

Lenguajes de Marcas

Contenido:

Introducción	2
Historia del lenguaje de marcas HTML	3
Editores HTML	6
Versiones de HTML	6
Arquitectura Servidor-Cliente	8
Servidores WEB	10
Navegadores WEB	12



Unidad 2– Lenguaje HTML

HTML, siglas de HyperText Markup Language («lenguaje de marcas de hipertexto»), hace referencia al lenguaje de marcado para la elaboración de páginas web. Es un estándar que sirve de referencia para la elaboración de páginas web en sus diferentes versiones, define una estructura básica y un código (denominado código HTML) para la definición de contenido de una página web, como texto, imágenes, etc. Es un estándar a cargo de la W3C, organización dedicada a la estandarización de casi todas las tecno-

logías ligadas a la web, sobre todo en lo referente a su escritura e interpretación. Es el lenguaje con el que se definen las páginas web.

El HTML se escribe en forma de «etiquetas», rodeadas por corchetes angulares (<,>). HTML también puede describir, hasta un cierto punto, la apariencia de un documento, y puede incluir o hacer referencia a un tipo de programa llamado script, el cual puede afectar el comportamiento de navegadores web y otros procesadores de HTML..

1.- Introducción

Como se comentó en el primer tema: a finales de los 80 se desarrolló el lenguaje de marcas SGML. En esa misma época Tim Bernes Lee utilizó SGML para definir un nuevo lenguaje de etiquetas que llamó Hypertext Markup Language (lenguaje de marcado de hipertexto) para crear documentos transportables a través de Internet en los que fuera posible el hipertexto; es decir la posibilidad que determinadas palabras marcadas de forma especial permitieran abrir un documento relacionado con ellas.

A pesar de tardar en ser aceptado, HTML fue un éxito rotundo y la causa indudable del éxito de Internet. Hoy en día casi todo en Internet se ve a través de documentos HTML, que popularmente se denominan páginas web.

Inicialmente estos documentos se veían con ayuda de intérpretes de texto (como por ejemplo el Lynx de Unix) que simplemente coloreaban el texto y remarcaban el hipertexto. Después el software se mejoró y aparecieron navegadores con capacidad más gráfica para mostrar formatos más avanzados y visuales.

Lógicamente desde 1989 hasta nuestros días HTML ha mejorado. Entre sus avances fundamentales:

- El lenguaje cada vez ha ido incorporando nuevas etiquetas más potentes, que permiten incluir en los documentos HTML, tablas, capas, marcos, imágenes, ...
- Se han añadido lenguajes de script (como JavaScript) con código incrustado en las páginas HTML que permiten añadir funcionalidades y dinamismo a las páginas web
- Se han añadido técnicas en el lado del servidor con la misma finalidad como aplicaciones CGI, PHP, ASP o JSP.
- Se incorporaron lenguajes de estilo (como CSS) para generar un formato de documento más avanzado
- Se han añadido utilidades para gestión avanzada de JavaScript con XML (AJAX)

- Se ha permitido la inclusión de elementos avanzados en las páginas como Flash o los applets de Java para dar mayor funcionalidad.
- Se permiten elementos semánticos para dar mayor significado al contenido
- Se permite el dibujo libre de elementos en la página mediante el elemento canvas.
- Es posible añadir vídeo, audio y otros elementos multimedia de forma fácil

En la actualidad HTML sigue siendo el lenguaje fundamental de las páginas web; pero ahora Internet es la web, es decir todo en Internet se ve a través de una página web. Por eso hoy en día HTML es la capa superficial bajo la que se agolpan tecnologías muy diversas y muy distintas de HTML.



2.- Historia del lenguaje de marcas HTML

La historia completa de HTML es tan interesante como larga, por lo que a continuación se muestra su historia resumida a partir de la información que se puede encontrar en la Wikipedia.

El origen de HTML se remonta a 1980, cuando el físico Tim Berners-Lee, trabajador del CERN (Organización Europea para la Investigación Nuclear) propuso un nuevo sistema de "hipertexto" para compartir documentos.

Los sistemas de "hipertexto" habían sido desarrollados años antes. En el ámbito de la informática, el "hipertexto" permitía que los usuarios accedieran a la información relacionada con los documentos electrónicos que estaban visualizando. De cierta manera, los primitivos sistemas de "hipertexto" podrían asimilarse a los enlaces de las páginas web actuales.

Tras finalizar el desarrollo de su sistema de "hipertexto", Tim Berners-Lee lo presentó a una convocatoria organizada para desarrollar un sistema de "hipertexto" para Internet. Después de unir sus fuerzas con el ingeniero de sistemas Robert Cailliau, presentaron la propuesta ganadora llamada WorldWideWeb (W3).

El primer documento formal con la descripción de HTML se publicó en 1991 bajo el nombre HTML Tags (Etiquetas HTML) y todavía hoy puede ser consultado online a modo de reliquia informática.

