

# Introducción a las Bases de Datos

Profesor:

*Máster Carlos Carranza Blanco*

*Diseño e Implementación de Bases de Datos*

*Ingeniería del Software*

8/3/2022

Basado en el libro Fundamentos de Bases de Datos 6ª Ed.

Silberschatz - Korth - Sudarshan

# Agenda



- Presentación del curso, profesor y estudiantes
- Emprendedurismo
- Video motivacional
- Entrega y aprobación programa curso
- Presentación Introducción a las Bases de Datos.

# Introducción a las Bases de Datos



- La gestión de las bases de datos ha evolucionado desde una aplicación informática especializada hasta convertirse en parte esencial de los entornos informáticos modernos. Por tanto, el conocimiento acerca de los sistemas de bases de datos se ha convertido en una parte imprescindible de la formación en Informática.
- Es importante conocer los conceptos fundamentales de la gestión de las bases de datos. Estos conceptos incluyen aspectos del diseño, de los lenguajes y de la implementación de los sistemas de bases de datos.

# Sistemas Gestores de Bases de Datos (SGBD)



- Consiste en una colección de datos interrelacionados y un conjunto de programas para acceder a dichos datos. La colección de datos, normalmente denominada base de datos, contiene información relevante para una empresa.
- El objetivo principal de un SGBD es proporcionar una forma de almacenar y recuperar la información de una base de datos de manera que sea tanto práctica como eficiente.

**Grandes Cantidades de Información** – Definición de Estructuras para almacenarla y manipularla – **Garantizar la fiabilidad** – **Recuperación** – **Seguridad** – **Concurrencia**

# Propósito de los SABD



- Surgieron en respuesta a los primeros métodos de gestión informatizada de los datos comerciales.
- Los sistemas operativos convencionales soportan este sistema de procesamiento de archivos típico. El sistema almacena los registros permanentes en varios archivos y necesita diferentes programas de aplicación para extraer y añadir datos a los archivos correspondientes. Antes de la aparición de los sistemas gestores de bases de datos (SGBD), las organizaciones normalmente almacenaban la información en sistemas de este tipo.

# Propósito de los SABD...



- Inconvenientes de los sistemas de procesamiento de archivos:
  - **Redundancia e inconsistencia de los datos:** Genera mayores necesidades de almacenamiento a un mayor coste. Además, puede que las distintas copias de los datos no coincidan.
  - **Dificultad en el acceso a los datos:** Se depende mucho de lo que el programa tenga desarrollado en cuanto a generación de reportes, por lo que no permiten recuperar datos de forma práctica y eficiente.
  - **Aislamiento de datos:** Datos dispersos en varios archivos, diferentes formatos...
  - **Problemas de integridad:** La responsabilidad de las restricciones de consistencia queda solo del lado del sistema y de los desarrolladores de este.
  - Problemas de atomicidad: Sistemas sujetos a fallos, ante estos se requiere restauración a estados consistentes



# Propósito de los SABD...



- Inconvenientes de los sistemas de procesamiento de archivos:...
  - **Problemas de atomicidad:** Sistemas sujetos a fallos, ante estos se requiere restauración a estados consistentes, por ello la necesidad de transacciones atómicas, es decir debe ocurrir en su totalidad o no ocurrir en absoluto. Resulta difícil asegurar la atomicidad en los sistemas convencionales de procesamiento de archivos.
  - **Anomalías en el acceso concurrente:** Sistema multiusuarios imprescindibles, millones de accesos simultáneos, producen actualizaciones recurrentes de información generando datos inconsistentes.
  - **Problemas de seguridad:** Controlar el acceso a ciertos datos según los privilegios de cada usuario.

