Imperial Journal of Interdisciplinary Research (IJIR)
Vol-2, Issue-10, 2016
ISSN: 2454-1362, http://www.onlinejournal.in

A Study about Non-Volatile Memories

Dileep Kumar
Department of Information Media, The University of Suwon, South Korea.

***Abstract**: This paper presents an upcoming nonvolatile memories (NVM) overview. Non-volatile memory devices are electrically programmable and erasable to store charge in a location within the device and to retain that charge when voltage supply from the device is disconnected. The non-volatile memory is typically a semiconductor memory comprising thousands of individual transistors configured on a substrate to form a matrix of rows and columns of memory cells. Non-volatile memories are used in digital computing devices for the storage of data. In this paper we have given introduction including a brief survey on upcoming NVM's such as FeRAM, MRAM, CBRAM, PRAM, SONOS, RRAM, Racetrack memory and NRAM. In future Non-volatile memory may eliminate the need for comparatively slow forms of secondary storage systems, which include hard disks.*

1. Introduction

Memory is divided into two main parts: volatile and nonvolatile. Volatile memory loses any data when the system is turned off; it requires constant power to remain viable. Most kinds of random access memory (RAM) fall into this category. Nonvolatile memory does not lose its data when the system or device is turned off. A nonvolatile memory (NVM) device is a MOS transistor that has a source, a drain, an access or a control gate, and a floating gate. In floating gate memory devices, charge or data is stored in the floating gate and is retained when the power is removed. All floating gate memories have the same generic cell structure. They consist of a stacked gate MOS transistor as shown in Figure 1. Floating gate devices are typically used in EPROM (Electrically Programmable Read Only Memory) and EEPROM's (Electrically Erasable and Programmable Read Only Memory) [1].

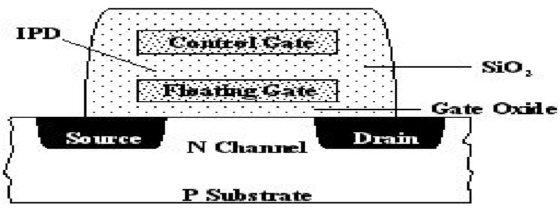

Figure 1. A typical floating gate memory structure.

Nonvolatile memory (NVM) retains its data without power. Today's ubiquitous Flash memory can be found in cell phones, personal digital assistants (PDAs), cameras, MP3 players, and a host of other portable consumer products. In a computer system, non-volatile memory is used for long-term storage of programs and data which seldom or never changes, and volatile memory devices are used for the short-term storage of program instructions and data during the execution of a program. According to the application functions, memories can be categorized into read only memory (ROM) and random access memory (RAM). As the name implies, the read only memory is only read accessible. A ROM device cannot be rewritten once it has been programmed. Embedded software applications use ROM to store embedded code and data records. The processor in an embedded software application retrieves each instruction from ROM and executes it. The random access memory can perform both write and read operations. Random access memory (RAM) also differs from ROM in that when power is disconnected from RAM, the data stored in random access memory is lost whereas when power is disconnected from ROM the data stored in read only memory remains. The read only memory is further categorized into a mask read only memory programmable read only memory (PROM), erasable programmable read only memory (EPROM), and electrically erasable programmable read only memory (EEPROM). Whereas, the random access memory can be further categorized into a static random access memory (SRAM) and a dynamic random access memory (DRAM). Static read/write random-access memory (SRAM) is a type of volatile memory in which the data, once it is written to a memory location, remains stored there as long as power is applied to the memory chip. The magneto resistive random access memory (MRAM) is an alternative memory device to dynamic random access memory (DRAM). An MRAM device uses magnetic orientations to retain data in its memory cells. MRAM devices are relatively fast, are nonvolatile, consume relatively little power, and do not suffer from a write cycle limitation. Programmable read only memory (PROM) allows the device manufacturer to program the embedded code. This allows for revisions in the code but still does not allow for modification or erasure of the ROM once it has been programmed. Erasable programmable read only memory (EPROM), electronically erasable programmable read only memory (EEPROM) and flash memory are a growing class of non-volatile storage integrated circuits based on floating gate transistors. Flash memory devices typically use a one-transistor memory cell that allows for high memory densities, high reliability, and low power consumption.

2. Background

One type of computer memory is Non-volatile memory, nonvolatile memory, NVM or non-volatile storage, which can retain the stored information even when not electrical supply is provided. Examples of non-volatile memory include read-only memory, flash memory, most types of magnetic computer storage devices (e.g. hard disks, floppy disk drives, and magnetic tape), optical disc drives, and early computer storage methods such as paper tape and punch cards. Non-volatile memory is typically used for the task of secondary storage, or long-term persistent storage. The most widely used form of primary storage today is a volatile form of random access memory (RAM), meaning that when the computer is shut down, anything contained in RAM is lost. Unfortunately, most forms of non-volatile memory have limitations that make them unsuitable for use as primary storage. Typically, non-volatile memory either costs more or performs worse than volatile random access memory. Several companies are working on developing non-volatile memory systems comparable in speed and capacity to volatile RAM. For instance, IBM is currently developing MRAM (Magnetic RAM). Not only would such technology save energy, but it would allow for computers that could be turned on and off almost instantly, bypassing the slow start-up and shutdown sequence. A number of conferences are held every year that focus specifically on non-volatile memory. One of the most prominent is the Non-Volatile Memory Technology Symposium.

Non-volatile data storage can be categorized in electrically addressed systems read only memory and mechanically addressed systems hard disks, optical disc, magnetic tape, Holographic memory and such. Electrically addressed systems are expensive, but fast, whereas mechanically addressed systems have a low price per bit, but are slow.

2.1. Electrically addressed systems

Electrically addressed non-volatile memories based on charge storage can be categorized according to their write mechanism:

Read only memory: ROM is a class of storage media used in computers and other electronic devices. Because data stored in ROM cannot easily and soon alter. In its strictest sense, ROM refers only to mask ROM, which is fabricated with the desired data permanently stored in it, and thus can never be modified. However, more modern types such as EPROM and flash EEPROM can be erased and re-programmed multiple times; they are still described as "read-only memory" because the reprogramming process is generally infrequent, comparatively slow.

