

Introducción a las bases de datos

Índice

1. **Introducción**
2. **Arquitectura de los SGBD**
3. **Modelos, lenguajes y administración de BD**

1. Introducció

BD y SGBD

Llamamos **bases de datos** (BD) a los **conjuntos de ficheros interrelacionados**, con estructuras complejas y compartidos por varios procesos de forma simultánea.

Una **base de datos de un SI** es la representación integrada de los conjuntos de entidades instancia correspondientes a las diferentes entidades tipo del SI y de sus interrelaciones. Esta representación informática debe poder ser utilizada de forma compartida por muchos usuarios de distintos tipos.

El software especializado en la gestión de las bases de datos se denomina **sistema de gestión de bases de datos** (SGBD).

Objetivos y servicios de los SGBD

- Permitir la realización de **consultas no predefinidas** y complejas.
- Ofrecer **flexibilidad a los cambios** (independencia).
- Facilitar la eliminación y la gestión de la **redundancia**.
- Asegurar el mantenimiento de la **calidad** de los datos (integridad).
- Gestionar la **concurrency** de acceso a los datos.
- Proveer servicios de **seguridad**.

2. Arquitectura de los SGBD

El esquema de la DB

Los SGBD necesitan que les demos una descripción o definición de la BD. Esta descripción recibe el nombre de **esquema** de la BD, y los SGBD la tendrán continuamente a su alcance.

Anteriormente, ya hemos hablado de la distinción entre dos niveles de representación informática:

- El **nivel lógico** nos oculta los detalles de cómo se almacenan los datos, cómo se mantienen y cómo se accede físicamente a ellos. En este nivel sólo se habla de entidades, atributos y reglas de integridad.
- Por cuestiones de rendimiento, nos podrá interesar describir elementos de **nivel físico** como qué índices tendremos, cómo y dónde queremos que se agrupen físicamente los registros, de qué tamaño deben ser las páginas, etc.

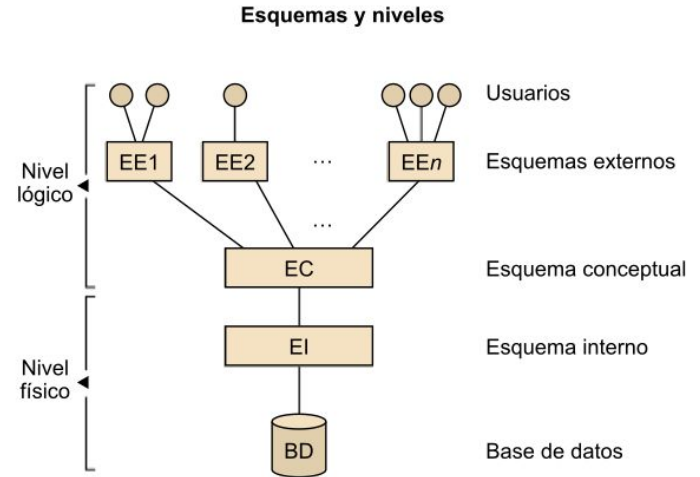
Arquitectura ANSI/SPARC: equivalencias

De acuerdo con la arquitectura **ANSI/SPARC**, debía haber **tres niveles** de esquemas (tres niveles de abstracción). La idea básica de ANSI/SPARC consistía en descomponer el nivel lógico en dos: el nivel **externo** y el nivel **conceptual**. Denominábamos nivel **interno** lo que aquí hemos denominado nivel físico.

Anteriormente	ANSI/SPARC
Lógico	Externo
	Conceptual
Físico	Interno

Arquitectura ANSI/SPARC: los tres niveles I

- En el nivel externo se sitúan las diferentes visiones lógicas que los procesos usuarios (programas de aplicación y usuarios directos) tendrán de las partes de la BD que utilizarán. Estas visiones se denominan **esquemas externos**.
- En el nivel conceptual hay una sola descripción lógica básica, única y global, que denominamos **esquema conceptual**, y que sirve de referencia para el resto de los esquemas.
- En el nivel físico hay una sola descripción física, que denominamos **esquema interno**.



Arquitectura ANSI/SPARC: los tres niveles II

En el **esquema conceptual** se describirán las entidades tipo, sus atributos, las interrelaciones y las restricciones o reglas de integridad.

Al definir un **esquema externo**, se citarán sólo aquellos atributos y aquellas entidades que interesen; los podremos renombrar, podremos definir datos derivados o redefinir una entidad para que las aplicaciones que utilizan este esquema externo creen que son dos, definir combinaciones de entidades para que parezcan una sola, etc.

El **esquema interno** o físico contendrá la descripción de la organización física de la BD: caminos de acceso (índices, hashing, apuntadores, etc.), codificación de los datos, gestión del espacio, tamaño de la página, etc.