# Niveles de presentación

- Una de las principales finalidades de los sistemas de bases de datos es ofrecer a los usuarios una visión abstracta de los datos. Es decir, el sistema oculta ciertos detalles del modo en que se almacenan y mantienen los datos.
- **Nivel físico:** El nivel más bajo de abstracción describe cómo se almacenan realmente los datos.
- **Nivel lógico:** describe qué datos se almacenan en la base de datos y qué relaciones existen entre esos datos. Describe toda la base de datos en estructuras relativamente simples. La implementación de estructuras simples en el nivel lógico puede involucrar estructuras complejas del nivel físico, los usuarios del nivel lógico no necesitan preocuparse de esta complejidad. **A esto se denomina independencia de los datos físicos.**



# Niveles de presentación...

- **Nivel de vistas:** El nivel más elevado de abstracción solo describe parte de la base de datos. El nivel de abstracción de vistas existe para simplificar su interacción con el sistema. El sistema puede proporcionar muchas vistas para la misma base de datos. En el nivel de vistas se definen varias vistas de la base de datos, y los usuarios de estas pueden verlas todas o solamente una parte. Además de ocultar los detalles del nivel lógico de la base de datos, las vistas también proporcionan un mecanismo de seguridad para evitar que los usuarios tengan acceso a ciertas partes.



Figura 1.1. Los tres niveles de abstracción de datos.

# Esquemas de Bases de Datos



- La colección de información almacenada en la base de datos en un momento dado se denomina ejemplar de la base de datos. El diseño general de la base de datos se denomina esquema de la base de datos.
- Los sistemas de bases de datos tienen varios esquemas, divididos según los niveles de abstracción. El **esquema físico** describe el diseño de la base de datos en el nivel físico, mientras que el **esquema lógico** describe su diseño en el nivel lógico. Las bases de datos también pueden tener varios **esquemas** en el nivel **de vistas**, a veces denominados subesquemas, que describen diferentes vistas de la base de datos.

# Esquemas de Bases de Datos...



- De estos, el esquema lógico es con mucho el más importante en términos de su efecto sobre los programas de aplicación, ya que los programadores crean las aplicaciones usando el esquema lógico.
- El esquema físico está oculto bajo el esquema lógico, y generalmente puede modificarse fácilmente sin afectar a los programas de aplicación.
- Se dice que los programas de aplicación muestran independencia física respecto de los datos si no dependen del esquema físico y, por tanto, no hace falta volver a escribirlos si se modifica el esquema físico.

# Modelos de Datos



- Bajo la estructura de las bases de datos se encuentra el modelo de datos: se trata de una colección de herramientas conceptuales para describir los datos, sus relaciones, su semántica y las restricciones de consistencia. Los modelos de datos ofrecen un modo de describir el diseño de las bases de datos en los niveles físico, lógico y de vistas.
- Los modelos de datos pueden clasificarse en cuatro categorías diferentes:
- **Modelo relacional:** utiliza una colección de tablas para representar tanto los datos como sus relaciones. Cada tabla tiene varias columnas, y cada columna tiene un nombre único. **El modelo de datos relacional es el modelo de datos más ampliamente utilizado.**

# Modelos de Datos...



- **Modelo entidad-relación (E-R):** consiste en una colección de objetos básicos, denominados entidades, y de las relaciones entre ellos. Una entidad es una «cosa» u «objeto» del mundo real que es distinguible de otros objetos. **El modelo entidad-relación se usa mucho en el diseño de bases de datos.**
- **Modelo de datos basado en objetos:** El modelo orientado a objetos se puede considerar como una extensión del modelo E-R con conceptos de encapsulación, métodos (funciones) e identidad de objetos.
- **Modelo de datos semiestructurados:** Permite la especificación de datos en el que los elementos de datos individuales del mismo tipo pueden tener diferentes conjuntos de atributos. Esto lo diferencia de los modelos mencionados anteriormente, en los que cada elemento de datos de un tipo particular debe tener el mismo conjunto de atributos. El lenguaje de marcas extensible (XML) se emplea mucho para representar datos semiestructurados.

# Diseño de una Base de Datos



- Los sistemas de bases de datos se diseñan para gestionar grandes cantidades de información. Esas grandes cantidades de información no existen aisladas. Forman parte del funcionamiento de alguna empresa.
- El diseño de bases de datos implica principalmente el diseño de su esquema.
- Los modelos de datos de alto nivel resultan útiles a los diseñadores de bases de datos al ofrecerles un marco conceptual en el que especificar, de manera sistemática, los requisitos de datos de los usuarios de las bases de datos y la manera en que se estructurará la base de datos para satisfacer esos requisitos.



