



8. CONCLUSIONES Y DESARROLLO FUTURO.

En este Proyecto Fin de Carrera se ha abordado el estudio de dos de los sistemas gestores de bases de datos más completos bajo entorno Linux: MySQL y PostgreSQL. Este estudio ha comprendido un análisis teórico de las características y funcionalidades de cada uno, un análisis comparativo de dichas características, un análisis del rendimiento de cada uno mediante el empleo de diversos bancos de pruebas y un análisis del rendimiento de dichos sistemas cuando funcionen con aplicaciones reales.

Del análisis teórico de MySQL y PostgreSQL se obtiene que ambos sistemas gestores de bases de datos son bastante completos, aptos para la mayoría de aplicaciones que se puedan imaginar. Ambos son sistemas relativamente sencillos de instalar, además, suelen venir incluidos en la mayoría de las distribuciones de Linux. El carácter “*Open Source*” de ambos sistemas es también uno de los principales causantes de la popularidad de la que gozan estos sistemas, que nada tienen que envidiar a otros sistemas gestores de bases de datos comerciales como Oracle. Ello implica también que existan muchas aplicaciones, interfaces y funcionalidades desarrolladas por terceras partes que se pueden emplear con estos sistemas.

Ambos sistemas gestores de bases de datos comparten bastantes similitudes, tales como el tipo de licencia, número de plataformas soportadas, el carácter cliente/servidor, interfaces de usuario parecidas (programas monitor), soporte para múltiples servidores en la misma máquina, juego de funciones y operadores amplio, gran número de herramientas de administración, soporte para zonas horarias, diferentes juegos de caracteres, etc.

Pero en general, de los dos sistemas, PostgreSQL aporta más funcionalidades y características adicionales. Algunas de estas características extra que aporta son el carácter orientado a objetos, soporte completo para esquemas, subconsultas, transacciones, etc. Además, también ofrece disparadores (*triggers*), vistas, restricciones, roles, un conjunto más rico de tipos de datos y la posibilidad de empleo procedimientos almacenados, que se pueden escribir en varios lenguajes. No obstante, las diferencias principales entre ambos sistemas son dos. La primera es que PostgreSQL soporta transacciones, mientras que en MySQL sólo están disponibles con determinados motores de almacenamiento. La segunda es la técnica que emplean para el control de concurrencia, que en MySQL está basada en bloqueos y en PostgreSQL es multiversión. El control de concurrencia multiversión es más seguro y proporciona más estabilidad e integridad en los datos, pero, como se ha visto en las pruebas de rendimiento, influye negativamente en el rendimiento del sistema.

MySQL destaca en algunos aspectos aislados, como por ejemplo el soporte para replicación y el número máximo de usuarios simultáneos que soporta (casi cuatro veces mayor que el número máximo de usuarios concurrentes de PostgreSQL), aunque este último aspecto depende mucho de la máquina en la que se ejecute el sistema gestor de bases de datos, pudiendo existir alguna configuración en la que esta diferencia sea inapreciable. Pero quizás la diferencia más notable es que proporciona una interfaz gráfica de usuario muy avanzada, que permite que tanto su uso como las tareas más comunes de administración se hagan de manera más sencilla.



Una vez visto que PostgreSQL es el sistema de los dos estudiados que más características y funcionalidades ofrece, el siguiente paso fue estudiar el rendimiento de cada uno de estos sistemas en funcionamiento. Este estudio se llevó a cabo desde dos puntos de vista diferentes. Primero se atacaron los dos sistemas con tres bancos de pruebas diferentes. Estos bancos de pruebas medían valores límite, tiempos de respuesta en determinadas operaciones y factores de concurrencia. El otro punto de vista ha consistido en estudiar el rendimiento de cada uno de los sistemas en aplicaciones reales, tanto en un entorno de un único usuario como en entorno multiusuario.

