Future Memory Technologies

Benjamin Klenk
Department of Computer Engineering
University of Heidelberg
Germany, 68131 Mannheim
Email: b.klenk@stud.uni-heidelberg.de

***Abstract*—Memory is becoming increasingly important to achieving high performance. The past has shown that the development of memory performance lags behind that of processor performance and thus results in a gap known as the Memory Wall. In order to reduce this gap, memory hierarchies try to address this, but the gap is still present and keeps growing. The most common memory is Dynamic Random Access Memory (DRAM) which provides a large memory capacity and speed. Even with recent technology, there are limits and they will be reached soon. Therefore, we need improved technologies. A survey about such innovations and their promises will be shown in this paper.**

***Keywords*—DRAM, PRAM, HMC, Racetrack, STTRAM**

I. Introduction

In the past, high performance computing meant making processors faster and faster. We have already reached a limit and now have to achieve high performance with multiple processing elements instead of only one CPU. However, there is another component limiting performance. The memory provides all the data for the processing elements to work with. One task is to offer enough data (memory capacity) and to deliver that data as fast as possible. In this case, fast means to provide a high memory bandwidth and low latency to access data. Dynamic Random Access Memory (DRAM) has been the standard for main memory technology for decades. Besides having a good density, DRAM enables quick access and has become a mass-produced product. However, there are reasons why DRAM has already reached its limits and why we need improved and promising technologies.

A. Memory Wall

Even with the development of efficient multi-core processors, the gap between processor performance continues to grow. Memory hierarchies try to address the memory gap, but the problem is still present and becoming more prevalent. DRAM maintains this gap by limiting bandwidth and capacity simultaneously. The bandwidth is calculated by formula 1.

$$BW\left[\frac{byte}{s}\right] = f[Hz] \cdot \frac{w[bit]}{8[\frac{bit}{byte}]} \tag{1}$$

It is obvious that bandwidth could be increased either by frequency or by a wider data bus. Unfortunately, higher frequency also means higher power consumption and less signal integrity. Broadening the data bus is also not possible due the pin limitation of DRAM modules and multi memory-channel processors. It is easy to understand that DRAM works almost at its limit concerning memory bandwidth.

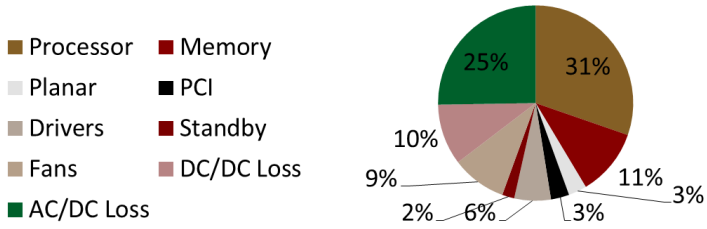

Fig. 1. Breakdown of power consumption [2]

B. Power Wall

Apart from the memory performance, the power also nearly reaches its limit. Currently, it is not possible to scale data centers as much as we would like due to power limitations. Regarding a computing system there are multiple power consumers, one of which is memory as shown in figure 1 [2].

If you keep in mind that the power supply will be much more efficient in the future to avoid losses, the memory increasingly becomes a main power consumer. While DRAM cells are power efficient, the interface is not. A 8GB DRAM module wastes about 17% of its power with refreshing and a 64GB module wastes about 50% [9].

The remainder of this paper is organized as follows: Section II will introduce DRAM technology as standard for main memory. Section III gives a short overview of Phase Change Memory. Following section IV deals with the Hybrid Memory Cube. The necessary physical background for further technologies is show in section V. Section V-C shows Racetrack memory and section V-D deals with Spin-Torque Transfer RAM. In the last section VI a comparison between all discussed technologies is shown.

