Memory Key Characteristics

Introduction to Memory

- **Charles Babbage started Difference engine in1821 but failed its test in 1833, Why?**

  Due to unavailability of Memory

- **What is Memory?**
- **A single separate storage structure that holds information in the form of bits called as Memory**
- **The binary information may be instructions and data**
- **Stored program concept was introduced with the advent of vacuum tubes by John Von Neumann 1940**

Memory Capacity

- Number of bytes that can be stored

Term	Normal Usage	Usage as Power of 2
K (Kilo)	10^3	$2^{10} = 1,024$
M (Mega)	10^{6}	$2^{20} = 1,048,576$
G (Giga)	10^9	$2^{30} = 1,073,741,824$
T (Tera)	10^{12}	$2^{40} = 1,099,511,627,776$

Key Characteristics

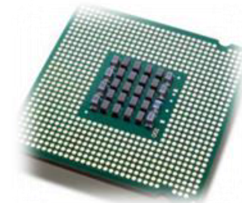
- Location
  - CPU
  - Internal (main)
  - External (secondary)
- Capacity
  - Word size
  - Number of words
- Unit of transfer
  - Word
  - Block
- Access methods
  - Sequential access
  - Direct access
  - Random access
  - Associative access
- Performance
  - Access time
  - Cycle time
  - Transfer rate

Key Characteristics contd.,

- Physical Type
  - Semiconductor
  - Magnetic surface
  - Optical
- Physical Characteristics
  - Volatile / Non-Volatile
  - Erasable / Non-erasable
- Organization

Location

- Three locations of memories
  - CPU
    - Registers – used by CPU as its local memory

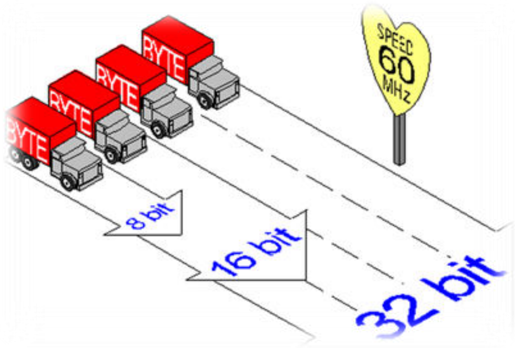


  - Internal memory
    - Main memory
    - Cache memory

  - External memory
    - Peripheral devices – disk, tape – accessible to CPU via I/O controllers

Capacity

- Internal memory capacity is expressed in terms of bytes or words.
- External memory capacity is expressed in terms of blocks (depends on words in memory)
- Total memory = number of words × word length
- Number of words = $2^{\text{address bus width}}$
- Word length = Data bus width

