

RESUMEN TEORÍA BDD

BASE DE DATOS

Una base de datos, por lo general, se puede definir como un conjunto de información relacionada que se encuentra agrupada y estructurada. Desde el punto de vista informático, podemos decir que es un sistema formado por un conjunto de datos que permiten que un conjunto de programas puedan manipularlos.

DATOS

Los datos son las representaciones informáticas de la información disponible. Es decir, cualquier elemento informativo que tenga relevancia para un usuario.

INFORMACIÓN

La información está formada por los datos que han sido organizados y tratados.

- Mundo real: Son los objetos materiales y no materiales de la realidad.
- Mundo conceptual: Son los conocimientos o informaciones que obtenemos del mundo real de cada persona.
- Mundo de las representaciones.

ENTIDADES

Las entidades son los objetos de la realidad que conceptualizamos. Puede haber dos tipos:

- Entidad tipo: Tipo genérico de entidad, se refiere a un objeto en general.
- Entidad instancia: Se refiere a un objeto concreto, distinguible por algún atributo.

ATRIBUTOS

Los atributos son las características que tiene una entidad. Un atributo puede tener:

- Valor nulo: El valor es desconocido o inexistente.
- Atributo identificativo clave o clave primaria: Es cuando uno o más atributos nos permiten identificar a una entidad concreta.

MUNDO DE LAS REPRESENTACIONES

La representación más frecuente es la representación tabular, estructurada en campos y registros.

EVOLUCIÓN DEL ALMACENAMIENTO DE LA INFORMACIÓN

- Archivos manuales: Los documentos se etiquetaban, ordenaban y clasificaban en ficheros, fue necesario reemplazarlos por un sistema informatizado.
- Sistemas de archivos: Los datos se guardaban en archivos en el ordenador.
- Sistemas gestores de bases de datos: Cualquier cambio en la estructura de los datos no afecta a los programas que los utilizan.

FICHEROS

Para el almacenamiento de la información de forma permanente se utilizan dispositivos de almacenamiento masivo llamados memorias secundarias, ya que los datos guardados en la memoria principal desaparecen cuando se desconecta el ordenador.

- Datos: Son hechos o aspectos que necesitamos guardar para obtener información.
- Campo: Carácter o conjunto de caracteres que tienen un significado específico.
- Registro: Conjunto de campos relacionados siguiendo una cierta lógica, que describen una persona, un lugar...
- Fichero: Conjunto de registros seleccionados.

TIPOS DE ARCHIVOS

Según la forma de acceso:

- Secuencial: Los registros se leen uno tras otro.
- Directo: Permite acceder directamente al registro mediante una clave.
- Indexado: Mediante un acceso directo al bloque de registro concreto de este permite leer de forma secuencial los registros de ese bloque.
- Dinámico: Permite el acceso directo o por índice a un registro y a partir de este se accede de forma secuencial.

INCONVENIENTES DE LOS FICHEROS

- Inconsistencia de la información: Un error en la actualización de esta información puede provocar inconsistencias en los datos del archivo.
- Redundancia: Se tienen datos que no aportan información y que se pueden calcular a partir de otros.
- Rigidez de búsqueda: Cada fichero tiene una organización determinada, según para el tipo de acceso para el que se haya definido.
- Dependencia de los programas: Cualquier cambio en la estructura del archivo conlleva tener que modificar los programas que lo utilizan.

SISTEMAS GESTORES DE BASES DE DATOS

Un SGBD es una colección de programas que permiten a los usuarios crear y mantener la BD.

Características:

- Abstracción: Los SGBD ahorran a los usuarios detalles del almacenamiento físico de los datos.
- Independencia de los datos: Es la capacidad de modificar la estructura física o lógica de una BD sin tener que cambiar las aplicaciones que acceden a ella.
- Redundancia: No hay datos repetidos. Problema de espacio pero, sobre todo, del mantenimiento de la calidad de los datos.
- Integridad y consistencia de datos: Si hemos permitido que exista redundancia, es necesario que el SGBD vigile que esta información repetida se actualice de forma coherente.
- Seguridad: Deben garantizar que la información está segura.

ARQUITECTURA DE LOS SGBD

El esquema de la BD es un elemento fundamental de la arquitectura de un SGBD que permite independizar el SGBD de la BD, cambiar el diseño de la BD, su esquema, sin tener que cambiar el SGBD y, en la medida de lo posible, los programas que accedan.

- Nivel externo y lógico: Describe los conceptos que forman la base de datos en forma de modelos o esquemas.
 - Esquemas externos: Son las diversas visiones que los usuarios tienen de la BD.
 - Esquema conceptual: En él se explicitan los datos, las interrelaciones y las restricciones que se han establecido. Permite trabajar con datos de forma sencilla.
- Nivel interno y físico: Contendrá la descripción de la organización física de la BD. Implica un conocimiento a bajo nivel de la implementación física de la organización de los datos y su acceso.

INDEPENDENCIA DE LOS DATOS

- Independencia física: Cuando los cambios en la organización física de la BD no afectan al nivel lógico (ni al esquema conceptual ni los esquemas externos).

- Independencia lògica: Cuando los usuarios no se ven afectados por cambios en el nivel lògico.

ETAPAS DE DISEÑO DE UNA BD

- Etapa 0 o previa “Anàlisis de requerimientos”
Consiste en entender qué datos tendrán que ser almacenados dentro de la BD.
- Etapa del “Diseño conceptual” (pasamos al modelo E/R)
Se obtiene una estructura de la información de la futura BD.
- Etapa del “Diseño lògico” (pasamos al modelo relacional)
Su objetivo es transformar los modelos obtenidos en un conjunto de datos propios del modelo de datos escogido.
- Etapa del “Diseño físico” (pasamos a SQL)
Será necesario conocer la funcionalidad de SGBD escogido.

