

Trabajo 1: Optimización de MySQL y PostgreSQL

Entorno de trabajo.

Para desarrollar el trabajo deberéis utilizar la máquina virtual de VMWare con sistema operativo CentOS 6.5 que se os ha proporcionado.

Introducción.

En teoría se han instalado las bases de datos MySQL y PostgreSQL y se han analizado algunos de sus parámetros de configuración que pueden incidir en el rendimiento de las mismas.

Como complemento al trabajo visto en teoría se pretende, en primer lugar, aprender a utilizar herramientas que permitan realizar tests de rendimiento sobre bases de datos en general y MySQL y PostgreSQL en particular, y con posterioridad utilizar dichas herramientas para verificar el impacto que posee sobre el rendimiento de las bases de datos la modificación de sus parámetros de configuración.

Existen herramientas propias de cada base de datos, como pueden ser `mysql-test` o `postgresql-test` que permiten ejecutar test de rendimiento sobre MySQL o PostgreSQL respectivamente, pero ofrecen unos resultados muy pobres.

Una herramienta de software libre que permite ejecutar test de rendimiento sobre diversas bases de datos es HammerDB, que permite ejecutar test sobre base de datos Oracle, SQL Server, TimeTen, PostgreSQL, Greeplum, Postgres Plus Advanced Server, MySQL y Redis. El programa y su documentación se encuentran en la URL:

<http://hammerora.sourceforge.net>

El programa puede descargarse para Windows y Linux en 32 o 64 bits y posee una amplia documentación sobre su instalación, configuración y funcionamiento en la URL indicada con anterioridad¹.

El programa HammerDB debéis instalarlo en el propio sistema Centos 6.5, para evitar problemas de conexión, etc., entre el ordenador anfitrión y el sistema virtualizado. Este hecho hará que el test sobre la base de datos se vea influido por el uso del propio sistema para ejecutar el programa que realiza el test de rendimiento, pero podemos considerar despreciable ese efecto².

¹ En nuestro caso deberemos descargar la versión de 32 bits para Linux.

² El efecto puede no ser despreciable, pero vamos a considerarlo despreciable con el fin de simplificar el trabajo a desarrollar.

Objetivo de la práctica.

El objetivo de la práctica es utilizar la herramienta HammerDB para comprobar el rendimiento de las bases de datos MySQL y PostgreSQL instaladas en nuestra máquina virtual.

Inicialmente deberéis realizar un test con la configuración por defecto y con posterioridad, debéis modificar parámetros de configuración de las bases de datos para ver como influyen dichos parámetros en el rendimiento de las mismas. Para evitar interacciones entre los parámetros, la modificación deberá realizarse parámetro a parámetro, esto es, se modifica un parámetro, se comprueba su influencia sobre el rendimiento de la base de datos y se vuelve al valor inicial antes de modificar otro parámetro.

No es necesario que modifiquéis y comprobéis todos los parámetros vistos en clase, basta con modificar un par de parámetros en cada base de datos y realizar las pruebas de rendimiento oportunas.

En la documentación del programa HammerDB veréis que es posible modificar el número de usuarios concurrentes que efectúan el test, etc., por lo que podéis cambiar parámetros del test para comprobar como afectan su cambio al rendimiento de las bases de datos³.

Por último, se os recuerda que debéis rearrancar la base de datos para que se cargue con los parámetros que hayamos modificado.

Memoria a entregar.

Debéis entregar un documento PDF en el que se refleje de forma clara y concisa la configuración de la base de datos, la configuración del test de rendimiento que se ha ejecutado y el rendimiento del mismo. Siempre que sea posible deberá mostrarse la información de forma gráfica con una breve explicación si se considera conveniente.

Es importante para que los test tengan sentido que se efectúen con la máquina virtual ejecutándose en el mismo ordenador anfitrión, pues las características del mismo influirán obviamente en el resultado del test. Además procurar que el ordenador anfitrión no ejecute otras tareas simultáneamente a la ejecución de los tests para intentar evitar interferencias en el uso del disco, etc.

Por último, indicaros que los valores absolutos de los tests no son importantes, pues obviamente dependerán del hardware sobre el que se ejecuta la máquina virtual, por lo que lo importante es la relación existente entre el test inicial y los test obtenidos una vez modificados algunos parámetros de la base de datos.

³ Esto puede ser importante en algunas modificaciones de parámetros de la base de datos que tengan que ver con número de usuarios concurrentes, etc.