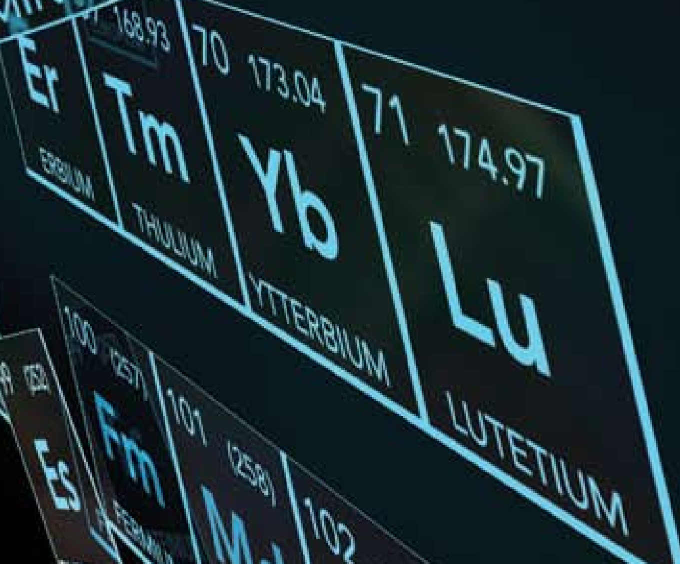
P
SULPHUR
35
Br
34
78.96
Se
BROMINE
SELENIUM
Si
PHOSPHORUS
SILICON
54
131.29
Xe
53
126.90

49
114.82
In
INDIUM

33
74.922
As
ARSENIC
(222)
RADON

31
69.723
Ga
GALLIUM

13
26.982
Al
ALUMINIUM



168.93
Er
Tm
ERBIUM
THULIUM
70
173.04
Yb
YTTERBIUM
71
174.97
Lu
LUTETIUM
100
(257)
Fm
101
(258)
102

Proof #4 COLOR 8/5/08 @ 5:15 pm BP

BEYOND SILICON'S ELEMENTAL LOGIC

IN THE QUEST FOR SPEED, KEY PARTS OF MICROPROCESSORS MAY SOON BE MADE OF GALLIUM ARSENIDE OR OTHER "III-V" SEMICONDUCTORS BY PEIDE D. YE

