

TEORIA DE BASES DE DATOS

Docentes:

Dra. Claudia Deco

M. Sc. Cristina Bender

Ezequiel Bisiach

OBJETIVO DE LA MATERIA

- Capacitar al alumno en los **conocimientos** fundamentales, **teóricos** y **prácticos**, necesarios para comprender el **funcionamiento** de los sistemas de **Bases de Datos**.
- Al completar este curso, el alumno deberá ser capaz de comprender y aplicar los **conceptos fundamentales** de la **teoría de BD**, desde el modelado de los datos hasta la evaluación y explotación eficiente de la administración de los distintos sistemas de BD.

Contenidos Conceptuales

UNIDAD TEMÁTICA I: Introducción.

- Objetivos de los sistemas de Bases de Datos.**
- Abstracción de datos: Niveles externo, conceptual e interno.**
- Modelos de datos.**
- Lenguaje de definición de datos.**
- Lenguaje de manipulación de datos.**
- Administración de Bases de Datos.**

Contenidos Conceptuales

UNIDAD TEMÁTICA II: El modelo relacional.

- Estructura del Modelo Relacional: dominios y atributos, relaciones, claves primarias y externas.**
- Restricciones de integridad: integridad de entidades, integridad referencial. Restricciones de dominio.**

Contenidos Conceptuales

UNIDAD TEMÁTICA III: Lenguajes relacionales

- **Álgebra Relacional: operaciones primitivas, operaciones derivadas.**
- **Cálculo Relacional: cálculo relacional de Tuplas y de Dominios.**
- **Complejidad Relacional.**
- **Álgebra vs Cálculo: lenguajes procedurales y no procedurales.**

Contenidos Conceptuales

UNIDAD TEMÁTICA IV: El lenguaje estándar SQL

- Lenguaje de manipulación de datos (DML).**
- Lenguaje de definición de datos (DDL).**
- Catálogo del Sistema. Vistas.**

Contenidos Conceptuales

UNIDAD TEMÁTICA V: Diseño de Bases de Datos Relacionales.

- Representación de la información. Pérdida de la información.
- Dependencias funcionales, Axiomas de Armstrong;
- Teoría formal de la Normalización de esquemas relacionales.
Definición formal de las tres primeras formas normales.
- Formas normales avanzadas. Teoría de las dependencias multivaluadas. Cuarta y quinta formas normales.
- Normalización vs. Performance.

Contenidos Conceptuales

UNIDAD TEMÁTICA VI: Integridad y Seguridad.

- Autorizaciones y Vistas.**
- Restricciones de Integridad.**
- Especificación de la seguridad en SQL.**

Contenidos Conceptuales

UNIDAD TEMÁTICA VII: Control de Concurrencia y Recuperación.

- Transacciones.**
- Protocolos para el Control de Concurrencia.**
- Problemas principales de concurrencia**
- Recuperación de Transacciones.**

Contenidos Conceptuales

UNIDAD TEMÁTICA VIII: Optimización de consultas

- Interpretación de consultas.**
- Equivalencia de expresiones.**

Contenidos Conceptuales

UNIDAD TEMÁTICA IX: Bases de datos distribuidas

- Estructura de las bases de datos distribuidas.**
- Diseño de las bases de datos distribuidas.**
- Control de Concurrency y Recuperación.**
- Optimización de consultas.**

BIBLIOGRAFÍA

Principal:

- **Korth, H.F.; Silberschatz, A. - Fundamentos de Bases de Datos. - McGraw-Hill.**
- **Date, C.J. - Introducción a los Sistemas de Bases de Datos. Vol I - Addison-Wesley Iberoamericana.**

Adicional:

- **Ullman, J.D.; Jeffrey, R. - Principles of Database Systems - Computer Science Press.**
- **De Miguel, A.; Piattini, M.- Concepción y Diseño de Bases de Datos Rama.**
- **Elmasri, R.; Navathe, S.B. - Sistemas de Bases de Datos - Conceptos Fundamentales- Addison-Wesley Iberoamericana.**

