

United States House of Representatives
Committee on Science, Space and Technology
Subcommittee on Energy

Hearing on America's Next Generation Supercomputer: The Exascale Challenge
May 22, 2013

Testimony

Roscoe C. Giles, Ph.D.
Chair, Advanced Scientific Computing Advisory Committee (ASCAC)
Office of Science, U.S. Department of Energy

Professor, Department of Electrical and Computer Engineering, Boston University
Boston University
roscoe@bu.edu

Good morning Madame Chair and members of the Subcommittee. Thank you for inviting me to testify today.

Background

I am Roscoe C. Giles, a Professor in the Department of Electrical and Computer Engineering at Boston University. I have been involved for many years in leadership roles in computational science and high performance computing and in computational science education. The primary perspective that I offer to you today comes from my role as a long time member and current chair of the U.S. Department of Energy (DOE) Office of Science Advanced Scientific Computing Advisory Committee (ASCAC).

The bulk of the written materials I wish to submit to the Subcommittee are in the form of published [ASCAC reports](#) listed on the accompanying citation page.

Two reports that this discussion particularly focuses on are:

- (1) "[Synergistic Challenges in Data-Intensive Science and Exascale Computing](#)", March 2013, and
- (2) "[The ASCAC Facilities Statement](#)," March 21st, 2013.

In addition, our Fall 2010 report on "[The Challenges and Opportunities of Exascale Computing](#)," is also especially relevant.

The following remarks are presented as responses the questions provided by the committee staff.

Q1: Summarize the work of the Advanced Scientific Computing Advisory Committee (ASCAC) in reviewing and advising DOE's Advanced Scientific Computing Research (ASCR) program. Specifically, please discuss significant issues and priorities confronting ASCR.

ASCAC was first constituted in 1999 and is chartered under the Federal Advisory Committee Act (FACA). ASCAC members are appointed by the DOE Undersecretary for Science and are experts in their fields. We report to the Director of the Office of Science in response to formal charges, and we are not paid for our work on ASCAC. Our purpose is to provide a useful external, community perspective on the impact, significance, and directions of ASCR efforts. Our committee meetings and reports are public.

Charges to ASCAC range from reviews of program management and effectiveness - for example we supervise regular Committees of Visitors for ASCR research program areas - to major reviews of strategic areas of emphasis and plans, such as the reports on the Exascale Computing and Data Intensive Science that we focus on today. Charges are generally handled by Subcommittees consisting of a few ASCAC members together with selected external experts chosen for their expertise in the specific area of the report.

ASCR Programs

ASCR's mission is "*...to discover, develop, and deploy computational and networking capabilities to analyze, model, simulate, and predict complex phenomena important to the Department of Energy (DOE).*"

In pursuit of this mission, ASCR has programs and investments that include: (1) Major computing and networking facilities for meeting the needs of current DOE Science programs; (2) Leadership Computing Facilities with unique high-end capabilities made available both to DOE and to all of the nation, including industry; (3) Applied Mathematics Research whose results provide the framework for future applications and systems; (4) Computer Science system and software research whose results both enable applications of current systems and chart the direction for future systems.

As part of its efforts, ASCR investments have built human expertise ('human capital,' if you will) in the scientific and technical staff at the labs and through attention to integrating the next generation of computational science leaders into the DOE programs and facilities.

ASCR's success is ultimately reflected in the scientific productivity, deepening insights, results, and technologies of DOE Science. ASCR has pioneered program partnerships to achieve these goals. The Scientific Discovery through Advanced Computing (SciDAC) programs linked ASCR computing experts with other DOE

scientists and programs through effective domain applications. INCITE (Innovative and Novel Computational Impact on Theory and Experiment) has made resources at leadership computing facilities available competitively to DOE and other scientists and engineers, including industry.

A recent IDC Interim Report on a Survey¹ undertaken for the EU, notes that the SciDAC and INCITE programs are the top two mentioned by respondents as models for the world's most successful High Performance Computing (HPC) programs. This is particularly noteworthy since the majority of respondents in this survey are European scientists who aim to compete with the US for leadership in High Performance Computing.