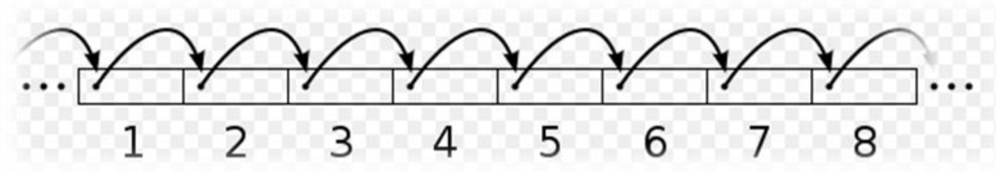


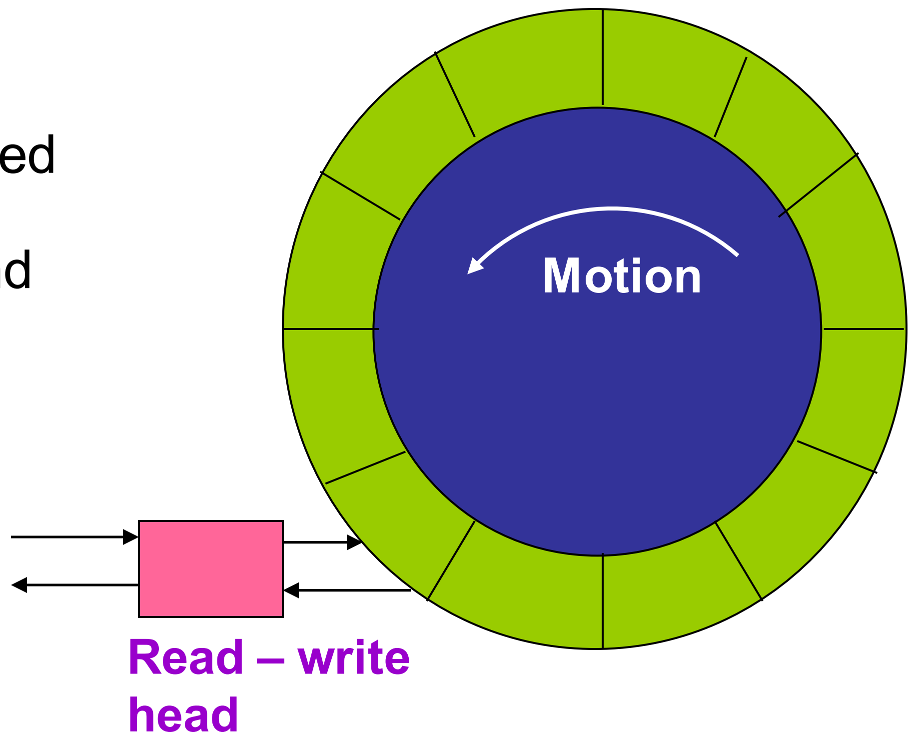
Unit of transfer

- Internal memory – number of data lines into and out of the main memory module
- External memory – blocks – longer units than a word

Access Methods

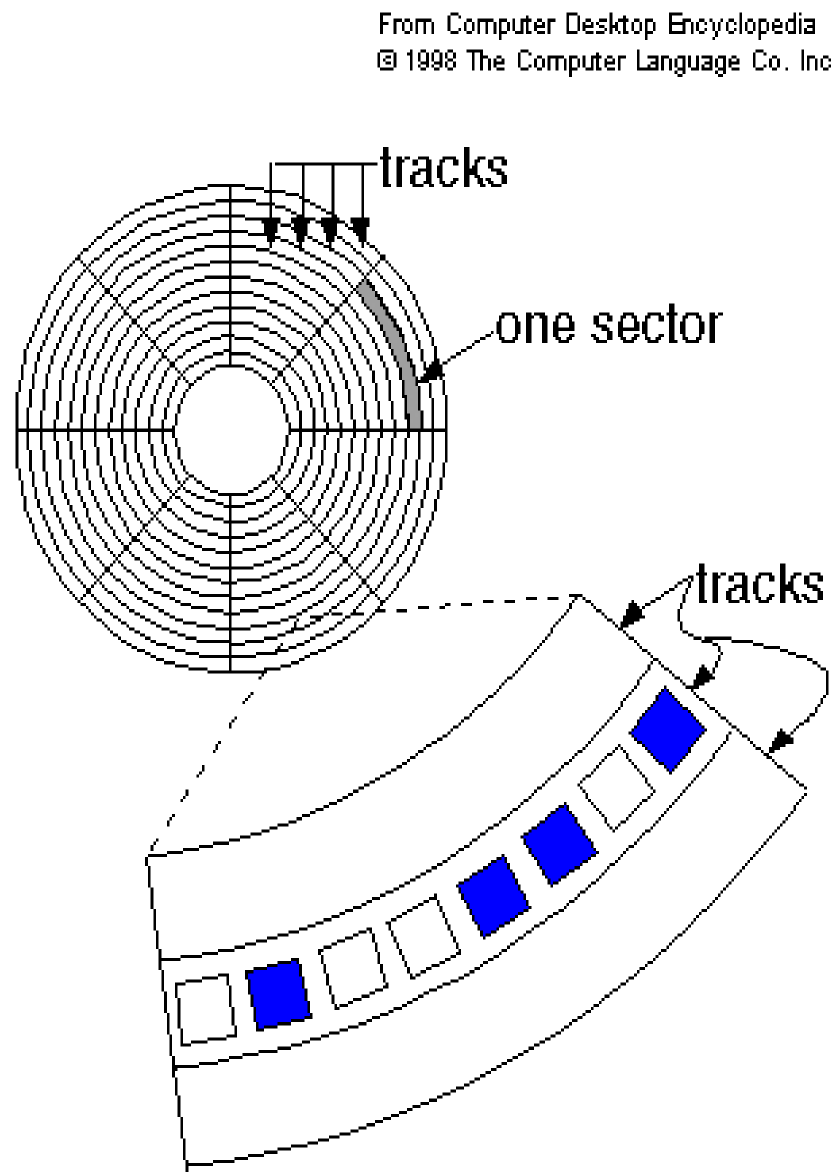
- Four types
  - Sequential Access
    - Accesses the memory in predetermined sequence
    - Shared read/write head is used, and this must be moved its current location to the desired location, passing and rejecting each intermediate record.
    - So, the time to access an arbitrary record is highly variable
    - Slower than random access memory
    - Ex: Magnetic Tapes, data in memory array





