

# IEEE SPECTRUM

THE MAGAZINE OF TECHNOLOGY INSIDERS

## Winners & Losers 2007 Beauty AND THE BEAST

**B**ig-circulation magazines do some things well, but serious coverage of technology generally isn't one of them. First there are the biz mags, for which technology coverage usually means accounts of financial developments at companies that happen to offer tech-related products or services, as opposed to, say, plumbing fixtures or hockey pucks. Then there are the popular science and technology magazines, which offer a monthly feast of gizmos, vehicles, and pipe dreams. One of them takes UFOs seriously; another put the Bugatti Veyron supercar on its cover recently to pitch its annual roundup of innovation.

Now, the Veyron is indeed a lovely piece of engineering, as we said when we wrote about it—10 months ago in our "Top 10 Tech Cars" report. But there is more to technological innovation than cars, gadgets, and fat IPOs. There's software and semiconductors, for starters, from which have sprung most of the important technical advances of the past 30 years. Without software and semiconductors, the gizmos, flying whatsits, and monster tech successes wouldn't be possible.

For confirmation, check out the five technology projects we've selected for this issue as "winners" in our annual roundup of win-

ners and losers. You'll find a heavy dose of remarkable semiconductor technologies and software used in brilliant ways.

**Innovative Silicon's new superdense memories promise to boost the performance of all kinds of microprocessors, and might even make them cheaper while doing so.** Vanu's cellular base-station project is pushing the boundaries of software-defined radio, allowing cellphone service providers to expand and enhance their services with a simple software upload, rather than a hugely expensive hardware overhaul. RapidMind, a plucky Canadian company, has put together a software development platform that helps programmers navigate the tricky task of fully unleashing the astounding power of IBM's radical new Cell microprocessor. And BT Group has seized the high ground in telecom by embarking on a bold plan to convert its telephony networks—every last meter of them—to one of the greatest triumphs of software innovation, the Internet Protocol.

Perhaps best of all, there's Magneti Marelli's new engine controller, which can automatically detect and accommodate any mixture of gasoline and ethanol—from 100 percent ethanol to none at all. It even lets the engine burn natural gas. It's all done in software, which varies the cylinders' fuel intake and timing. The controller is now fitted to a 1.4-liter engine in a version of the Siena sedan sold by Fiat in Brazil.

The Siena is so plain that it makes the Chevy Aveo look daring. Unlike the beastly Bugatti, the Siena can't go 408 kilometers per hour. And unlike the Terrafugia Transition concept vehicle, also described in this issue, it can't fly. What the Siena can do is provide affordable and economical transportation anywhere in the world, no matter what kind of fuel is available. If your definition of "innovation" goes beyond developments that immediately gratify investors, superwealthy car buffs, or gadget lovers, you've got to love the homely Siena. Beauty really isn't skin-deep. ■

## MASTERS OF Memory

Swiss firm crams 5 megabytes of RAM  
into the space of one



**T**he running joke at **Innovative Silicon**, a four-year-old Swiss firm on the shores of Lake Geneva, is that state-of-the-art microprocessor chips shouldn't be thought of as logic chips with embedded memory. Rather, they should be called memory chips with embedded logic.

Yes, the joke is so geeky it hurts. But it conveys more than a little truth—and that's pretty painful, too. On-chip memory already takes

up more than 50 percent of the surface area of any respectable microprocessor. It's expected to occupy a whopping 83 percent of the area of high-end processors made in 2008, and 90 percent by 2011. And that's no joke for designers who will be hard-pressed to cram in hundreds of megabytes of memory without making their chips any bigger.

That's why Innovative Silicon, which is essentially an intellectual property firm, may have the

last laugh. The company, which is in Lausanne, has developed what it says is the densest—and cheapest—embedded memory technology in the world. It's called **Z-RAM, for zero-capacitor dynamic random access memory**, and if it grabs even a little piece of the on-chip memory market, it will change the ground rules for microprocessor design and quickly make Innovative Silicon a company to be reckoned with.

