

# ALMACENAMIENTO DE LA INFORMACIÓN



# ALMACENAMIENTO DE LA INFORMACIÓN

- Ficheros (planos, indexados y acceso directo, de marcas, entre otros).
- Bases de Datos (BD). Conceptos, usos y tipos según el modelo de datos, la ubicación de la información.
- Sistemas gestores de bases de datos (SGBD): funciones, componentes y tipos.
- SGBD comerciales vs. SGBD libres.
- Bases de datos centralizadas y bases de datos distribuidas



# ALMACENAMIENTO DE LA INFORMACIÓN

- Ficheros (planos, indexados y acceso directo, de marcas, entre otros).
  - Un ordenador almacena muchos tipos de información (datos administrativos, contables, bancarios, música, películas, páginas webs, partidas de videojuegos, etc.)
  - Toda esta información está almacenada en los dispositivos de almacenamiento del ordenador, como son los discos duros, pen drives, etc.
  - Para poder organizar la información se utilizan los ficheros o archivos.
- **Los ficheros son estructuras de información que crean los sistemas operativos de los ordenadores para poder almacenar datos.** Suelen tener:
  - Nombre
  - Extensión, que determina tipo de información que contiene y el formato de la misma



# ALMACENAMIENTO DE LA INFORMACIÓN

- Ficheros (planos, indexados y acceso directo, de marcas, entre otros).
  - El formato y tipo de fichero determinan la forma de interpretar la información que contiene (lo que se almacena en un fichero es una ristra de bits, que es necesario interpretar para dar sentido a la información que almacena).
    - Por ejemplo, fichero bmp (binario para almacenar imágenes):
      - Vector de datos con los colores que tiene cada pixel que forma la imagen
      - Paleta de colores
      - Dimensiones

Todos estos datos se ordenan según un formato. La aplicación que trate los gráficos, debe conocer este formato para poder extraer los píxeles y mostrarlos por pantalla



# ALMACENAMIENTO DE LA INFORMACIÓN

- Clasificación de los ficheros.
  - Según su contenido:
    - Texto: los bits almacenados en un fichero pueden ser traducidos de forma directa a caracteres alfabéticos y números que entiende el ser humano
    - Binarios: constituidos en base a componentes de estructuras de datos más complejas, que almacenan sonido, vídeo, imágenes, etc.
  - Según su organización (la forma en que se han de acceder a los datos):
    - Secuencial: los datos están dispuestos siguiendo una secuencia ordenada (unos detrás de otros). Es necesario recorrer todos los datos anteriores para llegar a uno en concreto
    - Acceso directo: permiten acceder a un dato en concreto sin necesidad de acceder a todos los anteriores. Formados por "trozos" de igual tamaño, que se denominan normalmente registros, y para acceder a un registro se calcula su posición inicial a partir del número de registro y de su tamaño
    - Indexados: acceden a los datos consultando un índice (estructura de datos que permite acceder a la información rápidamente)



# ALMACENAMIENTO DE LA INFORMACIÓN

- Clasificación de los ficheros.
  - Según la utilidad (el uso que se va a hacer del fichero):
    - Maestro: información fundamental y/o estable de una organización
    - De movimientos: con las variaciones que se producen en el día a día y que periódicamente se incorporan a ficheros maestros
    - Históricos: con datos que no son necesarios para su proceso diario
  - Actualmente las dos últimas clasificaciones no se suelen considerar, el sistema operativo trata los ficheros según:
    - Su contenido (texto o datos binarios)
    - Su tipo (ejecutables, imágenes, vídeos, etc.)



# ALMACENAMIENTO DE LA INFORMACIÓN

- Ficheros de texto
  - También se denominan ficheros planos o ficheros ascii (American Standard Code for Information Interchange)
  - El código ascii: asigna un valor numérico a cada carácter → **tabla ascii**
    - Utiliza 1 byte para la representación de cada carácter → 256 posibles caracteres
  - Existen otras codificaciones (UTF, EBCDIC,...)
- Los ficheros de texto, aunque no necesitan un formato para ser interpretados suelen tener extensiones para conocer qué tipo de texto se halla dentro del fichero:
  - Ficheros de configuración: .ini, .inf, .conf
  - Código fuente: .c, .java, .sql
  - Páginas web: .html, .xml



