

Índice general

1. Nivel Interno	2
1.1. Medidas para evaluar un sistema de archivos	2
1.2. Registros y bloques	3
1.3. Organización de archivos y métodos de acceso	5
1.4. Evaluación del sistema	5

Nivel Interno

Recordemos que un sistema gestor de bases de datos (SGBD) tiene tres niveles distintos: interno, conceptual y lógico. En este tema nos vamos a centrar en el primero de ellos. El nivel interno está formado, a su vez, por otros dos niveles distintos:

- (a) Comunicación SO: asociado al sistema operativo donde se ejecute el SGBD.
- (b) Gestión de la información: asociado el SGBD en sí.

1.1. Medidas para evaluar un sistema de archivos

Al final, la información está almacenada en un disco duro, que tienen sistemas de ficheros asociados que les dan estructura. Hay que compararlos para elegir el más eficiente, para ello, tenemos una serie de parámetros:

Parámetro	Mide...
R	la memoria necesaria para almacenar un registro
T	el tiempo para encontrar un registro arbitrario
T_F	el tiempo para encontrar un registro por clave
T_W	el tiempo para escribir un registro cuando ya se tiene su posición
T_N	el tiempo para encontrar el siguiente registro a uno dado
T_I	el tiempo necesario para insertar un registro
T_U	el tiempo necesario para actuar un registro
T_X	el tiempo necesario para leer el archivo
T_Y	el tiempo necesario para reorganizar el archivo

1.2. Registros y bloques

Un SGBD almacena la información en tablas. La estructura de las tablas vienen determinadas por las columnas. La información de una tabla se almacena en filas. EL SGBD provee de una serie de tipos de datos para las columnas de una tabla (numéricos, enteros, etc).

Los ficheros en los que se almacenan las tablas de una base de datos se componen de un conjunto de bloques, que a su vez se componen de un conjunto de registros, que a su vez están compuestos por campos, y en cada campo se almacena un valor.

Los registros, según su longitud (R), pueden clasificarse en:

- (a) Longitud fija: todos los campos tienen longitud fija conocida a priori. Si denotamos por V_i a la longitud del campo i -ésimo, entonces la longitud del registro es:

$$R = \sum_i V_i$$

- (b) Longitud variable: hay por lo menos un campo del registro que tiene longitud variable. Con la siguiente notación

a) A : longitud media de los nombres de atributo.

b) V : longitud media de los valores de atributo.

c) a' : número medio de atributos.

d) s : número de separadores por atributo.

tenemos que la longitud del registro es:

$$R = a'(A + V + s)$$

Ejemplo 1.1. Calcular la longitud del siguiente registro de longitud variable:

Factura=325;	Línea=1;	Concepto=Análisis;	Cant=1;	Precio=300;
--------------	----------	--------------------	---------	-------------

$$\left. \begin{array}{l} A = (7 + 5 + 8 + 4 + 6)/5 = 6 \\ V = (2 + 2 + 8 + 2 + 6)/5 = 4 \\ s = 2 \end{array} \right\} \Rightarrow R = a'(6 + 4 + 2) = 12a'$$

Los **bloques** son las unidades de transferencia de información del disco a la memoria o viceversa (aunque en memoria se denominan *buffers*). El tamaño del bloque es fijo para toda la base de datos y es múltiplo del tamaño del bloque físico del SO (por eficiencia).

Según su estructura, los bloques se pueden clasificar en:

- (a) Estructura homogénea: todos los registros tienen la misma estructura, es decir, todos los registros son del mismo tipo y tienen el mismo número de campos.

- (b) Estructura heterogénea: los registros tienen distinta estructura (los registros tienen distinta longitud). Para usar este tipo, también necesitamos almacenar información sobre la propia estructura del registro.

La relación entre los registros en disco, los bloques del SO y los bloques del SGBD es:

$$\text{Registro disco} \Leftrightarrow \text{Bloque SO} \Leftrightarrow \text{Bloque SGBD}$$

Llamamos **factor de bloqueo** (Bfr), al número de registros que caben en un bloque y depende del tamaño del mismo bloque, B , y del tamaño de los registros, R . Este valor puede ser fijado a priori por el administrador del SGBD. Siempre incluye una cabecera **C**, con información útil para el sistema (referencias, fecha de actualización, etc), que tiene que restarse a B .

Denominamos **bloqueo** a la forma en la que se ajustan los registros a un bloque. Hay dos tipos de bloqueo:

- (a) Fijo o entero: se rellena el bloque con tantos registros como sea posible.

- a) Registro longitud fija:

$$Bfr = \left\lfloor \frac{B - C}{R} \right\rfloor$$

- b) Registro de longitud variable:

$$Bfr = \left\lfloor \frac{B - C}{R + M} \right\rfloor$$

M es el tamaño de las marcas de separación, ya que al ser de longitud variable, las necesitamos para diferenciar entre los registros.

En ambos se redondea hacia abajo ya que si no cabe un registro entero, no podemos guardar una fracción de él.

- (b) Partido o desencadenado: se escriben registros en un bloque hasta que no quede espacio. Cuando vayamos a insertar el último registro, puede caber entero o partirse en dos partes, cada una en un bloque distinto. Debido a eso, es necesaria la existencia de una referencia del bloque conteniendo el primer trozo, al bloque conteniendo el otro.

- a) Registro longitud fija:

$$Bfr = \left\lfloor \frac{B - C - P}{R} \right\rfloor$$

- b) Registro de longitud variable:

$$Bfr = \left\lfloor \frac{B - C - P}{R + M} \right\rfloor$$

P es el tamaño del puntero al siguiente bloque.

1.3. Organización de archivos y métodos de acceso

1.4. Evaluación del sistema