### Sistema de archivos

Implementación del sistema de archivos

### Estructura del sistema de archivos

Los archivos son unidades de almacenamiento lógicas.

El sistema de archivos reside en el almacenamiento secundario (discos) y proporciona un acceso eficiente y conveniente al dispositivo de almacenamiento al permitir que los datos se almacenen, ubiquen y recuperen fácilmente.

• Define cómo deben verse los archivos para el usuario (atributos de archivo, operaciones y organización).

• Proporciona algoritmos y estructuras de datos para mapear el sistema de archivos lógico en los dispositivos físicos de almacenamiento secundario.

Para mejorar la eficiencia de E / S, las transferencias de E / S entre la memoria y el almacenamiento masivo se realizan en unidades de bloques. Cada bloque en una unidad de disco duro tiene uno o más sectores. Dependiendo de la unidad de disco, el tamaño del sector suele ser de 512 bytes o 4.096 bytes.

### Nivel de aplicación

Gestiona la información de metadatos y la estructura del directorio. Mantiene la estructura de archivos a través de bloques de control de archivos (FCB) o inodos.

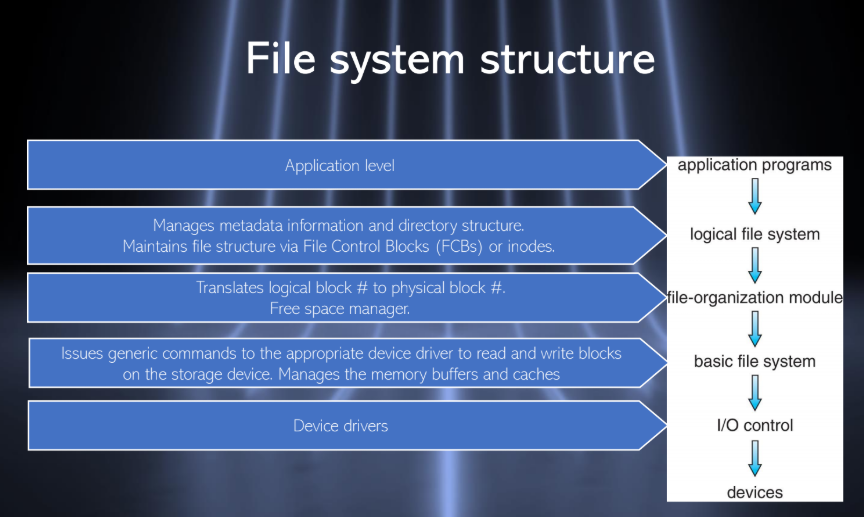
Traduce el bloque lógico # al bloque físico #.

Administrador de espacio libre.

Emite comandos genéricos al controlador de dispositivo apropiado para leer y escribir bloques en el dispositivo de almacenamiento. Gestiona las memorias intermedias y las memorias caché.

Controladores de dispositivo

### Estructura del sistema de archivos



### Archivos en el sistema de archivos

Muchos sistemas de archivos, a veces muchos dentro de un sistema operativo. Cada uno con su propio formato. Algunos ejemplos son FAT, FAT32, NTFS, ext3, ext4, ZFS.

Bloque de control de volumen: contiene información sobre los archivos y sus ubicaciones en el HDD.

• Tabla maestra de archivos para Windows NTFS

• Superbloque para sistemas de archivos Unix (Ext3, Ext4, etc.)

Si la partición se usa para arrancar, entonces el FS debe mantener un Bloque de control de arranque, por lo general, es el primer bloque de volumen.

• El MBR utiliza la información de esta partición para iniciar el sistema operativo.

• Si el gestor de arranque no puede leer el FS, entonces esta partición no es de arranque.

La tabla de montaje almacena montajes del sistema de archivos, puntos de montaje, tipos de sistemas de archivos

### Bloque de control de archivos (FCB)

Un FCB (un inodo en sistemas de archivos UNIX) contiene información sobre el archivo, incluida la propiedad, los permisos y la ubicación del contenido del archivo.

• típicamente número de inodo, permisos, tamaño, fechas

• NFTS se almacena en la tabla del archivo maestro utilizando estructuras de base de datos relacionales

### Implementación de directorio

Para cada directorio, el FS necesita saber qué archivos y subdirectorios están dentro.

Hay dos implementaciones principales:

• Lista lineal: contiene una lista de todos los archivos y carpetas. Si bien es fácil de implementar, requiere tiempo lineal para buscar un archivo.

• Tabla hash: crea un hash para cada archivo en el directorio. Disminuye el tiempo de búsqueda del directorio. Puede provocar colisiones.

### Asignación contigua de archivos

Asignación contigua: cada archivo ocupa un conjunto de bloques contiguos

• Mejor rendimiento en la mayoría de los casos.

