

Capitolo sesto	2
La memoria di massa.....	2
Premessa	2
1. La registrazione su superfici magnetiche	2
2. Pre-e post-elaborazioni; sincronizzazione.....	6
2.1 Controllo della sequenza di bit.....	6
2.2 Compressione	6
2.3 Rilevazione e correzione di errore.....	7
2.4 La sincronizzazione dei blocchi	8
3. Architettura delle memorie di massa.....	8
4. Interfacce e bus per le memorie di massa.....	10
4.1 Porta seriale e parallela.....	12
4.2 Interfacce IDE/ATA	13
4.3 Interfaccia SCSI.....	13
4.4 Bus esterno: USB e IEEE 1394.	15
4.5 Interfaccia PCMCIA.....	15
4.6 SSA e Fibre Channel	15
5. Memorie a nastro	16
5.1 Generalità	16
5.2 Nastri a 9 tracce	18
5.3 Nastri a serpentina	18
5.4 Nastri a scansione elicoidale	19
5.5 Nastri ottici trasversali.....	20
6. Memorie a dischi: caratteristiche meccaniche	20
7. Dischi magnetici: organizzazione dei dati.....	22
7.1 Cilindro, superficie, settore	22
7.2 Tempo di accesso al settore.....	22
7.3 Struttura del settore	24
7.4 Formattazione	24
7.5 Densità di registrazione	24
7.6 Il drive	25
8. I dischi magnetici	25
8.1 Hard disk.....	25
8.2 Array di dischi	26
8.3 Dischi rimovibili.....	26
9. Dischi ottici e magneto-ottici	27
9.1 Dischi ottici	27
9.2 Dischi magneto-ottici	29
10. Biblioteche di nastri e dischi	30
11. Dischi a stato solido (SSD) e schede di memoria	31
11.1 SSD.....	31
11.2 Schede di memoria ed apparati PCMCIA	31

Capitolo sesto

La memoria di massa

Premessa

I vertiginosi cambiamenti nella tecnologia delle memorie di massa, hanno portato sul mercato soluzioni tecnologiche a costi sempre più bassi, a dispositivi sempre più veloci e con elevata larghezza di banda. Il risultato di tutto ciò è la grande esplosione dei supporti di memorizzazione e di input/output.

È difficile seguire questa dinamica in un testo come questo, il cui obiettivo primario resta quello di introdurre i concetti fondamentali sui quali si basa l'ingegneria informatica. Pertanto, in questo capitolo si presentano da un lato i fondamenti delle memorie di massa, senza inseguire le innumerevoli varianti che ne hanno perfezionate le prestazioni, per le quali non basterebbe un apposito volume, e dall'altro si offre una panoramica dell'attuale produzione (1999), con particolare riferimento ai prodotti per PC.

1. La registrazione su superfici magnetiche

Il principio fisico delle memorie con registrazione su superfici magnetiche (nastri e dischi) è analogo a quello della registrazione audio/video analogica, benché dal punto di vista tecnico vi siano differenze a causa del fatto che occorre realizzare solo due distinti stati di magnetizzazione, corrispondenti ai due valori del bit e che ai requisiti di fedeltà vanno sostituite quelli di capacità e velocità.

In fig.1.1 è mostrato lo schema di principio di un gruppo testina-superficie magnetica; in fase di scrittura, nell'avvolgimento apposito viene inviato un impulso di corrente mentre la superficie è in movimento relativo rispetto alle testine con velocità costante; le linee di flusso che si stabiliscono nel circuito magnetico della testina si chiudono attraverso la superficie che, essendo costituita da un materiale ad alto valore di induzione residua e bassa forza coercitiva, resta così magnetizzata. Sulla superficie ne risulta allora una distribuzione delle linee di flusso del tipo di fig.1.1c) e si forma una piccola zona che si comporta come una sottile barretta di materiale ferromagnetico; la

distribuzione della densità di flusso lungo la superficie $B(l)$ assume un andamento del tipo indicato in fig.1.2b), supposto che l'impulso di corrente sia di tipo rettangolare (fig. 1.2a). La zona così magnetizzata viene detta cella elementare.

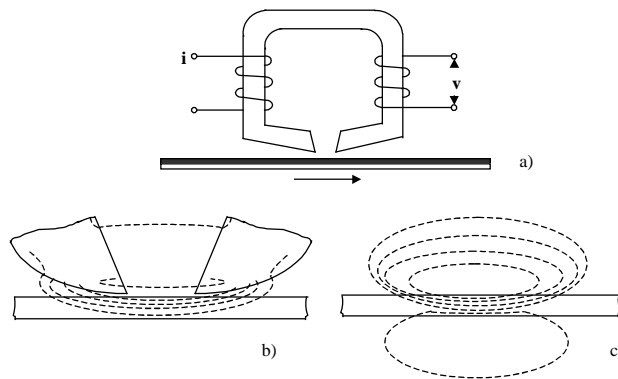


Figura 1.1 - Registrazione su superficie magnetica

In fase di lettura, allorché la cella elementare transita con velocità costante sotto la testina, la variazione del flusso $B(l)=B(t)$ che si determina nel circuito magnetico induce nell'avvolgimento apposito una tensione v che assume nel tempo la forma della derivata della funzione $B(t)$ (fig. 1.2c) e che identifica in qualche modo l'impulso di corrente che ha causato la magnetizzazione (teoricamente, mediante integrazione elettronica della tensione $v(t)$, si ottiene una curva che riproduce la funzione $B(t)$ e quindi la $i(t)$). La registrazione di un "1" potrebbe essere distinta da quella di uno "0" usando segnali di registrazione di polarità opposta (come si vedrà, questa è soltanto una delle tecniche possibili).

Sul piano tecnico, la lunghezza della cella elementare determina la densità di registrazione, intesa come numero di bit per unità di lunghezza. Questa è vincolata dalle caratteristiche elettromagnetiche del sistema superficie-testina ed in primo luogo dal fatto che, in presenza di celle adiacenti molto vicine, magnetizzate alla medesima polarità, si determina una distribuzione complessiva della densità di flusso del tipo illustrato in fig.1.2d), che tende a rimanere costante e dalla quale, quindi, risulta difficile trarre un segnale di rilevazione utile. Sol tanto in seconda istanza il limite alla densità di registrazione è dato dal limite al numero di inversioni di flusso per unità di lunghezza determinato dal materiale magnetico.

Diversi sono i metodi di registrazione adoperati; quelli fondamentali si possono così classificare:

- metodi a *modulazione di ampiezza*, *modulazione di fase* o *modulazione di frequenza*, a seconda della grandezza elettrica che viene usata per distinguere il segnale di registrazione di un 1 da quello di uno 0 (si tratta, peraltro in ogni caso di segnali impulsivi e quindi le grandezze elettriche non sono quelle relative a segnali continui analogici, ma la loro trasposizione nel discreto);

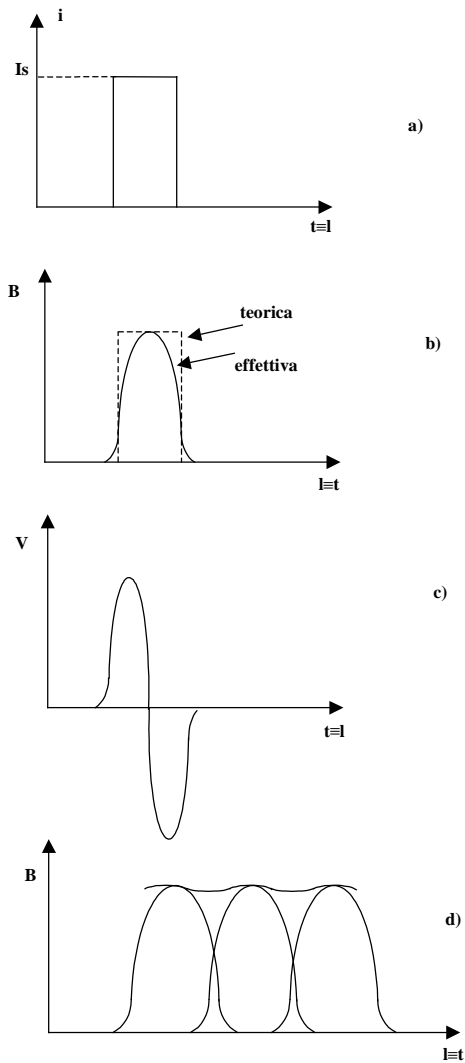


Figura 1.2 - Principi di funzionamento della memoria a superficie magnetica

- metodi *RZ* o *NRZ* (Ritorno o Non-Ritorno a Zero) a seconda che, nell'intervallo fra due escursioni successive, il segnale ritorna o non ad un livello di riferimento (ad esempio 0 ampère).

In fig.1.3 sono esemplificati i metodi fondamentali di registrazione. Il metodo *RZ* a modulazione di ampiezza (fig.1.3a) è il primo, caratterizzato da impulsi di polarità opposta per l'1 e lo 0, sempre di durata inferiore a quella della cella; la rilevazione avviene come già illustrato.

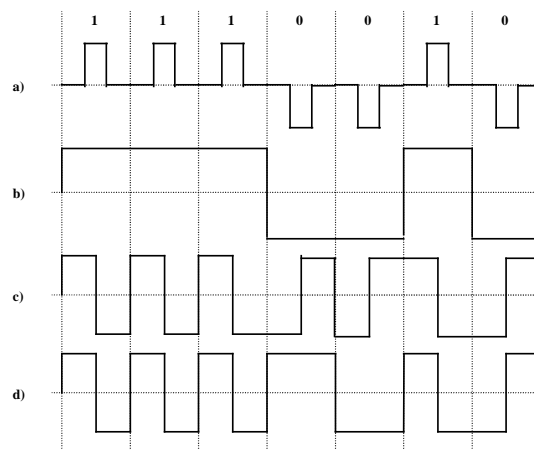


Fig.1.3. Metodi di registrazione su superficie magnetica con diversi tipi di modulazione: a) ampiezza, *RZ*; b) ampiezza, *NRZ*; c) fase; d) frequenza.