# ALMACENAMIENTO DE LA INFORMACIÓN

- Ficheros de texto
  - Tabla ascii

| Caracteres ASCII de control |      |                     | Caracteres ASCII imprimibles |         |    | ASCII extendido<br>(Página de código 437) |     |   |     |   |     |   |     |   |     |      |
|-----------------------------|------|---------------------|------------------------------|---------|----|-------------------------------------------|-----|---|-----|---|-----|---|-----|---|-----|------|
| 00                          | NULL | (carácter nulo)     | 32                           | espacio | 64 | @                                         | 96  | ` | 128 | Ç | 160 | á | 192 | Ł | 224 | Ó    |
| 01                          | SOH  | (inicio encabezado) | 33                           | !       | 65 | A                                         | 97  | a | 129 | ü | 161 | í | 193 | ł | 225 | ß    |
| 02                          | STX  | (inicio texto)      | 34                           | "       | 66 | B                                         | 98  | b | 130 | é | 162 | ó | 194 | Ł | 226 | Ò    |
| 03                          | ETX  | (fin de texto)      | 35                           | #       | 67 | C                                         | 99  | c | 131 | â | 163 | ú | 195 | ł | 227 | Ô    |
| 04                          | EOT  | (fin transmisión)   | 36                           | \$      | 68 | D                                         | 100 | d | 132 | ä | 164 | ñ | 196 | — | 228 | ö    |
| 05                          | ENQ  | (consulta)          | 37                           | %       | 69 | E                                         | 101 | e | 133 | à | 165 | Ñ | 197 | † | 229 | Õ    |
| 06                          | ACK  | (reconocimiento)    | 38                           | &       | 70 | F                                         | 102 | f | 134 | å | 166 | ª | 198 | ‡ | 230 | µ    |
| 07                          | BEL  | (timbre)            | 39                           | '       | 71 | G                                         | 103 | g | 135 | ç | 167 | º | 199 | ä | 231 | þ    |
| 08                          | BS   | (retroceso)         | 40                           | (       | 72 | H                                         | 104 | h | 136 | ê | 168 | ¿ | 200 | Å | 232 | þ    |
| 09                          | HT   | (tab horizontal)    | 41                           | )       | 73 | I                                         | 105 | i | 137 | ë | 169 | ® | 201 | Ł | 233 | Û    |
| 10                          | LF   | (nueva línea)       | 42                           | *       | 74 | J                                         | 106 | j | 138 | è | 170 | ™ | 202 | ł | 234 | Ü    |
| 11                          | VT   | (tab vertical)      | 43                           | +       | 75 | K                                         | 107 | k | 139 | ï | 171 | ½ | 203 | Ł | 235 | Ù    |
| 12                          | FF   | (nueva página)      | 44                           | ,       | 76 | L                                         | 108 | l | 140 | î | 172 | ¼ | 204 | ł | 236 | Ý    |
| 13                          | CR   | (retorno de carro)  | 45                           | -       | 77 | M                                         | 109 | m | 141 | ï | 173 | ¡ | 205 | = | 237 | Ÿ    |
| 14                          | SO   | (desplaza afuera)   | 46                           | .       | 78 | N                                         | 110 | n | 142 | Ä | 174 | « | 206 | ‡ | 238 | ˆ    |
| 15                          | SI   | (desplaza adentro)  | 47                           | /       | 79 | O                                         | 111 | o | 143 | Å | 175 | » | 207 | ¤ | 239 | ˜    |
| 16                          | DLE  | (esc.vínculo datos) | 48                           | 0       | 80 | P                                         | 112 | p | 144 | É | 176 | • | 208 | ø | 240 | ≡    |
| 17                          | DC1  | (control disp. 1)   | 49                           | 1       | 81 | Q                                         | 113 | q | 145 | æ | 177 | • | 209 | Ð | 241 | ±    |
| 18                          | DC2  | (control disp. 2)   | 50                           | 2       | 82 | R                                         | 114 | r | 146 | Æ | 178 | • | 210 | Ê | 242 | ≡    |
| 19                          | DC3  | (control disp. 3)   | 51                           | 3       | 83 | S                                         | 115 | s | 147 | ø | 179 | • | 211 | Ë | 243 | ¾    |
| 20                          | DC4  | (control disp. 4)   | 52                           | 4       | 84 | T                                         | 116 | t | 148 | ö | 180 | • | 212 | È | 244 | ¶    |
| 21                          | NAK  | (conf. negativa)    | 53                           | 5       | 85 | U                                         | 117 | u | 149 | ò | 181 | • | 213 | Ì | 245 | §    |
| 22                          | SYN  | (inactividad sinc)  | 54                           | 6       | 86 | V                                         | 118 | v | 150 | û | 182 | • | 214 | Í | 246 | ÷    |
| 23                          | ETB  | (fin bloque trans)  | 55                           | 7       | 87 | W                                         | 119 | w | 151 | ù | 183 | • | 215 | Î | 247 | °    |
| 24                          | CAN  | (cancelar)          | 56                           | 8       | 88 | X                                         | 120 | x | 152 | ÿ | 184 | • | 216 | Ï | 248 | ˆ    |
| 25                          | EM   | (fin del medio)     | 57                           | 9       | 89 | Y                                         | 121 | y | 153 | Ö | 185 | • | 217 | Ɔ | 249 | ˆ    |
| 26                          | SUB  | (sustitución)       | 58                           | :       | 90 | Z                                         | 122 | z | 154 | Ü | 186 | • | 218 | Ɔ | 250 | ˆ    |
| 27                          | ESC  | (escape)            | 59                           | ;       | 91 | [                                         | 123 | { | 155 | ø | 187 | • | 219 | Ɔ | 251 | ˆ    |
| 28                          | FS   | (sep. archivos)     | 60                           | <       | 92 | \                                         | 124 |   | 156 | £ | 188 | • | 220 | Ɔ | 252 | ˆ    |
| 29                          | GS   | (sep. grupos)       | 61                           | =       | 93 | ]                                         | 125 | } | 157 | Ø | 189 | • | 221 | Ɔ | 253 | ˆ    |
| 30                          | RS   | (sep. registros)    | 62                           | >       | 94 | ^                                         | 126 | ~ | 158 | × | 190 | • | 222 | Ɔ | 254 | ˆ    |
| 31                          | US   | (sep. unidades)     | 63                           | ?       | 95 | _                                         |     |   | 159 | f | 191 | • | 223 | Ɔ | 255 | nbsp |
| 127                         | DEL  | (suprimir)          |                              |         |    |                                           |     |   |     |   |     |   |     |   |     |      |





