

# Tema I: Ficheros

Santiago VELILLA

Dpto. de Informática e Ingeniería de Sistemas  
Universidad de Zaragoza



## Tema I: Ficheros

- 1 - Dispositivos de almacenamiento masivo: organización y gestión
- 2 - Conceptos fundamentales de ficheros. Organizaciones y acceso. Organizaciones secuenciales
- 3 - Indización. Ficheros con organización secuencial indexada
- 4 - Ficheros con organización dispersa

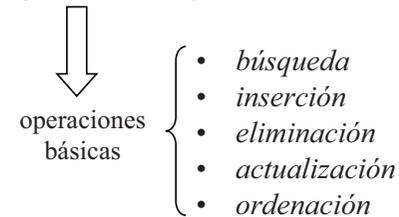
### Lección 1

#### Dispositivos de almacenamiento masivo: organización y gestión

- Introducción al problema de almacenamiento de información
- Cintas: organización y acceso. Parámetros que determinan su utilización
- Discos: organización y acceso. Parámetros que determinan su utilización
- Otros dispositivos de almacenamiento
- Jerarquía de almacenamiento: arquitectura global del sistema. Gestión de buffers

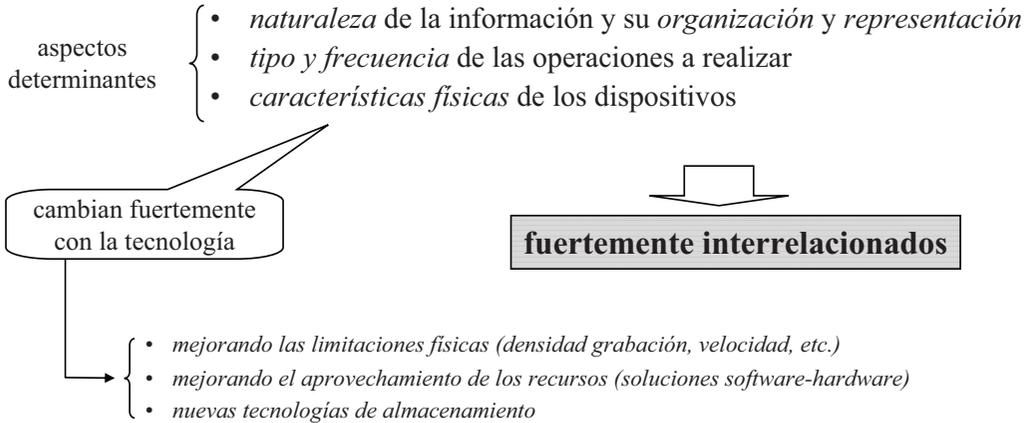
#### Introducción al problema de almacenamiento de información

- aspectos a considerar {
- Almacenamiento en memoria **no-volátil**
  - **gestión** de "grandes" cantidades de información

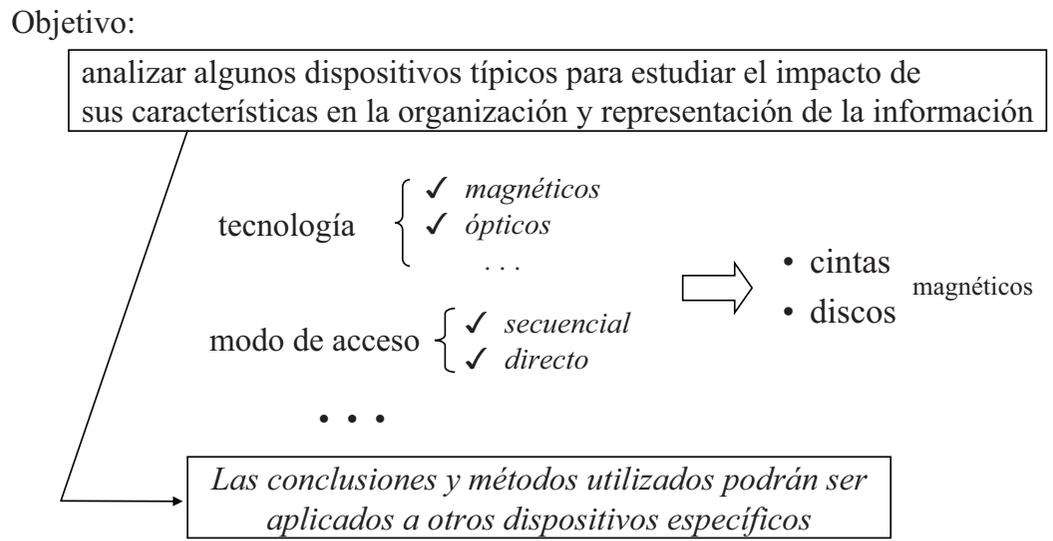


- características del sistema {
- ↓ coste de la realización
  - ↑ eficiencia de las operaciones
- ⇒ compromiso

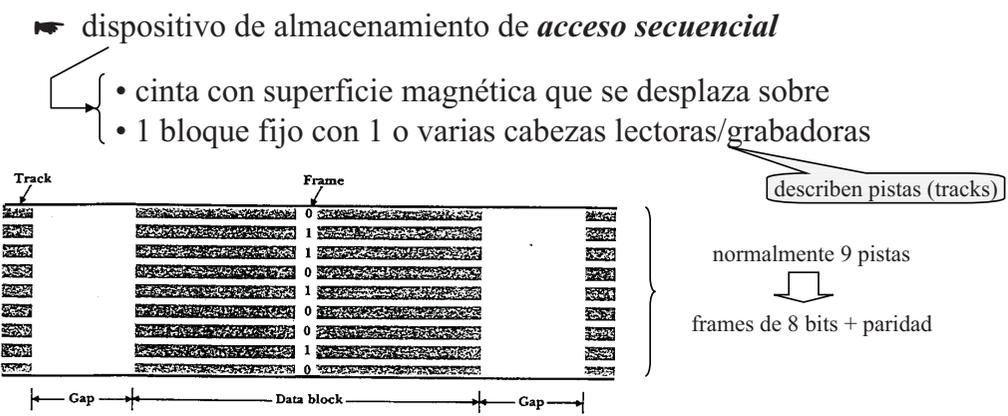
# Introducción al problema de almacenamiento de información



