

# Is Silicon's Reign Nearing Its End?

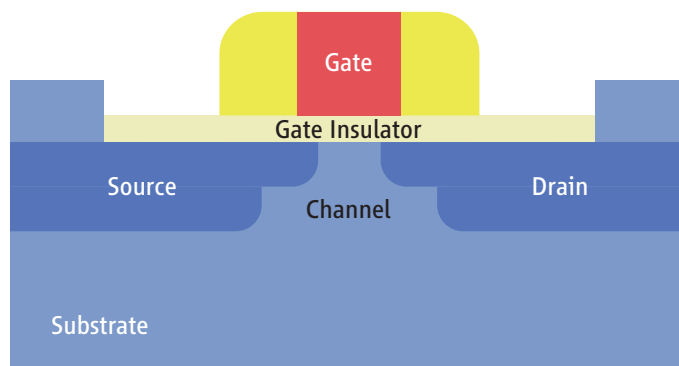
Silicon is almost synonymous with computer chips. But as the semiconductor struggles at the minute scales of today's devices, chipmakers are being forced to consider other materials

If you want to see real nanotechnology in action, check out Intel's Penryn computer chip. It contains some 820 million transistors, each with features just a few tens of nanometers across. These transistors are so small that more than 2 million can fit on the period at the end of this sentence. A device inside each one flips an electrical switch on and off as many as 300 billion times a second. In the time it takes for one such flip, light travels less than half a centimeter.

Amazing—but not good enough. Late this year or early next, Intel plans to introduce a new line of chips that shrinks the components even smaller. For the past half-century, this perpetual contraction has been at the heart of the industry's favorite trend: Moore's Law, which holds that the number of transistors on chips will double about every 2 years. For most of that history, this downsizing, known as scaling, came about as engineers developed ever sharper chip-patterning techniques. But the past few years have brought a quiet revolution to chipmaking. Because conventional materials have started misbehaving at such small scales, engineers have had to call in reinforcements. Since the 1990s, chipmakers have gone from making their devices out of about 15 chemical elements to making them out of more than 50, in hope that the

new additions will help them keep shrinking the devices.

Few doubt that Moore's Law will hold for a couple more generations of chips, and perhaps even longer. But continuing this trend will not be a straightforward enterprise. Researchers are looking at redesigning the way they make transistors, incorporating new insulators, and even replacing silicon as the semiconductor through which electrical charges flow in their circuitry. "There are very serious challenges ahead," says electrical engineer Jesús del Alamo of the Massachusetts Institute of Technology (MIT) in Cambridge. Adds electrical engineer Mark Rodwell of the University of California, Santa Barbara, "Regardless of which [technology] wins, there's a real sense that there aren't too many years left to play this game."



**Logical.** Conventional transistors, known as MOSFETs, keep getting smaller and better. Today, more than 30 million can fit on the head of a pin.

## Matchmaker, matchmaker

At the center of this game are transistors, or more precisely, metal-oxide-semiconductor field-effect transistors (MOSFETs). They work by sending a pulse of electricity to a central electrode called the gate (see figure, below). The charge on this gate causes a spike in the electrical conductivity of a semiconducting channel that sits below it. This opens an electrical doorway in the channel, allowing an electric current to flow between two other electrodes—known as the source and the drain—that sit at opposite ends of the channel. The high electrical output at the drain is a "1" in the on/off digital world. If you turn off the charge on the gate, the channel door closes and the current flow to the drain shuts off, giving you a digital "0."

Silicon MOSFETs have worked wonders in recent decades, thanks in large part to a unique marriage in the material world. In a MOSFET, the gate electrode and the conducting channel beneath it must be insulated from each other, so that when the gate is closed, any excess charges in the gate aren't able to open the channel and let current flow between the source and the drain. In traditional logic chips that are the brains of desktop computers and servers, that insulating layer has been provided by silicon dioxide (SiO<sub>2</sub>).

Silicon and silicon dioxide's

storybook marriage has thrived for decades because where they meet, they form a near-perfect union. At these interfaces, each silicon atom binds readily to four oxygen atoms. That's good, because it ties up unfilled bonds at the silicon channel's edge that can trap electrical charges in place and disrupt the transistor's ability to switch when prodded. Whenever defects occur, engineers can easily neutralize them by piping in hydrogen to latch onto any free bonds. The result has been that although other semiconductors are faster and stronger, the near-perfect union between silicon and silicon dioxide has made it possible to continually improve transistors.

