

Bases de datos

Conceptos

Introducción

1. Que es una base de datos
2. Definición de base de datos
3. Características de una base de datos
4. Tipos de bases de datos
5. Modelos de bases de datos
6. Bases de datos jerárquicas
7. Bases de datos de red
8. Bases de datos distribuidas
9. Bases de datos centralizadas
10. Bases de datos descentralizadas
11. Bases de datos transaccionales
12. Bases de datos multidimensionales
13. Bases de datos orientada a objetos
14. Bases de datos relacionales

Que es una base de datos?

Una base de datos, también llamado banco de datos es una colección de datos que pertenecen a un mismo contexto y que se almacenan de manera sistemática para un posterior uso.

Este banco de datos, al estar organizado disminuye la redundancia de datos, facilita el acceso y la administración de los mismos.

Las Bases de datos son accesibles en tiempo real, permiten usuarios concurrentes y el análisis de la variada información almacenada en forma simultánea.

Debido al volumen que llegan a alcanzar las bases de datos se hace necesario que las mismas sean administradas a través de Sistemas Gestores de Bases de Datos (SGDB).

Los Sistemas Gestores de Datos (en inglés DBMS, Database Management System) garantizan el almacenamiento, seguridad, respaldo, gestión y funcionalidad de la base de datos.



Definición de base de datos

Es un conjunto de información almacenada en memoria auxiliar que permite acceso directo y un conjunto de programas que manipulan esos datos

Base de Datos es un conjunto exhaustivo no redundante de datos estructurados organizados independientemente de su utilización y de su implementación.

Son accesibles en tiempo real y compatibles con usuarios concurrentes que analizan o requieren información diferente.

Características de una base de datos



- Independencia
- Reducción en los costos de mantenimiento
- Reducción en la redundancia de los datos
- Disminución de inconsistencias
- Consultas lógicas
- Incrementa la disponibilidad de los datos
- Políticas de seguridad
- Perfiles de Usuarios
- Operaciones con datos

Tipos de base de datos

Según la variabilidad de los datos almacenados:

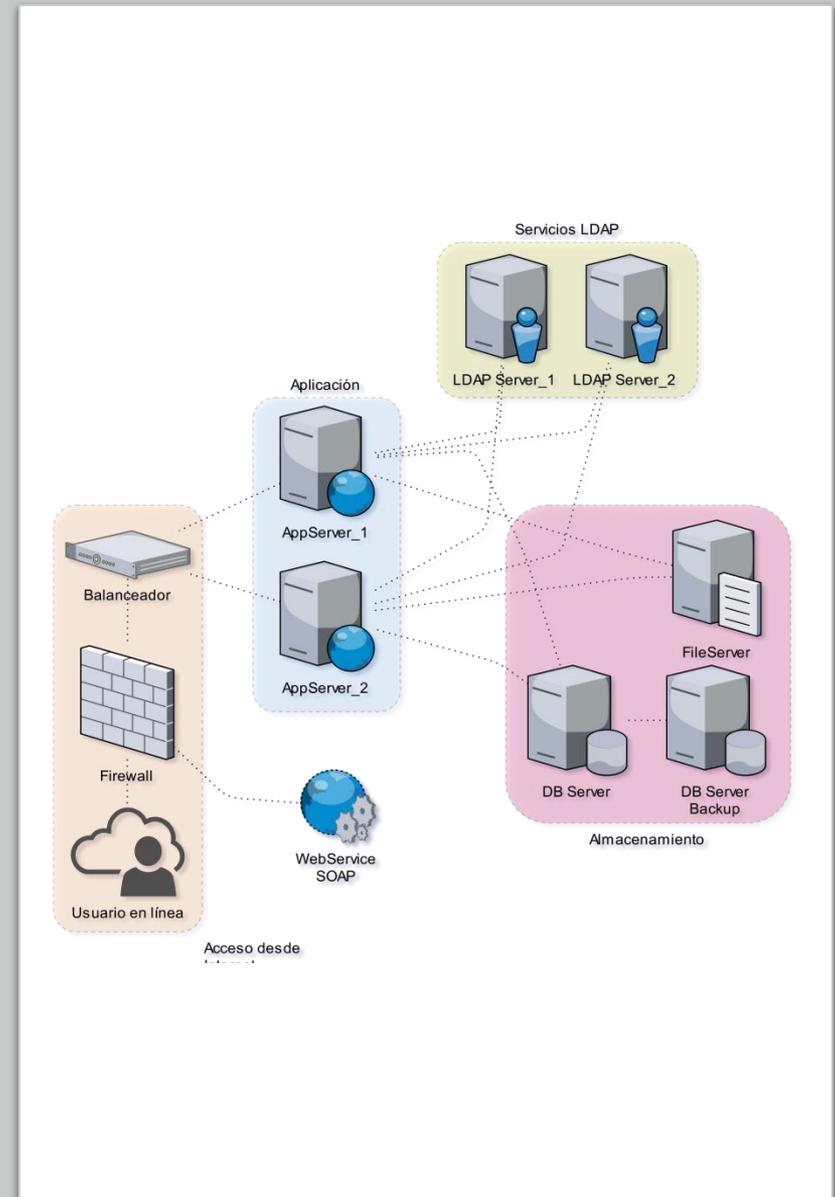
Bases de datos estáticas

Éstas son bases de datos de sólo lectura, utilizadas primordialmente para almacenar datos históricos que posteriormente se pueden utilizar para estudiar el comportamiento de un conjunto de datos a través del tiempo, realizar proyecciones y tomar decisiones.

Bases de datos dinámicas

Éstas son bases de datos donde la información almacenada se modifica con el tiempo, permitiendo operaciones como actualización, borrado y adición de datos, además de las operaciones fundamentales de consulta.

Un ejemplo de esto puede ser la base de datos utilizada en un sistema de información territorial (SIT), un Catastro.