EN 1992 la NCSA3 se interesa por la ya llamada web de Bernes Lee y crea el primer navegador: Mosaic. Entre sus creadores está Eric Bina y Marc Adreessen futuros millonarios gracias a la web. Lou Montulli desarrolla Lynx para los sistemas Unix, el primer navegador de texto en la web. Será ampliamente utilizado en los años siguientes, aunque luego quedará rápidamente superado por las capacidades de los navegadores gráficos

En 1993 se lanza comercialmente Mosaic por parte de la NCSA. Añade nuevos elementos al HTML original como por ejemplo la inclusión de imágenes.

El lenguaje HTML empieza a ser caótico porque aparecen

numerosas etiquetas puestas por cada entidad privada. Dan Connolly recopila las etiquetas HTML de la época más utilizada y se crea el borrador de HTML 2.

Marc Adreessen y Jim Clark abandona la NCSA y fundan Mosaic Communications (futura Netscape). Dejan también los estándares y crean elementos nuevos en el lenguaje HTML para crear páginas más vistosas para su navegador.

A finales del año se crea la World Wide Consortium (W3C) fichando a algunos de los principales impulsores de la web (incluido Tim Bernes Lee). Se convertirá en el principal organismo de estandarización de las tecnologías relacionadas con la web en general y de HTML en particular .

A En 1995 aparecen los primeros elementos de creación de hojas de estilo, raíz del lenguaje CSS que permite dar formato avanzado a las páginas web y que sigue siendo una de las tecnologías imprescindibles en la actualidad para crear páginas web. Sun Microsystems crea el lenguaje Java, que tendrá una enorme influencia en el desarrollo de Internet.

Los hermanos Allaire, crean ColdFussion, un lenguaje basado en HTML que se ejecuta en el servidor que aloja las páginas web (servidores compatibles con esta tecnología) de modo que el cliente no necesita tener un software especial que reconozca esta tecnología. Al cliente le llegan páginas web normales que ha preparado el servidor tras traducir este lenguaje. Fue la primera tecnología de script en el lado del servidor.

EN 1996 se crea Yahoo! la primera página exitosa que permite organizar la web para facilitar la búsqueda de otras páginas. La empresa será una de las de mayor valor en bolsa durante varios años. Netscape desarrolla JavaScript, un lenguaje basado en C y Java que se incrusta dentro del código HTML de las páginas para darles una mayor potencia. Todavía sigue siendo uno de los lenguajes más influyentes en el desarrollo de páginas y

aplicaciones para la web. En definitiva es una tecnología de script en el lado del cliente. Los navegadores, debido a su éxito, tuvieron que ir poco a poco incorporando plugins (software añadido) que permitieran ejecutar código JavaScript.

Rasmus Lerdorf crea el lenguaje PHP. Todavía sigue siendo la tecnología del lado del servidor más popular.

La empresa Macromedia crea el software Flash. Se trata de una tecnología del lado del cliente (requiere un plugin en el navegador) que permitió que las páginas incluyeran todo tipo de elementos multimedia e interactivos que hacían de las páginas aplicaciones ricas semejantes a las aplicaciones de un escritorio de ordenador personal.

Hakom Wum Lie crea el navegador Opera. Nunca ha alcanzado una gran cuota de público pero sigue presente después de todos esos años.

En 1997 aparece la especificación estándar HTML 3.2, la primera en ser ampliamente aceptada. Incluye tablas, applets (pensadas para añadir elementos Java a las páginas) y otros formatos avanzados de formato. Sun Microsystems crea Java Servlets y Microsoft crea el lenguaje ASP. Son dos de las tecnologías del lado del servidor que tendrán una gran influencia en los años siguientes.



La W3C lanza en 1998 como estándares a HTML 4.0 y a CSS2. Los estándares de la W3C cada vez se tienen más en cuenta y ambos alcanzan un gran éxito. La combinación HTML+JavaScript+CSS se conoce este año como DHTML (HTML dinámico). Alcanzará un enorme notoriedad y será la combinación habitual para hacer páginas web atractivas.

Aparece XML 1.0 por parte de la W3C, como el lenguaje que debió ser HTML (en palabras del propio Tim Bernes Lee). No ha llegado a suplantarse a HTML pero sigue teniendo una enorme influencia en todo tipo de tecnologías.

En 1999 Sun crea JSP (páginas de servidor en lenguaje Java) y la plataforma de trabajo J2EE (Java Enterprise) con lo que pretende crear un entorno poderoso de trabajo para crear aplicaciones y servicios de Internet en los servidores.

El navegador Internet Explorer de Microsoft domina el mercado poniendo fin a la primera guerra de navegadores.

Aparece RSS un formato de contenido basado en XML que permite sindicarse y obtener información de forma veloz. La W3C presenta HTML 4.01, indicando que será la última versión del HTML clásico.

En 2002 la fundación Mozilla recoge el testigo de Netscape y crea el navegador Firefox.

Microsoft crea la plataforma de aplicaciones .NET con vocación de competir con J2EE. Apple lanza al mercado el navegador Safari en 2003.

