EDITORIAL: Dr. Michael Strayer

Information Technology in the Next Decade

Dr. Michael Strayer

M. Strayer

At the end of the first decade of the 21st century the impact of information processing on society has been most dramatic. The amount of data, both personal and otherwise is exploding at an astronomical rate, far faster than society's ability to effectively manage or control. The changes that have been brought to people's lives by the availability of data and the communication of data effortlessly across the planet is often referred to as an "information processing" revolution. As in most revolutions there is the good and the bad. Economists speculate that by the second half of the 21st century more than half of the energy consumed worldwide will be for the storage, communication, and processing of digital information. Today massive data centers with hundreds of millions of processors are being built at an enormous rate and in aggregate consume as much power as a small town. At the current rate of investment it is not impossible that a handful of companies will manage most of the digital information in the world – with the exception of science. Perhaps manage is an inappropriate term, since the complexity of the growth appears widely out of control.

In modern technological societies it is acknowledged that innovations and discovery are brought through science; however scientific data seem to lag behind others in the sense that it is assuredly more zealously guarded. There are a number of reasons, but individual researchers tend to maintain ownership and control of the information that fuels their careers. Scientists are conservative and protective. In the arena of modeling and simulations, 60 years of growth has brought an increase of a factor of 15 orders of magnitude in computing power – every factor of a thousand has forced new concepts and ideas into computing and the role of computing in science. The growth of computer science as a discipline, together with applied mathematics and simulations, has led to a new interdisciplinary field of study – computational science. Many universities now grant degrees in computational science. Much of the growth in computational science can be traced to the DOE's mission-driven science and technology programs. The use of advanced simulations to more fully understand and unravel complexity in such areas of nuclear security and in science as materials, combustion, catalysis, and climate change has met with unparalleled success. Here teams of researchers are formed, comprised of computer scientists, applied mathematicians, and specific domain scientists to address mission problems – in nuclear security, the Advanced Simulation and Computing program, and in science the Scientific Discovery through Advanced Computing program. Both programs have promoted young investigator and graduate student programs in computational science. Notably, The DOE Computational Science Graduate Fellowship was established to promote computational science as an interdisciplinary field of study and has trained more than five generations of students who are active in universities, laboratories, and government today.

General-purpose exascale computer systems are expected to be technologically feasible within the next 15 years. These systems will likely have between 10 million and 100 million processing elements or cores. The major U.S. vendors of large-scale systems and processors are in general agreement that these systems will push the envelope of a number of important technologies, including processor architecture, scale of multicore integration (perhaps into the range of 1,000 cores per chip or beyond), power management, and packaging. The projected exascale systems themselves will have part counts comparable to those of today's largest systems (or slightly larger). A major source of uncertainty is how quickly the general marketplace will be able to adopt highly parallel, single-chip, multicore systems in normal information technology products. The current belief is that the broad market is not likely to be able to adopt multicore systems at the 1,000-processor level without a substantial revolution in software and programming techniques for the hundreds of thousands of programmers who work in industry and do not yet have adequate parallel programming skills. In particular these systems are expected to have in aggregate requirements of more than hundreds of millions of processors, and billions of threads. Furthermore, the expected component failure rate will limit the mean time to interruption to less than about 30 hours. None of the present-day concepts for both system software and application software will be applicable. New ideas that can express parallelism, locality, and fault tolerance in the software stack are needed. To realize science at the exascale will require a concerted effort to couple advances in algorithms, programming models, operating systems, file systems, I/O environments, and data analysis tools. In fact, exascale systems are likely to be so demanding that they will drive new working relationships between the disciplinary scientists and the computer science and mathematics communities.

More importantly, we need to begin now to train a new generation of investigators that can aid in developing new ideas in such disciplines as energy research, climate modeling, socioeconomics, environmental modeling, computer science and applied mathematics, basic sciences, and multiscale biology. The broad computational science community has a golden opportunity to accelerate the availability of usable exascale systems. To take full advantage of this opportunity to deliver exascale computing will require an integrated program of investments in hardware and software research and education. Computational science and engineering opportunities at the exascale are wide and deep and have an enormous potential impact on advancing energy technology and fundamental science. ●

Contributor Dr. Michael Strayer, SciDAC Program Director, and Associate Director of Science for Advanced Scientific Computing Research, DOE