

Unidad de trabajo 1

Almacenamiento de la Información

→ Sistema Gestor de Base de Datos

- Definición de SGBD
- Arquitectura de un SGBD
- Funciones de un SGBD
- Independencia de los datos
- Componentes de un SGBD
- Ventajas de los SGBD
- Tipos de SGBD
- SGBD comerciales
- SGBD libres

Definición de SGBD

Sistema Gestor de Base de Datos (SGBD): Conjunto coordinado de programas, procedimientos, lenguajes y técnicas, que ofrece a todos los usuarios (administrador, diseñador, programador y usuario final) los medios necesarios para describir y manipular los datos contenidos en la base de datos, manteniendo su integridad, confidencialidad y seguridad.

También llamado **GBD** (Gestor de Base de Datos), **DBMS** (DataBase Management System) o **RDBMS** (Relational DBMS) si se trata de una BD relacional.

El SGBD interacciona con el sistema operativo para utilizar los recursos hardware del sistema, ocultando su complejidad a las aplicaciones que utilizan la BD.

| |
|-------------------|
| Aplicaciones |
| SGBD |
| Sistema Operativo |
| Hardware |

Existen diversos **tipos de usuarios** que utilizan los SGBD:

Administrador: encargado de la implementación física de la BD. Define los tipos de ficheros, su ubicación, los índices necesarios, etc. Toma las decisiones relacionadas con el funcionamiento físico del almacenamiento de información y establece las políticas de seguridad.

Diseñador: encargado de diseñar cómo será la base de datos. Realizará un modelo partiendo del mundo real, identificando qué datos deben almacenarse y su formato, las relaciones entre ellos, sus restricciones, etc. Para ello han de conocer a fondo los datos y procesos a representar en la base de datos (reglas de negocio de la empresa).

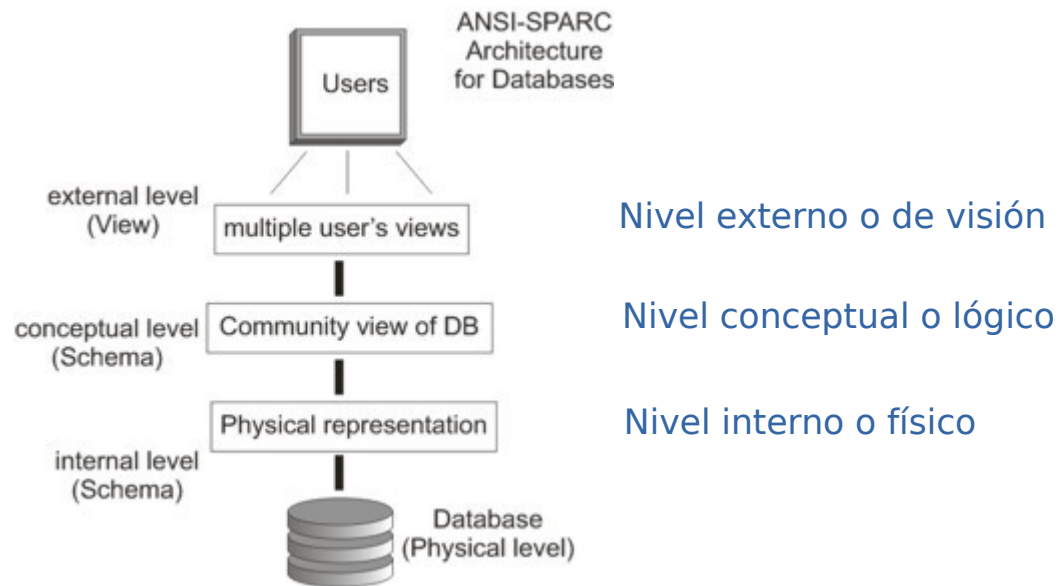
Programador: implementa los programas de aplicación que servirán a los usuarios finales. Ofrecerán la posibilidad de realizar consultas de datos, inserción, actualización o eliminación, sin necesidad de tener conocimiento técnico de bases de datos.

Usuario final: utiliza la información de la base de datos a través de una interfaz que no requiera tener conocimiento técnico ni informático, por ejemplo empleando un navegador web.

Arquitectura de un SGBD

La arquitectura **ANSI-SPARC** (American National Standards Institute, Standards Planning And Requirements Committee) es un estándar de diseño abstracto para un SGBD, propuesto por primera vez en 1975.

