Architettura di un elaboratore

Architettura di un elaboratore

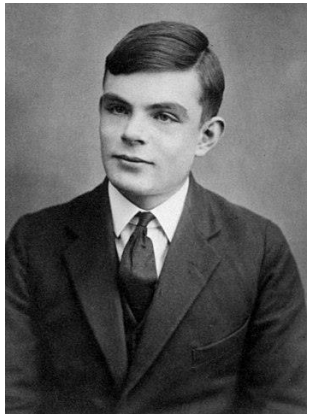
1936 – Turing, A.: «*On Computable Numbers, with an Application to the Entscheidungsproblem*»

- Calcolatore astratto «programmabile», in grado di eseguire qualunque funzione computabile. L'input è costituito non solo dai dati, ma anche dalla sequenza di operazioni da eseguire...

1945 – Nell'ambito dello sviluppo di calcolatori per il calcolo delle traiettorie di missili balistici e del Progetto Manhattan

- ENIAC: uno dei primi elaboratori digitali programmabili. Cambiare programma significa cablare nuovamente e riconfigurare la macchina
- EDVAC: il primo elaboratore digitale programmabile tramite software

Alan Turing 1912-1954



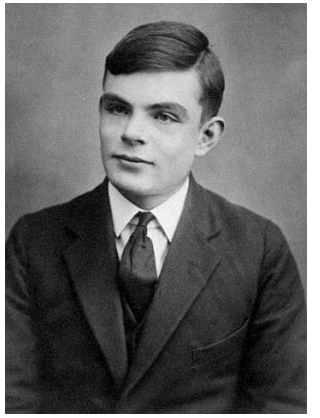
John Von Neumann 1903-1957

Architettura di un elaboratore

1945 – Von Neumann, J.: «*First Draft of a Report on the EDVAC*»

- John Von Neumann descrive l'architettura dell'EDVAC, cioè l'architettura base di un elaboratore digitale programmabile. Diventerà nota come architettura di Von Neumann

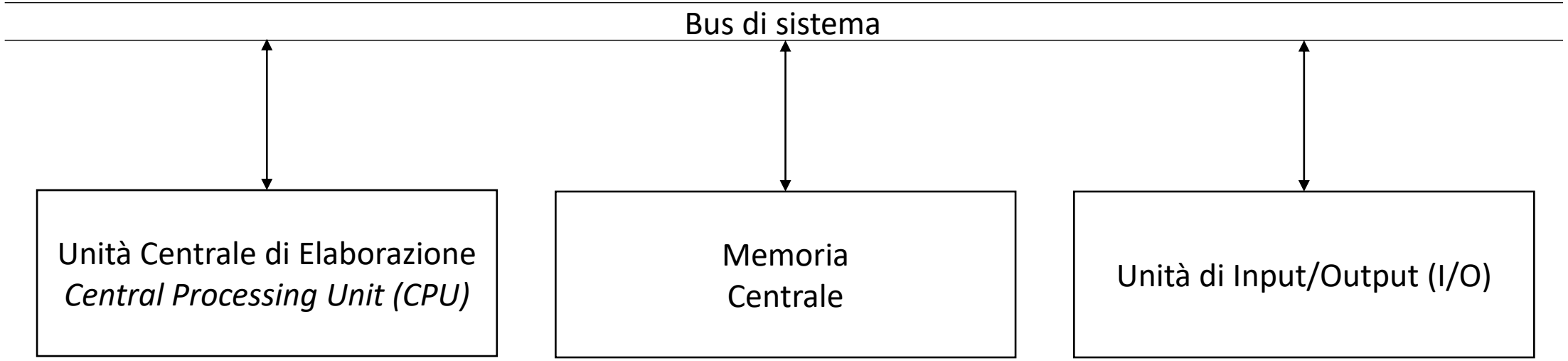
Alan Turing 1912-1954



John Von Neumann 1903-1957

Architettura di Von Neumann

Componenti fondamentali



Architettura di Von Neumann – CPU

Dirige l'esecuzione delle istruzioni. Specifiche della particolare famiglia di CPU, le istruzioni ne costituiscono il linguaggio macchina.

Contiene i dispositivi elettronici per acquisire ed eseguire le istruzioni che compongono un programma, elaborando i dati in ingresso in dati in uscita. In un unico circuito integrato è composta da:

- Unità di controllo (Control Unit – CU)
- Unità aritmetico-logica (Arithmetic Logic Unit – ALU)
- Registri

Architettura di Von Neumann – CPU

Dirige l'esecuzione delle istruzioni. Specifiche della particolare famiglia di CPU, le istruzioni ne costituiscono il linguaggio macchina.

- L'unità di controllo recupera la prossima istruzione da eseguire in memoria («Fetch»), la decodifica («Decode») e, dopo avere individuato i dati usati dall'istruzione, la esegue («Execute»)
- L'ALU esegue le operazioni aritmetiche sui dati
- I registri sono locazioni di memoria ad accesso molto veloce usate per memorizzare i dati su cui lavorano un'istruzione, indirizzi o istruzioni (per esempio la successiva istruzione da eseguire)

Architettura di Von Neumann - Memoria

La memoria centrale memorizza i dati da elaborare/elaborati e le istruzioni per eseguire l'elaborazione

E' composta da:

- RAM (Random Access Memory)
- ROM/EPROM (Read Only Memory/Erasable Programmable Read Only Memory)

Architettura di Von Neumann - Memoria

La memoria centrale memorizza i dati da elaborare/elaborati e le istruzioni per eseguire l'elaborazione

La RAM è una sequenza di locazioni (*celle*) di memoria identificate da un indirizzo. In queste celle vengono caricate le istruzioni dei programmi in esecuzione e i dati che i programmi elaborano.

