

Sistemas Operativos

Almacenamiento masivo

Armando Arce, Escuela de Computación, arce@itcr.ac.cr

Tecnológico de Costa Rica

Estructuras de almacenamiento masivo

Estructuras de Almacenamiento Masivo

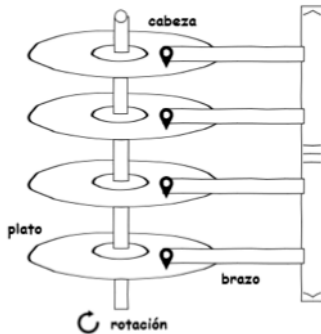


Figure 1:

Los discos magnéticos proporciona la parte principal del almacenamiento secundario en los modernos sistemas informáticos.

- Cada plato que conforma un disco magnético tiene forma circular plana, como un CD.
- Las dos superficies de cada plato están recubiertas de un material magnético. La información se almacena grabándola magnéticamente sobre los platos.

Discos magnéticos

Un cabezal de lectura/escritura “vuela” a muy poca distancia de cada una de las superficies de cada plato.

- Los cabezales están conectados a un brazo de disco que mueve todos los cabezales como un sola unidad.
- La superficie de cada plato está dividida desde el punto de vista lógico en pistas circulares, que a su vez se dividen en sectores.
- El conjunto de pistas que están situadas en una determinada posición del brazo forman un cilindro.

Discos magnéticos

Cuando se está usando el disco, un motor hace que gire a gran velocidad, normalmente entre 60 y 200 revoluciones por segundo.

- La velocidad del disco consiste realmente de dos velocidades independientes: la velocidad de transferencia y el tiempo de posicionamiento.
- La velocidad de transferencia consiste en el tiempo necesario para que fluyan los datos entre el disco y el computador. El tiempo de posicionamiento se divide a su vez en otros dos: tiempo de búsqueda y latencia rotacional.
- El tiempo de búsqueda consiste del tiempo necesario para mover el brazo hasta el cilindro requerido, el tiempo de latencia rotacional consiste del tiempo que dura el sector deseado en rotar hasta pasar por debajo del cabezal del disco.

Discos magnéticos

Aunque los platos de disco están recubiertos de una fina capa protectora, a veces el cabezal puede llegar a dañar la superficie magnética (aterrizaje de cabezales).

- Este tipo de accidente no se puede reparar y es necesario cambiar el disco completo.
- Para conectar una unidad de disco a un computador, se utiliza un conjunto de cables denominados *bus de E/S*.
- Hay disponibles varios tipos de buses, entre ellos: EIDE, ATA, SATA, USB, FC y SCSI.
- Las transferencias de datos en un bus son realizadas por procesadores especiales denominados controladores.

Estructura de un disco

Un disco se direcciona como un gran arreglo de bloques lógicos, en la que dicho bloque es la unidad más pequeña de transferencia.

- Luego es posible convertir el número de bloque lógico en una dirección de disco basado en: número de cilindro, número de pista y número de sector.
- Sin embargo, a veces esto resulta complicado pues pueden existir sectores defectuosos en el disco que sean ser reemplazados por sectores libres en otra zonas del disco.
- Además, en discos modernos, el número de sectores por pista no es constante (menos sectores en las pistas internas).

Estructura de un disco

Algunos dispositivos de almacenamiento como CD-ROM y DVD-ROM utilizan *velocidad lineal constante* (CLV, constant linear velocity) en la cual la densidad de bits por cada pista es uniforme.

- Por ello, con el fin de mantener un flujo constante de bits, el motor de la unidad debe decrementar su velocidad conforme el cabezal se mueve de las pistas internas hacia las externas (igual que un CD de música).
- Otros dispositivos de almacenamiento como los discos magnéticos utilizan *velocidad angular constante* (CAV, constant angular velocity) en la que la velocidad de rotación del disco puede permanecer constante pues la densidad física de bits se reduce al pasar de las pistas internas a las externas.

Las computadoras acceden a los datos almacenados en disco de formas distintas.

- Algunos métodos utilizan un acoplamiento fuerte (mediante el bus interno de la máquina) y otros utilizan métodos de acoplamiento mas débiles (mediante conexiones de red).

El almacenamiento conectado al *host* utiliza puertos de E/S locales para acceder a los discos.

- En arquitecturas IDE o ATA se soportan únicamente dos unidades de disco por puerto de conexión, otras arquitecturas como SCSI y FC (fiber channel) soportan 16 y 126 unidades, respectivamente.

