

Sistema de Archivos

Hugo J. Guevara, Jesús E. Hernández,

Marlem Martínez, Miryam Salazar

Facultad de Ciencias de la Computación

Benemérita Universidad Autónoma de Puebla

Resumen

Los sistemas de archivos en sistemas operativos son conjuntos de estructuras y reglas que se utilizan para organizar y gestionar los datos almacenados en dispositivos de almacenamiento, como discos duros, unidades flash, discos ópticos y sistemas de almacenamiento en red. Estos sistemas permiten a los usuarios y programas almacenar, recuperar, organizar y administrar archivos de manera eficiente. En resumen, los sistemas de archivos en sistemas operativos son fundamentales para la gestión de datos en dispositivos de almacenamiento. Proporcionan una estructura organizativa, nombres únicos, metadatos, permisos de acceso y operaciones de archivo, y existen varios tipos de sistemas de archivos adaptados a diferentes plataformas y necesidades.

1. Introducción

En el núcleo de cualquier sistema operativo, existe un componente esencial que suele operar bajo el radar de la atención pública, pero que es absolutamente vital para el funcionamiento de nuestros dispositivos y sistemas operativos: los sistemas de archivos. Estas infraestructuras digitales son como el sistema de clasificación en una biblioteca gigante, encargadas de organizar, almacenar

y administrar el torrente constante de datos que fluye a través de nuestros dispositivos, desde los archivos personales y multimedia hasta las aplicaciones que utilizamos y la configuración fundamental de nuestros sistemas.

A pesar de su papel central, los sistemas de archivos a menudo operan en silencio, detrás de la pantalla, permitiéndonos acceder y gestionar nuestros archivos sin que les prestemos atención. Sin embargo, este elemento subyacente esconde un mundo complejo de tecnología y diseño que es esencial para la informática moderna.

Descubriremos los secretos de la organización de nuestros archivos, la eficiencia en el acceso a la información y la garantía de la integridad de nuestros datos. Nos ayudará a comprender por qué los sistemas de archivos son una columna vertebral sólida en la informática, contribuyendo de manera crucial a la funcionalidad, eficiencia y seguridad de nuestros sistemas operativos.

En este trabajo se presenta un simulador, donde de una manera grafica nos mostrara el funcionamiento de los sistemas de archivos, el cual ayudara a la comprensión del alumno.

2. Simulador de sistema de archivos

El simulador de sistemas de archivos es una innovadora aplicación desarrollada en el lenguaje de programación Java, que presenta una interfaz gráfica amigable y efectiva para la entrada y visualización de datos relacionados con sistemas de archivos. Al iniciarlo, nos recibe con una interfaz principal de la primera versión del simulador, la cual está diseñada con la intención de proporcionar una experiencia de aprendizaje intuitiva y accesible.

2.1 Introducción al simulador

El simulador de sistema de archivos tiene por objeto mostrar el funcionamiento de algunos aspectos de los archivos.

El sistema operativo almacena datos en bloques de memoria usando una variedad de métodos de asignación de archivos. Eso permite que el disco duro se use de una manera eficiente y se acceda al archivo.

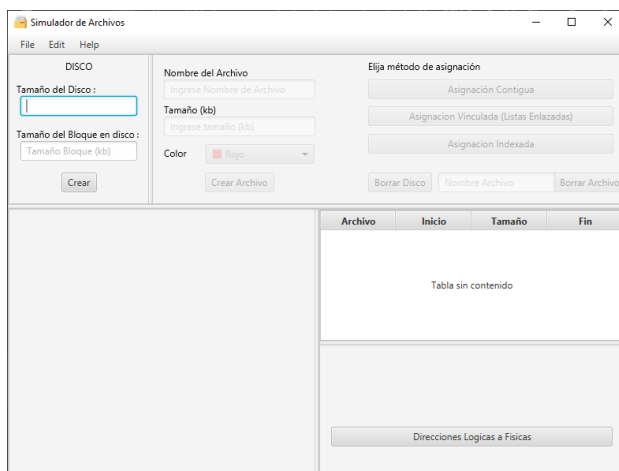


Figura 1. Esquema del simulador de Sistema de Archivos

3. Métodos de asignación de archivos en el sistema operativo

Cuando formateamos un disco duro, se divide el espacio en bloques, que son pequeñas unidades para almacenar archivos. Cada archivo se compone de uno o varios de estos bloques. Los bloques son esenciales para una gestión eficiente de datos, permitiendo un acceso más rápido y la optimización del espacio disponible en el disco. Además, facilitan la administración de archivos y la realización de copias de seguridad.