- E' ad accesso casuale, cioè il tempo di accesso non dipende dalla posizione della cella
- E' volatile, cioè si perde il contenuto quando non si spegne l'elaboratore digitale

Architettura di Von Neumann - Memoria

La memoria centrale memorizza i dati da elaborare/elaborati e le istruzioni per eseguire l'elaborazione

L'accesso alla RAM è sia in lettura che in scrittura. Ad esempio, i dati elaborati dai programmi vengono scritti in RAM. Al contrario della RAM, la EPROM:

- Permette la sola lettura (*Read Only*)
- Non è volatile (mantiene le informazioni quando non alimentata)
- Contiene il firmware dell'elaboratore digitale, cioè le istruzioni per il suo corretto avvio (come le istruzioni per controllare quali periferiche di I/O sono connesse e le istruzioni per avviare il *bootloader* del sistema operativo).

Architettura di Von Neumann - Memoria

La memoria centrale memorizza i dati da elaborare/elaborati e le istruzioni per eseguire l'elaborazione

Oltre a RAM e ROM, in un elaboratore digitale sono presenti anche memorie cache, pensate per ospitare dati di frequente utilizzo con tempo di accesso più rapido ai dati rispetto alla RAM

- È volatile
- Quando la CPU deve prelevare dati ed istruzioni dalla RAM, questi vengono copiati in cache insieme a dati ed istruzioni vicine (località spaziale)
- La cache contiene dati e istruzioni usate più frequentemente (località temporale)

Architettura di Von Neumann - Memoria

La memoria centrale memorizza i dati da elaborare/elaborati e le istruzioni per eseguire l'elaborazione

Quando la CPU deve caricare dati o istruzioni da elaborare, controlla se sono già presenti in cache

- Se ci sono, li preleva direttamente dalla cache, anziché dalla RAM (accesso più veloce)
- Se non ci sono, li carica dalla RAM alla cache (futuri accessi più veloci).

Memoria secondaria

- Non è volatile: i dati restano allo spegnimento del computer.
- Serve per immagazzinare *grandi quantità di dati* (ad oggi, contiene centinaia di GB o qualche TB).
- I dati e i programmi contenuti nei file negli hard disk (memoria secondaria), verranno caricati in RAM (memoria centrale) per essere elaborati ed eseguiti.
- Oltre a HDD e SSD, sono memoria secondaria pendrive USB, schede SD, DVD, CD…

Architettura di Von Neumann - Memoria

Perché è necessaria una tale organizzazione gerarchica della memoria in un elaboratore digitale?

Architettura di Von Neumann - Memoria

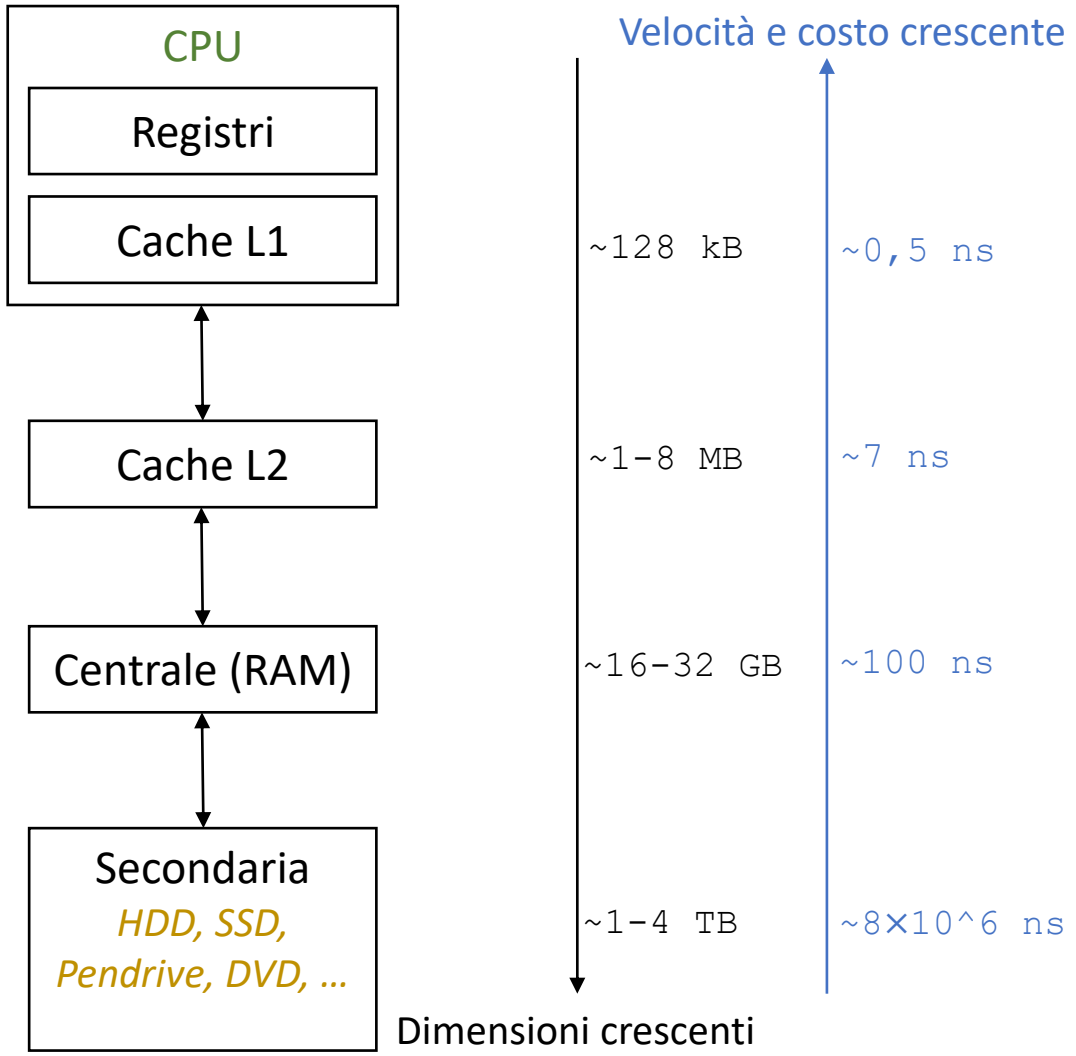
«Teoricamente si vorrebbe avere una memoria di capacità indefinitamente grande, tale che ogni particolare parola possa essere immediatamente disponibile. Siamo così costretti a riconoscere la possibilità di costruire una gerarchia di memorie, ognuna delle quali ha capacità superiore a quella precedente, ma a cui si può accedere con minor velocità.»

