

Introducción a los sistemas de bases de datos

Bases de datos II: Tema 1



Estructura del tema

- ¿Qué es un sistema SGBD?
- ¿Por qué utilizar un SGBD?
- Modelos de datos
- Niveles de abstracción en un SGBD
- Independencia de datos
- Consultas en un SGBD
- Control de concurrencia en un SGBD
- Atomicidad y sistema de recuperación
- Estructura de un SGBD
- Clasificación de los SGBDs



¿Qué es un SGBD?

- Una BD es un conjunto de datos que describen las actividades de una o varias organizaciones relacionadas.
 - Entidades (*p.e. estudiantes y cursos*)
 - Vínculos o relaciones (*p.e. María está en el curso de BDII*)
- Un sistema de Gestión de Base de Datos es un software específico que permite a los usuarios crear, mantener y manipular la BD (incorporando transparente e internamente los mecanismos necesarios para su funcionamiento correcto) .
- Un sistema de base de datos al conjunto formado por la BD más el SGBD.



¿Por qué utilizar un un SGBD?

- Acceso eficiente e independencia de los datos.
- Reducción del tiempo de desarrollo de aplicaciones.
- Integridad y seguridad de los datos.
- Administración de los datos de forma uniforme.
- Posibilidad de acceso concurrente
- Recuperación del sistema frente a fallos.

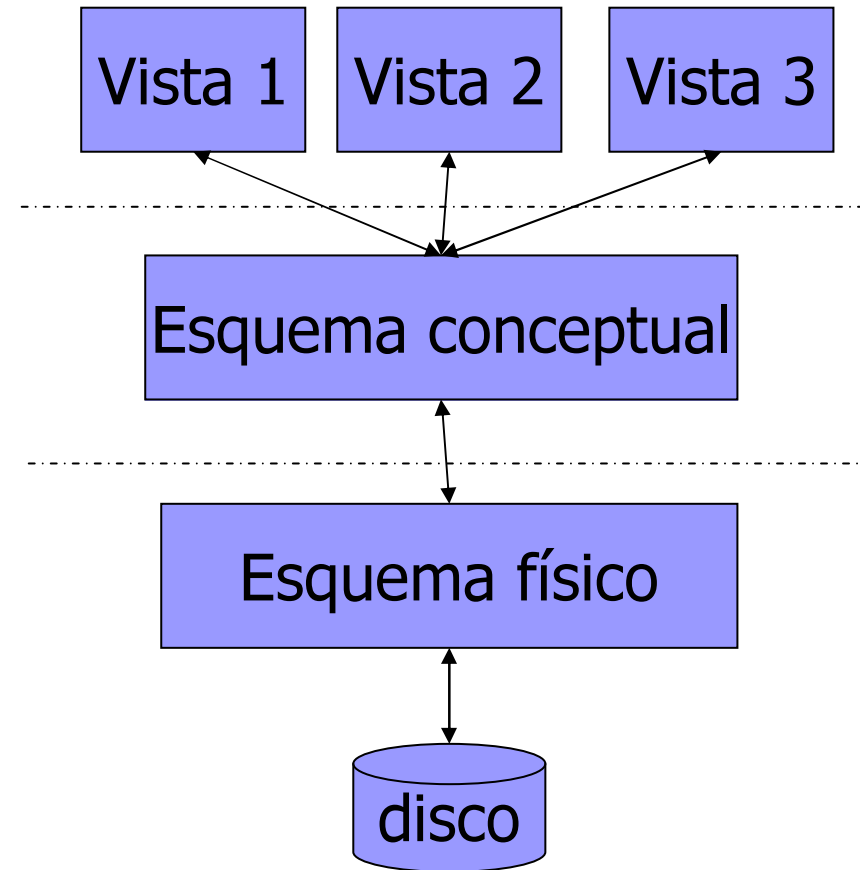


Modelos de datos

- Un **modelo de datos** es una colección de conceptos para la descripción de los datos.
- Un **esquema** es una descripción particular de datos utilizando un modelo dado.
- El **modelo relacional de datos** es el más utilizado en la actualidad:
 - Principal concepto: **relación** que es básicamente una tabla con filas y columnas (relación en el sentido de enumeración, no de correspondencia) .
 - Cada relación tiene un **esquema** que describe las columnas, o campos.