In its reviews of INCITE and the Research Programs in Mathematics and Computer Science and Networking, ASCAC has been impressed with the quality, significance and effectiveness of ASCR's work. A recent ASCAC review of the Computational Science Graduate Fellows Program found it to be "exceptional" in the quality of its participants and in its management, and "...unique in its focus on Computational Science."

ASCR Facilities Enable Science Accomplishments

ASCR facilities include the Leadership Computing Facility at Argonne National Laboratory (ALCF) and Oak Ridge National Laboratory (OLCF), National Energy Scientific Computing Center (NERSC) at Lawrence Berkeley National Laboratory, and the Energy Research Network (ESNET). The list of accomplishments of these facilities, just in the recent years, gives us a window into what can be achieved scientifically through advanced computing.

LCF

The Leadership Computing Facility at Argonne National Laboratory and Oak Ridge National Laboratory was established in 2004 with a mission to provide the world's most advanced computational resources to the open science community. Through the INCITE and Advanced Leadership Computing Challenge (ALCC) programs, computational resources are provided to scientists from the research community in industry, academia, and national laboratories. The LCF is a national resource as well as a DOE resource. Scientists and engineers using the LCF have achieved numerous wide-ranging research accomplishments and technological innovations. The full breadth and impact of science productivity cannot be adequately discussed in a few paragraphs. But as one measure, more than 500 peer-reviewed research articles based directly upon LCF projects were published in 2012 alone, including several in high-impact journals such as *Science*, *Nature*, and *The Proceedings of the National Academy of Sciences* (PNAS).

¹ "D2 Interim Report: Development of a Supercomputing Strategy in Europe (SMART 2009/ 0055, Contract Number 2009/ S99-142914)", IDC, May 7, 2010

Specific high-impact scientific achievements from the past year include some of the largest nuclear structure studies ever performed and the world's largest high-resolution cosmology simulation, modeling over one trillion particles. At OLCF, exploration of the nuclear landscape carried out by INCITE researchers and their theoretical prediction of isotopes was featured in *Nature* in 2012. They answered one of the fundamental questions of nuclear structure physics. By exploring the limits of nuclear stability, demonstrating there are approximately 7,000 possible combinations of protons and neutrons allowed in bound nuclei with up to 120 protons, the researchers provided fundamental insight into theoretical constraints on isotopes.

At ALCF, a trillion-particle Outer Rim cosmology simulation, which was 15 times larger than the largest simulation previously carried out in the US, is providing invaluable results for ongoing and upcoming DOE-funded sky surveys, such as the Dark Energy Survey and the Large Synoptic Survey Telescope. In addition, this simulation is setting new standards for computational performance, achieving 69.2% of peak performance (13.94 Pflop/s) on Sequoia. Researchers working on a climate end station INCITE project were the first to definitively show carbon dioxide as the major driver of planetary warming by producing a more comprehensive global paleoclimate proxy dataset coupled with the simulation of the Earth system's energy transport mechanisms during the last deglaciation, and in a separate modeling projection quantified the mechanisms driving sea-level rise.

NERSC

National Energy Scientific Computing Center is the main computing resource that supports production scientific computing within DOE Office of Science. In 2012 over 600 projects benefited from the high performance computing environment at NERSC, including fusion energy, materials science, lattice QCD, chemistry, climate science, earth science, astrophysics, biosciences, accelerator science, combustion, nuclear physics, engineering, mathematics and computer science. The impact of NERSC is highly visible – over 1500 peer reviewed journal publications are produced each year. I note here only a few of the many widely recognized breakthroughs: Nobel Prize awards in 2007 and 2012 from scientific simulations at NERSC; Supernova 2011fe was caught within hours of its explosion in 2011, and telescopes around the world were rapidly redirected to it; and the new approach developed by MIT researchers to desalinate sea water using sheets of graphene, a one-atom-thick form of the element carbon. The latter was *Smithsonian Magazine's* fifth “Surprising Scientific Milestone of 2012.”

