

6 La funzione dello storage.

La realizzazione di una Biblioteca Digitale (BD) si basa sulla creazione di un set di documenti consultabili in forma elettronica localmente o attraverso collegamenti telematici. È intuitivamente evidente che la quantità di dati generati da un processo del genere possa essere cospicua, ma per poterne dare una stima quantitativa, si esaminano le implicazioni associate alla digitalizzazione di diversi tipi di materiale.

6.1 Da dove nascono le grosse moli di dati?

Nel processo di creazione di una Biblioteca Digitale, sintetizzato nello schema a blocchi in figura 6.1, è prevista una prima fase di digitalizzazione che differisce a seconda del tipo di materiale da acquisire ma che ha come scopo comune quello di generare un set di dati digitali che rappresentino compiutamente il materiale di partenza.

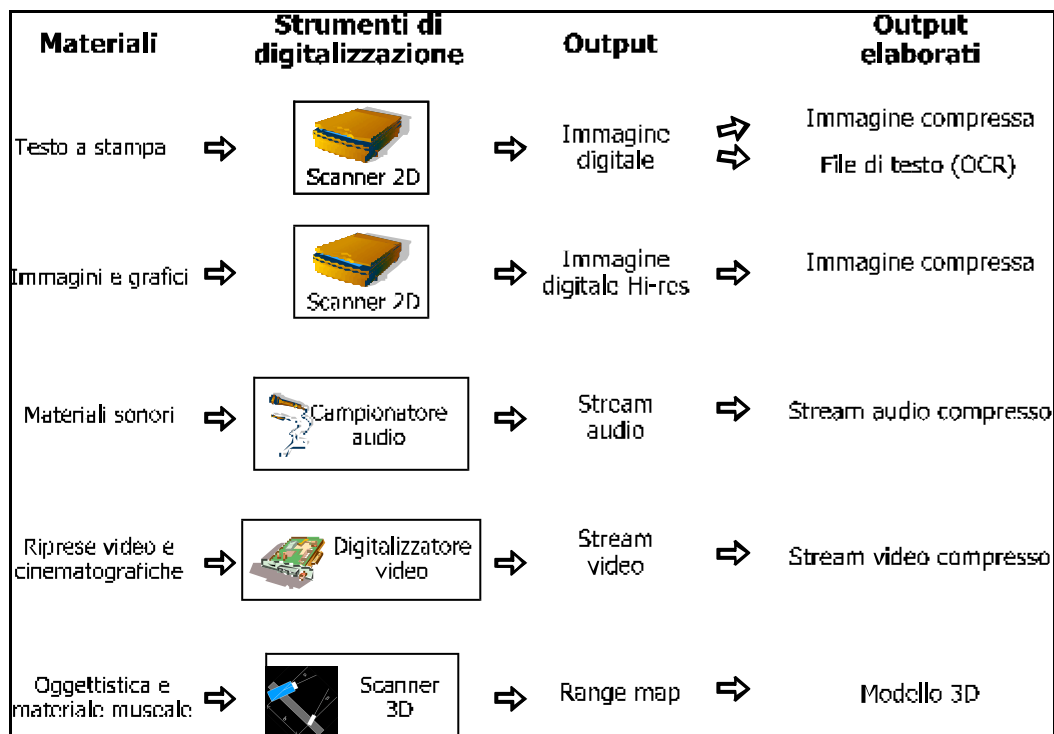


Figura 6.1 – Flusso di acquisizione dei diversi materiali: si evidenziano gli strumenti per la digitalizzazione, l'output grezzo di tali strumenti ed il dato utile da impiegare nella BD

6.1.1 La scansione 2D

Il materiale a stampa ad esempio può essere a contenuto prevalentemente testuale oppure avere una componente grafica più significativa di quella rappresentata dal testo. In ognuno dei due casi varia la strategia di scansione ed eventualmente anche il formato dei dati prodotti.

Nel primo caso la scansione sarà monocromatica ed a risoluzione non eccessivamente alta per evitare l'acquisizione di dettagli non necessari per l'interpretazione del testo. In tal caso è infatti il testo e non la sua immagine grafica quello che costituisce l'informazione utile. L'esperienza di alcuni progetti internazionali ha prodotto una serie di "best practices" che, per questo tipo di scansioni, indicano una risoluzione tipica di 300-600 punti per pollice o DPI (Dots per Inch). Nel caso di digitalizzazione di pagine testuali si deve poi tenere conto del fatto che la scansione può essere soltanto un passaggio intermedio per arrivare ad un file di testo attraverso un processo di riconoscimento automatico dei caratteri. In tal caso si impiegano programmi di OCR (Optical Character Recognition), che forniscono l'output utile del processo di acquisizione digitale. Per questo sottoprodotto il formato di memorizzazione è diverso da quello per la scansione "grezza" e la conservazione definitiva del dato può avvenire direttamente in questa forma. L'attuale sviluppo dei sistemi di riconoscimento automatico dei caratteri consente inoltre un accurato riconoscimento anche con un numero ridotto di punti per pollice. Permettendo scansioni meno onerose in termini di memoria e più rapide¹³⁵. Per immagini, disegni, incisioni ed in generale materiali in cui la componente iconografica costituisca l'elemento informativo più pregnante, la scansione sarà a colori e con risoluzione elevata, fino a 4000 DPI. L'immagine risultante non sarà sottoposta ad altre elaborazioni se non opportuni riduzioni adatte alla fruizione da parte dell'utente finale.

¹³⁵ A questo proposito vale la pena di citare il pacchetto "Active Paper Archive" della Olive Software utilizzato tra gli altri dalla sezione quotidiani della British Library. In tal caso il riconoscimento automatico fatto sul testo dei quotidiani per l'indicizzazione automatica, richiede soltanto una risoluzione di 200 DPI per fornire risultati altamente ripetibili.

Tabella 6.1 – Risoluzioni tipiche in fase di scansione

DPI	Punti per mm	Punti Per mm ²
300	11.8	140
600	23.6	558
1200	47.2	2232
2400	94.5	8928
4000	157.5	24800

Il campo di variazione del livello di risoluzione, riportato in tabella 6.1, evidenzia come, anche per i valori più bassi, il numero di punti acquisiti per ogni millimetro lineare di scansione sia elevato. Questo ci consente di effettuare la stima, attraverso un esempio, dell'occupazione di memoria di una pagina a stampa e, da lì, a quello di un numero di volumi compatibile con le dimensioni di una BD.

Per avere un'idea dell'ordine di grandezza delle moli di dati coinvolte in una BD a titolo di esempio consideriamo una pagina A4, analoga a quella su cui si effettua normalmente una fotocopia, e valutiamone l'occupazione di memoria nel caso si debbano utilizzare le risoluzioni riportate in tabella 6.1.

Le dimensioni di una pagina A4 sono 21 x 29.7 cm, che corrispondono a 210 x 297 mm ovvero 62370 mm². Il numero di punti che vengono acquisiti su tale superficie dipendono dalla risoluzione di scansione. Come già accennato sopra, a seconda del tipo di materiale si potrà passare da una risoluzione di 300 DPI ad una maggiore, con un conseguente numero di punti acquisiti che partendo da 8.7 milioni cresce con progressione geometrica al raddoppiare della risoluzione lineare, come riportato in tabella 6.2.

Una volta stabilito il numero di punti per pagina, o *pixel* (contrazione di *picture element*), si deve tenere conto che la digitalizzazione può essere portata a termine rappresentando ogni singolo punto con un numero di bit più o meno elevato. Naturalmente maggiore sarà il numero di bit per punto, e migliore sarà la fedeltà di rappresentazione (o più correttamente la gamma *dinamica*).

Tabella 6.2 – Occupazione di memoria di una pagina A4

DPI	Punti totali
300	8.7 M
600	35 M
1200	139 M
2400	557 M
4000	1.5 G

Anche in questo caso il tipo di materiale influenzerà questo parametro che può limitarsi ad 1 bit per pixel per il materiale in bianco e nero, oppure mantenersi sugli 8 bit per pixel nel caso di materiale con disegni e livelli di grigio, o spingersi ai 16 bit per pixel nel caso di fotografie in bianco e nero in cui sia importante apprezzare tutte le più lievi nuances di grigio.

Se poi la scansione deve tener conto del colore, poiché questo viene rappresentato come composizione di tre componenti fondamentali (vedi il successivo par. 7.3 per i dettagli), i valori in bit per pixel sopra riportati devono essere moltiplicati per tre. Si parla quindi di acquisizione a 24 o 48 bit per pixel.

Tornando al nostro esempio e supponendo di avere una pagina che rappresenti un'immagine in bianco e nero di elevata qualità, scannerizzata a 1200 DPI, 16 bit per pixel, si otterrà dalla tabella 6.2 il numero di punti (139 milioni), e moltiplicando per 16, il numero di bit totali:

$$139 \text{ Mpunti} \times 16 \text{ bit} = 2224 \text{ Mbit}$$

Infine, dividendo il numero totale di bit per la lunghezza in bit di un byte (8), avremo la dimensione in Megabyte della foto in formato A4:

$$2224 \text{ Mbit} : 8 = 278 \text{ MByte}$$

Un solo libro fotografico di elevata qualità, supponendolo formato da 100 pagine di questo genere, per essere adeguatamente rappresentato richiederebbe una memoria di:

$$100 \times 278 \text{ MByte} = 27 \text{ GByte}$$

Naturalmente una valutazione verosimile dell'impegno di memoria coinvolto nella digitalizzazione di materiale bibliografico non può prescindere dalla compressione dei dati, che verrà analizzata in dettaglio nel capitolo sui formati (7.1), ma considerando un fattore di compressione

medio attorno a 1:10, possiamo pensare di memorizzare una tale massa di dati su circa 3GByte.

Un materiale completamente diverso dal punto di vista della memorizzazione può essere costituito dalla monografia in cui la parte preponderante sia costituita da testo piuttosto che immagini. In tal caso ad una scansione a bassa risoluzione (2-300 DPI) può seguire una fase di OCR con il risultato incapsulato in forma compressa su files in formato PDF (vedi par. 7.2.4) per i quali l'occupazione di può andare dai 20 ai 200 kB per pagina a seconda del livello di grafica incorporata nel testo. Pertanto stimando una media di 500 pagine a volume si può ipotizzare ragionevolmente $20\text{kB} \times 500 = 10\text{ MB}$ per un volume di solo testo, fino ad un valore 100-200 MB per volumi con una parte grafica.