# ALMACENAMIENTO DE LA INFORMACIÓN

- **Ficheros binarios**
  - Son todos los que no son de texto. Requieren la consideración de un formato para ser interpretados
    - De imagen: jpg, gif, .tiff, .bmp, .wmf, .png, .pcx, ...
    - De vídeo: .mpg, .mov, .avi, .qt, ...
    - Ejecutables o compilados: exe, .com, .cgi, ...
    - Procesadores de textos: .doc, .odt, ...
  - Generalmente los ficheros que componen una base de datos son de tipo binario, ya que la información que hay en ellos debe tener una estructura lógica y organizada para que las aplicaciones puedan acceder a ella. Esta estructura es difícil de expresar mediante ficheros de texto, por tanto, la información de una base de datos se suele guardar en uno o varios ficheros binarios:
    - Oracle guarda la información en múltiples tipos de ficheros, llamados datafiles, tempfiles, logfiles, etc.
    - Un tipo de tablas del gestor MySQL guarda su información en 3 ficheros binarios, con extensión frm, myd y myi.
    - Access guarda toda la información de una base de datos en un fichero con extensión mdb.



# ALMACENAMIENTO DE LA INFORMACIÓN

- Bases de Datos
  - Una Base de Datos es una **colección de información perteneciente a un mismo contexto** (problema, asunto) que está almacenada de forma organizada en ficheros
  - El **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 así como proporcionar acceso controlado a la misma. Es una **herramienta que sirve de interfaz entre el usuario y las bases de datos**
  - Lo más común es que una base de datos está organizada mediante **tablas** que almacenan información. Estas tablas se relacionan formando vínculos (**relaciones**) entre ellas que ayudan a mantener la información de los diferentes objetos de forma ordenada y coherente (sin contradicciones)



# ALMACENAMIENTO DE LA INFORMACIÓN

## ■ Bases de Datos

- Cada una de estas tablas es una estructura que se parece a las hojas de cálculo. Está dispuesta mediante filas y columnas.
- Cada fila almacena un registro con tantos campos como columnas tenga la tabla. Por ejemplo, se podría tener una tabla de Clientes, donde cada **fila** o registro corresponde a un cliente de la empresa y cada **columna** o campo representa un trozo discreto de información sobre cada cliente, por ejemplo su dirección o el nombre de la persona de contacto.

