 By Brian Dipert, Technical Editor

Hitting their stride

Nonvolatile-memory upstarts draw near to established leaders

ONE NEXT GENERATION TECHNOLOGY IS FINALLY ON A LEADING-EDGE PROCESS, ANOTHER IS NEARING PRODUCTION AFTER YEARS OF DEVELOPMENT, AND A THIRD HAS ITS FIRST HIGH-VOLUME CUSTOMER. THE NONVOLATILE-MEMORY RACE HAS, FOR THE FIRST TIME IN A LONG TIME, BECOME COMPETITIVE AGAIN.

IT'S BEEN MORE THAN 15 YEARS since fast-random-access NOR flash memory's initial public unveiling by Intel in 1988. NOR quickly gobbled up all of the memory sockets that EPROM had previously serviced, despite that EPROM technology was more than a decade and a half *more* mature (**Figure 1**). Nowadays, NOR flash memory is big business; Semico analyst Jim Handy estimates this year's revenue at $13.362 billion.

NAND flash memory is even older than NOR—two decades old last year. Initially, NAND slowly ramped up its year-to-year incremental volume shipments, but of late it's been on a tear, hotly fueled by the success of markets its unique attributes helped create, such as digital audio players, digital still cameras, and USB memory sticks (**references 1** and **2**). This year, it and closely related AND flash memory together comprise a $4.442 billion market, according to Semico. Ironically, although the NAND-plus-AND revenue will be approximately only one-third that of NOR in 2005, NAND-plus-AND shipments measured in aggregate density will be *more than three times* those of NOR.

At a glance......56
Slow and steady......56
For more information..60

Flash memory sounds like a healthy, smooth-running business, right? Why, then, are all of the major NOR and NAND suppliers, along with a slew of other semiconductor manufacturers, busy developing and ramping production on next-generation nonvolatile memories? Intel's strategic marketing manager, Greg Komoto, explains, "Smart vendors begin rolling out next-generation technologies approximately five years before existing tech-

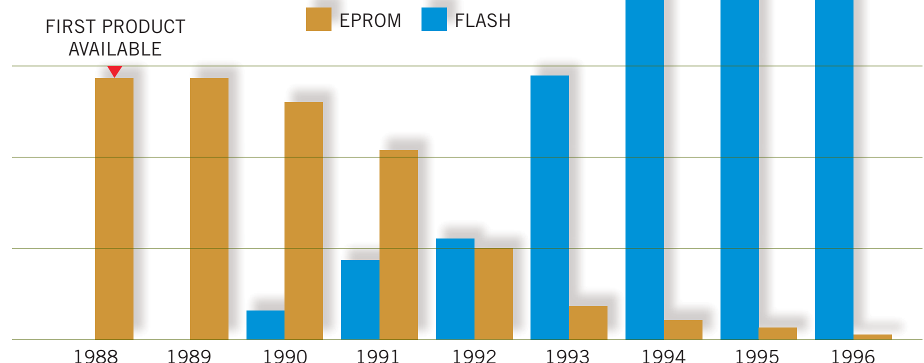


Figure 1

Within five years, NOR flash-memory shipments had surpassed those of EPROM; EPROM soon faded away (courtesy Intel).

nologies reach their end of life" (**Figure 2**). And, in fact, flash-memory suppliers are finding it increasingly difficult to develop robust chips as process lithographies dip below 100 nm.

Deep-submicron transistors cannot tolerate the high voltages employed during the CHE (channel-hot-electron) injection and FN (Fowler-Nordheim) tunneling program and erase processes used to place charge into and remove it from the flash-memory cell's floating gate. Overload of the thin, fragile, insulating oxide layer between the floating gate and the select gate and between the floating gate and the source, drain, and substrate lead to spurious disturbance of floating-gate electrons, causing inadvertent programming or erasing. The phenomenon is particularly acute with multilevel-cell flash memories that rely on a precise amount of stored charge on the floating gate to reference a particular 2-bit value combination.

AT A GLANCE

▷ Next generation nonvolatile memories are coming on strong but face stiff competition from established approaches.

▷ The FRAM transition to low-cost 1T1C (one-transistor one-capacitor) cells is nearly complete, and advanced processes and highly integrated derivative chips bode well for the technology.

