

Tema 4: Dispositivos de almacenamiento

1. Tipos de dispositivos de almacenamiento

Los dispositivos de almacenamiento se basan principalmente en tres tecnologías: magnética, estado sólido y óptica. En las siguientes secciones se estudian por separado.

1.1 Dispositivos magnéticos

Los dispositivos magnéticos almacenan la información modificando las propiedades magnéticas del material del que están hechos. Los principales dispositivos de este tipo son:

- **Disquetes:** En la actualidad el único formato habitual de estos dispositivos es el de 3.5" con capacidad de 1.44 MB.
- **Discos duros:** Son con mucho el tipo de dispositivo más importante y serán analizados con más profundidad en las siguientes secciones.
- **Discos ZIP:** Son un tipo de dispositivo que intentan ser una alternativa a los disquetes. Tienen mayor capacidad (hasta 750 MB). Su mayor inconveniente es que son un desarrollo propietario de la casa Iomega y eso impidió su uso estándar antes de que apareciesen otras tecnologías (los CD-R y los dispositivos de estado sólido) capaces de reemplazarlos.
- **Discos Jaz:** Son otro dispositivo desarrollado por Iomega. Tienen una capacidad de hasta 2 GB.
- **Unidades de cinta:** Son dispositivos de acceso secuencial, en lugar de aleatorio como el resto de los vistos hasta ahora. Son dispositivos lentos y se utilizan sobre todo para realizar copias de seguridad. La capacidad de las cintas está en el orden los de GB, sobre los 20 GB y hasta los 80 comprimidos. Hay muchos formatos distintos e incompatibles entre sí.

1.2 Dispositivos de estado sólido

Se basan en dispositivos semiconductores. En la actualidad la tecnología de este tipo que más se emplea es la Flash ROM, con capacidades de unos pocos GB. Con respecto a los discos duros, tienen la ventaja de no poseer partes y móviles, lo que incrementa la fiabilidad, y de poder realizarse en tamaños más pequeños. Es por esta razón por la que son el principal sistema de almacenamiento en dispositivos móviles. Como desventajas con respecto a los discos duros, la capacidad máxima es inferior; las velocidades, especialmente en escritura secuencial, pueden ser inferiores; tienen un número de escrituras limitado; y el coste por bit es superior.

También están apareciendo discos de estado sólido (*Solid State Disks*, SDDs) basados en memoria DRAM. Utilizan unas pequeñas baterías para no perder la alimentación cuando el sistema se apaga. Pueden tener también un disco duro magnético de respaldo al que copian la información cuando el equipo se apaga. Su gran ventaja es la velocidad. Como inconveniente, el coste por bit es muy

superior al de tecnologías tradicionales y no existen en tamaños tan grandes como un disco duro magnético.

1.3 Dispositivos ópticos

Los dispositivos ópticos utilizan propiedades ópticas de los materiales que están hechos para almacenar y leer la información. Los principales son:

- Basados en CD:
 - CD-ROM: Discos de sólo lectura con capacidad de 650 MB o 700 MB.
 - CD-R: Discos de lectura y escritura sólo una vez.
 - CD-RW: Discos de lectura y escritura. Hay que hacer una operación de borrado antes de volver a escribir.
- Basados en DVD: Existen básicamente dos líneas de dispositivos DVD grabables, los DVD+ y los DVD-, promovidos por distintas compañías. En el momento de su introducción esto causó muchos problemas porque los formatos no eran compatibles entre sí. Con el tiempo, todos los dispositivos se fueron haciendo compatibles con ambos formatos (los llamados dispositivos «duales») y se evitó una fragmentación del mercado.
 - DVD-ROM: Discos de sólo lectura. Los DVDs pueden tener varias capas y una o varias caras. Un DVD de una capa y una cara puede almacenar 4.7 GB (4.38 GiB). Hay DVDs hasta de doble capa y doble cara, con aproximadamente 18 GB de capacidad.
 - DVD-RAM: Discos pensados sólo para datos (no vídeo) y que no tienen gran uso fuera de ámbitos profesionales. Fue el primer formato de DVD grabable en aparecer.
 - DVD-R: Formato de DVD grabable sólo una vez introducido por Pioneer en 1998.
 - DVD+R: Es la alternativa de un consorcio llamado DVD Alliance al DVD-R.
 - DVD-RW: Es un formato de discos regrabables impulsado por Pioneer y el DVD Forum.
 - DVD+RW: Es una alternativa al DVD-RW.
- De alta densidad: Se propusieron dos alternativas de mucha mayor capacidad al formato DVD: Blu-Ray y HD-DVD. El primero almacena más información por capa que el segundo (25 GB frente a 15 GB). A principios de 2008 HD-DVD perdió los apoyos que tenía, quedando como estándar Blu-Ray.

Una de las características principales de estos dispositivos es la velocidad de lectura o de escritura. Se suele expresar como un multiplicador, por ejemplo 4x. El multiplicador base es distinto para los CDs y los DVDs. Un CD 1x tiene una tasa de transferencia de 150 KB/s, mientras que un DVD de 1x tiene una tasa de transferencia de 1.321 MB/s. Las velocidades que se dan se refieren

habitualmente a las que se obtienen en la parte externa del disco, que gira a más velocidad, y por lo tanto son las máximas que se pueden obtener.

