

Tema: Gestión del Sistema de Archivos

Objetivo:

- Analizar el funcionamiento y manejo del sistema de archivos y de los dispositivos de entrada/salida en un Sistema Operativo.

Concepto

Un sistema de gestión de archivos es aquel conjunto de software de sistema que proporciona servicios a los usuarios y aplicaciones en el uso de archivos. Típicamente, la única forma en la que un usuario o aplicación puede acceder a los archivos es a través del sistema de gestión de archivos. Esto elimina la necesidad de que el usuario o programador desarrolle software de propósito especial para cada aplicación. El sistema de archivos incluye:

- La colección de todos los archivos en el disco.
- Conjuntos de estructuras de datos utilizadas para administrar archivos (directorios de archivos, descriptores de archivos, tablas de asignación de espacio libre y ocupado en disco).
- Un complejo de software de sistema que implementa varias operaciones en archivos: creación, destrucción, lectura, escritura, nomenclatura, búsqueda.

Algunas operaciones del sistema de archivo incluyen:

- Crear
- Borrar
- Abrir
- Cerrar
- Leer
- Escribir

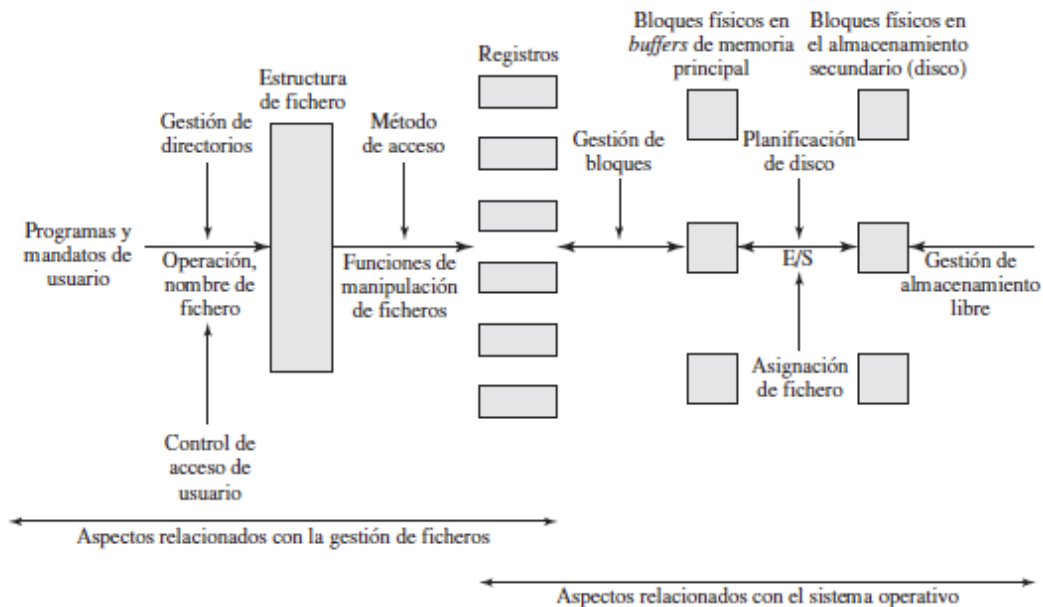


Figura 1. Organización y acceso a archivo

Estas tareas son complicadas en los sistemas operativos multitarea que están diseñados para funcionar como un solo usuario, pero proporcionan la capacidad de ejecutar múltiples procesos simultáneamente. Se agrega una nueva encomienda a las tareas enumeradas anteriormente: compartir un archivo de varios procesos. El archivo en este caso es un recurso compartido, lo que significa que el sistema de archivo (file system, FS) debe resolver todo el complejo de problemas asociados con dichos recursos. En particular: se deben proporcionar medios para bloquear el archivo y sus partes, hacer coincidir copias, evitar carreras, eliminar puntos muertos. En los sistemas multiusuario, surge otra tarea: proteger los archivos del acceso no autorizado de otro usuario.

Archivos

Nombre de archivo

Los usuarios necesitan poder referenciar un archivo mediante un nombre simbólico. Claramente, cada archivo en el sistema debe tener un nombre único a fin de que las referencias al mismo no sean ambiguas.

