Gestión de Bases de Datos

Unidad 1:

Sistemas de almacenamiento de la información

UNIDAD 1: SISTEMAS DE ALMACENAMIENTO DE LA INFORMACIÓN

- 1. Ficheros
 - 1.1 Tipos
 - 1.2 Inconvenientes
 - 1.3 Actividad 1
- 2. Bases de datos
 - 2.1 Conceptos
 - 2.2 Usos y tipos según el modelo de datos
 - 2.3 Ubicación de la información
 - 2.4 Usuarios
 - 2.5 Ventajas
 - 2.6 Actividad 2
- 3. Sistemas gestores de base de datos
 - 3.1 Funciones
 - **3.2 Componentes**
 - 3.3 Actividad 3
- 4. Bases de datos centralizadas y bases de datos distribuidas
- 5. Modelo de capas

1. Ficheros.

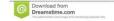
Un ordenador maneja y almacena una gran cantidad de información.

La información se almacena en dispositivos como discos duros, pendrives, DVDs, etc.

¿Qué características de los dispositivos de memoria han de tener necesariamente los sistemas de almacenamiento?

Esta información almacenada debe estar adecuadamente organizada.

Para poder organizar la información se utilizan los ficheros o archivos.



Concepto de fichero:

Es una estructura de información que crea y gestiona el sistema operativo de un ordenador, para almacenar información.

Los ficheros tienen un nombre y normalmente una extensión. La extensión determina el formato de la información almacenada

en el fichero.

TXT: texto

EXE: Ejecutable

BMP: Mapa de bits o imagen

MP3: Audio

AVI: Vídeo



¿Qué extensiones de archivo poco comunes conoces? ¿Qué ocurre si un archivo no tiene extensión? ¿Y si su extensión es la equivocada?

1.1 Ficheros. Tipos

Tipos de ficheros según la representación interna del contenido

□Ficheros de texto

También se denominan **ficheros planos.** Representan texto, números y otros caracteres. Almacenan cada carácter en un sistema de codificación de texto como ASCII o UNICODE. Son directamente legibles por programas editores de texto.

¿Qué editores de texto conoces que utilicen ficheros planos? ¿Qué formato de codificación es el más utilizado actualmente?

□Ficheros binarios

Usan sistemas de codificación diferentes a los de texto.

Almacenan información de muy diversos tipos.

Texto con formato, hojas de cálculo, presentaciones.

Imágenes, vídeo, audio.

Ficheros comprimidos

Ficheros ejecutables

Busca en tu ordenador un fichero docx y prueba a abrirlo con el bloc de notas, usando el menú contextual. Observa el resultado

1.1 Ficheros. Tipos

Tipos de ficheros según la organización o la forma de acceder a los datos:

□Secuenciales

Los datos se van almacenando uno tras otro, no admite inserciones. Para acceder a un dato hay que acceder previamente a todos los anteriores.

¿Qué soportes de almacenamiento secuencial conoces?

□De organización directa

Se accede a un registro o elemento del fichero de forma directa a partir de un valor, por ejemplo, a partir del número de un alumno accedemos directamente a la información del alumnos

□Ficheros indexados

Hay un fichero de índices secuencial desde el que se accede directamente al fichero con todos los datos

1.2 Ficheros. Inconvenientes

Problemas de usar sistemas de almacenamiento basados en ficheros.

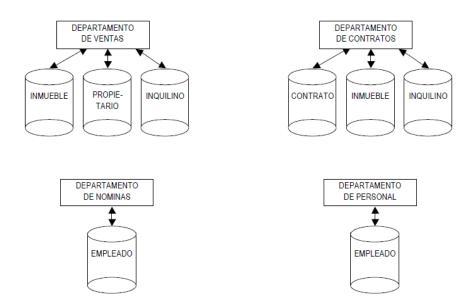
Datos redundantes.

Distintos formatos para los mismos datos.

Programas dependientes de la estructura física de los ficheros de datos.

Se pueden llegar a tener numerosos programas de aplicación para la manipulación de los ficheros de una organización.

Sistema de ficheros de una empresa INMOBILIARIA



1.3 Ficheros. Actividad 1

Actividad 1.

Agrupados en grupos de 4 miembros, investigad los diferentes sistemas de archivo utilizados en los principales sistemas operativos. Realizad una presentación en la que se detallen los siguientes puntos:

- 1. ¿Qué es un sistema de archivo?
- 2. ¿Cuales son sus principales funciones? Explicadlas.
- 3. ¿Qué tipos de sistemas de archivo existen?

