

# Tema 0: Introducción

---

1. Internet.
2. WWW.
3. HTTP.
4. Modelo cliente-servidor.
5. Páginas estáticas-dinámicas.
6. Generación dinámica en el lado del servidor.
7. La Web de 3ª generación

# 1. Internet.

---

- Internet nació como una red de computadoras del ejercito de USA (ARPANET).
- Posteriormente la red se extendió a varias universidades.
- Actualmente es una enorme red de computadoras conectadas a nivel mundial.
- Una de las claves del crecimiento de Internet ha sido la aparición de la Web.
- Recientemente aparecen los sitios web corporativos: Aunque en origen, Internet fue concebido como un medio de transmisión de información militar y científica. Con el paso del tiempo, la Red fue desarrollándose y dio cabida a distintas manifestaciones lúdicas. El uso puramente empresarial tardó algo más en extenderse. Hace solo cinco años pocas empresas ajenas al sector de las tecnologías de la información y la comunicación contaban con sitios web corporativos. Hoy en día, tener una sede web se está convirtiendo en una necesidad para la mayoría de las empresas, independientemente de su tamaño o su volumen de facturación.

## 1. Internet (II).

---

- Servicios estándar de Internet:
  - WWW.
  - Correo electrónico.
  - Chat.
  - FTP.
  - Telnet, etc..

## 2. WWW (I)

---

- Definiciones World Wide Web (WWW) :
  - Es un servicio de intercambio de información multimedia a nivel mundial (Liu, 1994).
  - Es un entorno global en el cual toda la información accesible a través de Internet (texto, imágenes, audio, video, etc.) es accesible de una manera consistente y simple utilizando un conjunto de convenciones para el nombre y el acceso (Revista WebMaster, 1996).
  - Es el universo de la información accesible a través de Internet, una representación de todo el conocimiento humano (W3C, 1999).

## 2. WWW (II)

---

- Algunas características :
  - Las aplicaciones que intercambian la información están basadas en un esquema cliente/servidor.
  - La información está inherentemente distribuida.
    - Cualquiera puede añadir más información.
  - Se accede a la información de una manera estándar (tanto a la hora de solicitar como a la hora de ofrecer dicha información) independiente del sistema específico donde resida o a la cual se envía (URL y HTTP).
  - La información es descrita en términos estándar (HTML).

## 2. WWW (III)

---

- La Web: basa su diseño y funcionamiento en:
  - Sistema de direcciones **URI** (Uniform Resource Identifier): identifica el recurso.
    - Protocolo://usuario:clave@máquina:puerto/directorio/fichero#subdirección?parámetros
    - <http://informatica.uv.es/it3guia/IST>
    - <http://pepe:cds3@servidor.com:8080/entrada/inicio>
    - <ftp://ftp.uca.es/imagenes/globo.gif>
  - **HTTP** (Hypertext Transfer Protocol): Protocolo estándar usado para la comunicación, y transferencia de los recursos, entre los servidores web y los clientes (navegadores).
  - **HTML (XHTML)** (Hypertext Markup Language): Lenguaje estándar utilizado para describir el contenido semántico (no la apariencia) de los documentos web.

## 2. WWW (IV)

---

- **Según un informe de la consultora Netcraft:**
  - En 2007 se crearon en todo el mundo nada menos que 50 millones de sitios web, lo que significa que en esos momentos había más de 155 millones de páginas en internet.
  - Este incremento es muy superior al alcanzado en 2006 (30 millones de páginas más), y se debe en gran parte a la creación de blogs personales por muchos usuarios.
  - Internet alcanzó los 186 millones de sitios Web en 2008. Sin embargo, el mundo digital registró un crecimiento menor a años anteriores. De acuerdo al estudio, en diciembre se contabilizaron **186.727.854** de páginas, un 17% más que en 2007.
  - En total, en 2008 se crearon unos 31,1 millones de portales nuevos. El número de sitios Web se duplicó desde 2005, cuando habían unas 72.251.156 millones de páginas.
  - El servidor de código abierto Apache es la ciberpágina más visitada del mundo.

