

TEMA 1 (PRIMERA PARTE). SISTEMAS CLÁSICOS DE GESTIÓN DE FICHEROS.

1.1.INTRODUCCIÓN.

1.2.CONCEPTOS BÁSICOS.

1.3.ALMACENAMIENTO DE LA INFORMACIÓN.

1.4.EVOLUCIÓN DE LA GESTIÓN DE DATOS.

1.5.LOS SISTEMAS GESTORES DE FICHEROS CONVENCIONALES.

1.6.BASES DE DATOS.

1.1. INTRODUCCIÓN.

El almacenamiento y manejo de grandes cantidades de datos se hace necesario **en cualquier empresa para el logro de sus objetivos**. Por ejemplo, se necesitan los datos de los empleados, de los clientes, de los proveedores, de los productos almacenados, etc.

Normalmente la gestión de estos datos **se ha venido realizando de forma manual**. Se organizaban en forma de **fichas**, informes o expedientes, colocándolos en **carpetas** y almacenando estas en un **archivador**. Por ejemplo, se tiene un archivo de clientes en el que cada ficha contiene todos los datos correspondientes a un cliente. Cuando se necesita consultar o modificar los datos de un cliente concreto será necesario realizar la operación manualmente.

La utilización de los ordenadores en la administración de las empresas supuso una revolución respecto al almacenamiento y gestión de sus datos dando lugar a los denominados: *ficheros de datos informáticos* y posteriormente a las *bases de datos*.

Las ventajas se derivan de las características de los soportes y de los ordenadores:

- ✓ Gran **capacidad** de almacenamiento en un reducido espacio.(tamaño)
- ✓ **Rapidez** en el proceso de los datos. (velocidad)
- ✓ **Precisión** de los resultados obtenidos del proceso. (calidad)

1.2. CONCEPTOS BÁSICOS.

Un **archivo** o **fichero** es una estructura de datos que reside en memoria secundaria, consistente en un conjunto de informaciones estructuradas en unidades de acceso denominadas **registros**, todos del mismo tipo y en número indeterminado. **Conjunto de elementos con características comunes (campos), elementos=registros, soportados en memorias auxiliares.**

Un **registro lógico** o simplemente registro, es cada uno de los componentes del archivo, conteniendo un conjunto de informaciones a la que se acceden y se tratan de manera unitaria. Está constituido por uno o más elementos denominadas **campos** que pueden ser de diferentes tipos.

Un registro puede tener un **campo clave**, cuyo valor sirve para identificar de forma única el registro, y por tanto, dicho valor no puede aparecer repetido en otro registro diferente. **(Univoca; un registro una clave, una clave un registro, ejemplo dni)**

Si un fichero contiene la información de un conjunto de individuos u objetos, entonces sus registros contienen la información de cada uno de ellos y los campos los datos de esta.

Ejemplo: Fichero de personal de una Empresa.

Registro Personal:

Nº D.N.I.	2.198.365
Nombre	Luis Ortiz Rubio
Dirección	Olmos, 16
Localidad	28099 Madrid
Departamento	Producción
Categoría Laboral	Técnico Auxiliar
Fecha de Ingreso	15-9-99

Un **registro físico** o bloque corresponde a la cantidad de información que se transfiere físicamente en cada operación de acceso (lectura o escritura).

Comentar el factor de bloque (numero de bytes...etc). El tamaño y formato del registro lógico los define el programador, mientras que el tamaño del registro físico viene dado por las características físicas del ordenador.

