

# Tema 6

## Ficheros

*Quizás los ficheros sean el aspecto del sistema operativo más cercano a los usuarios, ya que todo usuario está acostumbrado a trabajar con ellos. Sabemos que en los ficheros se almacenan los programas que utilizamos y los datos que éstos manejan.*

*En este tema se describen diversos aspectos básicos de los ficheros, tales como qué dispositivos se utilizan para almacenarlos, qué tipos de ficheros existen y cómo se organizan y almacenan en los dispositivos de almacenamiento.*

### 6.1 Introducción

#### \* Sistemas de almacenamiento primario y secundario

- Sistema de almacenamiento primario: Está constituido por la memoria, que en los sistemas modernos es un dispositivo jerarquizado (caché y memoria principal). Su objetivo es almacenar los programas durante su ejecución. La memoria sólo almacena información mientras el computador está en funcionamiento.
- Sistema de almacenamiento secundario: Está constituido por diversos tipos de dispositivos, tales como discos duros, disquetes, dispositivos ópticos y memorias flash. Su objetivo es proporcionar un sistema de almacenamiento permanente de la información. Es decir, dichos dispositivos mantienen la información aunque no se encuentren en funcionamiento.

#### \* Organización de la información en los sistemas de almacenamiento secundario: concepto de sistema de ficheros

- En los sistemas de almacenamiento secundario la información se organiza utilizando dos tipos de entidades lógicas, los ficheros y los directorios.
- Se conoce con el nombre de **sistema de ficheros** a la estructura organizativa utilizada para organizar la información en un dispositivo de almacenamiento secundario mediante ficheros y directorios.
- Ejemplos de aspectos organizativos definidos por un sistema de ficheros son los siguientes:
  - Reglas que deben cumplir los nombres de los ficheros y los directorios, como por ejemplo el número máximo de caracteres que pueden tener y el conjunto de caracteres que se permiten en su formación.
  - Tamaño máximo que puede alcanzar cada fichero.
  - Tamaño máximo que puede alcanzar la información total almacenada.

- Tipo de organización adoptada por los directorios (típicamente éstos se organizan mediante una estructura de árbol.)
- Método utilizado para localizar la posición que ocupa cada fichero en el dispositivo de almacenamiento.
- Etc.
- La plataforma Windows utiliza diversos tipos de sistemas de ficheros (cada uno de ellos con sus características particulares), según el dispositivo de almacenamiento al que van orientados. Los principales son los siguientes:
  - NTFS: Sistema de ficheros para el disco duro.
  - CDFS: Sistema de ficheros para los CD-ROM.
  - FAT: Sistema de ficheros para disquetes y memorias USB.
- El gestor de ficheros del sistema operativo es el encargado de gestionar y proporcionar acceso al sistema de ficheros de cada dispositivo de almacenamiento presente en un sistema.

## 6.2 Ficheros

### \* Concepto de fichero

Un fichero es un conjunto de información relacionada, grabada en el sistema de almacenamiento secundario y a la que se hace referencia mediante un nombre.

*Ejemplos de información relacionada son el conjunto de instrucciones y datos de un programa, o los caracteres de un documento. Esta información relacionada es la que se almacena en ficheros.*

*Debe resaltarse la necesidad de dar un nombre a los ficheros, con objeto de que los usuarios y programas puedan identificarlos y así acceder a ellos.*

### \* Nombre de los ficheros

- Las reglas que definen la forma de dar nombre a los ficheros varían en función del sistema operativo.
- En muchos sistemas, el NTFS de Windows por ejemplo, el tamaño máximo admitido para el nombre de los ficheros es de 255 caracteres.
- Muchos sistemas (el NTFS por ejemplo) soportan nombres de dos partes, que se separan mediante un punto. La primera parte es el nombre en sí mismo. La segunda parte recibe el nombre de extensión, e indica el tipo de información almacenada en el fichero.
- Ejemplos de extensiones de ficheros:
  - exe → Fichero ejecutable
  - c → Programa fuente en lenguaje C
  - cpp → Programa fuente en lenguaje C++
  - asm → Programa fuente en lenguaje ensamblador
  - html → Documento escrito en lenguaje de hipertexto

- jpg → Imagen codificada según el estándar JPEG
- mp3 → Sonido codificado según el estándar MPEG 3
- pdf → Documento en formato PDF (*Portable Document Format*)
- zip → Archivo comprimido.

#### \* Tipos de ficheros

Pueden organizarse en dos grandes categorías: ficheros de texto (o ASCII) y ficheros binarios.

