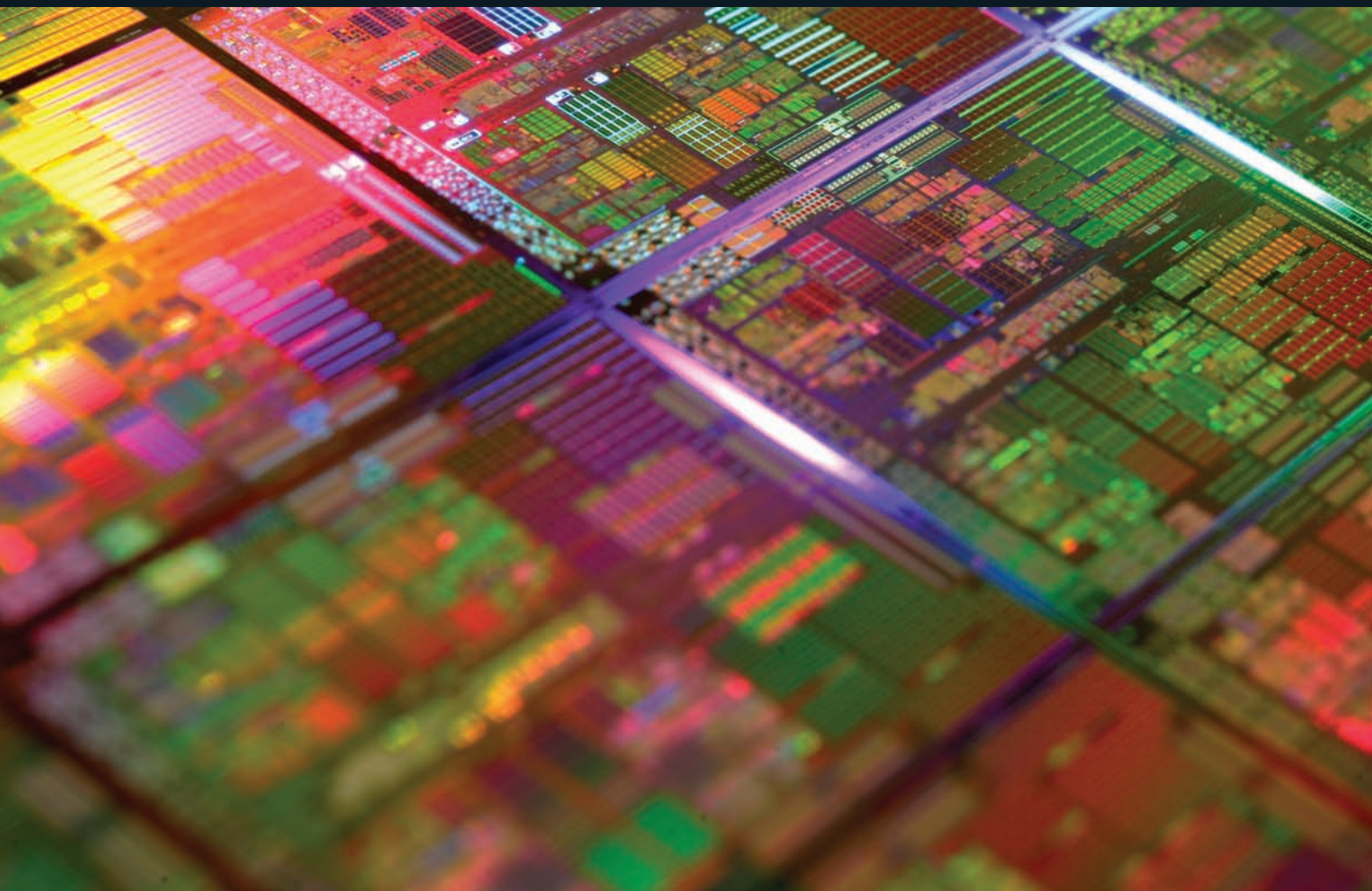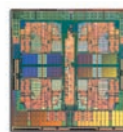# THE NEXT 20 YEARS OF
MICROCHIPS

Designers are pushing all the boundaries to make integrated circuits smaller, faster and cheaper □ □ □ **By the Editors**

I n 1975 electronics pioneer Gordon Moore famously predicted that the complexity of integrated-circuit chips would double every two years. Manufacturing advances would allow the chip's transistors to shrink and shrink, so electrical signals would have to travel less distance to process information. To the electronics industry and to consumers, Moore's Law, as it became known, meant computerized devices would relentlessly become smaller, faster and cheaper. Thanks to ceaseless innovation in semiconductor design and fabrication, chips have followed remarkably close to that trajectory for 35 years.

