



Universidad  
Politécnica  
de Cartagena

# Applications on the Internet

Lesson 1. The World Wide Web



*At the end of the lesson the student will be able to:*

- *Explain the operation of the Web: its technologies, protocols and resources.*
- *Describe the evolution of the basic components the Web, the reasons for its initial development and the functional needs that it covers.*
- *Enumerate other distributed systems operating over the Internet and describe their differences over the Web.*



1. What is the Web (World Wide Web)?
  1. Components
  2. Operation
2. Web history.
  1. Evolution.
3. Other distributed systems and services.
4. Quiz
5. Bibliography and extras



# What is the Web?

## World Wide Web (WWW) = The Web

- Definition: "An access service, via the Internet, to *hyperlinked* documents/**resources**. "

## Is Internet = The Web?

- NO. They are used interchangeably, but technically, **the Web is a service implemented on the Internet infrastructure.**

## Internet

- A global system of interconnected networks.
- It provides the physical (servers, routers, cables, antennas, etc.) and logical (software, protocols) infrastructure on which multiple services are offered.
  - Services that use it: Web, email, instant messaging, file sharing (P2P) ...



# What is the Web?

**The Web as a system is a set of simple technologies used to access, via Internet, to *linked resources*. It is constantly evolving.**

**These simple technologies are used to:**

- Identify and locate a resource (URI / URL)
- Represent a resource (HTML and others)
- Access, transfer or interact with a resource (HTTP)

**But what is a resource?**

- *Any information that can be named:* A document, an image, a video, a transitional service ( "time in Cartagena" "hotel in Mallorca"), etc.



# What is the Web?

## Activity

-  = "Enter in your browser bar the following string:"
- <http://www.aemet.es/es/eltiempo/prediccion/municipios/horas/cartagena-id30016>

URI / URL `http://www.aemet.es/es/eltiempo/prediccion/municipios/horas/cartagena-id30016`

### Representation (HTML)

```
<!DOCTYPE html PUBLIC ...  
<html> <Head>  
<title> Weather ... </title>  
...  
</html>
```

### Resource

"Weather in  
Cartagena"