Mask-programmed ROM: One of the earliest forms of non-volatile read-only memory, the mask-programmed ROM was prewired at the design stage to contain specific data; once the mask was used to manufacture the integrated circuits, the data was cast in stone (silicon, actually) and could not be changed.

Programmable ROM: PROM was invented in 1956 by Wen Tsing Chow, is a form of digital memory where the setting of each bit is locked by a fuse or antifuse. Such PROMs are used to store programs permanently. The Advantages are Reliability, Stores data permanently, moderate price, Built using integrated circuits rather than discrete components, and speed is between 35ns and 60ns.

Erasable PROMs: The EPROM was invented by Israeli engineer Dov Frohman in 1971. It is an array of floating-gate transistors individually programmed by an electronic device that supplies higher voltages than those normally used in digital circuits. Once programmed, an EPROM can be erased only by exposing it to strong ultraviolet light. There are two classes of non-volatile memory chips based on EPROM technology, UV-erase EPROM and OTP (one-time programmable) ROM.

Electrically erasable PROM: EEPROM is a type of non-volatile memory used in computers and other electronic devices to store small amounts of data that must be saved when power is removed, e.g., calibration tables or device configuration. Electrically erasable PROM's have the advantage of being able to selectively erase any part of the chip without the need to erase the entire chip and without the need to remove the chip from the circuit. While an erase and rewrite of a location appears nearly instantaneous to the user, the write process is slightly slower than the read process; the chip can be read at full system speeds. The limited number of times a single location can be rewritten is usually in the 10000-100000 range; the capacity of an EEPROM also tends to be smaller than that of other non-volatile memories.

Flash memory: Flash memory (both NOR and NAND types) was invented by Dr. Fujio Masuoka in 1980. The flash memory chip is a near contrast with the EEPROM; difference is that it can only be erased one block or "page" at a time. Flash memory has been widely used for high volume data storage in devices such as personal computers, personal digital assistants (PDAs), digital cameras, and cellular telephones. Program code and system data such as a basic input/output system (BIOS) are typically stored in flash memory devices for use in personal computer systems. A typical flash memory comprises a memory array, which includes a large number of memory cells. The cells are usually grouped into blocks. Each of the cells within a block can be electrically programmed in a random basis by charging the floating gate. A flash memory is provided with a high-voltage transistor and a low-voltage transistor for driving cells in view of a device's characteristic.

2.2. Mechanically addressed systems

Mechanically addressed systems have a low price per bit, but are slow.

Magnetic tape: Magnetic tape was first invented by Fritz Pfleumer in 1928 in Germany for recording sound. Magnetic tape is a source for magnetic recording generally consisting of a thin magnetizable coating on a long and narrow strip of plastic. Nearly all recording tape is of this type, whether used for recording audio or video or for computer data storage.

Hard disk drive: HDDs introduced in 1956 as data storage for an IBM accounting computer [3], which stores digitally encoded data on rapidly rotating platters with magnetic surfaces. Strictly speaking, "drive" refers to a device distinct from its medium, such as a tape drive and its tape, or a floppy disk drive and its floppy disk. In the 21st century, applications for HDDs have expanded to include digital video recorders, digital audio players, personal digital assistants, digital cameras and video game consoles.

Optical disk: The optical disc was invented in 1958. An optical disc is a flat, circular disc usually polycarbonate wherein data are stored in the pits or bumps in its flat surface — sequentially on the continuous, spiral track extending from the innermost track to the outermost track, covering the entire disc surface. The data are accessed in the disc when a special material illuminated with a laser diode that is aluminum. The pits distort the reflected laser light; hence, most optical discs characteristically have an iridescent appearance created by the grooves of the reflective layer. Write-once optical discs commonly use an organic dye, and re-writable discs use phase change alloys.

Millipede memory: Millipede is a non-volatile computer memory stored on nanoscopic pits burned into the surface of a thin polymer layer, read and written by a MEMS-based probe. Millipede storage technology is being pursued as a potential replacement for magnetic recording in hard drives, at the same time reducing the form-factor to that of Flash media.

Holographic data storage: Holographic data storage is a potential replacement technology in the area of high-capacity data storage currently dominated by magnetic and conventional optical data storage. Magnetic and optical data storage devices rely on individual bits being stored as distinct magnetic or optical changes on the surface of the recording medium. Additionally, whereas magnetic and optical data storage records information a bit at a time in a linear fashion, holographic storage is capable of recording and reading millions of bits in parallel, enabling data transfer rates greater than those attained by optical storage.[2]

3. Upcoming Non-Volatile Memories

Most of famous memory technologies have certain shortcomings: DRAMs are volatile means needs the power supply. Flash memory technologies are non-volatile but they have limited write endurance and low write speed. SRAMs are fast while these are volatile and comprises of large cell sizes. So related with these current technologies performance, now we have need of upcoming non-volatility memory technologies with high speed, high write endurance and a small size. Below we will discuss detailed overview of upcoming non-volatile memory technologies; the focus here is on the most widely pursued technologies: FeRAM, MRAM, CBRAM, PRAM, SONOS, RRAM, Racetrack memory, NRAM and PCRAM. These kinds of made of new materials not used ever before in IC technologies. Comparison of FRAM with other memory products is shown in Figure 2.

3.1. Ferroelectric random access memory (FeRAM):

A ferroelectric random access memory (FRAM) is a nonvolatile semiconductor memory device capable of operation without the need for refresh as in a dynamic random access memory (DRAM) device. A ferroelectric random access memory uses a ferroelectric capacitor as a storage element of each memory cell. Each ferroelectric memory cell stores a logic state based upon electric polarization of its ferroelectric capacitor. Ferroelectric random access memory (FRAM) devices are "nonvolatile" memory devices because they preserve data stored therein even in the absence of a power supply signal. The nonvolatile nature of a ferroelectric memory cell is a direct consequence of using a ferroelectric material as the dielectric of the cell's capacitor. A ferroelectric material has ferroelectricity. Structure of a FeRAM cell and Structure of a 1 transistor FeRAM cell and its working mechanism is shown in Figure 3 and Figure 4 respectively. The ferroelectricity is a physical property in which if an external voltage is applied to electric dipoles arranged in the ferroelectric material, a spontaneous polarization of the electric dipoles is generated. The FRAMs using the ferroelectric material are largely classified into two types; a first type which operates by detecting a change in a charge amount stored in a ferroelectric capacitor, and a second type which operates by detecting a change in resistance of a semiconductor due to spontaneous polarization of the ferroelectric material. In the FRAM, a memory cell is composed of a ferroelectric capacitor and an access transistor and stores logical data `1` or `0` depending on an electrical polarization state of the ferroelectric capacitor. When a voltage is applied across the ferroelectric capacitor, a ferroelectric material is polarized according to the direction of an electric field. FRAM can be used as main memory in various electronic equipment having file storage and search functions, such as portable computers, cellular phones and game machines, or as a recording medium for voice or images.

