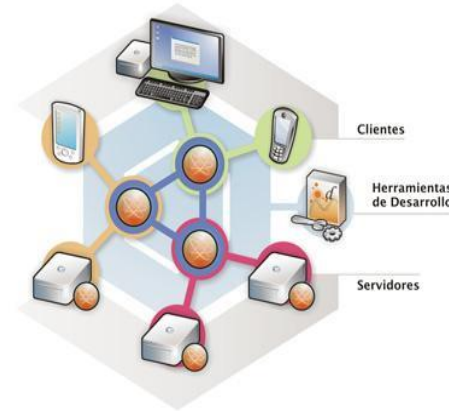
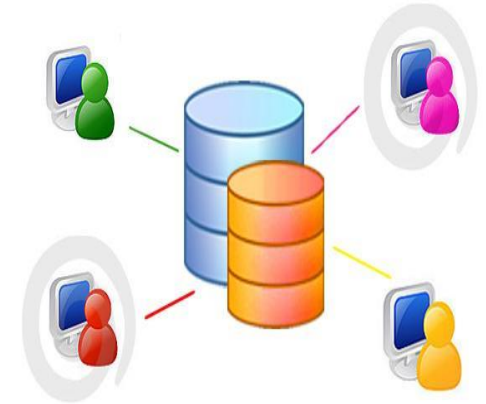


Tema 1

Conceptos Básicos

**Grado en
Ingeniería
Informática**



**Bases de
Datos**

2021/22

Departamento de Tecnologías de la Información
Universidad de Huelva

Objetivos

- ☐ Conocer el concepto de base de datos y sus características principales
- ☐ Conocer el concepto de Sistema de Gestión de Bases de Datos (SGBD)
- ☐ Comprender la importancia de la independencia de los datos respecto a los programas

Contenidos

- 1.1 Introducción. Conceptos: BD / SGBD
- 1.2 Sistemas de bases de datos vs. Sistemas de procesamiento de ficheros
- 1.3 Personas que interactúan con la BD
- 1.4 Ventajas de utilizar un SGBD
- 1.5 Arquitectura de un SGBD. Abstracción e independencia de los datos
- 1.6 Modelos de datos, esquemas e instancias
- 1.7 Lenguajes de los SGBD
- 1.8 Componentes de un SGBD
- 1.9 Clasificación de los SGBD

Duración

- ☐ 2 clases

Bibliografía

- ☐ Capítulos 1 y 2 de [Elmasri 02]
- ☐ Capítulo 1 de [Silberschatz 06]
- ☐ Capítulo 1 y 2 de [Connolly 05]

1.1 Introducción. Conceptos: BD / SGBD

Base de Datos (BD)

colección de datos relacionados con un significado implícito

Propiedades implícitas de una BD:

- Representa algunos aspectos del *mundo real*, también llamado **minimundo** o **universo del discurso**
- Se diseña, construye y carga con datos, para un **propósito específico**. Está destinada a un grupo de usuarios concreto y suelen existir aplicaciones preconcebidas en las que están interesados estos usuarios
- Es una **colección coherente de datos**. Un conjunto aleatorio de datos no puede considerarse como una BD

En resumen, una BD tiene alguna fuente de la que provienen los datos, algún grado de interacción con sucesos del universo de discurso y una serie de usuarios interesados en su contenido

Sistema de Gestión de Bases de Datos (SGBD)

(DBMS: Database Management System)

conjunto de programas que permiten a los usuarios crear y mantener una base de datos

Un **SGBD** es por tanto un *sistema software* que facilita los **procesos** de:

Definición: Especificar los tipos de los datos, las estructuras y las restricciones para los datos que se van a almacenar en la BD

Construcción: Proceso de almacenar los datos en algún medio de almacenamiento controlado por el SGBD

Manipulación: Engloba tareas tales como consultar los datos para obtener información específica, actualizar la base de datos (insertar, modificar o eliminar datos) para reflejar los cambios ocurridos en el minimundo y generar informes a partir de los datos

Un **Sistema de Base de Datos** es el conjunto formado por la Base de Datos y el SGBD

