EEE 3101: Digital Logic and Circuits

Memory and Storage

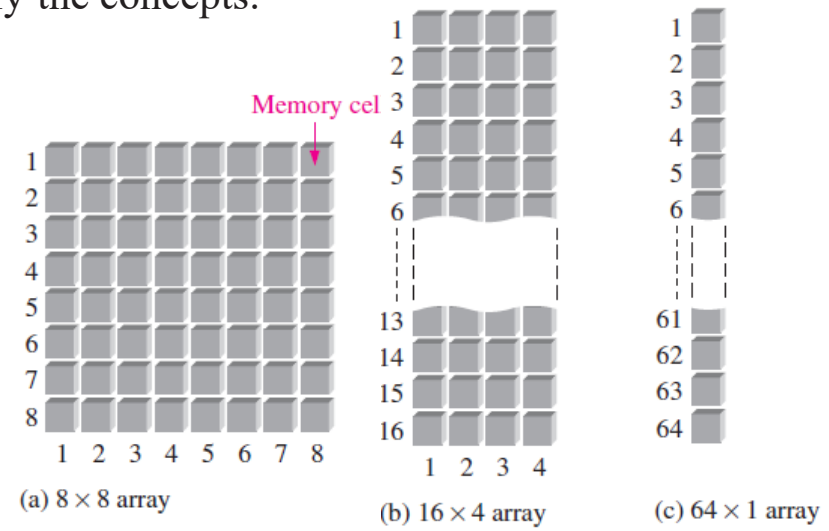
Course Teacher: Nafiz Ahmed Chisty

Head, Department of EEE
Associate Professor, Department of EEE & CoE
Faculty of Engineering
Room# D0105, D Building
Email: chisty@aiub.edu
Website: http://engg.aiub.edu/faculties/nafiz
Website: www.nachisty.com

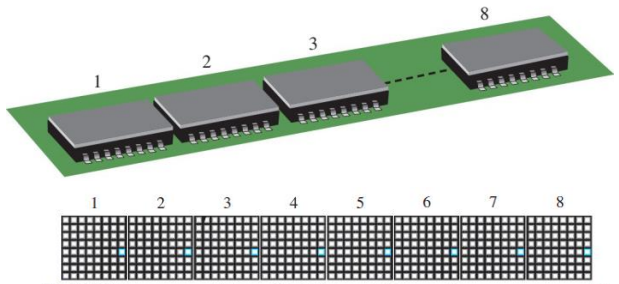
Basics of Semiconductor Memory

Memory is the portion of a computer or other system that stores binary data. In a computer, memory is accessed millions of times per second, so the requirement for speed and accuracy is paramount. Very fast semiconductor memory is available today in modules with several GB (a gigabyte is one billion bytes) of capacity. These large-memory modules use exactly the same operating principles as smaller units, so we will use smaller ones for illustration in this class to simplify the concepts.

- The smallest unit of binary data is the bit.
- Data are handled in an 8-bit unit called byte.
- The byte can be split into two 4-bit unit that are called nibbles.
- A complete unit of information is called a word. For a 32-bit architecture, word size is 32 bit and for 64-bit architecture word size is 64 bit
- Each storage element in a memory can retain either a 1 or a 0 is called a cell.
- Memories are made up of array of cells.
- Each block in the memory array represents one storage cell.
- It's location can be identified by specifying a row and a column.
- The location of a unit of data in a memory array is called its address.
- The capacity of a memory is the total number of data units that can be stored.



A 64-cell memory array organized in 3 different ways

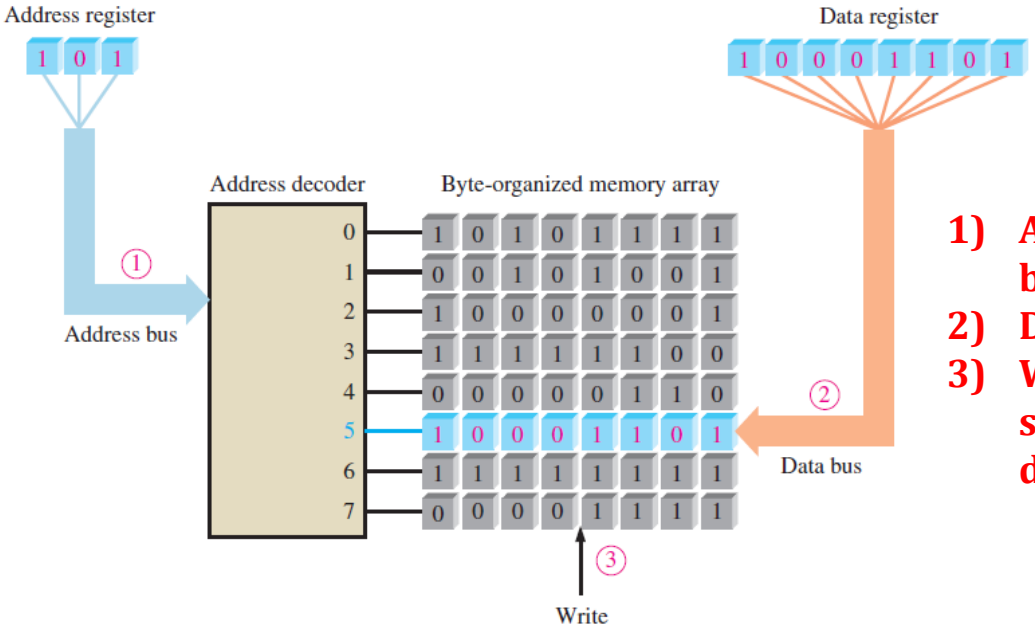


The 8 × 8 bit array expanded to a 64 × 8 bit array. This array forms a memory module.