# Introducción al problema de almacenamiento de información



## Cintas: Organización

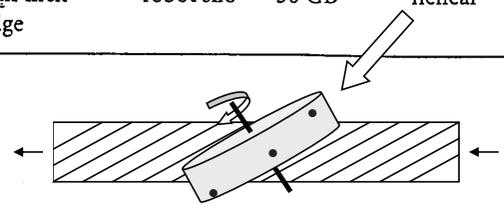


- ✓ información estructurada en bloques de datos, normalmente de igual tamaño, separados por
- ✓ bloques de control (gaps) de tamaño fijo → permiten parar la cinta y añadir redundancias

## Cintas: Organización

Comparación de algunos sistemas de cinta "actuales" (histórico)

Tape Model	Media Format	Loading	Capacity	Tracks	Transfer Rate
9-track	one-half inch reel	autoload	200 MB	9 linear	1 MB/sec
Digital linear tape	DLT cartridge	robot	35 GB	36 linear	5 MB/sec
HP Colorado T3000	one-quarter inch cartridge	manual	1.6 GB	helical	0.5 MB/sec
StorageTek Redwood	one-half inch cartridge	robot silo	50 GB	helical	10 MB/sec



## Cintas: Organización

### Parámetros físicos:

- *densidad de grabación nominal* (800 a 6250 bpi)
- *velocidad desplazamiento* (30 a 200 ips)
- **tasa de transferencia nominal** (bps) = densidad \* velocidad
- tamaño del "gap" entre bloques (0,3 a 0,75 i)
- tiempo de arranque/paro (4 a 20 mseg.)
- velocidad desplazamiento rápido

### Parámetros lógicos:

- *factor de bloque* (nº reg./bloque)
- densidad de grabación efectiva (bpi de datos)
- tasa de transferencia efectiva (bps de datos)

+  $\left\{ \begin{array}{l} \checkmark \text{ velocidad rebobinado} \\ \checkmark \text{ longitud cinta } (\approx 2400 \text{ pies} = 732\text{m}) \end{array} \right.$

1 inch = 2,54 cm

## Cintas: ejemplo de organización

Se desea almacenar la información de 1 millón de clientes en una cinta magnética de las siguientes características:

- densidad de grabación nominal: 6250 bpi
- velocidad desplazamiento: 200 ips
- tamaño del "gap" entre bloques: 0,3 i
- tiempo de arranque: 20 mseg.

Si para cada cliente son necesarios 100 bytes, estudiar las características de la implementación para factores de bloque 1 y 50

tasa de transferencia nominal

$$tt_{\text{Nom}} = dg_{\text{Nom}} (\text{bytes/inch}) \times v_c (\text{inch/seg}) = 6250 \times 200 = 1250 \text{ kb/seg}$$

ejemplo organización cinta: a) factor de bloque,  $f_{b1} = 1$

longitud de la cinta

$$l_{B1q} = \frac{f_{b1} (\text{reg./bloque}) \times O_{\text{reg}} (\text{bytes/reg})}{dg_{\text{Nom}} (\text{bytes/inch})} = \frac{1 \times 100}{6250} = 0,016 (\text{inch /bloque})$$

$$l_{\text{cinta}} = n_{B1q} (\text{bloque}) \times (l_{B1q} + l_{\text{gap}}) (\text{inch /bloque}) =$$

$$= 1000000 \times (0,016 + 0,3) = 316.000 \text{ inch} \approx 8\text{km.}$$

≈ 12 cintas

Densidad de grabación efectiva

$$dg_{\text{efc}} = \frac{f_{b1} (\text{reg./bloque}) \times O_{\text{reg}} (\text{bytes/reg})}{(l_{B1q} + l_{\text{gap}}) (\text{inch /bloque})} = \frac{1 \times 100}{0,316} = 316,5 \text{ bpi}$$

ejemplo organización cinta: a) factor de bloque,  $f_{b1} = 1$

tasa de transferencia efectiva

$$tt_{\text{efc}} = dg_{\text{efc}} (\text{bytes/inch}) \times v_c (\text{inch/seg}) = 316,5 \times 200 = 63,3 \text{ kb/seg}$$

rendimiento

$$\frac{l_{B1q}}{l_{B1q} + l_{\text{gap}}} = \frac{dg_{\text{efc}}}{dg_{\text{Nom}}} = \frac{tt_{\text{efc}}}{tt_{\text{Nom}}} = \frac{316,5}{6250} \approx 5\% \text{ aprovechamiento}$$

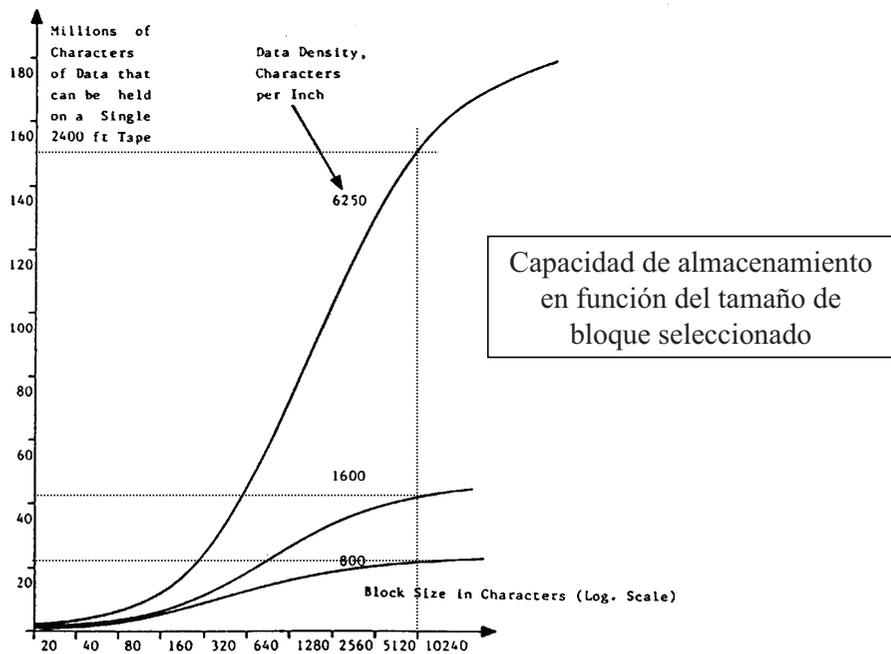
tiempo de lectura

$$t_{\text{lect}} = \frac{l_{\text{cinta}} (\text{inch})}{v_c (\text{inch/seg})} = \frac{316.000}{200} = 1580 \text{ seg.}$$

+ arranques:  $1000.000 \times 20 = 20.000 \text{ seg.}$

→ 21.580 seg.

