



UNIVERSITÀ DEGLI STUDI DI PALERMO

INFORMATIZZAZIONE DELLA DOCUMENTAZIONE SANITARIA

Anno Accademico 2019/2020

Docente: ing. Salvatore Sorce

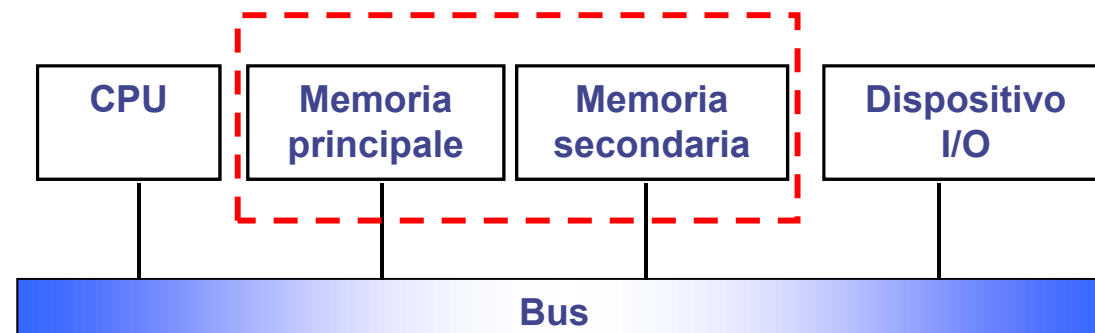
Architettura dei calcolatori

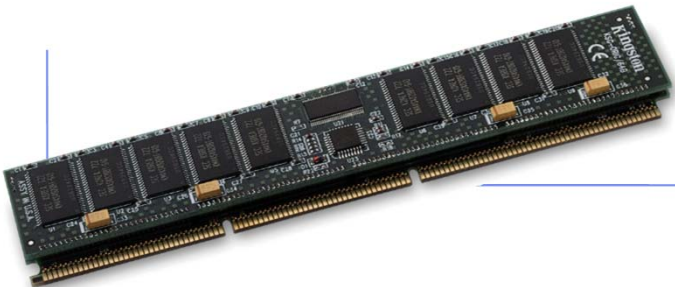
II parte – Memorie



Struttura di un calcolatore

- Componenti principali:
 - Unità di controllo di processo (Central Processing Unit)
 - **Memoria principale (Main memory)**
 - Memoria secondaria (Mass storage)
 - Dispositivi di ingresso/uscita







La memoria

- **Supporto alla CPU:** deve fornire alla CPU dati e istruzioni il più rapidamente possibile
- **Archivio:** deve consentire di archiviare dati e programmi garantendone la conservazione e la reperibilità anche dopo elevati periodi di tempo
- **Diverse esigenze:**
 - **velocità** per il supporto alla CPU
 - **non volatilità** ed **elevate dimensioni** per l'archivio
- **Diverse tecnologie**
 - **elettronica:** veloce, ma costosa e volatile
 - **magnetica** e **ottica:** non volatile ed economica, ma molto lenta



Criteri di caratterizzazione della memoria

- Velocità
- Volatilità
- Capacità
- Costo (per bit)
- Modalità di accesso
 - **diretta** (o casuale)
 - **sequenziale**
 - **mista**
 - **associativa**

The slide features a minimalist design with several thin blue lines. A vertical line on the left and a horizontal line at the top intersect at a small circle in the upper-left corner. Another horizontal line is positioned below the title. A vertical line on the right and a horizontal line at the bottom intersect at a small circle in the lower-right corner. A vertical line also runs along the right edge of the slide.

