

CAPÍTULO 1

Introducción a las bases de datos y a los SGBD

PRESENTACIÓN

- Tipos de bases de datos y aplicaciones de bases de datos
- Definiciones básicas
- Funciones típicas de un SGBD
- Ejemplo de base de datos (UNIVERSIDAD)
- Principales características del enfoque basado en bases de datos
- Tipos de usuarios de bases de datos
- Ventajas del enfoque basado en bases de datos
- Ampliación de las capacidades de las bases de

datos

- Cuándo no utilizar bases de datos

Tipos de bases de datos y aplicaciones de bases de datos

- Aplicaciones tradicionales:
 - Bases de datos numéricas y textuales
- Aplicaciones más recientes:
 - Bases de datos multimedia
 - Sistemas de Información Geográfica (SIG)
 - Bases de datos biológicos y genómicos
 - Almacenes de datos
 - Bases de datos móviles
 - Bases de datos activas y en tiempo real
- La primera parte del libro se centra en las aplicaciones tradicionales
- *Más adelante se describen algunas aplicaciones*

recientes (por ejemplo, en los capítulos 24, 25, 26, 27, 28 y 29).

Acontecimientos recientes (1)

- Las redes sociales empezaron a captar mucha información sobre las personas y sobre las comunicaciones entre ellas -posts, tweets, fotos, vídeos- en sistemas como:
 - Facebook
 - Twitter
 - Linked-In
- Todo lo anterior constituye datos
- Motores de búsqueda (Google, Bing, Yahoo):
recopilan su propio repositorio de páginas web para

realizar búsquedas.

Acontecimientos recientes (2)

- Están surgiendo nuevas tecnologías de los llamados proveedores de software ajeno a las bases de datos para gestionar ingentes cantidades de datos generados en la web:
- Sistemas de almacenamiento de Big Data con grandes clústeres de ordenadores distribuidos (capítulo 25)
- Sistemas NOSQL (Not Only SQL) (capítulo 24)
- Una gran cantidad de datos reside ahora en la "nube", lo que significa que está en

enormes centros de datos que utilizan miles de máquinas.

Definiciones básicas

- **Base de datos:**
 - Una recopilación de datos relacionados.
- **Datos:**
 - Hechos conocidos que pueden registrarse y tienen un significado implícito.
- **Mini-mundo:**
 - Parte del mundo real sobre la que se almacenan datos en una base de datos. Por ejemplo, las notas y expedientes académicos de los estudiantes de una universidad.
- **Sistema de gestión de bases de datos (SGBD):**
 - Paquete o sistema informático que facilita la creación y el mantenimiento de una base de datos informatizada.
- **Sistema de base de datos:**
 - El software DBMS junto con los propios datos. A veces, también se incluyen las aplicaciones.

Impacto de las bases de datos y la tecnología de bases de datos

- **Empresas:** Banca, seguros, comercio minorista, transporte, sanidad, industria manufacturera
- **Industrias de servicios:** Financiero, Inmobiliario, Jurídico, Comercio Electrónico, Pequeñas empresas
- **Educación :** Recursos para contenidos y entrega
- **Más recientemente:** Redes Sociales, Aplicaciones Científicas y Medioambientales, Medicina y Genética