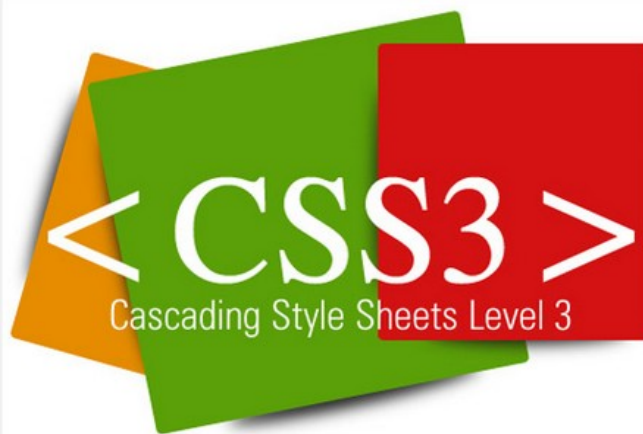
Se crea Wordpress el primer gestor de contenidos web (CMS). Permite crear páginas web (especialmente blogs) fácilmente y gestionar a diferentes usuarios que podrán editar contenidos de la web fácilmente.

En 2004 se comercializa el navegador Firefox, comienza la segunda guerra de navegadores. En 2005 aparece AJAX, tecnología que combina JavaScript, HTML, CSS y XML se populariza pasando a ser una de las tecnologías fundamentales para crear páginas web dinámicas.

Se lanza Joomla! se convertirá en el CMS más popular (aunque actualmente parece que Wordpress le está superando).

En 2008 aparece el primer borrador de HTML 5. Se acepta HTML 5 y Flash empieza a dejar de utilizarse (aunque sigue siendo muy influyente)

Aparecen numerosos borradores de CSS3



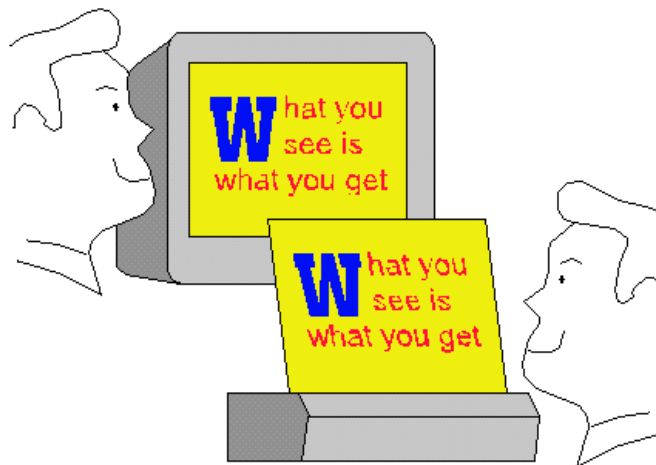
3.- Editores HTML

Para escribir HTML bastaría con un editor de texto plano como el bloc de notas de Windows o gedit de Linux. Pero los resultados se deben mostrar mediante un navegador o aún mejor, probar en varios navegadores para comprobar problemas (ya que hay elementos HTML que no son compatibles con todos los navegadores).

Sin embargo lo ideal es trabajar mediante editores de código capaces de entender el lenguaje y colorear de diferente manera las etiquetas HTML para distinguirlas del texto normal y así trabajar mejor. Algunos populares son Notepad++ o Sublime Text.

Los entornos de edición en XML como Oxygene también permiten trabajar en HTML y de hecho en el caso de utilizar XHTML son una gran opción.

Otra forma es trabajar con editores WYSIWYG (What You See Is What You Get, lo que se ve es lo que se obtiene), en los que se edita el documento al estilo de los procesadores de texto. De modo que se ve el resultado sin tocar el código y es el editor el que traduce el resultado al código correspondiente. Aunque en realidad no es posible un WYSIWYG real ya que los navegadores muestran el resultado de diferente forma.



4.- Versiones de HTML

Los programas capaces de traducir el código HTML y producir una salida en pantalla de los mismos son los navegadores (browsers en inglés). Se trata de un software gráfico que se inició con la creación de Mosaic a principios de los 90 y que poco a poco produjo más productos hasta llegar a una guerra de navegadores a finales de los 90 entre Internet Explorer de Microsoft y Navigator de Netscape que ganó Microsoft pero que ahora continúa con otros navegadores como el propio Internet Explorer, Mozilla Firefox (sucesor de código abierto de Netscape), Google Chrome, Opera o Apple Safari entre otros.

El problema surgió en cuanto unos navegadores incorporaron elementos HTML que el resto no traducía, con lo que aparecieron diferentes dialectos HTML. Así una página se podía mostrar de forma totalmente diferente según el navegador.

La solución pasó por intentar estandarizar el lenguaje. Por ello el propio Tim Bernes Lee fundó la World Wide Web Consortium (abreviado W3C) como organismo de estandarización del lenguaje HTML ante la industria.

En la actualidad las directrices de W3C son seguidas por la mayoría de navegadores aunque no al 100%, lo que sigue generando problemas a los creadores de páginas web.

