

# MEMORIA

Programación II

# Temas

---

- Subsistema de Memoria
- Organización de Memoria Principal
- Memoria Cache

# Contexto actual (1)

- Velocidad del procesador: se duplica cada 18 meses (sin variar su precio) la cantidad de instrucciones ejecutadas por segundo.
- Memoria: se cuadruplica su tamaño cada 36 meses (al mismo precio). Velocidad aumenta a razón de un 10% anual.

# Contexto actual (2)

- A medida que aumenta la brecha entre las velocidades del procesador y de la memoria, las distintas arquitecturas buscan tender un puente sobre esta brecha.
- Una computadora típica suele tener distintos tipos de memoria, desde una rápida y cara (registros) hasta una barata y lenta (discos).

# Contexto actual (3)

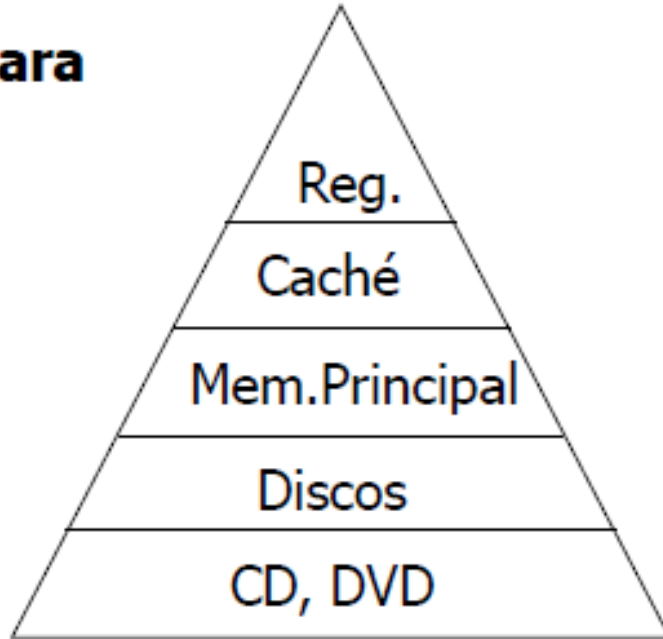
- La interacción entre los diferentes tipos de memoria se aprovecha de forma tal que se logra un comportamiento, por parte de la computadora, equivalente al que tendría con una memoria única, grande y rápida, cuando en realidad tiene distintos tipos de memoria trabajando en forma coordinada.

# Jerarquía de memorias (1)

- La forma en que se organizan estos distintos tipos de memoria es lo que se conoce como jerarquía de memoria.
- En la cima de la jerarquía están los registros.
- En la base, las memorias secundarias (discos magnéticos) y de almacenamiento “off line” (CD, DVD, cintas).

# Jerarquía de memorias (2)

**Rápida y cara**



**Lenta pero barata**

# Jerarquía de memorias (3)

- A medida que ascendemos tenemos mayor rendimiento y más costo por bit.
- Entre la memoria principal y la secundaria hay otro tipo de memoria para salvar la brecha.
- Cuando ascendemos, también aumenta la frecuencia de accesos a ese tipo de memoria



# Jerarquía de memorias (4)

## □ Memoria del computador :

- Tecnologías diferentes
- Fundamentos físicos distintos
- Localización en lugares distintos



## □ Objetivo :

- Capacidad de almacenamiento
- Tiempo de acceso reducido

# Jerarquía de memorias (4)

Tipos de memoria	Tiempos de acceso	Tamaño típico
Registros	1 ns	1 KB
Caché	5-20 ns	1 MB
Mem. Principal	60-80 ns	1 GB
Discos	10 ms	160GB

# Características (1)

## □ **Duración de la información:**

- Memorias volátiles: RAM
- Memorias no volátiles: discos, cintas
- Memorias permanentes: ROM, EPROM

## □ **Modo de acceso:**

- Acceso por palabra: memoria principal
- Acceso por bloque: discos, caché

# Características (2)

## □ Velocidad

### Memorias semiconductoras:

Tiempo de acceso: tiempo máximo que transcurre desde que se inicia la operación de lec/esc hasta obtener/almacenar el dato.

### Tiempo de ciclo:

Tiempo mínimo que tiene que haber entre dos operaciones sucesivas sobre la memoria

$$t_{\text{ciclo}} > t_{\text{acceso}}$$

# Características (3)

- **Memorias magnéticas: ¿como se calcula?**

**Tiempo de acceso:** tiempo de posicionar el cabezal + tiempo de latencia (+ tiempo de lectura)

**Velocidad de transferencia:** bytes/seg

# Características (4)

## □ **Métodos de acceso**

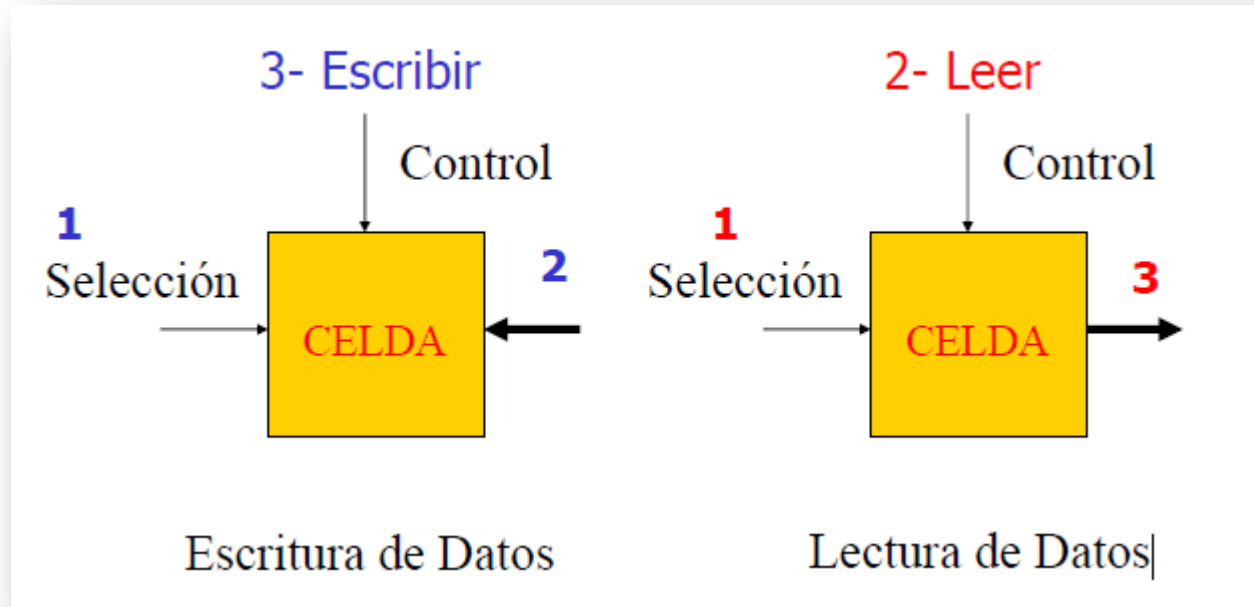
- **Acceso aleatorio:** el tiempo para acceder a una locación dada es independiente de la secuencia de accesos anteriores y es constante. Ejemplo la memoria principal.
- **Acceso secuencial:** el acceso debe hacerse en una secuencia lineal específica. Variable. Ejemplo son las unidades de cinta.