1.3. ALMACENAMIENTO DE LA INFORMACIÓN.

Existen dos formas básicas de almacenar los datos:

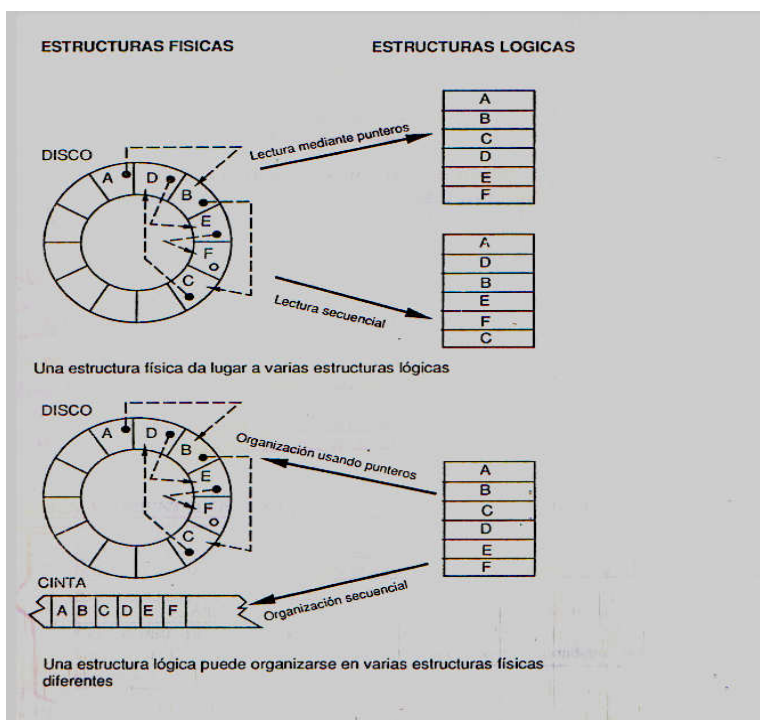
Convencional. Se utilizan los ficheros planos y por tanto los Sistemas Clásicos de Gestión de Ficheros.

No Convencional. Se utilizan las Bases de datos y por consiguiente los Sistemas Gestores de Bases de Datos.

1.4. EVOLUCIÓN DE LA GESTIÓN DE DATOS.

ESTRUCTURAS FÍSICAS Y LÓGICAS DE DATOS

De la física a la lógica accedemos de dos formas según si usamos o no los punteros y de la lógica a la física idem, teniendo en cuenta que hablaremos de soportes directo o secuenciales (en desuso).



SISTEMAS ORIENTADOS HACIA EL PROCESO.

A. Primera etapa: Datos integrados en el programa.

- ✓ No se diferencia entre estructura física y lógica.
- ✓ Sólo se analizan parcelas muy pequeñas por lo general del mundo científico.

B. Segunda Etapa: Aparecen los ficheros secuenciales.

- ✓ Se separan los datos al aparecer los ficheros.
- ✓ Siguen representándose aún parcelas pequeñas del mundo real.

C. Tercera Etapa: Ficheros de Acceso Directo.

- ✓ Las estructuras físicas y lógicas pueden ser diferentes.
- ✓ Se siguen analizando parcelas en mayor número.

SISTEMAS ORIENTADOS HACIA LOS DATOS.

- ✓ Las estructuras lógicas y físicas se separan claramente.
- ✓ La descripción de la estructura lógica se separa del lenguaje de programación.
- ✓ El concepto de relación pasa a ocupar un lugar destacado y los Sistemas de Bases de Datos facilitan instrumentos para describirlas y almacenarlas.

1.5. LOS SISTEMAS GESTORES DE FICHEROS CONVENCIONALES.

CLASIFICACIÓN DE LOS FICHEROS SEGÚN SU USO

Los archivos se clasifican según la utilización que se hace de ellos en tres grandes grupos:

ARCHIVOS PERMANENTES. Contiene información que varía poco a lo largo del tiempo. Pueden ser de tres clases:

- ✓ **Archivos Constantes.** Su información permanece prácticamente inamovible, utilizándose principalmente como archivos de consulta.

Ejemplo: Un archivo de este tipo puede ser el de la red de metro de una ciudad, que contiene la descripción, características, número de estaciones, numero de trenes, etc., de cada línea.

