

I.I.S. “Benvenuto Cellini”

Corso di formazione tecnica

Le memorie di massa: hard disk

Prof. Alessandro Pinto

v.2009



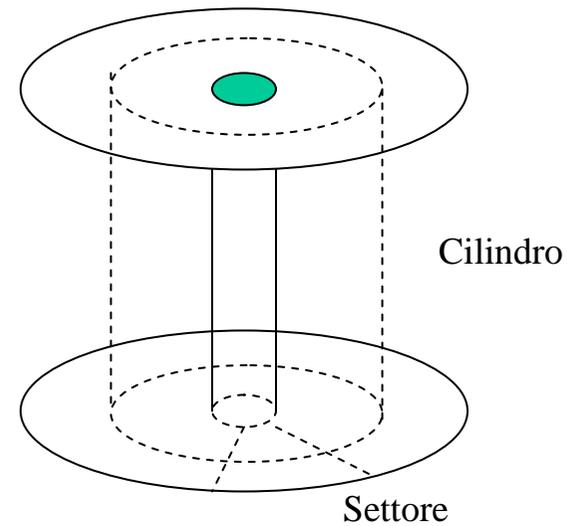
1956

Capacità: 5MB

Dimensioni: (1,5x1,7x0,7) m

Dischi:50 (24"), 2 testine

Tempo di accesso: 600ms



2008

Capacità: 500GB

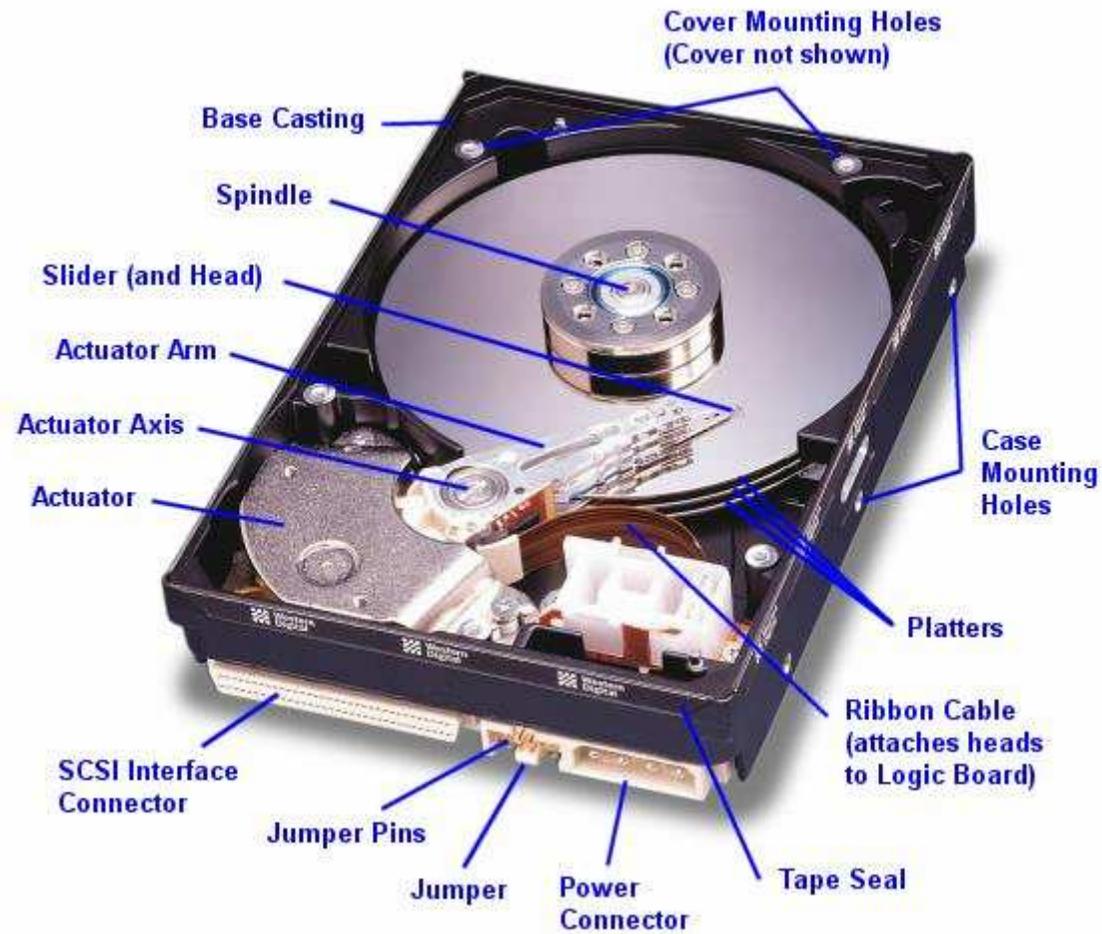
Dimensioni: (101x25x146) mm

Dischi:5 (3"1/2) , 10 testine

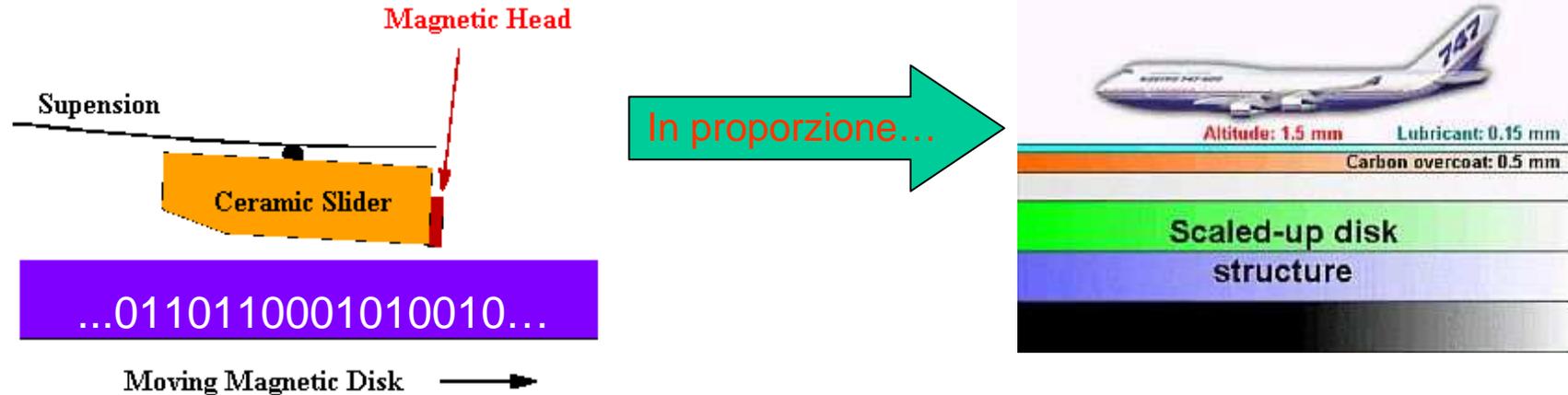
Tempo di accesso: 8,5ms



L'hard disk a livello fisico



L'hard disk a livello fisico



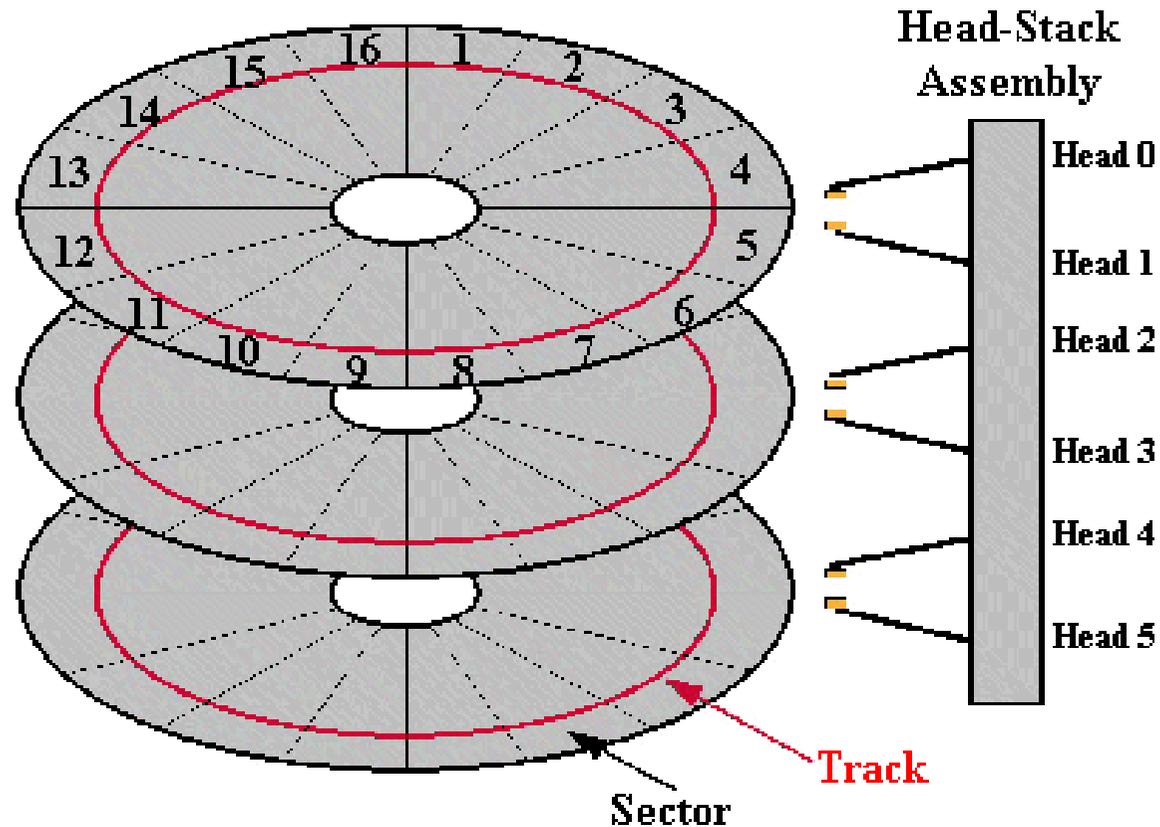
La testina “vola” su un cuscino d’aria senza mai toccare la superficie del disco ad una velocità che, alla periferia del disco, è intorno ai 100km/h.

In volo la distanza dalla superficie del disco è di qualche decina di nanometri (capello umano ~ 100um)

Quando il contatto avviene al di fuori della *landing zone* è un evento catastrofico, che prende il nome di *head crash*, che comporta perdita dei dati e danneggiamento fisico del disco.

