

1. Suponiendo que un sistema de bases de datos relacional que presenta un read-heavy workload y los queries son muy diferentes, explique detalladamente ¿por qué el uso de caches puede afectar el rendimiento del sistema de forma negativa? (30 pts)

El objetivo del cache es tener a la mano la información que más se está utilizando y mejorar el rendimiento en consecuencia, de esta forma no se tendría que estar accediendo a la capa de subyacente de almacenamiento que es mucho más lenta. Pero en el caso supuesto anteriormente no tendría ningún sentido y no se estaría cumpliendo el objetivo del cache que sería rapidez, pues tener un read-heavy y consultas con datos muy diferentes no habría ningún patrón de datos a guardar que fuera consultado con mayor frecuencia, en este caso se tendría que estar yendo a la memoria principal a consultar y adicionalmente el cache almacenaría datos aleatorios que puede ser que no se vayan a utilizar en otras consultas. En este caso, el cache no sería útil y bajaría el rendimiento.

2. El particionamiento de tablas en bases de datos relacionales es un concepto muy parecido al de shards en bases de datos NoSQL, explique detalladamente ¿cómo afecta el particionamiento y el sharding en el rendimiento de bases de datos SQL y NoSQL? (30 pts)

SQL: El particionamiento consiste en que tablas muy grandes sean divididas en varias más pequeñas. Teniendo eso en mente, con respecto al rendimiento este se vería afectado de manera positiva si tenemos un supuesto en donde, consultas que solo acceden a una fracción de los datos pueden correr más rápido porque hay menos datos que se deban recorrer. Pero obviamente se debe conocer bien la base de datos y el tipo de consultas que se puedan dar, para hacer esas particiones de manera correcta.

NoSQL: En este caso el sharding, sería algo muy parecido al particionamiento. Consiste en dividir grandes volúmenes de datos en partes más pequeñas. Lo principal es que el sharding pudiese brindarnos velocidad, no importaría si por ejemplo, se tiene un servidor en distintas zonas del mundo pues el sharding lo que hace es reducir el volumen de latencia y que el rendimiento sea mejor. Por otro lado, igualmente lo complicado sería saber cuándo se debe utilizar y bajo qué límites aprovecharlo.

3. En un sistema de bases de datos con Strong Consistency cuyo workload es de read-heavy y write-heavy, ¿cómo afectan los exclusive locks el rendimiento de las bases de datos NoSQL? (20 pts)

En este caso los exclusive locks son cerraduras que afectan especialmente a las transacciones de escritura, pues están hechas para bloquear el acceso de una transacción T2, mientras la transacción T1 está ejecutándose. En el supuesto anterior en donde hay una consistencia fuerte y su workload es de lectura y escritura, esto se vería afectado de manera negativa en su rendimiento pues al tener una consistencia fuerte, el read como en el write se verían perjudicados pues se debe asegurar que todas las réplicas tengan el mismo contenido por lo que se aplicarían cerraduras también a las lecturas.

4. Explique detalladamente, ¿cómo afecta la selección de discos físicos el rendimiento de una base de datos SQL y NoSQL? (20 pts)

Teniendo en cuenta que los discos físicos nos podían beneficiar más que tener los datos moviéndose en red hay que tener en consideración que los discos de estado sólido son más rápidos que los HDD, pero son más caros. En este caso se podría tener de los

dos tipos y dividir la información que más se utilice en el más rápido, para tener un mejor rendimiento.