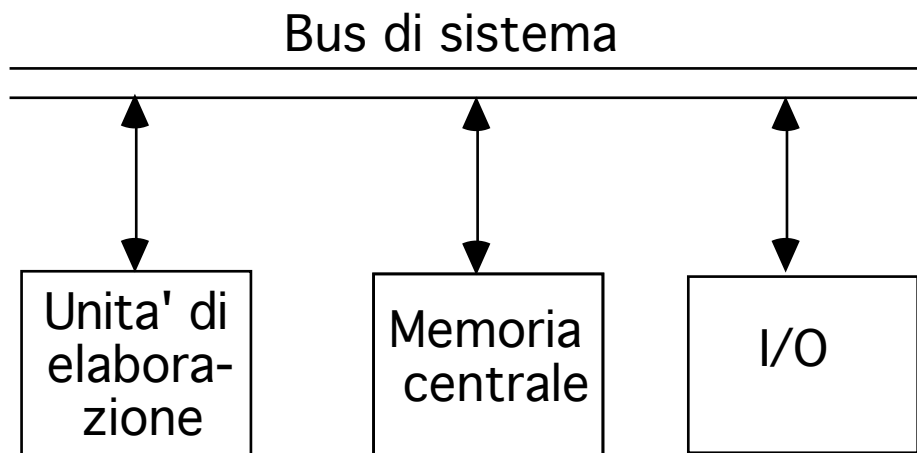


Architettura di un Elaboratore

Organizzata secondo il modello della **macchina di von Neumann** definita nei tardi anni '40 all'Institute for Advanced Study di Princeton.

È costituita da quattro elementi funzionali fondamentali:

- Unità di Elaborazione o unità centrale (CPU - *central processing unit*, o Microprocessore);
- Memoria Centrale (RAM, ma non solo);
- Unità di I/O (ingresso/uscita, seriali, parallele, eccetera);
- Bus di sistema (tra i vari tipi ISA, EISA, PCI).



Architettura di un elaboratore: elementi funzionali

La **CPU** contiene i dispositivi elettronici in grado di acquisire, interpretare ed eseguire il programma trasformando i dati.

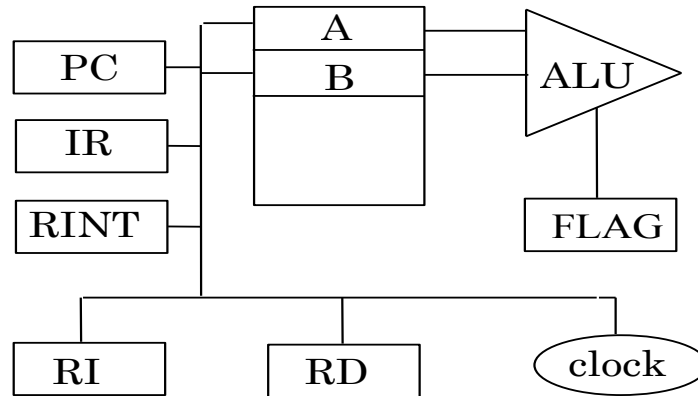
Ogni istruzione è *logicamente* eseguita **in sequenza**. Però i calcolatori moderni riescono a volte a eseguire più istruzioni in una volta.

La **memoria centrale** contiene sia le *istruzioni* sia i *dati* (informazioni per eseguire un programma). Il sistema operativo le carica dalla memoria di massa quando comandiamo l'esecuzione di un programma. Vengono trasferite alla CPU quando la CPU ne ha bisogno. La memoria centrale è suddivisa in celle di memoria, ognuna delle quali univocamente identificata da un indirizzo (i.e., un numero).

Le **unità di I/O** consentono uno scambio di informazione fra l'elaboratore e l'esterno (*ingresso/uscita, memorie di massa, periferiche*). Anch'esse sono identificate da appositi indirizzi.

Il **bus di sistema** collega questi elementi funzionali. Fornisce il supporto fisico per la trasmissione dei dati tra i vari elementi.