Cada grupo deberá elegir un sistema de archivo (por sistema operativo) diferente (FAT, HPFS, NTFS, EXT, XFS, JFS, HFS, APFS) y responder a las siguientes cuestiones:

- 4. ¿Cuáles son sus principales características?
- 5. ¿En qué sistemas operativos es utilizado?
- 6. ¿Qué características lo diferencian del resto?
- 7. ¿Cuáles son sus principales ventajas?
- 8. ¿Cuáles son sus mayores inconvenientes?
- 9. Haced una comparación con el sistema de archivos exFAT

2. Bases de Datos

Base de Datos: Es una colección de datos adecuadamente organizados, que constituyen información relevante para la administración o gestión de una empresa, organismo, asociación, etc., facilitando las tareas de administración y manipulación de esos datos.

¿Qué bases de datos utilizas a diario (aún sin saberlo)? Busca en Internet las mayores bases de datos.

Una base de datos se organiza en Tablas que se relacionan entre si para que la información esté almacenada de forma ordenada y coherente.

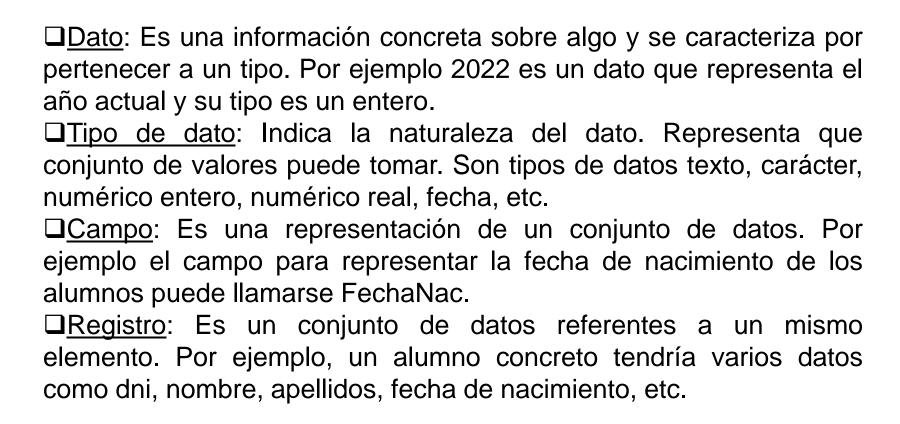
9

2. Bases de Datos

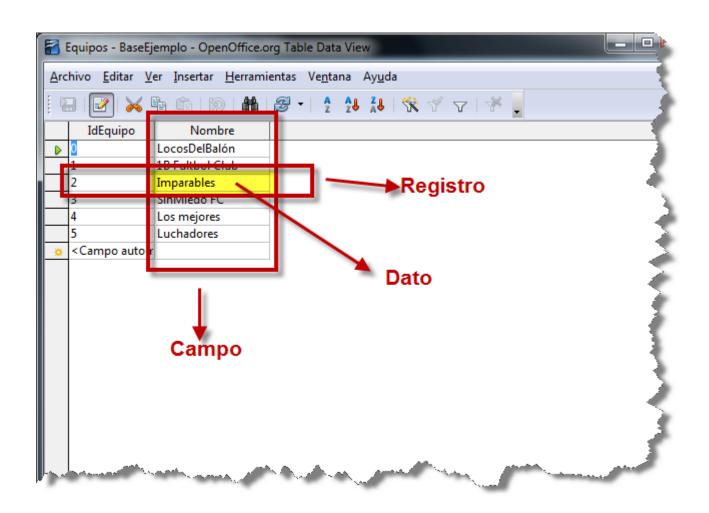
Algunos ejemplos de bases de datos pueden ser: □Base de datos para gestionar una empresa de alquiler de coches. Esta base de datos se dividirá en varias unidades informativas como coches, clientes, reservas, contratos, empleados, mantenimiento, etc. □Base de datos para gestionar una **biblioteca**. Esta base de datos se dividirá en unidades informativas para almacenar libros, socios, préstamos, reservas, pedidos de libros, compras de libros, etc. □Base de Datos para la **gestión de centro escolares en la** Consejería de Educación que se dividirá en unidades informativas para almacenar información relativa a Centros, profesorado, personal de administración y mantenimiento, directores, alumnos, consejos escolares, niveles educativos, medios audiovisuales,