3.1 Método de asignación contigua

En este método, a los archivos se les asignan los bloques de disco de manera contigua, lo que significa que, si la dirección del bloque de disco inicial de un archivo es x , entonces se le asignarán los bloques con la dirección $x+1$, $x+2$, $x+3$, ..., siempre que los bloques no estén ya ocupados. El sistema operativo también mantiene una tabla de directorios que incluye el nombre del archivo junto con la dirección inicial y la longitud de los bloques asignados. La longitud representa el número de bloques de disco requeridos por el archivo.

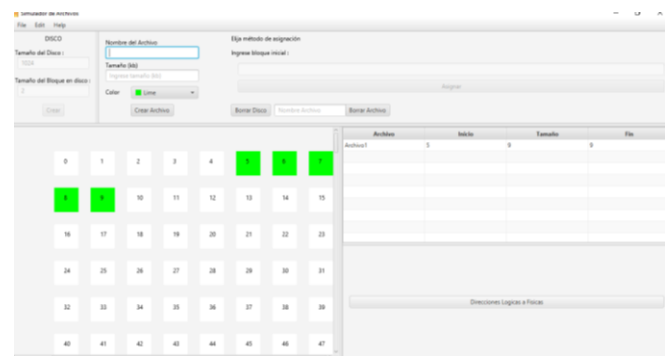


Figura 2. Asignación contigua

La asignación contigua tiene ventajas, como:

- Puesto que los bloques se asignan de forma consecutiva, es posible acceder a ellos en orden secuencial, dado que la ubicación de inicio y la extensión ya se encuentran registradas en la tabla de directorio.
- La asignación de bloques se asemeja al funcionamiento de una matriz. Con la dirección de inicio proporcionada, es posible desplazarse a cualquier dirección de bloque mediante la adición del tamaño del bloque a la dirección inicial, de manera similar a cómo accedemos a elementos individuales en una matriz. En consecuencia, la asignación contigua también habilita la posibilidad de acceder de forma no secuencial a los bloques.
- Es muy rápido, ya que el tiempo de búsqueda es menor debido a la asignación contigua de bloques.

Así mismo, tiene algunas desventajas, como:

- Sufre de fragmentación interna.
- Sufre de fragmentación externa.
- El tamaño del archivo se puede aumentar solo si hay bloques de disco contiguos libres disponibles.
- La gestión de la memoria es deficiente.

3.2 Método de asignación vinculada o ligada

En este método, la asignación vinculada funciona igual que las listas enlazadas. Aquí, los bloques se asignan de tal manera que cada bloque contiene un puntero al siguiente bloque que se asigna al archivo.

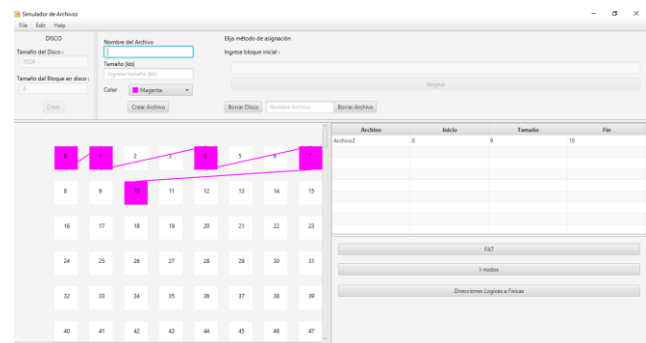


Figura 3. Asignación vinculada

En este contexto, se trata de ver cada archivo como si fuera una secuencia de bloques de disco vinculados entre sí, como una cadena. No obstante, los bloques de disco asignados a un archivo específico no tienen que estar uno al lado del otro físicamente en el disco. Cada bloque de disco asignado a un archivo contiene una referencia que indica cuál es el siguiente bloque de disco asignado al mismo archivo.

La asignación vinculada cuenta con algunas ventajas, como:

- No existe fragmentación externa debido a que los bloques pueden ser asignados de manera no secuencial gracias a la utilización de referencias o punteros.
- El tamaño del archivo se puede aumentar.

- Uso juicioso de la memoria.

Algunas desventajas son:

- No se posibilita el acceso aleatorio debido a que la memoria asignada no es contigua. Así que, con la dirección del bloque de inicio, no podemos desplazarnos directamente a otro bloque, de manera análoga a cómo no podemos saltar a cualquier nodo en una lista enlazada solo con el puntero principal.
- Es lento, por el mayor tiempo de búsqueda
- Cada bloque debe contener información adicional sobre el puntero al siguiente bloque.