Organizzazione del disco

Drive Physical and Logical Organization



Un insieme verticale di tracce (track) costituisce un *cilindro*. Cilindro, testina, settore identificano univocamente una porzione di disco (blocco). Più blocchi costituiscono un *cluster*.

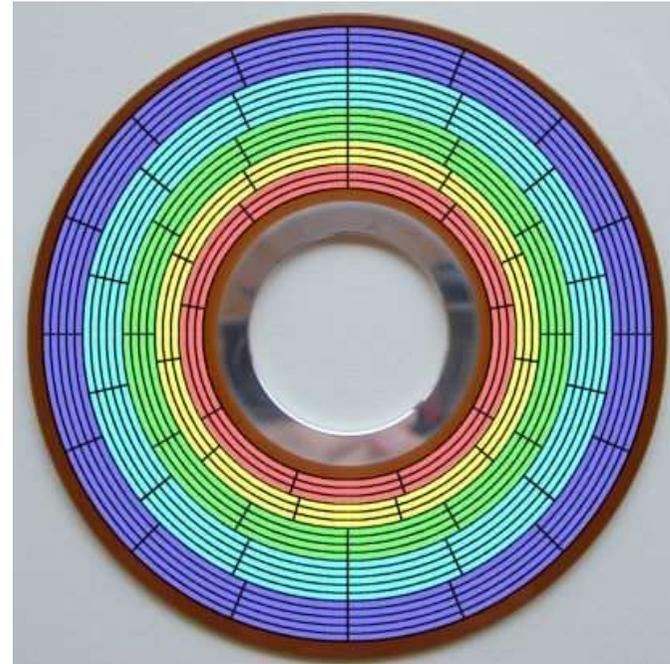
Organizzazione del disco



Ogni traccia contiene lo stesso numero di settori

- Ampiezza dei cluster variabile dal centro alla periferia
- densità di scrittura variabile (*write precompensation*)

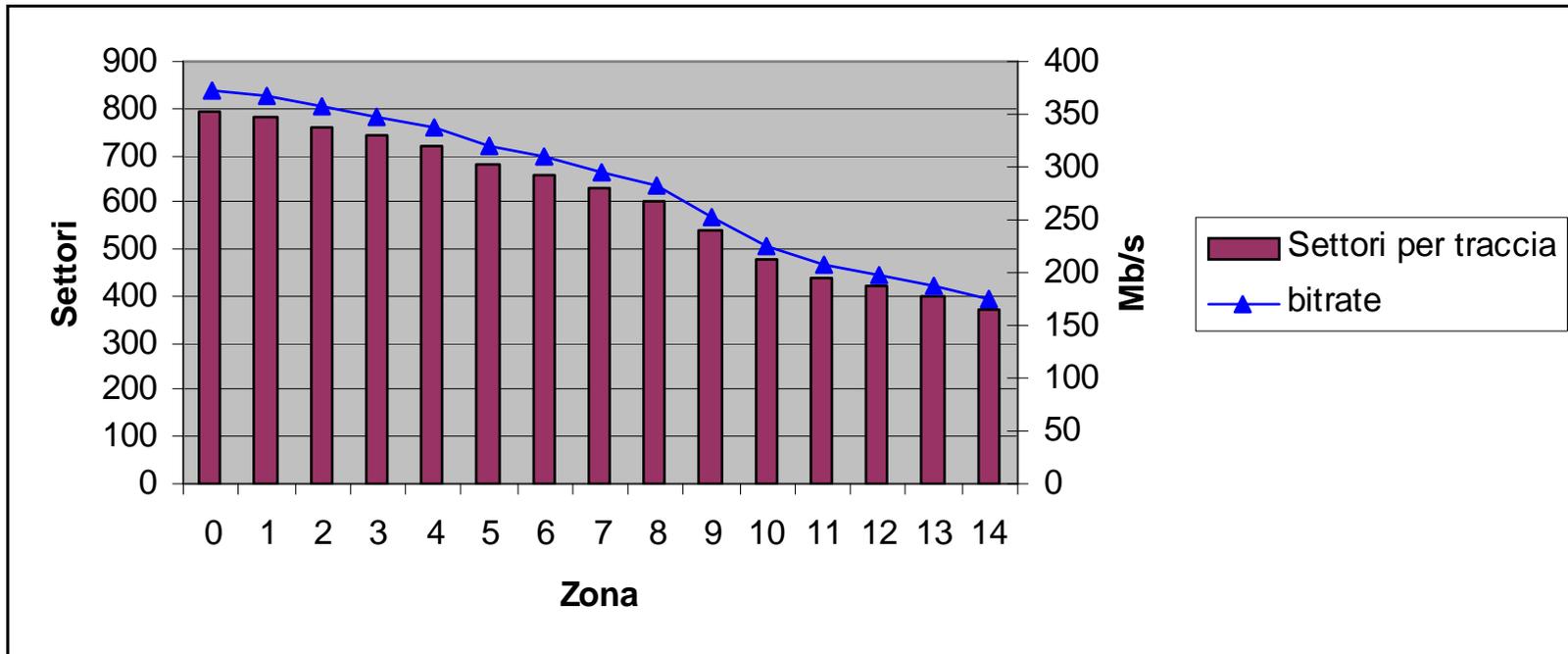
Tracce esterne sottoutilizzate
Velocità di lettura costante



Numero di settori variabile a seconda della zona (*zoned bit recording*)

- Ampiezza dei cluster mediamente *costante* dal centro alla periferia
- densità lineare di scrittura costante

Ottimizzazione dello spazio occupato
Velocità di lettura variabile



Description	HDP725016GLxxxx HCP725016GLxxxx
Physical Layout	
Label capacity (GB)	160
Bytes per Sector	512
Number of Heads	2
Number of Disks	1
Logical Layout²	
Number of Heads	16
Number of Sectors/ Track	63
Number of Cylinders ¹	16,383
Number of Sectors	312,581,808
Total Logical Data Bytes	160,041,885,696

Preparazione del disco

Formattazione a basso livello

Crea la struttura fisica sull'hard disk in termini di tracce e settori, definendo le informazioni di controllo che relative alla loro posizione.

Per i dischi attuali la formattazione di basso livello è effettuata (ed effettuabile) solo in fabbrica. Non è mai richiesto all'utente di provvedere in seguito.

Utility in grado di formattare a basso livello in realtà effettuano solo una inizializzazione (riempimento dei dati con pattern predefiniti es. zero). Nello sventurato caso che realmente effettuino una LLF, è molto probabile che il disco **venga danneggiato irreparabilmente** (perdita delle informazioni per il posizionamento del servo)

Partizionamento

Suddivide il disco fisico in unità logiche, ognuna delle quali può avere un proprio file system

Ottimizzazione della dimensione del cluster (*slack reducing*)

Obbligatoria se la dimensione del disco è superiore rispetto al limite imposto dal tipo di file system adottato

Separazione tra sistema operativo e dati

Formattazione ad alto livello

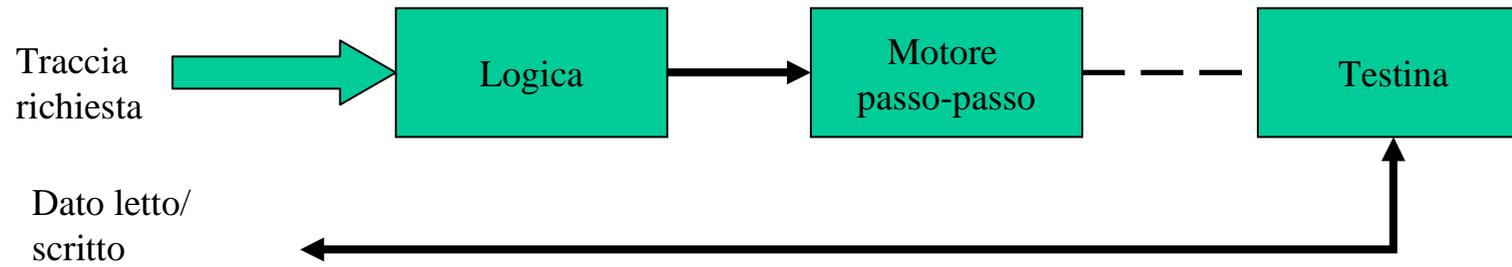
Creazione della struttura del file system (boot record, tavola di allocazione, root directory...)

Scansione della superficie per la ricerca di difetti

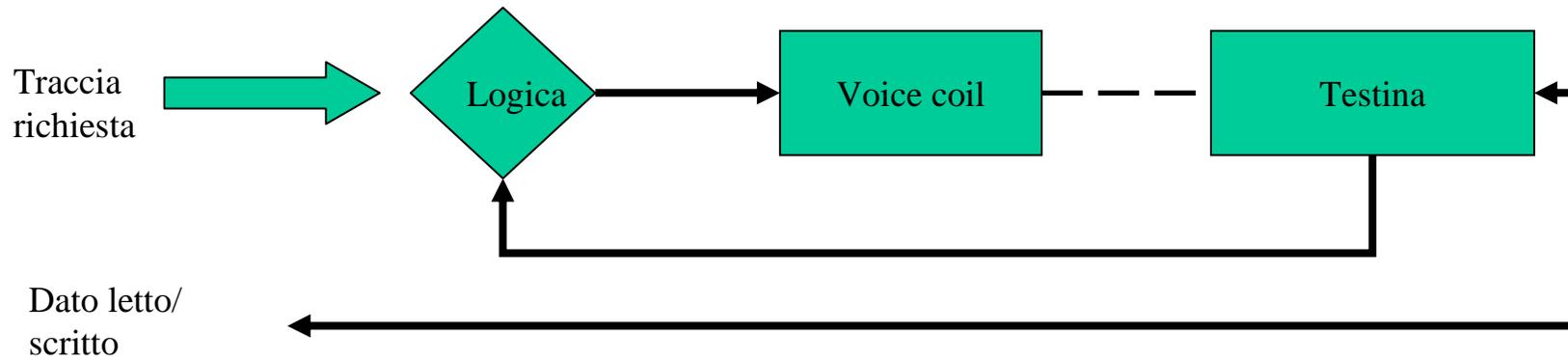
Eventuale operazione di *overwrite* incondizionato per la cancellazione dei dati precedentemente memorizzati

Movimento della testina

Stepper motor



Servocomando "Voice Coil"



Movimento della testina

Stepper motor

- Sistema ad anello aperto: la posizione della testina è determinata a priori in base (numero di passi di avanzamento) in funzione della geometria
- Avanzamento a passi discreti
- Non è possibile correggere errori di allineamento dovuti a variazioni meccaniche:
 - Temperatura
 - Posizione
- Accesso ai dati lento
- Necessita di procedure di parcheggio della testina

Voice Coil

- Sistema ad anello chiuso: la testina avanza fino a quando non trova la traccia richiesta
- Avanzamento continuo
- Correzione della posizione intrinseca (non risente di temperatura e posizione)
- Velocità
- Parcheggio automatico
- Le tracce sono definite in fabbrica con la formattazione a basso livello inserendo sui piatti stessi (embedded servo) o su un piatto dedicato (dedicated servo) le informazioni di posizionamento
- L'utente non deve effettuare formattazione a basso livello

Indirizzamento C-H-S

Standard

L'accesso ai dati memorizzati nel disco rigido avviene in base alla loro posizione fisica, quindi con una terna che identifica la traccia (cilindro – C), il piatto (testina – H) e il settore (S) occupato.