II. DRAM memory technology

DRAM has become the state of the art main memory technology. DRAM has found success in spite of Static Random Access Memory (SRAM), a memory cell that consists of six transistors. The design of DRAM leads to a bigger density and therefore more capacity per die area. However, this architecture reaches a leakage current. Charge flows out of the capacitor, resulting in data loss. DRAM avoids this type of data loss by refreshing the memory cells frequently. This technique creates an unnecessary power consumption as well as slower access time.

DRAM is organized like an array. There are both row access lines (called Word Lines, WL) and column address lines (called Bit Lines, BL). This is shown in figure 2 [11].

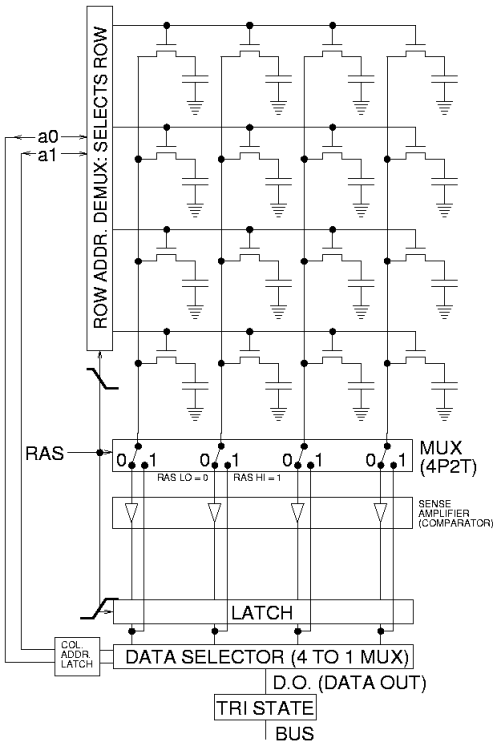

Fig. 2. DRAM 4x4 array organization [11]

While accessing a row, the capacitors charge the corresponding BLs. This current is amplified and a data selector chooses as much BLs as there are output pins and connects them afterwards. Then, the amplified current charges the capacitors again and the row is refreshed. If there is a write cycle the BLs are charged with the appropriate value and because the current loads the capacitors, the data is stored.

A special type of DRAM is used for main memory: Double Data Rate (DDR)-DRAM. It provides the double effective frequency and bandwidth. While DDR reached about 6.4 GBs, DDR3 currently reaches about 34 GBs. In near future, there will be the next generation DDR4 which will deliver about 60 GBs. Even so, this increase in speed does not come for free. If the frequency becomes higher, the signal integrity becomes more of a problem. Consequently, you can not use multiple DDR-DRAM modules due to reflections on the data bus. This inhibits the possibility of more capacity for main memory because high capacity DDR-DRAM modules are costly.

Based on these facts and the pursuit of faster and bigger memory, we need other architectures or technologies to satisfy our own requirements. Those will be presented in the following sections.

III. Phase Change Memory (PCM)

Phase Change Memory (often referred as PRAM or PCRAM) is based on chalcogenide glasses, which is also used for CD-ROMs. The basic idea behind this practice is to make use of two different states of the material: amorphous and crystalline. Depending on the state, the electrical resistance becomes low or high as shown in figure 3 [10]. If the material is crystal, the resistance changes linearly, but in the amorphous

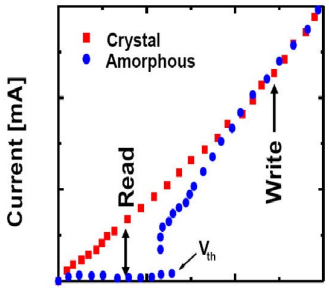

Fig. 3. Resistance of PCM depends on the phase of the material [10]

phase the current will not increase unless the voltage reaches a threshold. Through measuring the resistance the two different states can be determined as logical '1' or '0'.

Only two states are not enough to build a memory cell. Transitions between these phases are reached through electrical pulses heating the material. To get into the amorphous phase the temperature must reach the melting point to separate the crystal lattice but cool down again quickly. The atoms do not have enough time to reorder and locked in their position. To reach the crystalline state the temperature needs to reach a certain point to get the atoms moving. If the temperature is held for a amount certain time the atoms will build a crystal lattice. Unfortunately, the RESET signal (to get in the amorphous phase) needs a higher current, on the other hand the SET signal needs more time.