### 3. HTTP

---

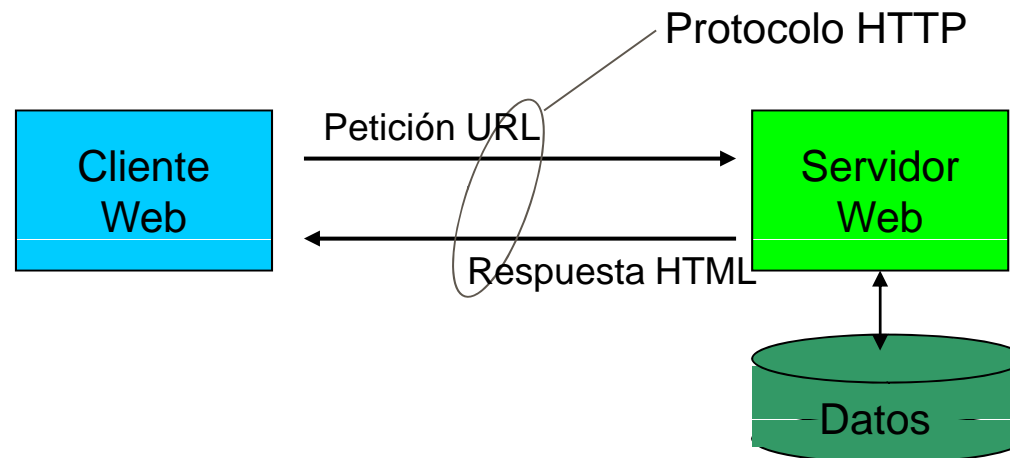
- Capa de aplicación situada sobre la capa TCP (OSI).
- Es un protocolo basado en mensajes (texto plano).
- HTTP define una secuencia cuatro pasos para cada petición de una página web:
  - Conexión: (normalmente a través del puerto 80).
  - Petición: `GET /index.html HTTP 1.0` (+ otros comandos)
  - Respuesta: código respuesta + cabecera mime + documento
  - cierre de la conexión:
- Soporta distintos métodos para la petición:
  - **GET**: Utilizado para solicitar al servidor que envíe un recurso web al cliente.
  - **POST**: Utilizado para enviar al servidor datos recogidos en el lado del cliente (formulario).



## 4. Modelo cliente-servidor.

---

- Esquema de comunicación: Cliente-Servidor.
  - **Cliente**: navegador.
  - **Servidor**: aplicación servidora HTTP.
  - **Red**: Canal de comunicación.
- Modelo de comunicación petición/respuesta:



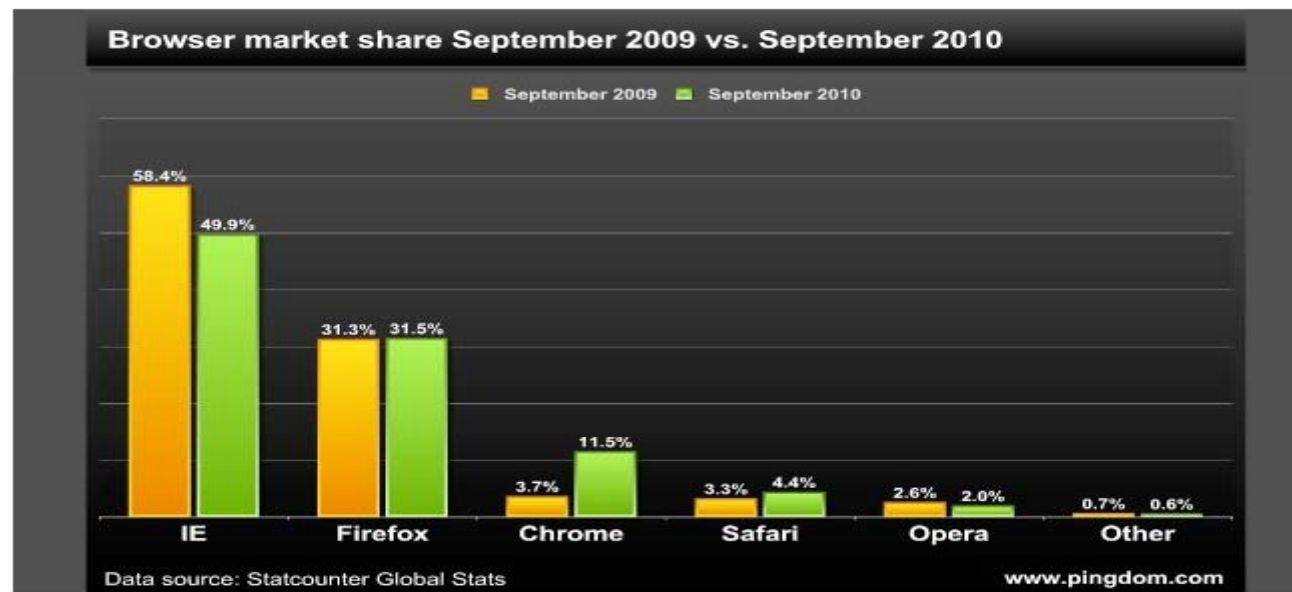
## 4. Clientes web (Navegadores).

