

Arquitectura del Computador y Sistemas Operativos

Decimosexta Clase



Almacenamiento secundario

Generalidades

Consiste en un conjunto de dispositivos de almacenamiento (generalmente no volátil) asociados a la memoria principal del ordenador.

Es difícil trazar una línea. Lo que es una constante es que las memorias más rápidas se colocan más cerca del procesador y a medida que se alejan van siendo más lentas:

- Cache nivel 1
- Cache nivel 2
- Memoria DRAM
- Discos
- Cintas almacenadas en una bóveda

Almacenamiento secundario

Discos magnéticos (1/5)

Los discos magnéticos consisten en uno o más discos de material magnético colocados uno sobre otro.

Cada uno de esos discos tiene cabezas lecto-grabadoras (generalmente dos, una por cara).

Cada cabeza es un electroimán, que producen campos magnéticos al aplicarles una corriente eléctrica (para grabación) y generan una corriente eléctrica al ser expuestas a un campo magnético (lectura)

Los discos giran y las cabezas se mueven el saltos discretos. Ésto hace que describan trayectorias cilíndricas a lo largo de la superficie magnética



Almacenamiento secundario

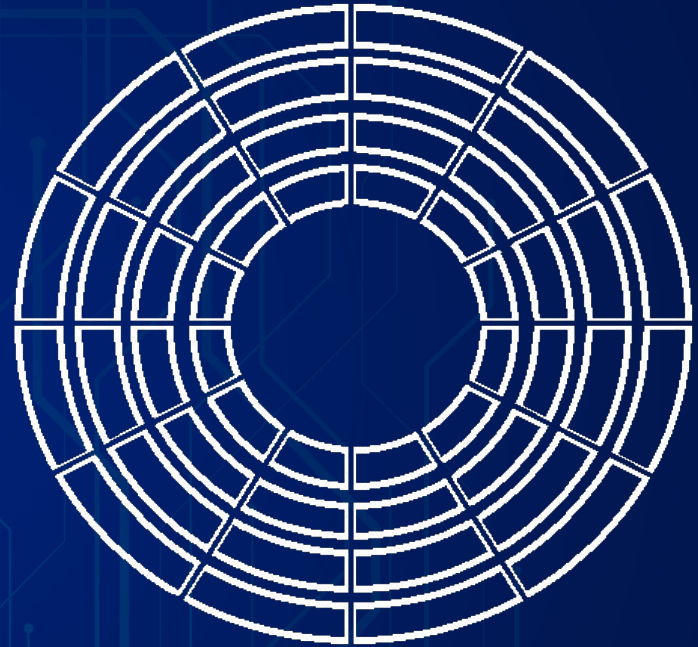
Discos magnéticos (2/5)

La información tampoco se guarda de forma continua dentro del cilindro, sino en bloques discretos llamados sectores.

La **geometría** de un disco es la cantidad de: cilindros, caras y sectores. A veces referidos como **CHS**, por Cylinder-Head-Sector.

Un **sector** es la menor cantidad de **información** que se puede **leer o escribir**.

A lo largo de la historia se fue aumentando la capacidad de almacenamiento de los dispositivos magnéticos de diversas formas. Se aumentó la cantidad de discos (o platos) y se bajó el lugar que ocupa un sector. Éste último enfoque implicaba grabar más cilindros más angostos y más sectores por cilindro, aumentando la "**densidad de datos**" en el material magnético.



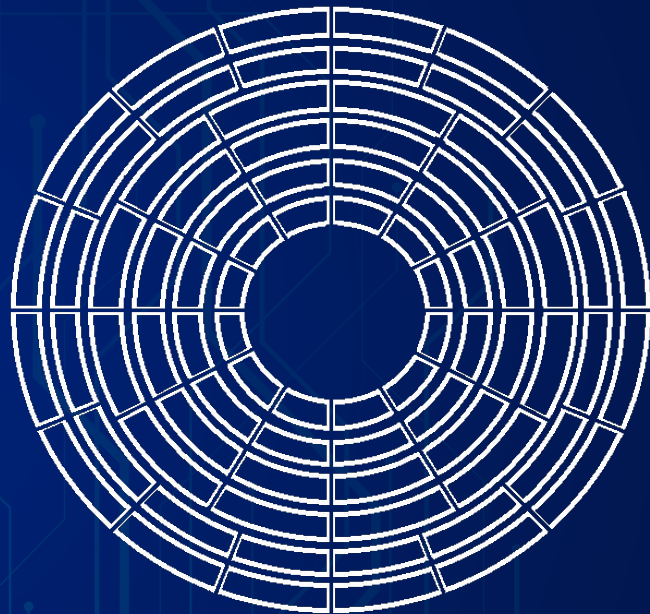
Almacenamiento secundario

Discos magnéticos (3/5)

Aumentar la “densidad de datos” es un límite tecnológico que se ha ido mejorando, pero pronto quedó claro que mantener la misma cantidad de sectores en los cilindros internos y externos era un desperdicio.

