

TEMARIO

UNIDAD 1. INTRODUCCIÓN A LAS BASES DE DATOS

El dato (en plural datos o como un sustantivo en masa) es cualquier secuencia de uno o más símbolos a los que se les da significado mediante actos específicos de interpretación. Los datos (o datos —una sola unidad de datos—) requieren una interpretación para convertirse en **información**. Para traducir datos a información, debe haberse considerado varios factores conocidos. Los factores involucrados están determinados por el creador de los datos y la información deseada. El término metadatos se usa para hacer referencia a los datos sobre los datos. Los **metadatos** pueden estar implícitos, especificados o dados. Los datos relacionados con eventos o procesos físicos también tendrán un componente temporal. En casi todos los casos, este componente temporal está implícito. (definición de los datos en la base de datos o programa por ejemplo una variable x su metadato es int[entero] -> int x)

Concepto de Base de Datos (BD)

Una base de datos es una colección de datos relacionados. Con la palabra datos nos referimos a los hechos (datos) conocidos que se pueden grabar y que tienen un significado implícito

La definición anterior de base de datos es muy genérica;

Una base de datos tiene las siguientes propiedades implícitas:

- Una base de datos representa algún aspecto del mundo real, lo que en ocasiones se denomina minimundo o universo de discurso (UoD, Universe of discourse). Los cambios introducidos en el minimundo se reflejan en la base de datos.
- Una base de datos es una colección de datos lógicamente coherente con algún tipo de significado inherente. No es correcto denominar base de datos a un surtido aleatorio de datos.
- Una base de datos se diseña, construye y rellena con datos para un propósito específico. Dispone de un grupo pretendido de usuarios y algunas aplicaciones preconcebidas en las que esos usuarios están interesados.

En otras palabras, una base de datos tiene algún origen del que se derivan los datos, algún grado de interacción con eventos del mundo real y un público que está activamente interesado en su contenido.

Sistemas de información.

El objetivo de los sistemas de información es entender y analizar cómo ocurre el impacto de la adopción de las tecnologías de información en los procesos de decisión gerenciales y administrativos de las empresas.

El sistema de información puede trabajar con diversos elementos. Entre ellos están software, hardware, base de datos, sistemas especialistas, sistemas de apoyo a la gerencia, entre otros.

Es decir, están incluidos todos los procesos informatizados, que pueden disponibilizar la información correcta y hacer que la empresa funcione de manera adecuada.

- **Relevancia**

El sistema debe generar informaciones relevantes y necesarias a la empresa, que deben ser generadas a tiempo y ser confiables.

- **Integración**

Hay que tener una integración entre el sistema de información y la estructura de la empresa

- **Flujo independiente**

Esa característica es bastante diferenciada, porque, al mismo tiempo en que hay un flujo de procesamiento de datos, que ocurre de manera interna y externa, también hay un flujo independiente de los sistemas de información.

- **Control**

No es obligatorio, pero los sistemas de información pueden contener herramientas de control interno, cuya finalidad es asegurar que las informaciones generadas son confiables y actuar de manera a proteger los datos controlados.

Tipos de sistemas de información

Desde un punto de vista empresarial u organizativo, los sistemas de información pueden clasificarse en:

- **Sistemas de Procesamiento de Transacciones (TPS).** También conocidos como sistemas de gestión operativa, recopilan la información pertinente a las transacciones de la organización, es decir, de su funcionamiento.
- **Sistemas de Información Ejecutiva (EIS).** Monitoriza las variables gerenciales de un área específica de la organización, a partir de la información interna y externa de la misma.
- **Sistemas de Información Gerencial (MIS).** Contemplan la información general de la organización y la comprenden como un todo.
- **Sistemas de soporte de decisiones (DSS).** Orientados al procesamiento de información intra y extra organizacional, para el apoyo en la conducción de la empresa.

Sistema de Base de Datos (SBD)

Un sistema de bases de datos es básicamente un sistema computarizado para llevar registros. Es posible considerar a la propia base de datos como una especie de armario electrónico para archivar; es decir, es un depósito o contenedor de una colección de

archivos de datos computarizados. Los usuarios del sistema pueden realizar una variedad de operaciones sobre dichos archivos.
por ejemplo:

- ≡ Agregar nuevos archivos vacíos a la base de datos;
 - ≡ Insertar datos dentro de los archivos existentes;
 - ≡ Recuperar datos de los archivos existentes;
 - ≡ Modificar datos en archivos existentes;
 - ≡ Eliminar datos de los archivos existentes;
- Eliminar archivos existentes de la base de datos.

Aplicaciones de los SBD.