Niveles de abstracción en un SGBD

- Existen muchas vistas de un único esquema conceptual y físico.
 - Las vistas describen como los usuarios ven los datos.
 - El esquema conceptual define la estructura lógica.
 - El esquema físico describe los ficheros e índices utilizados.





Niveles de abstracción

- Esquema externo aplicación 1 (Vista 1):

```
CREATE VIEW pas (dni, nombre, sal_men)
AS SELECT dni, nombre, salario/14
FROM emp WHERE tipo='pas';
```

- Esquema conceptual:

```
emp(dni, nombre, dirección, salario, tipo)
```

- Esquema interno:

Fichero ordenado empleado con índice sobre el campo dni



Niveles de abstracción

- Aplicación 1: accede a la información a través del esquema externo 1

```
SELECT nombre, sal_men  
FROM PAS  
WHERE dni='algo';
```

- SGBD: Control del acceso y resolución de la operación pedida
- SO: Manipulación de los drivers de los dispositivos de memoria secundaria

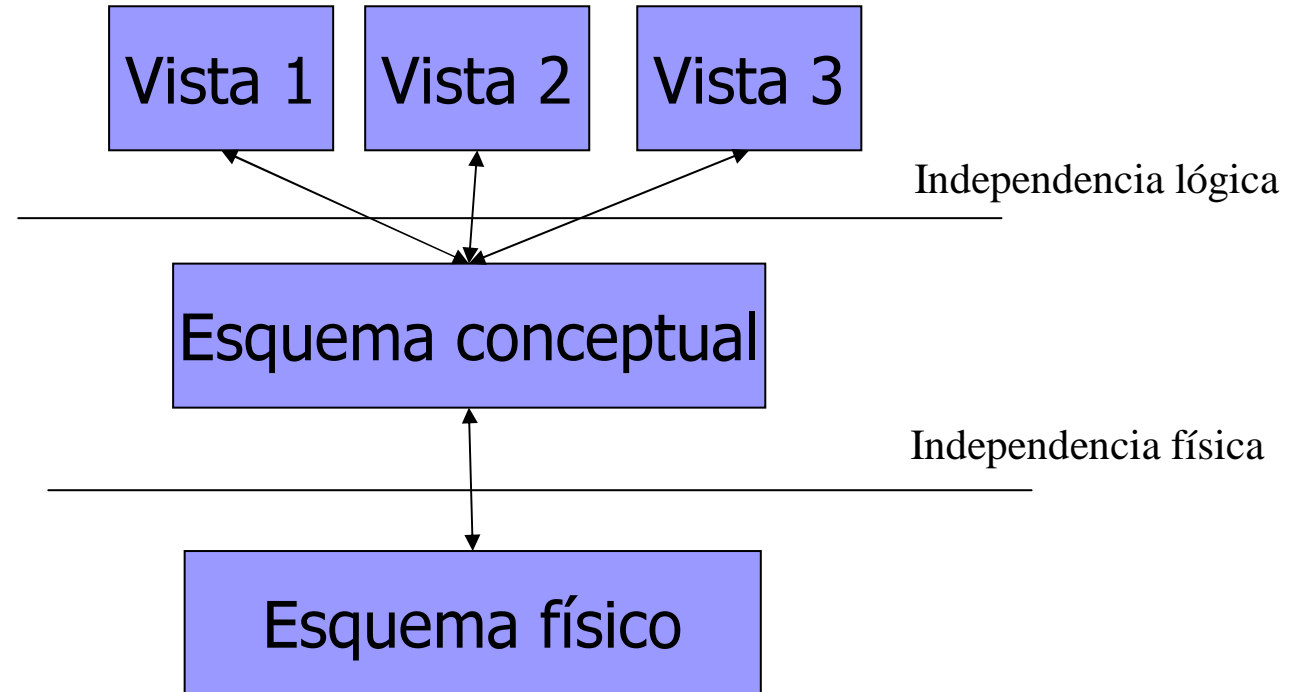


Independencia de los datos

(uno de los mayores beneficios de usar un SGBD)

- Los niveles de abstracción son útiles para explicar la independencia de datos: la capacidad para modificar el esquema en un nivel del sistema sin tener que modificar el esquema del nivel inmediato superior.
- Así, las aplicaciones están “aisladas” (no deben preocuparse) de cómo están los datos estructurados y almacenados.
- Independencia lógica de los datos:
 - Protección ante cambios en la estructura lógica de los datos.
- Independencia física de los datos:
 - Protección ante cambios en las estructuras físicas de los datos.