El uso de un directorio estructurado en forma de árbol minimiza la dificultad de asignar nombres únicos. Cualquier fichero del sistema se puede localizar siguiendo un camino desde el directorio raíz o maestro y bajando por las ramas hasta alcanzar el archivo. El conjunto de nombres de

directorios, finalizando en el nombre del archivo, constituye un nombre de camino para el archivo.

Estructura de un archivo

La estructura de un archivo involucra cuatro elementos, a saber: campo, registro, archivo, base de datos.

Campo. Es un elemento básico que contiene un valor único. Se caracteriza por su longitud y el tipo de datos (por ejemplo, ASCII, cadena de caracteres, decimal). Dependiendo del diseño del fichero, el campo puede tener una longitud fija o variable. En este último caso, el campo está formado normalmente por dos o tres subcampos: el valor real almacenado, el nombre del campo, y en algunos casos, la longitud del campo. En otros casos de campos de longitud variable, la longitud del campo se indica mediante el uso de símbolos de demarcación especiales entre campos.

Registro. Es una colección de campos relacionados que pueden tratarse como una unidad por alguna aplicación. Los registros también pueden ser de longitud fija o variable.

Archivo. Es una colección de campos relacionados. Se trata como una entidad única por los usuarios y por las aplicaciones.

Base de datos. Es una colección de datos relacionados. Los aspectos esenciales de una base de datos son que la relación que exista entre los elementos de datos sea explícita y que la base de datos se diseña para su uso por parte de varias aplicaciones diferentes. Una base de datos podría contener toda la información relacionada con una organización o proyecto, tal como información de negocio o de estudio científico. La base de datos está formada por uno o más tipos de archivos. Normalmente, hay un sistema de gestión de base de datos separado del sistema operativo, aunque hace uso de algunos programas de gestión de archivos.

Métodos de acceso

Los principales métodos de acceso a archivos son: secuencial, acceso directo, acceso indexado.

Acceso secuencial implica que sólo se leen los bytes en orden, empezando por el principio. No se puede saltar a una posición específica. Si hay que ir al inicio, se debe releer el archivo.

Acceso directo o aleatorio permite acceder a cualquier área o registro del archivo a partir de un puntero. Se asume que se puede mover aleatoriamente entre los distintos bloques que componen el archivo.

Acceso indexado. En el fichero indexado general, se abandonan los conceptos de secuencialidad y clave única. Los registros se acceden sólo a través de sus índices. El resultado es que no hay restricción en la colocación de los registros siempre que al menos un puntero en un índice se refiera a dicho registro. Además, se pueden emplear registros de longitud variable.

