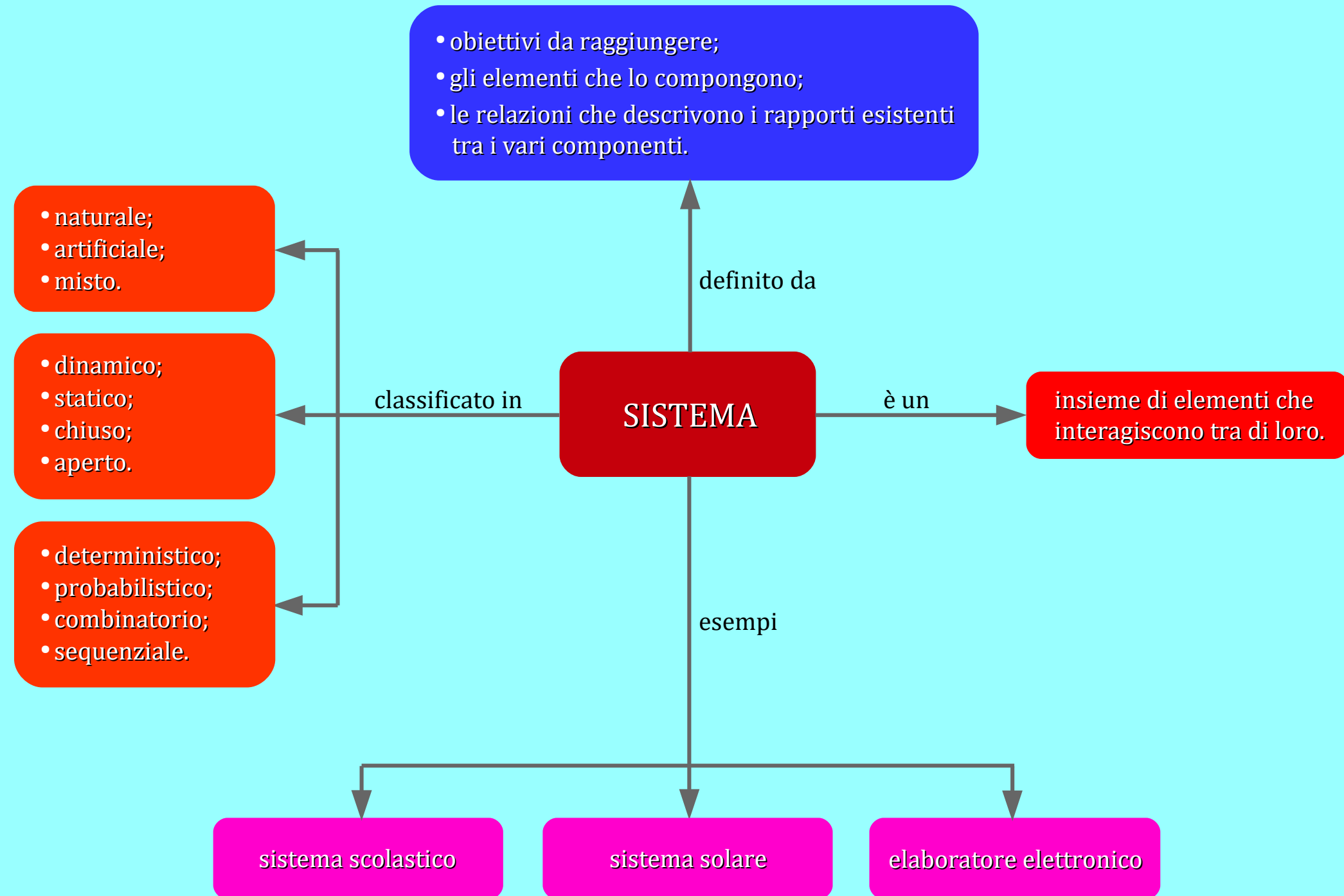


SISTEMI E RETI

Unità 1

Il sistema di elaborazione

Mappa concettuale



Mappa concettuale

ELABORATORE

è un

SISTEMA DI
ELABORAZIONE

in grado di

ricevere degli input dall'esterno e
fornire l'output dell'elaborazione.

basato su

MODELLO DI
VON NEUMANN

quindi è un

SISTEMA APERTO

descrive

Come è fatto

Come funziona

ARCHITETTURA DI
UN SISTEMA

I/O

CPU

MEMORIA

BUS

FETCH

DECODE

EXECUTE

■ Sistema

Un insieme di elementi (sottosistemi) che interagiscono al fine di raggiungere un preciso scopo. Sistema deriva dal greco “syn-istemi”: parti che “stanno insieme” perchè connesse da reciproche relazioni.

■ Sistema aperto

Un sistema che interagisce con l'ambiente esterno. Un calcolatore è un sistema aperto.

■ Sistema artificiale

Un sistema realizzato dall'uomo.

■ Sistema deterministico

Un sistema in cui ad una certa sollecitazione risponde con una sola e univoca risposta. Un programma per computer è deterministico in quanto a fronte di uno stesso input fornisce sempre lo stesso output (uscita predeterminata).

■ Sistema combinatorio

Un sistema in cui l'uscita dipende solo dagli ingressi e non dallo stato interno. Tali sistemi sono detti anche sistemi “senza memoria”.

■ Sistema sequenziale

Un sistema in cui l'uscita dipende sia dagli ingressi sia dallo stato interno.

■ Architettura di un sistema

Per architettura di un sistema si intende come è fatto e come funziona il sistema, ossia quali sono le componenti che lo formano e come interagiscono fra loro. Un particolare sistema di elaborazione è il computer: una macchina in grado di acquisire dall'esterno dati e istruzioni e produrre in uscita i risultati dell'elaborazione.

■ Programma

Una sequenza ordinata di istruzioni che trasforma i dati ricevuti e fornisce i risultati.

■ Modello di Von Neumann

Il modello di Von Neumann prevede che l'elaboratore sia composto da:

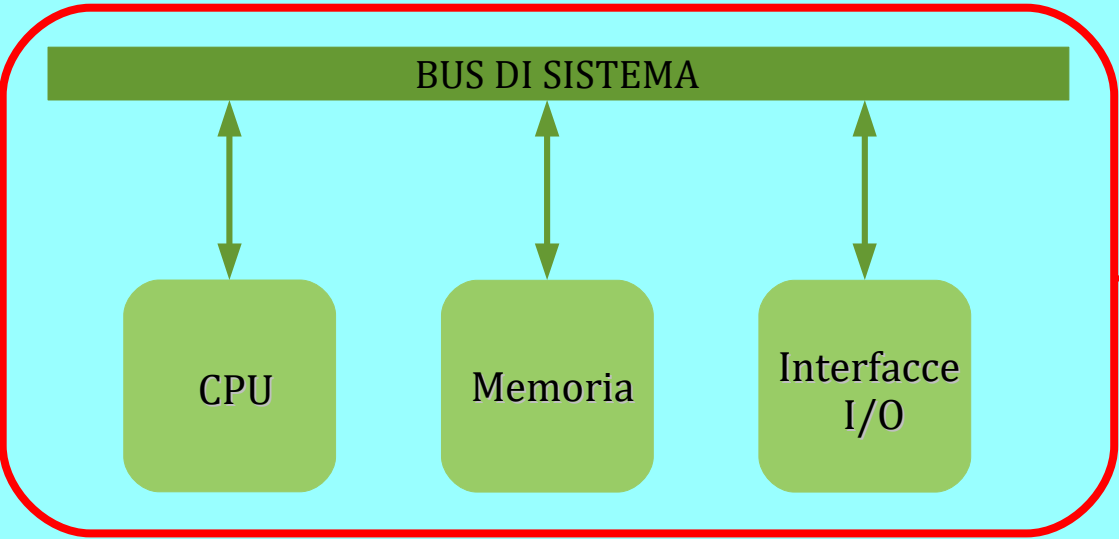
- **processore (CPU):** legge istruzioni e dati e fornisce i risultati dell'elaborazione;
- **memoria (RAM):** contiene informazioni, cioè istruzioni e dati;
- **interfacce di Input/Output(I/O):** collegamenti tra sistema e mondo esterno;
- **bus di sistema:** collega tra loro le componenti della macchina di Von Neumann.

■ Bus di sistema

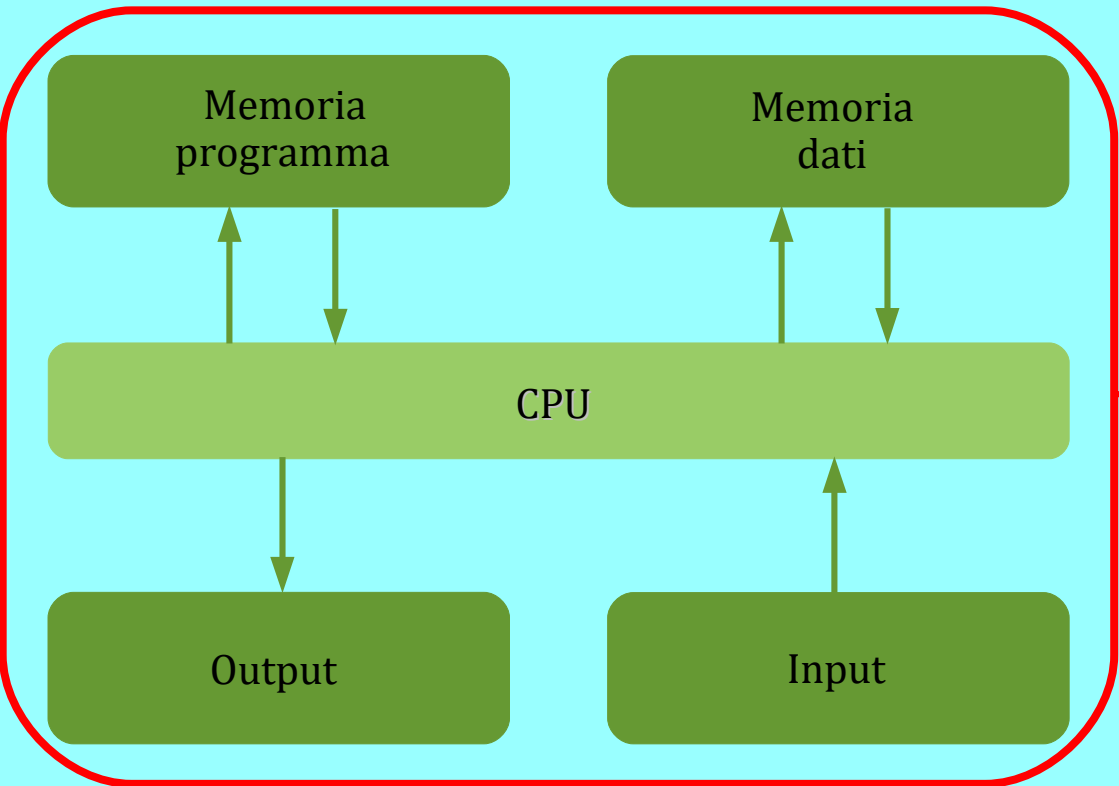
Il bus è costituito da linee che collegano la CPU, memoria e periferiche. Si suddivide in:

- **bus indirizzi:** serve alla CPU per attivare una cella di memoria o una periferica;
- **bus dati:** serve per trasferire dati tra la CPU e le periferiche e le periferiche stesse;
- **bus di controllo:** consente alla CPU e alle periferiche di stabilire e mantenere una comunicazione che consenta il trasferimento dati.

■ Confronto tra modello di Von Neumann e modello di Harvard



- Modello di Von Neumann
- Una sola memoria per dati e istruzioni.
 - Esecuzione sequenziale delle operazioni.
 - Architettura economica ma lenta.
 - Applicazioni: CPU general purpose.



- Modello di Harvard
- Due memorie diverse per dati e istruzioni.
 - Architettura ad elevate prestazioni.
 - Elevata complessità realizzativa.
 - Applicazioni: CPU specializzate.

1 Che cosa si intende per “sistema”?

Un insieme di elementi che cooperano al fine di raggiungere un obiettivo comune.

2 Che cos'è un sistema artificiale?

È un sistema costruito dall'uomo.

3 Che cos'è un sistema aperto?

È un sistema che interagisce con il mondo esterno.

4 Che cos'è un sistema deterministico?

È un sistema in cui l'uscita è predeterminata e prevedibile: a fronte dello stesso input fornisce sempre il medesimo output.

5 Fornisci almeno due esempi di sistema naturale, cioè che si presenta in natura.

Una catena montuosa, il sistema solare.

6 Fornisce almeno un esempio di sistema artificiale, individuando le parti di cui è costituito e lo scopo per cui è stato creato.

Lavatrice: usata per lavare indumenti e tessuti vari. Componenti: oblò, vaschetta per detersivo, filtro, serpentina,

Pianoforte: usato per produrre suoni. Componenti: tastiera, pedali, cassa, martelletti,

7 Spiega la differenza tra sistema e sottosistema.

Un sottosistema è uno degli elementi di un sistema.

8 Cosa significa che possiamo considerare un sistema come una “scatola nera”?

Significa che può comunicare con il mondo esterno (input e output).

9 Perché la macchina di Von Neumann può essere considerata un sistema deterministico?

Perché a fronte di un input, dato un programma, l'output è predeterminato.

10 Trova almeno tre sinonimi del termine “elaborare”.

Trasformare, modificare, processare.

11 Perché un computer può essere considerato un sistema?

Perché è costituito da sottosistemi (insieme delle componenti: CPU, memoria, periferiche, ...) collegati, che lavorano per un unico scopo.

12 Che cosa si intende per “architettura di un sistema”?

Come è fatto e come funziona.

13 Qual è stata l'innovazione di Von Neumann?

Creare una macchina per elaborare i dati in cui la CPU e la memoria sono componenti separate.

14 Quali sono le componenti dell'architettura di Von Neumann?

Processore, memoria centrale, interfacce di input/output, bus di sistema.

15 Come funziona la macchina di Von Neumann?

Dati e istruzioni vengono caricati nella RAM a partire dalla prima istruzione fino all'ultima in modo sequenziale; per ogni istruzione vengono effettuate la fase di fetch, decode, execute.

16 Che cos'è un programma?

Una sequenza ordinata di istruzioni scritte in un linguaggio di programmazione che trasforma i dati ricevuti e fornisce i risultati (le informazioni).

17 Perché un programma è un sistema deterministico?

Perché a fronte di uno stesso input fornisce sempre lo stesso output (uscita predeterminata).

18 Quali sono le funzioni del bus di sistema?

Collega tra loro le componenti del sistema.

19 Qual è la differenza tra dato e istruzione?

Un'istruzione è una operazione applicata ai dati.

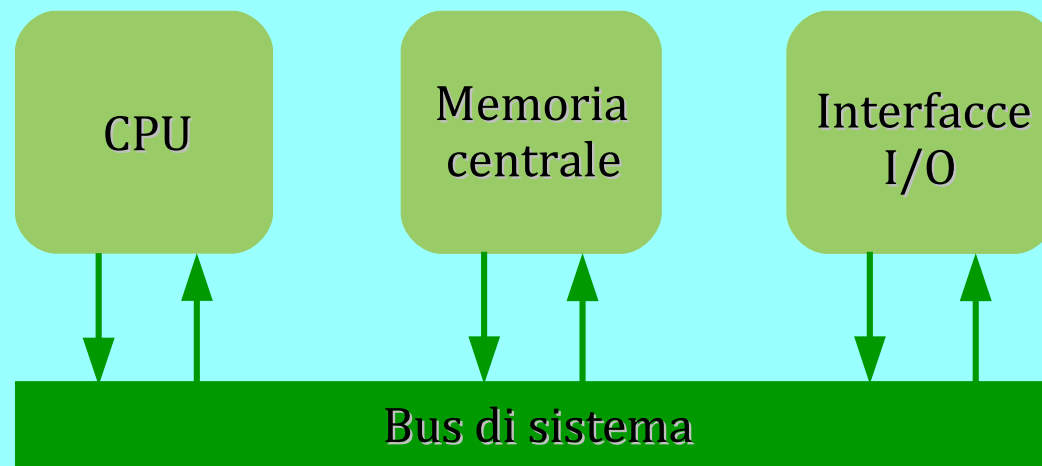
20 Per ogni tipo di bus, descrivere la direzione.

Bus dati: bidirezionale (dalla CPU alla memoria o alle periferiche e viceversa).

Bus indirizzi: monodirezionale (dalla CPU alla memoria o alle periferiche).

Bus controllo: monodirezionale, nel senso che ogni linea ha un'unica direzione entrante o uscente rispetto alla CPU (dipende dal segnale).

21 Esponi, avvalendoti di un disegno, come è fatta l'architettura di Von Neumann, mettendo in evidenza le quattro componenti.



22 Completa la tabella per classificare un sistema di elaborazione (metti X se la proprietà è vera).

Sistema	
Chiuso	
Aperto	X
A tempo continuo	
A tempo discreto	X
Artificiale	X
Probabilistico	
Dinamico	
Naturale	
A tempi discreti	X

23 Completa la seguente tabella inerente al bus di sistema.

Bus	Funzione	Direzione
Bus dati	Trasportare dati dalla CPU alla RAM e viceversa. Trasportare istruzioni dalla RAM alla CPU.	Bidirezionale
Bus indirizzi	Trasportare un indirizzo di memoria o periferica.	Unidirezionale
Bus controllo	Trasportare segnali precisi generati dalla CPU o dai dispositivi collegati.	Unidirezionale

Unità 2

L'elaboratore

Mappa concettuale

Processo attraverso il quale la CPU esegue un'istruzione ed è suddiviso in tre fasi.

MICROPROCESSORE

fisicamente
rappresentata
da

CPU

componenti
interni

ALU

Unità di controllo

Registri

uso speciale

uso generale

costruita
con architettura

ciclo
macchina

sincronizzato

Fetch

Decode

Execute

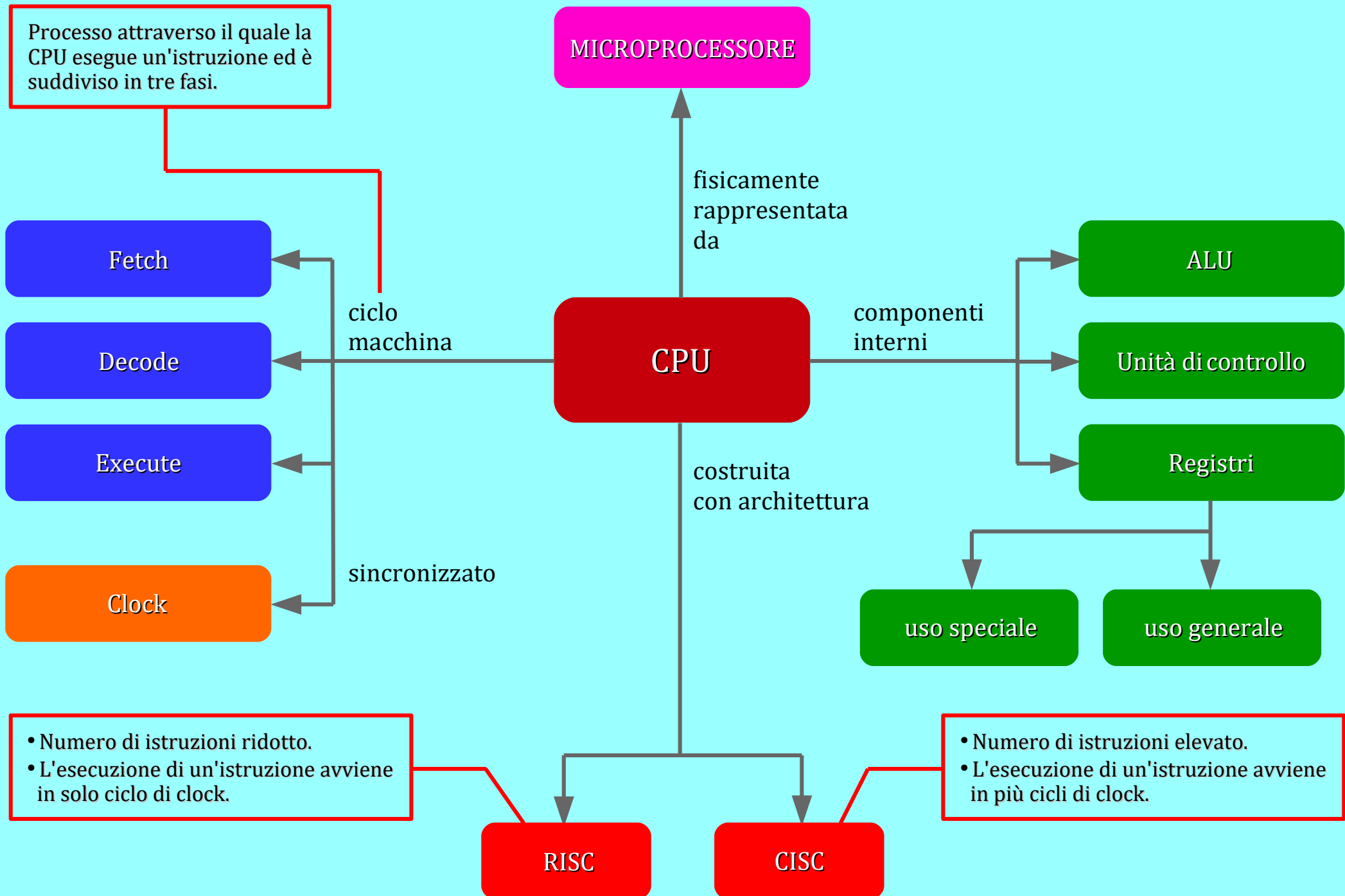
Clock

- Numero di istruzioni ridotto.
- L'esecuzione di un'istruzione avviene in solo ciclo di clock.

