

Memory Lane Change

Get ready for a walk down memory lane. Tom describes the major paradigm shifts that have occurred in memory technology over the past three decades and then readies you for what's ahead.

Nowhere is the march of silicon demonstrated so clearly than with memory chips, where it all began some three decades ago.

Everybody knows Intel created the 'x86 (8088) microprocessor that powered the original IBM PC. But do you remember that they got their start years earlier with the 1103 dynamic RAM (DRAM) chip? Simply put, no DRAM, no PC, or any of the other digital A/V gadgets (e.g., video game and HDTV) we take for granted.

I certainly don't design PCs, video games, or HDTVs, and I doubt many of you do either. Thus, to tell the truth, I haven't kept up with developments on the DRAM front. Time to, pardon the pun, refresh my memory.

Let's see, the latest and greatest DRAMs are 2-Gb chips that run from 500 to 800 MHz.^[1] That's roughly 2 million times the density and 200 times the speed of the 1-Kb, 300-ns access time 1103. You don't need a PC crunching those numbers to understand that a 400 million times improvement is a big deal. And that's not even considering issues like chip size, reliability, and power. (The 1103 needed 19- and 22-V supplies!) What the heck, let's just say modern DRAMs are a billion times better than the original. If you feel that's an exaggeration, patience, grasshopper. As with all things silicon, it's only a matter of when, not if.

The invention of the DRAM was a true example of a "paradigm shift" unlike most (ab)uses of that over-worked phase. Almost overnight, the

big computer companies of the era Dr. Jacked core memory in favor of DRAM (see Photo 1). It wasn't more than a couple of years after the 1103's introduction that it was the single biggest selling chip in the world.

Lest you think Intel's 1103 was a fluke, keep in mind that their memory credentials also include big winners like EPROM and the flash memory that replaced it (another paradigm shift). On the other hand, the ride down memory lane was not without some potholes. (Remember "Bubble Memory"?)

It's a bit ironic that core memory had features (i.e., the accessibility of RAM and nonvolatility of flash memory) that still elude mainstream memory technologies. But maybe that's about to change. Are there memory technologies on the horizon that could obsolete DRAM and flash memory just as they did their predecessors? Paradigm shift anyone?

LIGHT MAKES RIGHT

Given the history lesson above, what better place to start our tour down memory lane than at Intel? Although their 'x86 machinations may get all the headlines, behind the scenes, they're hard at work searching for the next big thing memory-wise.

For now, the current memory technologies (e.g., DRAM and flash memory) are doing just fine. But the chip wizards are saying it won't be long before the party is over for these historic designs as challenges to continued scaling (aka "the wall") loom ever closer.

Stefan Lai's "Non-Volatile

Memories: A Look Into the Future" puts the situation in stark perspective. According to Intel, the good news is that flash memory has gone through eight process migrations, shrinking cell size by almost a factor of 500 from the 1.8 μm original to the currently shipping 90 nm. The bad news is that there are only a couple of more shrinks in the pipeline (65 nm, which Intel just started sampling, and then 45 nm) before current designs hit "the wall." In other words, the party is 90% over.

That doesn't mean DRAM and flash memory are doomed to join vacuum tubes and core memory on the ash heap of computer history. The wizards may yet have a few tricks up their sleeves. For instance, there's talk of potentially increasing the number of bits per cell from 2 (i.e., a four-level cell) to 4 (i.e., a 16-level cell), although the tiny voltage margins (e.g., 0.1 V) would certainly pose a challenge. Nevertheless, march of silicon naysayers have been proven wrong many times in the past, and I suspect flash memory and DRAM will be around for a long, long time. But it will also take a long time for any successor to get ramped up, so there's no time to dillydally.

Although Intel is reportedly investigating many different approaches, the leader appears to be Ovonic Unified Memory (OUM), which is also known as Phase-Change Memory (PCM). One of the main things going for this new idea is that it isn't a new idea at all. Stanford Ovshinsky of Energy Conversion Devices first discovered the approach in the '60s. In 1970,

engineers from that company contributed to an article that appeared in *Electronics* magazine describing an "Amorphous Semiconductor Memory."^[2] The other coauthor was, surprise, one Gordon Moore of then brand new startup Intel.