---

- Browser o Navegador: cliente más habitual, encargado de visualizar la información.
- Canaliza las peticiones del usuario e interpreta la respuesta del servidor.
- Todos tienen funcionalidades similares (MOSAIC).
- Evolución:
  - Texto: Linemode (CERN) y Lynx.
  - Mosaic (NCSA). Aplicación de libre distribución.
  - Navigator (Netscape).
  - Internet Explorer (Microsoft).
  - Otros: Opera, Mozilla, Firefox, Konqueror, etc.

## 4. Clientes web (Navegadores) II

2010	<u>Internet Explorer</u>	<u>Firefox</u>	<u>Chrome</u>	<u>Safari</u>	<u>Opera</u>
Diciembre	27.5 %	43.5%	22.4%	3.8%	2.2%
Noviembre	28.6 %	44.0%	20.5%	4.0%	2.3%
Octubre	29.7 %	44.1%	19.2%	3.9%	2.2%
Septiembre	31.1 %	45.1%	17.3%	3.7%	2.2%
Agosto	30.7 %	45.8%	17.0%	3.5%	2.3%
Julio	30.4 %	46.4%	16.7%	3.4%	2.3%
Junio	31.0 %	46.6%	15.9%	3.6%	2.1%
Mayo	32.2 %	46.9%	14.5%	3.5%	2.2%
Abril	33.4 %	46.4%	13.6%	3.7%	2.2%
Marzo	34.9 %	46.2%	12.3%	3.7%	2.2%
Febrero	35.3 %	46.5%	11.6%	3.8%	2.1%
Enero	36.2 %	46.3%	10.8%	3.7%	2.2%



## 4. Servidores web (I).

---

- Los servidores sirven los recursos solicitados por el cliente.
- Otras tareas:
  - Recogen y procesan la información enviada por el cliente (formularios, cookies, etc.).
  - Generan dinámicamente la respuesta (PHP, ASP, JSP).
    - Se conectan a BD.
  - Se comunica con aplicaciones externas que generan dinámicamente la respuesta (CGI's).
  - Identifican al usuario.
  - Restringen el acceso a la información.
  - etc.

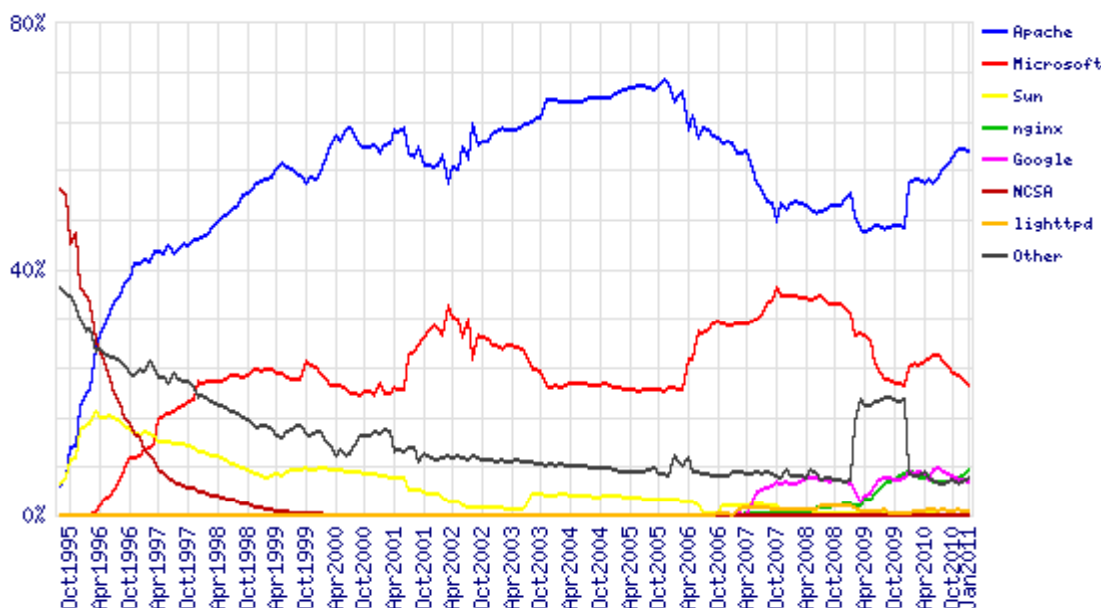
## 4. Servidores web (II).

