dei gruppi di discussione (newsgroup), mettendo in contatto sistemi Unix presenti in tutto il mondo.
- Con lo sviluppo della suite di protocolli TCP/IP (Transmission Control Protocol/Internet Protocol) Internet si evolve; vengono introdotti i servizi di e-mail, FTP e telnet che permettono l'utilizzo della rete anche ai non addetti ai lavori.
- Vista la crescita continua di Internet, si sente la necessità di creare degli strumenti per tenere traccia delle risorse disponibili. Nascono Archie (Peter Deutsch, McGill University in Montreal, Canada) e WAIS (Wide Area Information Service).



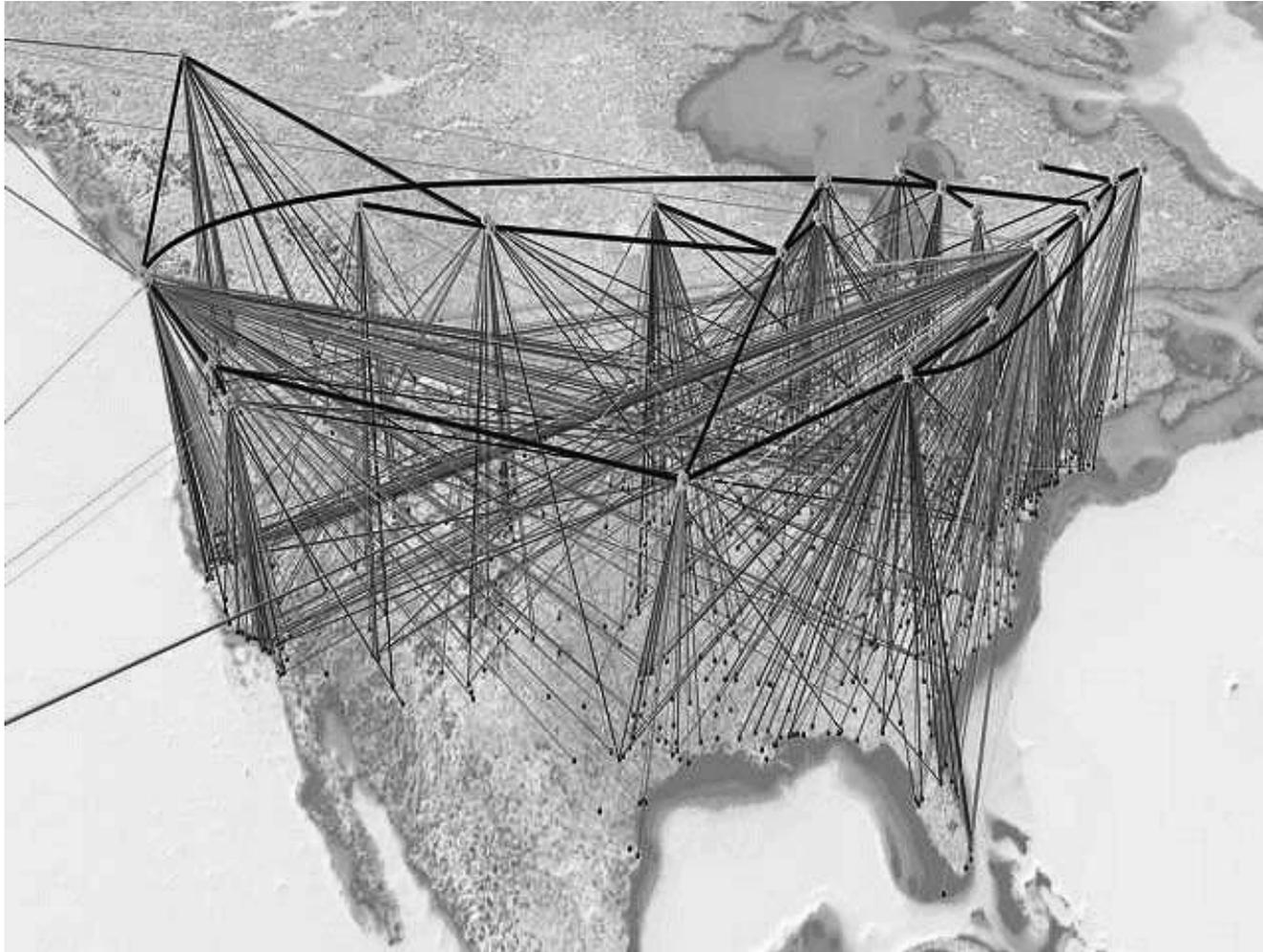
Gli anni '90

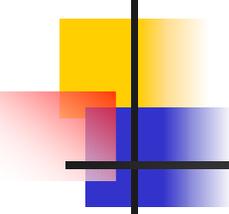
- 1991: presso l'Università del Minnesota viene creata la prima interfaccia "semplice" per Internet che trova una larga diffusione: Gopher (basato su un'architettura client-server).
- La popolarità di Gopher cresce a dismisura quando l'Università del Nevada (Reno) sviluppa un indice consultabile di siti gopher chiamato VERONICA: Very Easy Rodent-Oriented Netwide Index to Computerised Archives (indice di rete per roditori di facile uso degli archivi informatici).
- 1989: Tim Berners-Lee ed altri ricercatori del CERN propongono un nuovo protocollo (HTTP: HTML Transfer Protocol) per lo scambio di documenti: nasce il World Wide Web.
- 1993: Marc Andreessen (NCSA) sviluppa con la sua équipe Mosaic (il primo browser grafico).
- 1995: spariscono tutte le limitazioni all'uso commerciale di Internet.
- Nel 1993 i siti Web erano solo 130. Al giorno d'oggi ce ne sono milioni.

ARPANET (1969)



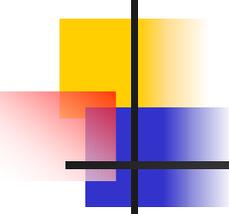
Internet oggi negli USA





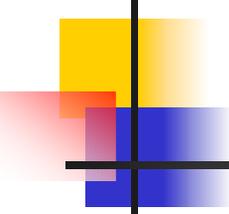
Indirizzi IP

- Per comunicare fra loro i computer (detti anche host) utilizzano dei protocolli (i.e., insiemi di regole). Tra questi il protocollo IP (Internet Protocol) stabilisce il formato degli indirizzi IP.
- Un indirizzo IP è formato da quattro numeri (ognuno può assumere un valore da 0 a 255) separati da punti.
- La funzione degli indirizzi IP è quella di identificare in modo univoco una macchina (host) all'interno di una rete (network).
- Quindi ogni host ha un indirizzo IP unico all'interno della rete in cui si trova, in modo da permettere di creare un **canale virtuale** di comunicazione fra due host.
- Esistono varie classi (tipologie) di indirizzi IP, a seconda del tipo di rete a cui l'host è connesso.