Una biblioteca di dimensioni medie dimensioni (50 000 volumi) trasformata in forma digitale, potrebbe quindi richiedere da $50000 \times 10\text{ MB} = 0.5\text{ TB}$ (TeraByte, ovvero migliaia di GigaByte) fino a $50000 \times 200\text{ MB} = 10\text{ TB}$, senza tenere conto di materiali fotografici. In quest'ultimo caso si può arrivare ad incrementare la richiesta di spazio di memoria di circa 1 TB ogni 300 volumi.

6.1.2 La scansione 3D

Sebbene ancora non sia diventata una procedura standard, la scansione digitale tridimensionale degli oggetti museali, quali statue, busti e bassorilievi, sta assumendo sempre di più un ruolo analogo a quello della ormai diffusissima scansione bidimensionale per quanto riguarda i materiali a stampa.

Una efficace documentazione delle opere scultoree è un aspetto essenziale nel campo della conservazione dei Beni Artistici. Mentre l'archiviazione di alta fedeltà delle opere pittoriche ha subito una decisiva evoluzione con l'applicazione delle tecnologie di acquisizione bidimensionale di immagini (colori, infrarossi, ultravioletto), i tradizionali metodi di documentazione di opere scultoree e architettoniche hanno subito fino ad oggi scarsi progressi, consistendo essenzialmente in fotografie e rilievi ripresi da

numerosi angoli di vista. Un archivio digitale di modelli tridimensionali di alta risoluzione e accuratezza costituirebbe un grande progresso in questo campo. Quando opportunamente mantenuti gli archivi digitali sono inalterabili nel tempo, e quindi possono essere usati come riferimento per monitorare sia il possibile degrado che il restauro delle opere. In caso di danneggiamento, le opere potrebbero infatti essere restaurate in modo da riprodurre fedelmente il modello digitale. Anche le tecniche di controllo possono beneficiare della disponibilità di archivi 3D: una periodica scansione dell'opera o di una sua parte può assicurare una rapida verifica di possibili alterazioni accidentali o dovute al degrado di qualunque origine.

Lo studio e la fruizione delle opere artistiche sono limitati dalla lontananza dei musei, dal limitato tempo di accesso, dal loro affollamento. Tale limitazione diventerà sempre più severa con l'affacciarsi al mercato del turismo artistico di popolazioni di grandi paesi in via di sviluppo. Spesso, inoltre, le opere sono collocate in posizioni non accessibili, come alla sommità di edifici monumentali, oppure sono smembrate e le diverse parti conservate in musei lontani (rimpatrio digitale). Talvolta, lavori di restauro che si protraggono per lungo tempo impediscono ai visitatori e agli studiosi di vedere opere importanti. In tutti questi casi, la disponibilità di modelli digitali offre nuove prospettive per lo studio e la fruizione remota di opere artistiche. La disponibilità di collezioni digitali tridimensionali tematiche può contribuire infine a introdurre cambiamenti significativi nel modo di insegnare la storia dell'Arte e le discipline artistiche.

Dal punto di vista delle occupazioni di memoria di materiale originato da scansione 3D così come per la scansione bidimensionale esiste una variabilità dipendente dalla risoluzione spaziale con cui viene effettuata. Come riferimento si possono considerare i risultati di alcuni importanti progetti di scansione portati a termine negli ultimi anni, che hanno evidenziato la necessità di spazi di memoria da mezzo GigaByte per statue a dimensione umana, o fino a 20 Gigabyte per statue alte cinque metri come il David di Michelangelo.

6.1.3 L'acquisizione di materiale multimediale

Per poter memorizzare in forma digitale un brano audio occorre prima provvedere alla sua digitalizzazione ovvero alla scomposizione in campioni del segnale analogico che costituisce il brano audio (campionamento), seguita dall'assegnazione ad ogni campione di un valore numerico rappresentabile con un limitato numero di bit (quantizzazione).

Dalla teoria del campionamento emerge che il numero di campioni al secondo necessari per costituire una esatta rappresentazione del segnale analogico è pari al doppio della massima frequenza del segnale analogico che stiamo campionando (teorema di Shannon).

Nel caso del segnale audio Hi-Fi la gamma di frequenze riproducibili va da 20 a 20000 Hz. Per questo per il Compact Disc Audio è stata scelta una frequenza di campionamento un po' maggiore del doppio di 20000 Hz, ovvero 44100 Hz (o campioni per secondo). Nel CD audio ogni campione viene codificato con 16 bit e vengono generati due flussi di dati distinti per i due canali stereofonici destro e sinistro. Pertanto risulta che per memorizzare un brano della durata di un minuto occorrono:

$$\begin{aligned} & (44.100 \text{ campioni/sec/canale}) \times \\ & (16 \text{ bit/campione}) \times \\ & (2 \text{ canali}) \times \\ & (60 \text{ sec}) \\ & = 84.672.000 \text{ bit} \\ & = 10.584.000 \text{ Byte ovvero circa } 10\text{MB}. \end{aligned}$$

Per questo un CD audio, che è in grado di contenere più di un ora di musica, ha dimensioni che vanno da 650 a 700 MB.

Digitalizzare una sequenza video implica la produzione quantità di dati ancora maggiore. Nello standard televisivo italiano (PAL-B) si ha una digitalizzazione senza perdita di informazione impiegando una risoluzione di 720x576 pixel ed una profondità di colore di 16 bit, il che produce una massa di dati di 1.58 MB per singolo fotogramma. Per evitare lo sfarfallio nelle transizioni tra un'immagine e la successiva lo stesso standard prevede 25 fotogrammi al secondo. Il flusso dati originato dal

campionamento di una scena video arriva pertanto alla ragguardevole cifra di $25 \times 1.58 = 39.5$ MB per ogni secondo di filmato senza tener conto dell'ulteriore "fetta" di flusso dati da dedicare all'audio.

6.2 Memorie digitali

Nonostante le tecniche di compressione dei dati possano in parte consentire una riduzione dei volumi di memoria impegnata, è evidente come nella realizzazione di una BD si debba misurare la memoria impegnata con un'unità di misura fino ad ora poco utilizzata: il TeraByte (TB), corrispondente a 1000 Gigabyte.

Per arrivare a gestire quantità così spinte di memoria occorre mettere in gioco una serie di tecnologie complementari tra loro, che sono state sviluppate negli anni per risolvere il problema dello storage già a livello di singolo Personal Computer. Analizziamo dunque gli aspetti tecnologici dello storage per arrivare ad individuare le scelte più adeguate nello sviluppo di una BD.

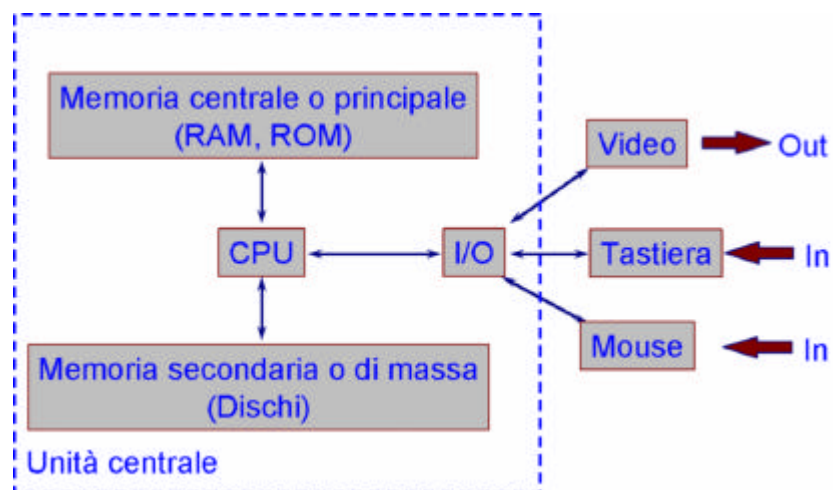


Figura 6.2 – Schema di interazioni tra la CPU e le periferiche all'interno di un Personal Computer.

Nella struttura di un PC, schematizzata in figura 6.2, si rappresenta tipicamente una CPU (Central Processing Unit) al centro di una serie di

dispositivi che consentano all'elaboratore di svolgere la sua principale funzione, ovvero eseguire un programma e presentare risultati.

Come noto un ruolo centrale è svolto dalle memorie, che vengono realizzate con differenti tecnologie a seconda del ruolo che devono svolgere.

6.2.1 Memorie volatili

La memoria in cui si alloca il programma al momento dell'esecuzione ed i relativi dati è indicata nello schema in figura come "memoria centrale". La CPU è in grado di elaborare la sequenza di istruzioni da eseguire con una cadenza scandita dalla cosiddetta "frequenza di clock" che per gli attuali processori si attesta attorno ai 2 GHz (2000 miliardi di cicli al secondo).

È evidente quindi che per fornire istruzioni alla CPU con un ritmo commisurato alla sua straordinaria velocità di elaborazione, sia necessaria una memoria da cui poter estrarre informazioni o istruzioni di programma in tempi estremamente brevi. Altro aspetto importante riguarda la possibilità di accedere alla memoria in maniera non sequenziale, in modo da consentire l'esecuzione di serie di istruzioni che prevedano salti del flusso di programma condizionati da test logici.

Questi due aspetti, legati alle caratteristiche di funzionamento della CPU, rendono necessario per la memoria centrale l'impiego di memorie RAM (Random Access Memory), che, oltre ad essere in grado di supportare sia la lettura che la scrittura di un dato, forniscono:

elevata velocità di accesso sia in lettura che in scrittura;

accesso arbitrario nello spazio di indirizzamento.

Il "prezzo" da pagare per l'impiego di questi potenti e flessibili componenti, oltre al costo elevato (400€/GB nel 2003), è la volatilità, ovvero l'incapacità di mantenere informazioni senza che sia fornita energia al componente. Per questo nasce la necessità di una memoria di massa in cui sia possibile allocare in maniera non volatile ed a costi più bassi, grosse moli di dati.

