

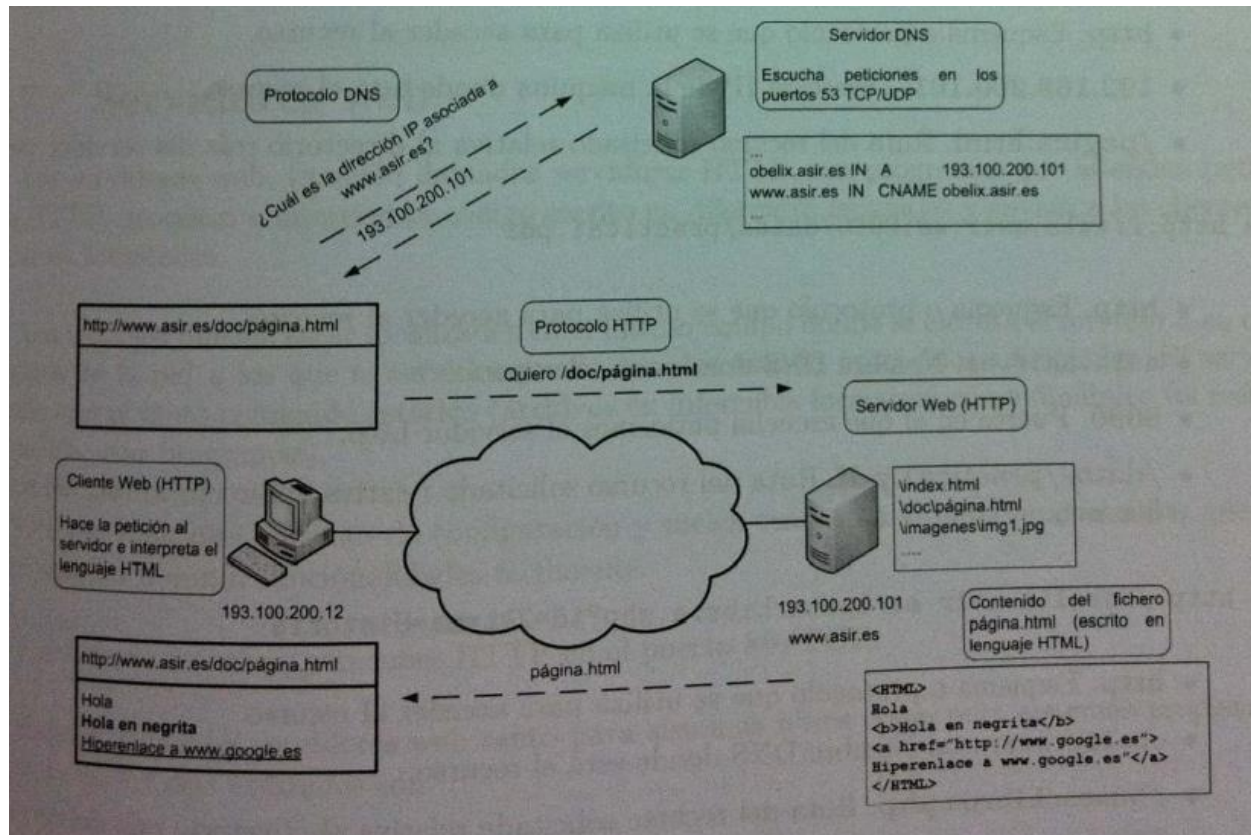
SERVIDORES WEB (HTTP)

- Servicio web y la WWW (World Wide Web).
 - Se basa en el protocolo HTTP (Hyper Text Transfer Protocol) de capa de aplicación que facilita el **acceso a información hipermedia remota** de sistemas conectados a red TCP/IP.
 - La WWW es un servicio de distribución de información que permite acceder a recursos electrónicos y aplicaciones distribuidos en servidores por todo Internet e identificados / localizados por direcciones (URIs o URLs).
- Hiperenlaces
- W3C controla actualmente el desarrollo de la WWW y desarrolla estándares Web



SERVIDORES WEB (HTTP)

- Componentes y funcionamiento



SERVIDORES WEB (HTTP)

- Nombres y direcciones (URIs y URLs)
 - Son las cadenas de caracteres que identifican de forma inequívoca los recursos que se acceden desde la web y que permiten su localización y acceso
 - URI, Uniform Resource Identifier → Permite gran variedad de sistemas de denominación y métodos de acceso
 - URL, Universal Resource Locator → Son tipos de URIs



SERVIDORES WEB (HTTP)

