

Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Introducción

Sistemas de información

Datos y tipos de datos

Ficheros

Bases de datos y SGBD

Diseño de bases de datos

Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Introducción

Un dispositivo de almacenamiento de datos es un conjunto de componentes electrónicos habilitados para leer o grabar datos en el soporte de almacenamiento de datos de forma temporal o permanente



La creación de datos, a punto de explotar

Cantidad real y prevista de datos generados en todo el mundo (en zettabytes)



@Statista_ES

Fuente: Statista Digital Economy Compass 2019

statista

MareNostrum



Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Introducción

Sistemas de información

Datos y tipos de datos

Ficheros

Bases de datos y SGBD

Diseño de bases de datos

Qué es el almacenamiento de datos

El almacenamiento de datos es el proceso tecnológico por el cual se archiva, organiza y comparten los bytes de información que componen los archivos que se utilizan en el día a día como documentos de texto, imágenes, configuraciones, vídeos, sonidos y cualquier otra información en formato digital.

Componentes

- Hardware: componentes físicos que constituyen la computadora que almacenará la información.
- Software: aplicaciones informáticas específicas encargadas de gestionar los datos.
- Datos: información que almacena y gestiona el sistema de información.
- Recursos humanos: personal que hace uso y hace el mantenimiento del sistema de información.

Las principales características que definen a un sistema de almacenamiento de datos son:

- Capacidad.**

- Mide la cantidad de datos que puede almacenar el sistema de almacenamiento, y es medida en bytes (Gigabytes o Terabytes, habitualmente, aunque con el Big Data se manejan incluso Petabytes).

- Rendimiento.**

- Cómo de rápido y eficiente es el sistema de almacenamiento de datos.

- Fiabilidad.**

- Es la disponibilidad de los datos cuando son solicitados, así como el hecho de disponer de una baja tasa de errores o fallos (por ejemplo, utilizando una configuración RAID).

- Recuperabilidad.**

- Mide la capacidad del sistema para recuperar datos tras una pérdida, borrado, corrupción o cualquier otro incidente que impida el acceso a los mismos.

Qué dispositivos de almacenamiento de datos existen



Discos



Cintas magnéticas



Almacenamiento en red

SAN. SAN (*Storage Area Network*)

Almacenamiento en la nube



Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Introducción

Sistemas de información

Datos y tipos de datos

Ficheros

Bases de datos y SGBD

Diseño de bases de datos

Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Datos y tipos de datos

Un dato es la representación de una variable que puede ser cuantitativa o cualitativa que indica un valor que se le asigna a las cosas y se representa a través de una secuencia de símbolos, números o letras.

2012	Mayas
Rafa	Tenista
DAW/DAM	Especialidades
30,5	Temperatura
Xxx-ccc-ddd	isbn

Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Datos y tipos de datos

Entero.

Dato numérico que no cuenta con parte decimal.
Engloba todos los número positivos, negativos y el 0

Ejemplos:

0, 120 , -110, -42001, 21120, etc..

Real.

Dato formado por una variable numérica que puede contar con parte decimal.

Ejemplos:

0.25, 120.30 , -110, -42001.12, 21120, etc..

Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Datos y tipos de datos

Lógicos.

0,1, verdadero o falso (Boolean)

Texto

Carácter.

Tipo de dato formado por una unidad o símbolo que puede ser una letra, un número, una mayúscula o un signo de puntuación. (#, a, A, ?)

Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Datos y tipos de datos

Un dato de tipo estructurado está compuesto por una serie de datos que guardan alguna relación entre ellos.

- Homogéneas: todos los datos que lo componen son del mismo tipo.
- Heterogéneos: los datos que lo componen son de diferente tipo.

Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

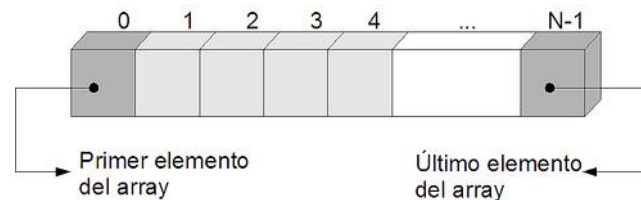
Estructuras de datos

Cadena o Strings.

Formado por un conjunto de caracteres dispuestos de forma consecutiva que se representa entre comillas. (“Hola”, “Adios”, “Star Wars”)

Un array.

Es un tipo de dato estructurado que permite almacenar un conjunto de datos homogéneo, es decir, todos ellos del mismo tipo y relacionados.



Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Estructuras de datos

Registro

También llamado fila o tupla representa un objeto único de **datos** implícitamente estructurados en una tabla. Un **registro** es un conjunto de campos **que** contienen los **datos que** pertenecen a una misma entidad.

