Unidad 3

Introducción al servicio HTTP

Índice

Internet, red de redes

Internet es la red de redes que interconecta ordenadores mediante diferentes protocolos.

Un ordenador puede conectarse a una red local y los ordenadores de dicha red local pueden conectarse a Internet mediante una puerta de enlace.

Existen ordenadores permanentemente conectados a Internet que ofrecen recursos (son los servidores)

¿Qué es el servicio HTTP?

HTTP es un protocolo, para la transferencia de contenido HTML "HyperText Markup Language" en su mayoría, entre un cliente y un servidor, indicado mediante una cadena de caracteres o URL "Uniform Resource Locator".

Existe una versión segura de HTTP, llamada HTTPS, que puede utilizar cualquier método de cifrado.

¿Qué es una URL?¿Qué indica una URL?

Una URL es una cadena que identifica los recursos de una red y se descompone en el siguiente formato:

esquema://host[:puerto][/ruta][?consulta]

Los **esquemas** más usados son los siguientes:

- http: Recursos HTML.
- https: HTTP encriptado, para realizar conexiones seguras, mediante SSL "Secure Sockets Layer".
- ftp: Transferencia de archivos.
- mailto: Direcciones de correo electrónico.
- File: Recursos disponibles en una red local.
- data - Agregar elementos en los documentos (Por ejemplo imágenes codificadas en base64).

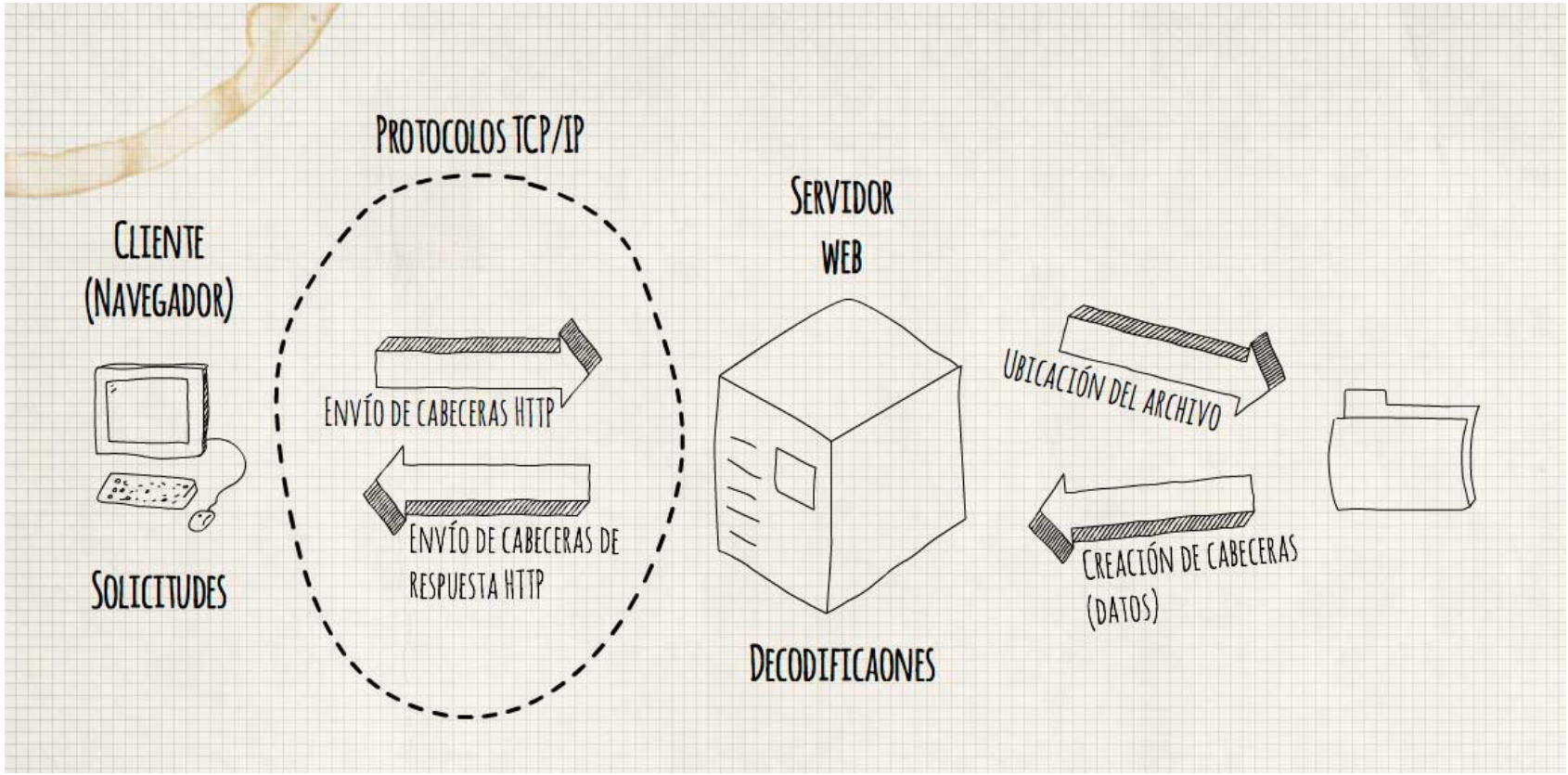
¿Qué indican el resto de los campos?

