Memory Elements

Digital Logic Design (CC131)

BSCS 2nd Semester Spring-2026

By

Syed Zulfiqar Ali Shah,

Associate Professor (CS) / HoD

Govt. College of Management Sciences, Kohat

www.linkedin.com/in/syedzulfiqaralishah

Memory

- Memory is the area of a digital device which stores data.
- This area is heavily used in all kinds of computations, so it is accessed millions of times per second.
- Hence speed and accuracy is very important in data storage.

Units of Binary Data

- **Bit**: It is the smallest unit of binary data.
- **Byte**: It is an 8-bit unit.
- **Nibble**: It is a 4-bit unit.
- **Word**: A group of bits or bytes that acts as a single entity that can be stored on one memory location.

Basic Memory Array

- **Cell**: A cell is a storage element in memory that can retain either a 1 or a 0.

- Memories are made up of arrays of cells. These cells can be organized in arrays in different ways.

- For example, 64-cell array can be organized in three different ways.

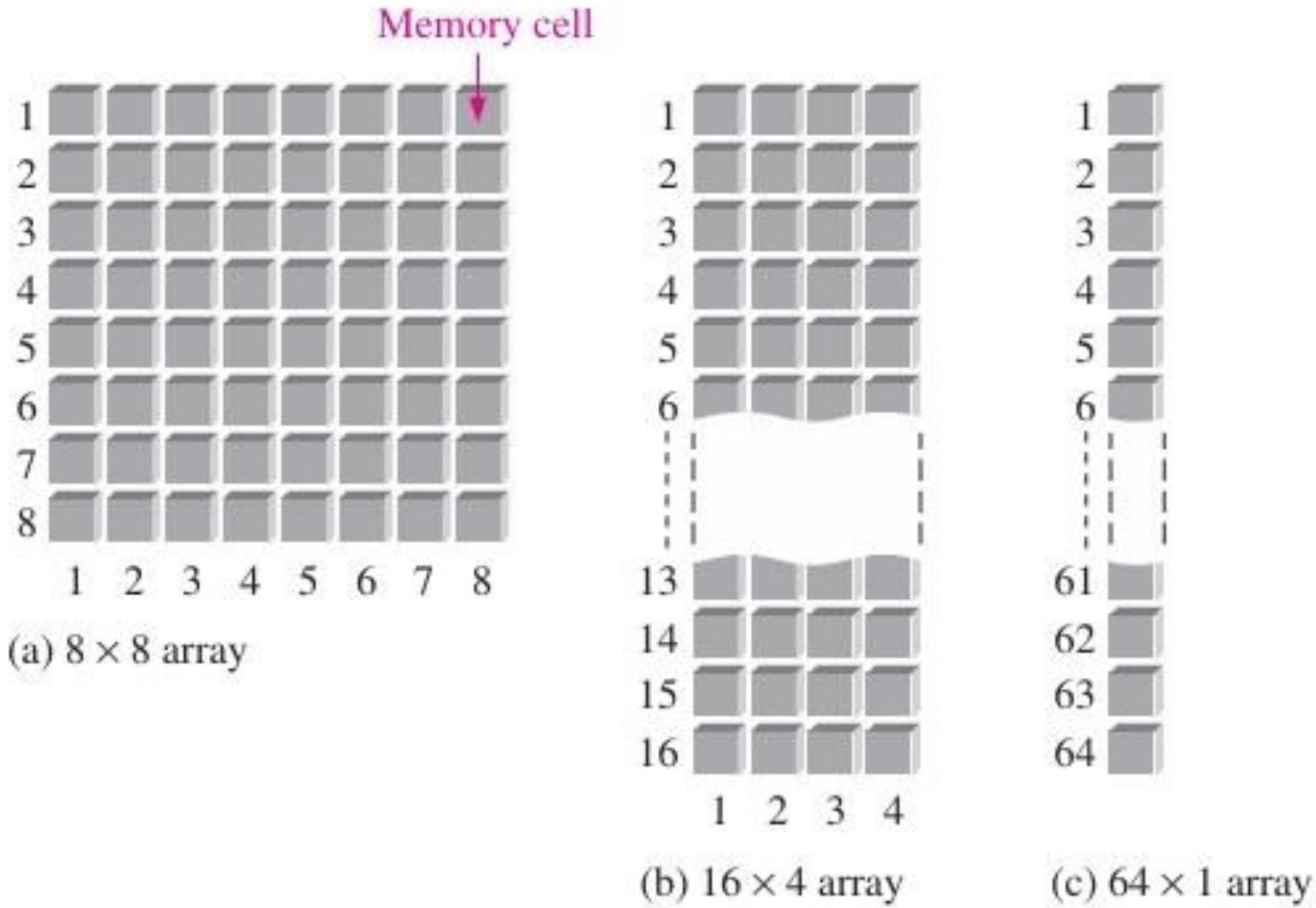
  8 X 8 – array,
  16 X 4 – array,
  64 X 1 – array

Basic Memory Array

- A memory is identified by the number of words it can store times the word size.

- For example,
  16K X 8 memory can store 16384 words
  of eight bits each.

Basic Memory Array

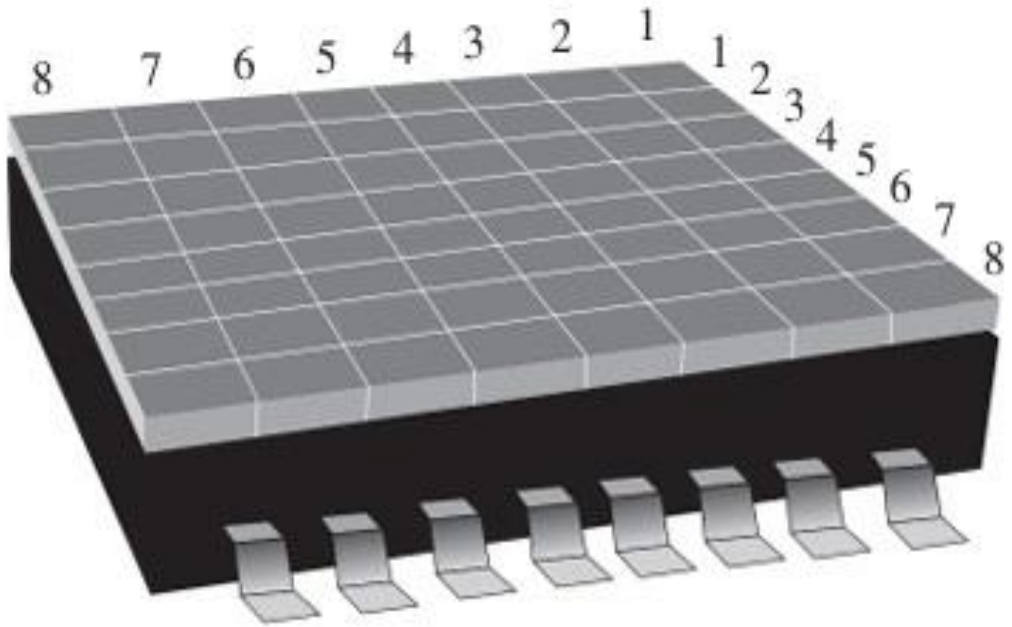


A 64-cell memory array organized in three different ways.

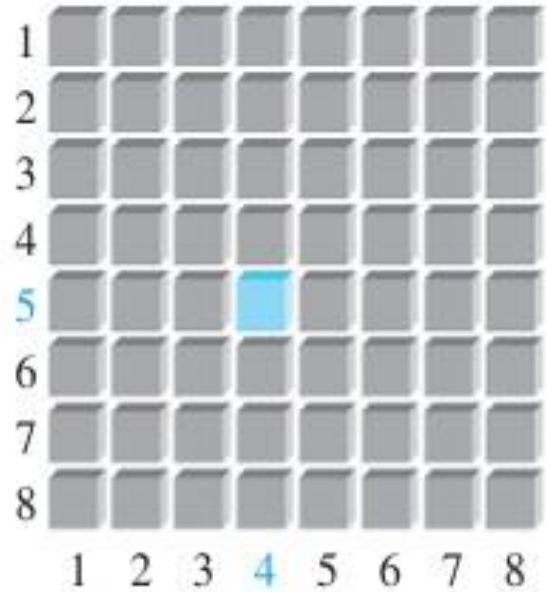
Memory Address & Capacity

- The location of a unit of data in a memory array is called its **Address.**
- The address of a bit in 2-dimensional array is specified by the row and column.
- The address of a byte is specified only by the row.
- The address depends on how the memory is organized into units of data.
- In personal computers the RAM is organized into Bytes.

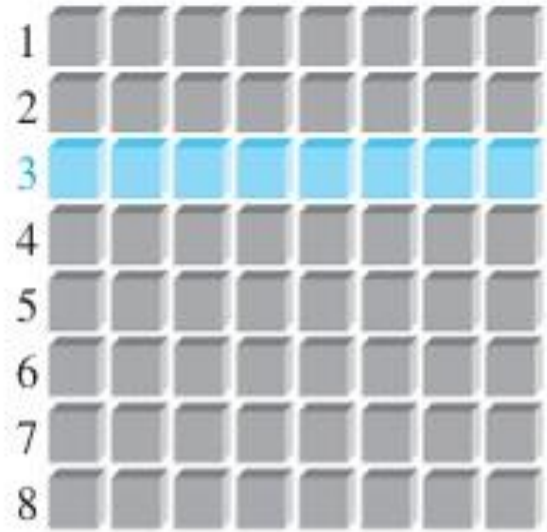
Memory Address & Capacity



(a) Physical structure of 64-bit memory.



(b) The address of the blue bit is row 5, column 4.

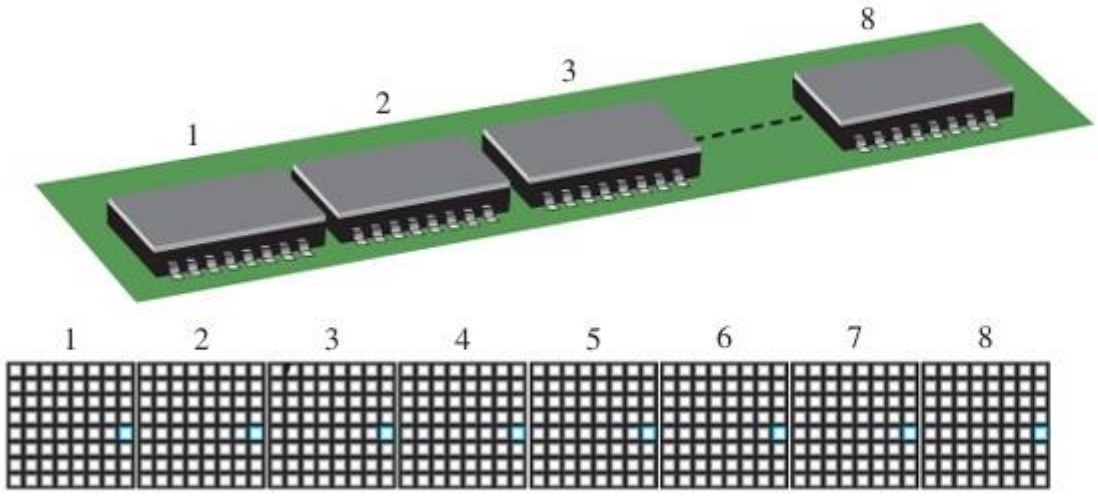


(c) The address of the blue byte is row 3.

