

World Wide Web

Angelo Di Iorio

angelo.diiorio@unibo.it

Cosa è il World Wide Web?

- Il World Wide Web è un sistema ipertestuale per la presentazione a schermo di documenti multimediali, e per l'utilizzo di link per la navigazione.
- Il sistema è distribuito e scalato su Internet
- Che relazione/differenza tra Internet e il World Wide Web?

Due tipi di architetture

- **Client/server:** se un calcolatore (client) ha bisogno di un servizio chiede ad un altro calcolatore (server) di fornire il servizio in questione. *Esempio: WWW*
- **Peer-to-Peer (P2P):** i nodi della rete che comunicano/collaborano rivestono ruoli interscambiabili e svolgono le stesse funzioni.
Esempio: eMule
 - i nodi della rete comunicano tra di loro direttamente e senza bisogno di intermediari

L'architettura del WWW

- L'organizzazione del WWW è basata su servizi di tipo ***client-server***:
 - I **server** mettono a disposizione le informazioni, cioè i documenti ipertestuali, che sono residenti sul server stesso.
 - I **client (browser)** accedono ai documenti multimediali memorizzati sui server, attraverso un indirizzo che individua univocamente i documenti multimediali sulla rete.
- Il protocollo di comunicazione è HTTP, HyperText Transfer Protocol

L'indirizzamento sul WWW: URL

- I dettagli di HTTP non ci interessano, ma è importante capire come indirizzare e recuperare le risorse sul Web
- Ogni risorsa è univocamente individuata da una sequenza di caratteri che ne costituisce l'indirizzo, chiamata URL (Uniform Resource Locator) o URI (Uniform Resource Identifier)

`http://www.a3.unibo.it/registrazione/index.html`

Protocollo

Nome simbolico di dominio

Nome locale del documento

Quando si usano questi URL?

- Gli URL possono essere forniti ad un browser direttamente dall'utente, inserendoli nella barra di navigazione
- Oppure possono essere scelti implicitamente cliccando su un link. Il "click" viene trasformato in una richiesta HTTP al dato URL
- Se un documento include altre risorse, infine, queste sono identificate (e recuperate) tramite i loro URL
 - immagini, video, etc.
 - programmi da eseguire sul browser (applets/javascript)

Ma come è scritta una pagina Web?

- In HTML (HyperText Markup Language), un linguaggio che “marca”:
 - struttura del documento
 - informazioni di presentazione
 - collegamenti ipertestuali (e URL della destinazione)
 - risorse multimediali (e URL)
- Il browser è in grado di interpretare queste informazioni e visualizzare la pagina finale

Il codice sorgente di una pagina HTML...

```
<HTML>
<HEAD>...</HEAD>
<BODY>
<H1> Questo &egrave; il titolo</H1>
<P>Questo documento descrive con un esempio l'uso di
HTML. </P>
<H2> Essenza di HTML </H2>
<P>Un documento HTML consiste di testo "immerso" in
comandi HTML (tecnicamente: tags). I comandi
descrivono la struttura del documento. HTML
permette:</P>
<UL>
<LI> di descrivere la struttura del documento;</LI>
<LI> inserire elementi non testuali, come questo
stemma: <IMG src="./logoUnibo.gif"
ALT="logo dell'universita di Bologna">;</LI>
```

[...]

...e il risultato sul browser

Questo è il titolo

Questo documento descrive con un esempio l'uso di HTML.

Essenza di HTML

Un documento HTML consiste di testo "immerso" in comandi HTML (tecnicamente: tags). I comandi descrivono la struttura del documento. HTML permette:

- di descrivere la struttura del documento;
- inserire elementi non testuali, come questo stemma:  ;
- inserire riferimenti ipertestuali: [Dipartimento di Scienze dell'Informazione](#)

Un esempio di caricamento di pagina Web

- Il **browser** include quindi un **client HTTP** e un **visualizzatore di pagine HTML**

<http://www.w3.org/Addressing/Activity>

The screenshot shows a Mozilla Firefox browser window with the address bar containing <http://www.w3.org/Addressing/Activity>. The page content includes the W3C logo and navigation links for 'Architecture domain' and 'URI home'. A left sidebar lists various W3C topics, with 'Activity Statements' highlighted. The main content area features the title 'Uniform Resource Identifier (URI) Activity Statement' and a paragraph explaining the goal of the Web. Below this, there are sections for 'Highlights Since the Previous Advisory Committee Meeting' and 'Upcoming Activity Highlights'. The browser's status bar at the bottom shows search and navigation controls.