Il corrispondente metodo *NRZ* (fig.1.3b) dà per scontato che in corrispondenza di stringhe di egual valore si tende a non ottenere variazione di flusso e quindi, in tal caso, mantiene costante il segnale di registrazione. La tecnica di lettura è basata sul fatto che il sistema di lettura tiene memoria del bit letto precedentemente e lo assume costante mentre il circuito di rilevazione segnala (con la corrispondente variazione di flusso) la eventuale variazione da 1 a 0 o viceversa,

I metodi a modulazione di fase e di frequenza (e *NRZ*), prendendo atto delle difficoltà derivanti da magnetizzazioni costanti in celle consecutive, impongono comunque variazioni di polarità nella registrazione.

Il metodo a modulazione di fase (PE: Phase Encoder), con tecnica NRZ, pone in ogni cella sempre un segnale di doppia polarità, in differente fase se per l'1 o per lo 0 (fig.1.3c).

Il metodo a modulazione di frequenza (fig.1.3d, sempre di tipo NRZ) prevede due segnali di polarità diversa per l'1, un unico segnale per lo 0, ma in ogni caso il primo segnale di una cella ha polarità opposta dell'ultimo della cella precedente. In sede di rilevazione, è possibile estrarre dalla registrazione un segnale che, opportunamente filtrato, fornisce da un lato il sincronismo e dall'altro il bit 1 o 0.

I metodi fondamentali sono stati poi variamente sviluppati. In particolare, con l'uso di materiali magnetici con particolari caratteristiche (p.e. materiali anisotropi) e con l'uso di particolari testine di lettura-scrittura (testine a film magnetico, testine magnetoresistive etc) si raggiungono oggi (1999) densità di registrazione longitudinale (cioè lungo la traccia di registrazione) di centinaia di migliaia di bit/pollice (bpi) a fronte dei 6.400 bpi dei nastri magnetici della generazione dei mainframe.

2. Pre-e post-elaborazioni; sincronizzazione

2.1 Controllo della sequenza di bit

I metodi di registrazione non consentono particolari sequenze di bit in quanto risulterebbero irriconoscibili alla lettura, ad esempio non consentono una stringa di oltre n bit eguali.

Per fare in modo che la stringa registrata non contenga le sequenze vietate, in fase di scrittura appositi algoritmi trasformano la stringa originaria aggiungendovi alcuni bit strategicamente posizionati; in fase di lettura avviene la trasformazione opposta.

I moderni drive di dischi e nastri sono tutti dotati di almeno un microprocessore all'interno dell'unità di controllo; a questo è affidato il compito di pre- e post-elaborazione della stringa di bit.

Ad esempio, lo standard di registrazione dei CD-ROM prevede che ogni byte venga registrato su 14 bit rielaborati secondo un algoritmo detto Reed Salomon e che fra un byte ed un altro vengano inoltre aggiunti altri 3 bit detti di merging: agli 8 bit originari se ne aggiungono così altri 9.

2.2 Compressione

Effettuare una *compressione* dei dati significa rappresentarli con un numero di bit inferiore a quello derivante dalla loro normale rappresentazione carattere per

carattere (se è un testo), pixel per pixel (se è una immagine), scena per scena (se è un film) e così via.

Ovviamente, le tecniche di compressione ben si adattano alle memorie di massa che possono essere in tal modo usate in modo più efficiente; esse sono anche raccomandate se bisogna trasmettere in rete un grosso file, in quanto si risparmia in tal modo tempo di trasmissione.

Appositi algoritmi si possono applicare alla stringa di bit originaria per comprimerla oppure, viceversa, per decomprimere la stringa compressa; essi peraltro esorbitano dai limiti di questo testo e pertanto ci limiteremo soltanto ad accennarne l'esistenza.

Gli algoritmi possono essere realizzate via software, usando appositi pacchetti di utilità (tipico è WinZip operante sotto Windows) oppure via hardware, ed è quello che spesso viene realizzato dalle memorie di massa: l'hardware del drive effettua la compressione e decompressione dei dati.

Un algoritmo di compressione si dice *distruttivo* se non rende possibile la decompressione: sono tipicamente distruttivi gli algoritmi che si applicano alle immagini per memorizzarle, riducendone anche di un fattore 200 lo spazio necessario senza ridurne significativamente la qualità (ad esempio riducendone la scala dei grigi).

Per la trasmissione e la memorizzazione di dati testuali si adoperano viceversa tecniche di compressione non distruttive, con le quali si riduce di un fattore 2 o 3 lo spazio di memoria (o il tempo di trasmissione) necessario.

2.3 Rilevazione e correzione di errore

La tecnica dell'apposizione di un bit di parità (cfr. § IV.2.4) può essere generalizzata con l'apposizione di un carattere oppure di una stringa di bit di controllo: così come il bit di parità è la somma modulo-2 di tutti i bit da controllare e l'ultimo carattere del codice fiscale è un carattere di controllo calcolato in funzione dei precedenti attraverso un algoritmo definito dalla Gazzetta Ufficiale, è possibile costruire una stringa di controllo dei dati contenuti in un blocco. Aumentando il numero di extra-bit aggiunti, aumenta la probabilità che non sfuggano errori al controllo.

Codici di controllo siffatti si dicono *checksum* oppure *LRC (Longitudinal Redundancy-Check)* oppure ancora *CRC, Cyclic Redundancy Check*¹. Essi hanno lo scopo di riconoscere l'esistenza di un errore nella stringa di bit costituente il blocco e vengono tipicamente apposti in coda a ciascun blocco.

La teoria dei codici fornisce anche altre possibilità: aumentando la ridondanza, è possibile costruire codici che siano in grado, sempre su basi

¹ Invero, il codice CRC deriva da una specifica tecnica di controllo, detta appunto di codice ciclico o polinomiale, diversa da quella del resto in modulo.

probabilistiche, di correggere un errore “piccolo” all’interno del blocco (ad esempio, l’errore di un solo bit). Codici di questo genere sono detti a correzione di errore (*ECC, Error Correcting Code*) e sono applicati in molte memorie di massa, che diventano allora capaci di autocorreggere alcuni errori. In coda ai blocchi viene posto, dopo il checksum, il codice ECC e l’hardware del drive corregge i bit del blocco se il checksum lo ha rivelato errato.

2.4 La sincronizzazione dei blocchi

L’inizio e la fine di un blocco sono individuati dal sistema di lettura-scrittura attraverso apposite registrazioni di controllo; in particolare, l’inizio è individuato da una particolare sequenza di caratteri, detta "di sincronismo", ad esempio un carattere speciale, detto SYNC ripetuto più volte e seguito da un carattere START. Analogamente, la fine può essere individuata da un particolare carattere, che diremo EOB (End Of Block). La struttura di un blocco assume dunque in generale il seguente aspetto:

SYNC-...-SYNC -START- (dati utili) -EOB- LRC- ECC

Il sistema di scrittura registra dunque i caratteri di controllo di cui sopra, quello di lettura li riconosce ed opera di conseguenza. La tecnica è concettualmente simile a quella usata per trasmettere un blocco di dati in rete.

3. Architettura delle memorie di massa

Dal punto di vista dell’architettura generale del sistema di elaborazione, una memoria di massa è una unità periferica ed è pertanto collegata al processore che la gestisce con uno degli schemi di cui al cap.IV (p.e. fig. IV.6.1), in particolare attraverso un’apposita unità di controllo della periferia; questa, nel caso specifico delle memorie di massa, è fondamentalmente composta da due sezioni, talora anche fisicamente distinte e costituenti altrettante unità parzialmente autonome: il trasportatore e l’unità di controllo (o governo) propriamente detta.

Il trasportatore contiene le parti meccaniche per realizzare il moto testina-superficie e i circuiti per le operazioni di scrittura e lettura; esso riceve ed esegue i comandi per l’avvio e l’arresto del moto, la lettura e la scrittura.

L’unità di controllo costituisce un’interfaccia logico-elettronica fra il trasportatore e l’interfaccia del processore; le sue funzioni sono molteplici e più o meno complesse. Fondamentalmente, essa svolge le seguenti funzioni:

- riceve dal processore i comandi relativi all'operazione da compiere, li decodifica e li invia al trasportatore, trasformandoli eventualmente in sequenza di comandi più elementari;
- controlla il transito dei dati, trasformandoli dalla forma fornita dalla CPU a quella propria del trasportatore e viceversa;
- comprime e decomprime i dati e in genere li pre-elabora e post-elabora;
- corregge eventuali errori in caso di codici ECC;
- riconosce le condizioni di malfunzionamento del trasportatore o della trasmissione (o, più in generale, lo "stato" della periferica) e le trasmette al processore con le regole del colloquio fra processore e periferica;
- adatta i livelli logico-elettronici dei segnali, controlla la tempificazione delle trasmissioni e sincronizza le due unità in connessione;
- in qualche caso svolge in hardware operazioni di selezione dei dati;
- in qualche caso gestisce una *cache* della memoria di massa.

Inoltre, un'unica unità di controllo può da un lato governare più trasportatori e dall'altro essere collegata a più processori; le sue funzioni comprendono allora la selezione del trasportatore e la gestione del collegamento con le CPU.

Spesso l'unità di controllo è strutturata come un vero e proprio sistema di elaborazione, con un processore per la pre- e post-elaborazione, una memoria con funzioni di buffer o di cache e addirittura una memoria di massa, per esempio in una unità a nastri con cache a dischi.

L'unità di registrazione su memoria di massa è il *record fisico o blocco*: una singola operazione di lettura, scrittura o selezione avviene, infatti, dal punto di vista hardware, su un insieme di caratteri, detti appunto record fisico o blocco.