1.4 Dispositivos híbridos

En 2007 Samsung introdujo en el mercado el concepto de disco híbrido (*Hybrid Hard Driver*, HHD), que consiste en un disco duro tradicional basado en tecnología tradicional combinado con un gran buffer de memoria Flash. Están destinados a los portátiles porque (en teoría) tienen un consumo menor que un disco tradicional.

2. Interfaces

2.1 IDE

2.1.1 Evolución

El interfaz IDE ha sido el más común en los PCs. Hasta el 2005 aproximadamente, fecha en la que empezó a ser habitual el interfaz SATA, casi todas las placas base han traído integradas en la placa dos ranuras adaptadoras para conectores IDE que conectan con el bus PCI. IDE es la abreviatura de *Integrated Drive Electronics* (electrónica controladora integrada), lo que quiere decir que la controladora del disco duro está integrada en el propio disco¹. La ventaja principal de los discos IDE es su bajo coste.

Este interfaz ha ido evolucionando con el tiempo. El nombre oficial del estándar correspondiente es ATA (*AT Attachment*, conexión a AT, porque en principio era un estándar para los IBM AT). Las nuevas versiones incorporan nuevas velocidades. La más alta alcanzada por esta interfaz es de 133 MB/s. Las velocidades inferiores utilizaban modos PIO (*Programmed Input/Output*) mientras que las modernas utilizan modos UltraDMA (*Direct Access Memory*).

El conector se ha mantenido estable con 40 *pins*, pero los cables han cambiado. A partir del estándar ATA-5 (también llamado ATA/66) los cables tienen 80 conductores: los 40 que no se conectan a ningún *pin* sirven de aislamiento. Las distintas versiones del estándar son compatibles hacia atrás.

ATAPI (*ATA Packet Interface*) es un estándar que se desarrolló para conectar CD-ROMs y unidades ZIP a través del interfaz IDE. Es más complicado que el ATA *normal* y se parece más al SCSI. El ATAPI es parte del ATA y se empezó a incluir en ATA-4.

¹ Interfaz IDE es el nombre más usado, pero es incorrecto en el sentido de que en la actualidad los discos SCSI o USB también tienen la controladora integrada y, por lo tanto, son IDE. Tras la aparición del Serial ATA (SATA), se empezó a llamar también Parallel ATA o PATA.

	ATA-1	ATA-2	ATA-3	ATA-4	ATA-5	ATA-6	ATA-7
Modos UltraDMA	-	-	-	0,1,2	3,4	5	6
Tasas máximas de transferencia	11.1 MB/s	16.6 MB/s	16.6 MB/s	33.3 MB/s	66.6 MB/s	100 MB/s	133 MB/s
Nº de cables	40	40	40	40/80	80	80	80
Año del estándar ANSI	1994	1996	1997	1998	2000	2001	2005
Características nuevas	-	Transferencias en bloque, LBA, identificación de disco	SMART, características de fiabilidad	CRC, cable de 80 conductores	-	LBA de 48 bit	SATA-1
Conocido también como...	ATA/IDE	ATA/IDE	ATA/IDE	UltraDMA/33	UltraDMA/66, ATA/66	UltraDMA/100, ATA/100	UltraDMA /133, ATA133

2.1.2 Configuración

El estándar ATA permite conectar hasta dos discos en el mismo cable. Esto tiene sus problemas ya que los dos discos tienen una controladora propia. Para seleccionar qué comandos van a cada disco, debe configurarse un disco como maestro (el 0) y el otro como esclavo (el 1). Esto se suele hacer a través de *jumpers*. Otra posibilidad es dejar que la posición en el cable determine qué dispositivo es maestro y qué dispositivo es esclavo; a esta opción se la suele denominar *cable select* y deben seleccionarla los dos dispositivos conectados a un cable para que funcione.

En el cable usado a partir del ATA-5 los conectores del cable están identificados por un código de colores:

- Azul: Conector para el interfaz en la placa base.
- Negro: Conector para el disco duro maestro. Si sólo hay una unidad, debe ir obligatoriamente en este conector.
- Gris: Conector para el disco duro esclavo.

Tradicionalmente, los cables IDE han sido planos, con todos los conductores puestos unos al lado de los otros. En los equipos modernos, donde la refrigeración es un factor importante, esto causa problemas porque la superficie del cable corta el flujo de aire. Por esta razón también existen cables redondeados.

La combinación de limitaciones en la BIOS y diferencias en el tamaño de los campos para cilindros, cabezas y sectores entre la BIOS y el estándar ATA provocan problemas que se explican más adelante en el apartado de geometría de los discos duros.

La relación entre la configuración de los dos dispositivos conectados a un cable y la velocidad es la siguiente:

- Con dispositivos actuales, si se conectan dos de distinta velocidad (por ejemplo, un DVD UltraDMA/33 y un disco duro UltraDMA/100), no hay penalización cuando se accede al más rápido (si se utiliza el cable adecuado), ya que las controladoras actuales se comunican con cada dispositivo a la velocidad más rápida que este acepta. Por supuesto, cuando se accede al más lento, el más rápido debe esperar. Lo que no se puede hacer es mezclar un

dispositivo en modo PIO con otro en modo UltraDMA porque las controladoras no pueden cambiar entre modos.

- Con respecto a la importancia de escoger qué dispositivo es maestro y cuál es esclavo, la única diferencia es que la controladora le pregunta antes al maestro que al esclavo, así que el maestro puede tener pequeñas ventajas en la latencia.