6.2.2 Memorie non volatili

Ad oggi i dischi rigidi o HD (Hard Disk) rappresentano ancora il miglior compromesso qualità prezzo (2€/GB nel 2003), con capacità di memoria che si attestano attorno agli 80GB per dischi di medie dimensioni.

Il principio di funzionamento di un disco rigido ricalca in sostanza quello di tutti gli altri supporti magnetici come ad esempio le videocassette.

Questo tipo di registrazione si basa sul fatto che il ferro, o altri materiali simili ad esso e per questo detti "ferromagnetici", quando vengono sottoposti ad un campo magnetico rimangono magnetizzati.

La magnetizzazione indotta dall'esterno ha una polarità che può assumere almeno due stati diversi. Questo suggerisce la possibilità di memorizzare informazioni binarie come quelle che compongono dati e i programmi.

Inoltre la possibilità di polarizzare il materiale in modo permanente senza che si debbano fornire ulteriori sollecitazioni fisiche o elettriche in tempi successivi a quelli dello stimolo iniziale, attribuisce a questa modalità di memorizzazione la non banale prerogativa di essere non volatile, caratteristica basilare per una memoria di massa.

In posizione intermedia tra RAM e dischi magnetici si vengono a trovare tutte quelle memorie integrate che hanno la prerogativa di non essere volatili:

- ROM (Read Only Memory), che peraltro non consentono la scrittura;
- EPROM (Erasable Programmable Read Only Memory), che consentono una scrittura solo dopo una fase di cancellazione effettuata tramite raggi ultravioletti, il che le rende inadatte all'uso corrente come strumento di memorizzazione dati;
- EEPROM, dette anche E²PROM (Electrically Erasable Programmable Read Only Memory), che alla possibilità di essere scritte dopo una fase di cancellazione, analoga a quella delle EPROM, associano la capacità di effettuare la cancellazione per via elettrica anziché tramite ultravioletti.

Le caratteristiche di quest'ultimo tipo di memorie le rende in parte simili ad una sorta di RAM non volatili sebbene con prestazioni inferiori.

Il fatto che le EEPROM siano comunque memorie integrate (ovvero allocabili su minichip), ne ha decretato il successo come supporti di memoria per applicazioni collaterali come ad esempio la fotografia digitale, in cui è necessario disporre di "serbatoi" di memoria abbastanza cospicui da contenere immagini digitali di grosse dimensioni (alcuni MB per immagine), senza vincolare il mantenimento dell'immagine alla disponibilità di una batteria carica (prerogativa della non volatilità).

Un'altra applicazione tipica delle EEPROM è quella dei memory stick, utilizzati ormai correntemente come piccoli dispositivi di storage (tipicamente da 64 a 256 MB), collegabili ad un PC attraverso interfaccia USB (Universal Serial Bus). La loro funzione è quella di un supporto di storage adatto allo scambio dati tra computer non collegati in rete tra loro.

Tabella 6.3 – Tipologie di memorie e loro caratteristiche

	Volatile	Scrivibile	Dimensioni	Tempo di accesso
RAM	Si	Si	64 kB – 4 GB	Decine di ns
ROM/EPROM	No	No	Centinaia di kB	Decine di ns
EEPROM	No	Si	Decine/centinaia di MB	Centinaia di ns
Dischi rigidi	No	Si	Decine/centinaia di GB	Alcuni ms

La tabella 6.3 riporta una sintesi riassuntiva sui diversi dispositivi di storage, con le caratteristiche salienti.

6.3 Hard Disk

Il meccanismo di registrazione magnetica digitale, sviluppato inizialmente per nastri derivati da quelli per applicazioni audio, si è perfezionato già dagli albori della IT (Information Technology) per consentire tempi di accesso all'informazione indipendenti dalla collocazione fisica all'interno del supporto.

Una cassetta magnetica può essere definita lineare, dato che le informazioni vengono registrate sequenzialmente dall'inizio alla fine; non esiste un metodo rapido per accedere direttamente a metà cassetta, ma è necessario farla scorrere fino a raggiungere il punto desiderato. In un disco che ruota, al contrario, qualsiasi zona della sua superficie può essere raggiunta direttamente dalla testina di lettura/scrittura.

6.3.1 Struttura

Per questo l'hard disk si compone di uno o più dischi solidali tra loro che ruotano attorno ad uno stesso asse con velocità costante. Sulla superficie di supporto di ogni disco, che può essere di alluminio o di vetro, è depositato un sottile strato di materiale ferromagnetico, costituito generalmente da ossido di ferro o composti simili come evidenziato in figura 6.3.

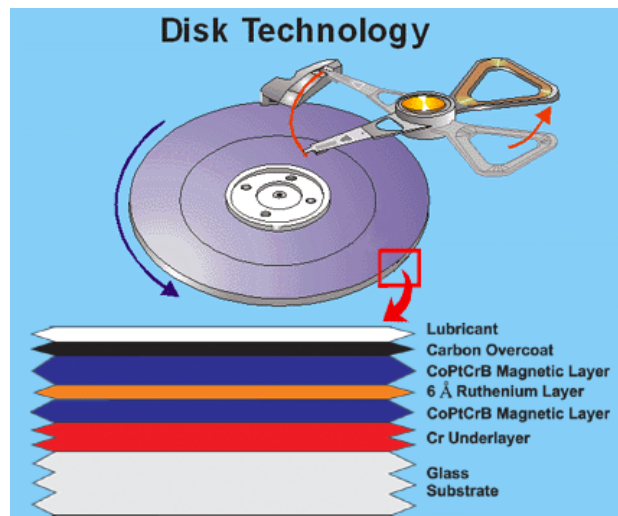


Figura 6.3 – Elementi mobili di un HD e struttura della superficie sensibile (fonte Hitachi).

In figura 6.4 è invece mostrato un HD smontato in cui si possono vedere, oltre al motore che consente di porre in rotazione il disco e l'elettronica di interfaccia (zona in basso nell'immagine), le testine di lettura e scrittura che possono traslare per raggiungere tutte le zone della superficie sensibile, dalle più interne, ovvero vicine all'asse di rotazione, alle più esterne.



Figura 6.4 – L'interno di un hard disk (IBM Deskstar 180GXP): oltre alla superficie per la memorizzazione si nota il gruppo di testine il motore e l'elettronica di interfaccia.

Poiché il raggiungimento di qualsiasi porzione del disco si ottiene combinando il movimento rotatorio del disco con quello traslatorio delle

testine, più rapido sarà il movimento di entrambi e minore sarà il tempo di accesso alle informazioni su di esso contenute. Per questo alcuni degli sforzi tecnologici per migliorare le prestazioni degli HD si sono concentrati sulla miniaturizzazione delle testine e sulla velocità di rotazione dei dischi. La miniaturizzazione delle testine ha consentito di renderle più leggere e quindi in grado di essere spostate più rapidamente a parità di energia impiegata, mentre la velocità di rotazione dei dischi è progressivamente aumentata dai 5400 RPM (Rotations per Minute) dei primi HD, ai 7200 RPM di quelli normalmente disponibili sul mercato, ai 15000 RPM degli HD più avanzati.

La registrazione magnetica digitale viene fatta suddividendo l'intera superficie del disco in tante piccole aree ciascuna delle quali viene gestita come memoria da un bit ed impostata all'equivalente magnetico di 0 oppure 1. Dato che la posizione di queste aree non è determinata precisamente lo schema di registrazione si serve di alcuni segni che aiutano il "registratore" a trovare l'esatta posizione. Questi vengono impostati nella fase di "formattazione" del disco, necessaria preliminarmente al primo utilizzo.

Come mostrato in figura 6.5 le aree sensibili sono organizzate in cerchi concentrici, o tracce (tracks), non collegati tra loro. La superficie di ogni disco è così suddivisa in tracce distinte numerate a partire da quelle più esterne verso quelle più interne. Il numero di tracce varia a seconda del tipo di disco ed a seconda della densità di tracce, identificata sulle specifiche di un HD dal parametro TPI (Track Per Inch), ovvero tracce per pollice. Tale parametro può anche essere ricavato dalla spaziatura tra le tracce (track pitch) che a sua volta dipende dalla larghezza assoluta della traccia (track width) e dallo spazio di sicurezza (guard band) che viene lasciato tra tracce adiacenti per evitare interferenze tra porzioni vicine di materiale magnetico.

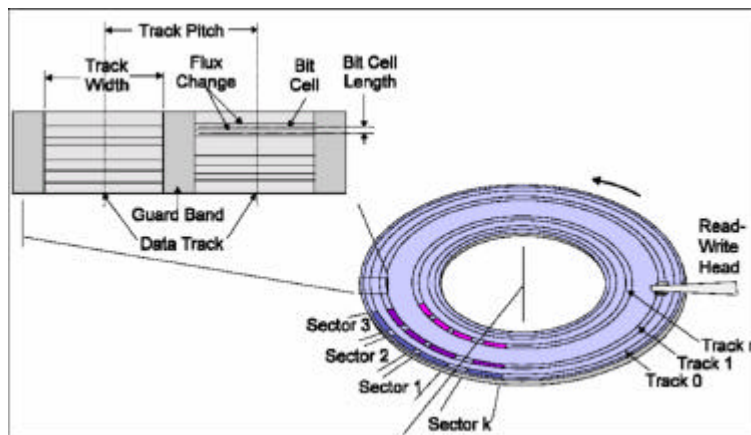


Figura 6.5 – Organizzazione delle aree magnetiche in tracce e settori, dove si evidenziano la spaziatura tra le tracce (o track pitch), la larghezza delle tracce (track width) e gli spazi di guardia tra tracce adiacenti (guard band).