PCM still provides a lower access time than DRAM, but needs more power today. Power scales with the size of the chalcogenide element. As a result of decreasing feature sizes, the power consumption also decreases. The main advantage about PCM is that the data can be kept even without power with a higher density than DRAM [6]. Therefore, PCM is more of a competitor with FLASH for mass storage products than with DRAM for main memory. However, there are some approaches to build even main memory with PCM. The most promising is a hybrid solution with DRAM and PCM, where DRAM acts as a buffer to hide the worse latency of PCM. Behind that, PCM is used for a high capacity memory. Figure 4 [6] shows this approach.

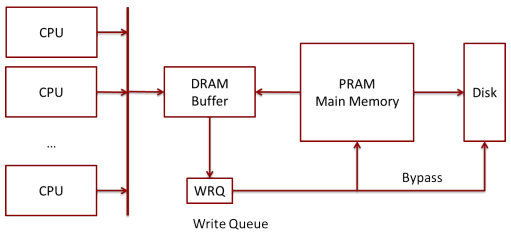

Fig. 4. Hybrid approach to make use of high capacity PCM and faster DRAM as a buffer to hide the slower access time of PCM. [6]

As always, the benefit of this approach depends on the application running on the system. Nonetheless, simulations show the performance of a 32GB PCM / 1GB DRAM hybrid system reaches nearly the same performance as a 32GB DRAM-only solution [6]. Based on this, PCM provides more capacity and in the future the hybrid solution could outperform a pure DRAM main memory.

IV. Hybrid Memory Cube (HMC)

In comparison to all other shown technologies in this paper, HMC is more a new architecture than a new technology. It still relies on DRAM as memory technology, but improves the access to it. While recent architectures consist of a 2D displacement of DRAM, the HMC is made up of a 3D displacement in the shape of a cube (figure 5 [3]). One important peculiarity is the processor memory interface of the HMC. Presently, the memory controller has to deal with all constraints of the DRAM module, which means there are a lot of signals which must be driven. Moreover, these signals are time critical. HMC pursues a more abstract protocol based on messages. This enables a higher data rate, but also an easier connection to the processor.

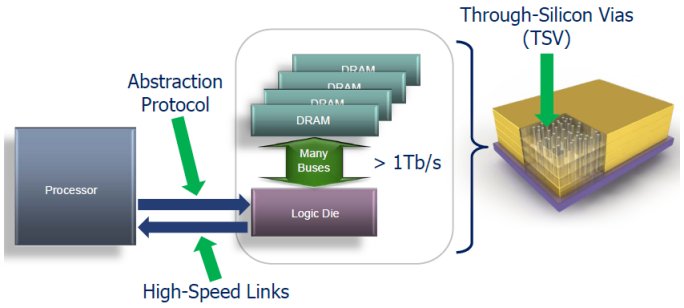

Fig. 5. HMC consists of multiple DRAM layers connected to the logic die via TSV technology. [3]

The main element of the HMC is the logic die. It combines the interface between the processor and the memory and furthermore between the control logic and the DRAM cells. The logic die and the DRAM slices above are connected by through-silicon vias (TSV), which provide a vertical connection through a silicon die. TSV also enables a bandwidth above 1 Tbps between the logic and the memory arrays [3]. Due to this, there is less logic necessary on each DRAM slice because the slices consist of more memory cells. Table I shows a comparison between DRAM technology and HMC. A HMCgen1 with 512MB total DRAM capacity and 128 GBs bandwidth was used for benchmarking [3].