In realtà, con i dischi attuali, si tratta comunque di una posizione logica: la traduzione logico/fisico è effettuata dal controller integrato

BIOS (L-CHS):

C: 10 bit (1024 cilindri)

H: 8 bit (256 testine)

S: 6 bit (63 settori)



C x H x S x 512 byte:
8 GB (7.8GiB) Indirizzabili

IDE (P-CHS):

C: 16 bit (65536 cilindri)

H: 4 bit (16 testine)

S: 8 bit (256 settori)



C x H x S x 512 byte:
137GB (128GiB) Indirizzabili



BIOS+IDE:

C: 10 bit (1024 cilindri)

H: 4 bit (16 testine)

S: 6 bit (63 settori)



C x H x S x 512 byte:
528 MB (504MiB)Indirizzabili

Indirizzamento C-H-S

Esteso (o *large mode*)

Per superare il limite dei 528MB dovuto alla incompatibilità tra routine del bios e standard ATA/IDE, viene introdotto il modo ECHS o Large Mode. Consiste in una traduzione, a livello di bios, tra il formato presentato dal disco (secondo lo standard IDE/ATA) ed una geometria equivalente riconosciuta dal bios stesso. Questo è possibile poiché un formato prevede più cilindri dell'altro, ma meno testine.

La traduzione consiste nel dividere il numero di cilindri per un numero fisso e moltiplicare, per lo stesso numero fisso, il numero di testine. Abbiamo due terne CHS compatibili per i rispettivi standard che rappresentano la stessa capacità del disco.



	Cilindri	Testine	Settori	Capacità
Limite IDE/ATA	65536	16	256	128GB
Geom.logica	6136	16	63	3.16GB
Traslazione	÷8	x8		
Geom.traslata	767	128	63	3.16GB
Limite Bios	1024	256	63	8GB

N.B.: in alcuni BIOS capitava che, nonostante la traslazione, non venissero riconosciuti più di 2GB. Questo era dovuto ad una errata implementazione che non riconosceva più di 12bit per la codifica dei cilindri (limite dei 2GB o del 4096° cilindro)

Indirizzamento LBA (Logical Block Addressing)

I singoli settori sono identificati da un indice sequenziale a partire da 0 (settore 1, testina 0, cilindro 0) fino all'ultimo settore del disco.

Nel modo LBA quindi è sufficiente un solo numero per individuare il blocco da leggere/scrivere, contrariamente dal modo CHS che richiede una terna di numeri.

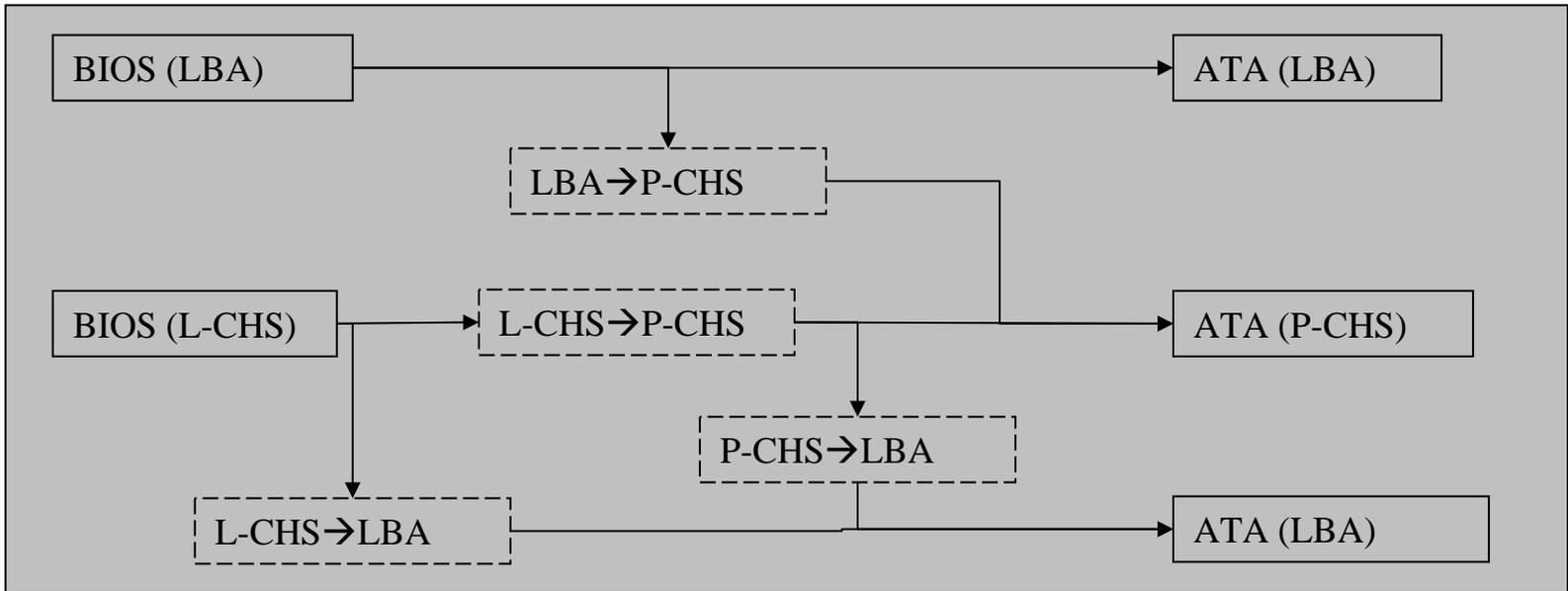
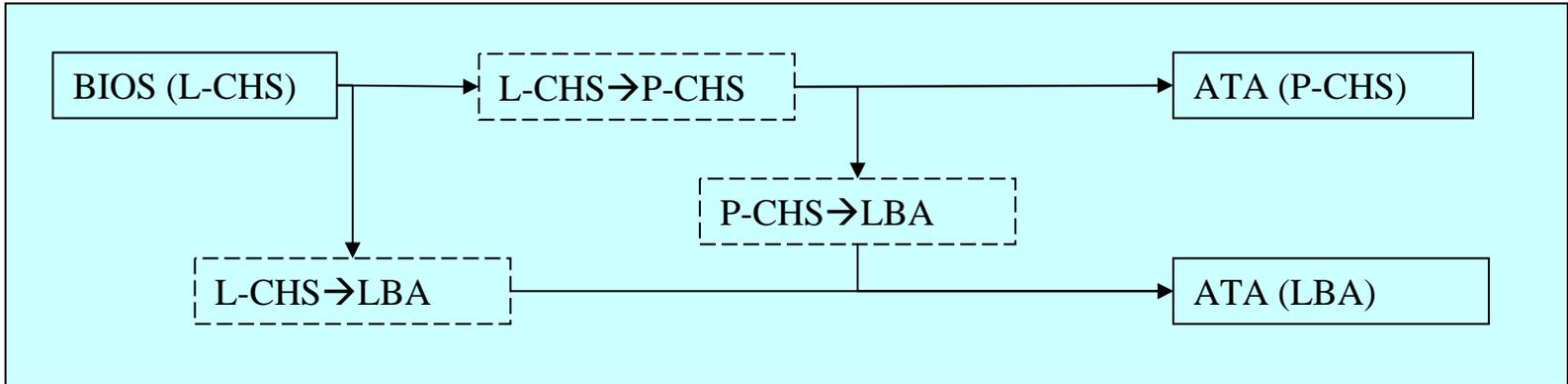
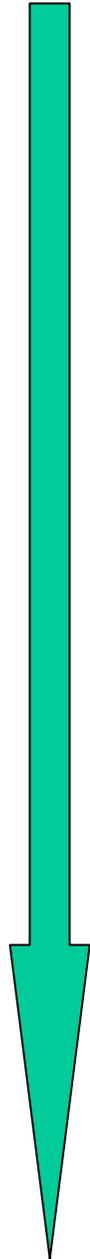
$$\text{LBA} = \text{C} \times \text{Num_Head} \times \text{Num_Sec} + \text{H} \times \text{Num_Sec} + (\text{S} - 1)$$

Nella prima implementazione (ATA-2) era previsto, come per il CHS, una lunghezza di 28 bit per l'indice. Questo limita la dimensione massima a 137GB.

La successiva estensione a 48 bit (ATA-6) ha portato il limite a 144PB (144000TB)

Per l'impiego "con successo" di dischi superiori a 137GB di capacità occorre che siano verificate le seguenti condizioni:

- BIOS compatibile LBA48
- controller compatibile LBA48
- Supporto 48bit da parte del sistema operativo (Win2k SP3 e succ., XP SP1 e succ,)