# Características (5)

---

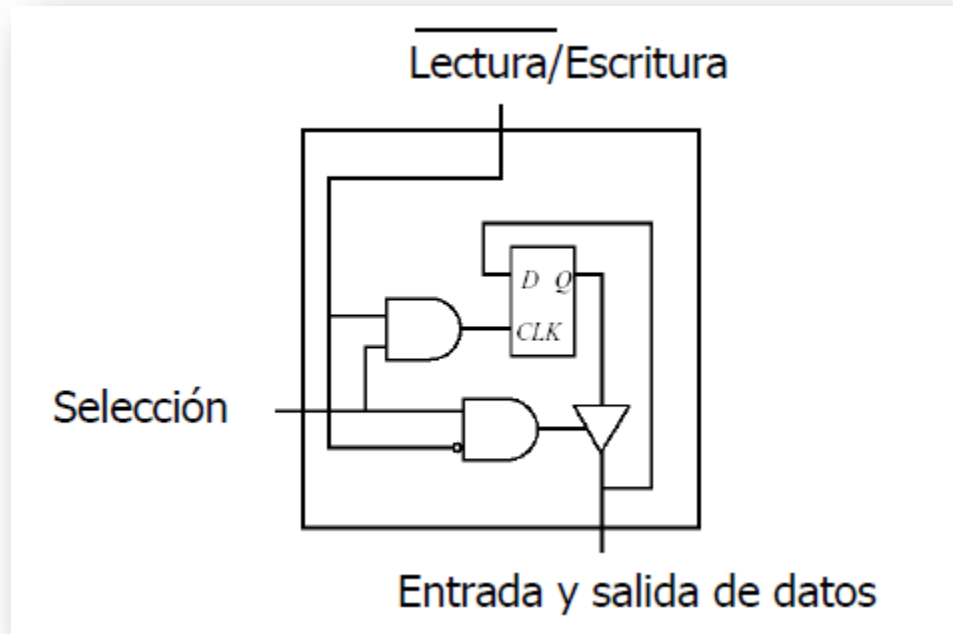
- **Acceso directo:** los bloques ó registros individuales tienen una dirección única que se basa en la localización física. Variable. Ejemplo los discos magnéticos.
- **Acceso asociativo:** memoria caché

# Organización de la memoria (3)





# Celda de memoria



# Memoria de acceso aleatorio

**RAM (Random Access Memory).** Aleatorio significa que se puede acceder a cualquier celda de memoria en el mismo tiempo, independientemente de la posición en la estructura de la memoria.

Hay dos tipos:

Basada en flip flops: memoria estática (SRAM)

Basada en transistores: memoria dinámica (DRAM). Cargas almacenadas en transistores (capacitores) *El acceso a las ROM también es de éste tipo*

# RAM estática o SRAM (1)



El almacenamiento en RAM estática se basa en circuitos lógicos denominados flip-flop, que retienen la información almacenada en ellos mientras haya energía suficiente para hacer funcionar el dispositivo .

Un chip de RAM estática puede almacenar tan sólo una cuarta parte de la información que puede almacenar un chip de RAM dinámica de la misma complejidad.

# RAM estática o SRAM (2)



La RAM estática no requiere ser actualizada y es normalmente mucho más rápida que la RAM dinámica (el tiempo de ciclo de la SRAM es de 8 a 16 veces más rápido que las DRAM).

También es más cara, por lo que se reserva generalmente para su uso en la memoria de acceso aleatorio (caché).

# RAM dinámica o DRAM

**DRAM** almacenan la información en circuitos integrados que contienen condensadores, que pueden estar cargados o descargados. Almacena más información que SRAM en la misma superficie. Los “capacitores” son más chicos que los flip flop.

❖ Como éstos pierden su carga en el transcurso del tiempo, se debe incluir los circuitos necesarios para "refrescar" los chips de RAM cada pocos milisegundos, para impedir la pérdida de su información

❖ Usada como memoria principal

# RAM dinámica o DRAM

- Algunas memorias dinámicas tienen la lógica del refresco en la propia pastilla, dando así gran capacidad y facilidad de conexión a los circuitos. Estas pastillas se denominan casi estáticas.
- Mientras la RAM dinámica se refresca, el procesador no puede leerla. Si intenta hacerlo en ese momento, se verá forzado a esperar. Como son relativamente sencillas, las RAM dinámicas suelen utilizarse más que las RAM estáticas, a pesar de ser más lentas.

# TIPOS DE MODULOS DE RAM

## Los módulos que se utilizan en la actualidad



**DDR2 RAM:** tienen 240 pines. Los zócalos no son compatibles con DDR RAM. La muesca está situada dos milímetros hacia la izquierda con respecto a DDR RAM. Se comercializan pares de módulos de 2 Gb (2x2= 4 GB). Pueden trabajar a velocidades entre 400 y 800 MHz.



**DDR RAM:** sucesora de la memoria SDRAM, tiene un diseño similar pero con una sola muesca y 184 contactos. Ofrece una velocidad entre 200 y 600 MHz. Se caracteriza por usar un mismo ciclo de reloj para hacer dos intercambios de datos a la vez.



**Rambus:** puede ofrecer velocidades entre 600 y 1.066 MHz. Tiene 184 contactos. Algunos de estos módulos disponen de una cubierta de aluminio (dispersor de calor) que protege los chips de memoria de un posible sobrecalentamiento. Debido a su alto coste, su utilización no se ha extendido mucho.

# SDRAM



**SDRAM:** se instalan sin necesidad de inclinarlos con respecto a la placa base. Se caracterizan porque el módulo tiene dos muescas. El número total de contactos es de 168. Pueden ofrecer una velocidad entre 66 a 133 MHz. En la actualidad casi ya no se comercializan.

## ❖ Synchronous DRAM (SDRAM)

### Actualmente en DIMMs

### Acceso sincronizado con un reloj externo

Se presenta una dirección a la RAM

RAM encuentra los datos (y CPU esperaría la DRAM)

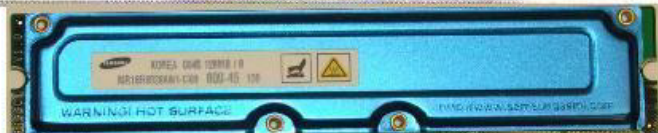
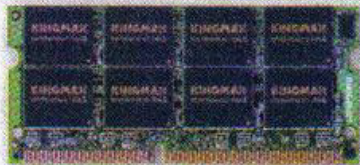
SDRAM mueve datos en tiempo del reloj del sistema, la CPU conoce cuando los datos estarán listos.

CPU puede hacer otra cosa mientras tiene que esperar

### Modo Burst permite SDRAM trabajar en bloques



# TIPOS DE MODULOS DE RAM



**SODIMM:** el tamaño de estos módulos es más reducido que el de los anteriores ya que se emplean sobre todo en ordenadores portátiles. Se comercializan módulos con capacidades de 512 MB y 1 GB. Los hay de 100, 144 y 200 contactos.