Hoy los discos tienen “**zonas**” con distintas cantidades de sectores por cilindro en cada zona.

Los primeros discos exponían su geometría al SO, y esta última técnica hubiera sido un problema. Sin embargo hoy los discos numeran los sectores en forma correlativa en lo que se ha llamado LBA (Logical Block Addressing). Cada disco tiene capacidad de cómputo propio que transforma este número secuencial en los tres números CHS de acuerdo a la geometría de cada zona del disco.



Almacenamiento secundario

Discos magnéticos (4/5)

La forma que vimos de guardar los datos en disco se conoce como “*Conventional Magnetic Recording*” o CMR. Ésta tecnología se topó con un límite físico que impidió seguir aumentando la capacidad:



Cuando se trata de grabar cilindros muy delgados, las líneas magnéticas van directo de un polo del imán al otro sin bajar lo suficiente como para orientar los dipolos magnéticos del material del disco y “grabar” la información.

Almacenamiento secundario

Discos magnéticos (5/5)

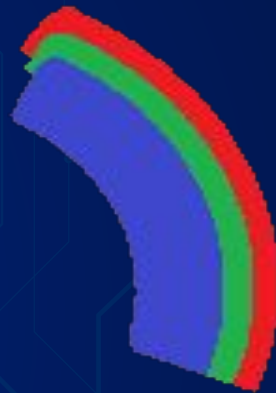
Como no se pueden generar cilindros muy angostos, se optó por escribir superpuesto. En la figura se ven tres cilindros superpuestos y luego un espacio para luego repetir el proceso. De esta forma, el producto resultante es el buscado.

Esta tecnología se conoce como "*Shingled Magnetic Recording*" o SMR.

Sin embargo esto tiene una clara desventaja: Si se deseara modificar el sector rojo del gráfico, no se podría escribirlo solo. Habría que escribirlo y luego reescribir el verde y luego el azul.

O sea re-escribir toda la zona.

Si bien el procesador dentro del disco resuelve ésto en forma transparente, la pérdida de tiempo es inevitable. Estos discos son mucho más lentos que los CMR en estos casos.



Almacenamiento secundario

Discos Ópticos

Otra tecnología usada fueron los discos ópticos. Éstos utilizaban un láser y un sensor de luz para medir la reflexión de la luz en un disco reflectivo. El láser usado a alta potencia calentaba y degradaba la superficie reflectiva para que no refleje.

Al igual que con la clave morse, el láser grababa puntos y rayas que se leían como ceros y unos.

EL problema era que la superficie degradada no se podía volver a hacer reflectiva fácilmente. Peor aún los procesos de borrado que se encontraron no podían borrar solamente un sector pequeño. El tener que pasar por un proceso de borrado total, unido a los errores generados por rayones, y a que las memorias alcanzaran densidades de grabación similares, hicieron que se dejaran de usar.





Almacenamiento secundario

Discos de estado sólido (1/2)

Hoy en día están en el mercado dos tecnologías de discos de estado sólido. Ambas están basados en memorias, pero están estructurados de forma diferente:

Característica	Memoria NAND	Memoria NOR
Parte	MT29F2G16ABBG4H-AIT:G	MT28FW02GBBA1HPC-0AAT
Densidad	Mayor	Menor
Costo	Menor (USD 3)	Mayor (USD 29)
Ciclos Borrado	1 M	100 K
Errores	Muchos, requiere ECC	Pocos
Acceso	Por página (2 KBytes)	Por Word (2 Bytes)
Tiempo Borrado	700 μ s	200 ms
Tiempo R	20 ns	100 ns
Tiempo W	20 ns	60 ns



Almacenamiento secundario

Discos de estado sólido (2/2)

Las memorias NAND tienen mayor densidad de almacenamiento de datos, por lo que se usan para discos SSD y Pen Drivers, aunque su gran tasa de fallas requiere que se apoyen en ECCs para mantener el TMEF (Tiempo Medio Entre Fallas) alto.

Las memorias NOR permiten acceder byte por byte (no de a páginas), por lo que se usan para conectarlas en lugar de ROM a procesadores, como en el caso de los celulares.

