

Pauta de Corrección

CERTAMEN #1

ICI-343

Wenceslao Palma <wenceslao.palma@ucv.cl>

1. (5 pts c/u) Responda V/F, justifique.
 - (a) Un índice multinivel es un índice denso.
F. La idea de un índice multinivel es reducir el espacio ocupado por el índice, luego no puede ser denso.
 - (b) En un índice basado en hashing lineal con $d = 2$, $n = 3$, la clave 0011 se inserta en el bucket 00.
F. Si $n = 3$ el último bucket tiene entrada 11, luego según el algoritmo de inserción la clave 0011 se inserta en el bucket $m - 2^{d-1} = 11 - 10 = 01$
 - (c) Un índice basado en hashing lineal puede responder la consulta $100k \leq sueldo \leq 1000k$.
F. La generación de buckets no permite establecer un orden entre los registros.
 - (d) Un índice denso permite responder una consulta sin acceder al disco.
V. Cuando se pregunta por la existencia de un registro la respuesta se encuentra en el índice siempre y cuando la búsqueda se realice usando el atributo utilizado para construirlo.
2. (20 puntos) Se tiene un archivo de 200000 registros, cada uno de 250 bytes, el cual se almacenará en un disco con bloques de 1KB. La clave de cada registro la cual se utiliza para construir un índice basado en árbol B+ es de 9 bytes. Si los nodos se llenan a un 70% de su capacidad. Cuantos accesos se necesitan para:
 - (a) recuperar todos los registros del archivo.
 - (b) recuperar un registro específico.

R.:

$fb = 1024/250 = 4$ [registros/bloque]
 $\#bloques = 200000/4 = 50000$ [bloques]

tamaño del puntero = $\log_2(50000) = 2$ bytes (5 ptos.)

orden del árbol

$$2p + (9+2)(p-1) \leq 1024$$
$$p=79$$

(5 ptos.)

si un nodo se encuentra lleno en un 70% $p=55$

#niveles del árbol

	#punteros	#registros
1er nivel	55	54
2do nivel	55^2	55×54
3er nivel	55^3	$55^2 \times 54$
4to nivel	55^4	$55^3 \times 54$

- (a) para recuperar todos los registros se necesitan como mínimo 50000 accesos. (5 ptos)
- (b) para recuperar un registro específico se necesitan como máximo 5 accesos. (5 ptos)

3. Suponga un archivo de 7500000 páginas almacenado en un disco de 7200 RPM, seek de 10 ms y transferencia de 1ms por bloque de 4KB. Si se dispone de 54 páginas en memoria principal.

- (a) Determine cuál es el tiempo que se necesita para ordenar el archivo considerando:
- (i) (6 ptos) 53 buffers de entrada de 1 página c/u y 1 buffer de salida de 1 página.
- (ii) (6 ptos) 12 buffers de entrada de 4 páginas c/u y 1 buffer de salida de 6 páginas.

Comente los resultados obtenidos.

(b) (8 pts) Cuántos buffers se necesitan para ordenar el archivo en 2 pasos?

R.

(a)

$$\text{latencia} = 1/2 \text{ vuelta} \rightarrow 4.2 \text{ [ms]}$$

en general el tiempo necesario para ordenar un archivo es:

$$T = (1 + \log_{\text{bufIN}}(N/\text{total_buf})) \times (T_r + T_w) \times \text{\#Pags}$$

$$(i) T_r = T_w = 4.2 + 10 + 1 = 15.2 \text{ [ms]}$$

$$\text{\#pasos} = 1 + \log_{53}(7500000/54) = 1 + (\log(138889)/\log(53)) = 4$$

La cantidad de pasos depende de la cantidad total de buffers (54) y la cantidad de buffers de entrada (53)

$$T = 4 \times 2 \times 15.2 \times 7500000 = 909600 \text{ [seg]} = 253 \text{ [hrs]} \quad (6 \text{ pts})$$

$$(ii) T_r = 4.2 + 10 + (1 \times 4) = 18.2 \text{ [ms]}$$

$$T_w = 4.2 + 10 + (1 \times 6) = 20.2 \text{ [ms]}$$

$$\text{\#pasos} = 1 + \log_{12}(7500000/54) = 1 + (\log(138889)/\log(12)) = 6$$

La cantidad de pasos depende de la cantidad total de buffers (54) y la cantidad de buffers de entrada (12)

$$T = 6 \times (18.2/4 + 20.2/6) \times 7500000 = 356250 \text{ [seg]} = 99 \text{ [hrs]} \quad (6 \text{ pts})$$

(b) para ordenar el archivo en 2 pasos se requiere una cantidad de buffers de entrada suficiente para almacenar todas las particiones. De este modo se busca $B - 1 \geq N/B \rightarrow B > \sqrt{N}$. En este caso se necesitan $\sqrt{N} = 2740$ buffers. (8 pts)