Engineers knew, however, they would hit a wall at some point. Transistors would become only tens of atoms thick. At that scale, basic laws of physics would impose limits. Even before the wall was hit, two practical problems were likely to arise. Placing transistors so small and close together while still getting a high yield—usable chips versus defective ones—could become overly expensive. And the heat generated by the thicket of transistors switching on and off could climb enough to start cooking the elements themselves.

Indeed, those hurdles arose several years ago. The main reason common personal computers now have the loudly marketed "dual-core" chips—meaning two small processors instead of one—is because packing the needed number of transistors onto a single chip and cooling it had become too problematic. Instead computer designers are choosing to place two or more chips side by side and program them to process information in parallel.

Moore's Law, it seems, could finally be running out of room. How, then, will engineers continue to make chips more powerful? Switching to alternative architectures and perfecting nanomaterials that can be assembled atom by atom are two options. Another is perfecting new ways to process information, including quantum and biological computing. In the pages ahead, we take a look at a range of advances, many currently at the prototype stage, that in the next two decades could keep computing products on the "smaller, faster, cheaper" path that has served us so well.

KEY CONCEPTS

- It may soon be impossible to make transistors on integrated-circuit chips even smaller. Alternative materials and designs will be needed for chips to continue to improve.

- Nanowires, graphene, quantum particles and biological molecules could all spawn new generations of chips that are more powerful than today's best.

—The Editors
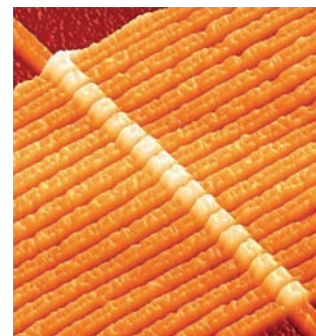
SIZE: CROSSING THE BAR

☐ ☐ ☐

The smallest commercial transistors now made are only 32 nanometers wide—about 96 silicon atoms across. The industry acknowledges that it may be extremely hard to make features smaller than 22 nanometers using the lithography techniques that have improved for decades.

One option that has circuit features of a similar size but offers greater computing power is known as crossbar design. Instead of fabricating transistors all in one plane (like cars packed into the lanes of a jammed silicon highway), the crossbar approach has a set of parallel nanowires in one plane that crosses over a second set of wires at right angles to it (two perpendicular highways). A buffer layer one molecule thick is slipped between them. The many intersections that exist between the two sets of wires can act like switches, called memristors, that represent 1s and 0s (binary digits, or bits) the way transistors do. But the memristors can also store information. Together these capabilities can perform a number of computing tasks. Essentially one memristor can do the work of 10 or 15 transistors.

Hewlett-Packard Labs has fabricated prototype crossbar designs with titanium and platinum wires that are 30 nanometers wide, using materials and processes similar to those already optimized for the semiconductor industry. Company researchers think each wire could get as small as eight nanometers. Several research groups are also fashioning crossbars made from silicon, titanium and silver sulfide.

MEMRISTOR, from Hewlett-Packard, is a new kind of circuit element created at each raised intersection of overlapping nanowires.
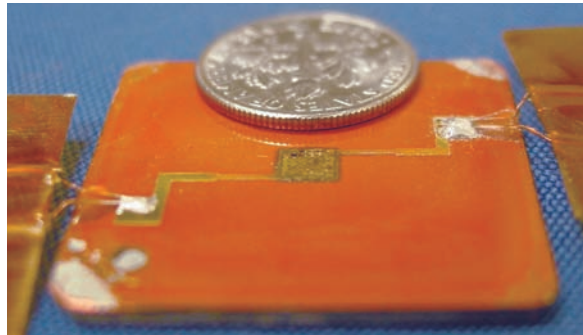
With as many as one billion transistors on a chip, getting rid of heat generated as the transistors switch on and off is a major challenge. Personal computers have room for a fan, but even so about 100 watts of power dissipation per chip is as much as they can cool. Designers are therefore devising some novel alternatives. The MacBook Air notebook computer has a sleek case made from thermally conductive aluminum that serves as a heat sink. In the Apple Power Mac G5 personal computer, liquid runs through microchannels machined into the underside of its processor chip.