John Von Neumann

Organizzazione gerarchica della memoria

- La memoria è più lenta della CPU e tende a limitarne le prestazioni
- Occorre un compromesso tra costo, prestazioni e dimensione della memoria

Poca memoria <u>veloce vicino alla CPU</u> e tanta memoria <u>lenta per memorizzare informazioni</u>



Gerarchia memorie – Principi di località

L'obiettivo ideale sarebbe fornire una quantità di memoria pari a quella disponibile nella tecnologia più economica garantendo una velocità di accesso pari a quella della memoria più costosa.

L'organizzazione gerarchica della memoria cerca il miglior compromesso tra costi, dimensione e prestazioni. In particolare, la cache è gestita sfruttando due principi di località nell'esecuzione delle istruzioni di un programma

- Località temporale
- Località spaziale

Gerarchia memorie – Principi di località

Località temporale: tendenza a riferirsi allo stesso elemento entro breve tempo (es. ripetizioni all'interno degli algoritmi)

Località spaziale: tendenza a riferire la successiva lettura/scrittura in memoria ad elementi che hanno indirizzo vicino a quello dell'elemento corrente (es. istruzioni sequenziali)

Architettura di Von Neumann – I/O

Periferiche di Input/Output (I/O): dispositivi che permettono l'interazione tra elaboratore digitale e utente, ambiente e altri elaboratori digitali.

- Periferiche di Input: interazione in ingresso per l'elaboratore digitale (es. tastiera, mouse…)
- Periferiche di Output: interazione in uscita dall'elaboratore digitale (es., schermo, stampante…)

Architettura di Von Neumann – Bus

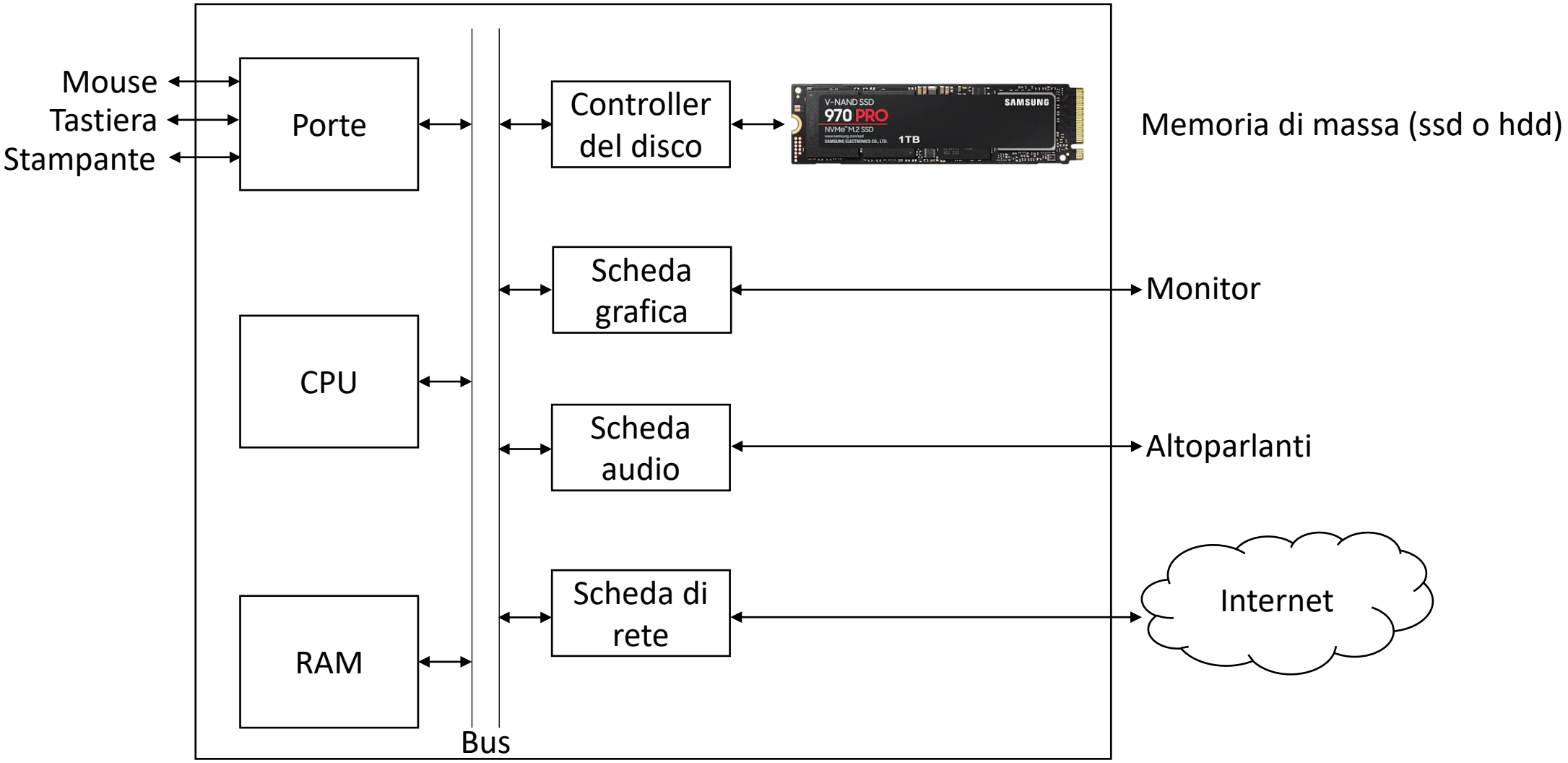
I byte fluiscono tra CPU, memoria e periferiche, sotto il controllo della CPU, attraverso il bus di sistema.