La memoria centrale



La memoria centrale (RAM)

- Mantiene al proprio interno i dati e le istruzioni dei programmi in esecuzione
- Memoria ad accesso “casuale”
- Tecnologia elettronica
 - veloce ma volatile e costosa
- Due “eccezioni”
 - ROM: elettronica ma permanente e di sola lettura
 - Flash: elettronica ma non volatile e riscrivibile



Indirizzi di memoria

- I bit nelle memorie sono raggruppati in celle:
 - tutte le celle sono formate dallo **stesso numero di bit**;
 - una cella composta da **k bit**, è in grado di contenere una qualunque tra **2^k combinazioni** diverse di bit.
 - Ogni cella ha un indirizzo:
 - serve come accesso all'informazione;
 - in una memoria con **N celle** gli indirizzi vanno da **0** a **N-1**.
 - La cella è l'unità indirizzabile più piccola.
- In quasi tutti i calcolatori è di 8 bit (un byte).
- I byte vengono raggruppati in parole (che oggi sono di 32/64 bit), su cui la CPU esegue le operazioni.

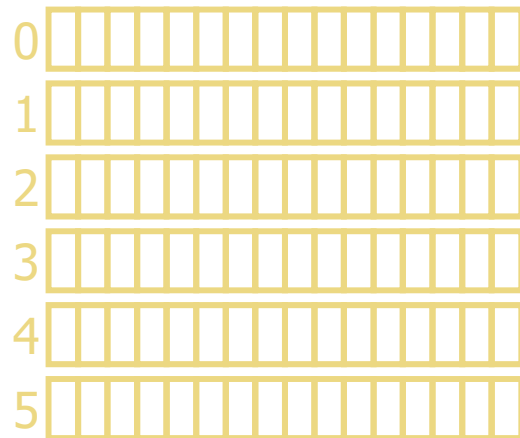


Organizzazione della memoria

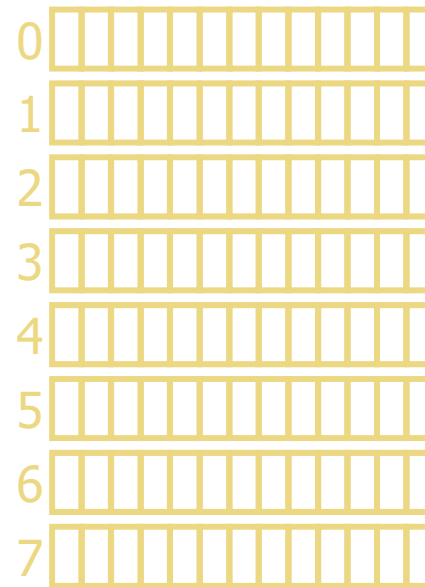
- Anche gli indirizzi della memoria sono rappresentati come numeri binari:
 - un indirizzo di **M** bit consente di indirizzare **2^M** celle;
 - per 6 o 8 celle bastano 3 bit, per 12 celle ne servono 4;
 - il **numero di bit nell'indirizzo** determina il **numero massimo di celle indirizzabili** nella memoria ed è indipendente dal numero di bit per cella (una memoria con 2^{12} celle richiede sempre 12 bit di indirizzo, quale che sia la dimensione di una cella).
- Una memoria può essere organizzata in diversi modi:
 - con 96 bit possiamo avere 6 celle di 16 bit ($6 \cdot 16 = 96$), o 8 celle di 12 bit ($8 \cdot 12 = 96$) o 12 celle di 8 bit ($12 \cdot 8 = 96$).