Le dimensioni del blocco sono dettate da esigenze, talora contrastanti, di ottimizzazione delle prestazioni, quali:

- limiti fisici del supporto di registrazione;
- dimensioni del buffer di memoria centrale nel (dal) quale trasferire il blocco;
- probabilità crescente di riscontrare un errore di registrazione al crescere della dimensione del blocco;
- diminuzione dei tempi complessivi di lettura e scrittura per blocchi di dimensioni crescenti.

Sul piano software o più in generale su quello delle applicazioni, dove è necessario far riferimento ad aspetti logici invece che fisici, si fa riferimento come unità di registrazione, invece, ad un *record logico*, le cui dimensioni sono definite dalle stesse applicazioni. Più record logici costituiscono un file (logico), che è l'entità cui fanno riferimento le applicazioni.

In genere, k record logici sono registrati in un blocco e k è detto *fattore di bloccaggio*; il software di base consente di trasformare le operazioni di accesso a record logici richieste dall'applicazione in operazioni di accesso ai blocchi.

Il supporto magnetico sul quale sono registrati i blocchi (bobina di nastro, disco, etc.) è detto *volume*. Un file può occupare parte di un volume oppure anche più volumi (*file multivolume*).

In una memoria di massa più che negli altri sistemi si sente l'esigenza di risolvere la problematica di gestione in un'architettura a livelli nella quale si procede da macchine semplici via via a macchine più complesse, l'una innestata nell'altra, quali tipicamente:

- *trasportatore e unità di controllo locale*, che costituiscono nel loro complesso il drive hardware dell'apparecchiatura;
- *driver* di gestione della periferica: è un sottoprogramma operante nella CPU, che esegue una singola operazione di lettura o scrittura, trasformandola in generale nella sequenza di codici necessari per la specifica unità di controllo;
- *metodo di accesso*: è un insieme di sottoprogrammi che consentono, mediante una o più chiamate al driver e ad altri eventuali sottoprogrammi di servizio, di "vedere" gli archivi organizzati logicamente secondo regole che ne semplificano le operazioni di selezione;
- *sistema operativo*: si interfaccia con i metodi di accesso da un lato e con tutto il sistema dall'altro, gestendo lo sfruttamento delle risorse necessarie alle operazioni su file e record (assegnazione di buffer, allocazione dei file, etc.);
- *sistema di programmazione*: al livello del singolo linguaggio o a livelli più elevati (ad esempio sistema data-base) definisce le operazioni di ingresso/uscita su memoria di massa dal punto di vista delle applicazioni.

Un'operazione di selezione, lettura o scrittura richiede quindi un complesso funzionamento interlacciato di gran parte del sistema calcolatore, sia come hardware che come software; tale complessità deriva dal numero notevole di operazioni elementari necessario per l'effettuazione di un'operazione completa di ingresso o uscita e dalla necessità di rendere completamente automatica la procedura di collegamento, scambio di messaggi, controllo di errori e sincronizzazione. Nei paragrafi che seguono si illustrano soltanto gli aspetti più elementari dei livelli inferiori dell'architettura.

4. Interfacce e bus per le memorie di massa

Le memorie di massa, così come le altre periferiche, sono tipicamente collegate su appositi bus che ne caratterizzano il comportamento sia dal punto di

vista del colloquio sia da quello delle prestazioni, tipicamente la velocità di trasferimento dati, sia infine da quello elettrico. Il bus di cui si è detto al cap. IV in uscita dalla CPU verso la memoria e la periferia (cfr. ad esempio figg. IV.9.1, IV.11.1, IV.12.1, IV.13.1) costituisce in effetti il cosiddetto *bus interno*, che è caratterizzato dal colloquio elementare ivi studiato e dalla massima velocità di trasferimento dati, quella verso la memoria.

In quella sede si è ipotizzato per semplicità che la periferica fosse collegata direttamente al bus interno, ma ciò non accade in realtà. Nell'architettura dei PC, spesso ripresa anche dai sistemi più evoluti, al bus interno sono connessi altri bus attraverso appositi *bridge* (per adattarne caratteristiche elettriche) o interfacce, per adattarne anche caratteristiche logico-funzionali e definirne nuovi protocolli di colloquio.

In primo luogo, al bus interno sono connessi appositi “*bus di espansione*”, che permettono di inserire all'interno del computer apparecchiature aggiuntive, quali adattatori video, acceleratori grafici, sound card e modem interni (*expansion boards*).

I primi PC facevano uso di un bus di espansione chiamato *bus AT* o bus *ISA* (*Industry Standard Architecture*), dotato di una frequenza di 8 MHz ed un parallelismo di 16 bit, poi evoluto nell'*Enhanced ISA* (EISA) a 32 bit. Per motivi storici, citiamo anche il bus *MCA* (Micro Channel Architecture) proprietario su macchine IBM. Questi bus, tuttavia, sono stati rapidamente superati in quanto la loro velocità di trasferimento costituiva un vero e proprio collo di bottiglia a seguito delle sempre migliori prestazioni dei sistemi di memoria di massa

L'evoluzione tecnologica ha fatto sì che, a partire dal 1992, fossero proposte soluzioni radicali che sostituivano il bus di espansione con i più veloci *bus locali*, dando vita allo standard *VESA Local Bus* (da Video Electronics Standards Association), o *VL Bus* e poi all'attuale (1999) più flessibile bus *PCI* (Peripheral Component Interconnect). Il bus *ISA* tipicamente permane, collegato con un bridge al *PCI*, per motivi di compatibilità con dispositivi più lenti che ben si adattavano alla vecchia tecnologia (cfr. fig. 4.1).

Il bus *PCI* ha prestazioni molto elevate: è dotato infatti di un parallelismo a 64 bit, anche se spesso viene proposto nella versione a 32 bits; ha una frequenza di 33 o 66 MHz: per dare un esempio delle prestazioni ottenibili con tali dati, con semplici calcoli è possibile ricavare che con un parallelismo a 32 bits e frequenza pari a 33 MHz, la velocità del bus è di 132 MB/s. Una delle ragioni del grande successo del bus *PCI* è che esso viene popolato da adattatori con velocità che sono prossime a quelle del bus locale della CPU.

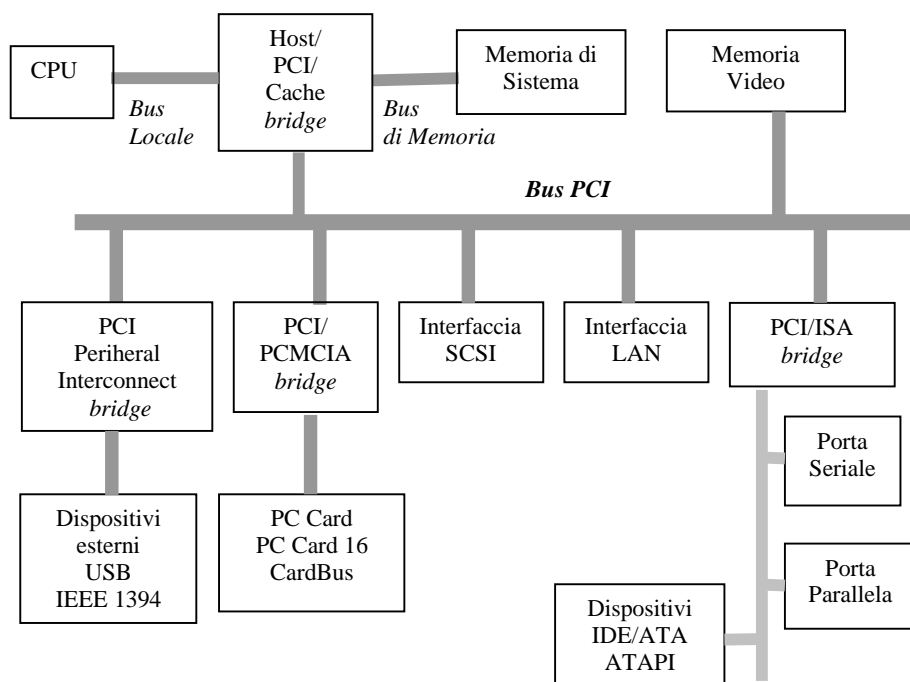


Figura 4.1: Architettura basata su bus PCI.

Nel seguito descriviamo brevemente le principali interfacce che ritroviamo su PCI e sulle espansioni attaccate ad esso.

4.1 Porta seriale e parallela

Le più antiche interfacce di collegamento verso l'esterno dei PC sono la *porta seriale* e la *porta parallela*, collegate sul bus di espansione ISA.

La porta seriale implementa lo standard *RS-232 C* o *RS-422* e viene usata, ovviamente, per comunicazioni in cui viene trasmesso 1 bit per volta. Costituisce una interfaccia *general purpose* che può essere usata per svariate tipologie di dispositivi, come modem, mouse, stampanti, collegamento ad altri computers.

La porta parallela implementa di solito lo standard di fatto *Centronics*, usa connettori a 25 pin detti DB-25 e viene usata per connettere stampanti. Porte parallele più veloci di circa 10 volte rispetto alla Centronics sono la EPP (*Enhanced Parallel Port*) e la ECP (*Extended Capabilities Port*), che supportano comunicazione bidirezionale.

Sia la porta parallela sia quella seriale stanno per essere sostituite dai bus esterni tipo *USB* e *IEEE 1394*.

4.2 Interfacce IDE/ATA

La famiglia di interfacce in questione viene chiamata *IDE* (*Intelligent Drive Electronics* o, secondo altri, *Integrated Drive Electronics*) oppure *ATA* (*AT Attach*, cioè “*attaccata*” ad un AT, il PC di seconda generazione, essendo quelli di prima ad 8 bit denominati XT) ed ha un parallelismo di 16 bit, esattamente come il bus di I/O dell’AT IBM originario.

