

Sistemas Operativos

Sistema de Archivos I

Curso 2024

Facultad de Ingeniería, UDELAR

1. Interfaz

Archivos

Directorios

Seguridad en archivos

2. Implementación

Interfaz del sistema de archivos

- Los dispositivos masivos (discos duros, cintas, etc.) permiten guardar información de forma no volátil.
- El sistema operativo se abstrae de las propiedades físicas de los dispositivos para definir una unidad lógica de almacenamiento: el archivo.
- Los archivos son *mapeados* por el sistema a los distintos dispositivos.
- Un archivo es una agrupación de información que es guardada en algún dispositivo no volátil.
- Desde la perspectiva del usuario, es la unidad mínima de almacenamiento que el sistema le provee.

- Los archivos poseen atributos:
 - **Nombre**: un nombre simbólico que permite identificar el archivo.
 - **Identificador**: símbolo que lo identifica de forma única a nivel global del sistema de archivos. Usualmente es un número.
 - **Tipo**: programa ejecutable, archivo de datos, etc.

- Los archivos poseen atributos:
 - **Ubicación**: puntero al dispositivo y lugar donde reside el archivo.
 - **Tamaño**: tamaño actual del archivo (en bytes, palabras o bloques).
 - **Protección**: información de control para el acceso al archivo. Ej.: usuarios que pueden acceder, en que forma, etc.
 - **Información de conteo**: fecha de creación, último acceso, etc.

Operaciones sobre los archivos

- El sistema operativo brinda servicios para la manipulación de archivos:
 - **Crear y abrir:** creación de un archivo en el sistema de archivos. Se debe especificar un nombre para el nuevo archivo. Además, se provee la apertura de un archivo ya existente para acceder o modificar su información.
 - **Escribir:** escribir información en un archivo previamente abierto.
 - **Leer:** leer información en un archivo previamente abierto.
 - **Reposicionar dentro de un archivo:** acceder a cualquier parte del archivo.
 - **Eliminar:** destruir el archivo a nivel del sistema de archivos.
 - **Truncar:** eliminar parte de la información del archivo.

Operaciones sobre los archivos

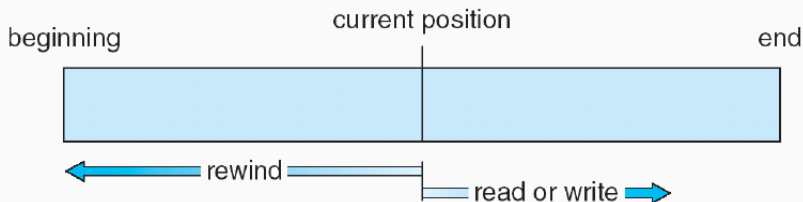
- Por lo general, los sistemas mantienen una tabla de archivos abiertos por proceso.
- Estos archivos se abren a través de un llamado al sistema y, de esa forma, se puede operar con ellos (leer, escribir, etc.).
- Finalmente, el archivo es cerrado antes que finalice la ejecución del proceso.
- Tener un archivo abierto para el sistema implica mantener una estructura que tenga por lo menos:
 - puntero de archivo (*file pointer*) para operaciones de lectura y escritura
 - ubicación del archivo en el dispositivo
 - permisos de acceso

Operaciones sobre los archivos

- Algunos sistemas proveen acceso exclusivo a un archivo (**lock**) por parte de los procesos.
- A su vez, algunos sistemas implementan el mapeo de archivos al espacio de memoria virtual de un proceso. De esta forma, no es necesario realizar **read** y **write** para operar sobre el archivo, sino que se puede acceder directamente a memoria principal. Esto trae el beneficio de no hacer el llamado a sistema para operar sobre el archivo.

Métodos de acceso: Secuencial

- La información es accedida en orden, registro a registro.
- El concepto de registro depende del tipo de archivo (en texto sería un carácter).
- Las operaciones de lectura acceden a la información en forma secuencial e incrementando el puntero de archivo.
- Basado en el modelo de cinta, y funciona en dispositivos secuenciales y de acceso directo.