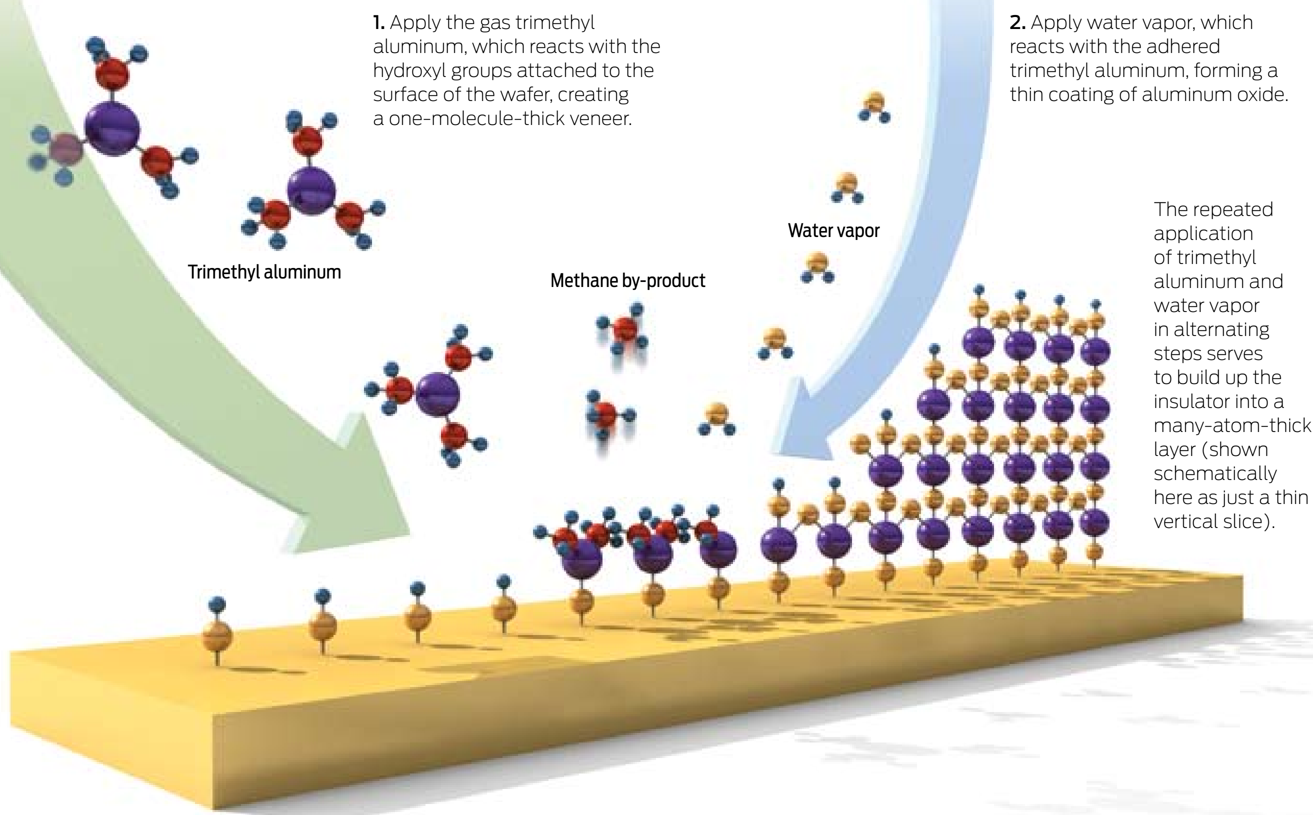
THE FIRST general-purpose microprocessor, the Intel 8080, released in 1974, could execute about half a million instructions per second. At the time, that seemed pretty zippy.

Today the 8080's most advanced descendant operates 100 000 times as fast. This phenomenal progress is a direct result of the semiconductor industry's ability to reduce the size of a microprocessor's fundamental building blocks—its many metal-oxide-semiconductor field-effect transistors (MOSFETs), which act as tiny switches. Through the magic of photolithography, billions of them are routinely constructed en masse on the surface of a silicon wafer.

As these transistors got smaller over the years, more could fit on a chip without raising its overall cost. They also gained the ability to turn on and off at increasingly rapid rates, allowing microprocessors to hum along at ever-higher speeds.

But shrinking MOSFETs much beyond their current size—a few tens of nanometers—will be a herculean challenge. Indeed, at some point in the next several years, it may become impossible to make them more minuscule, for reasons

ALL ARTWORK: BRYAN CHRISTIE DESIGN



of fundamental physics rather than nuts-and-bolts engineering. So people like me have been looking at other ways to boost their speed. In particular, we've been laboring to build them using compound semiconductors like gallium arsenide, which would allow such transistors to switch on and off much faster than their silicon cousins can.

This strategy is by no means new. Practically ever since the silicon MOSFET was invented in 1960, engineers have been attempting to come up with a gallium arsenide version suitable for large-scale integrated circuits. No one has yet succeeded. Those repeated failures have led to the oldest joke in Silicon Valley: gallium arsenide—it's the technology of the future, and it always will be.

But that perennial skepticism may be about to vanish. My colleagues and I at Purdue University's Birck Nanotechnology Center, in West Lafayette, Ind., along with other researchers in industry and academia, have recently made some advances that might soon allow transistors built with gallium arsenide or a related compound to be used for large-scale digital ICs. That capability would go a long way toward bringing us microprocessors that can blaze along at triple or even quadruple the speed of today's best. Achieving that goal will no doubt require other improvements in semiconductor technology to take place in parallel, but gallium arsenide or something close to it could be key. No wonder some of us have been unwilling to give up on this remarkable material.