L’interfaccia è una estensione naturale del bus di sistema ISA, in quanto la maggior parte dei segnali e della circuiteria necessari sono già presenti nella CPU; ne segue che uscite IDE/ATA sono spesso presenti sulla “scheda-madre” del PC, senza hardware aggiuntivo. Essa riproduce il concetto dei vecchi main frame di “periferica interna” al sistema; l’interfaccia standard prevede un singolo canale condiviso da due dischi, in modalità master/slave.

Il sistema ATA, peraltro, definisce un apposito *protocollo di comunicazione*, inizialmente rivolto ai soli dischi, poi ampliato a nastri e CD-ROM. In esso sono definiti i comandi ad alto livello dell’host, le risposte di stato da parte della periferica circa il completamento dell’operazione o eventuali condizioni di errore, le modalità di trasferimento a blocchi (DMA) e così via. Il protocollo è del tipo *programmed I/O*, o PIO (cfr. § IV.13); ovviamente, il drive della periferica contiene un’apposita unità di controllo locale progettata secondo le regole del protocollo ATA (sui manuali si legge che il drive è “intelligente” oppure che “possiede un controllore builte-in”).

La famiglia in esame, oltre ad ATA originario, possiede le evoluzioni ATA-2, ATA-3, Ultra-ATA, Fast-ATA, ATAPI, EIDE, che presentano diversi miglioramenti nella velocità, nell’indirizzamento dei blocchi, nell’affidabilità, nell’analisi dei guasti. Si segnala in particolare che Ultra ATA raggiunge velocità elevate (33.3 MB/s), usando a tale scopo una particolare logica di rilevazione e correzione di errore.

4.3 Interfaccia SCSI

L’interfaccia *SCSI* (pronuncia *scasi*, *Small Computer System Interface*) definisce un protocollo e presenta in uscita un bus di I/O parallelo (fig. 4.2) su cui possono essere collegati svariati dispositivi e diverse unità di controllo locale (controller). Progettata attorno al 1985, SCSI è il più vecchio sistema di interconnessione per apparecchiature “esterne” attualmente in uso e serve una grossa varietà di dispositivi: dischi, cassette, dischi ottici (*inclusi CD-ROM*,

DVD, WORM e magneto-ottici), dischi rimovibili, sottosistemi RAID, scanner, stampanti e plotters.

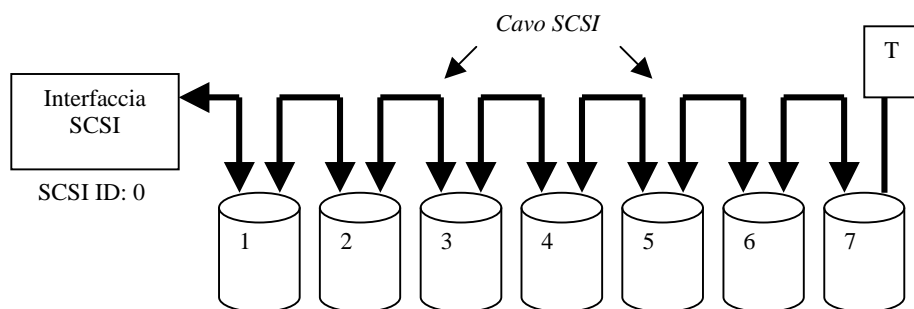


Figura 4.2: Il bus SCSI

Il primo standard SCSI, *parallel SCSI o SCSIII*, fu standardizzato nel 1986 e progettato per interconnettere fino a 7 dispositivi alla velocità di 5MB/s, con 8 linee dati + 1 di parità e con l'uso di un connettore a 50 pin. Ne furono implementate due differenti versioni, incompatibili tra loro, caratterizzate da distinte tecniche di pilotaggio elettrico del bus: Single Ended (*SE*) e High Voltage Differential, (*HVD*) o semplicemente *Differenziale*; quest'ultima, essendo meno sensibile al rumore, può essere utilizzata su distanze e velocità di bus più elevate.

Successivamente si sono avute le seguenti principali evoluzioni dello SCSI, ottenute anche attraverso tecnologie avanzate di tipo elettrico, quali la *LVD*, *Low Voltage Differential*:

- *SCSI-2* (1992), con 10 MB/s per parallelismo di 1 byte, ma con la possibilità di raddoppiare il parallelismo (16 bit-dati + 2 di parità, detta *wide SCSI*) arrivando così a 20 MB/s;
- *Ultra SCSI*, con 20 MB/s su 1 byte, 40 MB/s nella versione a 2 byte (*Wide Ultra SCSI*, o *Fast 40*);
- *Ultra 2 SCSI*, con 40 MB/s su 1 byte, 80 MB/s nella versione wide (*Wide Ultra 2 SCSI* o *Fast 80*);
- *Ultra 3 SCSI*, detta anche *SPI-3 (SCSI Parallel Interface 3)*, in fase di sviluppo. Una versione implementata è detta *Ultra160 SCSI* e, introducendo nuove tecnologie, arriva ad una velocità di 160 MB/s, solo nella versione wide bus.
- *SPI-4*, detta anche *Ultra 320 SCSI*, prevista per il 2001;
- *SPI-5*, detta anche *Ultra 640 SCSI*, prevista per il 2003;

4.4 Bus esterno: USB e IEEE 1394.

L'interfaccia *USB (Universal Serial Bus)* fu sviluppata nel 1995 con l'idea di individuare un meccanismo semplice di installazione di una periferica, a basso costo e semplice da usare, avente caratteristiche "Plug and Play" e di connessione "a caldo". Oggi (1999) è uno standard che supporta trasferimenti di dati a 12 Mbit per secondo e che può connettere per ciascuna porta fino a 127 dispositivi, compresi modem, mouse, tastiere; quasi tutti i PC di nuova produzione hanno una o più porte USB.

L'interfaccia *IEEE 1394* (che assume anche nomi commerciali *FireWire*, *Link* e *Lynx*) è più veloce (400 Mbit per secondo), una sua porta connette fino a 63 apparati e supporta anch'esso 'Plug and Play' e connessione a caldo. Lo standard supporta anche il trasferimento *isocrono*, ovvero fornisce garanzie che un flusso di dati sia trasmesso in un determinato tempo. Questa caratteristica rende il bus ideale per periferiche che richiedono trasferimento dati *real time*, come dispositivi video. Anche a causa del suo maggior costo rispetto all'USB, IEEE 1394 viene di solito usato per *video camere*,

4.5 Interfaccia PCMCIA.

PCMCIA sta per *Personal Computer Memory Card Internation Associaton*, e rappresenta una associazione con circa 500 società che ha sviluppato uno standard per dispositivi piccoli, di forma simile a carte di credito, chiamata anche PC cards.

Originariamente, PCMCIA è stato progettato per aggiungere memoria ai PC portatili; è stato poi espanso più volte ed ora può essere usato per diversi dispositivi. Vi sono tre tipi di carte PCMCIA, tutte di forma rettangolare (85.6 x 53 mm), ma di differente spessore (cfr. § 11.2):

- tipo I: spesse 3.3 mm, usate principalmente per collegare *RAM o ROM*;
- tipo II: spesse 5.0 mm, usate per *modem e modem fax*.
- tipo III: spesse 10.5 mm, usate per *dischi portabili*.

4.6 SSA e Fibre Channel

L'interfaccia *SSA (Serial Storage Architecture)* fu proposta da un gruppo di produttori come un nuovo standard, che doveva sostituire SCSI e altre interfacce più vecchie e più lente (Progetto ANSI X3T10.1). A differenza della configurazione a bus o catena prevista dallo SCSI, la SSA prevede una topologia ad anello con connessioni a *daisy chain*. Ogni nodo può supportare due anelli

controrotanti, avendo così la possibilità di trasmettere dati contemporaneamente nei due sensi. Il progetto prevede per ogni link 40 MB/s teorici (in realtà è più lento) e possibilità di supportare fino a 127 dispositivi con distanze fino a 15 metri tra gli stessi.

Il progetto FC (Fibre Channel) ebbe inizio nel 1988, supportato da un altro gruppo di produttori, con lo scopo di rimpiazzare le interfacce di memoria esistenti (Progetto ANSI X3T11) e di connettere sia sistemi sia supporti di memoria. Attualmente (1999), è presente sul mercato la tecnologia Fibre Channel Arbitrated Loop (FC-AL), sviluppata con particolare attenzione alle connessioni di supporti di memoria. La FC-AL connette fino a 126 apparati, con una banda teorica che va da 13 MB/s fino a 4 GB/s, con distanze fino a 10 Km.

5. Memorie a nastro

5.1 Generalità

Le memorie a nastro costituiscono le più classiche memorie ad accesso sequenziale. In esse, il supporto magnetico è costituito da una striscia di materia plastica ricoperta da uno strato di materiale magnetico, di lunghezza e larghezza variabile nei diversi modelli (oggi, larghezza da 4 a 19 mm, lunghezza di qualche centinaia di metri o di 1 kilometro). Il nastro è inizialmente avvolto su una bobina e, nei modelli più moderni, è incapsulato in un supporto (*cartridge*).

Le unità a nastro magnetico sono fra le memorie quelle che hanno subito le più grosse trasformazioni con lo sviluppo della tecnologia. Dal punto di vista funzionale, pur avendo sempre occupato il vertice della gerarchia di memorie, si sono via via trasformate da memorie di lavoro “in linea” (negli anni '60 si usavano come supporto alle operazioni di ordinamento dei file) a supporti per le operazioni di backup. Dal punto di vista meccanico, alle grosse bobine di 12” di diametro dei modelli dei vecchi mainframe si è passati a strutture molto simili alle audiocassette (e le stesse audiocassette sono state adoperate come unità a nastro).

