International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:4, No:10, 2010

Exploring the Potential of Phase Change Memories as an Alternative to DRAM Technology

Venkataraman Krishnaswami, Venkatasubramanian Viswanathan

***Abstract*—**Scalability poses a severe threat to the existing DRAM technology. The capacitors that are used for storing and sensing charge in DRAM are generally not scaled beyond 42nm. This is because; the capacitors must be sufficiently large for reliable sensing and charge storage mechanism. This leaves DRAM memory scaling in jeopardy, as charge sensing and storage mechanisms become extremely difficult. In this paper we provide an overview of the potential and the possibilities of using Phase Change Memory (PCM) as an alternative for the existing DRAM technology. The main challenges that we encounter in using PCM are, the limited endurance, high access latencies, and higher dynamic energy consumption than that of the conventional DRAM. We then provide an overview of various methods, which can be employed to overcome these drawbacks. Hybrid memories involving both PCM and DRAM can be used, to achieve good tradeoffs in access latency and storage density. We conclude by presenting, the results of these methods that makes PCM a potential replacement for the current DRAM technology.

***Keywords*—**DRAM, Phase Change Memory.

I. Introduction

THE scaling in memory technology demands increased storage density and reduced size of the capacitor cells in DRAM. The conventional DRAM technology requires inherently very efficient charge placement and control on the capacitor. On scaling down, the charge placement and control over the capacitor cells becomes more difficult. Hence the charge sensing mechanism becomes unreliable [2]. Furthermore in volatile main memory, DRAM must not only place charge in the capacitors, but also mitigate the charge leakage through the access device. Studies have shown that, DRAM memory energy conservation should focus on leakage energy reduction, since leakage grows with main memory size and it could dissipate as much energy as its dynamic energy [1].

Venkataraman Krishnaswami is with Delft University of Technology, Netherlands, pursuing his Masters in Computer Engineering in the faculty of EEMCS (phone: +31-(0)6-81334710e-mail: V.Krishnaswami-1@student.tudelft.nl).

Venkatasubramanian Viswanathan is with Delft University of Technology, Netherlands pursuing his Masters in Computer Engineering in the faculty of EEMCS(phone:+31-(0)6-16146626; e-mail: V.Viswanathan@student.tudelft.nl).

Fortunately, several non-volatile memories have emerged as potential solutions, due to their exceptional low leakage power and high scalability. However it should be comparable to, or better than DRAM in terms of performance, efficiency and scalability to replace them completely. Considering the above criteria, flash memories and Phase Change Memories (referred as PCM from now on), amongst several others has stood as a potential replacement to DRAM technology. In this paper we provide an overview of PCM technology as a possible alternative to DRAM technology.

PCM enjoys a number of advantages, when compared to DRAM. The advantages being,

- High Scalability
- High storage density
- Zero Leakage power
- Low 'burst read' latencies
- Immune to cross talk

On the other hand, it also suffers from a few drawbacks when compared to DRAM. The drawbacks

- Limited lifetime
- High access latencies
- High energy consumption

Several approaches have been adopted by making architectural improvements, to improve the performance gap between PCM and DRAM.

The rest of the paper is organized as follows. Section 2 gives a brief overview of alternate memory technologies which are cited as a suitable replacement to DRAM. Section 3 discusses the pros and cons of PCM as a replacement to DRAM. Sections 4 describe the techniques and methods that are used to overcome the cons. Section 5 explores hybrid PCM and its advantages. Section 6 includes the inferences and Section 7 summarizes the study.

II. Alternate Memory Technologies

The DRAM technology, suffers from a number of drawbacks. Several other memory technologies like flash memories, phase change memories etc are being explored as a suitable replacement for DRAM. In this section, we provide an overview of these memory technologies.

A. Flash Memories

Flash memory is a non-volatile memory which can be electrically erased and reprogrammed. Since it is non-volatile it does not need power to retain its data. Flash memory is classified into two types, based on how the individual cells are connected. Each cell is connected in a way which resembles NAND gate or NOR gate. Hence the name NAND flash and NOR flash respectively.