The semiconductor industry is absolutely awash with innovative memory technologies that were supposed to conquer the world but didn't. What's so great about Z-RAM? Basically, it does not require any new materials or extra processing steps in the fabrication process, says Pierre Fazan, one of the founders of Innovative Silicon and its chairman and chief technology officer.

That's extremely important to chip makers, who are reluctant to add any new



From left, Pierre Fazan, cofounder; Mark-Eric Jones, CEO; and Serguei Okhonin, cofounder of Innovative Silicon

materials to their already complex and delicate processes, for fear of how the additions may erode the proportion of working chips that emerge from their fabrication runs. Extra processing steps, too, are anathema, because they increase costs, often greatly.

That's the beauty of Z-RAM: no exotic semiconductors, no oddly structured parts, and no experimental insulators. Each memory cell is just a single transistor. That's it. For comparison, conventional on-chip memories typically use six transistors per memory cell. So you can fit as much as 5 megabytes of Z-RAM into the space occupied by a single megabyte of conventional embedded memory. That lets you greatly increase the amount of memory on the chip and thereby improve its performance, make the chip a lot smaller and cheaper, or do a good deal of both. Even engineers outside of Innovative Silicon acknowledge the idea's appeal. "It is an elegant technology," says Subramanian S. Iyer, an embedded-memory specialist at IBM.

"The transistor is the most studied device in the world," says Serguei Okhonin, Innovative Silicon's other founder and its chief scientist. "To make it work as a memory, we had to find something different." What they found was a way to temporarily store a bit as charge inside the body of a transistor made on a silicon-on-insulator (SOI) semiconductor wafer. Such wafers are gaining ground as the substrate for high-performance processors, such as the Cell microprocessor and Advanced Micro Devices' Opteron.

In 2000, Okhonin and Fazan, then at the École Polytechnique Fédérale de Lausanne, were researchers in one of just two groups to successfully create this type of memory, termed a floating-body cell. But since then, several large chip makers—notably Toshiba Corp. and Renesas Technology Corp., both in Tokyo—have followed suit.

Nevertheless, Okhonin and Fazan pulled way ahead of the pack last year by bagging one of the biggest fish in the microprocessor pond: Advanced Micro Devices (AMD), in Sunnyvale, Calif. The US \$5.8-billion-per-year company contracted to purchase a license for possible

integration of Z-RAM into future processors. "AMD is considering this technology for our highest-end microprocessors," Craig Sander, AMD's vice president of process technology and development, said in an interview. "However, we do believe if successful at that, we could see this migrate down to lower-cost versions as well."

Meanwhile, with Toshiba nipping at its heels with a similar type of fast, compact embedded memory, Innovative Silicon isn't standing still. Just weeks ago, it unveiled a radical new version of Z-RAM that will send the competition back to the laboratory.

**Z-RAM IS A** technology that straddles two great industry imperatives: ever more on-chip memory and transistors that operate faster and consume less power.

Increasingly, for makers of microprocessors (with the notable exception of Intel), gains in speed and efficiency are coming from the use of SOI wafers, which serve as the foundation on which a device is fabricated. Although SOI wafers are still a small fraction of semiconductor wafers sold, their shipments doubled in 2006 to 22 000 square meters and are expected to reach 39 000 square meters in 2008.

An SOI wafer differs from an ordinary silicon wafer in that it has a very thin layer of insulating silicon dioxide buried a few hundred nanometers or less below the surface. That layer of insulation cuts the transistor off from the vast bulk of the wafer—which, in turn, limits the amount of charge the transistor must move in order to switch on or off.

The result is to speed up circuits by as much as 30 percent, Fazan says. As transistors shrink, they increasingly leak current, even when they are turned off. But the insulation in SOI wafers blocks a major pathway for that current, thus reducing the power that transistors draw by 30 percent when they're switching and 50 percent to 90 percent when they're not.

