

UC San Diego's GreenLight Project to Improve Energy Efficiency of Computing

July 28, 2008

Doug Ramsey

The information technology industry consumes as much energy and has roughly the same carbon "footprint" as the airline industry. Now scientists and engineers at the University of California, San Diego are building an instrument to test the energy efficiency of computing systems under real-world conditions - with the ultimate goal of getting computer designers and users in the scientific community to re-think the way they do their jobs.

The National Science Foundation will provide \$2 million over three years from its Major Research Instrumentation program for UC San Diego's GreenLight project. An additional \$600,000 in matching funds will come from the UCSD division of the California Institute for Telecommunications and Information Technology (Calit2) and the university's Administrative Computing and Telecommunications (ACT) group.

The GreenLight project gets its name from its plan to connect scientists and their labs to more energy-efficient 'green' computer processing and storage systems using photonics - *light* over optical fiber. "As a leader in the field of information technology, UC San Diego has a responsibility to reduce the amount of energy required to run scientific computing systems," said UCSD Chancellor Marye Anne Fox. "Project GreenLight will train a new generation of energy-aware scientists, and it will produce energy consumption data to help investigators throughout the research community make informed choices about energy-efficient IT infrastructure."

The rapid growth in highly data-intensive scientific research has fueled an explosion in computing facilities and demand for electricity to power them. Energy usage per compute server rack is growing from approximately 2 kilowatts (KW) per rack in 2000 to an estimated 30 KW per rack in 2010. Every dollar spent on power for IT equipment requires that another dollar be spent on cooling - equivalent to double the cost of the hardware itself over three years. As a result, cooling and power issues are now becoming a major factor in system design. "If we are going to continue to allow ourselves the benefits of advances in computing, we need to understand power and cooling requirements much better," said Thomas A. DeFanti, a research scientist at Calit2 and principal investigator of the GreenLight project. "Scientists from all domains will choose more efficient systems as they invest in new cyberinfrastructure, and we expect that GreenLight will give them the data they need. Some scientific computing jobs need more powerful processors, some can do with less memory, some can use specialized processors: these are important requirements to understand so the optimally configured cluster can be chosen and scheduled through virtualization techniques each and every time."

To eliminate the need for air conditioning, each Sun MD's closed-loop water-cooling system uses built-in heat exchanges between equipment racks to channel air flow. This allows the unit to cool 25 kilowatts per rack, roughly five times the cooling capacity of typical datacenters. The industry-standard racks can also be placed close together, further reducing the structure's overall eco-footprint and increasing energy efficiency by eliminating dead space.

"Using the Sun Modular Datacenter as a core technology and making all measurements available as open data will form a unique, Internet-accessible resource that will have a dramatic impact on academic, government and private-sector computing," said Emil J. Sarpa, Director of External Research at Sun Microsystems, Inc. "By placing experimental hardware configurations alongside traditional rack-mounted servers and then running a

variety of computational loads on this infrastructure, GreenLight will enable a new level of insight and inference about real power consumption and energy savings."

The GreenLight Instrument will use sensors in the controlled datacenter environment to measure temperature (at 40 points in the air stream), humidity, energy consumption and other variables, in addition to monitoring the internal measurements of the servers. Researchers hope to use the data to find ways to minimize the power needed to run computers, to make use of novel cooling sources, and to develop software that automates the optimizing of power strategies for each given computing process.

The facility will provide computing and storage services to large-scale projects in five diverse scientific areas: metagenomics; ocean observing; microscopy; bioinformatics; and digital media. Researchers from these fields will be able to carry out quantitative explorations into energy-efficient cyberinfrastructure in a real-world environment. "We will be running full-scale applications on full-scale computing platforms, so we will be able to draw conclusions about the comparative amount of energy that is consumed by one workload versus another," said Calit2 Director Larry Smarr, co-principal investigator on GreenLight and PI of the Community Cyberinfrastructure for Advanced Marine Microbial Ecology Research and Analysis (CAMERA) metagenomics project. "We expect that this new approach will re-define the fundamentals of computer systems engineering and accelerate adoption of a transformative concept for the computer industry - green cyberinfrastructure."