Sistemas de gestión de BD

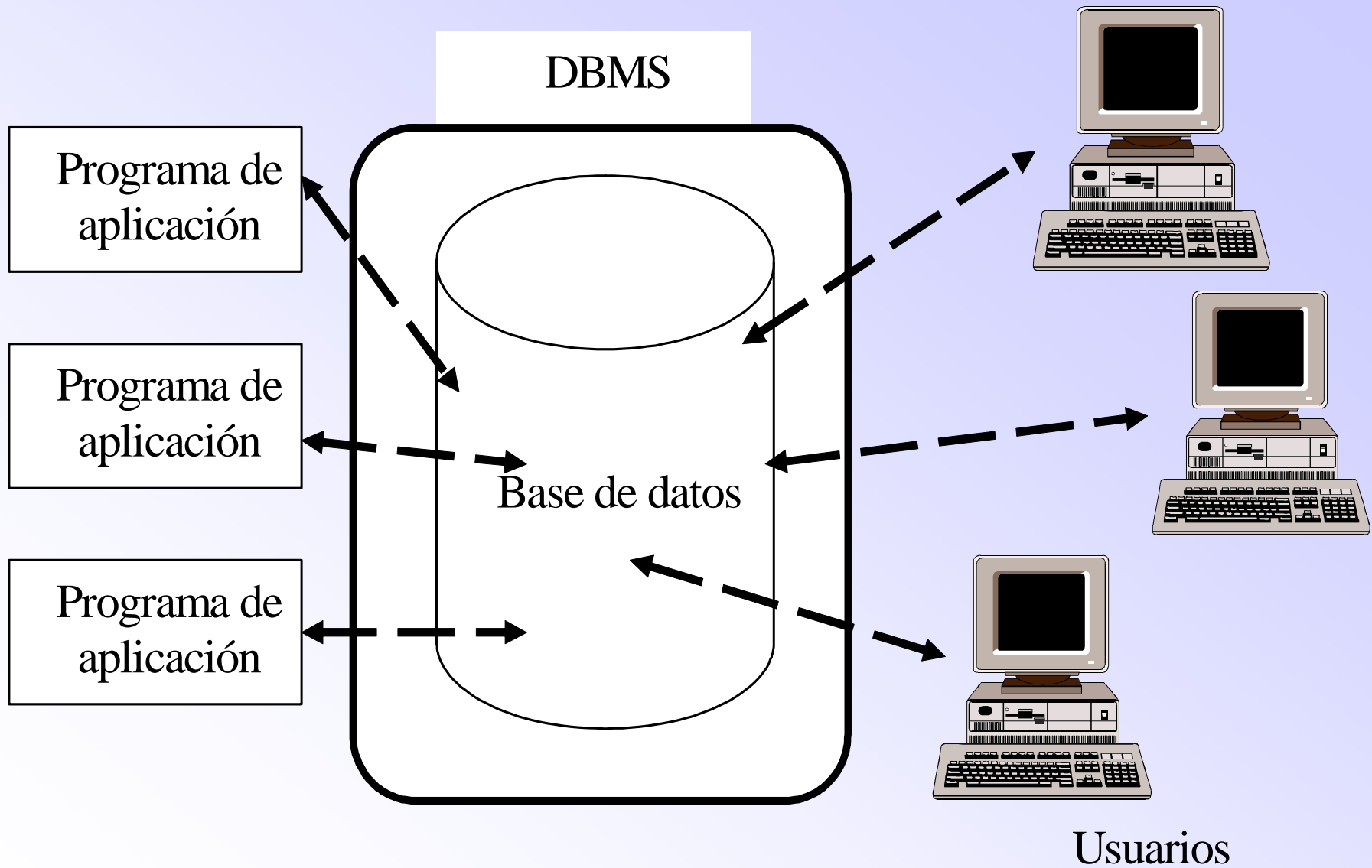
Data Base Management System (DBMS)

Consiste en:

- **una colección de datos interrelacionados y**
- **un conjunto de programas para acceder a esos datos.**

Objetivo:

proporcionar un entorno *conveniente y eficiente* para extraer y almacenar información.



Sistemas de gestión de BD

Están diseñados para gestionar grandes bloques de información:

- ***Definición de estructuras para el almacenamiento de información***
- ***Provee mecanismos para la gestión de información***
- ***Mantenimiento de la seguridad de la información almacenada (caídas del sistema, accesos no autorizados)***
- ***Control de concurrencia: Provee mecanismos que eviten posibles resultados anómalos en datos compartidos por varios usuarios***

Sistema de Procesamiento de Archivos

Redundancia e inconsistencia de datos:

Los **archivos** y los **programas** de aplicación son creados por **distintos programadores** en **distintos momentos**, por lo cual es probable que:

- los archivos tengan **diferentes formatos** y
- **programas** puedan estar **duplicados** en varios sitios

Sistema de Procesamiento de Archivos

Dificultad para tener acceso a los datos:

Si una **solicitud** no fue **prevista** al diseñarse el sistema original, no habrá **ningún programa** de aplicación que la satisfaga.