▷ MRAMs require specialized low-temperature manufacturing processes but deliver high-performance reads and writes—albeit at high power consumption for writes.

▷ XPM and 3DM memories prove that simplest is sometimes best.

▷ Today's memory-market leaders are taking a conservative approach to next-generation-technology development.

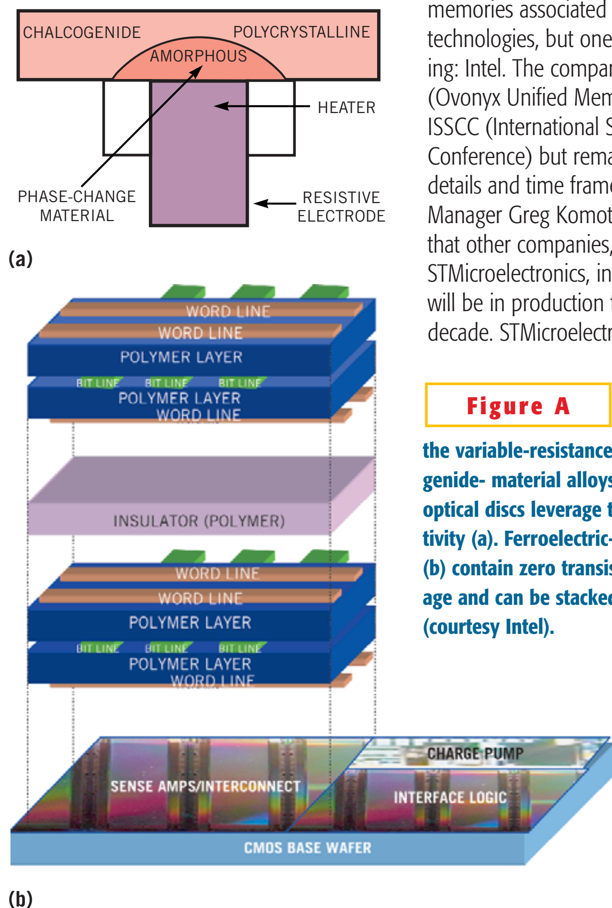
FRAM (ferroelectric-RAM), MRAM (magnetic-RAM), and other next generation technologies are all attempts to develop the "perfect" memory: One that is nonvolatile, whose bits you can fully alter, with ultrafast read and write rates and an infinite number of rewrite cycles. None of them succeeds in all areas, but all of them make key advancements in at least some of these important memory characteristics. (See **references 3** and **4** for more on these next-generation technologies.)

Technology and device architects are attempting to develop memories that not only exhibit the aforementioned features, but also achieve the key criteria of low cost per bit in its various manifestations (**Table 1**). These cost criteria include conventional CMOS-process compatibility both in materials and in manufacturing flow, few to no incremental processing steps over standard logic and memory structures, and no need for high or oth-

SLOW AND STEADY

The main text of this article mentions many of the biggest names in semiconductor memories associated with up-and-coming technologies, but one key company is missing: Intel. The company unveiled its OUM (Ovonyx Unified Memory) plans at the 2002 ISSCC (International Solid State Circuits Conference) but remains mum on product details and time frames. Strategic Marketing Manager Greg Komoto, however, points out that other companies, such as Samsung and STMicroelectronics, indicate that the devices will be in production toward the end of this decade. STMicroelectronics, in fact, delivered a paper on what it calls Phase Change Memory at last month's IEDM (International Electron Devices Meeting).

Intel intends OUM to be the successor to its NOR flash memory, thereby delivering high random read performance (**Figure Aa**). For high-density NAND-like data storage, Intel's "primary internal program" is ferro-electric-polymer memory (**Figure Ab**). But, clearly, nothing's set in stone; the company's literature points out that it is currently investing in "seven technologies, from university research to internal programs, covering all major classes of promising materials" and "focused on data- and code-storage technologies." In 2003 alone, the company claims, it reviewed more than 10 memory concepts from industry start-ups and academic research labs. However, Intel literature maintains that the company sees "no creditable challenger to flash memory through the end of the decade."