Una classificazione di tipo meccanico distingue le unità in:

- *cartridge* (ma per qualcuno questo termine è usato in senso più generale), che contengono la sola bobina alimentatrice e che, come le vecchie *bobine aperte*, devono essere riavvolte per essere smontate;
- *cassette a due ruote*, che, come le audiocassette, possono essere rimosse senza la necessità di riavvolgerle.

Su un nastro, i dati sono suddivisi in blocchi (record fisici) strutturati ciascuno con una testata (*header*), i dati veri e propri, una coda (*trailer*) ed appositi spazi fra i blocchi detti *inter-block gap* o *inter-record gap*. In qualche caso si usano blocchi a lunghezza fissa (tipicamente 512 byte), in altri variabili (da 1 a 64k, raramente più lunghi).

Si usano appositi blocchi detti *tape marker* per indicare la fine di un file. All'atto della lettura, all'unità di controllo del nastro si può chiedere di posizionarsi (*locate* è l'istruzione che si usa nello standard SCSI-2) all'inizio dello *i*-esimo file e quindi il posizionamento avviene in base al conteggio dei *tape-marker*. In alternativa, alcuni modelli di nastro prevedono informazioni per la localizzazione dei file registrate fra i blocchi di dati oppure su apposite tracce percorse a grande velocità.

La maggioranza dei nastri è del tipo "append-only", nel senso che su un nastro scritto non è possibile riscrivere da un certo punto in poi, aggiungendovi altri file.

Le caratteristiche meccaniche e strutturali del nastro lo rendono adatto alla lettura in sequenza dei blocchi ed al trasferimento, anche a buone velocità, di flussi di dati. Se il flusso di trasferimento è sospeso, la necessità di fermarsi, riallocarsi e riprendere il trasferimento peggiora anche di due ordini di grandezza la velocità di trasferimento dei dati. Altri limiti nell'uso dei nastri derivano dal fatto che esiste un tempo non indifferente per il caricamento del supporto, a causa di una notevole sequenza di operazioni preliminari di aggiustaggio meccanico ed elettronico che vengono avviate automaticamente.

I nastri usano intensivamente tecniche di codici a correzione di errori (ECC) per migliorare da un lato la velocità e dall'altro l'affidabilità.

Ad esempio, il blocco del nastro Exabyte 8505 contiene 14 byte di header, 1024 byte di dati utili, 2 byte per il checksum e 400 byte di ECC.

I nastri usano spesso tecniche di compressione dei dati, tipicamente in un fattore da 2 a 3. Di questa evenienza si deve tener conto in quanto talora i dati di targa delle apparecchiature riportano capacità e data transfer rate riferiti a blocchi compressi piuttosto che a quelli originari (detti anche parametri "nativi"), inducendo in errore l'utente.

I nastri possono essere classificati in base all'orientamento delle tracce di registrazione rispetto al nastro stesso nonché alle tecniche meccaniche di registrazione:

- nastri a 9 tracce
- nastri longitudinali a serpentina
- nastri a scansione elicoidale
- nastri con tracce trasversali.

5.2 Nastri a 9 tracce

È questo il modello originario (o quasi) dei nastri, nato negli anni '60 per i mainframe, sopravvissuto fino agli inizi degli anni '90, molto popolari perché spettacolari nei loro movimenti, tanto da essere stati spesso riprodotti in spettacoli cinematografici e televisivi.

Il modello, a bobina aperta di 12" e nastro largo ½", registra i dati su 9 tracce parallele (1 byte + 1 bit di parità), sviluppate nel senso longitudinale, ciascuna facente capo ad un circuito di lettura/scrittura.

Le densità di registrazione erano di 800, 1600, 6250 bpi, la capacità di una bobina (2400 piedi) di circa 140 MB. Poiché all'epoca i nastri venivano anche usati come memorie in linea, e quindi l'accesso era tipicamente al singolo blocco, particolari accorgimenti meccanici erano usati per garantire buone prestazioni (rispetto alla tecnologia dell'epoca) sulle proprietà start-stop del sistema (dimensionamento dell'inter-record gap, sistema di controllo a vuoto spinto per attenuare l'effetto delle forti accelerazioni, sistema per tenere il nastro a minima distanza dalle testine senza toccarle, etc.).

5.3 Nastri a serpentina

La tecnica di registrazione a serpentina è una evoluzione di quella longitudinale a 9 tracce: lungo il nastro è registrata una traccia (o un gruppo di tracce), poi le testine di lettura-scrittura vengono spostate trasversalmente, invertono la loro direzione e seguono un'altra traccia ancora lungo il nastro. Sul nastro si contano dunque parecchie tracce (ad esempio 48 oppure 256); in qualche caso esistono più gruppi di lettura-scrittura, in modo da ridurre l'escursione trasversale delle stesse. In ogni caso, si riducono così i tempi di accesso ai file.

I primi prodotti di questo tipo nascono presso i venditori di mainframe (IBM, Fujitsu, Storage Tek, etc.). Oggi sono ancora prodotti dagli stessi, ma anche da altri produttori specializzati.

Ad esempio, un "classico" è il nastro IBM 3480, del 1984, con 18 tracce, 200 MB). Nell'attuale produzione, il prodotto-top della IBM è la serie Magstar 3590 (nata nel 1995), a 128 o 256 tracce e testina magnetoresistiva che legge 16 tracce in parallelo, con capacità fino a 20 GB nativi e data transfer rate fino a 14 MB/s, destinata alla fascia alta di computer, dai mainframe fino ai server Windows NT.

Una serie di drive similari, prodotti da tutte le ditte di computer e specializzate nelle memorie, è costituita dai cartridge *DLT*, *Digital Linear Tape*, che allo stato sono nella fascia di prodotti-top fra le unità a nastro: sempre a ½", capacità nativa fino a 40 GB ed oltre, data transfer rate di 3-6 MB/s.

Ad esempio, Quantum DLT 8000: 208 tracce, nastro lungo 1828 piedi, capacità nativa 40 GB (80 GB con compressione), densità lineare di registrazione 98 kbp, data transfer rate 6 MB/s, operante con interfaccia SCSI-2.

A costi più contenuti e ovviamente a prestazioni più basse fanno riferimento le cassette *QIC*, *Quarter Inch Cassette*, esternamente di due dimensioni: 5.25” e 3.5”. Inizialmente disponibile da molti venditori come supporto per PC, spinto poi da 3M sul mercato di media potenza con la stima di 100 GB/cassetta per il 1999 (che, però, fino ad oggi non si è vista). Molti sviluppi si trovano su questa linea.

Ad esempio, Travan (™ di 3M) è una famiglia di minicassette che fornisce capacità addizionali usando una maggiore lunghezza (229 metri) ed una maggiore larghezza (.315”=8mm) dei QIC. I drive prodotti oggi leggono QIC e Travan indifferentemente.

Come altro esempio si ricorda il Tandberg Mod. SRL6: 144 tracce, nastro lungo 2000 m, capacità nativa 12 GB (24 GB con compressione),

5.4 Nastri a scansione elicoidale

Queste apparecchiature, derivate da tecnologie audio e video, sono realizzate mediante una testina costituita da un cilindro che ruota velocemente in direzione trasversale del nastro, mentre questo si muove longitudinalmente; l’effetto è che le tracce si dispongono diagonalmente lungo il nastro. La velocità di rotazione della testina contribuisce a rendere elevata la velocità di trasferimento dei dati, il relativamente lento movimento longitudinale preserva il nastro dalla usura.

Appositi artifici sono inoltre usati per rendere veloce il tempo di ricerca di un file lungo il nastro: vengono a tale scopo adoperate apposite tracce longitudinali oppure appositi indici scritti sulle tracce diagonali, che in ogni caso vengono letti a velocità di 50-100 volte superiori a quella di lettura dei dati.

Nati attorno al 1990, questi nastri ebbero subito un gran successo a causa del basso costo ed alta capacità, sono stati superati dai DLT per qualche tempo, ma restano competitivi con questi.

I primi drive di questo genere sono i *DAT* (*Digital Audio Tape*), derivati appunto dalle audiocassette, poi standardizzati come *DDS* (*Digital Data Storage*), e classificati in differenti classi (DDS-DC, DDS-2, DDS-3, un recente annuncio HP promette un DDS-4). Il prodotto originale aveva 1.3 GB su un nastro di 60 metri a circa 180 kB/s.

Fra i prodotti attuali si citano il Sun StorEdge di 12 GB, nastro di 125 m, 1 MB/s e l’annuncio HP del DAT 40 di 20 GB nativi.

I nastri “8 mm” derivano dalla tecnologia video e costituiscono un prodotto con prestazioni superiori. Sony sviluppò un trasportatore inizialmente destinato

al mercato video domestico. Exabyte ne acquisì i diritti di riproduzione negli USA.

Oggi il drive Sun Storage di questa categoria ha capacità di 20 GB, lunghezza di 170 m, velocità di 3 MB/s, tempo di ricerca di un file di 72 s

Nel 1994, Sony sviluppò un trasportatore inizialmente destinato al mercato video professionale, basato su un nastro a 19 mm, tratto dalla tecnologia broadcast televisiva ed adattato alle esigenze digitali con l'aggiunta di una traccia di controllo longitudinale e di codici ECC. Nacque così la classe delle memorie a nastro a 19 mm. Prodotti in diverse varietà (D-1, D-2, etc.) e destinati al mercato dei supercomputer, presentano capacità dell'ordine di 25-175 GB e velocità di trasferimento dell'ordine di 8-50 MB/s.

5.5 Nastri ottici trasversali

Nella fascia altissima di costo e di prestazioni si trova un prodotto a tecnologia ottica e tracce disposte trasversalmente sul nastro: il movimento longitudinale è solo per le operazioni di posizionamento, quello trasversale delle testine per la scrittura e lettura.