Memory products	*1 SRAM	*2 DRAM	*3 E2PROM	*4 FLASH	*5 FRAM
Memory type	Volatile back-up	Volatile	Non -volatile	Non -volatile	Non -volatile
Cell structure	6T	1T/1C	2T	1T	1T/1C 2T/2C
Read cycle (ns)	12	70	200	70	110
Internal write voltage (V)	3.3	3.3	20 (supply voltage 3.3V)	12 (supply voltage 3.3V)	3.3
Write cycle	12ns	70ns	3ms	1 sec.	110ns
Data write	Overwrite	Overwrite	Erase + Write	Erase + Write	Overwrite
Data erase	Unnecessary	Unnecessary	Byte (64 byte page)	Sector (8K /16K /32K /64K)	Unnecessary
Endurance	∞	∞	1E5	1E5	1E10 to 1E12
Stand-by current (μA)	7	1000	20	5	5
Read operation current (mA)	40	80	5	12	4
Write operation current (mA)	40	80	8	35	4

Notes: *1: 512K x 8bit, *2: 2M x 8bit, *3: 8K x 8bit, *4: 1M x 8bit, *5: 8K x 8bit

Figure 2. Comparison of FRAM with other memory products

Features of FRAM: There are conventional nonvolatile memories since as EEPROM and Flash. However, on the demand of high-speed and low-power consumption and high-rewriting endurance, FRAM has the superior performance as compared with those nonvolatile memories.

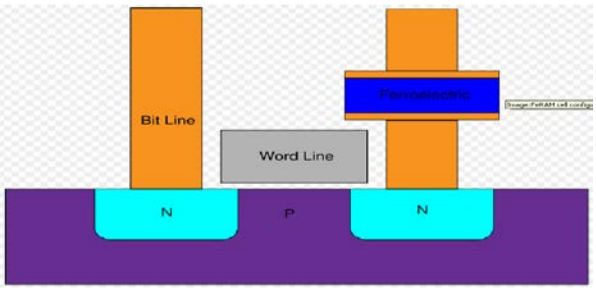

Figure 3. Structure of a FeRAM cell.

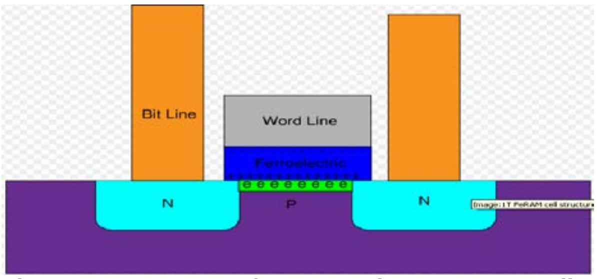

Figure 4. Structure of a 1 transistor FeRAM cell and its working mechanism.

3.2. Magnetoresistive Random Access Memory (MRAM):

A magnetoresistive random access memory (MRAM) is a non-volatile memory device using magnetic elements having magnetoresistance effects in a memory cell. The magnetoresistive random access memory (MRAM) has been developed as a next generation memory device capable of replacing a conventional dynamic random access memory (DRAM) in which it has a fast data write speed but data stored inside is erased once an electric power is off, and a flash memory having a data write speed 1,000 times slower than the DRAM. A magnetoresistance effect is a phenomenon that occurs when a magnetic field is applied to a ferromagnetic material, electric resistance changes in accordance with the direction of magnetization of the ferromagnetic material. The MRAM has a plurality of memory cells. Each memory cell is a multilayered structure in which two layers of magnetic material are separated by a layer of non-magnetic metallic conducting material. Each memory cell comprises a memory element (e.g., a giant magnetoresistance (GMR) element or a magnetic tunnel junction (MTJ) element) in electrical communication with a transistor through an interconnect stack. An MTJ element has a structure formed by sandwiching an insulating film by two magnetic films having conductivity. Two states are created depending on whether the spin directions in the two magnetic films that sandwich the insulating film are parallel or anti-parallel. In each memory cell of an MRAM, an MTJ element serving as an information storage element is formed at the interconnection between a bit line and a word line. In a data write mode, a current is supplied to each of a selected bit line and a selected word line. Data is written in the MTJ element of the selected cell located at the intersection between the selected bit line and the selected word line by a composed magnetic field generated by the currents. An MRAM using a tunneling magnetoresistance effect (TMR) in a magnetic tunnel junction has a sandwich structure in which an insulating layer (tunnel barrier layer) is inserted between two ferromagnetic layers (MTJ). In the Figure 5 we have shown the Comparison of MRAM with other memory products. And Figure 6 is discussed for MRAM's Chip Design, while Figure 7 is simplified structure for MRAM Cell.

Comparison to Other Memory Technologies: Comparison of MRAM with other memory technologies suggests that it can be competitive in overall performance. Since MRAM is nonvolatile, it retains the data when completely turned off. System power can be significantly reduced compared to DRAM by shutting down the MRAM when inactive since there is no background refreshing required. The straightforward integration scheme used for MRAM makes it easier to embed. Comparison with SRAM shows that MRAM will compete favorably in cost because of its smaller cell size. It also is non-volatile, which is only available in more complex and expensive battery backup solutions for SRAM. When compared with Flash, MRAM achieves much better performance in write characteristics since no high-voltage tunneling mode is required and MRAM write cycle is much faster. MRAM consumes much less energy in a write cycle because the energy/bit is several orders of magnitude lower than Flash. In addition, MRAM endurance is unlimited, with no known or expected deterioration mechanism, while typical Flash endurance is only 105 write cycles.