| Clientes       |                                    |                     |                          |                               |              |   |
|----------------|------------------------------------|---------------------|--------------------------|-------------------------------|--------------|---|
| Id. de cliente | Nombre de compañía                 | Nombre del contacto | Cargo del contacto       | Dirección                     | Ciudad       |   |
| ALFKI          | Alfreds Futterkiste                | Maria Anders        | Representante de ventas  | Obere Str. 57                 | Berlín       |   |
| ANATR          | Ana Trujillo Emparedados y helados | Ana Trujillo        | Propietario              | Avda. de la Constitución 2222 | México D.F.  |   |
| ANTON          | Antonio Moreno Taquería            | Antonio Moreno      | Propietario              | Mataderos 2312                | México D.F.  |   |
| AROUT          | Around the Horn                    | Thomas Hardy        | Representante de ventas  | 120 Hanover Sq.               | Londres      |   |
| BERGS          | Berglunds snabbköp                 | Christina Berglund  | Administrador de pedidos | Berguvsvägen 8                | Luleå        |   |
| BLAUS          | Blauer See Delikatessen            | Hanna Moos          | Representante de ventas  | Forsterstr. 57                | Mannheim     |   |
| BLONP          | Blondel père et fils               | Frédérique Citeaux  | Gerente de marketing     | 24, place Kléber              | Estrasburgo  |   |
| BOLID          | Bólido Comidas preparadas          | Martín Sommer       | Propietario              | C/ Araquil, 67                | Madrid       |   |
| BONAP          | Bon app'                           | Laurence Lebihan    | Propietario              | 12, rue des Bouchers          | Marsella     |   |
| BOTTM          | Bottom-Dollar Markets              | Elizabeth Lincoln   | Gerente de contabilidad  | 23 Sawassen Blvd.             | Tsawassen    | B |
| BSBEV          | B's Beverages                      | Victoria Ashworth   | Representante de ventas  | Fauntleroy Circus             | Londres      |   |
| CACTU          | Cactus Comidas para llevar         | Patricio Simpson    | Agente de ventas         | Cerrito 333                   | Buenos Aires |   |
| CENTC          | Centro comercial Moctezuma         | Francisco Chang     | Gerente de marketing     | Sierras de Granada 9993       | México D.F.  |   |
| CHOPS          | Chop-suey Chinese                  | Yang Wang           | Propietario              | Hauptstr. 29                  | Berna        |   |
| COMMI          | Comércio Mineiro                   | Pedro Afonso        | Asistente de ventas      | Av. dos Lusíadas, 23          | São Paulo    | S |
| CONSH          | Consolidated Holdings              | Elizabeth Brown     | Representante de ventas  | Berkeley Gardens              | Londres      |   |
| DRACD          | Drachenblut Delikatessen           | Sven Ottlieb        | Administrador de pedidos | Walserweg 21                  | Aachen       |   |
| DUMON          | Du monde entier                    | Janine Labrune      | Propietario              | 67, rue des Cinquante Otages  | Nantes       |   |
| EASTC          | Eastern Connection                 | Ann Devon           | Agente de ventas         | 35 King George                | Londres      |   |

Registro: 1 de 91 Sin filtro Buscar



# ALMACENAMIENTO DE LA INFORMACIÓN

- Bases de Datos. Conceptos (I)
  - **Dato**: es un trozo de información concreta sobre algún concepto o suceso. Por ejemplo, 2000 es un número que representa un año de nacimiento de una persona. Los datos se caracterizan por pertenecer a un tipo
  - **Tipo de Dato**: indica la naturaleza del campo (dato). Se puede tener datos numéricos (aquellos con los que se pueden realizar cálculos aritméticos) y los datos alfanuméricos que contienen caracteres alfabéticos y dígitos numéricos. Los datos alfanuméricos y numéricos se pueden combinar para obtener tipos de datos más complejos, como el tipo de dato Fecha contiene tres datos numéricos que representan el día, el mes y el año de una fecha
  - **Registro**: conjunto de datos referentes a un mismo concepto o suceso. Por ejemplo, los datos de una persona pueden ser su nombre, su NIF, año de nacimiento, su dirección, etc. A los registros también se les llama tuplas o filas



# ALMACENAMIENTO DE LA INFORMACIÓN

- Bases de Datos. Conceptos (II)
  - **Campo**: es un identificador de un tipo de información (dato dentro de los registros). Cada campo pertenece a un tipo de datos. Por ejemplo, el campo “FechaNacimiento” hace referencia a las fechas de nacimiento de las personas que hay en la tabla. Este campo pertenece al tipo de dato Fecha. Al campo también se le denomina columna.
  - **Campo Clave**: Es un campo especial que identifica de forma única a cada registro, por ejemplo el NIF de una persona es único, por tanto es campo clave
  - **Tabla**: Es un conjunto de registros bajo un mismo nombre que representa el conjunto de todos ellos. Por ejemplo, la información de todos los clientes de una base de datos se pueden almacenar en una tabla cuyo nombre es Clientes



# ALMACENAMIENTO DE LA INFORMACIÓN

- Bases de Datos. Conceptos (III)

- **Consulta:** Es una instrucción para hacer peticiones a una base de datos

Puede ser una búsqueda simple de un registro específico o una solicitud para seleccionar todos los registros que satisfagan unos criterios

Sin embargo es más apropiado considerar el término de partida en inglés **query**, que hace referencia a petición, y que además de las consultas de búsqueda de información incluye peticiones de eliminación, inserción o actualización de registros, cuya ejecución altera los valores de los mismos

- **Índice:** Es una estructura que almacena los campos clave de una tabla, organizándolos para hacer más fácil encontrar y ordenar los registros de esa tabla. Se guarda información de la posición de un elemento en la base de datos, de forma que para buscar un elemento que esté indexado el SGBD solo tiene que buscar en el índice dicho elemento y acceder a la posición marcada por el índice