2.2 Serial ATA

El estándar Serial ATA, aparecido a mediados de la primera década del siglo XXI, es el sustituto del ATA. La diferencia principal entre ATA y Serial ATA es que el segundo manda los datos en serie. La razón, aunque no parezca lógico en principio, es alcanzar mayores velocidades. En una comunicación en paralelo se manda más bits en cada ciclo, pero es muy difícil incrementar la frecuencia del reloj porque hay que sincronizar un conjunto de señales en paralelo. En una comunicación serie es mucho más fácil incrementar la frecuencia y conseguir tasas de transferencia mayores.

El estándar Serial ATA utiliza cables y conectores mucho más estrechos que el ATA. Además, permite tener cables de longitudes mayores y realizar conexiones en caliente. La tasa máxima de transferencia que permite su primera versión es 150 MB/s. La segunda versión puede llegar hasta 300 MB/s.

2.3 SCSI

2.3.1 Evolución

SCSI es la abreviatura de *Small Computer System Interface*. Es un interfaz para conectar no sólo discos duros sino también otros dispositivos como escáners, CD-ROMs, cintas, etc. Es el interfaz más habitual para discos duros de altas prestaciones en servidores. SCSI es un bus puede tener 7 ó 15 dispositivos como máximo.

Para la interconexión entre el bus SCSI y el bus del sistema (PCI) son necesarias unas controladoras de SCSI, llamadas *host adapters*. Las placas bases para servidores las tienen integradas, pero también se pueden comprar como tarjetas de expansión que se conectan al bus PCI. La controladora SCSI no se comunica directamente con el disco, sino que cada disco tiene una controladora propia integrada (por lo tanto, hablando con propiedad, los discos SCSI también son IDE).

En un bus SCSI puede haber 8 ó 16 dispositivos, cada uno con su ID. Uno de estos dispositivos es la controladora SCSI, por lo tanto sólo quedan 7 ó 15 para otros dispositivos.

El estándar SCSI define la parte hardware, no la software, así que hay problemas de *drivers*: habitualmente los *drivers* sirven sólo para un tipo concreto de controladora SCSI y de disco. Las controladoras SCSI tienen habitualmente su propia ROM BIOS.

La siguiente tabla resume las características principales de los distintos estándares SCSI que no están obsoletos.

Estándar SCSI	Tecnología SCSI	Nombre de marketing	Reloj (MHz)	Ancho de la transferencia	Velocidad (MB/s)	Núm. máx. dispositivos
SCSI-2	Fast-5/Wide	Wide	5	16 bits	10	15
SCSI-2	Fast-10	Fast	10	8 bits	10	7
SCSI-2	Fast-10/Wide	Fast/Wide	10	16 bits	20	15
SPI (SCSI-3)	Fast-20	Ultra	20	8 bits	20	7
SPI (SCSI-3)	Fast-20/Wide	Ultra/Wide	20	16 bits	40	7
SPI-2 (SCSI-3)	Fast-40	Ultra2	40	8 bits	40	7
SPI-2 (SCSI-3)	Fast-40/Wide	Ultra2/Wide	40	16 bits	80	15
SPI-3 (SCSI-3)	Fast-80DT	Ultra3 (Ultra160)	40	16 bits	160	15
SPI-4	Fast-160DT	Ultra4	80	16 bits	320	15
SPI-5	Fast-320DT	Ultra5			640	

El estándar SCSI es compatible hacia adelante y hacia atrás: un dispositivo SCSI-2 puede funcionar en un bus SCSI-3 (aunque no aprovechará las características nuevas) y un dispositivo SCSI-3 puede funcionar en un bus SCSI-2.

2.3.2 Cables y conexión

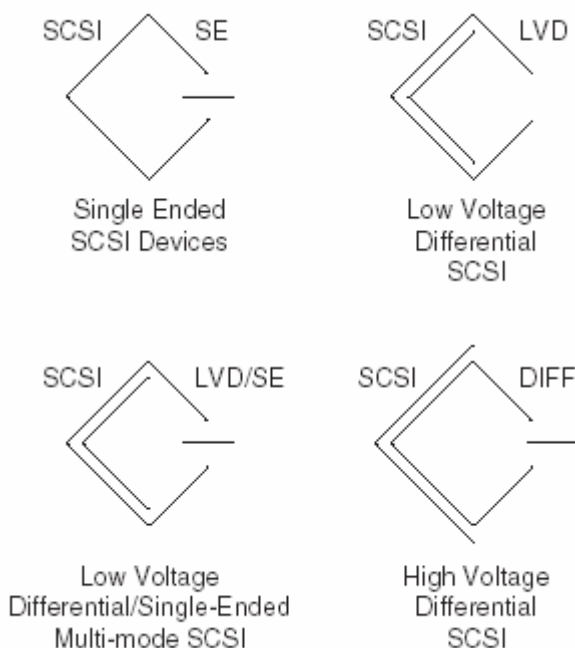
Hay dos tipos de cables SCSI:

- El tipo A: tiene 50 conectores.
- El tipo B: tiene 68 conectores. Es necesario para utilizar los estándares de 16 bits.

