Journey with Main Memory from Ram (1947) To 3d Xpoint (2015): An Overview

Narinder Kumar
S.B.S.B.M. University College Sardulgarh (Mansa),
Punjabi University Patiala, India

Amit Kumar Jain
M.Tech , Ambala Cantt.
Modern Institute of Engg. & Tech., India

Abstract: Memory is also main part of Computer, we can say that it is place without we can't initiate computer. In this paper we are will discuss the journey of memory till 2015.Beginners don't know the concepts behind main memory. Various types of memory are now available and we use them as per requirement. This paper elaborates various memory currently used in computer machines. It is a very challenging task to read a large number of papers, books and concerned articles to capturing essential information. The main aim of this paper is to know, learn and to provide basic terms of memory and which are we using.

Keywords: SSEM, PROM, TTL, ECL

I. INTRODUCTION TO MAIN MEMORY

Main memory : Main memory is memory which is internal to the computer which can be referred as a physical memory. To separate the meaning of main memory from mass storage device a special word is used i.e. known as 'main' which is used with memory as complete main memory. Word RAM is also used for main memory. Any type of execution which we want to perform through computer processor, data must present on main memory, due to this any type of instruction, file, and program is copied from storage device to main memory. How much data can be executed and how much data is available for read is depend on the size of main memory. We know that the amount of main memory is too small to load all data at once so a technique is used which is called swapping. In swapping technique, data or a part of data is swapped between main memory and storage memory when needed with the help of scheduling.
Technically we classify main memory in two categories which are RAM and ROM which are further divided.
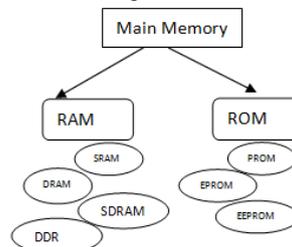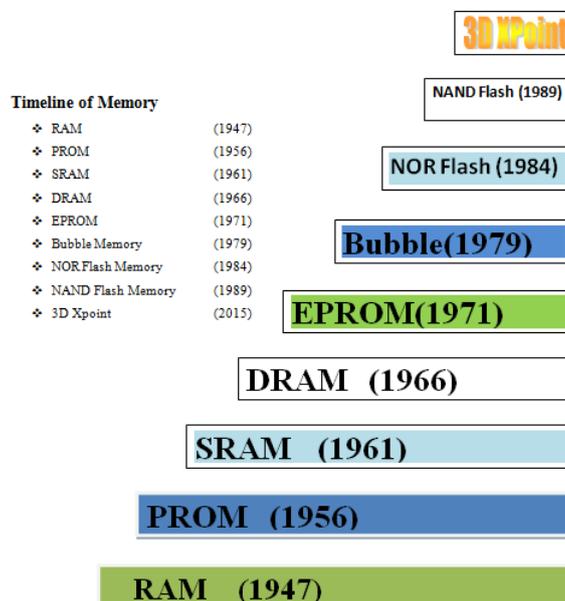

Fig: Main Memory Classification

II. CHANGES IN MAIN MEMORY

To provide a better computer changes are made time to time in hardware technologies and software technologies. As times passed changes are required in main memory also so the brief details of changes occurred in main memory are given below:

RAM (1947):

In 1947 William Tube starts a practical form of Random-access memory. Data is stored on the face of cathode ray tubes as electrically charged spots. The operations read and writes are performed by electron beams which are part of cathode rays tubes. Initially capacity of storage of RAM is very low; it cans store only few hundred bits near about to thousand bits. But in comparison vacuum tubes the size of RAM is very small   and these work very fast than vacuum tubes and more power efficient. Manchester Small Scale Experimental Machine (SSEM) computer is the first machine where first electronically stored memory program was implemented which is developed in University of Manchester in England by the William Tube.

Fig: William Tube

Williams Tube used a series of electronic CRT for digital storage. Its initial capacity is near about ~128 Byte.