Se definen tres niveles:



La finalidad es:

- Ocultar la complejidad de la base de datos a los usuarios
- Independizar los niveles de forma que se pueda operar en cada uno de ellos sin afectar al resto

Nivel externo o de visión: describe la visión individual de un usuario o grupo de usuarios (qué datos puede manipular y en qué formato los puede ver).

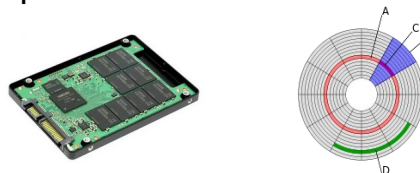
Pueden haber varios esquemas externos en una BD.



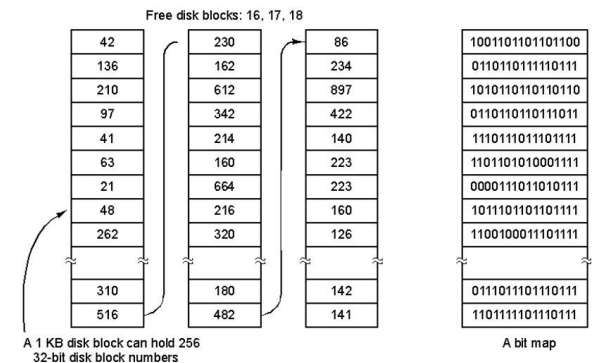
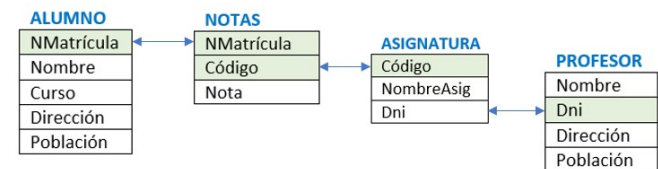
Nivel conceptual o lógico: describe la estructura de la BD (cómo se organizan y relacionan los datos según el modelo de datos). Solo hay un esquema conceptual en una BD.



Nivel interno o físico: describe los detalles de cómo se almacenan físicamente los datos (los archivos que contienen la información, su organización, los métodos de acceso a los registros, los tipos de registros, la longitud, los campos que los componen, etc). Solo hay un esquema interno en una BD.



| Curso | Nombre | Nombre de asignatura | Nota |
|-------|--------|---|------|
| 1 | Ana | Programación en lenguajes estructurados | 6 |
| 1 | Ana | Sistemas informáticos multiusuario y en red | 8 |
| 2 | Rosa | Desa. de aplic. en entornos de 4.ª Generación y H. Case | 5 |
| 2 | Juan | Desa. de aplic. en entornos de 4.ª Generación y H. Case | 7 |
| 1 | Alicia | Programación en lenguajes estructurados | 5 |
| 1 | Alicia | Sistemas informáticos multiusuario y en red | 4 |



Funciones de un SGBD

Un SGBD Un SGBD desarrolla tres funciones fundamentales:

→ **Definición de Datos:** Permite al diseñador de la base de datos crear las estructuras apropiadas para integrar adecuadamente los datos. Se realiza mediante el lenguaje de definición de datos o **LDD**.

Esta función es la que permite definir las tres estructuras de la base de datos (los tres niveles en la arquitectura ANSI-SPARC):

- A nivel interno (esquema físico) se indica el espacio de disco reservado para los datos.
- A nivel conceptual (esquema lógico) se crea o modifica la estructura de la base de datos. Según el mdello de BD, se definen las estructuras que contendrán los datos, sus relaciones y restricciones (por ejemplo las tablas, claves e índices en el modelo relacional).
- A nivel externo (esquemas de visión) se definen las vistas de los usuarios.

→ **Manipulación de Datos:** permite leer y actualizar los datos. Utilizado por los usuarios para realizar consultas, inserciones, eliminaciones y modificaciones de los datos. Se realiza mediante el lenguaje de manipulación de datos o **LMD**.

→ **Control de Datos:** permite gestionar los accesos a la base de datos. Empleado por los administradores para establecer las políticas de acceso. Adicionalmente, incorpora sistemas para la creación de copias de seguridad, carga de ficheros, auditoría, protección de ataques, configuración del sistema, etc. Se realiza mediante el lenguaje de control de datos o **LCD**.

