

Tema 4

El Nivel Interno



Apartado 1

Introducción

Contenidos

- Introducción
- Método de Acceso a la BD almacenada
- Representación de la BD en el nivel interno

Contenidos

- Introducción
- Método de Acceso a la BD almacenada
- Representación de la BD en el nivel interno

Introducción

Importancia de la organización de los datos

Una BD sirve para ...

almacenar de forma permanente grandes cantidades de datos.

Con el propósito principal de ...

gestionar de forma eficiente los datos y su almacenamiento para poderlos recuperar/consultar cuando sea necesario.

Esto tiene sus consecuencias tanto en la organización lógica de los datos, como en su organización física.



Introducción

Nivel Interno

Nivel Interno

- Expresa en última instancia, las operaciones sobre los datos (creación, alteración y recuperación) en términos de actuación sobre unidades mínimas de almacenamiento denominadas páginas o bloques de base de datos.
- Provee al administrador de mecanismos para optimizar el almacenamiento y el acceso a los datos.
- Se encuentra implementado en el SGBD.

Introducción

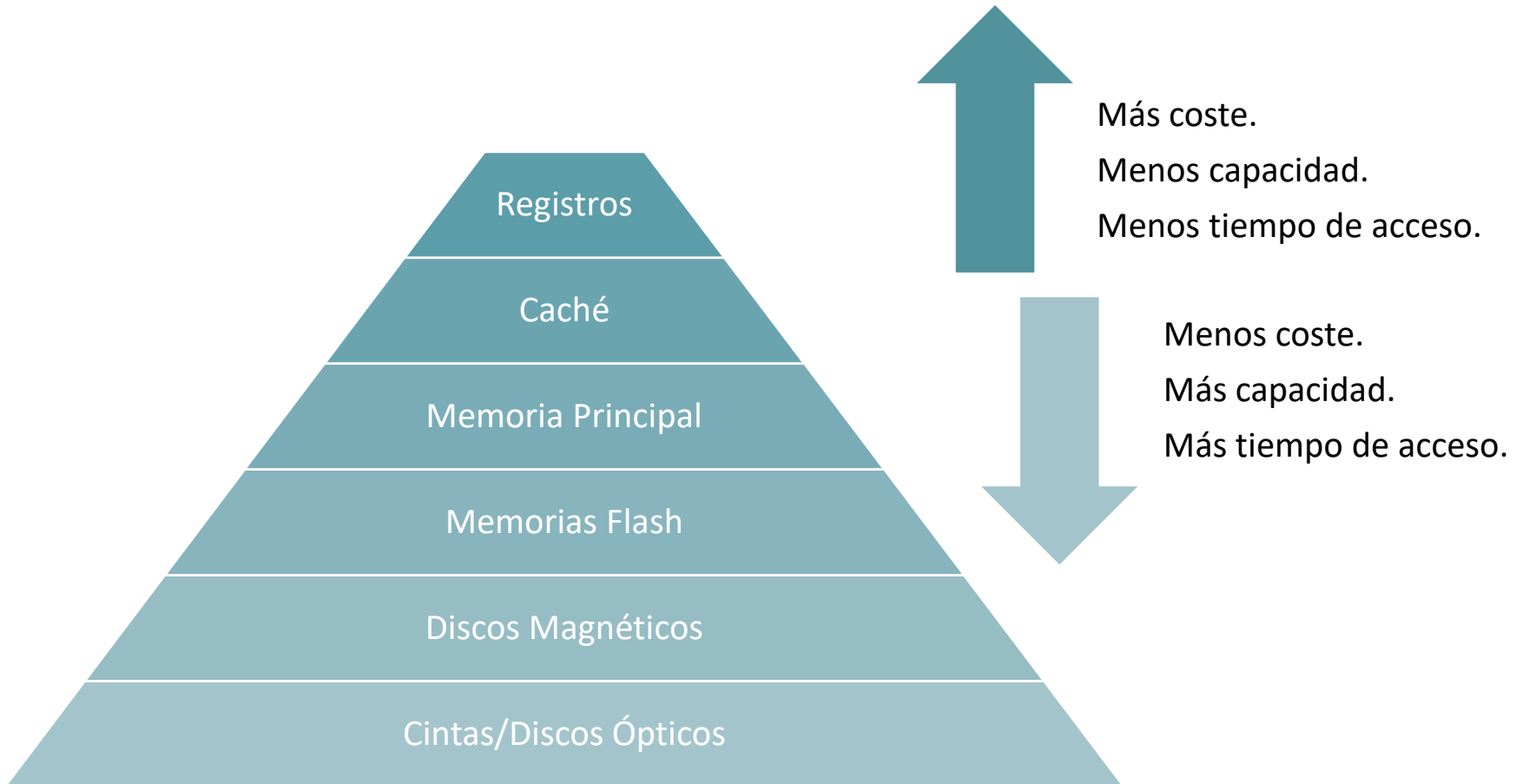
Nivel Físico

Nivel Físico

- Se encuentra implementado en el Sistema Operativo.
- Proporciona al SGBD una capa de abstracción sobre el hardware.
- Realiza el acceso a los medios de almacenamiento mediante llamadas a los servicios del sistema de archivos proporcionado por el SO.