Así, aunque las interfaces de usuario ocultan los detalles del acceso a las bases de datos, y la mayoría de la gente ni siquiera es consciente de que están interactuando con una base de datos, el acceso a las bases de datos forma actualmente una parte esencial de la vida de casi todas las personas. La importancia de los sistemas de bases de datos se puede juzgar de otra forma—actualmente, los fabricantes de sistemas de bases de datos como Oracle están entre las mayores compañías de software del mundo, y los sistemas de bases de datos forman una parte importante de la línea de productos de compañías más diversificadas como Microsoft e IBM.

Las bases de datos se usan ampliamente. Algunas de sus aplicaciones representativas son:

- Banca: para información de los clientes, cuentas, préstamos y transacciones bancarias.
- Líneas aéreas: para reservas e información de horarios. Las líneas aéreas fueron de las primeras en usar las bases de datos de forma distribuida geográficamente.
- Universidades: para información de los estudiantes, matrículas en las asignaturas y cursos.
- Transacciones de tarjetas de crédito: para compras con tarjeta de crédito y la generación de los extractos mensuales.
- Telecomunicaciones: para guardar un registro de las llamadas realizadas, generar las facturas mensuales, mantener el saldo de las tarjetas telefónicas de prepago y para almacenar información sobre las redes de comunicaciones.
- Finanzas: para almacenar información sobre compañías tenedoras, ventas y compras de productos financieros, como acciones y bonos; también para almacenar datos del mercado en tiempo real para permitir a los clientes la compraventa en línea y a la compañía la compraventa automática.
- Ventas: para información de clientes, productos y compras.

- Comercio en línea: para los datos de ventas ya mencionados y para el seguimiento de los pedidos Web, generación de listas de recomendaciones y mantenimiento de evaluaciones de productos en línea.
- Producción: para la gestión de la cadena de proveedores y para el seguimiento de la producción de artículos en las factorías, inventarios en los almacenes y pedidos.
- Recursos humanos: información sobre los empleados, salarios, impuestos sobre los sueldos y prestaciones sociales, y para la generación de las nóminas.

Propósito de los SBD.

- Los sistemas de bases de datos se diseñan para almacenar grandes cantidades de información. La gestión de los datos implica tanto la definición de estructuras para el almacenamiento de la información como la provisión de mecanismos para la manipulación de la información. Además, los sistemas de bases de datos deben preocuparse de la seguridad de la información almacenada, en caso de caídas del sistema o de intentos de acceso sin autorización. Si los datos deben compartirse entre varios usuarios, el sistema debe evitar posibles resultados anómalos.
- Uno de los propósitos principales de los sistemas de bases de datos es ofrecer a los usuarios una visión abstracta de los datos. Es decir, el sistema oculta ciertos detalles de la manera en que los datos se almacenan y mantienen.

Usuarios de la BD.

Las personas que trabajan con una base de datos se pueden clasificar como usuarios o administradores de bases de datos.

Hay cuatro tipos diferentes de usuarios de los sistemas de bases de datos, diferenciados por la forma en que esperan interactuar con el sistema. Se han diseñado diferentes tipos de interfaces de usuario para los diferentes tipos de usuarios.

Los usuarios normales son usuarios no sofisticados que interactúan con el sistema invocando alguno de los programas de aplicación que se han escrito previamente.

Los programadores de aplicaciones son profesionales informáticos que escriben programas de aplicación.

Los usuarios sofisticados interactúan con el sistema sin escribir programas.

Los usuarios especializados son usuarios sofisticados que escriben aplicaciones de bases de datos especializadas que no encajan en el marco tradicional del procesamiento de datos.

La persona que tiene ese control central sobre el sistema se denomina administrador de bases de datos (ABD). Las funciones del ABD incluyen:

- La definición del esquema.
- La definición de la estructura y del método de acceso.
- La modificación del esquema y de la organización física.

- La concesión de autorización para el acceso a los datos.
- El mantenimiento rutinario.

Diseñadores e implementadores de sistemas DBMS. Diseñan e implementan los módulos y las interfaces DBMS como un paquete software.

Desarrolladores de herramientas. Diseñan e implementan herramientas (paquetes de software que facilitan el modelado y el diseño de la base de datos, el diseño del sistema de bases de datos y la mejora del rendimiento).

Operadores y personal de mantenimiento (personal de administración del sistema). Son los responsables de la ejecución y el mantenimiento real del entorno hardware y software para el sistema de bases de datos.

Ciclo de vida.

UNIDAD 2. CONCEPTOS DE SISTEMA Y DE ARQUITECTURA DE BASES DE DATOS.

