



Fascículo 3

Arquitectura de sitios web

Cuaderno 3: Arquitectura de datos y bases de datos en la Web

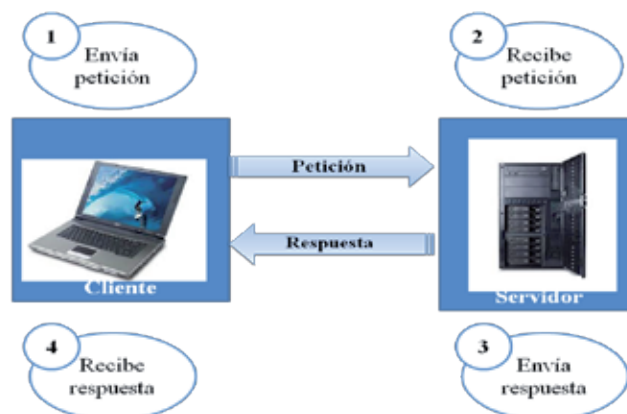
Servidores web: la arquitectura de datos en la Web

Como vimos en el cuaderno 1 de este mismo fascículo, un **servidor** es una **computadora** con **gran** capacidad de **almacenamiento** y **velocidad** que está conectada las 24 horas del día, de modo que permite a los administradores de la página y a los usuarios acceder al sitio web. Para que una computadora funcione como servidor es necesario instalarle un **programa** o **software servidor**.

En adelante cuando nombremos un **servidor web** nos estaremos refiriendo a un programa que permite a una computadora funcionar como servidor y no al hardware del servidor.

Los **servidores web HTTP** son aquellos programas que están diseñados especialmente para transferir hipertextos, páginas web o páginas HTML, a partir de la implementación de un **protocolo¹ HTTP**. El protocolo HTTP es el protocolo que se utiliza en todas las **transacciones de datos** de la Web.

El protocolo HTTP se rige por el **esquema petición-respuesta** entre un cliente y un servidor, y representa el **diseño clásico de arquitectura web**. Veamos un gráfico que representa este esquema:



¹ Ver definición de protocolo o protocolo de red en Glosario.



Servidores web: Apache

Apache es un **servidor web** de **código abierto**² compatible con plataformas Unix, Microsoft Windows y Macintosh que implementa el protocolo HTTP, y por lo tanto su arquitectura se rige por el **esquema petición-respuesta** que acabamos de describir.

El servidor web Apache presenta muchas ventajas, entre ellas que es altamente configurable y permite la autenticación de bases de datos. Sin embargo, fue muy criticado porque carece de una interfaz gráfica que permita que usuarios sin conocimientos técnicos avanzados accedan a su configuración.

Sabían que...

El nombre **Apache** se debe a que su creador quería que la nominación tuviera una connotación firme y enérgica, pero no agresiva. Este pueblo originario de América del Norte fue el último en rendirse durante la colonización, y esa resistencia inspiró al autor de Apache para invocar su nombre.

Los requerimientos mínimos para utilizar **Apache** son:



1. Un **sistema operativo compatible**, que tal como dijimos antes puede ser Unix, Microsoft o Macintosh.
2. Un **protocolo de control de transmisión**³ y un **protocolo de internet**⁴. El protocolo más difundido que reúne ambos requerimientos es el protocolo HTTP.
3. Bajar, instalar y configurar el **programa** de la [página oficial](#).

2 Ver definición de código abierto en Glosario.

3 Ver definición de protocolo de control de transmisión en Glosario.

4 Ver definición de protocolo de internet en Glosario.



Características generales de Apache

La arquitectura característica de un servidor que funciona con Apache es **modular**. Esto quiere decir que está compuesto por **partes** o **módulos** que utilizamos de acuerdo con las necesidades que se presentan. Algunos de los módulos más utilizados son:

- **mod_ssl:** es un módulo que permite transferencias de datos seguras, ya que ofrece un protocolo de capa de conexión segura o TLS (del inglés *Transport Layer Security*). Este protocolo hace que ante un requerimiento del servidor se autentifiquen los datos del cliente, y de este modo garantiza mayor seguridad al usuario.
- **mod_auth_Idap:** es un módulo que permite autenticar usuarios de una página web. Este módulo funciona cuando necesitamos autenticar los datos de usuario y contraseña que se ingresan del lado del cliente para acceder a una página web o servidor.
- **mod_cband:** es un módulo que permite configurar parámetros para controlar el tráfico que maneja el servidor y limitar el ancho de banda que utiliza cada requerimiento o demanda del lado del cliente.
- **mod_php:** este módulo permite trabajar con páginas dinámicas programadas con **PHP**.
- **mod_aspdotnet:** es un módulo que permite trabajar con páginas dinámicas programadas con **.Net** de Microsoft.

