

Práctica de Examen

Bases de Datos II

Luis Diego Delgado Muñoz

Prof. Nereo Campos

Fecha de Entrega: 19/05/2024

Vamos a abordar cada una de las preguntas del archivo, proporcionando respuestas detalladas y fundamentadas.

Pregunta 1

Solución para mejorar el rendimiento de la base de datos actual

Para mejorar el rendimiento de la base de datos actual en MariaDB sin causar un downtime significativo, se pueden implementar las siguientes estrategias:

1. **Indexación:** Crear índices adecuados para las columnas más consultadas. Dado que las consultas de lectura son altas, aplicar la indexación puede mejorar significativamente el tiempo de respuesta.
2. **Particionamiento:** Dividir las tablas grandes en particiones más pequeñas puede mejorar el rendimiento al facilitar la gestión y acceso a datos.
3. **Optimización de Consultas:** Revisar las consultas SQL para asegurarse de que están optimizadas. Utilizar EXPLAIN para entender cómo MariaDB está ejecutando las consultas y ajustar en consecuencia [1].

Recomendación de tipo de base de datos

Para un problema de esta naturaleza, recomendaría migrar a una base de datos NoSQL como MongoDB o una base de datos en la nube incluso. Un par de opciones serían:

1. **MongoDB:** Ideal para manejar datos semi-estructurados y permite una fácil escalabilidad horizontal.
2. **Elasticsearch:** Para mejorar las capacidades de búsqueda, Elasticsearch podría ser integrado para gestionar búsquedas rápidas y eficientes utilizando índices invertidos.

Cada uno de estos servicios permite escalabilidad y rendimiento superior al que actualmente se presenta.

Conveniencia de mantener la base de datos actual en la casa de un fundador versus moverla a un proveedor de nube

Mantener la base de datos en la casa de un fundador no es una opción ideal debido a que las soluciones en la nube como AWS ofrecen alta disponibilidad y redundancia. Los proveedores de la nube permiten una fácil escalabilidad tanto vertical como horizontal y sumado a eso también ofrecen medidas de seguridad avanzadas que son difíciles de implementar en una infraestructura local.

Reducción del memory footprint utilizando un índice invertido y Stemming

Para reducir el memory footprint, se puede implementar un índice invertido y utilizar Stemming con Elasticsearch para reducir las palabras a su raíz común. Esto permite almacenar menos variaciones de palabras y realizar búsquedas más rápidas y eficientes.

Pregunta 2

Impacto de los índices en el rendimiento de las bases de datos relacionales

Algunos de los beneficios que presentan los índices es que estos permiten un acceso más rápido a los datos, además de que las búsquedas y recuperaciones de datos se realizan de manera más eficiente. Sin embargo, los índices deben ser actualizados cada vez que los datos subyacentes cambian, lo que puede ralentizar las operaciones de escritura. Sumado a eso, los índices requieren almacenamiento adicional.

Incluso si el hardware no es un problema, no se deben crear índices innecesarios, ya que pueden degradar el rendimiento de las operaciones de escritura y ocupar espacio adicional.

Pregunta 3

Impacto de los componentes de hardware y software en el rendimiento de la base de datos

Los componentes de hardware y software afectan el rendimiento de las bases de datos de la siguiente manera:

1. **CPU:** Afecta la velocidad de procesamiento de las consultas.
2. **Memoria RAM:** Influye en la capacidad de la base de datos para mantener los datos en caché, mejorando el acceso a datos frecuentemente consultados.
3. **Disco de Almacenamiento:** Afecta la velocidad de lectura y escritura de datos. Los discos SSD son preferibles por su alta velocidad.
4. **Sistema Operativo:** La eficiencia del sistema operativo en la gestión de recursos y procesos afecta el rendimiento general de la base de datos.

Pregunta 4

Importancia de la Observabilidad y Métricas

La observabilidad es crucial para la escalabilidad automática, permitiendo monitorear y ajustar el rendimiento del sistema en tiempo real.

Importancia de la Observabilidad:

1. **Monitoreo en Tiempo Real:** Permite identificar y resolver problemas de rendimiento rápidamente.
2. **Predicción de Comportamiento Futuro:** Analizar las métricas ayuda a prever comportamientos que se deben de tratar antes de que puedan ser un problema.

Dentro de las métricas que se necesitan para asegurar que se presente una escalabilidad automática adecuada se encuentran los siguientes:

1. **Memoria:** Para asegurarse de que hay suficiente memoria para manejar las cargas actuales y futuras.
2. **CPU:** Para evaluar si el procesamiento es adecuado para las operaciones actuales.

3. **Disco:** Para asegurarse de que las operaciones de I/O no se conviertan en un cuello de botella.

Sin embargo, también sería beneficioso monitorear otras métricas como latencia de red, tiempos de respuesta de consultas, entre otras.

Bibliografía

[1] K. MariaDB, "EXPLAIN," MariaDB KnowledgeBase, <https://mariadb.com/kb/en/explain/> (accessed May 18, 2024).