# ¿Qué es un fichero? (2.1)

**Fichero o archivo: conjunto** de **información** **relacionada**, tratada como un todo y **organizada** de forma estructurada. Es una **secuencia** de **dígitos** **binarios** que **organiza** **información** **relacionada** con un mismo aspecto.

|  |
| --- |
| Los ficheros están formados por **registros lógicos** que contienen **datos** **relativos** a un **mismo** **elemento** u objeto (datos de usuarios).  Los datos almacenados se pueden añadir, suprimir, actualizar o consultar individualmente en cualquier momento.  **Solo** se **puede** **llevar** a la **memoria** **principal** **partes** de los ficheros porque son **muy** **pesados**.  La **cantidad** de **información** que es **transferida** entre el **soporte** donde está almacenado el fichero, y la **memoria** **principal**, en una sola operación de lectura/grabación, **se** **llama** **registro físico**o **bloque**.  En una **operación** de lectura/grabación se **transfiere** un **bloque**, con **varios** **registros** lógicos del fichero, a estos registros dentro de un bloque se les llama **factor de blocaje**, y a la operación de agruparlos **bloqueo de registros.** |

**Tipos de ficheros (2.2)**

|  |  |  |  |
| --- | --- | --- | --- |
| Ficheros permanentes | | | |
| Contienen… | | Ejemplo | |
| Maestros | **Estado** **actual** de los datos que pueden **modificarse** desde la aplicación. Es la parte **central** de la aplicación, su **núcleo**. | | Usuarios de una plataforma educativa |
| Constantes | **Datos** **fijos** para la aplicación. **No** **suelen** ser **modificados** y se **accede** a ellos para **realización** de **consultas**. | | Códigos postales |
| Históricos | **Datos** que **fueron** **considerados** como **actuales** en un periodo o **situación** **anterior**. Se utilizan para la **reconstrucción** de **situaciones**. | | Usuarios que han sido dados de baja en la plataforma educativa |

|  |  |
| --- | --- |
| Ficheros temporales | |
| Almacenan... | |
| Intermedios | Resultados de una aplicación que serán utilizados por otra |
| Maniobras | Datos de una aplicación que no pueden ser mantenidos en memoria principal por falta de espacio. |
| Resultados | Datos que van a ser transferidos a un dispositivo de salida. |

# Los soportes de información (2.3)

|  |
| --- |
| Los soportes más utilizados para almacenar los ficheros son:   * Las cintas magnéticas * Los discos (magnéticos, ópticos, o magneto-ópticos). |

|  |  |  |
| --- | --- | --- |
| Tipos de soportes | | |
| Acceso Directo a los datos | Son los **más** **usados**  Se puede **acceder** a los datos de **forma** **directa**  Podemos **colocarnos** en la **posición** que nos **interesa** y **leer** a partir de ella. | Discos |
| Acceso Secuencial | Usados para **copias** de **seguridad**  **No** se puede **acceder** a los datos de **forma** **directa**.  Para leer un dato, hay que **leer** **todo** lo que hay **hasta** **llegar** a esa **posición**. | Cintas magnéticas |

# Métodos de acceso (2.4-2.8)

|  |
| --- |
| Objetivos fundamentales de estas modificaciones |
| * **Proporcionar** un **acceso** **rápido** a los **registros**. |
| Conseguir **economizar** el **almacenamiento**. |
| **Facilitar** la **actualización** de los **registros**. |
| **Permitir** que la **estructura** refleje la **organización** **real** de la información. |

|  |  |
| --- | --- |
|  | |
| **Ficheros secuenciales** | **Definición:** Los **registros** están **almacenados** de **forma** **contigua**, para acceder a un registro hay que **leer** **todos** los registros **hasta** **llegar** a él.  **Características:**  La **lectura** siempre se realiza hacia delante.  Son **ficheros** **monousuarios**, **no** **permiten** el **acceso** **simultáneo** de varios usuarios.  Tienen una **estructura** **rígida** de campos. **Todos** los **registros** deben **aparecer** en **orden** y 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**, **al** **no** **dejar** **huecos** **vacíos**.  Se pueden **grabar** en **cualquier** **tipo** de **soporte**, secuenciales y direccionables.  Todos los **lenguajes** de programación pueden **trabajar** con este tipo de **ficheros**.  **No** se **pueden** **insertar** **registros** **entre** **los** que ya están **grabados**. |