**MEMORIAS RIMM** acrónimo de Rambus Inline Memory Module, designa a los módulos de memoria RAM que utilizan una tecnología denominada RDRAM, desarrollada por Rambus Inc. A pesar de tener la tecnología RDRAM niveles de rendimiento muy superiores a la tecnología SDRAM y las primeras generaciones de DDR RAM, debido al alto costo de esta tecnología no han tenido gran aceptación en el mercado de PCs. Su momento álgido tuvo lugar durante el periodo de introducción del Pentium 4 para el cual se diseñaron las primeras placas base, pero Intel ante la necesidad de lanzar equipos más económicos decidió lanzar placas base con soporte para SDRAM y más adelante para DDR RAM desplazando esta última tecnología a los módulos RIMM del mercado.

# Organización interna de la memoria

Una celda de memoria es capaz de almacenar un bit de información.

Por lo general, varias celdas se organizan en forma de arreglo.

En general la celda tiene 3 terminales

funcionales capaces de llevar una señal

eléctrica:

- ❖ **Selección:** selecciona una celda de memoria
- ❖ **Control:** especifica lectura ó escritura
- ❖ **Escritura/Lectura** de datos

# Las nuevas tecnologías en RAM (1)

## **Enhanced DRAM (ESDRAM)**

Para superar algunos de los problemas de latencia inherentes con los módulos de memoria DRAM standard, se incluye una cantidad pequeña de SRAM directamente en el chip, eficazmente creando un caché en el chip. Permite tiempos de latencia más bajos y funcionamientos de 200 mhz. La SDRAM oficia como un caché dentro de la memoria. 7. Una de las desventajas de estas memorias es que su valor es 4 veces mayor al de la memoria DRAM.

# Las nuevas tecnologías en RAM (3)

---

## **Synclink DRAM - SLDRAM**

Es una DRAM es una extensión más rápida y mejorada de la arquitectura SDRAM que amplía el actual diseño de 4 bancos a 16 bancos.

El ancho de banda de SLDRAM es de los más altos 3.2GB/s y su costo no sería tan elevado.

# Bibliografía recomendada

---

- Capítulo 4: Memoria Interna. Stallings. 5ta Ed.

- Link sugerido:

<http://www.pctechguide.com/14Memory.htm>

- 
- MEMORIA CACHE

# Memoria Caché

- Históricamente CPU han sido más rápidas que las memorias.
- El aumento de circuitos que es posible incluir en un chip
  - ❖ Diseñadores de CPU lo usaron para hacerla más veloz (ej. pipeline).
  - ❖ Los diseñadores de memoria lo usaron para aumentar la capacidad del chip (más memoria, más grandes decodificadores).

# Memoria Caché (2)

- Esta diferencia implica: después que la CPU „emite“ una solicitud de lectura a la memoria (bus de direcciones, bus de control) pasan muchos ciclos de reloj antes que reciba la palabra que necesita, por el bus de datos.



# Memoria Caché (3)

- En todos los ciclos de instrucción, la CPU accede a memoria al menos una vez, para buscar la instrucción y muchas veces accede a buscar operandos.
- La velocidad a la cual la CPU ejecuta instrucciones está limitada por el tiempo del ciclo de memoria.

# Memoria Caché (4)

---

- El problema no es tecnológico sino económico. Se pueden construir memorias tan rápidas como la CPU, pero para obtener la máxima velocidad tiene que estar dentro del chip de la CPU llegar a la memoria por el bus del sistema es "lento".

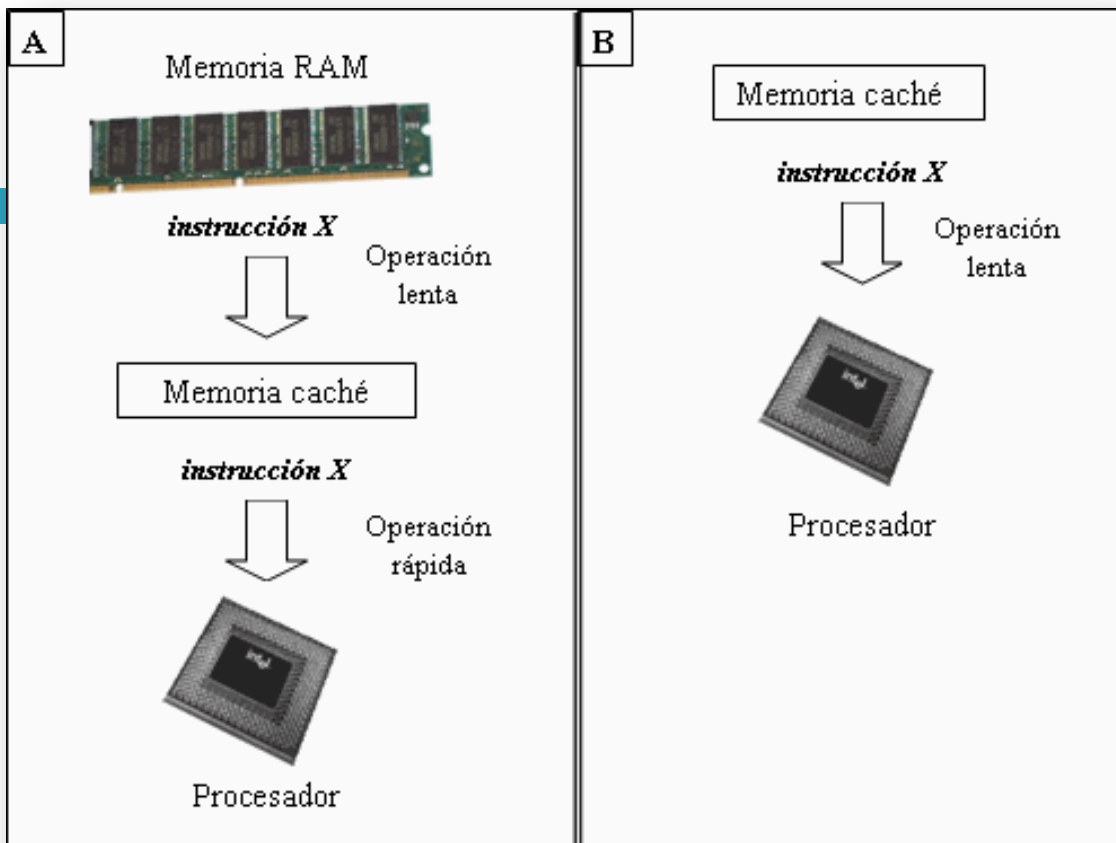
# Memoria Caché (5)

---

## ➤ Solución

Técnicas para combinar una cantidad pequeña de memoria rápida con una cantidad grande de memoria lenta, para obtener la velocidad de memoria "casi" rápida.

# Funcionamiento de la Memoria Caché



Si el procesador desea buscar la instrucción X (situada en cierta posición de la RAM), intentará primero hallar dicha instrucción en la caché. Si la encuentra allí, el acceso será rápido (caso B). Si no la encuentra, se añade un retraso extra al buscar la instrucción en la RAM y copiarla en la caché (caso A).

# Principios (1)

➤ El uso de la memoria caché se sustenta en dos principios ó propiedades que exhiben los programas:

1. Principio de localidad espacial de referencia

➤ cuando se accede a una palabra de memoria, es "muy probable" que el próximo acceso sea en la vecindad de la palabra anterior.