Nafiz Ahmed Chisty| Head (UG) & Associate Professor, Dept. of EEE, FE, AIUB|

Write Operation

- The write operation puts data into a specified address in the memory. Data units go into the memory during 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 in the data register is placed on the data bus and stored in the selected memory address, thus completing the write operation.
- When a new data byte is written into a memory address, the current data byte stored at that address is overwritten

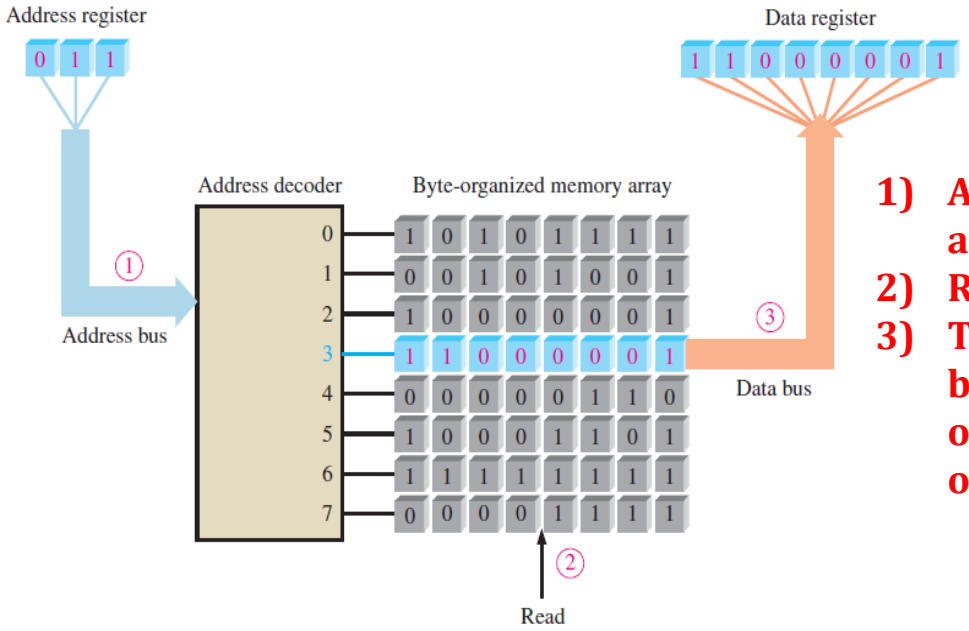


1) **Address code 101 is placed on the address bus and address 5 is selected.**
2) **Data byte is placed on the data bus.**
3) **Write command causes the data byte to be stored in address 5, replacing previous data.**

Nafiz Ahmed Chisty| Head (UG) & Associate Professor, Dept. of EEE, FE, AIUB|

Read Operation

- The read operation copies data out of a specified address in the memory. Data units come of the memory during a 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.
- Then memory then gets a read command and "copy" of the data byte is stored in the selected memory address is placed on the data bus and loaded into the data register, thus completing the read operation.
- When a data byte is read from a memory address, it also remains stored at that address. This is called nondestructive read.



1) **Address code 011 is placed on the address bus and address 3 is selected.**
2) **Read command is applied.**
3) **The content of address 3 is placed on the data bust and shifted into data register. The content of address 3 is not erased by the read operation.**

Nafiz Ahmed Chisty| Head (UG) & Associate Professor, Dept. of EEE, FE, AIUB|

RAMs and ROMs

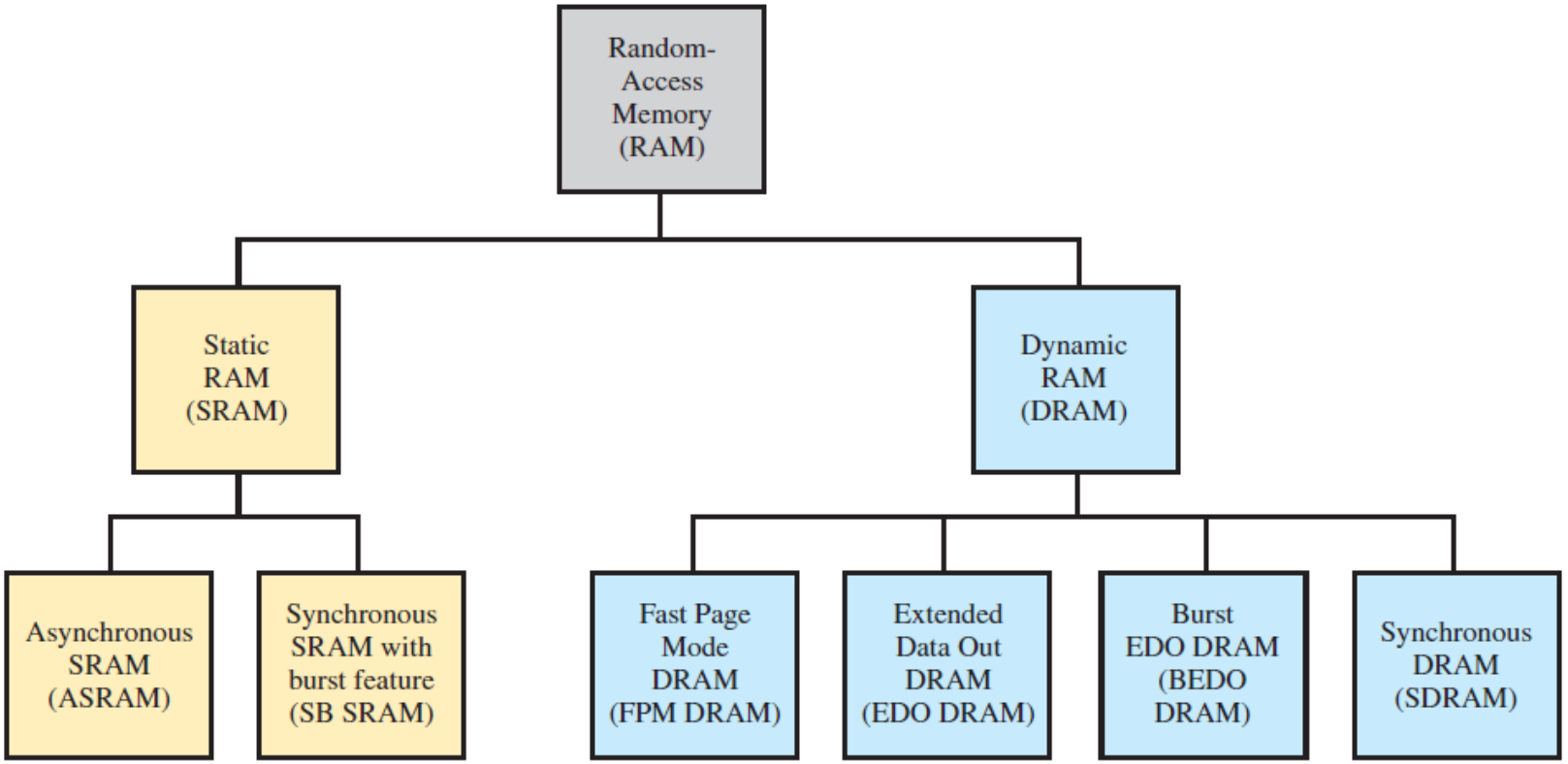
- The two major categories of semiconductor memories are the RAM and ROM.