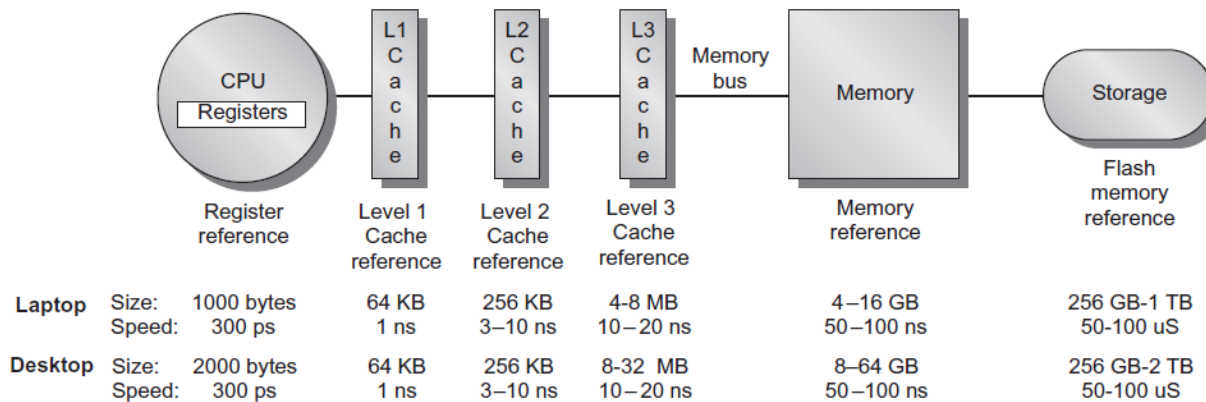
Introducción

Jerarquía de memoria

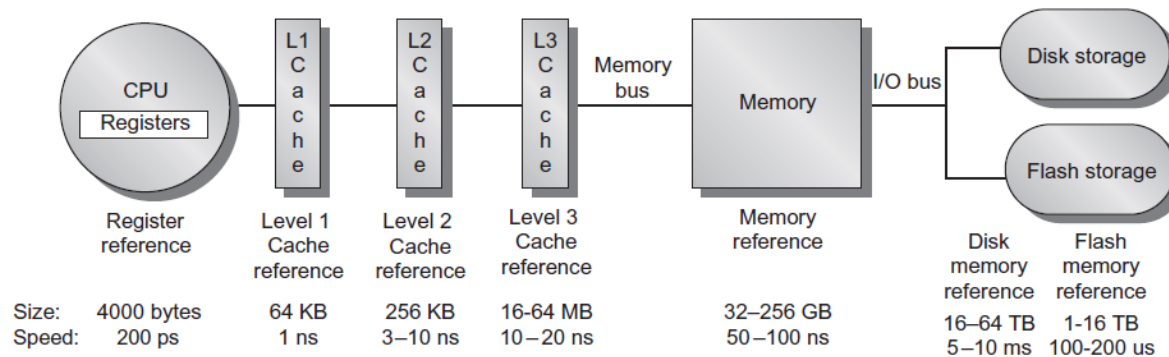


Introducción

Jerarquía de memoria



(B) Memory hierarchy for a laptop or a desktop



(C) Memory hierarchy for server

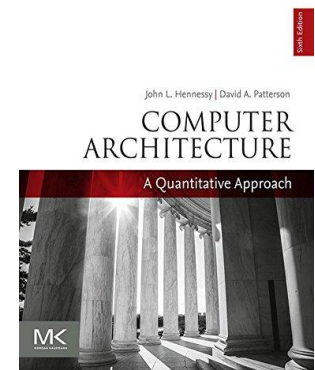


Figura tomada de John. L. Hennessy, David A. Patterson: Computer Architecture. A Quantitative Approach, 6th Edition, Morgan Kaufmann, 2017.

Introducción

Memoria principal

- Es el dispositivo de almacenamiento primario de los ordenadores.
 - Hace trabajos de caché de la porción de la BD de uso más reciente.
 - Elemento de almacenamiento intermedio que ubica de forma temporal los datos afectados por las operaciones.
- El nivel interno debe optimizar su uso:
 - Es rápida: acelera el procesamiento
 - Es cara: hay que optimizar su uso
- Es volátil, su información se pierde cuando se apaga el sistema o se cae. El nivel interno debe garantizar que la información contenida tenga un respaldo en almacenamiento secundario para que no se pierda.
- Se utilizan distintos niveles de caché para acelerar el acceso a los datos.

Introducción

Discos duros



- Dispositivo de almacenamiento más usado en BD.
 - Conjunto de discos magnéticos con dos caras.
 - Cada cara tiene un conjunto de pistas concéntricas (cilindro: la misma pista de todas las caras).
 - Cada pista se divide en sectores con la misma capacidad de almacenamiento (bloque).
- Localización de un bloque:
 - cilindro
 - superficie de disco
 - sector
- Parámetros:
 - Capacidad
 - Tiempo medio de acceso
 - RPM
 - Velocidad sostenida de lectura/escritura.

Imágenes tomadas de <https://pixabay.com/>

Introducción

- Medidas de rendimiento:
 - Tiempo medio de acceso (t_a): tiempo medio transcurrido entre una instrucción y la obtención de la información
 - Tiempo medio de búsqueda (t_b): tiempo medio de posicionamiento en pista
 - Tiempo de latencia rotacional (t_l): tiempo medio de posicionamiento en sector

