# AMERICA'S NEXT GENERATION SUPERCOMPUTER: THE EXASCALE CHALLENGE

## HEARING

BEFORE THE

SUBCOMMITTEE ON ENERGY

COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

HOUSE OF REPRESENTATIVES

ONE HUNDRED THIRTEENTH CONGRESS

FIRST SESSION

WEDNESDAY, MAY 22, 2013

**Serial No. 113–31**

Printed for the use of the Committee on Science, Space, and Technology

Available via the World Wide Web: http://science.house.gov

U.S. GOVERNMENT PRINTING OFFICE
81–195PDF WASHINGTON : 2013

For sale by the Superintendent of Documents, U.S. Government Printing Office
Internet: bookstore.gpo.gov Phone: toll free (866) 512–1800; DC area (202) 512–1800
Fax: (202) 512–2104 Mail: Stop IDCC, Washington, DC 20402–0001

## COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

HON. LAMAR S. SMITH, Texas, *Chair*

DANA ROHRABACHER, California
RALPH M. HALL, Texas
F. JAMES SENSENBRENNER, JR., Wisconsin
FRANK D. LUCAS, Oklahoma
RANDY NEUGEBAUER, Texas
MICHAEL T. McCAUL, Texas
PAUL C. BROUN, Georgia
STEVEN M. PALAZZO, Mississippi
MO BROOKS, Alabama
RANDY HULTGREN, Illinois
LARRY BUCSHON, Indiana
STEVE STOCKMAN, Texas
BILL POSEY, Florida
CYNTHIA LUMMIS, Wyoming
DAVID SCHWEIKERT, Arizona
THOMAS MASSIE, Kentucky
KEVIN CRAMER, North Dakota
JIM BRIDENSTINE, Oklahoma
RANDY WEBER, Texas
CHRIS STEWART, Utah
VACANCY

EDDIE BERNICE JOHNSON, Texas
ZOE LOFGREN, California
DANIEL LIPINSKI, Illinois
DONNA F. EDWARDS, Maryland
FREDERICA S. WILSON, Florida
SUZANNE BONAMICI, Oregon
ERIC SWALWELL, California
DAN MAFFEI, New York
ALAN GRAYSON, Florida
JOSEPH KENNEDY III, Massachusetts
SCOTT PETERS, California
DEREK KILMER, Washington
AMI BERA, California
ELIZABETH ESTY, Connecticut
MARC VEASEY, Texas
JULIA BROWNLEY, California
MARK TAKANO, California
ROBIN KELLY, Illinois

## SUBCOMMITTEE ON ENERGY

HON. CYNTHIA LUMMIS, Wyoming, *Chair*

RALPH M. HALL, Texas
FRANK D. LUCAS, Oklahoma
RANDY NEUGEBAUER, Texas
MICHAEL T. McCAUL, Texas
RANDY HULTGREN, Illinois
THOMAS MASSIE, Kentucky
KEVIN CRAMER, North Dakota
RANDY WEBER, Texas
LAMAR S. SMITH, Texas

ERIC SWALWELL, California
ALAN GRAYSON, Florida
JOSEPH KENNEDY III, Massachusetts
MARC VEASEY, Texas
MARK TAKANO, California
ZOE LOFGREN, California
DANIEL LIPINSKI, Illinois
EDDIE BERNICE JOHNSON, Texas

# CONTENTS

## Wednesday, May 22, 2013

### Opening Statements

### Witnesses:

### Appendix I: Answers to Post-Hearing Questions

# AMERICA'S NEXT GENERATION SUPERCOMPUTER: THE EXASCALE CHALLENGE

---

**WEDNESDAY, MAY 22, 2013**

HOUSE OF REPRESENTATIVES,
SUBCOMMITTEE ON ENERGY
COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY,
*Washington, D.C.*

The Subcommittee met, pursuant to call, at 10:05 a.m., in Room 2318 of the Rayburn House Office Building, Hon. Cynthia Lummis [Chairwoman of the Subcommittee] presiding.

LAMAR S. SMITH, Texas
CHAIRMAN

EDDIE BERNICE JOHNSON, Texas
RANKING MEMBER

Congress of the United States
House of Representatives
COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY
2321 RAYBURN HOUSE OFFICE BUILDING
WASHINGTON, DC 20515–6301
(202) 225–6371
www.science.house.gov

## Subcommittee on Energy

## *America's Next Generation Supercomputer: The Exascale Challenge*

Wednesday, May 22, 2013
10:00 a.m. – 12:00 p.m.
2318 Rayburn House Office Building

Witnesses

**Dr. Roscoe Giles**, Chairman, Advanced Scientific Computing Advisory Committee, Professor, Boston University

**Dr. Rick Stevens,** Associate Laboratory Director, Computing, Environment and Life Sciences, Argonne National Laboratory

**Ms. Dona Crawford,** Associate Director for Computation, Lawrence Livermore National Laboratory

**Dr. Daniel Reed,** Vice President for Research and Economic Development, University of Iowa.

# U.S. HOUSE OF REPRESENTATIVES
# COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY
# SUBCOMMITTEE ON ENERGY

## HEARING CHARTER

***America's Next Generation Supercomputer: The Exascale Challenge***

Wednesday, May 22, 2013
10:00 a.m. – 12:00 p.m.
2318 Rayburn House Office Building

## PURPOSE

The Subcommittee on Energy will hold a hearing entitled *America's Next Generation Supercomputer: The Exascale Challenge* on Wednesday, May 22, at 10:00 a.m. in Room 2318 of the Rayburn House Office Building. The purpose of the hearing is to examine high-performance computing research and development challenges and opportunities, specifically as they relate to exascale computing. The hearing will also explore advanced scientific computing research. The hearing will additionally examine draft legislation[1] directing the Department of Energy (DOE) to develop an exascale computing system.

## WITNESS LIST

- **Dr. Roscoe Giles**, Chairman, Advanced Scientific Computing Advisory Committee, Professor, Boston University.
- **Dr. Rick Stevens,** Associate Laboratory Director, Computing, Environment and Life Sciences, Argonne National Laboratory.
- **Ms. Dona Crawford,** Associate Director for Computation, Lawrence Livermore National Laboratory.
- **Dr. Daniel Reed,** Vice President for Research and Economic Development, University of Iowa.

## BACKGROUND

Scientific research is traditionally conducted through theory or experimentation, both of which generate data that requires the capacity to be processed and analyzed. The invention of computers permitted this data to be examined with increased speed and complexity. As computational technology advanced, this capacity increased in pace and capability, while the data generated from various sensors and experiments also increased in volume. The advent of

---

[1] Legislation is appended.

scientific discovery in which large volumes of data is gathered and mined to exploit information, sometimes referred to as "big data,"[2] has transformed computing technology needs.

The greater availability and utilization of these high-speed supercomputers allows increasingly complex scientific research to be achieved. Medical research, energy and environment system simulations, computational chemistry, and innumerable other scientific problems directly benefit from high-performance computing (HPC).

Computing speed is measured in floating-point operations per second, or flops. In the 1970's, the first supercomputers had a capacity of about 100 megaflops, or 100 million flops. Through forty years of technology advancement, computing capacity climbed through gigaflops ($10^9$ calculations per second) and teraflops ($10^{12}$), to current HPC capacity of petaflops ($10^{15}$). Exascale computing refers to computing systems capable of a thousand-fold increase over current petascale computers, or the capability to do a quintillion, $10^{18}$, calculations per second. To put this in context, there are currently about 1 sextillion ($10^{21}$) known stars in the universe – therefore "an exascale computer could count every star in the universe in 20 minutes."[3]

Currently, the fastest computer in the world is the Cray Titan, located at Oak Ridge National Lab, with a peak speed of 17.59 petaflops.[4] As of November 2012, the United States was home to five of the ten fastest supercomputers in the world (others include two from Germany and one each from Japan, China, and Italy).[5] Three of the top ten fastest supercomputers in the United States were developed and operated by the Department of Energy at Oak Ridge National Laboratory, Lawrence Livermore National Laboratory, and Argonne National Laboratory. The National Science Foundation also supports the development of supercomputers.

## Advanced Scientific Computing Research Program

DOE's Office of Science administers the Advanced Scientific Computing Research (ASCR) program. ASCR is the leading supporter of non-military high-performance computing program within the Federal government. ASCR's mission is to:

> "advance applied mathematics and computer science; deliver, in partnership with disciplinary science, the most advanced computational scientific applications; advance

---

[2] For more information on "Big Data" see April 24, 2013 House Science, Space, and Technology Subcommittees on Research and Technology hearing titled *"Next Generation Computing and Big Data Analytics."* http://science.house.gov/hearing/subcommittee-technology-and-subcommittee-research-joint-hearing-next-generation-computing

[3] Department of Energy, Advanced Scientific Computing Research "Leap to the extreme scale could break science boundaries," February 14, 2011. Accessible at: http://ascr-discovery.science.doe.gov/feature/exa_ov1.shtml

[4] http://wayback.archive.org/web/20130121075914/http://top500.org/blog/lists/2012/11/press-release/

[5] http://www.top500.org/lists/2012/11/

computing and networking capabilities; and develop, in partnership with the research community, including U.S. industry, future generations of computing hardware and tools for science."[6]

***Department of Energy (DOE) Advanced Scientific Computing Research Spending***
*(dollars in millions)*

| **Subprogram** | **FY12 Current** | **FY13 Annualized CR*** | **FY14 Request** | **FY14 Request versus FY12 Enacted** | |
|---|---|---|---|---|---|
| | | | | **$** | **%** |
| *Mathematical, Computational, and Computer Sciences Research* | *151.6* | -- | *172.4* | *20.8* | *13.7* |
| *High Performance Computing and Network Facilities* | *276.7* | -- | *293.1* | *16.4* | *5.9* |
| **Total, Advanced Scientific Computing Research** | ***428.3*** | ***443.6*** | ***465.6*** | ***37.3*** | ***8.7*** |

*FY 2013 amounts shown reflect the P.L. 112-175 continuing resolution level annualized to a full year. These amounts are shown only at the "congressional control" level and above; below that level a dash (--) is shown.

In addition to high-performance computing activities, DOE's ASCR program also supports other activities in applied mathematics, computer science, next generation networks for science, and computational partnerships. For example, DOE is funding the development of the Energy Sciences Network (ESNet), to provide high-bandwidth connections to link national laboratories, universities, and other research institutions to allow those entities to collaborate on scientific research.

ASCR's primary scientific computing facility is the National Energy Research Scientific Computing Center (NERSC). NERSC, managed by Lawrence Berkeley National Laboratory, performs basic scientific research over a wide range of disciplines, such as material science, high energy physics data analysis, and chemistry simulations.[7]

DOE Exascale Strategy

The FY 2012 Consolidated Appropriations Act[8] expressed Congressional support for exascale computing, and specifically noted exascale is a "crucial component of long-term U.S. leadership." However, the accompanying Conference Report stressed the need for an "integrated

---

[6] DOE FY14 Detailed Budget Request, Volume 4, SC-21.

[7] National Energy Research Scientific Computing Center, "About NERSC," Last edited, April 4, 2013. Accessible at: http://www.nersc.gov/about/

[8] P.L. 112-74

strategy and program plan" from DOE. Accordingly, the Conference Report directed DOE to submit to the House and Senate Committees on Appropriations, not later than February 10, 2012, "a joint, integrated strategy and program plan for the crosscutting effort to develop exascale computing that includes:

- a target date for developing an operational exascale platform;
- interim milestones toward reaching that target;
- minimum requirements for an exascale supercomputer system, including power consumption efficiency goals;
- multi-year budget estimates for the exascale supercomputer initiative and costs of meeting each interim milestone;
- clear roles and responsibilities for each office involved in exascale supercomputer research and development; and
- a complete listing of exascale supercomputer activities included in the fiscal year 2013 budget request broken out by program, project and activity with comparisons to the current year's funding levels."[9]

Despite the directive, DOE has not yet reported its plan to Congress. In the absence of the DOE exascale supercomputer strategy, ASCR's FY14 budget request for High Performance Computing and Network Facilities subprogram still includes funding to "expand investments in critical technologies for exascale." However, no specific budget is requested.

### Non-Civilian Exascale Uses

Should exascale computing be developed, a major beneficiary would be DOE's National Nuclear Security Administration (NNSA). NNSA supports a number of unclassified and classified computing activities to maintain the nuclear stockpile and develop new nuclear weapons. NNSA has previously held workshops in partnership with the Office of Science to examine research and development challenges to go from petascale to exascale computing systems.[10] ASCR is currently developing an exascale supercomputer development plan with the NNSA.[11]

### Exascale Challenges

While exascale supercomputers would serve as a breakthrough leap above current computing capacity, important scientific and technical obstacles currently exist. For example, to facilitate increased computing speed, current computers can scale up to 250,000 parallel

---

[9] P.L. 112-74, Conference Report , p. 846

[10] Lawrence Livermore National Laboratory, "From Petascale to Exascale: R&D Challenges for HPC Simulation Environments," Last updated: October 24, 2012. Accessible at: https://asc.llnl.gov/exascale/

[11] DOE FY14 Detailed Budget Request, Volume 4, SC – 26.

processors, known as "parallelism." However, an exascale system would require parallelism up to one billion processors, which is an extremely complex and daunting task.[12] Additionally, using today's computing technology, an exascale system would consume more than a gigawatt of electricity.[13] One gigawatt of power is equivalent to power demands for roughly 700,000 to 1,000,000 homes, or the power output of a single, dedicated nuclear reactor. Obviously, the operating costs for such a exascale supercomputer could cost hundreds of millions of dollars per year in electricity costs alone if technological advancements are not made.

ASCAC Exascale Report on Synergistic Challenges

In March 2013, the Advanced Scientific Computing Advisory Committee (ASCAC) Data Subcommittee issued a report *Synergistic Challenges in Data-Intensive Science and Exascale Computing*.[14] The report reviewed challenges facing both "Big Data" and exascale computing systems and commented on the relationship between them. The report notes, "data-intensive research activities are increasing in all domains of science, and exascale computing is a key enabler of these activities."[15] ASCAC identified four findings and made three accompanying recommendations:

Findings:

1. There are opportunities for investments that can benefit both data-intensive science and exascale computing.
2. Integration of data analytics with exascale simulations represents a new kind of workflow that will impact both data-intensive science and exascale computing.
3. There is an urgent need to simplify the workflow for data-intensive science.
4. There is a need to increase the pool of computer and computational scientists trained in both exascale and data-intensive computing.

Recommendations:

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.
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.

---

[12] DOE ASCR "Leap to the extreme scale."

[13] DOE ASCR "Leap to the extreme scale."

[14] Department of Energy, Advanced Scientific Computing Advisory Committee, Data Subcommittee Report, "Synergistic Challenges in Data-Intensive Science and Exascale Computing," March, 2013. Accessible at: http://science.energy.gov/~/media/ascr/ascac/pdf/reports/2013/ASCAC_Data_Intensive_Computing_report_final.pdf

[15] ASCAC Data-Intensive Computing report.

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.

## ADDITIONAL READING

DOE ASCAC Subcommittee Report: *The Opportunities and Challenges of Exascale Computing,* Fall 2010.
http://science.energy.gov/~/media/ascr/ascac/pdf/reports/exascale_subcommittee_report.pdf

DOE ASCAC Data Subcommittee Report: *Synergistic Challenges in Data-Intensive Science and Exascale Computing*, March 2012.
http://science.energy.gov/~/media/ascr/ascac/pdf/reports/2013/ASCAC_Data_Intensive_Computing_report_final.pdf

Chairwoman LUMMIS. Good morning. The Subcommittee will come to order. And we are delighted to have a terrific panel here this morning, so welcome to our hearing entitled "America's Next Generation Supercomputer: the Exascale Challenge." In front of you are packets containing the written testimonies, biographies, and truth-in-testimony disclosures for today's witness panel.

And now, I will recognize myself for five minutes for and opening statement followed by our Ranking Member Mr. Swalwell.

The development and expanded availability of supercomputers has enabled society to push the frontiers of nearly every scientific discipline, and accelerate applications of that science in countless fields. It has enabled modeling and simulation necessary to address national security needs. It drives the boundaries of medical research, reduces cost to develop new products, and improves materials design processes, just to name a few.

High performance computing has also revolutionized how the energy sector operates. Advanced modeling and simulation techniques, driven by computer algorithms and faster computing speeds, improve the efficiency of energy production and consumption technologies.

These advancements ultimately trace back to Federal investments in basic research that provided the foundation for most of today's computing technologies. From the first megaflop supercomputers of the 1960s, the Federal investments have led to push across each landmark thousand-fold speed barrier to gigaflops, teraflops, and petaflops. I always think of floppy-eared rabbits and when I was a kid showing critters in 4H, I should have named them Giga, Tera, and Peta, but I just didn't know about it back then because that proceeded the first megaflop.

Throughout this computing age, we have witnesses—we have witnessed yesterday's supercomputers become today's desktop computers and consumer devices often in incredibly short time frames. The spillover benefits to society are countless and immeasurable.

The Department of Energy, led by the Advanced Scientific Computing Research program, plays a critical role in driving these computing technology breakthroughs. DOE supports world-class computational science facilities, such as the National Energy Research Scientific Computing Center. Additionally, DOE funds cutting-edge applied mathematics research and next-generation networking activities.

DOE's next major computing challenge, constructing an "exascale" computer system that is a thousand times faster than current world-leading supercomputers, may be the most daunting. Key scientific and technical obstacles associated with the architecture and energy efficiency of an exascale system must be overcome, and an immense amount of resources and effort will be required.

As we head down this inevitable path to exascale computing, it is important we take time to plan and budget thoroughly to ensure a balanced approach that ensures broad buy-in from the scientific computing community. The Federal Government has limited resources and taxpayer funding must be spent on the most impactful projects. We need to ensure DOE efforts to develop an exascale system can be undertaken in concert with other foundational advanced scientific computing activities. This morning, we will hear testimony from expert witnesses regarding how best to achieve this balance.

I would like to recognize if he is here, yes, he has come in, a leader in this effort, my colleague on the Energy Subcommittee, Representative Randy Hultgren.

I would now like to yield the balance of my time to the gentleman from Illinois to summarize the discussion draft of his bill, "American High-End Computing Leadership Act."

[The prepared statement of Mrs. Lummis follows:]

### PREPARED STATEMENT OF SUBCOMMITTEE CHAIRMAN CYNTHIA LUMMIS

Good morning and welcome to today's Energy Subcommittee hearing to examine high performance computing research and development challenges and opportunities.

The development and expanded availability of supercomputers has enabled society to push the frontiers of nearly every scientific discipline, and accelerate applications of that science in countless fields. It has enabled modeling and simulation necessary to address national security needs. It drives the boundaries of medical research, reduces cost to develop new products, and improves materials design processes, just to name a few areas.

High performance computing has also revolutionized how the energy sector operates. Advanced modeling and simulation techniques, driven by complex algorithms and faster computing speeds, improve the efficiency of energy production and consumption technologies.

These advancements ultimately trace back to Federal investments in basic research that provided the foundation for most of today's computing technologies. From the first megaflop supercomputers of the 1960s, Federal investments have led the push across each landmark thousand-fold speed barrier-to gigaflops, teraflops, and petaflops. Throughout this computing age, we have witnessed as yesterday's supercomputers become today's desktop computers and consumer devices often in incredibly short time frames. The spillover benefits to society are countless and immeasurable.

The Department of Energy, led by the Advanced Scientific Computing Research program, plays a unique and critical role in driving these computing technology breakthroughs. DOE supports world-class computational science facilities, such as the National Energy Research Scientific Computing Center. Additionally, DOE funds cutting edge applied mathematics research and next generation networking activities.

DOE's next major computing challenge-constructing an "exascale" computer system that is a thousand times faster than current world-leading supercomputers-may be the most daunting. Key scientific and technical obstacles associated with the architecture and energy efficiency of an exascale system must be overcome, and an immense amount of resources and effort will be required.

As we head down this inevitable path to exascale computing, it is important we take time to plan and budget thoroughly to ensure a balanced approach that ensures broad buy-in from the scientific computing community. The Federal government has limited resources and taxpayer funding must be spent on the most impactful projects. We need to ensure DOE efforts to develop an exascale system can be undertaken in concert with other foundational advanced scientific computing activities. This morning, we will hear testimony from expert witnesses regarding how best to achieve this balance.

I would like to recognize a leader of this effort, my colleague on the Energy Subcommittee, Representative Randy Hultgren. I would now like to yield the balance of my time to the gentleman from Illinois to summarize the discussion draft of his bill, "American High-End Computing Leadership Act."

Mr. HULTGREN. Thank you, Madam Chair, for holding this hearing today. Exascale computing represents a brave new world of science for our Nation. The application of the next generation of supercomputers is vast. A thousand-fold increase in processing power will give us the intense computing tools necessary to ensure our national security by better testing our nuclear stockpile, revolutionized our understanding and treatment of complicated healthcare problems like neurological diseases or the genetics underpinning cancer with the ability to model new treatments and ensure our Nation's competitiveness in the big data economy of the 21st century by spilling over knowledge and expertise into industry and academia.

And while I can postulate further on some of the applied uses of faster machines, I also know that simply by making these investments in basic science needed to overcome challenges in the immensely massive parallelism, power management, new architecture, and programming models, we will enrich our Nation intellectually and ensure our labor force remains competitive.

I think at that point I will yield back, Madam Chair. Let me follow up if I have another minute. Do I?

Chairwoman LUMMIS. Mr. Hultgren, you do.

Mr. HULTGREN. Madam Chair, let me summarize my bill quickly. Thank you.

My bill would amend the existing statute by specifying the need to target the specific challenges and power requirements and parallelism required to make the leap to exascale. It also will instruct the Secretary of Energy to conduct a coordinated research program to develop exascale computing systems and require an integrated strategy and program management plan to ensure the health of existing research activities is not harmed.

The bottom line is we do not know all of the ways we will use this next-generation of supercomputers, but given the vast and unpredictable ways that computing technology has already enhanced every part of our lives and given the investments being made in other countries to deploy large-scale systems, it is more important than ever that we make this investment today.

I look forward to hearing the witnesses, what they think of this legislative proposal, areas we can improve it, challenges that we will face. And with that, I do thank you. I apologize for my confusion here but I yield back to the Chairwoman. Thank you very much, Madam Chair.

[The prepared statement of Mr. Hultgren follows:]

## PREPARED STATEMENT OF REPRESENTATIVE RANDY HULTGREN

Thank you, Madam Chair, for holding this hearing today.

Exascale computing represents an exciting new world of science for our nation. The applications for the next generation of super computers are vast.

A thousand fold increase in processing power will give us the intense computing tools necessary to ensure our national security by better testing our nuclear stockpile; revolutionize our understanding and treatment of complicated health care problems like neurological diseases or the genetics underpinning cancer with the ability to model new treatments; and ensure our nation's competitiveness in the big data economy of the 21st century by spilling over knowledge and expertise into industry and academia.

And while I can postulate further on some of the applied uses of faster machines; I also know that simply by making these investments in the basic science needed to overcome challenges in immensely massive parallelism, power management, new architectures and programming models, we will enrich our nation intellectually and ensure our labor force remains competitive.