¿Qué unidades informativas contendrá una base de datos de una empresa de desarrollo de aplicaciones multiplataforma? 10

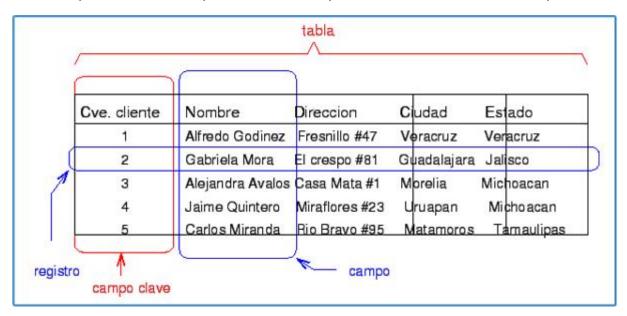
estadísticas de evaluación, etc.



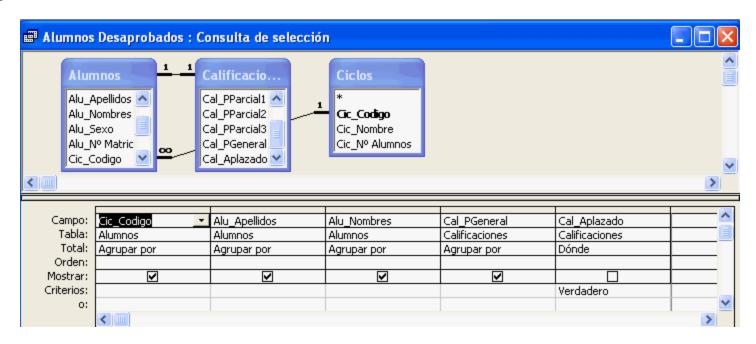
¿Qué campos y tipos de datos podrían contener las unidades informativas de la empresa de desarrollo de aplicaciones multiplataforma?



□<u>Tabla</u>: Es un conjunto de registros representado con un nombre que contiene toda la información de una parte del sistema de información. Una base de datos se organiza en tablas. Por ejemplo, una base de datos de un centro de estudios podría tener las tablas **profesores**, alumnos, módulos, matriculaciones, etc.

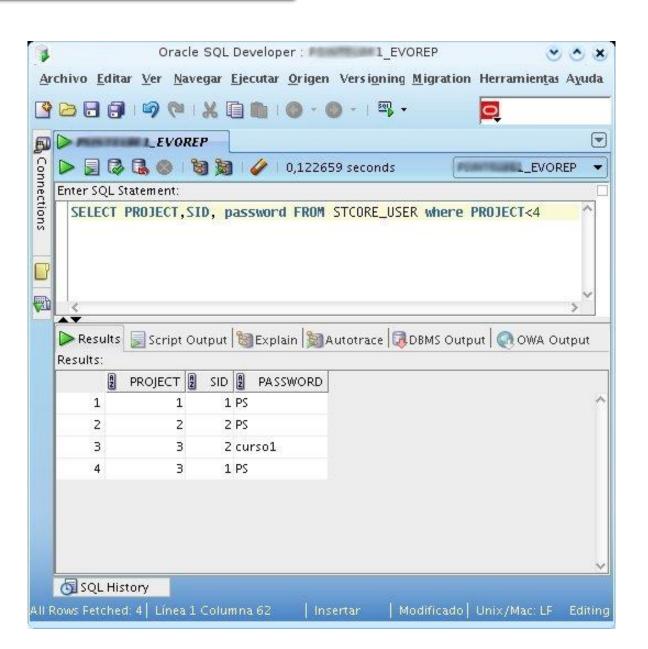


□Consulta: Se trata de una instrucción para hacer peticiones de datos a una base de datos para que se haga una búsqueda en la base de datos de los registros que cumplen con las condiciones expresadas en la instrucción.



□¿Cuáles serán las principales partes de una consulta?

□Consulta



□Índice: Es una estructura que almacena los campos clave de una tabla para que sea más rápido encontrar y ordenar los registros de la tabla.
□Vista: Es una transformación que se hace de una o más tablas para obtener una tabla que será visible para determinados usuarios. Esta tabla es virtual, no permanece almacenada. Por ejemplo, las direcciones de entrega en una tienda online podrán recogerse en una vista que contenga los datos personales del cliente (Tabla CLIENTE) y su dirección (Tabla DIRECCIÓN).
□Informe: Es un documento que se genera como resultado de una consulta a la base de datos y que es fácilmente legible para los usuarios.
□ Guiones o scripts: Son conjuntos de instrucciones que realizan

Por ejemplo, un script que autogenera el identificador de un objeto al insertarse en la tabla.

operaciones avanzadas de mantenimiento de los datos.