Hay distintos modos de transmitir la señal (*signaling*) en SCSI:

- Original: Sin balancear (*unbalanced* o *single-ended*, SE). Máxima velocidad: 20 MB/s.
- HVD (*High Voltage Differential*): Balanceado. Permite cables más largos. No ha sido muy usado porque la construcción de dispositivos era más cara y el formato original era suficiente para bajas velocidades. Se ha eliminado en SCSI-3. Si se conecta un dispositivo HVD a otro SE, éste puede quedar dañado porque HVD utiliza tensiones mucho más altas.
- LVD (*Low Voltage Differential*): Es similar a HVD pero con tensiones inferiores. Esto hace que no se dañe un dispositivo SE si se conecta a otro LVD. El estándar permite versiones mixtas LVD/SE. El LVD permite velocidades mayores que el SE y cables más largos. En un cable en el que haya conectados dispositivos LVD/SE y SE se trabajará con SE, es decir, a bajas velocidades.

Los conectores de los cables son iguales excepto en los símbolos que ponen. La siguiente tabla muestra estos símbolos y a qué se corresponden:



2.3.3 Configuración

La configuración de un disco SCSI tiene dos partes:

- **Fijar el ID SCSI:** Cada dispositivo tiene un ID, entre 0 y 7 para SCSI estrecho y entre 0 y 15 para SCSI ancho. La controladora SCSI suele coger el último valor del rango. Las controladoras antiguas sólo permiten que el disco de arranque sea el que tenga el ID 0, pero en las modernas puede serlo cualquiera. El ID se suele escoger mediante *jumpers* que lo codifican en binario.
- **Fijar los terminadores:** Un bus SCSI necesita un terminador en cada extremo. Si la controladora o un dispositivo está en un extremo, se debe activar la terminación del elemento. Si no tiene esta posibilidad, hay que conectar un terminador externo al conector SCSI de salida del elemento. Hay dispositivos que no tienen un conector SCSI de entrada y otro de salida. En estos dispositivos hay que utilizar un terminador especial que se coloca entre el cable SCSI y el conector del dispositivo (*pass-through terminator*).

Otros posibles aspectos a configurar son:

- **Inicio retrasado (*Delayed Start* o *Start on Command*):** Al empezar a girar, un disco duro puede consumir hasta 4 veces la potencia habitual. Si hay varios discos duros en el mismo bus esto puede causar problemas. Si se configura el inicio retrasado, el disco no empieza a girar hasta que recibe la orden de la controladora SCSI. Esta función no es necesaria si los discos son externos y tienen su alimentación propia.

- Paridad SCSI: Permite utilizar paridad para corregir errores. Sólo algunas controladoras muy antiguas no la soportan, así que lo mejor es tenerla activada.
- Alimentación del terminador (*Terminator Power*): Indica si se debe alimentar al terminador. Habitualmente lo hace la controladora SCSI. No hay ningún problema si varios dispositivos tienen esta opción seleccionada.
- Negociación síncrona SCSI: El bus SCSI puede funcionar de manera síncrona o asíncrona. Lo normal es tener esta opción habilitada para que los dispositivos negocien cómo hacerlo según lo que soporten. El problema es que algunos dispositivos antiguos no soportan la negociación. Sólo en este caso se debe deshabilitar esta opción.

2.3.4 SCSI Plug-and-Play

Un nuevo estándar es SCSI Plug-and-Play (PnP). Las características más importantes son:

- Configuración con un solo tipo conector.
- Terminación automática del bus SCSI.
- SCAM (*SCSI Configured Automatically*): Asigna automáticamente los IDs.
- Compatibilidad hacia atrás de los dispositivos PnP con los dispositivos SCSI existentes.

2.4 Otros interfaces

Aunque ATA y SCSI son los estándares que dominan el mercado en cuanto a la conexión de discos duros internos, en el apartado de discos duros externos se suelen utilizar otros interfaces como el puerto paralelo, USB (1.0 y 2.0) y FireWire. Con el puerto paralelo y con USB 1.0 se tiene una tasa de transferencia demasiado baja (hasta 1.5 MB/s con USB 1.0) para un trabajo habitual, pero con USB 2.0 (60 MB/s) y FireWire 400 (50 MB/s) y FireWire 800 (100 MB/s) se pueden tener tasas de transferencia adecuadas, teniendo en cuenta que en la actualidad las tasas de transferencias máximas de los discos duros son del orden de 70-80 MB/s.

Para mayores tasas de transferencia se utiliza el estándar Fibre Channel, que habitualmente se utiliza para conectar con cabinas de RAID o hacer redes de almacenamiento (*Storage Area Network*, SAN). Fibre Channel puede conseguir velocidades de hasta 4 Gbits/s, conectando miles de dispositivos en distancias de hasta 10 km. Admite gran cantidad de protocolos sobre la capa física, siendo SCSI el más habitual.

3. Discos duros

3.1 Geometría

Los discos duros están formados por varios platos que giran solidariamente alrededor de un eje. En ellos se guarda la información por medios magnéticos. Unas cabezas lectoras (una por cada superficie de cada plato que esté utilizada para datos) se encargan de leer y escribir la información. Para que esta información sea accesible debe tener un orden. Las superficies se dividen en pistas concéntricas. Al conjunto de todas las pistas concéntricas a la misma distancia del eje se le denomina cilindro. Las pistas están divididas en sectores. Todos los sectores tienen el mismo número de bytes: 512 bytes.