	MRAM	SRAM	DRAM	Flash	FeRam
Read Speed	Fast	Fastest	Medium	Fast	Fast
Write Speed	Fast	Fastest	Medium	Low	Medium
Array Efficiency	Med/High	High	High	Med/Low	Medium
Future Scalability	Good	Good	Limited	Limited	Limited
Cell Density	Med/High	Low	High	Medium	Medium
Non-Volatility	Yes	No	No	Yes	Yes
Endurance	Infinite	Infinite	Infinite	Limited	Limited
Cell Leakage	Low	Low/High	High	Low	Low
Low Voltage	Yes	Yes	Limited	Limited	Limited
Complexity	Medium	Low	Medium	Medium	Medium

Figure 5. Comparison of MRAM with other memory products.

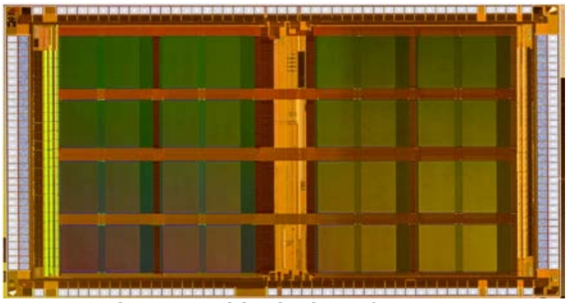

Figure 6. Chip design of MRAM.

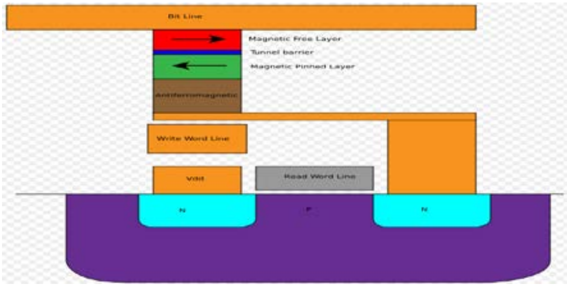

Figure 7. Simplified structure of an MRAM cell.

3.3. Conductive-Bridging RAM (CBRAM):

The CBRAM memory effect is based on a polarity-dependent, resistive switching at a low write threshold voltage Vth of ~250 mV with typically ~2 µA write current and an erase voltage threshold of ~80 mV. Key attributes are low voltage and current operations, excellent scalability, and a simple fabrication sequence. In Figure 8, the CBRAM switching mechanism is depicted. The ON-state of a CBRAM is achieved by applying a positive bias larger than the threshold voltage Vth at the oxidizable anode resulting in redox reactions driving Ag ions in the chalcogenide glass. This leads to the formation of metal rich clusters, which form a conductive bridge between both electrodes. The device can be switched back to the OFF state by applying an opposite voltage. In this case, the metal ions are removed, which in turn erases the conductive bridge.

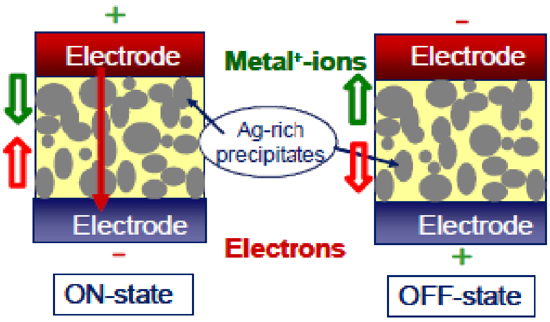

Figure 8. Schematic illustration of the CBRAM switching mechanism: ON state: Redox reaction drives Ag ions in chalcogenide glass, resulting in a conductive bridge. OFF state: Size and number of Ag-rich clusters is reduced breaking the conductive bridge [4].

CBRAM vs. RRAM: CBRAM differs from RRAM in that for CBRAM metal ions dissolve readily in the material between the two electrodes, while for RRAM, the material between the electrodes requires a high electric field causing local damage akin to dielectric breakdown, producing a trail of conducting defects (sometimes called a "filament").

Hence for CBRAM, one electrode must provide the dissolving ions, while for RRAM; a one-time "forming" step is required to generate the local damage.

The programmable metallization cell, or PMC, is a new form of non-volatile computer memory being developed at Arizona State University and its spinoff, Axon Technologies. PMC is one of a number of technologies that are being developed to replace the widely used flash memory, providing a combination of longer lifetimes, lower power, and better memory density.

3.4. Phase-change memory (PCM or PRAM):

This entirely new class of non-volatile memory brings together the best attributes of NOR, NAND and RAM. It simplifies memory and produces more capabilities within a single chip. It is also known as PCM, PRAM, PCRAM, Ovonic Unified Memory, Chalcogenide RAM and C-RAM. PCM holds promise to revolutionize applications through new capabilities and dramatic improvements to speed, durability, and power. PRAM is based on chalcogenide glass, which can be altered using the heat generation by an electric current. Heat changes the physical structure of the glass to either a crystalline or amorphous state. Each of these states has a distinct electrical resistance that is used to represent the ones and zeroes needed to represent stored data in binary terms. PRAM looks set to offer better read-write speed and durability than flash memory, which works by trapping electrons in a memory cell. Over time, electrons inevitably become trapped in these cells and can no longer be removed, rendering the memory chip useless. In the below Figure 9, A cross-section of two PRAM memory cells. One cell is in low resistance crystalline state; the other in high resistance amorphous state is mentioned.

Intel and other companies are counting on PRAM to replace both NOR and NAND flash memory to generate the demand required to produce the new memory chips in volume, and drive down costs.

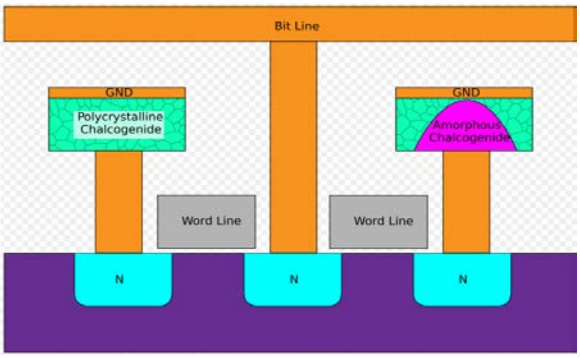

Figure 9. A cross-section of two PRAM memory cells. One cell is in low resistance crystalline state, the other in high resistance amorphous state.

