

Sistemas Informáticos



Tema 2: Sistemas distribuidos basados en la *World Wide Web*



Contenido

- 2.1 Introducción
- 2.2 Web hipertexto
 - HTTP
 - SGML
 - HTML
 - CSS
 - XML
 - JSON
- 2.3 Web interactiva (Aplicaciones Web)
 - Formularios HTML
 - CGIs
 - Web APIs e interfaces híbridas
 - Páginas dinámicas
 - Python
- 2.4 Ejecución de código en el cliente
 - *javascript, ajax, jquery*
- 2.5 HTML5
- 2.6 Aplicaciones Web
- 2.7 Servidores de aplicaciones
 - Modelo de arquitectura
- Bibliografía del tema



Sistemas Informáticos



Tema 2: Sistemas distribuidos basados en WWW

2.1 Introducción



Introducción

- La generalización del uso de la *World Wide Web* como mecanismo de intercambio de información ha producido un cambio en el modo de realizar las aplicaciones distribuidas.
- El cliente Web se establece como la interfaz de uso genérico para todo tipo de interacciones de los usuarios con un sistema distribuido.
 - Cliente ligero. Sus carencias se deben suplir con algunos elementos de programación complementarios en el cliente.
- El servidor Web pasa a ser el núcleo básico donde se gestiona la ejecución de programas en el sistema distribuido.
- Es necesario completar la funcionalidad inicial del servidor Web para:
 - Ejecutar programas.
 - Enlazar con programas ya existentes.
- En este tema se abordan los distintos modos que se han ido desarrollando para implementar este modelo de sistemas, hasta llegar a los servidores de aplicaciones actuales.



Internet, Intranet, Extranet

- Tres nombres para referirse a un sistema distribuido que emplea la tecnología web como su arquitectura básica.
 - Internet: Red pública, externa, comunicación con el mercado. Soporte *business to customer (b2c)*.
 - Intranet: Red IP privada. Soporte de comunicaciones de un grupo de usuarios: empresa, institución... Soporte *business to employee (b2e)*
 - Extranet: Acceso a una Intranet por usuarios externos: Clientes, proveedores... Soporte *business to business (b2b)*.
- A nivel de arquitectura básica no existen diferencias entre ellas.
 - Sus diferencias provienen fundamentalmente de los distintas necesidades de seguridad de las aplicaciones que ejecuten.



World Wide Web



- Colección de documentos o páginas.
- Cada página puede contener objetos de diversos tipos:
 - Texto, gráficos, imágenes, voz, vídeo, otros documentos...
- Los objetos se reconocen mediante su *Uniform Resource Locator, URL*.
- Cada documento puede tener referencias a cualquier otro objeto en la red. Se denominan Hiperenlaces (*Hyperlinks*).
- El cliente accede a las páginas a través de un visualizador (*browser*) o navegador.
- Los nodos de la red presentan las páginas a través de un servidor.
- Comunicación mediante protocolo HTTP (*HyperText Transfer Protocol*) y cada vez más HTTPS
- Contenido de las páginas descrito en lenguaje HTML (*HyperText Markup Language*). Actualmente HTML5



Uniform Resource Locator, URL

- Extensión del clásico nombre de un fichero que permite definir objetos en la red Internet.
- En ella se contienen datos sobre:
 - El tipo de protocolo (cómo se transmite la información).
 - El nodo de la red en el que se encuentra el objeto.
 - El fichero físico en dicho nodo que contiene el objeto.
- Dos tipos de URLs:
 - Absolutas: especifican un camino (*path*) completo.
 - Ej: `http://www.webcon.com/~tbrown/computer.html`
 - Relativas: especifican un *path* relativo a la URL del documento.
 - Ej: `img/br2.gif`



Estructura URL

Toda URL tiene la siguiente estructura:

[esquema]:[parte-dependiente-del-esquema]

- Esquema
 - Objetos asociados con alguno de los protocolos o servicios disponibles en Internet:
ftp, http, gopher, mailto, news, file, telnet...
- Parte-dependiente-del-esquema
 - Datos necesarios para localizar el objeto, dependientes, por tanto, del tipo del mismo.
 - En su mayoría responden a la estructura:
`//[usuario[:clave]@]host[:puerto]/[path-ur]`
 - El parámetro puerto se puede omitir, y se tomaría el válido por defecto para el protocolo utilizado.