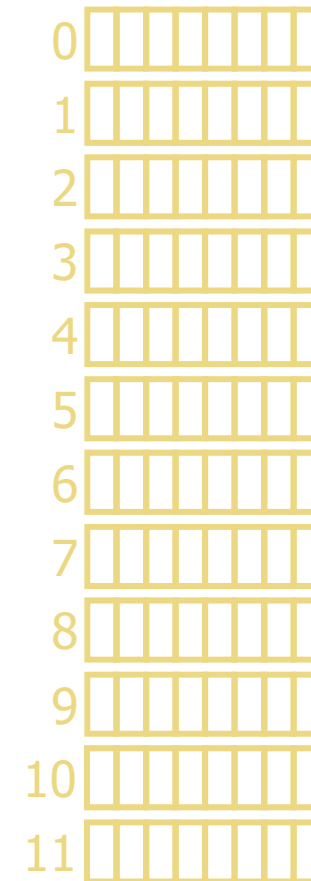
Organizzazione della memoria



6 parole da 16 bit



8 parole da 12 bit



12 parole da 8 bit



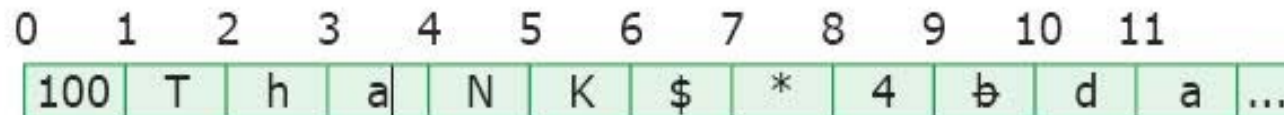
Memoria vs. CPU

- Le CPU sono sempre state più veloci delle memorie
 - l'aumento di integrazione ha consentito di realizzare **CPU** pipeline e super scalari, molto efficienti e **veloci**;
 - nelle memorie è aumentata la **capacità** più che la **velocità**.
- L'accesso alla memoria passa attraverso il bus
 - la **frequenza** di funzionamento **del bus** è molto **più bassa** di quella della CPU;
 - il **bus** può essere **impegnato** ad effettuare trasferimenti controllati **da dispositivi** di I/O "**autonomi**" (e.g. DMA).
- È difficile riordinare le istruzioni in modo da poter sfruttare i tempi di attesa della memoria.
- È possibile fare memorie molto veloci se stanno nel chip della CPU, ma sono piccole e costose.



Memoria principale - RAM

- Generalmente un diagramma della memoria di un computer rappresenta le locazioni discrete come piccoli rettangoli
- L'indirizzo delle locazioni è indicato sopra il rettangolo.
- I valori delle locazioni sono riportati nel rettangolo corrispondente.

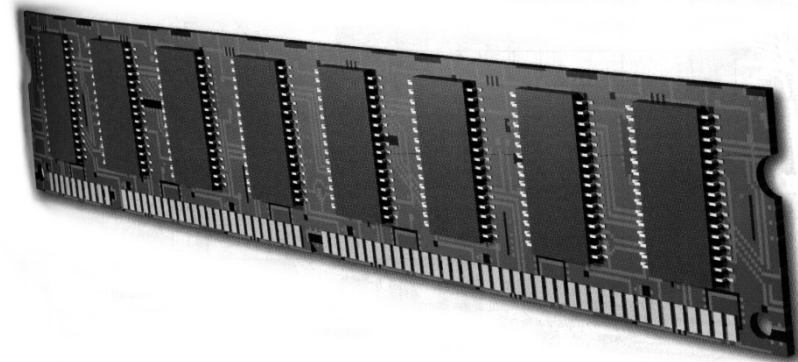


Un diagramma della memoria di un computer che illustra le sue caratteristiche chiave.



Memoria principale - RAM

- La memoria RAM è organizzata in byte (1 byte = 8 bit)
- Multipli del byte:
 - 1 KB = 2^{10} byte = 1024 byte
 - 1 MB = 2^{10} Kb = 1024 KB = 1.048.576 byte
 - 1 GB = 2^{10} MB = 1024 MB = 1.048.576 Kbyte
- "ad accesso casuale" significa che il computer può accedere direttamente a qualsiasi locazione di memoria, in un tempo che è indipendente dalla locazione