Almacenamiento conectado al host

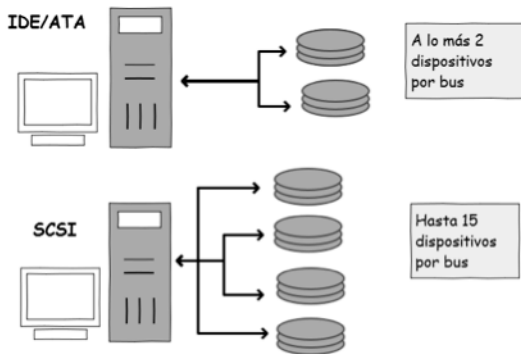


Figure 2:

Los dispositivos de almacenamiento que resultan adecuados para conectar mediante este esquema son: discos duros, matrices RAID, unidades de CD y DVD, y unidades de cinta.

Almacenamiento conectado a la red

Un dispositivo de almacenamiento conectado a la red (NAS, network-attached storage) es un sistema de almacenamiento de propósito especial al que se accede de forma remota a través de una red de datos.

- Los clientes acceden al almacenamiento conectado a la red a través de una interfaz de llamadas a procedimientos remotos, por ejemplo NFS sobre Unix.

Almacenamiento conectado a la red



Figure 3:

Almacenamiento conectado a la red

El almacenamiento conectado a la red proporciona una forma conveniente para que todas las máquinas de una LAN compartan dispositivos de almacenamiento con la misma facilidad que si trataran de dispositivos locales.

- Sin embargo, este sistema tiende a ser menos eficiente y proporciona menos velocidad que la conexión directa al dispositivo.
- Una de las desventajas de este esquema es que las operaciones de E/S para acceso a los dispositivos consumen ancho de banda de la red de datos, incrementando así la latencia de las comunicaciones a través de la red.

Una red de área de almacenamiento (SAN, storage-area network) es una red privada, que utiliza protocolos de almacenamiento en lugar de protocolos de red, que conecta los servidores con las unidades de almacenamiento.

- La principal ventaja de una SAN radica en su flexibilidad, ya que pueden conectarse múltiples *host* y múltiples matrices de almacenamiento a la misma SAN y los recursos de almacenamiento pueden asignarse de forma dinámica a los hosts.

Redes de área de almacenamiento



Figure 4:

Planificación del disco

Una de las responsabilidades del sistema operativo es utilizar de manera eficiente el hardware disponibles.

- En unidades de disco, se puede mejorar tanto el tiempo de acceso como el ancho de banda planificando en el orden adecuado el servicio de las distintas solicitudes de E/S del disco.
- Para un sistema de multiprocesamiento con varios procesos, la cola de disco pueden tener a menudo varias solicitudes pendientes.
- Para seleccionar cuál es la siguiente solicitud que será atendida se pueden usar varios tipos de algoritmos de planificación de disco.

La forma más simple de planificación de disco es el algoritmo FCS (first-come,first-served).

- Este algoritmo es intrínsecamente equitativo, pero generalmente no proporciona el tiempo de servicio más rápido.

Planificación FCFS

cola = 98,183,37,122,14,124,65,67

cabeza empieza en 53

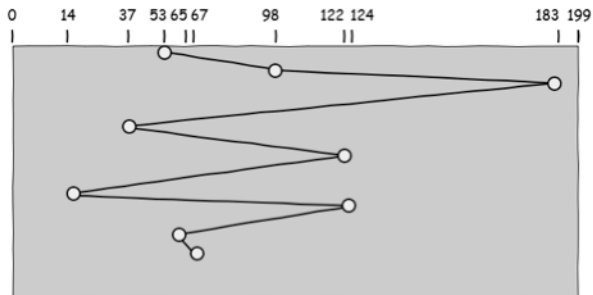


Figure 5:

La amplísima oscilación que se produce entre los cilindros representa el problema existente con esta planificación.

- Si se pudiera satisfacer de forma sucesiva las solicitudes que se encuentran cercanas, la cantidad total de movimiento de los cabezales podría reducirse sustancialmente y la velocidad mejoraría.

Parece razonable satisfacer todas las solicitudes que están próximas a la posición actual del cabezal antes de desplazarlo hasta una posición muy lejana para dar servicio a otras solicitudes.

- Esta suposición es la base para el *algoritmo del tiempo de búsqueda más corto* (SSTF, shortest-seek-time-first).
- Este algoritmo selecciona la solicitud que tenga el tiempo de búsqueda mínimo con respecto a la posición actual del cabezal.
- Este algoritmo proporciona una mejora sustancial de tiempo con respecto al algoritmo FCFS.