- ✓ **Archivos de Situación.** También denominados archivos maestros, contienen la información que refleja el estado o situación de una empresa en un determinado momento. Estos archivos se actualizan periódicamente para adaptarlo a cada nuevo situación.

Ejemplo: Un archivo de personal con contrato temporal en una empresa, o el archivos de existencias de un almacén.

- ✓ **Archivos Históricos.** Se obtienen de los anteriores cuando se dejan fuera de uso para futuros estudios estadísticos o consultas.

Ejemplo: El archivo que contiene la información de libros adquiridos por una biblioteca en la década de los ochenta.

ARCHIVOS DE MOVIMIENTO. En ellos se almacena la información que se utilizará para actualizar los archivos maestros. Sus registros, denominados movimientos o transacciones, son de tres clases: altas, baja y modificaciones.

Una vez realizado el proceso de actualización de un archivo de maestro por medio de un archivo de movimientos, éste pierde su validez y podemos deshacernos de él

Ejemplo: Un archivo de este tipo para actualizar el antes mencionado de personal contratado, es el que refleja las nuevas contrataciones, finalizaciones de contrato y modificaciones de los mismos producidas en la empresa durante el mes actual.

ARCHIVOS DE MANIOBRA O TRABAJO. Tienen una vida limitada, normalmente igual a la duración de la ejecución de un programa, y se utilizan como auxiliares de los anteriores.

Ejemplo: Se desea una lista alfabética de los nombres del personal contratado, se hará por medio de un archivo de maniobra en el que se almacene esta información a partir del archivo de personal. Este archivo desaparecerá una vez que tenga la lista impresa.

ORGANIZACIÓN DE LOS FICHEROS

Al diseñar un archivo, dependiendo del uso que se va a hacer del mismo y del soporte utilizado, se pueden elegir diferentes maneras de organizar sus registros, siendo sus principales organizaciones las siguientes:

- ✓ Secuencial
- ✓ Directa o Aleatoria
- ✓ Secuencial Indexada

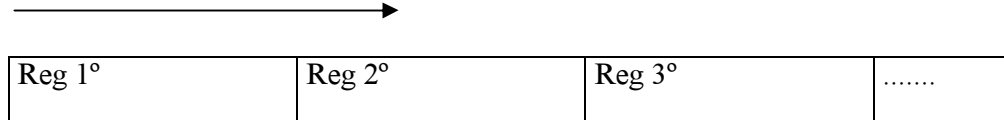
Organización Secuencial.

Es aquella en la cual los registros ocupan posiciones consecutivas de memoria, y sólo se puede acceder a ellos de uno en uno a partir de primero.

No se pueden hacer operaciones de escritura cuando se está leyendo, ni operaciones de lectura cuando se está escribiendo.

Para actualizarlos es preciso crear nuevos archivos donde se copien los registros que vayan a permanecer, modificados o no junto con los nuevos.

Acceso Secuencial



Organización Directa o Aleatoria

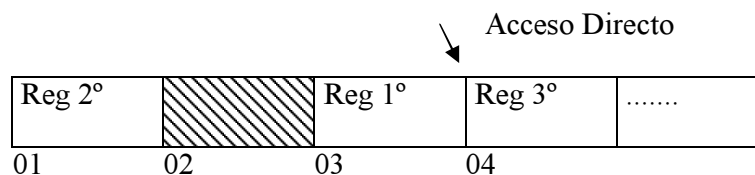
Esta organización también se denomina **relativa**. En estos archivos las informaciones se colocan y se acceden aleatoriamente mediante su posición,

es decir, indicando el lugar relativo que ocupan dentro del conjunto de posiciones posibles.

En esta organización se pueden leer y escribir registros, en cualquier orden y cualquier lugar.

Presenta el inconveniente de que es tarea del programador establecer la relación entre la posición que ocupa el registro y su contenido, además puede desaprovecharse parte del espacio destinado al archivo, ya que pueden quedar huecos libres entre unos registros y otros.

