

BASES DE DATOS

Conceptos

Un poco de historia ...

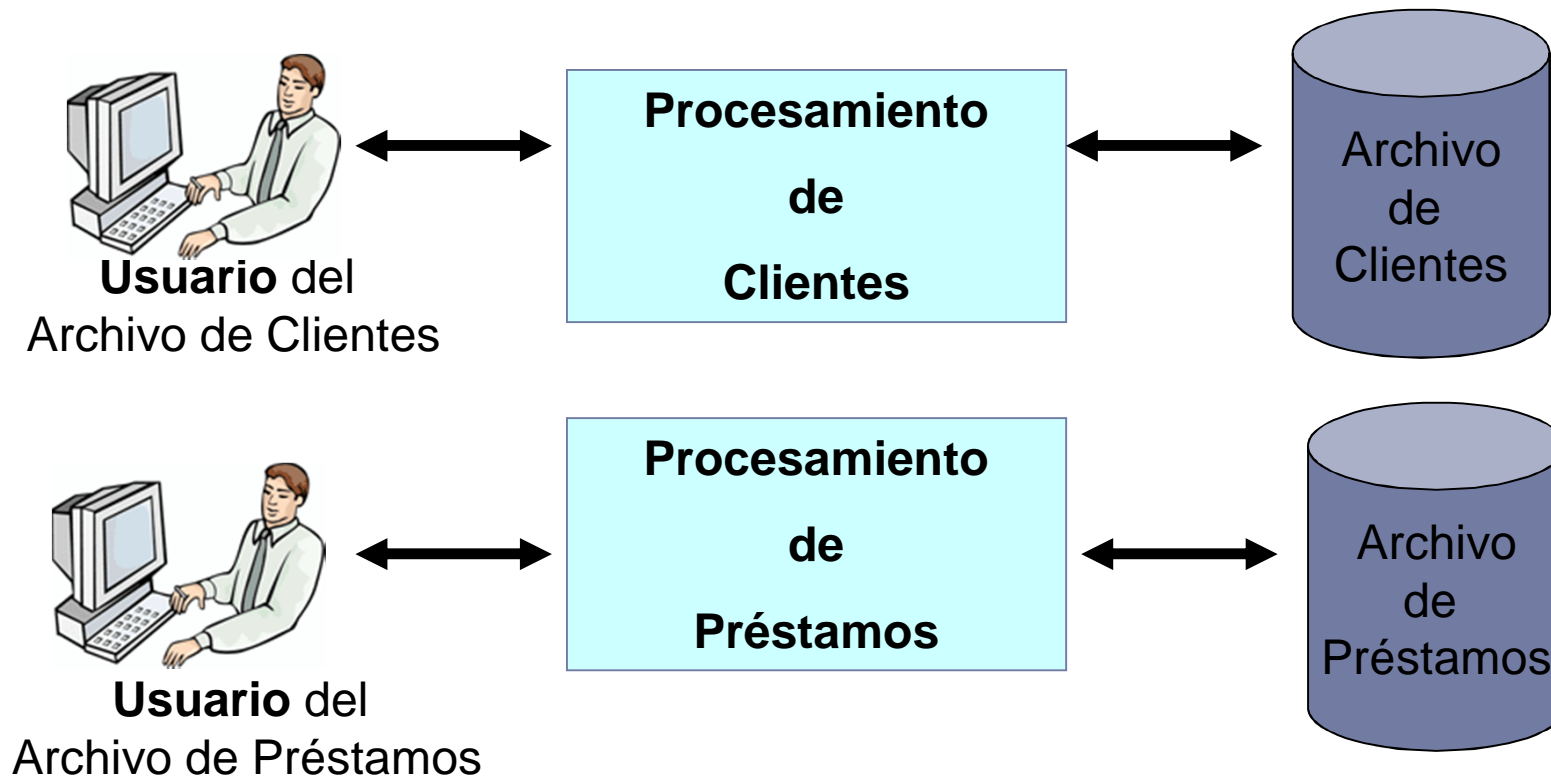
Sistemas de archivos:

- ❑ Surgieron a raíz de la necesidad de almacenamiento de la información para su correspondiente reutilización (**persistencia**)

Desventajas:

- ❑ Redundancia de los datos
- ❑ Los archivos están separados unos de otros (no se pueden “combinar” fácilmente)
- ❑ Alto costo para la propagación de cambios
- ❑ Inconsistencia debido a actualizaciones simultáneas

Un poco de historia ...



