

Mantenimiento de Equipos Informáticos

Tema 5: Sistemas de almacenamiento

PABLO ESCOBEDO ARAQUE

PABLOESCOBEDO@UGR.ES

DPTO. ELECTRÓNICA Y TECNOLOGÍA DE COMPUTADORES

A solid orange horizontal bar spanning the width of the slide at the bottom.

Contenido

1. Introducción
2. Discos magnéticos
3. Unidades de almacenamiento SSD
4. Mantenimiento
 - Parámetros de fiabilidad
 - Factores ambientales
 - Tecnología S.M.A.R.T.
 - RAID

Introducción

- El subsistema de almacenamiento, y específicamente el disco duro, surge de la necesidad de conservar grandes cantidades de información después de apagar el ordenador (información no volátil).
 - **Bajo coste (€/Gbyte):** Imprescindible por el creciente volumen de información que se maneja habitualmente.
 - **Tiempo de acceso:** es superior a la memoria volátil y su velocidad de transferencia menor.
 - Los sistemas de almacenamiento de alta capacidad incluyen **elementos móviles** (rozamiento, vibración, fatiga), lo que será determinante a la hora de su **fiabilidad** (i.e. mantenimiento).

Discos magnéticos

- En una memoria volátil, RAM, la unidad mínima de acceso es el **byte**. En un disco, la unidad mínima accesible es el **sector** (512 bytes).
- Un disco duro está formado por varios elementos, entre los que se encuentran varios **discos magnéticos** (platos) que sirven de soporte a los datos.
- El funcionamiento del disco magnético se fundamenta en una **superficie magnetizable** (medio de grabación) y un **cabezal** capaz de detectar y grabar variaciones magnéticas.



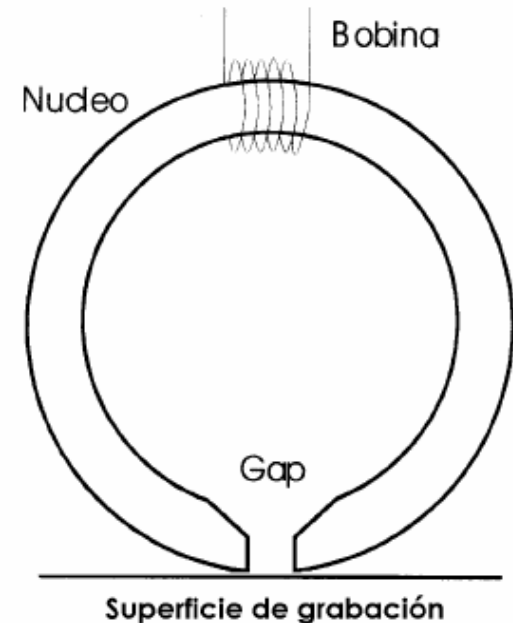
Discos magnéticos: Medio de grabación

- Sobre el plato de aluminio (sustrato) se encuentra la superficie magnetizable que contendrá los datos.
- **Sustrato:**
 - El plato suele ser de aluminio.
 - Debe tener un bajo coeficiente de dilatación, pues la fuerza centrífuga de la rotación y el calor en el interior de la unidad podrían deformarlo, dificultando la lectura y escritura.
 - La superficie debe ser extremadamente lisa.
- **Superficie magnetizable:**
 - Capa de material magnético y por tanto susceptible a los campos electromagnéticos.
 - En general se utilizan dos tipos de capas: óxido y [película delgada](#).

Discos magnéticos:

Cabezal de lectura-escritura

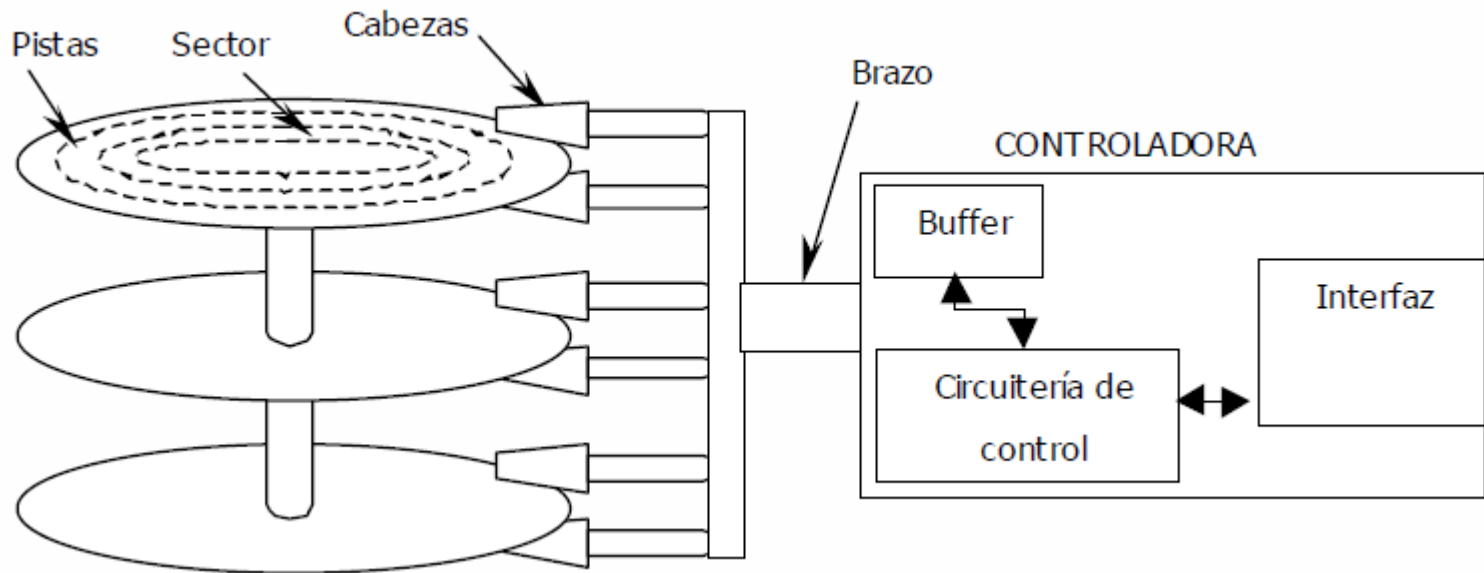
- **Escritura:** La corriente que circula por la bobina provoca un flujo magnético sobre el núcleo y por lo tanto sobre la superficie del disco.
 - Sentido de la corriente → dirección de magnetización → bit.
- **Lectura:** El flujo magnético del disco induce una corriente eléctrica sobre la bobina del cabezal.