In particolare, il drive EMASS/CREO è un nastro di 35 mm con capacità di 1 TB e velocità di 3 MB/s

6. Memorie a dischi: caratteristiche meccaniche

I modelli di unità a disco variano dagli economici dischetti flosci (floppy) dei PC, ai dischi rigidi (hard disk, HD) degli stessi PC, delle workstation oppure anche dei grossi sistemi, composti da uno o più piatti montati coassialmente, di dimensioni variabili e che tendono sempre più a miniaturizzarsi.

Il supporto è in materiale metallico, vetro o plastico ricoperto da un film sottile di materiale magnetico su una o su entrambe le facce, con le tracce per la registrazione di forma circolare e concentriche. Il numero di tracce ed il numero di bit per traccia sono variabili con le dimensioni ed i modelli: le capacità complessive vanno dall'ordine di grandezza del Mbyte per i più piccoli dischetti alle centinaia di Gbyte o al Terabyte per i grossi sistemi RAID.

Sul piano funzionale, le unità a dischi sono *rimovibili* (i dischetti e alcune unità realizzate per il backup) oppure *fissi*: nel primo caso, il disco è asportabile e funge da archivio fuori-linea, nel secondo è incorporato nel sistema di calcolo ed occorre un altro sistema di memoria di massa per l'archiviazione.

In ogni caso, è disponibile in genere una testina di lettura-scrittura per ciascuna superficie, e quindi 1 per sistemi monodisco a singola faccia, 18-20 per

i sistemi con 10 piatti impilati (per motivi meccanici talora si preferisce non usare le superfici esterne della pila). Le testine di lettura-scrittura, pilotate da un sistema di selezione, sono mobili radialmente, in modo da poter raggiungere le singole tracce. Come si vedrà meglio in seguito, questo movimento delle testine costituisce un vincolo alla realizzazione di sistemi veloci (vedi seek, § .7.1).

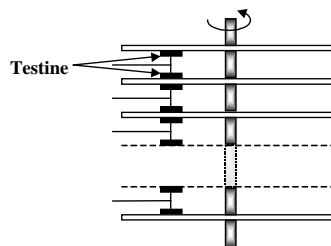


Figura 6.1 - Dischi a testine mobili

Nel caso di dischi rimovibili il meccanismo deve anche spostarsi al di fuori della pila di dischi per permetterne la sostituzione (fig.6.1). Le unità a dischi fissi, invece, non dovendo curare la meccanica del blocco lettura-scrittura per la rimovibilità del supporto, possono realizzare prestazioni migliori in termini di tempi di accesso e capacità. In particolare, per i dischi rigidi è possibile realizzare un sistema secondo il quale le testine “volano” su un cuscinio d’aria creato fra superficie e testina dalla rapida rotazione del disco (principio di Bernoulli)

In ogni caso, la migliore affidabilità del sistema dovuta alla rigidità della superficie magnetica consente l’impiego di densità di registrazione, di velocità meccaniche e di velocità di trasferimento dei dati più elevate rispetto ai nastri². Rispetto a questi, inoltre il sistema di selezione è migliore in quanto quasi casuale.

Il sistema di registrazione più diffuso è quello in cui i dati sono registrati serialmente su ciascuna traccia, suddivisi in *record fisici* o *settori*. Tipicamente, un settore contiene 512 byte (ma esistono anche sistemi a lunghezza di settore variabile), una traccia qualche decina di settori. In genere, un settore è anche l’unità fisica di trasferimento dati, in qualche caso invece un blocco (unità di trasferimento) è costituito da più settori consecutivi associati in un cluster.

La pila di dischi o il singolo disco sono in genere in continua rotazione sul proprio asse; invero, talora si arrestano fra una serie di operazioni e l’altra per diminuirne l’usura, ma quando operano restano comunque in rotazione. Ciò

² Invero, vincoli alla velocità di trasmissione dei dati sono spesso dettati dalle caratteristiche dell’interfaccia cui ima memoria di massa è collegata.

consente di ridurre notevolmente il gap necessario fra una registrazione e la successiva ed elimina la problematica dell'inter-block-gap, esistente sui nastri qualora si volesse effettuare un posizionamento a livello di blocco.

7. Dischi magnetici: organizzazione dei dati

7.1 Cilindro, superficie, settore

Dal punto di vista della selezione, una registrazione è identificata da:

- un *cilindro*, che è l'insieme delle tracce di egual raggio disposte sulle diverse superfici;
- una *superficie* (una delle facce degli n dischi);
- un *settore*, cioè l'arco di circonferenza con i dati.

Il sistema di selezione è dunque un sistema multiplo:

- il cilindro è identificato da un "indirizzo di cilindro" ed è selezionato con spostamento radiale del blocco di testine (selezione spaziale dinamica, cfr. § 4.2): il tempo di accesso associato è il *tempo di seek*,
- la superficie, anch'essa individuata da un indirizzo, è selezionata mediante semplice selezione elettronica di una fra le n testine (selezione spaziale statica, cfr. § V.4.2);
- il settore è selezionato con una delle tecniche per la selezione dinamica (cfr. §§ V.4.3, V.4.4); il tempo associato determina un *tempo di latenza*.

L'indirizzo di un settore è dunque dato dalla tripla (n.ro di cilindro; n.ro di superficie, detto anche n.ro di traccia; n.ro di settore); negli ultimi tempi va diffondendosi una tecnica per un indirizzamento dei settori con un unico indirizzo non distinto nelle tre componenti.

7.2 Tempo di accesso al settore

Il tempo di accesso ad un settore si compone dunque di tre componenti:

- tempo di ricerca del cilindro (tempo di seek);
- tempo di posizionamento sul settore (tempo di latenza);
- tempo di selezione della superficie, trascurabile rispetto agli altri.

Al tempo di accesso va poi aggiunto il:

- tempo di trasferimento del blocco.

In particolare, vengono definiti i seguenti parametri:

- *Tempo minimo di seek* TS_{min} : è il tempo necessario allo spostamento delle testine da un cilindro a quello adiacente, composto da una costante a , dipendente dalla inerzia, e da un tempo, b , del moto fra le due tracce:

$$TS_{min} = a + b$$

- *Tempo massimo di seek* TS_{max} : è il tempo necessario allo spostamento dal primo all'ultimo cilindro; se n sono i cilindri, è pari a:

$$TS_{max} = a + (n-1)b \quad (8.1)$$

- *Tempo medio di seek* TS : è il tempo medio di spostamento radiale, convenzionalmente fissato in³:

$$TS = a + bn/2 \quad (8.2)$$

- *Velocità di rotazione* n_0 , espressa in giri/min (rpm); ovviamente, il periodo di rivoluzione t_{rot} è:

$$t_{rot} = 1/n_0.$$

- *Tempo medio di latenza* TL : è il tempo necessario per raggiungere il settore, funzione della velocità e della tecnica di selezione angolare; per una tecnica associativa (cfr. § V.4.4), può essere assunto pari alla metà del periodo di rivoluzione.

$$TL = t_{rot} / 2 = 1 / (2 n_0) \quad (8.3)$$

- *Velocità di trasferimento dati* V , espressa in Mbyte/s detta anche *data transfer rate*; ovviamente, il tempo di trasferimento di un settore di B byte è:

$$TT = B/V \quad (8.4)$$

- *Tempo medio di accesso*: è la somma dei tempi medi di seek e di latenza

$$TA = TS + TL \quad (8.5)$$

³ Si è a lungo discusso e si discute ancora sul modo migliore di calcolare questo valore medio: il caso migliore è 0 (la testina sta sul cilindro desiderato), un caso abbastanza favorevole è quello del tempo minimo, il caso peggiore è quello della escursione massima. Una teoria sostiene che la media debba ritenersi pari all'escursione di 1/3 delle tracce, in quanto la differenza attesa fra due variabili aleatorie uniformemente distribuite in un intervallo di tempo è appunto 1/3.

7.3 Struttura del settore

Per la registrazione di un settore si usa oggi in prevalenza la tecnica già indicata per i nastri b (cfr. § 5.1): dopo lo *inter-sector gap* è registrato un *header*, contenente byte di sincronizzazione ed indirizzo, al quale seguono i bit utili e una coda (*trailer*) contenente *checksum* e codice *ECC*.

Questa tecnica di registrazione è anche detta di *dischi a settori software* (*soft sectorized*), in contrapposizione con una tecnica ormai abbandonata di dischi a settori hardware (*hard sectorized*), nella quale le tracce erano fisicamente divise in settori attraverso appositi segnali ottenuti su tracce di riferimento, che indicavano altresì la marca di "inizio traccia". La tecnica software fa riferimento alla selezione con informazioni di riferimento (cfr. § 4.4), quella hardware al metodo della selezione temporale (cfr. § 4.3).

Con la tecnica a settori software è possibile realizzare sia settori di lunghezza fissa sia variabile; essa è inoltre flessibile e veloce: la flessibilità le deriva dal fatto che nella testata del settore possono essere anche registrate informazioni quali "settore guasto", "prossimo settore", etc.; la velocità deriva dal fatto di essere una tecnica associativa (cfr. § V.4.4) e quindi di essere caratterizzata da un tempo massimo di latenza corrispondente ad 1 periodo di rivoluzione piuttosto che a 2 (cfr. § V.4.3).

7.4 Formattazione

Un disco, prima di essere usato, va "inizializzato" o "preparato" o "formattato"; questa operazione consiste nel registrare sulla superficie tutte le informazioni ausiliarie necessarie. Per la tecnica *soft sectorized* vanno in particolare definiti i settori, individuati i settori e le tracce eventualmente guaste e così via. In dipendenza, inoltre, dagli standard e dal sistema operativo, in fase di formattazione possono essere registrate altre informazioni utili, come ad esempio il nome del volume, l'ampiezza dei settori, le caratteristiche logico-elettroniche del supporto.