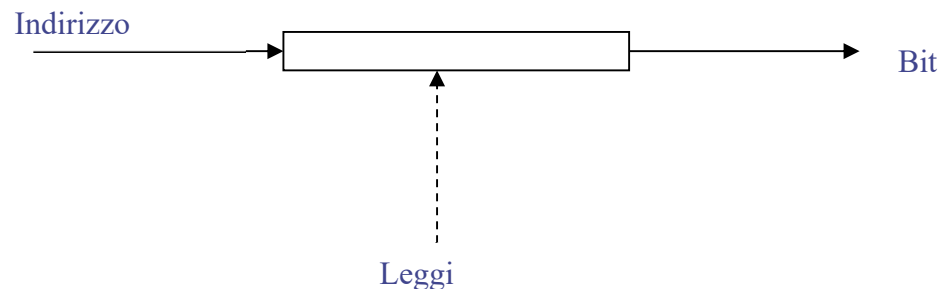


- La RAM è organizzata in moduli
 - SIMM (Single In-line Memory Module)
 - DIMM (Dual In-line Memory Module)
- Moduli di tipo SIMM vanno sempre aggiunti in coppie
- Chip da 1 Mbit, 8 chip =>
 - Modulo da 1 MB

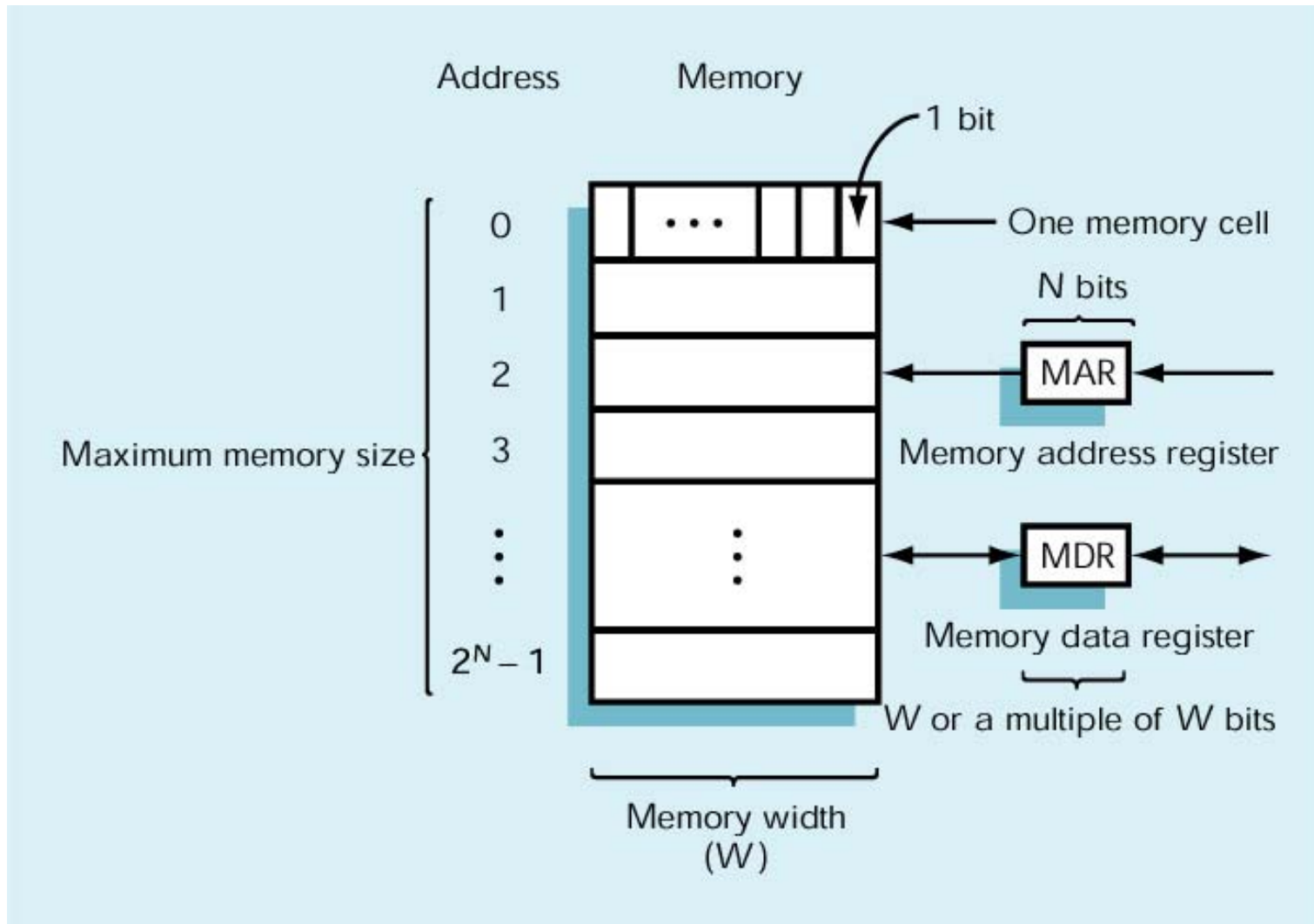


Memoria principale - RAM

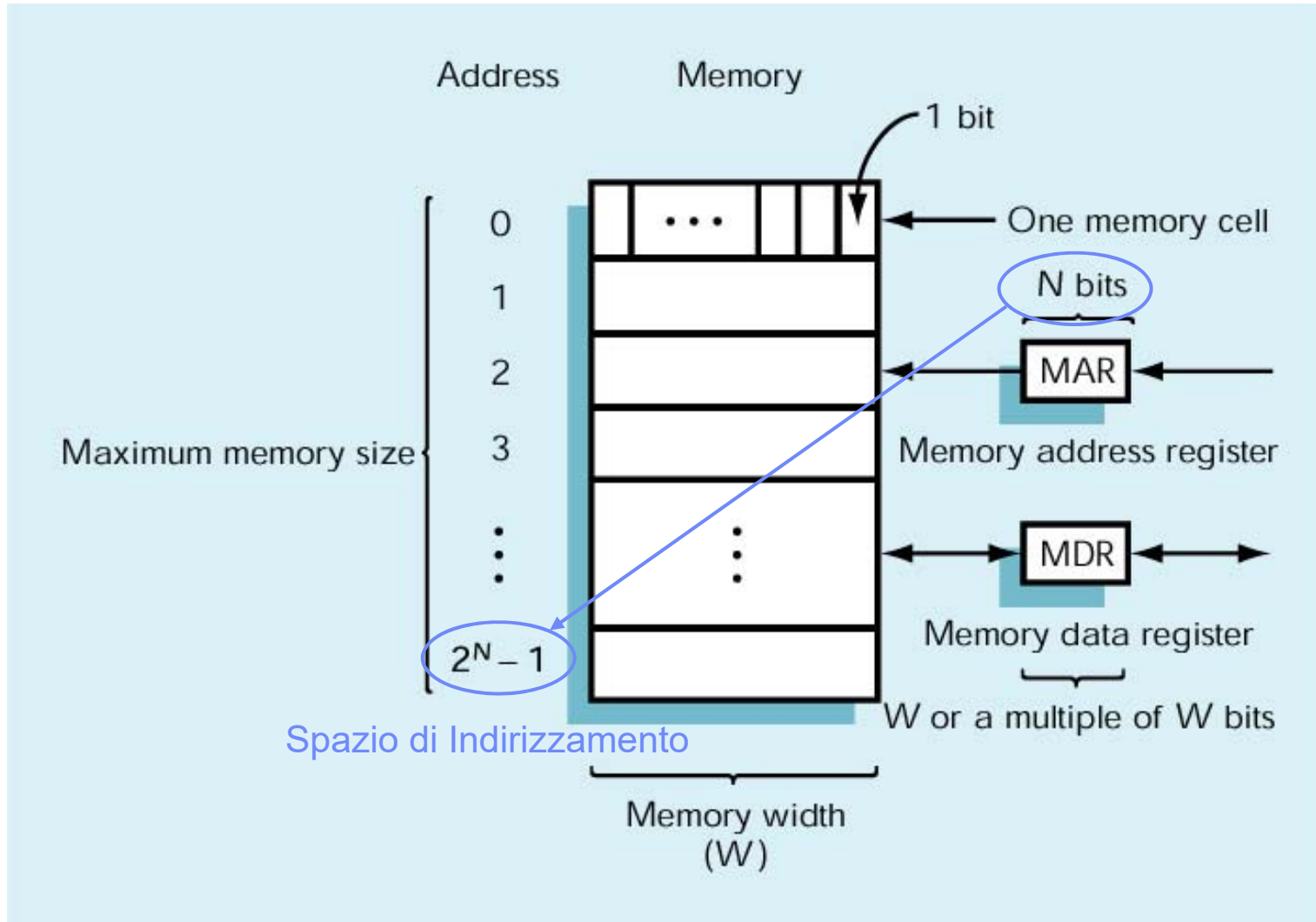
- Ciclo di lettura
 1. La memoria riceve un comando di lettura unitamente ad un indirizzo
 2. Il dato contenuto nell'indirizzo specificato viene ricercato e copiato in un buffer
 3. La memoria è pronta per un nuovo ciclo
- Tempo di accesso (o risposta) = tempo tra l'ingresso dell'indirizzo e l'uscita del contenuto
 - Per le RAM tempo di accesso = 60/70 ns
- Transfer rate = velocità di trasferimento (in Mbit) per ogni singola linea di memoria