Diccionario de datos o esquema: Es la definición de la estructura donde se almacenan los datos. Contiene todo lo necesario para organizar la información mediante tablas, registros y campos o columnas.

Es uno de los primeros pasos al crear una base de datos, definir su esquema.

- □Por ejemplo en MySQL hay una base de datos llamada information_schema que contiene una serie de tablas que sirven para definir la estructura de todas las bases de datos que se gestionen en MySQL.
- ☐ Tiene varias tablas como TABLES, ROUTINES, TRIGGERS, USER_PRIVILEGES, etc.

2.2 Usos y tipos según el modelo de datos

Las bases de datos se utilizan en diversos sistemas de información. Algunos de los usos más frecuentes de las bases de datos son:

Bases de datos administrativas de una empresa. Relacionan clientes, pedidos, ventas, facturas, productos, etc.

Bases de datos contables: Relacionan los datos contables de una empresa, entre ellos, pagos, ingresos, balances, declaraciones de IVA, etc.

Bases de datos para motores de búsqueda: Por ejemplo Google dispone de un gigantesca y distribuida.

Bases de datos científicas: Normalmente también almacenan grandes cantidades de datos como datos climáticos.

Otros usos son configuraciones de software, bibliotecas, censos, antivirus, deportes, etc.

2.2 Usos y tipos según el modelo de datos

Tipos de bases de datos según el modelo de datos:

- □ Bases de datos jerárquicas y en red (obsoletas). Tienen estructura de árbol.
- □ Bases de datos relacionales. Se basan en el algebra relacional. Organizan la información en tablas y relaciones entre estas. Son las más usadas. Sobre estas bases de datos se usa el lenguaje SQL para manipularlas.
- □ Bases de datos orientadas a objetos u objeto-relacionales. Añaden funcionalidad a las relacionales permitiendo el procesamiento mediante programas orientados a objetos.
- □ Las bases de datos documentales están diseñadas para almacenar datos semiestructurados, como documentos, normalmente en formato JSON o XML. Permiten expresar los datos en su forma natural

2.3 Bases de Datos. Ubicación

La ubicación de una base de datos respecto de los usuarios que acceden a ella puede ser:

Local: El acceso se produce en el mismo equipo que almacena la base de datos.

En Servidor: La base de datos se almacena en un ordenador servidor. Los usuarios acceden desde otros ordenadores conectados dentro de un red local o a través de Internet.

Distribuidas: La base de datos, por su gran tamaño o por necesidad de tenerlos distribuidos geográficamente se almacena en varios equipos distintos. Se almacenan partes de la base de datos en cada equipo, no la base de datos completa en todos los equipos.

Reflejadas: Cuando la misma base de datos se almacenan en varios equipos por seguridad.

2.3 Bases de Datos. Usuarios

Funciones y características de cada uno de los usuarios:

Administrador: Se encarga de la creación o implementación física de la base de datos. Establece la política de seguridad y de acceso. Toma las decisiones relacionadas con el funcionamiento físico del almacenamiento de información.

Diseñador: Identifica los datos, las relaciones entre ellos, sus restricciones, etc. Se encarga de diseñar cómo será la base de datos.

Programador: Se encarga de implementar los programas de aplicación que servirán a los usuarios finales. Estos programas de aplicación ofrecerán la posibilidad de realizar consultas de datos

Usuario final: Es el cliente final de la base de datos.

2.4 Bases de Datos. Ventajas

Ventajas fundamentales que ofrece el uso de bases de datos:

- Acceso múltiple: diversos usuarios o aplicaciones podrán acceder a la base de datos, sin problemas de acceso.
- **Utilización múltiple**: cada uno de los usuarios o aplicaciones podrá acceder sólo a la parte le corresponde.
- Confidencialidad y seguridad: el control del acceso a los datos podrá ser establecido para que unos usuarios o aplicaciones puedan acceder a unos datos y a otros no, impidiendo a los usuarios no autorizados la utilización de la base de datos.
- Protección contra fallos: en caso de errores en la información, existen mecanismos bien definidos que permiten la recuperación de los datos de forma fiable.
- Independencia física: un cambio de soporte físico de los datos (por ejemplo: el tipo de discos), no afectaría a la base de datos o a las aplicaciones que acceden a ellos.
- Independencia lógica: los cambios realizados en la base de datos no afectan a las aplicaciones que la usan.
- Redundancia: los datos se almacenan, por lo general, una única vez. Aunque si es necesario, podríamos repetir información de manera controlada.