Componenti di un PC

Architettura di un PC

Anche un moderno PC, come tutti gli elaboratori digitali, si basa sull'architettura di Von Neumann.

Architettura di un PC

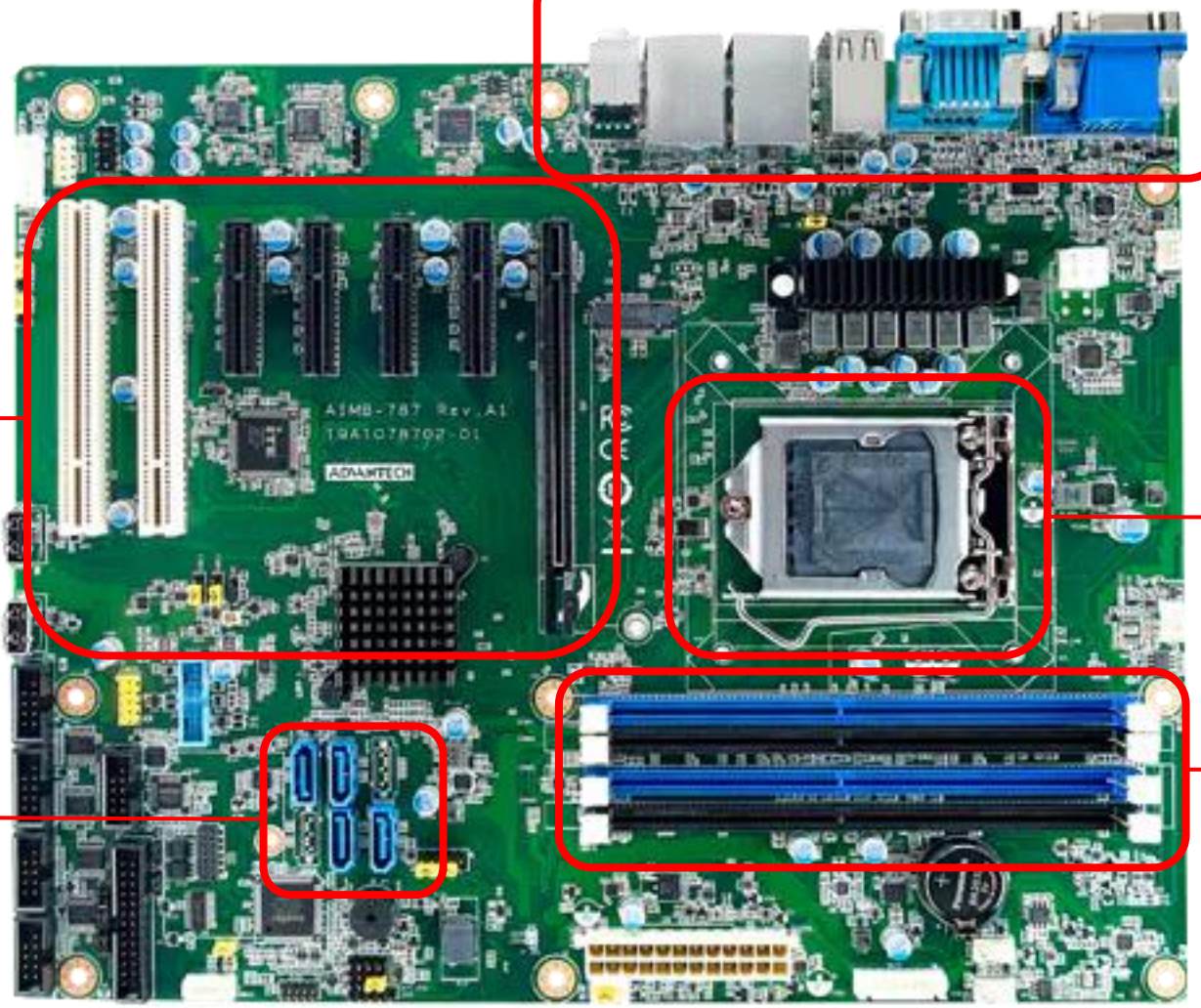


Architettura di un PC – Scheda Madre

E’ il circuito stampato principale di un elaboratore digitale «general purpose», cioè un computer.