3.3 Método de asignación indexada

Este método, se considera como una mezcla de asignación vinculada y contigua que es más eficiente en términos de espacio y tiempo.

Cada archivo cuenta con su propio conjunto de referencias, encapsulado en un bloque de índice que se encuentra en el disco. En lugar de alojar directamente el archivo, este bloque contiene enlaces a otros bloques que, a su vez, albergan el contenido del archivo. Este enfoque facilita el acceso

aleatorio a los archivos y previene la fragmentación externa.

La técnica de asignación indexada presenta ciertas ventajas, tales como:

- Evita la fragmentación externa.
- Facilita el acceso aleatorio a bloques de disco.
- Permite un acceso directo, reduciendo la complejidad.

Sin embargo, también presenta algunas desventajas, como:

- Su implementación es bastante compleja.
- Requiere memoria adicional para almacenar los bloques de índice.
- Genera una sobrecarga debido al uso de punteros grandes.

4. Tabla de asignación de archivos (FAT)

Una tabla de asignación de archivos (FAT) es un sistema de archivos desarrollado para discos duros que originalmente usaba 12 o 16 bits para cada entrada de clúster en la tabla de asignación de archivos. El sistema operativo (SO) lo utiliza para administrar archivos en discos duros y otros sistemas informáticos.

Microsoft desarrolló el sistema de archivos FAT (File Allocation Table) con el objetivo de adaptarse a un almacenamiento eficiente y a estructuras de directorios sencillas. Este sistema utiliza una tabla para rastrear los clústeres en un volumen de almacenamiento, y los archivos y directorios

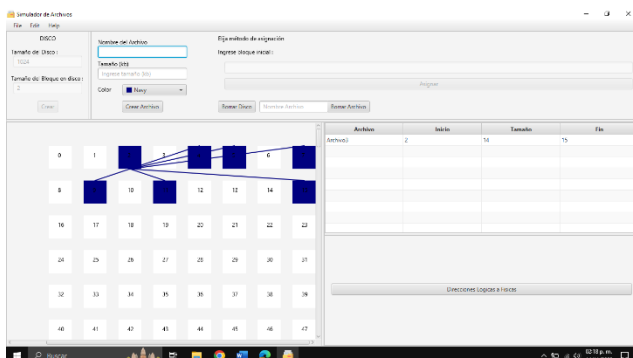
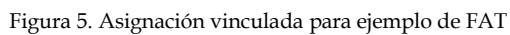


Figura 4. Asignación indexada

Teniendo en cuenta los siguientes datos, podemos observar nuestro simulador de archivos de la siguiente manera:

[illegible]

La FAT es una tabla que relaciona los nombres y números de las carpetas con las carpetas reales y sus datos. Cada entrada de la tabla se denomina “ubicación”. Cada directorio tiene una “ubicación-0” (L0) en la tabla. Cada carpeta tiene una “ubicación-1” en la tabla. Si copias o mueves un archivo o

La nueva entrada se denomina “mapeo” y el mapeo es necesario porque el sistema de archivos está organizado en una estructura de árbol. Cuando el sistema operativo lee datos de una ubicación, lee la tabla para obtener la ubicación de los datos. Si la ubicación de los datos se encuentra en el sistema de archivos, el sistema operativo accede a los datos directamente y no necesita volver a leer la tabla. De este modo, los datos del sistema de archivos se leen directamente y no se producen retrasos.

- Utilización completa de los bloques de disco para almacenar datos.
- La pérdida de bloques sucesivos no ocurre en su totalidad en caso de que uno esté defectuoso.
- Aunque hay acceso aleatorio, no es especialmente rápido.

- Se necesita una entrada en la tabla de asignación de archivos para cada bloque de disco.
- El tamaño del archivo FAT depende de la cantidad de entradas FAT.
- Aumentar el tamaño del bloque puede mejorar la fragmentación interna, pero disminuirá el número de entradas en la FAT.

5. I-nodo

En sistemas de archivos al estilo Unix, como en Linux, cada elemento, ya sea un archivo o un directorio, se representa mediante una estructura de datos conocida como i-nodo (nodo de índice). La información esencial de cualquier archivo en Linux, a excepción de su nombre y contenido, se guarda en un i-nodo. Cuando un usuario o programa solicita acceso a un archivo, el sistema operativo busca el i-nodo específico y único (identificado por un número de i-nodo) en una tabla dedicada de i-nodos. En esencia, este número de i-nodo, obtenido de la tabla, actúa como una guía que ayuda a la aplicación o al usuario a localizar y acceder al archivo deseado.