Su principal ventaja es su rapidez de acceso a un registro cualquiera, ya que para ello no es necesario pasar por los anteriores.



Organización Secuencial Indexada.

Un archivo de esta organización consta de tres áreas:

Área de Índices.

Área Primaria

Área de Excedentes (overflow)

El **Área Primaria** contendrá los registros de datos, clasificados en orden ascendentes por su clave.

El **Área de Índices** es un archivo secuencial creado por el sistema, en el que cada registro establece una división (segmento) en el área primaria, y contiene la dirección del comienzo del segmento y la clave más alta del mismo. De esta manera el sistema accede de forma directa a un segmento del área

primaria a partir del área de índices, de forma similar a la búsqueda de un capítulo de un libro a partir de su índice.

Por último se reserva un espacio, llamado **área de excedentes**. Para añadir nuevos registros que no pueden ser colocados en el área primaria cuando se produce una actualización del archivo.

Ventajas:

Rápido acceso por medio de la clave del registro

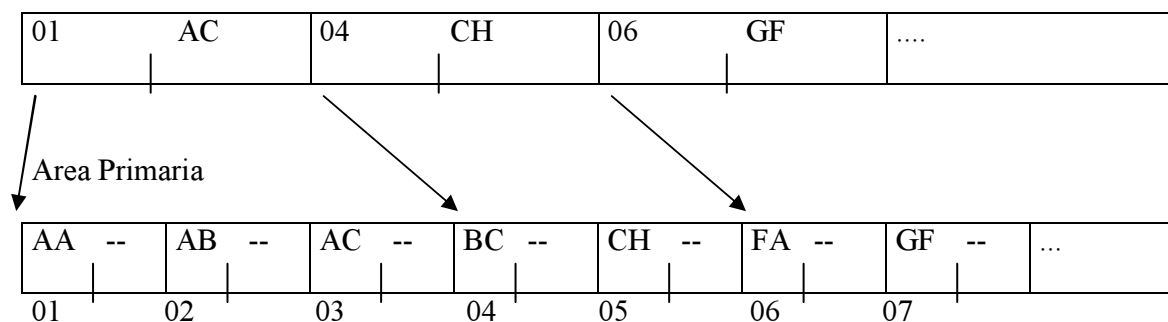
El sistema se encarga de relacionar la posición de cada registro con su contenido por medio del área de índices.

También es trabajo del sistema la gestión del área de índices y excedentes.

Inconvenientes:

Necesidad de espacio adicional para el área de índices

Área de Índices



Área de Excedentes

FH	--	AN	--
----	----	----	----	-------

MODOS DE ACCESO

Se denomina **modo de acceso** a la forma en que un dispositivo que maneja un soporte de información, que contiene un archivo, se posiciona en un

determinado lugar del mismo para realizar una operación de lectura o de escritura de un registro.

El modo de acceso lo decide el programador de la aplicación en función del tipo de organización.

Hay dos modos básicos: Secuencial y Directo.

Acceso Secuencial. Supone acceder inicialmente al primer registro del archivo y después, consecutivamente, a todos los sucesivos hasta llegar al registro deseado. Este modo de acceso se puede utilizar con cualquier organización.

Acceso directo. Consiste en el posicionamiento sobre cualquier registro sin necesidad de haber accedido antes a los anteriores.

Este modo de acceso se puede utilizar con la organización directa y con la indexada.

En los *archivos de organización directa*, este acceso se consigue proporcionando al dispositivo la posición del registro que se desea acceder. Normalmente se programa una función de aleatorización o hashing, la cual permite calcular la posición de cualquier registro a partir del valor de su clave.

En los *archivos indexados*, el acceso directo lo realiza de forma automática el sistema a partir del valor del campo clave; para ello utiliza, como se ha descrito en el apartado anterior, un conjunto auxiliar de índices que relacionan las claves con las posiciones de los registros.