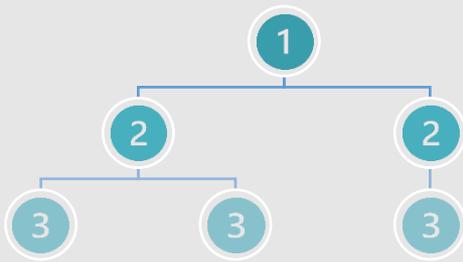
Modelos de bases de datos

Además de la clasificación por la función de las bases de datos, éstas también se pueden clasificar de acuerdo a su modelo de administración de datos.

Un modelo de datos es básicamente una "descripción" de algo conocido como contenedor de datos (algo en donde se guarda la información), así como de los métodos para almacenar y recuperar información de esos contenedores.

Los modelos de datos no son cosas físicas: son abstracciones que permiten la implementación de un sistema eficiente de base de datos; por lo general se refieren a algoritmos, y conceptos matemáticos.

Bases de datos Jerárquicas



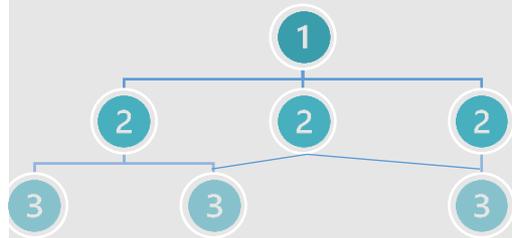
Éstas son bases de datos que, como su nombre indica, almacenan su información en una estructura jerárquica. En este modelo los datos se organizan en un diagrama de árbol invertido, en donde un nodo padre de información puede tener varios hijos.

El nodo origen (que no tiene padres) es llamado raíz, y a los nodos que no tienen hijos se los conoce como hojas.

Las bases de datos jerárquicas son especialmente útiles en el caso de aplicaciones que manejan un gran volumen de información y datos muy compartidos permitiendo crear estructuras estables y de gran rendimiento.

Una de las principales limitaciones de este modelo es su incapacidad de representar eficientemente la redundancia de datos.

Bases de datos de red



Éste es un modelo ligeramente distinto del jerárquico; su diferencia fundamental es la modificación del concepto de nodo: se permite que un mismo nodo tenga varios padres (posibilidad no permitida en el modelo jerárquico).

Fue una gran mejora con respecto al modelo jerárquico, ya que ofrecía una solución eficiente al problema de redundancia de datos; pero, aun así, la dificultad que significa administrar la información en una base de datos de red ha significado que sea un modelo utilizado en su mayoría por programadores más que por usuarios finales.

Bases de datos distribuidas

En un sistema distribuido de bases de datos se almacena la base de datos en varias computadoras. Varios medios de comunicación, como las redes de alta velocidad o las líneas telefónicas, son los que pueden poner en contacto las distintas computadoras de un sistema distribuido.

No comparten ni memoria ni discos. Las computadoras de un sistema distribuido pueden variar en tamaño y función pudiendo abarcar desde las estaciones de trabajo a los grandes sistemas.

Los sistemas distribuidos de bases de datos consisten en sitios débilmente acoplados que no comparten ningún componente físico. Además, puede que los sistemas de bases de datos que se ejecutan en cada sitio tengan un grado sustancial de independencia mutua.

Bases de datos centralizadas

Los sistemas de bases de datos centralizados son aquellos que se ejecutan en un único sistema informático sin interaccionar con ninguna otra computadora. Tales sistemas comprenden el rango desde los sistemas de bases de datos monousuario ejecutándose en computadoras personales hasta los sistemas de bases de datos de alto rendimiento ejecutándose en grandes sistemas. Por otro lado, los sistemas cliente-servidor tienen su funcionalidad dividida entre el sistema servidor y múltiples sistemas clientes.

Sistemas centralizados

Esta configuración consiste en una o varias unidades centrales de procesamiento y un número determinado de controladores para los dispositivos que se encuentran conectados a través de un bus común, el cual proporciona acceso a la memoria compartida unidades centrales de procesamiento y poseen memorias caché locales donde se almacenan copias de ciertas partes de la memoria para acelerar el acceso a los datos.

Bases de datos descentralizadas

La base de datos descentralizada es particularmente útil porque la administración centralizada presenta problemas como:

Degradación del desempeño provocado por un número creciente de ubicaciones remotas a mayores distancias.

Costos altos asociados con el mantenimiento y operación de grandes sistemas de bases de datos centrales (mainframe).

Problemas de confiabilidad creados por la dependencia en un sitio central

El ambiente de negocios dinámico y las deficiencias de las bases de datos centralizados demandaron aplicaciones que permitieran el acceso a los datos de diferentes fuentes de múltiples ubicaciones. Ese ambiente de base de datos de fuentes y ubicaciones múltiples, conocido como base de datos distribuida, es manejado por un sistema de administración de base de datos distribuidos.

