

Flash memory was invented by Fujio Masuoka at Toshiba in the early 1980s.^{[33][34]} Masuoka and colleagues presented the invention of NOR flash in 1984,^[35] and then NAND flash in 1987.^[36] Toshiba commercialized NAND flash memory in 1987.^{[37][38]}

Applications

MOS memory applications

MOS memory type	Abbr.	MOS memory cell	Applications
Static random-access memory	SRAM	MOSFETs	Cache memory, cell phones, eSRAM, mainframes, multimedia computers, networking, personal computers, servers, supercomputers, telecommunications, workstations, ^[39] DVD disk buffer, ^[40] data buffer, ^[41] nonvolatile BIOS memory
Dynamic random-access memory	DRAM	MOSFET, MOS capacitor	Camcorders, embedded logic, eDRAM, graphics card, hard disk drive (HDD), networks, personal computers, personal digital assistants, printers, ^[39] main computer memory, desktop computers, servers, solid-state drives, video memory, ^[40] framebuffer memory ^{[42][43]}
Ferroelectric random-access memory	FRAM	MOSFET, Ferroelectric capacitor	Non-volatile memory, radio-frequency identification (RF identification), smart cards ^{[39][40]}
Read-only memory	ROM	MOSFET	Character generators, electronic musical instruments, laser printer fonts, video game ROM cartridges, word processor dictionary data ^{[39][40]}
Erasable programmable read-only memory	EPROM	Floating-gate MOSFET	CD-ROM drives, embedded memory, code storage, modems ^{[39][40]}
Electrically erasable programmable read-only memory	EEPROM	Floating-gate MOSFET	Anti-lock braking systems, air bags, car radios, cell phones, consumer electronics, cordless telephones, disk drives, embedded memory, flight controllers, military technology, modems, pagers, printers, set-top box, smart cards ^{[39][40]}
Flash memory	Flash	Floating-gate MOSFET	ATA controllers, battery-powered applications, telecommunications, code storage, digital cameras, MP3 players, portable media players, BIOS memory, ^[39] USB flash drive, ^[44] digital TV, e-books, memory cards, mobile devices, set-top box, smartphones, solid-state drives, tablet computers ^[40]
Non-volatile random-access memory	NVRAM	Floating-gate MOSFETs	Medical equipment, spacecraft ^{[39][40]}

See also

- List of best-selling electronic devices
- Semiconductor industry

References

1. "The MOS Memory Market" (<http://smithsonianchips.si.edu/ice/cd/MEMORY97/SEC01.PDF>)

Fairchild Semiconductor in 1964.^{[14][15]} In addition to higher performance, MOS memory was cheaper and consumed less power than magnetic-core memory.^[14] This led to MOSFETs eventually replacing magnetic cores as the standard storage elements in computer memory.^[12]

In 1965, J. Wood and R. Ball of the Royal Radar Establishment proposed digital storage systems that use CMOS (complementary MOS) memory cells, in addition to MOSFET power devices for the power supply, switched cross-coupling, switches and delay-line storage.^[16] The development of silicon-gate MOS integrated circuit (MOS IC) technology by Federico Faggin at Fairchild in 1968 enabled the production of MOS memory chips.^[17] NMOS memory was commercialized by IBM in the early 1970s.^[18] MOS memory overtook magnetic core memory as the dominant memory technology in the early 1970s.^[14]

The term "memory" when used with reference to computers most often refers to volatile random-access memory (RAM). The two main types of volatile RAM are static random-access memory (SRAM) and dynamic random-access memory (DRAM). Bipolar SRAM was invented by Robert Norman at Fairchild Semiconductor in 1963,^[9] followed by the development of MOS SRAM by John Schmidt at Fairchild in 1964.^[14] SRAM became an alternative to magnetic-core memory, but required six MOS transistors for each bit of data.^[19] Commercial use of SRAM began in 1965, when IBM introduced their SP95 SRAM chip for the System/360 Model 95.^[9]

Toshiba introduced bipolar DRAM memory cells for its Toscal BC-1411 electronic calculator in 1965.^{[20][21]} While it offered improved performance over magnetic-core memory, bipolar DRAM could not compete with the lower price of the then dominant magnetic-core memory.^[22] MOS technology is the basis for modern DRAM. In 1966, Dr. Robert H. Dennard at the IBM Thomas J. Watson Research Center was working on MOS memory. While examining the characteristics of MOS technology, he found it was capable of building capacitors, and that storing a charge or no charge on the MOS capacitor could represent the 1 and 0 of a bit, while the MOS transistor could control writing the charge to the capacitor. This led to his development of a single-transistor DRAM memory cell.^[19] In 1967, Dennard filed a patent under IBM for a single-transistor DRAM memory cell, based on MOS technology.^[23] This led to the first commercial DRAM IC chip, the Intel 1103, in October 1970.^{[24][25][26]} Synchronous dynamic random-access memory (SDRAM) later debuted with the Samsung KM48SL2000 chip in 1992.^{[27][28]}

The term "memory" is also often used to refer to non-volatile memory, specifically flash memory. It has origins in read-only memory (ROM). Programmable read-only memory (PROM) was invented by Wen Tsing Chow in 1956, while working for the Arma Division of the American Bosch Arma Corporation.^{[29][30]} In 1967, Dawon Kahng and Simon Sze of Bell Labs proposed that the floating gate of a MOS semiconductor device could be used for the cell of a reprogrammable read-only memory (ROM), which led to Dov Frohman of Intel inventing EPROM (erasable PROM) in 1971.^[31] EEPROM (electrically erasable PROM) was developed by Yasuo Tarui, Yutaka Hayashi and Kiyoko Naga at Japan's Ministry of International Trade and Industry (MITI) Electrotechnical Laboratory in 1972.^[32]

- **MROM** (*Mask programmed ROM or Mask ROM*) – In this type the data is programmed into the chip when the chip is manufactured, so it is only used for large production runs. It cannot be rewritten with new data.
- **PROM** (*Programmable read-only memory*) – In this type the data is written into an existing PROM chip before it is installed in the circuit, but it can only be written once. The data is written by plugging the chip into a device called a PROM programmer.
- **EPROM** (*Erasable programmable read-only memory* or UVEPROM) – In this type the data in it can be rewritten by removing the chip from the circuit board, exposing it to an ultraviolet light to erase the existing data, and plugging it into a PROM programmer. The IC package has a small transparent "window" in the top to admit the UV light. It is often used for prototypes and small production run devices, where the program in it may have to be changed at the factory.
- **EEPROM** (*Electrically erasable programmable read-only memory*) – In this type the data can be rewritten electrically, while the chip is on the circuit board, but the writing process is slow. This type is used to hold firmware, the low level microcode which runs hardware devices, such as the BIOS program in most computers, so that it can be updated.
- **NVRAM** (*Non-volatile random-access memory*)
 - **FRAM** (*Ferroelectric RAM*) – One type of nonvolatile RAM.
 - **Flash memory** – In this type the writing process is intermediate in speed between EEPROMS and RAM memory; it can be written to, but not fast enough to serve as main memory. It is often used as a semiconductor version of a hard disk, to store files. It is used in portable devices such as PDAs, USB flash drives, and removable memory cards used in digital cameras and cellphones.