TABLE I. Comparison between DRAM technology and HMC [3]

Technology	VDD	BW [GBs]	Power [W]	mW/GBs	pJ/bit
DDR2-667 2GB	1.8	5.34	5.18	971.51	121.44
DDR3-1333 2GB	1.5	10.66	5.52	517.63	64.70
DDR4-2667 4GB	1.2	21.34	6.60	309.34	38.67
HMCgen1	1.2	128.00	11.08	86.53	10.82

HMC becomes a promising approach, especially in the current multi-core era. It delivers much more bandwidth than current architectures and enables more concurrency. In a system where plenty of threads work together, each thread has its own working set. While accessing the main memory each thread should access a different bank to get most concurrency. A common system has about eight DRAM devices with eight banks per device, which leads to a total of 64 banks which can be accessed in parallel. The HMCgen1 consists of four DRAM devices with 16 slices each and two banks per slice. This results in 128 banks, twice as many as the pure DRAM system shown before [3].

V. Spin-based memory technologies

Shown technologies are based on electrical charge or changing phases of a material. Recent research focuses on the ability to store data magnetically inside a small cell. To understand these technologies some physical background is necessary. After that two promising technologies will be shown.

An electron has properties like mass or charge. However, there is another important property of particles called spin. The spin can be either ”up” or ”down”, therefore only two states are possible. Inside a non-ferromagnetic material, the population of spin-up and down electrons is equal. Alternatively, ferromagnetic materials consist of an unequal population. If a current flows through a ferromagnetic material it becomes polarized and is called spin-polarized current. Putting two ferromagnetic materials together and then separating them with a insulator barrier results in a Magnetic Tunnel Junction (MTJ) element. A current flowing through a MTJ element becomes spin-polarized by the first magnetic element and has an impact on the magnetic momentum of the second one. The build-up is shown in figure 6.

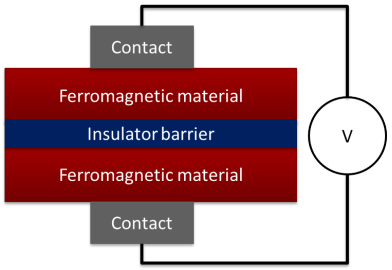

Fig. 6. MTJ element consisting of two magnetic layers separated by a insulator barrier.

There are two physical phenomena which are important for the use of a MTJ element as a memory technology.

A. Tunnelling Magneto-resistance (TMR)

One important feature of a memory technology is to have two possible states which can be clearly distinguished from each other. Regarding a MTJ element, the two states are defined by high and low electrical resistance between both contacts. This effect is called Tunnelling Magneto-resistance and is directly proportional to the direction of the magnetic moments of the ferromagnetic material inside the MTJ element. The direction of these magnetic moments can either be parallel or contrary like shown in figure 7.

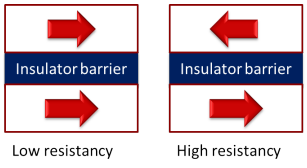

Fig. 7. The direction of the magnetic moments inside the MTJ elements determine the electrical resistance.

B. Spin-Torque Transfer (STT)

Besides the ability to distinguish two different states, it is also necessary to influence them. As mentioned above, a polarized current can impact the magnetic momentum of a ferromagnetic material. This is only possible if the material is thin enough (100-200nm) [5]. Therefore, MTJ elements are built with a thin ferromagnetic layer, called free layer (FL) and then a thicker one which is called pinned layer (PL). The direction in which the magnetic momentum of the FL will move depends on the direction of the polarized current which is injected. This way, the state is changed and the data written to a MTJ memory cell.

C. Racetrack Memory

The main idea behind Racetrack Memory is to store data on a ferromagnetic nano-wire. The wire is divided into magnetic domains, separated by domain walls (DW). Domains are either magnetized "up" or "down" and a DW separates different magnetized domains. To read or write data the physical phenomena of section V are used. The approach is shown in figure 8 [12].