- Aplicaciones personalizadas: basadas en dispositivos móviles inteligentes

Entorno simplificado del sistema de bases de datos

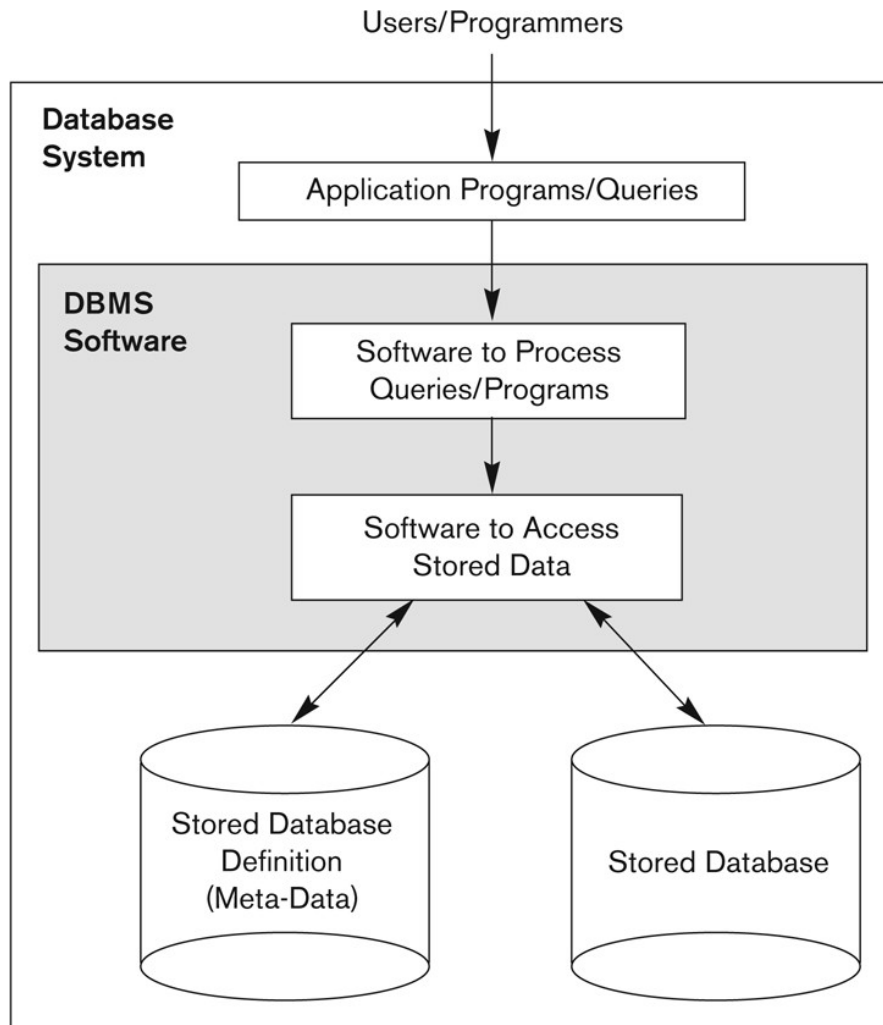


Figure 1.1
A simplified database
system environment.

Funciones típicas de un SGBD

- *Definir* una base de datos concreta en términos de tipos de datos, estructuras y restricciones.
- *Construir* o cargar el contenido inicial de la base de datos en un soporte de almacenamiento secundario
- *Manipulación* de la base de datos:
 - Recuperación: Consulta, generación de informes
 - Modificación: Inserciones, supresiones y actualizaciones de su contenido.
 - Acceso a la base de datos a través de aplicaciones Web
- *Procesamiento y compartición* por parte de un conjunto de usuarios y programas de aplicación simultáneos,

pero manteniendo la validez y coherencia de todos los datos.

Actividades de aplicación contra una base de datos

- Las aplicaciones interactúan con una base de datos generando
 - Consultas: que acceden a distintas partes de los datos y formulan el resultado de una petición.
 - Transacciones: que pueden leer algunos datos y "actualizar" ciertos valores o generar nuevos datos y almacenarlos en la base de datos.
- Las aplicaciones no deben permitir que usuarios no autorizados accedan a los datos
- Las aplicaciones deben adaptarse a las

necesidades cambiantes de los usuarios en función de la base de datos.

Funciones adicionales del SGBD

- DBMS puede proporcionar adicionalmente:
 - Medidas de protección o seguridad para impedir el acceso no autorizado
 - Tratamiento "activo" para realizar acciones internas sobre los datos
 - Presentación y visualización de datos
 - Mantenimiento de la base de datos y de los programas asociados durante toda la vida útil de la aplicación de base de datos.
 - Mantenimiento de bases de datos, software y sistemas

Ejemplo de base de datos (con un modelo conceptual de datos)

- **Mini-mundo para el ejemplo:**
 - Forma parte de un entorno UNIVERSITARIO.
- **Algunas *entidades* del minimundo:**
 - ESTUDIANTES
 - CURSOS
 - SECCIONES (de CURSOS)
 - DEPARTAMENTOS (académicos)
 - INSTRUCTORES

Ejemplo de base de datos (con un modelo conceptual de datos)

- **Algunas *relaciones* en el minimundo:**
 - Las SECCIONES *son de* CURSOS *específicos*
 - Los ESTUDIANTES *toman* SECCIONES
 - Los CURSOS *tienen como prerrequisito* CURSOS
 - Los INSTRUCTORES *imparten las* SECCIONES
 - Los CURSOS *son ofrecidos por* DEPARTAMENTOS
 - Los ESTUDIANTES *se especializan en* DEPARTAMENTOS
- **Nota:** Las entidades y relaciones anteriores se expresan normalmente en un modelo de datos conceptual, como el modelo de datos ENTITY-RELATIONSHIP (véanse los capítulos 3 y 4).