Identifies

Represents





# What is the web?

## Where is the resource?

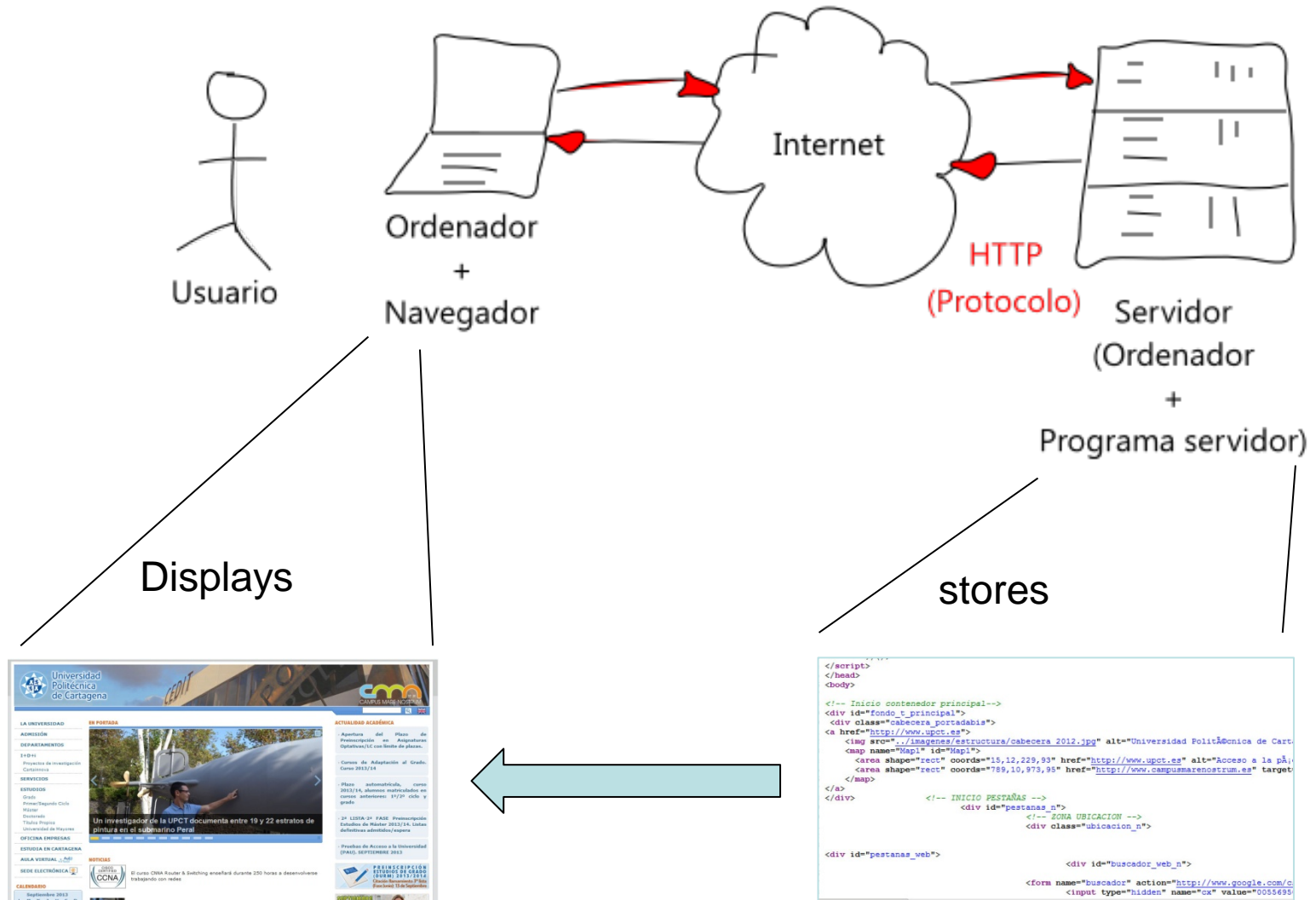
- Resources are *hosted* in Servers.
- **Server:** basically a computer running a **server program**, which serves requests and returns documents stored on its hard drive.

## Browser = User Agent

- A program (application software) that is responsible for:
  - Transform the URI / URL to a suitable request and send it to the server (connect to it).
  - Get the answer and convert it to a format appropriate for the user -> display on screen (such as text, video, audio ...)



# Components of the Web







# How do you work with the Web?

- Resources are represented by HTML.
- Resources are identified by URI / URL.
- HTML documents are stored on servers.
- The server runs an HTTP server program.
- Typical workflow:
  1. The user enters a URL in the browser.
  2. The browser generates a request to the server from the URL.
  3. The request is submitted using the HTTP protocol over TCP / IP.
  4. The server receives the request and returns the requested document (HTML + additional header code).
  5. The browser receives the data and displays them on the screen (HTML processes the received).



## **Initial conception: how to share scientific documents with colleagues?**

### **It starts in March 1989 at CERN:**

- Original proposal by physicist Tim Berners-Lee.
- Uses the concept of hypertext, conceptually created by Vannevar Bush in 1945.
- Berners-Lee developed the first prototypes of the 3 ideas: URLs, HTTP and HTML.
- Created the first prototype of a browser in 1990 (console mode, text only).

### **Mosaic: First graphical browser (1993)**

- Marc Andreessen (Univ. Of Illinois).
- Attaches graphical info to documents (images).

**In 1992, 50 servers worldwide.**

**Today, more than 1300 million users.**



# Evolution of the Web

## Initially the Web is static

- HTML documents stored on a server are retrieved by customers.
- In practice, it is only transfer of text files with images.

## The Web increasingly becomes more dynamic

- The user is able to send information to the server.
- HTML is generated dynamically by program before being served, usually based on the information provided by the user.
- New formats (video) and functionality (executable code on the client) are incorporated.

## The Web becomes a platform for applications

- Services and programs are developed specifically for use on the Web. Exploit the inherent *connectivity and scale* the Web.
- Web technologies support elements similar to that of a desktop application.