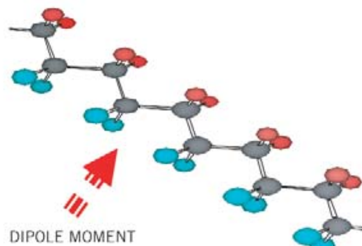


Figure A **Phase-change memory employs the variable-resistance attributes of chalcogenide- material alloys, whereas writable optical discs leverage their variable reflectivity (a). Ferroelectric-polymer memories (b) contain zero transistors per bit of storage and can be stacked in three dimensions (courtesy Intel).**

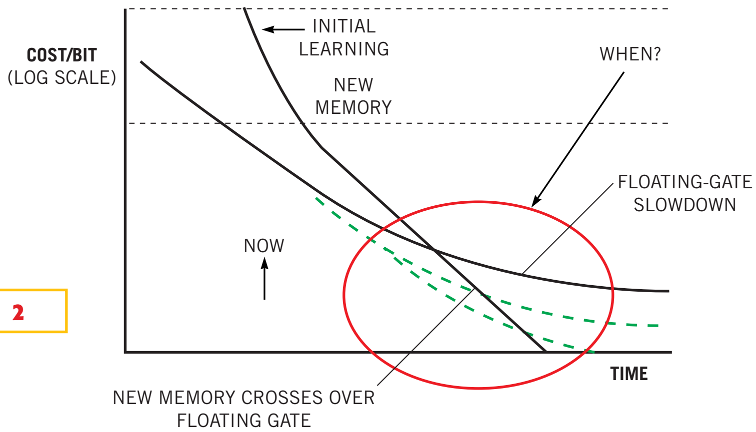
erwise-nonstandard internal voltages. Other, expense-minimization factors include the smallest possible cell dimensions at a given lithography's feature size and the ability to smoothly scale the memory's storage cells and other circuits from one lithography to another, smaller process.

FRAM ADVANCES, DIVERSIFIES

A perusal of Ramtron's corporate fact sheet from the third quarter of 1996 effectively explains why some potential customers might have viewed FRAM with skepticism over the past few decades. In 1996, Ramtron predicted that partners Fujitsu and Toshiba would both begin producing 1-Mbit FRAMs in 1998, and that 4- and 16-Mbit devices from multiple companies would follow shortly thereafter. The 1-Mbit FRAM didn't actually appear until February 2001, and then only in the form of a jointly authored ISSCC (International Solid State Circuits Conference) paper from Fujitsu and Ramtron. Fujitsu's 1-Mbit product announcement occurred just last November; Ramtron previewed its own 1-Mbit FRAM chip (different from the Fujitsu device) that same month at Europe's Electronica trade fair. Ramtron's past predictions may have been overenthusiastic, but there's still sufficient evidence to conclude that FRAM's future prospects are solid.

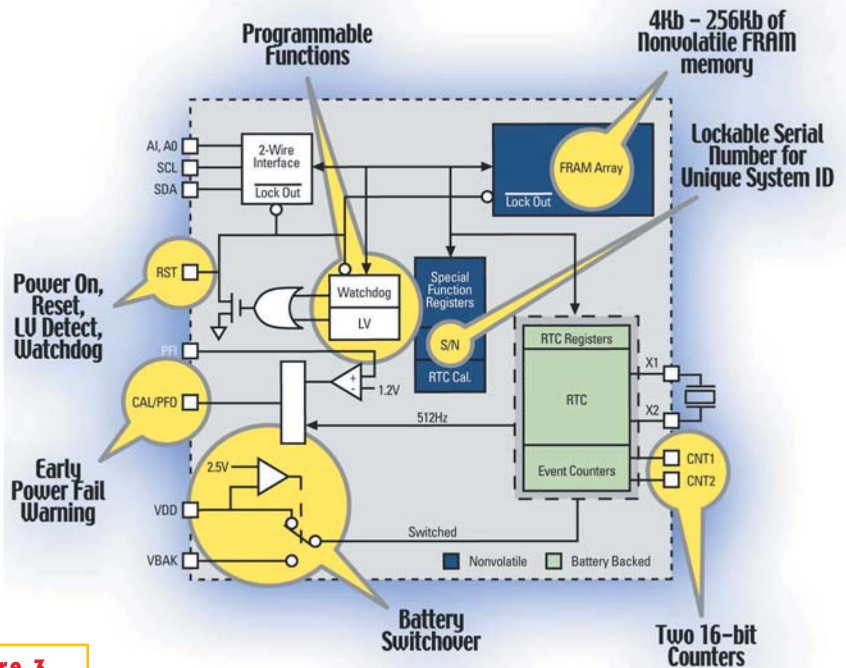
The previously mentioned 1-Mbit FRAM is a 1T1C (one-transistor, one-ferroelectric-capacitor) memory, much less expensive on a cost-per-bit basis than past 2T2C devices. Fujitsu's first-generation part, which comes in 8- and 16-bit-interface variants, targets data-storage applications; its specifications are conservative, with 250-nsec read and write timings. Ramtron asserts that its version of the 1-Mbit device will be more aggressive in its read and write specifications and therefore also be