Flash memories have some major limitations such as block erasure and memory wearing. Both NAND and NOR flash memories can be read or programmed either byte wise or word wise at a time using random access. However, they can only be erased and rewritten block-wise. The block sizes could vary anywhere from 16KB to 64KB, depending on the size of the flash memory. For instance, erasing a block will set all the bits in a block to 1. However, if one of the words or bits has been set to 0, the data can be rewritten only if the entire block is eased and reset again. Otherwise, data can be written only if the data to be written is a superset of data that is already written. This property of flash memory clearly shows the amount of overhead involved in rewriting data.

Flash memories can be erased and rewritten only for a limited number of times. The maximum number of times for which a cell can be erased is limited to 100000 times, after which cells start to wear out and the disk begins to lose its integrity. This is called as the wear-leveling. This could however, partially be offset by dynamically counting the erase cycles in each cell and remapping it to the other blocks which have been less written. This clearly involves write overhead.

Flash memory even though being non-volatile and four times denser than DRAM is clearly not an ideal replacement for DRAM owing to their drawbacks. This clearly puts the flash memory in lower levels of storage hierarchy on par with hard disks (referred as HDDs from now on) and USB device. Since flash memories have better latency power utilization than HDD, it is widely used as a disk cache [3].

B. Phase Change Memories

PCM uses a special type of material known as phase change material to store information. The phase change material can exist in two different states namely amorphous and crystalline. Since there are two states, which differ in their resistance values by several orders of magnitude, reading a value is just a simple sensing of resistance over cells. It has been found that a Phase change material with 60nm2 cross sectional area has a resistance of ~95 kΩ for SET state, and ~ 500 kΩ for RESET state [4].

Phase change materials are alloys such as Ge2Sb2Te5 (referred as GST from now on). There are also other alloys that could be used [10]. However, GST is the most widely used alloy for phase change materials due to its superior physical properties, when compared to that of the other alloys [6]. The material changes its physical state when heated above a particular temperature [4]. The crystallization temperature of the GST alloy is 300*C and its melting temperature is 600*C. The alloy, when heated to a temperature above its crystallization point (but below its melting point) changes into a crystalline state, which is referred as logic '1' or SET state. When heated above the melting point and cooled quickly, the alloy changes its state into amorphous state which is referred as logic '0' or RESET state.

Although the operating temperature of PCM is quite high, the usage of access device to heat individual cells makes the thermal crosstalk between cells to be very negligible [5]. PCM arrays can be fabricated quite similar to the DRAM arrays, only for the fact that we can use a phase change material instead of a capacitor [4].

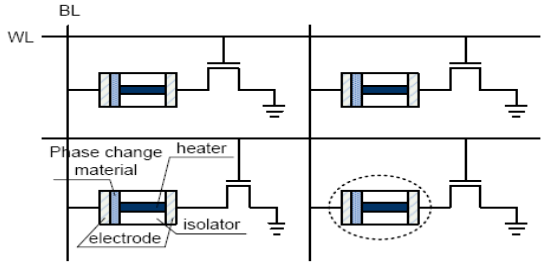

Fig. 1 PCM Cell Array [4].

The above figure shows the structure of a typical 2*2 PCM cell array. The PCM core is shown inside a circle. The phase change material is compressed by electrodes and a heating coil which is placed inside an insulator. Reading and writing operations have different access times and operating temperatures. Writing a '0' or RESET operation, bring the phase change material to amorphous state. It takes the highest operating temperature, since the operation involves heating the material above its melting point. However it relatively (with respect to DRAM) takes lesser time. The SET operation brings the material to its crystalline state. Writing a '1' or SET operation takes more time but lesser operating temperature, since the material takes a longer time to reach its crystalline state. However read operation takes the least amount of time and operating temperature.

As mentioned before, the PCM can endure only 108 to 109 writes. It has been shown that with such write endurance, a PCM used as a main memory can last only for ~100 days running a typical SPEC CPU program [4]. From the measurements that has been obtained, writing a single bit in a DRAM cell consumes 86.1fJ , whereas writing a '0' in a PCM needs 26808fJ and writing a '1' requires 13733fJ [4]. Although dynamic energy leakage is almost zero in PCM, it is quite important to reduce the dynamic energy consumption when compared to that of DRAM. Another important concern in PCM is that the 'read' and 'write' operations of a single cell are quite slower than that of a DRAM cell. Hence PCM as such is not an ideal replacement for DRAM technology. However, after implementing certain architectural changes and improvements the performance of PCM becomes quite comparable to that of DRAM.