# ALMACENAMIENTO DE LA INFORMACIÓN

- Bases de Datos. Conceptos (IV)
  - **Vista**: Es una transformación que se hace a una o más tablas para obtener una nueva tabla virtual que no está almacenada realmente en los dispositivos de almacenamiento del sistema, aunque sí se almacena su definición
  - **Informe**: Es un listado ordenado de los campos y registros seleccionados en un formato fácil de leer. Generalmente se generan como peticiones expresas de un tipo de información por parte de un usuario (ejemplo, un informe de las facturas impagadas de un determinado mes ordenado por nombre de cliente)
  - **Guiones o scripts**: Son un conjunto de instrucciones, preparadas para su ejecución ordenada, que realizan operaciones avanzadas de mantenimiento en la base de datos
  - **Procedimientos**: Son un tipo especial de scripts que están almacenados en la base de datos y que forman parte de su esquema





# ALMACENAMIENTO DE LA INFORMACIÓN

- Bases de Datos. Estructura
  - Una base de datos almacena los datos en base a un **esquema**
  - El esquema es la **definición de la estructura donde se almacenan los datos**, contiene todo lo necesario para organizar la información mediante tablas, registros (filas) y campos (columnas)
  - También contiene otros objetos necesarios para el tratamiento de los datos como procedimientos, vistas, índices, etc.
  - Al esquema también se le suele llamar **metainformación**, es decir, información sobre la información o metadatos
- Los gestores de bases de datos Oracle, MySQL y DB2, entre otros, almacenan el esquema de la base de datos en tablas como cualquier otro dato, de forma que el propio esquema de la base de datos se puede tratar como cualquier otro dato de la base de datos





# ALMACENAMIENTO DE LA INFORMACIÓN

- Usos de las bases de datos (I). Están presentes en prácticamente todos los ámbitos de sistemas informáticos, entre los más frecuentes...
  - Bases de datos Administrativas: para registrar y relacionar clientes, pedidos, facturas, productos, etc.
  - Bases de datos Contables: que permitir gestionar los pagos, balances de pérdidas y ganancias, patrimonio, declaraciones de hacienda. . .
  - Bases de datos para motores de búsquedas: buscadores como por ejemplo Google tienen una base de datos muy voluminosa donde se almacena información sobre todos los documentos de Internet
  - Científicas: para recolección de datos climáticos y medioambientales, químicos, genómicos, geológicos. . .



# ALMACENAMIENTO DE LA INFORMACIÓN

- Usos de las bases de datos (II)
  - Configuraciones: para datos de configuración de sistemas informáticos, como por ejemplo, el registro de windows
  - Bibliotecas: para almacenar información bibliográfica, como la tienda de amazon o las bibliotecas propiamente dichas
  - Censos: con información demográfica de pueblos, ciudades y países
  - Virus: para guardar información sobre todos los potenciales software maliciosos
  - Otros usos: Militares, videojuegos, deportes, etc.



# ALMACENAMIENTO DE LA INFORMACIÓN

- Tipos de bases de datos según el modelo de datos (I)
  - Cuando el volumen y la complejidad de la información de las bases de datos es grande se hace preciso el uso de modelos para facilitar el diseño de la mismas
  - **Un modelo de datos es una colección de herramientas conceptuales** para describir los datos, las relaciones que existen entre ellos y sus restricciones
- Existen diferentes tipos de modelos



# ALMACENAMIENTO DE LA INFORMACIÓN

- Tipos de bases de datos según el modelo de datos (II). Tipos de modelos
  - **Conceptuales**. Se usan para describir los datos a nivel global, se presentan los datos de forma parecida a como los captamos en el mundo real. El más utilizado de este tipo es el Modelo Entidad-Relación

El MER representa la realidad a través de:

- Entidades: objetos que existen y que se diferencian de otros por sus características (p.e. altura, nombre, etc.)
- Atributos: son las características de las entidades
- Relaciones: para indicar las asociaciones o relaciones que puede haber entre entidades (p.e. un jugador está relacionado con un equipo con una relación de pertenencia)

Este tipo de modelos utiliza una simbología para representar cada elemento



# ALMACENAMIENTO DE LA INFORMACIÓN

- Tipos de bases de datos según el modelo de datos (III). Tipos de modelos
  - **Modelos lógicos tradicionales.** También se usan para describir los datos a nivel global pero de un modo más lógico, con una descripción más estructurada y más próximo a la implementación y a la máquina. Utilizan tablas de registros para representar los objetos modelados y sus relaciones:
    - Modelo relacional: se representan los datos y las relaciones entre estos a través de una colección de tablas
    - Modelo jerárquico: almacenan su información en una estructura jerárquica, los datos se organizan en una forma similar a un árbol al revés, con nodos padres e hijos (que sólo pueden tener un padre). Son muy útiles en el caso de grandes volúmenes de datos que mantienen una estructura estable, permitiendo gran rendimiento
    - Modelo de red: similar al jerárquico con la diferencia de que se permite que un nodo tenga varios padres, permitiendo mejorar el problema de redundancia de datos del modelo jerárquico, sin embargo son más complicadas de administrar