GALLIUM ARSENIDE'S two main components come from the third (gallium) and fifth (arsenic) columns in the right-hand portion of the periodic table of elements, which is why cognoscenti refer to it as a III-V semiconductor. There are more than a dozen such compounds, including gallium nitride and indium phosphide, but gallium arsenide is the most common example and therefore the best studied. It currently accounts for about 2 percent of the semiconductor market.

Gallium arsenide devices cost a lot more than ones built of silicon—the raw materials are about 10 times as expensive—but they serve well for certain specialized applications, including high-efficiency solar cells, laser diodes, and one very special kind of field-effect transistor: the high-electron-mobility transistor, or HEMT, which is used in cellphones, communication systems, and radars, among other things.

HEMTs are remarkable devices because they overcome a fundamental problem of solid-state physics. Semiconductors, as their name implies, normally don't conduct electricity all that well. Usually,

they must be doped with other kinds of atoms to become electrically conductive. But those impurities tend to interfere with the movement of electrons through the semiconductor's crystal lattice, limiting the conductivity that can be obtained.

In HEMTs, electrons are introduced into a III-V semiconductor not by doping but by placing the material in contact with another III-V compound that *is* doped. In essence, electrons fall a short distance into the undoped material, allowing a thin layer of it—the channel—to conduct electricity extremely well whenever the transistor is switched on.

HEMTs can be used singly or in integrated circuits with, say, 100 or even 1000 of them clustered together, but they can't yet work for microprocessors. The problem is that too many of the electrons that are supposed to flow through the channel from the transistor's source electrode to its drain instead seep out the controlling input electrode—the gate—creating heat. With millions of leaky transistors crowded together on the same chip, things would quickly get hot enough to melt.

In a silicon MOSFET, a layer of intervening insulation (traditionally silicon dioxide) prevents electrons from slipping out of the channel into the gate. In a HEMT, the channel is separated from the gate by a semiconductor, which, as you might expect, is somewhat conductive. What's needed here, of course, is an insulator, but for decades there have been no good gate insulators available for gallium arsenide. From time to time over the years, researchers seem to uncover a promising material, but nothing ever really panned out—until recently.

It's easy enough to understand, at least in general terms, why finding a gate insulator for silicon was easier than it has been for gallium arsenide. Silicon dioxide is a native oxide of silicon—a naturally forming coat that grows when silicon is exposed to oxygen. By good fortune, silicon dioxide makes for an excellent chemical marriage with the silicon it covers: only one out of 100 000 silicon atoms at the interface fails to bond with the adjacent silicon dioxide, leaving what's called a dangling bond. These defects disrupt the flow of electrons in the channel, but they are rare enough that they don't materially degrade the performance of a transistor.

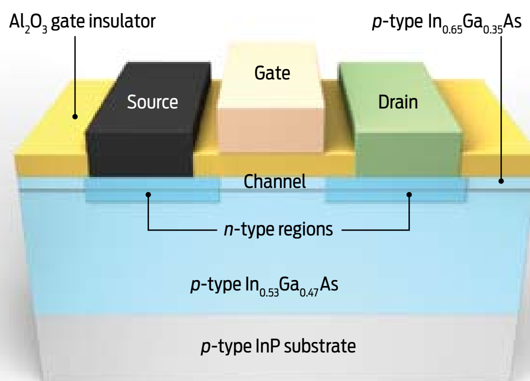
Gallium arsenide is a different story. When it oxidizes, it forms a complex mixture of Ga_2O_3, As_2O_3, and As_2O_5. Beginning in the 1960s, some researchers tried using these native oxides for a gate insulator, but that tactic proved worse than useless because the native oxides create all kinds of defects at the interface with the gallium arsenide, which destroy the electrical conductivity of the adjacent channel. Clearly, a better material needed to be found if there was to be any hope of making gallium arsenide MOSFETs for digital ICs.