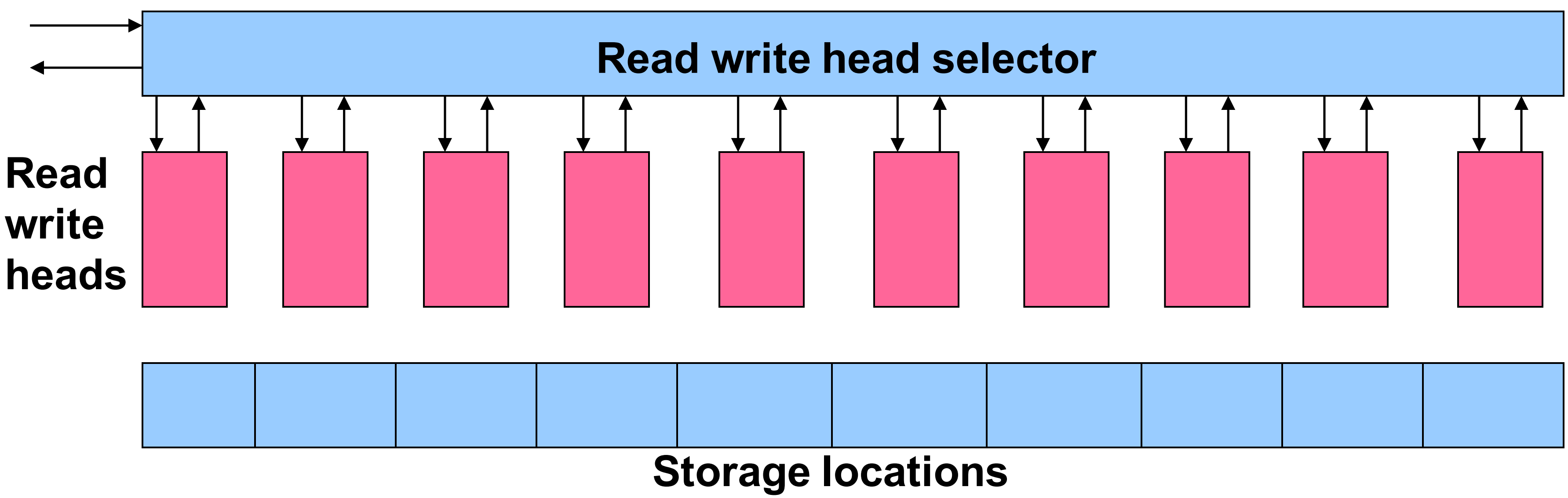
Access Methods contd.,

- Direct access
  - Also referred as semi random access memory
  - Access time is variable
  - The track is accessed randomly but access within each track is serial
  - Access is accomplished by general access to reach a general vicinity plus sequential searching, counting, waiting to reach the final location.
  - Ex: Magnetic Disk

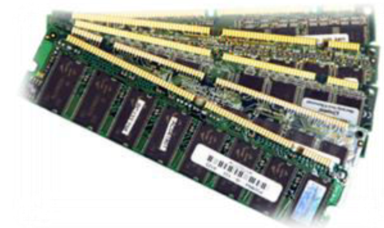


Access methods contd.,

- Random Access
  - Each addressable location in memory has unique, physically wired – in addressing mechanism
  - Time to access a location is independent of the sequences of prior access and is constant
  - Main memory systems are a random access
  - Storage locations can be accessed in any order
  - Semi conductor memories



Access Methods contd.,



- Associate Access
  - Word is retrieved based on portion of its contents rather than its address
  - This enables one to make a comparison of desired bit locations within a word for specific match
  - Has own addressing mechanism
  - Retrieval time is constant
  - Access time is independent of location or prior access patterns
  - Cache memories

