**TEMA 1 – ALMACENAMIENTO DE LA INFORMACIÓN**

1. **LOS FICHEROS DE INFORMACIÓN**
   1. **Fichero o Archivo**

Conjunto de información relacionada, tratada como un todo y organizada de forma estructurada, es una secuencia de dígitos binarios que organiza infomración relacionada como un mismo aspecto.

Estan formados por **registros lógicos** que contienen datos relativos a un mismo elemento y objeto.

Los registros estan divididos en **campos** que contienen cada una de las informaciones elementales que forman un registro.

Los datos están almacenados de tal forma que se puedean añadir, suprimir, actualizar o consultar individualmente en cualquier momento.

**Registro físico** o **Bloque**, una sola operación de lectura/grabación donde la cantidad de informacion que es transferida entre el sopoerte en que se almacena el fichero, y la memoria principal del ordenador.

**Factor de Blocaje**, el número de registros que entran en un bloque.

**Bloqueo de Registros**, operación de agrupar varios registros en un bloque.

* 1. **Tipos de Ficheros**
* **Ficheros Permanentes**, contienen informacion relevante para una aplicación, tienen un periodo de permanencia en el sistema amplio.
  + **Ficheros Maestros**, contienen el estado actual de los datos que pueden modificarse desde la aplicación, la parte central de la aplicación, su núcleo.
  + **Ficheros Constantes**, incluyen datos fijos para la aplicación.
  + **Ficheros Históricos**, contienen datos que fuerón considerados como actuales en un periodo o situación anterior, se utilizan para la reconstrucción de situaciones.
* **Ficheros Temporales**, utilizados para almacenar información útil para una parte de la aplicación, no para toda ella, tienen un corto periodo de existencia. Generados a partir de los ficheros permanentes.
  + **Ficheros Intermedios**, almacenan resultados de una aplicación que serán utilizados por otra.
  + **Ficheros de Maniobras**, almacenan datos de una aplicación que no pueden ser mantenidos en memoria principal por falta de espacio.
  + **Ficheros de Resultados**, almacenan datos que van a ser transferidos a un dispositivo de salida.
  1. **Los Soportes de Información**

Soportes de información manejados por dispositivos periféricos del ordenador, que permitne leer y grabar datos en el soporte.

Tipos de soportes de almacenamiento de datos:

* **Soportes de Acceso Directo a los Datos**, los más empleados y el acceso a los datos puede hacerse de forma directa, pudiendo colocarnos en la posición que nos interesa y leer de ella.
* **Soportes de Acceso Secuencial**, suelen usarse en copias de seguridad y para leer un dato de una posición de la cinta habrá que leer todo hasta llegar a esa posición.
  1. **Métodos de Acceso**

Las distintas formas de organizar un fichero en un soporte de memoria .

Las organizaciones secuencial, de acceso directo y de acceso indexado son las más comunes.

* 1. **Ficheros Secuenciales**

Caracterizados porque sus registros están almacenados de forma contigua, la única forma de acceder a él, es leyendo un registro tras otro de principio a fin.

**EOF** (End of File), marca indicativa del fin del fichero.

**Clave** o **Llave**, información ubicada en uno de los campos de los registros almacenados, la cual lo identifica. Si ordenamos un archivo secuencial por su clave, es más rápido realizar cualquier operación de lectura o escritura.

Características relevantes de los ficheros secuenciales:

* La lectura siempre se realiza hacia delante.
* No permiten el acceso simultáneo de varios usuarios.
* Estructura rígida de campos. Todos los registros deben aparecer en orden, es decir, la posición de los campos de cada registro siempre ha de ser la misma.
* El modo de apertura del fichero, condiciona la lectura o escritura.
* Aprovechan al máximo el soporte de almacenamiento, no dejan huecos vacios.
* Se pueden grabar en cualquier tipo de soporte, tanto secuenciales como direccionales.
* Todos los lenguajes tienen instrucciones para trabajar con ellos.
* No se pueden insertar registros entre los que están grabados.
  1. **Ficheros de Acceso Directo**

En estos se puede acceder a un registro indicado la posición relativa del mismo dentro del archivo, o a través de una clave que forma parte del registro como un campo más, estos deben almacenarse en dispositivos de memoria masiva de acceso directo, como son los disco magnéticos.