Jugador	
Campos	Tipo
Nombre	String
Edad	Entero
Altura	Real

Registro

Nombre	Edad	Altura
Jhon	25	1.35

Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Estructuras de datos

Registro

También llamado fila o tupla representa un objeto único de **datos** implícitamente estructurados en una tabla. Un **registro** es un conjunto de campos **que** contienen los **datos que** pertenecen a una misma entidad.

Lista .- conjunto de registros

Jugador	
Campos	Tipo
Nombre	String
Edad	Entero
Altura	Real

Registro/fila

Nombre	Edad	Altura
Jhon	25	1.35

Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Estructuras de datos


Pilas

ES una estructura LIFO (Last In. First OUT)

El último elemento que entra es el primer elemento que sale.

Tipo de lista en que se añaden y eliminan los elementos sólo en uno de sus extremos.

	Nombre	Edad	Altura
Registro/fila	Jhon	25	1.35



Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

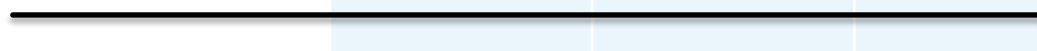
Estructuras de datos

Colas

First In. First OUT)

Una cola es una estructura de datos, caracterizada por ser una secuencia de elementos en la que la operación de inserción push se realiza por un extremo y la operación de extracción pull por el otro.

	Nombre	Edad	Altura
Registro/fila	Jhon	25	1.35

A horizontal line with an arrow pointing to the right, indicating the direction of data flow or processing in a queue.

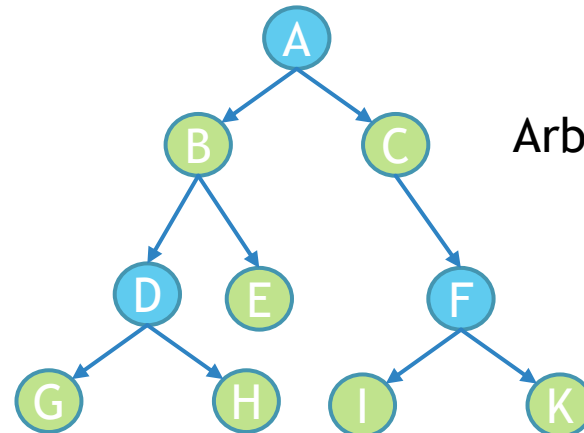
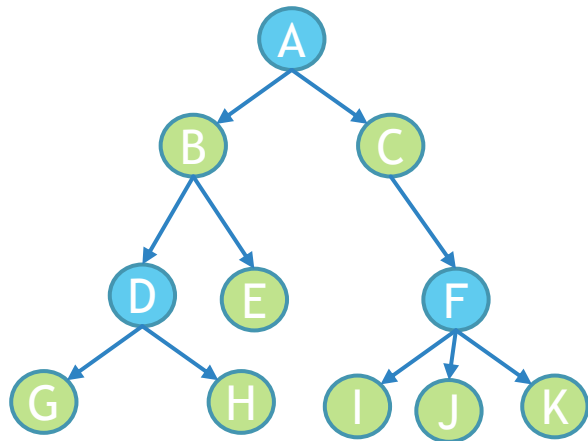
Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Estructuras de datos

Árbol

Los Árboles son las estructuras de **datos** mas utilizadas, pero también una de las mas complejas, Los Árboles se caracterizan por almacenar sus nodos en forma jerárquica y no en forma lineal como las Listas Ligadas, Colas, Pilas, etc



Arbol binario - máximo dos hijos

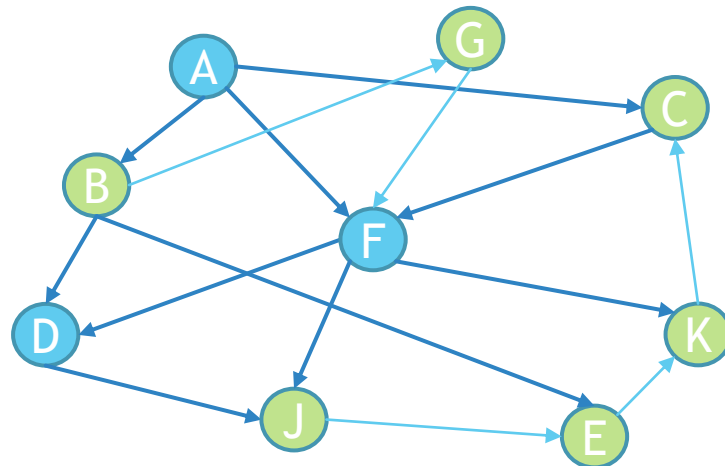
Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Estructuras de datos

Grafo

Estructura de datos que difiere de los árboles en que cada nodo puede tener varios nodos padre.



Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Introducción

Sistemas de información

Datos y tipos de datos

Ficheros

Bases de datos y SGBD

Diseño de bases de datos

Almacenamiento de la información

Primario

RAM, memorias caché

Secundario

Discos duros

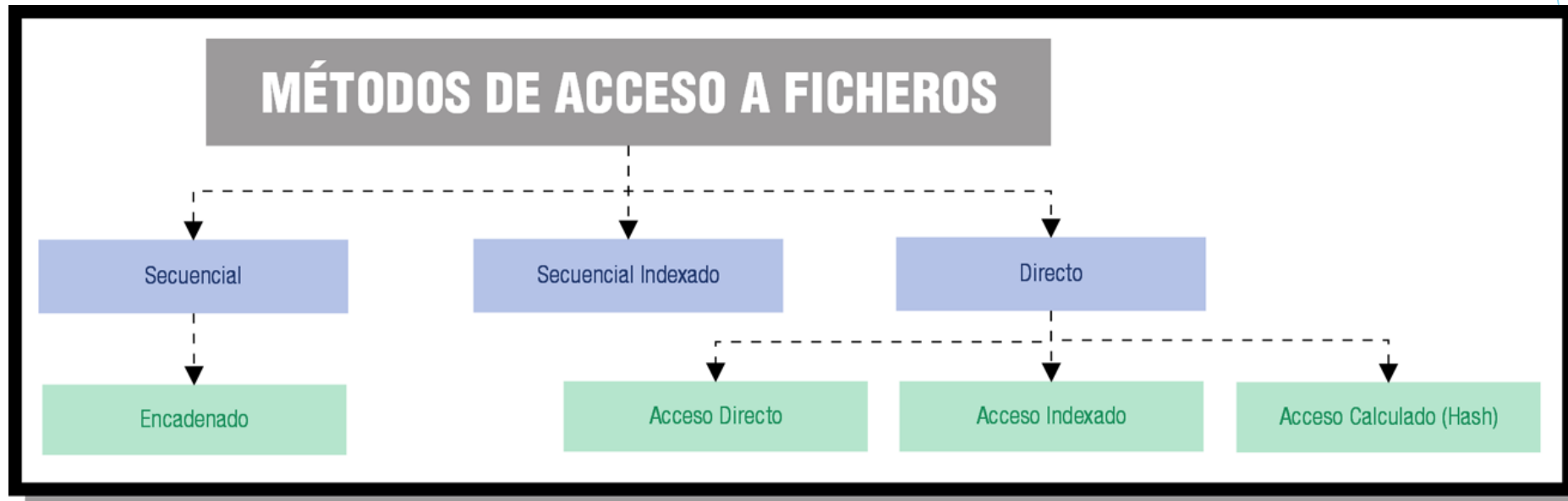
Intermedio

BUFFER

Ficheros

Conjunto de datos almacenados en un dispositivo.

- Residen en sistemas de almacenamiento secundario, memoria auxiliar, periférica o externa.
- Independencia de aplicación con que manejarlo. El mismo fichero se puede utilizar por diferentes programas.
- Gran capacidad de almacenamiento.

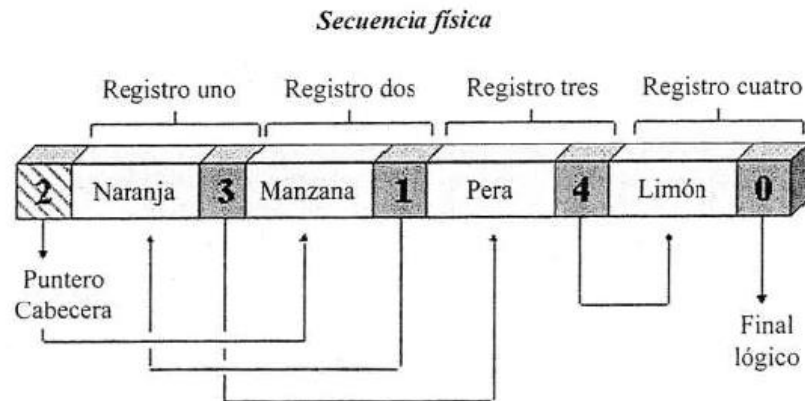


Fichero secuencial

Los registros están almacenados de forma contigua

Para acceder a un registro, es leyendo un registro tras otro desde el principio hasta el final