Unità di Elaborazione (*Central Processing Unit, CPU*)



Blocchi Componenti (in un unico Chip):

- l'**unità aritmetico-logica** (*Arithmetic Logic Unit, ALU*) che esegue le operazioni elementari necessarie per l'esecuzione delle istruzioni sui dati binari
- l'**unità di controllo** (*Control Unit, CU*), controlla e coordina l'attività della CPU. In particolare, è responsabile del trasferimento e della decodifica delle istruzioni dalla memoria centrale ai registri della CPU;
- il **clock** (orologio) “cadenza” le operazioni elementari, permettendo il sincronismo delle operazioni;
- vari **registri** (ad esempio, **A**, **B**, **PC**, **FLAG**, etc.): un registro è una locazione di memoria utilizzata per memorizzare all'interno della CPU: *dati* prelevati dalla memoria e se cui la CPU deve *operare*, *istruzioni* prelevate dalla memoria e che la CPU deve *eseguire*, o *indirizzi* di celle di memoria all'interno delle quali ci sono dei dati o istruzioni da *prelevare*. In altre parole, la CPU non lavora direttamente sulla memoria centrale, ma sposta temporaneamente nei suoi registri ciò che gli interessa in un certo momento.

Unità di controllo

Le istruzioni sono organizzate in memoria in sequenza. Le celle di memoria che contengono istruzioni sono indirizzate dal registro PC.

L'unità di controllo esegue le istruzioni secondo quello che viene detto ciclo *accesso-decodifica-esecuzione*. Si tratta di una sequenza di passi in cui:

1. **FASE DI FETCH.** si accede alla prossima istruzione (il cui indirizzo in memoria è indicato nel registro PC, *program counter*) portandola dalla memoria nel registro IR (*instruction register*);
2. **FASE DI DECODIFICA.** si decodifica il tipo dell'istruzione da eseguire;
3. si individuano i dati usati dall'istruzione (il cui indirizzo in memoria è indicato nell'istruzione stessa e viene riportato in uno dei registri indirizzo), che vengono prelevati dalla memoria portati nei registri opportuni (A, B, o altri registri di uso generale per i dati);
4. **FASE DI EXECUTE.** si esegue l'istruzione.

Istruzioni particolari possono alterare il flusso sequenziale (salto, chiamata di sotto-programmi).

Unità Logico-Aritmetica (ALU Arithmetic-Logic Unit)

Esegue operazioni aritmetiche e confronti sui dati della memoria centrale o dei registri.

In dipendenza dall'elaboratore può essere molto complessa e sofisticata. Qui consideriamo un caso semplice. ALU in grado di eseguire somma, sottrazione, prodotto, divisione con due operandi contenuti nei registri A e B.

1. I due operandi vengono caricati nei registri A e B;
2. La ALU viene attivata da un codice operativo inviato dalla CPU che specifica il tipo di operazione;
3. Nel registro A viene caricato il risultato dell'operazione eseguita dalla ALU;
4. Il registro FLAG riporta sui suoi bit indicazioni sul risultato dell'operazione (riporto, overflow, etc.).

Fra i bit di FLAG ricordiamo:

- bit di carry (indica la presenza di un riporto);
- bit di segno (indica il segno del risultato e consente di identificare la condizione di overflow che si verifica quando i due operandi hanno segno uguale fra loro ma diverso dal risultato).

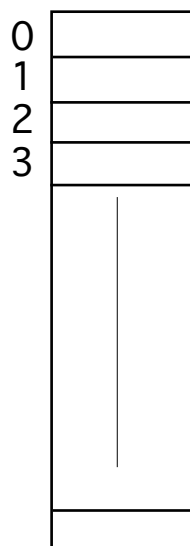
Memoria centrale

La memoria centrale contiene tutta l'informazione che deve essere elaborata "a breve termine". Ogni elemento di informazione, per poter essere elaborato, **deve** prima essere acquisito dalla memoria centrale e portato nei registri

È *volatile* (se ne perde il contenuto quando la macchina viene spenta).

