

Actividad 1: Sistemas de almacenamiento en disco y dispersión interna

- Objetivos
 - Ser capaz de explicar las características de los sistemas de almacenamiento en disco más utilizados por los gestores de bases de datos, y sus ventajas e inconvenientes.
 - Conocer y comprender el uso de la dispersión para los ficheros internos.

- Descripción de la actividad:

1) De **modo individual** debe realizarse la **lectura y estudio** de los tipos de almacenamiento secundario más empleados por los SGBDs, así como las técnicas de dispersión interna.

La información debe ser extraída del libro:

Elmasri, R.; Navathe, S. Fundamentos de Sistemas de Bases de Datos. Addison-Wesley (5ª edición)

Concretamente, deberán estudiarse, dentro del *Capítulo 13: Almacenamiento en disco, estructuras básicas de ficheros y dispersión*, los apartados:

13.4 Ubicación de los registros de fichero en disco

13.5 Operaciones sobre ficheros

13.6 Ficheros de registros desordenados (ficheros heap)

13.7 Ficheros de registros ordenados (ficheros ordenados)

13.8.1 Dispersión interna

2) **Realizar los ejercicios** cuyos enunciados se muestran a continuación.

Presentación de la actividad: Deberán subirse resueltos los ejercicios planteados escritos a mano (único medio aceptado) antes del **miércoles 20/09/2023 a las 21:00 h.**

Escanear o fotografiar las páginas y construir UN ÚNICO PDF.

Actividad 1: Sistemas de almacenamiento en disco y dispersión interna

EJERCICIOS A REALIZAR

1 – Supongamos un fichero con 12,000 registros de longitud fija. Cada registro tiene los siguientes campos y tamaños:

CÓDIGO	NOMBRE	DIRECCION	TLF	FECHA NACIM	SEXO	HORARIO	GRUPO	AULA
6 bytes	20 bytes	50 bytes	9 bytes	8 bytes	1 byte	2 bytes	2 bytes	4 bytes

Se utiliza 1 byte adicional como marcador de eliminación.

(a) Calcular el tamaño de un registro R en bytes.

(b) Calcular el factor de bloqueo bfr y el número de bloques del fichero b suponiendo una organización no extendida. El tamaño del bloque B es de 512 bytes.

2- Supongamos que solo el 75% de los registros anteriores tienen un valor en el campo TLF, el 60% en HORARIO, y el 85% en AULA, y que se usa un fichero de registros de longitud variable. Los campos no mencionados tienen valores para todos los registros.

Cada registro tiene un *tipo de campo* (un número entero que identifica el campo) de 1 byte por cada campo incluido en el registro, más el marcador de eliminación de 1 byte y un marcador de fin de registro de 1 byte.

Supongamos que se utiliza una organización extendida de registros, en la que cada bloque tiene un puntero de 4 bytes al siguiente bloque (este espacio no se utiliza para almacenar registros).

(a) Calcular el tamaño medio de un registro R en bytes.

(b) Calcular el número de bloques necesarios para el fichero, teniendo en cuenta que el tamaño del bloque es de 512 bytes.

3 – Supongamos un fichero CLIENTES con $Numcli$ como clave de direccionamiento calculado, que contiene registros con los siguientes valores de $Numcli$: 7107, 844, 1540, 4800, 826, 7110, 1821, 2376, 2002, 4981, 962, 2084, 2306. Los registros se almacenan en un array de tamaño 8.

Cargar estos registros en el array en el orden dado, empleando la función de direccionamiento calculado $h(K) = K \bmod 8$.

- Resolver las colisiones utilizando la técnica de encadenamiento.
- Resolver las colisiones por direccionamiento abierto.
- Resolver las colisiones por **dispersión múltiple** aplicando una segunda función de dispersión $h_2(K) = K \bmod 2$.