- Ficheros de texto (o ASCII): Contienen líneas de texto representadas en código ASCII.
  - Los ficheros de programas fuente escritos en cualquier lenguaje de programación son ejemplos de ficheros ASCII. No obstante, estos ficheros pueden contener cualquier tipo de información.
  - La gran ventaja de este tipo de archivos es que son estándar en todo tipo de sistemas (MS-DOS, Windows, MacOS, Unix, etc.). Son manejados mediante un tipo de programa, estándar en todo sistema, conocido como *editor de textos*. (Bloc de notas y EDIT en Windows.)
  - Debido a su uso estandarizado, la práctica totalidad de herramientas ofimáticas (procesadores de texto, hojas de cálculo, gestores de bases de datos, etc.) tienen la capacidad de manejar este tipo de ficheros.
- Ficheros binarios: Son aquellos que no son ficheros ASCII. Ejemplos típicos de ficheros binarios son:
  - Ficheros ejecutables: Contienen programas en formato ejecutable, como por ejemplo, los ficheros “.exe” en Windows.
  - Ficheros de datos de aplicaciones: Por ejemplo los “.doc” de Word, los “.xls” de Excel, etc. Estos ficheros utilizan un formato que es específico de la aplicación que los genera.
  - Ficheros multimedia: Almacenan información multimedia, es decir, imagen (jpeg por ejemplo), sonido (mp3 por ejemplo) o vídeo (mpeg por ejemplo).

#### \* Operaciones sobre ficheros

Son implementadas mediante el gestor de ficheros. Operaciones típicamente soportadas son crear, borrar, abrir, cerrar, leer y escribir.

## 6.3 Directorios

*Un sistema de ficheros almacenado en un disco duro puede contener fácilmente miles de ficheros. Debido a ello, es esencial disponer de un mecanismo que permita establecer una organización de los ficheros en el sistema, de forma que el usuario pueda encontrar y manejar los ficheros sin dificultad. Dicho mecanismo es proporcionado por los directorios.*

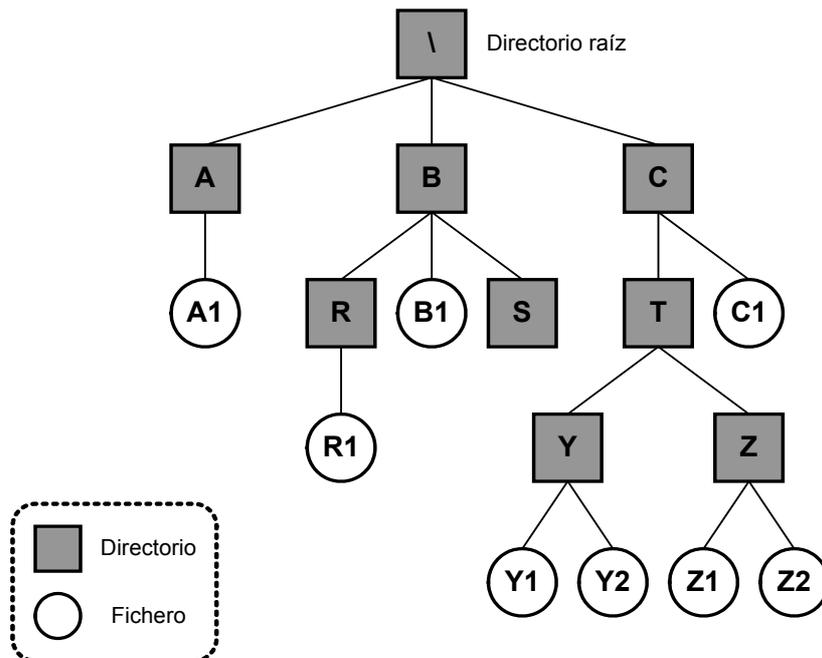
### \* Concepto de directorio

Es un fichero especial que contiene información sobre otros ficheros y directorios. Básicamente almacena los nombres de otros ficheros (o directorios) y la posición que ocupan en el sistema de ficheros.