**Campo Clave**, permite identificar y localizar un registro de manera ágil y organizada.

Cada registro se guarda en una posición física, que depende del espacio disponible en memoria masiva, de ahí que la distribución de los registros sea aleatoria dentro del soporte. Para acceder a la posición física de un registro se utiliza una dirección o índice, no siendo necesario recorrer todo el fichero para encontrar un determinado registro.

En el acceso directo la clave coincide con la dirección, debiendo ser numérica y comprendida dentro del rango de valores de las direcciones.

Características fundamentales de estos ficheros:

* Posicionamiento inmediato.
* Registros de longitud fija.
* Apertura del fichero en modo mixto (lectura y escritura).
* Permiten múltiples usuarios.
* Los archivos se crean con un tamaño definido (con un máximo de registros establecido durante la creación).
* Solo es posible en soportes direccionales.
* Usados cuando el acceso a los datos de un registro se hace siempre empleando la misma clave y la velocidad de acceso a un registro es lo más importante.
* Permiten la actualización de los registros en el mismo fichero, sin necesidad de copiar el fichero.
* Permiten realizar procesos de actualización en tiempo real.
  1. **Ficheros Indexados**

Se basan en la utilización de **Índices**, permiten el acceso a un registro del fichero de forma directa, sin leer los anteriores.

**Zona de Registros**, en la que estan los datos del archivo.

**Zona de Índices**, tabla con las claves de los registros y las posiciones donde se encuentran.

La tabla de incices será cargada en memoria principal para realizar la búsqueda de la fila correspondiente a la clave del registro a encontrar, obteniéndose la dirección donde se encuentra el registro, ya solo hay que acceder a la zona de registros en el soporte de almacenamiento y posicionarnos en la dirección indicada.

Caracteristicas más relevantes:

* **Campo Clave**, campo o combinación de ellos de un registro, que permite identificar cada registro de forma única.
* Permite utilizar el modo de **acceso secuencial** y **acceso directo** para leer la información guardada en sus registros.
* Para acceder a los ficheros con el modo acceso directo se hace conociendo el contenido del campo clave del registro a localizar.
* Para acceder utilizando el acceso secuencial los registros son leídos ordenados por el contenido del campo clave.
* Solo se puede grabar en un soporte direccionable, si no, no se podria usar el acceso directo.
  1. **Otros Ficheros**

**Ficheros Secuenciales Indexados**

Parcialmente indexados, tiene una zona de indices y otra de registro de datos, pero esta se encuentra dividida en segmentos (bloques de registros) ordenados.

En la tabla de índices, cada fila referencia a cada segmento, la clave corresponde al último registro y el índice apunta al registro inicial. Una vez que se accede al primer registro del segmento, dentro de él se localiza secuencialmente el registro buscado.

Características:

* Permite el acceso secuencial.
* Permite el acceso secuencial, los registros se leen ordenados por el campo clave.
* Se puede actualizar los registros en el mismo fichero.
* Ocupa mas espacio que los ficheros secuenciales debido al uso del área de índices.
* Solo utiliza soportes direccionales.
* Obliga a una inversion económica mayor.

**Ficheros de Acceso Calculado o Hash**

Técnica del acceso calculado o hash, permite accesos más rápidos, ya que se usa una transformación o función matemática conocida, que a partir de la clave genera la dirección de cada registro del archivo.

**Colisión**¸ problema que a partir de diferentes claves se obtenga la misma dirección al aplicar la función matemática o transformación. **Sinónimos**.

Métodos de transformación:

* **Módulo**, la dirección será igual al resto de la división entera entre la clave y el número de registros.
* **Extracción**, la dirección será igual a una parte de las cifras que se extraen de la clave.
  1. **Parámetros de Utilización**

Mediante los parámetros de referencia se puede determinar el uso de un fichero, estos son:

* **Capacidad o Volumen**, el espacio en caracteres que ocupa el fichero, puede calcularse multipicando el número previsto de registros por la longitud media de cada registro.
* **Actividad**, permite conocer la cantidad de consultas y modifcaciones que se realizan en el fichero, para especificar la actividad se deben tener en cuenta:
  + **Tasa de consulta o modificación**, el porcentaje de registros consultados o modificados en cada tratamiento del fichero, respecto al número total de registros contenidos en él.
  + **Frecuencia de consulta o modificación**, número de veces que se accede al fichero para hacer una consulta o modificación en un periodo de tiempo fijo.
* **Volatilidad**, mide la cantidad de inserciones y borrados que se efectúan en un fichero. Es necesario conocer:
  + **Tasa de renovación**, % de registros renovados en cada tratamiento del fichero, respecto al nº total de registros contenidos en él.
  + **Frecuencia de renovación**, el nº de veces que se accede al fichero para renovarlo en un periodo de tiempo fijo.
* **Crecimiento**, la variación de la capacidad del fichero:
  + **Medido con la tasa de crecimiento**, el % de registros en que aumenta el fichero en cada tratamiento.

1. **BASES DE DATOS**
   1. **Conceptos**

**Base de Datos**, colección de datos relacionados lógicamente entre sí, con una definición y descripción comunes, estructurados de determinada manera. Es un conjunto estructurado de datos que representa entidades y sus interrelaciones, almacenados con la mínima redundancia y posibilitando el acceso a ellos eficientemente por parte de varias aplicaciones y usuarios.

**Metadatos**, descripcion de los datos contenidos en la Bda, se almacena en el **diccionario de datos** o **catálogo** y permite que exita **impedancia de datos** lógica-física.

**Elementos** de una base de datos:

* **Entidades**, objeto con características diferenciadoras de otros, del que se almacena información en la Bda.
* **Atributos**, los datos almacenados de la identidad, cualquier propiedad o característica.
* **Registros**, donde se almacena la información de cada entidad, un conjunto de atributos que contienen los datos que pertenecen a una misma repetición de entidad.
* **Campos**, donde se almacenan los atributos de cada registro.

**Ventajas fundamentales**:

* **Acceso múltiple**.
* **Utilización múltiple**.
* **Flexibilidad**, diferentes formas de acceder a la información.
* **Confidencialidad y seguridad**, control de acceso a los datos.
* **Proteccion contra fallos**, mecanimos de recuperación de datos.
* **Independiencia física**, el cambio de soporte físico no afectaría a la base de datos o las aplicaciones que accedan a ellos.
* **Redundacia**, los datos se almacenan, por lo general, una vez.
* **Interfaz de alto nivel**, uso de lenguajes de alto nivel.
* **Consulta directa**, herramienta para poder acceder a los datos interactivamente.

1. **USOS**

* Cuatro tipos de personas que hacen uso de una Bda:
* **El Administrador.**

El encargado de la creación o implementación física de la Bda, generalmente toma las decisiones relacionadas con el funcionamiento físico del almacenamiento de información.

Junto a estas tareas, establecerá la politica de seguridad y de acceso para garantizar el menor número de problemas.

* **Los Diseñadores.**

Encargados de diseñar cómo será la Bda.

Han de conocer a fondo los datos y procesos a representar en la Bda, debe implicar en el proceso a todos los usuarios de la Bda.

* **Los Programadores de Aplicaciones.**

Se encargan de implementar los programas de aplicación que sirven a los usuarios finales, estos programas ofrecen la posibilidad de realizar consultas, inserción, actualización o eliminación de datos.

* **Los usuarios finales.**

Los clientes finales de la Bda.

* **Usos de las Bda**

Se usan en diversos campos como pueden ser: banca, líneas aéreas, universidades, transacciones de tarjeta de crédito, telecomunicadores, medicina, justicia, legislación, organismos públicos, cultura, etc.

* 1. **Ubicación de la Información**
* **Discos SATA**, interfaz de transcendencia de datos entre la placa base y algunos dispositivos de almacenamiento, como pueden ser disco duro, lectores y regrabadores de CD/DVD/BD.
* **Discos SCSI**, interfaces preparadas para discos duros de gran capacidad de almacenamiento y velocidad de rotación.
* **RAID** (**Redundant Array of Independent Disk**), se basa en el montaje en conjunto de dos o más discos duros, formando un bloque de trabajo para obtener desde una apliación de capacidad a mejoras de velocidad y seguridad de almacenamiento.
* **Sistemas NAS** (**Network Attahced Storage**), permiten compartir la capacidad de almacenamiento de un computador(Servidor) con ordenadores personales o servidores clientes a través de una red, haciendo uso de un SO optimizado para dar acceso a los datos a través de protocolos de comunicación específicos.
* **Sistemas SAN** (**Storage Area Network**), red concebida para conectar servidores, matrices(arrays) de discos y liibrerías de soporte. Su arquitectura permite que los recursos de almacenamiento estén disponibles para varios servidores en una red de área local o amplia.