Ironically, favorable prospects for OUM chip memory owe a debt to advances in rotating memory, namely read-write CDs and DVDs. Thanks to the materials refinement and production know-how gained with optical drives, OUM is now poised to live up to the promise made nearly 50 years ago.

As you may recall, read/write optical drives exploit the property of chalcogenide glasses to change state when heated. In a CD-RW or DVD-RW, a laser is used to provide the heat, and the property of the glass that is exploited is optical, namely the difference in reflectivity between the amorphous (dull) and crystalline (shiny) states.

An OUM uses basically the same material and is conceptually similar, but with two key differences (see Figure 1). First, IC process-formed heating elements replace the optical drive's laser. Second, the state change is measured electrically rather than optically (i.e., resistance is high in the amorphous state and low in the crystalline state).

As far as blue-sky stuff goes, OUM is a relatively real deal. Besides Intel, other heavyweights signed-up include Samsung and STMicroelectronics. Significant test chips that go well beyond the usual dog and pony show prototype have already been produced, such as a 256-Mb unit from Samsung.

Remember in *The Graduate* (1967)

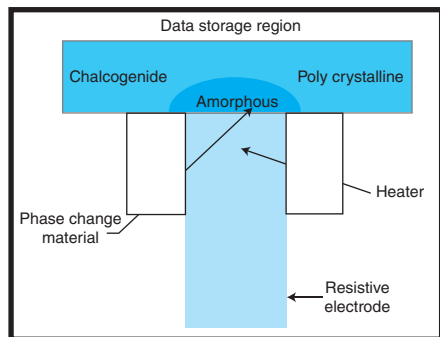


Figure 1—Look, Ma, no transistors. Ovonic Unified Memory and other transistor-less media technologies are preparing for the day DRAM and flash memory hit "the wall."

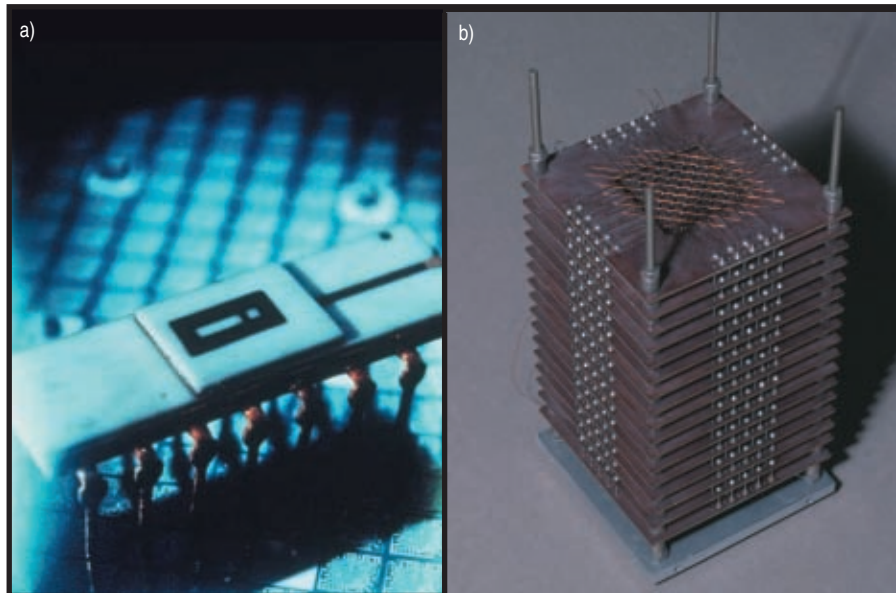


Photo 1—Almost overnight, DRAM (a) killed core memory (b). It's not a matter of if a new memory technology will return the favor, but it's definitely a matter of when. (1103 DRAM photo credit: www.cpu-museum.com/Memories_e.htm. Core photo credit: <http://physics.indiana.edu/~meyer1/historinstr/pics/coremem1/page1.htm>.)

when Dustin Hoffman gets the one-word career advice "plastics"? Plastic indeed may be the future if another memory technology Intel is exploring, Polymeric Ferroelectric RAM (PFRAM), pans out. The beauty of this scheme is that the memory cell itself, like OUM, has no transistors (and thus no transistor scaling issues). Furthermore, layers can be stacked for even higher density.