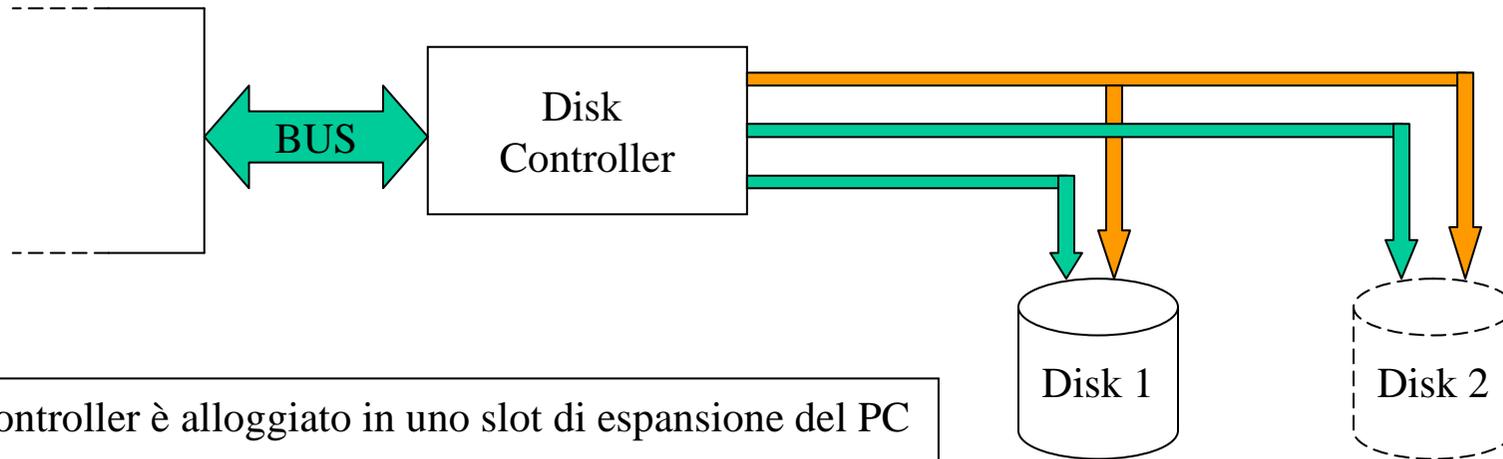
Le interfacce disco - computer

- ST506 / ST412 (obsoleta)
- ESDI (obsoleta)
- SCSI
- ATA/IDE (ora ridefinita PATA)
- Serial ATA (SATA)

ST506/ST412

Prima implementazione di unità disco rigido sui PC-IBM (anni '80)

(Talvolta indicata anche come “MFM” o “RLL” dai tipi di codifica impiegati per la scrittura sul disco)



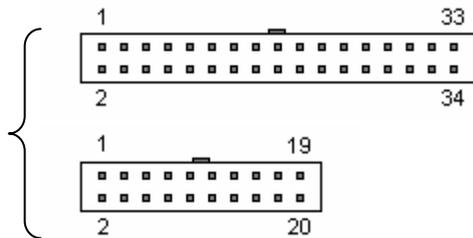
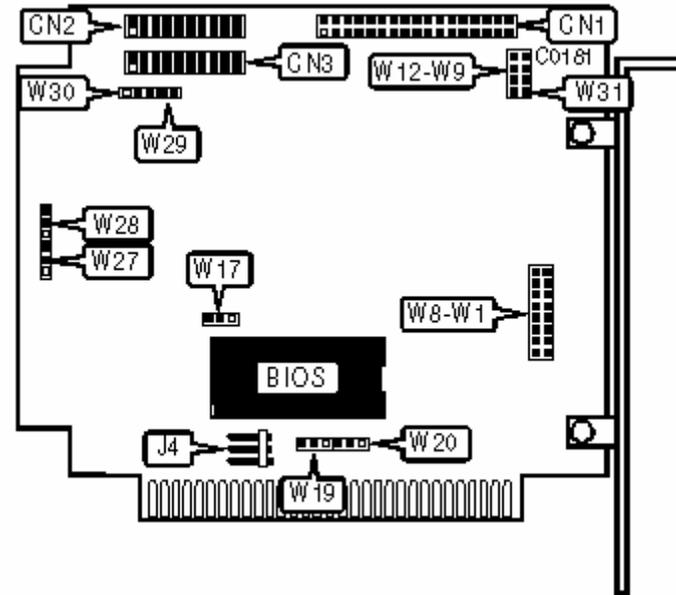
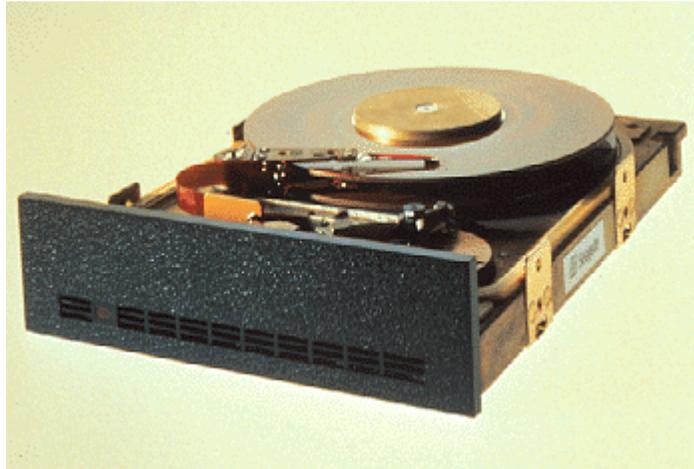
Il controller è alloggiato in uno slot di espansione del PC

La funzione del controller è quella di gestire il posizionamento della testina in corrispondenza di traccia e settore da leggere/scrivere.

Il disco non dispone di nessuna altra logica di controllo

Limiti:

- Gestione complessa (l'utente deve impostare la geometria fisica del disco e altri parametri)
- Affidabilità (Corruzione dei dati dovuti alla lunghezza del collegamento tra testine e controller)
- Compatibilità (dischi e controller costruiti in tempi diversi da produttori diversi)
- Velocità



Controllo (posizionamento testina)

Dati (trasferimento dei dati)

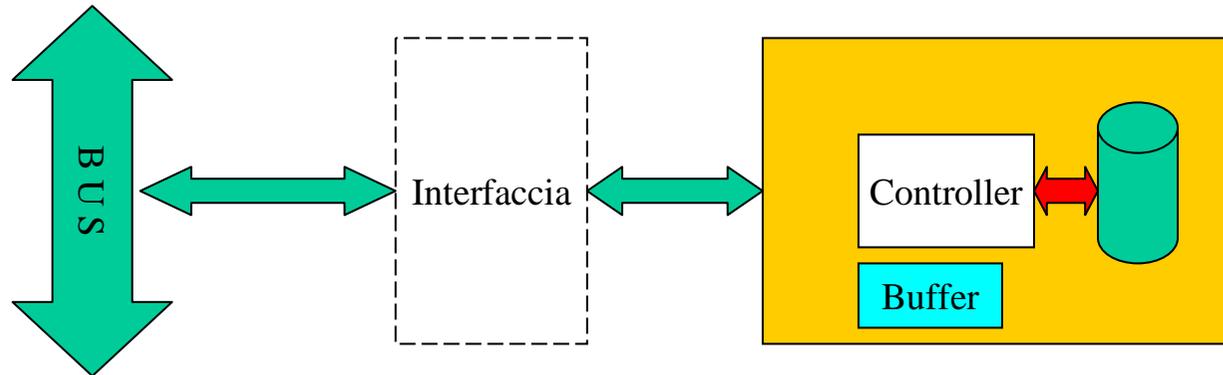
In un sistema con più dischi il cavo di controllo è unico, mentre ci sono tanti cavi dati quanti sono i dischi rigidi



Interfaccia ATA (AT-Attachment)

Il controller è integrato sul disco rigido assumendo l'aspetto di una periferica unica. IDE (Integrated Drive Electronic)

La periferica è connessa direttamente (da qui il nome) o tramite una scheda di interfaccia, al bus di sistema.



La geometria fisica del disco (Cilindri, settori, testine) è mascherata dal controller integrato. Al sistema è mostrata una geometria logica. La traslazione fisico-logico è mappata nel firmware del controller

E' necessaria per implementare la Zoned Bit Recording (numero di settori per traccia dipendente dalla distanza dal centro del disco).

Svincola il disco dal BIOS

Assenza di settori difettosi (Bad Cluster). Il disco ha un certo numero di settori di riserva (spares) cui rimappare in maniera trasparente gli eventuali settori difettosi. → La comparsa di settori difettosi va interpretata come imminente failure del disco

La formattazione a **basso livello**, viene effettuata solo la prima volta in fabbrica

Interfaccia ATA, le specifiche in ordine cronologico

Standard	Nome commerciale	Modo Trasferimento dati	Limite Disco	Caratteristiche	Periodo
pre-ATA	IDE	PIO 0	2.1GB	22 bit LBA	
ATA-1	ATA,IDE	PIO 0,1,2,3 SW DMA 0,1,2 MW DMA 0	137GB	28 bit LBA	1994-1999
ATA-2	EIDE, FastATA, FastIDE,UltraATA	PIO 3,4 MW DMA 1,2			1996-2001
ATA-3	EIDE	Spariscono i Single Word DMA		S.M.A.R.T.	1997-2002
ATA/ATAPI-4	ATA-4, UltraATA/33	Ultra DMA 0,1,2 (UDMA/33)		Supporto CD- ROM, TAPE...	1998-
ATA/ATAPI-5	ATA-5, UltraATA/66	Ultra DMA 3,4 (UDMA/66)		Cavo 80 poli	2000-
ATA/ATAPI-6	ATA-6, UltraATA/100	UltraDMA 5 (UDMA/100)	144PB	48 bit LBA	2002-
ATA/ATAPI-7	ATA-7, UltraATA/133	UltraDMA 6 (UDMA/133)		SATA 1.0	2005-

Programmed I/O (PIO) Modes

E' la modalit  di trasferimento dati prevista dallo standard ATA nelle prime implementazioni.

La CPU controlla il trasferimento dei dati tra il sistema e il disco rigido

→ Degradamento delle prestazioni della CPU nelle operazioni di I/O

Pio Mode	Durata ciclo (nS)	Max Transfer Rate (MB/s)	Standard
Mode 0	600	3.3	ATA
Mode 1	383	5.2	
Mode 2	240	8.3	
Mode 3	180	11.1	ATA-2
Mode 4	120	16.7	

Direct Access Memory (DMA) Modes

Il trasferimento dei dati tra il sistema e il disco rigido avviene direttamente, senza coinvolgere la CPU.