- El núcleo ferromagnético posee una separación llamada **gap** ($< 1 \mu\text{m}$) que permite que el flujo electromagnético se propague hasta la superficie de grabación.

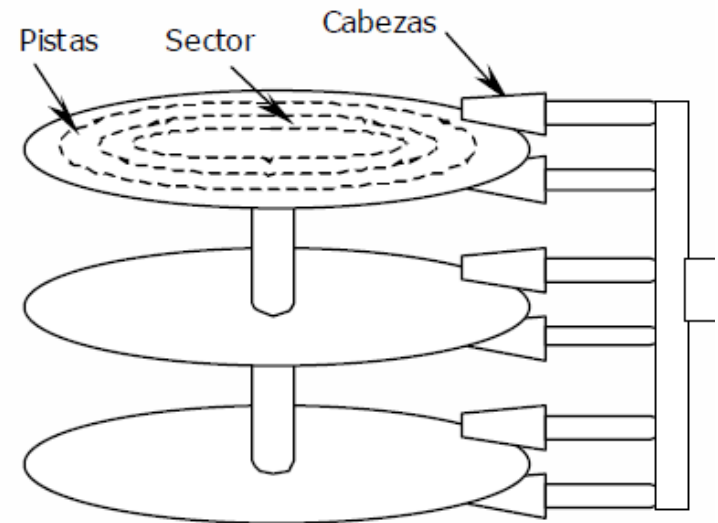
Discos magnéticos: Constitución

- El disco duro está formado por una **pila de platos** que son leídos por un **conjunto de cabezas**. Cada plato se divide en **pistas** y éstas, a su vez, se dividen en **sectores**.



Discos magnéticos: Constitución

- **Cabezas (Heads)**
 - Realizan la lectura física de cada superficie del disco.
 - Hay tantas cabezas como superficies útiles. En principio, dos cabezas por cada disco.
 - Las cabezas se mueven a la vez, pues están unidas al mismo brazo.
 - Un multiplexor se encarga de conmutar la lectura/escritura de una a otra cabeza.



Discos magnéticos: Constitución

- **Los platos**
 - **Pistas (Tracks):** Cada una de las líneas concéntricas en que se divide la superficie de un plato.
 - **Sectores:** Unidad mínima de información en la que se dividen las pistas.
 - **Cilindro (Cylinder):** Pistas de todos los platos que equidistan del eje. Es decir, pistas a las que acceden las cabezas cuando el brazo no se mueve.