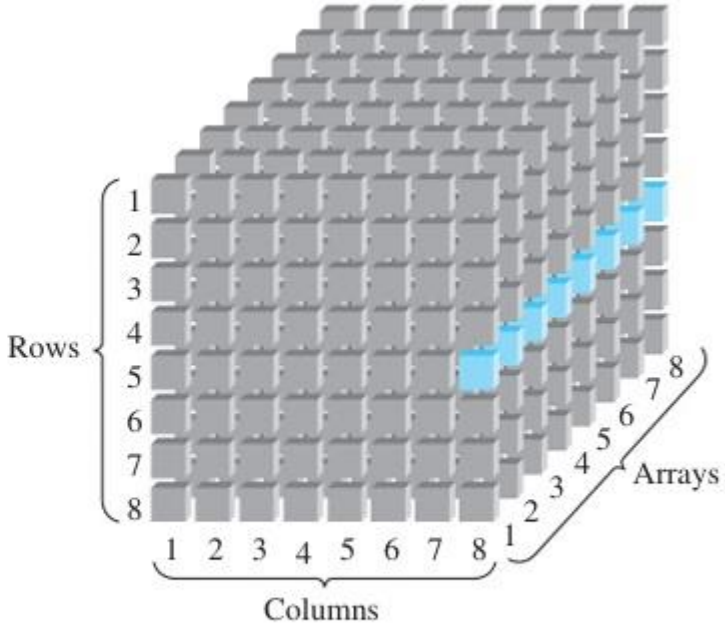
Examples of memory address in a 2-dimensional memory array.

Memory Address & Capacity

- **Capacity**: The total number of data units that can be stored in memory is called its Capacity.



(a) The 8 × 8 bit array expanded to a 64 × 8 bit array. This array forms a memory module.



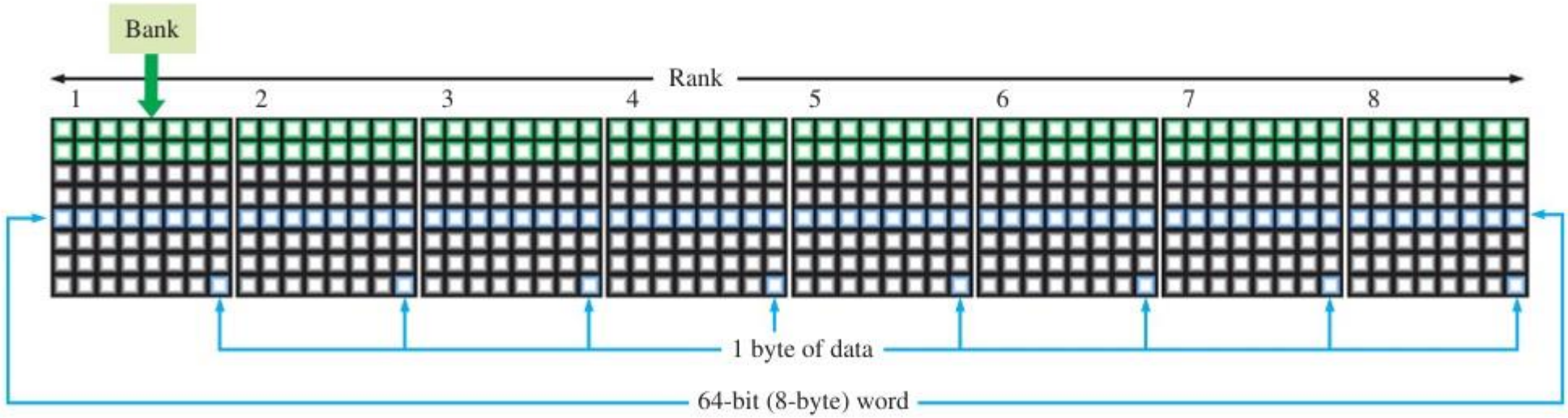
(b) The address of the blue byte is row 5, column 8.

Example of memory address in an expanded (multiple) array.

Memory Bank & Rank

- A **bank** is a section of memory within a single memory array (chip). A memory chip may have one or more banks.

- A **rank** is a group of chips that make up a memory module that stores data in units such as words or bytes.



Simple illustration of memory bank and memory rank.

Basic Memory Operations

- **Addressing** is the process of accessing a specified location in memory.
- Memory is accessed for two reasons either storing something or copying something that is already stored there.
- The **write** operation puts data into a specified address in the memory.
- The **read** operation copies data out of a specified address in the memory.
- The addressing operation, which is part of both the write and the read operations, selects the specified memory address.

Basic Memory Operations

- Data units go into the memory during a write operation and come out of the memory during a read operation on a set of lines called the data bus.
- In case of byte-organized memories, the data bus has at least eight lines so that all eight bits in a selected address are transferred in parallel.
- For a write or a read operation, an address is selected by placing a binary code representing the desired address on a set of lines called the address bus.

Basic Memory Operations

- The address code is decoded internally, and the appropriate address is selected.

- The number of lines in the address bus depends on the capacity of the memory.

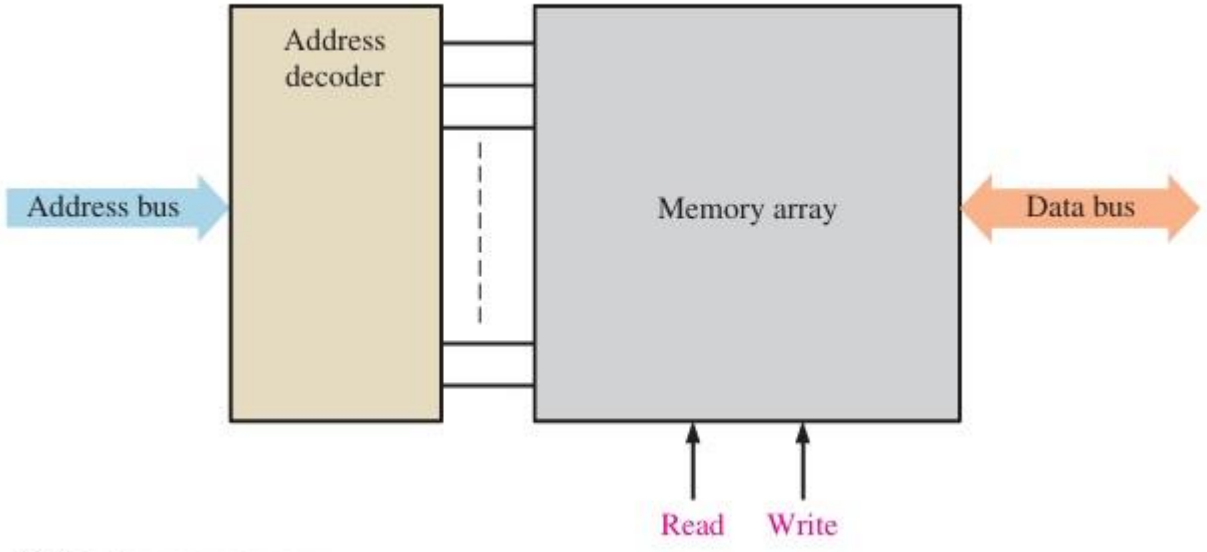
- For example, a 15-bit address code can select 32,768 locations (215) in the memory,

  a 16-bit address code can select 65,536 locations (216) in the memory, and so on.

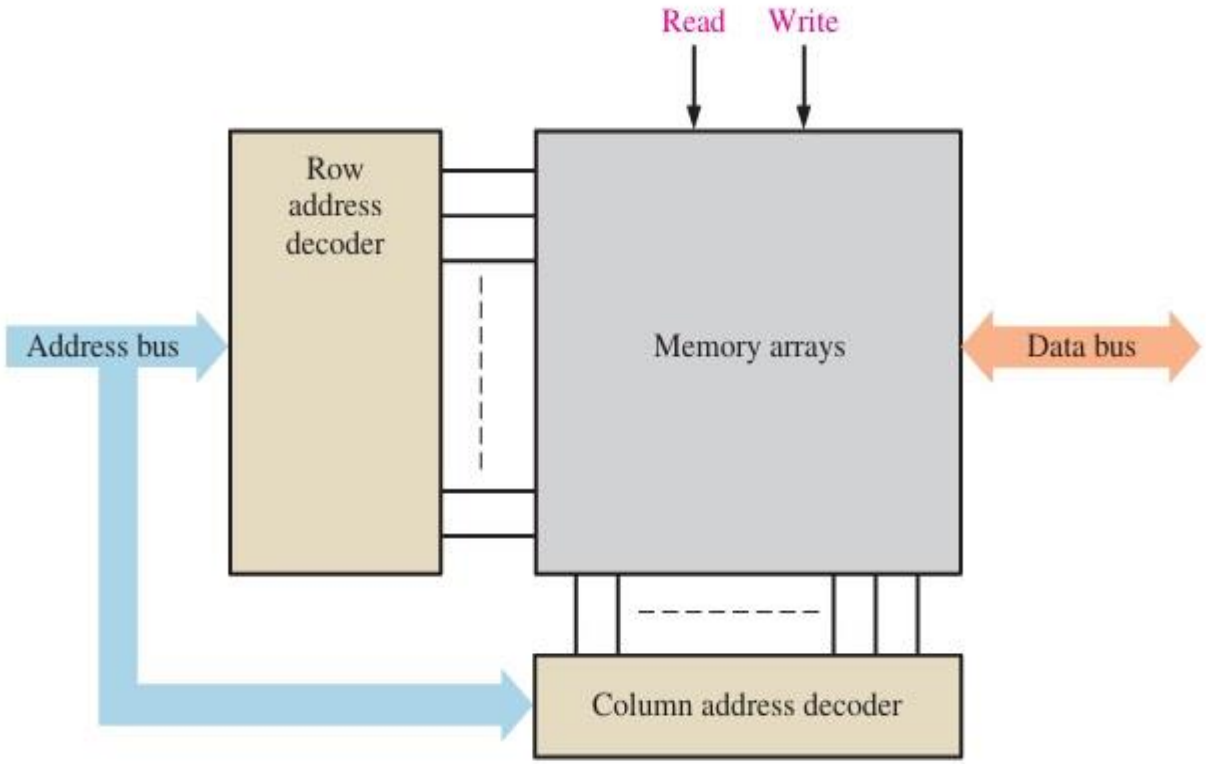
WRITE OPERATION

- To store a byte of data in the memory, a code held in the address register is placed on the address bus.
- Once the address code is on the bus, the address decoder decodes the address and selects the specified location in the memory.
- The memory then gets a write command, and the data byte held in the data register is placed on the data bus and stored in the selected memory address