# ALMACENAMIENTO DE LA INFORMACIÓN

- Tipos de bases de datos según el modelo de datos (IV). Tipos de modelos
  - **Modelos lógicos avanzados**. Son modelos más recientes, cada vez más utilizados, para aplicaciones que manejan tipos de datos más complejos:
    - Modelos de datos orientados a objetos: se utilizan sobre todo en aplicaciones programadas siguiendo el paradigma de la orientación a objetos. En la base de datos se almacenan datos (estado) y la funcionalidad (comportamiento) asociados a los objetos que a su vez están relacionados con otros objetos incluidos en la base de datos
    - Modelos de datos declarativos, suelen usarse en bases de conocimiento:
      - Deductivos
      - Funcionales
- Otros tipos de bases de datos:
  - Geográficas: para almacenamiento de información geográfica
  - Documentales: especializadas en el almacenamiento y búsquedas de texto



# ALMACENAMIENTO DE LA INFORMACIÓN

- Sistema Gestor de Base de Datos, SGBD
  - Aplicación que permite a los usuarios definir, crear y mantener la base de datos así como proporcionar acceso controlado a la misma. Es una **herramienta que sirve de interfaz entre el usuario y las bases de datos**
  - Un ejemplo de software Gestor de Base de Datos es Oracle 11g (la última versión de Oracle es la 21c), que incorpora un conjunto de herramientas software que son capaces de estructurar en múltiples discos duros los ficheros de una base de datos, permitiendo el acceso a sus datos tanto a partir de herramientas gráficas como a partir de lenguajes de programación (PL-SQL, php, c++. . . )



# ALMACENAMIENTO DE LA INFORMACIÓN

- Funciones de un SGBD (I)
  - Permiten a los usuarios **almacenar datos, acceder a ellos** y actualizarlos de forma sencilla y evitando al usuario la complejidad y la gestión directa de las características físicas de los dispositivos de almacenamiento
  - Garantizan la **integridad de los datos**, respetando las reglas y restricciones que dicte el programador de la base de datos. Es decir, no permiten operaciones que dejen cierto conjunto de datos incoherentes (incompletos o incorrectos)
  - Integran, junto con el sistema operativo, un sistema de **seguridad** en base al control de acceso de usuarios exclusivamente a la información para la que disponen de autorización
  - Proporcionan un **diccionario de metadatos**, que contiene el esquema de la base de datos (cómo están estructurados los datos en tablas, registros y campos, las relaciones entre los datos, usuarios, permisos, etc.). Este diccionario de datos debe ser también accesible de una forma similar al resto de datos





# ALMACENAMIENTO DE LA INFORMACIÓN

- Funciones de un SGBD (II)
  - Permiten el **uso de transacciones**, garantizan que todas las operaciones de la transacción se realicen correctamente. En caso de alguna incidencia en alguna de las operaciones que forman parte de la transacción se deshacen los cambios ejecutados hasta el momento evitando inconsistencias
  - Ofrecen herramientas para la **obtención estadísticas sobre el uso del gestor**, registrando operaciones realizadas, operaciones fallidas y cualquier tipo de incidencia. Permite **monitorizar** el uso de la base de datos, y permiten analizar hipotéticos malfuncionamientos.
  - Permiten **uso concurrente**, es decir, varios usuarios trabajando sobre un mismo conjunto de datos y proporciona mecanismos que permiten resolver situaciones conflictivas en el acceso o modificación de un dato al mismo tiempo por parte de varios usuarios
  - **Independizan los datos de la aplicación** o usuario que los utiliza, haciendo más fácil su migración a otras plataformas



# ALMACENAMIENTO DE LA INFORMACIÓN

- Funciones de un SGBD (III)
  - Ofrecen **conectividad con el exterior**, lo que facilita replicar y distribuir bases de datos, al incorporar herramientas estándar de conectividad (el protocolo ODBC está muy extendido como forma de comunicación entre bases de datos y aplicaciones externas)
  - Incorporan herramientas para la **salvaguarda y restauración de la información en caso de desastre**. Algunos gestores, disponen de sofisticados mecanismos para poder establecer el estado de una base de datos en cualquier punto anterior en el tiempo. También deben ofrecer sencillas herramientas para la importación y exportación automática de información



# ALMACENAMIENTO DE LA INFORMACIÓN

- Componentes de un SGBD (I).