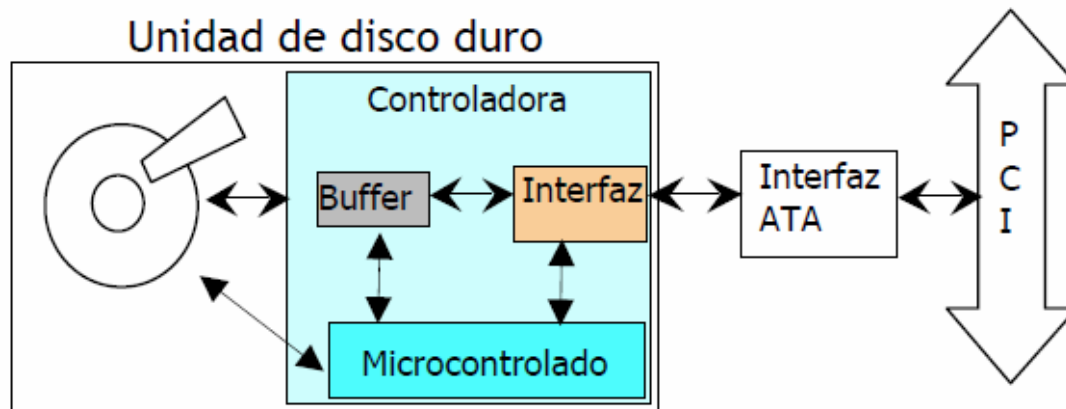


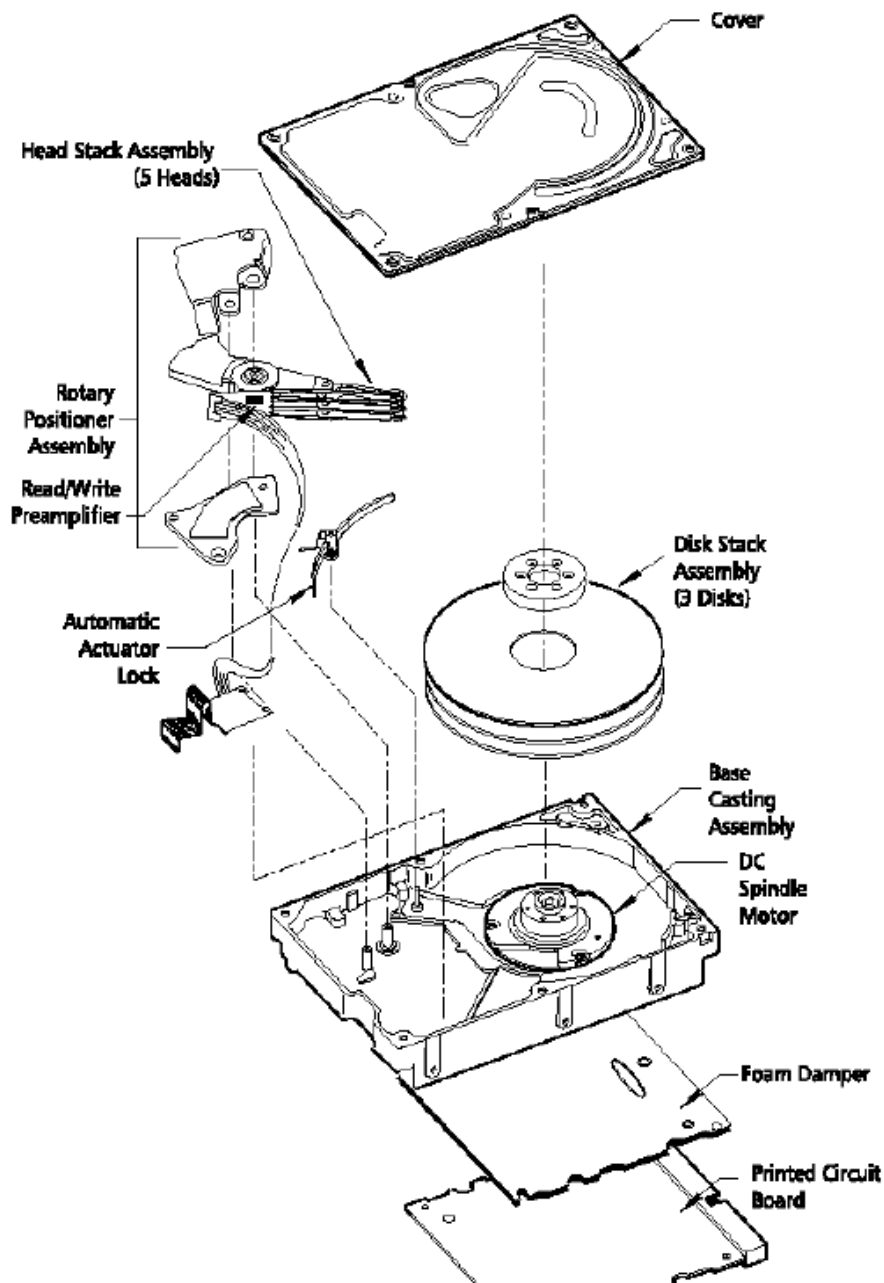
Discos magnéticos: Controladora

- **Dirección CHS (Cylinder Head Sector)**
 - Según lo visto las coordenadas que determinan un sector serían el cilindro, la cabeza y el sector.
 - Ejemplo: Parámetros de un disquete de 3,5 pulgadas:
80 Cylinders, 2 Heads, 512 Bytes/Sector, 18 Sectors/Track
¿Nº bytes?
- **Direccionamiento lógico (LBA, Logic Block Addressing)**
 - Actualmente los sectores se identifican mediante un número de 48 bits, desde 0 hasta el $2^{48}-1$, desde el centro del disco hacia el exterior.
 - La controladora contenida en la unidad de disco se encarga de la traducción **CHS \leftrightarrow LBA**

Discos magnéticos: Controladora

- Las unidades actuales, además de los platos y cabezas lectoras, incorporan toda la circuitería necesaria para el control de los discos.
- También incluyen un **buffer de sectores** que permite acelerar el flujo de datos hacia la interfaz hardware (SATA, SAS).
- La **interfaz** será la encargada de emitir/recibir las órdenes en el bus externo (SATA o SAS) y de recibir/servir la información de los discos.





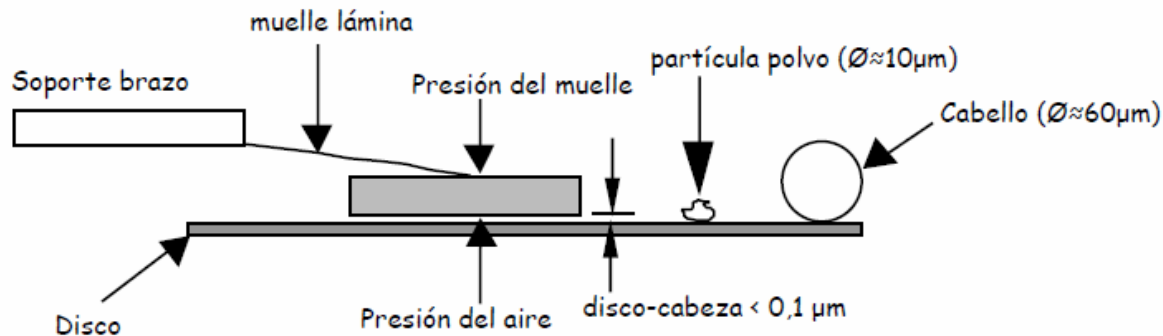
Por ejemplo el fabricante de discos Maxtor declara que ensambla los discos en un ambiente limpio [clase 100](#).

La presión interna y la externa del disco duro se mantienen equilibradas gracias a un orificio que incorpora un filtro de carbono de 0,3 micras.

