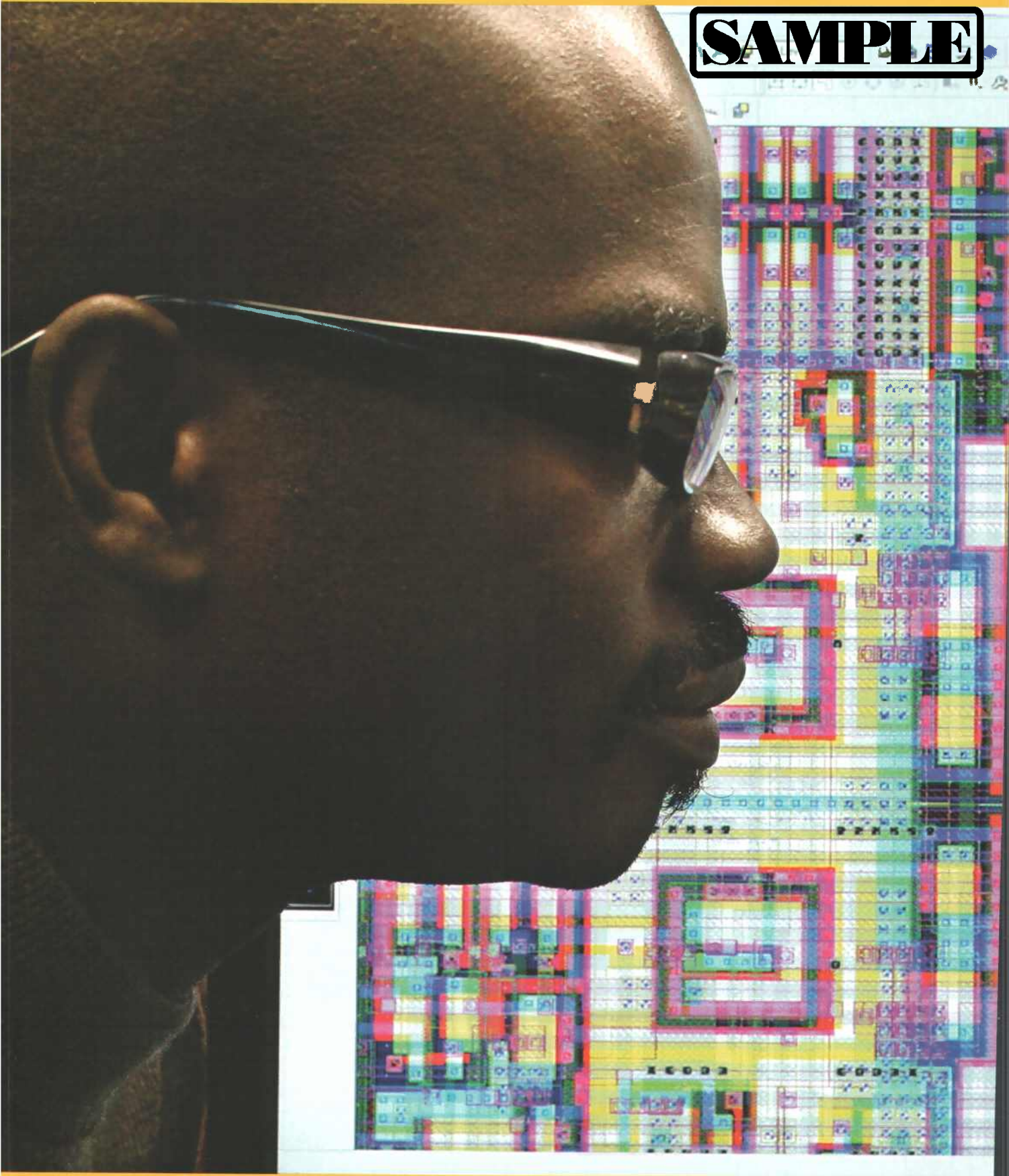


**SAMPLE**



CAN WE BUILD AN ARTIFICIAL BRAIN?