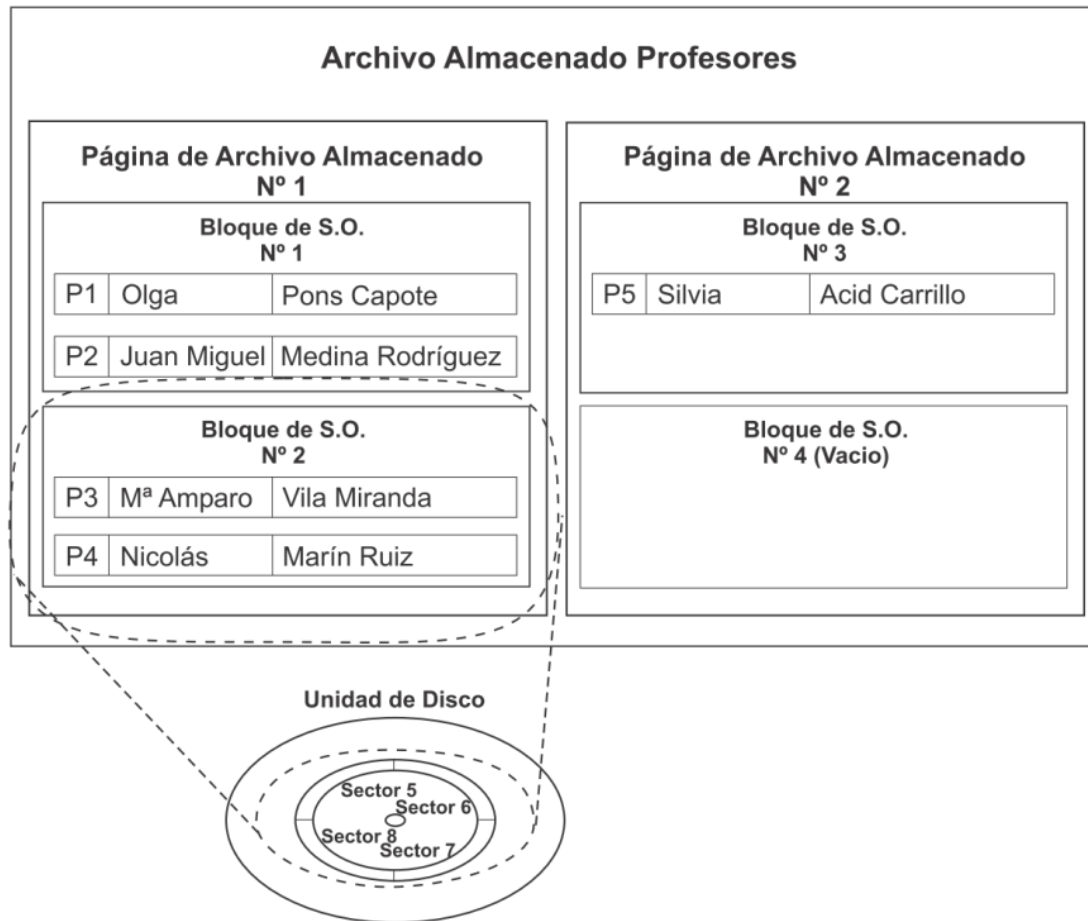
$$t_a = t_b + t_l$$

- Tiempo medio entre fallos (MTBF)

Contenidos