Por cada préstamo se guarda la información del cliente correspondiente: Redundancia

Un poco de historia ...

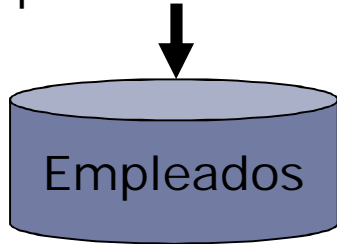
¿Por qué surgieron los sistemas de Bases de Datos?

- ▶ Necesidad de solucionar las debilidades de los sistemas de archivos
- ▶ Capacidades:
 - ▶ Manejo de persistencia
 - ▶ Soporte por lo menos de un modelo de datos
 - ▶ Soporte de un lenguaje de alto nivel que permita manipular y definir la estructura de la información
 - ▶ Control de acceso
 - ▶ Evitar inconsistencias al **compartir** la información

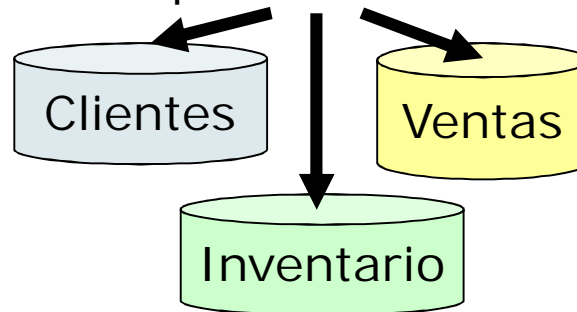
Antes



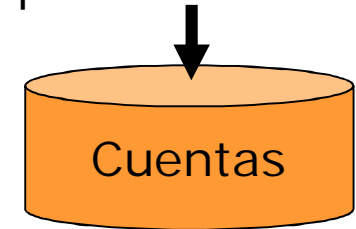
Dpto. Personal



Dpto. Ventas



Dpto. Contabilidad



... Ahora



Definición



- Una **base de datos** es un **conjunto** estructurado de datos coherentes

**Colección disponible de
información**

Definición



- Una **base de datos** es un conjunto **estructurado** de datos coherentes

Colección organizada en subconjuntos, en función de ligas y de relaciones entre las diferentes informaciones (estructura lógica)

Definición

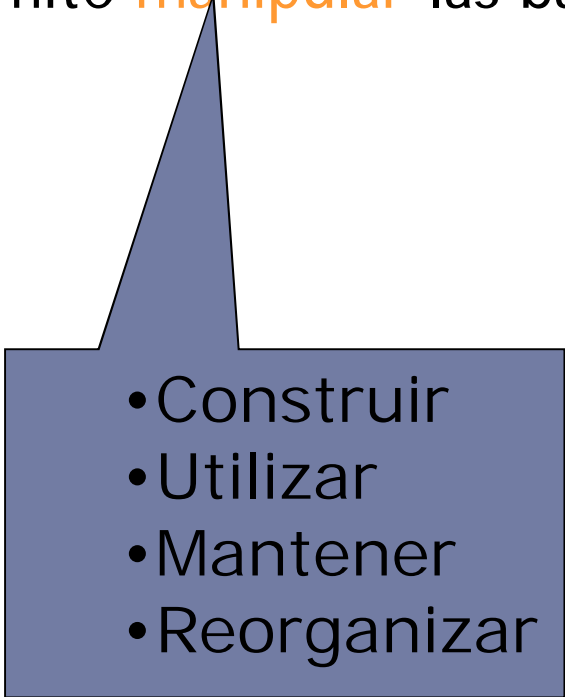


- Una **base de datos** es un conjunto estructurado de datos **coherentes**

No hay contradicción entre los datos ligados, no hay pérdida de información, aun sabiendo que hay una utilización compartida de los datos entre varios usuarios

Definición

- Un **Sistema de Gestión de Bases de Datos (SGBD)** es un software que permite **manipular** las bases de datos