Actividad 2.

Investiga acerca de las bases de datos SQL y NoSQL y responde a las siguientes preguntas:

- 1. ¿Qué es una base de datos SQL? ¿Y una NoSQL?
- 2. ¿Qué ventajas e inconvenientes presentan cada una de ellas?

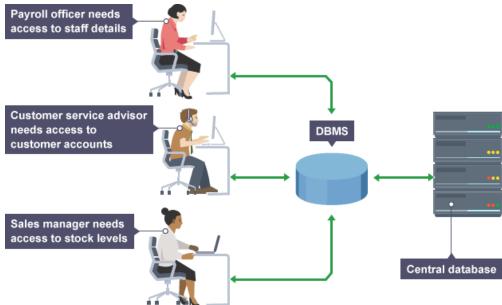
Decide qué tipo de base de datos elegirías en los siguientes casos y por qué?

- Un blog personal.
- Un gestor documental.
- Un videojuego en línea.
- Una entidad bancaria.

<u>Sistema Gestor de Bases de Datos (SGBD o DBMS):</u> Es un conjunto de datos interrelacionados y un conjunto de programas y recursos para administrar y gestionar esos datos y para que otros programas puedan usar esos datos.

En definitiva, el sistema gestor de base de datos es el conjunto de herramientas que nos permiten crear y gestionar bases de

datos.



3 Sistemas Gestores de Bases de Datos



















3.1 Funciones de los Sistemas Gestores de Bases de Datos

<u>Funciones de un SGBD (no todos las tienen todas):</u>

- Permiten a los usuarios almacenar datos, acceder a ellos y actualizarlos de forma sencilla.
- ▶ Garantizan la integridad de los datos.
- Disponen de un sistema de seguridad que garantiza el acceso a la información a los usuarios autorizados.
- Proporcionan un diccionario de metadatos.
- Ofrecen conectividad con el exterior.
- ▶Permiten la concurrencia.
- ▶Ofrecen estadísticas sobre el uso del gestor, registran operaciones efectuadas, consultas solicitadas, operaciones fallidas y cualquier tipo de incidencia.
- Independizan los datos de la aplicación o usuario que esté utilizándolos.
- Incorporan herramientas para la salvaguarda y restauración de los datos en caso de desastre.
- ▶Permiten el uso de transacciones. ¿Qué es una transacción?

3.1 Funciones de los Sistemas Gestores de Bases de Datos

Transacción:

Es un conjunto de operaciones a realizar sobre una estructura de datos compleja, que se deben ejecutar una tras otra. La transacción debe realizarse por completo, antes de que se permita la interacción con la estructura de datos implicada. En caso de error, todas las operaciones de la transacción se revierten, manteniendo la estructura de datos en su estado inicial, previo a la transacción.

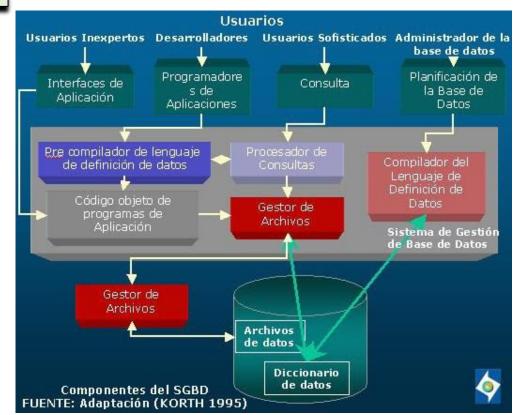
Las transacciones deben cumplir cuatro propiedades, denominadas ACID.

- Atomicidad (Atomicity): es la propiedad que asegura que la operación se ha realizado completamente.
- ▶ Consistencia (Consistency): es la propiedad que asegura que sólo se inicia aquellas transacciones que se pueden terminar.
- Aislamiento (Isolation): es la propiedad que asegura que una operación no afecta a otras ajenas a la transacción. Asegurando que la realización de dos transacciones sobre la misma estructura de datos no generará errores.
- Durabilidad (Durability): es la propiedad que asegura que una vez acabada la transacción, ésta persistirá ante un fallo del sistema.