Procesamiento descentralizado de datos

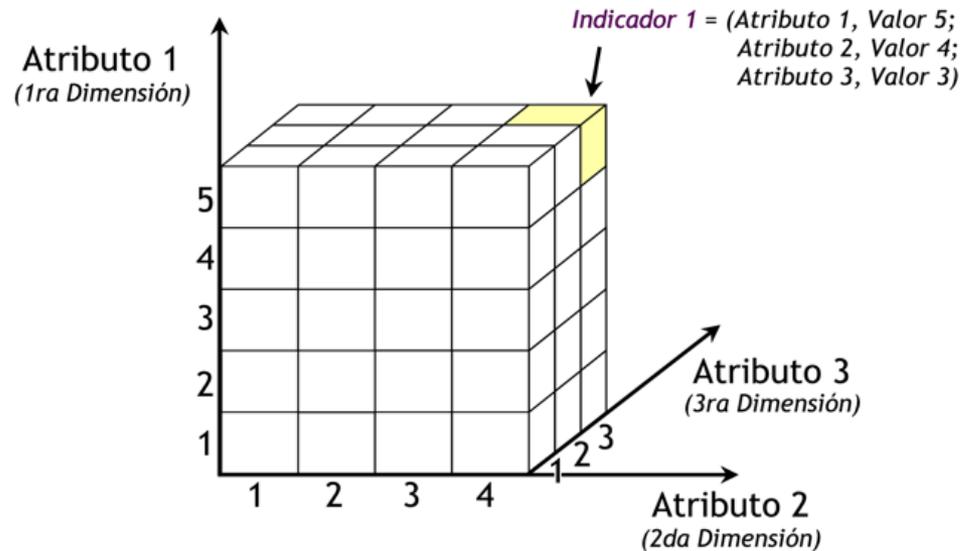
El desarrollo de la tecnología informática, la evolución de la tecnología de bases de datos, y el surgimiento de los sistemas de Procesamiento Distribuidos de Datos (DDP)., que son tres elementos interrelacionados permiten que y las bases de datos pueden "distribuirse" entre los sistemas de cómputo para mejorar la efectividad de los sistemas DDP.

Bases de datos transaccionales

Son bases de datos cuyo único fin es el envío y recepción de datos a grandes velocidades, estas bases son muy poco comunes y están dirigidas por lo general al entorno de análisis de calidad, datos de producción e industrial, es importante entender que su fin único es recolectar y recuperar los datos a la mayor velocidad posible, por lo tanto la redundancia y duplicación de información no es un problema como con las demás bases de datos, por lo general para poderlas aprovechar al máximo permiten algún tipo de conectividad a bases de datos relacionales.

Bases de datos multidimensionales

Son bases de datos ideadas para desarrollar aplicaciones muy concretas, como creación de Cubos OLAP. Básicamente no se diferencian demasiado de las bases de datos relacionales (una tabla en una base de datos relacional podría serlo también en una base de datos multidimensional), la diferencia está más bien a nivel conceptual; en las bases de datos multidimensionales los campos o atributos de una tabla pueden ser de dos tipos, o bien representan dimensiones de la tabla, o bien representan métricas que se desean estudiar.



Bases de datos orientadas a objetos (I)

Este modelo, propio de los modelos informáticos orientados a objetos, trata de almacenar en la base de datos los objetos completos (estado y comportamiento).

Una base de datos orientada a objetos es una base de datos que incorpora todos los conceptos importantes del paradigma de objetos:

Encapsulación - Propiedad que permite ocultar la información al resto de los objetos, impidiendo así accesos incorrectos o conflictos.

Herencia - Propiedad a través de la cual los objetos heredan comportamiento dentro de una jerarquía de clases.

Polimorfismo - Propiedad de una operación mediante la cual puede ser aplicada a distintos tipos de objetos.

Bases de datos orientadas a objetos (II)

En bases de datos orientadas a objetos, los usuarios pueden definir operaciones sobre los datos como parte de la definición de la base de datos.

Una operación (llamada función) se especifica en dos partes. La interfaz de una operación incluye el nombre de la operación y los tipos de datos de sus argumentos (o parámetros).

La implementación (o método) de la operación se especifica separadamente y puede modificarse sin afectar la interfaz.

Los programas de aplicación de los usuarios pueden operar sobre los datos invocando a dichas operaciones a través de sus nombres y argumentos, sea cual sea la forma en la que se han implementado. Esto podría denominarse independencia entre programas y operaciones..

Bases de datos relacionales

Una base de datos relacional es una base de datos que cumple con el modelo relacional, el cual es el modelo más utilizado en la actualidad para implementar bases de datos ya planificadas.

Permiten establecer interconexiones (relaciones) entre los datos (que están guardados en tablas), y a través de dichas conexiones relacionar los datos de ambas tablas, de ahí proviene su nombre: "Modelo Relacional".

Tras ser postuladas sus bases en 1970 por Edgar Frank Codd, de los laboratorios IBM en San José (California), no tardó en consolidarse como un nuevo paradigma en los modelos de base de datos.

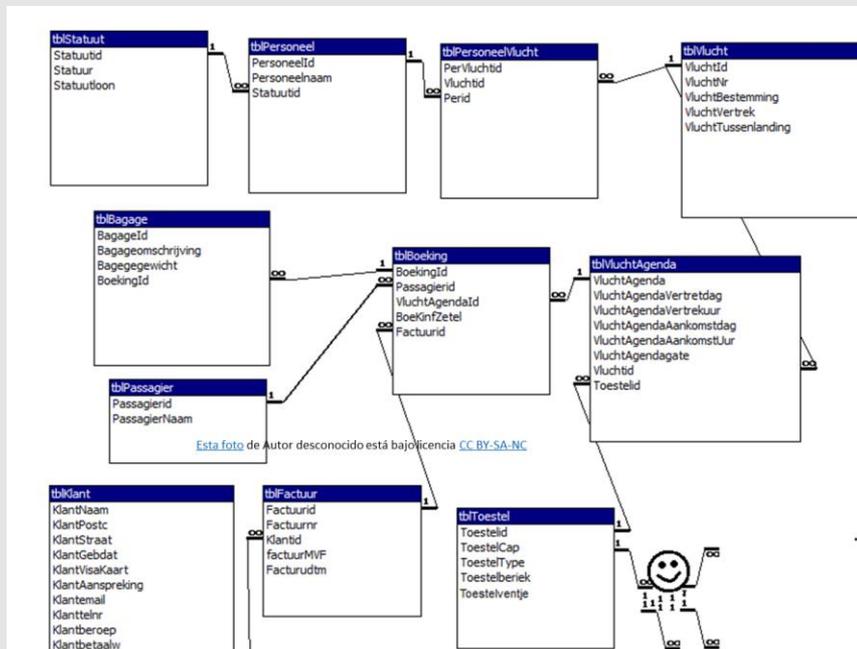
Características de las bases de datos relacionales

Una base de datos relacional se compone de varias tablas o relaciones.

