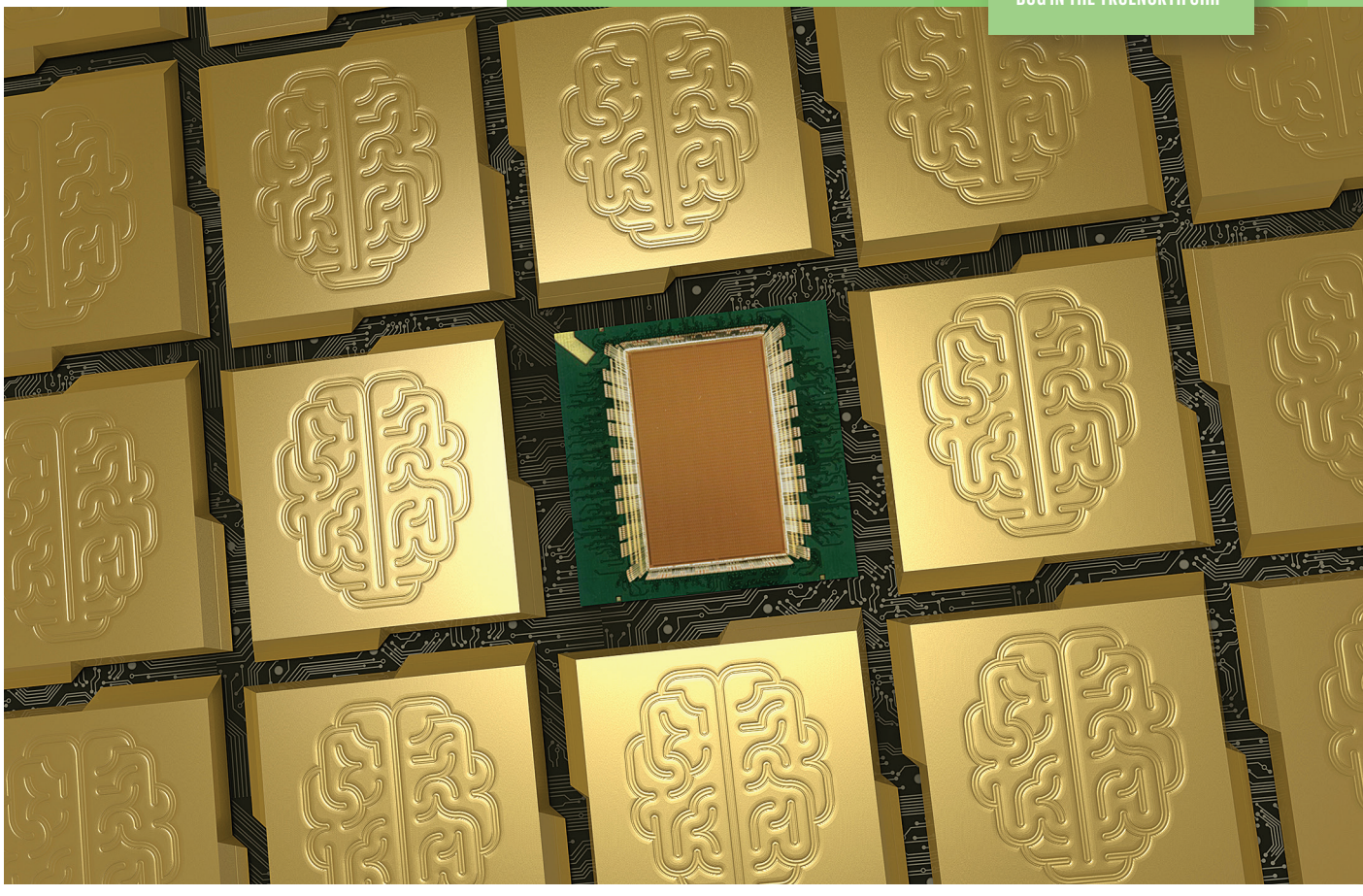
US $1,000: VALUE OF UNCLAIMED CHAMPAGNE OFFERED TO ANY IBM ENGINEER WHO COULD FIND A BUG IN THE TRUENORTH CHIP



IBM

IBM'S NEW BRAIN

The TrueNorth neuromorphic chip takes a big step toward using the human brain's architecture to reduce computing's power consumption

Neuromorphic computer chips meant to mimic the neural network architecture of biological brains have generally fallen short of their wetware counterparts in efficiency—a crucial factor that has limited practical applications for such chips. That could be changing. At a power density of just 20 milliwatts per square centimeter, IBM's new brain-inspired chip [above] comes tantalizingly close to such wetware efficiency. The hope is that it could bring brainlike intelligence to the sensors of smartphones, smart cars, and—if IBM has its way—everything else.