SGBD más utilizados
actualmente

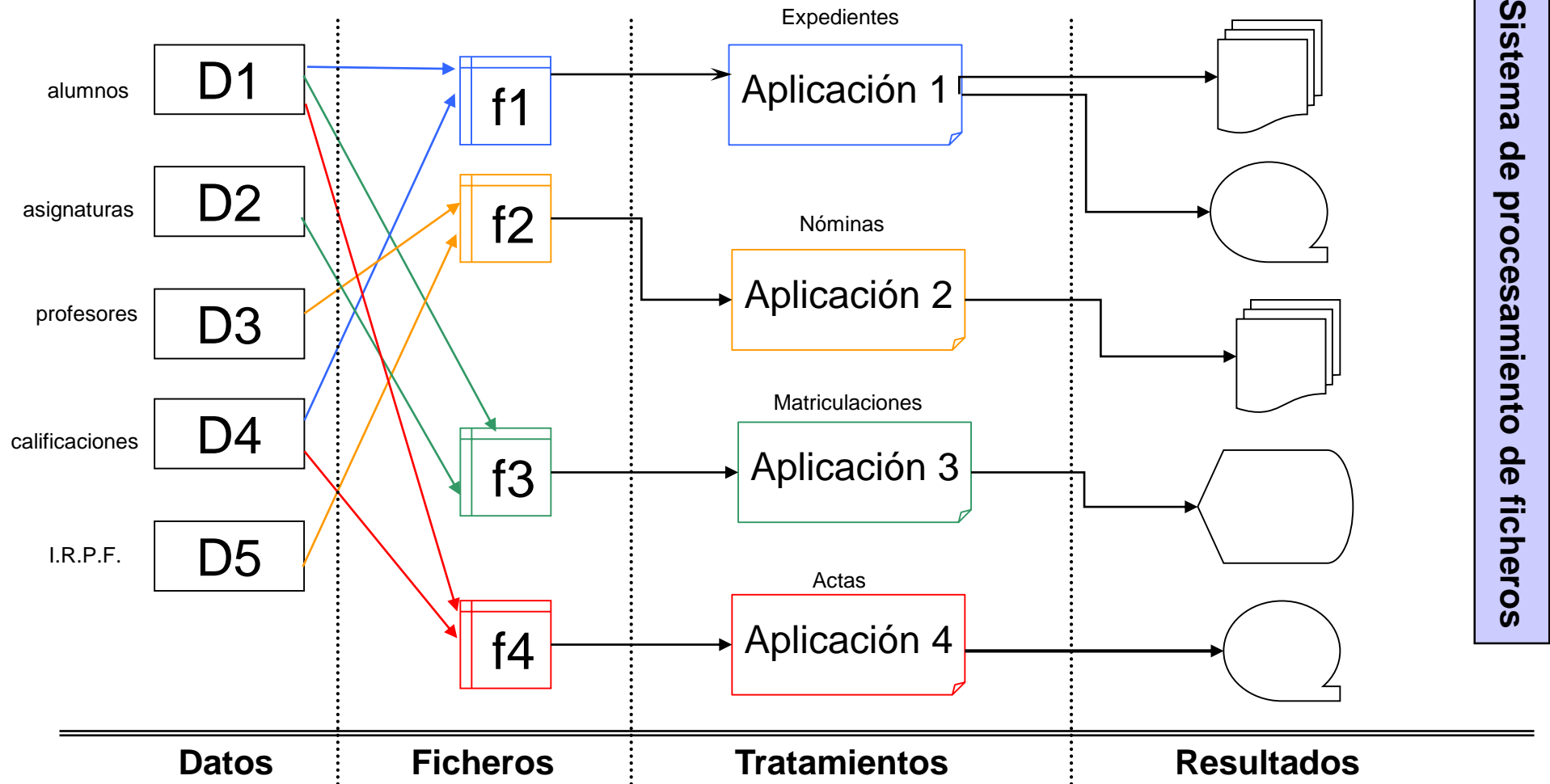


ORACLE®

SYBASE® | An **SAP** Company
Adaptive Server Enterprise

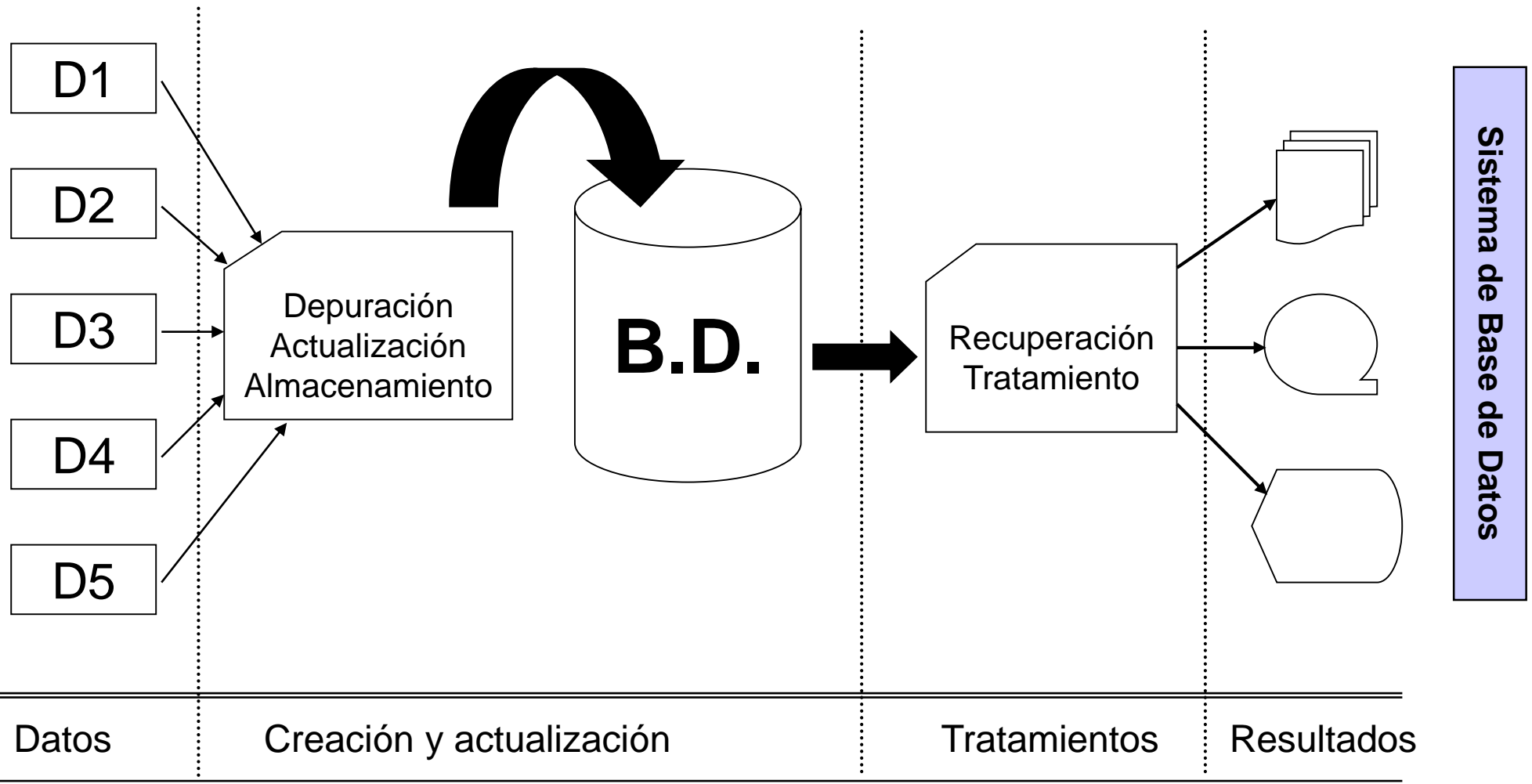


1.2 Sistemas de bases de datos vs. Sistemas de procesamiento de ficheros



Desventajas principales de los sistemas de *procesamiento de ficheros*

- Redundancia e inconsistencia de los datos
- Dificultad para recuperar datos de forma conveniente y eficiente
- Fallos en el acceso concurrente: Varios usuarios pueden actualizar datos de forma simultanea
- Problemas de seguridad: No todos los usuario deben poder acceder a la totalidad de los datos
- Problemas de integridad: Los valores de los datos almacenados deben satisfacer ciertos tipos de restricciones de consistencia