PRAM Comparison with Flash: the new phase change memory is about 100,000 faster than Flash Memory. The write speed has gone from 1ms to 10ns per byte. With its faster merit also it is more durable. Each time data is written to flash memory, the process will cause damage and degradation. Most flash devices will be able to erase count is 10,000 - 100,000 writes per sector. Phase change memory on the other hand can withstand 100,000,000 writes per sector. It can also retain data without it becoming corrupted for much longer; estimates range from years to a decade, but those are conservative estimates. Compared to the advantages of phase change memory, the disadvantages are minor. Perhaps the biggest disadvantage is that due to the high temperatures involved in the manufacturing process, it is impossible to solder pre-programmed phase change memory chips onto a board. The chips must be programmed after they have been soldered into place. Some might also consider that higher voltages are also required to write data to the phase change memory to be a disadvantage. Figure 10 is tells about the PRAM test Setup and Figure 11 is about chip Design for PRAM.

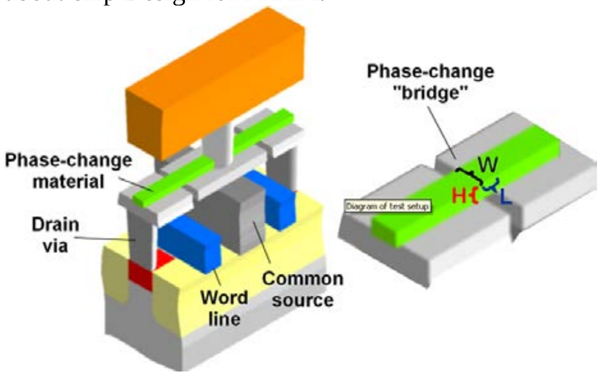

Figure 10. Diagram of test setup.

Phase change memory offers us faster and more durable memory. It could be ready for sale by the end of this year, though chances are that many manufacturers will wait a little before joining the cause. MP3 players and the like will probably continue to use flash memory for the time being. I Assuming that Intel manage to put their phase change memory to market within the next few months and Samsung come in early next year, we can expect the technology to proliferate as the older flash chips get used up.

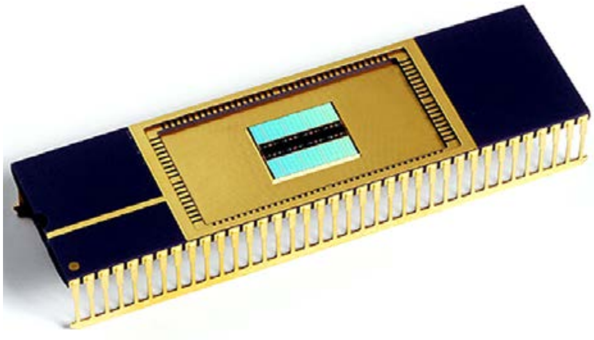

Figure 11. Chip Design of Phase Change RAM.

3.5. Silicon-Oxide-Nitride-Oxide-Silicon (SONOS):

Silicon-Oxide-Nitride-Oxide-Silicon is a kind of high-performance non-volatile computer memory. It is similar to the widely used Flash RAM, but offers lower power usage and a somewhat longer lifetime. SONOS is being developed as one of a number of potential Flash Memory replacements. Silicon/oxide/nitride/oxide/silicon (SONOS) is a memory technology that provides most of the advantages of floating gate without any of its disadvantages. It differs from floating gate in that the electrons used to store the data are distributed throughout a thin insulating layer of silicon nitride (Si3N4) rather than concentrated on a thick conducting layer of polysilicon. Figure 12 depicts a comparison of the cross-section of the cells. The stack height of floating gate can be twice the stacking of SONOS, primarily caused by the difference in the thickness of the storage layer as displayed below in Figure 12, and Schematic drawing of a SONOS memory cell is in Figure 13.

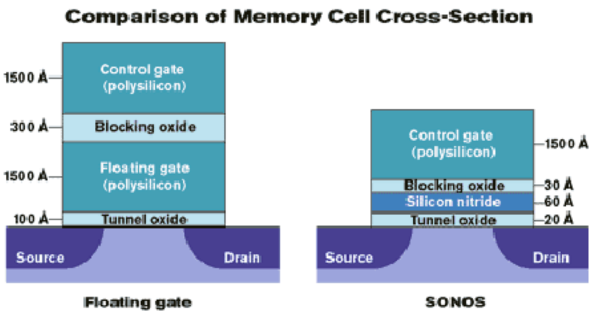

Figure 12. The stack height of floating gate can be twice the stacking of SONOS, primarily caused by the difference in the thickness of the storage layer.

Improvements in the technology have made SONOS a formidable competitor to floating gate in commercial-grade applications as well, with the latest SONOS devices offering faster read times, longer data retention and a significantly greater number of read/write cycles. Most importantly, SONOS doesn't have the scaling limitations of floating gate, so it may very well become the NVM of choice for the advanced applications of the future.

Alternatives to SONOS: SONOS is not the only alternative to floating gate technology. Recent developments have brought more exotic technologies into the limelight, including ferroelectrics (FRAM), magnetoresistive memories (MRAM) and phase-change memories (PCM). However, these technologies are still quite new and present a number of serious manufacturing hurdles. They also require unusual materials and processing equipment, adding significant costs to the manufacturing process. SONOS, on the other hand, uses materials that are readily available in any wafer fab, and the silicon nitride layer used to store the data can be deposited using standard processing equipment.

