I. INTRODUCTION

MAGNETORESISTIVE Random Access Memory, or MRAM, is showing promising potential for commercial competition among universal memory [1]. Existing semiconductor memories all have shortcomings and limitations, and cannot fulfill all the important attributes needed for memory with one solution. MRAM, however, possesses all of these key features. It is the only non-volatile, non charge-based, non destructive memory currently in production. The most important of its features are its non-volatility and unlimited endurance for reading and writing. In addition, the potential for very high densities while maintaining high speed, low power operation make MRAM attractive memory architecture for continual research.

The following discussion outlines the architecture and operation of MRAM. We track the progress within the industry and compare its state-of-the-art features with those of other competitive memory.

Challenges to be considered for MRAM production include improving manufacturability and scaling its cell density down. We also discuss scaling limitations later in this discussion. The paper concludes with a summary of MRAM performance and future potential, with a short discussion on whether or not it is worth investment.

II. ARCHITECTURE & OPERATION

MRAM data is stored by changing the magnetic polarity of the memory element. The storage element, called the Magnetic Tunneling Junction (MJT), is composed of two magnetic plates that are separated by a thin insulating layer. The thin insulating layer allows tunneling when a current is applied through the cell, allowing current to pass through the junction for reading purposes. The MJT operates similar to a capacitor such that data is stored between two metal plates. One plate is made of soft ferromagnetic material, and the other is permanent magnetic material. Essentially, we store data bits by controlling the plates’ magnetic polarities. A memory device consists of an array of these cells.

Word lines connecting each memory element are arranged perpendicularly above and below the cell array. To read the data in a cell, current is passed through the memory element and the bit is determined by measuring its resistance. This allows fast reading without changing the state of the bit. When the two magnetic plates have parallel polarity, the resistance through the MJT will be much lower than when they are aligned with anti-parallel polarity. A transistor is used to control the current flow to each cell, allowing the ability to read individual memory elements. Since there are no stored electrons, there is no dissipation of the data stored in the MJT. The magnetic polarization will retain its state for a prolonged length of time at high temperatures.

Writing to the cells is slightly more complicated. Current is passed through the word lines corresponding to the desired cell, which introduces an induced magnetic field at the tunneling junction. The soft or “free” magnet picks up this field and polarizes itself accordingly. The setting/erasing of the bit is done by the different applied currents on the word and bit lines. A major problem with writing to memory this way is that nearby cells often pick up some induced magnetization as well, which is quite undesirable. Later in the paper we will discuss how this affect can be countered.

III. INDUSTRY

In order to gain a better understanding of the future trends of MRAM we will now analyze the past and present MRAM industry. The origins of the MRAM industry began in 1989 when several important discoveries were made by IBM on the Giant Magnetoresistive (GMR) effect. It was soon realized that this effect could be used in a memory circuit and several other companies join the research and development of MRAM. Some of these companies were: IBM, Freescale, Infineon, NVE, and Cypress. The first chip containing MRAM was manufacture in 2003, it had 128Kb of memory and was built using 0.18um technology this chip was not made...
commercially available however. In 2004 Infineon developed a 16Mb prototype proving the technology was capable of storing large amounts of memory. This same year, Freescale was the first company to sell a chip with MRAM in its standard product line. In 2005 Honeywell joined the MRAM scene and began to sell chips with MRAM that passed high radiation and military standards, giving them an edge on the defense and aerospace industries. Also during this time a 2Ghz MRAM cell was created by Hans Schumacher of the German Physical-Technical Institute in Braunschweig (PTB) using a “ballistic addressing” technique. One of the most disappointing pieces of news from this year was that Cypress decided to leave the MRAM market stating, “We no longer believe that the 1T-1MTJ MRAM technology will be able to successfully attack the SRAM market, leaving MRAM as a niche technology with higher bit pricing than that of SRAM.”

2006 brought more optimism however, when Toshiba and NEC developed a 16Mb, 200Mb/s 1.8V MRAM memory cell in the lab. This same year Freescale began sales of their 4Mb MRAM chips priced at approximately $25 each, they announced that “We are now officially in the era of new memory technology!” MRAM development continued in 2007 when Toshiba found that Perpendicular magnetic anisotropy could be used to put MRAM memory densities in the gigabyte range and speed up to 250 MHz were achieved. 2008 has also proven to be a good year for MRAM so far; Korea has promised Samsung and Hynix $50 Billion over the next 3 years for development of new memory technologies, MRAM being the most promising on their list. And recently Freescale has won a bid to provide Siemens with MRAM memory for their industrial touch screen controller interface [3].