Fluids and electronics can be a dicey mix, however, and smaller, portable gadgets such as smart phones simply do not have room for plumbing—or fans. A research group led by Intel has crafted a thin-film superlattice of bismuth telluride into the packaging that encases a chip (*above*). The thermoelectric material con-

COOLING PATCH (*center, gold*) made of bismuth telluride would transfer heat away from a much larger chip fastened on top of it to a thin dissipation layer (*orange*). The patch and layer use less space and power than current heat sinks.

verts temperature gradients into electricity, in effect refrigerating the chip itself.
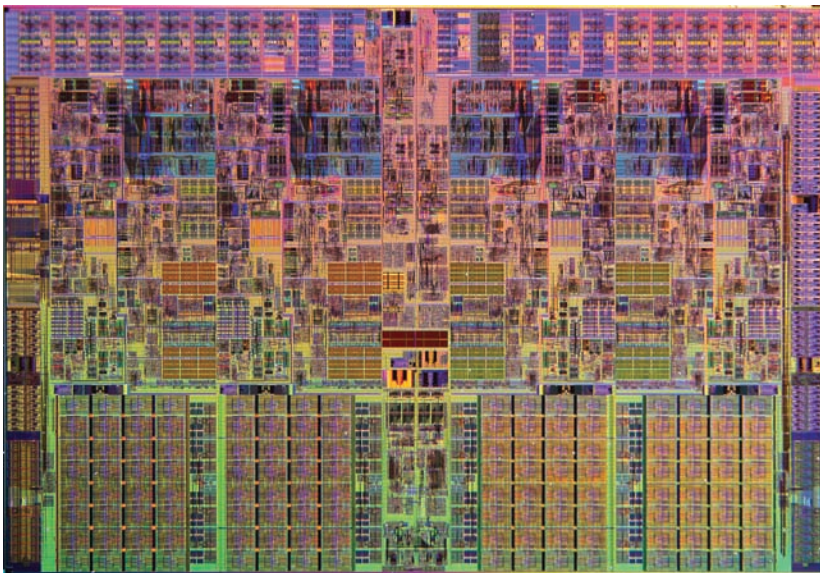
Based on work at Purdue University, start-up company Ventiva is making a tiny solid-state "fan" with no moving parts that creates a breeze by harnessing the corona wind effect—the same property exploited by silent household air purifiers. A slightly concave grating has live wires that generate a microscale plasma; the ions in this gaslike mixture drive air molecules from the wires to an adjacent plate, generating a wind. The fan produces more airflow than a typical mechanical fan yet is much smaller. Other innovators are crafting Stirling engine fans, still somewhat bulky, that create wind but consume no electricity; they are powered by the difference in temperature between hot and cool regions of the chip.

INTEL I7 processor has four cores (*bottom*) that work in parallel to quicken computation.

Smaller transistors can switch between off and on to represent 0 and 1 more quickly, making chips faster. But the clock rate—the number of instructions a chip can process in a second—leveled off at three to four gigahertz as chips reached the heat ceiling. The desire for even greater performance within the heat and speed limits led designers to place two processors, or cores, on the same chip. Each core operated only as quick-

ly as previous processors, but because the two worked in parallel they could process more data in a given amount of time and consumed less electricity, producing less heat. The latest personal computers now sport quadruple cores, such as the Intel i7 and the AMD Phenom X4.

The world's most powerful supercomputers contain thousands of cores, but in consumer products, using even a few cores most effectively requires new programming techniques that can partition data and processing and coordinate tasks. The basics of parallel programming were worked out for supercomputers in the 1980s and 1990s, so the challenge is to create languages and tools that software developers can use for consumer applications. Microsoft Research, for example, has released the F# programming language. An early language, Erlang, from the Swedish company Ericcson, has inspired newer languages, including Clojure and Scala. Institutions such as the University of Illinois are also pursuing parallel programming for multiple-core chips.

