PRIMER EXAMEN PARCIAL 18/03/2021

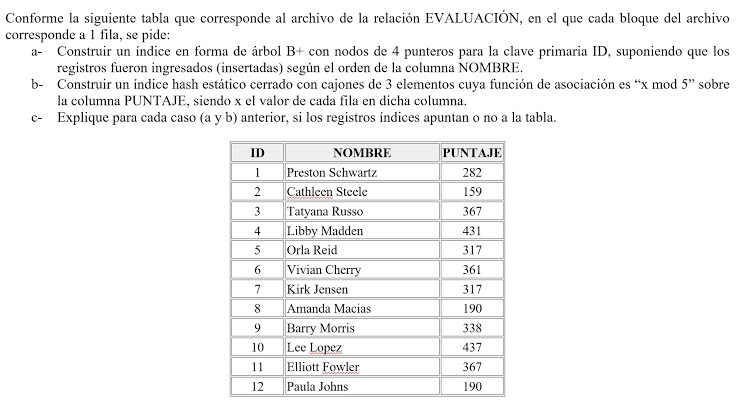
Indique para que tipo de consultas resulta más apropiado: (a) un índice ordenado, (b) un índice hash, (c) un índice bitmap

5/5

(a) Para búsquedas por rango de datos.

(b) Para búsquedas en condición de igualdad (mismo valor hash).

(c) Para búsquedas por varias claves en simultáneo.



Explique en sus términos las formas organización de archivos en posibles para la implementación de archivos de tablas/datos.

5/5

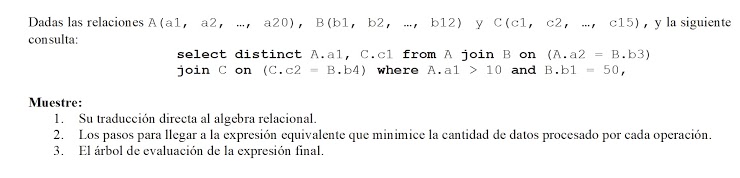
- Registros de longitud fija: Cada registro posee una longitud igual a los demás y se insertan al final.

- Lista libre: Cada registro tiene una cabecera sin datos pero que apunta a la siguiente cabecera vacía.