WRITE OPERATION

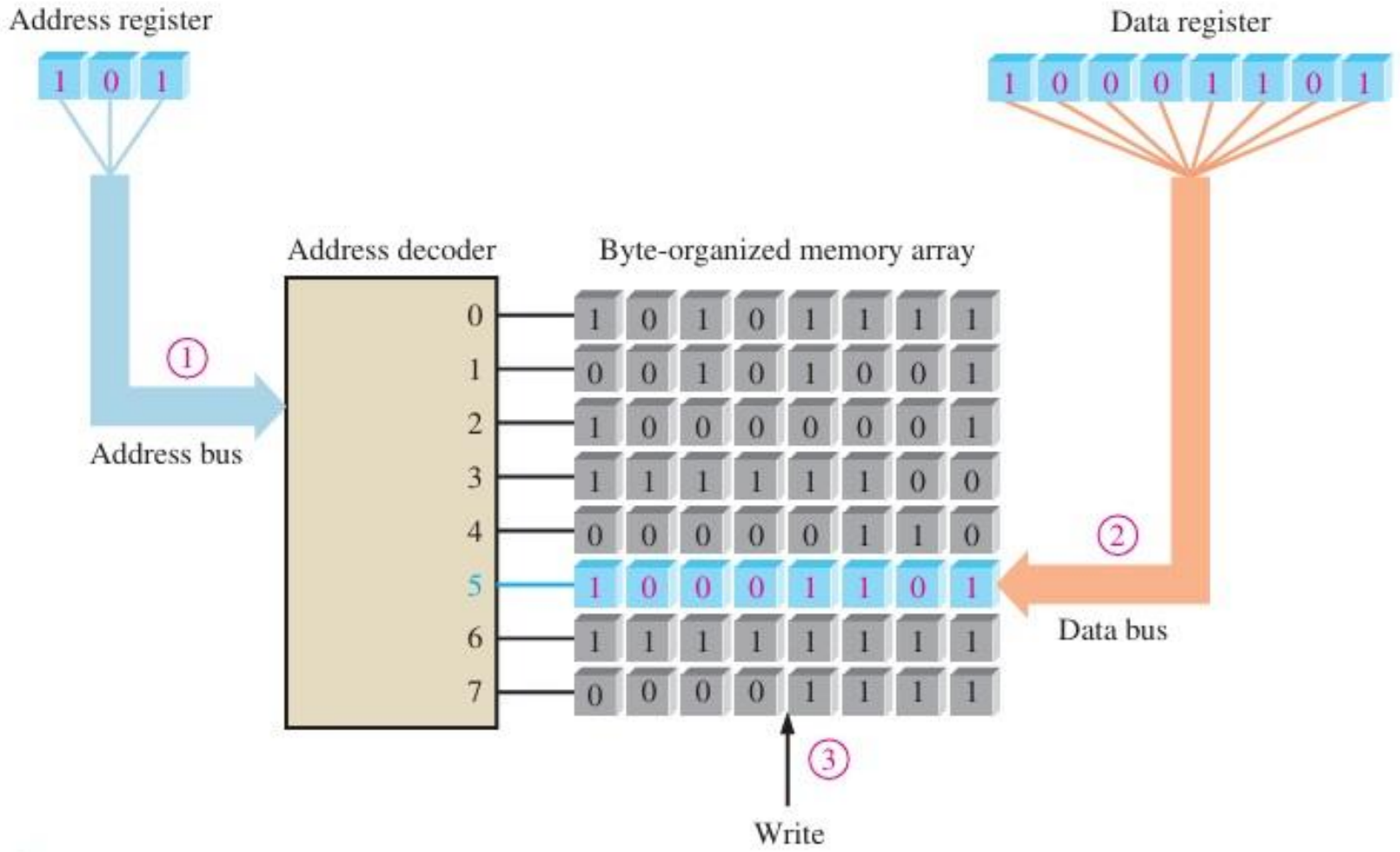


(a) Single-array memory



(b) Multiple-array memory

WRITE OPERATION

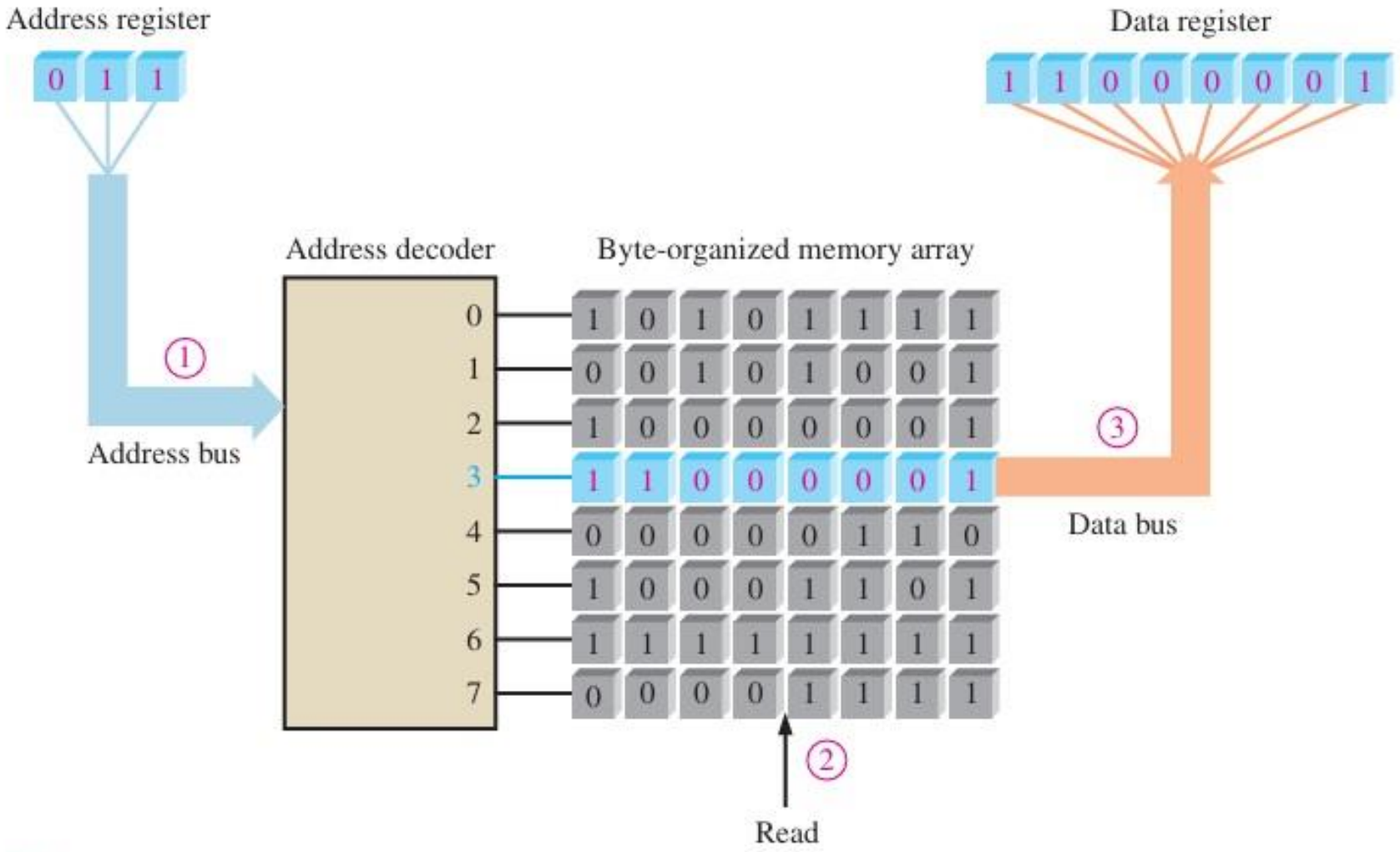


1. Address code 101 is placed on the address bus and address 5 is selected.
2. Data byte is placed on the data bus.
3. Write command causes the data byte to be stored in address 5, replacing previous data.

READ OPERATION

- A code held in the address register is placed on the address bus.
- Once the address code is on the bus, the address decoder decodes the address and selects the specified location in the memory.
- The memory then gets a read command, and a “copy” of the data byte that is stored in the selected memory address is placed on the data bus and loaded into the data register, thus completing the read operation.

READ OPERATION



① Address code 011 is placed on the address bus and address 3 is selected.

② Read command is applied.

③ The contents of address 3 is placed on the data bus and shifted into data register. The contents of address 3 is not erased by the read operation.

RAM & ROM

RAM & ROM

- RAM and ROM both are semiconductor memories.
- Both are regarded as Primary Memory.

Random Access Memory

- RAM (random-access memory) is a type of memory in which all addresses are accessible in an equal amount of time and can be selected in any order for a read or write operation.
- All RAMs have both read and write capability.
- RAMs lose stored data when the power is turned off, they are volatile memories.

Random Access Memory

- When a data unit is written into a given address in the RAM, the data unit previously stored at that address is replaced by the new data unit.

- When a data unit is read from a given address in the RAM, the data unit remains stored and is not erased by the read operation.

RAM Family

- The two major categories of RAM are the
  - Static RAM (SRAM)
  - Dynamic RAM (DRAM)

- SRAMs generally use latches as storage elements and can therefore store data indefinitely as long as dc power is applied.

- DRAMs use capacitors as storage elements and cannot retain data very long without the capacitors being recharged by a process called refreshing.

RAM Family

- Data can be read much faster from SRAMs than from DRAMs.
- DRAMs can store much more data than SRAMs for a given physical size and cost.

