

ELEMENTI DI ARCHITETTURA DI UN SISTEMA DI ELABORAZIONE

per le classi prime

vers2.0

Contenuti

Architettura di un sistema di elaborazione.....	2
Macchina di Von Neumann	2
Elementi grafici per rappresentare una architettura	2
Funzioni svolte da un sistema di elaborazione.....	4
Componenti di un sistema di elaborazione	4
Unità Centrale - CPU (CU, ALU).....	5
Clock.....	6
Bus di sistema	7
Memoria.....	8
Unità di Canale (input/output- I/O).....	10
Interruzioni - Interrupt	12
Dispositivi di archiviazione di massa.....	13
Disco rigido - HARD DISK.....	13
Nastri magnetici	15
CD/DVD	15
Prestazioni di un sistema di elaborazione	16
Istruzioni e programma	17
Programma eseguibile.....	17
Esecuzione di una istruzione.....	17
Struttura delle istruzioni.....	18
Insieme delle istruzioni eseguibili.....	18
Programma e linguaggi di programmazione.....	19
Esecuzione di un programma compilato.....	20

ARCHITETTURA DI UN SISTEMA DI ELABORAZIONE

Il termine architettura, riferito ai sistemi di elaborazione, indica l'organizzazione dei vari elementi presenti in un sistema di elaborazione ed il modo in cui essi cooperano per eseguire le operazioni richieste.

Macchina di Von Neumann

I moderni elaboratori elettronici, dai piccoli personal computer ai grandi computer, derivano da uno stesso “modello” di elaboratore chiamato Macchina di Von Neumann dal nome dello scienziato che nel 1946 pubblicò, insieme ad altri, un progetto per una nuova macchina di calcolo automatico.

L'elemento caratteristico e fondamentale introdotto da Von Neumann e che costituì una evoluzione rispetto agli elaboratori dell'epoca, è il “programma memorizzato”, cioè il fatto che il programma da eseguire viene immagazzinato nella memoria al pari dei dati su cui esso opera.

Pertanto, l'elaboratore è un sistema che può svolgere compiti molto diversi tra loro variando opportunamente i programmi.

Un programma è composto da un gran numero di istruzioni codificate in una forma eseguibile dal elaboratore.

Elementi grafici per rappresentare una architettura

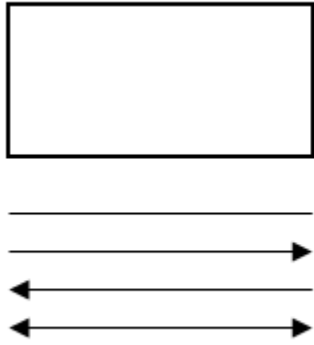
L'architettura di un sistema di elaborazione può essere visualizzata mediante una rappresentazione grafica evidenziando meglio gli elementi caratteristici dell'architettura.

In informatica, vengono usate altre rappresentazioni grafiche sia per descrivere i programmi (flow chart), sia per descrivere archivi di dati (entity relationship) etc...

Per rappresentazione graficamente un concetto, una idea è necessario indicare quali sono:

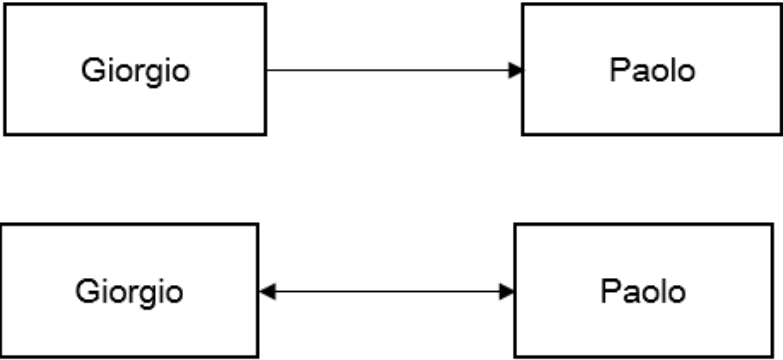
- i simboli utilizzati e cosa rappresentano(dizionario)
- le loro regole di composizione (sintassi).

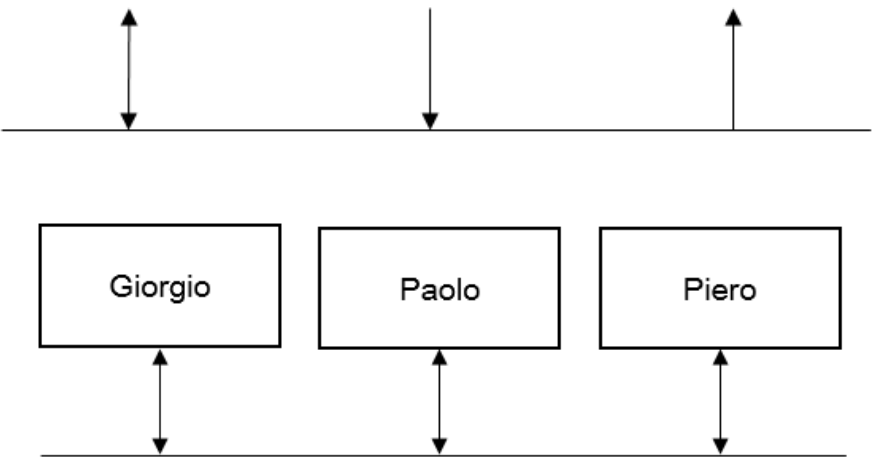
I simboli utilizzati sono:

<p>Rettangolo: rappresenta l'entità, l'elemento che intendiamo rappresentare (concetto, cosa, dispositivo, persona, etc..)</p> <p>Segmento con o senza freccia che collega due rettangoli, che evidenzia la relazione che si attua tra le entità/elemento (scambio d'informazione, scambio di materiale, dipendenza temporale), la relazione va nel senso della freccia</p>	
---	--

Le regole di composizione sono:

- o ad ogni rettangolo possono insistere più segmenti con freccia
- o ogni segmento con freccia deve essere collegato alle due estremità ad un rettangolo o ad un segmento senza frecce

<p>Esempio1 Giorgio parla e Paolo ascolta</p> <p>Esempio2 Giorgio e Paolo parlano tra loro</p>	
---	---

<p>Per dire che più persona parlano tra di loro utilizziamo la seguente rappresentazione</p>	
--	--

Per rappresentare una struttura minima di elaborazione associamo ai rettangoli dispositivi caratterizzati da precise funzionalità ed alle frecce la modalità con cui questi dispositivi scambiano informazione.

Funzioni svolte da un sistema di elaborazione

Le funzioni richieste da un sistema di elaborazione sono:

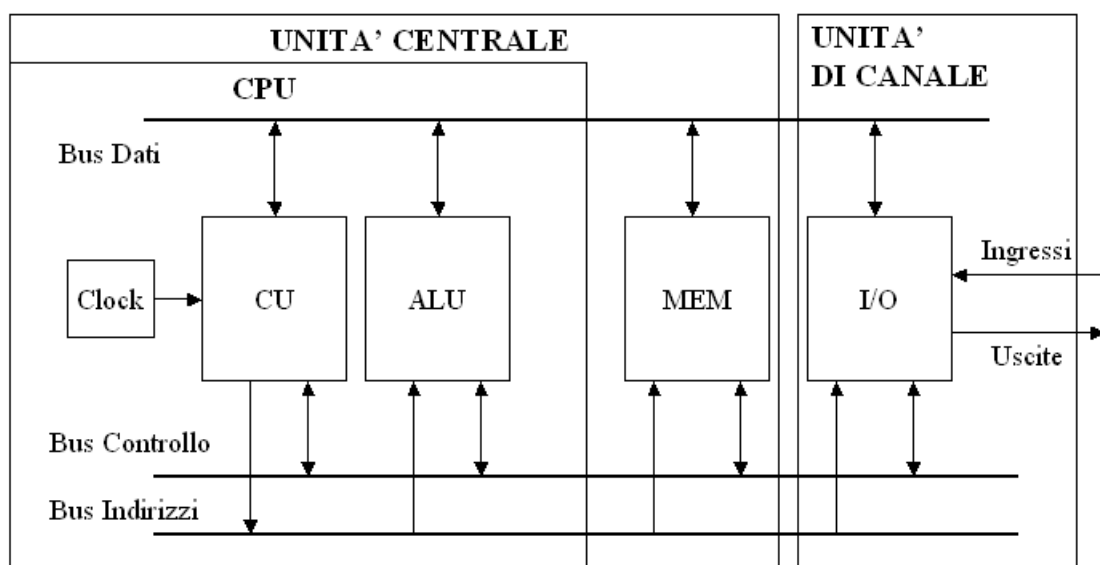
- funzione di memoria: codifica di informazione (programmi e dati) su supporti in grado di conservarla per il tempo necessario alla elaborazione
- funzione di esecuzione: esecuzione di una specifica gamma di operazioni, potendo determinare l'operazione da eseguire tra quelle a disposizione e specificando gli "operandi" (dati) su cui essa opera.
- funzione di controllo: riconoscimento delle istruzioni da eseguire e conseguente avvio della loro esecuzione mediante opportuni segnali e controllo della successione di istruzioni che compongono il programma
- funzione di ingresso o input: introduzione dei dati e dei programmi nel sistema
- funzione di uscita o output: comunicazione dei risultati all'esterno.

Componenti di un sistema di elaborazione

Le componenti di un sistema di elaborazione possono essere suddivise in due grandi blocchi funzionali:

- unità centrale, adibita all'elaborazione dei dati ed al controllo generale del sistema
- unità di canale (unità di I/O), adibite al collegamento con le periferiche, cioè i dispositivi che permettono il colloquio del sistema con l'ambiente esterno o la memorizzazione permanente di dati e programmi (tastiera, video, stampante, memorie di massa, ecc.).

Le componenti di un sistema di elaborazione sono evidenziate nella figura seguente:



Unità Centrale - CPU (CU, ALU)

La CPU (*Central Processing Unit*) costituisce il cuore del sistema di elaborazione è il luogo dove avviene l'esecuzione delle istruzioni di un programma; è costituita dalla unità di controllo CU (*Control Unit*), dalla ALU (*Arithmetic Logic Unit*) e da un certo numero di celle di memoria "locale", chiamate registri, che contengono i dati necessari per eseguire l'istruzione richiesta .

Unità di controllo (CU)

L'unità di controllo (CU) è un dispositivo logico progettato per coordinare tutte le attività del sistema di elaborazione, imponendo la corretta sequenzializzazione delle operazioni elementari che devono essere svolte per eseguire una istruzione.