Las **alternativas** serán:

- realizar una **extracción manual** desde informes ya existentes ó
- escribir un **nuevo programa de aplicación** necesario.

Ninguna es satisfactoria.

Deben desarrollarse sistemas de **recuperación de datos para uso general.**

Sistema de Procesamiento de Archivos

Aislamiento de los datos:

- Los **datos** están **repartidos** en varios archivos, y pueden tener **formatos diferentes**
- Es difícil **escribir nuevos programas** de aplicación para obtener los datos apropiados.

Sistema de Procesamiento de Archivos

Anomalías del acceso concurrente:

- Se puede **acceder a los datos** por medio de diversos **programas de aplicación** diferentes que **no** han sido **previamente coordinados**
- La supervisión es muy difícil de proporcionar.

Sistema de Procesamiento de Archivos

Problemas de seguridad:

- **No todos** los usuarios del sistema de BD **deben poder acceder a todos** los datos.
- Puesto que los programas de aplicación se añaden al sistema de una forma precisa, es **difícil implantar las restricciones** de seguridad.

Ventajas del enfoque de BD

➤ **Es posible disminuir la redundancia.**

➤ **Es posible evitar la inconsistencia.**

➤ **Es posible compartir datos.**

Aplicaciones ya existentes pueden compartir información de la BD.

Se pueden desarrollar aplicaciones nuevas para trabajar con los mismos datos almacenados.

➤ **Es posible hacer cumplir las normas.**

El **DBA** (control centralizado sobre la BD) puede garantizar la aplicación de normas para la representación de los datos.

Ventajas del enfoque de Base de Datos

➤ **Es posible aplicar restricciones de seguridad.**

El DBA puede **asegurar** que el **acceso** a la BD sea **sólo** a través de los canales apropiados

y, por tanto, puede **definir** las **verificaciones** de **seguridad** por realizar cuando se intente acceder a información restringida.

➤ **Es posible mantener la integridad.**

El DBA puede **definir verificaciones** de **integridad** que deben realizarse en toda operación de actualización de datos.

Abstracción de datos

Un objetivo importante del DBMS es:

- **dar** a los usuarios una **visión *abstracta*** de los datos,
y
- **extraerlos en forma eficiente.**

Abstracción de datos

√ **Nivel físico (ó interno).**

- Describe *cómo* se almacenan los datos.
- Se describen en detalle las estructuras de datos complejas.