- The basic types of SRAM are the asynchronous SRAM and the synchronous SRAM.
- The basic types of DRAM are;

  Fast Page Mode DRAM (FPM DRAM),

  Extended Data Out DRAM (EDO DRAM),

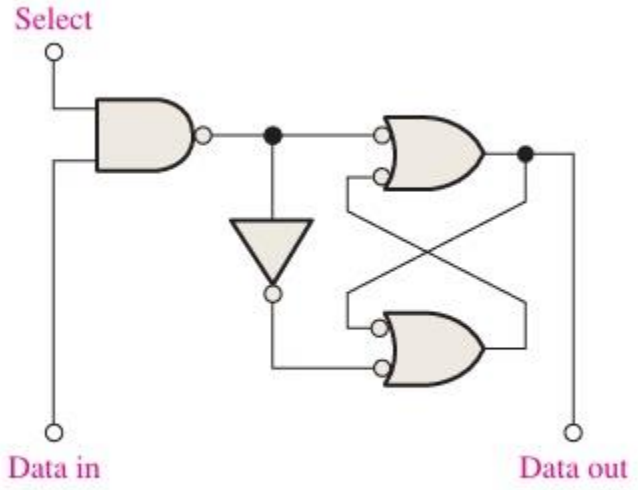
  Burst EDO DRAM (BEDO DRAM),

  Synchronous DRAM (SDRAM).

Static RAM

All SRAMs are characterized by latch memory cells.

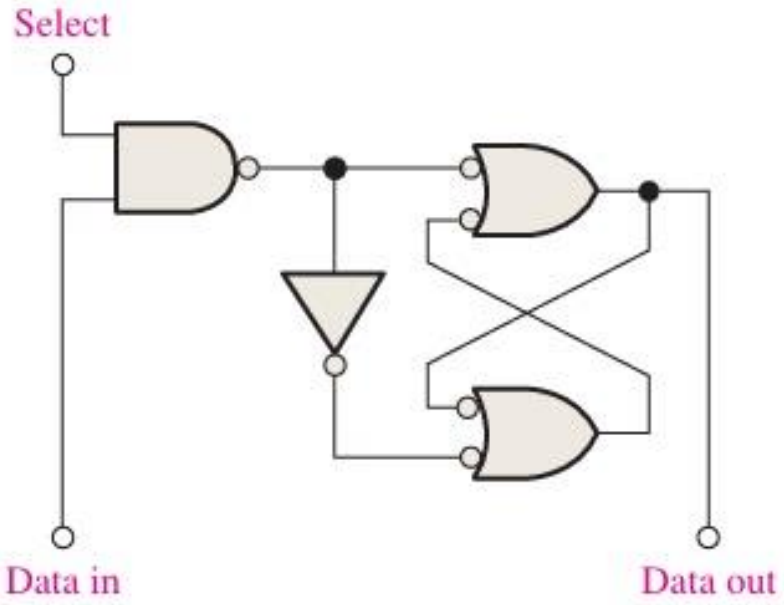
Memory Cell: A Static Memory Cell can retain either a 1 or a 0 state indefinitely as long as DC power is applied to it.
When the power is removed, data is lost.



A typical SRAM latch memory cell.

Static RAM

- The cell is selected by an active level on the Select line.
- Data bit (1 or 0) is written into the cell by placing it on the Data in line.
- A data bit is read by taking it off the Data out line.

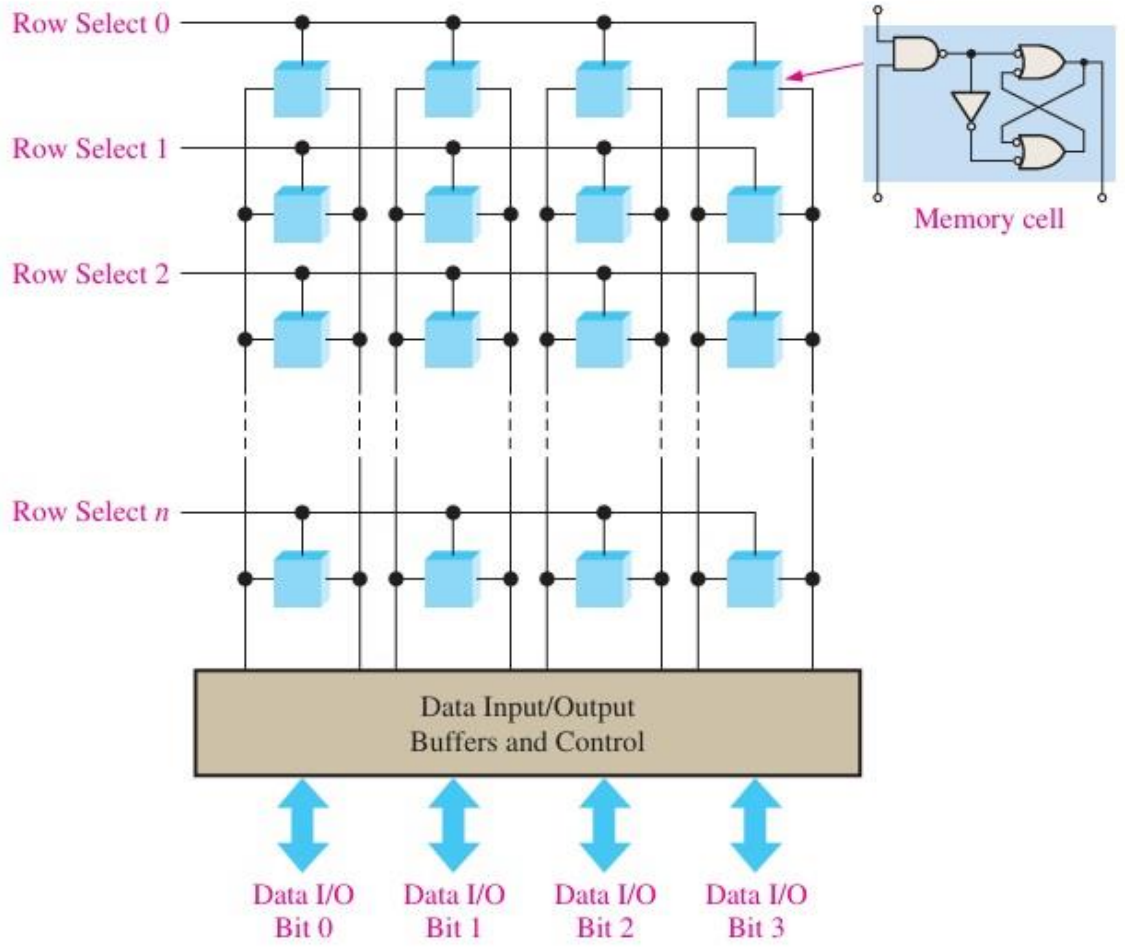


A typical SRAM latch memory cell.

Static Memory Cell Array

- The memory cells in a SRAM are organized in rows and columns.
- All the cells in a row share the same Row Select line.

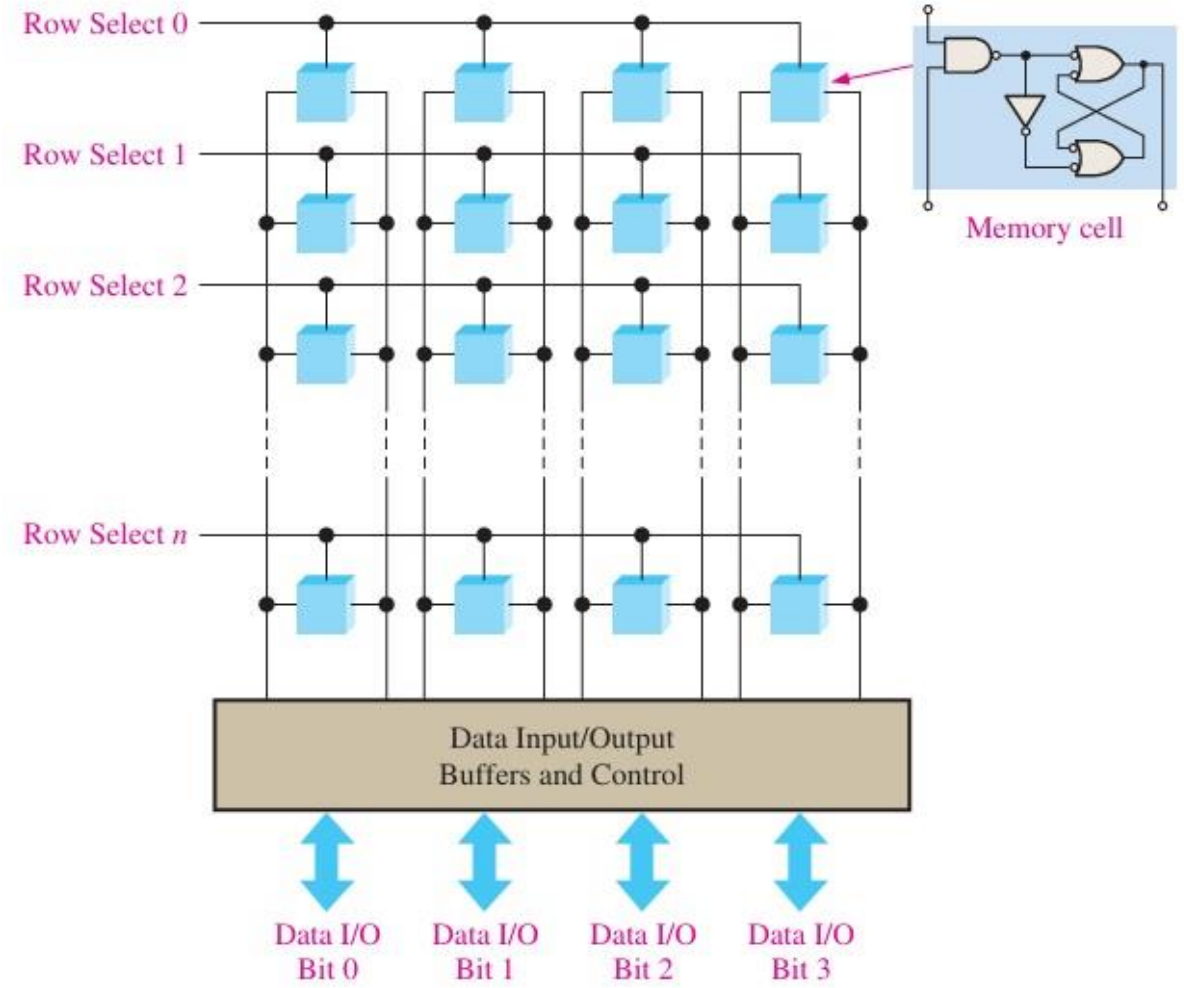
Think of it as a table of bits arranged in **rows and columns**.