¿Qué es una base de datos?

Las **bases de datos** rodean todo lo que es el mundo web: están cuando ingresamos a nuestro correo electrónico, cuando hacemos compras en línea, en los juegos en red, en las redes sociales, etc. Pero... ¿qué es una base de datos?

Una base de datos es una **colección organizada de información**. En el mundo analógico podemos decir que una agenda telefónica es una base de datos que tiene **información** (nombres, apellidos, teléfonos, direcciones) **organizada** con alguna **lógica** que nos permita acceder a ella de un modo rápido y eficaz (orden alfabético).

En el mundo digital, una base de datos es un **sistema de archivos electrónicos organizados**. La lógica y organización de la misma es fundamental ya que, dada la enorme cantidad de datos que se manejan, sin una **estructura lógica** adecuada sería imposible el intercambio de datos entre una página web y un servidor.

Tradicionalmente, las bases de datos se organizan en tres **categorías**: campos, registros y archivos.

- Un **campo** es una pieza única de información;
- un **registro** es un conjunto de campos relacionados entre sí; y



- un **archivo** es una colección de registros.

Veamos esto en un ejemplo sencillo, una base de datos realizada en una planilla de cálculo:

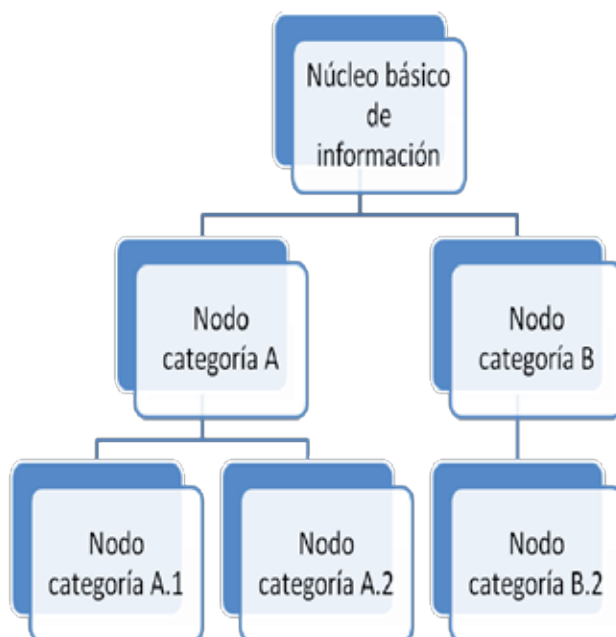
	A	B	C	D	E	F
21061	Rigoberto	Serrano	130 Melody Ln	Monett	MO	(417) 236-2812
21063	Rigoberto	Sevilla	28 Brandies Rd	Toms River	NJ	(732) 240-7274
21064	Rigoberto	Sevilla	41 Frederick Dr	Baville	NJ	(732) 606-0084
21065	Rigoberto	Serrano	No address	Sebastian 's	NV	(702) 733-6624
21066	Rigoberto	Serrano	No address	Coldwater	OH	(419) 678-0476
21067	Rigoberto	Serrano	5724 Coventry Park Dr	Haltom City	TX	(817) 503-8028
21068	Rigoberto	Serna	No address	Mercedes	TX	(956) 514-1518
21069	Rigoberto	Serrano	2520 Madison Ave, Apt 2	Ogden	UT	(801) 621-6056
21070	Rigoberto	Serna	2648 S 15th Pl	Milwaukee	WI	(414) 645-0041
21071	Rigoberto M	Sebastian	No address	Taft	CA	(661) 765-1090
21072	Robert	Veras	38 Catherine Dr	Attleboro	MA	(508) 223-5823

Hay diferentes tipos de base de datos, que difieren en la forma en la que estructuran la información. Los tipos más comunes y utilizados son:

- **Las bases de datos jerárquicas**

Tal como indica su nombre, este tipo organiza los datos a partir de una **jerarquización de la información**. Los datos se organizan a partir de un **núcleo básico de información**, del que se desprenden **nodos** con información relacionada y subsumida en el núcleo básico o central. Este tipo de base de datos es utilizada cuando se manejan grandes volúmenes de información.

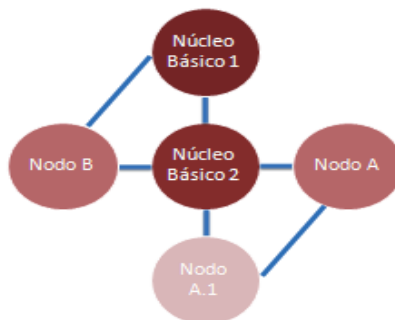
Sin embargo, tienen la limitación de no representar la redundancia de datos, que es cuando un mismo campo o registro pertenece a dos categorías diferentes.





