

1. – SISTEMAS DE ALMACENAMIENTO DE LA INFORMACIÓN.

- 1.1 – Ficheros
- 1.2 - Concepto de B.D.
- 1.3 - Sistemas gestores de B.D. (S.G.B.D.)
- 1.4 - Funciones de un S.G.B.D.
- 1.5 - Diccionario de datos.
- 1.6 - Arquitectura cliente / servidor.
- 1.7 - Tipos de usuario.
- 1.8 - Modelos de BD.

1.1 . FICHEROS.

Los ficheros son estructuras de información que crean los sistemas operativos de los ordenadores para poder almacenar los datos. Tienen un nombre y una extensión que determina el formato de la información que contiene.

El formato y el tipo de fichero determina la forma de interpretar la información que contiene, puesto que en definitiva, lo único que almacena es una ristra de bits de forma que es necesaria su interpretación para dar sentido a la información que almacena.

Tradicionalmente los ficheros se han clasificado de la siguiente forma:

- a) Según su **contenido** de un fichero puede ser tratado como texto o como datos binarios, es decir, los bits almacenados en el fichero pueden ser traducidos por el sistema operativo a caracteres alfabéticos y números que entiende el ser humano, o pueden ser tratados como componentes de estructura de datos más complejos.
- b) Según su **organización** de un fichero se dicta la forma de acceder a los datos, así pues los datos de un fichero con organización secuencial los registros están dispuestos siguiendo una secuencia ordenada, los de organización directa permiten el acceso a los datos directamente sin necesidad de pasar por los anteriores y los de organización indexada acceden a la información a través de un índice.
- c) Según su **utilidad**, indica el uso que se le va a dar al mismo, pudiendo ser maestros y de movimientos.

Ficheros de texto

Se les conoce también con el nombre de planos o **ascii** (American Standard Code for Information Interchange). La asignación de valores a caracteres viene dado por 2^X .

Los ficheros de texto se agrupan de la siguiente forma:

- Ficheros de **configuración**: se utilizan para configurar los SO. .ini, .inf, .conf.
- Ficheros de **código fuente**: su contenido es texto con programas informáticos. .sql, .c, .java.
- Ficheros de **páginas web**: que no dejan de ser ficheros de texto con hipertexto que interpreta el navegador. .html, .php, .css, .xml.
- Formatos **enriquecidos**: son textos que contienen códigos de control par ofrecer una visión del texto más elegante: .rtf, .ps, .tex.

Ficheros binarios

Son todos aquellos que no son de texto, y requieren un formato para ser interpretados, como son:

- De imagen: .jpg, .gif, .tiff, .bmp, .wmf, .png, .pcx, etc.
- De video: .mpg, .mov, .avi, .qt.
- Comprimidos o empaquetados: .zip, .Z, .gz, .tar, .lhz.

- Ejecutables o compilados: .exe, .com, .cgi, .o, .a.
- Procesadores de texto: .doc, .odt.

1.2 . CONCEPTO DE B.D.

Una BD es un conjunto de informaciones no redundantes las cuales pueden ser compartidas por varias aplicaciones.

Se considera redundante una BD cuando se repite un elemento que no es clave principal de una tabla muchas veces.

Cuando creábamos tablas en lenguajes de 3ª generación estas no podrían ser utilizadas por ningún otro proceso que no perteneciese al programa, o sea, que si en otro programa se necesitaba una tabla debía crear una nueva. Esto se ha eliminado en los lenguajes de 4ª generación, las tablas ya pueden compartirse según derechos preestablecidos por el administrador, y así se reducen enormemente las redundancias.

En los sistemas informáticos tradicionales, los datos se almacenan en ficheros diseñados específicamente para cada aplicación, y no suele compartirse dicho fichero entre distintas aplicaciones. En una BD estos se almacenan con el objetivo de que sirvan a distintas aplicaciones y que pueda ser compartidos por estas. Podemos decir que los sistemas tradicionales están orientados al proceso y los datos están subordinados a los mismos. En las BD son sistemas orientados a datos y estos se almacenan independientemente de las aplicaciones.

Ventajas de las B.D.

- Independencia de los datos y los procesos que la utilizan
- Coherencia de los resultados obtenidos
- Mayor y menor disponibilidad de los datos
- Mayor valor informativo
- Mejor y más normalizada la documentación de la información

1.3 . S.G.B.D.