What is the size of this memory?

- Number of **rows** = number of memory words.
- Number of **columns** = bits per word.

Static Memory Cell Array



Data is stored in memory cells

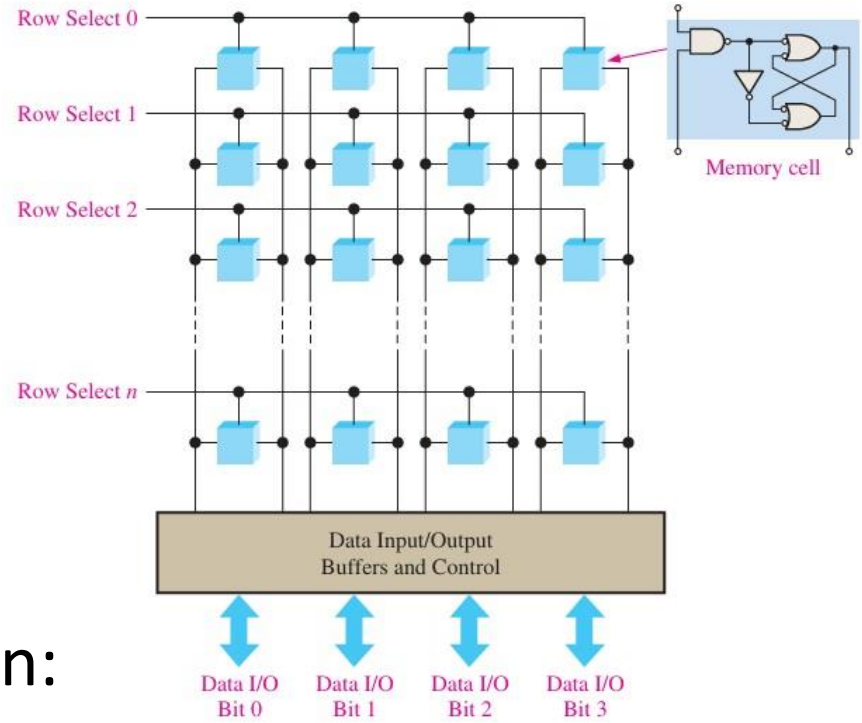
- Each blue block is a **memory cell**.
- One cell stores **1 bit** (either 0 or 1).
- The small circuit on the right is the internal SRAM memory cell that remembers the bit.

Static Memory Cell Array

How Rows are selected

- The signals **Row Select 0, Row Select 1, Row Select 2, ... Row Select n** are called **word lines**.
- Only **one row is activated at a time**.
- Activating a row connects all cells in that row to the column lines.

Static Memory Cell Array



Columns carry the data

- The vertical lines are **bit lines**.

- Each column corresponds to one data bit position:
  - Data I/O Bit 0
  - Data I/O Bit 1
  - Data I/O Bit 2
  - Data I/O Bit 3

Static Memory Cell Array

Reading data

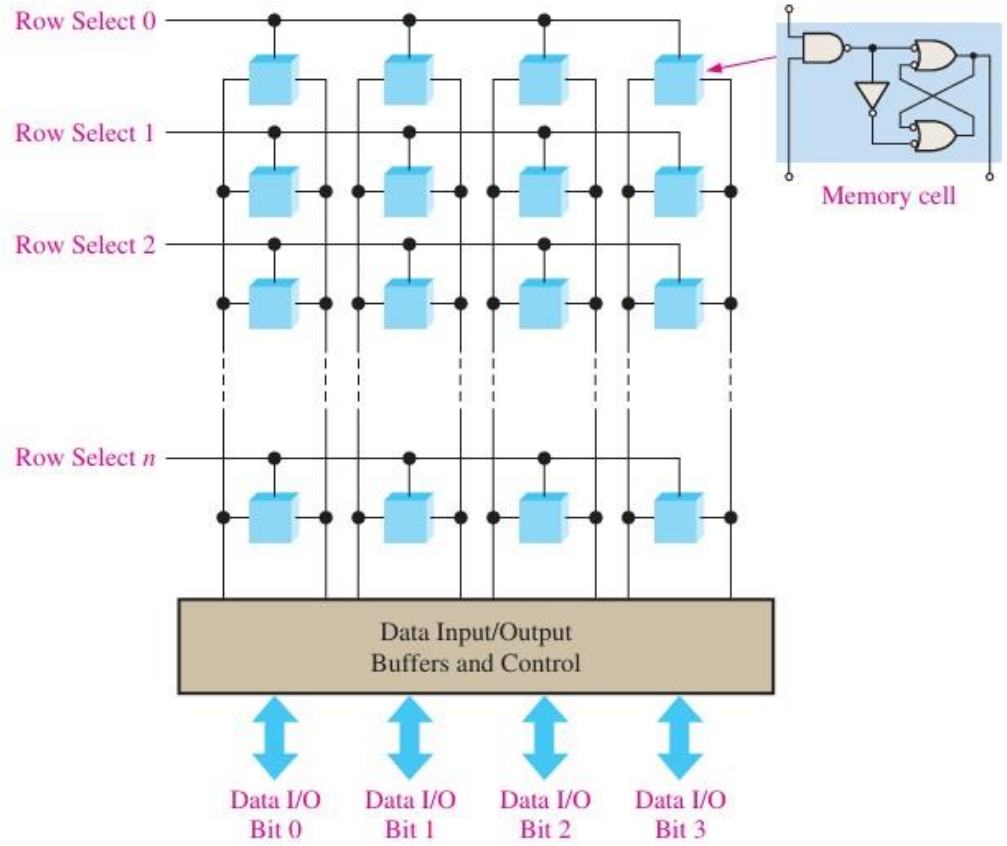
- Suppose **Row Select 1** becomes active.
- The four cells in that row place their stored values onto the four bit lines.
- The **Data Input/Output Buffers and Control** circuit collects these bits and sends them out.
- Example:
  - Cell values in Row 1 = **1 0 1 1**
  - Output = **1011**

Static Memory Cell Array

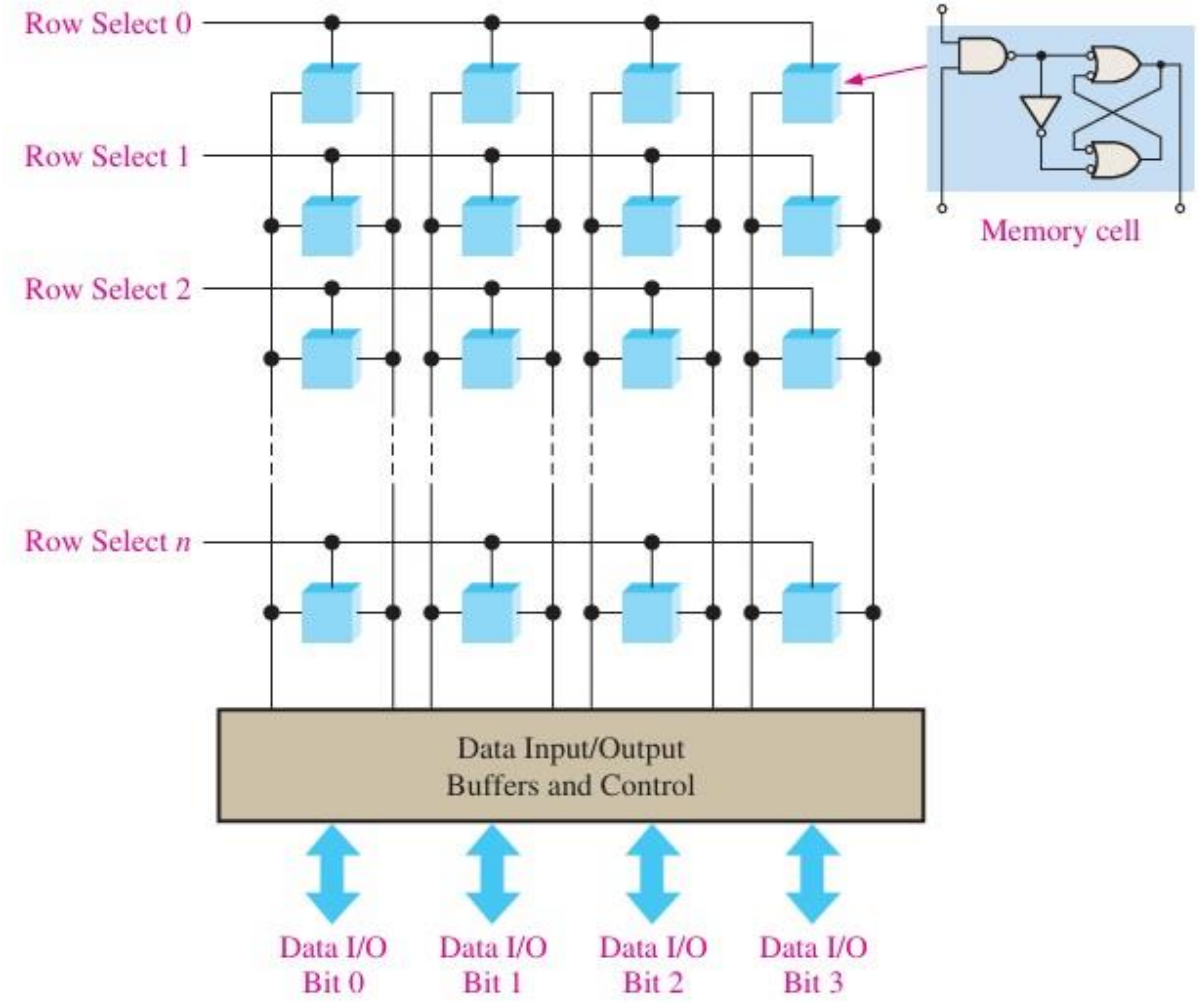
Writing data

- To store a new value:
  - Put the desired bits on the Data I/O lines.
  - Activate the required row.
  - The selected cells store the new bits.

- Example:
  - Write **0110** into Row 2.
  - Activate **Row Select 2**.
  - The four cells of Row 2 become **0, 1, 1, 0**.



Static Memory Cell Array



One-Line Summary

A row select signal chooses which memory word to access, and the column lines carry the bits of that word to or from the Data I/O circuitry.

Dynamic RAM Memory

- Dynamic RAM is a type of semiconductor memory which uses a single capacitor and a single transistor per storage cell.
- The capacitor stores electrical charge while transistor acts as a switch to access the capacitor.
- When capacitor is charged, a 1 is stored. When capacitor is discharged, a 0 is stored in the cell.
- The capacitor cannot hold charge forever and the charge leaks away, causing data to be lost.