Madam Chair, my bill would amend the existing statute by specifying the need to target the specific challenges in power requirements and parallelism required to make the leap to exascale. It would also instruct the Secretary of Energy to conduct a coordinated research program to develop exascale computing systems, and require an integrated strategy and program management plan to ensure the health of existing research activities is not harmed.

The bottom line is, we do not know all of the ways we will use the next generation of supercomputers, but given the vast and unpredictable ways that computing technology has already enhanced every part of our lives, and given the investments being made in other countries to deploy large scale systems, it is more important than ever that we make this investment today.

I look forward to hearing what the witnesses think of this legislative proposal, areas we can improve it, challenges we face, and with that I thank you and I yield back.

Chairwoman LUMMIS. The gentleman yields back.

And I might add on a personal note, today, my daughter is being awarded her master's degree in digital media from Columbia University. I unfortunately cannot be at her graduation because Congress is in session but I get to watch it on the computer, so I will get to see it. And I think to myself, first of all, what is a master's degree in digital media? Somebody my age doesn't even know what that is. And certainly, when I was her age, I could not have even begun to envision the career that would be open to her as of today, and the career that is open to her as of today is due in part to the investment that the people in this room and that the American people have made in computing, for science, and for the benefit of mankind. So this is a very important subject.

The fact that it is such an important subject leads me to let you all know that there will be several comings and goings by Committee Members this morning. There are concurrent meetings going on around the buildings. In my case, we have the IRS in front of us down in Oversight and Government Reform and I know there are other Members that may have to come and go from time to time. We deeply appreciate your testimony here today. In my absence, our Vice Chair Mr. Weber will be in the chair, and of course, Mr. Swalwell, who is our Ranking Member, who I will recognize now, the gentleman from California, Mr. Swalwell.

Mr. SWALWELL. Thank you, Chairman Lummis. And also congratulations to your daughter on this achievement. And thank you for holding this hearing today. And I want to thank the witnesses for being here. I also thank the witnesses who are not from the 15th Congressional District. We welcome you as well but especially welcome Ms. Crawford from Livermore, California.

I am excited to learn more about the work that the DOE is doing in partnership with industry and our national laboratories, including both Lawrence Livermore and Berkeley national laboratories in particular and are carrying out to maintain the United States' leadership in the critical area of high-performance computing.

As I am sure the witnesses will all describe in more detail, this capability enables our best and brightest minds to gain new insights into societal concerns ranging from Alzheimer's disease to climate change. Other examples of both industrial and academic research that benefit from our advanced, high-end computing capabilities include high-temperature superconductivity to significantly reduce energy losses in transmitting electricity; aerodynamic modeling for aircraft and vehicle design; pharmaceutical development; next-generation nuclear reactor design; fusion plasma modeling; and combustion simulation to guide the design of fuel-efficient clean engines such as work being carried out at the Sandia National Laboratory's combustion research facility.

In short, many of the most pressing issues of our time, whether it is how we find our energy resources, how we make our energy resources more efficient, or how we solve the rising cost of healthcare can be solved through investments in high-performance computing.

A focus of today's hearing is the development of an exascale computing capability. Now, my understanding is that exascale is often interchangeably used with extreme scale to refer to the next generation of supercomputers in general, but it also refers to a computing system that would be able to carry out a million trillion operations per second. Yes, a million trillion or a 1 with 18 zeros after it. That is about 500 times faster than the world's fastest computer today. Such a system would be critical to meeting the Nation's needs in a number of important research areas like combustion science, climate science, modeling of the human brain, and ensuring the reliability of our nuclear weapons stockpile.

That said, as we pursue the next generation of supercomputing capabilities, which I fully support, I want to ensure that the Nation is getting the most bang for buck out of our current world-leading facilities. It is noteworthy that while Lawrence Livermore, Argonne, and Oak Ridge national laboratories are three of the most powerful supercomputing centers in the world, and they are addressing incredibly important scientific issues that really require their advanced computing capabilities. Lawrence Berkeley's National Energy Research Scientific Computing Center actually serves thousands more users with only a fraction of those leadership machines' computing power.

The point is not every computational research effort requires the fastest most sophisticated system we can possibly build and I think we also need to work more to make sure that what is sometimes called capacity supercomputing is more accessible to both the academic and industrial research communities that could benefit.

I have always believed whether it was as a local city councilman or a sitting Member of Congress that the government works best when we can share our resources with the private sector. It doesn't serve anyone any good if we are just doing the research in the government and not transferring that research out to the private sector, and I think in high-performance computing we have already shown in our laboratories we are transferring it out. The transfer out makes us more efficient, can reduce healthcare costs, and also more importantly, especially in our area, it can create private-sector jobs on top of the thousands of jobs that already exist at our laboratories.

So with that, I look forward to discussing these important issues with each of you today and I yield back the balance of my time.

[The prepared statement of Mr. Swalwell follows:]

## PREPARED STATEMENT OF SUBCOMMITTEE RANKING MEMBER ERIC SWALWELL

Thank you Chairman Lummis for holding this hearing today, and I also want to thank the witnesses for being here—even the ones from outside of the 15th District of California!

I am excited to learn more about the great work that the Department of Energy in partnership with industry and our national laboratories, including both Lawrence Livermore and Lawrence Berkeley National Laboratories in particular, are carrying out to maintain and advance U.S. leadership in the critical area of high performance computing.

As I'm sure the witnesses will describe in more detail, this capability enables our best and brightest scientists to gain new insights into societal concerns ranging from Alzheimer's disease to climate change. Other examples of both industrial and academic research that benefit from our advanced high-end computing capabilities include: high temperature superconductivity to significantly reduce energy losses in transmitting electricity; aerodynamic modeling for aircraft and vehicle design; pharmaceutical development; next generation nuclear reactor design; fusion plasma modeling; and combustion simulation to guide the design of fuel-efficient clean engines, such as work being carried out at the Sandia National Laboratories' Combustion Research Facility.

A focus of today's hearing is the development of an exascale computing capability. Now, my understanding is that "exascale" is often used interchangeably with "extreme scale" to refer to the next generation of supercomputers in general, but it also refers to a computing system that would be able to carry out a million trillion operations per second. (Yes, a million trillion, or a 1 with 18 zeros after it.) That's about 500 times faster than the world's fastest computers at today. Such a system would be critical to meeting that nation's needs in a number of important research areas like combustion science, climate science, modeling of the human brain, and ensuring the reliability of our nuclear weapons stockpile.

That said, as we pursue the next generation of supercomputing capabilities–which I fully support–I also want to ensure that the nation is getting the most bang per buck out of our current world-leading facilities. It is noteworthy that while Lawrence Livermore, Argonne, and Oak Ridge National Laboratories are 3 of the most powerful supercomputers in the world, and they are addressing incredibly important scientific issues that really require their advanced computing capabilities, Lawrence Berkeley's National Energy Research Scientific Computing Center actually serves thousands of more users with only a fraction of those leadership machines' computing power. The point is, not every computational research effort requires the fastest, most sophisticated system we can possibly build, and I think we also need to do more to make what's sometimes called "capacity" supercomputing more accessible to both the academic and industrial research communities that could benefit.

With that, I look forward to discussing these important issues with each of you today, and I yield back the balance of my time.

Chairwoman LUMMIS. Thank you, Mr. Swalwell.

If there are Members who wish to submit additional opening statements, your statements will be added to the record at this point.

Well, at this time I would like to introduce our witnesses, and the fun part today is we have two Members here who have witnesses from their districts. So I will start by introducing Dr. Roscoe Giles, Chairman of the Advanced Scientific Computing Advisory Committee of the Department of Energy and Professor at Boston University. Dr. Giles—and I have that right, don't I, Dr. Giles? Thank you. He has served in a number of leadership roles in the community including Member of the Board of Associated Universities, Inc., Chair of the Boston University Faculty Council, and General Chair of the SC Conference in 2002. He received his Ph.D. in physics from Stanford University in 1975. That is a remarkable record of achievement, Dr. Giles. Thank you for being here.

At this time, I would like to yield to the gentleman from Illinois, Mr. Hultgren, to introduce our second witness.

Mr. HULTGREN. Thank you, Madam Chair.

Our second witness is Dr. Rick Stevens, Associate Laboratory Director for Computing, Environment, and Life Sciences at Argonne National Laboratory. He heads Argonne's Computational Genomics Program and co-leads the DOE's laboratory planning effort for exascale computing research. He is also Professor of computer science at the University of Chicago and is involved in several interdisciplinary studies at the Argonne University of Chicago

Computation Institute and at the Argonne University of Chicago Institute for Genomics and Systems Biology. He is doing amazing work at Argonne and at the University and the entire Illinois community is proud of his contributions to this cutting edge field of science. We are very glad to have you here, Dr. Stevens. Thank you.

I yield back.

Chairwoman LUMMIS. Thank you for your attendance today. That was my field although at a much lower level of academic achievement, Dr. Stevens. We are delighted you are here.

Now, I would like to yield to the gentleman from California, Mr. Swalwell, to introduce our third witness.

Mr. SWALWELL. Thank you, Chairman Lummis.

And I have been very eager on this Committee to have a witness from Lawrence Livermore laboratory.

Chairwoman LUMMIS. I can testify to that.

Mr. SWALWELL. I thank you for allowing this witness to be here today. Lawrence Livermore is the largest employer in my Congressional District and I have to really just commend the laboratory for their advocacy of the issues facing Lawrence Livermore. They are in constant contact with our office and this Committee so I am honored to today introduced Dona Crawford, who is the Associate Director of Computation at Lawrence Livermore National Laboratory.

Ms. Crawford is responsible for a staff of roughly 900 to develop and deploy an integrated computing environment for advanced simulations of complex physical phenomena like climate change, clean energy creation, biodefense, and nonproliferation. She has served on Advisory Committees for the National Academies and the National Science Foundation and currently serves as co-Chair of the Council on Competitiveness High-Performance Computing Advisory Committee, and is a member of IBM's Deep Computing Institute External Advisory Board. Ms. Crawford has a master's degree in operations research from Stanford University and a bachelor's degree in mathematics from the University of Redlands, California.

Ms. Crawford, thank you for being here today and I yield back the balance of my time.

Chairwoman LUMMIS. Thank you, Mr. Swalwell. And my first exposure to Livermore, I used to walk around the lab. My first job out of college was working for a rodeo company in Northern California, and we were putting on the rodeo at Livermore.

Mr. SWALWELL. It is the fastest rodeo in the world. Did you know that?

Chairwoman LUMMIS. You know, considering the rodeo company I worked for, I would believe that. Those rodeos ran like that and I used to go for walks around the lab just to get some exercise when I was there at Livermore putting on rodeos. So I know where you are, at least I knew where you are when I was a young college graduate in my first job.

Our final witness is Dr. Daniel Reed, Vice President of Research and Economic Development at the University of Iowa. Previously, he served as a Senior Leader at Microsoft serving as Microsoft's Computing Strategist to Corporate Vice President for Extreme Computing, I love that, and Technology Policy. He received his Ph.D. in computer science in 1983 from Purdue University.

As our witnesses should know, spoken testimony is limited to five minutes each, after which the Members of the Committee will have five minutes each to ask questions. I now recognize Dr. Giles for five minutes to present his testimony with deep gratitude to all of you for your attendance today. Dr. Giles.

## TESTIMONY OF DR. ROSCOE GILES, CHAIRMAN, ADVANCED SCIENTIFIC COMPUTING ADVISORY COMMITTEE

Dr. GILES. Yes, thank you, Chairman Lummis. And thanks to Members of the Committee for inviting me to testify today.

I think the bill you are considering is very, very important for our field and for maintaining the Nation's leadership in computing and computational science. I am testifying today in my role as Chair of the Advisory Committee to ASCR and I will try to reflect that committee's views of some elements of the ASCR program and hope to demonstrate that we are ready to move forward and sort of eager to move forward in this direction. And it is important that we do so.

The Office of Advanced Scientific Computing Research has programs and investments that include computer and networking facilities that support DOE's science programs; leadership computing facilities for which the exascale discussion is very directly relevant with unique high-end capabilities made available to DOE and to all the Nation, including industry; applied mathematics research whose results provide the framework for future applications and systems; computer science system and software research, whose results both enable applications of current systems and chart the direction for future systems.

And beyond this, ASCR investments—ASCR is the abbreviation for Advanced Scientific Computer Research—we get lost in acronyms sometimes. ASCR investments have also built human expertise in the scientific and technical staff at the labs and through attention to integrating the next generation of computational science leaders into DOE programs and facilities through programs like the Computational Science Graduate Fellows Program, which I also am involved with.

It is hard in these few minutes to state the breadth and depth of science productivity that is being enabled by these machines. We now see the initial results of the petascale era. As one measure, we might mention that more than 2,000 peer-reviewed research articles based directly on projects supported by ASCR computing facilities were published in 2012 alone. One I love is a trillion-particle simulation in cosmology, since I started out in the '80s struggling to do a million-particle simulation of molecular dynamics, and to go another factor of a million is astonishing.

In 2009 our advisory committee was charged with reviewing ASCR's body of work on exascale computing. We delivered the Rosner report, "The Opportunities and Challenges of Exascale Computing," in fall of 2010. We found the case for exascale computing compelling and recommended the DOE should proceed expeditiously with an exascale initiative so that it continues to lead in using extreme scale computing to meet important national needs.

As you have heard mentioned this morning, when we wrote that, we were talking about growing a factor of 1,000 forward in the future. Now, that is a factor of 500. I am glad to see that we are starting to in this bill to really move forward on this. And we have had a sense in the committee that we have been waiting for that forward motion from the system.

Some of the—during this time, ASCR has been busy doing foundational research to make this possible, so there—and we will hear more about it, I am sure, from other speakers. But the establishment of co-design centers, computing research, and applied mathematics research and some prototype projects with fast forward and design forward that are bringing us in this direction, and I think we are making progress but not the progress we should be making at the scale we should be making it, and hopefully, the bill will help deal with that issue.

Our committee has been asked to review ASCR facility plans for the relatively short-term future of the next ten years, not including exascale deployment, and we found those facility plans to be very sound and compelling that involve enhancements to the petascale systems. We have also recently examined the intersection of big data needs within the Department of Energy and ASCR's exascale program and found them quite convergent. The exascale technologies we are talking about developing will be essential in systems that analyze big data problems of the nature that come to the Department of Energy from both experiment and theory, and we have a—quite a long and detailed report about that.

I wanted to just summarize by saying I am very, very glad to see the legislation that we have here. I am very supportive of the direction we are going. I would only ask that the funding level be sure to be sufficient for the scope of our dreams. Thank you.

[The prepared statement of Dr. Giles follows:]

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 Survey[1] 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 ESNet 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.

## Citation List for Roscoe Giles Testimony

The cititations are all available from links on the report Section of the Advanced Scientific Computing Advisory Committee website: http://science.energy.gov/ascr/ascac/reports/

Specific reports which are part of the testimony are:

(1) "Synergistic Challenges in Data-Intensive Science and Exascale Computing", March 2013: http://science.energy.gov/~/media/ascr/ascac/pdf/reports/2013/ASCAC_Data_Intensive_Computing_report_final.pdf

(2) "The ASCAC Facilities Statement," March 21st, 2013: http://science.energy.gov/~/media/ascr/ascac/pdf/reports/2013/ASCAC_facilities_statement_final.pdf

Additional reference of interest:

(3) "The Opportunities and Challenges of Exascale Computing", Fall 2010: http://science.energy.gov/~/media/ascr/ascac/pdf/reports/Exascale_subcommittee_report.pdf

Roscoe C. Giles Short Biographical Statement

Roscoe Giles is a professor in the Department of Electrical and Computer Engineering at Boston University (BU) and Deputy Director of the Boston University Center for Computational Science. Giles' research focuses on the application of high performance and parallel computing to physics and materials problems.

Giles has served in a number of leadership roles in the community, including as a member of the DOE Advanced Scientific Computing Advisory Committee (ASCAC), the board of Associated Universities Inc, chair of the Boston University Faculty Council, and General Chair of the SC Conference in 2002.

Professor Giles has also worked to prototype and build computational and educational infrastructure that will enable broad participation of scholars and students in high-performance computing. For his work in increasing the participation of minorities in computer and computational science, in 2000 Giles received the A. Nico Habermann award from the Computing Research Association.

Professor Giles received his PhD in Physics from Stanford University in 1975 - the first in a long line of African-American PhD's in Physics and Applied Physics at Stanford. He received a Bachelor of Arts in Physics from the University of Chicago in 1970.

Mr. WEBER. [Presiding] Thank you, Dr. Giles.

Now, I recognize Dr. Stevens to present his testimony. Turn your mike on.

## TESTIMONY OF DR. RICK STEVENS, ASSOCIATE LABORATORY DIRECTOR FOR COMPUTING, ENVIRONMENT AND LIFE SCIENCES, ARGONNE NATIONAL LABORATORY

Dr. STEVENS. Oh, thanks. Thank you. Madam Chair, Ranking Member Swalwell, Members of the Subcommittee, I appreciate this opportunity to talk to you about the future of high-performance computing research and development and about the importance of U.S. leadership in the development and deployment of exascale computing.

In my own work at Argonne and the University of Chicago I split my time between trying to advance high-performance computing architectures and systems and doing research on how to do computational genomics in the pursuit of problems in energy, the environment, and infectious disease. And those projects have given me insight not only on the underlying technology but on the impact of applications.

I believe that advancing American leadership in high-performance computing is vital to our national interest. High-performance computing is a critical technology for the Nation, and it is also the underlying foundation for advancing progress in modeling and simulation and big data. It serves both of these needs. It is also needed by all branches of science and engineering for forward progress. It is used more and more by U.S. industry to maintain a competitive edge in the development of new products and services, and it is emerging as a critical policy tool for government leaders who can rely on simulations to add insight to policy or technical decisions.

Today, the United States is the undisputed leader in the development and use of high-performance computing technologies. However, other nations are increasingly aware of the strategic importance of HPC and are creating supercomputing research programs that challenge our leadership.

Japan has significant programs for over a decade in this area. They have fielded large-scale machines that are comparable to the machines in the United States. But China is emerging as a serious player as well and Europe has been investing in revitalization of their own high-performance computing sector. So we now have at least three sectors on the planet besides the United States making serious progress.

All have set their sights on the development of machines that are 1,000 times faster than those most powerful machines today. Everyone is looking at exascale. And achieving this goal is important. The drive to exascale will have a sustained impact on American competitiveness. It gives companies and researchers the means and the impetus for developing new processes, new services, and new products.

For example, we need increased compute power to enable first principle simulations of materials for energy storage that would give us access to a potential 500-mile battery pack for electric cars. We want to build end-to-end simulations of advanced nuclear reactors that are modular, safe, and affordable. We want to revolutionize small business manufacturing and digital fabrication and put in place a digital supply chain that would potentially revolutionize the economy in the United States.

We want to model controls for power grids that have significant amount of renewable energy, and we want to increase the resolution of climate models to provide more details on regional impacts. And finally, we want to create a personalized medicine that can incorporate an individual's genomic information into a specific customized plan for prevention or treatment of disease.

All of these challenges require machines that are thousands of times faster than the current machines. The development of practical exascale system, however, will also mean affordable petascale systems and broad deployment, broad accessibility.

The DOE Office of Science supercomputer centers at Argonne, Berkeley, and Oak Ridge are currently oversubscribed by at least a factor of three. This means that not all of the science that we could be doing on these machines is getting done. With current funding levels, these systems can only be upgraded about once every four to five years. And at current research—at current levels of research investment, the U.S. vendors are not likely to reach an exascale performance level that we can afford to deploy until considerably after 2020. This is a problem for us if we want to maintain our leadership.

Both China and Japan are working on plans to reach the level by 2020 or before. Japan is building a $1.1 billion investment program aiming to deploy exascale machines by 2020, and China has announced a goal to reach exascale before 2020. China is aggressively spending on infrastructure for supercomputing and succeeding in deploying large-scale systems rivaling the largest systems deployed in the United States. It is widely expected they will regain lead on this capability this year, although their designs are mostly based on incorporating U.S. components. In the future, they plan to deploy systems based on Chinese components.

I have been working since 2007 building a plan with my colleagues at the laboratories, academia, and DOE, and we identified five hurdles that we must cross in order to reach exascale. We have to reduce systems powered by a factor of 50; we must improve memory performance and cost by a factor of 100; we must improve our ability to program these systems; we must increase the parallelism in our applications; and we must improve reliability. These are not simple tasks but these are very important if we are to reach this goal. And I believe we have a duty to move as swiftly as we can on this objective.

Thank you. I would be more than happy to answer your questions. Thank you.

[The prepared statement of Dr. Stevens follows:]

## Statement of Rick Stevens

Associate Laboratory Director
for Computing, Environment and Life Science
Argonne National Laboratory

Before the
**Subcommittee on Energy**
**of the Committee on Science, Space and Technology**
U.S. House of Representatives
May 22, 2013

Thank you Chairman Lummis, Ranking Member Swalwell, and Members of the Subcommittee. I appreciate this opportunity to talk to you about the future of high performance computing research and development, and about the importance of continued U.S. leadership in the development of Exascale computing.

I am Rick Stevens, the Associate Laboratory Director responsible for Computing, Environment, and Life Sciences research at Argonne National Laboratory, one of America's first and largest multipurpose science and engineering laboratories. My research focuses on finding new ways to increase the impact of computation on science – from advancing new architectures to developing large-scale applications for computational genomics targeting infectious disease research, energy and the environment. I also am a Professor at the University of Chicago in the Department of Computer Science, where I hold senior fellow appointments in the University's Computation Institute and the Institute for Genomics and Systems Biology. In addition, I am the co-founder of the University of Chicago/Argonne Computation Institute, which advances interdisciplinary projects involving computation.

I believe that advancing American leadership in high-performance computing (HPC) is vital to our national interest. High-performance computing is a critical technology for the nation, it is needed by all branches of science and engineering, and it is a critical policy tool for government leaders. More importantly, its availability is a pacing item for much of science and for many technological developments on the horizon. Today the United States is the undisputed leader in the development of high-performance computing technologies both hardware and software. However, the nation that succeeds in leading in high-performance computing and large-scale data analysis for the long term will gain an insuperable competitive advantage in a wide array of sectors, including advanced manufacturing, energy, health care, space, defense, transportation, education, basic science and information technology.

The next stage of international leadership in computing is the development of systems capable of 100x to 1000x times more performance on both modeling and simulation tasks, as well as the analysis of ever-increasing large and complex datasets from commerce, the internet, healthcare and science. This next stage of development is known as Exascale (i.e. $10^{18}$ operations per second or $10^{18}$ bytes of storage). The availability of Exascale computing capabilities will improve our economic competitiveness, giving companies the capability to use modeling and simulation at unprecedented speed and fidelity to spur creativity and speed development of innovative new products. Just as importantly, Exascale computing will be an extraordinarily powerful scientific tool, enabling more accurate predictive models and facilitating analysis of massive quantities of data, making it possible for researchers to tackle and solve problems where experiments are dangerous, impossible, or inordinately costly. For example, it has been remarked that practical high-throughput genome sequencing would have been be impossible without high-performance computing. The rise of 3D printing and digital fabrication, which have the potential to transform our economy, likewise relies on computing in fundamental new ways.

Today, high-performance computing plays a central role in our nation's scientific, technical and economic enterprise, and the global community looks to our fiercely competitive HPC industry for cues on where this vital technology is heading next.

