

■ SISTEMA GESTOR DE BASE DE DATOS

Un sistema gestor de bases de datos (SGBD) consiste en un conjunto de programas para acceder a datos almacenados. Esta colección de datos, normalmente denominada base de datos, contiene información relevante para una institución.

El objetivo principal de un SGBD es proporcionar una forma de almacenar, recuperar y administrar la información de una base de datos de manera que sea tanto práctica como eficiente.

Los SGBD se diseñan para gestionar grandes cantidades de información. La gestión de los datos implica tanto la definición de estructuras para almacenar la información como la provisión de mecanismos para procesar la información, respondiendo a consultas, generando informes, etc.

Además, deben proporcionar la fiabilidad de la información almacenada, a pesar de las caídas del sistema o los intentos de acceso sin autorización. Si los datos van a ser compartidos entre diversos usuarios, el sistema debe evitar posibles resultados erróneos que podrían producirse si dos usuarios actualizan un mismo dato en forma simultánea.

Un SGBD también llamado motor de base de datos puede manejar en forma simultánea varias bases de datos. Como ejemplo un servidor WEB que mantiene varios sitios WEB donde cada uno de ellos almacena su información en una base de datos propia, necesita una única instalación de MySQL.

Las bases de datos forman una parte esencial de casi todas las empresas actuales que pusieron disponibles en línea muchos servicios a través de interfaces Web que acceden a sus bases de datos.

Las bases de datos permiten:

- Definición de la estructura de datos en forma independiente al programa y al usuario. Cada usuario definido tendrá accesos específicos brindando la confiabilidad necesaria. Por ejemplo si se almacenan datos personales y laborales de empleados, algunos usuarios podrían ver solo algunos de ellos con distintos permisos de lectura y modificación.
- Definir reglas que aseguren la integridad de los datos, es decir que sus valores estén en determinado rango, o que no sean nulos, o que no hagan referencia a elementos que no existen. Ej: los nombres sólo con letras y los teléfonos sólo con números.
- Realizar consultas para extraer o actualizar datos de la base con una sola instrucción. Si esto mismo se quisiera realizar en un lenguaje de programación serían necesarios extensos y complejos programas.
- Brindar mecanismos que permitan la recuperación de la base de datos si se producen fallas en hardware del equipo donde está instalada.

■ COMPONENTES DEL SISTEMA DE BASES DE DATOS

Un sistema de base de datos comprende

- Los datos almacenados.
- Lenguajes de definición y consulta.
- El diccionario de datos.
- El procesador de consultas.
- El gestor de almacenamiento.
- Las aplicaciones.
- Los usuarios.

Los puntos siguientes profundizan las características y funciones de los componentes de los SGBD. Su lectura es optativa y sus contenidos no serán evaluados en las comprobaciones.

1. La base de datos

Almacena los datos y sus relaciones en forma no redundante de forma tal de satisfacer cualquier consulta o modificación de los mismos en forma rápida y eficiente asegurando la integridad de los mismos.

Esto ha conducido al diseño de estructuras de datos complejas para la representación y el almacenamiento.

A fin de simplificar la interacción de los usuarios con la base de datos los desarrolladores han definido varios niveles de abstracción que ocultan la complejidad subyacente en el nivel inferior.

- *Nivel físico:* El nivel más bajo de abstracción describe cómo se almacenan realmente los datos. En el nivel físico se describen en detalle las estructuras de datos complejas de bajo nivel.
- *Nivel lógico:* El siguiente nivel de abstracción describe qué datos se almacenan en la base de datos y qué relaciones existen entre esos datos. La base de datos completa se describe así en términos de un número pequeño de estructuras relativamente simples. Aunque 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. Los administradores de bases de datos, que deben decidir la información que se mantiene en la base de datos, usan el nivel lógico de abstracción.

- *Nivel de vistas:* El nivel más alto de abstracción describe sólo parte de la base de datos completa. A pesar del uso de estructuras más simples en el nivel lógico, queda algo de complejidad, debido a la variedad de información almacenada en una gran base de datos. Muchos usuarios del sistema de base de datos no necesitan toda esta información. En su lugar, tales usuarios necesitan acceder sólo a una parte de la base de datos. Para que su interacción con el sistema se simplifique, se define la abstracción del nivel de vistas. El sistema puede proporcionar muchas vistas para la misma base de datos. Esto permite privacidad de los datos entre distintos tipos de usuarios.



Ejemplares y esquemas

Las bases de datos van cambiando a lo largo del tiempo conforme la información se inserta y borra. La colección de información almacenada en la base de datos en un momento particular se denomina un ejemplar de la base de datos.