URI Activity Statement - Mozilla Firefox

File Modifica Visualizza Cronologia Segnalibri Strumenti ?

http://www.w3.org/Addressing/Activity

W3C Architecture domain URI home

W3C Home
About W3C
W3C Activities
Activity Statements
Device Independence
Document Object Model
Extensible Markup Language (XML)
Graphics
HTML
Incubator
Internationalization
Math
Mobile Web Initiative
Multimodal Interaction
Patent Policy
Privacy
Quality Assurance
Rich Web Client
Security
Semantic Web
Style

Uniform Resource Identifier (URI) Activity Statement

One goal of the Web, since its inception, has been to build a global community in which any party can share information with any other party. To achieve this goal, the Web makes use of a single global identification system: the URI. URIs are a cornerstone of Web architecture, providing identification that is common across the Web.
—[Architecture of the World Wide Web, Volume One, section 2](#)

The [W3C URI Interest Group](#), in collaboration with the [W3C Technical Architecture Group](#) supports the [IETF URI and IRI specifications](#).

Highlights Since the Previous Advisory Committee Meeting

Milestones in the URI Activity:

- January 2005: [W3C Supports the URI Standard and IRI Proposed Standard](#) See [press release](#).
- February 2005: [W3C Launches URI Interest Group](#); [URI Interest Group](#) replaces the URI Coordination Group.

Upcoming Activity Highlights

This Activity has now closed. This page will not be updated any further.

The [URI Interest Group](#) is scheduled to continue through January 2007. Review of [Guidelines and Registration Procedures for new URI Schemes](#) is concluding. We expect it to update RFC 2718 in the coming months.