# Localidad Espacial

---

Se sustenta en:

- ☐ Ejecución secuencial del código

- ☐ Tendencia de los programadores a hacer próximas entre sí variables relacionadas

- ☐ Acceso a estructuras tipo matriz ó pila

# Principios (2)



2. Principio de localidad temporal de referencia cuando se accede a una posición de memoria, es “muy probable” que un lapso de “tiempo corto”, dicha posición de memoria sea accedida nuevamente.

# Localidad Temporal

---

Se sustenta en:

Formación de ciclos o bucles

Subrutinas (Procedimientos o Funciones)

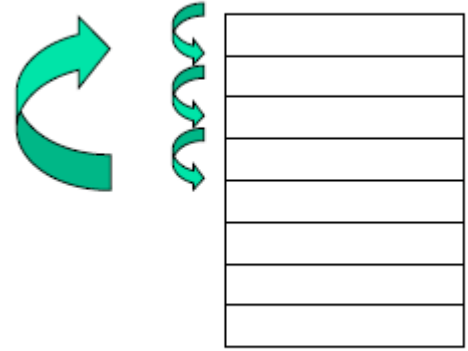
Pilas



# Ejemplo

Estas 2 sentencias exhiben los dos principios antes mencionados:

```
for i=1 to i=10, do  
  A[i]:=0;
```



□ En cada ciclo se consulta cuanto vale  $i$ .

□ Cada asignación  $A[i]:=0$  almacena un 0 en un elemento del arreglo (el siguiente).

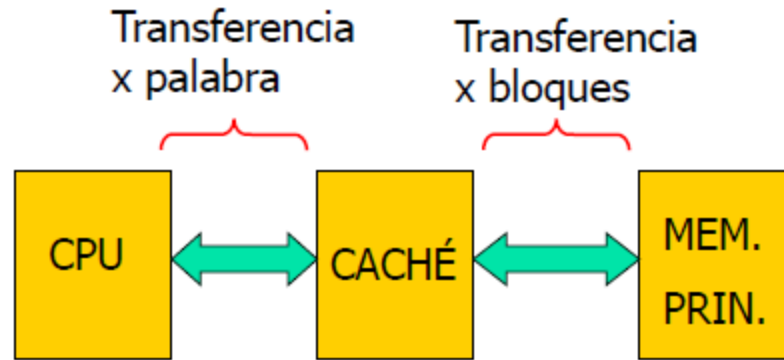
# Memoria Cache



La idea general es que cuando se hace referencia a una palabra, ella y alguna de las vecinas se traen de la memoria grande y lenta a la caché, para que en el siguiente acceso la palabra buscada se encuentre en el caché.

# Memoria Cache

• Ej.



# Aciertos y fallos (1)

□ La efectividad de la caché se expresa a través de la frecuencia de aciertos: es decir el número de veces que la caché acierta direcciones.

□ Un acierto de caché sucede cuando los datos que necesita el procesador están almacenados en la caché entonces la CPU obtiene los datos a alta velocidad.

# Aciertos y fallos (2)

---

□ Un fallo de caché ocurre cuando los datos buscados no se encuentran en la caché.

En este caso la CPU tiene que obtenerlos de la memoria principal, a una velocidad menor.

# MEMORIA EXTERNA

Programación II

# Tipos de memoria Externa

---

## Discos magnéticos

## Discos ópticos

- CD-ROM
- CD-R
- CD-RW
- DVD

## Cintas Magnéticas

# Discos Magnéticos

---

## Platos

- Superficies de Al cubiertos con óxido de Fe, material magnético.

☐ Ahora también se usa vidrio

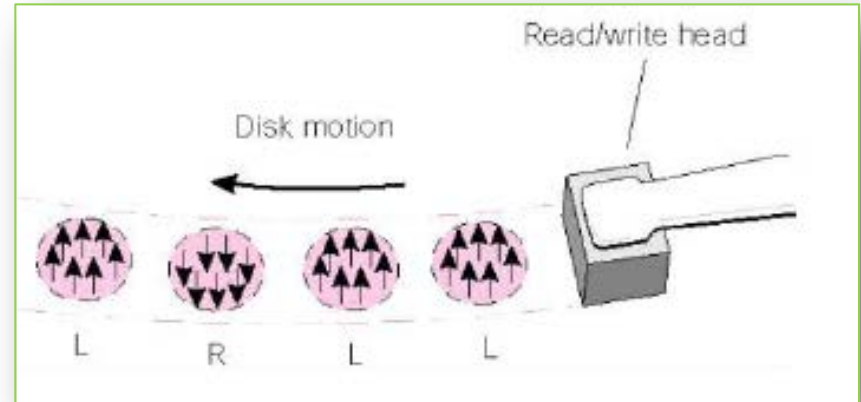
- Se dilata menos que el Al.
- Superficie más uniforme.
- Reducción de defectos superficiales.



# Principios Físicos

- Pequeñas áreas del disco son magnetizadas en diferentes direcciones por un transductor.
- Debe existir un movimiento relativo entre disco y el transductor al momento de la lectura/escritura.

Cambios en la dirección de magnetización es lo que se detecta en la lectura

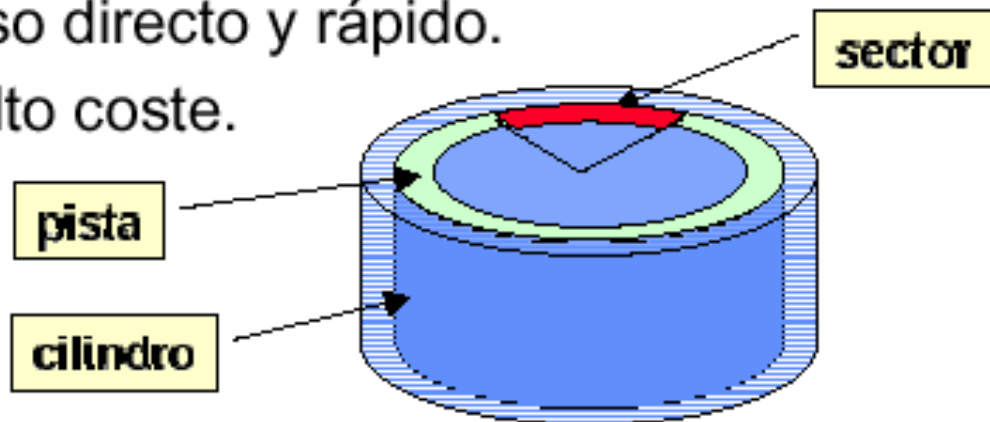


# Mecánica de Lectura y Escritura

- ❖ Lectura y escritura es a través de una cabeza transductora (bobina).
- ❖ Durante lectura/escritura, la cabeza es estacionaria y el plato gira.
- ❖ Se almacenan ceros y unos por medio de la magnetización de pequeñas áreas del material.

– Dispositivos de acceso directo (**discos**)

- **Grabación.** Los registros se agrupan en bloques denominados “**sectores**”, que a su vez se agrupan en “**pistas**” y “**cilindros**”. El acceso se efectúa de forma directa.
- **Recuperación.** Acceso a un bloque arbitrario.
- **Ventajas.** Acceso directo y rápido.
- **Desventajas.** Alto coste.

