Memoria Externa

Presenta: **Héctor Tejeda Villela**

Introducción

Se examinan distintos sistemas y dispositivos de memoria externa. El dispositivo más importante es el disco magnético, ya que constituyen la base de las memorias externas en la mayoría de los equipos. Se revisa también el uso de conjuntos de discos para conseguir mayores prestaciones, en particular la familia conocida como RAID ("Redundant Array of Independent Disks" conjunto redundante de discos independientes). Por su parte la memoria óptica externa se ha constituido en un componente cada vez más importante. También se hace una revisión de las cintas magnéticas.

Discos Magnéticos

Es un plato circular construido con metal o plástico cubierto por un material magnetizable. Los datos se graban y se recuperan a través de una bobina llamada *cabeza*. En una operación de lectura o escritura, la cabeza permanece quieta mientras el plato rota bajo ella.

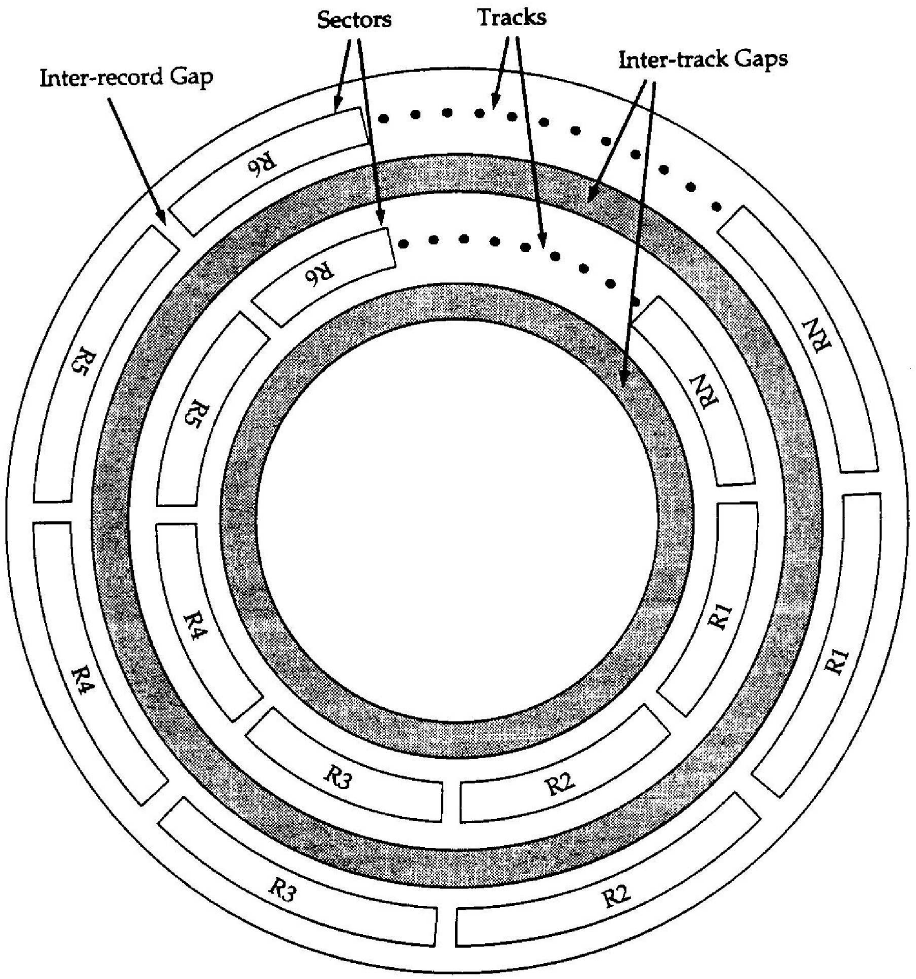
El mecanismo de escritura se basa en el campo magnético producido por el flujo eléctrico que atraviesa la bobina. Se envían pulsos a la cabeza, y se graban patrones magnéticos en la superficie bajo ella, con patrones diferentes para corrientes positivas y negativas. El mecanismo de lectura se basa en la corriente eléctrica que atraviesa la bobina, producida por un campo magnético que se mueve respecto a la bocina; la corriente que se genera es de la misma polaridad que la que produjo la grabación magnética.