PROM (1956)

The Programmable Read only memory is called PROM. This memory was invented by Wen Tsing Chow in 1956 at the request of United State Air Force, productized by American Bosch Arms Corporation in Garden City New York to come up with a more secure and flexible way to store data in computer. PROM is a type of read only memory that can be modified only once by a used with the help of special hardware/Machine called PROM programmer. An electronic current is supplied by machine to specific cell in the ROM that blows a fuse in them. This process called burning the ROM. The term "Burn" referring to the process of programming a PROM comes one of the original implementations, where the internal whiskers of diodes were literary burned with a current overload to produce a circuit discontinuity.
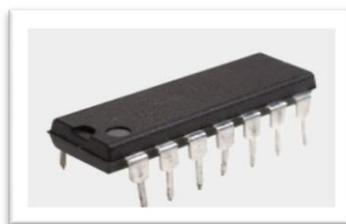

Fig: Programmable Read only Memory (PROM)

Initial capacity of PROM is ~128 Bytes. PROM is commonly used in video games consoles, mobile phones, and microcontrollers, Medical devices, HDMI and many other automotive electronic devices like

Fig: Devices where ROM is used.

SRAM (1961)

SRAM is Static Random Access Memory Static Random Access Memory is a type of Random Access memory which is introduced by Bob Norman in 1961. Flip Flops are used by this type of Semiconductor Memory, Flip Flop store bit...It is volatile memory means that data is lost when power is off. Transistor – transistor Logic means TTL and ECL are used.

Fig: SRAM

SRAM was productized by Jay Last. Its Initial capacity was ~128 Bytes. SRRAM is often used in Microprocessor registrar and primary Cache. It is Expensive and not dense enough for high capacity needs.

DRAM (1966)

Dynamic Random Access Memory is coined in 1966 at IBM Thoms J.W. Research Center.by Dr. Robert Dennard. DRAM is Dynamic Random Access memory. After that Honey well and productized by Intel in 1966.Data bits are stored in separated capacitors with integrated circuits in this memory. Capacitors are charged and discharged so 1 and 0 are sates on charged and discharged. Capacitors are refreshed periodically. DRAM typically used for CPU Cashes. It is volatile memory. On single chip millions of capacitors and transistors are placed. Transistors acts like witches that allows control circuitry on the memory chip read and capacitors or alters its states.

DRAM is simpler structurally, making it less expensive and capable much higher densities than SRAM. It is used in smart phones, personal computers, workstations, video games controller and tablets.

EPROM (1971)
Erasable programmable read only Memory (PROM) Is a type of non-volatile memory. Dov Frohman invented it in 1971. It is an integrated circuit composed by array of floating gate. There is a glass window on the top of Random access memory. From this glass window ultraviolet light is passed 10 to 20 minutes, which eras the Chip EPROM and can be reprogrammed again. Data can be read until its present on it.
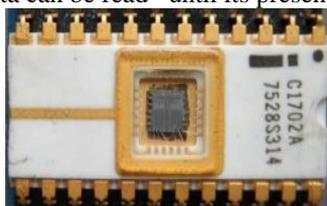

Fig: EPROM

Initial capacity of EPROM is ~128 bytes and it is most used in micro controllers. Normally it retains data ten to twenty years.

Bubble Memory (1979)

Magnetic type thin film material is used in bubble memory to hold small magnetized areas and it is a type of non-volatile memory. When a magnetic field is applied to circle area of this substances that is not magnetized in the same direction. The area is reduced to a smaller circle or bubble. Gadolinium gallium garnet material is used in this type of memory in which area can be easily magnetized in only one direction. Bubble memory uses a system similar to delay line memory system.

Fig: Bubble Memory

The initial capacity of Bubble memory was ~128 Bytes. First Andrew Bobeck work on this type of memory and it is productized by Texas instruments. It's used in video games system.

NOR Flash Memory (1984)

NOR flash memory was invented by Dr.Fujio Masuoka and productized by Intel in 1984. Name flash is suggested by doctor's colleagues. It is a type of non-volatile storage technology that does not require power to retain data. In NOR flash memory one end is connected to bit line and other is conned to ground directly due to this it act like NOR gate so its arrangement is called NOR flash. Default state is logically equivalent to binary value 1 because current will flow through the channel under application of an appropriate voltage to the control gate. It can be reprogrammed also.

Fig: NOR Flash Memory

NOR Flash memory is used in embed applications. Initial capacity was ~1MB. This type of memory is used in mobile phones, washing machines, mp3 players, DVD payers, digital Camera etc.