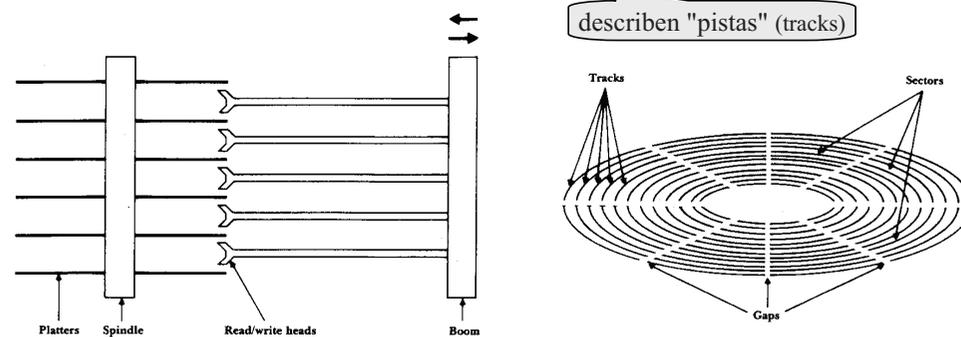




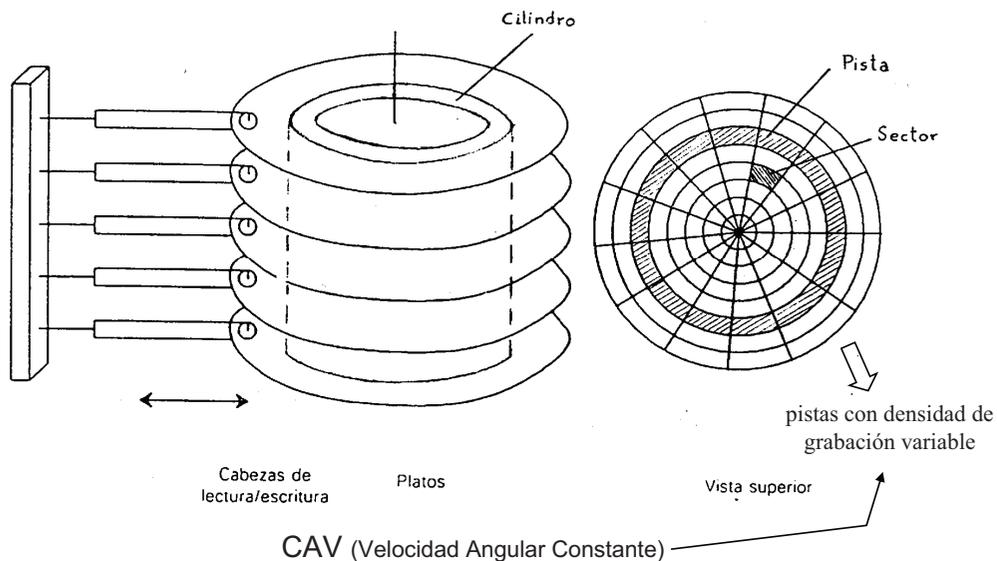
## Discos: Organización

dispositivos de almacenamiento de *acceso directo*

- bloque de 1 o varios platos (discos) con superficie magnética
- bloque de 1 o varias cabezas que se desplazan en sentido radial



## Discos: Organización



## Discos: Organización

- organización
- en **sectores**, o fragmentos de pista de igual cantidad de información
    - pb. fragmentación interna
    - gestión más simple
  - en **bloques**, o fragmentos de pista de cantidad de información variable
    - pb. fragmentación externa, mejor aprovechamiento
    - gestión más compleja

- direccionamiento
- método del **sector**  $\Rightarrow$  (nº de sector absoluto)  $\rightarrow$  **LBA**
  - 
  - método del **cilindro**  $\Rightarrow$  (nº cilindro, nº pista, nº sector)  $\rightarrow$  **CHS**



ejemplo organización disco: a) totalmente contigua

cálculo de la ocupación

$$512/100 = 5 \text{ clientes/sector} \Rightarrow O_{\text{tot}} = 12800/5 = 2560 \text{ sectores}$$

$$O_{\text{tot}} = 2560 \text{ sector} / 16 \text{ sector/pista} / 4 \text{ pista/cilindro} = 40 \text{ cilindros}$$

tiempo de lectura de un cilindro (una vez posicionada la cabeza)

$$t_{\text{Rrot}} = \frac{t_{\text{vuelta}}}{2} = \frac{60 \text{ (seg/min)} \times 1000 \text{ (mseg/seg)}}{3600 \text{ (vueltas/min)}} \times \frac{\text{vueltas}}{2} = 8,3 \text{ msec.}$$

$$t_{\text{trP}} = f_{\text{int}}(\text{vueltas}) \times \frac{60000 \text{ msec}}{3600 \text{ vueltas}} = 83,3 \text{ msec./pista} \quad \leftarrow 10\% = \frac{1}{2 \times f_{\text{int}}}$$

$$t_{\text{lcil}} = t_{\text{Rrot}} + 4 \times t_{\text{trP}} = 341,7 \text{ msec.}$$

ejemplo organización disco: a) totalmente contigua

tiempo de posicionamiento en el primer cilindro

$$t_{\text{dam}} = \left[ \frac{1}{3} \times 612 \right] \times v_{\text{dcab}} + t_{\text{arr}} = 204 \times 0,3 + 20 = 81,2 \text{ msec.}$$

tiempo de lectura de todos los datos

$$t_{\text{lfich}} = t_{\text{dam}} + n_{\text{cil}} \times t_{\text{lcil}} + (n_{\text{cil}} - 1) \times (v_{\text{dcab}} + t_{\text{arr}}) =$$

$$= 81,2 + 40 \times 341,7 + 39 \times (0,3 + 20) = 14,54 \text{ seg.}$$

tasa de transferencia efectiva

$$t_{\text{tefc}} = \frac{12800 \times 100}{14,54} = 88,0 \text{ kb/seg.} \approx 18\% \text{ de la nominal (491,5 kb/seg)}$$

