

Sistemas Operativos

Gestión del almacenamiento

Lic. R. Alejandro Mansilla

Licenciatura en Ciencias de la Computación
Facultad de Ingeniería
Universidad Nacional de Cuyo

Gestión del Almacenamiento: conceptos

- Generalmente el sistema de archivos es el aspecto más visible del sistema operativo.
- Muchos recursos se modelan como archivos.
- En CC se le llama **archivo** a un conjunto de información relacionada, definida por su creador y almacenada en la computadora.
- A la forma de organizar, nombrar, almacenar y manipular a los archivos se la denomina **sistema de archivos** (filesystem).
- campo: elemento de datos básico tal como nombre o fecha.
- registro: conjunto de campos relacionados que pueden tratarse como una unidad (entonces, un archivo será un conjunto de registros similares)
- bloque: es la unidad básica de una operación de entrada y salida entre el disco y los buffers de memoria del sistema de archivos (POSIX)

Gestión del almacenamiento: conceptos.

- El SO considera a un archivo como una colección de bloques lógicos.
- El disco en sí mismo es una colección de bloques físicos.
- Los archivos representan datos y programas. Son secuencias de bytes, líneas, registros, que tienen sentido para la aplicación y el usuario. No para el sistema operativo.
- Por eso, también se almacena información del tipo de archivo del cual se trata, bien dentro de él y/o como parte del nombre (*extensión*).
- Las extensiones pueden servir para indicar la estructura interna del archivo.
- Es común organizar a un grupo de archivos relacionados entre sí en un directorio o carpeta (*folder*). Como contenedores de archivos y subdirectorios, formando una estructura jerárquica.

Atributos

Los atributos de un archivo son los metadatos que describen o están asociados a un archivo: nombre, tipo, tamaño, ubicación, propietario, permisos, fechas de creación, acceso o modificación.

- El nodo-i (index node o i-node) de UNIX contiene información acerca del tipo de archivo (*normal o regular, directorio, enlace simbólico o link, tubería o pipe, archivo dispositivo, socket*), permisos o modos de acceso, el propietario, el grupo propietario, tamaño, fechas de creación de último acceso y de última modificación, y punteros a bloques de datos.
- El Master File Table o MFT de Windows describe el archivo mediante los siguientes atributos: cabecera, información estándar (por ejemplo fechas y conteo de enlaces), descriptor de seguridad, nombre de archivo y datos, el identificador de objeto (Object ID).

Atributos

- En MS-DOS la tabla de directorio (*directory table*) es un tipo especial de archivo que representa a un directorio. Cada archivo o directorio almacenado dentro de ella está representado por una entrada de 32 bytes en la tabla, cada entrada registra el nombre, extensión, los atributos (bits de archivado, directorio, oculto, sólo-lectura, sistema y volumen), la fecha y hora de creación, la dirección del primer cluster de los datos del archivo o directorio y el tamaño del archivo o directorio.
- La tabla de directorio está almacenada en la región de datos (Data Region) de la tabla de asignación de archivos (File Allocation Table o FAT).
- Una agrupación o cluster es un conjunto contiguo de sectores que componen la unidad más pequeña de almacenamiento de un disco.
- Son los bloques direccionables que usan muchos sistemas de archivos. El tamaño es siempre un múltiplo del tamaño del sector.
- Se utiliza el concepto de agrupación para administrar más eficientemente el espacio del disco.

Operaciones

Varían según los programas que las efectúan sobre los distintos sistemas operativos.
Están relacionadas con las distintas llamadas a sistema que ofrecen los SO:

- Crear
- Abrir
- Leer
- Buscar (seek)
- Escribir
- Actualizar
- Añadir (append)
- Truncar
- Trabar y destrabar
- Cerrar
- Colocar atributos o permisos
- Copiar
- Borrar
- Renombrar

Particiones

- Es un área separada de datos de un disco rígido físico.
- La tabla de particiones almacena el sector de comienzo de una partición, su tamaño y otras características y está ubicada en el mismo disco que la partición.
- Por cada HD puede haber una o más particiones.
- Cada partición puede tener un sistema de archivos diferente.
- Cada sistema de archivos puede alojar a un sistema operativo diferente.
- Cada SO puede usar más de una partición
- Se le llama volumen a objetos que representan sectores de una partición que los manejadores del sistema de archivos gestionan como una unidad.
- Los volúmenes multipartición son objetos que representan sectores tomados a partir de varias particiones y que los manejadores del sistema de archivos gestionan como una unidad.
- Existen gestores de volúmenes lógicos tales como Logical Disk Manager (LDM) de Veritas Software para Windows, y Logical Volume Manager (LVM) utilizado en Linux