|  |  |
| --- | --- |
| **Ficheros de acceso directo** | **Definición:** Se puede **acceder** a un **registro** indicando la **posición** **relativa** del mismo dentro del archivo o a través de una **clave** que forma parte del registro como un campo más. Estos archivos deben **almacenarse** en **dispositivos** de **memoria** masiva de **acceso** **directo**.  **Campo clave:** permite **identificar** y **localizar** un **registro** de manera **ágil** y **organizada**.  Cada uno de los **registros** se guarda en una **posición** **física**, que **dependerá** del **espacio** **disponible** en memoria masiva, de ahí que la distribución de los registros sea aleatoria dentro del soporte de almacenamiento. Para **acceder** a la **posición** **física** de un registro se **utiliza** una **dirección** o **índice.**  **Características:**  **Posicionamiento** **inmediato**.  Registros de **longitud** **fija**.  **Apertura** del fichero en **modo** **mixto**, para lectura y escritura.  Permiten **múltiples** **usuarios** utilizándolos.  Los registros se **borran** colocando un **cero** en la **posición** que **ocupan**.  Permiten la utilización de **algoritmos** de **compactación** de **huecos**.  Los **archivos** se crean con un **tamaño** **definido**, es decir, con un **máximo** de **registros** establecido durante la creación.  Esta organización sólo es posible en **soportes** **direccionables**.  **Se** **usan** 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 que más nos importa.  Permiten la **actualización** de los **registros** en el mismo fichero, **sin** **necesidad** de **copiar** el **fichero**.  Permiten realizar procesos de **actualización** en **tiempo** **real**. |

|  |  |
| --- | --- |
| **Ficheros indexados** | Utilizando **índices***,* permiten el acceso a un registro del fichero de forma directa.  Existe una **zona de registros** en la que se encuentran los datos del archivo y una **zona de índices,** que contiene una tabla con las claves de los registros y las posiciones donde se encuentran estos. La tabla de índices estará ordenada por el campo clave.  La **tabla de índices** se carga en **memoria principal** para **buscar** la fila correspondiente a la **clave del registro** a encontrar, obteniéndose así la **dirección** donde se **encuentra** el **registro**.  Una vez **localizada** la dirección, sólo hay que **acceder** a la **zona** de **registros** en el **soporte** de almacenamiento y **posicionarnos** en la **dirección** indicada. Puesto que la tabla debe prever la inclusión de todas las direcciones posibles del archivo.  Su principal **inconveniente** resulta **determinar** su **tamaño** y **mantenerla** **ordenada** por los valores de la clave.  **Características:**    El diseño del registro tiene que tener un campo, o combinación de campos, que permita identificar cada registro de forma única, es decir, que no pueda haber dos registros que tengan la misma información en él. A este campo se le llama campo clave y es el que va a servir de índice. Un mismo fichero puede tener más de un campo clave, pero al menos uno de ellos no admitirá valores duplicados y se le llama clave primaria. A las restantes se les llama claves alternativas.  **Permiten** **utilizar** el modo de **acceso** **secuencial** y el modo de **acceso** **directo** para leer la información guardada en sus registros.  Utilizando el modo de acceso directo:  Se hace **conociendo** el **contenido** del **campo** **clave** del **registro** que queremos localizar. Con esa información el **sistema** **operativo** puede **consultar** el **índice** y **conocer** la **posición** del **registro** dentro del **fichero**.  Utilizando el modo de acceso secuencial:  Los **registros** son **leídos** **ordenados** por el **contenido** del **campo** **clave**, **independientemente** **del** **orden** en que se **fueron** **grabando** (el orden lógico no es igual al orden físico), debido a que el **acceso** a los **datos** se hace **a** **través** del **índice**, que para **hacer** **más** **fácil** la **búsqueda** de los **registros**, permanece **siempre** **ordenado** por el **campo** **clave**.  **Solamente** se puede **grabar** en un **soporte** **direccionable.** |

|  |  |  |
| --- | --- | --- |
|  | | |
| **Ficheros Secuenciales Indexados** |  |  |
| **Ficheros de Acceso Calculado o Hash** |  |  |

