High Performance Computing – Scientific Discovery and the Importance of Collaboration

Raymond L. Orbach
Under Secretary for Science
U.S. Department of Energy

French Embassy
September 16, 2008

I have followed with interest the establishment of the Partnership for Advanced Computing in Europe (PRACE) which was launched in January, 2008 to create a persistent pan-European High Performance Computing service and infrastructure for research. Bringing together the know-how and resources of its 16 partners—Germany, United Kingdom, France, Spain, Finland, Greece, Italy, Netherlands, Norway, Austria, Poland, Portugal, Sweden, Switzerland, Turkey and Ireland—PRACE is the single European entity that will develop the top level of the European HPC ecosystem and provide European researchers with access to supercomputing capacities at a world-class level, transgressing those affordable at a national level. As a first step, PRACE recently selected of six prototype architectures for its Petaflop(s) class systems.

High Performance Computing is the third pillar for scientific discovery – along with theory and experiment. And, experiment and computational science are indelibly linked to technology development and engineering. There is a clear feedback loop. Science is advanced by new technologies, and, concomitantly, drives technology development, and the Department of Energy is at the forefront of this nexus. Through the combination of high performance computing facilities, applications expertise, applied mathematics and computer science research, the Department is delivering computational science breakthroughs today, and leading the way for tomorrow’s scientific discoveries in a wide range of areas. These include climate research, nanotechnology, energy and the environment.

These achievements have developed through President Bush’s first and second terms, and are a direct consequence of his commitment to basic research. The first step was to improve performance of the U.S. computational and networking infrastructure, as well as existing codes—similar to current efforts within PRACE and GEANT. At the beginning of President Bush’s Administration, the Department of Energy had largely switched from vector systems to massively parallel systems. And our application codes were getting, at most, 10-20 percent of theoretical peak – with many applications getting in the single digits. The largest computer available to the Office of Science was at the National Energy Research Scientific Computing (NERSC) facility, which at the time seemed quite capable at 3 Teraflops peak speed. Our highest sustained speed on, for example, fusion codes was 485 Gigafllops. Data was traveling into NERSC, and around the DOE system via ESnet at 622 megabits per second. Gains were being made with clock speed improvements in individual processors, but future gains would provide only marginal
improvements. It was clear at that time that we needed a radical new approach to scientific code development to get more out of our investments in computing. What was not clear was the profound impact our new approach would have on science.

The Department of Energy undertook an innovative program for accelerating progress in computational science by breaking down the barriers between disciplines, and forming dynamic partnerships between applications – such as astrophysics or biology – and the computer scientists and mathematicians who deeply understand the hardware and software available. This program, called Scientific Discovery through Advanced Computing or SciDAC, has a remarkable history of success in advancing code development, improving the performance of scientific applications by up to two orders of magnitude. Today, NERSC is in the process of upgrading to a 355 Teraflop quad-core Cray XT4 system, and ESnet delivers 10 gigabit per second core service with Metropolitan Area Networks, such as the one in the bay area that moves data into and out of NERSC, at 20 gigabits per second. In addition, the Department now also offers two Leadership Computing Facilities – at Oak Ridge and Argonne National Laboratory. Together these facilities offer architectural diversity with a 263TF Cray XT4 at Oak Ridge and the world’s fastest computer dedicated to open science, the 556 IBM Blue Gene P at Argonne. And SciDAC Fusion codes are getting up to 75% of theoretical peak on our 263 Teraflop Cray system at Oak Ridge.

The impact on our science was even more dramatic. Our climate models can now spontaneously generate key phenomena such as hurricanes; we are learning so much about things we can’t study any other way such as supernovae; improved modeling and simulation of sub surface flow allow us to better understand nuclear waste containment as well as facilitate carbon sequestration efforts; multifaceted collaborations between applied mathematics, computer science and fusion scientists are providing key elements in the research towards the success of ITER and the taming of nuclear fusion for our energy needs – the list goes on…

These facilities should be thought of as similar to our light sources and large scale accelerators. They benefit from open peer-reviewed competition on a global scale for use. Supercomputers should be regarded similarly. We opened access to these systems though a competitive peer reviewed process called the Innovative and Novel Computational Impact on Theory and Experiment, or INCITE, program, opening access to the Office of Science’s largest computing resources to all qualified researchers regardless of institutional affiliation or sponsor. This year, we made a quarter of a billion hours of computing time available through INCITE that will expand to over a half billion hours in 2009.

The response has been wonderful. Ever since we created INCITE, in 2006, we have been able to offer an order of magnitude more hours each year. However, the requests have always exceed what we have available by about a factor of three! I am particularly proud that this demand has been universal, with applications from, and awards to, industry and other agencies continuously growing.
I am also proud that through INCITE we are sharing our resources with international research teams. Last year we awarded an INCITE allocation to Thierry Poinsot and Gabriel Staffelbach from the European Centre for Research and Advanced Training in Scientific Computation (CERFACS) in Toulouse, France. In June, Dr. Poinsot briefed a small delegation from DOE, NSF and our Leadership computing facilities on his INCITE project which is focused on gaining an understanding of ignition, quenching and combustion instabilities in real gas turbines and demonstrating the usefulness of Large Eddy Simulations in the design phase of real gas turbines. Access to our 556 TF IBM Blue Gene at Argonne allowed the team to performed simulations of full combustion chambers not just a sector.