- Nombres y direcciones (URIs y URLs)

- `http://193.168.200.101/pagina.html`
 - `http`. Esquema o protocolo que se utiliza para acceder al recurso.
 - `192.168.200.101`. Dirección IP de la máquina donde está el recurso.
 - `/pagina.html`. Ruta del recurso solicitado relativa al directorio raíz del servidor web.
- `http://aula.asir.es:8080/datos/practica1.pdf`
 - `http`. Esquema o protocolo que se utiliza para acceder al recurso.
 - `aula.asir.es`. Nombre DNS donde está el recurso.
 - `8080`. Puerto en el que escucha peticiones el servidor `http`.
 - `/datos/practica1.pdf`. Ruta del recurso solicitado relativa al directorio raíz del servidor web.
- `http://obelix.asir.es/buscarLibros.php?id=2&tema=Historia`
 - `http`. Esquema o protocolo que se utiliza para acceder al recurso.
 - `obelix.asir.es`. Nombre DNS donde está el recurso.
 - `/buscarLibros.php`. Ruta del recurso solicitado relativa al directorio raíz del servidor web.
 - `?id=2&tema=Historia`. Parámetros y su valor enviados al servidor.



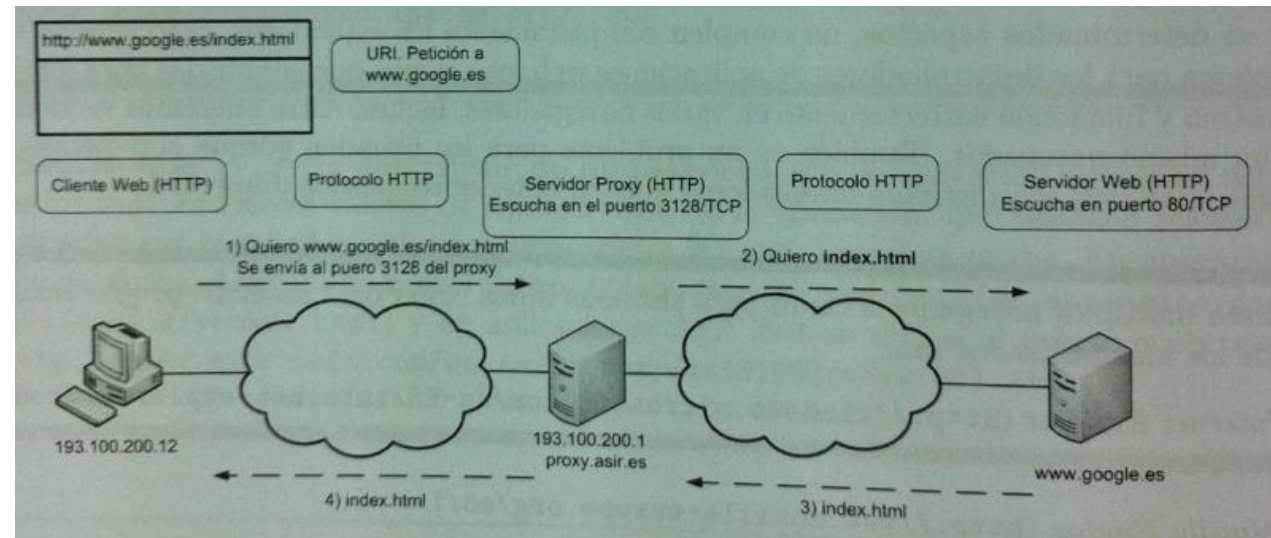
SERVIDORES WEB (HTTP)

- Clientes web
 - También conocidos como navegadores, son programas con los que interactúa el usuario, que le permiten introducir URIs / URLs para acceder a recursos disponibles en la red.
 - Pueden actuar como clientes de diferentes protocolos, pero su uso principal es como cliente HTTP.
 - Reciben los recursos de los servidores web, los procesan y muestran al usuario, permitiéndole la interacción si es necesario.
 - Mantienen una memoria caché para la información de uso reciente



SERVIDORES WEB (HTTP)

- Proxies web / proxy HTTP
 - Hace de intermediario entre un cliente web y un servidor web
 - **Proxy directo** (forward proxy): recibe una petición iniciada por un cliente web y se la traslada al servidor web. La petición es hacia el servidor web, no hacia el proxy que sólo hace de intermediario.
 - Su principal cometido es asegurar y controlar el acceso a redes externas de los clientes (web) de una organización



SERVIDORES WEB (HTTP)