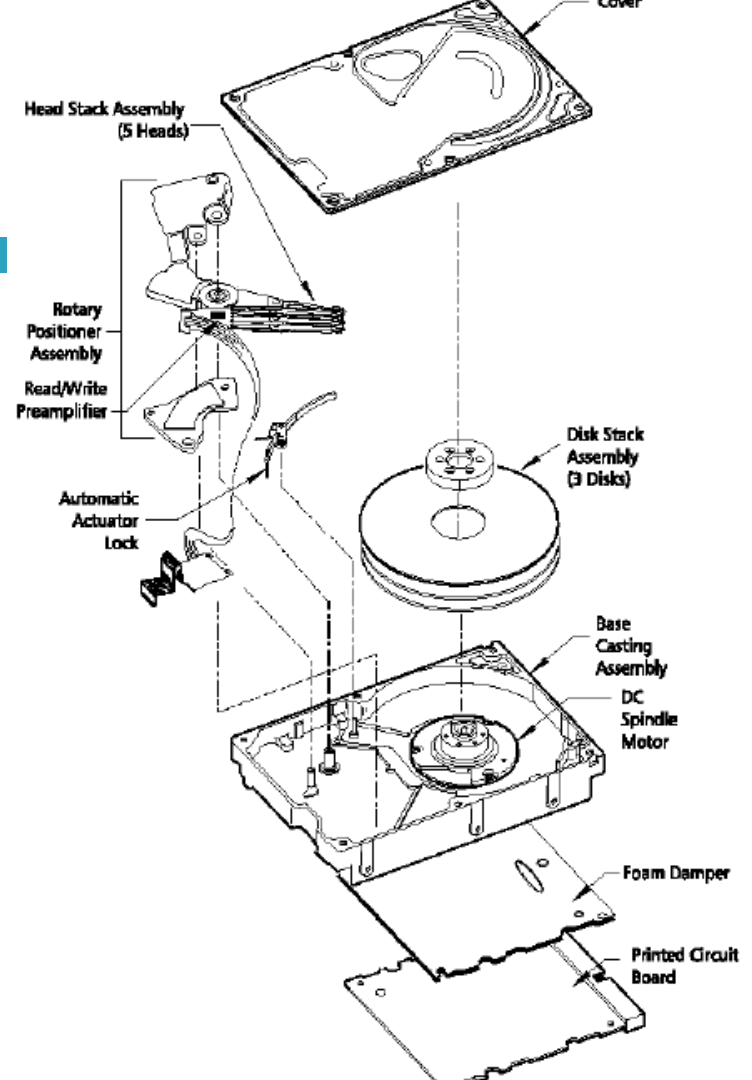


# Organización de los Datos

- ❖ Anillos concéntricos: **pistas** ó tracks.
  - Espacios (gaps) entre pistas
  - Reducir gaps para aumentar capacidad
  - Mismo No de bits por pista
  - Velocidad angular constante
- ❖ Pistas divididas en sectores
- ❖ Mínimo tamaño de bloque: **sector**
- ❖ Bloque: más de un sector (**cluster**)

# Características posibles

- Cabeza fija (raro) o móvil.
- Disco removible o fijo.
- Simple ó doble lado.
- Uno ó múltiples platos.
- Mecanismo de cabeza:
  - Contacto (Floppy)
  - Distancia Fija
  - Aerodinámica (Winchester)

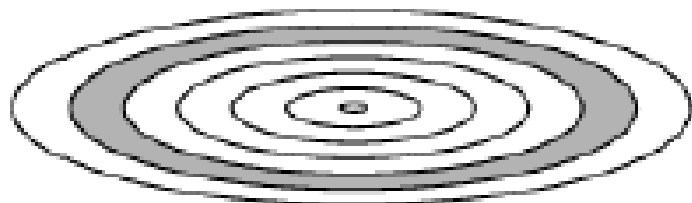
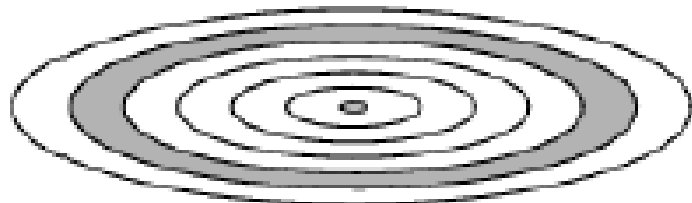
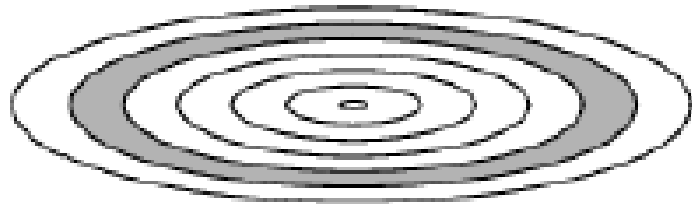
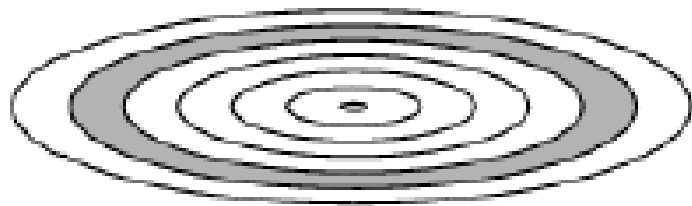


# Estructura de un disco

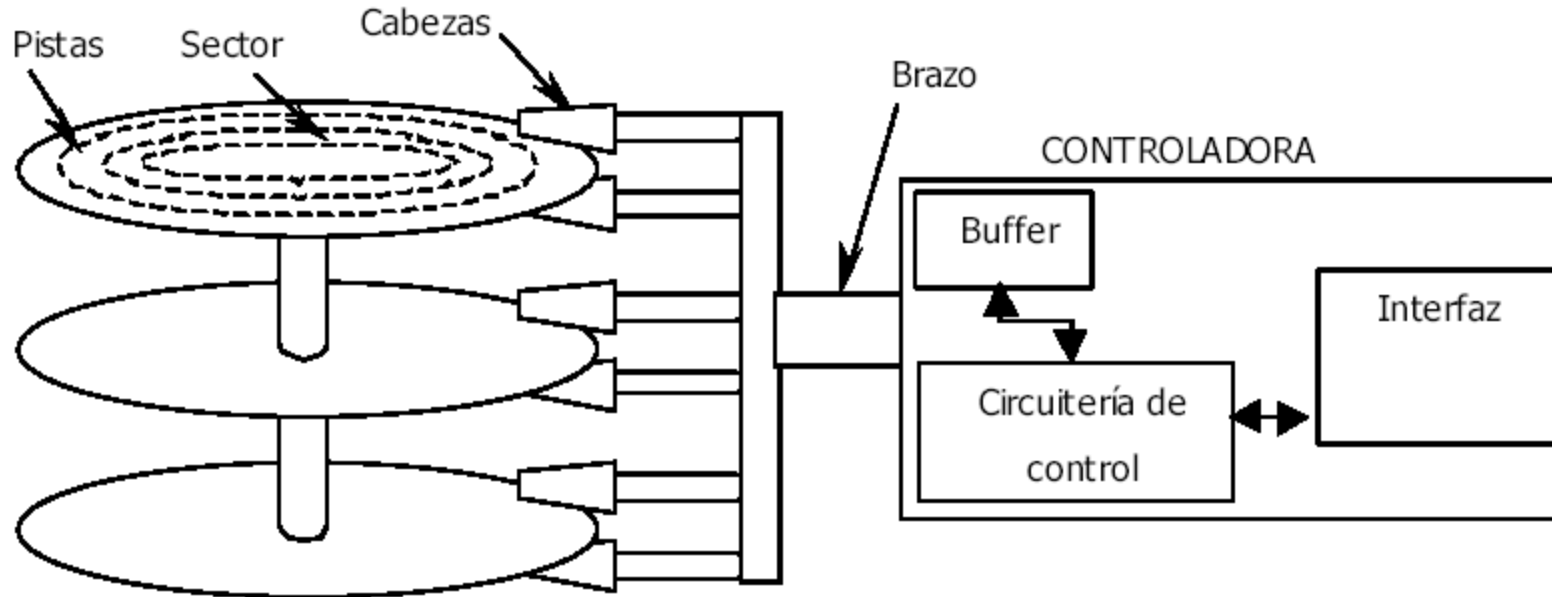
- ❖ Múltiples platos
- ❖ Una cabeza por cara
- ❖ Todas las cabezas se mueven solidariamente
- ❖ Pistas alineadas en cada plato forman **cilindros**
- ❖ Datos son almacenados por cilindros
- ❖ Reduce movimientos de cabezas
- ❖ Aumenta velocidad de respuesta