If the approaches can be perfected, desktop and mobile devices could contain dozens or more parallel processors, which might individually have fewer transistors than current chips but work faster as a group overall.

For a decade already, pundits have hailed nano-technology as the solution to all sorts of challenges in medicine, energy and, of course, integrated circuitry. Some enthusiasts argue that the semiconductor industry, which makes chips, actually created the nanotechnology discipline as it devised ever tinier transistors.

The higher expectation, however, is that nanotechniques would allow engineers to craft designer molecules. Transistors assembled from carbon nanotubes, for example, could be much smaller. Indeed, engineers at IBM have fabricated a traditional, complementary metal-oxide-semiconductor (CMOS) circuit that uses a carbon nanotube as the conductive substrate, instead of silicon (*right*). Joerg Appenzeller from that team, now at Purdue University, is devising new transistors that are far smaller than CMOS devices, which could better exploit a minuscule nanotube base.

Arranging molecules or even atoms can be tricky, especially given the need to assemble them at high volume during chip production. One solution could be molecules that self-assemble: mix them together, then expose them to heat or light or centrifugal forces, and they will arrange themselves into a predictable pattern.

IBM has demonstrated how to make memory circuits using polymers tied by chemical bonds. When spun on the surface of a silicon wafer and heated, the molecules stretch and form a honeycomb structure with pores only 20 nanometers wide. The pattern could subsequently be etched into the silicon, forming a memory chip at that size.

RING OSCILLATOR circuit is built on a single carbon nanotube that connects the circuit elements.

Traditional CMOS circuit

Chip base
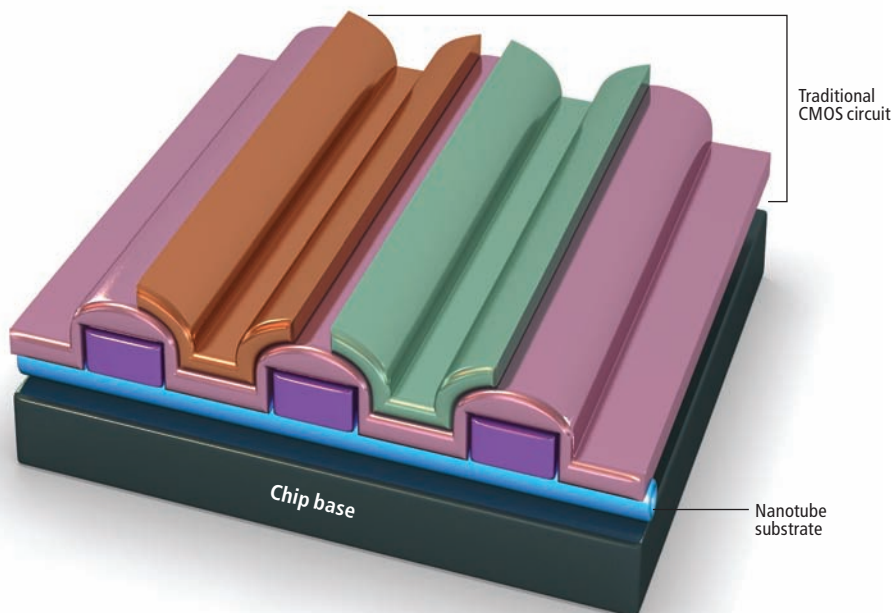
Nanotube substrate

The point of continually shrinking transistors is to shorten the distance that electrical signals must travel within a chip, which increases the speed of processing information. But one nanomaterial in particular—graphene—could function faster because of its inherent structure.

Most logic chips that process information use field-effect transistors made with CMOS technology. Think of a transistor as a narrow, rectangular layer cake, with an aluminum (or more recently, polysilicon) layer on top, an insulating oxide layer in the middle, and a semiconducting silicon layer on the bottom. Graphene—a newly isolated form of carbon molecule—is a flat sheet of repeating hexagons that looks like chicken wire but is only one atomic layer thick. Stacked atop one another, graphene sheets form the mineral graphite, familiar to us as pencil "lead." In its pure crystal form, graphene conducts electrons faster than any other material at room temperature—far faster than field-effect transistors do. The charge carriers also lose very little energy as a result of scattering or colliding with atoms in the lattice, so less waste heat is generated. Scientists isolated graphene as a material only in 2004, so work is very early, but researchers are confident they can make graphene transistors that are just 10 nanometers across and one atom high. Numerous circuits could perhaps be carved into a single, tiny graphene sheet.