Dynamic RAM Memory

- In order to retain data, charge in the capacitor must be maintained at the required level. This is done by recharging the capacitor, which is called refresh operation.
- During a refresh operation, first the stored data is read, then capacitor is recharged and finally the data is rewritten on the cell.
- This refreshing takes place 16 times per second normally.
- Due to this continuous refreshing, this memory is called Dynamic.

Address Multiplexing

- Address multiplexing is a technique used in DRAM in which the memory address is divided into row and column addresses and transmitted over the same address lines at different times.

- This done to reduce the number of address pins required on the memory chip.

- Instead of sending the entire memory address at once, the address is divided into two parts:
  1. Row Address
  2. Column Address

Address Multiplexing

Why Is It Needed?

- Suppose a DRAM chip requires a **16-bit address**.
- Without multiplexing:
- 16 separate address pins would be needed.

- With address multiplexing:
- The 16-bit address is split into:
  - 8-bit row address
  - 8-bit column address
- Only **8 address pins** are required.

This reduces Chip size, Pin count, Cost

FAST PAGE MODE DRAM

Fast Page Mode (FPM) DRAM is an improved type of DRAM that speeds up memory access when multiple data items are located in the **same row (page)**.

The key idea is:
Once a row has been selected, keep it open and access multiple columns within that row without repeatedly supplying the row address.

FAST PAGE MODE DRAM

Why Is It Faster?

In ordinary DRAM:
For every memory access:

- Send Row Address (RAS)
- Send Column Address (CAS)
- Read/Write data
- If several accesses are in the same row, the row address must normally be supplied repeatedly.

FAST PAGE MODE DRAM

In FPM DRAM:

- Send the row address once.
- Activate **RAS** and keep the row open.
- Change only the column address.
- Activate **CAS** for each new column access.
- Since the row-selection step is skipped for subsequent accesses, memory access becomes faster.

SDRAM

- SDRAM (Synchronous Dynamic Random Access Memory) is a type of DRAM whose operations are synchronized with a system clock.

- This synchronization allows predictable timing, pipelined operations, and burst data transfers, resulting in higher performance than earlier asynchronous DRAM technologies.

DDR SDRAM

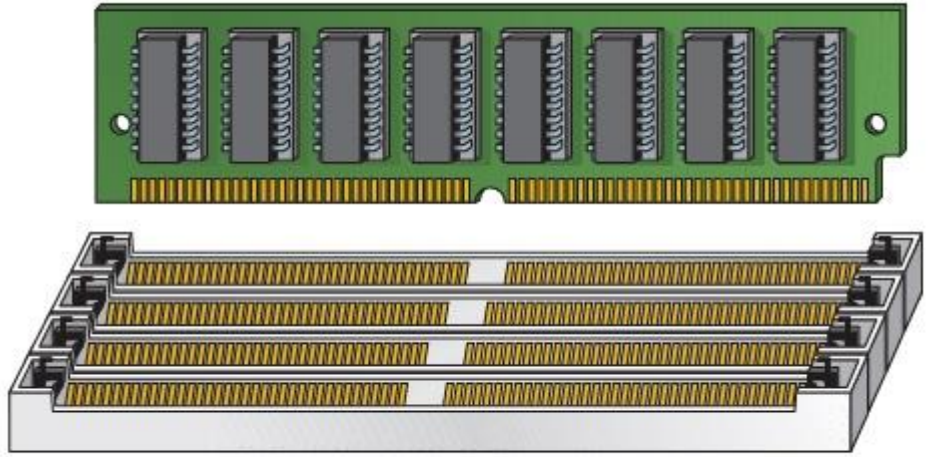
- DDR SDRAM is an improved version of SDRAM that transfers data twice per clock cycle instead of once.
- The term DDR means Double Data Rate.
- In standard SDRAM, data is transferred only on **one edge** of the clock signal (usually the rising edge).
- DDR SDRAM transfers data on both the Rising edge (↑) and the Falling edge (↓) of the same clock signal.

Memory Modules

- SDRAMs are available in modules consisting of multiple memory ICs arranged on a printed circuit board (PCB).
- The most common type of SDRAM memory module is called a DIMM (dual in-line memory module).
- A type of memory module, (now obsolete), is the SIMM (single in-line memory module).
- The SIMM has connection pins on one side of a PCB where the DIMM uses both sides of the board.

Memory Modules

- DDR means double data rate, so a DDR SDRAM transfers two blocks of data for each clock cycle rather than one like a standard SDRAM.
- Three basic types of modules are DDR, DDR2, and DDR3.
- The DDR, DDR2, and DDR3 have transfer data rates of 1600 MB/s, 3200 MB/s, and 6400 MB/s respectively.



A memory module with connectors.

Read Only Memory

- ROM (read-only memory) is a type of memory in which data are stored permanently.
- Data can be read from a ROM, but there is no write operation as in the RAM.
- The ROM, like the RAM, is a random-access memory but the term RAM traditionally means a random-access read/write memory.
- ROMs retain stored data even if power is turned off, they are nonvolatile memories.

PROGRAMMABLE ROMs

- **PROM (Programmable Read-Only Memory)** is a type of ROM that is manufactured **blank** and can be **programmed once by the user** after manufacturing.

- After programming, the stored data becomes permanent and normally cannot be changed.

PROGRAMMABLE ROMs

How Does PROM Store Data?

PROM contains a large number of **fusible links** (fuses).

Initially: All fuses intact

During programming:

- Selected fuses are blown (melted) using a higher-than-normal voltage/current.
- Intact fuse stores a 1 and blown fuse stores a 0.
- The pattern of intact and blown fuses represents the stored binary data.

PROGRAMMABLE ROMs

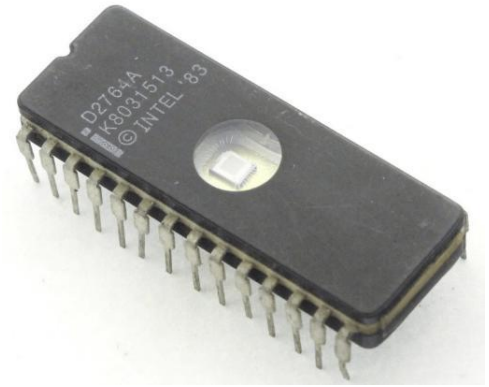
Programming Process

1. Start with a blank PROM.
2. Use a PROM programmer device.
3. Apply special programming voltages.
4. Blow selected fuses.
5. The resulting pattern permanently stores the desired data.
6. Once a fuse is blown, it cannot be restored.

EPROM (Erasable Programmable Read-Only Memory)

EPROM is a type of ROM that:

- Can be programmed by the user.
- Can be erased and reprogrammed multiple times.
- Retains data even when power is removed (**non-volatile**).



How It Stores Data

EPROM uses **floating-gate MOS transistors**.
Electrons trapped on the floating gate represent stored data.

How It Is Erased

The entire chip is erased by exposing it to **ultraviolet (UV) light** for several minutes.

EEPROM (Electrically Erasable Programmable Read-Only Memory)

- **EEPROM** is an improvement over EPROM.

- Like EPROM, it:
  - Is non-volatile.
  - Uses floating-gate transistors.
  - Can be reprogrammed many times.

- But unlike EPROM:
  EEPROM can be erased electrically without removing the chip from the circuit.

EPROM (Erasable Programmable Read-Only Memory)

Advantages of EEPROM

- No UV lamp required.
- No quartz window needed.
- Individual bytes can often be erased and rewritten.
- Can be updated while installed in a system.

This makes EEPROM much more convenient.

FLASH Memory

- **Flash memory** is a type of **non-volatile memory** that can be **electrically erased and reprogrammed**, while retaining its data even when power is turned off.
- It is a descendant of EEPROM technology but is designed to be **faster, denser, and cheaper** for large amounts of storage.
- Due to bulk erasure capability, it is named flash.

FLASH Memory

Flash memory is found in:

- USB flash drives
- Solid-State Drives (SSDs)
- SD and microSD cards
- Smartphones
- Tablets
- Cameras
- Firmware storage

