

Introducción a las bases de datos

Introducción

Desde los primeros días de las tecnologías de la información, los datos han estado en el centro de todo lo que realizan los ordenadores.

Los usuarios introducen datos, las aplicaciones analizan, procesan y realizan cálculos con esos datos. Y por último se produce la salida de resultados.

Toda esta información se debe guardar en algún lugar. Las técnicas empleadas para almacenar datos son sumamente importantes para la velocidad de acceso y recuperación de los mismos. El método más utilizado es el almacenamiento en **bases de datos**.

Bases de datos

De forma sencilla se define una **base de datos** como una serie de datos organizados y relacionados entre sí, los cuales son almacenados y explotados por los sistemas de información de una empresa. Este término fue utilizado por primera vez en 1963, en un simposio celebrado en California (EEUU).

Las bases de datos proporcionan las siguientes **funciones**:

Almacenamiento

Los datos deben mantenerse dentro de un sistema informático, por lo general de almacenamiento en disco, de manera que los datos están disponibles cuando sea necesario. Las técnicas de almacenamiento en disco pueden variar, pero deben de ser rápidas y con espacio suficiente.

Organización

Los datos deben ser almacenados de manera lógica en el disco para que puedan ser encontrados de forma rápida y eficiente.

Acceso

Es una función clave de las bases de datos la devolución rápida y eficiente de la información solicitada.

Seguridad

Una base de datos determina quién puede acceder y lo que se puede hacer con los datos.

Manipulación

Después de agregar la información a la base de datos, puede ser modificada y/o eliminada. El software de la base de datos gestiona las reglas de manipulación de los datos.

Salvaguarda

Las bases de datos deben mantener la información de una forma segura y que se encuentre disponible cuando ocurra un fallo como una pérdida de un disco. Las copias de seguridad (backups) y recuperación (recovery) son los procesos más eficientes para **asegurar la disponibilidad de la información**.

A continuación se exponen los diferentes **tipos de bases de datos**.

1. Bases de datos relacionales

Es un tipo de base de datos que cumple con el modelo relacional, siendo esta técnica la más utilizada en la actualidad para implementar bases de datos. Este modelo está basado en conceptos muy sencillos, permitiendo establecer relaciones (interconexiones) entre los datos (que están guardados en tablas) y, a través de dichas conexiones, obtener información de varias entidades.

El modelo tiene asociada la teoría de normalización de relaciones con objeto de eliminar comportamientos anómalos durante los procesos de manipulación de la información.

Tras ser postuladas sus bases en 1970 por Edgar Frank Codd, de los laboratorios IBM, no tardó en consolidarse como un nuevo paradigma en el modelado de base de datos.

Características:

- Una base de datos relacional se compone de varias tablas o relaciones.
- No pueden existir dos tablas con el mismo nombre.

- Cada tabla tiene a su vez un conjunto de registros (filas y columnas).
- La relación entre una tabla principal y una secundaria se lleva a cabo por medio de las claves primarias y externas (o foráneas).
- La clave primaria es el identificador principal de un registro dentro de una tabla y debe cumplir con la integridad de los datos.
- Las claves externas se colocan en la tabla secundaria, contienen el mismo valor que la clave primaria del registro principal; por medio de estas claves se establecen las relaciones.

La información puede ser recuperada mediante "consultas" que ofrecen una amplia flexibilidad y eficacia para administrar la información.

El lenguaje más utilizado para gestionar las bases de datos relacionales es SQL (Structured Query Language) o Lenguaje de Consulta Estructurado, un estándar implementado por los principales motores o sistemas de gestión de bases de datos relacionales.

2. Bases de datos jerárquicas

En este tipo de base de datos los datos se organizan utilizando estructuras con forma de árbol plano.

Un **ÁRBOL** es una estructura jerárquica en la que los elementos se suelen denominar **NODOS** y existen dependencias entre los nodos.

Bajo este modelo es habitual que se produzcan redundancias de datos en inconsistencias ya que varias aplicaciones pueden tratar el mismo dato.

Otra característica es que existe una alta dependencia de los datos y las aplicaciones. Modificar la estructura de algún archivo de datos obliga a modificar los programas que los tratan.

Generalmente los datos están repartidos en diferentes archivos. Archivos que pertenecen a una determinada aplicación y sólo es posible acceder a los datos a través de esa aplicación por lo tanto esto provoca aislamiento de datos.