Organización de datos y formato

Los datos se organizan en un conjunto de anillos concéntricos en platos, llamados *pistas*, que son del mismo ancho que la cabeza. Las pistas adyacentes están separadas por *bandas vacías*, para minimizar los errores debidos a desalineamientos de la cabeza o por interferencias del campo magnético. Se acostumbra almacenar el mismo número de bits en cada pista para simplificar la electrónica, por lo que la *densidad* (bits por pulgada lineal) aumenta conforme se va de la pista más externa a la más interna.

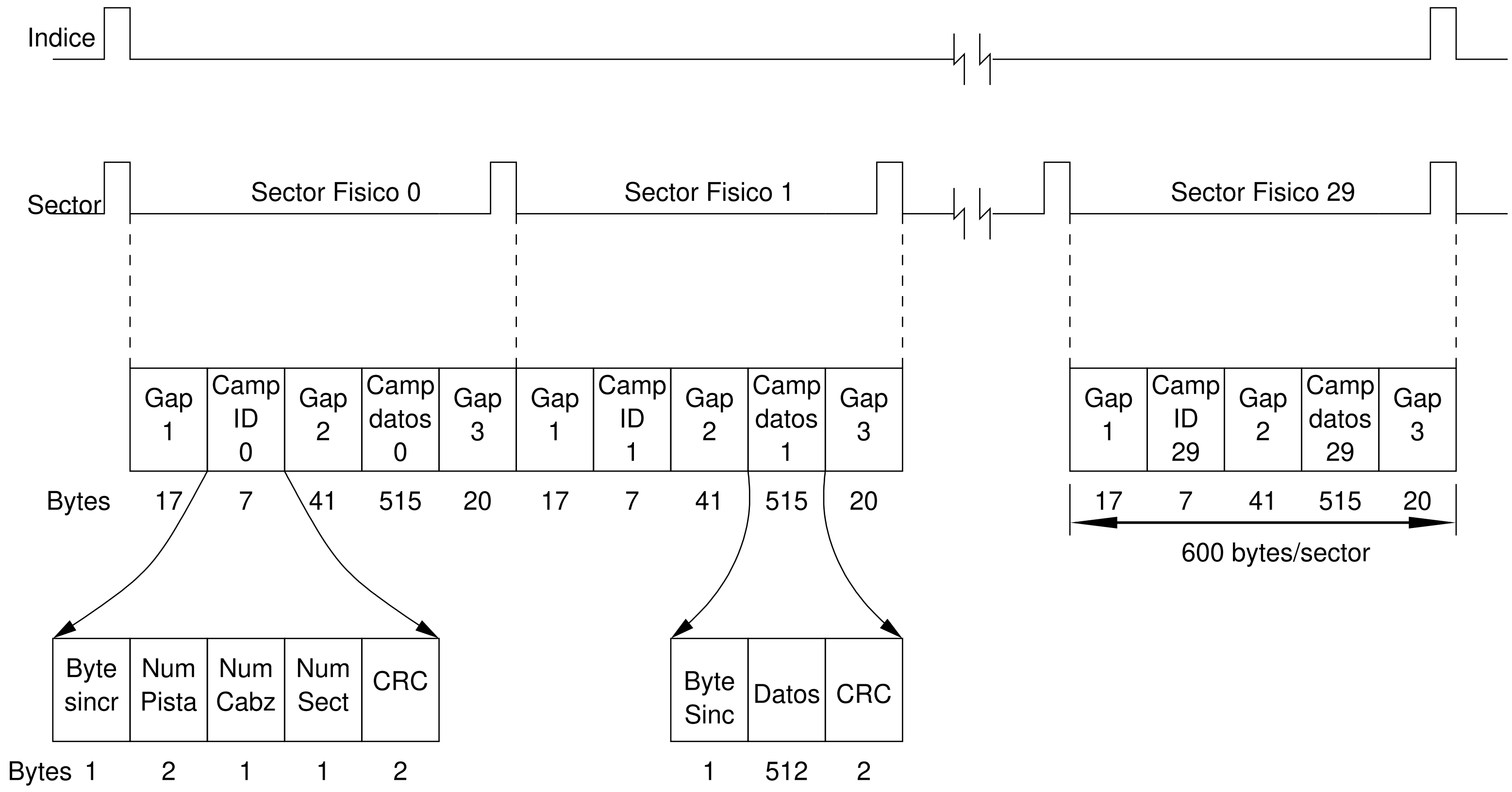
Los datos se transfieren en bloques. El bloque generalmente es de menor capacidad que una pista. Las regiones del tamaño de un bloque se llaman *sectores*. Hay entre 10 y 100 sectores por pista pudiendo ser su longitud fija o variable. Los sectores adyacentes se separan con interbloques (“inter-record-gaps”).