- 
- Construir
 - Utilizar
 - Mantener
 - Reorganizar

Sistema de Gestión de Bases de Datos

- ▶ Propiedades fundamentales
 - Independencia de los datos (OJO es muy importante)
 - Acceso eficiente a los datos
 - Integridad y seguridad de los datos
 - Administración de los datos
 - Acceso concurrente y recuperación en caso de « crash »
- ▶ Los SGBD deben cumplir con las propiedades ACID para las transacciones:
 - A**tomicity (Atomicidad: las transacciones son atómicas)
 - C**onsistency (Consistencia: una transacción transforma un estado consistente de la BD en otro)
 - I**solation (Aislamiento: las transacciones están aisladas entre si)
 - D**urability (Durabilidad: después que una transacción ha sido confirmada ella persiste)

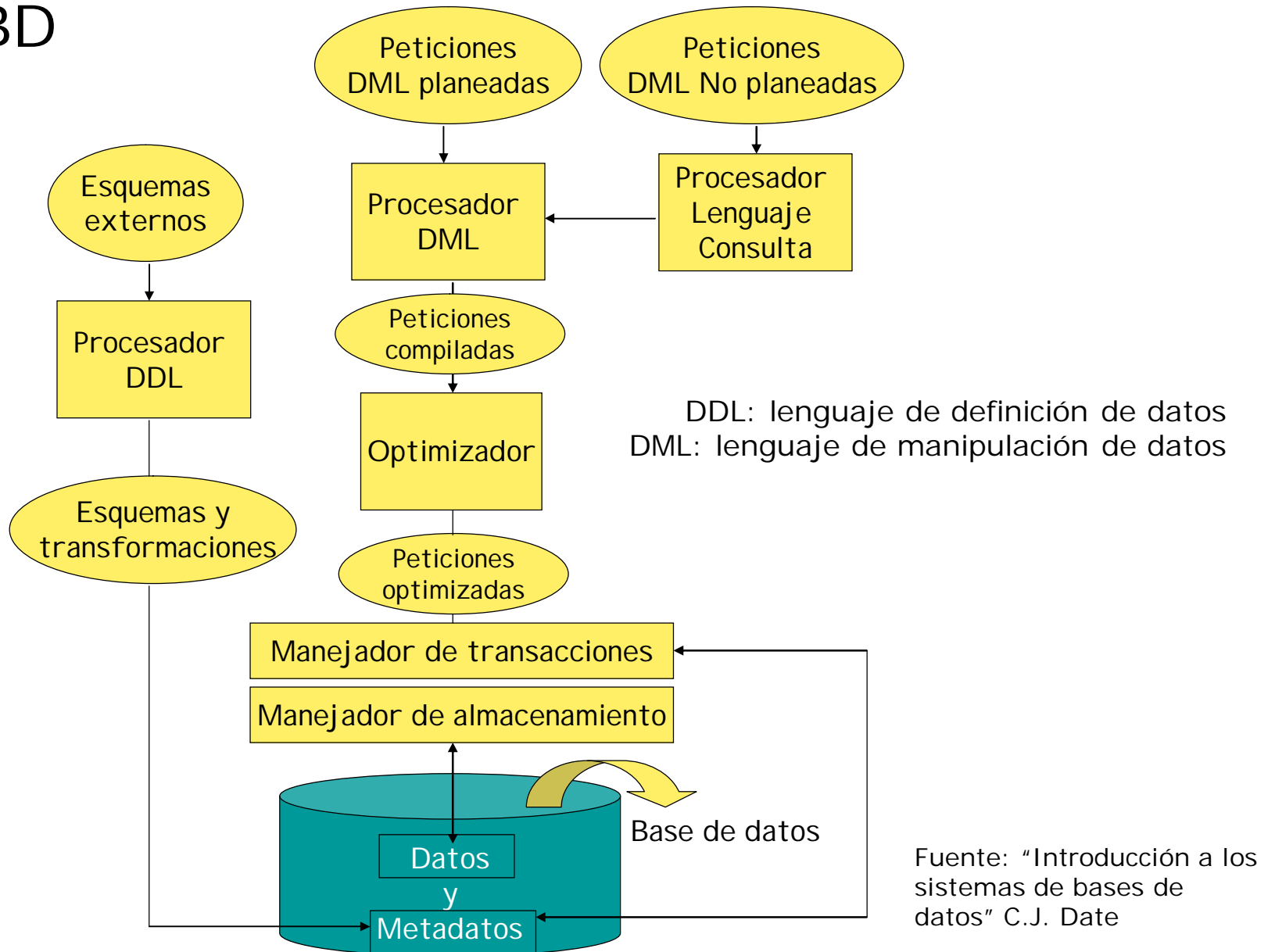
Algunas definiciones

- ▶ **Esquema de una Base de Datos:** Cuando una base de Datos es diseñada interesa definir una estructura para ésta.
Esta estructura permanece “estática” durante un gran período de tiempo, aunque puede sufrir modificaciones ocasionales
- ▶ **Instancia:** Es la información que en un determinado instante del tiempo posee la base de datos y que cambia permanentemente (excepto en algunos tipos de bases de BD particulares)