Indeed, the 3-D approach is an intriguingly simple, if admittedly brute-force, approach to cramming more functionality into less space. We're already seeing the benefits in the variety of "stacked die" parts appearing on the market. Matrix Semiconductor is an outfit that takes the concept one step further. They whimsically observe that silicon wafer costs are now approaching \$1 billion per acre. So, instead of stacking die, a better approach is to make "high-rise" chips that stack layers of memory on a single die (see Photo 2). The only downside is that their anti-fuse-based chips are only one-time programmable (OTP). Nevertheless, Matrix is on a roll, as evidenced by their recent acquisition by SanDisk (for a cool \$250 million) and a list of investors that includes heavyweights like Sony, Nintendo, Seagate, and Kodak. Note that the OTP limitation could actually

be a "feature" for certain applications such as piracy-proof prerecorded media or "black box" data loggers.

Like current memory chips, all of the approaches discussed so far are so-called "XY" addressable. That means lithography limits remain a challenge for accessing the memory elements even if the elements themselves are no longer lithography-bound (i.e., they aren't transistor based).

Once again, taking a page from the rotating memory playbook, so-called "seek-and-scan" memories separate the read/write and access function in much the same way as disk drives have a read/write head "seeking and

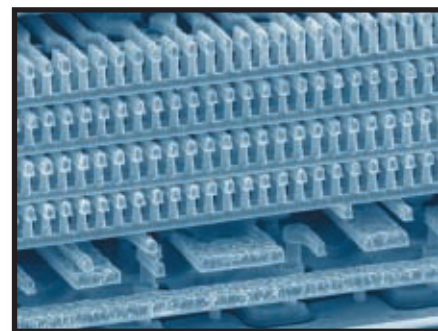


Photo 2—According to Matrix Semiconductor, which was recently acquired by SanDisk, when the real estate gets expensive, it's time to build up as shown here with one of their "four-story" chips. They cost less than flash memory, but they're only one-time programmable, which may be an acceptable or even desirable trade-off in certain applications.

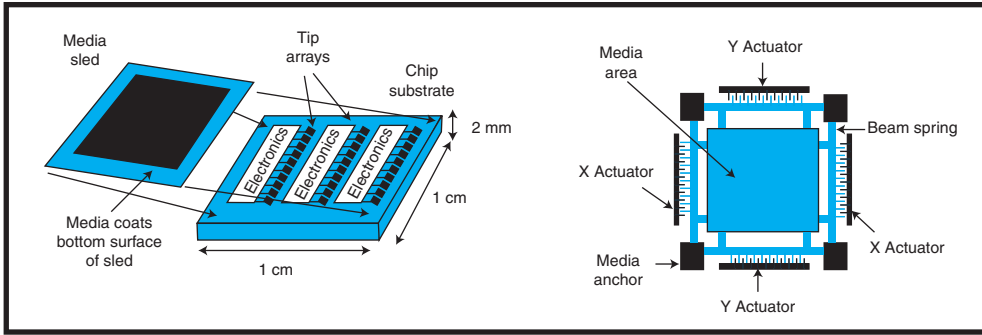


Figure 2—A disk drive is certainly an electromechanical structure, so why not use MEMS technology to micro-size it? MEMS disk drives use a “media sled” and array of probes to emulate the rotating platters and seeking read/write heads of a disk drive.

scanning” large platters of media.^[3]

Companies like IBM and HP have done pioneering experimentation with seek-and-scan. Another new company, Nanochip, has recently licensed the OUM technology described earlier. Although details are sparse (the company’s web site is under construction), by pairing OUM media with MEMS-based seek-and-scan, the company appears to be shooting for something like the IC equivalent of an optical drive. Instead of rotating platters, an OUM “media sled” will be XY positioned using MEMS “springs” (see Figure 2). It’s an interesting idea, and it’s also interesting to note that Intel is an early round investor in Nanochip.