factor intercalado  $\Rightarrow$  20%

ejemplo organización disco: b) sectores dispersos

$$\text{ocupación} = 2560 \text{ sectores}$$

tiempo de lectura de un sector

$$t_{\text{dam}} = \left[ \frac{1}{3} \times 612 \right] \times v_{\text{dcab}} + t_{\text{arr}} = 204 \times 0,3 + 20 = 81,2 \text{ msec.}$$

$$t_{\text{Rrot}} = \frac{t_{\text{vuelta}}}{2} = \frac{60000 \text{ msec}}{3600 \text{ vuelta}} \times \frac{1 \text{ vuelta}}{2} = 8,3 \text{ msec.}$$

$$t_{\text{trS}} = \frac{1 \text{ vuelta}}{16} \times \frac{60000 \text{ msec}}{3600 \text{ vuelta}} = 1,04 \text{ msec.}$$

$$t_{\text{lsec}} = t_{\text{dam}} + t_{\text{Rrot}} + t_{\text{trS}} = 90,58 \text{ msec.}$$

ejemplo organización disco: b) sectores dispersos

tiempo de lectura de todos los datos

$$t_{\text{lfich}} = 2560 \text{ sectores} \times t_{\text{lsec}} = 231,87 \text{ seg.}$$

16 veces más que contiguo

tasa de transferencia efectiva

$$t_{\text{tefc}} = \frac{12800 \times 100}{231,87} = 5,5 \text{ kb/seg.} \approx 1,1\% \text{ de la nominal (491,5 kb/seg)}$$

ejemplo organización disco: c) cluster dispersos

tamaño del cluster = 16 sectores



1 cluster = 1 pista  
ocupación = 160 cluster

tiempo de lectura de un cluster

$$t_{dam} = \left[ \frac{1}{3} \times 612 \right] \times v_{dcab} + t_{arr} = 204 \times 0,3 + 20 = 81,2 \text{ mseg.}$$

$$t_{Rrot} = \frac{t_{vuelta}}{2} = \frac{60000 \text{ mseg}}{3600 \text{ vuelta}} \times \frac{1 \text{ vuelta}}{2} = 8,3 \text{ mseg.}$$

$$t_{trP} = f_{int}^{(vueltas)} \times \frac{60000 \text{ mseg}}{3600 \text{ vueltas}} = 83,3 \text{ mseg.}$$

$$t_{lclus} = t_{dam} + t_{Rrot} + t_{trP} = 172,87 \text{ mseg.}$$

podría mejorarse el  $f_{int}$

ejemplo organización disco: c) cluster dispersos

tiempo de lectura de todos los datos

$$t_{lfich} = 160 \text{ cluster} \times t_{lclus} = 27,66 \text{ seg.}$$

8,4 veces menos que sector

tasa de transferencia efectiva

$$t_{tefc} = \frac{12800 \times 100}{27,66} = 46,3 \text{ kb/seg.} \approx 9,4\% \text{ de la nominal (491,5 kb/seg)}$$

factor intercalado  $\Rightarrow$  20%

Barracuda ATA Family Product Manual, Rev. B

Specification summary table

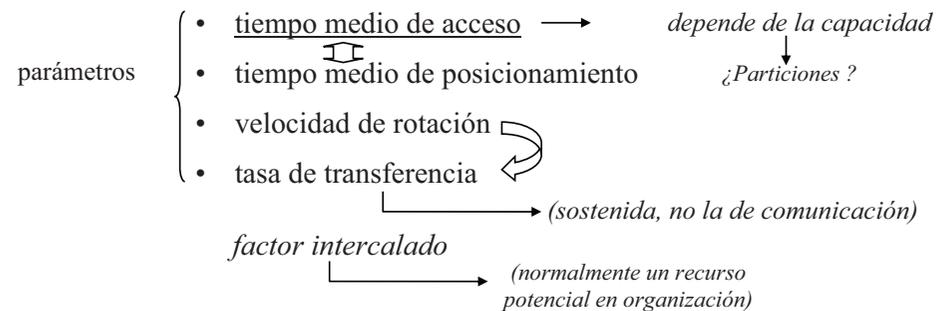
The specifications listed in this table are for quick reference. For details on specification measurement or definition, see the appropriate section of this manual.

Drive Specification	ST328040A ST327270A	ST320430A	ST313620A	ST310220A	ST36810A
Guaranteed Mbytes ( $\times 10^6$ bytes)	28,520 27,333	20,520	13,666	10,242	6,833
Guaranteed sectors	55,704,096 53,385,552	40,079,088	26,692,776	20,005,650	13,346,235
Bytes per sector	512				
Default sectors per track	63				
Default read/write heads	16	16	16	16	15
Default cylinders	16,383	16,383	16,383	16,383	14,123
Physical read/write heads	8	6	4	3	2
Discs	4	3	2	2	1
Recording density KBPI (bits/inch max)	305				
Track density TPI (tracks/inch)	16,420				
Areal density (Mbits/inch <sup>2</sup> max)	5,100				
Spindle speed (RPM)	7,200				
Internal data-transfer rate (Mbits/sec max)	323				
I/O data-transfer rate (Mbytes/sec max)	66.6				
ATA data-transfer- modes supported	PIO modes 0-4 Multiword DMA modes 0-2 Ultra DMA modes 0-4				
Cache buffer (Kbytes)	512				
Height (mm max)	26.1				

Drive Specification	ST328040A ST327270A	ST320430A	ST313620A	ST310220A	ST36810A
Width (mm max)	102				
Length (mm max)	147.0				
Weight (typical)	550 grams (1.2 lb)				
Average latency (msec)	4.16 msec				
Power-on to ready (sec typical)	12 sec				
Standby to ready (sec typical)	12 sec				
Startup current (typical) 12V (peak)	2.2 amps				
Track-to-track seek time (msec typical)	0.9				
Average seek time (read, msec typical)	8.0				7.6
Full-stroke seek time (msec typical)	14.5				14.2
Seek power (typical)	11.9 watts				
Read/Write power (typic.)	8.5 watts				7.5 watts
Idle mode (typical)	7.5 watts				6.8 watts
Standby mode (typical)	1.2 watts				
Sleep mode (typical)	1.0 watts				
Voltage tolerance (including noise)	5V $\pm$ 5% - 0.7 amps max 12V $\pm$ 10% - 2.4 amps max				
Ambient temperature	0 $^\circ$ to 55 $^\circ$ C (op.), -40 $^\circ$ to 70 $^\circ$ C (nonop.)				
Temperature gradient ( $^\circ$ C per hour max)	20 $^\circ$ C				
Relative humidity (op. and nonop.)	8% to 80% (op.) 5% to 90% (nonop.)				
Relative humidity gradient	30% per hour max				

