

1 MySQL

MySQL es un sistema gestor de bases de datos relaciones de código abierto. Es un software gratuito y libre bajo los términos de GNU¹. MySQL también se encuentra bajo licencias cerradas como es la de Oracle, quien adquirió a Sun Microsystems, uno de los primeros propietarios de MySQL. Tras esta adquisición surgió MariaDB otro RDBMS² que surgió debido a la privatización de MySQL.

MySQL es la base de datos de software libre más popular del mundo para distribuir con bajo coste económico aplicaciones web o e-commerce, procesamiento de transacción online y base de datos integrada confiables. Es una base de datos integrada, segura para transacción, de conformidad con las propiedades ACID, rollback, recuperación de falla a nivel de fila. MySQL proporciona facilidad de uso, posibilidad de dimensionamiento y alto desempeño, así como un paquete completo de drivers de base de datos y herramientas visuales para ayudarles a los desarrolladores y DBAs a crear y gestionar sus aplicaciones MySQL.

Entre las funcionalidades podemos destacar por ejemplo la capacidad de redimensionamiento para atender a las demandas de cargas de datos y usuarios en crecimiento exponencial. También se dispone de clusters de replicación con auto-recuperación para mejorar la escalabilidad, el desempeño y la disponibilidad. Otra de las funcionalidades destacables es el poder cambiar el esquema online para satisfacer exigencias de negocio que se cambian constantemente. Este gestor también ofrece herramientas de monitorización (a nivel aplicación y consumo de recursos) y de autenticación mediante PAM y encriptación. Otra característica fundamental es el acceso tanto SQL como NoSQL para realizar consultar y operaciones simples y rápidas de Clave-Valor.

Cuando hablamos de escalabilidad MySQL ofrece esta solución a través de diferentes métodos. Permite replicación sobre diferentes nodos para distribuir la carga de trabajo. También trata con IDs en las transacciones para facilitar la conmutación por error y la creación de topologías de replicación en anillo. Respecto a los suscriptor, son multiprocesos para garantizar la disponibilidad, además que son cash-safe, es decir, permiten recuperarse ante fallos. Finalmente ofrece la herramienta de particionado MySQL Fabric para distribuir los datos en diferentes servidores.[1]

Otro de los pilares fundamentales es la concurrencia. MySQL ofrece un mecanismo de hilos con un buen *throughput*. Provee un hilo por conexión y este hilo no se reasigna hasta que la conexión no acaba, además este método es escalable. Con este mecanismo es suficiente para la mayoría de webs, sin embargo, si se va a tratar con mucho tráfico es recomendable usar la versión enterprise que es ofertada con un *Pool de Threads*. Con esta última funcionalidad el gestor nos provee grupos de thread en función de la demanda, es útil una vez sobrepasamos las 256 conexiones simultaneas a la base de datos.

El control de la concurrencia se realiza en dos niveles, primero a nivel servidor y segundo a nivel almacenamiento. Cuenta con bloqueos de lectura y escritura con gran transparencia. Destacar que tiene bloqueos a nivel de tablas y de filas, esta última ofrece la mayor concurrencia posible, sin embargo, es la menos eficiente. [2]

MySQL está disponible en diferentes SO como son: Linux, Windows, Oracle Solaris, macOS, FreeBSD entre muchas otras. Además ofrece conectores para todos los lenguajes clave de desarrollo que incluyen PHP, Perl, Python, Java, C, C++, C#, Ruby, Node.js y Go. Cabe destacar que si optamos por la opción de MySQL enterprise podemos acceder a una amplia gama de herramientas de Oracle.

En cuanto a comunidad cuenta con un buen apoyo, en StackOverflow, existen mas de medio millón de consultas respecto a MySQL.

MySQL cuenta con el soporte de Oracle. Como MySQL está dedicado mas al ámbito de web es fundamental para Oracle para que pueda completar su catálogo, de hecho, el gestor aumentó su popularidad tras la adquisición y el mantenimiento de Oracle.

En cuanto a empresas que usand MySQL destacan los grande como Twitter , Facebook, YouTube, PayPal, Yahoo, Linkedin, Ticketmaster...