• Simple: solo se requiere la ubicación inicial (número de bloque) y la longitud (número de bloques)

• Los problemas incluyen encontrar espacio para el archivo, conocer el tamaño del archivo, fragmentación externa, necesidad de compactación fuera de línea (tiempo de inactividad) o en línea

### Asignación de archivos vinculados

Asignación vinculada: cada archivo es una lista vinculada de bloques

• El archivo termina con un puntero nulo

• Sin fragmentación externa

• Cada bloque contiene puntero al siguiente bloque

• Sin compactación, fragmentación externa.

• La localización de un bloque puede tomar muchas E / S y búsquedas de disco

### Tabla de asignación de archivos (FAT)

Variación de asignación vinculada.

• Una sección de almacenamiento al comienzo de cada volumen se reserva para contener la tabla de punteros. La tabla tiene una entrada para cada bloque y está indexada por número de bloque.

• Al igual que una lista vinculada, pero más rápido en disco y almacenable en caché

• Originalmente utilizado por MS-DOS.

• La implementación común es FAT 32.

### Asignación indexada

Cada archivo tiene su (s) propio (s) bloque (s) de índice de punteros a sus bloques de datos.

• Requiere la sobrecarga de un bloque de índice.

• Más útil para el acceso directo (aleatorio) sin crear fragmentación externa.

### Esquema combinado: UNIX UFS

Los primeros punteros del bloque de índice apuntan a bloques directos; es decir, contienen direcciones de bloques que contienen datos del archivo.

Los siguientes punteros apuntan a bloques indirectos (simple, doble o triple),

### Espacio libre

Para realizar un seguimiento del espacio libre en disco, el sistema mantiene una lista de espacio libre. La lista de espacio libre registra todos los bloques de dispositivos libres, aquellos no asignados a algún archivo o directorio.

No necesariamente implementado como una lista. Con frecuencia, la lista de espacio libre se implementa como un mapa de bits o un vector de bits. Cada bloque está representado por 1 bit. Si el bloque es libre, el bit es 1; Si el bloque está asignado, el bit es 0.

0 1 2 n-1

...

bit [i] =

1 ⇒ bloque [i] libre

0 ⇒ bloque [i] ocupado

### Registro de sistemas de archivos estructurados

Los sistemas de archivos estructurados de registro (o registro en diario) registran cada actualización de metadatos en el sistema de archivos como una transacción; guardar primero las acciones y luego ejecutarlas de forma asincrónica.

Si el sistema de archivos falla, todas las transacciones restantes en el registro aún deben realizarse.

### Interfaz del sistema de archivos

Descripción general de la interfaz del sistema de archivos

• Un archivo es un tipo de datos abstracto definido e implementado por el sistema operativo. Es una secuencia de registros lógicos. Un registro lógico puede ser un byte, una línea (de longitud fija o variable) o un elemento de datos más complejo. El sistema operativo puede admitir específicamente varios tipos de registros o puede dejar ese soporte para el programa de aplicación.

• Un directorio estructurado en árbol permite al usuario crear subdirectorios para organizar archivos. Las estructuras de directorio de gráficos acíclicos permiten a los usuarios compartir subdirectorios y archivos, pero complican la búsqueda y eliminación.

• Dado que los archivos son el principal mecanismo de almacenamiento de información en la mayoría de los sistemas informáticos, se necesita protección de archivos en sistemas multiusuario. El acceso a los archivos se puede controlar por separado para cada tipo de acceso: leer, escribir, ejecutar, agregar, eliminar, listar directorios, etc. La protección de archivos se puede proporcionar mediante listas de acceso, contraseñas u otras técnicas

Concepto de archivo

Un archivo es un tipo de datos abstracto definido e implementado por el sistema operativo. Puede contener:

• Datos (numéricos, caracteres, binarios)

• Programa Muchos tipos: archivo fuente, archivo de texto, archivo ejecutable, etc. Cada archivo tiene un conjunto de atributos:

• Nombre

• Identificador

• Tipo

• Ubicación

• Talla

• Proteccion

• Hora, fecha e identificación del usuario.

Operaciones de archivo

• Crear

• Escribir: en la ubicación del puntero de escritura

• Leer: en la ubicación del puntero de lectura

• Reposicionar dentro del archivo - buscar

• Eliminar

• Truncar: borra el contenido de un archivo pero conserva sus atributos.

• Abrir: busque la estructura de directorios en el disco y mueva el contenido a la memoria

• Cerrar: mueve el contenido de la memoria a la estructura de directorios en el disco

Abrir archivos

Se necesitan varios datos para administrar archivos abiertos:

• Tabla de archivo abierto:

• La tabla de archivos abiertos de todo el sistema contiene una copia del FCB de cada archivo y otra información

• La tabla de archivo abierto por proceso contiene punteros a las entradas apropiadas en la tabla de archivo abierto de todo el sistema, así como otra información

• Recuento de archivos abiertos: contador de la cantidad de veces que se abre un archivo, para permitir la eliminación de datos de la tabla de archivos abiertos cuando el último proceso lo cierra

• Ubicación del disco del archivo.