- RAM stands for Random Access Memory.
- It is a type of memory in which all addresses are accessible in 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.
- Because RAMs lose stored data when the power is turned off they are called volatile memories.

- ROM stands for Read Only Memory.
- It is where data are stored permanently or semi-permanently.
- Data can be read from a ROM, but there is no write operation as in RAM.
- ROMs retain stored data even if power is turned off, so they are non-volatile memory.

Nafiz Ahmed Chisty| Head (UG) & Associate Professor, Dept. of EEE, FE, AIUB|

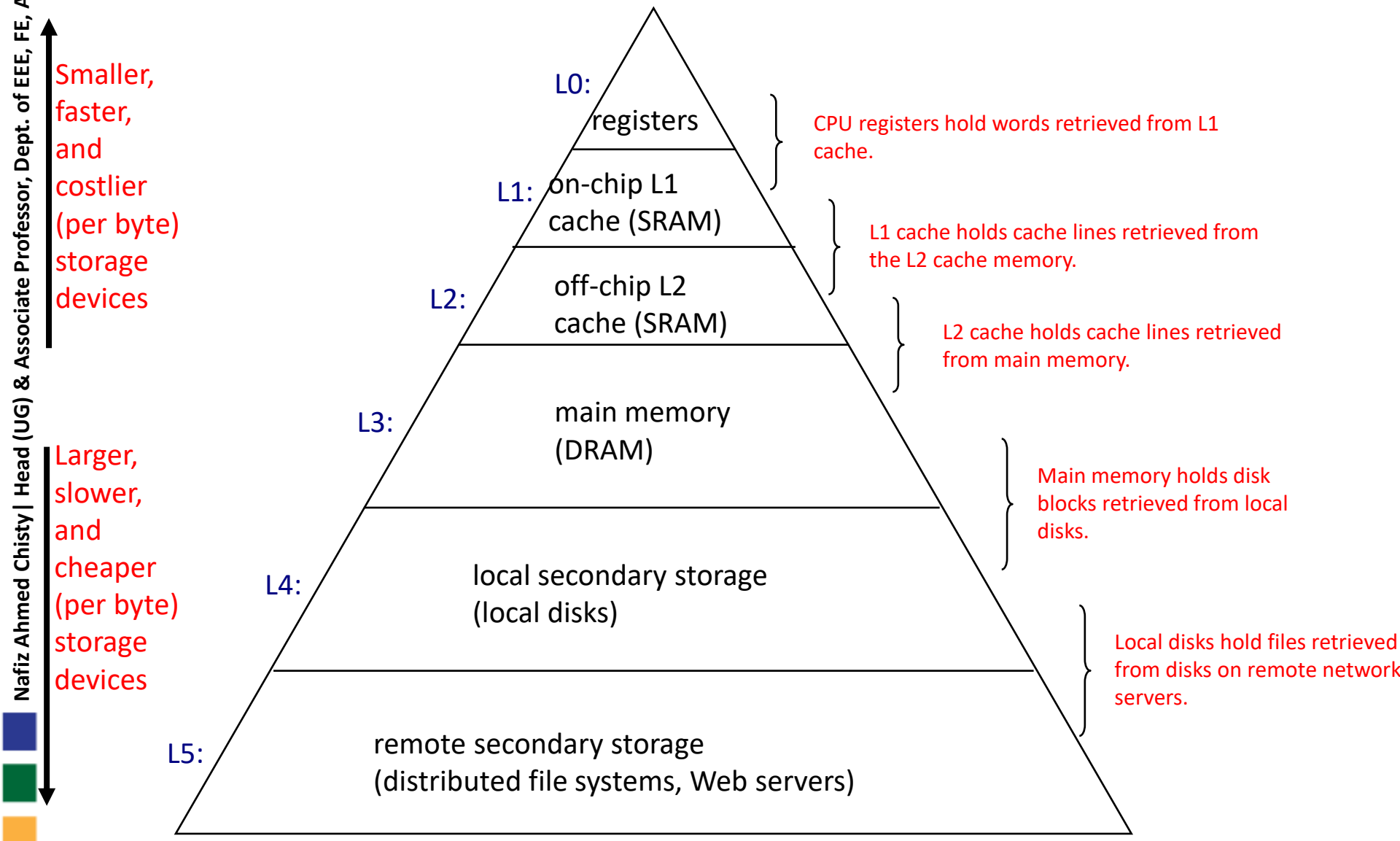
Random Access Memory (RAM)