Sistema gestor base de datos

es un conjunto de programas que se encargan de manejar la creación y todos los accesos a las bases de datos, está compuesto por:

DDL: Lenguaje de Definición de Datos

DML: Lenguaje de Manipulación de Datos

SQL: Lenguaje de Consulta.

* Definir una base de datos: consiste en especificar los tipos de datos, estructuras y restricciones para los datos que se almacenarán.

* Construir una base de datos: es el proceso de almacenar los datos sobre algún medio de almacenamiento.

* Manipular una base de datos: incluye funciones como consulta, actualización, etc. de bases de datos.

Consideraciones al elegir un DBMS

- Número de usuarios
- Número de transacciones
- Cantidad de datos para almacenar
- Consistencia en la información
- Presupuesto
- Experiencia propia o externa.

Control de la redundancia

La redundancia resultante de almacenar los mismos datos varias veces conduce a serios problemas. En primer lugar, las actualizaciones lógicas sencillas (como la introducción de los datos de un estudiante nuevo) hay que hacerlas varias veces: una por cada archivo donde se almacenen los datos de los estudiantes. Esto lleva a una duplicación del esfuerzo. En segundo lugar, se derrocha espacio de almacenamiento al guardar repetidamente los mismos datos, y este problema puede llegar a ser muy serio en las bases de datos más grandes. En tercer lugar, los archivos que representan los mismos datos pueden acabar

siendo incoherentes, lo que puede ocurrir cuando una determinada actualización se aplica a unos archivos y a otros no. Incluso si una actualización (por ejemplo, la adición de un estudiante nuevo) se aplica a todos los archivos adecuados, los datos relacionados con ese estudiante pueden ser incoherentes porque las actualizaciones han sido aplicadas por los distintos grupos de usuarios.

Restricción del acceso no autorizado

Hay que controlar el tipo de operación de acceso (recuperación o actualización). Normalmente, los usuarios o grupos de usuarios tienen números de cuenta protegidos mediante contraseñas, que pueden utilizar para tener acceso a la base de datos. Un DBMS debe proporcionar seguridad y un subsistema de autorización, que el DBA utiliza para crear cuentas y especificar las restricciones de las mismas. Después, el DBMS debe implementar automáticamente esas restricciones.

Almacenamiento persistente para los objetos del programa

Suministro de estructuras de almacenamiento

para un procesamiento eficaz de las consultas. Los sistemas de bases de datos deben proporcionar capacidades para ejecutar eficazmente consultas y actualizaciones. Como la base de datos normalmente se almacena en el disco, el DBMS debe proporcionar estructuras de datos especializadas para acelerar la búsqueda en el disco de los registros deseados. Con este fin se utilizan unos archivos auxiliares denominados índices, que están basados casi siempre en el árbol de estructuras de datos o en las estructuras de datos dispersas, convenientemente modificados para la búsqueda en disco.

Cuándo no usar un DBMS.

A pesar de las ventajas de usar un DBMS, hay algunas situaciones en las que su uso puede suponer unos sobrecostes innecesarios en los que no se incluiría el procesamiento tradicional de archivos. Los sobrecostes de utilizar un DBMS se deben a lo siguiente:

- Inversión inicial muy alta en hardware, software y formación.
- La generalidad de que un DBMS ofrece definición y procesamiento de datos.
- Costes derivados de las funciones de seguridad, control de la concurrencia, recuperación e integridad.

Es posible que surjan otros problemas si los diseñadores y el DBA no diseñan correctamente la base de datos o si las aplicaciones de sistemas de bases de datos no se implantan correctamente. Por tanto, puede ser más deseable utilizar archivos normales en las siguientes circunstancias:

- Aplicaciones de bases de datos sencillas y bien definidas que no es previsible que cambien.
- Requisitos estrictos y en tiempo real para algunos programas que no podrían satisfacerse debido al sobre coste de un DBMS.
- Inexistencia del acceso multiusuario a los datos.

Algunas industrias y aplicaciones prefieren no utilizar DBMSs de propósito general.

Transacciones.

Una transacción es cualquier ejecución de un programa de usuario en un SGBD. (Las ejecuciones subsecuentes del mismo programa genera varias transacciones.) Se trata de la unidad básica de modificación desde el punto de vista del SGBD: no se permiten transacciones parciales, y el efecto de un grupo de transacciones es equivalente a la ejecución en serie de todas las transacciones. A continuación se describe brevemente la manera en que se garantizan estas propiedades

Los protocolos de bloqueo son conjuntos de reglas que debe seguir cada transacción (y que el SGBD debe hacer que se cumplan) para garantizar que, aunque las acciones de varias transacciones se intercalan, el efecto neto sea idéntico a la ejecución de todas las transacciones en un orden consecutivo dado.