Come la superficie è suddivisa in tracce, così anche le tracce sono a loro volta suddivise in più parti dette settori. Ogni settore è generalmente composto da 512 bytes, pertanto il numero di settori per traccia dipende dalla dimensione fisica del disco e dalla spaziatura tra bit adiacenti lungo la traccia, detta Bit Cell Length. Questa influenza direttamente la lunghezza del blocco da 512 bytes, come evidenziato dall'immagine in figura 6.6. I segmenti verticali chiari e scuri rappresentano una sequenza di 0 ed 1. Nella parte superiore dove i segni sono più spaziati, i bit sono scritti con la attuale densità di 6.4 Giga bit per pollice quadrato, mentre nella traccia sottostante è stata usata una densità tripla (20 Giga bit per pollice quadrato). Al momento attuale (2003) il massimo livello di densità raggiunto è di 100 Gbit/pollice²

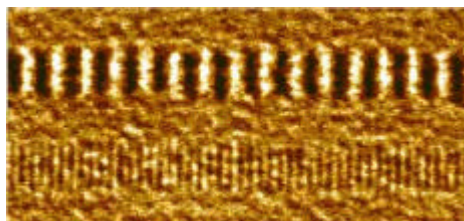


Figura 6.6 – Immagine al microscopio a scansione magnetica di due tracce scritte con una bit cell length di 6.4 Gbit/pollice² (sopra) e 20 Gbit/pollice² (sotto). Fonte IBM).

I dischi fissi, inoltre, spesso contengono più un disco (piatto) e possono quindi avere più di due lati. I lati di un disco sono ovviamente identificati da un numero che, come nel caso delle tracce, parte da zero. In alcuni

casi è conveniente far riferimento a tutte le tracce, una su ogni lato che si trovano alla stessa distanza dal centro del disco. Questo insieme "verticale" di tracce viene chiamato "cilindro". Quindi, se un disco fisso ha due piatti, ogni cilindro sarà costituito da 4 tracce.

Combinando tutte queste dimensioni, e quindi in ultima analisi il numero di cilindri il numero di piatti ed il numero di settori per traccia, si arriva alla capacità di immagazzinamento del disco.

L'evoluzione dei parametri descritti sopra ha portato all'aumento della capacità di immagazzinamento che risulta aumentata di 17 milioni di volte dalla creazione del primo HD ad oggi, come evidenziato da alcuni risultati pubblicati recentemente da IBM [0] e riportati in figura 6.7.

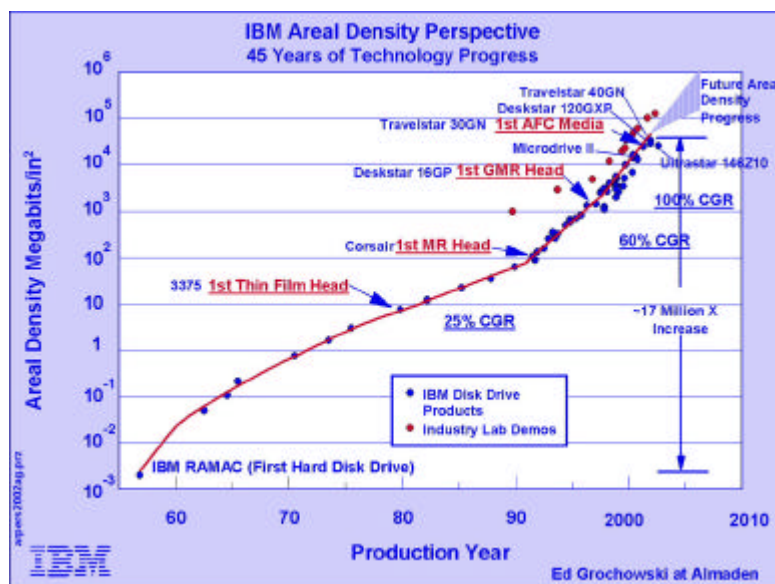


Figura 6.7 – Trend di crescita della densità di memorizzazione in Mbts per pollice quadrato, dal primo HD ad oggi (dati IBM).

Nella stessa sede è riportata un'interessante affermazione relativa al futuro dei sistemi di storage. Nonostante i grossi sforzi per produrre tecnologie alternative di memorizzazione, come ad esempio quelle non magnetiche e quelle olografiche [0], si afferma che "there are no alternative technologies which show promise for replacing hard disk storage in the next ten years".

Per valutare quali sono gli altri elementi tecnologici che caratterizzano le prestazioni di un HD si deve tenere in considerazione qual è l'iter che un'informazione compie dal momento in cui un PC effettua la richiesta, fino al momento della sua definitiva memorizzazione su disco.

Inizialmente i dati sono memorizzati in uno spazio di memoria (buffer) in cui sono codificati per renderli meno vulnerabili agli eventuali errori prodotti dal rumore elettronico. Quindi vengono selezionati alcuni settori liberi del disco e gli attuatori muovono le testine sulla superficie sensibile (tempo di accesso). Una volta che la testina è sulla traccia, questa non può scrivere dati finché i settori utilizzabili non ci passano sotto grazie alla rotazione del disco (tempo di latenza). Più è veloce la rotazione minore è la latenza. Al momento della scrittura un pattern di impulsi che rappresentano il dato codificato viene inviato alla testina che produce un campo magnetico variabile. Poiché la testina è molto vicino al disco questo altera l'orientazione magnetica del materiale sensibile su di esso depositato, registrando l'informazione.

È evidente quindi che oltre alla funzione "meccanica" svolta dal sistema disco testina, ed una elettromeccanica svolta dall'elettronica a bordo del disco per comandare motori ed attuatori di movimento, esiste una consistente parte elettronica preposta alla comunicazione tra PC e disco che analizziamo nel prossimo paragrafo.

Interfacce di connessione

Per effettuare l'invio dei dati da trasferire e l'Hard Disk vengono usate apposite interfacce indicate in genere con il termine *controller*. Si tratta di schede da inserire in uno slot del PC per consentire la comunicazione con la CPU.

Si deve tener presente che le problematiche di connessione sono diverse sull'interfaccia tra controller e PC e sull'interfaccia tra controller e Hard Disk.

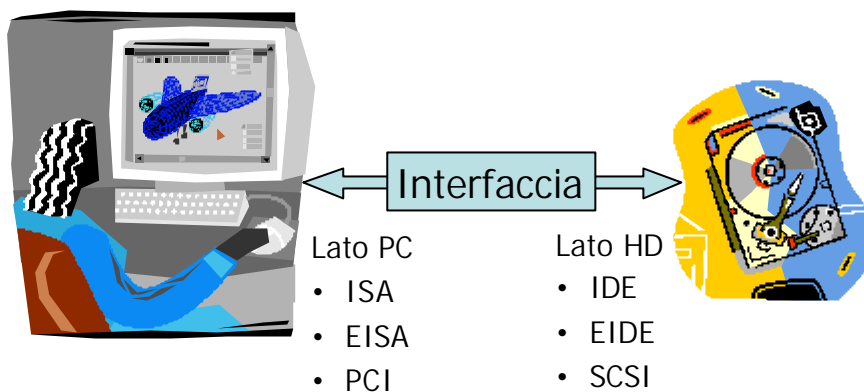


Figura 6.8 – Interfaccia PC-HD. Gli standard di interfaccia sui due lati si sono evoluti indipendentemente per far fronte alla crescita di prestazioni dei processori e degli HD.

Sul lato del controller che “guarda” il PC si ha tipicamente un’interfaccia non specializzata poiché deve consentire uno scambio di informazioni con qualsiasi dispositivo, a prescindere dalla sua funzione.

Ciò che definisce l’interfaccia è un insieme di segnali elettrici che vengono normalmente indicati con il termine BUS.

A parte le schede grafiche, che per la loro particolare natura impiegano un BUS dedicato, il resto dei dispositivi sono collegati alla piastra madre attraverso un’unica interfaccia. In passato numerosi standard sono stati proposti per svolgere questa funzione, come ad esempio ISA (Industry Standard Architecture), EISA (Extended ISA) o Microchannel. Attualmente i differenti costruttori hanno concordato un unico standard detto PCI (Peripheral Component Interconnect).

Uno degli aspetti fondamentali legati alla definizione dello standard di interfaccia riguarda le prestazioni, ovvero la rapidità con cui si riescono a far transitare dati dall’interno del PC verso i dispositivi periferici. Si è passati dagli 8 MByte/secondo del primo bus ISA (1982), ai 32 MByte/secondo del bus EISA (1988), fino ai 132 MByte/secondo dell’attuale bus PCI (versione su PC con trasferimenti a 32 bit).

Per l’impiego in un controller per HD, su cui la quantità di dati da memorizzare cresce regolarmente di anno in anno, è fondamentale che tali velocità siano il più possibile elevate, per minimizzare il tempo di accesso complessivo alla memoria di massa.

Sul lato del controller connesso con il disco l'interfaccia diviene più specializzata poiché è da questa che partono i comandi di lettura e scrittura per il disco che devono obbedire ad una logica ed ad un protocollo orientati al dispositivo specifico.

Così come per l'altro lato del controller anche in questo caso negli anni si sono sviluppati diversi tipi di interfaccia, con la differenza che in questo caso non si è giunti all'affermazione di un singolo standard. Sono due infatti le filosofie che si sono affermate: una più ampiamente diffusa ed a basso costo, e l'altra più costosa ma anche più adatta a gestire applicazioni spinte come ad esempio un file server.

La prima, denominata IDE (Integrated Drive Electronics), o più di recente ATA (AT Attachment), nella prima versione permetteva una connessione di 2 soli HD con una velocità di trasferimento tra interfaccia e periferica di 4 MB/secondo, ed una massima dimensione del disco collegabile di 512 MByte.

Dall'81 ad oggi l'interfaccia è evoluta di pari passo con l'evoluzione degli HD. Come facilmente intuibile, maggiore è la capacità di un HD e proporzionalmente maggiore è il numero di settori che lo compongono. Per cui oltre alla velocità di trasferimento altro parametro importante che incide sulla realizzazione fisica dell'interfaccia, è il numero di linee che si utilizzano per indirizzare il settore su cui memorizzare una sequenza di dati. Nell'ultima versione dell'interfaccia ATA si è giunti ad un indirizzamento a 48 bit, che consente di gestire dischi fino a 144 PB (PetaBytes=Migliaia di TeraBytes), superando così la barriera dei 137 GB imposta dall'indirizzamento a 28 bit. In tabella 6.4 è riportata l'evoluzione delle caratteristiche delle differenti versioni di bus ATA. Sebbene le prestazioni di tale interfaccia siano ragguardevoli, un aspetto limitante è costituito dal massimo numero di periferiche collegabili simultaneamente allo stesso PC, che rende questo standard più adatto per workstations personali che non per file server.