1. **MODELOS DE BASES DE DATOS**
   1. **Modelo Jerárquico**

La información se organiza con jerarquía en la que la relación entre las entidades siempre es del tipo padre/hijo, de tal manera que existen nodos que contienen atributos o campos y se relacionarán con sus nodos hijos, pudiendo tener cada nodo más de un hijo, pero un nodo tendra siempre un solo padre.

**Segmentos**, estructuras lógicas donde se almacenan los datos en este modelo, se relacionan entre sí utilizando arcos.

La forma visual de este modelo es de árbol invertido, en la parte superior están los padres y en la inferior los hijos.

El modelo jerárquico esta en desuso debido a sus limitaciones.

* 1. **Modelo en Red**

Organiza la información en registros (nodos) y enlaces.

En los registros se almacenan los datos.

Los enlaces permiten relacionar los datos

Este modelo es parecido al Jerárquico pero puede haber más de un padre.

Se pueden representar perfectamente cualquier tipo de relación entre datos, pero hace muy complicado su manejo.

Ahorra espacio de almacenamiento al no tener que duplicar información.

El sistema más popular es el IDMS.

* 1. **Modelo Relacional**

Este modelo es precibido a nivel lógico por el usuario como un conjunto de tablas, a nivel físico puede estar implementada mediante distintas estructuras de almacenamiento.

**Tablas bidimensionales** (relacionales), usadas para la representación lógica de los datos y las relaciones entre ellos. Cada relación(tabla) posee un nombre único y contiene un conjunto de columnas.

* **Registro, entidad** o **tubla**, cada fila de la tabla.
* **Campo** o **atributo**, cada columna de la tabla.
* **Dominio**, conjunto de valores que puede tomar un deter. atributo.
* **Clave**, un atributo o conjunto de ellos que identifique de forma única a una tubla.

Requisitos que deben cumplir las tablas:

* 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.

Durante su diseño, una Bda relacional pasa por un proceso que se conoce como normalizacion de una Bda.

* 1. **Modelo Orientado a Objetos**

Define una Bda en términos de **Objetos**, sus propiedades y operaciones.

Los objetos con la misma estructura y comportamiento pertenecen a una **clase**.

Las clases se organizanen **jerarquías**.

**Métodos** procedimientos predefinidos que especifican las operaciones de cada clase.

Modelos **Objeto-relacionales**

Conceptos más importantes del paradigma de objetos:

* **Encapsulación**, propiedad que permite ocultar la información al resto de objetos impidiendo accesos incorrectos o conflictivos.
* **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.
  1. **Otros Modelos**
* **Modelo Objeto-Relacional**.

Híbrido entre relacional y orientado a objetos.

Se busca obtener lo mejor del modelo relacional, incorporando las mejoras ofrecidas por la orientación a objetos.

Se siguen almacenando tuplas, sin restringir su estructura, sino que las relaciones pueden ser definidas en función de otras -> Herencia Directa.

Tienen capacidad de incorporar funciones que tengan un código de algun lenguaje de ptrogramación.

* **Modelo de Bases de Datos Educativas**.

Las Bda almacenan la información y permiten realizar deducciones a través de infernecias, es decir, se derivan nuevas informaciones a partir de las que se han introducido explícitamte en la Bda por parte del usuario.

Son tambien llamadas Bda lógicas, al basarse en la lógica matemática.

Surgieron para contrarrestar las limitaciones del modelo relacional para la respuesta a consultas recursias y la deduccion de relaciones indirectas entre los datos almacenados.

* **Bases de Datos Multidimensionales**.

Ideadas para desarrollar aplicaciones muy concretas.

Almacenan sus datos con varias dimensiones, encontramos varios valores dependiendo de los ejes definidos o una Bda de estructura basada en dimensiones orientada a consultas complejas y de alto rendimiento.