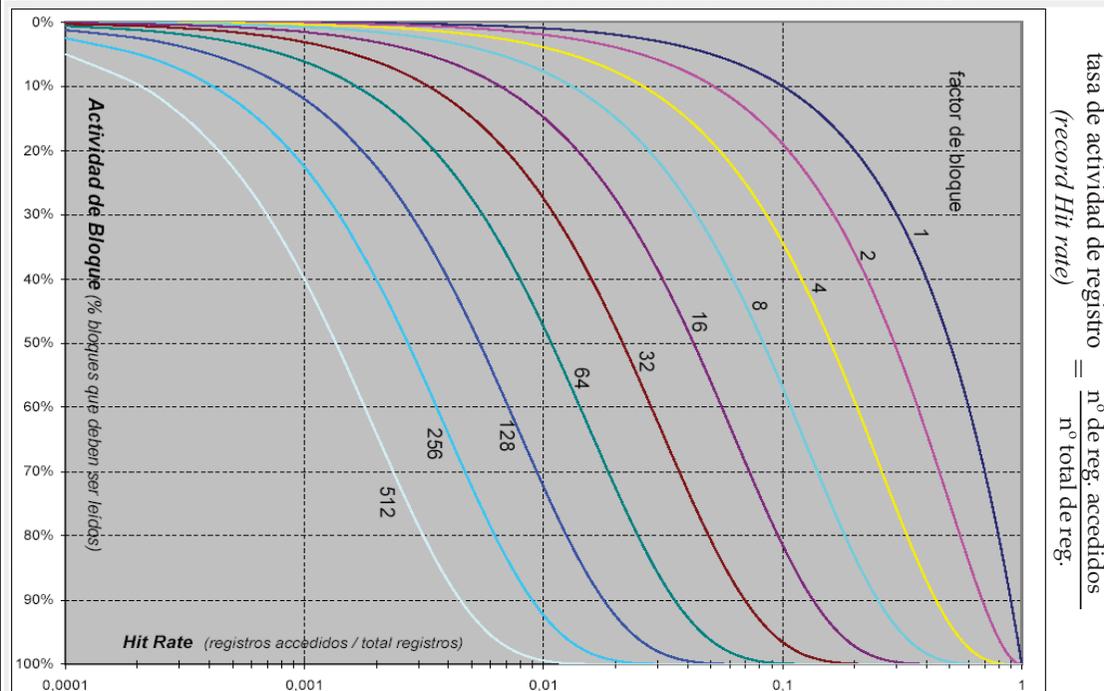
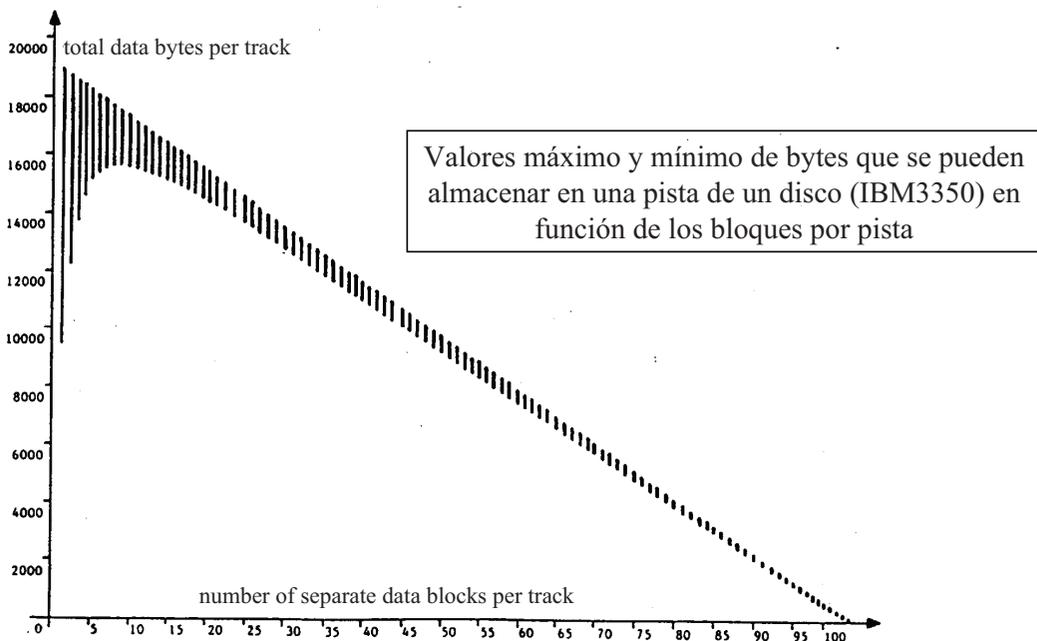
Drive Specification	ST328040A ST327270A	ST320430A	ST313620A	ST310220A	ST36810A
Wet bulb temperature (°C max)	29.4 (op.), 40.0 (nonop.)				
Altitude, operating	-122 m to 3,048 m (-400 ft to 10,000+ ft)				
Altitude (meters below mean sea level, max)	-122 m to 12,192 m (-400 ft to 10,000+ ft)				
Shock, operating (Gs max at 2 msec)	63				
Shock, nonoperating (Gs max at 2 msec)	300 Gs				
Vibration, operating	0.5 Gs (0 to peak, 5-300 Hz)				
Vibration, nonoperating	5 Gs (0 to peak, 5-300 Hz)				
Drive acoustics					
Sound power (Bels)					
Idle mode	3.5 (typical), 3.9 (max)				
Seek mode	4.0 (typical), 4.3 (max)				
Nonrecoverable read errors	1 per 10 <sup>14</sup> bits read				
Mean time between failures (power-on hours)	500,000				
Contact start-stop cycles (25°C, 40% relative humidity)	50,000				
Service life (years)	5				

