The National Energy Research Scientific Computing Center:
Forty Years of Supercomputing Leadership

Harvey J. Wasserman and Richard A. Gerber | National Energy Research Scientific Computing Center, Lawrence Berkeley National Laboratory

Back in 1974, the psychedelic aura of the "love generation" was about to yield to an exciting new fad called disco; *Blazing Saddles* and *The Towering Inferno* were blockbuster movie box office hits; and the oil embargo of the previous year had led to gasoline shortages and skyrocketing prices that shocked the American public.

But long lines at gas stations weren't the only energy-related story of this period: 1974 also saw the beginning of a special high-performance computing tale that continues to this day.

The need for alternative energy sources led to renewed interest in magnetic fusion energy, and the Department of Energy (DOE) funded a computer center in 1974 dedicated to simulating the behavior of plasma in a fusion reactor. That center—the Controlled Thermonuclear Research Computer Center (CTRCC), established at the Lawrence Livermore National Laboratory with an

 1521-9615/15/$31.00 © 2015 IEEE Copublished by the IEEE CS and the AIP

almost obsolete supercomputer "handed down" from defense research—later became the National Magnetic Fusion Energy Computer Center, and finally the National Energy Research Scientific Computing Center (NERSC; www.nersc.gov).

NERSC Today

Today, NERSC is located at the Lawrence Berkeley National Laboratory (Berkeley Lab), and its mission has expanded to scientific discovery in areas ranging from climate science to high-energy physics. Still, the center remains committed to supporting research aligned with DOE's mission to help ensure America's security and prosperity by addressing challenges in efficient energy use, reliable energy sources, improved environmental quality, and fundamental understanding of matter and energy.

But, as the first unclassified supercomputer center, NERSC was, and has become, the model for those that followed.

NERSC is one of several national user facilities provided by the DOE's Office of Science (SC), which is the nation's largest supporter of basic research in the physical sciences. As the principal provider of high-performance computing services to SC's programs in magnetic fusion energy, high-energy physics, nuclear physics, basic energy sciences, biological and environmental research, and advanced scientific computing research (ASCR), NERSC has grown to serve more than 4,000 researchers working on some 700 projects. Users routinely report about 1,500 refereed scientific publications each year derived from research using NERSC, and the center can count at least three Nobel Prize winners among its users.

Since its inception, NERSC's central mission has nevertheless remained consistent: to accelerate the pace of scientific discovery by providing high-performance computing, data, communications, information, and related services. Recently, scientific discovery has been characterized as comprising four key paradigms, with both simulation and data analysis being represented alongside the traditional pillars of theory and experiment (http://research.microsoft.com/en-us/collaboration/fourthparadigm). NERSC's role in accelerating discovery has been evident in both simulation and data analysis for many years.

Fundamental to NERSC's mission is enabling computational science of scale, in which large, interdisciplinary teams of scientists attack fundamental problems in science and engineering that require massive calculations or massive data and have broad scientific and often, economic, impact. Thus, resources available to NERSC customers have come a long way from that borrowed fusion energy supercomputer, a CDC 6600. Today, NERSC runs petascale systems and is actively working with DOE, industry, and other researchers to extend its capabilities into the exascale realm.

In This Issue

In celebration of NERSC's 40th anniversary in 2014, we present this special issue of *CiSE* with articles selected to describe both NERSC's past and give a glimpse into its future.

Two retrospective articles provide descriptions of key science enabled by NERSC resources over the past 15 years. Although research in both of these areas, gyrokinetic plasma simulation and Big Bang cosmology, continues today as robustly and with as much importance as ever, it's instructive to reflect on the unique way in which both the science and NERSC co-evolved so as to ensure maximum impact of the science and the facility.

NERSC strategic objectives going forward are in two key areas: meeting the ever-growing computing and data needs of our users by providing usable exascale computing and storage systems, and increasing the productivity, usability, and impact of other DOE user facilities and observational platforms by providing comprehensive data systems and services to store, analyze, manage, and share data.