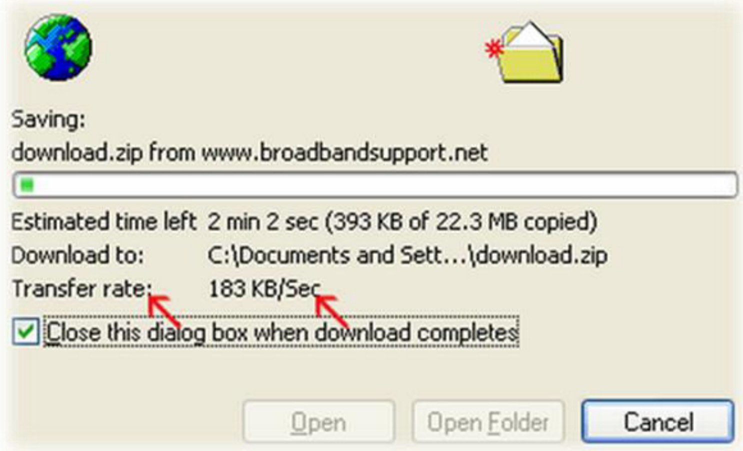
Performance

- Access time
  - The time required to read / write the data from / into desired record
  - Depends on the amount of data to be read / write
  - If the amount data is uniform for all records then the access time is same for all records.
  - Time from the instant that an address is presented to the memory to the instant that data have been stored or made available for use.
- Memory Cycle time
  - Access time + time required before a second access can commence

  - For Random access method ,this memory cycle time is same for all records
  - The sequential access and direct access ,the memory cycle time is different

Performance contd.,

- Transfer rate / Throughput
  - Rate at which the data can be transferred into or out of a memory unit
  - Random access memory
    - 1/cycle time
  - Non-Random access memory
    - $T_n = T_a + (N/R)$, where
      - T_n – average time to read or write N bits
      - T_a – average access time
      - N – Number of bits
      - R – Transfer rate, in bits per second (BPS)



Physical type

Semiconductor

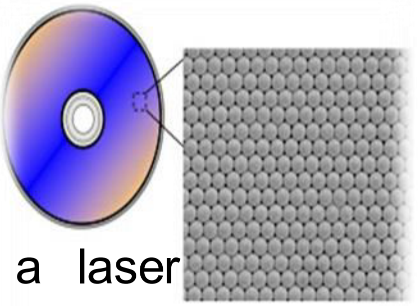
Semiconductor memory uses semiconductor-based integrated circuits to store information.

Magnetic surface

Magnetic storage uses different patterns of magnetization on a magnetically coated surface to store information.

Example:

Magnetic disk, Floppy disk, Hard disk drive



Optical

The typical optical disc, stores information in deformities on the surface of a circular disc and reads this information by illuminating the surface with a laser diode and observing the reflection.

Physical characteristics

- Volatile memory
  - Information decays naturally or lost when electrical power is switched off
- Non-volatile memory
  - Once recorded is retained until deliberately changed
  - No electrical power is needed to retain information
  - Magnetic surface memories
- Semiconductor memories may be either volatile or non-volatile
  - A type of non-volatile semiconductor memory known as flash memory
  - A type of volatile semiconductor memory is random access memory
- Non-erasable memory
  - Cannot be altered, except by destroying the storage unit (ROM)
  - A practical non-erasable memory must also be non-volatile
  - Ex: CD-R, Flash Memories
- Erasable memory
  - Erase the stored information by writing new information
  - Ex: Magnetic storage is erasable

Organization

- Physical arrangement of bits to form words
- 2 types
  - 1 dimensional
  - 2 dimensional

