

Materia: ADMINISTRACIÓN DE BASES DE DATOS

Maestro: Gerardo Rodríguez Rojano

Fecha: 22 de Mayo 2019

T A R E A N O . 9

Fecha de Entrega: 1 de junio 2019;

Modalidad: En equipos de 2 a 3 integrantes.

Forma de entregar: subir este archivo con las respuestas en cada pregunta en formato PDF al campus virtual del curso.

Datos de Integrantes:

Expediente	Nombre Completo (nombre y apellidos)
258849	Michael Brandon Serrato Guerrero

INSTRUCCIONES:

1. Leer el documento del curso titulado:
 - a. Indices en Bases de Datos.
 2. *Contestar el cuestionario* de esta tarea no 9.
-

Fuente: Indices en Bases de Datos

1. ¿Por qué crear un índice en una tabla de base de datos acelera los tiempos de acceso a los datos?

R: Cuando se crean índices en una tabla, al momento de realizar una consulta el Sistema Gestor de Base de Datos va a ese índice primero en una estructura de datos tipo árbol B+ o tabla de hash (**según sea el tipo de índice esto ya descarta una gran cantidad de datos**), obtiene el número de bloque de disco y un "puntero" para poder desplazarse dentro del bloque y una vez ahí recuperar los registros correspondientes directamente. Algo sin duda muy eficiente ya que al no hacer uso de los índices, el DBMS se ve forzado a recorrer todos los registros de la tabla para recuperar los registros deseados (un proceso conocido como "escaneo de tablas" que resulta ser muy lento).

2. Explique la diferencia entre un índice Primary Key y un índice UNIQUE.

R: En ambos índices los valores contenidos por el campo deben ser estrictamente únicos, sin embargo, en el caso de **los índices UNIQUE es posible ingresar valores nulos**, además de que **es posible definir uno o varios de ellos en la misma tabla**, por otro lado, los índices **PRIMARY KEY no permiten valores nulos en ningún momento y sólo existe un índice de este tipo por tabla**.

3. Explique la diferencia entre un índice Primary Key y un índice secundario.

R: Mientras que un índice **PRIMARY KEY** sirve para identificar de forma única los registros dentro de una tabla (sin permitir valores nulos) y teniendo presente sólo uno de estos índices por tabla, los índices **SECUNDARIOS** además de poder estar presentes en múltiples campos dentro de una misma tabla, permiten almacenar valores repetidos en sus campos, **su principal función es la de lograr acelerar las búsquedas de los registros que contengan el mismo valor, básicamente es una forma de agrupar ciertos registros que cumplan la misma condición tomando como ventaja la selectividad que estos tengan**.

4. Explique las características de un índice clúster.

R: Son índices que instruyen al DBMS para que ordene la tabla físicamente mediante el valor del campo, además de ello sólo puede haber un índice clúster por cada tabla debido a que por lógica las filas de los datos sólo pueden estar almacenadas de una forma.

5. ¿Por qué en sistemas de bases de datos que se utilizan para Data Warehouse se recomienda poner en cada tabla todos los índices posibles?

R: Debido a que los datos almacenados por los Data Warehouse son obtenidos por medio de cargas masivas de datos provenientes de sistemas transaccionales (generalmente por procesos nocturnos en batch) no hay transacciones (Insert, Update o Delete) de tal forma que **sólo se realizan consultas de tipo SELECT sobre ellos y es aquí donde el uso de índices juega un papel importante acelerando el acceso a grandes cantidades de datos sin tomar en cuenta lo ineficiente que podría ser para los otros tipos de consultas en donde se modifican los datos.**

6. ¿Por qué se prefiere escoger llaves primarias que tengan un tipo de dato numérico entero?

R: Porque al momento de indexar o de comparar los índices de los datos es más eficiente hacerlo por un tipo de dato numérico que por cualquier otro tipo de dato que pudiera ser más extenso (como por ejemplo un tipo “varchar”) debido a la simplicidad y ahorro de memoria que esto implica.

7. El documento explica que cuando la métrica de selectividad del índice tiene valores entre 1% y 10% se dice que el índice fue altamente selectivo; ¿por qué? Dé un ejemplo.

R: La selectividad es la relación que existe entre el número de registros que satisfacen una condición y el número total de registros en una tabla. Se dice que un índice tiene mayor efectividad cuando los valores de columna son altamente selectivos. Por lo que si un índice tiene valores entre 1% y 10% la cantidad de registros obtenidos será notablemente menor a la cantidad total de registros en la tabla logrando así una consulta mucho más rápida.

Un ejemplo podría ser una tabla de alumnos aspirantes a una beca estatal en donde la columna "promedio" es un índice cuyo valor es de tipo entero entre 0 y 100 (para evitar decimales), suponiendo que la beca sólo se entrega a aspirantes con promedio mayor o igual a 90 (9.0) entonces la consulta de los beneficiarios debería retornar un conjunto del 10% de los registros existente en la tabla, cumpliendo así la propiedad de ser un índice altamente selectivo.

F I N