Discos magnéticos: Funcionamiento

Rotación

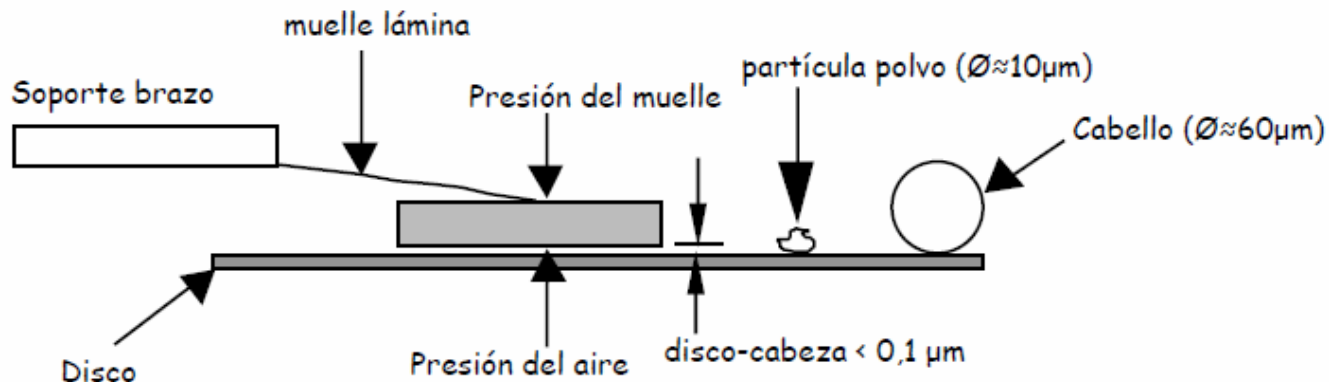
- Los discos se encuentran girando a velocidad elevada (típicamente 5400, 7200, 10000, 15000 r.p.m.) de forma continua, debido a dos razones:
 - La **inercia** ralentizaría excesivamente el proceso de arranque-parada (como en las disqueteras).
 - Gracias a la **forma aerodinámica** del cabezal, el aire al chocar a velocidad elevada provoca una fuerza hacia arriba sobre la cabeza. Esta fuerza equilibra la presión del muelle y evita el contacto disco-cabeza eliminando el rozamiento (desgaste).



Discos magnéticos: Funcionamiento

Latencia mecánica: Posicionamiento cabezas.

- Los brazos, que sujetan a las cabezas, son movidos por diminutos motores eléctricos servoasistidos
 - La **precisión** en el movimiento de las cabezas determinará la separación entre pistas.
 - El **movimiento de traslación del brazo** (seek time) y la velocidad de rotación determinan el tiempo de acceso a los sectores.



Discos magnéticos: Funcionamiento

Latencia mecánica: Posicionamiento cabezas.

- **Tiempo de búsqueda: (*seek time*)**
 - Tiempo empleado en mover la cabeza a la pista requerida. Suele especificarse su valor medio que actualmente no sobrepasa los 10 ms.
- **Latencia rotacional: (*Rotational latency*)**
 - Una vez el cabezal en la pista, es la tardanza en situar la cabeza en el sector. Se especifica el valor medio, que es igual al tiempo que tarda el disco en dar media vuelta. Por lo tanto mejora al aumentar la velocidad de rotación.
- **Tiempo total de acceso**
 - $t_{\text{total}} = t_{\text{búsqueda}} + t_{\text{latencia}}$
 - La mayor capacidad de los discos gracias al mayor diámetro de los platos, limita su velocidad de rotación a 7.200 rpm. La ventaja de ir más despacio es el menor consumo eléctrico y menor rozamiento que, a la postre, significa menos calor. Los platos más pequeños permiten brazos más cortos y ligeros y, por tanto, más rápidos en su movimiento de traslación

Discos magnéticos: Funcionamiento

Aparcamiento

- Cuando el disco se para, el cabezal es aparcado en una zona especialmente diseñada en la parte más cercana al eje del disco, conocida como zona de aparcamiento (***Landing Zone***).
- Esta zona tiene una superficie de textura especial con el fin de limpiar la suciedad que pueda arrastrar el cabezal.
 - Una capa de carbono cubre la superficie actuando de lubricante
 - El contacto con la superficie de aparcamiento solo se produce cuando el disco ya no gira a su velocidad máxima.
 - El área de aparcamiento se encuentra en la zona más interna del disco, lejos del último cilindro de datos.

Discos magnéticos: Formato de grabación

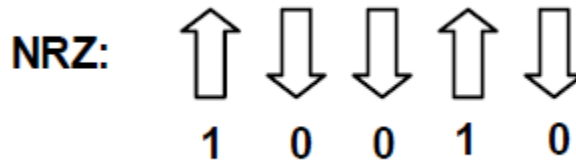
Formato

- Es el conjunto de reglas que determinan la grabación de datos en el disco. Este formato lo dividiremos en 3 niveles según cual sea su objetivo organizativo:
1. **1^{er} nivel:** Se organiza la superficie en dominios magnéticos (áreas) que representarán los bits (bit = cambio de flujo magnético en las partículas de la superficie)
 2. **2^º nivel:** Organización de la superficie en bloques de bits que las cabezas pueden identificar. División en sectores y pistas. Se añade información de identificación y control de errores.
 3. **3^{er} nivel:** Organización del disco para su utilización por el sistema operativo. División en particiones, directorios, tablas de asignación de archivos (FATs), etc.

Discos magnéticos: Formato de grabación

PRIMER NIVEL

- En este primer nivel se trata la información en su forma mas básica, es decir los estados binarios “0” y “1” como inversión del flujo magnético.
- El formato mas sencillo es el NRZ (Non Return to Zero), en el que simplemente cada estado de polarización significa “0” ó “1”.
- Este código tiene la pega de que si existen demasiados “0” ó “1” seguidos no se producen cambios de flujo por lo que necesitaremos una señal de sincronismo. Para evitarlo surgieron otros códigos que solucionan este problema, como son NRZI, PE, MFM, RLL, etc.

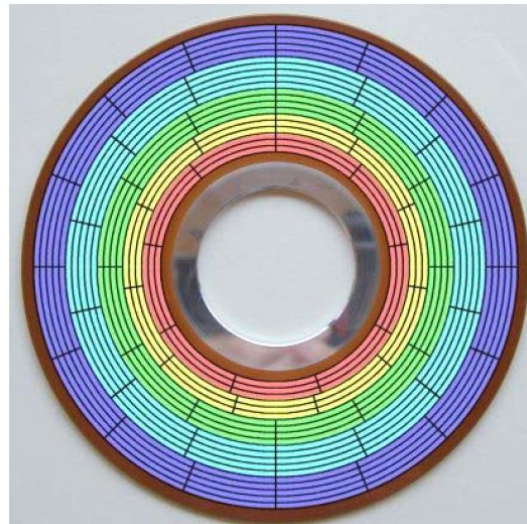


Discos magnéticos: Formato de grabación

SEGUNDO NIVEL

- Este nivel de formato se encarga de organizar todos esos bits individuales de la superficie del disco en bloques de datos a los que accederá la controladora: los **sectores**. Esto supone crear y marcar esos sectores en los que se dividen las pistas (Lo hace el fabricante).
- La **densidad de datos no es uniforme en todo el disco**. Por el contrario el disco se divide en zonas con un número de sectores por pista distinto.