Ventajas de las *BD* frente a los sistemas de *procesamiento de ficheros*

- Naturaleza autodescriptiva de los sistemas de bases de datos
 - La BD no sólo contiene *datos*, sino también una definición de la estructura de la BD y sus restricciones, que se almacena en el **catálogo** o **diccionario de datos (DD)**
- Independencia entre los programas y los datos
 - La estructura de los almacenes de datos se guarda en el DD del SGBD separadamente de los programas de acceso a los datos
- Coherencia en los resultados de las consultas
 - Como los datos están almacenados una sola vez y en un solo lugar, todos los programas harán uso de los mismos datos
- Mejor disponibilidad de los datos para los usuarios
 - Dado que los datos se comparten entre el conjunto de aplicaciones, todos los usuarios pueden disponer de ellos (siempre que tengan autorización)

- Mayor eficiencia en la recogida, validación e introducción de los datos en el sistema
 - Al no existir redundancias, los datos se recogen y validan una sola vez, aumentando el rendimiento del proceso previo al almacenamiento.
- Reducción del espacio de almacenamiento
 - Al no haber redundancia en el almacenamiento de los datos
- Soporte de múltiples vistas de los datos
 - Cada uno de los usuarios de una BD puede requerir una perspectiva o vista diferente de la BD global

1.3 Personas que interactúan con la BD

Administrador de base de datos

Persona encargada de gestionar y controlar la utilización de la base de datos, el SGBD y el software relacionado

- ✓ Concesión de autorización para el acceso a los datos
- ✓ Adquirir los recursos hardware y software necesarios
- ✓ Velar por la seguridad de la base de datos
- ✓ Garantizar un tiempo de respuesta adecuado

Diseñador de base de datos

Persona encargada de comunicarse con los usuarios de la base de datos, comprender sus necesidades y realizar un diseño que satisfaga esos requerimientos

- ☒ Definición del esquema
- ☒ Definición de las estructuras de almacenamiento
- ☒ Desarrollo de **vistas** adecuadas para cada grupo de usuarios

Usuario final

Persona cuyo trabajo requiere acceder a la base de datos para consultarla, actualizarla o generar informes. Categorías:

☒ Usuario ocasional

- Accede de vez en cuando pero puede que necesite información distinta en cada ocasión
- Utiliza un lenguaje de consulta avanzado para especificar sus solicitudes
- Ej: gerentes de nivel medio o alto

☒ Usuario simple o paramétrico

- Realiza consultas y actualizaciones constantes de la BD
- Utiliza tipos estándar de consultas y actualizaciones incluidas en programas de aplicación
- Constituye una parte considerable del total de usuarios finales
- Ej: cajeros de bancos, encargados de reservas (vuelos, hoteles, ...)

Usuario final

☒ Usuario avanzado

- Persona muy familiarizada con los recursos del SGBD como para implementar aplicaciones de requerimientos complejos
- Ej: ingenieros, científicos, analistas de negocio, ...

☒ Usuario autónomo

- Persona que mantiene BD personales mediante la utilización de software comercial que cuenta con interfaces de fácil uso basado en menús o gráficos

Analista de sistemas y programador de aplicaciones

☒ Analista de sistemas

- Determina los requerimientos de los usuarios finales (sobre todo de los paramétricos) y desarrolla especificaciones para programas de aplicación que satisfagan dichos requerimientos

☒ Programador de aplicaciones

- Implementa, prueba, depura, documenta y mantiene los programas de aplicación

1.4 Ventajas de utilizar un SGDB

❑ Control de la redundancia

- La redundancia consiste en el almacenamiento de los mismos datos varias veces
- 3 problemas principales:
 - Multiplicación del trabajo en los procesos de actualización de los datos
 - Ocupación innecesaria de espacio de almacenamiento
 - Elevada probabilidad de inconsistencia
- La solución es mantener cada dato en un único lugar de la Base de Datos
- A veces resulta conveniente la redundancia controlada para mejorar el rendimiento de algunas consultas. El SGBD debe controlar, de forma automática, dicha redundancia para no producir inconsistencias

❑ Restricción de los accesos no autorizados

- Generalmente, distintos tipos de usuarios acceden a la misma base de datos. No todos los usuarios tienen la misma autorización para acceder la información
- También es habitual que unos usuarios tengan autorización según el tipo de operación (recuperación, inserción, actualización y borrado)
- El SGBD debe tener un sistema de seguridad y autorización basado en cuentas de usuario con restricciones

❑ Capacidad de realizar acciones usando reglas

- Los SGBD actuales poseen la capacidad de almacenar reglas que se activan cuando se produce un determinado evento en la base de datos. Habitualmente se realizan con los disparadores (*triggers*)

❑ Posibilidad de ofrecer diferentes interfaces de usuario

- Lenguajes de consulta
- Interfaces de lenguaje de programación
- Formularios

❑ Mayor integridad en los datos

- Las aplicaciones de bases de datos suelen tener restricciones de integridad que deben cumplir los datos
- El SGBD debe ofrecer mecanismos para definirlas y garantizar que se cumplan
- La integridad de los datos depende de la semántica del minimundo que representa
- Las restricciones de integridad se deben detectar en la etapa de diseño
- Algunas restricciones se pueden especificar en el SGBD y otras más complejas pueden requerir controles especiales cuando se introducen los datos

❑ Facilidad para realizar copias de seguridad y recuperación

- Si el sistema falla durante la ejecución de una actualización compleja, el sistema debe recuperar la base de datos al estado en el que se encontraba antes de comenzar el proceso de actualización

❑ Mejor accesibilidad a los datos y mayor capacidad de respuesta

- La mayoría de los SGBD proporcionan lenguajes de consultas que permiten a los usuarios obtener información, de forma inmediata, sin necesidad de utilizar una aplicación realizada por un programador