Algunas definiciones

- ▶ **Usuario final:** El que interactúa con la base de datos, por lo general a través de aplicaciones e interfaces
- ▶ **Usuario especialista:** El que diseña y programa aplicaciones para usuarios finales.
- ▶ **DBA (Database Administrator):** El que administra la base de datos. ¿Qué hace?

Funciones y componentes principales de un SGBD



Componentes funcionales de un SGBD

- ▶ **Metadatos o Diccionario de Datos (D.D):** Contiene el esquema de la B.D, los usuarios, los permisos de acceso, etc. Son datos sobre los datos. Almacena la información que permite la traducción entre los 3 niveles de la Arquitectura ANSI/SPARC*
- ▶ **Optimizador de consultas:** Define el plan de ejecución de operaciones solicitadas por los usuarios, de tal manera que se lleven a cabo de la manera más eficiente posible
- ▶ **Manejador de transacciones:** Controla el acceso y la concurrencia de operaciones

*Se verá mas adelante en el curso

Componentes funcionales de un SGBD

- ▶ **Manejador de almacenamiento** tiene dos componentes:
 - Manejador de archivos** → recupera desde disco los bloques que contienen la información solicitada por una transacción.
 - Manejador de buffer** → mantiene en memoria principal la información más usada y decide cuando llevar a disco alguno de sus bloques.

SGBD - Funciones

- ▶ Soporta DML: Lenguaje para actualización, almacenamiento y recuperación de datos
- ▶ Ofrece optimización en la búsqueda de la información
- ▶ Soporta DDL : Lenguaje para definir los datos
- ▶ Metadatos (DD): Catálogo autodescriptivo, información sobre los objetos existentes en el sistema → *Datos sobre los datos*

Ventajas de un SGBD

1. Reutilización de datos y programas
2. Control de redundancia
3. Estandarización ¿En qué sentido?
4. Consistencia (No hay redundancia)
5. Es posible equilibrar las cargas de los requerimientos (establecer prioridades)