Ejemplo de base de datos sencilla

COURSE

Course_name	Course_number	Credit_hours	Department
Intro to Computer Science	CS1310	4	CS
Data Structures	CS3320	4	CS
Discrete Mathematics	MATH2410	3	MATH
Database	CS3380	3	CS

SECTION

Section_identifier	Course_number	Semester	Year	Instructor
85	MATH2410	Fall	04	King
92	CS1310	Fall	04	Anderson
102	CS3320	Spring	05	Knuth
112	MATH2410	Fall	05	Chang
119	CS1310	Fall	05	Anderson
135	CS3380	Fall	05	Stone

GRADE_REPORT

Student_number	Section_identifier	Grade
17	112	B
17	119	C
8	85	A
8	92	A
8	102	B
8	135	A

PREREQUISITE

Course_number	Prerequisite_number
CS3380	CS3320
CS3380	MATH2410
CS3320	CS1310

Figure 1.2
A database that stores
student and course
information.

Principales características del enfoque basado en bases de datos

- **Naturaleza autodescriptiva de un sistema de base de datos:**
 - Un **catálogo** DBMS almacena la descripción de una base de datos concreta (por ejemplo, estructuras de datos, tipos y restricciones).
 - La descripción se denomina **metadatos***.
 - Esto permite al software DBMS trabajar con diferentes aplicaciones de bases de datos.
- **Aislamiento entre programas y datos:**
 - En el procesamiento tradicional de archivos, la estructura de los archivos de datos está incrustada en los programas de aplicación, por lo que muchos cambios en la estructura de un archivo pueden requerir cambiar todos los programas que acceden a ese archivo. En cambio, los programas de acceso al SGBD no requieren tales cambios. La estructura de los datos se almacena en el catálogo del SGBD independientemente de los programas de acceso. Llamamos a esta propiedad independencia programa-datos

- * Algunos sistemas más recientes, como algunos sistemas NOSQL, no necesitan metadatos: almacenan la definición de los datos dentro de su estructura.

Ejemplo de catálogo de base de datos simplificado

RELATIONS

Relation_name	No_of_columns
STUDENT	4
COURSE	4
SECTION	5
GRADE_REPORT	3
PREREQUISITE	2

Figure 1.3

An example of a database catalog for the database in Figure 1.2.

COLUMNS

Column_name	Data_type	Belongs_to_relation
Name	Character (30)	STUDENT
Student_number	Character (4)	STUDENT
Class	Integer (1)	STUDENT
Major	Major_type	STUDENT
Course_name	Character (10)	COURSE
Course_number	XXXXNNNN	COURSE
....
....
....
Prerequisite_number	XXXXNNNN	PREREQUISITE

Note: Major_type is defined as an enumerated type with all known majors. XXXXNNNN is used to define a type with four alpha characters followed by four digits

Principales características del enfoque basado en bases de datos (continuación)

■ **Abstracción de datos:**

- Ahorra a los usuarios detalles sobre el almacenamiento de los datos. No importa si la base de datos ocupa uno o cientos de archivos, este hecho es transparente para el usuario.

■ **Soporte de múltiples vistas de los datos:**

- Cada usuario puede ver una vista diferente de la base de datos, que describe **sólo** los datos que le interesan.

Principales características del enfoque basado en bases de datos (continuación)

- **Uso compartido de datos y procesamiento de transacciones multiusuario:**
 - Permitir que un conjunto de **usuarios concurrentes** recupere y actualice la base de datos.
 - *El control de concurrencia* dentro del DBMS garantiza que cada **la transacción** se ejecuta correctamente o se cancela
 - El subsistema de *recuperación* garantiza que cada transacción completada tenga su efecto registrado permanentemente en la base de datos.
 - **OLTP** (Online Transaction Processing) es una parte importante de las aplicaciones de bases de datos.

Permite ejecutar cientos de transacciones simultáneas por segundo.

Usuarios de la base de datos

- Los usuarios pueden dividirse en
 - Los que realmente utilizan y controlan el contenido de la base de datos, y los que diseñan, desarrollan y mantienen las aplicaciones de la base de datos (denominados "Actores en escena"), y
 - Los que diseñan y desarrollan el software del SGBD y las herramientas relacionadas, y los operadores de los sistemas informáticos (llamados "trabajadores detrás de la escena").