Organización de los datos en un disco



Para identificar las posiciones de los sectores dentro de una pista dada, se hace a partir de un punto de comienzo y del identificador de principio y fin de cada sector. Lo anterior es gestionado con datos de control grabados, por lo tanto, el disco se graba con un formato que contiene algunos datos adicionales y que no son accesibles al usuario.

Formato de las pistas de un disco Winchester (Seagate ST506)



Características de los sistemas de discos

Desplazamiento de cabezas

—Cabeza fija (una x pista)

—Cabeza móvil (una por superficie)

Transportabilidad de disco

—Disco fijo

—Disco extraíble

Superficies

—Superficie única

—Superficie doble

Platos

—Plato único

—Múltiples platos

Mecanismo de la cabeza

—Contacto (disquete)

—Separación fija

—Separación aerodinámica

Tiempo de acceso a disco

Cuando la unidad de disco está operando, el disco rota a una velocidad constante. Para leer o escribir, la cabeza debe posicionarse en la pista deseada y al principio del sector deseado en la pista. La selección de la pista implica un movimiento de la cabeza, en un sistema de cabeza móvil, o una selección electrónica de una cabeza, en un sistema de cabezas fijas. En un sistema de cabeza móvil, el tiempo que tarda la cabeza en posicionarse en la pista se conoce como *tiempo de búsqueda*. Una vez que se ha seleccionado la pista, el sistema debe esperar hasta que el sector apropiado rote hasta alinearse con la cabeza. El tiempo que tarda el sector en alcanzar la cabeza se llama *latencia rotacional*. La suma del tiempo de búsqueda y la latencia rotacional se denomina *tiempo de acceso*.

RAID

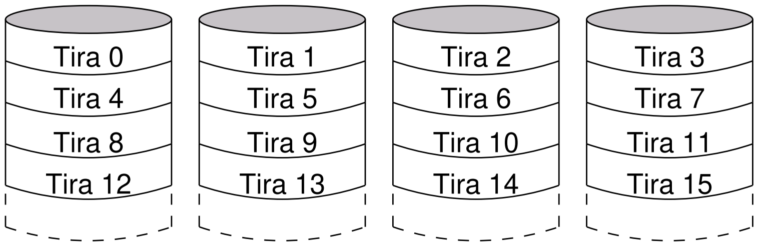
El sistema de memoria de disco es el principal foco de optimización en las prestaciones de las computadoras. Para tener una ganancia adicional en prestaciones se usan conjuntos de discos que operen indenpendientemente y en paralelo. En donde la estrategia es sustituir una unidad de disco de gran capacidad por varias unidades menores y distribuir los datos de forma que se habiliten accesos simltáneos a los datos.

Empleando varios discos hay varias formas en las que se pueden organizar los datos y a los que se puede agregar redundancia para mejorar la seguridad. El esquema RAID tiene seis niveles independientes (0-5). Las características comunes son:

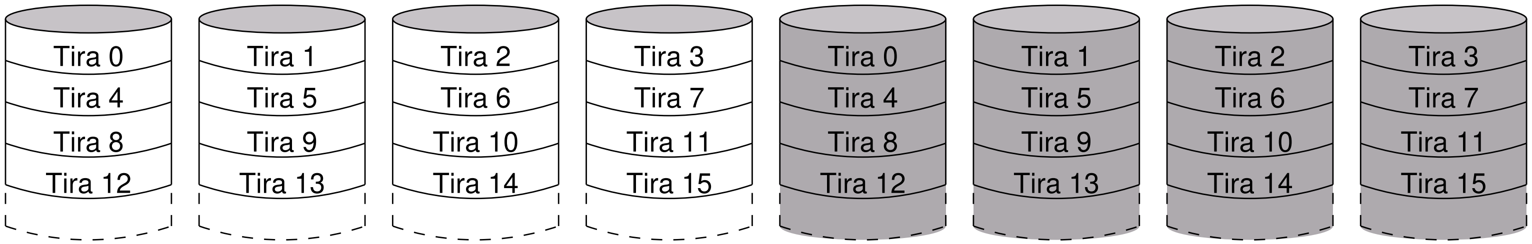
1. Es un conjunto de unidades físicas vistas por el sistema operativo como una única unidad lógica.
2. Los datos se distribuyen en las unidades.
3. Los discos redundantes son usados para almacenar información de paridad.

La 2^a y 3^a característica cambian según los distintos niveles RAID. RAID 0 no soporta la tercera característica.

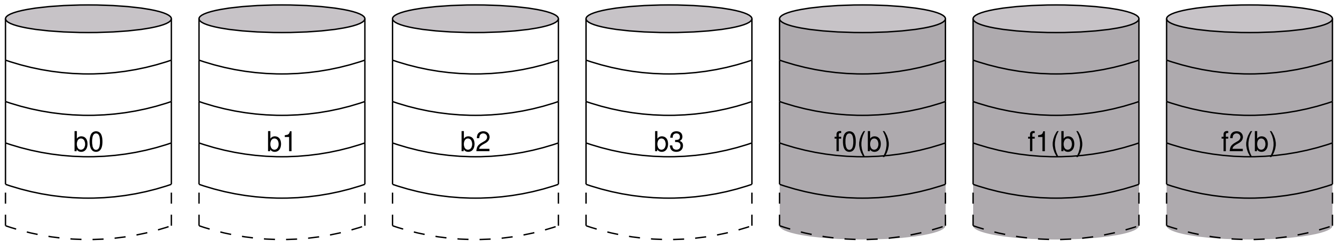
Niveles 0-2



RAID 0 (no redundante)



RAID 1 (Con espejo)



RAID 2 (Redundante con código Hamming)

Nivel 0 de RAID

No se considera como un verdadero miembro de la familia RAID por no incluir redundancia. El usuario y los datos del sistema están distribuidos a lo largo de todos los discos, con la ventaja frente al uso de un único y gran disco, ya que la E/S se puede emitir en paralelo para varios discos, con lo que se reduce el tiempo de cola.

Los datos se *organizan en forma de tiras de datos*. Las tiras se obtienen dividiendo el disco en bloques físicos, sectores o alguna otra unidad. Las tiras se proyecta cíclicamente, en miembros consecutivos del conjunto. Un conjunto de tiras lógicamente consecutivas, que proyectan exactamente, una tira en cada miembro del conjunto, se denomina franja.

RAID 0 para alta capacidad de transferencia

Los dos requisitos para lograrlo son:

1. Debe existir una capacidad de transferencia alta entre los discos individuales y los controladores de buses internos, buses de E/S, adaptadores de E/S y buses de memoria del anfitrión, y

2. La aplicación que haga peticiones de E/S controle el conjunto de discos eficientemente. Lo anterior se cumple si la petición típica es de una gran cantidad de datos lógicamente.

RAID 0 para altas velocidades de petición de E/S

En los entornos orientados a transacciones donde puede haber cientos de peticiones de E/S por segundo, la parte crítica es el tiempo de respuesta y no la velocidad de transferencia. Para una petición individual de una pequeña cantidad de datos, el tiempo de E/S se debe al tiempo de búsqueda y la latencia rotacional.

La forma de conseguir velocidades altas de ejecución es balanceando la carga de E/S a través de los distintos discos, donde hay varias aplicaciones independientes o una única aplicación orientada a transacción que es capaz de generar varias peticiones de E/S asíncronas. Si la franja es relativamente grande, donde una petición única requiera un único acceso a disco, entonces las peticiones pueden ser tratadas en paralelo.