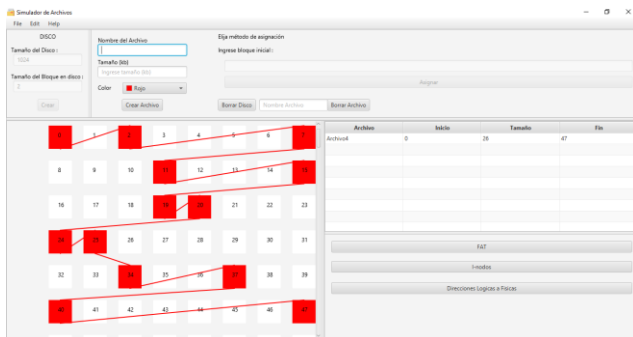


Figura 7. Asignación vinculada para ejemplo i-nodo

Nuestra tabla de i-nodos, se muestra de la siguiente manera:

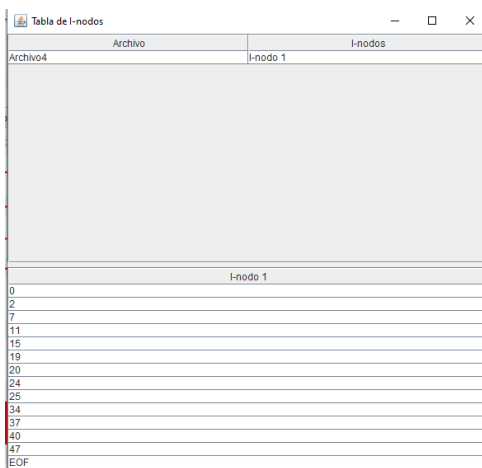


Figura 8. Tabla i-nodo para Archivo4

Si agregamos mas archivos, también la tabla de i-nodos aumenta.

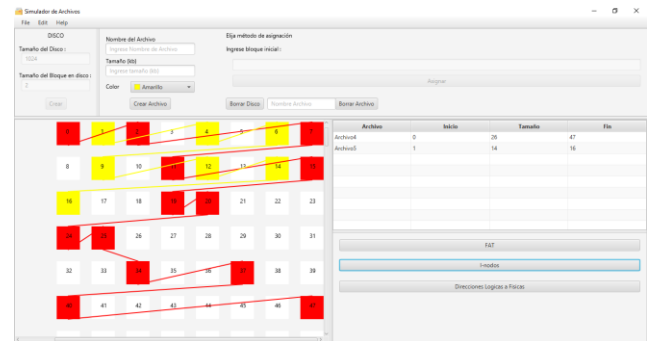


Figura 9. Asignación vinculada para ejemplo i-nodo con dos archivos

Teniendo dos archivos, podemos observar la tabla i-nodo de la siguiente manera:

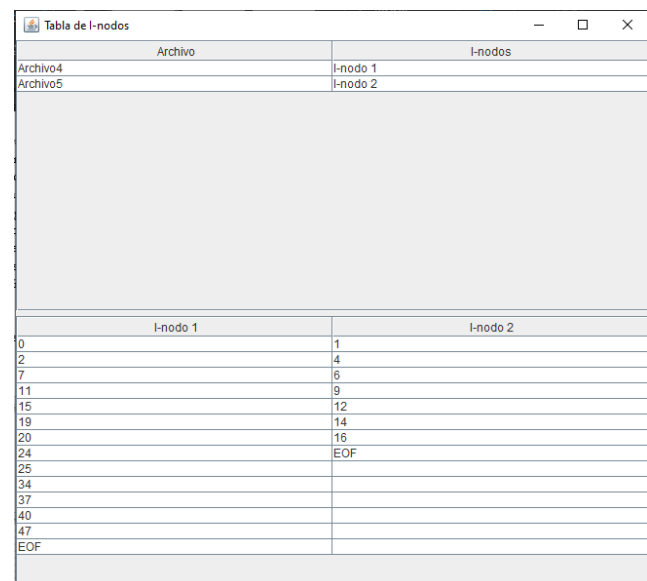


Figura 10. Tabla de i-nodos para Archivo4 y Archivo5

Algunas de las ventajas asociadas con los i-nodos son:

- Los nombres de archivo no son la representación real en los i-nodos.
- El acceso a la información de un i-nodo requiere permisos para acceder al archivo.