SISTEMAS DE FICHEROS TRADICIONALES VS SISTEMAS DE BASES DE DATOS

	Fitxers	Bases de dades
Entitats tipus	Les entitats instància d'un fitxer pertanyen a una sola entitat tipus.	Les BD contenen entitats instància d'infinitat d'entitats tipus interrelacionades.
Interrelacions	El sistema no interrelaciona fitxers.	El sistema té previstes eines per interrelacionar fitxers.
Redundàncies	És necessari crear fitxers a mida de cada aplicació, amb totes les dades necessàries, encara que estiguin repetides en altres fitxers.	Tècnicament, totes les aplicacions poden treballar amb la mateixa BD, la qual cosa evita la redundància de dades i els riscos que comporta.
Inconsistències	És possible que els valors d'unes mateixes dades en diferents fitxers no coincideixin, si els programadors no les han actualitzat degudament.	Si les interrelacions estan ben dissenyades, les dades només han d'estar emmagatzemades en la BD un sol cop. Per tant, no hi ha risc d'inconsistències.
Obtenció de dades	Si no hi ha una aplicació que obtingui les dades que volem, o bé s'ha de fer un programa a mida, o bé s'ha d'aprofitar la sortida d'un programa amb objectius similars, i fer els càlculs necessaris manualment.	Permeten obtenir qualsevol conjunt de dades, segons les necessitats, dels del seu propi entorn de treball, sense haver d'escriure, compilar i executar cap nou programa d'aplicació contra la BD.

Aïllament de dades	Les dades estan disperses i aïllades en diferents arxius, la qual cosa dificulta el desenvolupament de les aplicacions.	Totes les dades són en la mateixa BD, interconnectades, la qual cosa en facilita l'obtenció.
Integritat de dades	Els programes han d'implementar totes les restriccions sobre les dades, afegint el codi font corresponent. El manteniment és complicat quan la informació es conté en diferents fitxers utilitzats per diferents aplicacions.	La BD s'encarrega directament d'implementar les restriccions sobre les dades. Els programes no han d'incorporar codi font addicional per garantir-les.
Atomicitat	Alguns conjunts d'operacions sobre les dades s'han d'executar de manera indivisible (o tots o cap), independentment de les fallades que el sistema pugui presentar (com ara per un tall de subministrament elèctric). Però això és molt difícil de garantir amb un sistema d'informació basat en fitxers.	Les BD incorporen la tècnica de les transaccions per tal de garantir fàcilment l'execució atòmica d'una pluralitat de processos sobre les dades.
Accés concurrent	L'actualització simultània de dades d'un mateix fitxer per part de diferents usuaris o aplicacions en pot provocar fàcilment la inconsistència.	Amb la tècnica del bloqueig, les BD garanteixen automàticament la consistència de les dades, malgrat que més d'un usuari o més d'una aplicació les vulguin actualitzar simultàniament.
Seguretat	Habitualment, cada fitxer serveix per a un sol usuari o una sola aplicació (sobretot simultàniament), i ofereix una visió única del món real. Però no sempre tots els usuaris que utilitzen un fitxer haurien de tenir accés a totes les dades que conté.	Una BD pot ser compartida per molts usuaris de diferents tipus (fins i tot, simultàniament), els quals poden tenir diferents visions (vistes) del món real, en funció del seu perfil i dels permisos que s'hagin de concedir en cada cas.

MODELOS DE BASES DE DATOS

Los modelos son los instrumentos que se aplican a una parcela del mundo real para obtener una estructura de datos a la que llamamos esquema.

Un modelo de datos es un conjunto de herramientas conceptuales que permiten describir los datos, sus relaciones y las reglas de integridad que deben cumplir.

Se pueden agrupar en tres grupos:

- MODELOS LÓGICOS BASADOS EN OBJETOS: Se aplican al nivel lógico.
 - MODELO ENTIDAD-RELACIÓN: Formado por entidades.
 - MODELO ORIENTADO A OBJETOS: Definen a las BD en términos de objetos, de sus propiedades y de sus operaciones. Cada objeto está formado por: variables, clases, métodos y comentarios.
- MODELOS LÓGICOS BASADOS EN REGISTROS: Se aplican al nivel físico.
 - MODELO JERÁRQUICO
 - MODELO EN RED

- MODELO RELACIONAL: Los datos se estructuran en representaciones tabulares, llamadas tablas, que representan entidades tipos del mundo conceptual, y que están formadas por filas y columnas. Las columnas forman los campos y las filas son los registros.

Ventajas:

- Evita la duplicidad de registros.
- Vela por la integridad referencial.
- Favorece la comprensibilidad.

Los principales SGBD relacionales son:

- MySQL
- MariaDB
- SQLite
- PostgreSQL
- Microsoft SQL Server
- Oracle

- MODELOS FÍSICOS DE DATOS: Representa cómo se construirá el modelo a la base de datos físicamente en el sistema informático, además muestra todas las estructuras de la tabla.

Los modelos relacionales no son válidos para algunos programas que necesitan tratar un volumen de datos muy grande o desestructurados.

- BASES DE DATOS DOCUMENTALES: Son aquellas que utilizan un conjunto de programas que gestionan datos a partir de documentos o archivos con diferentes estructuras.
- SGBD NoSQL más usado:
 - MongoDB: Es el más popular y utilizado actualmente.