III. Potential of Phase Change Memories

The phase change memories can be seen as a potential alternative to DRAM technology. However, it has its own pros and cons. In the section we provide an overview of the advantages and disadvantages of phase change memories, in comparison to DRAM.

A. Advantages of PCM as Main Memory

1. Scalability: PCM works on the principle of change in resistance of a material, under different states [10]. The absence of a capacitor in PCM makes it readily scalable.

In addition to this, there are no problems relating to charge sensing or storage mechanisms as in DRAM. Furthermore there is no need of periodic refresh in case of PCM, unlike DRAM.

2. Density: PCM offers much higher density than that of DRAM [3]. The density of PCM is almost four times to that of DRAM. Hence more amount of information can be stored in a PCM, than that of a DRAM for a given size. For example the GST material used in PCM, can exhibit four different transitional states. Consequently four different values can be stored in a single memory cell.

3. Leakage power: PCM belong to the class of non-volatile memory. Hence there is almost no leakage power [4].

4. Burst reads: The access latencies in case of 'burst reads' (reads that hit the same bank of memory) is much lesser than that of DRAM [4]. This is due to the fact that the peripheral logic of PCM is faster than peripheral logic of DRAM. In addition to this, DRAM has destructive reads, which is not the case in PCM. Hence the difference in access latencies during burst read operations (at same bank).

5. Crosstalk: The PCM is almost free from cross coupling effects. This is in sharp contrast with DRAM, which suffers from cross coupling effects between capacitors. This is more prevalent below 65nm technology.

B. Drawbacks of PCM as Main Memory

1. Limited Lifetime: PCM just like other non-volatile memories like flash memory suffers from limited write endurance. This severely limits the lifetime of PCM. The number of writes to a PCM is limited (about 10^9), after which the memory cell begin to wear out. The wearing is due to the fact that the operation is temperature dependant. During write condition, the material undergoes a change in state leading to expansions and contractions within the storage elements. This may also lead to segregation of the constituents of the phase change material during long runs. In addition to this PCM may also be prone to wear out during malicious attacks (unwanted writes) [8].

2. High access latencies: PCM also suffers from high access latencies. This is due to the fact that, the material takes a significant amount of time to undergo a transition between states, during a write operation .i.e. during the SET operation. The latency is of the order of tens of nanoseconds.

3. High energy consumption: Though PCM enjoys the advantage of having almost zero leakage power, it suffers from higher dynamic power consumption. This mainly supported by the fact that the read and write operations are temperature dependant. Energy consumed, also depends on the operating value of the injection currents.

PCM can be seen as a potential replacement for DRAM in the future, owing to the advantages of scalability. However the major drawbacks of PCM have to be addressed. Section 4 and section 5 deals with techniques which can be employed to overcome the drawbacks of PCM.

IV. Overcoming Drawbacks in PCM

In order to make PCM much more useful, and comparable to that of DRAM in terms of performance, it becomes essential to mitigate these drawbacks. Several methods have been explored in this regard. In this section, we provide an overview of the different methods adopted.

A. Improving Lifetime

The following techniques help to improve the write endurance of a PCM cell and thereby increase the lifetime. The lifetime of a PCM can be viewed as the time from the start (of usage) until the first cell of the PCM starts to wear out.

1. Eliminating Redundant Writes

As a first step in improving the lifetime of PCM, Ping et. al. [4] suggests reducing the write frequency to a single PCM cell. In a typical DRAM write operation, the write update writes the entire row. It has been observed that most of these writes are redundant [4]. A write does not change all the bit values. This means that the redundant writes could possibly be eliminated. Fig. 2 shows the results of Ping et. al. experiments on number of redundant writes, after testing with various memory benchmark programs.

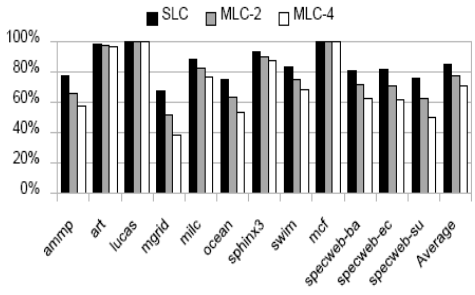

Fig. 2 Percentage of Redundant Bit Writes for single level cell (SLC) and multiple level cell (MLC) PCMs [4].