- La lectura siempre se realiza hacia delante.
- Son ficheros monousuario, no permiten el acceso simultáneo de varios usuarios.
- Tienen una estructura rígida de campos.
- Aprovechan al máximo el soporte de almacenamiento, al no dejar huecos vacíos.
- No se pueden insertar registros entre los que ya están grabados



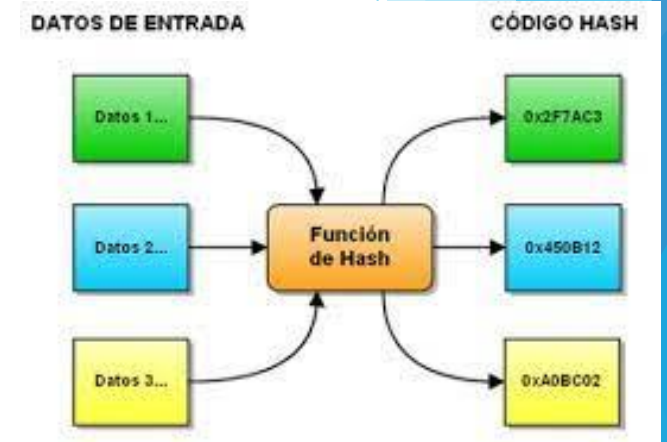
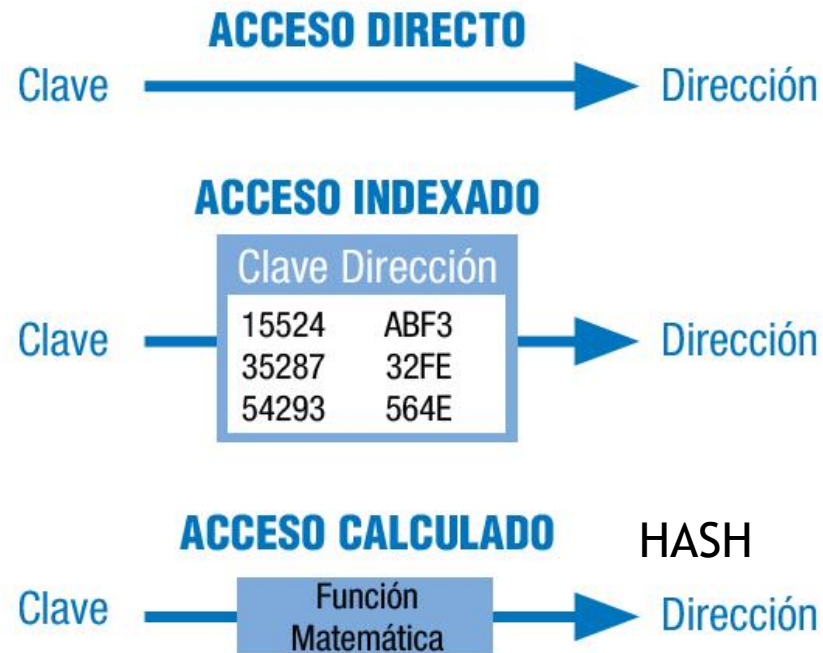
Fichero secuencial - indexado

Permite el acceso directo a los registros.

Fichero acceso directo

A los elementos de estos archivos se accede directamente
No se sitúan éstos en posiciones físicamente consecutivas

- Permiten la actualización de los registros en el mismo fichero, sin necesidad de copiar el fichero.
- Permiten realizar procesos de actualización en tiempo real.



Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Introducción

Sistemas de información

Datos y tipos de datos

Ficheros

Bases de datos y SGBD

Diseño de bases de datos

Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Bases de datos

Definición 1:

Una base de datos es un conjunto de datos relacionados y organizados con cierta estructura.

El modelo de bases de datos más extendido es el **relacional**.

Para su manipulación y gestión surgieron los sistemas gestores de bases de datos (**SGBD** en lo sucesivo).