The RAM Family

Nafiz Ahmed Chisty| Head (UG) & Associate Professor, Dept. of EEE, FE, AIUB|

An Example Memory Hierarchy



Random Access Memory (RAM)

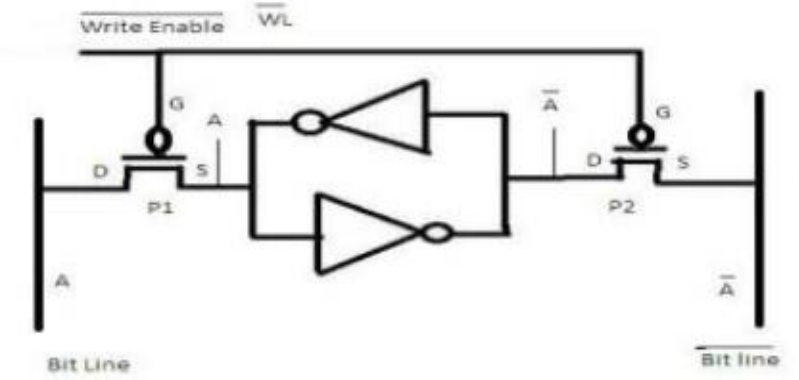
Comparison between SRAM and DRAM:

1. SRAM generally uses Latches as storage elements
2. It can store data indefinitely as long as power is applied.
3. SRAMs are faster.
4. For a given physical size and cost SRAMs can store less data.
5. SRAM cell is complex.
6. For a given physical size, less number cells can be crammed in the space.

1. DRAM generally uses capacitors as storage elements.
2. It cannot retain data very long without the capacitors being refreshed.
3. DRAMs are slower.
4. For a given physical size and cost DRAMs can store much more data.
5. DRAM cell is much simpler.
6. For a given physical size, a greater number of cells can be crammed in the space

Static Random-Access Memory (SRAM)

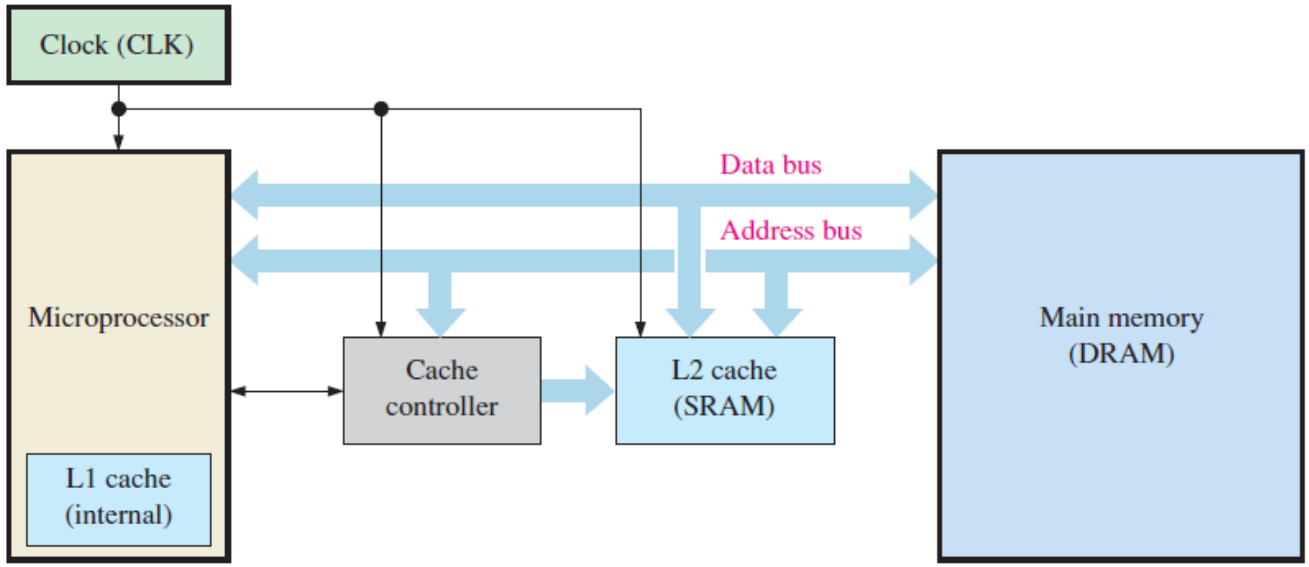
- All static RAM are characterized by latch memory cells.
- As long as dc power is applied to a static memory cell, it can retain a 1 or 0 indefinitely.
- If power is removed, the stored data bit is lost.
- Flip-flop storage cell are typically implemented in integrated circuits with several MOS transistors.



SRAM cell with inverter latch

Cache Memory

- One of the major applications of SRAM is the cache memories in computers.
- 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.
- Cache memory uses SRAM technology as it is faster.
- However, SRAM is faster but more expensive where as, DRAM is slower but cheaper.
- It is a cost-effective method of improving system performance without having to resort to the expense of making the overall memory faster.