Un bloqueo es un mecanismo empleado para controlar el acceso a los objetos de la base de datos. Los SGBD suelen soportar dos tipos de bloqueos: los compartidos sobre un objeto pueden establecerse dos transacciones diferentes al mismo tiempo, pero el bloqueo exclusivo de un objeto garantiza que ninguna otra transacción establezca un bloqueo sobre ese objeto.

Modelos de datos.

Una colección de herramientas conceptuales para describir los datos, sus relaciones, su semántica y las restricciones de consistencia. Los modelos de datos ofrecen un modo de describir el diseño de las bases de datos en los niveles físico, lógico y de vistas.

Los modelos de datos pueden clasificarse en cuatro categorías diferentes:

- **Modelo relacional.** El modelo relacional usa una colección de tablas para representar tanto los datos como sus relaciones. Cada tabla tiene varias columnas, y cada columna tiene un nombre único. El modelo relacional es un ejemplo de un modelo basado en registros. Los modelos basados en registros se denominan así porque la base de datos se estructura en registros de formato fijo de varios tipos. Cada tabla contiene registros de un tipo dado. Cada tipo de registro define un número fijo de campos, o atributos. Las columnas de la tabla se corresponden con los atributos del tipo de registro. El modelo de datos relacional es el modelo de datos más ampliamente usado, y una gran mayoría de sistemas de bases de datos actuales se basan en el modelo relacional.

- **El modelo entidad-relación.** El modelo de datos entidad-relación (E-R) se basa en una percepción del mundo real que consiste en una colección de objetos básicos, denominados entidades, y de las relaciones entre ellos. Una entidad es una “cosa” u “objeto” del mundo real que es distinguible de otros objetos. El modelo entidad-relación se usa mucho en el diseño de bases de datos y en el Capítulo 6 se examina detalladamente.

- **Modelo de datos orientado a objetos.** El modelo de datos orientado a objetos es otro modelo de datos que está recibiendo una atención creciente. El modelo orientado a objetos se puede considerar como una extensión del modelo E-R con los conceptos de la

encapsulación, los métodos (funciones) y la identidad de los objetos. En el Capítulo 9 se examina este modelo de datos.

- *Modelo de datos semiestructurados*. El modelo de datos semiestructurados permite la especificación de datos donde los elementos de datos individuales del mismo tipo pueden tener diferentes conjuntos de atributos. Esto lo diferencia de los modelos de datos mencionados anteriormente, en los que cada elemento de datos de un tipo particular debe tener el mismo conjunto de atributos. El lenguaje de marcas extensible (XML, eXtensible Markup Language) se emplea mucho para representar datos semiestructurados. Se estudia en el Capítulo 10. El modelo de datos de red y el modelo de datos jerárquico precedieron cronológicamente al relacional.

Arquitectura de tres niveles.

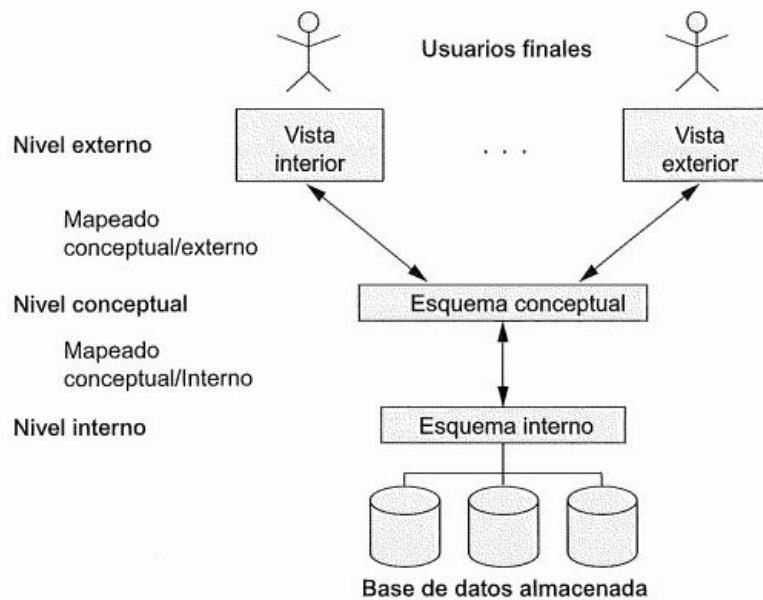
Arquitectura de tres esquemas

El objetivo de la arquitectura de tres esquemas, ilustrada en la Figura 2.2, es separar las aplicaciones de usuario y las bases de datos físicas. En esta arquitectura se pueden definir esquemas en los siguientes tres niveles:

1. El **nivel interno** tiene un **esquema interno**, que describe la estructura de almacenamiento físico de la base de datos. El esquema interno utiliza un modelo de datos físico y describe todos los detalles del almacenamiento de datos y las rutas de acceso a la base de datos.