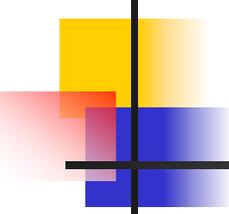
Classi di indirizzi IP

- Nel seguito 'n' indica un network bit, 'h' un host bit, 'm' un multicast bit e 'r' un bit riservato:
- Classi di tipo A - 0nnnnnnn.hhhhhhhh.hhhhhhhh.hhhhhhhh (primo bit: 0; 7 network bit; 24 host bit; primo byte: 0 – 127). Esistono quindi 126 classi di tipo A (le classi 0 e 127 sono riservate). Ci possono essere 16.777.214 host in ogni classe A .
- Classi di tipo B - 10nnnnnn.nnnnnnnn.hhhhhhhh.hhhhhhhh (primi 2 bit: 10; 14 network bit; 16 host bit; primo byte: 128 – 191). Esistono quindi 16.384 classi di tipo B ed ognuna di esse può contenere 65.532 host.
- Classi di tipo C - 110nnnnn.nnnnnnnn.nnnnnnnn.hhhhhhhh (primi 3 bit: 110; 21 network bit; 8 host bit; primo byte: 192 – 223). Esistono quindi 2.097.152 classi di tipo C ed ognuna di esse può contenere 254 host.
- Classi di tipo D - 1110mmmm.mmmmmmmm.mmmmmmmm.mmmmmmmm (primi 4 bit: 1110; 28 bit per il multicast address, i.e., trasmissione di dati a più host contemporaneamente; primo byte: 224 – 247).
- Classi di tipo E - 1111rrrr.rrrrrrrr.rrrrrrrr.rrrrrrrr (primi 4 bit: 1111; 28 bit per indirizzi riservati; primo byte: 248 – 255).



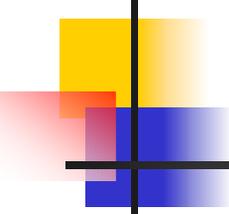
Classi di indirizzi IP - Esempi

- L'indirizzo IP del web server dell'Università di Udine è 158.110.1.44.
- Analizzando il valore del primo numero (158) si evince che l'host appartiene ad una rete con 65.532 host (indirizzo di classe B) in quanto $128 < 158 < 191$.
- Ogni host può "comunicare con se stesso" utilizzando l'indirizzo IP 127.0.0.1 (indirizzo di classe A). Infatti quest'ultimo rappresenta il local host (macchina locale).
- Tutti gli host che appartengono alla stessa rete (network) hanno un indirizzo IP tale che la parte di quest'ultimo che individua la rete stessa è invariante.
- Ad esempio, tutti i computer dell'università di Udine hanno indirizzi IP della forma 158.110.x.y, ovvero, variano soltanto gli ultimi due numeri (infatti i primi due individuano la rete di tipo B).



Nomi simbolici

- Nonostante gli indirizzi IP siano un modo efficace per individuare in modo univoco un host sulla rete, le persone li trovano scomodi in quanto sono dei dati numerici, difficili da ricordare.
- Quindi fin dagli inizi di Internet ad ogni indirizzo IP si associava un nome simbolico (il nome dell'host).
- In tal modo le persone potevano utilizzare il nome simbolico (facile da ricordare) per riferirsi ad un host; il software di rete poi si occupava di tradurre tale nome nell'indirizzo IP corrispondente in maniera da stabilire la connessione.
- Inizialmente le associazioni tra nomi simbolici ed indirizzi IP erano mantenute in un file di testo (una per riga). Quest'ultimo era memorizzato e mantenuto aggiornato presso il NIC (Network Information Center); ogni host doveva periodicamente scaricarlo per essere al passo con i cambiamenti di Internet (e.g., comparsa di nuovi host).

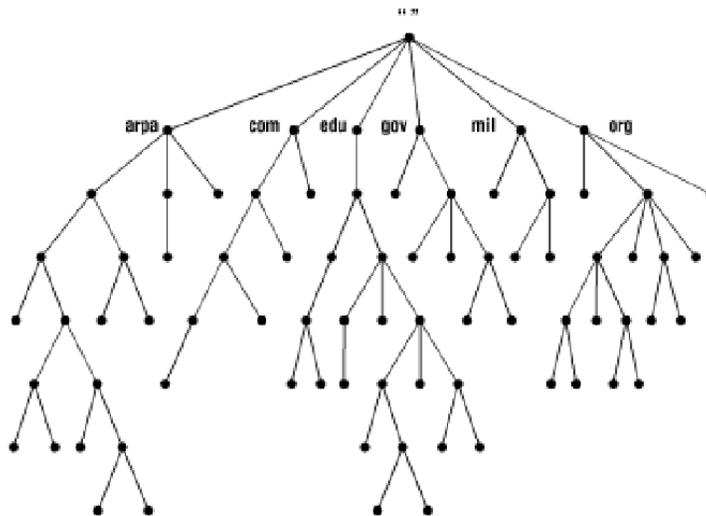


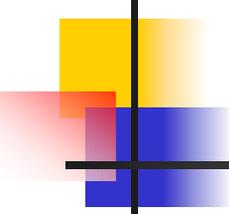
Nomi simbolici - DNS

- La soluzione **centralizzata** del NIC, con il rapido evolversi di Internet, cominciava a mostrare dei limiti:
 - Gli aggiornamenti si rendevano via via più frequenti a causa della forte crescita della rete (aggiunta di nuovi host).
 - La necessità di scaricare frequentemente il file aggiornato dal server INTERNIC creava dei problemi di traffico di rete.
 - Diventavano frequenti i casi in cui due host distinti (quindi con indirizzi IP diversi) si registravano presso l'INTERNIC con lo stesso nome simbolico.
- Per superare tali problemi si passò ad un database **distribuito** noto come DNS (Domain Name System). Tale soluzione permise un controllo **locale** sulle singole parti della mappatura nomi-indirizzi, mantenendo accessibile nel contempo l'informazione a tutta la rete.

DNS

- L'architettura su cui si basa il DNS è di tipo client-server:
 - I name server contengono parte delle informazioni sulle associazioni nomi-indirizzi.
 - I client (o resolver) formulano delle query ai name server per ottenere l'indirizzo IP corrispondente ad un dato nome.
- La struttura del database del DNS è quella di un albero rovesciato (con la radice in alto e le foglie in basso):



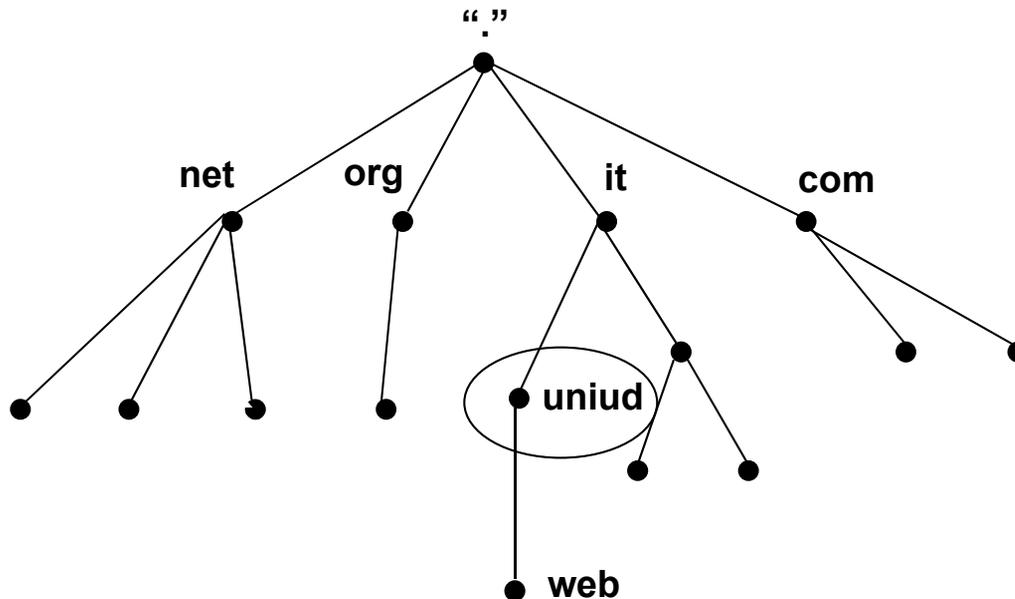


DNS

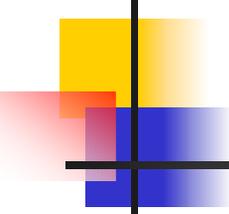
- Ogni nodo dell'albero può essere visto come la radice del sottoalbero ottenuto considerando soltanto i rami ed i nodi "sottostanti" al nodo considerato.
- Ogni sottoalbero corrisponde ad una porzione del database ed è chiamato **dominio**.
- Ogni dominio ha un nome unico, ottenuto concatenando tramite dei punti (".") le etichette dei nodi attraversati a partire dal nodo radice del dominio al nodo radice di tutto l'albero (tale nodo, detto anche "world", ha come etichetta il punto, che solitamente non si scrive nel nome risultante).