Son los elementos que proporcionan las funciones / servicios comentados anteriormente:

- **Lenguaje SQL** (Structured Query Language): la principal herramienta de un gestor de base de datos, implementa la interfaz de acceso y gestión del usuario, que consiste en un lenguaje muy sencillo mediante el cuál el usuario realiza peticiones al servidor. Este lenguaje está estandarizado por la ISO, lo que supone que todas las bases de datos que soporten SQL deben tener una sintaxis muy similar a la hora de aplicar el lenguaje. Se divide en 3 sublenguajes:
  - Lenguaje DML: lenguaje de manipulación (consulta, inserción, actualización y eliminación) de datos (Data Manipulation Language)
  - Lenguaje DDL: lenguaje de definición de datos (Data Definition Language). Este lenguaje permite crear la estructura de la base de datos completa (desde tablas hasta usuarios)
  - Lenguaje DCL: lenguaje de control de datos (Data Control Language). Incluye comandos que permiten al administrador gestionar el acceso a los datos contenidos en la base de datos.



# ALMACENAMIENTO DE LA INFORMACIÓN

- Componentes de un SGBD (II).
  - **Diccionario de datos**: esquemas que describen el contenido del SGBD incluyendo los distintos objetos y sus propiedades
  - **Objetos**: tablas base y vistas, consultas, tipos definidos de datos, restricciones de tabla, funciones y procedimientos almacenados, disparadores (triggers)
  - **Optimizador de consultas**: para determinar la estrategia óptima para la ejecución de consultas
  - **Gestor de transacciones**
  - **Planificador**: para programar y automatizar la realización de determinadas tareas
  - **Gestor de copias de seguridad**: creación, acceso y restauración en caso de necesidad
  - Herramientas para seguridad, integridad, control de concurrencia, control de recuperación, gestión del diccionario de datos, programación de aplicaciones, importación exportación de datos, distribución de datos, replicación y sincronización



# ALMACENAMIENTO DE LA INFORMACIÓN

- Tipos de SGBD (I).

Se pueden hacer diferentes clasificaciones:

- Según su capacidad:

- Gestores de **Bases de Datos ofimáticas**: manipulan bases de datos pequeñas (ofimáticas) orientadas a almacenar datos domésticos o de pequeñas empresas. Incluso permiten construir pequeñas aplicaciones para ayudar a un usuario inexperto a manipular los datos de una base de datos de forma sencilla e intuitiva. Un ejemplo de un SGBD ofimático es Microsoft Access que posee una interfaz de usuario muy sencilla
- Gestores de **bases de datos Corporativas**: tienen la capacidad de gestionar bases de datos enormes, de grandes o medianas empresas con una carga de datos y transacciones que requieren un servidor de grandes dimensiones. Estos gestores son capaces de manipular grandes cantidades de datos de forma muy rápida y eficiente y poder resolver la demanda de muchos (cientos) de usuarios. Un ejemplo típico de servidor de base de datos Corporativas es Oracle, DB2
- Soluciones **intermedias**: debido al coste elevado de estos últimos se ha recurrido a soluciones intermedias entre gestores de base de datos ofimáticas y corporativas, como es MySQL, que, además de ser gratuito y sencillo, es capaz de manipular gran cantidad de datos. Aunque implementa SQL, no tiene un lenguaje de programación propio como SQL Server u Oracle pero a cambio se integra fácilmente en las típicas soluciones XAMPP, que son paquetes que incluyen, además de MySQL, una versión del servidor Web Apache y varios lenguajes de script (php, perl. . . ) que dotan a MySQL de potentes herramientas para acceso y publicación de los datos



# ALMACENAMIENTO DE LA INFORMACIÓN

- Tipos de SGBD (II).
  - Según el número de sitios:
    - Centralizados
    - Distribuidos: en varios equipos que pueden ser homogéneos o heterogéneos
  - Ámbito de aplicación:
    - Propósito general
    - Propósito específico: centrada en un tipo de aplicaciones
  - Modelo lógico en el que se basan:
    - Jerárquico
    - En red
    - Relacional
    - Objeto-relacional
    - Orientado a objetos



# ALMACENAMIENTO DE LA INFORMACIÓN

- **SGBDs comerciales y libres (I)**
  - En los últimos años se viene produciendo una consolidación del software libre como alternativa técnicamente viable y económicamente más sostenible al software comercial, que permite ofrecer los mismo servicios con un coste menor
  - No obstante no siempre el software libre está exento de costes, ya que en muchas ocasiones puede ser necesario contratar un soporte experto (dependiendo de la relevancia del sistema para la organización)
  - En el ámbito de los SGDBs también hay alternativas de software libre como es MySQL frente a SQL Server o PostgreSQL frente a Oracle
  - Cada vez son más los fabricantes que ofrecen versiones gratuitas (que no libres) aunque con limitaciones en su funcionalidad para prueba y aprendizaje (se suelen denominar versiones express)