Figure 2



Introduce a new technology too fast, and your company wastes money; wait too long, and you lose your market momentum (courtesy Intel).

Figure 3



Processor Companions bundle a variable-sized FRAM array with various other essential system peripherals (courtesy Ramtron).

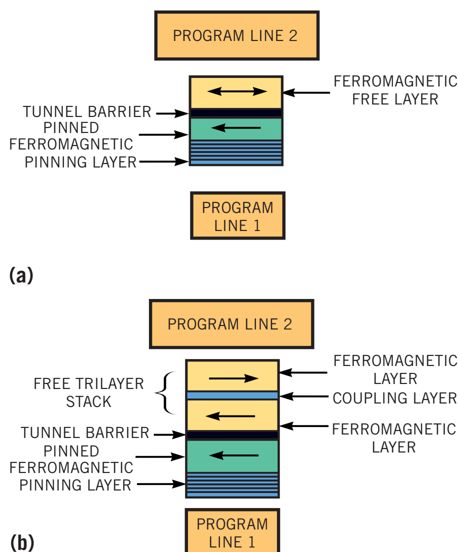
TABLE 1—MEMORY CHARACTERISTICS

	MRAM	FRAM	3DM	Ovonyx Unified Memory
Cell size	Large	Moderate	Very small	Small
CMOS-integration issues	Intolerant of high temperatures, requires specialized materials	Intolerant of hydrogen exposure, requires specialized materials	Requires specialized materials	Requires specialized materials
Read speed	Fast, nondestructive	Moderate, destructive	Slow (random), moderate (sequential), nondestructive	Moderate, nondestructive
Write speed and power consumption	Fast, high power	Moderate speed, moderate power	Slow, moderate power	Moderate speed, moderate power
Cycling endurance	Theoretically infinite	10 billion to 1 trillion, claimed "unlimited" on latest generation devices	Not applicable (one-time programmable)	As many as 1 trillion claimed

suitable for direct code execution.

The cycling specifications of FRAM are increasing from one product and process generation to another; this improvement is particularly notable in light of the simultaneous conversion to 1T1C cells. (One key function of 2T2C was to increase program and erase margins and thereby maximize a device's cycling capabilities.) Ramtron, in fact, claims *unlimited* cycling on its latest generation parts, even at extended-temperature operating conditions. Cycling is a particularly critical parameter for FRAM because of the technology's DRAM-like destructive read behavior: The device automatically rewrites the value previously stored in the ferroelectric capacitor at the conclusion of a read. There must be at least *some* reality behind the company's optimistic claims. Ramtron's 64-kbit SPI FRAM, for example, last October achieved AEC Q100 automotive qualification.

Finally, the industry has not yet felt the impact of Ramtron's partnership with Texas Instruments, first unveiled at the December 2002 IEDM (International Electron Devices Meeting). Whereas today's manufacturers produce leading-edge FRAMs on relatively trailing-edge 0.35-micron processes, a device currently under development by the two companies employs TI's state-of-the-art 0.13-micron process. The chip has completed tape-out, the two companies report, and first samples are due to appear by press time. The dual-mode device operates as a 1T1C, 8-Mbit memory or a 2T2C, 4-Mbit FRAM; Ramtron and TI claim minimal die-size impact due to the incremental dual-mode circuitry. The intent, if the process and product are "clean," is for Ramtron to take the 8-Mbit variant to production; TI will act as Ramtron's foundry and has rights to the technology for use in embedded arrays on ASICs.