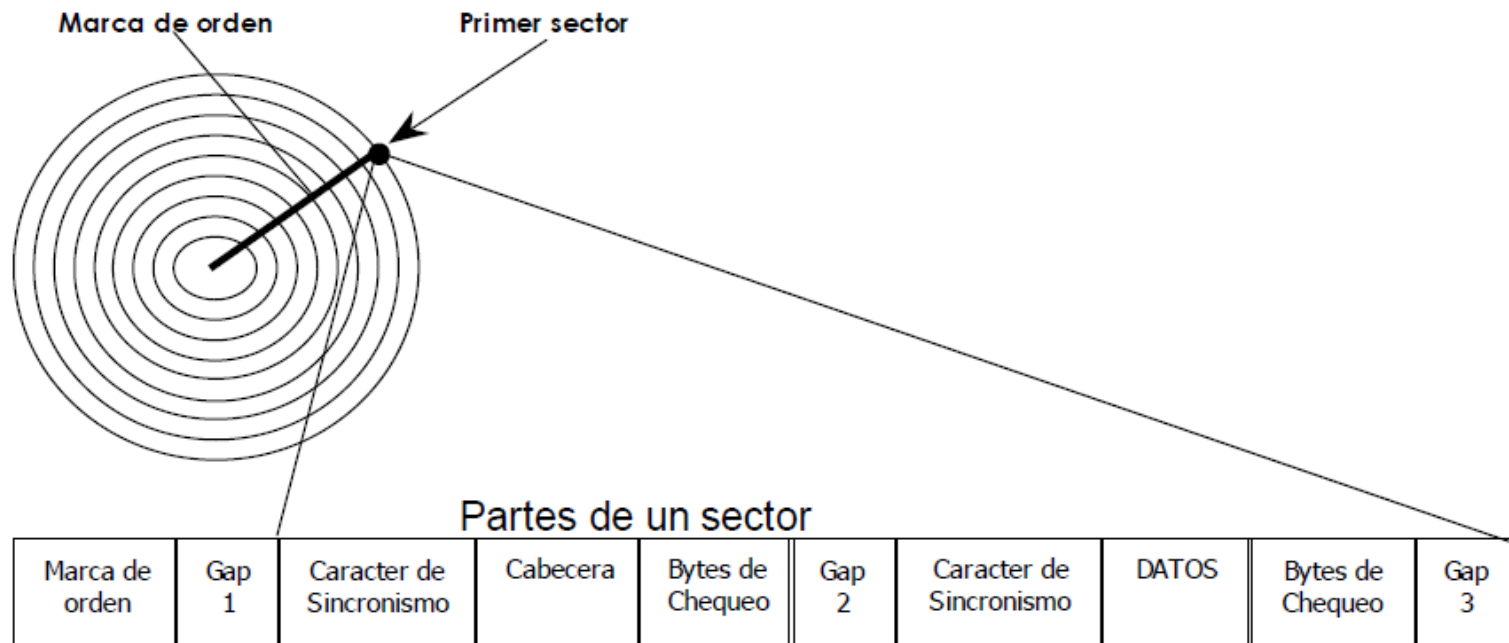
CYLINDER CONTENTS	ZONE ¹	NUMBER OF TRACKS	SECTORS PER TRACK	DATA RATE
System Data	0	68	299	469.23
	1	2582	837	551.28
	2	4412	826	543.59
	3	4123	814	536.20
	4	3539	805	530.22
	5	4068	789	520.51
	6	4380	784	509.62
	7	4388	744	489.88
	8	5022	700	461.54
	9	5190	658	434.62
	10	4406	616	407.05
	11	3332	588	387.96
	12	3985	548	360.95
	13	2933	522	344.62
	14	2914	490	323.53
	15	3696	457	304.20



Discos magnéticos: Formato de grabación

SEGUNDO NIVEL

- Partes de un sector



Discos magnéticos: Formato de grabación

TERCER NIVEL

- En este nivel de formato, **se organiza el disco para su utilización por el sistema operativo**. Esto incluye la división en particiones, en directorios, tablas de asignación de archivos (FATs, MFTs), etc.
- En este nivel de formato, el disco duro se divide en las siguientes partes:
 - Particiones donde se instalarán los sistemas operativos y/o residirán los archivos. Se formatean según el sistema de archivos que se elija (Ext, FAT, NTFS, ...)
 - Sector de particiones o Master Boot Record (MBR): En él se definen las particiones en que se divide el disco, y que pueden mostrarse al usuario para que elija desde un menú.

