

MEMORIA PRINCIPALE

(RAM, ROM E CACHE)

La memoria centrale

- La CPU possiede un'area per la memorizzazione dei dati, di dimensioni molto limitate, dove conserva le informazioni attualmente elaborate o i risultati intermedi di un'elaborazione: **memoria cache** e **registri**. Questa piccola area di memoria non è sufficiente per eseguire programmi → I PC dispongono di vari tipi di memoria una **gerarchia di memorie** — per poter ospitare i dati e i programmi
- La **memoria centrale** (o *memoria di sistema*), cui può accedere la CPU mediante il bus indirizzi, è una combinazione di **RAM** e **ROM**
- **RAM**, *Random Access Memory* – La CPU può leggere e scrivere informazioni dalla RAM; è una memoria *volatile*, che conserva le informazioni solo finché non si conclude la sessione di lavoro
- **ROM**, *Read Only Memory* - il contenuto non può essere modificato; è una memoria permanente, poiché i dati in essa contenuti vengono conservati anche a computer spento. Il BIOS è contenuto nella ROM (controllo hardware all'accensione)
- Ciascuna posizione di memoria ha un **indirizzo** unico
- **Nota:** Tutti i programmi applicativi ed i dati attualmente in elaborazione devono essere residenti in RAM; quando è satura, parte del contenuto informativo viene scaricato in un'area di *swap* sul disco fisso
 - Ⓞ Una RAM grande permette di non ricorrere a salvataggi temporanei frequenti sull'unità di memoria di massa, operazione che comporta un sensibile decadimento delle prestazioni

Tipologie di memoria: SRAM

Static Random Access Memory, ovvero RAM statica, ogni cella è costituita da un "flip flop" realizzato da due porte logiche. Consente di mantenere le informazioni per un tempo infinito, è molto veloce consuma poco e quindi dissipa poco calore. La necessità di usare molti componenti, però, la rende molto costosa. E' solitamente usata per le memorie cache, dove elevate velocità e ridotti consumi sono caratteristiche fondamentali.





Tipologie di memoria: DRAM

Dinamic Random Access Memory, ovvero RAM dinamica. Questo tipo di RAM è costituito, a livello concettuale, da un transistor che mantiene l'informazione sotto forma di carica elettrica. È così possibile usare un solo componente per ogni cella di memoria, con costi molto ridotti e la possibilità di aumentare notevolmente la densità di memoria.

Però questa memoria è soggetta al fenomeno della "scarica". A causa del non perfetto isolamento la memoria si scarica, quindi dopo un breve lasso di tempo il suo contenuto diventa inaffidabile. Si rende necessario perciò ricaricarlo eseguendo un'operazione di lettura e riscrittura entro il tempo massimo in cui il contenuto può essere considerato ancora valido. Queste operazioni sono eseguite da un circuito interno alle memorie stesse. Oltre a comportare un certo dispendio di energia rendono più lenta la memoria in quanto, mentre si sta eseguendo il "refresh", non è possibile accedere alla memoria.

Tipologie di memoria: DRAM



L'accesso in scrittura ed in lettura alla DRAM è comandato direttamente dai segnali in ingresso (provenienti dalla CPU).

Sono generalmente usate per la memoria principale del sistema perché consentono di ottenere una grande capacità e sono (relativamente) economiche.

Tipologie di memoria: SDRAM



Sincronous Dinamic Random Access Memory, ovvero DRAM sincrone. Si differenziano dalle DRAM normali per il fatto che l'accesso è sincrono, ovvero governato dal clock.

E' un tipo di RAM utilizzata nelle DIMM per la memoria principale dei personal di tipo Pentium e successivi. Un segnale di clock temporizza e sincronizza le operazioni di scambio di dati con il processore, raggiungendo una velocità almeno tre volte maggiore delle SIMM DRAMi asincrone.

Tipologie di memoria: DDR SDRAM



Sincronous Dinamic Random Access Memory Double Data Rate, ovvero SDRAM con Data Rate doppio. Si differenziano dalle SDRAM per il fatto che consentono il trasferimento dei dati sia sul fronte positivo del clock sia su quello negativo, consentendo così di raddoppiare la banda teorica. Sono ottenute organizzando la memoria in due banchi separati, uno contiene le posizioni pari, a cui si accede sul fronte positivo del clock, e l'altro le posizioni dispari, alle quali si accede sul fronte negativo del clock.

Tipologie di memoria: DDR-II SDRAM

(Double Data Rate Two Synchronous Dynamic Random Access Memory) Il vantaggio della DDR-II rispetto alla DDR-I è la sua capacità di funzionare con una velocità di clock maggiore.

