

February 28, 2011

Remapping Computer Circuitry to Avert Impending Bottlenecks

By **JOHN MARKOFF**

PALO ALTO, Calif. — [Hewlett-Packard](#) researchers have proposed a fundamental rethinking of the modern computer for the coming era of nanoelectronics — a marriage of memory and computing power that could drastically limit the energy used by computers.

Today the microprocessor is in the center of the computing universe, and information is moved, at heavy energy cost, first to be used in computation and then stored. The new approach would be to marry processing to memory to cut down transportation of data and reduce energy use.

The semiconductor industry has long warned about a set of impending bottlenecks described as “the wall,” a point in time where more than five decades of progress in continuously shrinking the size of transistors used in computation will end. If progress stops it will not only slow the rate of consumer electronics innovation, but also end the exponential increase in the speed of the world’s most powerful supercomputers — 1,000 times faster each decade.

However, in an article published in *IEEE Computer* in January, Parthasarathy Ranganathan, a Hewlett-Packard electrical engineer, offers a radical alternative to today’s computer designs that would permit new designs for consumer electronics products as well as the next generation of supercomputers, known as exascale processors.

Today, computers constantly shuttle data back and forth among faster and slower memories. The systems keep frequently used data close to the processor and then move it to slower and more permanent storage when it is no longer needed for the ongoing calculations.

In this approach, the microprocessor is in the center of the computing universe, but in terms of energy costs, moving the information, first to be computed upon and then stored, dwarfs the energy used in the actual computing operation.

Moreover, the problem is rapidly worsening because the amount of data consumed by computers is growing even more quickly than the increase in computer performance.

“What’s going to be the killer app 10 years from now?” asked Dr. Ranganathan. “It’s fairly clear it’s going to be about data; that’s not rocket science. In the future every piece of storage on the planet will come with a built-in computer.”

To distinguish the new type of computing from today’s designs, he said that systems will be based on memory chips he calls “nanostores” as distinct from today’s microprocessors. They will be hybrids, three-dimensional systems in which lower-level circuits will be based on a nanoelectronic technology called the memristor, which Hewlett-Packard is developing to store data. The nanostore chips will have a multistory design, and computing circuits made with conventional silicon will sit directly on top of the memory to process the data, with minimal energy costs.

Within seven years or so, experts estimate that one such chip might store a trillion bytes of memory (about 220 high-definition digital movies) in addition to containing 128 processors, Dr. Ranganathan wrote. If these devices become ubiquitous, it would radically reduce the amount of information that would need to be shuttled back and forth in future data processing schemes.

For years, computer architects have been saying that a big new idea in computing was needed. Indeed, as transistors have continued to shrink, rather than continuing to innovate, computer designers have simply adopted a so-called “multicore” approach, where multiple processors are added as more chip real estate became available.

The absence of a major breakthrough was referred to in a remarkable confrontation that took place two years ago during Hot Chips, an annual computer design conference held each summer at [Stanford University](#).

John L. Hennessy, the president of Stanford and a computer design expert, stood before a panel of some of the world’s best computer designers and challenged them to present one fundamentally new idea. He was effectively greeted with silence.

“What is your one big idea?” he asked the panel. “I believe that the next big idea is going to come from someone who is considerably younger than the average age of the people in this room.”

Dr. Ranganathan, who was 36 at the time, was there. He said that he took Dr. Hennessy’s criticism as an inspiration for his work and he believes that nanostore chip design is an example of the kind of big idea that has been missing.

It is not just Dr. Hennessy who has been warning about the end the era of rapidly increasing computer performance. In 2008, [Darpa](#), the Defense Advanced Research Projects Agency

assembled a panel of the nation's best supercomputer experts and asked them to think about ways in which it might be possible to reach an exascale computer — a supercomputer capable of executing one quintillion mathematical calculations in a second, about 1,000 times faster than today's fastest systems.

The panel, which was led by Peter Kogge, a [University of Notre Dame](#) supercomputer designer, came back with pessimistic conclusions. “Will the next decade see the same kind of spectacular progress as the last two did?” he wrote in the January issue of *IEEE Spectrum*. “Alas, no.” He added: “The party isn't over, but the police have arrived and the music has been turned way down.”

One reason is computing's enormous energy appetite. A 10-petaflop supercomputer — scheduled to be built by [I.B.M.](#) next year — will consume 15 megawatts of power, roughly the electricity consumed by a city of 15,000 homes. An exascale computer, built with today's microprocessors, would require 1.6 gigawatts. That would be roughly one and half times the amount of electricity produced by a nuclear power plant.

The panel did, however, support Dr. Ranganathan's memory-centric approach. It found that the energy cost of a single calculation was about 70 picojoules (a picojoule is one millionth of one millionth of a joule. The energy needed to keep a 100-watt bulb lit for an hour is 360,000 joules). However, when the energy costs of moving the data needed to do a single calculation — moving 200 bits of data in and out of memory multiple times — the real energy cost of a single calculation might be anywhere from 1,000 to 10,000 picojoules.

A range of other technologies are being explored to allow the continued growth of computing power, including ways to build electronic switches smaller than 10 nanometers — thought to be the minimum size for current chip-making techniques.

Last month, for example, researchers at Harvard and Mitre Corporation reported the development of nanoprocessor “tiles” based on electronic switches fabricated from ultrathin germanium-silicon wires.

I.B.M. researchers have been pursuing so-called phase-change memories based on the ability to use an electric current to switch a material from a crystalline to an amorphous state and back again. This technology was commercialized by Samsung last year. More recently, I.B.M. researchers have said that they are excited about the possibility of using carbon nanotubes as an a partial step to build hybrid systems that straddle the nanoelectronic and microelectronic worlds.

Veteran computer designers note that whichever technology wins, the idea of moving

computer processing closer to memory has been around for some time, and it may simply be the arrival of nanoscale electronics that finally makes the new architecture possible.

An early effort was called iRAM, in a research project at the [University of California, Berkeley](#), during the late 1990s. Today pressure for memory-oriented computing is coming both from computing challenges posed by smartphones and from the data center, said Christoforos Kozyrakis, a Stanford University computer scientist who worked on the iRAM project in graduate school.

This article has been revised to reflect the following correction:

Correction: March 2, 2011

An article on Tuesday about measures to rethink computer technology for the coming era of nanoelectronics misstated the amount of energy needed to keep a 100-watt bulb lighted for an hour. It is 360,000 joules, not eight million joules.