Tabella 6.4 – Evoluzione del BUS IDE/ATA

	ATA-1 (IDE)	ATA-2	ATA-3	ATA-4	ATA-5	ATA-6 (Ultra ATA 100)	Ultra ATA 133
Anno	1981	1994	1996	1997	1999	2000	2002
Max HD collegabili	2	2	2	4	4	4	4
Transfer rate (MB/s)	4	16	16	33	66	100	133
Max capacità	512 MB	8 GB	8 GB	8 GB	137 GB	137 GB	144 PB

La seconda tipologia di interfaccia è la cosiddetta SCSI (pronuncia "scasi") acronimo di Small Computer System Interface. Tale standard, introdotto inizialmente nel mondo Apple, è stato successivamente adattato nel campo dei PC.

A differenza dell'ATA nasce come bus indipendente da quello del PC per connettere diverse periferiche di memoria, come ad esempio dischi, nastri o unità CD-ROM. Questo ha implicato che, fino dalle sue prime realizzazioni, l'interfaccia SCSI permettesse di collegare fino ad 8 unità contemporaneamente. Poiché il PC è considerato nel novero delle unità collegabili (costituisce l'unità 0), il numero massimo di HD collegabili corrisponde al massimo meno uno (unità 1-7).

Come nel caso dell'ATA il trasferimento di informazioni attraverso l'interfaccia avviene in forma parallela, così come per l'interfaccia PC-stampante. Come evidenziato in tabella 6.5, a partire dalla SCSI-2 (1994), si è introdotta la possibilità di usare sia un parallelismo a 8 bit che a 16, fornendo un'utile opzione per produrre un raddoppio del transfer rate. Nell'ultima versione dell'interfaccia si prevede di aumentare il parallelismo a 32 bit, portando la velocità di trasferimento teorica ad un limite ragguardevole: 320 MByte/secondo.

Tabella 6.5 – Evoluzione del BUS SCSI

	SCSI-1	SCSI-2	SCSI-3 (Ultra SCSI)	Ultra 2 SCSI	Ultra 3 SCSI	Ultra 320 SCSI
Anno	1986	1994	1996	1999	2001	-
Max HD collegabili	7	7/15	7/15	7/15	7/15	31
Transfer rate (MB/s)	5	10/20	20/40	40/80	80/160	80/160/320

Una domanda comune riguarda la scelta dell'interfaccia: ATA o SCSI? In pratica non c'è una risposta sempre valida poiché ci sono vantaggi e svantaggi in entrambe che si evidenziano a seconda delle applicazioni. SCSI è una tecnologia più costosa e presenta qualche difficoltà aggiuntiva in fase di configurazione rispetto alla ATA. Quindi la ATA pare adatta a tutti le applicazioni di elaborazione personale d'ufficio. Le prestazioni sono superiori per la SCSI che diviene la scelta obbligata per file server o per qualsiasi attività che preveda un uso intensivo del disco, come ad esempio gestione di data base o acquisizioni video. Inoltre l'aspetto architetturale più caratterizzante della SCSI è costituito dalla possibilità da parte della CPU di inviare richieste all'interfaccia, senza gestire direttamente il protocollo di scambio dati con l'unità disco, ma demandando la richiesta all'interfaccia che si fa carico di una parte del processo di comunicazione, scaricando la CPU da questo onere. Ciò consente di sovrapporre richieste di accesso agli HD collegati sulla stessa interfaccia, cosa che diviene pressoché impossibile per la ATA quando si superino i 2 HD collegati. Questo rende la SCSI indispensabile qualora si realizzino sistemi di storage avanzati come i RAID (vedi par 6.4).

Parametri caratteristici

Come visto un HD può essere realizzato con differenti requisiti elettrici, meccanici e magnetici, ma l'utente deve essere in grado di selezionare un dispositivo senza doversi addentrare nei dettagli costruttivi.

Per questo vengono forniti una serie di parametri caratteristici che servono per aiutare un utente, come ad esempio il coordinatore di una BD che debba supervisionare un capitolato, nel complesso processo di scelta di questa delicata risorsa.

Questi sono:

- Capacità formattata. Si intende la quantità di GBytes che possono essere memorizzati sul disco;
- Tempo medio di accesso. Indica il tempo che mediamente intercorre tra una richiesta di accesso e l'arrivo del dato alla CPU. Si

compone di un tempo di latenza, legato alla velocità di rotazione del disco, ed un tempo di posizionamento delle testine;

- Velocità di rotazione. Più è elevata e minore è il tempo di latenza. Valori tipici sono 7200 RPM per dischi standard e 10000/15000 RPM per dischi SCSI di elevate prestazioni;
- Numero di cilindri.
- Numero di testine (ovvero di piani utili). In genere si possono avere da uno a 8 piani
- Numero di settori
- Tipo di interfaccia (Ultra ATA/SCSI)
- Velocità di trasferimento dati. Spesso viene data sia quella sul lato dell'interfaccia esterna, sia quella interna. È importante che quella verso l'interno sia più veloce di quella verso l'esterno
- Dimensioni Buffer. Più è grande e maggiore è il numero di dati che si scaricano sull'HD in un singolo passo di lavoro, minimizzando il tempo di trasferimento complessivo. Al momento attuale si trovano facilmente sul mercato dischi con 2 MBytes di buffer, mentre iniziano ad affacciarsi dischi con 8MBytes.
- Tracce per pollice o Track Per Inch (TPI). Tanto più questo numero è elevato, tanto maggiore è la densità delle stracce sul disco. A sua volta un elevata densità indica una tecnologia più spinta (e di conseguenza più recente).

Affidabilità

Come visto il sistema chiave su cui si basa la memorizzazione di massa è il disco rigido, che per il fatto di impiegare una tecnologia magnetica è realizzato tramite un insieme di elementi elettromeccanici connessi al mondo esterno attraverso un'interfaccia elettronica.

La gran parte degli elementi che si possono trovare all'interno di un PC, come la CPU o le memorie volatili, essendo costituite da circuiti integrati, hanno una probabilità di guasto estremamente limitata. Al contrario la forte presenza di elementi meccanici che tipicamente si trovano all'interno

di un hard disk implica una notevole riduzione di affidabilità della memoria di massa rispetto alla media degli altri componenti.

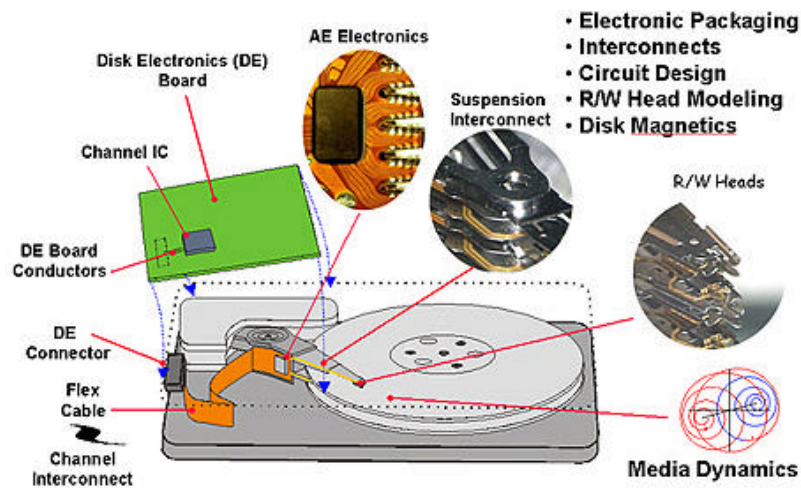


Figura 6.9 – Punti critici di un HD (fonte Hitachi).

Come evidenziato in figura 6.9 numerosi sono gli elementi critici all'interno del disco rigido che possono produrre malfunzionamenti, anche se la prima causa è costituita dall'usura degli elementi mobili come le testine o i perni di rotazione del disco che col passare del tempo possono produrre movimenti non più precisi con conseguente perdita di dati.

Per far fronte al rischio di guasto si stima l'affidabilità attraverso un parametro statistico detto Mean Time Between Failure (MTBF) che esprime il numero medio di ore tra due guasti consecutivi. I valori tipici sono nell'ordine delle 500 000 ore. In ogni caso si deve tenere presente che anche un elevato valore di MTBF non dà garanzie contro i guasti essendo un dato statistico. Pertanto un corretto approccio alla tutela delle informazioni deve prevedere una ulteriore riduzione dei rischi con l'introduzione di informazioni ridondanti che possano consentire di rigenerare l'informazione originaria nel caso di guasti dell'Hard Disk.

Informazioni ridondanti possono essere prodotte sia on-line, al momento della scrittura del dato, che off-line, eseguendo un'opportuna politica di backup del disco in modo da avere più versioni in tempi diversi delle stesse informazioni.

6.4 Ridondanza on-line (RAID)

Introdurre una ridondanza sui dati implica avere la possibilità di correggere gli errori di memorizzazione, oppure, se prodotta in misura adeguatamente elevata, riuscire addirittura a rigenerarla nel caso di eventuali guasti del disco.

Le modalità di collegamento che fanno uso di ridondanza sono state codificate e vengono identificate normalmente col termine RAID (Redundant Array of Independent Disks).

Le varie configurazioni RAID prevedono l'impiego simultaneo di più HD che lavorano congiuntamente per ridurre la probabilità di fermo macchina. Il concetto base consiste nel distribuire il dato da scrivere su diversi HD, secondo una logica che può variare a seconda del livello di ridondanza che si intende aggiungere al dato di partenza. Naturalmente più è alto il livello di ridondanza, maggiore è la probabilità di recuperare l'informazione in caso di guasto, ma per contro maggiori sono i costi di memorizzazione. Ad esempio nel caso più semplice di configurazione RAID il dato è scritto in copia su due HD identici, che vengono visti come un'unica unità logica. È evidente che a parità di spazio utile di memoria il costo è almeno il doppio di quello richiesto per memorizzare i dati su un unico HD.

In generale le tecniche RAID si realizzano soltanto con interfaccia SCSI poiché offre la possibilità di gestire l'accesso simultaneo a più dischi, operazione fondamentale per distribuire l'informazione da memorizzare, scaricando la gestione degli accessi alle singole unità sul controller e liberando la CPU da tale attività.

Esistono diverse configurazioni RAID possibili a seconda del livello di ridondanza del dato, come descritto di seguito.