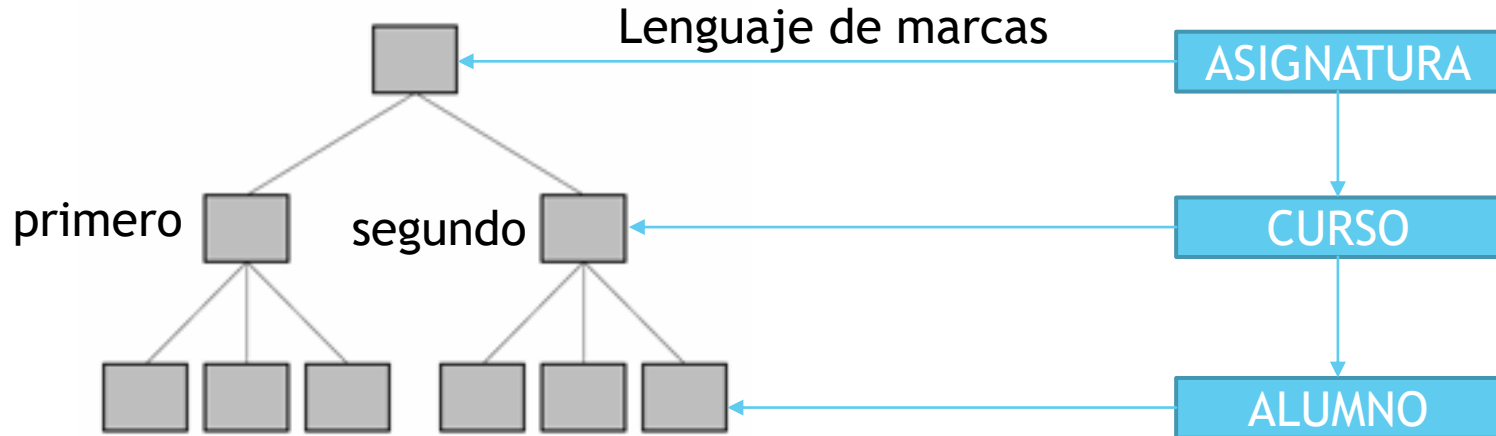


Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Bases de datos - Jerárquicas (*)

- Organiza los datos utilizando árboles.
- Cada nodo representa una entidad. A su vez, cada registro o segmento se compone por campos.
- Las relaciones entre entidades están representadas por las ramas

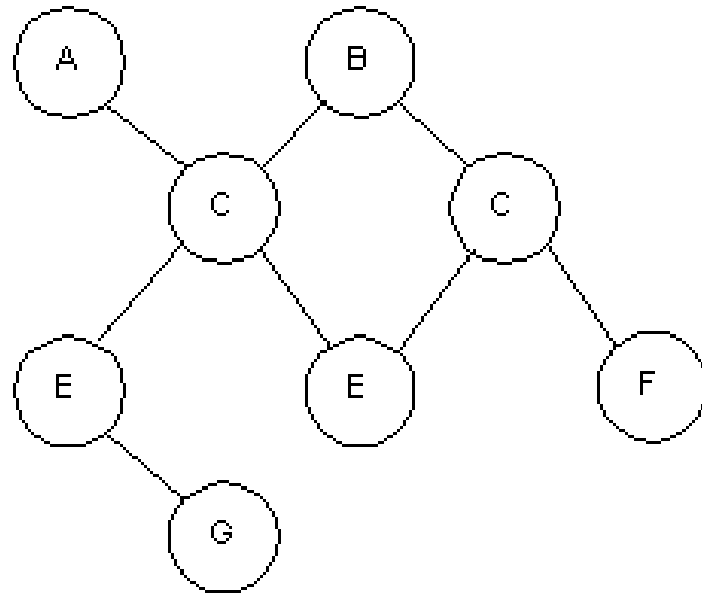


Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Bases de datos - Red (*)

- Intenta superar las deficiencias del enfoque jerárquico.
- Permite el tipo de relaciones de muchos a muchos, mediante la estructura de red o grafo.



Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Bases de datos - Relacional

- Todos los registros son del mismo tipo.
- La tabla sólo puede tener un tipo de registro.
- No existen campos o atributos repetidos.
- No existen registros duplicados.
- No existe orden en el almacenamiento de los registros.
- Cada registro o tupla es identificada por una clave que puede estar formada por uno o varios campos o atributos



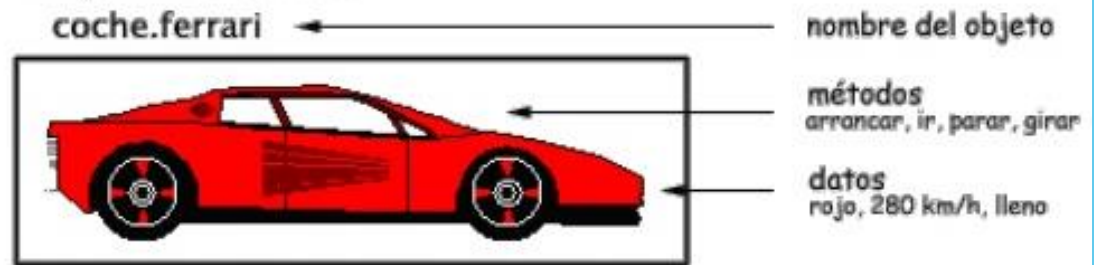
Unidad 1

Bases de datos - Orientada a objetos

Clase:
Coche



♦ Objeto: *Ferrari*



- Encapsulación** - Propiedad que permite ocultar la información al resto de los objetos, impidiendo así accesos incorrectos o conflictos.
- Herencia** - Propiedad a través de la cual los objetos heredan comportamiento dentro de una jerarquía de clases.
- Polimorfismo** - Propiedad de una operación mediante la cual puede ser aplicada a distintos tipos de objetos

Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Bases de datos - Orientada a objetos - Relacional



Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

Sistemas Gestores de Bases de Datos

- Definición y objetivos
- Funciones
- Componentes del SGBD
- Usuarios del SGBD
- Tipos de SGBD
- Sistemas gestores libres y comerciales

Bases de datos centralizadas y distribuidas

- Arquitectura de un DDBMS
- Técnicas de fragmentación, replicación y distribución

Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

SGBD

Definición 2:

El sistema de gestión de la base de datos (SGBD) es una aplicación que permite a los usuarios definir, crear y mantener bases de datos, proporcionando acceso controlado a las mismas.

Es una herramienta que sirve de interfaz entre el usuario y las bases de datos.



Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

SGBD

El sistema gestor de base de datos (SGBD) es el encargado de administrar y manipular la base de datos.

- Herramientas para administrar y crear la estructura física para almacenamiento de los datos.
- Herramientas para la manipulación de los datos (añadir, modificar, borrar o consultar datos).
- Herramientas de recuperación.
- Herramientas de copias de seguridad.
- Herramientas para la exportación e importación de datos.

ARQUITECTURA DE BASES DE DATOS

Nivel externo o vista de usuario

Independencia lógica

La *independencia lógica* es la capacidad de modificar el esquema conceptual sin tener que alterar los esquemas externos ni los programas de aplicación.

Nivel lógico o global

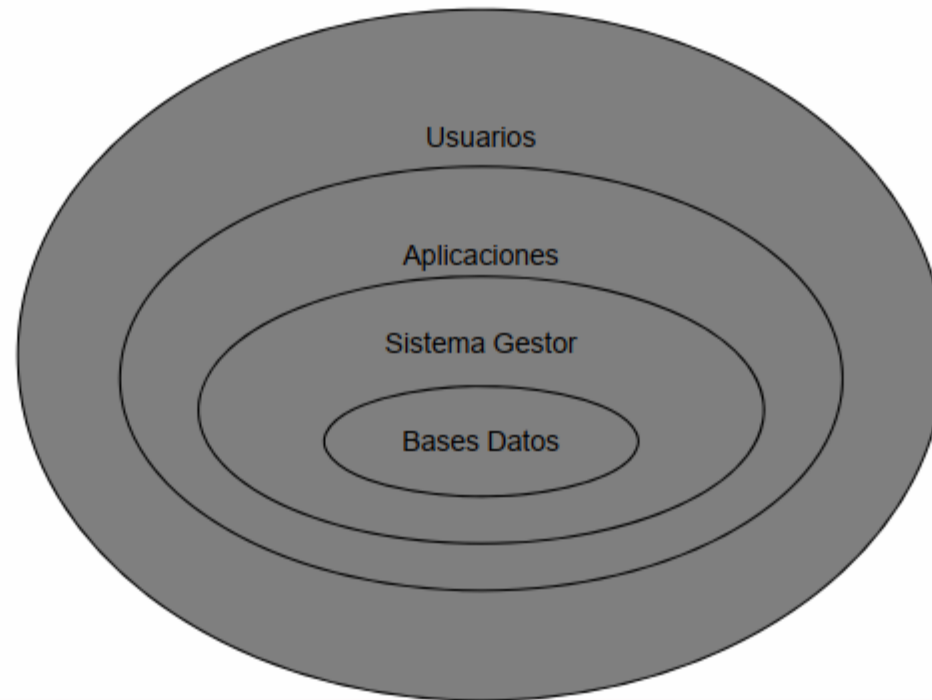
Independencia física

La *independencia física* es la capacidad de modificar el esquema interno sin tener que alterar el esquema conceptual (o los externos).

Nivel interno o físico

SISTEMAS GESTORES DE BASES DE DATOS I

Un sistema de gestión de la base de datos (SGBD) es una aplicación que permite a los usuarios definir, crear y mantener la base de datos, y proporciona acceso controlado a la misma. Es una herramienta que sirve de interfaz entre el usuario y las bases de datos.



Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

SGBD - Funciones

Por definición es importante indicar que una **función** es un conjunto de instrucciones que realizan una tarea específica de manera automática.

Las **funciones** fomentan la reutilización del código.