Visitors from around the world visit the DOE National Laboratories to learn about how we develop, deploy and use large-scale computers to attack the hardest problems in science and engineering. The DOE National Laboratories are recognized worldwide as thought leaders in the development and use of the largest systems, they are developers of the most advanced systems software and tools, and they pioneered the idea of open source math libraries long before open source was recognized as a key to rapid progress in software.

In the DOE Office of Science, the Advanced Scientific Computing Research program supports the operation of four national scientific user facilities:

- Energy Sciences Network (ESnet), a high-speed network serving thousands of Department of Energy researchers and collaborators worldwide. Managed and operated by the ESnet staff at Lawrence Berkeley National Laboratory, ESnet is one of the DOE's most widely based and successful cooperative efforts, providing direct connections to more than 30 DOE sites at speeds up to 10 gigabits per second. ESnet allows scientists to access unique DOE research facilities and computing resources independent of time and location with state-of-the-art performance levels.

- Oak Ridge National Laboratory Leadership Computing Facility (OLCF) serves all areas of the research community through programs such as the Department of Energy's Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program. Last year, INCITE awarded nearly a billion processor hours on OLCF's Jaguar system to 35 projects from universities, private industry, and government research laboratories.
- Argonne Leadership Computing Facility (ALCF) enables research spanning a diverse range of scientific areas, providing expertise and assistance in support of user's projects to achieve top performance of applications and maximize resources. The ALCF has allocated over 2.1 billion processor hours via INCITE to over 37 projects for the current year.
- National Energy Research Scientific Computing (NERSC) Center, a division of Lawrence Berkeley National Laboratory located at the University of California, Oakland Scientific Facility, serves over 3,000 scientists throughout the United States each year. These researchers work at DOE laboratories, universities, industrial laboratories and other Federal agencies. Computational science conducted at NERSC covers the entire range of scientific disciplines, but is focused on research that supports DOE's missions and scientific goals. The NERSC staff delivers critical computing resources, applications and information enabling users to optimize the use of their computer time allocation.

Over the years, the National Laboratories have formed important collaborations and partnerships that have strengthened our nation's computing leadership. Argonne and Lawrence Livermore National Laboratory (LLNL), along with industry partner IBM, pioneered a highly successful co-design approach to develop Petascale computers for grand challenge science resulting in the BG/L, BG/P and BG/Q systems. Additionally, Oak Ridge National Laboratory (ORNL), Argonne and LLNL have developed a partnership, known as CORAL (Collaboration of Oak Ridge Argonne Livermore), to develop the next round of pre-Exascale machines. Our success has been made possible by strong public/private partnerships, supported by thoughtful ongoing investments by our leaders in Washington.

In recent years, other nations have challenged our present dominance and future information superiority by making substantial, sustained investment in high performance computing research and systems deployments. Today, the Department of Energy's Titan supercomputing system at ORNL tops the current list of the world's most powerful supercomputers, and U.S. machines hold five of the top 10 spots. (Of those, three are sited at DOE National Laboratories: Titan at Oak Ridge, Sequoia at Lawrence Livermore, and Mira at Argonne.) But other nations are increasingly aware of the strategic importance of HPC and are creating supercomputing research programs that rival our own. (Most notably, China has announced plans to build more than a dozen supercomputing centers, with an announced goal of reaching Exascale capability by 2018.) Japan is planning a next-generation supercomputing project with an estimated budget of 110 trillion yen, and Europe has established PRACE (Partnership for Advanced Computing in Europe) to advance high-performance computing and to re-establish an HPC industry in Europe.

Right now, our competitors are relying primarily on American technology to create these powerful machines. But increasingly, other nations are developing the expertise and technology to build supercomputers that could rival or even surpass American-made high-performance computing systems.

I have been working over the past five years with my colleagues in the National Laboratory system, academia and private industry to develop an integrated, ambitious plan to keep the United States at the forefront of high-performance computing, a plan that will enable us to reach the Exascale performance level on critical DOE applications by the end of this decade. Our plan is based on a continuing, long-term partnership between computer hardware manufacturers and laboratory based teams of scientists, mathematicians and engineers who develop mathematical models, write parallel computer programs that implement those models on hundreds of thousands of processors, and develop the software that enables the computers operate efficiently. These teams are supported in turn by hundreds of scientists and engineers who develop and operate the machines.

Already, our joint efforts have led to major advances in nearly every area of computational science. For example:

- Researchers at the University of Chicago are using the Argonne Leadership Computing Facility's resources to pursue transformative breakthroughs in materials for batteries and fuel cells. At present, much of material science discovery still relies on the traditional "Edisonian", intuition-based "trial and error" approaches. The ALCF enables multi-scale modeling of charge transport processes in materials relevant to fuel cell and battery technologies – an effort that could have a significant impact on chemistry and material science while improving our nation's energy security. This research will have important implications for the work of the Joint Center for Energy

Storage Research, the new Battery and Energy Storage Hub headquartered at Argonne and funded through the DOE's Office of Basic Energy Sciences.

- Researchers at LLNL, working with colleagues at the IBM T. J. Watson Research Center in New York, used Sequoia to develop a highly scalable code, called Cardioid, that replicates the electrophysiology of the human heart. By accurately simulating the activation of each heart muscle cell and the cell-to-cell electric coupling, Cardioid will give researchers groundbreaking new insights into serious arrhythmia, a heart malfunction that causes 325,000 deaths in the United States each year.
- Pratt & Whitney is working with the Argonne Leadership Computing Facility (ALCF) to perform "virtual testing" of jet engine designs. Engine improvements based on these computer simulations have contributed to 15% improved fuel burn, saving $1.5 million per airplane each year while reducing carbon emissions by 3,000 tons. These design upgrades were made possible by ALCF researchers who were able to implement code improvements that resulted in a 10-fold performance advance in the Pratt & Whitney simulations.
- The Consortium for Advanced Simulation of Light Water Reactors (CASL), a group of national laboratories, universities and industry partners led by ORNL, is using the Titan system to develop a virtual nuclear reactor simulation toolkit that will model the interior of a nuclear reactor and gain information to aid design of next-generation reactors, and to improve the safety and performance of reactors currently in use.
- Procter & Gamble researchers used the ALCF to investigate the molecular mechanisms of bubble formation in foams, performing unprecedented computer simulations of the dissolving of soap and foaming of suds. By better understanding the process of "sudsing," Procter & Gamble can evaluate new materials more quickly and create better, less expensive consumer products, foods, and materials for fire control.
- A group at the National Center for Atmospheric Research is collecting weather data from the past 150 years to achieve a more detailed understanding of climate, with a goal of predicting severe events such as Superstorm Sandy. Using NERSC, the team is analyzing massive datasets that use the history of global weather patterns to validate climate and weather models and enable more accurate forecasting.
- Researchers from the University of California, San Diego are using the ALCF to target Parkinson's disease, a progressive and devastating disorder of the nervous system that affects more than 2 million Americans and costs an estimated $25 billion annually in medical treatments and lost productivity.

> Supercomputers accelerate drug development research by making it possible for medical researchers to model the proteins that cause Parkinson's and rapidly identify possible therapeutic interventions.

The HPC systems now being used by these researchers, and by thousands more nationwide, offer speeds that are measured in Petaflops – a quadrillion sustained floating-point operations per second. These systems took roughly five years from design to deployment, and have productive life spans of about five years. At present, the HPC industry in the United States is capable of delivering new frontier-level systems every three to four years, through the efforts of a complex and interwoven "ecosystem" that brings together industry, national laboratories and universities. As we look to the future, however, it is clear that the domestic demand for HPC access is outpacing supply, and the gap between our HPC capabilities and our national needs will widen dramatically as more and more leaders in private industry come to understand how high-performance computing can accelerate and support their R&D programs.

To better understand our nation's HPC needs for the future, DOE held twelve discipline-oriented workshops over the past few years to determine the scientific and engineering opportunities and priorities for HPC for this decade and beyond. These workshops brought together more than 1,200 scientists and engineers from around the country and around the world to discuss the central problems in each field, to identify important questions that might be addressed through advances in computation, and to determine the performance requirements for those computations.

The DOE workshops led to identification of pressing questions and important problems that will require Exascale computing capability to solve. For example, we want to use first principles to design new materials that will enable a 500-mile electric car battery pack. We want to build end-to-end simulations of advanced nuclear reactors that are modular, safe and affordable. We want to add full atmospheric chemistry and microbial processes to climate models, and to increase the resolution of climate models to provide detailed regional impacts. We want to model controls for an electric grid that has 30 percent renewable generation. We want to create personalized medicines that will incorporate an individual's genetic information into a specific, customized plan for prevention or treatment. In basic science, we would like to study dark matter and dark energy by building high-resolution cosmological simulations to interpret next generation observations. All of these challenges require machines that have more than 100x-1000x the processing power of current supercomputers.

The challenges of creating Exascale systems will require significant changes both in the underlying hardware architecture and in the many layers of software above it. To address those challenges, researchers at the National Laboratories and throughout the HPC community are working together to design and develop architectures, operating systems, runtime, storage, languages and libraries, and application codes. We are thinking about new algorithms that will be "Exascale ready," and we are building co-design teams that bring together computer scientists, mathematicians and scientific domain experts to work on solving these problems. We also are working with existing applications communities to guide their decisions about rewriting codes for near-term opportunities so their work will transition smoothly to Exascale systems. As we move ahead, this process will require increasing communication and continued refinement of ideas among a larger-than-normal group of stakeholders. In the past, architects could use rules of thumb from broad classes of applications or benchmarks to resolve design choices. But Exascale will require an aggressive co-design process that makes visible to the whole team the costs and benefits of each set of decisions on the architecture, software stack, and algorithms.

As we transition to Exascale, the hierarchy of systems will largely remain intact. That means the advances needed for Exascale will be hugely beneficial to Petascale computing and so on down the computing space. So, for example, if improved energy efficiency and better software solutions for resilience are developed as part of Exascale research, then it becomes possible to build Petascale computers out of less expensive components.

We have identified five major hurdles that must be overcome if we are to achieve our goal of pushing the computing performance frontier to the Exascale by the end of the decade:

- We must reduce power consumption by at least a factor of 50.
- We must increase the parallelism of our applications software and operating systems by at least a factor of 1,000.
- We must develop new programming methods to increase dramatically the number of programmers that can develop parallel programs.
- We must improve memory performance and cost by a factor of 100.
- We must improve systems reliability by at least a factor of 10.

Let me now address each of these major challenges in a bit more detail.

**Power Consumption**

The majority of performance improvements are achieved by increasing the number of processors, which requires an increase in power consumption. Today's most powerful supercomputer systems require a few megawatts of electricity; with current power prices at roughly $1 million per megawatt year, it is clear that we must find ways to reduce power consumption of both processors and memories by 50x over the next decade if we are to create Exascale systems that are affordable to operate. At present, industry is on track to lower power consumption by approximately 5x, but additional research is needed to lower it by another factor of 10, resulting in Exascale systems that will consume approximately 20 megawatts.

To reach our power consumption goals, we must redesign the processors and the memories that feed the processors, each of which contribute about equally to the power used. We will also need to replace copper wiring that supports communications between the processors with lower-power optics. Preliminary research work has indicated that major changes in processor design could save up to 20x on power consumption, so there is reason to be optimistic that we can achieve the necessary energy efficiency. These improvements will not only make it feasible to build and operate Exascale supercomputers for science; they also will have a positive cascading impact on energy consumption across the entire information technology (IT) sector. Incorporating these advances in products in the broad market would improve power consumption in large-data centers and extend battery life in laptops and handheld devices, making computer systems more power-efficient and therefore more affordable, while reducing their environmental impact. Given that an estimated 5 percent of global energy consumption, and of global carbon dioxide emissions, is attributed to computing services, energy efficiency improvements of this magnitude could have a significant impact on the environment.

**Increasing Parallelism**

In the past, supercomputer performance gains have been achieved by improving the speed of individual processors. Now, however, we have reached practical limits in both features sizes and power consumption, which means nearly all future performance improvements will come through increasing the number of CPU "cores," or processing units, that can be applied to a single problem. Today, your laptop has a few cores; the biggest systems have approximately 1,600,000 cores. In the future, we expect this number to grow linearly with overall performance. So to reach our goal of

100-fold to 1,000-fold improvement in performance, we will need to increase the number of cores by approximately 100x -1,000x, resulting in systems of 100,000,000 cores or more. This transition, from faster single processors to increased numbers of processors, will require an equal increase in the parallelism, or concurrency, of our applications and systems software. In response to this challenge, the community has been working hard to develop parallel programming technology and tools, parallel programming languages and new parallel algorithms all aimed at enabling this transition. This shift to increased concurrency as a means to improve performance will impact all sectors of IT, from business servers, to desktops and laptops, to cellphones and personal electronics. HPC will lead the way, but the transformation will be ubiquitous, impacting all forms of computing.

### New Programming Models

Today, programmers who develop codes for scientific applications must indicate to the computer precisely how information is divided among the processors in the system, and how the different parts of the problem are communicated between processors to enable all the processors to work together to solve the problem – a challenge roughly equivalent to writing out detailed instructions for managing all the traffic in New York City. The tool most commonly used for this part of the programming task is a language extension called the Message Passing Interface (MPI). It works well; nearly all-scientific groups use MPI to enable their programs to run on today's massively parallel systems. In the future, however, we would like to make it easier for programmers to develop parallel software, eliminating the need to explicitly manage millions of processors by developing new parallel programming languages and tools that will allow programs to be written at a higher level, making programmers more productive while increasing the scope of applications for highly parallel systems. This is an active area of computer science research and one that will impact industry broadly; indeed, Microsoft, Intel, IBM and others are now working closely with U.S. universities and national laboratories to address this challenge.

### Memory

To solve larger and more complex problems, computational scientists need computers that offer increased memory as well as increased speed. In fact, memory capacity is often the limiting factor in determining whether it is possible to solve a particular problem. For example, in climate modeling we might want to increase the resolution (by using a finer mesh) of the simulation to resolve details relevant to regional impacts – a challenge that requires more grid points, and therefore more memory.

Ideally scientists would like memory capacity to grow proportionally with the number of processors. However, that goal is not feasible in the near future, given that increasing the number of processor cores is a much simpler task than increasing memory capacity. We also need memory that is faster, to communicate with the processor at a higher throughput. Because increasing memory bandwidth consumes even more power, new ideas are needed to simultaneously improve memory performance and reduce memory power consumption. With balanced investments, it appears that we can increase memory capacity by approximately 100x as we move toward Exascale. However, it should be noted that the United States faces serious international competition in this challenge; of the top 10 global suppliers of memory, only one – Micron – is U.S.-based.

**Reliability**
As we build ever-larger systems, it becomes necessary to improve the reliability of every component; otherwise, the risk of diminished overall system reliability increases with the addition of each new component. While all modern large-scale computers have sophisticated mechanisms in place to identify and manage failures, these systems must be improved to ensure that our future systems will stay up long enough to do useful work. Overall reliability can be improved through new hardware designs that make fewer assumptions about individual components being failure-free, and through new ideas in software that can identify and isolate failed components, enabling the system to stay up while users' jobs are restarted on different parts of the machine.

In summary, Exascale computing represents a critical technological and economic opportunity for the United States. Right now, the HPC global market is estimated at $10 billion, and that market is expected to grow to $40 billion over the next decade. At present, we lead the world both in the development of HPC and in the use of HPC for advancing science and engineering, and we are working hard to achieve the next great milestone.

I believe that we can – and must – continue American leadership in HPC, to Exascale and beyond. But to fulfill the promise of Exascale, we must make sure that our efforts to develop the next-generation supercomputer are matched by increased outreach to American industry – to identify new industrial partners, to show them how HPC can support their work, and to address any gaps in industry-specific technologies. Exascale computing will create enormous opportunities in modeling and simulation, but it may have even more impact on the large-scale data problems at the heart of many enterprises. So as we go forward, we must continue to work with our partners in industry to identify sectors where big data will enable smarter, faster decisions and outcomes, and where highly accurate modeling and simulations could lead to better results – from choosing the most effective treatment for an individual patient with breast cancer, to improving car engine combustion efficiency, to creating new energy technologies that will protect our environment and our national security.

All of us who are working in this community believe that Exascale supercomputing will be a reality by the end of this decade, and that the next-generation machines will provide tremendous benefits in terms of scientific impacts, national security and international economic competitiveness. We also understand that reaching Exascale will require many breakthroughs in science and engineering, supported by a strong public/private sector partnership.

Ultimately, however, this is a race, not against our international competitors, but against ourselves. Exascale computing is necessary to the achievement of our most urgent goals in energy, in medicine, in science and in the environment. I believe we have a duty to move as swiftly as we can, and I sincerely hope that we will seize this opportunity to maintain and extend our record of success in HPC over the next decade by making a national commitment to Exascale computing.

Thank you. I would be happy to answer any questions you or other members of the committee may have.

**Rick L. Stevens**

Rick L. Stevens is Associate Laboratory Director of Computing, Environment, and Life Sciences at Argonne National Laboratory, which is the U.S. Department of Energy's (DOE's) oldest lab for science and energy research. He heads Argonne's computational genomics program and co-leads the DOE laboratories planning effort for exascale computing research. He is a professor of computer science at the University of Chicago (UChicago) and is involved in several interdisciplinary studies at the Argonne/UChicago Computation Institute and at the Argonne/UChicago Institute for Genomics and Systems Biology, where he holds senior fellow appointments.

Stevens is co-principal investigator, chief technical officer, and chief architect of the DOE Systems Biology Knowledgebase project, an emerging software and data environment designed to enable researchers to collaboratively generate, test and share new hypotheses about gene and protein functions, perform large-scale analyses on a scalable computing infrastructure, and model interactions in microbes, plants, and their communities. Stevens is also co-principle investigator for the NIAID Bioinformatics Resource Centers program where his group has developed computational tools and genomics databases to support infectious disease research.

Stevens is interested in the development of innovative tools and techniques that enable computational scientists to solve important large-scale problems on advanced computers. His research focuses on two principal areas: high-performance computer architectures, and computational problems in the life sciences. In addition to his research work, Stevens teaches courses on computer architecture, collaboration technology, parallel computing, and computational science. He serves on many national and international advisory committees and still finds time to occasionally write code and play with his 3D printer.

Mr. WEBER. Thank you, Dr. Stevens.

I recognize Ms. Crawford for her testimony.

## TESTIMONY OF MS. DONA CRAWFORD, ASSOCIATE DIRECTOR FOR COMPUTATION, LAWRENCE LIVERMORE NATIONAL LABORATORY

Ms. CRAWFORD. I thank you. I thank Chair Lummis and I thank you, Mr. Vice Chairman Weber and Ranking Member Swalwell, for inviting me to be here today. I ask that my full statement as submitted to the Committee made part of the hearing record, and if I may, I will summarize.

Mr. WEBER. Without objection, so ordered.

Ms. CRAWFORD. I am Dona Crawford, Associate Director of Computation at Lawrence Livermore National Laboratory. I will shorten that by saying LLNL or Livermore. Livermore is a national security laboratory of the National Nuclear Security Administration of the Department of Energy and home to Sequoia, one of the fastest computers in the world.

Livermore has the responsibility for maintaining the safety, security, and effectiveness of the Nation's strategic nuclear deterrent through the Stockpile Stewardship Program. High-performance computing has been a core competency of the lab to meet this mission need since over 60 years. In fact, the NNSA labs, working in close partnership with U.S. HPC industry, were at the forefront of the last revolutionary design shift in HPC computer architectures and applications development. That is the foundation of today's HPC systems.

Over the past 20 years, the NNSA labs learned many valuable lessons, including how to best structure R&D efforts to develop computing architectures that meet our demanding mission requirements while cost-effectively leveraging market-driven technology within industry. These lessons are very valuable in our efforts to develop exascale computing.

I applaud the Committee for its determined efforts to sustain U.S. leadership in this vitally important and increasingly competitive arena of high-performance computing. It is imperative that the United States embark on an R&D program to develop new technologies and computer architectures to support exascale computing.

My main point of emphasis today is straightforward. This pursuit must be a joint Office of Science/NNSA effort working in tandem through partnership with U.S. HPC industry to ensure system architectures that meet Office of Science and NNSA mission requirements. Working together, the Office of Science and NNSA have combined scarce resources and have already initiated a number of R&D efforts and contracts with industrial partners but lack the resources to invest at the magnitude necessary to assure success over the next decade.

Due to the technically challenging nature of developing exascale supporting technologies in computing capability, it is vitally important to ensure there are competitive teams each with Office of Science and NNSA laboratories partnered with U.S. HPC industry collaborators. Equally important is the development of an integrated strategy and program management plan.

Current U.S. leadership in HPC is a direct result of the Nation's investment in computational capability to support the Stockpile Stewardship Program. U.S. HPC investment has provided significant computing capability to maintain the U.S. nuclear deterrent and this computing capability enables us to simulate in 2–D at high resolution and high physics fidelity or simulate in 3–D at low resolution. Today, we cannot simulate in 3–D at high resolution and high physics fidelity which will be required for the stockpile mission needs. Therefore, a new architecture enabling exascale computing is required for the NNSA mission.

This will not be easy. Development of exascale-class systems cannot be achieved through a straightforward refinement of today's technologies. Surmounting multiple technical issues will require sustained research and development effort. But there is no doubt exascale computing will yield valuable benefits to near-term mission requirements, as well as to U.S. economic competitiveness.

Over the last two decades, supercomputers have transformed the way the world conducts scientific research and has enabled discovery and development across a broad set of disciplines. In a 2008 U.S. Council on Competitiveness report, the Council states, "supercomputing is part of the corporate arsenal to beat rivals by staying one step ahead of the innovation curve. It allows companies to design products and analyze data in ways once unimaginable."

In one example, Livermore is leveraging its HPC capabilities in the California Energy Systems for the 21st Century Initiative. The California Public Utilities Commission and state investor-owned utilities are collaborating with Livermore to improve and expand energy systems to meet our future energy needs. The owners, operators, regulators, and a joint team of technical experts will use the Nation's most advanced modeling simulation and analytical tools to gain unprecedented insight and generate new data to reduce risk and inform solutions to issues facing 21st-century energy systems such as renewable energy integration and use of smart grid technology.

There are many other examples that highlight the importance of supercomputing and reinforce the value of maintaining U.S. HPC leadership. For now, let me close again by saying thank you and I look forward to working with the Committee to ensure continued U.S. HPC leadership.

[The prepared statement of Ms. Crawford follows:]

# EXASCALE COMPUTING CHALLENGES AND OPPORTUNITIES

Hearing of the House Science, Space, and Technology Committee
Subcommittee on Energy
U.S. House of Representatives
May 22, 2013
Dona L. Crawford, Associate Director for Computation
Lawrence Livermore National Laboratory

## OPENING REMARKS

Thank you, Madam Chair, Ranking Member Swalwell, and Members of the Committee. I am Dona Crawford, Associate Director of Computation at Lawrence Livermore National Laboratory (LLNL). I welcome the opportunity to provide my perspective on high performance computing (HPC). There are major opportunities and challenges associated with developing exascale computing, the next generation of HPC capability. At the same time, exascale computing is critically needed to support national security priorities, advance science and technology, and enable greater innovation in U.S. industry.