Si distingue tra *third party dma* (quando c'è un terzo dispositivo, il dma controller, che gestisce il transito tra disco e memoria) e **bus master** (o *first party dma*) quando è la periferica stessa che prende il controllo del bus

DMA Mode	Durata ciclo (nS)	Max Transfer Rate (MB/s)	Standard
Single Word Mode 0	960	2.1	ATA (rimossi in ATA-3)
Single Word Mode 1	480	4.2	
Single Word Mode 2	240	8.3	
Multi Word Mode 0	480	4.2	ATA
Multi Word Mode 1	150	13.3	ATA-2
Multi Word Mode 2	120	16.7	

Word: 16 bit = 2 byte

→ Introduzione Bus Master

Non ha avuto molto successo:

Scarso supporto, driver difettosi. Più complesso del PIO 4 senza grandi vantaggi

Ultra DMA Modes

Vengono migliorate le prestazioni del trasferimento mediante DMA:

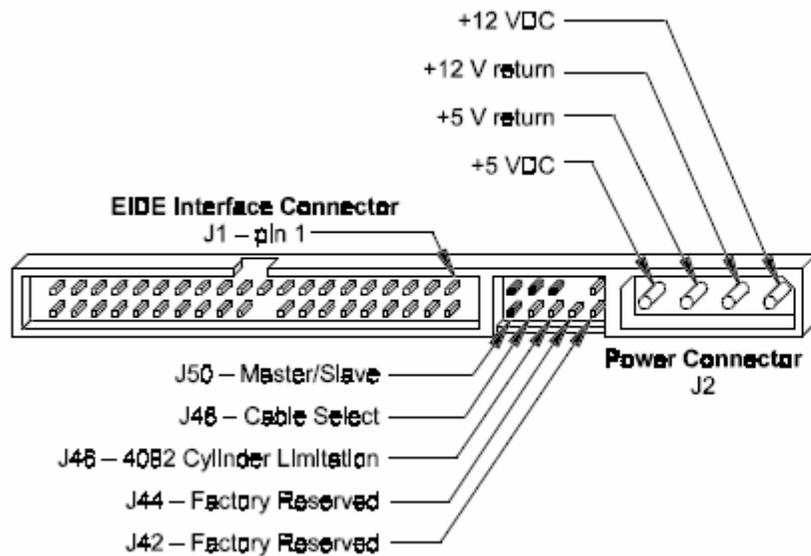
- Due trasferimenti per ciclo di clock, invece di uno (*double transition clocking*)
- Controllo degli errori (CRC):
 - Ritrasmissione del blocco danneggiato
 - Scalatura a velocità inferiore nel caso di eccessive ritrasmissioni
- Cavo da 80 poli (invece dei consueti 40): vengono aggiunti 40 conduttori posti a GND intercalati agli esistenti per migliorare la separazione dei segnali e ridurre le interferenze

UltraDMA Mode	Durata ciclo (nS)	Max Transfer Rate (MB/s)	Standard
Mode 0	240	16.7	ATA/ATAPI-4
Mode 1	160	25.0	
Mode 2	120	33.3	
Mode 3	90	44.4	ATA/ATAPI-5
Mode 4	60	66.7	
Mode 5	40	100.0	ATA/ATAPI-6

- Ciascuno percorso dei dati attraverso l'interfaccia IDE/ATA è definito "canale"
- Ogni canale può supportare fino a due periferiche (disco o altro dispositivo ATAPI)
- In teoria un sistema può supportare fino a quattro canali (limite: risorse, es. IRQ)

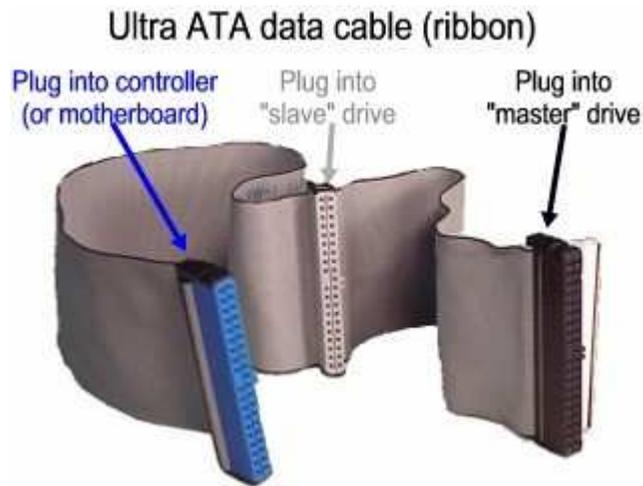
Le due periferiche su uno stesso canale vengono definite *master* e *slave*. La distinzione è necessaria per distinguere le due unità ed evitare conflitti (visto che uso lo stesso mezzo di comunicazione). Ma non c'è nessuna operazione di controllo dell'una sull'altra, o differenza di prestazioni.

In effetti la dizione corretta (e l'unica riportata nelle specifiche ATA) è, rispettivamente: **device 0** e **device 1**



JUMPER CONFIGURATION	J50	J48	J46	J44	J42
Master/Slave					
Only drive in single drive system*	C				
Master drive in dual drive system*	C				
Slave drive in dual drive system	O				
Cable Select					
Disabled*		O			
Enabled		C			
4092 Cylinder Limitation					
Disabled*			O		
Enabled			C		
Factory Reserved				O	
Factory Reserved					O
Key * = Default C = Closed (jumper installed) O = Open (no jumper installed)					

Interfaccia ATA: connessione



- Lunghezza massima del cavo: **46cm** (18 pollici) secondo le specifiche. E' **sconsigliato** l'uso di cavi di lunghezza superiore.
- Connettore IDC (Insulated Displacement Connector) 40 pin 0.100" spacing
- Lungo il cavo è evidenziato il bordo corrispondente al conduttore n.1 (tipicamente rosso, per i cavi grigi, bianco per i cavi neri)
- I cavi con funzione Cable Select (CS) presentano connettori colorati per distinguerne la funzione

• L'interfaccia ATA (Parallel ATA dopo l'introduzione del SATA) è di tipo **parallelo** a 16 bit.

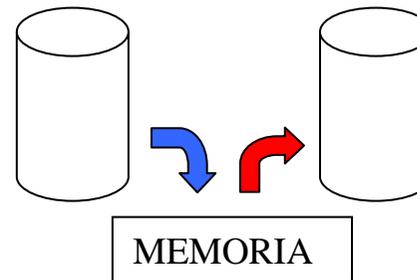
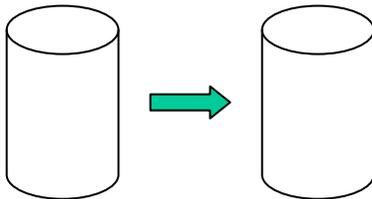
- Cavo a 40 conduttori / 40 pin: previsto dalle prime specifiche ATA, **non adatto per la modalità UltraDMA**
- Cavo a 80 conduttori / 40 pin: è compatibile pin-to-pin con il cavo standard, gli ulteriori 40 conduttori sono connessi a massa e servono a migliorare la schermatura tra i segnali. **E' indispensabile per operare in UDMA/33 e superiori.** In caso contrario il controller si predispose in un modo di trasferimento inferiore.
- Se il disco è impostato su *cable select* il master e slave sono riconosciuti automaticamente dalla posizione assunta sulla piattina: device 0 (o master) quello all'estremità, device 1 (o slave) quello in posizione intermedia. (Se viene scambiata la posizione del controller entrambi i dischi vengono visti come master, quindi conflitto)
- Se invece l'indirizzamento è stato prefissato con i jumper, la posizione è indifferente.
- N.B. se su una piattina c'è un solo disco, questo deve collegato all'estremità, per evitare problemi di riflessione.
- I cavi CS a 40 conduttori non sono compatibili (relativamente alla funzione cable select) con quelli a 80. Ovviamente se l'impostazione è da jumper, non c'è incompatibilità..

Master e Slave: prestazioni e criterio di configurazione

Due dispositivi a velocità diversa sullo stesso controller sono limitati alla velocità del più lento ?

- Se il controller supporta la *independent master/slave device timing* ogni dispositivo funziona alla propria velocità ottimale. Altrimenti, effettivamente, se il chipset non lo supporta e le velocità sono sensibilmente diverse, vengono entrambi fatti lavorare nella modalità più lenta compatibile.
- In ogni caso il mezzo di comunicazione è **condiviso**, solo un dispositivo alla volta può “parlare”. Questo significa che durante un massiccio impiego del canale il dispositivo più veloce dovrà attendere che l’altro abbia terminato. Al contrario, nelle maggior parte delle motherboard, due dispositivi su due canali distinti sono in grado di elaborare le richieste contemporaneamente

L’effetto “collo di bottiglia” assume rilevanza diversa a seconda del tipo di operazione effettuata:



Master e Slave: altre considerazioni

- Sequenza di boot: l'identificazione master/slave condiziona anche la scelta dell'unità dalla quale avviare il sistema operativo (In alcuni chipset più vecchi il boot è possibile solo dal canale primario)
- Assegnazione da parte del sistema operativo della lettera di unità (C: D:...)
- Hard disk e drive ATAPI:
 - Diversa struttura dei comandi
 - Più lenti
 - In alcuni casi incompatibilità con DMA Bus Mastering
- Presenza di dischi non recenti: limitazione della velocità di trasferimento (es. supporto solo PIO)
- Tipo di bus cui è connesso l'host adapter

Serial ATA (SATA)

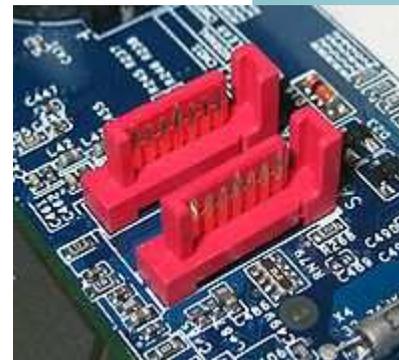
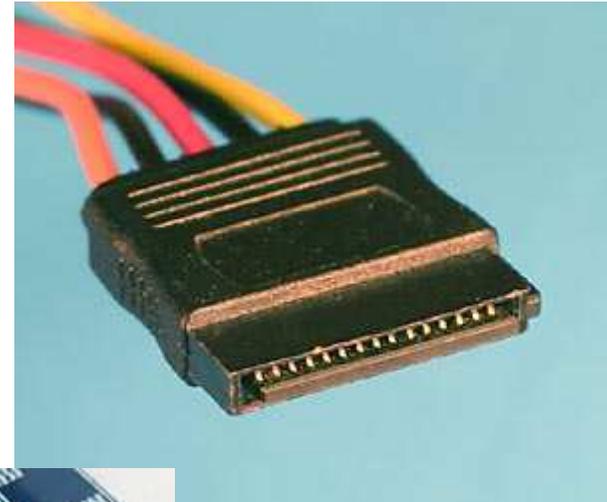
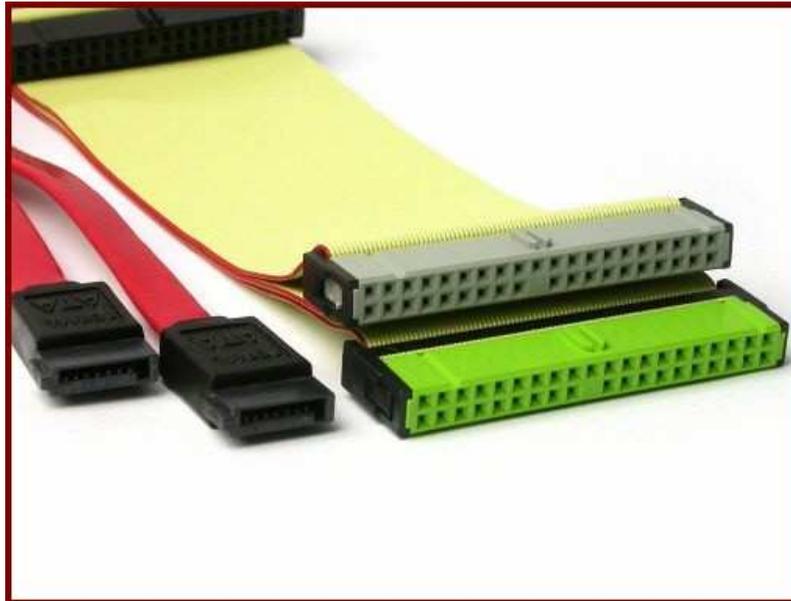
Trasmissione Seriale (1bit) Protocollo Punto-Punto