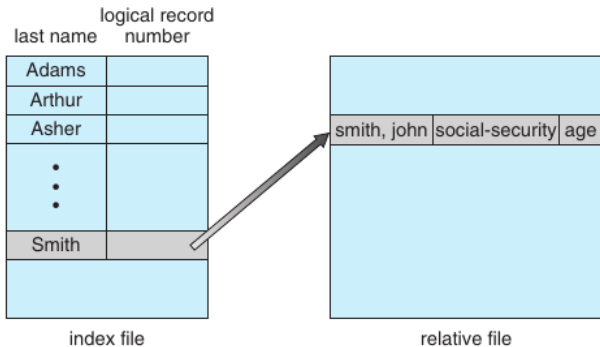


Métodos de acceso: Directo

- El archivo es presentado como una secuencia numerada de registros.
- La información es accedida libremente. No existe restricciones sobre el orden de escritura y lectura del archivo.
- Es basado en el modelo de disco duro, en donde los archivos están estructurados en bloques.

Directorios

- El sistema de archivos se encuentra, por lo general, estructurado en directorios que contienen archivos.
- Los directorios permiten a los usuarios del sistema tener una organización lógica del sistema de archivo.
- Un directorio puede verse como una tabla de símbolos que mapea nombres a bloques que forman un archivo.

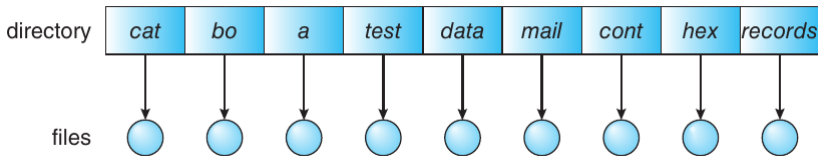


Operaciones sobre los directorios

- **Búsqueda:** es necesario poder buscar un archivo en un directorio
- **Crear un archivo:** archivos nuevos deben ser creados e incorporados al directorio.
- **Eliminar:** borrar un archivo del directorio.
- **Listar:** visualizar los archivos que están en un directorio.
- **Renombrar un archivo:** cambiar el nombre de un archivo dentro del directorio.
- **Permitir la navegación:** acceder a todos los directorios del sistema de archivos.

Estructura de directorios – Nivel único

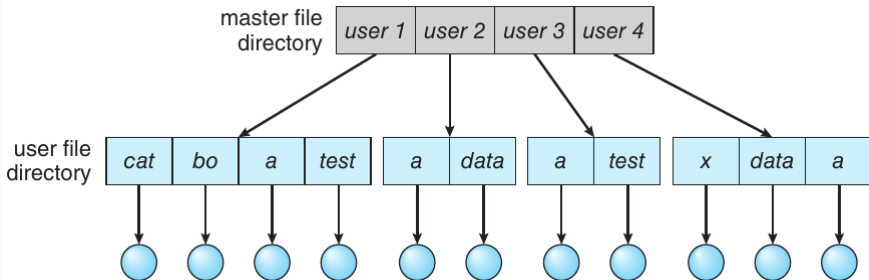
- El esquema más sencillo es tener un único nivel de directorios en el sistema de archivos.
- Todos los archivos están contenidos en el mismo directorio.



- A medida que el sistema crece trae limitaciones. Por ejemplo, los archivos deben tener nombre único.

Estructura de directorios – Dos niveles

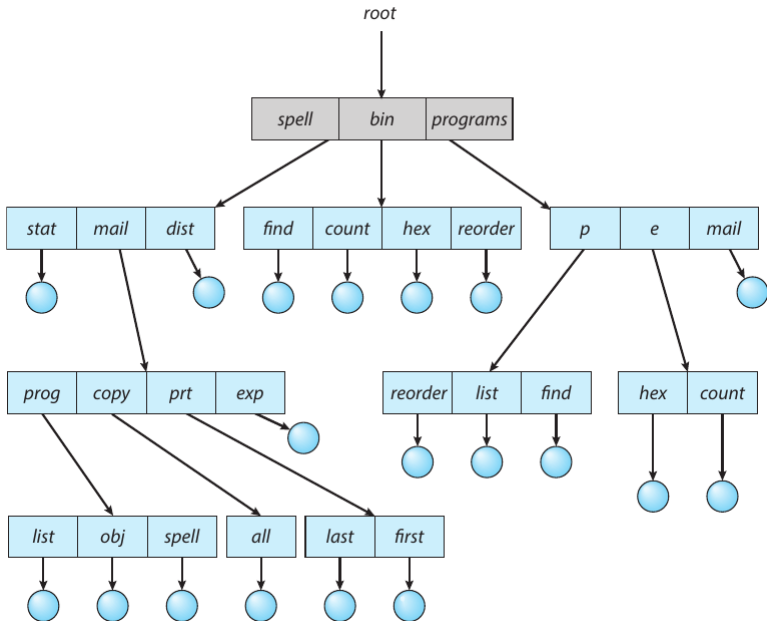
- Tener un único directorio genera confusión entre diferentes usuarios por los nombres de archivos.
- La solución estándar es tener un directorio independiente para cada usuario.