The above figure shows that all the benchmark programs exhibit high level of memory write redundancy. Here MLC-2 and MLC-4 stands for Multi Level Cell, where a single PCM

cell can hold two values and four values respectively.

Removing the redundant writes can be done by implementing a read before a write. The read operations are much faster than write operations in PCM. Therefore implementing a read before a write takes lesser time than what a complete write operation takes.

Therefore it is very beneficial to do a read before a write. This is illustrated in Fig. 3. This can be implemented by a simple XNOR gate on the 'write' path of the cell. This XNOR blocks the current value of the cell to be written if the value is the same as before. This has been illustrated from experiments performed by Ping et. al, the results of which are shown in Fig. 4.

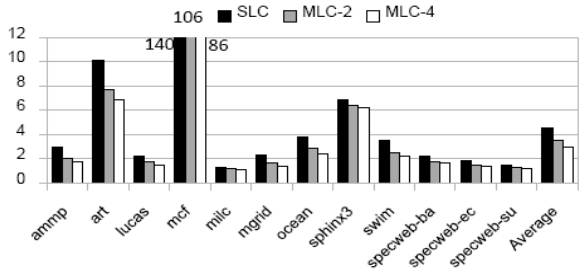

Fig. 3 Lifetime Improvement factor after redundant write Removal [4].

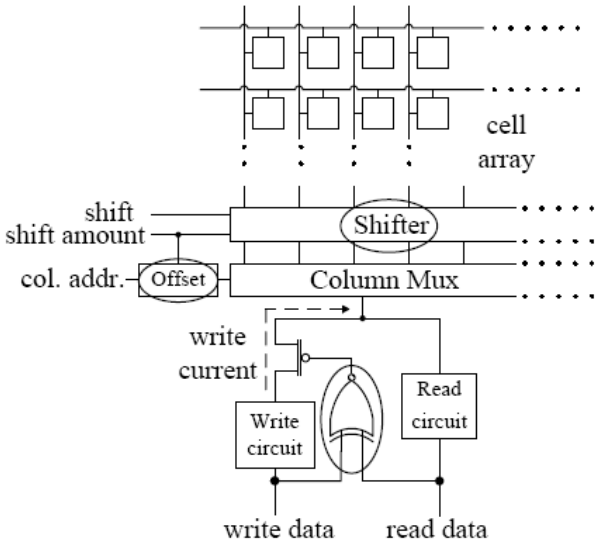

Fig. 4 XNOR gate on the 'write' path of the cell [4].

It appears at the first sight that the same technique could be used for DRAM as well. But DRAM doesn't benefit from this technique because, the read and write operations in DRAM take about the same time and energy. However in PCM write operations take about 5.x to 10.x times that of a read operation. Therefore this technique is very useful.

2. Row Shifting

The above method reduces the redundant writes up to 5 times. This results in a life time improvement to ~1.4 to 2.2 years [4]. Yet, this is very short for a main memory. This is because; most of the writes happen locally and therefore, certain hot cells get worn out soon. To avoid this, a technique called Row Shifting is used.

Row shifting mechanism aids in spreading out writes that tend to be localized to a few specific cells of a given row. After certain number of writes to a specific number of cells, new set of cells are chosen to write the data. This process helps to write evenly in all the cells in a row. Experimental results from the work done by Ping and Bo have proven that, shifting a byte at a time improves the performance very well [4]. However, the frequency of shifting also influences the lifetime. For instance, shifting quite often is not preferred because row once shifted, is difficult to be brought back into place and it involves a lot of overhead. This is due to the property of temporal locality. Therefore the frequency at which shifting is done, is to be carefully chosen. Furthermore, not all pages of the memory are written quite frequently. The pages can be sorted according to the number of times they are accessed. Therefore the best row shifting algorithm varies from page to page. Page classification is done based on the total write counts on the page and the standard deviation of writes among all lines in a page.