Los procesos de borrado de ambas tecnologías son por bloque. Ésto requiere que si se va a tocar un solo byte hay que **reescribir toda la zona**.

Almacenamiento secundario

Soporte del SO a los dispositivos con zonas

Si bien los discos tienen procesadores que “resuelven” el problema de reescribir una zona completa, si el SO interviene puede mejorar la performance.

Desde la versión 4.10.0 Linux tiene “Linux Kernel Zoned Storage Support” que es un módulo especialmente diseñado para optimizar accesos a discos magnéticos de tipo SMR y/o unidades SSD.

La tecnología permite varias optimizaciones:

- A nivel escritura de drivers
- A nivel filesystem
- A nivel de device-mapper



Estructuras lógicas dentro del disco

Como vimos, los discos, son un conjunto de sectores, todos del mismo tamaño.

Estos sectores, de acuerdo al estándar LBA, se enumeran secuencialmente de 0 hasta un determinado número que depende del tamaño del disco.

El desafío ahora es cómo convertimos ese enorme conjunto de sectores sueltos en una unidad utilizable por los programas.

La primer estructura que se guarda en el disco es la tabla de particiones. Permite al que la lee:

- Identificarlo como tal
- Entender divisiones que permitan tratar un disco grande como varios más pequeños.
- Iniciar el SO desde el disco



Tabla de Particiones

Alternativas

Hay dos formatos de tabla de particiones:

- MBR (Master Boot Record)
- GPT (GUID Partitioning Table)



Tabla de Particiones

Master Boot Record

El que lee la tabla debe leer el bloque de 512 bytes y verificar la firma, que debe valer 0x55 0xAA.

Si es correcta asume que es válida, y puede analizar las interpretas la tabla de particiones.

Cada entrada en la tabla de particiones se interpreta así:

Active	Libre	CHS Primer Sector	Tipo	CHS Último Sector	LBA Primer Sector	Número de sectores
1bit	7 bits	24 bits	8 bits	24 bits	32 bits	32 bits

Primero se usaba CHS, pero limitaba a:

$$2^{24} \cdot 512 = 8\text{Gb}$$

Se cambió a LBA, que limita a:

$$2^{32} \cdot 512 = 2\text{Tb}$$

Ejercicio:

Ejecute `hexdump -v -C -n512 /dev/sda` en su máquina Linux. Analice las particiones.

0 - 439

Bootloader
Program

Disk Id

Partición 1

Partición 2

Partición 3

Partición 4

Firma

440 - 443

444 - 445

446 - 461

462 - 477

478 - 493

494 - 510

510 - 511

Tabla de Particiones

GUID Partition Table

SO

Aparecieron en el año 2018, principalmente para sobrellevar el límite de 2 Tb.

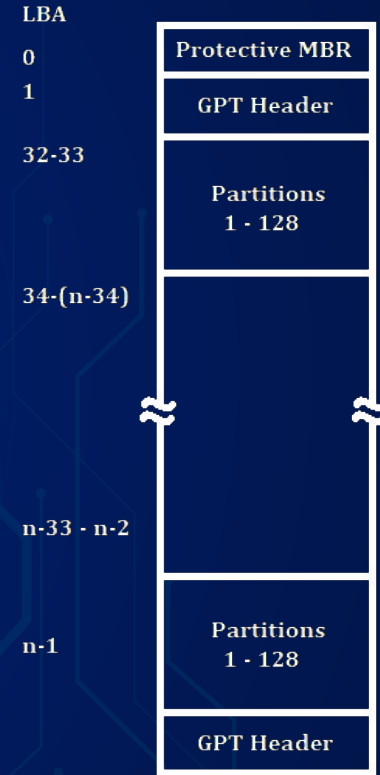
Incluye un MBR por compatibilidad con herramientas viejas, pero debe estar todo en cero salvo la firma y la primera partición que debe ocupar todo el disco y ser de tipo 0xEE.

Ya no tiene lugar para Bootloader Program.

Incluye una copia de la tabla al final del disco por si se corrompe la primera.

Permite un máximo de 128 particiones con el siguiente descriptor:

Tipo	GUID de la Partición	LBA Primer Sector	LBA Último Sector	Flags	Nombre
16 bytes	16 bytes	8 bytes	8 bytes	8 bytes	72 bytes



An abstract pattern of light blue lines and dots on a dark blue background, resembling a circuit board or a network diagram, located on the left side of the slide.

Fin
¿Preguntas?