Nell'unità di controllo sono presenti due registri fondamentali per l'esecuzione di una istruzione:

- il program counter (PC)
- l'instruction register (IR).

Il program counter (PC) contiene l'indirizzo della prossima istruzione e viene incrementato automaticamente di uno o può venire modificato durante l'esecuzione di opportune istruzioni.

L'instruction register (IR) contiene l'istruzione in esecuzione.

L'unità di controllo (CU) per eseguire una istruzione compie tre operazioni di base:

- prelievo (fetch)
- decodifica (decode)
- esecuzione (execute)

L'insieme di queste tre operazioni è chiamato ciclo istruzione.

ESEMPIO:

Nei primi computer il registro di program counter (PC) era composto da 16 bit ovvero 2 byte

Il registro di istruzione (IR) può essere composto da 8 bit ovvero da 1 byte

Unità aritmetica e logica (ALU)

La ALU è un dispositivo logico progettato per eseguire operazioni logiche (AND, OR, NOT) e aritmetiche (addizione (ADD), sottrazione (SUB) moltiplicazione (MUL), divisione(DIV)) sugli operandi memorizzati nei registri interni della ALU.

Le operazioni vengono svolte con l'ausilio di appositi registri interni e precisamente:

- due o più registri per immagazzinare gli operandi delle operazioni da eseguire ed il relativo risultato (chiamato anche registro accumulatore)
- un registro di stato (RS), nel quale vengono memorizzate alcune informazioni utili sul risultato dell'operazione eseguita.

Le informazioni memorizzate nel registro di stato sono un bit di:

- carry (CF-carry flag), indica la presenza di un riporto (carry) nell'istruzione di somma tra due numeri naturali
- zero (ZF-zero flag), indica che l'operazione eseguita vale zero (cioè tutti i bit nel registro accumulatore valgono 0)
- overflow (OF-overflow flag), indica il verificarsi di overflow, ovvero il risultato dell'operazione non è rappresentabile correttamente nel registro accumulatore
- segno (SF-sign flag), indica che l'operazione eseguita ha prodotto un risultato con il bit più significativo pari a 1, ovvero nel caso di numeri interi con segno si tratta di un numero negativo

ESEMPIO

I registri per memorizzare gli operandi e risultati sono formati da almeno 2 byte e le operazioni della ALU sono eseguite su almeno due Byte.

Clock

Il clock è un dispositivo che sincronizza i vari circuiti presenti nel sistema di elaborazione.

Il clock genera impulsi di durata costante ed emessi ad intervalli di tempo regolari.

L'intervallo tra due impulsi consecutivi del clock è detto periodo di clock ed è anche chiamato ciclo macchina ed è il passo elementare di esecuzione.

L'inverso del periodo di clock indica quanti impulsi vengono emessi nell'unità di tempo ovvero quanti cicli macchina possono essere eseguiti in un secondo, si chiama frequenza del clock e si misura in Hertz (Hz).

Maggiore è la frequenza del clock, maggiore è la velocità dei dispositivi presenti e quindi il numero di operazioni svolte in un certo intervallo di tempo.

Si usano i seguenti multipli del Hz:

- MHz = 10^6 Hz mega hertz
- GHz = 10^9 Hz giga hertz

Si usano i seguenti sottomultipli del sec.

- ms = 10^{-3} s millisecondo
- μ s = 10^{-6} s microsecondo
- ns = 10^{-9} s nanosecondo

ESEMPIO

Le prime CPU lavoravano con un clock di 1 Mhz ovvero con $1 \cdot 10^6$ cicli macchina al secondo, ogni ciclo macchina dura $1 \mu\text{s} = 1 \cdot 10^{-6}$ s

Fino a circa 10 anni fa le CPU lavoravano con un clock di 200 MHz svolge $200 \cdot 10^6$ cicli macchina al secondo, ogni ciclo macchina dura $5 \text{ ns} = 1 / (200) \cdot 10^{-6}$ sec

Attualmente le CPU lavorano con un clock maggiore di 2GHz.

Bus di sistema

I vari dispositivi del sistema di elaborazione si scambiano informazioni mediante opportuni canali chiamati bus.

In un sistema di elaborazione sono presenti tre tipi bus (interni):

- bus indirizzi (address bus - AB)
- bus dati (data bus -DB)
- bus di controllo (control bus CB)

BUS INDIRIZZI

L'informazione che viene scambiata nel bus indirizzi consente l'identificazione di un particolare dispositivo e/o cella di memoria.

Il bus indirizzi è caratterizzato dal numero di bit che è composto.

Un bus di indirizzi composto da n bit permette di individuare 2^n dispositivi/celle di memoria diverse.

ESEMPIO

Con un bus di indirizzi composto da 20 bit possiamo indirizzare:

- 2^{20} celle di memoria circa 1 MByte,
- la prima cella viene identificata con il numero 0 e l'ultima con il numero $2^{20} - 1$. Questo intervallo si chiama spazio di indirizzamento o intervallo d'indirizzamento.

Nei primi computer il bus di indirizzi era composto da 16 bit, poteva indirizzare $2^{16} = 65536$ celle di memoria

BUS DATI

L'informazione che viene scambiata nel bus sono dati o istruzioni

Il bus dati è caratterizzato dal numero di bit che è composto; solitamente il numero di bit corrisponde a quello dei bit elaborati dalla ALU/CU.