Presumably, beyond all of these lie various forms of “nano,” “molecular,”

and “atomic” memory. This is kind of fun to think about because there are on the order of 5 sextillion (5×10^{21}) atoms in a drop of water.^[4] Considering, for example, that a 500-GB disk drive holds roughly 5×10^{12} bits, that means a drop of bloatware-be-damned water could hold as much data as a billion disk drives!

SUPER RAM

It’s a bird! It’s a plane! It’s SuperRAM!

Er, Earth to Tom: time to get back to reality. OK, how about we take a look at some new-school memories you can buy today (see Figure 3), namely magnetoresistive RAM (MRAM) and ferroelectric RAM (FRAM).

I suppose there’s a bit of irony in the fact MRAM works quite like the core

memory of yore, with a magnetic tunneling junction transistor standing in for the little magnetic doughnuts.

Meanwhile, FRAM uses a special ferroelectric film that is electrically polarized, which kind of reminds me of that nonvolatile LCD I covered a while back (“Batteries Not Included,” *Circuit Cellar* 182, September 2005). MRAM and FRAM have advantages over flash memory when it comes

to writing in terms of speed, simplicity, and power consumption.

As one of the pioneering FRAM suppliers, Ramtron deserves credit for sticking with it through the tough challenges of pioneering a new technology. I’d say their new 1-Mb part is a good payoff, one that can actually make sense in real applications.

Notably, the Ramtron FM20L08 adopts the classic JEDEC “byte-wide” SRAM pinout and functionality. As a point of fact, this means it can target sockets for battery-backed SRAMs, which typically require large high-profile DIP packages to accommodate the battery and can have issues in manufacturing (e.g., cooking the battery). The FM20L08 not only ditches the battery, but also requires less board