ASCR High End Computing Development Activities

After an extraordinary series of community workshops, engaging DOE applications scientists and engineers, computer scientists, mathematicians, industry representatives and academics, ASCR developed the foundations for the “exascale” initiative: to build the technology – hardware, software, applications frameworks—that would allow for a machine to deliver a computational capability of 10^{18} operations per second and, along the way, enable remarkable advances at all

intermediate scales of computing of relevance to science and industry. The term “extreme-scale computing” is used more broadly to refer to leadership systems across these scales, ranging from embedded processors to leadership facilities that will host exascale computers. A recent report from the National Research Council entitled “The Future of Computing Performance: Game Over or Next Level?” highlighted the importance of leadership in extreme-scale computing for US competitiveness.

In 2009, ASCAC was charged with reviewing ASCR’s body of work on exascale computing. Dr. Robert Rosner of the University of Chicago led our subcommittee on this charge. We delivered our review report, “[The Opportunities and Challenges of Exascale Computing](#),” in fall of 2010. We found the case for exascale compelling and recommended that **“DoE should proceed expeditiously with an exascale initiative so that it continues to lead in using extreme scale computing to meet important national needs.”**

And, indeed, ASCR has been working in partnership with industry, the lab personnel, and the community to move us along the path to exascale. Some program elements have included:

- Establishment of Co-Design centers to exploit a key element of effective extreme computing applications – the guided interplay of application/hardware/software in the design of systems
- Computer Science Research: X-Stack software to develop tools for extreme scale systems, Advanced Architectures
- Applied Mathematics Research: Uncertainty Quantification, Extreme Scale Algorithms
- Prototypes: (joint with NNSA) FastForward, Design Forward
- Community: Exascale Research Conferences

ASCR Priorities and Significant Issues

I believe that ASCR has made very good progress – excellent progress under the circumstances – on advancing the frontier of high end computing.

ASCR has engaged the research programs, labs, and community effectively in working on extreme-scale computing issues that must be resolved on the path to exascale, all without explicit funding for traveling this path.

A risk is that, while focusing efforts toward extreme scale, budget constraints may force ASCR to underfund support for synergistic research activities that may be needed to support and sustain the long-term viability of our leadership in computational science. ***History has shown that success in computational science productivity depends on all the elements of the ecosystem that ASCR has nurtured over the years: leading edge hardware, software and tools for applications, mathematical methods, and professionals and students working in these areas.***

While funding the acquisition of the most advanced hardware is not, in itself, a guarantee of success, underfunding those developments and acquisitions is a likely

guarantor of failure. One of our competitive advantages over the rest of the world has been our ability to provide highly effective user support both for using the machines and for applications development. This allows us to lead in science whether or not at one moment or another we are number one in hardware. But this is a fragile lead, and if we fall consistently behind in new technologies, this lead will disappear.

I am very hopeful that the legislation being discussed today will address these issues and make it possible for ASCR, by continuing to carefully husband resources and prioritize its efforts, to be successful in the exascale initiative while continuing to nurture its productive system of research, support, and partnerships.

Q2: Summarize key findings and recommendations from recent ASCAC reports, including the "ASCAC Facilities Statement" and the report "Synergistic Challenges in Data-Intensive Science and Exascale Computing. "

ASCAC Facilities Statement

The "ASCAC Facilities Statement" is a letter report, drafted by a subcommittee I chaired, which was asked to comment on the decadal facilities plans being prepared by the ASCR office. We were specifically not asked to prioritize facilities, but rather to assess their impact and readiness. The full letter report is available online.

Here are excerpts from the report:

Facilities

ASCR computing, networking, storage, software and applications support are key underpinnings of the activities of the Office of Science. Although ASCR is renowned for fielding some of the most powerful computing available at any given time, the real impact of ASCR facilities is realized in the successes of the research programs of all the offices in SC—basic energy sciences, biological & environment science, fusion energy science, high energy physics, nuclear physics and ASCR itself. It is important to consider the proper balance between these underpinnings, and realize it changes over time. Identifying application and technology drivers is crucial, and ASCR facilities staff has considerable experience and expertise in this area.