But those advantages come at a cost. A 200-millimeter SOI wafer sells for about \$275, while a plain silicon wafer of the same size goes for \$65. Mark-Eric Jones, Innovative Silicon's president and CEO, points out that as a percentage of total cost, the price gap collapses considerably once you factor in the \$2400 in processing needed to turn a blank wafer of either substrate into a repeating mosaic of microchips. Still, all things being equal, the SOI chip will cost almost 9 percent more than the bulk-silicon chip.

But all things are no longer equal. The SOI wafer lets you substitute Z-RAM for the chip's conventional embedded memory. SOI's insulating layer is key to storing the bit in Z-RAM, so you

cannot build it on a plain wafer. By Innovative Silicon's estimates, if the conventional memory takes up half the area, replacing it with Z-RAM would let designers shrink a chip to 72 square millimeters from 120 mm<sup>2</sup>. That would boost the number of chips per wafer and cut the final cost of the chip almost in half. Suddenly, SOI looks like a bargain.

**WHETHER THEY'RE USING** SOI or not, microprocessor makers are compelled to continue boosting the amount of on-chip memory in their designs for the simple reason that they can't get the performance they need any other way. The other means of increasing processing rates—running clocks at higher speeds and putting more processor cores on a chip—are effective only if those processor cores have rapid access to data.

The standard form of embedded memory used on microprocessor chips is static random access memory. Designers incorporate the SRAM as blocks of memory called caches. The level 1 cache, or L1, is optimized for speed and located near the processor core on the chip. It stores the most frequently needed few kilobytes of data, so when the processor needs data, it looks there first. Then it checks in a larger but more distant and somewhat slower cache, called L2, which is usually about 16 MB these days but could balloon to 200 MB for next-generation chips.

If the data it requires are in neither of those caches, on some processors there is a still larger cache, L3, in which to look. Failing that, it's off to the computer's main memory, which consists of hundreds of megabytes of dynamic random access memory (DRAM), or as a last resort, the hard drive.

But going off chip is costly, both in the amount of time the processor spends spinning its wheels waiting for data and in the power expended exporting data across the computer's wiring. The performance boost you can get out of packing more memory into the CPU depends on the type of processor, but AMD's Sander estimates that adding memory gives the kind of performance boost that you'd get from a year's worth of improvements to the chip's transistors.

**Z-RAM DOESN'T JUST** have to be better than traditional embedded SRAM; it also has to be better than DRAM, a memory technology that has been slowly winning a place on logic chips. Though slower than SRAM, DRAM consumes about one-fifth as much power and is about four times as dense. DRAM is so much more dense than SRAM because it consists of a single transistor and a capacitor instead of SRAM's six transistors.

But that capacitor is the problem. Moore's Law doesn't apply to it, so it stays big while the transistors all around it continue their mad descent into the infinitesimal. The capacitor can't shrink because it must stay large enough

to store a detectable amount of charge.

The growing mismatch between the size of transistors and the size of capacitors has led to strange-looking arrangements, such as capacitors built as narrow trenches having a depth many times greater than the chips' transistors. Another configuration has relatively enormous fin-shaped capacitors built above the silicon in the area that usually holds the chip's wiring. Both arrangements are too expensive to put into many logic chips, requiring several extra manufacturing steps. Nevertheless, DRAM is a well-understood technology, and it is embedded in some memory-intensive chips such as IBM's BlueGene processor.

Add the speed of SRAM to DRAM and remove the capacitor and you get Z-RAM. Fazan and Okhonin removed the capacitor in DRAM by turning a bug into a feature. The bug is typically called the "floating-body effect." The term comes from the fact that the insulation layer in an SOI wafer electrically separates the body of the transistors from the rest of the silicon, letting its voltage vary, or "float." The result can be a slight difference in the way transistors built in SOI operate from those fabricated in bulk silicon. In particular, it can lead to the transistor's passing more current or less current for a given voltage signal, depending on how much current flowed in the recent past.