DNS - Esempio

- Consideriamo il nodo cerchiato nel seguente diagramma:



- Il nome del dominio corrispondente è: uniud.it (il punto finale, ovvero l'etichetta della radice dell'albero, non si scrive).

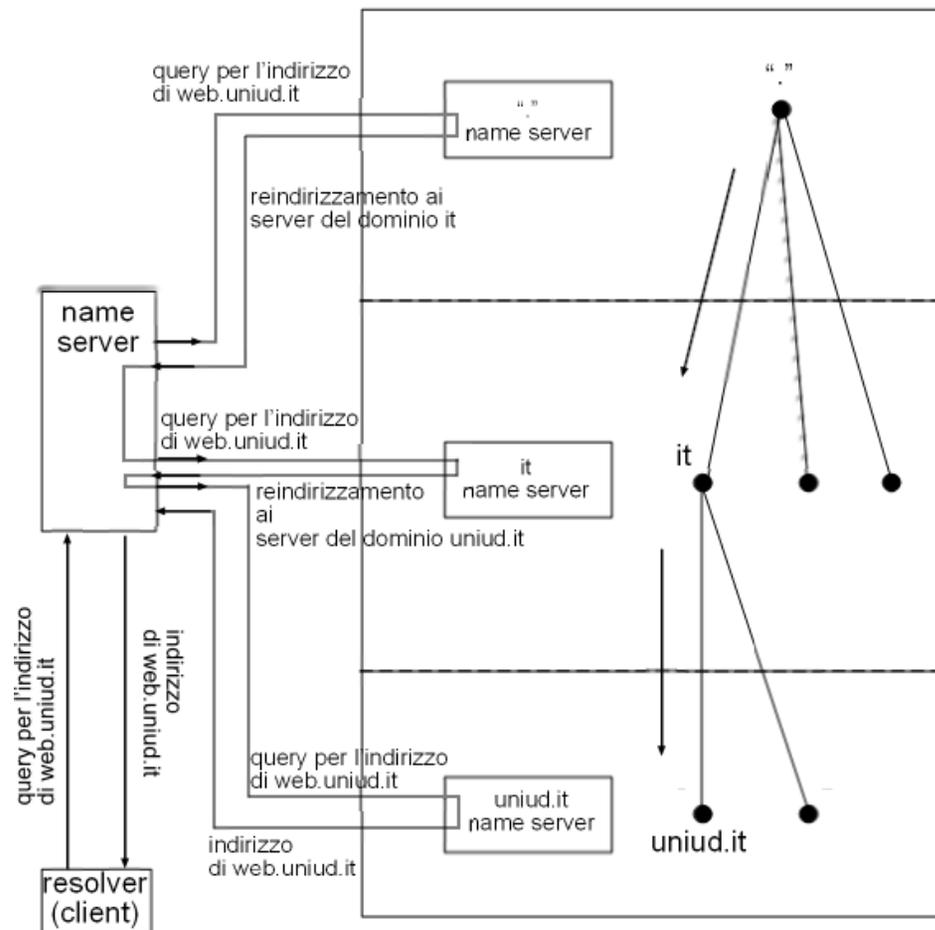


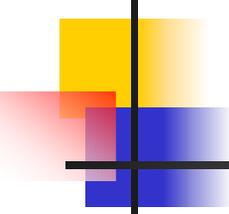
DNS

- I nodi figli della radice (".") sono i cosiddetti domini top level: **com** (e.g., IBM: ibm.com), **edu** (e.g., Università di Berkeley: berkeley.edu), **gov** (e.g., NASA: nasa.gov), **mil** (e.g., Esercito USA: army.mil), **net** (e.g., NSFNET: nfs.net), **org** (e.g., Electronic Frontier Foundation: eff.org), **int** (e.g., NATO: nato.int), **it** (e.g., Università degli Studi di Udine: uniud.it), ...
- Vantaggi del DNS:
 - Essendo il database distribuito, il carico di lavoro dovuto all'interrogazione, al mantenimento ed all'aggiornamento delle varie parti non va a gravare su un unico server.
 - Grazie alla struttura ad albero non vi sono rischi di duplicazione di nomi; infatti se anche due nodi appartenenti a domini distinti avessero la stessa etichetta, i rispettivi nomi completi (FQDN, ovvero, fully qualified domain name) sarebbero comunque diversi.

Risoluzione dei nomi

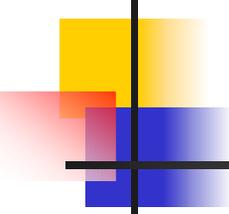
- Il seguente diagramma mostra i passi necessari per ottenere l'indirizzo (IP) del web server dell'Università di Udine (web.uniud.it):





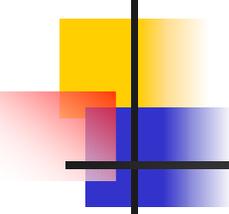
Risoluzione dei nomi

- La catena di eventi descritta nel precedente diagramma si verifica quando, ad esempio, vogliamo collegarci al sito web dell'Università di Udine (web.uniud.it):
 - Il resolver (e.g., il browser) richiede al name server della propria rete locale (e.g., al name server del proprio provider) l'indirizzo IP corrispondente al nome simbolico web.uniud.it
 - Se il name server non conosce la risposta a tale query (perché già eseguita recentemente), inoltra quest'ultima ad uno dei name server del mondo (".") di cui conosce l'indirizzo IP.
 - Il name server del mondo risponde fornendo l'indirizzo IP di uno dei name server del dominio it, a cui il name server della rete locale inoltra la query.
 - Il name server del dominio it risponde fornendo l'indirizzo IP di uno dei name server del dominio uniud.it, a cui il name server della rete locale inoltra la query.
 - Finalmente il name server del dominio uniud.it fornisce l'indirizzo IP corrispondente al nome simbolico web.uniud.it.
 - Tale indirizzo viene poi inviato al resolver.