Because of the unique and rapidly changing role of computing and data in all areas of science, we believe that investment in this area is critical to the overall mission of the Office of Science and to DOE and the nation. The facilities we comment on here represent the minimum necessary to support the needs of the Office of Science, DOE and the nation and do not explicitly incorporate a full scale commitment to exascale computing development and deployment as envisioned in ASCAC's Fall 2010 report.

The three major facilities brought to our attention by the ASCR AD reflect a balanced roadmap for upgrading existing ASCR computing and networking capabilities to meet the expected and emerging needs of DOE and the nation’s scientists:

1. Upgrading the production computing facility at NERSC, which supports more than 600 projects sponsored by the DOE Office of Science Program Offices.
2. Upgrading the Leadership Computing Facility (LCF) at ANL and ORNL, which advances the frontier of computational science and discovery for the nation and the world.
3. Increasing the network bandwidth of ESnet, which enables the large data flows needed for DOE computing, experiments and analysis in an expanding national and international collaborative environment.

In addition, to meet the emerging critical need to support and develop large-scale data science, we propose adding a fourth facility to the portfolio:

4. A Virtual Data Facility (VDF). This multi-site facility would add the data storage and analysis resources to the existing ASCR facilities to address the data challenges to all SC programs, and is being considered by the ASCR facilities leaders.

The table following summarizes our findings; a further discussion of each facility and justification for their categorization are provided in subsequent sections. Given the rapid pace of technology change, we feel it necessary to distinguish near-term (within 5 years) and far-term (towards the end of the 10-year timeframe covered by this charge) readiness levels, as described in the table.

Facility	Impact	Readiness	
		(2014-2017)	(2018-2024)
NERSC	A	A	B
LCF	A	A	B
ESnet	A	A	B+
VDF	A	B+	C

(The classifications used are those described in the charge letter,
 Impact: A=“absolutely central”, B=“important”, C=“lower priority”, D=“don’t know enough yet”.
 Readiness: A=“ready to initiate construction”, B=“significant science/engineering challenges to resolve before initiating construction”, C=“mission and technical requirements not yet fully defined”)

Impact

Each of the four facilities has a key role to play in a balanced ecosystem of DOE high performance computing, and each in its own way contributes in an essential way to DOE's ability to contribute to world leading science. We agree with the AD's assessment that the facilities she identified (1-3) are in the highest "(A): absolutely central" category. We also believe that the proposed Virtual Data Facility is in this category.

1. NERSC is the main engine that supports the breadth of scientific computing for the Office of Science. It provides a broad range of scientists and engineers with advanced technology and applications support and is the vehicle by which cutting edge computing technologies enter production. NERSC helps emerging fields of science and engineering take advantage of supercomputing and, at the same time, provides the production resources needed by all of the programs in the Office of Science.
2. The LCF must continue to address the most challenging computer- and data-intensive problems in the national research portfolio. It helps develop and use the most advanced computing systems for the open science community, including industry, and also works intensively with key science teams to enable breakthrough computations. The lessons learned about large scale computing systems and user support inform NERSC and others about how to broaden and extend the impact of advanced scientific computing to the wider community.
3. ESnet provides the key data linkage for instruments, people, and computational resources. The projected data growth in the next decade is exponential and in some cases faster than Moore's Law. ESnet has a leadership role in delivering highly resilient data transport services optimized for large-scale science. Upgrading to 400 Gb/s on the backbone will have a large impact in addressing this challenge.
4. The VDF will provide an integrated capability for data science across all SC computational and experimental facilities. The ASCAC report on data science and exascale computing notes the emerging impact of "big data" on computation, experiment and science as a whole. Key to DOE's leadership in computing is the development of data science at a scale commensurate with the needs of modern experiment, theory, and computation. This is the challenge VDF directly addresses.

