

1. Explique cómo afectan los siguientes componentes el rendimiento de un sistema de base de datos:

a. Disco: Es de los componentes más económicos, pero también es de los más lentos, ir a traer información a disco es muy costoso en términos de recursos, por lo que, si los datos se encuentran en disco y cada vez hay que ir a traerlos, el consumo de recursos será muy alto en comparación a si los datos estuvieran en memoria. Además, cuando toda la información está en disco, provoca que se tengan que hacer muchos context switch y esto provoca muchas interrupciones.

b. Memoria Virtual: La memoria virtual simula a la memoria principal en disco, trae consecuencias, porque si bien, la memoria virtual mejora rendimiento, puede hacer que se degrade el intercambio entre la memoria principal y la virtual. Por otro lado, genera una carga más en el disco y es más lenta que la memoria principal, por lo que el acceso a los datos también es más lento.

c. Memoria: Es muy rápida, sobre todo en comparación con el disco, sin embargo, es muy cara. Tener datos en memoria hace que sea más rápido acceder a estos y cuesta menos recursos, sin embargo, no se puede tener todos los datos en memoria porque es limitada y como se mencionó anteriormente, es muy costosa. La memoria también tiene modo privilegiado, donde el Sistema Operativo puede matar aplicaciones.

d. Caché de CPU: De esta depende el aprovechamiento de la velocidad que brinda el CPU. El CPU le pide información al cache debido a que este guarda instrucciones, si la información se encuentra en cache hace que se ahorre el viaje a memoria, esto se conoce como cache hit, cuando los datos o información solicitada no está, debe ir a buscar la

información, ocasionando interrupciones. El cache también permite el almacenamiento de consultas y resultados, aunque al ser limitada, no es lo más recomendable, pero puede usarse para reutilizar la información en los casos cuando las consultas no cambian.

e. CPU: El CPU “gobierna” aspectos de la computadora, es de los más importantes. Al igual que la memoria, es bastante costoso, pero es el más rápido. El CPU permite paralelismo, aunque esto depende del core. Un CPU multicore permite n cores, que permiten n usuarios a la vez y, por ende, n niveles de paralelismo, es decir, si un CPU tiene dos cores y cada uno puede correr dos hilos, puede haber cuatro procesos simultáneos.

2. ¿De qué forma se benefician las aplicaciones del uso de caches? Explique.

El cache es muy rápido y permite que el CPU pueda acceder a los datos almacenados en cache, para así no tener que ir a ver si están en memoria, esto permite que la CPU pueda hacer uso de su velocidad, debido a que la memoria cache también es bastante rápida y le ahorra ir a disco en busca de información. Todo esto permite que las aplicaciones sean más rápidas y hace que no se tenga que estar interrumpiendo procesos para buscar información.

3. Desde el punto de vista de Elasticsearch, ¿Que es un índice?

. El índice en Elasticsearch se puede ver como una colección de documentos que consisten en colecciones de campos llave valor (tipo diccionario), además, estos consisten en estructuras optimizadas. Los índices permiten que las búsquedas sean rápidas y eficientes.

4. ¿Qué es un mapping en Elasticsearch?

Al cargar un documento a Elasticsearch, se hace un mapping dinámico, este consiste en una detección automática de los campos, haciendo más fácil la indexación y exploración de datos, el mapping puede ser personalizado, permitiendo tomar la decisión de cómo guardar o indexar campos, esto permite hacer un tipo de formato que, en algunos casos, es bastante útil al implementar análisis de textos.