Block Diagram of cache memory in a computer system

Cache Memory

- The operation of cache memory is analogous to the refrigerator at our home.
- We can get all what we need from the supermarket, however, we store the most necessary items in our refrigerator to save time.
- So every time we need something, we first look at our refrigerator, if it is there, we save a lot of time.
- Commonly there are two levels of cache: L1 and L2
- L1 is the first level cache and is usually integrated in the microprocessor.
- It has a very limited storage capacity.
- L1 is also known as primary cache.
- L2 cache is a separate memory chip or set of chips external to the processor.
- L2 cache usually has a larger storage capacity than L1 cache.
- L2 cache is also known as secondary cache.
- It should be noted that some systems also have higher level of cache L3 and L4

Nafiz Ahmed Chisty| Head (UG) & Associate Professor, Dept. of EEE, FE, AIUB|

Dynamic Random-Access Memory

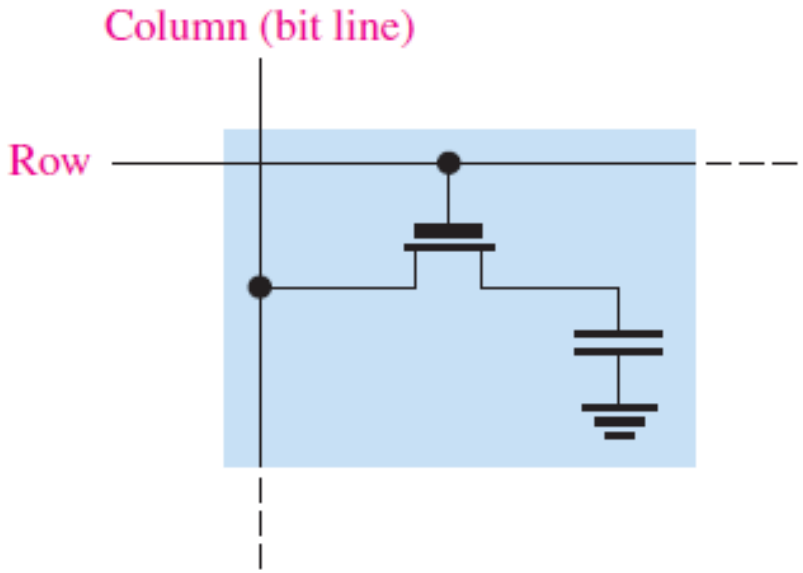
- Dynamic memory cells sore a data bit in a capacitor rather than a latch or flip-flop.
- Each cell consists of only one transistor and a capacitor.
- It is much simpler than the SRAM cell.
- It allows very large memory arrays to be constructed on a chip at a lower cost per bit than in a SRAM.

Disadvantages of DRAM

- The storage capacitors can not hold its charge over an extended period of time and lose the stored data bit unless its charge is refreshed periodically.
- To refresh it requires additional circuit and thus complicates the operation of DRAM

Dynamic Random-Access Memory

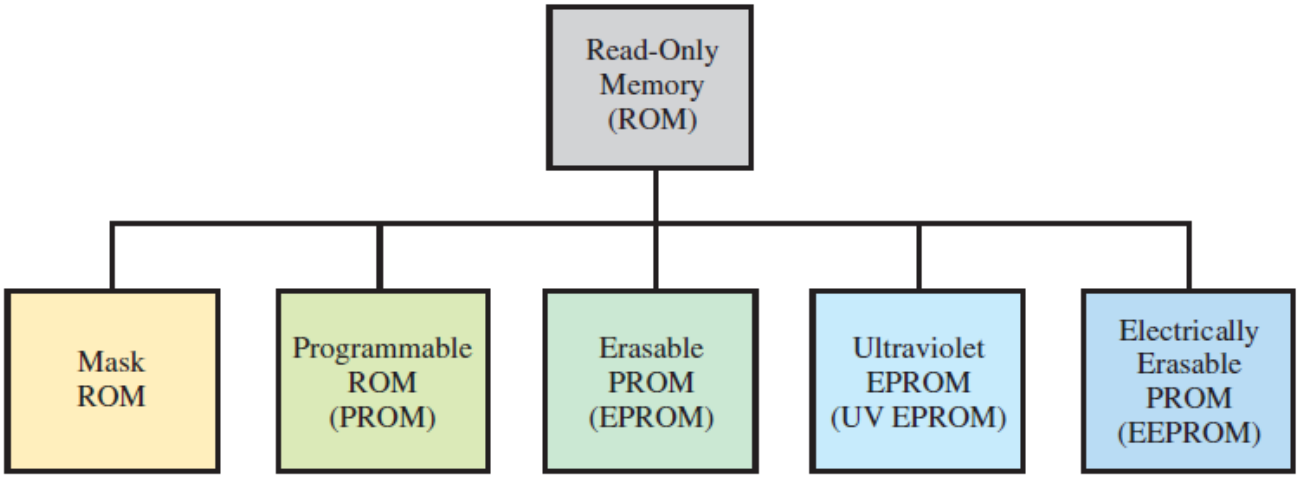
- The figure shows a DRAM cell.
- The transistor acts as a switch.
- The capacitor stores the charge.
- The row line is connected to the gate of the transistor.
- The column line is the bit line.
- When the transistor is on, the capacitor is connected to the bit line.
- Depending on the value of the capacitor and bit line the capacitor either charges or discharges.



Read-Only Memory (ROM)

- A ROM contains permanently or semi permanently stored data.
- It can be read from the memory.
- Either it cannot be changed at all or cannot be changed without specialized equipment.
- ROMs retain stored data when the power is off and are therefore nonvolatile memories.