RISC

CISC

- Numero di istruzioni elevato.
- L'esecuzione di un'istruzione avviene in più cicli di clock.



■ CPU (Central Processing Unit)

La CPU è la parte del computer che esegue l'elaborazione dei dati, ed è rappresentata fisicamente dal microprocessore. E' composta dai seguenti elementi:

- **unità aritmetico-logica** (ALU): esegue i calcoli elementari e le operazioni logiche;
- **unità di controllo** (CU): governa e impartisce gli ordini di esecuzione all'ALU;
- **registri di appoggio**: piccole aree di memoria molto veloci, usate per memorizzare provvisoriamente i dati utilizzati per l'esecuzione dei calcoli.

■ Registri di uso speciale

Sono registri che svolgono una specifica funzione e in particolare sono:

- **Program Counter** (PC): contiene l'indirizzo della prossima istruzione da eseguire;
- **Status Register** (SR): insieme di bit, ciascuno dei quali fornisce informazioni sullo stato in cui si trova il processore. Alcuni bit sono di stato, altri di controllo;
- **Stack Pointer** (SP): contiene l'indirizzo della cima dello stack. Lo stack è un'area di memoria in cui i dati possono essere inseriti solo dall'alto;
- **Instruction Register** (IR): contiene il codice operativo dell'istruzione;
- **Memory Address Register** (MAR): contiene l'indirizzo che seleziona la locazione di memoria oppure il dispositivo di I/O coinvolto nell'operazione;
- **Memory Data Register** (MDR): contiene i dati che devono essere scritti in memoria oppure i dati letti dalla memoria.

■ Registri di uso generale

Sono registri destinati a memorizzare gli operandi e i risultati parziali dell'elaborazione.

■ Unità aritmetica-logica (ALU)

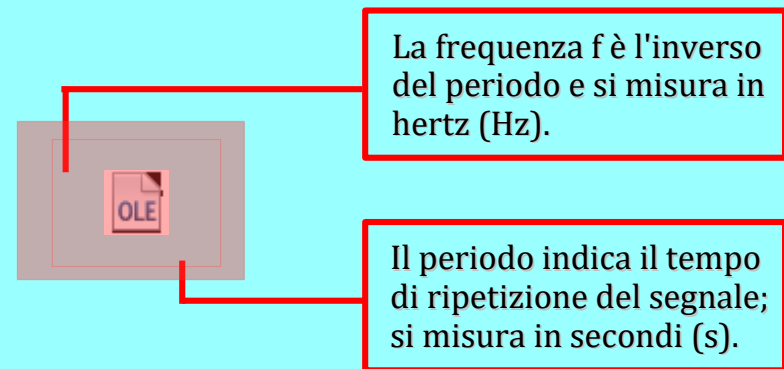
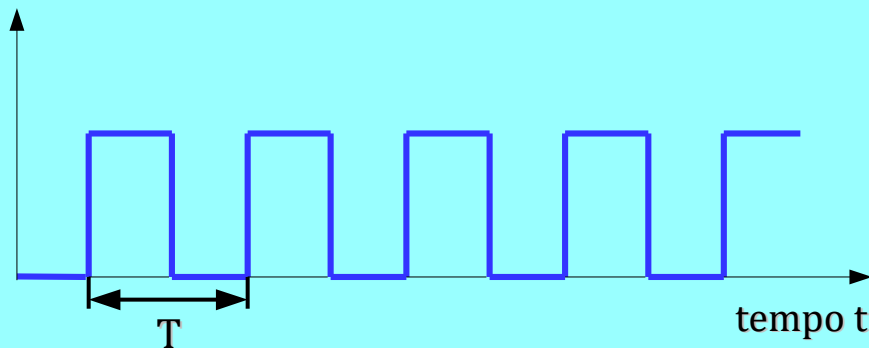
L'ALU è formata da porte logiche opportunamente collegate che eseguono le operazioni aritmetiche (somma, sottrazione, moltiplicazione, divisione) e logiche (AND, OR, NOT). I dati su cui opera possono essere presenti nei registri interni o in memoria e produce un risultato che può essere salvato su un registro o in memoria.

■ Unità di controllo (CU)

La CU coordina le operazioni svolte dalla CPU. I trasferimenti di dati o istruzioni tra CPU e memoria sono coordinati attraverso segnali di temporizzazione. I segnali sono ricevuti dalle componenti (memoria, I/O) al fine di eseguire il compito loro assegnato.

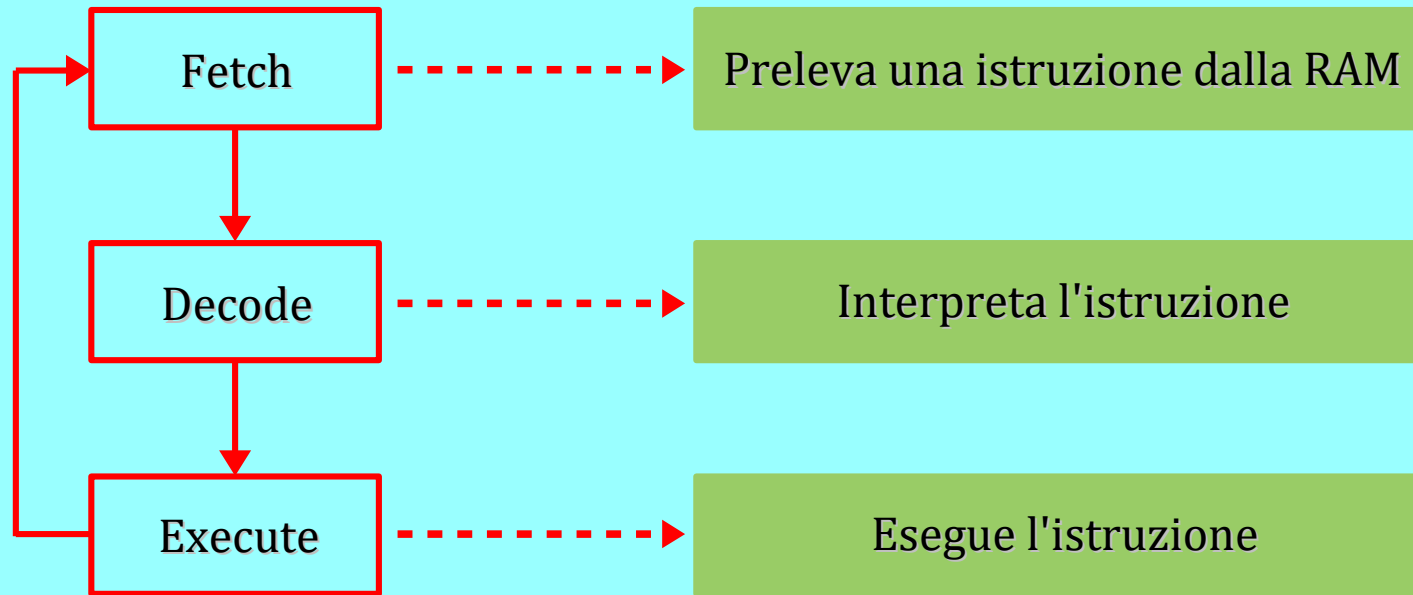
■ Clock

Il clock è un segnale in ingresso alla CPU che determina la temporizzazione della CPU. Si tratta di un segnale a onda quadra caratterizzata da una particolare frequenza generato da un opportuno circuito di temporizzazione.



■ Ciclo macchina

L'esecuzione di una istruzione è sempre suddivisa in tre fasi:



Le tre fasi si ripetono per ciascuna istruzione fino alla fine del programma. In generale il PC viene modificato prima della fase di execute nel seguente modo:

$$\text{PC} = \text{PC} + \text{lunghezza dell'istruzione corrente}$$

In questo modo il PC viene aggiornato in modo tale che contenga già l'indirizzo di memoria dell'istruzione successiva. Per esempio, se il valore attuale del PC è 0 e l'istruzione corrente occupa 2 byte, il PC viene aggiornato con il valore 2 (0+2).

Ogni ciclo macchina è scandito da un temporizzatore o clock, ossia un oscillatore al quarzo che emette segnali ad intervalli di tempo regolari.

■ Ciclo macchina (descrizione dettagliata)

1. **Preleva** l'istruzione in memoria di indirizzo uguale al **PC** e mettilo nel registro **IR**.

2. **Incrementa** il contenuto del registro **Program Counter** in modo da puntare all'istruzione successiva.

3. Determina il tipo di istruzione letta (**decodifica**).

4. Se l'istruzione richiede degli operandi, determina dove si trovano (memoria o registri).

5. Preleva, se necessario, dalla memoria gli operandi e ponili nei registri della **CPU**.

6. **Esegui** l'istruzione (**execute**).

7. **Salva** il risultato in un registro o in una cella di memoria.

8. **Torna** al punto 1.

Fetch istruzione

Decode istruzione

Fetch operandi

Execute

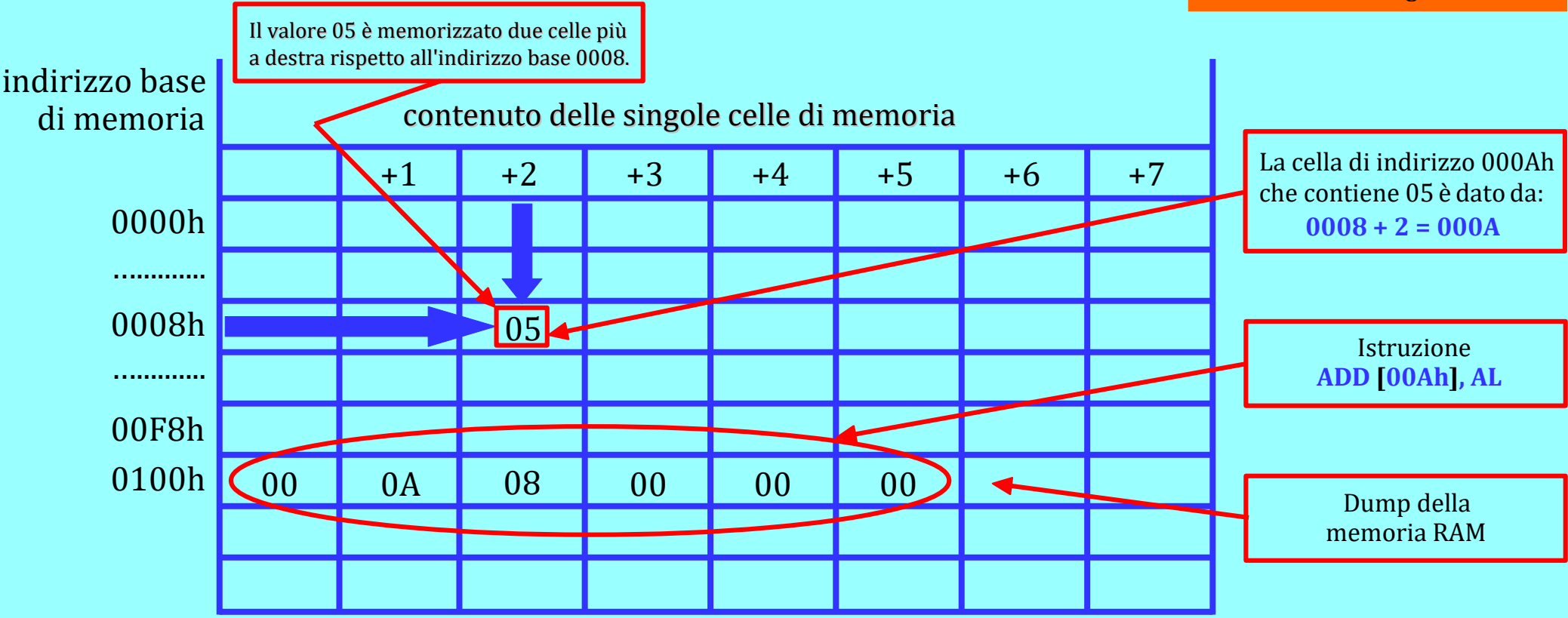
Memory dump

Identifica lo stato in cui si trovano le celle di memoria in un determinato momento. In alcuni casi viene anche mostrato il corrispettivo ASCII del contenuto della singola cella di memoria.

Esempio

Rappresentare in memoria l'istruzione assembly ADD [000Ah], AL.

Significato: somma il contenuto della cella di indirizzo 000Ah e il contenuto del registro AL.

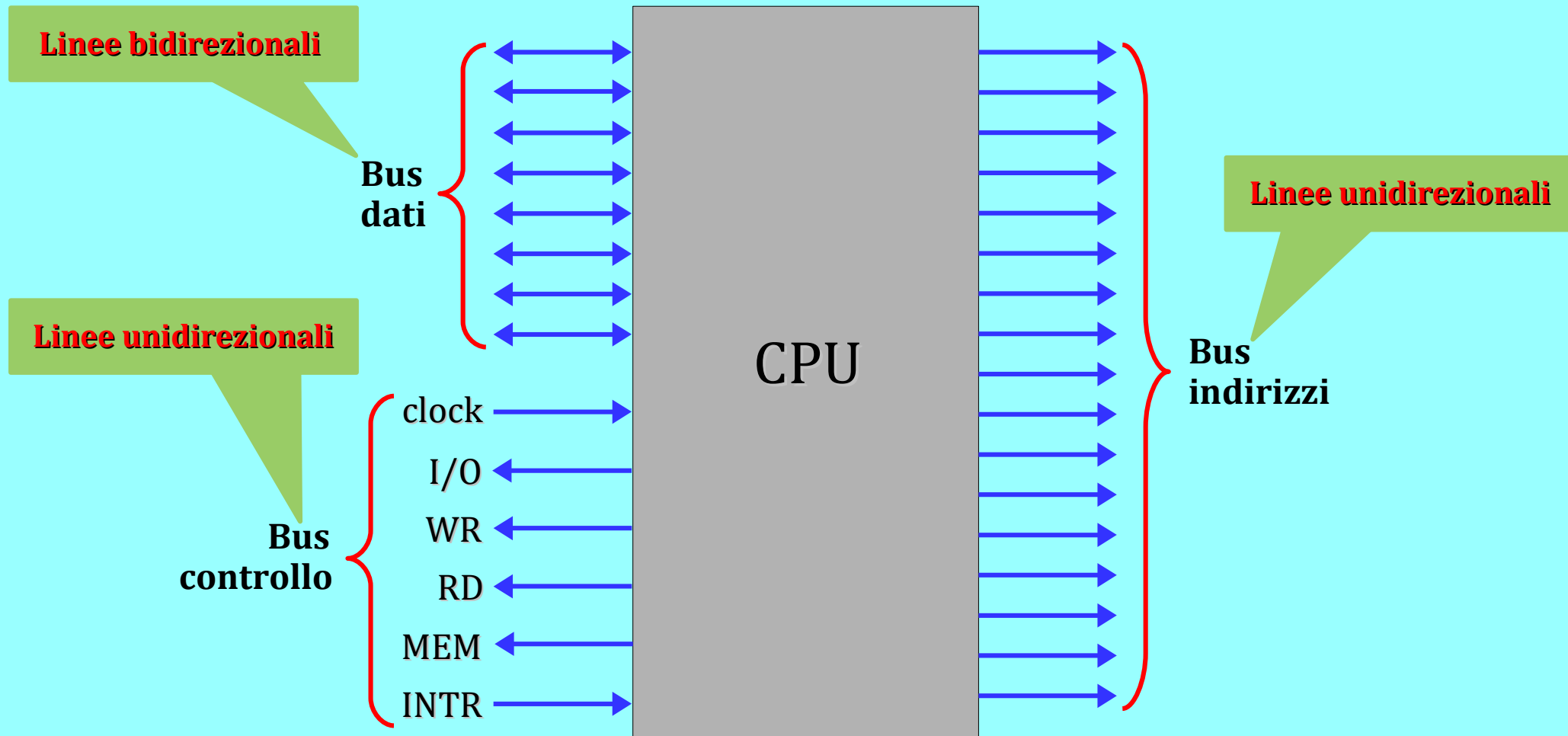


Operazioni svolte dal processore

1. **Fetch**: preleva l'istruzione formata da 6 byte (00, 0A, 08, 00, 00, 00), il cui indirizzo è contenuto nel registro PC (0100h). Il codice dell'istruzione viene inserito nel registro IR.
2. **Decode**: decodifica l'istruzione contenuta nel registro IR.
3. **Fetch degli operandi**: preleva il dato 05h contenuto nella cella 000Ah.
4. **Execute**: il valore 05h viene sommato con il contenuto del registro AL (si pone contenga il valore 0Fh). Il risultato 14h viene posto ancora nella cella di indirizzo 000Ah.

Nota. La lettera h posta dopo la cifra meno significativa rappresenta in assembly un numero espresso in esadecimale (base 16).

■ Architettura esterna di una CPU



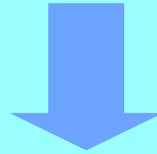
Principali segnali di controllo

- **Lettura (RD):** è attivo se la CPU esegua una operazione di lettura da memoria o da periferica; è in uscita dalla CPU.
- **Scrittura (WR):** è attivo se la CPU esegue una operazione di scrittura da memoria o da periferica; è in uscita dalla CPU.
- **Memoria (MEM):** è attivo se la CPU effettua un trasferimento dati che coinvolge una cella di memoria; è in uscita dalla CPU.
- **Input/Output (I/O):** è attivo se la CPU effettua un trasferimento dati che coinvolge una periferica; è in uscita dalla CPU.
- **Clock:** è il segnale di temporizzazione; è in ingresso alla CPU.
- **Interrupt Request (INTR):** è il segnale di richiesta di interruzione da parte di una periferica; è in ingresso alla CPU.

■ Prestazioni di un microprocessore

Per misurare le prestazioni di una CPU nel tempo ci si riferisce a diverse unità di misura:

- **MIPS** (Million Instruction Per Second): si tratta di una unità di misura media in quanto istruzioni diverse hanno tempi di esecuzione diversi;
- **FLOPS** (Floating Point Operations Per Second): unità di misura media analoga al MIPS, ma per le sole istruzioni floating point (si impiega il multiplo MegaFLOPS);
- **Dhrystone/Whetstone**: algoritmi standard di misura delle prestazioni rispettivamente per il calcolo numerico con valori interi e in virgola mobile (floating point).



La sola velocità del processore in uno specifico test non è particolarmente significativa. Da alcuni anni vengono utilizzati dei **benchmark**, cioè un insieme di test software atti a mettere sotto stress le componenti di un PC per valutarne le prestazioni.



Quindi il benchmark misura le prestazioni di un elaboratore e non del solo processore: meno tempo impiega a eseguire il benchmark, migliore è la prestazione.

■ Processori Multi-Core

Il termine multicore si usa per descrivere una CPU composta da più di due core.

IL CORE E' IL NUCLEO ELABORATIVO DI UN PROCESSORE.

Per esempio, CPU dual-core (doppio core) unisce due processori indipendenti in un unico package.

Vantaggi

- Aumenta le prestazioni e la potenza di calcolo.
- Minore surriscaldamento rispetto ad una CPU basata su singolo core.

E' tuttavia bene ricordare che le frequenze di clock dei diversi core della stessa CPU non possono essere sommate. Esempio: una CPU dual-core a 2 GHz non elabora a 4 GHz, ma semplicemente ogni core lavora in modo indipendente alla stessa frequenza di 2 GHz.

■ Case

E' un contenitore metallico nel quale sono installati i diversi componenti del computer, tra cui: alimentatore, dissipatore di calore (ventola) posto sopra la CPU e cavi che permettono il transito dei dati alla scheda madre verso i dispositivi periferici.

■ Alimentatore

E' in grado di fornire energia elettrica necessaria al funzionamento delle varie componenti del sistema. La corrente elettrica si muove su cavi che si servono di determinati connettori per alimentare la scheda madre:

- **ATX** (Advanced Technology Extended), connettore principale che può avere 20 o 24 pin;
- **AUX**, un connettore ausiliario dotato di 4 o 6 pin, da collegare alla scheda madre.

