

Pauta de Corrección

CERTAMEN #1

ICI-343

Wenceslao Palma <wenceslao.palma@ucv.cl>

1. (5 ptos c/u) Responda cada una de las siguientes preguntas.

- (a) Si se desea construir un índice basado en árbol B+ con punteros de 4 bytes y una clave de indexación de 12 bytes. Cuántos punteros y claves puede contener un bloque de 16384 bytes?

R:

| | |
|---------------------------|-------------------------------|
| | también se acepta: |
| $4p + 12(p-1) \leq 16384$ | $4p + (12+4)(p-1) \leq 16384$ |
| $16p \leq 16396$ | $20p \leq 16400$ |
| $p = 1024$ | $p = 820$ |

- (b) Considerando un índice basado en hashing lineal donde el último bucket es 101. En cuál bucket se debe almacenar la clave 11110?

R.: ya que $110 > 101$ el bucket donde se almacena 11110 es $m - 2^{d-1} = 110 - 100 = 010$

- (c) Cuál es la desventaja de los índices basados en hashing?

R.: no son eficientes para responder consultas tipo rango.

- (d) Bajo qué condiciones un índice tipo grid puede tener una gran cantidad de buckets vacíos?

R.: cuando existe una correlación entre los datos a indexar. Considerando un índice basado en dos dimensiones, la correlación entre los valores de los datos a indexar puede provocar que una gran cantidad de éstos se almacenen en buckets correspondientes a la diagonal.

2. (20 puntos) En un disco de 7200 RPM, seek de 8.9 ms, transferencia de 100 MB/seg y bloques de 4KB, se desea almacenar un archivo de 10.000.000 de registros de 100 bytes c/u.

- (a) (6 ptos) Si el contenido del archivo se encuentra desordenado, Cuanto es el tiempo de I/O máximo/mínimo requerido para recuperar un registro?

R.:

$$\text{latencia} = (60000/7200)/2 = 4.2 \text{ [ms]}$$

$$\text{factor de bloqueo fb} = \text{tamañoBloque}/\text{tamañoRegistro} = 4096/100 = 40$$

$$\text{numBloques} = \text{numRegistros}/\text{fb} = 10000000/40 = 250000$$

$$\text{transferencia} = (4096/(1024*1024*100))*1000 = 0.04 \text{ [ms]}$$

$$\text{Tmax} = (4.2 + 8.9 + 0.04)*250000 = 54.75 \text{ [min]}$$

$$\text{Tmin} = 4.2 + 8.9 + 0.04 = 13.1 \text{ [ms]}$$

- (b) (6 ptos) Si el contenido del archivo se ordena, Cuanto es el tiempo de I/O para recuperar un registro?

R.:

$$\begin{aligned} T &= (4.2 + 8.9 + 0.04)*\log_2(\text{numBloques}) = 13.1*\log_2(250000) \\ &= 13.1*17.9 = 234.5 \text{ [ms]} \end{aligned}$$

- (c) (8 ptos) Cuantas cintas se necesitan para respaldar todo el contenido del archivo? Considere cintas de 6250 bpp, 2400 pies, espacio entre bloques de 0.6 pulg, y factor de bloqueo=20.

R.:

$$\text{tamaño de un bloque} = (20*800)/6250 = 2.56 \text{ [pulg]}$$

$$\text{numBloques} = (2400*12)/(2.56+0.6) = 9113$$

$$\begin{aligned} \text{numCintas} &= \text{registrosArchivo}/(\text{numBloques}*\text{fb}) = 10000000/(9113*20) \\ &= 55 \end{aligned}$$

3. Considere un archivo de 25000 páginas y un disco que posee una latencia de 5 ms, seek de 10 ms y es capaz de transferir 1 página por ms.

- (a) Determine cuál es el tiempo que se necesita para ordenar el archivo considerando:

- (i) (6 ptos) 4 buffers de entrada de 4 páginas y 4 buffers de salida de 4 páginas.

$$\text{Tr} = 5+10+(1*4) = 19 \text{ [ms]}$$

$$\text{Tw} = 5+10+(1*4) = 19 \text{ [ms]}$$

$$\#\text{pasos} = 1+\log_4(25000/32) = 1+(\log(25000/32)/\log(4)) \approx 6$$

$$T = \#\text{pasos}*(\text{Tr}/4+\text{Tw}/4)*25000 = 1425000 \text{ [ms]} = 23.75 \text{ [min]}$$

- (ii) (6 ptos) Se utiliza 10 buffers de entrada de 2 páginas y 1 buffer de salida de 12 páginas.

$$T_r = 5 + 10 + (1 \times 2) = 17 \text{ [ms]}$$

$$T_w = 5 + 10 + (1 \times 12) = 27 \text{ [ms]}$$

$$\# \text{pasos} = 1 + \log_{10}(25000/32) \approx 4$$

$$T = \# \text{pasos} \times (T_r/2 + T_w/12) \times 25000 = 1075000 \text{ [ms]} = 17.91 \text{ [min]}$$

Comente los resultados obtenidos.

- (b) (8 ptos) Cuántas operaciones de E/S se realizan en cada caso ((i) y (ii))?

Operaciones de E/S = $2N \times \# \text{pasos}$

$$(i) \quad 2 \times 25000 \times (1 + \log_4(25000/32)) = 300000$$

$$(ii) \quad 2 \times 25000 \times (1 + \log_{10}(25000/32)) = 200000$$