RESEARCHERS continued the decades-long quest, testing silicon dioxide, silicon nitride, silicon oxynitride, and aluminum oxide, among other candidates. They also tried adding a third material, such as sulfur, silicon, or germanium, between the substrate and the insulator to negate the pernicious effects of dangling bonds. Yet the results always proved disappointing, and by the early 1990s most investigators had simply given up. Two notable exceptions were Minghwei Hong and Matthias Passlack at Bell Labs, who had developed a way of depositing a combination gallium oxide–gadolinium oxide insulator on a III-V substrate, a strategy that Passlack later refined at Motorola and at a company Motorola spun off in 2004, Freescale Semiconductor.

At that time, the engineers making silicon MOSFETs were beginning to experience trouble with their gate insulator too. As the dimensions of these transistors shrank, the silicon dioxide insulating the gate no longer functioned well. Indeed, it became so thin that electrons could pass through it as if it were a sieve. Great effort went into finding alternative materials with higher dielectric constants, which could be made physically thicker without compromising the electrical functioning of the transistor. Eventually, suitable compounds were found. Intel, for instance, is now using a hafnium-based gate insulator on some of its most advanced microprocessors [see "The High-*k* Solution," *IEEE Spectrum*, October 2007].

To control the thickness of these "high-*k*" dielectrics (a designation that refers to the symbol used for a material's dielectric constant, the Greek letter kappa), manufacturers apply them to the silicon substrate using a technique called atomic-layer deposition. It's quite ingenious, really. The trick is that you employ a chemical carrier molecule that sticks to the target surface but not to itself. Such a chemical thus deposits a one-molecule-thick veneer. A second treatment with another carrier molecule removes the first carrier, leaving a two-atom-thick layer of the desired material behind. Repeated application of the two carrier gases, one alternating with the other, allows chip makers to deposit various high-*k* gate insulators on silicon with atomic-level precision.

In 2001, Glen Wilk, then my colleague at Bell Labs, and I decided to try to put a high-*k* gate insulator—in this case, aluminum oxide (Al_2O_3)—on top of gallium arsenide using atomic-layer deposition, which was all the rage at the



AN EXPERIMENTAL TRANSISTOR of indium gallium arsenide (blue) is built on a bed of indium phosphide (gray). Positive voltage applied to the gate draws electrons into the channel between the silicon-doped *n*-type regions beneath the source and drain electrodes, allowing current to flow.

YEARS IN THE MAKING

1960 Dawon Kahng and Martin M. Atalla at Bell Labs invent the MOSFET.

1965 Hans Becke, Robert Hall, and Joseph White at RCA devise the first gallium arsenide MOSFET using silicon dioxide for the gate insulator.

1979 Takashi Mimura at Fujitsu Laboratories invents a type of gallium arsenide FET: the high-electron-mobility transistor (HEMT).

1995 Minghwei Hong and Matthias Passlack at Bell Labs deposit a gallium oxide–gadolinium oxide insulator on a III-V substrate.

2001 The author of this article, Peide Ye, and Glen Wilk deposit aluminum oxide insulator on a gallium arsenide substrate using atomic-layer deposition.

2005 Intel announces interest in III-V semiconductors for future microprocessors.

2007 The author and his colleagues measure record-breaking current for a III-V MOSFET.

time. After 2003, our team continued to study this approach at Agere Systems, in Allentown, Pa., a spin-off of Bell Labs and Lucent Technologies, where our group had been moved. The maneuver succeeded better than we could have dreamed.

That's not to say we were able to make a perfectly functioning MOSFET out of gallium arsenide straight off. Rather, what stunned us early on was that atomic-layer deposition allowed us to apply the Al_2O_3 despite having done nothing to remove the troublesome native oxide from the gallium arsenide. The reason, as researchers at the University of Texas at Dallas have recently detailed, was that the first carrier, a molecule called trimethyl aluminum, eats away at gallium arsenide's native oxides, which despite all reasonable precautions tend to cover the substrate. It's the atomic-scale equivalent of the mold on an old porch floor. And as any homeowner knows, if you want to repaint those boards, you'd better scrape off the gunk first.

Using trimethyl aluminum was like having an all-in-one product that strips, primes, and paints all at once. If you want additional coats—that is, a thicker film of Al_2O_3—just repeat the application of trimethyl aluminum and the second carrier, water vapor, in alternating steps.