NAND Flash Memory (1989)

NAND flash memory stores data in an array of memoy cells made from foating-gate transistors.Insulated by an oxide layer are two gates, the Control gate and the floating gate. NAND flash memory is invented by Dr.Fujio Masuoka and productized by Toshiba in 1989. Write cycles are used by NAND which are in finite number.
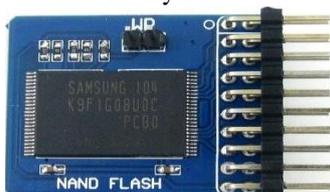

Fig: NAND Memory

Tunnel injection concept is used in NAND flash memory for wiring and tunnel release for erasing. Extra level of addressing is used in NAND flash memory by replacing single transistor with serial-linked grouped. NAND flash might address it by page, word and bit. NAND flash memory commonly used for high speed, high capacity storages in personal computers, Tablets, smart phones, digital Cameras etc. its initial capacity was ~16MB.

III. INTRODUCTION TO 3D X POINT MEMORY

3D X point memory which is pronounced three dee cross point memory. It is a new type of non-volatile memory jointly developed by Micron Technology and Intel Corporation and released in 2015.It is based on SSD. The first commercial solid state drives are yet to hit the market. 3D Xpoint is said to be up 1,000 times fasted and more durable than NAND flash storage that's currently used in mobile devices and solid state drives. After 25 year a new memory chips to come up in market. It is also calling that it is a major breakthrough in memory process technology.
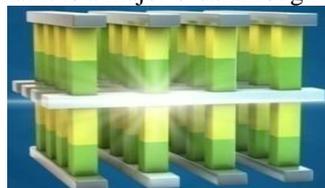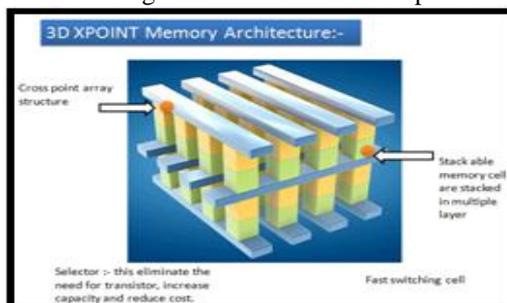

Fig: 3D Xpoint Memory

3D Xpoint was built to create high-performance, high storage capacity, memory solution and non volatile memory that was also affordable. Data can be stored closer to the processor and accessed quickly. 3D Xpoint memory is 10 times denser than the DRAM Chip and faster than Flash memory. Initial capacity ~16 GB per die across two memory layers, But the future will be able to increase the number of memory layers to improve system capacities.

History of 3D Xpoint Memory

Before 3D xpoint memory Intel and Micron had developed PCM(phase Challenge memory) technology which is non-volatile memory MarkDurcan of Micro said 3 D Xpoint Memory architecture differs from previously offering, and use materials that are faster and more stable. Development of 3D Xpoit memory began around2012

3D Xpoint Memory Technology Concept

Before starting to discuss about this emerging technology first of all we will discuss few point which are related to memory and also have drawback. Like the SRAM,DRAM,PRAM etc. flash memoty and DRAM (dynamic random access memory). DRAM offers excellent latency and store each bit of data in the separate capacitor within the circuit , but consume lot of power and large cell sizes while working with the large data file. DRAM doesn't work well for permanent storage because the cells don't retain data when they're in an off state. To remove all these drawback one more memory come into existence and that was PRAM (Phase –Change Random Memory) which is nonvolatile memory and also faster than flash memory and DRAM.in the PRAM the chalcogenide glass change between the two stae , polycrystalline and amorphous whaen the current is passed due to this heat is produced and phase is change of the substance so it called phase-changed memory.this memory have low resistance and electron can move easily through the crystalline structure and give high speed .but on the few advantage like (fast read performance , erase and write performance , scalability) there is also disadvantage like to detect the multiple bits and to make the commercialized.

So to overcome all these problem Intel and Micron came up with new advanced technology 3D XPOINT Technology . Unlike NAND, 3D Xpoint memory can write data into much smaller areas as compare to NAND flash memory. NAND flash memory required large block of memory to store the data. but IN 3D XPOINT Technology is non volatile storage and its structure is more efficient.. the way to store the data is very simple concept that we will discuss later and according to Intel and Micron this technology work 1,000 times faster than NAND flash memory, and on these days this is the most popular non-volatile memory in the marketplace .it provide the several feature like the performance, density,power,non volatile and the cost .intel also mention that use of this technology is very simple , user friendly and also interactive in social media and collaboration .For the gamming purpose it is most effective and reliable.

