

Resumen ejecutivo

Desde la aparición de las primeras computadoras, el almacenamiento y procesado de información ha sido una de sus tareas fundamentales. Los Sistemas Gestores de Bases de Datos son los encargados de satisfacer estas tareas cuando la complejidad de la información o de su tratamiento así lo requiere. Los siempre crecientes requisitos de las aplicaciones que utilizan bases de datos, generan una permanente necesidad de rendimiento, necesidad que se ha atacado tradicionalmente desde dos frentes: refinamiento continuo del *software*, y nuevos diseños de *hardware* con mayor capacidad o potencia.

En este proyecto he estudiado el rendimiento del Sistema Gestor de Bases de Datos MySQL desde los dos frentes expuestos, observando la interacción entre la aplicación y la máquina sobre la que se ejecuta, con un doble fin: observar el aprovechamiento que se hace del *hardware*, y documentar la utilización de la jerarquía de memoria como referencia para el diseño de futuras arquitecturas.

Para alcanzar tales fines, he realizado un estudio exhaustivo de las bibliotecas libres de medida de eventos del procesador a través de contadores *hardware*. Así mismo, he probado y analizado diferentes *benchmarks* de código abierto, buscando una carga realista y representativa, siendo finalmente TPC-C el seleccionado.

He realizado un largo proceso de optimización de los parámetros del servidor, que ha resultado en una configuración del sistema con un rendimiento similar al que obtiene el grupo de *benchmarking* de MySQL. Gracias a esto, he podido realizar las mediciones en un entorno en el que la práctica totalidad del trabajo es útil.

Finalmente, y gracias a las optimizaciones realizadas, he sido capaz de caracterizar MySQL con carga TPC-C para tamaños de base de datos inferiores a la memoria utilizada. Los resultados indican un buen aprovechamiento del procesador, con una sorprendentemente baja tasa de fallos en cache, que hacen que el sistema pueda suponerse efectivo para esta carga.