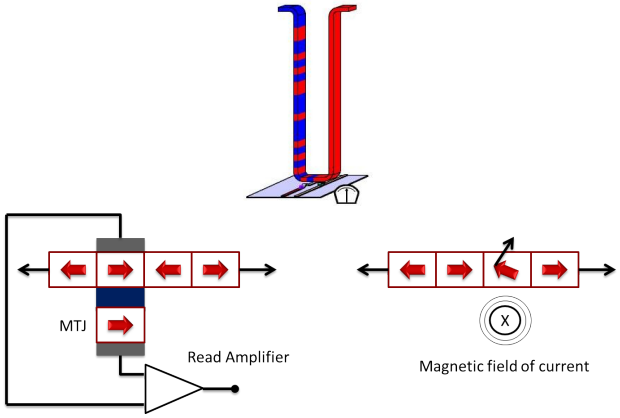

Fig. 8. A ferromagnetic nano-wire is the main element of the Racetrack Memory approach. Below is shown the read and write abilities. [12]

The memory acts like a shift register where data can be shifted in both directions. Polarized current pulses are injected to shift the data. The current becomes spin polarized and aligned with the magnetization of the material. At DW the orientation of the spin-polarized electrons flip and this results in a flip of neighbour atoms since the angular momentum must be conserved. When electrons flip from 0 to 1, the atoms flip from 1 to 0 and thus the domains are shifted along the racetrack [12].

Early studies show that Racetrack reaches about 20-30 ns of read/write time. However, the access time depends on the density and how many bits are stored on one nano-wire. DWs can be pushed along the racetrack at about 100-110 $\frac{m}{s}$ and therefore the more bits per nano-wire the slower the access time. Also here a tradeoff has to be find in future [12].

Racetrack memory needs a great deal more research but is already a promising technology for the future. It remains to be seen if this will become a serious option to be used for main memory. This question can not be answered at this point.

D. Spin-Torque Transfer RAM (STTRAM)

Another memory technology which is based on MTJ elements is STTRAM. STTRAM is a further development of Magneto-resistive Random Access Memory (MRAM). The memory cell is built up as shown in figure 9 [13]. The main element of the memory cell is a MTJ element. Spin-polarized current is used to manipulate the memory state and the state can then be determined by measuring the TMR.

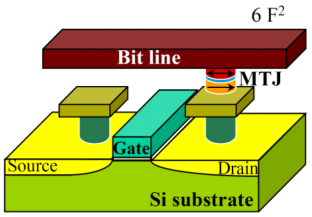

Fig. 9. Memory cell of STTRAM. [13]

Even if STTRAM is still in the research phase, it shows promising characteristics. A main advantage is the ability to write current scales with cell size which leads to less power consumption. Also, STTRAM provides the ability to store more than one bit in each cell, therefore a high density can be reached. Recently a TMR of 100% has been reached, but researchers claim the TMR can be significantly higher if a dual MTJ (DMTJ) is used. A DMTJ element consists of two PLs and a FL in the middle of them.

Although, there are many advantages, there are also challenges which must be solved in future. One of them is to achieve a low write current density. Currently, a write current density of 1-2 MA/cm^2 with a standard MTJ and about 1 MA/cm^2 with a DMTJ element is reached [7]. The smaller the device area the smaller the write current can be. Another problem is the thermal stability. Contrary to the write current density the thermal stability decreases as device area shrinks. Therefore, a tradeoff between write current density, thermal stability and device area must be found in near future.

VI. Comparison

The previous sections have shown the various approaches of new technologies. Especially Racetrack and STTRAM are still research projects and therefore characteristics are still not well known so far. Important characteristics are as follows:

- Cell size: The smaller the higher the density can be.
- Access time: Determines how fast data can be accessed by a processor.
- Energy per bit: Energy which is needed to store one bit in a memory cell.

Table II [6][7][12][14] shows a comparison between all shown technologies concerning features mentioned above. The access time is divided in write access time (first value) and read access time. Consider that the access time means the access to a memory cell which is not equal to latency. Also, energy per bit is not the same as power consumption. For example, memory cells of DRAM are very power efficient. However, the interface consumes a lot of power.