Permette la comunicazione tra CPU e RAM attraverso il bus e fornisce i connettori per le periferiche

Architettura di un PC – Scheda Madre



Connettori periferiche (USB, Rete, Monitor…)

Slot di espansione (per schede grafiche, audio…)

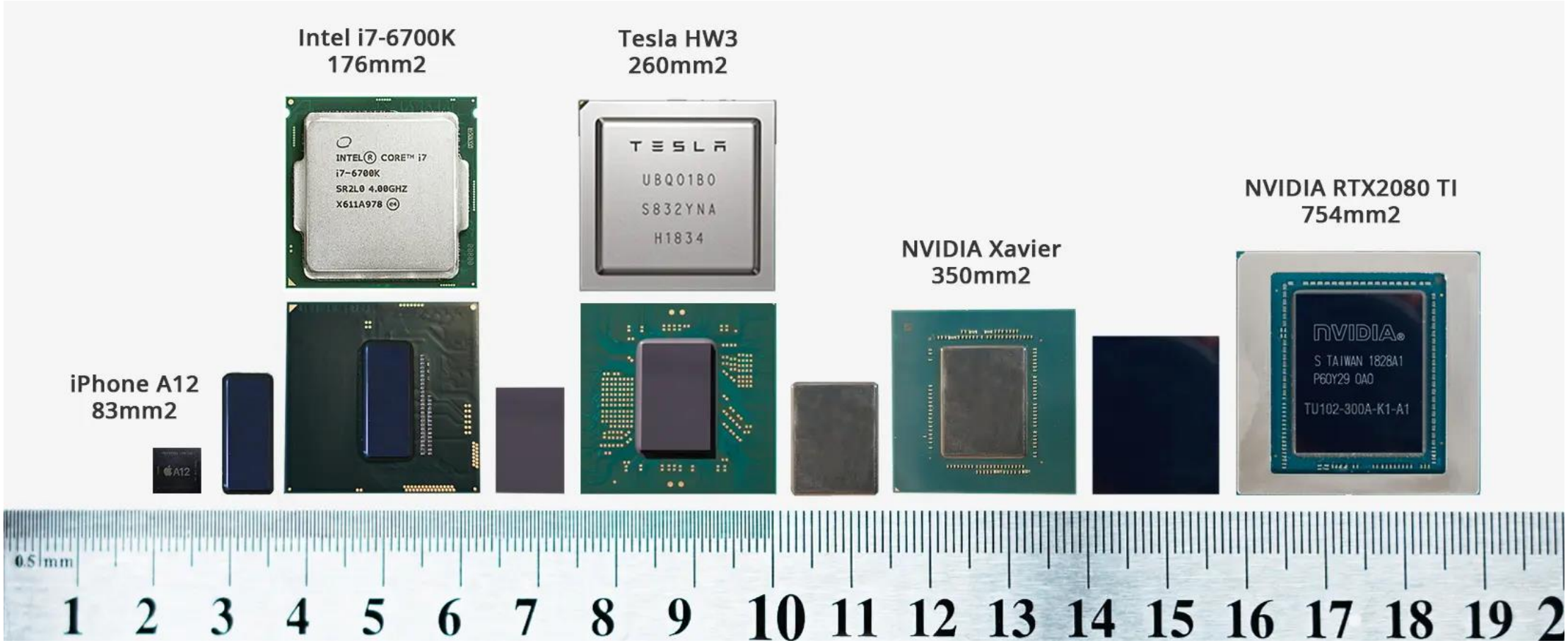
Socket CPU

Connettori SATA (per memoria di massa)

Slot RAM

CPU – Microprocessore

CPU – Microprocessore



CPU – Microprocessore

Una CPU per un comune PC ha all'interno miliardi di transistor (~10). I transistor sono i blocchi costitutivi di una CPU e, intuitivamente, sono microscopici interruttori (On/Off – 1/0) che formano le porte logiche e permetto ad una CPU di eseguire le istruzioni.

Frequenza di clock: Numero di commutazioni tra i due livelli «0» e «1» che i circuiti all'interno di una CPU (o di un componente elettronico) sono in grado di eseguire in un secondo. Una CPU a 3 GHz è in grado di eseguire 3.000.000.000 di commutazioni al secondo…

Più alto è il clock più «veloce» sarà la CPU…

CPU – Parallelismo

Un sistema multiprocessore è un elaboratore con più CPU collegate insieme per consentire l'elaborazione simultanea di più istruzioni («in parallelo»), aumentando la velocità di esecuzione, o per garantire ridondanza e quindi tolleranza ai guasti.

Processore Multi Core: microprocessore costruito in un unico circuito integrato costituito da due o più unità di elaborazione (i «core»), capaci di eseguire istruzioni in parallelo.

Ad un maggior numero di core corrisponde una CPU più «veloce»…

CPU – Microprocessore

Più alto è il clock più «veloce» sarà la CPU…

Ad un maggior numero di core corrisponde una CPU più «veloce»…

Vero, ma…

- Difficile comparare CPU di generazioni e produttori diversi
- La frequenza può fluttuare tra più valori, in base al carico di lavoro e l'energia richiesta (es. «*thermal throttling*»)
- Per valutare le CPU si aggiungono anche altre metriche come i FLOPS (FLoating point Operations per Second). CPU desktop sono sull'ordine delle centinaia di miliardi di FLOPS (1 miliardo di FLOPS = 1 GFLOPS).