OPERACIONES SOBRE LOS FICHEROS

Entre las operaciones más usuales que se realizan con archivos destacan la creación, copia, consulta, clasificación, concatenación, intersección, fusión, partición, actualización, reorganización, y borrado.

Creación: Consiste en la escritura o grabación en un soporte determinado de todos los registros que van a formar el archivo. Los datos pueden ser

introducidos desde teclado, pueden proceder de otro archivo, o ser obtenido como resultado de algún proceso.

Copia: Consiste en crear un nuevo archivo como duplicación de otro existente. La copia puede realizarse en el mismo o en diferente soporte de información. Un caso particular de esta operación es la impresión de un archivo (copia en impresora).

Consulta: Normalmente irá precedida de una búsqueda, se utiliza para obtener el contenido de uno o varios registros.

Clasificación u ordenación: Consiste en reubicar los registros de tal forma que queden ordenados respecto a los valores de un campo que denominamos clave de ordenación. Por ejemplo, puede ser útil tener clasificado el archivo de alumnos por el campo nombre, ya que la mayoría de las consulta se hará utilizando este campo.

Concatenación: Dados dos archivos con registros de igual estructura, obtendremos uno sólo en el que figuren todos los registros del primero y a continuación todos los del segundo. Esta operación se puede generalizar para más de dos ficheros.

Intersección: Dados dos archivos con registros de igual estructura, obtendremos otro en el que figuren los registros comunes a ambos. Ejemplo: Si queremos saber quienes son los repetidores del primer curso, haremos la intersección del fichero de alumnos matriculados en el centro en el primer curso de este año y del fichero que contiene los del año anterior.

Fusión o mezcla: Dados dos archivos con registros de igual estructura y clasificados por el mismo campo, obtendremos uno sólo en el que figuren los registros de ambos y que mantiene la ordenación.

Partición: Consiste en descomponer un archivo en dos, atendiendo a alguna característica de sus registros. Ejemplo: Podemos realizar una partición del archivo de alumnos en dos, según el valor del campo EDAD; el primero contendrá los mayores de edad y el segundo los menores.

Actualización: Es la operación de modificar un archivo de situación por medio de un archivo de movimientos, conteniendo altas, bajas y modificaciones que hay que realizar sobre el archivo maestro para ponerlo al día.

Reorganización: Operación consistente en reubicar los registros de un archivo que ha sufrido actualizaciones, de tal manera que se ocupen los posibles huecos libres intermedios resultantes de bajas de registros para optimizar la ocupación de la memoria, liberando la que no estaba aprovechada.

Borrado: Eliminación total del archivo, cuando ya no se necesite, dejando libre el espacio de memoria que ocupaba en el soporte utilizado.

1.6. BASES DE DATOS.

Los sistemas de gestión de ficheros clásicos se han mostrado insuficientes para cubrir las necesidades de almacenamiento y proceso de datos en la actualidad.

Los sistemas de información de las empresas son cada vez más grandes y complejos, y su gestión mediante el uso de archivos presenta algunos inconvenientes:

Los datos están repartidos por distintos departamentos de la empresa, lo que supone tener que desplazarse o solicitarlos cada vez que se necesitan en un lugar diferente del que se generan.

Además pueden estar repetidos varias veces con el inconveniente de que su actualización dará lugar a inconsistencias si no se realiza en todos los lugares donde se ubican.

Las aplicaciones que se creen deberán tener en cuenta las distintas estructuraciones de los datos.

Al existir dispersión de los datos es difícil establecer sistemas de control y seguridad de los mismos.

Los sistemas de Bases de Datos pretenden dar solución a los problemas expuestos.

DEFINICIÓN DE BASE DE DATOS

Una colección de datos que están lógicamente relacionados entre sí, que tiene una definición y una descripción comunes y que están estructurados de una forma particular. Podemos decir que una Base de Datos es también, un modelo del mundo real, y como tal, debe poder servir para toda una gama de empleos y aplicaciones.

