

NEWS CENTER :: FEATURE STORIES

April 2008

Green Computing***Exploring energy-efficient computing, from system designs to applications***

Contact: Ucilia Wang, uwang@lbl.gov

The staff of the National Energy Research Scientific Computing (NERSC) Center and the Computational Research Division (CRD) at Berkeley Lab are taking on energy-efficiency research that aims to influence the computing industry in designing and building computer and storage technologies that will benefit scientists, consumers, and the environment. In a coordinated set of projects, these researchers are exploring subjects in computer architecture, algorithms, and mass-storage-system designs to improve the energy efficiency of scientific computations.



Kathy Yelick
(Photo Roy
Kaltschmidt)

Kathy Yelick, NERSC Director, explains the importance of this work: “Power is the most important problem in computing today, not just at the high end, but from hand-held devices and laptops to data centers and computing centers like NERSC. Power density within chips has forced the entire processor industry to put multiple cores on a chip, and within centers, the total system power is a major component of cost and availability.”

She describes the NERSC and CRD energy-efficiency projects as a “multi-faceted attack” on the problem, starting with a blank slate on the architecture end and rethinking algorithms, applications, and software to make use of energy-efficient hardware. The first goal is to do more science with less energy, and the second is to enable the next generation of exascale computing systems (on the scale of quintillions, or 10^{18} , arithmetic operations per second), which will require technological breakthroughs to address the power issues at such extreme scales.

Climate modeling is the target application for the first of the four related projects. In this case researchers are taking a vertical slice through the problem space, looking at a single application domain and considering alternative algorithms as well as architectures for solving the problem. Climate modeling was selected because of its significance to science and the general public, and because it requires millions of central-processing-unit (CPU) computing hours to explore various climate scenarios and the possible impacts of changes in policy or alternate fuel sources.

The other three projects take a broader look at specific aspects of the problem, including energy-efficient computing components based on multicore technology, energy-efficient storage systems, and application characterizations that explore the ability of various key algorithms to adapt to energy-efficient hardware.

These research activities also complement the work by researchers at Berkeley Lab’s Environmental Energy Technologies Division (EETD) that looks at energy-saving technologies for data centers and other high-tech buildings.

An energy-efficient Climate Simulator

The development of multicore chips has been the computer industry’s solution to keeping power consumption in check. Two cores consume less power than a single core running twice as fast. But much of the computing industry is using a conservative approach to parallelism, starting with the relatively complex cores designed during the single-core era, then doubling the number of cores as transistor density doubles.

John Shalf, Lenny Oliker, and Michael Wehner are investigating a more radical approach, which combines a very large number of simple cores on each chip; this allows them to lower clock rate and save power while still obtaining high performance. They are also borrowing design techniques from the consumer electronics industry to tailor the chip design to the needs of applications.

Consumer electronics devices often run on battery power, and are therefore designed with power reduction as a first-order concern. The Berkeley Lab researchers anticipate this approach could achieve 100 times or more improvement in power efficiency and effective performance over business-as-usual supercomputing. The approach not only promises to be more power-efficient than the conventional path forward in high-performance computing (HPC), it also promises to be more cost-effective. And by considering massively parallel chips, the researchers are jump-starting a program to address problems that will have to be tackled for exascale computing—an increase in performance by a factor of more than 1,000, relative to the fastest HPC systems available today.

continued



From left, John Shalf, Lenny Oliker, and Michael Wehner (Photos Roy Kaltschmidt)

The Climate Simulator team is working with Tensilica, an embedded-processor design firm, and with NERSC user David Randall, a professor in the Atmospheric Sciences Department at Colorado State University. The use of embedded processors has exposed orders of magnitude more parallelism in applications; by exploring application and algorithm issues in conjunction with hardware design, the team believes they can design a more effective system. Randall's climate-modeling code, developed under DOE's SciDAC program, represents a new breed of such codes, capable of expressing enough parallelism to run kilometer-scale simulations a thousand times faster than real time on the machines envisioned by the Climate Simulator research team. Employing massive numbers of simple processors and applications like Randall's, designed for extreme scaling, will enable a new generation of HPC systems and keep energy use in check.

"We want to find compelling solutions to scientific problems that need petascale machines," Shalf says. "The use of these power efficient cores will help us achieve those goals."

Many cores for fewer watts

The development of multicore chips represents the most significant shift in microprocessor engineering in several decades, and it opens up opportunities for exploring innovative designs for high-performance computers.

Jonathan Carter, head of the User Services Group at NERSC, is leading the project to explore a wide range of multicore computer architectures, examining how efficiently those systems can perform challenging scientific computations.