RAID 0

Nella configurazione RAID 0 l'informazione da memorizzare è suddivisa in blocchi che vengono allocati su dischi diversi. Questa modalità di funzionamento, indicata anche come "stripe-set", consente un notevole miglioramento nelle prestazioni poiché il tempo necessario a scrivere un

dato viene approssimativamente diviso per il numero dei dischi che compongono il set.

In realtà in questo modo non viene creata nessuna ridondanza e la configurazione non si può dire “a prova di guasto” (fault tolerant). Un eventuale guasto su uno dei dischi implica la perdita dell'intera informazione. L'unico vantaggio nell'uso di diversi dischi implica conseguenze soltanto sulla velocità di accesso. Per questo non dovrebbe essere utilizzata per dati critici da conservare, ma solo per applicazioni ad alte prestazioni su dati temporanei.

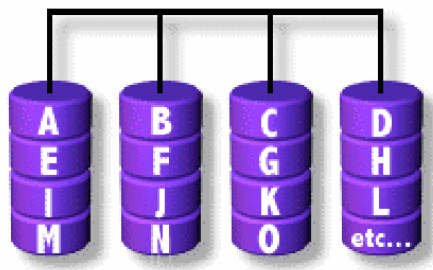


Figura 6.10 – Organizzazione delle informazioni u un gruppo di dischi in configurazione RAID 0 (Fonte: Advanced Computer & Network Corporation).

RAID 1 (Mirroring)

Con RAID 1 si privilegia la massima ridondanza delle informazioni creando coppie di HD su cui replicare le stesse informazioni. Su ogni coppia di dischi è possibile una scrittura o due letture alla volta. La velocità in lettura è quindi doppia rispetto all'impiego di un singolo disco. Grazie ad una ridondanza del 100% in caso di guasto non è necessaria nessuna complessa operazione di ricostruzione dei dati: è sufficiente utilizzare le informazioni disponibili sul disco ancora funzionante.

I principali svantaggi di questo approccio sono costituiti dall'inefficienza poiché si ha il massimo “spreco” di risorse a parità di spazio di memoria utilizzato. Inoltre l'impatto sul PC in termini di velocità di lavoro può essere pesante se si sceglie l'opzione di mirroring software offerta ormai anche

dai sistemi operativi Windows. È in generale preferibile una realizzazione in hardware.

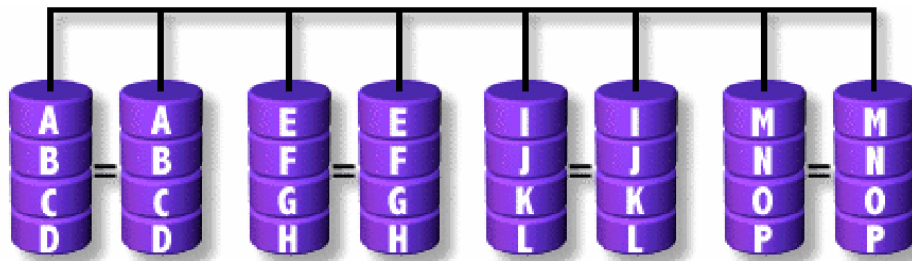


Figura 6.11 – Organizzazione delle informazioni su un gruppo di dischi in configurazione RAID 1 (Fonte: Advanced Computer & Network Corporation).

RAID 0+1

RAID 0+1 è una configurazione che prevede il mirroring di segmenti RAID 0, ed ha lo stesso livello di protezione dei dati del solo mirroring, con un aumento di prestazione legato allo “striping” ovvero alla distribuzioni di diverse porzioni del dato su dischi fisicamente differenti. Soluzione molto costosa anche se relativamente semplice da realizzare.

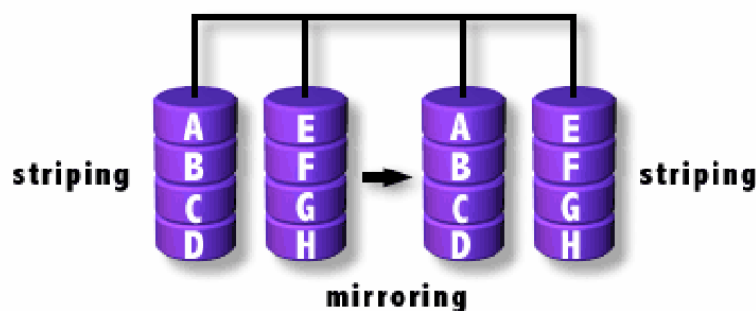


Figura 6.12 – Organizzazione delle informazioni su un gruppo di dischi in configurazione RAID 0+1 (Fonte: Advanced Computer & Network Corporation).

RAID 3

Il blocco dati è suddiviso (striped) e scritto su diversi dischi, ma al momento della scrittura viene generato un codice di ridondanza, memorizzato sul disco schematizzato in rosso nella figura, che viene

successivamente verificato in fase di lettura per scoprire ed eventualmente correggere possibili errori sui dati dovuti al malfunzionamento di uno dei dischi.

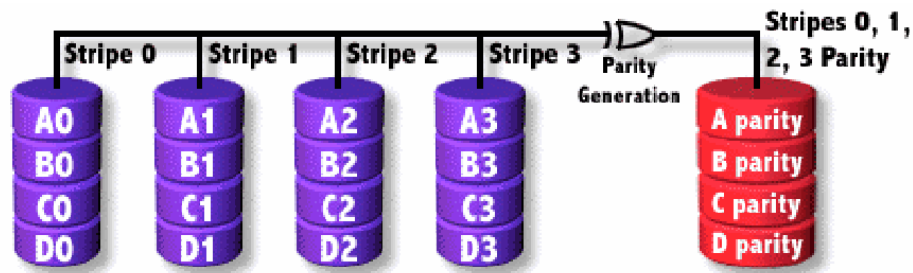


Figura 6.13 – Organizzazione delle informazioni su un gruppo di dischi in configurazione RAID 3 (Fonte: Advanced Computer & Network Corporation).

Servono almeno 3 dischi per realizzare questa configurazione.

L'efficienza è molto elevata sia in scrittura che in lettura, a patto che le attività dei dischi siano ben sincronizzate. Questo implica grosse difficoltà nella realizzazione software del RAID 3, o, per contro, la necessità di controllers piuttosto complessi per una efficiente realizzazione hardware.

RAID 5

Costituisce in un certo senso una generalizzazione del RAID 3, di cui riprende la filosofia di suddivisione in parti con l'aggiunta di un codice di ridondanza su un ulteriore disco. A differenza del RAID 3 in questo caso non c'è una specializzazione nella funzione dei dischi, in cui alcuni servono solo per i dati ed uno in particolare solo per i codici di ridondanza. Col RAID 5 la ridondanza viene infatti distribuita ciclicamente su tutti i dischi, in modo da "sparpagliare" il rischio legato ad un guasto in maniera uniforme su tutti i dischi che compongono l'array.

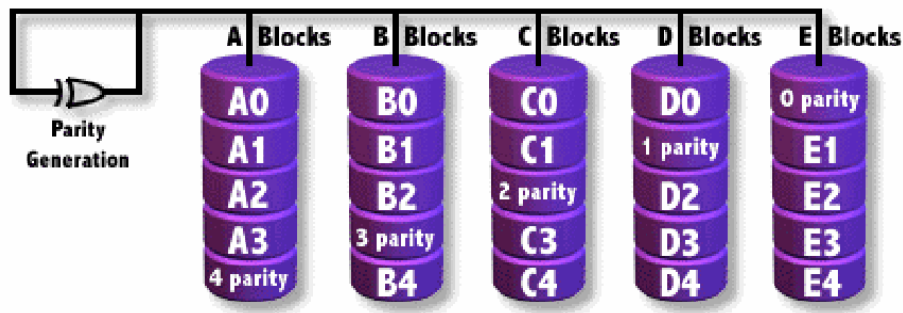


Figura 6.14 – Organizzazione delle informazioni su un gruppo di dischi in configurazione RAID 5 (Fonte: Advanced Computer & Network Corporation).

6.5 Ridondanza off-line (backup)

Un'operazione analoga a quella che si effettua imponendo una ridondanza sui dati al momento della memorizzazione, può essere realizzata creando una ridondanza in tempi distinti. Diverse strategie sono possibili in modo da coprire il periodo "storico" più lungo possibile a parità di unità di memoria impiegate (cartucce a nastro, dischi magnetottici, CD scrivibili, DVD scrivibili).

In genere si cerca di utilizzare una cadenza di backup che diminuisce all'aumentare della distanza temporale dall'istante attuale.

I media più largamente impiegati per le operazioni di backup sono quelli magnetici. Dai nastri impiegati fino agli anni '80, si è passati negli ultimi 10 anni all'impiego di cartucce a nastro digitali, che assumono forme e caratteristiche leggermente diverse a seconda dello standard a cui aderiscono. Vale la pena citare ad esempio:

- Digital Linear Tape (DLT) - Quantum
- Linear Tape Open (LTO) – HP, IBM, Seagate
- Advanced Intelligent Tape (AIT) - Sony

Sebbene tali standard rispondano più ad esigenze squisitamente commerciali piuttosto che offrire soluzioni tecnicamente complementari, l'ultimo dei tre citati (AIT) propone un'innovazione interessante che

permette di migliorare notevolmente le prestazioni durante la riletture del nastro.

Le cartucce AIT utilizzano infatti una memoria EEPROM (Electrically Erasable Programmable Read Only Memory) per conservare la directory del nastro. Questo fa sì che per ogni riletture non sia necessario – come invece accade nella maggior parte delle altre cartucce a nastro - riposizionare il nastro all'inizio per leggere la directory contenente la lista dei file e la loro posizione lungo il nastro. Questo naturalmente migliora le prestazioni del nastro ottimizzandone i tempi di posizionamento.

Gli ultimi anni hanno visto la crescita dei sistemi di memorizzazione ottici e magneto-ottici, come ad esempio:

- Compact Disk Recordable (CD-R)
- CD Recordable and Writable (CD-RW)
- Digital Versatile Disk Recordable (DVD-R)
- DVD Recordable and Writable (DVD-RW)

L'impiego di questi supporti per il backup dei dati potrebbe sembrare ottimale, vista la sempre più capillare diffusione di lettori di CD e DVD sui PC attuali, rispetto alla riletture di cartucce nastro che richiede la presenza di lettori dedicati.