*Como usuarios tenemos una percepción de los directorios como contenedores de ficheros o de otros directorios, pero más que esto son índices que permiten localizar a otros ficheros o directorios.*

### \* Organización de un sistema de ficheros mediante directorios

- En los sistemas de ficheros modernos los directorios se organizan en forma de árbol.
- Ejemplo de árbol de directorios



- Especificación del nombre de un archivo: sendas (*path names*)

*Cuando el sistema de ficheros está organizado mediante un árbol de directorios, para especificar un fichero no basta con indicar su nombre, ya que podría haber muchos ficheros con el mismo nombre en distintos directorios del árbol. Entonces hay que indicar su nombre y la posición que ocupa en el árbol de directorios. De esta forma, no hay ambigüedad posible.*

*Hay dos mecanismos para indicar la posición que ocupa un fichero en el árbol de directorios.*

- Sendas absolutas: Permiten especificar un archivo mediante la senda que hay que recorrer desde el directorio raíz hasta la posición que ocupa el archivo en el árbol. Para separar los directorios que integran la senda se utiliza un carácter de separación. En el caso de Windows, se utiliza el carácter '\

Ejemplo: Especificación del archivo Z1 del árbol de ejemplo: \C\T\Z1

- Sendas relativas: Se define un directorio activo en el sistema. Entonces la senda que especifica un archivo es relativa a dicho directorio.

*Debe indicarse que las sendas relativas se diferencian claramente de las absolutas en que no empiezan con el nombre del directorio raíz, ‘\’ en el caso de Windows.*

Ejemplo: Especificación del archivo Z1 siendo C el directorio activo: T\Z1

*Debe tenerse en cuenta que cuando se utiliza la interfaz de comandos gráfica de Windows, el usuario no necesita especificar sendas, ya que navega por el árbol de directorios utilizando las ventanas del explorador de archivos. Sin embargo, cuando utiliza la interfaz textual, o bien cuando en un programa se requiere abrir un archivo hay que utilizar sendas para especificar el archivo al que se desea acceder.*

#### \* Operaciones sobre directorios

Son implementadas mediante el gestor de ficheros. Operaciones típicamente soportadas son crear, borrar, abrir, cerrar.

## 6.4 El disco duro

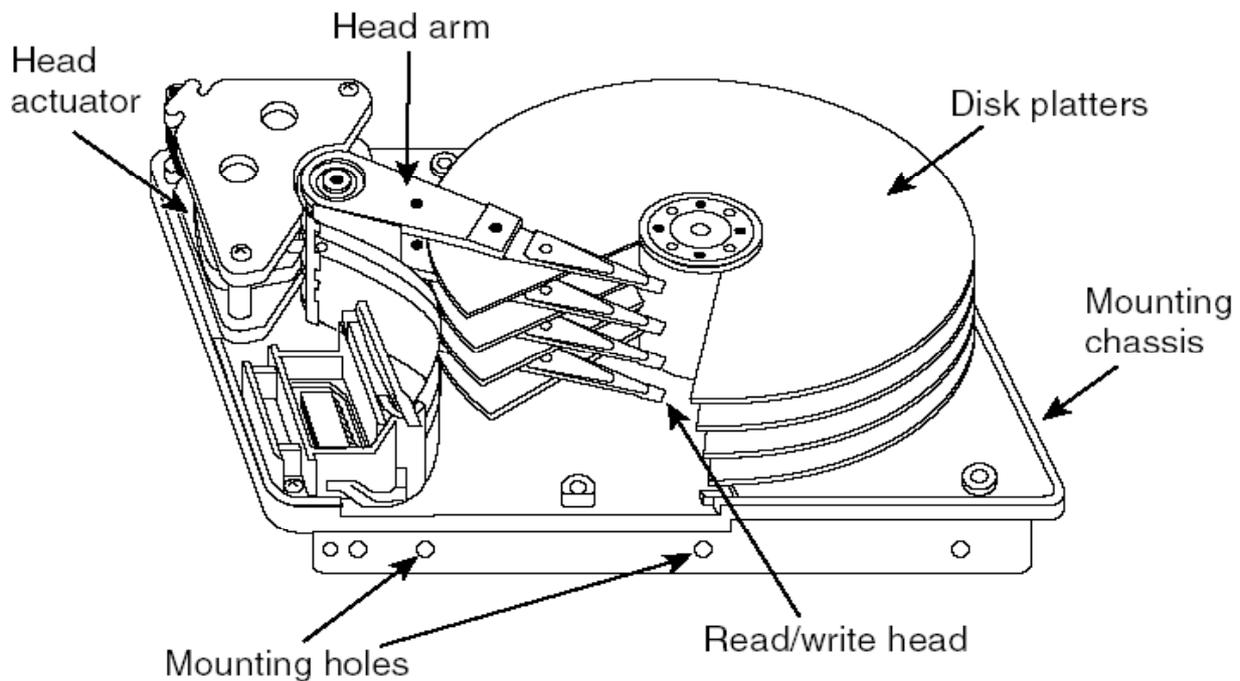
*Aunque ya se ha comentado previamente que hay diversos tipos de dispositivos de almacenamiento permanente, de todos ellos, el disco duro tiene una relevancia especial. Se trata del dispositivo de almacenamiento por excelencia debido a su gran capacidad y velocidad. Además, en un sistema común es el que almacena los ficheros del sistema operativo.*