3.2 Componentes de los SGBD

Componentes:

- ▶ Programas gestores:
 - Atienden consultas y buscan archivos.
- Diccionario de datos:
 - Contiene información para los analistas que usan en diseño. Flujos de datos, tamaños...
- ▶ Recursos humanos: Los usuarios, gestores, diseñadores, etc.
- Lenguaje de interacción con las bases de datos: sql, cql, oql, Xquery...



3.2 Componentes de los SGBD

Recursos Humanos:

- Administradores: Son los encargados del control central del SGDB y de las bases de datos gestionadas. Sus tareas son:
 - Definir esquema lógico de la base de datos.
 - Definir el esquema físico de la base de datos, estructuras y modos de almacenamiento, métodos de acceso a los datos.
 - Definir los subesquemas o vistas que de la base de datos va a tener cada usuario.
 - Establecer privilegios y permisos sobre los usuarios.
 - Establecer procedimientos de recuperación de datos en casos de fallo del sistema.
- Usuarios
- Diseñadores
- Analistas y Programadores

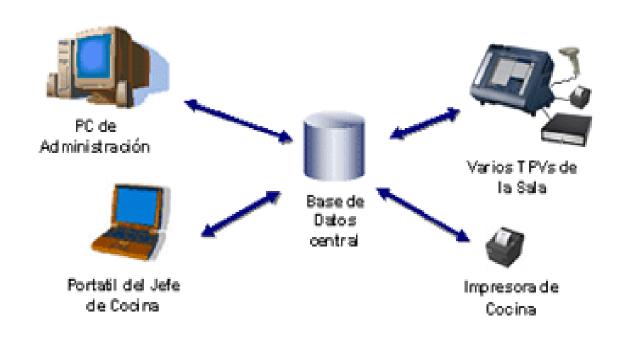
Lenguajes de interacción con la base de datos:

El lenguaje SQL es un estándar como lenguaje de interacción con las bases de datos relacionales. Se puede dividir en:

- ▶ **Lenguaje DDL** (de descripción de datos): Permite definir la estructura de la base de datos (tablas, relaciones entre tablas, tipos de datos de las tablas, etc.).
- ▶ **Lenguaje DML** (de manipulación de datos): Para realizar consultas, inserciones, borrados y modificaciones de los datos.
- ▶ Lenguaje DCL (de control de datos): Se usa para controlar la seguridad, integridad y privacidad de las bases de datos (crear usuarios, definir privilegios de usuario, dar permisos sobre tablas, establecer mecanismos de recuperación de la información, etc.).
- ▶ **Lenguaje TCL** (de control de transacciones): Permite el uso de transacciones y su control.

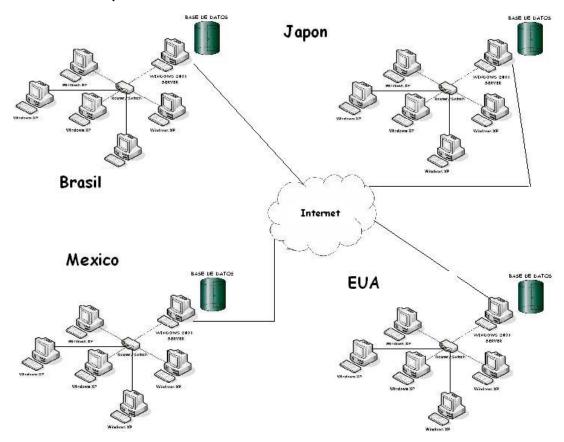
4 Bases de datos centralizadas y distribuidas

Una **base de datos centralizada** es una base de datos almacenada en su totalidad en un solo ordenador servidor. Eso no impide que se pueda acceder a esa base de datos desde otros equipos y que se pueda realizar acceso concurrente.



4 Bases de datos centralizadas y distribuidas

Una base de datos distribuida es una base de datos en la que los datos se encuentran distribuidos entre diferentes sitios interconectados por una red de comunicaciones.



Ventajas e inconvenientes de bases de datos centralizadas y distribuidas.