- No pueden existir dos tablas con el mismo nombre.
- Cada tabla es a su vez un conjunto de registros (filas y columnas).

La relación entre una tabla padre y un hijo se lleva a cabo por medio de las claves primarias y ajenas (o foráneas).

- Las claves primarias son la clave principal de un registro dentro de una tabla y éstas deben cumplir con la integridad de datos.
- Las claves foráneas se colocan en la tabla hija, contienen el mismo valor que la clave primaria del registro padre; por medio de éstas se hacen las relaciones.



Estructura de las bases de datos relacionales

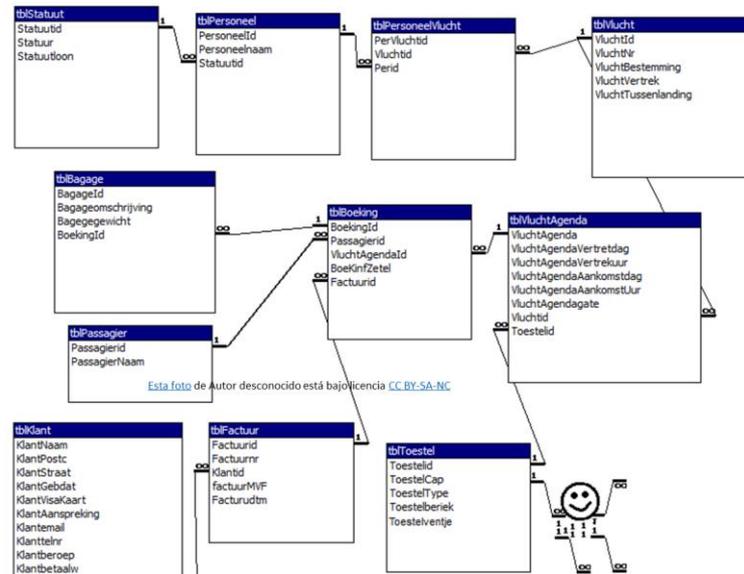
La base de datos se organiza en dos marcadas secciones; el esquema y los datos (o instancia).

El esquema es la definición de la estructura de la base de datos y principalmente almacena los siguientes datos:

- El nombre de cada tabla
- El nombre de cada columna
- El tipo de dato de cada columna
- La tabla a la que pertenece cada columna

Las bases de datos relacionales pasan por un proceso al que se le conoce como normalización, el resultado de dicho proceso es un esquema que permite que la base de datos sea usada de manera óptima.

Los datos o instancia es el contenido de la base de datos en un momento dado. Es en si, el contenido de todos los registros.



Bases de datos

Implementación

Introducción

1. Información y datos
2. Archivos vs. Bases de datos
3. Sistemas gestores de bases de datos
4. Bases de datos con información geográfica
5. Visión de los datos
6. Modelo relacional de bases de datos
7. Diseño de bases de datos
8. Lenguajes de consulta: QBE, SQL
9. Arquitectura de un SGBD
10. Rendimiento
11. Aplicaciones

Información y datos

Datos

- Datos son hechos, eventos, transacciones, etc., que han sido registrados.
- Es la entrada sin procesar de la cual se produce la información.

Información

- La información se registra mediante datos.
- Los datos se interpretan para obtener una información.
- La interpretación depende del humano

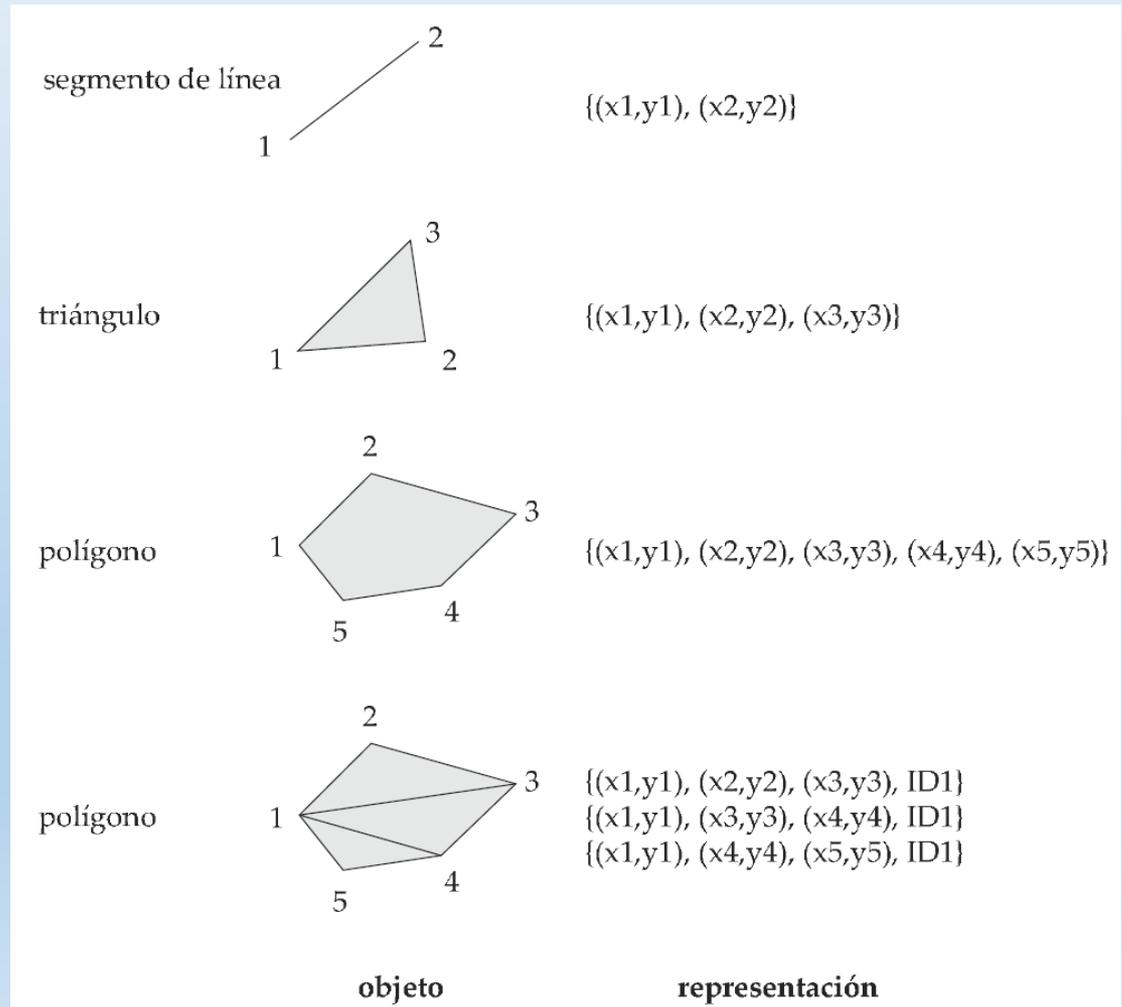
Es necesario un almacenamiento persistente de los datos.