Sistemas de archivos: directorio

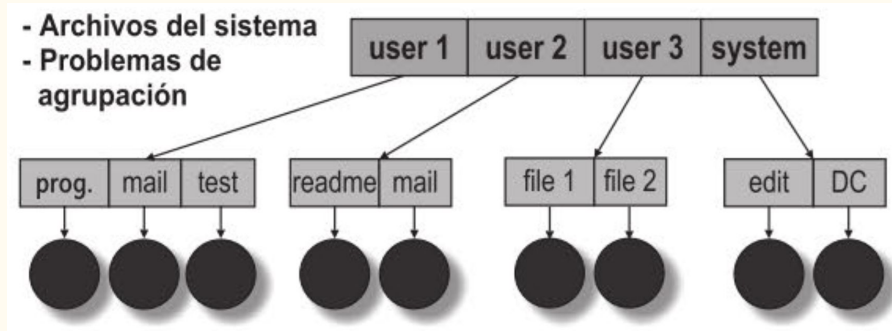
- Un directorio es un objeto que relaciona de forma unívoca el nombre dado por el usuario a un archivo y el descriptor interno del mismo usado por el sistema operativo.
- Un mismo nombre no puede identificar nunca a dos archivos distintos, aunque varios nombres se pueden referir al mismo archivo.
- Un directorio contiene tantas entradas como archivos son accesibles a través de él.
- Podemos pensarlo como una tabla que traslada los nombres de archivo en la referencia necesaria para ubicar el archivo dentro de la estructura.
- Cuando se abre un archivo, el SO busca en el sistema de directorios hasta que encuentra la entrada correspondiente al nombre del archivo.
- A partir de dicha entrada, el SO obtiene el identificador interno del archivo y algunos de sus atributos.
- Todas las referencias posteriores al archivo se resuelven a través de su descriptor (nodo-i, registro MFT)
- La estructura creada por la colocación de nombres en directorios puede tomar un número de formas: directorios de un nivel, de dos niveles, árboles de múltiples niveles, grafos acíclicos o grafos generales

Sistemas de archivos: directorios de un nivel

Código Usuario	Nombre Archivo	Tipo	Grado	Conteo	Nros de bloques de disco
1 byte	8 bytes	3 bytes	1 byte	1 byte	16 bytes

- En este sencillo enfoque todos los archivos están contenidos en un único directorio, es decir que no existe el concepto de subdirectorio
- Los archivos deben tener nombres únicos.
- Los SO CP/M (1975) y MP/M (1979) utilizaban este enfoque.

Sistemas de archivo: directorio de dos niveles



- Hay un directorio para cada usuario. Cada usuario tiene su UFD (*user file directory*).
- Existe un MFD (*master file directory*) que está indexado por nombre de usuario y apunta al UFD.
- Esta estructura aísla un usuario de otro, no pueden cooperar.
- El MFD se usó en el sistema ITS (Incompatible Timesharing System) a finales de la década del 60.

Directorios con estructura de árbol

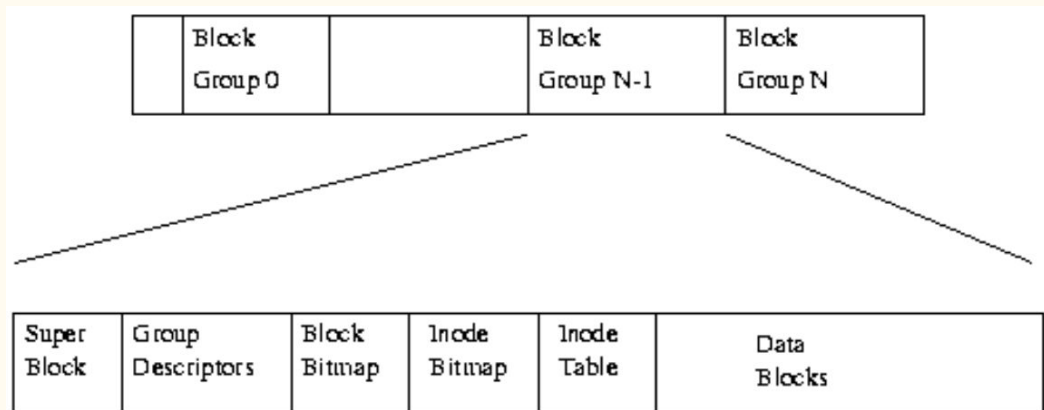
Nombre	Extensión	Atributos	Reservado	Tiempo	Fecha	1er. nro. bloque	Tamaño
8 bytes	3 bytes	1 byte	10 bytes	2 bytes	2 bytes	2 bytes	4 bytes

- Es la evolución lógica de los niveles vistos previamente. Un árbol multinivel permite el crecimiento del árbol más allá del segundo nivel.
- Cualquier directorio puede contener nombres de más directorios como así también objetos que no sean directorios.
- Puede utilizarse tanto en sistemas monousuarios como multiusuarios.
- La primera implementación de la estructura de árbol multinivel (*o sistema de archivos jerárquico*) data del SO Multics.