Un SGBD es un conjunto de programas que permiten a los usuarios describir, recuperar y manipular los datos almacenados en la BD, protegiendo dichos datos contra todas aquellas acciones intencionadas o no que los puedan destruir.

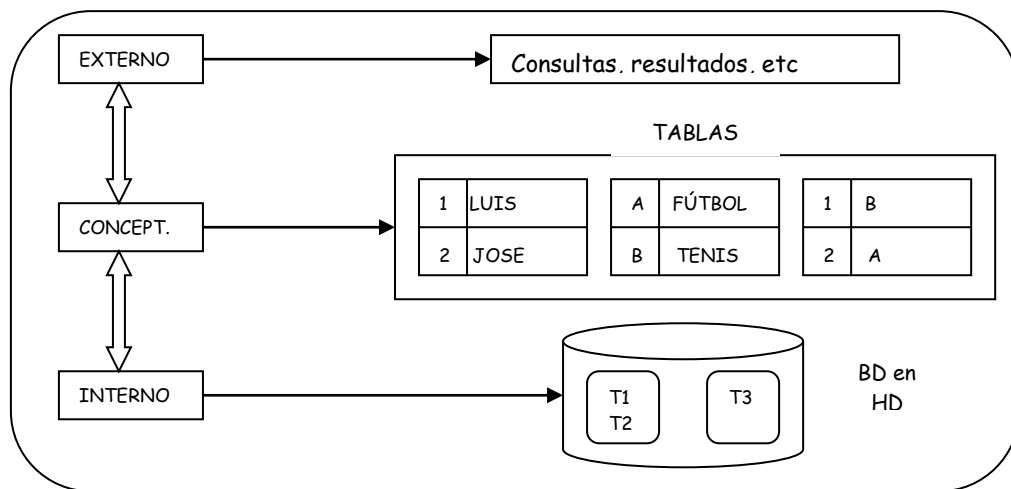
En principio se puede observar una estructura lógica que corresponde a la visión que de los datos tiene el usuario y una estructura física o forma en que están almacenados dichos datos.

El grupo ANSI/X3 definió 3 niveles de BD:

- EXTERNO (lo que el usuario ve y maneja)
- CONCEPTUAL (el modelo E – R)
- INTERNO (información organizada por el SGBD)

1.4 . FUNCIONES QUE REALIZA UN S.G.B.D.

- *Descripción de datos:* implica a decir que datos componen la base, su estructura, interrelaciones y restricciones.
- *Manipulación de Datos:* se pueden actualizar, crear, modificar, borrar e insertar. Las personas capacitadas para realizar esta labor son los usuarios, mientras que el administrador da permisos para que puedan realizar dichas acciones.
- *Transformación de Datos:* existen 2 niveles de transformación:
 - Transformación conceptual interna: permite pasar los datos desde su representación conceptual a su representación interna, y viceversa.
 - Transformación conceptual externa: permite pasar los datos desde su representación en nivel conceptual al esquema externo y viceversa.



- *Protección de Datos:*
 - Seguridad: Se trata de una protección contra los fallos lógicos o físicos que destruyan los datos total o parcialmente.
 - Integridad: el objetivo de la integridad es lograr que los datos sean correctos. Para ello el SGBD debe detectar y corregir las operaciones incorrectas que introduzcan inconsistencias en la BD. La integridad dice si los datos introducidos son correctos. Existen 2 operaciones que pueden afectar a la integridad:
 1. Las que afectan a las propias restricciones: se deben introducir datos válidos, y el sistema no debe aceptar los incorrectos.
 2. Las operaciones concurrentes: evitar este tipo de operaciones.
 - Confidencialidad: Los datos que se crean deben estar protegidos del acceso de otros usuarios, excepto si el propietario lo permite.
- *Utilidades de administración:* Van a englobar tareas de seguridad y de mantenimiento, y que solamente el administrador podrá realizar estas funciones. También podrá realizar tareas como cambiar ficheros, obtención de copias de seguridad, recuperación tras la caída del sistema, etc.