4 Bases de datos centralizadas

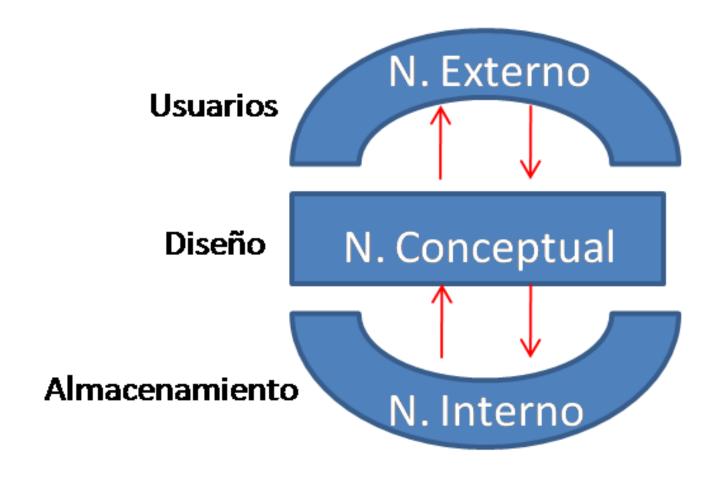
VENTAJAS	INCONVENIENTES
Se evita la redundancia debido a la posibilidad de inconsistencias y al desperdicio de espacio.	Un mainframe en comparación de un sistema distribuido no tiene mayor poder de computo.
Se evita la inconsistencia. Ya que si un hecho específico se representa por una sola entrada, la no-concordancia de datos no puede ocurrir.	Cuando un sistema de bases de datos centralizado falla, se pierde toda disponibilidad de procesamiento y sobre todo de información confiada al sistema.
La seguridad se centraliza.	En caso de un desastre o catástrofe, la recuperación es difícil de sincronizar.
Puede conservarse la integridad.	Las cargas de trabajo no se pueden difundir entre varias computadoras, ya que los trabajos siempre se ejecutarán en la misma máquina.
El procesamiento de los datos ofrece un mejor rendimiento.	Los departamentos de sistemas retienen el control de toda la organización.
iviantenimiento mas parato, ivietor uso de los recursos	Los sistemas centralizados requieren un mantenimiento central de datos.

4 Bases de datos distribuidas

VENTAJAS	INCONVENIENTES
El acceso y procesamiento de los datos es más rápido ya que varios nodos comparten carga de trabajo.	La probabilidad de violaciones de seguridad es creciente si no se toman las precauciones debidas.
Desde una ubicación puede accederse a información alojada en diferentes lugares.	Existe una complejidad añadida que es necesaria para garantizar la coordinación apropiada entre los nodos.
Los costes son inferiores a los de las bases centralizadas.	La inversión inicial es menor, pero el mantenimiento y control puede resultar costoso.
Existe cierta tolerancia a fallos. Mediante la replicación, si un nodo deja de funcionar el sistema completo no deja de funcionar.	Dado que los datos pueden estar replicados, el control de concurrencia y los mecanismos de recuperación son mucho más complejos que en un sistema centralizado.
El enfoque distribuido de las bases de datos se adapta más naturalmente a la estructura de las organizaciones. Permiten la incorporación de nodos de forma flexible y fácil.	El intercambio de mensajes y el cómputo adicional necesario para conseguir la coordinación entre los distintos nodos constituyen una forma de sobrecarga que no surge en los sistemas centralizados.
Aunque los nodos están interconectados, tienen independencia local.	Dada la complejidad del procesamiento entre nodos es difícil asegurar la corrección de los algoritmos, el funcionamiento correcto durante un fallo o la recuperación.

Actividad 3.

Con el aula dividida en dos grupos, cada grupo deberá reunirse y defender un modelo de bases de datos (centralizadas o distribuidas) y después desarrollar un debate para ver que equipo defiende mejor su modelo.



La arquitectura de una base datos se basa en tres niveles de visión de los aspectos de la base de datos por los usuarios:

Nivel externo o de vistas: Los usuarios no conocen los detalles de las tablas que hay en la base de datos, los campos y tipos que contienen. Ven únicamente los datos que se les permite.

Nivel lógico o conceptual: En este nivel el usuario ve todos los datos de la base de datos y los detalles del diseño (tablas, campos, relaciones).

Nivel físico interno: En este nivel, el usuario trabaja sobre detalles de como se almacenan realmente los datos en las unidades de almacenamiento y cuales son los sistemas de acceso. Decide donde se almacena la información, como se almacena, etc.