HTML

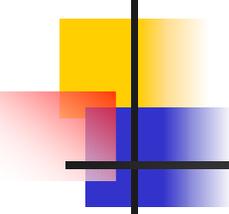
- Il World Wide Web (WWW) è uno dei servizi disponibili su Internet e rappresenta un nuovo modo per fornire informazione accessibile su scala mondiale.
- L'informazione viene codificata in pagine rese disponibili sui Web server. Ogni pagina è un documento **ipertestuale** e **multimediale**.
- Il linguaggio di specifica per le pagine Web è l'HTML (HyperText Markup Language). Si tratta di un linguaggio che utilizza dei marcatori (detti anche istruzioni) inseriti opportunamente nel testo "normale" in modo da conferire ad un documento una natura ipertestuale (i.e., la possibilità di una lettura non necessariamente sequenziale) e multimediale (inclusione di immagini, suoni, filmati).
- Quindi le pagine HTML non sono altro che dei file di testo in cui sono presenti dei marcatori.
- La visualizzazione di una pagina web è il risultato finale dell'interpretazione dei marcatori HTML da parte del browser (tale processo viene denominato talvolta "rendering").
- Nei prossimi lucidi si vedranno degli esempi di utilizzo dei marcatori principali.



Caratteri speciali

Nelle pagine HTML, bisogna utilizzare dei codici appositi per essere sicuri che certi caratteri vengano visualizzati correttamente da tutti i browser:

Entità carattere	Carattere corrispondente	Descrizione
<	<	Segno di minore
>	>	Segno di maggiore
&	&	Ampersand
"	“	Apice doppio
 		Spazio unificatore
©	©	Copyright
®	®	Marchio registrato
~	~	Carattere corrispondente al numero 126 nel set di caratteri di default (ISO Latin-1)



Intestazioni

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
```

```
<HTML>
```

```
<HEAD>
```

```
<TITLE> Esempi di intestazioni </TITLE>
```

```
</HEAD>
```

```
<BODY>
```

```
<H1>Intestazione di livello 1</H1>
```

```
<H2>Intestazione di livello 2</H2>
```

```
<H3>Intestazione di livello 3</H3>
```

```
<H4>Intestazione di livello 4</H4>
```

```
<H5>Intestazione di livello 5</H5>
```

```
<H6>Intestazione di livello 6</H6>
```

```
</BODY>
```

```
</HTML>
```

Intestazione di livello 1

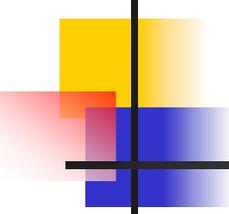
Intestazione di livello 2

Intestazione di livello 3

Intestazione di livello 4

Intestazione di livello 5

Intestazione di livello 6



Esempio

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
```

```
<HTML>
```

```
<HEAD>
```

```
<TITLE> Declaration of independence </TITLE>
```

```
</HEAD>
```

```
<BODY>
```

```
<H1 ALIGN="center">Declaration of Independence</H1>
```

```
<H5 ALIGN="right">(Adopted in Congress 4 July 1776)</H5>
```

```
<H3>The Unanimous declaration of the Thirteen United States of America</H3>
```

```
<P><H4 ALIGN="center">When, in the course of human events, it becomes necessary  
for one people to dissolve the political bonds which have connected them with another,  
and to assume among the powers of the earth, the separate and equal station to which  
the laws of nature and of nature's God entitle them,<BR> a decent respect to the  
opinions of mankind requires that they should declare the causes which impel them  
to the separation.</H4></P>
```

```
</BODY>
```

```
</HTML>
```

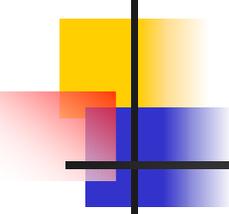
Declaration of Independence

(Adopted in Congress 4 July 1776)

The Unanimous declaration of the Thirteen United States of America

When, in the course of human events, it becomes necessary for one people to dissolve the political bonds which have connected them with another, and to assume among the powers of the earth, the separate and equal station to which the laws of nature and of nature's God entitle them,

a decent respect to the opinions of mankind requires that they should declare the causes which impel them to the separation.



Testo preformattato

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
<TITLE> Blocchi di testo </TITLE>
</HEAD>
```

```
<BODY>
```

In HTML per andare a capo
bisogna usare `
` in questo modo.

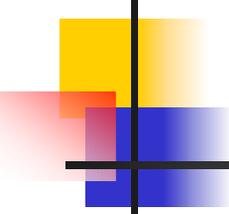
```
<PRE>
oppure bisogna racchiudere
il testo all'interno l'istruzione PRE.
</PRE>
```

```
</BODY>
</HTML>
```

Risultato:

In HTML per andare a capo bisogna usare `
`
in questo modo.

oppure bisogna racchiudere
il testo all'interno l'istruzione PRE.



Blocchi di testo

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
<TITLE> Blocchi di testo </TITLE>
</HEAD>
```

```
<BODY>
```

```
<P ALIGN="center">Esempio di paragrafo allineato <STRONG>centralmente</STRONG></P>
```

Citazione:

```
<BLOCKQUOTE>
```

Il flusso del tempo che corre veloce porta il tormento dell'intuire.

```
</BLOCKQUOTE>
```

```
<ADDRESS>
```

H.P. Lovecraft

Providence, Rhode Island

```
</ADDRESS>
```

```
</BODY>
```

```
</HTML>
```

Risultato:

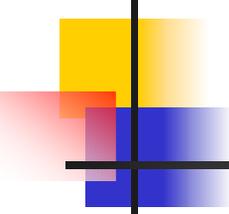
Esempio di paragrafo allineato **centralmente**

Citazione:

Il flusso del tempo che corre veloce porta il tormento dell'intuire.