In generale, di dimensioni ridotte, costo elevato ed un **tempo di accesso molto basso** (il tempo necessario per scrivervi o leggerne un elemento può scendere, per certi calcolatori, fino a pochi *nano*-secondi (10^{-9} secondi)).

È una sequenza di celle, o locazioni, di memoria, ciascuna identificata da un *indirizzo* e contenente una sequenza di cifre binarie, o **bit** (**b**inary **d**igit) chiamata **parola** (word).



Memoria centrale

Per ogni calcolatore la capienza della cella di memoria è fissa. Calcolatori diversi possono avere capienze differenti (8, 16, 32, 64 bit).

Unità di misura:

byte, equivale ad 8 bit, **word** equivale a 16 bit

Unità successive, potenze in base 2 del byte (con esponente potenza intera di 10):

Kilobyte	2^{10} byte	1024 byte	KB
Megabyte	2^{20} byte	1048576 byte	MB
Gigabyte	2^{30} byte	$\sim 10^9$ byte	GB
Terabyte	2^{40} byte	$\sim 10^{12}$ byte	TB

Memoria centrale: indirizzamento

L'elaboratore seleziona una particolare cella di memoria ponendone l'indirizzo in uno dei registri-indirizzo.

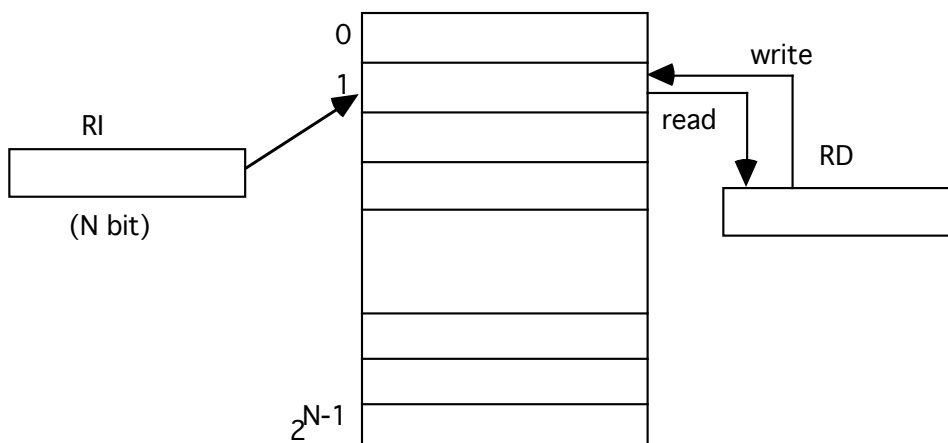
Se i registri-indirizzi sono lunghi N bit, si possono indirizzare 2^N celle di memoria da 0 a $2^N - 1$ (per esempio, se $N=10$ si possono indirizzare 1024 celle).

Operazione di **lettura** (*load*)

Copia in un registro della CPU (IR, A, B, etc.) il contenuto della cella di memoria indirizzata dal registro PC o da un altro registro indirizzo .

Operazione di **scrittura** (*store*)

Copia il contenuto di un registro della CPU (A, B, etc.) in una cella di memoria indirizzata da uno dei registri indirizzo. Le operazioni avvengono sotto il controllo della CPU.



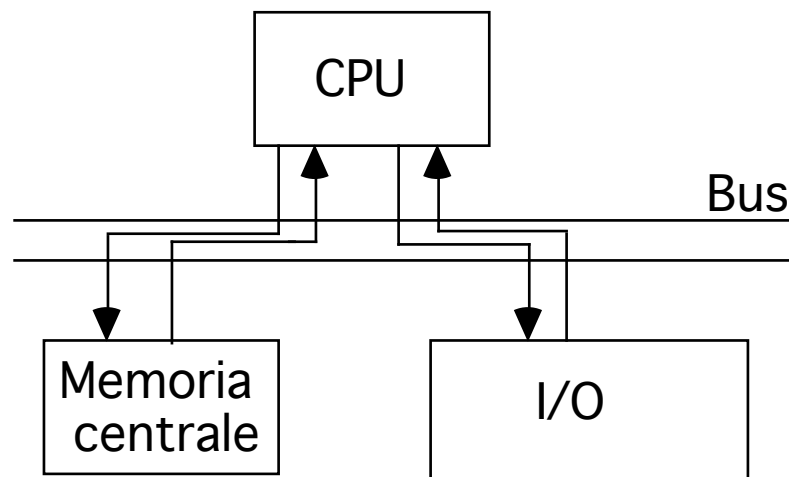
Dispositivi fisici utilizzati per la memoria centrale