Para terminar vamos a hablar del coste. Si optamos por la versión estandar, el coste es de 2000 euros anuales, en cambio, si optamos por la versión enterprise este precio se eleva a 5000 euros anuales la versión de MySQL con hasta 4 sockets en el servidor.[3]

¹General Public License

²Relational Database Manager System

2 Caché

Caché es un SGBD comercial que pertenece a *InterSystems*. Está destinado principalmente a servicios de salud, administración, banca, finanzas, gobierno y otros sectores parecidos. *Caché* permite usar la base de datos SQL y orientado a objetos, al igual que permite a los desarrolladores manipular directamente las estructuras de datos almacenados en arrays multidimensionales. *MUMPS* es el lenguaje que inspiró a *Caché*³.^[4]

*Caché*⁴ deriva gran parte de su potencial a su arquitectura única. El core de *Caché* es una BBDD que provee los servicios más comunes - incluyendo almacenamiento, concurrencia, transacciones, administración de procesos - necesario para un gestor en condiciones. Sin embargo, donde encontramos lo mejor es en su variedad de modelos de datos, este combina objetos, tablas relacionales y estructuras multidimensionales, todos accediendo a los mismos datos que se describen en una única vez. Pero *Caché* no es solo un SGBD, también incluye un servidor de aplicaciones con capacidades de POO.

Caché proporciona diferentes lenguajes de script para creación y acceso de datos (*Caché ObjectScript*, *Caché Basic*...). Lenguajes como *Java*, *C#*, *C++* tienen soporte para llamada directa, pero también permite el uso de *ODBC*, *JDBC*, *.Net* e interfaces que provee *Caché* para acceder a la BBDD. Como hemos dicho anteriormente, *Caché* incluye mas funcionalidades a parte de un SGBD, incorpora herramientas para desarrollar aplicaciones Web basadas en navegador a través de *InterSystem Zen* o *Caché Server Pages*. También tiene soporte para aplicaciones que no se basan en web, por ejemplo, aquellas cuya IU está programada en *Java*, *.net*, *C++* o *C#*, brindando grandes rendimientos si se implementa el resto de la aplicación con *Caché*. También es compatible con otras tecnologías como Angular a través de la *API Rest*.

InterSystem Caché no es muy abierto cuando hablamos de SO. Tiene licencias para unos SOs y hardware específicos que se pueden encontrar en esta web. Además, nos ofrecen dos tipos de licencias, *Independiente de Plataforma* o *Específico para Plataforma*, destacar que si se adquiere la específica y se cambia de SO, se aplican cargos.

En⁵ relación a las características, *Caché* incluye un sistema de bloqueos muy potente que hace poder tener acceso concurrente a la BBDD. Ofrece operaciones atómicas sin necesidad de bloqueos a nivel aplicación, esto lo realiza asignando ids únicos a cada objeto o fila. También posee la capacidad de realizar *locks* a nivel lógico, es decir, no bloquea grandes cantidades de datos mientras se realizan actualizaciones. Los bloqueos tiene gran granularidad y es capaz de realizar bloqueos a nivel fila u objeto. Por último, también dispone de bloqueos distribuidos.

Este gestor ofrece una buena escalabilidad debido a que usa un protocolo propio, *ECP (Enterprise Cache Protocol)*. Este protocolo permite que diferentes máquinas distribuidas usen unos a otros como bases de datos. Lo más destacable de este protocolo es que a nivel aplicación no hay cambios, es transparente por lo que para las aplicaciones es como acceder de forma local. Otro punto a destacar de este proceso, es que si un cliente necesita un dato de otro servidor de datos, el servidor local lo obtiene, se lo proporciona y lo cachea para futuros accesos. Tener datos cuyo acceso es habitual de manera cacheada, hace que el trafico en red se reduzca y de esta manera se brinde un servicio mucho mejor.

Cuando se habla de disponibilidad, *Caché* ofrece grandes ventajas. La primera de ellas es que hace uso de *Write-Image Journaling* que consiste en una técnica en dos fases. Al hacer una operación primero se escriben de memoria a un *"transitional journal"* local y luego en la base de datos, de esta manera si la segunda falla, se dispone de una copia en el journal. Es tolerante a fallos de hardware y hasta fallos eléctricos. Podemos añadir otras funcionalidades que usa para mejorar la disponibilidad como es el *Database Mirroring* (que consiste en replicar en un disco a parte en tiempo real), *ECP distribuido*, y los *Failover Clusters*.

Por último, del presupuesto no tenemos mucha información. Nos pusimos en contacto con el servicio de ventas de *Caché* pero fue tarea complicada obtener el precio, ya que nos hacían llamar a un servicio de venta y no nos ofrecían documentación.

Sin embargo, nos avisaron de que *Caché* está evolucionando a *Iris*, un servicio que ofrece gestor de bases de datos, análisis y servicios de machine learning.