Description: SONOS "cells" consist of a standard NMOS transistor with an additional layering of insulators on the gate (the transistor's "switch"). The layering consists of an oxide layer approximately 2 nm thick, a silicon nitride layer about 5 nm, and a second oxide layer with a thickness between 5 and 10 nm. When the gate is biased positively, electrons from the source-drain circuit "above" the layer tunnel through the oxide layer and get trapped in the silicon nitride. This results in an energy barrier between the drain and the source, raising the threshold voltage Vt (the gate-source voltage necessary for current to flow through the transistor). The electrons can be removed again by applying a negative bias on the gate. A SONOS memory device is constructed by fabricating a grid of SONOS transistors along with a small amount of control circuitry. After storing or erasing the cell, the controller can measure the state of the cell by passing a small voltage across the source-drain pair; if current flows the cell must be in the "no trapped electrons" state, which is considered to mean "0". If no current is seen the cell is in the "1" state. The needed voltages are normally about 2 V for the erased state, and around 4.5 V for the programmed state.

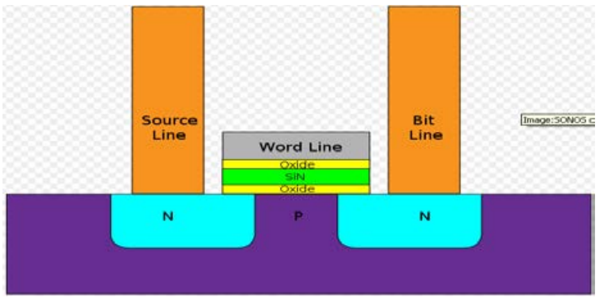

Figure 13. Schematic drawing of a SONOS memory cell.

Comparison with Flash: SONOS and Flash have same theory. But Flash requires the construction of a very high-performance insulating barrier on the gate leads of its transistors, often requiring as many as nine different steps, whereas the oxide layering in SONOS can be more easily produced on existing lines. SONOS, on the other hand, requires a very thin layer of insulator in order to work, making the gate area smaller than Flash. This allows SONOS to scale to smaller line width, with recent examples being produced on 40 nm fabs and claims that it will scale to 20 nm.[5] Flash devices can only be written to between 10,000 and 100,000 times, depending on the type. SONOS devices require much lower write power, typically 5 to 8 V, and do not degrade in the same way. SONOS does suffer from an unrelated problem, however, where electrons become strongly trapped in the ONO later and cannot be removed again. Over long usage this can eventually lead to enough trapped electrons to permanently set the cell to the "1" state, similar to the problems in Flash. However, in SONOS this requires on the order of a 100,000,000 write cycles, 1000 to 10,000 times better than Flash.

3.6. Resistive random-access memory (PRAM):

Resistive Random Access Memory (RRAM) is a non-volatile memory based on resistance switching caused by internal stoichiometry changes in compound materials. The basic idea is that a dielectric, which is normally insulating, can be made to conduct through a filament or conduction path formed after application of a sufficiently high voltage. RRAM based on transition metal oxide and compatible electrode metals for more than a million cycles of continuous use. Endurance of 10 yrs at 85°C has been demonstrated as well. Literature data are giving more indications that RRAM is closest to becoming a universal memory. Compared to PRAM, RRAM operates at a faster timescale (switching time can be less than 10 ns), while compared to MRAM, it has a simpler, smaller cell structure (a 4-8 F2 MIM stack). Compared to flash memory, a lower voltage is sufficient. Compared to DRAM, the data is retained longer (10 years). PRAM uses very simple cell structure as designed in Figure 14.

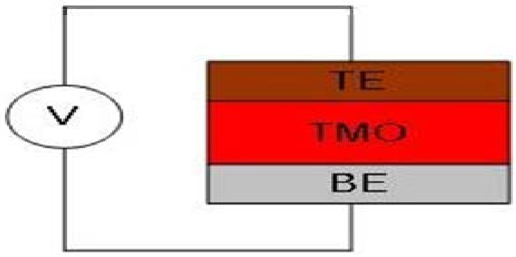

Figure 14. RRAM uses simple cell structure.

It is still difficult to predict the scalability of RRAM, when the underlying mechanism is not well understood yet. However, it is believed that if a filament is responsible, it would not exhibit direct scaling with cell size.[6] Instead, the current compliance limit (set by an outside resistor, for example) could define the current-carrying capacity of the filament.[7]

Applications: PRAM's applications are Cellular Phone, Embedded Memory, Data storage, Hard drive replacement and Stand-alone RAM.

3.7. Racetrack memory:

As both silicon-based microelectronic devices and HDDs are essentially two-dimensional (2D) arrays of transistors and magnetic bits, respectively, the conventional means of developing cheaper and faster devices relies on reducing the size of individual memory elements or data storage bits. An alternative approach is to consider constructing truly 3D devices. One such approach is "racetrack" memory (RM) [8], in which magnetic domains are used to store information in tall columns of magnetic material arranged perpendicularly on the surface of a silicon wafer as shown in Figure. 15.

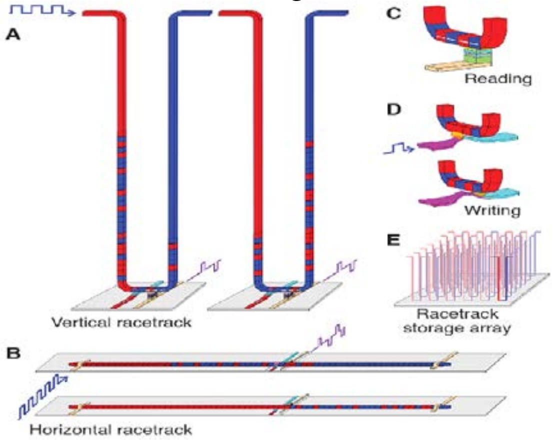

Figure. 15. The racetrack is a ferromagnetic nanowire, with data encoded as a pattern of magnetic domains along a portion of the wire.

(A) A vertical-configuration racetrack offers the highest storage density by storing the pattern in a U-shaped nanowire normal to the plane of the substrate. The two cartoons show the magnetic patterns in the racetrack before and after the domain walls (DW) have moved down one branch of the U, past the read and write elements, and then up the other branch.

(B) A horizontal configuration uses a nanowire parallel to the plane of the substrate.

(C) Reading data from the stored pattern is done by measuring the tunnel magnetoresistance of a magnetic tunnel junction element connected to the racetrack.