Directorio con estructura de grafo

- En esta estructura podemos tener un objeto con varios nombres.
- Como dos enlaces (*nombres*) llevan a uno de los objetos, la estructura deja de ser un árbol para pasar a ser un grafo.
- Si bien son más flexibles, las estructuras de nombres de grafos agregan complejidad.
- La solución de estos problemas es mucho más difícil cuando la estructura de nombres es un grafo que contiene ciclos.
- Cada vez que se crea un directorio se agregan dos nombres dentro del directorio nuevo:
- El punto “.”: es un alias para el directorio creado recientemente.
- Los dos puntos “..”: es un alias para el directorio padre del directorio creado recientemente.
- Los nombres “.” y “..” crean ciclos

Segundo sistema de archivos extendido (ext2)



El sistema de archivos EXT2 se basa en la premisa de que los datos contenidos en los archivos se guardan en bloques de datos. Estos bloques de datos son todos de la misma longitud y se establece cuando se lo crea. El tamaño de cada archivo se redondea hacia arriba a un número entero de bloques. Si el tamaño del bloque es de 1024 bytes, entonces un archivo de 1025 bytes ocupará dos bloques de 1024 bytes.

ext2 (*cont.*)

- No todos los bloques del sistema de archivos contienen datos, algunos deben utilizarse para contener la información que describe la estructura del sistema de archivos.
- EXT2 define la topología del sistema de archivos describiendo cada archivo del sistema con una estructura de datos de nodos índice (inodo).
- Un inodo describe qué bloques ocupan los datos dentro de un archivo, así como los derechos de acceso al mismo, los tiempos de modificación del archivo y el tipo de archivo.

ext2 - Entradas de directorio

- La información acerca de los archivos está almacenada en una entrada de directorio.
- La información almacenada depende del sistema operativo.
- En algunos sistemas, la entrada de directorio apunta a una estructura de archivo distinta en la que está almacenada la información acerca del archivo.

Numero de nodo-i	Nombre del archivo
2 bytes	X bytes

- En Linux la información del archivo está almacenada en el nodo-i.
- La entrada de directorio contiene sólo el nombre del archivo y su número de nodo-i, es decir que el directorio es un archivo que asocia nodos índice a nombres de archivo.
- Un nodo índice está asociado con cada archivo y mantiene toda la información referida a él excepto su nombre y sus datos.
- Otra estructura de datos importante en el sistema de archivos de Linux es el súper bloque (super block). Contiene información acerca de cada sistema de archivos montado.

ext2 - Superbloque

El súper bloque (superblock) contiene una descripción del tamaño y la forma básica del sistema de archivos. La información que contiene permite al administrador del sistema de archivos utilizarlo y mantenerlo. Normalmente sólo se lee el súper bloque del Grupo de Bloques 0 cuando el sistema de archivos está montado, pero cada Grupo de Bloques contiene una copia duplicada en caso de corrupción del sistema de ficheros.

ext2 - descriptor de bloque

Cada grupo de bloques tiene una estructura de datos que lo describe. Al igual que el superbloque, todos los descriptores de grupo para todos los Grupos de bloque se duplican en cada Grupo de bloque en caso de corrupción del sistema de archivos. Cada descriptor de grupo contiene la siguiente información:

- Bloques de mapa de bits: El número de bloque del mapa de bits de asignación de bloque para este grupo de bloques. Esto se usa durante la asignación de bloques y la desasignación,
- Inode Bitmap: El número de bloque del mapa de bits de asignación de inodo para este grupo de bloques. Esto se usa durante la asignación de inodo y la des-asignación,
- Tabla de inodo: El número de bloque del bloque inicial para la tabla de inodo para este grupo de bloques.
- Conteo de bloques libres, conteo de Inodes libres, conteo de directorios usados