- Proxies web / proxy HTTP
 - **Proxy inverso** (reverse proxy): recibe una petición de un cliente web y se la reenvían a uno o varios servidores web. La petición es hacia el proxy (que para los clientes es un servidor normal).
 - Se usan para dar acceso a servidores web que están detrás de un firewall y no son accesibles directamente, para balancear la carga entre varios servidores web, para optimización de tiempos de respuesta mediante el uso de caché



SERVIDORES WEB (HTTP)

- Protocolo HTTP (HyperText Transfer Protocol)
 - Es de facto el protocolo de comunicación en la Web
 - Define las reglas de comunicación entre los componentes software involucrados (clientes, servidores y proxies)
 - Es un protocolo sin estado que utiliza TCP como protocolo de transporte (no en HTTP/3 que usa UDP)
 - Determina los tipos de peticiones que los clientes pueden enviar, así como el formato y la estructura de las respuestas
 - También define una estructura de metadatos que se incluyen en las cabeceras y que se envían tanto en las peticiones como en las respuestas



SERVIDORES WEB (HTTP)

- Funcionamiento básico HTTP
 - El usuario introduce una URI (o URL) en el navegador o hace clic en un hipervínculo
 - El navegador establece una conexión TCP con el servidor web (si se invoca con nombre de dominio previamente debe resolverla). Si no se indica número de puerto se conecta al puerto 80 del servidor
 - Una vez establecida la conexión TCP, el navegador envía un mensaje HTTP de petición
 - El servidor envía un mensaje de respuesta que depende de la petición realizada y del estado del servidor
 - Se cierra la conexión TCP



SERVIDORES WEB (HTTP)

- Códigos de estado y error HTTP:
 - Códigos que envía el servidor en su respuesta HTTP al cliente y que informan de cómo ha sido procesada la petición
 - 100 – 199: Informativos
 - 200 – 299: Éxito
 - 300 – 399: Redirección
 - 400 – 499: Errores del cliente
 - 500 – 599: Errores en el servidor



SERVIDORES WEB (HTTP)

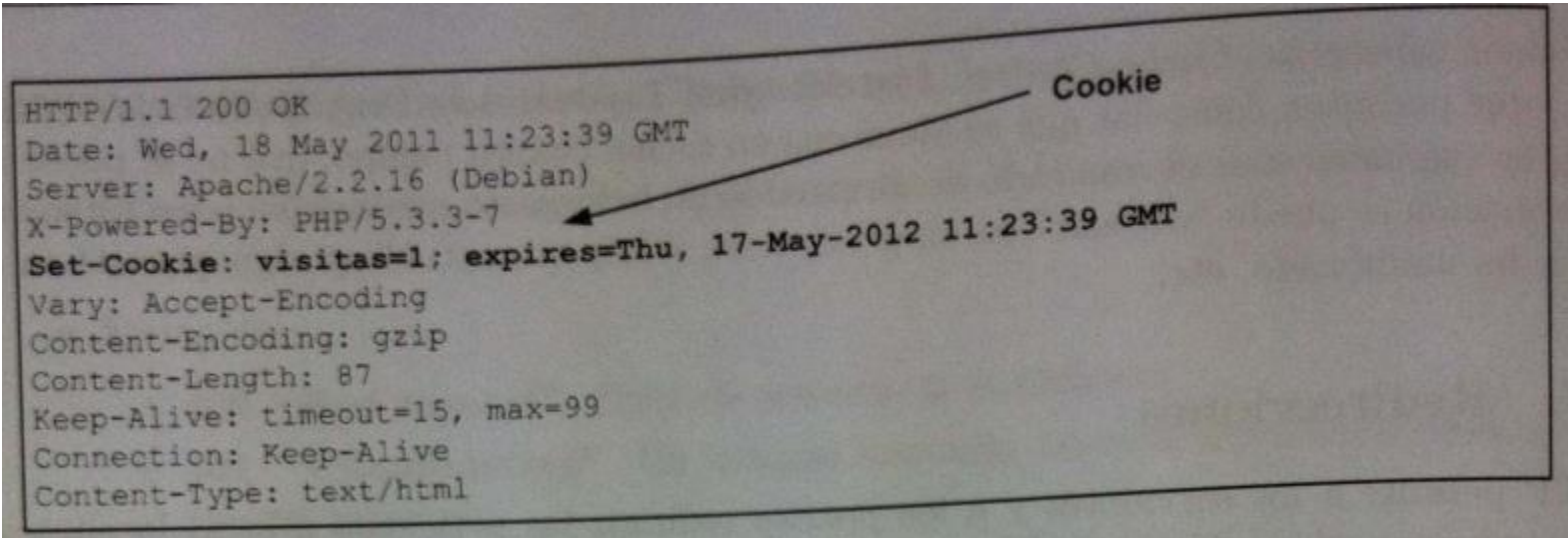
- Almacenamiento en caché (caching)
 - Los servidores web contienen información que no cambia durante ciertos periodos de tiempo, algunos de ellos en periodos muy largos
 - HTTP soporta almacenamiento en caché para evitar tráfico innecesario y aumentar el rendimiento
 - Así los navegadores y proxies pueden almacenar localmente recursos durante un tiempo y evitar nuevas descargas innecesarias
 - Se definen cabeceras que permiten controlar lo que se puede almacenar en caché y si se puede cachear en un proxy
- Compresión
 - Es posible que los servidores compriman los recursos antes de enviarlos a los clientes (usando gzip p.e.) para reducir el tráfico de red
 - Los clientes usan cabeceras en los mensajes de petición para indicar que soportan compresión
 - Los servidores en las respuestas también usan cabeceras para indicar al cliente que envían datos comprimidos