Memory Organization

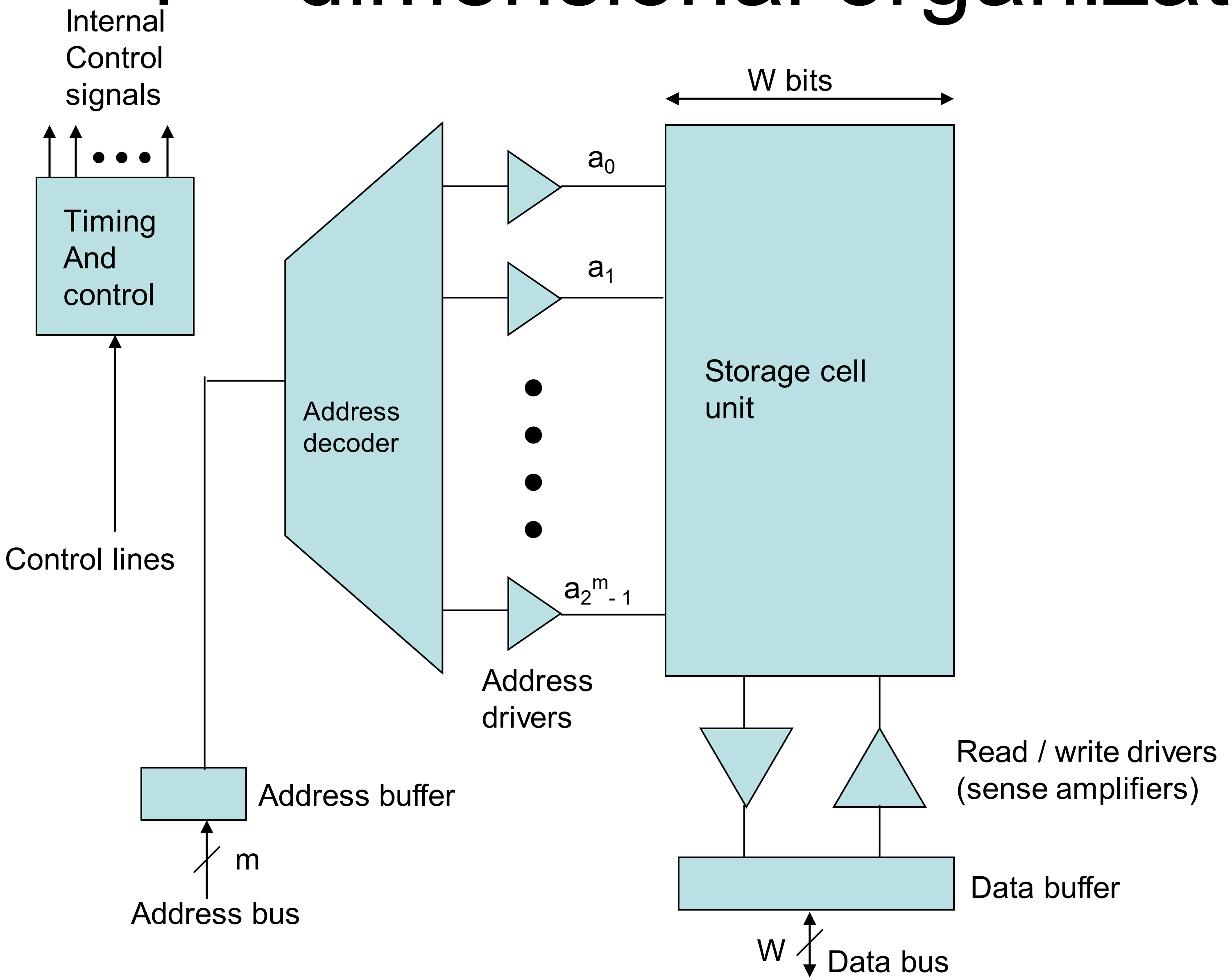
- Basic element = memory cell
- Properties of Memory cell:

  - They exhibit two stable states, which can be used to represent binary 1 and 0.