Los descriptores de grupo se colocan después de otro y juntos forman la *tabla de descriptores de grupo*. Cada grupo de bloques contiene la tabla completa de descriptores de grupo después de su copia del superbloque. Solo la primera copia (en el Grupo de bloques 0) es utilizada realmente por el sistema de archivos EXT2. Las otras copias están allí, como las copias del superbloque, en caso de que la copia principal esté dañada

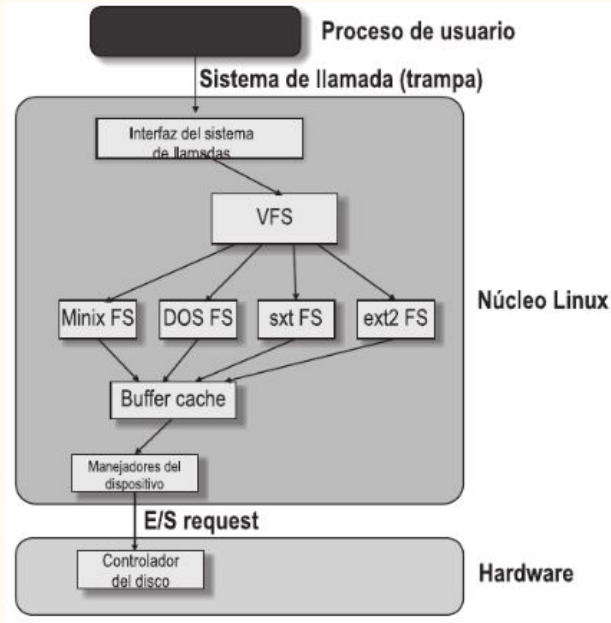
ext2 - directorios

En el sistema de archivos ext2, los directorios son archivos especiales que se utilizan para crear y mantener rutas de acceso a los archivos en el sistema de archivos. Un archivo de directorio es una lista de entradas de directorio, cada una con la siguiente información:

- **Inodo:** El inodo para esta entrada de directorio. Este es un índice en el conjunto de inodos contenidos en la tabla de inodo del grupo de bloques. En la figura, la entrada del directorio para el archivo llamado file tiene una referencia al número de inodo i1.
- **longitud del nombre:** La longitud de esta entrada de directorio en bytes.
- **nombre :** El nombre de esta entrada de directorio.

Las dos primeras entradas para cada directorio son siempre las entradas estándar “.” y “..” que significan “este directorio” y “el directorio padre” respectivamente.

Sistemas de archivo Virtual - VFS



La capa del VFS del núcleo de Linux se usa durante las llamadas a sistema que actúan sobre archivos, es una capa de indirección que maneja las llamadas a sistema orientadas a archivos y llama a las funciones necesarias en el código del sistema de archivo físico para realizar la entrada y salida. Cuando un proceso efectúa una llamada a sistema orientada a operar un archivo, el núcleo llama a una función contenida en el VFS. Esta función se encarga de las manipulaciones independientes de la estructura y redirige la llamada a una función contenida en el código del sistema de archivos físico, el cual es responsable de encargarse de las operaciones dependientes de la estructura.

Otros sistemas de archivos

ext3fs: Agrega capacidades de journaling (bitácora) de transacciones mediante el archivo `.journal`. Es retrocompatible con ext2. Permite crecimiento on-line. Utiliza una versión especial de B-tree denominada H-tree. El número máximo de subdirectorios es de 32000.

ReiserFS: Fue diseñado por Hans Reiser. Utiliza journaling. Puede contener hasta un máximo de 2^{32} archivos, en un sistema que puede llegar a tener 16 TiB (Tebibits) y el tamaño máximo que puede llegar a tener un archivo es de 2^{32} , es decir, 2 GiB en la versión 3.5 y de 1 EiB (Eibit) en la versión 3.6. También soporta listas de control de acceso (Access Control Lists – ACLs).

ext4fs: La siguiente versión de ext3fs, compatible hacia adelante y hacia atrás. El direccionamiento del bloque es de 48 bits, por lo tanto el sistema de archivos puede llegar a ocupar hasta 1 EB (1 EB=1024 PB, 1 PB=1024 TB, 1 TB=1024 GB). El tamaño máximo que puede llegar a tener un archivo es de 16 TB. El número de subdirectorios es ilimitado. Usa suma de comprobación (checksum) del journal. Utiliza extensiones (extents) en vez de indirectos.

btrfs: Abreviatura de b-tree file system. Puede contener hasta 2^{64} archivos. Un archivo puede llegar a ocupar hasta 16 EiB Exbibyte (1 EiB=260 bytes).

zfs: Se apoya en espacios de almacenamiento virtuales (virtual storage pools), o vdevs, locales o remotos.