## discos magnéticos: resumen

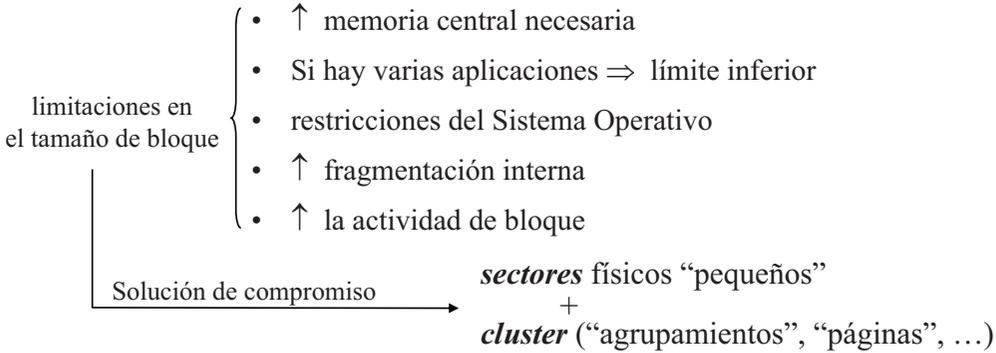


interesa procesamiento lo más secuencial posible ⇒ tamaño bloque grande

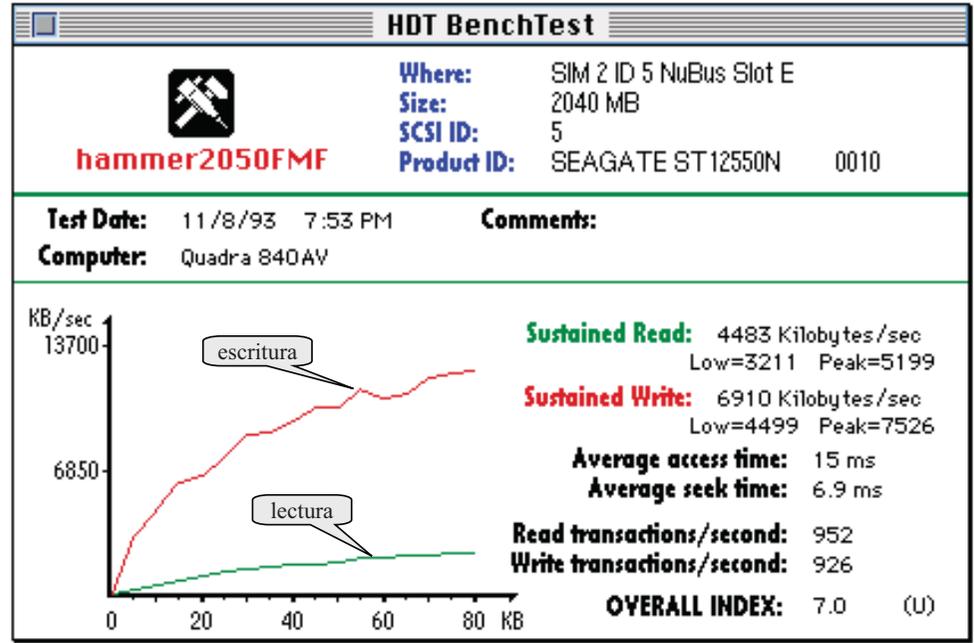
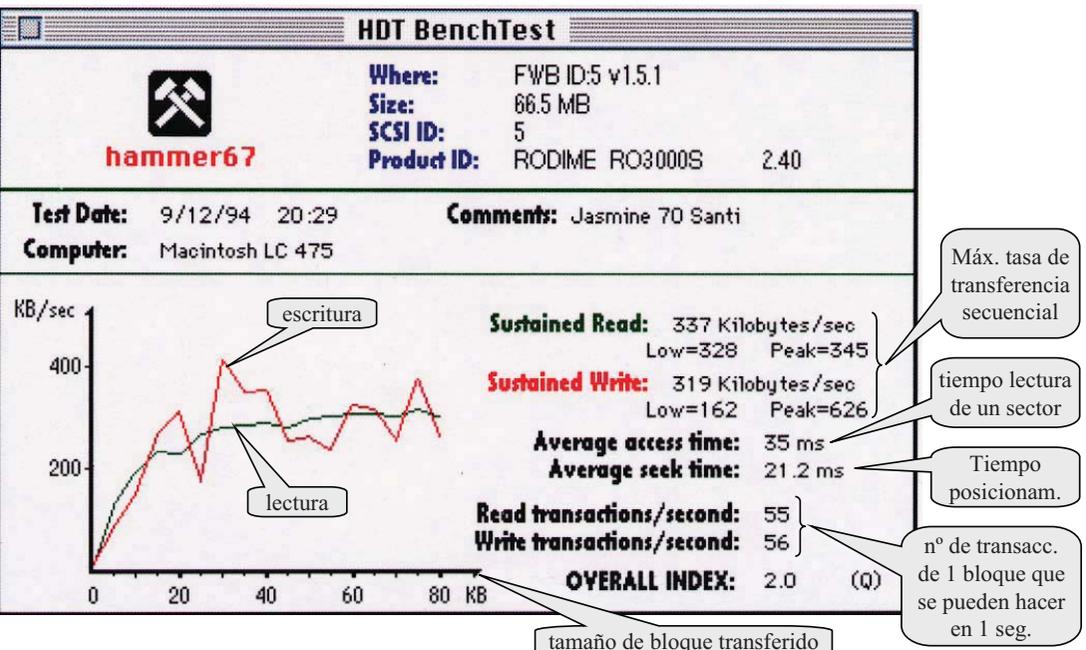
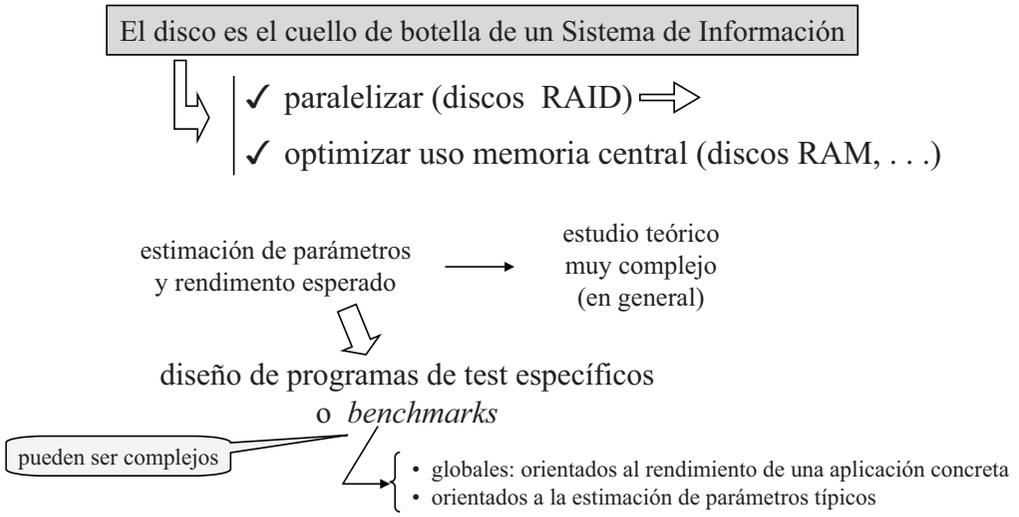
adecuar las organizaciones (si es posible)