Once you grow a suitably thick layer of aluminum oxide on gallium arsenide in this way, you use traditional lithography to construct the drain, source, gate, and other components of a MOSFET. No special processes are required. The rub is that the transistor you'll end up with will be a dud: it won't pass any more current through its channel than did some of the failed designs of decades past.

WHEN I CAME TO Purdue three and a half years ago, Yi Xuan, a postdoctoral investigator in my research group, and I took on this problem of dismally poor current capacity. At about that time, Intel announced that its engineers were seriously considering the use of III-V semiconductors in its future chips. IBM, too, made its interest in this technology known. The quest for speed, it seemed, was driving a renaissance in research on how to make III-V semiconductors for digital applications. But despite all this attention from some of the biggest guns in the industry, nobody had a clear idea about how to achieve sufficient current-carrying capacity for III-V MOSFETs. The challenge was greatest for those operated in enhancement mode, meaning that electrons flow from source to drain only when a voltage is applied to the gate, as is the case for the silicon MOSFETs found in digital ICs.

Based on published work carried out almost a decade earlier at Bell Labs and on my own research on depletion-mode MOSFETs, which switch off when voltage is applied to the gate, I realized that a related III-V semiconductor—indium gallium arsenide—would serve better for the channel. In this compound, indium atoms substitute for galliums to a degree that can be adjusted arbitrarily. That is, you can have mostly indium atoms, mostly gallium atoms, or a 50:50 mix of the two bonding to the arsenic atoms.

Tinkering with the indium content allowed us to engineer the substrate's electronic properties as required, instead of trying to work around the givens of a particular material. After much experimentation, we settled on a composition that had a 65:35 ratio of indium to gallium. With it we were able to build a MOSFET that carried more than 1 ampere per millimeter of channel width—the highest current density ever produced in four decades of work on gallium arsenide MOSFETs. Indeed, it was so large that it initially sent our semiconductor parameter analyzer off scale!

One well-known difficulty with this approach is that indium gallium arsenide has very poor mechanical properties, so poor that it would be problematic, if not impossible, to use it to make wafers. Pure gallium arsenide is much more robust. Our wafer supplier, a UK-based company called IQE, was able to overcome this hurdle by growing a thin layer of indium gallium arsenide on a thick base of indium phosphide. These two compounds have crystal lattices of similar sizes, so they bond reasonably well together. And the mechanical properties of indium phosphide, while not ideal, proved good enough to allow us to construct various test transistors.

Passlack and his co-workers at Freescale Semiconductor and the University of Glasgow have also been experimenting with indium gallium arsenide over the past few years, using a gallium oxide–gadolinium oxide insulator. Hong, who is now at National TsingHwa University in Taiwan, continues work on this combination as well. Although such MOSFETs have shown a reasonably good ability to carry current, they would be difficult to manufacture. The problem is that they require two applications of a high-vacuum deposition technique called molecular-beam epitaxy: one to lay down the indium gallium arsenide and then a second to coat it with the gate oxide. Doing molecular-beam epitaxy twice, all the while keeping things under high vacuum, is possible in the lab, but it would be

a challenge for industrial-scale production.

Research groups at the National University of Singapore and at IBM are pursuing yet another design that has lately shown promise, one that uses chemical means to add a layer of amorphous silicon between the semiconductor and the gate insulator. This approach resembles a strategy that was attempted two decades ago and is similar to work going on now at the University of Texas at Austin and at the State University of New York in Albany.

What's more, some researchers are looking to build a very different kind of III-V field-effect transistor suitable for digital applications, one that can function without a gate oxide at all. These devices operate similarly to HEMTs in that a semiconductor provides the barrier between the gate and a highly conductive, undoped channel. Intel and UK-based QinetiQ in particular have over the past few years achieved impressive performance with a transistor fashioned this way using an indium antimony channel. Jesús A. del Alamo and his colleagues at MIT are also investigating how to make such HEMT-like transistors smaller and less prone to gate leakage so that they may one day serve for digital applications.