---

- Destacan: Apache e Internet Information Server (IIS) de Microsoft.
  - Otros: Tomcat, AOLserver, SunONE, Zeus, IBM-Lotus Domino, etc.
- **Apache:** lider del mercado.
  - SO: Linux, Windows, Digital Unix, BSDI, FreeBSD, AIX, HP-UX, IRIX, Solaris, etc.
  - Ventajas: barato, potente, seguro, robusto, modular, etc.
  - Inconvenientes: instalación, mantenimiento.
- **IIS:**
  - SO: Windows NT/XP.
  - Ventajas: fácil de administrar, potente, fiable, robusto, gran número de herramientas de desarrollo, análisis, mantenimiento, administración remota, etc.,
  - Inconvenientes: SO, precio (frente a apache).

## 4. Servidores web (III).

Market Share for Top Servers Across All Domains  
August 1995 - January 2011



Developer	December 2010	Percent	January 2011	Percent	Change
Apache	151,516,152	59.35%	161,591,445	59.13%	-0.23
Microsoft	56,723,544	22.22%	57,392,351	21.00%	-1.22
nginx	16,910,205	6.62%	20,504,634	7.50%	0.88
Google	14,933,865	5.85%	15,112,532	5.53%	-0.32
lighttpd	1,308,935	0.51%	1,866,872	0.68%	0.17

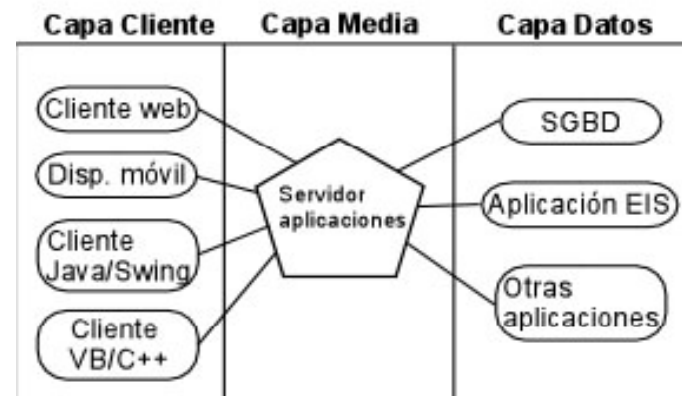
En el 2010, la lista de servidores web más utilizados, siguió siendo encabezada por Apache, seguida de Microsoft.

Sin embargo, el uso de Microsoft respecto Apache parece distanciarse de nuevo, a favor de este último.