H.P. Lovecraft

Providence, Rhode Island



Tabelle

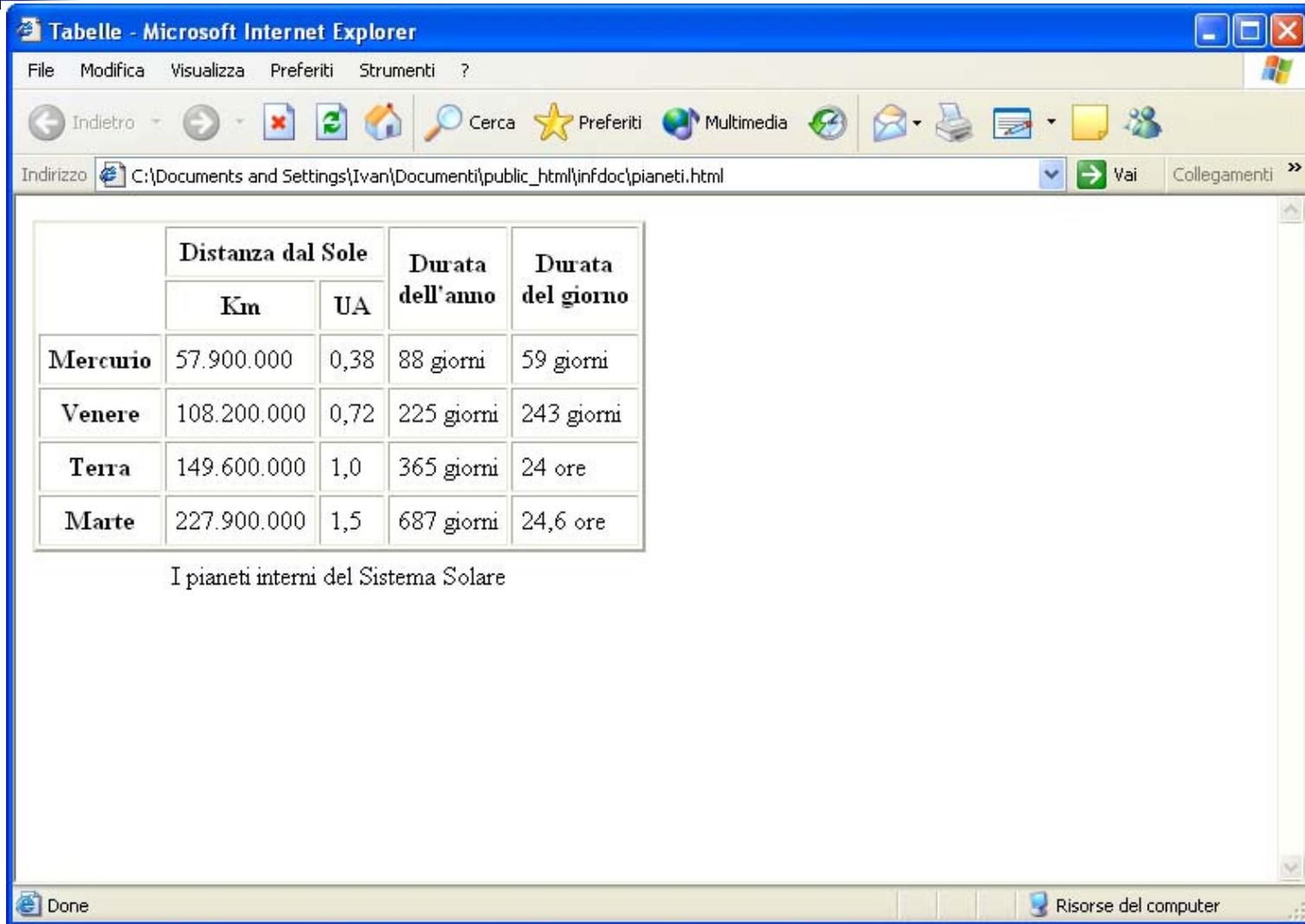
```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
<TITLE> Tabelle </TITLE>

</HEAD>

<BODY>
<TABLE CELLPADDING="5" BORDER="2">
<CAPTION ALIGN="bottom">I pianeti interni del Sistema Solare</CAPTION>
<TR>
    <TH ROWSPAN="2"></TH>
    <TH COLSPAN="2">Distanza dal Sole</TH>
    <TH ROWSPAN="2">Durata<BR>dell'anno</TH>
    <TH ROWSPAN="2">Durata<BR>del giorno</TH>
</TR>
<TR>
    <TH>Km</TH><TH>UA</TH>
</TR>
<TR>
    <TH>Mercurio</TH>
    <TD>57.900.000</TD>
    <TD>0,38</TD>
    <TD>88 giorni</TD>
    <TD>59 giorni</TD>
</TR>
<TR>
    <TH>Venere</TH>
    <TD>108.200.000</TD>
    <TD>0,72</TD>
    <TD>225 giorni</TD>
    <TD>243 giorni</TD>
</TR>
<TR>
    <TH>Terra</TH>
    <TD>149.600.000</TD>
    <TD>1,0</TD>
    <TD>365 giorni</TD>
    <TD>24 ore</TD>
</TR>
<TR>
    <TH>Marte</TH>
    <TD>227.900.000</TD>
    <TD>1,5</TD>
    <TD>687 giorni</TD>
    <TD>24,6 ore</TD>
</TR>
</TABLE>
</BODY>
</HTML>
```

continua a lato ⇒

Tabelle



The screenshot shows a Microsoft Internet Explorer browser window titled "Tabelle - Microsoft Internet Explorer". The address bar displays the URL "C:\Documents and Settings\Ivan\Documenti\public_html\infdoc\pianeti.html". The main content area contains a table with the following data:

	Distanza dal Sole		Durata dell'anno	Durata del giorno
	Km	UA		
Mercurio	57.900.000	0,38	88 giorni	59 giorni
Venere	108.200.000	0,72	225 giorni	243 giorni
Terra	149.600.000	1,0	365 giorni	24 ore
Marte	227.900.000	1,5	687 giorni	24,6 ore

I pianeti interni del Sistema Solare

The browser interface includes a menu bar (File, Modifica, Visualizza, Preferiti, Strumenti, ?), a toolbar with navigation and utility icons, and a status bar at the bottom showing "Done" and "Risorse del computer".

I righelli orizzontali

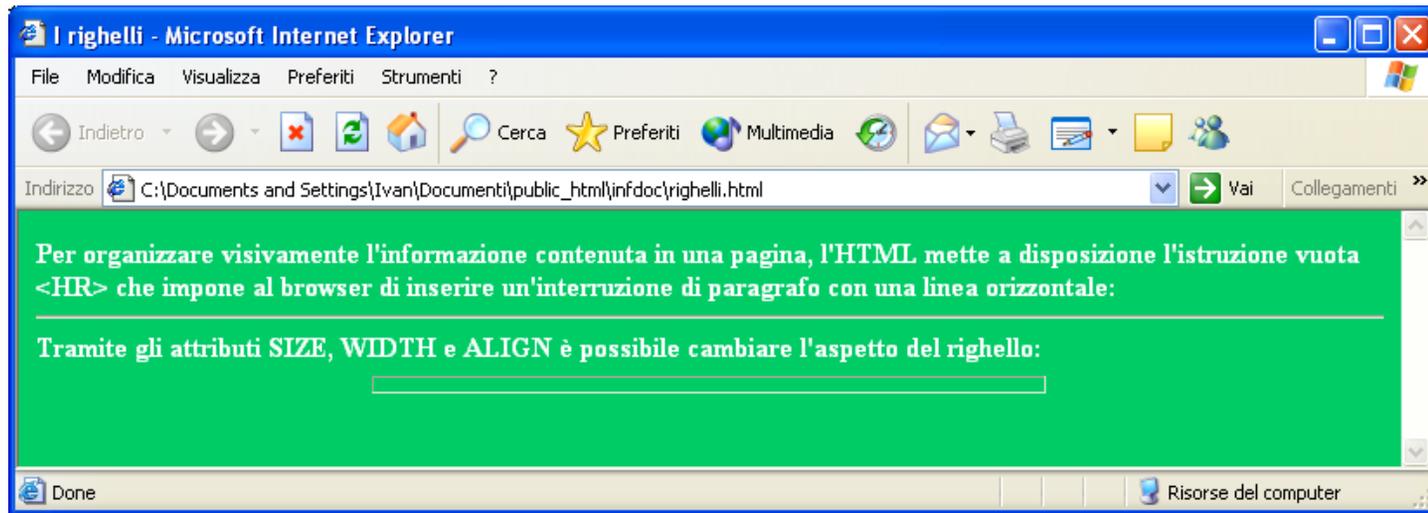
```
<HTML>
<HEAD>
<TITLE> I righelli </TITLE>
</HEAD>
```

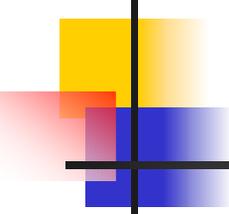
```
<BODY BGCOLOR="#00CC66">
<FONT COLOR="#FFFFFF">
<STRONG>
```

Per organizzare visivamente l'informazione contenuta in una pagina, l'HTML mette a disposizione l'istruzione vuota `<HR>` che impone al browser di inserire un'interruzione di paragrafo con una linea orizzontale: `<HR>` Tramite gli attributi `SIZE`, `WIDTH` e `ALIGN` è possibile cambiare l'aspetto del righello:

```
<HR SIZE="10" WIDTH="50%" ALIGN="center">
```

```
</STRONG>
</FONT>
</BODY>
</HTML>
```