Datos (I)

- Tipo de información representada con datos:
 - Atómica:
 - Números (enteros, coma flotante)
 - Cadenas de caracteres
 - Estructurada:
 - Números complejos (partes real e imaginaria)
 - Registros
 - Ej: coche(matrícula,modelo,color)

Datos (II)

- Información geográfica y espacial:
 - Datos vectoriales



Datos (III)

- Información geográfica y espacial:
 - Datos raster:



2. Archivos vs. Bases de datos

- Archivo: almacenamiento persistente (no volátil) de datos usado para representar información.
 - Ejemplos: Archivos de texto, documentos XML
- Base de datos: colección persistente de datos *relacionados* usada para representar información.
 - Ejemplos MS Access, MS SQL Server, MySQL, PostgreSQL, Oracle, IBM DB2

Problemas de las organizaciones de archivos

- Redundancia e inconsistencia de datos
- Costes de mantenimiento de los programas:
 - Dificultad en el acceso. Cada consulta de datos implica generalmente escribir un nuevo programa
 - Aislamiento de datos. Formatos diferentes y en medios diferentes
 - Integridad. Para implementar restricciones de integridad es necesario modificar todos los programas que accedan a los datos
- Atomicidad.
 - Cuando hay un fallo informático (corte de corriente, error de disco, ...) se puede producir una inconsistencia en una transferencia bancaria. La transferencia debe ser una operación atómica (ocurre totalmente o no ocurre).

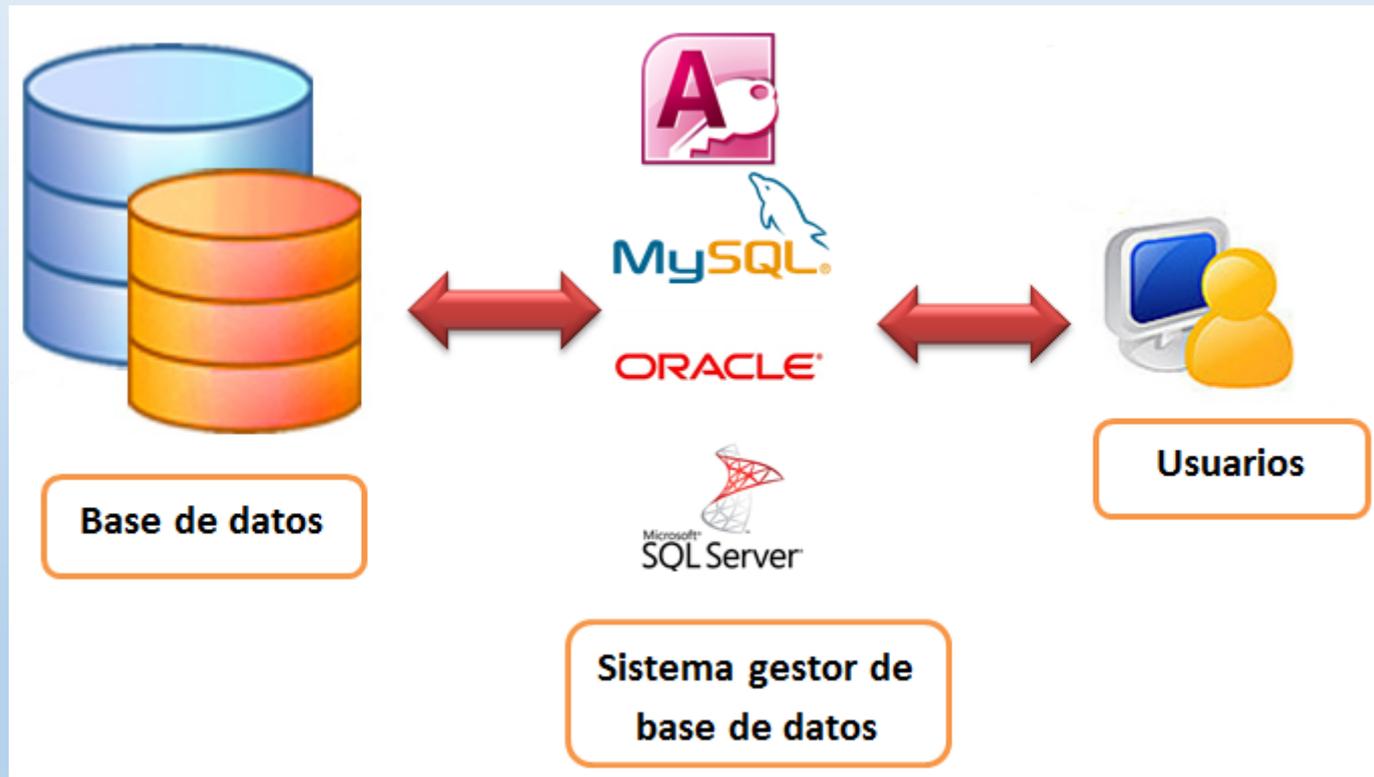
Problemas de las organizaciones de archivos

- Las bases de datos solucionan los problemas de los archivos.
- Sin embargo, no sustituyen a los archivos en todas las aplicaciones.
 - Por ejemplo: archivos temporales, datos no relacionados.
 - Además, las bases de datos se implementan con archivos.