Two articles in this special issue will highlight recent work in these areas. Given the breadth of NERSC's customer workload in terms of science and resource requirements, no single work can provide a truly comprehensive treatment of either of these topics, so in both articles, a case study approach is used to survey representative samples with a view toward teasing out the salient lessons learned for a wider community.

The first of these articles presents a look at the challenges of transitioning codes to manycore processor architectures, using an Intel Knights Corner (Intel Xeon Phi) product as an example. Energy-efficient computing is a cornerstone technology of exascale computing and represents the only way of continuing NERSC's historic performance growth in response to science needs. Not all SC computing problems require computing at the exascale level but all SC computing problems must be aware of the myriad issues affecting programming models, vastly increased levels of on-node parallelism, and deeper memory hierarchies. NERSC's next big supercomputer, named Cori, after American Nobel Laureate Gerty Cori, will be based on a new and enhanced Intel manycore product called Knights Landing. As part of its "Application Readiness" initiative, the NERSC staff has been examining what will be required from users to fully extract the science potential of emerging computing systems, and this article captures much of the knowledge gained from this effort.

Although sometimes viewed as being at odds, data analysis and simulation are actually closely intertwined.[1] Although it's becoming increasingly evident that extreme data science is also a driver for exascale computing technologies, much more capable data-intensive supercomputers and storage systems must be accompanied by a rich suite of services to facilitate extreme data science. The second forward-looking article in this special issue examines one technology, SciDB, being deployed at NERSC (and elsewhere) that's designed to provide a unified, scalable, easy-to-use, extendable system to tackle extreme data challenges.

The Future of NERSC

The demands from larger and more detailed simulations, massive numbers of simulations, and the explosion in the size and number of experimental datasets mean there's no end in sight to the need for NERSC resources. As we look forward to the center's second 40 years, we're excited about the multitude of challenges and opportunities that await. The center will move into a new building, the highly energy-efficient, state-of-the-art Computational Research and Theory (CRT) facility on the main Berkeley Lab site that will provide the DOE research community with greatly expanded resources in support of their work. The Cori system, with its exciting processor and storage technologies, will vastly extend the limits of both simulation and data analysis. New and continuing deep collaborations with experimentalists will assure ever-increasing scientific impact of Office of Science national scientific user facilities.

So while the legacy of disco might be subject to some debate, the future impact of NERSC remains secure.

Acknowledgments

The work described in the four articles comprising this special issue used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the US Department of Energy under contract number DE-AC02-05CH11231. We're grateful to our DOE SC sponsors for their continued support, to our user customers for the science they produce using NERSC resources, to the many coauthors of the four articles in this special issue, and to the NERSC management and staff whose dedication makes NERSC such a successful high-performance computing center. Part of this introduction is based an article, "NERSC History, Powering Scientific Discovery Since 1974," by Jon Bashor, Lawrence Berkeley Lab Computing Sciences, available on the NERSC website. ■

Reference

1. S. Dosanjh et al., "Extreme Data Science at the National Energy Research Scientific Computing (NERSC) Center," *Advances in Parallel Computing*, vol. 25, 2014, pp. 3–18.

Harvey J. Wasserman is a computer systems engineer at Lawrence Berkeley National Laboratory. He has been involved in supercomputer workload characterization, benchmarking, system evaluation, and user support at NERSC and at Los Alamos National Laboratory for over 30 years. Wasserman has a PhD in chemistry from SUNY, Buffalo. Contact him at hjwasserman@lbl.gov.

Richard A. Gerber is Senior Science Advisor and User Services Group Lead at NERSC. His research interests include understanding how scientists use supercomputers and big data systems; monitoring application performance on HPC systems; and, in his spare time, dynamics and interactions of galaxies. Gerber has a PhD in physics, specializing in computational astrophysics, from the University of Illinois at Urbana-Champaign. Contact him at ragerber@lbl.gov.