When a transistor is "on," electric current runs from the transistor's source to its drain. By the time those accelerating electrons get to the boundary of the drain, they are moving so quickly that some will whack into silicon atoms energetically enough to ionize them. This impact ionization, as it's called, generates pairs of electrons and holes. The electrons exit the transistor through the drain, which is connected to a positive voltage. But the holes are repelled by the drain. In a bulk silicon crystal, this extra positive charge would harmlessly drift out into the silicon, but in SOI, the insulating layer traps it in the transistor, forming a body of charge that floats above the insulator.

That floating charge, a minor annoyance to advanced processor designers, could be a wonderful thing, Fazan and Okhonin realized. You can use it to store data—in fact, a transistor with such a floating charge is basically a Z-RAM cell storing a 1 [see illustration, "Z-RAM"]. To erase the 1 and store a 0, increase the voltage on the transistor's gate. That pushes the holes out of the transistor through the source electrode and even leaves a slight negative charge behind.

Reading a bit from a Z-RAM cell is simple. All you need to do is turn the transistor on and measure the amount of current flowing through it. For a field-effect transistor like the ones you'd find in a typical processor, turning the device on involves applying a voltage to the transistor's gate. The voltage forces open a conductive channel between the source and the drain, allowing current to flow. More current will flow through a cell with a 1 than through one with a 0, because

## Z-RAM



**GOAL:** Make the densest, cheapest embedded memory in the world.



**WHY IT'S A WINNER:** Its technology could quintuple the amount of memory incorporated into microprocessor chips and make them both cheaper and faster.



**PLAYER:** Innovative Silicon.



**WHERE:** Lausanne, Switzerland.



**STAFF:** 32 (from 16 countries) as of September 2006.



**BUDGET:** US \$22 million in venture capital.



**MORE:** <http://www.innovativesilicon.com>

the floating body charge that makes up the bit exerts its own force on the channel and acts almost like a second gate, amplifying the effect of the real gate. If you compare the current of the memory cell to that of a reference cell, you can tell a 1 from a 0.

In Fazan and Okhonin's first devices, the difference between a 1 and 0 was just 3 microamps, or 15 microamps per micrometer of channel width. During the next few years, Innovative Silicon improved on that difference only incrementally until the company began a project aimed at a different goal.

The new version of Z-RAM, called Gen2, which the company unveiled just weeks ago, started out as an attempt to make a low-power Z-RAM suitable for cellphone chips. But Okhonin's group "came back with something not just dramatically lower in power but significantly faster and denser," says Jones, the Innovative Silicon CEO. Okhonin had found a way to store vastly more charge in a smaller transistor.

During my September visit to Lausanne, as Fazan, Jones, and I crowded into Okhonin's laboratory, Okhonin and an engineer set up an oscilloscope and a test program to show the difference between the original Z-RAM—Gen1—and Gen2. For the Gen1 test, I saw what I expected, a small rounded bump of current when the device was storing a 0 and a slightly larger bump for a 1.

Then, using the same transistor but a different set of input voltages, he demonstrated Gen2.

"That's zero," Okhonin said, pointing, it seemed, to nothing at all. "That's one." The 1 signal was a sharp rectangle that to my eye was at least 10 times the size of either Gen1 signal. "It's basically digital," he said, unable to stifle a huge grin.

Because the difference between the two states, the margin, is so enormous, there's no need for the reference cells Gen1 required to interpret 1s and 0s. The much larger margin also means a Gen2 cell can be read about twice as fast as a Gen1 cell.

What's more, Gen2 uses considerably less power—75 percent less to read and 90 percent less to write. Part of the power savings comes because Gen2 cells hold their data longer. Z-RAM is similar to DRAM in that over a period of time the cell will inevitably lose its data. That is, 1 and 0 will become indistinguishable. For that reason, an array of Z-RAM must periodically be refreshed by reading and rewriting the data, a subtle but persistent drain on power. But Gen2 requires fewer refreshes, because it takes about 8 seconds to lose its bit, and that's 10 times as long as Gen1 or DRAM takes.