Para hacer referencia a un sector concreto dentro del disco los programas, las BIOS y los estándares antiguos utilizan un sistema conocido como CHS por las iniciales de cilindro (*cylinder*), cabeza (*head*) y sector (*sector*). Hay un número de bits limitado para cada uno de estos campos, lo que pone un límite máximo al tamaño de disco que se puede manejar. Esto se ve agravado por que el número de bits no es igual en la BIOS y en el estándar ATA. La siguiente tabla resume los datos fundamentales:

Parámetro	Límite en la BIOS	Límite en ATA	Límite combinado
Cilindros	1024 (10 bits)	65536 (16 bits)	1024
Cabezas	256 (8 bits)	16 (4 bits)	16
Sectores	63 (6 bits)	255 (8 bits)	63
Nº total de sectores	16 515 072	267 386 880	1 032 192
Tamaño máximo	8.4 GB	139.9 GB	528 MB

El número de sectores en la BIOS es 63 y no 64 porque los sectores se empiezan a numerar en 1. Por la misma razón, el número de sectores posibles en ATA es 255 y no 256.

Como se puede ver, el máximo valor de CHS posible combinado es 1024/16/63, lo que da un total de 528 MB, que se queda muy lejos del tamaño de los discos actuales. Para solucionar este problema, las BIOS comenzaron a integrar, a partir de 1994, dos sistemas de traducción (*translation*) en el servicio INT 13h que permite acceder al disco:

- *ECHS* (*Extended CHS*, a veces también llamado *Large*): Modifica el número de cilindros y cabezas para que quepan en los campos correspondientes pero el producto dé lo mismo.
- *LBA* (*Logical Block Address*, dirección lógica de bloque): Utiliza un número absoluto para hacer referencia a cada sector, empezando por el cilindro 0, cabeza 0, sector 1 y asignando números consecutivamente hasta el último sector físico del disco. Se soporta a partir de ATA-2 y es la forma habitual de referenciar en discos duros SCSI.

Es mejor utilizar LBA porque es más eficiente. Lo que no se debe hacer nunca es cambiar de uno a otro una vez que se ha instalado un disco porque se pueden perder los datos.

Los sistemas de traducción transforman los parámetros CHS físicos en otros lógicos que quepan dentro de los límites. Por ejemplo, si un disco tiene 12000 cilindros (más de los 1024 soportados por el servicio 13h de la BIOS), 16 cabezas y 63 sectores, la BIOS dice que tiene 750 cilindros ($12000/16$), 256 cabezas (16×16) y 63 sectores. A los primeros parámetros se les llama físicos (aunque, de hecho, tampoco son los datos físicos reales porque la controladora del disco hace otra traducción interna) y a los segundos lógicos. De esta forma se puede llegar hasta los 8.4 GB.

La siguiente tabla resume todos los tipos de traducciones (excepto la interna de la controladora al disco porque es totalmente transparente) que se pueden hacer y las limitaciones a las que están sometidas:

Modo de la BIOS	OS a BIOS	BIOS a disco	Límite
Estándar o normal	Parámetros CHS físicos	Parámetros CHS físicos	528 MB
ECHS o <i>Large</i>	Parámetros CHS lógicos	Parámetros CHS físicos	8.4 GB
LBA	Parámetros CHS lógicos	Parámetros LBA	8.4 GB
LBA puro (EDD BIOS)	Parámetros LBA	Parámetros LBA	144 PB

Tanto en modo ECHS como en LBA se utiliza un número para comunicarse entre la BIOS y el disco y otro número para comunicarse entre la BIOS y los programas que la llaman. De esta manera se rompe el límite de los 528 MB y se pasa al de los 8.4 GB debido al tamaño de los campos CHS en el servicio 13h de la BIOS. Los discos actuales son mayores que este valor, así que fue necesario un nuevo método para acceder a discos de más de 8.4 GB.

La solución fue un nuevo estándar para la BIOS denominado *BIOS Enhanced Disk Drive Specification (EDD)* que se aprobó en 1998 y permite llegar hasta 2^{64} sectores, lo que da un total de 144 PB (Petabytes). Al ritmo de crecimiento actual de la capacidad de los discos (se dobla cada año y medio o dos años) será suficiente hasta el año 2057 por lo menos. Pero hay que tener en cuenta que el hasta el estándar ATA-5 todavía se tiene el límite de los 139.9 GBs (porque aunque ya no recibe los datos en formato CHS sigue recibiendo el mismo número de bits, 28). El estándar ATA-6 (también conocido como ATA/100) ya utiliza 48 bits, con lo que se pueden direccionar 2^{48} sectores. El estándar EDD hace que los programas se comuniquen con la BIOS directamente con los datos LBA y esta con el disco también con LBA, así que el límite será el número de bits que intercambien la BIOS y el disco.

Otro lugar donde el utilizar el sistema CHS causa problemas es en el proceso de arranque. El cargador de DOS y el de Windows NT 4.0 utilizan el sistema CHS para buscar la partición desde la que deben arrancar y, por el número de bits de los campos que utilizan, si la partición empieza más allá de los 2 GB, no podrán arrancar el sistema.

3.2 SMART

En cualquier sistema hay dos tipos de fallos: predecibles e impredecibles. Los primeros se pueden predecir midiendo ciertas características. Por ejemplo la temperatura del motor de un disco duro puede dar una indicación de si fallará en el futuro. El disco informaría de este hecho y el usuario todavía tendrá tiempo a salvar sus datos.