Recently, however, the stress of relentless scaling has proved too great for SiO<sub>2</sub>, which is only a moderately effective insulator. As transistors continued to shrink, engineers were forced to make their SiO<sub>2</sub> layer as thin as 1 nanometer, or about three atomic layers thick. That was so thin that it started to leak. "We were just running out of atoms," says Suman Datta, an electrical engineer and Intel veteran, who now works at Pennsylvania State University, University Park.

In the mid- to late 1990s, researchers at Intel, IBM, and elsewhere realized the days for silicon dioxide were just about up and that they needed to ditch silicon's devoted partner for a new trophy wife. Their goal was to find a material with a higher insulating value, known as its dielectric constant and denoted by the Greek letter kappa ( $\kappa$ ). Researchers tested dozens of "high-k" alternatives, finally settling on a new bride called hafnium dioxide (HfO<sub>2</sub>).

"It started off innocently, saying we'll just swap out the SiO<sub>2</sub> with a high dielectric constant [material] and be on our way," says Supratik Guha, a materials scientist and senior manager at IBM in Yorktown Heights, New York. Unfortunately, the change wasn't so simple. Among many other problems, the new insulator didn't form a clean interface with the gate above it, which was made from polycrystalline silicon. So engineers were forced to replace the polysilicon gate with titanium-based alloys that performed better. That change raised new problems, because the metal alloys couldn't handle the high temperatures used in manufacturing some of the other components. Ultimately, Intel worked out a manufacturing strategy, and last year the company began shipping chips made with the new metal gates and HfO<sub>2</sub> insulator. At the

time, Gordon Moore, co-founder of Intel and eponymous drafter of the law, called the switch "the biggest change in transistor technology since ... the late 1960s."

What is sobering is that this relatively straightforward change took the industry a good 10 years to accomplish, and the challenges ahead appear to be far greater. "The complexity of the problems is only going to grow," Datta says. For starters, as

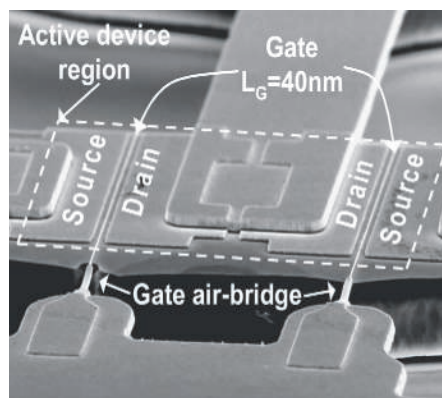
Intel and other chip companies prepare to contract from the 45 nanometer scale to 32 nanometers (the numbers refer to half the distance between adjacent lines of memory cells), it appears that silicon dioxide is set to cause a new round of trouble. Even though SiO<sub>2</sub> was supposedly

replaced as the gate insulator, a little bit has remained behind because it is nearly impossible to eliminate. When silicon is placed next to hafnium dioxide, Datta explains, some oxygen atoms at the interface invariably break their bonds to hafnium and hook up with silicon. This isn't all bad, he says, because the clean interface with that ultrathin layer of SiO<sub>2</sub> seems to improve the conduction of electric charges through the semiconductor channel. But with all the device components set to shrink again, the

## Online sciencemag.org

Podcast  
interview  
with author  
Robert F. Service

Ti/Au Source	Ti/Au Gate	Ti/Au Drain
6nm p-doped		low resistance cap
10nm p-doped		Al <sub>x</sub> In <sub>1-x</sub> Sb top barrier
3nm undoped		Al <sub>x</sub> In <sub>1-x</sub> Sb top barrier
Be $\delta$ -doping		
7nm undoped		Al <sub>x</sub> In <sub>1-x</sub> Sb top spacer
5nm InSb quantum well		
3 $\mu$ m undoped Al <sub>x</sub> In <sub>1-x</sub> Sb bottom barrier		
200nm Al <sub>y</sub> In <sub>1-y</sub> Sb interfacial layer		
Semi-insulating GaAs substrate		



**Speed boost.** Novel high-speed transistors (*above*) from Intel push positive charges through a thin layer of InSb (*top, red*).