La situación, sin embargo se ha complicado en estos últimos años con la aparición de diferentes estándares, en concreto actualmente se consideran estándares a HTML 4.01, XHTML 1.0 y 1.1 y ya se considera de la misma forma a HTML 5. Las diferencias entre ellos son:

HTML 4.01. Se trata de la versión estándar del HTML tradicional hecha en el año 1999 y que sigue teniendo mucha vigencia actualmente. Hay tres versiones: la transicional (que permite seguir usando algunas etiquetas que se consideran obsoletas), la estricta (que elimina numerosas etiquetas y atributos para forzar a crear un HTML con menos formato y más significado) y la frameset orientada a usar los ya muy poco utilizados marcos. La versión más popular es la transicional al ser más libre.

Para avisar de que nuestro documento sigue las normas de HTML 4.01, se coloca una etiqueta DOCTYPE que permite llegar al DTD estándar de esta versión. Para HTML 4.01 transicional es:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN" "http://www.w3.org/TR/html4/loose.dtd">
```

HTML 4 no es compatible con XML y por ello a los antiguos diseñadores web les gusta más; pero precisamente el no ser compatible con XML no permite utilizar muchas aplicaciones creadas para ese lenguaje XHTML.

XHTML era planteado como el sustituto de HTML, la primera versión, la 1.0 sigue siendo la más usada y en la fecha de escritura de este texto es la versión HTML más utilizada. Hay también versión estricta, transicional y (muy poco utilizada) frameset; la más utilizada, otra vez es la transicional porque permite el uso de numerosos atributos y elementos que se consideran obsoletos, pero que muchos diseñadores utilizan.

Las páginas XHTML obligan a que la escritura de HTML siga las reglas del lenguaje XML bien formado.

Para indicar que utilizamos XHTML 1.0 transicional se debe poner este DOCTYPE:

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
```

Además la etiqueta raíz html debe usarse de este modo (especificando el espacio de nombres y el lenguaje de la página):

```
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="es">
```

En el caso de XHTML 1.1 el DOCTYPE sería:

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
```

Se puede observar que es más fiel al lenguaje XML. La etiqueta raíz html sería igual que la anterior

HTML 5. Se considera el estándar de este año 2012. Poco a poco se está convirtiendo en el HTML más utilizado. Se creó fuera del World Wide Web Consortium en la entidad llamada WHATWG (Web Hypertext Application Technology Working Group), auspiciado por Apple y Mozilla entre otras empresas a las que se le han ido añadiendo muchas más.

La W3C decidió unirse a la WHATWG en 2008 para buscar un estándar. Actualmente el proyecto está en fase de borrador planificándose su paso a recomendación oficial en 2014. La WHATWG va por su lado mediante una filosofía de "estándar vivo" en la que ya hay una recomendación de uso por su parte, pero que va cambiando continuamente según se consideren nuevas mejoras.

Aunque no es una recomendación oficial y por ahora los navegadores aceptan sólo partes de HTML 5 (algunos casi todas, pero otros muy pocas) el hecho es que ya se anima a usarlo y la mayoría de páginas nuevas son HTML5. Para ello marcan un DOCTYPE mucho más sencillo:

```
<!DOCTYPE html>  
<html lang="es">
```

La etiqueta raíz HTML también es más sencilla al no indicar un espacio de nombres.

XHTML 5. Se trata de una especificación en proceso. Es simplemente HTML 5 pero obligado a usar las reglas de XML bien formado. Es decir es HTML 5 siguiendo lo estricto de XML.

5.- Arquitectura Servidor-Cliente

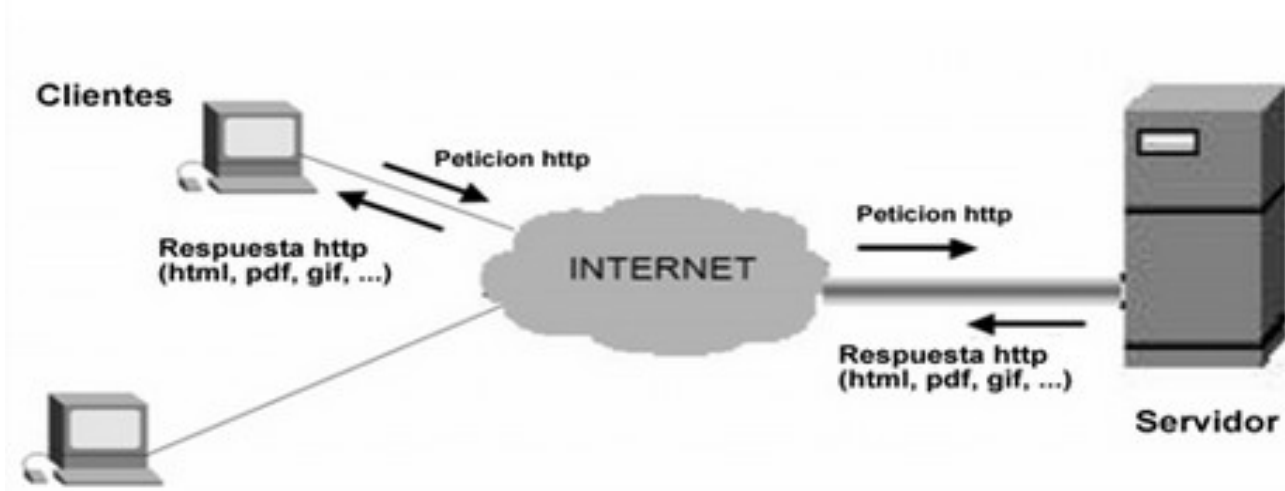
Comenzaremos por definir lo que es un servidor. Como su nombre indica, son los dispositivos de red que brindan un servicio a otros dispositivos, a los cuales se les conoce como clientes. En general quien realiza esta tarea es un software especializado, pero comúnmente se conoce como servidor al equipo físico donde se ejecuta, el cual es el centro de la infraestructura de la red.

Los servidores están integrados en una arquitectura servidor-cliente. El servidor es un ordenador que tiene instalado algún tipo de servicio como servicio web, ftp, archivos, DNS, etc. El cliente es ordenador conectado al servidor y que solicita los servicios que éste proporciona.

Un mismo ordenador puede servir varios servicios distintos. Cada servicio tiene asignado un protocolo de comunicación y un puerto. Así, por ejemplo, el servidor web tiene asignado por defecto el puerto 80 y utiliza el protocolo HTTP para comunicarse con el cliente.

Puerto	Desc.	Estado	Observaciones
20	FTP	cerrado	Utilizado por FTP
21	FTP	cerrado	Utilizado por FTP
22	SSH	cerrado	Secure Shell.
23	TELNET	cerrado	Acceso remoto
25	SMTP	cerrado	Servidor de correo SMTP
53	DNS	cerrado	Servidor DNS
79	FINGER	cerrado	Servidor de información de usuarios de un PC
80	HTTP	cerrado	Servidor web
110	POP3	cerrado	Servidor de correo POP3
119	NNTP	cerrado	Servidor de noticias

Modelo Cliente/Servidor



DENOMINACIÓN DEL SERVIDOR	DESCRIPCIÓN
Servidor de Correo	Es el servidor que almacena, envía, recibe y realiza todas las operaciones relacionadas con el e-mail de sus clientes.
Servidor Proxy	Es el servidor que actúa de intermediario de forma que el servidor que recibe una petición no conoce quién es el cliente que verdaderamente está detrás de esa petición.
Servidor Web	Almacena principalmente documentos HTML (son documentos a modo de archivos con un formato especial para la visualización de páginas web en los navegadores de los clientes), imágenes, videos, texto, presentaciones, y en general todo tipo de información. Además se encarga de enviar estas informaciones a los clientes.
Servidor de Base de Datos	Da servicios de almacenamiento y gestión de bases de datos a sus clientes. Una base de datos es un sistema que nos permite almacenar grandes cantidades de información. Por ejemplo, todos los datos de los clientes de un banco y sus movimientos en las cuentas.
Servidores DHCP	Son servidores especializados en asignar direcciones IP de forma dinámica.
Servidores DNS	Un servidor DNS o de nombres de dominio, es como un directorio telefónico, solo que este en vez de tener números de teléfono y sus respectivos titulares, tiene nombres de dominio y sus respectivas direcciones IP. Es más fácil que recuerdes www.google.com , que 74.125.227.19.
Servidores de imágenes	Recientemente también se han popularizado servidores especializados en imágenes, permitiendo alojar gran cantidad de imágenes sin consumir recursos de nuestro servidor web en almacenamiento o para almacenar fotografías personales, profesionales, etc.

6.- Servidores WEB

Los **servidores web** son aquellos cuya tarea es alojar sitios y/o aplicaciones, las cuales son accedidas por los clientes utilizando un navegador que se comunica con el servidor utilizando el protocolo HTTP.