Problemas de las organizaciones de archivos

- Los archivos se prefieren a los SGBD cuando:
 - BD y aplicaciones simples, bien definidas y sin visos de cambio.
 - Requisitos de tiempo real que no pueden cubrir los SGBD.
 - No se requiere acceso concurrente.

Sistemas gestores de bases de datos (SGBDs)



Tipos de SGBDs

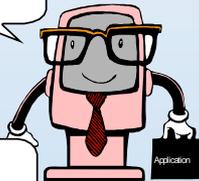
- Relacionales.
- Orientados a objetos.
- Geográficos.
- Temporales.
- XML.

Características de los SGBDs

- Concurrencia
- Integridad
- Seguridad

Concurrencia a nivel de fila

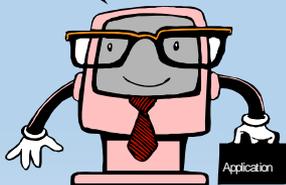
¡Estoy leyendo la 3ª fila!



¡Estoy cambiando la 6ª fila!



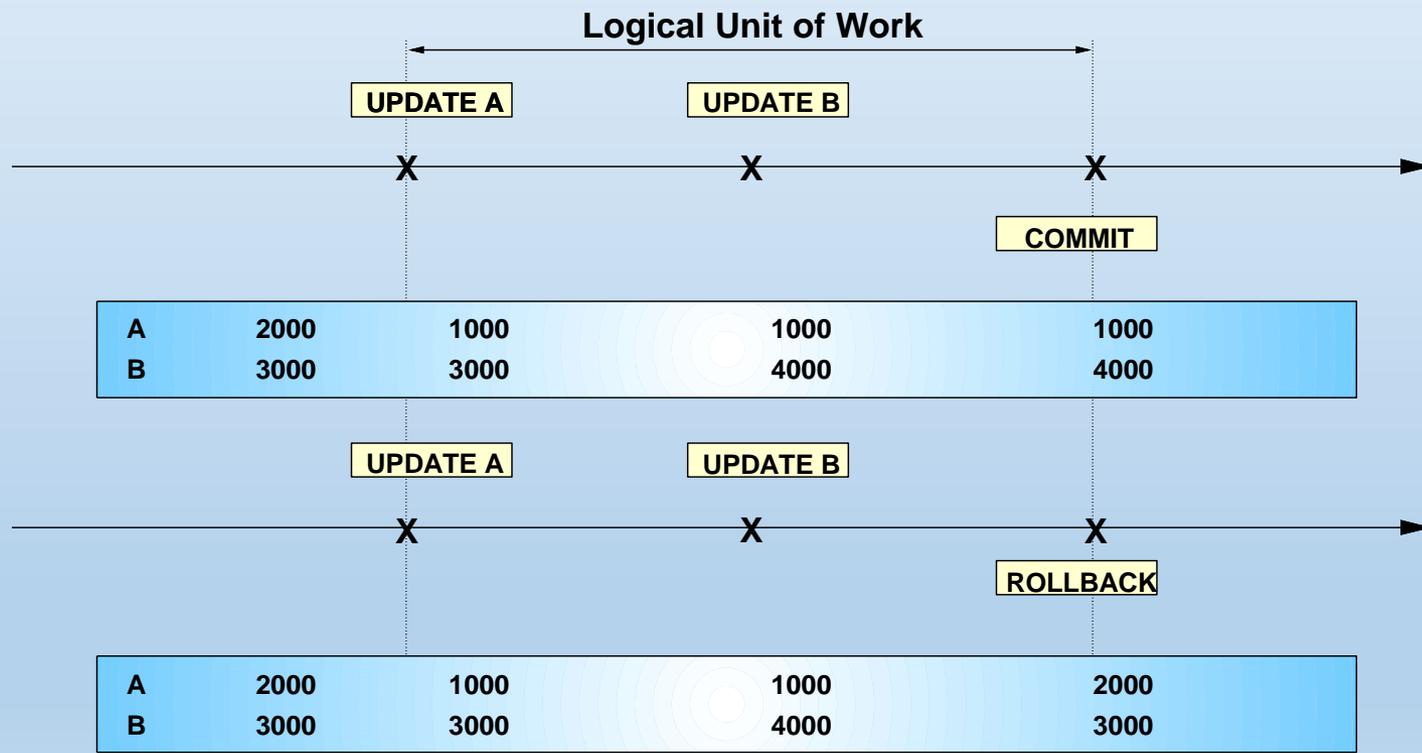
¡Estoy leyendo la 8ª fila y luego la modificaré!