Indeed, much work goes on around the world on bringing III-V semiconductors into what has long been the sole domain of silicon. In addition to the efforts being mounted in industry, academic teams have formed centers of research at the University of California at Santa Barbara, the University of Glasgow, and the University of Tokyo, specifically to carry out these investigations.

CONSIDERABLE progress will yet have to be made before any of these new kinds of field-effect transistors replace their slower silicon counterparts in microprocessors, memory chips, and other digital ICs. In particular, device engineers will have to optimize many parameters besides current-carrying capacity and gate leakage. They'll also want these transistors to be able to operate at low voltages, for instance, so as to reduce another troublesome source of heating: the power expended at the moment the transistors switch states. (Indium-rich indium gallium arsenide holds great promise in this regard.) Designers will also want to ensure that very little current flows when a transistor is nominally "off," so that power isn't expended—and heat isn't generated—uselessly. Doing so, all while making these transistors as tiny as today's silicon wonders, will be no small feat.

In addition, it is likely that manufacturers will have to find ways to place III-V semiconductors on top of a silicon wafer. That is, chip makers will surely aim to use the compound semiconductors only where they're needed rather than trying to replace silicon entirely, although getting all these materials to work properly together turns out to be a tricky undertaking and the subject of much research.

One reason for keeping as much silicon as possible around is that it has considerably better physical properties for making the large wafers used in semiconductor manufacturing. Also, silicon is cheap and environmentally friendly, whereas gallium arsenide is expensive and, because it contains arsenic, potentially quite toxic.

Another reason not to expect an all–gallium arsenide microprocessor anytime soon is that III-V semiconductors can speed up only half the transistors in a CMOS chip: the *n*-channel ones, which carry current in the form of negative charges—electrons. CMOS integrated circuits require a combination of both *n*-channel and *p*-channel MOSFETs, which together draw power only when they switch states, such as when an *n*-channel transistor turns on and the *p*-channel transistor that's wired in series with it turns off. When not switching between states, such a complementary pair draws no power, which is what makes CMOS chips so energy efficient.

Although gallium arsenide allows electrons to move through it especially easily, it doesn't offer any advantage over silicon for positive charge carriers—the "holes," which are sites in the semiconductor's crystal lattice that are deficient in outer-shell electrons. So it would be very difficult to make a high-performance *p*-channel MOSFET using gallium arsenide or another III-V compound. The current consensus is that the semiconductor industry will probably employ germanium for those transistors. The Duallogic academia-industry consortium in Europe, for example, is working to combine germanium and III-V semiconductors in this way.

The III-V devices that my Purdue colleagues and I have recently constructed represent a whopping leap forward, as these MOSFETs are both easy to fabricate and able to carry record currents. The competing designs offer some attractive features too. Still, many barriers stand in the way of their widespread use. In particular, chip makers will have to learn to mix and match some very different kinds of semiconductors on a single wafer. Perhaps chip makers will have to weave together a patchwork quilt of indium gallium arsenide and germanium on a bed of silicon, or maybe it will be something even more complicated. But if there's any lesson to be drawn from the past four decades of dizzying advances in computing power, it's that this industry thrives on a challenge. ❏

TO PROBE FURTHER Details of the author's work in this area are available in "High-Performance Inversion-Type Enhancement-Mode InGaAs MOSFET With Maximum Drain Current Exceeding 1 A/mm," by Y. Xuan, Y.Q. Wu, and P.D. Ye, IEEE Electron Device Letters *29:294, April 2008.*

To learn more about the use of a gallium oxide–gadolinium oxide insulator on a III-V substrate, see "High Mobility III-V MOSFET Technology" by M. Passlack, R. Droopad, K. Rajagopalan, J. Abrokwah, P. Zurcher, R. Hill, D. Moran, X. Li, H. Zhou, D. Macintyre, S. Thoms, and I. Thayne at http:// www.gaasmantech.org/Digests/2007/2007%20Papers/12c.pdf.