4M EPROM, showing transparent window used to erase the chip

History

Early computer memory consisted of magnetic-core memory, as early solid-state electronic semiconductors, including transistors such as the bipolar junction transistor (BJT), were impractical for use as digital storage elements (memory cells). The earliest semiconductor memory dates back to the early 1960s, with bipolar memory, which used bipolar transistors.^[9] **Bipolar semiconductor memory** made from discrete devices was first shipped by Texas Instruments to the United States Air Force in 1961. The same year, the concept of solid-state memory on an integrated circuit (IC) chip was proposed by applications engineer Bob Norman at Fairchild Semiconductor.^[10] The **first single-chip** memory IC was the BJT 16-bit IBM SP95 fabricated in December 1965, engineered by Paul Castrucci.^{[9][10]} While bipolar memory offered improved performance over magnetic-core memory, it could not compete with the lower price of magnetic-core memory, which remained dominant up until the late 1960s.^[9] Bipolar memory failed to replace magnetic-core memory because bipolar flip-flop circuits were too large and expensive.^[11]

MOS memory

The advent of the metal–oxide–semiconductor field-effect transistor (MOSFET),^[12] invented by Mohamed M. Atalla and Dawon Kahng at Bell Labs in 1959,^[13] enabled the practical use of metal–oxide–semiconductor (MOS) transistors as memory cell storage elements, a function previously served by magnetic cores in computer memory.^[12] MOS memory was developed by John Schmidt at

- **DDR SDRAM** (*Double data rate SDRAM*) – This could transfer twice the data (two consecutive words) on each clock cycle by double pumping (transferring data on both rising and falling edges of the clock pulse). Extensions of this idea are the current (2012) technique being used to increase memory access rate and throughput. Since it is proving difficult to further increase the internal clock speed of memory chips, these chips increase the transfer rate by transferring more data words on each clock cycle
 - **DDR2 SDRAM** – Transfers 4 consecutive words per internal clock cycle
 - **DDR3 SDRAM** – Transfers 8 consecutive words per internal clock cycle.
 - **DDR4 SDRAM** – Transfers 16 consecutive words per internal clock cycle.
- **RDRAM** (*Rambus DRAM*) – An alternate double data rate memory standard that was used on some Intel systems but ultimately lost out to DDR SDRAM.
- **XDR DRAM** (*Extreme data rate DRAM*)
- **SGRAM** (*Synchronous graphics RAM*) – A specialized type of SDRAM made for graphics adaptors (video cards). It can perform graphics-related operations such as bit masking and block write, and can open two pages of memory at once.
- **GDDR SDRAM** (*Graphics DDR SDRAM*)
 - **GDDR2**
 - **GDDR3 SDRAM**
 - **GDDR4 SDRAM**
 - **GDDR5 SDRAM**
 - **GDDR6 SDRAM**
- **HBM** (*High Bandwidth Memory*) – A development of SDRAM used in graphics cards that can transfer data at a faster rate. It consists of multiple memory chips stacked on top of one another, with a wider data bus.
- **PSRAM** (*Pseudostatic RAM*) – This is DRAM which has circuitry to perform memory refresh on the chip, so that it acts like SRAM, allowing the external memory controller to be shut down to save energy. It is used in a few game consoles such as the Wii.
- **SRAM** (*Static random-access memory*) – This stores each bit of data in a circuit called a flip-flop, made of 4 to 6 transistors. SRAM is less dense and more expensive per bit than DRAM, but faster and does not require memory refresh. It is used for smaller cache memories in computers.
- **CAM** (*Content-addressable memory*) – This is a specialized type in which, instead of accessing data using an address, a data word is applied and the memory returns the location if the word is stored in the memory. It is mostly incorporated in other chips such as microprocessors where it is used for cache memory.

Non-volatile memory

Non-volatile memory (NVM) preserves the data stored in it during periods when the power to the chip is turned off. Therefore, it is used for the memory in portable devices, which don't have disks, and for removable memory cards among other uses. Major types are:^{[7][8]}

- **ROM** (*Read-only memory*) – This is designed to hold permanent data, and in normal operation is only read from, not written to. Although many types can be written to, the writing process is slow and usually all the data in the chip must be rewritten at once. It is usually used to store system software which must be immediately accessible to the computer, such as the BIOS program which starts the computer, and the software (microcode) for portable devices and embedded computers such as microcontrollers.

The two basic operations performed by a memory chip are "read", in which the data contents of a memory word is read out (nondestructively), and "write" in which data is stored in a memory word, replacing any data that was previously stored there. To increase data rate, in some of the latest types of memory chips such as DDR SDRAM multiple words are accessed with each read or write operation.

In addition to standalone memory chips, blocks of semiconductor memory are integral parts of many computer and data processing integrated circuits. For example, the microprocessor chips that run computers contain cache memory to store instructions awaiting execution.

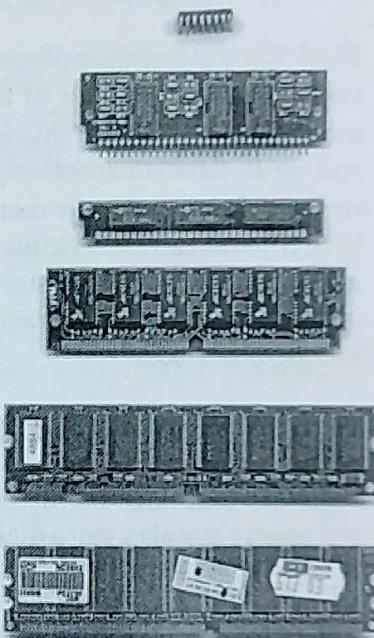
Types

Volatile memory

Volatile memory loses its stored data when the power to the memory chip is turned off. However it can be faster and less expensive than non-volatile memory. This type is used for the main memory in most computers, since data is stored on the hard disk while the computer is off. Major types are:^{[7][8]}

RAM (*Random-access memory*) – This has become a generic term for any semiconductor memory that can be written to, as well as read from, in contrast to ROM (*below*), which can only be read. All semiconductor memory, not just RAM, has the property of random access.