DDR-II raddoppia la velocità di clock rispetto a DDR, che a sua volta raddoppia la velocità della SDRAM. Come DDR, DDR-II trasferisce i dati ad ogni fronte di salita e di discesa, ottenendo così una velocità di trasferimento doppia. Altri miglioramenti sono ottenuti attraverso un aumento dei buffer, un miglioramento del prefetch, richieste elettriche ridotte e packaging migliorato. DDR3



Il software della ROM – 1

- La ROM è una memoria di sola lettura, permanentemente registrata nei circuiti dei chip del PC; contiene il software e i dati necessari ad inizializzare il computer ed a far funzionare i dispositivi periferici
 - Ⓜ **Vantaggio**: routine che forniscono le funzionalità di base incorporate nel sistema, con garanzia di continua disponibilità
- Il nucleo del software della ROM è costituito dalle **routine di avviamento** che comprendono il **caricatore di boot-strap** ed il **ROM BIOS**

Il software della ROM – 1

- Le routine di avviamento si occupano dell'inizializzazione del PC:
 - @Ne effettuano un rapido controllo di affidabilità, per accertare che tutte le componenti hardware siano perfettamente funzionanti
 - @Inizializzano i chip ed i dispositivi standard collegati al computer
 - @Inizializzano la tabella dei *vettori di interrupt*
 - @Controllano quali dispositivi opzionali sono collegati
 - @Caricano il sistema operativo dal disco (caricatore di boot-strap)

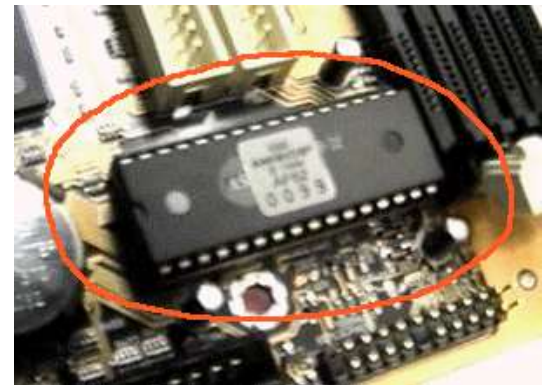
Memorie non volatili: la Rom

Nel corso degli anni sono stati sviluppati diversi tipi di Rom:

- **Rom** **Read Only Memory**
- **Prom** **Programmable Read Only Memory**
- **EProm** **Erasable Programmable Read Only Memory**
- **EEprom** **Electrically Eras. Progr. Read Only Memory**

L'unico tipo *veramente* di sola lettura è il primo, e il più vecchio. Le flash ROM, aggiornabili via software, sono quelle di uso corrente.

Le ROM sono utilizzate anche in dispositivi diversi dai personal computer (ad esempio nei cellulari)