# Cilindros

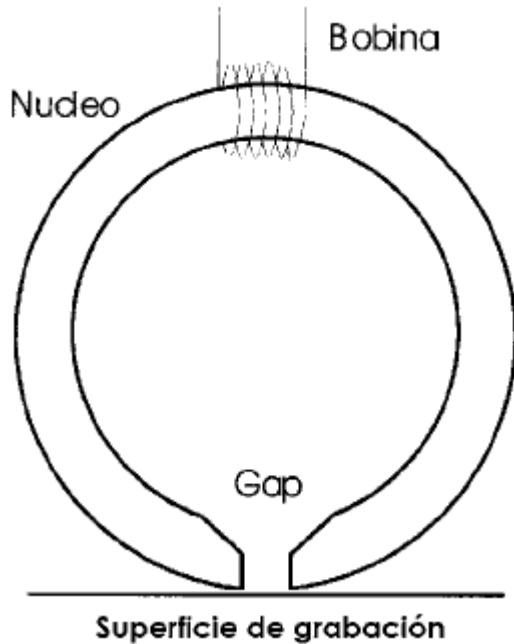


# Constitución de un disco





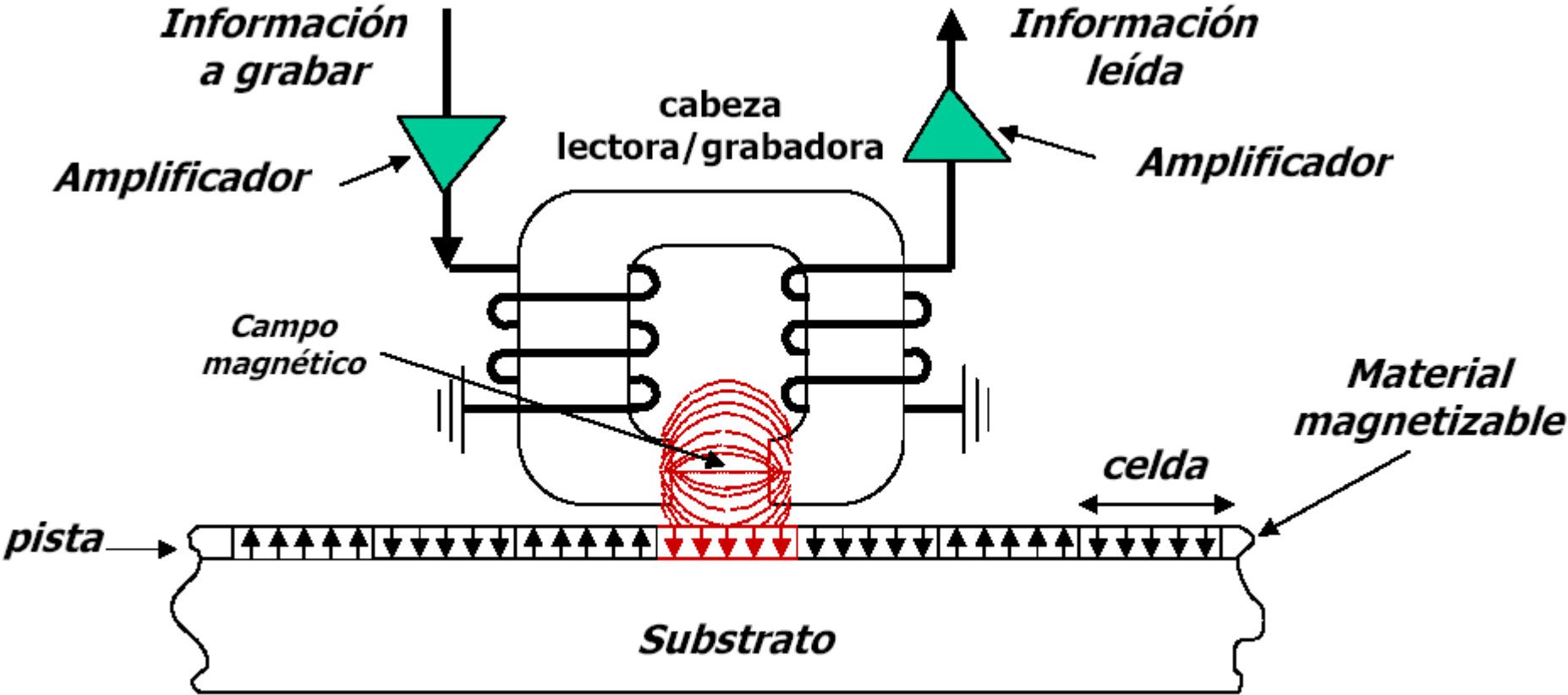
# Escritura/Lectura del disco magnetico



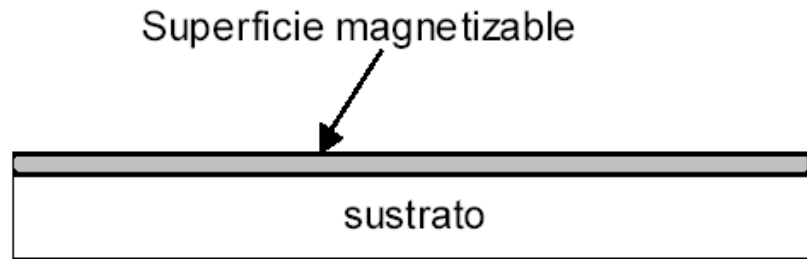
**Lectura:** El flujo magnético del disco se transforma en corriente eléctrica sobre la bobina del cabezal.