La información se representa como **Cubos** -> matrices multidimensionales, cuadros de múltiples entradas o funciones de varias variables sobre conjuntos finitos.

Facilita el manejo de grandes cantidades de datos dándo una amplia aplicación dentro de varias áreas y diferentes campos del conocimiento.

* **Bases de Datos Transaccionales.**

Caracterizadas por su velocidad para gestionar el intercambio de información, utilizadas sobre todo en sistemas bancarios, análisis de calidad y datos de producción industrial.

Muy fiables, ya que cada una de las operaciones de insercción, actualización o borrado se realizan completamente o se descartan.

* **Bases de Datos Orientadas a Documentos.**

En este el principal objeto de gestión es el documento que contiene datos semiestructurados que podrán estar almacenados en algun formato.

1. **TIPOS DE BASES DE DATOS**

* Según su Contenido.
  + Con Información Actual.
  + Directorios.
  + Documentales.
    - De texto completo.
    - Archivos electrónicos de imágenes.
    - Referenciales.
* Según su Uso.
  + Individual.
  + Compartida.
  + Acceso Público.
  + Propietarias o Banco de Datos.
* Según la Variabilidad de la Información.
  + Estáticas.
  + Dinámicas.
* Según la Localización de la Información.
  + Centralizadas.
    - Basada en Anfitrión.
    - Basada en Cliente/Servidor.
  + Distribuidas.
* Según el Organismo Producto.
  + De Organismos Públicos y de la Administración.
    - De acceso público.
    - De uso interno.
  + De Instituciones Sin Ánimo de Lucro.
  + De Entidades Privadas o Comerciales.
    - Para facilitar la circulación de info dentro de la empres.
    - Que ocasionalmente ofrecen servicio hacia el exterior.
    - Diseñadas especificamente para user externos.
  + Realizadas por Cooperación en Red.
* Según el Modo de Acceso.
  + De Acceso Local.
  + En CD-ROM.
  + En Línea.
    - Acceso Vía Telnet o Mediante Línea de Internet.
    - Acceso vía Web.
* Según Cobertura Temática.
  + Científico-tecnológicas.
    - Multidisciplinares.
    - Especializadas.
    - Económico-empresariales.
    - De Medios de Comunicación.
    - Del Ámbito político-administrativo y Jurídico.
    - Del Ámbito Sanitario.
    - Para el Gran Público.

1. **SISTEMAS GESTORES DE BASES DE DATOS**

**DataBase Management System**, conjunto coordinado de programas, procedimientos, lenguajes, …, que suministran los medios necesarios para describir y manipular los datos contenidos en la Bda, manteniendo su integridad, confidencialidad y seguridad.

Ventajas

* Proporcionan una visión abstracta de los datos, ocultando complejidades relacionadas al almacenamiento y mantenimento de los datos.
* Ofrencen **Independencia Física**, la visión de la información y la manipulación de los datos almacenados es independiente de cómo estén almacenados físicamente.
* Disminuyen la redundancia e inconsistencia de los datos.
* Aseguran la integridad de los datos.
* Facilitan el acceso a los datos, aportanto rapidez y evitando la perdida de datos.
* Aumenta la seguridad y la privacidad de los datos.
* Mejoran la eficiencia.
* Permiten compartir datos y accesos concurrentes.
* Facilitan el intercambio de datos entre sistemas.
* Incorporan mecanimos de backup y recoupreacion en casos de fallos en el sistema.
  1. **Funciones**

**Función de Descripción o Definición.**

Permite al disñador crear las estructuras apropiadas para integrar adecuadamente los datos, permite difinir la estructura interna, conceptual y externa.

Se raliza mediante el **lenguaje de descripcion de datos o DDL**.

Se especificarán las características de los datos a cada uno de los 3 niveles y el SGBD se ocupara de la transformación de las estructuras externas orientadas a los usuarios a las estructucturas conceptuales y de la relación de ésta y la estructura física.

**A nivel interno**, se ha de indicar el espacio de disco reservado para la Bda, la longitud de campos, su modelo de representación.

**A nivel conceptual**, se proporcionan herramientas para la definición del esquema de la base de datos.

**A nivel externo**, se deben definir las vistas de los distintos users a través del lenguaje para la definición de estructuras externas.