RAID 1

RAID 1 se diferencia de los niveles 2 al 5 en la forma de obtener la redundancia. Para logra la redundancia se duplican todos los datos. Cada franja lógica se proyecta en dos discos físicos separados.

Las ventajas que se tienen son:

1. Una petición de lectura puede ser servido por cualquiera de los discos que contenga los datos, con cualquiera de ellos se tiene un tiempo de búsqueda mínimo más la latencia rotacional.
2. La escritura requiere que dos tiras se actualizen, lo que se puede hacer en paralelo.
3. La recuperación tras un fallo es sencilla. Cuando una unidad falla, se puede acceder a los datos desde la otra unidad.

La principal desventaja es el costo, se requiere el doble de espacio, por lo que es empleado en almacenar software del sistema y datos altamente críticos, ya que se proporcion una copia de seguridad en tiempo real.

Se consiguen altas velocidad de petición si la mayoría son lecturas. Si las peticiones son de escritura puede no ser significativa respecto a RAID 0. También tiene una mejora respecto a RAID 0 si en aplicaciones de transferencia intensiva se tiene un alto porcentaje de lectura.

Nivel 2 de RAID

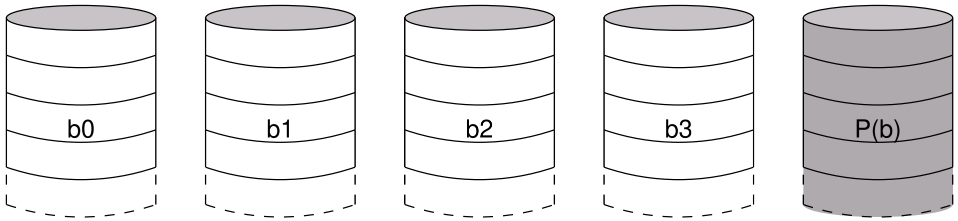
Se usa una técnica de acceso paralelo, donde todos los discos participan en la ejecución de cada petición de E/S, el giro de cada unidad individual está sincronizado, por lo que cada cabeza de disco está en la misma posición para un instante dado.

Se usa la descomposición de datos en tiras muy pequeñas (como un único byte o palabra). El código de corrección de errores se calcula a partir de los bits de cada disco, y los bits del código se almacenan en las correspondientes posiciones de bit en varios discos de paridad. Se usa el código Hamming para corregir errores en un bit y detectar errores en dos bits.

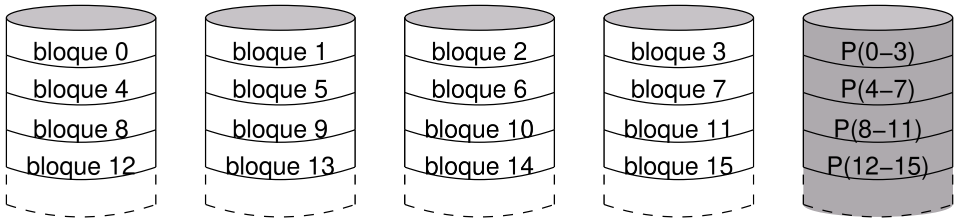
El número de discos redundantes es proporcional al logaritmo del número de discos de datos. El controlador del conjunto proporciona los datos pedidos y el código de corrección de errores asociado. Si hay un error en un solo bit, el controlador lo puede reconocer y corregir instantáneamente, con lo que el tiempo de acceso a lectura no se ralentiza.

Se recomienda en un entorno en el que haya muchos errores de disco. Si hay una alta seguridad en los discos individuales y en las unidades de disco, es excesivo y no se aconseja implementar.