**Escritura:** La corriente que circula por la bobina provoca un flujo magnético sobre el núcleo y por lo tanto sobre la superficie del disco. El sentido de la corriente  $\leftrightarrow$  dirección de magnetización. El núcleo ferromagnético posee una separación llamada gap que permite que el flujo electromagnético se propague hasta la superficie de grabación. Podemos hacernos una idea de su tamaño:  $\text{Gap} \leq 40 \mu\text{m}$ .

# Escritura/Lectura del disco magnetico



# Medio de grabación



## Sustrato:

Suelen ser de aluminio ya que debe cumplir varios requisitos:

- Debe tener un bajo coeficiente de dilatación, pues la fuerza centrífuga de la rotación podría deformar el soporte dificultando la lectura y escritura.
- La superficie debe ser muy lisa.

## Superficie magnetizable:

Se trata de una capa de material magnético y por tanto susceptible a los campos electromagnéticos.

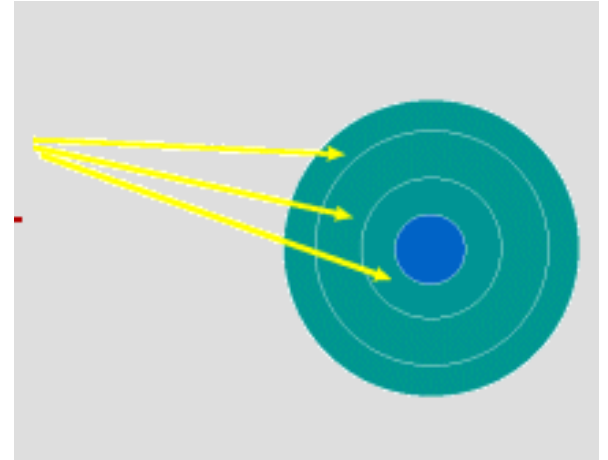
- En general se utilizan dos tipos de capas: óxido y película delgada.

# Pistas

## Pistas (Tracks)

Cada una de las líneas concéntricas en que se divide la superficie de un disco.

Una pista es un anillo circular sobre un lado del disco. Cada pista tiene un número. El diagrama muestra 3 pistas.



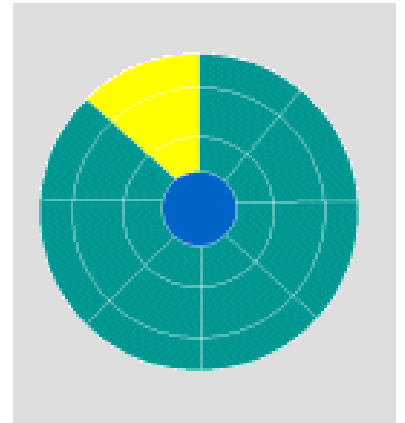
# Sectores

Se llama sector de un disco, una parte en forma de cuña del mismo. Cada sector está numerado.

En un disquete de 5 1/4" hay 40 pistas con 9 sectores cada una.

En un disquete de 3 1/2" hay 80 pistas con 9 sectores cada una.

De manera que, en un disquete de 3 1/2" hay el doble de lugares que en un disquete de 5 1/4" .



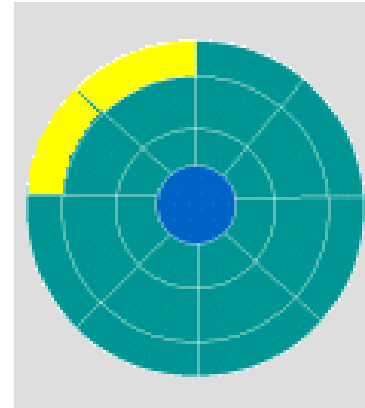
# Bloques

## Bloque (Clusters)

Un cluster es un conjunto de sectores de pista , desde 2 a 32 o más, dependiendo del esquema de formateo que se use.

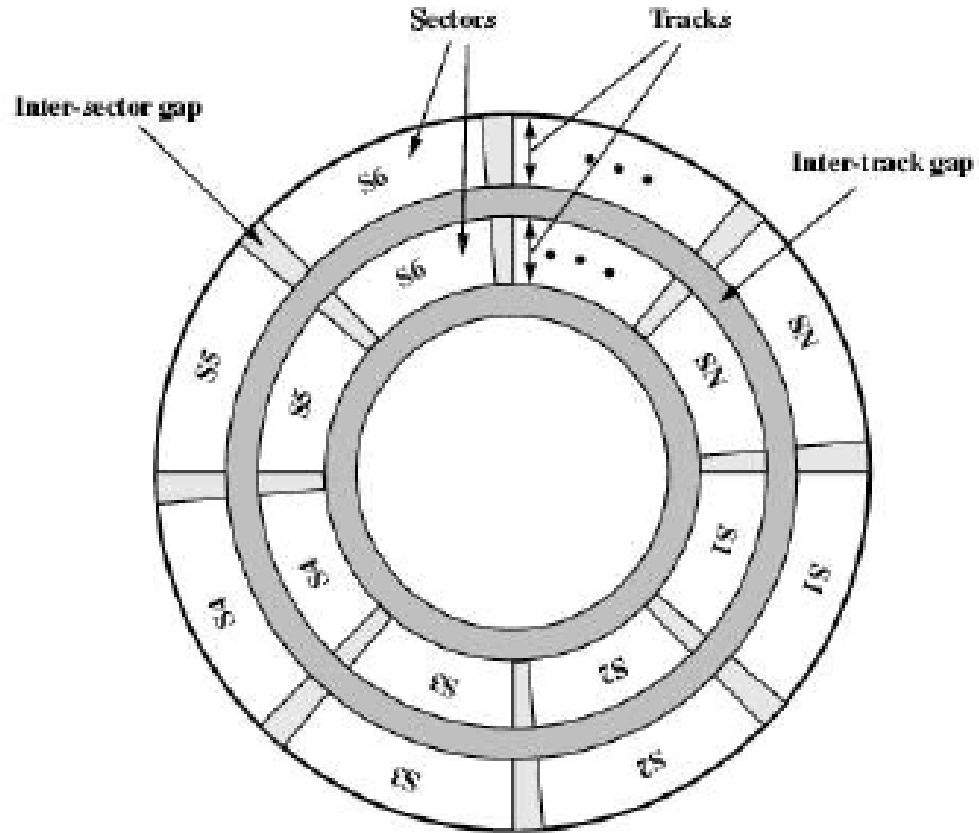
El esquema más común para PC determina la cantidad de sectores de pista por cada cluster sobre la base de la capacidad del disco.

Un disco rígido de 1.2 gigabytes tendrá el doble de clusters que un disco rígido de 500 MB.