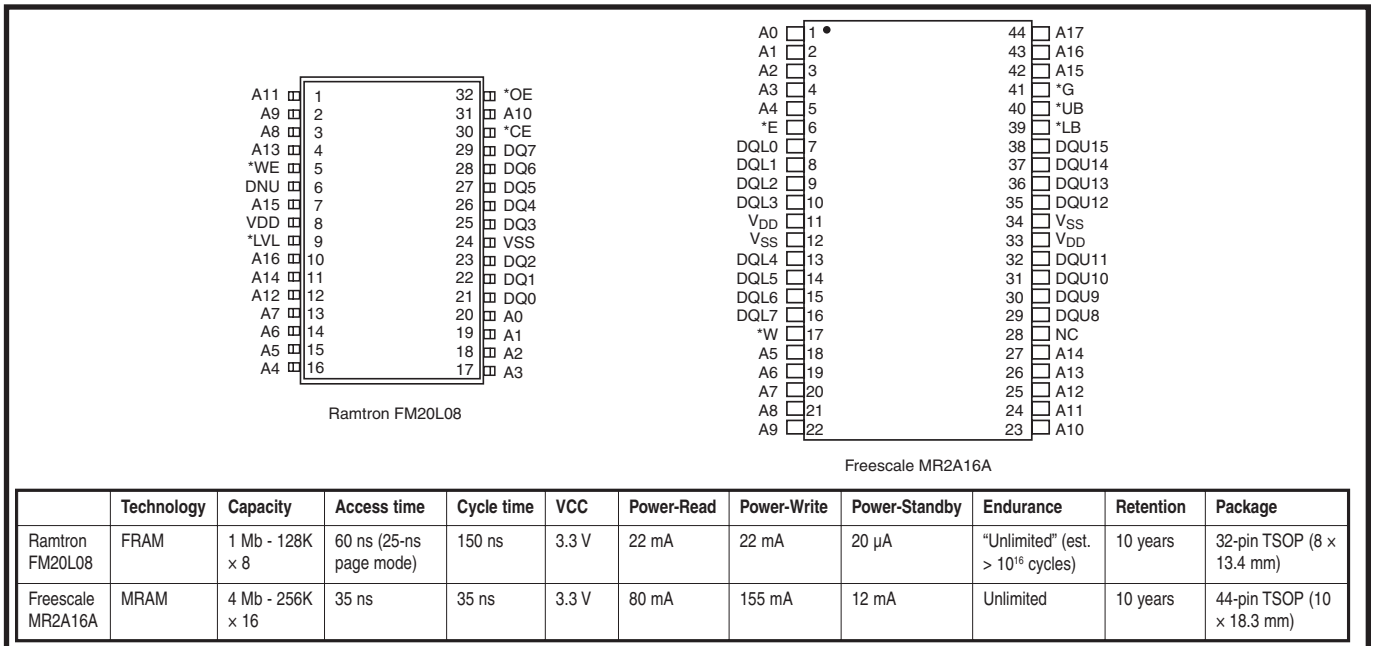


Figure 3—Although not dense or cheap enough to challenge DRAM and flash memory in mainstream applications, FRAM and MRAM are now viable options for designs that can take advantage of their unique capabilities.

space and goes through manufacturing just like any other chip.

The FM20L08 does a decent job impersonating an SRAM with the familiar pins (i.e., *CE, *OE, *WE, etc.) working more or less the usual way, although there are a few differences. For an SRAM, the cycle time and access time are usually the same. But the FM20L08 access time is 60 ns, while the cycle time is a more leisurely 150 ns. Performance-oriented applications can take advantage of a DRAM-like page mode to burst up to 8 bytes at 33 MHz. Another embellishment is a low-voltage lockout function to prevent spurious writes during power failure. After all, what good is a nonvolatile memory if the data that gets written into it is erroneous? Finally, there's a version of the chip available that includes a software write protect scheme as an additional way to harden the data against software and hardware glitches.

Meanwhile, Freescale is in volume production with their 4-Mb (256K × 16) MR2A16A MRAM chip. Sharing the FRAM's 3.3-V supply and low-voltage lockout features, the MR2A16A feels even more at home in SRAM sockets thanks to equal—and at 35 ns, fast—access and cycle times. Although it's an x16 part, separate upper- and lower-byte select pins allow single byte access.

Although their SRAM-like façade is familiar enough, these parts require a different way of thinking when it comes to minimizing power consumption. For example, the Freescale MRAM standby current spec of 12 mA is so much higher than the microamps of an SRAM that I thought it was a typo. What seems a problem has a simple brute force solution, which is just to cut the power to the chip entirely. Yes, there is a 1- μ s delay required after power-up, but that's not a very long time, which means turning off the power for even very short idle periods (e.g., milliseconds) can pay off.

Meanwhile, the FM20L08 features a more typical (i.e., SRAM-like) standby current of 20 μ A. On the other hand,

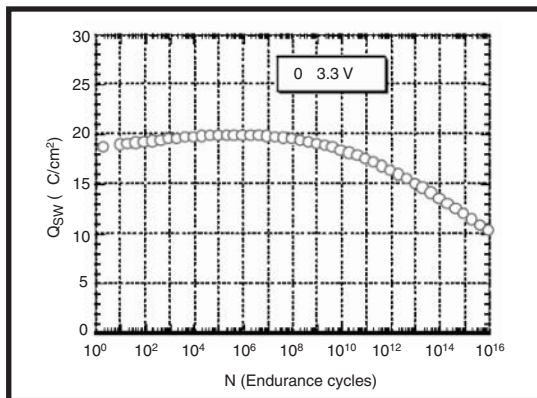


Figure 4—If not truly a concern in practice, the “destructive read” characteristics of FRAM inspired nagging doubts about endurance for read-mostly-and-often applications. According to Ramtron, those doubts should be laid to rest for their new 0.5- μ m, 3.3-V parts. As shown here, they predict that even after 10^{16} cycles, the polarization density of $10 \mu\text{C}/\text{cm}^2$ will safely exceed the minimum sensing threshold of $6 \mu\text{C}/\text{cm}^2$.

its power-up delay spec is much more significant at up to 5 ms. Thus, the “standby” versus “power-down” trade-off is quite different for the FM20L08 FRAM than for the Freescale MRAM.