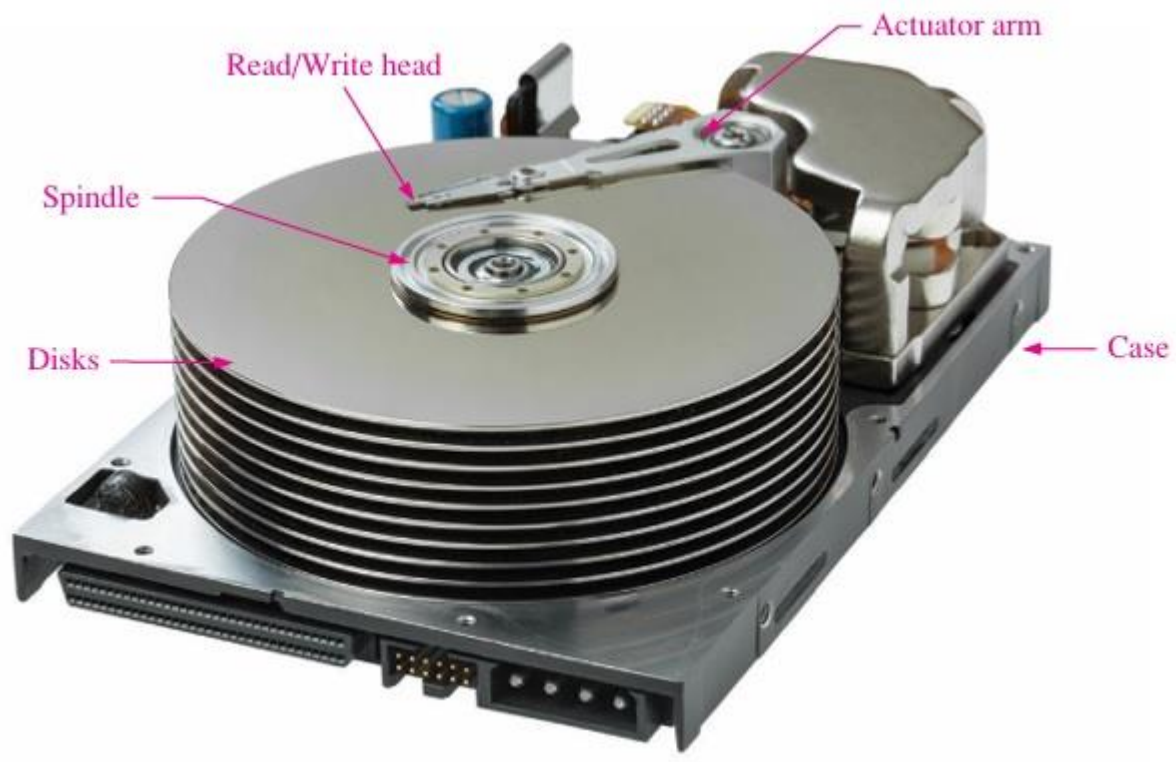
Magnetic Storage

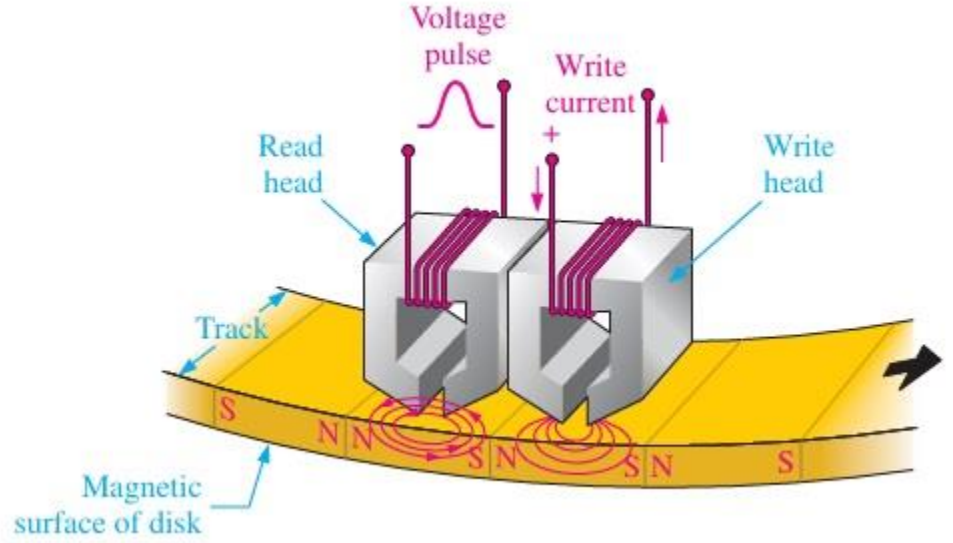
- **Magnetic storage** stores data by magnetizing tiny regions on a magnetic surface.

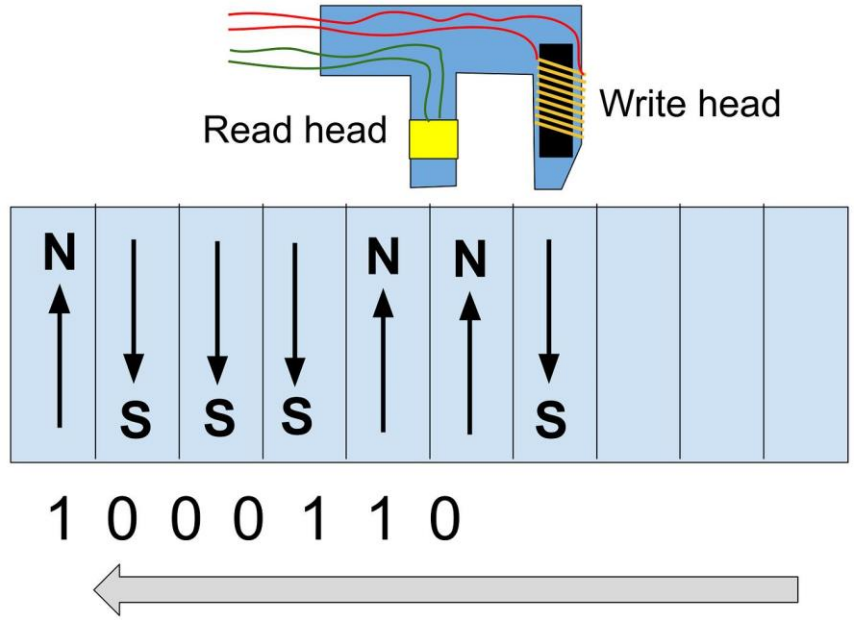
- Examples: Hard Disk Drives (HDDs) , Floppy disks , Magnetic tapes

- The disk surface is coated with a magnetic material that contains billions of tiny magnetic domains. Each domain can be magnetized in one of two directions.
  - Direction A → Binary 1
  - Direction B → Binary 0
- Thus, data is stored as patterns of magnetic polarities rather than electrical charges.

Hard Disk Drive







Hard Disk Drive

A seek operation is the movement of the read/write head to the desired track.

The **seek time** is the average time for this operation to be performed.

The **latency period** is the time it takes for the desired sector to spin under the head once the head is positioned over the desired track.

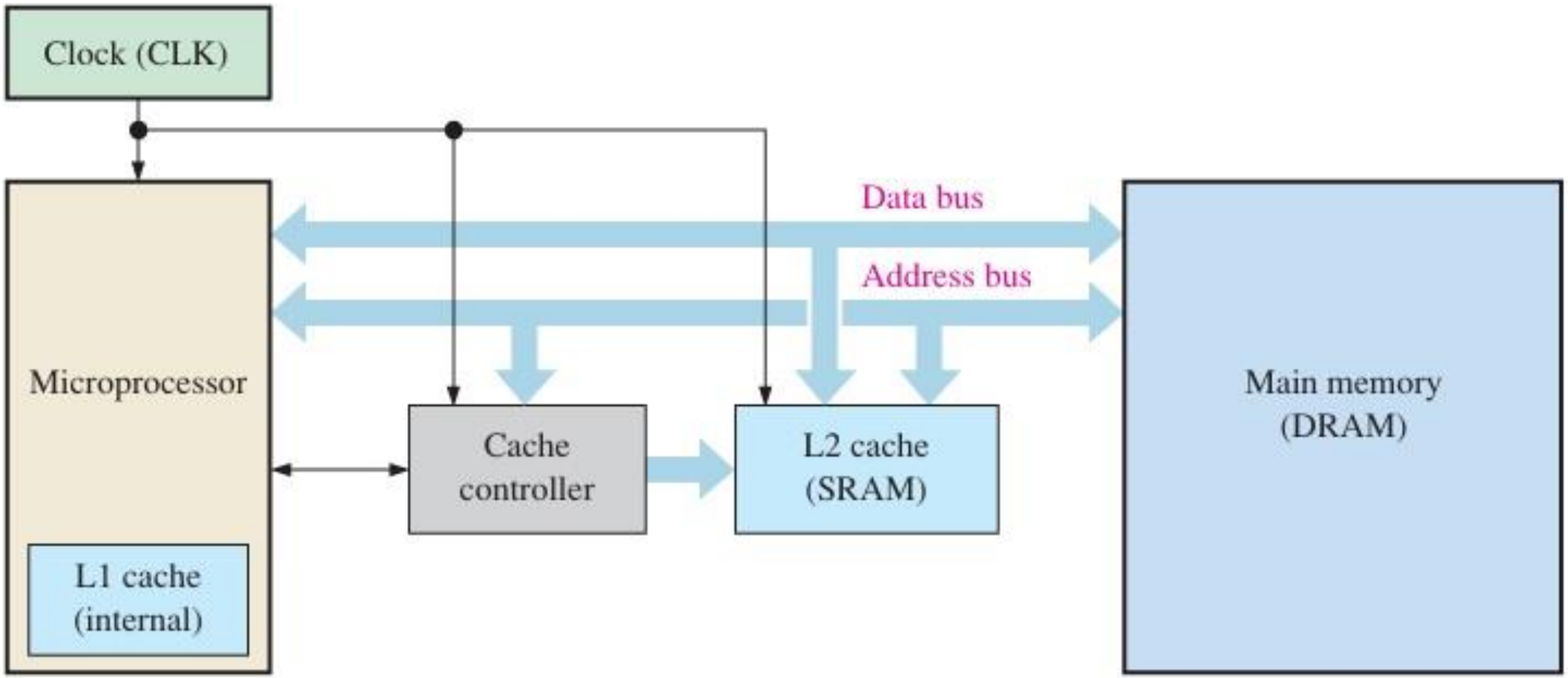
Cache Memory

- Cache memories are built using SRAM.
- Cache memory is a relatively small, high-speed memory that stores the most recently used instructions or data from the larger but slower main memory.
- Basically, the cache controller “guesses” which area of the slower memory the CPU will need next and moves that data to the cache.
- If the cache controller guesses right, the data are immediately available to the microprocessor.
- If the cache controller guesses wrong, the CPU must go to the main memory and wait much longer for the correct instructions or data.

L1 & L2 Cache

- A first-level cache (L1 cache) is usually integrated into the processor chip and has a very limited storage capacity.
- A second-level cache (L2 cache) may also be integrated into the processor or as a separate memory chip.
- L2 has a larger storage capacity than an L1 cache.
- Some systems may have higher-level caches (L3, L4, etc.).

L1 & L2 Cache


Clock (CLK)
Microprocessor
L1 cache
(internal)
Cache
controller
L2 cache
(SRAM)
Data bus
Address bus
Main memory
(DRAM)

Cache Memory Hit/Miss

- A miss is a failed attempt by the processor to read or write a block of data in a given level of memory (such as the cache).
- A miss causes the processor to have to go to a lower level of memory (such as main memory), which has a longer latency.
- A successful attempt to read or write a block of data in a given level of memory is called a hit.
- The hit rate is the percentage of memory accesses that find the requested data in the given level of memory.
- The time required to access the requested information in a given level of memory is called the hit time.