7.5 Densità di registrazione

Il miglioramento delle prestazioni dei dischi è legato alla densità di registrazione sostenibile: da un lato interessa migliorare la densità lineare di registrazione (bit per pollice, bpi) per migliorarne la velocità di trasmissione, dall'altro la densità radiale (tracce per pollice, tpi) per migliorarne la capacità. Si usa anche indicare la *densità superficiale*, espressa in bit/cm².

Per quanto attiene alla distribuzione delle densità di registrazione lungo la superficie del disco, occorre considerare quanto segue. Molti dischi (*hard*, *floppy* e ottici) operano in modalità CAV (*Constant Angular Velocity*), con le

tracce interne ed esterne contenenti il medesimo numero di settori: le tracce più esterne sono quindi a minore densità lineare di registrazione, presentando in qualche modo uno spreco di spazio.

I CD ed i primi floppy Mc Intosh, invece, operano in modalità CLV (Constant Linear Velocity), cioè la cella-bit è dovunque quasi della medesima dimensione e la velocità angolare varia con la posizione della testina. Ciò produce la migliore densità superficiale in bit, sacrificando i tempi di seek poiché ogni seek richiede un aggiustamento della velocità di rotazione.

Va ora diffondendosi la modalità ZCAV (Zoned Constant Angular Velocity): sul disco sono definite alcune zone ed il numero di settori per traccia è diverso in ogni zona; in questo modo, i dati sono impacchettati a maggiore densità, trovando un buon equilibrio fra CAV e CLV. Nello ZCAV il transfer rate varia in quanto il tempo/traccia è costante e i bit/traccia variano (p.e. il drive Seagate ST12450W Barracuda varia da 68 a 113 Mbits/s)

7.6 Il drive

Il drive dell'unità a disco comprende, come per quella a nastro, unità di controllo e trasportatore. La prima comprende tipicamente un microprocessore che ne gestisce la cache, gestisce l'indirizzamento dei record, ottimizza l'allocazione dei settori sul disco e così via.

8. I dischi magnetici

8.1 Hard disk

Gli hard disk singoli (in contrapposizione con i RAID, di cui alla voce seguente) sono apparecchiature essenziali, presenti in tutti i sistemi di calcolo, dai portatili ai grossi sistemi. Si costruiscono in versione interna ed esterna alla CPU ed hanno le dimensioni fisiche di 5.25", 3.5" e 2.5". Le loro prestazioni coprono un ampio raggio e sono comunque in rapida crescita.

I modelli per PC e server o workstation di medie dimensioni variano oggi (1999) da 4 a 20 GB di capacità, con 5400-7200 rpm, tempo di seek 7-12 ms, cache di 512 kB e costo di 20 Euro/GB.

Ad esempio, nella produzione Seagate si trova il drive Medalist 17240, interfaccia Ultra ATA, capacità di 17.2 GB, 5400 rpm, tempo medio di seek di 9.3 ms ed il Medalist Pro, interfaccia Ultra SCSI, capacità di 9.1 GB, 7200 rpm, tempo medio di seek di 9.5 ms, destinato a server e workstation di medie dimensioni.

Le prestazioni dei dischi per grossi sistemi e file server non si discostano sostanzialmente da quelle per i PC, ma si raggiungono i 50 GB di capacità, i 5

ms di seek time, i 10.000 rpm e i 4MB per la cache. È da segnalare la significativa crescita della densità superficiale di registrazione, che negli ultimi anni è aumentata anche del 60% all'anno. È del 1999 un annuncio IBM che ha raggiunto i 20 GB/inch² (contro i 3-5 GB/inch² attuali), con 490.000 bpi e 41.400 tpi. È ancora da notare che la velocità di trasferimento dati è fissata dall'interfaccia e varia da 3 a 80 MB/s.

Ad esempio, il drive Ultrastar 36ZX della IBM presenta diverse interfacce, capacità di 136.7 GB (su 10 piatti e 20 testine), 10.000 rpm, tempo medio di seek di 5.4 ms

Per i vecchi mainframe si usava tipicamente l'acronimo DASD (Direct Access Storage Device), ma ne è poi stato coniato un altro: SLED: Single Large Expensive Disk.

8.2 *Array di dischi*

Un *array di dischi* è una schiera di dischi opportunamente organizzati in modo da essere un sistema di dischi. In particolare, il più diffuso sistema, definito anche dagli standard, è il *RAID (Redundant Array of Inexpensive Disks)*, che è appunto un array di dischi visto come un unico volume di dati dal sistema centrale, ma in effetti composto da qualche decina di dischi (di 5.25" o 3.5") pilotati da una unità di controllo "intelligente". Il sistema è realizzato secondo possibili "livelli" di RAID che vanno dal più semplice (livello 0) nel quale un file è distribuito su più dischi e viene utilizzato il parallelismo fra i rispettivi drive per migliorare i tempi di accesso, a quelli più sofisticati, nei quali si utilizza la ridondanza fra i dati per consentire la massima affidabilità al sistema. In particolare, si usa la tecnica del *mirroring* (livello 1), secondo la quale ogni file è duplicato su un diverso disco e diverse tecniche di codici ECC (livelli 2-5), che con una minore ridondanza (20-30%) consentono affidabilità anche superiori. Spesso questi sistemi sono inoltre dotati della proprietà dello *hot swapping*, cioè la possibilità di sostituire in esercizio un drive malfunzionante. Altri livelli (RAID6, RAID7) sono evoluzioni di quelli fondamentali sopra definiti oppure sigle commerciali

Ad esempio, il sistema SUN StorEdge A5000 realizza un RAID di 509 GB mettendo insieme 56 dischi da 9.1 GB ciascuno.

Esistono anche soluzioni software che realizzano la tecnica RAID, ovviamente con minori prestazioni e costi.

8.3 *Dischi rimovibili*

Fra i dischi rimovibili vanno in primo luogo annoverati i classici dischetti da PC nei loro formati:

- 8", con rivestimento di plastica, floscio, in disuso da tempo;

- 5.25”, con analogo rivestimento, in disuso;
- 3.5”, con rivestimento rigido, ma pur sempre di materiale floscio.

La capacità standard dei dischetti da 3.5” è di 1.4 MB, la struttura è di un’unica superficie a doppia faccia. Tipicamente sono caratterizzati da una densità di traccia di 135 tpi ed una densità lineare di 0.8 MB/s , una velocità angolare di 300 rpm e tempo medio di accesso di 500 ms.

Le minori prestazioni rispetto ai dischi rigidi dipendono dal fatto che sono più lenti in quanto le testine di lettura-scrittura non “volano” sulla superficie del disco, ma vi si pongono a contatto.

La loro capacità, un tempo notevole, oggi è insufficiente per usarli, come una volta, come supporti per la distribuzione di software (oggi in gran parte sostituiti dai CD-ROM) oppure come supporti da backup (sostituiti dai nastri o dagli HD rimovibili); tuttavia, sono ancora ampiamente usati per trasferire dati fra i sistemi o per piccoli archivi.

Questo motivo ha suggerito ad alcuni produttori la produzione di *dischetti ad elevata capacità*, che mantengono la compatibilità in lettura dei floppy da 1.4 MB, con tecnologie varie che consentono di aumentare la densità di registrazione. Fra questi, si segnalano alcuni prodotti detti *floptical* per l’uso di tecnologie miste magnetiche ed ottiche, che tuttavia fino ad ora non hanno avuto la diffusione prevista a causa dell’elevato costo. Sono però recenti annunci di prodotti a prezzi competitivi con le altre memorie di backup.

Ad esempio, un floppy disk di tal genere da 120 MB è prodotto da Iomega con data transfer rate di 4 MB/s e costo di circa 150 Euro.

La storia degli hard dischi rimovibili è interessante. Negli anni ’60 e ’70 si può dire che tutti gli hard dischi erano rimovibili: si trattava di grossi padelloni (*disk-pack*) estraibili che avevano tanto la funzione di archivi in linea quanto quella di archivi di backup. I dischi rimovibili furono poi sostituiti dai dischi fissi, più efficienti, ed ai nastri fu data la funzione di memorie di backup. Relativamente da pochi anni, sono stati rilanciati gli HD con funzioni di backup. I più diffusi sono allo stato i cosiddetti Zip e Jaz della Iomega e i Syquest, con capacità di 100-200 MB.

9. Dischi ottici e magneto-ottici

9.1 Dischi ottici

Un disco ottico è una apparecchiatura simile al disco magnetico, salvo l’uso di principi ottici piuttosto che magnetici per la registrazione e rilevazione dei bit

registrati: i bit sono registrati in forma di “macchie nere” dette *pit*, lette da un raggio laser e misurando l’intensità della luce riflessa

I dischi possono essere di tipo:

- ROM (cfr. § V.7.1): usano una superficie che funziona come uno specchio, con apposite fossette o buchi “stampate” su di essa;
- WORM (§ V.7.2): scrivono un pit per ablazione, bruciando un pit sulla superficie attraverso un raggio laser;
- read/write (§ V.7.3): usano tecnologie reversibili di registrazione, quali quella del *phase change* (che distingue la registrazione in una forma cristallina ed una amorfa, commutabili con l’effetto laser) o quella del *dye-polimer* (che gioca sulla espansione di un polimero che copre una sostanza colorante).

Vengono usate due tecniche di organizzazione delle tracce: quella a tracce concentriche e velocità angolare costante (CAV) tipica dei dischi magnetici, con le varianti dell’organizzazione a zona (ZCAV, cfr. § 7.5) e quella a velocità lineare costante (CLV), nella quale esiste un’unica traccia a spirale che si sviluppa su tutta la superficie.