# THINKING MACHINE

The future of computing may depend on embracing the chaos that defines the human brain. BY DOUGLAS FOX

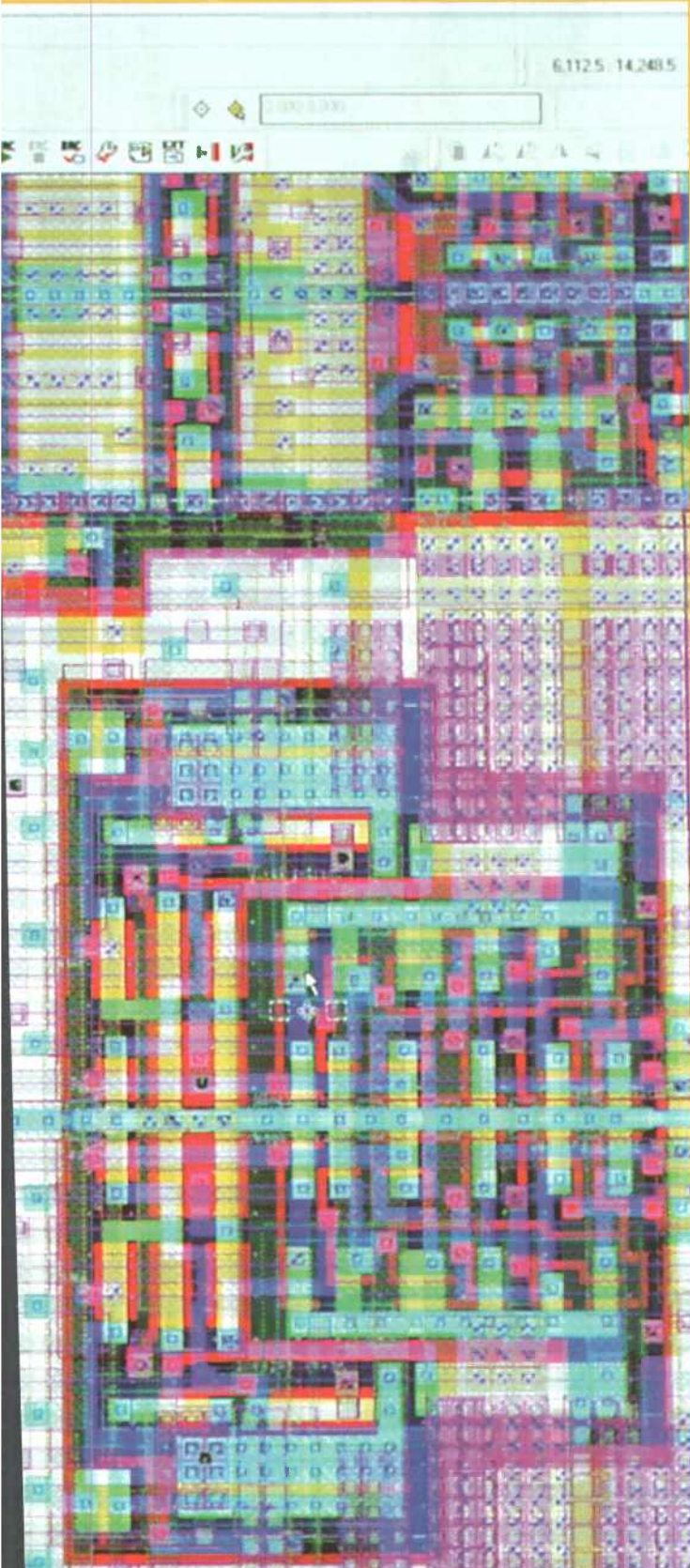
**K**wabena Boahen's love affair with digital computers began and ended in 1981, when he was 16.

Boahen lived outside the city of Accra in the West African nation of Ghana. His family's sprawling block house stood in a quiet field of mango and banana trees. One afternoon Boahen's father rolled down the driveway with a surprise in the trunk of his Peugeot: a RadioShack TRS-80—the family's first computer—purchased in England.