A commonplace MRAM cell is susceptible to disturbance from nearby cells' program and erase operations (a), whereas an advanced cell design is more immune to these effects (b). Conventional (c) and crosspoint (d) cells have different cost and performance characteristics, suitable for diverse applications (a and b courtesy Freescale, c and d courtesy Infineon).

Whether Ramtron will spend the time, money, and effort to go back and design a 1T1C, 4-Mbit device is unclear. Ramtron Vice President Mike Alwais points out, "A 2T2C, 4-Mbit FRAM on 0.13 micron is still much cheaper on a cost-per-bit basis than a 1T1C 0.35-micron, 1-Mbit FRAM." The company and its partners will continue to focus on discrete FRAMs and embedded-FRAM arrays; in parallel, Ramtron has rolled out a line of FRAM-based Processor Companion chips that include circuits such as real-time clocks and alarms on some devices, watchdog timers, power monitors, and embedded serial numbers (**Figure 3**). They're conceptually similar to the battery-backed SRAM- and EEPROM-based multifunction chips that other companies sell.

MRAM INCHES TOWARD PRODUCTION

One longstanding FRAM problem that Ramtron and its partners claim to have finally overcome involves hydrogen, a common byproduct of semiconductor processing that degrades the ferroelectric capacitor's capabilities. Encapsulating the capacitor, thereby mitigating its hydrogen exposure, reportedly did the trick. FRAM's primary upstart competitor, MRAM, has its own weakness: The magnetic material it employs for storage within the cell is sensitive to high temperatures commonly found in subsequent processing steps. According to Michael Haight, Freescale's strategic marketing manager, "Some silicon-CMOS back-end processing steps, like

FOR MORE INFORMATION...

For more information on products such as those discussed in this article, contact any of the following manufacturers directly, and please let them know you read about their products in *EDN*.

Altis Semiconductor
www.altissemiconductor.com

Cypress Semiconductor
www.cypress.com

Freescale Semiconductor
www.freescale.com

Fujitsu
www.fujitsumicro.com

Infineon Technologies
www.infineon.com

Intel
www.intel.com

Kilopass Technology
www.kilopass.com

Matrix Semiconductor
www.matrixsemi.com

Mattel
www.mattel.com

Ramtron
www.ramtron.com

Samsung
www.samsungsemi.com

Semico
www.semico.com

STMicroelectronics
www.st.com

Texas Instruments
www.ti.com

Toshiba
www.toshiba.com/taec

TSMC
www.tsmc.com

dielectric deposition, are at 400°C. MRAM does require some modification to back-end process steps to accommodate a maximum-temperature limitation after deposition of magnetic material. Today, we are limiting the temperature to around 250°C after magnetic materials deposition and have already developed lower temperature dielectrics that are compatible with the MRAM process."

Freescale (then Motorola) at the February 2001 ISSCC unveiled a 256-kbit MRAM test chip, built on 0.6-micron technology. A 1-Mbit test chip followed a year later, and, in October 2003, the company began shipping samples of its first generation 0.18-micron, 4-Mbit device. Freescale is now preparing to ramp into production a redesigned 4-Mbit device, with qualification scheduled for late in the first quarter or early in the second. Freescale builds the chip on a five-layer metal process; like TI with FRAM, Freescale has its long-term sights on both discrete memories and embedded arrays within ASICs.