Architecture of 3D xpoint memory is given below:-
The 3D XPoint is designed to overcome the shortcoming in the NAND and DRAM memory technology , 3D XPoint is innovative, transistor-less cross point architecture creates a three-dimensional checkerboard where memory cells are shown in the diagrams, allowing the cells to be addressed individually. As a result, data can be written and read in small sizes, leading to fast and efficient read/write processes.
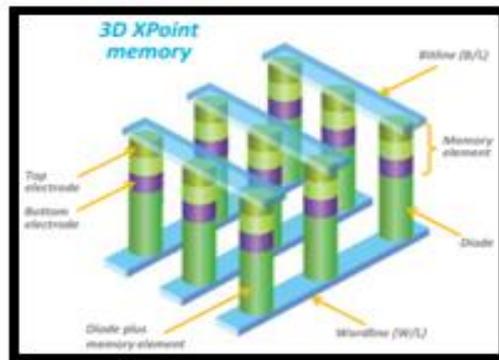
Cross Point Array
Cross-point memory array offers high device density , The three-dimensional cross-point array structure suitable for use with resistance-change non-volatile memory. The resistance-change cell serves as both the access element and the memory element, eliminating the need for individual selection devices..
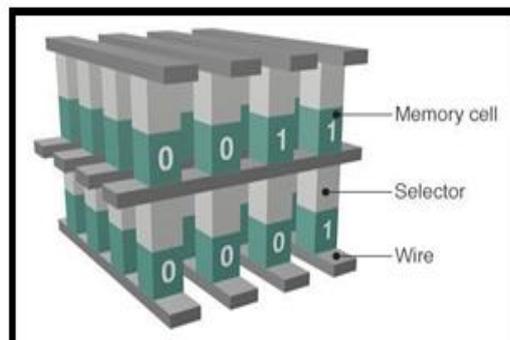
Stackable
        The initial technology stores 128Gb per die across two stacked memory layers. Future generations of this technology can increase the number of memory layers and/or use traditional lithographic pitch scaling to increase die capacity.

Selector
        Memory cells are accessed and written or read by varying the amount of voltage sent to each selector. This eliminates the need for transistors, increasing capacity and reducing cost.

Fast Switching Cell
        With a small cell size, fast switching selector, low-latency cross point array, and fast write algorithm, the cell is able to switch states faster than any existing nonvolatile memory technologies today.

How it works:-
        The 3D XPoint structure consists of a selector and memory cell, which are embedded between two electrode (hence the "crosspoint" name). to maximize the intensity we design the whole grid into the 3D structure.Applying a specific voltage on the electrode that will activate a single selector and enables the cell underneath to be either written (i.e. a bulk property change in the memory cell material) or read (allows the current through to check whether the memory cell is in high or low resistance state.

As the name suggests, the cells can be stacked in a 3D fashion to further improve density and the first generation die that is currently sampling a a two-layer design. Two layers does sound few compared to 3D NAND that is already at 32 layers with 48 being close to production, but the way 3D XPoint array is built is fundamentally quite different.
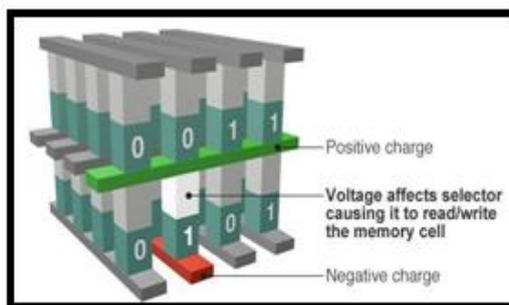
3D NAND is manufactured by first depositing alternating layers of conducting and insulating materials on top of each other. It's only after all layers have been deposited that the "cell towers" are lithographically defined, followed by a high aspect ratio etch with the hole being filled by channel material to gain access to a memory cell in each layer. In 3D XPoint, however, each layer needs to be lithographically patterned and etched (i.e. repeating the same process for each layer) before another layer can be deposited, which does take away some of the economic benefits that 3D NAND has (i.e. very few lithography steps), but the 3D XPoint approach still provides higher density than what pure lithography based scaling would.

Intel and Micron said that scaling in future will happen through both lithography and 3D stacking of layers. Being scalable both horizontally and vertically is the key in enabling future-proof scalability because traditional Argon Fluoride based immersion lithography with multi-patterning is getting quite uneconomical at ~10nm and there is still no clear successor in the pipeline..