**Función de Manipulación.**

Permite a los users buscar, añadir, suprimir o modificar los datos, siempre de acuerdo con las especificaciones y las normas de seguridad dictadas por el admin.

S llevá a cabo por medio de un **Lenguaje de Manipulación de Datos** (**DML**) que facilita los instrumentos necesarios para la realización de estas tareas.

Se encarga también de definir **la vista externa**, de todos los users o vistas parciales que cada user tiene de los datos definidos con el DDL.

Para la manipulacion de datos entenderemos:

* **Consultas**, la recuperación de la información almacenada.
* La inserción de información nueva.
* El borrado de información.
* La modificación de información almacenada.

**Función de Control.**

Permite al admin establecer mecanismos de protección de las diferentes visiones de los datos asociadas a cada user.

El lenguaje que implementa esta función es el **lenguaje de control de datos** o **DCL**.

* 1. **Componentes**.
     1. **Lenguajes de la Bda.**

A través de ellos se pueden especificar los datos que componen la Bda, su estructura, relaciones reglas de integridad, control de acceso, características físicas y vistas externas de los users.

* **DDL**, Data Definition Lenguaje.
* **DML**, Data Management Lenguaje.
* **DCL**, Data Control Lenguaje.
  + 1. **El diccionario de Datos.**

Descripción de los datos almacenados, lugar donde se deposita la información sobre la totalidad de los datos que forman la Bda.

Contiene las características lógicas de las estructuras que almacenan los datos, su nombre, descripción, contenido y organización.

En una Bda relacional, nos aportará información sobre:

* Estructura lógica y física de la Bda.
* Def. de tablas, vistas, indices, disparadores, procedimientos, etc.
* Cantidad de espacio asignado y utilizado por los elementos.
* Descripción de las restricciones de integridad.
* Información sobre los permisos asociados a cada perfil del user.
* Auditoría de acceso a los datos, utilización, etc.
  + 1. **El gestor de la Bda.**

La parte del software encargada de garantizar el correcto, seguro, íntegro y eficiente acceso y mantenimiento de datos.

El encargado de proporcionar una interfaz entre los datos almacenados y los programas de la aplicación que los manejan.

Es un intermediario entre el user y los datos.

Encargado de garantizar la privacidad, seguridad e integridad de los datos, controlando los accesos concurrentes e interactuando con el S.O.

* + 1. **Usuarios de la Bda.**
* **Database Administrator**, persona o conjunto de ellas encargadas de la función de administración de la Bda, tiene el control centralizado de la Bda y es el responsable de su buen funcionamiento.

Encargado de autorizar el acceso a la base de datos, de coordinar y vigilar su utilización y adquirir los recursos soft/hardware necesarios.

* **Usuarios de la Bda**, los diferentes users con diferentes necesidades sobre los datos, con diferentes accesos y privilegios:
  + **Diseñadores.**
  + **Operadores y personal de mantenimiento.**
  + **Analistas y programadores de aplicaciones.**
  + **Users finales: ocasionales, simples, avanzados y autónomos.**
    1. **Herramientas de la Bda.**

Conjunto de aplicaciones que permiten a los administradores la gestión de la base de datos, de los usuarios y permisos, genedarores de formularios, informes, interfaces gráficas, generadores de aplicaciones, etc.

* 1. **Arquitectura**

El objetivo fundamental es separar los programas de aplicación de la Bda física.

La arquetectura propuesta proporciona 3 niveles de abstracción:

* **Nivel interno o físico**, describe la estructura física a través de un esquema interno encargado de detallar el sistema de almacenamiento y sus métodos de acceso.
* **Nivel lógico o conceptual**, describe la estructura completa a través de un esquema que detalla las entidades, atributos, relaciones, operaciones de los usuarios y restricciones.

Los detalles relacionados con las estructuras de almacenamiento se ocultan, permitiendo realizar una abstracción a más alto nivel.

* **Nivel externo o visión del usuario**, se describen las diferentes vistas que los usuarios percibirán de la base de datos.