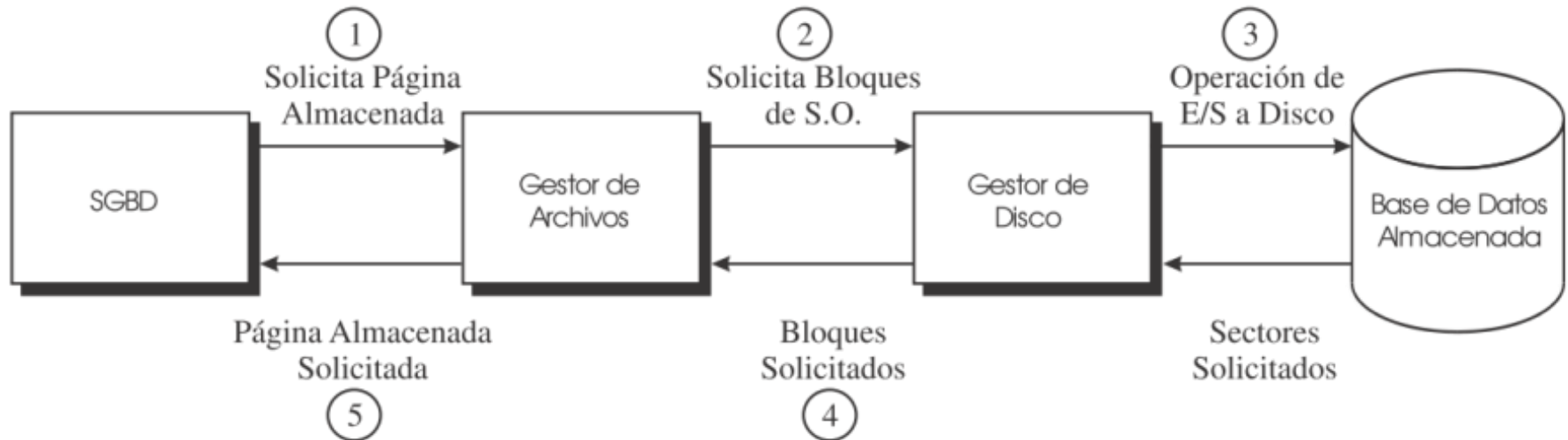
- Introducción
- Método de Acceso a la BD almacenada
- Representación de la BD en el nivel interno