Vi sono poi i connettori per alimentare i cavi SATA che hanno 15 pin per collegare gli hard disk, unità ottiche o un qualsiasi dispositivo che abbia un connettore di questo tipo.

■ Scheda madre (motherboard)

E' l'elemento principale di un computer che consente di mettere in comunicazione tra loro i diversi componenti. In particolare contiene chipset, schede aggiuntive, bus di espansione e interfacce di collegamento con il mondo esterno.

■ CPU socket

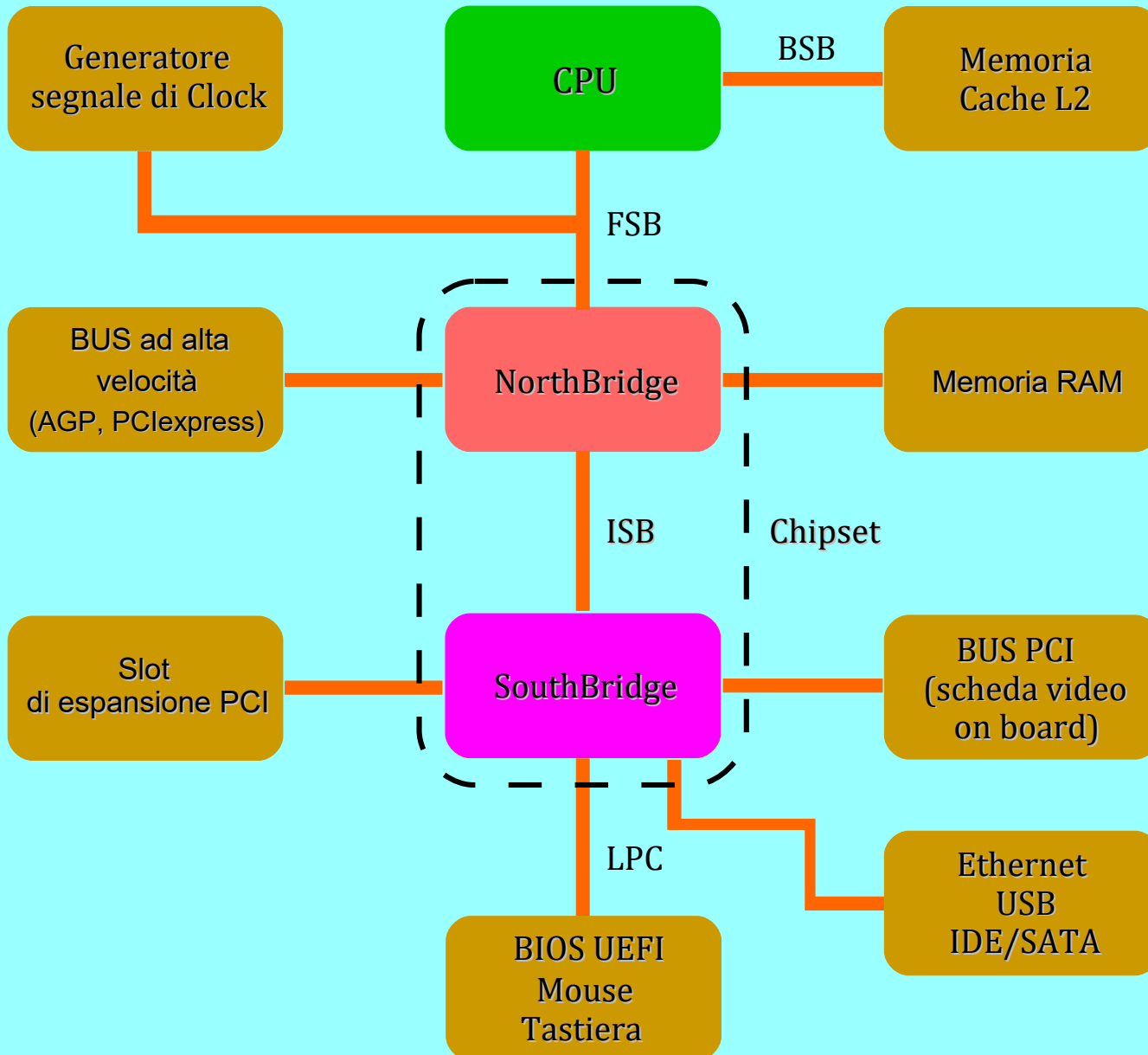
Apposito connettore che accoglie i piedini della CPU ed è di tipo **ZIF** (Zero Insertion Force) per il modo dolce con cui la CPU viene inserita.

ZIF (Zero Insertion Force)

Il termine ZIF indica un alloggiamento particolare nel quale non è necessario utilizzare la propria forza per collocare la CPU in sede. Infatti permette di facilitare il suo inserimento semplicemente sollevando una piccola leva.

■ Chipset

Circuito integrato che gestisce il flusso dei dati tra i principali componenti del computer: la CPU, la RAM, la cache di livello 2 e i dispositivi collegati al bus PCI e ai canali PATA e SATA.



- **NorthBridge**: gestisce il flusso dei dati verso i dispositivi più veloci.
- **SouthBridge**: gestisce il flusso dei dati verso i dispositivi più lenti.
- **ISB** (Internal Side BUS): connette i due chipset tra loro.
- **FSB** (Front Side BUS): consente il trasferimento dei dati tra la CPU e il chipset NorthBridge.
- **BSB** (Back Side BUS): consente il collegamento tra il processore e la memoria cache di livello 2.

■ Banda passante (bandwidth) del FSB

Il numero massimo di Byte al secondo (B/s) che si possono trasmettere attraverso il canale al NorthBridge. Tale valore si determina mediante la seguente formula:

$$\text{Bandwidth} = \text{Larghezza BUS} \times \text{Frequenza clock} \times \text{Numero di data transfer}$$

Esempio

Si determini la banda passante di un sistema con:

- processore a 16 bit (2 byte),
- FSB a 100 MHz,
- due trasferimenti a ciclo,

possiede una bandwidth di:

$$2 \text{ (byte)} \times 100 \text{ (FSB)} \times 2 \text{ (tc)} = 400 \text{ MB/s (Megabyte al secondo)}$$

Esempio

Si determini la banda passante di un sistema con:

- processore a 32 bit (4 byte),
- FSB a 100 MHz,
- quattro trasferimenti a ciclo,

possiede una bandwidth di:

$$4 \text{ (byte)} \times 100 \text{ (FSB)} \times 4 \text{ (tc)} = 1600 \text{ MB/s (Megabyte al secondo)}$$

LARGHEZZA BUS: NUMERO DI LINEE INDIPENDENTI PER LA TRASMISSIONE DI DATI.

Verifica delle abilità

1 Come avviene l'esecuzione di un programma?

Per ogni istruzione la CPU esegue ripetutamente un ciclo di fetch, decode, execute.

2 Dove devono essere i dati e le istruzioni che formano un programma che deve essere eseguito?

Nella memoria centrale.

3 Che cosa significa che una CPU lavora a 1 MHz?

Significa che la frequenza di clock è di 1 MHz.

4 La CPU da dove legge le istruzioni da eseguire?

Dalla memoria centrale.

5 Descrivi quali sono le funzioni di una CPU.

Leggere le istruzioni, decodificarle, eseguirle e scrivere il risultato in memoria.

6 Descrivi le principali componenti di una CPU.

Unità di controllo, unità aritmetico logica, registri (di uso speciale e di uso generale).

7 Che cosa sono i registri di uso speciale? Quali sono?

Sono registri che svolgono una particolare funzione. Ogni processore è dotato di propri registri, ma sono presenti in generale il program counter, lo stack pointer e il registro di stato.

Program Counter (PC): contiene l'indirizzo della prossima istruzione da eseguire.

Stack Pointer (SP): punta alla cima dello stack.

Il registro di stato è composto da una sequenza di bit, ciascuno dei quali rappresenta lo stato in cui si trova la CPU. Alcuni bit forniscono informazioni sul risultato di una operazione aritmetica logica (bit di stato), altri segnalano lo stato del processore (bit di controllo).

8 Da dove deriva il termine settare e cosa significa settare un bit?

Deriva dall'inglese e significa assegnare. Settare un bit significa porre il valore 1 al bit.

9 Cos'è e come viene usato il registro MAR?

Il MAR (Memory Address Register) contiene l'indirizzo della cella di memoria da cui leggere o in cui scrivere un dato.

10 Nell'ambito dell'architettura degli elaboratori, cos'è l'indirizzo e qual è la sua funzione?

Un indirizzo è un numero che indica la posizione di una cella di memoria.

11 Spiegare la differenza tra un'operazione di lettura da memoria e un'operazione di scrittura in memoria facendo riferimento alle componenti del sistema coinvolte.

Una operazione di lettura è un trasferimento da memoria o periferica alla CPU, una operazione di scrittura è un trasferimento da CPU a memoria o periferica.

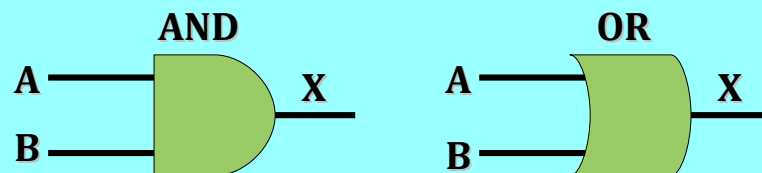
12 Esporre la differenza tra la funzione svolta dal bus esterno e quella svolta dal bus interno.

Il bus interno trasporta i dati all'interno del processore, mentre il bus esterno trasporta i dati tra il processore e la memoria (o i dispositivi di I/O).

13 Perché si dice che l'esecuzione di un programma è un ciclo di esecuzione?

Perché la CPU effettua le operazioni di fetch, decode, execute ciclicamente per ogni istruzione.

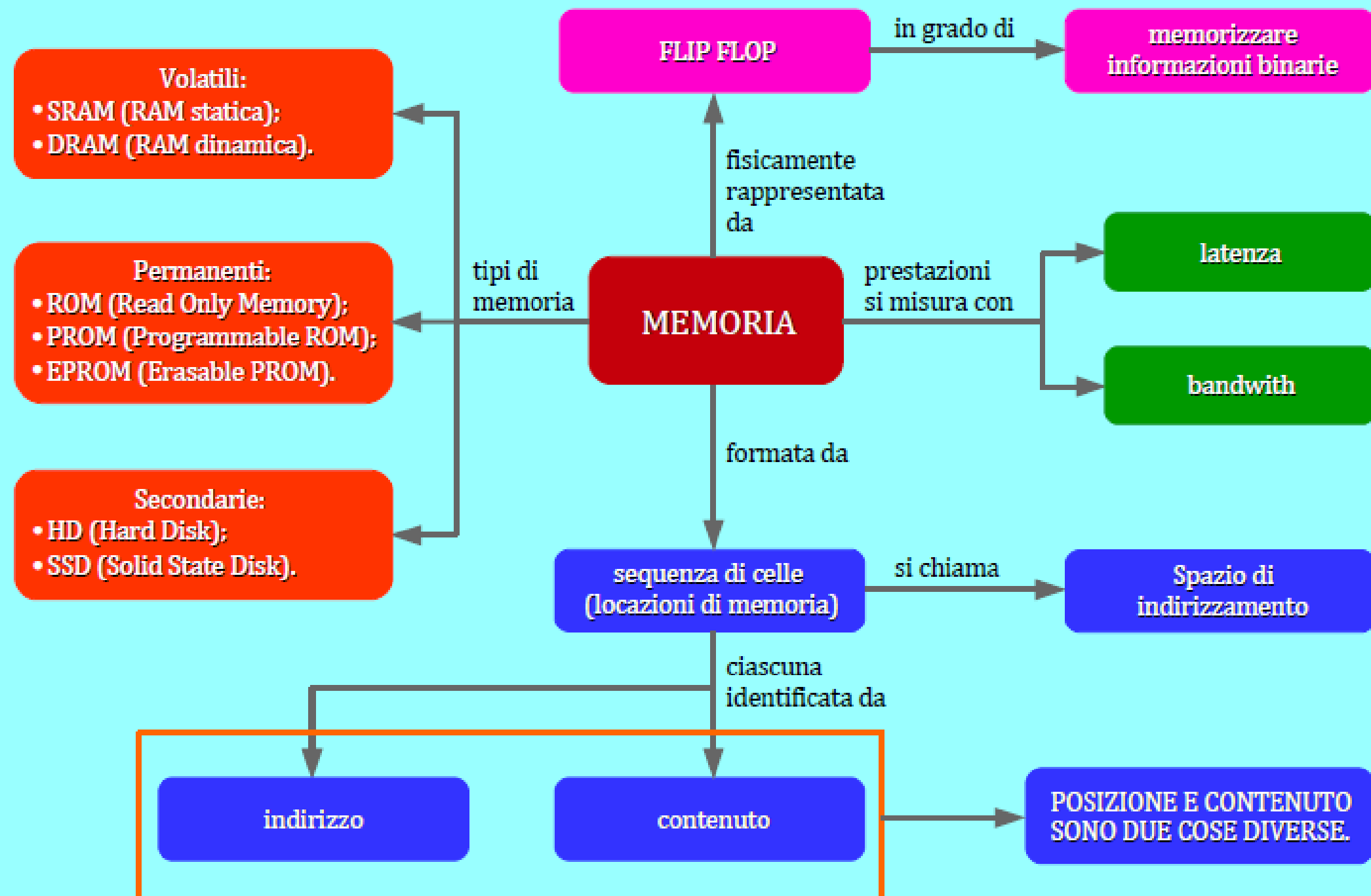
14 Rappresenta graficamente le porte logiche per le funzioni AND e OR.



Unità 3

Le memorie

Mappa concettuale



■ Bit

E' un simbolo che codifica lo stato fisico di un segnale e può assumere solo due valori, che corrispondono i valori logici VERO o FALSO (o, indifferentemente e più comunemente 1 e 0), rappresentati di solito da due diversi valori di tensione (ad esempio 0 V, +5V).

■ Byte

E' l'unità di misura della capacità della memoria e corrisponde a 8 bit. Per convenzione si usa indicare il byte con B per distinguerlo dal bit, che è indicato con b.

■ Capacità di memoria

Indica la quantità di byte che la memoria è in grado di immagazzinare e si misura dunque in byte o in multipli del byte (è sempre una potenza di 2).

■ RAM (Random Access Memory)

E' la parte del calcolatore che svolge la funzione di memorizzare temporaneamente i dati, le istruzioni, i risultati parziali e quelli finali prima di essere mandati in output. Il termine random viene tradotto impropriamente con casuale, ma non nel senso che l'informazione è conservata dove capita, ma perchè il tempo di accesso non dipende dalla sua posizione.

■ Locazione di memoria

Una parte di memoria che può essere referenziata (ossia identificata). In genere, la prima locazione di memoria ha indirizzo 0. Per fare riferimento ad una locazione di memoria, la CPU deve utilizzare il suo indirizzo.

■ Indirizzo di una locazione di memoria

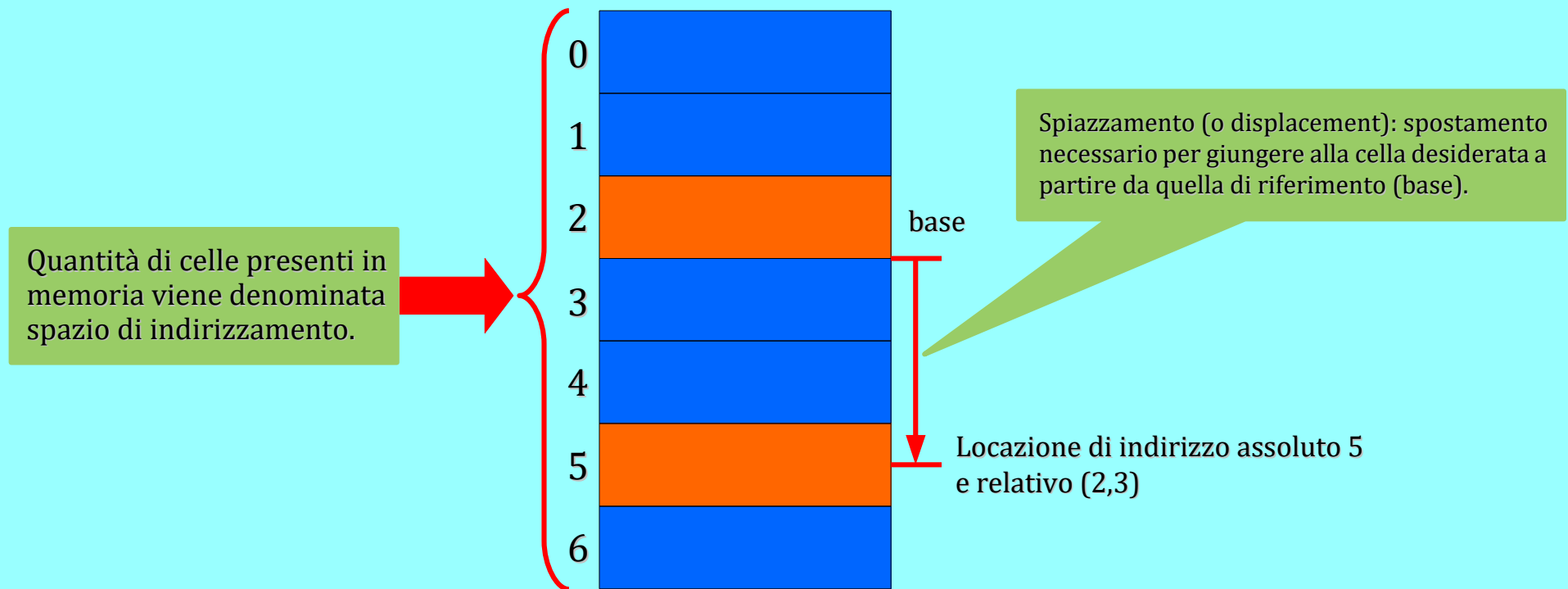
Indica la posizione di una locazione di memoria, ovvero il modo in cui la cella di memoria viene referenziata (o identificata).

■ Indirizzo assoluto (o fisico)

Indica la posizione di una locazione di memoria a partire dalla cima della memoria.

■ Indirizzo relativo (o logico)

E' la posizione di una locazione di memoria relativamente ad un'altra. E' costituita da due numeri: base e spiazzamento. Lo spiazzamento è la distanza rispetto alla base.



Ciascuna locazione possiede un solo indirizzo assoluto, ma più indirizzi relativi: a un indirizzo assoluto possono corrispondere più indirizzi relativi.

■ Spazio di indirizzamento

Per scrivere o leggere dalla memoria, la CPU deve inviare, tramite il bus indirizzi, l'indirizzo della cella di memoria su cui scrivere o da cui leggere, deve trasferire o ricevere sul bus dati il dato che vuole scrivere o prelevare e deve abilitare, tramite il bus controllo, il comando di scrittura (WR) o di lettura (RD). Il numero di linee che formano il bus indirizzi determina la dimensione massima della memoria.

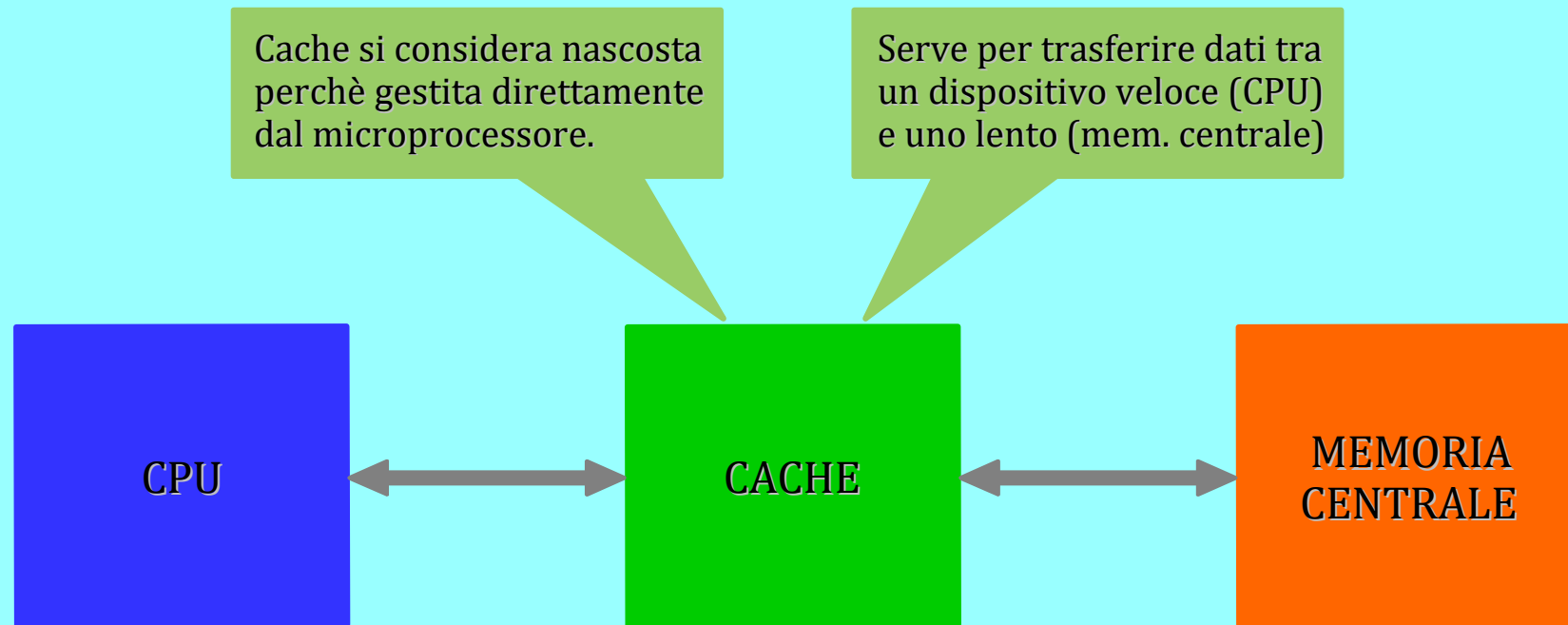
Esempio

- **8 linee indirizzi:** $2^8 = 256$ locazioni di memoria. Gli indirizzi vanno da 0 a 255, cioè $2^8 - 1$ (in binario, da 00000000 a 11111111; in esadecimale, da 0h a FFh).
- **16 linee indirizzi:** $2^{16} = 65536$ locazioni di memoria, cioè 64 KB.
- **20 linee indirizzi:** $2^{20} = 1048576$ locazioni di memoria, cioè 1 MB.
- **32 linee indirizzi:** $2^{32} = 4294967296$ locazioni di memoria, cioè 4 GB.