On varying row shift interval from 0 to 256 writes and averaging the resulting lifetime from all the sample pages, it was found that results for various benchmarks varied greatly. On plotting the results it was observed that the shift interval of 256 writes generates the highest lifetime for all the means for mcf benchmark [4] as shown in the Fig. 5. The write intervals were not extended for 2 reasons 1) Geometric and harmonic means have leveled off 2) increasing the write count will increase the hardware complexity.

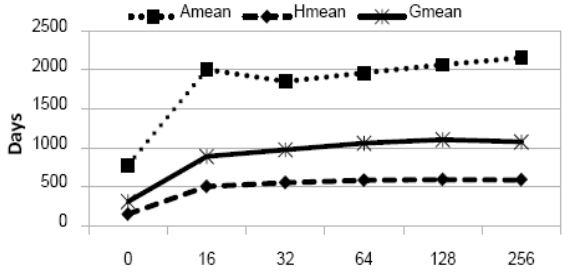

Fig. 5 Lifetimes over different row shift intervals arranged over different benchmarks [4].

3. Segment Swapping

The row shifting mechanism only improves the lifetime of each row. However, this technique has to be implemented on a granularity level big enough to be applicable for memory segments such as hot pages that are written quite often. This can be done with a technique called segment swapping. However the important parameters to be considered here are the size and swap interval of the segments.

The main problem lies in choosing the page size is the metadata that has to be sorted every time in order to determine the cold pages and hot pages. For instance, having a 4GB

memory with 4KB page size may require 1MB page counter size. Although this is not a big memory overhead, it requires long latency running times for running through the entire page counter. Therefore it is better to have bigger page sizes.

Different benchmark programs were run with different page sizes with varying swap intervals. The averages of all the results are plotted in a graph with harmonic mean, as in Fig. 6. It is inferred that a segment size of 1 MB with swap interval of 2X is the most efficient. This is because; the bigger page sizes incur more overhead for the extra writes. For example, the overhead for 1MB, 4MB, and 16MB segments on their base swap intervals are 2.8%, 5.6% and 5.2% respectively. This has been shown in the following figure The swap interval is in terms of base interval 'X' because; the swap size should be based on the page size. Larger pages should use larger page swap intervals.

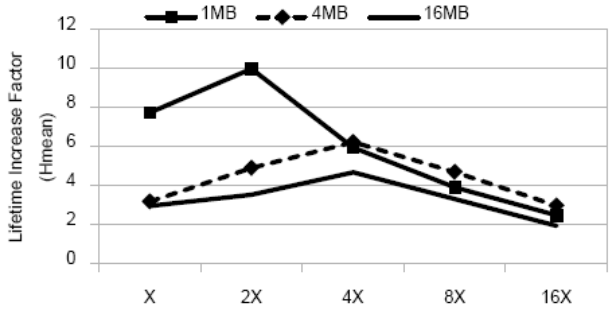

Fig. 6 Effect of segment swap interval on lifetime in HMean [4].

The life time improvement achieved, after implementing all the above described methods, (as obtained by Ping and Bo) are presented below in the form of a table, as shown below.

TABLE I
LIFETIME IMPROVEMENT ON SLC PCM AFTER IMPLEMENTING ALL THE WEAR LEVELING METHODS (IN YEARS)

Benchmark	SLC (Segment Swapping only)	SLC (after all the methods)
Ammp	4.3	32.5
Art	1.3	24.8
Lucas	1.1	15
Specweb-banking	4.4	22.1
Specweb-ecommerce	8.0	42.9

4. Partial writes

In case of a main memory system involving PCM, partial write technique reduces the number of writes by tracking the dirty data in the L1 cache. An extra state is added to each cache line which keeps track of stores using fine grained dirty bits. The data is written back to PCM only when the data in the cache is modified or evicted from the cache. Consequently number of writes to the PCM, are mitigated. This incurs a small overhead of latches which are used for this implementation. The partial level writes can be done in two levels of granularity, i.e. cache line size and word size.

B. Improving latency and Power

PCM suffers from very high access latencies (5-10 times that of a DRAM) in its operation. This limits the performance of the system. During write operation, latency in PCM is mainly attributed to the time taken by the phase change material to undergo a transition in its state i.e. from crystalline to amorphous or vice versa. The 'GS' phase change material, offers the best in terms of achieving a lesser latency [6]. These latencies can be hidden or tolerated to a certain extent by bringing about changes at the architectural level.