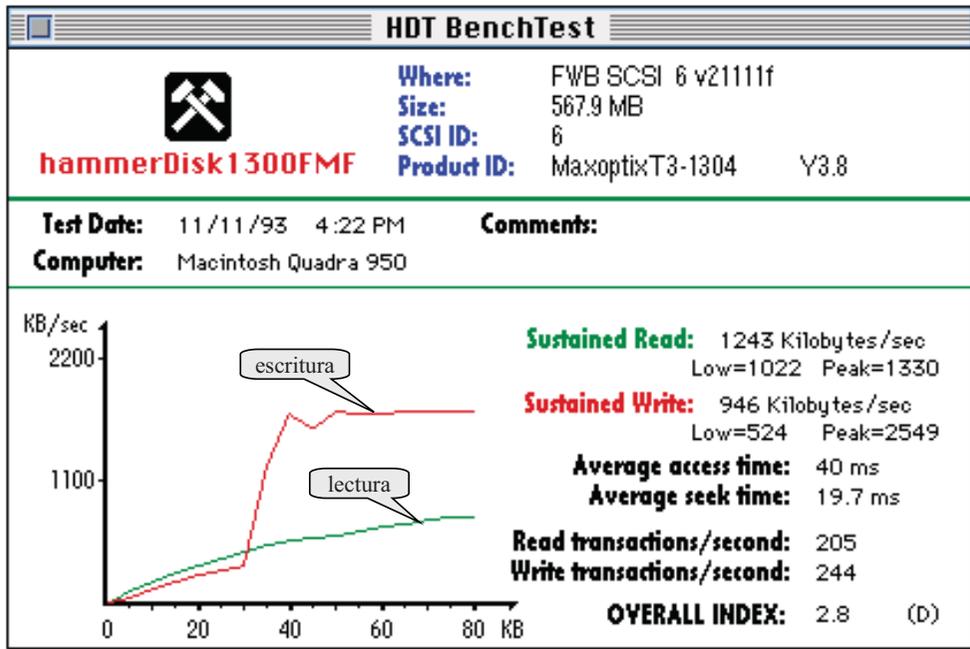


# discos magnéticos: resumen

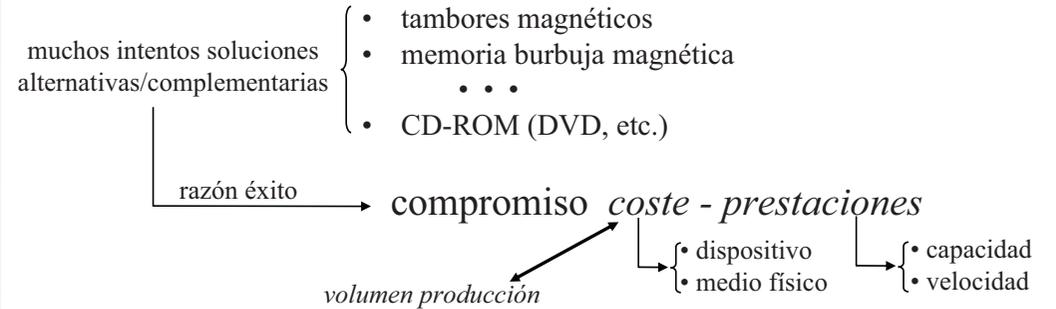


# discos magnéticos: evaluación de prestaciones





## Otros dispositivos de almacenamiento: CD-ROM



Ej. CD-ROM : objetivo ≈ 1 hora de música de calidad y relativamente “barato”

44.100 muestras/seg. ⇒ 88.200 words/seg. = **176.400 bytes/seg.**

## Otros dispositivos de almacenamiento: CD-ROM

aprovechar capacidad medio físico ⇒ densidad grabación constante ⇒ velocidad lineal constante

velocidad angular variable

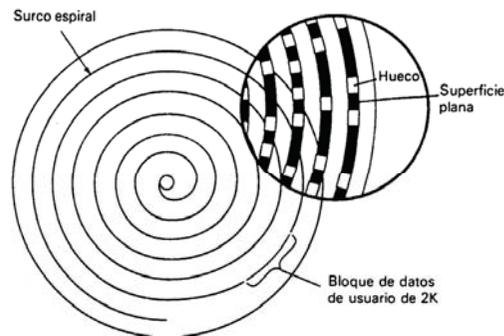
tecnología láser:

densidad grabación  
2352 bytes/inch

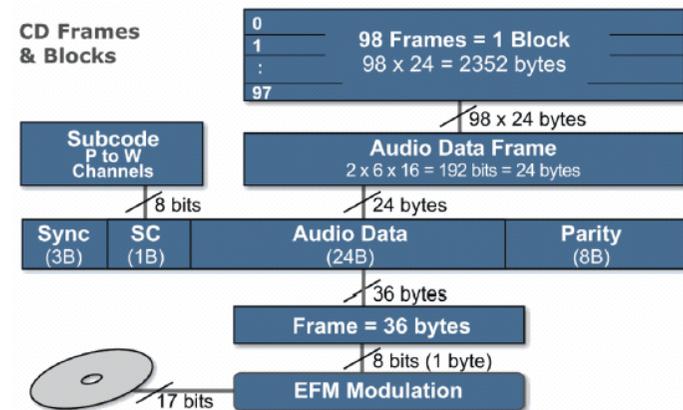
75 inch/seg.  
(176.400/2.352)

≈ 270.000 bloques  
(60\*60\*75)

≈ 635 Mb  
(270.000\*2352)