CPU – Microprocessore

- Un altro parametro è il TDP (Thermal Design Power), che fornisce un'indicazione del «calore» dissipato da una CPU
- Esistono risorse online per confrontare CPU e i diversi parametri con test per valutarne le prestazioni («benchmark»). Esempi:
  - https://www.cpubenchmark.net/
  - https://cpu.userbenchmark.com
  - https://www.tomshardware.com/reviews/cpu-hierarchy,4312.html

CPU – Microprocessore

Come si produce un microprocessore?

L'elemento base è il silicio: secondo elemento per abbondanza nella crosta terrestre (27% del peso).

Nei processori deve essere puro: viene fuso e purificato per ottenere lingotti cilindrici (30 cm di diametro).

Da lingotto, con una «affettatrice» vengono ricavati dischi spessi 1mm: i «wafer».

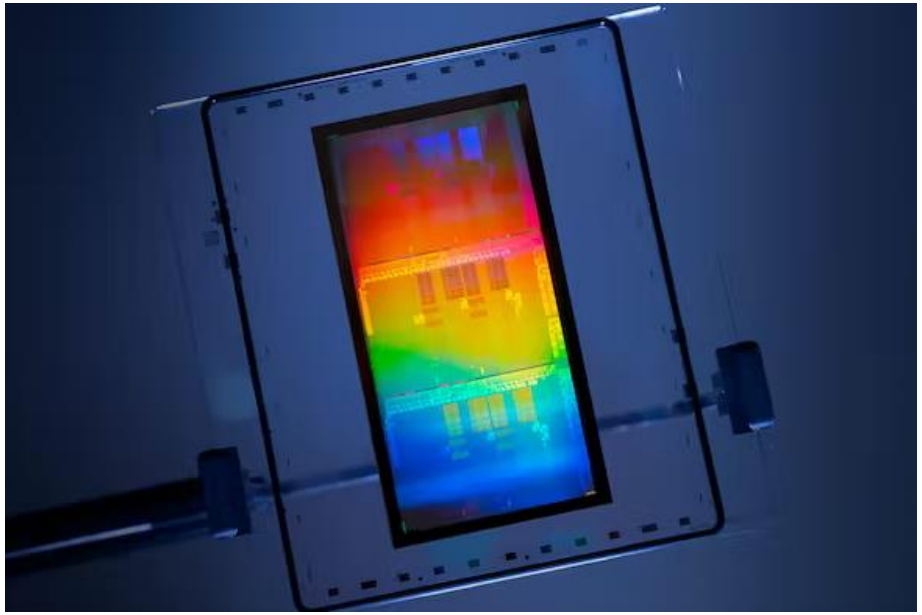
CPU – Microprocessore

Una volta progettato un microprocessore, il produttore crea delle «maschere», degli stampati che rappresentano i transistor e i blocchi di un processore.

La maschera è un blocco di quarzo serigrafato

Le maschere vengono usate per imprimere gli schemi nel wafer mediante fotolitografia

Il wafer viene quindi rivestito di un materiale fotoresistente in maniera uniforme



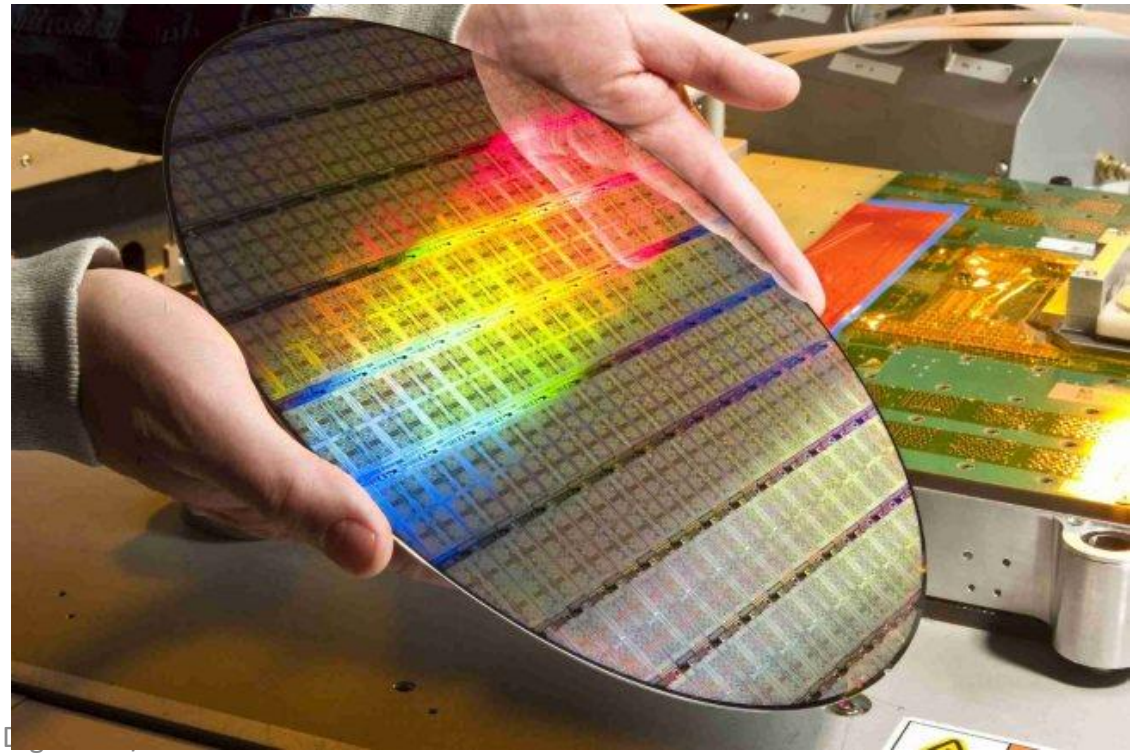
CPU – Microprocessore

Una raggio da un proiettore di luce colpisce il wafer illuminando il materiale fotoresistente attraverso le maschere.