Independencia de los datos





Consultas en un SGBD

- Un SGBD proporciona un lenguaje especializado que permite obtener datos de la BD.
- El SGBD se preocupa de evaluar las consultas de la forma más eficiente que sea posible.
- El SGBD permite a los usuarios (¡a los usuarios autorizados!) crear, modificar y consultar los datos almacenados mediante un lenguaje de manipulación de datos.

Lenguajes DDL Y DML



Transacción

- Concepto clave en SGBD es la **transacción**: es una secuencia atómica de acciones en una BD iniciada por un usuario.
(atómica = no se puede dividir, o se ejecuta todo, o nada (se retrotrae todo)) .
- Cada transacción, ejecutada completamente, debe dejar la BD en un estado consistente (si la BD es consistente cuando la transacción comenzó). *(sign. consistencia)*
 - Los usuarios pueden especificar algunas restricciones de integridad sobre los datos y el SGBD controlará estas restricciones (y si no se cumplen, retrotraerá la transacción) .
 - El significado de las restricciones es sólo conocido por el usuario, por lo que el asegurar que una transacción preserve la consistencia de la BD es responsabilidad de los usuarios.



Control de concurrencia en un SGBD

- La ejecución concurrente de varios programas de usuario es necesaria para tener un buen rendimiento del SGBD.
- La concurrencia (intercalación) de acciones de diferentes programas de usuario podría conducir a inconsistencias en la BD.
- El SGBD asegura que estos problemas no surjan: los usuarios trabajarán con el sistema como si fuera un sistema monousuario.



Ejecución de transacciones concurrentes

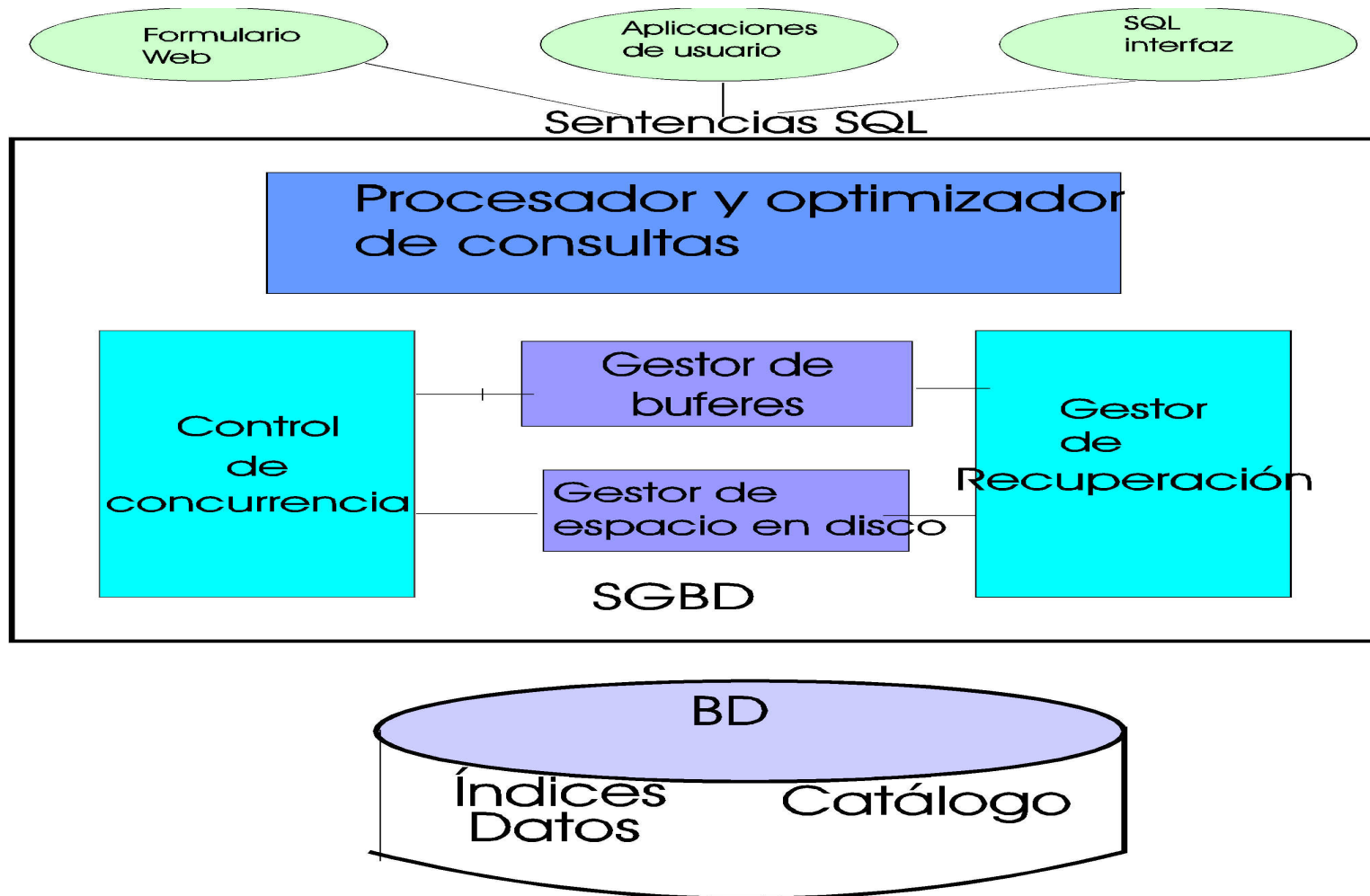
- Un SGBD asegura que la ejecución de transacciones de forma concurrente sea igual a alguna ejecución secuencial de esas transacciones (y evita que se ejecuten en un orden inadecuado o ineficiente) .
- Existen varios algoritmos que permiten lo anterior:
 - Bloqueos,
 - Marcas de tiempo,
 - Algoritmos optimistas...