Planificación SSTF

cola = 98,183,37,122,14,124,65,67

cabeza empieza en 53

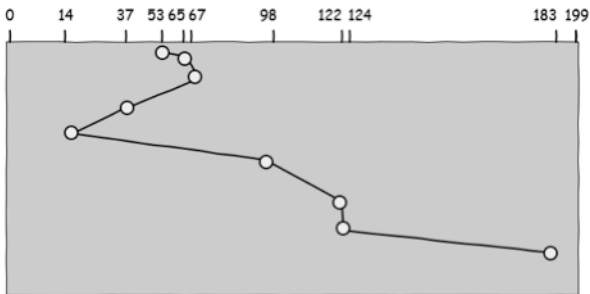


Figure 6:

La planificación SSTF es, esencialmente, un tipo de planificación SJF (shortest-job-first) y, al igual que esta, puede provocar la muerte por inanición de algunas solicitudes, pues podría llegar un flujo continuo de solicitudes que impidan que se atienda una solicitud del final o inicio del disco.

El algoritmo SCAN (o algoritmo de exploración), el cabezal comienza en uno de los extremos del disco y se mueve hacia el otro extremo, dando servicio a las solicitudes a medida que pasa por cada cilindro, hasta llegar al otro extremo del disco.

- En ese otro extremo se invierte la dirección de movimiento del cabezal y se continúa dando servicio a las solicitudes.
- El algoritmo SCAN se denomina también en ocasiones *algoritmo del ascensor* ya que el brazo del disco se comporta como un ascensor dentro de un edificio.

Planificación SCAN

cola = 98,183,37,122,14,124,65,67

cabeza empieza en 53

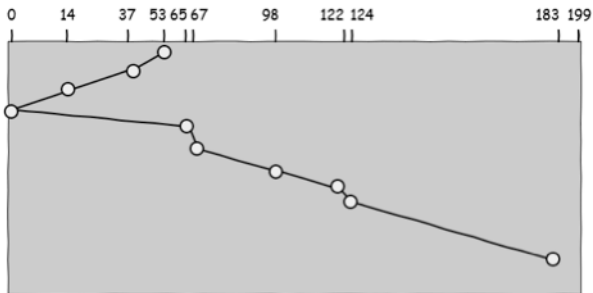


Figure 7:

Algoritmo C-SCAN

La planificación SCAN circular (C-SCAN) es una variante de SCAN diseñada para proporcionar un tiempo de espera más uniforme.

- Al igual que SCAN, C-SCAN mueve el cabezal de un extremo del disco al otro, prestando servicio a las solicitudes a lo largo de ese trayecto.
- Sin embargo, cuando el cabezal alcanza el otro extremo, vuelve inmediatamente al principio del disco, sin dar servicio a ninguna solicitud en el viaje de vuelta.

Algoritmo C-SCAN

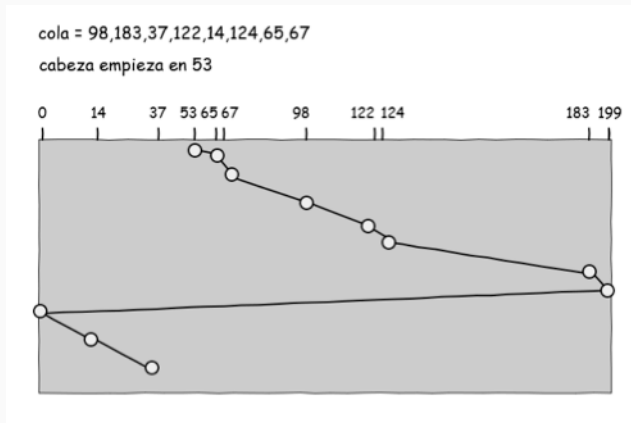


Figure 8:

Tanto el algoritmo SCAN como el C-SCAN mueven el brazo del disco a través de la anchura completa del disco, sin embargo, en la práctica, ninguno de los dos algoritmos se suele implementar de esta manera.

- Lo más común es que el brazo sólo vaya hasta el cilindro correspondiente a la solicitud final en cada dirección.
- Entonces, invierte su dirección inmediatamente, sin llegar hasta el extremo del disco.
- Las versiones de SCAN y C-SCAN que siguen este patrón se denominan planificación LOOK y C-LOOK ya que los algoritmos miran (look) si hay una solicitud antes de continuar moviéndose en una determinada dirección.