I applaud the Committee for its determined support to maintain U.S. leadership in HPC. The U.S. has benefited immensely from the investments the nation has made in developing and applying HPC capability. The U.S. has enjoyed a long record of success and unparalleled leadership in computing for many years. The world took notice, and many other nations started investing in HPC and now strive to challenge U.S. leadership in this critical arena. It is imperative that we continue to make smart investments of our limited resources to maintain U.S. HPC leadership.

In view of today's fiscal constraints, we must apply the lessons learned in our past successes to strategically target investments to make crucial early steps on the road to exascale-level computing. To cost effectively and efficiently maintain U.S. leadership in HPC, the nation must build upon and leverage programmatic and technical approaches that established the U.S. as the leader in innovative HPC systems over the past half-century. In particular, next generation HPC must be developed through an integrated partnership of the Department of Energy's (DOE) National Nuclear Security Administration (NNSA) and the Office of Science (SC). Both technical and fiscal responsibilities must be shared, taking advantage of the core capabilities of the partners—and working closely with industry. This includes balanced investments in both the ongoing core HPC computing programs and breakthroughs necessary to achieve exascale HPC. Leading-edge HPC is vital for U.S. national security and science missions and to advance U.S. economic competitiveness goals.

I thank the Committee for its support for the development of next generation supercomputing and its recognition that the dialogue must integrate NNSA and its laboratories in this effort.

## NNSA's and LLNL's Role in and Reliance on Supercomputing

HPC has been a core competency of the nation's nuclear weapons enterprise from the birth of HPC in the 1950s and has been essential to the nation's ability to develop and maintain a nuclear deterrent. LLNL is one of the DOE/NNSA's national security laboratories with responsibility for maintaining the safety, security, and effectiveness of our nation's strategic deterrent. HPC has played an increasingly important role in LLNL's nuclear deterrence mission since the cessation of nuclear testing over 20 years ago and the creation of the Stockpile Stewardship Program (SSP).

Current U.S. leadership in HPC is a direct result of the nation's investment in computational capability for the support of the SSP. After the U.S. decided to forego underground nuclear testing, DOE Defense Programs (NNSA's predecessor) embarked on a focused effort to develop vastly improved computational capabilities, along with advanced experimental capabilities, as a foundation for the SSP that would provide the scientific basis for maintaining the nation's nuclear deterrent without nuclear testing. This was and continues to be a grand technical challenge.

Nuclear weapons are extremely complex devices, with thousands of components made from a variety of materials that must work together seamlessly to produce a nuclear detonation. As they age, nuclear weapons are subject to environmental conditions that pose a number of challenges that affect performance of components and the weapon itself. Plastics can break down and give off potentially destructive gases, metals can corrode and weaken, and coatings can deteriorate. Some materials may change properties unpredictably in response to the high radiation fields, fluctuating temperatures, and other environments to which nuclear weapons are subject.

Congress acknowledged the challenges associated with maintaining the nuclear deterrent without testing by creating in the mid-1990s an initiative within the SSP to rapidly develop substantially more powerful computational, simulation, and modeling capabilities. At the time, DOE was not an HPC technology driver as it had been in earlier decades. If the Accelerated Strategic Computing (ASCI) Program had not been formed and aggressively funded, the HPC industry would have continued to evolve toward serving its consumer base, toward business and industry focused server solutions of relatively limited capability—and later toward gaming applications—and would not have enabled the success of the SSP.

ASCI was highly successful and has since evolved into the current program, called Advanced Simulation and Computing (ASC). The NNSA laboratories (LLNL, Los Alamos National Laboratory and Sandia National Laboratories) and our industry partners worked to develop computer architectures that would enable the laboratories to run large-scale, high-fidelity simulations integrating data from past underground nuclear tests and experimental capabilities to continue to assess and certify the safety, security, and reliability of the nation's nuclear deterrent. This effort spearheaded the revolutionary design shift in HPC computer architecture and applications development that occurred over the last two decades. ASCI and ASC in the SSP—and later Office of Science—pushed the extreme limits of what was possible, and this ultimately led the way for more competitive business and industrial related computing and simulation. The research and development (R&D) has led to unprecedented advances in HPC and remarkably capable computing systems that now are becoming ubiquitous and are impacting scientific discovery and industrial competitiveness.

Time-urgent questions about the safety and reliability of the nuclear stockpile drove the DOE and its NNSA laboratories to invest in and develop supercomputers. These systems, developed by industry in response to national security demands, have dominated global HPC performance. In the last forty Top500 lists, the U.S. has held the top position 26 times, with NNSA HPC systems in 23 of these cases. Through investments in the ASCI and ASC programs, computing has become the single *integrating* element in assessment and certification of the U.S. nuclear weapons stockpile. As the global nuclear security landscape has evolved, these same computational tools are now being continuously applied to combating nuclear proliferation and bolstering counterterrorism—both nuclear and conventional.

The enormous success of ASCI/ASC has been a result of:

- sustained support for a Congressional initiative to develop HPC simulation as a pillar with the Stockpile Stewardship Program,
- leadership in DOE and NNSA driving the development of high-end computer architectures and associated simulation software designed for our unique mission requirements and targeted at specific challenges associated with the stockpile, and
- unprecedented level of cooperation between the national laboratories and industry to co-design/co-develop software and computational platforms required for the mission.

The importance of these three elements should not be underestimated. It is only through this initiative's combination of commitment, leadership, and cooperative R&D that U.S. computing made a revolutionary design shift in computer simulation required to ensure the safety, security, and reliability of the nuclear stockpile. As a result, computing capability used by NNSA's national laboratories increased over a million fold.

Each time the laboratories and their industry partners to develop a new generation of supercomputers, we discover new science that helps understand material phenomena and performance of the aging nuclear stockpile with a higher degree of accuracy. Today, the Sequoia machine at LLNL—a breakthrough ASC system with over 1.5 million processor units or cores, and 1.6 petabytes of memory—serves as a bridge between supercomputers of the past decade and exascale computers of the future. Indeed, it is arguably the first of the new era of daunting "many-core" computers. The machine's extraordinary capabilities are being used to improve models of weapons physics, particularly in the areas of hydrodynamics, radiation transport, and the properties of materials at extreme pressures and temperatures. In addition, Sequoia is able to run large suites of calculations designed to characterize uncertainties in weapon performance resulting from small variations in the weapon system and uncertainties in the physics models used.

Improved capabilities for uncertainty quantification (UQ) are essential for assessing the impact on performance of physical changes in aging weapons and for certifying stockpile Life Extension Programs (LEPs). Sequoia, with its 20-petaflop capability, can effectively address many stockpile issues through the use of UQ with two-dimensional (2-D) applications. However, the system can only provide "entry-level" capabilities to run suites of three-dimensional (3-D) weapons physics simulations for UQ. It remains the role of exascale-class systems to address the full breadth of issues that will arise as the stockpile ages, as significant findings are identified, and as even more advanced safety and security features are added. In the future, 3-D predictive UQ analysis will become essential for essentially all aspects of work.

### National Security Mission Need for Exascale

With the planned modernization of the stockpile and simultaneous decrease in both its overall size and composition, advanced computing and simulation will play an increasingly critical role. A thousand-fold improvement over today's modeling and simulation capability (exascale technology) is required over the long term to assure with confidence the safety, security, and performance of the nation's nuclear stockpile. These more capable computers are needed to run large suites of high-fidelity simulations to fully map out the impact of uncertainties.

Nuclear weapons are engineered 3-D systems with complex materials that change over time as they age. Today, we do not have the computing power to simulate weapons performance in 3-D at the required resolution and incorporating detailed physics and age-aware material models. Additionally, we do not have the computing power to conduct the tens of thousands of high-resolution 3-D simulations to quantify the variation on weapon performance taking into account uncertainties in our modeling capabilities.

Through our success in developing and applying advanced HPC, we have resolved the energy balance problem—one of the physics issues remaining unresolved at the cessation of the underground nuclear test program. Today's available technology allows us to simulate in two dimensions at high resolution and physics fidelity, or simulate in 3-D at moderate (not high) resolution, but today's available technology does not enable our weapons specialists to simulate at high resolution and 3-D simultaneously. We are looking forward to HPC systems that are also capable of bringing to closure the grand challenge of modeling the physics phenomena of boost.

Examples of the role exascale computing will play in the continued maintenance of the stockpile include:

- **Warhead assessment and certification of smaller stockpiles**: As the stockpile decreases in size, the performance of individual weapons becomes increasingly important. Higher fidelity, 3-D simulation of warheads including detailed representation of initial conditions, engineering features, safety features, and security features are required to ensure the safety and performance of each weapon. These simulations will each require between 0.5-10 exaflops.
- **Material aging, compatibility, and acceptance of modern efficient manufacturing processes:** As weapons continue to age, the complexity of issues to be resolved could increase exponentially. To address potential material related issues in the life-extended U.S. stockpile, more detailed weapons science calculations will be required. Simulations at increased resolution that capture real materials (micro-structure,

interfaces, kinetics) as opposed to simulations based on simple models of bulk material properties require between 0.5-100 exaflops.

- **LEPs**: Confidence in the assessment and certification of future life-extended warheads will be informed by suites of high-resolution integrated weapons simulations to quantify and bound uncertainties in performance. Detailed uncertainty quantification (UQ) will require routinely running between 1,000 and 100,000 simulations per study in order to rapidly converge LEP design options. Each of these UQ suites and the confirmatory steps will require between 10 and 1000 exaflops.
- **Safety and Security:** Enhancing weapon safety and security to address the 21st-century threat environment (including non-state actors) is just one example of a potential LEP goal. The development and certification of even more advanced safety, security, and use control (surety) features that can be embedded in a nuclear warhead as part of an LEP may require at least 10-100 exaflops.
- **Boost:** Boost—the process of boosting the fission yield of weapon primaries—is key weapons performance, not well understood, and among the most challenging physical phenomena to model. Greater computational power is needed to apply the improved physics models developed within the ASC *Predictive Capability Framework* (PCF) in large ensembles of weapons simulations. Large ensembles are required for rigorous UQ explorations using those models. We estimate 10-100 exaflops is required for resolving the largest known uncertainty associated with boost.

### Technological Challenges to Achieving Exascale Must be Overcome

The development of exascale-class systems cannot be achieved through a straightforward refinement of today's technologies. Surmounting multiple technical issues will require sustained research and development and some key breakthroughs.

Succeeding generations of microprocessors, standard in computers of every scale, have grown faster by increasing the speed and shrinking the size of transistors, effectively packing more calculations into every unit of time and space on a computer. But now transistors are reaching a lower limit in size and an upper limit in speed. Although individual transistors could be pushed to run faster, speeding up millions of transistors on a microprocessor would drive energy demands and operational costs to unsupportable levels. To make exascale computing practical, the electrical power requirements must be reduced at least ten-fold per floating point operation. Without this reduction, exascale computers would need hundreds of megawatts—enough to power a small city—at an unacceptable cost of hundreds of millions of dollars a year to pay the electricity bill. Developing a low power system encompasses changes to every component of the HPC system: memory (e.g. stacked memory), networks (e.g. optics), and processors (including accelerators).

Exascale systems will be comprised of tens to hundreds of millions of components. Calculations run on these systems will require tolerance for component failure as well as the management of up to a billion separate, but coordinated threads of execution—in short, the systems must become self-aware and must compensate, in real time, for failures.

In addition, memory and storage will be challenging. Science and nuclear stockpile applications require large amounts of memory per core. At exascale levels, data movement—not attaining a higher amount of flops—will be the performance bottleneck. This includes movement of data from memory to the processing unit as well as movement of data across the machine, from one processing unit to another processing unit. Current memory bandwidth and network technology is too slow, costly, and unreliable to support the millions of trillions of calculations per second required in an exascale machine. At the exascale level, more components mean higher failure rates. In addition, the faster the data moves, the more error-prone it becomes.

Finally, exascale computing will require new programming models that allow software developers to exploit unprecedented parallelism. Applications in the future may have to support upwards of a billion separate independently-executed instructions to efficiently use the hardware. This will require our scientists to find ways to break their problems into many more independent pieces than even today's largest computers support. They will have to re-think their entire solution approach to meet this challenge.

The U.S. has invested heavily over the last two decades to develop nuclear weapons simulation codes to maintain the deterrent. As we develop the next generation of supercomputers, it is of paramount importance that NNSA partners with SC to minimize disruption in the utility of current codes. If HPC technology were to evolve in a less controlled manner, there is the potential that our existing codes will become ineffective and possibly obsolete before we have time to rebuild them from scratch to operate efficiently on new architectures.

### R&D Program Required to Attain Exascale HPC

Attaining and harnessing exascale computing will require an integrated, decadal program of technology development and testing that balances technology push with applications pull. This requires the combination of a sustained federal investment, commitment to deliver hardware and software tuned for core missions, and strong partnerships between American industry and DOE laboratories tasked to define the exascale ecosystem. That ecosystem needs to support the combined national security, economic competitiveness, and science missions.

The overriding theme in an efficient and effective R&D program for exascale is co-design and co-development. Design and development are linked because mission application requirements influence computer architecture design and architecture technology constraints influence the formulation of algorithms and software in mission applications. To be effective in addressing stockpile mission requirements, future HPC systems must be capable of simulating the complex physics and material changes in a nuclear weapon. It is essential the NNSA laboratories are involved from the start in this daunting but necessarily seamless transition from today's systems to exascale. It is not sufficient to assume that next generation HPC computing developed for either mass consumers or for science applications will meet national security mission needs. The U.S. government must invest in the technology development effort in a way that ensures an architecture that meets national mission requirements.

The mission drive of NNSA and the leveraging of NNSA's outstanding systems-engineering track record in delivering leading edge HPC will serve to focus technology development on the path to exascale and lead to an architecture that will serve our national security mission requirements. I endorse SC's early focus on long lead-time research and development in advanced technologies. I strongly support NNSA's continued focus on investments that support the design and the delivery of well-balanced and well-architected HPC systems used to meet mission requirements. Both agencies today are working together to maintain their current HPC capabilities that are of vital importance to meeting near-term mission deliverables for NNSA and SC. But these agencies have very limited resources for advanced technology efforts required to achieve exascale. Thus, a balanced alliance between SC and NNSA in the next three years, supported by Congress, in pursuing R&D of balanced and innovative systems is a cost effective strategy.

The HPC R&D program required to meet NNSA mission needs in the next three years includes:

- Development of a scalable design for advanced system architecture at 100-200 petaflops. This effort couples NNSA systems requirements with industry expertise and best technologies. By funding advanced systems architectures and scalability, development and design through this integrated research and system delivery program, we can minimize the potential drift of next generation hardware away from our mission application needs.
- Acquisition of 100-200 petaflop systems in the FY2016-2017 time frame will enable prototype builds on the path toward exascale and will deliver NNSA required mission capability. Through these interim prototype systems, potential new technologies can be evaluated in the context of new architectures that utilize them to solve the various challenges. To maximize cost effectiveness of this approach, NNSA laboratories and SC laboratories are combining forces. Argonne National Laboratory, Oak Ridge National Laboratory, and LLNL as one team and Los Alamos, Sandia, and Lawrence Berkeley National Laboratories as another are partnering to pursue system acquisitions in 2017 and 2015, respectively. For cost effective risk reduction, an exascale initiative should pursue a minimum of two technology tracks in future acquisitions.
- Co-development of new exascale algorithms, applications, tools and runtime environments are needed by programmers to achieve maximum sustained performance on exascale systems. Today's largest systems have greater than a million cores, while an exascale system is expected to require billion-way parallelism. To be able to harness these new hardware technologies and architectures for our nuclear weapons mission, we must immediately begin to develop software tools that enable scalability, programmability, fault tolerance and code portability. As described earlier, this kind of investment and

involvement in co-design is essential to ensure we do not have to rewrite millions of lines of computer code. This would add substantial cost and threaten NNSA's ability to maintain the deterrent.

**Broad Benefits of Supercomputing Capability**

Over the last two decades, supercomputers have transformed the way the world conducts scientific research and enabled discovery and development across a broad set of disciplines from physics, chemistry, and bioscience to engineering. Simulation—the ability to virtually mimic physical phenomena with great accuracy—is now considered a peer to theory and experiment, and a pillar of the scientific method pioneered by Isaac Newton more than 300 years ago. HPC simulations have advanced medicine, energy, aviation, and manufacturing domains. In the 2008 U.S. Council on Competitiveness report "The New Secret Weapon," the Council states "Supercomputing is part of the corporate arsenal to beat rivals by staying one step ahead of the innovation curve. It allows companies to design products and analyze data in ways once unimaginable." Forefront HPC has moved from a tool developed and used by the national security laboratories like LLNL, to a critical tool for the U.S. science laboratories and Fortune 50 companies. The massive, complex simulations that run on today's HPC allow us to explore fields such as global food, water, and energy supplies, as well as tackle problems for which experiments are impractical, hazardous, or prohibitively expensive.

LLNL has certainly leveraged its high performance computing capability and applying it beyond the nuclear weapons program to other important mission area. New tools and expertise developed in other mission areas at LLNL can then be brought to bear on maintaining the nuclear deterrent. The multi-program utility of HPC capability and the joint benefits of applications to the weapons program and other mission areas are illustrated by three examples:

In 2010, the Department of Defense urgently tasked LLNL to develop an advanced conventional munition in record time. Based on the Mark 82 steel case form factor, LLNL combined a novel explosive design with a carbon fiber case that met the military need for a lightweight weapon that could deliver lethal effects with low collateral damage. Using LLNL's HPC simulation codes, originally developed in the nuclear weapons program, the team was able to accurately predict warhead performance under dynamic conditions and achieve desired strength properties by optimizing critical design features and tailoring fiber-composite winding patterns. This extensive use of HPC, allowing accurate design simulation, eliminated costly and lengthy iterated cycle of developing and testing prototypes. After only one set of proof tests, the BLU129/B advanced conventional munition was deployed in theater only 10 months after a Joint Urgent Operational Need was identified. The historic average of new munitions development is 4.5 years.

In the last year, Sequoia demonstrated its great scalability with a 3-D simulation of the human heart's electrophysiology. Using a code called Cardioid, created in a partnership between LLNL and IBM scientists, researchers are modeling the electrical signals moving throughout the heart. Cardioid has the potential to be used to test drugs and medical devices, paving the way for tests on humans. The code, running in nearly real time across the 20-petaflop system (an astonishing 60 beats a minute) predicted an arrhythmia that was known to occur with the injection of a drug. The fact that a calculation of this complexity ran at 59% of peak on Sequoia is astonishing. Cardioid demonstrated outstanding scalability and time-to-solution over 1200 times faster than previous state of the art, and the simulation is performing within 12% of real-time. The work showed the promise of advanced computing, but it also demonstrated the extreme level of specificity and technical acuity required to achieve this result. The insights gained and techniques employed by the code team are proving useful to Sequoia's national security applications.

Through DOE economic competitiveness initiatives, a number of American HPC centers and laboratories—including Lawrence Livermore—are making large-scale HPC resources available to U.S. companies both large and small. In LLNL's California Energy Systems for the 21st Century (CES-21) initiative, the California Public Utilities Commission and state investor-owned utilities are collaborating with LLNL to improve and expand energy systems to meet future needs. The owners, operators, regulators, and a joint team of technical experts will use the nation's most advanced modeling, simulation, and analytical tools to gain unprecedented insight and generate new data—information that can reduce risk and inform solutions to issues facing 21st-century energy systems, such as renewable energy integration and use of smart-grid technology. CES-21 will benefit from LLNL's extensive experience in national security supercomputing. Our expertise will be utilized to perform realistic and verifiable tests of how utilities will need to operate in the future.

## The U.S. Must Lead in High Performance Computing

The U.S. is entering the 21st century as the global leader in HPC, with the vast majority of high-end computer systems produced worldwide using U.S. technologies. However, our leadership is not undisputed. U.S. leadership faces an unprecedented challenge, principally from China, but also other global players.

China is steadily increasing its investment and marshaling its technological capabilities and its state-owned industries to quickly develop next-generation supercomputers. China has indicated its intent to use HPC for oil exploration, aircraft design, and weapon system design. China's 5-year roadmap demonstrates its commitment to assume and maintain a leadership position in HPC. A report from China's Academy of Sciences has stated that China has fast-tracked the development of exascale computing with the intention of using indigenous technology to field an exascale system by the end of this decade. They are funding a smart, systematic strategy to develop three different approaches to key hardware technologies such as chips, operating systems, file systems, and networks.

This represents a dramatic change from just a few years ago. Over the years, from 1993 until November 2001, the Chinese had zero, one or two systems on the Top500 list. Press reports called it a "Sputnik moment" for the U.S. following the release in November 2011 of the Top500 list of the world's fastest supercomputers. For the first time, Chinese supercomputers vaulted to the number one and three rankings on the list, displacing HPC systems at SC and NNSA labs. China has continued its focused investment in HPC and now has 72 systems in the Top500 list. This investment strategy has allowed them to surpass Japan and Europe in the numbers of systems on the Top500 list. Without dedicated U.S. investment, we risk ceding our leadership to China and eroding the U.S. HPC industrial base so important to our national security and overall economy.

Given our current global leadership position in HPC and our proven capability to innovate HPC simulation advances, the U.S. is positioned to continue HPC leadership if there is sustained R&D investment. The country that develops the key intellectual property for the next generation supercomputers will also control the high ground for standards-based decisions to be made over the coming decades: instruction sets, programming tools, I/O, visualization and protocols. This is a level of control that the U.S. has taken for granted due to its past domination in this arena. In addition, the low power-consuming technologies to be developed can be applied to the entire hierarchy of server offerings, down to the cell phone, providing almost unthinkable leverage and intellectual property value. Consider the position China could have in military operations if it controlled the low-power technology, next generation un-manned surveillance, and propagation of intelligence analysis to inform battlefield awareness in hand-held devices. Trusting U.S. nuclear weapon technology and even more sensitive national security technologies to Chinese-built HPC systems is untenable. We must invest in leadership HPC to support our national security interests and maintain a healthy U.S. HPC industrial base.

**Comments on Draft Exascale Legislation**

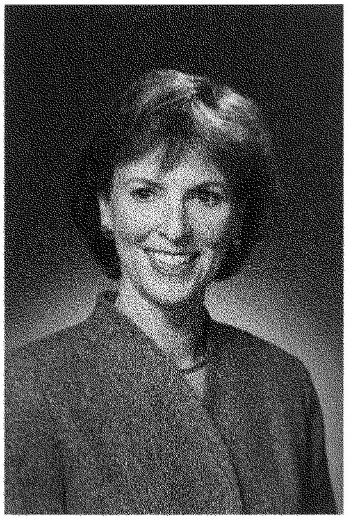
I am very encouraged by the demonstration of bi-partisan and bicameral support for continued U.S. investment in HPC and efforts to develop exascale-capable computing. This is a critical capability that is essential to U.S. competitiveness in multiple fields. I commend Representative Hultgren for his leadership in moving forward an authorization bill for the exascale R&D effort. It is imperative that the U.S. embark on an R&D program to develop new technologies and computer architectures to support exascale computing. This needs to be a joint Office of Science (SC)-NNSA effort leveraging the strengths and expertise of SC and NNSA laboratories in close partnership with the U.S. HPC industry. Recognizing the jurisdictional boundaries of Congressional committees, I note that the draft bill only authorizes funding for SC participation in the exascale effort. It is imperative that SC and NNSA move in tandem in this R&D effort for cost effective leverage of resources across the Department and to ensure systems architecture to meet SC and NNSA mission requirements. Due to the technically challenging nature of developing exascale supporting technologies and computing capability, it is vitally important to ensure that there are competitive teams of SC and NNSA laboratories partnered with U.S. HPC industrial collaborators. Equally important is the development of an integrated strategy and program management plan. I believe the draft bill, by and large, addresses these critical elements of a successful exascale R&D effort and opens a door for further discussion on how to most effectively structure a joint SC-NNSA exascale R&D effort.