2. El **nivel conceptual** tiene un **esquema conceptual**, que describe la estructura de toda la base de datos para una comunidad de usuarios. El esquema conceptual oculta los detalles de las estructuras de almacenamiento físico y se concentra en describir las entidades, los tipos de datos, las relaciones, las operaciones de los usuarios y las restricciones. Normalmente, el esquema conceptual se describe con un modelo de datos representativo cuando se implementa un sistema de bases de datos. Este esquema conceptual de implementación se basa a menudo en un diseño de esquema conceptual en un modelo de datos de alto nivel.

Figura 2.2. Arquitectura de tres esquemas.



3. El **nivel de vista** o **externo** incluye una cierta cantidad de esquemas externos o vistas de usuario.

Un esquema externo describe la parte de la base de datos en la que un grupo de usuarios en particular está interesado y le oculta el resto de la base de datos. Como en el caso anterior, cada esquema externo se implementa normalmente mediante un modelo de datos representativo, posiblemente basado en un diseño de esquema externo de un modelo de datos de alto nivel.

Independencia de datos.

Una ventaja muy importante del empleo de un SGBD es que ofrece independencia con respecto a los datos. Es decir, los programas de aplicación quedan aislados de las modificaciones debido al modo en que se estructuran y se guardan los datos. La independencia con respecto a los datos se consigue mediante el empleo de los tres niveles de abstracción de los datos; en concreto, el esquema conceptual y el externo ofrecen claras ventajas en este campo.

El concepto de independencia de los datos, que puede definirse como la capacidad de cambiar el esquema en un nivel de un sistema de bases de datos sin tener que cambiar el esquema en el siguiente nivel más alto. Se pueden definir dos tipos de independencia de datos:

1. Independencia lógica de datos. Es la capacidad de cambiar el esquema conceptual sin tener que cambiar los esquemas externos o los programas de aplicación. Es posible cambiar el esquema conceptual para expandir la base de datos (añadiendo un tipo de registro o un elemento de datos), para cambiar las restricciones o para reducir la base de datos (eliminando un tipo de registro o un elemento de datos). En el último caso, no deben verse afectados los esquemas externos que sólo se refieren a los datos restantes. Por ejemplo, el esquema externo de la Figura 1.5(a) no debe verse afectado por

cambiar el archivo INFORME_CALIF (o tipo de registro) de la Figura 1.2 por el mostrado en la Figura 1.6(a). Sólo es necesario cambiar la definición de la vista y los mapeados en un DBMS que soporta la independencia lógica de datos. Una vez que el esquema conceptual sufre una reorganización lógica, los programas de aplicación que hacen referencia a las estructuras de esquema externo deben funcionar como antes. En el esquema conceptual se pueden introducir cambios en las restricciones sin que se vean afectados los esquemas externos o los programas de aplicación.

2. Independencia física de datos. Es la capacidad de cambiar el esquema interno sin que haya que cambiar el esquema conceptual. Por tanto, tampoco es necesario cambiar los esquemas externos. Puede que haya que realizar cambios en el esquema interno porque algunos archivos físicos fueron reorganizados (por ejemplo, por la creación de estructuras de acceso adicionales) de cara a mejorar el rendimiento de las recuperaciones o las actualizaciones. Si en la base de datos permanecen los mismos datos que antes, no hay necesidad de cambiar el esquema conceptual. Por ejemplo, el suministro de una ruta de acceso para mejorar la velocidad de recuperación de los registros de sección (véase la Figura 1.2) por semestre y año no debe requerir modificar una consulta del tipo "listar

todas las secciones ofrecidas en otoño de 2004", aunque el DBMS ejecutará la consulta con más eficacia utilizando la ruta de acceso nueva.

Arquitectura de los SGBD.

Un sistema gestor de bases de datos (SGBD) consiste en una colección de datos interrelacionados y un conjunto de programas para acceder a dichos datos. La colección de datos, normalmente denominada base de datos, contiene información relevante para una empresa.

La Figura 1.3 muestra la arquitectura (con algunas simplificaciones) de un SGBD típico basado en el modelo relacional de datos.