Planificación LOOK

cola = 98,183,37,122,14,124,65,67

cabeza empieza en 53

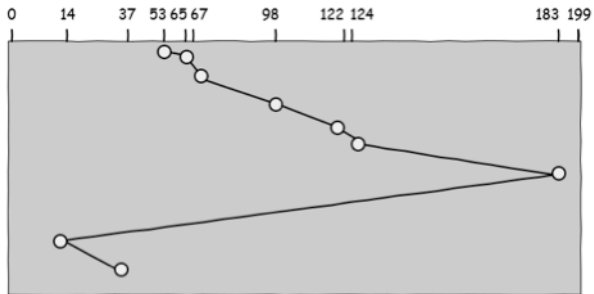


Figure 9:

Selección de un algoritmo de planificación

SSTF resulta bastante común y tiene un atractivo natural, porque mejorar la velocidad que puede obtenerse con FCFS. SCAN y C-SCAN tienen un mejor compartamiento en aquellos sistemas donde el disco está sometido a una intensa carga, porque es menos probable que provoquen problemas de muerte por inanición.

El sistema operativo es responsable también de varios otros aspectos de la administración del disco.

- Por ejemplo, la inicialización del disco, del arranque desde el disco y la recuperación de bloques defectuosos.

Antes de poder almacenar datos en el disco, es necesario dividir éste en sectores que la controlador de disco pueda leer y escribir.

- Este proceso se denomina formato de bajo nivel o formateo físico.
- El formateo de bajo nivel llena el disco con una estructura de datos especial para cada sector.
- La estructura de datos para un sector consta típicamente de una cabecera, un área de datos y una cola.
- La cabecera y la cola contienen información utilizada por la controlador del disco, como el número de sector y un código de corrección de errores (ECC, error-correcting code).

- Cuando la controladora escribe un sector de datos durante la E/S normal, el ECC se actualiza con un valor que se calcula a partir de todos los bytes contenidos en el área de datos.
- Cuando se lee el sector, se recalcula el ECC y se lo compara con el valor almacenado.
- Si los valores almacenado y calculado son diferentes, esa diferencia indica que el área de datos del sector se ha corrompido y que el sector de disco puede ser defectuoso.

Para utilizar un disco para almacenar archivos, el sistema operativo sigue necesitando poder grabar sus propias estructuras de datos en el disco y para ello sigue un proceso de dos pasos.

- El primer paso consiste en particionar el disco en uno o más grupos de cilindros.
- El sistema operativo puede tratar cada partición como si fuera un disco distinto.
- Por ejemplo, Una partición puede albergar una copia del código ejecutable del sistema operativo, mientras que otra puede almacenar los archivos de usuario.

Después del particionamiento, el segundo paso es el formato lógico (o creación de un sistema de archivos).

- En este paso, el sistema operativo almacena las estructuras de datos iniciales del sistema de archivos en el disco.
- Estas estructuras de datos pueden incluir mapas de espacio libre y asignado (una tabla FAT o una serie de inodos) y un directorio inicial vacío.

Para aumentar la eficiencia, la mayoría de los sistemas de archivos agrupan los bloques en una serie de fragmentos de mayor tamaño, que frecuentemente se denominan *clusters*.

- La E/S de disco se realiza mediante bloques, pero la E/S del sistema de archivos se realiza mediante *clusters* garantizando así que la E/S tenga un acceso con características más secuenciales y menos aleatorios.

Bloque de arranque

Para que una computadora empiece a operara debe tener un programa inicial para ejecutar. Este programa inicial de arranque tiende a ser muy simple.

- Se encarga de inicializar todos los aspectos del sistema, desde registros de la CPU hasta las controladores de dispositivo y el contenido de la memoria principal, y luego arranca el sistema operativo.
- Para llevar a cabo su tarea, el programa de arranque localiza el *kernel* del sistema operativo en el disco, carga dicho kernel en memoria y salta hasta una dirección inicial con el fin de comenzar la ejecución del sistema operativo.

Para la mayoría de la computadoras, el programa de arranque está almacenado en memoria de sólo lectura (ROM, read-only memory).

- El programa cargador completo se almacena en los “bloques de arranque”, en una ubicación fija del disco. Un disco que tenga una partición de arranque se denomina *disco de arranque* o *disco del sistema*.

Bloque de arranque

El código de la ROM de arranque ordena al controlador de disco que lea los bloques de arranque en memoria (todavía no hay ningún controlador de dispositivo cargado en ese punto) y luego comienza a ejecutar dicho código.

- El programa de arranque completo es más sofisticado que el cargador de arranque almacenado en la ROM; el programa de arranque completo es capaz de cargar todo el sistema operativo desde una ubicación no fija dentro del disco y luego iniciar la ejecución del sistema operativo.
- A pesar de ello, el código del programa de arranque completo puede ser pequeño.