Las bases de datos en red

La diferencia fundamental de este tipo de base de datos, es que permite que un **nodo** esté **subsumido** en **más de un núcleo básico** o central de información. Esta característica permite representar la redundancia de datos pertenecientes a diferentes categorías.



Las bases de datos relacionales

Este modelo, a diferencia del modelo jerárquico y el modelo en red, no requiere la definición de categorías. El modelo relacional se basa en la **formación lógica de conjuntos de datos basados en el binomio campo y registro, formando tablas.**

Num	Empresa	Dirección	Telefono
1	A & B Hidrosistemas SAC	Av. Tomás Marsano 1265	2710263
2	AJEPER	Av. La Paz Lt.30, Sta. María de Huachipa	3711812
3	Courier Cristiano Jeshua	Jr. Risco 538	96464860
4	Distribuidora Mensajil Hnos.	Jr. Camaná 1163	4239453
5	Global Medica	Jr. Pablo Bermudez 192	4336470
6	Juegos Emperatriz	Teodoro Cárdenas 790, Sta. Beatriz	4714603
7	Lima International School	Av. La Molina 1255, Sol de La Molina	4790046
8	Macro Work SAC	Calle La Habana 280	4222122
9	Notaria Beatriz Zevallos G.	Av. Túpac Amaro 762, 2º Piso	4852526
10	Par Systems SA	Av. Arenales 2523	4226322
11	Pesacon	Jr. Los Astrónomos Mz. C 11 Lt. 13	3872056
12	Publmedios	Av. Independencia 879, Huamanga	312436
13	Quipu Digital	Carret. Central Km. 21, 730	3584531
14	Sandwiches Passatore	Jr. Toribio Seminario 696	4508684
15	Serv. Ledesma SAC	Prolg. Iquitos 2200	2659006
16	Sonido Profesional Arriola	Alicia Alarcón 400	4820496
17	Stockholm Hipermedia S.A.	Los Calderos 140, Urb. Vulcano	3494222
18	Transporte Torre Fuerte	Calle Costa Azul Mz. B Lt. 16	2511749

Fecha: 12/09/2007
Hora: 11:00:00

Sistemas de gestión de bases de datos: MySQL

Los **sistemas de gestión de bases de datos** son **programas** específicos que se utilizan para manejar una **base de datos**, cuando el **usuario** o las **aplicaciones** o programas de una página requieren la información que está en dicha base.

El sistema de gestión de bases de datos más difundido es **MySQL**, sencillo de usar, rápido y gratuito para aplicaciones no comerciales. Para trabajar con este sistema de gestión de datos es preciso conocer el lenguaje de programación **SQL**.⁵

MySQL es una **base de datos relacional**, de modo que los datos se almacenan en tablas entre las cuales se establecen relaciones. Esta característica permite un manejo eficiente de los datos.

⁵ Ver definición de SQL en Glosario.



Glosario

Código abierto: es el término con el que se conoce al código de programación que es distribuido y desarrollado libremente, o sea que es una tecnología que no tiene patentes ni se comercializa. Sin embargo sí puede estar regido por licencias no comerciales.

Protocolo de control de transmisión: el protocolo garantiza que los datos serán entregados en su destino sin errores y en el mismo orden en que se transmitieron.

Protocolo de internet: conocido como IP, es un protocolo no orientado a conexión usado tanto por el origen como por el destino para la comunicación de datos.

Protocolo de red: un protocolo es un conjunto de reglas y procedimientos para la comunicación. Cuando dos o más computadoras están conectadas en red, se requiere –para que puedan comunicarse e intercambiar información– un determinado conjunto de reglas y procedimientos, que se denominan protocolo de red.

SQL: de las siglas en inglés para lenguaje de consulta estructurado (*Structured Query Language*). Es un lenguaje de programación declarativo para el acceso y gestión de bases de datos relacionales; se caracteriza por el manejo del álgebra y el cálculo relacional.

Fuentes

- Álvaro Martínez Echevarría, *Manual práctico de HTML*, Universidad Politécnica de Madrid, España, 1995. Versión digital en: <http://www-app.etsit.upm.es>
- <http://www.apache.org/>
- <http://www.desarrolloweb.com/>
- <http://www.guiaweb.gob.cl/>
- <http://www.javascriptya.com.ar/>
- <http://www.nic.ar>
- <http://www.programacion.com/>
- <http://www.tawdis.net/>
- <http://www.webtaller.com/>



Autores: Sebastián Otero
Carla Maglione
Coordinación editorial: Mara Mobilia