Young Boahen parked the machine at a desk on the porch, where he usually dismantled radios and built air guns out of PVC pipe. He plugged the computer into a TV set to provide a screen and a cassette recorder so he could store programs on tapes, and soon he was programming it to play Ping-Pong. But as he read about the electronics that made it and all other digital computers work, he soured on the toy.

Moving the Ping-Pong ball just one pixel across the screen required thousands of 1s and 0s, generated by transistors in the computer's processor that were switching open and shut 2.5 million times per second. Boahen had expected to find elegance at the heart of his new computer. Instead he found a Lilliputian bureaucracy of binary code. "I was totally disgusted," he recalls. "It was so brute force." That disillusionment inspired a dream of a better solution, a vision that would eventually guide his career.

Boahen has since crossed the Atlantic Ocean and become a prominent scientist at Stanford University in California. There he is working to create a computer that will fulfill his boyhood vision—a new kind of computer, based not on the regimented order of traditional silicon chips but on the organized chaos of the human brain. Designing this machine will mean rejecting everything that we have learned over the past 50 years about building computers. But it might be exactly what we need to keep the



Kwabena Boahen in front of a model of the biology-inspired Hippocampus chip, which is far more efficient than conventional chips.

information revolution going for another 50.

The human brain runs on only about 20 watts of power, equal to the dim light behind the pickle jar in your refrigerator. By contrast, the computer on your desk consumes a million times as much energy per calculation. If you wanted to build a robot with a processor as smart as the human brain, it would require 10 to 20 megawatts of electricity. "Ten megawatts is a small hydroelectric plant," Boahen says dismissively. "We should work on miniaturizing hydroelectric plants so we can put them on the backs of robots." You would encounter similar problems if you tried to build a medical implant to replace just 1 percent of the neurons in the brain, for use in stroke patients. That implant would consume as much electricity as 200 households and dissipate as much heat as the engine in a Porsche Boxster.

"Energy efficiency isn't just a matter of elegance. It fundamentally limits what we can do with computers," Boahen says. Despite the amazing progress in electronics technology—today's transistors are  $1/100,000$  the size that they were a half century ago, and computer chips are 10 million times faster—we still have not made meaningful progress on the energy front. And if we do not, we can forget about truly intelligent humanlike machines and all the other dreams of radically more powerful computers.

Getting there, Boahen realized years ago, will require rethinking the fundamental balance between energy, information, and noise. We encounter the trade-offs this involves every time we strain to hear someone speaking through a crackly cell phone connection. We react instinctively by barking more loudly into the phone, trying to overwhelm the static by projecting a stronger signal. Digital computers operate with almost zero noise, but operating at this level of precision consumes a huge amount of power—and therein lies the downfall of modern computing.

In the palm of his hand, Boahen flashes a tiny, iridescent square, a token of his progress in solving that problem. This silicon wafer provides the basis for a new neural supercomputer, called Neurogrid, that he has nearly finished building. The wafer is etched with millions of transistors like the ones in your PC. But beneath that veneer of familiarity hides a radical rethinking of the way engineers do business.

Traditional digital computers depend on millions of transistors opening and closing

with near perfection, making an error less than once per 1 trillion times. It is impressive that our computers are so accurate—but that accuracy is a house of cards. A single transistor accidentally flipping can crash a computer or shift a decimal point in your bank account. Engineers ensure that the millions of transistors on a chip behave reliably by slamming them with high voltages—essentially, pumping up the difference between a 1 and a 0 so that random variations in voltage are less likely to make one look like the other. That is a big reason why computers are such power hogs.

Radically improving that efficiency, Boahen says, will involve trade-offs that would horrify a chip designer. Forget about infinitesimal error rates like one in a trillion; the transistors in Neurogrid will crackle with noise, misfiring at rates as high as 1 in 10. "Nobody knows how we're going to compute with that," Boahen admits. "The only thing that computes with this kind of crap is the brain."