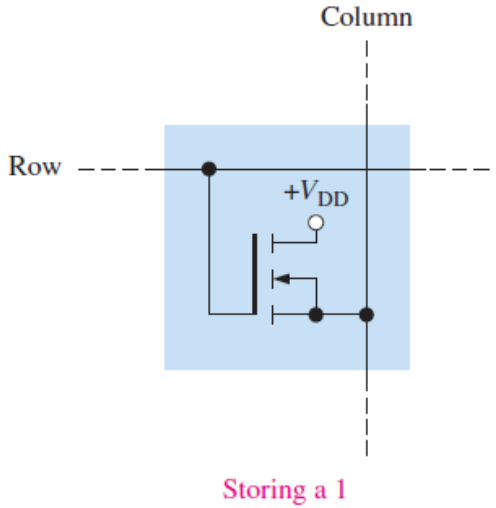
ROM Family



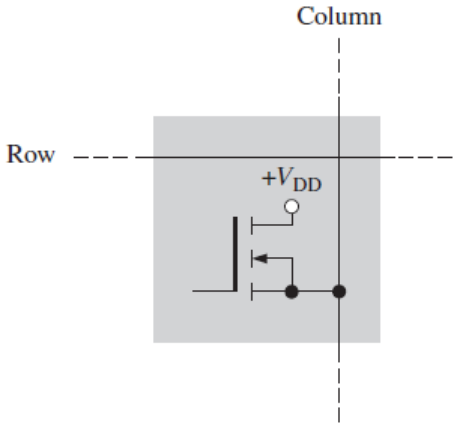
Nafiz Ahmed Chisty| Head (UG) & Associate Professor, Dept. of EEE, FE, AIUB|

Mask ROM

- The mask ROM is usually referred to simply as a ROM.
- It is permanently programmed during the manufacturing process.
- It can provide widely used standard functions, such as popular conversions, or to provide user-specified functions.
- Once memory is programmed it cannot be changed.
- A connection with the Row line to the gate of the transistor represents a stored 1.
- A missing connection with Row line to gate of the transistor represents a stored 0.



Storing a 1

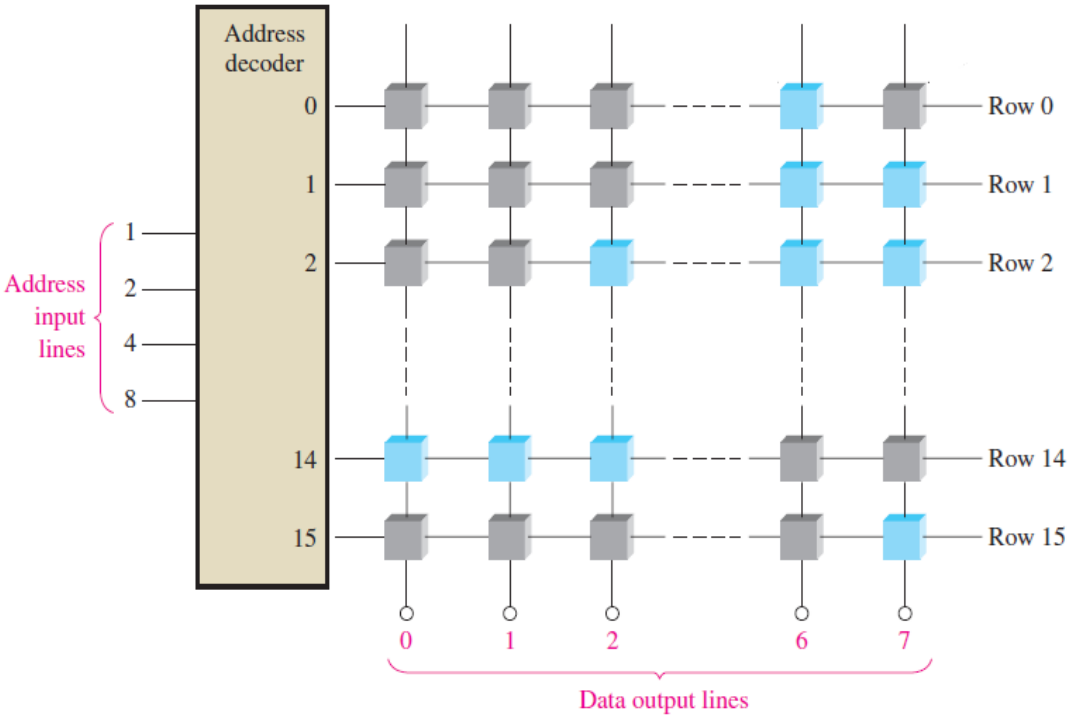


Storing a 0

Mask ROM

Read Operation:

- The blue squares represents stored 1.
- The gray squares represents stored 0.
- When binary address code is applied, corresponding row line becomes HIGH.
- Where there is 1 stored, the HIGH gets connected to the column line.
- Where there is 0 stored, the column line remains LOW.
- The column lines form the output line.
- The data stored in the row appears in the data line.