Gate electrode

Drain

Source

Electron

Quantum dot
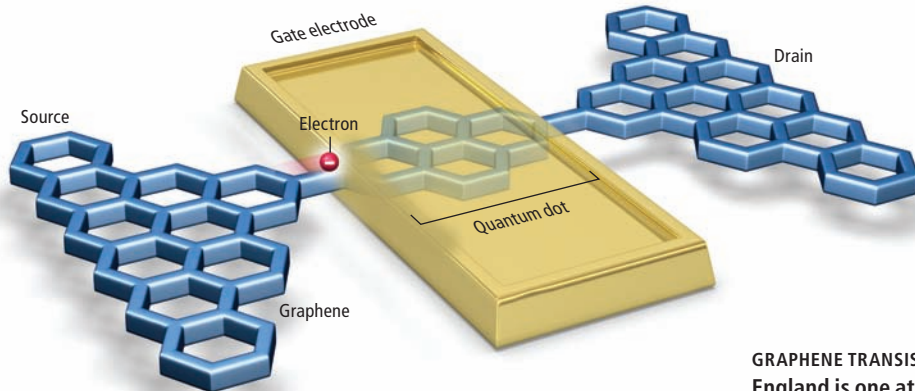
Graphene

GRAPHENE TRANSISTOR fabricated at the University of Manchester in England is one atom thick. A quantum dot allows only a single electron to move from source to drain, registering a 1 or 0.

OPTICAL COMPUTING: QUICK AS LIGHT □ □ □

Radical alternatives to silicon chips are still so rudimentary that commercial circuits may be a decade off. But Moore's Law will likely have run its course by then, so work is well under way on completely different computing schemes.

In optical computing, electrons do not carry information, photons do, and they do so far faster, at the speed of light. Controlling light is much more difficult, however. Progress in making optical switches that lie along fiber-optic cables in telecommunications lines has helped optical computing, too. One of the most advanced efforts, ironically, aims to create an optical interconnect between the traditional processors on multicore chips; massive amounts of data must be shuttled between cores that are processing information in parallel, and electronic wires between them can become a bottleneck. Photonic interconnects could improve the flow. Researchers at Hewlett-Packard Labs are evaluating designs that could move two orders of magnitude more information.

Other groups are working on optical interconnects that would replace the slower copper wires that now link a processor chip to other components inside computers, such as memory chips and DVD drives. Engineers at Intel and the University of California, Santa Barbara, have built optical "data pipes" from indium phosphate and silicon using common semiconductor manufacturing processes (*below*). Completely optical computing chips will require some fundamental breakthroughs, however.

OPTICAL CHIP **can compute fast if it has an internal, controllable light source. Electrons and holes in indium phosphate layers recombine at the center to generate light that spreads down a silicon waveguide and through a glass layer.**
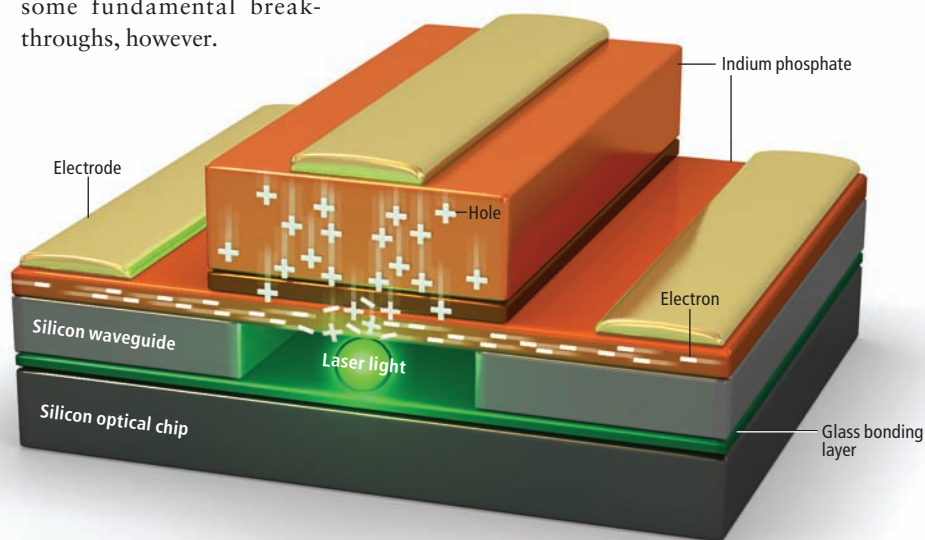
## MOLECULAR COMPUTING: ORGANIC LOGIC □ □ □