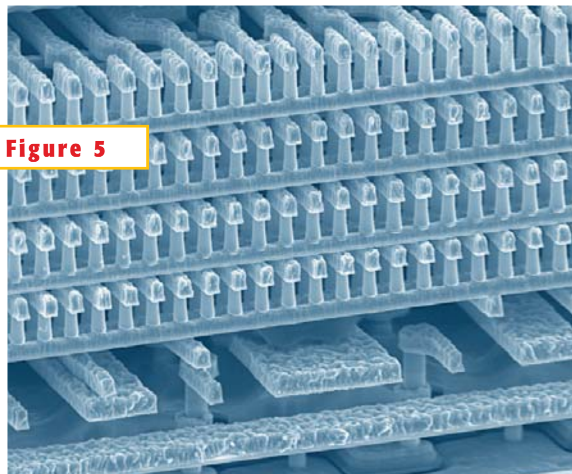
Two key advancements made their way into this second-generation device. First, the company added ATD (address-transition-detection) circuitry. You need no longer toggle the chip-select input between accesses; the MRAM acts just like an SRAM. Freescale confidently boasts that this MRAM, also like SRAM, is capable of infinite rewrite cycling. Unlike FRAM, MRAM does not suffer from cycle-count-boosting destructive reads.

The second advancement addresses the program and erase disturbances that plagued the first generation 4-Mbit MRAM: Altering cell data would affect the integrity of adjacent cells' stored information. Freescale has incorporated the Savtchenko Bit Cell, an advanced storage-element approach named after its late Russian inventor Leonid Savtchenko (**Figure 4**). The Savtchenko Bit Cell replaces the single free layer of a conventional FRAM with a three-layer alternative containing two ferromagnetic layers at opposite polarities. The approach increases program and erase margin. MRAMs are noteworthy not just for their fast read and write speeds (Freescale's part specifies 25-nsec cycle times for both), but also for their high write current; Haight estimates 10 mA per bit programmed or erased. For a 16-bit-interface device, the total current requirement for a word of data could approach 200 mA. Statistically, however, such a situation is unlikely; not *all* bits within a given word are normally "flipped" during each write cycle. The current draw also takes place over a short time; the write *energy* consumption of MRAM is comparable with that of other, slower memory technologies.

Although Freescale asserts a firm commitment to MRAM, and Texas Instruments is similarly devoted to FRAM, other companies, especially those with deep pockets, are hedging their bets by developing multiple next generation memory technologies in parallel. At last June's IEEE Symposium on VLSI technology, Infineon presented a paper on an ultrasmall FRAM cell with a vertical capacitor structure and size comparable with a NAND flash-memory cell, the result of co-development work with Toshiba. Also in June, at the related IEEE Symposium on VLSI Circuits, the company unveiled a 30-nsec, 16-Mbit MRAM, built on a 0.18-micron process and co-developed with IBM. It will subsequently transfer this device to French joint-venture company Altis, which will evaluate its feasibility and fine-tune its characteristics for production.

Infineon's Memory Chief Technical Officer Wilhelm Beinvogl, at a briefing in Dresden, revealed that the company had as of last June developed three FRAMs: a 4-Mbit device on a 0.25-micron process and 16- and 32-Mbit FRAMs on a 0.2-micron process. Beinvogl demonstrated functional FRAMs running on an evaluation board but cautioned that the company is at least several years away from high-volume production. Beinvogl also confirmed that Infineon is evaluating both conventional MRAMs with fast cycle times and resistive XPC (crosspoint-cell) devices with NAND-like slow random-access speeds that, by discarding the per-cell transistor, are ultradense and stackable.

Figure 5



Three-dimensional memories place the periphery circuitry below the storage arrays; conventional approaches place them alongside each other (courtesy Matrix Semiconductor).

OTP TAKES CENTER STAGE

What is the perfect memory? One attribute that Intel's Komoto mentions is the ability "to eliminate the memory cell transistors, thereby eliminating the potential for transistor scaling issues at advanced lithographies and enabling 3-D array stacking." Komoto alludes to the ferroelectric polymer memory his company is evaluating, but he could have been describing XPC MRAM (see **sidebar** "Slow and steady"). He also could have been describing Matrix Semiconductor's 3DM, now ramping into production in Mattel's Juice Box personal media player and targeting high-density data-storage applications that might otherwise be candidates for mask ROM or a writable optical disk with a system-housed drive.

