

Administración de archivos en UNIX

Los **archivos** son estructuras que permiten mantener información en el disco de tal manera que esta pueda ser manejada en forma individual por los usuarios que tienen derechos de acceso a dicha información.

Como dicha estructura usa recursos del sistema, es función del sistema operativo disponer los mecanismos necesarios para su administración.

La información del archivo se encuentra en los **bloques** del disco. Cada bloque está formado por uno o más sectores consecutivos, y constituye la unidad de asignación, ya que un archivo solo podrá tener un número entero de bloques.

Como ya vimos, el sistema operativo es el encargado de administrar dichos recursos, y por lo tanto su función será asignar bloques a los archivos, desasignar bloques para que dejen de pertenecer a un archivo, y localizarlos cuando un proceso acceda a ellos.

Por otra parte cada archivo, además de su información, tendrá atributos, mediante los cuales el sistema operativo lleva a cabo su administración.

En UNIX los archivos se localizan mediante su nombre, que figura en una entrada del directorio que los contiene, y en la misma entrada se encuentra un número de dos bytes que permite localizar una estructura de datos que se almacena en el disco, que contiene el resto de los atributos, y los punteros que permiten localizar a los distintos bloques del disco. Esta estructura se llama **inodo**, y debe copiarse en la memoria cuando se abra el archivo.

Cuando se abre un archivo (para su posterior acceso) el sistema operativo lleva un control de las operaciones posibles de dicho archivo mediante atributos adicionales que se suman a los del inodo.

Por dicha razón al inodo lo llamamos **inodo en disco**, y a todos los atributos que residen en la memoria se los llama **inodo en memoria**.

Estructura del inodo en disco

Contiene los siguientes campos:

- **UID** del usuario propietario. El UID es un número que le asigna el sistema operativo a cada usuario cuando este se registra en el sistema. Cada usuario para poder trabajar en la terminal (o computadora) debe registrarse en el sistema. Cada archivo tiene un propietario, que es el usuario que creó el archivo.
- El **GID** del grupo propietario. Este GID es un número que el sistema operativo le asigna a un grupo de usuarios, al cual pertenece el usuario (por defecto es el **grupo primario** del usuario propietario, que luego de creado el archivo podrá cambiarse). Todo usuario pertenece a un grupo primario, normalmente definido por el administrador del sistema (dicho grupo por defecto es el grupo primario de login. Sin embargo el usuario podrá cambiarlo por otro grupo al cual también pertenezca). El **administrador del sistema** es un usuario que tiene una jerarquía especial que le permite realizar operaciones que normalmente no están permitidas a usuarios normales. Dicho administrador se llama genéricamente **Superusuario** en el sistema operativo UNIX, aunque su nombre específico es **root**.
- **Tipo** de archivo. Ya mencionamos los distintos tipos de archivos. Cuando el inodo no tiene asignado ningún archivo, el tipo se define como *tipo 0*. De esa forma se reconoce cuando un inodo está libre, o sea no asociado a ningún archivo.
- Protección **rwX** para el propietario del archivo; protección **rwX** para los integrantes del grupo propietario; protección **rwX** para el resto de los usuarios. Los archivos se protegen contra lectura (**r**); contra escritura (**w**) y contra ejecución (**x**).
- Fecha y hora del último acceso al archivo.

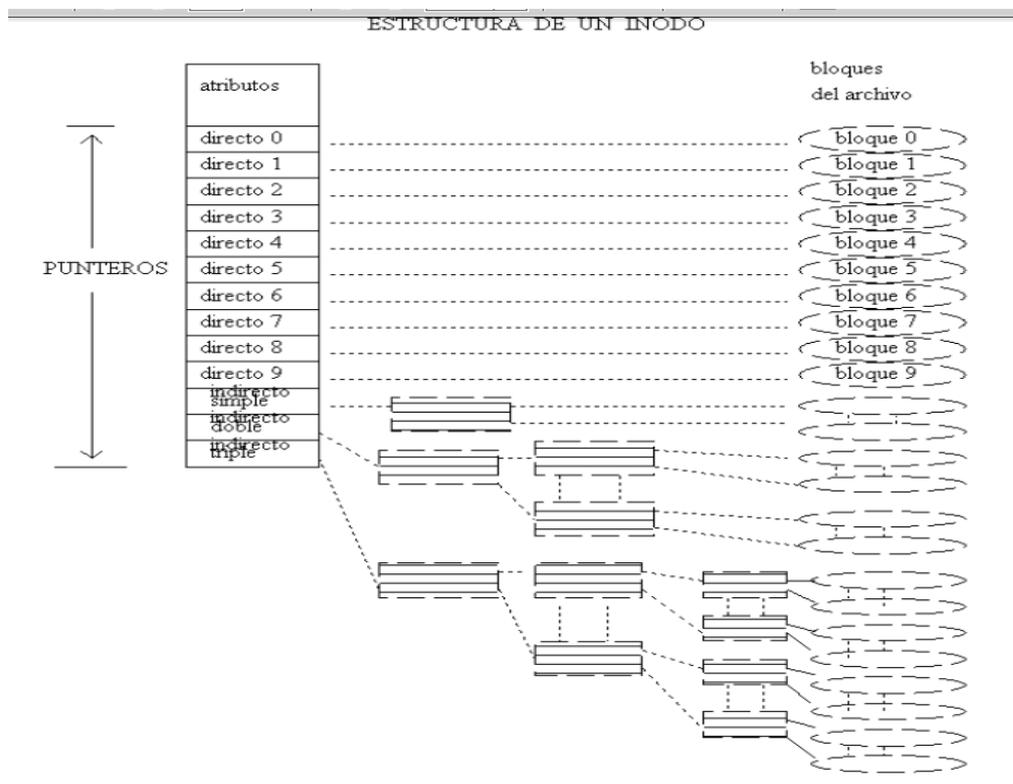
- Fecha y hora de la última modificación (o creación) del archivo.
- Fecha y hora de la última modificación del inodo.
- Numero de enlaces duros (**links**). Pueden haber varios nombres (enlaces) asociados con un mismo inodo. En otras palabras varios nombres de archivos comparten el mismo archivo físico.
- Punteros a los bloques del archivo. Son números de 32 bits que permiten localizar a un bloque físico del disco que contiene información del archivo.
- Tamaño del archivo en bytes.