In molecular computing, instead of transistors representing the 1s and 0s, molecules do so. When the molecule is biological, such as DNA, the category is known as biological computing [see "Biological Computing: Chips That Live," on opposite page]. To be clear, engineers may refer to computing with nonbiological molecules as molecular logic, or molectronics.

A classic transistor has three terminals (think of the letter Y): source, gate and drain. Applying a voltage to the gate (the stem of the Y) causes electrons to flow between source and drain, establishing a 1 or 0. Molecules with branchlike shapes could theoreti-

cally cause a signal to flow in a similar way. Ten years ago researchers at Yale and Rice universities crafted molecular switches using benzene as a building block.

Molecules can be tiny, so circuits built with them could be far smaller than those made in silicon. One difficulty, however, is finding ways to fabricate complex circuits. Researchers hope that self-assembly might be one answer. In October 2009 a team at the University of Pennsylvania transformed zinc and crystalline cadmium sulfide into metal-semiconductor superlattice circuits using only chemical reactions that prompted self-assembly (*below*).

Zinc Cadmium Sulfur Nanowire
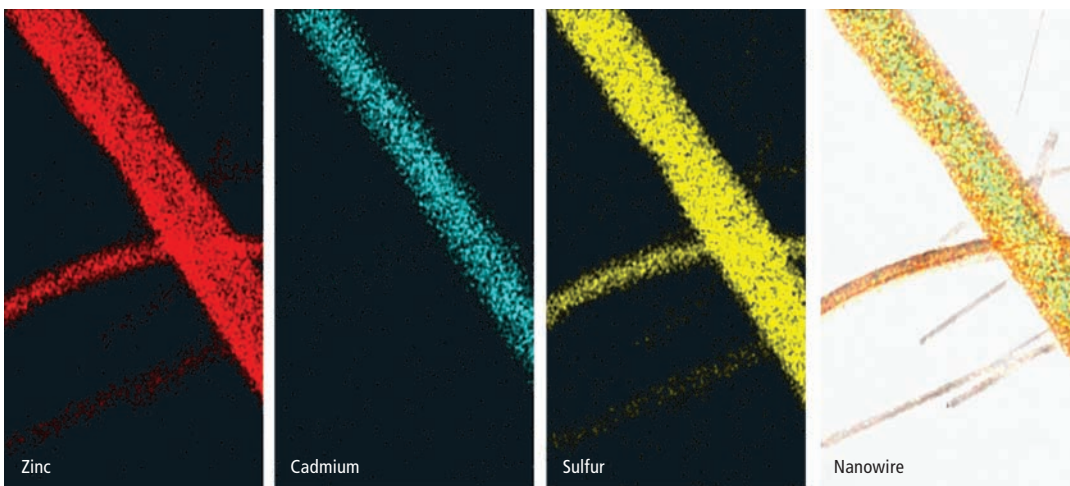
THIN ZINC, **cadmium and sulfur nanowires self-assemble into a thicker nanowire and shell (*far right*) appropriate for a circuit, when exposed to a half-second pulse of dimethylzinc vapor.**

## QUANTUM COMPUTING: SUPERPOSITION OF 0 AND 1 ▫ ▫ ▫

Circuit elements made of individual atoms, electrons or even photons would be the smallest possible. At this dimension, the interactions among the elements are governed by quantum mechanics—the laws that explain atomic behavior. Quantum computers could be incredibly dense and fast, but actually fabricating them and managing the quantum effects that arise are daunting challenges.