Estructura de directorios – Árbol

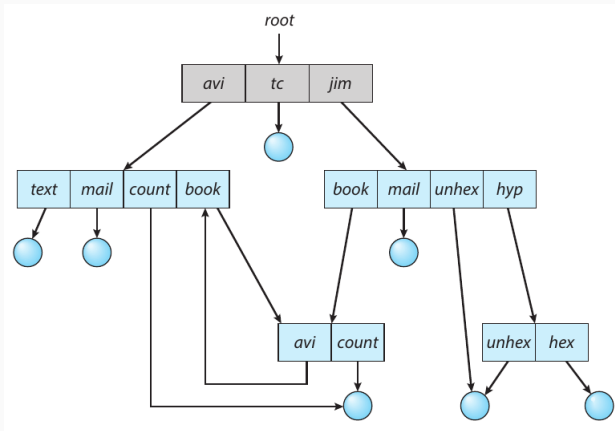
- Un mejor enfoque es permitir varios niveles de directorios.
- Se logra teniendo directorios (o subdirectorio) dentro de un directorio.
- Se genera una estructura jerárquica de directorios en forma de árbol.
- Se denomina ruta (*path*) absoluta de un archivo, al camino desde la raíz hasta el archivo.

Estructura de directorios – Árbol



- La estructura anterior se potencia con un esquema de grafo mediante enlaces (*links*).
- Dos estrategias:
 - Enlace simbólico a un archivo o directorio mediante archivos de enlace (**soft link**).
 - Duplicar la información de un archivo en diferentes directorios (**hard link**).

Estructura de directorios – Grafo



- En un sistema multiusuario es necesario proteger la información de cada usuario.
- En muchos casos los usuarios se agrupan según el uso que tienen sobre un sistema.
- Se definen permisos sobre los archivos tanto a nivel de usuario como de grupo.
- Los permisos más comunes son de:
 - Escritura
 - Lectura
 - Ejecución
 - Eliminar
 - Listar

Seguridad: ejemplo UNIX

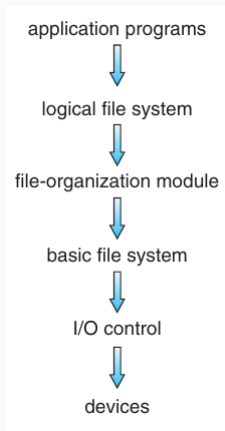
- Seguridad de archivos y directorios se maneja idéntico.
- Cada archivo y directorio tiene asociado un usuario dueño y un grupo.
- También tiene asociado tres campos (permisos dueño, permisos grupo, permisos demás) y tres bits (rwx).

```
-rw-rw-r-- 1 pbg staff 31200 Sep 3 08:30 intro.ps
drwx----- 5 pbg staff 512 Jul 8 09:33 private/
drwxrwxr-x 2 pbg staff 512 Jul 8 09:35 doc/
drwxrwx--- 2 jwg student 512 Aug 3 14:13 student-proj/
-rw-r--r-- 1 pbg staff 9423 Feb 24 2017 program.c
-rwxr-xr-x 1 pbg staff 20471 Feb 24 2017 program
drwx--x--x 4 tag faculty 512 Jul 31 10:31 lib/
drwx----- 3 pbg staff 1024 Aug 29 06:52 mail/
drwxrwxrwx 3 pbg staff 512 Jul 8 09:35 test/
```