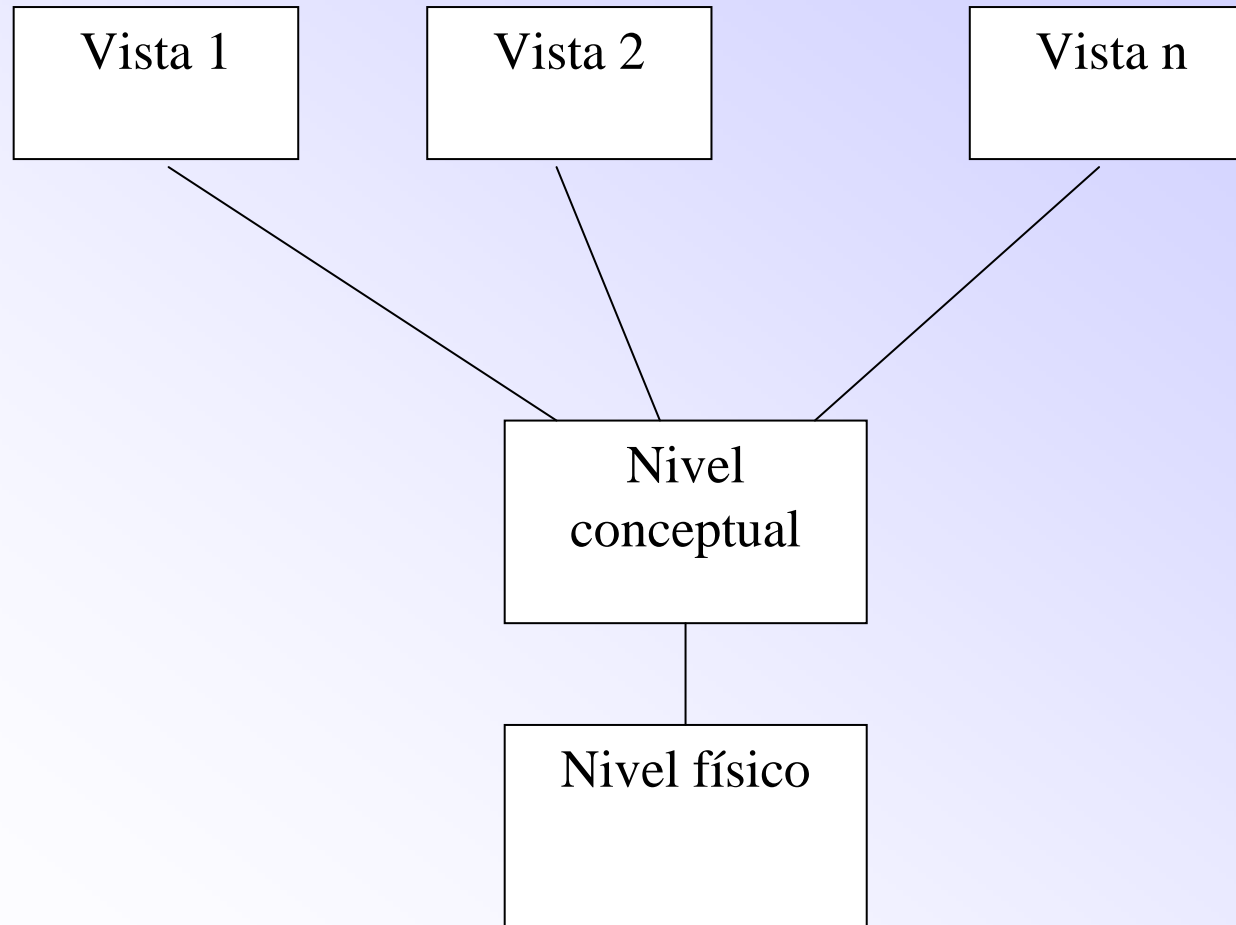
√ **Nivel conceptual.** (usado por los **administradores de BD**)

- Describe qué datos son realmente almacenados en la BD y las relaciones que existen entre los datos.

√ **Nivel de visión (ó externo).** (usado por el **usuario**)

- Describe sólo parte de la BD completa.

Abstracción de datos



Modelos de datos

Modelo de datos es:

- Colección de **herramientas conceptuales** para describir **datos**, **relaciones** entre ellos, **semántica** asociada a los datos y **restricciones** de consistencia.
- Permiten describir la **estructura de una BD**.

Modelos de datos

a) Modelos lógicos basados en objetos :

Se usan para describir datos en los **niveles conceptual y de visión.**

b) Modelos lógicos basados en registros:

Se usan para describir datos en los **niveles conceptual y físico.**

Modelos de datos

El modelo entidad-relación (E-R)

Se basa en una percepción de un mundo real
que consiste en
una colección de objetos básicos llamados *entidades*
y
relaciones entre estos objetos.

Modelos de datos

El modelo entidad-relación (E-R)