The latest IBM neurosynaptic computer chip, called TrueNorth, consists of 1 million programmable neurons and 256 million programmable synapses conveying signals between the digital neurons. Each of the chip's 4,096 neurosynaptic cores includes the entire computing package: memory, computation, and communication. Such architecture helps to bypass the bottleneck in traditional von Neumann computing, where program instructions and operation data cannot pass through the same route simultaneously.

"This is literally a supercomputer the size of a postage stamp, light like a feather, and low power like a hearing aid," says Dharmendra Modha, IBM »

fellow and chief scientist for brain-inspired computing at IBM Research-Almaden, in San Jose, Calif.

Such chips can emulate the human brain's ability to recognize different objects in real time; TrueNorth showed it could distinguish among pedestrians, bicyclists, cars, and trucks. IBM envisions its new chips working together with traditional computing devices as hybrid machines, providing a dose of brainlike intelligence. The chip's architecture, developed together by IBM and Cornell University, was first detailed in August in the journal *Science*.

"The impressive aspects of TrueNorth are the integration density—a million neurons on a single, admittedly very big, chip—and the very low power consumption for this many neurons," says Steve Furber, a professor of computer engineering at the University of Manchester, in England, who is behind a competing effort [see "To Build a Brain," *IEEE Spectrum*, August 2012].

With a total of 5.4 billion transistors, the computer chip is one of the largest CMOS chips ever built. Yet it uses just 70 mW in operation and has a power density about 1/10,000 that of most modern microprocessors. That brings neuromorphic engineering closer to the human brain's marvelous efficiency as a grapefruit-size organ that consumes just 20 W.

IBM minimized power usage in several ways. For one, it traded the traditional processor's clock—used to trigger and coordinate computational processes—for a more biological concept called event-driven computing. TrueNorth's digital neurons can work together asynchronously without a clock by reacting to signal spikes, which are the output of both real neurons and silicon ones.

IBM also saved on power through the design of an on-chip network that interconnects all the chip's neurosynaptic cores instead of using extra power to communicate with off-chip memory. And finally, it made the chip using a process technology meant for producing low-power mobile processors.

One brainlike feature that IBM did not mimic to reduce power consumption was to make TrueNorth's neurons analog instead of digital. The choice to go all-digital led to a number of advantages: First, IBM dodged the problem of slight differences in the manufacturing process or temperature fluctuations that have an outsize effect on analog circuits.

BIG-BRAIN CHIPS

TrueNorth IBM, DARPA SyNAPSE

Project features Low-power neuromorphic chip designed for applications in mobile sensors, cloud computing, and so on.

Chip specs 1 million neurons; 256 million synapses

Largest current configuration 16 chips: 16 million neurons; 4 billion synapses

Power density 20 mW/cm²

Final configuration 10 billion neurons; 100 trillion synapses

Neurogrid Stanford University

Project features Simulates biologically realistic brain models; team has moved on to autonomous robot applications.

Chip specs 65,536 neurons; 375,000 synapses

Largest current configuration 16 chips: 1 million neurons; 8 billion synapses

Power density 50 mW/cm²

Final configuration Info not available

HRL neuromorphic chip HRL Laboratories, DARPA SyNAPSE

Project features Emphasizes biological realism with learning capabilities on chip; has flexible, programmable routing approach.

Chip specs 576 neurons; 73,000 synapses

Largest current configuration 4 chips: 2,304 neurons; 292,000 synapses

Power density 120 mW/cm²

Final configuration Info not available

SpiNNaker University of Manchester, Human Brain Project

Project features Enables low-power, large-scale digital model of brain; helps improve models of brain diseases.

Chip specs Up to 16,000 neurons; 16 million synapses

Largest current configuration 1,152 chips: Up to 20 million neurons; 20 billion synapses

Power density 1,000 mW/cm²

Final configuration Up to 1 billion neurons; 1 trillion synapses

HiCANN Heidelberg University, Human Brain Project

Project features Simulates brain activity at 10,000 times normal speed to compress a day into 10 seconds; focuses on studying brain learning and plasticity.