Last Name	First Name	Address	Phone
Adams	Bobby	1426 Main	555-1423
Adams	Sandy	214 Cisco Lane	555-4297
Brown	Terry	45 River Walk	555-8495
Caldwell	Simone	122 42ND ST	555-5367
Fridrich	John	956 Billings	555-4311
Gaines	Lois	6094 Beaker	555-0945
Jones	Larry	567 Yale Ave.	555-6932
Jones	Larry	240 Boston Blvd.	555-5390
Jones	LaVerne	South Star Route	555-2345
Mitchell	Jean	504 1ST Ave.	555-1777
Mitchell	Ted	504 1ST Ave.	555-1778
Smith	Diana	2253 Dahlia	555-6098
Smith	Grace	983 Fairfax	555-8764
Wise	Sara	9023 Colorado Blvd.	555-3242

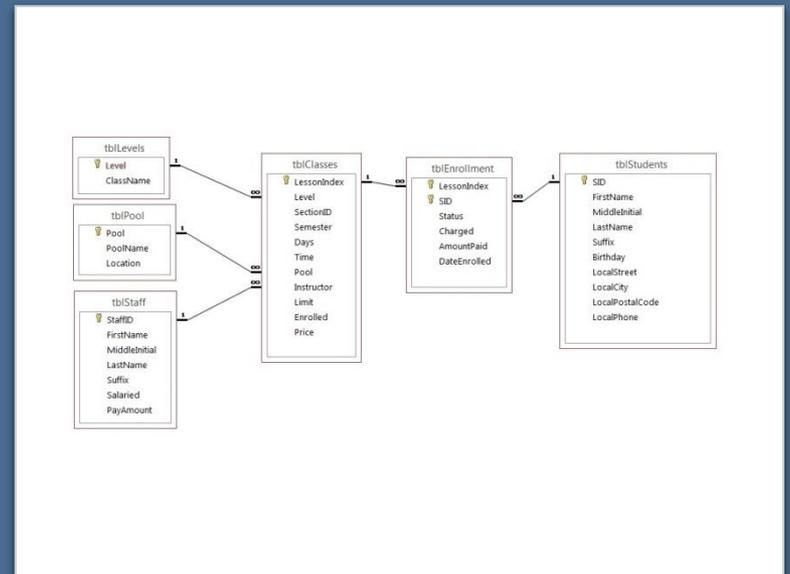
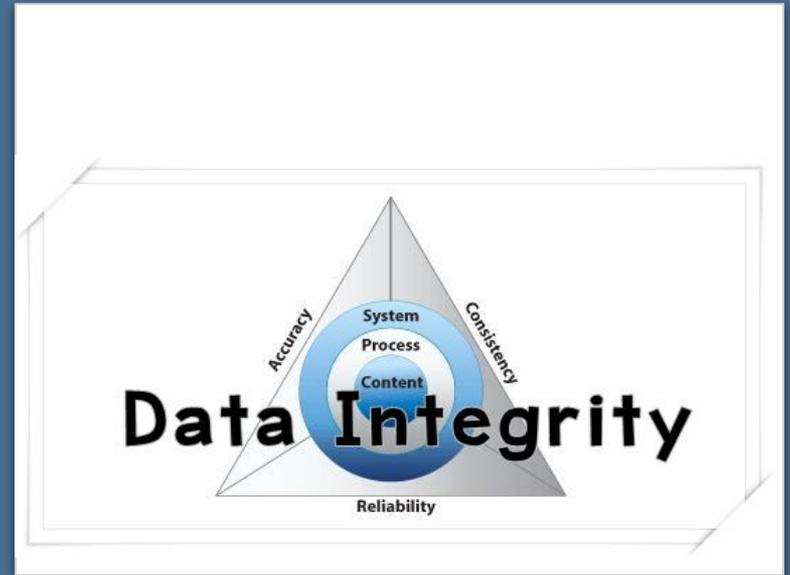
Transacción (unidad de trabajo)

El programa transfiere 1000€ desde la cuenta A a la B



Integridad

- Restricciones de integridad:
 - Dominio.
 - Clave primaria.
 - Clave externa.
 - Definidas por el usuario.



Seguridad

- Un SGBD es un sistema software capaz de manejar grandes colecciones de datos:
 - Relacionados.
 - Compartidos.
 - Persistentes.
- Aseguran su fiabilidad y privacidad.
 - Usuarios.
 - Roles.
 - Restricciones de acceso.

Ventajas

- Como consecuencia de estas características se obtienen las siguientes ventajas:
 - Normalización.
 - Tiempo de desarrollo de aplicaciones reducido.
 - Flexibilidad.
 - Facilidad para añadir o eliminar datos necesarios.
 - Disponibilidad inmediata de las actualizaciones de datos para todos los usuarios.