It sounds cockamammy, but it is true. Scientists have found that the brain's 100 billion neurons are surprisingly unreliable. Their synapses fail to fire 30 percent to 90 percent of the time. Yet somehow the brain works. Some scientists even see neural noise as the key to human creativity. Boahen and a small group of scientists around the world hope to copy the brain's noisy calculations and spawn a new era of energy-efficient, intelligent computing. Neurogrid is the test to see if this approach can succeed.

Most modern supercomputers are the size of a refrigerator and devour \$100,000 to \$1 million of electricity per year. Boahen's Neurogrid will fit in a briefcase, run on the equivalent of a few D batteries, and yet, if all goes well, come close to keeping up with these Goliaths.


**T**he problem of computing with noise first occurred to a young neuroscientist named Simon Laughlin three decades ago. Laughlin, then at the Australian National University in Canberra, spent much of 1975 sitting in a black-walled, windowless laboratory with the lights off. The darkness allowed him to study the retinas of blowflies captured from Dumpsters around campus. In hundreds of experiments he glued a living fly to a special plastic platform under a microscope, sunk a wisp-thin electrode into its honeycombed eye, and recorded

**The human brain runs on only about 20 watts of power, equal to the dim light behind the pickle jar in your refrigerator.**

how its retina responded to beams of light. Laughlin would begin recording at noon and finish after midnight. As he sat in the gloomy lab, watching neural signals dance in green light across an oscilloscope, he noticed something strange.

Each fly neuron's response to constant light jittered up and down from one millisecond to the next. Those fluctuations showed up at every step in the neurons' functioning, from the unreliable absorption of light by pigment molecules to the sporadic opening of electricity-conducting proteins called ion channels on the neurons' surfaces. "I began to realize that noise placed a fundamental limit on the ability of neurons to code information," Laughlin says.

Boosting a crackly signal so that it stands above background noise requires energy. Whether you are a neuron or the operator of a ham radio, doubling your signal-to-noise ratio demands quadrupling your energy consumption—a law of rapidly diminishing returns. "The relationship between information and energy is rather deep, and grounded in thermodynamics," says Laughlin, who now works at the University of Cambridge in England. He has spent the last 12 years studying how brains perform the three-way balancing act among information, energy, and noise.



A rendering of the circuitry on a “neuromorphic” chip, which mimics the noisy workings of the human brain. It was designed by John Arthur, a postdoc in Boahen’s lab.

That balance is critical to survival. Neurons are far more efficient than computers, but despite that, the brain still consumes a tremendous amount of energy. While accounting for just 2 percent of our body weight, the human brain devours 20 percent of the calories that we eat.

Functionally, most neurons have a lot of properties similar to those of transistors. Both act as switches that can either transmit or not transmit electrical pulses, depending on signals they receive. The trade-offs that have evolved in humans could not be more different from those that engineers made in designing conventional computers, however. Engineers chose accuracy. Brains, shaped by natural selection, minimize energy consumption at all costs. Skinny neurons require less energy, so evolution shrank them, and brains adapted to operate barely above the noise threshold.

With great efficiency, though, came a lot

of mistakes. Ideally, for example, neurons should fire off electric spikes only when they receive signals from other cells telling them to do so. But the brain’s skinniest neurons sometimes send out random spikes triggered by ion channel proteins’ popping open accidentally. The smaller the neuron, the more sensitive it is to these random channel openings, and the more often these hiccups occur. The brain’s smallest neurons operate “at the limit of biophysics,” Laughlin says. In 2005 he found that shrinking those neurons a tiny bit more meant they would burp out more than 100 random spikes per second.