# Parámetros de utilización (2.9)

|  |  |  |
| --- | --- | --- |
|  | | |
| **Capacidad o volumen** | Espacio, en caracteres, que ocupa el fichero. La capacidad podrá calcularse multiplicando el número previsto de registros por la longitud media de cada registro | |
| **Actividad** | Permite conocer la cantidad de consultas y modificaciones que se realizan en el fichero | |
| **Tasa de consulta o modificación** | 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 | |
| **Tasa de renovación** | Tanto por ciento de registros renovados en cada tratamiento del fichero, respecto al número total de registros contenidos en él |
| **Frecuencia de renovación** | Número de veces que se accede al fichero para renovarlo en un periodo de tiempo fijo |
| **Crecimiento** | Variación de la capacidad del fichero y se mide con la tasa de crecimiento, que es el porcentaje de registros en que aumenta el fichero en cada tratamiento | |

# Conceptos (3.1)

**Base de datos: Es una colección de datos relacionados lógicamente entre sí, con una definición y descripción comunes y que están estructurados de una 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.**

|  |  |  |
| --- | --- | --- |
| Elementos de una Base de Datos | |  |
| **Entidades** | **Objeto** **real** o **abstracto** con **características** **diferenciadoras** de otros, del que se **almacena** **información** en la base de datos. | En una clínica veterinaria, podría ser: ejemplar, doctor, consulta |
| **Atributos** | **Datos** que se **almacenan** **de** la **entidad**. **Cualquier** **propiedad** o **característica** de una **entidad** puede ser atributo. | Raza, color, nombre, número de identificación |
| **Registros** | **Donde** se **almacena** la **información** de **cada** **entidad**. Es un **conjunto** de **atributos** que **contienen** los **datos** que **pertenecen** a una **misma** **repetición** de entidad. | 2123056, Sultán, Podenco, Gris, 23/03/2009 |
| **Campos** | **Donde** se **almacenan** los **atributos** de **cada** **registro**. | El valor Podenco |

|  |  |
| --- | --- |
| Ventajas | |
| **Acceso múltiple** | **Varios** **usuarios** o **aplicaciones** podrán **acceder** a la **base** de **datos**, **sin** que existan **problemas** en el acceso o los datos |
| **Utilización múltiple** | **Cada** **uno** de los **usuarios** o **aplicaciones** podrán disponer de una **visión** **particular** de la **estructura** de la **base** de **datos**, de tal manera que **cada** **uno** de ellos **accederá** **sólo** a la **parte** que **realmente** le **corresponde**. |
| **Flexibilidad** | La **forma** de **acceder** a la **información** puede ser **establecida** de **diferentes** **maneras**, ofreciendo **tiempos** de **respuesta** muy **reducidos**. |
| **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 (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** |
| **Interfaz de alto nivel** | **Mediante** la **utilización** de **lenguajes** de **alto** **nivel** puede **utilizarse** la **base** de **datos** de **manera** **sencilla** |
| **Consulta directa** | **Existe** una **herramienta** para poder **acceder** a los **datos** **interactivamente** |