Per contro studi internazionali, sembrano indicare che la dinamica di deterioramento sia decisamente peggiore per quanto riguarda i media ottici e quelli magneto-ottici, rispetto a quelli magnetici. Inoltre i media per backup si deteriorano in maniera diversa a seconda delle condizioni di conservazione: particolarmente importante è la temperatura di stoccaggio. Nella tabella 6.6 sono riportati i dati di uno studio della British Library [0] relativo al comportamento di diversi media al variare della temperatura.

Tabella 6.6 – Stima dei tempi medi di affidabilità di diversi supporti per il backup

	Temperatura di conservazione				
	10° C	15° C	20° C	25° C	28° C
Tape	50	25	15	3 anni	1 anno

	anni	anni	anni		
Cartuccia	75 anni	40 anni	15 anni	3 anni	1 anno
DVD	75 anni	40 anni	20 anni	10 anni	2 anni
CD-ROM	30 anni	15 anni	3 anni	9 mesi	3 mesi

La banda larga di Internet ha fatto nascere anche numerosi servizi on-line rivolti alla conservazione affidabile dei dati. Ad es:

<http://www.amerivault.com>

<http://www.usdatatrust.com>

<http://www.backupusa.com>

sono alcuni esempi di siti dedicati allo storage su Internet. In tal caso la risorsa di memoria viene garantita e tutto il problema della manutenzione dei dati si sposta in outsourcing.

6.6 La virtualizzazione dello storage

A causa della rapidissima evoluzione tecnologica a cui assistiamo quotidianamente, gli utenti come le BD, legati alla memorizzazione di grosse moli di dati, sono spinti a rendere sempre più scalabili e riconfigurabili le loro soluzioni di storage.

La virtualizzazione dello storage va in questa direzione ponendosi come obiettivo la separazione tra unità di memoria ed unità di elaborazione. In quest'ottica la memoria di massa viene vista come una risorsa a cui accedere, eventualmente attraverso una rete locale o geografica, piuttosto che come un pezzo del sistema di elaborazione.

Il ruolo della virtualizzazione è simile a quello che svolge un sistema operativo nei confronti dell'hardware di un personal computer, consentendo ai programmi applicativi di occuparsi soltanto di aspetti relativi alla gestione dei dati che non dei dettagli relativi a come questi vengono visualizzati o memorizzati. Quando questo processo viene messo

in atto si dice che le risorse vengono viste “ad un più alto livello di astrazione”. Quindi, in altre parole, la virtualizzazione consiste nell’astrazione dello storage.

Le sigle tipicamente utilizzate in questo contesto sono DAS (Direct Attached Storage), NAS (Network Attached Storage) e SAN (Storage Area Network), in cui significato è spiegato nei tre paragrafi seguenti.

DAS – Direct Attached Storage

Prima del 1990 la memoria di massa era una periferica collegata alla CPU internamente al computer. Solo successivamente sono iniziati ad apparire i primi Hard Disk removibili, nelle prime versioni collegati attraverso semplici interfacce non dedicate ai dischi, come ad esempio la porta parallela, nata per le stampanti.

La diffusione di questo tipo di periferiche, pur senza affermarsi in maniera completa, è andata progressivamente a crescere. Portando all’esterno anche altri elementi come l’alimentazione elettrica, si evita di limitare il numero massimo di dispositivi collegabili a causa delle limitazioni in termini di spazio fisico all’interno dell’unità centrale, o della potenza dell’alimentatore ad essa collegato. L’Hard Disk removibile è quello che ha consentito di introdurre il concetto di separazione tra unità centrale ed unità di storage, consentendo di separare anche l’adattamento alle necessità e l’investimento: in una parola la scalabilità [0].

Un’altra innovazione importante riguarda l’accoppiamento del concetto di separazione tra storage e CPU con quello di RAID per aumentare l’affidabilità. Collegare un set di dischi RAID piuttosto che un singolo disco “mostra” dal lato dell’utente sempre lo stesso dispositivo di memoria di massa, con la differenza che questa è internamente dotata di capacità di segnalazione e correzione degli errori che permettono di porre rimedio a guasti parziali del sistema di memorizzazione prima che questi divengano irreparabili.

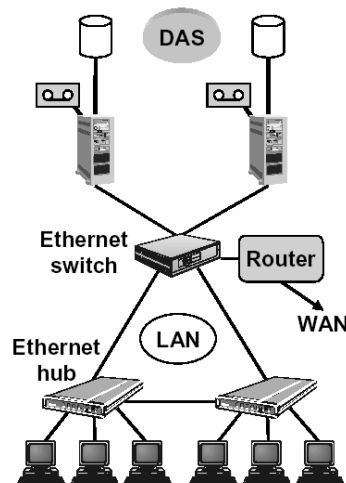


Figura 6.15 –Architettura Direct Attached Storage (DAS). Fonte McKinsey.

Dalla metà degli anni '90 le crescenti necessità di elaborazione hanno portato all'estensione del concetto di ridondanza anche alle unità di elaborazione. Server interfacciati direttamente con dispositivi di storage (DAS) in configurazione RAID e collegati tra loro attraverso una rete locale (server clusters), hanno iniziato a sostituire il singolo server (vedi figura 6.x). In questo modo è stato possibile arrivare a gestire complesse applicazioni di Data Base, come quelle tipicamente richieste dai motori di ricerca per Internet, grazie alla capacità di elaborazione parallela. Oltretutto in tal modo si è anche aumentata l'affidabilità del sistema di elaborazione complessivo che può continuare a funzionare anche in presenza di guasti di uno dei server.

In un contesto del genere la condivisione dei dati impone un enorme carico di lavoro ai server che devono dedicare una notevole percentuale del tempo di CPU a tutte le richieste di accesso ai dati provenienti dagli altri server.

Oltretutto la natura distribuita dei DAS pone significativi problemi di gestione: mantenere i sistemi di memorizzazione di pochi server connessi in cluster è senza dubbio fattibile, ma laddove i server raggiungano le centinaia quando non addirittura le migliaia, diviene pressoché impossibile.

NAS – Network Attached Storage

Le periferiche NAS sono la combinazione di una unità di elaborazione con una grande quantità di storage in configurazione RAID a cui si può accedere attraverso una rete locale. Questo insieme forma una risorsa di storage progettata per essere facile da mantenere, di elevate prestazioni, ed in grado di interagire con diversi sistemi operativi (interoperabilità).

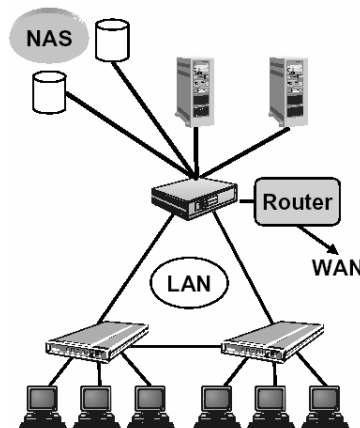


Figura 6.16 –Architettura Network Attached Storage (NAS). Fonte McKinsey.

L'elemento chiave di un dispositivo NAS è un server dedicato che agisce da interfaccia tra la rete locale e la risorsa di memoria, ottimizzato per operazioni di memorizzazione ad alte prestazioni. Questo server, denominato "filer", costituisce il "cervello" di tutto il sistema ed è inoltre il luogo dove si trova il file system, ovvero l'infrastruttura software per interagire con i files e con gli elementi che li compongono. Il file system è generalmente incorporato all'interno del sistema operativo di ogni singola unità di elaborazione che si affaccia sulla rete. Spostarlo sull'unità remota permette di ottenere una serie di vantaggi:

- Interoperabilità - i sistemi NAS sono intrinsecamente interoperabili poiché le richieste di accesso ai dati avvengono ad un più alto livello di astrazione rispetto alla scrittura su un disco fisico. Questa modalità per così dire "semplificata" di accedere ai dati naturalmente deve seguire delle regole e dei protocolli, che costituiscono il "Network file system". Sul mercato ne esistono diversi di cui due più diffusi: il Common Interface File System

(CIFS) per il mondo Windows NT, ed il Network File System (NFS) in Unix. Un dispositivo NAS come minimo supporta entrambi, pertanto una richiesta di accesso ad un NAS può arrivare indifferentemente da un PC o una workstation Unix, senza che questo cambi di fatto il meccanismo di lettura e scrittura sull'unità.

- Manutenibilità – Poiché tutte le risorse di memoria sono concentrate su un unico dispositivo di elaborazione, queste sono anche fisicamente contigue. Questo rende praticamente molto più semplice il lavoro di controllo e riparazione di eventuali guasti sul dispositivo di memoria, soprattutto in confronto ad una struttura DAS della memoria di massa, quando applicata ad una molteplicità di server distribuiti su una grossa area.
- Scalabilità – Alla possibilità di mantenere facilmente la risorsa si aggiunge quella di poterla riconfigurare aggiungendo risorse di memoria a seconda delle necessità, in maniera trasparente ai server che ne fanno uso. Su alcuni sistemi NAS quest'operazione può essere fatta anche senza spegnere l'unità di elaborazione del NAS, riducendo a zero i costi per fermo macchina.
- Facilità di backup – Analoga considerazione può essere fatta per quanto concerne la salvaguardia dei dati. La possibilità concentrare tutto lo storage in una sola unità permette una gestione semplice ed eventualmente automatizzata del backup. Tutto questo in modo trasparente ed indipendente dall'utente.

Nell'ambito di una BD una soluzione come questa potrebbe costituire un buon compromesso tra qualità e quantità di spazio di storage, e costi di gestione.

SAN – Storage Area Network

Una SAN è una rete locale ad alte prestazioni, specificamente progettata per fungere da sistema di memoria. In una installazione SAN una serie di sottosistemi di storage sono collegati tra loro attraverso fibre ottiche in modo da essere "visti" sulla rete come un'unica risorsa di memoria.

