

Prueba Corta #5 y #6

Edgar André Araya Vargas 2020142856

1. Explique en que consiste unclustered index y cuál es la diferencia entre este y un índice non-clustered que utiliza INCLUDE para agregar columnas al índice.(25pts)

Los índices en las bases de datos relacionales mejoran el rendimiento de las consultas al permitir una búsqueda rápida y eficiente de los datos. Hay dos tipos de índices principales: los índices no agrupados y los índices agrupados. Un índice no agrupado almacena las claves de los registros y un puntero al registro en la tabla que contiene el resto de la información. Por otro lado, un índice agrupado también almacena las claves de los registros, pero los datos se almacenan físicamente en la misma estructura de índice, ordenados por la clave de índice. En cuanto a la diferencia entre un índice no agrupado y uno que utiliza INCLUDE, el último tipo también incluye en la estructura del índice los valores de las columnas que se han agregado con INCLUDE. Esto significa que las consultas que solicitan esos valores pueden ser respondidas completamente a través del índice, sin tener que buscar en la tabla subyacente, lo que puede mejorar significativamente el rendimiento de las consultas que requieren datos de esas columnas.

¿Cuál es la relación entre un memory footprint alto y la paginación a disco? (25 pts)

Cuando un programa tiene un memory footprint alto, puede ocurrir que el sistema operativo no tenga suficiente memoria disponible y deba recurrir a la paginación a disco para liberar espacio y permitir la ejecución de otros programas. La paginación a disco es un procedimiento mediante el cual se trasladan los datos de la memoria RAM a un disco duro, lo que puede ser más lento y afectar negativamente el desempeño del programa. Me parece que un memory footprint elevado puede impactar negativamente en la rapidez y el rendimiento del programa, y generar la necesidad de recurrir a la paginación a disco en caso de que la memoria disponible sea insuficiente.

¿Que tipo de base de datos recomendaría a TooSlow para almacenar sus datos? (50pts)

En este específico caso que maneja una gran cantidad de registros y necesita agilizar las búsquedas en ellos, la solución que le recomendaría a la empresa TooSlow y a mi parecer la mas clara seria usar una base de datos no relacional en este caso podria ser Cassandra. Mi recomendacion proviene de la idea de que Cassandra esta diseñada para manejar grandes cantidades de datos y proporcionar un alto rendimiento y escalabilidad en la recuperación de datos. Tambien permite el uso de varios nodos. Esto aumenta la disponibilidad del sistema y la tolerancia a fallas. Ahora para organizar la informacion seria sensato utilizar índices en las columnas de país, ciudad y fecha para permitir búsquedas más eficientes, incluso quiza utilizar un orden de datos en arbol B. Estos tipos de estructuras son más eficientes cuando se navega por grandes conjuntos de datos y se pueden aplicar técnicas de partición para dividir los datos en segmentos más pequeños para una recuperación más rápida. A mi parecer una base de datos SQL no seria la respuesta correcta para esta pregunta ya que necesita ingresar y trabajar con gran cantidad de registros, si se quisiera buscar en multiples columnas se necesitaria hacer complicados queries que necesiten de muchos recursos y las SQL no trabajan tan bien al manejar datos no estructurados como los documentos JSON en la columna de payload.