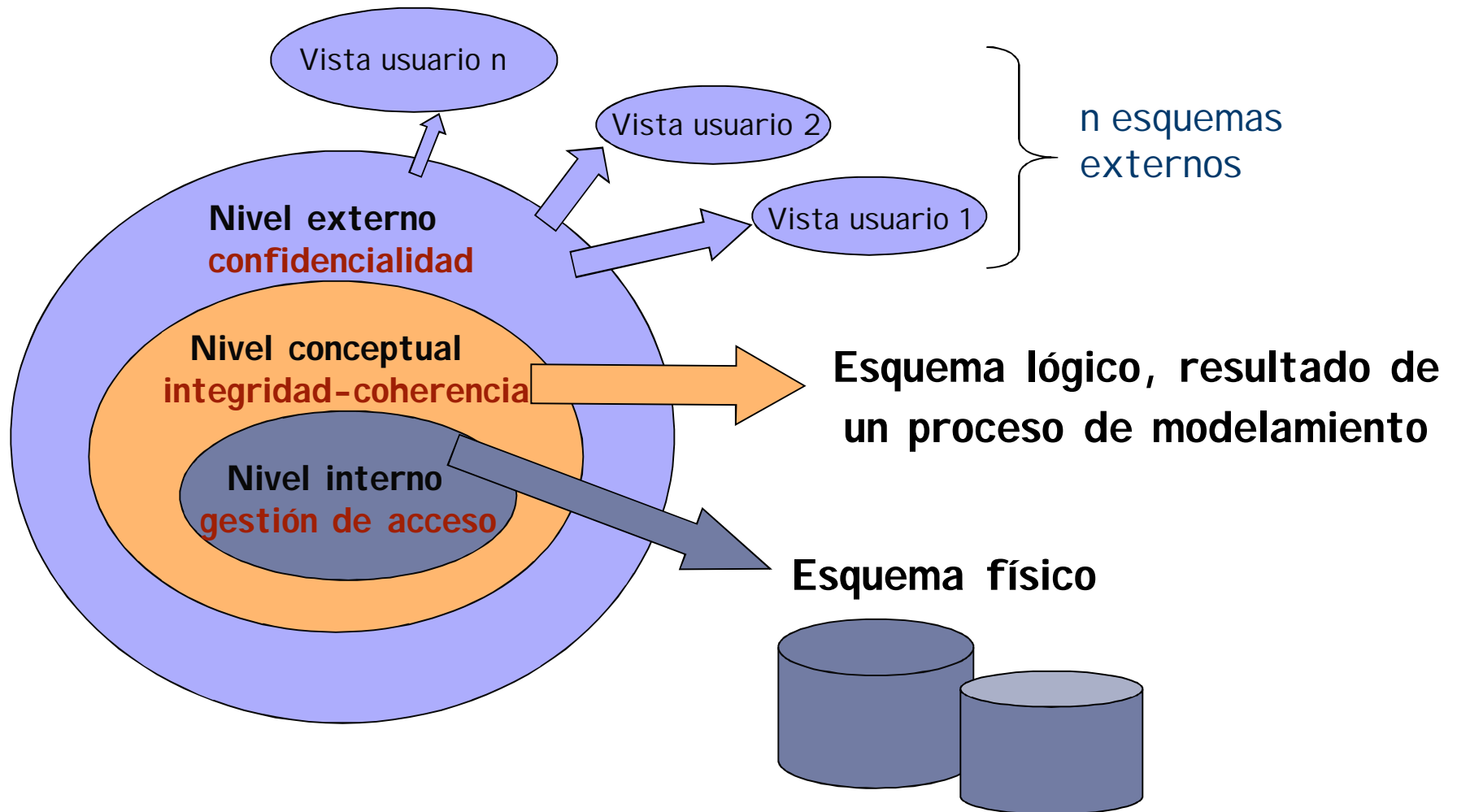
Ventajas de un SGBD

- 6. Integridad (Se cumplen las reglas establecidas → ¿Por quién?)
- 7. Seguridad
- 8. Rapidez de desarrollo
- 9. Mantenimiento y reingeniería: cambios en la estructura de datos sin cambiar los programas que los usan (hasta cierto punto)

Desventajas de un SGBD

1. Tamaño
2. Susceptibilidad a fallas (*Discutible*)
3. Complejidad en la recuperación a fallas (*Discutible*)
4. Puede llegar a trabajar en forma “lenta” debido a la cantidad de **verificaciones** que debe hacer

Niveles de abstracción (ANSI/SPARC)



Nivel de Visión

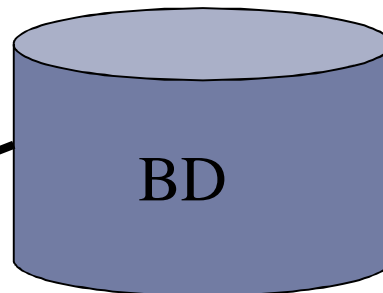
- ▶ Externo
- ▶ El más cercano a los usuarios finales
- ▶ Percepción de la base de datos por parte de los usuarios finales
- ▶ *Tantas visiones como tipos de usuarios*
- ▶ Cada visión de usuario final se puede caracterizar como un *subesquema*

Nivel de Visión

Cada visión puede proporcionar diferentes representaciones de los mismos datos

Visión Vendedor

Fechas con formato:
(dd-mm-yy)



Visión Contador

Fechas con formato:
(yyyy-dd-mm)



Nivel de Visión

Algunas *visiones* de usuario pueden incluir:

- ▶ Datos Agrupados: ***Totales por Dpto.***

- ▶ Datos Derivados:

Sueldo total = básico + comisión

- ▶ Datos Calculados:

Edad de una persona (inferida a partir de su fecha de Nacimiento)

Nivel de Visión

- ▶ Los conceptos de dato derivado y calculado podrían usarse indistintamente para significar, en general, que un dato se genera a partir de otro(s)
- ▶ Pueden haber a su vez datos agregados derivados, datos agregados calculados etc.