Trova: _____

Completato

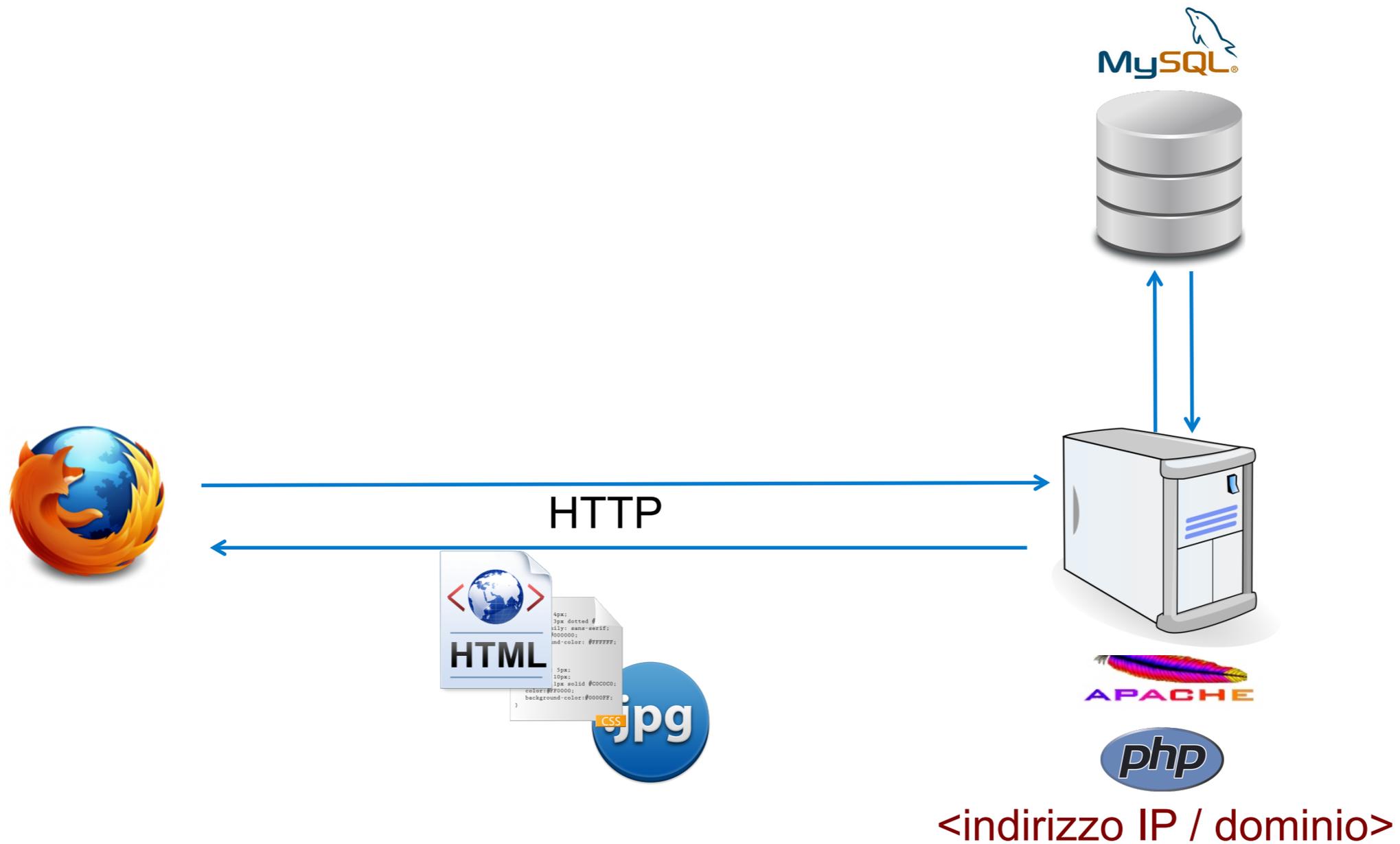
Le fasi del caricamento

1. Il browser determina la correttezza formale dello URL;
2. il browser sfrutta i livelli inferiori di TCP/IP (in particolare: il DNS per trasformare `www.w3.org` in un indirizzo IP; il protocollo TCP per inoltrare una richiesta al server che risiede su `www.w3.org`, ecc.);
3. Il browser inoltra al server di `www.w3.org` la richiesta del documento `/Addressing/Activity`;
4. il server invia il file richiesto e rilascia la comunicazione;
5. il browser interpreta il file ricevuto, rileva che la sua visualizzazione completa richiede la presenza di una figura, contraddistinta dallo URL `http://www.w3.org/Icons/WWW/w3c_home`
6. il browser inoltra la richiesta della figura al server indicato;
7. e così via per ogni altra componente del documento.

Contenuti statici



Contenuti dinamici e database



Cos'è il markup?

- Definiamo markup ***ogni mezzo per rendere esplicita una particolare interpretazione di un testo.***
- Per esempio, tutte quelle aggiunte al testo scritto che permettono di renderlo più fruibile.
- Oltre a rendere il testo più leggibile, il markup permette anche di specificare ulteriori usi del testo.
- Con il markup per sistemi informatici (il nostro caso), specifichiamo le modalità esatte di utilizzo del testo nel sistema stesso.
- Il markup esiste da prima dell'informatica!
- HTML è un linguaggio di markup (basato su SGML)

Struttura di una pagina Web (HTML)

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

```
<html>
```

```
  <head>
```

```
    <title>Titolo del documento</title>
```

```
  </head>
```

```
  <body>
```

```
    <p class='primo'>Testo di un paragrafo</p>
```

```
  </body>
```

```
</html>
```

- Gli elementi sono le parti di documento dotate di un senso proprio.
- Il titolo, l'autore, i paragrafi del documento sono tutti elementi.
- Un elemento è individuato da un tag iniziale, un contenuto ed un tag finale.
- **Non confondere i tag con gli elementi!**

```
<p>Tre uomini in barca</p>
```

- Gli attributi sono informazioni aggiuntive sull'elemento che non fanno effettivamente parte del contenuto (meta-informazioni).
- Essi sono posti dentro al tag iniziale dell'elemento. Tipicamente hanno la forma nome="valore"

```
<div class="indirizzo">...</div>
```

Possiamo considerare i tag di HTML organizzati secondo alcune categorie:

- Tag della struttura complessiva (HTML, HEAD, BODY)
- Tag inline (B, I, SPAN, ecc.)
- Tag di blocco e liste (P, H1, H2, DIV, UL, OL, ecc.)
- Tag speciali (A, IMG, HR, BR)
- Tabelle (TABLE, TR, TD, TH)

Noi guardiamo solo pochi elementi e attributi, per avere un'idea generale del funzionamento di HTML

I tag di blocco definiscono l'esistenza di blocchi di testo che contengono elementi inline.

Elementi base:

- P (paragrafo),
- DIV (generico blocco),
- PRE (blocco preformattato),

Blocchi con ruolo strutturale

- H1, H2, H3, H4, H5, H6 (intestazione di blocco)

Le liste di elementi sono contenitori di elementi omogenei per tipo.

- UL: Lista a pallini di ; Attributo type (disc, square, circle)
- OL: lista a numeri o lettere di ; attributi start (valore iniziale) e type (1, a, A, i, I).

```
<UL>  
  <LI>Primo</LI>  
  <LI>Secondo</LI>  
  <LI>Terzo</LI>  
</UL>
```

I link sono definiti attraverso l'elemento A (anchor interna al documento). <A> è sintatticamente un elemento inline (non interrompe il flusso di testo)

Attributo principale:

- HREF: specifica l'URI di una destinazione. Quindi ... è un *estremo di partenza* di un link.

```
<p>Un link al <a href="B.html"> documento  
B</a></p>
```

- Scrivere il codice HTML per visualizzare una lista puntata di link. La lista contiene 4 item, ognuno dei quali corrisponde ad una stagione e punta alla pagina `<nomestagione>.html`