- **Host**: Dominio o IP del dispositivo a que queremos acceder.
- **Puerto**: Si no se indica, el puerto por defecto para http es 80.
- **Ruta**: Indica la carpeta y el recurso al que queremos acceder, los servidores suelen tener configurado recursos por defecto cuando no se indica en la URL, el más común es **index.html**. Otros ejemplos que suelen estar preconfigurados en los servidores son : index.php, index.shtml, index.jsp, default.asp.
- **Consulta o cadena de búsqueda**: Indica variables a través de la url (?nombre=helena&apellido=sanchez)

¿Cómo se realiza la comunicación HTTP?

La comunicación se realiza en dos pasos:

- El navegador manda una **petición HTTP** y solicita un archivo.
- El servidor responde con la información, que es descifrado por el navegador



El proceso es más complejo, ya que la comunicación no se realiza en una única etapa. La información se divide y se envía en pequeños paquetes que son unidos en el destino final, que es lo que establecen las normas o el protocolo **TCP.**

¿Cómo funciona el protocolo HTTP?

HTTP es un protocolo cliente/servidor cuyo funcionamiento está basado en el envío de mensajes y el mecanismo es el siguiente:

- El usuario especifica en el cliente web (navegador) la dirección de la página que quiere consultar.
- El cliente web descodifica la información de la URL diferenciando el protocolo de acceso, IP o nombre de dominio del servidor, puerto,etc.

- El cliente web establece una conexión (socket) con el servidor web y solicita la página (mensaje request).
- El servidor envía dicha página (sino existe, envía un código de error) y el cliente web interpreta el código HTML recibido (mensaje response del servidor).
- Se cierra la conexión.

El protocolo HTTP es un protocolo sin estado, es decir, no recuerda nada relativo o conexiones anteriores a la actual. En consecuencia, cuando el cliente web solicita un documento HTML pueden ocurrir dos casos:

- Si el servidor lo encuentra, lo envía.
- Si no existe, envía un código de error.

En ambos casos, y por ser un protocolo sin estado, al final se libera la conexión. La conexión solo tiene la duración necesaria para la transmisión del documento solicitado.

Para cada objeto que se transfiere se realiza una conexión independiente. Por ejemplo, si el cliente web solicita una página que incorpora varias imágenes, se realizan las siguientes conexiones: una para el documento HTML y una por cada una de las imágenes.

Para resolver estas situaciones de "falta de memoria" del protocolo, además de recordar la información de la sesión actual (cuya información se pierde, por ejemplo, al cerrar el navegador), se utilizan las cookies.

Tipos de mensajes HTTP

En HTTP existen 20 comandos o métodos para acceder a una url. Los comandos más importantes son los siguientes:

- GET: Obtiene información del servidor. Esta es la operación que se ejecuta cada vez que se pulsa sobre un enlace y se accede a una página web. Se pueden enviar datos en la URL.
- POST: Envía información desde el cliente web al servidor, por ejemplo, los datos introducidos en un formulario web.
- HEAD: Es similar a GET, pero solo se pide la cabecera de la página con información como el tamaño, tipo, fecha de modificación, etc.

- **PUT:** Almacena el documento enviado en el cuerpo del mensaje.
- **DELETE:** Elimina el documento referenciado en la URL.
- **TRACE:** Rastrea los intermediarios por los que pasa la petición.
- **OPTIONS:** Averigua los métodos que soporta el servidor.
- En una caché sólo se guardan las respuestas de las peticiones realizadas con GET y HEAD (POST no).

