

CONSTRUCCIÓN DE PORTALES

Juan Antonio Pastor Sánchez

1. Introducción

La Gestión de los contenidos informativos de los portales en Internet ha evolucionado desde la elaboración de páginas WEB hasta la gestión a través de sistemas de Bases de datos. Los motivos que han motivado este cambio han sido principalmente:

- Necesidad en la estructuración de contenidos informativos.
- Gestión descentralizada de la información a través de tecnologías cliente/servidor.
- Elaboración de sistemas de búsqueda de información especializados en categorías temáticas concretas.
- Aplicación de estilos homogéneos a las distintas páginas WEB.
- Heterogeneidad en los conocimientos del personal que mantiene la información.
- Simplificación en los procesos de gestión de servidores.

Por todo ello se han tenido que desarrollar tecnologías que permitan la integración del protocolo HTTPD (utilizado en los servidores WEB) y nuevas posibilidades de programación y desarrollo. Este documento tiene como objetivo ofrecer una visión general del panorama actual en el desarrollo y gestión de portales.

2. Esquema General de Funcionamiento

El protocolo HTTPD en un principio solo permitía la consulta de páginas elaboradas en lenguaje HTML. Sin embargo, antes de la creación definitiva de dicho estándar se introdujo un cambio fundamental: la posibilidad de incluir enlaces a programas CGI (Common Gateway Interface). Estos CGI están ubicados en el servidor y son ejecutados por el mismo. Las acciones de dichos CGI son muy variables, desde enviar mensajes de correo hasta interactuar con Bases de datos para introducir, borrar o modificar registros. En cualquier caso, los resultados se devuelven en formato HTML para que sean interpretados por el cliente WEB. La ventaja principal de las tecnologías dependientes del servidor radica en la seguridad que tiene el programador sobre su código, ya que éste se encuentra únicamente en los archivos del servidor que al ser solicitado a través del web, es ejecutado, por lo que los usuario no tienen acceso más que a la página resultante en su navegador.

Con el paso del tiempo los lenguajes de programación utilizados han sufrido continuos cambios. En un principio se utilizaban CGI escritos en lenguaje C++ que se compilaban y creaban programas ejecutables. No obstante, esta opción ha caído en desuso debido a la complejidad de desarrollo y el alto coste de tiempo que conlleva. La aparición de lenguaje PERL supuso una revolución en el mundo de los CGI. Se pasó del desarrollo de programas compilados y ejecutados por el servidor a otros en donde no se realizaba una compilación, sino que eran interpretados (sin necesidad de compilación) lo cual permitía un desarrollo rápido de aplicación, aunque su ejecución fuera algo más lenta y consumiera más recursos del servidor.

La expansión de la tecnología WEB y su difusión entre el público en general demandó una mayor flexibilidad en el desarrollo de aplicaciones. PERL aunque bastante flexible demandaba aún bastantes conocimientos de programación y solamente los expertos en la materia podían desarrollar CGI. Además, la gran desventaja de estos sistemas era que los ficheros con el código en PERL estaba separado de la propia página HTML por lo que su gestión y mantenimiento resultaba bastante complejo. Por ello se desarrollaron otras opciones que permitían la inclusión de código de programación dentro de las propias páginas HTML y que agilizaban la interacción entre los usuarios y la información depositada en los servidores WEB.