(D) Writing data is accomplished, for example, by the fringing fields of a DW moved in a second ferromagnetic nanowire oriented at right angles to the storage nanowire.

(E) Arrays of racetracks are built on a chip to enable high-density storage.

Comparison to other memory devices: Flash, in particular, is a highly asymmetrical device. Although read performance is fairly fast, especially compared to a hard drive, writing is much slower. Flash works by "trapping" electrons in the chip surface, and requires a burst of high voltage to remove this charge and reset the cell. In order to do this, charge is accumulated in a device known as a charge pump, which takes a relatively long time to charge up. In the case of "NOR" flash, which allows random bit-wise access like IBM Racetrack Memory, read times are on the order of 70 ns, while write times are much slower, about 2,500 ns. To address this concern, "NAND" flash allows reading and writing only in large blocks, but this means that the time to access any random bit is greatly increased, to about 1,000 ns. Additionally, the use of the burst of high voltage physically degrades the cell, so most flash devices allow on the order of 10,000 writes to any particular bit before their operation becomes unpredictable Wear leveling and other techniques can spread this out, but only if the underlying data can be re-arranged.

DRAM has a cell size of about 6 F^2, SRAM is much worse at 120 F^2. NAND flash is currently the densest form of non-volatile memory in widespread use, with a cell size of about 4.5 F^2, but storing two bits per cell for an effective size of 2.25 F^2. NOR is slightly less dense, at an effective 4.75 F^2, accounting for 2-bit operation on a 9.5 F^2 cell size.[9]

IBM Racetrack Memory appears to scale to much smaller sizes than any current memory device. In the vertical orientation (U-shaped) about 128 bits are stored per cell, which itself can have a physical size of at least about 20 F^2. No other near-term solid-stage technology appears to be able to scale anywhere near these densities, representing a storage density about 100 times that of Flash.[9] The caveat here is that bits at different positions on the "track" would take different times (from ~10 ns to nearly a microsecond, or 10 ns/bit) to be accessed by the read/write sensor, because the "track" is moved at fixed speed (~100 m/s) past the read/write sensor.

IBM Racetrack Memory is one of a number of new technologies aiming to replace Flash, and potentially offer a "universal" memory device applicable to a wide variety of roles. Other leading contenders include MRAM, PCRAM and FeRAM. Most of these technologies offer densities similar to Flash, in most cases worse, and their primary advantage is the lack of write endurance limits like those in Flash. Field-MRAM offers excellent performance as high as 3 ns access time, but requires a large 25 to 40 F^2 cell size. It might see use as a SRAM replacement, but not as a mass storage device. The highest densities from any of these devices is offered by PCRAM, which has a cell size of about 5.8 F^2, similar to Flash, as well as fairly good performance around 50 ns. Nevertheless, none of these can come close to competing with IBM Racetrack Memory in overall terms, especially density. For example, 50 ns allows about 5 bits to be operated in an IBM Racetrack Memory device, resulting in an effective cell size of 20/5=4 F^2, easily exceeding the speed-density product of PCM.

3.8. Nano-RAM (NRAM):

Nano-RAM is a proprietary computer memory technology from the company Nantero and NANOMOTOR is invented by University of bologna and California nano systems. NRAM is a type of nonvolatile random access memory based on the mechanical position of carbon nanotubes deposited on a chip-like substrate. In theory the small size of the nanotubes allows for very high density memories. Nantero also refers to it as NRAM in short, but this acronym is also commonly used as a synonym for the more common NVRAM, which refers to all non-volatile RAM memories. Nanomotor is a molecular motor which works continuously without the consumption of fuels. It is powered by sunlight. Figure 16 is simple NANORAM structure and Figure 17 is NANORAM with Carbon Nanotubes.

Storage in NRAM works by balancing the on ridges of silicon. Under differing electric charges, the tubes can be physically swung into one or two positions representing one and zeros. Once in position the tubes stay there until a signal resets them. The bit itself is not stored in the nano tubes, but rather is stored as the position of the nanotube. Up is bit 0 and down is bit 1. Bits are switched between the states by the application of the electric field.

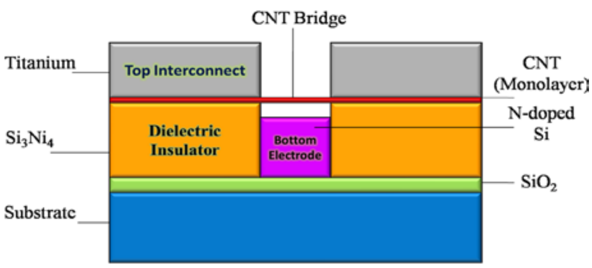

Figure 16. NANO-RAM

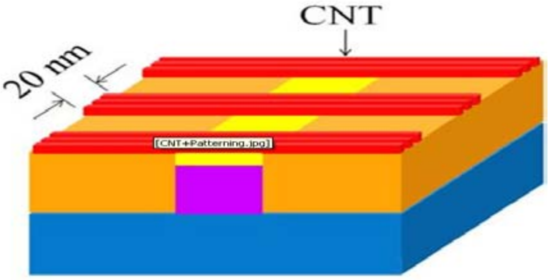

Figure 17. NRAM with Carbon Nanotubes (CNT)