## Is the Web declining today?

- [http://www.cs.virginia.edu/~Robins/Long\\_Live\\_the\\_Web.pdf](http://www.cs.virginia.edu/~Robins/Long_Live_the_Web.pdf)





# Organization and governance

---

**The Web, as a service, does not need governing bodies, unlike the Internet. But it is necessary to establish international standards.**

**The technologies used are specified in standards, mostly published by the World Wide Web Consortium (W3C) and the Internet Engineering Task Force (IETF)**

-  <http://www.w3.org/>
-  <http://www.ietf.org/>



# Summary and key to its success

**Resource:** anything, information that can be identified

**The Web provides technologies to access resources linked.**

**Based on three simple technologies whose goal is respectively:**

- Identify resources
- Represent resources
- Transfer resources

**The key to its success: Opening**

- Can be **extended** and implemented in different ways without changing its functionality.
- The Web as a system **imposes no restrictions on** the type of resources it can accommodate.
- Simply defines how these resources can be exchanged between computers (and thus between people).
- Anyone can add new resources. Anyone can link them together.



# Other distributed systems / services 1

---

**There are a multitude of systems / services operating on the Internet infrastructure which are conceptually different from the Web.**

- There are shared elements. For example, URLs are used in many of them. The HTTP protocol also.
- **email**
- **Peer-to-peer networks (P2P)**
- **Instant messaging**





# Integration of services on the Web

## Many services have been integrated into the Web

- An intermediary "**Web Interface**" service is offered. HTTP requests are transformed into the service requests and the response is transformed to HTML.
- Example: web mail or email via the Web, as gmail, hotmail, yahoo, etc.
- Other services are directly offered /reimplemented as web applications, for example, Usenet (news) is a precursor of web forums, although they are technically different.
- The integration of many business applications is done by exporting its functionality through web interfaces for machines (Web Services). Equivalent to the remote procedure call (RPC) but implemented on the Web.
- Example: metasearch of flights, hotels, etc.



## HTML (HyperText Markup Language)

- Markup language used to represent resources on the Web.

### hyperlink

- It is a reference (pointer) to external data. In HTML it is a mark that includes a URL identifier pointing to an external document (or parts of the document itself) and that the user can follow.

### hypertext

- Hyperlinked text



## HTTP (HyperText Transfer Protocol)

- The protocol used to transfer data on the Web. It is an (1) application-level, (2) text-based, (3) request / response and (4) stateless protocol

## URI (Uniform Resource Identifier)

- Character string used to identify a resource on the Web.
- AURI can be classified into URL or URN.
- URLs are **locators** and indicate the operation to act on them.
- The URN **uniquely identify** a resource but do not specify how to locate it or act upon it.



## HTTP server

- A program (application software) serving requests done with the HTTP protocol.

## Server

- A computer running a HTTP server, which stores documents / files / resources that can be served via HTTP.

## Browser

- A program (application software) that implements a **HTTP client** and is capable of **process and display HTML documents** and other formats used as resources on the Web.



# Questions and exercises

---

- 1. Describe the differences and similarities regarding the operation, resources, identification and protocol between a) the Web and bittorrent b) the Web and Whatsapp.**
- 2. Describe 1) the steps you have to perform as a user and 2) the components of the web involved to: a) publish a document on the web b) access to that document.**
- 3. Describe in what way many new applications for web and mobile applications are no longer open, as opposed to the model of the web.**



# Bibliography and extras

- Web architecture: <http://www.w3.org/TR/2004/REC-webarch-20041215/>
- Representational State Transfer (REST) is the formal architecture on which the Web is based:  
[http://www.ics.uci.edu/~fielding/pubs/dissertation/rest\\_arch\\_style.htm](http://www.ics.uci.edu/~fielding/pubs/dissertation/rest_arch_style.htm)
- Peer-to-peer: Buford, J., Yu, H., Iua, E., *P2P Networking and Applications*, Morgan Kaufman 2009
- L. Shklar R. Roses, "Web Application Architecture: Principles, Protocols and Practices"., 2nd ed, Wiley, 2009.