Le immagini (sorgente)

```
<HTML>
```

```
<HEAD>
```

```
<TITLE> Le immagini </TITLE>
```

```
</HEAD>
```

```
<BODY>
```

```
<IMG SRC="C:\Documents and
```

```
Settings\Ivan\Documenti\Immagini\compagnia0.jpg"
```

```
ALT="Il Signore degli Anelli: La Compagnia dell'Anello" BORDER="10"
```

```
WIDTH="30%"
```

```
HEIGHT="70%">
```

```
<IMG SRC="..\..\Immagini\compagnia0.jpg"
```

```
ALT="Il Signore degli Anelli: La Compagnia dell'Anello" BORDER="10"
```

```
WIDTH="15%"
```

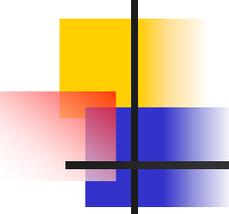
```
HEIGHT="35%">
```

```
</BODY>
```

```
</HTML>
```

Le immagini (risultato visivo)





I collegamenti

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
<TITLE> I collegamenti </TITLE>
</HEAD>
```

```
<BODY>
<A NAME="top"><H1>Inizio pagina</H1></A>
<P>Link ad un documento esterno su un altro web server (indirizzamento assoluto):<BR>
<A HREF="http://www.ilsignoredeglianelli.it ">Il Signore degli Anelli: Sito ufficiale</A></P>
<P>Link ad un documento sullo stesso web server in cui si trova questa pagina (indirizzamento relativo):<BR>
<A HREF="immagini.html">Il Signore degli Anelli: La compagnia dell'Anello</A></P>
<TABLE>
<TR><TH>Il Signore degli Anelli: Le due Torri</TH></TR>
<TR><TD><IMG SRC="../../Immagini/duetorri0.jpg" WIDTH="562" HEIGHT="600" BORDER="0"></TD></TR>
</TABLE>
```

```
<P>
Link ad un altro punto in questo stesso documento:<BR>
<A HREF="#top">Inizio pagina</A>
</P>
```