Método de acceso a la BD



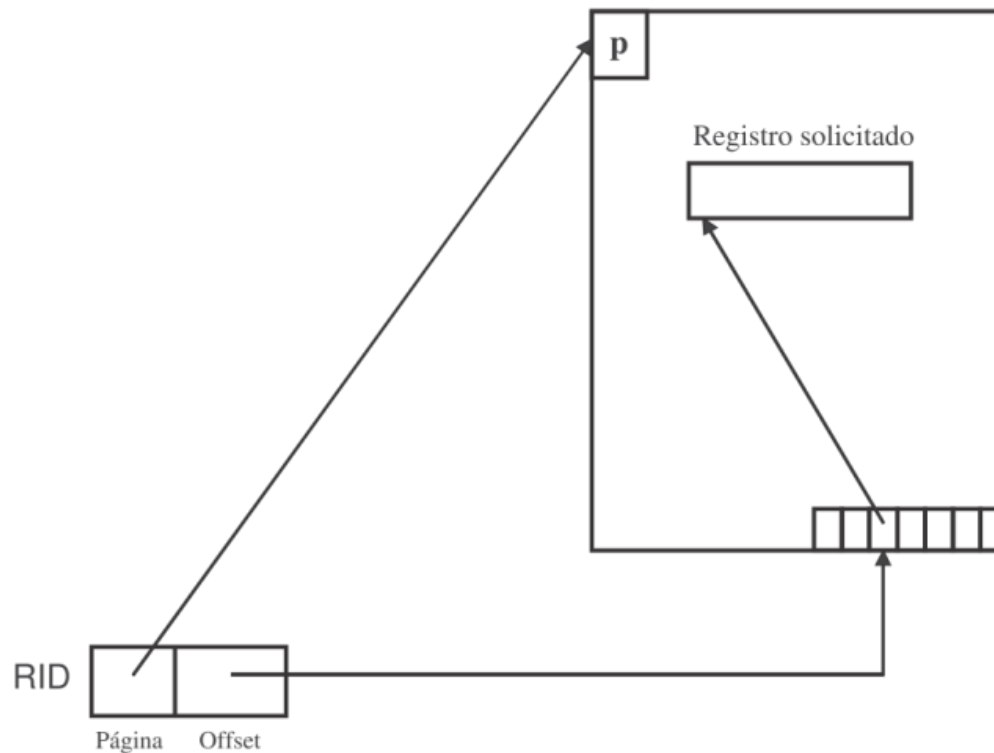
Método de acceso a la BD

¿Cómo se transforma un registro almacenado en una representación física en el almacenamiento secundario?



Método de acceso a la BD

- Para que el gestor de almacenamiento pueda localizar un registro almacenado, se utiliza el RID (Record Identifier):



Método de acceso a la BD

- Cada registro almacenado:
 - Cabecera: Número y tipo de columnas que lo integran.
 - Datos: Contenido de las columnas.
- Las páginas o bloques de la BD deben tener un tamaño múltiplo de las páginas del sistema operativo (mínima unidad de E/S).
- Para recuperar un registro almacenado hay que determinar la página de BD que lo contiene y entonces recuperar los bloques de disco que la integran.
- Hay que organizar la estructura de almacenamiento y los métodos de acceso, de forma que se optimice la interacción con los dispositivos de almacenamiento secundario.
- **Deben minimizarse las operaciones de E/S al almacenamiento secundario.**

Método de acceso a la BD

- El Gestor de Disco del SO
 - Normalmente el SGBD interactúa con la BD almacenada en el sistema de almacenamiento secundario a través del gestor de disco del SO.
 - El gestor de disco organiza los datos en conjuntos de bloques o archivos de SO. Una BD puede utilizar uno o varios de estos archivos para almacenar su contenido.
 - También se encarga de gestionar el espacio libre en el disco.