Independencia de los datos

En un sistema informático basado en ficheros existe una dependencia de los datos física-lógica, o lo que es lo mismo, la estructura física de los datos (definición de archivos y registros) se encuentra codificada en los programas de aplicación. Cualquier cambio en esa estructura implica al programador identificar, modificar y probar todos los programas que manipulan esos archivos.

En cambio, en un SGBD existe un diseño realizado por capas que constituyen distintos niveles de abstracción. Se consigue una independencia entre la organización lógica y física de la información.

→ **Independencia física:** capacidad de modificar la forma en la que se almacena la información en los ficheros y soportes físicos (nivel físico) sin alterar la estructura lógica de los datos (nivel conceptual) ni su visualización (nivel externo).

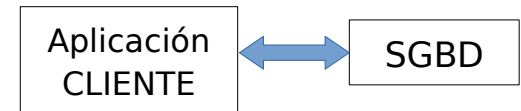
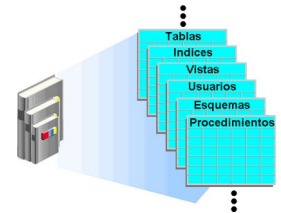
Por ejemplo, se pueden reorganizar los archivos físicos con el fin de mejorar el rendimiento de las operaciones de consulta y actualización, o se pueden añadir nuevos archivos de datos porque los que había han crecido demasiado.

→ **Independencia lógica:** capacidad de modificar la estructura lógica de los datos (nivel conceptual) sin afectar a las aplicaciones que usan la BD (nivel externo).

Componentes de un SGBD

Un SGBD es un paquete de software complejo compuesto por una serie de componentes:

- ➔ **Motor de base de datos (engine):** se utiliza para crear, leer, actualizar y eliminar datos de una base de datos. Muchos SGBD admiten múltiples motores de almacenamiento. Ejemplos: InnoDB, XtraDB, Aria...
- ➔ **Diccionario de datos:** descripción de los datos que pueden almacenarse en la base de datos: tipos de estructuras, significado, relaciones, uso y formato. No son los datos en sí, sino la descripción de qué puede contener la base de datos (son metadatos).
- ➔ **Lenguajes de base de datos:** representan la interfaz que se proporciona al exterior para interactuar con los datos.
 - **Lenguaje de Definición de Datos (LDD o DDL):** crea o modifica la estructura de la BD.
 - **Lenguaje de Manipulación de Datos (LMD o DML):** consulta y actualiza los datos.
 - **Lenguaje de Control de Datos (LCD o DCL):** gestiona los accesos a la base de datos.
 - **Lenguaje de Control de Transacciones (LCT o TCL):** confirma las operaciones LMD.
- ➔ **Herramientas:** conjunto de aplicaciones que ofrecen funcionalidades adicionales: administración de la base de datos, generadores de formularios, generadores de informes, interfaces gráficas, espacios de trabajo, desarrollo de aplicaciones, etc.
- ➔ **Usuarios:** existen diferentes perfiles de usuario, cada uno de ellos con una serie de permisos sobre los objetos de la base de datos. Conceptos relacionados: privilegios, perfiles y roles.



Ventajas de los SGBD

- Proporcionan al usuario una visión abstracta de los datos, ocultando parte de la complejidad relacionada con cómo se almacenan y mantienen.
- Ofrecen Independencia física y lógica. Flexibilidad (modificaciones de las estructuras). Escalabilidad (dimensionamiento según las necesidades)
- Disminuyen la redundancia y la inconsistencia de datos.
- Coherencia de los datos (transacciones)
- Eficiencia en el almacenamiento (ahorro de espacio de almacenamiento)
- Facilitan el acceso a los datos, aportando rapidez y eficiencia. Permiten búsquedas complejas de información.
- Aumentan la seguridad y privacidad de los datos. Integridad (autenticación, permisos, vistas, cifrado)
- Permiten compartir datos y accesos concurrentes. Facilitan el intercambio de datos entre distintos sistemas.
- Incorporan mecanismos de copias de seguridad y recuperación para restablecer la información en caso de fallos en el sistema. Disponibilidad.
- Registro de operaciones. Logs

Tipos de SGBD

Una clasificación según varios criterios:

- **Según el modelo lógico de base de datos:** modelo jerárquico, en red, relacional, noSQL, orientado a objetos... Los dos primeros se han quedado obsoletos.
- **Según el número de usuarios:**
 - Monousuario: bases de datos de escritorio.
 - Multiusuario: la mayoría de SGBD comerciales.
- **Según la ubicación de los datos:**
 - Centralizado: en una máquina.
 - Distribuido: en varias máquinas. Clusters.
- **Según el propósito o finalidad de los datos:**
 - General: puede emplearse para cualquier ámbito (médico, comercial, industrial...).
 - Específico: diseñado para un ámbito concreto. Se dará prioridad a ciertos aspectos clave según el sector del que se trate. Por ejemplo, si se van a soportar consultas masivas remotas y concurrentes, puede focalizarse el diseño en minimizar el retardo de la respuesta.
- **Según la licencia:** sistemas comerciales y libres. Existen diversos modelos de licencia que incluyen distintos tipos de pagos, tarifas y libertades en el uso del software.

SGBD comerciales

En el mercado de software existen multitud de sistemas gestores de bases de datos comerciales. La decisión de elegir uno de ellos para un proyecto depende de múltiples factores: volumen de carga, sistema operativo, tipo de acceso de los usuarios, aplicaciones adicionales, presupuesto, soporte...

A continuación se enumera una muestra de SGBD comerciales. La mayoría son relacionales:



Primera base de datos comercial.
Usa ficheros dbf, que tienen registros de longitud fija.



Base de datos de escritorio incluida en el paquete Microsoft Office.



Utilización en sistema operativo Windows.



Muy extendido a empresarial.



Express Edition
(versión gratuita)



Producto de IBM



Integra documentos XML de forma nativa



Desarrollada por Borland Software Corporation. Requiere poca administración



Varios productos de BD, propiedad de SAP.



Servicio de base de datos noSQL ofrecido por Amazon como parte de Amazon Web Services (AWS). Prioriza el rendimiento.

GPL General Public License (Licencia Pública General):

licencia regula los derechos de autor de los programas de software libre (free software). Su objetivo es doble: recalcar que el software cubierto por esta licencia es libre y, por otro lado, protegerlo (mediante una práctica conocida como copyleft) de intentos de apropiación que restrinjan esas libertades a nuevos usuarios cada vez que la obra es distribuida, modificada o mejorada.

Un programa es sw libre si los usuarios tienen las cuatro libertades:

- ▶ **LIBERTAD 0.** La libertad de ejecutar el programa como se desee, con cualquier propósito.
- ▶ **LIBERTAD 1.** La libertad de estudiar cómo funciona el programa, y cambiarlo a conveniencia. El acceso al código fuente es una condición necesaria para ello.
- ▶ **LIBERTAD 2.** La libertad de redistribuir copias para ayudar a otros.
- ▶ **LIBERTAD 3.** La libertad de distribuir copias de sus versiones modificadas a terceros. Esto le permite ofrecer a toda la comunidad la oportunidad de beneficiarse de las modificaciones. El acceso al código fuente es una condición necesaria para ello.





Muy utilizado ya que se puede integrar con mucha facilidad en páginas web.

Desarrollado inicialmente por Sun Microsystems, fue adquirido por Oracle cuando ésta adquirió a Sun. En estos momentos MySQL se mantiene con **licencia doble**, por una parte es GPL y código abierto, gratuita para uso por la comunidad, por otra parte tiene licencia comercial. Esta dualidad de licencia ha sido la que provocó la creación de MariaDB, un sistema gestor de bases de datos creado a partir de MySQL, donde no se depende de la tutorización de Oracle



BASE. Base de datos de escritorio incluida en el paquete LibreOffice.



PostgreSQL

Sistema relacional orientado a objetos



NoSQL distribuida basada en documentos sin estructura y asociados por colecciones



BD en memoria volátil aunque puede usarse como persistente. Tabla hash clave-valor. NoSQL.



NoSQL distribuida. Usada por Facebook y Netflix. Permite P2P. Apache.



Apache Derby

Relacional. Tamaño reducido, muy portable.



NoSQL distribuida orientada a páginas web (JavaScript, JSON). Apache



NoSQL basado en BigTable. Usada por Tuenti, Facebook y Twitter. Distribuida.



Distribuida NoSQL con clave-valor. Diseño para LinkedIn



Firebird®