Por otra parte, Google se mantiene en los niveles alcanzados en 2009

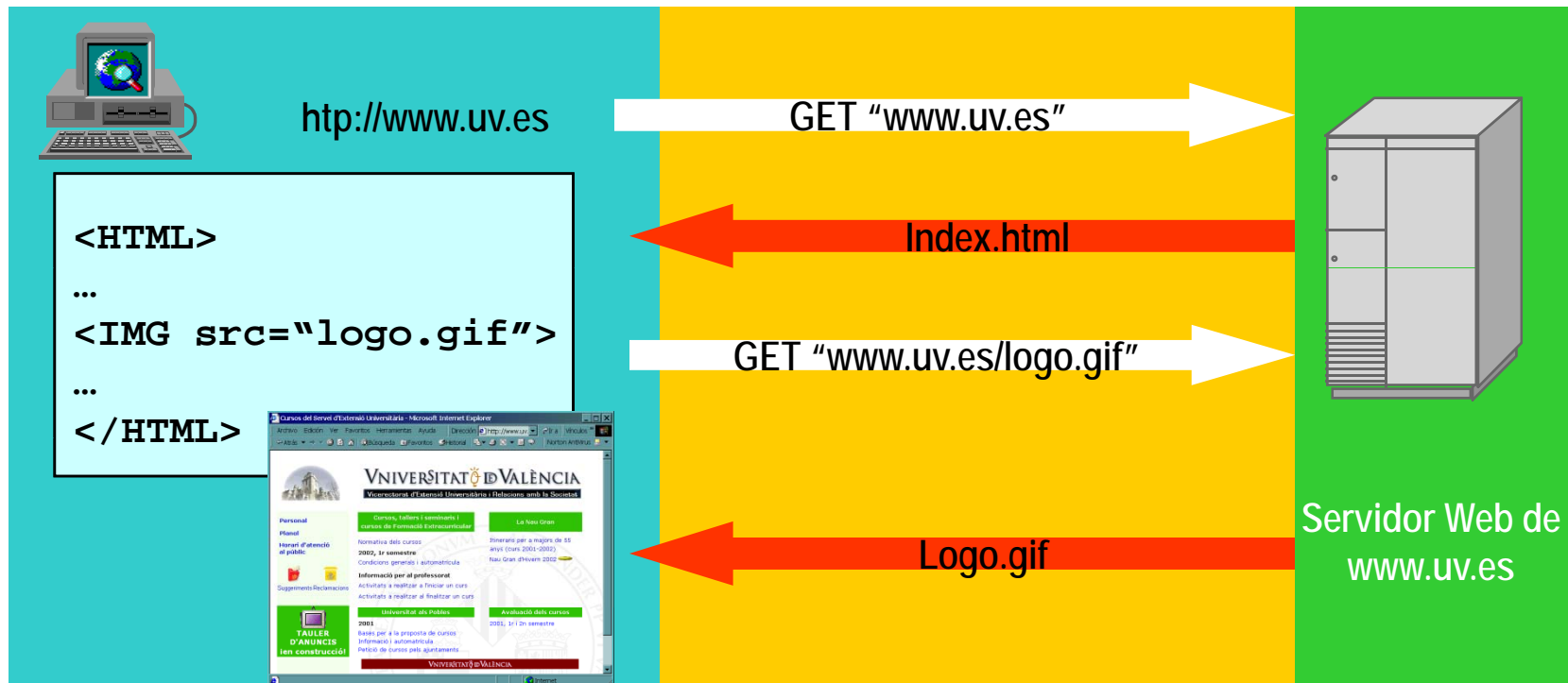
## 4. Servidores de aplicaciones

- En informática se denomina **servidor de aplicaciones** a un servidor en una red de computadores que ejecuta ciertas aplicaciones
- Usualmente se trata de un dispositivo de software que proporciona servicios de aplicación a las computadoras cliente. Un servidor de aplicaciones generalmente gestiona la mayor parte (o la totalidad) de las funciones de lógica de negocio y de acceso a los datos de la aplicación. Los principales beneficios de la aplicación de la tecnología de servidores de aplicación son la centralización y la disminución de la complejidad en el desarrollo de aplicaciones. Si bien el término es aplicable a todas las plataformas de software, hoy en día el término *servidor de aplicaciones* se ha convertido en sinónimo de la plataforma J2EE de Sun Microsystems.
- Los más utilizados
  - BEA WebLogic
  - IBM WebSphere
  - Sun-Netscape IPlanet
  - Sun One
  - Oracle IAS
  - Borland AppServer
  - HP Bluestone
  - S.A. Libres : JOnAS y JBOSS



## 5. Páginas estáticas

- Los datos se organizan en forma de páginas.
- En un principio eran estáticas (HTML + gráficos como gif, jpg).