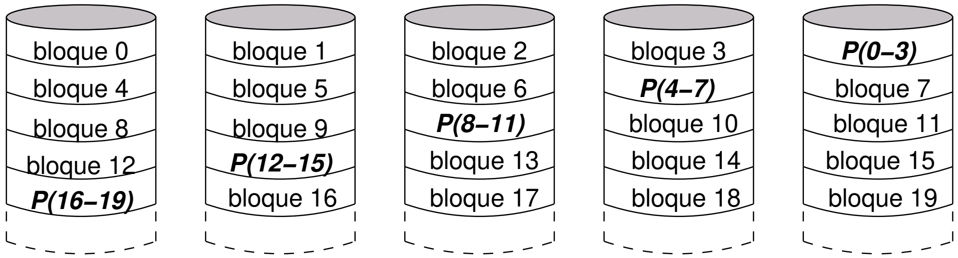
Niveles 3-5



RAID 3 (Bit de paridad intercalado)



RAID 4 (Paridad en bloques)



RAID 5 (Paridad distribuida a nivel de bloques)

Nivel 3 de RAID

Organizado en forma similar a RAID 2, con la diferencia de que se requiere sólo un disco redundante, no importando el tamaño del conjunto de discos. Utiliza acceso paralelo, con datos distribuidos en pequeñas tiras. En vez del código de corrección se calcula un bit de paridad para el conjunto de bits individuales en la misma posición en todos los discos de datos.

En el caso de un fallo en una unidad, se accede a la unidad de paridad y se reconstruyen los datos desde el resto de los dispositivos. Una vez que se sustituye la unidad que ha fallado. los datos que faltan se restauran en la nueva unidad y se reanuda la operación.

Suponiendo un conjunto de 5 discos donde los primeros son de datos y el último de paridad, para el i-ésimo bit se calcula como:
$X4(i) = X3(i) \oplus X2(i) \oplus X1(i) \oplus X0(i)$
Si la unidad X1 falla, se suma $X4(i) \oplus X1(i)$ a ambos miembros de la ecuación y se tiene:
$X1(i) = X4(i) \oplus X2(i) \oplus X1(i) \oplus X0(i)$

Prestaciones

Como los datos se dividen en tiras muy pequeñas, se tienen velocidades de trasferencia de datos muy altas. Cualquier petición de E/S implicará una trasferencia de datos paralela desde todos los discos de datos.

Sólo se puede ejectuar a la vez una petición de E/S. Para un entorno orientado a transacciones, el rendimiento baja.

RAID 4

Los niveles 4 y 5 de RAID usan una técnica de acceso independiente, donde cada disco opera en forma separada, por lo que las peticiones de E/S separadas se atienden en paralelo. Es adecuado para aplicaciones que requieren velocidades de petición de E/S altas, y son menos adecuados para aplicaciones que requieren velocidades altas de transferencia de datos.

Se usan tiras de datos relativamente grandes. Se calcula una tira de paridad la cual se guarda en el disco de paridad. Cuando se realiza una petición de escritura de E/S pequeña se incurre en una penalización, ya que se deben leer los datos del usuario y la paridad, y después se deben escribir los datos y la paridad.

Las operaciones de escritura usan al disco de paridad, por lo que se hace un cuello de botella.

Nivel 5 de RAID

Organizado de manera similar a RAID 4, la diferencia es que se distribuyen las tiras de paridad a lo largo de todos los discos, con lo que se evita el potencial cuello de botella de E/S de RAID 4.

Niveles RAID

Nivel	*Categoría*	*Descripción*	*Grado de E/S*
0	Estructura en tiras	No redundante	Tiras largas: excelente
1	Estructura en espejo	Espejo	Bueno/Regular
2	Acceso Paralelo	Redundante con cdigo Hamming	Pobre
3	Acceso Paralelo	Bit de paridad intercalado	Pobre
4	Acceso Independiente	Bloque de paridad intercalado	Excelente/Regular
5	Acceso Independiente	Paridad distribuida en bloques intercalados	Excelente/Regular

Niveles RAID (continuación)

Nivel	*Transferencia de datos (Lectura/Escritura)*	*Aplicación Típica*
0	Pequeñas tiras: excelente	Altas prestaciones con datos no críticos
1	Regular/Regular	Controladores de sistemas; ficheros críticos
2	Excelente	
3	Excelente	Aplicaciones con muchas E/S (imágenes, CAD)
4	Excelente/Pobre	
5	Excelente/Pobre	Grado de petición alto, lectura intensiva, consulta de datos.