Rather than focusing on a single design, as in the Climate Simulator project, Carter's effort will look at several competing multicore designs and evaluate their effectiveness for scientific computations. The computing industry is moving rapidly to provide a variety of architectures that exploit multiple cores, but the designs vary significantly, and trade-offs in performance, power and programmability are not well understood. They include heterogeneous designs, such as the Cell processor developed by IBM, Sony, and Toshiba; graphics processing units (GPUs); and processors for the embedded market. There are also homogeneous designs, such as dual and quad core chips by Intel, IBM, and AMD that replicate conventional microprocessors on a single chip, and Sun's rather different approach based on large-scale multithreading. In many cases, multicore technologies offer higher absolute performance and more energy-efficient computation.



Jonathan Carter (Photo Roy Kaltschmidt)

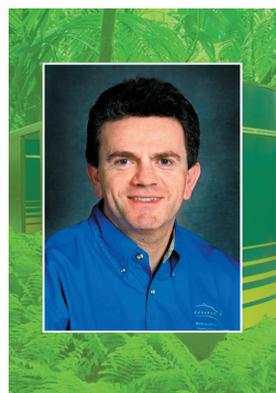
"This project provides a breadth of architecture coverage to our whole ultra-efficient research thrust—we want to identify candidate algorithms that map well to multicore technologies and document the steps needed to re-engineer programs, to take advantage of these architectures," Carter says. "In addition, perhaps there are design elements in multicore chips that we can influence to help design a better high-performance system."

Equal-opportunity energy efficiency

The goal of the overall energy efficiency effort is to maximize the amount of science that can be done for a given investment in hardware and energy. A clever algorithm may allow a scientist to save both energy and time by reaching a solution quickly. Thus energy-efficiency must be viewed from the application perspective rather than solely as a hardware problem.

Erich Strohmaier is addressing the algorithmic challenges of gaining energy efficiency through massive parallelism. Rather than using a single-application focus, as in the Climate Simulator, he is developing a test bed for benchmarking some of the key algorithms across a wide range of scientific applications. This will allow energy-efficient designs to be evaluated on a broad range of applications and permits hardware and algorithm codesign.

Strohmaier will take advantage of computational science expertise from across Berkeley Lab and the U.C. Berkeley campus and will build on a categorization of algorithms known as the "dwarfs," which was developed as a campus/Lab collaboration. The name refers to an observation that there are seven computational methods ("the seven dwarfs") that dominate much of scientific computation. The list has roughly doubled to cover applications that run on embedded devices, databases, and personal computers, but it still provides a compact way of exploring a broad application space.



Erich Strohmaier (Photo Roy Kaltschmidt)

Strohmaier's goal is to devise ways of using a set of algorithms to gauge the performance of systems from personal computers to high-performance systems. History has shown the importance of understanding performance of commodity hardware, since the commodity market will determine what processors are available as a building block for HPC systems.

continued

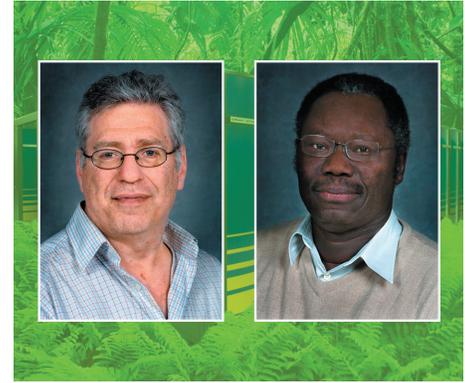
Saving data while saving energy

Led by CRD scientists Ekow Otoo and Doron Rotem, the “Energy-Smart, Disk-Based Mass Storage System” project sets out to investigate energy-efficient disk storage configurations that also provide quick access to massive amounts of data.

Today’s storage systems in data centers use thousands of continuously spinning disk drives. These disk drives and the necessary cooling components use a substantial fraction of the total energy consumed by the data center. As the need for reliable long-term storage of data grows, so will the associated energy costs.

Otoo and Rotem have set out to explore new configurations that divide the disks into active and passive groups. The active group contains continuously spinning disks and acts as a cache for most frequently accessed data. The disks in the passive group would power down after a period of inactivity. Besides looking at optimal disk configurations and file placement algorithms, the researchers will also develop simulation models for analyzing energy use.

Pursuing energy-efficient computing research at Berkeley Lab makes sense not only to save energy but also to shift NERSC resources away from energy costs and towards systems and services that directly benefit the scientific computing. With over 3,000 users and an insatiable demand for NERSC computing and storage facilities, results from these projects could indirectly benefit scientists in cosmology, climate, life sciences, material sciences, and all others disciplines that rely heavily on computing.



*Doron Rotem and Ekow Otoo
(Photos Roy Kaltschmidt)*



Additional information

More about the challenges of climate modeling may be found at <http://www.lbl.gov/CS/Archive/news020808.html>.

A slide presentation on “Climate Modeling at the Petaflop Scale” is at http://vis.lbl.gov/~jshalf/SIAM_CSE07/LBL_SIAMCSE2007.pdf.

More about parallel computing and multicore chips may be found at <http://www.lbl.gov/CS/Archive/news030207.html>.

Learn about research collaboration among Berkeley Lab scientists and engineers at <http://lbl.gov/cs/html/energy%20efficient%20computing.html>.