Unidades de almacenamiento SSD: Diagrama de bloques

- Colocar las etiquetas en las flechas 1, 2, 3, 4 y 5



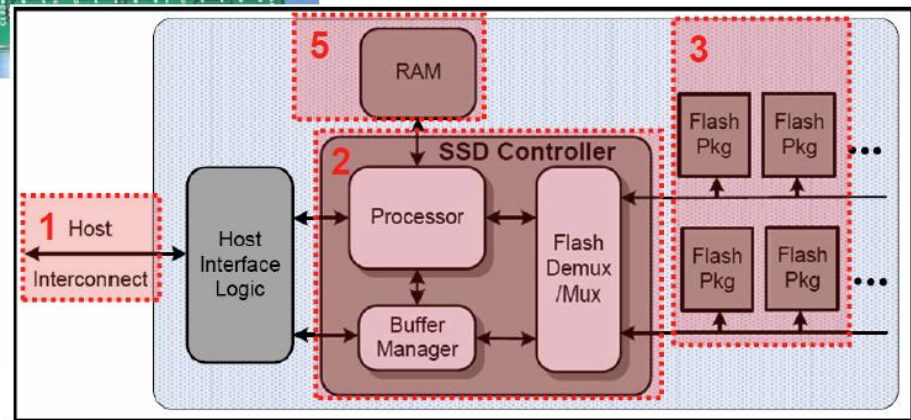
SATA Interface

Buffer

Flash memory

Controller

Power conector

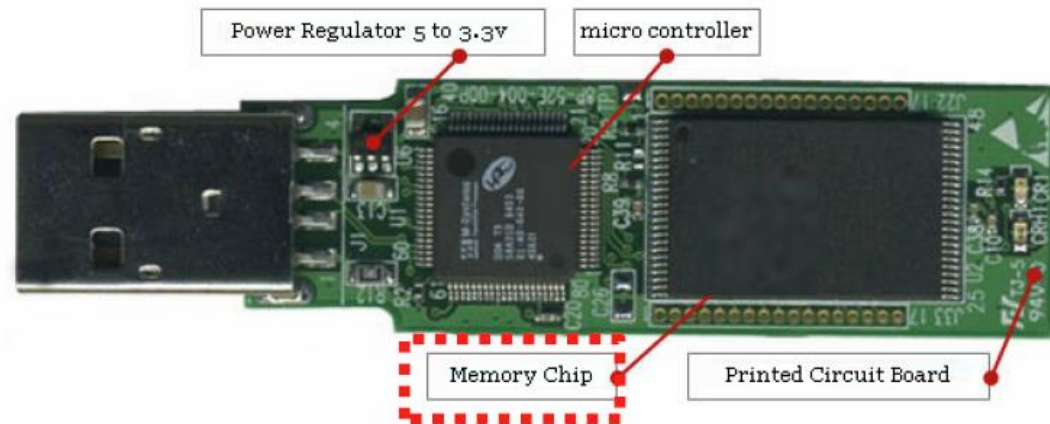


Unidades de almacenamiento SSD: Tecnología de fabricación

- **NAND Flash:** Derivada de la Flash EEPROM usada para la ROM-BIOS:
 - Antes de escribir en una celda, debe borrarse → **Latencia**
 - Ciclos de borrado/escritura LIMITADOS → **Durabilidad**
 - Ejemplo: Intel SSD DC S3710 de 1,2 TB: “La durabilidad se mantiene en 10 escrituras por día durante cinco años, como es costumbre en las unidades destinadas al mercado empresarial”

Imagen: La electrónica de un Pendrive al descubierto

imgarcade.com/1/usb-flash-drive-inside/

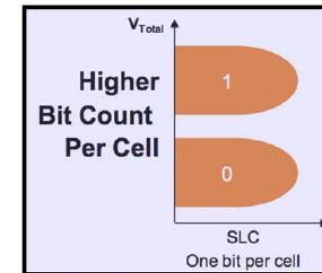


Unidades de almacenamiento SSD: Tecnología de fabricación

- La tecnología de fabricación NAND Flash permite implementaciones distintas:

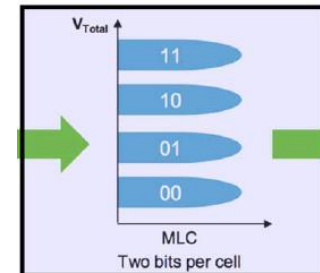
1. SLC flash (single level cell flash)

- 1 bit/celda → 2 estados (niveles de tensión)
- Durabilidad de 100.000 ciclos de borrado.



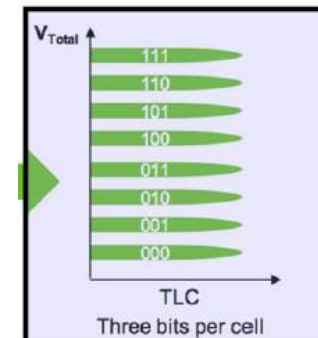
2. MLC flash (Multiple level cell flash)

- 2 bits/celda → 4 estados (niveles de tensión)
- Durabilidad de 10.000 ciclos de borrado (10 veces < SLC)



3. TLC flash (triple level cell flash)

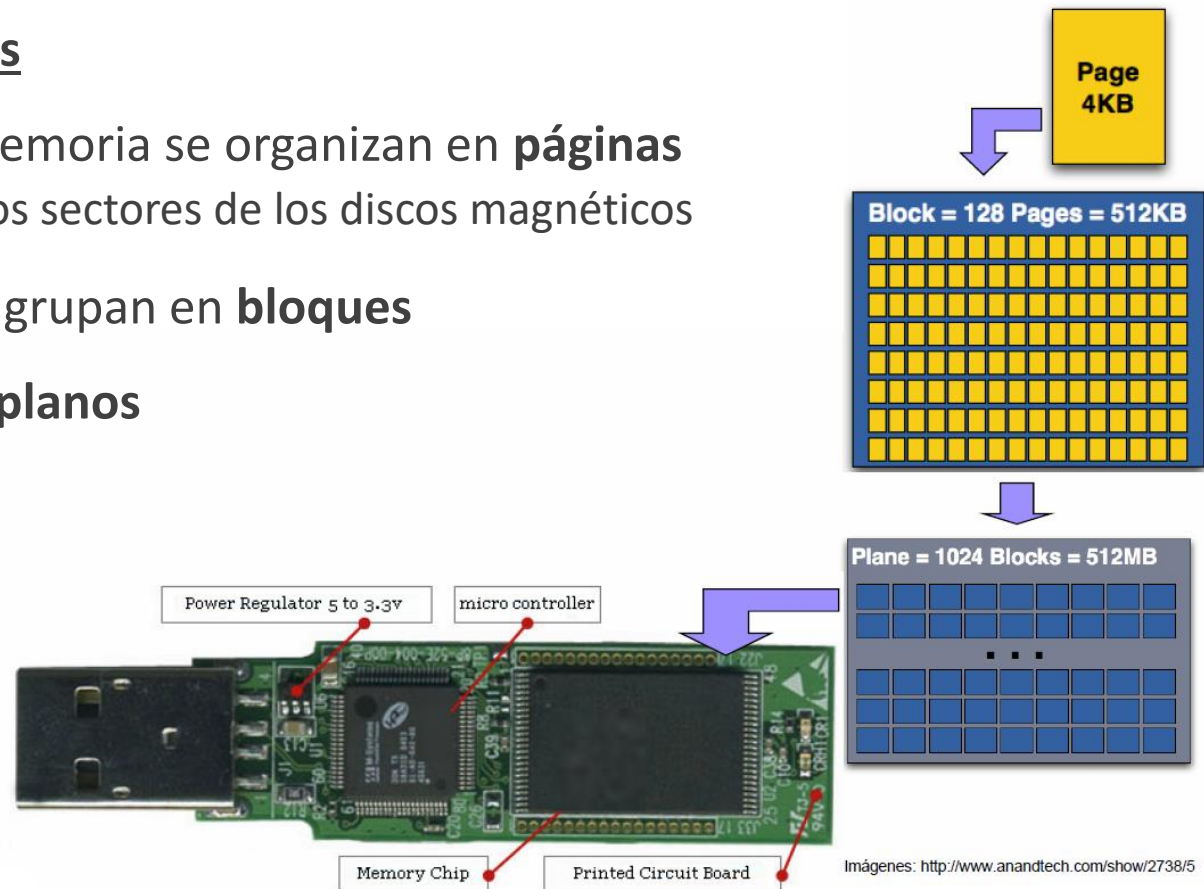
- 3 bits/celda → 8 estados (niveles de tensión)
- Durabilidad menor: 1.000 ciclos de borrado (10 veces < MLC)
- Mayor Bit Error Rate (BER) → menor fiabilidad.
- Prestaciones menores que SLC y MLC