Dall'esempio si nota che un indirizzo di memoria è espresso in binario o in esadecimale, che è la forma più compatta. Ciascuna locazione di memoria contiene dati e/o istruzioni (o parte di istruzioni), anch'esse espresse in binario.

■ La cache memory

La cache è una memoria estremamente veloce e di piccole dimensioni, la cui funzione è di memorizzare temporaneamente i dati e i programmi che si prevede debbano essere usati nell'immediato futuro. Spesso la cache è integrata nel chip contenente il microprocessore, altre volte è contenuta in chip separati.



Attualmente la cache è suddivisa in tre livelli: una cache integrata direttamente nel processore, detta **cache di primo livello** (L1); una cache esterna ma collegata direttamente al processore, detta **cache di secondo livello** (L2); una cache esterna sulla motherboard, chiamata **cache di terzo livello** (L3).

La memoria cache è realizzata con la tecnologia **SRAM** (Static Random Access Memory), mentre la RAM è progettata con la tecnologia **DRAM** (Dynamic Random Access Memory). Le due tecnologie sono fatte in modo diverso e con costi diversi (la SRAM è più veloce ma molto più costosa della DRAM).

■ Operazione di lettura (cache memory)

Quando la CPU richiede un dato, lo cerca prima nella cache: se è presente lo trasferisce nei registri di CPU e lo utilizza, altrimenti preleva dalla memoria centrale un blocco di dati che contiene il dato cercato e lo porta dalla cache nei registri per utilizzarlo.

■ Operazione di scrittura (cache memory)

Quando la CPU deve effettuare una scrittura, la cache può operare in due modalità:

- **modalità write-through**: il dato viene scritto sia nella cache sia nella RAM. Questo metodo è semplice da realizzare e non richiede che il blocco presente in cache venga riscritto nella memoria centrale quando viene sostituito;
- **modalità write-back**: il dato viene scritto soltanto nel blocco in cache; il blocco modificato viene scritto nella memoria centrale soltanto quando deve essere sostituito. A ogni blocco è associato un bit (**dirty bit**) che indica se il blocco nella cache è stato modificato o meno e quindi va copiato in memoria in caso di sostituzione.

■ Cache hit e cache miss

Quando il dato viene trovato nella cache si parla di **cache hit**, per denotare il successo nella lettura, altrimenti si parla di **cache miss**, cioè il dato non è presente nella cache. Il rapporto tra cache hit e accessi totali alla memoria, chiamato **hit rate** (rapporto di successo), misura l'efficacia della memoria cache.

■ LRU (Least Recently Used)

Si tratta di un criterio per eliminare dalla cache i dati che non sono usati da più tempo ed è prevalentemente applicato nel metodo write-back quando lo spazio nella memoria cache è insufficiente per caricare nuovi dati.

■ Tipologie di memorie

Le memorie si possono classificare in **volatili** e **permanenti**.

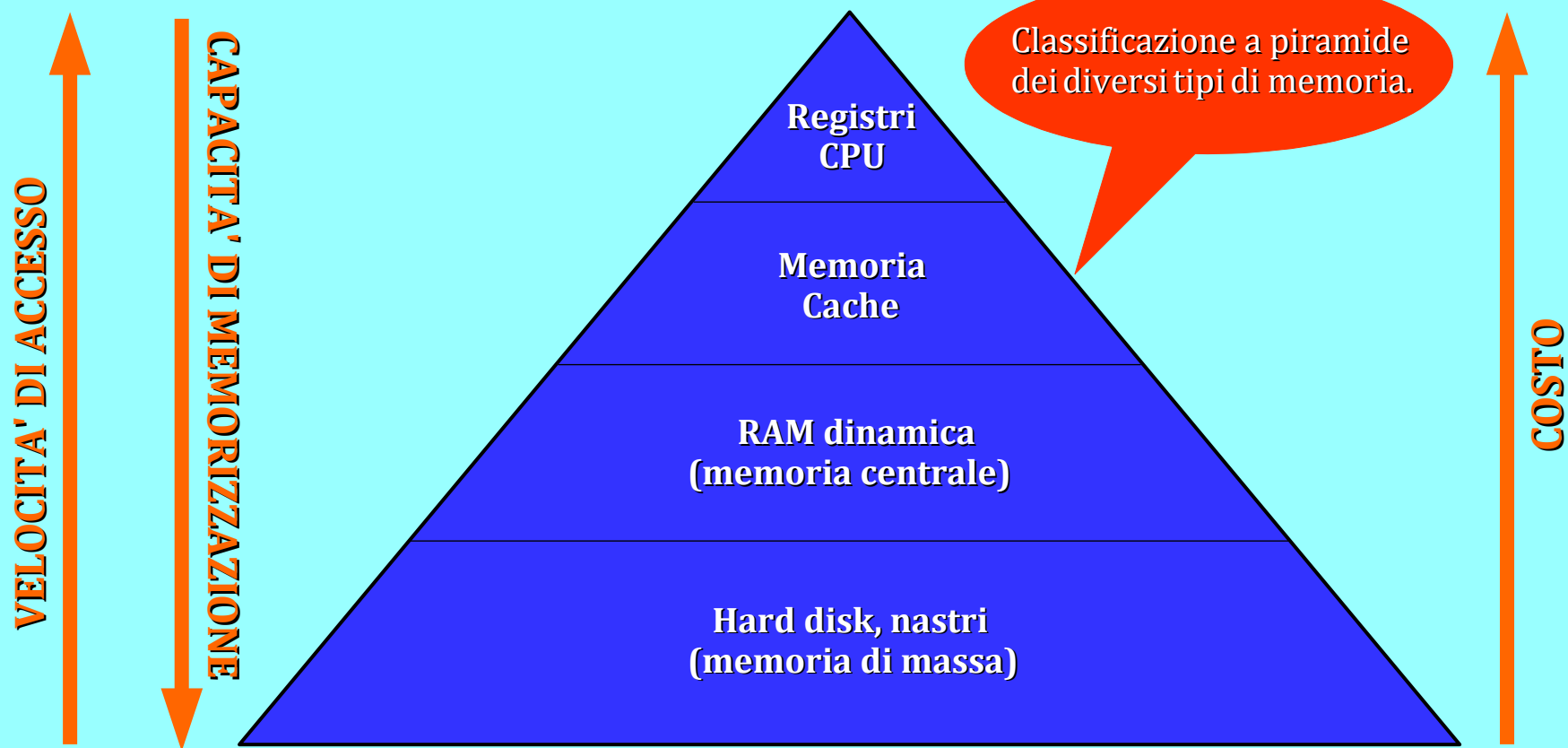
- Memoria **RAM**: memoria volatile (ovvero non mantiene le informazioni quando manca l'alimentazione del computer); si suddivide in due sottocategorie:
 - **SRAM** (Static RAM): ogni cella è costantemente alimentata per mantenere il dato memorizzato; i tempi di risposta alla CPU sono però più brevi.
 - **DRAM** (Dynamic RAM): ogni cella è formata da un condensatore che mantiene la carica elettrica per un tempo limitato; un circuito presente in memoria provvede costantemente a ricaricare le celle prima che perdano il dato memorizzato.
- Memoria **ROM**: memoria permanente (ovvero mantiene le informazioni anche quando manca l'alimentazione del computer); esistono differenti tipi di ROM:
 - **PROM** (Programmable ROM): possono essere programmate dall'utente solo una volta per mezzo di uno speciale dispositivo.
 - **EPROM** (Erasable PROM): possono essere anche cancellate mediante dispositivi un certo numero limitato di volte e riprogrammate.

Le prestazioni della memoria

1. **Latenza**: tempo richiesto per trasferire un dato da/verso la memoria;
2. **Velocità di trasferimento** (bandwidth): numero di byte trasferiti in un secondo.
3. **Frequenza di funzionamento**: blocchi di dati trasferiti al secondo.

■ Gerarchia di memoria

La memoria all'interno della scheda madre di un PC è organizzata in livelli caratterizzati da velocità, capacità di memorizzazione e costi diversi. In ogni istante le informazioni vengono copiate soltanto tra due livelli adiacenti: il livello superiore (quello più vicino al processore) è più piccolo e più veloce. L'obiettivo è quello di tenere i blocchi di informazioni più usati di frequente nei livelli di memoria più vicini al processore, allo scopo di evitare il più possibile accessi a tipi di memorie molto capienti ma lente.



Andando dall'alto al basso, aumenta la capacità di memoria, ma di conseguenza diminuisce la velocità di accesso e i costi per byte.

■ Principio di località

Definisce la ragione per cui i dati nella memoria cache sono usati con maggior frequenza di quelli della memoria centrale. Si distinguono due tipi diversi di località:

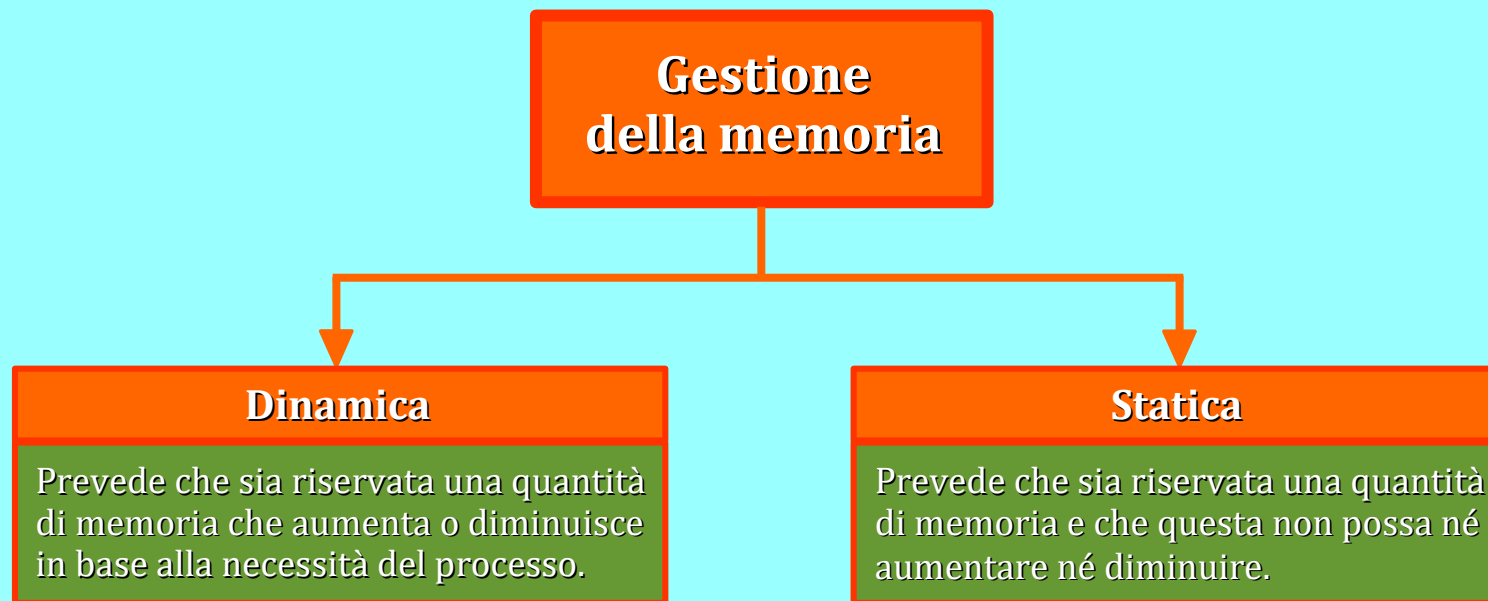
- **Località temporale:** se un dato è richiesto in un certo istante, è probabile che lo stesso dato venga nuovamente richiesto entro breve, come per esempio nelle iterazioni in cui le stesse istruzioni vengono ripetute frequentemente;
- **Località spaziale:** se un dato è richiesto in un certo istante, è probabile che dati situati nelle celle di memoria vicine vengano richieste entro breve, come per esempio accesso a vettori oppure matrici attraverso un indice.

Tabella. Caratteristiche della gerarchia delle memorie.

Tipo di memoria	Capacità	Tempo di accesso
Registri di memoria	< 1 KB	1 – 3 ns
Memoria cache	512 KB – 4 MB	3 – 10 ns
Memoria centrale	1 – 4 GB	50 – 200 ns
Disco magnetico	50 GB - 1TB	20 – 30 ms
Nastro	4 GB – 300 GB	> 1 ns
Dischi ottici	650 MB – 4,7 GB	> 1 ns

■ Stack

E' una parte della memoria centrale che può memorizzare solo dati in situazioni particolari. Caratteristica dello stack è la **gestione dinamica con politica di tipo LIFO (Last In First Out)** e significa che l'ultimo elemento che entra è il primo a uscire, come in una pila di libri.



Gli elementi dello stack sono situati in posizioni adiacenti della memoria centrale ed è prevista solo la presenza di uno specifico registro di CPU, lo **stack pointer** (SP) che contiene l'indirizzo della cima dello stack. E' bene prestare attenzione a non confondere **stack** con **stack pointer**: lo stack è un'area di memoria, lo stack pointer è un registro speciale della CPU.

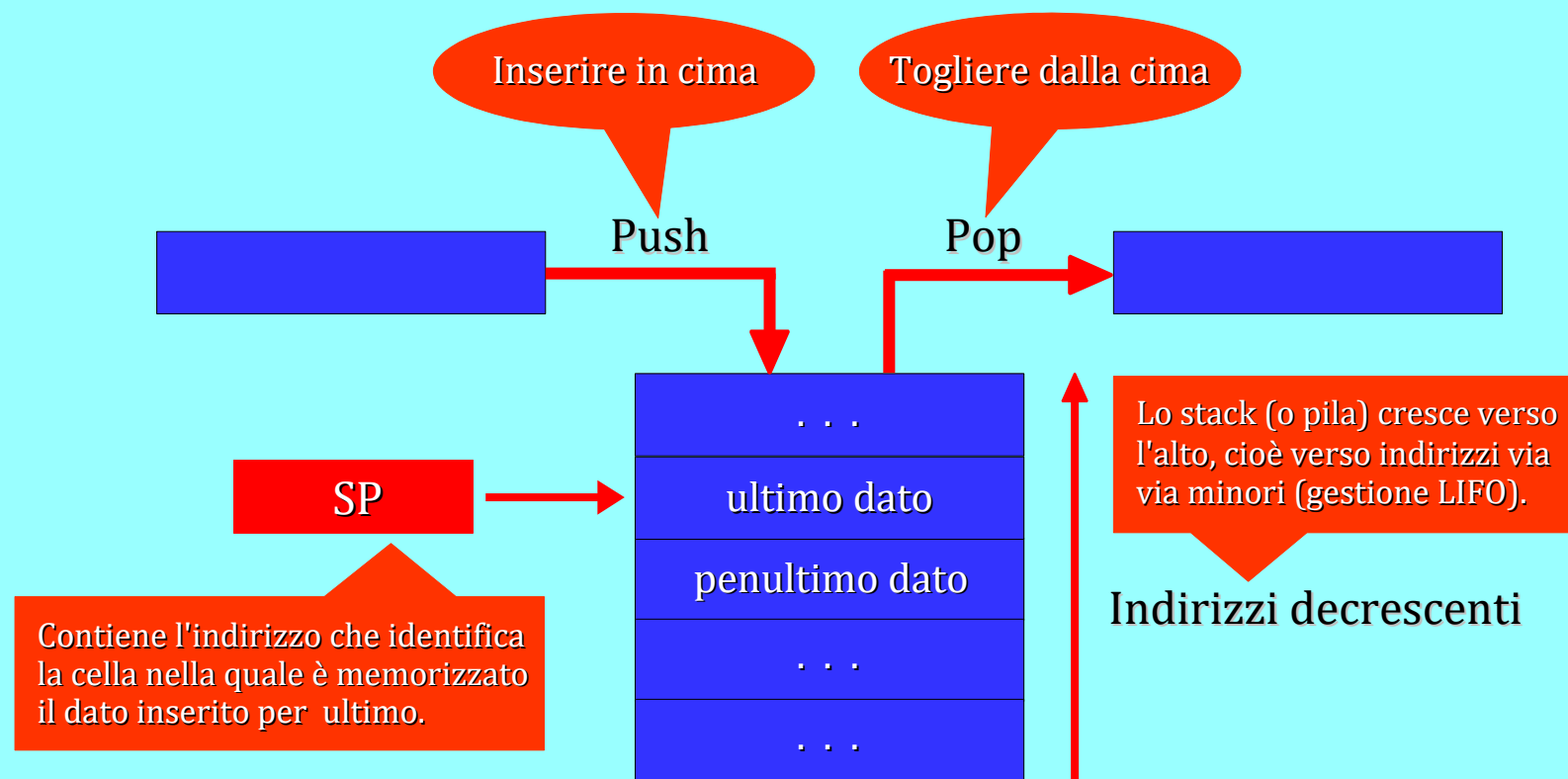
Analogia

Lo stack è una pila che ha una cima e un fondo. Gli elementi tra la parte superiore e quella inferiore mantengono un ordine, come la pila di libri sulla scrivania: è possibile inserire un nuovo libro sulla cima della pila o prelevare solo il libro inserito per ultimo.

■ Operazioni sullo stack

Per inserire o leggere dei dati in uno stack esistono due istruzioni particolari:

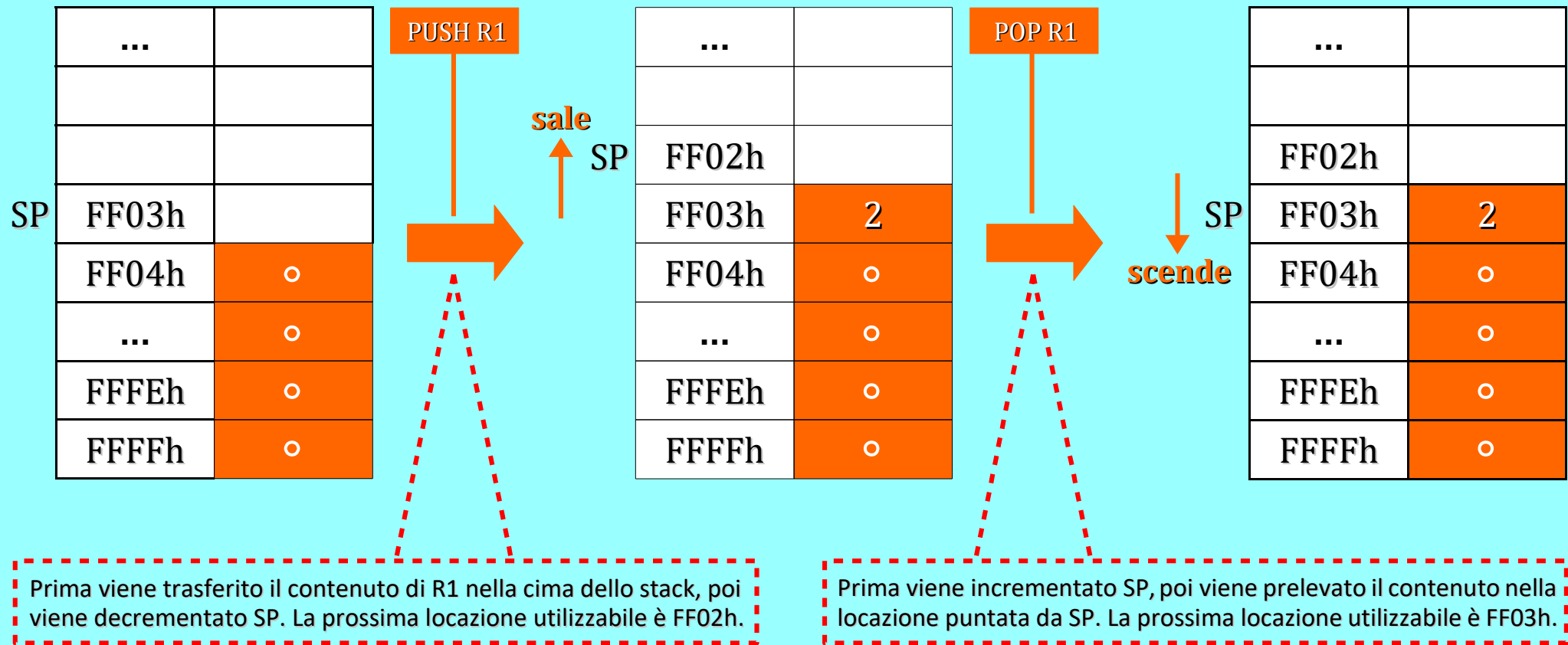
- **PUSH** scrive un dato nello stack all'indirizzo indicato da SP come prima locazione vuota; il contenuto di SP viene decrementato;
- **POP** incrementa il contenuto di SP e preleva l'elemento situato nella locazione di memoria all'indirizzo puntato da SP.