❑ Mayor productividad

- Los SGBD proporcionan muchas funciones que el programador debería incluir en una aplicación basada en archivos
- Esto permite que el programador se concentre en la funcionalidad requerida por el usuario
- Con esto se mejora la productividad de los programadores y se reducen los tiempos de desarrollo y, por consiguiente, los costes

1.5 Arquitectura de un SGBD. Abstracción e independencia de datos

Arquitectura de tres esquemas

Objetivo: separar las aplicaciones del usuario y la base de datos física

1. Nivel interno

- Describe la estructura física de almacenamiento de la BD
- El **esquema interno** emplea un modelo de datos físicos y describe todos los detalles del almacenamiento

2. Nivel conceptual

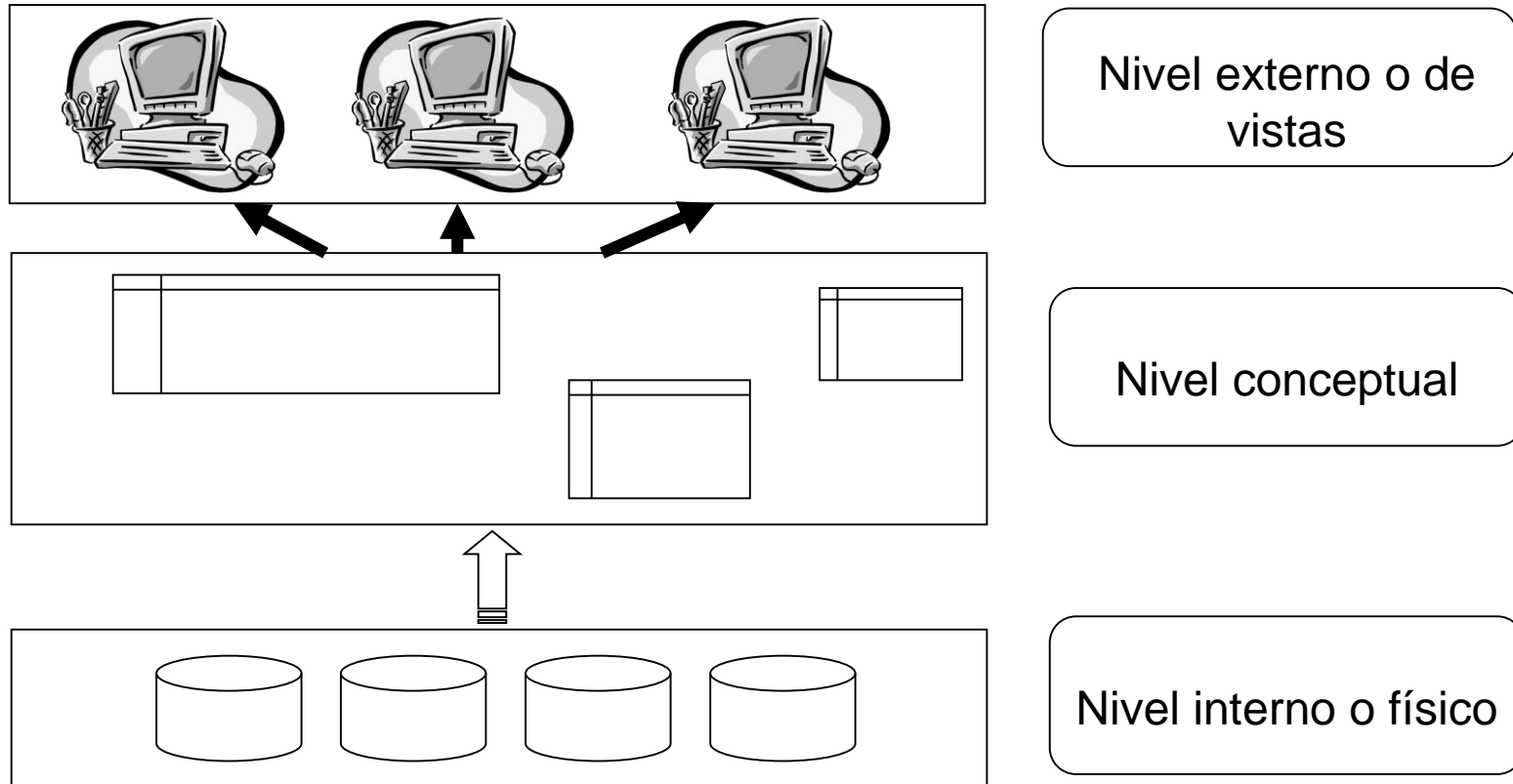
- Describe la estructura completa de la BD
- El **esquema conceptual** describe entidades, tipos de datos, vínculos y restricciones, ocultando los detalles de las estructuras físicas de almacenamiento

Arquitectura de tres esquemas

3. Nivel externo o de vistas

- Incluye varios **esquemas externos o vistas de usuario**
 - Cada vista describe la parte de la BD que interesa a un grupo de usuarios determinado y oculta a ese grupo el resto de la BD
- ☑ La mayoría de los SGBD soportan la arquitectura en tres capas, aunque no separan los niveles completamente
- Algunos SGBD incluyen detalles de nivel físico en el esquema conceptual
 - En casi todos los SGBD, los esquemas externos se especifican en el mismo modelo de datos que se describe la información a nivel conceptual
- ☑ Los esquemas son *descripciones* de los datos. Los datos están realmente en el nivel físico