I look forward to working with the Committee to ensure the U.S. retains its HPC leadership.

## Conclusion

Although supercomputing may lack the glamour of the space race, U.S. leadership is critically needed to achieve key national priorities. Failing to maintain U.S. HPC leadership has consequences beyond national security, reaching much further and more broadly into our economic future. Supercomputers have become a differentiating tool for discovery and innovation, with profound impacts on science, national security, and industrial competiveness. HPC at the exascale level will be a powerful lever to influence outcomes and foster prosperity and security as we face uncompromising competition in an uncertain world. In its 2008 report "The New Secret Weapon," the U.S. Council on Competitiveness said, to "out compute is to out compete." If we are to be partners in a world of global competition, I want us to come from a position of strength based on the best U.S. industry, academia, and the national laboratories have to offer. That is what put us and has kept us in the leadership role we enjoy today in supercomputing. It is imperative we now begin to push forward on the necessary technology to ensure a continued leadership position. The stakes are very high. A robust multi-year effort harnessing a partnership between DOE's Office of Science and NNSA with industry is key to national security and science, which underpins competitiveness.

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.



**Dona L. Crawford**
**Associate Director of Computation**
**Lawrence Livermore National Laboratory**

Ms. Crawford is responsible for a staff of roughly 900 who develop and deploy an integrated computing environment for petascale simulations of complex physical phenomena such as understanding global climate warming, clean energy creation, biodefense, and non-proliferation. This environment includes high performance computers, scientific visualization facilities, high-performance storage systems, network connectivity, multi-resolution data analysis, mathematical models, scalable numerical algorithms, computer applications, and necessary services to enable Laboratory mission goals and scientific discovery through simulation. An icon for the computing environment provided is the Advanced Simulation and Computing (ASC) Program's BlueGene/Q Sequoia machine (peak 20 quadrillion floating-point operations per second (PF)). This is among the fastest computers in the world

Ms. Crawford has served on advisory committees for the National Research Council and the National Science Foundation. She is Co-Chair of the CRDF Global Board, Co-Chair of the Council on Competitiveness High Performance Computing Advisory Committee, and a member of IBM's Deep Computing Institute's External Advisory Board. She is a member of the Institute of Electrical and Electronics Engineers (IEEE) and the Association for Computing Machinery (ACM), and participates in community outreach activities to promote math and science. She has a master's degree in operations research from Stanford University and a bachelor's degree in mathematics from the University of Redlands, California.

Dona was named 2005 Woman of the Year in Science in Alameda County and received the Computerworld Honors Award in 2006. In November 2010, Dona was featured as one of *insideHPC's* "Rock Stars of HPC." Her undergraduate alma mater, presented her with the "Alumni Career Achievement Award" in 2012, and HPCwire named her among their 2013 People to Watch.

Mr. WEBER. Thank you, Ms. Crawford.

Dr. Reed, I recognize you for your testimony.

## TESTIMONY OF DR. DANIEL REED, VICE PRESIDENT FOR RESEARCH AND ECONOMIC DEVELOPMENT, UNIVERSITY OF IOWA

Dr. REED. Thank you. Chair Lummis, Vice Chair Weber, Ranking Member Swalwell, Members of the Subcommittee, my name is Dan Reed and I am the Vice President for Research and Economic Development at the University of Iowa. Thank you for the opportunity to share my perspectives on exascale computing and to respond to your questions regarding the American High-End Computing Leadership Act.

Today, I would like to make four points regarding the exascale and high-performance computing program followed by a set of specific recommendations for the future. They are drawn from my nearly 30 years of experience in high-performance computing as a researcher, as an academic and corporate leader, as a Director of the National Science Foundation Supercomputing Center, and as a participant in national science and technology policy.

First of all, as others have noted, high-performance computing is unique among scientific instruments. It is distinguished by its universality as an intellectual amplifier. New, more powerful supercomputers and computational models yield insights across all scientific and engineering disciplines. Advanced computing is also essential for analyzing the torrent of experimental data produced by scientific instruments and sensors, but it is about more than science. With advanced computing, real-time data fusion, and powerful numerical models, we have the potential to predict the tracks of devastating tornadoes such as the recent one in Oklahoma, saving lives, and ensuring the future of our children.

My second point is that we face an uncertain future of computing and in particular high-performance computing leadership in this country. As others have noted, today, HPC systems from Oak Ridge, Lawrence Livermore, and Argonne National Laboratories occupy the first, second, and fourth place on the list of the world's fastest computers. From this, one might surmise that all is well, yet U.S. leadership in both deployed HPC capability and in the technologies needed to create future systems is under challenge.

Also, as others have noted, other nations are investing strategically in high-performance computing to advance national priorities. And the U.S. research community has repeatedly warned of the potential and actuality of eroding U.S. leadership in computing and in high-performance computing and emphasized the need for sustained and strategic investment. I have had the privilege of chairing many of those studies personally as a member PITAC, of PCAST, of National Academies' boards, and yet many of these warnings have been largely unheeded.

This brings me to my third point: the deep interdependence a basic research of vibrant U.S. computing industry and high-performance computing capability. It has long been axiomatic that the United States is the world leader in information technology. Our global leadership is not a birthright. As Andy Grove, the former CEO of Intel, noted in his famous aphorism "only the paranoid survive." U.S. leadership has been repeatedly earned and hard-fought based on continued Federal Government commitment to basic research, translation of that research into technological innovations, and the creation of new products by vibrant U.S. industry.

This brings me to my fourth point. Computing is in deep transition to a new era with profound implications for the future of U.S. industry and HPC. My colleague Mr. Stevens touched on many of the issues around energy management and low-power devices and they are key to this topic. U.S. consumers and businesses are an increasingly small minority of the global market for mobile devices and for cloud services.

We live in a post-PC world, as we all know, where U.S. companies compete in a global device ecosystem. Unless we are vigilant, these economic and technical changes could further shift the center of enabling technology R&D away from the United States with profound implications for our future HPC capability. Given this, what are my recommendations for the future? First and most importantly, we need to change our model for HPC research and deployment if the United States is to maintain its leadership. This must include deep and sustained interagency collaborations defined by a regularly updated strategic R&D plan and associated, verifiable metrics, commensurate budget allocations, and accountability to realize the plan's goals.

DOE's partners—it needs the National Science Foundation, the Department of Defense, NIST and NIH, and other agencies to fulfill their important and complementary roles to DOE as engaged partners and supporters of basic research in technology development. We also need long-term industry engagement.

Second, advanced HPC deployments are crucial, but the computing R&D journey is as important as any single system deployment. A vibrant U.S. ecosystem of talented and trained people and technical innovation is the true lifeblood of sustainable exascale computing.

Finally, we must balance and embrace dual-use technology R&D supporting both high-performance computing and ensuring U.S. industry competitiveness. Neither HPC nor big data R&D can be sacrificed to advance the other, nor can hardware R&D dominate investments in algorithms, software, applications, and people. All are crucial.

Finally, let me again commend this Committee for its continued commitment to high-performance computing. It has been my privilege to testify here many times. I appreciate the support of the Committee. And like my colleagues, I would be delighted to answer questions at the appropriate time.

[The prepared statement of Dr. Reed follows:]

# Written Testimony
**Professor Daniel A. Reed**
**Vice President for Research and Economic Development, The University of Iowa**

**Before the**
**Subcommittee on Energy**
**Committee on Science, Space and Technology**
**U.S. House of Representatives**

## Hearing on "America's Next Generation Supercomputer: The Exascale Challenge"

May 22, 2013

Chair Lummis, Ranking Member Swalwell, and Members of the Subcommittee, my name is Dan Reed, and I am the Vice President for Research and Economic Development at the University of Iowa. Thank you for the opportunity to share perspectives on the opportunities and challenges surrounding exascale computing and to respond to your questions regarding the American High-end Computing Leadership Act. I appreciate the time and attention that the Committee is spending on this topic, and I commend you for advancing the dialogue on computational science and high-performance computing.

My testimony begins by emphasizing the importance of high-performance computing as an enabler of scientific discovery and innovation across all disciplines, which distinguishes it from other scientific instruments. It summarizes key points in the shifting technology base of high-performance computing and the critical dependence of that base on continued investments in basic research. It then outlines some of the key recommendations from past reviews of the U.S. advanced computing investment strategy. It also emphasizes the interdependence of high-performance computing and the broader computing ecosystem, with implications for the future of U.S. competitiveness. Finally, it concludes by providing a set of recommendations and next steps for the Federal government and others to allow the U.S. to develop next-generation high-performance computing systems and to maintain its global leadership.

### High-Performance Computing: A Universal Amplifier

The English scientist Sir Humphrey Davy could well have been speaking about high-performance computing when he said, two centuries ago:

> *Nothing tends so much to the advancement of knowledge as the application of a new instrument. The native intellectual powers of men in different times are not so much the causes of the different success of their labors, as the peculiar nature of the means and artificial resources in their possession.*

In a phrase – success accrues to the talented and trained who have access to the most effective and powerful tools, whether computers, telescopes, particle accelerators, or genetic sequencers. Computing, and particularly high-performance computing, is unique among these and other scientific instruments, distinguished by its universality as an intellectual amplifier.

New telescopes advance astronomy and deepen our understanding of the universe's origins and cosmological future, but do not illuminate biological processes and the origins of life. New particle accelerators test the limits of the Standard Model and our understanding of fundamental physics, but do not yield new insights into the Earth's geological processes or and the exogeology of other planets in our solar system.

In contrast, **new, more powerful supercomputers and improved computational models yield new insights into all scientific disciplines, for they breathe life into the underlying mathematics of scientific models, allowing us to understand nuanced predictions and to shape experiments more efficiently. They also help capture and analyze the torrent of experimental data being produced by a new generation of scientific instruments and sensors, themselves made possible by advances in computing and microelectronics. Consequently, high-performance computing has emerged as the third pillar of the research portfolio, complementing theory and experiment across all disciplines.**

### High-Performance Computing: Past and Present

At any moment, high-performance computing (HPC) is most accurately defined by its impact – those computing systems with transformative power to enable breakthrough scientific discoveries, ensure defense preeminence and maintain international competitiveness. Thus, these HPC systems integrate the most advanced microprocessors and computational accelerators, the highest speed, lowest latency networks and the highest capacity storage systems. Their system software also embodies advanced techniques for resource management and systemic resilience, and the applications integrate complex numerical techniques that span a wide range of temporal and spatial scales. In short, they embody the most advanced computing technology currently available.

**In the past thirty years, we have seen repeated shifts in HPC hardware and software technologies, themselves consequences of long-term, U.S. government-funded basic research, with concomitant changes in computing systems deployments across industry, academia and our national laboratories.** In the 1980s, vector supercomputing dominated, as embodied in the eponymously named systems designed by the late Seymour Cray. The 1990s saw the rise of massively parallel processing (MPPs) and shared memory multiprocessors (SMPs), built by Thinking Machines, Silicon Graphics and others. In turn, clusters of commodity (Intel/AMD x86) and purpose-built processors (e.g., IBM's BlueGene), dominated the previous decade. Today, those clusters have been augmented with accelerators and GPUs. Each of these technology transitions brought dramatically higher performance – from gigaflops ($10^9$ arithmetic operations (flops) per second) through teraflops ($10^{12}$ flops/second) to petaflops ($10^{15}$ flops/second) – and new scientific and technical insights via higher fidelity computational models.

**Today, leading edge HPC systems at the Department of Energy and the National Science Foundation allow researchers to explore the frontiers of phenomena in scientific and engineering domains as diverse as high-energy physics, materials science, combustion dynamics, biophysics and computational chemistry, structural mechanics, and molecular biology. From understanding the subtleties of airflow in turbomachinery and underhood cooling through chemical molecular dynamics for consumer products to biomass feedstock for fuel cells, these and other systems also support advanced design and manufacturing, in partnership with U.S. industry.**

Across government agencies, these systems have also played an essential role in ensuring the safety and reliability of our nuclear stockpile and in protecting our national security in an uncertain and dangerous world. Large-scale data analytics also now enable extraction of insights from the unprecedented volumes of scientific and biomedical data being created by scientific instruments, as well as helping ensure information superiority for national security. High-speed networking and the global Internet also facilitate research collaboration and information sharing.

### High-Performance Computing: Looking to the Future

With every new generation of high-performance computing technology, the Department of Energy and its national laboratories have been at the forefront, collaborating with universities, other agencies and industry in the design, deployment and operation of the world's most powerful supercomputers. DOE's Advanced Scientific Computing Research Program (ASCR) has been a crucial element of this activity, funding basic and applied computing research and system development, and developing new computational science applications. ASCR has also worked closely with its DOE partner, the National Nuclear Security Agency (NNSA), on advanced technologies and system deployments.

**Today, HPC systems from DOE's Oak Ridge National Laboratory, Lawrence Livermore National Laboratory and Argonne National Laboratory occupy the first, second and fourth places on the list of the world's fastest computers, based on the Top500 list. From this, one might surmise that all is well. After all, in today's 21st century knowledge economy, the importance of U.S. leadership in high-performance computing and computational science would seem self-evident.**

**Yet today's U.S. leadership in both deployed HPC capability and in the underlying technologies that are needed to create the future generations of HPC systems is now under unprecedented challenge.** Other nations are investing strategically in HPC and computational science to advance their national and regional priorities. The U.S. research community has repeatedly issued warnings and alarms about this erosion of U.S. leadership in information technology and high-performance computing.

In 2004, I testified to this committee on this same topic while serving as the Director of the NSF-funded National Center for Supercomputing Applications (NCSA), which for twenty-five years has provided HPC services to the national science and engineering community, most recently via the NSF Blue Waters petascale HPC system. At the time of my 2004 testimony, I had recently chaired the 2003 community workshop on the *Roadmap for the Revitalization of High-end Computing*,[1] which had been convened in response to a request from the interagency High-end Computing Revitalization Taskforce (HECRTF). The workshop report's executive summary noted,

> *The common theme throughout these recommendations is the need for sustained investment in research, development, and system acquisition. This sustained approach also requires deep collaboration among academic researchers, government laboratories, industrial laboratories, and computer vendors. ... Rather,* ***multiple cycles of advanced research and development, followed by large-scale prototyping and product development, will be required to develop systems that can consistently***

---

[1] Community workshop on the *Roadmap for the Revitalization of High-end Computing*, 2003, organized by the Computing Research Association (CRA), available at http://archive.cra.org/Activities/workshops/nitrd/

> ***achieve a high fraction of their peak performance on critical applications, while also being easier to program and operate reliably.***

In 2005, as a member of the President's Information Technology Advisory Committee (PITAC), I chaired a review of U.S. computational science capabilities, which produced a report to the President entitled *Computational Science: Ensuring America's Competitiveness*.[2] The report's principal finding was

> *Computational science is now indispensable to the solution of complex problems in every sector, from traditional science and engineering domains to such key areas as national security, public health, and economic innovation. Advances in computing and connectivity make it possible to develop computational models and capture and analyze unprecedented amounts of experimental and observational data to address problems previously deemed intractable or beyond imagination. Yet, despite the great opportunities and needs, universities and the Federal government have not effectively recognized the strategic significance of computational science in either their organizational structures or their research and educational planning. These inadequacies compromise U.S. scientific leadership, economic competitiveness, and national security.*

Based on this finding, the PITAC's principal recommendation was the following:

> *To initiate the required transformation,* ***the Federal government, in partnership with academia and industry, must also create and execute a multi-decade roadmap directing coordinated advances in computational science and its applications in science and engineering disciplines.***

Today, we are poised on the threshold of a new era, one defined by exascale computing ($10^{18}$ flops/second) and trans-petascale data analysis. It brings the promise of new scientific discoveries and insights, but also difficult technical and engineering challenges. Exascale system design and construction will require solutions to some deep and fundamental problems in semiconductor processes, energy-efficient computing and data movement, primary and secondary memory design, packaging and cooling, resilience and reliability, resource management and programming. It will also require development of new numerical algorithms, data analysis techniques and scientific and engineering applications.

These solutions will not be simply incremental extensions of current technologies, nor will those solutions be derived from current industry research and development paths alone. Equally importantly, the fruits of such collaboration can have far deeper benefits than simply the construction of an exascale computing platform. They will be the innovative disruptions that will help position the U.S. information technology industry for the future. Our global competitors are well aware of this disruption opportunity -- there are now active and well-funded initiatives for hardware, software and applications in the European Union, Japan, China and India.

---

[2] *Computational Science: Ensuring America's Competitiveness*, President's Information Technology Advisory Committee (PITAC), June 2005

## Basic Computing Research: Partnerships and Innovative Disruption

In the United States, it has long been axiomatic that we are the world's leader in information technology and the application of that technology to business, science, engineering and government. In the 1960s, the birth of System/360 mainframe computing and its support for business processes made IBM a global leader in computing. In the 1970s and 80s, minicomputers such as the PDP-11 and VAX brought computing to research laboratories and smaller businesses, making Digital Equipment (DEC) a global brand. In the 1980s and 90s, personal computing made Intel and Microsoft large and successful companies. Today, Apple, Google and Amazon are icons of the smartphone and Internet age.

Each of these companies has been the beneficiary of Federal investments in long-term basic research, including investments in high-performance computing. The microprocessors and software in our PCs and smartphones embody architectural research, resource management and programming abstractions developed over four decades of research. Indeed, many of these ideas were first tested and validated in systems designed for high-performance computing.

**Today's Internet originated from DARPA network research investments in the 1970s and 80s, and from NSFnet, which the National Science Foundation (NSF) created to connect NSF's supercomputing centers and provide open access to high-end computing facilities. This environment spawned the Mosaic graphical web browser at the University of Illinois's NSF-funded National Center for Supercomputing Applications (NCSA) sparking the 1990s dot.com boom and the explosive growth of the Internet. That environment and further investment in search and indexing research led to the search engines, social networks and cloud services that define our daily interactions.**

**Make no mistake; global computing leadership is not a U.S. birthright**; it has been repeatedly earned and hard fought, based on a continued commitment to basic research investments by the Federal government, translation of those basic research insights into technological innovations, and strategic investment and business acumen to create and deliver new computing systems and products.

Andrew Grove, the former CEO of Intel, highlighted the importance of continual innovation in his famous computing aphorism, "only the paranoid survive." What he really said is far more subtle and important, "Success breeds complacency. Complacency breeds failure. Only the paranoid survive." Simply put, past success can lull one into complacency at precisely the time that changes and strategic investment are needed to ensure future success.

The computing industry is replete with telling examples of Grove's maxim, when technology breakthroughs spawned disruptive innovations. The rise of the personal computer made Microsoft and Intel large and successful, but it also required IBM to reinvent itself to continue to prosper. In that same period, DEC failed to make that transition successfully, despite its talented people and technology base. More recently, the birth of the Internet and the rise of smartphones and tablets have had similar disruptive effects on the computing ecosystem, with important consequences for our future.

**The Internet and web services revolution is now global and U.S. influence, though still substantial, is being diluted. Notwithstanding Apple's phenomenal success, the majority of smartphones and tablets are now designed, built and purchased outside the U.S., and the annual sales volume of smartphones and tablets already exceeds that of PCs and servers. In short, this exploding "post-PC" market is international in scope, with U.S. consumers an increasingly small minority of users.**

This ongoing shift in consumer preferences and markets is accompanied by another seismic technology shift. Smartphones and tablets are based on low power, energy-efficient microprocessors (a key component of proposed exascale computing designs) and systems-on-a-chip (SoCs) using the U.K.-created ARM architecture. Unlike Intel and AMD, which design and manufacture the x86 chips found in today's PCs and most leading edge servers and HPC systems, ARM does not manufacture its own chips. Instead, it licenses the design to others, who incorporate the architecture into custom SoCs that are manufactured by global semiconductor foundries such as TSMC.[3] Thus, the ARM hardware ecosystem is global in scope, and U.S. vendors, led by NVIDIA, Qualcomm and Texas Instruments, are but three of the international competitors in the ARM SoC market.

As a member of the National Academies Board on Global Science and Technology (BGST), in 2012, I chaired a study on this and other shifts and their implications for the United States and its future defense capabilities. The resulting report, entitled *The New Global Ecosystem in Advanced Computing: Implications for U.S. Competitiveness and National Security,*[4] made several salient points relevant to this discussion, of which the following is notable:

> *Over time, the increasing presence and establishment of foreign markets that are larger, are potentially more lucrative, and have better long-term growth potential than in the United States and other developed countries could also have significant implications.* ***Any shift in the global commercial center of gravity could lead to a shift in the global R&D center of gravity as international firms are required to locate in these markets if they are to remain competitive and to meet the requirements of government regulations in the target markets.***

These observations are equally apt for the future of HPC and exascale programs. U.S. competitiveness and continued HPC leadership are predicated on a vibrant U.S. computing industry that can continue to invest in the development of new technologies – advanced chips and architectures, novel networks and hardware systems, and new system software and applications – while leveraging continued investment in basic computing research by Federal research agencies, universities and national laboratories.

### Actionable Recommendations

The global computing ecosystem is in flux, and other nations are investing strategically in high-end computing. In the U.S., we also face difficult decisions about Federal investment priorities, given current economic realities. Thus, it has never been more important that we act strategically and thoughtfully as we consider the future of funding for basic computing research and for high-end computing. I believe the following are essential elements of a successful U.S. strategy.

1. **Advanced HPC system deployments are crucial, but the computing research and development journey is more important than any single system deployment by a pre-determined date.** The basic and applied research in algorithms, software, applications, semiconductor technologies, storage systems, energy management, integration and packing, resilience and scaling, among others, will produce unexpected discoveries and technology breakthroughs, as well as enable design

---

[3] Taiwan Semiconductor Manufacturing Company (TSMC), http://www.tsmc.com/english/default.htm

[4] *The New Global Ecosystem in Advanced Computing: Implications for U.S. Competitiveness and National Security*, National Academies Board on Global Science and Technology, 2012, available at http://www.nap.edu/catalog.php?record_id=13472

and deployment of effective exascale systems. Those discoveries and the people who made them are the lifeblood of computing innovation and future U.S. competitiveness. They are the true enablers of sustainable exascale computing.

2. **High-end data analytics (big data) and high-end computing (exascale) are both essential elements of an integrated computing research and development agenda; neither can be sacrificed or minimized to advance the other.** From web search, social networks and business processes, through government efficiency and service optimization to large-scale scientific instrumentation and sensors, big data has been and will be transformational. Cloud computing infrastructure and services and high-performance computing systems and services have much in common, and insights from each can benefit the other. Global leadership in both is essential to the our future.

3. **Research and development of next-generation algorithms, software and applications is as crucial as investments in semiconductor devices and hardware, and we have historically underinvested in these areas relative to hardware.** Despite this underinvestment, experience has shown that over the past thirty years performance increases in high-performance computing systems has been due in equal parts to hardware and software advances. The massive and unprecedented scale of current and future high-performance computing systems is bringing new challenges in programmability and systemic resilience, resource scheduling and numerical stability. We must invest in a balanced way.

