

FBD

Ejercicios de autocomprobación Raúl Castro Moreno 2ºB

Tema 4

1. Describe los medios de almacenamiento físico que hay actualmente en tu ordenador, indicando capacidad y velocidad de transferencia de cada uno de ellos.

| | Capacidad | Velocidad de transferencia |
|-------------------------------|-----------|---|
| Caché | 6 MB | L1: 1150 MB/s L2: 570 MB/s L3: 220 MB/s |
| Memoria Principal (DDR4-2400) | 16 GB | 19 200 MB/s |
| Disco Magnético (SATA 3) | 1 TB | 600 MB/s |
| Unidad de Estado Solida(SSD) | 256 GB | 2700 MB/s |

2. Explicad por qué la organización de los registros en bloques es un elemento importante dentro del rendimiento de una base de datos.

Porque no hay que recuperar ni almacenar la cantidad mínima de bytes posibles, sino que a la hora de almacenar, se traduzca el almacenamiento en las mínimas unidades de E/S secundarias, es decir, en la mínima cantidad de bloques de base de datos involucrados tanto en escritura, es decir, desde la memoria al almacenamiento secundario, como en recuperación, recuperando del almacenamiento secundario a la memoria asignada al SGBD.

3. Disponemos de una base de datos sobre profesores, asignaturas e impartición, almacenada como muestran la Figura 1 y la Figura 2, usando un registro por página. Complete la Figura 3 y la Figura 4 después haber realizado la siguiente secuencia de actualizaciones: Se inserta la asignatura A7, se suprime el profesor P1, se eliminan las imparticiones del profesor P1.

| | | | | | | | | | | | |
|----|----------------------|------------------------------------|----|------------------------------------|----|------------------------------------|----|------------------------------------|----|------------------------------------|----|
| 0 | X | 1 | 2 | 2 | 3 | 3 | 4 | 4 | 5 | 5 | 24 |
| | | A₁ | | A₂ | | A₃ | | A₄ | | A₅ | |
| 6 | 12 | 7 | 8 | 8 | 9 | 9 | 10 | 10 | 11 | 11 | X |
| | | P₂ | | P₃ | | P₄ | | P₅ | | P₆ | |
| 12 | 18 | 13 | 14 | 14 | 15 | 15 | 16 | 16 | 17 | 17 | 19 |
| | | A₁/P₂ | | A₁/P₃ | | A₁/P₄ | | A₁/P₅ | | A₁/P₆ | |
| 18 | 25 | 19 | 20 | 20 | 21 | 21 | 22 | 22 | 23 | 23 | X |
| | | A₃/P₂ | | A₃/P₆ | | A₄/P₂ | | A₄/P₄ | | A₄/P₅ | |
| 24 | X | 25 | 26 | 26 | 27 | 27 | 28 | 28 | 29 | 29 | X |
| | A₇ | | | | | | | | | | |

| 0 | X | | | | | | | | | | |
|---|-------------------|------------------|-------------------|----------------|---|-------------|---|------------|---|---------|----|
| <table> <tr> <th>Conjunto Páginas</th><th>Dirección 1ª Pág.</th></tr> <tr> <td>Páginas Libres</td><td>6</td></tr> <tr> <td>Asignaturas</td><td>1</td></tr> <tr> <td>Profesores</td><td>7</td></tr> <tr> <td>Imparte</td><td>13</td></tr> </table> | | Conjunto Páginas | Dirección 1ª Pág. | Páginas Libres | 6 | Asignaturas | 1 | Profesores | 7 | Imparte | 13 |
| Conjunto Páginas | Dirección 1ª Pág. | | | | | | | | | | |
| Páginas Libres | 6 | | | | | | | | | | |
| Asignaturas | 1 | | | | | | | | | | |
| Profesores | 7 | | | | | | | | | | |
| Imparte | 13 | | | | | | | | | | |

4. Dado que los índices agilizan el procesamiento de las consultas. ¿Por qué no se mantiene un índice para cada campo del fichero? Enumera tantas razones como se te ocurran.