## Organización de un CD musical



# Otros dispositivos de almacenamiento: CD-ROM

escasa fiabilidad ⇒ codificación redundante (14 bits/byte + cod. Reed-Solomon)  
 ↳ 1 byte / 2 discos ⇒ *codificación redundante adicional para datos*  
 ↳ 1 byte no corregible / 20.000 discos

- Tiempo medio de posicionamiento 0,5 seg. Pb. sincronización de lectura
- Tasa transferencia efectiva 153,6 kb/seg. ≈ 87% de la nominal
- *sólo lectura*

12 bytes synch	4 bytes sector ID	2048 bytes user data	4 bytes error <i>detection</i>	8 bytes null	276 bytes error <i>correction</i>
----------------	-------------------	----------------------	--------------------------------	--------------	-----------------------------------

Estructura de un sector de un CR-ROM (2352 bytes)

# Jerarquía de almacenamiento: arquitectura global del sistema. Gestión de buffers

pb. Intercambio información CPU-dispositivo {
 

- organización de información en bloques
- diferente velocidad de procesamiento

memoria auxiliar intermedia (“tampón” o **buffer**)

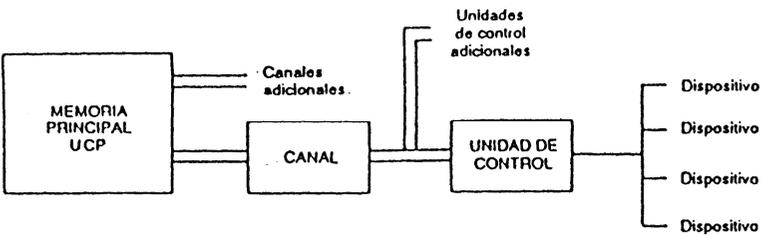
además → capacidad de anticipación ⇨ ↑ rendimiento global

tipos {
 

- hardware (ligados al dispositivo)
- software (ligados a procesos usuario y S.O.)

# Jerarquía de almacenamiento: arquitectura global del sistema. Gestión de buffers

gestión comunicación: procesador de E/S (canal)

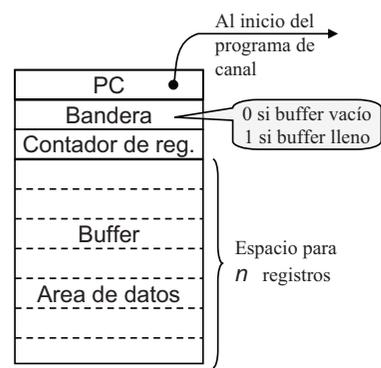


- ✓ como mínimo 1 buffer
- ✓ habitualmente 2 o más → {
  - simultanear E/S
  - mejorar prestaciones

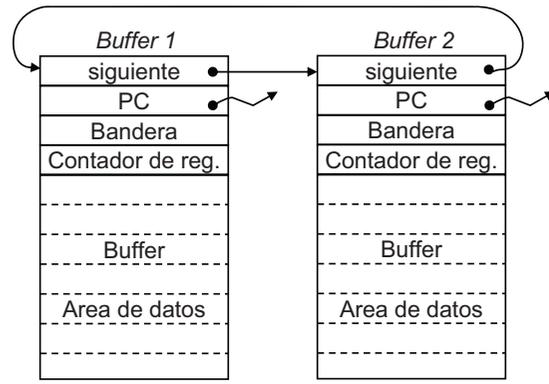
☛ a veces asociados a cada proceso

☛ también **pool de buffers** → política de reemplazo de páginas

# Gestión de buffers



estructura de un buffer

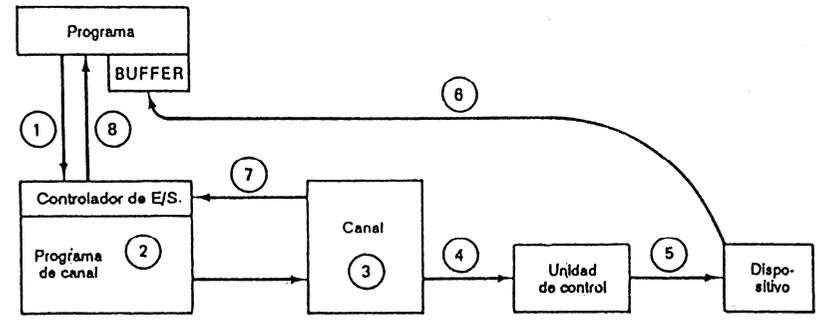


estructura de doble buffer

# Jerarquía de almacenamiento

Tipos de memoria	Dispositivos y medios	Tiempos de acceso (seg)	Capacidades (bytes)	Costo (centavos de dólar/bit)
<b>Primaria</b>				
Registros	Núcleo de ferrita y semiconductores	$10^{-8}$ – $10^{-6}$	$10^0$ – $10^7$	$10^0$ – $10^{-2}$
Memoria RAM				
Disco en RAM y caché disco				
<b>Secundaria</b>				
Acceso directo	Discos magnéticos	$10^{-2}$ – $10^{-1}$	$10^4$ – $10^9$	$10^{-2}$ – $10^{-5}$
Serie	Cintas y almacenamiento masivo	$10^1$ – $10^2$	$10^6$ – $10^{11}$	$10^{-5}$ – $10^{-7}$
<b>Fuera de línea</b>				
Respaldo y archivo	Discos magnéticos removibles, discos ópticos y cintas	$10^0$ – $10^2$	$10^4$ – $10^9$	$10^{-5}$ – $10^{-7}$

# arquitectura global del sistema.



Secuencia de eventos en el proceso de lectura de un fichero

# arquitectura global del sistema.

## Procesamiento de una lectura

Cuando un programa requiere una *LECTURA* del fichero, ocurre la siguiente secuencia de eventos:

1. El programa emite una instrucción de lectura, la cual interrumpe al controlador de E/S.
2. El controlador de E/S construye un programa de canal en memoria principal.
3. El programa de canal se ejecuta por el canal direccionado.
4. Señales apropiadas se transmiten a la unidad de control direccionada.
5. Estas señales se interpretan por la unidad de control y se usan para controlar las operaciones del dispositivo para leer los datos solicitados.
6. Los datos solicitados fluyen del dispositivo, a lo largo de la ruta hasta llegar al área de almacenamiento temporal del fichero en el espacio de memoria del programa.
7. Una interrupción se emite por el canal para señalar la continuación de la ejecución del programa.
8. El control regresa al programa.