This flaky behavior places a fundamental limit on how we function. Compensating for random neural noise has shaped the human brain—and human intelligence—from the bottom up: the size and shape of neurons, the wiring pattern of neural circuits, and even the language of spikes that encodes information. In the most basic

sense, the brain manages noise by using large numbers of neurons whenever it can. It makes important decisions (such as “Is that a lion or a tabby cat?”) by having sizable groups of neurons compete with each other—a shouting match between the lion neurons and the tabby cat neurons in which the accidental silence (or spontaneous outburst) of a few nerve cells is overwhelmed by thousands of others. The winners silence the losers so that ambiguous, and possibly misleading, information is not sent to other brain areas.

The brain also filters out errors using a neural code based on coincidences in timing. Consider the “Bill Clinton cells” that neuroscientists have found in the brain’s medial temporal lobe. These neurons fire whenever you see a picture of Bill Clinton, hear his voice, or read his name. (You have similar neurons for each of the hundreds of other people you are familiar with.) A Clinton neuron might give off a spike whenever

it receives, say, 100 or more simultaneous spikes from other neurons. Even if the false-positive rate for each incoming spike is as high as 1 in 2, the collective false-positive rate for 100 spikes arriving at the same time is considerably less.

Laughlin and David Attwell at University College London estimate that neural signaling accounts for 80 percent of the brain's energy use, whereas keeping those neurons charged and ready to fire takes only 15 percent. This finding has major implications. It means that the brain can save energy by containing large numbers of neurons that it rarely uses.

With many extra neurons lying around, each spike can travel along any one of many different routes through the brain. Each of these power-consuming spikes can transmit information along multiple

than 1 to 15 percent—and often just 1 percent—of its neurons at a time. “People hadn’t considered that most neurons in the brain have to be inactive most of the time,” Laughlin says.

The Neurogrid chip mimics the brain by using the same analog process that neurons use to compute. This analog process occurs until a certain threshold is reached, at which point a digital process takes over, generating an electric spike (the spike is like a 1, and lack of a spike is like a 0).

Instead of using transistors as switches the way digital computers do, Boahen builds a capacitor that gets the same voltage a neuron makes. “By using one transistor and a capacitor, you can solve problems that would take thousands of transistors in a modern digital computer,” Boahen says.

**F**ollowing the discovery in the 1980s of the brain’s amazingly efficient method of noisy computing, an engineer and physicist named Carver Mead tried to do the same thing using transistors. Mead, now professor emeritus at Caltech and one of the fathers of modern silicon chips, wanted to find more-efficient ways to compute. When he applied low voltages to a regular transistor, he could coax it to produce currents that had the same dependence on voltage as neuronal membrane currents had. The field now known as neuromorphic engineering was born.

Boahen arrived at Mead’s laboratory in 1990 to pursue his doctorate. Mead’s lab has produced many leaders in the field of neuromorphic electronics, including Boahen, Rahul Sarpeshkar (now at MIT), Paul Hasler (now at Georgia Tech), and Shih-Chii Liu (now at the Institute of Neuroinformatics in Zurich). Mead’s grad students wore sandals and cowboy boots, worked until 1 a.m., and often spent seven or eight years, rather than the usual four or five, earning their Ph.D.s. “It was a fantastically creative environment,” says Sarpeshkar, who graduated a year after Boahen. “We were all having a good time. We weren’t necessarily in a great hurry to graduate.”

Mead’s students read biological journals religiously and then attempted to construct silicon versions of the neural circuits that brain scientists were mapping out. One of Sarpeshkar’s first chips was an early analog of the cochlea, which processes sound in the inner ear. Boahen was working on

retina chips, which produced fuzzy signals and grainy, salt-and-pepper images. These silicon-chip mimics faced the same problems of noise that real neurons face. In silicon the noise arises from manufacturing imperfections, random variations, and thermal fluctuations in the devices. This problem is exacerbated by large variations in electronic currents. “The currents of two transistors are supposed to be identical,” Boahen says, “but at low power they can differ by a factor of two, and that makes everything quite random.”