4. **Much deeper and sustained interagency collaboration is needed. The Department of Energy, particularly the Advanced Scientific Computing Research Program (ASCR), has led the development of an exascale computing research and development agenda, but it cannot succeed alone.** In the past, the National Science Foundation, the Department of Defense, the National Institute of Standards and Technology, and the National Institutes of Health have been active and engaged partners in the high-performance computing research and development agenda. Today, that is much less true.

   **The historical strength of the U.S. research strategy in high-performance computing has been the complementarity and diversity of its participating research agencies. We must renew and reenergize that partnership, given the unique role that each agency plays:**
   - NSF – basic computer science research in the enabling technologies; data management and sustainable cyberinfrastructure for national science and engineering academic community
   - DoD – advanced technology research and prototyping; mission-oriented deployments
   - NIST – standards and cybersecurity
   - NIH – computational modeling, big data analytics and biomedical applications for higher quality, lower cost health care
   - DOE – computational science, systems research and prototyping; large-scale system deployments, building on the research and operations staff of the Office of Science and NNSA laboratories

5. **We must change the model for research, development, acquisition and deployment of high-end computing systems if the U.S. is to sustain the leadership needed for future scientific discovery and national security.** As the HECRTF report noted, we must support and sustain multiple cycles of advanced research and development, followed by large-scale prototyping and product development. In a budget-constrained world, we must work more efficiently and collaboratively, which will require new and deeper interagency prioritization and budget allocations, along with

long-term industry partnerships. To ensure that, **there should be verifiable metrics of interagency collaboration, community engagement and technical progress that are tied to agency funding.**

6. Finally, the global information technology ecosystem is in flux, with the transition to a new generation of low power, mobile devices and cloud services. **We must recognize and embrace the need for "dual use" technology research and development that enables high-performance computing systems and scientific discovery while also ensuring the competitiveness of U.S. industry, both in information technology and in the use of computing to advance U.S. businesses. Our long-term national security depends on this.**

I believe we face both great opportunities and great challenges in high-performance computing. Scientific discovery via computational science truly is the "endless frontier" of which Vannevar Bush spoke so eloquently in 1945. The challenges are for us to sustain the research, development and deployment of the high-performance computing infrastructure needed to enable those discoveries. To do so, we must adapt our model of collaboration, retaining the strength of its diversity while focusing our resources efficiently.

Finally, let me again commend this committee and its continued leadership and commitment to high-performance computing, including the American High-end Computing Revitalization Act.

# Daniel A. Reed
## Vice President for Research and Economic Development
## The University of Iowa

Daniel A. Reed is Vice President for Research and Economic Development, as well as University Chair in Computational Science and Bioinformatics and Professor of Computer Science, Electrical and Computer Engineering and Medicine, at the University of Iowa. In this role, he is responsible for the university's research processes and compliance, campus-wide research and policy centers, technology evaluation and licensing, and economic development.

Prior to joining the University of Iowa, from 2007 to 2012, he was a senior leader at Microsoft, first serving as Microsoft's scalable and multicore computing strategist and later as Corporate Vice President for Extreme Computing and Technology Policy. In these roles, he helped shape Microsoft's long-term vision for technology innovations in parallel and cloud computing and the company's associated policy engagement with governments and institutions around the world.

Previously, he was the Chancellor's Eminent Professor at UNC Chapel Hill, as well as the Director of the Renaissance Computing Institute (RENCI) and the Chancellor's Senior Advisor for Strategy and Innovation for UNC Chapel Hill. Prior to that, he was Gutgsell Professor and Head of the Department of Computer Science at the University of Illinois at Urbana-Champaign (UIUC) and Director of the NSF National Center for Supercomputing Applications (NCSA). He was also one of the principal investigators and chief architect for the NSF TeraGrid. He received his PhD in computer science in 1983 from Purdue University.

In addition to his technical activities, Reed has been deeply involved in policy initiatives related to science, technology and innovation, both in the United States and internationally. He has served as a member of the U.S. Federal Communications Commission's Technical Advisory Committee, as a member of the U.S. President's Council of Advisors on Science and Technology (PCAST) and chair of the computational science subcommittee of the President's Information Technology Advisory Committee (PITAC). He currently serves as chair of the capability computing review committee for Los Alamos National Laboratory, as a consultant to Department of Energy, and as a member of the National Academies Board on Global Science and Technology.

Mr. WEBER. Thank you, Dr. Reed. And thank you all for your testimony. Man, lots of questions come to mind. And, you know, I guess I am an old-timer. I grew up back in the '60s and we didn't have computers, actually, we did. There was a flat table, you put a quarter in, and you chased a little Pac-Man around. Those were our computers.

So I have a question here, and I think you kind of alluded to it, Dr. Reed, but I will ask this maybe starting with Dr. Giles. Is it Giles?

Dr. GILES. Giles actually.

Mr. WEBER. Giles, there you go. Thank you.

In December 2011 Congress directed DOE to provide a strategic roadmap relating to the development of an exascale computing system. However, it is my understanding that after 15 months of the mandated completion date, the report is not yet finalized. Are you aware of this report?

Dr. GILES. I am aware of it but my position is as an external representative of the community relative to ASCR so I am actually not an insider and I have not seen the report.

Mr. WEBER. Okay.

Dr. GILES. My understanding is exactly what you said.

Mr. WEBER. Okay.

Dr. GILES. But I don't have anything unfortunately to add to that.

Mr. WEBER. Nothing that you want to admit here publicly?

Dr. GILES. No, actually nothing that I know. Anything I know, I will tell you.

Mr. WEBER. Of course, our goal is to get it. Dr. Stevens, how about you?

Dr. STEVENS. Well, I am aware of the report. I think it is a fine plan. I think that the internal process of getting that report out is what has blocked it, and I hope it reaches you quickly.

Mr. WEBER. Okay. Any help you can give us in that endeavor?

Dr. STEVENS. I don't have any specific recommendation except to just reemphasize that this is a critical plan that must be delivered and must be understood and articulated and executed.

Mr. WEBER. Okay. Thank you. And Ms. Crawford, I don't mean to put you on the hot seat, but you are on the hot seat.

Ms. CRAWFORD. I have nothing to add to what Dr. Stevens said. We are—we work at the laboratories and we are not part of the formal process between the DOE and the OMB to get that report out. I do support what is written in the report. The labs had a lot of input. I have not seen the final report.

Mr. WEBER. And Dr. Reed, since you came to us with four points followed by recommendations, and I love that by the way. One of the things you said in your recommendation was the Department of Energy needs partners and long-term industry engagement. How do we expedite this? How do we make this happen?

Dr. REED. Well, I think there are several points relevant. One is to recognize that, as I said, it is a false dichotomy to pit investment and some of these big data issues against high-performance computing, and I think frankly that is the root of some of the issues that we have to resolve in terms of moving forward.

In terms of the agencies, I believe, as I pointed out in my written testimony, that they each fulfill and historically have fulfilled important and complementary roles. The Department of Energy has been crucial in terms of advanced prototyping and deploying of the largest scale systems. The other agencies, though, provide support for enabling technology research. The National Science Foundation is one of the key enablers of that long-term research.

What is important is that all those players be at the table and be engaged in supporting this integrated agenda. I think from the industry's perspective to sort of answer your specific question, that is where industry—and I speak now again from my industry experience—it is important that the government be a committed and not fickle partner because the cost of money and the time planning for companies to execute is really crucial. And as I was saying, that combination is key to the future of the U.S. industry not just for high-performance computing but for how much information technology means to the U.S. economy.

Mr. WEBER. All right. Thank you. And Ms. Crawford, let me come back to you. I think you said that this exascale computing either can't or won't be achieved through refinement. What did you mean by that?

Ms. CRAWFORD. What I mean by that is the current system architectures today can't simply be scaled up to produce a usable and cost-effective system. In principle, one could scale it up and you would have a system that would fill the room and would take 100 megawatts of power, so that is not a cost-effective system. So the technologies have to change and we have to change in memory, in processors, in storage and networking and the programming models. And so that is what I mean by we can't simply scale-up the programs of today.

Mr. WEBER. Let me send that over to Dr. Stevens. And you mentioned about more or less power, I guess explain, you said less power.

Dr. STEVENS. Right. We need to develop processors and memory and network components that consume considerably less power than current systems in order to scale-up. Right now, if we took a current kind of 10 petaflop system and scaled it up to an exascale, it would consume nearly a gigawatt of power, which is not feasible from a physical infrastructure standpoint or a——

Mr. WEBER. Right.

Dr. STEVENS. —cost standpoint. So we need much more power-efficient devices. We also need better programming models because we are going to have to have a lot more parallelism inside these machines, 1 million—or 1,000 times more parallelism than we have now and we need ways of accessing that parallelism easily for programmers. So we have a lot of work to do. We know what to do. The DOE's plan includes all of these activities so it is—I think the United States has a good position to do this; we just need the resources and the long-term commitment.

Mr. WEBER. All right. Thank you. And I just want to make an observation before I yield to the Ranking Member and that is that I am glad to hear you say that national security is involved in this and tied up in this. That is very crucially important. And I think it will carry a lot of weight with Congress. Hopefully, it will. So I thank you for your testimony. And with that, I yield to Mr. Swalwell.

Mr. SWALWELL. Thank you, Mr. Chair. And my questions will principally be for Ms. Crawford.

First, does research in high-performance computing require the United States Government to make investments? And what I mean is why can't we simply rely on the private sector to innovate and invent the next supercomputing architecture and software and then the government can just buy off-the-shelf technology?

Ms. CRAWFORD. The short answer is, yes, the U.S. Government does need to invest in order to shape the exascale architectures for our mission needs. I can use an old example. When we started the Accelerated Strategic Computing Initiative in the mid-1990s for the Stockpile Stewardship Program, industry and the consumer base was driving computing in a direction that would not meet our needs. And without our investment and our sustained investment and focused on cooperation and developing those processors that would meet our needs, we wouldn't have had the computers and the computing capability that we have today. And so today, it is essential that we work together with the Office of Science laboratories and the NNSA laboratories to meet this mission needs.

A shorter answer perhaps is that we are going to follow industry technologies. We can't afford our own, you know, brand-new fab or our brand-new machines. What we want to do is pay on the margins to make those machines viable for our particular applications, which is mimicking the, you know, physical phenomena around us.

Mr. SWALWELL. And when we look at our global competitors—Japan, China, India, Brazil, Russia—are they allowing or relying solely upon the private market or are they also having government investment at the table as well?

Ms. CRAWFORD. There is strong government investment in Japan, China, Russia, the European Union. It is about $1.1 billion of investment in Japan. I would have to do the translation but the Ministry of Science and Technology five-year plan within China is investing and again not just in the hardware technologies but they are investing in the low-level software and the applications and making sure that they have the ecosystem in order to be able to deploy these systems effectively to make a difference to their underlying national security and economic competitiveness. So——

Mr. SWALWELL. So it sounds like——

Ms. CRAWFORD. —they are going to be large investments.

Mr. SWALWELL. It sounds like for the United States to keep its edge in high-performance computing, we will need to continue to have the Federal Government make investments in these programs, is that right?

Ms. CRAWFORD. Absolutely.

Mr. SWALWELL. You talked a little bit about the joint partnership that must take place between NNSA and the Office of Science. Why is exascale capability so critical to DOE's National Nuclear Security Administration?

Ms. CRAWFORD. So then I will take a more focused view on just what is going on within the NNSA laboratories. It is our duty to assess the state of the stockpile on an annual basis, and the stockpile is being decreased in the numbers of weapons and the types of weapons. That makes each single weapon remaining in the stockpile critically important to understand what is going on——

Mr. SWALWELL. Going toward a more leaner and meaner model, right?

Ms. CRAWFORD. Leaner and meaner, and so those systems, as they age, they are being modernized as parts begin to fail, and so there are a number of things that we need to understand, you know, physically. You know, nuclear weapons are very complex. Think about parking your car in the garage and not turning it on but then wanting to be able to use it when you have to. You know, there are special materials that are changing over time and all kinds of things that go on just sitting there.

We need high fidelity 3–D simulations to understand, you know, the initial conditions, the engineering features, safety features, the security features, and today, we cannot simulate at that high fidelity. So we have a number of—what we do is look at the kinds of calculations we are going to do and the kinds of computing that is required to do those calculations and so—for stockpile assessment, for the life extension programs for materials aging, for safety and surety, we have a range of exascale needs for the kinds of calculations that will have to go on in high fidelity, high-resolution 3–D, and they range from half-an-exascale to 1,000 exascales over the period of the next 10 years.

Mr. SWALWELL. And Ms. Crawford——

Ms. CRAWFORD. Starting in about 2018.

Mr. SWALWELL. Can you tell me more about Livermore's work to address industrial and medical research needs, for example, your groundbreaking simulation of the human heart and your recent work with the California Energy Commission to improve energy management throughout the State and how exascale and HPC have affected our ability to do this?

Ms. CRAWFORD. I would be glad to. Having developed these capabilities for our mission drivers, then they are applicable, as Dr. Reed has said, to many other activities. Last year, we worked with IBM to develop a code called cardioid and it does—it models the electrical signals of the heart and it has the potential to be used to test drugs or medical devices, the code ran in nearly real-time across our 20 petaflop machine at Livermore beating an astonishing 60 beats per minute, so this is almost, you know, 12 percent of real-time. This calculation ran at 59 percent of peak of this machine, and that is—you know, it is very incredible and amazing thing to take a new code and put it on a new machine and run at this scale. It runs in a time to solution over 1,200 times faster than the previous state-of-the-art and this work shows promise for what advanced computing can do for understanding the human body. But it also demonstrated the extreme level of specificity and technical acuity required to achieve this result. And of course, these insights that we gain there will then be applied back to the stockpile.

Mr. SWALWELL. Great. Well, thank you so much, Ms. Crawford. And thank you again to our other witnesses. Thank you, Mr. Chair.

Mr. HULTGREN. [Presiding] Thank you. And I will recognize myself for five minutes for a few questions.

Part of our challenge as a Subcommittee is certainly to understand the right thing to do but also to present it to the larger Committee and even beyond that to Members of Congress, so a couple of questions. Just if you have been messaging or how to present how important this is and why this is so important so I would address this first question to Dr. Stevens and Ms. Crawford. Wonder if you could just discuss the expected breadth of applications for the exascale computing. Is this something that could be used for a wide range of important disciplines from material to chemistry to medicine to nuclear science similar to the current supercomputers or is the expected range of disciplines more narrow such as climate science modeling or for weapons development?

Dr. STEVENS. So the range of applications for exascale are no less broad than the current machines. In fact, there are many problems that haven't been tried in the past, particularly in biomedical science where we were just afraid to try them. We didn't have enough compute power. This idea of trying to build, say, detailed models in the human body, not just the heart but now include the lungs, include the nervous system, include the gastrointestinal system and build a virtual human, that is a problem that will require 1,000 times current machines. It is not really feasible so people haven't tried it. So my sense is that we will find more and more applications as we build more capable systems.

We are also going to increase the ability for these systems to deal with data, and so a new class of applications that is emerging in both national security and in engineering research is this idea of doing modeling simulation with uncertainty quantification, this idea that not only will you get a result, you will get some confidence measure on that result. And that is something that requires hundreds of times more compute power than the current capability which means you can only do one simulation.

Ms. CRAWFORD. I second everything that Dr. Stevens said. And it is limitless. Computing is so foundational. Anything that—any physical process that you can represent mathematically, which are most of them, you can then model in the computer with great fidelity. And the greater the fidelity we have, the better we can understand the world around us. And so I can just go on and on and on but, you know, we work at our laboratory in a number of areas with industry, with other national laboratories, with academia to make sure that we are applying these to the breadth of possibilities.

Mr. HULTGREN. Well, Ms. Crawford, if I can get into just a little bit more specific and really following up on the Ranking Member's discussion and also on the Vice Chairman's of what does speak to Members of Congress and inspire us to make a commitment, especially a financial commitment at a time like this, and certainly, one of those is national security.

So I wondered if you could just talk briefly. Is exascale computing considered critical to advancing national security, and if so, has the National Nuclear Security Administration gone on record to say that? If so, how is the NNSA prepared to financially contribute to this effort and what would be an appropriate percentage contribution to an exascale computer from NNSA, would you say?

Ms. CRAWFORD. There is a lot of questions there so——

Mr. HULTGREN. Yes.

Ms. CRAWFORD. —let's see if I can remember them.

Mr. HULTGREN. The first thing is have they gone on record of saying that this is a key component? And then basically then what kind of commitment should we expect from them?

Ms. CRAWFORD. Computing is the integrating element of maintaining the safety, security, and reliability of the nuclear weapons stockpile without returning to underground tests. So by integrating element, what I mean is we have the old test data, we have aboveground small experiments that we are doing, and we have a lot of theory and we have our new models. And we are bringing this all together in the computer. So this is an integrating element and this is the only way that we know to understand what is going on in the nuclear weapon. And so for that reason, we believe that it is extremely important.

The NNSA is making an investment in the Advanced Scientific Computing Program. To maintain leadership, you need to have a base program. You need to have, you know, sort of a near-term program and you need to have a far-reaching program. Currently, the Office of Science and the NNSA both have a very strong base program. We have heard about the wonderful facilities at the laboratories, and of course it is not just the computer hardware itself but it is the applications that help us understand the world around us.

We are investing with the Office of Science in some near-term research with industrial partners to look at some of those long lead time technology changes that need to be made. We need to make additional investments that are not in our current budgets in the programming environments for the exascale computing and in the math libraries so that we can actually use this billion-way parallelism.

Mr. HULTGREN. Okay. I see my time is expired. At this point, I hope that we will have an opportunity to have a second round of questioning as well.

Mr. SWALWELL. I don't have any objection.

Mr. HULTGREN. We can talk about that. Well, let's go ahead and we will recognize Mr. Veasey from Texas. Okay. Then Mr. Lipinski from Illinois is recognized for five minutes.

Mr. LIPINSKI. Thank you, Mr. Chairman.

I wanted to ask everyone on the panel a question about international partnerships. You know, obviously this cuts both ways. You can reduce the cost of reaching exascale capabilities with international partnerships but then there is the issue of, you know, damaging our Nation's economic competitiveness, potentially our national security, because we are not doing this on our own. Now, where do you come down on this? Is it worthwhile and how far should we go in international partnerships and at what point is it still an advantage? At what point does it become a disadvantage for us economically, giving up our lead on high-end computing? So whoever wants to start with that one. Dr. Stevens?

Dr. STEVENS. I will start. So I think the primary opportunity in international collaboration is in software, and in particular, the components of software that are open source that right now most of the software that runs on these machines other than the applications is built on—based on open source technologies developed largely in the United States. That is a significant lift to move all of that software to next-generation platforms, and international collaboration can help there provided that the software is—stays in the open.

I think where we don't want to go at least in the near term is in deep hardware partnerships internationally. I think that is a place where we want to maintain our competitive edge. We have significant advantages with the U.S. vendor community and we want to maintain that as long as we can.

Mr. LIPINSKI. Thank you. Anyone else? Ms. Crawford?

Ms. CRAWFORD. Yes. I would like to add that it is very important that the United States maintain the key intellectual property for the next supercomputer levels. If we control that, we have the high ground for the standards space base that will make all the decisions in the coming decade, and I would not want to cede that to another country. I cannot trust the U.S. nuclear weapons technology to a system built in China, say. That is untenable. I would like to not consider that those low-power technologies are developed ahead of time in other countries that we will use embedded in our intelligence systems. To me, it is very important that the United States take a very strong leadership position in this technology arena.

Mr. LIPINSKI. Thank you. Dr. Reed?

Dr. REED. Yes, if I might add to that. It is part of the reason in my testimony I spoke very specifically to the importance of U.S. industry engagement. And as we move into this increasingly mobile device, low-power world, which is one of the key enabling technologies for future exascale systems, it is really crucial that the U.S. vendors maintain the competitive edge and strike a balance, as we do in terms of investment, between the global market and maintaining the unique capabilities for U.S. national security.

Now, that is part of the role of the Federal Government in terms of, as Ms. Crawford said in her testimony, about shaping the direction of industry to ensure that we have the technology capability that we need.

And I would echo that there are other uses as well. As we have talked about the rise of data analytics and its importance for national security and signal intelligence and other domains, that is another area where we must think carefully about many of the enabling technologies of which hardware is one, but the algorithms and other pieces need careful scrutiny also.

Mr. LIPINSKI. Thank you.

Dr. GILES. Yes, I would agree with what has been said. Just two points: I think it would be truly shameful for us to give up the elements of leadership that we have. And one of the things we pointed out and we asked in our exascale report was the criticality of time and of seizing the opportunity that in some way is presented uniquely to us to advance this field. But many, many countries will want to do that and we have a little bit of a time advantage because of our starting place.

The other point, which is—it goes sort of in the direction of the open source software is the observation that a lot of the open science that is done in the world is done with international collaboration and with international connections. And we would, I think, like to still be in the position of having a lot of influence on the under-layer of that on which we will all build. But there certainly is international collaboration in science and I wouldn't want to minimize that—the importance of that for the open science community.

Mr. LIPINSKI. Thank you very much. And I want to ask a question if the Chairman would give me just a few extra seconds here. I just want to also echo what I know some of my colleagues have stated. I know exascale computing is important but we have to make sure that we don't pursue that at the expense of other important R&D activities that ASCR is doing. And I yield back.

Mr. HULTGREN. The gentleman yields back. We will go through a second round of questioning if anyone would have other questions, so I will begin by recognizing myself for five minutes.

And I would address this first to Dr. Stevens but also ask if any of you would have other thoughts on this and really following up on Mr. Lipinski's questions of timing and competitiveness. And I wondered, Dr. Stevens, if you would have some thought of what level of investment is needed for the United States to maintain global leadership in scientific and technical computing for the next decade? And then something specific of if we maintain current investment, at what point would China surpass us in computing capabilities? And then also just looking at dates, what type of approach and how much investment would be necessary for us to lead to a deployable system by 2020?

Dr. STEVENS. Okay. So on the first one in terms of the resources required to do this, in the plans developed by the laboratories, we estimated that in addition to the current funding levels that we have, we would need an increment over time of approximately $400 million a year. That would be split between the two partners, the Office of Science and the NNSA. At that funding level, we think it is feasible—not guaranteed but feasible—to deploy a system by 2020. Of course, we made those estimates a few years ago when we had more runway than we have now. And that investment would go to both hardware and software and some applications of them—more applications would be needed by that time.

At the current funding level, not including the bill——

Mr. HULTGREN. Right.

Dr. STEVENS. —that is in front of us, it is estimated that we would not reach an exascale capability until middle of the next decade. We don't have accurate estimates of precisely what China will do but my guess is they will probably exceed us by the end of the decade if we were in that scenario. I don't remember your——

Mr. HULTGREN. I think that covered it. So really it is, you know, without the investment, it is going to be probably 2025 before we would reach that level?

Dr. STEVENS. Absolutely.

Mr. HULTGREN. Do you think with the investment, is it a possible——

Dr. STEVENS. We have——

Mr. HULTGREN. —expectation to reach exascale levels by 2020?

