

Prueba Corta #3/4 Bases de datos II

Isaac Araya Solano 2018151703

1. Comente de qué forma un query como:

SELECT col1, col2, ..., colN FROM table;

N es menor al (número total de columnas /2).

Se puede ver beneficiado por un columnar storage (25 pts).

Al estar haciendo un SELECT de la mitad de las columnas de una tabla, la información de la otra mitad de columnas no es necesaria. Si se almacenara la información en filas, para acceder a cada columna que solicito tendría que “traer” la información de las columnas que no son necesarias también. En cambio, al almacenar la información por columna, puedo “traer” la información solamente de las columnas que se necesitan y así no se accede a datos innecesarios, incrementando la eficiencia de la consulta.

En resumen, la utilización de un almacenamiento columnar puede mejorar significativamente el rendimiento de este tipo de consultas, reduciendo el tiempo de respuesta y la carga en el sistema de almacenamiento.

2. Una base de datos MariaDB o MySQL se encuentra bajo mucha presión debido a un alto número de lecturas (heavy read), de qué forma se podría lograr bajar esta presión sobre la base de datos, justifique su respuesta (25 pts).

Si una base de datos MariaDB o MySQL tiene un alto número de lecturas lo que se puede hacer es agregar al sistema réplicas de lectura, o read replicas. Estas ayudaran a aliviar la carga ya que le proporcionarán a los usuarios otros servidores a los cuales conectarse al hacer una lectura y liberará la carga de lecturas de cada servidor repartíendola entre más réplicas. Además, al ser servidores de solo lectura podrán dedicar todo su poder computacional a manejar lecturas y no tendrán que manejar transacciones o escrituras de ningún tipo más que las que son para mantener la consistencia de la base.

3. Durante la semana santa 2023 un grupo de estudiantes de Bases de Datos 2 decidieron desarrollar un juego retro de la plataforma Atari llamado Breakout, el mismo en su versión original es de un sólo jugador, pero se toma la decisión de implementar una versión multi-jugador, el backend se encuentra almacenado en una base de datos NoSQL que se encuentra ubicada en Azure Cloud en una región en Estados Unidos, el juego se volvió muy popular en menos de una semana, tanto que una gran cantidad de usuarios en otros continentes se han estado quejando de retrasos en el juego lo cual causa una mala experiencia, gracias a las herramientas de observabilidad configuradas, se ha descubierto que la naturaleza del workload es pocas escrituras y muchas lecturas a los mismos datos, presente dos recomendaciones (con su debida explicación) desde el punto de vista de base de datos que daría para mejorar el rendimiento de la misma (50 pts).

1) Al tener problemas por los tiempos de respuestas de personas de otros continentes lo que se puede hacer es crear réplicas de la base de datos en distintas Availability Zones más cercanas a los usuarios que se están quejando y distribuidas alrededor de los continentes. Esto haría que los usuarios puedan tener respuestas más rápidas al tener los servidores más cerca, además de que las escrituras se harían también de forma óptima.

2) Además, en esas AZ se pueden crear múltiples replicas de Lectura. Estas replicas serían de especial utilidad ya que se identificó que la principal carga de trabajo de los servidores es de lectura. Esto agilizaría los procesos de lectura haciendo que el rendimiento general del juego para los usuarios sea mejor y que los tiempos de respuesta disminuyan.