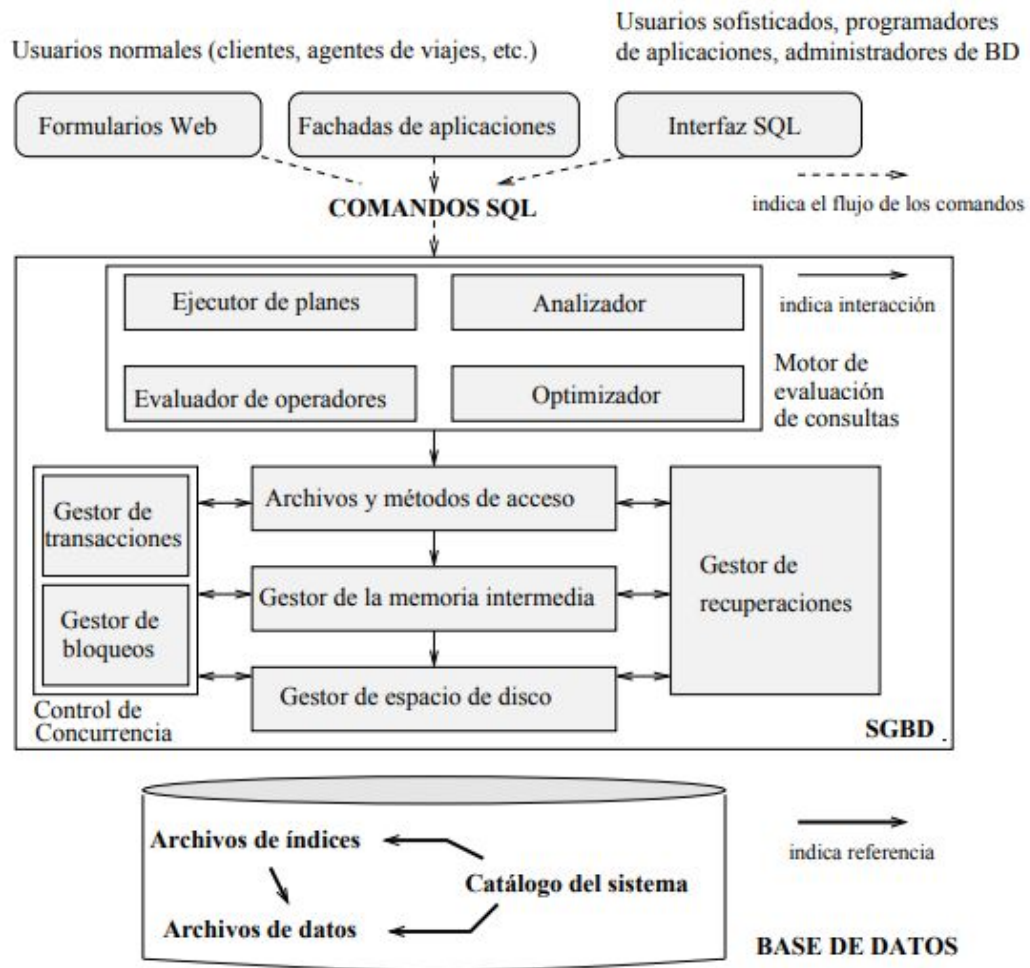


Figura 1.3 Arquitectura de un SGBD

<https://www.youtube.com/watch?v=h5FggYdjdMA>

El SGBD acepta las órdenes SQL generadas por gran variedad de interfaces de usuario, produce planes de evaluación de consultas, ejecuta esos planes contra la base de datos y devuelve las respuestas.

Cuando un usuario formula una consulta, se analiza y envía esta consulta a un optimizador de consultas, que utiliza información sobre el modo en que se guardan los datos para producir un plan de ejecución eficiente para la evaluación de esa consulta. Un plan de ejecución es un plan detallado para la evaluación de la consulta, representado habitualmente como un árbol de operadores relacionales (con anotaciones que contienen información detallada adicional sobre los métodos de acceso que se deben emplear, etcétera). Los operadores relacionales son como los elementos constitutivos de la evaluación de las consultas planteadas a los datos.

El código que implementa los operadores relacionales se sitúa por encima de la capa de los archivos y los métodos de acceso. Esta capa soporta el concepto de archivo, que es un conjunto de páginas o de registros en los SGBD. Se admiten tanto los archivos en montículos, o archivos de páginas sin ordenar, así como los índices. Además de realizar el

seguimiento de las páginas de los archivos, esta capa organiza la información en el interior de cada página. El código de la capa de archivos y métodos de acceso se sitúan por encima del gestor de la memoria intermedia, que lleva las páginas desde el disco a la memoria principal según va haciendo falta, en respuesta a las solicitudes de lectura. La capa inferior del software de los SGBD se ocupa de la administración del espacio de disco, donde se almacenan los datos. Las capas superiores asignan, desasignar, leen y escriben las páginas (mediante las oportunas rutinas) a través de esta capa, denominada gestor del espacio de disco.