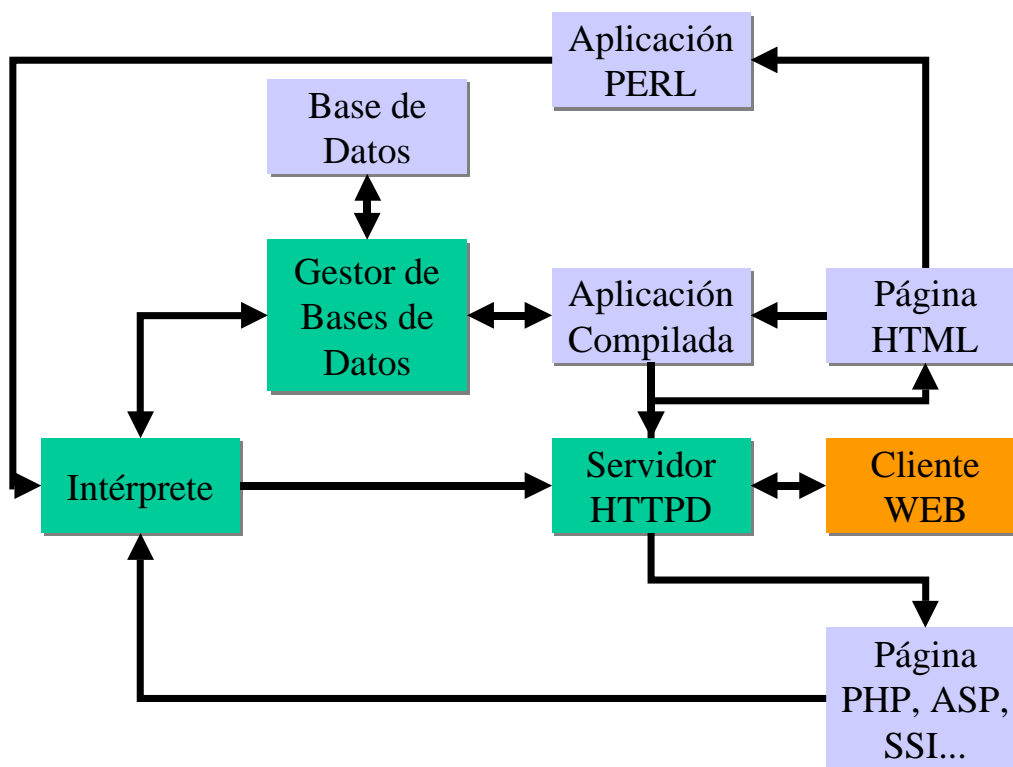
Las páginas WEB que incluyen códigos de programación que se ejecutan en el servidor son interpretados previamente desde el servidor a través de un programa Intérprete que ejecuta las acciones especificadas y devuelve una página HTML con los resultados. Estas páginas se caracterizan por tener en su nombre una extensión distinta a la de .htm o .html. El principal precursor de esta tecnología fue Microsoft con ASP (Active Server Page) que permitía la elaboración de páginas WEB que incluyen comandos en este lenguaje y que son interpretados por el software ASP. No obstante el mercado demandaba otras soluciones que no estuvieran limitadas a las tecnologías de bases de datos de Microsoft. Es por ello por lo que se desarrollo PHP (Preprocesed Hypertext Page). PHP está desarrollado para su uso público y gratuito y está totalmente integrado con sistemas de gestión de bases de datos relacionales como MySQL, mSQL, POSTGRESS, etc.

En definitiva para utilizar este tipo de lenguajes deben insertarse dentro del código HTML comandos de este tipo

```
<script language="php">
    $destinatario = "gestion@tubitacora.com";
    $asunto = "Lazarillo";
    $remite = $lazarilloemail;
    $mensaje0=$lazarillotitulo."\n\n";
    $mensaje1=$lazarillourl."\n\n";
    $mensaje2=$lazarillotexto.\n\n";
    $mensaje3="URL enviada por ".$lazarillonombre. " ( ".$lazarilloemail. ")\n\n";
    $mensaje=$mensaje0.$mensaje1.$mensaje2.$mensaje3;
    mail($destinatario, $asunto, $mensaje, "From: ".$remite);
</script></p>
```

Para diferenciar las marcas HTML del código de programación (PHP en este caso) se incluyen las marcas <script> y </script> que indican el inicio y el fin del código que debe ejecutar el intérprete PHP. Cuando un usuario desde su navegador WEB, carga una página con extensión .php (esto último dependerá de la configuración del sistema), el servidor no devuelve directamente la página, sino que interpreta y ejecuta el código PHP y lo sustituye por la salidas generadas en formato HTML. Las páginas HTML se convierten en "plantillas" que dan formato a los resultados dinámicos.

Por supuesto, existen comandos dentro de estos lenguajes que permiten su interacción con Bases de Datos, sistemas de correo electrónico, etc. En definitiva el funcionamiento general de los portales basados en esta tecnología es el siguiente:



Esquema de funcionamiento de la tecnología basada en el servidor

El cliente WEB solicita una página WEB, esta página WEB puede estar escrita en HTML, desde donde puede utilizar una aplicación compilada o un fichero en lenguaje PERL. Si se solicita una página PHP, ASP, SSI, etc, el código de programación será ejecutado por un programa Intérprete. En cualquier caso las conexiones con bases de datos se realizan a través del Sistemas Gestor de Bases de Datos y en todo momento es el servidor el que envía la salida al cliente WEB. Todo lo anterior se complementa con la evolución del HTML y la

aparición de JavaScript y otros lenguajes de script que son interpretados desde el propio cliente y no desde el servidor.