Una **función** acepta entradas en forma de parámetros y devuelve un valor y son las que permiten la creación, modificación y depuración de los **datos** dentro de una base de datos

- ❑ Incorporar catálogo
- ❑ Integridad
- ❑ Concurrencia de accesos
- ❑ Recuperación de datos
- ❑ Integración

SISTEMAS GESTORES DE BASES DE DATOS III

Componentes

Lenguaje de definición de datos (ddl: *Data Definition Language*)

Sencillo lenguaje artificial para definir y describir los objetos de la base de datos, su estructura, relaciones y restricciones.

Diccionario de datos

Esquemas que describen el contenido del **sGbd** incluyendo los distintos objetos con sus propiedades.

Objetos: tablas base y vistas (tablas derivadas)

- Consultas.
- Dominios y tipos definidos de datos.
- Restricciones de tabla y dominio y aserciones.
- Funciones y procedimientos almacenados.
- Disparadores o *triggers*.

Optimizador de consultas

Para determinar la estrategia óptima para la ejecución de las consultas.

Planificador (*scheduler*)

Para programar y automatizar la realización de ciertas operaciones y procesos.

Copias de seguridad

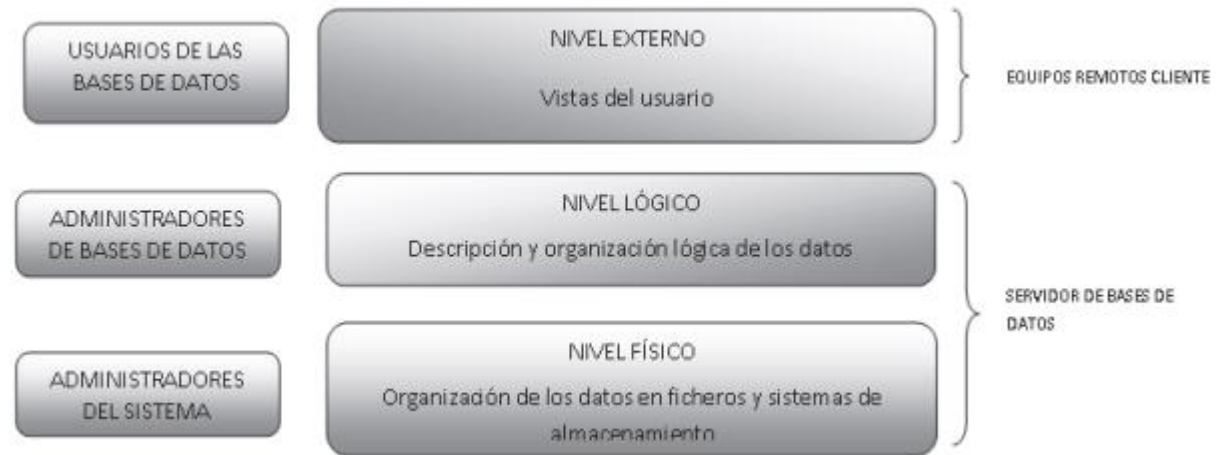
Para garantizar que la base de datos se puede devolver a un estado consistente en caso de que se produzca algún fallo o error grave.

Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

SGBD - Usuarios

- ❑ Administradores
- ❑ Diseñadores
- ❑ Programadores
- ❑ Usuarios finales



SISTEMAS GESTORES DE BASES DE DATOS V

Tipos de SGB

❑ Según modelo lógico:

- Jerárquico
- En red
- Relacional
- Objeto-relacional
- Orientado a objetos

❑ Según número de sitios:

- Centralizado
- Distribuido

❑ Según tipo de datos:

- Relacionales
- XML
- Objeto-relacionales
- Orientados a objetos

❑ Según número de usuarios:

- Monousuario
- Multiusuario

❑ Según Ámbito de aplicación

- Propósito general
- Propósito específico

❑ Según lenguajes soportados:

- SQL
- NoSQL

Unidad 1

Sistemas de Almacenamiento de la Información y SGBD

SGBD - Libres y comerciales

Libres

Disponibles de forma abierta en la red (no necesariamente gratuitos) y modificables según licencia GPL