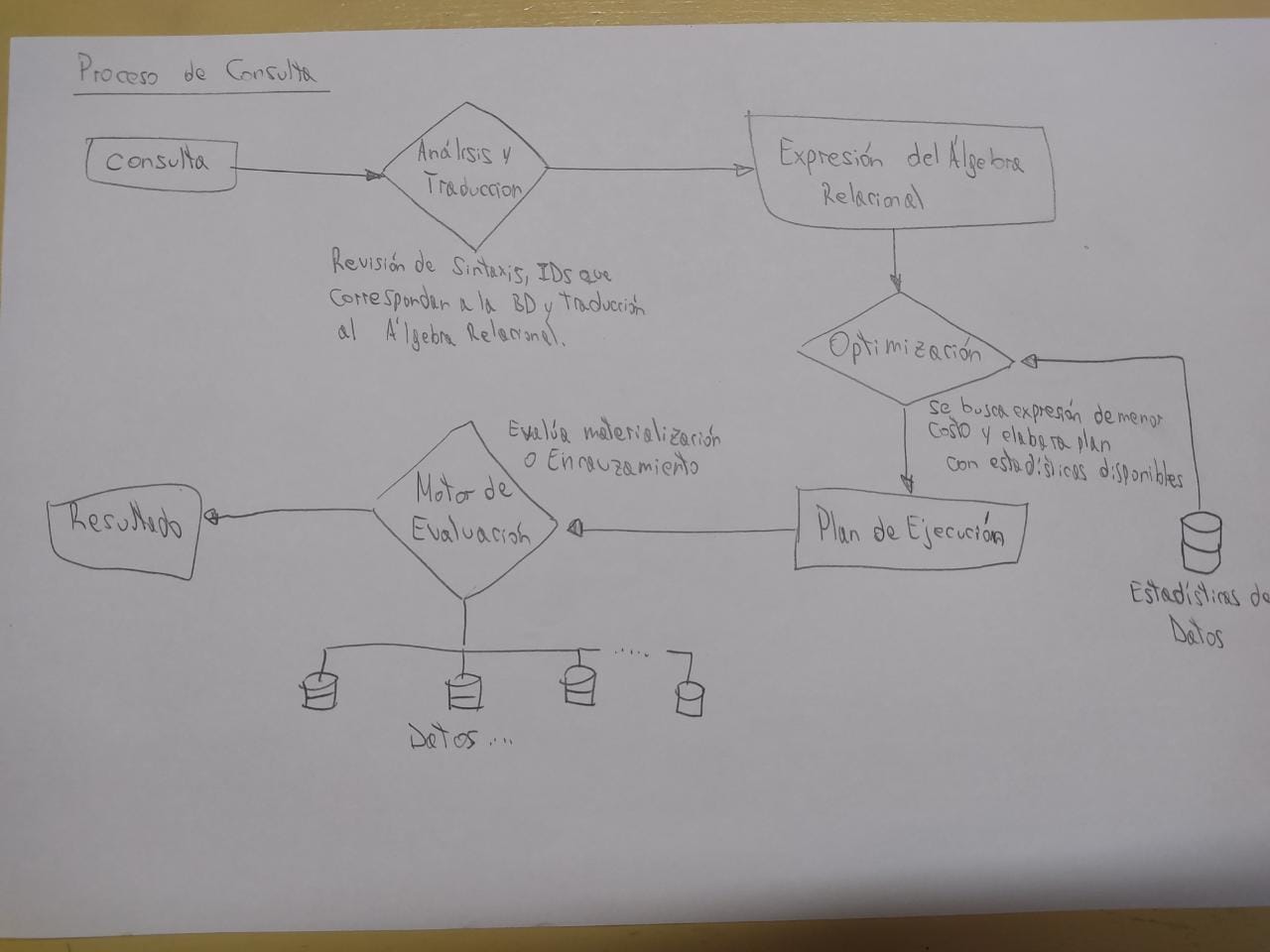
- Registros de longitud variable: Cada registro posee longitud variable y en general un espacio libre disponible.

- Archivos secuenciales: Registros se introducen uno detrás del otro.

- Organización en agrupación: Se juntan varias columnas que se suelen solicitar en conjunto para recuperarlas más rápido.



Grafique el Diagrama de Procesamiento de consultas y explique los pasos del mismo.



Explique los niveles de RAID 0, 1 y 5. A su criterio, ¿cual es la principal razón para la implementación de unidades de almacenamiento redundante?

3/5

RAID 0: Sin redundancia.

RAID 1: Redundancia de 1 disco redundante por cada disco existente.

RAID 5: Redundancia por paridad distribuida en diferentes discos.

La principal razón es la búsqueda de un sistema de backup lo más eficiente posible dado que el hardware puede fallar por el desgaste.

Explique las medidas de rendimiento que deben tener en cuenta para la elección de unidades de discos magnéticos. ¿A su criterio, cuál de las mismas sería la más determinante?

5/5

- Tiempo de acceso: Tiempo que lleva leer o escribir en el disco. Depende del tiempo de búsqueda y de la latencia rotacional.

- Tasa de transferencia: Cantidad de datos que pueden escribirse o leerse por unidad de tiempo.

- Tiempo de Fallo (TMDF): Tiempo de duración del disco hasta empezar a tener fallos de hardware. A mi criterio, lo más importante sería hoy día el tiempo de fallo, ya que el tiempo de acceso y la tasa de transferencias ya se encuentran prácticamente al límite físico de lo que permite esta tecnología.

Indique para que tipo de consultas resulta más apropiado: (a) un índice ordenado, (b) un índice hash, (c) un índice bitmap

5/5

(a) Para búsquedas por rango de datos.

(b) Para búsquedas en condición de igualdad (mismo valor hash).

(c) Para búsquedas por varias claves en simultáneo.