El diseño completo de la base de datos se llama el esquema de la base de datos.

Tipos de bases de datos

Existen distintos tipos de bases de datos que difieren tanto en la estructura a nivel físico como a nivel lógico.

El primer tipo de base de datos fue la jerárquica seguida posteriormente por la de red. Estos modelos ligaron fuertemente la implementación física con la estructura lógica lo que complicaba la tarea del modelado de datos, por lo que fueron reemplazados por las base de datos relacionales.

Las bases de datos relacionales son las más ampliamente utilizadas.

En el modelo relacional se utilizan tablas para representar los datos y las relaciones entre ellos.

Cada tabla está compuesta por varias columnas, y cada columna tiene un nombre único. Logra un alto nivel de abstracción en el nivel lógico.

Han surgido también bases de datos orientadas a objetos y bases de datos relacionales orientadas a objetos que están recibiendo una atención creciente.

2. Los lenguajes de base de datos

Un sistema de bases de datos proporciona un lenguaje de definición de datos (LDD) para especificar el esquema de la base de datos y un lenguaje de manipulación de datos (LMD) que permite a los usuarios acceder o manipular los datos organizados mediante el modelo de datos apropiado.

En la práctica, los lenguajes de definición y manipulación de datos no son dos lenguajes separados; en su lugar simplemente forman parte de un único lenguaje de bases de datos, tal como SQL, ampliamente usado.

3. El diccionario de datos o directorio de datos.

Un diccionario de datos contiene metadatos, es decir, datos acerca de los datos. El esquema de una tabla es un ejemplo de metadatos. Un sistema de base de datos consulta el diccionario de datos antes de leer o modificar los datos reales.

Especificamos el almacenamiento y los métodos de acceso usados por el sistema de bases de datos por un conjunto de instrucciones en un tipo especial de LDD denominado lenguaje de almacenamiento y definición de datos.

Estas instrucciones definen los detalles de implementación de los esquemas de base de datos, que se ocultan usualmente a los usuarios. Los valores de datos almacenados en la base de datos deben satisfacer ciertas restricciones de consistencia. El LDD proporciona facilidades para especificar tales restricciones. Los sistemas de bases de datos comprueban estas restricciones cada vez que se actualiza la base de datos.

4. El Procesador de consultas.

Una consulta es una instrucción que permite recuperar información. La parte de un lenguaje de administración de datos que implica recuperación de información se llama lenguaje de consultas.

Sin embargo, como el usuario no especifica cómo conseguir los datos, el sistema de bases de datos tiene que determinar un medio eficiente de acceder a los datos.

El componente procesador de consultas del sistema de bases de datos traduce las consultas LMD en secuencias de acciones en el nivel físico del sistema de bases de datos para dar respuesta a la misma.

5. El gestor de almacenamiento

Un gestor de almacenamiento es un módulo de programa que proporciona la interfaz entre los datos de bajo nivel en la base de datos y los programas de aplicación y consultas emitidas al sistema.

El gestor de almacenamiento es responsable de la interacción con el gestor de archivos. Los datos en bruto se almacenan en disco usando un sistema de archivos, que está disponible habitualmente en un sistema operativo convencional.

El gestor de almacenamiento traduce las diferentes instrucciones LMD a órdenes de un sistema de archivos de bajo nivel. Así, es responsable del almacenamiento, recuperación y actualización de los datos en la base de datos.

El gestor de almacenamiento implementa varias estructuras de datos como parte de la implementación física del sistema:

- Archivos de datos, que almacenan la base de datos en sí.
- Diccionario de datos, que almacena metadatos acerca de la estructura de la base de datos, en particular, el esquema de la base de datos.
- Índices, que proporcionan acceso rápido a elementos de datos que tienen valores.

Los componentes del gestor de almacenamiento incluyen:

- *Gestor de autorización e integridad*: comprueba que se satisfagan las restricciones de integridad y la autorización de los usuarios para acceder a los datos.
- *Gestor de transacciones*: asegura que la base de datos quede en un estado consistente (correcto) a pesar de los fallos del sistema, y que las ejecuciones de transacciones concurrentes ocurran sin conflictos.
- *Gestor de archivos*: gestiona la reserva de espacio de almacenamiento de disco y las estructuras de datos usadas para representar la información almacenada en disco.
- *Gestor de memoria intermedia*: es responsable de traer los datos del disco de almacenamiento a memoria principal y decidir qué datos tratar en memoria caché. El gestor de memoria intermedia es

una parte crítica del sistema de bases de datos, ya que permite que la base de datos maneje tamaños de datos que son mucho mayores que el tamaño de la memoria principal.