Dr. STEVENS. I think it is possible. I think we would have to get moving faster than we are now and of course the industry is ready to do this. Labs are ready to do it; academia is ready to do it. We just need the resources and the commitment and also to do it in a way that doesn't cannibalize the current program. We need the base—we have to build on the base both in the Office of Science and in NNSA, and so this is really, really looking at incremental resources unfortunately to do it.

Mr. HULTGREN. Okay. Thank you. Do any of the others have any thoughts or disagreement?

Ms. CRAWFORD. I would just add that understanding what the sustained commitment is, whatever that dollar level turns out to be, is critical because then we can plan into the future. And not knowing whether, you know, the base budget is cut and the exascale R&D budget is cut and we have got a commitment to do this and then we are—now, we must do that because we have a contract and yet that prevents us from doing something else. So not knowing is really difficult to plan ahead and manage it effectively. So understanding that and sustaining that is very important.

Mr. HULTGREN. I absolutely agree and it is one of the things I am passionate about. I know other Members of our Subcommittee and Committee are as well of bringing some certainty specifically to research and to science. When we are looking to advance these programs it is so important that we are not budgeting month-to-month, which this place, Washington D.C., has kind of fallen into the habit of doing, but it has incredible detrimental impact, I know, on the great work that you all are doing.

So I for one and I know my colleagues on both sides of the aisle would love to see some of that change. We are going to be fighting for that.

Let me switch gears just a little bit and address this to Dr. Giles if I could and also to Dr. Stevens. But with respect to achieving an extraordinary number of computations per second, exascale appears to be a somewhat arbitrary goal. With current budgetary constraints, could DOE consider slower systems that would still be by far the fastest in the world or how do you see that fitting into this challenge of kind of keeping up with the rest of the world if DOE were to say, well, you know, we want to do some advancement but we are not going to go for that larger goal. We will just kind of settle for a lesser goal. How do you see that impacting the work that you are doing and the work that other nations are doing?

Dr. GILES. Okay. Well, I think the key research to lower power consumption, to identify the pathway that takes us to exascale is one that is defined by that goal but which is a sort of—has a certain integrity of its own. Okay. If you do that—if one does that and makes that commitment to do their research and to do that beginning development, then how far you take it is part of the deployment question of how big a machine you build with the technology that you have done the research for. It—so—at least that is my take on it. I am not the technologist that Rick is and you may have a comment on that.

Dr. STEVENS. Well, what I can say is that the laboratories are excellent stewards of the Nation's money——

Dr. GILES. Yes.

Dr. STEVENS. —and we will buy the most capable systems that we can afford to buy when we have to replace and when we can replace the current systems. So I think that the question of, you know, can we settle is really a question of do we want to settle for not being able to do all the science or the most impactful engineering or address the most important national security challenges? We will do the best we can with what is provided to us. There is no question. I think lowering our sights though is not in our DNA.

Mr. HULTGREN. Right.

Dr. STEVENS. Right.

Mr. HULTGREN. No, that is helpful. Thank you. My time is expired. I will recognize the Ranking Member, Mr. Swalwell.

Mr. SWALWELL. Thank you, Mr. Chair. And I appreciate your comments about providing more certainty to our national laboratories. And we know that it is not just the laboratories who need the certainty but also private industry or any contractors who depend on work from the laboratories.

One of the first lessons I learned when I was a planning commissioner years ago on a local sign ordinance issue from a local small business owner was vote for me, vote against me, but just give me certainty and, you know, do not have, you know, month-to-month sign regulations that give us no certainty at all, which now I have learned here, as the Chair said, month-to-month budgets also don't serve our laboratories well or private industry well. And so I join you in hoping that we can find ways to provide more certainty.

I was hoping to just go witness by witness briefly and if you could just tell me for my own edification, and I am sure many others are curious, what are the private/public partnerships that you have at your laboratories through the exascale program?

Dr. GILES. Well, let's see. I don't run a laboratory.

Mr. SWALWELL. Sure.

Dr. GILES. But I would note things like you do run a lab that does the INCITE program in ASCR that invites researchers from outside DOE and from industry and with the particular emphasis on some industries to use the most advanced facilities that we have, so I think that would be one that I would identify coming out of ASCR.

Mr. SWALWELL. Great. And Dr. Stevens?

Dr. STEVENS. Well, just a few that we have done in the recent past. We have got a collaboration with Pratt & Whitney developing more efficient turbine engines, with Procter & Gamble on a variety of improving consumer products, with Cummins in improving diesel engines, and Caterpillar improving their ability to model whole vehicles and including the transmission systems and so forth, with the Mayo Clinic in applying computations and larger-scale problems in metagenomics, and so on. There is a long list. Some of these are collaborations with end-user companies and some are collaborations with companies like IBM or with Intel and with Cray in developing next-generation technologies, and we also work with small businesses.

So the laboratories have collaborations on both the end-user component of this technology and the company is developing the technology itself.

Mr. SWALWELL. And when I hear some of those companies, IBM, Intel, Cummins, Caterpillar, I think of billions and billions and billions of dollars of exports. Those are some of the largest exporters in the United States, and if we are going to truly achieve our goal of doubling our exports over the next five years, making sure that those companies can continue to play a part in reducing that trade deficit—we have about $40 billion every month—is crucial and it sounds like the laboratories are helping them to do that so they can sell their goods and services to the marketplace outside the United States.

Dr. STEVENS. Absolutely. And we are also working with companies like Dow and DuPont and Johnson Controls. And it is a long list, right? And I think we exactly get this idea of helping American industry be more competitive.

Mr. SWALWELL. Great. And Ms. Crawford?

Ms. CRAWFORD. So rather than going through the long list, let me talk about the barriers for industrial adaptation of advanced computing. There have been a number of studies and there are three main barriers. One is the cost of establishing a supercomputing facility, the computer itself, the computing room, et cetera. The second one is the expertise, you know, having the skilled workforce that understands how to use these computers in a meaningful way for their products. And then the third is the software itself that helps them understand their products and how to improve those products. So the kind of partnerships that Dr. Stevens is talking about and that we have in our laboratory are helping to demonstrate to industry how to overcome those barriers so that they can in fact utilize this. And once they have firsthand demonstration and know the value, then they will start making the investments themselves at a higher level to drive their own productivity and competitiveness.

Mr. SWALWELL. Great. And Dr. Reed, I mean also just like Dr. Giles I know you do not run a laboratory but any public/private partnerships you are familiar with that are working right now and also helping the innovation economy?

Dr. REED. Certainly. And I have been in similar roles in the past. As I mentioned, I used to run an NSF supercomputer center and we did very similar things in Illinois when I was there. Advanced manufacturing was certainly a target, logistics and supply chain optimization. But in Iowa now, there are many issues around advanced biological modeling and how we think about the future of healthcare in terms of everything from modeling the characteristics of lungs and what the implications are for drug delivery, how we might work with companies about those issues.

I would echo what Ms. Crawford said, though. What is really crucial in those engagements and use of high-performance computing is simplicity of use because the domain experts are interested in advancing either the technology or the science or its applications and less interested in understanding what those of us in the technology business might view as the really cool stuff.

Mr. SWALWELL. Right.

Dr. REED. It is a means to an end and so those software user interface issues are really important.

When I was at Microsoft, I spent a great deal of time working with the community in science on exactly those issues. How do we bring the power of advanced computing into small companies and into individual's hands where, from their perspective, the ease-of-use that they find familiar in their mobile device or their PC extends seamlessly and apparently magically to exploit those advanced capabilities?

Mr. SWALWELL. Great. Well, thank you. Thank you, Mr. Chair. This has been a great hearing. You know, I didn't pay enough attention to this stuff when I was in high school. I am learning a heck of a lot now in Congress and I could sit here for another few hours but I know our witnesses and our panel have other places to be. But thank you again.

Mr. HULTGREN. Thank you. Thank you. And I do want to thank each one of you for being here today on a very busy day on Capitol Hill. And with that, I just want to thank you for your valuable testimony and I want to thank the Members for the questions that they have had. The Members of the Committee may have additional questions, especially with competing hearings that were going on at the same time, so we will ask if you would be willing to respond in writing to questions that we would submit.

And with that thought, we will keep the record open for two weeks for additional comments and written questions from Members and request for your response to those.

With that, I again want to thank you so much for your time and for your wisdom and information today. With that, the witnesses are excused and this hearing is adjourned. Thank you.

[Whereupon, at 11:20 a.m., the Subcommittee was adjourned.]

# Appendix I

## ANSWERS TO POST-HEARING QUESTIONS

## ANSWERS TO POST-HEARING QUESTIONS

*Responses by Dr. Roscoe Giles*

**Boston University** College of Engineering
Department of Electrical and Computer Engineering

8 Saint Mary's Street
Boston, Massachusetts 02215
T 617-353-2811 F 617-353-7337

BOSTON
UNIVERSITY

## EXASCALE COMPUTING CHALLENGES AND OPPORTUNITIES
## RESPONSE TO QUESTIONS FROM THE COMMITTEE

Prof. Roscoe C. Giles
Chair, Advanced Scientific Computing Advisory Committee
Professor, Boston University College of Engineering

## Hearing Questions for the Record
## The Honorable Cynthia Lummis

1. **Please summarize some of the varying views and differences within the computing communities regarding challenges and opportunities associated with the pursuit of an exascale computing system.**

Computing communities I will discuss include: computer scientists and engineers who research, design and develop new systems; application scientists whose principle interest is solving their particular scientific engineering problems; and computational scientists and applied mathematicians whose primary interest is in effective use of computers to solve problems. Each community has its own perspective on the exascale initiative.

Computer researchers and designers see the many of the key elements of exascale technologies as shared research and development challenges with mainstream computing ranging from mobile systems to data centers. Emerging changes to computing include effective use of large scale parallelism on multicore chips, power aware computing, and new aspects of error management, all of which require new programming models, new approaches to correctness and robustness, and better software tools. Progress toward exascale exacerbates the need overall progress in computing. Exascale has additional challenges that are different from the mainstream – emphasis on floating point performance (which is also important for some computer graphics) and tightly coupled systems needing fast and frequent data exchange and synchronization for example – for which progress toward exascale complements and extends the mainstream direction.

Application specialists whose science and engineering problems are central to next generation computing systems have expressed their positive vision for the capabilities of exascale since the inception of the program – see for example the materials from the original nine exascale computing applications workshops (summarized in the ASCAC Report "The Opportunities and Challenges of Exascale Computing", 2010). The common vision is the transformation of modeling and simulation from primarily an analytic mode to a predictive mode – through comprehensive code validation with experiments – giving us capabilities for designing as well as understanding increasingly complex dynamic systems that better represent the real world. This radically impacts DOE science and engineering as well as industry. Concerns for the applications community include: (1) that there is adequate systems and software support for bringing new architectures into production mode smoothly and without losses of productivity during transition; and (2) that the new capabilities are brought into production soon enough to have a timely impact on applications.

Computational scientists and applied mathematics researchers are very aware of some of the less visible challenges that are required for success of next generation computing facilities at the exascale. For example, we need new algorithms for solving large systems of connected equations on the new architectures that minimize data movement and memory requirements, and we need a new understanding of uncertainty and fault tolerance for very large-scale systems with different processor/memory/communication characteristics than previous computers. We also need to investigate alternative programming models that facilitate expressing parallelism and controlling data layout in discretizations of large systems of connected equations. We also need new algorithms for in situ analytics/visualization and uncertainty quantification given limited I/O bandwidth. These algorithms must share data structures with the solver and an optimized execution model for end-to-end exascale workflows needs to be developed. Some applied mathematics and computational science researchers are excited by elements of these frontiers. Others are challenged by the need to develop systems that work for applications on a relatively short time horizon.

The DOE co-design centers, bringing together these communities address some of the development issues the communities face and connect them in working together for solutions.

2. **What is the role and importance of the Department of Energy's applied mathematics research program in supporting scientific discovery utilizing DOE computers? To what extent are new applied mathematics discoveries required to support the evolving scientific requirements, and to what extent can DOE rely on existing applied mathematics techniques to meet its mission requirements?**

It is important to note that the applied mathematics research program has had and continues to have a substantial impact on scientific discovery beyond its role in enabling the efficient use of any particular computing system. The same ambitious agenda of scientific discovery that drives the need for exascale computing requires advances in computational and applied mathematics. One driver is the need to simulate multiscale, multiphysics models with unprecedented ranges of physical scales and resolution. A second driver is in the analysis of increasingly large amounts of data from experiment and simulation. These scientific requirements directly drive the need for research in applied mathematics. The twin needs for applied mathematics research and exascale computing should be viewed as independent and complementary to advance scientific discovery.

At the same time, as the size of computer systems and the problems they address scale up and as the architectures change, additional mathematical challenges emerge as a result of the need to make effective use of the hardware. For example, the report of the 2012 DOE Workshop on "Extreme-Scale Solvers: Transition to Future Architectures" identified a number of such areas. These areas include fault tolerant/resilient algorithms, mix-precision arithmetic algorithms, energy-efficient algorithms, and communication/latency hiding in algorithms.

3. **As supercomputing capability advances, there are an increasing number of beneficiaries within the Federal government. For example just within the Department of Energy, applied energy programs are taking advantage of supercomputing to advance modeling and simulation of energy research. The National Science Foundation supports a significant scientific computing research program and the National Institutes of Health funds medical research that uses increasingly fast machines. What areas of opportunity exist for cooperation within various Federal government agencies to continue to push the envelope of computing speeds?**

The agencies which support high performance computing enabled research in one way or another are aware of each other and discuss issues of overlap through a variety of formal and informal mechanisms. NITRD formally is charged with coordinating Federal IT programs including High End Computing, Big Data, and Large Scale Networking. DOE ASCR and NSF Cyberinfrastructure leaders also regularly meet to discuss their plans and activities. DOE ASCR also regularly meets with NNSA ASC.

From the perspective of national science and industry progress, it is essential that the federal investments consistently and effectively allow researchers and developers to push the frontiers. Coordination among the agencies is important. More important is nurturing the ecosystem of computing and computational science research across the universities, national laboratories and industry.

As we move into a new generation of computing, the various agencies have new opportunities for collaboration on development, applications, and deployment of exascale technologies.

a) Exascale technology deployment: DOE continues to push the high end leadership computing and will explore exascale first. DOE mission needs (including both science and security) include clear drivers for exascale applications. Exascale technologies will be the basis for systems at all scales in the future computing infrastructure. NSF cyberinfrastructure brings high end systems to a broad swath of the nations' researchers. NIH is focused on the use of computing technologies to further their mission, rather than on the development of those technologies. DOE facilities can be shared, as for example in the INCITE program and thereby advance research across the nation.

b) Computing and Applied Mathematics Research: DOE and NSF have research programs that impact computer science and software development for advanced computing. NSF invests in fundamental research in computing, applied math, and basic science applications that will contribute to exascale and in the deployment of some shared infrastructure, but it is only DoE that combines it all into high-end systems that stress the technology for the sake of applications important to DoE's mission (and useful for other mission agencies as well). The research communities are well aware of these programs and take good advantage of them.

c) Application Areas: computationally intensive science, engineering, and industrial applications garner support across agencies and the private sector. As we deploy exascale technologies, it is important that they be widely available for applications. Each application area has its own pattern of support and interaction with the agencies. Our challenge is to adopt policies between the agencies that allow for applications to have access to the most effective technologies as they emerge.

d) Computational Science Workforce Development: the nation has a compelling need for a workforce educated in computational science at all levels. Coordination among agency program that address this need is useful. Taking advantage of unique facilities and capabilities in individual programs (such as in the DOE science labs with ASCR's Computational Science Graduate Fellows program) should be encouraged.

4. **With respect to achieving an extraordinary number of computations per second, Exascale speed appears to be a somewhat arbitrary goal. With current budgetary constraints, should DOE consider "slower" systems that would still be by far the fastest in the world?**

    a. **Should we consider pursuing capabilities that would be more than an order of magnitude increase over current computing power, such as several hundred petaflops? Would pursuing these computing speeds be more achievable, and perhaps allow us to avoid deep cuts to other areas of DOE's computing portfolio?**

Over the last 50 years, computer power (as measured in operations per second) has been exponentially increasing. Underlying this steady overall increase, there have been key theoretical and technological 'steps' that enabled continued improvement – the path forward has never been exactly smooth.

In the case of exascale computing, however, we have to simultaneously face an intertwined set of challenges – the end of Moore's law and the large and widening gap between compute speed and communication speed – that amount to a paradigm shift. If the research isn't done we'll never get to exascale, and if we don't employ a design-experiment-revise approach by building systems at scale, we might not get there at all. See the 2012 NRC Report "The New Global Ecosystem in Advanced Computing: Implications for U.S. Competitiveness and National Security" for a deeply thoughtful analysis and discussion of this situation.

For us, the exascale goal was originally proposed for 2018-2019 time frame, which would have established the U.S. as the unchallenged global leader in advanced computing. We have now seen that time frame likely slip to the 2021-2022. We have helped lead the world in the research needed to establish the likelihood that exascale can be achieved and now we are not alone among nations in actively pursuing exascale. (For example, as of June 2013 China's National University of Defense Technology supercomputer is the fastest in the world).

If we attempt to focus on leveling off our development at the several hundred petaflop level, then even if that were the 'fastest in the world' when it was deployed, we would be passed by in a few years. To me, this amounts to abandoning our investment in computing leadership and consoling ourselves with the hope that some other nation will pioneer exascale computing and that the costs will come down so it will be cheaper for us to buy it from them. I do not think this is a position we should take.

***The benefits of going to exascale far outweigh the costs.***

With the necessity of redoing computer applications comes the opportunity to include new predictive and analytical capabilities that will ideally underpin all mathematical models going forward. Two such capabilities deserve emphasis because they involve fundamental advances: (1) sophisticated uncertainty quantification can be embedded within each application, providing an overall statement of accuracy for computational forecasts and predictions; and (2) the application framework may be designed from the beginning as multi-scale – say, from molecules to planet-level. The consistent presence of both these crosscutting technologies will represent a major advance, long discussed but never before possible, in the very character of simulations.

Exascale will have a broad and positive impact on U.S. industrial competitiveness. As already noted, exascale technology breakthroughs will affect leadership from laptop to exaflop systems, because programmers at every scale will be faced with issues of performance and programming. As with high-performance computing in the past, high-tech industries such as transportation, aerospace, nuclear energy, and petroleum will rapidly acquire exascale applications and technology, especially those that allow accurate representation of multiple scales. Science breakthroughs at exascale may also lead to exponential growth in new industries such as renewable energy and materials by design.

5. **The recent Advanced Scientific Computing Advisory Committee ASCAC report on computing facilities identified a "Virtual Data Facility" as a long-term need for DOE's computing research activities: Will you please describe what this facility would be and why it is necessary to develop?**

The proposal for the Virtual Data Facility in the ASCAC Facilities letter responds to two key findings of the Data Intensive Science / Exascale report: (1) "Integration of data analytics with exascale simulations represents a new kind of work flow that will impact both data-intensive science and exascale computing"; (2) "There is an urgent need to simplify the work flow for data-intensive science." That is, there are needs both for exascale simulations to more effectively handle large model datasets and for the large data sets associated with DOE experimental facilities to exploit exascale technologies for their analysis.

The Virtual Data Facility has not been formally designed yet. The vision we have for it is that it would provide "the ability to effectively capture, store, filter, analyze, curate and archive data across all SC facilities..". "This facility would upgrade NERSC, LCF and ESnet resources to provide coordinated storage, archival, analysis and networking capabilities for extremely large data sets." This would involve both new dedicated storage, mid-range computing facilities optimized for working with large dataflows, and use of ESnet to handle large data flows among the centers and the experimental facilities.

The result would be a foundation that allow for the exploration of new workflows for data analytics and work with DOE's big data from experiment and simulation. This would also offer some economies of scale for addressing data needs across the DOE facilities.

6. **How is the development of an exascale computing system related to the President's Big Data initiative? Would the development of an exascale system address other scientific computing challenges that are not exclusively regarded as exascale issues?**

As discussed in the ASCAC report "Synergistic Challenges in Data-Intensive Science and Exascale Computing," there are areas of data-intensive science of interest to DOE that are directly connected to exascale computing. These include both the analysis and visualization and modeling of 'Big Data' generated by large experimental facilities and exascale modeling and simulation that leads to large computed datasets.

In addition, many of the technical challenges of exascale (energy aware computing, fault tolerance, high thread parallelism, optimization of data movement) are also key to effective data-centric computing. Along the road to an exascale machine, exascale technology will be important for data-centric, mobile, and general computing. Conversely, some technologies developed for data-intensive computing will have application to exascale.

# Hearing Questions for the Record
# The Honorable Randy Neugebauer

1. **How can we effectively pursue exascale computing systems in a flat budget environment without negatively impacting core Department of Energy computing efforts, such as computational science, network upgrades and applied mathematics activities? If related computational science activities are hindered, is exascale still worth pursuing? In .other words, do its potential benefits outweigh potential drawbacks?**

Core DOE computing efforts involving computational science, facilities, and applied research are part of the foundation that has enabled us to envision, design and plan for exascale. Exascale computing systems represent the future foundation of the core computational science needed by DOE and the nation. This mutual interdependence is a virtue rather than a conflict, and I strongly resist the idea of shortchanging one for the other.

As noted in previous answers, I believe that the benefits of going to exascale far outweigh the costs, both for DOEs' mission and for its positive impact on U.S. industrial competitiveness. The overall cost/benefit consideration should include these impacts beyond direct ASCR research and facilities.

*Responses by Dr. Rick Stevens*

**America's Next Generation Supercomputer: The Exascale Challenge**

**Response to questions from The Honorable Cynthia Lummis**

**From Rick Stevens**

1. Important scientific research is conducted at all levels of computing systems, from the fastest world-class machines to relatively "slower" supercomputing systems. How will technology developed as part of an exascale program affect systems at smaller scales?

The technological advances needed to create exascale computer systems that are affordable to own and operate (e.g. lower power, more concurrency, more memory bandwidth, improved data store, improvements in reliability, improved programming models, etc.) will impact the whole range of computing systems from commercial servers used in data centers and clouds, to desktops, laptops, tablets and phones. Improvements in price performance due to these technological advances will mean that US companies will be able to install Petascale systems to replace todays Terascale systems for roughly the same cost as the systems they have today. Providing a dramatic increase in capability for engineering, data analysis and design.

Improvements for consumer devices will include improved battery life in mobile devices, new capabilities such as better voice and image recognition and improvements in graphics and network performance. In the world of big data and data analytics the improvements from the R&D needed for exascale will enable faster and more complex analysis and improve affordability of large-scale data systems. In short the advances needed to reach the exascale will positively impact nearly all areas of information technology.

2. As supercomputing capability advances, there are an increasing number of beneficiaries within the Federal government. For example, just within the Department of Energy, applied energy programs are taking advantage of supercomputing to advance modeling and simulation of energy research. The National Science Foundation supports a significant scientific computing research program and the National Institutes of Health funds medical research that uses increasingly fast machines. What areas of opportunity exist for cooperation within various Federal government agencies to continue to push the envelope of computing speeds?