Varios fabricantes han desarrollado un estándar para que el disco duro indique este tipo de errores. El estándar se denomina S.M.A.R.T. (*Self-Monitoring, Analysis and Reporting Technology*, Tecnología de auto-monitorización, análisis e informes).

Las monitorizaciones más habituales que se hacen en un disco duro son:

- Altura de vuelo de las cabezas lectoras.
- Rendimiento de tasa de transferencia de datos.
- Tiempo para comenzar a girar (*spin-up time*).
- Cuenta de sectores resituados (*reallocated sector count*) o libres (*spare*). Los discos duros actuales tienen sectores extra libres que en principio no se utilizan. Cuando un sector *normal* falla, se deja de utilizar y se sustituye por uno de estos sectores libres. La cuenta de sectores que han fallado (o la de los que quedan libres si se sabe la cantidad inicial) da una idea del número de fallos que ha habido en los sectores del disco.
- Tasa de errores de búsqueda.
- Rendimiento del tiempo de búsqueda.
- Cuenta de los reintentos de comienzo de giro (*drive spin-up retry count*).
- Cuenta de los reintentos de calibración del disco.

3.3 RAID

3.3.1 Introducción

Mientras las prestaciones de los procesadores han crecido a un ritmo de un 50% al año, los tiempos de acceso de los discos, que dependen de sistemas mecánicos, han mejorado menos de un 10% al año. Los tiempos de transferencia, que siguen las mejoras de los sistemas mecánicos y la densidad de los medios magnéticos, han mejorado aproximadamente un 20% al año, todavía lejos de los procesadores. Esto hace que la diferencia entre la velocidad de los discos y los del procesador se agrande con el tiempo.

Una solución a este problema es utilizar matrices de discos (*disk arrays*). Una matriz de discos es un conjunto de varios discos independientes organizados para que el sistema operativo los vea como un solo disco lógico. En 1988 Patterson, Gibson y Katz, un grupo de investigadores de Berkeley, publicaron un artículo que definió la nomenclatura de las principales organizaciones de matrices de discos y es la referencia básica en este campo. En este artículo, los autores denominaban a las matrices de discos Redundant Arrays of Inexpensive

Disks (RAID) para dar idea de que con una matriz de discos económicos se puede conseguir tanto rendimiento (o más) que con un solo disco grande y costoso. Sin embargo, los beneficios del RAID van más allá de simple uso de discos económicos y se ha extendido a todo tipo de discos, por lo que se ha cambiado el significado del término RAID por Redundant Array of **Independent** Disks.

Los sistemas RAID mejoran la velocidad gracias a una técnica denominada *striping*, que consiste en dividir los datos en bandas (*stripes*) que se distribuyen entre los discos que forman la matriz. Si en un momento dado se requieren datos que estén en dos discos distintos, se podrán traer en paralelo, lo que mejorará el tiempo de respuesta.

Esta solución trae sin embargo un nuevo problema: en una matriz sin redundancia, el fallo de un disco significa el fallo de todo el sistema de discos. Por lo tanto, cuantos más discos se tengan más probabilidad hay de que falle el sistema de discos. La solución pasa por utilizar información redundante, de tal forma que aunque falle un disco el sistema pueda seguir funcionando.

Añadir información redundante tiene efectos sobre el rendimiento. Cuando se expliquen las distintas organizaciones de RAID se verán estos efectos. Por otra parte, la información redundante también tiene un coste: los bits que se utilicen para almacenar información redundante no se utilizan para almacenar información útil (datos, desde el punto de vista del usuario). Se denomina sobrecarga (*overhead*) al cociente nº de bytes redundantes / nº de bytes de datos.

3.3.2 Niveles RAID

En el artículo se definían 5 formas distintas de organizar las matrices de disco que se denominaron “niveles RAID”, del 1 al 5. Hay que tener en cuenta que a pesar de que el uso de la palabra “nivel” parece implicar una jerarquía, esto no es así; es decir, el nivel 5 no es mejor que el 1: tiene ventajas en algunos aspectos y desventajas en otros. Por lo tanto, la elección de un nivel u otro involucrará un **compromiso** entre las ventajas y las desventajas, teniendo que elegir, por ejemplo, entre mayor fiabilidad y menor rendimiento. En la práctica, sólo se usan habitualmente tres niveles: 0, 1 y 5. Se describen a continuación.

3.3.2.1 RAID 0: No redundante

Este tipo no es estrictamente RAID, ya que no incorpora información redundante, sino solamente *striping*: simplemente distribuye los datos en bandas a lo largo de varios discos (ver Figura 1).

Como ventajas, es el sistema más barato (ya que tiene sobrecarga cero) y tiene un rendimiento muy elevado, al no necesitar ni leer ni escribir la información redundante en cada petición de E/S. La gran desventaja de este sistema es su baja fiabilidad:

$$MTTF_{array} = \frac{MTTF_{disco}}{N}$$

Siendo N el número de discos del array y $MTTF$ el *Mean Time To Failure* (tiempo medio hasta el fallo). Por ejemplo, si se tiene un $MTTF_{\text{disco}}$ de 10 000 horas (es decir, 1.14 años), en un array de 6 discos el $MTTF_{\text{array}}$ será de sólo 1 666.67 horas (0.19 años).