Una razón es que ralentizan las demás operaciones como son inserción, borrado que requieren actualizar el archivo secuencial y todos los índices que están definidos sobre ese archivo. Otra razón es que si no se realizan consultas muy frecuentes sobre ese índice no tiene sentido mantenerlo. Hay que tener básicamente si va a ser útil o no ese índice para teniendo en cuenta su frecuencia de uso en consultas y su frecuencia a la hora de realizar operaciones de mantenimiento.

5. ¿Es posible tener dos índices primarios con diferentes claves sobre el mismo archivo almacenado?

No es posible debido a que un índice primario tiene almacenado el mismo número de registros que el archivo secuencial que indexa, y este índice primario guarda todas las claves del registro, 1 por cada registro y su RID. Por lo que este índice tiene ya todas las claves, por lo que no sería posible crear otro índice primario con diferentes claves puesto que tiene todas las claves ya el primer índice primario de ese mismo archivo.

6. ¿Por qué es preferible utilizar un índice no-denso a uno denso? ¿Por qué no se utilizan índices no densos para todas las claves de búsqueda?

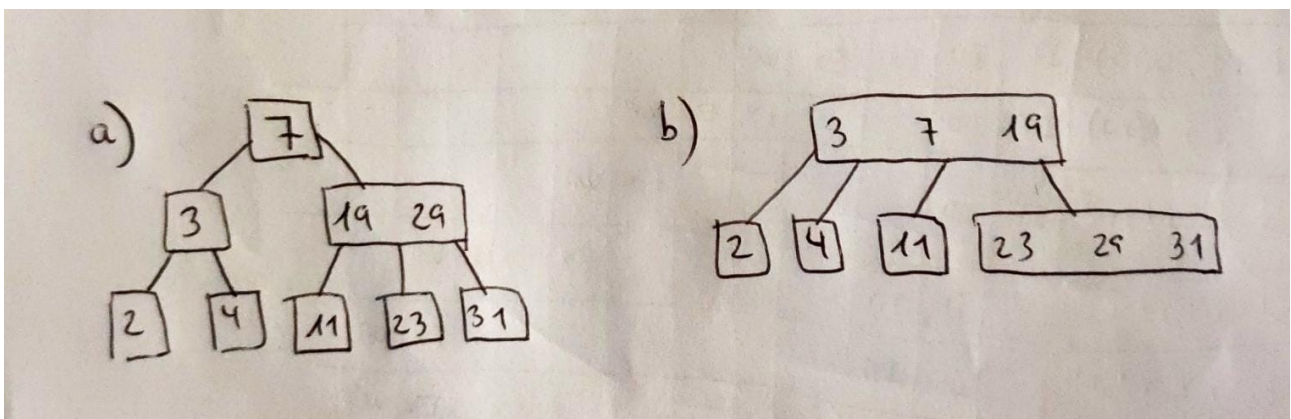
Porque se ahorra mucho tamaño a la hora de mantener el índice en la memoria principal debido a que los índices densos siguen siendo muy grandes. Con el índice no denso, tendremos un archivo que tendrá el mismo número de registros, que bloques tiene el fichero secuencial, reduciéndose mucho el tamaño.

Porque no nos permiten realizar una comprobación de existencia de una clave solo a través del índice puesto que en un no denso, tendrías que recuperar el bloque donde supuestamente está esa clave, mientras que un denso, bastaría con mirar el índice solo.

7. Construir un árbol B con el siguiente conjunto de valores de la clave:

a) Orden 3

b) Orden 4



8. Para el árbol del ejercicio anterior mostrad los pasos involucrados en las consultas siguientes:

a) Encontrar los registros con un valor de la clave de 11.