Lunghezza dei cavi prevista dalle specifiche: max 1m (è superiore rispetto al PATA)

Velocità: 1.5Gb/s (SATA/150) 3.0Gb/s (SATA/300) 6Gb/s (SATA/600)

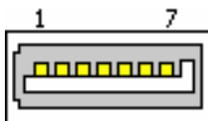
4 linee di segnale, trasmissione differenziale, tensione ridotta (0.5V contro 5V del PATA)

Può supportare l'*hotplugging* (verificare l'host adapter)



SATA: connettori dati e power supply

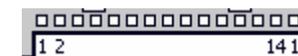
Connettore DATI



Pin	Nome	Funzione
1	GND	Ground
2	A+	Transmit+
3	A-	Transmit-
4	GND	Ground
5	B-	Receive-
6	B+	Receive+
7	GND	Ground
-		chiave

Pin	Mating	Funzione
-		chiave
1	3 [^]	3.3V*
2	3 [^]	
3	2 [^]	
4	1 [^]	Ground
5	2 [^]	
6	2 [^]	
7	2 [^]	5V
8	3 [^]	
9	3 [^]	
10	2 [^]	Ground
11	3 [^]	Reserved
12	1 [^]	Ground
13	2 [^]	12V
14	3 [^]	
15	3 [^]	

Connettore Power Supply



Per supportare l'**hot plugging** i piedini del connettore maschio e femmina sono disposti in modo da entrare in contatto secondo la sequenza indicata

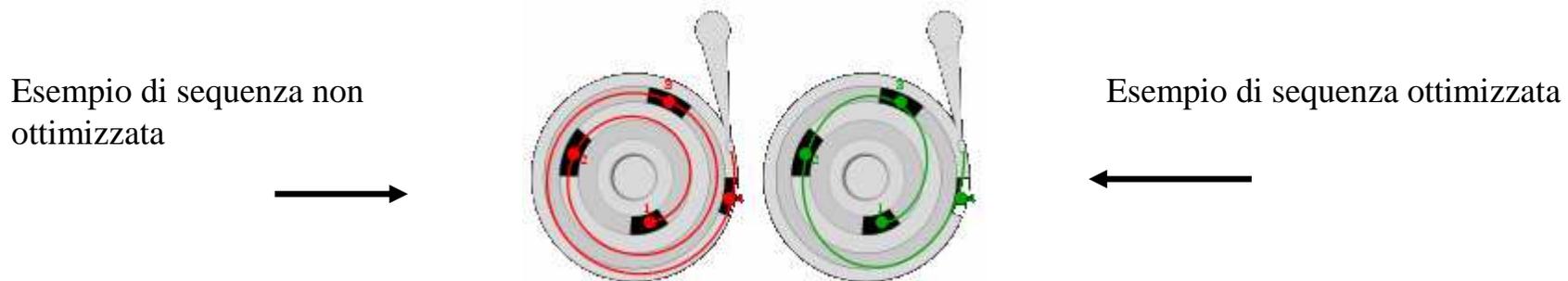
Rispetto al connettore tradizionale è possibile la presenza di una terza tensione di alimentazione (3.3V)

I contatti sono parallelati a gruppi di tre per assicurare sufficiente portata in corrente.

Esistono adattatori dal connettore Molex standard 4 poli. (Solo +5V e +12V)

SATA vs PATA

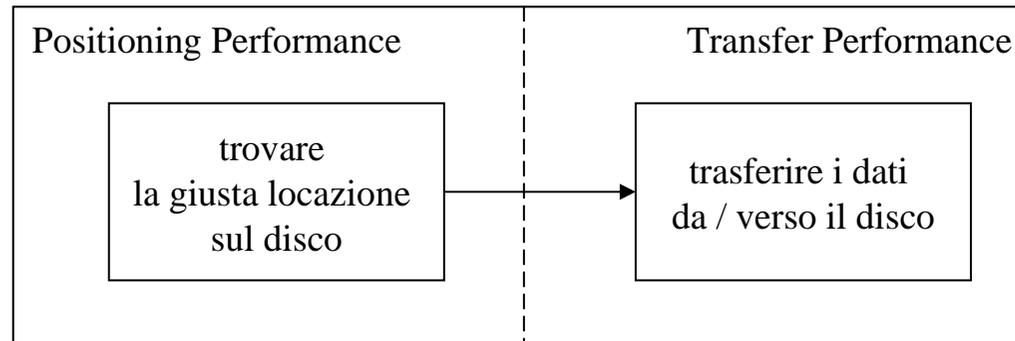
- Protocollo Punto-Punto: non occorre impostare master e slave
- Cavi più sottili e di maggior lunghezza rispetto il PATA: facilità di cablaggio, meno ostruzione alla circolazione d'aria nel case
- Velocità comparabile al PATA(SATA/150) o superiore (SATA/300-SATA/600)
- Tensione ridotta (0.5V contro 5V del PATA): minor emissione disturbi, consumo comparabile
- Può supportare l'hotplugging (verificare la disponibilità sul host adapter)
- Native Command Queing (NCQ): incremento delle prestazioni permettendo al disco di ottimizzare l'ordine con il quale i comandi di lettura/scrittura ricevuti sono eseguiti



“Non tutto quel ch'è oro brilla,/ Né gli erranti sono perduti”, ovvero:

- Prestare attenzione ai connettori dati che hanno la tendenza a sfilarsi.** (Eventualmente impiegare i connettori dotati di fermo di chiusura)
- Fragile: Non piegare / tirare il cavo lateralmente, pena la rottura del connettore sulla motherboard**

Prestazioni: parametri significativi



Seek time: è la misura del tempo impiegato per spostare la testina dalla traccia corrente ad una nuova traccia

Latenza: è il tempo che occorre attendere, una volta posizionata la testina sulla traccia desiderata, affinché il settore cercato si presenti sotto di essa. E' direttamente correlato con la velocità di rotazione del disco.

Acces time: rappresenta il tempo di accesso ad una locazione sul disco. E' una grandezza composta.

Internal Media Transfer Rate: misura la velocità con cui vengono letti i bit dalla superficie del disco (singola traccia)

Internal Sustained Transfer Rate: misura la velocità alla quale è possibile trasferire dati sequenziali da più cilindri sul disco. Significativo per file di grandi dimensioni. Risente della frammentazione

External Transfer Rate (Burst Rate): rappresenta la velocità di trasferimento dei dati dal buffer all'host attraverso l'interfaccia.

S.M.A.R.T. (Self-Monitoring, Analysis and Reporting Technology)

- E' uno strumento di monitoraggio dei dischi allo scopo di evitare la perdita di dati prevedendo eventuali malfunzionamenti
- Introdotto ufficialmente come specifica di comunicazione host-disco a partire da ATA-3 (1997)

Guasti prevedibili e imprevedibili

- **Imprevedibili:** dovuti a cause elettriche (es. sbalzi di tensione) o meccaniche senza manifestazione di segni premonitori. L'incidenza dei guasti imprevedibili può essere ridotta operando sulla qualità in sede di progettazione e produzione.
- **Prevedibili:** caratterizzati dal progressivo degrado di un parametro misurabile nel tempo prima che il disco manifesti il malfunzionamento. Nella maggior parte dei casi si tratta di guasti di tipo meccanico.(es. usura, stress da riscaldamento)

Circa il 60% dei guasti tipicamente sono di origine meccanica (fonte Seagate)

Attributi

Ogni costruttore definisce un insieme di parametri (attributi) da tenere sotto osservazione. Ciascun attributo ha:

- un valore grezzo (*raw value*), il cui significato è strettamente dipendente dal costruttore (Sotto questo aspetto la tecnologia SMART non è uno "standard"), es. temperatura in gradi, tempo di funzionamento in minuti o ore, etc
- un valore *normalizzato* in una scala da 0 a un valore massimo (100-200-253).
Regola empirica: valori alti=buon segno, valori bassi=cattivo presagio.
- Il costruttore stabilisce anche la soglia (*threshold*) sotto la quale, fintanto che il disco funziona regolarmente, il parametro osservato non deve scendere.

	Raw read error rate	Tasso di errori di lettura
	head flying height	Altezza di volo fuori specifiche (numero di eventi)
	data throughput performance	Prestazioni generali
	spin-up time	Tempo di avvio
	re-allocated sector count	Conteggio dei settori difettosi riallocati in settori di riserva
	seek error rate	Tasso di errori di ricerca (problemi nel posizionamento corretto della testina)
	seek time performance	
	Power on hour count	Tempo di funzionamento (da confrontare con il MTBF)
	spin retry count	Conteggio dei tentativi di avvio del disco
	drive calibration retry count	Numero dei tentativi di ricalibrazione
	Temperature	Temperatura del disco
	Reallocation Event Count	Numero di operazioni di rimappatura dei settori (sia riuscite che fallite)
	Current Pending Sector Count	Numero di settori "instabili" (non ancora rimappati)
	Uncorrectable Sector Count	Errori non recuperabili
	UltraDMA CRCError Rate	Sintomo di problemi di comunicazione host/disco (es.cavi difettosi o fuori specifica)
	Write Error Rate	Tasso di errore in scrittura