L'immersione del wafer in una soluzione apposita rimuove tutte le parti dove la luce ha colpito.

Il processo viene ripetuto più volte, fino a formare diversi strati sul wafer

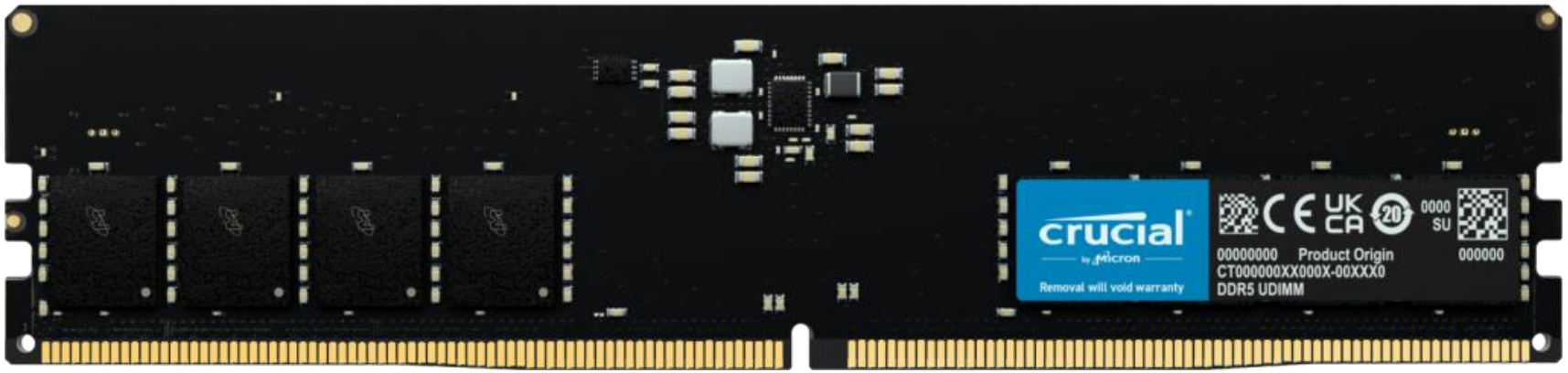
Alla fine del processo, il wafer contiene più chip, che vengono tagliati



CPU – Microprocessore

Al microscopio: https://www.youtube.com/watch?v=Fxv3JoS1uY8

RAM



La RAM è realizzata mediante tecnologia SDRAM (Synchronous Dynamic Random Access Memory). Semplificando, RAM possono essere confrontate per

- Dimensione – **Oggi**, un buon sistema desktop a 16 GB
- Standard DDR – DDR1, DDR2, DDR3, DDR4, DDR5. DDR5 è appena uscito e, tra una famiglia e la successiva, c'è tipicamente un aumento della banda (aumento della capacità di trasferire informazioni in un secondo)
- Frequenza di clock – Occorre assicurarsi che quella RAM sia compatibile con processore e scheda madre

Memoria di massa – Hard Disk

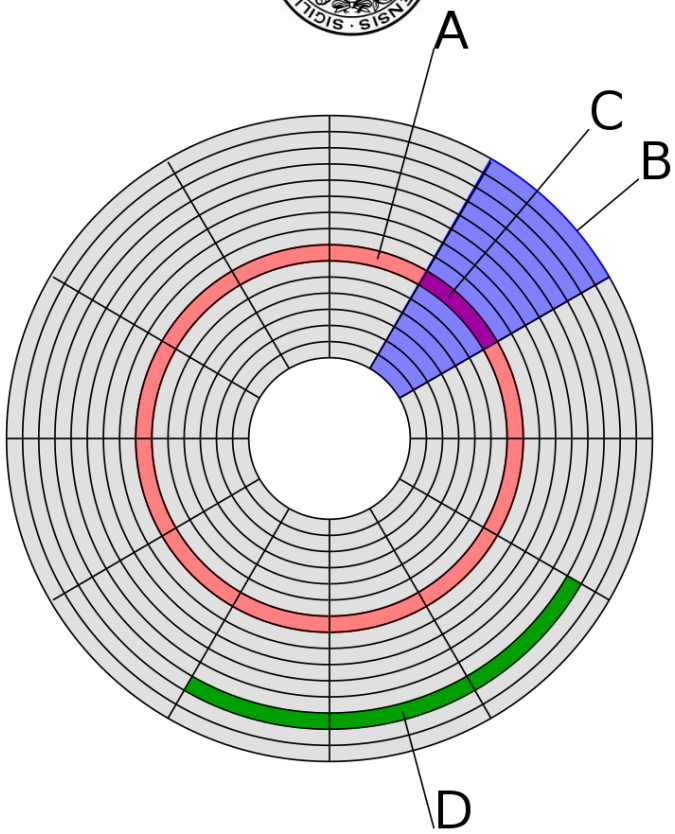
Memoria di massa – Hard Disk

- HDD – Hard Disk Drive
- Sono dischi magnetici costituito da uno o più piatti sovrapposti di alluminio/vetro rivestiti da materiale ferromagnetico
- Una testina per lato legge/scrive i settori leggendo/cambiando il verso della magnetizzazione

Parametri

- Dimensione (ormai alcuni Terabyte)
- Giri per minuto – RPM – (5400, **7200**, 10400, 15000). Influisce su tempo di accesso
- Dimensione Cache