Atoms and electrons have traits that can exist in different states and can form a quantum bit, or qubit. Several research approaches to handling qubits are being investigated. One approach, called spintronics, uses electrons, whose magnetic moments spin in one of two directions; think of a ball spinning in one direction or the other (representing 1 or 0). The two states can also coexist in a single electron, however, creating a unique quantum state known as a superposition of 0 and 1. With superposition states, a series of electrons could represent exponentially more information than a string of silicon transistors that have only ordinary bit states. U.C. Santa Barbara scientists

have created a number of different logic gates by tapping electrons in cavities that are etched into diamond.

In another approach being pursued by the University of Maryland and the National Institute of Standards and Technology, a string of
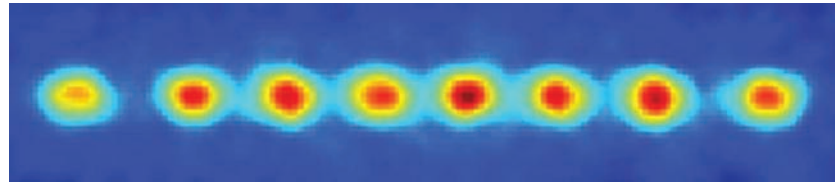

ions is suspended between charged plates, and lasers flip each ion's magnetic orientation (their qubits). A second option is to detect the different kinds of photons an ion emits, depending on which orientation it takes.

In addition to enjoying superposition, quantum elements can become "entangled." Information states are linked across many qubits, allowing powerful ways to process information and to transfer it from location to location.

LEVITATED STRING of calcium ions in a vacuum chamber can perform quantum calculations.

BIOLOGICAL COMPUTING: CHIPS THAT LIVE ▫ ▫ ▫
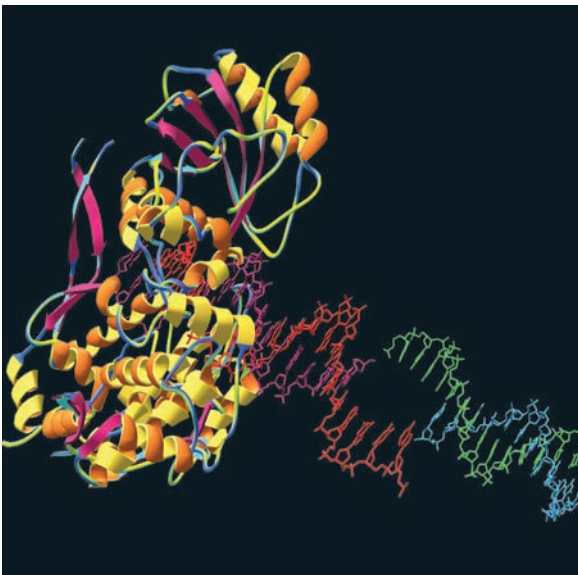
COMPUTATION can occur when a DNA molecule (*right, green*) provides data to software DNA molecules (*center, red*) that a FokI enzyme (*colored ribbon*) can process.

Biological computing replaces transistors with structures usually found in living organisms. Of great interest are DNA and RNA molecules, which indeed store the "programming" that directs the lives of our cells. The taunting vision

is that whereas a chip the size of a pinky fingernail might contain a billion transistors, a processor of the same size could contain trillions of DNA strands. The strands would process different parts of a computing task at the same time and join together to represent the solution. A biological chip, in addition to its having orders of magnitude more elements, could provide massively parallel processing.

Early biological circuits process information by forming and breaking bonds among strands. Researchers are now developing "genetic computer programs" that would live and replicate inside a cell. The challenge is finding ways to program collections of biological elements to behave in desired ways. Such computers may end up in your bloodstream rather than on your desktop. Researchers at the Weizmann Institute of Science in Rehovot, Israel, have crafted a simple processor from DNA (*above*), and they are now trying to make the components work inside a living cell and communicate with the environment around that cell.

➡ **MORE TO EXPLORE**

A Future of Integrated Electronics: Moving Off the Roadmap. Edited by Daniel J. Radack and John C. Zolper. Special issue of *Proceedings of the IEEE,* Vol. 96, No. 2; February 2008.

Carbon Wonderland. Andre K. Geim and Philip Kim in *Scientific American*, Vol. 298, No. 4, pages 90–97; April 2008.

Molecular Implementation of Simple Logic Programs. Tom Ran et al. in *Nature Nanotechnology,* Vol. 4, pages 642–648; October 2009.