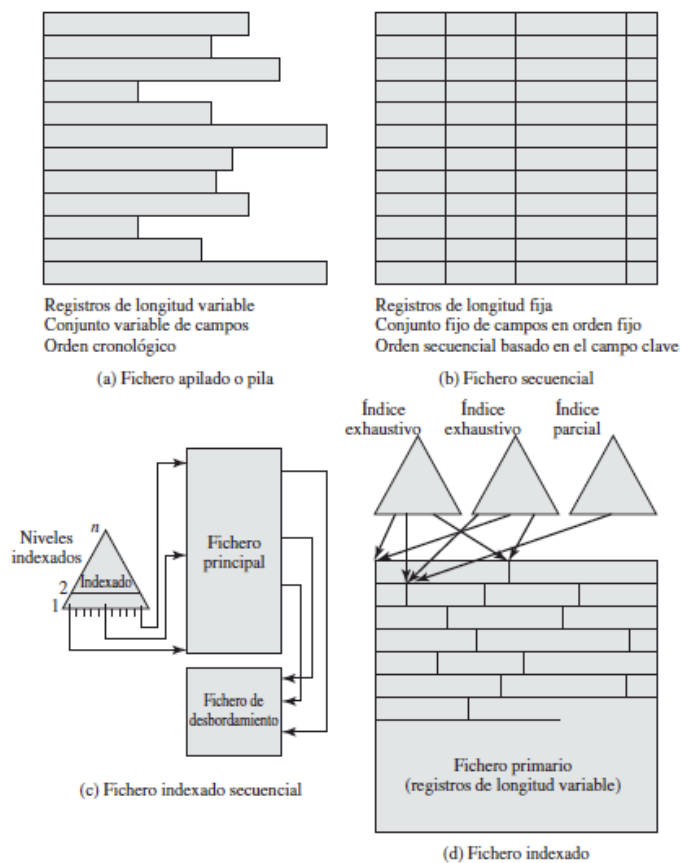


Figura 2. Organización y métodos de acceso

Directorios

Concepto

El directorio contiene información sobre los ficheros, incluyendo atributos, ubicación y propiedad. Gran parte de esta información, especialmente la que concierne a almacenamiento, la gestiona el sistema operativo. El directorio es a su vez un fichero, accesible por varias rutinas de gestión de ficheros. Aunque parte de la información de los directorios está disponible para los usuarios y las aplicaciones, esto se proporciona generalmente de forma indirecta por las rutinas del sistema.

Estructura

La forma más sencilla de estructura para un directorio es una lista de entradas, una por cada fichero. Esta estructura se podría representar como un fichero secuencial simple, con el nombre del fichero actuando como clave. En algunos sistemas iniciales monousuario, se ha utilizado esta técnica. Sin embargo, esta técnica es inadecuada cuando múltiples usuarios comparten el sistema o cuando un único usuario tiene muchos ficheros.

Una primera solución para resolver estos problemas sería pasar a un esquema de dos niveles. En este caso, hay un directorio por cada usuario y un directorio maestro. El directorio maestro tiene una entrada por cada directorio usuario, proporcionando información sobre dirección y control de acceso. Cada directorio de usuario es una lista simple de los ficheros de dicho usuario. Esto implica que los nombres deben ser únicos sólo dentro de la colección de los ficheros de un único usuario y que el sistema de ficheros puede fácilmente asegurar las restricciones de acceso de los directorios. Sin embargo, aún no proporciona ayuda a los usuarios para estructurar su colección de ficheros.

Una técnica más potente y flexible, que es casi universalmente adoptada, es utilizar una estructura jerárquica en forma de árbol. Como en la técnica anterior, hay un directorio maestro, que tiene bajo dicho directorio varios directorios de usuario. Cada uno de estos directorios de usuario, a su vez, podría tener subdirectorios y ficheros como entradas. Esto se cumple para todos los niveles: es decir, en cada nivel, un directorio podría estar formado por subdirectorios y/o ficheros.

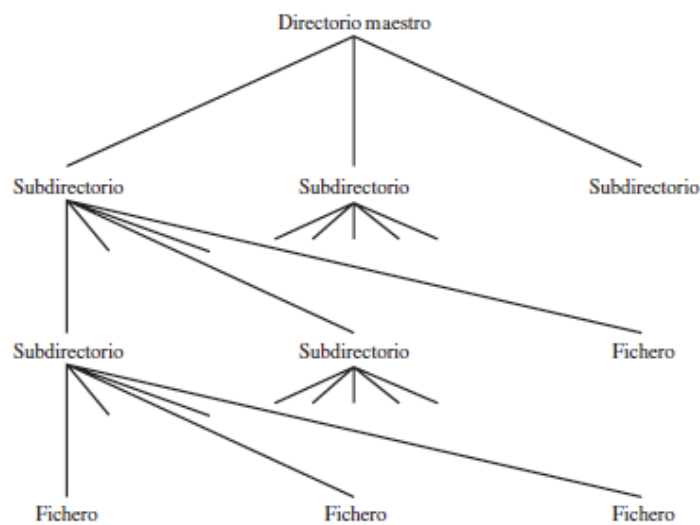


Figura 3. Estructura de directorio en forma de árbol

Nombres jerárquicos

El uso de un directorio estructurado en forma de árbol minimiza la dificultad de asignar nombres únicos. Cualquier archivo del sistema se puede localizar siguiendo un camino desde el directorio raíz o maestro y bajando por las ramas hasta alcanzar el archivo. El conjunto de nombres de directorios, finalizando en el nombre del archivo, constituye un nombre de camino para el archivo.

Se utiliza la barra (/) para delimitar nombres en una secuencia. El nombre del directorio maestro es implícito, porque todos los nombres comienzan en dicho directorio. Es aceptable tener varios archivos con el mismo nombre, siempre que ambos ficheros tengan nombres de camino únicos,

lo que equivale a decir que el mismo nombre de archivo se puede utilizar en diferentes directorios.

