

Introduzione alle memorie

Matteo Sonza Reorda

Politecnico di Torino
Dip. di Automatica e Informatica

1

M. Sonza Reorda - a.a. 2001/2002

Introduzione

Ogni sistema di elaborazione contiene dispositivi per la memorizzazione di dati ed istruzioni.

L'insieme di tali dispositivi, e degli algoritmi per la loro gestione, costituisce il *sotto-sistema di memoria*.

Tale sotto-sistema deve essere realizzato in modo che il processore debba attendere il meno possibile per accedere a dati o istruzioni. Poichè d'altro canto il costo delle memorie è proporzionale alla loro velocità, si deve cercare un compromesso tra il costo del sistema e le sue prestazioni.

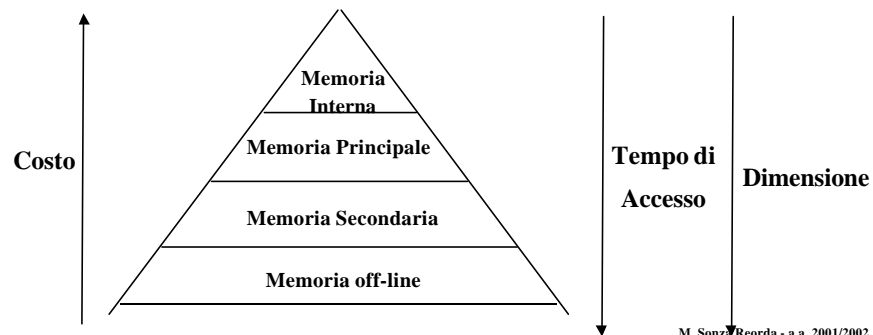
2

M. Sonza Reorda - a.a. 2001/2002

Livelli di memorie

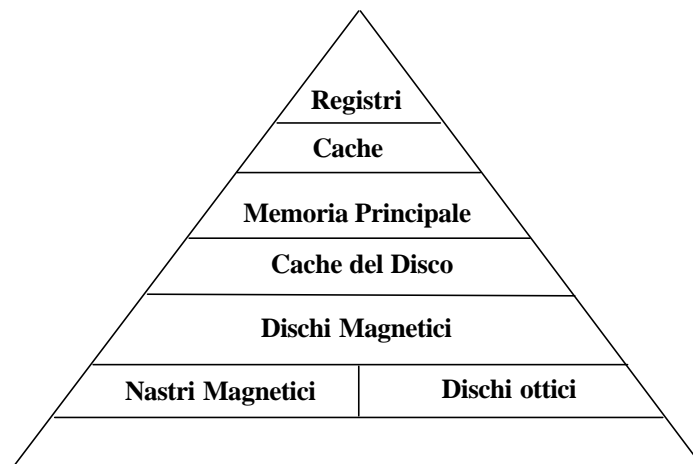
La memoria di un calcolatore è normalmente organizzata in *livelli*. Ogni livello è caratterizzato da un tipo di memoria (in termini di velocità, costo e dimensione).

In tal modo si raggiungono prestazioni comparabili con quelle della memoria più veloce, a prezzi comparabili con quelli della memoria più lenta.



3

Situazione Attuale



4

M. Sonza Reorda - a.a. 2001/2002

Memoria interna al processore

Corrisponde ai registri interni del processore.

È caratterizzata da:

- **alta velocità (comparabile con quella del processore)**
- **limitate dimensioni (al più qualche migliaio di byte).**

È di solito realizzata tramite celle di RAM statica.

5

M. Sonza Reorda - a.a. 2001/2002

Memoria Principale

Ha dimensioni molto maggiori (fino a qualche Gbyte) ma tempi di accesso più elevati.

È accessibile in modo diretto tramite indirizzi.

È normalmente realizzata sotto forma di circuito integrato.

È di solito realizzata tramite circuiti di RAM dinamica.

6

M. Sonza Reorda - a.a. 2001/2002

Memoria Secondaria

Ha dimensioni e tempi di accesso ancora maggiori.

Viene usata per memorizzare dati e programmi non immediatamente richiesti dal processore

L'accesso è gestito da appositi programmi di interfaccia.

È normalmente realizzata sotto forma di dischi magnetici.

7

M. Sonza Reorda - a.a. 2001/2002

Memoria Off-line

Permette di memorizzare grandi moli di dato (dai Gbyte ai Terabyte) con tempi di accesso elevati (decine di secondi).

In taluni casi l'accesso alla memoria off-line richiede l'intervento di un operatore.

Di solito è composta da dischi ottici o nastri.

8

M. Sonza Reorda - a.a. 2001/2002

Cache

Le *cache* sono memorie estremamente veloci che si interpongono tra il processore e la memoria principale.

All'interno di una cache risiedono temporaneamente i dati/programmi utilizzati in quel momento.

Il loro uso è trasparente al programmatore e al Sistema Operativo.

Le cache permettono di aumentare la velocità di accesso ai dati nella memoria principale senza ricorrere per essa a memorie di tipo più costoso.

Strategia Generale

Nella progettazione di un sistema di memoria vengono tenuti in conto i seguenti punti:

- conviene che il sistema complessivo sia composto da memorie di tipo e costo diversi, rispondenti ai diversi usi che della memoria vengono fatti (*gerarchia di memoria*)
- la gestione della memoria deve essere il più possibile trasparente per il programmatore e l'utente (*memoria virtuale*)
- se il sistema è di tipo *multiprocessore*, ogni processore deve poter lavorare con la memoria al massimo della velocità e senza interferire con il lavoro degli altri.