ESEMPIO

Il bus dati può essere rappresentato da 1 byte, 2 byte, 4 byte.

Nei primi computer il bus di dati era composto da 8 bit

BUS CONTROLLO

L'informazione che viene scambiata nel bus di controllo consente l'identificazione e la sincronizzazione delle varie operazioni in corso tra le diverse parti dell'architettura.

ESEMPIO

Nel caso di lettura e scrittura di una locazione di memoria è presente un bit chiamato READ/WRITE il cui valore indica se l'operazione da eseguire è di lettura o scrittura della locazione della memoria

Memoria

Struttura logica delle memorie

La memoria è un dispositivo in grado di conservare informazioni.

Dal punto di vista logico la memoria si caratterizza per i seguenti fattori:

- è organizzata in celle o locazioni di uguali dimensioni ovvero un “gruppo di bit”, individuabile ed accessibile mediante un indirizzo
- la cella o locazione di memoria rappresenta l'unità di informazione scambiata tra i vari elementi funzionali di cui si compone l'architettura
- il contenuto di una locazione, ovvero l'informazione in essa registrata, si chiama parola di memoria, le cui dimensioni tipiche sono 8, 16, 32, 64 e 128 bit.
- la posizione che la locazione occupa nella memoria rispetto alla prima locazione, che ha indirizzo zero, si chiama indirizzo di una locazione
- quando vengono compiute operazioni di lettura o di scrittura su una determinata locazione di memoria si verifica un “accesso” alla memoria

Modalità e tempo di accesso alla memoria

Una caratteristica importante di una memoria è la modalità di accesso, parleremo di :

- memorie ad accesso uniforme: una cella è accessibile indipendentemente dalla sua posizione ed in un tempo uguale per tutte le celle presenti nella memoria (esempio memoria centrale);
- memorie ad accesso diretto: una cella è accessibile indipendentemente dalla sua posizione, ma il tempo impiegato dipende da quest'ultima (esempio: memorie ausiliarie su disco);
- memorie ad accesso sequenziale: una cella è accessibile solo passando attraverso tutte le celle che la precedono (esempio: memorie ausiliarie su nastro).

Memoria centrale/primaria e memoria secondaria

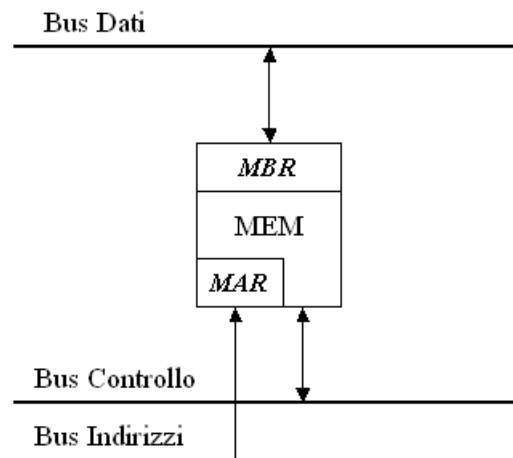
In un sistema di elaborazione sono sempre presenti due tipi di memoria:

- la memoria centrale, interna al calcolatore, direttamente accessibile dalla CPU realizzata da componenti a semiconduttore;
- la memoria secondaria o di massa, esterna al calcolatore realizzata da componenti magnetici o ottici (dischi, nastri) ed accessibile mediante l'unità di canale.

Memoria centrale del sistema di elaborazione

La memoria centrale è composta per poter essere acceduta da:

- un registro di indirizzamento in memoria MAR (*Memory Address Register*) contenente l'indirizzo cui accedere
- un decodificatore degli indirizzi che riceve il suo input dal MAR
- un registro di lettura e scrittura MBR (*Memory Buffer Register*) per immagazzinare la parola interessata al trasferimento



Ciascuna locazione di memoria è selezionata specificando nel registro degli indirizzi (MAR Memory Address Register) il suo indirizzo, ossia la sua posizione rispetto alla prima cella di memoria, a cui viene attribuita per convenzione la posizione (ovvero l'indirizzo) zero.

Il numero di locazione di memoria, chiamata anche capacità della memoria centrale, è un parametro che caratterizza la memoria ed è uno degli elementi che determina le "potenza di calcolo dell'elaboratore"

Si definisce spazio di indirizzamento l'insieme (o il numero) delle celle indirizzabili direttamente.

L'unità di misura della memoria è il byte (B) con i suoi multipli:

- KB=1.024 B Kilobyte
- MB=1.048.576 Be Megabyte
- GB= 1.073.741.824 B Gigabyte
- TB= 1.099.511.627.776 B Terabyte

Il numero di celle indirizzabili è una potenza di due. Con:

- 16 bit si indirizzano $2^{16} = 65.536$ celle
- 32 bit si indirizzano $2^{32} = 4.294.967.296$ celle
- 64 bit si indirizzano $2^{64} = 18.446.744.073.709.600.000$ celle

ESEMPIO

il registro degli indirizzi (MAR) è composto da k bit pertanto può selezionare una cella di memoria tra indirizzare 2^k celle di memoria i cui indirizzi variano da 0 a 2^k-1 .

il registro MAR è composto da 10 bit, può selezionare un indirizzo tra 2^{10} , ovvero 1024, locazioni di memoria.