SiO<sub>2</sub> layer must also get thinner—and that could disrupt the flow of charges in the silicon. "It's a very tough problem," Datta says.

At the International Electron Devices Meeting (IEDM) in San Francisco in December 2008, Intel researchers reported that they had solved this problem and many others. The company announced that it had completed the development phase of its 32-nanometer-manufacturing process and would begin turning out the new chips by the end of 2009. Intel researchers haven't revealed their latest tricks, but Datta calls the performance data they have presented "excellent." Datta says he expects Intel to stick with HfO<sub>2</sub> for now, because it was so expensive to make the switch in the first place.

### A new look

So if the next step in shrinking silicon electronics is on track, what's next? After 32 nanometers, the next step down is 22 nanometers. "This is where things get very interesting," Datta says. At this dimension, it's likely that HfO<sub>2</sub> will also begin to fail to contain current within its walls. If that happens, one option is to change to a material with an even higher dielectric constant, says Darrell Schlom, a high-k expert at Cornell University. Schlom's group has tested more than a dozen. Among the best, he says, is crystalline lanthanum lutetium oxide, which has a dielectric constant of 40, more than 10 times that of SiO<sub>2</sub> and nearly double that of HfO<sub>2</sub>. A big challenge, however, is that the more elements make up the insulating material, the harder it is to keep the perfect order of the material at the interfaces.

Another option at this scale is to redesign the architecture of transistors altogether. One alternative, considered most likely for 22-nanometer scale devices, is to move away from layered, sandwichlike devices and stand the silicon channel on its side. Such devices, known as FinFETs (because the vertical silicon channel looks a bit like a fish fin), in principle would allow engineers to surround the channel on three sides with dielectrics and gate materials. Then using several gates in concert to trigger the flow of current in the channel could make it easier for engineers to control when the devices flip on and off and how much current they put out when they do.

Academic researchers have been turning out FinFETs and other exotic-shaped devices for years. Now, however, even the big chip players are looking at exotic designs. At IEDM, for example, a consortium of researchers from Toshiba, IBM, and Advanced Micro Devices reported making



the world's smallest FinFET transistors with high-k dielectrics and metal gates. The new transistors were only half the size of previous FinFET record holders, and studies revealed that their geometry made them more reliable than planar versions, according to company press materials. But, at least for now, the new devices suffer from poor electrical contacts between the source and drain electrodes, Datta says.

the silicon wafer beneath it is a problem, equally difficult is the interface to a high-k dielectric above it. And so far, efforts to neutralize defects have proven underwhelming. Unlike silicon and SiO<sub>2</sub>'s perfect marriage, "there's no wonderful interface with the III-Vs," Schlom says.

Another challenge is making transistors that conduct positive charges, called holes. That's important for chip designers, as it

als to sit next to each other with relatively few problems. Researchers elsewhere have also had success with InGaAs. In one example, researchers led by Peide "Peter" Ye of Purdue University in West Lafayette, Indiana, reported in the April 2008 issue of *IEEE Electron Device Letters* that they had made high-speed InGaAs n-type transistors that had a large output current when switched on, an important hurdle for the field.

Still, Ye and others acknowledge that III-Vs aren't there yet. Ye's InGaAs devices, for example, work splendidly when switched on, says materials scientist Paul McIntyre of Stanford University in Palo Alto, California; but try to switch one off and current still leaks through, like water overtopping a levee. In addition, neither Ye's group nor any other has had much luck in making high-quality p-type transistors from InGaAs. At IEDM, Chau and colleagues at Intel did report novel InSb p-type transistors that are the fastest to date. But, for now, they too remain leaky when switched off.

There might be a workaround. Researchers at IBM and elsewhere have also had some success with making high-speed p-type transistors using germanium (Ge). And though Ge isn't itself a III-V compound, it seems to integrate fairly well with them. One hope, Ye and others say, is that researchers will be able to make integrated circuitry using InGaAs n-type transistors next to Ge or InSb p-type transistors. Even if they succeed, however, the circuits could be prohibitively expensive to manufacture, as they require patterning very different materials in the same layer of the chips. Still, despite these challenges, Ye remains optimistic. "I think the future is still bright because there is no other choice," he says.