Memoria di massa – Hard Disk

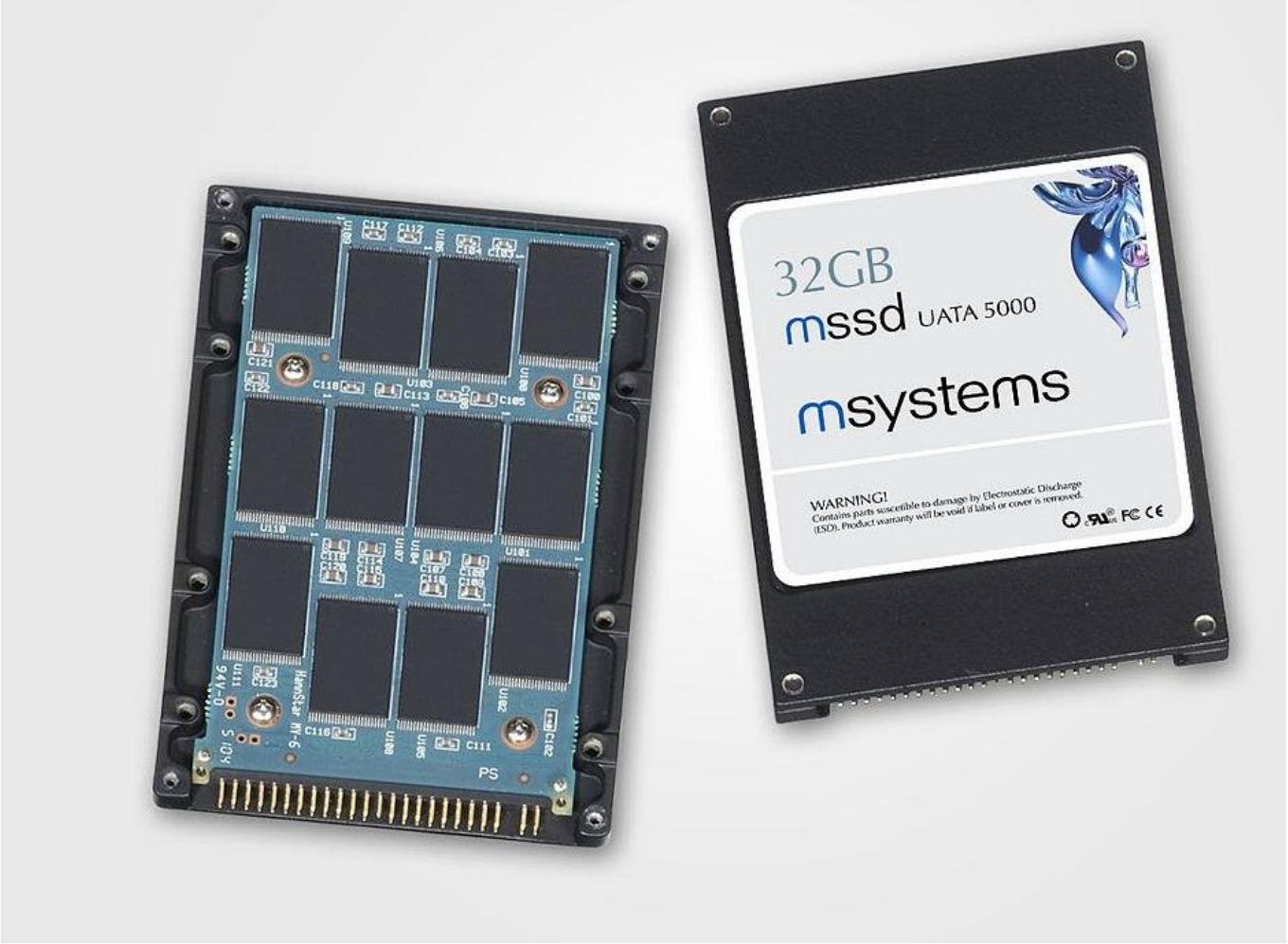


- A – traccia
- B – settore
- C – settore di una traccia
- D – cluster

Memoria di massa – Hard Disk

- Componente meccanico: può rompersi o smagnetizzarsi
- Superati in velocità da unità a stato solido (SSD – Solid State Disk): in media 520 MB/s per lettura e scrittura contro 125 MB/s
- Oggi, la ridotta differenza di prezzo tra SSD e HDD non giustifica l'uso di un HDD a discapito di un SDD per un sistema desktop, se non per ragioni come lo storage di grandi quantità di dati a scopo di backup.

Memoria di massa – SSD



Memoria di massa – SSD

- Basato su memorie flash «NAND» realizzate mediante *transistor*. Non necessita di parti meccaniche (a differenza degli HDD)

Parametri

- Velocità di lettura e scrittura sequenziale (tipicamente 560 MB/s in lettura e 520 MB/s in scrittura per SSD Sata, oltre i 2000 MB/s in lettura e scrittura per SSD PCIE M.2)
- Resistenza, espressa in vari modi MTBF (Mean Time Between Failures – es. 1.800.000 h), quantità totale di dati che un SSD può scrivere nel corso della sua vita (es. 180 TB), Classe di resistenza (es. 5 anni a 195 GB al giorno)
- Interfaccia di comunicazione (SATA o PCIE) e form factor (2.5'', mSATA, M.2). Gli SSD PCIE sono molto più veloci degli SSD sata, a loro volta più veloci dei tradizionali HDD

Memoria di massa – SSD