Lo stack viene usato nei linguaggi di alto livello per memorizzare **variabili locali** di una funzione. In questo caso in presenza di chiamate di funzioni nidificate, le ultime variabili allocate sono le prime a essere deallocate. Lo stack è anche utilizzato per salvare lo **stato di un processore** al verificarsi di un **interrupt** (segnale inviato dalla periferica alla CPU).

Esempio: operazioni sullo stack

Consideriamo la situazione di partenza mostrata in figura, in cui la prima cella utilizzabile è all'indirizzo FF03h (cima dello stack). Supponiamo di avere a disposizione della CPU, R1, che contiene il dato 2: $R1 = 2$.



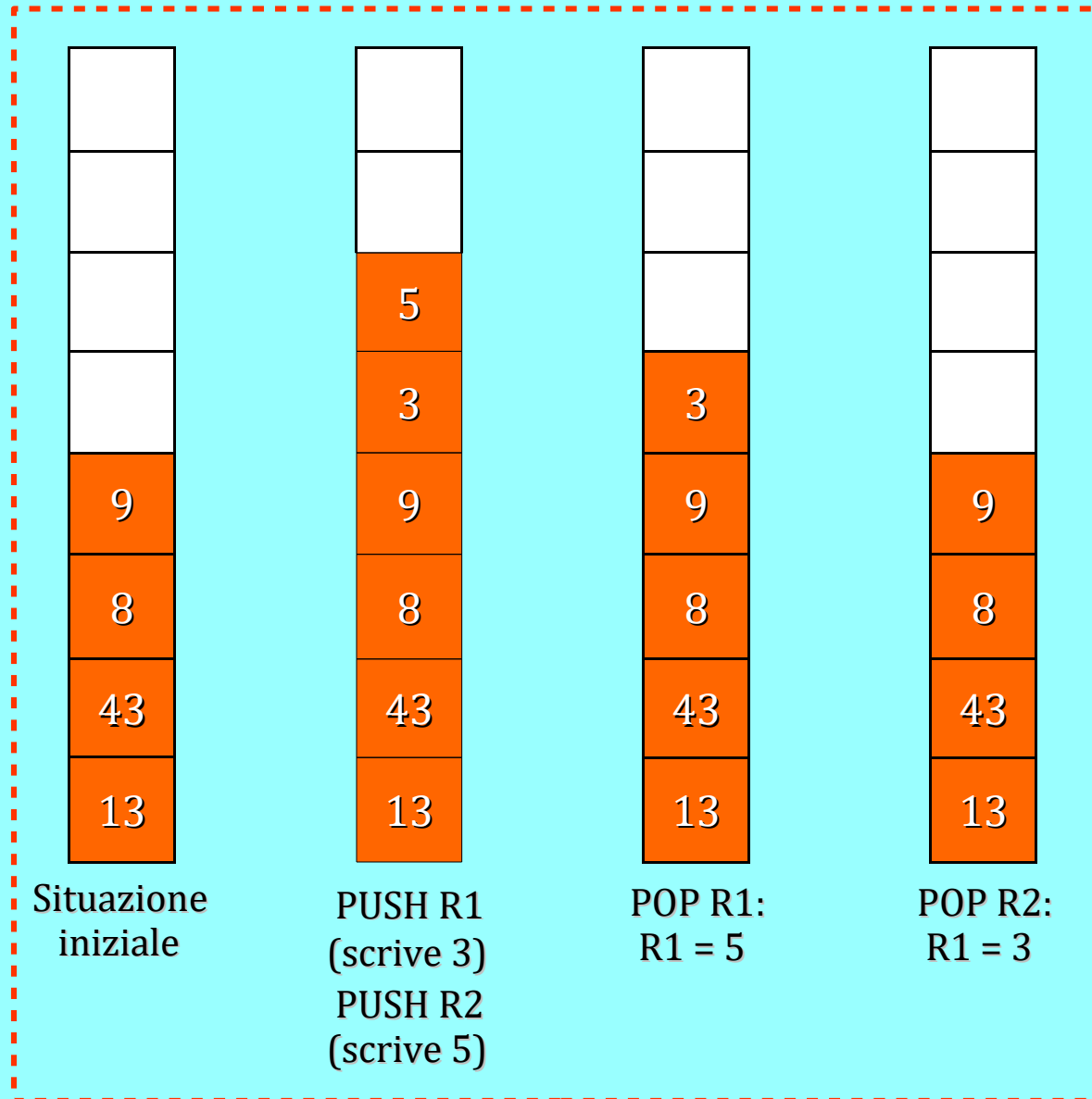
Notando le figure si vede che il contenuto della memoria non è stato modificato; infatti l'istruzione POP preleva il dato, ma non lo cancella dallo stack. Quello che cambia è il valore di SP. Dopo la POP $SP = FF03h$, e ciò significa che la prossima PUSH userà quella cella, eventualmente sovrascrivendo il dato. E' lo stack pointer che indica quale locazione è possibile utilizzare.

Esempio: utilizzo dello stack

Siano $R1 = 3$ e $R2 = 5$. Cosa contengono $R1$ e $R2$ dopo l'esecuzione delle seguenti istruzioni: PUSH $R1$, PUSH $R2$, POP $R1$, POP $R2$.

La situazione è rappresentata nella figura seguente:

Si badi bene che un uso errato di push e pop può causare notevoli danni!



DOPO
ESECUZIONE
ISTRUZIONI:
SCAMBIO DEI
VALORI DI $R1$
E $R2$.

1 Che cosa è l'indirizzo di una locazione di memoria?

Identifica la posizione di una locazione di memoria.

2 Che cosa è la memoria centrale?

E' la memoria volatile da cui la CPU legge dati e/o istruzioni e su cui la CPU scrive i risultati.

3 Quale funzione ha il Program Counter?

Contiene l'indirizzo della prossima istruzione da eseguire.

4 Che cos'è un bit?

E' una codifica dello stato fisico di un segnale digitale.

5 Che cos'è un byte?

L'unità di misura della capacità di memoria e corrisponde a 8 bit.

6 Come si misura la capacità di memoria?

In byte.

7 A che cosa corrisponde 1 kbyte?

A 1024 byte.

8 Perché i multipli dei byte usano la notazione k (kappa), M (mega), G (giga), ... , pur non essendo potenze di 10?

Perché la capacità di memoria si misura in potenze di 2. Per esempio, 1 kbyte è pari a 1024 byte perché è la potenza di 2 più vicina a 1000.

9 Qual è la differenza tra indirizzo di memoria e locazione di memoria?

L'indirizzo di memoria è la posizione di una locazione di memoria. La locazione di memoria è una parte della memoria centrale.

10 Che cos'è un puntatore?

E' un indirizzo di memoria in quanto punta alla locazione corrispondente.

11 Che cos'è un registro puntatore?

E' registro che contiene un indirizzo di memoria.

12 Che cosa significa "politica di tipo LIFO"?

LIFO è l'acronimo di Last In First Out e significa che l'ultimo elemento che entra è il primo ad uscire, come una pila di libri.

13 Che differenza c'è tra stack e stack pointer?

Lo stack è una parte della memoria centrale, lo stack pointer è un registro di uso speciale della CPU che punta alla cima dello stack.

14 Quale funzione ha lo stack pointer?

Punta alla base dello stack.

15 Quali operazioni possono essere fatti nello stack?

La PUSH per scrivere un dato nello stack all'indirizzo indicato da SP e la POP per leggere il dato contenuto nella locazione di memoria puntata da SP.

16 Qual è il ruolo del bus indirizzi nell'esecuzione di un'istruzione?

Sul bus indirizzi viene posto l'indirizzo che attiva la locazione di memoria nella quale prelevare l'istruzione da eseguire. L'indirizzo è quello presente nel Program Counter.

17 Qual è la funzione della cache memory?

Serve per velocizzare le operazioni di accesso ai dati da parte del processore.

18 Come avviene la scrittura nella cache memory?

La scrittura nella cache memory avviene in due modalità:

- write-through: scrittura sia nella cache sia nella RAM;
- write-back: scrittura nella cache con dirty bit.

18 Esporre la differenza tra processore e memoria in termini di funzioni svolte.

La memoria contiene dati e istruzioni, la CPU elabora i dati tramite le istruzioni.

19 Da dove deriva il termine "referenziare" e cosa significa referenziare una locazione di memoria?

Deriva dall'inglese e significa fare riferimento. Referenziare una locazione di memoria significa identificarla, cioè fare riferimento a quella locazione di memoria tramite il suo indirizzo.

20 In che senso la memoria di un sistema di elaborazione può essere paragonata ad una cassettera?

Perché funzionalmente la cassettera è un contenitore e ogni cassetto è analogo ad una cella di memoria identificata tramite la sua posizione.

21 Quando viene modificato il contenuto del registro Program Counter e perché?

Prima della fase di execute poiché deve puntare all'istruzione successiva a quella in corso.

22 Esporre il significato di "gerarchia di memoria".

Per gerarchia di memoria si intende una classificazione delle memorie dalla meno capiente (accesso più veloce) alla più capiente (accesso meno veloce).

23 Se $SP = \text{FFFBh}$, quanto vale SP dopo l'esecuzione di PUSH R1 , PUSH R2 , POP R2 , POP R3 , PUSH R1 , PUSH R1 , POP R3 ? ($R1, R2, R3$ sono dei generici registri).

Si suppone che l'operando della PUSH e della POP sia a 1 byte.

$\text{PUSH R1: } SP = SP - 1 = \text{FFFBh} - 0001\text{h} = \text{FFFAh}$

$\text{PUSH R2: } SP = SP - 1 = \text{FFFAh} - 0001\text{h} = \text{FFF9h}$

$\text{POP R2: } SP = SP + 1 = \text{FFF9h} + 0001\text{h} = \text{FFFAh}$

$\text{POP R3: } SP = SP + 1 = \text{FFFAh} + 0001\text{h} = \text{FFFBh}$

$\text{PUSH R1: } SP = SP - 1 = \text{FFFBh} - 0001\text{h} = \text{FFFAh}$

$\text{PUSH R1: } SP = SP - 1 = \text{FFFAh} - 0001\text{h} = \text{FFF9h}$

$\text{POP R3: } SP = SP + 1 = \text{FFF9h} + 0001\text{h} = \text{FFFAh}$

24 Siano istr1 (lunga 3 byte) e istr2 (lunga 3 byte) due istruzioni (in questo caso non serve conoscerne il significato).

F000h	istr1
	istr2

Determina:

- l'indirizzo della locazione di memoria contenente l'istruzione istr2 ;**
- il contenuto di PC prima della fase di fetch di istr1 ;**
- il contenuto di PC dopo la fase di fetch di istr1 .**

a. $PC = \text{F000h} + 3 \text{ (byte)} = \text{F0003h}$

b. $PC = \text{F000h}$

c. $PC = \text{F003h}$ (dopo la fetch, prima di execute)

Unità 4

Le periferiche di Input/Output

Mappa concettuale

Sono dispositivi che mettono in comunicazione il computer con il mondo esterno.

TECNICHE

tramite

- POLLING;
- INTERRUPT;
- DMA.

possono essere gestite con

PERIFERICHE

dotate di

interfacce programmabili

classificate in

input

output

gestiscono

porte seriali

porte parallele

Dati inviati molto più velocemente rispetto alla seriale.

contengono registri

dati

di controllo

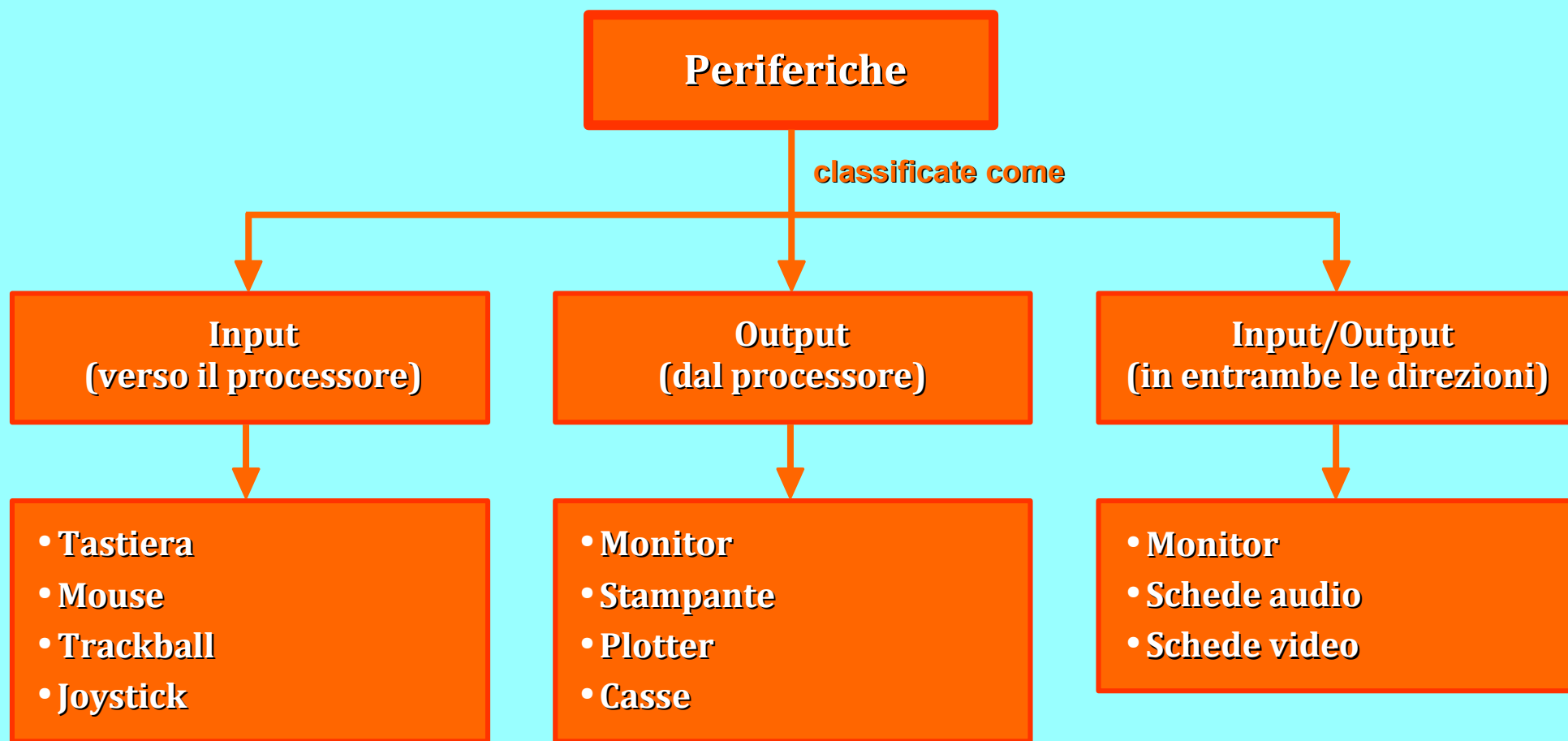
di stato

I bit transitano uno dopo l'altro.

I bit transitano su più linee parallele.

■ Periferiche di Input/Output

Servono per comunicare dall'esterno con il sistema e viceversa. Sono collegate al sistema attraverso particolari circuiti, detti **interfaccia**. Ogni periferica necessita di un dispositivo hardware (**controller**) e di software specifici (**driver**).

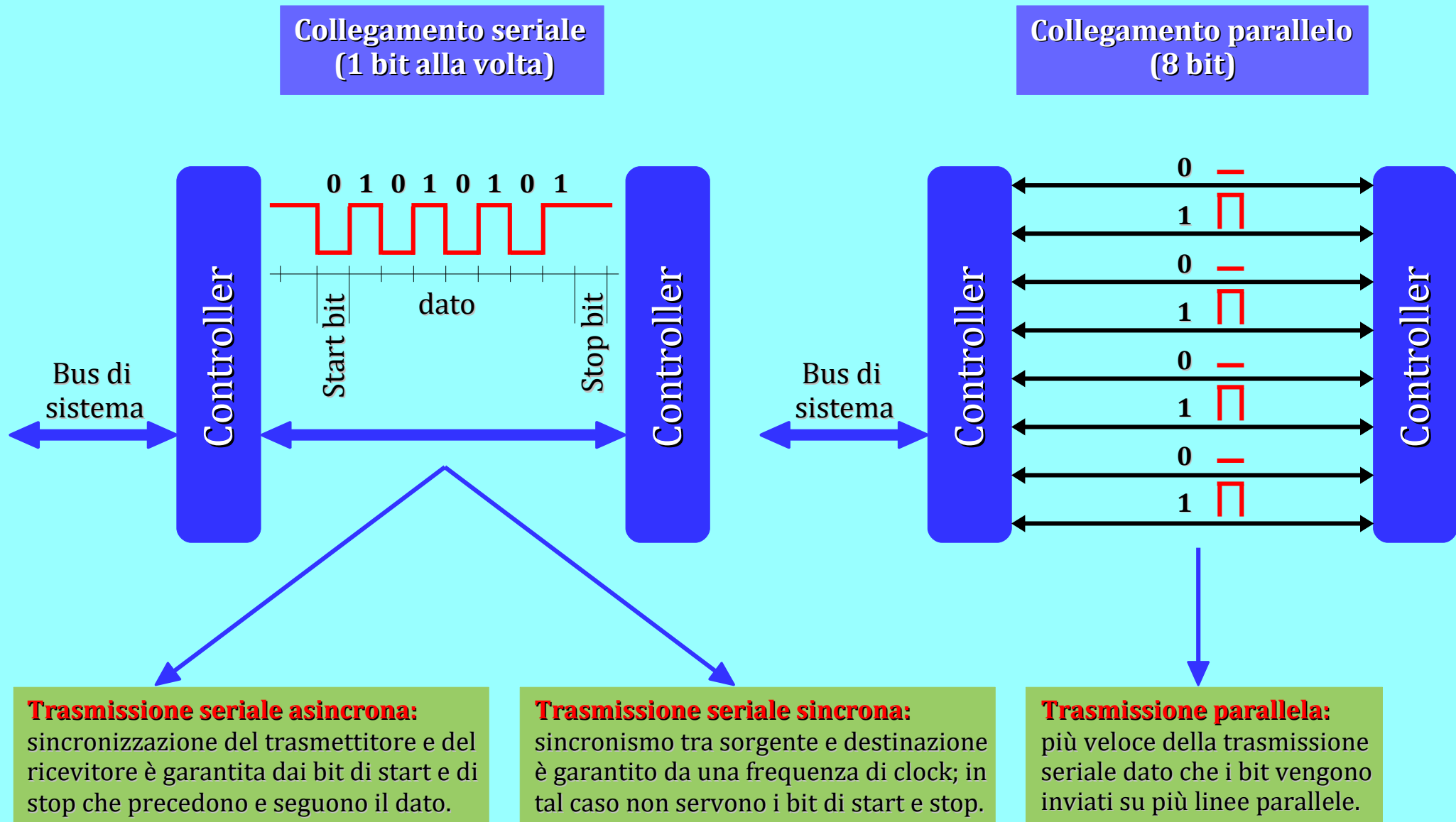


Nota bene

Con Input/Output (Ingresso/Uscita, In/Out) sono denotate le interfacce che permettono lo scambio di informazioni tra il PC e il mondo esterno. Alle interfacce sono collegate le periferiche di I/O quali tastiere, mouse, monitor, touchscreen, modem ecc.