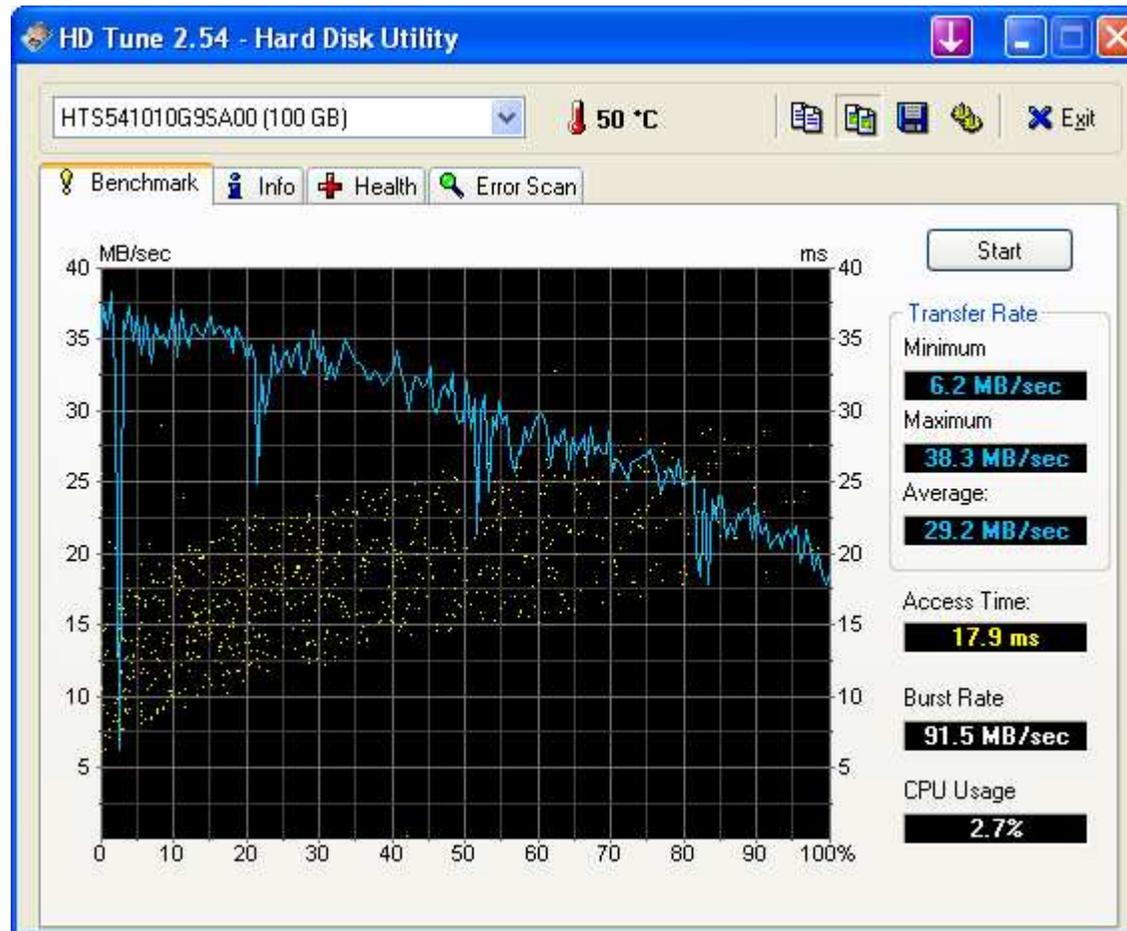
Utility per monitoraggio...

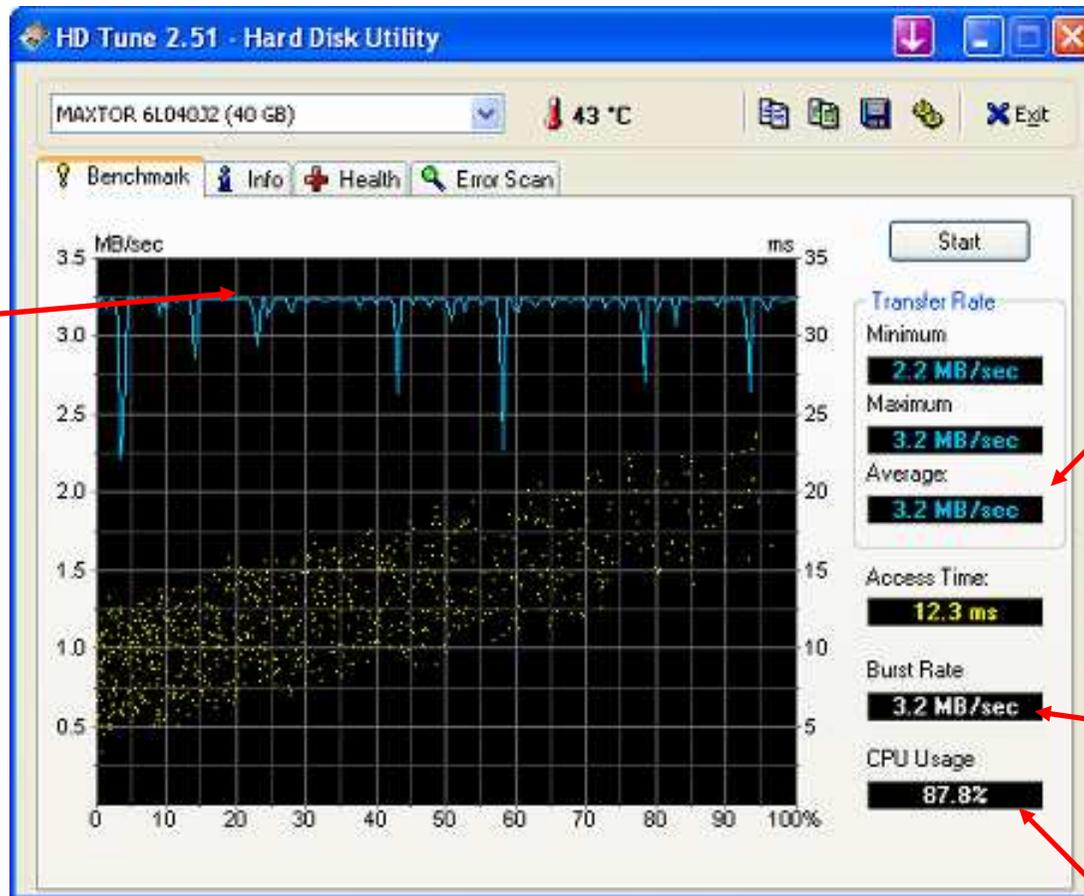
The screenshot shows the HD Tune 2.54 - Hard Disk Utility application window. The title bar reads "HD Tune 2.54 - Hard Disk Utility". The main window displays the drive "HTS541010G9SA00 (100 GB)" and a temperature of "49 °C". The "Health" tab is selected, showing a table of SMART attributes. The table has columns for ID, Current, Worst, Threshold, Data, and Status. The status for all attributes is "Ok". At the bottom, the "Power On Time" is 8635 and the "Health Status" is "Ok".

ID	Current	Worst	Threshold	Data	Status
(01) Raw Read Error Rate	100	100	62	0	Ok
(02) Throughput Performance	100	100	40	0	Ok
(03) Spin Up Time	253	253	33	0	Ok
(04) Start/Stop Count	9	9	0	144081	Ok
(05) Reallocated Sector Count	100	100	5	0	Ok
(07) Seek Error Rate	100	100	67	0	Ok
(08) Seek Time Performance	100	100	40	0	Ok
(09) Power On Hours Count	81	81	0	8635	Ok
(0A) Spin Retry Count	100	100	60	0	Ok
(0C) Power Cycle Count	99	99	0	2522	Ok
(BF) G-sense Error Rate	100	100	0	0	Ok
(C0) Power Off Retract Count	100	100	0	18	Ok
(C1) Load Cycle Count	65	65	0	350545	Ok
(C2) Temperature	112	112	0	1048625	Ok
(C4) Reallocated Event Count	100	100	0	0	Ok
(C5) Current Pending Sector	100	100	0	0	Ok
(C6) Offline Uncorrectable	100	100	0	0	Ok
(C7) Ultra DMA CRC Error Count	200	253	0	0	Ok

Power On Time: **8635** Health Status: **Ok**

...e rilievo delle prestazioni





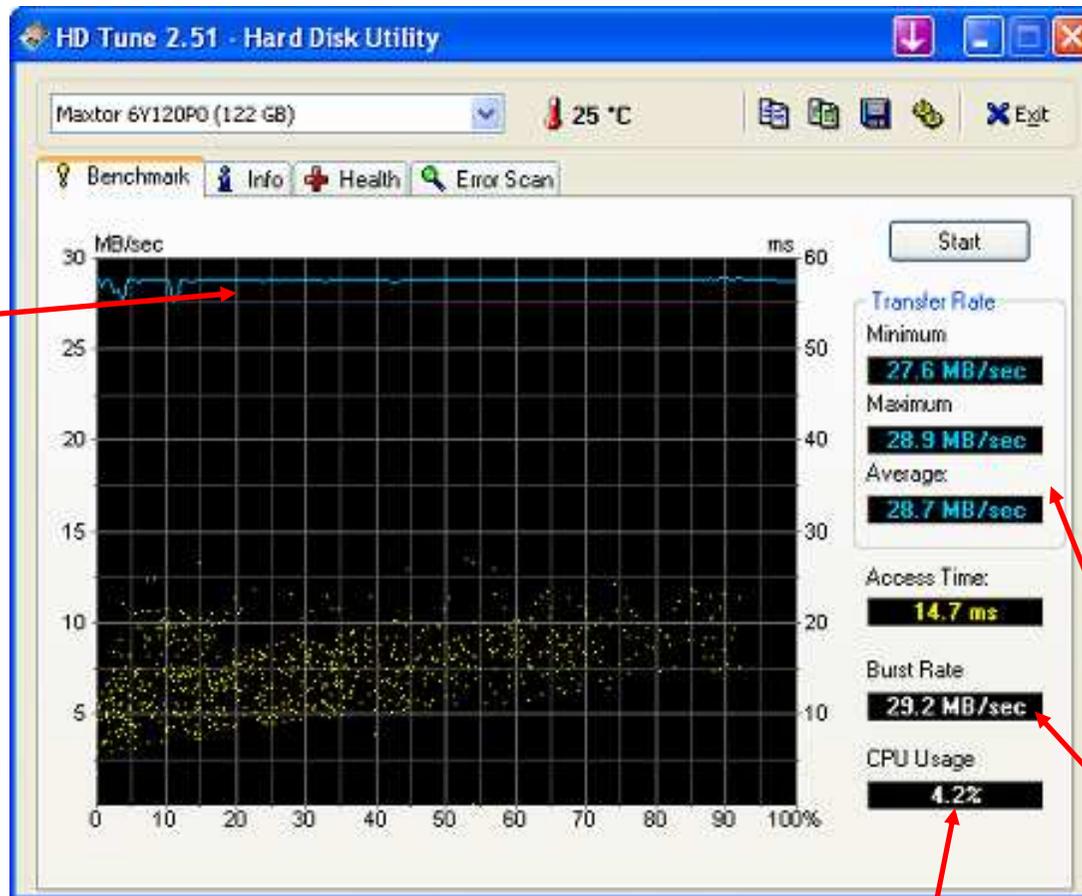
Banda piatta

Basso transfer rate

Basso transfer rate di picco (cache→host)

Elevato carico della CPU

Interpretazione: modo UDMA non attivo



Banda piatta

Interpretazione: modo non superiore a UDMA2. Se il disco supporta modi superiori verificare che non sia stato usato un cavo da 40 poli invece che 80

Basso carico della CPU

Transfer rate \leq 30MB/s

RAID (Redundant Array of Inexpensive/Independent Disk)

Sistema mediante il quale l'informazione è, in maniera trasparente all'utente, divisa e/o replicata tra più dischi

Gli obiettivi che si pone un sistema RAID sono:

- Incrementare l'affidabilità/integrità dei dati
- Incrementare la tolleranza ai guasti
- Incrementare le prestazioni in termini di tempi di accesso/transfer rate.
- Incrementare lo spazio disponibile

A seconda del tipo di architettura RAID implementata è possibile raggiungere uno o più degli obiettivi elencati

Implementazione:

Software – Il RAID è realizzato, sfruttando i normali controller, a livello di sistema operativo. Soluzione a basso costo (non richiede hardware addizionale), ma nel complesso lento.