Puesto que los discos tienen partes móviles y tolerancias muy pequeñas son bastante propensos a los fallos.

- Lo más frecuente es que uno o más sectores pasen a estar defectuosos.
- Dependiendo del disco y de la controladora que se utilice, estos bloques se gestionan de diversas formas.

Sustitución de sectores

La controlador mantiene una lista de bloques defectuosos en el disco.

- La lista se inicializa durante el formateo de bajo nivel y se actualiza a todo lo largo de la vida del disco.
- El formateo a bajo nivel también reserva una serie de sectores adicionales que no son visibles para el sistema operativo.
- Se puede ordenar a la controlador que sustituya lógicamente cada sector defectuoso por uno de los sectores adicionales reservados.
- Este esquema se denomina *reserva de sectores* o *sustitución de sectores*.

Reserva de sectores en cada cilindro

La *sustitución de sectores* podría invalidar las optimizaciones realizadas por el algoritmo de planificación del disco del sistema operativo.

- Por esta razón, la mayoría de los discos se formatean de modo que se proporcionen unos cuantos sectores reservados en cada cilindro y también un cilindro reservado.
- Cuando se remapea un bloque defectuosos, la controlador utiliza un sector reservado del mismo cilindro, siempre que sea posible.

Deslizamiento de sectores

Como alternativa al mecanismo de reserva de sectores, algunas controladores ofrecen la posibilidad de sustituir un bloque defectuoso por el procedimiento de *deslizamiento de sectores*.

- Si el bloque 17 estuviera defectuoso y el primer sector reservado disponible está situado después del sector 202, entonces, el procedimiento de deslizamiento de sectores “remapea” todos los sectores entre el 17 y el 202, moviéndolos una posición.
- O sea, el sector 202 se copia en el sector reservado, luego el sector 201 se copia en el sector 202, el 203, etc.; hasta que el sector 18 se copia en el sector 19.
- Deslizar los sectores de esta forma libera el espacio del sector 18, por lo que se puede mapear el sector 17 sobre el espacio que aquel ocupaba.

La administración de *espacio de intercambio* es otra tarea de bajo nivel del sistema operativo.

- La memoria virtual utiliza el espacio de disco como una extensión de la memoria principal.
- Puesto que el acceso a disco es mucho más lento que el acceso a memoria, la utilización del espacio de intercambio reduce significativamente las prestaciones del sistema.

El espacio de intercambio se utiliza de diversas formas en los distintos sistemas operativos, dependiendo de los algoritmos de administración de memoria que se empleen.

- Los sistemas de paginación pueden, sencillamente, almacenar las páginas que hayan sido expulsadas de la memoria principal.

Ubicación del intercambio

El espacio de intercambio puede residir en uno de dos lugares: puede construirse a partir del sistema de archivos normal o puede residir en una partición de disco separada.

- Si el espacio de intercambio es simplemente un archivo de gran tamaño dentro del sistema de archivos, puede usarse las rutinas normales del sistema de archivos para crearlo, nombrarlo y asignarle el espacio.
- Esta técnica, aunque resulta fácil de implementar, también es poco eficiente.
- Se puede mejorar el rendimiento almacenando en caché la información de ubicación de los bloques dentro de la memoria física y utilizando mecanismos especiales para asignar bloques físicamente contiguos para el archivo de intercambio.

Ubicación del intercambio

Alternativamente, puede crearse un espacio de intercambio en una partición sin formato separada, ya que en este espacio no se almacena ningún sistema de archivos ni estructura de directorio.

- En lugar de ello, se utiliza un administrador de almacenamiento independiente para el espacio de intercambio, con el fin de asignar y desasignar los bloques de la partición sin formato.
- Sin embargo, para añadir más espacio de intercambio, es necesario volver a particionar el disco o añadir otro espacio de intercambio en algún otro lugar.

Las unidades de disco han ido evolucionado para hacerse más pequeñas y más baratas, por lo que ahora es económicamente factible conectar muchos discos a un mismo sistema informático.

- Disponer de gran número de discos en un sistema permite mejorar la velocidad con la que los datos pueden leerse o escribirse, siempre y cuando los discos se operen en paralelo.
- Además, esta configuración también permite mejorar la fiabilidad del almacenamiento de los datos, ya que puede almacenarse información redundante en ese conjunto de múltiples discos.
- De este modo, el fallo de uno de los discos no conduce a la pérdida de datos.