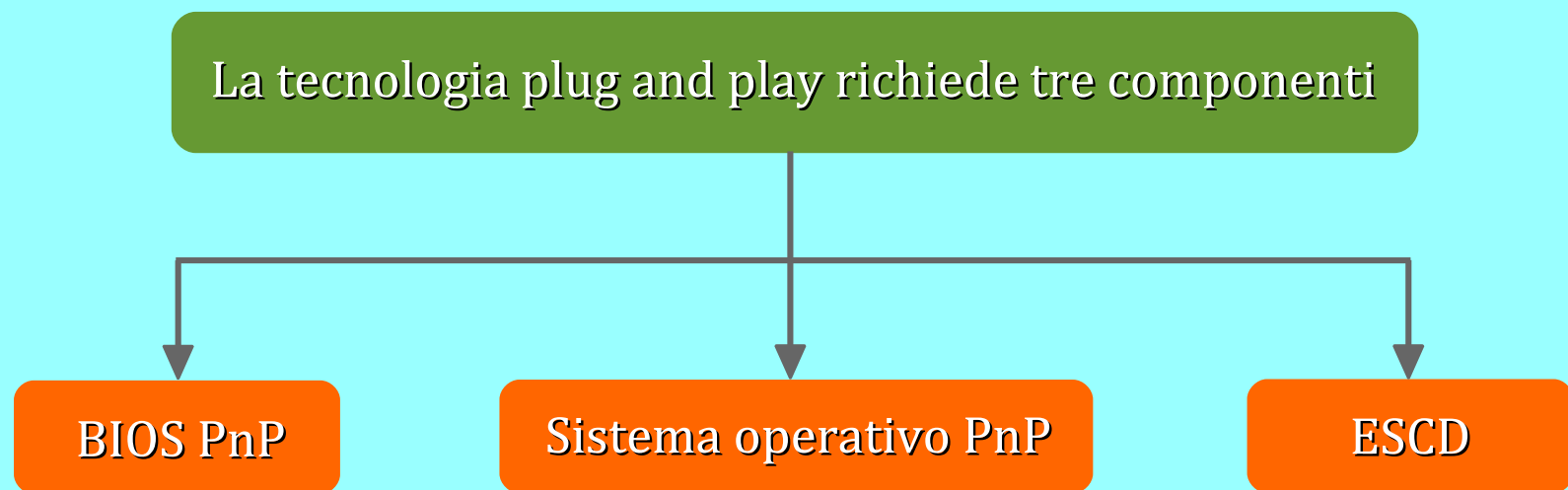
■ Porte seriali e parallele

La scheda di interfaccia è collegata al bus tramite porte che possono essere **seriali**, quando i dati vengono trasmessi un bit dopo l'altro, oppure **paralleli**, quando i bit dei dati vengono trasmessi contemporaneamente.



■ Periferiche plug and play

Termine che tradotto dall'inglese vuol dire “**connetti e utilizza**” e denota le periferiche che possono essere connesse a caldo (**hot swap**): ciò significa che in qualsiasi istante il sistema le riconoscerà e le configurerà automaticamente.



BIOS PnP: si tratta di un software implementato in ROM del sistema che abilita e determina quando una periferica si è connessa.

ESCD (Extended System Configuration Data): si tratta di un file che contiene tutte le informazioni necessarie per l'installazione delle periferiche plug and play.

Sistema operativo PnP: è il sistema operativo (da Windows 95 in poi) capace di supportare la gestione dei driver per le periferiche plug and play.

Driver: insieme delle procedure, spesso scritte in Assembly, che consente al sistema operativo di pilotare un dispositivo hardware.

■ USB (Universal Serial BUS)

E' una **porta seriale** progettata per consentire a più periferiche di essere connesse usando una sola interfaccia e un solo tipo di connettore e possono collegare fino a 127 periferiche senza dover spegnere il computer. Lo standard USB consente di **collegare** e **configurare** le periferiche plug and play al computer.

La distanza massima a cui può trovarsi una periferica è di cinque metri.

Le periferiche che comunicano con un computer tramite USB sono molteplici. Ad esempio: mouse, stampante, hard disk esterno e tastiera.

La porta USB viene usata anche come connessione di ricarica elettrica per dispositivi alimentati a batteria.

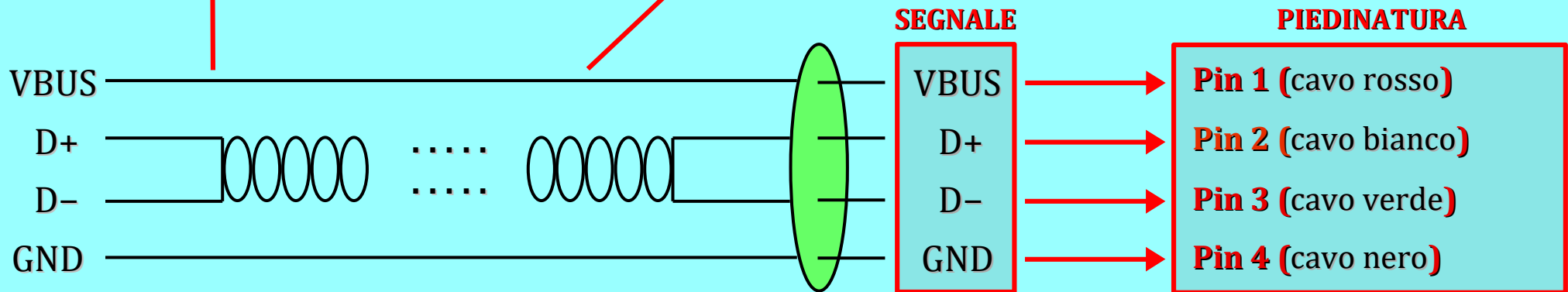
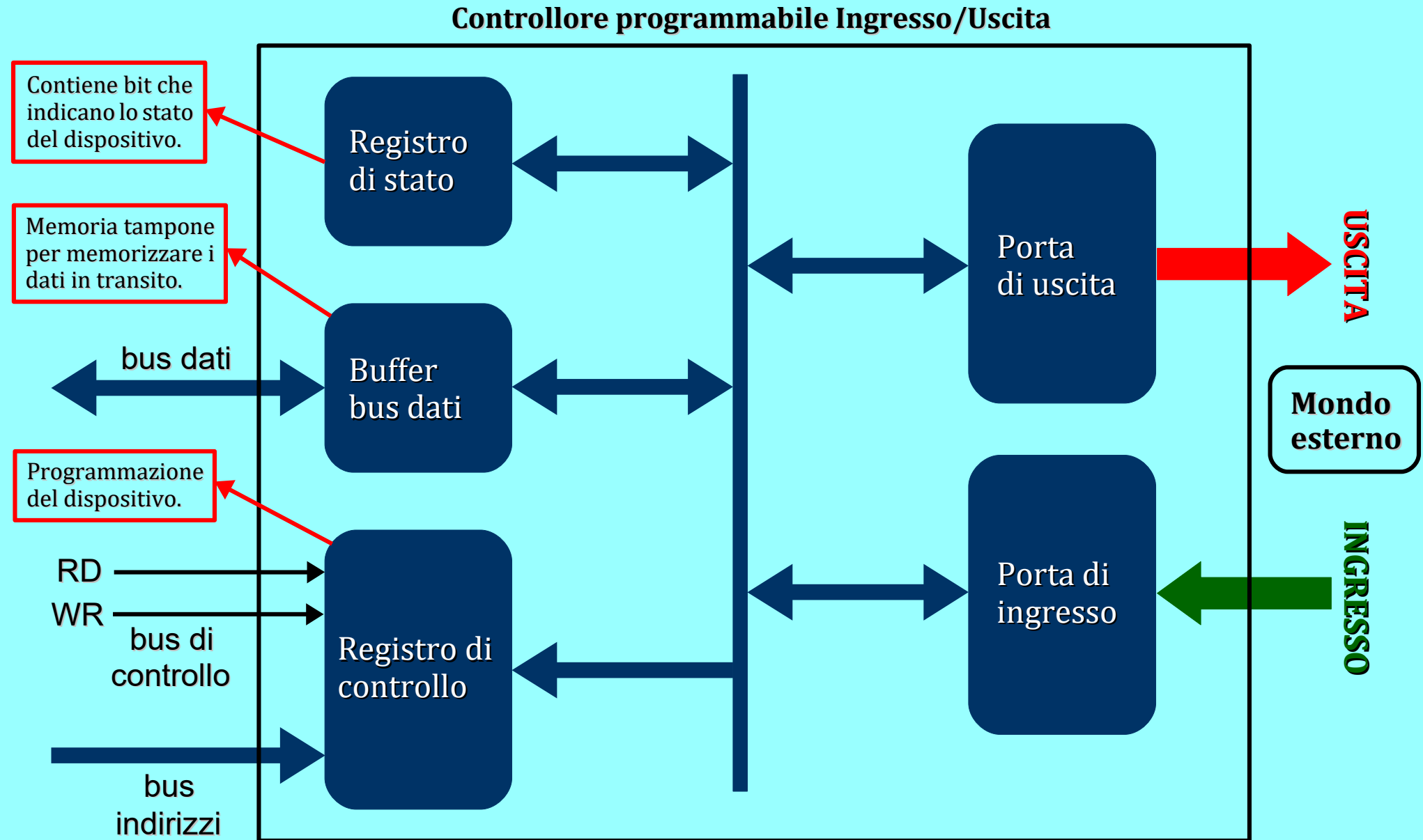


Tabella. Velocità di trasmissione dello standard USB.

Standard USB	Anno di uscita	Velocità
USB 1.0	1996	1,5Mbps
USB 2.0	2000	480 Mbps
USB 3.1	2017	20 Gbps

■ Struttura di un I/O

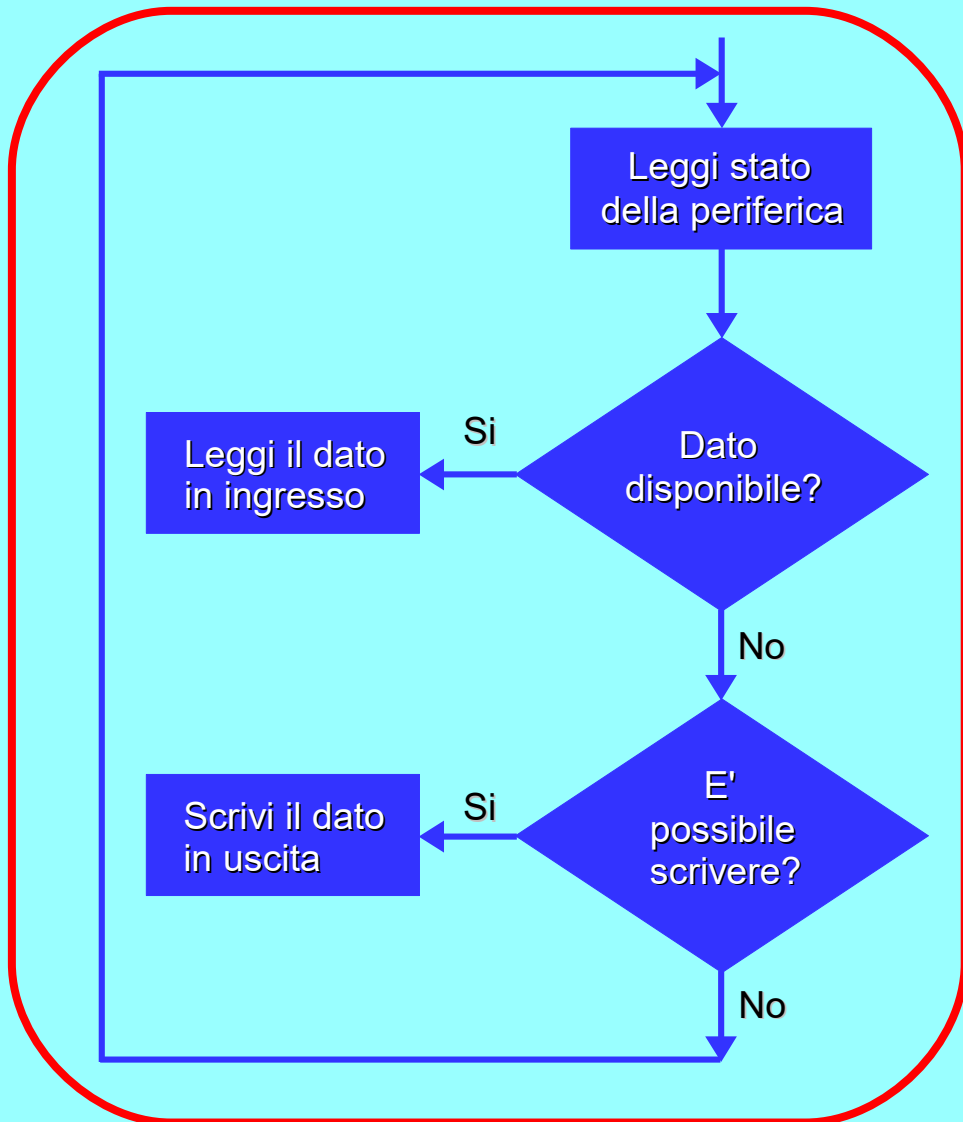
Un sistema di I/O è costituito da **dispositivi programmabili** che gestiscono le porte seriali (SIO – Serial Input Output) e parallele (PPI – Parallel Port Interface). Il dispositivo contiene un **buffer di dati**, un **registro di controllo** e un **registro di stato**. Questi registri sono gestiti dalla CPU mediante particolari istruzioni.



■ Tecniche per la gestione delle periferiche

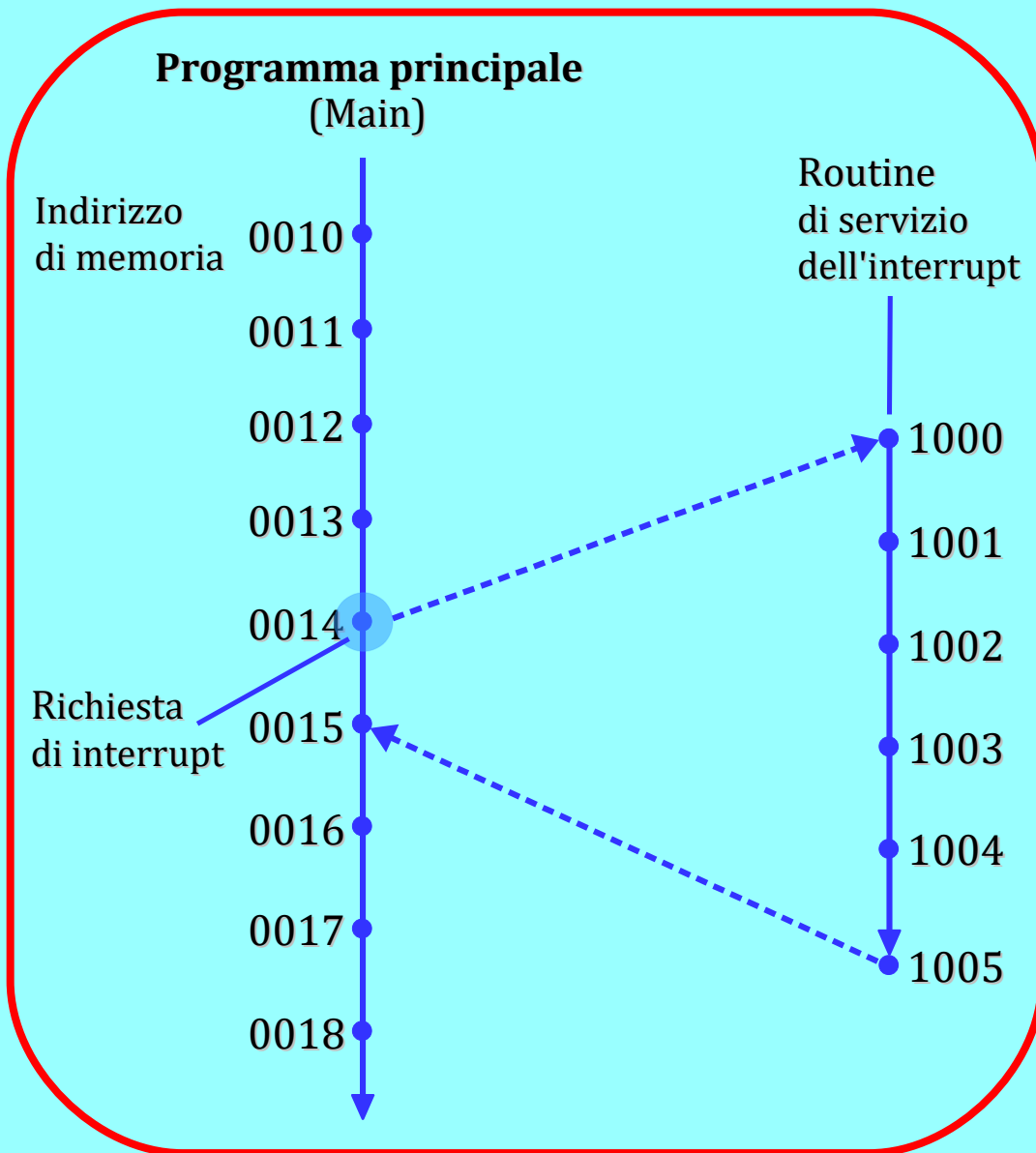
Per accedere ai registri della periferica e di conseguenza eseguire un trasferimento di dati, la CPU può utilizzare tre tecniche di gestione: **polling**, **interrupt**, **DMA**.

■ La tecnica del polling (o controllo da programma)



- Gestione prettamente software in polling.
- La CPU verifica ciclicamente lo stato della periferica tramite il registro di stato.
- Qualora la periferica è pronta a inviare o ricevere un dato, viene eseguito il codice relativo al trasferimento del dato.
- Incoveniente: spreco di tempo della CPU per interrogazione ciclica.

■ La tecnica dell'interrupt (interruzione)

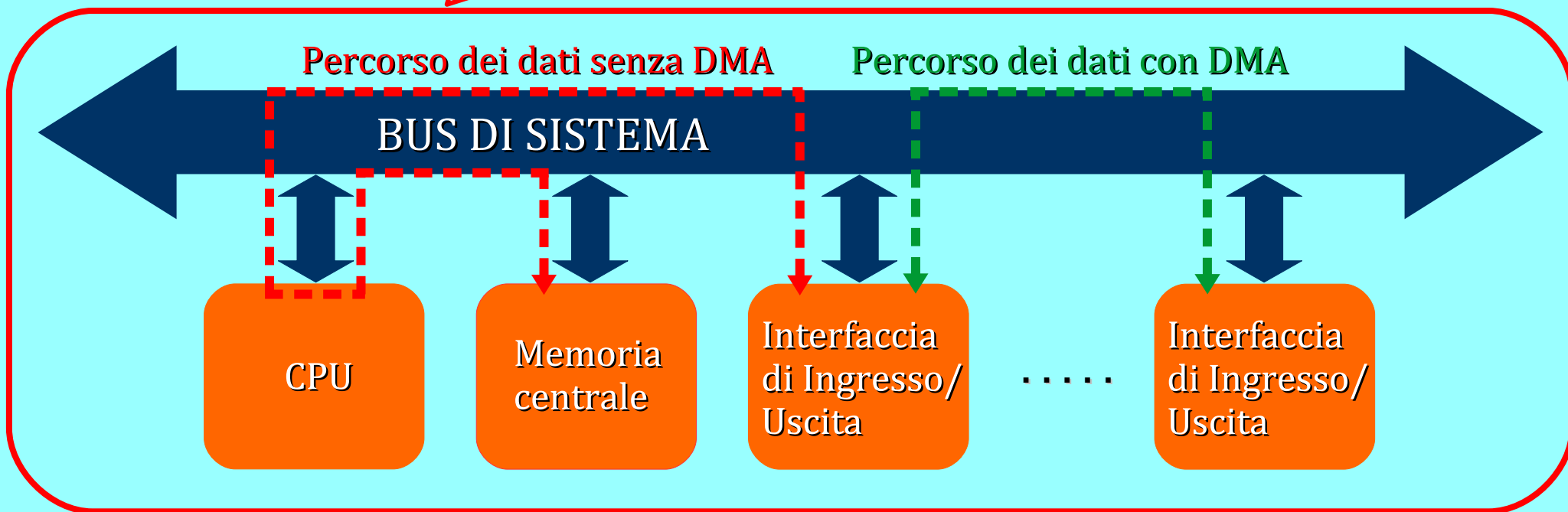


- Gestione con interruzione (interrupt).
- La periferica richiede alla CPU di essere servita inviando un segnale di interrupt.
- La CPU interrompe quindi il programma in corso di esecuzione e “salta” a eseguire un programma (routine) diverso.
- Alla fine della routine la CPU riprende dal punto successivo all'interruzione.
- Vantaggio: la CPU non spreca tempo per interrogare ciclicamente le periferiche.