- **DRAM** (*Dynamic random-access memory*) – This uses memory cells consisting of one MOSFET (MOS field-effect transistor) and one MOS capacitor to store each bit. This type of RAM is the cheapest and highest in density, so it is used for the main memory in computers. However, the electric charge that stores the data in the memory cells slowly leaks out, so the memory cells must be periodically refreshed (rewritten) which requires additional circuitry. The refresh process is handled internally by the computer and is transparent to its user.
- **FPM DRAM** (*Fast page mode DRAM*) – An older type of asynchronous DRAM that improved on previous types by allowing repeated accesses to a single "page" of memory to occur at a faster rate. Used in the mid-1990s.
- **EDO DRAM** (*Extended data out DRAM*) – An older type of asynchronous DRAM which had faster access time than earlier types by being able to initiate a new memory access while data from the previous access was still being transferred. Used in the later part of the 1990s.
- **VRAM** (*Video random access memory*) – An older type of dual-ported memory once used for the frame buffers of video adapters (video cards).
- **SDRAM** (*Synchronous dynamic random-access memory*) – This added circuitry to the DRAM chip which synchronizes all operations with a clock signal added to the computer's memory bus. This allowed the chip to process multiple memory requests simultaneously using pipelining, to increase the speed. The data on the chip is also divided into banks which can each work on a memory operation simultaneously. This became the dominant type of computer memory by about the year 2000.



RAM chips for computers usually come on removable memory modules like these. Additional memory can be added to the computer by plugging in additional modules.

WIKIPEDIA
The Free Encyclopedia

Semiconductor memory

Semiconductor memory is a digital electronic semiconductor device used for digital data storage, such as computer memory. It typically refers to devices in which data is stored within metal-oxide-semiconductor (MOS) memory cells on a silicon integrated circuit memory chip.^{[1][2][3]} There are numerous different types using different semiconductor technologies. The two main types of random-access memory (RAM) are static RAM (SRAM), which uses several transistors per memory cell, and dynamic RAM (DRAM), which uses a transistor and a MOS capacitor per cell. Non-volatile memory (such as EPROM, EEPROM and flash memory) uses floating-gate memory cells, which consist of a single floating-gate transistor per cell.

Most types of semiconductor memory have the property of random access,^[4] which means that it takes the same amount of time to access any memory location, so data can be efficiently accessed in any random order.^[5] This contrasts with data storage media such as hard disks and CDs which read and write data consecutively and therefore the data can only be accessed in the same sequence it was written. Semiconductor memory also has much faster access times than other types of data storage; a byte of data can be written to or read from semiconductor memory within a few nanoseconds, while access time for rotating storage such as hard disks is in the range of milliseconds. For these reasons it is used for primary storage, to hold the program and data the computer is currently working on, among other uses.

As of 2017, semiconductor memory chips sell \$124 billion annually, accounting for 30% of the semiconductor industry.^[6] Shift registers, processor registers, data buffers and other small digital registers that have no memory address decoding mechanism are typically not referred to as *memory* although they also store digital data.

Description

In a semiconductor memory chip, each bit of binary data is stored in a tiny circuit called a memory cell consisting of one to several transistors. The memory cells are laid out in rectangular arrays on the surface of the chip. The 1-bit memory cells are grouped in small units called words which are accessed together as a single memory address. Memory is manufactured in word length that is usually a power of two, typically $N=1, 2, 4$ or 8 bits.

Data is accessed by means of a binary number called a memory address applied to the chip's address pins, which specifies which word in the chip is to be accessed. If the memory address consists of M bits, the number of addresses on the chip is 2^M , each containing an N bit word. Consequently, the amount of data stored in each chip is $N2^M$ bits.^[5] The memory storage capacity for M number of address lines is given by 2^M , which is usually in power of two: 2, 4, 8, 16, 32, 64, 128, 256 and 512 and measured in kilobits, megabits, gigabits or terabits, etc. As of 2014 the largest semiconductor memory chips hold a few gigabits of data, but higher capacity memory is constantly being developed. By combining several integrated circuits, memory can be arranged into a larger word length and/or address space than what is offered by each chip, often but not necessarily a power of two.^[5]

Memory Classification

- (i) Registers, Main memory & Secondary memory.
- (ii) Sequential Access & Random access.
- (iii) Static & Dynamic Memory
- (iv) Volatile & non-volatile memory
- (v) Magnetic & Semiconductor memory.

PLD → Programmable Logic Device: ROM, PROM, EEPROM
EEPROM, FPGA.

FPGA → is a reprogrammable gate array that uses anti-fuse.

⇒ The difference between PLD & FPGA is that FPGA incorporates logic blocks instead of fixed AND - OR gates and is faster with low power dissipation.

ROM:

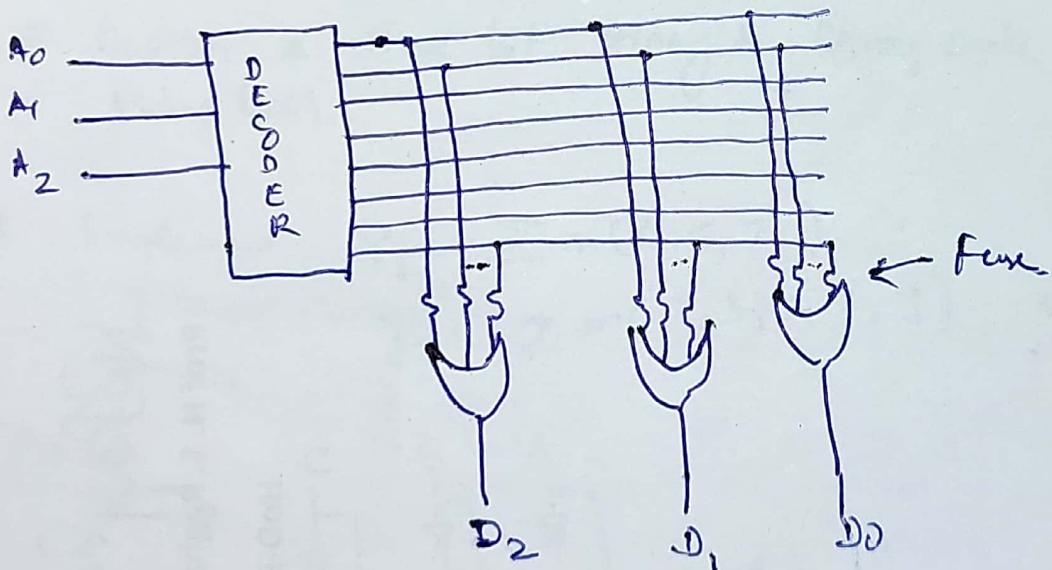
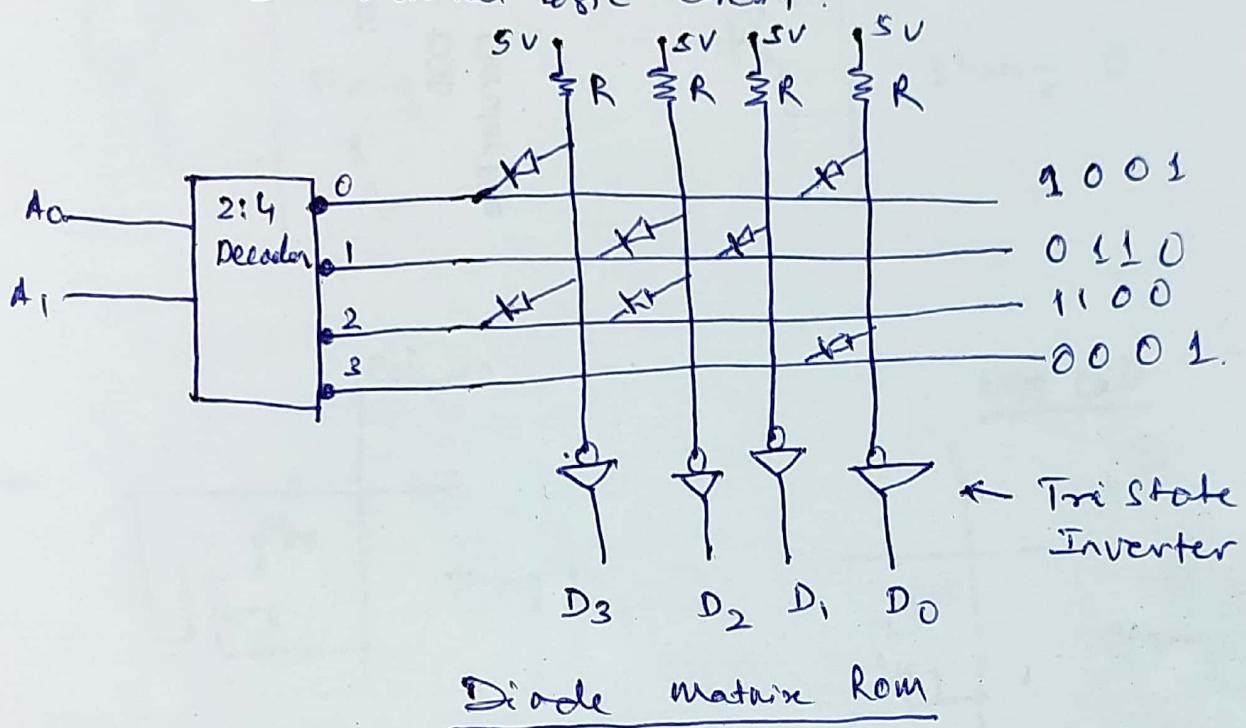
- (A) Used to store data, program as well as can implement any logic circuit / Truth Table.

- (B) Uses fixed AND gate array [Decoder] & programmable OR gate array with fuses.

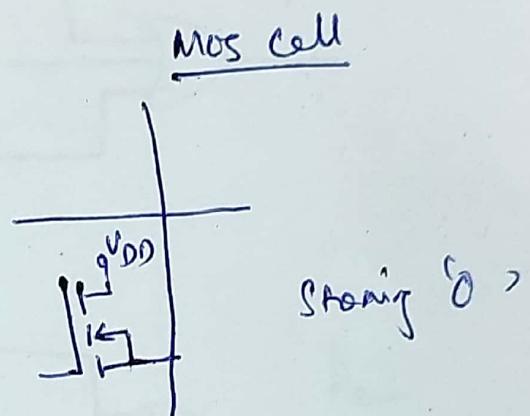
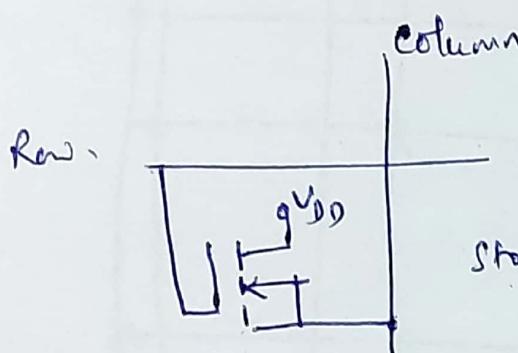
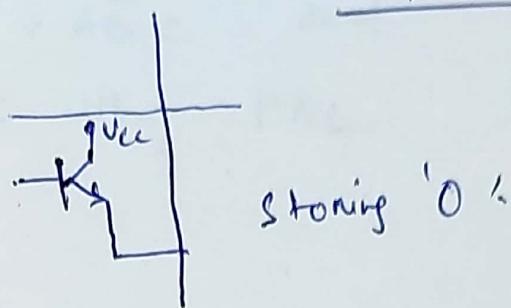
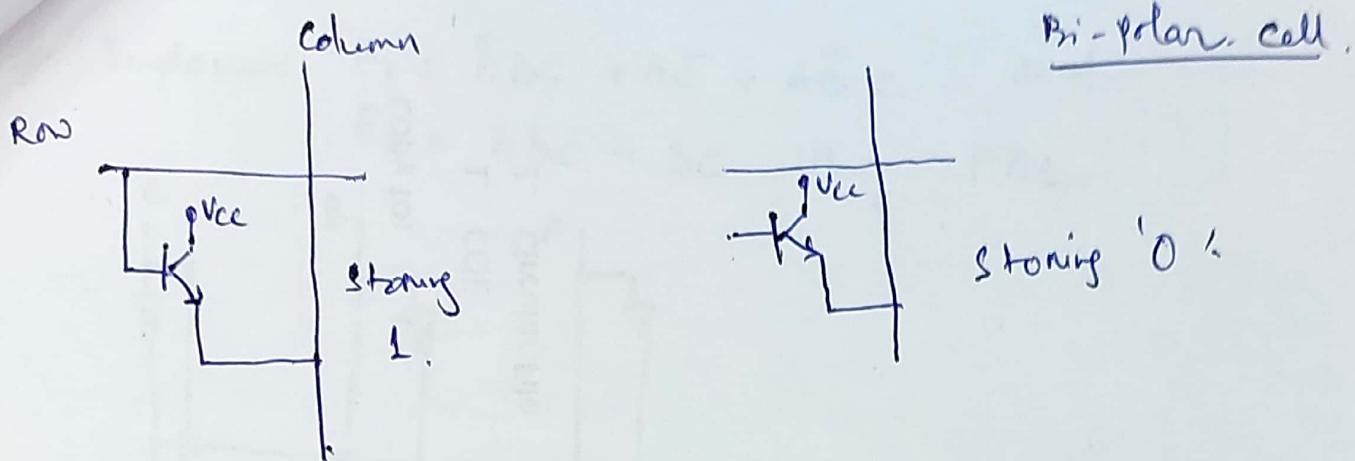
- (C) PAL - Programmable AND gate array & fixed OR gate array.