Básicamente un servidor WEB consta de un intérprete HTTP el cual se mantiene a la espera de peticiones de clientes y le responde con el contenido según sea solicitado. El cliente, una vez recibido el código, lo interpreta y lo exhibe en pantalla.

Además los servidores pueden disponer de un intérprete de otros lenguajes de programación que ejecutan código embebido dentro del código HTML de las páginas que contiene el sitio antes de enviar el resultado al cliente. Esto se conoce como programación de lado del servidor y utiliza lenguajes como ASP, PHP, Perl y Ajax. Las ventajas de utilizar estos lenguajes radica en la potencia de los mismos ejecutando tareas más complejas como, por ejemplo acceder a bases de datos abstrayendo al cliente de toda la operación.

Originalmente la Web únicamente contaba con contenidos estáticos. La necesidad de ofrecer servicios más sofisticados a través de este medio llevó a soluciones tecnológicas que permitiesen cierto grado de interacción con el servidor más allá de solicitar una página en concreto. Los CGI fueron las primeras opciones en este sentido. Permitían interactuar con código ejecutable en el servidor desde el navegador. A través de la URL solicitada por el navegador se pasaban los parámetros de entrada para el código ejecutable y este devolvía una página HTML con la respuesta. Este mismo método se utiliza en tecnologías similares como ISAPI, NSAPI o Java Servlets.

A medida que la sofisticación de los servicios ofrecidos vía Web ha ido aumentando, se han ido añadiendo, al menos a nivel lógico y no necesariamente a nivel físico, nuevos elementos en la arquitectura de las aplicaciones Web (por ejemplo el servidor de aplicaciones, bases de datos,...).

Al aumentar la complejidad de los procesos se acaba con el denominado problema del "Cliente Pesado". Al poner la mayor parte del código necesario para llevar a cabo los procesos en el cliente, este debe descargar del servidor los datos necesarios para llevarlos a cabo. Esto es muy ineficiente por dos razones principales:

- La red sufre una gran carga debido a las múltiples descargas de cada uno de los clientes.
- La gran dependencia en el rendimiento del hardware en el lado del cliente.

Las arquitectura en tres niveles divide la funcionalidad para optimizar el uso de recursos. Se consiguen soluciones mucho más flexibles y escalables. Los tres niveles son:

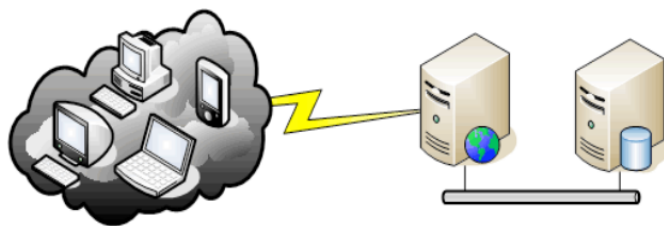
Cliente. Contiene los componentes de usuario que son únicos para cada uno de ellos. Esto es la lógica de aplicación específica del usuario y la interfaz.

Aplicación. Constituye un entorno multiusuario y mantiene las partes compartidas de la aplicación. Las operaciones con un uso intensivo de datos deben ejecutarse en este nivel. También es el punto donde se puede llevar a cabo la coordinación de transacciones (operaciones de múltiples usuarios).

Almacenamiento. Es nivel de la base de datos. Se especializa en dar un servicio de persistencia a los datos de la aplicación y permite manejar grandes volúmenes de ellos.



Arquitectura básica Cliente/Servidor



Arquitectura en tres capas: cliente, aplicación y almacenamiento

Al mantener buena parte del código en el nivel intermedio, la aplicación se aísla de la interfaz de usuario y de la base de datos. Por lo tanto, se pueden hacer cambios en el código en el nivel de aplicación sin que esto afecte al resto. Además, los datos están centralizados y fácilmente accesibles sin necesidad de moverlos completamente hasta el cliente.