Nivel Conceptual

- ▶ Mediador entre los otros 2 niveles
- ▶ Interesante para el usuario especialista
- ▶ Se ocupa de los datos almacenados en la base de datos física y las *relaciones* entre ellos
- ▶ Descripción semántica de los datos que conforman la base de datos
- ▶ Soporta a cada visión de usuario externa

Nivel Conceptual

- ▶ Es una visión completa de todos los *requerimientos* y elementos de interés para la organización
- ▶ Debe incluir *restricciones* sobre los datos
- ▶ La descripción del nivel conceptual no debe tener detalles dependientes del almacenamiento*
- ▶ Tiene asociado un lenguaje de alto nivel

* Sin embargo en algunos SGBD esto no se logra expresar de esta manera...

Nivel Físico

- ▶ Interno
- ▶ Más cercano a la máquina
- ▶ Interesa al Administrador y al usuario Especialista
- ▶ Esquema físico: Descripción y *tipos de datos*: tamaño y precisión, tipos de índices y de estructuras de almacenamiento concretas que se manejan, de acuerdo con un *SGBD particular*

Nivel Físico

- ▶ Describe **cómo** los datos son almacenados en términos de estructuras de datos particulares
- ▶ Se encarga de:
 - Reservar espacio para datos e índices
 - Compresión de datos
 - Técnicas de encriptamiento de datos

Independencia de los datos

- ▶ Es uno de los objetivos de la arquitectura ANSI/SPARC
- ▶ Permite modificar la definición de un nivel sin afectar (en lo posible) el nivel inmediatamente superior
- ▶ Sin independencia de datos se requeriría mucho esfuerzo para cambiar las aplicaciones de tal forma que se adaptasen a la nueva estructura de la base de datos.
- ▶ Hay dos tipos: física y lógica

Independencia de los datos

- ▶ Física: inmunidad que tienen los usuarios y las aplicaciones ante los cambios en la forma de almacenar físicamente los datos.
- ▶ Conceptual o lógica: inmunidad que poseen los usuarios y las aplicaciones ante los cambios en la estructura lógica de la base de datos

Independencia Física

- ▶ Se presenta entre el nivel conceptual y el nivel físico
- ▶ Un cambio en el esquema físico (usar otras estructuras de almacenamiento) no conduce a cambios en el esquema conceptual

Ej. Cambio en la forma de almacenar un índice: *B+* por *Hashing*, ¿con qué propósito se hace un cambio de este tipo?

- ▶ **Inmunidad** del esquema conceptual ante cambios del esquema físico

Independencia Lógica

- ▶ Se presenta entre el nivel de visión y el nivel conceptual
- ▶ Significa que un cambio en el nivel conceptual no debe conllevar a un cambio en el nivel de visión
- ▶ Es más difícil de lograr. ¿Por qué?

Independencia Lógica

Algunos de los posibles cambios en el nivel conceptual:

- ▶ Adición de nuevos elementos (atributos, entidades etc.)
- ▶ Eliminación de elementos → *Puede afectar a los subesquemas externos*

Tendencias de los SGBD

- ⇒ Sistemas para el soporte para toma de decisiones (Data Warehouse) y/o con capacidades deductivas y bases de datos temporales
- ⇒ Manejar información multimedial: Imágenes, audio, videos
- ⇒ Bases de datos orientadas a objetos y objeto-relacionales
- ⇒ Manejar información georeferenciada: Sistemas de Información Geográficos (SIG)
- ⇒ Manejo de información documental: Motores de búsqueda, sistemas para el manejo de documentos.
- ⇒ Bases de datos nativas para XML

RESUMEN

- ▶ La descripción completa de una base de datos se denomina *esquema*
- ▶ Cada visión de usuario tiene su esquema (subesquema), existe **un** esquema conceptual y **uno** interno
- ▶ Existen correspondencias (“mappings”)
 - Entre cada subesquema externo y el conceptual
 - Entre el esquema conceptual y el interno
- ▶ Estas correspondencias le permiten al SGBD por ejemplo saber un registro conceptual con qué registro(s) del nivel físico se corresponde

RESUMEN

- ▶ En *teoría* debería existir un lenguaje de descripción para cada nivel*
- ▶ En la práctica sin embargo en muchos SGBD el esquema conceptual queda fusionado con la especificación del esquema interno.

* No siempre sucede así en los SGBD actuales...

Próximas clases

► Modelos:

- Modelo conceptual: el diseño de un modelo conceptual parte de la especificación de requisitos
 - Ingeniería de requisitos
 - Modelos conceptuales: E/A, semántico
 - Modelos Lógicos: Relacional
 - Modelos físicos: dependientes del SGBD