Classificazione delle Memorie

È basata su vari parametri:

- costo
- tempo di accesso
- modi di accesso
- alterabilità
- durevolezza del contenuto
- affidabilità
- tempo di ciclo e velocità di trasferimento
- caratteristiche fisiche.

11

M. Sonza Reorda - a.a. 2001/2002

Costo

Comprende anche il costo della circuiteria ed eventualmente del software per la gestione delle interfacce necessarie all'uso della memoria.

È normalmente misurato in dollari/bit o \$/Mbyte.

12

M. Sonza Reorda - a.a. 2001/2002

Tempo di Accesso

Coincide normalmente con il *Tempo di Accesso in lettura*, essendo questo quasi sempre uguale a quello in scrittura. È il tempo che intercorre tra l'istante in cui all'unità di memoria giunge la richiesta di un dato, e quello in cui tale dato viene fornito.

Il Tempo di Accesso è di solito inversamente proporzionale al costo della memoria.

13

M. Sonza Reorda - a.a. 2001/2002

Evoluzione delle Memorie

L'evoluzione della tecnologia presenta 2 tendenze:

- riduzione del *costo* per bit
- riduzione (meno marcata) del *tempo di accesso*.

14

M. Sonza Reorda - a.a. 2001/2002

Costo/Prestazioni nel 1997

	tempo di accesso	\$/Mbyte
SRAM	5-25 ns	100-250
DRAM	60-120 ns	5-10
Dischi magnetici	10-20 ms	0.10-0.20

15

M. Sonza Reorda - a.a. 2001/2002

Modi di Accesso

Sono principalmente:

- *sequenziale*: le informazioni possono essere lette/scritte solo in un ordine prefissato; è il caso delle memorie ottiche, e dei nastri.
- *diretto*: ogni blocco ha un indirizzo, e la ricerca nel blocco è sequenziale; è il caso dei dischi.
- *random*: ogni unità di dato ha un indirizzo, e le informazioni possono essere lette/scritte in qualsiasi ordine; il tempo di accesso è uguale per tutte le locazioni di memoria; è il caso delle memorie a semiconduttore;
- *associativo*: l'accesso avviene tramite un confronto tra il contenuto di ogni cella e quello specificato in una maschera; è il caso di taluni tipi di cache.

16

M. Sonza Reorda - a.a. 2001/2002

Alterabilità

Vi sono memorie il cui contenuto può essere scritto una volta sola (*Read Only Memory - ROM*). È il caso delle ROM.

In alcuni casi il contenuto di una memoria ROM può essere modificato off-line (*Programmable ROM - PROM*).

17

M. Sonza Reorda - a.a. 2001/2002

Durevolezza del Contenuto: Destructive Readout

Vi sono alcune tecnologie particolari, nelle quali l'operazione di lettura causa la cancellazione del dato memorizzato (*Destructive Readout - DRO*). In tal caso dopo ogni lettura è necessario eseguire un'operazione di riscrittura del dato.

18

M. Sonza Reorda - a.a. 2001/2002

Durevolezza del Contenuto: refreshing

- In alcune tecnologie (ad esempio quelle delle memorie dinamiche) dopo un certo tempo i bit a 1 si trasformano in 0.
- Questo effetto si ha ad esempio quando il bit è memorizzato sotto forma di carica all'interno di un condensatore, a causa delle correnti di scarica.
- È quindi necessario che periodicamente si provveda a leggere ogni bit e a riscrivere gli 1 (*refreshing*).

19

M. Sonza Reorda - a.a. 2001/2002

Durevolezza del Contenuto: volatilità

Alcune memorie, come quelle a semiconduttore, perdono il loro contenuto quando non sono alimentate (memorie *volatili*).

20

M. Sonza Reorda - a.a. 2001/2002

Tempo di Ciclo

In taluni casi (ad esempio se la memoria è di tipo DRO) la singola operazione di lettura non è terminata sino a quando non si è eseguita la corrispondente operazione di riscrittura (*ciclo di memoria*).

La frequenza massima degli accessi alla memoria è quindi diversa da $1/(\text{tempo di accesso})$, ed è data da $1/(\text{tempo di ciclo})$.

Una volta che il dato è stato estratto dalla memoria, il processore può procedere, mentre la memoria esegue la riscrittura.

La frequenza di accesso alla memoria può inoltre essere limitata dalla *banda* del bus.

Velocità di Trasferimento

Misura la velocità con cui si può accedere ad un blocco di dati contigui. Il tempo medio T_N per la lettura o scrittura di un blocco di N bit è dato da

$$T_N = T_A + N/R$$

dove T_A è il tempo di accesso medio, ed R la frequenza di trasferimento, in bit al secondo (bps).

Caratteristiche Fisiche

- **Tipo della memoria (elettronica, magnetica, meccanica, ottica)**
- **Consumo (può comportare la necessità di sistemi di raffreddamento)**
- **Portabilità**
- **Densità di immagazzinamento**
- **Affidabilità; è di solito misurata dal tempo medio tra 2 guasti (*Mean Time Between Failure - MTBF*).**