A cross-section of 3DM reveals in its bottom layer conventional CMOS transistors that implement address and control signal decoding, along with various state-machine functions (**Figure 5**). Above those transistors are multiple layers of storage arrays with an antifuse and a diode at each crosspoint link (conceptually similar to 3D-ROM, which **Reference 3** mentions). A layer of metal lines at the top provides signal interconnection. Experience with antifuse FPGAs might lead you to incorrectly deduce that their tight reliability constraints and manufacturing challenges similarly restrict antifuse memories. In an FPGA, the antifuse passes a logic signal and must therefore be robustly "grown" and have ultralow impedance. In 3DM, however, a "short" or "open" measures the presence or absence of the antifuse, more simplistically determining whether the stored bit has been programmed or is, as shipped from the factory, still erased.

The 3DMs that Matrix is now shipping represent the third generation of the technology. The company's first chips were developed in a Cypress Semiconductor research fab through a partnership. Matrix built the second generation of 3DM, and the first devices that saw limited production, on a 0.25-micron process. These latest chips, from TSMC's 0.15-micron foundry, come in both NAND flash-memory, pinout-compatible TSOPs and MMCs (multimedia cards), and they derive from four die: a 64-Mbyte, four-layer device; two- and four-layer 32-Mbyte devices; and a 16-Mbyte, two-layer device. They're fully read-function-compatible with NAND flash memory, and they resolve read-performance differences using the ready/busy pin and status register bit. (The 3DM devices can sustain 2 Mbytes/sec.)

Write differences between 3DM and NAND flash memory reflect 3DM's OTP nature. You must issue a unique multicommand sequence to the 3DM before you unlock it for writes unless you override the requirement by tying an input pin high. Reflecting the devices' internal EDC (error-detection and -correction) scheme, you also must write data in 64-bit packets minimum, with the data octal-byte-aligned. As many as 643 bad pages can exist within each 16-Mbyte zone—a characteristic reminiscent of the bad-block specification of NAND flash memory that media-management hardware or software can handle. (Matrix also designed the controller in its MMCs.) The company claims that 3DM's sustained write performance is 1 Mbyte/sec. Matrix believes that 3DM can operate at extended temperatures but has yet to fully characterize the memory beyond 0 to 70°C. The company has successfully constructed 90-nm test circuits and feels that the technology is scalable at least down to the 45-nm-process generation.

Kilopass takes the same general concept of antifuse-based OTP memory in a different application direction with its XPM technology, which it intends to hold not only data, but also directly executable code and to embed in arrays within ASICs. Kilopass employs an unspecified semiconductor physics phenomenon to break down a conventional CMOS transistor's gate oxide, creating a "short circuit" that a sense amplifier can subsequently detect. This transformation directly programs the memory cell and thereby does not require a separate floating gate, storage capacitor, or other components. The company points out that, unlike XPM, you cannot easily integrate polysilicon-dominated flash memory within a metal-rich CMOS logic chip, and logic-friendly EEPROM requires multiple transistors per bit, along with high voltage and other costly characteristics. XPM memory contents are also difficult to reverse-engineer, leading to improved design security (**Reference 5**). Kilopass' estimate of 1.5 transistors per cell for today's XPM arrays reflects the technology's included EDC circuitry; the company hopes to reduce or even eliminate this overhead logic and memory requirement for future arrays. The 10V programming voltage, which the device can internally generate or which you can externally apply to the ASIC, will also drop as the manufacturing lithography dips below 0.18 micron.□

References

1. Dipert, Brian, "Flash forward to the future," *EDN Global Report*, Nov 2004, pg 62.
2. Dipert, Brian, "Pick a card," *EDN*, July 8, 2004, pg 53.
3. Dipert, Brian, "Exotic memories, diverse approaches," *EDN*, April 26, 2001, pg 56.
4. Dipert, Brian, "FRAM: ready to ditch niche?" *EDN*, April 10, 1997, pg 93.
5. Dipert, Brian, "Cunning circuits confound crooks," *EDN*, Oct 12, 2000, pg 103.

Author's biography

Technical Editor Brian Dipert is utterly astounded that it's now possible to buy 1-Gbyte flash-memory cards, after rebate, for less than $50. The road to success for next generation competing memories won't be smooth. Contact him at 1-916-454-5242, fax 1-617-558-4470, bdipert@edn.com, and www.bdipert.com.

Talk to us

Post comments via TalkBack at the online version of this article at www.edn.com.