```
</BODY>
</HTML>
```

I collegamenti

I collegamenti - Microsoft Internet Explorer

File Modifica Visualizza Preferiti Strumenti ?

Indietro Cerca Preferiti Multimedia

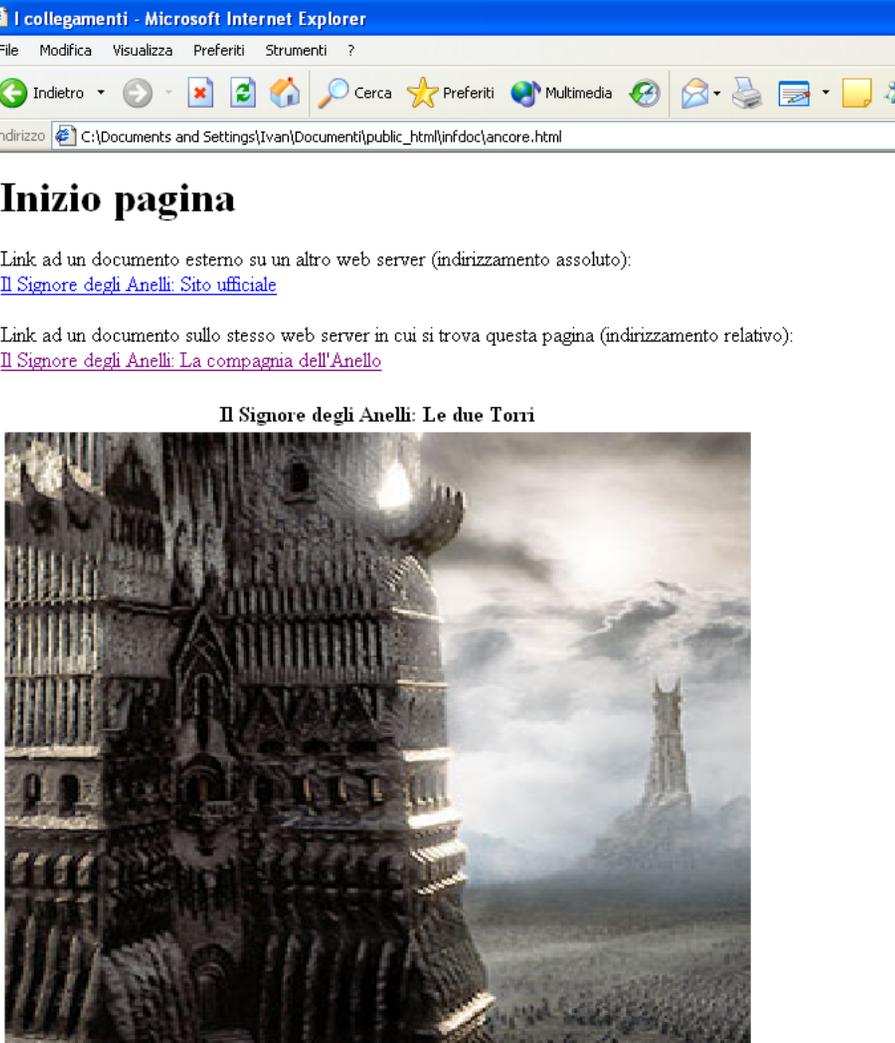
Indirizzo C:\Documents and Settings\Ivan\Documenti\public_html\infdoc\ancore.html

Inizio pagina

Link ad un documento esterno su un altro web server (indirizzamento assoluto):
[Il Signore degli Anelli: Sito ufficiale](#)

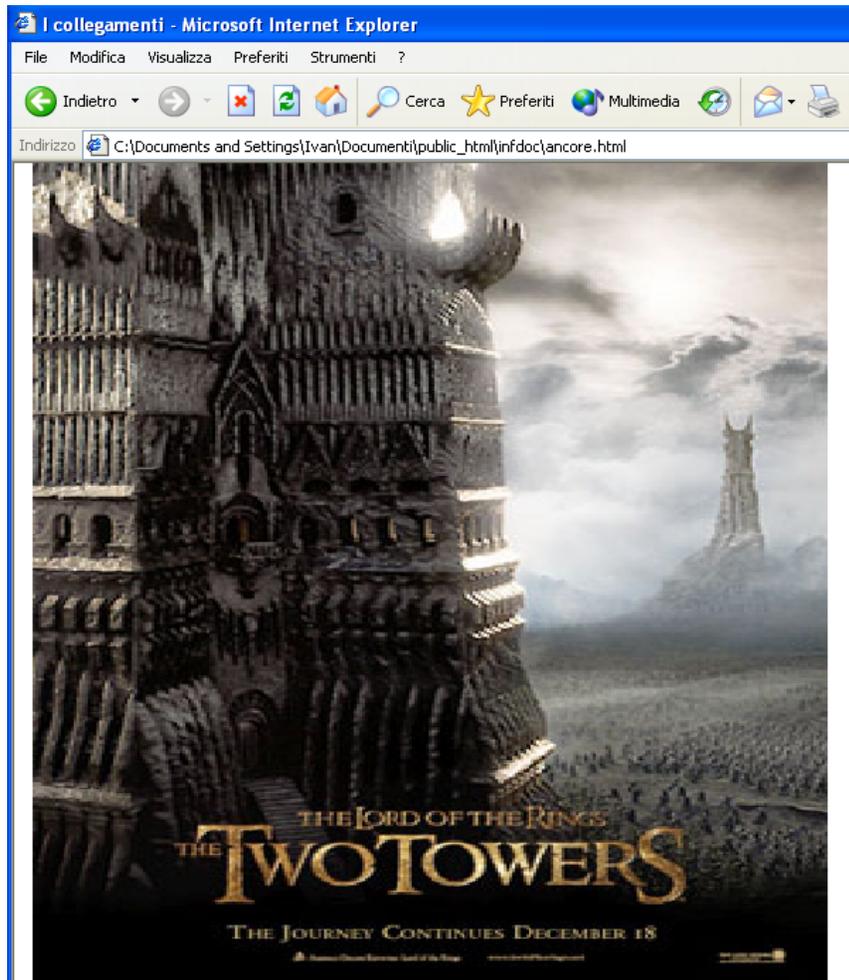
Link ad un documento sullo stesso web server in cui si trova questa pagina (indirizzamento relativo):
[Il Signore degli Anelli: La compagnia dell'Anello](#)

Il Signore degli Anelli: Le due Torri



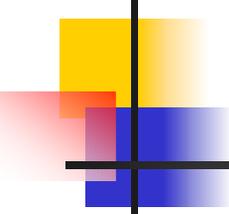
The image shows a screenshot of a Microsoft Internet Explorer browser window. The title bar reads 'I collegamenti - Microsoft Internet Explorer'. The menu bar includes 'File', 'Modifica', 'Visualizza', 'Preferiti', and 'Strumenti ?'. The toolbar contains icons for 'Indietro', 'Cerca', 'Preferiti', and 'Multimedia'. The address bar shows the local file path 'C:\Documents and Settings\Ivan\Documenti\public_html\infdoc\ancore.html'. The main content area features a heading 'Inizio pagina', followed by two links: one absolute link to 'Il Signore degli Anelli: Sito ufficiale' and one relative link to 'Il Signore degli Anelli: La compagnia dell'Anello'. Below the links is a sub-heading 'Il Signore degli Anelli: Le due Torri' and a large image of the Dark Tower in Minas Morgul from 'The Lord of the Rings'.

I collegamenti



Link ad un altro punto in questo stesso documento:

[Inizio pagina](#)



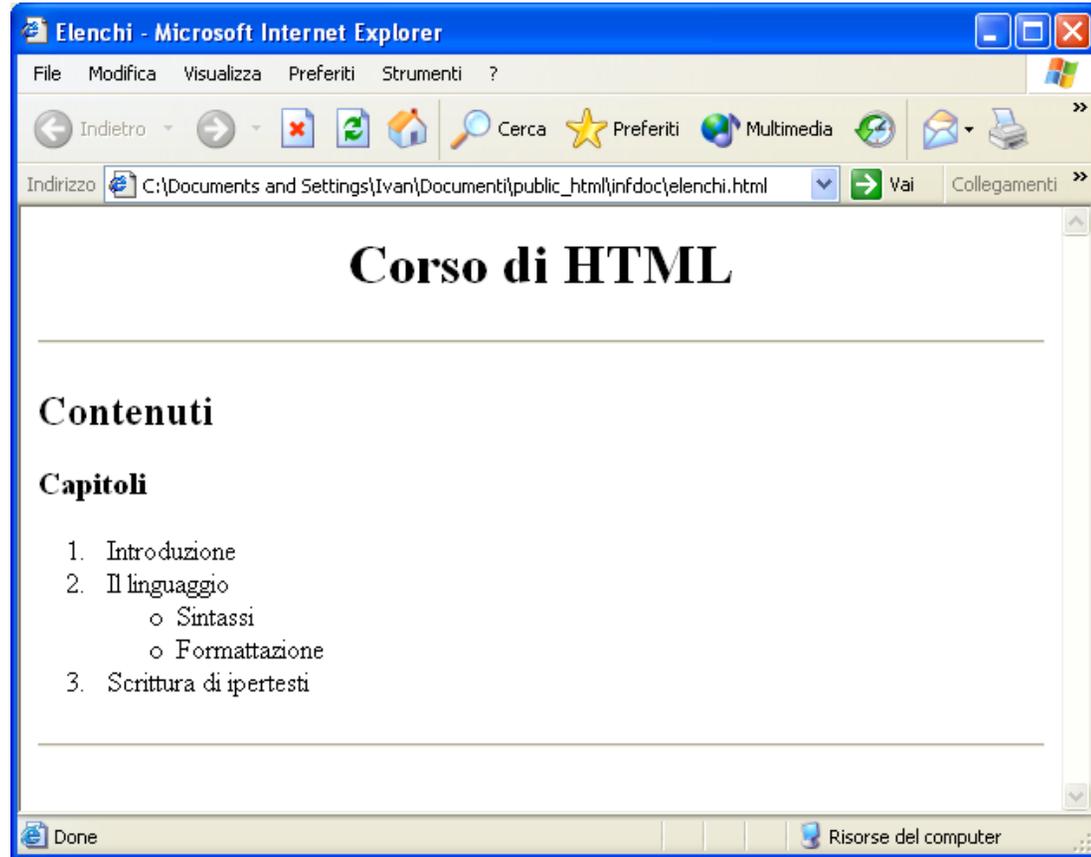
Gli stili del testo

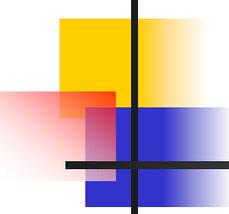
- Le istruzioni di stile cambiano l'aspetto tipografico del testo.
- Si possono individuare due categorie di stili:
 - Stili logici: rendono lo stile del testo relativo, dicendo, ad esempio, che una certa parola deve essere enfatizzata rispetto al contesto.
 - Stili fisici: permettono un controllo più preciso di come viene reso visivamente il testo dal browser.
- Esempi di istruzioni di stili logici:
 - `` (enfasi)
 - `` (maggiore enfasi)
 - `<TT></TT>` (Typewriter Text)
 - `<BIG></BIG>`, `<SMALL></SMALL>` (testo un po' più grande/piccolo)
 - `<STRIKE></STRIKE>` (testo barrato)
 - ``, `` (pedici e apici)
- Esempi di istruzioni di stili fisici:
 - `` (grassetto)
 - `<I></I>` (corsivo)
 - `<U></U>` (sottolineato)

Elenchi ordinati e non ordinati

