

## Bloque 2

---

## 3. Base de datos Relacional: MySQL

---

Las bases de datos relacionales son, sin lugar a dudas, las más utilizadas, y las estadísticas señaladas dejan bien a las claras que para una institución de propósito general, estas bases de datos son la opción preferida; entre los principales motivos están: la madurez de la tecnología y el amplio soporte con el que cuentan.

Si bien las soluciones comerciales de Oracle, Microsoft e IBM copan gran parte del mercado, MySQL consigue un porcentaje importante, resaltando sobre manera que se trata de una solución de código abierto, siendo éste el motivo principal por el que se ha elegido este desarrollo como opción para implementar una solución y analizar sus características.

### 3.1 Introducción

---

MySQL es el sistema gestor de bases de datos relacional, (SGBDR), de código abierto más popular. El soporte y apoyo de la comunidad en internet han sido decisivos para que sea uno de los productos más valorados.

Se trata de un SGBDR rápido, sólido y flexible, que ha alcanzado un grado de madurez tecnológica que lo hace del todo confiable. Como lenguaje de consulta utiliza SQL, que es el más desarrollado y utilizado actualmente. Tiene una serie de ventajas competitivas que lo hacen una elección ideal para este proyecto:

- Puede ser descargado gratuitamente haciendo uso de la licencia GPL
- Se puede personalizar y adaptar a las necesidades del usuario. Permitiendo la consulta y modificación de su código fuente
- Es un sistema cliente/servidor multiusuario: Permite el acceso simultáneo de un gran número de usuarios, realizando un control de acceso sobre los mismos
- Se trata de un sistema portable, puede ser utilizado en multitud de plataformas informáticas.

MySQL permite trabajar con multitud de lenguajes de programación, que permiten a distintas aplicaciones escritas con diferentes lenguajes el acceso a la base de datos. Cada uno de estos lenguajes hará uso de una interfaz de programación de aplicaciones, (API), específica, lo que dota de gran versatilidad al sistema.

Es muy habitual ver paquetes software preparados donde se hace uso de la combinación MySQL y PHP, entre las que destaca LAMP (Linux + Apache + MySQL + PHP), que facilita la labor de integración e instalación a los más profanos en la materia.

MySQL cuenta con dos programas principales: el servidor, y el cliente. El servidor está siempre a la espera de peticiones de los clientes; el cliente es el que actúa de interfaz frente al usuario.

A la hora de trabajar con MySQL puede hacerse desde una ventana de comandos, o bien a través de una interfaz gráfica phpMyAdmin, que resulta realmente cómoda; está desarrollada en PHP y nos permite realizar multitud de funciones en la base de datos, como la creación de una tabla, la gestión de privilegios de un usuario, o la importación/exportación de datos de una base de datos a otra. Para acceder a phpMyAdmin debe ser un usuario autorizado a ello.

## 3.2 Características

---

### 3.2.1 Motor de Almacenamiento

MySQL permite trabajar con varios motores de almacenamiento que siguen estrategias diferentes, y que van a incidir en el rendimiento del sistema.

El motor MyISAM, que es el motor por defecto de MySQL, siendo diseñado originariamente para trabajar con muchas operaciones de lectura, pero no demasiadas operaciones de escritura; esto es lo que suele ocurrir con frecuencia en portales web.

No cumple con las propiedades ACID, (Atomicidad, Consistencia, Aislamiento, Durabilidad), por lo que no soporta transacciones, y tendrá como consecuencia un aumento en la complejidad del entorno debido a consultas no ejecutadas correctamente.

Utiliza un archivo de índices y otro de datos para guardar la información, con lo que se consiguen consultas muy rápidas relacionadas únicamente con índices.

No cuenta con una funcionalidad específica para la gestión de la memoria, por lo que será el sistema operativo el que decida qué archivos se guardan en memoria.

Otros motores de almacenamiento son: InnoDB, Merge, Memory, Example, Federated, Archive y CSV, (comma-separated values en inglés; los datos se almacenan en ficheros de texto separados por comas).

### 3.2.2 Tipos de tabla

De acuerdo con el motor por defecto con el que trabaja MySQL, MyISAM, las tablas podrán ser estáticas, dinámicas o comprimidas.