Timelines and Readiness

The committee believes that all existing ASCR facilities – NERSC, LCF and ESnet - are ready to upgrade their facilities in the near-term (2014-2017). That is, there are no significant scientific or engineering challenges as yet unresolved. Specifically, the NERSC CRT building is under construction and scheduled for completion in 2015 and the CD (Critical Decision) process is underway for power upgrades for LCF to accommodate its next generation systems.

Beyond the near-term, there is considerable uncertainty in the performance of systems for a given footprint, power envelope and cost. Therefore, we have divided the 2014-2024 report period into the near-term (2014-2017) and far-term (2018-2024). The out-year uncertainty is manageable if there is a significant, robust exascale program addressing issues in hardware, software and applications.

Additional elements required for effective facilities

Effective computing facilities are comprised of hardware, software, and scientific computing expertise, including applications development and support. Support for applications development and support must come from all offices of SC, particularly in light of the ongoing fundamental change in computing and programming, pioneered by the LCF and soon to be embraced by NERSC. Hardware lifecycles are short (3-4 years) and predictable within known technologies - shorter than the decadal horizon of this report. Application development and support has a significantly more complex and nuanced timeline - starting with early adopters who embrace technology advances, often with significant support from the LCF, and then expanding to include a broader community, including NERSC. It is important to consider all these components in thinking about the timeline.

Broad support for the scientific and engineering applications of the future must be an integral part of the future of these facilities. The LCF has very successfully implemented a relatively small collection of important applications on the next generation of energy-efficient petascale computing systems. However, significant additional domain-specific support to migrate the broad range of DOE applications relying on NERSC systems is required for future success. This is a responsibility of SC as a whole, not only of ASCR.

Synergistic Challenges in Data-Intensive Science and Exascale Computing Report

This report responded to a major charge from the Office of Science to consider the challenges of meeting the new needs for managing data rates and movement of data in an exascale computing environment. It examined how “Big Data” and “Exascale” meet and interact in the DOE context.

Professor Vivek Sarkar of Rice University chaired a distinguished panel to address this issue. Their report is online and has also been submitted with this testimony. Here are salient excerpts from the report:

Introduction

Historically, the two dominant paradigms for scientific discovery have been theory and experiments, with large-scale computer simulations emerging as the third paradigm in the 20th century.

Over the past decade, a new paradigm for scientific discovery is emerging due to the availability of exponentially increasing volumes of data from large instruments such as telescopes, colliders, and light sources, as well as the proliferation of sensors and high-throughput analysis devices. Further, data sources, analysis devices, and

simulations are connected with current-generation networks that are faster and capable of moving significantly larger volumes of data than in previous generations.

However, generation of data by itself is of not much value unless the data can also lead to knowledge and actionable insights. Thus, the *fourth paradigm*, which seeks to exploit information buried in massive datasets to drive scientific discovery, has emerged as an essential complement to the three existing paradigms. For example, experiments using the Large Hadron Collider (LHC) currently generate tens of petabytes of reduced data per year, observational and simulation data in the climate domain is expected to reach exabytes by 2021, and light source experiments are expected to generate hundreds of terabytes per day.

Analysis of this large volume of complex data to derive knowledge, therefore, requires *data-driven* computing, where the data drives the computation and control including complex queries, analysis, statistics, hypothesis formulation and validation, and data mining.

The report considers exemplar use cases from High Energy Physics, Climate Science, Combustion, Biology and Genomics, Light Sources, and Neutron Science all of which illustrate elements needed for exascale data science.

Findings & Recommendations

Finding 1: There are opportunities for investments that can benefit both data-intensive science and exascale computing. There are natural synergies among the challenges facing data-intensive science and exascale computing, and advances in both are necessary for next-generation scientific breakthroughs.

Finding 2: Integration of data analytics with exascale simulations represents a new kind of workflow that will impact both data-intensive science and exascale computing.

Finding 3: There is an urgent need to simplify the workflow for data-intensive science. Analysis and visualization of increasingly larger-scale data sets will require integration of the best computational algorithms with the best interactive techniques and interfaces.