- (D) PLA - Both are Programmable.

- ROM:
- Does not have a write capability.
 - Programmed for a purpose during manufacturing.
 - Combinational logic circuit.



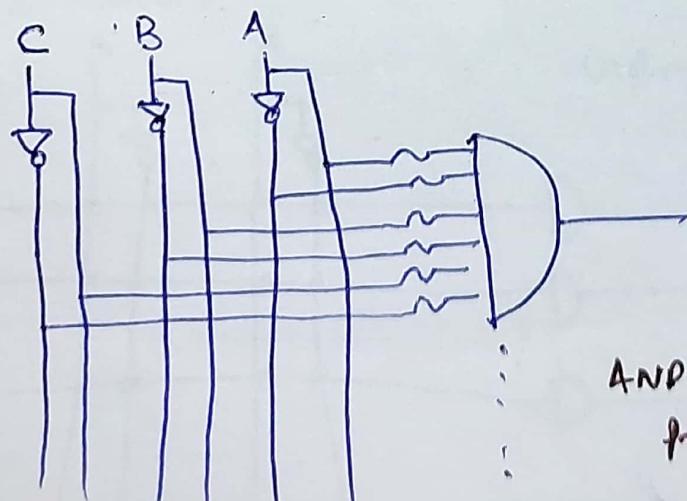
Rom Using OR gate array



Design a three bit Binary to Gray Code Converter Using ROM.

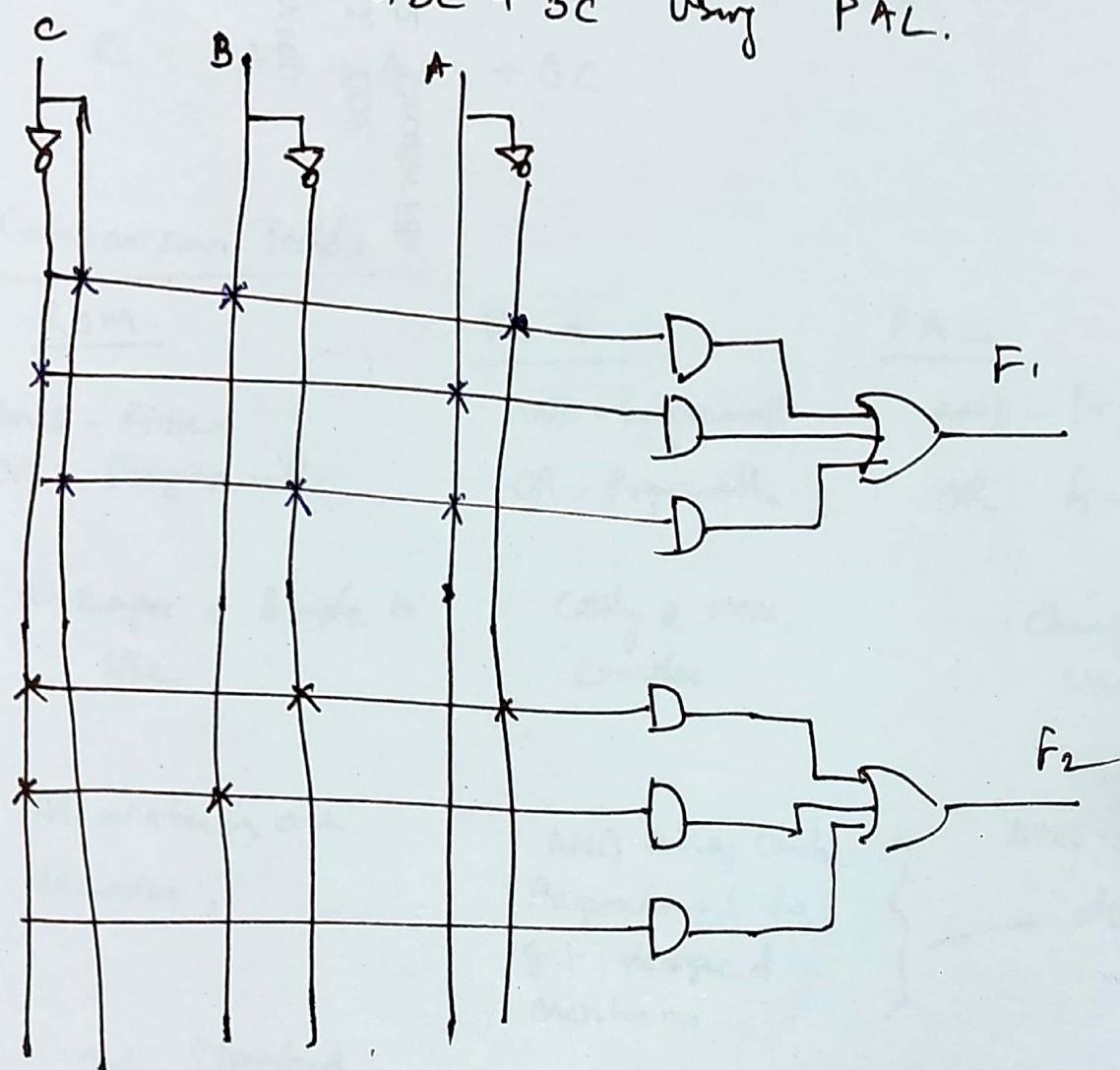
Implement $y_1 = \sum m(4, 5, 7, 9)$
 $y_2 = \sum m(1, 3, 5, 7, 9)$ using Rom.