Tipi di memoria centrale del sistema di elaborazione

La memoria centrale è composta da memoria RAM (*Random Access Memory*) e da memoria ROM (Read Only Memory), cache

RAM

La memoria RAM è una memoria ad accesso uniforme di lettura e scrittura; è *un* dispositivo logico, i cui circuiti sono progettati per conservare informazioni in codifica binaria fino allo spegnimento della macchina; si dice che è una memoria volatile

Il tempo di accesso della memoria RAM dipende dal tipo di tecnologia utilizzata, mediamente è una memoria veloce.

ESEMPIO

Per leggere/scrivere una cella sono richiesti, in media 5-30 nanosecondi (millesimi di milionesimi di secondo 10^{-9} secondi).

ROM

La memoria ROM è una memoria ad accesso uniforme non volatile, a sola lettura; viene scritta in fabbrica e non permette in nessun modo di essere riscritta.

Memoria CACHE

La cache è una memoria più piccola della memoria centrale e più veloce.

Nei moderni personal si trovano almeno due memorie cache: quella di primo livello (L1) grande alcuni KB, inglobata all'interno della CPU e veloce quanto la CPU, e quella di secondo livello (L2), grande qualche MB, situata esternamente alla CPU, meno veloce, ma comunque più veloce della memoria centrale.

Il loro ruolo è quello di interporre tra la CPU e la memoria centrale allo scopo di immagazzinare dati ed istruzioni che si ritiene possano servire alla CPU per le operazioni immediatamente successive, onde evitare che debbano essere letti nella più lenta memoria centrale.

Per questo la CPU cerca dati e istruzioni prima nella memoria cache L1, poi nella L2 e, solo in caso di insuccesso, accede alla memoria centrale; inoltre, quando una cella viene letta dalla RAM, il suo contenuto viene portato, oltre che alla CPU, anche nelle memorie cache, insieme a quello delle celle limitrofe.

A seconda del processore, la memoria cache può essere di varie dimensioni.

Valori tipici sono: 64 KB, 128 KB, 256 KB, 512KB, 1MB, 2MB.

Unità di Canale (input/output- I/O)

Un sistema di elaborazione è generalmente collegato a vari dispositivi di ingresso ed uscita chiamati periferiche, esempi sono la tastiera, il mouse, lo schermo, le stampanti, il modem le memorie di massa quali le unità disco ed il lettore di CD-ROM, il DVD

I dispositivi di Input/Output consentono l'interazione uomo-macchina e l'interazione macchina-macchina.

L'unità di canale ha il compito di interfacciare i vari dispositivi con il sistema di elaborazione, ovvero di tradurre i segnali interni del sistema di elaborazione in segnali accettati dalla periferica, e viceversa

Le interfacce di I/O, a seconda del tipo di periferica a cui devono essere collegate, sono dotate di:

- un registro dati per scambiare dati con la periferica
- un registro comandi per impartire comandi alla periferica
- un registro di stato per conoscere lo stato della periferica

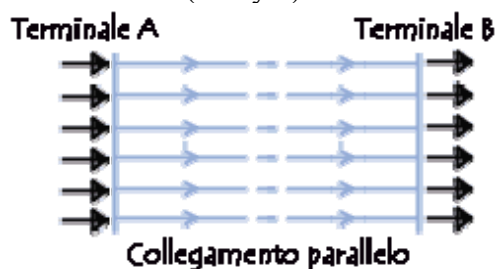
Il registro dati viene collegato al bus dati, il registro comandi ed il registro al bus controllo

Tradizionalmente un sistema di elaborazione è stato da subito dotato di dispositivi di comunicazione standard con modalità di trasmissione:

- trasmissione parallela
- trasmissione seriale

Trasmissione parallela

La trasmissione parallela è una modalità di comunicazione tra dispositivi digitali nella quale le informazioni sono inviate simultaneamente su più canali (fili). Le porte parallele presenti sui PC permettono di inviare simultaneamente 8 bit (un byte) attraverso 8 fili.

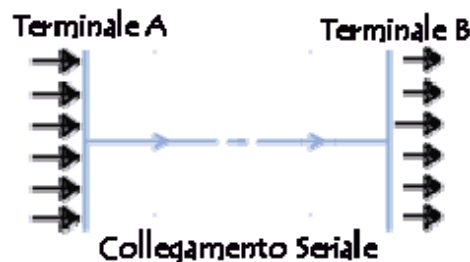


Questo collegamento era tipico delle stampanti degli scanner ora è stato abbandonato a favore della comunicazione seriale veloce USB.

Trasmissione seriale

La trasmissione seriale è una modalità di comunicazione tra dispositivi digitali nella quale le informazioni sono inviate una di seguito all'altra (ovvero in sequenza).

Le porte seriali rappresentano le prime interfacce che hanno permesso ai computer di scambiare delle informazioni con il "mondo esterno".



Le velocità di una trasmissione seriale si misura in bit/sec ovvero dal numero di bit trasmessi in un secondo

Universal Serial Bus USB

Il bus USB (Universal Serial Bus, in italiano Bus seriale universale) è basato su un'architettura di tipo seriale molto veloce rispetto alle porte seriali standard.

Direct Memory Access - DMA

La tecnica di collegamento DMA prevede che le unità di Input/Output siano collegate direttamente alla memoria centrale attraverso opportuni bus.