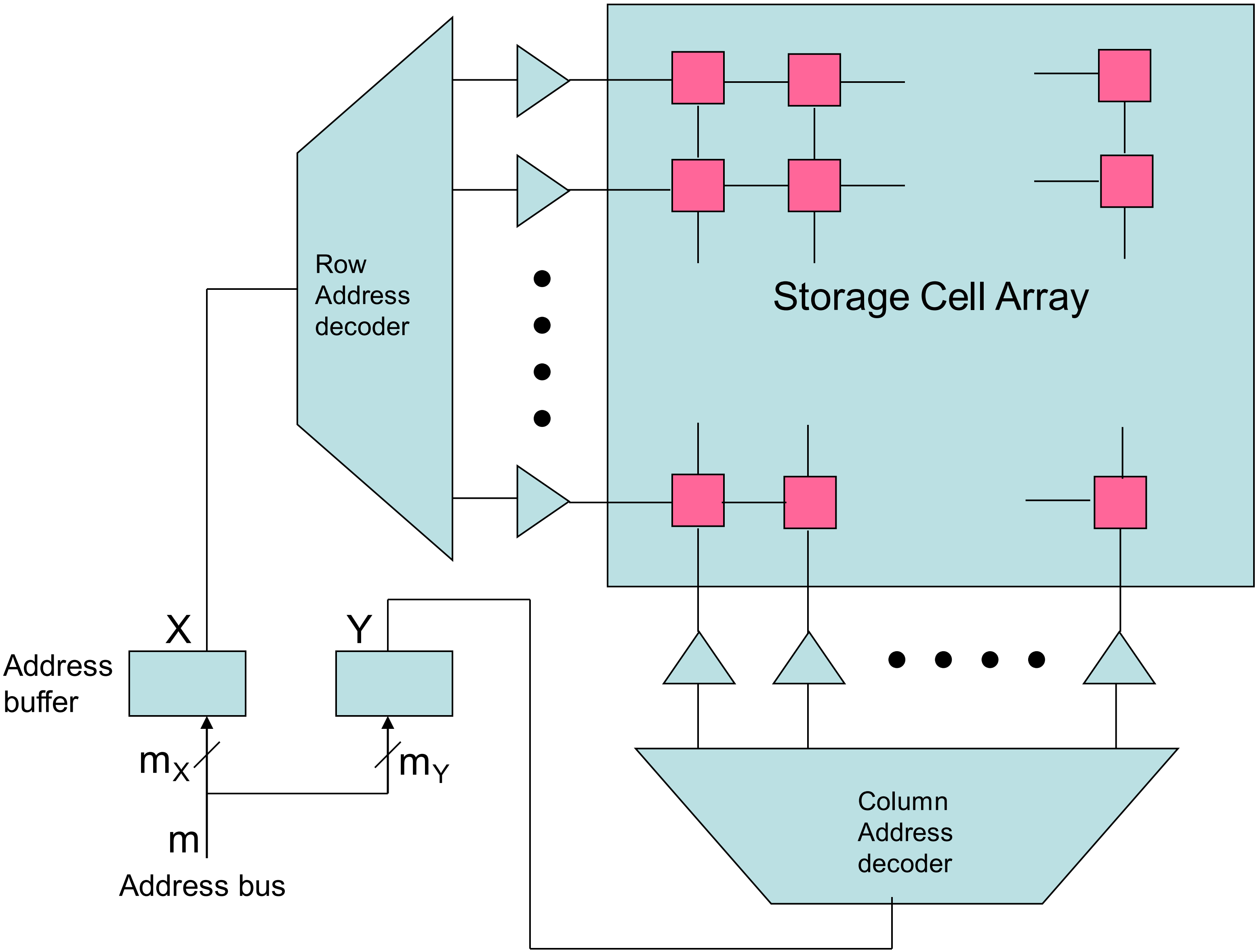
  - They are capable of being written into (atleast once) to set the state.

  - They are capable of being read to sense the state.

1 – dimensional organization



2 – dimensional organization



Byte Storage Methods

- Big-Endian
  - Assigns MSB to least address and LSB to highest address
  - Ex: 0 × DEADBEEF

Memory Location	**Value**
Base Address + 0	DE
Base Address + 1	AD
Base Address + 2	BE
Base Address + 3	EF

Byte Storage Methods contd.,

- Little Endian
  - Assigns MSB to highest address and LSB to least address
  - Ex: 0 × DEADBEEF

Memory Location	**Value**
Base Address + 0	EF
Base Address + 1	BE
Base Address + 2	AD
Base Address + 3	DE

Byte Storage Methods contd.,

- Little Endian
  - Intel × 86 family
  - Digital equipment corporation architectures (PDP – 11, VAX, Alpha)
- Big Endian
  - Sun SPARC
  - IBM 360 / 370
  - Motorola 68000
  - Motorola 88000
- Bi-Endian
  - Power PC
  - MIPS
  - Intel’s 64 IA - 64

Example

- **Example**: Show the contents of memory at word address 24 if that word holds the number given by 122E 5F01H in both the big-endian and the little-endian schemes?

Big Endian

	MSB	---------->		LSB
	24	25	26	27
Word 24	12	2E	5F	01

Little Endian

	MSB	---------->		LSB
	27	26	25	24
Word 24	12	2E	5F	01

References

- William Stallings “Computer Organization and architecture” Prentice Hall, 7th edition, 2006