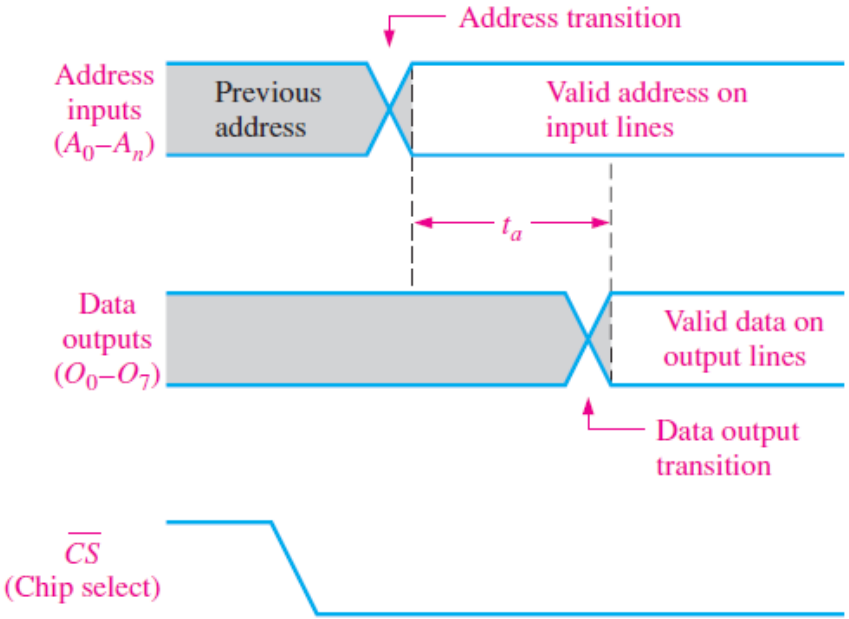


A 16 X 8-bit ROM array

Mask ROM

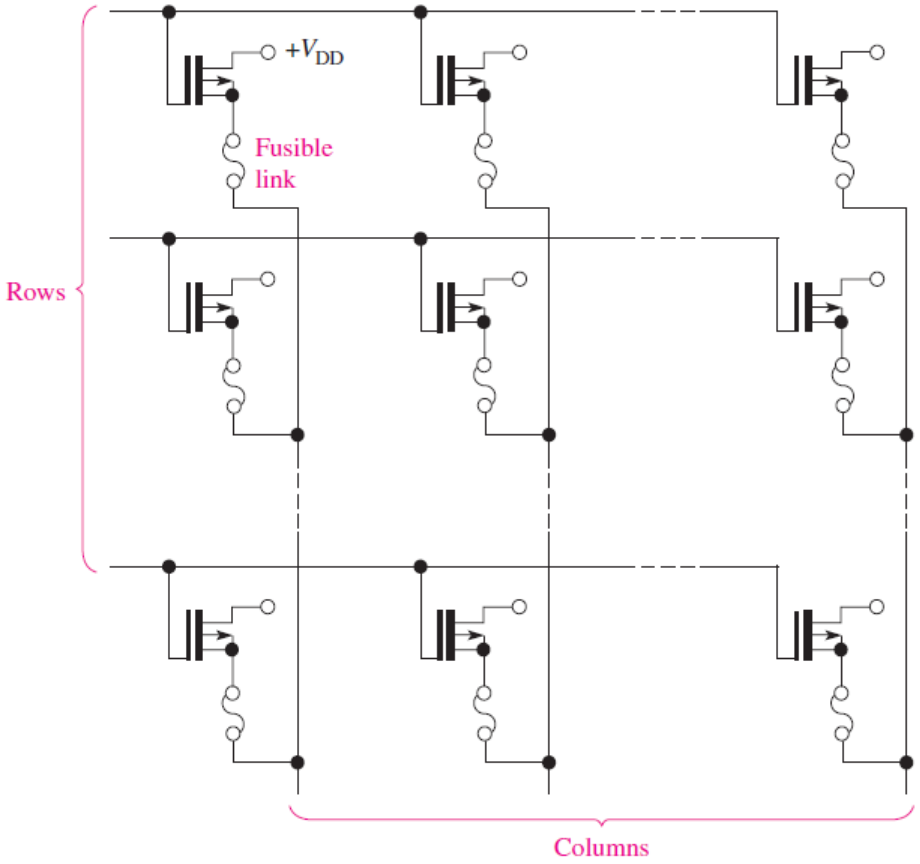
ROM Access Time

- The access time, $\mathbf{t_a}$, of a ROM is the time from the application of a valid address code on the input lines until the appearance of valid output data.
- Access time can also be measured from activation of the chip select ($\overline{CS}$) input to the occurrence of valid output data when a valid address is already on the input lines.



Programmable ROM (PROM)

- PROMs are basically the same as mask ROMs, once they have been programmed.
- The difference is that the PROM comes from the manufacturer unprogrammed.
- Thus they can be custom programmed in the field to meet the users demand.
- A PROM uses some sort of fusing process to store bits, in which a memory link is burned open or left intact.
- The presence of fuse stores a 1 and the absence stores a 0.
- The fusing process is irreversible.
- In the programming process, a sufficient amount of current is injected though the fusible link to burn open a fuse link.



MOS PROM array with fusible links

Nafiz Ahmed Chisty| Head (UG) & Associate Professor, Dept. of EEE, FE, AIUB|

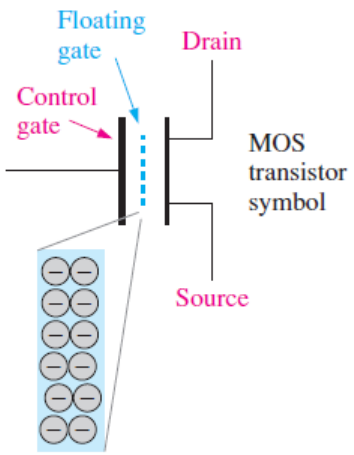
Erasable PROM (EPROM)

- An EPROM is an erasable PROM.
- Unlike an ordinary PROM, an EPROM can be reprogrammed if an existing program in the memory array is erased first.
- Two basic types of EPROM are the ultraviolet erasable PROM (UV EPROM) and the electrically erasable PROM (EEPROM)
- UP EPROM: Erasure is done by exposure of the memory array chip to high-intensity ultraviolet radiation through the quarts window on top of the package.
- EEPROM: An electrically erasable PROM can be both erased and programmed with electrical pulses. Since it can be both electrically written into and electrically erased, the EEPROM can be rapidly programmed and erased in-circuit for reprogramming.