Método de acceso a la BD

- Funciones elementales:

- Crear un nuevo archivo de sistema operativo.
- Eliminar un archivo de sistema operativo existente.
- Añadir un bloque nuevo al conjunto de bloques c.
- Eliminar el bloque b del conjunto de bloques c.
- Devolver el bloque b del conjunto de bloques c.
- Reemplazar el bloque b dentro del conjunto de bloques c.

Método de acceso a la BD

- El Gestor de Archivos del SGBD
 - Componente del SGBD que se encarga de:
 - Hacer la transformación entre:
 - Campos, registros y archivos almacenados
 - y
 - bloques y conjuntos de bloques (que pueda entender el gestor de disco).
 - Organizar los datos de manera que se minimice el tiempo de recuperación:
 - Minimizar las E/S a disco.

Método de acceso a la BD

- Funciones básicas:
 - Crear un nuevo archivo almacenado.
 - Asociar al archivo un conjunto de páginas o bloques de la BD.
 - Eliminar un archivo almacenado.
 - Recuperar el registro almacenado r del archivo almacenado a.
 - Normalmente, el SGBD proporciona el RID.
 - Sólo hay que obtener en memoria la página que contiene el registro para extraerlo.
 - Añadir un nuevo registro almacenado al archivo almacenado a. Hay que localizar la página de BD más apropiada de las pertenecientes al archivo almacenado.
 - Si no se pudiera, se solicita una nueva página.
 - Se devuelve al SGBD el RID nuevo.

Método de acceso a la BD

- Funciones básicas (continuación):
 - Eliminar el registro r del archivo almacenado a .
 - Hay que recuperar la página de BD que contiene dicho registro y marcar el espacio ocupado por el registro en dicha página como disponible.
 - Actualizar el registro r en el archivo almacenado a .
 - Recupera la página de la BD que contiene el registro que se desea actualizar.
 - Trata de sustituir la información. Si no puede, se intenta ubicar en otra página.

Contenidos

- Introducción
- Método de Acceso a la BD almacenada
- Representación de la BD en el nivel interno

Representación de la BD en el nivel interno

- La BD se representa de diferentes formas en los diferentes niveles de la arquitectura del SGBD.
 - Su representación en el nivel interno no tiene por qué coincidir con su representación en el nivel conceptual.
 - Cada conjunto de registros del mismo tipo no tiene por qué ser un fichero.
- El nivel interno debe traducir las estructuras del nivel conceptual a otras estructuras intermedias más cercanas al almacenamiento real de los datos (nivel físico).

Representación de la BD en el nivel interno

Agrupamiento

La BD en el nivel interno

Conjunto de páginas en las que se van ubicando los registros.

Agrupamiento Intra-Archivo

Ubicar en una página registros del mismo tipo.

Es el más frecuente.

Agrupamiento Inter-Archivo

Ubicar en una página registros de distinto tipo.

Ha de existir relación (por ejemplo entidades fuerte-débil)

Representación de la BD en el nivel interno

Ejemplo

Se inserta una nueva asignatura con código **A6**.

- Se localiza la primera página libre (la **24**).
- Se inserta el registro correspondiente.
- Se añade esta página al conjunto de páginas de asignaturas.

0		1		2		3		4		5	24
			A₁		A₂		A₃		A₄		A₅
6		7		8		9		10		11	X
	P₁		P₂		P₃		P₄		P₅		P₆
12		13		14		15		16		17	
	A₁/P₁		A₁/P₂		A₁/P₃		A₁/P₄		A₁/P₅		A₁/P₆
18		19		20		21		22		23	X
	A₃/P₁		A₃/P₂		A₃/P₆		A₄/P₂		A₄/P₄		A₄/P₅
24	X	25		26		27		28		29	X
	A₆										

Representación de la BD en el nivel interno

Ejemplo

Se borra la asignatura con código **A2**.

- La página que contiene a esta asignatura pasa al conjunto de páginas libres.
- Se reorganiza la lista correspondiente a Asignaturas.