Implementación del sistema de archivos

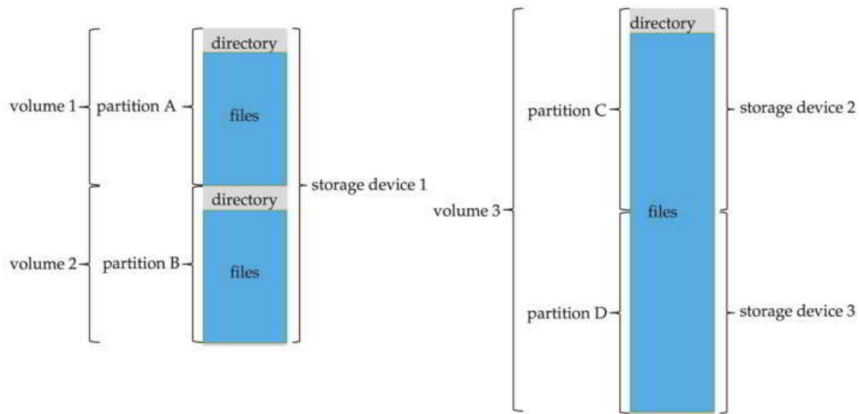
Diseño

- El diseño de un sistema de archivo enfrenta dos problemas:
 - Visión del usuario: implica definir los atributos y las operaciones válidas de los archivos, y la estructura de directorios para la organización de los archivos.
 - La creación de algoritmos y estructura de datos para hacer corresponder el sistemas de archivos lógico con los dispositivos físicos de datos.
- El sistema de archivos está compuesto de varias capas. Cada una de las cuales utiliza la funcionalidad de la capa inferior.



- Un sistema puede tener múltiples dispositivos físicos de almacenamiento.
- Cada dispositivo puede estar dividido en particiones.
- Las particiones contienen volúmenes que pueden extenderse por varias particiones.
- Los volúmenes contienen los directorios y archivos.

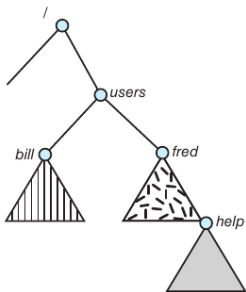
Organización



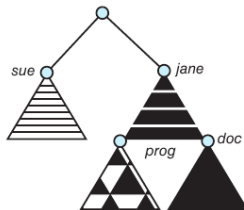
Montado del sistemas de archivos

- Un sistema de archivos debe ser montado para poder ser usado.
- La estructura de directorio del sistema puede estar compuesta por múltiples volúmenes.
- La operación de montaje es simple: recibe un volumen y un directorio (típicamente vacío) como argumentos.
- El volumen raíz es seleccionado montado al momento de inicio del sistema y típicamente contiene el kernel.

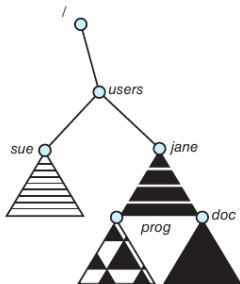
Montado del sistemas de archivos



(a)



(b)



Estructura: en dispositivo

- Los dispositivos físicos contienen las estructuras:
 - **Bloque de control para el arranque** por volumen (*boot control block*): es necesario para iniciar el sistema operativo en el volumen. Este bloque es típicamente el primer bloque del volumen y puede estar vacío.
 - **Bloque de control de volumen** por volumen (*volume control block*): contiene información del volumen como cantidad de bloques, tamaño del bloque, bloques usados y libres, etc.
 - **Estructura de directorios** por sistema de archivos: para la organización de los archivos, incluye información de directorios, nombres de archivos, punteros al archivo, etc.
 - **Bloque de control del archivo** por archivo (*file control block (FCB)*): contiene los detalles del archivo, en particular debe indicar que bloques del disco contienen la información del archivo.

- El sistema operativo mantiene **en memoria** las estructuras:
 - La tabla de montado con información de los volúmenes montados.
 - La caché de estructura de directorio con la información de los directorios accedidos recientemente.
 - Tabla de descriptores de archivos abiertos a nivel global del sistema: contiene una copia del FCB de cada archivo abierto.
 - Tabla de descriptores de archivos abiertos por cada proceso del sistema: contiene punteros a la tabla global de archivos abiertos.

- Cada archivo en el sistema tendrá un bloque de control (*file control block (FCB)*).
- El bloque de control contiene varios atributos de conteo, permisos y donde están los datos del mismo:
 - Fechas (creación, acceso, modificación).
 - Usuario propietario y grupo propietario.
 - Permisos del archivo.
 - Tamaño del archivo.
 - Bloques de datos del archivo.