6. Las aplicaciones

Los programas de aplicación son programas que se usan para interaccionar con la base de datos.

Los programas de aplicación se escriben usualmente en un lenguaje anfitrión, tal como C, C++, HTML o Java.

Para acceder a la base de datos, las instrucciones de consulta necesitan ser ejecutadas desde el lenguaje anfitrión.

Hay dos maneras de hacer esto:

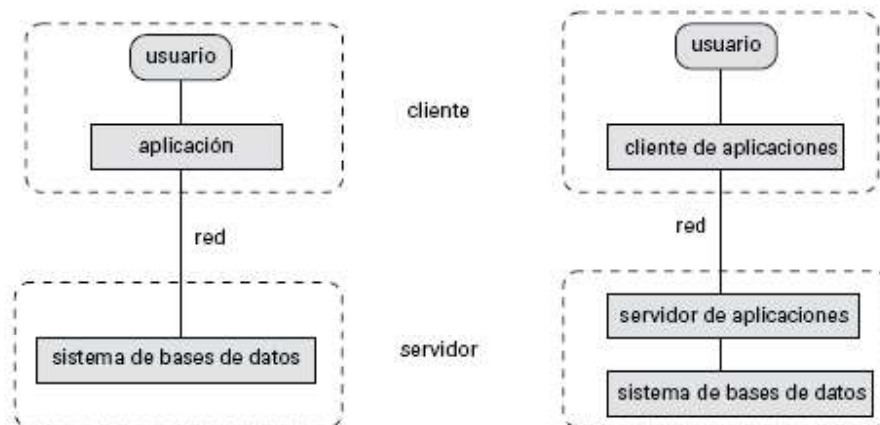
- Proporcionando una interfaz de programas de aplicación (conjunto de procedimientos) que se pueden usar para enviar instrucciones LMD y LDD a la base de datos, y recuperar los resultados.
- Extendiendo la sintaxis del lenguaje anfitrión para incorporar llamadas LMD dentro del programa del lenguaje anfitrión.

Normalmente los usuarios de un sistema de bases de datos se conectan a él a través de una red. Se puede diferenciar entonces entre las máquinas cliente, en donde trabajan los usuarios remotos de la base de datos, y la máquina servidor, en la que se ejecuta el sistema de bases de datos.

Las aplicaciones de bases de datos se dividen usualmente en dos o tres partes.

En una arquitectura de dos capas, la aplicación se divide en un componente que reside en la máquina cliente, que llama a la funcionalidad del sistema de bases de datos en la máquina servidor mediante instrucciones del lenguaje de consultas.

En cambio, en una arquitectura de tres capas, la máquina cliente actúa simplemente como frontal y no contiene ninguna llamada directa a la base de datos. En su lugar, el cliente se comunica con un servidor de aplicaciones, usualmente formularios html. El servidor de aplicaciones, a su vez, se comunica con el sistema de bases de datos para acceder a los datos. Las aplicaciones de tres capas son más apropiadas para grandes aplicaciones, y para las aplicaciones que se ejecutan en World Wide Web.



7. Usuarios

Podemos distinguir cuatro tipos diferentes de usuarios de un sistema de base de datos, diferenciados por la forma en que interactúan:

- *Usuarios normales:* usuarios no sofisticados que interactúan con el sistema mediante la invocación de alguno de los programas de aplicación permanentes que se ha escrito previamente.
- *Usuarios sofisticados:* interactúan con el sistema sin programas escritos. En su lugar, ellos forman sus consultas en un lenguaje de consulta de bases de datos.
- *Programadores de aplicaciones:* profesionales informáticos que escriben programas de aplicación para que sean utilizados por los usuarios normales. Pueden utilizar herramientas para desarrollar interfaces de usuario.
- *Administrador de la base de datos (ABD o DBA):* tiene un control centralizado tanto de los datos como de los programas que acceden a esos datos.

Las funciones del ABD o DBA (en inglés) incluyen:

- *Definición del esquema:* crear el esquema original de la base de datos escribiendo un conjunto de instrucciones de definición de datos en el LDD.
- *Definición de la estructura y del método de acceso.*
- *Modificación del esquema y de la organización física:* para reflejar las necesidades cambiantes de la organización, o para alterar organización física para mejorar el rendimiento.

- *Concesión de autorización para el acceso a los datos*: permite determinar a qué partes de la base de datos puede acceder cada usuario. La información de autorización se mantiene en una estructura del sistema especial que el sistema consulta cuando se intenta el acceso a los datos.
- *Mantenimiento rutinario*, como:
 - Copia de seguridad periódica de la base de datos.
 - Control del espacio libre.
 - Supervisión de los trabajos que se ejecuten