Estáticas	Cada registro tiene asignados 10 Bytes
	Son muy rápidas
	Requieren más espacio en disco
	Son sencillas de almacenar en caché y reconstruir tras fallo
Dinámicas	Atributos con distintos tamaños
	Ocupan menos espacio en disco que las estáticas, en general
	Requieren mantenimiento para evitar su fragmentación
Comprimidas	No resulta sencillo reconstruir tras fallo
	Son de sólo lectura
	Ocupan mucho menos espacio en disco
	Carga de acceso reducida
	Cada atributo puede utilizar un algoritmo de compresión diferente

Figura 10: Tipos de tablas MySQL

Con este motor de almacenamiento, el tamaño máximo de las tablas es de  $(256^7 - 1)$  bytes, lo que supone una cantidad enorme de datos; es por ello que el límite real lo van a establecer los ficheros de los sistemas operativos y no los límites internos de MySQL

### 3.2.3 Tipos de atributos

Utilizaremos tipos de atributos en función de las necesidades de los datos; hay que tener en cuenta que dentro de cada tipo de datos, existen varios subtipos, que varían principalmente en su tamaño, por lo que habrá que alcanzar un compromiso en el tamaño de los mismos, ya que a mayor tamaño, menor velocidad de lectura, pero tampoco podemos irnos a un tamaño muy pequeño, porque puede significar la pérdida de parte de la información.

Numérico	Cadena	Fecha y Hora
Smallint	Char	Datetime
Int	Varchar	Date
Integer	Blob	Timestamp
Float	Text	Time
Double	Longtext	Year

Figura 11: Tipos de atributos más destacados

### 3.2.4 Tablas e índices

No todos los tipos de índices están disponibles para las distintas opciones de tabla, por lo que habrá que hacer un estudio de necesidades y elegir la mejor opción posible.

Cada vez que se realice una operación de inserción o actualización será necesario volver a organizar el índice para incluir los cambios, lo que puede acarrear una saturación de trabajo para el sistema. Los índices podrán ser:

- Clave primaria: Cada valor es exclusivo y ninguno puede ser nulo
- Índice exclusivo: Permiten valores duplicados
- Índice de texto completo: Diseñados para la búsqueda de palabras clave
- Índice ordinario: En caso de necesidad para algunas consultas. Puede provocar degradación en el rendimiento.

Tabla MyISAM	<p>Índices almacenados con extensión .MYI</p> <p>Los índices de número se almacenan con el bit más significativo primero</p> <p>Se permiten valores nulos en los índices</p> <p>Los datos y el índice se pueden incluir en directorios diferentes</p>
--------------	---

Figura 12: Relación de los índices y las tablas MyISAM

### 3.2.5 Sistemas Operativos soportados

MySQL es portable, funcionando en diferentes plataformas que tengan un compilador C++ y una implementación funcional de subprocesos POSIX, entre las que destacan:

- Unix
- Linux
- Windows
- Mac OS X

No todas las plataformas son igualmente aptas para ejecutar MySQL, y en palabras del Manual de Referencia de MySQL, las mejores plataformas para ejecutar MySQL son x86 con SuSE Linux, SPARC con Solaris, y FreeBSD.

En cualquier caso, en este proyecto se van a realizar implementaciones en Linux y Windows para comprobar las diferencias de rendimiento en ambos casos.

### 3.2.6 APIs disponibles

Llamamos API a una interfaz de programación de aplicaciones; se trata de un conjunto de funciones y procedimientos que ofrece una biblioteca para ser utilizado por otro software como una capa de abstracción. La API C fue desarrollada por el equipo de MySQL, por lo que es la más desarrollada, sin embargo existen otras posibilidades de gran eficiencia como: C++, PHP, Perl y Python.

### 3.2.7 Lenguaje

El lenguaje utilizado para trabajar con MySQL es SQL, (structured query language), lenguaje de consulta estructurado altamente eficaz para este tipo de bases de datos, que basan toda su operatividad en este lenguaje. La extensión de SQL es tremendamente amplia, y su estudio escapa de este proyecto, por lo

que puede consultarse la bibliografía recomendada para avanzar en su conocimiento.

### 3.3 Escalabilidad

En MySQL, cuando el tamaño de la base de datos crece considerablemente y hay que mejorar la solución se opta por la escalabilidad del sistema, que supone dotarlo de más recursos en cuanto a número de equipos.

Suelen aplicarse técnicas de optimización, que identifican qué datos son los más utilizados, si existen cuellos de botella en el sistema que haga que se deteriore su rendimiento, etc.

Cuando se opta por utilizar múltiples máquinas para aumentar el rendimiento global del sistema se forma lo que es llamado un clúster, que son un conjunto de máquinas vistas como una única entidad de cara al exterior. Se utilizan balanceadores de carga que permiten compartir el volumen de datos del sistema entre los distintos equipos del mismo.

Dependiendo del volumen de datos con los que se trabaje, el trabajo de escalado de MySQL será más o menos complejo, siendo en cualquier caso un trabajo poco intuitivo que requiere de una gran cantidad de pruebas para afinar el comportamiento del sistema, adaptándolo a las necesidades de la organización.

En este capítulo que finaliza se ha profundizado en una solución de base de datos relacional de código abierto, MySQL, posiblemente la más popular en este sentido. Se han conocido aspectos relevantes de su arquitectura y cómo trabaja con los datos que almacena, aportando una visión global a esta solución de la compañía Oracle. El siguiente capítulo se adentrará en las bases de datos que tal vez sean menos conocidas, las XML nativas.