

Guía de Problemas Nro. 1: Archivos e Índices

Bases de Datos

1- Un archivo tiene 20.000 registros estudiante de longitud fija. Cada registro tiene los siguientes campos: NOMBRE (30 bytes), DNI (9 bytes), DIRECCIÓN (49 bytes), TELEFONO (9 bytes), FECHANACIMIENTO (8 bytes), SEXO (1 byte), CODIGODEPTO(4 bytes), CODCLASE(4 bytes, entero) y especialidad (3 bytes). Se utiliza un byte adicional como marcador de eliminación. EL archivo de almacena en un disco cuyo tamaño de bloque B=512 bytes.

- Calcule el tamaño de un registro en bytes.
- Calcule el factor de bloqueo del archivo y el número de bloques del archivo.
- Calcule el tiempo medio que tarda la búsqueda de un registro si se utiliza una búsqueda lineal en el archivo.
- Suponga que el archivo esta ordenado por DNI, calcule el tiempo que tarda una búsqueda binaria de un registro dado su número de DNI.

2- Un Archivo Componente con numComp como clave hash contiene registros con los siguientes valores claves: 2369-3760-4692-4871-5659-1821-1074-7115-1620-2428-3943-4750-6975-4981-9208 El archivo emplea 8 cubetas, numeradas de la 0 a la 7, donde cada cubeta corresponde a un bloque del disco que contiene dos registros. Cargar estos registros en el archivo en el orden dado, empleando la técnica de hashing $h(k) = K \bmod 8$.

3- Asuma que se está usando una tabla hash con 7 buckets de 1 registro cada uno, cada registro tiene los siguientes campos: NOMBRE, NUMERO, DEPORTE. Nuestra función hash se llamará h. Se desea almacenar los siguientes registros:

IDAAF # 1234 # futbol # h=5
LHC # 1244 # volei # h=6
LFF # 1202 # futbol # h=5
JKLM # 2333 # basquet # h=7
IFKN # 2441 # futbol # h=5

- Si los desbordes son tratados a través de direccionamiento abierto (derrame consecutivo) y si almacena los registros mostrados previamente, muestre como quedaría la tabla resultante.
- Si los desbordes son tratados por encadenamiento (cadena de desbordes), muestre como queda la tabla luego de insertar los registros.

4- Considere un disco con tamaño de bloques B=512 bytes. Un puntero a un bloque tiene P=6 bytes de longitud, y un puntero a registro tiene $P_R = 7$ bytes de longitud. Un archivo tiene 30.000 registros EMPLEADO de longitud fija.

Cada registro tiene los siguientes campos: NOMBRE(39 bytes), DNI(9 bytes), CODIGODEPTO(9 bytes), DIRECCION (40 bytes), TELEFONO(9 bytes), FECHANACIMIENTO(8 bytes), sexo (1 byte), CODIGOTRABAJO(4 bytes), SALARIO(4 bytes, real). Se utiliza un byte adicional como marcador de eliminación.

- Calcule el tamaño del registro R en bytes.
- Calcule el factor de bloqueo del archivo y el número de bloques del archivo.
- Suponga que el archivo está ordenado, según el campo clave DNI, y que se desea construir un índice primario sobre el DNI. Calcule el factor de bloqueo del índice. El número de entradas y bloques del primer nivel de

índice. El número de niveles que se necesitan si lo convertimos en un índice multinivel y el número de accesos necesarios para buscar y recuperar un registro del archivo.

d. Suponga que los registros no están ordenados y que se quiere construir un índice basado en el DNI. Haga los cálculos solicitados en el ítem c. para este caso y compárelos.

e. Suponga que el archivo no está ordenado por el campo CODIGODEPTO y deseamos construir un índice secundario sobre dicho campo, empleando un nivel de indirección. Supongamos que hay 1000 valores distintos de CODIGODEPTO y los registros empleados están distribuidos de manera uniforme entre esos valores. Calcule:

- I. El factor de bloqueo del índice
- II. El número de bloques necesarios para el nivel de indirección que almacena punteros a registros
- III. El número de entradas y de bloques del primer nivel de índice
- IV. EL número de bloques que necesitaremos si lo convertimos en un índice multinivel.
- V. El número total de bloques que requiere el índice multinivel y el nivel adicional de indirección
- VI. El número aproximado de accesos a bloque necesarios para buscar y recuperar todos los registros del archivo que tienen un valor determinado de CODIGO DEPTO empleando el índice.