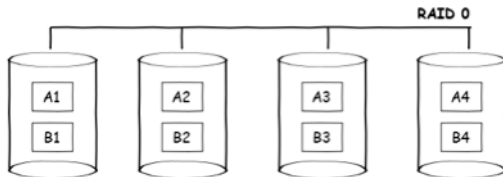
Existen diversas técnicas de organización de discos, colectivamente denominadas matrices redundantes de discos de bajo costo (RAID, redundant array of inexpensive disks), que se utilizan comúnmente para resolver estas cuestiones de velocidad y de fiabilidad.

Se han propuesto numerosos esquemas para proporcionar redundancia a un menor coste utilizando la idea de distribución en bandas combinada con bits de “paridad”.

- Estos esquemas tienen diferentes tipos de compromisos entre el coste y las prestaciones y se clasifican de acuerdo con una serie de niveles denominados niveles RAID.
- A continuación se describirán los diversos niveles existentes:

El mecanismo RAID nivel 0 hace referencia a matrices de discos con distribución de bandas en el nivel de bloques, pero sin ninguna redundancia (no tiene ni mecanismos de duplicación en espejo ni bits de paridad).

RAID nivel 0



Características:

- * Sin redundancia
- * Sin duplicación en espejo
- * Sin bits paridad

Ventajas:

- * Aumenta rendimiento
- * Crea un disco lógico grande a partir de varios discos físicos pequeños

Figure 10:

La distribución en bandas consiste en dividir los bits de cada byte o bloque entre múltiples disco; dicha distribución se denomina *distribución en bandas de nivel de bit*.

- Con este tipo de organización, todos los discos participan en cada acceso.

RAID nivel 1

RAID nivel 1 hace referencia a la duplicación en espejo de los discos.

- Con este sistema se almacena información adicional que no es normalmente necesarios pero que puede utilizarse en caso de que falle el disco, para reconstruir la información perdida.
- Esta técnica duplica cada uno de los discos, así cada disco lógico está compuesto por dos disco físicos y toda escritura se llevará a cabo en ambos discos a la vez.
- Si uno falla, pueden leerse los datos del otro disco. Los datos sólo se perderán si el segundo disco fallara antes de que el primer disco estropeado se reemplace.

RAID nivel 1

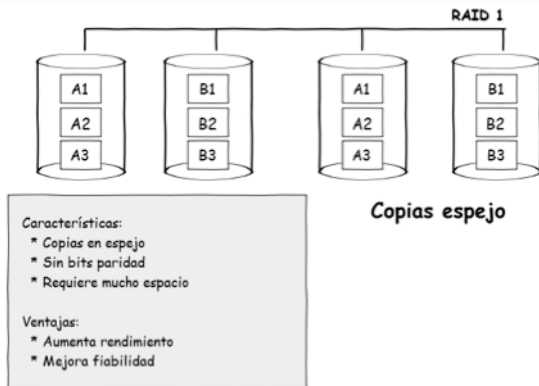


Figure 11:

RAID nivel 2 también se conoce como organización con códigos de corrección de errores (ECC) de tipo memoria.

- En los sistemas de memoria hace ya tiempo que se detectan ciertos errores utilizando bits de paridad.
- Cada byte de un sistema de memoria puede tener un bit de paridad asociado que registra si el número de bits del byte con valor igual a 1 es par (paridad = 0) o impar (paridad = 1).

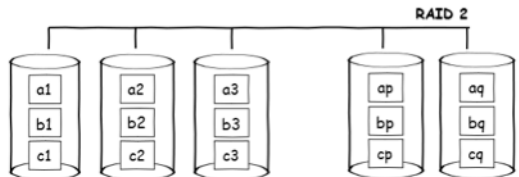
Si uno de los bits del byte resulta dañado (porque un 1 se transforma en un 0 o un 0 se transforma en un 1), la paridad del byte cambia y, por tanto, no se corresponderá con el valor de paridad almacenado.

- De forma: similar, si resulta dañado el bit de paridad almacenado, no se corresponderá con el valor de paridad calculado.
- De este modo, el sistema de memoria puede detectar todos los errores de un único bit.

Los esquemas de corrección de errores almacenan dos o más bits adicionales y pueden reconstruir los datos si resulta dañado un único bit.

- La idea de los códigos ECC puede usarse directamente en las matrices de discos distribuyendo los bytes entre los distintos discos.