TABLE II. COMPARISON BETWEEN ALL SHOWN TECHNOLOGIES [6][7][12][14]

Techn.	Cell size	State	Access time	Energy/bit	Retent.
DRAM	$6\ F^2$	Product	10/10 ns	2 pj/bit	64ms
PCM	$5\ F^2$	Prototype	100/20 ns	100 pj/bit	years
Racetrack	$\frac{20F^2}{Bits} \simeq 5F^2$	Research	20-30 ns	2 pj/bit	years
STTRAM	$4\ F^2$	Prototype	2-10 ns	0.02 pj/bit	years

HMC is not listed in the table II because it is more of an architecture than a new technology and is based on DRAM yet. Possibly, in the future HMC will also consist of new technologies as shown before.

VII. CONCLUSION

First of all, this paper shows the problems of current memory technology. Problems like the Memory or Power Wall must be solved with new architectures or technologies. These were shown earlier in this paper. The main question now is what we can expect in the future. Racetrack and STTRAM are still in their infancy and a lot of research will be necessary to determine if these could ever replace DRAM as standard for main memory. HMC is a sophisticated technology, which shows much promise even though this is more of an architectural change than a new technology. PCM or PRAM only makes sense in combination with DRAM as a hybrid solution. PCM has problems performing well due to the low access time. Regarding Table II, STTRAM shows most promise but it remains to be seen if it could fulfil the expectations many people have nowadays. In the near future, HMC will be an option for use as main memory. In the distant future, STTRAM and possibly Racetrack memory are other promising technologies, but still a lot of research is necessary. DRAM will probably remain as memory technology for this decade. Every technology has its own challenges and there is no final solution to fulfil all requirements simultaneously.

ACKNOWLEDGEMENT

The author would like to thank Prof. Dr. Holger Froening at University of Heidelberg for his role as the supervisor of this work. He has always been on hand with help and advice for the author.

REFERENCES

[1] B. Jacobs, *The Memory System*, 1st ed. Morgan & Claypool Publishers, 2009

[2] L. Minas, *The Problem of Power Consumption in Servers* Intel Inc, 2012

[3] T.J. Pawlowski, *Hybrid Memory Cube (HMC)* Micron Inc, 2011

[4] J. Jeddeloh and K. Brent, *Hybrid Memory Cube: New DRAM Architecture Increases Density and Performance* IEEE Symposium on VLSI Technology Diges of Technical Papers, 2012

[5] L. Gao, *Spin Polarized Current Phenomena in Magnetic Tunnel Junctions* Dissertation at Stanford University, 2009

[6] M. QURESHI and S. Gurumurthi and B. Rajendran, *Phase Change Memory*, 1st ed. Morgan & Claypool Publishers, 2012

[7] M. T. Krounbi, *Status and Challenges for Non-Volatile Spin-Transfer Torque RAM (STT-RAM)* International Symposium on Advanced Gate Stack Technology, Albany (NY), 2003

[8] P. Kogge et al, *ExaScale Computing Study*, Public Report, 2008

[9] J. Lia et al, *RAIDR: Retention-Aware Intelligent DRAM Refresh* ISCA '12 Proceedings of the 39th Annual International Symposium on Computer Architecture Pages 1-12

[10] Agiga Tech, *Website* http://agigatech.com/blog/pcm-phase-change-\ \memory-basics-and-technology-advances/ from 01/24/2012

[11] VLSI and Embedded System Technical Library Blog *Website* http://vetechlib.blogspot.de/2008/09/random-access-memory.html from 01/24/2012

[12] S. Parkin et al., *Magnetic Domain-Wall Racetrack Memory* Scientific Magazine, January 14, 2011

[13] A. Driskill-Smith, *Latest Advances and Future Prospects of STT-RAM* Grandis Inc Non-Volatile Memories Workshop at University of California, 2010

[14] M.H. Kryder and K. Chang Soo, *After HDD Drives - What Comes Next?* IEEE Transactions On Magnetics Vol 45 No 10, 2009