Schema di funzionamento della **gestione di periferica attraverso interrupt**. La periferica chiama la CPU, che passa all'esecuzione della routine di servizio.

■ La tecnica del DMA (Direct Memory Access)

La CPU rimane bloccata nello stato di attesa (WAIT) fino a quando la periferica non completa il trasferimento dei dati. La tecnica del DMA è comunemente utilizzata per il trasferimento di dati ad alta velocità tra interfaccia di I/O e memoria centrale.



Nel caso di **accesso diretto alla memoria (DMA)**, i dati transitano direttamente tra interfaccia e memoria. In questo caso la CPU “si stacca” dal bus rimanendo in attesa fino alla fine dell'operazione di trasferimento di dati (**percorso tracciato in verde**). In caso contrario, senza DMA, i dati transitano dall'interfaccia di I/O alla memoria centrale sotto il controllo della CPU (**percorso tracciato in rosso**).

1 Che cosa si intende per Input?

Ciò che dall'esterno entra nel sistema.

2 Che cosa si intende per Output?

Ciò che viene elaborato dal sistema e che viene fornito in output.

3 Che differenza c'è tra un'interfaccia di I/O e una periferica di I/O?

La periferica è un dispositivo che serve per inserire dati e programmi (input) o fornire all'utente il risultato dell'elaborazione (output). L'interfaccia è un circuito elettrico che collega il sistema di elaborazione e la periferica. L'interfaccia fa parte del sistema, la periferica no.

4 Fai almeno tre esempi di periferica di Input.

Tastiera, mouse, microfono.

5 Fai almeno tre esempi di periferica di Output.

Monitor, stampante, cassa audio.

6 Descrivi l'architettura interna di una generica interfaccia di I/O.

Un interfaccia di I/O contiene un registro dati, un registro di controllo e un registro di stato. Tali registri sono gestiti dalla CPU tramite opportune istruzioni.

7 Quali sono i motivi per cui è necessario collegare le periferiche tramite interfacce?

E' necessario perchè la velocità delle periferiche di I/O e quella del processore sono diverse, per le caratteristiche elettriche e meccaniche differenti (le periferiche utilizzano elementi meccanici che ne limitano fisicamente il funzionamento) e dal tipo di informazioni elaborate: digitali per il computer, analogiche per il mondo esterno.

8 Che differenza c'è tra trasmissione seriale e trasmissione parallela?

- Trasmissione parallela: i bit vengono trasmessi in parallelo uno per ogni filo.
- Trasmissione seriale: i bit vengono trasmessi in sequenza, uno dopo l'altro.

9 Esponi le tecniche studiate per la gestione delle periferiche.

- Interrupt: la periferica invia un segnale al processore quando deve essere servita. Il processore interrompe il processo in esecuzione a favore di una routine di servizio associata alla periferica.
- Polling: il processore interroga ciclicamente le periferiche di I/O per controllare se necessitano di essere servite.
- DMA: accesso diretto alla memoria da parte delle interfacce di I/O, senza passare dalla CPU.

10 Elenca vantaggi e svantaggi della gestione delle periferiche tramite interrupt.

- Vantaggi: la CPU lavora in modo più efficiente in quanto non perde tempo ad interrogare le periferiche, ma continua ad eseguire il programma in corso finché viene interrotta qualora una periferica necessiti di essere servita.
- Svantaggi: è più complessa da realizzare.

11 Classifica le seguenti periferiche:

tastiera – mouse – scanner – microfono – trackball – touchpad – webcam – monitor – altoparlanti – stampante – plotter – cuffie – modem.

- Input: tastiera, mouse, scanner, microfono, trackball, touchpad, webcam.
- Output: monitor, altoparlanti, stampante, plotter, cuffie.
- Input/Output: modem.

12 Che cosa sono e che funzioni hanno le “porte”?

Le porte sono parti del sistema usati per interconnettere le periferiche di I/O.

13 Perchè si dice che le interfacce di I/O sono dispositivi programmabili?

Perchè sono dotate di un registro di controllo a cui inviare le istruzioni da eseguire.

14 Quali sono i registri interni a un dispositivo di I/O programmabile?

Registro dati, registro di controllo, registro di stato.

15 Le periferiche sono interne o esterne al sistema?

Possono essere interne o esterne.

16 Cosa permette di fare lo standard plug and play?

Permette ai dispositivi di autoconfigurarsi.

17 Cosa sono i driver software?

Sono programmi utili al funzionamento delle periferiche. Ogni periferica ha il suo driver.

18 Cosa significa DMA?

Direct Access Memory.

19 E' più efficiente una gestione di periferiche in polling o a interrupt? Motivare la risposta.

E' più efficiente una gestione a interrupt in quanto la CPU può lavorare fino a quando non viene interrotta.

20 Completa la tabella.

Caratteristiche	Porta seriale	Porta parallele
Trasmissione bit	Uno dopo l'altro	In parallelo

20 Classifica le periferiche in base alle loro caratteristiche.

Classificazione	Esempio di periferica
Input	Mouse
Output	Monitor
Input/Output	Disco

21 Riassumi, usando la tabella, le caratteristiche delle tecniche di gestione delle periferiche.

	Polling	Interrupt	DMA
Descrizione	La CPU interroga ciclicamente le periferiche per verificare se necessitano di essere servite.	La periferica invia un segnale alla CPU quando necessita di essere servita.	La CPU rimane bloccata nello stato di attesa (WAIT), finchè la periferica non esaurisce il trasferimento dei dati.
Vantaggi	Semplice e via software.	La CPU ottimizza i tempi.	Utilizzata per trasferimento dati ad alta velocità.
Svantaggi	La CPU impiega tempo per interrogare le periferiche.	Complessa.	Necessità di hardware.

Unità 5

Introduzione alle reti di computer

■ Rete informatica

L'insieme dei dispositivi (computer, router, switch) interconnessi tra loro mediante mezzi trasmissivi. Lo scopo principale di una rete informatica è di **codividere risorse hardware** (memorie, dischi, periferiche, ecc) e **risorse software** (programmi applicativi e file di dati). Le reti offrono svariati vantaggi, tra cui **affidabilità, efficienza e velocità, riduzione dei costi di esercizio e scalabilità del sistema.**

Tabella. Vantaggi e svantaggi delle reti.

Vantaggi	Svantaggi
<ul style="list-style-type: none">• Periferiche condivise• Gestione remota di documenti condivisi• Salvataggio dei dati centralizzato• Scalabilità• Costi ripartiti	<ul style="list-style-type: none">• Amministrazione della rete• Gestione della sicurezza• Difficoltà nella ricerca dei guasti

■ Classificazione delle reti

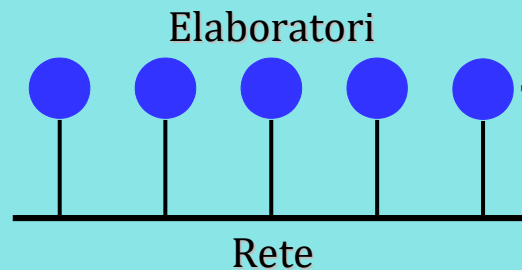
Le reti possono essere classificate in base a determinati requisiti, quali:

- l'**estensione geografica**, che comprende il raggio d'azione su cui operano;
- l'**architettura**, cioè come sono organizzate;
- la **topologia**, cioè il modo in cui è disposta la rete sul territorio, cioè la configurazione dell'hardware e il tipo di connessioni adottate.

■ Tecnologia trasmissiva

Reti broadcast

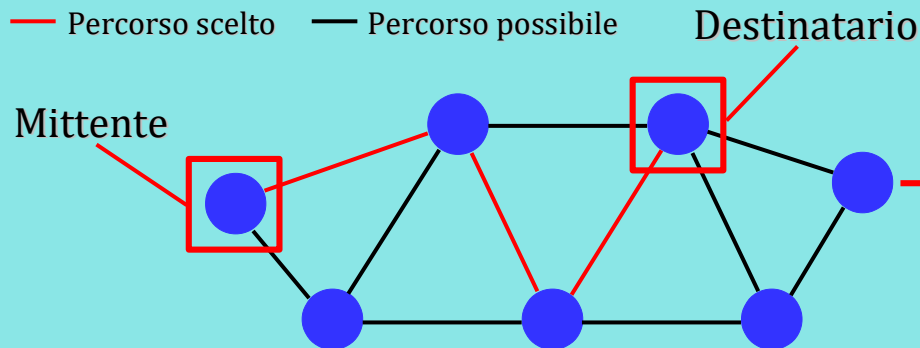
- I vari elaboratori sono connessi tra loro tramite un canale di trasmissione condiviso su cui può trasmettere un dispositivo alla volta (**broadcast**). Un breve messaggio, di solito detto **pacchetto**, trasmesso da un elaboratore è ricevuto da tutti gli altri. Un indirizzo all'interno del pacchetto specifica il vero destinatario.
- Le reti broadcast consentono anche di inviare un pacchetto di dati a più destinatari simultaneamente usando un opportuno indirizzo. In tal caso tutti prendono in considerazione il pacchetto (**multicast**).



Quando un computer riceve un pacchetto dati esamina l'indirizzo del destinatario; se questo corrisponde con il proprio indirizzo lo accetta e lo elabora, altrimenti lo ignora.

Reti punto a punto

- Consistono di un insieme di collegamenti tra due elaboratori o tra due gruppi di elaboratori. Per arrivare alla corretta destinazione un pacchetto di dati viene trasmesso dal mittente all'host a cui è direttamente connesso, che lo rigenera al successivo fino a giungere all'host destinatario.



Il percorso che il messaggio deve compiere non è noto a priori. La scelta dei percorsi viene fatta da appositi algoritmi di instradamento.

■ Throughput e bandwidth

Velocità di trasmissione: numero di bit trasmessi in un secondo ed è misurata in bit per secondo (bps).

La velocità di trasmissione viene espressa attraverso due concetti:

- **throughput:** quantità di dati (reale) che sono transitati in un canale in un certo periodo di tempo e viene espressa come numero di bit trasmessi in un secondo;
- **bandwidth** (larghezza di banda): quantità di dati massima (teorica) che può fluire mediante un canale in un certo periodo di tempo e viene espressa come la quantità di dati trasmissibili in un secondo.

Tabella. Unità di misura della velocità di trasmissione di una linea.

Unità di misura	Simbolo	Equivalenza
bit per secondo	bps	Unità di misura
Kilobit per secondo	Kps	$1 \text{ Kpbs} = 10^3 \text{ bps} = 1000 \text{ bps}$
Megabit per secondo	Mpps	$1 \text{ Mpbs} = 10^6 \text{ bps} = 1000000 \text{ bps}$
Gigabit per secondo	Gpps	$1 \text{ Gpbs} = 10^9 \text{ bps} = 1000000000 \text{ bps}$
Terabit per secondo	Tpbs	$1 \text{ Tpbs} = 10^{12} \text{ bps} = 10000000000 \text{ bps}$

Nota bene

Bandwidth è la velocità teorica raggiungibile mentre **throughput** è la velocità effettivamente misurata che risulta sempre minore della bandwidth per diversi motivi: numero di utenti in rete, topologia della rete, dispositivi di rete, disturbi presenti nel mezzo trasmissivo (rumore).

**Generalmente chiamata
NIC
(Network Interface Card)**

Scheda di rete

Circuito stampato che collega il PC al cavo. Detta anche LAN adapter.

Si dividono in:

- Ethernet (10 Mbps);
- Fast Ethernet (100 Mbps);
- Gigabit Ethernet (1 Gbps).

Nota con il nome di scheda Ethernet: fornisce porta RJ45 alla quale si collega il cavo di rete.

Indirizzo fisico formato da 6 byte (48 bit). I primi tre byte detti OUI denotano la azienda che ha progettato la scheda di rete.

Identificata da un codice univoco detto MAC (Media Access Control).

Hub

Ripete o trasmette tutti i dati ricevuti a tutte le porte.

- Definito ripetitore multiporta.
- Provoca un traffico inutile sulla rete.

- Hub attivi: necessitano di alimentazione poiché amplificano il segnale.
- Hub passivi: servono solo come punto di connessione fisica dei cavi.

Dispositivi di rete

Switch

Valuta il contenuto di un dato ricevuto e lo spedisce solo al reale destinatario.

Router

Permette l'accesso di una rete LAN ad una altra rete (per esempio Internet).

- Riduce il traffico inutile nella rete.
- Prestazioni della rete migliori.

■ Classificazione per estensione

In relazione all'**area geografica** su cui operano, possiamo distinguere le reti in:

- **PAN** (Personal Area Network), che agiscono in una area di pochi metri ad uso personale;
- **LAN** (Local Area Network – Reti Locali), localizzate nello stesso edificio oppure in edifici limitrofi; sono molto veloci e affidabili;
- **MAN** (Metropolitan Area Network), che coprono l'area di una città o di una provincia;
- **WAN** (Wide Area Network – Reti Geografiche), che si estendono a notevoli distanze fino a coprire l'intero globo come Internet.

Tabella. Tipi di rete e loro caratteristiche.

Distanza	Dove si trovano	Sigla	Velocità tipiche
1 – 10 m	Vicina all'utente	PAN	1 – 3 Mbps (Bluetooth)
10 – 100 – 1000 m	Stanza, ufficio, edificio	LAN	100 Mbps – 10 Gbps
10 – 100 km	Città, area regionale	MAN	100 Mbps – 10 Gbps
100 – 10000 km	Nazione, continente, pianeta	WAN	Connessioni a banda larga: fino a 1 Gbps

- **LAN**: una rete di computer e/o dispositivi confinata in un unico edificio, o in più edifici vicini. Lo scopo di una rete LAN è la condivisione di risorse hardware e software eseguita mediante il trasferimento dei dati da un computer (o da un dispositivo) all'altro.
- **WAN**: una rete di computer e/o dispositivi distribuiti su un'area geografica vasta. Lo scopo di una WAN è quello di permettere la comunicazione tra computer e dispositivi a livello globale.

■ Reti locali

Sviluppate su aree limitate tramite collegamenti fisici che non attraversano il suolo pubblico. Sono **economiche**, **flessibili**, **affidabili** e **facilmente installabili**. Progettate per essere veloci su aree normalmente inferiori al kilometro.



Si distinguono dagli altri tipi di reti per tre caratteristiche:

- dimensione**: non può andare oltre un certo limite che dipende dal tipo di cavo usato per la trasmissione;
- tecnologia trasmissiva**: sono in genere reti **broadcast**; velocità di trasmissione tipiche da 10 a 100 Mbps;
- topologia**: forma geometrica con la quale i nodi (**host**) sono tra loro connessi.



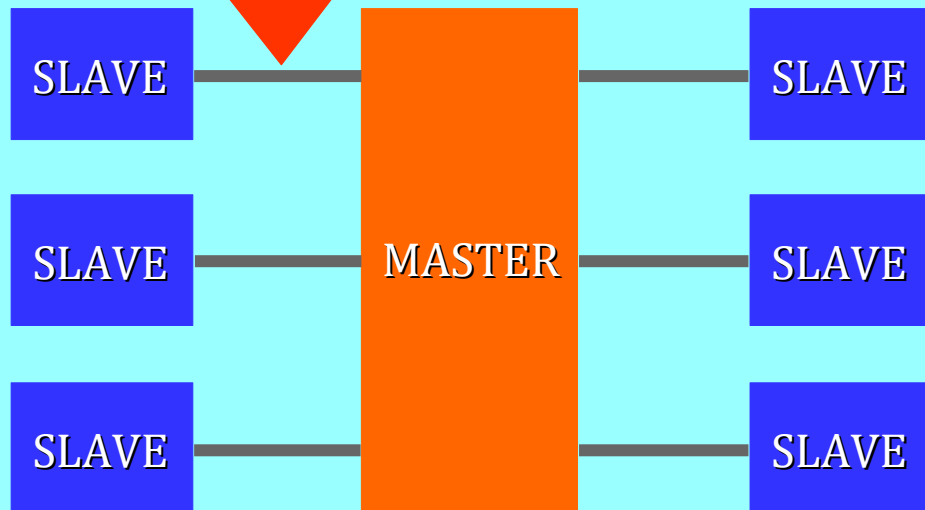
Il sistema di rete locale più diffuso è l'**Ethernet** basato sul protocollo **CSMA/CD**, con topologia a **bus** o a **stella**. Risponde ad un insieme di standard indicati sinteticamente con il codice **IEEE 802.3**. L'utilizzo di Ethernet consente l'inserimento, la modifica e l'eliminazione di dispositivi senza interrompere o modificare i servizi di rete. L'efficienza della rete degrada rapidamente all'aumentare del numero di nodi a causa dell'elevata probabilità di **collisioni**.

■ Classificazione per architettura

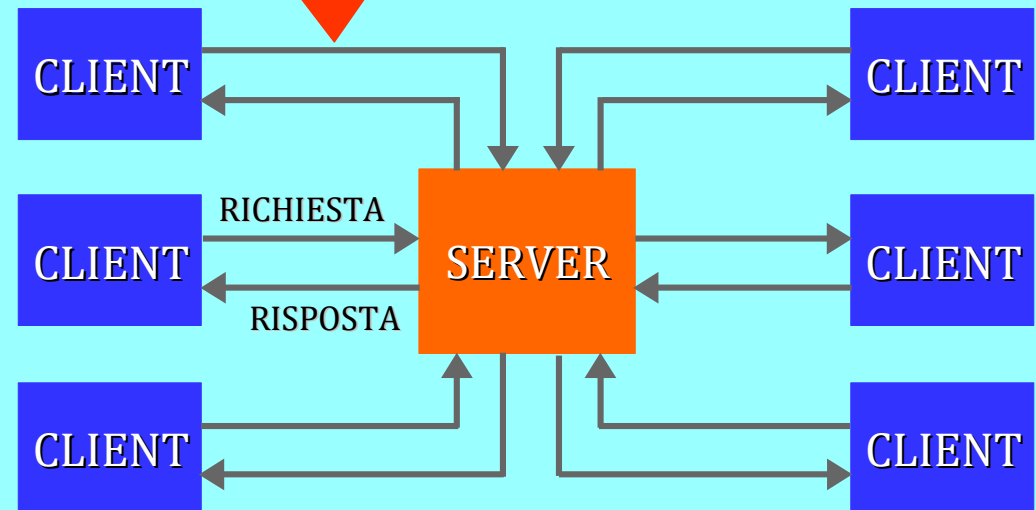
In funzione di **come sono organizzate**, le reti si distinguono in:

- **Master-Slave**: in questo modello comunicativo viene delegato a un dispositivo o processo software (master/mainframe) il controllo degli altri dispositivi connessi in rete (slave). Il master interroga o invia comandi alle varie stazioni;
- **Client-Server**: è il modello più diffuso sia nella rete Internet che nella maggior parte delle reti aziendali. E' richiesta la presenza di un sever che si occupa di rispondere alle richieste che gli vengono inoltrate dalle componenti client;
- **Peer-to-Peer**: in questo modello i dispositivi sono tutti “paritetici”, senza nessun punto di controllo centralizzato (niente Server). Solitamente tali architetture di rete si applicano a pochi dispositivi, come quelle presenti nelle reti domestiche.

Un singolo processo (master) ha il controllo su uno più processi (gli slave). I processi slave non comunicano tra loro, ma devono passare dal master.