Hardware – Fa uso di un apposito controller RAID. Può essere una scheda di espansione oppure un risiedere sulla scheda madre.. Questa soluzione ha prestazioni migliori ed è più semplice da gestire rispetto alla soluzione software

Livelli singoli:

RAID 0

RAID 1

RAID 5

Livelli nidificati:

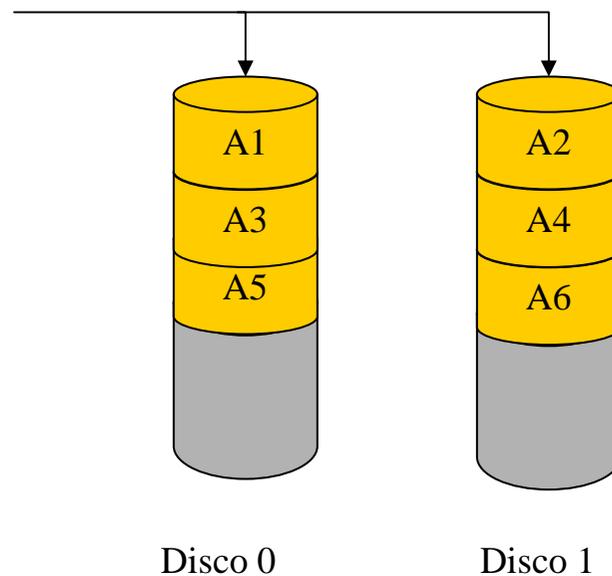
RAID 0+1

RAID 1+0

RAID 0 (Striped set without parity - Striping)

I blocchi logici sono divisi equamente tra le unità che costituiscono l'array. Non c'è nessuna informazione di parità o ridondanza.

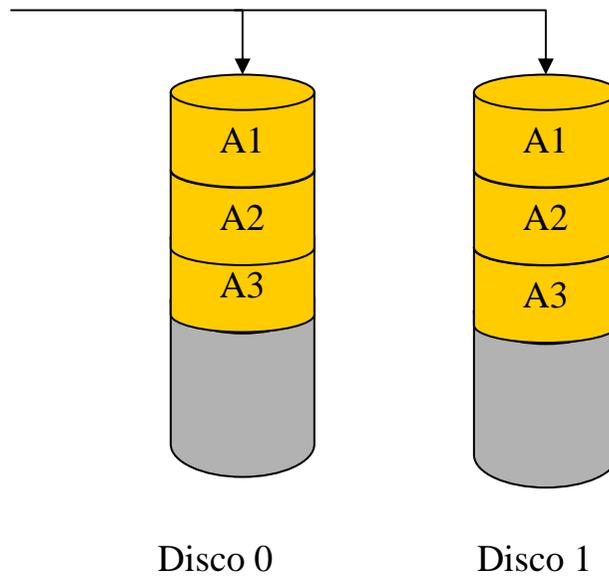
- Numero minimo di dischi: 2
- Incremento delle prestazioni
- Incremento dello spazio
- Non è "fault tolerant": la perdita di un disco comporta la distruzione dell'intero array (perdita dei dati)
- L'affidabilità è inversamente proporzionale al numero di dischi
- N.B.: originariamente non previsto nella definizione di RAID



RAID 1 (Mirrored set without parity - Mirroring)

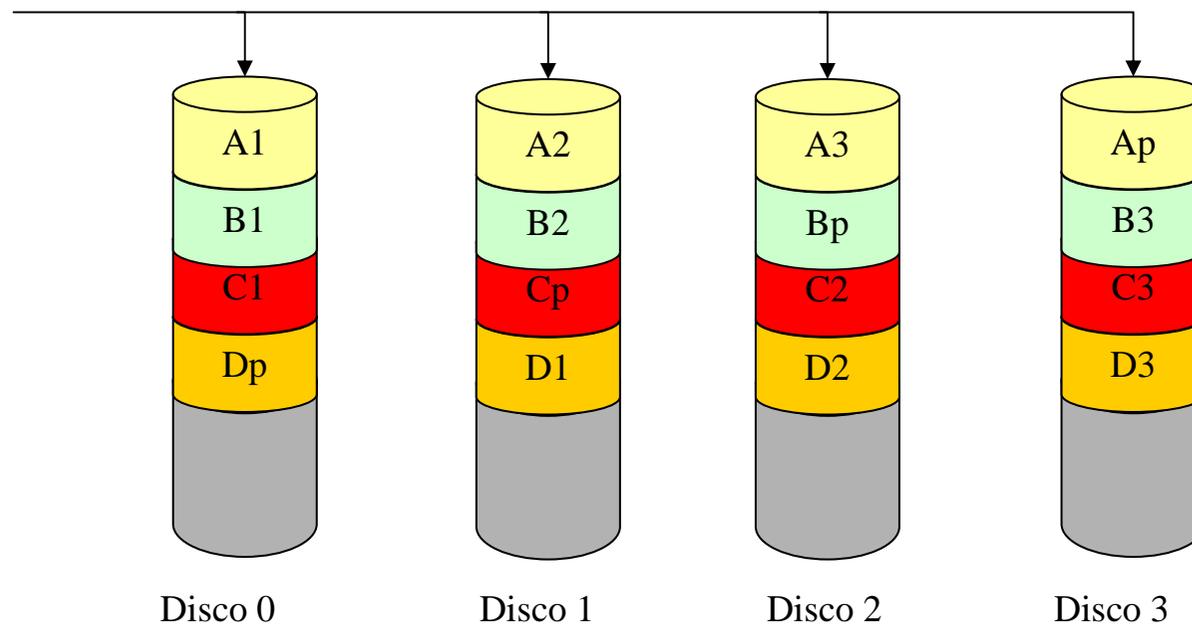
I dati sono copiati (copia esatta) su tutte le unità dell'array. Ridondanza completa.

- Numero minimo di dischi: 2
- La capacità dell'array è quella di un singolo disco.
- L'affidabilità aumenta linearmente con il numero di dischi impiegati
- Lieve incremento di prestazioni in lettura rispetto al singolo disco, peggiore rispetto altri livelli RAID
- Lieve peggioramento delle prestazioni in scrittura rispetto al singolo disco, migliori rispetto altri RAID



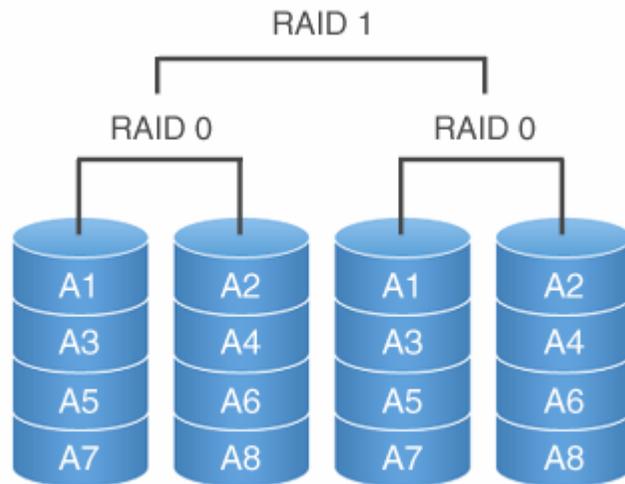
RAID 5 (Block level striping with distributed parity)

- E' presente un controllo di parità che consente il recupero degli errori.
- Sia i blocchi di dati che i blocchi di parità sono distribuiti tra le unità che costituiscono l'array.
- Numero minimo di dischi: 3
- La capacità totale dell'array è $(N-1)$ *singolo disco (se i dischi sono uguali).
- Fault tolerant: in caso di guasto i dati possono essere recuperati dai rimanenti grazie alla parità (con riduzione delle prestazioni)
- In presenza di guasto l'array è vulnerabile finché non è sostituito il disco difettoso (un secondo guasto distrugge l'array).
- E' considerato un buon compromesso tra prestazioni, affidabilità e capienza



RAID Nidificati

RAID 0+1



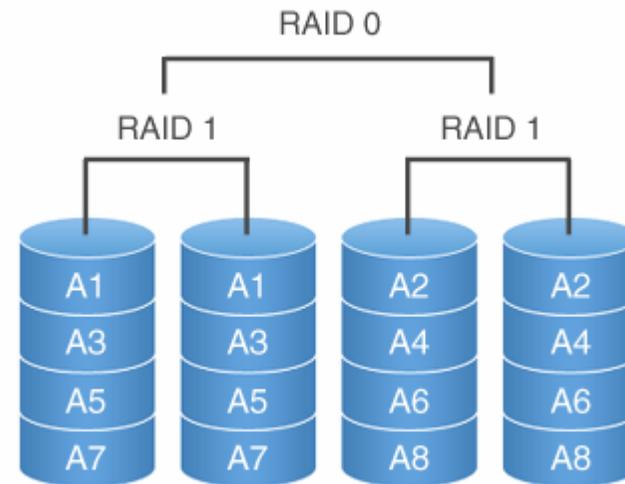
“Mirrored of striped”

Elevata velocità di trasferimento

Stesso livello di affidabilità di un RAID5

Il guasto di un singolo disco degrada l'array a RAID0

RAID 1+0



“Striped of mirrored”

Elevata affidabilità con alte prestazioni

Stesso livello di affidabilità di un RAID1

Può tollerare il guasto di due dischi in mirror diversi