1.5 . DICCIONARIO DE DATOS

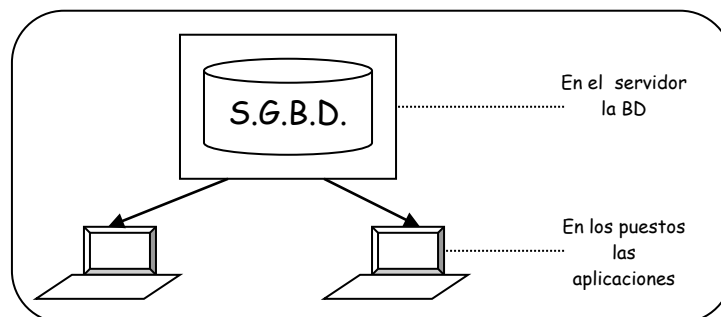
Contiene información sobre los datos almacenados en la base desde el punto de vista del usuario. En este estará almacenado lo que los usuarios necesitan para comprender el significado de los datos y poder manejarlos adecuadamente como descripciones, seguridad, estructuras, etc.

REPOSITORIO: Es un diccionario que almacena información manejada por una herramienta CASE. Se almacena información textual y gráfica relativa a las distintas fases del C.V.S.

1.6 . ARQUITECTURA CLIENTE-SERVIDOR

Un sistema de BD posee una estructura compuesta de 2 partes: un servidor o SGBD y un conjunto de clientes con las distintas aplicaciones ejecutadas en torno al gestor. Las aplicaciones de usuario suelen ser proporcionadas por el proveedor del SGBD. El software adecuado para una arquitectura cliente-servidor posee varios componentes, que se pueden asociar al cliente o al servidor. Son los siguientes:

- *Software de gestión de datos:* lleva a cabo la manipulación y gestión de los datos (SGBD)
- *Software de interacción con el usuario:* implementa funciones que se asocian a una interfaz gráfica de usuario (SQL). Reside en el cliente.
- *Software de desarrollo:* se usa para desarrollar aplicaciones. Reside en el cliente (Developer, Designer).



La arquitectura cliente-servidor permite las siguientes configuraciones:

1. **ANFITRIÓN:** S.G.B.D. y cliente son la misma máquina
2. **CLIENTE-SERVIDOR:** El S.G.B.D. reside en el servidor y los clientes se conectan a la misma a través de una red de área local.
3. **PROCESO DISTRIBUIDO:** El S.G.B.D. está repartido en más de una máquina servidora y los usuarios, no conociendo la ubicación física de los datos, acceden al servidor.

Tipos de usuarios

- **Usuarios finales:** Son aquellos no muy expertos que manejan normalmente aplicaciones y realizan labores de:

1. Manejar aplicaciones diseñadas exclusivamente para ellos.

2. También utilizan un lenguaje que les permite realizar operaciones de manipulación, actualización y recuperación.
 3. Realizar consultas y modificaciones a la BD sin necesidad de conocer ningún lenguaje.
- Desarrolladores de aplicaciones: Crean aplicaciones que han sido solicitadas por los clientes siendo sus principales funciones:
1. Escribir en un lenguaje de 3ª generación programas que intercalen sentencias de un lenguaje de manipulación de datos.
 2. Utilizar generadores de formatos de pantalla, generadores de informes y lenguajes de 4ª generación para desarrollar aplicaciones.
- Administrador de la BD: Es el responsable de diseño, creación y mantenimiento de la base.

1.7 . MODELOS DE B.D.

➤ Jerárquica

La representación de los datos se hace de forma arborescente y está compuesto por nodos (entidades), que son segmentos. Cada segmento guarda información homogénea. El grado jerárquico es uno a muchos.

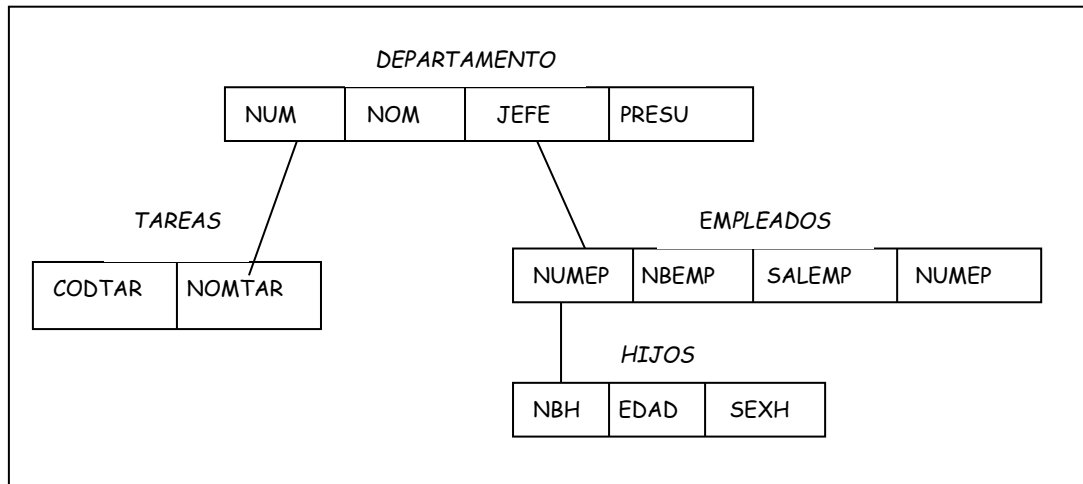
Características:

1. Que un segmento tiene hijos (1-M).
2. Un segmento atómico es aquel que no tiene hijos
3. Al conjunto de nodos del mismo nivel se les llama nodos hermanos.
4. Un segmento hijo sólo puede tener un padre.
5. Se recorre el árbol en forma de pre-orden.

Inconvenientes:

1. No admite más de una relación entre 2 segmentos.
2. No puede tener relaciones M-N (sino que transformada en 1-M, según Martín).
3. No se permite que un hijo tenga más de un padre.
4. Se suele recorrer en un orden.
5. Para acceder a cualquier segmento hay que pasar por el nodo raíz.

Un ejemplo de BD jerárquicas pueden ser: IMS de IBM ó SYSTEM 2000 de INTEL.



➤ En Red

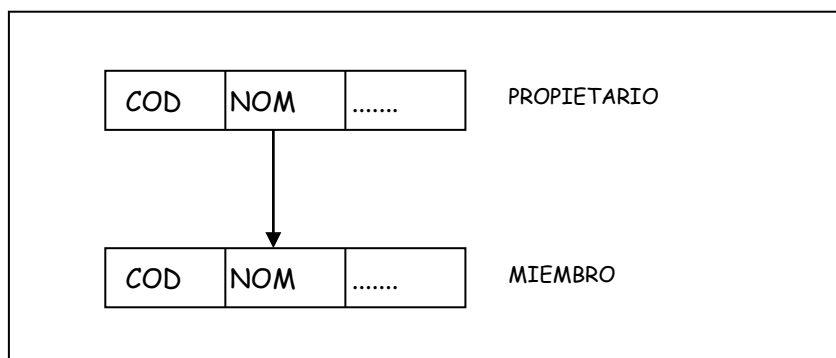
Las entidades se representan por nodos y las líneas comunican los nodos. El nodo se llama tipo de registro. Al campo se le llama elemento. Agregado de datos es un conjunto de datos independientes (fecha día / mes / año). Conjunto es la relación entre 2 tipos de registros.

Representa relaciones de tipo 1-M ó 1-1, aunque es factible representar las de M-N, siempre y cuando se cree un registro de enlace que conecte con los tipos registro.

Características:

1. Va a necesitar punteros que marquen la dirección que se va leer.
2. Por cada ocurrencia se obtienen grafos.
3. Al tipo de registro del que se parte la flecha se llama propietario.
4. Al tipo de registro al que llega la flecha se llama miembro.

Un ejemplo de BD en red es: DBOMP de IBM ó PHOLAS de PHILIPS.



➤ Relacional

Un S.G.B.D. de enfoque relacional utiliza tablas bidimensionales para la representación lógica de los datos y las relaciones entre ellos. Está basado la teoría matemática de las

relaciones. Fue Codd quien desarrolló en IBM-San José el modelo de datos relacional. Se trata de un modelo lógico de datos que hace referencia a representación lógica de la información; por tanto, no es directamente aplicable a la representación física.

Algunas de las ventajas del modelo relacional son las siguientes:

- Ofrece una visión conceptual sencilla del conjunto de datos existentes en la base.
- Puede entenderlo y usarlo cualquier usuario, no es preciso que sea un experto informático.
- Los usuarios no necesitan saber dónde se encuentra físicamente los datos.
- Se puede ampliar el esquema conceptual sin modificar los programas de aplicación.

El elemento principal del modelo relacional es la relación. Cada relación se representa mediante una tabla.

Actualmente, es el modelo más utilizado. Ejemplos de base de datos relacionales son: SYSTEM R y QBE de IBM, ORACLE de RSI, INFORMIX y DBASE IV.

Este tipo de BD se verá en el tema siguiente de forma detallada.