Es un modelo en **desuso** en la actualidad. Aunque la más extendida fue IMS (Information Management System). Presentado por IBM en 1968.

3. Bases de datos en red

Es una base de datos similar a la Jerárquica pero que admite a un hijo tener relación con varios padres. Es un modelo que ha quedado en sistemas mainframe y la más extendida es IDMS.

4. Bases de datos orientadas a objetos

Este modelo, bastante reciente y propio de los sistemas informáticos orientados a objetos, permite almacenar en la base de datos objetos completos (estado y comportamiento).

Una base de datos orientada a objetos se diseña para trabajar en conjunción con lenguajes de programación orientados a objetos como Java, C#, Visual Basic.NET y C++, incorporando todos los conceptos fundamentales del paradigma de orientación a objetos.

Características:

- **Encapsulación:** Permite ocultar la información (atributos y métodos) al resto de los objetos, impidiendo así accesos incorrectos o conflictivos.
- **Herencia:** A través de la cual los objetos heredan comportamiento dentro de una jerarquía de clases.
- **Polimorfismo:** Característica aplicada a un método (operación) mediante la cual el método puede ser aplicado a distintos tipos de objetos.

Sistemas gestores de Bases de Datos

Los **Sistemas Gestores de Base de Datos (SGBD)** pueden definirse como un paquete integral de software, que se ejecuta en un sistema servidor, centralizando los accesos a los datos y actuando de interfaz entre los datos físicos y el usuario.

Las principales funciones que debe cumplir un SGBD se relacionan con la creación y mantenimiento de la base de datos, el control de accesos, la manipulación de datos de acuerdo con las necesidades del usuario, el cumplimiento de las normas de tratamiento de datos, evitar redundancias e inconsistencias y mantener la integridad.

Prácticamente es un Sistema Operativo diseñado para el control del acceso a los datos. Para conseguir este control, todo RDBMS (Relational DataBase Management System) posee una serie de subsistemas que se encargan de gestionar cada servicio.

Ver [listado](#) de **SGBD** más comunes.

Algunos de estos **subsistemas** son:

Sistema de gestión de la memoria

Encargado de decidir que parte de la memoria se dedica a cada tarea del RDBMS. Su función es que haya suficiente memoria para que el RDBMS funcione eficazmente y a la vez no dejar menos memoria de la que necesita el Sistema Operativo para que la máquina funcione.

Gestión de Entrada y Salida

Consigue que los accesos a los datos sean adecuados.

Procesador de lenguajes

Interpreta las instrucciones SQL (o de otros lenguajes válidos) que los usuarios lanzan a la base de datos.

Control de procesos

Gestiona los programas en ejecución necesarios para el funcionamiento de la base de datos

Control de la red

Controla las conexiones a la base de datos desde la red y evita problemas en caso de desconexión.

Control de transacciones

Permite gestionar las transacciones (conjunto de operaciones de manipulación de datos que se pueden validar o anular).

Terminología en bases de datos

DBA: Database Administrator es el rol encargado de gestionar la o las bases de datos de un proyecto/organización. Un proyecto que precise almacenamiento de datos puede ir desde lo más sencillo (una instancia de base de datos) hasta lo más complejo, un entorno distribuido con múltiples tecnologías de bases de datos diferentes ofreciendo escalabilidad, redundancia, configuraciones zero-downtime deployment, etc.

DBMS: Database Management System es el sistema gestor de base de datos, es la tecnología que proporciona interacción a los usuarios finales con la base de datos a través de interfaces y que permite administrar y gestionar la información contenida en ella. El DBMS es el que proporciona el acceso a la base de datos.

Database: Database o base de datos, en su significado más amplio es una colección de datos organizados. Es la tecnología de persistencia que permite almacenar y gestionar la información.

Resumen

- A lo largo de esta unidad hemos definido base de datos como una serie de datos organizados y relacionados entre sí, que son almacenados y explotados por los sistemas de información de una empresa.
- Hemos descubierto los diferentes tipos de bases de datos que operan en el mercado, y conocido sus características y funcionalidades: jerárquicas, orientadas a objetos, de datos en red y relacionales. Además hemos aprendido a elegir la más adecuada para nuestros proyectos.
- Por último, hemos aprendido qué son los Sistemas Gestores de Bases de Datos (SGBD), paquetes integrales de software, que se ejecutan en un sistema servidor, centralizando los accesos a los datos y actuando de interfaz entre los datos físicos y el usuario.