# Usos (3.2)

|  |  |
| --- | --- |
| Ventajas | |
| **El administrador** | **Persona** **encargada** de la **creación** o **implementación** **física** de la **base** **de** **datos**. Escoge los **tipos** de **ficheros**, los **índices** que, la **ubicación** de éstos, también, establecerá la **política** de **seguridad** y de acceso para **garantizar** el **menor** **número** de **problemas**. |
| **Los diseñadores** | **Personas** **encargadas** de **diseñar** cómo será la **base** de **datos**.  **Identificarán** los **datos**, las **relaciones** entre ellos, sus **restricciones**...  El **diseñador** de la base de datos debe **implicar** en el **proceso** a **todos** **los** **usuarios** de esta. |
| **Los programadores de aplicaciones** | Se **encargan** de **implementar** los **programas** de **aplicación** que **servirán** a los **usuarios** **finales**.  **Programas** que **realizan** **consultas** de **datos**, **inserción**, **actualización** o **eliminación** de los mismos.  Para **desarrollar** estos **programas** se utilizan [**lenguajes** de **tercera** o **cuarta** **generación**](https://moodlefp.educacion.navarra.es/pluginfile.php/83432/mod_resource/content/2/BD01_Contenidos_Imprimible/index.html#t149edf5d-e218-6122-0d61-973b85aaceb0). |
| **Los usuarios finales** | Son los **clientes** **finales** de la **base** de **datos**.  Al diseñar, implementar y mantener la base de datos se busca cumplir los requisitos establecidos por el cliente para la gestión de su información. |

# Ubicación de la información

|  |  |
| --- | --- |
| Ventajas | |
| **Discos SATA** | **Interfaz** de **transferencia** de **datos** entre la **placa** **base** y algunos **dispositivos** de **almacenamiento:**  Disco duro, lectores y grabadores de CD/DVD/BD, Unidades de Estado Sólido...  SATA proporciona:   * **Mayores** **velocidades** * **Mejor** **aprovechamiento** cuando hay **varias** **unidades** * **Mayor** **longitud** del **cable** de **transmisión** de **datos** * **Capacidad** para **conectar** **unidades** al **instante**, es decir, insertar el dispositivo **sin** tener que **apagar** el **ordenador**   SATA 150 MB/s SATA II 300 MB/s SATA III 600 MB/s. |
| **Discos SCSI** | **Interfaces** para **discos** **duros** de **gran** **capacidad** de **almacenamiento** y **velocidad** de **rotación**.  Su **tiempo** **medio** de **acceso** puede llegar a **7 milisegundos** y su velocidad de **transmisión** **secuencial** de información es de:   * SCSI Estándar (Standard SCSI) - 5 MB/s en los discos * SCSI Rápido (Fast SCSI) - 10 MBps en los discos * SCSI Ancho-Rápido (Fast-Wide SCSI) - 20 MBps en los discos   Un controlador SCSI puede manejar hasta 7 discos duros SCSI. |
| **RAID** | Acrónimo de **Redundant** **Array** of **Independent** **Disks** o matriz de discos independientes,  Es un **contenedor** de **almacenamiento** **redundante**.  Se basa en el **montaje** en **conjunto** de **dos** **o más** **discos** duros, **formando** un **bloque** de trabajo, para **obtener** desde una **ampliación** de **capacidad** a **mejoras** en **velocidad** y **seguridad** de **almacenamiento**.  **Según** las **características** que queramos primar, se establecen **distintos** **sistemas** de RAID. |

|  |  |
| --- | --- |
| **Sistemas NAS** | Acrónimo de **Network** **Attached** **Storage** ó sistema de almacenamiento masivo en red.  Estos sistemas de almacenamiento **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 **sistema** **operativo** optimizado para **dar** **acceso** a los **datos** **a** **través** de **protocolos** de **comunicación** **específicos**.  Suelen ser **dispositivos** para **almacenamiento** **masivo** de datos con **capacidades** **muy** **altas**, de varios Terabytes, generalmente superiores a los discos duros externos y además se **diferencian** de estos al **conectar** por **red**. |
| **Sistemas SAN** | Acrónimo de **Storage** **Area** **Network** o red de área de almacenamiento. Se trata de una **red** **concebida** para **conectar** **servidores**, matrices (arrays) de discos y librerías de soporte.  La arquitectura de este tipo de sistemas **permite** que los **recursos** de **almacenamiento** estén **disponibles** para **varios** **servidores** **en** **una** **red** de **área** **local** o **amplia**.  Debido a que la **información** **almacenada** **no** **reside** directamente en **ninguno** de los **servidores** de la **red**, se **optimiza** el poder de **procesamiento** **para** **aplicaciones** **comerciales** y la **capacidad** de **almacenamiento** se puede **proporcionar** en el **servidor** donde **más** se **necesite**. |