Atomicidad

- El SGBD asegura la atomicidad de las transacciones (propiedad de todo o nada) incluso si el sistema se cae en mitad de una transacción.
- Idea: guardar una traza (*log*) de todas las acciones realizadas por el SGBD cuando se ejecuten transacciones sobre la BD.

Estructura de un SGBD





Partes de un SGBD (1)

- El SGBD:
 - Acepta consultas generadas por programas diversos.
 - Produce planes de ejecución de consulta.
 - Ejecuta consultas
 - Devuelva la respuesta.

- El optimizador busca la mejor forma de ejecutar la consulta.



Partes de un SGBD (2)

- El gestor de búferes trae páginas de disco a memoria principal.
- El módulo de control de concurrencia evita que se produzcan inconsistencias en la BD.
- El módulo de recuperación mantiene la información necesaria para que el sistema se recupere frente a fallos.



Clasificación de los SGBD (1)

- Criterios utilizados para la clasificación:
 - Modelo de datos utilizado.
 - Número de usuarios.
 - Número de lugares en que está distribuida la BD.



Clasificación de los SGBD (2)

■ Modelo de datos:

- Relacionales
- Objeto-relacionales.
- Orientados a objetos.
- Jerárquicos,
- Red,
- Otros...

■ Número de usuarios:

- Monousuario
- Multiusuario

■ Número de lugares:

- Centralizados
- Distribuidos.



Bibliografía

- R. Ramaskrishnan, J. Gehrke. Database Management Systems. McGraw Hill. Capítulo 1.
- Elmasri y Navathe. Fundamentos de Sistemas de Bases de datos. Addison Wesley. Capítulo 1.
- T. Connolly, C. Begg. Database Systems. Addison Wesley. Capítulo 1.



Tema 1: Cuestionario - resumen

1. ¿Por qué elegirías un sistema de BD en lugar de almacenar la información en ficheros del sistema operativo?
2. ¿Qué es la independencia lógica y por qué es importante?
3. Explica la diferencia entre independencia lógica e independencia física.
4. Explica la diferencias entre esquema externo, conceptual y físico.
5. Asumiendo que el administrador de la BD no va a ejecutar consultas, ¿debe entender la optimización de consultas? ¿Por qué?
6. ¿Qué es una transacción?
7. ¿Por qué un SGBD mezcla acciones de varias transacciones en lugar de ejecutar una tras otra?
8. ¿Qué debe un usuario garantizar con respecto a una transacción y la consistencia de la BD?