- **RAM:** Random Access Memory (ad accesso casuale): su di essa si possono svolgere operazioni sia di lettura sia di scrittura.
 - **ROM:** Read Only Memory (a sola lettura): non volatili e non scrivibili; in esse vengono contenuti i dati e programmi per inizializzare il sistema.
 - **PROM:** Programmable ROM. Si possono scrivere soltanto una volta, mediante particolari apparecchi (detti programmatori di ROM).
 - **EPROM:** Erasable-Programmable ROM (si cancellano sottoponendole a raggi ultravioletti).
- ☞ **Firmware** : è costituito dal software memorizzato nelle ROM (codice microprogrammato). Tipicamente i pezzi di base del sistema operativo (*BIOS*).

Bus di sistema

Interconnette la CPU, la memorie e le interfacce verso dispositivi periferici (I/O, memoria di massa, etc.)

Collega due unità funzionali alla volta: una trasmette e l'altra riceve. Il trasferimento avviene sotto il controllo della CPU.



Su questo supporto (costituito da più linee, cioè da insiemi logici diversi di fili) viaggiano dati, indirizzi e comandi. Si distinguono spesso tali linee in:

- bus dati (data bus)
- bus indirizzi (address bus)
- bus comandi (command bus)

Bus di sistema

Bus dati: bidirezionale. Serve per trasmettere dati dalla memoria al registro dati o viceversa.

Bus indirizzi: unidirezionale. Serve per trasmettere il contenuto del registro indirizzi alla memoria. Tramite l'indirizzo sul bus, la memoria seleziona una specifica cella per successive operazioni di lettura o scrittura.

Bus comandi: unidirezionale. Ad esempio, comando di lettura o scrittura verso la memoria; comando di stampa verso una periferica (interfaccia).

Se la dimensione (numero di bit) del bus dati è uguale alla dimensione della parola si può trasferire in parallelo un intero dato. Altrimenti occorrono più trasferimenti.

Interfacce di Ingresso/Uscita

Consentono il collegamento dell'elaboratore con le varie periferiche (dischi, terminali, stampanti, ...).

Le interfacce sono molto diverse tra loro, e dipendono dal tipo di unità periferica.

Un'interfaccia standard ed elementare contiene:

- Un registro-*dati*;
- Un registro-*comando*;
- Un'informazione sullo stato della periferica.

Tipi di Periferiche:

Terminali, hanno una tastiera e un video. Possono essere *alfanumerici* o *grafici*.

Mouse, è un dispositivo di ingresso che può essere utilizzato in sostituzione della tastiera.

Stampanti, sono caratterizzate da *velocità* di stampa (byte o carattere per secondo), *risoluzione* (numero di punti per pollice al quadrato), capacità grafica, rumorosità.

Tecnologia ad aghi, a margherita, a getto di inchiostro, laser.

Anche la memoria di massa è vista come dispositivo periferico.