Scheda video

La scheda video gestisce il segnale video e la rappresentazione dei testi e delle immagine sullo schermo di un terminale video sia grafico che testuale.

Interruzioni - Interrupt

Nel caso in cui un programma preveda una operazione di I/O le attività richieste sono:

- verificare lo stato della periferica
- comunicare con la periferica
- attendere il completamento dell'operazione

Durante il periodo di attesa il processore non esegue alcuna operazione utile, mentre potrebbe fare altre cose.

Perché ciò sia possibile è necessario che il dispositivo di I/O possa avvertire il processore che ha concluso l'operazione richiesta.

La possibilità di avvertire il processore avviene per mezzo di un segnale chiamato segnale di interruzione o interrupt ed appartiene al bus di controllo.

Se è presente il segnale di interruzione le azioni che il processore esegue sono le seguenti:

- comunicare con la periferica e sospendere il programma
- ricercare ed eseguire altre attività fino a quando non arriva una segnalazione di interrupt.
- ripristinare il programma sospeso per proseguire l'attività

In generale la gestione dell' interruzione è demandata al sistema operativo della macchina che gestisce il segnale di controllo dell'interruzione secondo i seguenti passi:

1. interruzione dell'esecuzione del programma
2. verifica la natura dell'evento che ha generato l'interruzione
3. attivare il programma che gestisce l'evento
4. ritornare al programma interrotto se compatibile alla gestione dell'evento.

Un segnale di interruzione può essere generato da eventi di natura diversa:

1. interruzione di I/O: una periferica di I/O che informa su un suo particolare stato (p.es. pronta a ricevere dati) al fine di sincronizzarsi con il processore
2. interruzione per guasti al sistema rivelati da apposite sonde
3. interruzione per riportare l'intero sistema in uno stato noto (reset)
4. interruzione periodica generata da un orologio interno (p.es. ogni 10ms) per permettere al sistema operativo di calcolare il tempo speso da una applicazione e di cedere eventualmente il controllo ad un'altra applicazione (multiprogrammazione)
5. interruzione per errori nel programma correntemente eseguito (p. es. overflow, esecuzione di un'istruzione inesistente o privilegiata, etc.); tali interruzioni vengono anche chiamate traps
6. interruzioni programmate (software interrupt) generate da un programma che voglia accedere una risorsa condivisa (es. periferiche di I/O) e per questo chiede la mediazione del sistema operativo.

Dispositivi di archiviazione di massa

Disco rigido - HARD DISK

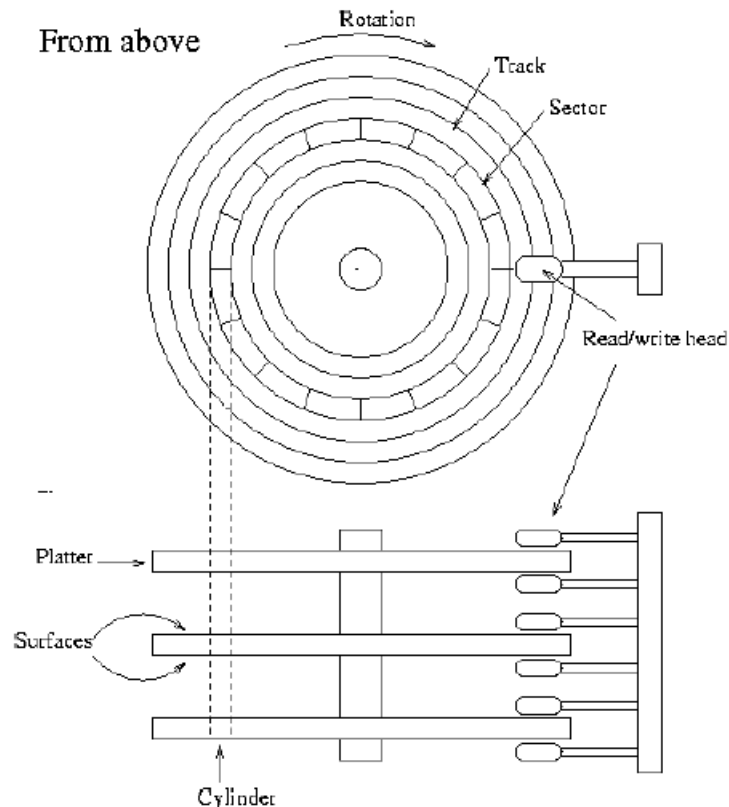
Il disco rigido è composto da:

- dischi rotanti rivestiti con una superficie di materiale magnetico
- testine mobili che accedono ai dischi per la lettura e scrittura di dati
- un dispositivo di interfaccia per comunicare con la memoria centrale e con l'unità di controllo.

Ogni superficie del disco è suddivisa in cerchi concentrici denominate tracce (track), ogni traccia è divisa in settori (Sector), in ogni settore sono memorizzati un identificativo del settore nel disco, i dati e un codice di correzione degli errori.

Nel caso di più dischi l'insieme di tutte le tracce che si trovano alla stessa distanza dal centro è chiamato cilindro, le testine sono relative a ogni superficie, sono collegate insieme e si spostano in modo solidale, in un dato istante ogni testina si trova sulla stessa traccia di ogni superficie.