Cache Memory Hit/Miss

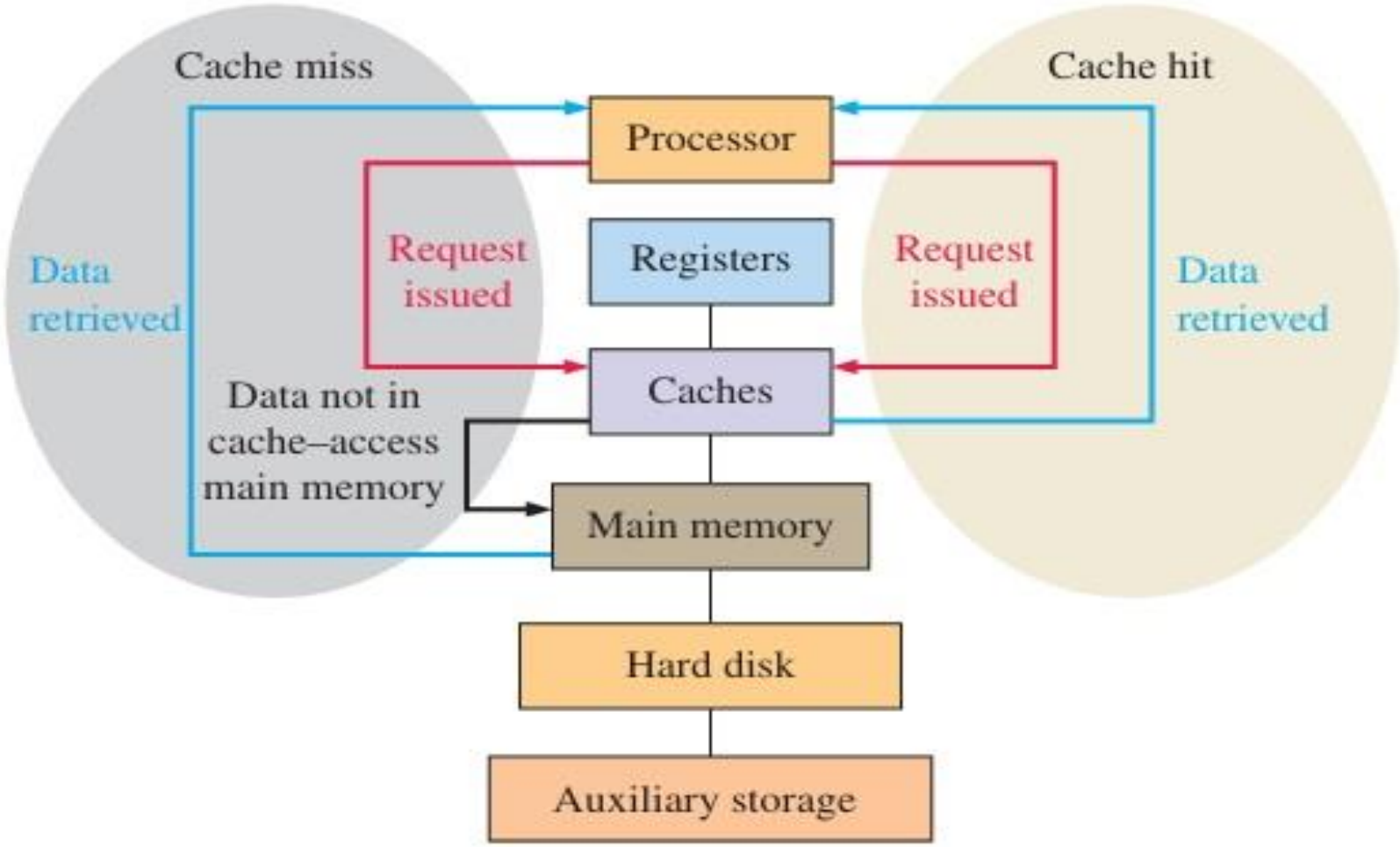
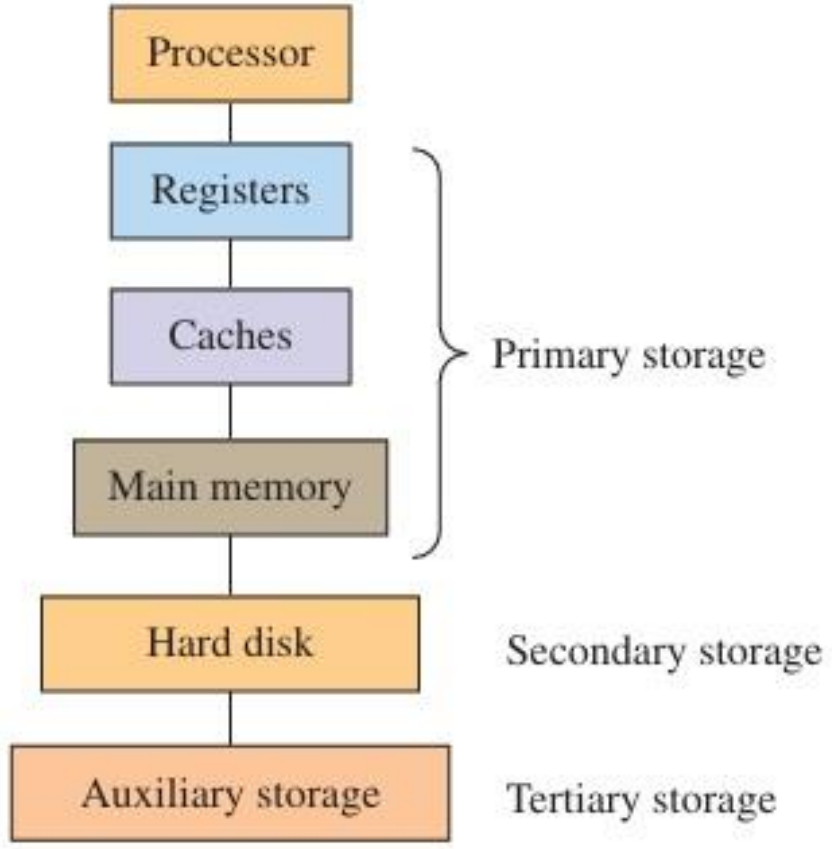


Illustration of a cache hit and a miss.

Memory Hierarchy

- Memory hierarchy is the arrangement of various memory elements within the computer architecture to maximize processing speed and minimize cost.

- Memory can be classified according to its "distance" from the processor in terms of the number of machine cycles or access time required to get data for processing.

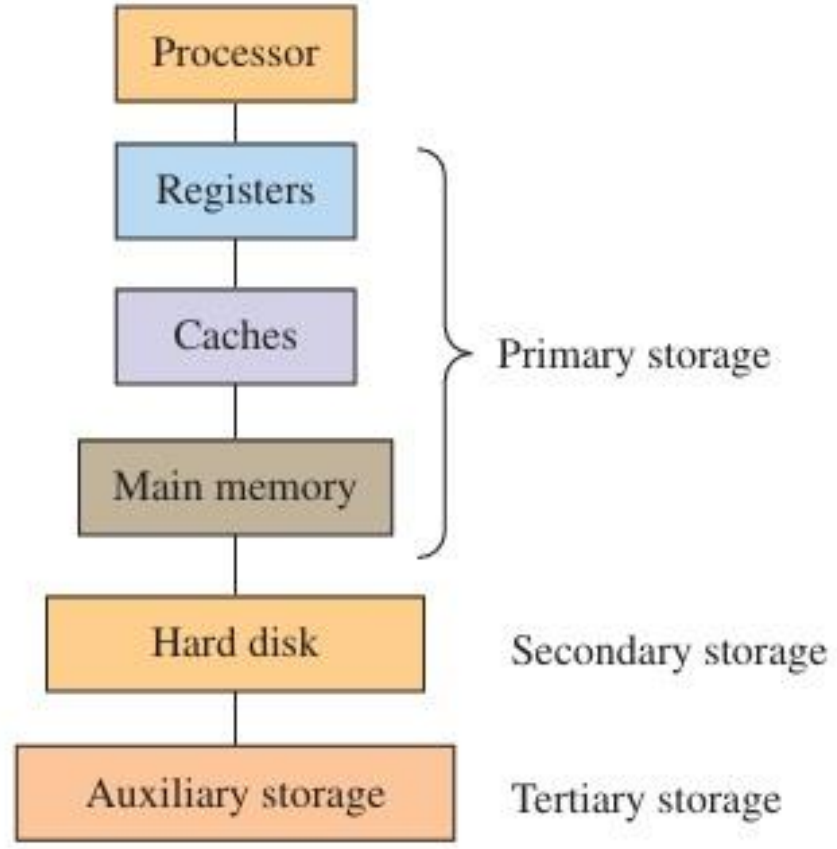


Typical memory hierarchy.

Memory Hierarchy

- Distance is measured in time, not in physical location. Faster memory elements are considered closer to the processor compared to slower types of memory elements.

- Also, the cost per bit is much greater for the memory close to the processor than for the memory that is further from the processor.

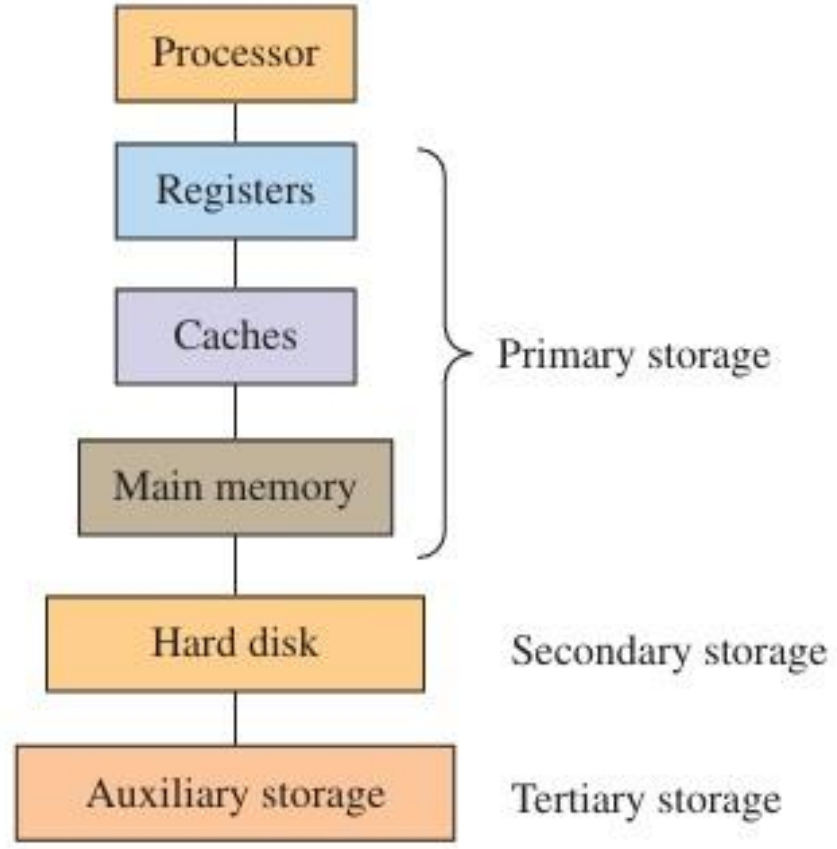


Typical memory hierarchy.

Memory Hierarchy

- A primary distinction between the storage elements in Figure is the time required for the processor to access data and programs.

- This access time is known as **memory latency**.



Typical memory hierarchy.

Thank You