Arquitectura de tres esquemas

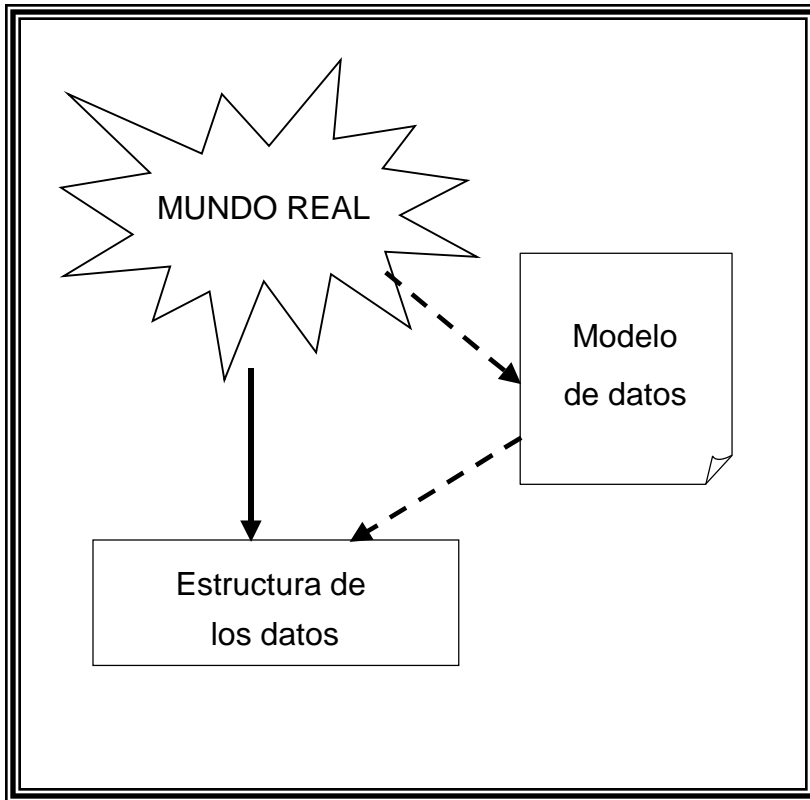


- ✓ El proceso de transformar solicitudes y resultados se denomina **correspondencia** o **transformación**

Independencia de datos

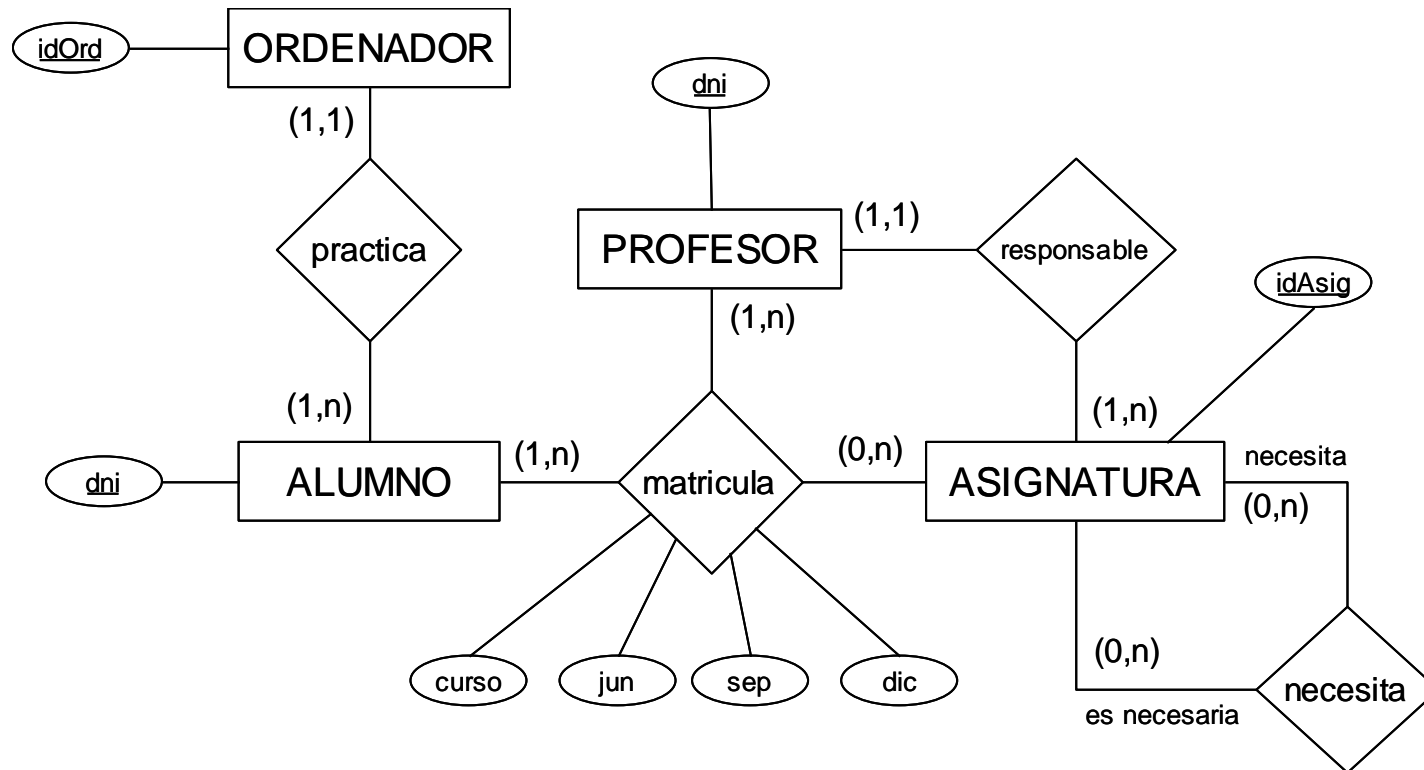
- ☑ Capacidad para modificar el esquema en un nivel del sistema de BD sin tener que modificar el esquema del nivel inmediato superior
 - ☑ La independencia es posible porque cuando cambia un esquema sólo se modifica la información almacenada en el diccionario de datos sobre la correspondencia entre niveles
1. Independencia lógica de los datos
 - Capacidad de modificar el esquema conceptual sin tener que alterar los esquemas externos ni los programas de aplicación
 - En algunos casos, las modificaciones pueden afectar a los esquemas y programas que usen los elementos modificados, pero no a los restantes
 2. Independencia física de los datos
 - Capacidad de modificar el esquema interno sin tener que alterar el esquema conceptual

1.6 Modelos de datos, esquemas e instancias



- Modelos de datos conceptuales o de alto nivel
 - Modelo entidad-relación
 - Modelo orientado a objetos
- Modelos de datos lógicos (de representación o implementación)
 - Modelo relacional
 - Modelo de red
 - Modelo jerárquico
- Modelos de datos de bajo nivel o físicos