Los SGBD soportan la concurrencia y la recuperación de fallos mediante la cuidadosa programación de las solicitudes de los usuarios y el mantenimiento de un registro de todas las modificaciones de la base de datos. Entre los componentes del SGBD asociados al control de la concurrencia y la recuperación están el gestor de transacciones, que garantiza que las transacciones soliciten y liberen los bloqueos de acuerdo con el correspondiente protocolo de bloqueo y programa la ejecución de las transacciones; el gestor de bloqueos, que realiza un seguimiento de las solicitudes de bloqueo y concede los bloqueos sobre los objetos de la base de datos cuando quedan disponibles; y el gestor de recuperaciones, que es responsable del mantenimiento de un registro y de la restauración del sistema a un estado consistente tras los fallos. El gestor del espacio de disco, el gestor de la memoria intermedia, y las capas de archivos y métodos de acceso deben interactuar con estos componentes.

Arquitectura cliente/servidor – centralizada y distribuida.

1. Centralizado

Los sistemas de bases de datos centralizados son aquellos que se ejecutan en un único sistema informático sin interactuar 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.

Los terminales conectados a un sistema central han sido suplantados por computadoras personales. De igual forma, la interfaz de usuario, que solía estar gestionada directamente por el sistema central, está pasando a ser gestionada, cada vez más, por las computadoras personales. Como consecuencia, los sistemas centralizados actúan hoy como sistemas servidores que satisfacen las peticiones generadas por los sistemas clientes. En la Figura 20.2 se representa la estructura general de un sistema cliente–servidor

1. Distribuido

En un sistema distribuido de bases de datos se almacena la base de datos en varias computadoras. Los medios de comunicación como las redes de alta velocidad o las líneas telefónicas 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

- Datos compartidos. La principal ventaja de construir un sistema distribuido de bases de datos es poder disponer de un entorno donde los usuarios puedan acceder desde una única ubicación a los datos que residen en otras ubicaciones.
- Autonomía. La principal ventaja de compartir datos por medio de distribución de datos es que cada ubicación es capaz de mantener un grado de control sobre los datos que se almacenan localmente.
- Disponibilidad. Si un sitio de un sistema distribuido falla, los sitios restantes pueden seguir trabajando.

Preguntas y ejercicios

Indíquese cuatro aplicaciones que se hayan usado que sea muy posible que utilicen un sistema de bases de datos para almacenar datos persistentes.

Indíquese cuatro diferencias significativas entre un sistema de procesamiento de archivos y un SGBD.

El DBMS (Sistema de gestión de bases de datos) y el Sistema de archivos son dos formas que podrían usarse para administrar, almacenar, recuperar y manipular datos. Un sistema de archivos es una colección de archivos de datos sin procesar almacenados en el disco duro, mientras que DBMS es un paquete de aplicaciones que se dedica a administrar los datos almacenados en bases de datos. Es el sistema integrado utilizado para administrar bases de datos digitales, que permite el almacenamiento de contenido de la base de datos, creación / mantenimiento de datos, búsqueda y otras funcionalidades. Ambos sistemas se pueden utilizar para permitir al usuario trabajar con datos de manera similar. Un sistema de archivos es una de las primeras formas de administrar los datos. Pero debido a las deficiencias presentes en el uso de un Sistema de archivos para almacenar datos electrónicos, los Sistemas de gestión de bases de datos comenzaron a usarse en algún momento más tarde, ya que proporcionan mecanismos para resolver esos problemas. Pero debe tenerse en cuenta que, incluso en un DBMS, los datos eventualmente se almacenan (físicamente) en algún tipo de archivo.

Explíquese la diferencia entre independencia de datos física y lógica.

Indíquese cinco responsabilidades del sistema gestor de bases de datos. Para cada responsabilidad, explíquese los problemas que surgirían si no se asumiera esa responsabilidad.

Indíquense al menos dos razones para que los sistemas de bases de datos soportan la manipulación de datos mediante un lenguaje de consultas declarativo como SQL, en vez de limitarse a ofrecer una biblioteca de funciones de C o de C++ para llevar a cabo la manipulación de los datos.

Explíquese los problemas que causa el diseño de la tabla de la Figura 1.5. 1.11 ¿Cuáles son las cinco funciones principales del administrador de bases de datos?