Aunque los nombres de camino facilitan la selección de los nombres de archivo, sería complicado para el usuario tener que escribir el camino completo cada vez que se hace una referencia a un archivo. Normalmente, un usuario interactivo o un proceso está asociado con un directorio actual, que se suele denominar directorio de trabajo.

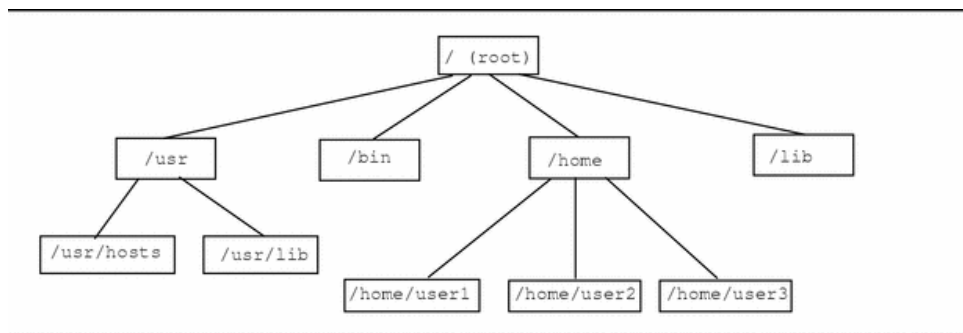


Figura 4. Jerarquía del sistema de archivos

Estructura y Almacenamiento del archivo y directorio

Los archivos por lo general se almacenan en discos. Existen dos estrategias para almacenar un archivo de n bytes. Esta misma situación se da en los sistemas de administración de memoria, caso segmentación y la paginación.

Asignación contigua: Al archivo se le asignan n bytes consecutivos de espacio de disco. Si el archivo crece (hay que tener en cuenta que los archivos son estructuras de datos con una alta volatilidad) probablemente tendrá que ser movido en el disco.

Asignación no contigua: El archivo se divide en m bloques de tamaño fijo que se almacenan en disco en bloques no necesariamente contiguos.

La segunda estrategia es más común. Una vez se ha decidido almacenar los archivos en bloques de tamaño fijo, se debe decidir el tamaño de éstos. El sector, la pista y el cilindro; y en el caso de un sistema de paginación, la página, son los referentes utilizados.

Después de decidir el tamaño del bloque, se debe considerar la forma de registrar los bloques libres. Una forma es llevar una lista ligada de bloques de disco en la cual cada uno contiene tanto bloques libres como sea posible. Otra técnica es llevar un mapa de bits. Los bloques libres se representan con el valor 1 en el mapa y los asignados con el valor 0. Algunos sistemas multiusuarios pueden establecer cuotas en el disco para evitar el uso excesivo de ciertos usuarios.

Sistemas de archivos

Un sistema de archivos son los métodos y estructuras de datos que un sistema operativo utiliza para seguir la pista de los archivos de un disco o partición; es decir, es la manera en la que se organizan los archivos en el disco. El término también es utilizado para referirse a una partición o disco que se está utilizando para almacenamiento, o el tipo del sistema de archivos que utiliza.

Algunos sistemas de archivos son:

- **Ext3.** El sistema de archivos ext3 posee todas las propiedades del sistema de archivos ext2. La diferencia es que se ha añadido una bitácora (journaling). Esto mejora el rendimiento y el tiempo de recuperación en el caso de una caída del sistema.
- **Vfat.** Esta es una extensión del sistema de archivos FAT conocida como FAT32. Soporta tamaños de discos mayores que FAT. La mayoría de discos con MS Windows son vfat.
- **Nfs.** Un sistema de archivos de red que permite compartir un sistema de archivos entre varios ordenadores para permitir fácil acceso a los archivos de todos ellos.

Referencia:

- Stallings, W. **Sistemas Operativos**. 5a. Edición. Editorial Prentice-Hall. (TEXTO). Marzo 2014.

Infografía:

Sistema de archivos: <https://sites.google.com/site/materiasisoperativo/unidad-5-sistemas-de-archivos/5-1-concepto>

Trabajar con archivos y directorios: <https://docs.oracle.com/cd/E19620-01/805-7644/6j76kloon/index.html>

Estructura de sistemas de archivos: <https://www.ibm.com/docs/es/aix/7.2?topic=tree-file-system-structure>