The particulars of the technology also let designers tweak it toward maximizing either speed or power efficiency. A Gen2 array can be made to access data at 400 megahertz or more—about 60 percent faster than a comparable embedded DRAM array but still not as speedy as the fastest SRAM. Or it can be designed to draw a remarkably low 10 microwatts per megahertz—about 10 percent of comparable embedded DRAM.

**INNOVATIVE SILICON'S GEN2** Z-RAM comes amid a flurry of zero-capacitor DRAM developments, some of them the culmination of years of laboratory work. In 2000, Toshiba's Takashi Ohsawa had almost the same idea as Fazan and Okhonin. The Innovative Silicon team reported its memory first, but Ohsawa won the race to the patent office. Still, Innovative Silicon does not believe its patents infringe on Toshiba's.

Of all the floating-body memories in development, Toshiba's most closely matches Innovative

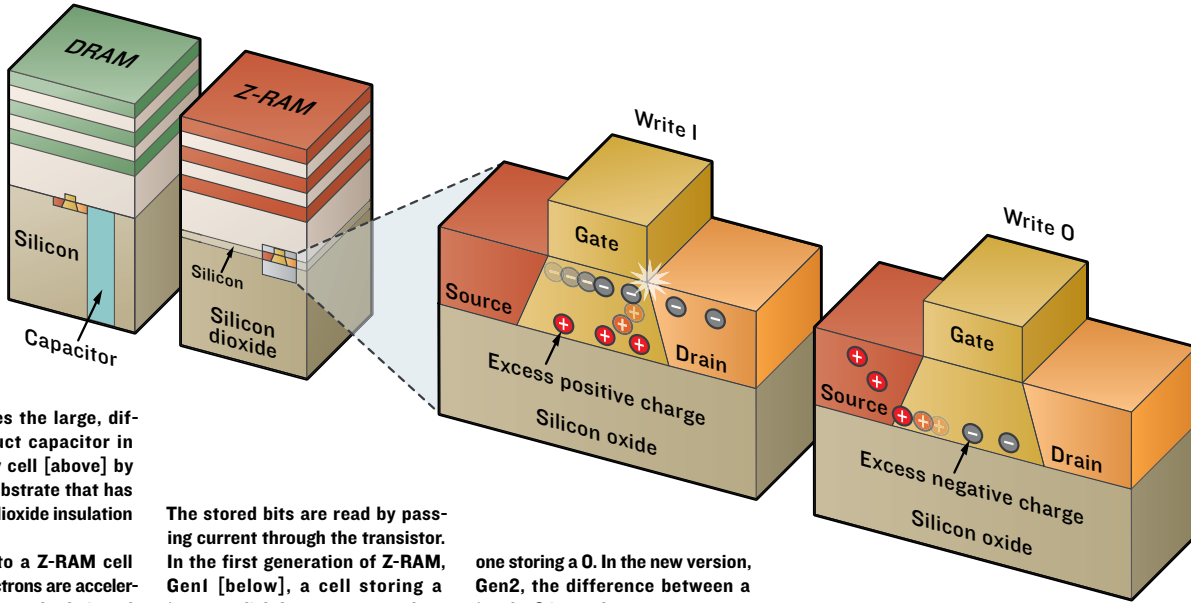
### What the Experts Say



**T.J. RODGERS:** Unfortunately, as the most advanced team at Toshiba published papers on bringing the Z-RAM to production, it became clear that storing a tiny amount of charge in a leaky environment with multiple disturb mechanisms would compromise the design so as to negate many of the hypothetical advantages of the technology.

**GORDON BELL:** This won't be without pain...but it will be a Big Deal if it can happen.



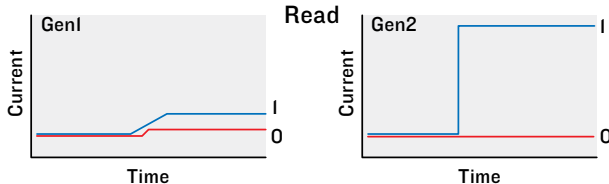


Z-RAM eliminates the large, difficult-to-construct capacitor in a DRAM memory cell [above] by using a silicon substrate that has a layer of silicon dioxide insulation buried within it.