As one can see there are many companies pursuing the MRAM market and pushing the technology forward, but currently Freescale is leading the market in MRAM production and sales. Much of their success can be attributed to the fact that they created the first commercially available MRAM device which was received by the electronic industry very well. In fact, there 4Mb device has won several awards since its introduction, including: “The Electronic Products' 2006 product of the year, EE Times China’s 2007 Memory product of the year, LSI of the Year's Award of Excellence, the 2007 In-Stat/Microprocessor Report's Product of the Year Award in innovation, it was also selected as a finalist in EDN's 2006 Innovation Awards and EE Times' 2006 ACE Awards.”[5] With all these recognitions from so many different groups and associations it looks as if the world understands and appreciates the features of MRAM when it is produced properly. Another reason Freescale’s MRAM is looking very promising can be found in the Price vs. Performance graph below. As the graph shows only nvSRAM has faster cycle times, but MRAM is by far the cheapest technology per Megabyte.

Freescale has also been working on creating a robust line of MRAM products. They currently have devices that work in the extended temperature range of -40° C to +105° C and are attempting to extend this range by working together with automotive companies to meet automotive standards and continue to broaden their MRAM market.

IV. PERFORMANCE

As one can see from Table 1 above MRAM is the only type of memory that has no drawbacks in any of the characteristics listed. One of the main arguments for MRAM as the universal memory solution is the program endurance (the number of times it can be written). A reliability test was performed by Slaughter et al, in which a bit was exposed to a pulsed switching current with 20 ns duration. The bit reversed every time with no significant change in the critical device parameters to greater than 10^11 write cycles. The approximate lifetime for FLASH memory is a much more limited 10^4–10^6 write cycles. These results therefore demonstrate the robust program endurance of MRAM that enables its potential use as a universal memory [1].

Non-volatility is the other strong argument for MRAM as the universal memory. For non-volatile memory options, the memory to beat is FLASH. FLASH will continue scaling for only a few more years but MRAM will scale at least that long because it is a newer technology.

Desikan et al. performed a benchmark for MRAM, SDRAM and stacked SDRAM. They are trying to evaluate the effect of
off-chip physical DRAM being replaced with on-chip MRAM. The following data is attributed to those measured and collected in [6]. Their setup is shown in the following table:

<table>
<thead>
<tr>
<th>Processor Core</th>
<th>Alpha 21264 pipeline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fetch Width</td>
<td>8</td>
</tr>
<tr>
<td>Issue Width</td>
<td>8 - Integer, 4 Floating Point</td>
</tr>
<tr>
<td>Commit Width</td>
<td>22</td>
</tr>
<tr>
<td>Processor Frequency</td>
<td>3.8 GHz (Clock period of 10 FO4 inverters per stage)</td>
</tr>
<tr>
<td>Size of Structures</td>
<td>Fetch Queue - 8</td>
</tr>
<tr>
<td></td>
<td>Issue Queue - 40 Integer, 30 Floating Point</td>
</tr>
<tr>
<td></td>
<td>Reorder Buffer - 160</td>
</tr>
<tr>
<td></td>
<td>Load/Store queue - 64 Load, 64 Store</td>
</tr>
<tr>
<td></td>
<td>Physical Registers - 62 Integer, 82 Floating Point</td>
</tr>
<tr>
<td>Functional Units</td>
<td>Integer - 8 ALUs, 4 multipliers</td>
</tr>
<tr>
<td></td>
<td>Floating Point - 2 ALUs, 2 multipliers</td>
</tr>
<tr>
<td>Branch Predictor</td>
<td>Alpha 21264 Tournament Predictor</td>
</tr>
<tr>
<td>L1 D-Cache</td>
<td>16 KB, 2-way set associative, 64 byte line size, 2 cycles lat latency</td>
</tr>
<tr>
<td>L1 I-cache</td>
<td>16 KB, 2-way set-associative, 64 byte line size, 1 cycle lat latency</td>
</tr>
</tbody>
</table>

The factors influencing the MRAM design space were number of banks, cache line size, and page placement policy. A large number of banks in the system increases the concurrency in the system, and ensures fast hits to the closest banks. However, the network traversal cost to reach the farthest bank also increases due to the increased number of hops. Large line sizes can result in increased spatial locality but they also result in an increase in the access time of the cache. Thus, there is a trade-off between increased locality and increased hit latency which determines the optimal line size when bandwidth is not a constraining factor. In addition, the line size has an effect on the number of bytes written back into an MRAM array, which is important due to the substantial amount of power required to perform an MRAM write compared to a read.

They use an area and timing model to get the access latency for different sizes of sub-banks. After choosing the optimal sub-bank size, they computed the latency for banks with different number of sub-banks. This latency was used to compute the performance in Instructions per Cycle (IPC).