I più diffusi tra i dischi ottici sono i CD-ROM (per il loro uso sui PC) e i CD da questi derivati, CD-R e CD-RW; la tecnologia dei supporti è diversa:

- Il CD-ROM è composto di uno strato di alluminio sul quale vengono “stampati” i pit; sono caratterizzati da un alto costo del complesso masterizzatore (la macchina che ne produce il master) e stampante (la macchina che ne realizza le copie) e da un basso costo delle copie se stampate in numero elevato); è un supporto ampiamente usato per la distribuzione di software, enciclopedie multimediali, etc.
- Nei CD-R (*Recordable*) il foglio di alluminio è sostituito da uno strato d’oro ricoperto di pittura traslucida che viene “bruciata” in fase di scrittura; la macchina per scrivere i CD-R si dice ancora masterizzatore, anche se produce direttamente il dischetto; il suo costo è contenuto, il costo dei dischi vergini è maggiore di quelli a sola lettura; sono particolarmente adatti per operazioni di archiviazione e per riproduzioni in poche copie.
- I CD-RW (*ReWritable*) usano la tecnologia del phase-change; il masterizzatore ed il supporto sono più cari (oggi, un supporto CD-RW costa circa 5 volte un CD-R).

La capacità tipica di un CD è di 650 MB, i tempi di accesso sono lenti (50-150 ms), le velocità di trasferimento dati, usualmente espresse in multipli della velocità della porta parallela del PC (150 kB/s) sono da 16x a 40x. (i tempi di scrittura sono maggiori di quelli di lettura, quelli di riscrittura ancora maggiori). Le prestazioni più ridotte in termini di capacità e velocità rispetto ai dischi

magnetici sono ripagate dalla rimovibilità del supporto (soltanto di recente si sono sviluppati dischi magnetici rimovibili a 100-200 MB (cfr. Zip e Jaz, § 8).

Ad esempio, lo HP CD-Writer 8100i Plus ha 4x in scrittura su CD-RW e CD-R, 2x in riscrittura, 24x in lettura; il tempo di accesso medio è di 150 ms.

I DVD (Digital Versatile Disk o Digital Video Disk), derivati dalla tecnologia video, costituiscono i CD-ROM del futuro. A uno o due strati, singola o doppia faccia, la capacità assume valori che vanno dai 4 ai 17 GB, velocità di trasferimento di 6x ed oltre. Oltre alla versione ROM è già disponibile sul mercato la versione riscrivibile, detta anche DVDRAM

Ad esempio, il lettore DVD Korea 6X ha capacità di 3.95-5.2 GB, transfer rate di 7.2 MB/s, tempo di accesso di 90 ms.

In laboratorio è allo studio la tecnologia “GIANT CD-ROMs”, che forse vedrà la luce nel 2000: dovrebbe avere una capacità di oltre 165 GB su un singolo disco. L'aumento di capacità dovrebbe derivare da un miglioramento della tecnologia laser.

9.2 Dischi magneto-ottici

Si tratta di dischi di plastica o vetro rivestiti di un materiale (spesso TbFeCo) che gode di proprietà speciali. Il disco è letto mediante un laser a bassa intensità ed esaminando la polarizzazione della luce riflessa. Per scrivere, si usa una maggiore intensità per riscaldare il materiale oltre il suo punto di Curie, ove esso è sensibile ad un campo magnetico; quando il materiale si raffredda nuovamente, il suo stato è “congelato”. La polarità della luce riflessa durante la lettura dipende dalla polarità del campo magnetico sotto il quale il materiale fu raffreddato l'ultima volta. Una volta raffreddato, il materiale non è più sensibile ai campi magnetici, sicché esso può essere comparato in un certo senso con il paleomagnetismo. Per migliorare le prestazioni, spesso si usano sistemi con più testine e più raggi laser che operano in parallelo.

La soluzione magneto-ottica è oggi la preferita per la memorizzazione di grosse capacità con le caratteristiche di accesso diretto, stabilità e durata. Le grosse capacità e le velocità che tendono a raggiungere quelle dei dischi magnetici (almeno per la lettura, mentre per la scrittura il processo richiede in generale tempi maggiori) rendono gli MO particolarmente adatti ad applicazioni “data intensive”, quali quelle audio-video. In particolare, le migliorate prestazioni in termini di velocità dei prodotti moderni (3000-3600 rpm, transfer rate >5 MB/s) forniscono accesso veloce a grossi file tipici delle applicazioni di grafica, multimedialità, immagini mediche ed editing audio/video.

I dischi sono organizzati in tracce e settori, spesso con tecnica ZCAV (cfr. § 7.5), su molti formati diversi (14", 12", 5.25", 3.5", 2.5"), nelle due versioni di WORM e riscrivibili. Si noti che la versione WORM può essere funzionalmente preferita nel caso si vogliano archiviare dati che non si vuole rischiare di perdere. I drive scrivono i dischi riscrivibili e leggono anche i WORM.

Ad esempio:

- VFD 16000 di ATG è un WORM a 12" di 16 GB, seek di 90 ms, e transfer rate di 3.3 MB/s
- Sony SMO F5441 è un MO riscrivibile di 5.25", di 5.25 GB, seek di 25 ms, latenza di 10 ms, data transfer rate in lettura 4.6 MB/s, che si dimezza in scrittura.

10. Biblioteche di nastri e dischi

I supporti rimovibili a grossa capacità (fondamentalmente nastri magnetici e supporti ottici) ben si prestano alla realizzazione di grosse biblioteche di nastri e dischi con funzioni di archivi di backup o comunque di memoria "terziaria" (secondo questa classificazione, la memoria primaria è quella centrale, la secondaria è quella dei dischi in linea, la terziaria è quella degli archivi ad un livello gerarchico sopra le memorie secondarie).

I sistemi sono formati da una schiera di drive ed una schiera ancora maggiore di nastri o dischi, con un sistema automatico che seleziona e preleva dalla biblioteca il supporto desiderato. Sono in proposito adoperati sistemi robotizzati che automaticamente caricano le cassette o i dischi nei drive ed avviano la trasmissione dati: vengono detti appunto *biblioteche (library), robot, juke-box o autoloader*. Sistemi di questo genere operano come memorie terziarie "quasi in linea" con il sistema centrale.

Ad esempio, una biblioteca potrebbe essere costituita da 25 dischi MO di 2.5 GB formando così un sistema iniziale che può poi essere ampliato aggiungendovi ancora altri dischi. Oppure può essere costituita da uno "stacker" (un rack motorizzato) che raccoglie 10 cassette di 4mm, 8mm oppure DLT oppure caroselli di 100 cassette e 4 drive. Si arriva anche a sistemi con centinaia di silos e migliaia di cassette. Il controllo del sistema è spesso effettuato via software e le cassette portano una etichetta, leggibile ad esempio attraverso codice a barre, che ne identifica la data di ingresso nella biblioteca. Esistono anche biblioteche con un migliaio di CD-ROM.

Gli algoritmi di distribuzione, riordino, indirizzamento, smontaggio e caricamento dei supporti non sono banali e tendono ad ottimizzare i tempi di accesso ai dati di un supporto, che sono comunque di parecchi secondi per dischi ottici veloci, decine di secondi per le unità a nastro di buone prestazioni e minuti per le unità a nastro economiche.

Ad esempio:

- HP SureStore CD/DVD-ROM Server U4152A è un server da collegare su Ethernet di 18 GB (20 CD-ROM), che gestisce CD e DVD-ROM elettronicamente (senza robot, con 7 drive ed una memoria propria di 18 GB ove i CD sono riversati).
- Quantum DLT 4700 è un autoloader con interfaccia SCSI che gestisce 7 cartridge DLT con 1 drive, raggiungendo 140 GB nativi (280 compressi); si può realizzare un sistema di autoloader.
- TimberWolf 9740 di StorageTek è una libreria automatizzata che monta fino a 10 drive DLT 4000 o 7000 e 494 “slot” per le cartridge, per una capacità complessiva di 395 GB nativi per cartridge di 800 MB, che diventano 25 TB per cartridge di 50 GB.

11. Dischi a stato solido (SSD) e schede di memoria

11.1 SSD

Con il nome di dischi a stato solido (SSD) sono individuati alcuni dispositivi che emulano il funzionamento dei dischi, essendo peraltro realizzati con tecnologie a stato solido. Si tratta quindi di supporti RAM esterni che operano funzionalmente come i dischi, con prestazioni di velocità delle RAM. Il supporto in sé è volatile, come la RAM, ma viene anche prodotto in una versione associata ad una memoria di backup, che salva il contenuto della RAM allorché l'alimentazione dovesse venir meno: in questo caso si dice che la SSD è *non volatile*.

Il sistema SSD emula in tutto e per tutto i dischi e quindi si collega ad esempio ad interfacce SCSI.

11.2 Schede di memoria ed apparati PCMCIA

Le *schede di memoria* (Memory cards) sono memorie a stato solido come gli SSD, ma non emulano i dischi. Sono piuttosto veri e propri chip di memoria

Per i PC sono diffuse le schede con interfaccia PCMCIA, usate per applicazioni speciali oppure come supporti per lo scambio di dati con apparecchiature speciali. Delle dimensioni di una carta di credito ma di maggior spessore, sono usate ad esempio per videocamere digitali, registratori audio e telefoni portatili.

Esiste in realtà una varietà notevole di memorie a stato solido sulle quali non è possibile soffermarci, quali ad esempio ROM, schede OTP, Schede SRAM, Schede flash.

Per quanto attiene alle schede PCMCIA (cfr. § 4), una grossolana classificazione può essere fatta in funzione del loro spessore: le schede di tipo I sono memorie a stato solido usate come periferiche; le schede di tipo II contengono anche un modem (si usano soprattutto per i PC portatili), quelle di

tipo III un hard disk reale miniaturizzato (con spessore di 1,8”) da collegare ad un PC attraverso l’interfaccia PCMCIA piuttosto che attraverso la SCSI.