Finding 4: There is a need to increase the pool of computer and computational scientists trained in both exascale and data-intensive computing.

Recommendation 1: The DOE Office of Science should give high priority to investments that can benefit both data-intensive science and exascale computing so as to leverage their synergies.

The findings in this study have identified multiple technologies and capabilities that can benefit both data-intensive science and exascale computing. Investments in

such dual-purpose technologies will provide the necessary leverage to advance science on both data and computational fronts. For science domains that need exascale simulations, commensurate investments in exascale computing capabilities and data infrastructure are necessary. In other domains, extreme-scale components of exascale systems will be well matched for use in different tiers of data analysis, since these processors will be focused on optimizing the energy impact of data movement. Further, research in applications and algorithms to address fundamental challenges in concurrency, data movement, and resilience will jointly benefit data analysis and computational techniques for both data-intensive science and exascale computing. Finally, advances in networking (as projected for future generations of ESN technology) will also benefit both data-intensive science and exascale computing.

Recommendation 2: DOE ASCR should give high priority to research and other investments that simplify the science workflow and improve the productivity of scientists involved in exascale and data-intensive computing.

We must pay greater attention to simplifying human-computer-interface design and human-in-the-loop workflows for data-intensive science. To that end, we encourage the recent proposal for a Virtual Data Facility (VDF) because it will provide a simpler and more usable portal for data services than current systems. A significant emphasis must be placed on research and development of scalable data analytics, mathematical techniques, data mining algorithms and software components that can be used as building blocks for sophisticated analysis pipelines and flows. We also recommend the creation of new classes of proxy applications to capture the combined characteristics of simulation and analytics, so as to help ensure that computational science and computer science research in ASCR are better targeted to the needs of data-intensive science.

Recommendation 3: DOE ASCR should adjust investments in programs such as fellowships, career awards, and funding grants, to increase the pool of computer and computational scientists trained in both exascale and data-intensive computing.

There is a significant gap between the number of current computational and computer scientists trained in both exascale and data-intensive computing and the future needs for this combined expertise in support of DOE's science missions. Investments in ASCR such as fellowships, career awards, and funding grants should look to increase the pool of computer and computational scientists trained in both exascale and data-intensive computing.

Q3: Comment on the draft legislation attached, and provide your general views on the importance of exascale and high end computing to U.S. scientific and economic competitiveness. Please also provide your recommendations regarding how ASCR can, in a constrained budget environment, best advance program goals and activities in a balanced manner.

(1) First, I appreciate that this bill is being discussed and introduced. On behalf of ASCAC, I would say that after the numerous ASCR workshops and our own endorsement of pursuing the opportunities and challenges of exascale computing, we have been gratified to see ASCR and the DOE working within existing frameworks to move along the path to exascale. It is essential that the effort to develop High End Computing be aggressively pursued and we expect that this legislation will be a key enabler of this.

(2) As our reports note, many areas of science, engineering, and industry – particularly data intensive sciences – will be impacted by extreme-scale computing technologies developed on the path to actual exascale machines. We appreciate that the legislation calls for outreach to industry as well as researchers.

(3) We would add to your consideration the importance of continued and increased attention to developing the next generation of computational and data scientists and computing- savvy industry leaders through education programs tightly coupled to the unique capabilities and needs of DOE. The longstanding and successful Computational Science Graduate Fellowship (CSGF) program is a model program that we should keep within ASCR and build upon.

(4) Key to being able to balance and prioritize ASCR activities in a slow growth budget environment is to ensure that the overall ASCR budget is sufficient. Prioritizing, while essential to effective management of an organization, is no remedy for inadequate funding. A great danger to the program in a very restricted budget environment would be the temptation to underinvest in longer-term research in math and computing that, as history shows, provide essential benefits to applications. A second danger is to forgo leadership opportunities because we might be able to do something whose short-term costs are smaller but which, in the long term, sacrifice global leadership.