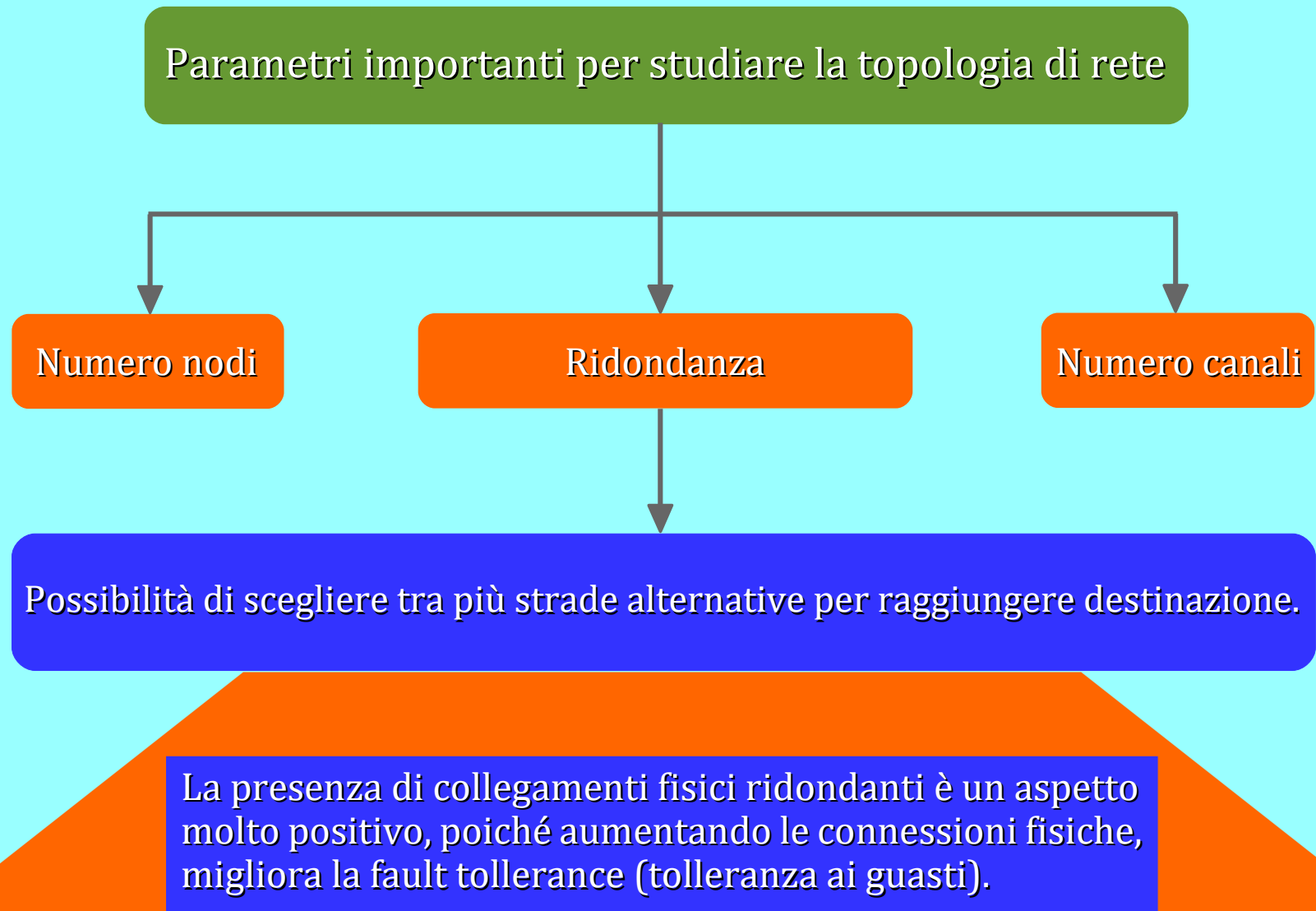


Il server svolge le operazioni necessarie per fornire un servizio, mentre il client richiede un servizio inviando le richieste al server.



■ Classificazione per topologia

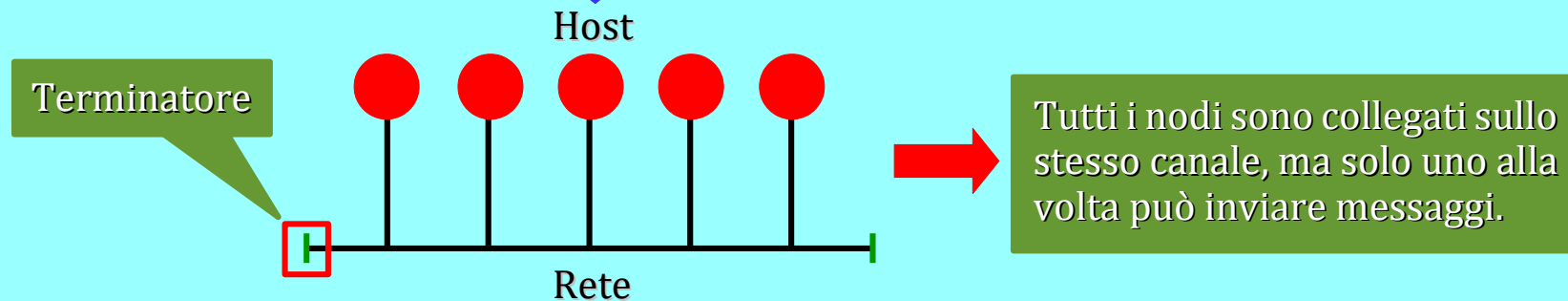
La topologia definisce la struttura di una rete, cioè il modo con cui i dispositivi (nodi) sono fisicamente collegati tra loro. Essa quindi determina la **dimensione e la forma**, l'**affidabilità**, i **costi**, l'**espandibilità e la complessità** della rete.



■ Topologia a bus

La **topologia a bus** usa un **singolo backbone** (linea principale), detto bus, a cui si collegano tutti gli host; alle **due estremità del cavo** è posta una resistenza terminale (**terminatore**).

Un generico dispositivo della rete. Può essere un personal computer o un dispositivo mobile, un server o un client, una workstation o una stampante, ma anche dispositivi per la comunicazione tra reti come un router o uno switch.

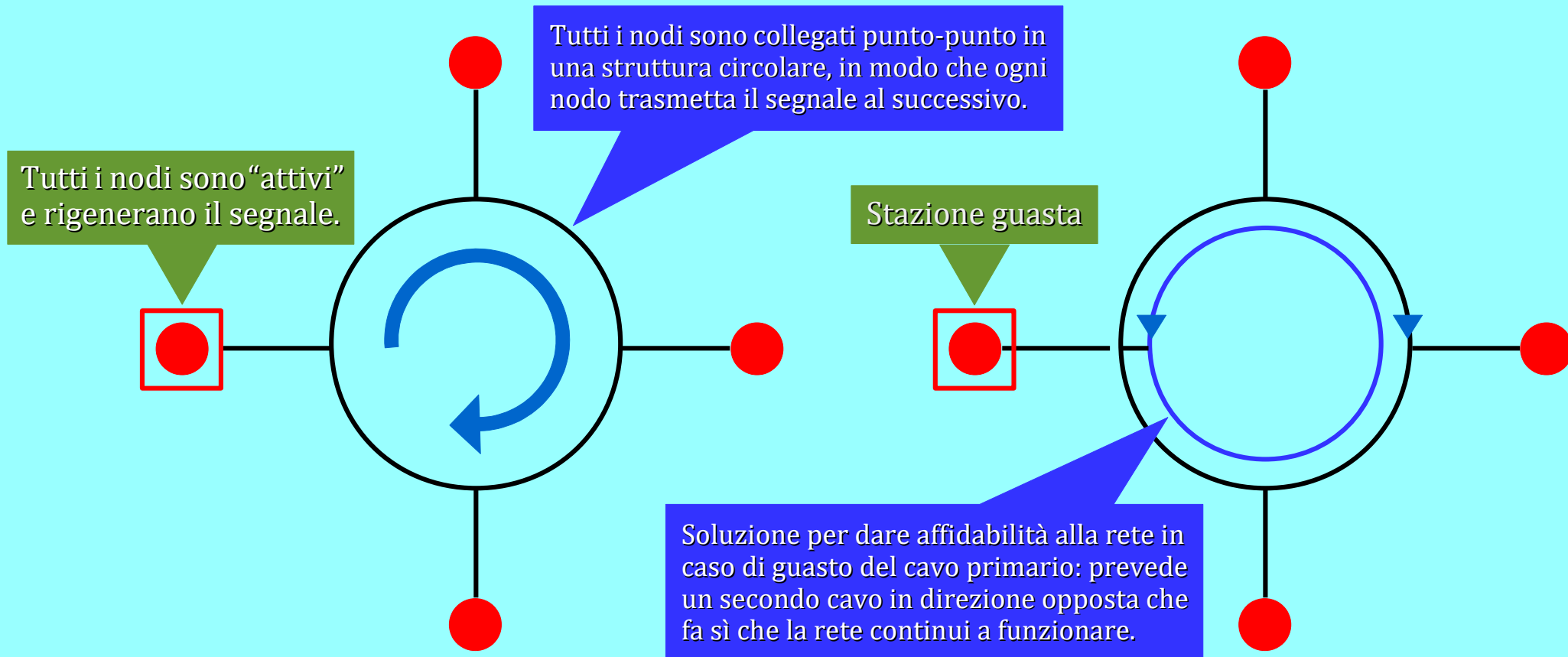


Tutti i nodi sono collegati sullo stesso canale, ma solo uno alla volta può inviare messaggi.

- Topologia a **basso costo** in quanto esiste un solo canale che collega tutti i nodi.
- I segnali passano lungo i cavi tra i due terminatori e vengono controllati da tutti gli host connessi al bus: solo se l'indirizzo di destinazione del messaggio coincide con quello dell'host, il messaggio viene ricevuto ed elaborato dall'host.
- Si tratta quindi di una **trasmissione di tipo broadcast** (cioè inviata a tutti).
- Se un host non funziona la rete continua a funzionare.
- **Svantaggio**: un guasto sul cavo provoca il malfunzionamento dell'intera rete.
- Questa topologia è **tipica delle reti locali (LAN) e metropolitane (MAN)**.
- Molto usata in passato, attualmente non viene più realizzata per la sua **bassa tolleranza ai guasti**.

■ Topologia ad anello

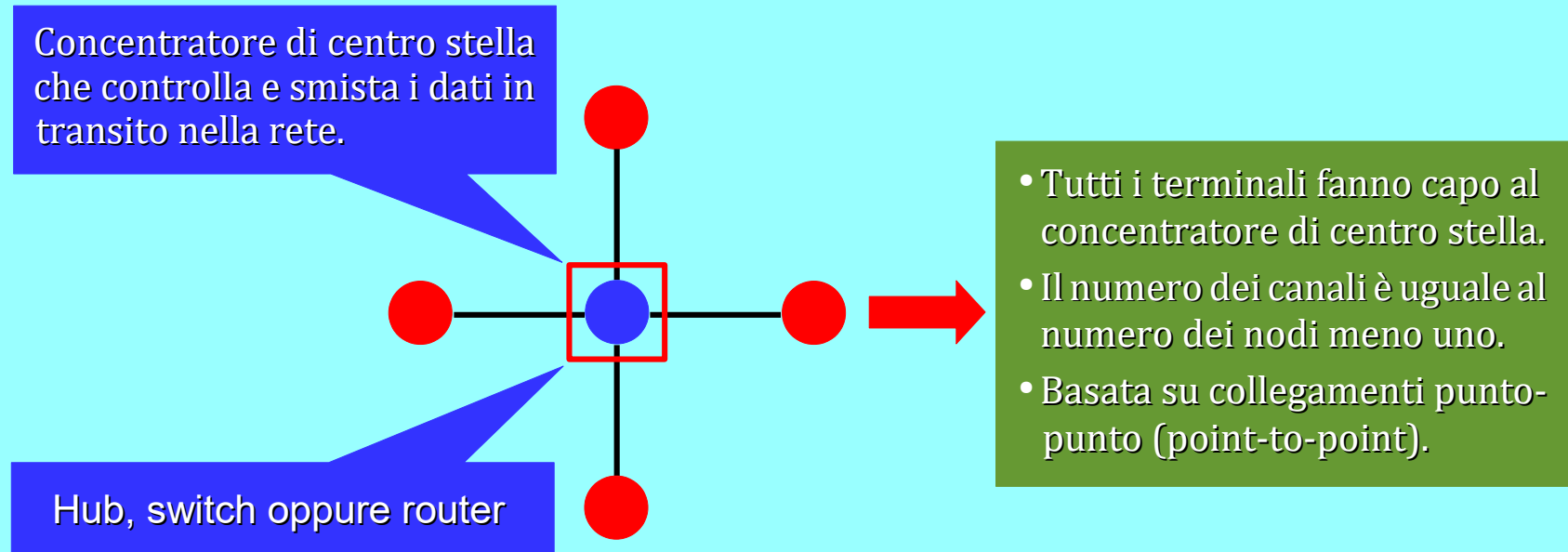
Nella **topologia ad anello (ring)** un **cavo collega un host al precedente e al successivo** creando un **circuito di rete continuo** su cui sono trasmessi i dati.



- La topologia ad anello può essere **unidirezionale** (ossia esiste un solo percorso possibile tra ogni coppia di nodi) e **bidirezionale** (ossia esistono due percorsi possibili tra ogni coppia di nodi).
- Per aumentare l'affidabilità della rete si adotta la soluzione della rete ad anello bidirezionale per garantire il suo funzionamento anche in presenza di un guasto al cavo primario.
- Questa topologia è utilizzata nelle **reti locali (LAN)** e soprattutto in quelle **metropolitane (MAN)**.

■ Topologia a stella

In questa topologia **tutti gli host sono collegati a un punto centrale**, chiamato **centro stella**, che di solito è un hub, switch o router e costituisce il punto di collegamento comune in maniera che i computer siano in comunicazione l'uno con l'altro.



- Anche se questa topologia porta ad aumento del numero di cavi essa offre **vantaggi** in termini di:
 - **fault tolerance**: il guasto di un canale o nodo della rete non ne compromette il funzionamento;
 - **flessibilità ed espandibilità**: lo spostamento di un host da un punto ad un altro della rete oppure l'inserimento di uno nuovo non richiedono il fermo della rete;
 - **semplicità di gestione**.
- **Svantaggi**: se si guasta il centro stella, **la rete smette di funzionare**.
- La topologia a stella è utilizzata nelle **reti locali (LAN)** e nelle **reti geografiche (WAN)**.

■ Protocollo di comunicazione

Un **insieme di regole** che definiscono l'ordine e il formato dei messaggi scambiati tra entità diverse di una rete, e quindi le azioni che si vengono ad attuare.

Per esempio, Susy entra nel bar per bere un caffè.
Per prima cosa saluta Mario, il barista che conosce da tanti anni.

“Ciao Mario.”

“Ciao Susy.”

“Mi fai un buon caffè ristretto, per favore? Oggi ho mal di testa...”

“Subito!”

Mario serve il caffè...

“Grazie molto buono. Quant'è?”

“Un euro.”

“Ciao.”

“Ciao Susy.”

Saluto
(apertura comunicazione)

Richiesta e fornitura servizio
(mantenimento comunicazione)

Saluto finale
(chiusura comunicazione)

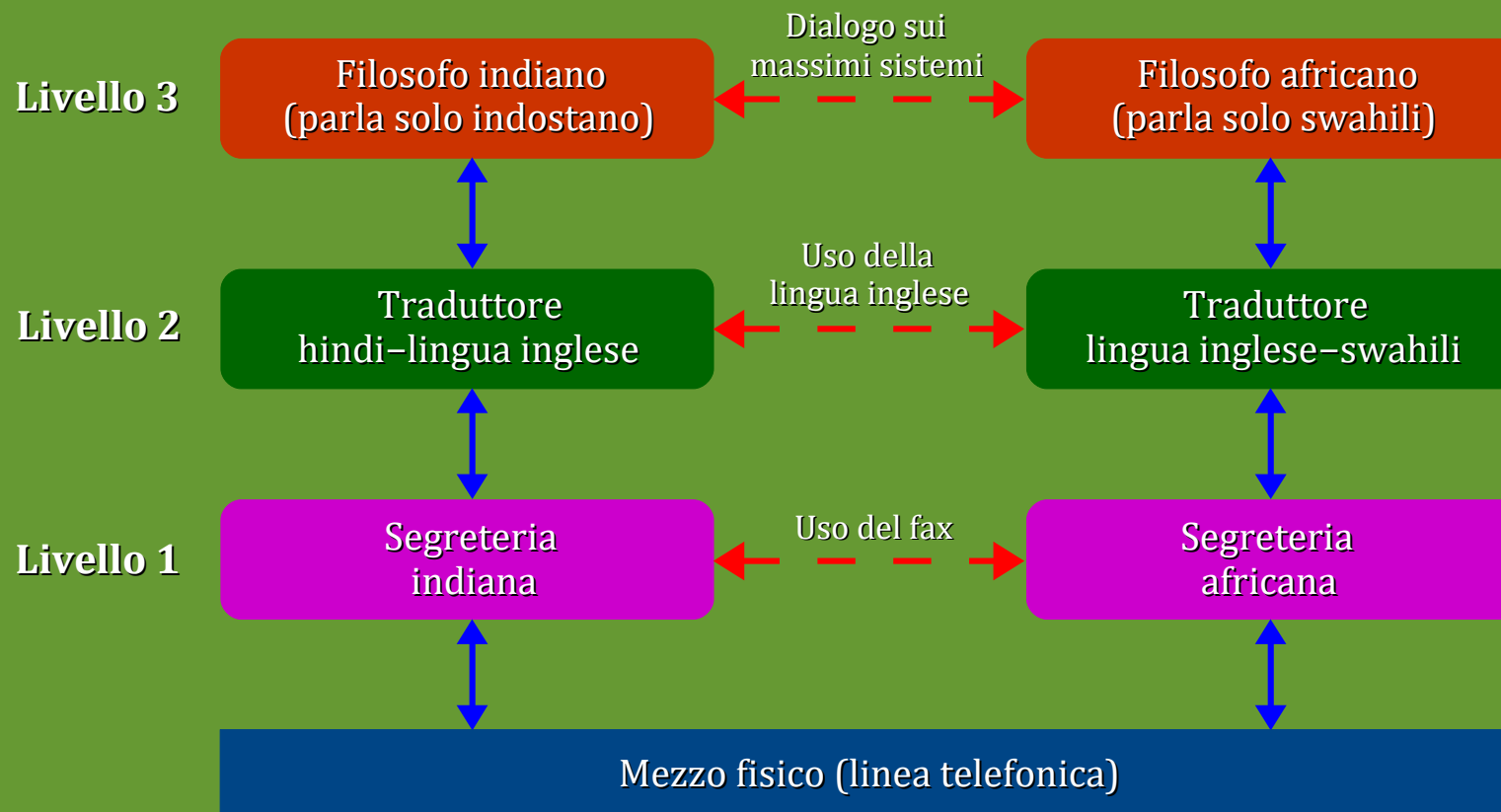
Prerequisito essenziale per la buona riuscita della comunicazione è che vengano rispettate certe regole di educazione e correttezza come il saluto, la richiesta, la risposta e il servizio.

■ Architettura di rete

L'insieme dei **livelli** e **protocolli** costituisce l'architettura di rete. Le tre architetture di rete più note sono: **SNA** di IBM (architettura di tipo proprietario), **OSI**, **TCP/IP**.

Analogia

Per comprendere i meccanismi basilari del funzionamento dell'architettura di rete si può pensare alla seguente analogia umana, in cui un filosofo indiano vuole conversare con un filosofo africano per scambiarsi opinioni filosofiche. Poichè i due filosofi parlano due lingue differenti (indostano e swahili), entrambi decidono di servirsi di una lingua comune (inglese) con la quale comunicare e di assumere a questo scopo dei traduttori hindi-inglese e swahili-inglese.



■ Il modello OSI

Modello standard di riferimento proposto dall'ISO per l'interconnessione di sistemi aperti.

Garantisce l'interoperabilità tra i sistemi e dispositivi prodotti da aziende diverse.

Utilizza una struttura formata da una pila (stack) di protocolli divisa in sette livelli.

Livello: un insieme di software utili per la soluzione di determinati problemi. La caratteristica di ogni livello è che deve fornire servizi a livello superiore nascondendo i dettagli dell'implementazione.

Livelli inferiori (Media Layers): si occupano del trasferimento delle informazioni tra i due sistemi.

Livelli superiori (Host Layers): si occupano delle applicazioni.

Dispositivo terminale: sorgente o destinatario delle informazioni.

Dispositivo intermedio: assicura la connessione fisica o logica tra due o più sistemi terminali.

Livelli superiori
(vicini all'utente)

Livelli inferiori
(vicini all'hardware)

Host layer

Media layer

Applicazione

Presentazione

Sessione

Trasporto

Rete

Collegamento

Fisico

Nei nodi intermedi sono presenti solo i primi tre livelli **OSI** mentre i nodi terminali hanno tutti i livelli **OSI**.

■ I livelli OSI

Livello	Funzionalità	Unità dati
7	Applicazione Contiene i programmi applicativi che consentono all'utente di svolgere le sue attività in rete: trasferimento, accesso dei file, posta elettronica, terminale virtuale ecc.	Dati
6	Presentazione Traduce i dati trasmessi in formati standard; gestisce anche la riservatezza e la protezione dei dati mediante algoritmi di criptaggio e autenticazione.	Dati
5	Sessione Ha lo scopo di aprire e chiudere una comunicazione tra due host in rete, nonché riaprirli nel caso di eventi indesiderati: perdita di dati, caduta della linea ecc.	Dati
4	Trasporto Si occupa di rendere affidabile la trasmissione dei pacchetti garantendo che arrivino al destinatario nell'ordine corretto, senza errori, duplicazioni o perdite.	Segmenti
3	Rete Trasmissione dei pacchetti mediante la rete, dal dispositivo sorgente al destinatario; a questo livello la trasmissione dei dati non è affidabile (non gestisce gli errori nei frame).	Pacchetti
2	Collegamento Trasmissione/ricezione di pacchetti di bit chiamati frames, garantendo che i bit arrivino a destinazione correttamente, senza errori, duplicazioni o perdite.	Struttura
1	Fisico Trasmissione o ricezione delle sequenze binarie sul mezzo trasmissivo; definisce le proprietà meccaniche ed elettriche dei collegamenti di rete (cavi, prese, modulazioni).	Bits