

# PAUTA DE CORRECCION

## CERTAMEN #1

### ICI-343

Wenceslao Palma <wenceslao.palma@ucv.cl>

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

- (a) Si un archivo esta almacenado en 50000 bloques, tiene registros de 200 bytes c/u con un atributo clave de 11 bytes y el disco tiene bloques de 2KB, cuántos punteros contiene un nodo de un índice basado en árbol B+?

R.:

$$\text{Tamaño del puntero a bloque} = \lceil \log_2 50000/8 \rceil = 2$$

$$\text{Punteros de un nodo} = 2p + 11(p - 1) \leq 2048 \rightarrow p = 158$$

- (b) Cual es la ventaja del uso de buckets de indirección en un índice secundario?

R.:

Evita accesos a disco cuando se realizan consultas basadas en 2 o más criterios de selección de registros. Lo anterior se obtiene intersectando las listas de punteros asociadas a los buckets de indirección.

- (c) Considerando un índice basado en hashing lineal donde el último bucket es 110. En cuál bucket se debe almacenar la clave 110111?

R.:

$$\text{Ya que } 111 > 110 \text{ la clave } 110111 \text{ se debe almacenar en el bucket } m - 2^{d-1} = 111 - 100 = 011$$

- (d)Cuál es la principal diferencia entre un índice basado en Hashing Extendido y Hashing Lineal?

R.:

El uso de directorios y el uso de un factor de carga para gestionar el overflow.

2. Suponga un archivo de 16600000 registros, c/u de 200 bytes, almacenado en un disco con bloques de 4KB. La clave de los registros es de 4 bytes y el puntero a bloque es de 8 bytes.

- (a) (7 ptos) Cuánto tiempo se requiere para recuperar un registro: (i) si el archivo no se encuentra indexado?, (ii) si el archivo no se encuentra indexado pero es ordenado usando el valor de la clave de los registros?. Considere que un acceso a disco tarda 13 ms.

R.:

$$\text{Factor de bloqueo (fb)} = 4096/200 = 20$$

$$\text{Cantidad de bloques para almacenar el archivo} = 16600000/20 = 830000$$

$$\text{(i) Tiempo} = 13 \times 830000 = 10790000[ms]$$

$$\text{(ii) Tiempo} = 13 \times \log_2 830000 = 255.62[ms]$$

- (b) (7 ptos) Si se construye un índice basado en un árbol B+ con nodos ocupados en un 75% y la raíz se almacena en memoria principal, determine cuánto tiempo se requiere para recuperar un registro. Considere que un acceso a disco tarda 13 ms.

R.:

$$\text{Orden del árbol} = 8p + 4(p - 1) \leq 4096 \rightarrow p = 341$$

$$\text{Número de niveles} = \lceil \log_{256} 16600000 \rceil = 3$$

$$\text{Tiempo} = (2 + 1) \times 13 = 39[ms]$$

- (c) (6 ptos) Cuántas cintas se necesitan para respaldar todo el contenido del archivo? Considere cintas de 6250 bpb, 2400 pies, espacio entre bloques de 0.6 pulg, y factor de bloqueo=200.

R.:

Cantidad de bloques a almacenar =  $16600000/200 = 83000$

Tamaño de un bloque =  $(200 \times (200 \times 8))/6250 = 51.2$ [pulg]

Pulgadas necesarias para almacenar todos los bloques =  $83000 \times (51.2 + 0.6) = 4299400$ [pulg]

Luego, la cantidad de cintas es  $4299400/28800 = 150$

3. (20 puntos) Considerando las siguientes configuraciones para un algoritmo de ordenamiento externo:
- (a) un archivo de 10000 páginas y 3 buffers en memoria principal.
  - (b) un archivo de 20000 páginas y 5 buffers en memoria principal.
  - (c) un archivo de 2000000 páginas y 17 buffers en memoria principal.

Responda para cada caso:

- (i) (7 ptos.) Cuántas particiones se generan luego del primer paso del algoritmo?

R.:

(a)  $\lceil 10000/3 \rceil = 3334$

(b)  $\lceil 20000/5 \rceil = 4000$

(c)  $\lceil 2000000/17 \rceil = 117648$

- (ii) (7 ptos.) Cuántos pasos se necesitan para ordenar el archivo considerando un buffer de salida?

R.:

(a)  $1 + \lceil \log_2 3334 \rceil = 13$

(b)  $1 + \lceil \log_4 4000 \rceil = 7$

(c)  $1 + \lceil \log_{16} 117648 \rceil = 6$

- (iii) (6 ptos.) Cuántas operaciones de E/S se realizan durante todo el proceso de ordenamiento?

R.:

(a)  $2 \times 10000 \times 13 = 260000$

(b)  $2 \times 20000 \times 7 = 280000$

(c)  $2 \times 2000000 \times 17 = 68000000$