RAID nivel 2



Discos de paridad

Características:

- * Usa entrelazado a nivel de bits
- * Usa múltiples bits de paridad
- * La E/S requiere de todos los discos
- * Requiere de mucho espacio

Ventajas:

- * Aumenta rendimiento
- * Mejora fiabilidad

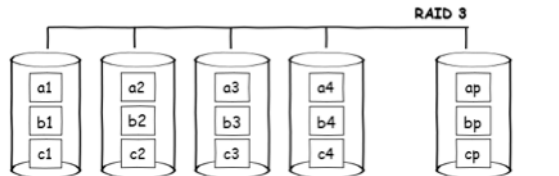
Figure 12:

El esquema de organización RAID nivel 3, u organización de paridad con entrelazado de bits, representa una mejora con respecto al nivel 2, porque tiene en cuenta el hecho de que, a diferencia de los sistemas de memoria, las controladoras de disco pueden detectar si un sector ha sido leído correctamente, por lo que puede utilizarse un único bit de paridad tanto para detección como para corrección de errores.

La idea es la siguiente: si uno de los sectores está dañado, se sabe exactamente qué sector es, y se puede determinar si algún bit del sector es un 1 o un 0 calculando la paridad de los bits correspondientes de los sectores almacenados en los otros discos.

- Si la paridad de los bits restantes es igual a la paridad almacenada, el bit que falta será un 0; en caso contrario, será un 1. RAID nivel 3 ofrece

RAID nivel 3



Disco de paridad

Características:

- * Usa entrelazado a nivel de bits
- * Usa un único bit de paridad
- * La E/S requiere de todos los discos

Ventajas:

- * Aumenta rendimiento
- * Mejora fiabilidad

Figure 13:

El esquema de organización RAID nivel 4, u organización de paridad con entrelazado de bloques, utiliza una distribución en bandas de nivel de bloque, como RAID 0, y además mantiene un bloque de paridad en un disco separado, con información de paridad para los bloques correspondientes de los otros N discos.

- Si uno de los discos falla, puede usarse el bloque de paridad con los bloques correspondientes de los otros discos, para restaurar los bloques del disco fallido.

RAID nivel 4



Características:

- * Usa entrelazado de bloque
- * Usa códigos de corrección de errores
- * El disco de paridad es un cuello de botella en las escrituras

Ventajas:

- * Permite que falle un disco
- * Mejora fiabilidad

Figure 14:

Una lectura de un bloque sólo accede a un disco, permitiendo que los otros discos procesen otras solicitudes.

- Así, la tasa de transferencia de datos para cada acceso es más lenta) pero pueden realizarse múltiples accesos de lectura en paralelo, lo que proporciona una mayor tasa de E/S global.
- Las tasas de transferencia para las lecturas de grandes volúmenes de datos son altas, ya que pueden leerse todos los discos en paralelo; las escrituras de grandes volúmenes de información también tienen altas tasas de transferencia, ya que los datos, y la paridad pueden escribirse en paralelo.

Las escrituras independientes de pequeños volúmenes de datos no pueden realizarse en paralelo.

- Una escritura por parte del sistema operativo de un conjunto de datos cuyo tamaño sea inferior a un bloque requerirá leer el bloque, modificarlo con los nuevos datos y volverlo a escribir.
- También será necesario actualizar el bloque de paridad.
- Esto se conoce con el nombre de ciclo de lectura-modificación-escritura.
- Así, una única escritura requiere cuatro accesos de disco. Dos para leer los dos bloques antiguos y dos para escribir los dos nuevos bloques.

RAID nivel 5

El esquema de organización RAID nivel 5, o paridad distribuida con entrelazado de bloque, difiere del nivel 4 en que se distribuyen los datos y la información de paridad entre los $N + 1$ discos, en lugar de almacenar los datos en N discos y la información de paridad en un disco.

- Para cada bloque, uno de los discos almacena la información de paridad y los otros almacenan los datos.
- Por ejemplo, con una matriz de cinco discos, la paridad del n -ésimo bloque se almacena en el disco $(n \bmod 5) + 1$; los bloques n -ésimos de los otros cuatro discos almacenan los datos reales correspondientes a dicho bloque.

Un bloque de paridad no puede almacenar información de paridad para bloques situados en el mismo disco, porque un fallo del disco provocaría una pérdida de los datos, además de perderse la información de paridad, por lo que dicha pérdida no sería recuperable.

- Distribuyendo la información de paridad entre todos los discos del conjunto, RAID 5 evita el uso potencial excesivo de un único disco de paridad que puede tener lugar en RAID 4.
- RAID 5 es el sistema RAID con paridad que más comúnmente se usa.