Other co-PIs on the project include UCSD Center for Networked Systems Director Amin Vahdat; Philip Papadopoulos of the San Diego Supercomputer Center (SDSC); as well as Computer Science and Engineering professor Ingolf Krueger. Other faculty members on the project include CSE faculty Pavel Pevzner, Falko Kuester, Tajana Simunic Rosing and Rajesh Gupta, as well as Biology professor Steve Briggs, and Electrical and Computer Engineering professor Bhaskar Rao.

Some of the research groups participating in GreenLight will re-locate servers, switches, computer clusters and related equipment to be deployed inside the first Sun Modular Datacenter. The scientists will continue to operate their equipment virtually and remotely over UCSD's high-performance network, just as if the computers were still in their labs. Indeed, many researchers may not even know where the computers are located.

"If the networking is transparent, the scientists won't care where the computers are as long as the data gets from their devices and back to their screens without delay," said DeFanti. "The full-scale GreenLight Instrument will measure, monitor and make publicly available real-time sensor outputs using a service-oriented architecture methodology, empowering researchers anywhere to study the energy cost of at-scale scientific computing."

Although the IT industry has begun to develop strategies for 'greening' major corporate data centers, most of the cyberinfrastructure on a university campus involves a complex network of *ad hoc* and suboptimal energy environments, with clusters placed in small departmental facilities.

According to DeFanti, the project decided to build the GreenLight Instrument around the Sun Modular Datacenter because, "it's the fastest way to construct a controlled experimental facility for energy research purposes." The modular structure also means the GreenLight Instrument can be cloned - unlike bricks-and-mortar computer rooms that cannot be ordered through purchasing.

The GreenLight Instrument will enable an experienced team of computer-science researchers to make deep and quantitative explorations in advanced computer technologies, including graphics processors, solid-state disks, photonic networking, and field-programmable gate arrays (FPGAs). Jacobs School of Engineering computer science professor Rajesh Gupta and his team will explore alternative computing fabrics from array processors to custom FPGAs and their respective models of computation to devise architectural strategies for efficient computing systems.

"Computing today is characterized by a very large variation in the amount of effective work delivered per watt, depending upon the choice of the architecture and organization of functional blocks," said Gupta. "The project

seeks to discover fundamental limits of computing efficiency and device organizing principles that will enable future system builders to architect machines that are orders-of-magnitude more efficient modern-day machines, from embedded systems to high-performance supercomputers."

The computing and systems research will yield new quantitative data to support engineering judgments on comparative "computational work per watt" across full-scale applications running on full-scale computing platforms.

Researchers in GreenLight advocated for full-scale computational and storage configurations in order to entice faculty and graduate students to run their computational work through GreenLight. "We are asking 25 faculty and researchers to work with us with no money for students, summer salary or even system administrative support," said DeFanti. "We wouldn't be able to retain them and get usable data if we were only offering toy-scale computation."

Computer scientists will use GreenLight to study topics ranging from virtualization for optimizing resource utilization, to power and thermal management. Jim Hollan, a professor of cognitive science at UCSD, will study how access to energy costs may influence the behavior of scientists in using shared computational resources - especially when the energy use is visible to the wider community.

Rather than give scientists physical access to the GreenLight Instrument, OptIPortal tiled display systems will serve as visual termination points - allowing researchers to "see" inside the instrument. Users will also be able to query and visualize all sensor data in real time and correlate it interactively and collaboratively in this immersive, multi-user environment.

Once a virtual environment of the system has been created, scientists will be able to walk into a 360-degree virtual reality version in Calit2's StarCAVE. Users will be able to zoom into the racks of clusters as well as see and hear the power and heat, from whole clusters of computers down to the smallest instrumented components, such as computer processing and graphics processing chips.

The GreenLight project aims to reach out to other campuses, including Indiana University. IU professor of computer science, informatics and physics and director of the Community Grids Laboratory Geoffrey Fox will organize a workshop in the next year, under the auspices of MSI-CIEC - the Minority Serving Institutions Cyberinfrastructure Empowerment Coalition - which reaches over 335 minority-serving institutions. "The workshop will feature lectures on the frontiers of green cyberinfrastructure, the hardware, software and middleware of the GreenLight Instrument," said Fox, "and participants will engage in working groups to provide feedback on the project's development and outreach plans going forward."

In addition to Sun Microsystems, other industry partners in GreenLight include San Jose, CA-based Cisco Systems, Inc. Cisco is contributing use of its CiscoWave network on the National LambdaRail (NLR) on an as-available basis.

Media Contact: Doug Ramsey, 858-822-5825