• Derechos de acceso: información del modo de acceso por proceso El sistema operativo puede proporcionar el bloqueo:

• obligatorio

• Asesoría

Controladores / descriptores de archivos

Para cada proceso que accede a un archivo, su PCB contiene un puntero a los bytes leídos o escritos. En Windows, esta estructura en la PCB se conoce como el administrador de archivos, en Linux como el descriptor de archivos.

El acceso a un archivo puede ser secuencial o directo.

Métodos de acceso

Acceso secuencial:

• leer a continuación

• escribe a continuación

• restablecer acceso directo: el archivo tiene registros lógicos de longitud fija

• leer n

• escribir n

• reescribir n

Directorio

La información sobre los archivos se mantiene en la estructura de directorios, que se mantiene en el disco. Operaciones:

• Buscar un archivo

• Crear un archivo

• Eliminar un archivo

• Listar un directorio

• Cambiar el nombre de un archivo

• Recorrer el sistema de archivos

Estructura de directorios

Dentro de un sistema de archivos, es útil crear directorios para permitir la organización de los archivos. Un directorio de un solo nivel en un sistema multiusuario causa problemas de nombres, ya que cada archivo debe tener un nombre único. Un directorio de dos niveles resuelve este problema creando un directorio separado para los archivos de cada usuario. El directorio enumera los archivos por nombre e incluye la ubicación del archivo en el disco, la longitud, el tipo, el propietario, la hora de creación, la hora del último uso, etc.

La generalización natural de un directorio de dos niveles es un directorio estructurado en árbol. Un directorio estructurado en árbol permite al usuario crear subdirectorios para organizar archivos. Las estructuras de directorio de gráficos acíclicos permiten a los usuarios compartir subdirectorios y archivos, pero complican la búsqueda y eliminación. Una estructura gráfica general permite una flexibilidad completa en el intercambio de archivos y directorios, pero a veces requiere recolección de basura para recuperar espacio en disco no utilizado.

Directorios estructurados en árbol

Nombre de ruta absoluto o relativo. Concepto actual de directorio.

Directorios de gráficos acíclicos

Permite subdirectorios y archivos compartidos

Enlaces

Nuevo tipo de entrada de directorio.

• Enlace: otro nombre (puntero) a un archivo existente

• Resolver el enlace: siga el puntero para localizar el archivo

Linux ln comando para crear enlaces blandos y duros.

Directorios de gráficos generales

Requiere recolección de basura para recuperar espacio en disco no utilizado

Compartir archivos y protección

El intercambio se puede realizar a través de un esquema de protección Si el sistema multiusuario:

• Las ID de usuario identifican a los usuarios, lo que permite que los permisos y protecciones sean por usuario. Las ID de grupo permiten a los usuarios estar en grupos, lo que permite los derechos de acceso de grupo

• Propietario de un archivo / directorio

• Grupo de un archivo / directorio

El propietario / creador del archivo debe poder controlar lo que se puede hacer y quién lo hace. Tipos de acceso:

• Leer

• Escribir

• Ejecutar

Protección UNIX

Unix separa a los usuarios en tres:

• Propietario: el creador del archivo (se puede cambiar con el comando Cambiar propietario)

• Grupo: conjunto de usuarios con los mismos permisos. Solo un grupo por archivo.

• Público: todos los demás

Los permisos también son tres, representados por nueve bits.

• R-Read, indicado con el primer bit (4).

• W-Write, indicado con el segundo bit (2).

• Ejecutar X, indicado con el tercer bit (1)

Listas de control de acceso

Se mantiene una lista para cada usuario del sistema para indicar sus permisos. El valor predeterminado es negar todo. Utilizado por el sistema operativo Windows

Estructura del disco

• El disco se puede subdividir en particiones, también conocidas como minidiscos, cortes.

• El disco o la partición pueden usarse sin formato, sin un sistema de archivos, o formateados con un sistema de archivos

• Entidad que contiene el sistema de archivos conocido como volumen. Cada volumen que contiene el sistema de archivos también rastrea la información de ese sistema de archivos en el directorio del dispositivo o en la tabla de contenido del volumen

• Además de los sistemas de archivos de propósito general, hay muchos sistemas de archivos de propósito especial, frecuentemente todos dentro del mismo sistema operativo o computadora

Montaje del sistema de archivos

• Es necesario montar un sistema de archivos en el sistema operativo para que el usuario pueda acceder a él.

• El montaje vincula una partición a un directorio. Normalmente, un punto de montaje es un directorio vacío.

• Las particiones se pueden montar en el momento del arranque o en el tiempo de ejecución.