PAL

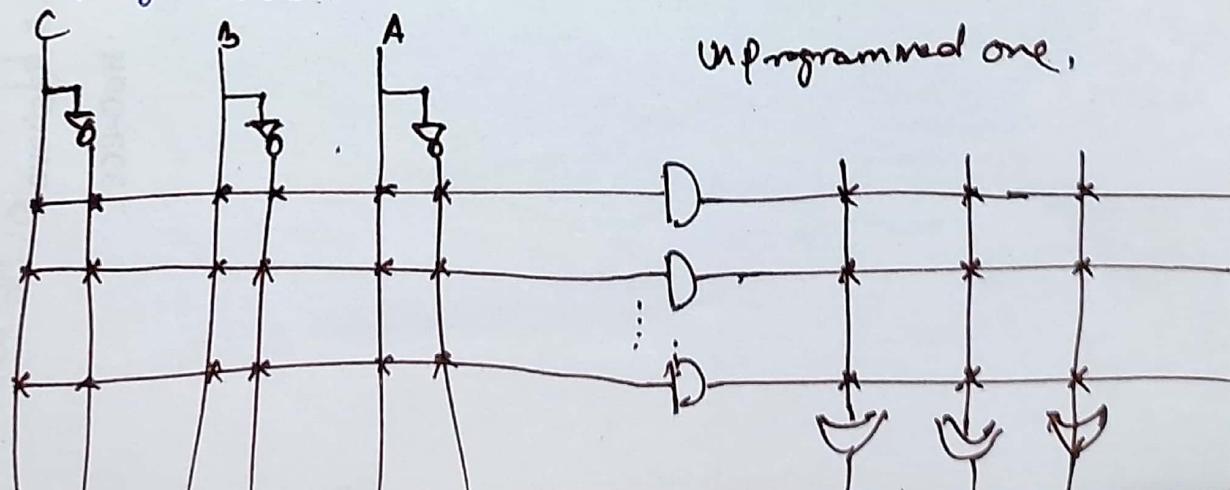


4

Implement $F_1 = ABC + A\bar{C} + A\bar{B}C$ and $F_2 = \bar{A}\bar{B}\bar{C} + BC$ using PAL.



PLA Both OR gate array & AND gate array are programmable.



• Implement a full adder circuit using PLA

$$S = \bar{A}\bar{B}C + \bar{A}B\bar{C} + A\bar{B}\bar{C} + ABC$$

$$C = AB + AC + BC$$

Comparison Table

	<u>ROM</u>	<u>PLA</u>	<u>PAL</u>
(i)	AND - Fixed OR - Programmable	AND - Programmable OR - Programmable	AND - Programmable OR - Fixed
(ii)	Cheaper & simple to use	Costly & more complex	Cheaper & simple
(iii)	All minterms are decoder.	AND array can be programmed to get desired minterms	AND → d0,
(iv)	Only Standard SOP can be implemented.	Any boolean function in SOP can be implemented	→ d0,

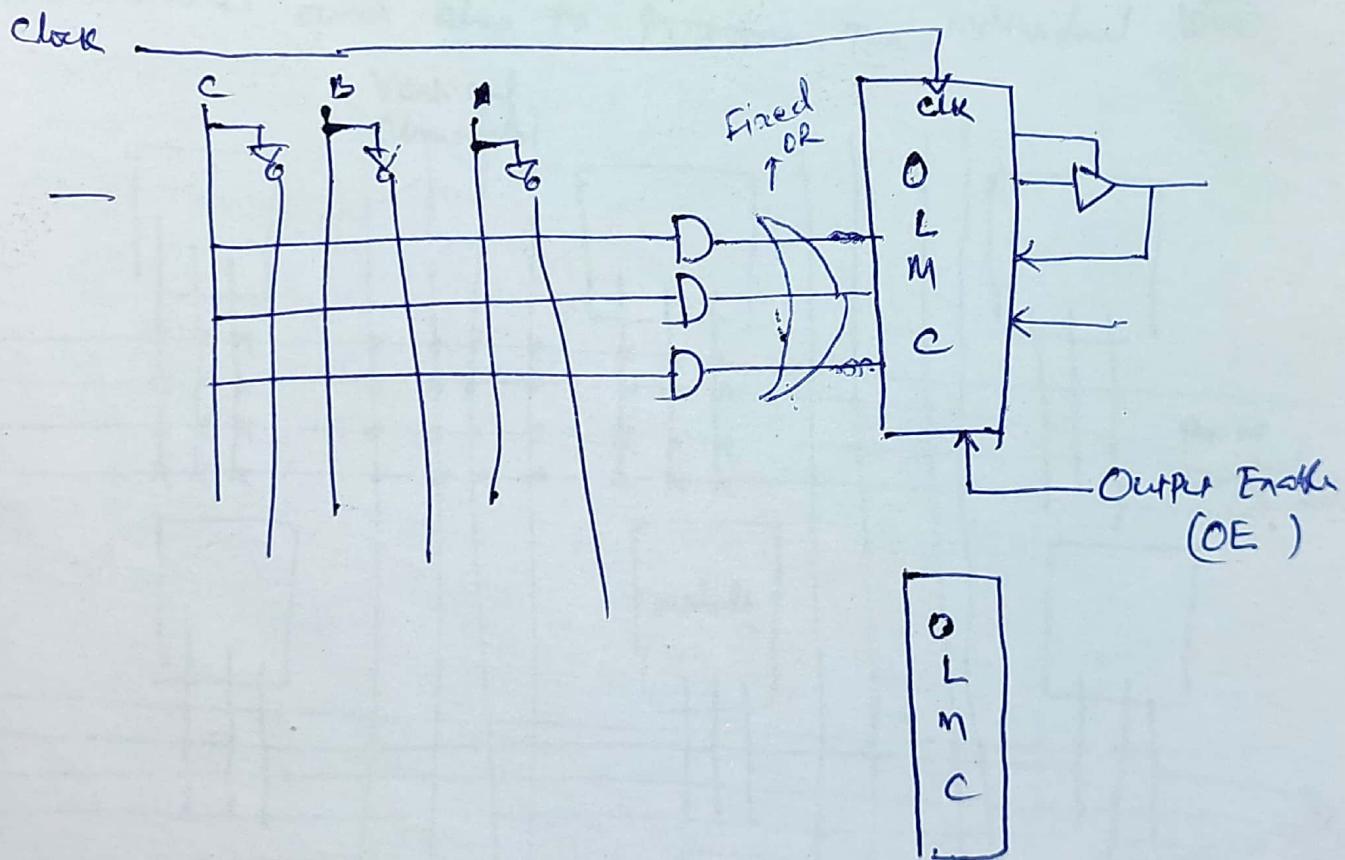
Generic Array Logic (GAL)

It contains Programmable AND array, a fixed OR array and an output stage.

The Output stage is called Output logic Macro Cell (OLMC). It contains MUX, flipflops, Tristate output buffers.

It is EEPROM i.e. Can be reprogrammed while still in the circuit board.

GTL 16 V 8A has 20 inputs & 8 outputs.