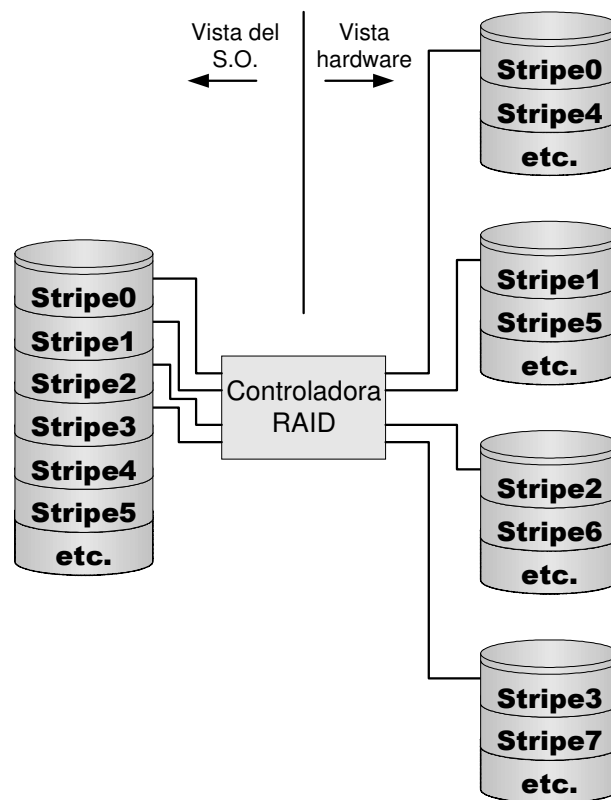


Figura 1. RAID 0

3.3.2.2 RAID 1: En espejo

En esta organización, cada vez que se escribe un bloque en un disco de datos, también se escribe en un disco redundante, con lo que se tienen dos copias de toda la información (ver Figura 2). Por lo tanto, la sobrecarga es del 100%, la más alta de todas las configuraciones. Como ventaja, es muy fácil reconstruir el array si falla un disco: simplemente hay que copiar la información de su disco espejo.

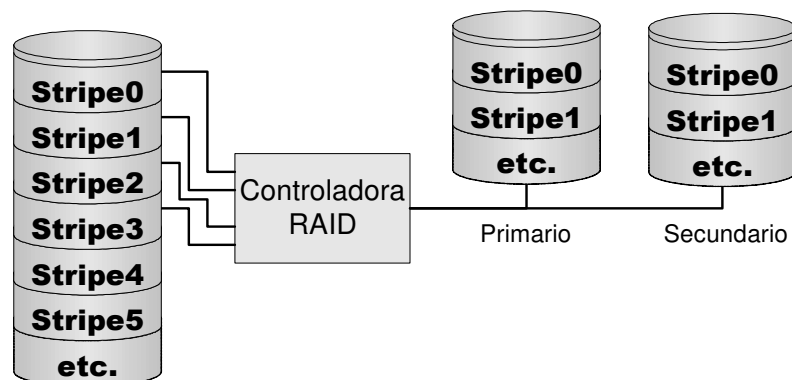


Figura 2. RAID 1

La velocidad de RAID 1 en lectura puede ser mayor que la de RAID 0 ya que se puede leer la información de dos discos a la vez y escoger el que menos tarde o se pueden dividir peticiones que vayan al mismo disco entre él y su disco espejo. En escrituras el tiempo será un poco mayor porque hay que hacer dos escrituras, que aunque vayan en paralelo requerirán el doble de tiempo en la controladora (suponiendo que sea única). Sin embargo, este tiempo es despreciable comparado con el de posicionamiento y transferencia.

El MTTF es:

$$MTTF_{array} = \frac{MTTF_{disco}^2}{2N * MTTR_{disco}}$$

con N el número de discos (sin incluir los de espejo). El MTTR es el tiempo medio de reparación. Por ejemplo, si se tiene, como antes, un $MTTF_{disco}$ de 10 000 horas, un MTTR de 5 horas y 6 discos, tendremos un $N=3$ y un $MTTF_{array}$ de 3 333 333.33 horas (380.52 años).

3.3.2.3 RAID 5: Paridad entrelazada a nivel de bloque y distribuida

En este sistema se crean tiras de paridad que tienen los datos de la paridad de las tiras de datos (ver Figura 3). Las tiras de paridad se reparten entre varios discos.

Este sistema es una solución intermedia entre RAID-1 y RAID-5: Tiene menos sobrecarga que RAID-1 y más fiabilidad que RAID-0. El MTTF es:

$$MTTF_{array} = \frac{MTTF_{disco}^2}{N * (N + 1) * MTTR_{disco}}$$

Con N el número de discos de datos. Por ejemplo, si se tiene, como antes, un $MTTF_{disco}$ de 10 000 horas, un MTTR de 5 horas y 6 discos (5 con datos y 1 con información de redundancia, aunque en realidad los 6 discos tienen datos e información de redundancia), tendremos un $MTTF_{array}$ de 666 666.67 horas (76.1 años).