That's not entirely true. If III-Vs don't pan out, or perhaps even if they do and last only a generation or two, there are plenty of far-out ideas for reinventing microelectronics. Among them: transistors made with single-layer carbon sheets called graphene, carbon nanotubes, or III-V nanowires. But for now, these upstarts still need plenty of work to have a shot at dethroning more-conventional approaches.

In any case, it's clear that this work needs to happen soon, or the steady progress of Moore's Law will begin to slow if not stop. "My own view is that we will not see another decade of scaling," McIntyre says. Perhaps not. But MIT's del Alamo and others point out that the \$260-billion-a-year chip industry has been jumping over roadblocks for years. "The potential payoff is gigantic," del Alamo says. Adds Datta: "One thing I've learned is, don't predict the end of Moore's Law."

—ROBERT F. SERVICE

Speed of Charges in Different Materials (cm<sup>2</sup>/V·s)

Charges	Si	GaAs	In <sub>0.53</sub> Ga <sub>0.47</sub> As	InAs	InSb
Electrons*	300	7000	10,000	15,000	30,000
Holes*	450	400	200	460	1250

\*Electron carrier mobilities measured in transistor channels with electron concentration of 1x10<sup>12</sup> cm<sup>-2</sup>. Hole mobilities in bulk.

**Racing form.** Engineers can speed up transistors by using higher-speed materials. But although negatively charged electrons move quickly in most III-V materials, positively charged "holes" typically do not.

### Beyond silicon

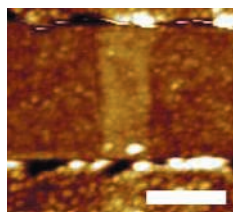
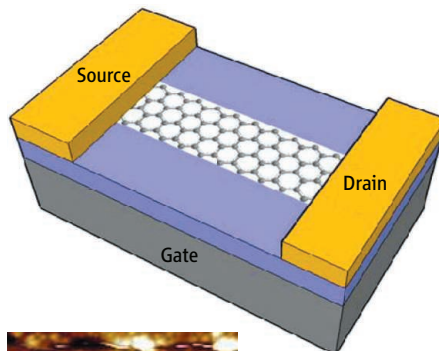
Even that switch, however, looks simple compared with the changes further ahead. After the 22-nanometer scale comes 15 nanometers, and it appears—at least for now—that improvements in transistor speed and performance will have to come from new materials rather than from scaling silicon further. "Now that you've gotten rid of the SiO<sub>2</sub>, why not have the high-mobility channel and high-k materials to put on top of it?" Schlom asks. In other words, why not replace the silicon with a better semiconductor?

Top candidates for this switch are semiconductor alloys known as III-Vs, due to the position of their elements in the periodic table. Examples include gallium arsenide, indium gallium arsenide (InGaAs), and indium antimonide (InSb), all of which can ferry electrical charges much faster than silicon can (see table, above).

But there are lots of problems in marrying III-Vs with modern logic-chip technology. For starters, it is not currently possible to make large-scale III-V wafers on which to construct the devices from the ground up. That means engineers need to integrate III-V materials on the silicon wafers that are the industry standard. That's no easy task, because the atomic spacing in III-Vs is significantly different from silicon. Try to grow InGaAs, for example, directly on silicon and you get an interface riddled with free atomic bonds that drastically slow the flow of charges in the device. If the interface between a III-V semiconductor and

allows them to alternate transistors that carry negative charges (n-type) and positive charges (p-type), which, when working together, give clearer electronic readouts and use less power. Silicon ferries holes nearly as well as it does electrons. But most III-Vs are far better at conducting electrons than holes. "There are a lot of problems that may seem insurmountable," Schlom says, "but I'm cautiously optimistic that they can be solved."

Some steps toward those solutions are already evident. In 2007, for example, researchers at Intel led by Robert Chau reported growing InGaAs transistors atop a thin buffer layer sitting on a silicon wafer. The buffer allowed the two dissimilar materi-



**New kid.** Transistors made from graphene nanoribbons could be blinding fast. But can they perform on an industrial scale?