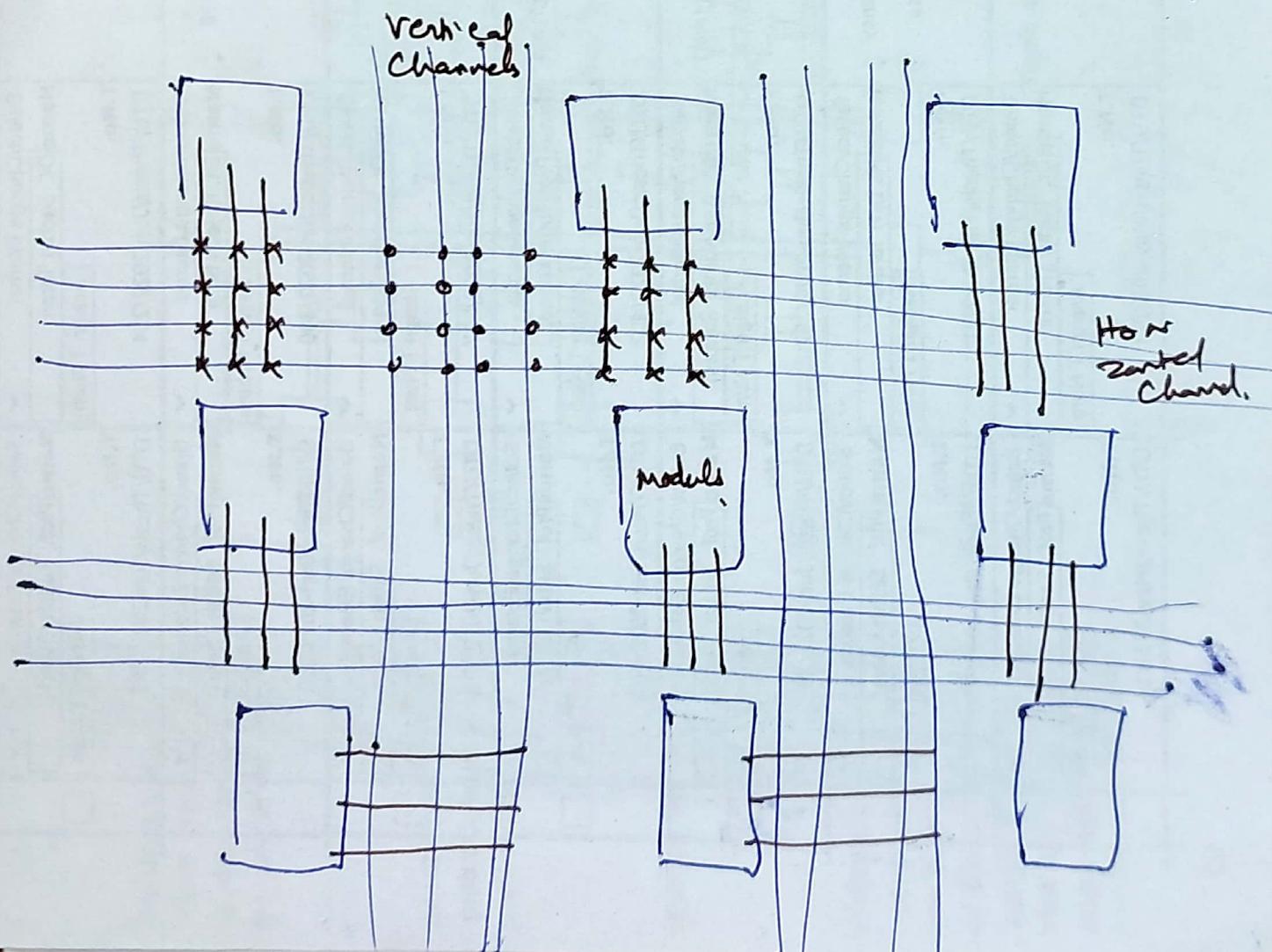
FPGA

It is a flexible architecture Programmable logic device. It is a single VLSI Circuit Constructed on a single Piece of Silicon.

It Consist of identically individually Programmable rectangular modules.

These modules are separated in both horizontal & vertical directions by horizontal & vertical metallic Conductors called CHANNELS.

Programmable links are used to interconnect the modules and also to program the individual ~~modules~~ modules.



becomes a high
saves typically
encoding for fast
one simple state

ndent levels of
patterns, which
programmed can be

unning voltage

rchitecture—the
, programmable
ave been added
re functionality
known as *field-*
implementation
ose of CPLDs.
modern standards