³<https://en.wikipedia.org/wiki/MUMPS>

⁴<http://www.intersystems.com/wp-content/uploads/sites/11/CacheTechnologyGuide.pdf>

⁵link

3 Microsoft SQL Server

Es un gestor de bases relacional desarrollado por *Microsoft*. Este gestor dispone de diferentes ediciones como es el *Enterprise*, que a parte de proporcionar el gestor, añade una serie de servicios y herramientas. La edición *standard* que incluye el gestor y una serie de servicios mas básicos y con limitación en el soporte de nodos del cluster. Otra de las versiones a destacar es *Azure*, que ofrece el servicio de este gestor pero como un *PaaS*.

La arquitectura que usan se basa en una capa externa que actúa como interfaz. Todas las operaciones que pueden ser invocadas en *SQL Server* se comunican a través de *Data Stream (TDS)*, este protocolo de nivel aplicación se encargará de transferir los datos entre cliente y servidor.

Microsoft SQL Server almacena tablas con columnas tipadas, al igual que permite el uso y definición de *UDTs*. El lenguaje de consulta que se usa no es SQL sino una variante, *T-SQL*. Este expande SQL incluyendo programación *procedimental*.

Una de las ventajas que tiene este gestor es la cantidad de servicios que ofrece, como por ejemplo, *servicio de Broker*, *servicio de análisis*, *SQLCMD* (permite ejecución de consultas desde terminal), y un *servicio de replicación*. Respecto al servicio de replicación incluye replicación transaccional, replicación por mezcla (cambios tanto en publicador como suscriptor, y luego se sincronizan de manera cruzada), y replicación por instantánea (se realiza una copia de la base y los suscriptores replican estos datos).

En cuanto a soporte de lenguajes, ofrece drivers para muchos, desde *Java* y *C++* hasta *Ruby*, en odemos encontrar todos.

Como es de esperar es un gestor que garantiza las propiedades *ACID* en las transacciones. Para garantizar una alta disponibilidad nos ofrece replicación, envios de log, *mirroring*, *clustering* y una funcionalidad llamada *allwaysON Availability Groups*. La replicación usa una serie de *jobs*, donde interactuan un *publicador(publisher)*, un *distribuidor*, y un *suscriptor(subscriber)*. El log se lleva a cabo a través de tareas que realizan transacciones del log con copias de la base desde un *primario* a un *secundario*, al igual que con el servicio de *mirroring*. El clustering consiste en que los diferentes servicios en funcionamiento realizan almacenamiento en diferentes putos compartidos que son altamente fiables. Por último el servicio de *AllwaysON* consiste en realizar copias en diferentes *grupos de disponibilidad* que creamos, estos grupos realizan la conmutación por error conjuntamente.

Respecto a concurrencia ofrece tres tipos de soluciones para mejorarla. La primera es *solo lectura*, permite leer pero no actualizar o borrar, esto brinda una alta concurrencia. Sin embargo, el gestor puede que realice bloqueos a nivel de filas para reforzar ciertas operaciones, también es capaz de realizar bloques de lectura en lugar de escritura. La segunda es a través del *locking*, el cursor usa el menor nivel de bloqueos que garantiza poder garantizar actualizaciones y borrados en el conjunto de datos. Esta opción mencionada ofrece baja concurrencia pero mayor operabilidad. Por último tenemos, la *concurrencia optimista*, consiste en actualizar o borrar solo si tras la última lectura no han sido cambiados. Esta última solución ofrece gran concurrencia pero no asegura las modificaciones.

Respecto al coste de las licencias, la versión *Enterprise* tiene su coste de licencia por núcleos, su precio inicial es de 13.000 euros en adelante. La versión *Standard* tiene un coste de 3.000 euros y al igual que el anterior es en función al número de núcleos. Destacar que ambos comienzan con un total de dos núcleos. Para terminar, destacar que dispone de una licencia Express para bbdd rápidas con coste gratuito, y de una licencia de desarrollador con limitaciones de uso y almacenamiento también sin coste.

References

- [1] MySQL, “Mysql enterprise edition product guide.” <https://www.oreilly.com/library/view/high-performance-mysql/9780596101718/ch01.html>. concurrencia y lock.
- [2] MySQL, “Ficha técnica mysql enterprise edition.” <https://www.mysql.com/why-mysql/white-papers/ficha-tecnica-mysql-enterprise-edition/>.
- [3] MySQL, “Para obtener precios.” <https://shop.oracle.com/apex/product?p1=MySQL>.
- [4] Wikipedia, “Intersystems cache.” https://en.wikipedia.org/wiki/InterSystems_Cache/.