La maggior parte dei controllori di disco includono una cache in cui vengono memorizzati i settori, a mano a mano che la testina li scorre;



L'accesso alle informazioni sul disco è così effettuato:

1. posizionamento della testina sulla traccia voluta: l'operazione è detta ricerca (seek) e il tempo necessario è detto tempo di ricerca, seek time; i produttori indicano valori minimo, massimo e medio del seek time (tempo medio: calcolato come la somma di tutti i possibili tempi di ricerca divisa per il numero di ricerche possibili).
2. tempo di attesa che il settore voluto giunga sotto la testina, detto tempo di latenza di rotazione. Latenza media è pari a metà di un giro.
3. trasferimento dei dati, caratterizzato dal tempo di trasferimento, ovvero dal tempo necessario per trasferire un blocco di bit. È funzione della dimensione del settore, della velocità di rotazione e della densità di registrazione lungo la traccia.

ESEMPIO DI CARATTERISTICHE

Diametro del disco che varia da un pollice a circa 3,5 pollici.

Velocità di rotazione: da 5.400 a 15.000 giri al minuto (RPM)

Numero di dischi:

Capacità supera abbondantemente gli 80 GB

Numero tracce per superficie: da 10.000 a 50.000 tracce.

Numero di settori per traccia è divisa in settori: 100 a 500 per traccia

Dimensione tipica di un settore è di 512 byte (fino a un massimo di 4096 byte).

Tempo d'accesso di un hard disk è dell'ordine dei millisecondi,

Tempo di ricerca viene indicato fra i 3 ms e i 14 ms, ma il tempo medio reale può essere solo il 25% -33% di quello indicato grazie alla località delle ricerche su disco.

Tempo di attesa

Frequenze di trasferimento (2004): da 30 a 80 MB/sec.

Nastri magnetici

I nastri magnetici sono caratterizzati da accesso sequenziale ai dati, e pertanto sono molto lenti. Vengono generalmente usati per i backup di grosse quantità di dati.

CD/DVD

I CD/DVD sono supporti “ottici” abbastanza simili agli hard disk: sono anch’essi organizzati in blocchi ma con la differenza che i dati sono scritti/letti tramite un laser.

La capacità di un CD arriva a 700 MB, mentre quella di un DVD a 4.7 GB.

PRESTAZIONI DI UN SISTEMA DI ELABORAZIONE

La prestazione di un sistema di elaborazione dipende da molti differenti fattori; *alcuni* indicatori per misurarli sono:

- MIPS (Milion Istructions Per Second): indica quanti cicli istruzione in assembler vengono svolti ogni secondo dal processore.
- MFLOPS – Millions of FLOating point operations Per Seconds (milioni di operazioni in virgola mobile al secondo): indica quante operazioni in virgola mobile vengono eseguite dal calcolatore. Tale indicatore è utilizzato soprattutto per misurare le prestazioni dei supercalcolatori.

ESEMPIO

Una CPU ha una potenza di calcolo da 10 GFlops.

ISTRUZIONI E PROGRAMMA

Programma eseguibile

Un programma eseguibile dalla macchina di Von Neumann consiste in un elenco di istruzioni registrate nella memoria centrale e scritte in codice binario. Tali istruzioni vengono eseguite in successione, che può venire alterata da eventuali istruzioni di controllo presenti nel programma.

Esecuzione di una istruzione

L'istruzione viene attivata dalla unità di controllo mediante il ciclo istruzione, il ciclo è suddiviso in tre operazioni base:

- prelievo (fetch)
- decodifica (decode)
- esecuzione (execute)

PRELIEVO (FETCH)

Consiste nel prelevare una istruzione dalla memoria centrale e trasferirla nell' instruction register (IR), secondo le seguenti operazioni elementari:

- unità di controllo (CU) scrive il contenuto del registro program counter (PC), che contiene l'indirizzo della cella di memoria contenente l'istruzione da eseguire nel registro indirizzi memoria (MAR) della memoria centrale.
- il contenuto della cella di memoria, che contiene l'istruzione da eseguire, è copiato nel registro dati memoria (MDR) e, successivamente, nel registro istruzione (IR) attraverso il bus dati.
- unità di controllo (CU) incrementa di una unità il contenuto del PC in modo che esso contenga l'indirizzo della prossima istruzione da eseguire.

DECODIFICA (DECODE)

Consiste nell'individuare, in base all'istruzione trasferita, l'operazione da effettuare e con quali eventuali operandi; l'operazione viene svolta da una unità di decodificazione denominata decodificatore di istruzioni presente nella unità di controllo (CU).

ESECUZIONE (EXECUTE)

In base alle segnalazioni della unità di decodifica di istruzioni, l'unità di controllo invia le direttive di esecuzione ai dispositivi coinvolti.

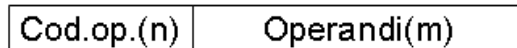
Le unità interessate all'esecuzione dell'istruzione vengono attivate.

Eventualmente vengono caricati gli operandi necessari (come nella fase di fetch) e/o scritti i risultati dell'operazione.

Struttura delle istruzioni

In generale la struttura di una istruzione eseguibile dalla unità di controllo (CU) è composta di due parti:

- la prima parte rappresenta il codice operativo dell'istruzione ed è formata da n bit
- la seconda parte rappresenta uno o più operandi necessari alla esecuzione dell'istruzione ed è formata formata da m bit



Il codice operativo specifica l'operazione da compiere, gli operandi specificano il dato oppure, in base a varie modalità, le locazioni di memoria in cui questo risiede.