# Diseño de una Base de Datos...



- **Proceso de diseño:**
  - Caracterizar completamente los requisitos de datos de los hipotéticos usuarios de la base de datos. El resultado de esta fase es la especificación de los requisitos de los usuarios.
  - Escoger un modelo de datos y traduce los requisitos en un esquema conceptual de la base de datos. El esquema desarrollado en esta fase de diseño conceptual ofrece una visión general detallada de la empresa. En este punto, la atención se centra en describir los datos y sus relaciones, más que en especificar los detalles del almacenamiento físico.
  - En términos del modelo relacional, el proceso de diseño conceptual implica decisiones sobre **qué** atributos se desea capturar en la base de datos y **cómo** agruparlos para formar las diferentes tablas.

# Diseño de una Base de Datos...



- Proceso de diseño:...

- Hay dos vías principales para afrontar el problema del **cómo**. La primera supone usar el modelo entidad-relación; la otra es emplear un conjunto de algoritmos (denominados colectivamente como normalización) que toma como entrada el conjunto de todos los atributos y genera un conjunto de tablas.
- Un esquema conceptual completamente desarrollado también indica los requisitos funcionales de la empresa. En la especificación de requisitos funcionales, los usuarios describen el tipo de operaciones que se llevarán a cabo con los datos.
- Dos fases de diseño finales. En la fase de diseño lógico, se relaciona el esquema conceptual de alto nivel con el modelo de implementación de datos del sistema de bases de datos que se va a usar. En la fase de diseño físico se especifican las características físicas de la base de datos. Entre esas características están la forma de organización de los archivos y las estructuras de almacenamiento interno.

# Usuarios y administradores de bases de datos



- Los **usuarios normales** son usuarios no sofisticados que interactúan con el sistema invocando alguno de los programas de aplicación que se han escrito previamente.
- Los **programadores de aplicaciones**.
- Los **usuarios sofisticados** interactúan con el sistema sin escribir programas. En su lugar, formulan sus consultas en un lenguaje de consultas de bases de datos o con herramientas como el software de análisis de datos.
- Los **usuarios especializados** son usuarios sofisticados que escriben aplicaciones de bases de datos especializadas que no encajan en el marco tradicional del procesamiento de datos.

# Usuarios y administradores de bases de datos...



- La persona que tiene ese control central sobre el sistema se denomina **administrador de bases de datos (ABD)**.
- Las funciones del ABD incluyen:
  - La definición del esquema.
  - La definición de la estructura y del método de acceso.
  - La modificación del esquema y de la organización física.
  - La concesión de autorización para el acceso a los datos.
  - El mantenimiento rutinario:
    - Copia de seguridad periódica de la base de datos, bien sobre cinta o bien sobre servidores remotos.
    - Supervisión y gestión del espacio libre en disco para las operaciones normales.
    - Supervisar los trabajos que se ejecuten en la base de datos y monitorear el rendimiento.

# Otros Conceptos Importantes



- **Lenguaje de manipulación de datos (LMD):** permite a los usuarios tener acceso a los datos o manipularlos. Los LMD no procedimentales solo necesitan que el usuario especifique los datos que necesita, sin aclarar exactamente la manera de obtenerlos.
- **Un lenguaje de definición de datos (LDD):** es un lenguaje para la especificación del esquema de la base de datos y otras propiedades de los datos.
- Un sistema de bases de datos consta de varios subsistemas:
  - – El subsistema **gestor de almacenamiento** proporciona la interfaz entre los datos de bajo nivel almacenados en la base de datos y los programas de aplicación, así como las consultas remitidas al sistema.
  - – El subsistema **procesador de consultas** compila y ejecuta instrucciones LDD y LMD.

# Otros Conceptos Importantes...



- **El gestor de transacciones:** garantiza que la base de datos permanezca en un estado consistente (correcto) a pesar de los fallos del sistema. El gestor de transacciones garantiza que la ejecución de las transacciones concurrentes se produzca sin conflictos.
- La arquitectura de un sistemas de base de datos se ve muy influida por el sistema computacional subyacente en que se ejecuta el sistema de la base de datos. Pueden ser centralizados o cliente-servidor. Los sistemas de bases de datos también se diseñan para explotar arquitecturas computacionales paralelas. Las bases de datos distribuidas pueden extenderse geográficamente en múltiples computadoras separadas.