Códigos de estado

1xx: Mensaje informativo.

2xx:Éxito
- 200 OK
- 201 Created
- 202 Accepted
- 204 No Content

3xx: Redirección
- 300 Multiple Choice
- 301 Moved Permanently
- 302 Found
- 304 Not Modified

4xx: Error del cliente
- 400 Bad Request
- 401 Unauthorized
- 403 Forbidden
- 404 Not Found

5xx: Error del servidor
- 500 Internal ServerError
- 501 Not Implemented
- 502 Bad Gateway
- 503 Service Unavailable

¿Qué es un servidor web?

Un servidor web es un programa que haciendo uso del protocolo HTTP, atiende las peticiones de los navegadores o clientes web y les proporciona los recursos solicitados.

La arquitectura utilizada es cliente/servidor. El equipo cliente hace una petición de página web al servidor, y éste atiende dicha solicitud. El objetivo de un servidor web es servir páginas web a los clientes (navegadores) que las solicitan.

El equipo cliente hace la petición a través del navegador o cliente web. Esta aplicación:

1. Es la interfaz de usuario: atiende sus peticiones de páginas web, muestra los resultados y proporciona al usuario herramientas que facilitan la comunicación con el servidor.
2. Se comunica con el servidor web, es decir, transmite las peticiones de los usuarios.

El cliente web

El cliente web o navegador web (web browser) es una aplicación que permite visualizar páginas web alojadas en servidores web. Esta aplicación es capaz de interpretar el código HTML mostrando el contenido de la página en la pantalla y si la página lo permite, interactuando con ella o navegando a través de enlaces a otras páginas web.

También es posible ejecutar determinadas aplicaciones complementarias que aumentan la funcionalidad de la página web. Por ejemplo:

- **Aplicaciones que se ejecutan en el cliente web**, es decir, en el equipo del usuario. El servidor envía el código al navegador (Java o Javascript) y este lo ejecuta. El navegador debe ser capaz de interpretarlo, a menudo requiere de la instalación de extensiones.

- **Aplicaciones que se ejecutan en el servidor** (suelen ser en PHP, Perl, Python, etc.) y generan el código HTML que es enviado al navegador. Este lo interpreta y lo muestra al usuario.

El servidor web seguro

Un servidor web decimos que es seguro cuando garantiza la comunicación con el cliente web con autenticación y confidencialidad.

El protocolo SSL permite establecer una comunicación segura y codificada entre el servidor web y el navegador. SSL trabaja conjuntamente con el protocolo HTTP, creando un protocolo de transmisión de hipertexto seguro llamado HTTPS.

El protocolo HTTPS se basa en dos tipos de criptografía:

- Criptografía simétrica o de clave compartida.
- Criptografía asimétrica o de clave pública/privada.

Se utilizan ambos tipos, porque n función del objetivo se aprovechan las ventajas de cada una y se evitan sus inconvenientes.

La criptografía asimétrica se utiliza en los procesos de autenticación, ya que cada usuario protege su clave secreta, pero es lenta para el cifrado.

La criptografía simétrica es rápida para el cifrado y menos adecuada para la gestión de claves.

¿Cómo funciona SSL (Secure Socket Layer)?

El protocolo SSL garantiza que todo el intercambio de información realizada en una sesión de conexión se lleva a cabo de forma segura mediante encriptación.

El uso del protocolo SSL en una conexión TCP/IP proporciona:

- **Confidencialidad**: cifra la información transferida.
- **Integridad del mensaje**: controla cualquier modificación intencionada o accidental en la información mientras se transmite por Internet.

- **Autenticación del servidor**: asegura la identidad del servidor con el que se establece la conexión y al que el usuario puede enviar información personal. Esto se debe al certificado del servidor.
- **Autenticación del cliente**: el servidor conoce la identidad del usuario y le permite o no su acceso a áreas protegidas. El cliente debe tener instalado un certificado de cliente en su ordenador (navegador web), con el que identifica el servidor al usuario.

Comunicación SSL

La comunicación se realiza en dos fases:

1. **Fase de saludo**: el cliente y el servidor se identifican mediante el intercambio de claves públicas (encriptación asimétrica) y llevan a cabo la autenticación utilizando certificados digitales.
2. **Fase de comunicación**: se realiza el intercambio de datos utilizando generalmente una clave de tipo simétrico.

Los navegadores llevan ya incorporados un módulo SSL que se activa de forma automática cuando es necesario.