#### \* Ejemplo de un disco duro real



*Las fotografías superiores muestran un disco duro real. Se trata de un modelo típico para PC. Tiene un diámetro de 3,5'', que es el mismo que un disquete. Se monta mediante tornillos en el bastidor interno de la caja y se conecta mediante un cable, usando su conector de datos, a una interfaz interna del computador que lo integra en el sistema de buses, para que pueda comunicarse con la CPU, memoria, etc.*

## \* Estructura física de un disco duro



*El disco está constituido por uno o más platos (disk platters) sobre los que se graba la información. Normalmente la información se graba en ambas caras de cada plato. Un motor está haciendo rotar los platos constantemente. Para escribir en el disco y leer información de él se utilizan las cabezas de lectura/escritura (read/write heads), que están soportadas por unos brazos (head arms). Éstos son movidos por un actuador (head actuator) en dirección radial. Entonces gracias a este movimiento y al de giro, una cabeza puede posicionarse sobre cualquier punto de la superficie de un plato.*

## \* Componentes de un disco duro

- **Platos (disk platters):** están formados por una aleación rígida de aluminio y recubiertos por una capa de material magnético sobre la que se graba la información.
- **Motor de giro:** su objetivo es hacer girar los platos a velocidad constante. Una velocidad típica es 7200 r.p.m.
- **Cabezas de lectura escritura (read/write heads):** su objetivo es escribir información sobre las superficies de los platos y leer de ellos. Hay una cabeza por cada superficie.
- **Brazo (head arm):** sirven de soporte a las cabezas.
- **Actuador (head actuator):** es un servomotor encargado de mover los brazos para posicionar las cabezas en las posiciones de los platos requeridas.

## \* Principios básicos de funcionamiento de un disco duro

- **Grabación:** La información se graba sobre el material magnético de la superficie de los platos. La grabación es realizada por la cabeza haciendo pasar una corriente eléctrica que magnetiza la superficie del plato situada debajo de ella.

- **Lectura:** Se basa en las variaciones de resistencia eléctrica que los campos magnéticos presentes en el plato inducen en la cabeza.
- **Posicionamiento de las cabezas:**
  - La combinación del movimiento de rotación de los platos y el movimiento radial de las cabezas sobre los platos hacen posible que una cabeza pueda llegar a cualquier punto de la superficie de un plato.
  - Las cabezas no tocan el disco mientras éste gira. El aire desplazado debido a la rotación de los platos separa las cabezas de éstos una distancia de 0,01 micras.

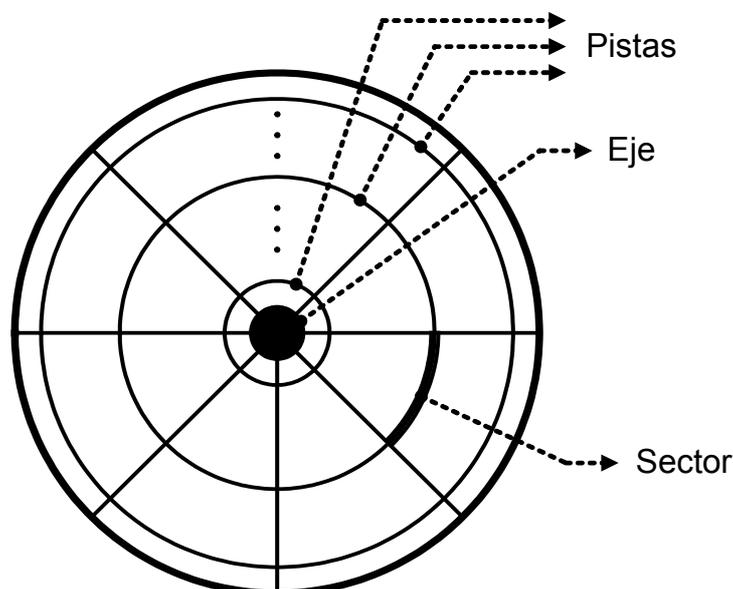
*Respecto al hecho de que las cabezas no toquen el disco, debe indicarse que esto es fundamental, ya que si no se produciría un efecto de abrasión que destruiría enseguida la superficie magnética. Hay que tener en cuenta la elevada velocidad a la que gira el disco.*

*Hay una zona del disco que se utiliza para el aparcamiento de las cabezas, es decir, sobre la que se dejan caer las cabezas cuando el disco deja de girar (por ejemplo al apagar el sistema). Esta zona se ubica al lado del eje de giro y en ella no se almacena información.*