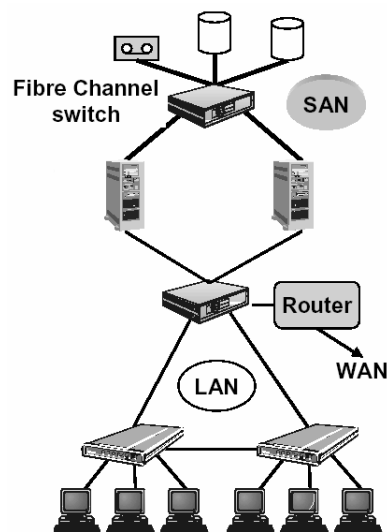


Figura 6.17 –Architettura Storage Area Network (SAN). Fonte McKinsey.

A differenza dell'approccio NAS non esiste una macchina privilegiata che opera da "filer" ed i file systems sono collocati sui server su cui viene eseguita l'applicazione.

Per evitare che file systems appartenenti a server diversi possano sovrasciversi a vicenda i dati, ogni sottosistema di memoria è partizionato in modo tale che un sottoinsieme delle proprie risorse di memoria sia assegnato rigidamente ad un server specifico. Perciò, nonostante la risorsa SAN sia vista da tutti i server simultaneamente, questa costituisce un collage di risorse di memoria dedicate.

I principali vantaggi di una SAN sono costituiti dall'affidabilità e dalle prestazioni. Rispetto ad un NAS la SAN prevede più unità di elaborazione dedicate alla gestione dello storage, e nel caso del guasto di una di queste possono essere messe in atto procedure "tampone" per evitare il fermo macchina (Backup servers). Inoltre la separazione dei flussi di dati su più canali permette una velocizzazione nell'accesso ai dati che, non dovendo "accodarsi" lungo un unico canale come succede nel NAS, possono arrivare più rapidamente a destinazione nel caso di accessi concorrenti alla risorsa di memoria.

Poiché le unità che fanno parte della rete SAN vengono generalmente collocate in un'area limitata (spesso nella stessa stanza), si hanno una

buona manutenibilità e scalabilità del sistema complessivo. Anche il backup può essere realizzato semplicemente potendo accedere ad ogni singola porzione di storage attraverso la rete locale.

Gli aspetti che possono per il momento frenare una BD dall'adottare una soluzione SAN al momento attuale sono essenzialmente tre: gli alti costi di realizzazione, l'immatunità della tecnologia SAN, e la mancanza di uno standard consolidato.

6.7 Virtualizzazione statica e dinamica

Nella virtualizzazione dello spazio di memoria si opera una ulteriore distinzione tra virtualizzazione *statica* e *dinamica*.

Nella virtualizzazione statica lo spazio di storage è tutto visibile e disponibile da parte del client. Spazio fisico e logico coincidono.

La virtualizzazione dinamica prevede che la mappatura tra spazio fisico e indirizzo logico avvenga durante l'esecuzione di un applicazione. Al momento dell'uso viene copiata la parte utile dello spazio di storage da un contenitore a basse prestazioni allo spazio di lavoro realizzato con memorie più veloci, analogamente a quanto avviene nel meccanismo di memoria virtuale dove la RAM è la risorsa veloce ed il file di "swap" su disco è il sistema d'appoggio a prestazioni inferiori.

Questo concetto può utilmente essere associato anche a meccanismi di backup estremamente elaborati, come appare in alcuni sistemi NAS dell'ultima generazione in cui ai dischi RAID che costituiscono la risorsa primaria su cui memorizzare i dati si affiancano sistemi robotizzati di cassette a nastro che rappresentano la risorsa di archivio "definitivo". Una volta che un dato viene scritto nel NAS, questo viene allocato sui dischi RAID. Quando trascorre un certo tempo senza che quel dato venga riutilizzato, questo viene "migrato" su una cassetta a nastro, mentre sull'unità RAID viene lasciato il solo riferimento alla cassetta. Man mano che una cassetta a nastro viene riempita, un robot la sostituisce con una diversa. In questo modo si può arrivare facilmente a gestire alcuni TB di

storage [0]. Soluzioni di questo tipo appaiono estremamente attraenti nell'ambito di una BD.

Riferimenti Bibliografici

H. Sunahara, M. Imai, R.S. Atarashi, H. Hida, T. Nishimura, N. Yokoyya, K. Chiara, **Beyond Digital Libraries: the 2nd generation of the NAIST digital library and its future**, *International Conference on Digital Libraries: Research and Practice*, 2000 Kyoto, pp. 101-104

Elizabeth J. Shawa, **Building a digital library: a technology manager's point of view**, *The Journal of Academic Librarianship*, Volume 26, Issue 6, November 2000, Pages 394-398.

Ogawa, S., Report on the provision of contents of Kyoto University Digital Library and its method of digitalizing rare materials, *International Conference on Digital Libraries: Research and Practice*, 2000 Kyoto, pp. 350-357

Beraldin J.-A., Blais F., Cournoyer L., Rioux M., Bernier F., Harrison N., **Portable digital 3-D imaging system for remote sites**, Proceedings of the *IEEE International Symposium on Circuits and Systems*, v 5, 1998, p V-488-V-493.

Bernardini F., Rushmeier H., Martin I.M., Mittleman J., Taubin G., **Building a digital model of Michelangelo's Florentine Pieta**, *IEEE Computer Graphics and Applications*, Volume: 22 Issue: 1, Jan.-Feb. 2002 Page(s): 59 –67

Levoy, M., Rusinkiewicz S., Ginzton M., Ginsberg J., Pulli K., Koller D., Anderson S., Shade J., Curless B., Pereira L., Davis J., Fulk D., **The digital Michelangelo project: 3D scanning of large statues**, Proceedings of the *ACM SIGGRAPH Conference on Computer Graphics*, 2000, p 131-144

Guidi G., Pieraccini M., Ciofi S., Damato V., Beraldin J.-A., Atzeni C., **Tridimensional digitizing of Donatello's Maddalena**, Proceedings of the *IEEE International Conference on Image Processing*, v 1, 2001, p 578-581

Anne R. Kenney, Louis H. Sharpe II, Barbara Berger, Rick Crowhurst, D. Michael Ott, and Allen Quirk, **Illustrated Book Study: Digital Conversion Requirements Printed Illustrations**, *Report to the Library of Congress - Preservation Directorate*, Contract #IN97C-22/97CLCCT7021 , July 1999 (<http://www.loc.gov/preserv/rt/illbk/ibs.htm>)

Abby Smith, **Strategies for Building Digitized Collections**, *Digital Library Federation - Council on Library and Information Resources*, Washington, D.C., September 2001.

Journal Storage (JSTOR). (<http://www.JSTOR.org>).

Don Williams, **Selecting a Scanner**, *Guides to Quality in Visual Resource Imaging*, Council on Library and Information Resources, 2000, (<http://www.rlg.org/visguides/visguide2.html>).

D. A. Thompson, J. S. Best, **The future of magnetic data storage technology**, *IBM Journal of Research and Development*, Volume 44, Number 3, pp. 311-321, 2000.

J. Ashley, M.-P. Bernal, G. W. Burr, H. Coufal, H. Guenther, J. A. Hoffnagle, C. M. Jefferson, B. Marcus, R. M. Macfarlane, R. M. Shelby, and G. T. Sincerbox, **Holographic Data Storage**, *IBM Journal of Research and Development*, Volume 44, Number 3, pp. 341-368, 2000.

Seamus Ross, Ann Gow, **Digital Archaeology: Rescuing Neglected and Damaged Data Resources**, *JISC/NPO Study within the Electronic Libraries (eLib) Programme on the Preservation of Electronic Materials*,

Humanities Advanced Technology and Information Institute (HATII),
University of Glasgow, February 1999.

Thomas Kraemer, James Berlino, John Griffin, Doug Haynes, Thomas Herbig, Peter Stern, Alberto Torres, **The Storage Report - Customer Perspectives & Industry Evolution**, *Joint Industry Study by McKinsey & Company and Merrill Lynch's Technology Research Group*, June 2001.

Arms, Caroline R. 2000. **Keeping Memory Alive: Practices for Preserving Digital Content at the National Digital Library Program of the Library of Congress**. *RLG DigiNews* 4(3). Disponibile all'indirizzo <http://www.rlg.org/preserv/diginews/diginews4-3.html>

Brancolini, Kristine R. 2000. **Selecting Research Collections for Digitization: Applying the Harvard Model**. *Library Trends* 48 (Spring): 783-798.

California Digital Library. 2000. **Library of Congress Grant: California Cultures**. *Executive Summary*. Oakland, Calif.: California Digital Library.

Cornell University. 2001. **Report of the Digital Preservation Working Group on Establishing a Central Depository for Preserving Digital Image Collections** (versione 1.0). Disponibile all'indirizzo: <http://www.library.cornell.edu/imls/image%20deposit%20guidelines.pdf>

Council on Library and Information Resources. 2001. *The Evidence in Hand: the Report of the Task Force on the Artifact in Library Collections*. Disponibile su: <http://www.clir.org/activities/details/artifactdocs.html>

Digital Library Federation. 2001. *Registry of Digitized Books and Serial Publications*. Available at <http://www.diglib.org/collections/reg/reg.htm>

Digital Library Federation and Research Libraries Group. 2000. *Guides to Quality in Visual Resource Imaging*. Disponibile all'indirizzo: <http://www.rlg.org/visguides/>

De Stefano, Paula. 2001. **Selection for Digital Conversion in Academic Libraries**. *College & Research Libraries* 62 (January):58-69.

Flecker, Dale. 2000. **Harvard's Digital Library Initiative: Building a First Generation Digital Library Infrastructure**. *D-Lib Magazine* 6(11). Disponibile su: <http://www.dlib.org/dlib/november00/flecker/11flecker.html>

Gertz, Janet. 2000. **Selection for Preservation in the Digital Age (Overview)**. *Library Resources & Technical Services* 44(April): 97-104.

Hazen, D., J. Horrell, and J. Merrill-Oldham, 1998. **Selecting Research Collections for Digitization**. Washington, DC: *Council on Library and Information Resources*. Disponibile all'indirizzo: <http://www.clir.org/pubs/reports/hazen/pub74.html>

Neil Beagrie, Maggie Jones, 2001, *Preservation Management of Digital Materials: A Handbook*, Published by THE BRITISH LIBRARY, ISBN 0 7123 0886 5

Terasystem (Azienda specializzata nello storage)
Web Page: www.terasystem.it