To write a 1 to a Z-RAM cell [above, right], electrons are accelerated from the source to the drain and crash into silicon atoms, creating pairs of electrons and positive holes. The holes accumulate in the transistor body and act as the stored bit. To write a 0 to the cell, the holes are drawn out through the source electrode, leaving excess negative charge.

The stored bits are read by passing current through the transistor. In the first generation of Z-RAM, Gen1 [below], a cell storing a 1 passes slightly more current than

one storing a 0. In the new version, Gen2, the difference between a 1 and a 0 is much greater.



TOPDOG ILLUSTRATION

16-megabit memory arrays built with the most advanced manufacturing process in use today, a 65-nanometer process, and using a next-generation process, 45 nm.

**AS TRANSISTORS GET SMALLER**, the economics tilt more and more in Z-RAM's favor, because more sectors of the chip industry will turn to SOI to solve their power and performance problems. "Graphics chips have all the same problems that microprocessors have," Jones notes. We may soon see the impact of Z-RAM in that sector. In fact, Jones's primary customer, AMD, recently purchased graphics-chip maker ATI Technologies, in Markham, Ont.

Cellphone chip makers, the original target for the work that led to Gen2, are also exploring SOI, because the insulator prevents power-sapping swirls of current from forming in radio chips. Cellphone chip makers, who are highly cost conscious, would look to offset the costs of a switch to SOI with something like Z-RAM, Jones argues.

One day, even commodity stand-alone DRAM chips might be made with Z-RAM. Initially, Jones did not expect DRAM makers to be interested until Z-RAM was well established as an embedded memory. "But—without giving away any details—perhaps [Z-RAM's adoption] is likely to be a lot earlier than we'd thought," he says.

—Samuel K. Moore

Silicon's Gen1 product, though it uses a proprietary method of writing data. Ohsawa told *IEEE Spectrum* that Toshiba plans to license its technology to chip developers.

Meanwhile, Renesas Technology is pushing a two-transistor cell. Sreedhar Natarajan, president and CEO of Ottawa-based Emerging Memory Technologies, looked at Innovative Silicon's Gen1 but decided to license Renesas's two-transistor approach. "The only disadvantage I saw [with the Renesas design] was the area penalty," he says.

Fujitsu has developed a single-transistor memory that uses a write mechanism different from Z-RAM's. And some university groups have built their own versions.

Even Intel, whose disdain for SOI is no secret, planned to present a floating-body cell of its own last month at the IEEE International Electron Devices Meeting. According to Michael C. Mayberry, Intel's director of components research, if the company needs such a memory, his group will look for a way to build it on ordinary bulk silicon.

Indeed, at the same meeting, Geneva-based

STMicroelectronics was expected to report a cell built on bulk silicon. It works by building a special layer beneath the transistors to substitute for the insulator of an SOI; its main drawback is that the difference between a 0 and a 1 is quite small.

"We feel Gen1 is better than those, but it's comparable," Jones says. "But Gen2 is such a radical step forward, particularly in power consumption, that it puts it in a unique league."

AMD's Sander is one of the few outside of Innovative Silicon who have had a close look at Gen2, and he agrees. "Gen2 looks significantly better than anything else that I'm aware of—higher density, lower power, better performance as well," he says. Sander notes that Gen2 should scale to smaller dimensions even more easily than Gen1.

Okhonin says that as the Z-RAM transistor shrinks, it stores less charge, but the effect of that charge on the current through the transistor only grows stronger. And because it's the current, not the charge, that's read as a bit, Z-RAM should work even with much smaller transistors, he adds. Innovative Silicon has already produced

(#13260) Adapted with permission from the January 2007 issue of *IEEE Spectrum*. © 2007 IEEE, Inc. For more information about reprints from *IEEE Spectrum*, contact PARS International Corp. at 212-221-9595.



www.z-ram.com



World's Lowest Cost Memory Technology