# Otros Conceptos Importantes...



- Las técnicas de descubrimiento de conocimientos intentan descubrir automáticamente reglas y patrones estadísticos a partir de los datos. El campo de la minería de datos combina las técnicas de descubrimiento de conocimientos desarrolladas por los investigadores de inteligencia artificial y analistas estadísticos con las técnicas de implementación eficiente que les permiten utilizar bases de datos extremadamente grandes.

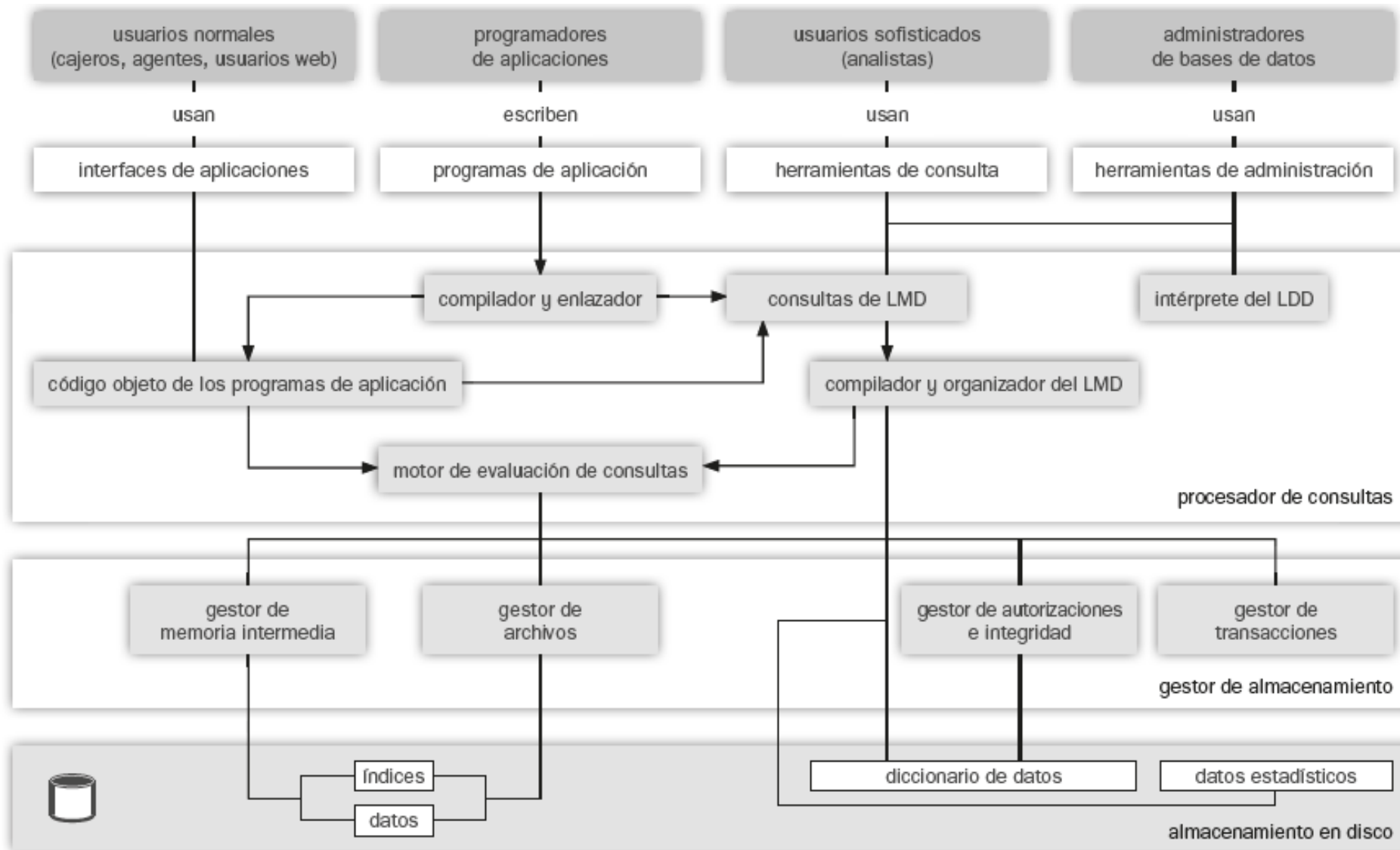


Figura 1.5. Arquitectura del sistema.