Il codice operativo è sempre presente; gli operandi possono mancare.

Insieme delle istruzioni eseguibili

Le istruzioni eseguibili dalla unità di controllo (CU) sono operazioni molto elementari e possono essere raggruppate nelle seguenti classi:

- istruzioni aritmetico-logiche
- istruzioni di controllo del flusso
- istruzioni di interazione con la memoria centrale
- istruzioni di I/O

ISTRUZIONE ARITMETICO/LOGICO

Rientrano in questa categoria le operazioni aritmetiche e logiche. Nelle architetture di elaborazioni più complesse possono includere anche le operazioni in virgola mobile. Il dispositivo che viene attivato è l'ALU (Arithmetic Logic Unit).

ISTRUZIONE DI CONTROLLO DI FLUSSO

Le istruzioni di controllo di flusso opera sul program counter (PC) e sono:

- confronto
- salto
- salto e ritorno da subroutine

Confronto

Esegue il confronto tra due dati ed aggiorna solo alcuni bit del registro di stato della ALU.

Salto

Può essere assoluto o condizionato. Se è condizionato è preceduto da un'altra istruzione che verifica la condizione, di solito un'istruzione di confronto.

Salto a subroutine, ritorno da subroutine.

La subroutine è un segmento di programma che può essere attivato a seguito di opportune condizioni.

Il salto a subroutine è un'istruzione che deve salvare tutte le informazioni correnti relative al programma in esecuzione ed andare ad eseguire un altro segmento di programma. Quando incontra invece un'istruzione di ritorno da subroutine il calcolatore riprende ad eseguire il programma interrotto a partire dal punto (istruzione) dove lo aveva lasciato precedentemente.

ISTRUZIONE DI ITERAZIONE CON LA MEMORIA

Si tratta ad esempio del caricamento di registri dalla memoria e viceversa (si parla di istruzioni LOAD per i trasferimenti da memoria a registro e di istruzioni STORE per i trasferimenti da registro a memoria), del trasferimento di dati da registro a registro.

ISTRUZIONE DI I/O.

Sono istruzioni che servono a trasferire dati all'esterno o ad acquisire dati dall'esterno del calcolatore. Si possono realizzare in due modi:

- quando incontra un'istruzione di I/O, l'unità centrale informa le periferiche che intende interagire con dispositivi esterni e, fornendo l'indirizzo, indica anche con quale dispositivo intende interagire; in questo caso si utilizzano codici specifici e indirizzi particolari;
- in alternativa è possibile riservare una zona nello spazio di indirizzamento della memoria in modo che ogni volta che si va ad agire in questa zona (con un'istruzione di trasferimento dati con la memoria) in realtà si interagisce con l'esterno

Programma e linguaggi di programmazione

Scrivere un programma direttamente in linguaggio macchina è difficile, in quanto ogni istruzione è espressa in codice binario, pertanto è nata l'esigenza di usare versioni simboliche dei linguaggi macchina, detti linguaggi assembly.

Tali linguaggi, non potendo essere eseguiti direttamente dai circuiti dell'unità di controllo, richiedono la traduzione nella equivalente versione in linguaggio macchina.

ESEMPIO

```
001 00000          LOADI 0
```

Dal momento che, comunque, anche scrivere programmi in linguaggio assembly è molto noioso e complicato, poiché ogni istruzione svolge solo funzioni elementari, è nata l'esigenza di progettare strumenti più potenti e più vicini al linguaggio naturale. Tali linguaggi sono comunemente detti linguaggi simbolici ad alto livello.

ESEMPIO

FORTRAN, C, C++, PYTHON, PASCAL, JAVA, BASIC, etc.

La traduzione di questi linguaggi ad alto livello in codice eseguibile può essere fatta con due diverse modalità:

- mediante compilatore
- mediante interprete

COMPILATORE

Traduce l'intero programma sviluppato dall'utente (denominato codice sorgente) in linguaggio assembler e successivamente in linguaggio macchina o direttamente in linguaggio macchina lasciando "traccia" della traduzione, ossia producendo una copia fisica (in memoria) del programma tradotto (si tratta del cosiddetto codice oggetto). Se non si modifica il codice sorgente, ogni volta che si vuole eseguire il programma e' necessario avere a disposizione il solo codice oggetto.

INTERPRETE

E' un programma tramite il quale si esegue un altro programma. Esso preleva una istruzione alla volta, la esamina, la "interpreta" e la fa eseguire dal calcolatore, quindi passa alla successiva istruzione, senza lasciare "traccia" del programma tradotto.

Ogni volta che si vuol fare eseguire il programma e' necessario sia il programma sorgente che l'interprete.

Esecuzione di un programma compilato

I programmi ed i dati risiedono normalmente in memoria secondaria.

Quando si lancia un programma questo viene copiato dalla memoria secondaria in memoria centrale ad un determinato indirizzo; questa operazione si chiama caricamento ed è eseguita dal sistema operativo.

Poi il controllo viene passato al programma che deve essere eseguito.

Ogni tanto interviene il sistema operativo per valutare il corretto funzionamento del medesimo.

Quando il programma conclude la sua esecuzione, il sistema operativo rilascia la memoria occupata del programma.