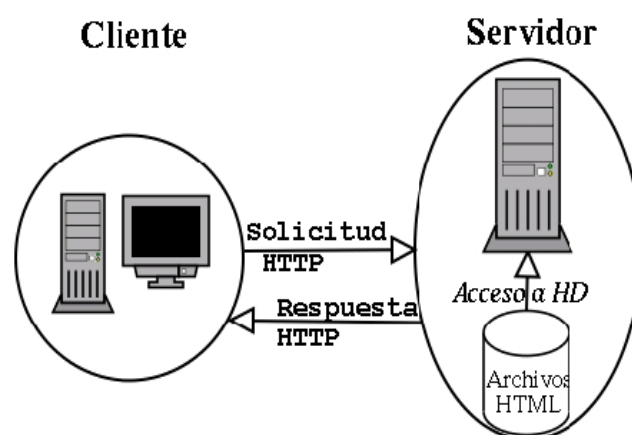


- - Características generales de un servidor web.
- - Protocolo HTTP.
- - Tipos MIME.
- - Apache
- - El cliente web. Parámetros de apariencia y uso.
- - Registro y monitorización.
- - Herramientas gráficas de configuración.
- - Bibliografía

Características generales de un servidor web

Un servidor web o servidor HTTP es un programa informático que procesa una aplicación del lado del servidor, realizando conexiones bidireccionales y/o unidireccionales y síncronas o asíncronas con el cliente y generando o cediendo una respuesta en cualquier lenguaje o Aplicación del lado del cliente.

El código recibido por el cliente suele ser compilado y ejecutado por un navegador web. Para la transmisión de todos estos datos suele utilizarse algún protocolo. Generalmente se usa el protocolo HTTP para estas comunicaciones, perteneciente a la capa de aplicación del modelo OSI.



Podemos decir que servidor web es un software instalado en el equipo con todas las condiciones necesarias para servir o entregar páginas web que le sean solicitadas por un navegador, asegurando que se muestren y representen todos los elementos necesarios para su correcto funcionamiento y visualización.

Existen varios tipos de servidores web, Apache es un software de código abierto, libre de uso y totalmente configurable, es en este momento el más utilizado en la red, ya sea en plataformas Linux o Windows.

El Servidor web se ejecuta en un ordenador manteniéndose a la espera de peticiones por parte de un cliente (un navegador web) y que responde a estas peticiones adecuadamente, mediante una página web que se exhibirá en el navegador o mostrando el respectivo mensaje si se detectó algún error. A modo de ejemplo, al teclear `www.wikipedia.org` en nuestro navegador, éste realiza una petición HTTP al servidor de dicha dirección. El servidor responde al cliente enviando el código HTML de la página; el cliente, una vez recibido el código, lo interpreta y lo exhibe en pantalla. Como vemos con este ejemplo, el cliente es el encargado de interpretar el código HTML, es decir, de mostrar las fuentes, los colores y la disposición de los textos y objetos de la página; el servidor tan sólo se limita a transferir el código de la página sin llevar a cabo ninguna interpretación de la misma.

Además de la transferencia de código HTML, los Servidores web pueden entregar aplicaciones web. Éstas son porciones de código que se ejecutan cuando se realizan ciertas peticiones o respuestas HTTP. Hay que distinguir entre:

- Aplicaciones en el lado del cliente: el cliente web es el encargado de ejecutarlas en la máquina del usuario. Son las aplicaciones tipo Java "applets" o Javascript: el servidor proporciona el código de las aplicaciones al cliente y éste, mediante el navegador, las ejecuta. Es necesario, por tanto, que el cliente disponga de un navegador con capacidad para ejecutar aplicaciones (también llamadas scripts). Comúnmente, los navegadores permiten ejecutar aplicaciones escritas en lenguaje javascript y java, aunque pueden añadirse más lenguajes mediante el uso de plugins.
- Aplicaciones en el lado del servidor: el servidor web ejecuta la aplicación;

ésta, una vez ejecutada, genera cierto código HTML; el servidor toma este código recién creado y lo envía al cliente por medio del protocolo HTTP.

Las aplicaciones de servidor muchas veces suelen ser la mejor opción para realizar aplicaciones web. La razón es que, al ejecutarse ésta en el servidor y no en la máquina del cliente, éste no necesita ninguna capacidad añadida, como sí ocurre en el caso de querer ejecutar aplicaciones javascript o java. Así pues, cualquier cliente dotado de un navegador web básico puede utilizar este tipo de aplicaciones.

El hecho de que HTTP y HTML estén íntimamente ligados no debe dar lugar a confundir ambos términos. HTML es un lenguaje de marcas y HTTP es un "protocolo".

Protocolo HTTP

Un servidor web opera mediante el protocolo HTTP, de la capa de aplicación del Modelo OSI. Al protocolo HTTP se le asigna habitualmente el puerto TCP 80. Las peticiones al servidor suelen realizarse mediante HTTP utilizando el método de petición GET, en el que el recurso se solicita a través de la url al servidor Web.