Las respuestas a las preguntas de repaso pueden hallarse en los apartados indicados.

¿Cuáles son las ventajas principales del empleo de SGBD para administrar los datos en aplicaciones que impliquen un amplio acceso a los datos? (Apartados 1.1, 1.4)

¿Cuándo se deben guardar los datos en SGBD en lugar de hacerlo en los archivos del sistema operativo y viceversa? (Apartado 1.3)

¿Qué son los modelos de datos? ¿Qué es el modelo de datos relacional? ¿Qué es la independencia con respecto a los datos y como la soportan los SGBD? (Apartado 1.5)

Explíquense las ventajas de emplear un lenguaje de consultas para procesar los datos en lugar de los programas habituales. (Apartado 1.6)

¿Qué son las transacciones? ¿Qué garantía ofrecen los SGBD con respecto a las transacciones? (Apartado 1.7)

¿Qué son los bloqueos de los SGBD y por qué se emplean? ¿Qué es el registro previo a la escritura y por qué se utiliza? ¿Qué son los puntos de control y por qué se usan? (Apartado 1.7)

Identifíquense los principales componentes de los SGBD y explíquense brevemente lo que hace cada uno de ellos. (Apartado 1.8)

Explíquense los diferentes papeles de los administradores de bases de datos, los programadores de aplicaciones y los usuarios finales de las bases de datos. ¿Quién necesita saber más sobre los sistemas de bases de datos? (Apartado 1.9)

EJERCICIOS

Ejercicio 1.1 ¿Por qué elegir un sistema de bases de datos en lugar de limitarse a guardar los datos en los archivos del sistema operativo? ¿Cuándo tendría sentido no emplear un sistema de bases de datos?

Ejercicio 1.2 ¿Qué es la independencia lógica con respecto a los datos y por qué es importante?

Ejercicio 1.3 Explíquense la diferencia entre la independencia lógica con respecto a los datos y la física.

Ejercicio 1.4 Explíquense las diferencias entre los esquemas externo, interno y conceptual. ¿Cómo están relacionadas estas capas de esquemas con los conceptos de independencia lógica y física con respecto a los datos?

Ejercicio 1.5 ¿Cuáles son las responsabilidades de los DBA? Si se supone que el DBA no está interesado en ejecutar nunca sus propias consultas, ¿sigue necesitando comprender la optimización de las consultas? ¿Por qué?

Ejercicio 1.6 Avaro Puñocerrado desea guardar la información (nombres, direcciones, descripciones de momentos embarazosos, etc.) sobre las muchas víctimas de su lista. Como era de esperar, el volumen de datos le impulsa a comprar un sistema de bases de datos. Para ahorrar dinero, desea comprar uno con las mínimas características posibles, y piensa ejecutarlo como aplicación independiente en su ordenador personal. Por supuesto, Avaro no piensa compartir esa lista con nadie. Indíquese por cuáles de las siguientes características de los SGBD debe pagar Avaro; en cada caso, indíquese también el motivo de que Avaro deba (o no) pagar por esa característica del sistema que va a comprar.

1. Un dispositivo de seguridad.
2. Control de la concurrencia.
3. Recuperación de fallos.
4. Un mecanismo de vistas.
5. Un lenguaje de consultas.

Ejercicio 1.7 ¿Cuál de los siguientes elementos desempeñan un papel importante en la representación de la información sobre el mundo real en las bases de datos? Explíquese brevemente.

1. El lenguaje de definición de datos.
2. El lenguaje de manipulación de datos.
3. El gestor de la memoria intermedia.
4. El modelo de datos.

Ejercicio 1.8 Describase la estructura de un SGBD. Si se actualiza el sistema operativo para que soporte alguna función nueva en los archivos del sistema operativo (por ejemplo, la posibilidad de obligar a que se guarde en disco alguna secuencia de bytes), ¿qué capa(s) del SGBD habría que volver a escribir para aprovechar esas funciones nuevas?

Ejercicio 1.9 Respondan las siguientes preguntas:

1. ¿Qué son las transacciones?
2. ¿Por qué los SGBD intercalan las acciones de las diferentes transacciones en lugar de ejecutar una transacción después de otra?
3. ¿Qué deben garantizar los usuarios con respecto a las transacciones y a la consistencia de las bases de datos? ¿Qué deben garantizar los SGBD con respecto a la ejecución concurrente de varias transacciones y a la consistencia de las bases de datos?
4. Explíquese el protocolo de bloqueo estricto de dos fases.

5. ¿Qué es la propiedad WAL(registro de escritura previa [Write-Ahead Log, WAL]) y por qué es importante?