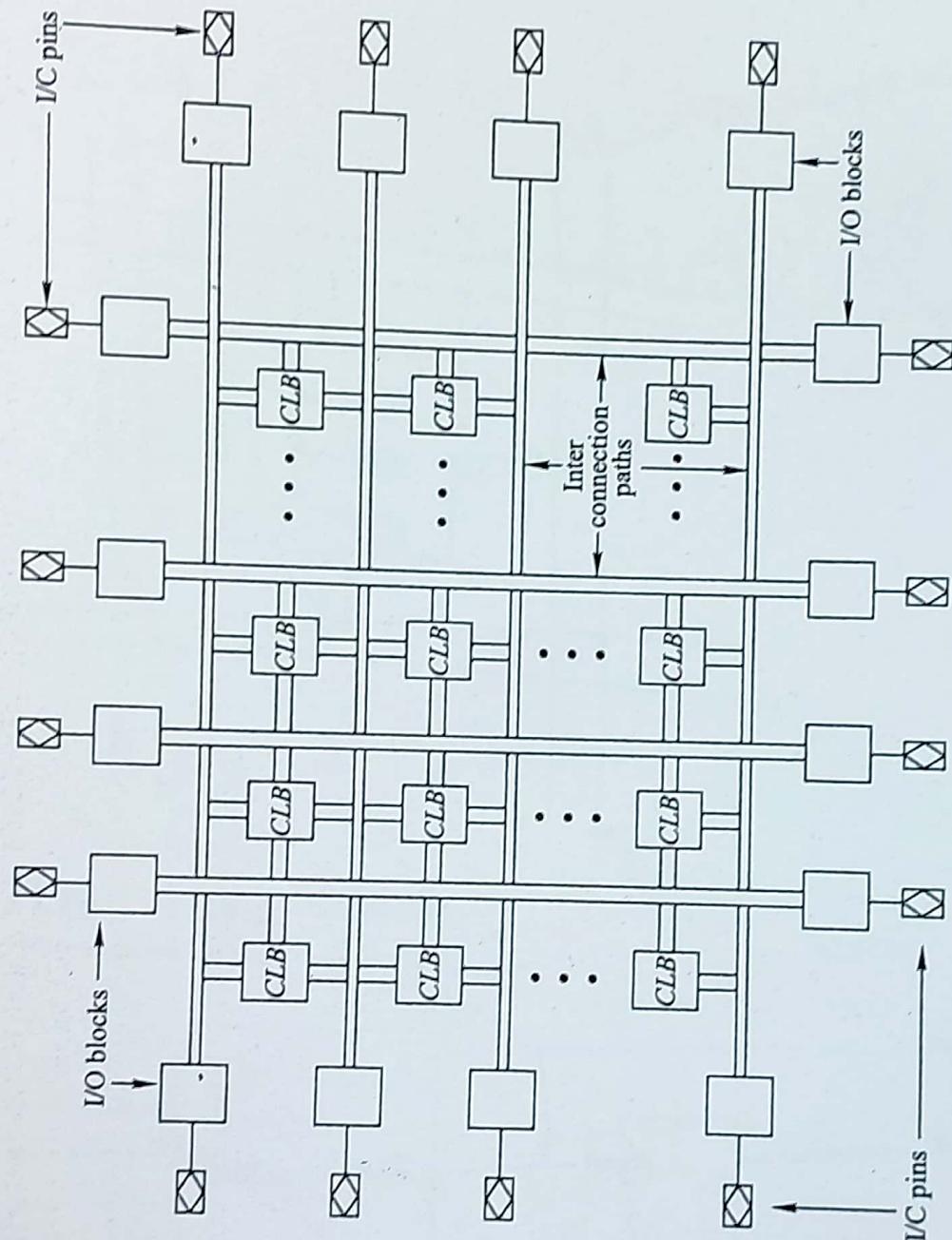
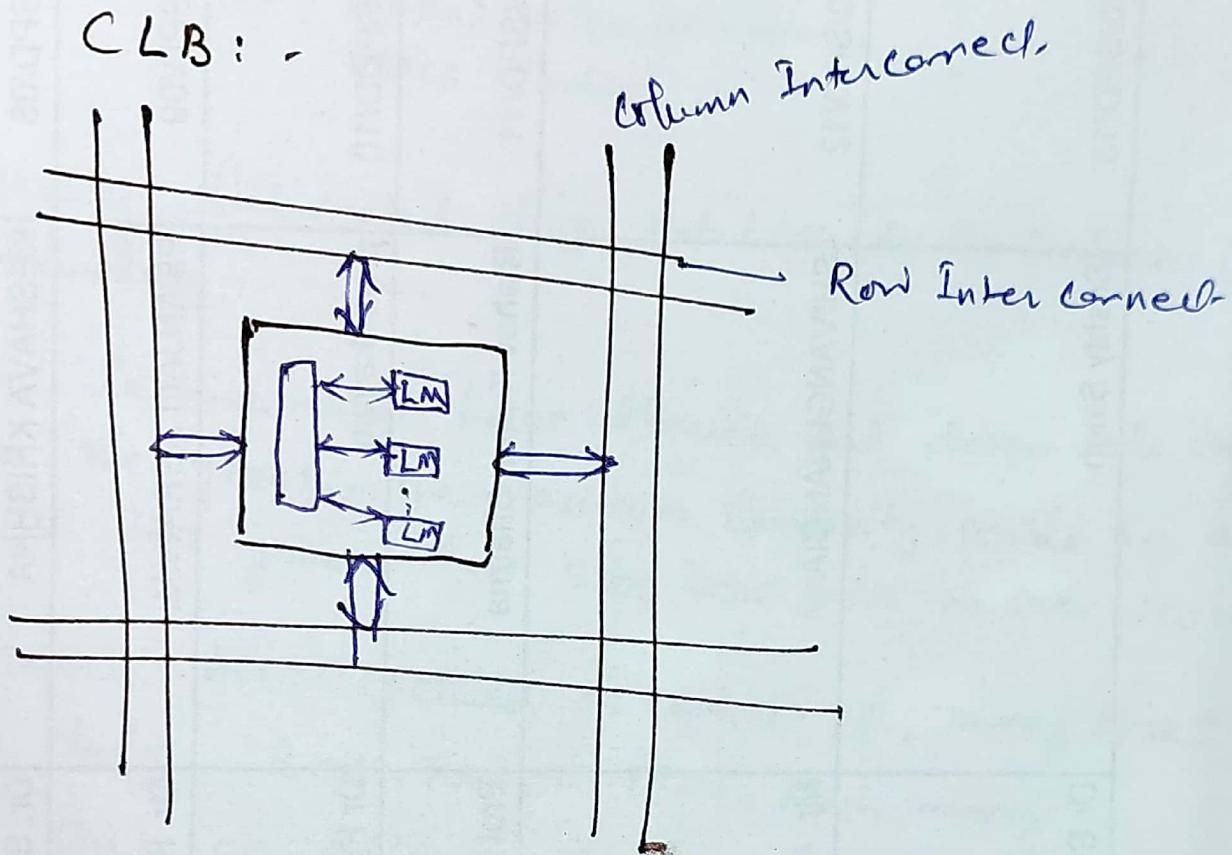


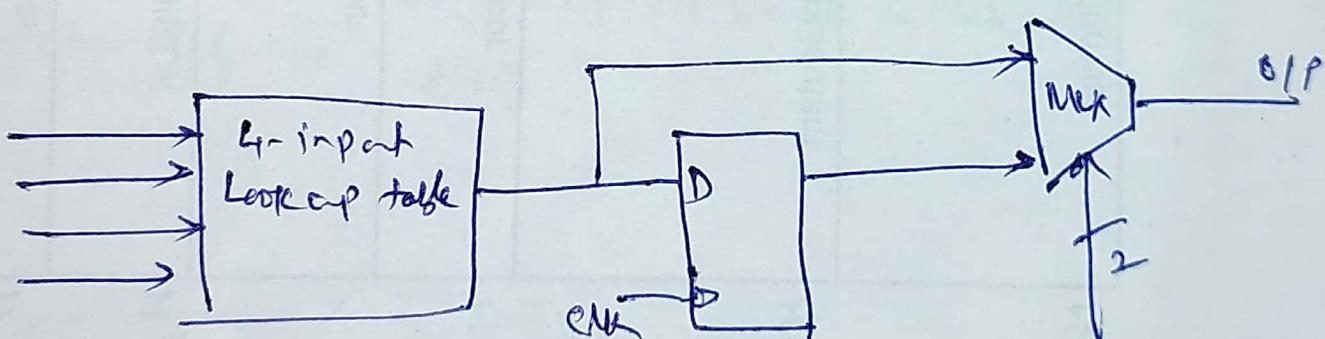
Fig. 12.45 Basic Architecture of FPGAs

It contains two types of modules, Combinational & sequential

CLB:-



LM: Logic module:



Xilinx in 1984,

Xilinx Vertex-5 family are latest.

ASIC

(Q)

- An application specific Integrated Circuit is an IC Chip Customised for a Particular Use, rather than intended for general purpose use.

Ex: Digital voice recorder,
Video Codec

- Usually fabricated Using Metal oxide Semiconductor (MOS) technology.
- As feature sizes have shrunk and design tools improved, over the years, the maximum complexity possible in an ASIC has grown from 5000 logic gates to over 100 million.
- Modern ASIC includes the entire CPU, memory blocks including ROM, RAM, EEPROM, Flash memory.
Such an ASIC is often termed as SOC (System on chip).
- It uses HDL like Verilog or VHDL to describe the functionality of ASICs.