Benjamin, Engin et. al., suggest reorganization of buffers. Buffers can be reorganized to reduce application execution time from 1.6x to 1.2x, considering that it takes 1.0x in case of DRAM [2]. The buffers are made narrower and arranged in multiple rows. This is done in such a way that the total area remains the same, because area directly translates into cost involved. Multiple rows exploit locality to coalesce writes and hence hides their write latency to a certain extent.

Benjamin, Enjin et. al. also suggest using narrow buffers. Narrow buffers also contribute in mitigating the energy. This is because number of sense amplifiers, required decreases linearly, with buffer width. During development of PCM, nitrogen doping helps in increasing the resistivity and lowering programming current [7]. Process scaling also helps in saving energy i.e. as the size of the memory cell scales down, the volume of the material stored inside the cell also decreases. Consequently, lesser amount of material has to undergo a phase transition during the write operation. Decrease in area, also contributes to increase in the resistance value (since resistance is proportional to ratio of length and area). As feature size scales down by a factor say 'k', the area decreases by a factor of $1/k^2$, there by leading to an increase in resistance value, by a factor 'k'. The injection current value decreases by a factor 1/k. This leads to decrease in energy consumption. On the other hand, increasing resistivity by decreasing contact area also has the risk of reducing the signal strength. However, the sense circuitries are capable enough of sensing these signals successfully [2].

The performance of PCM under different benchmark programs, after implementing the above mentioned architectural changes is shown in Fig. 7.

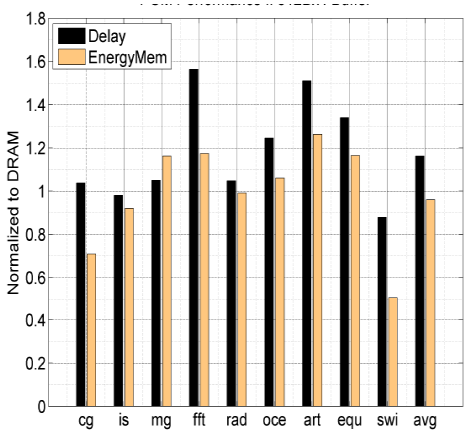

Fig. 7 Application delay and energy when using PCM with optimized buffering as a DRAM replacement [2].

V. Hybrid Memories

Moinuddin et. al. suggest the concept of hybrid memories. The hybrid PCM architecture is proposed to exploit the benefits of scalability and latency from PCM and DRAM respectively. This is achieved by using DRAM as a small buffer in front of the PCM main memory. Hybrid memories have latencies and performance very close to that of DRAM [3].

The traditional memory organization composed of DRAM as the main memory. Flash memories, were then used in conjunction with DRAM, to reduce the latency and power requirements of the disk. PCM can be used in place of DRAM, but it has a number of disadvantages to cope up with. Hence in the hybrid memory system, DRAM acts as a buffer as an interface between the processor system and PCM main memory. The different memory organizations are as shown in Fig. 8.

The hybrid PCM memory functions are organized similar to the DRAM main memory system. The OS manages the PCM main memory by means of a page table. The DRAM is organized just like a cache to the PCM main memory. The DRAM is not visible to the OS and is managed by a memory controller. The set up helps in reducing the read latency (as it is very less in case of a DRAM). In order to manage the write latency, write queues are placed between the DRAM and the PCM main memory. Several techniques are used by this system to overcome the drawbacks of PCM main memory [3].

A. Lazy Write Technique

The lazy write technique is used to improve the write endurance, in case of hybrid memories [3]. This is used to hide the slow write speed of PCM, without incurring any performance overhead. In case of a page fault, the page is fetched from the HDDs and stored in DRAM cache. But this data is not immediately written on to the PCM main memory, though a space in the memory is created for this data. The DRAM cache is provided with a tag, or an extra bit known as the 'presence bit'. The presence bit is made to be 0 when the data is fetched into the DRAM, from the HDDs. If the page is fetched from main memory it is set to 1. The data is written to PCM only when the dirty bit of an evicted data (from DRAM) becomes 1 or if the presence bit is 0. Hence the lazy write technique avoids writing the page fetched from HDDs directly to the PCM on a page fault. This improves write endurance, which ensures longer lifetime of PCM. This technique is almost similar to that of partial writes.