Advantages: NRAM has a density, at least in theory, similar to that of DRAM. DRAM consists of a number of capacitors, which are essentially two small metal plates with a thin insulator between them. NRAM is similar, with the terminals and electrodes being roughly the same size as the plates in a DRAM, the nanotubes between them being so much smaller they add nothing to the overall size. However it seems there is a minimum size at which a DRAM can be built, below which there simply not enough charge is being stored to be able to effectively read it. NRAM appears to be limited only by the current state of the art in lithography. This means that NRAM may be able to become much denser than DRAM, meaning that it will also be less expensive; if it becomes possible to control the locations of carbon nanotubes at the scale the semiconductor industry can control the placement of devices on silicon. Additionally, unlike DRAM, NRAM does not require power to "refresh" it, and will retain its memory even after the power is removed. Additionally the power needed to write to the device is much lower than a DRAM, which has to build up charge on the plates. This means that NRAM will not only compete with DRAM in terms of cost, but will require much less power to run, and as a result also be much faster (write speed is largely determined by the total charge needed). NRAM can theoretically reach speeds similar to SRAM, which is faster than DRAM but much less dense, and thus much more expensive. In comparison with other NVRAM technologies, NRAM has the potential to be even more advantageous. The most common form of NVRAM today is Flash RAM, which combines a bistable transistor circuit known as a flip-flop (also the basis of SRAM) with a high-performance insulator wrapped around one of the transistor's bases. After being written to, the insulator traps electrons in the base electrode, locking it into the "1" state. However, in order to change that bit the insulator has to be "overcharged" to erase any charge already stored in it. This requires high voltage, about 10 volts, much more than a battery can provide. Flash systems thus have to include a "charge pump" that slowly builds up power and then releases it at higher voltage. This process is not only very slow, but degrades the insulators as well. For this reason Flash has a limited lifetime, between 10,000 and 1,000,000 "writes" before the device will no longer operate effectively. NRAM potentially avoids all of these issues. The read and write process are both "low energy" in comparison to Flash (or DRAM for that matter), meaning that NRAM can result in longer battery life in conventional devices. It may also be much faster to write than either, meaning it may be used to replace both. A modern cell phone will often include Flash memory for storing phone numbers and such, DRAM for higher speed working memory because flash is too slow, and additionally some SRAM in the CPU because DRAM is too slow for its own use. With NRAM all of these may be replaced, with some NRAM placed on the CPU to act as the CPU cache, and more in other chips replacing both the DRAM and Flash.

Comparison with other proposed systems: NRAM is one of a variety of new memory systems, many of which claim to be "universal" in the same fashion as NRAM – replacing everything from Flash to DRAM to SRAM. The only system currently ready for commercial use is ferroelectric RAM (FRAM or FeRAM). FeRAM adds a small amount of a ferro-electric material in an otherwise "normal" DRAM cell, the state of the field in the material encoding the bit in a non-destructive format. FeRAM has all of the advantages of NRAM, although the smallest possible cell size is much larger than for NRAM. FeRAM is currently in use in a number of applications where the limited number of writes in Flash is an issue, but due to the massive investment in Flash factories (fabs), it has not yet been able to even replace Flash in the market. Other more speculative memory systems include MRAM and PRAM. MRAM is based on a magnetic effect similar to that utilized in modern hard drives, the memory as a whole consisting of a grid of small magnetic "dots" each holding one bit. Key to MRAM's potential is the way it reads the memory using the magneto-restrictive effect, allowing it to read the memory both non-destructively and with very little power. Unfortunately it appears MRAM is already reaching its fundamental smallest cell size, already much larger than existing Flash devices. PRAM appears to have a small cell size as well, although current devices are nowhere near small enough to find if there is some practical limit.

4. Discussion

In this paper we have discussed an upcoming non-volatile memories (NVM) overview in detail with regarding their cell structures, architectures, advantages and comparisons with current memory technologies. We have also introduced the previous non-volatile memories in short overview. As Non-volatile memory devices are electrically programmable and erasable to store charge in a location within the device and to retain that charge when voltage supply from the device is disconnected. The non-volatile memory is typically a semiconductor memory comprising thousands of individual transistors configured on a substrate to form a matrix of rows and columns of memory cells. Non-volatile memories are used in digital computing devices for the storage of data. In this paper we have given introduction including a brief survey on upcoming NVM's such as FeRAM, MRAM, CBRAM, PRAM, SONOS, RRAM, Racetrack memory and NRAM. In future Non-volatile memory may eliminate the need for comparatively slow forms of secondary storage systems, which include hard disks.

5. Acknowledgements

This research was supported by “The University of Suwon”.

6. References

[1] Jitu J. Makwana, Dr. Dieter K. Schroder “A Nonvolatile Memory Overview”.

[2] "Holographic data storage." Retrieved on 2008-04-28.

[3] IBM 350 disk storage unit.

[4] Michael Kund and etal; “Conductive bridging RAM (CBRAM): An emerging non-volatile memory technology scalable to sub 20nm.”

[5] Samsung unwraps 40nm "charge trap flash" device.

[6] I. G. Baek et al.,IEDM 2004.

[7] C-Y. Lin et al., J. Electrochem. Soc., 154, G189-G192.

[8] S. S. P. Parkin, U.S. Patents 6,834,005, 6,898,132, 6,920,062, 7,031,178, and 7,236,386 (2004 to 2007).

[9] Parkin, et al., Magnetic Domain-Wall Racetrack Memory, Science, 320, 190 (11 April 2008), DOI: 10.1126/science.

[10] http://zenov.wordpress.com/2008/04/28/skool-art/

[11] http://en.wikipedia.org/wiki/Non-volatile_memory

[12]http://www.itri.org.tw/eng/EOL/research-and-development-detail.asp?RootNodeId=020&NavRootNodeId=02042&NodeId=0204223&ArticleNBR=1728

[13] http://www.smso.net/Non-volatile_memory

[14]http://www.electronics-manufacturers.com/products/computer-hardware/computer-memory

[15] Gerhard Müller(a), Nicolas Nagel (b), Cay-Uwe Pinnow(a), Thomas Röhr, “Emerging Non-Volatile Memory Technologies”.

[16] Anselmo Lastra, “Nonvolatile Memories and Programmable Logic”.

[17] Roberto Bez, Livio Baldi, Emilio Camerlenghi and Paolo Cappelletti, “An Overview of Flash Memories”.

[18] Daniel Elmhurst, “Non-Volatile Memories”.

[19] Stefan Lai, “Nonvolatile Memory Technologies: A Look into the Future”.

[20] Todd Wallinger, Simtek Corp., Colorado Springs, Colo, “SONOS Eases Non-Volatile Memory Integration in SoC”.

[21] Scott Wilkinson, “NANO-RAM” Jul 1, 2005.

[22] Sumner Lemon, “Intel set for first public demo of PRAM”.

[23] Nano-RAM - A proprietary Computer Memory Technology.

[24] Stuart S. P. Parkin, Masamitsu Hayashi, Luc Thomas,

“Magnetic Domain-Wall Racetrack Memory”.