Modelo Entidad-Relación



Modelos de datos lógicos

Modelo Relacional

Nombre	Dirección	Ciudad	Número
García	C/ Rábida	Huelva	158
Pérez	C/ Rascón	Huelva	125
Martín	C/ Arroyo	Cádiz	402
Martín	C/ Arroyo	Cádiz	158

tabla CLIENTE

Número	Saldo
158	125000
125	75000
402	255000

tabla CUENTA

Instancias y esquemas

- **Esquema de una BD**

- Diseño global de la base de datos a nivel conceptual (datos, relaciones, restricciones, etc.)
- Una vez definido, no suele cambiar

Director (nombreDirector, nacionalidad, fechaNacimiento, numPelículas)

Película (títuloPelícula, director, género, guión, añoRodaje, nacionalidad, duración)

- **Instancia o ejemplar de una BD**

- Información que se encuentra almacenada en una base de datos en un momento determinado
- Es la parte cambiante del sistema

1.7 Lenguajes de los SGBD

❑ Lenguaje de definición de datos (LDD) o *Data Definition Language* (DDL)

- Se utiliza para crear la estructura de la BD y, generalmente, para crear las vistas del nivel externo
- Mediante el LDD se describen y nombran las entidades, atributos y relaciones. También se definen las restricciones de integridad y seguridad
- Al compilarse las instrucciones LDD se genera información que se almacena en el **catálogo del sistema o diccionario de datos**
- El diccionario de datos contiene los **metadatos**, que describen los objetos de la base de datos y que hacen más fácil el acceso y manipulación de dichos objetos
- A nivel teórico se distinguen LDD diferentes para cada esquema de la arquitectura: un LDD para los esquemas externos, otro para el conceptual y otro para el interno. En la práctica suele existir un único lenguaje LDD que permite la especificación de, al menos, los esquemas externo y conceptual

❑ Lenguaje de manipulación de datos (LMD) o *Data Manipulation Language* (DML)

- Lenguaje que proporciona un conjunto de operadores para permitir las manipulaciones básicas de los datos de una base de datos
- Operaciones de manipulación:
 - inserción de nuevos datos
 - modificación de los datos ya almacenados
 - borrado de datos
 - extracción de los datos almacenados en la base de datos
- La parte del LMD relacionada con la extracción de los datos se suele denominar **lenguaje de consulta**
- Dos tipos fundamentales de LMD:
 - **Procedimentales**: hay que especificar cómo obtener los datos
 - **No procedimentales**: sólo hay que especificar qué datos se quieren obtener y no cómo obtenerlos. Los SGBD incluyen algún lenguaje de este tipo, generalmente **SQL** (*Structured Query Language*) o **QBE** (*Query By Example*)

❑ ¿Cómo interactuar con una BD desde un lenguaje de propósito general?

- **SQL embebido (incrustado) en el lenguaje**

- se intercalan sentencias SQL en un lenguaje de alto nivel
- se necesita un precompilador que traduce las sentencias SQL, insertando llamadas a unas librerías específicas y generando el programa fuente modificado para que se pueda compilar, enlazar y ejecutar
- Oracle dispone de precompiladores para Cobol, C y C++

- **ODBC (*Open Database Connectivity*)**

- permite la integración de SQL en un lenguaje de propósito general
- se proporciona una biblioteca de funciones para que puedan ser invocadas desde el software de aplicación
- estas bibliotecas incluyen funciones como la conexión con la base de datos, la ejecución de sentencias SQL, la extracción de tuplas, etc.
- mediante el estándar ODBC se puede escribir código independiente del SGBD
- cada SGBD necesita su *driver* específico para trasladar las llamadas del ODBC en llamadas específicas del sistema gestor

❑ ¿Cómo interactuar con una BD desde un lenguaje de propósito general?

- **JDBC (conexión con bases de datos mediante Java)**

- con JDBC puede utilizarse Java como lenguaje anfitrión para escribir aplicaciones de BD
- está formado por una biblioteca de clases que permite la conexión con bases de datos que soporten SQL mediante Java
- la aplicación será independiente de la plataforma y se puede transportar de un SGBD a otro
- el código Java es independiente del SGBD utilizado

❑ Acceso a las Bases de Datos desde la Web

- **Lenguajes de *scripts*.** Permiten la creación de funciones incrustadas dentro del código HTML, lográndose automatizar procesos y acceder a diversos objetos (entre ellos a las bases de datos)
 - **JavaScript.** Lenguaje de *script* interpretado. Se puede ejecutar desde el cliente o desde el servidor (en este caso tienen la posibilidad de conectarse a una base de datos)
 - **VBScript.** Tiene la misma funcionalidad que JavaScript pero su sintaxis es mas parecida a Visual Basic
 - **PHP (*Hypertext Preprocessor*).** Lenguaje de *script* que se incrusta en el código HTML y que es soportado por la mayoría de los servidores *web*. Actualmente, la combinación más usada para diseñar páginas *web* dinámicas es el servidor HTTP de Apache, PHP y el SGBD MySql
 - **JSP (*JavaServer Pages*).** Lenguaje de *script* del lado del servidor, basado en Java, que permite mezclar HTML estático y dinámico
 - **ASP (*Active Server Page*).** Modelo de programación que permite crear páginas *web* interactivas y dinámicas en el servidor *web*.

❑ Acceso a las Bases de Datos desde la Web

- **CGI (*Common Gateway Interface*).**

- esta interfaz fue una de las primeras técnicas para crear contenido dinámico
- mediante un CGI un servidor *web* envía las peticiones de los clientes a un programa externo
- la salida del programa se envía al cliente como página estática
- los programas CGI pueden ser compilados (C, C++, Java, etc.) o *scripts* ejecutables (Perl, tcl, etc.)

1.8 Componentes de un SGBD

Un SGBD está estructurado en diversos componentes o módulos, cada uno de los cuales hace una función específica.