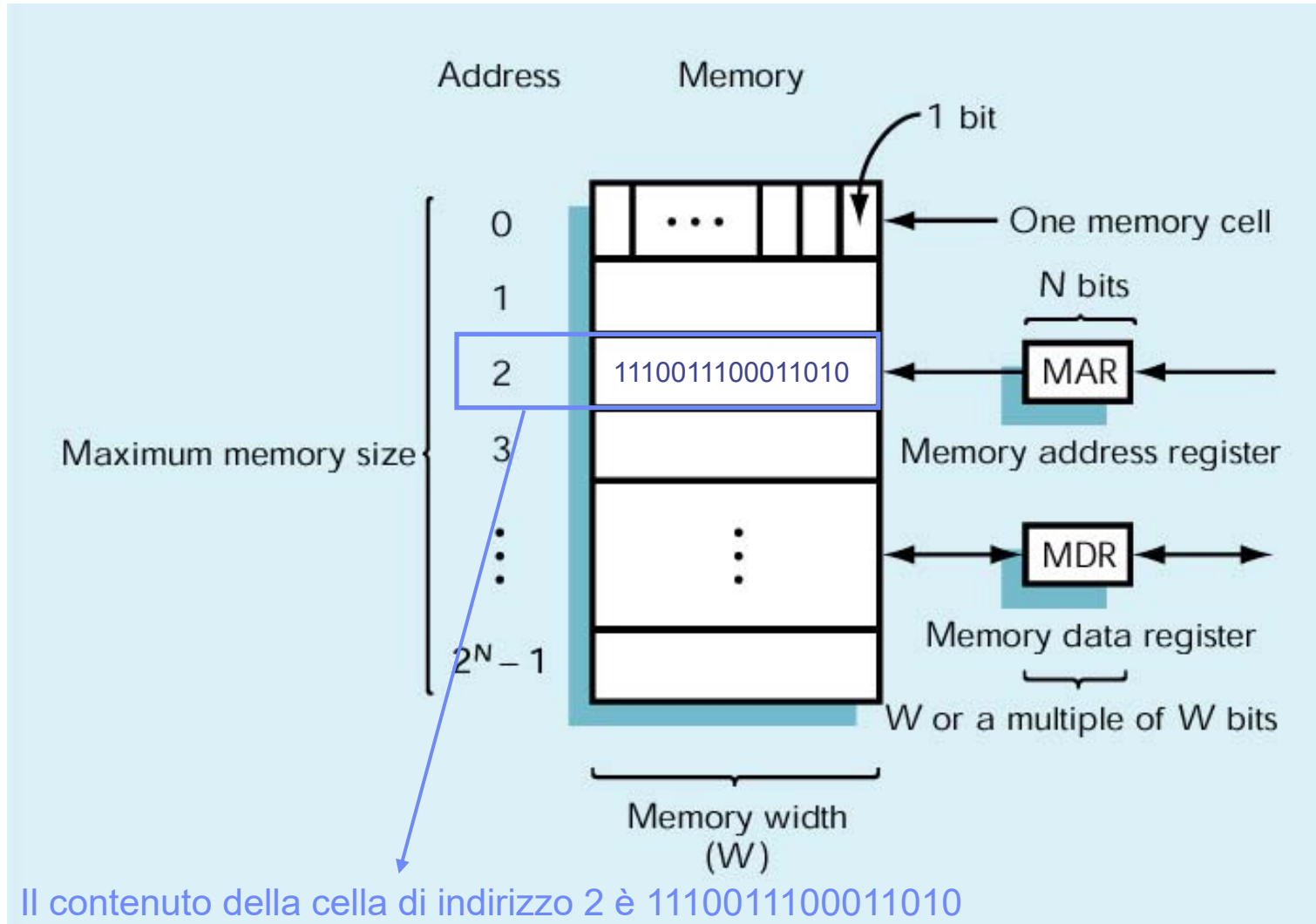
Memoria principale (RAM)