## 5. Páginas dinámicas.

---

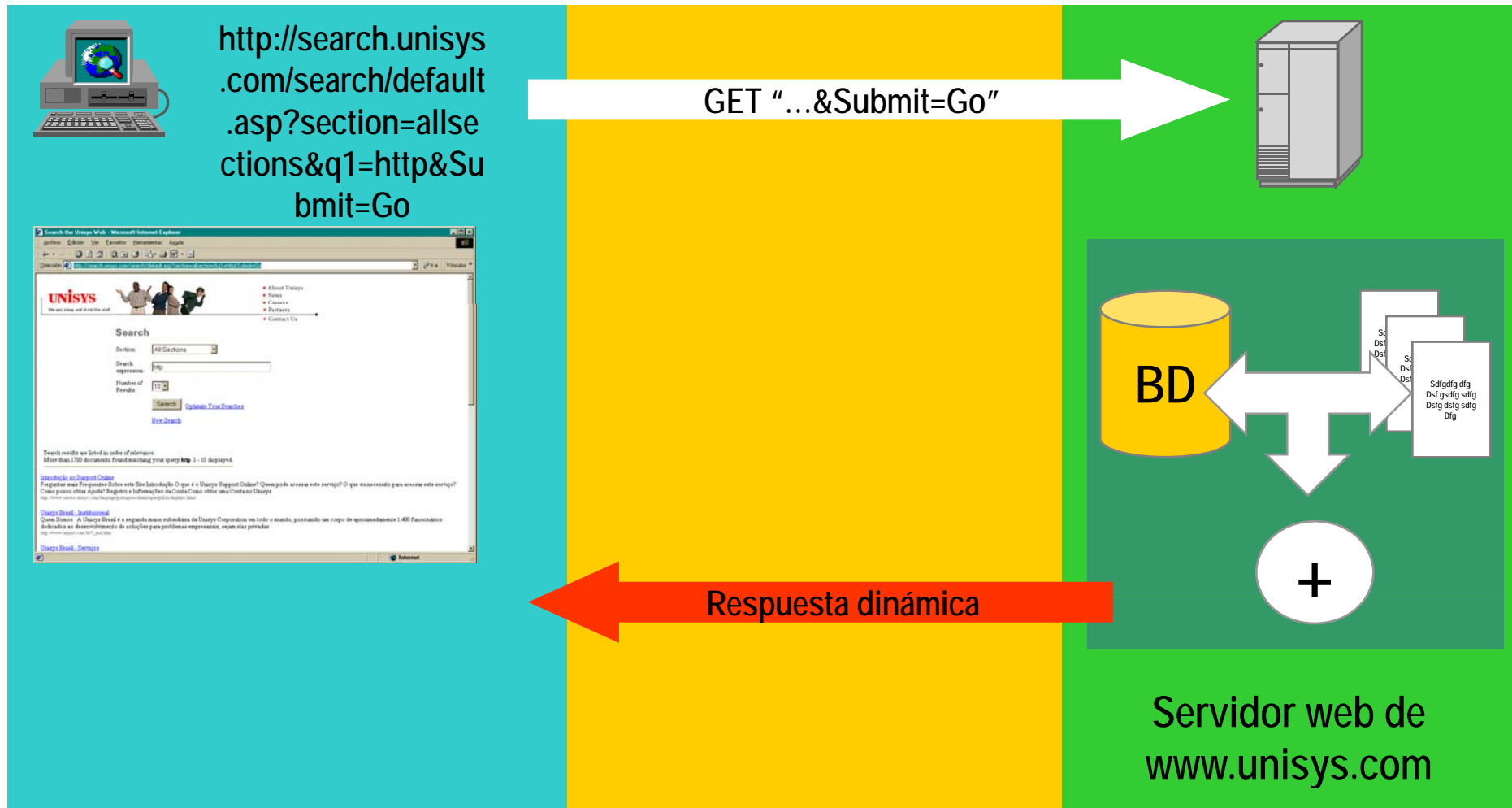
- **Más interactividad con el usuario (formularios) y dinamismo.**
  - **Lenguaje de script** en el cliente. Se integran dentro del código. Javascript y VBscript.
  - **Hojas de estilo (CSS):** Definen plantillas, características no exactamente dinámica. Combinadas con javascript (y su modelo de objetos) crean páginas dinámicas.
  - **DHTML:** páginas con contenido multimedia más rico.
    - DHTML = CSS + javascript + Modelo de objetos del documento (DOM) + Flash.
    - Problema: no es estándar.
  - **Applets.**
  - **Generación dinámica en el lado del servidor:** CGI, ASP, PHP, JSP, etc.