El S.O. proporciona servicios básicos

Módulos:

☐ **Procesador de consultas.**

- ☐ Se encarga de transformar las consultas en instrucciones de bajo nivel

☐ **Gestor de BD.**

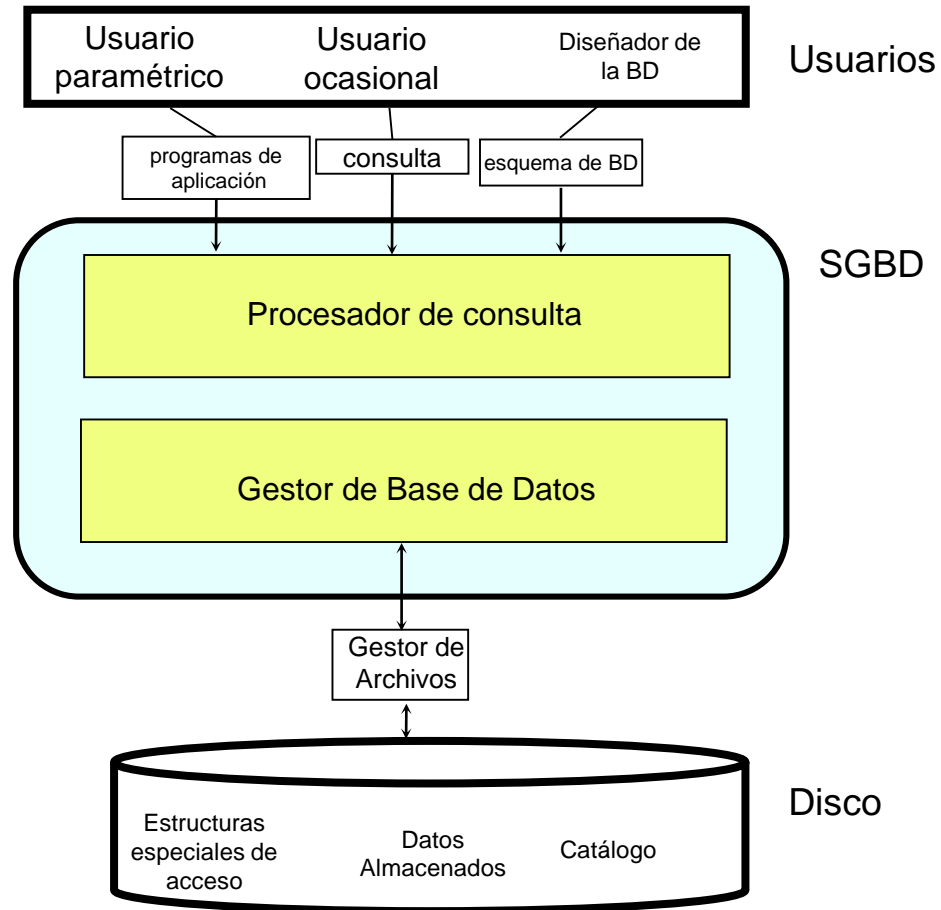
- ☐ Acepta consultas, examina los esquemas conceptual y externo para determinar qué registros conceptuales se necesitan y llama al gestor de archivos para satisfacer esa solicitud.

☐ **Gestor de Archivos.**

- ☐ Manipula los archivos almacenados y gestiona la asignación de espacio en disco.