In theory, 3D XPoint also supports multiple bits per cell, but that's not a path Intel and Micron are pursuing at the moment. While it's relatively easy to demonstrate multiple resistance levels in a lab, it's far more difficult to produce tens of thousands of wafers with each die having the necessary characteristics for proper multi-level cell operation. For comparison, it took nearly two decades before a second bit per cell was introduced to NAND, so for now Intel and Micron will focus on lithography and 3D scaling to increase density and cost efficiency, but multi-level cells may become a viable alternative in the future.

One of the big architectural differences to NAND is the fact that 3D XPoint is accessible at the bit-level. In NAND a whole page (16KB for the latest nodes) had to be programmed at once in order to save just one bit of data and to make matters worse you could only erase at the block level (a couple hundred pages at least). As a result, NAND requires sophisticated garbage collection algorithms for efficient performance, but regardless of the level of sophistication there is still performance degradation as a drive enters steady-state because of the inherent read-modify-write cycle that is needed to erase invalid pages within a block. With each cell being individually accessed 3D XPoint doesn't necessarily require any garbage collection to work effectively, which simplifies the controller and firmware architecture, and even more importantly enables higher performance and lower power consumption

IV. WHY IT DIFFER FROM OTHERS TECHNOLOGY?

Intel and Micron Technology of 3D XPoint memory, which will sit somewhere between DRAM and NAND flash in future systems, has everyone thinking about the architectural, economic, and performance So what's different about 3D XPoint? The companies said they invented "unique material Compounds" and a unique cross-point architecture that is 10 times denser than conventional memory and is able to scale in ways PCM cannot. Additionally, XPoint's memory cells are written to or read by varying the amount of voltage sent to each "selector," meaning if the voltage is high or low, it's a 1 or 0, respectively. "This eliminates the need for transistors, increasing capacity and reducing cost,". "The majority of resistive RAMs out there today are filamentary, so you have this statistical variation in how the filaments form each time you program it. On top of that it's difficult to scale and it's very slow," Meyers said. Intel and Micron aren't yet talking about the materials they're using to create the new memory. "The companies carefully skirted such questions. They were intentionally unclear about what it is," Handy said. Gregory Wong, an analyst with Forward Insights, said Intel and Micron appear to be about two years ahead of other memory makers in the development of cross-point memory.


3D XPoint technology uses its new material to switch the resistance state, so it doesn't rely on less reliable and more expensive elements, such as Memristor's titanium dioxide and platinum films or PCM's silver filaments, which wear out over time. And, in that one difference, Intel and Micron said they were able to bring their product to the manufacturing phase. "In addition, the combination of the architecture and the unique set of materials in

both the memory cell and selector enable 3D XPoint technology to achieve increased density, along with significant improvements in performance and endurance," Micron says.

3D Xpoint in future

The second-generation 3D XPoint memory will be introduced in 2016, just like the second-generation 3D NAND. The "new memory type B" is expected to be announced in 2017.
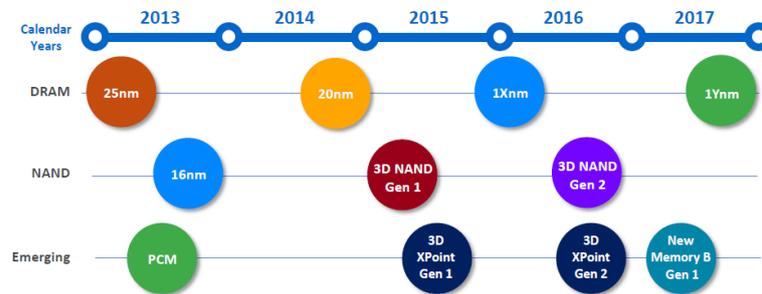
The differences between the first and the second generations of 3D XPoint are unknown. The first-gen 3D XPoint implementation is tens or hundreds times faster than NAND flash (in terms of read/write speeds, latency and IOPS peformance) and is dramatically more durable. One of the things where NAND flash excels the new type of memory is capacity. Modern NAND flash memory chips can store up to 256Gb of data, whereas the first 3D XPoint chips will store 128Gb. It is possible that with the second-generation 3D XPoint developers will focus on improving capacities rather than trying to further increase performance (still, keep in mind that increases of density usually also increase performance). Higher capacities will help to make solid-state drives featuring 3D XPoint a little cheaper (in terms of per-gigabyte costs), which will expand their market reach.