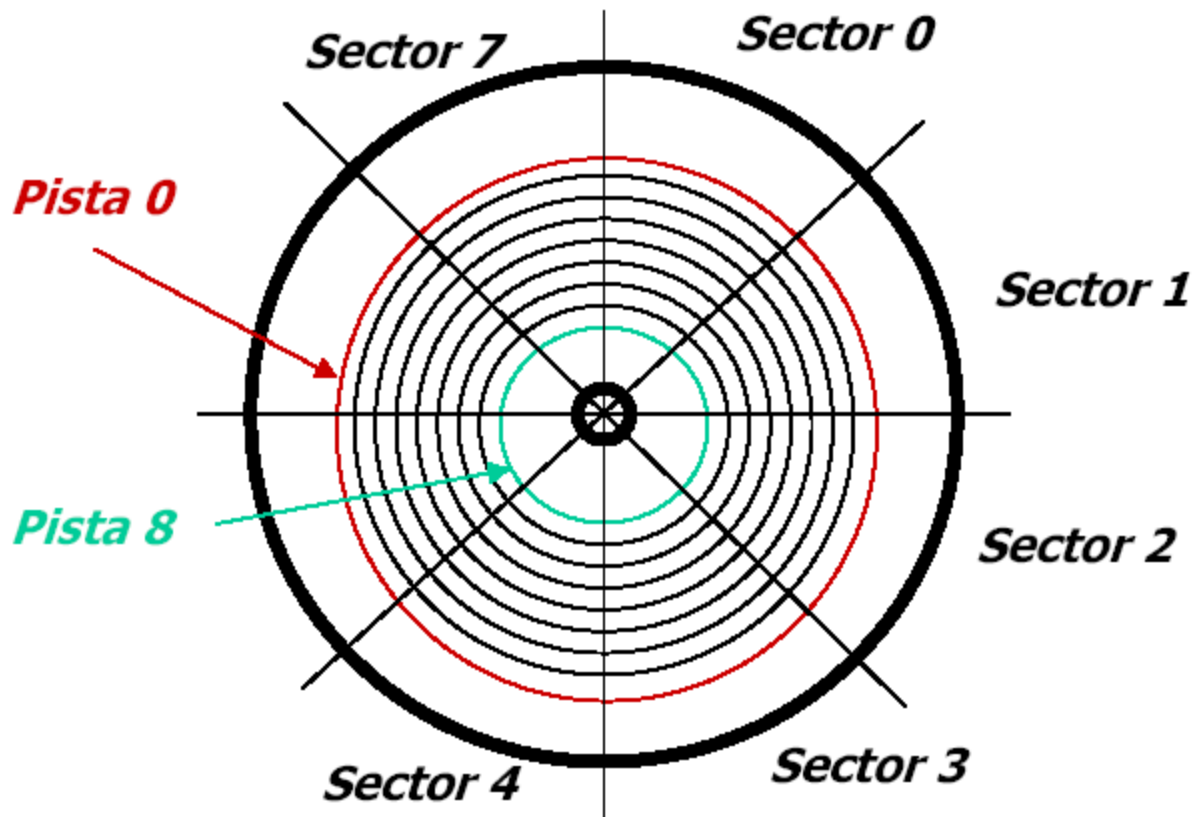


# Pista y Sector

- ❖ Un número entero de sectores se graban en una pista.
- ❖ El **sector** es la unidad de transferencia de/hacia el disco.



# Disposición de datos en el disco



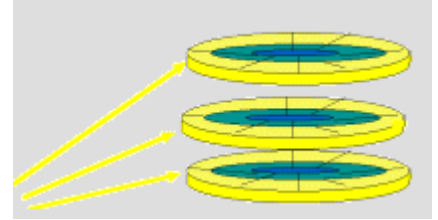


# Cilindros

## 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.



*Ejemplo:*

*Parámetros de un disquete de 3'5 pulgadas, 1'44MBytes:*

- 80 Cylinders, 2 Heads, 512 Bytes/Sector, 18 Sectors/Track*
- Multiplicando:  $80 \times 2 \times 18 \times 512 = 1.474.560$  bytes*

# Capacidad del disco

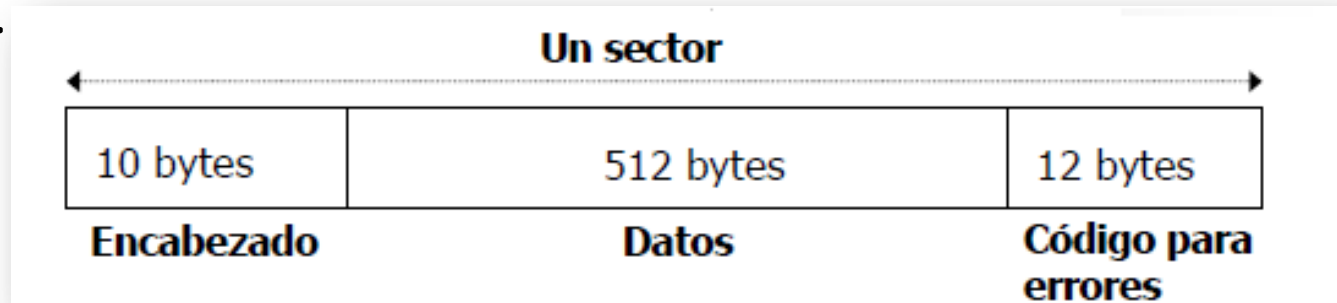
$$\text{Capacidad} = \frac{\text{bytes}}{\text{sector}} \times \frac{\text{sectores}}{\text{pista}} \times \frac{\text{pistas}}{\text{superficie}} \times \# \text{ de superficies}$$

- ❖ Se desperdicia espacio en pistas externas.
- ❖ Hoy en día se usan zonas para incrementar la capacidad
  - c/zona tiene fija la cantidad de bits/pista.
  - requieren circuitos más complejos.

# Sector típico

Sucesión o serie de bits divididos en campos

- Encabezado con información para sincronizar la lectura e identificar el sector.
- Datos con longitud en bytes expresada usualmente como potencia de 2.
- Código para errores con información para detectar y/o corregir posibles errores.



# Velocidad de giro del disco

- ❖ Rota a velocidad angular constante - CAV (**rpm**)
  - ✓ Un bit más cercano al centro gira más lento que un bit en la periferia.
  - ✓ Los bits (sectores) ocupan distinto espacio en diferentes pistas.
- ❖ En lectura o escritura se requiere sólo mover la cabeza hasta una pista y **esperar un sector**.
- ❖ Determina la velocidad de transferencia de datos (**bits/seg**).

# Velocidad de giro del disco

El procedimiento para leer/escribir datos en un disco duro tiene 4 pasos:

1. Búsqueda (seek)
2. Rotación (Latencia)
3. Ubicación
4. Transferencia de datos

Ver animación en el blog o desde:

<http://www.jegsworks.com/Lessons-sp/lesson6/lesson6-5.htm>

# Tiempos



## 1. Tiempo de seek (búsqueda)

✓ Mover al cilindro (o pista) correcto

## 2. Tiempo de latencia (por rotación)

✓ Esperar que el sector “pase” por debajo de la cabeza

## 3. Tiempo de Acceso: $T.\text{seek} + T.\text{latencia}$

## 4. Tiempo Total: $T.\text{de Acceso} + T.\text{de Transferencia de datos}$

# Formato



Define cantidad, tamaño y función de distintos campos en cada pista

- ❖• Hardware: tamaño de sector fijo por marcas físicas.
- ❖• Software: tamaño de sector determinado por S.O.

# Bibliografía



Capítulo 5: Memoria Externa. Stallings. 5ta Ed.

**Link sugerido**

<http://www.pctechguide.com/02Storage.htm>