Hypertext Transfer Protocol o HTTP (en español protocolo de transferencia de hipertexto) es el protocolo usado en cada transacción de la World Wide Web. HTTP fue desarrollado por el World Wide Web Consortium y la Internet Engineering Task Force, colaboración que culminó en 1999 con la publicación de una serie de RFC, el más importante de ellos es el RFC 2616 que especifica la versión 1.1. HTTP define la sintaxis y la semántica que utilizan los elementos de software de la arquitectura web (clientes, servidores, proxies) para comunicarse. Es un protocolo orientado a transacciones y sigue el esquema petición-respuesta entre un cliente y un servidor. Al cliente que efectúa la petición (un navegador web o un spider) se lo conoce como "user agent" (agente del usuario). A la información transmitida se la llama recurso y se la identifica mediante un localizador uniforme de recursos (URL). El resultado de la ejecución de un programa, una consulta a una base de datos, la traducción automática de un documento, etc.

HTTP es un protocolo sin estado, es decir, que no guarda ninguna información sobre conexiones anteriores. El desarrollo de aplicaciones web necesita frecuentemente

mantener estado. Para esto se usan las cookies, que es información que un servidor puede almacenar en el sistema cliente. Esto le permite a las aplicaciones web instituir la noción de "sesión", y también permite rastrear usuarios ya que las cookies pueden guardarse en el cliente por tiempo indeterminado.

Transacciones HTTP

Una transacción HTTP está formada por un encabezado seguido, opcionalmente, por una línea en blanco y algún dato. El encabezado especificará cosas como la acción requerida del servidor, o el tipo de dato retornado, o el código de estado.

El uso de campos de encabezados enviados en las transacciones HTTP le dan gran flexibilidad al protocolo. Estos campos permiten que se envíe información descriptiva en la transacción, permitiendo así la autenticación, cifrado e identificación de usuario.

Un encabezado es un bloque de datos que precede a la información propiamente dicha, por lo que muchas veces se hace referencia a él como metadato —porque tiene datos sobre los datos—.

Si se reciben líneas de encabezado del cliente, el servidor las coloca en las variables de entorno de CGI con el prefijo HTTP_ seguido del nombre del encabezado.

El servidor puede excluir cualquier encabezado que ya esté procesado, como Authorization, Content-type y Content-length. Ejemplos de esto son las variables HTTP_ACCEPT y HTTP_USER_AGENT.

- HTTP_ACCEPT. Los tipos MIME que el cliente aceptará, dados los encabezados HTTP. Otros protocolos quizás necesiten obtener esta información de otro lugar.
- HTTP_USER_AGENT. El navegador que utiliza el cliente para realizar la petición. El formato general para esta variable es: software/versión biblioteca/versión.

El servidor envía al cliente:

- Un código de estado que indica si la petición fue correcta o no. Los códigos de error típicos indican que el archivo solicitado no se encontró, que la petición no se realizó de forma correcta o que se requiere autenticación para

acceder al archivo.

- La información propiamente dicha. Como HTTP permite enviar documentos de todo tipo y formato, es ideal para transmitir multimedia, como gráficos, audio y video. Esta libertad es una de las mayores ventajas de HTTP.
- Información sobre el objeto que se retorna.

Hay que tener en cuenta que la lista no es una lista completa de los campos de encabezado y que todos ellos sólo tienen sentido en una dirección.

Métodos de petición.

HTTP define un conjunto de **métodos de petición** para indicar la acción que se desea realizar para un recurso determinado. Aunque estos también pueden ser sustantivos, estos métodos de solicitud a veces son llamados *HTTP verbs*. Cada uno de ellos implementan una semántica diferente, pero algunas características similares son compartidas por un grupo de ellos.

GET

El método GET solicita una representación de un recurso específico. Las peticiones que usan el método GET sólo deben recuperar datos.

HEAD

El método HEAD pide una respuesta idéntica a la de una petición GET, pero sin el cuerpo de la respuesta.

POST

El método POST se utiliza para enviar una entidad a un recurso en específico, causando a menudo un cambio en el estado o efectos secundarios en el servidor.

PUT

El modo PUT reemplaza todas las representaciones actuales del recurso de destino con la carga útil de la petición.

DELETE

El método DELETE borra un recurso específico.

CONNECT

El método CONNECT establece un túnel hacia el servidor identificado por el recurso.

OPTIONS

El método OPTIONS es utilizado para describir las opciones de comunicación para el recurso de destino.

TRACE

El método TRACE realiza una prueba de bucle de retorno de mensaje a lo largo de la ruta al recurso de destino. Se utiliza con fines de comprobación y diagnóstico.

PATCH

El método PATCH es utilizado para aplicar modificaciones parciales a un recurso.