Con esta arquitectura se consigue la **independencia de datos** a 2 niveles:

* **Independencia lógica**, podemos modificar el esquema conceptual sin alterar los esquemas externos ni los programas de aplicación.
* **Independencia física**, podemos modificar el esquema interno sin necesidad de modificar el conceptual o el externo.
  1. **Tipos de Sistemas Gestores de Datos**
     1. **Según el modelo lógico en que se basan:**
        1. **Modelo jérarquido.**
        2. **Modelo en Red.**
        3. **Modelo Relacional.**
        4. **Modelo Orientado a Objetos.**
     2. **Según el número de usuarios:**
        1. **Monousuario**, un usuario a la vez, se usa principalmente en ordenadores personales.
        2. **Multiusuario**, varios usuarios al mismo tiempo.
     3. **Según el número de sitios en los que está distribuida:**
        1. **Centralizados**, sus datos se almacenan en un solo computador, pueden atender a varios usuarios.
        2. **Distribuidos**(Hetereo/homogeneos), la Bda real y el propio software pueden estar distribuidos en varios sitios conectados por una red.
     4. **Según el coste**.
     5. **Según el propósito o finalidad:**
        1. **Propósito General**, pueden ser utilizados para el tratamiento de cualquier tipo de Bda y app.
        2. **Propósito Específico**, para un propósito especial para una aplicación específica.
  2. **SGBD comerciales**
  3. **SGBD Libres**

1. **BASES DE DATOS CENTRALIZADAS**

Es aquella estructura en la que el SGBD está implantando en una sola plataforma y ordenador desde donde se gestiona directamente, de modo centralizado, la totalidad de los recursos.

Es la arquitectura de los centros de proceso de datos tradicionales, se basa en tecnologia sencillas, muy experimentadas y de gran robustez.

Principales caracteristicas:

* Almacenada en una ubicación central.
* No posee múltiples elementos de procesamiento ni mecanismos de intercomunicación.
* Los componentes seon: los datos, el software de gestion de Bda y los dispositivos de almacenamiento secundario asociados.
* Su seguridad puede verse comprometida más fácilmente.

1. **BASES DE DATOS DISTRIBUIDAS**

**Base de Datos Distribuida – DDBA**

Conjunto de múltiples Bda lógicamente relacionadas, las cuales se encuentran distribuidas entre diferentes nodos interconectados por una red de comunicaciones.

**Sistema de Bases de Datos Distribuida – DDBAS**

Sistema en el cual múltiples sitios de Bda están ligados por un sistema de comunicaciones, de tal forma que, un user de cualquier sitio puede acceder a los datos en cualquier parte d la red exactamente como si los datos estuvieran almacenados en su sitio propio.

**Sistema Gestor de Bases de Datos Distribuida – MSDDBA**

Aquel encargado del manejo de la DDBA y proporciona un mecanismo de acceso que hace que la distribución sea tranparente a los usuarios, es decir que la aplicación trabajará como un solo MSBDA ejecutando una sola máquina administradora de datos.

1. **FRAGMENTACIÓN**

3 Reglas fundamentales al llevar a cabo una fragmentación; si una relacion R se descompone en fragmenteo R1, R2, .., Rn,:

* **Completitud**, cada elemento de datos que pueda encontrarse en R deberá poder encontrarse en uno o varios fragmentos Ri.
* **Reconstrucción**, la reconstrucción de la relación a partir de sus fragmentos asegura que se preservan las restricciones definidas sobre los datos.
* **Disyunción**, Si una relación se descompone verticalmente, sus atributos primarios clave normalmente se repiten en todos sus fragmento.

Existen 3 tipos de fragmentación:

* **Horizontal**, se realiza sobre las tuplas de la relacion, dividiendo la relación en subrelaciones que contienen un subconjunto de las tuplas que alberga la primera. (F.H.Primaria, F.H.Derivada)
* **Vertical**, se basa en los atribuos de la relación para efectuar la división.

Cada uno de los Ri contiene un subconjunto de los atributos R así como la llave primarioa de R.

Su objetivo es particionar una relación en un conjunto de relaciones más pequeñas de manera que varias aplicaciones de users se ejecutarán sobre un fragmento.

En esta, una fragmentación optima es la que produce un esquema que minimiza el tiempo de ejecucion de las consolas de usuario.

* **Híbrida o Mixta**, combinación de ambas.
  + **Fragmentación HV**, tras una vertical se realiza una horizontal.
  + **Fragmentacion VH**, tras una horizontal se realiza una vertical.