Para Orden 3: Se empieza desde el nodo raíz el cual es 7, como no hemos encontrado la clave 11, bajamos de nivel por su hijo derecho, debido a que 11 es mayor que 7. Nos encontramos con el nodo 19|29 en el cual tampoco esta la clave 11. Como tiene hijos, no paramos todavía el proceso y bajamos de nivel por el hijo izquierdo puesto que 11 es menor que 19. Nos encontramos en un nodo hoja el cual es 11. Hemos terminado.

Para Orden 4: Empezamos en el nodo raíz el cual es 3|7|19, como no esta 11 y tiene hijos, pasamos al siguiente nivel en el cual escogemos el tercer hijo empezando por la izquierda, es decir el hijo que esta entre 7 y 19 puesto que 11 es mayor que 7 y menor que 19. Llegamos a un nodo hoja el cual es el que tiene la clave 11. Hemos terminado.

b) Encontrar todos los registros con un valor de la clave entre 7 y 17.

Para Orden 3: Como en el apartado anterior, buscamos un registro el cual sea el que contenga el valor inferior o mas se aproxime a el. En este caso al ponernos en la raíz, encontramos que es 7 luego ya tenemos el nodo con el valor inferior y a partir de el recorremos todos los nodos hasta encontrar 17 o un valor mayor. Pero ademas si ese valor tiene hijos izquierdos hay que mirarlos puesto que pueden estar dentro del rango. Como ocurre en este arbol al bajar de 7 por su hijo derecho puesto que 17 es mayor que 7, encontramos el nodo 19|29 pero ahi no paramos aunque sea mayor que 17 puesto que tiene hijos izquierdos, por lo que recorremos ese hijo menor que 19 y encontramos 11. Finalmente concluimos que los registros entre ese rango son 7 y 11.

Para Orden 4: Igual que en el anterior, nos situamos en el nodo raíz, y encontramos 3|7|19, donde encontramos el valor inferior 7 y un valor superior a 17 el cual es 19, pero resulta que entre 7 y 19 hay un hijo, por lo que hay que recorrerlo, bajamos por ese hijo y encontramos el nodo hoja de 11. Ya hemos terminado concluyendo que los nodos entre ese rango son 7 y 11.

9. ¿Por qué una organización basada en el acceso directo no es la mejor idea para aquellos ficheros que se consulten por rangos de valores de la clave física, (entre c1 y c2)?

Esto es debido a por como es el acceso directo, la ubicación de los registros no va a ser contigua respecto al valor de la clave, debido a las transformaciones que realiza esta organización y además teniendo que hacer una búsqueda por cada clave.

10. Enumera las ventajas e inconvenientes que tienen, a tu juicio, la organización basada en índices y la que se basa en el acceso directo. Haz un análisis comparativo.

Organización basada en índices

Ventajas

- Disminuye el tiempo de acceso a los datos.
- Son más pequeños que los ficheros secuenciales a los que indexan.
- Las consultas por intervalos de claves son muy óptimas.
- Si son **no densos** el tamaño que ocupan es mucho más reducido y su mantenimiento es menos costoso.
- Si son **jerárquicos** se reduce el número de accesos al disco para localizar registros.

Inconvenientes

- Si no se tienen en cuenta la frecuencia de consulta o de las operaciones mantenimientos se ralentizan mucho las operaciones por mantener índices que se van a usar muy poco.
- Si el índice es **no denso** no nos permiten realizar una comprobación de existencia de una clave .
- Si el índice es **jerárquico**, se complica el mantenimiento del índice.

Organización basada en el acceso directo

Ventajas

- No se usa ninguna estructura adicional por lo cual no crearemos ningún archivo adicional.

Inconvenientes

- No es muy conveniente usarlo para aquellos ficheros que se consulten por rangos de claves puesto que la ubicación de los registros no van a ser contiguas respecto al valor de la clave.
- Se producen colisiones que hay que mirar con una gestión mediante listas de colisión
- Se producen huecos en el fichero de datos debido a las zonas no asignadas por el algoritmo en la salida.