- ***Entidad:***

objeto distinguible de otros
por medio de un conjunto específico de atributos

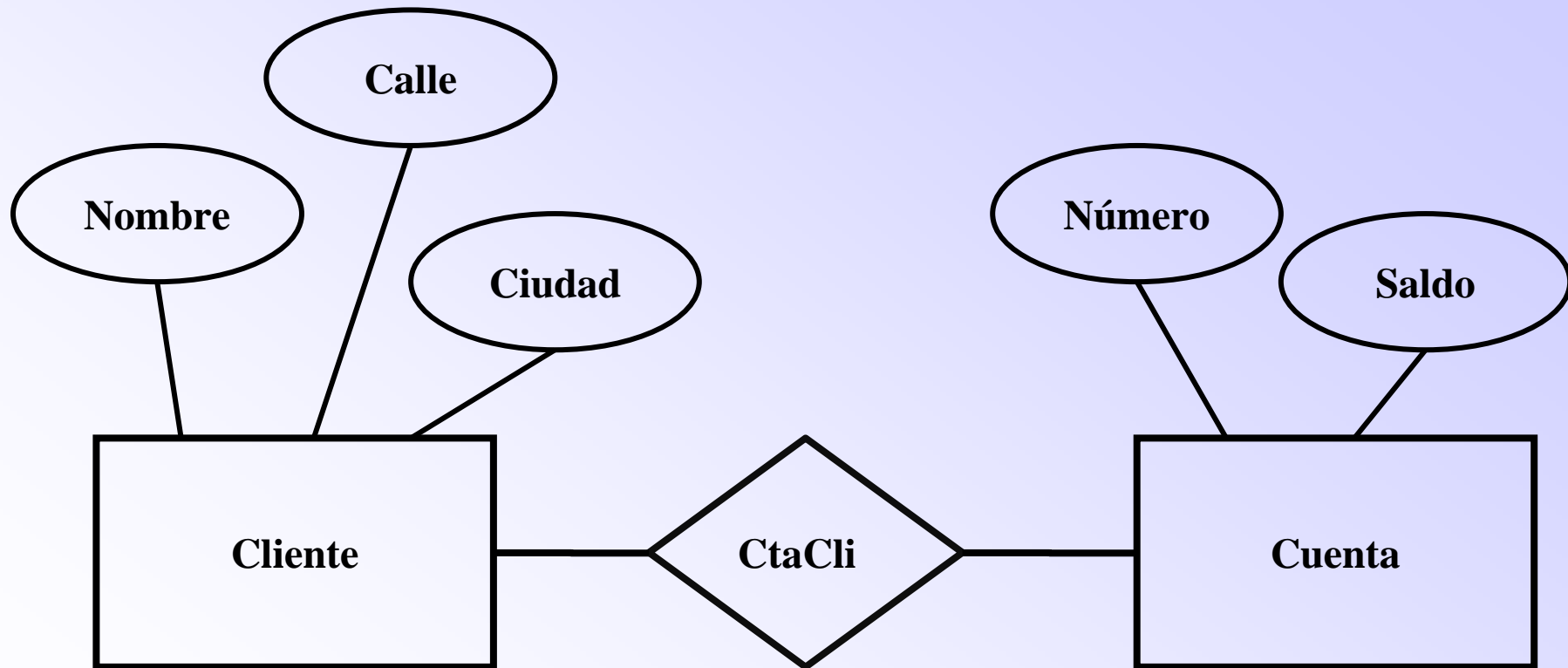
Ej.: **número** y **saldo** describen una cuenta de un banco.

- ***Relación:***

asociación entre entidades.

Ej.: la relación **CtaCli** asocia a un cliente con cada una de las cuentas que tiene.

Diagrama E / R



Modelos de datos

El modelo relacional

- Representa los **datos** y las **relaciones** entre los datos mediante una colección de **tablas**,
- cada una de las cuales tiene un número de columnas con nombres únicos.

El modelo relacional

Nombre	Calle	Ciudad	Número
Lowery	Maple	Queens	900
Shiver	North	Bronx	556
Shiver	North	Bronx	647
Hodges	Sidehill	Brooklyn	801
Hodges	Sidehill	Brooklyn	647

Instancias y esquemas

Instancia de la BD es

la colección de información almacenada en la BD

en un determinado momento en el tiempo.

Instancias y esquemas

Esquema de la BD es
el diseño global de la BD.

Los sistemas de BD tienen **varios** esquemas:

- un **esquema físico** (nivel de abstracción más bajo),
- un **esquema conceptual** (nivel intermedio) y
- uno ó más **subesquemas** (nivel más alto - vistas).

Independencia de datos

Capacidad de **modificar una definición de un esquema** en un nivel sin afectar la definición de un esquema superior siguiente.

Niveles:

√ Independencia **física** de datos:

capacidad de modificar el esquema **físico** sin tener volver a escribir los programas de aplicación.

√ Independencia **lógica** de datos:

capacidad de modificar el esquema **conceptual** (alterar la estructura lógica de la BD), sin tener que volver a escribir los programas de aplicación

Ej.: añadir un nuevo campo.

Lenguaje de definición de datos

Lenguaje de definición de datos (DDL):

conjunto de definiciones que especifica un esquema de BD.

Diccionario de datos (catálogo ó directorio):

- archivo que almacena un conjunto de tablas.
- contiene **metadatos**, y
- se consulta **antes** de leer ó modificar los datos reales en el sistema de BD.

Lenguaje de manipulación de datos

Manipulación de datos significa:

recuperación - inserción - supresión - modificación
de datos almacenados en la BD.

