

# Sistema de gestión de bases de datos

Un sistema de gestión de bases de datos (SGBD) es un conjunto de programas que permiten el almacenamiento, modificación y extracción de la información en una base de datos, además de proporcionar herramientas para añadir, borrar, modificar y analizar los datos. Los usuarios pueden acceder a la información usando herramientas específicas de interrogación y de generación de informes, o bien mediante aplicaciones al efecto.

Estos sistemas también proporcionan métodos para mantener la integridad de los datos, para administrar el acceso de usuarios a los datos y para recuperar la información si el sistema se corrompe. Permiten presentar la información de la base de datos en variados formatos. La mayoría incluyen un generador de informes. También pueden incluir un módulo gráfico que permita presentar la información con gráficos y tablas.

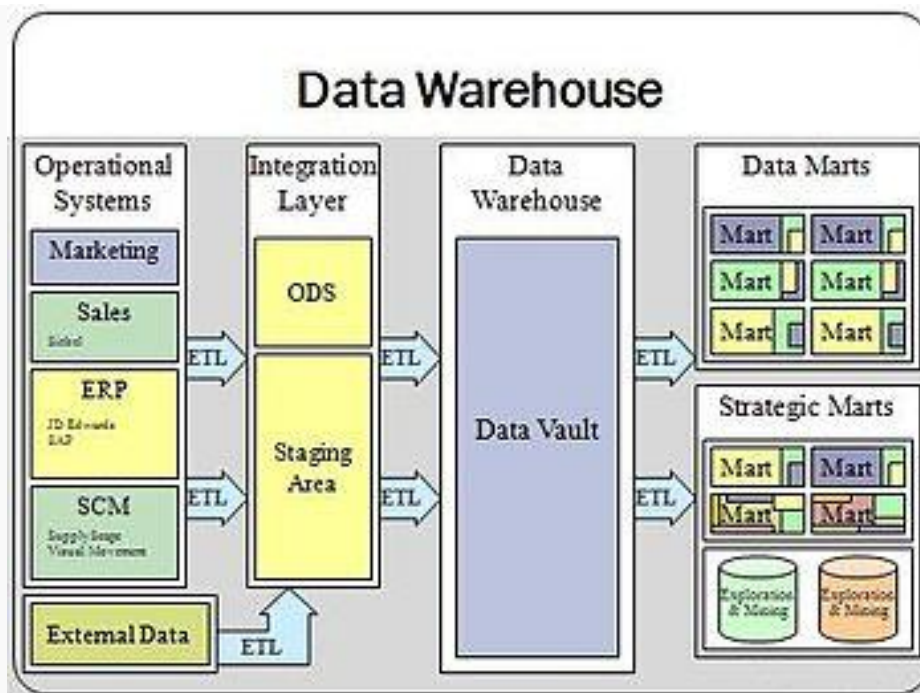
Hay muchos tipos distintos según cómo manejen los datos y muchos tamaños distintos de acuerdo a si operan en computadoras personales y con poca memoria o grandes sistemas que funcionan en *mainframes* con sistemas de almacenamiento especiales.

Generalmente se accede a los datos mediante lenguajes de interrogación, lenguajes de alto nivel que simplifican la tarea de construir las aplicaciones. También simplifican la interrogación y la presentación de la información.

Un SGBD permite controlar el acceso a los datos, asegurar su integridad, gestionar el acceso concurrente a ellos, recuperar los datos tras un fallo del sistema y hacer copias de seguridad. Las bases de datos y los sistemas para su gestión son esenciales para cualquier área de negocio, y deben ser gestionados con esmero.

Las bases de datos generalmente funcionan en computadoras dedicadas de forma exclusiva a este campo. Por las prestaciones requeridas, generalmente funcionan en computadoras multiprocesador con abundante memoria. Para el almacenamiento de los datos puede contar con sistemas de disco propio (DAS), puede conectarse a una red de almacenamiento (SAN) o conectarse a un sistema de almacenamiento en red (NAS). Existen aceleradores *hardware*, usados en grandes sistemas de proceso de transacciones. Los SGBD se encuentran en el corazón de toda aplicación que maneje datos. Los SGBD se basan en sistemas operativos estándar para efectuar dichas funciones.

## Almacén de datos



## Descripción de un Data Warehouse.

En el contexto de la informática, un almacén de datos (del inglés *data warehouse*) es una colección de datos orientada a un determinado ámbito (empresa, organización, etc.), integrado, no volátil y variable en el tiempo, que ayuda a la toma de decisiones en la entidad en la que se utiliza. Se trata, sobre todo, de un expediente completo de una organización, más allá de la información transaccional y operacional, almacenado en una base de datos diseñada para favorecer el análisis y la divulgación eficiente de datos (especialmente OLAP, *procesamiento analítico en línea*). El almacenamiento de los datos no debe usarse con datos de uso actual. Los almacenes de datos contienen a menudo grandes cantidades de información que se subdividen a veces en unidades lógicas más pequeñas dependiendo del subsistema de la entidad del que procedan o para el que sea necesario.

## Cubos de información

---

Artículo principal: **Cubo OLAP**

Los cubos de información o cubos OLAP funcionan como los cubos de rompecabezas en los juegos, en el juego se trata de armar los colores y en el data warehouse se trata de organizar los datos por tablas o relaciones; los primeros (el juego) tienen 3 dimensiones, los cubos OLAP tienen un número indefinido de dimensiones, razón por la cual también reciben el nombre de hipercubos.