#### \* Organización de la información en un disco duro

*Un disco duro, como todo dispositivo físico, es un dispositivo tridimensional. En el caso de los discos, sus tres dimensiones físicas dan lugar también a una estructura tridimensional de organización de la información. Veamos cómo es esa organización:*

- Dimensiones de organización de la información:
  - **Superficies:** un disco puede estar constituido por uno o más platos y en cada plato la información se graba en ambas superficies.
  - **Pistas:** cada superficie se organiza en pistas que son circunferencias concéntricas.
  - **Sectores:** cada pista se organiza en sectores. Un sector almacena 512 bytes. El sector es la unidad mínima de información que puede ser grabada en un disco o leída de él.
- Organización de la información en una superficie:



*La figura anterior muestra cómo se organiza una superficie en pistas y sectores.*

*Los ficheros se almacenan sobre los sectores de los discos. Un fichero pequeño ocupará pocos sectores y un fichero grande ocupará muchos sectores.*

*El sector es la unidad mínima de información manejada por el disco. Leer un byte de un fichero significa leer el sector entero en el que dicho byte se encuentra y luego acceder al byte concreto dentro del sector.*

\* Cálculo de la capacidad de almacenamiento de un disco duro (en bytes)

*La idea es multiplicar las dimensiones de las que hemos hablado antes*

• **Fórmula:**

$$\text{Capacidad} = \text{Superficies} \times \frac{\text{Pistas}}{\text{Superficie}} \times \frac{\text{Sectores}}{\text{Pista}} \times \frac{\text{Bytes}}{\text{Sector}}$$

• **Ejemplo:** determina la capacidad de almacenamiento expresada en bytes de un disco con 6 superficies, 10334 pistas por superficie y 63 sectores por pista.

$$\text{Capacidad} = 6 \times 10334 \times 63 \times 512 = 2000001024 \text{ bytes}$$

• **Expresión de la capacidad de almacenamiento:** actualmente los fabricantes de discos duros expresan la capacidad de estos dispositivos utilizando unidades basadas en potencias decimales en lugar de potencias binarias.

*En la tabla que se muestra a continuación se comparan las potencia decimales y binarias correspondientes a las unidades de almacenamiento más comúnmente utilizadas.*

	Potencias decimales	Potencia binarias
KB	$10^3$	$2^{10}=1024$
MB	$10^6$	$2^{20}=1024 \times 1024$
GB	$10^9$	$2^{30}=1024 \times 1024 \times 1024$
TB (Terabyte)	$10^{12}$	$2^{40}=1024 \times 1024 \times 1024 \times 1024$

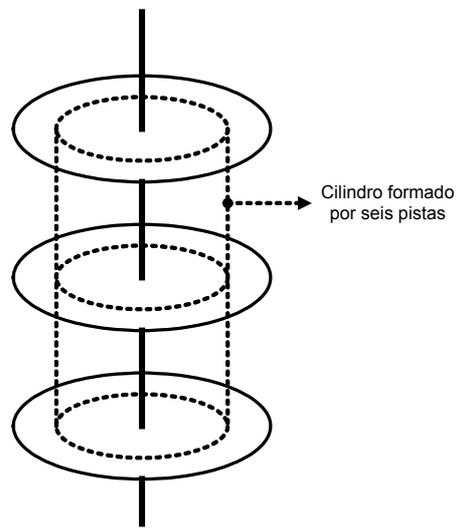
• Expresión en GB de la capacidad del disco del ejemplo anterior

$$\text{Capacidad (en GB)} = 2000001024 \text{ bytes} \times \frac{1 \text{ GB}}{10^9 \text{ bytes}} \cong 2 \text{ GB}$$

\* Nomenclatura utilizada por los fabricantes

• Concepto de cilindro: conjunto de todas las pistas que se ubican en la misma posición respecto al eje de giro.

*A continuación se muestra una figura representando el concepto de cilindro. En realidad, el número de cilindros de un disco coincide siempre con el número de pistas que hay en cualquiera de sus superficies.*



- Utilizan el término cilindro en vez de pista
- Utilizan el término cabeza den vez de superficie
- Ejemplo anterior: 6 cabezas, 10334 cilindros y 63 sectores

**NOTA:**

Los párrafos escritos sobre fondo gris y recuadrados mediante línea discontinua contienen información complementaria al resto del contenido de estos apuntes. De cara al examen, los conocimientos fundamentales que el alumno debe adquirir son los que se encuentran fuera de estos recuadros.