Recently Sarpeshkar adapted one of his audio chips into a biologically inspired radio frequency cochlea chip, which enables applications for cognitive and ultrahigh-band radios in the future. The chip, unveiled in June, will allow radios to simultaneously listen to a wide range of frequencies—spanning all radio and television broadcasts, along with all cell phone traffic—the way that ears listen to and analyze many sound frequencies at once. Boahen and his students have developed increasingly realistic silicon chips for the retina, which provides primary input to the visual cortex (which identifies objects that we see) and several other brain areas. These chips might one day provide a foundation for medical implants that restore vision in people with eye or brain injuries. For now, they serve as research tools for learning, by trial and error, how the brain encodes information and manages noise.

These chips use anywhere from  $1/10,000$  to  $1/500$  of the power that would be used by equivalent digital circuitry. Still, they represent mere baby steps on the road to building a brain-inspired computer. Until this year, the largest neural machines contained no more than 45,000 silicon neurons. Boahen’s Neurogrid supercomputer, in contrast, contains a million neurons. That tremendous increase in computing power will allow him to test ideas about the brain, and about how to manage noise, on an entirely new scale.

On a sunny Friday afternoon, Boahen walks, in jeans and Ghanaian sandals, into the computer lab at Stanford where his team is putting the final touches on Neurogrid. One of the computer stations is ringed by a shrine of empty Peet’s coffee cups, evidence of the serious amount of caffeine consumed here. “We’re on a chip deadline,” Boahen says, “so we’re pulling 15-hour days.”

John Arthur, an engineer and former

Consider the “Bill Clinton cells” in the medial temporal lobe. These neurons fire whenever you see his picture or read his name.

paths, so your brain can project the same amount of information by firing fewer of them overall. (Think about it: If you are writing in a language that has only two letters, each word has to be pretty long in order to have a unique spelling; if you have 26 letters to choose from, your words can be shorter, and a given sentence, or paragraph, or novel will also contain fewer keystrokes overall.) The brain achieves optimal energy efficiency by firing no more

a computer scientist named Dharmendra Modha recently used 16 digital Blue Gene supercomputer racks to mathematically simulate 55 million neurons connected by 442 billion synapses. The insights gained from that impressive feat will help in the design of future neural chips. But Modha's computers consumed 320,000 watts of electricity, enough to power 260 American households. By comparison, Neurogrid's 1 million neurons are expected to sip less than a watt.

Neurogrid's noisy processors will not have anything like a digital computer's rigorous precision. They may, however, allow us to accomplish everyday miracles

Boahen on the Stanford campus.

Ph.D. student of Boahen's, sits at the Peet's shrine. Arthur's computer monitor displays a schematic of triangles and squares: part of the Neurogrid chip design. The on-screen blueprint of transistors and capacitors represents a single neuron. "It's 340 transistors per neuron," he says.

These circuits are simple compared with living neurons, but they are advanced enough to illustrate the vast gulf in efficiency between digital and neural computing. The mathematical equations that Arthur and others are using to simulate the chip's behavior and test its blueprint for flaws would quickly bog down a regular digital computer. At full speed, even the high-end Dell Quad-core computers in Boahen's lab cannot simulate more than one of Neurogrid's silicon neurons at a time—and the complete chip contains 65,536 neurons.

A few months ago Boahen's team received the first batch of newly fabricated Neurogrid chips. On that pivotal day everything changed. The group finally said goodbye to the pesky equations they had been compelled to run, for months and months, using unwieldy software on energy-hogging conventional computers. At last they could take the leap from simulating neurons using software to embodying those neurons in a low-power silicon chip.

When the first of the Neurogrid chips was plugged in, its silicon neurons came to life, chattering back and forth with trains of millisecond electric spikes, which were then relayed onto a computer monitor through a USB cable. Just as spikes ripple down the branching tendrils of a neuron, pulses of electricity cascaded like flash floods through the chip's transistors and

## Neurogrid's noisy processors may not be as precise, but they will allow us to accomplish miracles that digital computers struggle with.

nanowires. This activity had no more to do with equations or programming than does water tumbling down Yosemite's Bridal Veil Falls. It happened automatically, as a result of the basic physics of electricity and conductors.