Este modelo se utiliza ampliamente en el entorno Web. Se incorpora la figura del servidor de aplicaciones que permite que el sistema gestione la lógica de negocio y el estado. Al ser una aplicación Web, la interfaz con el cliente sigue siendo básicamente HTTP y por lo tanto el servidor de aplicaciones incorpora un servidor Web o utiliza uno externo.

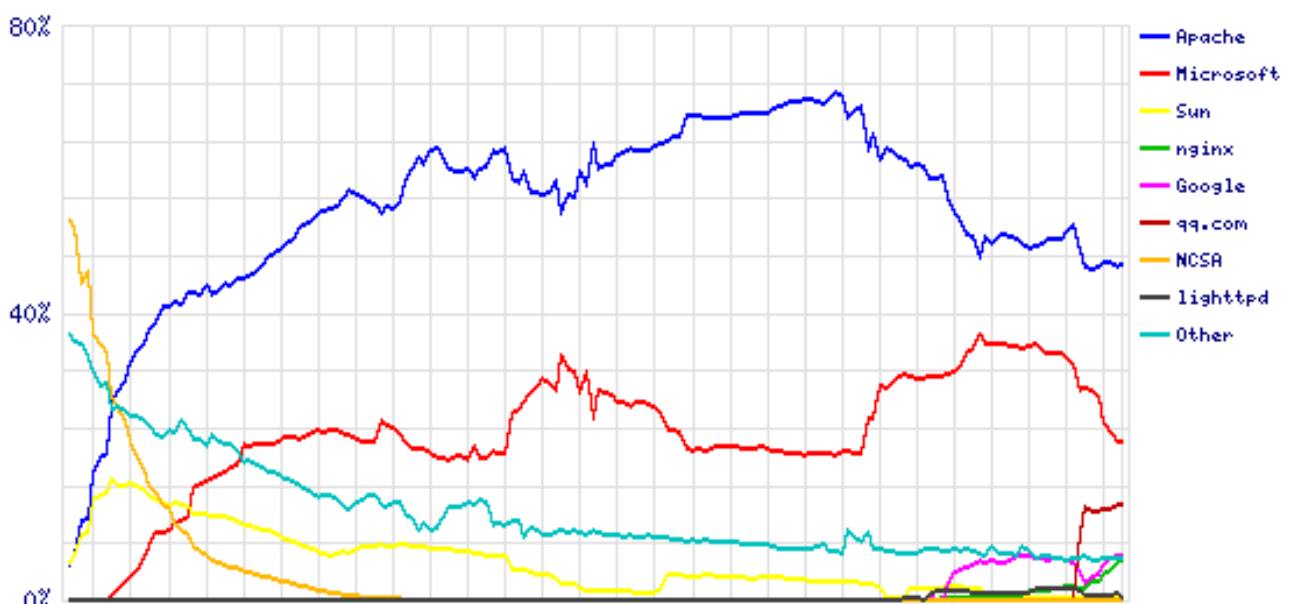
La división de la aplicación en niveles supone la necesidad de establecer interfaces entre ellas. Si además se quiere construir una arquitectura independiente de los fabricantes la mejor opción es utilizar elementos que implementen interfaces estándar.

Para la interfaz entre el servidor de aplicaciones y la base de datos las opciones son múltiples, dependiendo del lenguaje de programación (Java, C, PHP, VisualBasic,...), el tipo de base de datos (relacional, XML,...), la base de datos concreta MySQL, Oracle, eXist,...), etc. En cualquier caso con cada configuración hay más de una opción y lo normal es plantearse la elección tras definir la BD o el lenguaje de programación a utilizar.

Por lo que respecta a la interfaz entre el cliente y el servidor de aplicaciones, la opción básica sigue siendo HTTP.

La grafica siguiente muestra el resultado de la encuesta realizada por Netcraft en donde se muestra los porcentajes de los servidores web usados en Internet. La encuesta fue realizada a 226,099,841 sitios web.

Observamos que más del 40% de los servidores en el mundo son Apache, después de sigue el servidor IIS de Microsoft con más del 20% y el resto se compone de otros servidores web como nginx, google, lighttpd entre otros.



7.- Navegadores WEB

Un navegador o navegador web, o browser, es un software que permite el acceso a Internet, interpretando la información de archivos y sitios web para que éstos puedan ser leídos.

La funcionalidad básica de un navegador web es permitir la visualización de documentos de texto, posiblemente con recursos multimedia incrustados. Además, permite visitar páginas web y hacer actividades en ella, es decir, podemos enlazar un sitio con otro, imprimir, enviar y recibir correo, entre otras funcionalidades más.

Los documentos que se muestran en un browser pueden estar ubicados en la computadora en donde está el usuario, pero también pueden estar en cualquier otro dispositivo que esté conectado en la computadora del usuario o a través de Internet, y que tenga los recursos necesarios para la transmisión de los documentos (un software servidor web).

Tales documentos, comúnmente denominados páginas web, poseen hipervínculos que enlazan una porción de texto o una imagen a otro documento, normalmente relacionado con el texto o la imagen.

El seguimiento de enlaces de una página a otra, ubicada en cualquier computadora conectada a Internet, se llama navegación, de donde se origina el nombre navegador (aplicado tanto para el programa como para la persona que lo utiliza, a la cual también se le llama cibernauta). Por otro lado, hojeador es una traducción literal del original en inglés, browser, aunque su uso es minoritario.

El primer navegador, desarrollado en la CERN a finales de 1990 y principios de 1991 por Tim Berners-Lee, era bastante sofisticado y gráfico, pero sólo funcionaba en estaciones NeXT.

El navegador Mosaic, que funcionaba inicialmente en entornos Unix sobre X11, fue el primero que se extendió debido a que pronto el NCSA preparó versiones para Windows y Macintosh. Sin embargo, poco más tarde entró en el mercado Netscape Navigator que rápidamente superó en capacidades y velocidad a Mosaic. Este navegador tuvo la ventaja de funcionar en casi todos los Unix, así como en entornos Windows.