11. Supongamos que se desea diseñar una estructura de "hashing" dinámico para un archivo que contiene los valores de clave:

$\{2, 3, 4, 7, 11, 19, 23, 29, 31\}$

Mostrad la organización que se obtiene si la función de direccionamiento es $f(x) = x \bmod 8$ y las páginas pueden contener hasta tres registros.

| Mod 8 (bin) | Nº | Proj local | Puntero |
|----------------|-----------|---------------|---------|
| 000 | | NA | X |
| 001 | | | X |
| 010 | 2 | 3 | Cubo B1 |
| 011 | 3, 11, 19 | 3 | Cubo B2 |
| 100 | 4, 29 | 2 | Cubo C |
| 101 | | | |
| 110 | 7, 23, 31 | 2 | Cubo D |
| 111 | | | |

12. Mostrad como cambia la estructura del ejercicio anterior como resultado de los siguientes pasos:

a) Borrar el 11 y el 31.

b) Insertar el 1, el 15, el 36, el 40, el 46, el 25.

a)

| Mod 7 (bin) | Nº | Proj Local | Puntero |
|----------------|----------|---------------|---------|
| 000 | 2, 3, 19 | 1 | Cubo A |
| 001 | | | |
| 010 | | | |
| 011 | | | |
| 100 | 4, 29 | 2 | Cubo B |
| 101 | | | |
| 110 | 7, 23 | 2 | Cubo C |
| 111 | | | |

b)

| Mod 8 (bin) | Nº | Proj Local | Puntero |
|-----------------------|-----------|---------------|---------|
| 000 | 1, 25, 40 | 2 | Cubo A |
| 001 | | | |
| 010 | 2, 3, 19 | 2 | Cubo B |
| 011 011 | | | |
| 100 | 4, 29, 36 | 2 | Cubo C |
| 101 | | | |
| 110 | 46 | 3 | Cubo D1 |
| 111 | 7, 15, 23 | 3 | Cubo D2 |

13. Supongamos que disponemos de una estructura de "Hashing" Dinámico que alberga hasta tres registros por página y que usa una función de dispersión $f(x)=x \bmod 8$.

Dibujad la organización resultante después de:

a) Insertar los registros con siguientes los valores para la clave:

{12, 16, 19, 26, 29, 32, 35, 41, 44, 64}

b) Eliminar de la estructura anterior el registro con valor 32

c) Insertar sobre la estructura anterior los registros con siguientes los valores para la clave: {22, 39, 46, 55}

a)

| Mod 8 (bin) | Nº | Prof Local | Puntero |
|----------------|------------|---------------|---------|
| 000 | 16, 32, 64 | 3 | Cubo A1 |
| 001 | 41 | 3 | Cubo A2 |
| 010 | 26, 19, 35 | 2 | Cubo B |
| 011 | | | |
| 100 | 12, 44, 29 | 2 | Cubo C |
| 101 | | | |
| 110 | | NA | X |
| 111 | | | X |

b)

| Mod 8 (bin) | Nº | Prof Local | Puntero |
|----------------|------------|---------------|---------|
| 000 | 16, 64, 41 | 2 | Cubo A |
| 001 | | | |
| 010 | 26, 19, 35 | 2 | Cubo B |
| 011 | | | |
| 100 | 12, 44, 29 | 2 | Cubo C |
| 101 | | | |
| 110 | | NA | X |
| 111 | | | X |

c)

| Mod 8 (bin) | Nº | Prof Local | Puntero |
|----------------|------------|---------------|---------|
| 000 | 16, 64, 41 | 2 | Cubo A |
| 001 | | | |
| 010 | 29, 19, 35 | 2 | Cubo B |
| 011 | | | |
| 100 | 12, 44, 29 | 2 | Cubo C |
| 101 | | | |
| 110 | 22, 46 | 3 | Cubo D1 |
| 111 | 39, 55 | 3 | Cubo D2 |