0	1	3	2	25	3	4	5	24
	A₁				A₃	A₄	A₅	
6	7	8	9	10	11	X		
P₁	P₂	P₃	P₄	P₅	P₆			
12	13	14	15	16	17			
A₁/P₁	A₁/P₂	A₁/P₃	A₁/P₄	A₁/P₅	A₁/P₆			
18	19	20	21	22	23	X		
A₃/P₁	A₃/P₂	A₃/P₆	A₄/P₂	A₄/P₄	A₄/P₅			
24	X	25	26	27	28	29	X	
A₆								

Representación de la BD en el nivel interno

Ejemplo

Se introduce un nuevo profesor con código **P7**.

- Se ubica en la primera página libre disponible (la segunda).
- Se ajustan las cadenas de punteros.

0	1	3	2	28	3	4	5	24
	A₁	P7		A₃		A₄		A₅
6		7	8	9	10	11	X	
P₁	P₂	P₃	P₄	P₅	P₆			
12	13	14	15	16	17			
A₁/P₁	A₁/P₂	A₁/P₃	A₁/P₄	A₁/P₅	A₁/P₆			
18	19	20	21	22	23	X		
A₃/P₁	A₃/P₂	A₃/P₆	A₄/P₂	A₄/P₄	A₄/P₅			
24	X	25	26	27	28	29	X	
A₆								

Representación de la BD en el nivel interno

Ejemplo

- Se borra **A4**.
 - Su página pasa al conjunto de páginas libres.
 - Se reorganiza la cadena de punteros de las Asignaturas.

0	1	3	2	X	3	5	4	25	5	24
	A₁	P₇		A₃					A₅	
6	7	8	9	10	11	2				
P₁	P₂	P₃	P₄	P₅	P₆					
12	13	14	15	16	17					
A₁/P₁	A₁/P₂	A₁/P₃	A₁/P₄	A₁/P₅	A₁/P₆					
18	19	20	21	22	23	X				
A₃/P₁	A₃/P₂	A₃/P₆	A₄/P₂	A₄/P₄	A₄/P₅					
24	X	25	26	27	28	29	X			
A₆										

Representación de la BD en el nivel interno

Ejemplo Punteros para el recorrido secuencial lógico

0	X	1	3	2	X	3	5	4	25	5	24
		A₁		P₇		A₃				A₅	
6	7	7	8	8	9	9	10	10	11	11	2
P₁		P₂		P₃		P₄		P₅		P₆	
12	13	13	14	14	15	15	16	16	17	17	18
A₁/P₁		A₁/P₂		A₁/P₃		A₁/P₄		A₁/P₅		A₁/P₆	
18	19	19	20	20	21	21	22	22	23	23	X
A₃/P₁		A₃/P₂		A₃/P₆		A₄/P₂		A₄/P₄		A₄/P₅	
24	X	25	26	26	27	27	28	28	29	29	30
A₆											

Representación de la BD en el nivel interno

Índice en la página 0

0		X
----------	--	----------

Conjunto	Dirección a Primera Pág
Pág. Libres	4
Asignaturas	1
Profesores	6
Imparte	12

Representación de la BD en el nivel interno

Factor de bloqueo N

p	Información de cabecera	
	A1 BASES DE DATOS	A2 ALGEBRA
	A3 COMPUTABILIDAD	A4 METODOLOGÍA
	A5 PROGRAMACION DE BD	Espacio libre

En realidad, las páginas contienen más de un registro

Representación de la BD en el nivel interno

Ejemplo

p	Información de cabecera		
A1	BASES DE DATOS	A3	COMPUTABILIDAD
A4	METODOLOGÍA	A5	PROGRAMACION DE BD
A7	MATEMÁTICA DISCRETA	A9	SISTEMAS DE BD
Espacio libre			

Inserción de A9, borrado de A2 e inserción de A7

Representación de la BD en el NI

Consideraciones finales

La organización descrita es un ejemplo general. Cada SGBD comercial utiliza su variante concreta, aunque la idea subyacente es la misma.

No existe una relación directa fichero-almacenado/fichero-físico, ya que todos los conjuntos de páginas irán almacenados, con toda probabilidad, en uno o varios ficheros físicos.

Contenidos

- Introducción
- Método de Acceso a la BD almacenada
- Representación de la BD en el nivel interno