Un cubo OLAP contendrá datos de una determinada variable que se desea analizar, proporcionando una vista lógica de los datos provistos por el sistema de información hacia el data warehouse, esta vista estará dispuesta según unas dimensiones y podrá contener información calculada. El análisis de los datos está basado en las dimensiones del hipercono, por lo tanto, se trata de un análisis multidimensional.

A la información de un cubo puede acceder el ejecutivo mediante "tablas dinámicas" en una hoja de cálculo o a través de programas personalizados. Las tablas dinámicas le permiten manipular las vistas (cruces, filtrados, organización, totales) de la información con mucha facilidad. Las diferentes operaciones que se pueden realizar con cubos de información se producen con mucha rapidez. Llevando estos conceptos a un data warehouse, éste es una colección de datos que está formada por «dimensiones» y «variables», entendiendo como dimensiones a aquellos elementos que participan en el análisis y variables a los valores que se desean analizar.

## Middleware

Artículo principal: **Middleware**

Middleware es un término genérico que se utiliza para referirse a todo tipo de software de conectividad que ofrece servicios u operaciones que hacen posible el funcionamiento de aplicaciones distribuidas sobre plataformas heterogéneas. Estos servicios funcionan como una capa de abstracción de software distribuida, que se sitúa entre las capas de aplicaciones y las capas inferiores (sistema operativo y red). El *middleware* puede verse como una capa API, que sirve como base a los programadores para que puedan desarrollar aplicaciones que trabajen en diferentes entornos sin preocuparse de los protocolos de red y comunicaciones en que se ejecutarán. De esta manera se ofrece una mejor relación costo/rendimiento que pasa por el desarrollo de aplicaciones más complejas, en menos tiempo.

La función del middleware en el contexto de los data warehouse es la de asegurar la conectividad entre todos los componentes de la arquitectura de un almacén de datos.

## Diseño de un almacén de datos

---

Para construir un Data Warehouse se necesitan herramientas para ayudar a la migración y a la transformación de los datos hacia el almacén. Una vez construido, se requieren medios para manejar grandes volúmenes de información. Se diseña su arquitectura dependiendo de la estructura interna de los datos del almacén y especialmente del tipo de consultas a realizar. Con este criterio los datos deben ser repartidos entre

numerosos data marts. Para abordar un proyecto de data warehouse es necesario hacer un estudio de algunos temas generales de la organización o empresa, los cuales se describen a continuación:

- Situación actual de partida.- Cualquier solución propuesta de data warehouse debe estar muy orientada por las necesidades del negocio y debe ser compatible con la arquitectura técnica existente y planeada de la compañía.
- Tipo y características del negocio.- Es indispensable tener el conocimiento exacto sobre el tipo de negocios de la organización y el soporte que representa la información dentro de todo su proceso de toma de decisiones.
- Entorno técnico.- Se debe incluir tanto el aspecto del hardware (mainframes, servidores, redes,...) así como aplicaciones y herramientas. Se dará énfasis a los Sistemas de soporte a decisiones (DSS), si existen en la actualidad, cómo operan, etc.
- Expectativas de los usuarios.- Un proyecto de data warehouse no es únicamente un proyecto tecnológico, es una forma de vida de las organizaciones y como tal, tiene que contar con el apoyo de todos los usuarios y su convencimiento sobre su bondad.
- Etapas de desarrollo.- Con el conocimiento previo, ya se entra en el desarrollo de un modelo conceptual para la construcción del data warehouse.
- Prototipo.- Un prototipo es un esfuerzo designado a simular tanto como sea posible el producto final que será entregado a los usuarios.
- Piloto.- El piloto de un data warehouse es el primero, o cada uno de los primeros resultados generados de forma iterativa que se harán para llegar a la construcción del producto final deseado.
- Prueba del concepto tecnológico.- Es un paso opcional que se puede necesitar para determinar si la arquitectura especificada del data warehouse funcionará finalmente como se espera.

#### ACTIVIDAD EN CLASE:

1. En bases de datos se denomina ACID a un conjunto de características necesarias para que una serie de instrucciones puedan ser consideradas como una transacción. Así pues, si un sistema de gestión de bases de datos es ACID que podemos concluir de ello?
2. MySQL soporta cinco tipos de tablas relacionadas con las transacciones. Cuáles son? Y para qué sirven?
3. En pocas palabras que es la INTEGRIDAD REFERENCIAL?
4. DATAMINING, DATAWAREHOUSE, OLAP, bases de datos multimedia y las BDOO, son las tendencias más importantes actualmente en BD. En que consiste cada una de ellas (Explique brevemente)
5. Que es la Migración de datos entre diferentes aplicaciones de BD?

#### TEMAS PROYECTO FINAL:

\*\*\*Organizar pareja para exposición / muestra practica Diferencias entre SGBD

\*\*\*Practicas modelo CONCEPTUAL – Herramientas CASE (DIA-Workbench)

\*\*\*Conceptos Diagrama y modelo ER y creación BD mysql

\*\*\*Realizacion de PROYECTO BD online (cloud computing)