Memoria principale (RAM)



Memoria principale (RAM)

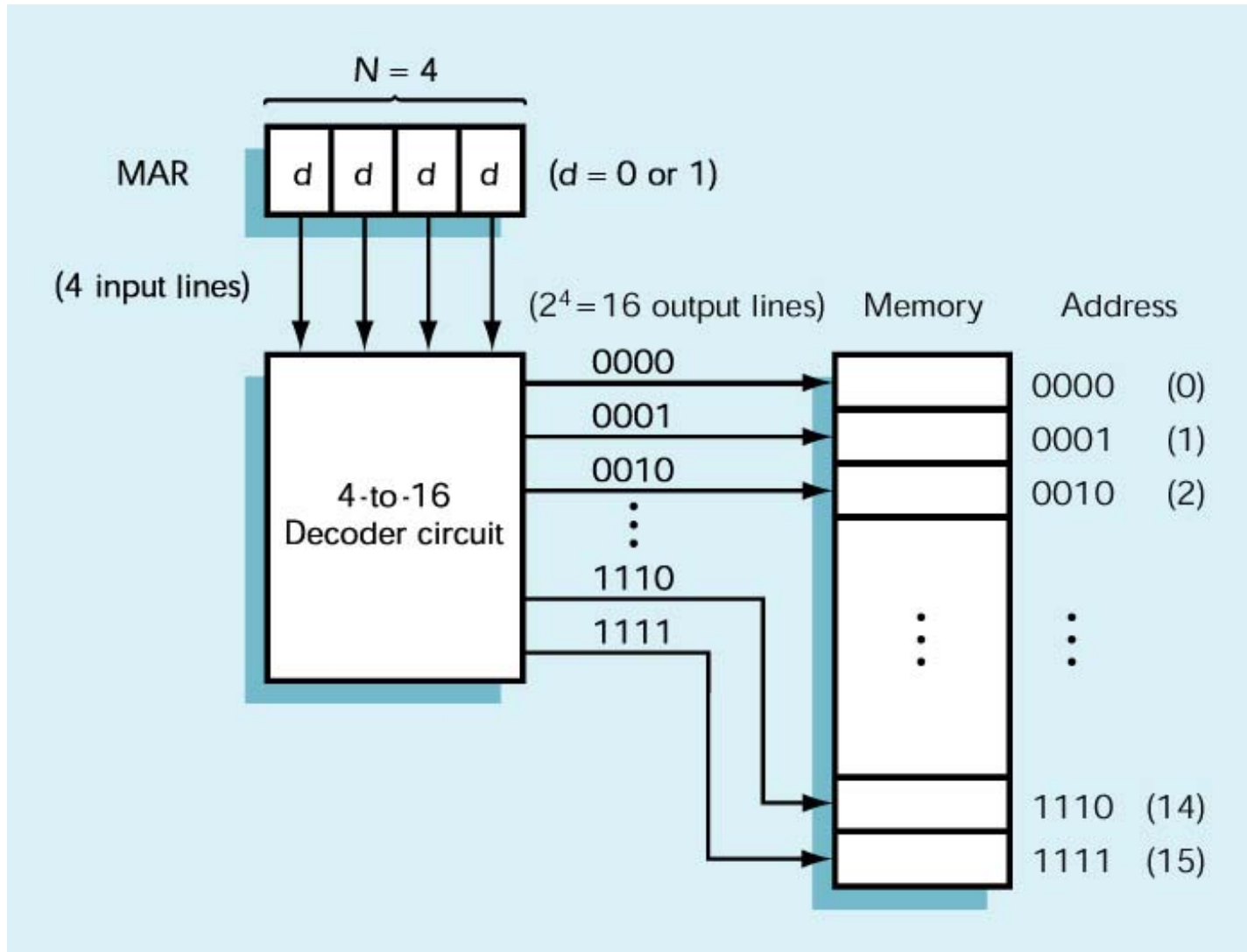




Memoria principale (RAM)

- Operazioni sulla memoria
 - Fetch(indirizzo)
 - ◆ Carica *indirizzo* nel MAR
 - ◆ Decodifica *indirizzo* del MAR
 - ◆ Copia contenuto della locazione in MDR
 - Store(indirizzo, valore)
 - ◆ Carica *indirizzo* nel MAR
 - ◆ Carica *valore* in MDR
 - ◆ Decodifica *indirizzo* del MAR
 - ◆ Copia *valore* di MDR nel contenuto della locazione

Memoria principale (RAM)





Memoria Cache

- Inizialmente un solo tipo di memoria -> CPU spesso inattiva
- Principio di località...:
 - ...temporale: quando un programma accede ad un dato o istruzione, c'è un'alta probabilità che nell'immediato futuro acceda alla stesso dato o istruzione
 - ...spaziale: quando un programma accede ad un dato o istruzione, c'è un'alta probabilità che nell'immediato futuro acceda anche ai dati o istruzioni vicini a quelli appena letti
- Non è quindi necessario avere tutta la memoria ad alta velocità, basta che lo sia anche solo una parte

Memoria Cache

- La **Memoria Cache** è una memoria ad alta velocità (più alta della RAM), e di capacità ridotta (più piccola della RAM)
- Mantiene i dati e le istruzioni momentaneamente usate dal processore + i loro "vicini" in memoria RAM
- Ciclo di lettura effettivo (con cache):
 - Ricerca del dato/istruzione nella cache;
 - Se la ricerca fallisce, viene rieseguita nella RAM e i dati vengono copiati nella cache per le prossime letture
- $T_a = (\text{hit ratio} \times t_{\text{cache}}) + (1 - \text{hit ratio}) \times (t_{\text{cache}} + t_{\text{RAM}})$

Gerarchia di memoria

- Per consentire al microprocessore di lavorare alla velocità più alta possibile, la memoria è organizzata in modo gerarchico
- I diversi livelli della gerarchia (procedendo dal microprocessore a scendere) sono caratterizzati da
 - Velocità decrescente
 - Dimensione crescente
- Forniscono l'illusione di una memoria infinitamente grande e veloce.