Explique el costo asintótico espacial y temporal del Algoritmo de Programación Dinámica: (a) sin la optimización de evaluación de arboles en profundidad por la izquierda. (b) con la optimización de evaluación de arboles en profundidad por la izquierda

Dada una operación de selección cualquiera basada en una condición de igualdad, explique en qué casos un SGBD utilizará los siguientes algoritmos e indique cual es el costo asintótico de los mismos: (a) Búsqueda Lineal. (b) Búsqueda Binaria. (c) Búsqueda en índice primario para un atributo clave. (d) Búsqueda en índice secundario para atributo no clave.

3/5

(a) Cuando se tiene una tabla desordenada. Costo = br (cantidad de bloques)

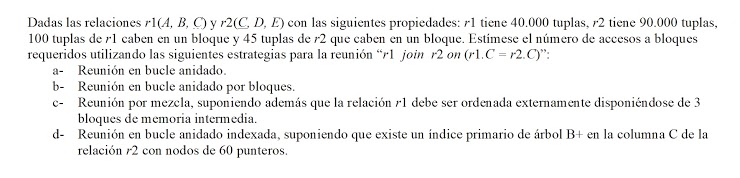
(b) Tabla está ordenada físicamente. Costo = log2(br) + CS(A,r)/fr -1 (localizar un bloque + bloques subsecuentes)

(c) Cuando el índice se contruye por un atributo clave en la tabla. Costo = AAi + 1 (recorrido en el árbol + lectura de bloque)

(d) Cuando no se tiene ningún índice de clave ni atributo clave Costo = AAi + 1 + CS(A,r)

Comentarios individuales: *d) la explicación del ítem es incorrecta.*

Indique los resultados del siguiente ejercicio.



SEGUNDO EXAMEN PARCIAL 20/05/2021

(a) Explique que es una Planificación Secuencial. (b) Explique que es una Planificación Secuenciable. (c) Fundamente a su criterio la importancia de la Secuencialidad en SGBD.

4/5

a) Una planificación secuencial es aquella cuyas operaciones transaccionales se ejecutan de seguido, una tras otra. Se considera correcta y se usa para el manejo de concurrencia.

b) Es una planificación que puede se transformada a una secuencial equivalente.

c) La secuencialidad permite a los SGBDs ordenar las transacciones concurrentes, a modo de poder asegurar que las operaciones cumplan con los principios de atomicidad, consistencia, aislamiento y persistencia para con los datos. Sin la secuencialidad, tendríamos importantes cuellos de botella, lecturas sucias y varios otros problemas costosos de solucionar.

(a) Defina el concepto de Transacción en SGBDs. (b) Indique y defina las fases del Ciclo de Vida de una Transacción. (b) Indique y defina las propiedades que un SGBD debe asegurar para las Transacciones.

4/5

a) Una transacción en un SGBD es un conjunto de operaciones concurrentes de manipulación de datos de la BD.

b) Activa: Estado inicial de la ejecución.

Parcialmente confirmada: Última operación ejecutada. Control de concurrencia.

Fallida: Si falla un chequeo o si es abortada.

Confirmada: Transacción finalizada de forma correcta.

Abortada: Terminación abrupta de la transacción y vuelta al estado anterior a la ejecución.

Terminada: Estado final.

c) Atomicidad: O se realiza toda la transacción o no se realiza nada.

Consistencia: La BD debe pasar de un estado consistente a otro al realizar la transacción.

Aislamiento: La transacción actual no debe ser visible a otras transacciones antes de finalizarse.

Persistencia: Al confirmarse los cambios hechos por la transacción, estos no deben perderse posteriormente.

(a) Describa la función del Componente de Gestión de Concurrencia en un SGDB. (b) Describa la estructura de datos utilizada para la gestión y concesión de bloqueos.

5/5

a) El gestor de concurrencia comprende a los métodos de acceso, administradores de búfer, de bloqueos y de logs. Se encarga de aplicar los protocolos adecuados para las operaciones concurrentes mediante la concesión de bloqueos sobre datos.

b) El gestor de bloqueos maneja una Tabla de Bloqueos, en la cual se almacenan los bloqueos concedidos y las peticiones solicitadas. Se trata de una tabla hash en memoria cuyos elementos son listas enlazadas con registros de bloqueos y peticiones.