Features Technology 3D XPoint:
➢ Perpendicular wires combined 128 billion memory cells. Each memory cell stores one bit of data. This allows for high speed and high density.
➢ In addition to the location in the cross-shaped structure, memory cells arranged in several layers. The original technology allows you to store 128 GB on a single die for the two layers of memory. Future generations of technology will increase the number of layers to scale capacity.
➢ Access and read or write in the memory cells are made by changing the voltage directed to each selector. This eliminates the need for the transistors, which increases capacity and reduces the cost.
➢ Due to the small size of the cells, high-speed selectors, low latency and fast recording cell can switch state faster than any non-volatile memory technology. Trial delivery of products based on the technology 3D XPoint for individual customers will begin later this year. In addition, Intel and Micron * developing its own products based on this technology.

The speed of technology 3D XPoint

Memory   Technology   3D   XPoint   to   1   thousand.   Times   faster   than   memory   NAND.
• The average travel time to work of the US population would be reduced from 25 minutes to 1.5 seconds.
• Flight   from   San   Francisco   to   Beijing   would   take   about   43   hours   instead   of   12.
• The Great Wall could be built in 73 days instead of 200 years.

The delay hard drives and NAND memory is measured in microseconds and memory 3D XPoint – in nanoseconds (one billionth of a second).
In the time it takes the hard drive to overcome the basketball court, memory NAND overcome the marathon distance, and 3D XPoint can circumnavigate the globe.
If we compare the data storage on a computer travel:
• Hard   drives   could   carry   you   from   New   York   to   Los   Angeles   for   4   days   (4000   km).
• During   this   time,   SSDs   could   take   you   to   the   moon   (386,000   km).
• During   this   time,   the   memory   3D   XPoint   could   take   you   to   Mars   and   back   (about   450   million   kilometers).collection of the Library of Congress (20 terabytes) 2 times every day. After 5 years, the volume would correspond to 1.46 billion standard archival cabinets in which to store text documents.

V. CONCLUSIONS
The new architecture for advanced memory technology

In the technical and research area, the demand of 3D XPoint technology sharply increase because of its high speed/performance , its non-volatile property , less expansive and one of the most things that it meet the   user requirement.. The technology uses a new class of non-volatile memory which significantly reduces delays, enabling to store a larger amount of data close to the processor.
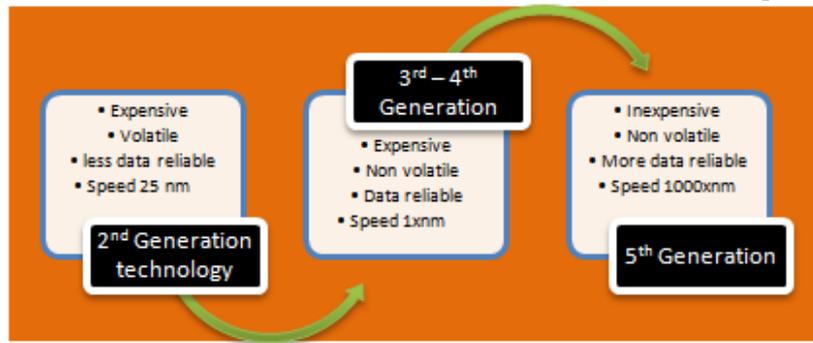
The innovative cross-architecture creates a three-dimensional "cube like chessboard" on which a memory cell located at the intersection of numerical lines, electrode and bits line, which allows for an independent right to perform addressing. As a result, data can be written and read with small latency and memory block size also reduced as compared to others memory(it mean that to store data it needed less memory), one more point to be noted that its processing speed is faster than others memory and also very effective in gamming environment or in high speed data processing calculation .

REFERENCES

[1]   http://www.eetimes.com
[2]   http://www.intelsalestraining.com
[3]   http://www.micron.com
[4]   https://en.wikipedia.org
[5]   https://www.youtube.com
[6]   www.intel.com
[7]   Design and performance Evaluation of a Main Memory Relation Database System by Tobin Jon Lehman
[8]   Fundamentals of Digital logic and Microcomputer Design By M. Rafiquzzaman
[9]   How Computer works: Processors and main memory by Roger Youbg
[10]   In-Memory Data Management: Technology an dapplications by Hasso Plattner, Alexender Zeier
[11]   The Cache Memory Book by Jim Handy
[12]   The Essentials of Computer Organization and Architecture by Linda Null, Pennsylvania State University Linda Null, Julia Lobur
[13]   Vertical 3D Memory Technologies by Betty Prince