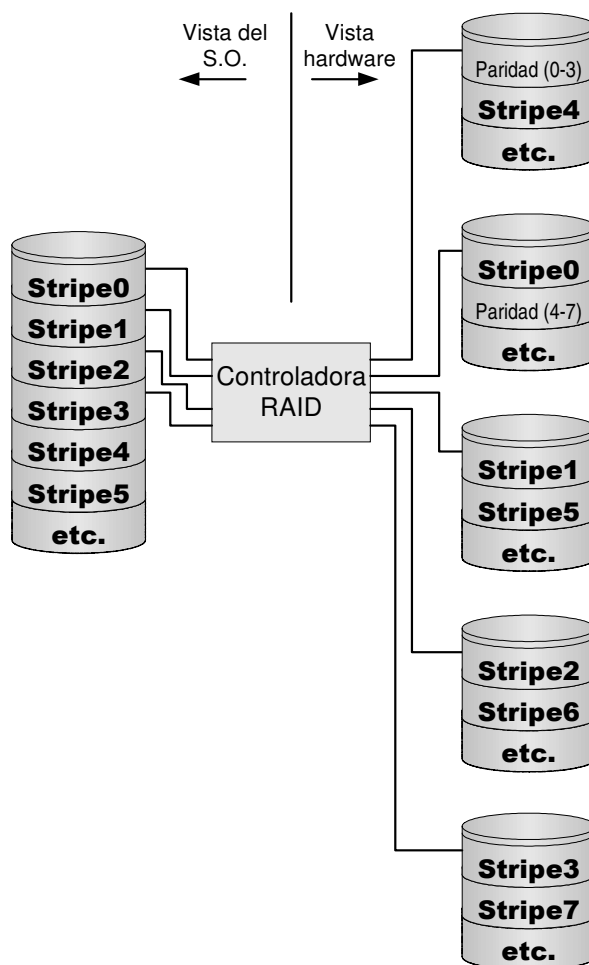


Figura 3. RAID 5

3.3.2.4 RAID de varios niveles

Se han desarrollado soluciones que utilizan una combinación de varios niveles de RAID, intentado combinar las ventajas de cada uno de los niveles empleados. Las más comunes son RAID 0+1 y RAID 1+0. A veces los productos comerciales no utilizan de manera consistente estas denominaciones. Incluso a veces se usa la denominación RAID 10.

3.4 El administrador de discos de Windows

Windows tiene una herramienta para administrar discos llamada “Administrador de discos”. Distingue dos tipos de discos:

- **Discos básicos:** Un disco básico es un disco físico que contiene particiones primarias y extendidas. Es el tipo de disco tradicional.
- **Discos dinámicos:** Un disco dinámico es uno o varios discos físicos. Los discos dinámicos no contienen particiones sino volúmenes dinámicos que se crean con el administrador de discos de Windows. La ventaja es que permiten cambiar el tamaño (extenderlos) una vez creados, utilizar RAID y cambiar la configuración sin tener que reiniciar el ordenador. Este tipo de discos no está soportado en ordenadores portátiles.

Los discos por defecto son de tipo básico. Para pasarlos a tipo dinámico se utiliza el administrador de discos. Primero hay que eliminar todas las particiones y luego se pasa a disco dinámico. Una vez que se pasa un disco de básico a dinámico, no se puede volver atrás (sin perder los datos). Los discos dinámicos no pueden ser vistos por Windows 98/Me. Además, Windows no puede ser instalado sobre un disco dinámico.

Windows divide los volúmenes en los discos dinámicos en estos tipos:

- Simple: Está en un solo disco físico. Se puede extender (aumentar de tamaño) usando espacio libre en el disco. También se puede extender usando otros discos, pasando a ser un volumen distribuido.
- Distribuido (*spanned*): Utiliza espacio de varios discos físicos (hasta 32). Se puede extender.
- Reflejado: Son dos discos físicos en RAID 1. No se puede extender.
- Seccionado (*stripped*): Son varios discos físicos en RAID 0. No se puede reflejar ni extender.
- RAID 5: Son varios discos físicos (3 como mínimo) en RAID 5.

El RAID que hace Windows es RAID software. Si se combinan discos de varios tamaños en un RAID, se escoge el menor de todos. Es decir, si se tienen dos discos de 80 GB y uno de 40 GB en un RAID 0, el tamaño total del RAID será $40 \times 3 = 120$ GB.

Windows sólo permite RAID en los operativos para servidor o en las versiones Enterprise y Ultimate de Windows Vista.

Otras acciones importantes que se pueden llevar a cabo con el *Administrador de discos* son:

- Cambiar la letra de unidad. Como se explica en el tema de instalación del sistema operativo, en Windows, al contrario que en Windows 98, las letras no se recalculan en cada arranque.
- Formatear. El formateo es a alto nivel.
- Activar particiones. Sólo en particiones primarias.
- Defragmentar. Esta operación elimina la fragmentación de ficheros en el disco duro. El administrador permite primero hacer un análisis y, según sus resultados, decidir si se quiere hacer una defragmentación o no. Microsoft recomienda defragmentar una vez a la semana. El defragmentador de Windows no defragmenta ni el MFT (*Master File Table*) ni el archivo de *swap* porque están en uso siempre. Para defragmentar estos archivos es necesario utilizar una herramienta de otra compañía que realice la defragmentación en el arranque de Windows. Para poder hacer la defragmentación total se necesita suficiente espacio para copiar contiguos los archivos grandes. Microsoft recomienda tener sobre un 30% del disco libre para que la defragmentación sea efectiva.
- Cambiar el tamaño de una partición: Esta opción apareció con Windows Vista.