Transistors Go Vertical

The semiconductor industry fights silicon sprawl by building up, not out

NEWS Through all the decades of microchip evolution one thing has remained constant: the silicon transistors of which they're made are basically flat. But that is very likely to change in the next five years. The semiconductor industry is facing a problem that can be solved only with a fundamental transistor redesign. Transistors are no longer the clean on-off switches they once were; instead, current leaks through them even when they are supposed to be off. As transistors shrink with each new generation of microchip, this errant current increases, draining batteries and heating chips up.

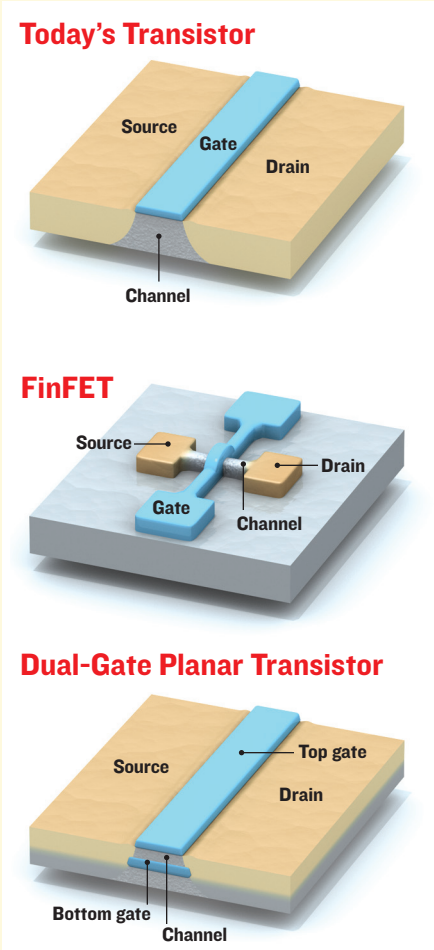
Major chip manufacturers will show off their latest proposed solutions from 10 to 12 December at IEEE's International Electron Devices Meeting, in Washington, D.C. What many of these experimental transistors have in common is that they are decidedly not flat.

Going from flat to three-dimensional in the conservative microchip industry is a radical shift, but as Leo Mathew, a research scientist at Freescale Semiconductor, says, "the payoff will be substantial."

The semiconductor industry has fed the consumer appetite for better electronics performance by shrinking the transistors' structure to cram more of them onto a chip. (However, some of the materials involved are changing for the first time in 40 years; see "The High-*k* Solution," *IEEE Spectrum*, October.) Normally, you can picture a transistor in four parts, the source and drain, connected by a channel and topped off by a gate. Most of the transistor is in one plane, built into the silicon substrate of the microchip. Only the gate and its extremely thin insulating layer, which lie directly above the channel, protrude slightly above the flat plane of silicon. Voltage on the gate causes a conductive path to form in the channel, allowing current to flow between the source and the drain.

However, shrinking this structure further means that removing the voltage on the gate no longer completely stops the flow of electrons. Even in today's transistors the source and drain are separated by mere tens of nanometers, a short enough distance for electrons to leak through the lower part of the channel, farthest from the gate. The result is wasted power and heat. It's one reason the battery power on an unplugged laptop seems to evaporate quickly and why companies have to spend huge sums to cool their server rooms.

Realizing that source-drain leakage will only get worse as chips shrink, researchers have sought to plug the leak by raising the channel, source, and drain out of the substrate. The gate is then draped over the channel on three sides like a lowercase "n." Now the current is constrained only to the raised channel, and electrons no longer have a path through which to leak. This general class of transistor is called multigate, because the wrapped gate is like having three gates instead of one. But "the shift to multigate transistors requires a fundamental change in transistor design," says John Pellerin, director of logic technology development at Advanced Micro Devices. That's because to make multiple gates, nearly everyone agrees you have to go vertical. According to Infineon Technologies' principal scientist for CMOS devices Klaus Schrüfer, multigate technology is a game changer: "Multigate is the only device architecture for scaling" into the foreseeable future, he says.



PRESENT, FUTURE, AND FUTURE: Today's CMOS transistor is planar, but chip makers are exploring more power-efficient three-dimensional structures (FinFET) as well as a planar structure with two gates.

The most common multigate transistor design is a structure called a FinFET. In a FinFET, the channel connecting the source and drain is a thin, finlike wall jutting out of the silicon substrate. The drawback to this design is the difficulty of etching it out of the silicon in the minuscule detail needed. Several researchers say that because the design changes are so substantial, the transition from planar to multigate will be more difficult than the transitions between any other past technology nodes.

The FinFET's co-inventor, University of California, Berkeley, professor Tsu-Jae King Liu, says that the early adopters of multigate technology will likely be DRAM manufacturers; logic device makers may follow eventually. There are two reasons for this: first, DRAM has much more stringent leakage current requirements than logic circuits do. In DRAM, a bit of information is stored as a puddle of electrons in a capacitor, with a transistor acting as the gatekeeper. If those electrons dribble out across the transistor, what's lost is not just heat or power—it's information. Another reason DRAM makers will likely move to multigate first is that they already have expertise etching the steep features needed, King Liu says. Because DRAM capacitors are already constructed in deep narrow trenches, the FinFET's form doesn't intimidate memory-chip makers.