3. Pasos en la construcción del portal

Es recomendable establecer una serie de pasos que permitan una gestión y diseño coherentes de los contenidos del portal. Una organización del trabajo podría ser la siguiente:

1. Identificar la tecnología disponible: Servidores, aplicaciones instaladas en los mismos, Sistemas de gestión de bases de datos disponibles, ancho de banda, cuota de disco duro, posibilidad de conexiones seguras...
2. Establecer la temática del portal (especializada en un tema concreto, en una área geográfica específica, si va a ser un portal de carácter general, etc.)
3. Escoger un nombre y registrar el dominio
4. Identificar los contenidos y estructurarlos en secciones.
5. Establecer los criterios de búsqueda de información
6. A continuación habrá que identificar los contenidos que son susceptibles de mantenerse a través de bases de datos. Los criterios para decidir esto son básicamente tres: cantidad de información, grado de estructuración y necesidad de actualización permanente (cuanto más altos sean las necesidades en cada uno de estos criterios más podremos decantarnos por su gestión en Bases de Datos).
7. Realizar las estructuras de las diferentes tablas o bases de datos necesarias.
8. Elaborar el sistema de mantenimiento de las diferentes tablas o bases de datos a través de formularios WEB.
9. Elaborar el sistema de búsqueda de información. El sistema deberá contemplar la búsqueda por apartados o secciones, categorías dentro de cada uno de ellos y permitirá el uso de operadores booleanos.
10. Comenzar la introducción de información en las bases de datos.
11. Decidir si se van a implantar servicios de correo electrónico, mensajes a móviles o similares y comenzar a tramitar su contratación y puesta en marcha.
12. Realizar el diseño gráfico del portal, así como los estilos y formato de visualización de los contenidos de las bases de datos
13. Creación de las páginas PHP, ASP y HTML necesarias que realicen las consultas a bases de datos o que contengan contenidos relativamente estáticos. Es evidente que se aplicará dicho diseño.
14. Evaluación continua del portal.

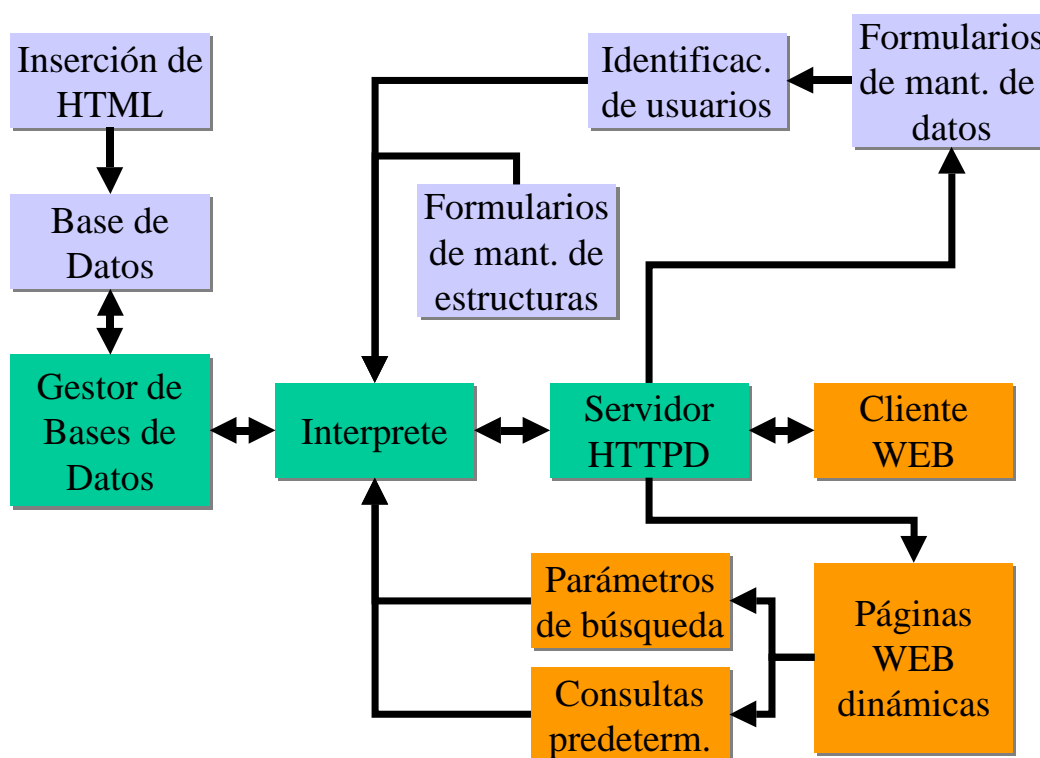
Por supuesto sobra indicar la necesidad de un reparto de tareas, así como de una especialización en el trabajo. Antes de la puesta en marcha definitiva habrá que realizar una de los resultados cada vez que se finalice un paso. Finalmente sería recomendable establecer una planificación temporal de las actividades con plazos de finalización de cada una de las tareas anteriormente expuestas, así como aquellas que son susceptibles de solaparse, lo cual puede ayudar a un ahorro considerable de tiempo.