Lenguaje de manipulación de datos (DML):

permite a los usuarios a acceder ó manipular datos.

- **Procedimentales**, el usuario debe especificar **qué** datos se necesitan y **cómo** obtenerlos.
- **No procedimentales**, el usuario debe especificar **qué** datos se necesitan **sin** especificar cómo obtenerlos.

Una **consulta** es una sentencia que solicita la recuperación de información. (**lenguaje de consultas**).

Gestor de BD (DBMS)

Un **gestor de BD** es un
módulo de programa que proporciona
el **interfaz** entre los datos de bajo nivel almacenados en la BD
y los programas de aplicación y consultas hechos al sistema.

El gestor de BD es responsable de las siguientes **tareas**:

Gestor de BD (DBMS)

- **Definición y manipulación de datos.**

-El DBMS **traduce** las distintas sentencias DDL y DML a comandos del sistema de archivos de bajo nivel.

-Es responsable del verdadero almacenamiento, recuperación y actualización de los datos en la BD.

Gestor de BD (DBMS)

- **Implantación de la integridad.**

-Los valores de los datos que se almacenan en la BD deben satisfacer ciertos tipos de restricciones de consistencia.

Ejemplo: saldo de una cuenta bancaria no negativo.

-El **DBA** debe **especificar** explícitamente estas restricciones.

-El **DBMS** **controla** si se violan estas restricciones y toma las acciones apropiadas.

Gestor de BD (DBMS)

- **Implantación de la seguridad.**

-No todos los usuarios de la BD necesitan tener acceso a todo su contenido.

-El **DBMS** hace que se **cumplan** los requisitos de seguridad **definidos** por el **DBA**.

Gestor de BD (DBMS)

- **Copia de seguridad y recuperación.**

- Es responsabilidad del DBMS detectar fallos (rotura de disco, corte de energía, etc.) y

- restaurar la BD al estado que existía **antes** de ocurrir el fallo.

Gestor de BD (DBMS)

- **Control de concurrencia.**

-Cuando varios usuarios actualizan la BD concurrentemente, es posible que no se conserve la consistencia de los datos.

-El **DBMS** controla la **interacción** entre los **usuarios concurrentes**.

Administrador de BD (DBA)

El **administrador de datos (DA)** es la persona que toma las decisiones estratégicas y de política con respecto a la información de la empresa.

El **DBA** es quien proporciona el apoyo técnico necesario para poner en práctica las decisiones del DA. Está encargado del **control general del sistema** en el nivel técnico.

Administrador de BD (DBA)

Las **funciones** del DBA incluyen:

- **Definición del esquema conceptual.**
 - Realizado el diseño lógico de la BD, el esquema conceptual se crea escribiendo un conjunto de definiciones DDL.
 - El DBMS traduce estas sentencias a un conjunto de tablas.

Administrador de BD (DBA)

- **Definición del esquema interno.** (Definición de la estructura de almacenamiento y del método de acceso).
 - El DBA debe decidir cómo se representará la información en la BD almacenada (diseño físico).
 - **Ejemplo:** disco donde almacenar datos, definición de índices, etc.

Administrador de BD (DBA)

- **Modificación del esquema y de la ubicación física.**
 - El DBA debe supervisar el desempeño y realizar los ajustes apropiados cuando cambien los requerimientos.

Administrador de BD (DBA)

- **Concesión de autorización para el acceso a los datos.**
 - Permite al DBA regular qué partes de la BD van a poder ser accedidas por varios usuarios.

Administrador de BD (DBA)

- **Especificación de las restricciones de integridad.**
 - Los valores de los datos que se almacenan en la BD deben satisfacer ciertos tipos de restricciones de consistencia.
 - **Ejemplo:** el saldo de una cuenta bancaria no debe ser negativo.
 - El DBA debe especificar explícitamente estas restricciones.

Administrador de BD (DBA)

- Definición de procedimientos de respaldo y recuperación.**

Usuarios de BD

1. Programadores de aplicaciones.

- Interaccionan con el sistema por medio de **llamadas en DML** incorporadas en un programa escrito en un lenguaje principal.
- Un ***precompilador de DML***, convierte las sentencias DML a llamadas normales a procedimientos en el lenguaje principal.

2. Usuarios que escriben sus preguntas con un lenguaje de consulta de BD.

3. Usuarios que interactúan con el sistema invocando a uno de los programas de aplicación existentes.

Componentes funcionales de un DBMS