In fact, the world's largest DRAM maker, Samsung Electronics Co., in Seoul, South Korea, has indicated that it may use a transistor structure like the FinFET for DRAM. But it will take longer for logic manufacturers to get on board, STMicroelectronics' Thomas Skotnicki explains. "Logic people are very much planar," he says. "For us, there are many barriers—including psychological ones."

Still, even the most conservative major chip makers, the foundries, are working on multigate transistors. Taiwan Semiconductor Manufacturing Co. (TSMC) hired FinFET co-inventor Chenming Calvin Hu to develop TSMC's proprietary FinFET. "A big foundry like TSMC generally tends to be more conservative," King Liu says. "They wouldn't go to a multigate transistor unless their major customers asked for it."

At the December conference, researchers from Freescale, IBM, Infineon, Intel, and others will feature their multigate devices. According to Intel director of

BRYAN CHRISTIE DESIGN

components research Michael C. Mayberry, Intel's transistor uses the archetypal FinFET structure. Freescale's new transistor bears some resemblance to the FinFET, says the company's Mathew, who invented it, but the finlike channel is shaped more like an inverted T.

If the shape of the transistor varies from company to company, the plans on when to introduce it do not seem to. "The earliest I see multigate transistors," says IBM Research Division's Wilfried Haensch, "is at 22 nanometers," which is planned for 2011. Infineon and Intel seem to be in agreement.

But does multigate have to be synonymous with 3-D? STMicroelectronics thinks not. The Franco-Italian chip maker has vowed to stay planar even beyond the 22-nm node, says Skotnicki, with a dual-gate device.

Imagine the ST transistor as a garden-variety FinFET but lying on its side. From the top you'd see a gate that looks like the one in today's transistor, but there is a second "shadow" gate buried beneath it, sandwiching the channel. The problem was that to make a double-gate transistor you needed to align the two gates with absolute precision. Perfectly aligning tiny gates is even harder than etching tiny fins. But Skotnicki says that ST has figured out how to do it, and the company showed a chip with the planar self-aligned double-gate structures in June: "It has the same electrical advantages as FinFET, but with probably the highest performance ever published," he says. "I don't know if FinFET can deliver that." He predicts that companies will abandon their multigate research and return to planar. And indeed, other firms such as IBM are working on planar dual-gate transistors alongside their 3-D multigate development. "We expect this to be a turning point," Skotnicki says.

Whether Skotnicki is right or not, multigate transistors are practically right around the corner. Usually companies have their technology figured out and in the pipeline a good two years before full production starts. By the time they take multigate transistors to market, they may have to dream up yet another design. At technology nodes beyond 22 nm, after 2013 or so, says Berkeley's King Liu, the FinFET might not deliver "any better performance" than a shrunk-down version of today's flat transistor. **—SARAH ADEE**

TRANSFER OF POWER: As this train in Oslo slows to a stop, its motors feed power back into the grid for other trains to use.

Oslo Metro Takes Greener Track

New trains share power captured by regenerative braking

Oslo has one of the world's smallest carbon footprints for a city of its size, but it wants to get even greener. To that end, it's replacing 63 of the T-bane Metro's trains with new three-car trains from Siemens that are 30 percent more energy efficient than the best cars currently in service there. The key is in the trains' ability to generate electricity while braking and transfer that power to other trains.

When a train's operator applies the brakes, the four 140-kilowatt, 750-volt dc electric motors are engaged as generators that use the kinetic energy of the turning wheels to send current back into the metro's power grid. This technique, called regenerative braking, allows the trains to recover up to 44 percent of the energy used to bring the trains up to speed.

Slowing vehicles down by transforming their inertia into electric current is by no means a new idea. Hybrid-electric cars use regenerative braking to charge onboard battery packs and help boost their fuel economy. The challenge with train systems is that the energy generation occurring in one train must be timed to coincide with a demand for power from a nearby train that is accelerating. The more these stops and starts can be paired, the less electricity the operating authority has to draw from the grid.

Regenerative braking is not to be confused with dynamic braking, employed in many diesel-electric trains to limit wear on the mechanical brakes. In dynamic braking, the current generated by a train's motors during deceleration goes to a set of large onboard resistors. They release the energy as waste heat or use it to warm the passenger compartments.

The environmental benefits from the new trains do not stop with the regenerative braking system. The 94-metric-ton, 54-meter-long MX3000 trains are made mainly of aluminum, so they are lightweight, and therefore require less energy than the average steel-bodied train to accelerate from a dead stop. What's more, 85 percent of the materials used to build each train are recyclable. Much of the rest can be burned at thermal energy plants.

Because of the MX3000's higher efficiency, plus the fact that most of Oslo's electricity is generated by hydroelectric plants, as little as 2.6 grams of carbon dioxide will be added to the atmosphere per kilometer traveled and per metric ton of vehicle weight, Siemens estimates. In other cities, the average electric train or tram contributes upward of 25 grams per kilometer traveled.

Two prototypes delivered to Oslo in 2005 for testing lived up to energy-efficiency expectations. The city has so far received a quarter of its 63-train order. By 2009 all of Oslo's metro system will rely on regenerative braking. **—WILLIE D. JONES**

SIEMENS