# ALMACENAMIENTO DE LA INFORMACIÓN

- SGBDs comerciales y libres (II)
  - Los SGBDs libres:
    - Presentan menores costes de adquisición y mantenimiento ya que el coste de las licencias y su mantenimiento es nulo
    - Suelen tener mayor disponibilidad de documentación y actualización de la misma
    - Por el contrario requieren de personal más cualificado, lo que en general supone mayores costes en personal aunque eso supone también una mayor independencia con respecto a proveedores externos en el uso y desarrollo de aplicaciones
  - Los sistemas comerciales suelen ser más cerrados y rígidos. Sin embargo existen contratos con tiempos bien definidos de resolución de problemas técnicos y de seguridad
  - En cualquier caso todas las tecnologías suelen disponer de servicios de soporte y documentación de calidad





# ALMACENAMIENTO DE LA INFORMACIÓN

- Bases de Datos y SGBD distribuidos (I)
  - Cuando una **base de datos está controlada por más de un servidor** se dice que está distribuida. Estos servidores están interconectados mediante una red de telecomunicaciones y los SGBD proporcionan mecanismos para poder consultar la información independientemente del servidor y la ubicación de los datos.
  - Por ejemplo, una consulta simultánea sobre dos tablas como puede ser la tabla de pedidos y la de clientes puede estar accediendo a dos servidores con dos instancias de bases de datos distintas. Incluso, si las tablas están particionadas es posible que la consulta sobre una única tabla esté accediendo a más de dos servidores.



# ALMACENAMIENTO DE LA INFORMACIÓN

- Bases de Datos y SGBD distribuidos (II)
  - La distribución de una base de datos proporciona una especie de **balanceo de carga** para que la información pueda ser accedida lo más rápidamente posible. De esta manera se consiguen los siguientes beneficios:
    - Más rapidez en el acceso a los datos y capacidad de almacenamiento, ya que **aumenta** el número de ordenadores donde se depositan los datos y la **capacidad de computación**, y por tanto, la capacidad de búsqueda de la información
    - Acceso más flexible ya que se incrementa el número de usuarios conectados que puede soportar el sistema al haber más capacidad de cómputo y gracias a la red de la BBDD distribuida, los usuarios pueden acceder desde más ubicaciones
    - **Escalabilidad**, al distribuir una base de datos en distintas ubicaciones se permite incrementar de forma más sencilla los recursos en cualquier momento para acomodar el sistema a mayores demandas de usuarios y aplicaciones
    - **Fiabilidad y cierta capacidad de tolerancia a fallos**, ya que si se realiza una adecuada política de fragmentación de la base de datos es posible conseguir que si algún nodo de la base de datos distribuida sufra algún percance, el resto del sistema siga funcionando con normalidad



# ALMACENAMIENTO DE LA INFORMACIÓN

- Bases de Datos y SGBD distribuidos (III)
  - Sin embargo no todo son beneficios:
    - Hay que tener en cuenta que estos sistemas son **más costosos de implementar** que los centralizados no solo por la duplicidad de recursos en términos de red y servidores
    - Presentan una **mayor dificultad de la administración** de los datos ubicados en distintas localizaciones
    - Se introduce un elemento adicional como son las redes de comunicaciones para el correcto funcionamiento del intercambio de los datos.
    - Además, la atomicidad de las operaciones es más complicada de conseguir por lo que la gestión transaccional tiene dificultades añadidas



# ALMACENAMIENTO DE LA INFORMACIÓN

- Bases de Datos y SGBD distribuidos (IV)
  - Tipos. Dependiendo del tipo de componentes elegidos en la formación de una base de datos distribuida se pueden tener:
    - BBDD distribuidas homogéneas: el SGBD utilizado es el mismo para todas las BBDD locales. Generalmente, el SGBD proporciona la tecnología para el enlace de las distintas BBDD
    - BBDD distribuidas heterogéneas: los SGBDs no son iguales para todas la BBDD locales, en estos casos las aplicaciones que consultan la base de datos distribuida realizan papel de mediador



# ALMACENAMIENTO DE LA INFORMACIÓN

- Bases de Datos y SGBD distribuidos (V)

- Técnicas de fragmentación:

El contenido de una tabla puede separarse en varios fragmentos que deben contener toda información como para poder reconstruir la tabla completa. Cada fragmento se encontrará en un nodo diferente, pero esta distribución a través de distintos nodos no implica duplicidad de los registros, solo existirá una copia de cada elemento en detrimento de la disponibilidad y fiabilidad de los datos en caso de caída de algún nodo si el diseño no es el adecuado

La fragmentación puede ser:

- **Horizontal**: los fragmentos son subconjuntos de una tabla y se definen a través de una operación de selección (condiciones sobre determinados atributos). La tabla original se reconstruye en base a una operación de unión de los fragmentos componentes
- **Vertical**: los fragmentos son subconjuntos de los atributos con sus valores. Para poder recomponer la tabla original, cada fragmento debe incluir la clave primaria de la tabla. Es útil por ejemplo en el caso de tener un conjunto de atributos más accedidos, se pueden situar estos en un servidor de mayor capacidad de proceso
- **Mixta**: se mezclan las dos técnicas anteriores