They consider four different sizes of MRAM banks with 25, 49, 100, and 196 banks respectively. The benchmarks use six SPEC2000 floating-point benchmarks, five SPEC2000 integer programs, 4 scientific applications – 3 from the NAS suite and smg2000, and a speech recognition program, sphinx. Their results show that MRAM systems perform 15% better than conventional SDRAM systems and 30% better than stacked SDRAM systems.

These results and all others discussed thus far in the paper have been positive for the case to invest in MRAM. Like all other things in life however MRAM does have its drawback, these will be discussed in the following section.

V. LIMITATIONS

MRAM has several limitations that need to be considered when comparing it to other substitutes. A few of the major issues that affect the viability of MRAM is cell density, temperature compatibility during manufacturing with existing processes, reliability and economical feasibility.

Density

Most basic MRAM cell sizes are limited to 180 nm or more due to half-select problem. For toggle mode MRAM, this becomes a problem near 90 nm. For feasible production, MRAM cell size needs to be in the region of 65 nm, which will require the spin-torque-transfer mode. This mode requires switching more current through the control transistor than conventional MRAM, requiring larger MRAM.

Temperature Compatibility

One of the challenges involved in MRAM application is the temperature compatibility with CMOS processes, requiring the magnetic materials to withstand standard process temperature or development of low temperature processes for MRAM technology.

Figure 5 shows the response of PSV-GMR bits to bake temperature. Each data point represents the average of about 100 bits across two wafers, exposed for 5 minutes each. Good stability is observed at 300°C, followed by a sharp drop in signal of more than 30% by 400°C, likely a result of increasing bit resistance caused by intermixing of the magnetic layers with the Cu spacer layer.
Figure 5: Temperature response of PSV-GMR bits: (a) Change in Magneto-Resistance vs. Temperature, (b) Change in Resistance vs. Temperature

Figure 6: Thermal Response of MTJ Bits: (a) Resistance vs. Temperature, (b) Magneto-Resistance vs. Temperature

Studies have also shown that the mean resistance initially increases on annealing, but begins to decrease between 240°C and 300°C. Figure 26 shows the response to anneal temperature. Figure 6a shows junction resistance decreasing steadily until 300°C and then beginning to increase. Figure 6b shows a slow increase of mean resistance behavior, which peaks at 275°C, followed by a sharp drop after 300°C, reaching ½ of the original value by 330°C and nearing zero at 390°C.

Figure 7 represents the change in the Magneto-Resistance transfer curve due to annealing, displaying the dramatic difference between as-deposited and low-temperature (T<300°C) annealed material. In this case, Magneto-Resistance doubles and resistance-area product decreases slightly after a 250°C anneal.

With temperatures affecting the characteristics of the magnetic material to this extent it is clear that fabrication is one of the most challenging setbacks MRAM needs to overcome before being the universal memory element.

**Manufacturing**

Manufacturing the required cell structure for MRAM on an industrial scale will require very tight control over the chemical composition and thickness of the various layers of the stack structure, notably the ultra-thin tunneling barrier. The cell layers also need to fit economically into a standard CMOS flow with minimum additional masks, and MRAM structures must not be damaged by later process steps. Improving the MRAM cell yield is an important and necessary goal, as the profit is directly proportional to the device yield once a fabrication has been decided upon. More work is required to determine fundamental reliability issues. No wear-out mechanism is known yet, allowing high write cycle endurance. Long test times and lack of accelerating processes make it difficult to establish the standards experimentally. Stray magnetic fields and magnetic disturbances caused by adjacent cells might also affect reliability, and are better investigated in large production samples.

**Economic**

Success of MRAM will ultimately depend on the economical viability of the product. Process equipment manufacturers need to be convinced to make the necessary investments to develop new fabrication equipment needed for the required processing. It also needs to be able to find a competitive application in which it can survive and grow competitively among other technologies currently in place, such as flash memory.

**VI. CONCLUSION**

The functionality, industry achievements, performance parameters, and limitations of MRAM have been thoroughly analyzed, followed by a comparison of MRAM against several commercial memories. We conclude that a company’s investment should be largely dependent upon the MRAM application. For small memory requirements in extreme environments where devices will be inaccessible for long periods of time, such as satellites and industrial applications, MRAM is currently a great investment. If a company needed large amounts of memory but does not require a large number of writes, flash is still the best available technology. From a foundry’s point of view, investing in equipment to fabricate MRAM would make them leaders in the MRAM industry. Being the first major foundry with the ability to place MRAM as the primary memory on a processor would be very beneficial. As stated in [6], MRAM has been proven to outperform SDRAM, the leading memory in processors. Perfecting this MRAM integration could lead to a new wave of products that were never realizable before.
VII. REFERENCES