RAID nivel 5

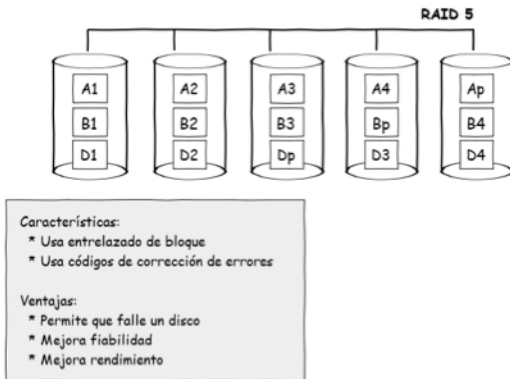
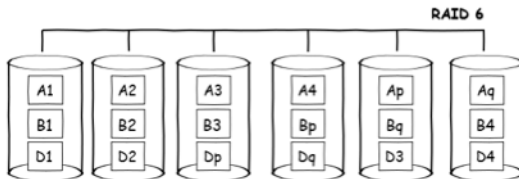


Figure 15:

El esquema de organización RAID nivel 6, también denominado esquema de redundancia $P + Q$, es bastante similar a RAID nivel 5, pero almacena información redundante adicional para proteger los datos frente a múltiples fallos de disco.

- En lugar de emplearse información de paridad, se utilizan códigos de corrección de errores, tales como los códigos Reed-Solomon.

RAID Nivel 6



Características:

- * Usa entrelazado de bloque
- * Usa códigos dobles de corrección de errores

Ventajas:

- * Permite que fallen hasta dos discos
- * Mejora fiabilidad
- * Mejora rendimiento

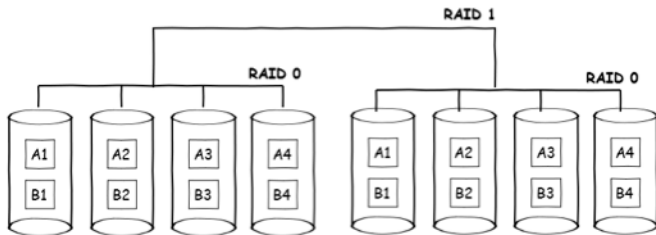
Figure 16:

RAID nivel 0 + 1.

El esquema de organización RAID nivel 0 + 1 hace referencia a una combinación de los niveles RAID 0 y 1. RAID 0 proporciona la mejora en las prestaciones, mientras que RAID 1 proporciona la fiabilidad.

- Generalmente, este nivel proporciona un mejor rendimiento que RAID 5. Resulta común en aquellos entornos en los que son importantes tanto las prestaciones como la fiabilidad.
- Desafortunadamente, este esquema dobla el número de discos necesarios para el almacenamiento, al igual que RAID 1, por lo que también es más caro.
- En RAID 0 + 1, se distribuye en bandas un conjunto de discos y luego ese conjunto se duplica en espejo en otro conjunto equivalente.

RAID nivel 0 + 1.



Características:

- * Combina RAID nivel 0 y RAID nivel 1
- * Duplica el espacio necesario

Ventajas:

- * Aumenta rendimiento
- * Mejora fiabilidad

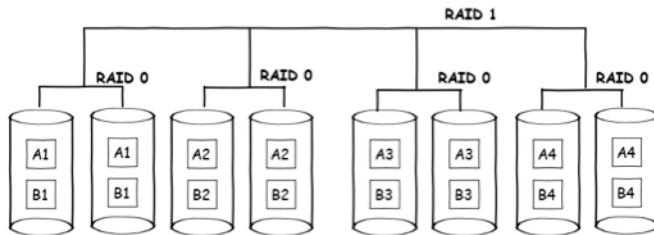
Figure 17:

RAID nivel 0 + 1.

Otra opción RAID que también está disponible comercialmente es RAID nivel 1 + 0, en el que los discos se duplican en parejas y luego se distribuyen en-bandas los datos entre las parejas de discos resultantes.

- Esta organización RAID tiene algunas ventajas teóricas sobre RAID 0 + 1.
- Por ejemplo, si falla un único disco en RAID 0 + 1, la banda completa deja de estar accesible, siendo accesible únicamente la otra banda.
- Con un fallo en RAID 1 + 0, ese único disco dejará de estar disponible, pero su pareja duplicada en espejo seguirá estando operativa, al igual que el resto de los discos.

RAID nivel 0 + 1.



Características:

- * Combina RAID nivel 1 y RAID nivel 0
- * Duplica el espacio necesario

Ventajas:

- * Aumenta rendimiento
- * Mejora fiabilidad

Figure 18:

A.SILBERSCHATZ, P. GALVIN, y G. GAGNE, Operating Systems Concepts, Cap.13, 9a Edición, John Wiley, 2013.