In its first experiment, Boahen's team coaxed the neurons on a single chip to organize themselves into the familiar "gamma rhythm" that scientists pick up with EEG electrodes on a person's scalp. Like members of a 65,536-member chorus, each silicon neuron adjusted its spiking rate to match the 20- to 80-wave-per-second gamma tempo. The researchers recently mounted 16 Neurogrid chips on a single board to emulate 1 million neurons, connected by a tangle of 6 billion synapses. By 2011 they hope to create a second-generation Neurogrid containing 64 million silicon neurons, about equal to the total brain of a mouse.

Just a few miles down the road, at the IBM Almaden Research Center in San Jose,

that digital computers struggle with, like prancing across a crowded room on two legs or recognizing a face.

The lessons of Neurogrid may soon start to pay off in the world of conventional computing too. For decades the electronics industry has hummed along according to what is known as Moore's law: As technology progresses and circuitry shrinks, the number of transistors that can be squeezed onto a silicon chip doubles every two years or so.

So far so good, but this meteoric growth curve may be headed for a crash.

For starters, there is, again, the matter of power consumption. Heat, too, is causing headaches: As engineers pack transistors closer and closer together, the heat they generate threatens to warp the silicon wafer. And as transistors shrink to the width of just a few dozen silicon atoms, the problem of noise is increasing. The

CONTINUED ON PAGE 75

random presence or absence of a single electricity-conducting dopant atom on the silicon surface can radically change the behavior of a transistor and lead to errors, even in digital mode. Engineers are working to solve these problems, but the development of newer generations of chips is taking longer. "Transistor speeds are not increasing as quickly as they used to with Moore's law, and everyone in the field knows that," Sarpeshkar says. "The standard digital computing paradigm needs to change—and *is* changing."

As transistors shrink, the reliability of digital calculation will at some point fall off a cliff, a result of the "fundamental laws of physics," says Sarpeshkar. Many people place that statistical precipice at a transistor size of 9 nanometers, about 80 silicon atoms wide. Some engineers say that today's digital computers are already running into reliability problems. In July a man in New Hampshire bought a pack of cigarettes at a gas station, according to news reports, only to discover his bank account had been debited \$23,148,855,308,184,500. (The error was corrected, and the man's \$15 overdraft fee was refunded the next day.) We may never know whether this error arose from a single transistor in a bank's computer system accidentally flipping from a 1 to a 0, but that is exactly the kind of error that silicon-chip designers fear.

"Digital systems are prone to catastrophic errors," Sarpeshkar says. "The propensity for error is actually much greater now than it ever was before. People are very worried."

Neurally inspired electronics represent one possible solution to this problem, since they largely circumvent the

heat and energy problems and incorporate their own error-correcting algorithms. Corporate titans like Intel are working on plenty of other next-generation technologies, however. One of these, called spintronics, takes advantage of the fact that electrons spin like planets, allowing a 1 or 0 to be coded as a clockwise versus counterclockwise electron rotation.

The most important achievement of Boahen's Neurogrid, therefore, may be in re-creating not the brain's efficiency but its versatility. Terrence Sejnowski, a computational neuroscientist at the Salk Institute in La Jolla, California, believes that neural noise can contribute to human creativity.

Digital computers are deterministic: Throw the same equation at them a thousand times and they will always spit out the same answer. Throw a question at the brain and it can produce a thousand different answers, canvassed from a chorus of quirky neurons. "The evidence is overwhelming that the brain computes with probability," Sejnowski says. Wishy-washy responses may make life easier in an uncertain world where we do not know which way an errant football will bounce, or whether a growling dog will lunge. Unpredictable neurons might cause us to take a wrong turn while walking home and discover a shortcut, or to spill acid on a pewter plate and during the cleanup to discover the process of etching.

Re-creating that potential in an electronic brain will require that engineers overcome a basic impulse that is pounded into their heads from an early age. "Engineers are trained to make everything really precise," Boahen says. "But the answer doesn't have to be right. It just has to be approximate." ■