## 6. Generación dinámica en el servidor (I)

---

- Limitaciones de los documentos HTML estáticos:
  - No pueden adaptar los documentos a clientes individuales.
  - Hay problemas para actualizar los documentos, especialmente si parte de los datos están replicados en varios documentos.
- Solución: Generar los documentos Web dinámicamente en el lado del servidor, extrayendo la información de una Base de Datos en el momento de la petición.
  - Se pueden seleccionar la información en función del usuario.
  - La información presentada está actualizada.

## 6. Generación dinámica en el servidor (II)



## 6. Generación dinámica en el servidor (CGI)

---

- Common Gateway Interface (CGI): estándar que define como interactúan los servidores Web y programas independientes.
- Comunicación con el servidor en los dos sentidos. Estos programas reciben los datos pasados por el cliente (pe: a partir de un formulario), y devuelven una cabecera texto más código HTML.
  - Cada petición genera un nuevo proceso corriendo de forma independiente.
- La interfaz es independiente de la aplicación, la cual puede estar programada en: PERL, C, C++, C Shell, Python, Java, etc.
- Estos programas actúan de interfaz con otras aplicaciones (pe: BD).

## 6. Generación dinámica en el servidor (PHP).

---

- PHP: Lenguaje de scripts inmerso en el código HTML. El módulo PHP analiza cada instrucción dentro de la página. Devuelve sólo código HTML.
  - Estilo de programación similar a C++ y PERL.
- Código fuente abierto y libre distribución (muy barato).
- Soportado por la gran mayoría de servidores (sobre todo Apache).
- Soporte y gran flexibilidad para acceder a BD: mSQL, MySQL, PostgreSQL, Oracle, Informix, etc.
- Ventajas sobre CGIs: está integrado perfectamente en el servidor Web -> no ejecuta una copia del programa por cada petición.

## 6. Generación dinámica en el servidor (ASP).

---

- ASP y ASP.NET: Alternativa Microsoft a PHP.
- Separa el lenguaje de script del modelo de objetos.
- Acceso a base de datos (Oracle, SQL server, Access, etc.) a través de ODBC u OLEDB.
- Ventajas frente PHP:
  - Microsoft da soporte con ASP de herramientas de apoyo, depuradores, utilidades, etc.
  - Flexible (varios lenguajes de programación).
- Desventajas:
  - Menos abierto (sólo soportado por IIS).

## 7.- La Web de 3ª generación

---

Limitaciones de la web de 2ª generación:

- Son los usuarios quien interpretan en última instancia la información, no los programas.
  - Interacción con una aplicación web dirigida por un usuario.
  - Búsqueda de información supervisada por un usuario.
- No permite la transferencia automática de información entre aplicaciones web (servicios web) distintas.
- Las aplicaciones no saben interpretar la información.

Tecnologías de 3ª generación:

- Basadas en XML.
- Servicios Web
- Web semántica

## 7. La web de 3ª generación (Servicios Web I)

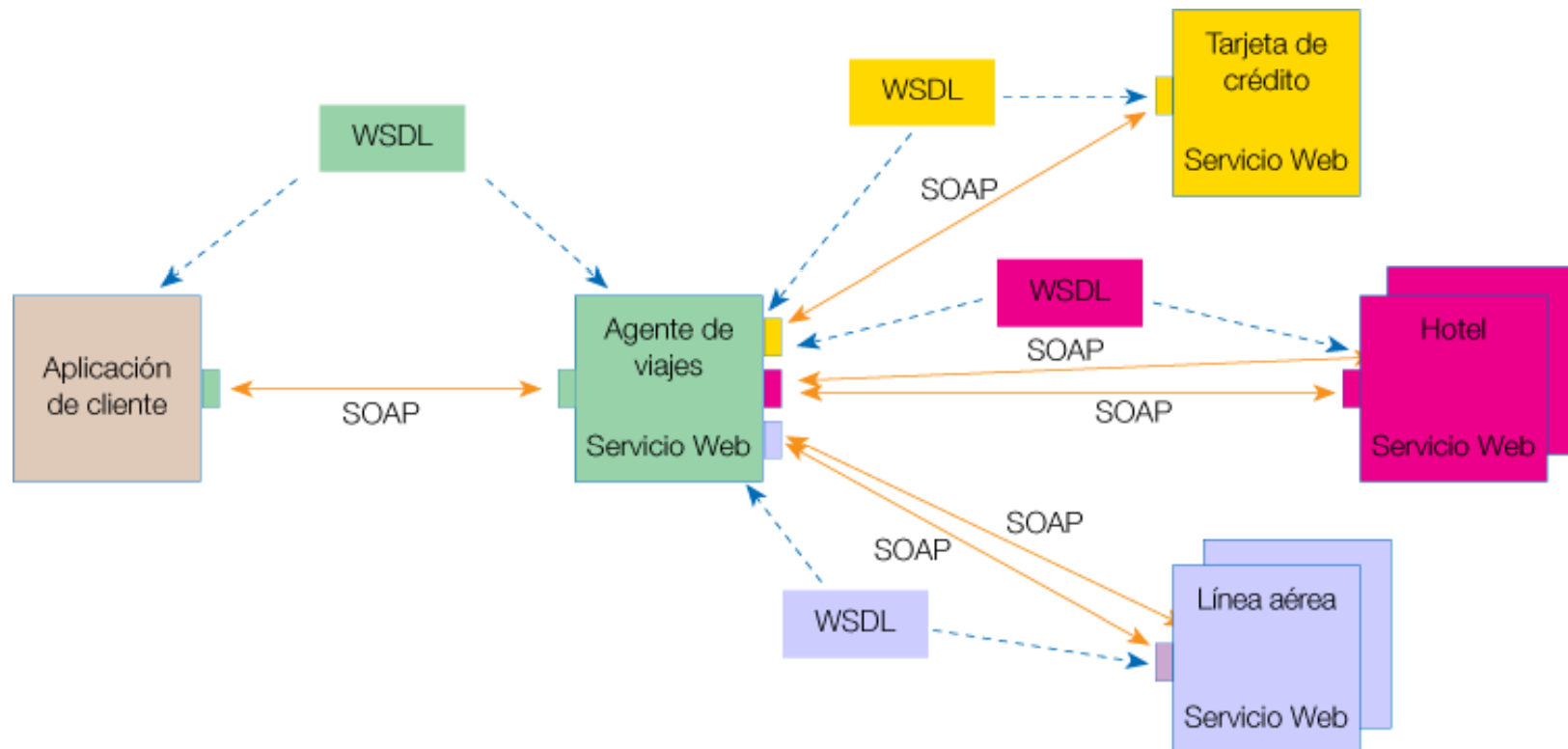
---

- Un **servicio Web** es una colección de protocolos y estándares que sirve para intercambiar datos entre aplicaciones.
- Características:
  - Uso de formato estándar, XML, para intercambiar datos.
    - Protocolo SOAP.
    - Transportado sobre HTTP.
  - Lenguaje de interfaz pública para los servicios web → WSDL
    - Descripción XML de cómo se establece la comunicación.
  - Comprobación de disponibilidad de servicios:
    - Uso del protocolo UDDI para publicarlos.



## 7. La web de 3ª generación (Servicios Web II)

---



## 7. La web de 3ª generación (Web Semántica)

---

- HTML o XHTML son incapaces de describir el contenido ( su significado o semántica).
- La **web semántica** se centra en el contenido.
  - Proporciona nuevos lenguajes (RDF, OWL, etc.) para representar la información publicada en Internet (W3C).
- **Objetivos:**
  - Mejorar la organización de la información (por contenido)
  - Automatizar la recopilación de información y el filtrado de contenidos (agentes inteligentes)
    - Mejorar de forma drástica la búsqueda y recuperación de información.
- Adaptan técnicas de Representación del Conocimiento (utilizadas en IA) a las características de la web como sistema de información.