4. Funcionamiento de Sistemas de Gestión de Información basados en WEB

El problema fundamental en el mantenimiento de información en bases de datos es la descentralización geográfica y de recursos humanos en este proceso. Es decir, las personas que mantengan la información pueden estar a gran distancia del sistema informático y disponer de recursos tecnológicos limitados. Por ello, la gestión de información se está realizando a través de clientes WEB, de esta manera, los navegadores se están convirtiendo en Clientes Universales que adoptan una interfaz muy similar para el mantenimiento de información en bases de datos. No obstante, con independencia de los diferentes flujos de trabajo y estructuras de trabajo en grupo el esquema fundamental de funcionamiento de estos sistemas se ajusta a unos principios básicos:

- Mantenimiento de la información a través de una serie de formularios WEB
- Los sistemas convencionales de mantenimiento de ficheros como el FTP no se utilizan o se usan exclusivamente para incluir en el servidor ficheros gráficos o multimedia.
- No es necesario el uso del HTML, aunque los sistemas están preparados para admitir marcas HTML en los diferentes campos de las bases de datos
- La estructura general se establece en: Categorías > Secciones > Elementos > Páginas. Los elementos pueden enlazar con otras páginas.

- Cada sección, elemento o página se ajusta a un estilo determinado, definido por el diseñador de las páginas o personalizable en cada registro. Además en las bases de datos se pueden definir estilos comunes como colores o tipos de letra para las secciones y páginas.



Esquema de Gestión de Información a través de WEB

En definitiva los sistemas de gestión de información para portales se realizan a través de formularios WEB. El diseñador del sistema debe contemplar la identificación y acceso de usuarios autorizados a módulos de gestión. No todos los usuarios tendrán los mismos derechos de acceso. Mientras que habrá usuarios que solo podrán gestionar ciertas bases de datos, otros usuarios (los administradores) podrán gestionar las estructuras de las bases de datos y estructuras lógicas de los contenidos. En cualquier caso la tecnología HTML que tanto preocupa a la hora de la visualización de los resultados es compatible mediante la introducción de marcas HTML en ciertos campos de texto de los registros. Sin embargo, esto último debe evitarse en la medida de lo posible mediante un buen análisis previo de las estructuras informativas. De esta manera, cuando el usuario acceda a páginas dinámicas, tanto si permiten la especificación de parámetros de búsqueda de información como si se trata de parámetros predefinidos, obtendrá información actualizada constantemente.

4. Sistemas Integrales de Gestión de Información en Web

Hoy día está apareciendo un nuevo paradigma en la gestión de información para portales. El desarrollo de aplicaciones utilizando tecnología que permite la programación de aplicaciones de gestión de información está encontrando una nueva alternativa en la aparición de software comercial que permite esta serie de procesos. Sin embargo, la gran novedad que aportan es la posibilidad de desarrollar portales de información y servicios sin necesidad de programación, utilizando una serie de funciones y características que agilizan la creación de formularios, bases de datos, gestión de usuarios, correo electrónico, listas de distribución, etc. todo ello integrado con tecnologías como Java y Javascript.

El principal exponente de esta tecnología es Lotus Domino que integra la gestión de usuarios, con la creación y mantenimiento de bases de datos y la elaboración de vistas a partir de búsquedas predefinidas de información. Todo ello a través del WEB y ofreciendo posibilidades de correo electrónico. Hoy día ya no es necesario elaborar páginas HTML, ni programar en PHP, ASP o PERL para consultar bases de datos o mantenerlas a través de un cliente WEB. Lotus Domino permite especificar la estructura de las bases de

datos, los formularios de consulta y mantenimiento y las definición de páginas que se elaboran dinámicamente con el contenido de las bases de datos. El gran inconveniente que presenta este tipo de tecnologías es que la formación requiere cierto grado de esfuerzo para pasar de unos esquemas de desarrollo basados en soluciones parciales y con un alto grado de gestión de los ficheros de script y CGIs en el servidor, a otro en donde estos aspectos no son importantes para la creación de los portales de información, centrándose en aspectos de estructuras y contenidos.