Unidades de almacenamiento SSD: Arquitectura funcional

Páginas vs Bloques

- Las celdas de memoria se organizan en **páginas**
 - Recordemos los sectores de los discos magnéticos
- Las páginas se agrupan en **bloques**
- Los bloques en **planos**
- Etc.



Unidades de almacenamiento

SSD: Ventajas respecto a HD

- ✓ **Mayor fiabilidad:**
No tiene partes móviles

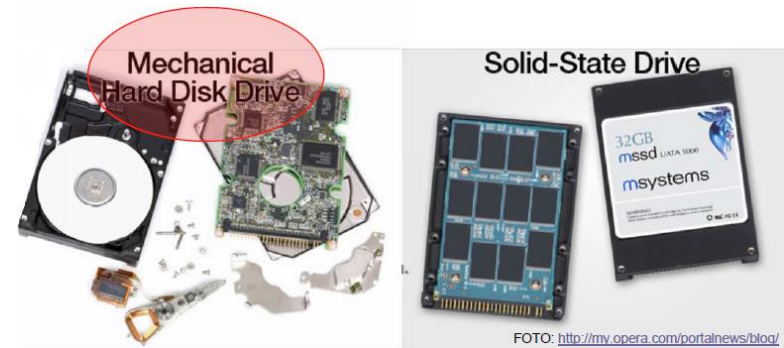
- ✓ **Arranque más rápido:**
No hay que vencer inercias de rotación.

- ✓ **Latencias muy bajas:**
No hay que posicionar el cabezal sobre la pista que contiene el sector

- ✓ **Consumo mucho menor:**
Ejemplo consumo en W:

Tipo unidad	Idle	Seek	Start-up
SATA	8 W	10 W	20 W
SSD	0.08 W	0.15 W	N/A

- ✓ **Velocidad de Lectura determinista:**
No depende de la localización del dato



Mantenimiento:

Parámetros de fiabilidad

- Según los fabricantes, los discos duros no requieren mantenimiento preventivo, pero su **disponibilidad** viene especificada básicamente por 2 parámetros ligados entre sí:

1. MTBF

Valores típicos de 500.000 horas para discos destinados a sobremesa y de más de 1.000.000 horas para los orientados a misión crítica. No olvidar el carácter estadístico de este parámetro.

2. Vida operativa

Es un valor estadístico que especifica el tiempo durante el que su disponibilidad viene dada por el MTBF. Suele encontrarse entre 3 y 5 años. (Más los SAS que los SATA).

Lectura recomendada: <https://www.xataka.com/tecnologiaz/en/como-saber-cuanto-va-a-durar-nuestro-disco-duro-sin-fallos-la-fecha-mas-temida-por-todos>

Mantenimiento: Parámetros de fiabilidad

La forma de saber el tiempo de vida de un disco magnético o la cantidad de datos escritos en un disco flash es tan fácil como instalar un programa que lea estas cifras en los registros internos de las unidades.

Una buena herramienta es CrystalDiskInfo, que permite saber tanto la cantidad de datos escritos en un SSD como las horas que un HDD ha estado encendido. Y con estos datos, podemos hacernos una idea acerca del estado de nuestras unidades de almacenamiento.

The screenshot shows the CrystalDiskInfo 8.0.0 application window. The main display shows the health status of the selected drive (C:) as 'Bueno' (Good) with a '100 %' health indicator. The temperature is shown as '41 °C'. The drive is identified as 'SAMSUNG MZVLB256HAHQ-00000 256.0 GB'. The interface includes a menu bar (Archivo, Editar, Función, Tema, Disco, Ayuda, Lengua(Language)) and a toolbar. The drive's specifications are listed in a table:

Item	Value
Estado de salud	Bueno 100 %
Temperatura	41 °C
Firmware	EXD7101Q
Número de serie	S444NF0K500680
Interfaz	NVM Express
Modo de transferencia	PCIe 3.0 x4 PCIe 3.0 x4
Letra de unidad	C:
Estándar	NVM Express 1.2
Características soportadas	S.M.A.R.T.
Total Host Reads	1357 GB
Total Host Writes	1440 GB
Velocidad rotación	---- (SSD)
Nº encendido	256 veces
Horas encendido	82 horas

Below the specifications, there is a table of S.M.A.R.T. attributes:

ID	Detalles ID	Valores en crudo
01	Critical Warning	00000000000000
02	Composite Temperature	0000000000013A
03	Available Spare	000000000000064
04	Available Spare Threshold	00000000000000A
05	Percentage Used	000000000000000
06	Data Units Read	000000002A6F96
07	Data Units Written	000000002D01D2
08	Host Read Commands	00000001D5C2B8
09	Host Write Commands	00000002EB5DC5
0A	Controller Busy Time	000000000000079

Mantenimiento: Factores ambientales

- Esta disponibilidad teórica se ve reducida por varios factores, como los **ciclos de arranque parada** (un valor típico puede situarse entre 40.000 y 50.000) y los **factores ambientales**:

1. Temperatura

- Las cada vez más elevadas velocidades de giro provocan un mayor **calentamiento** que puede requerir **ventilación añadida** para no perjudicar la vida del disco
- **Rango de temperatura no operativo:** Es el rango de temperaturas a las que puede someterse sin entrar en funcionamiento (Almacenaje). Un rango típico puede ser entre -40°C y 70°C .
- **Temperatura operativa:** Rango de temperaturas dentro de las que el disco debe funcionar para que puedan cumplirse sus expectativas de vida en funcionamiento. Aunque se respete el rango, no deben someterse a grandes variaciones de temperatura ya que la deformación brusca de los discos podría inutilizarlos. Un valor típico especificado es de 20°C/hora .

Mantenimiento: Factores ambientales

- Esta disponibilidad teórica se ve reducida por varios factores, como los **ciclos de arranque parada** (un valor típico puede situarse entre 40.000 y 50.000) y los **factores ambientales**:

2. Altitud

- Debido los principios aerodinámicos en los que se basa su funcionamiento, no pueden superarse unos 3.000 metros respecto el nivel del mar (Presión atmosférica).

3. Humedad

- Suele especificarse la tolerancia a la humedad en % de humedad relativa y también es habitual señalar que debe evitarse la condensación, ya que esta puede ser fatal para cualquier dispositivo electrónico.

Mantenimiento:

Factores ambientales

- Esta disponibilidad teórica se ve reducida por varios factores, como los **ciclos de arranque parada** (un valor típico puede situarse entre 40.000 y 50.000) y los **factores ambientales**:

4. Vibraciones

- Aunque no existe un criterio uniforme, es típico especificar la vibración en valores RMS.

5. Impactos

- Se miden en los tres ejes.
- Se expresan en unidades relativas a la aceleración de la gravedad: $1\text{ G} = 9.8\text{ m/s}^2$.
- Normalmente, se aplican impactos con forma de un semiperiodo sinusoidal de 2 ms de duración.
- Los discos para ordenadores portátiles suelen tener mejor resistencia al impacto.

Mantenimiento: Tecnología S.M.A.R.T.

S.M.A.R.T. = Self-Monitoring Analysis and Reporting Technology

- Permite a una unidad de disco monitorizar parámetros de funcionamiento, analizarlos e informar sobre incidencias que pudieran afectar a su funcionamiento.
- En unidades SATA, la capacidad S.M.A.R.T. puede habilitarse y deshabilitarse desde el BIOS del sistema o desde software del fabricante del disco.
- Cuando se detecta probabilidad de fallo, S.M.A.R.T. realiza un informe de estado disponible para que el sistema informe al usuario y pueda, por ejemplo, realizar un backup.

Mantenimiento: RAID

- El término original suele traducirse como ***Redundant Array of Inexpensive Disks***, aunque actualmente también puede verse traducido como ***Redundant Array of Independent Disks***.
- Básicamente se trata de aumentar la capacidad y/o la fiabilidad del almacenamiento en disco agrupando varios discos como si fueran uno.
- La interfaz de discos debe soportar esta característica, sin embargo los discos no necesitan de ninguna propiedad especial, salvo que suele ser necesario/recomendable que sean idénticos.



Figura 10: Tarjeta de interfaz SATA RAID Promise SuperTrak EX4350. Bajo el disipador se encuentra el micro, que gestiona el conjunto de discos según el firmware que ejecuta.

Mantenimiento: RAID

- **Capacidad de almacenamiento**

El hecho de agrupar varios discos como si fuera uno solo puede considerarse un aumento de capacidad, pero el uso de redundancia hace que la capacidad total del conjunto sea menor que la suma individual.

- **Rendimiento**

Los datos se distribuyen sobre los discos de forma que las peticiones de lectura/escritura puedan ejecutarse en paralelo, aumentando así el rendimiento.

- **Fiabilidad**

Los discos redundantes permiten almacenar información de paridad, que se usará en caso de fallar alguno. Tras cambiar el disco defectuoso por otro nuevo, la interfaz RAID usará esta información para la reconstrucción del conjunto.

Mantenimiento: RAID

- **Disco de reserva**

El hecho de agrupar varios discos como si fuera uno solo puede considerarse un aumento de capacidad, pero el uso de redundancia hace que la capacidad total del conjunto sea menor que la suma individual.

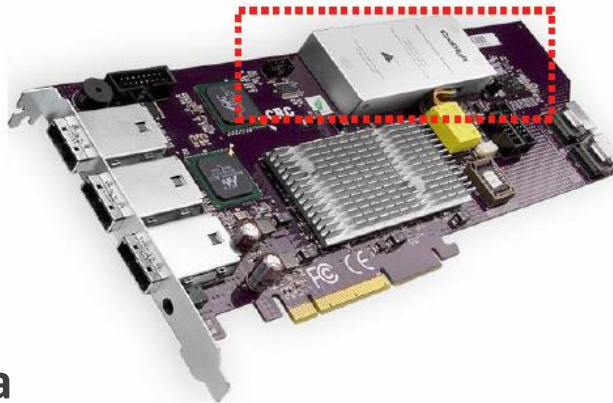


Figura 13: Tarjeta de interfaz RAID con batería de reserva (resaltada con línea discontinua).

- **Batería de reserva**

Algunas tarjetas de interfaz RAID ofrecen la posibilidad de incorporar una batería de respaldo que en caso de un fallo de alimentación permite continuar su funcionamiento durante un breve intervalo.