Meanwhile, active power consumption characteristics for the two parts are also quite different. The 'L08 FRAM gets relatively decent mileage with power consumption of 22 mA (maximum) for both reading and writing. For the MR2A16A MRAM, read power consumption is up to 80 mA (maximum 55 mA typical), while write current is nearly double that at as high as 155 mA (105 mA typical).

However, note that actual system power consumption may differ from that implied by the raw datasheet specs due to the functional differences between the parts, starting with the fact the MRAM has four times the number of bits. In addition, power consumed by the MR2A16A MRAM is potentially reduced because it's twice as fast (35 versus 60 ns access time) and twice as wide. Thus, it theoretically may be less “active” than the FM20L08 FRAM. For example, the MRAM requires only a single 35-ns cycle to move 16 bits, while the FRAM needs two 8-bit cycles to accomplish the same task, which requires at least 100 ns (page mode) or 300 ns (non-page mode).

In either case, it's an improvement over flash memory. Whether handled

by a coprocessor or software driver, writing a flash memory chip is a long and complicated process that consumes numerous clock cycles to erase and rewrite entire sectors. And don't forget what the high-voltage (e.g., 15 V) flash memory requires for writing. More instructions, more clocks, more bits, and more volts: it all means more power.

Besides the power consumption itself, flash memory machinations (e.g., busy bit babysitting) are rather an annoying and perhaps untimely chore for the processor. One key concern is what to do if the lights go out midway through the relatively long and unforgiving window of

flash memory vulnerability during a sector erase/write sequence. By contrast, to write an MRAM or FRAM you just, well, write it. One instruction, one cycle, one less thing to worry about.

Beyond the Memory 101 basics (speed, power, and price), the Holy Grailness of next-generation memories is all about the nonvolatility. Are these things robust or doomed to silicon senility? On the one hand, early prototypes were definitely prone to a variety of data-disrupting fatigue, stuck bit and write disturb phenomena. Indeed, there are nontrivial challenges like simply figuring out how reliable the memory is short of running a memory test for 10 years.

A notable spec has to do with the “destructive” characteristics of read cycles for FRAM. Like a DRAM, the ferroelectric capacitor-based FRAM requires data to be refreshed (i.e., rewritten) following a read cycle. Although the cycle endurance might be huge, this could prove problematic for applications that are “read mostly and often,” such as using an FRAM to store code run by a high-performance processor.

Early FRAMs had endurance specs on the order of 10^8 , or 100 million, cycles. The good news is even this first-generation spec was 100 to 1,000 times the right endurance spec for a typical flash memory. But crank the numbers and you can see the potential problem of using FRAM for execute-

in-place (i.e., op-code fetch) applications. Consider that 100-million cycle endurance spec in light of the fact that the FM20L08 could theoretically deliver more than 6 million cycles per second. Even assuming a more leisurely 1 million cycles per second, trouble could be mere minutes away.

Subsequent work to extend FRAM endurance paid off with improved specs on the order of 10^{12} , or 1 trillion, cycles. Now our 1-million-cycles-per-second application can run for 1 million seconds, which sounds good until you realize that is less than 12 days!

But, as represented by the FM20L08, the new generation of FRAMs have made dramatic progress with the latest endurance estimates on the order of 10^{16} ! Now we're talking something like 12,000 days, or 30-plus years, of 1-million-cycle-per-second operation (see Figure 4).

For MRAM, endurance is less an issue because access (flipping magnetic polarity like a disk drive) has no intrinsic wear-out mechanism. Nevertheless, earlier parts suffered from a "single-line disturb" phenomenon, which is essentially a crosstalk problem in which writing one bit could change others. With their introduction of the MR2A16A, Freescale solved the problem with a clever "toggle" approach. As the name implies, the memory is not set to 1 or 0 directly, but it's read and then toggled if necessary to achieve the desired state. The current waveforms applied on the word and bit lines for the toggle approach are less likely to disturb adjacent bits.