- MySQL
- PostgreSQL
- LibreOffice Base
- Redis
- MongoDB
- Apache couchDB

Comerciales

Requieren pago de licencias aunque suelen ofrecer versiones gratuitas limitadas

- ORACLE
- SQL server
- Microsoft Access

Bases de datos centralizadas y distribuidas

Ventajas

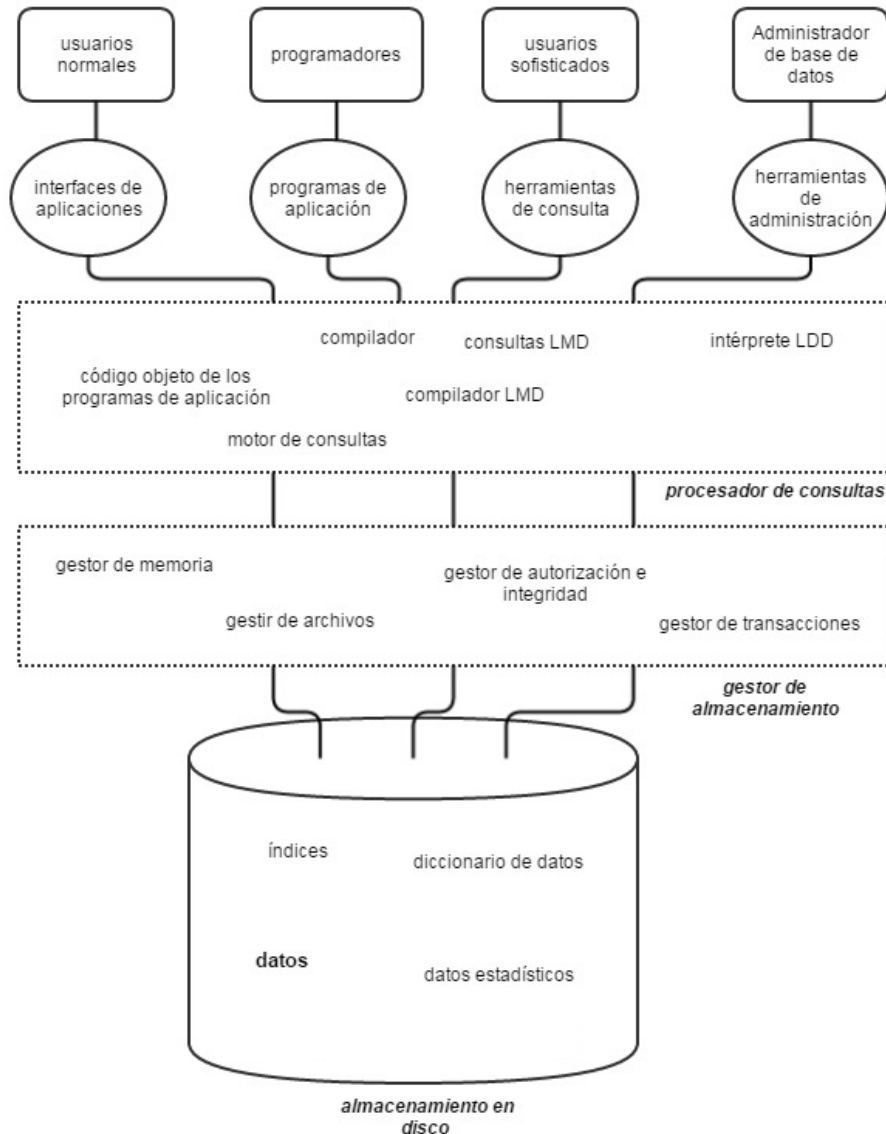
- ❑ Mejora de rendimiento
- ❑ Fiabilidad
- ❑ Disponibilidad
- ❑ Aplicaciones

Bases de datos centralizadas y distribuidas

Inconvenientes

- ❑ Necesidad de acceso remoto
- ❑ Mantenimiento catálogo
- ❑ Optimización de consultas
- ❑ Integridad de permisos
- ❑ Consistencia de copias replicadas
- ❑ Garantizar recuperación sistema

Arquitectura de un RDBMS



Almacenamiento en disco

Son los datos en sí, el diccionario de datos que contiene información sobre los datos.

Gestor de almacenamiento

El primer grupo de programas se encarga de gestionar el almacenamiento.

Procesador de consultas

El otro grupo de programas se encarga de las consultas.

LDD (Lenguaje de Definición de Datos) que interpreta las instrucciones del LDD y actualiza el diccionario de datos.

LMD (Lenguaje de manipulación de datos) que traduce las instrucciones del LMD que están en un lenguaje de consultas.

Usuarios

Normales.-sin conocimientos de base de datos

Programadores.- desarrollan aplicaciones para interactuar con la BBDD

Sofisticados.- utilizan la BBDD a través de consultas

Administrador.- se encarga de crear, modificar y autorizar (BBDD)

Fragmentación, replicación y distribución de datos

❑ Fragmentación

- Horizontal: separamos filas
- Vertical: separamos campos

Fragmentación, replicación y distribución de datos

□ Replicación

- Facilita la distribución de la carga
- Mejora la disponibilidad
- Sirve de copia de seguridad