Internet Explorer (anteriormente Spyglass Mosaic) fue la apuesta tardía de Microsoft para entrar en el mercado y consiguió desbancar al Netscape Navigator entre los

usuarios de Windows, debido a la integración del navegador con el sistema operativo, llegando a poseer cerca del 95% de la cuota de mercado. Netscape Communications Corporation liberó el código fuente de su navegador, naciendo así el proyecto Mozilla.

Finalmente Mozilla Firefox fue reescrito desde cero tras decidirse a desarrollar y usar como base un nuevo conjunto de widgets multiplataforma basado en XML llamado XUL y esto hizo que tardara bastante más en aparecer de lo previsto inicialmente, apareciendo una versión 1.0 de gran calidad y para muchísimas plataformas a la vez el 5 de junio de 2002.

El 7 de enero de 2003, Apple lanza al mercado el navegador web Safari. Este Navegador se hace con casi la totalidad del mercado de las microcomputadoras Mac, debido a su velocidad y gran cantidad de actualizaciones. Asimismo, también entra al mercado del sistema operativo Windows.

A finales de 2004 aparece en el mercado Firefox, una rama de desarrollo de Mozilla que pretende hacerse con parte del mercado de Internet Explorer. Se trata de un navegador más ligero que su hermano mayor.

El 2 de septiembre de 2008 Google Chrome vio la luz. Es el navegador web desarrollado por Google y compilado con base en componentes de código abierto como el motor de renderizado de WebKit y su estructura de desarrollo de aplicaciones (Framework). Está disponible gratuitamente bajo condiciones de servicio específicas. El nombre del navegador deriva del término usado para el marco de la interfaz gráfica de usuario ("chrome"). En diciembre de 2011, Chrome superó a Internet Explorer 8.0 como el navegador más utilizado a nivel mundial.

La comunicación entre el servidor web y el navegador se realiza mediante el protocolo HTTP, aunque la mayoría de los ojeadores soportan otros protocolos como FTP, Gopher, y HTTPS (una versión cifrada de HTTP basada en Secure Socket Layer o Capa de Conexión Segura (SSL)).

La función principal del navegador es descargar documentos HTML y mostrarlos en pantalla. En la actualidad, no solamente descargan este tipo de documentos sino que muestran con el documento sus imágenes, sonidos e incluso vídeos streaming en diferentes formatos y protocolos. Además, permiten almacenar la información en el disco o crear marcadores (bookmarks) de las páginas más visitadas.

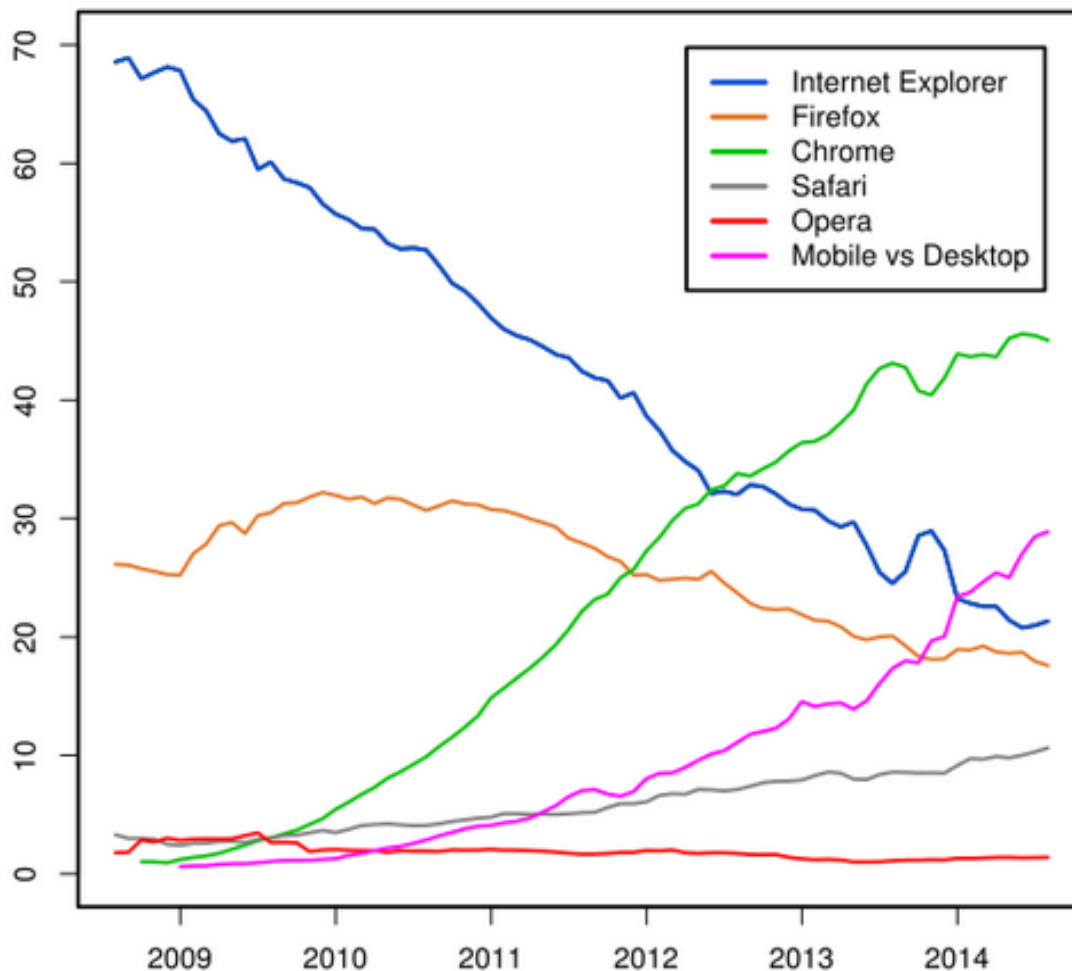
Algunos de los navegadores web más populares se incluyen en lo que se denomina una Suite. Estas Suite disponen de varios programas integrados para leer noticias de Usenet y correo

electrónico mediante los protocolos NNTP, IMAP y POP.

Los primeros navegadores web sólo soportaban una versión muy simple de HTML. El rápido desarrollo de los navegadores web propietarios condujo al desarrollo de dialectos no estándares de HTML y a problemas de interoperabilidad en la web. Los más modernos (como Google Chrome, Amaya, Mozilla, Netscape, Opera e Internet Explorer 9.0) soportan los estándares HTML y XHTML (comenzando con HTML 4.01, los cuales deberían visualizarse de la misma manera en todos ellos).



Usage share of web browsers



IES MIGUEL ROMERO ESTEO
C/MARTIN CARRIÓN S/N
29006 MALAGA

Teléfono: 951298668

Fax: 951298670

Estos apuntes forman parte del curso “Lenguajes de Marcas y Sistemas de Gestión de la Información” que forma parte del currículum del ciclo superior de “Administración de Sistemas Informáticos en Red”.

El temario completo se encuentra en la web del instituto Miguel Romero Esteo.



www.romeroesteo.es