Componentes de un SGBD

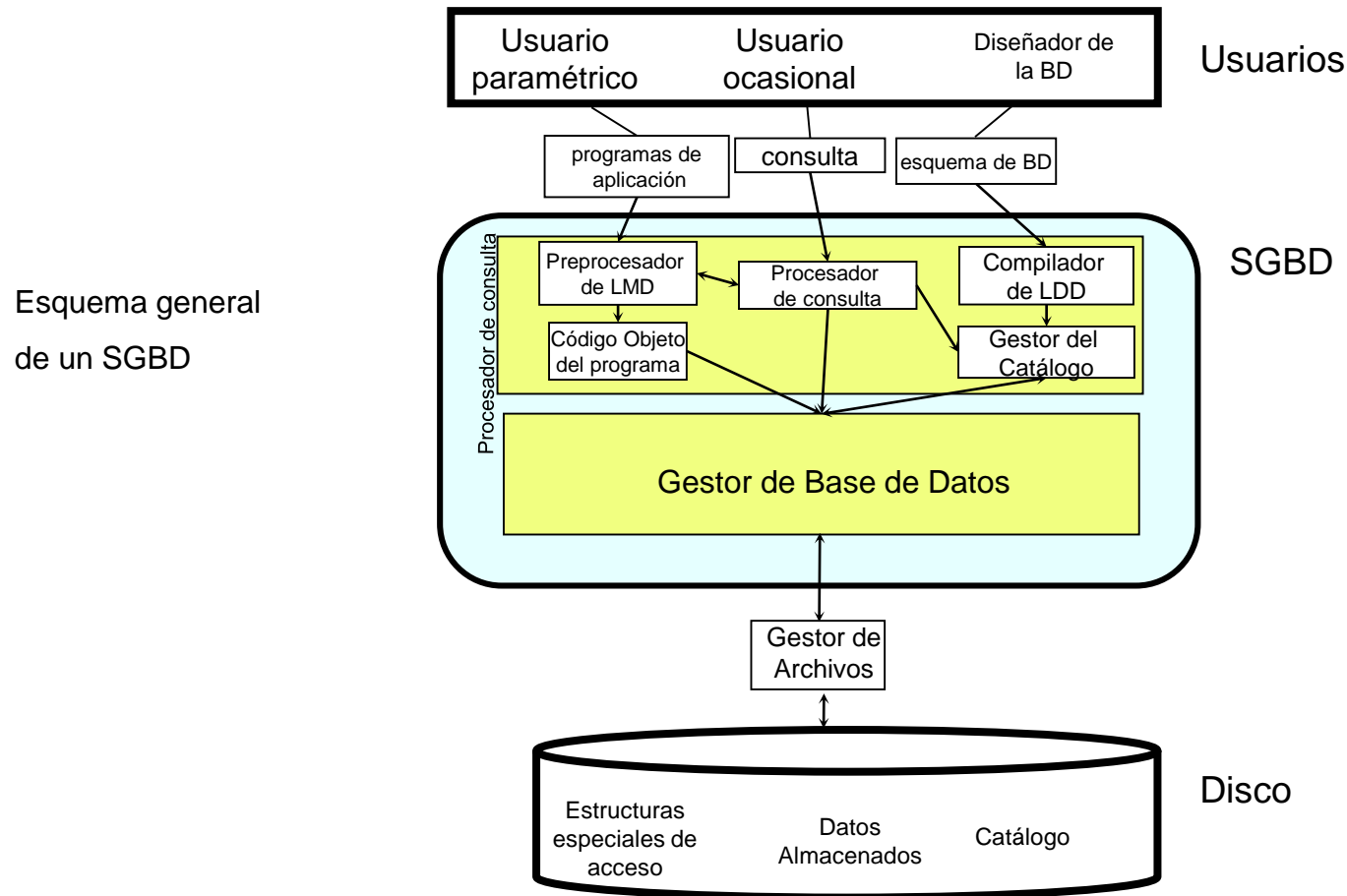
Esquema general
de un SGBD



❑ Procesador de consulta:

- **Preprocesador de LMD:** convierte las instrucciones de LMD de un programa de aplicación en llamadas a funciones en el lenguaje *host*. Debe interactuar con el procesador de consultas para generar el código apropiado.
- **Compilador LDD.** Convierte las instrucciones LDD en una serie de tablas que contienen *metadatos*. Estas tablas se almacenan en el catalogo.
- **Gestor del catalogo.** Gestiona y mantiene el catalogo del sistema.

Componentes de un SGBD

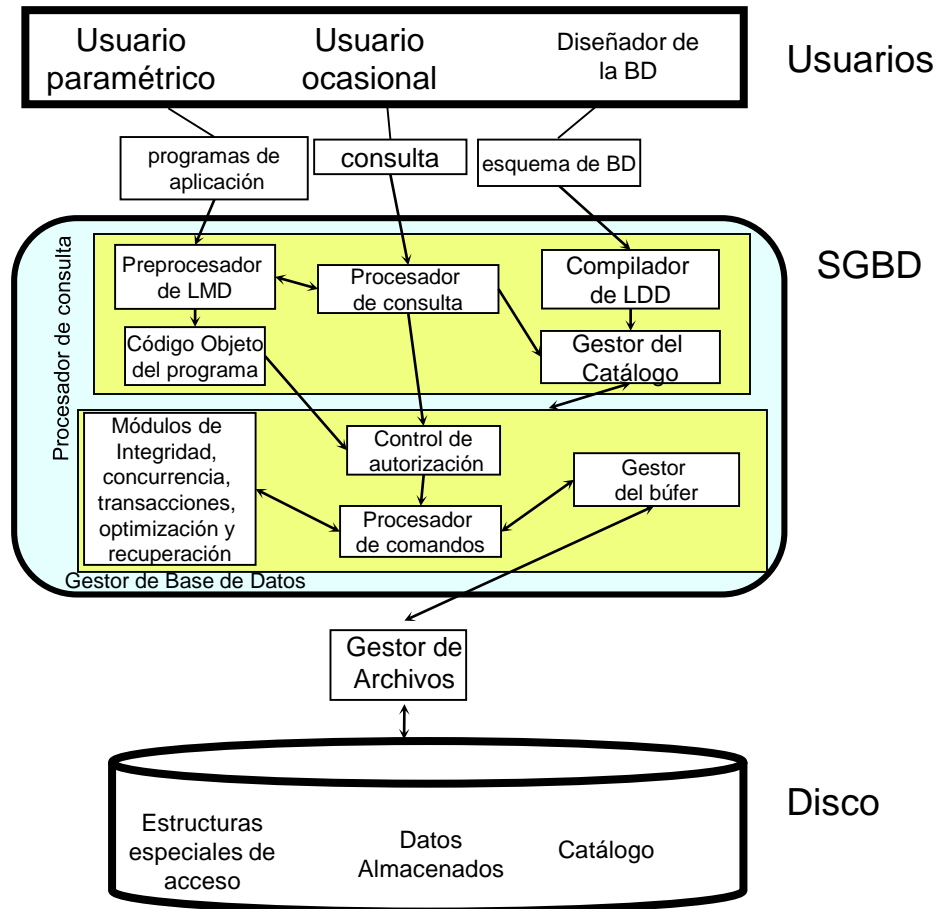


❑ **Gestor de la BD:**

- **Control de autorización.** Comprueba que el usuario tiene autorizaciones necesarias para realizar la operación requerida.
- **Procesador de comandos.** Una vez que el sistema ha comprobado los permisos del usuario, se pasa el control al procesador de comandos
- **Módulo de integridad.** Verifica que la operación satisface las restricciones de integridad.
- **Control de concurrencia.** Este módulo es el responsable de asegurar que las operaciones que se realizan concurrentemente sobre la base de datos tienen lugar sin conflictos.
- **Gestor de transacciones.** Realiza el procesamiento requerido para las operaciones que recibe de las transacciones.
- **Gestor de recuperación.** Garantiza que la BD permanezca en un estado coherente cuando se produzcan fallos.
- **Optimizador de consultas.** Determina la estrategia óptima para la ejecución de las consultas.
- **Gestor del *bufer*.** Responsable de la transferencia de datos entre la memoria principal y el almacenamiento secundario.

Gestión de almacenamiento

Esquema general
de un SGBD



1.9 Clasificación de los SGBD

Modelo de datos

Relacional: colección de tablas

Orientado a objetos: conjunto de objetos con sus propiedades y operaciones

Objeto-relacional: incorporan conceptos de las BD orientadas a objetos

Jerárquico: representa los datos como estructuras jerárquicas de árbol

Red: representa los datos como tipos de registro y los vínculos entre ellos mediante el tipo conjunto

Nº de usuarios

Monousuario: sólo un usuario a la vez (ordenadores personales)

Multiusuario: varios usuarios al mismo tiempo

Nº de sitios

Centralizado: datos almacenados en un solo sitio

Distribuido

Homogéneo: mismo software de SGBD en múltiples sitios

Heterogéneo: más de un SGBD distinto