```
<HTML>
<HEAD>
<TITLE> Elenchi </TITLE>
</HEAD>

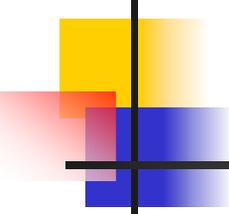
<BODY>
<H1 ALIGN="center">Corso di HTML</H1>
<HR>
<H2>Contenuti</H2>
<H3>Capitoli</H3>
<OL>
  <LI>Introduzione</LI>
  <LI>Il linguaggio</LI>
  <UL>
    <LI>Sintassi</LI>
    <LI>Formattazione</LI>
  </UL>
  <LI>Scrittura di ipertesti</LI>
</OL>
<HR>
</BODY>
</HTML>
```





Attributi degli elenchi

- Elenchi ordinati:
 - TYPE="1" (numerazione normale: default)
 - TYPE="A" (lettere maiuscole: A, B, C, D, ...)
 - TYPE="a" (lettere minuscole: a, b, c, d, ...)
 - TYPE="I" (numeri romani maiuscoli: I, II, III, IV, ...)
 - TYPE="i" (numeri romani minuscoli: i, ii, iii, iv, ...)
- Elenchi non ordinati:
 - TYPE="circle" (cerchi vuoti: default)
 - TYPE="square" (quadrati pieni)
 - TYPE="disc" (cerchi pieni)

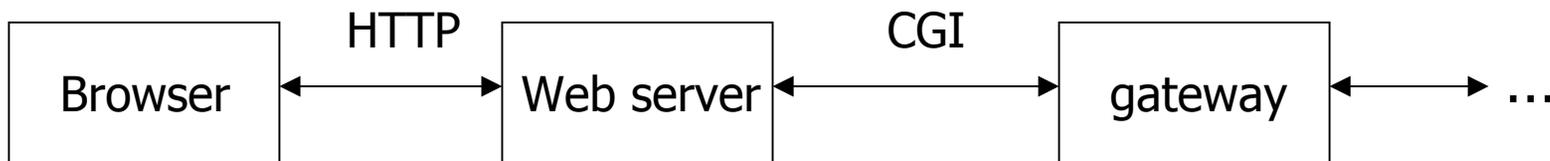


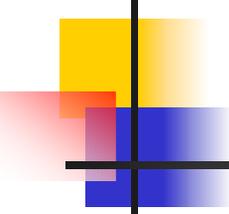
Il protocollo HTTP

- Il World Wide Web funziona per mezzo degli HTTP (HTML Transfer Protocol) server (detti anche Web Server) che forniscono le pagine richieste ai client (browser).
- Il protocollo HTTP specifica quattro fasi di interazione fra client e server:
 - Apertura della connessione: il browser richiede al server (specificato nell'URL) l'apertura di una connessione TCP (Transport Control Protocol) per verificare la disponibilità del servizio.
 - Connessione: il server (se possibile) accetta la connessione ed invia una conferma al browser.
 - Richiesta: il browser invia al server la richiesta di una risorsa (e.g., pagina HTML).
 - Risposta (reply): il server comunica al browser se la risorsa richiesta è disponibile ed in caso affermativo la invia, terminando nel contempo la connessione.
- Appare evidente quindi che il protocollo HTTP è **stateless**, i.e., non ha modo di tener traccia delle operazioni portate a termine per uno stesso client.

Gateway (programma CGI)

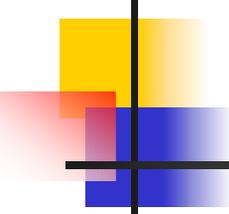
- Un gateway è un programma richiamato dal Web server che permette di stabilire una connessione fra Web ed un ambiente esterno (e.g., un DBMS).
- La comunicazione fra Web server ed i gateway è regolata da un meccanismo noto come CGI (Common Gateway Interface):
 - L'utente richiede l'esecuzione di un gateway cliccando su un apposito link o inviando i dati tramite un form.
 - Il Web server lancia il gateway passandogli i parametri inviati dall'utente.
 - Il gateway termina la sua esecuzione (nel corso della quale può interagire con altri elementi di un sistema informativo, e.g., un database) e restituisce i risultati al Web server.
 - Il Web server trasmette i risultati ricevuti dal gateway al client (browser).





Web Information System (WIS)

- Il Web sta progressivamente imponendosi come l'interfaccia standard per l'accesso ai sistemi informativi esistenti e futuri.
- L'informazione viene resa disponibile agli utenti sotto forma di pagine HTML.
- Siccome il contenuto delle pagine HTML è stabilito dai dati contenuti nel sistema informativo e questi ultimi sono soggetti a variazioni nel corso del tempo, si rende necessario aggiornare periodicamente le pagine HTML. In questo modo gli utenti possono accedere ad informazioni sempre aggiornate.
- Per quanto riguarda l'aggiornamento delle pagine Web, se quest'ultimo viene fatto manualmente (pagine statiche) alla lunga possono nascere delle incoerenze (dovute alla ridondanza dei dati presenti nelle varie pagine) ed al volume (i.e., numero di pagine) del sito.
- Per superare tale problema e fornire informazioni corrette ed aggiornate, si ricorre alle pagine generate dinamicamente dai programmi CGI, che leggono le informazioni dai database e **generano** la pagina da fornire al browser in tempo reale in base ai dati letti).



XML

- XML significa **EX**tensible **M**arkup **L**anguage.
- Si tratta di un linguaggio che fa uso di marcatori (tag) analogamente a quanto visto per l'HTML.
- Lo scopo dell'XML è di descrivere dati.
- Diversamente dall'HTML, l'utente può definire un suo insieme di marcatori, a seconda dei dati che vuole descrivere; quindi in XML i marcatori **non** sono predefiniti.
- La filosofia alla base di XML è che un documento scritto in tale linguaggio deve essere auto-descrittivo.
- Al giorno d'oggi XML è principalmente usato come "lingua franca" per lo scambio di dati fra sistemi incompatibili.

Esempio di file XML

```
<?xml version="1.0" encoding="ISO-8859-1"?>
```

```
<messaggio>
```

```
  <mittente>Paolo Bianchi</mittente>
```

```
  <destinatario>Mario Rossi</destinatario>
```

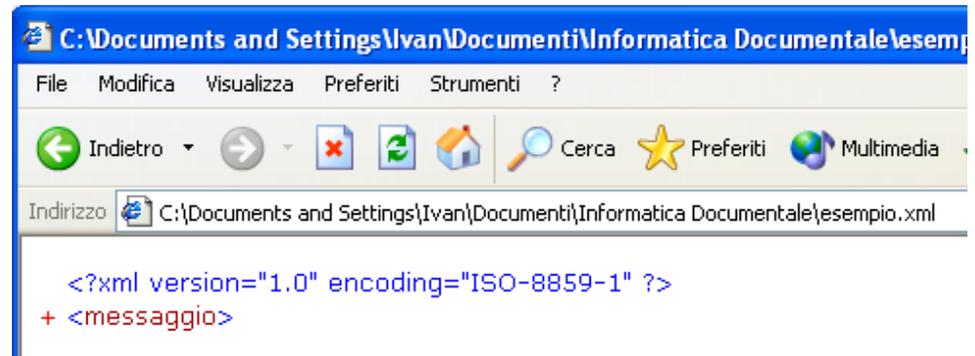
```
  <oggetto>Riunione di lavoro</oggetto>
```

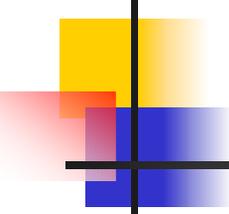
```
  <testo>
```

La riunione di lavoro fissata per domani 17/05/2003 è stata rinviata.

```
</testo>
```

```
</messaggio>
```





Dati in formato XML

- Un file XML può essere usato per memorizzarvi dei dati (analogamente a quanto avviene con i database relazionali).
- A differenza del modello relazionale, l'XML si rivela più flessibile; infatti, mentre la definizione di uno schema relazionale vincola rigidamente il numero dei campi presenti in ogni record (tupla), con un file XML possiamo averne un numero variabile.

Ad esempio se consideriamo lo schema Progetto(Denominazione, Finanziatore1, Finanziatore2), ogni record potrà avere al massimo due finanziatori.

Se per un certo progetto si volesse specificare 3 finanziatori, bisognerebbe modificare lo schema. In XML invece è sufficiente aggiungere il terzo finanziatore:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<progetti>
  <progetto>
    <denominazione>Ponte autostradale</denominazione>
    <finanziatori>
      <finanziatore>Autostrade S.p.A.</finanziatore>
      <finanziatore>ANAS</finanziatore>
      <finanziatore>Regione Lazio</finanziatore>
    </finanziatori>
  </progetto>
</progetti>
```