Usuarios de bases de datos - Actores en escena

- Actores en escena

- **Administradores de bases de datos:**

- Responsable de autorizar el acceso a la base de datos, de coordinar y supervisar su uso, adquirir recursos de software y hardware, controlar su uso y supervisar la eficacia de las operaciones.

- **Diseñadores de bases de datos:**

- Se encargan de definir el contenido, la estructura, las restricciones y las funciones o transacciones con la base de datos. Deben comunicarse con los usuarios finales y comprender sus necesidades.

Usuarios finales de bases de datos

- Actores en escena (continuación)
 - **Usuarios finales:** Utilizan los datos para consultas, informes y algunos actualizan el contenido de la base de datos. Los usuarios finales pueden clasificarse en:
 - **Casual:** acceso ocasional a la base de datos cuando sea necesario
 - **Ingenuos** o paramétricos: constituyen una parte considerable de los usuarios finales de bases de datos.
 - Utilizan funciones previamente bien definidas en forma de "transacciones enlatadas" contra la base de datos.
 - La mayoría de los usuarios de aplicaciones móviles pertenecen a esta categoría.

- Los cajeros son usuarios paramétricos que realizan esta actividad durante todo un turno de operaciones.
- Los usuarios de las redes sociales publican y leen información de

Copyright © 2017 Rwameez Eblmsasirti eandsShamkant B. Navathe

Diapositiv

a 1- 21

Usuarios finales de bases de datos (continuación)

■ **Sofisticado:**

- Entre ellos figuran analistas empresariales, científicos, ingenieros y otras personas que conocen a fondo las capacidades del sistema.
- Muchos utilizan herramientas en forma de paquetes de software que trabajan en estrecha colaboración con la base de datos almacenada.

■ **Independiente:**

- La mayoría mantiene bases de datos personales utilizando aplicaciones listas para usar.
- Un ejemplo es el usuario de un paquete de software financiero que almacena diversos datos financieros personales.

- Otro ejemplo es un usuario que mantiene una base de datos de fotos y vídeos personales.

Usuarios de bases de datos - Actores en escena (continuación)

- **Analistas de sistemas y desarrolladores de aplicaciones**
 - **Analistas de sistemas:** Comprenden los requisitos de los usuarios ingenuos y sofisticados y diseñan aplicaciones que incluyen transacciones enlatadas para satisfacer dichos requisitos.
 - **Programadores de aplicaciones:** Implementan las especificaciones desarrolladas por los analistas y las prueban y depuran antes de su despliegue.
 - **Analistas empresariales:** Cada vez es más necesario contar con este tipo de personas capaces de analizar grandes cantidades de datos empresariales y en tiempo real ("Big Data") para tomar mejores decisiones relacionadas con la planificación, la

publicidad, el marketing, etc.

Usuarios de bases de datos - Actores entre bastidores

- **Diseñadores e implementadores de sistemas:** Diseñan e implementan paquetes de SGBD en forma de módulos e interfaces y los prueban y depuran. El SGBD debe interactuar con aplicaciones, compiladores de lenguajes, componentes del sistema operativo, etc.
- **Desarrolladores de herramientas:** Diseñan e implantan sistemas de software denominados herramientas para el modelado y diseño de bases de datos, la supervisión del rendimiento, la creación de prototipos, la generación de datos de prueba, la creación de interfaces de usuario, la simulación, etc., que facilitan la creación de aplicaciones y permiten utilizar las bases de datos con eficacia.
- **Operadores y personal de mantenimiento:** Gestionan el funcionamiento real y el mantenimiento del

entorno de hardware y software del sistema de base de datos.

Ventajas del enfoque basado en bases de datos

- Control de la redundancia en el almacenamiento de datos y en las tareas de desarrollo y mantenimiento.
 - Intercambio de datos entre varios usuarios.
- Restricción del acceso no autorizado a los datos. Sólo el personal del DBA utiliza comandos e instalaciones privilegiados.
- Proporcionar estructuras de almacenamiento (por ejemplo, índices) para un procesamiento

eficaz de las consultas (véase el capítulo 17).