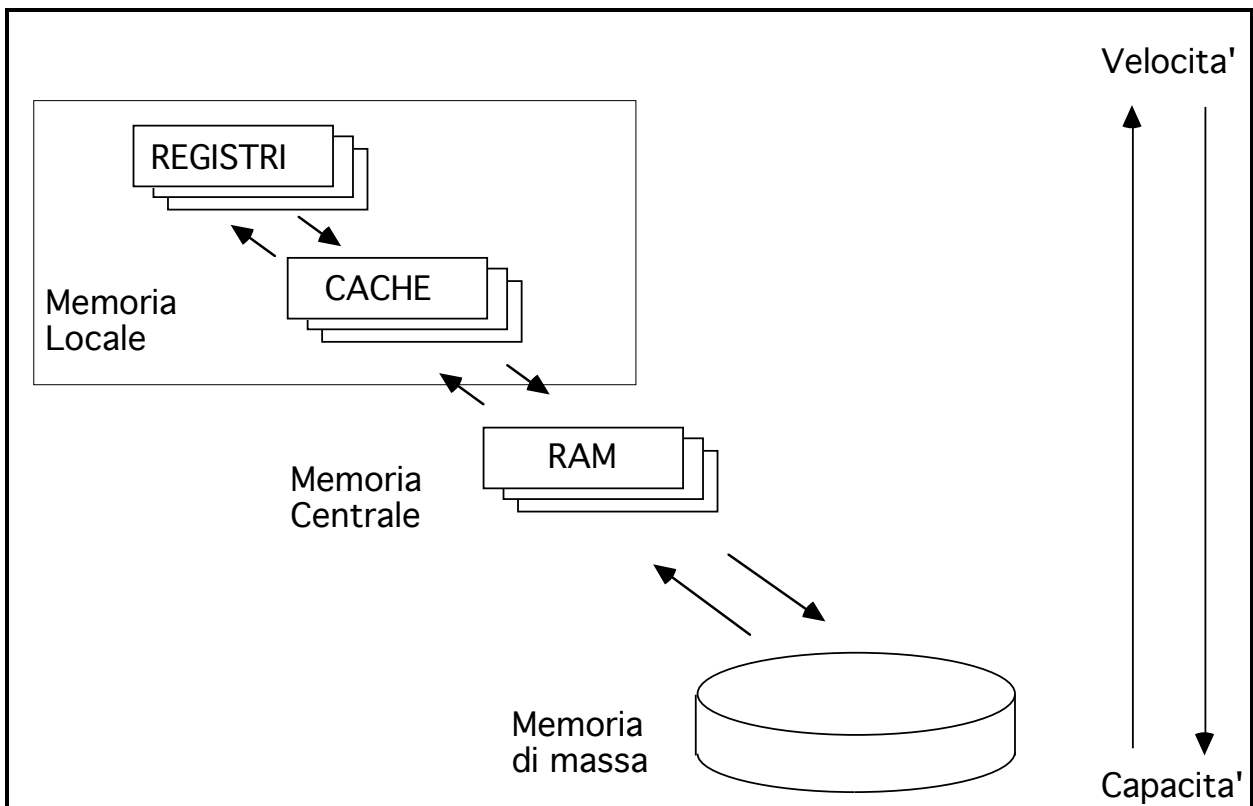
Estensioni della macchina di Von-Neumann

La macchina di Von Neumann ha come problema la stretta sequenzialità imposta alle operazioni.

Esistono proposte che introducono varie forme di **parallelismo**.

- Processori dedicati (*co-processor*) al calcolo numerico o alla gestione della grafica. Processori dedicati all'ingresso/uscita.
- Esecuzione *in parallelo* delle varie fasi di un'istruzione. Mentre si esegue un'istruzione è già in corso l'acquisizione ed interpretazione delle successive (Pipeline).
- Esistono anche *architetture non von Neumann*:: sistemi multi-processore, macchine dataflow, LISP machine, macchine sistoliche, reti neurali.

Gerarchia delle memorie



Memoria cache

Memoria di limitata capacità (es. 1Mb), formata da RAM molto **veloce** posta tra il microprocessore e la memoria centrale. In molti casi è presente internamente al microprocessore stesso.

La prima volta che il microprocessore carica dei dati dalla memoria centrale, tali dati vengono caricati anche sulla cache. Successivamente, i dati vengono letti dalla cache invece che dalla memoria centrale più lenta.