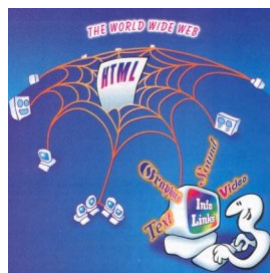


Ejemplos de *URLs*

- De transferencia de archivos:
 - ftp://[usuario[:contraseña]@]host[:puerto]/[path-del-archivo]
 - Ejp: ftp://ftp.ibm.com/programs/download/testip.zip
- De paginas HTML (protocolo HTTP):
 - http://[host][:puerto]/[path-de-la-página] [?consulta]
 - Ejp: http://search.yahoo.com/bin/search?p=url+uri
- De referencias de correo:
 - mailto:[Dirección-RFC822]
 - Ejp: mailto:bobsponja@fondobikini.com



Sistemas Informáticos



Tema 2: Sistemas distribuidos basados en WWW

2.2 Web hipertexto

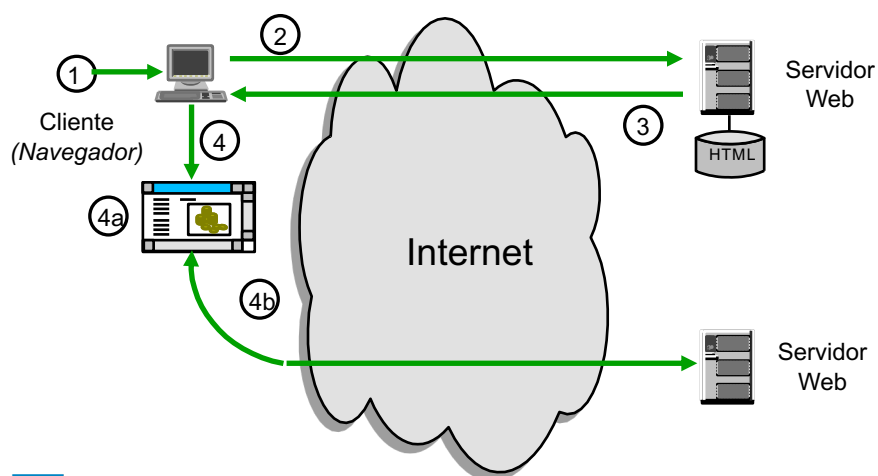


Web hipertexto

- Es el modelo más sencillo y original de la Web.
- Documentos distribuidos en cualquier servidor Internet.
 - Identificados por una URL.
- Pueden contener otros objetos o documentos en su interior.
- Enlaces desde cualquier documento a cualquier otro en la red.
- El documento se define atendiendo a los elementos que contiene, y no a la forma de presentarlo.
- Clientes estándar para visualizar documentos.



Funcionamiento básico (I)



Funcionamiento básico (II)

1. El usuario selecciona una URL que desea visualizar desde su navegador.
2. El navegador genera una petición HTTP y la envía al servidor Web solicitado.
3. El servidor recibe la petición y envía la página solicitada.
 - Una sesión TCP por cada solicitud.
 - La página se codifica utilizando HTML.
4. El cliente (navegador) visualiza la página recibida.
 - a. Puede tener asociados nuevos elementos en su interior.
 - b. Nueva solicitud a los servidores que los contienen para su recuperación (originalmente nueva sesión TCP, el protocolo HTTP actual permite reutilizar sesión TCP).



Hypertext Transfer Protocol, HTTP

- Protocolo de comunicaciones entre cliente y servidor WWW.
- Transporte: Conexión TCP sobre el puerto 80 (por omisión).
- Intercambio de mensajes ASCII entre ambos:
 - Cliente realiza una petición.

```
GET /hypertext/www/TheProject.html HTTP/1.0
```

- El servidor responde con un mensaje MIME:

```
HTTP:/1.0 200 Document follows
MIME-Version: 1.0
Server:CERN/3.0
Content-Type: text/html
Content-Length: 8247
```

```
<HEAD><TITLE>The World Wide Web Consortium (W3C)</TITLE></HEAD>
<BODY>
<H1><IMG ALIGN=MIDDLE ALT="W3C" SRC="icons/WWW(w3c_96x67.gif)>
The World Wide Web Consortium</H1><P>
```



Peticiones HTTP

Formato de las peticiones:

[método] URL [protocolo]

- URL: Identifica el objeto al que aplicar el método.
- Protocolo: El del mensaje enviado: HTTP/1.0, HTTP/1.1.
- Método: Acción a realizar:
 - GET: Lectura del objeto.
 - HEAD: Lectura de la cabecera del objeto.
 - PUT: Almacenamiento del objeto.
 - POST: Añadir al objeto la información enviada.
 - DELETE: Eliminar el objeto.
 - LINK: Conectar el objeto con otro determinado.
 - UNLINK: Desconectar el objeto de otro determinado.



Familia de lenguajes de la WWW