Yes, there may be issues for MRAM and FRAM lurking under the hood that will only become apparent over time. But for now, all you can really do is take the datasheets at their word, including the assertion on the Ramtron FRAM datasheet that it features "unlimited read/write cycles." In terms of data retention, both the Freescale and Ramtron parts specify 10 years. That's a spec I've got application concerns about, although it does not by any means imply that the chips get amnesia after 10 years and a day. Actual environmental and use characteristics have a big impact on retention, and extremes are unlikely in

actual practice. Because that 10-year spec is still widely used by flash memory suppliers (although there's a trend to increase it), let's leave the subject aside for now. The bottom line is that the specs for both the Freescale and Ramtron parts say you can treat the thing like a RAM, but it will hold the data like a ROM, at least for 10 years.

THANKS FOR THE MEMORIES

The latest generation of MRAMs and FRAMs work great, but they are by no means the Holy Grail memories that will lay the old to rest overnight. The new chips are just getting over teething pains and simply don't offer the density or cost per bit to displace DRAM or flash memory in mainstream markets. For now, they're niche players. But with that being said, they're still kind of cool.

Stand-alone memory chips are just one part of the story, and frankly the lesser part from my embedded perspective. When it comes to replacing existing memories, I'm just as interested in the memory found on high-integration MCUs and SoCs (notably including FPGAs) that package everything in a single chip. Yes, OUM, FRAM, MRAM, and all the others face challenges when it comes to integration with standard (i.e., CMOS) silicon processes. In the meantime, stacked die is an option. In any case, suppliers claim that none of the issues are showstoppers and all are targeting future integration with standard logic functions. Doing so would highlight the advantages of the new memory technologies (e.g., flexibility, speed, endurance, and power consumption) while diluting the disadvantages (density and price per bit).

Flash memory and DRAM will be around a long time. But that doesn't mean the new technologies have to wait around. It's not a zero sum game, and there are plenty of existing and as yet unforeseen applications to support a variety of new technologies. I say bring 'em on. ☒

Tom Cantrell has been working on chip, board, and systems design and marketing for several years. You may reach him by e-mail at tom.cantrell@circuitcellar.com.

REFERENCES

- [1] "Elpida Memory Ships Samples of 2 Gigabit DDR2 SDRAM using 80 nm Process Technology," www.physorg.com/news9189.html.
- [2] R. Neale, D. Nelson, G. Moore, "Non-Volatile and Reprogrammable, the Read-Mostly Memory is Here," *Electronics*, September 1970.
- [3] J. L. Griffin et al, "Operating Systems Management of MEMS-based Storage Devices," Fourth USENIX OSDI Symposium, www.usenix.org/publications/library/proceedings/osdi2000/full_papers/griffin/griffin_html/index.html.
- [4] M. Onken, Admin, MadSci Network, 2000, www.madsci.org/posts/archives/2000-10/971190308.Ch.r.html.
- [5] F. Chu and T. Davenport, "The Endurance Performance of 0.5um FRAM Products," Ramtron International Corp., www.ramtron.com/lib/literature/techpapers/Endurance%20of%200.5u%20FRAM%20Products.pdf.

RESOURCES

- Ovonic Unified Memory and Polymeric Ferroelectric RAM, Intel Corp., www.intel.com.
- S. Lai, "Non-Volatile Memories: A Look Into the Future", Intel Corp., <ftp://download.intel.com/technology/silicon/Stefan-SRI-Stanford-Berkeley-MIT-022604.pdf>.
- MEMS Ovonic Unified Memory Disk, Nanochip, Inc., www.nanochip.com.
- Ovonic Unified Memory, Ovonyx, Inc., www.ovonyx.com.
- Matrix 3D One-Time Programmable Memory, SanDisk Corp., www.sandisk.com.

SOURCES

- Magneto-resistive RAM**
Freescale Semiconductor, Inc.
www.freescale.com
- Ferroelectric RAM**
Ramtron International Corp.
www.ramtron.com