SERVIDORES WEB (HTTP)

- Cookies

- Una cookie es un fragmento de información que envía un servidor web en una respuesta HTTP y es almacenada por el navegador para su uso futuro en solicitudes posteriores al mismo servidor
- Los servidores envían las cookies usando cabeceras



The image shows a screenshot of an HTTP response header. The text is as follows:

```
HTTP/1.1 200 OK
Date: Wed, 18 May 2011 11:23:39 GMT
Server: Apache/2.2.16 (Debian)
X-Powered-By: PHP/5.3.3-7
Set-Cookie: visitas=1; expires=Thu, 17-May-2012 11:23:39 GMT
Vary: Accept-Encoding
Content-Encoding: gzip
Content-Length: 87
Keep-Alive: timeout=15, max=99
Connection: Keep-Alive
Content-Type: text/html
```

An arrow labeled "Cookie" points to the **Set-Cookie** header line.



SERVIDORES WEB (HTTP)

- Cookies

- Cuando un navegador realiza una nueva solicitud a un servidor HTTP consulta las cookies que tiene almacenadas y se las envía al servidor usando cabeceras
- Las cookies son utilizadas por los servidores para segmentar usuarios y conexiones
- Permite, por ejemplo, que aunque HTTP sea un protocolo sin estado recuerde datos de la navegación durante todo el tiempo que dura la misma
- Permite también que se muestren anuncios, productos, noticias, ... relacionadas con las preferencias del usuario



SERVIDORES WEB (HTTP)

- Autenticación
 - HTTP soporta mecanismos de autenticación para el control del acceso a los recursos que ofrece el servidor
 - No obstante la autenticación que ofrece no es segura y por ello se traslada a las aplicaciones web
- Conexiones persistentes
 - El uso de conexiones persistentes en HTTP consiste en que varias peticiones y respuestas sean transferidas usando la misma conexión TCP, lo que reduce el número de conexiones y un menor consumo de recursos de CPU y memoria y consiguiente reducción de tiempos de respuesta
 - Em HTTP 1.1 se usan conexiones persistentes por defecto (no así en HTTP 1.0)



SERVIDORES WEB (HTTP)

- MIME (Multipurpose Internet Mail Extensions)
 - Consiste en una serie de extensiones orientadas al intercambio de todo tipo de recursos (texto, audio, video, imágenes,...) en Internet usando protocolos HTTP y SMTP. Inicialmente fueron usadas para el correo electrónico, pero en la actualidad se extendió su uso a otros servicios
 - Definen:
 - Tipos y subtipos que determinan el contenido de los recursos enviados: text/html, text/css, image/gif, audio/x-mpeg, video/mpeg-2, multipart/form-data
 - Un conjunto de reglas para codificar mensajes no ASCII
 - Un conjunto de cabeceras para informar a los clientes y servidores sobre los recursos transmitidos
 - Se utilizan las cabeceras tipo MIME en:
 - Los mensajes de respuesta para informar al cliente de los recursos que le llegan. El navegador en función del tipo MIME visualiza el recurso, lo abre con una aplicación externa o pregunta al usuario cómo proceder
 - Los mensajes de petición de los clientes para informar al servidor de los tipos MIME que aceptan
 - Para encapsular uno o más recursos en el cuerpo del mensaje utilizando un tipo MIME denominado multipart