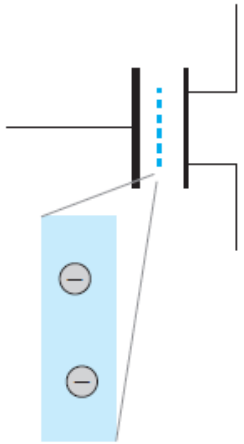
Nafiz Ahmed Chisty| Head (UG) & Associate Professor, Dept. of EEE, FE, AIUB|

Flash Memory

- The ideal memory has high storage capacity, nonvolatility, in-system read and write capability, comparatively fast operations and cost effectiveness.
- The traditional memory technologies such as ROM, PROM, EPROM, EEPROM, SRAM and DRAM individually exhibits one or more characteristics.
- However, none of them have all these characteristics except for Flash Memory.
- The stacked gate MOS transistor consists of a control gate and floating gate in addition to the drain and source.
- The floating gate stores electron as a result of a sufficient voltage applied to the control gate.
- A 0 is stored when there is more charge and a 1 is stored when there is less charge.
- The amount of charge in the floating gate determines if the transistor will turn on and conduct current from the drain to source when a control voltage is applied during a read operation.



Many electrons = more charge = stored 0.

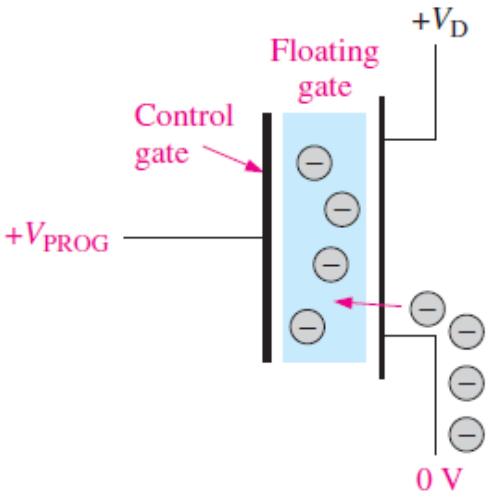


Few electrons = less charge = stored 1.

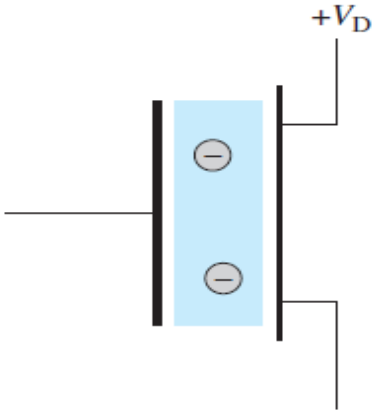
Nafiz Ahmed Chisty| Head (UG) & Associate Professor, Dept. of EEE, FE, AIUB|

Flash Memory Store Operation

To store a 0, sufficient positive voltage is applied to the control gate with respect to the source to add charge to the floating gate during programming.

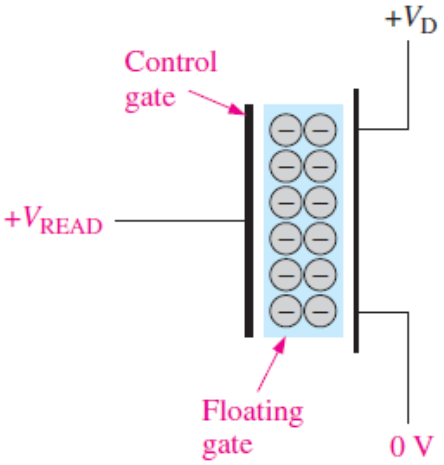


To store a 1, no charge is added, and the cell is left in the erased condition.

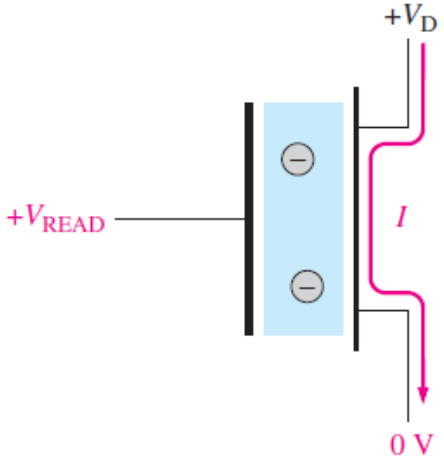


Flash Memory Read Operation

When a 0 is read, the transistor remains off because the charge on the floating gate prevents the read voltage from exceeding the turn on threshold



When a 1 is read, the transistor turns on because the absence of charge on the floating gate allows the read voltage to exceed the turn-on threshold.



Nafiz Ahmed Chisty| Head (UG) & Associate Professor, Dept. of EEE, FE, AIUB|

Comparison of types of memories.

Memory Type	Nonvolatile	High-Density	One-Transistor Cell	In-System Writability
Flash	Yes	Yes	Yes	Yes
SRAM	No	No	No	Yes
DRAM	No	Yes	Yes	Yes
ROM	Yes	Yes	Yes	No
EEPROM	Yes	No	No	Yes
UV EPROM	Yes	Yes	Yes	No

Nafiz Ahmed Chisty| Head (UG) & Associate Professor, Dept. of EEE, FE, AIUB|

Reference:

[1] Thomas L. Floyd, "Digital Fundamentals" 11th edition, Prentice Hall.
[2] M. Morris Mano, "Digital Logic & Computer Design" Prentice Hall.
[3] Mixed contents from Vahid And Howard.

Thanks


AIUB
EEE
FACULTY OF ENGINEERING
AIUB
Faculty of Engineering