Ventajas del enfoque basado en bases de datos (continuación)

- Optimización de las consultas para un procesamiento eficaz.
- Prestación de servicios de copia de seguridad y recuperación.
- Ofrecer múltiples interfaces a diferentes clases de usuarios.
- Representación de relaciones complejas entre datos.
- Aplicación de restricciones de integridad en la base

de datos.

Implicaciones adicionales del uso del enfoque de base de datos

- Posibilidad de hacer cumplir las normas:
 - Esto es crucial para el éxito de las aplicaciones de bases de datos en las grandes organizaciones. **Las normas** se refieren a nombres de elementos de datos, formatos de visualización, estructuras de informes, metadatos (descripción de datos), diseños de páginas web, etc.
- Reducción del tiempo de desarrollo de aplicaciones:
 - Se reduce el tiempo necesario para añadir

cada nueva aplicación.

Implicaciones adicionales del uso del enfoque de base de datos (continuación)

- Flexibilidad para cambiar las estructuras de datos:
 - La estructura de la base de datos puede evolucionar a medida que se definen nuevos requisitos.
- Disponibilidad de información actualizada:
 - Extremadamente importante para los sistemas de transacciones en línea, como compras, reservas de avión, hotel o coche.

Ampliación de las capacidades de las bases de datos (1)

- Se están añadiendo nuevas funciones a los SGBD en los siguientes ámbitos:
 - Aplicaciones científicas - Física, Química, Biología - Genética
 - Ciencias de la Tierra y de la Atmósfera y Astronomía
 - XML (lenguaje de marcado extensible)
 - Almacenamiento y gestión de imágenes
 - Gestión de datos de audio y vídeo
 - Almacenamiento de datos y minería de datos: un área muy importante para el desarrollo futuro utilizando nuevas tecnologías (véanse los capítulos 28-29).
 - Gestión de datos espaciales y servicios basados en la localización
 - Gestión de datos históricos
- Lo anterior da pie a *nuevas investigaciones y desarrollos* para incorporar nuevos tipos de datos, estructuras de

datos complejas, nuevas operaciones y esquemas de almacenamiento en los sistemas de bases de datos.

Ampliación de las capacidades de las bases de datos (2)

- Antecedentes desde la llegada del siglo ^{XXI}:
 - La primera década del ^{siglo XXI} ha sido testigo de un enorme crecimiento de los datos generados por los usuarios y recopilados a partir de aplicaciones y motores de búsqueda.
 - Las plataformas de medios sociales como Facebook y Twitter generan millones de transacciones al día y las empresas están interesadas en aprovechar estos datos para "entender" a los usuarios.

- El almacenamiento en nube pone a disposición de usuarios y aplicaciones una cantidad ilimitada de almacenamiento

Ampliación de las capacidades de las bases de datos (3)

- Aparición de las tecnologías Big Data y las bases de datos NOSQL
 - Fue necesaria una nueva tecnología de almacenamiento, gestión y análisis de datos para hacer frente a la avalancha de datos en petabytes al día (1000 terabytes) en algunas aplicaciones, lo que empezó a denominarse comúnmente "Big Data".
 - Hadoop (originario de Yahoo) y el enfoque de programación MapReduce para el procesamiento de datos distribuidos (originario de Google), así como el sistema de archivos de Google, han dado lugar a las tecnologías de Big Data (Capítulo 25).
 - Los sistemas NOSQL (Not Only SQL- donde SQL es el lenguaje estándar de facto para los SGBD relacionales) se han diseñado

para la búsqueda y recuperación rápidas a partir de documentos, el procesamiento de enormes gráficos que se producen en las redes sociales y otras formas de datos no estructurados (Capítulo 24).

Cuándo no utilizar un SGBD

- Costes de utilización de un SGBD:
 - Inversión inicial elevada y posible necesidad de hardware adicional.
 - Gastos generales de seguridad, control de concurrencia, recuperación e integridad.
- Cuándo puede ser innecesario un SGBD:
 - Si la base de datos y las aplicaciones son sencillas, están bien definidas y no se espera que cambien.
 - Si no se requiere el acceso a los datos por parte de varios usuarios.

Resumen del capítulo

- Tipos de bases de datos y aplicaciones de bases de datos
- Definiciones básicas
- Funciones típicas de un SGBD
- Ejemplo de base de datos (UNIVERSIDAD)
- Principales características del enfoque basado en bases de datos
- Tipos de usuarios de bases de datos
- Ventajas del enfoque basado en bases de datos
- Ampliación de las capacidades de las bases de

datos

- Cuándo no utilizar bases de datos