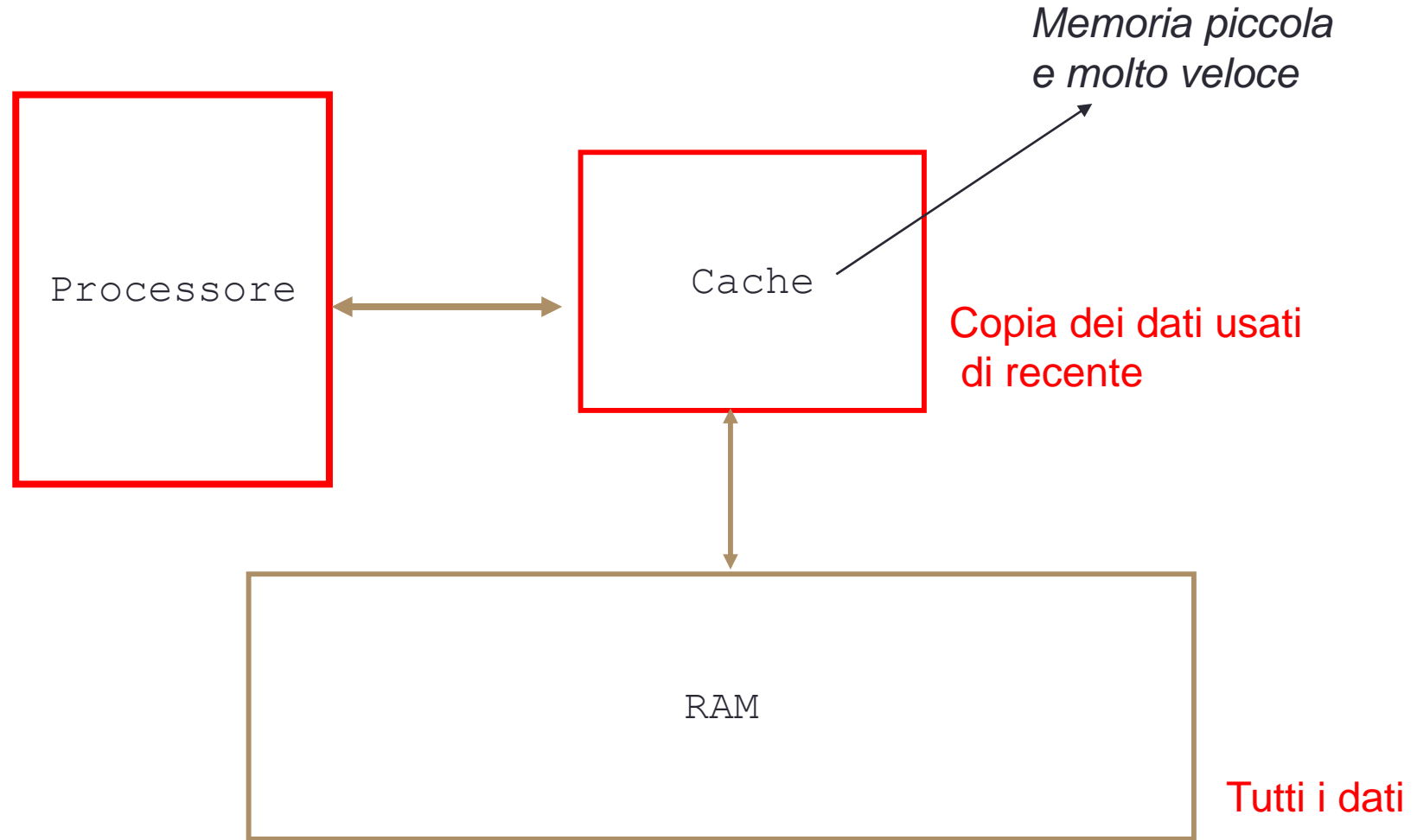
Il software della ROM – 2

- Il **caricatore di boot-strap** ha la funzione di leggere un programma di lancio dal disco e di trasferire ad esso il controllo
 - Ⓞ Su un **disco di sistema** (*bootable*), il programma controlla l'esistenza di un sistema operativo; in caso affermativo, i file necessari all'inizializzazione vengono caricati in memoria assieme al relativo file interprete dei comandi
- Il **ROM BIOS** è la parte della ROM attiva quando il computer è al lavoro: il suo ruolo è quello di fornire un insieme di servizi di base richiesti per il funzionamento delle periferiche
- **BIOS** (*Binary Input–Output System*): comandi che controllano i dispositivi periferici (es., traducono un comando, come "leggi dal disco", nella successione di passi elementari necessari a eseguire effettivamente l'operazione, incluso il rilevamento di errori)

La memoria CMOS

- Tutti i PC sono dotati di una memoria permanente speciale, alimentata da una batteria
- A questa memoria si fa riferimento con il termine **CMOS** (*Complementary Metal–Oxide–Silicon*) ed in essa sono immagazzinati i dati relativi al clock in tempo reale ed alla configurazione del sistema e delle sue periferiche
- Parte della memoria CMOS viene impiegata per conservare le specifiche relative al tipo di disco fisso e all'ammontare di memoria disponibile; al momento del *boot*, il BIOS legge le informazioni di configurazione in CMOS e le utilizza per inizializzare il sistema

Memoria cache 1



Memoria cache 2

- La cache è una memoria molto veloce posta tra CPU e memoria centrale. Serve per aumentare le prestazioni del sistema, ed è trasparente alla CPU.
- Si hanno generalmente 3 livelli di cache:
 - L1 → integrata nel chip della CPU
 - L2 → esterna al chip, ma collegata direttamente alla CPU mediante il Bus Back Side che lavora alla stessa frequenza
 - L3 → esterna, sulla M/B
- La cache è organizzata in blocchi (linee). Ad ogni operazione di lettura le informazioni vengono cercate prima in cache, e solo successivamente in RAM. Ad ogni operazione di scrittura la cache può usarsi come:
 - Write-through → scrittura sia in cache che in RAM
 - Write-back → scrittura solo in cache, solo su sostituzione viene scritta in RAM

Memoria cache 3

Algoritmo seguito:

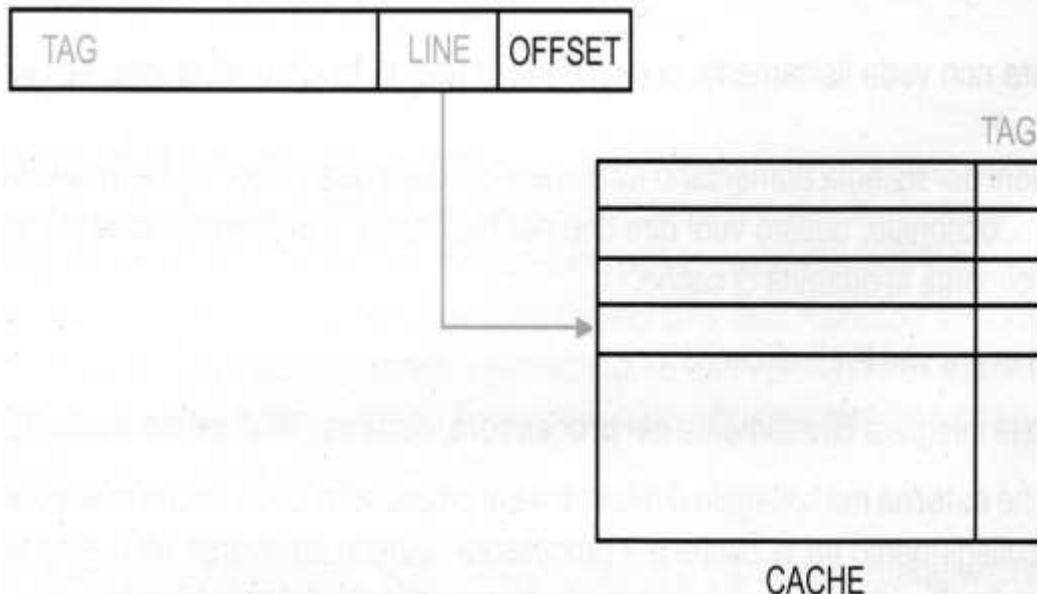
- Il dato viene cercato prima nella cache
- Se è presente abbiamo finito (*cache hit*)
- Se non è presente (*cache miss*) si legge la RAM e si copia in cache
- Per le scritture generalmente si scrive la RAM e si aggiorna la copia se c'è

E se devo inserire un nuovo dato e non c'è più posto nella cache?

- Questo accade perché le memorie cache sono molto più piccole delle RAM (< qualche MB)
- La soluzione è sovrascrivere una posizione già in uso.
- Ci sono diverse strategie : di solito si sovrascrive la posizione con dati utilizzati meno di recente

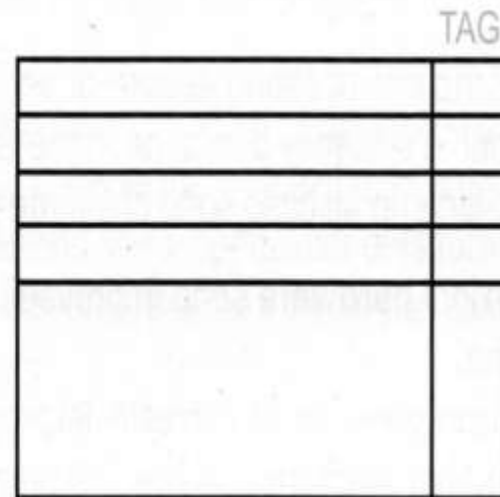
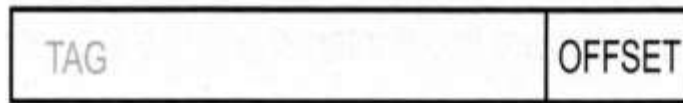
Metodi di gestione della cache 1

- **Metodo Diretto** → ogni blocco di RAM viene allocato in un preciso blocco di Cache. L'indirizzo di memoria è diviso in 3 parti, **LINE** indica la linea di cache dove è allocato il blocco. Tutti i blocchi con **TAG** differenti vengono allocati nella stessa linea, ed identifica quale blocco sta in cache. L'**OFFSET** indica la posizione dell'informazione nella linea



Metodi di gestione della cache 2

- **Metodo completamente Associativo** → ogni blocco di RAM viene allocato in un qualsiasi linea della Cache. L'indirizzo di memoria è diviso in 2 parti, il **TAG** serve per trovare il blocco nella cache. Ricerca sequenziale, svantaggi e vantaggi.



CACHE

Metodi di gestione della cache 3

Metodo Associativo a n vie → Cache divisa in gruppi di n linee. L'indirizzo di memoria è diviso in 3 parti, SET indica il gruppo in cache dove è allocato il blocco. Nel blocco vengono confrontati i TAG fino a trovarlo o ad esaurimento del gruppo. Se non c'è allora bisogna andare ad allocarlo, e se il gruppo non ha linee libere, bisogna rimpiazzare la linea di un determinato blocco secondo una determinata strategia.