Independencia de los datos

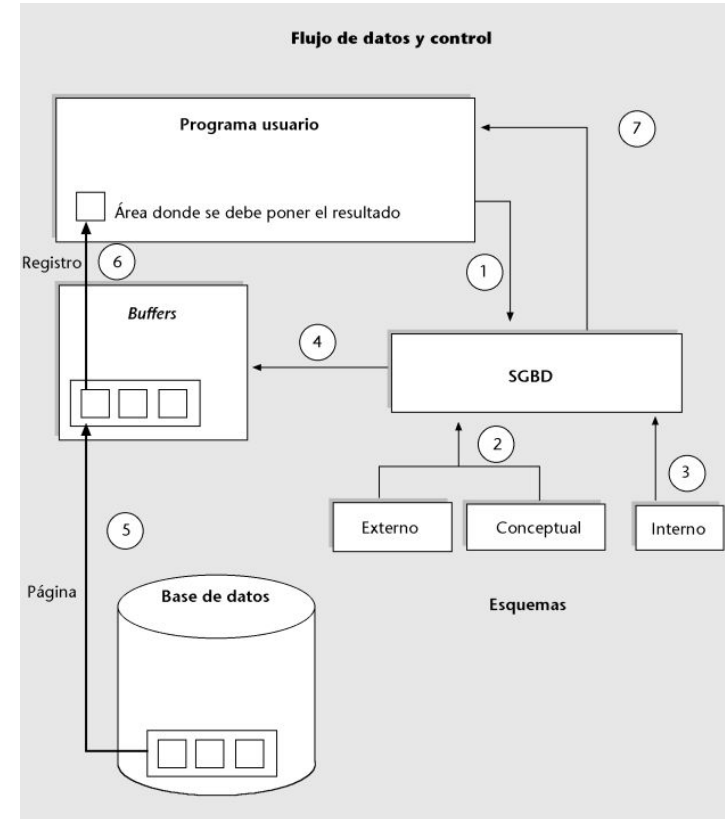
Hay **independencia física** cuando los cambios en la organización física de la BD no afectan al mundo exterior (es decir, los programas usuarios o los usuarios directos).

Hay **independencia lógica** cuando los usuarios no se ven afectados por los cambios en el nivel lógico.

Flujo de datos y de control

Los principales pasos de la ejecución de una consulta enviada al SGBD por un programa de aplicación son:

- A. El programa envía la consulta al SGBD (1) que consulta los esquemas conceptual y externo para validarla (2).
- B. El SGBD determina cómo ejecutar la consulta consultando el esquema interno (3).
- C. El SGBD comprueba si la página de interés está en los buffers y, si no lo está, la carga desde el disco (5).
- D. El SGBD copia los datos al área de trabajo del programa (6) y retorna el control a este (7).



3. Modelos, lenguajes y administración de BD

Modelos de BD

El conjunto de componentes o herramientas conceptuales que un SGBD proporciona para modelar recibe el nombre de **modelo de BD**.

Algunos ejemplos de modelos de BD son el modelo **relacional**, el modelo **jerárquico**, el modelo en **red** y el modelo **relacional con objetos**.

Los modelos de BD nos proporcionan tres tipos de herramientas:

- **Estructuras de datos** con las que se puede construir la BD: tablas, árboles, etc.
- Diferentes tipos de **restricciones** (o reglas) de **integridad** que el SGBD tendrá que hacer cumplir a los datos: dominios, claves, etc.
- Una serie de **operaciones** para trabajar con los datos.

Lenguajes

Para comunicarse con el SGBD, el usuario, ya sea un programa de aplicación o un usuario directo, se vale de un **lenguaje**. Hay muchos lenguajes diferentes, según el tipo de usuarios para los que están pensados (expertos o no informáticos) y el tipo de cosas que los usuarios deben poder expresar con ellos (expresiones complejas o simples).

- Hay lenguajes especializados en la escritura de esquemas; es decir, en la **descripción de la BD**. Se conocen genéricamente como **DDL o Data Definition Language**.
- Otros lenguajes están especializados en la **utilización** de la BD (consultas y mantenimiento). Se conocen como **DML o Data Management Language**.
- Sin embargo, lo más frecuente es que el mismo lenguaje disponga de construcciones para las dos funciones, DDL y DML.
 - Por ejemplo, SQL tiene instrucciones DDL (como CREATE TABLE), instrucciones DML como (SELECT) y adicionalmente instrucciones de control de entorno (p.e. COMMIT).

Administración de bases de datos

Una empresa o institución que tenga SI contruidos en torno a BD necesita que alguien lleve a cabo una serie de funciones centralizadas de gestión y administración, para asegurar que la explotación de la BD es la correcta. Este conjunto de funciones se conoce con el nombre de administración de BD, y los usuarios que hacen este tipo especial de trabajo se denominan **administradores de BD** (ABD).

Un administrador de base de datos se encarga de tareas como:

- Mantenimiento, administración y control de los esquemas.
- Asegurar la máxima disponibilidad de los datos (backups, logs, etc.)
- Resolución de emergencias.
- Vigilancia de la integridad y de la calidad de los datos.
- Diseño físico, estrategia de caminos de acceso y reestructuraciones.
- Control del rendimiento.
- Normativa y asesoramiento a los programadores y a los usuarios finales.
- Control y administración de la seguridad: autorizaciones, restricciones, etc.