Tipos MIME

MIME es un acrónimo de extensiones multipropósito de correo de internet (Multipurpose Internet Mail Extensions)

Se trata de un standard que especifica como debe un programa (inicialmente un programa de correo o un navegador web) trasferir archivos multimedia (video, sonido, por extensión cualquier archivo que no esté codificado en US-ASCII). Con anterioridad al desarrollo de las extensiones MIME, cualquier archivo que no se limitase a texto ascii debía ser codificado a estos caracteres (uuencode uudecode).

MIME adjunta un archivo de cabecera a cada archivo, especificando el tipo y el subtipo del contenido del archivo principal. Gracias a esta información tanto el servidor como el navegador pueden manejar y presentar correctamente los datos.

La importancia de los MIME Types

En el uso diario de internet estamos beneficiandonos (y a veces sufriendo) los MIME TYPES. Cada vez que solicitamos una página de internet se abre un diálogo entre nuestro navegador y el server. Nuestro navegador pide la página. El servidor, antes de enviarla, nos confirma que ese recurso existe, y el tipo de datos que contiene. Esto último, mediante referencia al tipo MIME al que corresponde. Este diálogo, oculto al usuario, es parte de las cabeceras HTTP, protocolo que se sigue en la web.

Apache

Apache es un servidor web de software libre desarrollado por la Apache Software Foundation (<http://www.apache.org/>) cuyo objetivo es servir o suministrar páginas web (en general, hipertextos) a los clientes web o navegadores que las solicitan.

La arquitectura utilizada es cliente/servidor, es decir, el equipo cliente hace una solicitud o petición al equipo servidor y éste la atiende.

En el equipo cliente se ejecuta una aplicación llamada 'navegador o cliente web' que:

- Sirve de interfaz con el usuario: atiende sus peticiones, muestra los resultados de las consultas y proporciona al usuario un conjunto de herramientas que facilitan su comunicación con el servidor.
- Se comunica con el servidor web: transmite las peticiones de los usuarios.

El protocolo utilizado para la transferencia de hipertexto es **HTTP** (HyperText Transfer Protocol) que está basado en el envío de mensajes y establece el conjunto de normas mediante las cuales se envían las peticiones de acceso a una web y la respuesta de esa web.

HTTP es un protocolo sin estado, es decir, no recuerda nada relativo a conexiones anteriores a la actual. La conexión sólo tiene la duración correspondiente a la transmisión de la página solicitada si la encuentra, y si no la encuentra devuelve un código de error.

El servidor web Apache proporciona contenidos al cliente web o navegador como:

- 1. Páginas estáticas:** es el uso más generalizado que se hace de un servidor web. De esta forma se transfieren archivos HTML, imágenes, etc y no se requiere un servidor muy potente en lo que al hardware se refiere.
- 2. Páginas dinámicas:** la información que muestran las páginas que sirve Apache cambia ya que se obtiene a partir de consultas a bases de datos u otras fuentes de datos. Son, por tanto, páginas con contenido dinámico, cambiante.

El curioso nombre de este servidor hace referencia a sus orígenes. Cuando el proyecto inicial (Centro Nacional de Actividades de Supercomputación, NCSA, Universidad de Illinois) fue abandonado por su principal desarrollador, Rob McCool, diferentes webmasters comenzaron a desarrollar 'parches' para el código fuente de este servidor inicial y mediante el correo electrónico sincronizaban sus aportaciones. De esta forma apareció el proyecto Apache, cuyo nombre se debe a: **A PAtCHy server.**

El cliente web. Parámetros de apariencia y uso.

El cliente web por excelencia es el navegador.

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

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.

Funcionamiento

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).

Los estándares web son un conjunto de recomendaciones dadas por el World Wide Web consortium W3C) y otras organizaciones internacionales acerca de cómo crear

e interpretar documentos basados en la web. Su objetivo es crear una web que trabaje mejor para todos, con sitios accesibles a más personas y que funcionen en cualquier dispositivo de acceso a Internet.

Registro y monitorización

AWStats es una herramienta open source de informes de análisis web, apta para analizar datos de servicios de Internet como un servidor web, streaming, mail y FTP. AWstats analiza los archivos de log del servidor, y basándose en ellos produce informes HTML. Los datos son presentados visualmente en informes de tablas y gráficos de barra. Pueden crearse informes estáticos mediante una interfaz de línea de comando, y se pueden obtener informes on-demand a través de un navegador web, gracias a un programa CGI.

AWStats soporta la mayoría de los formatos de archivos log de servidor web conocidos, entre ellos Apache (formato de log NCSA combinado/XLF/ELF o formato común/CLFt), WebStar, IIS (formato de log del W3C) y muchos otros formatos comunes de Internet.

Herramientas gráficas de configuración.

Podemos utilizar una herramienta gráfica como webmin para instalar y configurar el servidor Apache.

Bibliografía.

(<http://recursostic.educacion.es/observatorio/web/eu/software/servidores/580-elvira-mifsud>)

<https://es.wikipedia.org/wiki/Wikipedia>