- Los i-nodos contienen información sobre los IDs de usuario y grupo asociados con el archivo.

Sin embargo, algunas desventajas incluyen:

- Nuevos archivos y carpetas serán rechazados cuando el sistema de archivos se quede sin i-nodos.
- La pérdida de datos ocurre una vez que los i-nodos están completamente ocupados.
- Existe el riesgo de que las aplicaciones se bloqueen cuando los i-nodos alcanzan su capacidad máxima.

6. Gestión de bits

La gestión del espacio libre es un aspecto crítico en la administración eficiente de sistemas de almacenamiento. Se refiere a la supervisión y optimización del espacio disponible en dispositivos de almacenamiento, como discos duros o unidades de estado sólido. La tarea principal consiste en monitorear la cantidad de espacio no utilizado y aplicar estrategias para asegurar un uso eficiente. Esto implica la identificación y eliminación de archivos innecesarios, así como la prevención de la fragmentación del disco. Una gestión efectiva del espacio libre contribuye no solo a maximizar la capacidad de almacenamiento, sino también a mejorar el rendimiento del sistema al garantizar un acceso rápido y fluido a los datos.

Un sistema de archivos es responsable de asignar los bloques libres al archivo, por lo

tanto, debe realizar un seguimiento de todos los bloques libres presentes en el disco.

Existen principalmente dos enfoques mediante los cuales se administran los bloques libres en el disco.

La elección entre vector de bits y lista ligada depende de los requisitos específicos del sistema de archivos.

6.1 Vector de bits

Un vector de bits es una estructura de datos que se utiliza para llevar un registro del estado de ocupación de bloques de almacenamiento en un dispositivo, como un disco duro. Cada bit en el vector representa el estado de un bloque particular, indicando si está ocupado (1) o libre (0).

Este enfoque proporciona una representación compacta y eficiente del espacio disponible en el dispositivo. Al examinar el vector de bits, el sistema puede identificar rápidamente los bloques libres contiguos, facilitando la asignación eficiente de espacio para nuevos archivos o datos. Además, el uso de un vector de bits ayuda a minimizar la sobrecarga de almacenar información sobre cada bloque de manera individual, ya que se puede representar de manera concisa con bits.

6.2 Lista ligada

Una lista ligada es una estructura de datos que registra bloques de almacenamiento disponibles en un dispositivo. A diferencia del vector de bits, utiliza nodos interconectados para

representar dinámicamente bloques libres, facilitando la asignación eficiente de espacio para nuevos archivos. La lista ligada optimiza el uso del espacio, aunque puede tener cierta sobrecarga debido a la información adicional en cada nodo.

7. Conclusiones

En este trabajo, se ha presentado una herramienta dirigida a la simulación de Sistema de Archivos. Dicha herramienta ha sido realizada en lenguaje Java, lo que facilita su utilización en entornos web y su portabilidad a diferentes sistemas operativos. Así mismo proporciona una interfaz visual que favorece su uso por parte de los alumnos.

Nuestro simulador de archivos emerge como una herramienta invaluable para comprender y optimizar la gestión del espacio en sistemas de almacenamiento. Permitiendo la experimentación sin riesgos, este simulador ofrece una visión detallada de cómo diferentes estrategias impactan en la asignación y liberación de espacio. Desde vectores de bits hasta listas ligadas, la simulación proporciona una plataforma eficaz para evaluar y perfeccionar los métodos de gestión de espacio libre, contribuyendo así al diseño y mejora continua de sistemas de archivos más eficientes y robustos.

La herramienta, invita al alumno a interactuar con la aplicación y permite fácilmente la modificación de algunos de los parámetros, como borrar un archivo o crear uno nuevo.

Se espera que con dicha herramienta el aprendizaje y conocimiento de los alumnos se vea beneficiada, para que así, puedan implementar ejercicios de clase o bien, explorarlo por su cuenta.

Bibliografía

- Team, K. (2023, 6 marzo). ¿Qué es un sistema de ficheros? | KeepCoding Bootcamps. *KeepCoding Bootcamps*. <https://keepcoding.io/blog/que-es-un-sistema-de-archivos/>
- Greyrat, R. (2022, 5 julio). *Métodos de asignación de archivos – Barcelona Geeks*. <https://barcelonageeks.com/metodos-de-asignacion-de-archivos/>
- Greyrat, R. (2022a, julio 5). *Gestión del espacio libre en el sistema operativo – Barcelona Geeks*. <https://barcelonageeks.com/gestion-de-espacio-libre-en-el-sistema-operativo-1/>