Por lo que respecta a los bancos de pruebas, en la inmensa mayoría de los test realizados MySQL presentaba unos tiempos de respuesta más bajos y un mayor número de operaciones realizadas por segundo. Además, en las pruebas de concurrencia se obtuvo que MySQL se comporta bastante mejor que PostgreSQL cuando la carga de usuarios simultáneos en el sistema aumenta. Como MySQL posee varios motores de almacenamiento, las pruebas con este sistema gestor de bases de datos se hicieron empleando los dos motores más comunes: MyISAM e InnoDB. De estos motores, sólo el segundo soporta transacciones. Los mejores resultados se obtuvieron para el motor de almacenamiento MyISAM, siendo los resultados obtenidos con el motor InnoDB muy similares a los obtenidos con el sistema PostgreSQL. Por tanto, se pudo concluir que en este aspecto MySQL es el sistema más rápido de los dos estudiados. PostgreSQL obtuvo un peor rendimiento, pues acusa de alguna manera el soporte transaccional y las técnicas de control de concurrencia más avanzadas de las que dispone.

En cuanto a las aplicaciones, se desarrolló un analizador de bases de datos con interfaz Web. Este analizador se empleó para evaluar los tiempos de respuesta ante determinadas operaciones realizadas a través de una interfaz Web. Como era de suponer en vista de los resultados obtenidos en los bancos de pruebas, lo mejores tiempos de respuesta los obtuvo MySQL. Este analizador Web realizaba las medidas en un entorno de un usuario único. Para ver cómo se comportaría el sistema con una aplicación real que fuese utilizada por varios usuarios simultáneamente, se escogió una aplicación existente y se modificó para que funcionase con ambos sistemas gestores de bases de datos. Al analizar dicha aplicación con Apache Bench, se vio que la misma funcionaba mejor con MySQL que con PostgreSQL, sobre todo a medida que el número de usuarios aumentaba considerablemente. Por tanto, dado que dicha aplicación es completamente funcional con los dos sistemas gestores de bases de datos a estudio, se recomendaría su uso con el gestor de bases de datos MySQL.

Después de haber realizado los estudios teóricos de cada sistema gestor de bases de datos y de haber realizado una batería de pruebas bastante completa con ambos, se llega a la conclusión de que el empleo de uno u otro sistema gestor de bases de datos depende de los requisitos que se quieran cumplir, y del ámbito en el que se utilicen. Ninguno de los dos sistemas gestores de bases de datos es perfecto. Por tanto, a la hora de emplear uno u otro se tratará de escoger el más conveniente en cada caso. En este sentido, resultaría apropiado emplear MySQL para aplicaciones en las que se requiere gran velocidad a la hora de hacer operaciones, tanto en entornos con pocos usuarios como en entornos con alto nivel de concurrencia. También resultaría apropiado cuando el servidor de bases de datos ha de alojarse en máquinas con escasos recursos, debido a su bajo consumo.



Por el contrario, se recomendaría el uso de PostgreSQL cuando se requieren aplicaciones que necesiten claramente un funcionamiento más eficaz, lo que se puede conseguir con el uso de *rollbacks*, subconsultas, transacciones, etc. Además, como es capaz de comprobar la integridad referencial, así como de almacenar procedimientos en la propia base de datos, resultaría una excelente opción para muchos campos de aplicación que requieran una cierta especialización, sobre todo para aquellos programadores que provienen de otros sistemas gestores que proporcionan estas características. Como consume gran cantidad de recursos, su uso se recomendaría en máquinas grandes con muchos recursos.

En cuanto al desarrollo futuro, este Proyecto Fin de Carrera podría constituir la “primera piedra” de un amplio estudio comparativo sobre todos los sistemas gestores de bases de datos mas potentes que se emplean en la actualidad, tanto comerciales como de libre distribución. Por tanto, el estudio que se ha realizado aquí se podría extender a otros sistemas como Oracle, Informix, Sybase, etc. Además, también se podría extender el estudio realizado sobre MySQL y PostgreSQL evaluando su comportamiento en máquinas y entornos diferentes a los empleados en este Proyecto Fin de Carrera. Por otra parte, también se podrían modificar las aplicaciones desarrolladas y empleadas para su uso con otros sistemas gestores de bases de datos.

De esta manera, tanto este Proyecto Fin de Carrera como su desarrollo futuro podrían convertirse en una excelente guía que permita evaluar las ventajas y desventajas del uso de un determinado sistema gestor de bases de datos en un ambiente determinado, con objeto de facilitar al lector la elección del sistema más adecuado para el cumplimiento de unos determinados requisitos.