(a) Describa el protocolo de Bloqueo de 2 Fases. (b) Indique cuales son las variantes del mismo.

5/5

a) Las transacciones realizan peticiones de bloqueo y desbloqueo. Hay una fase de crecimiento (subir) en la cual se obtiene un bloqueo exclusivo y una fase de decrecimiento (bajar) en la que se liberan bloqueos (compartidos). Inicialmente una transacción está en fase de crecimiento, con lo que se mantiene secuencialidad al ir cambiando el tipo de bloqueo entre transacciones.

b) Las variantes del protocolo de bloqueo de dos fases son: el estricto y el riguroso. El primero libera los bloqueos exclusivos al finalizar la transacción, y el segundo libera todos los bloqueos al final de la transacción.

(a) Detalle el protocolo de Control de Concurrencia de Quorum de Consenso. (b) Explique las implicancias de la condición que debe tenerse en cuenta para la definición de los valores de Quorum. (c) Indique los valores de Quorum que permiten emular el Protocolo de Mayoría y el Protocolo Sesgado.

3/5

a) Cada sitio de la BDD tiene un peso determinado, y la suma de todos los pesos es S. Se tiene para cada dato Q un quórum de lectura (Qr) y uno de escritura (Qw), los cuales deben cumplir las condiciones: Qr + Qw > S y 2\*Qw > S. Para realizar lectura en uno de los sitios, se deben bloquear un número de réplicas tal que la suma de sus pesos sean >= Qr. Para escritura, de manera similar, se deben bloquear las réplicas tal que la suma de sus pesos sean >= Qw.

b) Debe tenerse en cuenta que al ser un protocolo generalizado, el mismo puede ajustarse para, con ciertos valores Qr y Qw, llegar a casos de otros protocolos.

c) Para emular el protocolo de Mayoría se requieren Qr y Qw iguales a (S/2 +1 ) y para Protocolo Sesgado Qr=1 y Qw=S.

Comentarios individuales: *b) incorrecto*

(a) Detalle los pasos de cada fase del protocolo C2F en Bases de Datos Distribuidas. (b) Indique como proceden los participantes en caso de falla del coordinador. (c) Indique como se procede un sitio en caso de falla del mismo.

0/5

a)

b) En caso de falla del coordinador, los sitios participantes deciden qué hacer en base a su bitácora y se comunican entre si.

c) En caso de fallo, se debe rehacer o deshacer T, dependiendo de si está comprometido o abortado, respectivamente.

Explique detalladamente las formas de almacenamiento distribuido en Bases de Datos Distribuidas Relacionales.

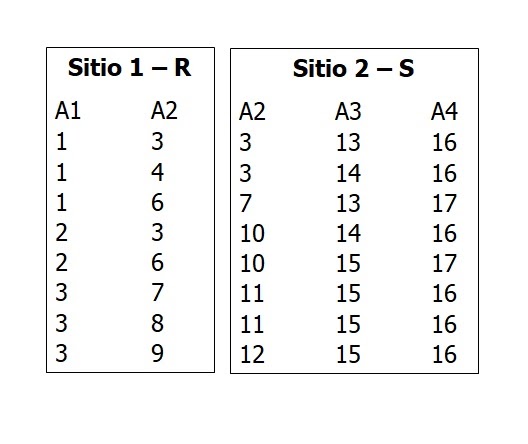
3/5

Replicación: Varios sitios mantienen copias de la misma relación. Esto permite alta disponibilidad, consultas en paralelo y una transferencia de datros reducida. Así también representa un costo muy incrementado de actualización y mayor complejidad de control de concurrencia.

Fragmentación: Varias partes o trozos de una relación se distribuyen en diferentes sitios. Puede ser horizontal (conjuntos de filas de una relación) o vertical (columnas de la relación). Esto permite mejor uso del almacenamiento y menor costo de mantenimiento, así como también procesamiento en paralelo de distintos fragmentos y transparencia según el sitio. Ambas fragmentaciones pueden ser combinadas.