Chip specs 512 neurons; 144,688 synapses

Largest current configuration 6 wafers: 1.2 million neurons; 300 million synapses

Power density 3,000 mW/cm²

Final configuration 5,000 wafers; 5 billion neurons; 1.3 trillion synapses

SOURCES: HEIDELBERG UNIVERSITY; HRL LABORATORIES; IBM RESEARCH-ALMADEN; STANFORD UNIVERSITY; UNIVERSITY OF MANCHESTER

Second, the lack of analog circuitry allowed the IBM team to dramatically shrink its hardware. Many experimental neuromorphic chips still use analog circuits that must be built using a process that on the Moore's Law curve is more than a decade behind the process used today, Furber explains. By comparison, IBM fabricated its chip using Samsung's 28-nanometer process technology—typical for manufacturing chips for today's mobile devices.

And finally, the digital design enabled TrueNorth's hardware to become functionally equivalent to its software—a factor that allowed the IBM software team to build TrueNorth applications on a simulator before the chip itself had been built.

The chip represents the culmination of a decade of Modha's personal research and almost six years of funding from the U.S. Defense Advanced Research Projects Agency (DARPA). Modha continues to lead DARPA's SyNAPSE project, a global effort that has committed more than US $100 million since 2008 to making computers that can learn. "Our long-term end goal is to build a 'brain in a box' with 100 billion synapses consuming 1 kilowatt of power," Modha says.

But the goal of that brain is different from the goals of other, similar projects. TrueNorth's digital circuits are designed with commercial applications in mind. Other projects have the goal of better understanding how the brain works.

SpiNNaker, a neural network based on digital circuits and led by Furber, uses a general-purpose parallel computing system tuned to run neurons based in software rather than hardware. The SpiNNaker team, which soon plans to simulate 100 million neurons using 100,000 chips, sacrificed the efficiency of dedicated hardware to gain the flexibility of software: "Our primary goal is to understand biology, so the flexibility is important in understanding the biological brain," Furber says. For applications-focused IBM, "the efficiency and the density of their chip is perhaps more important than the flexibility they've retained." —JEREMY HSU

BRETT HOKR

IDENTIFYING EXPLOSIVES AT A DISTANCE

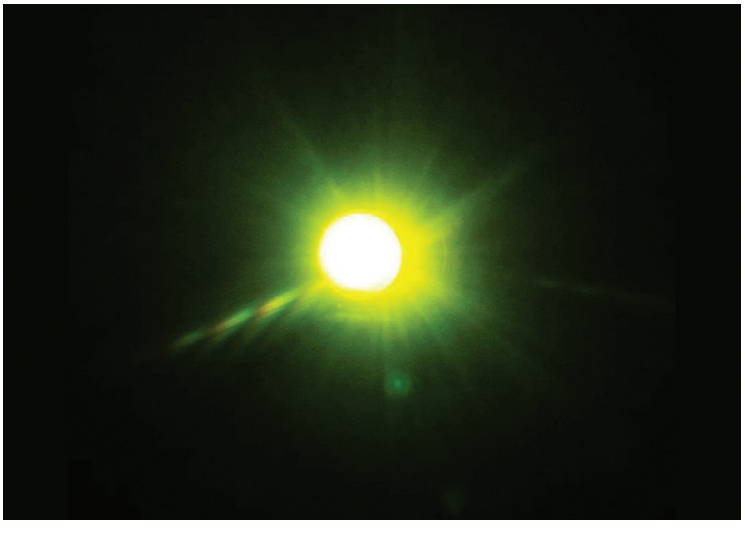
The random Raman laser is the latest technology to detect explosives and other nasty stuff from a safe vantage

Being standoffish is usually frowned upon—that is, unless what you're standing off from might be an explosive or a cloud of anthrax spores. That's why efforts have accelerated to develop standoff detection techniques that use lasers to identify chemicals and biological substances from a safe distance.

The newest entry in the field is called random Raman spectroscopy. Shine a laser beam into a loose material—say, a powder—and if the density is right, the photons will bounce around among the powder's particles until they stimulate a new laser emission. Such a random laser, as it is known, works much the same way as a more traditional laser cavity, only without mirrors.

Normally, about 1 in 10 million photons undergoes a process called spontaneous Raman scattering, in which it drops to a lower frequency determined by the particular molecule it's bouncing off. The random laser enhances this Raman scattering, producing a signal strong enough for a detector to pick up at a distance. By measuring the shift in frequency, scientists can tell the chemical makeup of the powder.

Marlan Scully and Vladislav Yakovlev of Texas A&M University, in College Station, demonstrated such a setup. Scully says they can perform spectroscopic analysis of a material at a distance of a kilometer, and that 10 kilometers should be possible. That would be useful for, say, a drone flying over an



RANDOM RAMAN LASER LIGHT: A laser beam fired at a powder causes the powder itself to become a laser, beaming out information about the material's molecular structure.