B. Page level bypass

Page level bypass, is another technique that is used to prevent excessive write operations over PCM [3]. In case of streaming operations, (where data is continuously changing, with respect to time) the write operations are avoided by making the data directly pass through DRAM and the OS invalidates entries in the PLB. The PLB (Page Level Bypass) is used to invalidate entries in the Page table. This prevents writing of data, in the PCM. The OS should be able to change this PLB, according to the application. A configuration bit may be used to achieve this. Similarly other techniques like line level write backs (writing in smaller chunks), wear leveling
Algorithms etc; can be used to improve write endurance.

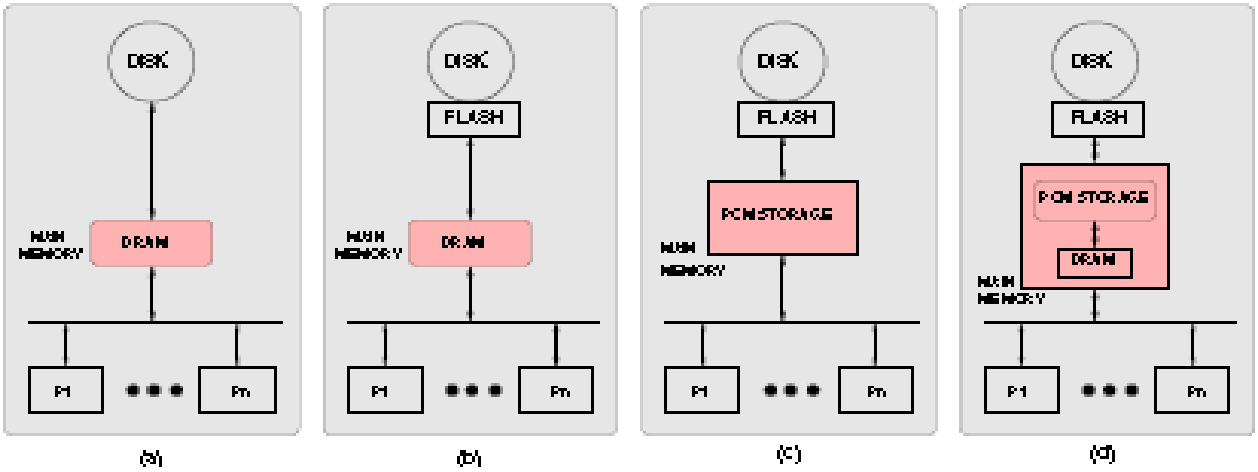

Fig. 8 a) Traditional memory system b) Flash based DRAM c) PCM d) Hybrid memory [3].

VI. Inferences

PCM is an attractive alternative to DRAM, provided their main drawbacks are addressed. The problem of latency as discussed earlier is mainly due to the phase transition time of the phase change material. The improvement of write latency can be brought about by developing new materials or optimizing the existing materials, like chalcogenide glass which can change their state quickly.

In order to optimize energy, power gating could be employed to the transistors in peripheral circuits. This is supported by the fact that data is not lost in non-volatile memories like PCM in absence of power. Process scaling also helps in improving energy. The memory state in PCM is detected based on the resistance values of the phase change material at different states. The injection current value depends on these resistance values. Since each state is associated with a resistance value, more number of states can be accommodated over a larger range of resistances exhibited by the phase change material. This enables us to store more number of bits per cell, thereby increasing the density. On the other hand such an attempt also bears the drawback of increase in the injection current and hence the energy. Hence a trade off should be considered between density and energy requirements.

The life time of the PCM is dependent on the number of write operations that we perform. Hence the PCM should be protected from unwanted malicious writes [8].For this purpose a pseudo random number generator along with a complex address translation mechanism can be used, to map the address to different locations for different writes.

The PCM has a very good storage density, and it can be used for storing back-up data, which will mostly require a read, once written. PCM should be avoided for streaming or multimedia applications, where a number of writes may be involved, with data changing continuously with respect to time. PCM can also make use of heuristic algorithms, to predict, what data could be fetched or read beforehand. The data can be retrieved at an earlier stage and stored in a buffer. This can help in hiding the read latencies in PCM.