Comentarios individuales: *incompleto*

Conforme la figura de abajo y la consulta "select \* from R join S". Indique los pasos y el costo total de trasmisión de realizar la Estrategia de la Semireunión si la consulta fue recibida: (a) en el Sitio 1, (b) en el Sitio 2.



Explique en sus términos: ¿Cuál es la diferencia entre bases de datos OLTP y OLAP?. Así mismo, cite al menos dos ejemplos de aplicación para cada tipo.

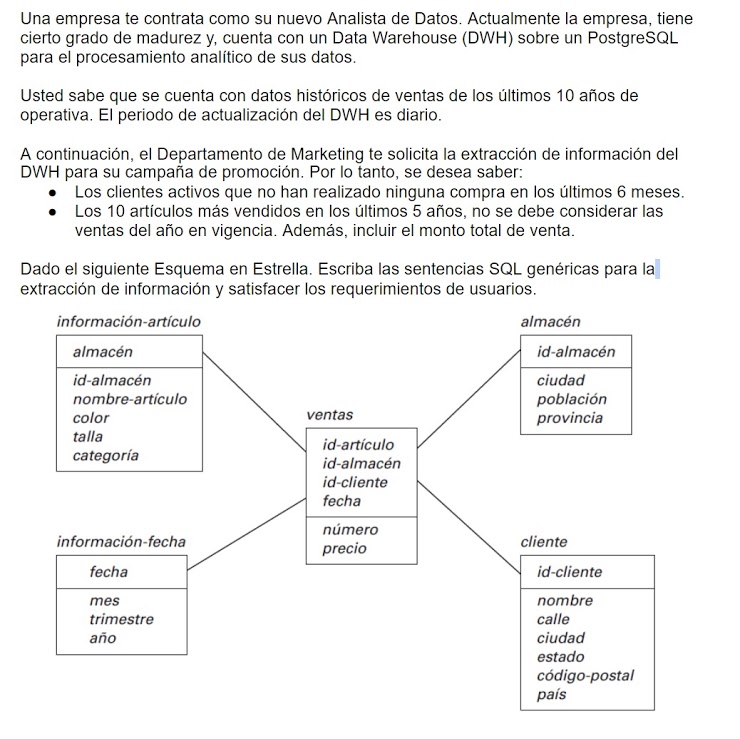
5/5

Las BDs OLTP son las utilizadas para las operaciones y transacciones de los datos de los procesos del negocio, mientras que las OLAP se utilizan para procesar datos e información para la toma de decisiones del negocio.

Ejemplos OLTP: Sistemas financieros de transacciones y administración de datos de compra/venta de una empresa.

Ejemplos OLAP: Business Inteligence y Data Mining.

Resuelva el siguiente planteamiento



a) SELECT c.id-cliente, c.nombre FROM cliente c LEFT JOIN (SELECT \* FROM ventas v WHERE id-cliente = c.id-cliente) WHERE c.estado = "activo" AND (DATEADD.MONTH - v.fecha.MONTH > 6)

b) SELECT \* FROM informacion-articulo a LEFT JOIN (SELECT \*FROM ventas v WHERE a.id-articulo)

Comentarios individuales: *Consultas SQL incompletas.*

Explique en sus términos en qué consisten las tablas de dimensiones, tablas de hechos y medidas en el modelado multidimensional. Así mismo, enumere al menos tres ejemplos para cada uno.

5/5

Tablas de dimensiones: Son las que restringen los rangos de los datos y brindan detalles adicionales para los usuarios. Múltiples tablas de dimensión se utilizan para definir los parámetros de los que dependen las tablas de hechos y medidas.

Ejemplos: Clientes, Proveedores, Productos, etc.

Tablas de hechos y medidas: Son los datos duros a ser analizados. Son la intersección de de sus dimensiones definidas.

Ejemplos: Ventas, Compras, Envíos, etc.

PRIMER EXAMEN FINAL 17/06/2021