Vemos entonces que el inodo **no** tiene registrado su número ni el nombre del archivo Tampoco especifica la posición del archivo en el árbol de directorio.

Vimos que el número del archivo esta indicado en el directorio que contiene al archivo, y se usa como índice en la **lista de inodos** que tiene una posición bien definida en el disco.

Vamos a aclarar un poco más sobre los punteros a los bloques del archivo.

La figura muestra al inodo como una estructura que contiene básicamente dos tipos de campos: **atributos** y **punteros**.



El orden de los campos donde están los punteros define la secuencia lógica de los bloques, o sea el número lógico de los bloques (bloque 0, bloque 1, bloque 2 etc.). El contenido de estos campos es el número de 32 bits que indica el número físico del bloque, comúnmente llamado **número de bloque**.

Los primeros diez números de bloque apuntan a bloques del archivo.

Suponiendo que cada bloque tenga 1024 bytes, estos diez bloques representan los 10 Kbytes iniciales del archivo (suponiendo que tenga más de 10 KB).

Cuando el archivo tiene mas de 10 KB se usa el casillero número 11 que contiene el puntero a un bloque del disco que no contiene datos del archivo pero que contiene 256 números de bloque que contienen datos del archivo. Por lo tanto este bloque del disco se

usa como expansión del inodo. El número que se guarda en el casillero 11 se denomina **puntero de indirección simple**.

El casillero 12 contiene un puntero de **indirección doble** que apunta a un bloque que contiene 256 números de bloque de indirección simple, los que a su vez apuntan a los respectivos bloques que contienen (cada uno de ellos) 256 números de bloques del archivo.

El casillero 13 contiene un puntero de **indirección triple** que apunta a un bloque que contiene 256 punteros de indirección doble que apuntan a bloques que contienen 256 punteros de indirección simple, y estos apuntan a bloques que contienen 256 punteros de los bloques del archivo.

Estructura del inodo en memoria

Cuando un archivo es accedido por algún proceso, es necesario que tenga todos sus atributos en la memoria. Estos se encuentran en el **inodo en memoria**.

Todas las operaciones que se deban realizar con un archivo, harán necesaria la ubicación del inodo en memoria.

Para que esta búsqueda sea rápida es necesario disponer una estructura en donde se administren los inodos y se facilite su localización.

Para ello los inodos en memoria se organizan en un **cache de inodos** similar al buffer cache.

Además de toda la información del inodo en disco (se copia el inodo del disco a la memoria), el inodo en memoria contiene lo siguiente:

- **Status**. Es un campo de la estructura que contiene detalles de la situación o estado en que se encuentra el inodo de acuerdo a las modalidades de la administración. Contiene la siguiente información:

- Indicación si el inodo está o no bloqueado. Si está bloqueado significa que un proceso que trate de acceder a dicho inodo deberá bloquearse hasta que el inodo se desbloquee.
- Indicación si existen procesos bloqueados a la espera que se desbloquee el inodo.
- Indicación si el inodo en memoria fue modificado y tiene su actualización en el disco, pendiente.
- Indicación si el archivo fue modificado en memoria pero todavía no se actualizó en el disco.
- Indicación si el archivo es un **punto de montaje**. Esto será aclarado con posterioridad.

- **Número del inodo**.

- **Número del dispositivo lógico** al que pertenece el archivo.

- **Punteros a los inodos de la lista**.

- **Punteros de la lista de inodos libres en memoria**. Aquellos inodos cuyo contador de referencia sea 0, estarán en esta lista. Si el contador de referencia es 0 significa que no hay ningún archivo que se esté accediendo y que use dicho inodo.

- **Contador de referencia**. Es un contador que lleva la cuenta de los accesos al archivo, autorizados por el sistema operativo.

Varios procesos pueden acceder a un mismo archivo. Y un proceso puede tener múltiples accesos autorizados en un mismo archivo.