Hybrid memories can be used to exploit the good features of both the DRAM and PCM. The latency and energy consumption for hybrid memories are almost same as that of the DRAM [3]. Table 2 provides a comparison between DRAM, PCM & hybrid memories.

The scalability of DRAM is achieved up to 42 nm, whereas for PCM it is 20nm [9]. The density of the PCM and Hybrid memories can be 2 to 4 times as that of DRAM.

TABLE II
Comparison between different memory organizations

no	Parameter	DRAM	PCM	HYBRID
1	Scalability	Less	High	Limited
2	Density	Less	High	High
3	Latency(read)	Less	High	Medium
4	Write speed	High	Low	Medium
5	Dynamic Power	Less	High	Medium
6	Static Power	High	Nil	Medium
7	Crosstalk effect	High	Nil	Less

The maximum read latency in case of DRAM may be around 60 ns, where as it is about 200ns-300ns in case of PCM [3]. The write speed, for DRAM can be as high as 1Gbps, whereas it is about 100MBps for PCM. The static power is almost nil in case of PCM as no periodic refresh is required, due to absence of capacitors. This also makes PCM, prone to almost negligible crosstalk, whereas this effect becomes predominant in case of DRAM, below 65nm technology.

VII. Summary

Non-volatile memories, like the PCM offers to be a very promising alternative technology to DRAM, owing to its high density and scalability advantages. However, it also suffers drawbacks of high access latencies, limited write endurance, and high dynamic power consumption. The access latency of a PCM can be improved by developing better materials. This is because; the access time is mainly dependent on the phase transition time of any given material. So the key lies in optimizing for better materials. The write endurance can be improved by techniques like partial writes (lazy write for hybrid architectures), redundant bit wires, wear leveling algorithms, page level bypass etc. Hence write endurance can be improved by modifying the architecture suitably. The dynamic power consumption can be reduced by decreasing the value of injection current (by increasing resistance). This can be done by suitably modifying the buffer architecture and through process scaling. Nitrogen doping could also be used to increase the resistance. Hybrid memories also offer an attractive alternative by exploiting the advantages of both the DRAM and PCM technology.

References

[1] S. Thoziyoor, J. H. Ahn, M. Monchiero, J. Brockman, N. P. Jouppi, "A Comprehensive Memory Modeling Tool and its Application to the Design and Analysis of Future Memory Hierarchies," The 35th International Symposium on Computer Architecture, pp.51-62, 2008.

[2] Benjamin C.Lee, Egin Ipek et.al. "Architecting Phase Change Memory as Scalable DRAM Alternative". ISCA'09, Austin, Texas, USA.

[3] Moinuddin K. Qureshi, Vijayalakshmi Srinivasan Jude A. Rivers "Scalable High Performance Memory Systems Using Phase Change Memory Technology" Proceedings of the 36th annual international symposium on Computer architecture 2009.

[4] Ping Zhou et.al, "A durable and energy efficient Main Memory using Phase Change Memory Technology" ISCA 2009

[5] A. Pirovano, A. Lacaita, A. Benvenuti, F. Pellizzer, S.Hudgens, R. Bez, "Scaling Analysis of Phase-Change Memory Technology," IEEE International Electron Devices Meeting, pp. 29.6.1-29.6.4, 2003.

[6] Y. Chen et al. "Ultra-thin phase-change bridge memory device using GeSb." In International Electron Devices Meeting, 2006

[7] H. Horii et al. "A novel cell technology using N-doped GeSbTe films for phase change RAM". In Symposium on VLSI Technology, 2003.

[8] Andre Seznec, "A Phase Change Memory as a Secure Main Memory," IEEE Computer Architecture Letters, vol. 9, no. 1, pp. 5-8, Jan.-June 2010, doi:10.1109/L-CA.2010.2

[9] Matthew J. Breitwisch "Phase Change Memory", IBM/Macronix PCRAM joint project aibm t j Watson Research Centre Yorktown heights, New York USA 2008.

[10] Matthias Wutting "Towards Universal Memories
Exploring the Potential of Phase Change Memories" International Symposium on VLSI Technology, Systems and Applications, 2007.