5 Modelo de capas

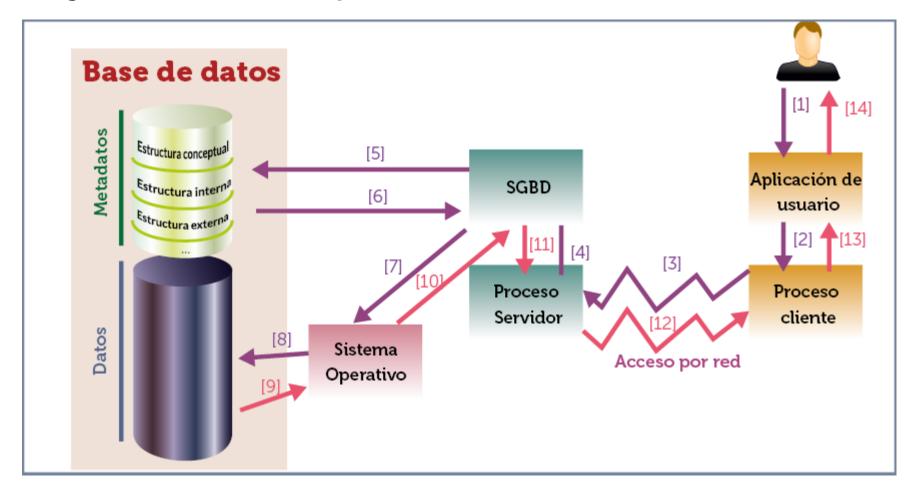
En base a esos tres niveles de visión de las bases de datos, los SGBD realizan todos los procesos operando conforme a un modelo de cinco capas:



5 Modelo de capas

- ☐ facilidades de usuario: Son las herramientas que proporciona el SGBD a los usuarios para permitir un acceso más sencillo a los datos.
- □ capa de acceso a datos: Traduce las peticiones del usuario para que lleguen de forma correcta a la base de datos y ésta pueda responder de forma adecuada.
- ☐ diccionario de datos: Las solicitudes de los clientes (que son conceptuales antes de llegar aquí) se traducen en instrucciones que hacen referencia al esquema interno de la base de datos.
- núcleo :El núcleo de la base de datos es la capa encargada de traducir todas las instrucciones requeridas y prepararlas para su correcta interpretación por parte del sistema operativo.
- □ sistema operativo: Es la capa que realmente accede a los datos. Realmente lo hace el Sistema Operativo y no el SGBD.

Esquema de procesamiento en una petición de usuario al SGBD según el modelo de capas:



5 Modelo de capas

Del Usuario al Sistema Gestor de Bases de Datos

- •[1] Los usuarios utilizan una aplicación para acceder a los datos. Estamos en el nivel externo de la base de datos, por lo que la propia aplicación traduce la petición que hizo el usuario de forma sencilla, a una petición entendible por la capa de acceso a los datos.
- •[2] El proceso cliente es el software de acceso a la base de datos y que está instalado en el lado del cliente. Se encarga simplemente de recoger y enviar la petición (comprobando antes si hay comunicación con el servidor de la base de datos).
- •[3] A través de la red (normalmente) el proceso cliente se comunica con el proceso servidor, que es el software de comunicación instalado en el lado del servidor. Ambos procesos (cliente y servidor) forman la capa de acceso a los datos.
- •[4] Estando ya en el servidor, la petición pasa al software del Sistema Gestor de Bases de Datos (habrá aquí, como se ha visto en el apartado anterior una traducción de datos, desde el nivel externo al nivel interno).
- •[5] El SGBD, comprobando el diccionario de datos, comprueba si la petición es correcta.
- •[6] El SGBD también revisa el diccionario de datos (si la petición es correcta) para saber con exactitud en qué archivos y en qué parte dentro de ellos, se encuentran los datos requeridos.

5 Modelo de capas

Petición al Sistema Operativo y recepción de los datos en el Sistema Gestor de BBDDs

- •[7] Con la información sobre dónde están los datos, el SGBD hace una petición al Sistema Operativo, que es el que tiene capacidad realmente de acceder a los archivos de datos. Por ello la petición del SGBD se traduce al formato utilizado por el Sistema Operativo.
- •[8] El Sistema Operativo accede a los datos.

Se Accede a la Base de Datos Físicamente

- •[9] El Sistema Operativo recibe los datos.
- •[10] Se entregan los datos al Sistema Gestor de Bases de Datos o, si ha habido un error al acceder a los datos, se indica el error ocurrido.

Del Sistema Gestor de Bases de Datos al usuario

- •[11] El SGBD traduce los datos a una forma más conceptual y se los entrega al proceso servidor.
- •[12] Los datos se entregan al proceso cliente.
- •[13] Los datos llegan a la aplicación.
- •[14] La aplicación de usuario traduce los datos recibidos en información presentada de la forma más conveniente para el usuario.