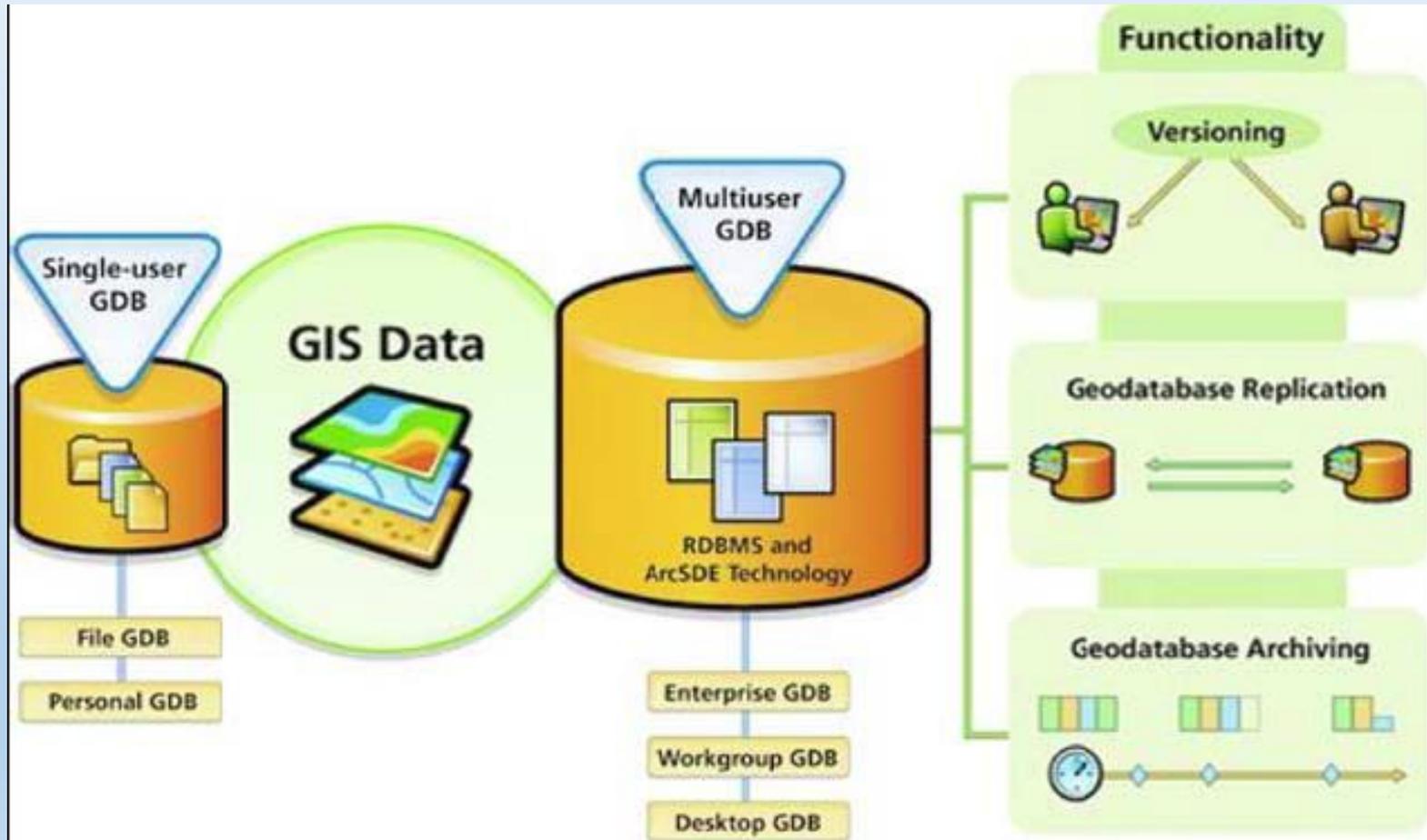
SGBDs

- ◆ MS SQL Server 2008. Libre/Comercial
- ◆ Oracle Spatial. Comercial
- ◆ PostgreSQL (PosGIS). Open-source
- ◆ MySQL Spatial Extensions. Open-source
- ◆ IBM DB2 Spatial Extender. Comercial
- ◆ IBM DB2 Geodetic Extender. Comercial
- ◆ Informix Spatial DataBlade. Comercial
- ◆ Informix Geodetic DataBlade. Comercial
- ◆ Spatial Query Server for Sybase. Comercial

SGBDs con extensiones espaciales

- ◆ MS SQL Server 2008. Free/Commercial
- ◆ Oracle Spatial. Commercial
- ◆ PostgreSQL (PostGIS). Open-source
- ◆ MySQL Spatial Extensions. Open-source
- ◆ IBM DB2 Spatial Extender. Commercial
- ◆ IBM DB2 Geodetic Extender. Commercial
- ◆ Informix Spatial DataBlade. Commercial
- ◆ Informix Geodetic DataBlade. Commercial
- ◆ Spatial Query Server for Sybase. Commercial

4. Bases de datos con información geográfica



Tipos de datos SGBDRs

- Texto
- Memo
- Numérico
- Autonumérico
- Fecha/Hora
- Moneda
- Sí/No
- Objeto OLE
- Hipervínculo

Visualización de los datos

- Objetivo de un SGBD:
 - Proporcionar una visión abstracta de los datos
 - Los detalles de cómo se almacenan y mantienen los datos no son visibles para los usuarios.
- Niveles:
 - Físico
 - Cómo se almacenan los datos.
 - Lógico
 - Qué datos y qué relaciones se almacenan.
 - Vistas
 - Subconjunto del nivel lógico.

Modelo relacional

- Los modelos de datos se usan para poder definir una interpretación de los datos conforme a su estructura, relaciones y operaciones que se les pueden aplicar.
- Por definición, es una colección de conceptos que se usan para describir la estructura de una BD.
- La estructura son los tipos de datos, las relaciones y las restricciones.
- Llevan asociados operaciones básicas.

Modelo relacional

Se usa una colección de tablas para representar los datos y sus relaciones.

TABLA DE CLIENTES				
C.I.	Nombre	Identif./C.	Dirección	Teléfono
16325825	Rivas, Luis	RL708	23654 Santa Rosa	15325948
12035824	Torres, Yeazy	TY011	2536 Calle Roma	12369581
10356528	Cruz, Carlos	CC125	2514 Av. Urdaneta	10256985

CAMPO CLAVE DE:

TABLA DE PEDIDOS					
Núm. Pedido	Identif./C.	Fecha	Monto	Embarque	Cargo envío
000454	RL708	11/02/2005	4.000.080,50	E401	10
000455	TY011	06/05/2005	1.032.200,00	E406	15
000456	CC125	07/05/2005	7.000.230,20	E900	10

Código	Nombre	Apel	Ape2	Depart
1	Juan	García	García	1
2	Pepe	García	Sanchez	1
3	Carlos	Sanchez	Sanz	3
4	Ana	Sanz	Lopez	3
5	Juana	Fernandez	Lopez	2
6	Lucía	Gomez	Lozano	1
7	Pablo	Lozano	García	3
8	Pedro	Heras	Gomez	4
9	Tomas	Alonso	Santos	5

Primary Key

Fila

Columna

Foreign Key

Diseño de bases de datos

- Proceso de creación de un esquema de la base de datos.
- Fases:
 - Conceptual.
 - Lógico.
 - Físico.

Conceptual: Modelo entidad-relación