The ability to study larger, more complex problems using the Department’s supercomputing resources is a common theme among INCITE projects. For example, the ability to simulate over a billion different conformations of the unstructured protein, alpha-synuclein, has enabled Igor Tsigelny of the San Diego Supercomputer Center (SDSC) at the University of California-San Diego (UCSD) and Eliezer Masliah of UCSD to determine which conformation causes Parkinson’s disease. Over the past three years, INCITE simulations run on the IBM Blue Gene L and P machines at ANL identified the ring-like conformation of alpha-synuclein that penetrates the cell membrane to form a pore leading to Parkinson. Because the concentration of calcium outside cells is much greater than inside, calcium flows through the pore and triggers a cascade of events leading to cell death. Electrophysiological studies support the pore theory by showing current flow caused when ions traverse membranes. Halting pore formation may stop the disease from progressing once it has begun or even prevent the disease from developing. Laboratory tests with mice show promise, with drugs blocking aggregation of alpha-synuclein.

As so here we are today. SciDAC and INCITE are already transforming Science. The announcement of the PRACE prototypes and the current installation of the Department’s petaflop Cray XT5 at Oak Ridge create an opportunity for further international collaborations. In late September, Dr. Michael Strayer, the Associate Director for the Department’s Office of Advanced Scientific Computing Research will lead a delegation from our supercomputing facilities at Argonne, Oak Ridge and Lawrence Berkeley National Laboratory to meet with a French delegation comprised of representatives from Electricité de France (EDF) R&D, CERFACS, the Commissariat à l’Energie Atomique (CEA) and the Institut National de Recherche en Informatique et en Automatique (INRIA) to continue the discussions started in series of meetings initiated last June by Jean-Philippe LaGrange, the former Science Attaché here at the French Embassy. The goal of the September meeting is to develop a short list of priority projects and participants to form the first phase of a broad US/French program. Specifically the delegations will discuss potential joint SciDAC-like collaborations in the areas of applications software and algorithms for peta-scale computing with a particular emphasis in the areas of energy, environment and basic science.

One of the reasons of the success of our SciDAC program was the corresponding research investment we made in the technologies that enable scientists to effectively use
our supercomputers the day they are installed. These same investments need to be continued for the petaflop resources and the US/French team will discuss the formation and coordination of an international open source software consortium focused on research in the enabling technology areas of systems software, I/O, data management, visualizations and libraries of all forms, again targeting petascale computers.

But that is just the beginning. There are exaflop systems looming on the horizon. Based on the three exascale workshops we conducted last year, there are huge opportunities at this speed that will transform the landscape of our world, from integration of human behavior into climate codes, to bringing convergence to the partial differential equations that govern our physical universe. Realizing the potential of exascale resources however represents a significant challenge because of their complexity. Many of the tools and techniques that we have worked so hard to develop may need to be re-invented or replaced because they won’t scale. In the Department, we recognize that with great risk comes great reward and are already investing in the long-term research that we hope will help meet those challenges. But the challenge is so great that it will take funding from more than one agency or government to effectively address it. We must act now to identify our leading computer scientists and mathematicians to work collaboratively on the research and development of new programming models and tools to address extreme scaling issues, hundreds of cores per node, heterogeneous systems with a mix of processor types, and performance. As we approach the exascale, the computers will become more expensive in cost and power. We need to start high level policy discussions to identify areas where governments and agencies can cooperate in larger-systems deployments for attacking global challenges such as energy and the climate.

And those are the easy problems. As our industries and laboratories tell us, there is a worldwide shortage of trained computational or simulation scientists. Each of us has programs that we have or are going to implement in this area.

- Over fifteen years ago, when the Department recognized its growing need for computational scientists, it created the successful Computational Science Graduate Fellowship program. At that time the fellowship also addressed the lack of computational science academic programs. Today graduates from that program are working at the Department’s national laboratories, in industries like Pratt & Whitney, Procter and Gamble, IBM and Microsoft and in academic institutions such as Massachusetts Institute of Technology and Stanford.

- As a result of the of the US/ French team meetings in June was the Pierre Wolf, a CERFACS PhD student will start a six week visit at Argonne National Laboratory in September to work on scaling up combustion code to fully utilize the 32 cabinets of the IBM Blue Gene/P their and to use the system on a set of benchmark problems related to his thesis.

- Last October, DOE and NSF in conjunction several French universities and research centers sponsored a US France Symposium for Young Engineering students. This symposium brought together young US and French investigators to foster potential computational science collaborations. A second symposium is
planned for this October in Oak Ridge with an expanded partnership that includes the US, France, Germany, Italy, United Kingdom and Japan.

However to take full advantage of our petascale and eventually exascale resources, we need to move beyond standalone successful programs and develop a common strategy that includes joint academic and corresponding research experience programs to educate, train and bring together the next generations of computational scientists.