Supercomputing is being used by nearly every federal agency. Most agencies benefit from access to HPC technology either via their own computing systems or those operated by contractors or service providers. With the exception of a few agencies (e.g. DOE, DOD and NSF) most agencies are users of supercomputers rather than developers of supercomputers. However each agency has a unique set of mission needs or distinct user communities. There are primarily three avenues of cooperation and collaboration between agencies that provide opportunities. First those agencies that have the capability to invest in R&D to create new supercomputing capabilities should coordinate those efforts to insure that resources are being leveraged in the most effective way. This probably applies to agencies such as DOE, DARPA, DOD and perhaps NSF. Second, those agencies that are investing in design of next generation machines should be representing the needs and requirements of all the federal agencies in those efforts. For example NIH, NOAA, NASA, EPA, etc. should be collaborating with DOE and others through a multilevel co-design process to ensure that those agencies needs are factored into the design of future machines. Finally, within the requirements of mission needs the Agencies need to be more aggressive in sharing access to machines and training for their staff and to increase the level of cooperation in the development of applications software.

3. With respect to achieving an extraordinary number of computations per second, Exascale speed appears to be a somewhat arbitrary goal. With current budgetary constraints, should DOE consider "slower" systems that would still be by far the fastest in the world?
    a. Should we consider pursuing capabilities that would be more than an order of magnitude increase over current computing power, such as several hundred petaflops? Would pursuing these computing speeds be more achievable, and perhaps allow us to avoid deep cuts to other areas of DOE's computing portfolio?

The exascale performance objective was selected to be a stretch goal for roughly a 10-year program. It also represents a capability that dramatically impacts the outcomes of science applications (factors of 10 certainly make a difference and a fact of 1000 was selected at the time (2008) to represent three generations of systems worth of improvement. Each system improved over the previous by a factor of 10. Since 2008 we have deployed machines in the 10's of Petaflops range and are on track to deploy 100 Petaflops systems in the 2017 time frame. Exascale simply represents the next level beyond that goal. Historically it has taken roughly 10 years to improve performance by a fact of 1000x over the last 50 years. So this objective is rooted in historical trends and what the community believes is possible. It also in a performance level that offers an opportunity for many new qualitatively different capabilities in science, ranging from more accurate hurricane forecasts to markedly improved materials and chemistry design capabilities to the ability to quickly analyze data from the largest scientific instruments without falling behind.

It is widely expected that we will eventually reach a exascale performance level within four generations of systems at the current rate of investment, so the more fundamental issue who will get to that level before us if the US does not increase its level of investment both in the technology as well in the deployment of the technology. Today China's biggest computer is roughly a factor of two faster than the biggest machines in the US. However it is built with the same generation of technology. So it points out that if we want to not only have the best technology but have a deployed capability that is also the fastest we will need to be prepared to scale up our investment in the deployment of systems to keep pace with competitors as well as the research and development needed to enable those systems.

**Response to the question from The Honorable Randy Neugebauer.**

**From Rick Stevens**

1. How can we effectively pursue exascale computing systems in a flat budget environment without negatively impacting core Department of Energy computing efforts, such as computational science, network upgrades and applied mathematics activities? If related computational science activities are hindered, is exascale still worth pursuing? In other words, do its potential benefits outweigh potential drawbacks?

If we are in a flat budget environment we will need to make choices. Computing is an area that the US has clear leadership and it impacts nearly all walks of life. It is one of the primary drivers to our scientific and technological competitiveness. If the choice is a flat budget within the DOE computing portfolio then one should invest in a well-selected long-term research agenda while maintaining some balance between math, computer science, applications domains and facilities. Ensuring that facilities are upgraded and well supported but at the scale we can afford without trading off opportunities for the next generation of scientists and mathematicians. Progress towards exascale should occur at the rate the market place can support in this model.

However…

Before we settle on that solution, I believe we should make a hard search for other areas in the broader federal budget that are likely to have less impact in both the short and long term and consider reprioritization and rebalancing across larger budget categories to find the resources to invest in HP.   The impact and importance of areas of science and investment do change over time.  Budget levels that made sense 10 years ago might no longer be the optimal resource allocations for today.  In times of flat budgets one needs to take a whole portfolio optimization approach that rebalances investments between disciplines as well as within them.

High-performance computing is a critical US technology, one where we have the ability and know how to maintain our leadership edge and in my view it is hard to find other areas with a clear plan for progress and a track record in broad economic, scientific and national security impacts.

*Responses by Ms. Dona Crawford*

## EXASCALE COMPUTING CHALLENGES AND OPPORTUNITIES

## RESPONSE TO QUESTIONS FROM THE COMMITTEE

Hearing of the House Science, Space, and Technology Committee
Subcommittee on Energy
U.S. House of Representatives
May 22, 2013
Dona L. Crawford, Associate Director for Computation
Lawrence Livermore National Laboratory

Response submitted on June 20, 2013

### Questions from Chair Cynthia Lummis

1. **As supercomputing capability advances, there are an increasing number of beneficiaries within the Federal government. For example, just within the Department of Energy, applied energy programs are taking advantage of supercomputing to advance modeling and simulation of energy research. The National Science Foundation supports a significant scientific computing research program and the National Institutes of Health funds medical research that uses increasingly fast machines. What areas of opportunity exist for cooperation within various Federal government agencies to continue to push the envelope of computing speeds?**

**Answer**

Supercomputing is a vital national asset. As supercomputing capabilities increase, the scope of the impact increases—from national security to economic competitiveness, energy security, and health care. However, it is only after the pioneering work of the leaders in the design and development of computer architectures, operating and data storage systems, and software and data visualization tools that the leading-edge systems become practical and affordable for use by a wider customer base. Ensuring that a broad community can take advantage of technology improvements is challenging and would benefit from greater cooperation. The reason there are "an increasing number of beneficiaries" to supercomputing today is that the remarkable advances—originally led by the Department of Energy (DOE) National Nuclear Security Administration's (NNSA) laboratories partnered closely with industry, and recently including DOE Office of Science (SC) laboratories—are now being much more widely used.

This pattern of trail blazing followed by wider dissemination has been demonstrated through many decades of supercomputing R&D. Other nations see the benefit and are emulating the U.S. in its time-proven strategy of a focused federal investment in supercomputing. Since the date of the hearing, China again took the lead in the TOP500 list of the world's most powerful supercomputers with Tianhe-2, developed at their National University of Defense Technology. This computing system is part of the sustained, balanced investment being made by the Chinese federal government, and one demonstration, among others, of their intent to surpass the U.S. in this arena. The challenge ahead for the U.S. is to lead in the advancement of the technologies for exascale-capable, balanced, computing platforms that meet important national needs and ensure rapid, widespread utilization of those advances. Ensuring the rapid dissemination of supercomputing improvements is an area ripe for cooperation.

Cooperation with other federal agencies (for example, the DoD for mission oriented deployments, NIST for standards and cyber security, NSF for basic computer science research, and NIH for biomedical applications) is important. This helps to maximize the benefits of the advancements by expediting the development and dissemination of supercomputing applications for use across the U.S. The technically challenging pioneering work to develop the exascale capable infrastructure has to be a focused effort, and I firmly believe that responsibility should rest with DOE because of its history of delivering these solutions and managing complex integrated projects. In view of today's fiscal constraints, we can most effectively leverage the expertise of the NNSA and Office of Science laboratories with industry in the co-design and co-development of systems architectures, tools, software, and DOE application codes. This includes balanced investments in both the ongoing core advanced computing programs and breakthroughs necessary to achieve exascale capabilities. If this critically important responsibility is distributed to a host of agencies, I fear U.S. leadership in supercomputing will be plagued by lack of focus and loss of momentum while others move ahead with highly focused programs.

2. **Recently, Los Alamos National Lab retired IBM's Roadrunner computer. At a cost of $100 million, the machine was the world's first petaflop computer and was only in operation five years prior to retirement. A computing expert noted that Roadrunner was "created by the artificial goal of the petaflop milestone." Do you agree with this assessment?**
    a. **With the expected cost to develop exascale considerably higher than $100 million, how can we ensure the first exascale system does not experience a similarly short lifecycle?**

**Answer**

The petaflop milestone was not artificial. The Stockpile Stewardship mission of the National Nuclear Security Administration (NNSA) required petascale computing capabilities to maintain the U.S. nuclear deterrent. The Roadrunner procurement responded to early recognition that technology changes were necessary to continue the advancement of computing capabilities, and the machine provided a capability whereupon NNSA scientists developed and tested codes to better predict the performance of aging systems in the nuclear stockpile.

A five-year life span for advanced technology is fairly typical. It is hard to think about the life cycle of a supercomputer being comparable to that of a laptop or an iPad, but there are many similarities, especially in the low level components that change rapidly. More important than its actual life span is the work accomplished during the time Roadrunner was in operation. Roadrunner enabled all three NNSA laboratories to achieve mission deliverables necessary for sustaining the deterrent that otherwise would have not been possible or much more costly. Moving forward, even greater supercomputing capabilities (and greater energy efficiency) are required. Follow-on NNSA investments after Roadrunner and continuing work with IBM led to Sequoia, which is the next step to the capabilities that are ultimately needed.

IBM's Sequoia machine at LLNL is now providing 20-petaflops of computational capability to the nation's nuclear weapons stewards to meet their mission requirements for stockpile life extension programs and advancing Stockpile Stewardship science understanding. Sequoia can effectively address many stockpile issues through the use of two-dimensional (2-D) applications. The system also provides "entry- level" capabilities to run suites of 3-D weapons physics simulations. Today's available technology allows us to simulate in 2-D at high resolution and physics fidelity, or simulate in 3-D at moderate (not high) resolution, but today's available technology does not enable our weapons specialists to simulate at high resolution and 3-D simultaneously. It remains the role of exascale-class systems to address the full breadth of issues that will arise as the stockpile ages, as significant findings are identified, and as even more advanced safety and security features are added to the U.S. stockpile in recognition of today's threat environment.

NNSA and the Office of Science envision that a sequence of increasingly powerful supercomputing systems would be developed on the road to exascale. Due to the technically challenging nature of developing exascale-supporting technologies and computing capabilities, it is vitally important to ensure there are at least two competitive teams each consisting of Office of Science and NNSA laboratories partnered with U.S. high-performance computing industrial collaborators on two or more alternative machine architectures. This competition of ideas and delivery of new systems every few years is critically important to ensure the successful development of an enduring architecture, as well as to ensure that the U.S. leads in the development of ever evolving technologies needed for advanced computing platforms and their utilization. Although the first-of-a-kind system will inevitably be the most expensive, cost-effective leveraging of Office of Science and NNSA capabilities and expertise through competitive teams can keep the costs to a reasonable level.

3. **With respect to achieving an extraordinary number of computations per second, Exascale speed appears to be a somewhat arbitrary goal. With current budgetary constraints, should DOE consider "slower" systems that would still be by far the fastest in the world?**
    a. **Should we consider pursuing capabilities that would be more than an order of magnitude increase over current computing power, such as several hundred petaflops? Would pursuing these computing speeds be more**

**achievable, and perhaps allow us to avoid deep cuts to other areas of DOE's computing portfolio?**

**Answer**

Acquisitions of 100-200 petaflop systems in the FY2016-2017 time frame are part of DOE's plan. As demonstrated by China reclaiming the lead in the June 2013 TOP500 list of the world's most powerful supercomputers, it is clear these plans for slower systems are not ambitious enough to produce systems "that would still be by far the fastest in the world." The DOE's planned slower-than-exascale systems will enable prototype builds on the path toward exascale. Through these interim prototype systems, potential new technologies can be evaluated in the context of new architectures that utilize them to meet DOE mission challenges. These interim systems will be based on current architectures, which we know cannot scale to cost-effectively support exascale-level computing.

The ultimate purpose of exascale, however, is to meet important mission requirements—not to attain a particular speed or to set speed records. Exascale research and development (R&D) aims to overcome the myriad technical issues facing the advancement of supercomputing and lead to the development of a new computing infrastructure that will provide capability with affordable operational costs required for the Department of Energy's (DOE) missions including (but not limited to) sustaining the nuclear deterrent. Efficient exascale-capable computers are not the goal *per se*, but they are metrics of success in a very real sense. Effective investment in exascale R&D also has the potential to sustain U.S. high-performance computing (HPC) leadership.

With the planned modernization of the stockpile and simultaneous decreases in both its overall size and composition, advanced computing and simulation will play an increasingly critical role. Nuclear weapons are complex, three-dimensional (3-D) engineered systems with special materials that change over time as they age. Higher fidelity, 3-D simulation of warheads including detailed representation of initial conditions, engineering features, safety features and security features are required to ensure the safety and performance of each weapon. A thousand-fold improvement over today's modeling and simulation capability (exascale technology) is required to assure with confidence over the long term the safety, security, and performance of the nation's nuclear stockpile.

Surmounting the multiple technical issues needed to scale to exascale-class capabilities will require sustained research and development and some key breakthroughs in technology, including but not limited to low power systems, memory and storage. If the U.S. government does not invest in the development of the next computing architecture and associated software, there will be no way to ensure exascale supercomputing architectures that meets national mission requirements. Investments toward a new architecture are essential for the U.S. to maintain leadership in the increasingly competitive HPC arena.

**Questions from *Representative Randy Neugebauer***

1. **How can we effectively pursue exascale computing systems in a flat budget environment without negatively impacting core Department of Energy computing efforts, such as computational science, network upgrades and applied mathematics activities? If related computational science activities are hindered, is exascale still worth pursuing? In other words, do its potential benefits outweigh potential drawbacks?**

Answer

It is imperative that the U.S. embarks on a research and development (R&D) program to develop new technologies and computer architectures to support exascale computing. The benefits are definitely worth the investment. The nation's strategic nuclear deterrent is maintained through the science-based Stockpile Stewardship Program. This vital national security program relies on high performance computing as the primary integrating capability to assess the safety, performance, and reliability of the U.S. nuclear weapons stockpile in a world without nuclear testing. With the planned and expected decrease of the stockpile in both overall size and composition, advanced computing and simulation will play an increasingly critical role. A thousand-fold improvement over today's modeling and simulation capability (exascale technology) is required to enable long term success in this effort. High performance computing underpins our ability to scientifically resolve outstanding weapons performance issues, address material aging and compatibility challenges, conduct future warhead Life Extension Program activities, and rapidly address results from Significant Finding Investigations.

A sustained investment is required to maintain U.S. leadership in high performance computing. The U.S. needs to commit to an effort to develop architectures for the next generation of high performance computing. China has reclaimed the lead in the TOP500 list of the world's most powerful supercomputers with Tianhe-2, developed at their National University of Defense Technology. This computing system is a result of the sustained and balanced investment being made by the Chinese federal government, and is one demonstration, among others, of their intent to surpass the US in this arena. Our global leadership in high performance computing requires three complementary efforts: 1) continuation of today's base program to utilize today's best capabilities to meet current mission, 2) development of interim architectures at the 100-200 petascale level as prototypes to evaluate promising technologies for exascale, and 3) sustained R&D investment in technologies required to achieve exascale. We must simultaneously invest in all three areas to maintain balance while making a conscious effort to ensure a skilled workforce pipeline. If the U.S. does not invest in maintaining leading-edge computation facilities that serve as magnets for a skilled workforce to pursue innovative research opportunities, our talented teams will dissipate and other nations will lead in supercomputing. Innovation and progress are underpinned by science and technology and only leadership capabilities combined with a capable talent pool can move us forward.

The most cost effective way to make investment in next generation supercomputing on the path to exascale is through a fully integrated program, balancing operation and application of today's systems, near-term upgrades in computational capabilities and an exascale R&D initiative that is executed through a joint Office of Science–National

Nuclear Security Administration (SC–NNSA) effort leveraging the strengths and expertise of SC and NNSA laboratories in close partnership with the U.S. high-performance computing industry. An effort that enables SC and NNSA to move in tandem in this R&D effort will enable cost effective leveraging of resources across the Department to ensure system architectures to meet SC and NNSA mission requirements. Due to the technically challenging nature of developing exascale supporting technologies and computing capabilities, it is vitally important to ensure there are at least two competitive teams each consisting of Office of Science and NNSA laboratories partnered with U.S. high-performance computing industrial collaborators on two or more alternative machine architectures.

2. **What are the challenges related to increasing the energy costs of an exascale computing system? If such a system were to use today's processing and chip technology, how much would the exascale "electric bill" cost? What is your confidence that research can make the breakthrough necessary to achieve it? What is your estimate of the time and expense to realize such a solution?**

**Answer**

The development of exascale-class systems cannot be achieved through a straightforward refinement of today's technologies. Surmounting multiple technical issues will require sustained research and development (R&D) and some key breakthroughs, which I am confident we can achieve through a properly structured, integrated R&D program.

Succeeding generations of microprocessors, standard in computers of every scale, have grown faster by increasing the speed and shrinking the size of transistors, effectively packing more calculations into every unit of time and space on a computer. But now transistors are reaching a lower limit in size and an upper limit in speed. Although individual transistors could be pushed to run faster, speeding up millions of transistors on a microprocessor would drive energy demands and operational costs to unsupportable levels. To make exascale computing practical, the electrical power requirements must be reduced at least ten-fold per floating point operation. Without this reduction, exascale computers would need hundreds of megawatts—enough to power a small city—at an unacceptable cost of hundreds of millions of dollars per year to pay the electricity bill. Developing a low-power system encompasses changes to every component of the computer system: memory (e.g. stacked memory), networks (e.g. optics), and processors (including accelerators).

I am confident we can develop exascale-capable computing architectures to meet our mission requirements through a properly structured and fully integrated program. It is imperative that the U.S. embarks on exascale R&D through a joint Office of Science–National Nuclear Security Administration (SC–NNSA) effort that leverages the strengths and expertise of SC and NNSA laboratories in close partnership with the U.S. high-performance computing industry. The partnerships are required to make the necessary advancements and achieve the breakthroughs required to develop an exascale-capable architecture. Due to the technically challenging nature of developing exascale supporting technologies and computing capabilities, it is vitally important to ensure there are at least two competitive teams each consisting of Office of Science and NNSA laboratories partnered with U.S. high-performance computing industrial collaborators on two or more alternative machine architectures.

The mission drivers of NNSA and the leveraging of NNSA's outstanding systems-engineering track record in partnering with industry to deliver leading edge computing systems will serve to focus technology development on the path to exascale and lead to an architecture that will serve our national security mission requirements. I endorse the Office of Science's early focus on long lead-time R&D in advanced technologies, and I strongly endorse NNSA's continued focus on investments that support the design and delivery of well balanced and well-architected high performance computing systems used to meet mission requirements.

Many in the U.S. high performance computing community believe it will take a decade of effort to achieve an exascale-capable system. The timeline depends upon available funding and will be heavily influenced by the structure of the program. I believe that a three-year R&D effort could produce a development plan with more accurate estimates.

*Responses by Dr. Daniel Reed*

**Responses from Dr. Daniel Reed**
**Vice President for Research and Economic Development, The University of Iowa**

From the House Committee on Science, Space and Technology's Subcommittee on Energy's Hearing on *America's Next Generation Supercomputer: The Exascale Challenge*, held May 22, 2013

**To Questions for the Record from The Honorable Cynthia Lummis**

1. *Please summarize some of the varying views and differences within the computing communities regarding challenges and opportunities associated with the pursuit of an exascale computing system.*

The computing community agrees that the U.S. competitive advantage in advanced computing is eroding, both in enabling technologies and in deployed systems. Since this hearing in May, China has announced a new high-performance computing system that was ranked number one in the world on the June 2013 Top 500 list of the world's fastest computing systems. There is also deep agreement that greater investment in computing research and development is needed to ensure continued U.S. competitive advantage.

The primary community differences surround the need for a specific performance level (i.e., an exascale system) at a particular time, versus the need for a coordinated research and development program, along with staged system deployments as application needs, technology capabilities and economics dictate. No one debates the need for greater research investment and coordination across research and mission agencies. Nor does anyone debate the need for continued deployment of supercomputing systems with leading edge capability.

2. *As supercomputing capability advances, there are an increasing number of beneficiaries within the Federal government. For example, just within the Department of Energy, applied energy programs are taking advantage of supercomputing to advance modeling and simulation of energy research. The National Science Foundation supports a significant scientific computing research program and the National Institutes of Health funds medical research that uses increasingly fast machines. What areas of opportunity exist for cooperation within various Federal government agencies to continue to push the envelope of computing speeds?*

Historically, the Federal agencies have collaborated and coordinated their research computing activities via the National Coordination Office (NCO) for the Networking and Information Technology Research and Development (NITRD) program. Recently, that cooperation has been less effective as agencies have pursued separate agendas, driven by their priorities and constituencies. It is important that we rekindle that interagency collaboration, leveraging the unique roles and capabilities of each agency. As I noted in my written testimony, these each have important roles:

- NSF – basic computing research in the enabling technologies; data management and sustainable cyberinfrastructure for the national science and engineering academic community
- DoD – advanced technology research and prototyping; mission-oriented deployments
- NIST – standards and cybersecurity
- NIH – computational modeling, big data analytics and biomedical applications for higher quality, lower cost health care

- DOE – computational science, systems research and prototyping; large-scale system deployments, building on the research and operations staff of the Office of Science and NNSA laboratories

Research problems in both enabling technologies and applications are increasingly interdisciplinary, requiring diverse skills. We need multidisciplinary teams working together to develop leading edge scientific and engineering applications and the advanced computing systems on which they operate. This principle of co-design is a deep and integral part of the DOE exascale plan.

3. *How is the development of an exascale computing system related to the President's Big Data initiative? Would the development of an exascale system address other scientific computing challenges that are not exclusively regarded as exascale issues?*

Without doubt, development of exascale technologies would address scientific and technical challenges other than just those constrained by computation speed. Most of the enabling technologies for exascale computing and big data analytics are the same. Today, the technologies inside commercial cloud data centers and those found in advanced supercomputing systems are very similar, and they both share the design challenges of large scale. Moreover, the system-level technical challenges in building next-generation cloud infrastructure -- low power processors and memories, resilience at scale, programmability and simplicity – are almost identical. The differences between exascale computing and clouds (big data) are primarily in the software technologies, software development culture and potential applications.

The Big Data Initiative poses a complementary but equally necessary set of research and development challenges as exascale computing. Innovative science and engineering increasingly requires both advanced (high-performance) computing and analysis of large volumes of data. The latter includes data produced by computational simulations and data captured from scientific instruments. The recent Executive Order that research agencies develop data management plans reflects that reality. As I noted in my written and oral testimony, I believe it is crucial that neither exascale computing nor the big data initiative be sacrificed for the other; both are important as research and development activities and for economic and national security.

### To Questions for the Record from The Honorable Randy Neugebauer

1. *What role does the private industry have in advancing supercomputing capabilities? What possibilities exist to effectively leverage public-private partnerships?*

U.S. high-technology companies play a crucial role in advancing supercomputing capabilities. Vibrant semiconductor, hardware, software and applications industries are key to not only U.S. global economic and technical competitiveness and economic security, but also to our national security. These companies translate basic research ideas into commercial practice, creating both the component hardware and software technologies (e.g., microprocessors, software tools and applications) and the supercomputing systems that contain them. They also develop new products based on their own research and development. This combination of public sector basic research, supported by the U.S. government, and private sector innovation and productization, has long been the envy of the world.

The keys to this public-private partnership are robust investment in basic research at U.S. research universities and national laboratories, translation of those ideas to industry, and a long-term, strategic plan that targets research and development on those technologies that are needed to build the next generation of advanced computing systems. These include, but are not limited to, low-power processors and high-bandwidth memory systems, scalable software and next-generation applications. Finally, it includes a predictable acquisition plan for the products that ensures a market and creates incentives for industry participation.